

עיבוד תМОנות

מיכאל אלעד



הקדמה

ספר זה נכתב כאוסף הרצאות לקורס מתקדם בעיבוד תמונות בפקולטת להנדסת חשמל ולאחר מכן מודיעי המחשב בטכניון. בגרסתה החדשה זו (אוקטובר 2003) של הספר נוספו מספר פרקים המכסים נושאים בטומוגרפיה ממוחשבת, ייצוגים מולטי-רזולוציוניים, עיבוד תמונות צבע, ונספח בו מובאות הרחבות בנושאים שונים כגון קוונטיזציה, שחזור והתרמות.

אין בכוונתי לחדש בספר זה אלא לרכז באופן קרייא ונוח (ובעיקר בשפה העברית) דיון בנושאים מוגוונים בעיבוד תמונות. דגש מרכזי מושם על עדכניות החומר המובא כאן, ועל ריבוי בדוגמאות לשם המבנה. הדוגמאות במהלך הפרקים נוצרו ע"י תוכנת MATLAB.

בנימית ספר זה מתבססת על מספר מקורות המתוארים בהמשך. זהה עדיין גירסה לא סופית ולצער עדיין ישן שגיאות במהלך הפרקים. אני מקווה לתקן ולעדכן ספר זה לכל ספר נוח ואמין ללימוד עיבוד תמונות ולהוסיף לתוכנו, כך שיוכל להיות בסיס לשני קורסים – מתחיל ומתקדם בתחום.

אני רוצה להודות לדורי דורו שקד על סיוע בקביעת תוכן הספר, ושיחותיו עמי על נושאים שונים הנסקרים בו.

מקורות

- [1] Anil. K. Jain, ***Fundamentals of Digital Image Processing***, Prentice Hall Englewood Cliffs, New-Jersy, 1989.
- [2] Kenneth R. Castleman, ***Digital Image Processing***, Prentice Hall Englewood Cliffs, New-Jersy, 1996.
- [3] Brend Jahne, ***Digital Image Processing – Concepts, Algorithms, and Scientific Applications***, Springer Verlag, 1995.
- [4] Boaz Porat, ***A Course In Digital Signal Processing***, Wiley, New-York, 1997.
- [5] Dan E. Dudgeon and Russell M. Mersereau, ***Multi-Dimensional Digital Signal Processing***, Prentice-Hall, Englewood-Cliffs, New Jersey, 1984.
- [6] George Wolberg, ***Digital Image Warping***, IEEE Computer Society Press, Los-Alamitos California, 1990.
- [7] Nahum Kiriaty, Doron Shaked, and Nir Sochen, ***Lecture Notes in Digital Image Processing***, The Computer-Science Department, The Technion, Israel Institute of Technology (Obtained by personal communication).
- [8] Alfred M. Bruckstein, ***Image Processing – Lecture Notes***, the Computer Science Department, The Technion, Israel Institute of Technology.
- [9] Michael Lindenbaum, ***Lecture Notes in Digital Image Processing***, The Computer-Science Department, The Technion, Israel Institute of Technology. Web-site: <http://www.cs.technion.ac.il/~mic/isp.html>
- [10] Hagit Zabrodzki Hel-Or, ***Lecture Notes in Image Processing***, The Computer-Science Department, Haifa University. Web-site: http://cs.haifa.ac.il/courses/image_p.

.ומאמרים שונים.

עיבוד תמונות

תוכן עניינים

פרק 1 – כללי	1
1.1 תמונה דיגיטלית	2
1.2 קלט אפלט לתמונות דיגיטליות	5
1.3 סוגיות בטיפול בתמונות	8
1.4 עיבוד תמונה ותchromים סמכים	10
1.5 יישומים לעיבוד ואנליזה של תמונות	11
פרק 2 – עיבודאות בדו-מימד	15
2.1 מערכות ליניאריות בדו-מימד – אוותות מעלה הרץ	15
2.2 התמרת פוריה דו-מימדית ותכונותיה	19
2.3 מערכות ליניאריות בדו-מימד – אוותות דיסקרטיים	23
2.4 התמרת פוריה לאוותות דיסקרטיים ותכונותיה	25
2.5 אוותות ומערכות ליניאריות בתמך סופי	29
2.6 התמרת פוריה דיסקרטית (DFT) לאוותות בתמך סופי	39
פרק 3 – דגימה ושחזור של תמונות	41
3.1 דגימת איחוד של אוותות דו-מימדיים	41
3.2 שחזור מדגימה איחוד	44
3.3 תופעת הקיפול בתמונות	46
3.4 משפטים דגימה מוכלים לדו-מימד	51
פרק 4 – קוונטייזציה לתמונות	55
4.1 קוונטייזציה סקלרית	55
4.2 קוונטייזר מקס-לoid	59
4.3 קוונטייזציה עם שיקולים חזותיים	67
4.4 קוונטייזציה וקטוריית קוונטייזציה צבע	71

פרק 5 – עיבודים בסיסיים לתמונות	75
5.1 פועלות נקודה והיסטוגרמה תמונה	76
5.2 פועלות עיצוב היסטוגרما	83
5.3 פועלות מרחביות בסיסיות	96
5.4 פועלות מרחביות מסתגלות	107
פרק 6 – שחזור תמונה	113
6.1 בעיית השחזור הקלאסית	113
6.2 מושערכים סטטיסטיים – 1- ML	116
6.3 שחזור מבוסס מושערך ML	122
6.4 שחזור מבוסס מושערך MAP	125
6.5 שיפור פונקציות ההסתברות לתמונות	132
6.6 בעיות היפוך אחריות והקשר לבניית השחזור הקלאסית	140
פרק 7 – התמורות דיסקרטיות בדו-מימד	143
7.1 התמורות ליניאריות - בסיס	143
7.2 התמרת ה- DFT הדו-מימדית ותכונותיה	148
7.3 התמרת ה- DCT הדו-מימדית ותכונותיה	151
7.4 התמרת Hadamard דו-מימדית ותכונותיה	156
7.5 התמרת Haar ומבוא ל- Wavelet	160
7.6 התמורות מבוססות תוכן תמונה	168
7.7 Over-Complete	171
פרק 8 – דחיסת תמונה	175
8.1 יסודות בתורת האינפורמציה	175
8.2 יתרויות בתמונה וניצול	181
8.3 קידוד אנטרופיה - אלגוריתם הופמן	184
8.4 קידוד עם שגיאות - גישת החיזוי	188
8.5 קידוד עם שגיאות - גישת ההתרמה	191
פרק 9 – עיבוד סדרות של תמונות	195
9.1 תנואה בין תמונות	196
9.2 שערוך תנואה בין תמונות - גישת בלוקים	201
9.3 שערוך תנואה בין תמונות גישה מלאה	205

210 9.4 ניקוי רעש בסדרות של תמונות
215 9.5 דחיסת סדרות של תמונות וմבוא ל- MPEG
219	פרק 10 - שחזור תמונה מהיטלים
220 10.1 בעיות שחזור מהיטלים וניסוחן המתמטי
221 10.2 התמרת ראדו
223 10.3 משפט החतך המרכזי והיפוך התמרת ראדו
226 10.4 שחזור מהיטלים ע"י גישות תדר
228 10.5 שחזור מהיטלים ע"י קונבולוציה
229 10.6 שחזור מהיטלים ע"י גישות אלגבריות
235	פרק 11 - מבנים פירמידליים ושימושיהם בתמונות
236 11.1 הפירמידה הגאומית
240 11.2 הפירמידה הלפלסיאנית
249 11.3 התמרת Wavelet ע"י Lifting Scheme
257	פרק 12 - מבוא לתורת הצבע
258 12.1 צבע – מהפיזיקה לעין האנושית
261 12.2 התאמות צבע
265 12.3 שיטות ייצוג צבע בתמונות
271 12.4 ייצוג תמונות צבע
276 12.5 שיפור ושחזור תמונות צבע
283 12.6 דחיסת תמונות צבע
285	<u>נספחים:</u>
286 נספח א' – הרחבות לקוונטייזר אופטימלי
294 נספח ב' – התמרת KLT כהתמורה אופטימלית והרחבותיה
299 נספח ג' – הקצאת סייביות אופטימלית בדחיסה
302 נספח ד' – נושאים מתקדמים בשחזור תמונות

פרק 1 – כללי

פרק 1 – כללי	
1
2
5
8
10
11
	1.1 תמונה דיגיטלית
	1.2 קלטומפלט לתמונות דיגיטליות
	1.3 סוגיות בטיפול בתמונות
	1.4 עיבוד תמונה ותחומים סטטוטים
	1.5 יישומים לעיבוד ואנליזה של תמונות

במסגרת פרק זה ננסה להגדיר מהי תמונה דיגיטלית, ומה ניתן ורצוי לבצע עליה. שאלת זו והתשבות האפשרות לה מגדירים תחום מחקר ופיתוח מרתק, המוכר בשם הכלול "עיבוד תמונות". חשוב לציין כי תחום זה קיים מלכתחילה בעיקר בשל ההתקדמות המטאורית בנושא המחשב. כל העיבודים אותם נזכיר במסגרת ספר זה הינם "דיגיטליים", דהיינו – אלגוריתמים מבוססי מחשב, וזאת בגין עיבוד אנלוגי או אופטי. לפיכך, שם ראוי יותר לספר זה יהיה "עיבוד וניתוח ספרתי של תמונות".

איפוא אנו פוגשים תמונות? מדובר כלל חשוב לדון בעיבודן?אלבום תמונות משפחתי, הטלוויזיה, האינטרנט, קריאה בעיתון, פרסום רחוב ועוד – אלה הם מקורות אשר מתרכזים בהעברת מידע ויוזאלי – אלינו. מכנה משותף למרביתם אלו ואחרים הוא רצונם לנצל את שתי המצלמות המוצבות בפנינו – העיניים. חוש הראייה מספק לנו בני האדם ולמגוון יצורים אחרים את עיקר המידע על-פיו אנו מתחממים במרחב, פעילים, ואפילו חושבים. מחקרים מדדים שחקל ניכר במוח האדם משמש לעיבוד התמונה הנקלחת בעיניהם, וניתווה. בהקשר זה, האמירה "תמונה אחת שווה אלף מילים" ממחישה את כמות האינפורמציה שנינתן לדלות מתמונה אופיינית.

כיצד המוח "מעבד" את התמונה? איזה פעולות התמונה עוברת בדרך מהחישנים בעיניהם ועד לגמר עיבוד זה? שאלות מסוג זה שאלו את עצם חוקרם במהלך מהלך מהה זה. על אף ריבוי המחקרים בתחום זה, האנושות רוחקה מלהבין את התהליכי המוחיים הכרוכים בטיפול במידע הויזואלי המזון למוח. מוטיבציה ראשונה, אם כן, למחקר ב"עיבוד תמונות" הוא הניסיון להבין, ולאחר מכן לחקות את דרך פעולה המוח ולפתור בעיות הקשורות בעיבוד מידע ויוזאלי, אשר ידוע לנו שהמוח פותר. כאמור, הבנתנו את דרך פעולה המוח בהקשר זה הינה בסיסית ביותר.

מכיוון הפק, "עיבוד תМОנות" עוסקת גם בהגדרת הקשורות בתמונות ופתרונות ע"י אלגוריתמים מוחשבים, תוך ניתוק מהשאלה על דרך פעולתו של המוח. דוקא כיון זה מניב תוצאות רבות ואיכותיות, וזהו הכוון בו "עיבוד תМОנה" זכתה בכך כלל להוות פריצת דרך.

בספר זה אנו סוקרים פרקי יסוד ב-"עיבוד תМОנות", וגישה לנו תהיה כזו המנותקת מערכות הראייה. "עיבוד תМОנה" אינו תיאוריה מגובשת! הפרקים אוטם נסקרו נשענים על תורות שונות ומקורות שונים. ניתן להתייחס ל- "עיבוד תМОנה" כארוג כלים מתוחכמים לפתרון בעיות המתעוררות לפרקים בטיפול בתמונות דיסקרטיות. זו גם הסיבה שלימוד אמיתי של "עיבוד תМОנות" חייב להיות מלאה בתרגילי מחשב בהם מתרגלים את האלגוריתמים (כלិ העובדה מארגז הכללים) הנלמדים.

בפרק זה אנו מתחילה את הדיוון בהגדרת מהוותה של תМОנה דיסקרטית, אפיון מקורוטיה ויעדיה. לאחר מכן אנו מזכירים בקצרה את מגוון הפעולות הנחוצות על תמונות, כאשרנו מסוגים משימות אלה לתחומי מחקר שונים שהתגבשו במהלך שני העשורים האחרונים. לסיום פרק זה אנו מזכירים מספר מצומצם של יישומים המערבים ברמה זו או אחרת עיבוד תמונות.

1.1 תМОנה דיגיטלית

על מנת שוכל להפעיל אלגוריתם על תМОנה, נדרש קודם להביאה אל המחשב. אלגוריתם אופייני יניח כי התМОנה הינה קובץ מחשב לכל דבר! איך נוצר קובץ זה, ואיך הוא קשור לתМОנה אמיתי? ובכן, נניח כי נתונה לנו תМОנה על נייר – נניח בשלב זה כי התМОנה היא תМОנת שחורה-לבן, כפי שמתואר בציור 1.1. הרעיון הבסיסי הינו להתייחס לתМОנה זו כפונקציה, שכל מקום במישור (x,y) תייחס "רמת אפור" $f(x,y)$. אנו נניח כי גוון שחורה מוחלט הוא גובה 0 ולבן מוחלט הוא גובה 1, וגונו האפור שביניהם יקבלו ערכים ממשיים בתחום האינטראול $[0,1]$. ציור 1.1 מראה כיצד נראה התМОנה הן כתМОנה והן כפונקציה.

על-פי הציור הנ"ל, הפינה השמאלית הילונה הינה שחורה מוחלט ולבן גובהה הוא אפס, והפינה הימנית התחתונה בעלת גובה 1 בשל היותה לבן מוחלט. בתМОנה המתוארת ישנו מעבר חלק משחור לבן, והדבר ניכר גם בתМОנה וגם ביריעה בתיאור התלת-מימדי.

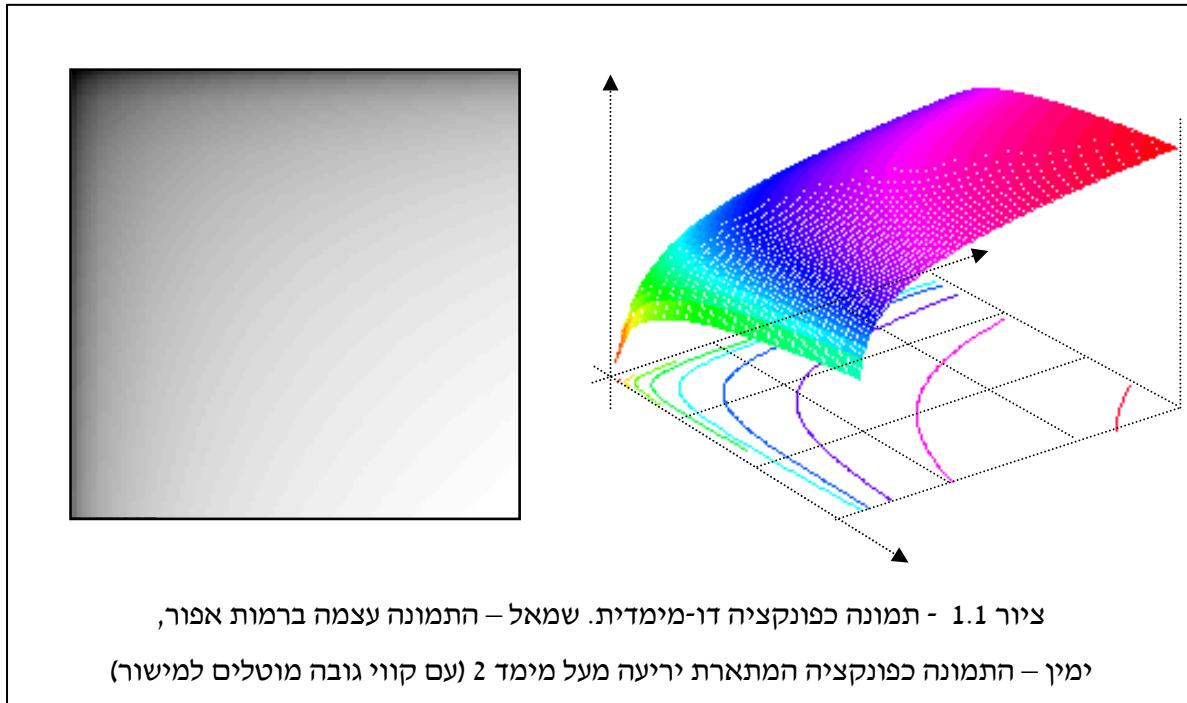
הערה: יתכן שהדפסת התМОנה (צד שמאל של ציור 1.1) ניכרת מדרגות בה. מדרגות אלה אינן חלק מההתМОנה ומקורן ב מגבלות תהליך ההדפסה, עליו עוד נדבר בהמשך.

ברור כי הפונקציה $f(x,y)$ מוגדרת מעל תחום הגדרה מוגבל Ω (תחומי התמך של התМОנה), וחסרת משמעותה לו. באופן פורמלי יותר נאמר כי התМОנה היא התאמה מהצורה הבאה:

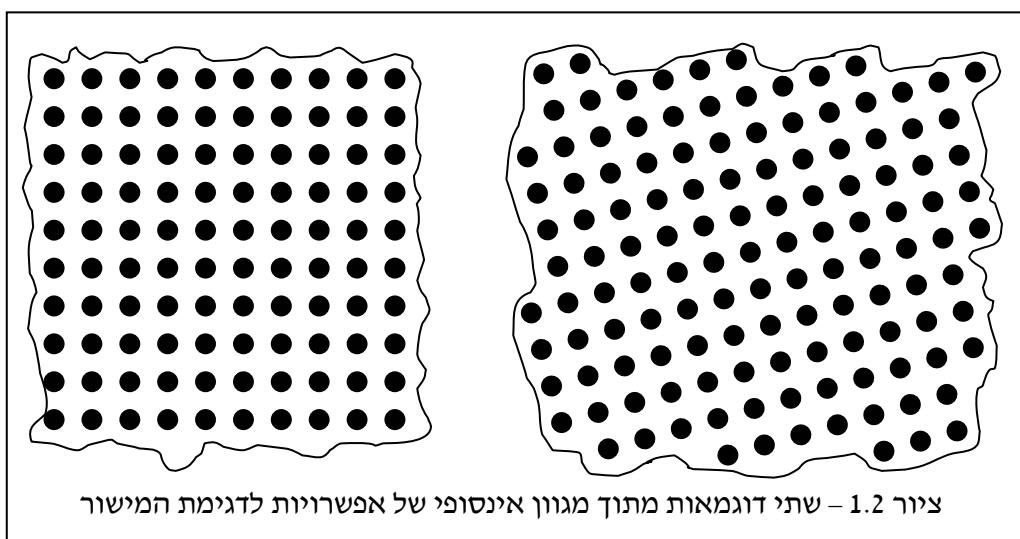
$$f : \mathcal{A}^{[0,1]} \rightarrow \mathcal{A}$$

כל זה טוב ויפה, אך בודאי לא אפשר טיפול ספרתי, בשל שתי בעיות:

1. התמונה מוגדרת מעל כל נקודה (y,x) ממשיים בתחום ההגדרה, זה אומר אינסוף בלתי ניתן למנות של נקודות.
2. בכל נקודה מקבלת הפונקציה ערך ממשי באינטרוול $[0,1]$ – ערך זה ידרוש ייצוג מדויק ע"י אינסוף סיביות.



שתי בעיות אלו מיתרגמות לצורך בכמויות אינסופיות של סיביות לשם ייצוג מלא של הפונקציה. הפתרון לבעה הראשונה הינו דגימה מרחבית. מושג הדגימה מוכר לנו בהקשר לאוטות חד-מימדיים, ואנו ראיינו שבמצבים מסוימים (כשהאות חסום תדר, ומרוחך הדגימה קבוע בהתאם לכך) ניתן לקבוע כי דגימה אינה מאבדת מידע. גם על אוטות דו-מימדיים ניתן ליחס דגימה, כשהפעם ישנו חופש גדול הרבה יותר לקביעת אופן הדגימה. ציור 1.2 ממחיש שתי צורות דגימה אפשריות, כאשר שmbdiff ביןיהם הוא ייחוס הציריים. בשני המקרים מדובר בדגימה אחידה בשני הציריים, אך בסיבוב שונה ביחס לציריהם הראשיים.



קיימות כמובן מוגוֹן אופציות אחרות לדגימה שלא מוזכרות כאן. נושא הדגימה חשוב לנו נקדים לו דיון נפרד. בשלב זה נניח כי מטעמי נוחיות בחרנו לדגום את הפונקציה במרוחה D בשני הזרים וכך קיבלנו את האות הדו-מיידי הדיסקרטי הבא:

$$f(x, y) \Big|_{x=mD, y=nD} = f(mD, nD) = f[m, n]$$

אנו נסמן בסוגרים מלכינים אתות דיסקרטיים, ובסוגרים עגולים אתות מעל הרץ' (כשאנו מאמצים, באופן זה, את סימונו של פרופי' בווע פורט בסיפורו על עיבוד אותן ספרתיים). אם תחום ההגדרה של התמונה מלכני - $[0, (N-1)D] \times [0, (M-1)D]$, אז ההפונקציה הדיסקרטית מוגדרת מעל הערכים

$$f : [m, n] \in [0, M-1] \times [0, N-1] \rightarrow \mathbb{R}^{[0,1]}$$

כך שהתמונה בנויה כמערך דו מימי של N על M אלמנטים ממשיים בתחום 0 עד 1. כאשר מרוחה הדגימה D נבחר להיות קטן, גודל התמונה גדול, ופרטים קטנים יותר ניתנים להכרה. במקרה זה אנו נאמר כי הרזולוציה של התמונה טובה יותר. רזולוציה מתייחסת לאו דווקא לגודל התמונה אלא ליכולת האבחנה בפרטים קטנים.

אם נחזור לצייר 1.1, נtake התמונה מסויים על ערכיו מתוארים בצייר 1.3. אנו רואים ערכים ממשיים כך שבפינה השמאלית העליונה הערך הנמוך ביותר, ועם ההתקדמות ימינה ולמטה, הערך גדל קלות.



נפתרו כעת את הבעיה השנייה – במקומות ליצג את רמות האפור כמספר ממשי בתחום $[0, 1]$, נמיר אותן ליצוג בעל מספר סיביות קבוע וידוע מראש. לדוגמה, נוכל להחליט כי כל רמת אפור תיויצג ב- L

סיביות ע"י המירה - $\text{round}(\frac{2^L - 1}{2})f[m, n]$. פועלה זו מעבירה כל ערך בתחום $[0, 1]$ לערך שלם בתחום $[0, 2^L - 1]$. פועלה זו קרוייה קוונטיזציה (בעברית - כימוי). ברור כי פעולה הקוונטיזציה מאבדת מידע! אבל, בזכות אובדן זה, כeut התמונה הינה מקבץ סופי של סיביות.

התמונה הדיגיטלית אותה בנו ע"י דגימה וקוונטיזציה מיוצגת ע"י MN נקודות תמונה הקרוויות פיקסלים (Pixel) - קיצור של Picture Element, וכל אחד מהם מיוצג ע"י L סיביות. לכן, סך של MNL סיביות נחוץ לאחסונה של התמונה. גדלים אופייניים לגודל תמונה יהיו $M = 50, N = 50$ לתמונה קטנה במיוחד, ועד $M = 10^5, N = 10^5$ יותר להדמאות לוין ותצלומי אויר. באשר לעומק הקוונטיזציה, ערך אופייני מאד הוא $L = 8$, אך ישנו מקורות עם ערכי L גבוהים יותר. כך יצא שטף תמונות מינימליסטיות שהוא 2,500Byte, תמונה בגודל סביר תצרוך 1Mbyte – 10Gbytes.

תיאור נפחי המידע הניל מחייב אותנו לאמירה מתחילה פרק זה על הקשר ההדוק בין ההתקדמות בעולם המחשב ו- "עיבוד תמונות". גם יכולת לאגור כמויות אלו, וכמוון יכולת לעבד נתונים רבים לגובה למחשבים. לדוגמה, גם פעולה פשוטה הכרוכה ב- 20 מכפלות לפיקסל בתמונה, הופכת להיות עצומה כאשר מדובר במליאוני פיקסלים.

כל שנאמר עד כה התייחס לתמונות שחזור-לבן. אחת הדרכים המקובלות ביותר לייצוג של תמונות צבע היא מבנה RGB. כל נקודה (x, y) בתחום ההגדלה של התמונה מקבלת לא ערך אחד, כמוום, אלא שלושה ערכים המתייחסים לעוצמות האדום (Red), הירוק (Green) וכחול (Blue). ע"י צירופים בעוצמות שונות של שלושת צבעי יסוד אלה אנו מסוגלים לייצר מגוון רחב מאוד של גוונים. לכן, לאחר דגימה וקוונטיזציה, במקומות מסוים אחד בין L סיביות, תמונה צבעונית תדרוש שלושה ערכים כאלו, ולכן יהיה גדול פי שלוש. ישנו גם ייצוגים אחרים לתמונות צבע, וכיום דיון עמוק בכלל בנושא צבע. במסגרת ספר זה נתמקד בתמונות שחזור-לבן עם ה策ות חוטפות לצבע מפעם לפעם.

כנוקדה אחרונה במסורת סעיף זה נזכיר כי כל האמור לעיל מתייחס לתמונה בודדה – Still. מה באשר לסרטים? ובכן, סרט הינו פונקציה של הזמן (t, y, x) אך גם של הזמן t . לכן, סרט הינו פונקציה במרחב 3 מהצורה $f(x, y, t)$. ממש כמו שדגמנו ב- (x, y) נוכל גם לדוגם בזמן הזמן, וכך הפוך סרט רצף – לסדרה ארוכה של תמונות Still. קצב הדגימה בזמן נקבע לפי תוכנות מערכת הרהה שלנו – מקובל כי הצגת 30 תמונות בשניה תיצור לעין אנושית (ולמוח שעבוד מאחוריה) תחושה של רצף. לכן, למשל אם תמונה בודדה הסרט היא בגודל של 10Kbyte (תמונה קטנה של 100 פיקסלים עם 8 סיביות לפיקסל), הרי סרט כזה באורך של שעה ידרוש נפח אחסון של כ- 1Gbyte. נפחים אלו הם הסיבה לעיכוב שלם ב- "עיבוד תמונה" בנושא הטיפול בוידאו. רק בעשור האחרון מיעזים להיכנס לנושא זה ולהציג כלים ואלגוריתמים לעובדות מורכבות בוידאו.

2.1 קלט/פלט לתמונות דיגיטליות

תיארנו כיצד מייצגת תמונה דיגיטלית במחשב, וראינו כי נחוצות פעולות דגימה וקונטיזציה על מנת להביאה ממצב של פונקציה רצף לקובץ מחשב תקין. איזה התקנים מבצעים פעולות אלו? כיצד? באופן דומה, בהינתן תמונה כקובץ מחשב, נרצה לעיתים להתבונן בה, אם על מנת לקבל את תוכאת העיבוד, או בשל רצוננו לראות את התמונה במהלך שלבי עיבודה. גם כאן עולה השאלה כיצד עושים זאת? במסגרת סעיף זה נציג באופן שטחי מענה לשאלות הנ"ל. המטרה המרכזית היא להבין כיצד נוצרים מסלולים שלמים המתחילה בהרכשת תמונה, העברתה למחשב לשם עיבוד, ומסתיימים בהפקת התקן פלט.

מקור אפשרי ראשוני של תמונה הוא מצלמה דיגיטלית, כגון מצלמות הוידאו הביתיות האופייניות, ומצלמות ה- Still החדשות בהן במקום סרט צילום ישנה מדיה דיגיטלית. מצלמות כאלו מורכבות סביר מטריצת חיישנים הרגיסטים לאור (כגון גלאי CCD), אשר קובעים מתחזרים חשמלי באופן ייחסי ישיר לכמויות הפוטונים הפוגעת בהם. כל חיישן מייצר ערך עבור פיקסל בודד, ולכן מטריצות חיישנים אלו צרכות להיות גדולות (מאות על מאות של חיישנים על chip בודד). המתחזרים הנוצרים מומרים ישירות ל- L הסיביות המייצגות את רמת האפור בנקודה זו וכך מתקיים ערכו של הפיקסל. פרמטר מרכזי בקביעת איקותה של מצלמה כזו הוא מספר הגלאים במטריצת החיישנים. מצלמות Still דיגיטליות עם מערך של 640 על 480 נחשבות לרמה הבסיסית ביותר (זוהי רזולוציית VGA), ואיכותן רוחקה מלהתרחות במצלמת סרט קונונציונלית. לשם קבלת איכות מתחרה הסרט צילום, על המערך להכיל כ- 2000 על 2000 חיישנים ויותר.

באלטרנטיבה למערך דו-מימדי של חיישנים, ניתן להציג מערך חד-מימדי וסרייקה מכנית (של המערך, מראה מטילה או פריזמה) לקבלת אוסף שורות היוצרות את התמונה. גישה זו מקובלת כאשר נדרשת רזולוציה גבוהה מאוד ולפיכך מערך חיישנים ארוך (בסדר גודל של אלפיים). הבעיה העיקרית עם גישה זו הוא הצורך במרכיבי מכנית מדויקת במלטה.

על מנת לקלוט תמונה צבע נדרש בעיקרו שימוש בשלושה מערכי חיישנים, שככל אחד מהם מחובר למסנן אופטי לבצע הרלוונטי. כך ייצא למשל כי שכבת ה- R תיזכר ע"י מערך גלאים עם פילטר אופטי המונע מכל צבע למעט אדום לחדרו. לכן, מצלמת צבע כזו תהיה בפועל שלוש מצלמות שונות על אותו ציר אופטי לשם קבלת התאמה מוחלטת ביןיה במובן הגיאומטרי.

כל האמור לעיל מתייחס בדרך בה פועלת מצלמה דיגיטלית אופיינית זו לתמונות Still וכן לוידאו (סדרות של תמונות ברצף). חשוב לציין כי מצלמה אלקטронית אינה מוגבלת לחישה של אור נראה – בהינתן גלאים החשים כל תופעה פיזיקלית אחרת וממיריהם אותה לאותות חשמליים, ניתן לסדרם במערך דו-מימדי ולהגדיר את תוכאת החישה בתמונה דיגיטלית לכל דבר. כך הדבר בצלום אינפרה-אדום, המספק תמונות של אורן גל אחר הקשור לחום של עצמים, כך גם בצלום רנטגן, צילומים סיסמיים של רעדות על-פני כדור הארץ, ועוד.

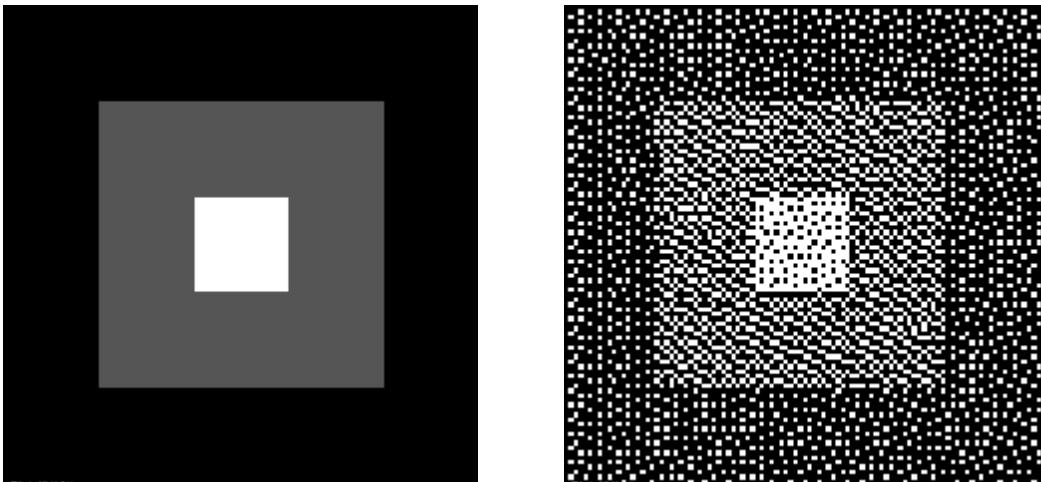
מקור מקובל אחר לתמונות הוא עותק קשה, דהיינו תצלום על נייר או על סרט צילום. הכללי המקובל לשם המרתת של תמונה כזו לקובץ הוא סורק (Scanner). סורק אינו אלא מצלמה דיגיטלית הפעלת בתנאים נוחים, ولكن מצלמה זו יכולה להיות פשוטה. התנאים הנוחים מתייחסים לכך שהעטם המצלום אינו זו, וכן אין סיבה להשתמש במערך דו-מימדי יקר – די במערך חד-מימדי הנע מצד אחד של התמונה לצידה الآخر. כמו כן, אין צורך ל-”לכוד” את התמונה בחטף, ונitin לעשות זאת בקצב קראת המידע ע”י המחשב. הצלום נעשה בתנאי תאורה מבוקרים ולבן אין צורך בעיבודים אלקטרוניים להפוך את היוצרות ולהמיר את השחור לערך גובה ואת הלבן לערך נМОך.

פרמטר מרכזי באפיון סורק היא רזולוציית הסריקה שלו, דהיינו המרווח המינימלי אשר הוא יכול לקבוע בין דגימות. סורקים אופייניים לשימושים ביוניים מספקים רזולוציה של עד 600 dpi, דהיינו 600 נקודות באינטש' אחד. לא פעם בפרסומים על סורקים מוזכרות רזולוציות בסדר גודל של dpi 9600 – חשוב להבהיר כי לרוב ערכים אלו מתייחסים לרזולוציה מלאכותית. התמונה נסרקת בפועל ברזולוציה של dpi 600, וזאת מוגדלת הלאה ע”י אינטרפולציה לגודל הנקוב. אינטרפולציה אינה תחליף לרזולוציה אמת. אנו נעמוד על ההבדלים בדיון בנושא דגימה, ולאחר מכן שכנדו בסוגיות הגדלת התמונה.

מקור שלישי ואחריו שיוזכר כאן לתמונות הוא תמונות מחושבות, או תמונות סינטטיות. צילום מכ’ם (SAR), צילום אולטראסאונד, צילום MRI וצילום CT – כל אלה יוצרים תמונות. תמונות אלו נוצרות ע”י מדידת תופעה פיזיקלית כלשהי לאו דווקא במערך דו-מימדי. עיבוד ספציפי של האות המתתקבל יוצר תמונה מלאכותית אשר חשובה ליישומים רבים. לדוגמה, בצלום MRI התמונה הנוצרת מתארת חתך (שלא ניתן היה לקבלו בדרכים אחרות) בגוף של המצלום, כשרמות האפור הקשורות לצפיפות המים בכל רקמה. גם לו חתכנו את הגוף המצלום באותו מקום בדיק, לא היינו מקבלים את אותה תמונה. כדוגמה שנייה, בצלום SAR, החזרים מכ’מים מתועדים כתמונה. מוט ברזל בשטח יתבטא בנקודה לבנה חזקה, ולעומתו, גם אם נפרוס בד ענק בצעב לבן בשטח, לא נראה אותו כלל.

עד כה סקרונו אפשרות להזנת תמונה אל המחשב. CUT נדון בקצרה בהתקני הפלט – המרת קובץ תמונה ל视窗. האפשרויות העיקריים הינם הצגה למסך או הדפסה לנייר. באשר לשימוש במסך, הגישה האופיינית היא שימוש ב- CRT (קיימות גם אפשרות אחרת כגון מסכי LCD) בהם קרן אלקטرونים פוגעת במשטח המmir פגיעה אלקטرونים בפליטות פוטוניים, תכמה קרואה פלוואוריצנטיות. ע”י שליטה אלקטронית בעוצמת קרן האלקטרונים ומיקום פגיעהה במסך, מתקבלת נקודת חומר בעוצמה נשלטת – זהה לרמת האפור הנדרשת. המסך עצמו מכוסה במערך דו-מימדי של נקודות חומר פלוואוריצנטי בצפיפות גבוהה הנותנות את התמונה. חשיבות רבה יש לשני פרמטרים במסך מעין זה – מהירות סריקת הקמן, אשר מיתרגמת ל מהירות בניתה של התמונה על המסך, ורזולוציית המסך – המרווח בין נקודות סמכות (פיקסלים) על פני המסך. הצגת תמונה צבע נעשית ע”י שימוש בשלושה צבעים – חומר פלוואוריצנטי (R, G, B) בשלשות של נקודות סמכות, ופגיעה בכל אחד מאלו בעוצמה הנכונה לשם קבלת שילוב הצבע הנחוץ.

ובאשר להדפסה, הדפסה איקוונית של תמונה פירושה המרת מטריצת הפיקסלים למגוון כל הגוונים הנחוצים על פני נייר ההדפסה. קיימות מדפסות Continuous-Tone אשר עשוות פעולה זו בדיקוק, אך השימוש בהן נדיר ויקר. מדפסות אלו פועלות כמו פיתוח על נייר צבע בצלום ביתי אופייני – חשיפת נייר רגיסטר לאור בGRADE מבקרת בכל נקודה. כתחליף לגישה יקרה זו המחייבת לא פעם נייר פיתוח, מדפסות אופייניות (InkJet ו-Laser) מציאות גישה שונה לפיה רק מקבץ מצומצם מאוד של צבעים אגוניים ניתן להפקה. על מנת ליצור את התמונה הנтונה, נעשה שימוש בגישה ה-Half-Tone, אשר יוצרת אשליה של מגוון כל הצבעים אגוניים ע"י פיזור של נקודות בגודל, פיזור וצבע (מתוך המקבץ המצומצם האפשרי) בשלטים. ציר 1.4 ממחיש גישה זו ע"י דוגמה. עבור מדפסת היכולה לצבע נקודות בשחור בלבד בגודל נקודה ובמיקום נשלט, כך נראה החרה מתמונת רמת אפור לתמונה בינהרית – ציפויות גדולה לرمות נמוכות ונמוכה לرمות גבוהות. גישה זו נהוגה במדפסות ביתיות, בתמונות עיתון ועוד.



ציר 1.4 – תמונה ברמות אפור (משמאל), ותמונה בינהרית בה נקודות שחומות על רקע לבן הבאה לחקותה

ביצוע Halftonning לצבע מורכב יותר, אך בניו על אותו עיקרון. אנו נקדים דיוון לשיטות-Half-toning Tone בהמשך.

3.1 סוגיות בטיפול בתמונות

כבר מהדיוון על בנייתה של תמונה כקובץ ע"י דגימה וקוונטיזציה, והפקתה לעותק קשה ע"י Halftonning, פגנו מספר אפליקציות המחייבות מענה ונדרונות במסגרת עיבוד תמונות. בנוסף לנושאים אלו, בהינתן תמונה דיגיטלית, ניתן לדון במגוון רחב של בעיות מעניינות אחרות ודריכים לפתרונן.

בשלב ראשון, ניתן להתייחס לתמונה כאוות דו-מימדי, ולהרחב את שלמדנו במסגרת עיבוד אותות דו-מימדי. באופן זה נוכל לשאול כיצד להציג ולבצע התמרות פוריה דו-מימדית (רציפה וdiskretiyta, כמו באוטות חד-מימדיים), כיצד לבצע סינון לתמונה, וכיצד לתכנן מסנן להגשה יעדים

מוגדרים מראש על התמונה. גם נושא הדגימה הדו-מימדית נופל למסגרת זו של עיבוד אוטות, וכפי שראינו, עלות אפשרויות מעניינות בהרחבת משפט הדגימה לדו-מימד. חשוב עם זאת להבהיר כי עיבוד אוטות דו-מימדיים אינו "עיבוד תמונה". "עיבוד תМОונות" כולל בתוכו הרבה יותר מאשר עיבוד אוטות גרייה - אנו נפגש המונע בעיות להן אין מענה מסגרת עיבוד אוטות, ויידרש טיפול ספציפי הנשען על פרקי מחקר אחרים.

בעיבוד תМОונות מקובל להגיד מספר שימושים יסוד בהם משפחות שונות של בעיות ואלגוריתמים.

נושא ראשון על עיבוד תМОונות הוא הפרק הדן בשיפור תמונה. לא אחת מתקבלת תמונה ירודה באיכותה בשל תנאי צילום לא מוצלחים, בשל פעולות לא נכוןות בעת המרתה לקובץ, או בשל תכונות בעיות של העצם המצלום. במקרים אלו נרצה להציג אלגוריתם המשפר את איכות התמונה ויוצר תМОונת פلت טוביה יותר באיזה שהוא מובן. הנקודת הבעייתית בהקשר זה היא קביעת ממד לאיכות תמונה. מגוון כלים רחב מאוד פותח במהלך השנים כדי לתת מענה לבעיות שונות בשיפור תМОונות. בדיאו על שיפור תמונה נעשית אבחנה בין צורך לשיפור תמונה שגורם קלוקלה או מוגדר, ובין בעיות שחזור בהן ידוע מה עברת התמונה בתהליך הקלקלול, ונדרשת מערכת הופכת את התהיליך במידת האפשר. אנו נקדים מספר פרקים לדיאו בנושאים אלו, ונסקור את הגישות הרווחות.

נושא שני בעיבוד תמונה הוא הפרק הדן בדיחיסט תМОונות. מהדיאו עד כה ברור כי תМОונות זו צריכה זיכרון טובעני. שאלת טבעיות שעולה לפיקח היא האם ניתן לייצג תמונה ללא פגיעה באיכותה בפחות סיביות. בדיחיסט מידע בכלל ובדיחיסט תМОונות בפרט מבחינים בין דיחיסה ללא אובדן, בה לאחר פרישה חוזרים בדיקות למקור, ובין דיחיסה אובדןית, אשר בפרישה חוזרת קרוב אך לא בדיקות למקור. אנו נראה כי ניתן בקלות לדוחס תמונה לכדי עשיירית ופחות מנפח המ庫ורי ללא שינויים ניכרים לעין. כל הדיאו בדיחיסה נסמך על תורת האינפורמציה – זה פרק חדש יחסית בתקשורות ספרטיט אשר, מطبع הדברים, ייסkr בשטחים רבים במסגרת ספר זה.

נושא שלישי בעיבוד תמונה זו בייצוגה של תמונה. פגשו כבר נושא אחד מדיאו זה – נושא ההמרה ל-Half-Tone. נושא הקוונטייזציה חיווני הן לבנייתה של תמונה דיגיטלית והן לצורך דיחיסט תמונה – נושאים אלו ואחרים ייסקו במסגרת ספר זה.

בכל הפעולות עד כה קיבלנו קובץ תמונה, ביצענו עיבוד, וקיבלו קובץ תמונה כפלט (בדיחיסה, אנו מתייחסים למערכת שדוחסת ולאחר מכן פורסת). לא אחת נרצה להכניס תמונה לעיבוד שימושיו אינם תמונה, אלא החלטה. למשל – האם יש אנשים בתמונה הנתונה, ואם כן כמה והיכן? גלה את מיקומם של הקווים היסרים בתמונה וכו'. סוגיות אלו בעיבוד תמונה משויות לתהום הקורי ראייה ממוחשבת, ואני נרחב על האבחנה בין התחומים הללו בסעיף הבא.

נושא אחרון שיידן במסגרת ספר זה הוא הטיפול בסדרות של תМОונות (סרטוי וידאו). גם בסדרות של תМОונות מדובר על שיפור, שחזור, דיחיסה, יצוג, ופעולות ראייה ממוחשבת. עניין ייחודי ומרתק שעולה בעיבוד סדרות הוא נושא התנועה בין תМОונות, ושימוש בו לשם פתרון בעיות הניל בטיב גבוה

יותר, בהשוואה לטיפול פרטני בתמונות בודדות. אנו נקדים פרק אחד לנושא זה, וניגע באופן זה בקצתו של קרחון של תוצאות מחקריות ואלגוריתמים.

חשוב להבהיר כי ספר זה רחוק מלבסות את מגוון הנושאים בעיבוד תמונות. המטרה היא סקירה ראשונית, ומטרת הפניה לעבודות המשך בתחום. עם זאת מושם דגש על עדכניות החומר, והבאת תוצאות חדשות ככל האפשר.

4. עיבוד תמונה ותחומים סטטיסטיים

הטיפול הממוחשב בתמונות אינו נחלתו הבלעדית של התחום הקרוי "עיבוד תמונות". קיימים פרקים סטטיסטיים אשר אף להם זיקה ישירה לעיבוד ספרתי של תמונות. חלוקה מקובלת כבר נרמזה בסעיף קודם, בו עמדנו על האבחנה בין עיבוד תמונה וראיה ממוחשבת. בעוד שבעיבוד תמונה, הקלט לאלגוריתם והפלט הם תמונות, הרי שראיה ממוחשבת דנה באלגוריתמים בהם הקלט הוא תמונה והפלט הינו מידע אחר.

בתוך תחום הראיה ממוחשבת מקובל להגדיר משימות ברמה נמוכה (Low-Level Vision) ובינונית (Mid-Level Vision) ו גבוהה (High-Level Vision). ראייה ממוחשבת ברמה נמוכה דנה בהפקת מסקנות ראשוניות מתוך צורה של מידע לא מעובד וחסר משמעות פיזיקלית ישירה. כך הדבר בהפקת וקטורי תנועה בין צמד תמונות, מיקום שפות או צורות פרימיטיביות אחרות כגון ישרים, פינות וכו'. ראייה ממוחשבת ברמה בינונית אלגוריתמים מסוימים מוסיקים מידע בעל ערך מתקדם יותר בעל ערך פיזיקלי, כגון מפת עומק תלת-מימדי מצמד תמונות, או תיאור של תנועת המצלמה במרחב בהתאם על התמונות הנתונות. בראייה ממוחשבת ברמה גבוהה האלגוריתמים מסוימים מסקנות דומות אנו ש כגון בתמונה זו רואים שולחן, בתמונה זו ישנים שלושה אנשים וכדומה.

בתוך הפרק הרחב הקרוי ראייה ממוחשבת מצוי פרק העוסק בזיהוי צורות – Pattern Recognition. פרק זה דן בשאלת כיצד מוצאים עצם מסוים בתמונה באופן יעיל ואמין. יש לזכור כי תמונה מהוות היטל מעולם תלת-מימדי למשור דו-מימדי, ובכך ישנו אובדן רב. בשל כך, גם משימות אלמנטריות לכוראה כגון זיהוי פניו של אדם מהתמונה, תחת האפשרות שהוא פונה קדימה או לצדדים, הינה בעיה קשה מאוד. פרק זה של זיהוי צורות שואב מתייעזרות רחבות יותר כגון "למידה חישובית", רשתות נוירונים, בינה מלאכותית (Computational Learning) ועוד.

קיימים יישומים רבים לראייה ממוחשבת ולזיהוי צורות, וניכרת חידרה חזקה של תחום זה לתעשייה ולמוסרים בעשור האחרון. נזכיר כמה דוגמאות כהן מחולקות לקטגוריות. בתחום הקרוי זיהוי תווים ניתן להזכיר קריאה אוטומטית של טפסים, איתור וקריאת בר-קוד, קריאה ממוכנת של צ'קים, המרת דף כתוב (מודפס או כתב יד) לקובץ ASCII, ומיון אוטומטי של מכתבים עפ"י קריאת המיקוד. בטיפול בתמונות רפואיות נזכיר ספירת תא דם, ספירת תא זרע ומציאת פגמים בהם, ואיתור גידולים בתמונות רפואיות. באוטומציה תעשייתית ורוביוטיקה נזכיר בקרת איכות למעגלים מודפסים, זיהוי חלקים בפס ייצור ומילויים, בקרת תנועה ע"י משוב חזותי, איתור מכשולים בתנועה ועוד. במיפוי

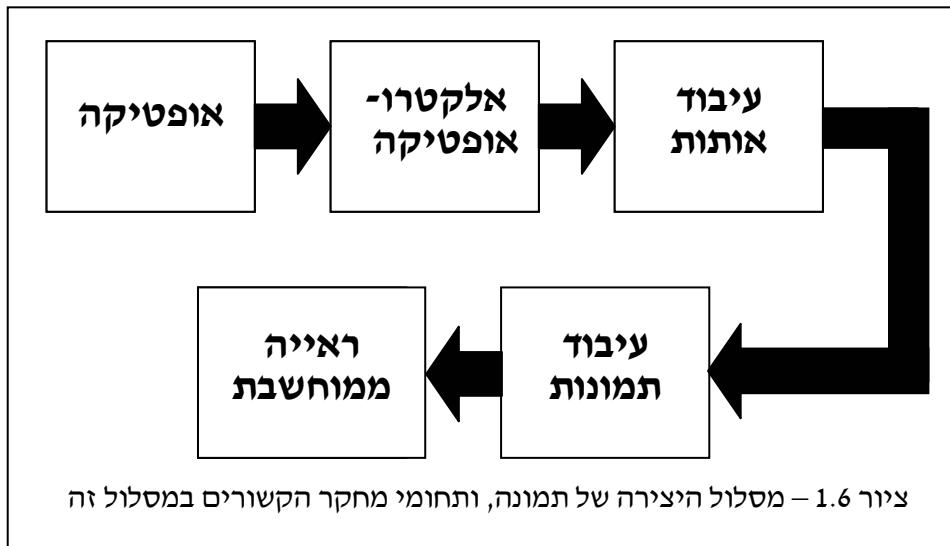
וחישה מרחוק (צילום ע"י מטוס אווני) ישנים ערכות מפות מתוך תצלום אויר, קישור תמונות שונות וואו מפות זו לזו מבחינה גיאומטרית, גילוי תנואה ושינויים בשטח, איתור מטרות ועוד. בזיהוי פלילי דנים בזיהוי טבעת אכבע תוך ייחוס למאגר קיים, זיהוי חתימה, זיהוי פנים ועוד.

בתחילת סעיף זה אמרנו ש- "שימוש תמונה" מתמקדת במשימות בהן הכניסה והיציאה הם תמונות. "ראיה ממוחשבת" לעומת זאת דנה במשימות בהן הכניסה היא תמונה והיציאה אינם תמונה. בהמשך טبعי להגדרות הניל', התהום בו הקלט אינו מידע כלשהו, והפלט הוא תמונה, קרוי גרפייה ממוחשבת. במידה רבה קיימת דואליות בין ראייה ממוחשבת וגרפייה ממוחשבת, ולהרבה בעיות שניתן להציג בפרק אחד, ניתן להציג בעיות דואליות בפרק השני. למשל, אם בראייה ממוחשבת מתוך תמונות נרצה חילוץ של מפת התרת-מידד של העצים המעורבים, הרי שבגרפייה נשאל כיצד מגופים תלת-ממדים ביצוג כלשהו בונים תמונה ריאלית (Rendering). ציור 1.5 ממחיש את המפגש בין שימוש תמונות, ראייה ממוחשבת, וגרפייה ממוחשבת.



ניתן גם להציג חלוקה אחרת של פרקים הקשורים בשימוש תמונות, המבוססת על תהליך השימוש של תמונת מרגע יצרתה ועד סיום הטיפול בה במערכת דמוית ראיית אדם. בשלב ראשון מצלמה (העין) אמורה לרכוש את התמונה, וזאת ע"י אופטיקה מתאימה (העדשה בעין), לאחר מכן התקנים אלקטרוניים (אוטם גלאי CCD למשל במכשיר, או החישנים בעין הקרויים Rods ו-Cones), ובסיום שימוש אופטיים (אותם דינמי, דגימה וקונטיוציה. ואmens, לאופטיקה, אלקטרוני, אלקטרוני-אופטיקה) ואותות אלמנטרי של תיקון תחום דינامي, דגימה וקונטיוציה. ואmens, לאופטיקה, אלקטרוני-אופטיקה. ושימוש אוטות קשור ישיר וחשוב לשימוש תמונות (שלבים אלו מתבצעים ברטינה בעין).

לאחר בנייתה של התמונה, יבואו אלגוריתמי דחיסה ושיפור איכות השיים ל- Main Stream בעיבוד תמונות (פעולות אלו מבוצעים בדרך מהעין למוח, ובמוח הראייה עצמו). לאחריהם מתחיל פועלת ניתוח (ראייה ממוחשבת) לשם הפקת מידע מהתמונה, כגון זיהוי האנשים, אבחנה בתנועה וכו'. כך אנו רואים מעבר מאופטיקה ועד ראייה ממוחשבת ולמעשה בין מלאכותית. ציור 1.6 סוקר את המסלול הנ"ל על הפרקים אותם הוא רותם.



1.5 יישומים לעיבוד ואנליזה של תמונות

לסיום פרק זה רצוננו להמחיש את חשיבותו ומעשיותו של התחום הקרויה "עיבוד תמונות". זאת נעשה ע"י מספר דוגמאות לモיצרים מערכות בהן משולבים אלגוריתמים לעיבוד תמונה, וראייה ממוחשבת.

דוגמה 1 – וידאו-פון ושיחה ע"י Net-Meeting: הרעיון ששיחת טלפון תלולה בצלום הדדי, כך שהדוברים ישוחחו זה עם זה, מלhibaנה חברות ומהנדסים קרוב לעשרים שנה, אך לאחרונה חל מפנה ממשמעותי בתחום זה, ונitinן לראות מערכות של ממש המימוש רעיון זה. הסיבה לעיכוב עד כה הייתה רוחב הפס הנדרש להעברת אותן וידאו. במיללים אחרות, כפי שראינו כבר בפרק זה, כמות הסיביות הנדרשת לתיאור סדרת תמונות היא עצומה, ולא מעשית בעוצמי תקשורת טלפון נייר גילים. לדוגמה, 30 תמונות צבע בגודל של 100 על 100 פיקסלים ידרשו שידור של קרוב ל- 1Mbyte לשניה, וזה עברו תמונה קטנה מאוד. עורך טלפון נדייב לעמודת זאת ייתן 64Kbits לשניה. ההתקנתוות האחורנית בנושא דחיסת וידאו אפשרו להמיר את הקצב המקורי בקצב הטלפון האופיני, ולהציג דחיסה ביחס של מעל 100, תוך שהסרט הנפרס משמר הרבה מכוחתו המקורי. ביום נתן לשוחח זה עם זה מעל פני האינטרנט תוכן שימושי בסיסי זה, כשהוא מלאה בפתרונות בעיות מורכבות בנושא של תזמון חבילות השידור לשם השגת שידור אפקטיבי בזמן אמיתי.

דוגמה 2 – מצלמה דיגיטלית ואלבום תמונות משפחתי ממוחשב: חברות רבות מציעות כיום מצלמה דיגיטלית כתחליף למצלמת היד הקונבנציונלית. גם במקרה זה, מדובר ב מוצר שעוכב בשל מגבלות טכנולוגיות. עד לאחרונה מקובל היה לטעון שמערכי CCD בגדים של מעל 1000 על 1000 יקרים מדי. גם נושא המדיה השומרת את התמונות זכה להתקדמות טכנולוגית בשל יכולת לדחוס את התמונות. כיום, מציעות מצלמות ברזולוציה המתקרבת לביצועי סרט צילום רגיל. הרעיון הוא לצלם ישירות לקובץ הנשמר כשהוא דחוס, וכך ניתן לשמרו של 10 עד 200 תמונות (תלו依 ביצרן, ובעוד גורמים). כל מצלמה כזו יכולה בתוכה סידרת תהליכיים בהם התמונה מעובדת לתיקון הצבעים, תיקון ההארה ועוד – אלגוריתמים קונבנציונליים המבוצעים בחומרה על ASIC.

אחת המוגבלות בשימוש בצלמות יד דיגיטליות לצלומי Still הוא הצורך בהדפסת התמונות. עד לפני כ- 5 שנים מקובל היה שהדפסה באיכות טוביה כדוגמת פיתוח מסרט לניר פיתוח היא בלתי אפשרית. ביום של ההתקדמות במדפסות הזרקת דיו, ניתן להפיק תמונות באיכות טוביה על ניר רגיל. במקביל, בשל העובדה ספרטטיבית, מוצע לנו宦 בתוכנות לאלבומים וירטואליים. באופן זה ניתן לבצע חיפושים חכמים באלבום כגון – מצא את דוד בכל התמונות באלבום, הראה לי את כל התמונות בהן דוד נלחם עומדת ליד דוד משה, וכו'ם. כל הרעיון הוא שתוכנת האלבום תבצע חיפוש מבוסס תוכן בתמונות, ולא על-פי מאפייני טקסט המלווים כל תמונה. אפליקציות מתוחכמתות לחיפושים מבוססי תוכן מוצעות בעת האחרון גם בהקשר של חיפוש תמונות ב-WEB.

דוגמה 3 – מדפסות: כבר הזכרנו שככל מדפסת מכילה בתוכה אלגוריתם להמרת תמונה מערכי הרצף – Half-Tone – ל- Continuous Tone. במדפסות סטנדרטיות גם משלבים אלגוריתמים אחרים לתיקון צבעים בתמונה, למתיחת תחום דינמי ועוד.

דוגמה 4 – גילוי פגמים בפס ייצור: נתהיל בדוגמה פשוטה – פס ייצור נושא ועליו תפוזים. על התפוזים להיכנס לארגזים כשם מסווגים לפי גודלם. מצלמת וידאו, המוצבת מעל פס הייצור ומתחברת למחשב PC רגיל ובו כרטיס דגימה, ותוכנת זמן אמת לעיבוד התמונה הנקלטת, יכולים לתת בקרה על פועלות הפס באופן שסיווג התפוזים יעשה באופן אוטומטי לחלוון. תבישה זו חלה מסתבר לא רק על תפוזים. תעשיות ה الكرטיסים המודפסים (PCB) ותעשייה המיקרו-אלקטטרונית (ייצור מעגלים על שבבי סיליקון) כבר מזמן עושות שימוש בקרה מבוססת ראייה ממוחשבת לשם פיקוח על פגמים בייצור.

דוגמה 5 – טלוויזיה דיגיטלית: כל עוד אין מוצרים בייטים המשלבים שימוש תמונה באופן המוני, הרי שנitinן לקבוע כי תחום זה אינו מצאה את הפוטנציאל שתולים בו. מגבלה זו עומדת לפוג ברגע שבכל בית תהיה טלוויזיה דיגיטלית. טלוויזיה כזו פירושה שההעברת השידורים נעשו באמצעות דיגיטלי, ולכן דחיסה ופרישה מעורבים בתהליך. יתרה מזו – ברגע שההתמונות מגויות באופן דיגיטלי אין מניעה מלבצע עליהם שימוש מתקדם. למשל, שיפור של איכות שידור במקרים בעייתיים, המרת התמונות לשידור סטריאו-איסוקופי שיאפשר ראיית עמוק, הקפאת תמונה, הגדלתה והדפסתה, אחסון סרטים במדיה דיגיטלית, ועוד.

פרק 1 - כללי

לרשימה הנ"ל ניתן להוסיף עוד יישומים כגון עיריכת וידאו ואפקטים, שירותים אינטראקטיביים לשיפור תצלומים מהאלבום המשפחתי, שחזור תמונות מצלולות כתוצאה מצלול המצלמה, ועוד. כמו כן הזכרנו יישומים רבים בהקשר לראייה ממוחשבת בסעיף קודם.

פרק 2 – עיבוד אותות בדו-מימד

15	פרק 2 – עיבוד אותות בדו-מימד
15	2.1 מערכות ליניאריות בדו-מימד - אותות מעל הרץ
19	2.2 התמרת פוריה דו-מימדית ותכונותיה
23	2.3 מערכות ליניאריות בדו-מימד – אותות דיסקרטיים
25	2.4 התמרת פוריה לאותות דיסקרטיים ותכונותיה
29	2.5 אותות ומערכות ליניאריות בתמך סופי
39	2.6 התמרת פוריה דיסקרטית (DFT) לאותות בתמך סופי

הזכירנו בפרק הקודם שימושות אינן רק עיבוד אותות דו-מימדי. עם זאת, רבים ממושגי היסוד בעיבוד תמונה, כמו מסננים, מערכות ליניאריות ושאינם ליניאריים, התמרת פוריה ועוד, מוקרמים בעיבוד אותות. בפרק זה נתמקד בהרחבת המוכר לנו מעיבוד אותות חד-מימדיים לדו-מימד. בסיסיס פרק זה ישנה ההנחה כי קיימים לקורא רקע בעיבוד אותות חד-מימדיים.

2.1 מערכות ליניאריות בדו-מימד – אותות מעל הרץ

עבור אות דו-מימיidi מעל הרץ, $f(x,y)$, המזון למערכת H , אנו מגדירים את מוצא המערכת ע"י

$$g(x,y) = H\{f(x,y)\}$$

$$H\{\alpha f_1(x,y) + \beta f_2(x,y)\} = \alpha H\{f_1(x,y)\} + \beta H\{f_2(x,y)\}$$

לכל צמד אותות $(f_1(x,y), f_2(x,y))$, ופרמטרים α, β . תכונה זו מוכרת לנו היבט מסגרת של מערכות חד-מימדיות.

מערכת H תיקרא חסרת זיכרון אם מוצאה במקום (y, x) תלוי א' ורק בערכו של א'ות הcnisha f במקומות (y, x) . בעיבוד אוטות חד-מימדיים, המשך טבעי למושג הזיכרון הוא הסיבתיות. עם זאת, סיבתיות אינה טبيعית בדיאון על אוטות דו-מימדיים, כיון שלא קיים סדר בין נקודות במישור. אם נגיד ר' מראש סדר סריקה למשור לפי ציר זמן (כגון, סריקה שורה אחר שורה), ניתן לדבר על סיבתיות, אך זו תהיה רלוונטיות רק לסריקה הנדונה. במקרה כזה בפועל המרנו את התמונה לאוטות חד-מימדיים, אז מושג הסיבתיות מתלבך עם סיבתיות חד-מימדית.

מערכת H תיקרא קבועה במקום אם תוצאה הז'אות האות במקומות בשיעור (X_0, Y_0) והז'נותו

למערכת מניבה תוצאה זהה להז'נות האות המקורי והז'נות מוצאת המערכת בשיעור דומה, דהיינו

$$H\{f(x, y)\} = g(x, y) \Rightarrow H\{f(x - X_0, y - Y_0)\} = g(x - X_0, y - Y_0)$$

דוגמה 2.1: נציג מספר דוגמאות להמחשת תוכנת הליניאריות, זיכרון, וקביעות במקום. בכל המקרים המתוירים ניתן להוכיח את הטענות השונות ישירות מתוך ההגדרות:

א. המערכת $H\{f(x, y)\} = f(x, y) + 3f(x - 2, y + 5)$ היא מערכת ליניארית. זהה המערכת קבועה במקום, עם זיכרון.

ב. המערכת $H\{f(x, y)\} = f(x, y) + g(x, y)f(x - 2, y + 5)$ (עבור g(x, y) קבוע וידוע מראש) היא ליניארית. זו מערכת שאינה קבועה במקום, וגם לה יש זיכרון.

ג. המערכת $H\{f(x, y)\} = f(x, y) \cdot f(x - 2, y + 5)$ אינה ליניארית. זו מערכת עם זיכרון והיא קבועה במקום.

ד. המערכת $H\{f(x, y)\} = f(x, y) + 1$ אינה ליניארית למקרה ההפתעה. זו מערכת חסרת זיכרון הקבועה במקום.

ה. המערכת $H\{f(x, y)\} = \text{Log}\{f(x, y)\}$ אינה ליניארית. זו מערכת חסרת זיכרון הקבועה במקום.

ו. המערכת $H\{f(x, y)\} = g(x, y) \cdot \text{Log}\{f(x, y)\}$ אינה ליניארית. בנוסף, זו מערכת תלויות מקום וחסרת זיכרון.

מערכת ליניארית H המזנת פונקציה הלאם דו-מימדית $(x - X_0, y - Y_0)$ מוגדרת במקומות את פונקציית התגובה להלאם $h(x, y, X_0, Y_0)$. במקרה הכללי של מערכת ליניארית, התגובה להלאם היא פונקציה באربعة משתנים כיון שכל כניסה דו-מימדית יש מושך דו-מימדי (ולבן שני המשתנים הראשונים), ופונקציה זו עשויה להיות תלויות מקומו הדו-מימדי של הלאם המזן למערכת (X_0, Y_0) .

מסתבר כי כל מערכת ליניארית מאופיינית לחלוון ע"י התגובה להלאם, h. דהיינו, בהינתן הפונקציה $h(x, y, X_0, Y_0)$, נוכל לקבוע מהו מוצאה של המערכת H לכל כניסה שהיא. תוכנה זו רואים מתוך כך שככל אותן $f(x, y)$ ניתן לתיאור כצירוף ליניארי של פונקציות הלאם ע"י

$$f(x, y) = \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot \delta(x - \alpha, y - \beta) d\alpha d\beta$$

בשל ליניאריות המערכת H , מתקבל מוצאה לאות $f(x, y)$ נתון ע"י האינטגרל הבא אשר זכה לשם קונבולוציה

$$\begin{aligned} H\{f(x, y)\} &= H\left\{\int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot \delta(x - \alpha, y - \beta) d\alpha d\beta\right\} = \\ &= \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot H\{\delta(x - \alpha, y - \beta)\} d\alpha d\beta = \\ &= \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot h(x, y, \alpha, \beta) d\alpha d\beta \end{aligned}$$

על פי האמור לעיל, כאשר המערכת H קבועה במקומות צריכה להתקיים כי

$$\begin{aligned} H\{f(x, y)\} = g(x, y) &= \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot h(x, y, \alpha, \beta) d\alpha d\beta \Rightarrow \\ 1. \quad H\{f(x - X_0, y - Y_0)\} &= \int_{\alpha} \int_{\beta} f(\alpha - X_0, \beta - Y_0) \cdot h(x, y, \alpha, \beta) d\alpha d\beta = g(x - X_0, y - Y_0) \\ 2. \quad g(x - X_0, y - Y_0) &= \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot h(x - X_0, y - Y_0, \alpha, \beta) d\alpha d\beta \end{aligned}$$

והשווות שתי המשוואות התחתונות מניבה לאחר אלגברה אלמנטרית

$$\begin{aligned} \forall x, y, \alpha, \beta, X_0, \& Y_0 : \quad h(x - X_0, y - Y_0, \alpha, \beta) &= h(x, y, \alpha + X_0, \beta + Y_0) \\ &\Rightarrow h(x, y, \alpha, \beta) = h(x - \alpha, y - \beta) \end{aligned}$$

משמעות תוצאה זו היא שהתגובה להלם אינה משנה את צורתה במקומות אלא זהה באותו שיעור בו זו ההלם שהוון למערכת. לכן, במקרה זה, תגובה ההלם היא פונקציה של שני משתנים בלבד. במקרה זה מוצאה המערכת מתקבל ע"י הקשר

$$\begin{aligned} H\{f(x, y)\} &= \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot h(x, y, \alpha, \beta) d\alpha d\beta = \\ &= \int_{\alpha} \int_{\beta} f(\alpha, \beta) \cdot h(x - \alpha, y - \beta) d\alpha d\beta = f(x, y) \otimes h(x, y) \end{aligned}$$

וזהו פועלות הקונבולוציה למקרה של מערכת קבועה במקומות.

דוגמה 2.2: בהמשך לדוגמה הקודמת, נתייחס למספר מקרים ונקבע את התגובה להלם עבורם :

פרק 2 – עיבוד אותות בדו-מימד

- א. ראיינו כי המערכת $H\{f(x,y)\} = f(x,y) + 3f(x-2,y+5)$ ליניארית וקבועה במקום. תגובתה להלם נתונה ע"י $h(x,y) = \delta(x,y) + 3\delta(x-2,y+5)$.
- ב. המערכת $H\{f(x,y)\} = f(x,y) + g(x,y)f(x-2,y+5)$ ליניארית, אך אינה קבועה במקום. $h(x,y,\alpha,\beta) = \delta(x-\alpha,y-\beta) + g(x,y)\delta(x-2-\alpha,y+5-\beta)$ תגובתה להלם נתונה ע"י וכפי שאנו רואים, זו אינה פונקציה של ההפרשים $(\alpha-x, \beta-y)$ בלבד.

כיצד מתבטאת מערכת חסרת זיכרון בתגובתה להלם? בהתבסס על הנאמר קודם, נובע כי על תגובת ההלם להיות $h(x,y,\alpha,\beta) = h(x,y)\delta(x-\alpha,y-\beta)$, ואז

$$\begin{aligned} H\{f(x,y)\} &= \int_{\alpha} \int_{\beta} f(\alpha,\beta) \cdot h(x,y,\alpha,\beta) d\alpha d\beta = \\ &= \int_{\alpha} \int_{\beta} f(\alpha,\beta) \cdot h(x,y)\delta(x-\alpha,y-\beta) d\alpha d\beta = f(x,y)h(x,y) = g(x,y) \end{aligned}$$

כלומר, המוצא מתקיים ע"י הכפלת הכניסה בסקלר כלשהו שעשויה להיות תלוי מקום. במקרה של מערכת קבועה במקום וחסרת זיכרון מקבלים כי המוצא הוא המכפלה בסקלר קבוע h של הכניסה.

מערכת ליניארית וקבועה במקום תיקרא פרידה (ספרבילית) אם מתקיים כי ניתן לפרק את תגובת ההלם הדו-מימדיות למכפלה של שני פונקציות חד-מימדיות $h = h_1(x)h_2(y)$. במקרה זה, מוצא המערכת נתון ע"י

$$\begin{aligned} H\{f(x,y)\} &= \int_{\alpha} \int_{\beta} f(\alpha,\beta) \cdot h(x-\alpha,y-\beta) d\alpha d\beta = \\ &= \int_{\alpha} \int_{\beta} f(\alpha,\beta) \cdot h_1(x-\alpha)h_2(y-\beta) d\alpha d\beta = \int_{\alpha} h_1(x-\alpha) \int_{\beta} f(\alpha,\beta) \cdot h_2(y-\beta) d\alpha d\beta \end{aligned}$$

ומושמות תוצאה זו היא שיש לבצע קונבולוציה חד-מימדית על כל שורה בפונקציה $f(x,y)$, ועל התוצאה שוב קונבולוציה חד-מימדית, הפעם על כל עמודה. אם גם האות f פריד נקבל

$$H\{f(x,y)\} = \int_{\alpha} f_1(\alpha)h_1(x-\alpha) d\alpha \int_{\beta} f_2(\beta) \cdot h_2(y-\beta) d\beta$$

ופירוש הדבר שבפועל בוצעו שתי קונבולוציות חד-מימדיות.

שרשור של מערכות ליניאריות מניב מערכת ליניארית – זאת ניתן להוכיח לפי הגדרה. כאשר מדובר במערכות קבועות במקום, ניתן להחילף את סדר פעולה ללא גרים מתינוי במוות המערכת הכלולת. אנו נפגש תכונה זו בהמשך, כשנדון במערכות דיסקרטיות עם תנק סופי, ונקשר אותה למכפלה מתחלפת של מטריצות.

2.2 התמרת פוריה דו-מימדית ותכונותיה

התמרת פוריה דו-מימדית לפונקציות מעל הרצף מוגדרת באופן הבא

$$F(u, v) = F\{f(x, y)\} = \iint_{x, y} f(x, y) \exp\{-j2\pi(ux + vy)\} dx dy$$

המשתנים u ו- v הם תדרים מרוחביים אופקי ואנכי. ההתמרה ההיפוכה נתונה ע"י

$$f(x, y) = F^{-1}\{F(u, v)\} = \iint_u v F(u, v) \cdot \exp\{+j2\pi(ux + vy)\} du dv$$

מלכתחילה אנו מגדירים את ההתמרה (קדימה ואחורה) עבור אותות חסומי אנרגיה המצוים ב- L^2 , ואינו מרשימים בפונקציות הנדנות אי-רציפות מסווג הראשון, בהן ערך הפונקציה קופץ בנקודת אחת לאחר מכן ומיד חוזר. הסיבה לכך היא שהמרחב בין שתי פונקציות זהות למעט סידרה (אפילו אינסופית אך ברת מניה) של אי-רציפות מסווג זה הוא 0, כך שבמרחב הפונקציות בו אנו פועלים הן נדמות זהות לכל דבר.

בטבלה 2.1 מתוארות מספר פונקציות בסיסיות והתמורות. ניכר דמיון רב בין התמרת פוריה החד-מימדית, וזו הנזונה כאן. משמעותה של הפונקציה $F(u, v)$ היא תיאור האות הדו-מימדי כהרכב של תדרים מרוחביים. התדר המרובי (v, u) מופיע באות הנדון בעוצמה $|F(u, v)|$. לכן, תמונה חלקה יחסית תורכב ממרכיבי תדר נוכחים בדרך כלל, ותמונה מרובה מעברים מהירים תורכו התדרים גבוהים.

$f(x, y)$	$F\{f(x, y)\}$
$\delta(x, y)$	1
$\delta(x - X_0, y - Y_0)$	$\exp\{-j2\pi(X_0 u + Y_0 v)\}$
$\exp\{j2\pi(xU_0 + yV_0)\}$	$\delta(u - U_0, v - V_0)$
$\exp\left\{-\pi(x^2 + y^2)\right\}$	$\exp\left\{-\pi(u^2 + v^2)\right\}$
$-1 \leq x, y \leq 1 = 1$ for $\text{rect}(x, y)$	$4 \sin(u) \sin(v) / uv$
$\text{tri}(x, y) = \text{rect}(x, y) \otimes \text{rect}(y, y)$	$\text{Sinc}^2(u, v)$
$\sum_k \sum_j \delta(x - k, y - j)$	$\sum_k \sum_j \delta(u - k, v - j)$

טבלה 2.1 – צמדים של פונקציות בדו-מימד וההתמרה הפוריה שלהם

להתמרת פוריה הדו-מימדית מספר תכונות חשובות אשר נסקור כאן ללא הוכחה. מרבית תכונות אלה מוכרות לנו מההתרמת פוריה בחד-מימד.

1. **יחידות:** לכל פונקציה $f(x,y)$ התמורה אחת ויחידה $F(u,v)$, ובביצוע ההתרמתה הפוכה, חוזרים להוותה פונקציה בדוק. אין אובדן אינפורמציה בהתרמתה קדימה או אחרת. כל זאת בכפוף לנאמר קודם על מרחב הפונקציות בו אנו פועלים.

2. **ליינאריות:** התמרת פוריה היא ליינארית, ולכן הפעלה על צירוף ליינארי של תמונות זהה להפעלה על כל תמונה בנפרד וצירוף ליינארי של התוצאות

$$\forall \alpha, \beta, f_1, f_2 : F\{ \alpha f_1(x,y) + \beta f_2(x,y) \} = \alpha F\{ f_1(x,y) \} + \beta F\{ f_2(x,y) \}$$

3. **התמרת תמונה צמודה:** הפעלת התמרת פוריה על תמונה צמודה מניבה התמורה צמודה ומהופכת בציריים, דהיינו

$$F\{f(x,y)\} = F(u,v) \Rightarrow F\{f^*(x,y)\} = F^*(-u,-v)$$

מתכוна זו עולה שעבור תמונה ממשית מתקיים כי התמרת הפוריה מקיימת $(-v, -u) = F(u, v) = F^*(-u, -v)$

4. **היפוך ציריים:** התמרת פוריה לתמונה עם היפוך ציריים מניבה תוצאה הפוכה לזו

$$F\{f(x,y)\} = F(u,v) \Rightarrow F\{f(-x,-y)\} = F(-u,-v)$$

5. **ספרביליות ההתרמה:** ההתרמה ניתנת ליישום בשני שלבים בהם פועלים בחד-מימד. חשוב להבין כי תכונה זו לא מניחה דבר על התמונה $f(x,y)$ עצמה. ניתן לראות זאת ישירות מההגדרה, לפי

$$F(u,v) = \int_x \left[\int_y f(x,y) \exp\{-j2\pi vy\} dy \right] \exp\{-j2\pi ux\} dx$$

בשל התכונה הנ"ל, כאשר אותן הכניסה פריד מתקיים כי ההתרמה אף היא פרידת, והינה מכפלת שתי התמורות פוריה חד-מימדיות

$$F(u,v) = \left[\int_y f_1(x) \exp\{-j2\pi ux\} dx \right] \left[\int_y f_2(y) \exp\{-j2\pi vy\} dy \right] = F_1(u)F_2(v)$$

.6. **תגובה הלם:** למערכת ליניארית קבועה במקום המאפיינת ע"י התגובה להלם $h(x,y)$, תגובה התדר של המערכת הליניארית נתונה ע"י $H(u,v) = F\{h(x,y)\}$. אם למערכת כזו מוכנסת תמונה הרמוניית

$$f(x,y) = \exp\{j2\pi(xU_0 + yV_0)\}$$

$$\begin{aligned} g(x,y) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{j2\pi(\alpha U_0 + \beta V_0)\} h(x-\alpha, y-\beta) d\alpha d\beta = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{j2\pi((x-\alpha)U_0 + (y-\beta)V_0)\} h(\alpha, \beta) d\alpha d\beta = \\ &= \exp\{j2\pi(xU_0 + yV_0)\} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp\{-j2\pi(\alpha U_0 + \beta V_0)\} h(\alpha, \beta) d\alpha d\beta = \\ &= \exp\{j2\pi(xU_0 + yV_0)\} \cdot H(U_0, V_0) \end{aligned}$$

משמעות תוצאה זו היא שלאות הכניסה המוצע, המוצא הינו אותו אותן כשהוא מוכפל בתגובה התדר. פיזיקלית פירוש הדבר שלכינסה הרמוניית בתדר מרחבי מסוים, המוצא ממתקנת ליניארית זהה בתדרו ומשנה רק את הפאזה והאמפליטודה בהתאם לאופי המערכת.

.7. **הגדלה והקטנה:** תמונה מוגדלת בהתאם a ו- b בשני הציריים בהתאם, תניב התמרה מוקטנת באוטם יחסים, עם הכפלה בגורם קבוע

$$F\{f(x,y)\} = F(u,v) \Rightarrow F\{f(ax,by)\} = \frac{F\left(\frac{u}{a}, \frac{v}{b}\right)}{|ab|}$$

.8. **הזזה ואפנון:** הזזה תמונה גורמת לאיפנון בתדר, ואפנון במקום יוצר הזזה בתדר

$$\begin{aligned} F(u,v) &= F\{f(x,y)\} \\ \Rightarrow F\{f(x-X_0, y-Y_0)\} &= \exp\{j2\pi(uX_0 + vY_0)\} \cdot F(u,v) \\ \Rightarrow F\{\exp\{j2\pi(xU_0 + yV_0)\} f(x,y)\} &= F(u-U_0, v-V_0) \end{aligned}$$

.9. **משפט הקונבולוציה:** קונבולוציה במקום מומרת להיות מכפלה בתדר, דהיינו,

$$\left. \begin{array}{l} F_1(u,v) = F\{f_1(x,y)\} \\ F_2(u,v) = F\{f_2(x,y)\} \end{array} \right\} F\{f_1(x,y) \otimes f_2(x,y)\} = F_1(u,v)F_2(u,v)$$

באופן דומה, מכפלה של שני אוטות דו-מימדיים במקום מותמרת לקונבולוציה בתדר, לפי הקשר

$$\left. \begin{array}{l} F_1(u,v) = F\{f_1(x,y)\} \\ F_2(u,v) = F\{f_2(x,y)\} \end{array} \right\} F\{f_1(x,y)f_2(x,y)\} = F_1(u,v) \otimes F_2(u,v)$$

10. **משפט הקורלציה:** קורלציה בין שני אותות דו-מימדיים במקומות, המוגדרת באופן הבא

$$f_1(x, y) * f_2(x, y) = \int_{\alpha} \int_{\beta} f_1(\alpha, \beta) f_2(\alpha + x, \beta + y) d\alpha d\beta$$

מומרת להיות מכפלה בתדר עם היפוך קווארדיינטות

$$\left. \begin{array}{l} F_1(u, v) = F\{f_1(x, y)\} \\ F_2(u, v) = F\{f_2(x, y)\} \end{array} \right\} F\{f_1(x, y) * f_2(x, y)\} = F_1(u, v) F_2(-u, -v)$$

באופן דומה, מכפלה של שני אותות דו-מימדיים במקומות עם היפוך ציריהם מותמרה לקורלציה בתדר.

11. **שיעור מכפלה פנימית:** מכפלה פנימית בין שתי תכונות היא בעלת אותו ערך בתדר ובמשורר המקום, דהיינו

$$\begin{aligned} F_1(u, v) &= F\{f_1(x, y)\}, \quad F_2(u, v) = F\{f_2(x, y)\} \\ \Rightarrow \langle f_1(x, y), f_2(x, y) \rangle &= \int_x \int_y f_1(x, y) f_2^*(x, y) dx dy = \\ &= \langle F_1(u, v), F_2(u, v) \rangle = \int_u \int_v F_1(u, v) F_2^*(u, v) du dv \end{aligned}$$

במקרה פרטי של תכונה זו מקבלים את משפט פרסלבל לפיו האנרגיה של האות זהה במקומות ובתדר

$$F(u, v) = F\{f(x, y)\} \Rightarrow \int_x \int_y |f(x, y)|^2 dx dy = \int_u \int_v |F(u, v)|^2 du dv$$

מדוע התמרת פוריה חשובה כל כך בדיאןמערכות ליניאריות קבועות במקומות? תשובה מסוימת לכך נרמזה בתכונת הנ"ל, ובעיקר תכונה 6 בהתייחס לתגובה התדר. בהינתן תמונה $f(x, y)$ ניתן להתייחס אליה כפונקציה הפרוסה על פני הבסיס הטריוויאלי של פונקציות דלתה (של דיראך) בכל מקום (y, x) . קבוצה זו של פונקציות מהוות בסיס אורתונורמלי הפרוס את מרחב הפונקציות ב- L_2 . התמרת פוריה מציעה החלפת בסיס זה בקבוצת פונקציות אורתונורמליות אחרות הפרוסות את מרחב הפונקציות – הפונקציות הרמוניות. יתרונו המשמעותי של הבסיס החדש הוא היותו עצמי למערכת ליניארית קבועה במקומות. פירוש הדבר שביצוג החדש, כל שעליינו לעשות על מנת לתאר את פעלת המערכת הוא לתאר כיצד "מוגברת" כל פונקציה הרמנית כזו, וזוו בדיקוק פונקציית תגובה התדר אותה פגשנו. ההסבר הנ"ל נשען על תחום במתמטיקה הקרוי אנליזה פונקציונלית. הסבר "נעימים" יותר של התכונה הנ"ל ייינתן בדיאן באמצעות דיסקרטיים, ואז נקשר את הדברים לאלגברה ליניארית, הקללה יותר להבנה.

2.3 מערכות ליניאריות בדו-מימד – אותות דיסקרטיים

כבר רأינו כי ע"י דגימה עברנו מאות מעל הרצף $(y, f(x, m))$. בקורסון, כל שאמרנו בתחילת פרק זה תקף גם כאן עם שינויים מתאימים. נחזור לפיקח בקצרה על עיקרי הדברים.

לאות $[f(m, n)] = H\{f(m, n)\}$, המזון למערכת H , אנו מגדירים את מוצא המערכת ע"י מערכת ליניארית אמ"ס $H\{\alpha f_1(m, n) + \beta f_2(m, n)\} = \alpha H\{f_1(m, n)\} + \beta H\{f_2(m, n)\}$. מערכת ליניארית H המזנת פונקציה הלם דו-מימדית $\delta[m - M_0, n - N_0]$ מניבה במושגתה את פונקציה התגובה להלם $[h(m, n, M_0, N_0)]$. במקרה למקרה הרציף, פונקציה דלתה כאן פירושה '1' במקומות השיא ואפס אחריו. מערכת ליניארית דיסקרטית מאופיינת לחולוטין ע"י תגובתה להלם. זאת רואים ע"י ייצוגה של התמונה כצירוף ליניארי ממושך של פונקציות הלם דיסקרטיות

$$f(m, n) = \sum_k \sum_j f(k, j) \cdot \delta[m - k, n - j]$$

ובשל ליניאריות המערכת H , מתקבל מושגתה לאות $[n, m] f$ נתון ע"י הסכום

$$H\{f(m, n)\} = H\left\{ \sum_k \sum_j f(k, j) \delta[m - k, n - j] \right\} = \sum_k \sum_j f(k, j) h(m, n, k, j)$$

כאשר המערכת H קבועה במקומות צריך להתקיים כי התגובה להלם היא פונקציה שאינה תלויות במקומות, ולכן היא פונקציה של שני משתנים בלבד $h[m - k, n - j] = h[m - k, n, k, j]$. מושגתה המערכת במקרה זה נתון ע"י $H\{f(m, n)\} = \sum_k \sum_j f(k, j) h[m - k, n - j] = f(m, n) \otimes h[m, n]$ וזו הפעולה הקונבולוציה למקרה של מערכת קבועה במקומות.

דוגמה 2.3: נדגים כיצד מיושמת קונבולוציה דו-מימדית דרך דוגמה. נניח לצורך פשטות קונבולוציה בין שני האותות הבאים :

$$f_1[m, n] = \begin{cases} 1 & -1 \leq m, n \leq 1 \\ 0 & \text{Otherwise} \end{cases}, \quad f_2[m, n] = \begin{cases} 1 & 0 \leq m, n \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

$$\sum_k \sum_j f_1[k, j] f_2[m - k, n - j] = f_1[m, n] \otimes f_2[m, n] = g[m, n]$$

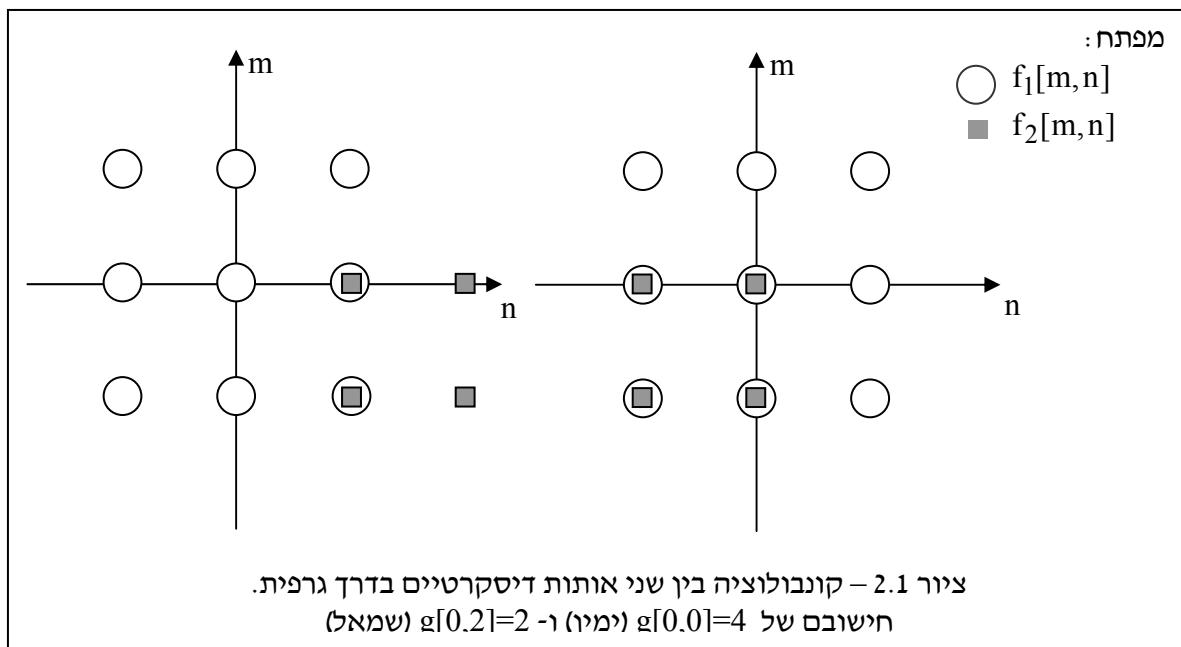
הפעולה הניל' יכולה להיעשות באופן גרפי ע"י ציור הפונקציה f_1 כמוות שהיא, ועל-גבייה ציור של הפונקציה f_2 כשהיא מהופכת בשני ציריה ומוחזק. גודל ההזזה קובע את מיקום המוצא. ערכו של המוצא

נקבע ע"י סיכון של מכפלות בין שני האותות המצויריים. למשל, כדי חישובם של $[0,0]$ ו- $[0,2]$ מתואר בציור 2.1. בגישה אלגברית נקבל

$$\begin{aligned} g[0,0] &= \sum_k \sum_j f_1[k,j] f_2[0-k,0-j] = \\ &= f_1[-1,-1]f_2[1,1] + f_1[-1,0]f_2[1,0] + f_1[-1,1]f_2[1,-1] + \\ &\quad f_1[0,-1]f_2[0,1] + f_1[0,0]f_2[0,0] + f_1[0,1]f_2[0,-1] + \\ &\quad f_1[1,-1]f_2[-1,1] + f_1[1,0]f_2[-1,0] + f_1[1,1]f_2[-1,-1] = 1+1+1+1 = 4 \end{aligned}$$

ובאופן דומה, $.g[0,2]=2$

מערכת ליניארית חסרת זיכרון תואופיין ע"י הקשר $f[m,n]h[m,n]=g[m,n]$. מערכת חסרת זיכרון וקבועה במקומות תהיה פשוט הכפלת הכניסה בקבוע. סדרביליות של מערכת זו מימדיות פירושה כי מתקיים

$$h[m,n] = h_1[m]h_2[n]$$


- דוגמה 2.4 :** נציג מספר מערכות דיסקרטיות ונדון בתכונותיהן. נעשה זאת בתאיימות לדוגמה 2.1.
- המערכת $H\{f[m,n]\} = f[m,n] + 3f[m-2,n+5]$ היא מערכת ליניארית. זהה ליניארית במקומות, עם זיכרון. תגובתה להלם היא $h[m,n] = \delta[m,n] + 3\delta[m-2,n+5]$.
 - המערכת $H\{f[m,n]\} = f[m,n] + g[m,n]f[m-2,n+5]$ (עבור $g[m,n]$ קבוע וידוע מראש) היא ליניארית. זו מערכת שאינה קבועה במקומות, וגם לה יש זיכרון. תגובתה להלם היא $h[m,n] = \delta[m,n] + g[m,n]\delta[m-2,n+5]$.
 - המערכת $H\{f[m,n]\} = f[m,n] \cdot f[m-2,n+5]$ אינה ליניארית. זו מערכת עם זיכרון והיא קבועה במקומות.

- א. המערכת $H\{f[m,n]\} = f[m,n] + 1$ אינה ליניארית. זו מערכת חסרת זיכרון הקבועה במקומות.
- ב. המערכת $H\{f[m,n]\} = \text{Log}\{f[m,n]\}$ אינה ליניארית. זו מערכת חסרת זיכרון הקבועה במקומות.
- ג. המערכת $H\{f[m,n]\} = g[m,n] \cdot \text{Log}\{f[m,n]\}$ אינה ליניארית. בנוסף, זו מערכת תלויית מקום וחסרת זיכרון.
- ד. המערכת $H\{f[m,n]\} = f[m + 0.5, n + 0.5]$ אינה מוגדרת כלל, כי אין משמעות להזזה בחצי מרוחך דגימה באוטות דיסקרטיים.

מערכת תוגדר כמערכת FIR (Finite Impulse Response) אם תגובתה להלם בעלת תמק סופי. מערכות כאלה לא קיימת בעיות יציבות בשל היותן עם אפסים בלבד. מערכת שאינה FIR מוגדרת כמערכת IIR (Infinite Impulse Response). מערכת IIR כוללת גם קטבים ולכון מתוערת בעיות יציבות. בעוד שבמערכות בימד אחד ניתן לחשב את הקטבים ע"י פקטורייזציה של פולינום, הרי שבימד זה בלתי אפשרי, והדבר מצביע בעיה קשה. כפי שאנו נראה, בשל כך מסנני FIR מאוד פופולריים בתמונה, בעוד שמסנני IIR נדרירים יותר.

4.2 התמרת פוריה דיסקרטית ותבונותיה

נניח כי מתוך האות $f[x,y]$ נוצר האות הדיסקרטי $\tilde{f}[m,n]$ על ידי פעולה דגימה במרוחך ייחידה בשני הצירים. נניח בנוסף כי דגימה זו התקבלה ע"י הכפלת התמונה הרציפה במרחב הלמים דו-מימדי. לכן

$$\tilde{f}[m,n] = f(x,y) \cdot \sum_m \sum_n \delta(m-x, n-y)$$

לפי רישום זה, כל דגימה היא פונקציה דلتה בגובה של ערך הפונקציה, ולכון הסימון השונה עם ה- \tilde{f} התמרת פוריה הרציפה עלאות זה תיתן

$$\begin{aligned} F\{\tilde{f}[m,n]\} &= \int_x \int_y f(x,y) \sum_m \sum_n \delta(m-x, n-y) \exp\{-j(ux+vy)\} dx dy = \\ &= \sum_m \sum_n \int_x \int_y f(m,n) \delta(m-x, n-y) \exp\{-j(ux+vy)\} dx dy = \\ &= \sum_m \sum_n f[m,n] \exp\{-j(um+vn)\} \end{aligned}$$

על בסיס זה, התמרת פוריה דו-מימדית לפונקציות דיסקרטיות $f[m,n]$ מוגדרת באופן הבא

$$F(\theta_1, \theta_2) = F_D\{f[m, n]\} = \sum_m \sum_n f[m, n] \exp\{-j(\theta_1 m + \theta_2 n)\}$$

המשתנים θ_1 ו- θ_2 הם תדרים מרוחביים אופקי ואנכי, כפי שקיבלו גם עבור המקרה הרציף. עם זאת קיים הבדל אחד מהותי – בעוד שבמקרה הרציף נעו תדרים אלו בתחום $[-\infty, +\infty]$, הרי שהפעם תדרים אלו מצויים בתחום $[\pi, -\pi]$, ולכל ערך מוחזק בתחום זה, מתיקיות מוחזירות. ההתרמה ההפוכה נתונה ע"י

$$f[m, n] = F_D^{-1}\{F(\theta_1, \theta_2)\} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(\theta_1, \theta_2) \cdot \exp\{+j(\theta_1 m + \theta_2 n)\} d\theta_1 d\theta_2$$

כשندון במשפט הדגימה אנו נרחב על הקשר הקיים בין הגדרות אלו ובין ההגדרות שניתנו להתרמת פוריה לאותות רציפים. לכל התכונות שהציגו בסעיף קודם יש תואם בהתרמת הפוריה הנדונה כאן לאותות דיסקרטיים. נعبر על תכונות אלו בקצרה. בחרנו להשמיט את הדיוון בתכונת ההקטנה והגדלה, כיוון שבה לא מתיקיינית דואליות פשוטה.

1. **יחידות:** לכל פונקציה $f[m, n]$ התרמה אחת ויחידה $F(\theta_1, \theta_2)$, ובביצוע ההתרמת ההפוכה, חוזרים להאותה פונקציה בדיקוק. אין אובדן אינפורמציה בהתרמת קדימה או אחרת. צריך להיזהר ולומר כי תיתכנה שתי פונקציות שונות $f_1(x, y)$ ו- $f_2(x, y)$ אשר לאחר פעולה הדגימה הן זהות, ולכן התמרת הפוריה על האותות הדוגמים יהיו זהים. אנו נזכיר תופעה זו בהמשך, כשהנדון בדגימות אחרות דו-ממדיים.

2. **ליניאריות:** התרמת פוריה היא ליניארית, ולכן הפעלה על צירוף ליניארי של תמונות זהה להפעלה על כל תמונה בנפרד וצירוף ליניארי של התוצאות

$$F_D\{\alpha f_1[m] + \beta f_2[n]\} = \alpha F_D\{f_1[m]\} + \beta F_D\{f_2[n]\}$$

3. **התמרת תמונה צמודה:** הפעלת התרמת פוריה על תמונה צמודה מניבה התרמה צמודה ומהופכת בצירום, דהיינו

$$F_D\{f[m, n]\} = F(\theta_1, \theta_2) \Rightarrow F_D\{f^*[m, n]\} = F^*(-\theta_1, -\theta_2)$$

ולכן, עבור תמונה ממשית מתקיים כי התרמת הפוריה מקיימת $F(\theta_1, \theta_2) = F^*(-\theta_1, -\theta_2)$

4. **היפוך צירום:** התרמת פוריה לתמונה עם היפוך צירום מניבה תוצאה הפוכה לזו

$$F\{f[m, n]\} = F(\theta_1, \theta_2) \Rightarrow F\{f[-m, -n]\} = F(-\theta_1, -\theta_2)$$

.5. **ספרביליות ההתרמה:** ההתרמה ניתנת ליישום בשני שלבים בהם פועלים בחד-מיינד. כאמור, תכונה זו לא מניחה דבר על התמונה $[n][m]$ עצמה. ניתן לראות זאת ישירות מההגדרה, לפי

$$F(\theta_1, \theta_2) = \sum_n \left[\sum_m f[m, n] \exp\{-j\theta_1 m\} \right] \exp\{-j\theta_2 n\}$$

בשל התכונה הניל, כאשר אות הכניסה פריד מתקיים כי ההתרמה אף היא פרידה, והינה מכפלת שתי התרמות פוריה חד-מימדיות

$$F(\theta_1, \theta_2) = \left[\sum_m f_1[m] \exp\{-j\theta_1 m\} \right] \left[\sum_n f_2[n] \exp\{-j\theta_2 n\} \right]$$

.6. **תגובה הלם:** למערכת ליניארית דיסקרטית קבועה במקומות המאופיינת ע"י התגובה להלם $[n][m], h[m, n]$, תגובה התדר של המערכת הליניארית נתונה ע"י $H(\theta_1, \theta_2) = F_D\{h[m, n]\}$. אם למערכת כזו מוכנסת תמונה הרמוניית $f[m, n] = \exp\{j(mU_0 + nV_0)\}$

$$g[m, n] = \sum_k \sum_i h[k, i] \exp\{j((m-k)U_0 + (n-i)V_0)\} = \exp\{j(mU_0 + nV_0)\} \cdot H(U_0, V_0)$$

משמעות תוצאה זו היא שלאות הכניסה המוצעת, המזקע הינו אותו אותן כשהוא מכפל בתגובה התדר. פיזיקלית פירוש הדבר שלכניסה הרמוניית בתדר מריחבי מסוים, המזקע המערכת ליניארית זהה בתדרו ומשנה רק את הפאזה והאמפליטודה בהתאם לאופי המערכת.

.7. **הזה ואפנון:** הזזה תמונה גורמת לאיפונו בתדר, ואפנון במקום יוצר הזזה בתדר

$$\begin{aligned} F(\theta_1, \theta_2) &= F_D\{f[m, n]\} \\ \Rightarrow F_D\{f(m - M_0, n - N_0)\} &= \exp\{j(\theta_1 M_0 + \theta_2 N_0)\} \cdot F(\theta_1, \theta_2) \\ \Rightarrow F_D\{\exp\{j2\pi(mU_0 + nV_0)\}f[m, n]\} &= F(\theta_1 - U_0, \theta_2 - V_0) \end{aligned}$$

.8. **משפט הקונבולוציה:** קונבולוציה במקומות מומרת להיות מכפלה בתדר, דהיינו,

$$\left. \begin{aligned} F_1(\theta_1, \theta_2) &= F\{f_1[m, n]\} \\ F_2(\theta_1, \theta_2) &= F\{f_2[m, n]\} \end{aligned} \right\} F\{f_1[m, n] \otimes f_2[m, n]\} = F_1(\theta_1, \theta_2)F_2(\theta_1, \theta_2)$$

באופן דומה, מכפלה של שני אותות דו-מימדיים במקומות מותמרת לקונבולוציה בתדר, לפי הקשר

פרק 2 – שימוש אוטות בדו-מימד

$$\left. \begin{array}{l} F_1(\theta_1, \theta_2) = F\{f_1[m, n]\} \\ F_2(\theta_1, \theta_2) = F\{f_2[m, n]\} \end{array} \right\} F\{f_1[m, n]f_2[m, n]\} = \frac{1}{4\pi^2} F_1(\theta_1, \theta_2) \otimes F_2(\theta_1, \theta_2)$$

. 9. **משפט הקורלציה:** קורלציה בין שני אוטות דו-מימדיים במקום המוגדרת באופן הבא

$$f_1[m, n]^* f_2[m, n] = \sum_k \sum_j f_1[k, j] f_2[m+k, n+j]$$

מומרת להיות מכפלה בתדר עם היפוך קווארדיינטות

$$\left. \begin{array}{l} F_1(\theta_1, \theta_2) = F\{f_1[m, n]\} \\ F_2(\theta_1, \theta_2) = F\{f_2[m, n]\} \end{array} \right\} F\{f_1[m, n]^* f_2[m, n]\} = F_1(\theta_1, \theta_2) F_2(-\theta_1, -\theta_2)$$

באופן דומה, מכפלה של שני אוטות דו-מימדיים במקום עם היפוך ציריים מותאמת לקורלציה בתדר.

. 10. **שמור מכפלה פנימית:** מכפלה פנימית בין שתי תമונות היא בעלת אותו ערך בתדר ובמשור המוקום, דהיינו

$$\begin{aligned} F_1(\theta_1, \theta_2) &= F_D\{f_1[m, n]\}, \quad F_2(\theta_1, \theta_2) = F_D\{f_2[m, n]\} \\ \Rightarrow \langle f_1[m, n], f_2[m, n] \rangle &= \sum_m \sum_n f_1[m, n] f_2[m, n] = \\ &= \langle F_1(\theta_1, \theta_2), F_2(\theta_1, \theta_2) \rangle = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F_1(\theta_1, \theta_2) F_2^*(\theta_1, \theta_2) d\theta_1 d\theta_2 \end{aligned}$$

כמקרה פרטי של תכונה זו מקבלים את משפט פרסל לפי האנרגיה של האות זהה במקום ובתדר

$$F(\theta_1, \theta_2) = F_D\{f[m, n]\} \Rightarrow \sum_m \sum_n |f[m, n]|^2 = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} |F(\theta_1, \theta_2)|^2 d\theta_1 d\theta_2$$

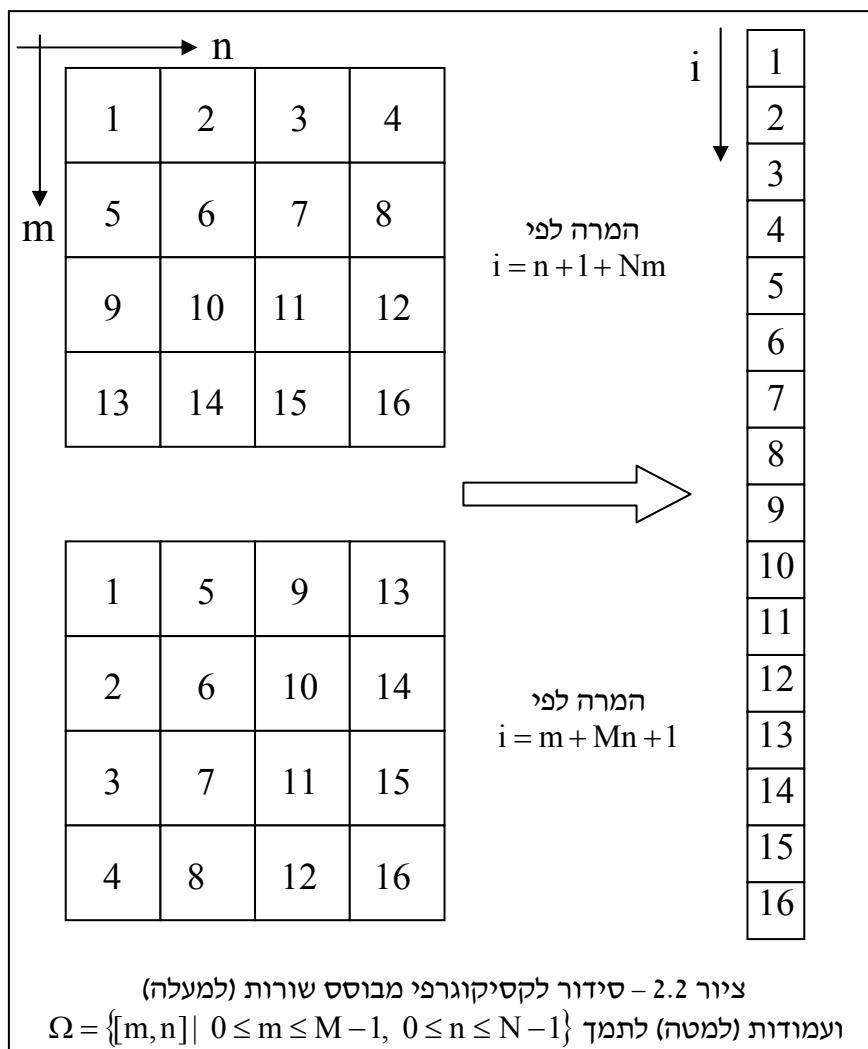
. 11. **מחזוריות:** כפי שכבר אמרנו קודם, הפונקציה $F(\theta_1, \theta_2)$ מחזורת ב- 2π בכל אחד מציריה, ולכן
ערכיה מעניינים אותנו רק בリיבוע $[-\pi, +\pi] \times [-\pi, +\pi]$

2.5 אוטות ומערכות ליניאריות בתמך סופי

עד כה עסקנו באוטות מעל השrieg הדגום $[n,m]$, אך לא הגבלנו את tamk התמונה. נניח כי התמונה עליה רצוננו לפעול נתונה לנו כ- $[n,m]$ ומוגדרת מעל השrieg הכוול $\Omega = \{[m,n] \mid 0 \leq m \leq M-1, 0 \leq n \leq N-1\}$ נקודות בדיק. נניח כי נתונה לנו מערכת ליניארית דיסקרטית H המאופיינת ע"י התגובה להלם $h[m,n,k,j]$. מוצא המערכת נתון ע"י

$$g[m,n] = H\{f[m,n]\} = \sum_k \sum_j f[k,j] h[m,n,k,j]$$

עד כה לא ראיינו כל שוני ביחס לנאמר בסעיף קודם. כיוון שהאות $f[m,n]$ כולל כמות סופית של דגימות, נוכל לסדרו כוקטור באורך MN . זאת נוכל לעשות בין היתר גם ע"י שרשור השורות או שרשור העמודות בזו אחר זו. תהליך זה קריוי column/row stack או סידור לקסיקוגרפי (lexicographic ordering) עפ"י שורות העמודות. ציור 2.2 ממחיש תהליך סידור זה.



פרק 2 – עיבוד אותות דו-מימדי

היתרון הגדול בדיאו על תכונות דיסקרטיות עם תמק סופי הוא יכולת לעשות שימוש בכלים מאלgebra ליניארית, כפי שנראה מיד. נטען כאן את הטענה הבאה: כל מערכת ליניארית H (אשר כזכור מאופיינת ע"י תגובתה להלם $[j, m, h]$) ניתנת לתיאור כמטריצה. בהינתן תכונת כניסה בגודל $M_1 \times N_1$, מומרת תמונה זו לוקטור בגודל N_1 . על תמונה זו מופעלת מערכת ליניארית ומוצאה היא תמונה בגודל $N_2 \times M_2$, המומרת לוקטור בגודל N_2 . המטריצה המתארת את פועלות המערכת תהיה בגודל של $M_2 \times N_1$ שורות ו- N_1 עמודות. תוכן מטריצה זו נקבע באופן חד-חד ערכי ע"י תגובת הלם של המערכת. רעיון זה של ייצוגה של מערכת ליניארית ע"י מטריצה אינו חדש ונדון כבר באאותות ומערכות חד-מימדיים. כאן נרחיב את הדיאו לאותות דו-מימדיים.

טענה זו אינה כה דרמטית כפי שהיא נשמעת! אם נסמן ב- \underline{f} את התמונה לאחר הסידור הלקסיקוגרפי, וב- \underline{H} את מטריצת המערכת, המכפלה $\underline{H}\underline{f}$ יוצרת וקטור חדש שיסומן $\underline{g} = g$. דבר ראשון בו נבחין הוא שלמערכת נכנסת תמונה ויצאה תמונה, כשתיהן בגדים הנכונים, ושתייהן מסודרות בסידור לקסיקוגרפי. האיבר ה- i בוקטור \underline{g} בניו מכפלה של השורה ה- i מהמטריצה H בוקטור \underline{f} , לפי

$$g_i = \sum_{k=1}^{M_1 N_1} H_{i,k} f_k$$

ובכך מתקיים שכל פיקסל בתמונה המוצא הוא צירוף ליניארי של פיקסלים מהתמונה הכניסה. מה באשר לתוכן של המטריצה H ? כיצד נקבע אותו? ענה על שאלת זו דרך דוגמה בסיסית.

דוגמה 2.5: נניח כי אנו עוסקים בתכונות $f[m, n]$ ו- $g[m, n]$ מעל השרג $\Omega = \{(m, n) | 0 \leq m \leq 1, 0 \leq n \leq 1\}$ (כלומר, תכונות של 2 על 2 פיקסלים). ונניח אופרטור כללי

$$g[m, n] = H\{f[m, n]\} = \sum_k \sum_j f[k, j] h[m, n, k, j]$$

נניח כי אנו עוסקים בסידור לקסיקוגרפי לפי עמודות. נמחיש תחילה כיצד תיראה השורה הראשונה. שורה זו בונה את האיבר ה- $[0, 0]$ של $[n, m, g]$. לפי הקשר הנ"ל מתקיים

פרק 2 – עיבוד אותות בדו-מימד

$$g[0,0] = \sum_k \sum_j f[k,j] h[0,0,k,j] = \\ = f[0,0]h[0,0,0,0] + f[1,0]h[0,0,1,0] + f[0,1]h[0,0,0,1] + f[1,1]h[0,0,1,1] =$$

$$= \begin{bmatrix} h[0,0,0,0] & h[0,0,1,0] & h[0,0,0,1] & h[0,0,1,1] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

לא קשה להשתכנע שהמטריצה \mathbf{H} בכללה נותנת את הקשר הבא

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[0,1] \\ g[1,1] \end{bmatrix} = \begin{bmatrix} h[0,0,0,0] & h[0,0,1,0] & h[0,0,0,1] & h[0,0,1,1] \\ h[1,0,0,0] & h[1,0,1,0] & h[1,0,0,1] & h[1,0,1,1] \\ h[0,1,0,0] & h[0,1,1,0] & h[0,1,0,1] & h[0,1,1,1] \\ h[1,1,0,0] & h[1,1,1,0] & h[1,1,0,1] & h[1,1,1,1] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

כהמשך של דוגמה זו, נניח כעת כי את תמונה המוצא אנו רוצים לקבל מעל תמק' שונה, בו יש שישה ערכי מוצא במקומות ארבעה - $\Omega_g = \{[m,n] \mid 0 \leq m \leq 2, 0 \leq n \leq 1\}$.

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[2,0] \\ g[0,1] \\ g[1,1] \\ g[2,1] \end{bmatrix} = \begin{bmatrix} h[0,0,0,0] & h[0,0,1,0] & h[0,0,0,1] & h[0,0,1,1] \\ h[1,0,0,0] & h[1,0,1,0] & h[1,0,0,1] & h[1,0,1,1] \\ h[2,0,0,0] & h[2,0,1,0] & h[2,0,0,1] & h[2,0,1,1] \\ h[0,1,0,0] & h[0,1,1,0] & h[0,1,0,1] & h[0,1,1,1] \\ h[1,1,0,0] & h[1,1,1,0] & h[1,1,0,1] & h[1,1,1,1] \\ h[2,1,0,0] & h[2,1,1,0] & h[2,1,0,1] & h[2,1,1,1] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

כשניתן לראות כי ארבע מהשורות במשוואת מטריצית זו חוזרות בדיק על קשרים מהמשמעות הקודמת.

мотוך הדוגמה הניל גם ברור מה יקרה בתמונות גדולות יותר. $H[m,n,k,j]$ עבר $[n,m,k,j]$ קבועים יملאו את השורה המתייחסת לאיבר ה- $[n,m]$ במוצא. מספירה של שורה זו יהיה במקרה הכללי $m+1+Mn = i$. הערכים בתוך השורה יסודרו לפי הסידור הלקסיקוגרפי של האינדקסים של תמונות הכניסה.

תכוна מעניינת שנובעת מהתייאור הניל היא נושא הצירוף בין מערכות. נניח כי נתונות לנו שתי מערכות ליניאריות (בתיאור מטריצי שלهن ולגודל זהה של תמונה כניסה ווצאה) \mathbf{H}_1 ו- \mathbf{H}_2 . הפעלת שתיהן המערכות על אותה כניסה וסכום המוצאים ייתן

$$\underline{g}_1 = \mathbf{H}_1 \underline{f} , \quad \underline{g}_2 = \mathbf{H}_2 \underline{f} \Rightarrow \underline{g}_1 + \underline{g}_2 = \mathbf{H}_1 \underline{f} + \mathbf{H}_2 \underline{f} = [\mathbf{H}_1 + \mathbf{H}_2] \underline{f}$$

כך שסיכום מוצאי המערכות כרוך בעצם בסיכום המטריצות. זה לא מפתיע כי סיכום המערכות כמוותו כמו סיכום פונקציות התגובה להלם בשל הליניאריות.

באופן דומה, עבור מערכות ריבועיות, דהיינו, מערכות שתमונת הכניסה באותו גודל של תമונת היציאה, שרשור המערכות בזו אחר זו נוטן

$$\underline{g}_1 = \mathbf{H}_1 \underline{f} , \quad \underline{g}_2 = \mathbf{H}_2 \underline{g}_1 = \mathbf{H}_2 \mathbf{H}_1 \underline{f}$$

והפעם קיבלנו כי יש להכפיל את המטריצות זו בזו כדי לקבל מערכת אפקטיבית אחת עם אותה פעולה. תcona ידועה באלגברה ליניארית היא שמטריצות ריבועיות בד"כ לא מתחלפות במכפלה, כלומר $\mathbf{H}_1 \mathbf{H}_2 \neq \mathbf{H}_2 \mathbf{H}_1$. אמירה זו מתרגם לכך שלא ניתן במקרה הכללי להחליף סדר הפעלתן של מערכות ליניאריות – החלפה כזו צפוייה לתת תוצאה שונה. עם זאת, הזכרנו קודם כי מערכות ליניאריות קבועות במקומות ניתנות להחלפה ללא פגעה בתוצאה. זה מעורר את השאלה מה מיוחד במטריצות המייצגות מערכות ליניאריות קבועות במקומות? כזכור, מערכת קבועה במקומות מאופיינת בקשר- $h[m,n,k,j]=h[m-k,n-j]$. נחזר לדוגמה הקודמת ונראה כיצד מטריצה זו לתמונה קטנה.

דוגמה 2.6: בהמשך לדוגמה 2.5, נניח כיuta שהמערכת קבועה במקומות ולבן מתקיים הקשר שהזוכר לעילו: $h[m,n,k,j]=h[m-k,n-j]$. הצבת תcona זו במטריצה שהתקבלה (לקלט ופלט מעל אותו תמצ) תיתן את המטריצה הבאה

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[0,1] \\ g[1,1] \end{bmatrix} = \begin{bmatrix} h[0,0,0,0] & h[0,0,1,0] & h[0,0,0,1] & h[0,0,1,1] \\ h[1,0,0,0] & h[1,0,1,0] & h[1,0,0,1] & h[1,0,1,1] \\ h[0,1,0,0] & h[0,1,1,0] & h[0,1,0,1] & h[0,1,1,1] \\ h[1,1,0,0] & h[1,1,1,0] & h[1,1,0,1] & h[1,1,1,1] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix} = \\ = \begin{bmatrix} h[0,0] & h[-1,0] & h[0,-1] & h[-1,-1] \\ h[1,0] & h[0,0] & h[1,-1] & h[0,-1] \\ h[0,1] & h[-1,1] & h[0,0] & h[-1,0] \\ h[1,1] & h[0,1] & h[1,0] & h[0,0] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

עבור החלק השני של דוגמה 2.5 נקבל:

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[2,0] \\ g[0,1] \\ g[1,1] \\ g[2,1] \end{bmatrix} = \begin{bmatrix} h[0,0] & h[-1,0] & h[0,-1] & h[-1,-1] \\ h[1,0] & h[0,0] & h[1,-1] & h[0,-1] \\ h[2,0] & h[1,0] & h[2,-1] & h[1,-1] \\ h[0,1] & h[-1,1] & h[0,0] & h[-1,0] \\ h[1,1] & h[0,1] & h[1,0] & h[0,0] \\ h[2,1] & h[1,1] & h[2,0] & h[1,0] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

היחود במטריצות הניל ניכר אם מסתכים על ארבעת חלקי בגודל 2 על 2 במקורה הראשון, ובлокים בגודל 3 (שורות) על 2 במקורה השני. כל תת-בלוק כזה הוא בעל מבנה טופליץ, בו ערכי האיברים על האלכסונים הראשיים זהים, כמפורט בציור 2.3. המטריצה אותה קיבלנו בדוגמה היא בעל מבנה בלוק-טופליץ. מבנה זה מתקבל אם כל בלוק במטריצה מהוועה מטריצת טופליץ, ואם בנוסף, הבלוקים עצם ערוכים בצורה טופליץ (שלאורך אלכסונים ראשיים הבלוקים זהים).

במקורה הכללי, אנו צפויים לקבל כי מערכת קבועה במקום הפעלת על תמונה בגודל $M_1 \times N_1$ ומוגבהת תמונה בגודל $M_2 \times N_2$ תתואר במטריצה בגודל $M_2 N_2$ שורות על $M_1 N_1$ עמודות. מטריצה זו תחולק ל- $N_2 \times N_1$ בלוקים שככל אחד בגודל $M_2 \times M_1$ ובעל מבנה טופליץ, והסדר ביניהם הוא כשל מבנה טופליץ.

$$\begin{bmatrix} a & b & c & d & e \\ f & a & b & c & d \\ g & f & a & b & c \\ h & g & f & a & b \\ i & h & g & f & a \end{bmatrix}$$

ציור 2.3 – מבנה כללי של מטריצת טופליץ בגודל 5 על 5. רק 9 ערכים מאפיינים את כל 25 כניסה של המטריצה. אם כל איבר במטריצה הניל הוא עצמו מטריצת טופליץ, הרי שהמבנה המתkeletal הוא מבנה של מטריצת בלוק-טופליץ.

לסיכום ביניים: המטריצה המתארת מערכת קבועה במקום הינה בעל מבנה בלוק טופליץ. האם מבנה זה מספק את תכונת החלפה – התשובה היא לצערנו עדין לא. נראה זאת דרך הדוגמה הבאה.

דוגמה 2.7: ניקח שתי מטריצות טופליץ ונכפול אותן זו בזו בשני הכיוונים.

$$\begin{pmatrix} 3.0 & 1.0 & 0.8 & 0.0 \\ 1.0 & 3.0 & 1.0 & 0.8 \\ 0.8 & 1.0 & 3.0 & 1.0 \\ 0.0 & 0.8 & 1.0 & 3.0 \end{pmatrix} \begin{pmatrix} 2.0 & 1.0 & 0.5 & 0.0 \\ 1.0 & 2.0 & 1.0 & 0.5 \\ 0.5 & 1.0 & 2.0 & 1.0 \\ 0.0 & 0.5 & 1.0 & 2.0 \end{pmatrix} = \begin{pmatrix} 7.4 & 5.8 & 4.1 & 1.3 \\ 5.5 & 8.4 & 6.3 & 4.1 \\ 4.1 & 6.3 & 8.4 & 5.5 \\ 1.3 & 4.1 & 5.8 & 7.4 \end{pmatrix}$$

$$\begin{pmatrix} 2.0 & 1.0 & 0.5 & 0.0 \\ 1.0 & 2.0 & 1.0 & 0.5 \\ 0.5 & 1.0 & 2.0 & 1.0 \\ 0.0 & 0.5 & 1.0 & 2.0 \end{pmatrix} \begin{pmatrix} 3.0 & 1.0 & 0.8 & 0.0 \\ 1.0 & 3.0 & 1.0 & 0.8 \\ 0.8 & 1.0 & 3.0 & 1.0 \\ 0.0 & 0.8 & 1.0 & 3.0 \end{pmatrix} = \begin{pmatrix} 7.4 & 5.5 & 4.1 & 1.3 \\ 5.8 & 8.4 & 6.3 & 4.1 \\ 4.1 & 6.3 & 8.4 & 5.8 \\ 1.3 & 4.1 & 5.5 & 7.4 \end{pmatrix}$$

כפי שאנו רואים, המכפלה קרובה מאוד אך לא זהה. בהמשך נבין מהיון נובע שינוי זה, ומדובר לא התקבלה יcollת החלפה כמקובל באותות בעלי תמק אינטובי.

נמשיך בקו אותו בנינו עד כה, ונדון בעת במקרה בו המערכת הליניארית פרידה. במקרה זה מתקיים $h[m,n] = h_1[m]h_2[n]$ וזה.

דוגמה 2.8: בהמשך לדוגמה 2.6 (החלק הראשון), נניח בעת שהמערכת גם פרידה. המטריצה תהיה

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[0,1] \\ g[1,1] \end{bmatrix} = \begin{bmatrix} h_1[0]h_2[0] & h_1[-1]h_2[0] & h_1[0]h_2[-1] & h_1[-1]h_2[-1] \\ h_1[1]h_2[0] & h_1[0]h_2[0] & h_1[1]h_2[-1] & h_1[0]h_2[-1] \\ h_1[0]h_2[1] & h_1[-1]h_2[1] & h_1[0]h_2[0] & h_1[-1]h_2[0] \\ h_1[1]h_2[1] & h_1[0]h_2[1] & h_1[1]h_2[0] & h_1[0]h_2[0] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

מה מיוחד במבנה הנ"ל? קשה לראות זאת אך ישנה כאן מכפלת קronecker (Kronecker) בין שתי מטריצות בגודל 2 על 2 בצורה

$$\begin{aligned} & \begin{bmatrix} h_1[0]h_2[0] & h_1[-1]h_2[0] & h_1[0]h_2[-1] & h_1[-1]h_2[-1] \\ h_1[1]h_2[0] & h_1[0]h_2[0] & h_1[1]h_2[-1] & h_1[0]h_2[-1] \\ h_1[0]h_2[1] & h_1[-1]h_2[1] & h_1[0]h_2[0] & h_1[-1]h_2[0] \\ h_1[1]h_2[1] & h_1[0]h_2[1] & h_1[1]h_2[0] & h_1[0]h_2[0] \end{bmatrix} = \\ & = \begin{bmatrix} h_2[0] & h_2[-1] \\ h_2[1] & h_2[0] \end{bmatrix} \otimes \begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \end{bmatrix} = \\ & = \begin{bmatrix} h_2[0] \cdot \begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \end{bmatrix} & h_2[-1] \cdot \begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \end{bmatrix} \\ h_2[1] \cdot \begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \end{bmatrix} & h_2[0] \cdot \begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \end{bmatrix} \end{bmatrix} \end{aligned}$$

באופן דומה, עבור החלק השני של הדוגמה יתקבל

$$\begin{bmatrix} h_1[0]h_2[0] & h_1[-1]h_2[0] & h_1[0]h_2[-1] & h_1[-1]h_2[-1] \\ h_1[1]h_2[0] & h_1[0]h_2[0] & h_1[1]h_2[-1] & h_1[0]h_2[-1] \\ h_1[2]h_2[0] & h_1[1]h_2[0] & h_1[2]h_2[-1] & h_1[1]h_2[-1] \\ h_1[0]h_2[1] & h_1[-1]h_2[1] & h_1[0]h_2[0] & h_1[-1]h_2[0] \\ h_1[1]h_2[1] & h_1[0]h_2[1] & h_1[1]h_2[0] & h_1[0]h_2[0] \\ h_1[2]h_2[1] & h_1[1]h_2[1] & h_1[2]h_2[0] & h_1[1]h_2[0] \end{bmatrix} =$$

$$= \begin{bmatrix} h_2[0] & h_2[-1] \\ h_2[1] & h_2[0] \end{bmatrix} \otimes \begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \\ h_1[2] & h_1[1] \end{bmatrix}$$

(יש לשים לב ולהיזהר מכך שבחרנו באותו סימון לكونבולוציה ולמכפלת קרוןקר).

במקרה הכללי, אנו צפויים לקבל כי מערכת ליניארית קבועה במקום ופרידה הפעלת על התמונה בגודל $M_1 N_1$ ומיניבה תמונה בגודל $M_2 N_2$ תתואר כמטריצה בגודל $M_2 N_2$ שורות על $M_1 N_1$ עםודות. מטריצה זו תיבנה מכפלת קרוןקר של מטריצות בגודל $N_1 \times N_2$ ו- $M_2 \times M_1$ כשל אחת מהן בעלת מבנה טופלי. ברור כי למטריצה הכלולית יהיה מבנה בлок טופלי, בשל דרך בנייתה.

הפיוק הנ"ל אינו מקרי. אם נחזר ליסודות, הרי שكونבולוציה עם תגובת הלם פרידה פירושה

הקשר הבא

$$g[m, n] = H\{f[m, n]\} = \sum_k \sum_j f[k, j] h[m - k, n - j] = \sum_k h_1[m - k] \sum_j f[k, j] h_2[n - j]$$

וכבר נאמר בפרק זה כי קשר זה ממשמו פעולה נפרדת על עםודות ושורות. כיצד תכונה זו באה לידי ביטוי ביצוג מטריצי? ובכן,שוב דרך דוגמה, נמחיש כיצד הדבר פועל.

דוגמה 2.9: ניקח את תמונה הכניטה מדוגמה קודמת, $f[m, n]$, והפעם נחזק אותה כמטריצה בגודל 2 על 2, לפי

$$f = \begin{bmatrix} f[0,0] & f[0,1] \\ f[1,0] & f[1,1] \end{bmatrix}$$

נניח כי רצוננו להפעיל את המערכת היליניארית הקבועה במקום ופרידה כפי שתואר בדוגמה הקודמת, לשם יצירת תמונה מוצא בגודל 3 שורות על 2 עםודות. הקשר הבא מתאר כיצד تعمل מערכת זו, וניתנו ע"י פתיחת המכפלות לוודא את נכונותה של המשוואה

$$\begin{bmatrix} h_1[0] & h_1[-1] \\ h_1[1] & h_1[0] \\ h_1[2] & h_1[1] \end{bmatrix} \begin{bmatrix} f[0,0] & f[0,1] \\ f[1,0] & f[1,1] \end{bmatrix} \begin{bmatrix} h_2[0] & h_2[-1] \\ h_2[1] & h_2[0] \end{bmatrix} = \begin{bmatrix} g[0,0] & g[0,1] \\ g[1,0] & g[1,1] \\ g[2,0] & g[2,1] \end{bmatrix}$$

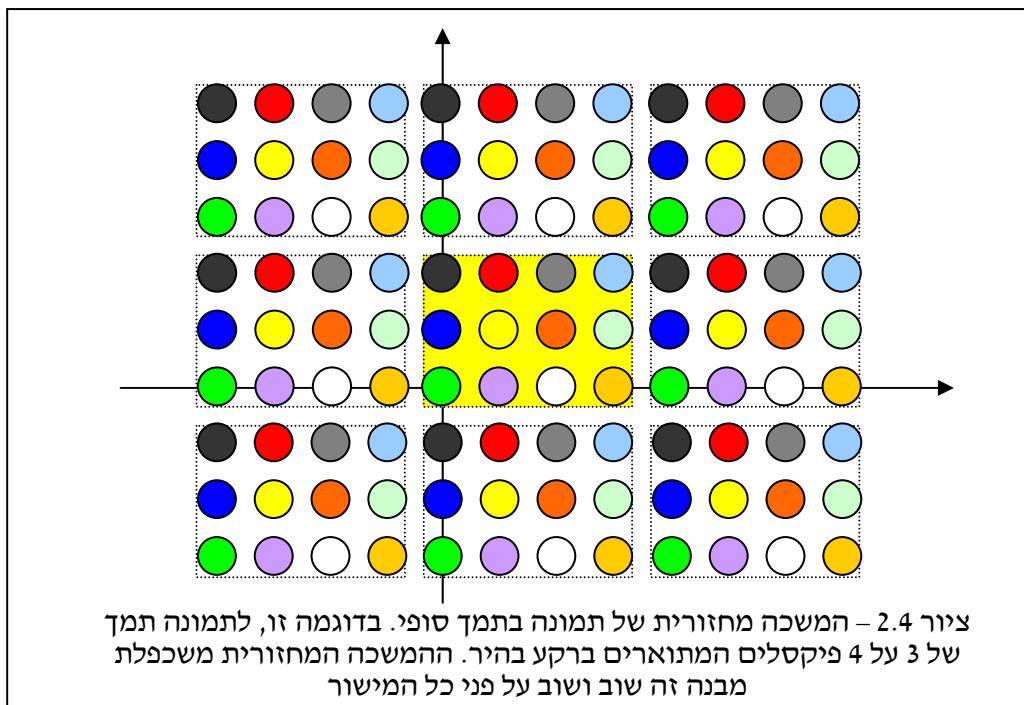
המטריצה כופלת את התמונה מצד ימין (בגודל 2 על 2) מבצעת פעולה סיכון שורות לתמונה – ניתן לראות זאת ע"י פתיחת המכפלת, ואבחנה כי נוצרו צירופים ליניארים של פיקסלים באותה שורה. באופן דומה, המכפלת מצד שמאל מבצעת סיכון עםודות.

פרק 2 – עיבוד אותות בדו-מימד

נושא אחרון בו נדון במסגרת סעיף זה היא פועלות הקונבולוציה הциיקלית. עד כה, כשלSPA^תתgebת ההלם מהתמק, הנחנו כי מוחוצה לו ישנים אפסים. עצת מציע גישה שונה הקרויה המשכחה מחזוריות של התמונה. בהינתן תמונה המוגדרת על התמק $\Omega = \{[m, n] | 0 \leq m \leq M-1, 0 \leq n \leq N-1\}$, נגדיר את ערכיה בכל מקום במישור באופן הבא:

$$f[m, n] = \begin{cases} f[m, n] & 0 \leq m \leq M-1, 0 \leq n \leq N-1 \\ f[m \bmod M, n \bmod N] & \text{otherwise} \end{cases}$$

ציור 2.4 מתאר תהליך זה לתמונה בגודל 3 על 4.



פועלות הקונבולוציה בין תמונה צו (לאחר המשכחה מחזוריית) ובין תגבות הלם (ואנו מניחים כאן מערכת ליניארית קבועה במקום), תתקבל ע"י הביטוי הבא

$$g_c[m, n] = \sum_{k=0}^{M-1} \sum_{j=0}^{N-1} f[k, j] h[(m-k) \bmod M, (n-j) \bmod N]$$

הקרויה קונבולוציה ציקלית. הפעולה \bmod נוטנת שכפול לפי הקשר $m \bmod M = m - \lfloor m/M \rfloor M$, ופועלות $[-]*$ היא פועלות הערך השלים הגדול ביותר מלמטה לערך. כדוגמה, עבור m בתחום 0 עד $M-1$, הערך לא משתנה כלל. באופן דומה, $M \bmod M = 0$, $M-1 \bmod M = 1$, $M-2 \bmod M = -1$, $M-3 \bmod M = -2$, וכו'. כך יוצאה שכאשר התגובה להלם פולשת מצד אחד של התמק, היא בפועל כופלת איברים מצד אחד אחר. נזכיר על דוגמאות קודומות, כשהפעים אנו מבצעים קונבולוציה ציקלית

דוגמה 2.10: אנו מתייחסים לתמונות כניסה $f[m,n]$ מעל תמצ $\Omega = \{(m,n) | 0 \leq m \leq 1, 0 \leq n \leq 1\}$ ותउין בערכיו המוצא $g[m,n]$ מעל השרג $\Omega_g = \{(m,n) | 0 \leq m \leq 2, 0 \leq n \leq 1\}$. נניח כי אנו עוסקים בסידור לקסיקוגרפי לפי עמודות. נמchio תחילת כיצד תיראה השורה הראשונה. שורה זו בונה את האיבר $h[0,0]$ של $[0,0] g$. לפי הקשר הניל מתקובל

$$\begin{aligned} g[0,0] &= \sum_k \sum_j f[k,j] h[-k \bmod 2, -j \bmod 2] = \\ &= f[0,0]h[0,0] + f[1,0]h[1,0] + f[0,1]h[0,1] + f[1,1]h[1,1] = \\ &= \begin{bmatrix} h[0,0] & h[1,0] & h[0,1] & h[1,1] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix} \end{aligned}$$

והמטריצה H בכללה נותנת את הקשר הבא

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[2,0] \\ g[0,1] \\ g[1,1] \\ g[2,1] \end{bmatrix} = \begin{bmatrix} h[0,0] & h[1,0] & h[0,1] & h[1,1] \\ h[1,0] & h[0,0] & h[1,1] & h[0,1] \\ h[0,0] & h[1,0] & h[0,1] & h[1,1] \\ h[0,1] & h[1,1] & h[0,0] & h[1,0] \\ h[1,1] & h[0,1] & h[1,0] & h[0,0] \\ h[0,1] & h[1,1] & h[0,0] & h[1,0] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

כפי שראויים, גם הפעם למטריצה מבנה בлок טופלי, כאמור. למעשה, המבנה אפילו יותר ייחודי מאשר בлок טופלי כללי – המבנה כתע בлок סיבובי (Block-Circulant). במבנה זה, שורה ראשונה של המטריצה מכתיבה אתcola, ע"י סיבוב משמאלי לימיין. ערך שנפלט ימינה נכנס משמאלו. ציור 2.5 ממחיש דרך בנייתה של מטריצה כזו. גם הפעם, מטריצה בעלת מבנה בлок-סיבובי מתקובלת ע"י בлокים שככל אחד מהם הוא בлок-סיבובי, וסדר הבlokים הוא בлок-סיבובי.

נמשיך את הדיון בדוגמה, כאשר מניחים כי במערכת פרידה. במקרה זה מתקובל

$$\begin{bmatrix} g[0,0] \\ g[1,0] \\ g[2,0] \\ g[0,1] \\ g[1,1] \\ g[2,1] \end{bmatrix} = \begin{bmatrix} h_1[0]h_2[0] & h_1[1]h_2[0] & h_1[0]h_2[1] & h_1[1]h_2[1] \\ h_1[1]h_2[0] & h_1[0]h_2[0] & h_1[1]h_2[1] & h_1[0]h_2[1] \\ h_1[0]h_2[0] & h_1[1]h_2[0] & h_1[0]h_2[1] & h_1[1]h_2[1] \\ h_1[0]h_2[1] & h_1[1]h_2[1] & h_1[0]h_2[0] & h_1[1]h_2[0] \\ h_1[1]h_2[1] & h_1[0]h_2[1] & h_1[1]h_2[0] & h_1[0]h_2[0] \\ h_1[0]h_2[1] & h_1[1]h_2[1] & h_1[0]h_2[0] & h_1[1]h_2[0] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix} =$$

$$= \begin{bmatrix} h_2[0] & h_2[1] \\ h_2[1] & h_2[0] \\ h_2[0] & h_2[1] \end{bmatrix} \otimes \begin{bmatrix} h_1[0] & h_1[1] \\ h_1[1] & h_1[0] \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ f[0,1] \\ f[1,1] \end{bmatrix}$$

$\begin{bmatrix} a & b & c & d & e \\ f & a & b & c & d \\ g & f & a & b & c \\ h & g & f & a & b \\ i & h & g & f & a \end{bmatrix}$	$\begin{bmatrix} a & b & c & d & e \\ e & a & b & c & d \\ d & e & a & b & c \\ c & d & e & a & b \\ b & c & d & e & a \end{bmatrix}$
---	---

ציור 2.5 – מבנה כללי של מטריצה סיבובית בגודל 5 על 5. אם בטופלי (צד שמאל) איפיינו 9 ערכים את המטריצה, הרי שהפעם (צד ימין) 5 בלבד מאפיינים אותה. אם כל איבר במטריצה הניל הוא עצמו מטריצה בלוק-סיבובית, הרי שהמבנה המתקיים הוא מבנה של מטריצות בלוק-סיבוביות.

והפעם יצא כי המטריצות הקטנות המתארות פעולות חד-מימדיות הינן סיבוביות עצמן. מטריצות סיבוביות או מטריצות בלוק-סיבוביות מקיימות את תוכנת החלפה. הדוגמה הבאה ממחישה תוכנה זו.

דוגמה 2.11: ניקח שתי מטריצות סיבוביות ונכפול אותן זו בזו בשני הכיוונים.

$$\begin{pmatrix} 3.0 & 1.0 & 0.8 & 0.0 \\ 0.0 & 3.0 & 1.0 & 0.8 \\ 0.8 & 0.0 & 3.0 & 1.0 \\ 1.0 & 0.8 & 0.0 & 3.0 \end{pmatrix} \begin{pmatrix} 2.0 & 1.0 & 0.5 & 0.0 \\ 0.0 & 2.0 & 1.0 & 0.5 \\ 0.5 & 0.0 & 2.0 & 1.0 \\ 1.0 & 0.5 & 0.0 & 2.0 \end{pmatrix} = \begin{pmatrix} 6.4 & 5.0 & 4.1 & 1.3 \\ 1.3 & 6.4 & 5.0 & 4.1 \\ 4.1 & 6.3 & 6.4 & 5.0 \\ 5.0 & 4.1 & 1.3 & 6.4 \end{pmatrix}$$

$$\begin{pmatrix} 2.0 & 1.0 & 0.5 & 0.0 \\ 0.0 & 2.0 & 1.0 & 0.5 \\ 0.5 & 0.0 & 2.0 & 1.0 \\ 1.0 & 0.5 & 0.0 & 2.0 \end{pmatrix} \begin{pmatrix} 3.0 & 1.0 & 0.8 & 0.0 \\ 0.0 & 3.0 & 1.0 & 0.8 \\ 0.8 & 0.0 & 3.0 & 1.0 \\ 1.0 & 0.8 & 0.0 & 3.0 \end{pmatrix} = \begin{pmatrix} 6.4 & 5.0 & 4.1 & 1.3 \\ 1.3 & 6.4 & 5.0 & 4.1 \\ 4.1 & 6.3 & 6.4 & 5.0 \\ 5.0 & 4.1 & 1.3 & 6.4 \end{pmatrix}$$

כפי שננו רואים, התוצאה זהה. יתרה מכך, מתקיים כי מטריצת התוצאה אף היא סיבובית. עובדה זו מתלכדת עם כך שרשרור מערכות המתארות קונבולוציה ציקלית הינו ייחודי קונבולוציה ציקלית חדשה.

2.6 התמרת פוריה דיסקרטית (DFT) לאותות בתמך סופי

כהמשך ישר לתיאור התמרת הפוריה הרציפה, וניוונה לטיפול באוטות דיסקרטיים, נעבור כעת לדzon באוטות דיסקרטיים בעלי tamk סופי. נתונה לנו תמונה $f[m,n]$ מעלה התמך $\Omega = \{[m,n] | 0 \leq m \leq M-1, 0 \leq n \leq N-1\}$. שימוש ישר בהגדרת התמרת פוריה לאוטות דיסקרטיים, וניצול העבודה שהtamk סופי, מניבה

$$F(\theta_1, \theta_2) = F_D\{f[m,n]\} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f[m,n] \exp\{-j(m\theta_1 + n\theta_2)\}$$

וההתמורה ההפוכה נתונה ע"י

$$f[m,n] = F_D^{-1}\{F(\theta_1, \theta_2)\} = \frac{1}{4\pi^2} \int_{-\pi}^{\pi} \int_{-\pi}^{\pi} F(\theta_1, \theta_2) \exp\{j(m\theta_1 + n\theta_2)\} d\theta_1 d\theta_2$$

ובכל הנאמר לעיל אין כל חדש, זהה חזרה על טיפול באוטות דיסקרטיים ללא tamk סופי. אבל, נוצר כאן תהיליך מעניין שניינו לניצול: בתחילת, לאוטות רציפים נוצרה התמורה רציפה. ברור כי אוט דיסקרטי מכיל פחות מידע מאשר רציף, ולכן אך טבעי שהוא שנקבל פישוט מה להתמרת הפוריה שלו. ואמנם, ההתמרת פוריה של אוט דיסקרטי נמצאה להיות מחזוריית, כך שעליינו להתעניין בה רק ברכיבו $[\pi, -\pi] \times [\pi, +\pi]$. בעת, עברנו לדבר על באוטות דיסקרטיים בעלי tamk סופי שפירושו כי ישם מספר סופי וידעו של דוגמאות המגדירות לחלוtin את האות, וכך גם את ההתמורה שלו. עם זאת, ההתמורה נותרה ביןתיים כפונקציה רציפה מעלה הריבוע $[\pi, -\pi] \times [\pi, +\pi]$, וב証明 כי ישנו כאן בזבוז. אם נגדיר את ההתמורה במקרה זה כך שיתקבלו MN ערכים בלבד בתדר, נראה כי אז שימרנו את עושרו של האות. ואמנם, מגדירים ההתمرة פוריה הדיסקרטית (Discrete Fourier Transform – DFT) באופן הבא

$$F[k, \ell] = F_{DFT}\{f[m,n]\} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f[m,n] \exp\left\{-j2\pi\left(\frac{mk}{M} + \frac{n\ell}{N}\right)\right\}$$

ההגדרה הניל' בוחרת להתעניין בערכי $F(\theta_1, \theta_2)$ רק בנקודות $N/\theta_1 = 2\pi k/M, \theta_2 = 2\pi\ell/N$. כפי שניתן לראות, לערכים $[k, \ell]$ מחוץ לתחום התמך מתתקבלת מחזוריות, כך שלעניןינו, יוצא כי רק MN ערכים מגדירים במלואה את ההתمرة. ההתמורה ההפוכה נתונה ע"י

$$f[m,n] = F_{DFT}^{-1}\{F[k, \ell]\} = \frac{1}{NM} \sum_{k=0}^{M-1} \sum_{\ell=0}^{N-1} F[k, \ell] \exp\left\{j2\pi\left(\frac{mk}{M} + \frac{n\ell}{N}\right)\right\}$$

כשדנים ב- DFT מקובל לשימוש בסימון $W_N^k = \exp\left\{j\frac{2\pi k}{N}\right\}$. אז ההתרמה וההתרמה ההפוכה נתונות

ע"י

$$F[k, \ell] = F_{DFT}\{f[m, n]\} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f[m, n] W_M^{-km} W_N^{-n\ell}$$

$$f[m, n] = F_{DFT}^{-1}\{F[k, \ell]\} = \frac{1}{NM} \sum_{k=0}^{M-1} \sum_{\ell=0}^{N-1} F[k, \ell] W_M^{km} W_N^{n\ell}$$

מדובר MN הערכים המוצעים מספייקים להגדרת ההתרמה? מסתבר כי ההתרמה המוצעת הינה אורתונורמלית (עד כדי קבוע), דהיינו, ברישום מטריצי מהצורה

$$\begin{bmatrix} F[0,0] \\ F[1,0] \\ \vdots \\ F[M-1, N-1] \end{bmatrix} = \begin{bmatrix} W_{MN}^0 & W_{MN}^0 & \cdots & W_{MN}^0 \\ W_{MN}^0 & W_{MN}^1 & \cdots & W_{MN}^{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ W_{MN}^0 & W_{MN}^{M-1} & \cdots & W_{MN}^{(M-1)(N-1)} \end{bmatrix} \begin{bmatrix} f[0,0] \\ f[1,0] \\ \vdots \\ f[M-1, N-1] \end{bmatrix}$$

$$\Rightarrow \underline{F} = \underline{W} \underline{f}$$

מטריצת ההתרמה הינה מטריצה יוניטרית (שוב – חסר הקבוע). לכן, היפוכה קיים ופשוט (נתון ע"י שחloff וצמוד). לכן, סידרת MN ערכי ההתרמה מתארים את N דגימות הכניסה באופן מלא. מטריצה זו בגודל MN על MN מלכשנות מטריצות בлок סיבוביות, כאמור, אם \mathbf{H} מטריצה בлок סיבובית, נקבל כי $* \mathbf{W} \mathbf{H} \mathbf{W}^*$ היא מטריצה אלכסונית. מדובר זה מעניין? ובכן, נניח כי \mathbf{H} מייצגת מטריצה תוגבה להלם של מערכת ליניארית וקבועה במקום הפעלת כקונבולוציה ציקלית. אז, מוצא המערכת לתמונות כניסה \underline{f} נתונה ע"י $\underline{g} = \underline{H} \underline{f}$. כיוון שהמטריצה \mathbf{H} בлок-סיבובית, נקבל כי מתקאים הקשר הבא

$$\underline{g} = \underline{H} \underline{f} \Rightarrow [\mathbf{W} \mathbf{H} \mathbf{W}^*] \underline{W} \underline{f} = \underline{W} \underline{g} \Rightarrow \Delta_H \underline{F} = \underline{G}$$

כאשר המטריצה Δ_H מטריצה אלכסונית. לכן, כל BIN בהתרמת התדר בין כניסה ליציאה קשורים ביניהם ע"י מכפלה בסקלר שהוא תוגב התדר בנקודה זו – בדיק כמי שקיבלו במקרה הרציף.

המטריצה \mathbf{W} המבצעת התרמת פוריה ניתנת לכתיבה כמכפלת קרוונקר של שתי מטריצות קטנות יותר, כשהכל אחת מבצעת התרמת פוריה חד-מימדית. בזה אין כל חדש כי אלו יודעים כי אפילו התרמת הפוריה הרציפה הינה פרידה. הייחוד הפעם הוא בכך שפרידות זו מיתרגרמת למכפלת קרוונקר, כיוון שלראשונה אנו עוסקים בייצוג מטריצי להתרמה. בשלב זה נעצור את הדיוון. ניתן כמובן להמשיך ולתאר את תוכנות ה- DFT, ייצוגו כפעולה מטריצית, ועוד. אנו נשוב לדון בנושאים אלו בפרק הדן בהתרמות דיסקרטיות בדו-מימד – פרק 4.

פרק 3 – דגימה ו שחזור של תמונות

41	פרק 3 – דגימה ו שחזור של תמונות
41	3.1 דגימה אחידה של אותן דו-מימדיים
44	3.2 שחזור מדגימה אחידה
46	3.3 תופעת הקיפול בתמונות
51	3.4 משפטי דגימה מוכללים לדו-מימד

בפרק זה נציג בהרחבה את הנושאים דגימה ו שחזור לאותות דו-מימדיים. כזכור מפרק 1, ע"י דגימה אנו עוברים מאות רציף המחייב שימוש באינסוף סיביות לשם ייצוג מלא, לאות דיסקרטי. חשוב להבהיר כי יש לבצע על תוצאת הדגימה פעולה קוונטיזציה על מנת לקבל אותן המיצג בכמות סופית (אך גדולה בד"כ) של סיביות. נושא הקוונטיזציה ייסקר במסגרת הפרק הבא.

1.3 דגימה אחידה של אותן דו-מימדיים

נתונה לנו תמונה המיצגת אותן דו מימי (x,y). רצוננו לדגם תמונה זו במרוחקים D_x ו- D_y

ולקבל תמונה דוגמה

$$f[m,n] = f(x,y) \begin{cases} x = mD_x \\ y = nD_y \end{cases} = f(mD_x, nD_y)$$

השאלה המרכזית בה עוסקת סעיף זה היא אילו תנאים יש להציב על המרווחים D_x ו- D_y כך שנייתן יהיה לשחזר את $f(x,y)$ מתוך דגימותיו $[f[m,n]]$? לצורך מענה לשאלה זו, נעשה שימוש בתיאור מתמטי נוח יותר לפעולות הדגימה – דגימה ע"י מכפלה במסרך הלמים. נקבע כי האות נדגם ע"י המכפלה

פרק 3 – דגימה ושחזור של תמונות

$$\tilde{f}[m,n] = f(x,y) \cdot \sum_{m,n} \delta(mD_x - x, nD_y - y)$$

בפרק קודם עשינו שימוש בביטוי זה בדיקות קביעה שרירוטית של מרוחקי הדגימה להיות 1) בובאנו להסביר את המקור להגדלת התמרת הפוריה לאותות דיסקרטיים. בעת נבעצ' מניפולציה פשוטה אחרת – כיוון שהאות בנוי מכפלה של שני אותות רציפים בערך, התמרת הפוריה שלו היא קונבולוציה של ההתמורות. לכן

$$\begin{aligned} F\{\tilde{f}[m,n]\} &= F\left\{f(x,y) \cdot \sum_{m,n} \delta(mD_x - x, nD_y - y)\right\} = \\ &= F\{f(x,y)\} \otimes F\left\{\sum_{m,n} \delta(mD_x - x, nD_y - y)\right\} = \\ &= \frac{1}{D_x D_y} F(u,v) \otimes \sum_{m,n} \delta\left(u - \frac{m}{D_x}, v - \frac{n}{D_y}\right) = \\ &= \frac{1}{D_x D_y} \sum_{m,n} F\left(u - \frac{m}{D_x}, v - \frac{n}{D_y}\right) \end{aligned}$$

התמרה על סידרת ההלמים מוכרת בגרסתה החד-מיידית בשם נוסחת פואסון. מסרך ההלמים החד-מיידי יוצר את המסרק הדז-מיידי ע"י מכפלה ספרטילית:

$$\begin{aligned} \text{Comb}(x,y,D_x,D_y) &\stackrel{\Delta}{=} \sum_{m,n} \delta(mD_x - x, nD_y - y) = \text{Comb}(x,D_x) \cdot \text{Comb}(y,D_y) \\ \Rightarrow F\{\text{Comb}(x,y,D_x,D_y)\} &= \frac{1}{D_x D_y} \text{Comb}(u,D_x^{-1}) \cdot \text{Comb}(v,D_y^{-1}) \end{aligned}$$

קיבלו Ci הפונקציה $F(u,v)$ משוכפלת שוב ושוב במרוחקים קבועים בשני הצירים - $(1/D_x, 1/D_y)$. ציור 3.1 ממחיש תופעה זו. מציר זה גם ניכרת התכונה הבאה – כאשר $F(u,v)$ חסום פס בתחום המלבן המשוכפל, מתקבל Ci השכפולים לא דווקאים זה על זה. נניח Ci אמורים $F(u,v)$ מקיימים

$$F(u,v) = \begin{cases} F(u,v) & |u| \leq U_{\max}, |v| \leq V_{\max} \\ 0 & \text{otherwise} \end{cases}$$

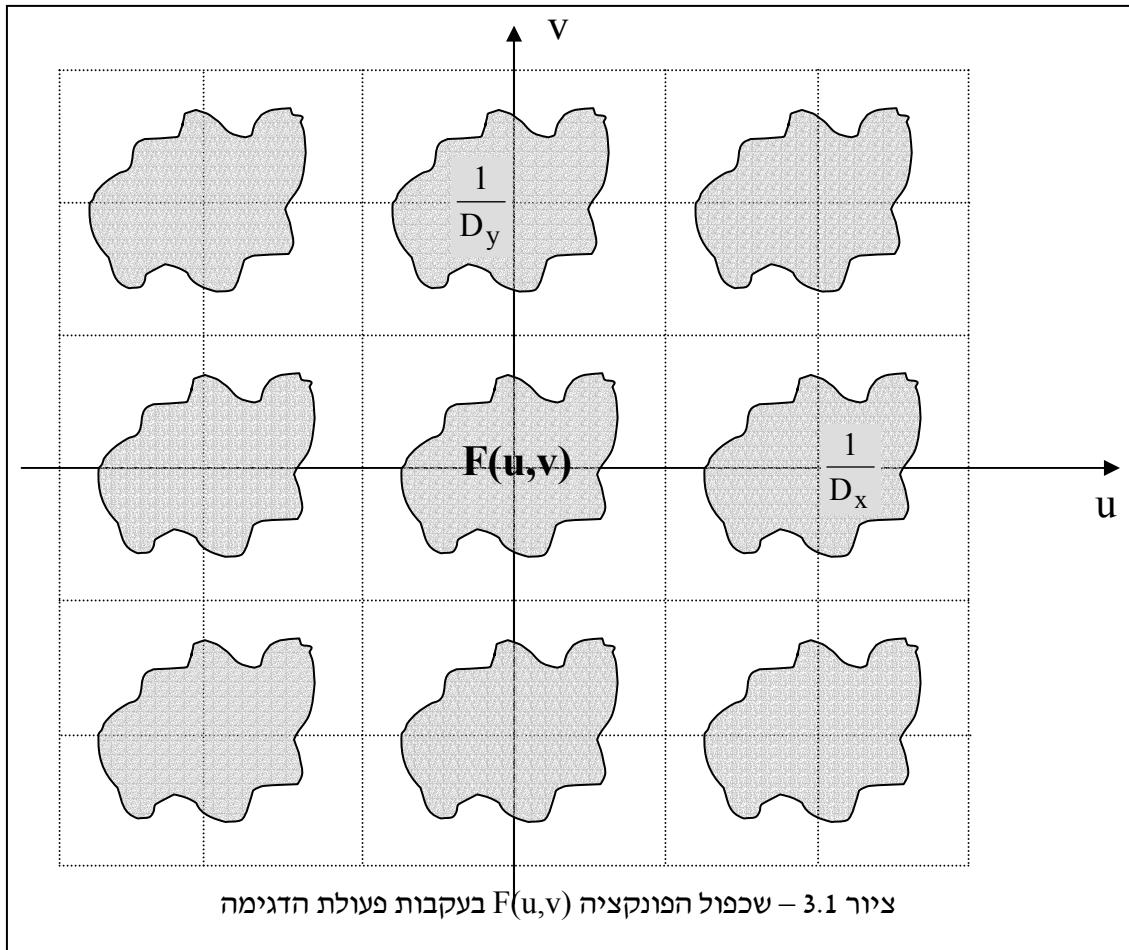
אז, על מנת שהשכפולים לא יעלו זה על זה נדרש

פרק 3 – דגימה ושחזור של תמונות

$$U_{\max} < \frac{1}{2D_x} \Rightarrow D_x < \frac{1}{2U_{\max}}$$

$$V_{\max} < \frac{1}{2D_y} \Rightarrow D_y < \frac{1}{2V_{\max}}$$

הדרישה זו אינה זרה לנו – פירושה הוא שעליינו לדגום מספיק צפוף כדי לכלול כמהות מספקת של דגימות לייצוג מלא של האות הדו-מימדי.



מדוע חשוב לנו למנוע מהשכפולים לעלות זה על זה? ובכן, בהינתן האות הדגום $\tilde{f}[m,n]$ (שכל דגימה בו היא פונקציית הלם), זהו אות רציף הנitin לסינון. ע"י הפעלת מסנן ריבועי LPF מהצורה

$$H(u,v) = \begin{cases} D_x D_y & |u| \leq \frac{1}{2D_x}, |v| \leq \frac{1}{2D_y} \\ 0 & \text{otherwise} \end{cases}$$

פרק 3 – דגימה ושחזר של תמונות

על אות זה נקבע בדיקות את אות המקור, כיון שהסינון הינו מכפלה בתדר, וע"י ההכפלה במסנן הנ"ל אנו מבטלים את כל השכפולים למעט זו הראשית, וכיון שהיא בדיקת התמרת של אות המקור, הרי שביצעו בשלמות פעולה הפוכה לפעולות הדגימה.

בחינה מעשית, לא נתנו לנו האות $\tilde{f}[m,n]$, אלא אות דגום $f[m,n]$. لكن השאלה המתבקשת היא כיצד בפועל מבצעים את הסינון המשוחרר? נענה לשאלת זו ע"י כתוב מפורש של פעולה השחזור.

$$\begin{aligned}\hat{f}(x,y) &= F^{-1} \left\{ H(u,v) F \left\{ \tilde{f}[m,n] \right\} \right\} = \tilde{f}[m,n] \otimes h(x,y) = \\ &= \int \int \sum \sum f[m,n] \delta(\alpha - mD_x, \beta - nD_y) h(x-\alpha, y-\beta) d\alpha d\beta = \\ &= \sum_m \sum_n f[m,n] h(x - mD_x, y - nD_y) = \sum_m \sum_n f[m,n] \operatorname{sinc} \left[\frac{x - mD_x}{D_x} \right] \operatorname{sinc} \left[\frac{y - nD_y}{D_y} \right]\end{aligned}$$

כלומר, השחזור נעשה ע"י צירוף ליניארי של פונקציות sinc דו-מימדיות, המוגדרות ע"י

$$\operatorname{sinc}(x) = \frac{\sin(\pi x)}{\pi x}$$

הנitorה שהוצג לעיל אינה אלא הרחבת טבעית של משפט הדגימה לדו-מימד, לפיו אותן העובר דגימה איחודית בשני הzcיריים, ובמקביל להם, מרוחקי הדגימה הנדרשים צריכים להיות קטנים מהיפוכו של פעמיים התדר המרבי בכל ציר.

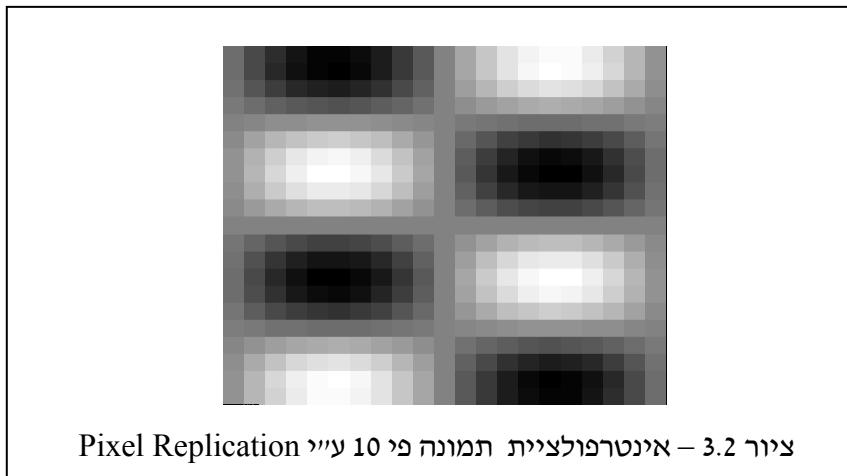
3.2 שחזור מדגימה איחודית

פעולות השחזור נועדה לקחת אות דגום ולהחזיר אל הרצף. פעולה זו נדירה למדי שכן מעט מאוד יכול להיעשות על תמונה ברצף (אפילו לא הצגתה). עם זאת, ניתן להתייחס לפעולות השחזור כניסיונו להגדיל רזולוציה, דהיינו, לקחת תמונה דוגמה וליצור דוגמאות נוספות לכך ששיריג הדגימה הסופי יהיה צפוף יותר וייתן תחושה של רצף. תהליך של ייצור תמונה גדולה יותר אך בעלת מספר גדול יותר של דוגמאות תicker אינטראקטיבית.

שחזור מדגימה בנקודה כלשהי עפ"י נוסחת השחזור שהתקבלה קודם לכן מחייב ביצוע צירוף ליניארי המערב את כל אינסוסף הדגימות הנתונות לנו. בפועל, כמשמעותם חייבים לקרב פעולה זו ע"י שימוש בצירוף ליניארי בתמק מוגבל וסופי. קיימות מספר אפשרויות מעשיות, בדומה לנלמד בעבר אחרות חד-ממדיים.

פרק 3 – דגימה ושהזהר של תמונות

השהזהר הפשטוט ביותר מציע החלפת כל דגימה בפונקציה מלכנית בגודל $D_x \times D_y$ ממורכזת סביב הראשית, בעלת גובה קבוע שווה לערך הדגימה. זהו ה-Zero Order Hold, אך עברו תמונות אין דרישת לסייעיות וכן הפונקציה סימטרית סביב הראשית. בפועל, האות הדיסקרטי (כשהוא מיוצג כמטריצת הלמים בגבהים שונים) עבר קונבולוציה עם פונקציה זו. כשההlixir זה מבוצע לאינטראולציה לשם הגדלת התמונה פי 2 בכל ציר, נקבל כי פשוט עליינו לשכפל כל שורה וכל عمودה פעמיים – תהליך הידוע Nearest Neighbor, או Pixel Replication, כיוון שהערכים החסרים ממולאים ע"י העתקת השכן הקרוב ביותר. ציור 3.2 ממחיש כיצד נראה תוצאה תהליכי זה עברו תומנת סינוס דו-מימדי בגודל 21 על 21 שעברה הגדלה פי 10 בכל ציר.

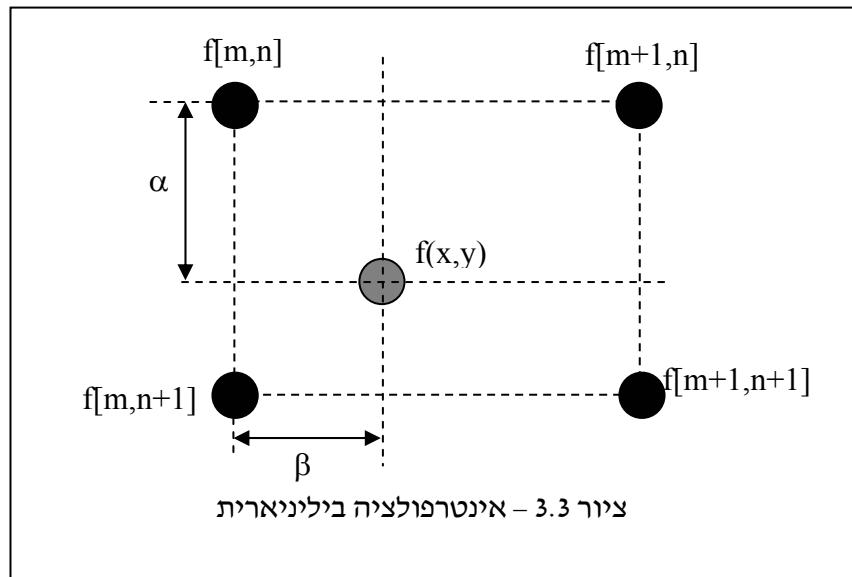


אפשרות טוביה יותר היא הפונקציה המשולשת הדו-מימדית, הבננית ממכפלת שתי פונקציות משולשות חד-מימדיות סימטריות סביב הראשית, הkoplotot זו את זו לייצרת פונקציה דו-מימדית ספרבילית. הפונקציה המשולשת עברו הציר האופקי מקבלת גובה 1 בראשית ו- 0 בנקודת סוף. שימוש בפונקציה זו לאינטראולציה קרווי Bilinear Interpolation, והשימוש בו מאוד פופולרי. עברו 4 דוגמאות סמוכות $\{f[m,n], f[m+1,n], f[m,n+1], f[m+1,n+1]\}$ ניתן להראות כי ערך המצוין בתוך הריבוע נתון ע"י

$$f(x,y) = [1-\alpha \quad \alpha] \begin{bmatrix} f[m,n] & f[m,n+1] \\ f[m+1,n] & f[m+1,n+1] \end{bmatrix} \begin{bmatrix} 1-\beta \\ \beta \end{bmatrix}$$

כאשר β, α הם ערכים בין 0 ל- 1 הנוגעים את מיקום ה- x וה- y יחסית לאורך הדגימה, כפי שמתואר בציור 3.3.

תוצאה הגדלת תמונה ע"י אינטראולציה ביליניארית נתונה בציור 3.4 על אותה דוגמה מצויר 3.2. כפי שניתן לראות השיפור הוא עצום. התוצאה עברו דוגמה זו מציינת כיון שהתמונה המקורית חלקה מאד, ואינטראולציה ביליניארית ממלאת את החסר בקווים אלכסוניים הנוגעים תחושה של רצף.



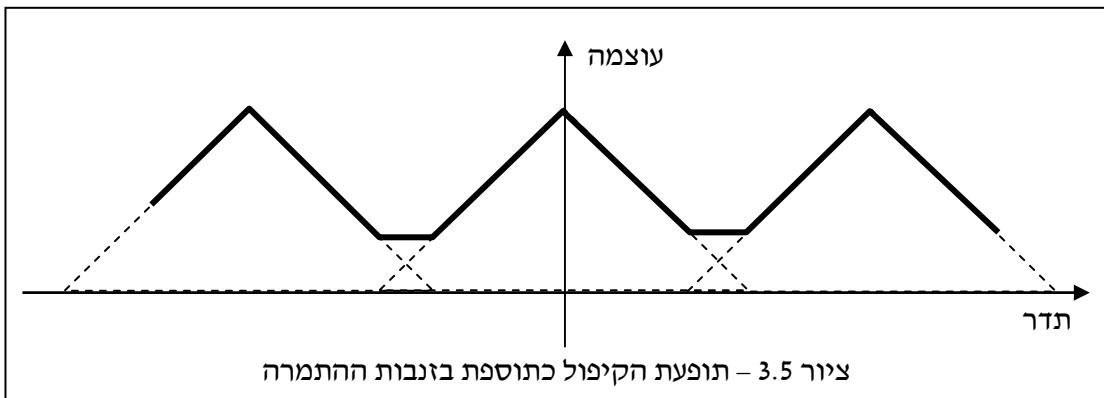
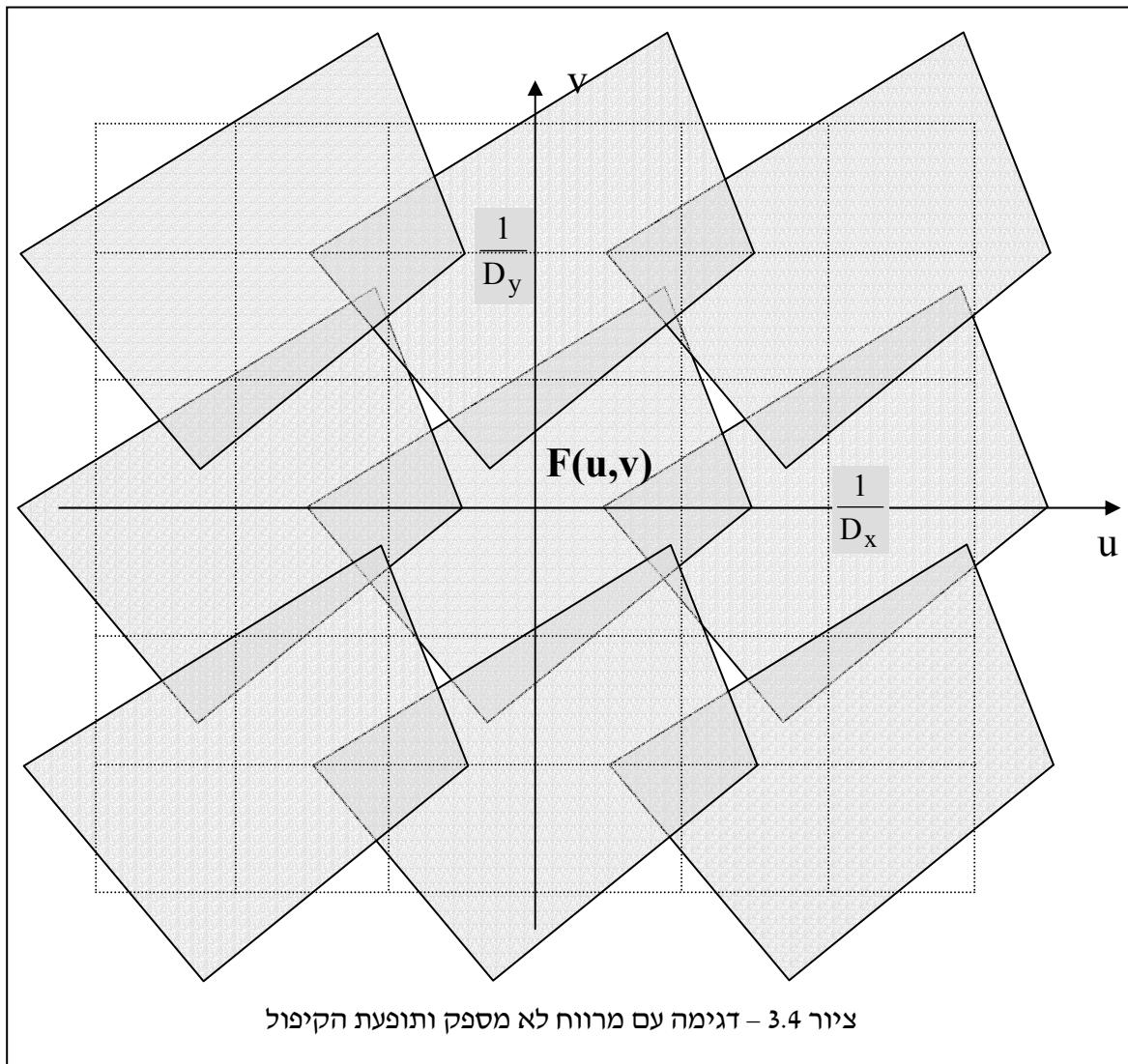
ניתן להמשיך ולתאר עוקמים מסדר גבוה יותר לאינטראפולציה, אך אנו נעצור את הדיזון בשלב זה. המשך ישיר של השיטות שהוזכרו כאן יובא בהמשך כشنשאל כיצד להגדיל תמונה תוך שימור שפota ופרטים.

3.3 תופעת הקיפול בתמונות

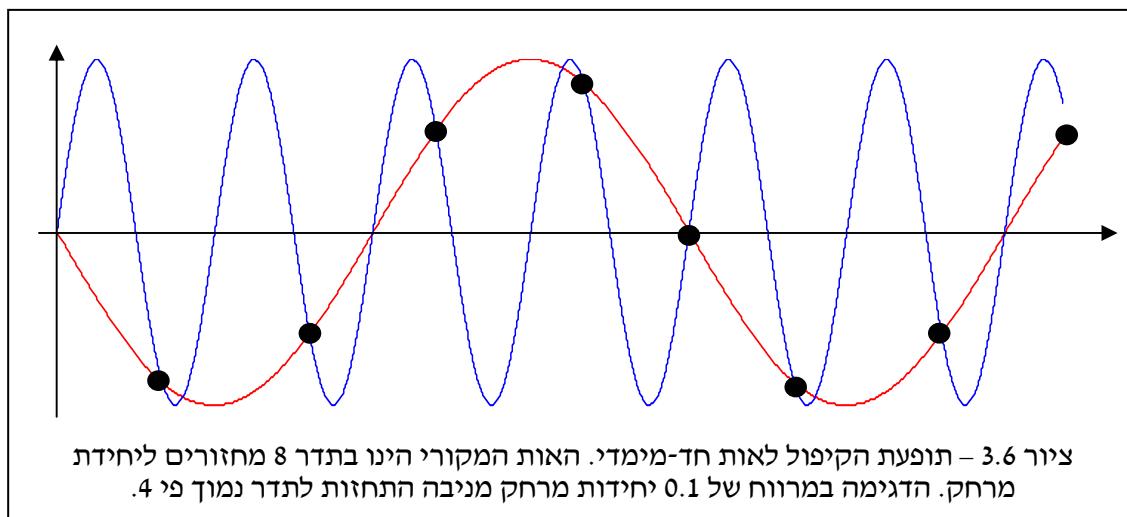
תופעת הקיפול – Aliasing – הינה פלישה של שכפולים בתדר כתוצאה מדגימה, זה לטעז זה, ובOLOR כי במקרה כזה לא ניתן לשחזר את האות המקורי ללא מידע נוסף. תופעת הקיפול מתרכשת כאשר מרוחה הדגימה גדול מדי, ואין מקיים את דרישת הסף שהוזכרה בסעיף קודם. למשל, כפי שנראה בהמשך, אותן בעל תמך סופי בمكان אין יכול להיות חסום תדר. לכן, דגימותיו בכל מרוחה שהוא חייבות להניב את תופעת הקיפול. צירור 3.4 ממחיש מקרה עם תופעת קיפול.

פרק 3 – דגימה ושחזור של תמונות

באופן טבעי נטפל בתמונות ובכאליה בעלי תמקס סופי, ותמיד נרצה לדגום אותן! האם פירוש הדבר שנסbold מטופעת הקיפול? ואם כן, כיצד תתבטא אותה תופעה בתמונות אלו אחר דגימה? כמה גדול הנזק שייגרם (אם בכלל)? ובכן, לאותות מעשיים לא ניתן להניח התנחות חסומת תדר. אולם, בהחלט ניתן להניח כי העוצמה של התמרת הפוריה דועכת ואפלו מהר כשמתרחכים מראשית היצרים (ה-DC). לכן, גם אם חלה פלישה של השכפולים אין פירוש הדבר נזק גדול בהכרח, כיון שפלישתו של זנב ההתמרה תשנה רק במקצת, כפי שמתואר לדוגמה חד-מיידי בציור 3.5.

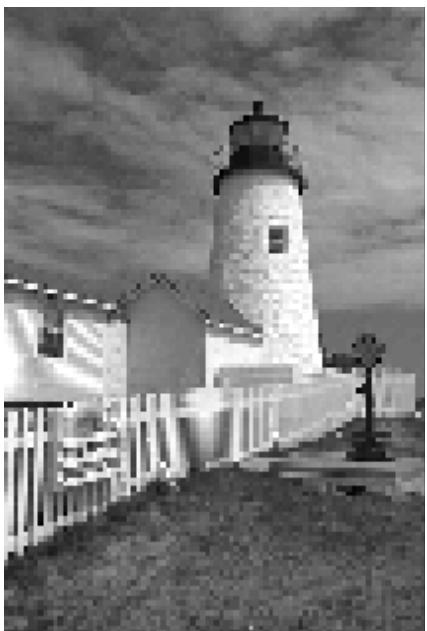


אם בכלל זאת הזוג אינו כה חלש אנרגטי, השפעתו תתבטא בתמונה הדוגמה, ותתקבלנה תופעת המוכרות שהיבטים שונים של תופעת הקיפול. תופעת הקיפול הינה במהותה תופעת התחזות בה תדר מרוחבי מעל הסף האפשרי מתחזזה לתדר נמוך המצויה בתמונה ראי לו. לשם דוגמה, אם הדגימה נעשית במרוחך של 0.1 (ביחידות אורך כלשהן), פירוש הדבר שהתדר המירבי המותר לאות הוא 5 מהזרים ליחידת אורך. נניח כי האות הנדגם כולל גם מרכיב בתדר 8 מהזרים בשניה. תדר זה יופיע בדגימה כשהוא מתחזזה ל- 2 מהזרים לשניה (תמונה הראי סביר 5 לערך 8). ציור 3.6 ממחיש תהליך זה.



בציר 3.7 נתונה תמונה דיסקרטית אשר נדגמה בczęיפות מספקת (כלומר, אין בה תופעת קיפול), ולצדיה תתי-דgesmoths שלה ע"י דילול פשוט ביחסים שלמים. לדוגמה, לקיחת כל פיקסל שני מכל שורה שנייה יוצרת תמונה קטנה פי ארבע המתיחס למרווח דגימה כפול. התמונות השונות הין בגודל שונה, אך מוצגות בגודל אחד ע"י מתייחה פשוטה הגורמת לשכפול פיקסלים. כפי שניתן לראות על פניו של הביתן בצד שמאל של התמונה, נוצר תדר שונה לכל מרוחך – אפקט קיפול קליני. אפקט נוסף ניתן לאבחנה בצד, שם ניכר כי איזורים יוצאים במרוחך מסוים מטושטשים ובמרוחכים גדולים יותר חוזרים להיות פעילים – גם אפקט של קיפול. מלבד שתי תופעות אלה ניכרת גם תופעת שינוי של שפות חדות המצוויות בזווית אלכסונית כמו במגדלור ובגיגות.

תופעת הקיפול היא תופעה מטרידה כשהיא בעוצמה חזקה. על מנת להימנע ממנה בתהליך הדגימה, נדרש להפעיל מסנן מעביר נמוכים (LPF) אשר יהפוך את התמונה הנתונה לחסומה לתדר באופן אפקטיבי, ורק לאחר מכן לבצע את הדילול. צייר 3.8 ממחיש אפקט זה. הפעם חרנוו על הניסוי של צייר 3.7, כשהפננו פועלות הדילול ביצענו סינון. במסגרת זו לא נאמר מהו המסנן שהופעל – דיוון בנושא תכנון מסנן נאות יינתן במסגרת פרק אחר. נאמר רק כי על המסנן זה להעביר טוב ככל האפשר את התדרים הנמוכים עד סף מסוים, ולהසומם טוב ככל האפשר מתחדר סף זה. כמו כן ברור כי בשל אופייה של הדגימה, המסנן הנדרש ניתן לתוכנו כמסנן ספרטילי, כאשר הסף בכל ציר עשוי להיות שונה. כפי שניתן לראות כי כל התופעות שהוזכרו נעלמו או דוכאו בצורה משמעותית.



ציור 3.7 – המחששת וופעת הקיפול: למעלה בצד ימין – תמונה מקורה. למעלה בצד שמאל – דגימה במרוחך כפול. למטה בצד שמאל – דגימה במרוחך גדול פי 3. למטה בצד ימין – דגימה במרוחך גדול פי ארבע.



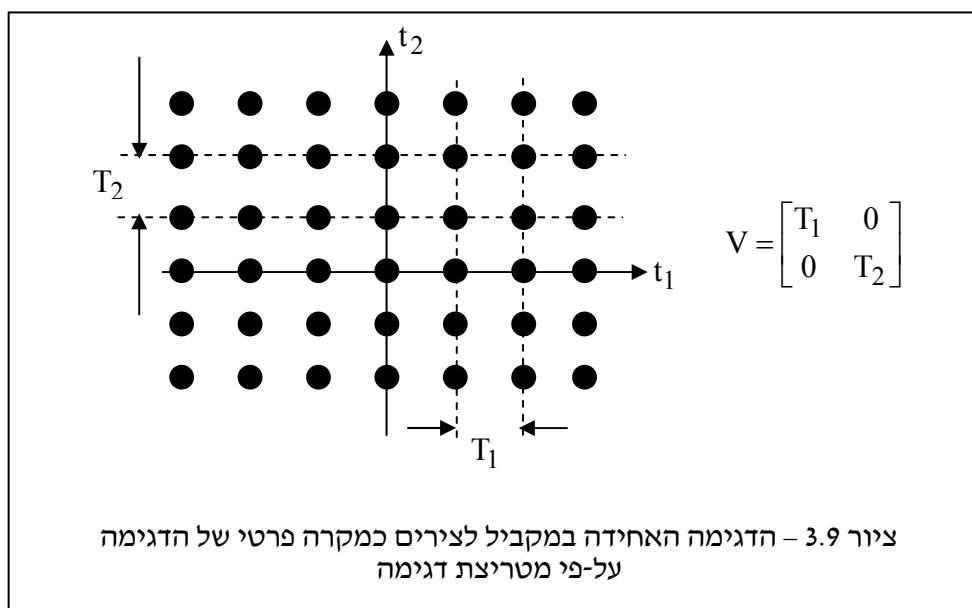
ציור 3.8 – חזרה על צייר 3.7 כשלפנוי פעולת הדילול מבוצע מסנן LPF נאות
לקיצוץ התדרים הגבוהים

3.4 משפט דגימה מוכללים לדו-מימד

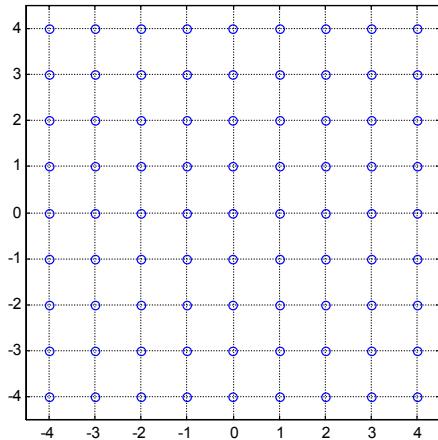
הכללה טبيعית וישראל למשפט הדגימה שתואר בסעיפים קודמים מוצעת באופן הבא: ניקח שני וקטורים בלתי תלויים ליניארית t_1 ו- t_2 , בימיד 2. שני הוקטורים בנויות מטריצה 7 כך שהוקטורים הניל הם عمודותיה. נציג דוגום את המרחב בנקודות

$$\begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{bmatrix} \begin{bmatrix} n_1 \\ n_2 \end{bmatrix} \quad n_1, n_2 \in \mathbb{N}$$

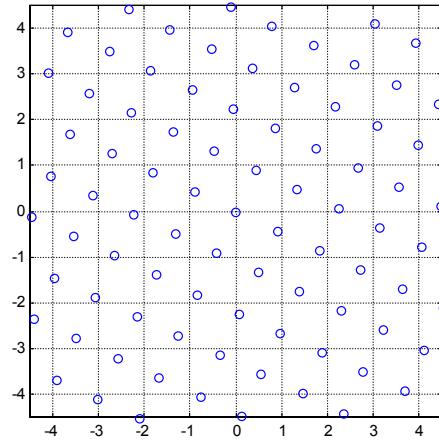
כלומר, השרגיג הרגיל של המספרים הטבעיים מומך ליניארית לנקודות $[t_1, t_2]$. שרגיג הדגימה המתkeletal מסודר במרוחקים אחידים אך עם סיבוב ביחס למישור ועובד עיוות בין הצירים. המקרה הטרייזואלי של דגימה איחודית מתואר בציור 3.9. מספר דוגמאות של המטריצה 7 והשריג הנובע ממנה מוצגים בציור 3.10. זאת כיוון ניתן להראות כי צפיפות הדגימה נתונה ע"י היפוך ערכו המוחלט של דטרמיננט המטריצה 7. זאת אומרת שוקטורים t_1 ו- t_2 יוצרים מעוין במישור אשר שטחו הוא בדיק הדטרמיננט. במקרה של דגימה איחודית ומקבילה לצירים מתkeletal כי הדטרמיננט הוא $|V|$, יש דגימה אחת לכל תא שטח בגודל $|V|$.



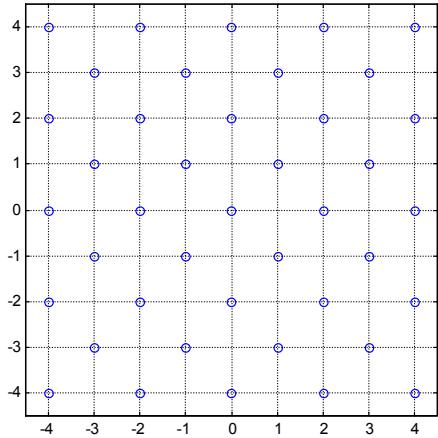
במקרה הכללי של מטריצה 7, מתkeletal שרגיג הלמים בעל מבנה מסודר המתאר את הדגימה. בדומה לתהליך שבוצע עבור הדגימה האיחודית, אם ניציג את האות הדגום כמכפלה בשרגיג מתאים של הלמים נקבל



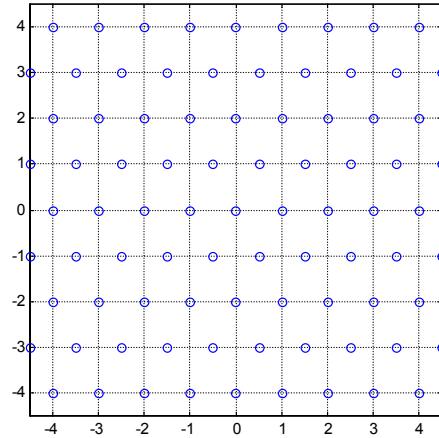
$$V = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$



$$V = \begin{bmatrix} \cos(25^\circ) & \sin(25^\circ) \\ -\sin(25^\circ) & \cos(25^\circ) \end{bmatrix}$$



$$V = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$



$$V = \begin{bmatrix} 0.5 & 0.5 \\ 1 & -1 \end{bmatrix}$$

ציור 3.10 – מטריצות דוגמאות שונות ושריג הדגימה המתתקבל מהן. יש לשים לב לכך שבשלוש דוגמאות (כולן למעט שורה תחתונה שמאליה), צפיפות הדגימה היא 1. בדוגמה החריגה, צפיפות הדגימה היא 0.5 (ביחידות של דגימות ליחידת שטח).

$$\tilde{f}[t] = f(x, y) \cdot \sum_{\underline{n}} \delta(V_{\underline{n}}) = f(x, y) \cdot \text{Comb}(x, y, V)$$

כל שעשינו הוא להכפיל את מושג המסרך. לכן, כאמור, כל שעליינו לדעת היא ההתמרה של המסרך הכללי הנ'יל. מתקובל (וأنנו לא נביא את ההוכחה)

$$\begin{aligned} F\{\tilde{f}[m, n]\} &= F\{f(x, y) \cdot \text{Comb}(x, y, V)\} = F(u, v) \otimes F\{\text{Comb}(x, y, V)\} = \\ &= \frac{1}{|V|} F(u, v) \otimes \text{Comb}(u, v, V^{-T}) = F(\Omega) \end{aligned}$$

וכך עשינו שימוש בהכללה של נוסחת פואסון. כל לראות כי המקרה הפרטני המוכר לנו של דגימה אחידה ומקבילה לצירים מסתדר היטב עם התוצאה הכללית שניתנה. במקרה הכללי מתקבל מחזוריות מהצורה

$$F(V^T \Omega) = F\left(V^T [\Omega + U \underline{k}] \right)$$

ובעבר, יש לדאוג לציפיות דגימה כזו שתבטיח כי השכפולים השונים לא עלים זה על זה, למניעת תופעת הקיפול. השחזר פועל בדרך דומה – יש לבדוק החוצה את השכפול הראשי המצו依 סיבוב הראשית. צורת המשנן במקרה הקודם הייתה מלכנית – עובדה שהביאה לפונקציה שחזר מבוססת \sin . הפעם, צורת המשנן תלולה בצורה שריג (ולא באופן פשוט).

מדוע חשוב להכיר את ההכללה שניתנה כאן? ובכן, חשוב להבהיר כי יש אינסוף מטריצות דגימה אפשריות. ע"י עיון בהתרמת הפוריה של האות הנטו ניתן להתחאים מטריצת דגימה אופטימלית. מטריצה זו תאפשר בцеיפיות הדגימה המינימלית, וכן ייצא שחסכנו בכמות הסיביות הנוחצות לאיפיונו המלא של אותן. בפועל, דגימה כללית מעין זו המתוארת כאן אינה קלה לביצוע, והדגימות היחידות המקובלות הן דגימה אחידה מקבילה לצירים או דגימה משושה. דגימות משושה מתקבלת ע"י הבחירה

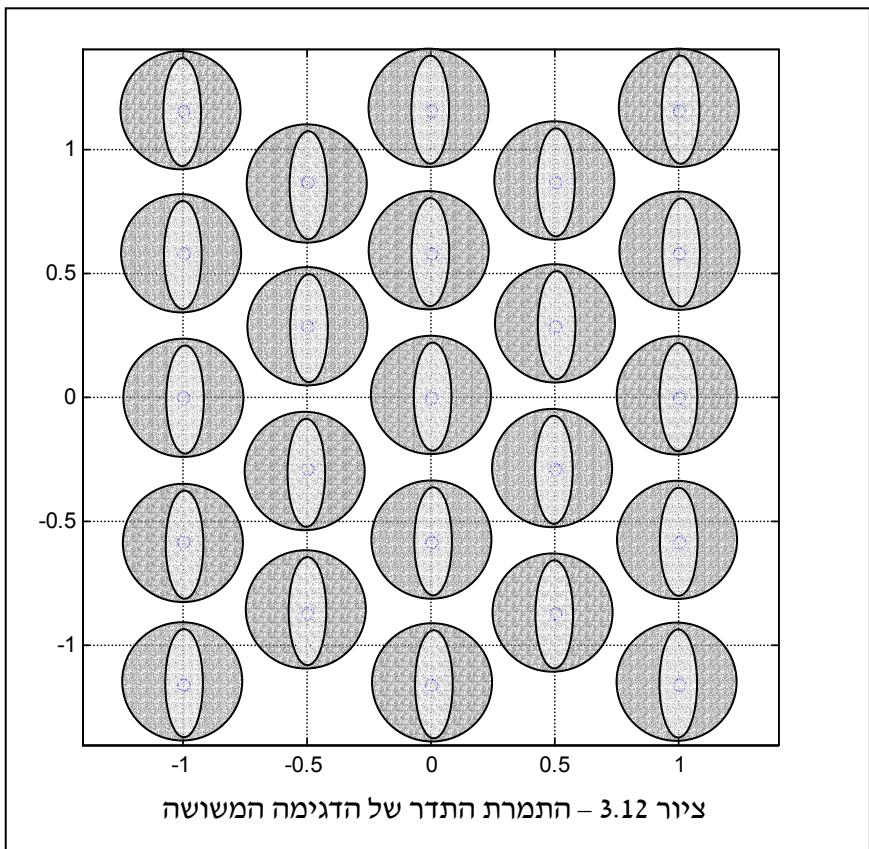
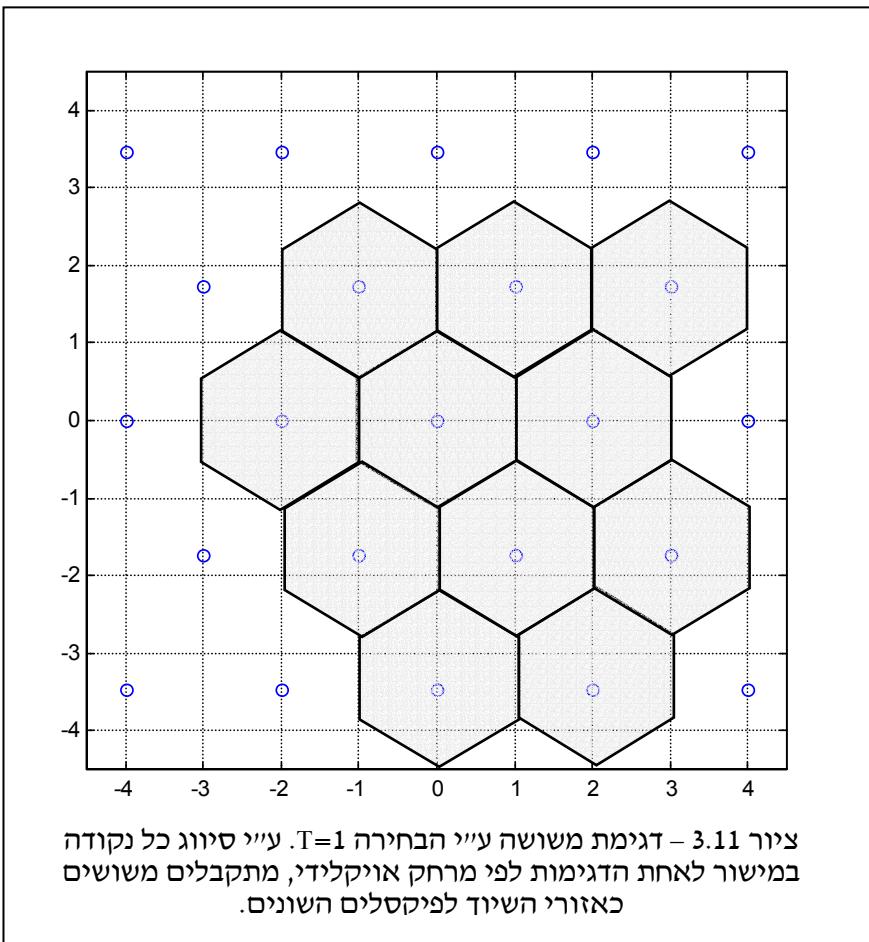
$$V = \begin{bmatrix} T & T \\ \sqrt{3}T & -\sqrt{3}T \end{bmatrix}$$

ומבנה שריג הדגימה המתואר בציור 3.11 עברו הבחירה $T=1$, ולכן, הציפיות היא דגימה אחת לשטח של 1.3.461. שריג מסרק ההלמים בתדר מתואר בציור 3.12, ומתקבל ע"י המטריצה

$$V^{-T} = \begin{bmatrix} 0.5 & 0.5 \\ 0.2887 & -0.2887 \end{bmatrix}$$

והציפיות של שריג זה הפוכה, דהיינו -0.2887 . בציור 3.12 אנו מניחים צורה כלשהי להתרמת התדר של האות המקורי $(x,y)f$ (מעגל ובתוכו אליפסה), ולכן מצליחים לתאר את האות הדגום בתדר במלואו. כפי שניתן לראות אין תופעת קיפול, כיון שאין חפיפה בין השכפולים השונים. עברו אותנו רציף, דגימה אחידה במקומות מחייבת מרוחך של 0.6 (בתדר) בין השכפולים – הן בציר x , והן בציר y . לכן, הציפיות בתדר נדרשת להיות $0.36 = 0.6^2$, ואז נקבל שהשכפולים נוגעים זה לזה. במקומות נקבל כי נדרשת דגימה אחת ל-2.8 יחידות שטח. כפי שאנו רואים דרך דוגמה זו, אותן בעל תמק מעגלי, דגימה ע"י שריג משושה יعلاה יותר.

פרק 3 – דגימה ושחזור של תמונות



פרק 4 – קוונטיזציה לתמונות

55	פרק 4 – קוונטיזציה לתמונות
55	4.1 קוונטיזציה סקלרית - בסיס
59	4.2 קוונטיזר מקס-לויד
67	4.3 קוונטיזציה עם שיקולים חוזתיים
71	4.4 קוונטיזציה וקטורית וקוונטיזציה צבע

בහינתם תמונה בראצ'ף, הראיינו בפרק הקודם כיצד יש לבצע עליה דגימה באופן נאות. לאחר פעולה הדגימה באהה בד"כ פועלה נספת בה ערכי הדגימה ממשיים מומרים למקבץ סופי ומוגדר מראש של ערכים אפשריים – פועלה זו קרויה קוונטיזציה. במסגרת זו אנו נעשו אבחנה בין קוונטיזציה סקלרית וקוונטיזציה וקטורית. הקוונטיזציה הסקלרית נחוצה לטיפול בתמונות שחור-לבן, והוקטורית לתמונות צבע.

חשוב לציין כי פעולה הקוונטיזציה (סקלרית או וקטורית) אינה נחלתו הבלעדית של נושא קליטת התמונה הרציפה בו דנים בהמרת התמונה הנтונה לקובץ מחשב. כפי שנראה בפרקים מתקדים, לקוונטיזציה קשר הדוק לדחיסת אותות, וכן להמרתם לצורך הדפסה. במסגרת פרק זה נציג את העקרונות תוכן שימת דגש על יישום אחד – קליטת תמונה והמרתה מראצ'ף.

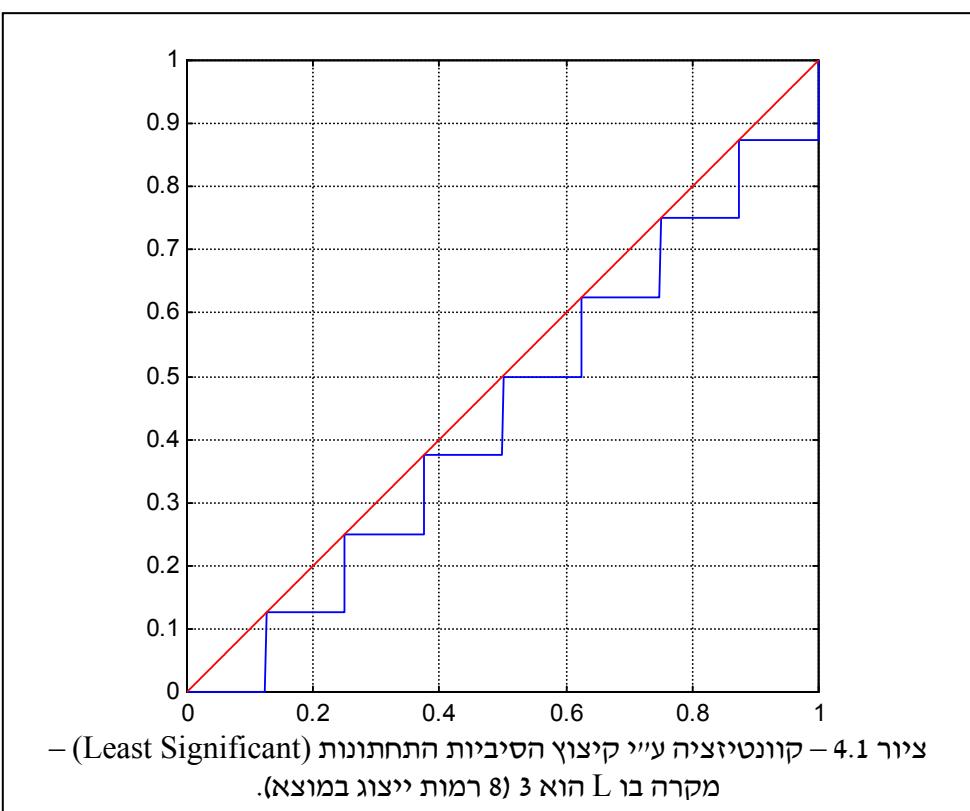
4.1 קוונטיזציה סקלרית - בסיס

נניח כי נתונה לנו תמונה דוגמה אשר מסומן כ- $[n,m]$. ערכי הדגימות הינם מעל המשיים וטוחה הערכים מצוי באינטראול $[0,1]$. כזכור, הערך 0 מתייחס לגוון שחור מוחלט, והערך 1 לגוון לבן מוחלט. ברור כי לשם ייצוג נאות של ערכי רמות האפור נידרש לאינסוף סיביות. מוצע אם כך להשתמש ב- L סיביות (L הוא ערך קבוע מראש וערכים הסבירים יהיו בתחום 16-1). השאלה המתבקשת היא כמובן כיצד להמיר את הערכים הרציפים לערכים הדיסקרטיים. כיוון ש- L סיביות מייצגות L^2 ערכים שונים, אנו נדרשים לבחור L^2 רמות ייצוג שונות.

אחת הפעולות הפחותות ביותר אותן ניתן להציג היא פועלת קיצוץ לסיביות הנמוכות. בגישה זו, עליינו לייצג את רמת האפור בקבוצת L הסיביות הגבוהות ביותר (Most Significant). מתמטית, פעולה זו נעשית ע"י המשוואה

$$Q(f) = \left\lfloor \frac{2^L \cdot f}{2^L} \right\rfloor$$

לדוגמה, עבור $L=2$, נקבל כי הערכים בתחום $[0, 0.25]$ ימופו לאפס, הערכים באינטרול $(0.25, 0.5]$ ימופו לגובה 0.25 וכן. באופן דומה ניתן להראות כי במצב הקוונטייזר ייכננו רק 4 ערכים אפשריים – 0, 0.25, 0.5 ו- 0.75. לערכי הכניסה בהם חל שינוי במצב נקרא רמות החלטה, ולערך המוצע האפשרי נקרא רמות הייצוג. ציור 4.1 ממחיש את הקשר בין הכניסה למוצע של הקוונטייזר המוצע עבור $L=3$. רמות הייצוג הם מיקומי הקווים האופקיים בפונקציית המדרגות על הציר האנכי, ורמות ההחלטה הם מיקומי הקווים האנכיים על הציר האופקי. כפי שהוא rõ, הקשר מאופיין ע"י פונקציית מדרגות.

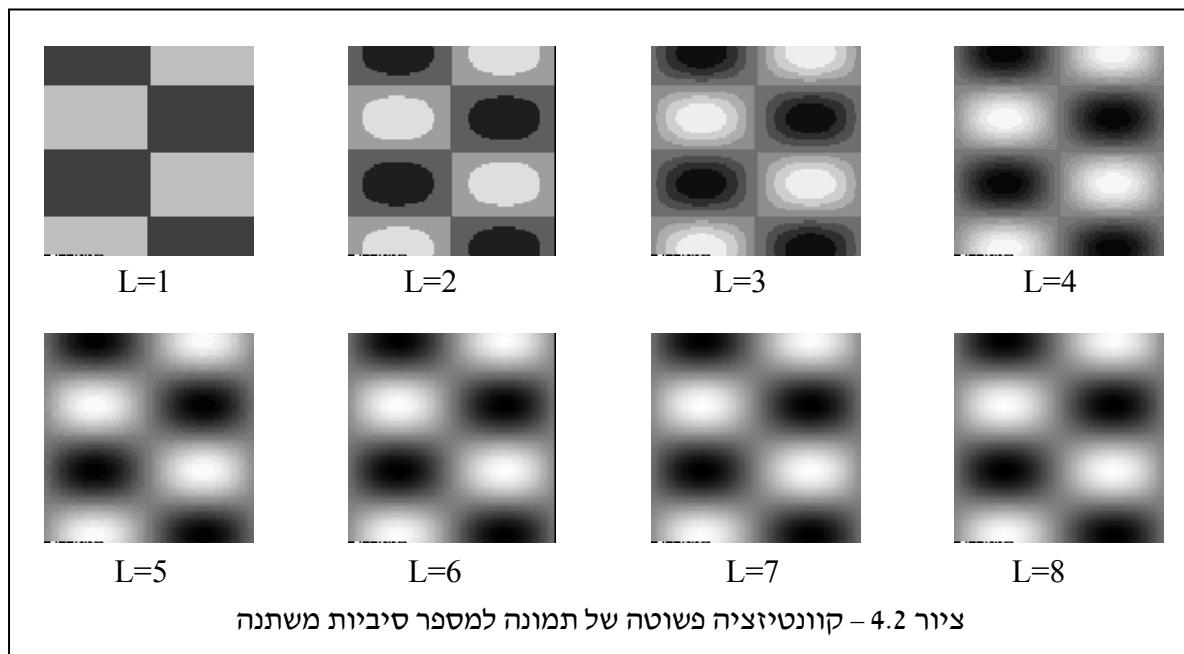


קוונטייזר מושלם היה קוונטייזר בעל מצב אחד זהה לכינסה, כיוון שאז אין שגיאה. ברור כי שימוש במספר סופי של סיביות לא מאפשר קוונטייזר כזה. כפי שניתן לראות, בגישה המוצעת תמיד נקבל כי מצב הקוונטייזר נמוך מכיניסתו. השגיאה בין הכניסה ליציאה תהיה אפס בנקודות המפגש בין הדרוג לקו האלכסוני (המייצג את מערכת היחידה), ומקסימלי בנקודות הברך שביניהם. ניתן לשפר את הפעולה הנילוליתית את השגיאה ע"י הכנסת תוספת למוצע:

$$Q(f) = \frac{\lfloor 2^L \cdot f \rfloor}{2^L} + 2^{-(L+1)} = \frac{\lfloor 2^L \cdot f \rfloor + 0.5}{2^L}$$

התוספת הנgrossה לגראף המדרגות לנوع מעלה ולהתמקם סביב מערכות היחידה (הקו האלכסוני), ואمنם, השגיאה בין הכניסה ליציאה תהיה קטנה יותר.

הגישה שתוארה פשוטה למדי ולכן פופולרית. בתמונות אופייניות, ביצוע קוונטייזציה כזו ליותר מ- 5 סיביות (32 רמות ייצוג) מניבה על פי רוב תוצאה בה הקוונטייזציה לא מרגשת לעין אנושית. כשירדים מ- 5 סיביות ניתן לראות בפועל את פועלן הקוונטייזציה, ונוצרות תופעות צורמות. אלו נדונם בדרכים לתיקון תופעות אלו בהמשך. צירור 4.2 מראה תמונה פשוטה (תמונה הסינוס הדווימדי), עם קוונטייזציה למספר סיביות משתנה, וממחישה תופעה זו. הקוונטייזציה שבוצעה משתמש בנוסחה עם העדכנית (הכוללת את התוספת להזות הגראף לפני מעלה).



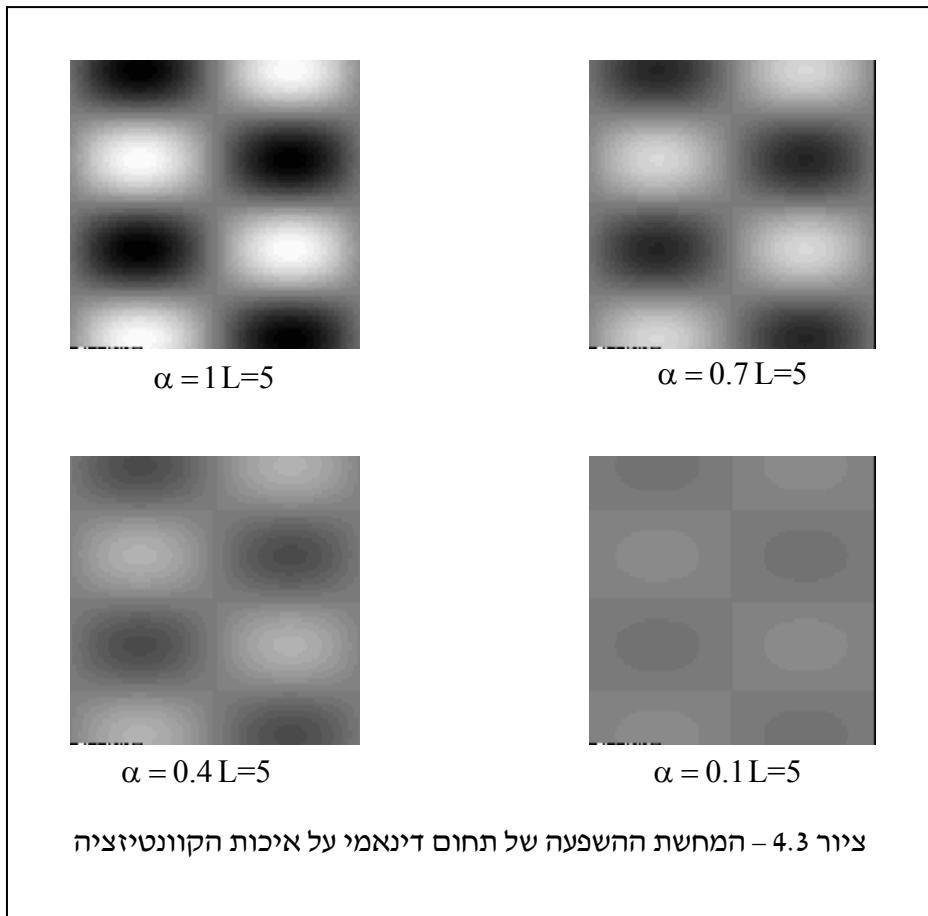
נzieי בעט ניסוי מעניין. ניקח את אותה תמונה הממלאת ערכי בתחום $[0,1]$ ומכוז את התחום הדינامي שלו עיי הקשר

$$\hat{f}[m,n] = \alpha[f[m,n] - 0.5] + 0.5$$

כאשר α בתחום $[0,1]$. עבר $\alpha = 0.5$ יתקבל כי התמונה בעט מקבלת ערכי רמות אפור בתחום $[0.25,0.75]$, ככלומר התחום הדינامي של התמונה קטן. ביצוע קוונטייזציה עם כמה סיביות נאותה (נניח 5 סיביות) בגישה שתוארה לעיל תיצור בעיה. קוונטייזציה על התמונה המקורית תיתן איות נאותה, אך

פרק 4 – קוונטייזציה לתמונות

עם כיווץ התחום הדינامي נקבל כי הקוונטייזציה מניבה תוצאה צורמת. צייר 4.3 מראה תוצאה זו עבור שימוש ב- 5 סיביות ושינוי α .



ניכר שוב כי עבור ערכי α קטנים, 5 סיביות אינן מספיקות. הסיבה לכך ברורה – עבור תחום דינامي קטן, רק חלק מ- 32 רמות הייצוג האפשרות באוט לידי ביתוי, וכל האחרות מתבזבזות! רעיון נאות יהיה, לפיכך, ביצוע קוונטייזיה שתלויה בתוכן התמונה, כך שבכל התמונות הניל' יונצלו כל רמות הייצוג באופן מיטבי. זהו בדיקת הרעיון מאחוריו לקוונטייזר מקס-לoid שיעוצג בסעיף הבא.

הקוונטייזיה שתוארה לעיל (על שתי גרסאותיה) קרויה אחדה (יוניפורמיית) כיון שהמרוחחים בין רמות הייצוג השונים או בין רמות החלטה השונות קבועים. נשאלת השאלה: מדווד לא לאמץ תפיסה בה פונקציה המדורגת כללית ומאופיינת ע"י רמות החלטה וייצוג כלשהם? ואם כך, כיצד נחליט מה תהינה הרמות הנאותות? במקרה הכללי, פעולה הקוונטייזיה תהיה פונקציה מדרגות המיפה את האינטראול $[0,1]$ לשידרה של $L^2 = N$ גבהים שונים באופן המתוואר בצייר 4.4. את רמות הייצוג ורמות ההחלטה נסמן ב- $\{f_k\}_{k=0}^N$, $\{f_k\}_{k=1}^N$, כמפורט בצייר.



4.2 קוונטייזר מקס-לoid

קוונטייזר מקס-loid (הקרויה על שם של שני חוקרים שפיתחוו בנפרד ובאופן בלתי תלוי) מציע קוונטייזציה התלויה בתמונה המוונת לה Mach, וקוונטייזציה שתרשה פונקציה מדרגות לא יוניפורמתית מאידך. הרעיון המרכזי הוא להציג את הקוונטייזר הכי טוב במובן זה שהמרחב בין כניסה ליציאה יהיה הקטן ביותר האפשרי, בכפוף לכך ש모פעלת פונקציה מדרגות על הכניסה. המרחק נמדד כ_qty ריבועית ממוצעת – ממד נוח מתמטי שלא תמיד מתלכד עם תוכנות מערכות הרהצה. הנעלמים בבעיה הם הפרמטרים של פונקציית המדרגות, דהיינו, רמות ההחלטה ורמות הייצוג. בניסוח מתמטי זה ייראה כך:

$$\underset{\{f_k, r_k\}_k}{\text{Minimize}} \quad \frac{1}{|\Omega|} \sum_{m,n \in \Omega} (f[m,n] - Q\{f[m,n]\})^2$$

בניסוח המתואר קשה להאמין שהבעיה בת פתרון מעשי. עם זאת ע"י מניפולציות פשוטות יחסית אנו נראה כי הבעיה פשוטה לאין ערוך מהנדסה. נגידר את הקבוצה Ω_k zusätzlich כל הנקודות $[n,m]$ בהן רמת האפור היא בתחום $(r_{k-1}, r_k]$, דהיינו:

$$\forall k = 1, 2, 3, \dots, N: \quad \Omega_k = \{[m,n] \mid r_{k-1} \leq f[m,n] < r_k\}$$

נוכל לפרך את הסכימה בביטוי שהוגדר ל- N נתחים באופן הבא :

$$\sum_{m,n \in \Omega} (f[m,n] - Q\{f[m,n]\})^2 = \sum_{k=1}^N \sum_{m,n \in \Omega_k} (f[m,n] - Q\{f[m,n]\})^2$$

אבל, ברור כי כל ערכי הקוונטיזציה של $f[m,n]$ יהיו הערך הקבוע f_k , ולכן הביטוי הופך להיות :

$$\sum_{m,n \in \Omega} (f[m,n] - Q\{f[m,n]\})^2 = \sum_{k=1}^N \sum_{m,n \in \Omega_k} (f[m,n] - f_k)^2$$

נתבונן בביטוי המופיע בסכימה הפנימית. ערכי $f[n,m]$ בו יגעו בתחום $(r_{k-1}, r_k]$. לערך מסוימים באינטראול זה, r^* , כל הנקודות $[n,m]$ המתקבלות ערך זה יניבו אותו ערך בדיקון $- (r^* - f_k)^2$, וערך זה יוכפל במספר הנקודות $[n,m]$ שערכן הוא r^* – מספר זה יסומן ב- $P(r^*)$. לכן, ניתן לסקום את הביטוי הפנימי באופן אחר, בו אנו נעים על כל הערכים בתחום $(r_{k-1}, r_k]$ בקפיצות אינפיטיסימליות, וכן נגד כל ערך לחבר את הביטוי הניל כשהוא מוכפל בכמות הנקודות הנאותה. כיוון שנעים על פני ערכים ממשיים בקפיצות אינפיטיסימליות, הסכימה הינה אינטגרציה מהצורה :

$$\sum_{m,n \in \Omega_k} (f[m,n] - f_k)^2 = |\Omega| \cdot \int_{r_{k-1}}^{r_k} (r - f_k)^2 P(r) dr$$

ההכפלה ב- dr פשוט מנורמלת את הביטוי הכלול, ולכן נדרש ההכפלה בגודל התמך. בעצם, הפונקציה $P(r)$ היא פונקציה ההסתברות לקבלת רמת אפור כלשהי, ולכן עליה לקיים

$$\forall f, \quad P(r) \geq 0, \quad \int_{-1}^1 P(r) dr = 1$$

הביטוי המתקבל אם כך הוא

$$\text{Minimize}_{\{r_k\}_{k=1}^{N-1}, \{f_k\}_{k=1}^N} \sum_{k=0}^N \int_{r_k}^{r_{k+1}} [r - f_k]^2 P(r) dr$$

פרק 4 – קוונטייזציה לתמונות

וזהו הביטוי אותו נרצה להביא למינימום. נשים לב לכך ש- r_0 ו- N ידועים כיון שאלה ערכי הקצה. כדי להשיג את המינימום علينا לגזר את הביטוי ביחס לנעלמים (רמות ההחלטה ורמות הייצוג) ולהשווות לאפס. גזירה זו מניבה :

$$\frac{\partial}{\partial f_k} \sum_{k=0}^N \int_{r_k}^{r_{k+1}} [r - f_k]^2 P(r) dr = \int_{r_k}^{r_{k+1}} [f_k - r] P(r) dr = 0$$

$$\frac{\partial}{\partial r_k} \sum_{k=0}^N \int_{r_k}^{r_{k+1}} [r - f_k]^2 P(r) dr = -[r_k - f_{k+1}]^2 P(r_k) + [r_k - f_k]^2 P(r_k) = 0$$

בשתי הנзорות, הסכימה החיצונית נופלת כיון שהנעלם נתיחס לגורם אחד (במקרה הראשון) או שניים (במקרה השני), וכל היותר מהווים לעניין הגזירה קבוע. הגזירה לפי f_k נכנסת פנימה לאינטגרל ומתבצעת בתוכו. הגזירה לפי r_k מניבה שני איברים כיון שנעלם זה מהוות פעמי גבול תחתון ופעמי גבול עליון באינטגרלים. גזירה לפי גבול האינטגרל הנה גובחה של הפונקציה הפנימית לאינטגרל בנזודה זו, וכך מתתקבלת התוצאה. אנו עושים שימוש בנוסחאות הבאות :

$$\frac{\partial}{\partial y} \int_{-\infty}^y f(x) dx = f(y)$$

$$\frac{\partial}{\partial y} \int_y^{\infty} f(x) dx = -\frac{\partial}{\partial y} \left[\int_{-\infty}^{\infty} f(x) dx - \int_{-\infty}^y f(x) dx \right] = -f(y)$$

$$\frac{\partial}{\partial y} \int_a^b f(x, y) dx = \int_a^b \frac{\partial}{\partial y} f(x, y) dx$$

ביצוע מספר פעולות אלגבריות מתקבל :

$$k = 1, 2, \dots, N: \quad f_k = \frac{\int_{r_k}^{r_{k+1}} r P(r) dr}{\int_{r_k}^{r_{k+1}} P(r) dr}$$

&

$$k = 1, 2, 3, \dots, N-1 \quad r_k = \frac{f_{k+1} + f_k}{2}$$

משמעות תוצאה זו היא שבהינתן רמות הייצוג, רמות ההחלטה מתקבלות ע"י מיצוע פשוט, כך שבין כל שתי רמות ייצוג עוקבות קיימת באמצעות בדיקת רמת החלטה. הביטוי الآخر (מצד שמאל) קובע כי

בහינתן רמות ההחלטה, רמות הייצוג מתכבלים ע"י מרכז מסה של פונקציית הפילוג P באינטראול

$$[r_k, r_{k+1}]$$

ברור כי משוואות אלה סתומות, ולא ניתן לחץ ביטויים סגורים עבור הנעלמים השונים. באופן מעשי, מציאת פתרון נעשית באופן איטרטיבי, בו מתחילה נניח שמודול של רמות ההחלטה (או רמות הייצוג – זה לא משנה), וمعدכנים את רמות הייצוג לפי המשוואות הנ"ל. עם הערכים המעודכנים חזרים ומשנים את רמות ההחלטה, וחזור חלילה. אתחול סביר לתהlik הוא בחירת רמות איחודות כפי שתואר קודם, וביצוע איטרציות ממש. התהlik האיטרטיבי המוצע מובטח להקטין את השגיאה הכלולית (ערך הפונקציה המובאת למינימום) בכל איטרציה, מובטח להתכנס לערכים יציבים כלשהם, אך בפירוש לא מובטח שהפתרון המתתקבל יהיה המינימום הגלובלי. בהחלט יתכן שהפתרון ייתקע על מינימום מקומי, ולכן חשוב לבצע את התהlik מספר פעמים עם איתחולים שונים ולבחרו את הטוב שבין הפתרונות.

בעיית המינימזה הלוקלית תליה באופן חזק באופן פונקציית ההסתברות $(z)P$. כאשר פונקציה זו מאוד לא חלקה, ישן מינימות מקומיות רבות, ובדרך כלל ההתקנסות תהיה לפתרון סמוך לערכי האיתחול עם שיפור עדין. פתרון מעשי לבעיה זו היא חילוק הפונקציה $(z)P$ בטרם עשיית שימוש בה.

שאלה חשובה שנותנה עד כה היא מקורה של הפונקציה $(z)P$. כזכור, התהlik מתחילה עבורנו בקבלת תמונה $[m,n]$ שערכיה ממשיים ובתחום $[0,1]$. נניח כי כמות הפיקסלים סופית ונתונה ע"י $|\Omega|$. אז, הפונקציה $(z)P$ נתונה כסכום של $|\Omega|$ פונקציות הלם חד-מידיות הממוקמות בערכי רמות האפור של כל פיקסל, דהיינו :

$$P(r) = \frac{1}{|\Omega|} \sum_{[m,n] \in \Omega} \delta(r - f[m,n])$$

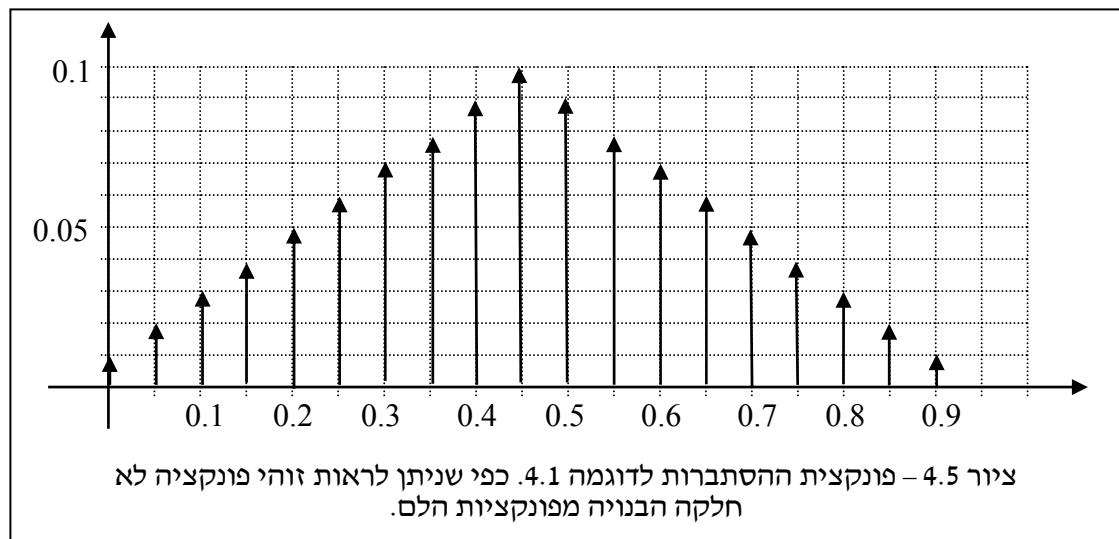
חלוקת ב- $|\Omega|$ נעשית לשם נרמול הפונקציה. וכפי שאנו רואים, במקרה של תמונה דוגמה, סביר למדי שפונקציה זו תהיה מאוד לא חלקה. חילוקה אפשרית ע"י ביצוע קוונבולוציה עם גאוסיאן ממורכו ברוחב σ , גודל הקובע את עצמת החלוקת, אז :

$$P_G(r) = \frac{1}{|\Omega|} \sum_{[m,n] \in \Omega} G(r - f[m,n], \sigma)$$

דוגמה 4.1: נניח כי נתונה תמונה בת 10 על 10 פיקסלים מעל התמך $[0,9] \times [0,9]$. רמת האפור בתמונה בנקודה $[m,n]$ נתונה ע"י $f[m,n] = 0.05 \cdot (m + n)$

$$f[m,n] = \frac{1}{20} \begin{bmatrix} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 \\ 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \\ 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 \\ 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 \\ 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 \\ 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 \\ 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 \\ 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & 18 \end{bmatrix}$$

בנייה עבורה תמונה זו את פונקציית ההסתברות. רמת אפור 0 מתקבל רק במקומות אחד – [0,0], רמת אפור 0.05 מתקבל בשני פיקסלים – [0,1] ו-[1,0], וכך הלאה עד רמת 0.45 שתתקבל בעשרה פיקסלים. הערך 0.5 יתקבל בתשעה פיקסלים וכוכי, ועד 0.9 שיתקבל בפיקסל אחד. תיאור גרפי של הפונקציה r נתון בציור 4.5.

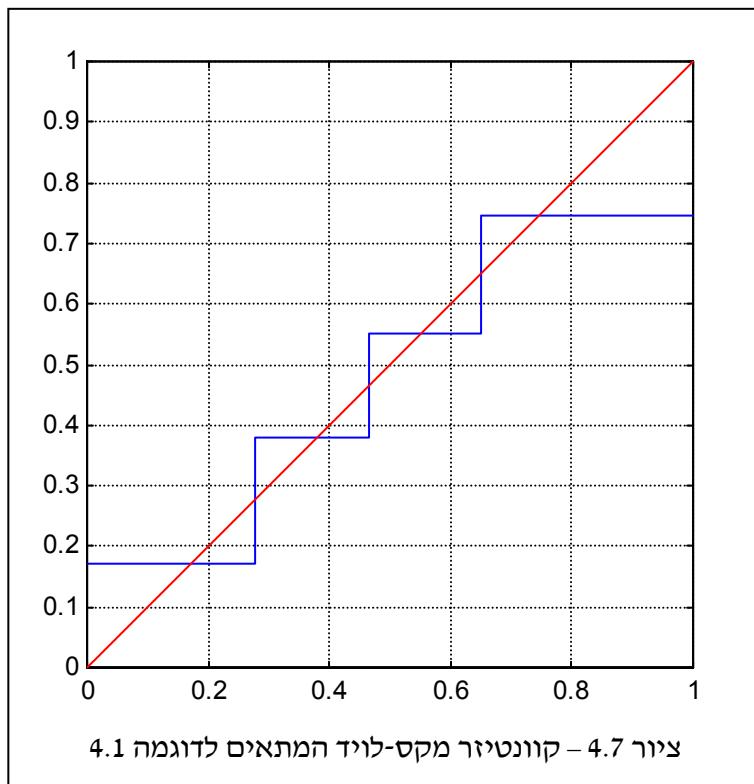
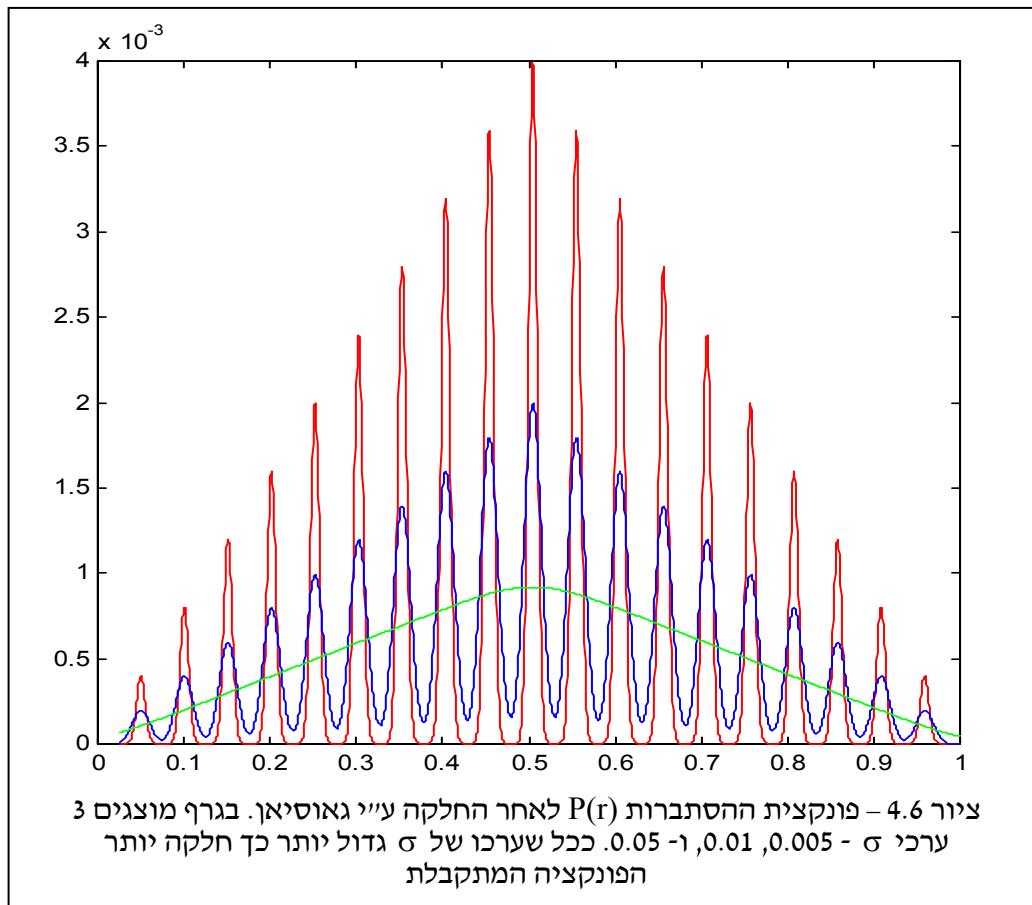


ציור 4.6 מציג תוצאת MATLAB לפונקציה זו תוך שימוש בגאוסיאן עם ערכים שונים של השונות σ , ואמנם ניתן לראות שההחלקה מצלילה יפה (תופעה מעניינת هي א-עם הגדלת σ – ס גובה הפונקציה הכללי יורד, וזאת כי יובו שסק השטח מתחת לגרף צריך להישאר 1).

נניח כעת כי רצוננו בקווונטייזר בין ארבע רמות ייצוג (כולם 2 סיביות לייצוג כל רמה). תוך שימוש בקירוב המוחלך של פונקציית רמות האפור, מהו הקווונטייזר האופטימלי? תוכנית לחישוב איטרטיבי כמפורט לעיל הניבה את התוצאה המתוארת בציור 4.7. ערכי רמות ההחלטה והיצוג לאחר 10 איטרציות הם

$$r: [0 \quad 0.2761 \quad 0.4656 \quad 0.6487 \quad 1.0000]$$

$$f: [0.1728 \quad 0.3794 \quad 0.5519 \quad 0.7455]$$



ניתן לראות מספר תופעות מעניינות בתוצאה שהתקבלה. תופעה ראשונה היא שהקוונטייזר חותר להיות סימטרי סביב ערך הכניסה 0.45 – תוצאה צפואה כיון שההסתברות סימטרית סביב ערך זה.

פרק 4 – קוונטייזציה לתמונות

תופעה שנייה היא שהמרוחחים בין רמות ההחלטה באמצע התחום קטנים משמעותית מהמרוחחים בקצבות. זאת כיון שבתחום סביבה 0.45 יש הרבה פיקסלים בעלי רמות אפור אלה, ולכן כדאי להעניק להם רמת ייצוג מוצלחת יותר אשר תקטין את השגיאה. לעומת זאת, בקצבות ישנה כמעט מועטה של פיקסלים ושגיאה בקוונטייזציה עבורה פחות כואבת.

דוגמה 4.2: נטפל בעת בתמונה שלה פונקציית הסתברות איחידה בתחום [0.2,0.6] :

$$P(r) = \begin{cases} 2.5 & 0.2 \leq r \leq 0.6 \\ 0 & \text{Otherwise} \end{cases}$$

מהו הקוונטייזר האופטימלי? מסתבר כי במקרה זה (ובעצם בכל מקרה של פילוג אחד מעל אינטראול כלשהו), הקוונטייזר האופטימלי הוא גם הקוונטייזר היוניפורמי שפוגשנו בתחילת פרק זה. לכן, במקרה זה, הקוונטייזר האופטימלי יהיה בעל רמות ההחלטה [1 0.1 0.3 0.4 0.5 0] ורמות הייצוג יהיו [0.25 0.35 0.45 0.55]. מדוע זה נכון? ובכן, אם נציב לתוך נוסחאות התנאים ההכרחיים את העובדה שהפילוג אחד קיבל כי באיזור התמך של הפילוג :

$$f_k = \frac{\int_{r_k}^{r_{k+1}} r P(r) dr}{\int_{r_k}^{r_{k+1}} P(r) dr} = \frac{\int_{r_k}^{r_{k+1}} r dr}{\int_{r_k}^{r_{k+1}} dr} = \frac{r_{k+1}^2 - r_k^2}{2(r_{k+1} - r_k)} = \frac{r_{k+1} - r_k}{2}$$

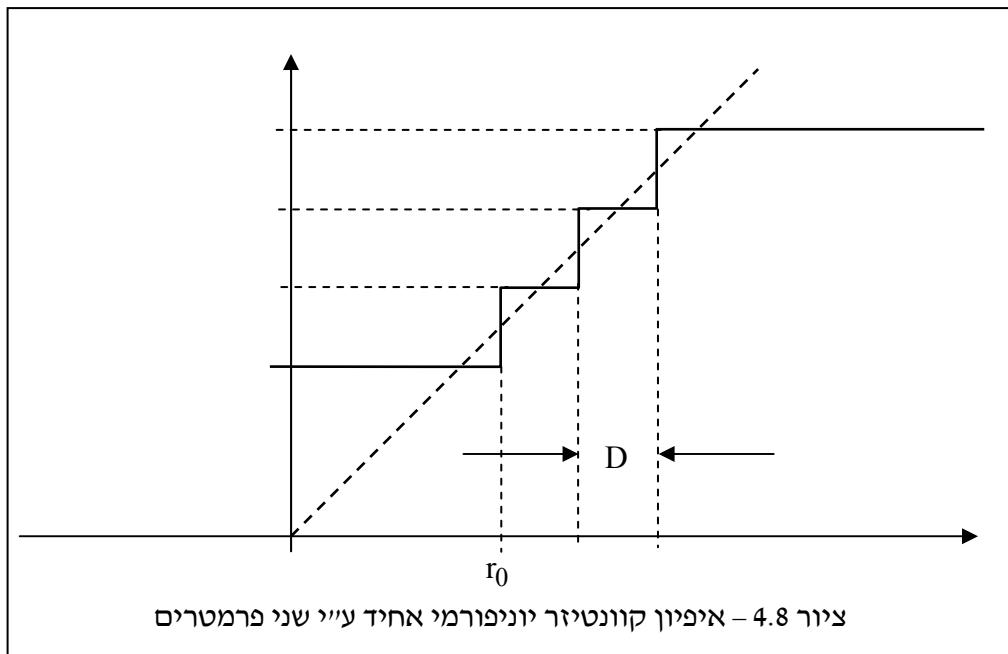
כלומר, גם רמות הייצוג הינם ממוצע של רמות ההחלטה (כזכור, קיימים כבר קשר הפוך שקובע כי רמות ההחלטה הן ממוצע של רמות הייצוג).

נושא אחרון במסגרת הדיון בקוונטייזרים אופטימליים הוא השימוש בקוונטייזר יוניפורמי אופטימלי. קוונטייזר יוניפורמי קל מאוד למימוש, ולכן מפתחה מאוד לעשوت בו שימוש, גם כשהוא אינו מתאים לפילוג הנתון. במקום $2N-1$ ערכים חופשיים (רמות ההחלטה ורמות הייצוג), יהיו בידינו עת רק 2 פרמטרים – מיקום המדרגה הראשונה וגודלו המרוחקים. השגיאה הריבועית תהיה :

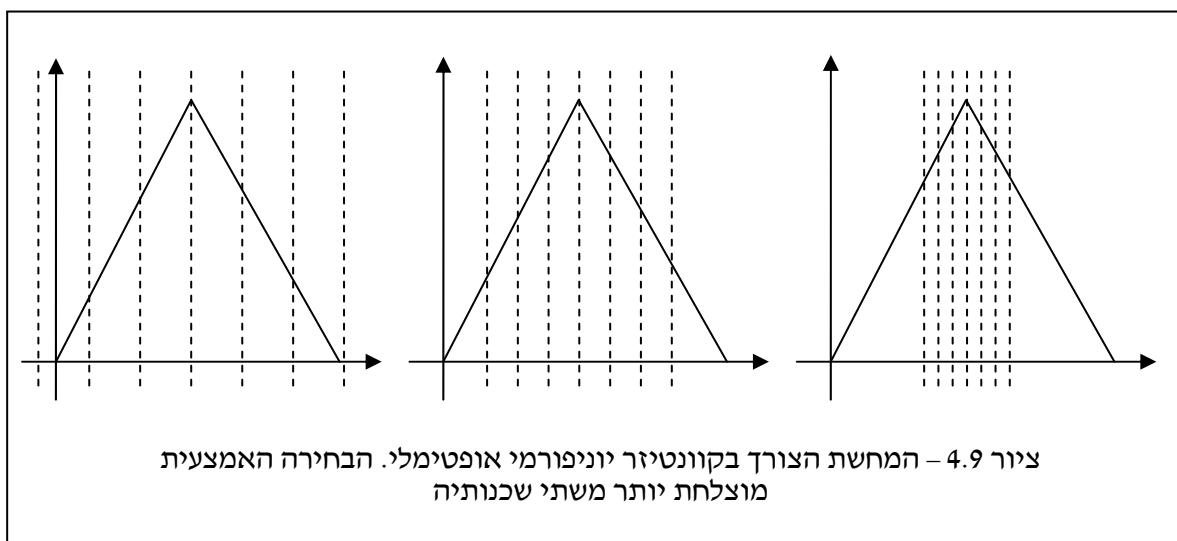
$$\begin{aligned} & \int_{-\infty}^{r_0} \left(r - \left[r_0 - \frac{D}{2} \right] \right)^2 P(r) dr + \\ & + \sum_{k=1}^{N-2} \int_{r_0 + (k-1)D}^{r_0 + kD} \left(r - \left[r_0 + kD - \frac{D}{2} \right] \right)^2 P(r) dr + \\ & + \int_{r_0 + (N-2)D}^{\infty} \left(r - \left[r_0 + (N-2)D + \frac{D}{2} \right] \right)^2 P(r) dr \end{aligned}$$

פרק 4 – קוונטייזציה לתרמונות

כל יתר הקוונטייזר מוגדר היטב מותוך שני ערכים אלה. בעית אופטימיזציה מענינית היא למצוא את הערכים D_0 , r_0 לקבלת שגיאה מינימלית. ציור 4.8 מראה כיצד נראה קוונטייזר כזה לארבעה רמות ייצוג.



ע"י גזירת הביטוי הנ"ל ניתן לקבוע באופן דומה תנאים הכרחיים על שני הנעלמים הנ"ל. ציור 4.9 ממחיש את השימוש בקוונטייזר יוניפורמי אופטימלי. הרעיון הבסיסי הוא לשחק עם ה- Scale של פונקציית הקוונטייזר ומיקומה היחסי לשם השגת שגיאה טובה יותר. בציור מתואר פילוג כלשהו ועליו את המיקומים של מדרגות הקוונטייזציה. ברור כי הבחירה בציור המרכזי תניב שגיאה קטנה יותר, בהשוואה לשתי האפשרויות האחרות.



4.3 קוונטיזציה עם שיקולים חזותיים

בתחילת פרק זה רأינו כי בקוונטיזציה עם מספר סיביות נמוך מתאפשרת תופעת Contouring, בה רואים עוקמים המתחמים איזורים שווי רמת אפור. קוונטיזר מקס-לoid יכול לסייע בפתרונה של הבעיה זו, אך מחייב תכנון של קוונטיזר לכל תמונה מחדש, דבר מטריד מדי. בפועל, מוביל להשתמש בקוונטיזר אחיד ושרירוני שאינו תלוי תמונה, ולפיכך נחוצים קרוב ל- 8 סיביות לקבלת איכות גבוהה. עם זאת, במצבים בהם לא ניתן להקצות 8 סיביות (אם מסיבות חומרה, זכרו או אילוצי תצוגה), הבעיה מתעורררת ומטרידה.

במסגרת סעיף זה נדגים כיצד ניתן לפתרו (לפחות חלקית) את בעיית ה- Contouring ללא שימוש בקוונטיזר מקס-לoid, ולאו שימוש בפעולות תלויות תמונה, תוך שימוש לפחות בסיביות. הרעיון המרכזי הוא לנצל את תכונות מערכת הראייה האנושית, כדי להשיג את השיפור המוחלט. תופעת ה- Contouring בולטות כל-כך כיוון שמערכת הראייה האנושית רגישה לשפות ואפילה עדינות. העוקמים המתקבלים הינם שפות לכל דבר, ולכן הם צורמים כל-כך. הפתרון, לפיכך, הוא שבירת שפות אלו, וזה מושג ע"י הוספת רעש לתמונה לפני הקוונטיזציה.

הreas הנוסף נדרש להיות רעש המפולג אחיד וסימטרי, חסר קורלציה מרחבית (זהינו – דגימותיו הנן תהליך iid – מפולגים באופן זהה וחסר תלות). פילוג אחיד נתון ע"י:

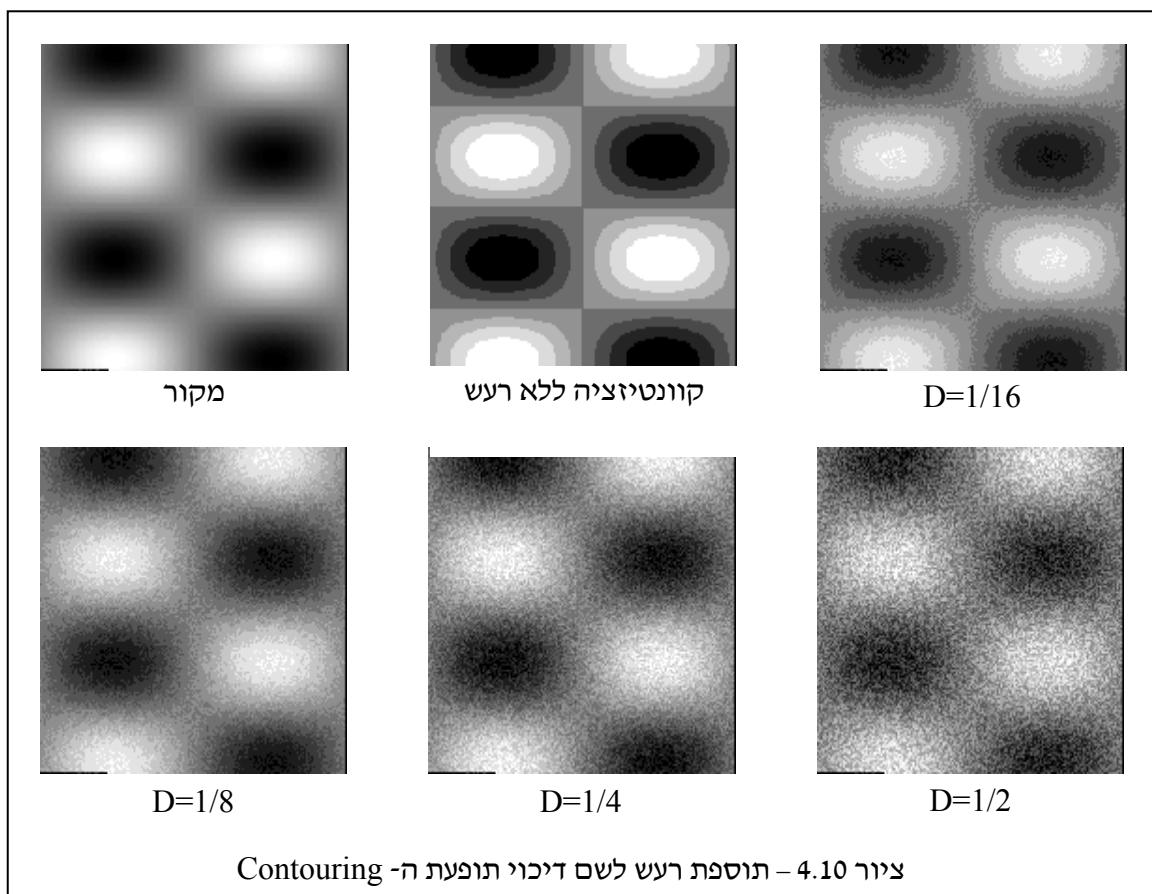
$$P(n) = \begin{cases} \frac{1}{2D} & -D \leq n \leq D \\ 0 & \text{Otherwise} \end{cases}$$

כאשר D היא אמפליטודת הרעש. אם התמונה עברת קוונטיזציה ל- R סיביות, האמפליטודה המומלצת לשימוש תהיה 2^{-R-1} . אם נסמן תמונה רעש בעלת פילוג אחיד וסימטרי סביר הראשית באמפליטודה 0.5 ע"י $[g[m,n]$, פועלות הקוונטיזציה המוצעת היא :

$$Qf[m,n] = \left\lfloor \frac{(f[m,n] + g[m,n] \cdot 2^{-R}) \cdot 2^R}{2^R} \right\rfloor + 2^{-(R+1)} = \frac{\lfloor 2^R f[m,n] + g[m,n] \rfloor}{2^R} + 2^{-(R+1)}$$

בציור 4.10 מתוארת תוצאה של תמונה הסינוס הדוו-מיידי לאחר קוונטיזציה ל- 3 סיביות עם עצמות שונות של הרעש (רעש המקורי בעל אמפליטודה של 0.5), וניכר שעבור עצמות רעש כפומלץ התוצאה נראהיה הרבה יותר טוב מלאה רעש כל.

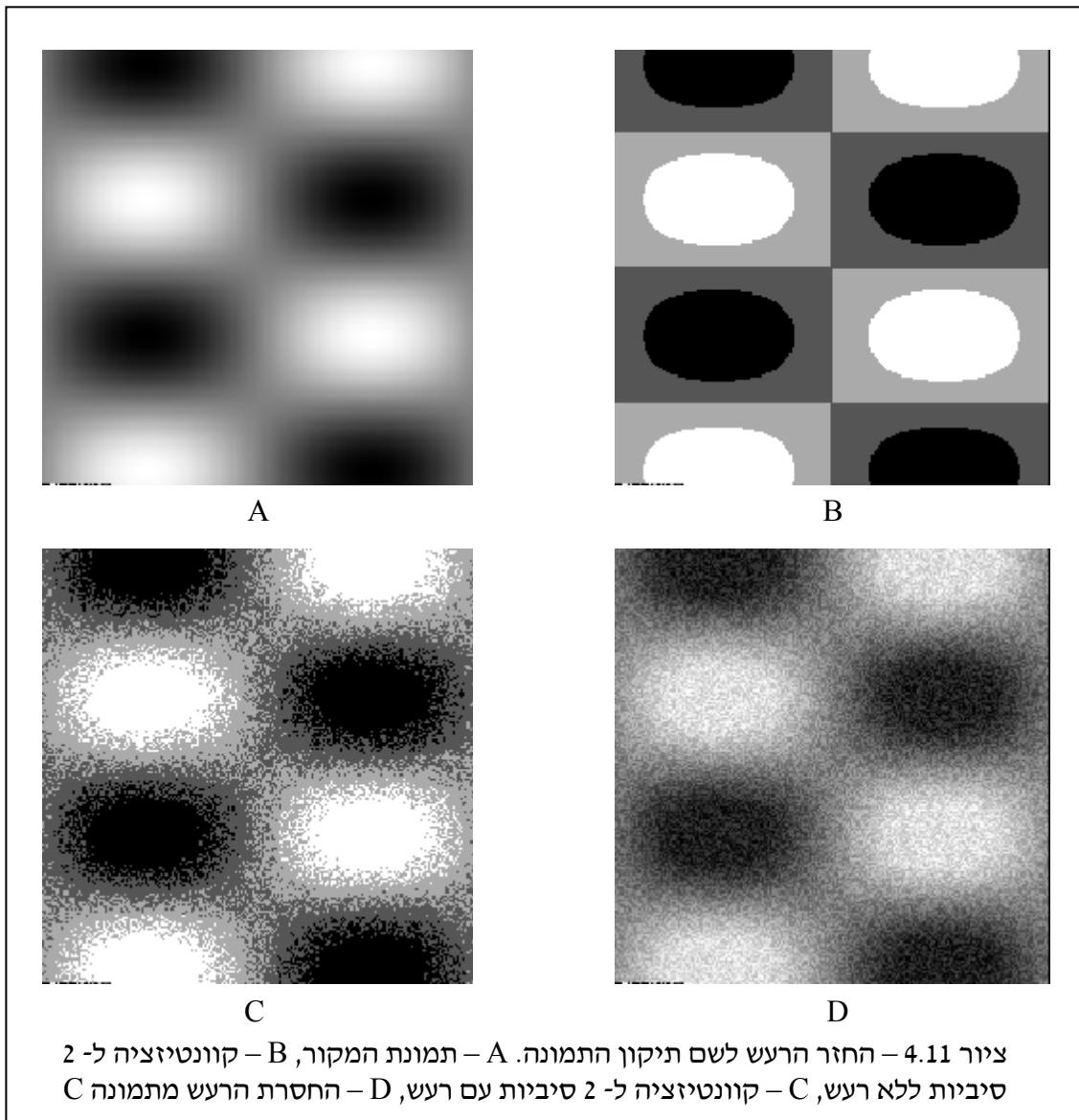
תוצאה מעניינת נוספת היא שאם עושים שימוש ברעש פסבדו-אקראי (כלומר, רעש מיוצר מחשב, הנitin לבניה מחדש בדיקוק מוחלט) לפי המאפיינים שתוארו, ניתן לחסיר את הרעש ממקור הקוונטיזר



ולקבל אותן קרוב למדוי למקור, בטרם הקוונטייזציה. תופעה זו ניתנת לניצול באופן הבא (וזהו רק דוגמא אחת אפשרית): נניח כי נתונה לנו תמונה ואנו צריכים להדפסה ולהעבירה לגורם אחר. נניח כי המדפסת יכולה להציג רק עם R סיביות (מדפסת אופיינית יכולה עם סיבית אחת). נבצע את הקוונטייזציה עם רעש פסבדו-אקראי ונקבל הדפסה שלה. מתקבל התמונה יוכל לסרוקה ולקבל קבוע תמונה לאחר קוונטייזציה. הוא יוכל לשחזר במדוק את הרעש ולהחסירו מהתמונה, וכך לקבל תוצאה קרובה לתמונה המקורית. צייר 4.11 ממחיש את התמונות בתהליך זה לשימוש בסיבית אחת.

מדוע הרעש מסיע? כיצד נוכלה כאן מערכת הראיה? כפי שנאמר קודם, המרכיב הצורם ביותר הם קווים השפה בין אזורים שווים רמת אפור, וקוויים אלו חלקיים למדוי, ולכן קופצים לעין. הוספת רעש אקראי עם ערכים חיוביים ושליליים בעוצמה של מרווה אחד שלם בקונטניזר המופעל גורמת לקוונטייזציה לkapo'ץ רמות מעלה מטה באופן אקראי מסביב לשפה. קפיצות אלו שובורות את קו התפר, ולכן מעליות אותו. מאידך, כיוון שהרעש סימטרי סביב האפס, הגוון הממוצע יותר כבמוקר. למעשה, תכונה נוספת של מערכת הראיה כאן לידי ביתוי. העין מחליקה את הרעש ע"י דיכוי התדרים הגבוהים. החלקה זו נותנת כי איןנו רואים בפועל את הקפיצות המהירות מרמה לרמה, אלא את הגוון הסביבתי הכלול, והוא זה נאמן יותר למקור.

דוגמה 4.2: נתונה תמונה עם רמת אפור קבועה ושויה 0.3, ולתמונה זו מבוצעת קוונטייזציה ל- 4 רמות ללא רעש. קוונטיזור זה מפה את כל הערכים בתחום $[0.25, 0.5]$ לערך 0.375. לכן נקבל במוואה הקוונטיזור תמונה קבועה שרמתה היא 0.375, והשגיאה בין המקור לתוצאה קבועה ושויה 0.075. כאשר נסיף רעש אקראי בעוצמה של 0.125, ערכי התמונה ינעו באקראי בתחום $[0.175, 0.425]$. הערכים



ציור 4.11 – החזר הרעש לשם תיקון התמונה. A – תמונה המקורי, B – קוונטייזציה ל- 2 סיביות ללא רעש, C – קוונטייזציה ל- 2 סיביות עם רעש, D – החזרת הרעש מהתמונה

בתחום $[0.175, 0.25]$ ימופו ל- 0.125, והערכים בתחום $[0.25, 0.425]$ ימופו ל- 0.375. אם ניקח איזור המכיל 10 על 10 פיקסלים, בשל הפילוג האחדיד קיבל שבסמוצע, שלוש עשרירות מהפיקסלים יקבלו ערך 0.125, ושבע עשרירות מהם יהיו 0.375. חלוקה זו נובעת מיחסיו אורכי האינטראולים. רמת האפור הממוצעת באיזור תהיה לפיכך:

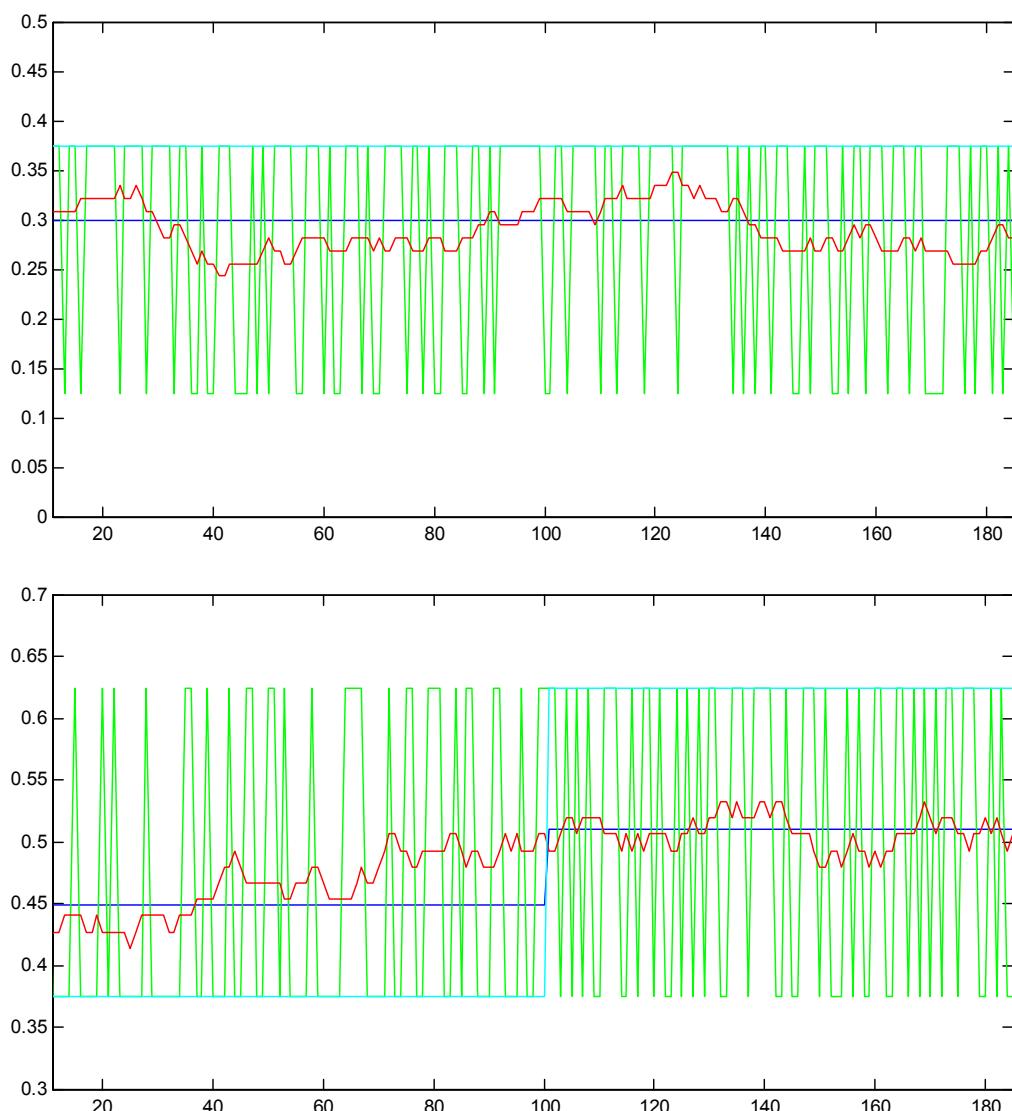
$$\frac{3 \cdot 0.125 + 7 \cdot 0.375}{10} = 0.3$$

וזהו בדיק הערך המקורי. הרווחנו כאן דיווק טוב יותר ברמת האפור על חשבון אובדן דיווק ברזולוציה, כיוון שכעת רמת האפור מוכבתת מסביבה של פיקסלים.

ציור 4.12 מראה גרפ' חד-מימדי של תהליך הקוונטייזציה שתואר בדוגמה, וממחיש באופן זה את שהתרחש. אנו רואים כי האות הממוצע (באדום) קרובה למקור הרבה יותר, בהשוואה לקוונטייזציה ללא רעש. לעומת זאת, הגרפ' השני נותן אות בעל מדרגה במקומם כלשהו מ- 0.45 ל- 0.51. רואים הפעם כי

פרק 4 – קוונטיזציה לתמונות

מיקום המדרגה מתואר بصورة מדוייקת ע"י הקוונטיזר ללא הרעש, אך עם סטיות רמה גדולות. הקוונטיזציה שעשוה שימוש ברעש כמעט ולא נותרת ביטוי למדרגה שהיתה.



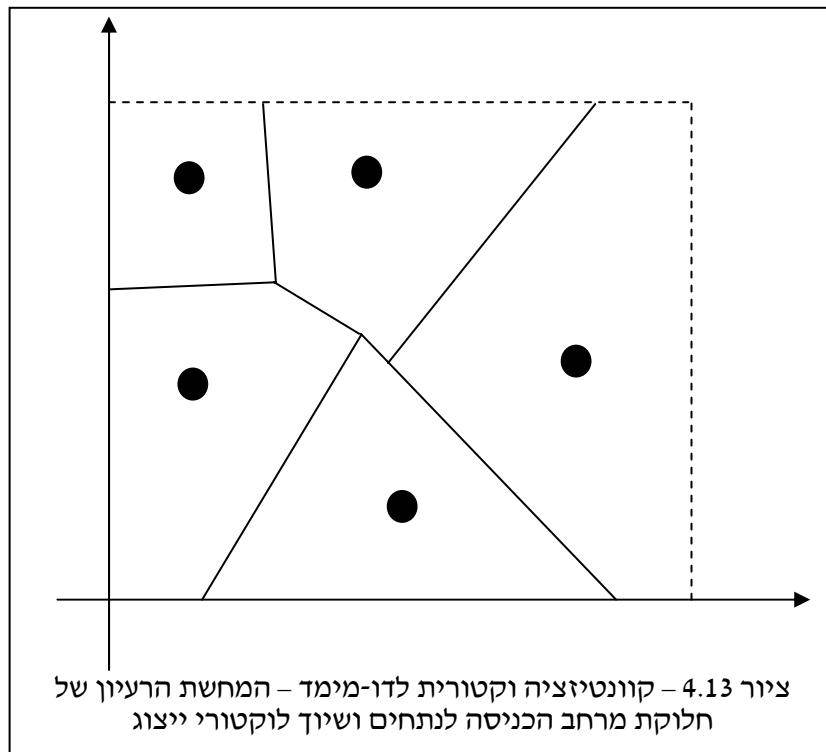
ציור 4.12 – קוונטיזציה עם רעש כאמצעי להמיר רוזולוציה מרוחבית בדיק רמות אפור. בשני הגרפים אותן מקור (כחול), קוונטיזציה ללא רעש (תכלכל), קוונטיזציה עם רעש (ירוק), ותוצאת החלקת הגרף הירוק ע"י מיצוע אחד על סביבה של 19 (אדום).

כפי שנראה במסגרת פרקים שיבאו בהמשך, למתואר כאן קשר ישיר לתהליך ה-Half-Tonning הנחוץ במדפסות. גם שם, שימוש בתוכנות מערכת הראית מניב ירידה כמות הסיבות הנחוצות לייצוגה של התמונה.

4.4 קוונטיזציה וקטוריית קוונטיזציה צבע

עד כה עסקנו בתמונות שחור-לבן, בהן הקוונטייזר מوزן בערכים סקלריים - רמות האפור. תמונת צבע, לעומת זאת, מאופיינית בכך שהיא כוללת מוקום [a,m] ע"י שלושה ערכים של רמות אפור – R, G, B, כאשר אחד מהם מסמן ממשי בתחום [0,1]. כשאין בעיה של כמות סיביות, מקובל לבצע קוונטיזציה יוניפורמיית לכל שכבת צבע בנפרד – 8 סיביות, וכך לקבל כי 24 סיביות מייצגות את מרחב הצבעים. קוונטיזציה צבע חשיבות מעשית רבה. במוגן גדול של מסכי מחשב תומך ברטיס המשך במספר צבעים מצומצם (בג"כ 256 צבעים שונים – 8 סיביות בלבד). לכן, ביצוע קוונטיזציה סקלרית לכל שכבת צבע בנפרד – 8 סיביות אינה מתאימה.

קוונטיזציה צבע היא קוונטיזציה וקטוריית, בה הקוונטייזר מזון בוקטור בעל שלושה מרכיבים, ומণיב וקטורי אלטרנטיבי בעל שלושה מרכיבים. הרעיון הבסיסי בקוונטיזציה וקטוריית הוא חלוקת מרחב הכניסה ל- N נתחים שונים, ובוחן לכל נתח כמה "רמת ייצוג" (ובעצם – וקטור ייצוג) אחד. בהינתן כניסה כלשהי, המרתת לוקטור הייצוג תיקבע לפי הקבוצה אליה היא שייכת. ציור 4.13 ממחיש רעיון זה לוקטורים דו-מימדיים. בציור, כל וקטור דו-מימדי מומר לאחד מתוך 5 וקטורי ייצוג אפשריים, ואלו הם ה-"צבעים" האפשריים (מפת השיווק המתקבלת נקראית דיאגרמת Voronoi).



לקוונטיזציה וקטוריית בכלל יש יישומים רבים גם לצורך דחיסה וגם לצורך הדפסה, בדומה לנאמר על קוונטיזציה סקלרית – בסוגרת פרק זה נתמקד בנושא קוונטיזציה כתהילך שבא אחרי דגימה.

פרק 4 – קוונטייזציה לתמונות

הפתרון הפשוט ביותר ליותר לקוונטייזציה וקטוריית הוא הפרדת הטיפול בין שלושת שכבות הצבע השונות, וביצוע קוונטייזיה יוניפורמת לכל אחת מהן בנפרד. כך למשל, אם נדרשים 256 צבעים ייצוג, יוכל לקחת את התמונה ולבצע קוונטייזיה לשתי שכבות ממנה – 6 תחומיים, ולשכבות הצבע השלישי – 7 תחומיים. קוביית היחידה התלת-מימדית $[0,1] \times [0,1] \times [0,1]$ תחולק בצורה כזו – $252 = 7 \cdot 6 \cdot 6$ מלבנים – תלת-מימדיים שווים נפח הממלאים את הקובייה. ארבעת הצבעים הנוטרים לא חייבים להיות מבוזזים – بد"כ במסכי מחשב, הצבעים הנוטרים משמשים כצבעים שמורים לרקע, חלונות וכו'.

הגישה המתוארת מקבילה להצעתו לשימוש בקוונטייזר יוניפורמי לתמונות שחור-לבן. חסרונו מרכזי של גישה זו הוא הטיפול גרוע בתמונות בהן ישנו ריכוז של צבעים מסוימים. בדומה לדיוון על קוונטייזיה סקלרית, עברו תמונות שרומות האפור בה (בשלושת השכבות) מרווחת בתת-אינטראול קטן, נקבע ייצוג גרוע של התמונה. היתרון הבולט של שיטה זו הוא פשוטה, והיותה בלתי תלולה בתמונה עלייה היא מופעלת.

על מנת לשפר את השיטה, מוצע לאמץ קוונטייזר שאינו יוניפורמי, ולעשות זאת תוך נתיחה של תוכן התמונה. גישה זו מقلילה את רعيון קוונטייזר מקס-לויד לטיפול בוקטוריים. אם נזכיר בדרך פעולה של קוונטייזר מקס-לויד, הרי שתהליך בנייתו כרוך בשני צעדים המבוצעים איטרטיבית. נזכיר מהם צעדים אלו ונראה כיצד להכלילם למימד גובה מ-1:

1. בהינתן N רמות הייצוג, יש לקבוע את רמות ההחלטה. במקרה הסקלרי עשינו זאת ע"י קביעת רמת ההחלטה כאמצע הדרך בין רמות הייצוג הסמכות, לפי הקשר

$$r_k = \frac{f_{k+1} + f_k}{2}$$

כתחליף לגישה זו, אם נתיחס לרמות ההחלטה כגבולות בין האינטראולים השונים, הרי שציר הממשיים מחולק לאזורים לפי שיטת השכן הקרוב, בה לכל נקודה על הציר נבדק מי רמת הייצוג הקרובה ביותר ואליה היא משוכנת.

בהתלהה לדו-מימד נאמר כי נתונים לנו N וקטורי ייצוג המסומנים ב- $\underline{f}_k \in \Re^S$. אנו לוקחים את מרחב הכניסה ה- S מיידי במקרה הכללי ($S=3$ לתמונות צבע), וכל נקודה בו משוכנת לוקטור הייצוג הקרוב ביותר.

אם במקרה הסקלרי בדקנו מרחק ע"י ריבוע ההפרש של הערכים, הרי שכן יש לנו חופש מוחלט בבחירה המטריקה. ניתן להשתמש במטריקה אויקלידית (סכום ריבועי הפרשי הקואורדינטות של הוקטוריים). לעיתים ניתן להציג מטריקה ייחודית המחשבת את המרחק בין הצבעים לפי תפיסתה של מערכת הראייה, וזה כמובן מצליח יותר.

2. בהינתן $N+1$ רמות ההחלטה, יש לקבוע את רמות הייצוג העדכניות. במקרה הסקלרי זה נעשה ע"י החישוב

$$f_k = \frac{\int_{r_k}^{r_{k+1}} r P(r) dr}{\int_{r_k}^{r_{k+1}} P(r) dr}$$

משמעות הביטוי – חישוב מרכז המאסה של כל קבוצה המשויכת לאחת מ- N רמות הייצוג. מרכז המאסה כאן בניו באינטגרל על פני כל ערכי הרצף ושימוש בפונקציית ההסתברות של רמות האפור. כאשר עוסקים בתמונה דוגמה $[m,n]$, נניח כי קבוצת הפיקסלים Ω משוויכת לרמת הייצוג f_k , ונניח כי כמות הפיקסלים בקבוצה זו סופית (הנחה סבירה למדי, כיון שסך כל הפיקסלים בתמונה היא כמות סופית). אז, תוך שימוש בעובדה שפונקציית ההסתברות היא סדרת הלמים בערכי רמות האפור הנתונות, ניתן להראות כי מרכז המאסה נתון ע"י:

$$f_k = \frac{\sum_{r \in \Omega_k} \sum_{[m,n] \in \Omega} \delta(r - f[m,n]) r dr}{\sum_{r \in \Omega_k} \sum_{[m,n] \in \Omega} \delta(r - f[m,n]) dr} = \frac{1}{|\Omega_k|} \sum_{[m,n] \in \Omega_k} f[m,n]$$

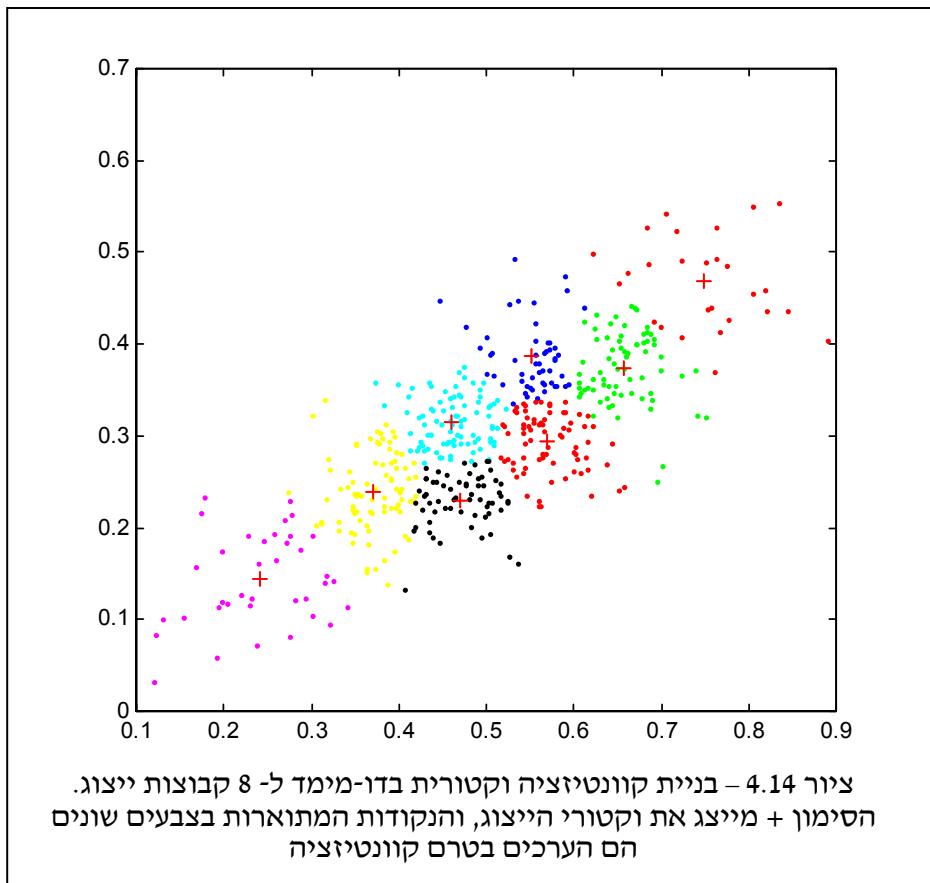
וכך יצא כי חישוב זה אינו אלה חישוב ממוצע הוקטוריים בקבוצה. בהכללה לוקטוריים, נרצה לחשב את מרכז המאסה ע"י חישוב ממוצע דומה של כל הצבעים השיכים בקבוצה.

למעשה, האלגוריתם שתואר מביא למינימום את השגיאה הבאה:

$$\varepsilon^2 \left(\{f_k\}_{k=1}^N \right) = \sum_{[m,n] \in \Omega} \min_{1 \leq k \leq N} d(f_k, f[m,n])$$

כאשר הפונקציה d היא המטריקה המ適用ת, וכפי שרואים, יש דמיון רב לנעשה במקרה הסקלרי.

דוגמה 4.3: נמחיש את האלגוריתם שתואר עבור מקרה דו-מימדי. נניח כי נתונות לנו נקודות בדו-מימד בפייזור כמתואר בצייר 4.14. רצוננו למצוא קוונטייזר ל- 8 וקטורי ייצוג דו-מימדיים. נפעיל את האלגוריתם המתואר. אתחול האלגוריתם נעשה ע"י בחירה אקראית של 8 נקודות במישור בקרבת מרכז המאסה (הממוצע) של הנקודות הנתונות. ביצוע של 10 איטרציות של האלגוריתם שתואר הניב את התוצאה המתוארת בצייר 4.14. לשם נוחות נקבע כל קבוצה בצבע שונה.



פרק 5 - שימושים בסיסיים לתמונות

פרק 5 – שימושים בסיסיים לתמונות	75
5.1	76
5.2	83
5.3	96
5.4	107

בפרק זה נתחיל את הדיוון בפעולות עיבוד על תמונות. העיבודים בהם עוסוק בפרק זה יתמקדו בניסיון לשפר את מראהה של התמונה הנתונה. חשוב להבהיר כבר outset כי בעיבוד תמונות קיימים עיבודים בעלי מטרת אחרת, כגון דחיסה, ניתוח תוכן התמונה ועוד, אך אלה יידחוו לפרקדים מאוחרים יותר.

מדוע לשפר תמונה? ובכן, במקרים רבים מתקבלת תמונה שאינה באיכות הרצויה, אם בשל דרך השגתה, או בשל קלוקלים שעברה בתהליכי עיבוד קודמים. בעיות אופייניות הן תמונה מרוחקת הדורשת חידוד, תמונה בה נדרש למתוח את הקונטраст, תמונה בה העצמים החשובים אינם נראים טוב ונדרש להבליטם, תמונה בה יש רעש נלווה ונחוץ לנוקוטו, ועוד. בכך כל שימוש צו ניתן להציג אלגוריתמים שונים ומגוונים. לכן, תחת הכותרת "שיפור תמונה" נכנס מגוון רחב מאוד של אלגוריתמים אשר אמורים לשפר את התמונה אך במובנים שונים, ולעתים קרובות אין כל קשר ביניהם.

מבחינתנו היה נפלא לו יכולנו להציג אלגוריתם כללי המקבל תמונה ומוסיאה גרסה משופרת שלה, ללא מעורבותו של המשתמש. המכשלה העיקרית לקיומו של תהליך כזה היא חוסר בתיאור מתמטי של מהי תמונה טובה. כיוון שאין מدد כמותי כזה, שיפור תמונות הוא עניין של מומחיות וטעם אישי – צריך לԶהות את הבעיה בתמונה ואז ליפור אלגוריתם ספציפי של פתרונה. לכן, התחום הקרי שיפור תמונות הוא תחום עם הרבה היורטיקות, ובו פעולות התבבססה על ניסוי וטעייה. במסגרת פרק זה ננסה להציג את היסודות וכלי העבודה הבסיסיים לשיפור תמונות.

פרק 5 – שימושים בסיסיים לתמונות

כנוקודה אחרונה במסגרת זו, חשוב להבהיר את ההבדל בין שיפור תמונה ושהזור תמונה. קיימים תחום נפרד הקרויה "שהזור תמונה", ובו לוקחים תמונה שעברה קלקל ידוע ומשחררים את המקור. אנו לדון באлогריטמים כאלה במסגרת הפרק הבא. במסגרת פרק זה איןנו מניחים כי הקלקל ידוע, ולכן דרך הפתרון משתנה באופן מהותי.

5.1 פעולות נוקודה והיסטוגרמה לתמונה

אחת הפעולות פשוטות ביותר שניתנו לבצע על תמונה היא פעולה חסרת זיכרון. כפי שכבר רأינו בפרק 2, פעולה זו בצורתה הכללית ביותר תהיה $\{f[m,n] = T[n,m]\}$, כאשר המוצא בנוקודה הדיסקרטית $[n,m]$ תלוי רק בערך הכניסה באותה נוקודה. מערכת זו ניתן לתאר ע"י פונקציה מעבר פשוטה הקובעת כיצד כל ערך כניסה עובר לערך יציאה, כמפורט בציור 5.1. אם נניח למשל כי תמונות הכניסה והמוצא מיוצגות ע"י 8 סיביות לפיקסל, הרי שלפנינו מערכת המעבירת כל מספר בתחום $[0,255]$ לערך מוצא בתחום זה, ולכן המערכת ניתנת לתיאור ע"י טבלת התאמה פשוטה הקרויה Look Up Table (LUT).



השתכנענו כי אכן מדובר במערכת פשוטה, והשאלה המותבקשת היא – מה ניתן להשיג באמצעות מערכת זו? מסתבר כי לעיתים תכופות יכולה מערכת LUT לספק שיפור ממשי לתמונה הכניסה. לפני שנמחיש את הפעולות הללו נציג את מושג ההיסטוגרמה לתמונה, אשר ישיע בהבנתן של פעולות אלו.

ההיסטוגרמה לתמונה היא פונקציה השכיחות של רמות האפור בתמונה, בדומה לפונקציות שתוארו במסגרת פרק קודם בדיון על קוונטייזר מקס-לויד. בהינתן תמונה בה כל פיקסל מיוצג ע"י 8 סיביות, כל פיקסל יכול לקבל ערך רמת אפור שלם בתחום $[0,2^8 - 1]$. הערך (x) הוא כמות הפיקסלים

בתמונה שערכם הוא x . לכן, הפונקציה $(x) h$ היא פונקציה חד-מידית המקבלת ערך חיובי כנגד כל x בתחום $[0,2^n]$. בדרכּ כל מתקבל לחלק את ערכי הפונקציה בכמות הפיקסלים הכלולות בתמונה, לקבלת פונקציה שסכום איבריה הוא 1. במקרה זה, הפונקציה $(x) h$ נראית כפונקציית הסתברות לכל דבר המקיים:

$$1. \forall x \in \mathbf{Z}, 0 \leq x < 2^n, h(x) \geq 0$$

$$2. \sum_{x=0}^{2^n-1} h(x) = 1$$

ההיסטוגרמה של תמונה מתארת תכונה מסוימת של התמונה – מידת אכלוס רמות האפור בתמונה. להיסטוגרמה אין מרכיב המתאר התנהגות מקומית, כלומר, מההיסטוגרמה לא ניתן לומר מהי רמת האפור במקומות מסוימים בתמונה, וברור כי מההיסטוגרמה לא ניתן לבנות מחדש את התמונה. בציור 5.2 מוצגות מספר תמונות וההיסטוגרמה שלהן. בכל המקרים מדובר בתמונות המיוצרות ב- 8 סיביות לפיקסל. כפי שמודגס בציור זה (בשורה האמצעית), שינוי סדר הפיקסלים בתמונה אינם משפיע כלל על ההיסטוגרמה. ע"י שינוי סדר הפיקסלים בתמונה מתקבלת תמונה שונה לחלווטין. למרות זאת, ברור כי ההיסטוגרמה נותרת ללא שינוי כלל.

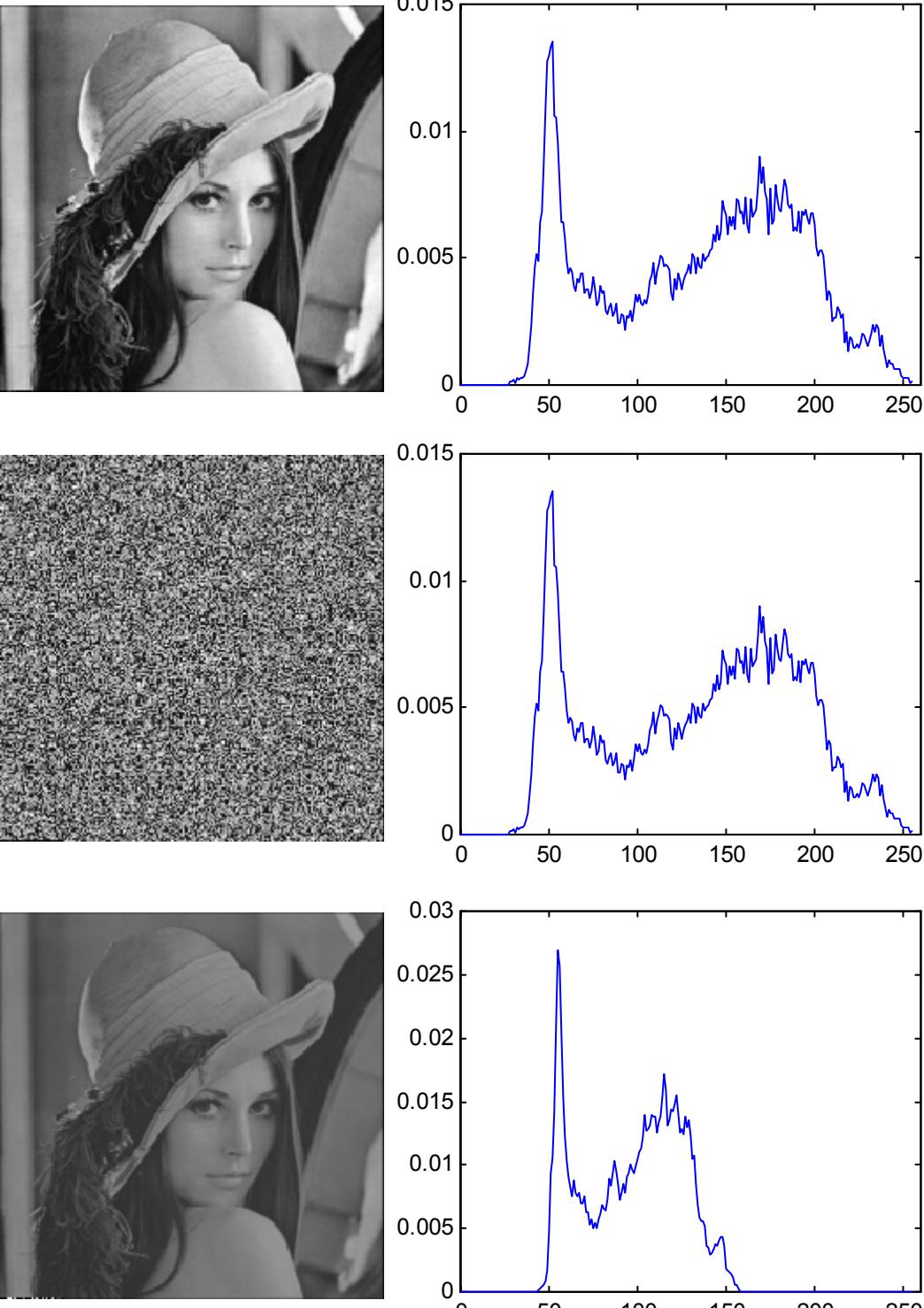
מדוע אנו מתעניינים בהיסטוגרמת תמונה? מסתבר כי עיון בהיסטוגרמת התמונה חשוב מידי בעיות של ניצול תחום דינامي. כך למשל, התמונה החשוכה מעט שמרתוארה בתחרתיות ציור 5.2 נראית כך כיון שערכי רמות האפור בה מאכלסים בקרוב את התחום $[50,150]$ – זהו תמונה שمبرזות תחום דינامي. ע"י פעולה LUT על תמונה זו מלהצהרה

$$\text{Out} = 2.55 \cdot [\text{In} - 50]$$

מתקבלת תמונה אשר מאכלסת את כל רמות האפור מאפס ועד 255. תוצאה זו ופעולה ה- LUT התואמת לה מתראים בציור 5.3. פעה זה תקרא מתייחת קוונטרסט (Contrast Stretching). נשים לב כי פעה זו ממפה לערכים שליליים את רמות האפור הקטנות מ- 50, ולערכים גבוהים מ- 255 את ערכי הכניסה מעל 150. אבל בתמונה הנתונה אף פיקסל כמעט אינו בתחום ערכים אלו, ולכן אין בעיה. על מנת להבטיח תוצאה נאותה, אנו מרתקים כל ערך שלילי לאפס, וכל ערך מעל 255 ל- 255.

במקרה הכללי, בהינתן תמונה בה רמת האפור המינימלית מסומן f_{\min} ורמת האפור המקסימלית f_{\max} , פעולה המתייחס לשם ניצול כל הרמות בתחום $[0,255]$ תהיה:

$$\text{Out} = 255 \cdot \frac{\text{In} - f_{\min}}{f_{\max} - f_{\min}}$$



ציור 5.2 – מספר דוגמאות לתמונה וההיסטוגרמה שלה. בשורה העליונה נתונה התמונה Lena וההיסטוגרמה המתאימה לה. בשורה האמצעית נלקחה התמונה Lena
ועורבלה (שינוי סדר אקריאי של הפיקסלים בה), וczfpo, ההיסטוגרמה שלה
זהה. בשורה התחתונה התמונה המקורית הוכפלה ב- 0.5 ונוסף לה
הערך 30, וכן, ההיסטוגרמה דומה במראה הכללי,
אך צרה יותר (ולכן גם גובהה יותר) ומוסטת.

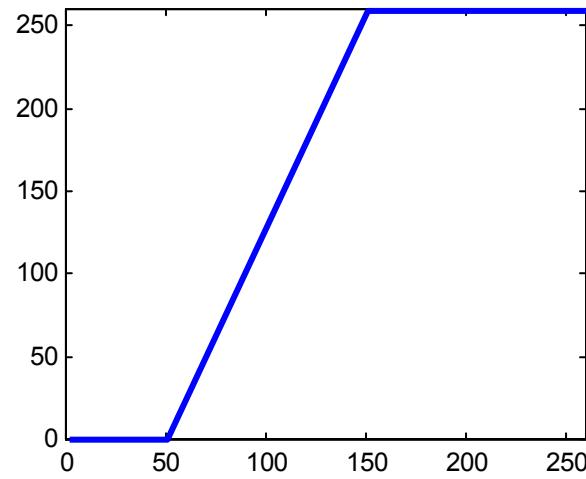
פרק 5 – שימושים בסיסיים לתמונות

לעתים שימוש בערכים הקיצוניים (f_{\min} ו- f_{\max}) אינה מוצלחת. לדוגמה, אם בתמונה כל רמות האפור מעט אחת מצויות בתחום $[0,100]$, וישנו פיקסל בודד בעל ערך 255, הרו שלפי המתוואר לעיל, מתייחס קונטרסט לא תנסה את התמונה, בשל אותו פיקסל חריג. כדי להתגבר על בעיה זו, ניתן להגדיר את f_{\min} ו- f_{\max} באופן שונה מעט. נזכיר כי פונקציית ההיסטוריה היא פונקציה צפיפות פילוג, ולכן ניתן להגדיר פונקציית הצלברות הפילוג ($H(x)$ עי):

$$H(x) = \sum_{k=0}^x h(k) \Rightarrow H(0) = 0; H(2^n - 1) = 1$$

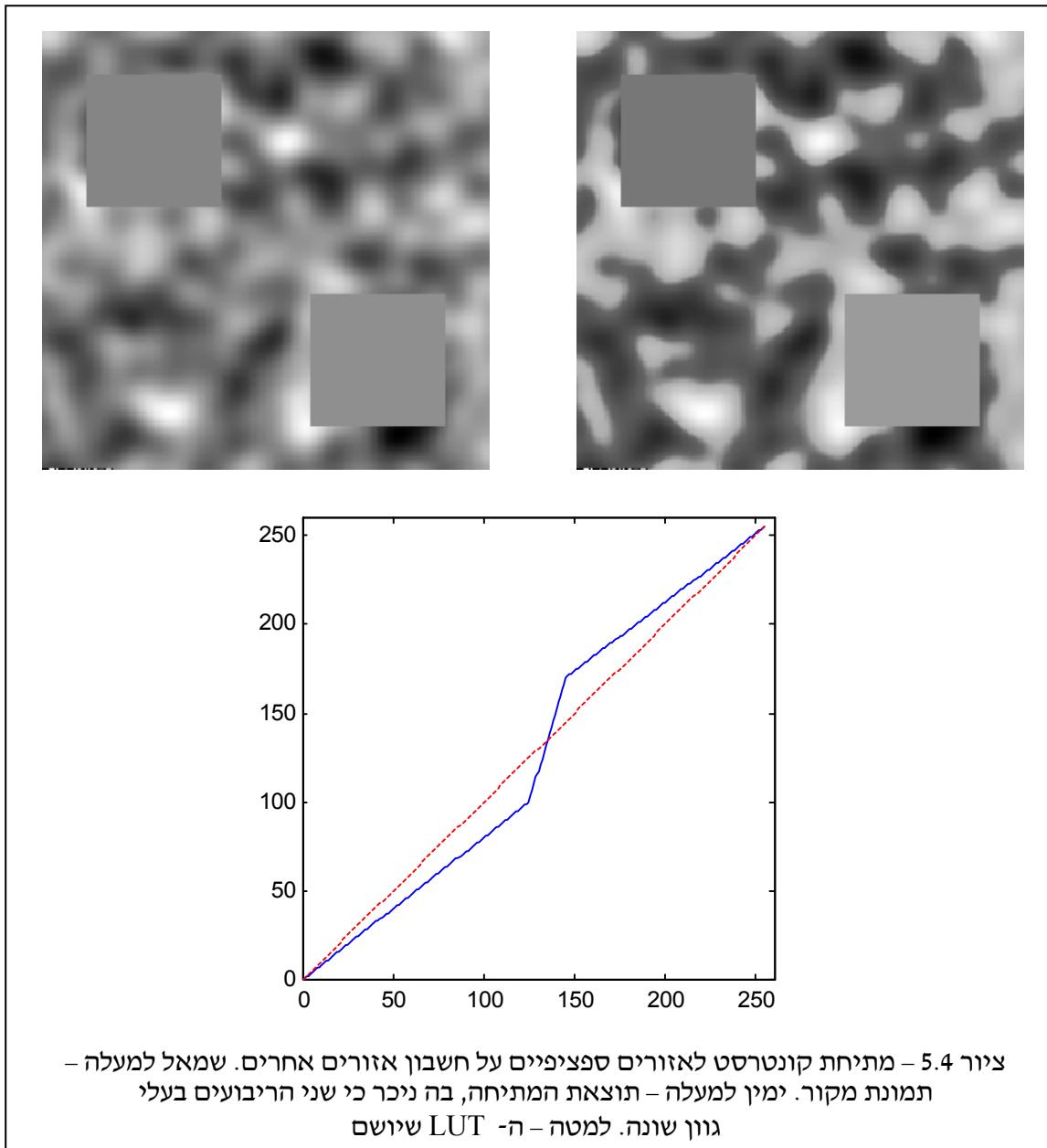
(ז – כמות הסיביות לפיקסל). אזי, אם קודם הגדרנו את f_{\min} להיות הנקודה בה $H(x) = 0$, נגידר אותה הפעם להיות הנקודה בה $H(x) = 1$. משמעות בחירה זו היא של-ע מהפיקסלים הכי נמכרים נתייחס חריגים ונבחר את הגדל שביניהם בערך מינימלי. באופן דומה, בחירת f_{\max} בנקודה בה $H(x) = 1$ תוחלף בנקודה בה $H(x) = 0$. מתייחס על-פי שני ערכים אלו תגרום לכמות יחסית של 2 מהפיקסלים בתמונה להיכנס לרוויה (מעל 255) ולQUITו (מתחת 0), אך זו קבוצה מצומצמת של נקודות, ולעומתה, עיקר התמונה יעבור שיפור מוצלח.

כדוגמה אחרת לפעולות מתייחס קונטרסט, בציור 5.4 מתוארת תמונה בה ישנים שני אובייקטים בעלי רמות אפור קרובות. נניח כי הם העיקר בתמונה ורצוננו להגדיל את יכולת האבחנה ביניהם. בציור 5.4 מוצع גם LUT להגשمت יעד זה. רמת האפור של המטריות הן 130 ו- 140. פעולה ה- LUT מרחיקה שתי רמות אלה לערכים 100 ו- 170 בהתאם, תוך פגיעה ביותר התמונה (אשר כאמור, נחשב לפחות חשובה במשימה שלפנינו). גם זהה פעולה מתייחס קונטרסט, אך רק בתחום מסוים של רמות האפור. ב圖 5.4 ניכר כי אמנים שני הריבועים הינם בעלי גוון שונה, כנדרש.



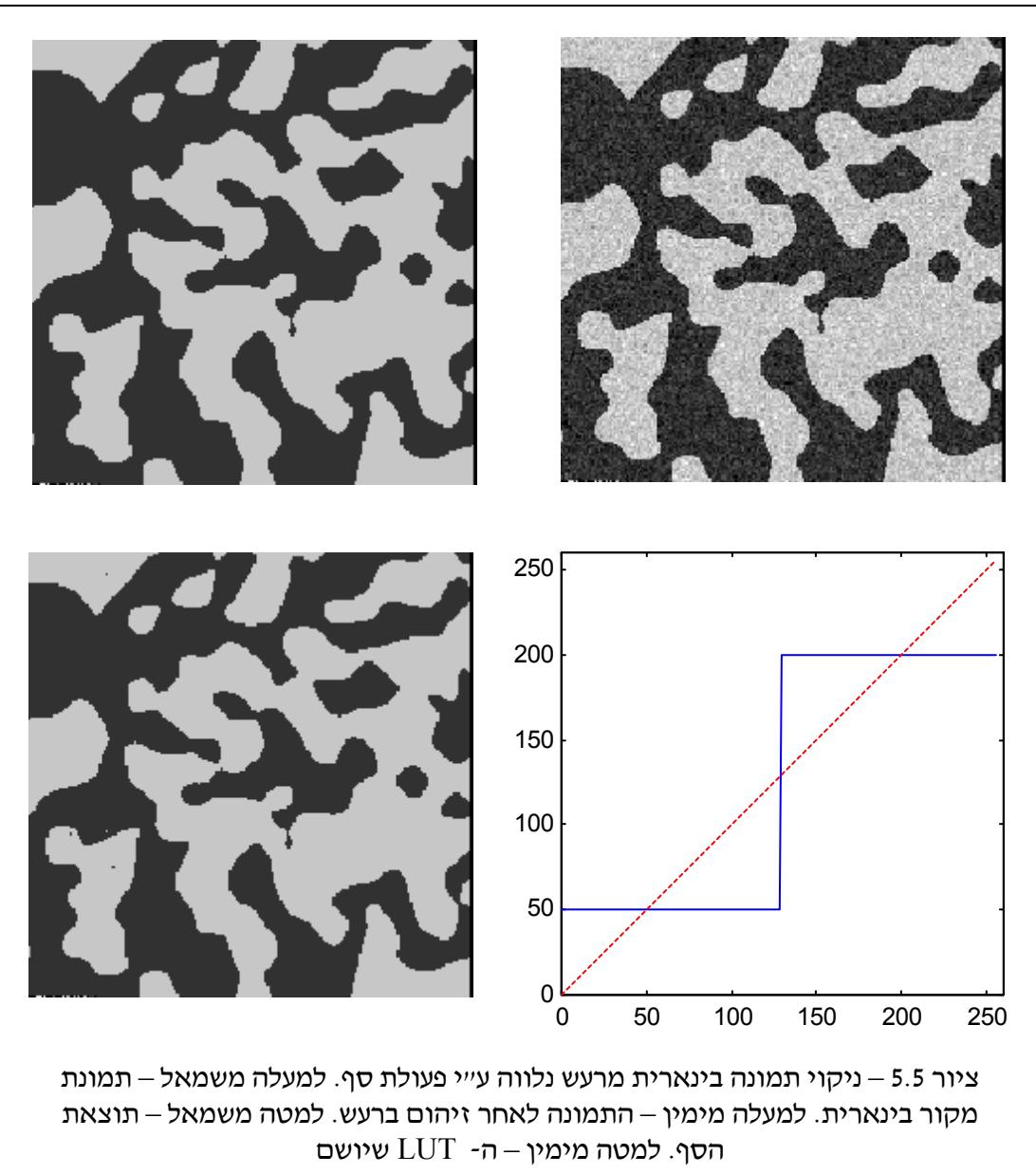
ציור 5.3 – שיפור התמונה בציור 5.2 (שורה תחתונה עי) פעולה מתייחס קונטרסט המתואמת כ- LUT מצד ימין

פעולה אחרת הניתנת לתיאור כ- LUT היא פועלות סף. נניח כי נתונה תמונה אשר ידוע כי במקורה היא תמונה בינהרית (נניח ערכי רמת אפור 50 או 200 בלבד!). לדוגמה זו צורף רוש בתהליך הרכשתה ורצוננו לנקוטה. ע"י פועלות סף נוכל להעלים חלק גדול מהרעש. ציור 5.5 ממחיש תהליך זה ותוצאותיו כפי שניתן לראותו אנו מצליחים לקבל תמונה כמעט זהה למקור לפני הרוש.



ציור 5.4 – מתיחת קוונטרסט לאזוריים ספציפיים על חשבון אזוריים אחרים. שמאל למעלה – תמונות מקור.ימין למעלה – תוצאה המתיחה, בה ניכר כי שני הריבועים בעלי גוון שונה. למטה – ה- LUT שיטם

הראינו כיצד אנו משנים קוונטרסט – כתע נבור להראות שינוי בהירות. בהירות היא בעצם רמת האפור הממצעת בתמונה. לכן, בחירה אינטואיטיבית לפועלות שינוי בהירות תהיה הוספת ערך קבוע לרמת האפור. בחירה כזו אינה מוצלחת בכך כל כיוון שהוא גורמת להוזמת הערכים מהתחום המותר. כתחליף, מקובל להשתמש בתיקון גמא (Gamma Correction), אשר נראה כך:



ציור 5.5 – ניקוי תמונה בינהרית מרעש נלווה ע"י פועלות סף. למטה משמאל – תמונה מקור בינהרית. למטה מימין – התמונה לאחר זיהום ברעש. למטה מימין – תוצאה הסף. למטה מימין – ה- LUT שיווש

$$\text{Out} = \text{In}^{1/\gamma}$$

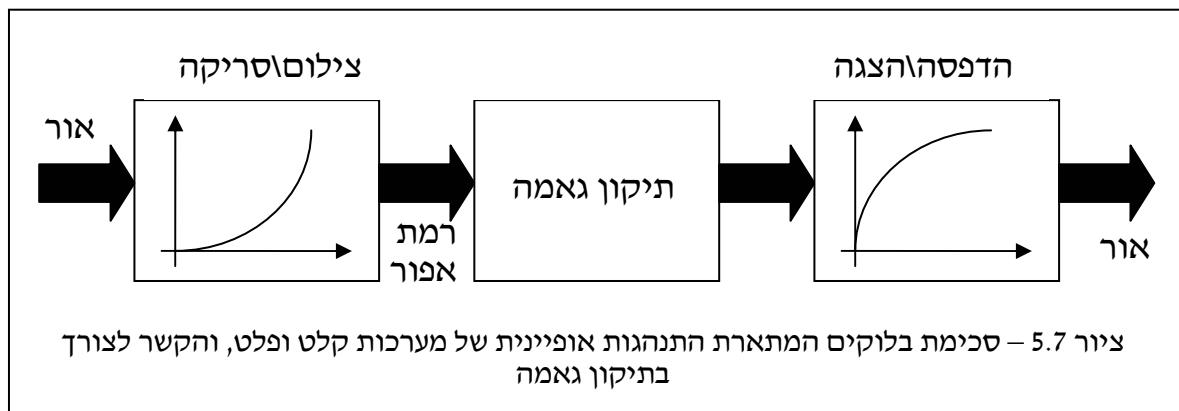
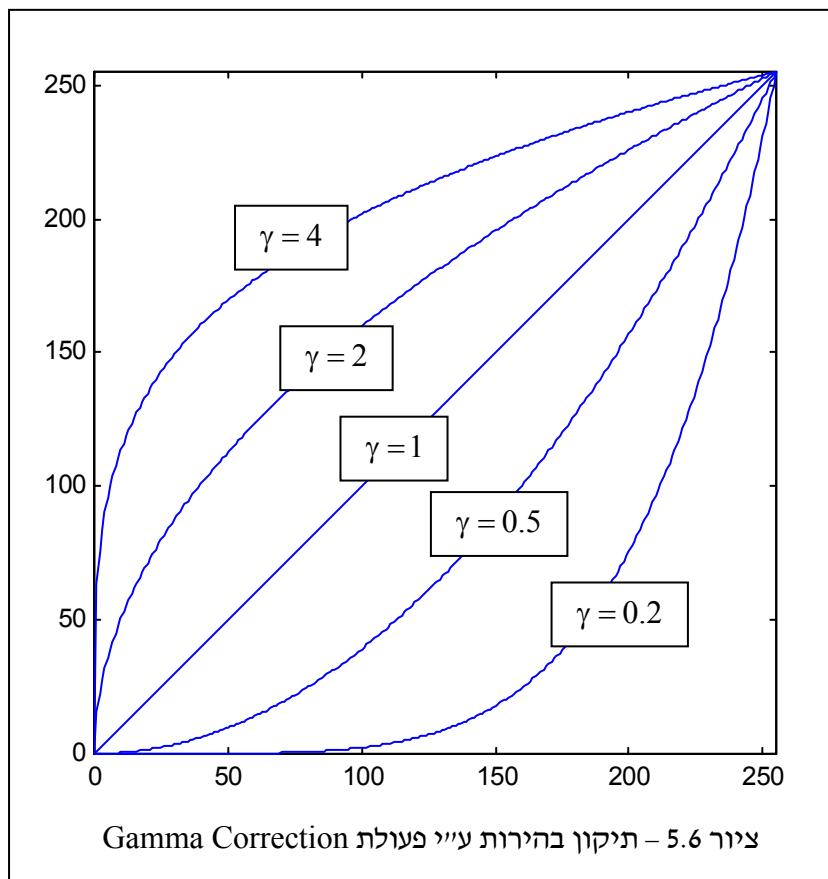


עבור ערכי גاما גדולים מ- 1 נקבל הבהרה של התמונה, ועבור ערכי גاما קטנים מ- 1 התמונה תוכחה.
ציור 5.6 מראה את צורת ה- LUT למגוון ערכי גاما.

לפועלות תיקון הגاما שורשים מוצקים בתחום החילוק וההדפסה. מסתבר כי בצלום תמונה (לסרט צילום או לגלאי CCD) רגישות החישה לאור אינה ליניארית וניתנת לתיאור כללי כפועלות חזקה על הכניסה, עם מעיריך חיובי וגדול מ- 1. לעומת זאת, תהליכי ההדפסה (או ההציגה למסך) מאופיינים אף הם בעקומה לא ליניארית של הצגת התמונה, כשבורם מתקבל להניח כי הפעולה היא פועלות חזקה עם מעיריך חיובי קטן מ- 1. ברור כי עבור ערכים توأمים (יחס של היפוך בין השניים) יתקבל ביטול הדדי ואין

פרק 5 – שימושים בסיסיים לתמונות

צורך בתיקון נוסף. לעומת זאת, כאשר התקני הקלט ופלט אינם מתואימים, נדרש פועלות תיקון ביניהם, וזאת כדי לבדוק פועלות הגاما. ציור 5.7 מראה סכימת בלוקים של המערכת אשר תוארה.



5.2 פעולות עיצוב היסטוגרמה

עד כה עסקנו בפעולות LUT פשוטות למדי, שנובעות מתוך שיקולים בסיסיים הניכרים מההיסטוגרמה. במסגרת סעיף זה נציג דרכי מורכבות יותר לתוכנו פעולה ה- LUT. העד הראשון ואנו נרצה להגשים הוא הבא – בהינתן תמונה כלשהי, נשאל מהי פעולה ה- LUT אשר תמיר את ההיסטוגרמהacha. מדו"ע איחוד? ובכן, אינטואיטיבית נראה כי תמונה שמנצלת כל רמתה לאפשר להיות ההיסטוגרמה אחת. אמור באופן אחד תהיה תמונה מאוזנת ונעימה לצפייה.

משימה שנייה בה עוסקת דומה בעירה – בהינתן תמונה כלשהי, נשאל מהי פעולה ה- LUT אשר תמיר את ההיסטוגרמהalla להיות ההיסטוגרמה אחרת נתונה? מדוע בעיה זו חשובה? ביישומים מסוימים כגון מציאת תנועה בין שתי תמונות, פסיפס של תמונות ועוד, משימה זו חשובה. הרעיון אם כך מقلיל את הבעיה הקודמת, כך שהפעם אנו צריכים לעצב את ההיסטוגרמה לפי צורה כלשהי.

האם פעולה LUT יכולה להגשים את היעדים הללו? מסתבר כי בכפוף לכמה דרישות סבירות, פעולות זו אפשריות. לשם בניית אלגוריתם למימוש משימה זו, נניח כי פונקציית ההיסטוגרמה הינה פונקציה רציפה מעלה רצף הערכים בתחום $[0,1]$. לכן, כל הניתוח שנטאר יתמקד בטיפול בפונקציות מעלה הרצף. לשם פתרון הבעיה שהוצעו, נזדקק לתוצאה הסתברות על הדרך בה משתמש חוק פילוג ע"י מערכת חסרת זיכרון.

נציג את הבעיה המתמטית באופן ממוקד: נתון משתנה אקראי x (זהה רמת האפור שלנו) עם פונקציית צפיפות הסתברות $(x) P_f$ (זהה כМОן ההיסטוגרמת תמונה הכנינה f). המשתנה האקראי x עובר דרך מערכת לא ליניארית חסרת זיכרון כלשהי T והפיכה (כלומר חד-חד ערכית, ולמעשה, מונוטונית עולה) ומתקבל המוצא $\{x\} = y$ (זהה פעולה ה- LUT). השאלה היא מהי פונקציית צפיפות הסתברות של המוצא האקראי y , אותה נסמן ב- $P_g(y)$? הפונקציה $(y) P_g$ היא הנגזרת של פונקציית הצטברות הסתברות, כמפורט במשוואה הבאה:

$$P_g(\alpha) = \frac{d}{d\alpha} \int_{-\infty}^{\alpha} P_g(\xi) d\xi$$

כיוון ש- T הפיכה, נוכל לומר כי $\{T^{-1}\{\alpha\}\} = \alpha$ וכך:

$$P_g(\alpha) = \frac{d}{d\alpha} \int_{-\infty}^{\alpha} P_g(\xi) d\xi = \frac{d}{d\alpha} \int_{-\infty}^{T^{-1}\{\alpha\}} P_g(\xi) d\xi = \frac{d}{d\alpha} \left[\int_{-\infty}^{T\{\beta\}} P_g(\xi) d\xi \right]_{\beta=T^{-1}\{\alpha\}}$$

המעבר האחרון הינו פשוט פעולה הצבה, ואין בו כל צעד אלגברי מהותי. אבל קיימים הקשר הבא:

פרק 5 – שימושים בסיסיים לתרומות

$$T \int_{-\infty}^{\beta} P_g(\xi) d\xi = \int_{-\infty}^{\beta} P_f(\xi) d\xi$$

כיוון שאנו יmino קובע מהו "מספרם" של הפיקסלים בעלי רמת אפור קטנה מ- β בתמונה הcnisha f , ואילו אף שמאל פירושו "מספרם" של הפיקסלים בעלי רמת אפור קטנה מ- β בתמונה המוצא g .

כיוון שהפעולה T מונוטונית עולה, הרי שני ערכיהם אלו זהים לחלוtin. הצבת קשר זה נותנת:

$$P_g(\alpha) = \frac{d}{d\alpha} \left[\int_{-\infty}^{\beta} P_f(\xi) d\xi \right]_{\beta=T^{-1}\{\alpha\}} = \frac{d\beta}{d\alpha} \frac{d}{d\beta} \left[\int_{-\infty}^{\beta} P_f(\xi) d\xi \right]_{\beta=T^{-1}\{\alpha\}}$$

כשעשינו שימוש בגזירה בשרשראת. הנגזרת על-פי β פשוטה כיוון שמתקיים:

$$P_f(\beta) = \frac{d}{d\beta} \int_{-\infty}^{\beta} P_f(\xi) d\xi$$

ולכן, הקשר הנ"ל הופך ל:

$$P_g(\alpha) = \frac{d\beta}{d\alpha} \frac{d}{d\beta} \left[\int_{-\infty}^{\beta} P_f(\xi) d\xi \right]_{\beta=T^{-1}\{\alpha\}} = \frac{dT^{-1}\{\alpha\}}{d\alpha} \cdot P_f(T^{-1}\{\alpha\})$$

גזירה של היפוכה של פונקציה נתונה ע"י הנוסחה הבאה:

$$\frac{d}{d\alpha} T^{-1}\{T\{\alpha\}\} = \frac{d\alpha}{d\alpha} = 1 \quad \stackrel{\uparrow}{=} \quad \frac{d}{d\beta} T^{-1}\{\beta\} \Big|_{\beta=T\{\alpha\}} \cdot \frac{dT\{\alpha\}}{d\alpha}$$

Chain Rule

$$\Rightarrow \frac{d}{d\beta} T^{-1}\{\beta\} = \left[\frac{1}{\frac{dT\{\alpha\}}{d\alpha}} \right]_{\alpha=T^{-1}\{\beta\}} \Rightarrow \frac{dT^{-1}\{\alpha\}}{d\alpha} = \frac{1}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}}$$

ולכן, התוצאה הסופית היא הביטוי הבא:

$$P_g(\alpha) = \frac{P_f(T^{-1}\{\alpha\})}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}}$$

דוגמה 5.1: נמחיש את הביטוי שהתקבל דרך מספר דוגמאות. עבור הפונקציה $T\{x\} = x + 1$, הנוסחה חניל' נוותנת

$$P_g(\alpha) = \frac{P_f(T^{-1}\{\alpha\})}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}} = \frac{P_f(\alpha - 1)}{1} = P_f(\alpha - 1)$$

וזו תוצאה צפואה ואינטואיטיבית, לפיו ההסתברות לקבל ערך כלשהו בmoזא זהה לחולtin לקבלת אותו ערך פחות 1 בכניסה. באופן דומה, עבור הקשר $T\{x\} = 2x + 1$ מתקבל:

$$P_g(\alpha) = \frac{P_f(T^{-1}\{\alpha\})}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}} = \frac{P_f(0.5[\alpha - 1])}{2} = \frac{1}{2} P_f\left(\frac{\alpha - 1}{2}\right)$$

עבור הקשר $T\{x\} = \exp(3x)$ מתקבל:

$$P_g(\alpha) = \frac{P_f(T^{-1}\{\alpha\})}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}} = \frac{P_f\left(\frac{1}{3}\log\{\alpha\}\right)}{\left[3\exp\{x\}\right]_{x=\frac{1}{3}\log\{\alpha\}}} = \frac{P_f\left(\frac{1}{3}\log\{\alpha\}\right)}{3 \cdot \exp\left\{\frac{1}{3}\log\{\alpha\}\right\}}$$

נזהר כעת ליעד הראשון אותו רצינו להגשים: רצוננו למצוא את T אשר ייתן כי היסטוגרמה המוצאת קבועה. עבור ערכי מוצא בתחום $[0,1]$ מתקבל כי נדרש $P_g(\alpha) = 1$. לכן:

$$1 = \frac{P_f(T^{-1}\{\alpha\})}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}} \Rightarrow P_f(T^{-1}\{\alpha\}) = \left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}$$

כיוון שני אגפי המשוואה הם פונקציות של $T^{-1}\{\alpha\}$, ומכוון שפונקציה זו חד-חד-ערכית, נוכל להשנית ולקבל את הקשר:

$$P_f(\alpha) = \frac{dT\{\alpha\}}{d\alpha} \Rightarrow T\{\alpha\} = \int_{-\infty}^{\alpha} P_f(x) dx$$

דוגמה 5.2: ננסה להמחיש כיצד פילוג לא אחד מומר לאחד באמצעות הקשר שהתקבל. נניח כי

$$P_f(x) = \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

נשים לב לכך שהפונקציה בעלת אינטגרל כולל השווה 1 כנדרש. לפי האמור לעיל, עלינו להפעיל LUT מהצורה:

$$T\{\alpha\} = \int_{-\infty}^{\alpha} P_f(x) dx = \begin{cases} 0 & \alpha < 0 \\ \alpha^2 & 0 \leq \alpha \leq 1 \\ 1 & 1 < \alpha \end{cases}$$

על מנת להשתכנע ש-LUT זה "ישף את הסchorה", נמchioז זאת דרך סימולציה Matlab. ניצור תמונה בעלת היסטוגרמה השווה בקירוב לו הנטונה כאן (אלכסונית), נסתכל על היסטוגרמת התמונה, ועל היסטוגרמת התמונה ביריעו. לשם ניטרול תופעת הקוונטייזציה בבדיקה זו, בחרנו 500 ערכים במרוחקים של 0.002 באינטראול [0,1]. תוצאות בדיקה זו מופיעות בציור 5.8, וכיום שניתן לראות, מתקבל שיוויי היסטוגרמה. הסיבה לחריגות מהתיאוריה היא הדיסקרטיזציה של התהיליך,علاיה עוד נדבר בהמשך.

דוגמה 5.3: בדוגמה הקודמת, נטפל הפעם במקרה המנוון בו הכניסה בעלת היסטוגרמה אחידה. לכן:

$$P_f(x) = \begin{cases} 1 & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases}$$

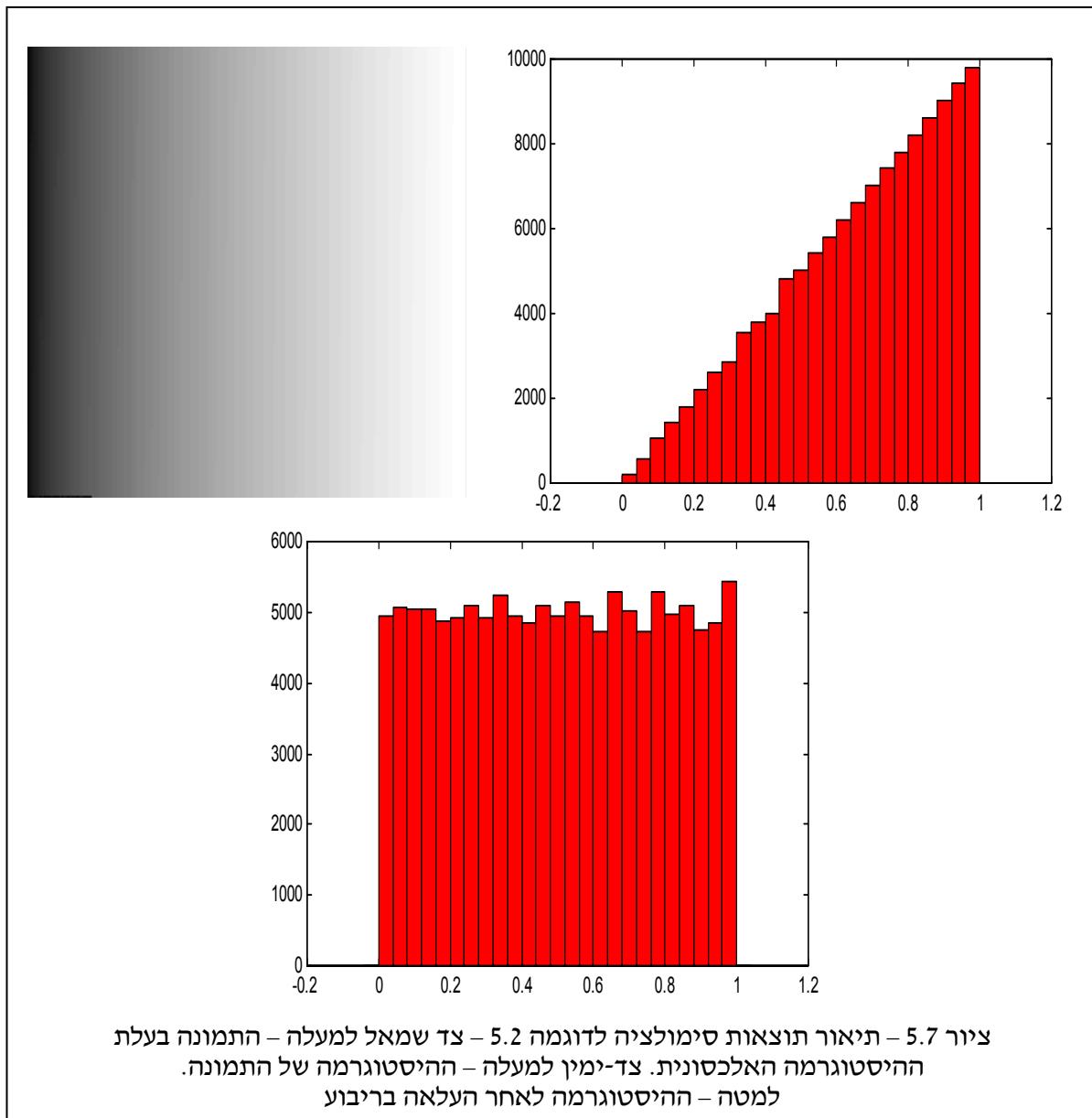
ה-LUT הנדרש הוא:

$$T\{\alpha\} = \int_{-\infty}^{\alpha} P_f(x) dx = \begin{cases} 0 & \alpha < 0 \\ \alpha & 0 \leq \alpha \leq 1 \\ 1 & 1 < \alpha \end{cases}$$

ולעניןינו, זהה בדיקת מערכת יחידה אשר אינה משנה את התמונה. תוצאה זו רצואה כיון שהМОצא מובטח אף הוא להיות בעל היסטוגרמה אחידה.

נזכור את אל התיאוריה, ונטפל בבעיית התאמת היסטוגרמות בין תמונות. הפעם, בהינתן תמונה בעלת היסטוגרמת כניסה $P_f(\alpha)$ רצוננו לעצב את היסטוגרמת המוצא לפי בחירה קבועה מראש $P_g(\alpha)$. שימוש בקשרים שהושגו קודם מטיב:

$$P_g(\alpha) = \frac{P_f(T^{-1}\{\alpha\})}{\left[\frac{dT\{x\}}{dx} \right]_{x=T^{-1}\{\alpha\}}} \Rightarrow P_g(T\{\alpha\}) = \frac{P_f(\alpha)}{\frac{dT\{\alpha\}}{d\alpha}} \Rightarrow \frac{dT\{\alpha\}}{d\alpha} P_g(T\{\alpha\}) = P_f(\alpha)$$



כששוב עשינו שימוש בעובדה שההתאמת T הפיכה. אינטגרציה על שני האגפים מיתן:

$$\int_{-\infty}^{\alpha} \frac{dT\{x\}}{dx} P_g(T\{x\}) dx = \int_{-\infty}^{\alpha} P_g(T\{x\}) dT\{x\} \stackrel{\substack{\uparrow \\ y=T\{x\}}}{=} \int_{-\infty}^{T\{\alpha\}} P_g(y) dy = \int_{-\infty}^{\alpha} P_f(x) dx$$

את הקשר הניל ראיינו כבר בעבר. נשתמש בסימונים הבאים לפונקציות הצבורות ההפוכות:

$$\int_{-\infty}^{\alpha} P_g(y) dy = F_g(\alpha) \quad \int_{-\infty}^{\alpha} P_f(x) dx = F_f(\alpha)$$

ונקבל :

$$\begin{aligned} \int_{-\infty}^{\alpha} P_g(y) dy &= F_g(T\{\alpha\}) = \int_{-\infty}^{\alpha} P_f(x) dx = F_f(\alpha) \\ \Rightarrow T\{\alpha\} &= F_g^{-1}(F_f(\alpha)) \end{aligned}$$

עבור המקרה המנוון בו ההיסטוגרמת המוצא איחודית, מתקבל כי $F_g(x) = x$, ולכן המיפוי הינו פשוט פונקציית הצלבות היחסטרות, כפי שקיבלנו קודם. נמחיש את התוצאה שהתקבלה במקרה מורכב מעט יותר.

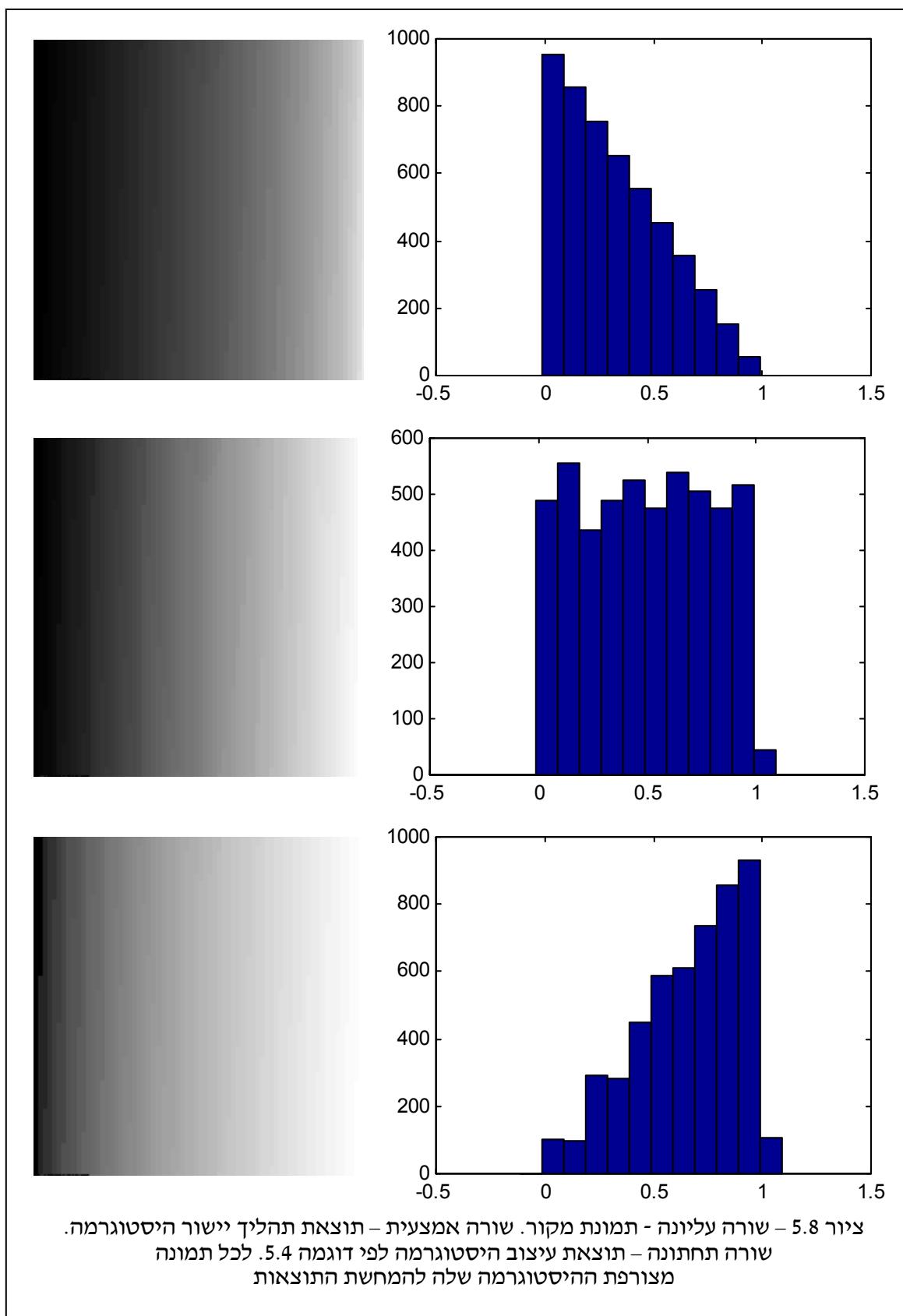
דוגמה 5.4: עבור שתי פונקציות היחסטרות הבאות :

$$\begin{aligned} P_g(x) &= \begin{cases} 2x & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} & \Rightarrow F_g(x) &= \begin{cases} x^2 & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} \\ P_f(x) &= \begin{cases} 2(1-x) & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} & \Rightarrow F_f(x) &= \begin{cases} 2x - x^2 & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} \end{aligned}$$

מתקובל :

$$\begin{aligned} F_g(x) &= \begin{cases} x^2 & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} & \Rightarrow F_g^{-1}(x) &= \begin{cases} \sqrt{x} & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} \\ \Rightarrow F_g^{-1}(F_f(x)) &= \begin{cases} \sqrt{2x - x^2} & 0 \leq x \leq 1 \\ 0 & \text{Otherwise} \end{cases} \end{aligned}$$

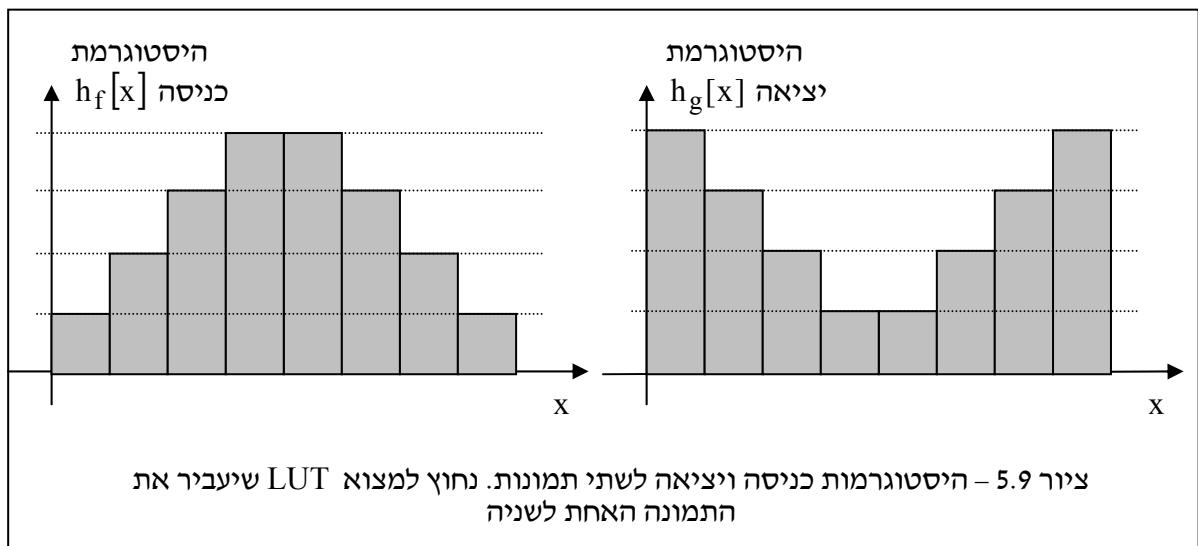
ציור 5.8 מראה את תמונה המקור בעלת ההיסטוגרמה האלכסונית, בה הכהמים דומיננטיים (שורה עליונה), תוצאת יישור ההיסטוגרמה ע"י ה- $Out_1 = T_1(In) = In \cdot (2 - In)$ LUT (שורה שנייה), ותוצאת ה- $Out_2 = T_2(In) = \sqrt{In \cdot (2 - In)}$ LUT לעיצוב ההיסטוגרמה כנדרש בדוגמה . ההיסטוגרמות נראהות כנדרש, למעט הפרעות הנובעות הדיסקרטיזציה של ערכי רמות האפור.



פרק 5 – שימושים בסיסיים לתמונות

נדון כעת בעיצוב היסטוגרמות לתמונות אחרי קוונטיזציה. כפי שראינו דרך הדוגמאות, דיסקרטיזציה (גם במובן של קיומם של ערכי רמות אפור דיסקרטיות, וגם במובן של כמה פיקסלים סופית) גורמת לביעות ולחיריגות. נמחיש את דרך העבודה, ואת הביעות המתעוררות דרך דוגמה, וכן נפתח אלטרנטיבת דיסקרטית לגישה האנליטית שפותחה.

דוגמה 5.5: נתונות שתי תמונות בנות 20 פיקסלים כל אחת, עם 8 רמות אפור אפשריות (כלומר – כל פיקסל מיוצג ב- 3 סיביות). ההיסטוגרמות לשתי תמונות אלה מtauור בציור 5.9. כל ריבוע מייצג פיקסל אחד, וניכר שישנם בדיקות 20 ריבועים כאלה.

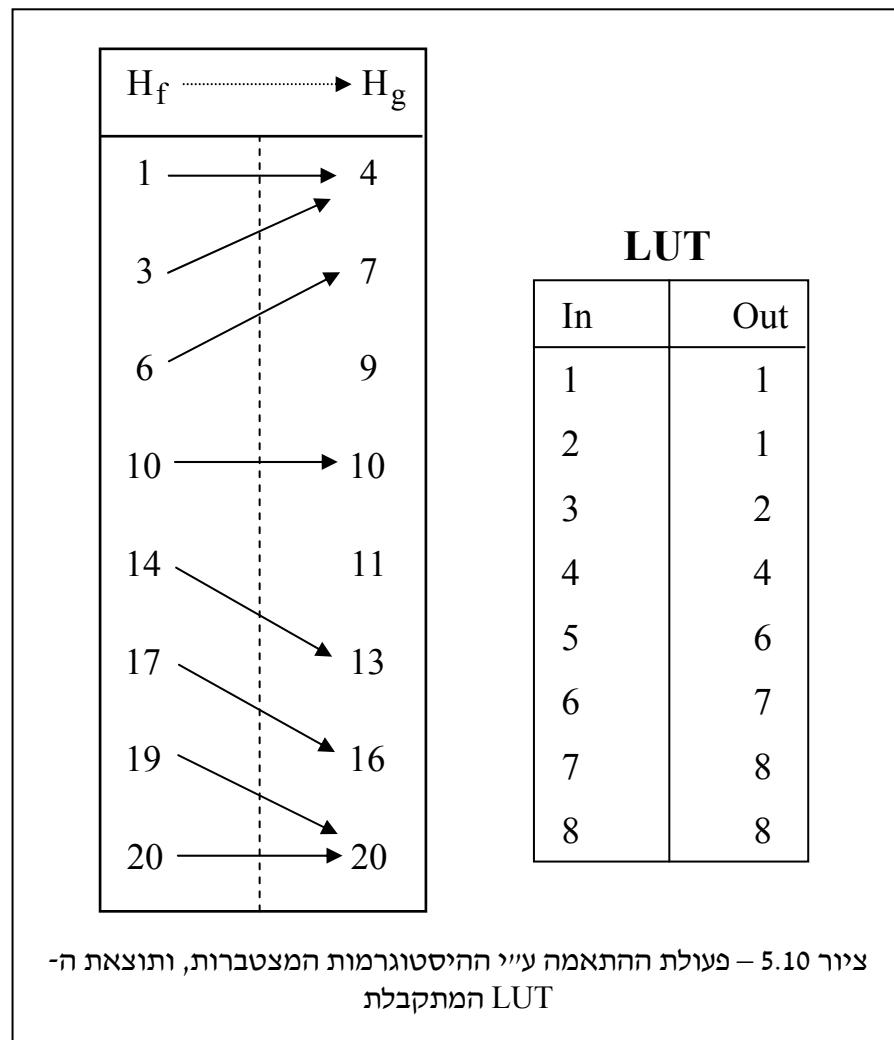


פועלות ה- LUT במקרה זה היא טבלת התאמה פשוטה בה יש 8 שורות – לכל ערך כניסה אפשרי (ויש 8 כללו), علينا להציג מה יהיה המוצא. בנית ה- LUT תעשה באופן סידרתי מהערך הנמוך לגבוה. נוח יותר ונכון יותר לעבוד עם ההיסטוגרמות המצטברות:

$$H_f[x] = \{1 \quad 3 \quad 6 \quad 10 \quad 14 \quad 17 \quad 19 \quad 20\}$$

$$H_g[x] = \{4 \quad 7 \quad 9 \quad 10 \quad 11 \quad 13 \quad 16 \quad 20\}$$

הרעין המרכזי הוא למצוא LUT בעל שגיאה מצטברת מינימלית. ערכי כניסה 1 ו- 2 (שני הנמכרים) ימופו לרמת מוצא 1. כך נקבל כי ההיסטוגרמת המוצאה המצטברת בעלת גובה 3 בעמודה ראשונה – סטיה של 1 מהדרוש. הוספת רמת כניסה 3 למוצא 1 תיתן גובה של 6 לעמודת המוצא הראשונה, ולכן לחיריגה של 2 – לכן עדיף למפות כניסה 2 למוצא 2. כך נמשך התהליך כאשרנו בודקים בכל שלב מהי ההחלטה שתיתן שגיאה מינימלית. למעשה, גישה פשוטה לבניית ה- LUT היא כתיבת שתי ההיסטוגרמות המצטברות זו מול זו, ומיפוי של כל ערך בהיסטוגרמות הכניסה (המצטברת) לערך הקרוב ביותר בהיסטוגרמת המוצאה המצטברת. ציור 5.10 מראה תהליך זה. בהינתן המייפוי, ה- LUT פשוט קשור את האינדקסים, כך למשל, החץ הראשון קשור את האיבר הראשון בכניסה לאיבר הראשון בכניסה, וכך ב- 1,LUT עברו ל- 1.

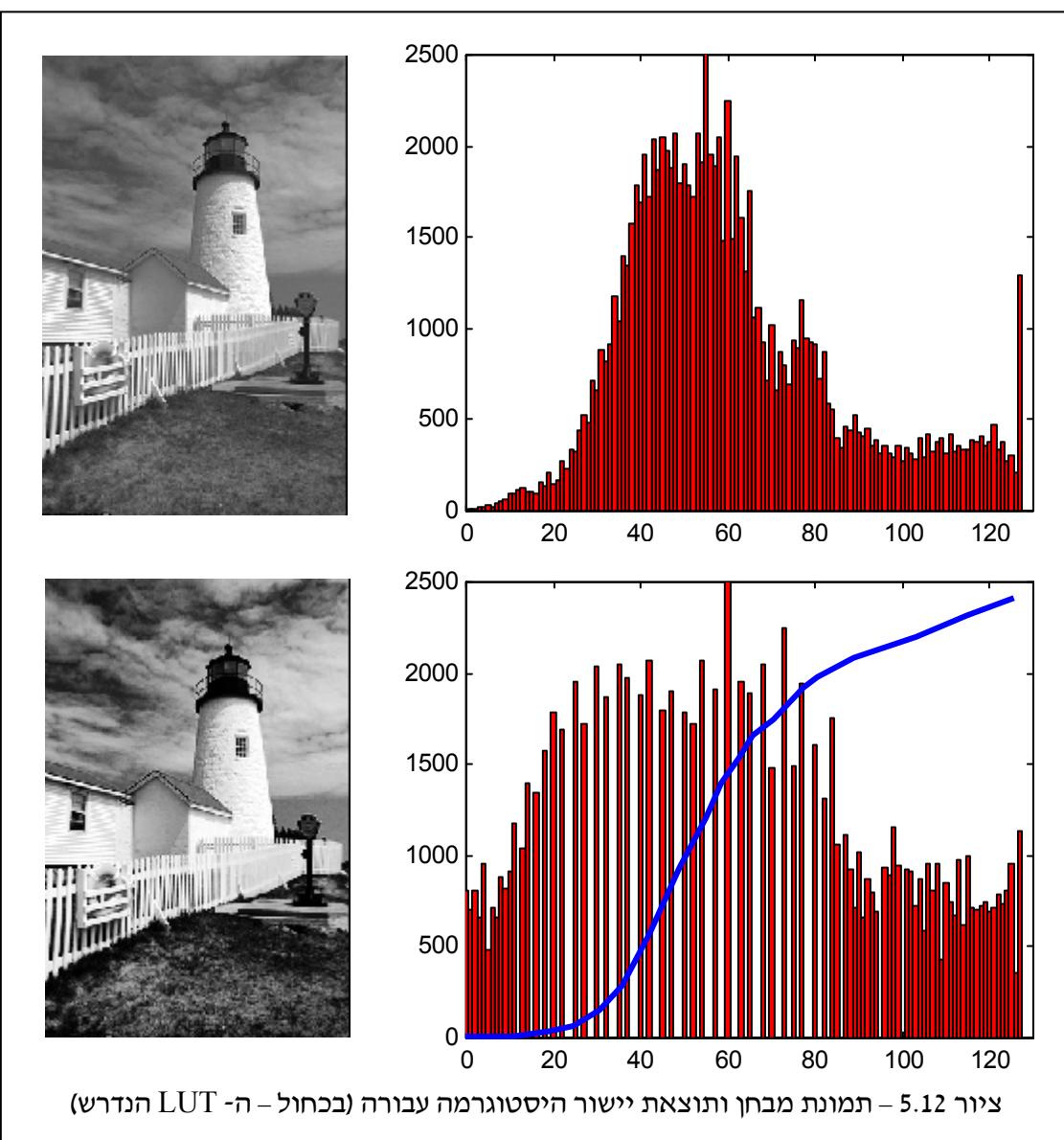
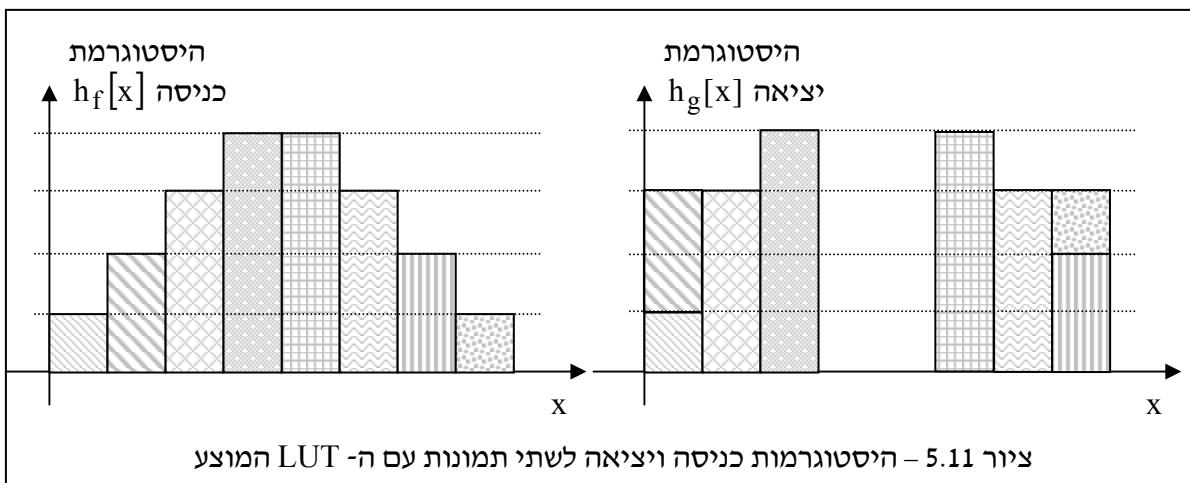


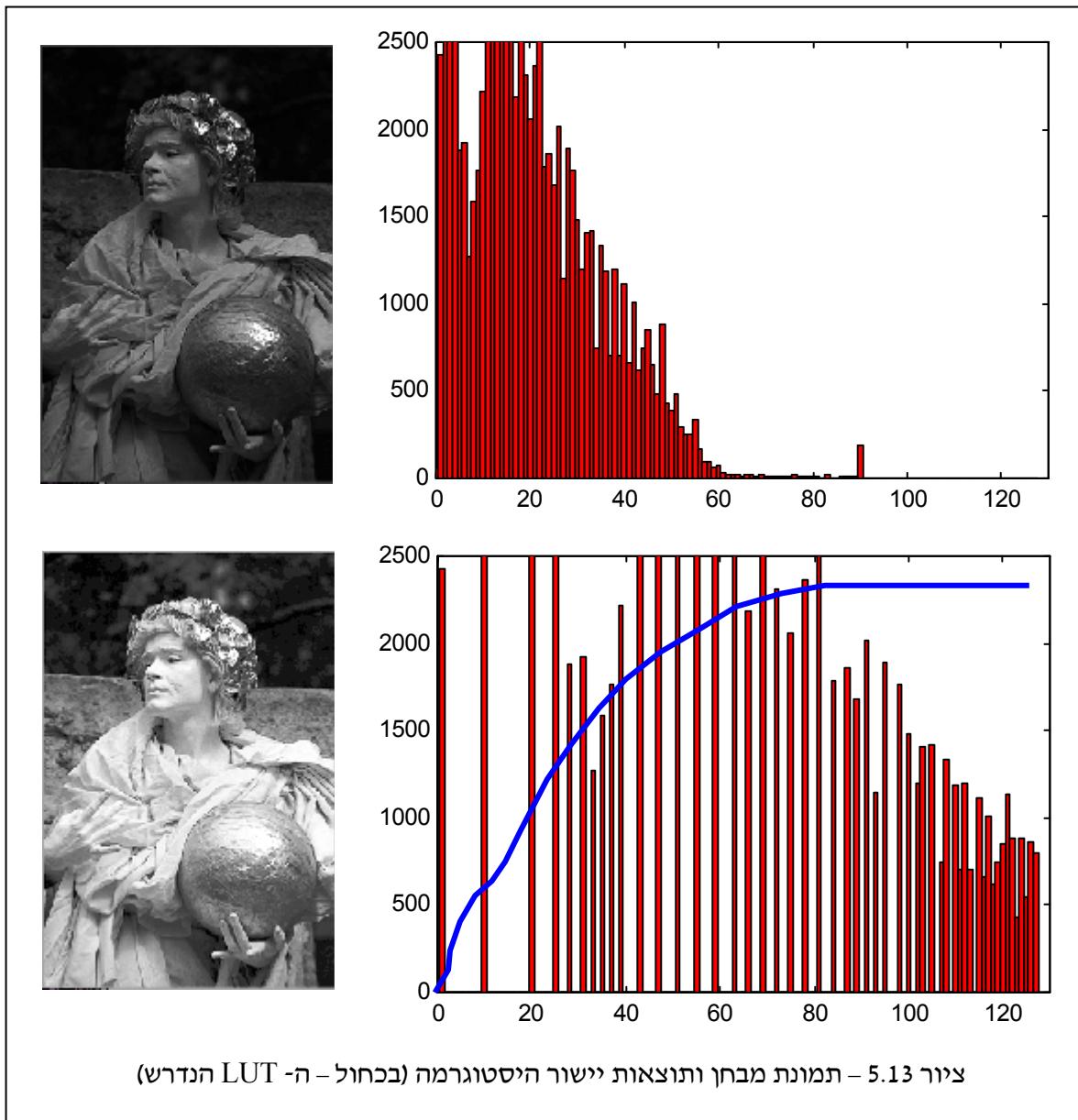
צייר 5.11 מראה את תוצאת ההיסטוגרמה ע"י ה-LUT המוצע. כפי שראויים, התוצאה אינה מושגית את ההיסטוגרמת היעד הנדרשת, בשל מגבלות הדיסקרטיזציה. באופן כללי, האלגוריתם שתואר לעתים קרובות נותן אפקטים של False Contouring (היווצרם של עקומים מלאכותיים) בתמונה. זה קורה כאשר שתי רמות אפור סמוכות יכולות כהתחליק להתרחק זו מזו ולתת שפה בתמונה.

קיימות מספר אלטרנטיבות לאלגוריתם הניל', הבאות לשפר את האלגוריתם הניל', ולמנוע אפקטים לא רצויים הנובעים מהיות התמונה דיסקרטית. פתרון אפשרי אחד הוא תוספת רעש לתמונה. התוספת רעש יכולה להיעשות לפני או אחרי הפעולה T, ובאופן דומה לדיוון בקונטיזציה, גם כאן מושג אפקט של ניטול העקומים המלאכותיים. אפשרות שנייה היא להכניס מרכיב הסתברותי בפעולת ה-LUT. למשל, אם נתיחס לדוגמה 5.5, לאחר מיפויים של 1 ו- 2 למצא 1, חסר פיקסל אחד בעמודת התוצאה (שגובהה 4 פיקסלים). רמת הכניסה 3, לעומת זאת, מואכליסת ב- 3 רמות אפור. לכן, אם בהסתברות 0.3333 נעביר את הכניסה 3 ל- 1, ובהתברות 0.6667 נעביר למצא 2, קיבל ההיסטוגרמת מוצא כנדרש (בממוצע). שם עוד פתרונות מגוונים, אך אנו נסתפק באלה.

פרק 5 – שימושים בסיסיים לתמונות

צעד אחרון בפרק זה, נראה מספר דוגמאות להמחשת תוצאות תהליכי יישור התאמה ההיסטוגרמתית. צירורים 5.12 ו- 5.13 מספקים תמונות מקור ותוצאת יישור היסטוגרמתית. הגרף הכחול מותאר את ה- LUT הנדרש (בכל המקרים, פיקסל מיוצג ע"י 7 סיביות).

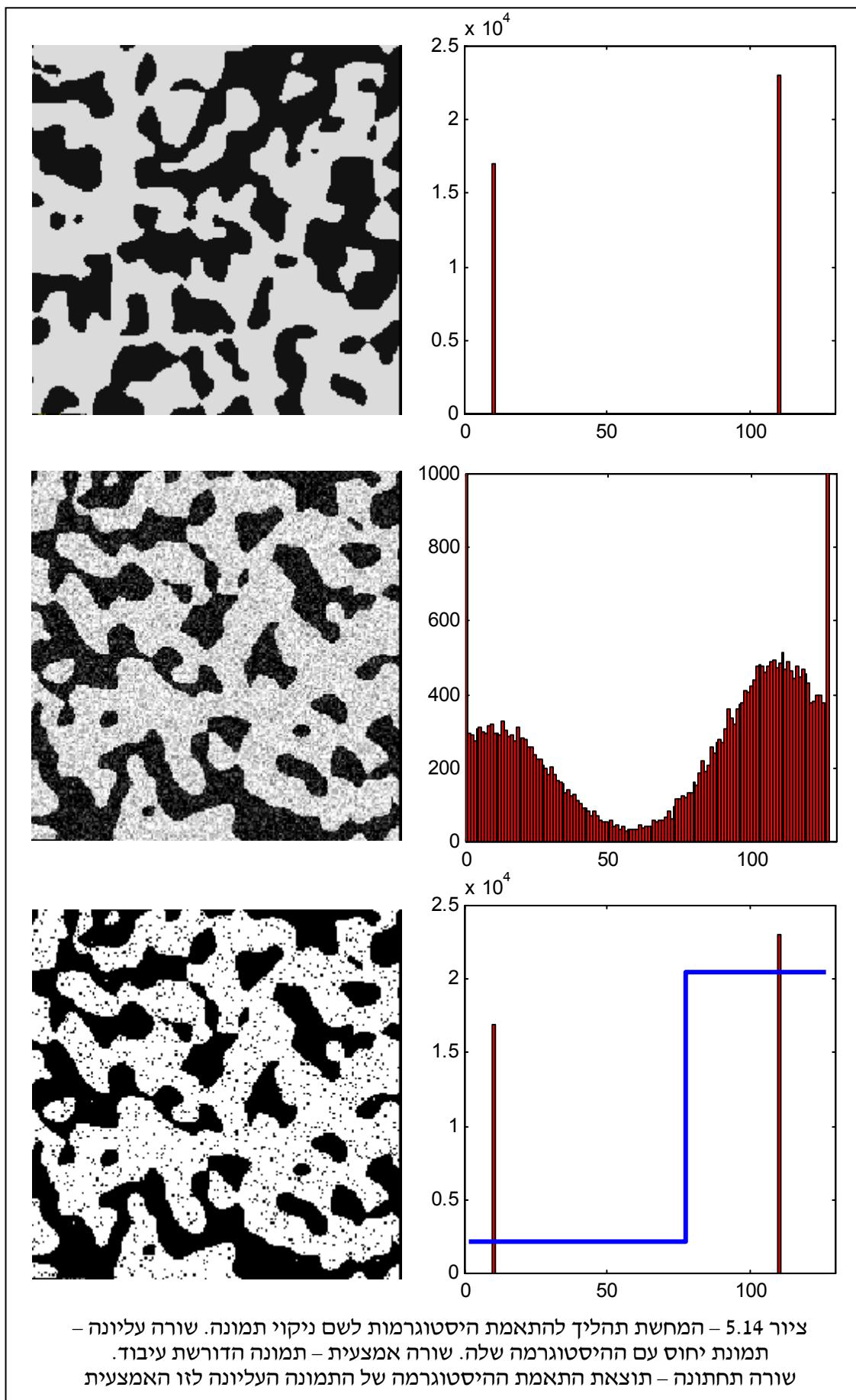




ציור 5.13 – תמונה מבחרן ותוצאות יישור היסטוגרמה (בכחול – ה- LUT המדרש)

צייר 5.14 מראה תהליכי של התאמת היסטוגרמות. נתונות שתי תמונות בעלות אופי דומה (שורות עליונה ואמצעית בציור), כשאחת מהן מזוהמת, ונדרש לנಕותה. לוקחים את התמונה המקולקלת ומבצעים עליה פעולה LUT לשם קבלת היסטוגרמה זהה זו של התמונה הנקייה. התוצאה מוצגת בשורה התחתונה, וראים כי אמם היסטוגרמה ב="#">התהליכי זהה להיסטוגרמת הייחוס. התמונה נקייה יותר, ובעצם, הפעולה שנמצאה היה השוואה לשפך כלשהו, כמוואר בגרף הכהול בציור.

לסיקום דיון זה חשוב לומר כי יישור התאמת היסטוגרמה אינו בהכרח משפר את התמונה, והוא לא אחיד שהתמונה אף ניזוקה. תמונות בהן לא צפוי פיזור אחד של רמות אפור הן תמונות בהן תהליכי יישור היסטוגרמה צפוי להזיק. זה נכון למשל לתמונות בהן צילום של גרפיקה טקסט הניתנות כמעט בינהיות, תמונות בהן גוון אחד דומיננטי מאוד, תמונות בהן מטרות קטנות יחסית אך חשובות (בשל גודלן, אין להן השפעה על היישור, אך הן מושפעות ממנו בפועל).



3. פעולות מרחביות בסיסיות

עד כה עסקנו בפעולות LUT אשר הין פעולות חסרות זיכרון מרחבי - ערך מוצא המערכת תלוי בערך הסקלארי של ערך הכניסה באותו מקום. יתרונה של פעולה LUT פשוטה, אך ברור כי מגוון התוצאות הנינטנות להשגה באמצעותה מוגבל למדי. פעולה כללית מעט יותר אשר עדין תהיה פשוטה יחסית היא פעולה מרחבית הפועלת על סביבת שכנים מצומצמת. למשל, מוצא המערכת הפעם יהיה תלוי בפיקסל הכניסה ומספר פיקסלים שכנים לו. פעולה זו יכולה להיות ליניארית או לא ליניארית, קבועה או משתנה במקומות, ופעולה LUT הופכת להיות מקרה פרטי של מערכת מעין זו. במסגרת סעיף זה נתמקד בפעולות קבועות במקומות (LINIARITY ו-AFFINITY).

באופן כללי, אם אנו דנים במערכת ליניארית על סביבה סמוכה לפיקסל, אנו בעצם דנים במסנן FIR דו-מימדי. ואננס, ברוב מוחלט של המקרים, סינון ליניארי של תמונה נעשה ע"י מסנן FIR. מסנן זה יאפשר ע"י תגובתו להלם אשר ניתנת לתיאור כסדרת המשקלות המכפילות כל פיקסל בסביבה, המתוארכות במטריצה דו-מימדית הקרויה גרעין (Kernel). נמחיש זאת דרך דוגמה.

דוגמה 5.6: אם פעולה הסינון על הפיקסל $[n,m]$ תהיה

$$\text{Out}[m,n] = \text{In}[m,n]\alpha_1 + \text{In}[m-1,n]\alpha_2 + \text{In}[m,n-1]\alpha_3 + \text{In}[m+1,n]\alpha_4 + \text{In}[m,n+1]\alpha_5$$

אזי הגרעין של מסנן ה- FIR נתון ע"י

$$K = \begin{bmatrix} 0 & \alpha_2 & 0 \\ \alpha_3 & \alpha_1 & \alpha_5 \\ 0 & \alpha_4 & 0 \end{bmatrix}$$

כאשר הראשית במישור האינדקסים מצויה למעלה מצד שמאל, והגרעין מופעל על התמונה ע"י הנחטו במסנן $[n,m]$ וחישוב סכום המכפלות בין המשקלות וערכי הפיקסלים של התמונה המטופלת. איבר האמצע בגרעין מתיחס למיקום $[n,m]$.

השימוש במסנן לשיפור תמונות רוח לשלווה יישומים עיקריים - ניקוי תמונה מרעש, חידוד תמונה, ומציאת שפות בתמונה. באשר לדיכוי רעש בתמונה: רעש יכול להציג לתוכה כתוצאה מiscalculon על תמונה שאמורה להיסרק, הפרעות שידור לתמונה מועברת בערוצ תקשורת, ועוד. בדרך כלל מתקבל להניח מודל אדיטיבי (חיבור) לפיו התמונה הרועשת $[n,m]$ היא סכום של התמונה המקורית $[n,m]$ ותמונה רעש $[n,d]$, כלומר $[n,m] = s[n,m] + r[n,m]$. ישנו גם מערכות צילום בהן הרעש הינו דזוקא מולטיפליקטיבי (מכפלתי), בו הרעש כופל את התמונה האידיאלית, אך אנו לא עוסקים בהן במסגרת זו.

פרק 5 - עיבודים בסיסיים לתמונות

בדרך כלל אות הרעש יונח להיות אקראי בעל ממוצע אפס, הומוגני ולבן. משמעות הנחה זו היא שערכי הפיקסלים בתמונה הרעש f מפולגים לפי אותו חוק פילוג בכל מקום בתמונה, לחוק פילוג זה ממוצע אפס ועוצמה (ווריאנס) קבועה, ואין קורלציה בין פיקסלים סמוכים. במלים אחרות - מתוך ידיעת ערכו של $[f]_{m,n}$ לא נוכל להזכיר דבר על ערci שכניו. בדרך כלל אף נהוג להניח כי הרעש הוא בעל חוק פילוג גאוסי.

הערה: המושג רעש מתאים יותר לאות שמע, וזהו באמת מקورو. הפרעה באות דיבור למשל תהווה "רעש". בתמונות, לפי הגיון זה, הפרעה הייתה צריכה להיקרא "לכלוך". בפועל, המושג רעש תפס גם בתמונות, והכוונה לאות אקראי לא רצוי.

הבעיה אם כך היא לנסות לבנות את התמונה f מתוך ידיעת התמונה s , ואת זאת לעשות ע"י סינון ליניארי. בנויגוד לרעש, ידיעת ערכו של הפיקסל $[f]_{m,n}$ מלמדת הרבה על ערci שכניו, כיוןSCPFI שכך ראיינו בפרק ראשון, ערci פיקסלים סמוכים הם לרוב קרובים (ואולי אפילו זהים). זהה בדיקת האבחנה שbezochotah אלו יכולים לנוקות רעש מהתמונה.

נתחיל את הניתוח מתוך הנחה שערci השכנים צריכים להיות זהים, ונניח שהמשנן פועל בגרעין בגודל של 3 על 3 פיקסלים. כאשרנו משתמשים בחלוון של 3 על 3 סביב הנקודה $[s]_{m,n}$, אנו מחפשים מהו "ערך האחד" $[f]_{m,n}$ שיחליף את $[s]_{m,n}$ ואשר יהיה קרוב לכל תשעת הערכאים בחלוון. למעשה, אנו מחפשים את המינימום לביטוי הבא:

$$\varepsilon\{f[m,n]\} = \sum_{k=-1}^1 \sum_{j=-1}^1 (s[m-k, n-j] - f[m, n])^2 \rightarrow \text{Minimum}$$

והתוצאה נתונה ע"י

$$\frac{\partial \varepsilon\{f[m,n]\}}{\partial f[m,n]} = -2 \sum_{k=-1}^1 \sum_{j=-1}^1 (s[m-k, n-j] - f[m, n]) = 0$$

$$\Rightarrow f[m, n] = \frac{1}{9} \sum_{k=-1}^1 \sum_{j=-1}^1 s[m-k, n-j]$$

כלומר, יש להפעיל על התמונה מסנן ליניארי בעל הגרעין

$$K = \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

פרק 5 - שימושים בסיסיים לתמונות

כשמשנן זה פועל על $[m,n]$ (אשר כזכור הוא $f[m,n] = f[m,n] + d[m,n]$), הוא נותן סכום של שני ממוצעים - ממוצע מקומי של התמונה f וממוצע מקומי של התמונה d . אם אמנס ב- f הערכים בסביבת 3 על 3 קרובים מאוד זה לזה, הרי שממוצע מקומי של f כמוותו כ- f עצמו. לעומת זאת, ממוצע של 9 משתנים אקריאים שווי פילוג הוא משתנה אקראי חדש בו השונות קטנה פי $\sqrt{9} = 3$, וכך יצא שעוצמת הרעש הוקטנה.

מדוע אם כך להסתפק בחלון של 3 על 3 - אם נעבד על חלון של 5 על 5 נקבל דיכוי עוצמת הרעש בפקטור 5. אבל - עם הגדלת החלון, הנחتانו על ערכי רמות אפור קבועים ב- f נשברת והולכת. וכך נקבל כי אמנס הרעש ייקטן יפה, אך עמו נקבל גם אפקט של טשטוש התמונה. ציור 5.15 מראה תוצאות הסינון עם חלונות בגודל שונה, וכי שרוואים, הרעש אמנס פוחת, אך עמו גם פירטי התמונה. האзорים בהם הכי גדול הנזק אם אזורי המעברים (שפויות בתמונה), בהם הנחת רמת אפור קבועה היא הנחה שגوية.



ציור 5.15 - ניקוי רעש ע"י מסנן ליניארי - שמאל למעלה - תמונה מקור. ימין למעלה - תמונה עם רעש גאוסי ולבן 15×15 לרמות אפור עד 255. שמאל למטה - ניקוי ע"י חלון 3 על 3. ימין למטה - כנ"ל עם חלון 5 על 5.

כדי לראות אם שיפרנו, נוכל לחשב את ההפרש בין תמונת המקור $[f,m]$ לכל אחת מההתמונות הנ"ל (עם רעש, סינון ע"י 3 על 3 וסינון ע"י 5 על 5), ולסכם את ריבועי הפרשים אלו. רצוי גם לחלק במספר

הפיקסלים בתמונה ולהוציא שורש מההתוצאה, על מנת לקבל מספר ביחידות של רמות אפור. הנוסחה אם :

$$\text{Error} = \sqrt{\frac{1}{MN} \sum_{m=1}^M \sum_{n=1}^N (f[m,n] - \hat{f}[m,n])^2}$$

כאשר $\hat{f}[m,n]$ היא התוצאה שלנו המנסה לקרב את $f[m,n]$. קритריון זה קרי שגיאה ריבועית ממוצעת (Mean Squared Error - MSE), והוא רוח בעיבוד תמונות כմדד לאיכות תהליכי עיבוד. ברור כי שימוש בו אפשרי רק כאשר זמינה לנו תמונה המקורי (טרם הטרופות הרעש). ברור כי ככל שערכו של סכום זה קטן יותר השיפור טוב יותר. התוצאה לקרה הניל נוותנת כי השגיאה המוחסנת לתמונה הרועשת היא 15.01 - זהה כמעט בדיקת השונות בה השתמשנו. לתמונה לאחר סינון 3 על 3 התקבלה שגיאה של 9.8, ולסינון 5 על 5 שגיאה של 13.38. בשני המקרים חל שיפור, אך סינון 3 על 3 טוב יותר. נשים לב כי אם התמונה הייתה קבועה למורי הינו צפויים לקבל שגיאות של 5 ו- 3 בקירוב טוב (לפי יחסית דיכוי רעש בלבד 3 ו- 5). בשל העובדה התמונה לא קבועה, השגיאה גדלה יותר, וזו הולכת ומחזיפה עם הגדלת החלון.

שאלה מעניינת היא השאלה הבאה: עברנו מחלון של 1×1 (זהו המקרא בו בעצם לא הופעל סינון), לחלון בגודל 3×3 , וממנו לחלון בגודל 5×5 , כשהכל המקרים קיבלו מסן אחד (ערכים זהים בגרעין). האם אנו יכולים להציג דרך אחרת גישה מגוון רחב יותר של פיטרים? התשובה לכך נועצה בביטוי שהוא הבנוו למינימום. אם נכניס לשגיאות הריבועיות משקלות לפי החשיבות לכל פיקסל

$$e\{f[m,n]\} = \sum_{k=-L}^L \sum_{j=-L}^L c[k,j] (s[m-k, n-j] - f[m,n])^2 \rightarrow \text{Minimum}$$

נקבל:

$$\frac{\partial e\{f[m,n]\}}{\partial f[m,n]} = -2 \sum_{k=-L}^L \sum_{j=-L}^L c[k,j] \cdot (s[m-k, n-j] - f[m,n]) = 0$$

$$\Rightarrow f[m,n] = \frac{\sum_{k=-L}^L \sum_{j=-L}^L c[k,j] \cdot s[m-k, n-j]}{\sum_{k=-L}^L \sum_{j=-L}^L c[k,j]}$$

והפעם הגרעין יהיה:

$$K = \frac{1}{\sum_{k=-L}^L \sum_{j=-L}^L c[k,j]} \begin{bmatrix} \ddots & & \vdots & & \ddots \\ & c[-1,-1] & c[-1,0] & c[-1,1] & \\ \cdots & c[0,-1] & c[0,0] & c[0,1] & \cdots \\ & c[1,-1] & c[1,0] & c[1,1] & \\ \ddots & & \vdots & & \ddots \end{bmatrix}$$

פרק 5 - עיבודים בסיסיים לתמונות

כיצד נקבע את המשקלות? גישה אפשרית אחת היא כערכיהם הדועכנים לפי המרחק האוקלידי מהמרכז. קצב הדעיכה יקבע בפועל את התמך האפקטיבי של המסלן שייווצר (רווחו). למשל, לחלון 3 על 3 ניתן לקבוע את המשקלות כ:

$$c[-1,-1]=c[-1,1]=c[1,-1]=c[1,1]=0.25, c[0,1]=c[1,0]=c[-1,0]=c[0,-1]=0.5, c[0,0]=1$$

והגרעין יהיה:

$$K = \frac{1}{4} \begin{bmatrix} 0.25 & 0.5 & 0.25 \\ 0.5 & 1 & 0.5 \\ 0.25 & 0.5 & 0.25 \end{bmatrix}$$

את התהליך הנ"ל יכולים לתרגם מתוך שיקולים ספקטוריים, כלומר תכוננו המסלן כمسلן LPF דו-מימדי במישור התדר. נראה את עיקרי הדברים כאן: לומר שהתמונה $[m,n]$ קבועה בקרוב במקומות שונים לאשרה שתמונה זו בעלת התמורה תדר הדועכת מהר לאפס כשמטרה קרים מהראשית (כלומר יש ספקטרום עם תדרים נוכחים בלבד). לעומת זאת, הרעש הוא בעל ספקטרום אחיד בכל התדרים (מכאן שימושה בא המושג "רעש לבן" - לבן הוא צבע העושה שימוש בכל אורך הגל באופן אחיד). לכן, סינוון הרעש תוך כדי אי-פגיעה בתמונות המקור מחייב מסלן LPF שאינו מזיק כמעט בכלל לתמונה $[m,n]$. לכן, המסלן לא יכול להיות LPF צר כרצונו, אלא תלוי בהtransformה התדר של התמונה $[m,n]$.

בקriterיוון שהוצע לעיל לבנית המסלן העלינו כל שגיאה בריבוע. ניתן כתחליף להגדיר את השגיאה ע"י ערך מוחלט כ:

$$\epsilon\{f[m,n]\} = \sum_{k=-L}^L \sum_{j=-L}^L |s[m-k, n-j] - f[m, n]| \rightarrow \text{Minimum}$$

השינוי הוא במהותו שינוי המטריקה המודדת את המרחק בין שתי תמונות, והמטריקה החדשה אינה פחות נכונה מקודמתה. הבאת ביטוי זה למינימום מניבה את המסלן הידוע בשם סינוון חציון (Median).

$$\frac{\partial \epsilon\{f[m,n]\}}{\partial f[m,n]} = -2 \sum_{k=-L}^L \sum_{j=-L}^L \text{sign}(s[m-k, n-j] - f[m, n]) = 0$$

נניח כי אוסף ה- $(2L+1)^2$ הפיקסלים בחלון סודרו בסדר עולה לפי ערכיהם (ללא כל חשיבות למקום). אז, הערך שיביא את הביטוי הנ"ל למינימום הוא הערך האמצעי - $(2L^2 + 2L + 1)$. למשל, עבור חלון 3 על 3 ($L=1$), האיבר האמצעי הוא החמישי מותוך 9. בחלון 5 על 5 יהיה זה האיבר ה- 13. מדוע האיבר האמצעי דואק? כיון שאז מחצית מהאיברים בסכימה הנ"ל בעלי סימן שלילי, מחצית בעלי סימן חיובי, וアイבר האמצע נותן ערך אפס. סכימת הסימנים תיתן התאפסות הנזרת כנדרש.

הMSN שהתקבל - סינון חציון - אינו ליניארי. MSN זה שיך למשפחה של MSNs סדר (Order Filters), בהם מוצא MSN תלוי בסידור (Sorting) של ערכי הכנסה. ניתן גם להציג בחירה של המקסימלי/מינימלי בחלוון, וישנם יישומים העשויים שימוש בMSNs מעין אלו.

אז איזה סינון יותר טוב - ליניארי או חציון? התשובה תלולה בטיבו של הרוש האדיטיבי (מהו חוק הפילוג שלו), מהי הגישה לבחון את התוצאות (האם MSE או מדד אחר) ועוד. בכלל, מקובל כי ברעשי شيئاם סינון חציון יותר טוב. רעש شيئاם, הקרויגס רעש מלח ופלפל Salt & pepper (נשמע יותר אמין ממש מה), הוא רעש בו פיקסלים במקומות אקראי משנים את ערכם לשחור/לבן מוחלט באופן אקראי. בדוגמה הבאה ננסה להמחיש מדו"ע סינון מדין טוב כל כך בטיפול ברעשי شيئاם.

דוגמה 5.7: נניח כי אנו משווים סינון ליניארי אחד וסינון חציון בחלוון של 3 על 3. בחלוון מסוימים התקבלו הערכים הבאים - {50 60 65 70 75 84} (תשעה ערכים ומיקומם לא חשוב). ערך הממוצע יהיה ≈ 64 והחציון יהיה 65 - אין הבדל גדול. נניח כעת כי הערך 75 השתנה כתוצאה מרוש וערך גדול יותר. הממוצע החדש יהיה 84, בעוד שהחציון לא השתנה כלל! כל שינוי באחד המספרים יגרום לשינוי מיידי (אך מוקטן) בתוצאה הממוצעת. לעומת זאת, בחישוב החציון שינוי של אחד מארבעת הערכים הנמוכים מטה או שינוי של אחד מארבעת הגבוהים מעלה לא תנסה דבר. במקרה שני, נניח כי חוזרים לערכים המקוריים והערך 60 עבר להיות 255. במקרה זה, החציון יהיה עדין 65, ואילו הממוצע האלגברי יהיה 85 - שוב שינוי ממשמעותי.

בצירור 5.16 מתוארת תמונה אשר אליה צורף רעש Salt & pepper בה 20% מהפיקסלים "נגועים" ברעש شيئاם לשני הכוונים (שחור ולבן בהסתברות שווה). כפי שניתן לראות, MSN חציון מצליח להתגבר על מרבית הרוש תוך שמירת מרבית הפרטים בתמונה, בעוד שMSN הליניארי מוביל לתוצאה ירודה למדי באיכותה.

חשוב לציין כי כאשר מניחים ידע כלשהו על סטטיסטיקת הרוש האדיטיבי (כגון האמירה - רוש לבן גאוסי וכו'), דיכוי הרוש יכול להיעשות מתוך שיקולים תיאורטיים מוצקים יותר, בהשוואה לגישה שתוארכה כאן. אנו נראה זאת במסגרת פרק הבא, בו נדבר על שחזור תמונות.

משמעותו שנייה של MSNs (LINARIES ו-LINARIES) מגשיים היא פעולה חידוד תמונה. המידע העיקרי בתמונה הם שפנות העצמים המופיעים בה, ולכן הדגשת שפנות היא משימה חשובה. בהסתכלות בתחום התדר, שפנות הם מידע בעל תדרים גבוהים (בגלל המעברים החדים), בעוד שפניהם העצמים מיוחס לתחומים נמוכים. לכן, חידוד שפנות פירושו MSN אשר יעבד על התדרים הגבוהים וידגיש אותם. זהה בדיקת פועלה הפוכה לפעולות ניקוי הרוש ע"י LPF כפי שתוארה קודם. לכן, חידוד תמונה בא בחשבון רק עבור תמונות נקיות מרוש. העיקרי/algoritm חידוד תמונה הוא להעביר את התמונה דרך HPF (MSN מעביר גבהים) ולהחבר את מוצא MSN זה לאחר משקלן נאות לתמונה המקורית. השאלה המ提בשת היא - כיצד ממשיכים HPF ? ובכן, HPL הוא עצם 1-LPF, ולמעשה LPF הוא כבר יודעים. לדוגמה, עבור ה- LPF פשוט בו מוצע גרעין של 3 על 3 אחד, ה- HPL יהיה



ציור 5.16 - סינון רעש עי"י מסנן ליניארי ומסנן חציון. עמודה שמאלית - מקור (למעלה) ותמונה רועשת. עמודה אמצעית - חציון 3 על 3 (למעלה) וחציון 5 על 5 (למטה).
עמודה ימנית - סינון ליניארי 3 על 3 (למעלה) וסינון ליניארי 5 על 5 (למטה)

$$K_{HPF} = 1 - K_{LPF} = 1 - \frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

מסנן HPF יהיה תמיד בעל סכום איברים השווה 0 (למה?). לכן, המסנן הכלול שיבצע את פעולה החידוד במקרה זה יהיה:

$$K_{Sharp} = 1 + K_{HPF} = 1 + \alpha \frac{1}{9} \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} = \frac{1}{9} \begin{bmatrix} -\alpha & -\alpha & -\alpha \\ -\alpha & 9+8\alpha & -\alpha \\ -\alpha & -\alpha & -\alpha \end{bmatrix}$$

עבור $\alpha = 0$ קיבל את תמונה המקור בדיק, וכש- α חיובי וגדל, קיבל אפקט חידוד חזק יותר ויותר ומןוקודה מסויימת ואילך, תתקבל תוצאה כה חזקה שתיראה לא טבעיות ולא יפה. ציור 5.17 מראה תוצאה של חידוד תמונה למגוון ערכי α עם המסנן הנ"ל. כתמונה מקור בחרנו לחתות את Lena לאחר טשטוש כל.



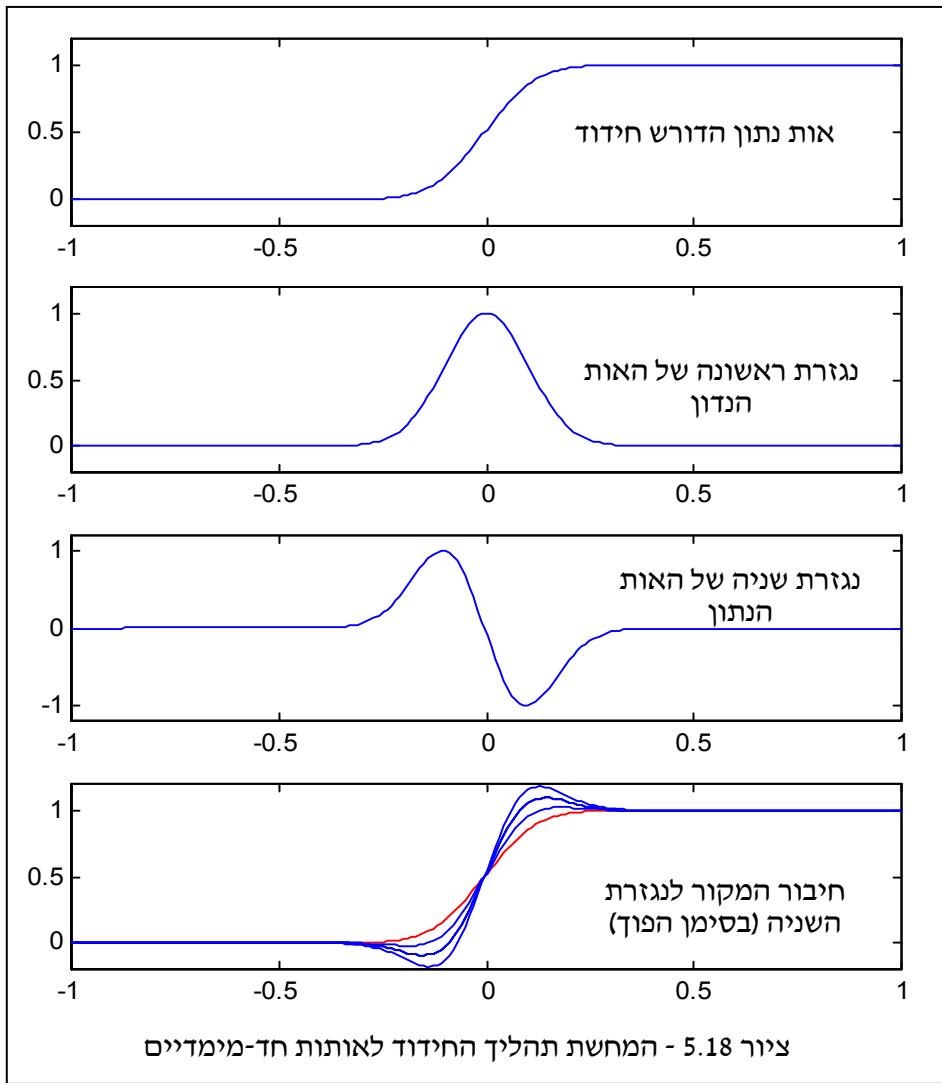
ציור 5.17 - תמונה מקור (שمال למעלה) ותוצאות חידוד שפות למגוון ערכי α - 0.1 (ימין למעלה), 0.3 (שمال למטה), 1 (ימין למטה).

מסנן שרואה בו השימוש במסנן HPF לטובת חידוד שפות הוא אופרטור הלפלסיאן הנטוון עיי'

$$K_{\text{Laplacian}} = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}$$

כפי שנראה בהמשך, לאופרטור זה קשור הדוק לאופרטור הנגזרת השנייה הרציף.

מה קורה בתהיליך החידוד? צייר 5.18 ממחיש את התהליך עבור אותן חידודים. אנו רואים שפה מטושטשת, ורואים כי נגזרת שנייה המוחכרת מאות המקור נתנת אפקט של מעבר חד יותר, כנדרש. מדוע סימן שלילי לנגזרת? מדובר נגזרת שנייה? לעת עתה שאלות אלו פתוחות כיון שלא ענינו על שאלה בסיסית יותר - מה הקשר בין נגזרות (ראשונה ושניה) במסנן ה- HPF שרצינו מלכתחילה? בכל אופן, אינטואיטיבית נוכל לומר את הדבר הבא - הפעלת מסנן כדוגמת הלפלסיאן נתנת ערך קרוב לאפס באזוריים בהם רמת האפור מהוות מישור מוטה. במקומות אלו לא יורגש כלל החידוד. במקומות בהם תוצאת הלפלסיאן גבוהה משמעותית מאפס, פירוש הדבר הוא שהפיקסל הנוכחי גבוה יותר מממוצע ארבעת שכניו. ככלומר - ישנו שם פרט. על מנת לחזדו, יש להגדיל עוד את ההפרש בין הפיקסל לבין שכניו, וזהי בדיקת הפעולה המתבצעת. תהליך דומה קורה לערכים שליליים.



נעבור כעת לדון במציאת שפות בתמונה, וקירוב נגזרות מעלה שריג דיסקרטי דו-מימדי. כזכור מפרק 1, אנו יכולים להתייחס לתמונה כמשטח (פונקציה) מעלה המשיר, כשרמת האפור ($f(x,y)$) היא גובהו של המשטח בנקודה (x,y) . לפונקציה זו אנו יכולים לחשב בנקודה (x,y) את הגרדיינט בכל נקודה. הגרדיינט הוא וקטור הבנוי מהנגזרת האופקית והנגזרת האנכית, אשר בrzצף מוגדר לפי:

$$\nabla f(x,y) = \frac{\partial f(x,y)}{\partial x} \hat{x} + \frac{\partial f(x,y)}{\partial y} \hat{y}$$

לקטור הגרדיינט חשיבות רבה. חישוב ערכו המוחלט יכול לומר לנו האם הנקודה בה אנו עומדים היא שפה או לא, שהרי נורמה גדולה לגרדיינט פירושה שיפוע חזק. נורמה זו נתונה ע"י:

$$\|\nabla f(x,y)\| = \sqrt{\left(\frac{\partial f(x,y)}{\partial x}\right)^2 + \left(\frac{\partial f(x,y)}{\partial y}\right)^2}$$

$$\Delta f(x, y) = \nabla^2 f(x, y) = \frac{\partial^2 f(x, y)}{\partial x^2} + \frac{\partial^2 f(x, y)}{\partial y^2}$$

כך שams נחזור למה שהציגנו קודם, חידוד הפונקציה, לו היה נעשה ברכף, היה מתבצע ע"י

$$g(x, y) = f(x, y) - \alpha \Delta f(x, y)$$

כל האמור לעיל טוב ויפה אך אינו מועיל בובנו לטפל בתמונות דיסקרטיות. לכן, علينا למצוא דרך לקרב נגורות באות דגום. הרעיון הבסיסי פשוט עד טרייאלי. כיוון שנגורת ברכף מייצגת הפרשים בין ערכים סמוכים למרחק אינפיטיסימלי וחולקה למרחק זה, הצעה מתבקשת היא לקבוע את הנגורת האופקית, למשל, כהפרש בין דגימות סמוכות באותה שורה, תוך הנחה שהמרחב ביןיהם הוא 1. הבעיה היא שכשאנו עומדים על הדגימה $[m, n]$, ישנן שתי אפשרויות :

$$\frac{\partial f}{\partial x} \rightarrow f[m, n] - f[m-1, n] \text{ or } f[m+1, n] - f[m, n]$$

האפשרות הראשונה קרויה גזירה אחורית, והשנייה גזירה קדמית. באופן דומה לגזירה אנכית :

$$\frac{\partial f}{\partial y} \rightarrow f[m, n] - f[m, n-1] \text{ or } f[m, n+1] - f[m, n]$$

כתחליף להגדרות הנ"ל ניתן להציג גזירה מרכזית לפי :

$$\begin{aligned} \frac{\partial f}{\partial x} &\rightarrow \frac{1}{2}[f[m+1, n] - f[m-1, n]] \\ \frac{\partial f}{\partial y} &\rightarrow \frac{1}{2}[f[m, n+1] - f[m, n-1]] \end{aligned}$$

בהמשך ישיר לקו הגיון זה, גזירה שנייה היא הפרש על הנגורת עצמה. לכן, גזירה שנייה אופקית תהיה :

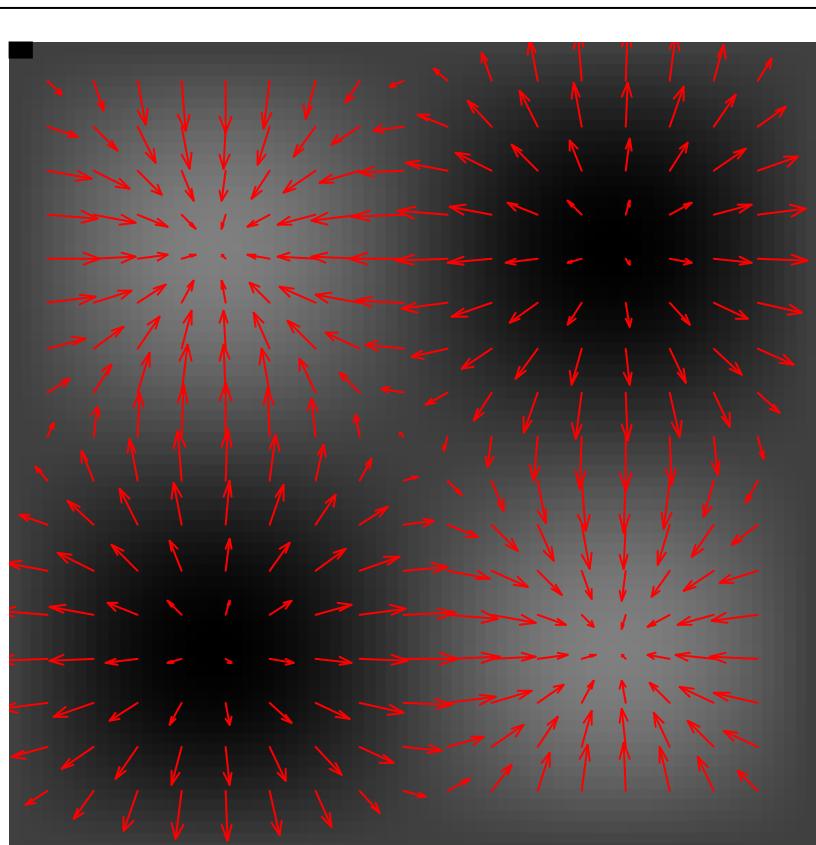
$$\frac{\partial^2 f}{\partial x^2} \rightarrow (f[m+1, n] - f[m, n]) - (f[m, n] - f[m-1, n]) = f[m+1, n] + f[m-1, n] - 2f[m, n]$$

כשהפעם עשינו שימוש בגזירה הקדמית והאחורית במשולב על מנת לקבל נגורת שנייה מאוונת (כלומר צו שמדובר רצוף במקום הנקו). לפי האמור לעיל, אופרטור הלפלסיאן הדיסקרט מתקובל ע"י :

$$\begin{aligned}\Delta f[m,n] &= (f[m+1,n] - f[m,n]) - (f[m,n] - f[m-1,n]) + \\ &+ (f[m,n+1] - f[m,n]) - (f[m,n] - f[m,n-1]) = \\ &= f[m+1,n] + f[m-1,n] + f[m,n+1] + f[m,n-1] - 4f[m,n]\end{aligned}$$

וכבר רأינו מסן זה קודם לנו, כשדנו בחידוד.

בצ'ור 5.19 מוצגת תמונה ועליה חיצי הגרדיאנט, אשר חושבו ע"י נזורת מרכזית. צפוי, חיצי הגרדיאנט מצביעים לכיוון העלייה המרבית במשתח התמונה.



צ'ור 5.19 - תמונה ועליה מפת וקטורי הגרדייאט שהושבו נומריית לפי נזורת מרכזית

צ'ור 5.20 מראה תמונה ומציאות שפותיה ע"י חישוב הגרדיאנט והשוואת הנורמה שלו לסק. הסך נבחר כך שאחוז שנקבע מראש של ערכים יעברו אותו. הצ'ור מראה העברה של 10%, 15% ו- 20% מכלל הפיקסלים, וניכר כי השפות בתמונה מתגלוות יפה, ובאופן רובטיבי לערכו של הסך.

כנוקודה אחרונה בסעיף זה נזכיר כי חישוב נזורות רגיש מאוד לרעש אדיטיבי. אין זה מפתיע כיון שרעש מtbody חזק יותר בתדרים הגבוהים, בהם עיקר האינפורמציה היא הרעש ולא תמונות המקור. פועלות גזירה היא פועלות חיזוק התדרים הגבוהים, ולכן אם רעש קיים, הוא יוגבר ויוזיק. על מנת להתגבר על בעיה זו מקובל לבצע סינון מקדים על התמונה לשם דיכוי הרעש בה, לפני חישוב הנזורות. אם



ציור 5.20 - גילוי שפות עיי' נורמת הגרדיאנט. תמונה מקור (שמאל למעלה), תמונה גראדיאנט לאחר סוף המעבר 20% מהפיקסלים (ימין למעלה), העברת 15% מהפיקסלים (שמאל למטה), ו- 10% מהפיקסלים (ימין למטה).

הסינון המקדים הוא ליניארי (כפי שמקובל) והגזירה ליניארית (וכל הנגורות שהראינו הין ליניאריות!), הרי שהמשנן הכלול ליניארי וניתן לביצוע כمسנן אחד. למשל, אם הסינון המקדים הוא 3 על 3 אחד, ורצוננו לבצע גזירה אופקית ממורכזת, המשנן הכלול יהיה:

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \otimes [1 \ 0 \ -1] = \frac{1}{9} \begin{bmatrix} 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \\ 1 & 1 & 0 & -1 & -1 \end{bmatrix}$$

ואנו רואים כאן הפרש אופקי בין ממצאים אзорים במקום בין פיקסלים בודדים - תוצאה זו צפואה.

5.4 פעולות מרוחביות מסתגלות

במסגרת הסעיף הקודם הוכיחו פועלות סינון ליניאריות ושהן ליניאריות, אך משותפת לכלן הייתה התוכונה שהן קבועים במקומות. באופן עקרוני ברור כי עם פועלות שונות במקומות ניתן להשיג תוצאות טובות יותר, אך הקביעה כיצד המשך יהיה תלוי במקומות היא שאלה קשה, ולא מענה מוצלח עלייה, לא השגנו דבר. במסגרת סעיף זה נראה מספר מצומצם של מסננים תלויים ממקום ונציג תוצאות אופייניות להן.

האלגוריתם הראשון אותו נציג קריי מסנן Wallis. הבעיה בה מסנן זה לטפל היא תיקון תחום דינامي וחידוד התמונה. כבר ראינו שכאשר בתמונה מנוצלים מספר מצומצם של ערכי רמות אפור, פועלות מתיחה קונטראסט או יישור היסטוגרמה מצלחים להשיג שיפור. מה קורה אם בתמונה מנוצלות הרמות כולן, אך באופן שאינו מוצלח, כך שלמשל באזור מסוים ישנו רק גוונים כהים ולא רואים פרטיהם, ובאזור אחר רק גוונים בהירים ושוב לא רואים פרטיים? במצב כזה כל הגוונים מצויים, אך בפרטם קשה להבחין. הפתרון לבעה זו הינו מתיחה לוקלית של רמות האפור - פעולה לוקלית במקומות פעולה גלובלית.

מתיחה היסטוגרמה לוקלית יכולה להיעשות ע"י חישוב בכל פיקסל של היסטוגרמות הסביבה, ותיקון ערך הפיקסל על-פי היסטוגרמה זו, כך שאם רמת האפור של הפיקסל המרכזי מצויה מצד התיכון של היסטוגרמה הלוקלית, יש להקטין את רמת האפור עוד, ובכך להשיג חידוד מקומי, ובאופן דומה, אם היא מצד העליון יש להעלותה. הבעיה עם גישה פשוטית זו היא מורכבותה החישובית, שכן בנית היסטוגרמה סיבוב כל פיקסל היה פעולה בלבד.

כתחליף, מסנן Wallis מציע להניח כי היסטוגרמות לוקליות הן בקירוב גאוסיות, ולכן, במקומות לבנותם במלואן, נדרש רק חישוב של שני פרמטרים - ממוצע ושותנות. נתחיל את תיאור האלגוריתם בתיאור דרך החישוב של הממוצע והשותנות באופן מקומי בזורה נוחה. אם נניח כי התמונה הנתונה היא $[n,m]$, חישוב הממוצע בכל פיקסל נעשה ע"י קונבולוציה עם מסנן בצורת גאוסיין רחב אשר רוחבו $\sigma = 3$ עד כמה הפעולה לוקלית (ערך מומלץ הוא $\sigma = 3 - 10 = 5$ פיקסלים). מסנן זה נבנה כמטריצה של 15 על 15 (בחירה שרירותית) ע"י המשוואה:

$$G_{\sigma}[m,n] = \frac{\exp\left\{ \frac{(m-8)^2 + (n-8)^2}{2\sigma^2} \right\}}{\sum_{k=1}^{15} \sum_{j=1}^{15} \exp\left\{ \frac{(k-8)^2 + (j-8)^2}{2\sigma^2} \right\}} \quad 1 \leq n, m \leq 15$$

לכן, תמונה הממוצע תהיה:

$$M_{In}[m,n] = x[m,n] \otimes G_{\sigma}[n,m]$$

קירוב השונות בכל פיקסל נעשה ע"י החסרת תמונה הממוצע מהתמונה המקורית, העלאת התוצאה בריבוע, ושוב קונבולוציה עם אותו מסנן גauss, כלומר:

$$S_{In}[m,n] = (x[m,n] - M_{In}[m,n])^2 \otimes G_\sigma[n,m]$$

ידוע כי בהינתן משתנה אקראי x בעל ממוצע M ושותות S , הערך $S/M-x$ הוא משתנה אקראי חדש בעל ממוצע אפס ושותות 1. באופן דומה, אם נתון משתנה אקראי x בעל ממוצע אפס ושותות 1, ורצוננו להמיריו להיות בעל ממוצע M ושותות S , علينا לבצע את הפעולה $-Mx+S$. לכן, אם רצוננו להגיע בכל פיקסל לממוצע יעד M_0 (למשל - 128), ולשותות יעד S_0 (למשל 40), אזי علينا לבצע את הפעולה הבאה:

$$x_{new}[m,n] = \frac{x[m,n] - M_{In}[m,n]}{S_{In}[m,n]} S_0 + M_0$$

כפי שראויים, תחילת נורמל המשטנה האקראי להיות בעל ממוצע 0 ושותות 1, ומשם הומר לערכי היעד. הפעולה הניל היא פעולה מקומית, אך תלואה (בשל השימוש ב-תמונה M ו- S) בסביבת הפיקסל. מסן זה אינו ליניארי ואיןו חסר זכרו!

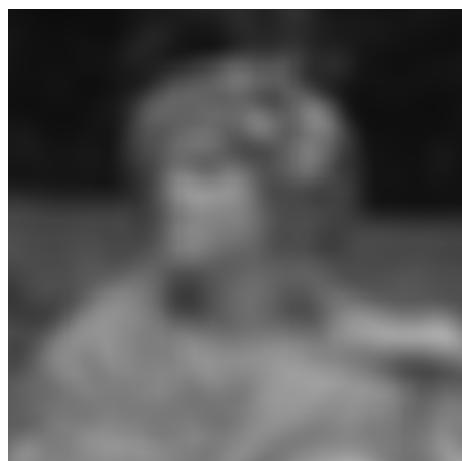
למעשה שיטה זו מניבה תוצאות קיצונית מידי, כיוון שכפיה ממוצע ושותות יעד (S_0, M_0) זהים בכל מקום היא דרישה חזקה. ריסון לאלגוריתם זה מוצע ע"י קביעת תמונות לממוצע ולשותות היעד, ממוצע משוקל של הערכים המקוריים וערבי היעד הקבועים. לכן:

$$\begin{aligned} M_{Dest}[m,n] &= (1-\alpha)M_{In}[m,n] + \alpha M_0 \\ S_{Dest}[m,n] &= (1-\alpha)S_{In}[m,n] + \alpha S_0 \\ \Rightarrow x_{new}[m,n] &= \frac{x[m,n] - M_{In}[m,n]}{S_{In}[m,n]} S_{Dest}[m,n] + M_{Dest}[m,n] \end{aligned}$$

בגישה זו, שימוש ב- $\alpha = 0$ לא משנה את התמונה, ושימוש ב- $\alpha = 1$ נותן את האלגוריתם הקיצוני. עבור ערכיBINIIM נקלט אפקט נאה של חידוד ומתייחת קוונטרסט לוקליים. ציור 5.21 מראה תמונה מקור ותוצאות העיבוד עם מסן Wallis למגוון ערכי α .

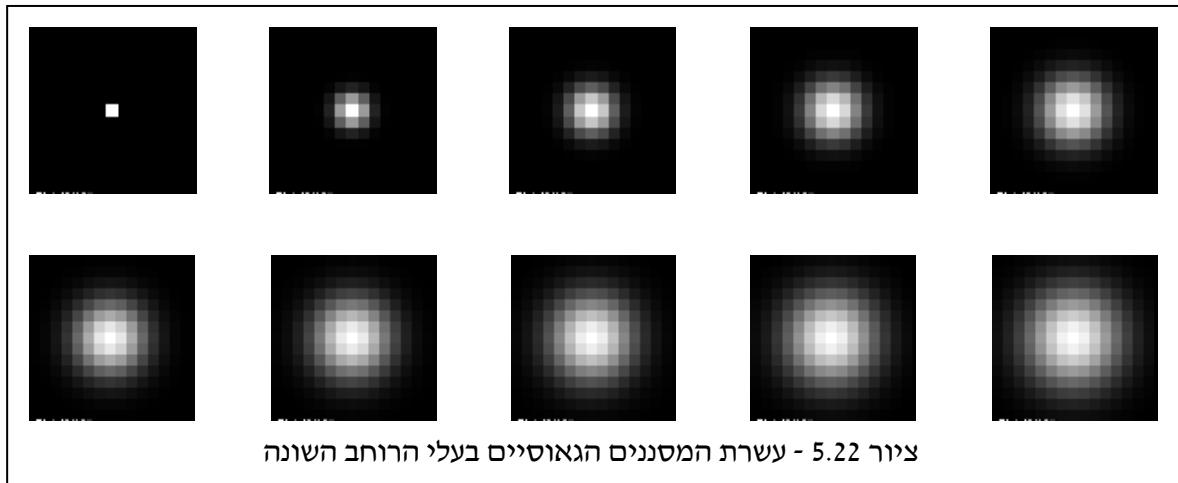
האלגוריתם הבא בו נדון נועד לנוקות תמונה מרעש אדיטיבי. כבר רأינו כי החלקה (LPF) מצליחה יפה בניקוי רעש כאשר מדובר באזורי חלקים. לעומת זאת, באזורי שפות, החלקה גורמת לתופעה הצורמת לצופה האנושי. הרעיון אם כך הוא להפעיל מסן גאוסי בעל שנות תלויות מקום - באזור חלק המсан יהיה רחב, ובאזור שפה או מרקם המсан יהיה צר. כך קיבל כי רעש יידחהיפה במקומות החלקים, וכמעט לא יידחה כלל באזורי שפה.

למעשה, במערכת הראייה האנושית קיימת תופעה הקרויה "טופעת המיסוך", אותה אנו מנצלים כאן. מסתבר כי קיומו של הרעש צורם לנו הרבה יותר באזורי חלקים, ושם אנו עושים עבודה טובה של דיכויו. לעומת זאת, באזורי שפות איןנוכח רגשים לקיומו של רעש, ולכן ניתן להשאירו.



ציור 5.21 - תוצאות סינון Wallis - שורה עליונה (שמאל) תמונה מקור, (ימין) תמונה הממוחע המקומיי, שורה אמצעית (שמאל) תמונה השונות המקומית, (ימין) תוצאה הסינוון עם $\alpha = 0.1$, שורה תחתונה (שמאל) תוצאה הסינוון עם $\alpha = 0.4$, (ימין) תוצאה הסינוון עם $\alpha = 1$.

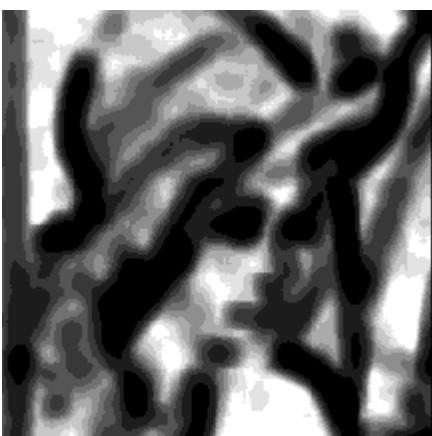
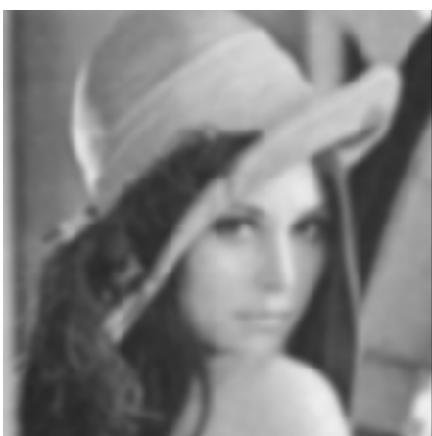
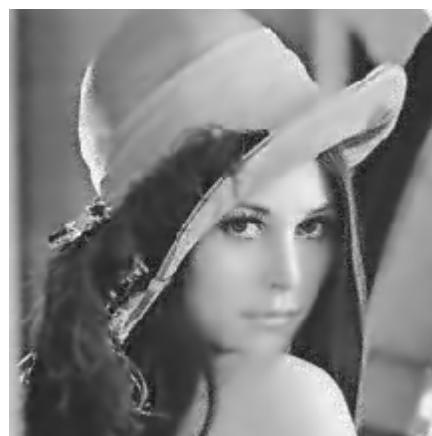
כיצד נקבע אם אзорו הוא כולל שפה? אפשרות אחת היא לבצע גזירה כפי שתואר בסעיף קודם, ולהתיחס לערכים המוחלטים של הגרדיאנט. ברור כי גישה זו חייבת לבוא עם החלקה מקדימה של התמונה לשם מניעת השפעה רעה של הרעש המזוי בתמונה. גישה אחרת בא אנו נעשה שימוש היא להשתמש בתמונה שוניות מקומיות, בדיקת כmo זו שנבנתה עבור מסנן ה-Wallis. נדרש לבנות סט של מסננים גאוסיים בעלי רוחב משתנה, ולבחר מסנן לפי ערך השונות (ביחס הפוך לה - שונות גובהה - רוחב נסוך). צייר 5.22 מראה את המסננים הגאוסיים בהם בחרנו לעשوت שימוש, וצייר 5.23 מראה תוצאות של ניקוי רעש ע"י מסנן אדפטיבי כזה. בצייר זה מתוארת גם תמונה בחירות המסננים (לפי אינדקסים בין 1 ל-10, כמספר המסננים הקיימים).



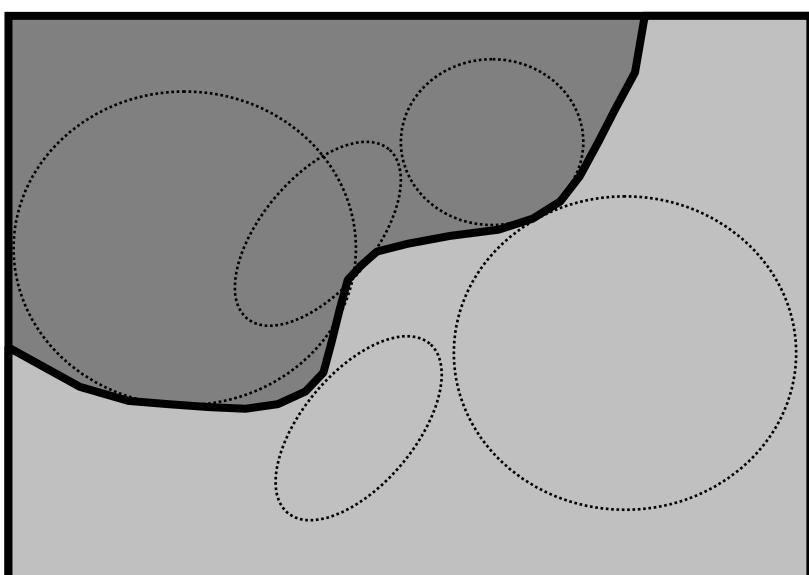
צייר 5.22 - עשרה המסננים הגאוסיים בעלי הרוחב השונה

מסתבר כי האלגוריתם הנ"ל אינו הכי טוב שנוכל לעשות, כאשרנו מפעילים מסנן תלוי פעילות מקומית. כאשרנו באיזור שקט אנו פועלים נכון כאשרנו מטשטשים ע"י מסנן רחוב תמק, וקשה יהיה להתעלות על כך. באוריינטציה, לעומת זאת, אנו בוחרים לא לטשטש, וכתוכאה מכך יותר הרעש כבמקורו. כתחליף לכך, היה ראוי להפעיל באיזור כזה מסנן בעל צורת שאינה עיגול אשר היה מכסה רק פיקסלים מהצד הנכון של השפה. צייר 5.24 ממחיש רעיון זה באופן גרפי. תמק המסנן נבחר להיות עיגול אם אין שפה באיזור התמק. באשר מתקרבים לשפה, מתחווות המסנן ומנה להשיק לקו השפה, ולא לחצותו. לכן, ממוצע משוקלל המבוצע עם מסנן כזה מערב רק אוכולוסייה מצד אחד של השפה.

לרעיון אלו של סינון כיווני عمוק תאורטי שלא יוצג כאן. מסתבר כי ניתן לבסס תכנון סינון שכזה ע"י כלים מחשבוני וריאציות וגיאומטריה דיפרנציאלית, ולהראות כי התוצאה הינה נקודת מינימום של פונקציה על מידת חלקות מסוימת של התמונה. אנו נזכיר גישה זו באופן מפשוט יותר במסגרת הפרק הבא, הדן בשחזור תמונה.

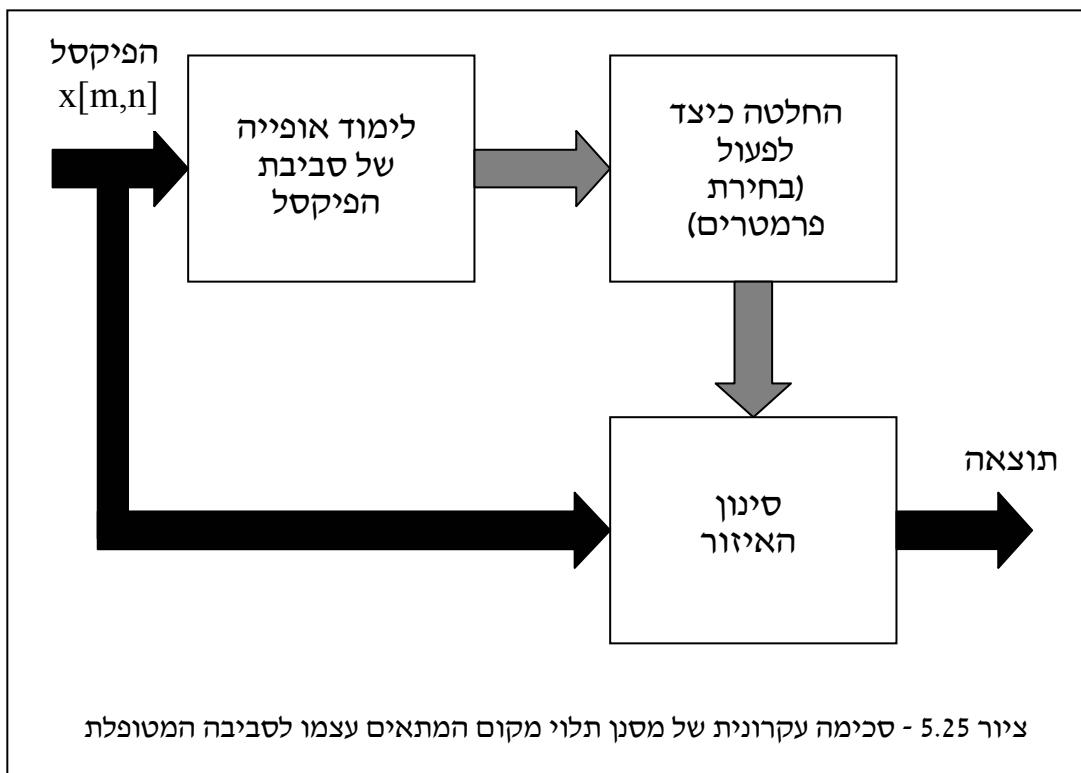


צירר 5.23 - תוצאות סינון אדפטיבי עם גאוסיאן ברוחב משתנה. למעלה (שמאל) תמונה מקור, (ימין) תוצאה אלגוריתם הסינון, למטה (שמאל) תוצאה סינון קבוע במקום ברוחב ממוצע זהה, (ימין) תומנת האינדקסים של המנסנים לפי המקום



צירר 5.24 - שימוש במסנן בעל חתך לא איזוטרופי לשם השגת דחיפות רעש גם ליד שפנות

ניתן המשיך ולתאר עוד מוגוון של מסננים אשר מתאימים עצם לסביבה הפיקסל לשם קביעת אסטרטגיית הטיפול בו, אך אלו נעצור את הדיוון כאן. למרבית המサンנים מהסוג שתואר כאן מבנה עקרוני כמתואר בציור 5.25, בו נעשית בדיקה של אופי המידע בסביבת הפיקסל, עפ"י מידע זה נלקחת החלטה איזו פעולה לבצע, ולבסוף מבוצעת פעולה תלויות תוכן. חשוב להבהיר שאבילי אם המсанן שימושם בסוף הוא ליניארי (כלומר, צירוף משוקל של רמות אפור), אין המسانן הכלול ליניארי, בשל שני הבלוקים האחרים בסכימה.



פרק 6 – שחזור תמונה

113	פרק 6 – שחזור תמונה
113 6.1 בעיית השחזור הקלאליסטית
116 6.2 שיטות שחזור בסיסיות
122 6.3 מושערכים סטטיסטיים – ML ו- MAP
125 6.4 שחזור מבוסס מושערץ ML ו- MAP
132 6.5 שיפור פונקציות ההסתברות לתמונות
140 6.6 בעיות היפוך אחריות והקשר לעייטת השחזור הקלאליסטית

במסגרת פרק זה נדון בשחזור תמונות מקלוקול מסוימים - טשטוש ורעש. הבעיה חשובה ומתעוררת במערכות רבות. אנו נניח כי הן האופרטור הממשתש והן אופי הרעש האדיטיבי ידועים לנו, ובבסיס שיטות לבנייתה של התמונה המקורית אשר קדמה לקלוקולים. זהה בעיה קלאליסטית בעיבוד אותן בכלל, וביעבוד תמונות בפרט, אשר זכתה לטיפול עשיר. למروת זאת, כמשמעותם על התוצאות המתתקבלות, נדמה כי עדין לא נאמרה המילה בתחום.

תיאור נאות של תחום זה מחייב היכרות עם כלים מתחום השערוך (פרק בסטטיסטיקה והסתברות). אנו נציג את הכלים הנחוצים, כך שפרק זה אינו מחייב קראיה נוספת לשלה הבנתו.

1. בעיית השחזור הקלאליסטית

נתחיל בתיאור הבעיה, ונמחייב אותה דרך דוגמה. בידינו מצלמה דיגיטלית ורצוננו לצלם יעד כלשהו. לו הייתה המצלמה שלנו אידיאלית (נטולת הפרעות), היינו מצלמים את התמונה $[n,m]X$ בגודל $L \times L$. לצערנו, בשל בחירתנו לעשות שימוש במצלמה מאיכות ירודה, אנו מקבלים תמונה אחרת $[n,m]Y$ בגודל $L \times L$ (ניתן להכניס לתוכה הסיפור גם מצב של צירמת תמונה קטנה יותר). מהם הקלוקולים שהכנסה המצלמה? ובכן, אנו נניח כי המצלמה עשויה שימוש באופטיקה מאיכות ירודה (כלומר עדשות ירודיות באיכותן) אשר גורמת לטשטוש H . אנו נניח כי טשטוש זה הוא ליניארי וקבוע במקומות. לכן, גורעינו הטשטוש $[n,m]H$ מספר את סיפורה המלא של פעולה זו.

ישנים מקרים בהם מקור הטעטוש אינה האופטיקה אלא תנועה של המצלמה. כאשר המצלמה זהה בזמן החשיפה, תתקבל מריחה כיוונית הקרוייה Motion Blur. מתוך התמונה המקולקלת ניתן לרוב לחשב את גרעין הטעטוש אשר הינו ליניארי וקבוע במקום.

נזכיר לסיפור על המצלמה: בנוסף, מטריצת הגלאים בה המצלמה עשו שימוש סובלת מרעים, וכן התמונה לאחר הטעטוש מקבלת מרכיב אדיטיבי $[n,m]N$ של רעש אקראי בעל ממוצע אפס. אנו נניח לכל אורךו של פרק זה כי רעש זה הינו רעש לבן והומוגני בעל פילוג גאוסי. פירוש הדבר שככל פיקסל מצטרף מספר אקראי גאוסי בעל שונות N קבוע במקומות, וממוצע אפס.

לכן, הקשר בין התמונה הרצiosa $[n,m]x$ לתמונה המתקבלת $[n,m]Y$ נתון ע"י הביטוי:

$$Y[m,n] = X[m,n] \otimes H[m,n] + N[m,n]$$

לשם נוחות מקובל לסדר את התמונות בסידור לקסיקוגרפי לפי עמודות (בחירה שרירותית), ואז הקשר הנ"ל הוא:

$$\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$$

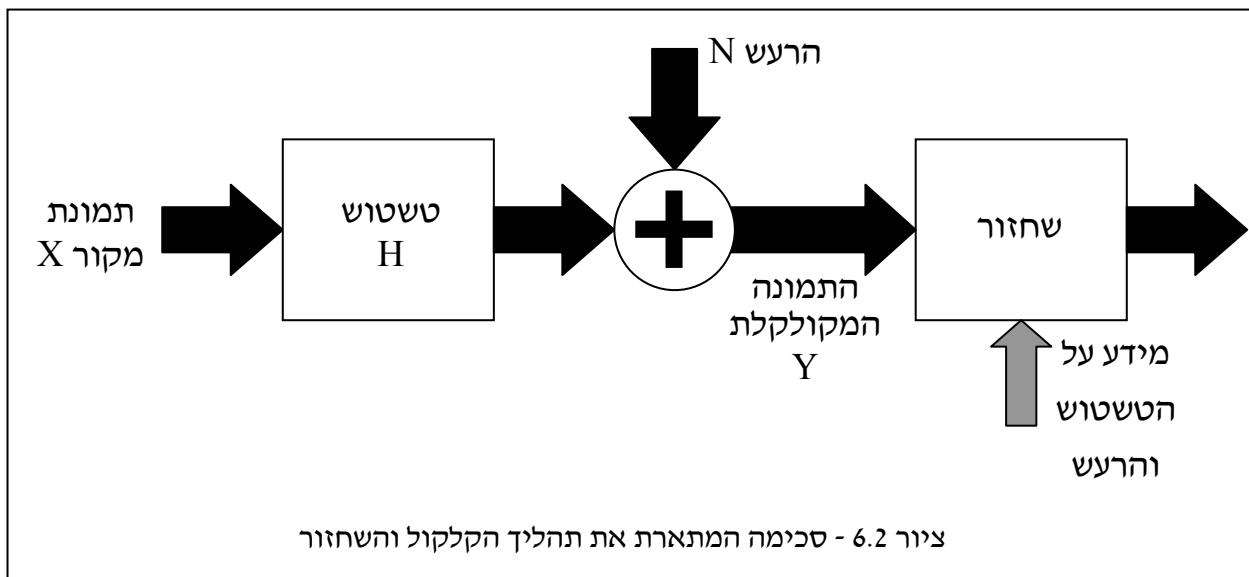
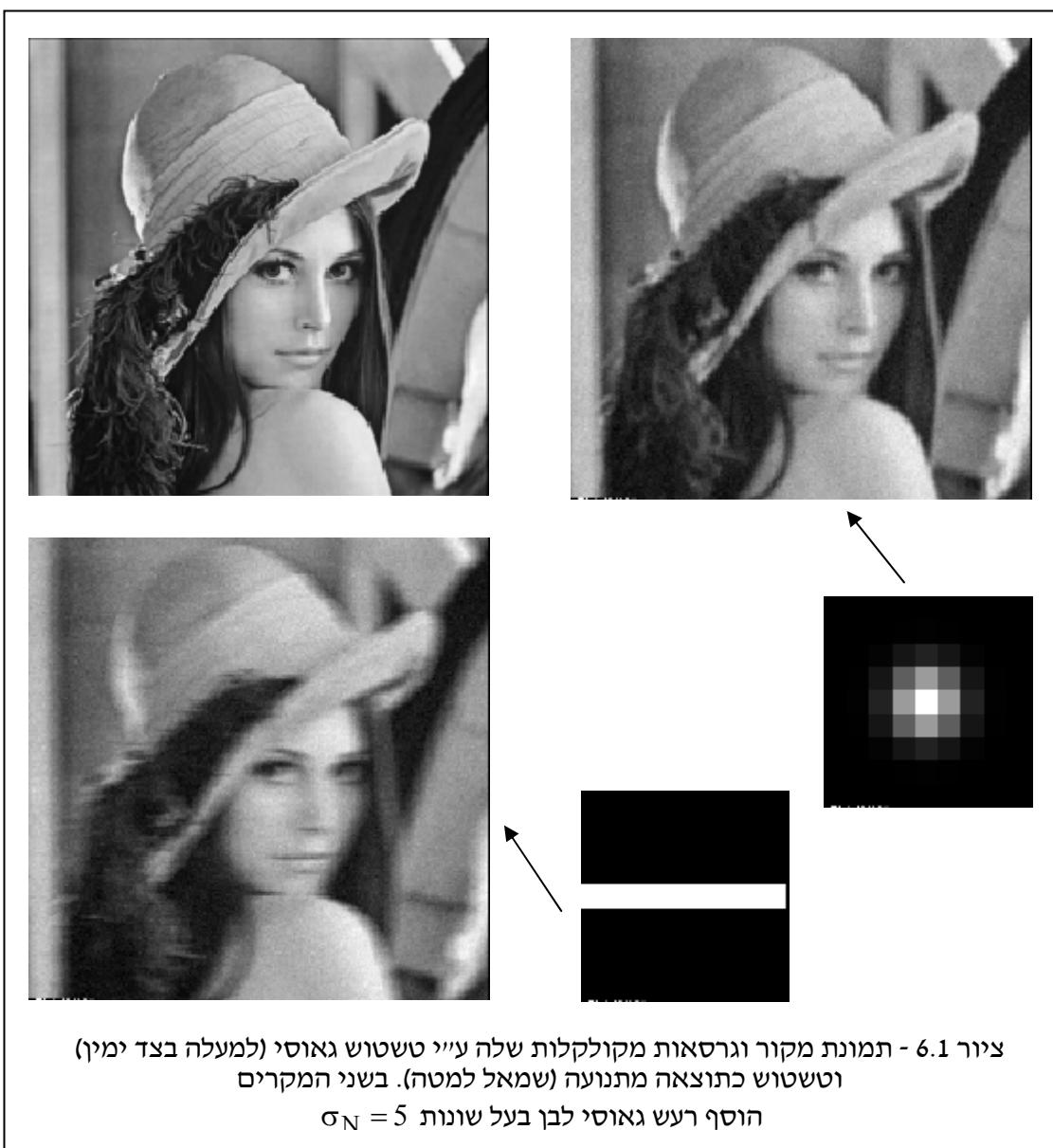
המטריצה \mathbf{H} היא אופרטור הטעטוש בגודל $[L^2 \times L^2]$. בהמשך טבעי לדיוון מפרק 2, נוכל לומר כי כיוון \mathbf{H} מייצגת אופרטור קבוע במקומות, הרי שמטריצה זו בעלת מבנה בלוק טופלי. אם מטריצה זו גדולה מאוד ביחס לגודל הגרעין המטשטש, נוכל אף להגיד בקירוב טוב כי יש לה מבנה בלוק סיבובי. אנו נזדקק להנחות אלו בהמשך.

ציור 6.1 מראה תמונה מקור X וגרסה מקולקלת שלה ע"י טשטוש אופיני לאופטיקה (טעטוש גאוסי בעל חתך רדילי), וגרסה של מריחה הנובעת מתנועה. שני המקרים הוסף רעש אדיטיבי. כאמור,

המטרה היא לחתוך את התמונה המקולקלת, את המידע על גורמי הקלקול, ולשחזר את תמונה המקור.

ציור 6.2 מראה סכימת בלוקים של התהיליך במלואו (קלקול ותיקו). המידע המוזן לתהיליך השחזר הוא התמונה המקולקלת עצמה, גרעין הטעטוש, ואופיו הסטטיסטי של הרעש (במקרה שלנו רק הפרמטר N).

נשים לב כיippi המודל הנ"ל יכולנו בקלות לטפל בטשטוש ליניארי תלוי מקום. גם אז אופרטור הטעטוש ניתן לתיאור כמטריצה \mathbf{H} , אלא שאז אין לה מבנה בלוק טופלי. הבעיה עם טשטוש תלוי מקום היא הדרוש לדעת אותו, ולבנות ממנו את המטריצה \mathbf{H} . אם בשל הבנה עמוקה של התהיליך הפיסיקלי של ייצור התמונה טשטוש זה ידוע, הרי שרוב הדברים שייאמרו בהמשך רלוונטיים גם לטשטוש תלוי מקום.



לעתים קרובות מקובל להציג את בעיית השחזור בתחום התדר. נציג את המטריצה T_L אשר תהיה פועלת DFT חד-מימדית על וקטור באורך L . אזי המטריצה $T_L \otimes T_L = T_{L^2}$ (מכפלת קרונקר של המטריצה עצמה) תהיה מטריצה בגודל $[L^2 \times L^2]$ (כמו H) המייצגת פועלת DFT דו-מימדי על תמונה בגודל $L \times L$, כשהיא נתונה בסידור לקסיקוגרפי. לכן, המכפלה $T_{L^2} X$ היא בעצם התמרת פורייה דיסקרטית על התמונה $[m,n] \times [X]$, אשר תהיה גם היא וקטור באורך L^2 (עם מספרים קומפלקטיים!). ניקח את המשווה הוקטורית ונכנס לתוכה התמרת פורייה (נשתמש בסימון מקוצר T) באופן הבא:

$$\underline{Y} = H \underline{X} + \underline{N} \quad \Rightarrow \quad T \underline{Y} = T H \underline{X} + T \underline{N} = T H T^H T \underline{X} + T \underline{N}$$

$$\text{For: } T \underline{Y} = \underline{Y}_T, \quad T \underline{X} = \underline{X}_T, \quad T \underline{N} = \underline{N}_T, \quad T H T^H = H_T$$

$$\Rightarrow \underline{Y}_T = T \underline{X}_T = H_T \underline{X}_T + \underline{N}_T$$

כאשר הגדים עם סימון T פירושם התמרת תדר. עבור H בлок סיבובי, כבר רأינו בפרק 2 כי התמרת תדר תביא לכלISON, ולכוון H_T מטריצה אלכסונית (זו הסיבה שאנו אהבים כל כך פועלות קבועות במקומות ציקליות). איברי האלכסון הראשי הם הערכים העצמיים של המטריצה המקורית H , ואלה הם גם הערכים שהיו מתקיים בתמרת DFT דו-מימדית ישירה על גרעין הטשטוש שבסנה את H (לאחר שרופף באפסים לגודל המלא של התמונה).

2. שיטות שחזור בסיסיות

אם נתעלם לרגע מהרעש, כיוון שהקלculus הוא טשטוש H , מדובר לא ליישם אופרטור הפוך לו, H^{-1} , לשם השחזור? אם קיים ההיפוך, הרי שאם $-H$ מבנה בлок סיבובי, אזי גם להיפוך מבנה בлок סיבובי. זאת קל לראות כיון שמתקיים:

$$H_T = T H T^H \Rightarrow H_T^{-1} = [T H T^H]^{-1} = T H^{-1} T^H$$

כלומר, רואים כי התמרת DFT על האופרטור הפוך גם היא מטריצה אלכסונית, ולכוון האופרטור H^{-1} גם הוא בлок סיבובי. פירוש הדבר הוא שהמסנן הפוך אף הוא קבוע במקומות, ולכוון ניתן להפעלה כקונבולוציה פשוטה.

אבל, מתעוררות שתי בעיות: ראשית, כלל לא ברור כי $-H$ קיים היפוך. למעשה, במרבית המקרים ניתן לקבוע בודאות כי לא קיים היפוך! קל יותר לראות זאת דזוקא בייצוג בתדר. אם H לא

הפייה (כלומר סינגולריות), איזי גם H_T סינגולריות. לכן, באיברי אלכסונה הראשי יש אפסים. מטריצה כזו לא ניתנת להיפוך! מבחינה אינטואיטיבית ניתן להסביר זאת באופן הבא: נניח כי קיים אפס כזה בשורה K כלשהו. אפס זה כפל את האיבר ה- K של התמורה הגדולה של תמונה הכנסה X. לכן, ערך זה אבד ולא ניתן יהיה לשחזרו.

תיקון לבעה זו יהיה (ושוב, אנו "חושבים" בתדר) לחתך כל ערך אלכסון ראשי שהוא אפס, ולהמירו בערך אחר קרוב לאפס - נניח 1e-10. גישה דומה אך מעט שונה היא להשתמש במסנן המשוחרר הבא (הנתון בתדר) :

$$G_T[k] = \frac{H_T^*[k]}{|H_T[k]|^2 + \text{Const}}$$

כאשר הקבוע קטן ממש מועט מ- 1 (סדר גודל מומלץ - 1.001). לכן, בתדרים נמוכים בהם $H_T^*[k]$ בערכו המוחלט קרוב ל- 1, הקבוע זניח ואנו מקבלים התנאות של מסנן הפוך. עבור תדרים גבוהים בהם $H_T^*[k]$ מתקרב לאפס, הקבוע משמש כהגנה מחלוקת באפס.

עד כה בחרנו להתעלם מקיומו של הרעש. מסתבר כי אפילו אם הרעש חלש, השפעתו בתהליכי של סינון הפוך היא דרמטית. השתמש שוב בתיאור אינטואיטיבי של הבעיה ע"י ייצוג הבעיה בתדר: אם נניח כי בשורה ה- k של H_T יש ערך נמוך יחסית (ניחה 3-1e-3). נניח כמו כן כי בנקודות תדר זו, עצמות האות המקורי ועצמות הרעש נמוכות מאוד אך הרעש חזק יותר מהאות. תופעה זו מתרחשת כמעט תמיד בתדרים גבוהים בהם כמעט אין נוכחות לאות המקור ואז יוצא שהרעש דומיננטי. נניח לנו - $X_T[k] = 1e-3$ ואילו הרעש חזק יותר - $\underline{X}_T[k] = 1e-2$. הפעלת המסנן הפוך פירושה הכפלת שני ערכים אלו בערך 1e3. גרמו בכך זה להכנסה חזקה מאוד של רעש לתמונה התוצאה.

המסנן המשוחרר שهواتו קודם לתנאי יותר טוב. במקרה כזה, נקבל שהן הרעש והן האות מועברים כshedם מוכפלים בגודל :

$$\frac{1 \times 10^{-3}}{1 \times 10^{-6} + \text{Const}} \underset{\text{Const}=0.001}{\approx} 1$$

תוצאה זו לא נראית כמו הכפלת ב- 1000, אך אינה טובה לכשעצמה. בנקודות תדר זו היה ראוי שנכפול את הכנסה בערך נמוך מ- 1 לשם דיכוי הרעש. מAMILIA האות טובל שם ברעש גדול ממנה, ולכן אין מה להפסיד. נציג אם כך את הגישה הבאה: בכל נקודת תדר נברר מהו היחס בין עצמות האות לעצמות הרעש, ולבצע מסנן הפוך רק בתדרים בהם יחס זה גבוה. במקומות בעלי יחס נמוך, בהם הרעש דומיננטי, כתחליף למסנן הפוך, יש להפעיל דיכוי (כלומר הכפלת בערך נמוך מ- 1. דיכוי זה יקטין את האות שם, אך גם את הרעש שAMILIA דומיננטי יותר.

הבעיה עם גישה זו היא הצורך בשני גדים שאותם אין לנו - עוצמת האות המקורי ועוצמת הרעש בכל נקודת תדר. כתחליף, ניתן להניח כי עוצמת הרעש קבועה וזזה בכל נקודת תדר, ושויה N^S . הנחה זו נסונה כאשר הרעש לבן. באשר לאות המקורי, הנחה סבירה היא שעוצמתו יחסית הפוך למרחק מהראשית בתדר. כלומר, עבור תדרים נמוכים יותר צפופה להיות גבוהה, ועם העלייה בתדר עוצמתו יורדת. ביצוע התמורות פוריה למגוון גדול של תמונות מוגלה כי כלל זה נכון ברוב מוחלט של המקרים. ההבדל בין התמונות מתבטא בכך כלל כמעט רק בפaza, ומזו אנו מתעלמים בה.

לסיכום, התהליך המוצע הוא ביצוע התמרת פוריה לאופרטור הטשטוש, ולקבל את H_T . אם

נסמן את ערכי האלכסון הראשי של מטריצה זו ב- $[k] H_T$, מסנן השחזור בתדר, G_T יתקבל ע"י:

$$G_T[k] = \frac{H_T^*[k]}{|H_T[k]|^2 + \alpha \sigma_N^2(r[k])^2}$$

כאשר * מייצג צמוד קומפלקס, ו- $[k]$ הוא המרחק אל הראשית במישור התדר. כך למשל, עבור $k=1$ זהו איבר ה- DC ולכן $=0[1]$. גישה זו דומה להצעתנו הקודמת, למעט העובדה שכעת ה-"קבוע" שלנו גדול עם התדר, ולכן הופך דומיננטי מהר יותר.

מדוע ביטוי זה? בשל העובדה גודל מאפס לכל k , ברור כי לא נקבל היפוך של 0, וכן נפרטרה הבעיה הראשונה. שנית, עבור k נמוכים, כאשר האות חזק מאוד ביחס לרעש, הגודל $[k] z$ קטן מאוד, ונקבל כי המسان הוא בקירוב טוב היפוך מדויק כיון שהאיבר הימני במכנה יהיה זניח. כאשר הרעש דומיננטי מאד, איבר זה (הימני במכנה) יהיה הדומיננטי במכנה ולכן המسان המשוחריר יהיה קרוב לאפס,-CNDRSH.

לאחר שהתקבל G_T (מטריצה אלכסונית) ניתן לבצע התמרת תדר הפוכה לקבלת G . הפעלת G על התמונה \underline{Y} תניב את השחזור. כתחליף, ניתן לבצע את השחזור בתדר ע"י הכפלת G_T ב- \underline{Y} וביצוע התמרת תדר הפוכה על התוצאה - שתי גישות אלו זהות לחולוטין.

הMSN המוצע מוכר מאד וקרויMSN Wiener. בניתו באופן מדויק וקבלת הביטוי שהוצע נעשית מתוך שיקולים הסתברותיים, שלא הצגנו כאן. ציור 6.3 מראה תמונה מקור מושטשת ורוועשת, ושהзор שלה ע"יMSN היפוך. בעוד שההמראק (במונט MSE) בין תמונה המקור לתמונה המוקולקת הוא 103, הרי שההמראק בין תמונה המקור לשחזור כאן הוא 67! התוצאה לכל ספק גראעה!

בציור 6.4 אנו מראים אותה תמונה בדיק את תוכנת MSN Wiener, עבור שלושה ערכי א' שונים. עבור $6 = 5e - 5e$ מתקבלת שגיאה של $1e - 87.1$, עבור $6 = 1e - 6$ השגיאה היא 78.8, ועבור $7 = 5e - 7$ השגיאה עולה ל- 94.8. אנו רואים כי חל שיפור בתהליכי השחזור, הן באופן ייזואלי, והן באופן

כמומי. עם זאת, קשה להתנתק מתחושה של החמוצה כמשמעותם בתוצאות - השיזור אינו מתקרב כלל לאיוכות של תמונה המקורי, ונשאלת השאלה האם זה אכן טוב שניתן להשיג? למולנו, התשובה על כך היא לא! ניתן להציג פתרונות איוכתיים יותר, אך הדבר מחייב שיטות מורכבות יותר, אשר יידונו בסעיף הבא.



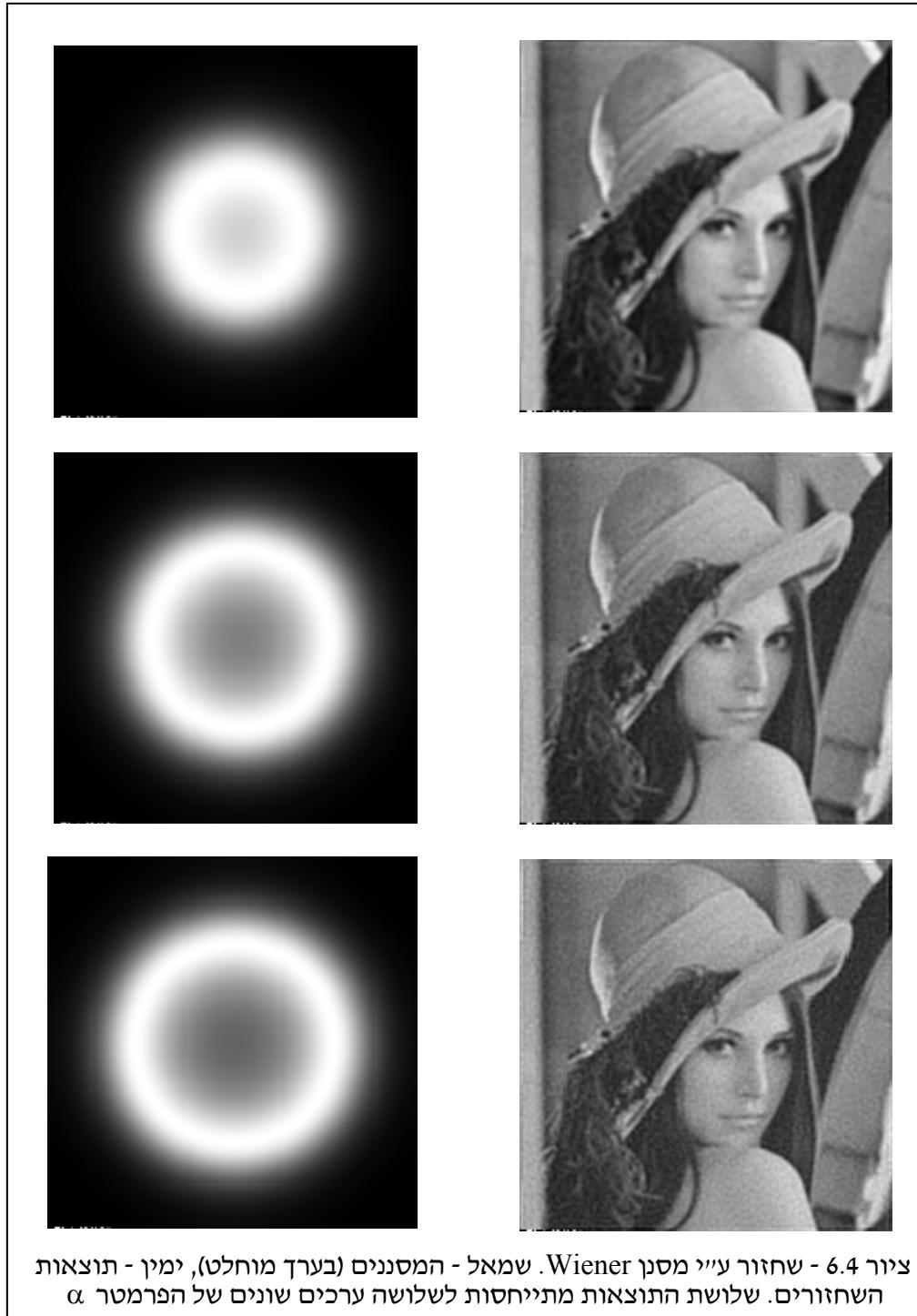
ציור 6.3 - שחזור ע"י מסנן הפוך. תמונה המקורי (שמאל למעלה), תמונה מקולקלת (ימין למעלה), תמונה המסנן בתדר (שמאל למיטה), ותוצאת השחזור (ימין למיטה).

דוגמה 6.1: נניח כי אנו עוסקים באוטות חד מימדיים, הטשטוש הוא הפונקציה הבאה (בתדר) :

$$H_T[k] = \frac{10}{k^2 + 10} \quad k = 0, 1, \dots, 100.$$

המסנן המשוחרר מהו רצוע ראשון בסעיף זה יהיה :

$$G_T[k] = \frac{\frac{10}{k^2 + 10}}{\left[\frac{10}{k^2 + 10} \right]^2 + 0.001} = \frac{10(k^2 + 10)}{100 + 0.001(k^2 + 10)^2}$$

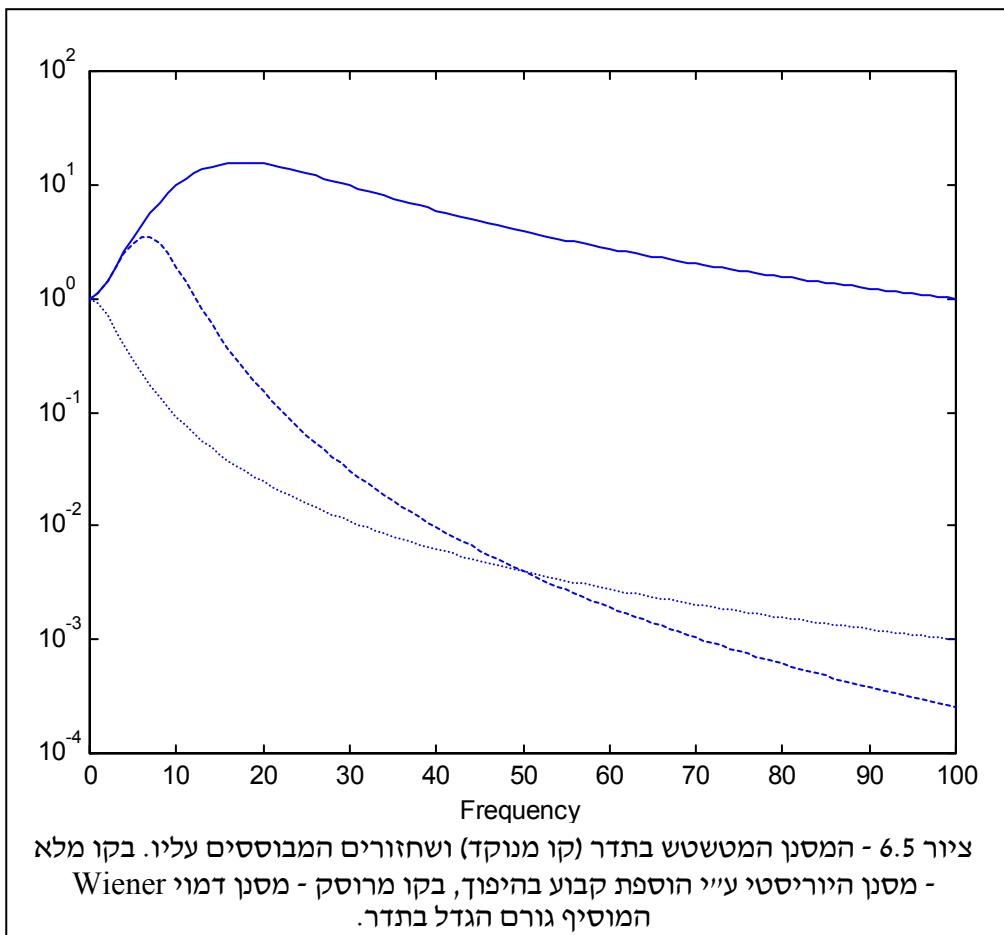


ציור 6.4 - שחזור ע"י מסנן Wiener. שמאל - המסננים (בערך מוחלט), ימין - תוצאות השחזורים. שלושת התוצאות מתיחסות לשולשה ערכים שונים של הפרמטר α

ואילו המסנן המשוחרר השני יהיה (נבחר כאן קבוע באופן שרירותי) :

$$G_T[k] = \frac{\frac{10}{k^2 + 10}}{\left[\frac{10}{k^2 + 10} \right]^2 + 0.001[0.4k^2]} = \frac{10(k^2 + 10)}{100 + 0.0004(k^2 + 10)^2 k^2}$$

צירור 6.5 מראה את שלושת מסננים אלו (הMSN המטשטש, ושני המשחזרים).



אנו רואים כי המSN הבניי על הוספה קבוע נוטן אמנים דעיכה בהגבר אחרי תדר מסוים, אך בכל מקרה הוא גבוה. מסן ה- Wiener מצליח לרדת חזק יותר ואף לתת ניחות חזק יותר משל המSN המטשטש - תוצאה רצiosa כיון שאנו מניחים שבאזור זה חזק הרעש מהאות. בשני המקרים, התנהוגות באיזור התדרים הנמוכים דומה - התנהוגות של מסן הפוך אשר אמורה לשחזר יפה מרכיבי תדר אלה.

6.3 משרכים סטטיסטיים – ML ו- MAP

בסעיף זה נציג באופן תיאורתי שני משלדים סטטיסטיים, בהם נעשה שימוש לטובת בניית אלגוריתמי שחזור תמונה טובים יותר. משלדים אלו הם משליך הסבירות המירבית (ML - Maximum Likelihood A-posteriori), ומשליך ההסתברות הסופית המקסימלית (MAP - Probability). בשני המקרים הסיפור מתחיל ידי דומה - נתנו לנו סט נתונים שיטומן \bar{Y} , ורצוננו לשערך ערך לא ידוע X . ידוע לנו קשר בין X ל- \bar{Y} באופן כללי $\bar{Y} = f(X)$. נמחיש את הסיפור הניל זרך דוגמה פשוטה, ורצוננו להציג דרך "חכמה" לחישוב של הנעלם X .

דוגמה 6.1: נניח נתונים לנו סידרה בת 100 מספרים y_k אשר ידוע כי הוגלו עפ"י חוק פילוג גאוסי חסר זיכרון עם ממוצע לא ידוע ושונות 1. הוקטור \bar{Y} יהיה סט מספרים אלו. הנעלם יהיה M - הממוצע. בrama אינטואיטיבית ידוע לנו כי שערוך ממוצע במצב כזה יכול להיעשות פשוט ע"י ממוצע המדידות, דהיינו :

$$\hat{M} = \frac{1}{100} \sum_{k=1}^{100} y_k$$

משליך הסבירות המירבית דורש מأتנו לבנות את פונקציית ההסתברות $P(Y/X)$, כלומר פונקציה שתקבע מהי הסבירות שהמדידות תהינה \bar{Y} עבור ערך הנעלם X . פונקציה זו היא פונקציה של המשתנה X . אם הצלחנו לנתח פונקציה כזו, המשערך מציע כי נמצא את ערכו של X המביא את הסתברות זו למקסימום :

$$\hat{X}_{ML} = \operatorname{ArgMax}_X P(Y/X)$$

נחוור לדוגמה הקודמת ונראה כיצד פונקציה זו מתאפשרת.

דוגמה 6.1 (המשך): חוק ההסתברות של הוקטור \bar{Y} הוא פשוט מכפלה של 100 הסתברויות לכל אחד מאיבריו, כאשר כל איבר כזה הוא הסתברות גאוסית. אם נניח כי אנו יודעים שהממוצע הוא M נקבל :

$$\begin{aligned} P(Y/M) &= \prod_{k=1}^{100} P(y_k/M) = \prod_{k=1}^{100} C \cdot \exp \left\{ -\frac{(y_k - M)^2}{2} \right\} = \\ &= C^{100} \cdot \exp \left\{ -\frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2 \right\} = L(M) \end{aligned}$$

הפונקציה שהתקבל נקראת פונקציית הסבירות L , והיא מעלה המשתנה M - הממוצע הלא נודע. כזכור, רצוננו למצוא את M שיביא פונקציה זו למקסימום, ולכן :

$$\underset{M}{\operatorname{ArgMax}} \left\{ C^{100} \cdot \exp \left\{ -\frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2 \right\} \right\} = \underset{M}{\operatorname{ArgMin}} \left\{ \sum_{k=1}^{100} (y_k - M)^2 \right\} = \frac{1}{100} \sum_{k=1}^{100} y_k$$

ואנו רואים כי ה- ML נתן את התוצאה שהצענו קודם מנקודת אינטואיציה.

דוגמה 6.2: נניח מקרה דומה, אלא שהפעם סט המדידות נובע מנקודת התפלגות אקפוננציאלית מהצורה :

$$P(y_k / M) = C \cdot \exp \left\{ -\lambda |y_k - M| \right\}$$

מהו משערץ ה- ML למוצע M במקרה זה? הפתרון בדומה לקודם יהיה לכתוב את ההסתברות המותנית :

$$\begin{aligned} P(Y/M) &= \prod_{k=1}^{100} P(y_k / M) = \prod_{k=1}^{100} C \cdot \exp \left\{ -\lambda |y_k - M| \right\} = \\ &= C^{100} \cdot \exp \left\{ -\lambda \sum_{k=1}^{100} |y_k - M| \right\} = L(M) \end{aligned}$$

מינימיזציה הפעם תניב :

$$\underset{M}{\operatorname{ArgMax}} \left\{ C^{100} \cdot \exp \left\{ -\lambda \sum_{k=1}^{100} |y_k - M| \right\} \right\} = \underset{M}{\operatorname{ArgMin}} \left\{ \sum_{k=1}^{100} |y_k - M| \right\} = \text{MED}\{y_1, y_2, \dots, y_{100}\}$$

התוצאה הפעם פחות צפואה. אנו רואים כי חישובו של ממוצע הפילוג תלוי למעשה בחוק הפילוג עצמו. יכולנו לשערץ את המוצע ע"י מיצוע אלגברי, אך במקרה זה שערוך זה צפוי להיות פחות טוב (פחות ב_errno).

עד כה דיברנו על משערץ ה- ML. לפי גישה זו הנחנו כי הנעלם הוא פרמטר בחוק הפילוג ולא משתנה אקראי בעצמו. מה אם ידוע לנו כי הנעלם X הוא אכן משתנה אקראי, וידוע לנו חוק הפילוג שלו? למשל, עבור הדוגמה הקודמת, מה אם ידוע לנו כי המוצע M הוא משתנה אקראי גaussiy בעצמו עם ממוצע 5 ושונות 2. כיצד נרתום מידע זה על מנת לשערץ טוב יותר את הנעלם?

התשובה היא משערץ ה- MAP, המציע את השימוש בחוק Bayes. משערץ ה- MAP מציין להביא למסתומים את ההסתברות $P(X/Y)$, שהינה ההסתברות לקבלת הערך X כשתנותן המדידות Y (בדיקת ה琐 ל- ML). נוסחת Bayes קובעת כי :

$$P(X/Y) = \frac{P(X,Y)}{P(Y)} = \frac{P(X)P(Y/X)}{P(Y)}$$

ולכן משערך ה- MAP ינסה להשיג מקסימום להסתברות זו ביחס לערכו של X. מקסימיזציה ביחס ל- X פירושה שהמכנה $P(Y)$ אינו משפיע כלל, ולכן :

$$\hat{X}_{MAP} = \operatorname{ArgMax}_X P(X/Y) = \operatorname{ArgMax}_X \{P(X)P(Y/X)\}$$

ישנו הרבה דמיון לשערוך ה- ML. ההבדל הוא בכך שפונקציית הסבירות $P(Y/X)$ מוכפלת בפונקציית ההסתברות של X הקרויה לעיתים Prior (כקיצור ל- Prior-Knowledge - מידע מוקדים). מעניין לציין, שכאשר אין לנו מידע מוקדים על X, אנו נניח כי הפילוג שלו מתפזר אחד על פני כל הערכים האפשריים, ואז נקבל כי משערך ה- MAP הופך להיות משערך ה- ML.

דוגמה 6.3: בהמשך לדוגמה 6.1, נניח כי ידוע כי הממוצע הלא נודע מפולג גאוסית לפי :

$$P(M) = C_0 \exp\left\{-\frac{(M-5)^2}{2\sigma_M^2}\right\}$$

במקרה זה, שערוך ה- ML לא משתנה. לעומת זאת שערוך ה- MAP יהיה :

$$\begin{aligned} \operatorname{ArgMax}_M & \left\{ C^{100} \cdot \exp\left\{-\frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2\right\} \cdot C_0 \cdot \exp\left\{-\frac{(M-5)^2}{2\sigma_M^2}\right\} \right\} = \\ & = \operatorname{ArgMin}_M \left\{ \frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2 + \frac{(M-5)^2}{2\sigma_M^2} \right\} = \frac{\frac{5}{\sigma_M^2} + \sum_{k=1}^{100} y_k}{100 + \frac{1}{\sigma_M^2}} \end{aligned}$$

עבור $0 \rightarrow M$ קיבל כי התוצאה שואפת ל- 5. זה צפוי כי כאשר $0 \rightarrow M$ פירוש הדבר שבודאות מוחלטת אנו יודעים כי הממוצע שווה 5 (עוד בטרם התקבלו מדידות). עבור $\infty \rightarrow M$ קיבל כי הממוצע האלגברי הוא השערוך, כמו שהציגו ה- ML. גם זו תוצאה צפואה כיון שעבור שונות גדולות מאוד, פירוש הדבר הוא שאין לנו בעצם מידע מוקדים על הנעלם, ולכן יש לדלותו מהמדידות בלבד.

6.4 שחזור תמונות מבוסס ML ו- MAP

נחוור את לביעית השחזור שלנו, כאשר מנסים להשתמש במשערץ ה- ML. וקטור המדידות הוא \underline{Y} , וקטור הנעלמים שלנו הוא \underline{X} . הקשר ביניהם נתון ע"י $\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$, כאשר \mathbf{H} מטריצת טשטוש ו- \underline{N} וקטור רעש אקראי גaussiano בעל חוק פילוג:

$$P_N(\underline{N}) = C \cdot \exp\left\{-\frac{\underline{N}^T \underline{N}}{2\sigma_N^2}\right\}$$

עפ"י ביטוי זה, ההסתברות $P(\underline{Y}/\underline{X})$ אינה אלא:

$$P_N(\underline{Y}/\underline{X}) = C \cdot \exp\left\{-\frac{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]}{2\sigma_N^2}\right\}$$

זאת כיון שם מתיחסים ל- \underline{X} כוקטור פרמטרים נתון, אזי הגורם $\mathbf{H}\underline{X}$ מהויה תוספת דטרמיניסטית, והוקטור החדש $\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$ הינו גausiano גם כן, בעל אותה שונות, ועם ממוצע חדש שהינו הגדל שהוסף. משערץ ה- ML יהיה הוקטור $\underline{\hat{X}}$ אשר יביא את ההסתברות הנ"ל למקסימום. לכן:

$$\begin{aligned} \underline{\hat{X}}_{ML} &= \underset{\underline{X}}{\operatorname{ArgMax}}\{P_N(\underline{Y}/\underline{X})\} = \underset{\underline{X}}{\operatorname{ArgMax}}\left\{\exp\left\{-\frac{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]}{2\sigma_N^2}\right\}\right\} = \\ &= \underset{\underline{X}}{\operatorname{ArgMin}}\left\{\frac{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]}{2\sigma_N^2}\right\} = \underset{\underline{X}}{\operatorname{ArgMin}}\{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]\} = \\ &= [\mathbf{H}^T \mathbf{H}]^{-1} \mathbf{H}^T \underline{Y} \end{aligned}$$

התוצאה שהתקבלה נcona רק אם קיימת מטריצה הפוכה ל- $\mathbf{H}^T \mathbf{H}$, שפירושה של \mathbf{H} עצמה להיות לא סינגולרית. למעשה, משערץ ה- ML מציין לנו להביא למינימום את הפונקציה $f(\underline{X}) = [\mathbf{H}\underline{X} - \underline{Y}]^T [\mathbf{H}\underline{X} - \underline{Y}]$ - הצעה בהחלט אינטואטיבית שפירושה, מצא את התמונה שאם טוטוטש ע"י \mathbf{H} היא תהיה קרובה ככל האפשר ל- \underline{Y} .

התוצאה שהתקבלה אינה אלא ביצוע של מסנן הפוך כיון שם \mathbf{H} ריבועית מתקבל:

$$\underline{\hat{X}}_{ML} = [\mathbf{H}^T \mathbf{H}]^{-1} \mathbf{H}^T \underline{Y} = \mathbf{H}^{-1} \mathbf{H}^{-T} \mathbf{H}^T \underline{Y} = \mathbf{H}^{-1} \underline{Y}$$

כאשר \mathbf{H} אינה ריבועית (בניגוד לקרה שלנו), התוצאה הניל הינה היפוך מוכפל (Pseudo-Inverse) המוכר לנו מאלgebra ליניארית. תוצאה מעניינת שהתקבלה כאן היא שלעוצמת הרוש אין משקל בקביעת המשנן המשוחרר.

בכל מקרה, ההנחה שקיים היפוך ל- \mathbf{H} היא הנחה מגבילה. על מנת להיפטר מה הצורך בהנחה זו, מקובל לעשות שימוש ברגולרייזציה. הרעיון די פשוט - המטריצה $\mathbf{H}^T \mathbf{H}$ חיובית חצי מוגדרת - עובדה שפירושה שככל ערכיה העצמיים ממשיים ואי-שליליים. לו היו כל הערכים העצמיים מעל האפס ממש, הרי $\mathbf{H}^T \mathbf{H}$ הייתה הפיכה. לכן, נוסיף ל- $\mathbf{H}^T \mathbf{H}$ מטריצת יחידה (I) מוכפלת בקבוע חיובי קטן מאוד ϵ^2 .² תוספת זו תזיז את כל הערכים העצמיים מעלה, ותביא למצב בו $\mathbf{I} + \epsilon^2 \mathbf{H}^T \mathbf{H}$ הפיכה. לכן, השחזור יהיה:

$$\hat{\underline{\mathbf{X}}}_{ML} = [\mathbf{H}^T \mathbf{H} + \epsilon^2 \mathbf{I}]^{-1} \mathbf{H}^T \underline{\mathbf{Y}}$$

ההנחה היא שם ϵ^2 קטן די לא להשפיע על ההתנהגות הכללית, ועם זאת, בעל ערך גבוה די למניעת בעיות יציבות נומריות, נקבל שחזור לא רע. בפרט, יכולנו להוסיף למטריצה $\mathbf{H}^T \mathbf{H}$ כל מטריצה חיובית מוגדרת \mathbf{M} ולהציג את אותו אפקט, אך בעצם אנחנו יודעים אייזו מטריצה לבחור. גילינו, אם כך, כי משיקולי ML יוצא שיש להפעיל מסנן היפוך לשם שחזור, וניתן ליצבו ע"י בחירה שרירית של מטריצה רגולרייזציה.

נשתמש כעת במשערץ ה- MAP. לצורך השימוש בו علينا לעשות שימוש בפונקציית הסתברות $P(\underline{\mathbf{X}})$. פונקציית הסתברות זו פירושה שבהתמונה $\underline{\mathbf{X}}$, נוכל לומר מהי מידת סבירותה. מה עשויה תמונה לסבירה יותר? שאלה קשה אשר עד היום אין תשובה מוצלחת עליה. אבל, ישן בהחלט תשובות חלקיים בהן נוכל לעשות שימוש.

ה策עה מאוד פופולרית לאמדן הסתברות של תמונה היא ה策עה לעשות שימוש במידת חלקותה של התמונה. ניתן לטעון כי תמונה חלקה (כמשטח בתלת-מימד) יותר סבירה מתמונה שאינה חלקה. ברור כי באזוריים חלקיים בתמונה הנחה זו נכון, ובאזורים שפות הנחה זו אינה נכון. נתעלם בשלב זה מפגם זה (מאוחר יותר נראה דרכי להתגבר על סטייה זו). את חלקותה של התמונה נוכל למדוד ע"י הפעלת נזרת כלשי (גרדיאנט, לפלייאן ...) וחישוב סכום ריבועי ערכי התוצאה. אם התמונה חלקה, הרי שתוצאות הנזרות נמוכה בכל מקום ולבן הערך הכלול יהיה נמוך. אם בתמונה יש הרבה שפות נקלט כי הנזרות גבוהות שכן הערך הכלול יהיה גבוה. אם נניח כי המטריצה \mathbf{D} מייצגת פועלות נזרות מסדר שנבחר, פונקציית ההסתברות המוצעת היא:

$$P(\underline{\mathbf{X}}) = C \cdot \exp \left\{ - \frac{[\mathbf{D}\underline{\mathbf{X}}]^T [\mathbf{D}\underline{\mathbf{X}}]}{2\sigma_D^2} \right\}$$

הbianioי $\underline{\mathbf{X}}$ מקבל את התמונה \mathbf{X} ומבצע עליה פעולה גזירה, והbianioי $[\underline{\mathbf{D}} \underline{\mathbf{X}}]^T [\underline{\mathbf{D}} \underline{\mathbf{X}}]$ לוקח את תוצאת הגזירה מעלה בריבוע כל איבר ומסכם. לכן, ערך זה יהיה נמוך לתמונות חלקות, ובוגה לתמונות שאינן חלקות. השימוש במבנה של פונקציית אקספוננט נובע מנוחות שתתבהר בהמשך. הפרמטר σ_D קובע את שוננותה של הסטברות זו. בכלל אופן, רואים שעבור תמונה חלה ההסתברות בגובה, ועבור תמונה עם ריבוי שפות, ההסתברות נמוכה, כנדרש.

משערך ה- MAP מביא את מכפלת פונקציית הסבירות (הלקוחה מה- ML) ופונקציית ה- $\hat{\underline{\mathbf{X}}}_{MAP} = \operatorname{ArgMax}_{\underline{\mathbf{X}}} \{P_N(\underline{\mathbf{Y}} / \underline{\mathbf{X}})P_X(\underline{\mathbf{X}})\} =$

הניל למקסימום, וכך:

$$\begin{aligned} \hat{\underline{\mathbf{X}}}_{MAP} &= \operatorname{ArgMax}_{\underline{\mathbf{X}}} \left\{ P_N(\underline{\mathbf{Y}} / \underline{\mathbf{X}})P_X(\underline{\mathbf{X}}) \right\} = \\ &= \operatorname{ArgMax}_{\underline{\mathbf{X}}} \left\{ \exp \left\{ - \frac{[\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}]^T [\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}]}{2\sigma_N^2} \right\} \cdot \exp \left\{ - \frac{[\underline{\mathbf{D}}\underline{\mathbf{X}}]^T [\underline{\mathbf{D}}\underline{\mathbf{X}}]}{2\sigma_D^2} \right\} \right\} = \\ &= \operatorname{ArgMin}_{\underline{\mathbf{X}}} \left\{ \frac{[\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}]^T [\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}]}{2\sigma_N^2} + \frac{[\underline{\mathbf{D}}\underline{\mathbf{X}}]^T [\underline{\mathbf{D}}\underline{\mathbf{X}}]}{2\sigma_D^2} \right\} = \\ &= \operatorname{ArgMin}_{\underline{\mathbf{X}}} \left\{ [\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}]^T [\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}] + \frac{\sigma_N^2}{\sigma_D^2} [\underline{\mathbf{D}}\underline{\mathbf{X}}]^T [\underline{\mathbf{D}}\underline{\mathbf{X}}] \right\} = \\ &= \left[\mathbf{H}^T \mathbf{H} + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{D} \right]^{-1} \mathbf{H}^T \underline{\mathbf{Y}} \end{aligned}$$

אנו רואים כי הפעם, המסנן התקבל עם רגולרייזציה שנבעה מתוך שיקולי ה- MAP. ככלומר, אם רצוננו בשחזור מוצלח, מוטב שנבחר לעשות שימוש מטריצת גזירה כלשהי. כמו כן, אנו רואים כי מנת שתי השונוויות פועלת במשקל לגורם הרגולרייזציה הנוסף.

בחשוואה ל- ML, משערך ה- MAP מציע מינימיזציה של הפונקציה

$$f(\underline{\mathbf{X}}) = [\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}]^T [\underline{\mathbf{Y}} - \mathbf{H}\underline{\mathbf{X}}] + \frac{\sigma_N^2}{\sigma_D^2} [\underline{\mathbf{D}}\underline{\mathbf{X}}]^T [\underline{\mathbf{D}}\underline{\mathbf{X}}]$$

בפונקציה זו המרכיב הראשון זהה לזה של ה- ML, והרכיב השני מושך את הפתרון לעבר תמונות חלקות יותר. אם תמונה מסוימת מצלילה להביא לערך מאוד נמוך את הגורם הראשון אין בכך די לשם קבלתה כפתרון - המחיר השני גם הוא משחק תפקיד בקביעת הפתרון. כאשר σ_D נמוך מאוד פירוש הדבר שאנו דורשים שהتوزאה תהיה חלקה מאוד, ואמנם בפונקציה f הגורם השני יהיה דומיננטי, ולכן הפתרון יהיה חלק. עבור $0 \rightarrow \sigma_D$ הבעה דורשת חלקות מוחלטת, וכך התמונה תהיה קבועה עבור גזירה

ראשונה ומישור מותה עבר גירה שנייה. כאשר $\rightarrow \sigma$ אנו מחלשים את דרישת החלוקות ואז יתנקז הפתרון ליצאת ה-ML.

בנוקודה אחרונה בסעיף זה נתייחס לפתרון מעשי של בעית השחזר ע"י הכלים שהוצעו לעיל. ניקח את תוצאת ה- MAP כיון שהיא מכלילה את ה- ML, ונראה כיצד מעשית מושגים שחשזר באמצעותה. נניח כי לשם רגולריזציה אנו משתמשים במסנן 3 על 3 לפלייאן. אם נניח כי התמונה בגודל 200×200 , הרי שהמטריצות המעורבות (H, D) תהיה בגודל של $40,000 \times 40,000$. מטריצות אלו דليلות (מרבית איבריהם אפסים), אך היפוך מטריצה בגודל כזה אינו בא בחשבון. שתי אלטרנטיבות קיימות הן פתרון במישור התדר, ושיטות איטרטיביות.

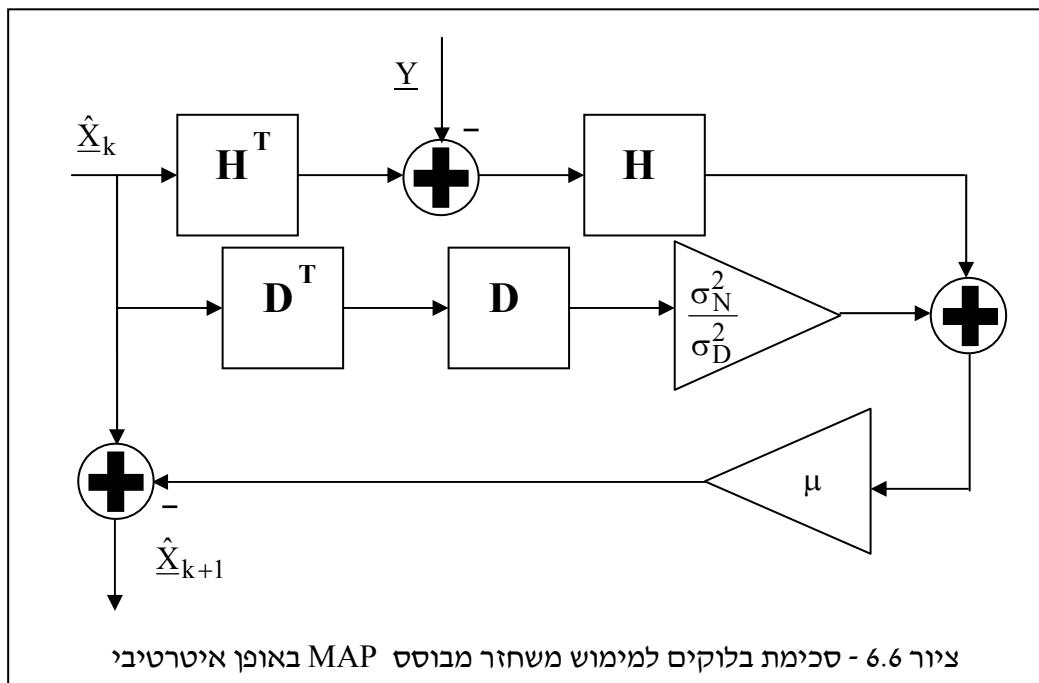
באשר לפתרון איטרטיבי, הרעיון הבסיסי ביוטר הוא אלגוריתם השיפוע המירבי (Steepest Descent). הרעיון הוא להתחיל מניחוש מוצלח לפתרון (למשל $\hat{X}_0 = \hat{X}$) ולשפרו שוב ושוב ע"י ירידה בכיוון הפוך לכיוון הגרדיאנט של הפונקציה (X_f). הגרדיאנט מצביע על הכוון בו חלה העלייה הגדולה ביותר, ולכן הליכה בכיוון הפוך פירושה ירידה מירבית. לכן, משוואת האיטרטציה תהיה:

$$\hat{X}_{k+1} = \hat{X}_k - \mu \frac{\partial f(\hat{X})}{\partial \hat{X}} \Big|_{\hat{X}_k} = \hat{X}_k - \mu \left[\mathbf{H}^T (\mathbf{H} \hat{X}_k - \underline{Y}) + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{D} \hat{X}_k \right]$$

הגודל μ הוא גודל הצעד לעדכון - ערך גדול מדי ייתן התבדרות, ערך קטן מדי ייתן התכנסות איטית מדי, ויש לבחור אותו בחוכמה (קיימות דרכים לקביעתו באופן אוטומטי, שלא ידונו כאן). התהליך המתואר מובטח להתכנס לפתרון במדדיק של הבעה בשל העובדה של הפונקציה f קמורה, ובשל תכונות האלגוריתם המוצע.

כיון ש- H ו- D מסננים ליניאריים וקבועים במקום, הרי שגם שחלוף פעולות אלו הין ליניאריות וקבועות במקום. יתרה מזו, כאשר הטשטוש בעלת גרעין סימטרי, מתקבל כי $\mathbf{H}^T = \mathbf{H}$ (גם הפלסיאן מקיים תכונה זו). לכן, אין כלל צורך לעבור עם תמונות בסידור לקסיקוגרפי או מטריצות כה גדולות. ניתן פשוט להתייחס למשווה הנ"ל כסדרת צעדי סינון שעל הפתרון הקודם לעבור לשם עדכונו. ציור 6.6 מראה את סכימת הבלוקים של תהליך העדכון.

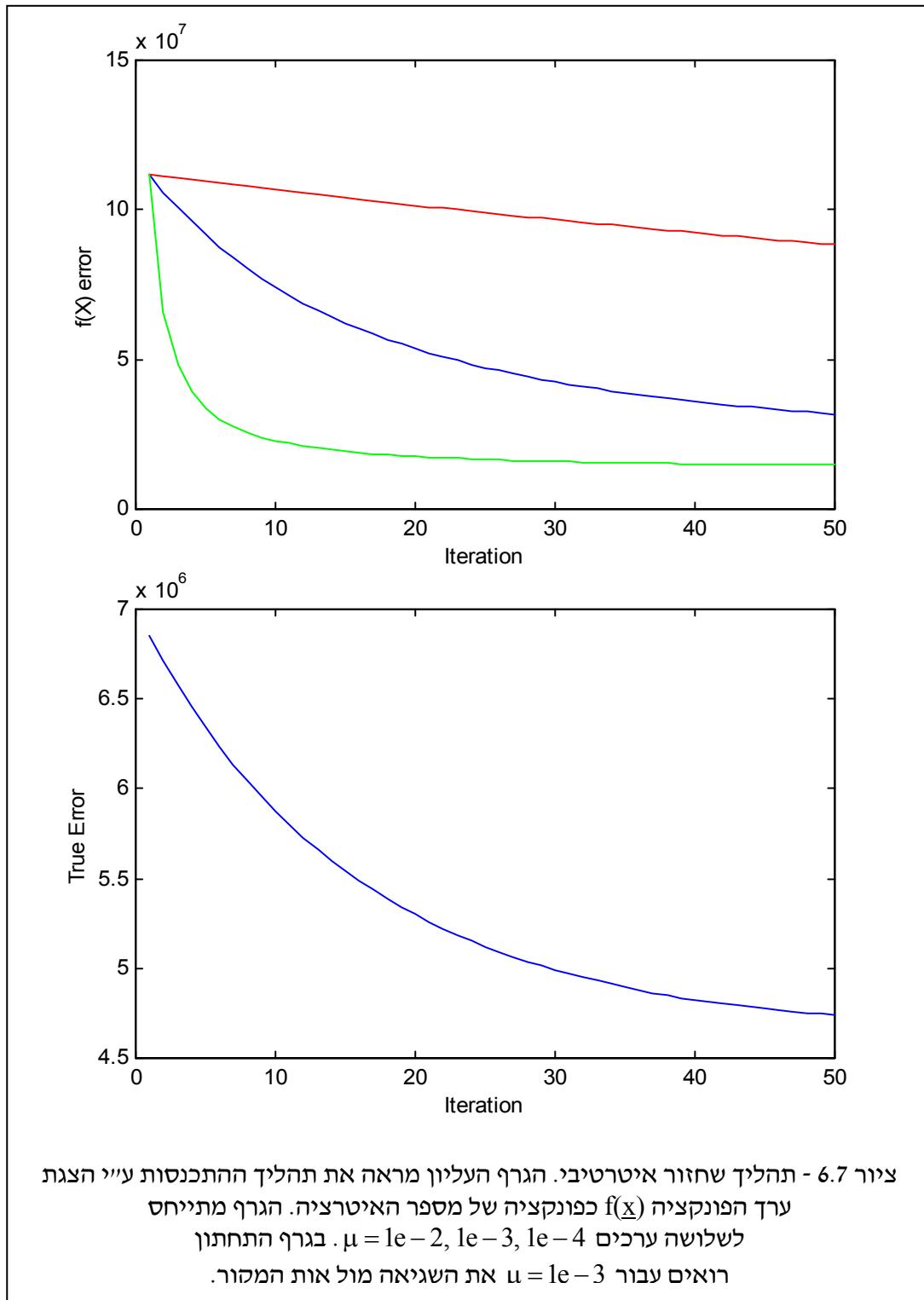
ציור 6.7 מראה גרף של ערך הפונקציה (X_f) שמובאות למינימום בתהליך האיטרטיבי כפונקציה של מספר האיטרטציה. בהרצה זו נעשה שימוש בפלסיאן לצורך רגולריזציה, ו- $\sigma = 25 = \frac{2}{D}$. רואים כי חלה ירידה עקבית עד להתכנסות. הגрафים השונים מתייחסים לערכי μ שונים. כפי שכבר נאמר, עבור ערך נמוך מדי ההתכנסות איטית, ועבור ערך גדול מדי חלה התבדרות.



הגרף השני בציור זה מראה את השגיאה $\|\underline{X} - \hat{\underline{X}}\|^2$. כפי שראאים, גם כאן מושגת ירידת עקבית, אם כי אין זו סטירה לティוריה לקבל גם עלייה - נזכיר כי רצוננו האמתי הוא בהשגת מינימום לפונקציה זו, אך התמונה המקורית אינה קיימת בידינו. לכן, הפונקציה אותה אנו מביאים למינימום (\underline{X}) הינה חיקוי לשגיאה האמיתית.

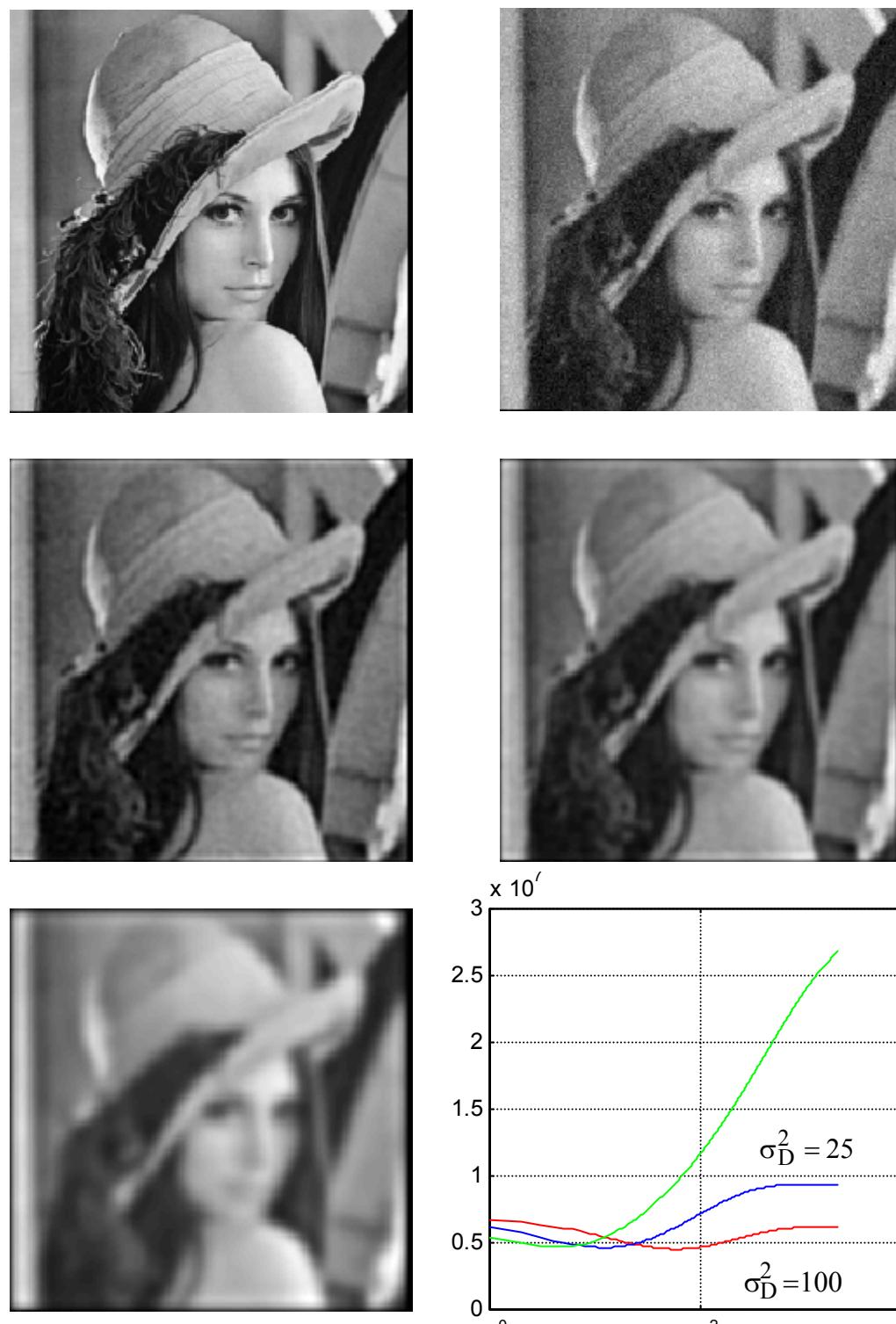
ציור 6.8 מראה תוצאות שחזור MAP ע"י הגישה האיטרטיבית למגוון ערכי $D^T\sigma$. אנו רואים כי עבור בחירת ערך גבוה התוצאה רועשת מיידי (טופעת ה- "צלצלים" - Ringing) בשל הגבר הרעש. עבור ערך נמוך התוצאה חלקה. בציור זה גם מובאת השגיאה האמיתית (מול התמונה האידיאלית) לשם השוואה. רואים כי לשולשת ערכי $D^T\sigma$ מתkowski אפקט דומה בו השגיאה יורדת באיטרציות הראשונות ואז עולה ומתייצבת על ערכה הסופי (עבור 2000 איטרציות!).

נציג כעת את הפתרון המבוסס על התמרת תדר, כמשמעותו מתייחסים לפתרון ה- MAP. נזכיר קודמים ומתחילת פרק זה, המטריצה T , המייצגת את התמרת הפורייה הדיסקרטית הדו-מיידית, מלכשנת מטריצות בלוק סיבוביות. במקרה שלנו, הן H והן D בעלות מבנה כזה בשל העובדה ייצוגם פעולה ליניארית וקבועה במקום. לכן :



ציור 6.7 - תהליכי שחזור איטרטיבי. הגרף העליון מראה את תהליכי ההתקנסות ע"י הצגת ערך הפונקציה ($f(\hat{X})$) כפונקציה של מספר האיטרציה. הגרף מתויחס לשולשה ערכים $\mu = 1e-2, 1e-3, 1e-4 \mu$. בגרף התיכון רואים עבור $\mu = 1e-3$ את השגיאה מול אותן המהו.

$$\begin{aligned}
 & \left[\mathbf{H}^T \mathbf{H} + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{D} \right] \hat{\mathbf{X}} = \mathbf{H}^T \underline{\mathbf{Y}} \\
 \Rightarrow & \left[\mathbf{T} \mathbf{H}^T \mathbf{T}^H \mathbf{T} \mathbf{H} \mathbf{T}^H + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{T} \mathbf{D}^T \mathbf{T}^H \mathbf{T} \mathbf{D} \mathbf{T}^H \right] \mathbf{T} \hat{\mathbf{X}} = \mathbf{T} \mathbf{H}^T \mathbf{T}^H \mathbf{T} \underline{\mathbf{Y}} \\
 \Rightarrow & \left[\mathbf{H}_T^H \mathbf{H}_T + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}_T^H \mathbf{D}_T \right] \hat{\mathbf{X}}_T = \mathbf{H}_T^H \underline{\mathbf{Y}}_T
 \end{aligned}$$



ציור 6.8 - תמונה מקור (למעלה בצד שמאל), תמונה מוקלקלת (למעלה בצד ימין), תוצאות שחזור MAP עם ערכים שונים של σ_D^2 - שורה שנייה בצד שמאל, 25 - שורה שנייה בצד ימין, 1 - שורה תחתונה בצד שמאל). כן מצורף גраф המתאר את השגיאה האמיתית כפונקציה של מספר האיטרצייה

הוקטור $\hat{\underline{X}}_T$ הוא התמונה הנדרשת במשור התדר. באופן דומה, $\hat{\underline{Y}}_T$ היא התמונה הנמדדת, גם בתדר. הקשר בין שתי תמונות אלו נתון ע"י המשוואה הניל' \underline{Y}_T , כשהמטריצות הkoplot (בשני האגפים) אלכסוניות! לכן, נוכל לקבל קשר סקלרי פשוט בין האיבר ה- k של $\hat{\underline{Y}}_T$ לאיבר ה- k של $\hat{\underline{X}}_T$:

$$\left[\mathbf{H}_T^H \mathbf{H}_T + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}_T^H \mathbf{D}_T \right] \hat{\underline{X}}_T = \mathbf{H}_T^H \underline{Y}_T \Rightarrow \hat{X}_T[k] = \frac{(\mathbf{H}_T[k])^*}{\left| \mathbf{H}_T[k] \right|^2 + \frac{\sigma_N^2}{\sigma_D^2} |\mathbf{D}_T[k]|^2} \underline{Y}_T[k]$$

קשר זה הוא בדיקת הקשר שהוצע כMSN Wiener, כך שהפעם בינו ממש את המסלן מתוך שיקולים תיאורטיים מבוססי שערוך MAP. עם זאת, לגישה אשרנדונה בסעיף זה יש יתרון חשוב - הכללה לשחזור ע"י סינון לא ליניארי תהיה קלה ופנוטה. נשים לב כי פעולה הלפלסיאן היאMSN מעביר גבוהים (HPF), ולכן התמרת התדר של פעולה הלפלסיאן הינה גורם שעולה עם התדר. לכן, במכנה בביטוי הניל' הגורם של הרגולריזציה יהיה זניח בתדרים נמוכים (כנדרש) ומשמעותי בתדרים גבוהים.

לסיכום בינויים - גישת ה- MAP הביאה לנו יכולת לבנות אתMSN Wiener מתוך שיקולים עמוקים יותר. אבל - לא קיבלו בינויים משחו חדש, והביצועים הצפויים (וזה ניכר מצייר 6.8) אינם מלבבים. היתרונו של שיטת ה- MAP הוא בכך שהוא ניתנת להכללה לסינון לא ליניארי, כפי שנראה בסעיף הבא.

6.5 שיפור פונקציית ההסתברות לתמונות

גילינו כי פעולה הרגולריזציה המבוססת על חלוקות גורמת להחלה יתר של התמונה. כאשר מקטינים את השפעת ה- Prior, אנו מתקרבים לביצועיMSN מהף, וזה כמובן גרווע. כך יוצאה שעלינו למצוא ערך בינויים אשר מנסה ליהנות מכל העולמות, ובפועל סובל מכל המחלות, והתוצאות מאכזבות. הסיבה לאי-ההצלחה כאן היא העובדה שהנחה לא נכונה במודע - שככל שביבה בתמונה تعدיף להיות חלקה. ברור שבאזורים חלקיים הדבר נכון, אך אזורים שפויות הם העיקרי בתמונה, וביהם פגעו. הפתרון לפיך הוא לשנות את חוק ההסתברות שהנחנו על תמונות. נזכיר כי הנחתנו המקורית היא:

$$P(\underline{X}) = C \cdot \exp \left\{ - \frac{[\mathbf{D}\underline{X}]^T [\mathbf{D}\underline{X}]}{2\sigma_D^2} \right\}$$

גישה אפשרית אחת לשפר את חוק ההסתברות זה היא להישאר במסגרת פילוג גאוסי, אך להרשות שנות תלויות מוקום. אפשרות אחרת היא לשנות את חוק הפילוג מגאוסי למשהו אחר נכון יותר. אנו נזכיר שתי גישות אלו, ונראה כי קיימים דמיון ביניהן.

נתחיל באפשרות של הרחבת ההסתברות הגאוסית לשונות תלויות מקום. ההסתברות הגאוסית במקורה הכללי ביותר לוקטור \underline{X} היא :

$$P(\underline{X}) = C \cdot \exp \left\{ -\frac{1}{2} [\underline{X} - \underline{M}]^T \mathbf{Q}^{-1} [\underline{X} - \underline{M}] \right\}$$

כאשר \underline{M} הוא הממוצע והמטריצה \mathbf{Q} היא מטריצת הקוריאנס. לפי הגירסה הקודמת בחרנו ממוצע אפס, ומטריצת שונות

$$\mathbf{Q}^{-1} = \mathbf{D}^T \mathbf{D}$$

באשר לממוצע, לא נציגו שינוי. חסרונו בולט במטריצת הקוריאנס היא הייתה בעלת מבנה בлок טופלי (כיוון ש- \mathbf{D} בעלת מבנה כזה). בambilים אחרים, מטריצת הקוריאנס מניחה כי בכל פיקסל יש אותה התנאות (מקביל לתכונת ה- Linear Space Invariance). על מנת לגוזן מעט את הבחירה, נציג

הנכשלה של מטריצת משקל אלכסונית \mathbf{W} , ונקבל

$$\mathbf{Q}^{-1} = \mathbf{D}^T \mathbf{W} \mathbf{D}$$

כאשר מטריצה זו תכיל ערכים זהים על אלכסונה, נתלבד עם הבחירה הקודמת, ובאשר הערכים ישתנו לאורך האלכסון, נקבל התנאות תלויות מקום. משערך ה- MAP הפעם יתקבל נקודת המינימום של הביטוי

$$f(\underline{X}) = [\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} [\mathbf{D}\underline{X}]^T \mathbf{W} [\mathbf{D}\underline{X}]$$

ומביטוי זה נוכל להסיק כיצד علينا לפעול בבחירה ערכי \mathbf{W} . הווקטור $\underline{D}\underline{X}$ היא תמונה נגזרת (כלשהי) של תמונה המקור \underline{X} . באזוריים שאמורים להיות חלקים נכהה חלוקת ע"י בחירת משקל 1. במקומות בהם צפואה שפה, נבחר ערך הקרוב לאפס. באופן זה נכהה חלוקות באופן תלוי מקום. באשר לבנייתה של \mathbf{W} , נוכל לחת את התמונה המוקלחת \underline{Y} , ולהשתמש בה כדי לקבוע בכל פיקסל האם ישנה שפה אם לאו.

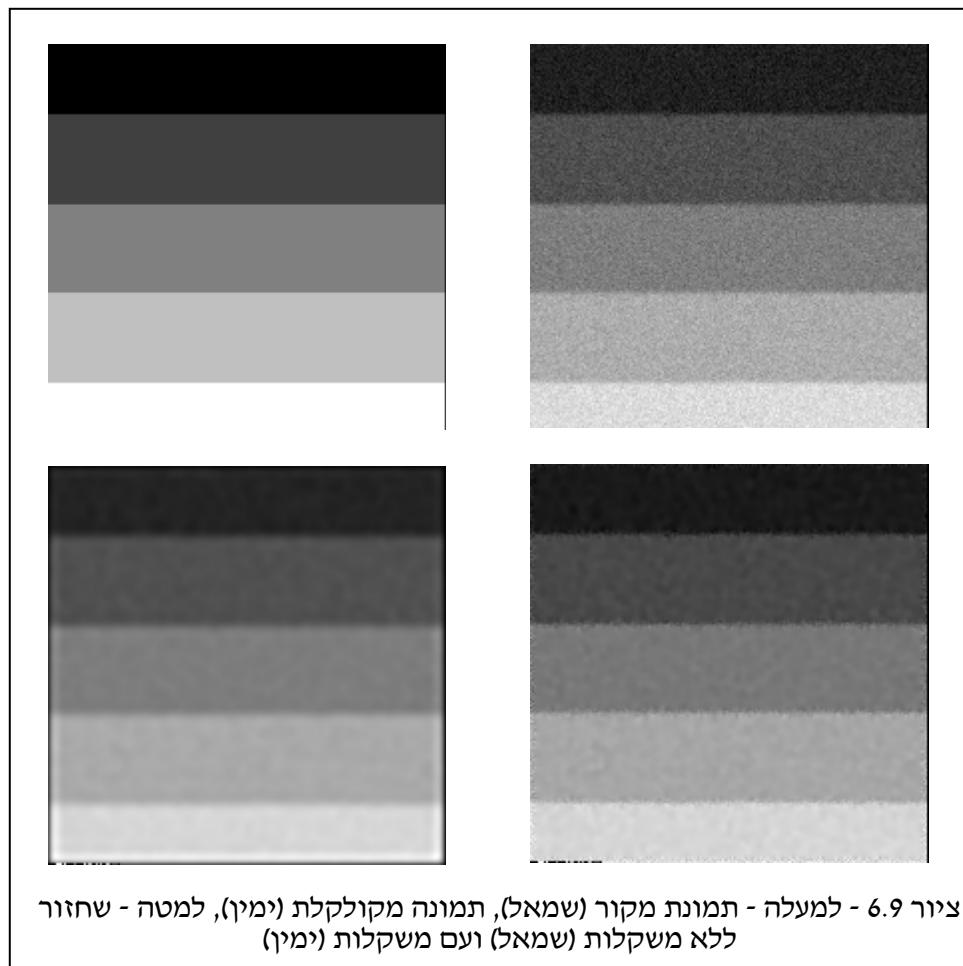
בניגוד למקורה הקודם, הפעם לא ניתן לעבוד בתדר כיוון שהחלק מהפעולות אין קבועות במקום. גישת האיטרציות פועלת באופן הבא :

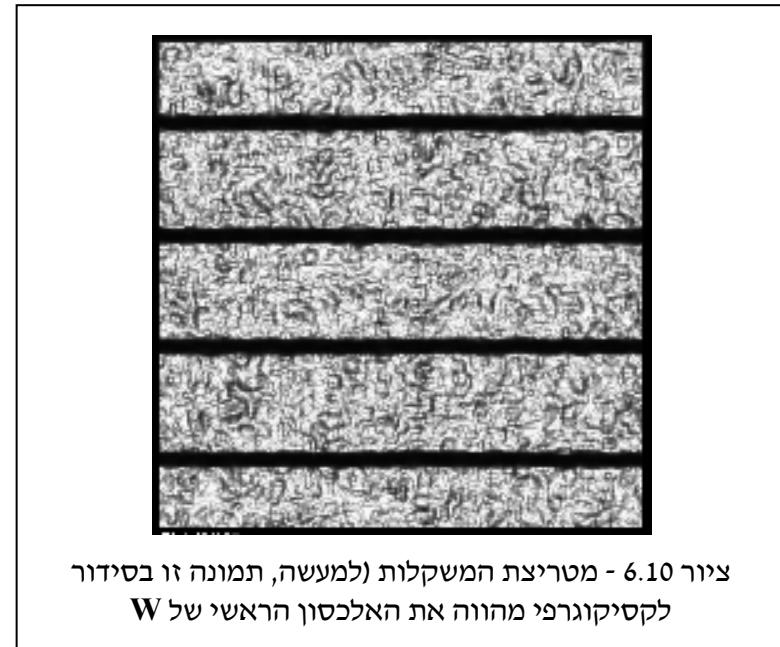
$$\hat{\underline{X}}_{k+1} = \hat{\underline{X}}_k - \mu \frac{\partial f(\underline{X})}{\partial \underline{X}} \Big|_{\hat{\underline{X}}_k} = \hat{\underline{X}}_k - \mu \left[\mathbf{H}^T (\mathbf{H}\hat{\underline{X}}_k - \underline{Y}) + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{W} \mathbf{D} \hat{\underline{X}}_k \right]$$

כך שההסכמה דומה מאד לקודמת, עם שינוי אחד - לאחר הפעלת מסנן גוזר על התמונה, יש לכפול את התמונה בתמונות מסוימתה שהינה בדיקת המשקלות שנבחרו.

לעתים ניתן אף לבחור ערך משקל שלילי במקומות בו צפואה שפה לשם השגת פועלה הפוכה של חידוד. בחירת הערך השלילי מסוכנת משום שאז יתכן שהפונקציה הניל' כבר אינה קמורה, ובעלת נקודות מינימה רבות. כמו כן הדבר מסכן את יציבות התהליך האיטרטיבי שתואר לעיל.

בציר 6.9 נראה תהליק השחזור עבור תמונה סינטטית. בציור מובאת התמונה המקורית, התמונה המקולקלת, שחזור עם משקל קבוע (הגישה הישנה) ושחזור עם משקל תלו依 מקום. שגיאת השחזור בני המקרים היא - ללא משקלות - $MSE=27.06$, ועם משקלות - $MSE=21.001$. ערכים אלו התקבלו ע"י החסרת התמונה המקורית מוצאת השחזור וסיכום ריבועי ההפרשים. לצורך זה הושמטה רצועה של 5 פיקסלים מڪנות התמונה, מכיוון שמיילא שם פועלותינו לא מוצלחות בכל שיטה. רואים שימוש שיפור משמעותי ע"י הכנסת המשקלות לשחזור. חשוב להבהיר כי השיפור הפוטנציאלי טוב אף יותר כיון שניתן היה להגדיל את המקדם של הרגולרייזציה ולכפות חלקות חזקה יותר באזורי החלקים, כך שפוטנציאלית גישה זו לא מוצאת עד תום. ציר 6.10 מראה את מטריצת המשקל שנבחרה.





המשקלות נבנו ע"י סידרת הצעדים הבאה :

1. לקיחת התמונה המקולקלת והחילקה ע"י מסנן 5 על 5. צעד זה נועד לדכא את הרעש בתמונה לפני חישוב נגורות.
 2. חישוב הנגורות האופקית והאנכית ע"י מסנן מהצורה $[1 \ 0 \ 0 \ 1]$ (בשני הצדדים), וחישוב נורמת הגרדיאנט (סכום ריבועי שני הנגורות). כעת יש לנו תמונה בה ערך גובה מעיד על שפה, וערך נמוך מתיחס לאזור חלק.
- בנייה משקל בכל נקודה ע"י הנוסחה $- [Grad+10]/[Grad+10] -$. כך קיבלנו כי כאשר $Grad$ הוא 0, המשקל הוא 0.99, ואשר הוא נוטה לאינסוף המשקל הוא -0.01.

עד כה תיארנו כיצד נכליל את הנחת הגאוסיות שלנו ע"י הכנסת מומנט סדר שני מורכב יותר. נראה כעת גישה אחרת שימושותה ברירה מהנחת הגאוסיות אל חוק פילוג אחרים. נזהר לבסיס - אנו לוקחים את התמונה \mathbf{X} ומבצעים עליה פעולה גזירה כלשהי \mathbf{D} . נסתכל על פיקסל מסוים ונניח כי ערך הנגורות הוא g . בשיטה הקודמת, g עולה בRibou והתוצאה היא "מחירו" של פיקסל זה. מדובר עלות בריבוע ולא בריבועית? מדוע לא ערך מוחלט? וניתן כմובן להציג מגוון אפשרויות אחרות.

נססה להבין מה המשמעות של בחירות שונות של החזקה בדרך הבאה. נניח שרצוננו ליציר פונקציה $[k]$ (חד-מיידית) חלקה וディיסקרטית אשר תעבור מ-0 ל-1 ידוע לנו כי הפונקציה ב- $k=0$ היא f , וב- $k=10$ היא g . נציג שתי חלופות - לפונקציה f ונבחן את מהירן. חלופה ראשונה תהיה $-f$ עליה $k=0$, וב- $k=10$ מדרגות בגובה 0.1 כל אחת, וחלופה שנייה בה f עליה ב- $k=5$ בverts אחת מ-0 ל-1. נניח כי פעולה הנגורות היא פשוט הפרשים של דגימות סמכות. לפי השיטה המקורית קיבל

$$\sum_{k=1}^{10} (f_1[k] - f_1[k-1])^2 = 0.1^2 + 0.1^2 + \dots + 0.1^2 = 10 \cdot 0.1^2 = 0.1$$

$$\sum_{k=1}^{10} (f_2[k] - f_2[k-1])^2 = 0^2 + \dots + 1^2 + \dots + 0^2 = 1$$

לכן ברור כי תמחור כזה יעדיף עליה הדרגתית. נניח כעת כי החלפנו את החזקה בערך מוחלט. נקבל:

$$\sum_{k=1}^{10} |f_1[k] - f_1[k-1]| = |0.1| + |0.1| + \dots + |0.1| = 10 \cdot 0.1 = 1$$

$$\sum_{k=1}^{10} |f_2[k] - f_2[k-1]| = 0 + \dots + |1| + \dots + 0 = 1$$

והפעם קיבלונו כי מחיר שני הגישות זהה. נמשיך בмагמה זו ונציג שורש על הערך המוחלט ונקבל:

$$\sum_{k=1}^{10} |f_1[k] - f_1[k-1]|^{0.5} = |0.1|^{0.5} + |0.1|^{0.5} + \dots + |0.1|^{0.5} = 10 \cdot 0.1^{0.5} = 3.16$$

$$\sum_{k=1}^{10} |f_2[k] - f_2[k-1]|^{0.5} = 0 + \dots + |1|^{0.5} + \dots + 0 = 1$$

ובמקרה זה הפונקציה تعدיף עליה חזקה על-פני עליה מתונה.

מדוע כל זה חשוב? ובכן, אנו מניחים שכאשר מבוצעת הפעולה \underline{D} , ערכיהם (ערך מוחלט) בסביבת האפס מתייחסים לאזור נמוך, ולכן הגישה הריבועית יפה להם, ככלומר נרצה שמעבר בין רמות יקרה בהדרגה. לעומת זאת, מערך מסוים ואילך נראה שמדובר בשפות, ולכן נרצה שמעבר בין רמות יקרה באופן חד. לכן, נציג באופן כללי כי פונקציית המחיר לחלוקת התמונה תהיה:

$$\sum_k \rho(\underline{D}\underline{x}|_k)$$

כלומר, נבצע כמוקדם פעולה גזירה \underline{D} , וכל ערך בתמונה התוצאה יוכנס לפונקציה $\rho(x)$. אם x^2 הוא לגרסת המקורית. אם נבחר (x) שיתנהכג כפונקציה ריבועית סביב הראשית, וימתו את עלייתנו מנקודה מסוימת, הרי שנשיג את שרצינו. ישנו מספר אפשרויות - נציג שתיים מהן:

$$\rho(x) = \frac{x^2}{x^2 + \alpha} \quad \text{or} \quad \rho(x) = \begin{cases} -T(2x + T) & x < -T \\ x^2 & -T \leq x \leq T \\ T(2x - T) & x > T \end{cases}$$

הבחירה הראשונה ידועה בשם פונקציית קושי, והשנייה בשם פונקציית Huber-Markov. שימוש בפונקציות (x) מוכרת בטטיטיקה בשם שערוך רוביטי (Robust Estimation), והפונקציות (x) זכו לשם M-functions.

הבעיה עם גישת השערוך הרוביטי היא שעבור בחירה של פונקציות שעולות בקצב נזוק מליניארי (כמו הדוגמה הראשונה מבין השתיים במשואה הקודמת), הפונקציה שמובאות למבאים כבר אינה קמורה, ואז מתאפשרות תופעות של נקודות מינימנה רבות, ובעיות התכנסות למינימה הגלובלית בהתאם לכך. בשל כך, בחירת פונקציית ערך מוחלט $|x| = (x)$ זוכה לפופולריות רבה.

עבור שימוש בפונקציית המחיר החדשנית, הפונקציה הכוללת שישי להביא למינימים היא:

$$f(\underline{X}) = [\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{D}\underline{X}]$$

כאשר לפי הסימון שלנו, \mathbf{C} על וקטור הוא פעלת (x) על כל אחד מאיבריו, וסיכון התוצאות. מתוך ביטוי זה ניתן לכת Achorah ולקבוע כי בעצם הנחנו כי פונקציה ההסתברות של מרחב התמונות נתונה עלי:

$$P(\underline{X}) = C \cdot \exp \left\{ -\frac{\rho[\mathbf{D}\underline{X}]}{2\sigma_D^2} \right\}$$

פונקציות פילוג בעלות מבנה זה קרויות פילוג Gibbs. המשותף לכלן הוא ההסתברות ל- \underline{X} יחסית הפוך דרך אקספוננט למידת האנרגיה שלו, והמעיריך הוא סוג של אנרגיה.

אם יש קשר בין גישה זו לגישה הקודמת של בחירת מטריצת משקל? לשם גילוי קשר כזה, נכתוב את משוואת האיטרציה שיש להפעיל לשם מינימיזציה של הביטוי שנכתב לעלה:

$$\hat{\underline{X}}_{k+1} = \hat{\underline{X}}_k - \mu \frac{\partial f(\underline{X})}{\partial \underline{X}} \Big|_{\hat{\underline{X}}_k} = \hat{\underline{X}}_k - \mu \left[\mathbf{H}^T (\mathbf{H}\hat{\underline{X}}_k - \underline{Y}) + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \rho'[\mathbf{D}\hat{\underline{X}}_k] \right]$$

משמעות משווה זו היא שיש לגזר את \underline{X} כמוקדם, ואז להעביר כל פיקסל דרך נגזרת הפונקציה (x) , ואת התוצאה שוב לגזרו. לשם דוגמה - עבור $x(x) = 0.5x^2$ ברור כי $x'(x) = x$, ואז קיבל כי פעולה ה- $'$ משaira את הערכים כמוותיהם. עבור $|x| = (x)$ הנגזרת תהיה $x(x) = \text{sign}(x)$, כך שכל פיקסל לאחר נגזרת יושווה לאפס - שליליים יהפכו ל- -1 , וחוביים ל $+1$. בהשוואה לשווה האיטרציבית במקרה של השימוש במשקל, שני הגורמים הנ"ל מתחלפים:

$$\mathbf{WD}\hat{\underline{X}}_k \Leftrightarrow \rho'[\mathbf{D}\hat{\underline{X}}_k] \quad (\text{גישת } \rho(x)) \quad (\text{משקל})$$

בשני המקרים התהילה מתחליל זהה - התמונה $\hat{\underline{X}}_k$ עוברת דרך פעולה הגדירה \mathbf{D} . בשני המקרים הפעולה הבאה היא פעולה על כל פיקסל בנפרד. האחת כופלת במשקל כלשהו, והאחרת מעבירה דרך פונקציה כלשהי. לכן, אם בגישה המשקלות לא היינו משתמשים במטריצת משקל קבוע לאורך כל התהילה האיטרטיבי, אלא משנים אותה בכל איטרציה להיות

$$\mathbf{W}|_{\text{pos}} = \frac{\rho'[\mathbf{D}\hat{\underline{X}}_k]|_{\text{pos}}}{[\mathbf{D}\hat{\underline{X}}_k]|_{\text{pos}}}$$

הינו מקבלים אותה תוצאה בדיק. מה שימושות בחירה זו? לדוגמה עבור $x^2 = 0.5x$ יתקבל כי \mathbf{W} אינה אלא 1 לכל הפיקסלים. עבור $|x| = 0$ נקבל כי המשקל יחס ההפוך לגודל הנגורת בערכה המוחלט, כיוון :

$$\mathbf{W}|_{\text{pos}} = \frac{\rho'[\mathbf{D}\hat{\underline{X}}_k]|_{\text{pos}}}{[\mathbf{D}\hat{\underline{X}}_k]|_{\text{pos}}} = \frac{\text{sign}(\mathbf{D}\hat{\underline{X}}_k|_{\text{pos}})}{|\mathbf{D}\hat{\underline{X}}_k|_{\text{pos}}} = \frac{1}{|\mathbf{D}\hat{\underline{X}}_k|_{\text{pos}}}$$

אם נחשב על מה שקיבלנו, הרי שלו יכולנו לקבוע את מטריצת המשקלות לא לפי \underline{Y} אלא לפי \underline{X} הינו צפויים לקבל תוצאה טובה יותר. לכן, כתחליף לחישוב \mathbf{W} מבוסס \underline{Y} , מציעה התוצאה שהתקבלה שנחשב את \mathbf{W} בכל איטרציה על סמך תוצאה האיטרציה הקודמת $\hat{\underline{X}}_k$, וזה קרוב יותר לשימוש ב- \underline{X} המקורי.

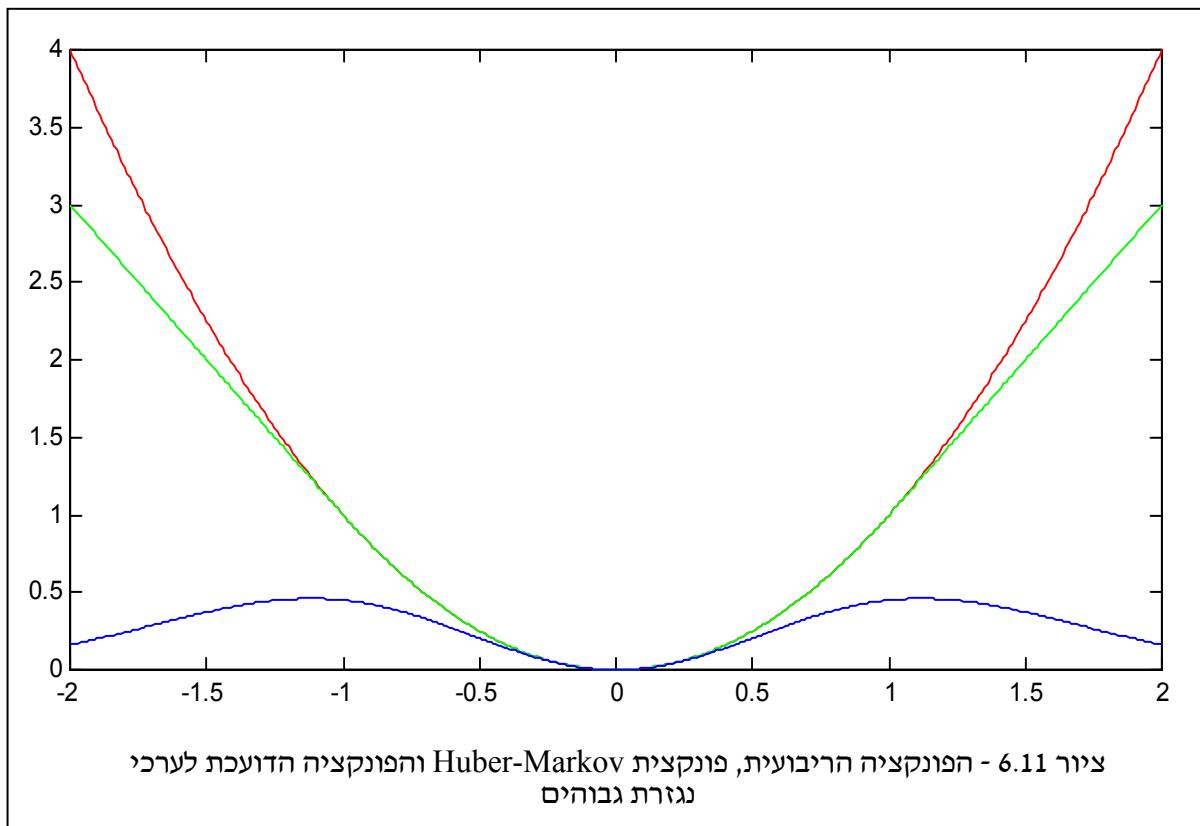
רעיון מעניין הוא הרעיון הבא : מדוע לא לעשות שימוש בפונקציות שאיפילו יורדות עבור ערכים גבוהים. למשל, נציג את הפונקציה הבאה :

$$\rho(x) = x^2 \cdot \exp\left\{-\alpha x^2\right\}$$

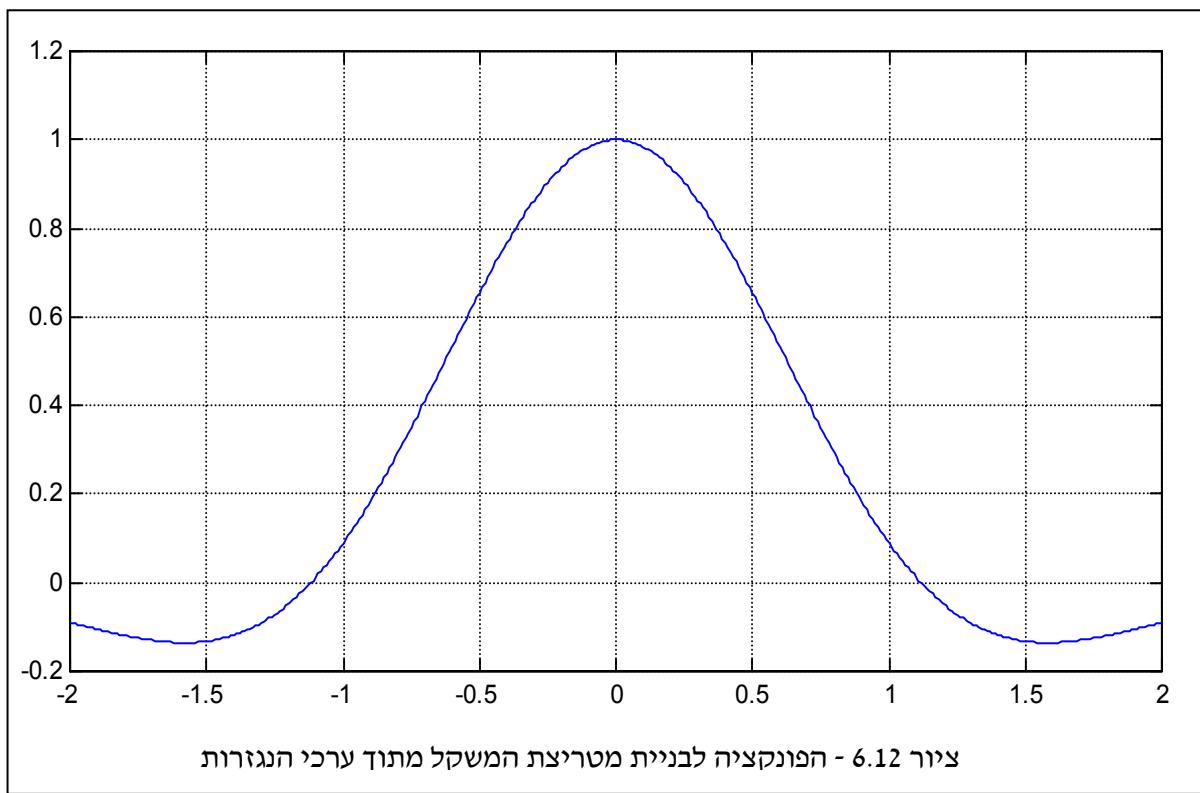
בחירה זו תנתנה כפונקציה ריבועית לערבים נומקיים, אך לאחר עלייה מסוימת תתקבל ירידה לאפס עבור ערכים גבוהים. ציור 6.11 מראה את הפונקציה הריבועית, פונקציית Huber-Markov, ובחרית פונקציה חדשה זו. בכל פונקציה שעולה מונוטונית אנו גובים מחיר שלו (שולוי - מתייחס לנגורת הפונקציה) חיובי.

עבור פונקציות יורדות נועד היוצרים של שפות אם נגורתם גדולה. עבור הבחירה הניל, המשקל שמקביל לשימוש יהיה :

$$w(x) = \frac{\rho'(x)}{x} = 2 \exp\left(-\alpha x^2\right) \left[1 - \alpha x^2\right]$$



עבור הגרף הנ"ל מתקבל כי חישוב משקל נעשה לפי הגרף הבא, המציג כי רעיון בחירת ערכים שליליים אינו נטול בסיס :



6.6 בעיות היפוך אחריות והקשר לעויהת השחזרה הקלאסית

אם נסכם את שראינו עד כה, הרי שעבור בעיות בעלות המבנה $\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$, בהן \mathbf{H} מערכת ליניארית כלשהי ידועה, \underline{X} תמונה רצiosa שלא ניתנת למדידה, \underline{Y} תמונה אחרת שנמדד במקומה, ו- \underline{N} וקטור רעש אקראי נאוסי לבן בעל עצמה Σ_N , הצעת שחזרה לא רעה ל- \underline{X} היא מינימלית של הביטוי:

$$f(\underline{X}) = [\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho [\mathbf{D}\underline{X}]$$

כאשר \mathbf{D} אופרטור גזירה ליניארי כלשהו, הפונקציה (x) היא פונקציה שערוך רובטית שנבחרת מראש, ו- σ_D פרמטר השולט בעוצמת החלוקות שנרצה לכפות.

עד כה השתמשנו בתוצאה הניל' לשם פתרונה של בעיית השחזרה הקלאסית של בניית שחזרה לתמונה מוטשטשת ורואה. נשאלת השאלה - האם ניתן לעשות שימוש בתוצאה הניל' בעיות שחזרה אחריות? התשובה היא כן, והוא נציג מספר בעיות כאלה.

האפשרות הראשונה היא ניקוי תמונה מרעש. במקרה זה, המודל נותר כמו שהוא למעט העובדה ש- $\mathbf{I} = \mathbf{H}$. לכן, במקרה זה, علينا להביא למינימום את

$$f(\underline{X}) = [\underline{Y} - \underline{X}]^T [\underline{Y} - \underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho [\mathbf{D}\underline{X}]$$

וכמוון התחליק שמתקיים פשוט יותר בהשוואה לזה המקורי (כיוון שאנו פטורים מהפעלת \mathbf{H} בתהליך האיטרטיבי).

אפשרות שנייה היא הגדלת תמונה. בהינתן תמונה \underline{Y} , נשאל כיצד להגדילה פי z בכל ציר. אפשרות אחת טריוויאלית היא שכפול פיקסלים, כך שכל פיקסל יוחלף בבלוק של z על z פיקסלים זהים לו.YSIS זו מניבה תמונה באיכות ירודה למדי. בפרק 3 בו דנו בשחזרה מדגימה תיארנו הגדלה ע"י אינטרפולציה ביליניארית. התוצאה בגישה זו טוביה יותר, אך עדין, באזורה שפה בתמונה קיבל מריחה לא נעימה של המעבר, ושוב מתקיים תמונה באיכות ירודה.

אפשרות חדשה היא להתייחס לעויהת ההגדלה כבעית שחזרה. אנו נניח כי הייתה תמונה גדולה ויפה \underline{X} אשר עברה תהליך של טשטוש \mathbf{H} ודילול \mathbf{S} לקבלת התוצאה הירודה \underline{Y} . לכן, העיה אותה נרצה לפטור תהיה מינימלית של:

$$f(\underline{X}) = [\underline{Y} - \mathbf{SH}\underline{X}]^T [\underline{Y} - \mathbf{SH}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{D}\underline{X}]$$

באופן זה עשינו שימוש בעובדה שדילול היא פעולה ליניארית הניתנת לתיאור כמכפלת מטריצה בוקטור. מטריצה זו מלכנית ולא ריבועית - אם התמונה בכניםתה היא 100 על 100, ובמוצאה 25 על 25, גודל המטריצה יהיה 625 שורות על 10,000 עמודות. פעולה הדילול היא פעולה ספרטילית וניתן לפיק'ס לכוטבה כמכפלת קרוונקר בין שתי פעולות חד-מימדיות.

באשר בדרך בחירת האופרטורים **H** ו- **S** להגדלת תמונה, **S** יקבע ישירות לפי יחס ההגדלה אותו רצוננו להשיג. **H** יכול להיבחר כ- **I** אם אין רצוננו באפקט של חידוד. בחירת קיומו של טשטוש תגרום לתוצאה להיות חדה יותר. במקרים רבים ניתן לקבוע מדו"ע התמונה קטנה מידי מתוך שיקולים פיזיקליים, ואז **H** יכולה להיבחר בהתאם לדרך היוצרה של התמונה.

בעיה אחרת לגמרי אשר גם היא נופלת לקטגוריה של משפחת הביעות הנ"ל היא בעיית הסופר-רזולוציה. נתונה לנו מצלמה ירודה באיכותה ורצוננו לצלם יעד כלשהו. היעד מיוצג בתמונה אידיאלית **X**, ואילו מה שהמצלמה מניבה היא התמונה **X** לאחר סידרת פעולות מסוימות - הזזה, טשטוש, דילול, ולבסוף רעש. במקומות לצלם פעם אחת אנו נצלם **M** פעמים, בהזוזות שונות ולבן נקבל:

$$\begin{aligned} \underline{Y}_1 &= \mathbf{SHF}_1 \underline{X} + \underline{N}_1 \\ \underline{Y}_2 &= \mathbf{SHF}_2 \underline{X} + \underline{N}_2 \\ &\vdots \\ \underline{Y}_M &= \mathbf{SHF}_M \underline{X} + \underline{N}_M \end{aligned}$$

הזוזת תמונה (כל עיוות גיאומטרי שהוא!) ניתנת לתיאור כפעולה ליניארית ולבן כמטריצה. זאת כיון שאם ניקח צירוף ליניארי של שתי תמונות ונזיז את שילובן, יהיה זה כמו להזיז כל אחת מהן ואז לצרפן ליניארית.

ב- **M** הפעמים בהם צילמנו עשינו שימוש באותו מצלמה ירודה באיכותה, ולבן **H** ו- **S** זהים לכל התמונות שנוצרו. לעומת זאת, תוכן כדי צילום זזנו ולבן האוריינטציה הגיאומטרית בין התמונות השתנתה. זו מיוצגת כאן ע"י האופרטורים הליניאריים **F**. כך התקבל סט תמונות **Y**. נוכל לנקות את כל השוואות הנ"ל ולצרפן לאחת לקבלת:

$$\begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_M \end{bmatrix} = \begin{bmatrix} \mathbf{SHF}_1 \\ \mathbf{SHF}_2 \\ \vdots \\ \mathbf{SHF}_M \end{bmatrix} \underline{X} + \begin{bmatrix} \underline{N}_1 \\ \underline{N}_2 \\ \vdots \\ \underline{N}_M \end{bmatrix} = \tilde{\mathbf{H}} \underline{X} + \tilde{\underline{N}}$$

ושוב, אלו רואים אותו מבנה. לכן, שחזור תמונה הסופר-רזולוציה יעשה ע"י מינימיזציה של הביטוי

$$f(\underline{X}) = [\underline{Y} - \tilde{\mathbf{H}}\underline{X}]^T [\underline{Y} - \tilde{\mathbf{H}}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{D}\underline{X}] = \sum_{k=1}^M \|\underline{Y} - \mathbf{SHF}_k \underline{X}\|^2 + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{D}\underline{X}]$$

בדומה לנאמר בעבר הגדלת תמונה, על מנת להפעיל תהליך סופר-רזולוציה, علينا לדעת את האופרטורים המעורבים. ידעת אופרטורי התנועה כרוכה שתהליכי שערוך תנווה בין התמונות - נושא שיידן בפרקם מתקדמים יותר. כאמור, \mathbf{S} יכול להיקבע ישירות מתוך יחס ההגדלה הרצוי. \mathbf{H} ייקבע ע"י הבנת מאפייני המצלמה בה נעשה שימוש, או פשוט בבחירה שריורותית לפי מידת החידוד שרצוננו להשיג בתוצאה הסופית.

פרק 7 – התמורות דיסקרטיות בדו-מימד

143	פרק 7 – התמורות דיסקרטיות בדו-מימד	
143	7.1 התמורת ליניארית - בסיס
148	7.2 התמרת ה-DFT הדו-מימדית ותכונותיה
151	7.3 התמרת ה-DCT הדו-מימדית ותכונותיה
156	7.4 התמרת Hadamard דו-מימדית ותכונותיה
160	7.5 התמרת Haar ומבוא ל-Wavelet
168	7.6 התמורות מבוססות תוכן תמונה
171	7.7 Over-Complete

כבר פגשנו במסגרת פרקים קודמים דיון בתמורות פוריה (רציף ודיסקרטי), ואנו שבים במסגרת פרק זה לנושא זה בשל חשיבותו הרבה. הרעיון הבסיסי שלולה אותנו לכל אורכו של הדיון כאן הוא שההינתן תמונה בגודל $N \times M$ כמטריצת ערכי רמות אפור, מבליל ששמנו לב נקופה עליינו ייצוג מאוד מסויים של התמונה מעל הבסיס הטוריוני-אליגרי $\sum_{j,k} a_{jk} e_j^* e_k$ (תמונה שכולה אפס למעט האיבר ה- a_{jk} השווה 1). למה שבחירה זו תהسب טובה יותר מאשר אחרת? מדוע לדבוק בה, למעט פשוטותה? אנו נראה כי אכן קיימות התמורות לבסיסים אחרים, עימם ניתן לפחות בעיות, לחסוך בחישובים ועוד.

עוד עבדה שתறחף מעל רוב פרק זה כאקסומה היא שאנו מעוניינים לעבוד עם התמורות ליניאריות בלבד, בהן התמונה תיזוג כצירוף ליניארי של תמונות בסיס כלשהו. ברור כי בחירה זו מצמצת באופן דרמטי את מקבץ האפשרויות, אך מאידך, מבטיחה לנו טיפול מתמטי קל יותר. אנו נתחיל את הדיון בבניית הקשר בין בחירת בסיס אלטרנטיבי וההתמרה ליניארית.

7.1 התמורות ליניאריות - בסיס

התמרה ליניארית קונציאונלית הינה פעולה הניתנת לתיאור כמכפלה במטריצה. אם אנו עוסקים בתמונה \underline{f} בגודל $N \times M$, נוכל לסדרה לוקטור אורך \underline{NM} ואז תמונה התמרה \underline{U} נתונה ע"י:

$$\underline{U} = A \underline{V}$$

כשאנו מניחים ש- A מטריצה ריבועית בגודל $[MN \times MN]$. במקרה זה נקבל כי תמונה התמורה זהה בגודלה לתמונה המקור. אנו נדרש כי A תהיה הפיכה על מנת לאפשר התמורה הפוכה מ- \underline{U} ל- \underline{V} . במקרים רבים אנו נלק' צעד נוסף בPGA הפשטת התמורה ונשתמש רק בתמורות שביצוען קדימה- אחרת קל עד טריויאלי - התמורות יוניטריות, עבוריון מתקיים:

$$A^{-1} = A^H = [A^T]^*$$

הזכירנו כי התמורה מהוות ייצוג עפ"י בסיס חדש. מהו הקשר בין בסיס חדש זה ומהמטריצה הנבחרת A ? נניח כי נתון לנו סט תמונות בסיס \underline{b}_k איתו אנו רוצים לעבוד. נניח בשלב זה כי מספר פונקציות הבסיס הוא בדיק MN (מספר הדגימות בתמונה המקור \underline{V}), וכן נניח כי בסיס זה פורס את המרחב. لكن נציג את \underline{V} כצירוף ליניארי של פונקציות בסיס אלה ונקבל:

$$\underline{V} = \sum_{k=1}^{MN} u_k \underline{b}_k = B \underline{U}$$

כאשר B היא מטריצה המכילה את \underline{b}_k כעמודותיה, והוקטור \underline{U} הוא מקדמי הצירוף הליניארי. כיוון שהבסיס שלים (פורס את המרחב) ובעל MN אלמנטים, קיים רק צירוף ליניארי אחד שייתן שיוויון. צירוף זה מתקבל ע"י:

$$\underline{U} = B^{-1} \underline{V} = A \underline{V}$$

ומכיון שהבסיס שלים נקבל כי המטריצה B לא סינגולרית והפיכה. למעשה, קיבלנו כי התמורה שיש לבצע על \underline{U} לקבלת \underline{V} היא ההיפוך של B , כלומר B מייצג את התמורה הפוכה. קשר פשוט יותר יתאפשר אם נעשה שימוש בוקטורים \underline{b}_k אורטונורמליליים, כלומר כאלה שמקיימים

$$\forall k, j \quad \underline{b}_k^T \underline{b}_j = \delta(k - j)$$

או יתאפשר כי המטריצה B יוניטרית, ולכו

$$\underline{U} = B^{-1} \underline{V} = B^H \underline{V} = A \underline{V}$$

ויצא כי שורותיה של המטריצה A מכילות (לאחר פעולת צמוד קומפלקס) את הוקטורים המהווים את הבסיס החדש.

להתמרה יוניטרית תכונה חשובה אותה אנו מכירים בשם "משפט פרסלבל" עבור המקרה הפרט של התמורה פוריה. אם נחשב לרוגע על התמורה יוניטרית כפעולה גיאומטרית המקיימת וקטור ממנו וקטור חדש, אזי, התמורה יוניטרית לא משנה את עצמתו (הנדצתה בnormה) של הוקטור בתהליך, אלא פועלת כסיבוב ברב מימד. קל לראות זאת אלגברית ע"י הקשר :

$$\underline{U} = \underline{A} \underline{V} \Rightarrow \|\underline{U}\|^2 = \|\underline{A} \underline{V}\|^2 = \underline{V}^H \underline{A}^H \underline{A} \underline{V} = \underline{V}^H \underline{V} = \|\underline{V}\|^2$$

\uparrow
 $\underline{A}^H \underline{A} = I$

במקרים מסוימים יש עניין בהתמורות בהן יש יותר פונקציות בסיס מאשר נקודות בתמונה. בסיסים אלו נקראים Over-Complete. במקרה זה, המטריצה B המכילה את וקטורי בסיס אלה בעמודותיה, ולכן יש לה יותר עמודות משורות, ולכן, הוקטור \underline{U} יהיה ארוך יותר מ- \underline{V} . במקרה זה הקשר המתאר את ההתמרה ההפוכה הוא :

$$\underline{V} = \sum_{k=1}^{MN} \underline{u}_k \underline{b}_k = B \underline{U}$$

התמרה - תהליך חישוב \underline{U} מתוך \underline{V} - מרכיבת יותר הפעם. כיוון שזו מערכת ליניארית בת NM משוואות ויותר מ- MN געלמים (איברי \underline{U}), ישנו אינסוף אפשרויות של פתרונות \underline{U} , ונשאלת השאלה את מי מהם לבחור. נוכל להביע את ה- \underline{U} בעל הנורמה המינימלית, \underline{U} בעל חלקות מירבית בין איבריו, \underline{U} הדليل ביותר האפשרי ועוד. אנו עוסקים לkrarat סופו של הפרק בבעיה זו ונציג לה פתרונות מעניינים. נקודה מעניינת היא שלמרבית המקרים לא ניתן יהיה להציג את בנית \underline{U} מתוך B ו- \underline{V} כפולה ליניארית, דהיינו מכפלה במטריצה.

נזכיר את התמורות ריבועיות בהן מספר הבסיסים כמספר הדגימות בתמונה הנתונה. ראיינו תהליכי בו התחילו עם התמרה כלשהי, עברנו להתמורות ליניאריות, ופישטנו הלה ע"י שימוש בהתמורות יונטריות. הפשתה אפשרית אחרת היא השימוש בהתמורות פרידוט (ספרביליות). בהתמורות אלו, במקומות העסוק ב- MN תמונות בסיס דו-מימדיות, אנו נניח כי יש לנו שתי משפחות של וקטורים דו-מימדיים :

$$\{\underline{c}_k\}_{k=1}^M \in \Re^M , \quad \{\underline{d}_k\}_{k=1}^N \in \Re^N$$

כלומר, M וקטורים באורך M כל אחד, ו- N וקטורים אחרים באורך N כל אחד. שני סטים אלו נבנה הבסיס הדו-מימי לפיה :

$$b_{[j,k]} = \underline{c}_j \underline{d}_k^T \Rightarrow b_k = \underline{c}_j \otimes \underline{d}_k$$

הbianioי הראשוני בונה את \underline{b}_k כתמונה בגודל M על N ע"י הכפלת וקטור عمودה בוקטור שורה. הביאioי השני השקול לו עושים שימוש בסידור לקסיקוגרפי לפי עמודות, ויוצר את \underline{b}_k ע"י מכפלת קרוונקר. את מלאה הבסיס B ניתן לייצר ע"י הפעולה:

$$B = C \otimes D$$

נראה דוגמה פשוטה להבנה הרעיון זהה.

דוגמה 7.1: נתונים לנו שני סטים של בסיסים חד-מימדיים באורך 3 ו- 4 בהתאם. בסיסים אלו מתוארים בעמודות המטריצות C ו- D :

$$C = \begin{bmatrix} 1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} \quad D = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

בננה בעזרתו שני בסיסים אלו בסיס לתמונות דו-מימדיות בגודל 3 (שורות) על 4 (עמודות). יהיו לנו 12 תמונות בסיס בנות 12 אלמנטים כל אחת. לדוגמה, וקטור בסיס ראשון יהיה:

$$\underline{b}_1^T = [1 \ -1 \ 0] \otimes [1 \ 0 \ 1 \ 0] = [1 \ -1 \ 0 \ 0 \ 0 \ 0 \ 1 \ -1 \ 0 \ 0 \ 0 \ 0]$$

באופן דומה, בניית מלאה הבסיס B תיתן:

$$B = C \otimes D = \left[\begin{array}{ccc|ccc|ccc|ccc} 1 & 1 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & -1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & -1 & 1 \\ \hline 1 & 1 & 0 & -1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 & 0 & 1 & -1 \end{array} \right]$$

מדוע בסיס פריד (ספרבילי) נחשב פשוט יותר? בשל התכונה הבאה (המתוארת על ההתרמה ההפוכה):

$$\underline{V} = B \underline{U} = [C \otimes D] \underline{U} = [C \cdot U \cdot D]_{CS}$$

משווה זו אומרת שבמוקם לכפול את הוקטור \underline{U} במטריצה התמורה המלאה, התמורה ניתנת להשגה ע"י פעולה על התמונה במבנה דו-מימדי U , שכפולים משמאלי במטריצה C ומימין במטריצה D . הפעלת C משמאלי מבצעת התמורה חד-מימדית על העמודות של U , ולאחר מכן הפעלת D מימין משלימה ביצוע התמורה חד-מימדית על השורות של התוצאה. השוואת כמהות המכפלות תחשוף מדוע זה כה אטרקטיבי: בגישה הישירה علينا לכפול מטריצה בגודל $[MN \times MN]$ בוקטור בגודל MN . לכן, סך המכפלות הוא M^2N^2 . בגישה הספרטילית, הפעלת C ב- U כרוכה ב- N^2 מכפלות (מטריצת התוצאה היא בגודל M על N , וכל איבר בה חושב ע"י M מכפלות), והכפלה נוספת של התוצאה ב- D דורשת M^2N מכפלות נוספת (מטריצת התוצאה היא בגודל M על N וכל איבר בה דורש N מכפלות). היחס בין מכויות החישובים הוא:

$$\text{Ratio} = \frac{M^2N^2}{NM(M+N)} = \frac{MN}{M+N}$$

לדוגמה, עבור $M=N=1$, יש יתרון לגישה הישירה (היחס הוא 0.5), עבור $M=N=2$ היחס הוא 1 - שתי הגישות זהות בmorכבותן. עבור $M=N=3$ היחס הוא 1.5, ולמעשה, מערך זה ומעלה יש יתרון הולך ומתחזק לטובת ביצוע התמורה במבנה ספרטילי על פני ביצועה ישיר. עבור הבחירה $N=M$ קיבל כי היחס הנ"ל הוא פשוט $0.5N$ לעומת N^2 ליניארית עם N .

עד כה עסקנו ביצוע התמורה הפוכה מ- \underline{U} ל- \underline{V} . מה באשר לתמורה הישירה עבור המקרה הספרטילי? נעשה שימוש בתכונה הבאה למכפלת קrongker ונקבל את הקשר הבא:

$$[C \otimes D]^{-1} = [C^{-1} \otimes D^{-1}] \Rightarrow B^{-1} \underline{V} = \underline{U} = [C^{-1} \otimes D^{-1}] \underline{V} = C^{-1} V D^{-1}$$

כלומר, קיבלנו כי גם התמורה הישירה היא ספרטילית, ולכן גם בה קיים אותו מרכיב של חסכו ביחסובים. אם בנוסף להיותה ספרטילית, התמורות C ו- D יוניטריות, קיבל כי גם פועלות ההיפוך פשוטה, והתמורה הכוללת קלה מאוד לביצוע. במקרה זה, התמורה הכוללת גם היא יוניטרית.

ראינו אם כך סיבות מצוינות להעדיף התמורות ריבועיות, יוניטריות ופרידות. ואמנם, אלה הן התמורות הפופולריות - התמרת ה- DFT, התמרת ה- DCT, התמרת ה- HAAR, וההתמרת ה- Wavelet. אבל, כפי שנאמר כבר בסעיף זה, ישנים מקרים בהם נרצה התמורה כללית יותר בשל סיבות ייחודיות.

7.2 התמרת ה- DFT הדו-מיידית ותכונותיה

במסגרת פרק 2 כבר פגשנו את התמרת ה- DFT הדו-מיידית. במסגרת סעיף זה נחזר ונציגו, והפעם באורינטציה של התمرة יוניטרית, ספרטילית דו-מיידית. ייחודה וחסרונה המרכזיות של התمرة זו, בהשוואה להtamורות אחרות, היא עובדת היותה קומפלקסית. ברוב מוחלט של המקרים עוסקים בתמונות מעלה הממשיים, והמעבר לערכי התمرة קומפלקסים יהווה עבורנו מטרד.

התמרת ה- DFT הדו-מיידית צומחת מתוך התמרת חד-מיידית. התمرة זו מוגדרת באופן הבא: לוקטור $\Omega_v = \{m | 0 \leq m \leq M-1\}$ מעלה התמך $v[m]$ מוגדרת ע"י נתונה ע"י

$$V[k] = F_{ID-DFT}\{v[m]\} = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} v[m] \exp\left\{-j\frac{2\pi mk}{M}\right\}$$

עבור התמך $\Omega_V = \{k | 0 \leq k \leq M-1\}$. ההתמרה ההפוכה מוגדרת ע"י

$$v[m] = F_{ID-DFT}^{-1}\{F[k]\} = \frac{1}{\sqrt{M}} \sum_{k=0}^{M-1} F[k] \exp\left\{j\frac{2\pi mk}{M}\right\}$$

עבור תמונה $f[m,n]$ מעלה התמך $\Omega_f = \{[m,n] | 0 \leq m \leq M-1, 0 \leq n \leq N-1\}$, התمرة ה- DFT הדו-מיידית נתונה ע"י ביצוע התمرة על שורתותה של התמונה, ולאחר מכן על عمودותיה. סדר זה יכול להתחלף ללא שינוי, כיון שהפעולה המבוצעת פרידית. לכן, בשלב ראשון מתקבל:

$$F[m, \ell] = F_{ID-DFT}\{f[m, n]\} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} f[m, n] \exp\left\{-j2\pi\left(\frac{n\ell}{N}\right)\right\}$$

ובשלב שני מתקבל:

$$\begin{aligned} F[k, \ell] &= F_{ID-DFT}\{F[m, \ell]\} = \frac{1}{\sqrt{M}} \sum_{m=0}^{M-1} F[m, \ell] \exp\left\{-j2\pi\left(\frac{mk}{M}\right)\right\} = \\ &= F_{2D-DFT}\{f[m, n]\} = \frac{1}{\sqrt{MN}} \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} f[m, n] \exp\left\{-j2\pi\left(\frac{mk}{M} + \frac{n\ell}{N}\right)\right\} \end{aligned}$$

באופן דומה, ניתן לראות כי ביצוע ההתמרה ההפוכה בשני שלבים של התمرة הפוכה חד-מיידית מניבים את הקשר הבא:

$$f[m,n] = F_{DFT}^{-1}\{F[k,\ell]\} = \frac{1}{\sqrt{NM}} \sum_{k=0}^{M-1} \sum_{\ell=0}^{N-1} F[k,\ell] \exp\left\{j2\pi\left(\frac{mk}{M} + \frac{n\ell}{N}\right)\right\}$$

כאשר התמק במשורר תדר נתון ע"י $\Omega_F = \{[k,\ell] | 0 \leq k \leq M-1, 0 \leq \ell \leq N-1\}$, כך שלעניןינו, יוצא כי רק MN ערכים מוגדרים במלואה את ההתמרה. כל שבדיל בין ההתמורות קדימה ואחרורה הוא סימנו של המערך המרוכב. הגדרות המוצעות שונות במקצת מאילו המופיעות בפרק 2, והבדל זה הינו בקבוע לפני ההתמרה אשר מבטיח כי מדובר בהתמרה יוניטרית.

כשדים ב- DFT מקובל להשתמש בסימנו $W_N^k = \exp\{-j2\pi k/N\}$. אז ההתמרה וההתמרה ההופוכה נתונות ע"י ייצוג מטריצי מהצורה (נתיחה להתמרה חד-מימדית):

$$\begin{bmatrix} V[0] \\ V[1] \\ \vdots \\ V[M-1] \end{bmatrix} = \frac{1}{\sqrt{M}} \begin{bmatrix} W_M^0 & W_M^0 & \dots & W_M^0 \\ W_M^0 & W_M^1 & \dots & W_M^{M-1} \\ \vdots & \vdots & \ddots & \vdots \\ W_M^0 & W_M^{M-1} & \dots & W_M^{(M-1)^2} \end{bmatrix} \begin{bmatrix} v[0] \\ v[1] \\ \vdots \\ v[M-1] \end{bmatrix} \Rightarrow \underline{F} = \mathbf{W}_M \underline{v}$$

מטריצת ההתמרה \mathbf{W} הינה מטריצה יוניטרית, ולכן, היפוכה קיים ופשט (נתון ע"י שלוח וצמוד). ההתמרה הדו-מימדית ניתנת אף היא לתיאור מטריצי וקטורי. למעשה ישנן שתי אפשרויות לייצגה. בדרך הראשונה, כשאנו מתייחסים ל- f כמטריצת הפיקסלים בגודל M שורות על N עמודות, הכפלת משמאלי במטריצה \mathbf{W}_M תביא לביצוע ההתמרה על עמודותיה של התמונה. הכפלת התוצאה מימין במטריצה \mathbf{W}_N משלימה ע"י ביצוע ההתמרה על שורות התמונה, וכך קיבלנו את F - תמונה ההתמרה הדו-מימדית. לכן:

$$F = \mathbf{W}_M f \mathbf{W}_N$$

בגישה השנייה, נמיר את התמונה למבנה וקטורי ע"י סידור לקסיקוגרפי לפי עמודות ונקל וקטור \underline{f} באורך MN . וקטור זה יוכפל באופן הבא:

$$\underline{F} = \mathbf{W}_{2D-DFT[M,N]} \underline{f} = [\mathbf{W}_M \otimes \mathbf{W}_N] \underline{f}$$

מטריצה זו בגודל MN על MN מבצעת את פעולה ההתמרה במלואה על הוktor f מניבה את תמונה ההתמרה כッグ היא במבנה וקטור עפ"י סידור לקסיקוגרפי לפי עמודות. קשר זה הכרנו עוד בפרק 2 כאשר אינו כי פעולות ליניאריות דו-מימדיות פרידות ניתנות לתיאור כמכפלת קרונקר של שתי פעולות חד-מימדיות. ברור כי המטריצה הכוללת מהוות אף היא מטריצה יוניטרית (ניתן להוכיח זאת ע"י שימוש בתכונות מכפלת קרונקר).

נסקרו מספר תכונות חשובות של ה- 2D-DFT:

1. **מחזריות:** בהינתן תמונה $[f]_{m,n}$ מעל התמך $\{0 \leq m \leq M-1, 0 \leq n \leq N-1\}$, הינה $\Omega_f = \{[m,n] | 0 \leq m \leq M-1, 0 \leq n \leq N-1\}$. הtransform ה- DFT הדו-מימדית הנגינה כ- $F[k,\ell]$ מוגדרת מעל התמך $\{0 \leq k \leq M-1, 0 \leq \ell \leq N-1\}$. לכל נקודה דיסקרטית $[k,\ell]$ מחוץ לתחום זה מתקבלת מחזריות לפי הקשר:

$$\forall p,q \in \mathbb{Z}, \quad F[k,\ell] = F[k+pM, \ell+qN]$$

2. **ליניאריות, יוניטריות, וספריליות:** פועלות ה- 2D-DFT המתווארת ע"י המטריצה $W_{2D-DFT[M,N]}$ מהויה התמורה ליניארית, ריבועית, יוניטרית, ופרידה. כל תכונות אלה קלות לאבחנה מתוך כל שנאמר עד כה. מטריצה זו הינה גם מטריצה סימטרית.

3. **יישום מהיר:** ביצוע פועלות ה- DFT הדו-מימדית על תמונה בגודל $M \times N$ יכול להיעשות ע"י $M \times N$ פועלות FFT חד-מימדיות לאורץ N , ולאחר מכן, $N \times N$ פועלות FFT חד-מימדיות לאורץ M . כיוון שפעולות FFT על N איברים דורשת $N \log_2 N$ פעולות (כל פעולה היא חיבור ומכפלה), סך הפעולות הנדרש להמרת מלאה הוא:

$$MN \log_2 N + NM \log_2 M = (NM) \log_2 (NM)$$

זאת בהשוויה לביצוע ישיר של ההמרת ע"י M^2 פעולות. נשים לב שלו לקחנו את התמונה והמרנו אותה לסדר לקסיקוגרפי, פועלות FFT חד-מימדית הייתה דורשת אותה כמות של פעולות.

4. **לכזון מטריצות בлок סיבוביות:** מטריצת התמורה ה- DFT החד-מימדית מלכשת מטריצות בעלות מבנה סיבובי (ראה פרק 2 להגדירה מדויקת של מבנה זה. בשל כך, מטריצת התמורה ה- DFT הדו-מימדית מלכשת מטריצות בлок סיבוביות).

5. **סימטרייה צמודה:** לתמונה ממשית $[f]_{m,n}$, ההמרת $F[k,\ell]$ מקיימת את הקשרים הבאים:

$$F[k,\ell] = F^*[M-k, N-\ell] \quad \forall 0 \leq k < M, \quad 0 \leq \ell < N$$

כתוצאה לכך, למROT שההמרת מכילה MN ערכים קומפלקסיים (כלומר שווה ערך ל- $2MN$ ערכים ממשיים), הרי שרק MN מתוכם בלתי תלויים. ראיינו כבר מקבילה של תכונה זו עבור התמורות הפוריות שנסקרו בפרק 2.

6. **משפט הקונבולוציה:** בהינתן שתי תמונות $[n][m]$ ו- $f_2[m,n]$ בגודל זהה $M \times N$, תוצאה ה-DFT הדו-מימדי של הקונבולוציה החיקילית ביניהם זהה לוצאה המכפלה של התמורותיהן. כלומר, אם:

$$z[m,n] = \sum_{k=0}^{M-1} \sum_{j=0}^{N-1} f_1[k,j]f_2[(m-k) \bmod M, (n-j) \bmod N]$$

$$\text{אז } Z[k,\ell] = F_1[k,\ell]F_2[k,\ell]$$

7. **משפט פרסלבל:** בהינתן שתי תמונות $[n][m]$ ו- $f_2[m,n]$ בגודל זהה $M \times N$, תוצאה מכפלתם הפנימית זהה המקיים ובתדר, דהיינו:

$$\sum_{m=0}^{M-1} \sum_{n=0}^{N-1} f_1[m,n]f_2^*[m,n] = \sum_{k=0}^{M-1} \sum_{\ell=0}^{N-1} F_1[k,m]F_2^*[k,n]$$

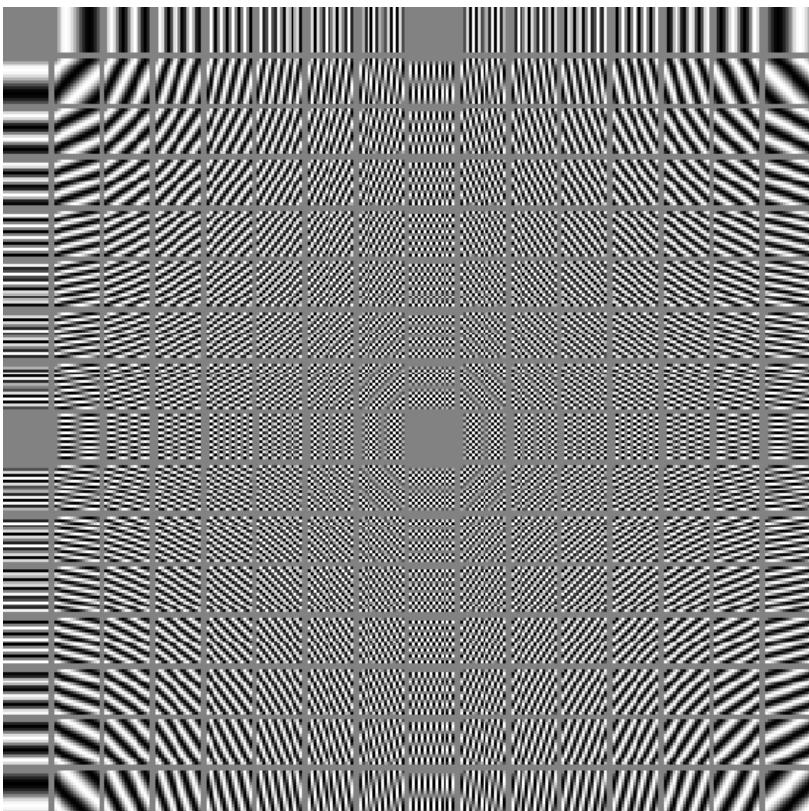
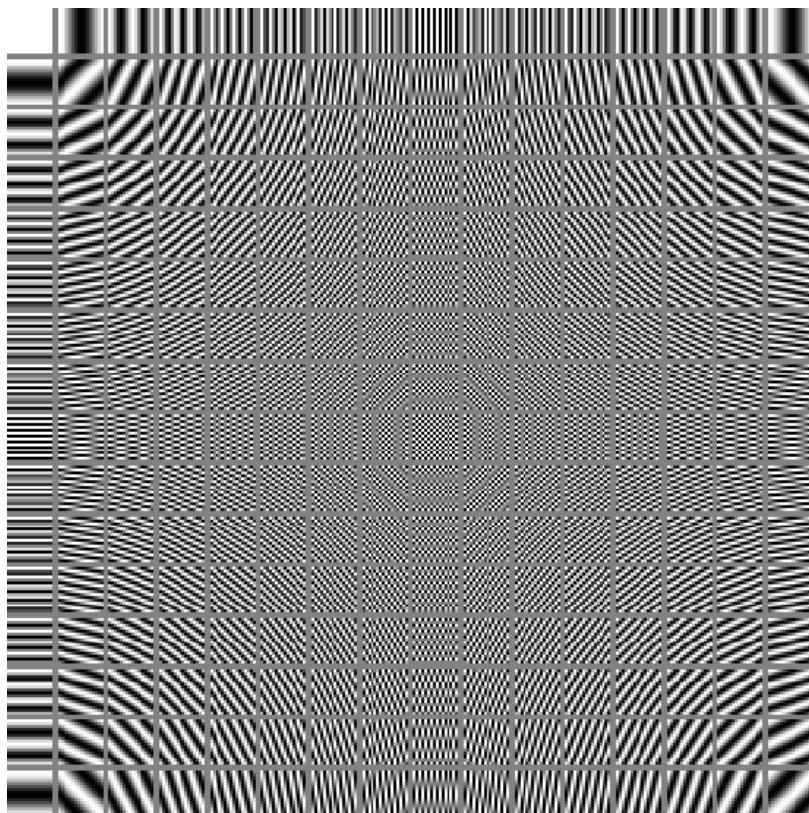
תוצאה זו לא אמורה להਪתיע אותנו כיון שיפוי שנאמר כבר בתחילת פרק זה, כל התמורה יוניטרית ממלאת תוכנה זו, ובפרט ה-2D-DFT.

8. **וקטור הבסיס להטמרה:** התמורה ה-DFT הדו-מימדי מציעה שימוש ב- MN תמונות בסיס בגודל $M \times N$ כל אחת, לצורך ייצוגה של תמונה נתונה. על מנת לקבל את תמונות הבסיס ה- k , علينا לבנות וקטור מלאכותי \underline{F} בו כל הערכים 0 למעט האיבר ה- k שייהי 1, ולבצע התמורה פוריה הפוכה עליו, לקבלת \underline{f} . מכאן נובע באופן פשוט כי עמודותיה של המטריצה $\mathbf{W}_{2D-DFT[M,N]}^H = \mathbf{W}_M^H \otimes \mathbf{W}_N^H$ הם וקטורי הבסיס של ההטמרה.

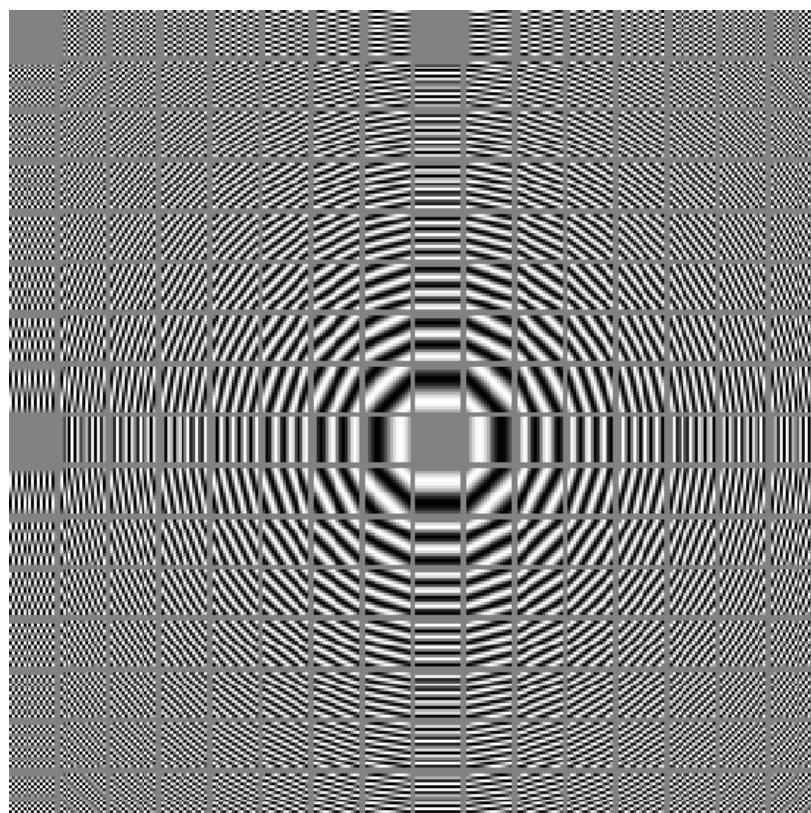
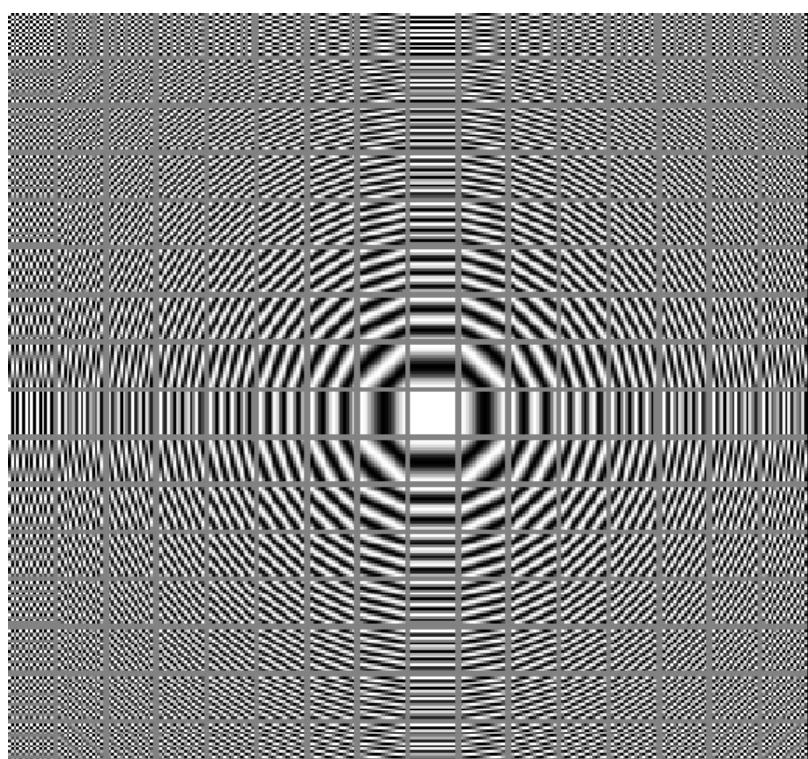
דוגמה 7.1: נציג את תמונות הבסיס של התמורה ה- 2D-DFT לתמונות בגודל 16 על 16. תהיינה 25 תמונות בסיס, אך כיון שכל תמונה בסיס כזו היא קומפלקסית, יש להציג אותה כשתי תמונות עבור החלק ממשי והחלק המדומה. ציור 7.1 מראה את משפחת הבסיסים זו.

ניתן לראות כי תמונה הבסיס הראשונה קבועה, וממנה התדר הולך ומתחזק כשהוא הרמוני (פונקציית קוסינוס או סינוס) טהור, ועם מגוון כיוניים. מעבר לחצי (אחרי האינדקס 9) מתקבלת ירידיה שוב.

נוח לראות את תמונות הבסיס לאחר פעולה זוזה למרכז, כך שהתדרים עולים כשמתרחקים באופן רדילי מהמרכז - ציור 7.2 נותן ראייה זו. כפי שניתנו לראות, סביבה נקודת המרכז התדר עולה ובמגוון זוויות לפי כיוון ההתקדמות. באופן זה כל שפה בתמונה יכולה להיות מיוצגת.



ציור 7.1 - וקטורי הבסיס של התמרת ה-2D-DFT. למעלה - החלק הממשי, למטה - החלק המדומה



ציור 7.2 - וקטורי הבסיס של התמורה ה- 2D-DFT לאחר פעולה הזזה למרכז (FFTSIFT). לעלוה - החלק ממשי, למטה - החלק המדומה

7.3 התמרת ה- DCT הדו-מימדית ותכונותיה

התמרת ה- DCT החד-מימדית לוקטור $[m] \times [M]$ והההתמרה הההפוכה לה נתונות ע"י שני

הקשרים הבאים :

$$V[k] = \begin{cases} \sqrt{\frac{1}{M}} \sum_{m=0}^{M-1} v[m] \cos\left[\frac{\pi(2m+1)k}{2M}\right] & k = 0 \\ \sqrt{\frac{2}{M}} \sum_{m=1}^{M-1} v[m] \cos\left[\frac{\pi(2m+1)k}{2M}\right] & k \neq 0 \end{cases}$$

$$v[m] = \sqrt{\frac{1}{M}} V[0] + \sqrt{\frac{2}{M}} \sum_{k=1}^{M-1} V[k] \cos\left[\frac{\pi(2m+1)k}{2M}\right]$$

גם הפעם מדובר בההתמרה ליניארית אך ללא ערכים קומפלקסיים. ניתן, בדומה ל- DFT, להצעה מבנה מטריצי לתיאור הההתמרה, וכל לראות מה יהיה תוכנה של מטריצה מעין זו, אשר תסומן ב- C_M . מטריצה זו אינה סימטרית, אך היא ממשית ויוניתרית. לכן, היפוכת נתון פשוט ע"י פעולה השחלוף (ללא צמוד קומפלקס).

בדומה ל- DFT, התمرة דו-מימדית לתמונה f בגודל $M \times N$ תיעשה ע"י ביצוע התمرة לשורות ולעמודות, כולם ביצוע פעולה פרידזה. לכן :

$$F = C_M f C_N$$

או בגישה של תמונה במבנה וקטורי ע"י סידור לקסיקוגרפי לפי עמודות :

$$\underline{F} = \mathbf{C}_{2D-DCT[M, N]} \underline{f} = [\mathbf{C}_M \otimes \mathbf{C}_N] \underline{f}$$

ומטריצה זו הינה שוב יוניתרית, המיצגת התمرة פרידזה. נסקור בקצרה ולא הוכיח את תכונותיה של התمرة זו.

1. **LINIARITY, YONIERTY, AND SPERIBILITY:** פועלות ה- 2D-DCT המתווארת ע"י המטריצה $\mathbf{W}_{2D-DCT[M, N]}$ מהוות התمرة ליניארית, ריבועית, יוניתרית, ופרידזה. כל תכונות אלה כוללות לאבחן מתוך כל שנאמר עד כה.

2. **יישום מהיר:** התמרת ה- DCT קשורה באופן הדוק לפעולות ה- DFT. כתוצאה מכך, גם ביצוע ה- DCT הדו-מימדי על תמונה בגודל M על N יכול להיעשות באופן מהיר ע"י שימוש בפעולות FFT. סך הפעולות הדרוש זהה לביצוע התמרת DFT דו-מימדי בגודל זהה.

3. **לכソン מטריצות בлок מركוביות:** מטריצת התמרת ה- DCT החד-מימדיות מלכסנת במדויק מטריצות תלת-אלכסוניות, בעלות המבנה הבא :

$$Q = \begin{bmatrix} 1-\alpha & -\alpha & 0 & & \\ -\alpha & 1 & -\alpha & 0 & \\ \ddots & \ddots & \ddots & & \\ 0 & -\alpha & 1 & -\alpha & \\ & 0 & -\alpha & 1-\alpha & \end{bmatrix}$$

בשל תכונה זו, ניתן להראות כי מטריצות טופליץ מרכוביות הבנוויות באופן הבא :

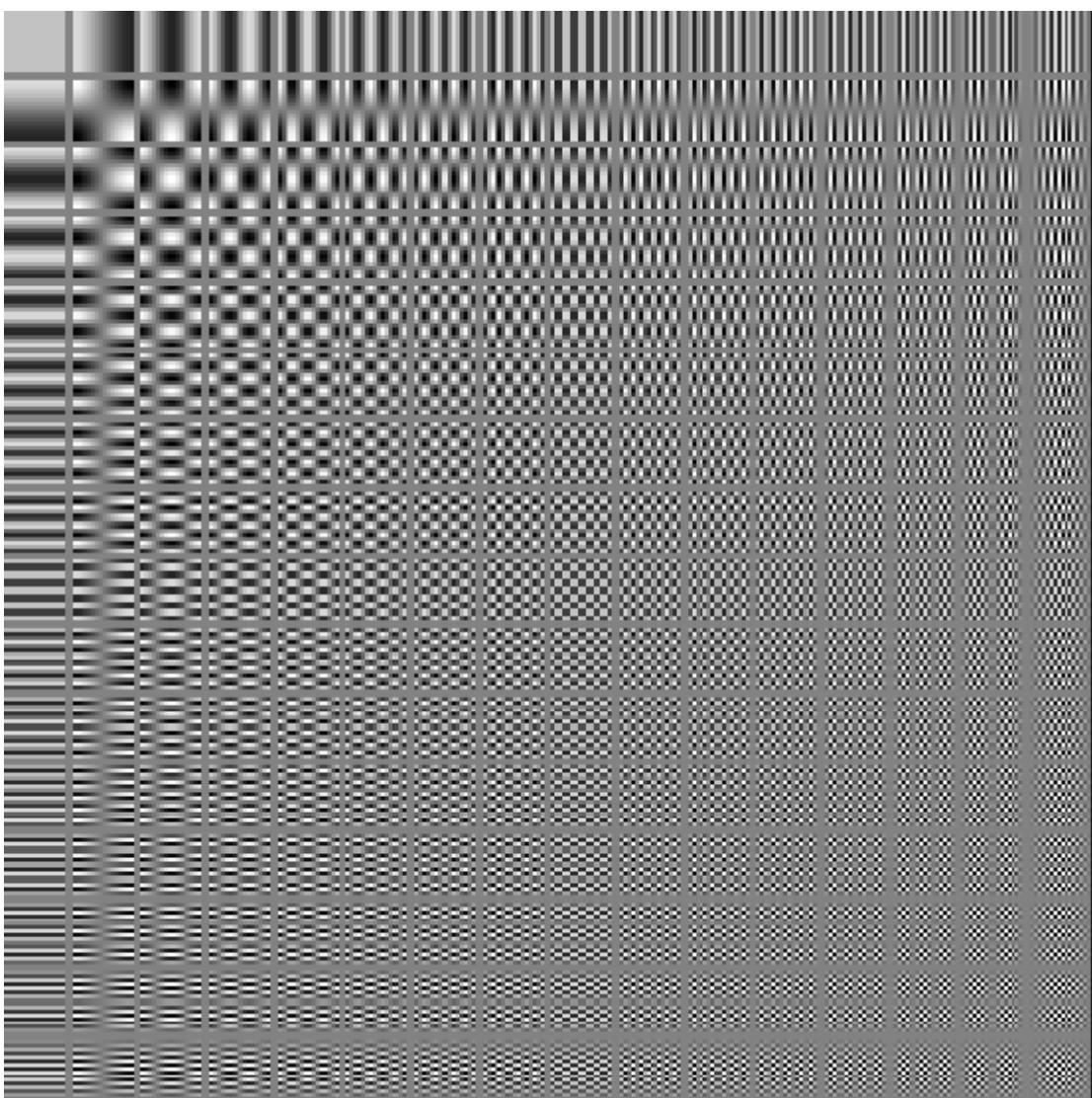
$$R = \begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{M-1} \\ \rho & 1 & \rho & \rho^2 & \dots \\ \ddots & \ddots & \ddots & & \\ \rho^2 & \rho & 1 & \rho & \\ \rho^{M-1} & \rho^2 & \rho & 1 & \end{bmatrix}$$

מלכסנות ע"י התמרת ה- DCT בקירוב טוב, אם $1 \rightarrow k$. הסיבה לכך היא שעבור $1 \rightarrow k$ מתקבל כי היפוכה של R היא בדיקות טוב, ולכן ברור שה- DCT מהווה מטריצה מלכסנת. תכונה זו חיונית כיון שכפי שנראה בהמשך, מטריצת טופליץ מרכובית מתארת התנהוגות אופיינית של אוטות בטבע, ולכsonoה חשוב לצורך ייצוג נאות או דחיסה. אנו נשוב לתכונה חשובה זו כshedבר על התמרת KLT.

בשל כל האמור לעיל, כאשר למטריצה יש מבנה בлок טופליץ מרכובי, מתקבל כי התמרת ה- DCT הדו-מימדיות מלכסנת אותה בקירוב טוב, וזה נכון כאשר $1 \rightarrow k$.

4. **קונבולוציה:** בשל הקשר הדוק בין ה- DFT ל- DCT, ניתן להראות כי ניתן בעזרת DCT לבצע קונבולוציה מהירה יותר, בדומה לכך בה הדבר מוצע ב- DFT. עם זאת, הקישרים מורכבים יותר, ואין זה הצד החזק של ה- DCT.

5. **וקטורי הבסיס להטמרה:** ציור 7.3 מראה את וקטורי הבסיס של התמרת ה- DCT הדו-מימדיות לתמונות בגודל 16 על 16. ניכר הדמיון הרב להטمرة ה- DFT.



ציור 7.3 - תМОונות הבסיס של התמורת ה- DCT לתМОונות בגודל 16 על 16.

7.4 התמורת הדוו-מיימדית ותכונותיה

החלנו את הדיוון בהtamורות ספציפיות בהtamורת DFT, וראינו כי היא כופה עלינו חשבונות עם מספרים קומפלקסיים. התמורת ה- DCT הקלה בעיה זו בהצעעה התמורה יוניטרית ממשית.Cut נמשיך קו זה של הפשטת ונציג התמורת בה אין כלל מכפלות במספרים כלשהם - כל איברי המטריצה הממירה יהיו 1 או -1. זהה התמורת Hadamard-Walsh, המוכרת גם בשם התמורת Hadamard.

במקרים, נתחיל בתיאור ההtamורה החד-מיימדית לוקטור באורך M , ונניח הפעם כי M מוגדרת כזוקה שלמה של 2 (דרישה זו מוכרת לנו גם בהקשר ל- DFT, שם היא חשובה לשם יישום יעיל של ה-

FFT). בנויגוד לכך בה נקטנו בדיון בהtamורות קודומות, הפעם נבנה ישירות את מטריצת ההמרה H . התהליך הוא רקורסיבי - נתחילה במטריצה בסיסית בגודל 2 על 2:

$$H_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

התמרה זו תקיפה לוקטורים באורך 2, וניכר כי היא יוניטרית. לשם בניית התמרה לאורך 4 נבנה את H_2 :

$$H_2 = H_1 \otimes H_1 = \frac{1}{2} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix}$$

בשל דרך בנייתה, ברור כי התמרה זו יוניטרית. התהליך יכול להמשיך שוב ושוב, ומתקבל כי :

$$H_n = H_{n-1} \otimes H_1$$

והתוצאה היא מטריצה המכילה 1 ו- (-1), שהינה יוניטרית וסימטרית. בשל כך, ההtamורה ההפוכה נתונה ע"י אותה מטריצה ממש.

ברור כי התמרה זו לתמונות בגודל M על N (שוב, בהנחה כי אלו הם חזקות של מומנות של 2) מתקבלת ע"י מכפלת קrongenker בין שתי מטריצות התמרה חד-מימדיות. עובדה זו מבטיחה שההתמרה הכוללת לתמונה תהווה מטריצה סימטרית, יוניטרית, בעלת איברים 1 ו- (-1), ולמעשה, זהה לחלווטין להtamורה לוקטור חד-מימדי באורך NM .

כמשמעותם על שורה מסוימת במטריצת ההמרה H , מספר החלפות הסימן מתאר מקבילה לתדר. מספר החלפות הסימן מוכרך בשם Sequency, וניתן לראות כי שורות המטריצה H אינן מסודרות בסדר Sequency מונוטוני. מקובל לעיתים לחת את מטריצת ההמרה שנבנתה ולסדר את שורותיה כך שאמנם יתקבל Sequency גדול מונוטוני.

דוגמה 7.2: נראה מהי מטריצת Hadamard לוקטורים באורך 8, וכי怎樣 לסדרה. המטריצה נתונה ע"י :

$$\left(\begin{array}{cccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1- & 1 & 1- & 1 & 1- & 1 & 1- & 1 \\ 1- & 1- & 1 & 1 & 1- & 1- & 1 & 1 \\ 1 & 1- & 1- & 1 & 1 & 1- & 1- & 1 \\ 1- & 1- & 1- & 1- & 1 & 1 & 1 & 1 \\ 1 & 1- & 1 & 1- & 1- & 1 & 1- & 1 \\ 1 & 1 & 1- & 1- & 1- & 1- & 1 & 1 \\ 1- & 1 & 1 & 1- & 1- & 1- & 1- & 1 \end{array} \right)$$

ה-Sequency לפי שורות הוא - 0,7,3,4,1,6,2,1. המטריצה לאחר סידורה תהיה:

$$\left(\begin{array}{cccccccc} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1- & 1- & 1- & 1- & 1 & 1 & 1 & 1 \\ 1 & 1 & 1- & 1- & 1- & 1- & 1 & 1 \\ 1- & 1- & 1 & 1 & 1- & 1- & 1 & 1 \\ 1 & 1- & 1- & 1 & 1 & 1- & 1- & 1 \\ 1- & 1 & 1 & 1- & 1- & 1- & 1 & 1 \\ 1 & 1- & 1- & 1- & 1- & 1 & 1- & 1 \\ 1- & 1 & 1 & 1- & 1- & 1- & 1- & 1 \end{array} \right)$$

נסקרו בקצרה ולא הוכחות את תכונותיה של התמורה זו.

1. **ליניאריות, יוניטריות, וספרביליות:** פועלות ה- 2D-Hadamard המתווארת ע"י המטריצה $[N,M]_{\text{Had.}} H_{2\text{D}}$ מהוות התמורה ליניארית, ריבועית, יוניטרית, ופרידה. כל תכונות אלה קלות

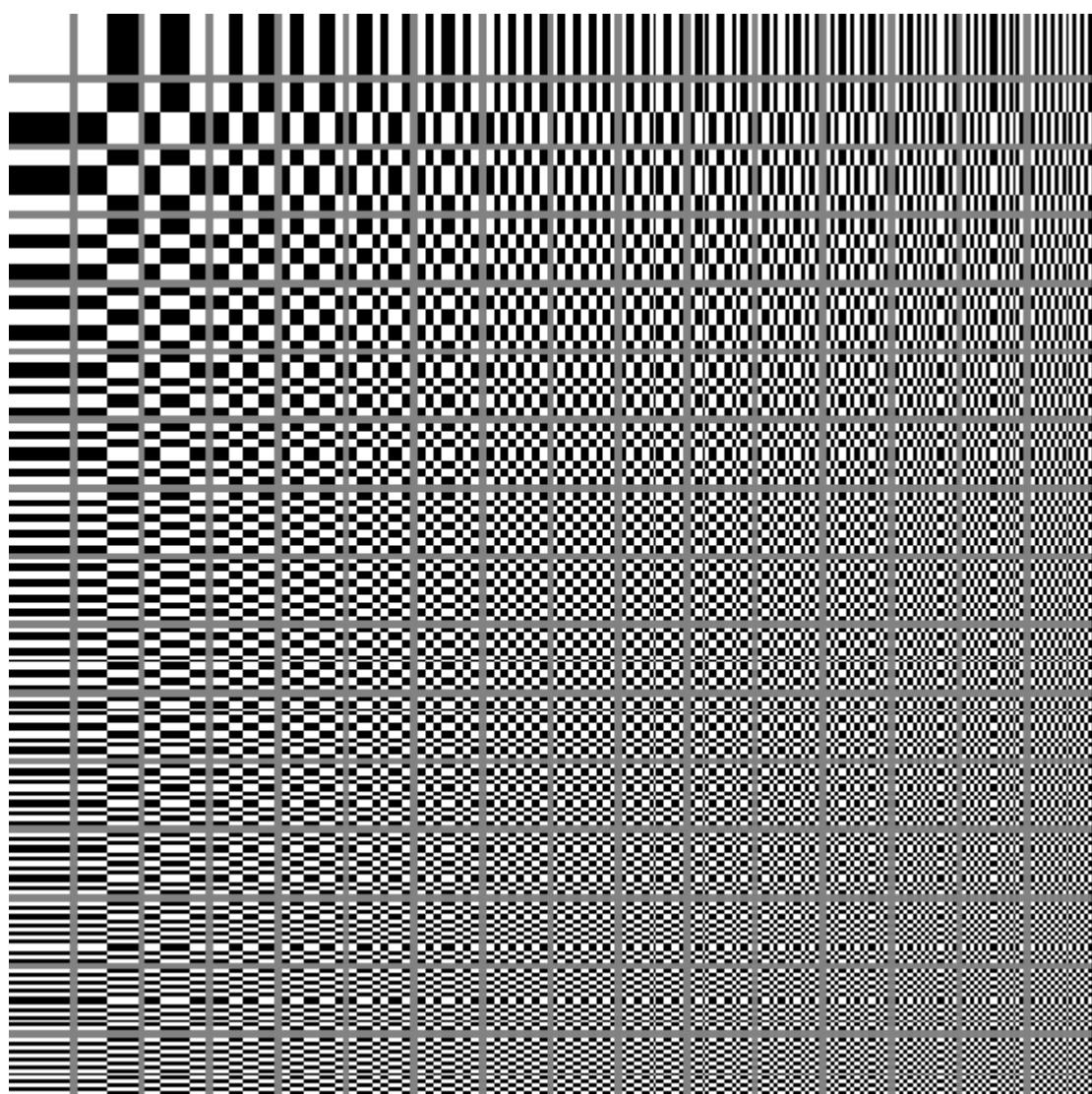
לאבחנה מותך כל שנאמר עד כה. יתרה מזו, התמורה דו-מימדית זו זהה בדיק להמרת הדמדוד חד-מימדית על וקטורים באורך MN . המטריצה $H_{2\text{D}-\text{Had.}}[M,N]$ סימטרית ומורכבת מאיברי $+1$ ו- 1 בלבד.

2. **יישום מהיר:** גם להמרת זו ישום מהיר, בדומה למבנה FFT, אך מהיר אף יותר בשל שימוש בחיבורים בלבד.

3. **לכסוון מטריצות בלוק מרקוביות:** בדומה להמרת ה- DCT, מטריצת התמרת הדמדוד מקרבת באופן סביר ליכסוון של מטריצות בעלות מבנה בלוק טופלי עם ערך קורלציה קרוב לאחת, $\rightarrow 0$.

4. **וקטורי הבסיס להמרת:** ציור 7.4 מראה את וקטורי הבסיס של התמרת הדמדוד הדו-מימדית לתמונות בגודל 16 על 16. ניכר הדמיון הרב להמרת ה- DCT. תכונה מעניינת היא שאם לוקחים את

בבסיס ה- DCT לאחר פעולה sign (כלומר שליליים ל- -1 וחוביים ל- $+1$), מקבלים התמורה קרובה מאוד להטמרה הדמරד, עובדה המצביעת על הקשר והדמיון הקרוב ביניהם.



ציור 7.4 - תМОונות הבסיס של התמורה Hadamard לתמונות בגודל 16 על 16.

7.5 התמרת Haar ומבוא ל-

מבין כל המטריצות היוניטריות שיכולה לשמש אותנו כהתמורה, ראיינו בתחילת מטריצה בה הערכים מרכיבים (DFT), שיפרנו ע"י מציאת התמורה ממשית (DCT), והקלנו את המורכבות החישובית לאין ערוך ע"י שימוש במטריצה בה האיברים הם $m+1$ או -1 . האם ניתן להמשיך במנגמת הפשטה זו? התשובה היא כן! נציג עצת התמורה אשר גם בה הערכים הם $m+1$ או -1 , אך מרבית האיברים מטריצת התמורה יהיו אפס, ובכך נשיג התמורה הדורשת פחות חישובים. התמורה זו היא התמורה Haar. נתחיל בהציגת התמורה זו לחדר מימד ונראה כיצד ממנה נובעת התמורה דו-מימדית.

חשוב להבהיר כי יתרונה של התמורה Haar אינו מותמצא בהיותה בעלת מטריצת התמורה דיליה (בها יש הרבה אפסים). התמורה זו מהוות מקרה פרטי פשוט של משפחה רחבה של התמורות הקrüיות, אשר נמצאו להיות ייעילות בשל מגוון סיבות לטיפול באירועים טבעיות. במסגרת זו נציג את רק את היסודות בתחום רחב ועשיר זה.

כאמור, נתחיל בתיאור התמורה חד-מימדית לוקטור באורך M זוגי. נתחיל במטריצה בסיסית בגודל 2 על 2. מטריצה זו זהה לתמורה Hadamard :

$$D_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \Rightarrow D_1^{-1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

התמורה זו תקיפה לוקטורים באורך 2, וניכר כי היא יוניטרית. משמעות התמורה זו היא שבהינתן שני מספרים a ו- b הבונים את הוקטור הנכנס לתמורה, התמורה תהיה :

$$\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} a+b \\ a-b \end{bmatrix} = \begin{bmatrix} c \\ d \end{bmatrix} \Rightarrow \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} c \\ d \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} c+d \\ c-d \end{bmatrix} = \begin{bmatrix} a \\ b \end{bmatrix}$$

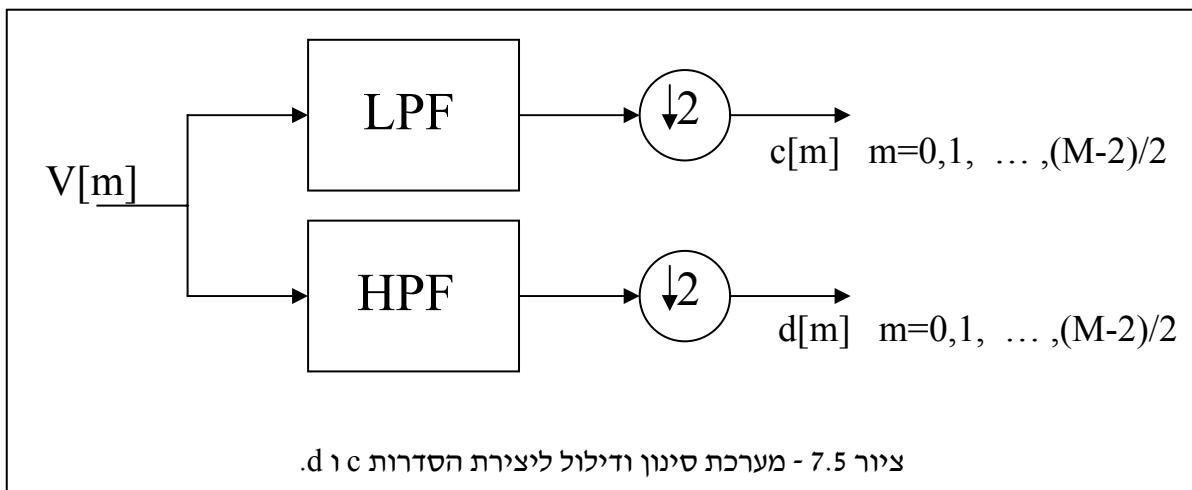
ככלומר, צמד המספרים הוחלף בתמורה בסכומים ובפרשנים (ובהכפלה בקבוע הנובע מרצונו בתמורה יוניטרית). התמורה הפוכה ביצעה אותו דבר בדיזוק.

בהינתן סידרה בת M (זוגי) איברים, נוכל לחלקם לזוגות סמוכים, ולהמיר כל צמד בסכומים והפרשנים כפי שבוצע לצמד הקודם. אם נסמן את סידרת הכניסה ב- $[m]_v^a$, קיבל את סידרת המספרים הבאים :

$$c[m/2] = \frac{1}{\sqrt{2}} [v[m] + v[m+1]] \quad m = 0, 2, 4, \dots, M-2$$

$$d[m/2] = \frac{1}{\sqrt{2}} [v[m] - v[m+1]] \quad m = 0, 2, 4, \dots, M-2$$

וכך התקבלה סידרת מספרים בת M איברים המייצגת באופן מלא את סידרת המקור. למעשה, המערכת המתווארת בציור 7.5 שköלה לתהיליך שתואר.



במערכת זו, עroz ה- LPF מחשב ממוצעים מקומיים לפי :

$$c[m] = \frac{1}{\sqrt{2}} [v[m] + v[m+1]] \quad m = 0, 1, 2, 3, 4, \dots, M-2, M-1$$

אך כיוון שאנו נזקקים רק לזוגיים שביהם, ישנה פעולה דילול ביחס 2, שזורקת את כל התוצאות במקומות אי-זוגיים. באופן דומה, עroz ה- HPF מחשב הפרשיים מקומיים לפי :

$$d[m] = \frac{1}{\sqrt{2}} [v[m] - v[m+1]] \quad m = 0, 1, 2, 3, 4, \dots, M-2, M-1$$

ושוב דילול הוא זה שמסיר ממחצית המספרים (מהמקומות האי-זוגיים). על מנת לקבל את הסדרות c ו- d ולהזור למקור $[m]$, علينا לחתך כל צמד של ערכים מתאימים, ולבצע עליהם סיכום והפרש, כמפורט, לפחות את הערכים לפי הסדר הנכון. ניתן לראות כי סידרת הפעולות הבאה מושגת אותו אפקט :

1. קחוי את שתי הסדרות c ו- d והכנסי אפס בין כל שני ערכים שלחן (איבר ראשון יהיה גם הוא 0).
2. העברוי את הסדרה המוגדלת c דרך LPF זהה לזה בו נעשה שימוש קודם.
3. העברוי את הסדרה המוגדלת d דרך HPF זהה לזה בו נעשה שימוש קודם.
4. חסרו את התוצאות, וקבל את v .

נמחייב תהליך זה דרך דוגמה :

דוגמה 7.1: ניקח את הסדרה הבאה ונפעיל עליה את התהליכיים שתוארו.

y=[1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16]

הסדרות ו- ד תתקבלנה ע"י

$$c = \frac{1}{\sqrt{2}} [(1+2) \quad (3+4) \quad (5+6) \quad (7+8) \quad (9+10) \quad (11+12) \quad (13+14) \quad (15+16)] = \\ = \frac{1}{\sqrt{2}} [3 \quad 7 \quad 11 \quad 15 \quad 19 \quad 23 \quad 27 \quad 31]$$

$$\begin{aligned} d &= \frac{1}{\sqrt{2}} [(1-2) \quad (3-4) \quad (5-6) \quad (7-8) \quad (9-10) \quad (11-12) \quad (13-14) \quad (15-16)] = \\ &= \frac{1}{\sqrt{2}} [-1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1] \end{aligned}$$

נבעו בקטע את תהליכי השזהר המוצע. מילואן באפסים כמשמעותם באפס:

$$c = \frac{1}{\sqrt{2}} [0 \ 3 \ 0 \ 7 \ 0 \ 11 \ 0 \ 15 \ 0 \ 19 \ 0 \ 23 \ 0 \ 27 \ 0 \ 31]$$

הפעלת LPF ו- HPF על שתי סדרות אלו תיתן :

$$\text{LPF}\{c\} = \frac{1}{2} [(0+3) \quad (3+0) \quad (0+7) \quad (7+0) \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots \quad \dots] =$$

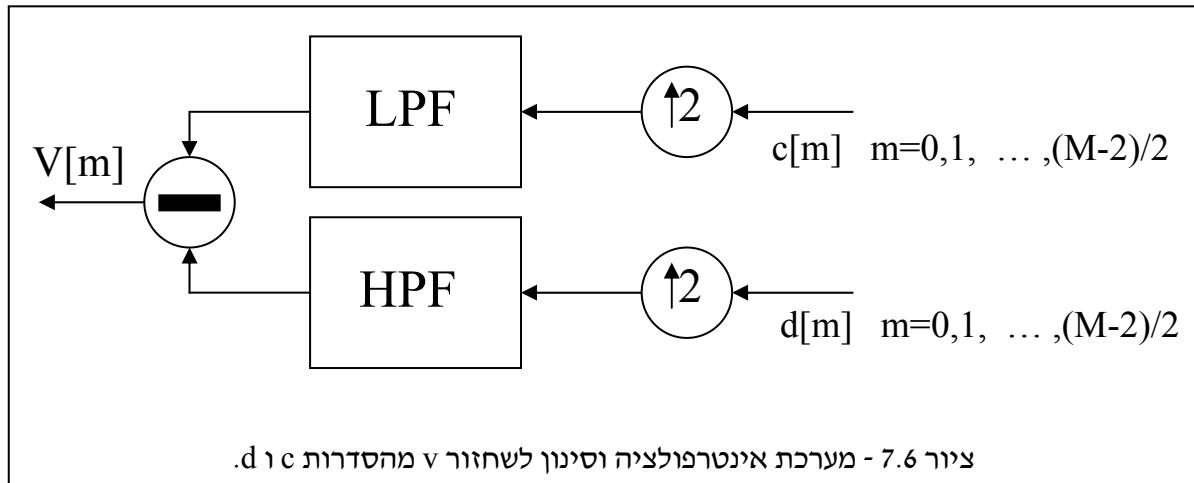
$$= \frac{1}{2} [3 \quad 3 \quad 7 \quad 7 \quad 11 \quad 11 \quad 15 \quad 15 \quad 19 \quad 19 \quad 23 \quad 23 \quad 27 \quad 27 \quad 31 \quad 31]$$

וחיסוך שטי סדרות אלו ייתן את סידרת המקור בדיק:

$$\text{LPF}\{c\} - \text{HPF}\{d\} =$$

$$\begin{aligned}
 &= \frac{1}{2} [3 \ 3 \ 7 \ 7 \ 11 \ 11 \ 15 \ 15 \ 19 \ 19 \ 23 \ 23 \ 29 \ 29 \ 31 \ 31] + \\
 &\quad - \frac{1}{2} [-1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1 \ -1 \ 1] = \\
 &= [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14 \ 15 \ 16].
 \end{aligned}$$

גם את תהליך השחזר ניתן לתאר כסיכום בלוקים הדומה לציור 7.5. סכימת השחזר מתוארת בציור .7.6



כיוון שבכל התחליך הנ"ל בוצעו פעולות ליניאריות על איברי $[m]$, נוכל לייצג מכפלת הוקטור V במטריצה. עבור וקטור בעל ארבעה איברים, מטריצה זו תהיה הבאה :

$$D_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}$$

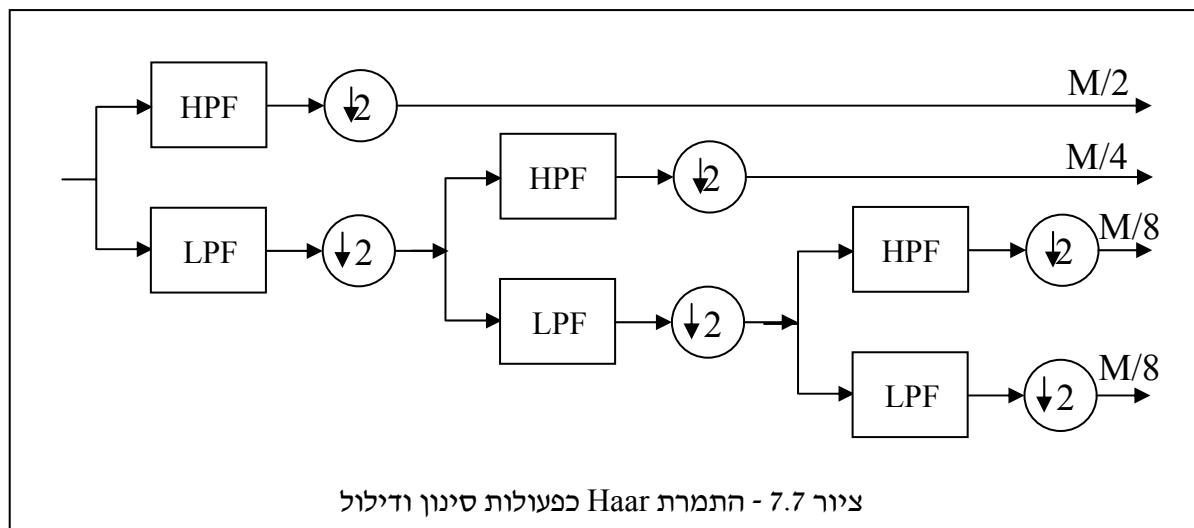
כך שתתי השורות הראשונות בונות את הממוצעים (LPF) ושתי השורות האחרונות בונות את ההפרשים (HPF). במטריצה זו רק חצי מהאיברים שונים מאפס. באופן דומה, לוקטור בעל שמונה איברים נקבל :

$$D_3 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}$$

וكل לראות כי מטריצה זו יוניטרית, ובה רק רביע מהאיברים שונים מאפס. חשוב להבהיר - זו עדין לא התמרת Haar.

עד כה הتمקדנו במרקם החשוב הראשון של ההתמרה - ההפרדה לסדרות LPF ו- HPF מדוילות. נציג כעת את המוטיב השני המרכזי והוא נושא המולטי-רזולוציה. הרעיון העיקרי פשוט: את סידרת הפרשים d נותר לא שינוי, ועל סידרת הממוצעים c נבצע אותו תהליך קודם, כלומר בנית שתי סדרות של ממוצעים והפרשים עבורה. ברור כי כל סידרה כזו כבר תהיה באורך רבע מהאורך המקורי, וכך כל האיברים יסתכם ל- M . תהליך זה יחוור רקורסיבית כל פעם על סידרת ה- LPF, עד לקבלת סדרות באורך 1. התוצאה היא התמרה Haar. סכימת בלוקים של דרך בניתה מתוארת בציור 7.7.

7.7



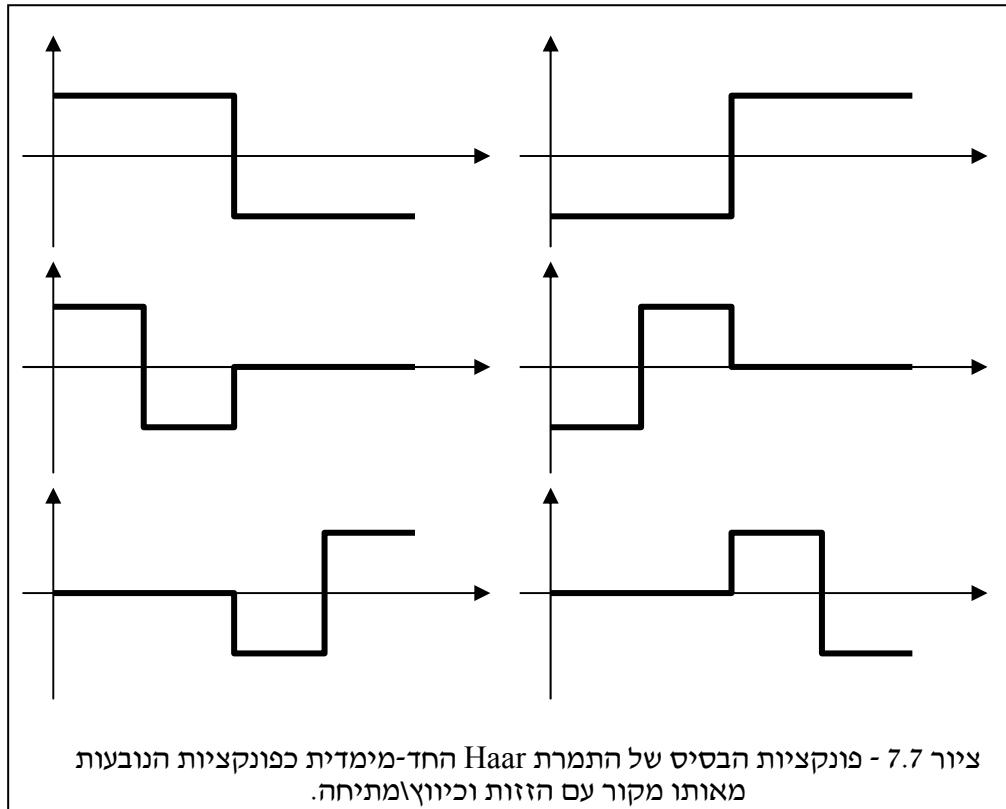
לדוגמה, התמרה וקטור בעל 4 אלמנטים יתבצע ע"י המטריצה הבאה:

$$D_2^H = \frac{1}{\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ \sqrt{2} & -\sqrt{2} & 0 & 0 \\ 0 & 0 & \sqrt{2} & -\sqrt{2} \end{bmatrix}$$

המטריצה הימנית במכפלה היא המטריצה שכתובה קודם. מטריצה זו מוכפלת במטריצה שמותירה את שתי השורות התחthonות ללא שינוי, וביצעת התמרה על שני האיברים הראשונים כפי שתואר קודם. ניתן באופן דומה לבנות את המטריצה המתאימה להתמרה וקטור באורך 8 אלמנטים:

$$D_3^H = \frac{1}{\sqrt{8}} \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & -1 & -1 & -1 & -1 \\ \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sqrt{2} & \sqrt{2} & -\sqrt{2} & -\sqrt{2} \\ 2 & -2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 2 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2 & -2 \end{bmatrix}$$

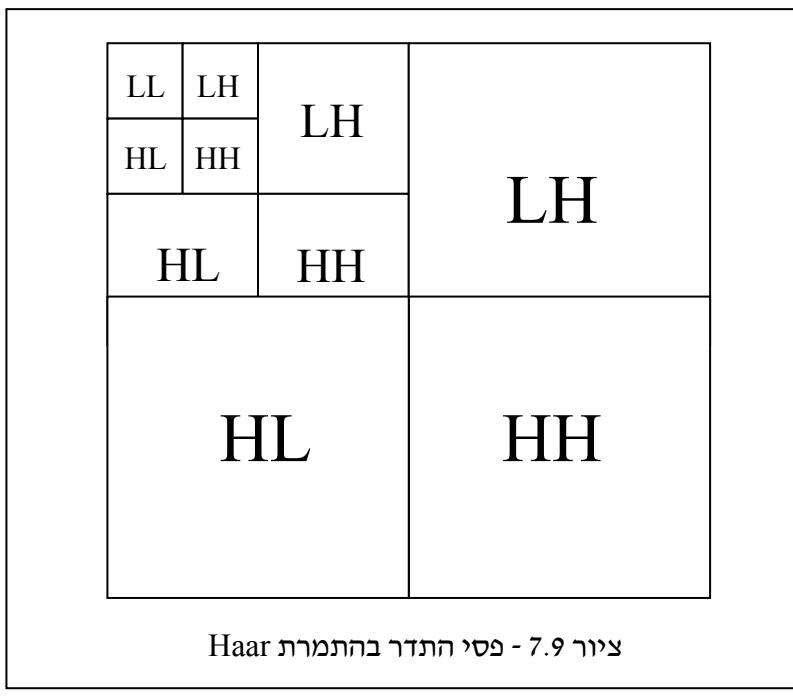
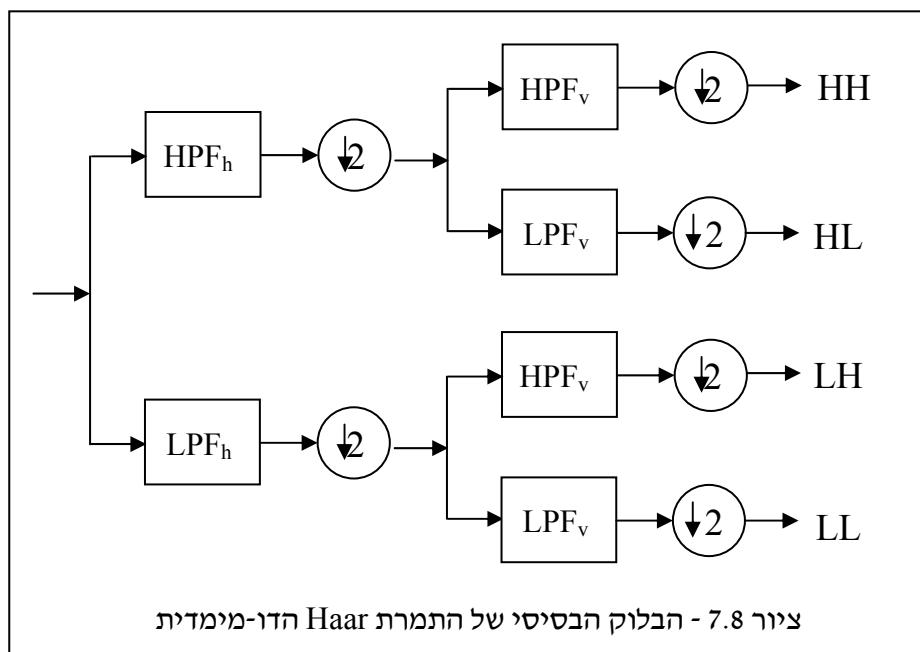
תכונה מעניינת וחשובה של מטריצת ההתרמה היא ששורותיה (למעט הראשונה) מהוות תולדה של אותו גל המעוות ע"י הזזה ומתייחה. גלים אלו נקראים "גלוניים" וזו הם התרגומים לשםWavelet. צייר 7.8 מראה גלים אלו כגרפים.



באשר לדו-מימד, כמו בכל המקדים הקודמים, ההתרמה על תמונה בגודל $M \times N$ מתקבלת ע"י הפעלה באופן ספרבילי. זהה התמורה יוניתרית קלה במיוחד הניתנת לביצוע ע"י $\{MN\}O$ פועלות (בניגוד לכל ההתרמות הקודומות הדורשות $\{MN\}O \cdot \log\{MN\}O$). ביצוע ההתרמה בדרך עילית נעשה ע"י מערכת הסינונים והדילולים שתוארה קודם לכן, עם הבלוק הבסיסי המתואר בצייר 7.8.

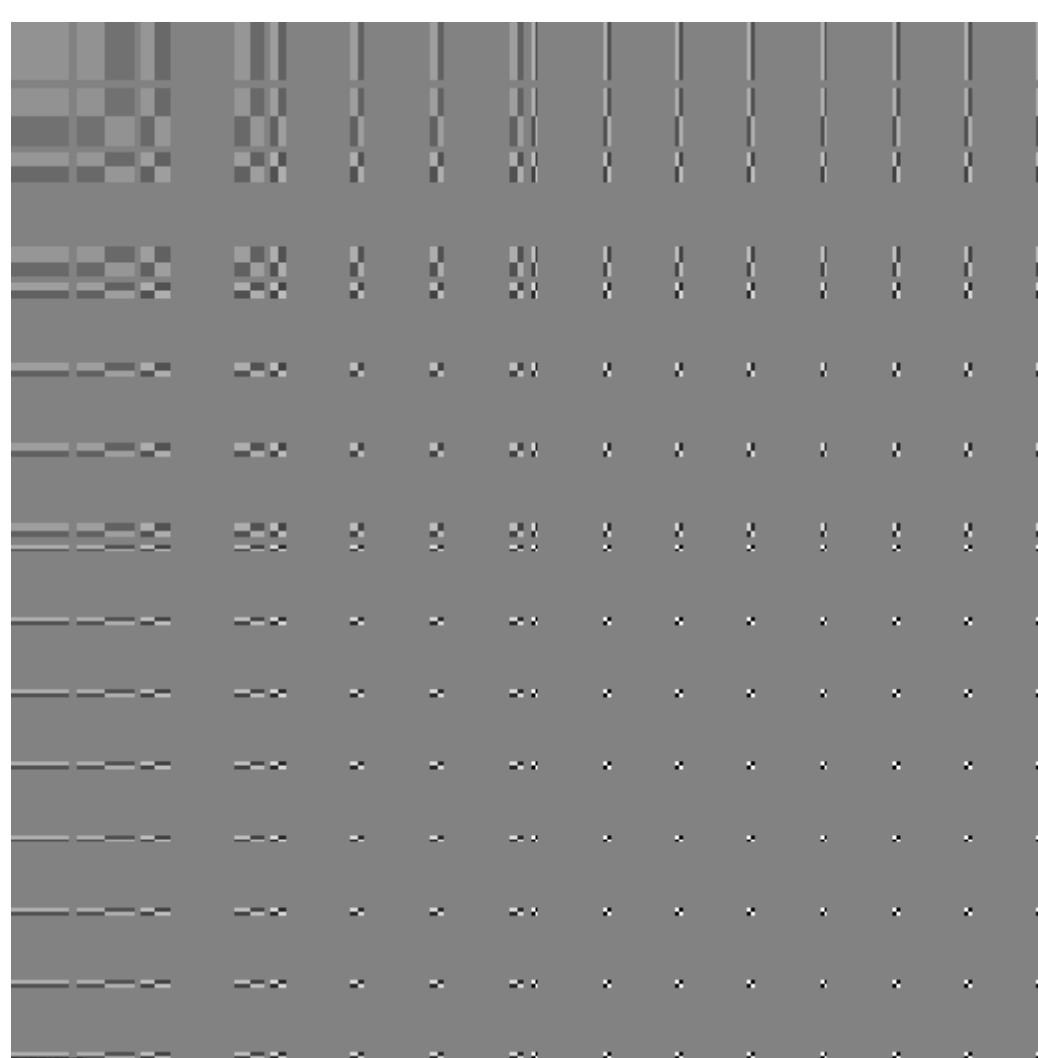
בבלוק זה, לאחר ביצוע סינון LPF ו- HPF על שורות (ולכן הסימונו h-Horizontal), כל מוצא עבור גם סינון LPF ו- HPF על עמודות. אפקטיבית, מתקבל כי תחום התדר מתחלק ארבעה אזורים שוויים המסומנים ב- LL, LH, HL ו- HH. באזור ה- LL המתייחס לתדרים הנמוכים מתבצעת שוב חלוקה לארבעה אזורים וחוזר שוב ושוב. צייר 7.9 מראה כיצד בסופו של דבר נראה החלוקה להתרמה בת שלושה שלבים (המתיאחסת לתמונות בגודל 8 על 8).

אם נתייחס לבлок LH הגדל, בלוק זה נבנה ע"י סינון תמונה המקורי במסנן LPF אופקי ו- HPF אנכי. כיוון שכעת התוצאה ממלאת רק רבע תחום בתדר, היא מדוילת ביחס 1:2 בכל ציר (כלומר צורכת רביע מכם הפיקסלים בתמונה המקורי). לכן, לתמונה של 8 על 8, בלוק זה מוחזק במטריצה של 4 על 4. באופן דומה, בלוק ה- LH הפנימי יותר מוחזק ע"י 2 על 2, וכו'.



פונקציות הבסיס של התמורה דו-מימדית זו מתוארים בציור 7.10. ניכר כי פונקציות אלה שונות ב��ילית מכל התמורות אחרות אותן פגשו קודם. הייחודה בפונקציות בסיס אלה הוא עובדת העובדה דילולות, כולמר תמונות בהן מעט איברים שונים מאפס. למעשה, אנו רואים כאן גלונים (Wavelet) הדומים זה לזה עד כדי מיקום וגודל.

בכל האמור עד כה התקדנו בתתمرة Haar. התמורה זו מהויה מקורה פרטיו של משפחת התמורות Wavelet. בתתمرة Wavelet כללית מוחלף המסלן [1] אשר שימש ל-HPF במסנן אחר, ארוך



ציור 7.10 - פונקציות הבסיס של התמרת Haar לתמונות בגודל 16 על 16

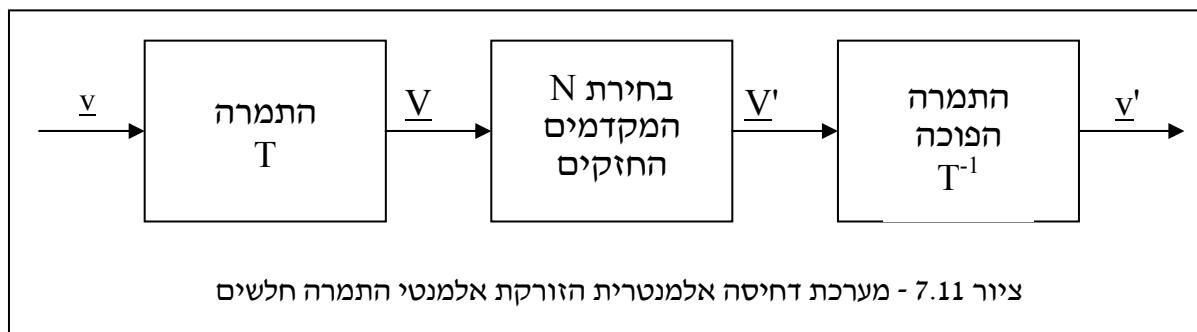
יוטר בדרך כלל, וכך שיכול לבצע הפרדה טובה יותר של פסי התדר. מסנן זה הנקרא "פונקציה האם" יוצר את כל מאפייני ההתמרה, וממנו נגזרות כל פונקציות הבסיס של ההתמרה. תכונו מסנן HPF כזה שייתן כי ההתמרה הכוללת היא אורתו-נורמלית אינה בעיה כלל, וישנן פונקציות מסוימות ידועות היכולות להיות פתרון.

לטיכום, התמרת Haar שונה באופייה מההתמורות הקודמות בהן עסקנו. זהה התמרה יוניטרית ומשנית. התמרה זו ניתנת לביצוע מהיר (ב- $O(M^2)$ פעולות) בשל העובדה דיליה. בנוסף לקלות ביצועה הנומירי, יש לה תכונה אטרקטיבית חשובה שלא נידונה כאן - היכולת לטפל באזות המשנים את אופיים הספקטרלי במקומם באופן נאות.

7.6 התמורות תלויות תמונה

עד כה עסקנו בתמורות קבועות שאינן תלויות בתמונה עליה הן מבוצעות. נשאלת השאלה - האם ניתן להרוויח משלו מבנית התמורות הקשורות לתמונה בה הן מטפלות. נציג אם כך את הבעיה הבאה - נתנו לנו סט של תמונות ורצוננו לתקן את ההתרמה הטובה ביותר עבורן. ברור כי מיד עולה השאלה - מהו קритריון הטיב בו נשמש? במסגרת זו נציג תוכנה חדשה בה לא עסקנו עד כה - ייצוג קומפקטי ע"י חלק מסויסי ההתרמה. תוכנה זו תשמש אותנו לקביעת טיבת התמורה.

הרעינו הוא הבא - נקבל וקטור (אשר יכול להיות תמונה לאחר סידור לקסיקוגרפי) \underline{v} שאורכו M , וביצע עליו התמורה T , לקבלת \underline{v}' . לאחר ההתרמה ניקח את M איברי הוקטור \underline{v}' ומ恣ום נבחר את N (ברור כי $M \leq N < 0$) האיברים החזקים, ונאפס את האחרים. פעולה זו מסומן ב- \underline{X}_N כלומר, הוקטור מוכפל במטריצה בגודל M על M הדומה למטריצת יחידה I , למעט העובדה שrank N איברים ראשונים על האלכסון הם 1. את התוצאה שתסומן ב- \underline{v}'' נזין להתרמה הפוכה של T , ונקבל וקטור המתאים להיות קירוב של \underline{v} , ולכן יסומן ב- \underline{v}'' . ציור 7.11 מראה תהליך זה בסכימת בלוקים. התמורה מוצלחת תהישב התמורה שלכל N שנבחר ניתן כי \underline{v}' ו- \underline{v}'' קרובים ככל האפשר.



בעיה אחרת לכואורה שונה היא הבעיה הבאה - נתיחס לסדרה בציור 7.11 ונדרוש כי לאחר איפוס $N-M$ האיברים החלשים, וקטור התוצאה \underline{v}'' יהיה בעל האנרגיה המרבית האפשרית. ברור כי אם T תיבחר להיות יוניטרית, אנרגיה זו תהיה קטנה מאנרגיית וקטור המקור \underline{v} , ולכן דרישת זו דומה במקצת לדרישת הקודמת כיוון ש- T נסה להיות כזו שתרכז את האנרגיה של הוקטוריים המוזנים לה במקדים הראשונים של התמורה.

נציג שתי בעיות אלו באופן פורמלי - סט התמונות (למעשה וקטורים) עבורן מתוכנת התמורה האופטימלית יסומן ב- $\{\underline{v}_1, \dots, \underline{v}_L\}$. עבר N כלשהו נקבע כי:

$$\underline{v}' = I_N T \underline{v}, \quad \underline{v}'' = T^{-1} I_N T \underline{v}$$

לכן, הבעיה הראשונה אותה הגדרנו תהיה:

$$\underset{T}{\text{Minimize}} \quad \varepsilon_1^2 \{T\} = \sum_{k=1}^L \left| T^{-1} I_N T \underline{v}_k - \underline{v}_k \right|^2$$

כמשמעות הביטוי היא שלכל אחד מהוקטורים עבורם נרצה את ההתמרה, על ההפרש בין כניסה המערכת למוצאה להיות קטן ככל האפשר. עבור הבעיה השנייה שהגדכנו קיבל את בעיית האופטימיזציה הבאה:

$$\underset{T}{\text{Maximize}} \quad \varepsilon_2^2 \{T\} = \sum_{k=1}^L \left| I_N T \underline{v}_k \right|^2$$

כשכאן אנו דורשים אנרגיה מרבית לאחר ההתמרה והקיצוץ.

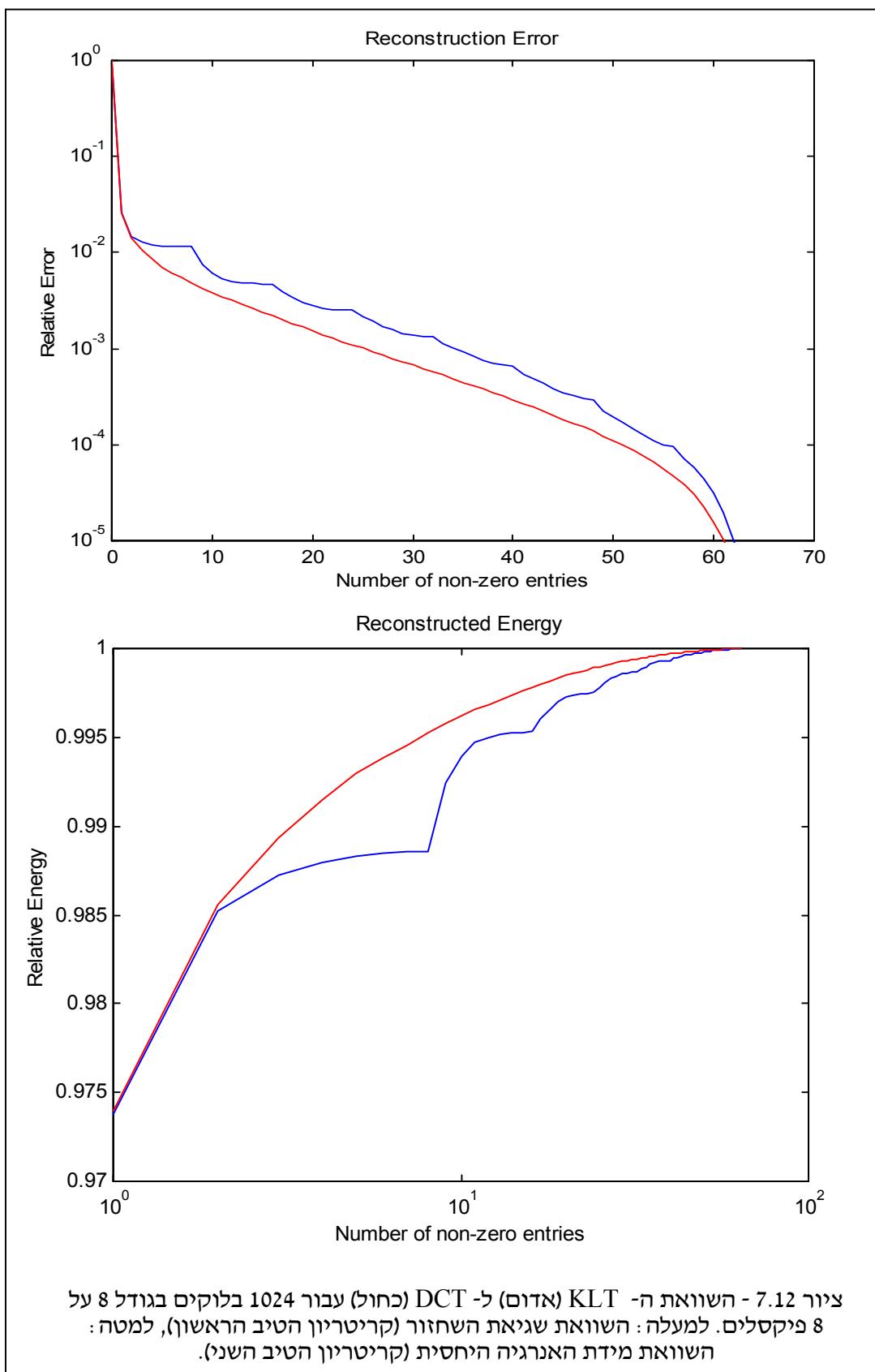
התקבלו שתי בעיות אופטימיזציה מאוד מעניינות אשר לא נפתרו במסגרת זו! מסתבר כי הפתרון לשתי הבעיהות זהה. ההתמרה האופטימלית בשני המובנים הניל' היא המטריצה היוניטרית המלכנת את המטריצה הבאה:

$$R = \sum_{k=1}^L \underline{v}_k \underline{v}_k^H$$

כאשר יש לסדר את עמודותיה לפי סדר יורך של הערכים העצמיים. לכן, קיבלנו ההתמרה יוניתריה חדשה התלויה בתמונות עליהן היא מבוצעת. כאשר אנסמל התמונות איסופי ומיצג את חוק פילוג התמונות, ההתמרה זו קרויה ההתמרה KLT - על שם של Loeve Karhunen ו- אשר הציעו אותה לראשונה. למעשה, כאשר מדובר באנסמל כזה, המטריצה R היא מטריצת האוטוקורלציה של הוקטור האקראי \underline{v} .

אם ההתמרה זו מעשית? מסתבר כי במקרים רבים לא! הצורך לחשב ההתמרה לתמונות מהוות מעסמה כבדה ומטרידה. בධיסת תמונות, שם נעשה שימוש ניכר בהתמורות, גישה זו גם אינה אפשרית כיון שיש לשדר למקלט גם את ההתמרה בה נעשה שימוש, וזה הורס את חס הדחיסה. שאלת טבעיות המתעוררת אם כן היא האם למשפחות מסוימות של תמונות ניתן לקרב את ה- KLT ע"י ההתמרה פשוטה ומכירת? למצלנו התשובה לכך היא חיובית. לתמונות טבעיות, עבר בלוקים קטנים יחסית (עד 16 על 16 פיקסלים), ההתמרה ה- DCT מהוות קירוב טוב מאוד להתמרה ה- KLT. אין זו הפתעה כיון שכבר אמרנו כי DCT מלכון בקירוב מטריצות בלוק-מרקוביות. מסתבר כי בלוקים קטנים זהו בדיקת המבנה שיש למטריצה R. בשל האמור לעיל, ה- DCT מהוות ההתמרה פופולרית מאוד בעיבוד תמונה בכלל ובධיסת תמונות בפרט.

בצירוף 7.12 מובאת השוואה בין ה- DCT ו- KLT. השוואה זו מתყבשת ל- 1024 בלוקים של תמונות בגודל 8 על 8 שנלקחו מהתמונה Lena. בהשוואה זו מוצגים שני קריטריוני הטיב אותן הצגנו, וניכר כי על-פי שניהם ה- KLT טוב יותר, אך גם ניכרת קירבה גדולה בין השניים.



Over-Complete התמורות

בפתיחה פרק זה העלנו את האפשרות לעשות שימוש בהtamורות ליניאריות בהן יש יותר פונקציות בסיס מאשר נקודות בתמונה. מסנו את כמות הפיקסלים (אורך וקטור הכניסה להtamורה) ב- M , ואת מספר הבסיסים כ- N , וכן, $M > N$. בסיסים מסוג זה נקראים Over-Complete. ברור כי במקרה כזה וקטורי הבסיס תלויים ליניארית, ושנה יתרות. נסמן ב- B מטריצה המכילה את וקטורי בסיס אלה כעמודותיה - למטריצה זו יש יותר עמודות משוראות (היא מלכנית אופקית). בהינתן וקטור \underline{v} באורך M , התמורתו תהיה הוקטור \underline{V} באורך N . במקרה כזה הקשר המתאר את התמורתה הפוכה הוא :

$$\underline{v} = \sum_{k=1}^{MN} V_k b_k = B\underline{V}$$

כלומר, ביצוע התמורתה הפוכה פשוט למדי - איברי \underline{V} מהווים את מקדמי הצירוף הליניארי של פונקציות הבסיס לבנייתה של \underline{v} . הסוגיה המרכזית היא כיצד לקבוע את \underline{V} , ככלומר לבצע את התמורתה הישרה (בניגוד להפוכה). מסתבר כי תהליך זה מורכב יותר הפעם. אם נתבונן בקשר הנ"ל, הוא מייצג מערכת ליניארית בת M משואות ו- N גורמים, ולכן ישנו אינסוף אפשרויות של פתרונות \underline{V} . כיוון שיש אינסוף גורמים, מהו הפתרון מביניהם שאוטו נעדיף? נראה מספר אפשרויות ונדון בהם.

בחירה אפשרית אחת היא להעדיף את הפתרון \underline{V} שיהיה גם בעל אנרגיה מינימלית. סיבה לבחירה זו היא העדפה של ערכאים נמכרים באיברי \underline{V} משיקולים שונים. במקרה כזה יהיה علينا לפתור את בעיית האופטימיזציה הבאה :

$$P_1 : \underset{\underline{V}}{\text{Minimize}} \quad \|\underline{V}\|^2 \quad \text{Subject To: } B\underline{V} = \underline{v}$$

פתרון בעיה זו ייעשה ע"י כופלי לגרנזי ומתקבל :

$$L\{\underline{V}, \lambda\} = \|\underline{V}\|^2 + \lambda^T [B\underline{V} - \underline{v}]$$

$$\begin{aligned} \Rightarrow 2\underline{V} + B^T \lambda &= 0 \Rightarrow \underline{V} = -\frac{1}{2} B^T \lambda \\ \Rightarrow -\frac{1}{2} BB^T \lambda &= \underline{v} \Rightarrow \underline{V} = B^T [BB^T]^{-1} \underline{v} \end{aligned}$$

בשורה ראשונה כתבנו את פונקציית הלגרנזי'ין כצירוף של פונקציית המחיר והאלומות המוכפלים בוקטור כופלי לגרנזי'. בשורה שנייה גזרנו לפי \underline{V} והשוונו לאפס, וכך התקבל ביתוי לפתרון התלו依 בקופלי הלגרנזי'. בשורה תחתונה הצנו את התוצאה לאלוץ לחילוצו של וקטור כופלי הלגרנזי' וככך התקבל

הפתרון המלא לבעה. התוצאה שהתקבלה היא בפועל היפוך מוכלל של B (Pseudo-Inverse). נשים לב כי החתמונה ליניארית - כלומר בנית \underline{V} מותך ע"י הכפלת במטריצה כלהי.

אפשרות שנייה לביצוע החתמונה היא העדפת וקטור \underline{V} חלק. סיבה לבחירה זו יכולה להיות תהליכי דחיסה שיבוצע עליו לאחר מכן (שיהיה אולי מוצלח יותר אם הווקטור יהיה חלק). בעית האופטימיזציה הפעם תהיה:

$$P_2 : \underset{\underline{V}}{\text{Minimize}} \|D\underline{V}\|^2 \text{ Subject To : } B\underline{V} = \underline{v}$$

כאשר D אופרטור גזירה דיסקרטי. נניח כי D בנויה כך שאינה סינגולרית. פתרון בעיה זו ע"י כופלי לגרנזי נתונים:

$$\begin{aligned} L\{\underline{V}, \underline{\lambda}\} &= \|D\underline{V}\|^2 + \underline{\lambda}^T [B\underline{V} - \underline{v}] \\ \Rightarrow 2D^T D\underline{V} + B^T \underline{\lambda} &= 0 \Rightarrow \underline{V} = -\frac{1}{2} [D^T D]^{-1} B^T \underline{\lambda} \\ \Rightarrow -\frac{1}{2} B [D^T D]^{-1} B^T \underline{\lambda} &= \underline{v} \Rightarrow \underline{V} = [D^T D]^{-1} B^T [B [D^T D]^{-1} B^T]^{-1} \underline{v} \end{aligned}$$

קיבלו הפעם הכללת התוצאה הקודמת, כשהמטריצה $D^T D$ מהוות מעין מטריצה משקל. גם הפעם החתמונה ליניארית ויש מטריצה המייצגת אותה.

אפשרות אחרתונה שאותה נזכיר היא האפשרות לבנות החתמונה דיליה. הרעיון הוא לחפש מבין הפתרונות את הפתרון בו מרבית איברי החתמונה הם אפס. סיבה טובה להעדיף וקטוריים דיליליים הינה כיון שהוא יציג קומפקטי ונוח, וייצוג זה גם טוב לדחיסה כפי שניתן לראות מהדוגמה הבאה. ישנה טענה רוחחת כי מערכת הראייה האנושית (במוח האדם) מבצעת חתומות מעין אלו לשם ייצוג טוב של המידע הנקלט במוח.

7.2: אם נניח כי לוקטור כניסה באורך 64 אברים בני 8 סיביות כל אחד נעשו שימוש בסיס בן ²⁰ 2 פונקציות בסיס, ונצליח (ע"י קירוב) למפות כל וקטור כניסה לוקטור בסיס אחד ויחיד (כלומר, כל יתר מקדמי הצירוף הליניארי הם 0), הרי שלצורך שידור וקטור הניסה נזדקק ל- 20 סיביות בלבד, במקומות 512 במקור (יחס דחיסה של 0.04)

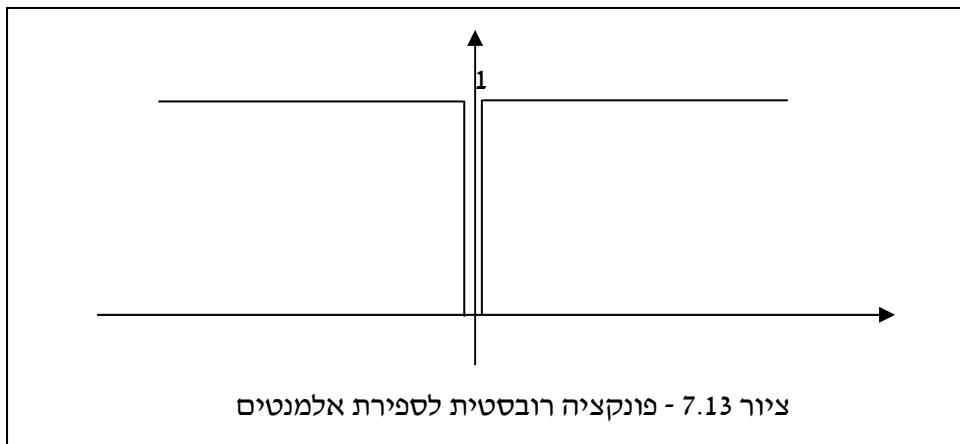
בטרם ניגש לשאלת כיצד מוצאים את החתומה הדיליה ביותר, חשוב להציג רעיון בסיס חשוב שנובע מהאמור לעיל. כשהצגנו את ה- KLT, טענו כי הוא האופטימלי במובן של ייצוג קומפקטי, אך שם הגדרת המושג "ייצוג קומפקטי" הייתה שונה בתכלית. בעוד שם רצינו של חלק מוקטור החתומה נצליח להעביר ככל האפשר יותר מהמידע הנוכחי, כאן ההגדרה מתרכזת בחתומות מסווג Over-Complete והציג הוא על דليلות וקטור הייצוג. ניסיונות נומריים למצוא את החתומה האופטימלית במובן חדש זה

הראו כי מתקנים בסיסים דמיי Wavelet, ותמונה זו מהווה חיזוק משמעותית לטענה כי התמורה היא אטרקטיבית לייצוג תמונות.

נחוור לבעה של בניית התמורה דليل. עליינו לכתוב פונקציה מחר ש-”תמדווד” את מידת הדילות של הוקטור \underline{v} . סוגיה זו מחייבת אותנו לשערוך רובסטי ולפונקציות $\rho(\underline{v})$. בהינתן וקטור \underline{v} , ברור כי הפונקציה

$$\sum_{k=1}^N \rho(\underline{v}_k)$$

תחווה אומדן טוב למידת דילולתה אם $\rho(\underline{v})$ נראית כבציר 7.13.



לכן, בעיית האופטימיזציה המוצעת היא:

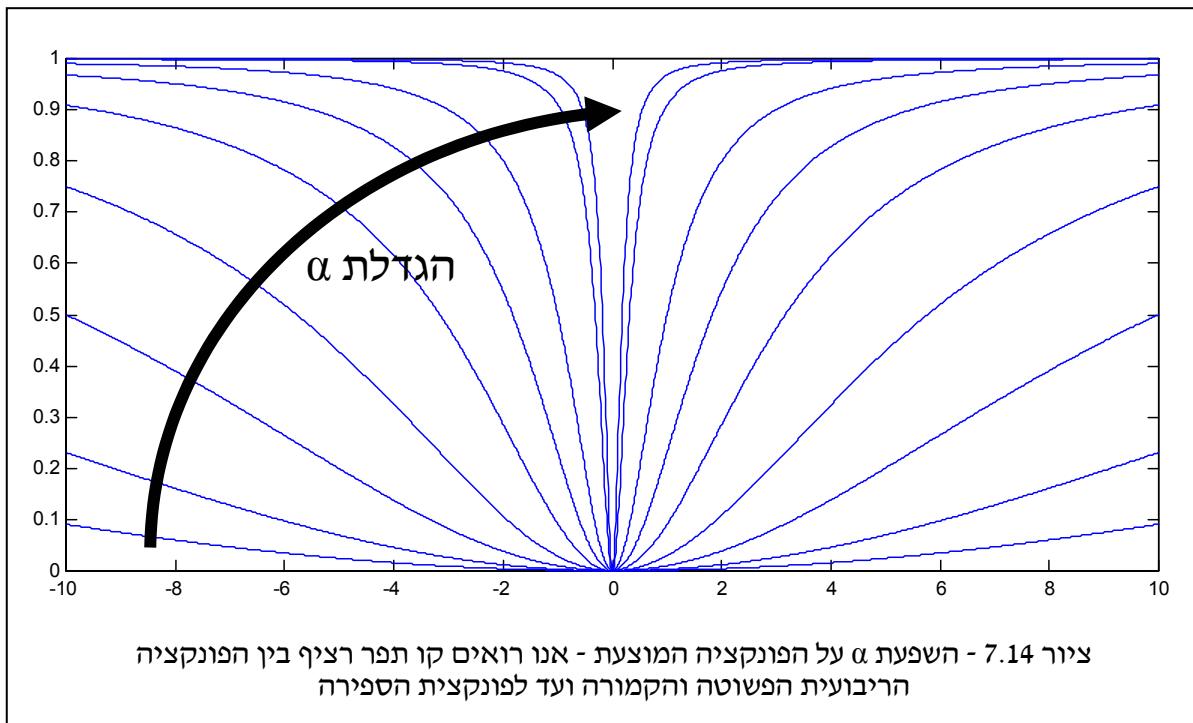
$$P_3 : \underset{\underline{v}}{\text{Minimize}} \sum_{k=1}^N \rho(\underline{v}_k) \text{ Subject To: } B\underline{v} = \underline{y}$$

הפונקציה המוצעת קשה מאוד לפתורו איטרטיבי בשל העובדה מאד לא קמורה. לכן, בסביבות גבוהה יתקבל כי ניפול לפתורו שיווה נקודת מינימום לקליטת של הפונקציה הניל. על מנת להתגבר על בעיה זו ניתן לעשות שימוש בפונקציה $\rho(x) = |x|$. בחירה זו תהפוך את הבעיה לקמורה, ולאחר בעלת פתרון ייחיד קל חישית לקבלה, ולא נקודות מינימום לוקליות. מסתבר כי בחירה זו גם מביאה פתרון דليل וגם ניתנת לייצוג כבעית תכונות ליניארי (זוהי משפחה של בעיות אופטימיזציה בעלות מבנה ייחודי המאפשר פתרון יעיל וחסיט).

אם הבחירה הניל של $\rho(x) = |x|$ אינה משביעת רצון, ניתן להציג תהליך הנקרא Non Convexity (GNC) לשם פתרון הבעיה תוך ניסיון לחמוק מנקודות מינימום מקומיות. הרעיון הוא לעשות שימוש בפונקציה מהצורה:

$$\rho(x) = \frac{\alpha x^2}{1 + \alpha x^2}$$

כאשר פותרים תחילה עם α קטן מאוד. עבור אפס נקבל את הפתרון של בעיית המינימום נורמה - P_1 . לאחר הפתרון נגדיל מעט את הערך של α ונפתרו שוב, וחזור חלילה. מכיוון שהתחלנו בעיה קמורה לאות אנו מעוניינים אותה לפונקציה שאינה קמורה, אנו מושגים אתחול טוב בכל שלב עבור השלב הבא, ולכן יש סיכוי טוב שהפתרון יגלוש לכיוון של תוכאה דיליה מקסימלית. נשים לב כי עבור $\infty \rightarrow \alpha$ אנו מושגים פונקציה שסופרת אלמנטים שונים מאפס. ציור 7.14 ממחיש זאת.



פרק 8 – דחיסת תМОנות

175	פרק 8 – דחיסת תМОנות
175 8.1 יסודות בתורת האינפורמציה
181 8.2 יתרוות בתמונות וניצולן
184 8.3 קידוד אנטרופיה - אלגוריתם הופמן
188 8.4 קידוד עם שגיאות - גישת החיזוי
191 8.5 קידוד עם שגיאות - גישת ההתמרה

בפרק זה אנו נכנסים לתוךם של דחיסת תМОנות. כבר בפרק ראשון עמדנו על הבעייתיות של נפח האופייני של תМОנות. תMOנה ממוצעת בגודל של 1000 על 1000 פיקסלים בצבע צורכת נפח של 3MB. הרעיון המרכזי מאחריו מגוון השיטות והאלגוריתמים שיידנו בפרק זה הוא שפח זה אינו מחייב המיציאות. בדחיסה אנו מבחנים בין דחיסה לא אובדן, בה לאחר דחיסה ופרישה מתקבלת אותה תMOנה בדיק, ובין דחיסה אובדןית, בה מתרירים שגיאות קלות לשם השגת רוח ביחס לדחיסה. כפי שנראה במסגרת פרק זה, ניתן לפחות מהנפח המלא לאחסן את אותה תMOנה בדיק, ונitin בהרבה פחות לאחסן תMOנה דומה לה מאוד. השאלה היא אם אכן עושים זאת, ועל שאלה זו ננסה לענות במסגרת פרק זה.

8.1 יסודות בתורת האינפורמציה

תורת האינפורמציה הינה תורה חדשה יחסית אשר פותחה לראשונה ע"י Shannon בשנות הארבעים של מאה זו. בסיסically, תורה זו בא לכתות מידע, ולא פיין מערכות בהקשר של העברת מידע. תורה זו מhoeva את התשתית התיאורטית לנושא דחיסת מידע בכלל ודחיסת תMOנות בפרט. ברור כי במסגרת זו לא ניתן יותר מסקירה ראשונית לא מוצה בשלה נפח העצום של תחומי זה וריבוי התוצאות בו.

נתחיל עם מושג האנטרופיה – מידת אי-זדאותו של מקור מידע. נתון לנו מקור המייצר מספרים אקראיים דיסקרטיים x חסרי תלות ביניהם (d.i.i.) עם חוק פילוג (x).P. האנטרופיה של מקור זה מוגדרת ע"י הביטוי:

$$H(x) = \sum_x P(x) \log_2 \frac{1}{P(x)} \left[\frac{\text{bits}}{\text{symbol}} \right]$$

משמעות גודל זה היא כמות הסיביות הניתנת לייצוג מדויק שלתו מקור זה בפועל. נראה מספר דוגמאות להמחשת מידת זו.

דוגמה 8.1: מקור שמייצר מספרים שלמים בתחום [1,16] בהסתברות זהה - 1/16 יהיה בעל אנטרופיה

$$H(x) = 16 \cdot \frac{1}{16} \log_2 16 = 4 \text{ bits/symbol}$$

תוצאה זו אינטואטיבית לנו כיון שעבור 16 אפשרויות, ניתן לייצג כל אחת ב- 4 סיביות, כפי שהתקבל בפועל.

דוגמה 8.2: עבור אותו מקור כשהפעם ההסתברות נתונה כ- $P(x) = x/136$ (המספר 136 מנורמל את ההסתברות לקבל סכום הסתברויות שווה 1), האנטרופיה תהיה:

$$H(x) = \frac{1}{136} \left[1 \cdot \log_2 \frac{136}{1} + 2 \cdot \log_2 \frac{136}{2} + \dots + 16 \cdot \log_2 \frac{136}{16} \right] \approx 3.76 \text{ bits/symbol}$$

תוצאה זו מעניינת כיון שעדיין יש לנו 16 אפשרויות, אך לטענת התיאוריה, די ב- 3.76 סיביות לתיאורן באופן מלא. ניתן למצוא בדבר הגיון אם נחשוב על הדברים באופן הבא - כאשר כל ההסתברויות זהות, אי הودאות היא מירבית, לעומת, אין לנו כל דרך לדעת משווה בטרם קיבלנו את התו. עבור הסתברות לא אחידה, עוד בטרם קיבל את התו נוכל לומר כי ישנה נטייה חזקה לכיוון הערכים הגדולים, בהיותם בעלי הסתברות גבוהה יותר. לכן, מידת אי-הודאות המתקבלת נמוכה יותר.

דוגמה 8.3: נניח כי ל- 16 התווים האפשריים, הסתברותו של 1 הוא 1, ושל כל היתר 0. במקרה כזה האנטרופיה תהיה:

$$H(x) = 1 \cdot \log_2 1 + 15 \cdot 0 \log_2 0 = 0 \text{ bits/symbol}$$

אין כאן כל הפתעה: במקרה זה, לפני שנקבל את התו נדע בדוק מה יקרה, ולכן אין כל אי-ודאות.

למעשה, דרך דוגמה זו המחשנו שתי תכונות חשובות של פונקציית האנטרופיה:

1. למקור בעל N תווים, האנטרופיה H חסומה מלמעלה לפי $H \leq \log_2 N$. שיוויון מתקיים כאשר ההסתברות זהה ושויה $1/N$ לכל تو.

2. למקור בעל N תווים, האנטרופיה H חסומה מלמטה לפי $-0 \leq H$. שיוון מתקיים כאשרתו אחד בעל הסתברות 1, וכל יתר התווים בעלי הסתברות 0.

משפט הצפינה של Shannon קובע כי עבור מקור בעל אנטרופיה $(x)H$, לא ניתן לדחוס מידע ממנה לפחות מהאנטרופיה ללא הכנסת שגיאות. יתרה מזו: לכל ϵ חיובי וקטן כרצוננו, קיימים **בaprion** אלגוריתם דחיסה אשר יתנו דחיסת המקור לקצב של $\epsilon+(x)H$. תוצאה זו קובעת כי האנטרופיה מהוות אומדן לביצועו של אלגוריתם דחיסה, אם רצוננו לדחוס ללא שגיאות. משפט זה של Shannon לא מציע את אלגוריתם הדחיסה עצמו, אלא מוכיח את קיומו באופן תיאורטי. בהמשך נכיר דרכי מעשיות למימוש טענה זו.

חשוב להבהיר כי מושג האנטרופיה אינו מוגבל לטיפול בתווים בודדים. לדוגמה, עבור מקור המיצר מספרים בתוחום $[1,16]$, ניתן להתייחס לצמדים תווים כ-"**סופר-תווים**" חדשים, ואז יש לנו 256 סוגי תווים אפשריים, וניתן לכל צירוף קבוע את הסתברותו, לחשב אנטרופיה כ precedent וגו'. תהליך זה חשוב ונחוצץ כאשר יש קשר סטטיסטי בין התווים. נמחיש זאת דרך דוגמה.

דוגמה 8.4: עבור המקור שתואר בדוגמה 8.1, לκיחת צמדים תביא לכך שיהיו 256 תווים אפשריים שווי הסתברות. האנטרופיה תהיה :

$$H([x_1, x_2]) = 256 \cdot \frac{1}{256} \log_2 256 = 8 \text{ bits/symbol}$$

נזכיר כי הפעםתו אחד הוא בעצם שני תווים, ולכן אין כל שינוי ביחס לטיפול בתווים בודדים.

דוגמה 8.5: נתון מקור המיצר את המספרים +1 ו-1 לפי חוק הפילוג הבא :

$$P(x_k / x_{k-1}) = \begin{cases} 0.8 & x_k = x_{k-1} \\ 0.2 & x_k \neq x_{k-1} \end{cases}$$

כלומר, לדוגמה, אם התו הקודם היה -1, סבירות גבוהה יש להיותו התו הבא -1 גם כן. בטיפול בתווים בודדים נקבל כי $P(+1) = P(-1) = 0.5$ כיון ש :

$$\begin{aligned} P(+1) &= P(+1/+1)P(+1) + P(+1/-1)P(-1) \\ P(-1) &= P(-1/+1)P(+1) + P(-1/-1)P(-1) \\ \Rightarrow &\begin{bmatrix} P(+1/+1)-1 & P(+1/-1) \\ P(-1/+1) & P(-1/-1)-1 \end{bmatrix} \begin{bmatrix} P(+1) \\ P(-1) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \Rightarrow P(+1) &= P(-1) = 0.5 \end{aligned}$$

לכן האנטרופיה תהיה :

$$H(x) = 0.5 \log_2 2 + 0.5 \log_2 2 = 1 \text{ bits/symbol}$$

עבור טיפול בצדדי תווים נקבל כי חוק ההסתברות הוא $P([+1,+1]) = P([-1,-1]) = 0.4$ ו- $P([-1,+1]) = P(+1,-1) = 0.1$, כיון ש:

$$P([x_k, x_{k-1}]) = P(x_k / x_{k-1})P(x_{k-1}) = \begin{cases} 0.8 \cdot 0.5 & [+1,+1] \\ 0.2 \cdot 0.5 & [+1,-1] \\ 0.2 \cdot 0.5 & [-1,+1] \\ 0.8 \cdot 0.5 & [-1,-1] \end{cases}$$

במקרה זה האנטרופיה תהיה:

$$H([x_1, x_2]) = 2 \cdot 0.4 \log_2 2.5 + 2 \cdot 0.1 \log_2 10 = 1.7219 \text{ bits/symbol}$$

כך שלתו יחד נדרשים לפחות מסיבית אחת, וברור כי הושגה עילוות גבוהה יותר כי נוכל לחבר בין התווים. למעשה, ניתן המשיך במגמה זו ולטפל בשלשות, רביעיות וכו', כצפוי שיפור כל שעובדים עם בלוקים גדולים יותר.

למעשה, יכולנו לנצל את הקשר הסטטיסטי בין התווים המקוריים בצורה יעילה יותר لو פעלנו על ההסתברויות המותניות בצורה ישירה. נניח כי אנו נמצאים בתו ה- L וכבר קיבלנו אותו בمفענה, ונניח כי הוא היה $+1$. לכן, בבואה לדוחס (ובמקלט - לפטוס) את התו הבא, אנו יודעים כי בסבירות גבוהה הערך הבא הוא $+1$ גם כן. למעשה, האנטרופיה של התו הבא תהיה:

$$H(X_{L+1}) = 0.8 \log_2 \frac{1}{0.8} + 0.2 \log_2 \frac{1}{0.2} = 0.7219$$

כך שקיבלו תוצאה טובה יותר מזו של העבודה בצדדים. גישה זו אינה גישה עבודה בבלוקים אלא גישה חלופית של חיזוי - על פי התו הנוכחי שכבר שודר (ונקלט), נדחס התו הבא, שימושים שימוש בהסתברות המותנית.

למעשה, לו קידדנו את המקור הזה בבלוקים שאורכם שואף לאינסוף היינו מקבלים תוצאה זו בדיקות. כאשר אנו עובדים בצדדי תווים אנו מנצלים את הקשר ביניהם, אך לא את הקשר לצמד הקודם או לצמד הבא, ולכן התוצאה פחותה טוביה. כאשר הבלוקים מתארכים אנו מצטחים טוב יותר לנצל את הקשרים בין התווים, כאשר מאבדים רק את הקשרים בתפרים בין הבלוקים. אם הבלוקים באורך אינסופי - אין אובדן. ברור כי הגישה המותנית עליילה יותר לאין-ערוך.

נציג עוד מספר תכונות המאפיינות את פונקציית האנטרופיה, הנוגעת לטיפול בבלוקים:

$$, \quad H([x_1, x_2, \dots, x_N]) \leq \sum_{k=1}^N H(x_k) \quad 1. \text{ קיימים אי-השוויון הבא :}$$

כששוויון מתקיים עבור חוק פילוג פריד - $P([x_1, x_2, \dots, x_N]) = P(x_1)P(x_2) \cdots P(x_N)$ רוחה או לפחות אי-הפסד בעובדה על בלוקים במקומות על תווים בודדים.

$$, \quad H([x_1, x_2, \dots, x_N]) \geq \max_k \{H(x_k)\} \quad 2. \text{ קיימים אי-השוויון הבא :}$$

כששוויון מתקיים עבור מקרה בו כל המשתנים קבועים באופן דטרמיניסטי מתוך אחד מהם. פירוש הדבר הוא שכאשר הינו בלוק של תווים אך למעשה רק אחד מהם אקראי וכל היתר נזירים ממנו בצורה מדויקת, די לקודדתו אקראי זה. במקרה כללי, אם נחשב אנטרופיה לכל אחד מהתווים בנפרד, לא יתכן שהאנטרופיה המשותפת תהיה קטנה מהගורע שביהם (గורע במובן של מכיל הרבה אינפורמציה).

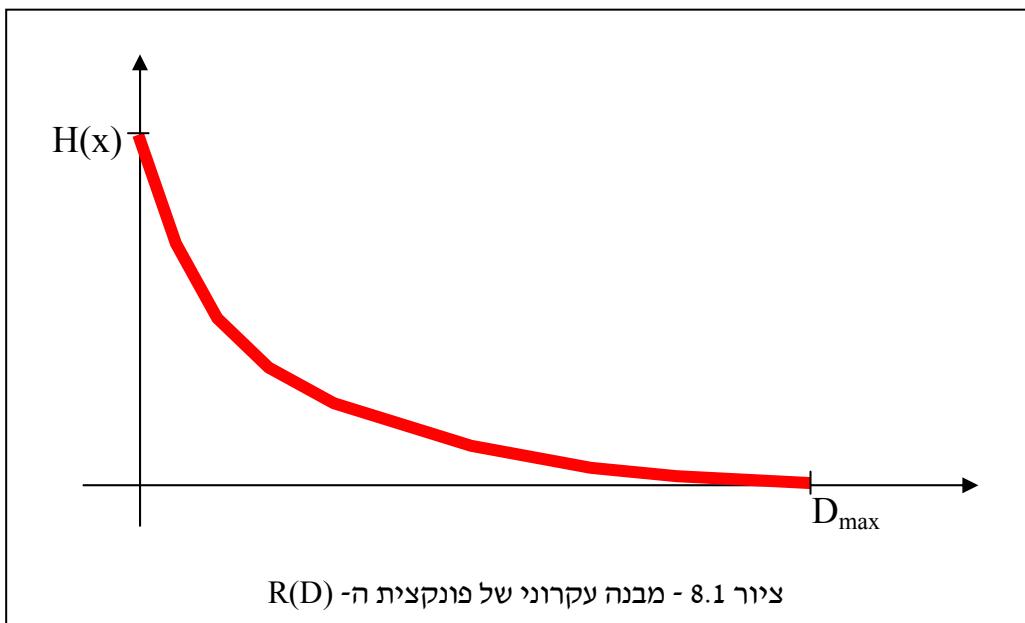
$$, \quad H([x_1, x_2]) \geq 2 \cdot H(x_1/x_2) \quad 3. \text{ קיימים אי-השוויון הבא :}$$

כששוויון מתקיים עבור מקרה בו אין קשר סטטיסטי בין התווים.

עד כה עסקנו בהגדרת מידת האינפורמציה של מקור כאשר רצוננו לדוחסו ללא שגיאות. מה קורה כאשר אנו מוכנים להתייר שגיאות? מסתבר שם לישום זה הצע Shannon אומדן אינפורמציה, הידוע בכינוי פונקציית קצב-היעיות (Rate-Distortion Function) - $R(D)$. הרעיון מכיל בצורה ישירה את מושג האנטרופיה מנקודות באופן הבא. נניח כי אנו מוכנים לקבל כי יוצר מרחק של עד D בין סידורת המקור x והסדרה הנפרשת לאחר דחיסה y . פער זה מסומן ב- $D \leq \{y, x\}_d$ וברור כי עלינו להגדיר מהי המטריקה בה אנו עושים שימוש. אז, מידת האנטרופיה של המקור x תחת שגיאה של עד D תהיה $R(D)$. נציג מספר תכונות של פונקציה זו :

1. עבור $D=0$ אנו בפועל לא מרשימים שגיאה ולכן ברור כי $R(0) = H(x)$.
2. עבור D מעל סף מסוים התלו依 בעוצמתו של x יתאפשר כי אם פשוט נשלח את ממוצע המספרים לכל TWO כניסה תתקבל שגיאה קטנה מ- D . במקרה זה נקבל כי $R(D) = 0$.
3. פונקציית ה- $R(D)$ מונוטונית לא עולה. עלייה לא תיתכן כיון שאז לשגיאה גדולה יותר נדמה כי נדרש קצב גבוה יותר - זה לא יתכן.
4. פונקציית ה- $R(D)$ היא פונקציה קמורה.

চিত্র 7.1. מראה כיצד נראה פונקציית ה- $R(D)$ באופן עקרוני.



חשיבות להבהיר כי בעוד שאנטרופיה של מקור קלה לחישוב, חישובה של $R(D)$ (באופן אליליטי או נומרי) קשה מאוד, ולרוב בלתי אפשרי. חישובה של האנטרופיה דורש את חוק פילוגם של תוכי הממקור, חישוב ה- $R(D)$ דורש בנוסף את הגדרת המטריקה בין אוטות הכניסה והיציאה ממערכת הדחיסה-פריסה.

דוגמה 8.6: נניח כי המקור שלנו מייצר מספרים שלמים בתחום $[0,127]$ בהסתברות אחידה. לכן האנטרופיה של המקור זה היא 7 סיביות לפחות. בעת נניח כי אנו מתירים לקידוד לבצע סטייה של 1, כלומר,

$$D(x,y) = (x-y)^2 \leq 1$$

במילים אחרות, אנו מתירים לסטות ברמה אחת במקלט בהשוואה למשדר. נציג הצעה פשוטה שתגשים שגיאה זו תוך שידור פחות סיביות - ניקח את המספר ונחלף אותו במספר הזוגי הקרוב אליו מלמטה. כך יקרה כי מספרים זוגיים ייוותרו כמותם, ומספרים אי-זוגיים יומרו למספרים זוגיים. כיוון שכעת קיימים רק מספרים זוגיים, נדרשים 6 סיביות במקומות 7. בהתאם דומה, עבור שגיאה של 2 נוכל לבחור את השכן הקרוב ביותר מבין המספרים המתחלקים ב- 3 ($0,3,6,9, \dots, 126$). בעת יש לנו רק 43 ערכים אפשריים שווים הסתברות וכלן האנטרופיה היא 5.246.

נוכל להמשיך תהליך זה ולהציג מוקודד עם שגיאה למספרים אלו, ובעצם לקבל פונקציה המקרבת את ה- $R(D)$. הפונקציה המוצעת היא פונקציה החוסמת מלמעלה את פונקציית קצב-היעיות כיוון שאין כל הבטחה כי לא קיימת אסטרטגיה טובה יותר לשידור עם שגיאות במקרה המוצע עם פחות אנטרופיה.

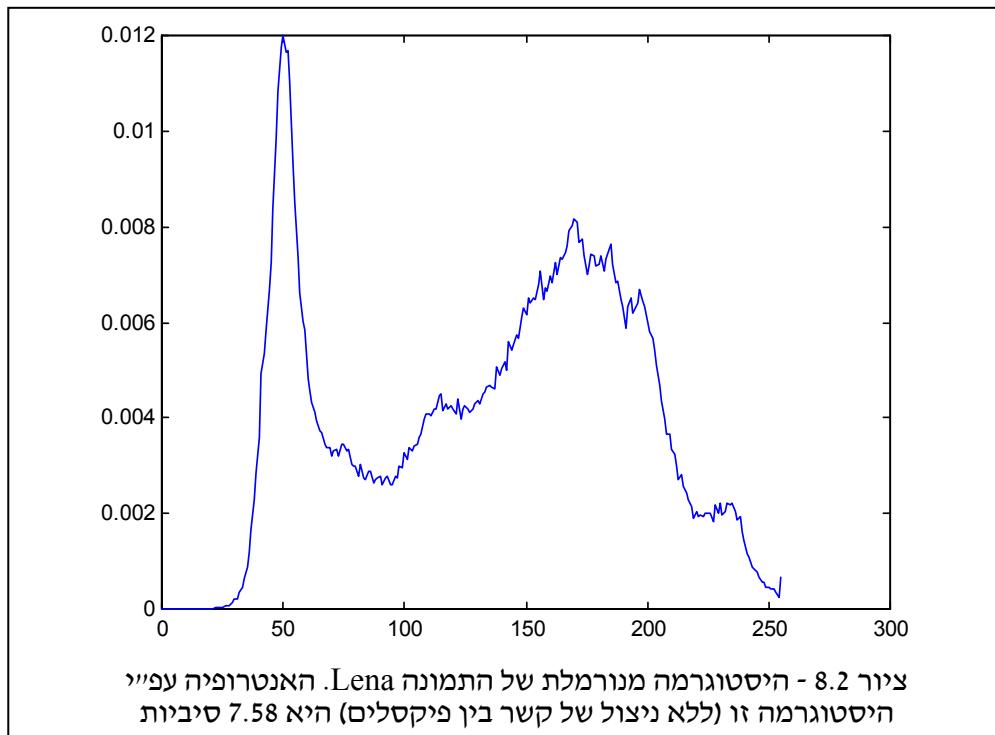
8.2 יתירות בתמונות וניצולן

מהדיוון בנושא האנתרופיה ופונקציית קצב-העיות ברור כי דחיסה מצילהה בשל יתרות באות הנדחס. יתרות פירושה שהמידע מכיל כפליות מיותרות הניתנות לניצול. עד כה ראיינו מספר סוגים יתרויות עקרוניים. נזכיר אותם ונראה כיצד הם ישימים בתמונות:

1. יתרות הסתברותית - אפילו אם המקור שלנו מנביב נתונים מתוך סט נתון של N נתונים אפשריים ובלי תלות ביניהם, אם נתונים שונים הסתברות שונה נקבל כי האנתרופיה נמוכה מ- $N \log_2 N$.
אם נחזור לרגע לתמונות, ונזכיר כי היסטוגרماה של תמונה היא פונקציית ההסתברות של רמות האפור בה, הרי שאם היסטוגרמאה אינה איחודת, ישנה כאן יתרות הסתברותית הניתנת לניצול לשם קבלת דחיסה. למעשה, משיקולים אלו ברור כי ככל שההיסטוגרמאה רחוקה מלהיות איחודת, כך התמונה דחיסה יותר.

דוגמה 8.7: ניקח את התמונה Lena בה יש 512 על 512 פיקסלים המיוצגים ע"י 8 סיביות כל אחד. עפ"י ההיסטוריה המנורמלת שלה נחשב את האנתרופיה. ההיסטוריה מתוארת בציור 8.2, והאנתרופיה המתקבלת היא:

$$H_{\text{Lena}}(\text{pixel}) = \sum_{k=0}^{255} h(k) \log_2 \frac{1}{h(k)} = 7.58 \text{ bits/pixel}$$



כלומר, רק מתוך ההיסטוריה ברור כי ניתן לפחות 8 סיביות לייצג כל פיקסל. כמו כן אנו רואים כי הרוחח אינו גדול - נוכל (טיאורטיות בשלב זה) לחסוך חצי סיבית לפיקסל.

2. יתרות בין תווים - ראיינו כי כאשר עובדים על בלוקים של תווים וקיימת תלות סטטיסטייה ביניהם, מתקבלת עילوت טוביה יותר של הדחיסה. בהשלה לתמונות נאמר כי כאשר אנו עומדים לקודד את ההחלטה - [ז'] בתמונה, אם ננצל את המידע שמשפקים שכנוו של פיקסל זה, נצליח טוב יותר בדחיסה. זאת כיוון שקיימת תלות בין פיקסלים סמוכים בתמונה - הם אמורים להיות בעלי רמות אפור קרובות, ולכן קיימת תלות סטטיסטייה חזקה ביניהם.

דוגמה 8.8: נחזור לתמונה Lena. על מנת לעבוד בבלוקים נידרש לבנות היסטוגרמה של צמדים. היסטוגרמה זו תדרוש כ- 65000 תאים (256 בריבוע), וברור כי מלאכת בניית היסטוגרמה כזו לא תהיה נעימה. כמו כן ברור כי לבlokים גדולים יותר נקבל היסטוגרמות בנפחים בלתי-אפשריים. מעבר לכך, אפילו אם נתעלם מהנפחים הנוצרים, הרי שבתמונה זו יש² 512 פיקסלים, לעומת, למשל, במאוץ - רק פי ארבע יותר צמדים מכמות התאים לעובדה על צמדים. לכן, אין מספיק מידע בתמונה לשם בניית פונקציית הסתברות דו-מימדית.

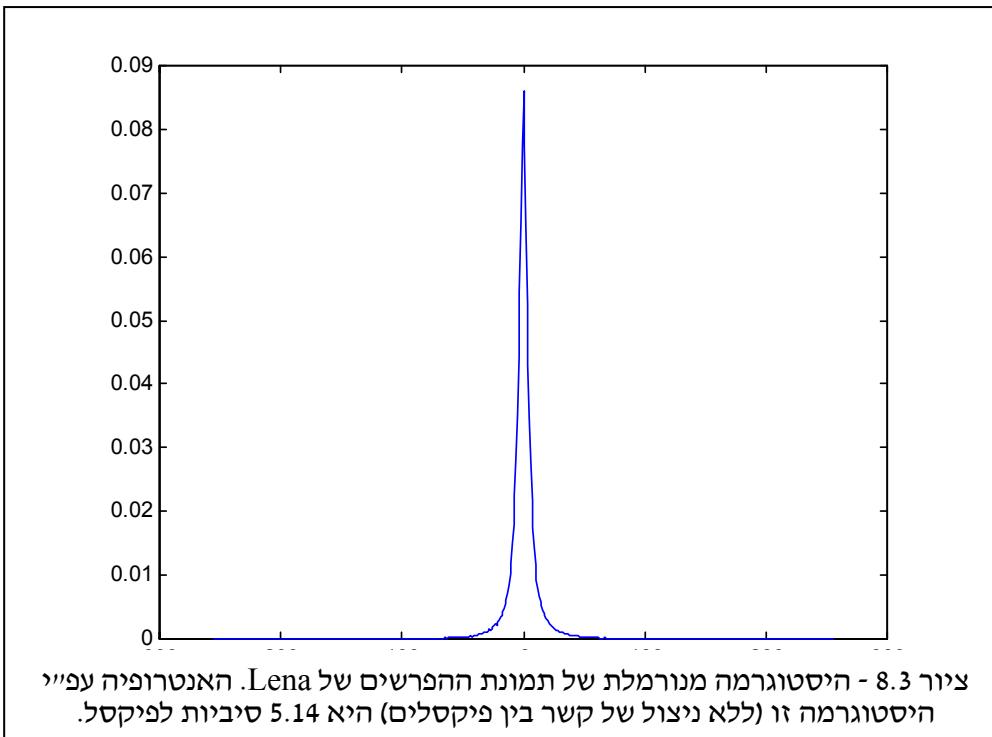
כתחליף לעובדה בבלוקים, ראיינו כי ניתן להציג עבודה עם חייזי. נמחיש רעיון זה על דחיסת התמונה הנתונה. נסדר את כל הפיקסלים בתמונה בעמודה בסידור לקסיקוגרפי, ונחליט כי במקום לשדר את הפיקסל $[k]X$, נשדר את ההפרש שלו משכנו הקודם - $[k-1]X$. במקלט כבר נקלט מידע על-פיו יודעים את $[k]X$, וכך בהינתן ההפרש - $E[k]=X[k]-X[k-1]$, ניתן יהיה לקבוע את $[k]X$ עצמו. נתבונן אם כך בהיסטוגרמה של הוקטור $E[k]$ - היסטוגרמה זו נתונה בציור 8.3. ברור כי ערכי ההפרש נעים בתחום 255 ועד +255, כך שהתחום הדינامي גדול. לעומת זאת, עיון בהיסטוגרמה זו מגלת כי רוב השגיאות מרוכזות קרוב לראשית (לאפס), והפלוג נראה פילוג לפלט בעל הצורה :

$$P(e) = \frac{\lambda}{2} \cdot \exp\left\{-\frac{|e|}{\lambda}\right\}$$

מסתבר כי שגיאות חייזי בתמונות בדרך כלל בעלות חוק פילוג כזה.

האנטロפיה במקרה זה היא 5.14. סיביות לפיקסל. ברור כי החיזוי המוצלח ביותר הוא לעשות שימוש בפונקציית ההסתברות המותנית, אך זו קשה לקבלה (היא גם משתנה במקומות) וקשה לשימוש. כפי שאנו רואים, ע"י חיזוי פשוט למדי (אנו נראה בהמשך שניתן להציג טוביים ממנו שייהיו מעשיים) קיבלנו אנטロפיה נמוכה בכ- 3 סיביות מהמשמעות המקורית.

כאשר נדון בדחיסת סדרות של תמונות אנו ננסה לנצל גם את הקשר הדוק בין פיקסלים בין תמונות. מסתבר קשר זה אף חזק יותר בהשוואה לחבר בין פיקסלים שכנים מאותו תמונה, כיוון שתמונות עוקבות בוידאו מאוד דומות. זהו בעצם המקור ליכולת הדחיסה החזקה כל-כך בוידאו.



3. **יתירות פסикו-ויזואלית** - שתי היתירות הניל מתייחסות למושג האנטרופיה. בעת לנצל את שלמדנו על פונקציית קצב-העיוות. מערכת הראייה שלנו אינה מושלמת, וניתן לנצל עובדה זו לשם השגת דחיסה טובה יותר.

לכן, אם רמות אפור בתמונת הנפרשת (לאחר דחיסה אובדןית) שונות במקצת מרמות האפור בתמונה המקורי, במרבית המקרים אין לנו מסוגלים להזוהות שינוי זה. כיון שכן, במקרים של תורת האינפורמציה, אין מתירים מרחק D כלשהו בין תМОונת הכניסה ותМОונת היציאה של תהליך הדחיסה-פרישה. כבר我们知道 כי מרחק כזה מניב אנטרופיה קטנה יותר - קלומר - המידע מכיל אפקטיבית יותר יתרות.

לסיכום דיוון זה על היתירות, היתירות הסטברותית סקלרית מציעה חסכוון קטן ביצוג האות. אם מוסיפים לה התיחסות להסתברויות בין פיקסלים, אין בעצם עושים שימוש ביתירות מרחבית (או זמנית) בין פיקסלים, ואז ניתן לדחוס בערך פי שתים. כאשר מכניםים למשחק את היתירות הפסיקו-ויזואלי ומતירים שגיאות קלות, יחס הדחיסה יכול לעלות לפי 10 ועד מאות - תלוי באיכות הרצואה ובאופן התמונאות המטופל.

עד כה עסקנו באנטרופיה תיאורטית. נשאלת השאלה - כיצד בונים מוקוד מעשי המצלחה להגשים את האנטרופיה או לפחות להתקרב לכך. באשר לקידוד ללא שגיאות בו אנו חותרים אל האנטרופיה, הכליל אותו נציג מיד הוא קוד הופמן. באשר לקידוד עם שגיאות, נזכיר כי לרוב לא ניתן כלל לקבוע מהי האנטרופיה האפקטיבית. אין נכיר מספר כלי יסוד המשמשים באlgorigתמי דחיסה אובדןית.

8.3 קידוד אנטרופיה - אלגוריתם הופמן

נניח כי נתנו לנו מקור X המציג תווים מעל אלף-בית של N תווים אפשריים (למשל - קובייה המייצרת מספרים בתחום - 1 עד 6). עבור מקור זה ידוע לנו חוק ההסתברות $(X)_P$, ועל פי פונקציה זו אנו יכולים לחשב את האנטרופיה התיאורטית - $H(X)$.

רצוננו לבנות קוד אשר יגשים קצב שיידור שהייתה האנטרופיה או קרוב אליה. לשם דוגמה, בהינתן תמונה בת מיליון פיקסלים כשל פיקסל בעל רמת אפור בתחום $[0,255]$, ברור כי ע"י הקצתה 8 סיביות לפיקסל בנוינו קוד מעשי, אך אם האנטרופיה של רמות האפור (קסקלרים) - אנו עדין לא מדברים על בלוקים או על חיזוי) ידועה להיות 4.5 סיביות לפיקסל - כיצד נתכנן קוד שירד לרמה זו? לשם מתן תשובה, נציג דוגמה פשוטה וממנה נכיר שני עקרונות חדשים - קוד בעל אורך מילה משתנה, וקוד Prefix.

דוגמה 8.9: נתנו מקור המציג 6 תווים אפשריים - $\{A,B,C,D,E,F\}$ עם ההסתברויות הבאות :

$$P\{A\} = 1/2, P\{B\} = 1/8, P\{C\} = 1/8, P\{D\} = 1/8, P\{E\} = 1/16, P\{F\} = 1/16$$

רואים כי סכום הסתברויות אלו הוא 1, כנדרש. בקידוד טריוויאלי נוכל להציג הקצתה 3 סיביות לכל תו ויהיה בכך בזבוז כיון שייתרתו שני צירופים ללא שימוש. תוצאה זו צפואה כיון שתיאורטי, עבור 6 תווים די ב- $2 \log_2 6 = 2.585$ סיביות לטו. במקרה שלנו האנטרופיה היא 2.125 סיביות לטו.

אם הגישה הטריוויאלית מקצת זהה של מילוט קוד לכל تو, גישה ייעלה יותר תהיה להקצות מילים באורך שונה לתווים שונים - تو בעל הסתברות גדולה קיבל מילה קצרה, ותו בעל הסתברות נמוכה קיבל מילה ארוכה. בדוגמה שלנו - נשתמש בקוד הבא :

טו	הסתברות	קוד	אורך
A	1/2	0	1
B	1/8	100	3
C	1/8	101	3
D	1/8	110	3
E	1/16	1110	4
F	1/16	1111	4

הארכאים אותם בחרנו הם בדיק OG לפי בסיס 2 של היפוך ההסתברות של כל تو. לכן, אורךה של המילה המומוצעת בקוד יהיה :

$$E[L] = \sum_{k=1}^N P\{X_k\}L_k = 1 \cdot \frac{1}{2} + 3 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} + 3 \cdot \frac{1}{8} + 4 \cdot \frac{1}{16} + 4 \cdot \frac{1}{16} = 2.125 \text{ bits/pixel}$$

וזהו בדיקת האנתרופיה, כך שיש לנו קוד מוצלח למדוי. השאלה הנותרת היא האם הוא ניתן לפענו? בקוד בו יש מספר קבוע של סיביות לכל TWO, כל של המקלט לעשות הוא לפרק את סידרת הסיביות המגיעה לקבוצות שוות סיביות וכל קבוצה תיצג TWO. עתה, כשמגיעה סידרת סיביות נדרש לפרק אותה לאורכים שונים, והשאלה אן זה אפשרי. נשים לב לתכונה הבאה המתקיימת בקוד שהוצע - אף מילת קוד קצרה אינה נהווה את הפתיחה למילוט ארכות יותר. תכונה זו נקראת **Prefix Code**, והיא זו המבטיחה פענו יחיד ונכון.

לדוגמה - נניח כי רצוננו לשדר את הסדרה - AABEACAADAFDBACA - נבנה עבור צירוף תווים אלו את סידרת הסיביות שתשדר - 0,0,100,1110,0,101,0,0,110,0,1111,110,100,0,101,0. סימני הפיסוק הם סימנים אותם המקלט לא יכול לקבל - עליו לפרק את הסדרה לבדו, ובזאת להיות הקוד בעלת תכונת ה- **Prefix** הוא יכול. התו הראשון הוא 0, ולכן המילה היחידה האפשרית היא A. וכך גם לסיבית השנייה. בסיבית השלישית יש 1 ולאחר זה יכול להיות B עד F. 1 כשלאחריו 0 יכולים להיות רק B ו- C, ו- 1 כשלאחריו שני אפסים זה רק B. כך התהליך נמשך כשהשנו רק פיסק אפשרי אחד - וזהו הפיסוק הנכון.

לxicoms דוגמה זו - ראיינו כי קיים קוד המגשים את האנתרופיה בדיקת זזאת ע"י קוד בעל אורך מילה משתנה - מיללים קצרות לתווים הסבירים יותר ומיללים ארוכים לתווים נוכחי החסתברות. כמו כן ראיינו כי היהת הקוד Prefix מאפשרת פענו יחיד ונכון של הסדרה. סוגיה שלא זכתה לمعנה עד כה - מאייפוא נולד הקוד שהוצע, וכייז נוכל תכנן קוד למקרים אחרים?

קוד הופמן דו במקורה הכללי ומנסה להכפיל את התוצאות הניל. אנו דנים במקור X המציין תווים מעל אלף-בית של N תווים אפשריים, ויש בידנו את חוק החסתברות (X)P, והאנתרופיה (X)H.

אלגוריתם הופמן מציע תהליך פשוט לפיו יבנה קוד Prefix מוצלח. התהליך הוא תהליך של בניית עצ' :

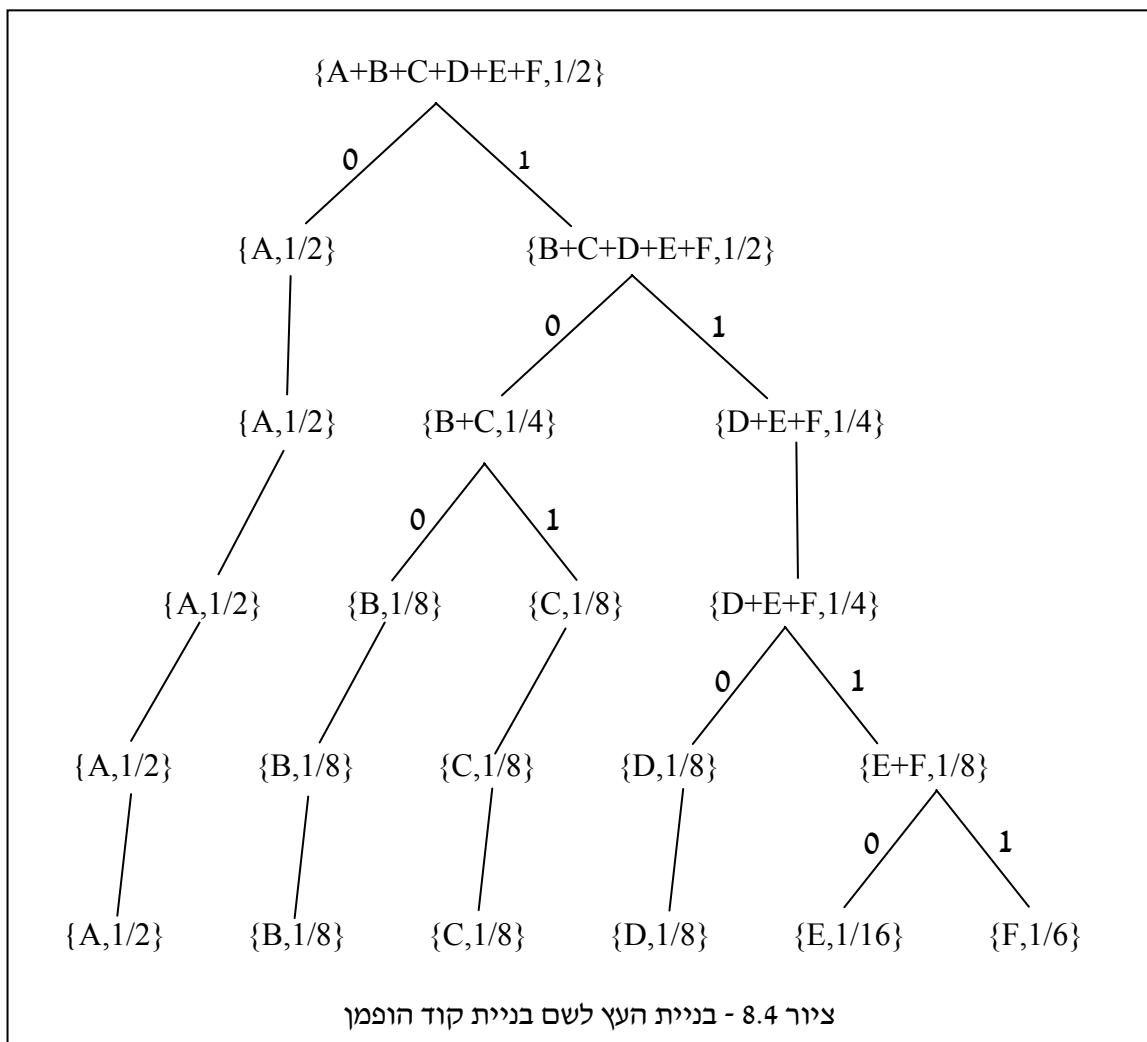
1. כתוב את סט N החסתברויות בשורה. קבוע $N=k$.
2. מתוך k התווים הקיימים, אחד את שתי התווים בעלי החסתברויות הנמוכות ביותר. האיחוד ייעשה ע"י שני ענפי-עץ המתאחדים לצומת אחת. צומת זאת מגדרה TWO חדש שהוא צירוף שני התווים שאוחדו, והסתברותו היא סכום החסתברותם. בסיום שלב זה מתקבל סט של $k-1$ תווים עם החסתברות לכל אחד, כסכום החסתברויות הוא 1.
3. אם $k > 1$ חזר לסעיף 2. ואחרת - סיום.

התהליך שתואר בונה עצ'. הקוד מתקיים ע"י הקצת 0 ו- 1 לכל לשני הענפים היוצאים מכל צומת. סיבירת החסתברות המגדירה מסלול משורש העץ ועד לעלה היא מילת הקוד של התו המתיחס לעלה זה. נמחיש על תהליך זה דרך דוגמה.

דוגמה 8.9: נחזר לדוגמה הקודמת בה נתון מקור המיצג 6 תווים אפשריים - $\{A,B,C,D,E,F\}$ עם $P\{A\} = 1/2$, $P\{B\} = 1/8$, $P\{C\} = 1/8$, $P\{D\} = 1/8$, $P\{E\} = 1/16$, $P\{F\} = 1/16$. ציור 8.4 מראה את העץ הנוצר.

בשורה התחתונה ישנה רשימה של כל התווים על הסטבריותיהם. בשלב ראשון מאוחדים E ו- F ביטם בעלי הסטברות הנמוכה ביותר. נוצרתו חדש בנקרא E+F והסתברותו שמיינית. בשלב שני יש ארבעה תווים בעלי הסטברות שמיינית - בחרנו לאחד את D על (E+F). כך התהליך ממשיך עד לקבלתתו אחד בעל הסטברות 1, שהוא למעשה מעשה כל התווים כולם.

בשלב שני אנו מקצים סיביות לכל שני ענפי צומת. כך למשל, על מנת לצאת משורש העץ (למעלה) ועד התו E עליינו ללקט במסלול שמותווה ע"י החלטות 1, 1, 0, ולבסוף 0 - זהה בדיקת המילה שהוצאה במסגרת דוגמה קודמת. למעשה - בדיקה קצרה מגלת כי עז זה מניב במדוק את הקוד שהוצע בדוגמה קודמת.

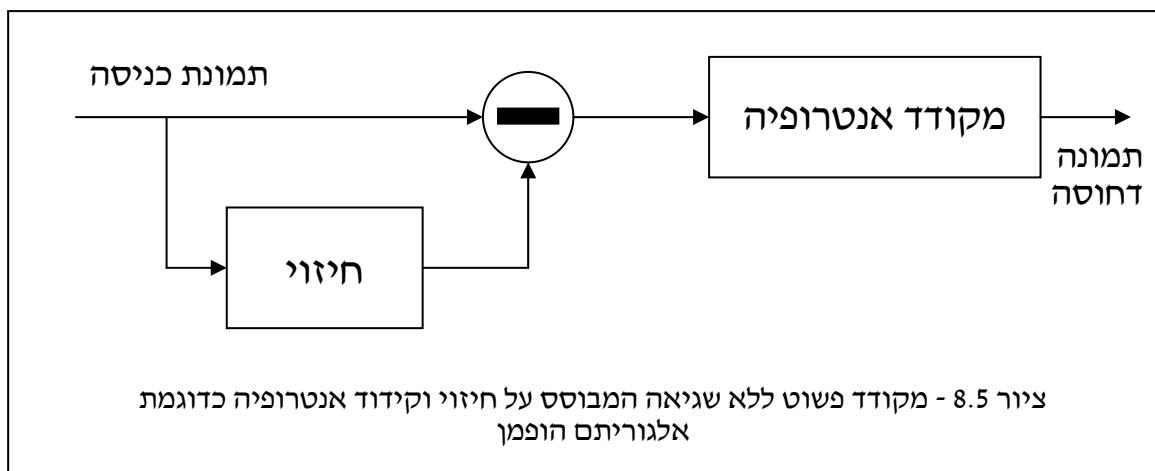


במקרה בו ההסתברויות מהוות חזקות שלמות שליליות של 2 (כמו הדוגמה הנ"ל) מתקבל כי קוד הופמן מගשים בצורה מדויקת את האנטרופיה, דהיינו, אורך המילה המומוצעת היא האנטרופיה. במקרה כלל יותר קוד הופמן ייתן אורך מילה שיהיה גבוהה מהאנטרופיה בעד סיבית אחת לטו, כלומר :

$$H(X) \leq L_{\text{Average}} < H(X) + 1$$

כאשר נתנו סט הסתברויות בו קוד הופמן סוטה בקרוב לסיבית אחת, ניתן להתגבר על כך ע"י עבודה בבלוקים. גם אם אין כל קשר סטטיסטי בין התווים, אם נתיחס לדוגמה לצמד תווים כ-"סופר-טו" חדש, הרי שנתקבל סטייה של עד סיבית אחד ל-"סופר-טו". לכן, לטו בודד תתקבל סטייה של עד חצי סיבית. במקרה הכללי, עבודה על בלוק של N תווים תיתן סטייה של עד $N/1$ סיביות. ברור כי תרגיל זה בא עם מחיר - עבודה על בלוקים קשה יותר. לדוגמה - עבור מקור בין 100 תווים, קוד הופמן המקורי דורש בניית של 100 מילוט קוד ושימוש בהן. עבור צמדים נידרש לתכnuן ואחסון של 10,000 מילוט קוד, ותהליך הדחיסה/פרישה מסתבך בהתאם.

בהתמך על כל האמור לעיל נוכל להציג אלגוריתם דחיסה פשוטי ללא שגיאות לתמונה נתונה. אלגוריתם זה יקודד את ההפרשין בין פיקסלים סמכים, וכבר ראיינו כי האנטרופיה הסקלרית של מידע זה היא 5.14 סיביות. אם נשתמש בקוד הופמן נוכל לקבוע כי במקרה החמור ביותר נצליח להגיע ל- 6.14 סיביות לפיקסל - דחיסה לא חזקה אך סבירה, ובמיוחד לנוכח פשtotות האלגוריתם. אם נשתמש בחיזוי מוצלח יותר, וואנו נקודד את שגיאות החיזוי בבלוקים נוכל לשוחט דחיסה טובה יותר. ציור 8.5 ממחיש בסכימת בלוקים את אלגוריתם הדחיסה המוצע.

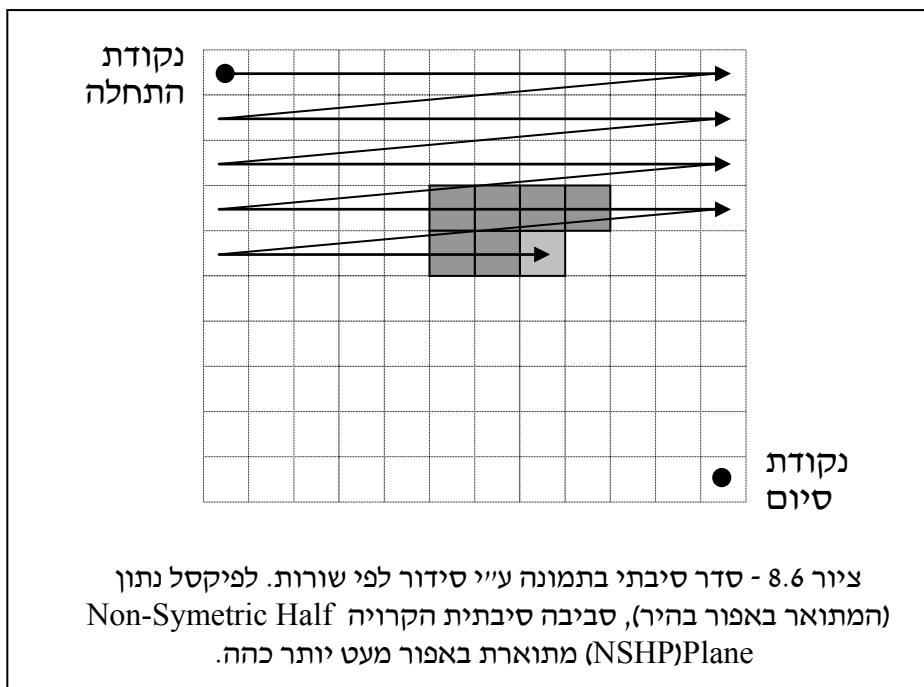


נקודה אחרת בדיעון זה על קודים בעלי אורך משתנה היא מידת רגשותם לשגיאות הארץ. אם מופעל קוד הופמן על סידרת תווים ומתקבלת סידרת סיביות, די בשינוי של סיבית אחת בסדרה כדי לשבש את הפענוח של כל מה שיבוא אחריה. ניתן להילחם בתופעה זה במגוון דרכים, אך אלו לא ניכנס לנושא זה.

8.4 דחיסה עם שגיאות - גישת החיזוי

באופן כללי ישן שתי גישות לביצוע דחיסה עם שגיאות - גישת החיזוי וגישת ההתרמה. אלו נציג את שתי גישות אלו במסגרת הדיוון הבא. בשתי גישות אלו, המרכיב האחראי על אובדן המידע הוא קוונטייזר, כפי שכבר הוצג בפרק 4. ההבדל המהותי הוא בשאלת מהי הפעולה המקדימה שיש לבצע על הפיקסלים בטרם יוכנסו לקוונטייזר - חיזוי (כפי שכבר עשינו בסעיפים קודמים), או התרמה. שתי אפשרויות אלו בעלות מכנה משותף חשוב - שתיהן ממירות את הפיקסלים עתירי התוירות במספרים בעלי אנטרופיה נמוכה יותר, וקשר סטטיסטי דל בין האלמנטים השונים. כתוצאה לכך, המידע לאחר פעולה החיזוי התרמה נוח יותר לדחיסה, ומאפשר עבודה בפיקסלים בודדים במקום בקטורים.

נתחל את הדיוון בסכימה המבוססת על חיזוי. עד ראשון לשם יישום של אלגוריתם דחיסה עם חיזוי הוא קביעת סדר סידור לפיקסלים. החיזוי הינו חיזוי סיבתי - ככלمر ערך החיזוי נבנה ע"י פיקסלים שכבר טופלו, וזאת על-מנת שגם המקלט יוכל לבנות את החיזוי ולצרכו לשגיאת החיזוי הנקלט. בעוד שבאות חד-מימדי סיבתיות היא עובדה טبيعית, בתמונה (אות דו-מימדי) אין זה כך. נניח לעת עתה כי התמונה מסודרת שורה לאחר שורה כמתואר בציור 8.6. בציור זה גם מתוארת סביבת פיקסלים סיבתיים עפ"י סדר זה, אשר יכולה לשמש לשם בניהת החיזוי. ישנו גישות שונות לסדר הפיקסלים אשר לא תידונה במסגרת זו.



בציור 8.5 הראיינו כיצד פועל אלגוריתם עקרוני של דחיסה מבוססת חיזוי ללא שגיאות. על מנת להכניס שגיאות (שבচוך, צפויות לשפר את עומק הדחיסה), علينا להכניס לפני קידוד האנטרופיה בлок חדש - קוונטייזר. הקוונטייזר יזון בשגיאת החיזוי אשר יכולה לקבל ערכים שליליים וחוביים. במקום לקודד את השגיאות במדוק, הקוונטייזר יבצע "העגלה" של שגיאות אלו, וכותזאה לכך תיווצר שגיאה בתהיליך הדחיסה. בזכות "העגלה" זו, מקודד האנטרופיה יזדקק להרבה פחות סיביות לשידור שגיאות אלו, וכך יושג הרוח בדחיסה.

דוגמה 8.10: נניח כי החיזוי שלנו לפיקסל $[j,k]$ הוא ממוצע של שכניו הסיבטיים לפי

$$\hat{X}[k,j] = \frac{1}{4} (X[k-1,j-1] + X[k-1,j] + X[k-1,j+1] + X[k,j-1])$$

בחיזוי זה עשינו שימוש בשלושת הפיקסלים המציגים מעליו, ובזה המציג באותה שורה לשמאלו. בהנחה שכל פיקסל מקבל רמת אפור המיצגת ע"י 8 סיביות, ערכי חיזוי זה ינעו בתחום - 255 עד + 255 ויקבלו כל ערך בתחום זה בקופיות של 0.25. לכן, יהיו 2041 ערכים אפשריים. דחיסה ללא שגיאה של אות שגיאה זה יחייב בניית קוד הופמן ל- 2041 ערכים אפשריים.

כתחליף, אם נפעיל קוונטיזר המעגל את השגיאה לערך השלם הזוגי, נקבל במצב קוונטיזר ערכי שגיאה חדשים באותו תחום אך עם 255 ערכים בלבד.

$$Q\{E\} = 2 \cdot \left\lfloor \frac{E+1}{2} \right\rfloor$$

קוונטיזר זה ימפה את הערכים $(-1,+1)$ לערך $(0,+1)$ ול- 2 וכן הלאה. נמחיש רעיון זה על הפיקסלים בתמונה Lena. נבנה לכל פיקסל את החיזוי כמתואר ונחשב את האנטרופיה ללא קוונטיזציה ועם קוונטיזציה כמתואר. לצורך חישוב האנטרופיה נזורך את העמודות והשורות הקיצונית, בהן החזאי נופל משטח התמונה. האנטרופיה ללא קוונטיזציה היא 4.86 סיביות לפיקסל, ואילו עם קוונטיזציה מתקבלת אנטרופיה של 2.78 סיביות לפיקסל.

במקודד ללא שגיאה, הפיקסלים הנקלטים הם הפיקסלים המקוריים (ללא שגיאה מעצם הגדרים ככלה). לכן, המקלט מצליח לבנות חיזוי זהה בערכו לחיזוי בו נעשה שימוש במשדר. בדחיסה עם שגיאה נוצרת בעיה - אם נבנה את החיזוי ע"י הפיקסלים הסיבטיים המקוריים במשדר, ובמקלט נשנה לחיקות תהליך זה ע"י בניית חיזוי על הערכים הנקלטים, נקבל ערכי חיזוי שונים. יתכן שהסת�性 תהיה קטנה, אך זו תשפיע על הערכים המפוענחים הבאים ותיזכר תופעה של שגיאה מצטברת. בעבר מספר עשרות פיקסלים המפענח יניב ערכים שווים לחלוtin ולא קשור לערכי המקור.

על מנת להתגבר על בעיה זו, על המשדר לבנות את החיזוי על פי ערכי רמות האפור כפי שיופיעו נח. במקלט. לכן, על המשדר להכיל בתוכו תת-מערכת שתהיה מקלט לכל דבר אשר מפענה את הערכים, ותת-מערכת זו היא אשר תניב את ערכי הפיקסלים על-פייהם ייבנה החיזוי. תהליך זה מתואר במלואו בצייר 8.7. צייר זה מראה סכימה עקרונית בה האות המקורי מזון לחזאי המנeba את הדגימה $[j,k]$ על סמך שכניה הסיבטיים. חיזוי זה מסומן ב- \tilde{X} . כאמור, על החיזוי במשדר להיות מבוסס על הערכים כפי שהוא ולא ע"י הערכים המקוריים. אנו נסמן את ערכי הדגימות הנקלטות במקלט ב- \hat{X} . החיזוי הוא למשל פונקציה מהצורה:

$$P\{X[k, j]\} = \tilde{X}[k, j] = f\{\hat{X}[k-1, j-1], \hat{X}[k-1, j], \hat{X}[k-1, j+1], \hat{X}[k, j-1]\}$$

בשלב השני מחוסר החיזוי מהערך האמתי ומתקבלת השגיאה

$$E[k, j] = X[k, j] - \tilde{X}[k, j]$$

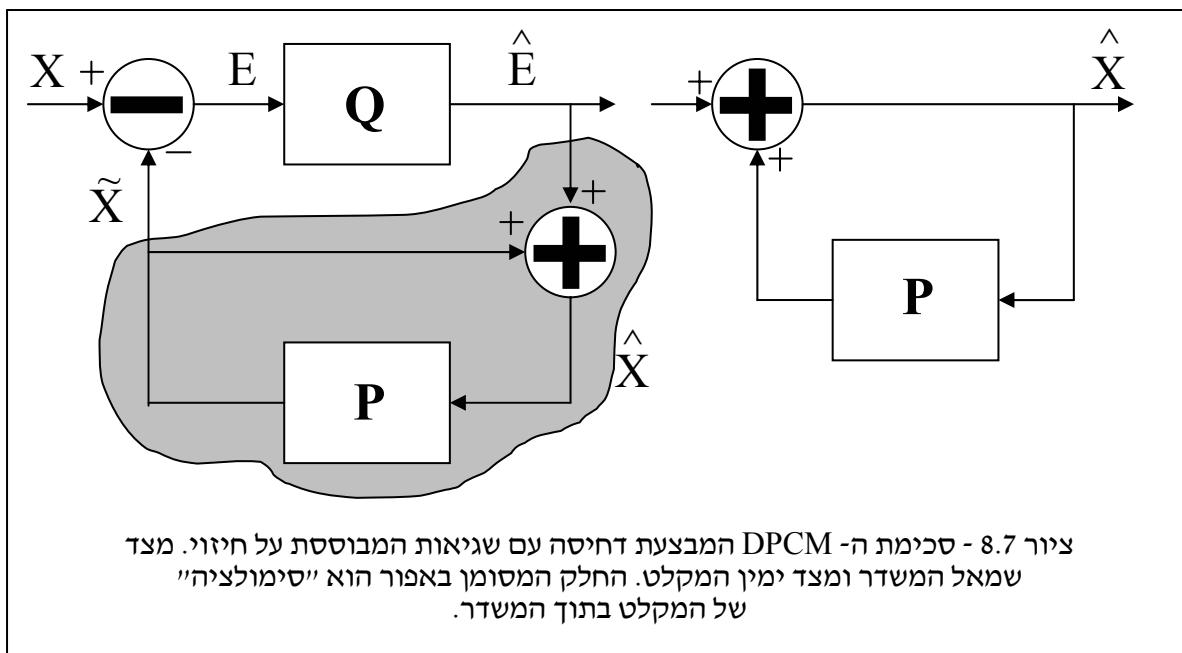
שגיאה זו מוגנת לקוונטייזר Q אשר ממיר את ערכי השגיאה בערכים מוקובים שלהם. זאת כיוון שאנו מוגנים על שידור השגיאות במדויק ורצוננו לשדר ערכים דומים להם. ברור כי הקוונטייזר היא פונקציה מדרגות, בה ככל שמדרגות גדולות יותר, שגיאת הקוונטייזציה גדלה יותר. בМОץ הקוונטייזר מתקבל

$$\hat{E}[k] = Q\{E[k]\}$$

גודל זה הוא המשודר החוצה, ולצורך כך נעשה שימוש במקודד אנטרופיה כדיימת קוד הופמן. אם הקוונטייזר אינו מאבד מידע (כלומר מהויה מערכת יחידה), אנו מקבלים מערכת דחיסה ללא שגיאה. המקלט מקבל את $\hat{E}[k]$ ומחבר אליו את החיזוי $\tilde{X}[k, j]$. לכן :

$$\begin{aligned}\hat{X}[k, j] &= \hat{E}[k, j] + \tilde{X}[k, j] = \\ &= (E[k, j] + \tilde{X}[k, j]) - (E[k, j] - \hat{E}[k, j]) = \\ &= X[k, j] - (E[k, j] - \hat{E}[k, j])\end{aligned}$$

כלומר, המוצא של המקלט הוא הערך האמתי של הפיקסל עם תוספת שגיאה שהיא בדוק שגיאת הקוונטייזציה. נשים לב כי המשדר מכיל בתוכו סימולציה של המקלט לשם בניית הערכים $\hat{X}[k, j]$.



דוגמה 8.11: אם נחזור לדוגמה 8.9, הרי שקיבלו ירידה בשתי סיביות בקירוב לפיקסל ע"י קוונטיזציה ששגיאתה היא לכל היותר רמת אפור אחת - שגיאה שאנו לא נצליח להבחין בה. בהמשך לדוגמה זו, עבור הקוונטיזור הבא :

$$Q\{E\} = 4 \left\lfloor \frac{E+2}{4} \right\rfloor$$

יתקבל 0 עבור כניסה בתחום [-2,2], 4 עבור כניסה בתחום (2,6], וכו'. במקרה זה האנתרופופיה המתقبلת היא 2.1 סיביות לפיקסל, והשגיאה המרבית הפעם היא 2 רמות אפור.

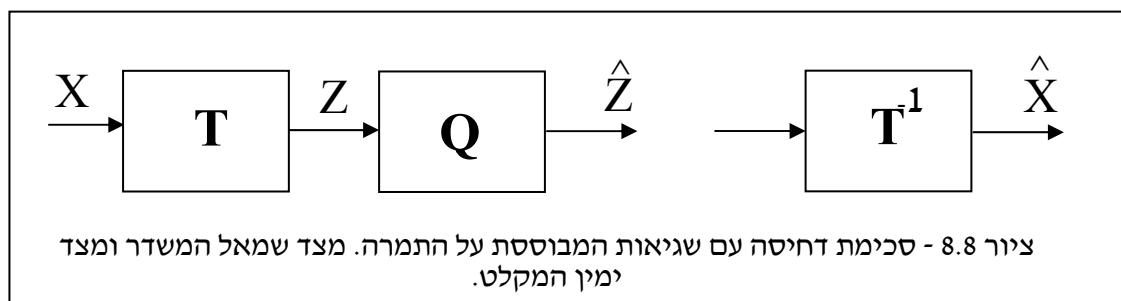
עד כה הוכחנו קוונטיזר פשוט וווניפורמי. למעשה, ניתן לישם קוונטיזר כללי יותר בעל מדרגות ברוחב משתנה, אשר יתוכנן ע"י אלגוריתם מקס-לOID. קוונטיזר כזה ינסה ליצור מדרגות צרות יותר סיביות הראשית כיון שלערכים אלו הסתברות גבוהה. לכן, השגיאה הכלולות בין אותן המשודר לאות הנקלט יהיה קטן יותר. ניתן גם להשתמש בשיקולים הנובעים מתוך מערכת הראיה בובאנו לתכנן את הקוונטיזור, אך אנו נעצור את הדין בנקודה זו.

אם נסכם, הרי שלמרכזו של האלגוריתם היה תהליך החיזוי והחזרתו אשר הניב תמונה חדשה בה הפיקסלים בעלי שתי תכונות חשובות :

1. אין קשר בין הפיקסלים - תכונה זו מאפשרת לנו לעבוד על כל פיקסל בנפרד מבלי לחושש שאנו מושתרים על יתרות מרחיבית.
2. ערכי הפיקסלים יוצרים פילוג עמוק אנטרופופיה (לפני קוונטיזור ובוודאי שגם אחריו) . לכן, אותן דחיס בצורה עילית.

8.5 דחיסה עם שגיאות - גישת התמרה

כתחליף לגישת החיזוי מוצעת גישה אחרת - גישת התמרה. גם גישה זו מנסה להשיג את שתי התכונות הנ"ל אך בדרך שונה. ניקח את התמונה הנדרשת - נניח כי היא בגודל M שורות על N עמודות - וביצע עליה התמרה. נניח כי אנו משתמשים בתמרה ריבועית יוניטרית ומתקבלת לפיכך תמונה התמרה בגודל זהה. אם ערכי התמרה יהיו חסרי קשר סטטיסטי ביניהם, ובנוסף בעלי אנטרופופיה עמוקה, יוכל להעביר כל אלמנט כסקלר לקידוד אנטרופופיה, ולהציג כתוצאה לכך דחיסה. על מנת להכניס שגיאות לתהליך ולשפר את יחס הדחיסה, יוכל לנקוט את ערכי התמרה ולהעבירם דרך קוונטיזור בטרם העברתם לקידוד אנטרופופיה. ציור 8.8 מראה סכימת בלוקים של מקודד ומפענה הבנויים על עקרון זה.



נותרה השאלה המרכזית - מיהי ההתרמה אשר תגשים יעד זה של ביטול הקשר בין הפיקסלים והקטנת האנתרופופיה שלהם. שאלת זו מ חוזירה אותנו לפרק 7 ולתוכנות שייצרו את התרמתה ה- KLT. ראיינו כי התרמתה ה- KLT היא ההתרמה היוניטרית הטובה ביותר במובן שאם נזוק חלק ממקדמי התרמתה, השגיאה בין תמונה המשוחזרת לאחר התרמתה הפוכה תהיה מינימלית. זריקת חלק מהאיברים כמו ביצוע של קוונטייזציה גסה מאוד עבור האיברים הנזוקים. יתרה מזו - הזכרנו כי ה- KLT מלכשנת את מטריצת האוטו-קורלציה של המידע בתמונה, ומסתבר (לא הוכח) כי/licson זה כמו ביצוע ביטול הקשר הסטטיסטי בין הפיקסלים השונים.

ברור כי אסור לשכך את החסרון המרכזי של התרמתה ה- KLT - היותה תלולה בתמונה בה אנו משתמשים. בשל תוכנה זו לא ניתן להשתמש ב- KLT לדחיסה. אבל ראיינו כי התרמתה ה- DCT מהוות קירוב טוב מאוד ל- KLT, והדבר נכון במיוחד לתמונות קטנות (16 על 16 ומטה פיקסלים). לכן, כתחליף ל- KLT ניתן לחת את התמונה, לחלק אותה לבlocks קטנים (גודל אופייני לשימוש הוא 8 על 8), לבצע על כל block כזו התרמת DCT דו-מימדית. התרמה זו יכולה להיעשות ע"י אלגוריתם מהיר המבוסס על FFT.

לאחר התרמתה, כל מקדם התרמתה מעבר דרך קוונטייזר. מתוך 64 הערכים המתקבלים (בשל התרמתה על 8 על 8), הקרובים לראשית הם החשובים בהיותם מייצגים מידע עמוק יותר. עבורם תיעשה קוונטייזציה עדינה (סקול לאמרה כי להם תינגן שגיאה קטנה יחסית). לערכים הרחוקים מהראשית תבוצע קוונטייזציה גסה יותר, ויתכן שאף ערכים אלו יזרקו. גישה מקובלת היא לבצע קוונטייזציה יוניפורמית לערכי התרמתה עם גודל מדרגה התלויה במיקום המקדם - קטן עבור מקדם הקרוב לראשית, וגדל עבור מקדם הרחוק מהראשית. גם כאן ניתן להשתמש בקוונטייזר מקס-לoid או קוונטייזציה הבנית מתוך שיקולי מערכת הרਆה.

האלגוריתם שתואר הוא אלגוריתם ה- JPEG הידוע בינוו הסטנדרט לדחיסת תמונות. אלגוריתם זה מבצע דחיסה ע"י התרמת DCT דו-מימדית וקוונטייזציה איחודית תלויות מקדם.

דוגמה 8.12: נדחוס את התמונה Lena ע"י אלגוריתם בסיסי מבוסט התרמתה. בשלב ראשון כל block של 8 על 8 מותמר ע"י DCT דו-מימדי. כל block של 8 על 8 עבר קוונטייזה ל- 8 סיביות באיבר ה- (0,0), ל- 3 סיביות לאיברים (*,0,0) ו- (0,*,0), (סה"כ 14 איברים ככל המיצגים את השורה הראשונה והעמודה הראשונה), ול- 2 סיביות לאיברים (1,1), (1,2), (2,1) ו- (2,2). לכן, לכל block שבמקור יוצג ע"י $2^{512} = 64^2$ סיביות אנו משתמשים ב- 64 סיביות, לעומת יחס דחיסה של 1 ל- 8 בקירוב. גודל המדרגה קבוע ע"י חישוב השונות של כל מקדם התרמתה ע"י מיצוע הבלוקים בכל התמונה. התקבלה התמונה הבהא, המופיע בציור 8.9. השגיאה בין שתי התמונות (מחושבת ע"י סכום ריבועי ההפרשיות מחולק במספר הפיקסלים ולאחר מכן שורש) היא 13.14 (ביחידות של רמות אפור).

מעיוון בתמונה המתקבלת ניכרת תופעה האופיינית לאלגוריתמי דחיסה הפעלים בבלוקים - אפקט הבלוקיות. גבולות הבלוקים ניכרים וצורמים - תופעה שנובעת מכך שהפעולות שבוצעו זרקו מידע

גודל מהבלוק, ולא נעשה כל ניסיון לשמר על רצף בין הבלוקים. קיימים אלגוריתמי דחיסה מוצלחים יותר אשר מטפלים יפה בבעיה ומציליםים להציג יחס דחיסה של 1 ל- 10 בלבד אנו שיען אනושית תוכל להבחן בהבדל. קיימים גם אלגוריתמים הנובעים מتوزע תורה השחוור שנדונה בפרק 6 אשר יודעים לקחת תמונה עם אפקט קלקל זה ולתקנה תזוז כפיה חלקות בין הבלוקים ושימוש שפות באופן כללי.



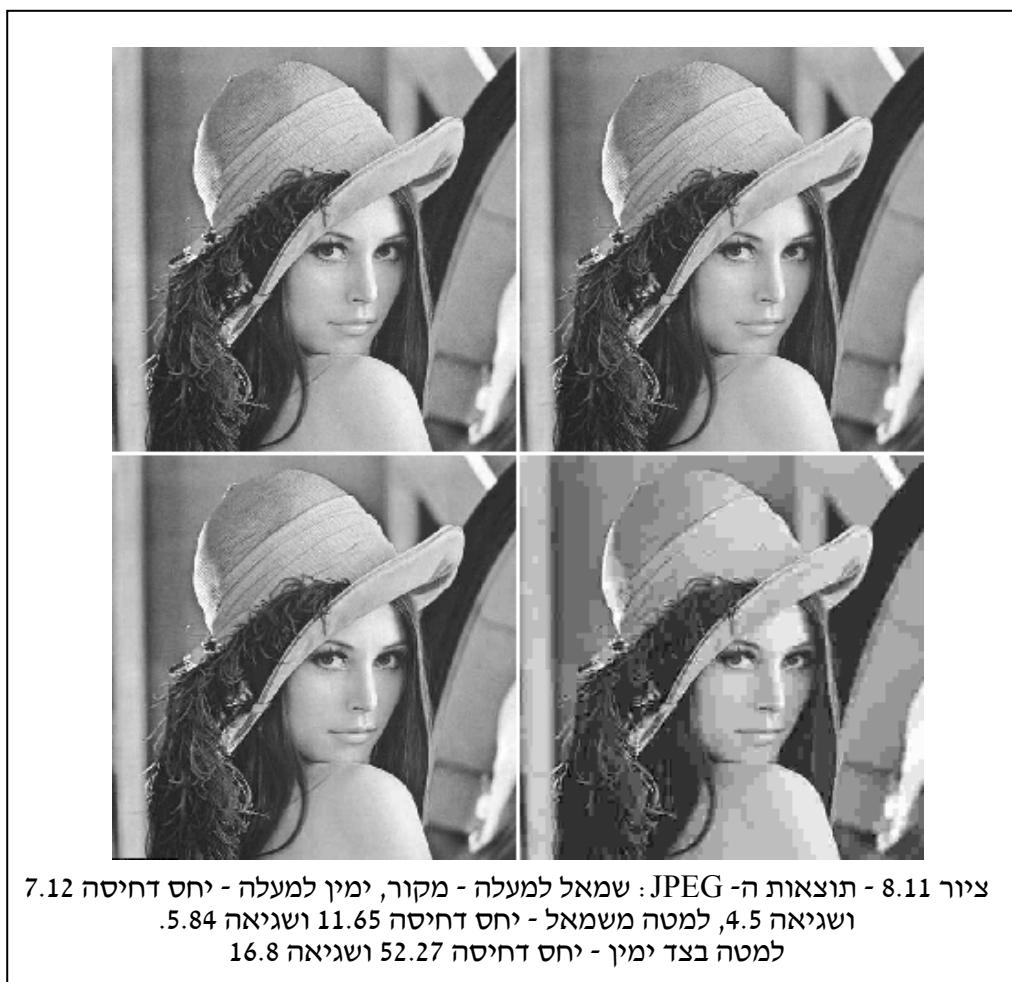
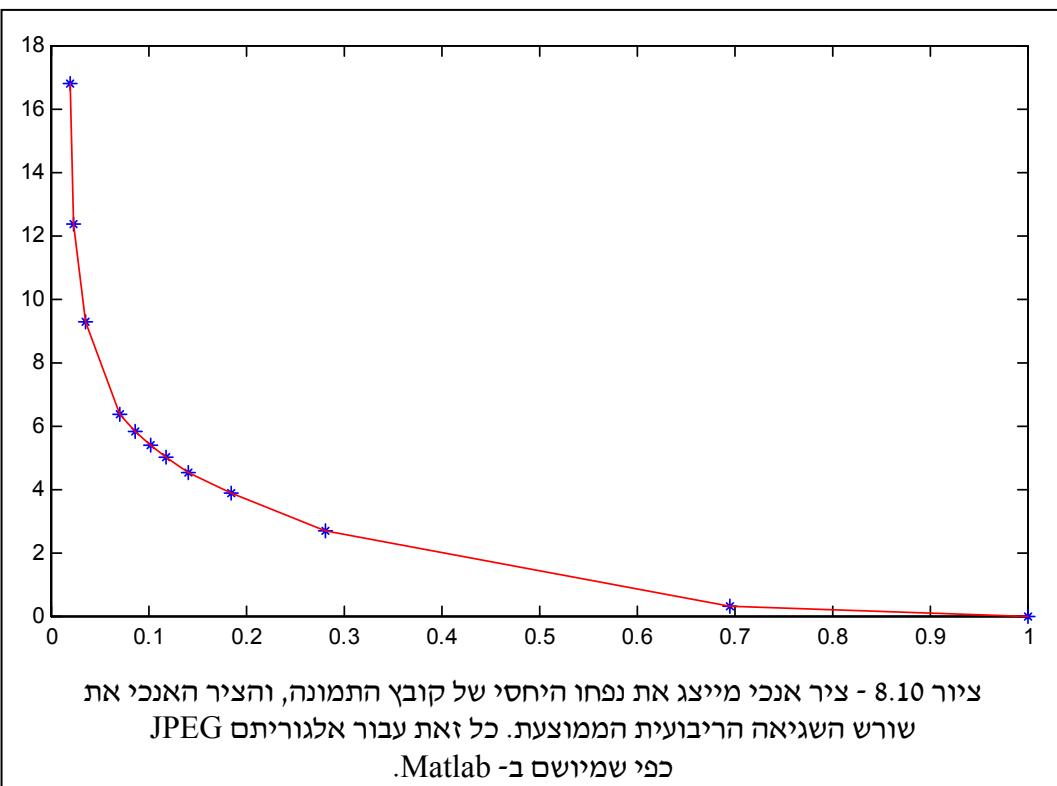
ציור 8.9 - תמונה מקור (משמאל) וגרסה דחוסה ופרוסת שלה ביחס 1 ל- 8 עיי
אלגוריתם דחיסה פשוטי המבוסס על התמרה, ולא קידוד אנטרופיה במווצאו

לשם השוואה, נפעיל את אלגוריתם ה- JPEG כפי שמיושם ב- Matlab. הפעלת אלגוריתם זה נעשית עיי פקודה שמירת התמונה (imwrite) באופן הבא - בהינתן תמונה כמערך ב- Matlab הקרויה (ונתון כ- uint8), הפקודה

```
imwrite(A,'temp.jpg','quality',70)
```

שומרת תמונה זו לקובץ בשם temp.jpg (השם מסגיר איזה פורמט התמונה מקבלת - במקרה זה דחיסת JPEG). באלגוריתם ה- JPEG, לא קובעים מראש את יחס הדחיסה אלא את מידת אובדן המידע. עבור איות של 100% מקבלים דחיסה ללא אובדן, ועבור ערך נמוך יותר מתירים אובדן. כאשר ערך זה מאוד נמוך יתקבל יחס דחיסה מצוין על חשבון שגיאה גדולה בין התמונה המקורית לתמונה לאחר פרישה.

ציורים 8.10 ו- 8.11 מראים תוצאות של דחיסת ה- JPEG. ציור 8.10 מראה גרפ' שהתקבל עבור מגוון ערכי האיכות (quality), וכגンド כל ערך כזו מוצג שורש השגיאה הריבועית המומוצעת כפונקציה של נפח היחסי של הקובץ ביחס למקור. רואים כי עבור התמונה המקורית בנפח יחידה מקבלת שגיאה 0, ירידת לנפח של 0.7 מהמקור נותנת שגיאה זעומה של כ- 0.3 ורמת אפור. תוצאה זו מעניינת כיון שהיא חושפת תכונה שלילית של קידוד מבוסס התמרה - יכולנו לקבל נפח של כ- 0.5 ללא שגיאה כלל, لوעשינו שימוש בדחיסת אנטרופיה. לעומת זאת, בהתרמה, מתקבלת שגיאה גם לנפחים גדולים מזה. עבור נפח של כ- 0.1 מתקבלת שגיאה של 5.4 - הרבה יותר טוב מהשגיאה אותה קיבלנו בדוגמה לעיל.



פרק 9 – עיבוד סדרות של תמונות

195	פרק 9 – עיבוד סדרות של תמונות
196 9.1 תנוצה בין תמונות
201 9.2 שערוך תנוצה בין תמונות - גישת בלוקים
205 9.3 שערוך תנוצה בין תמונות - גישה מלאה
210 9.4 ניקוי רעש בסדרות של תמונות
215 9.5 דחיסת סדרות של תמונות ומבוא ל- MPEG

בפרק זה אנו ננסים לתהום חדש של טיפול בסדרות של תמונות, למשל, סרטוי וידאו. כבר אמרנו בפרק ראשון כי סרט אינו אלא סידרה ארוכה של תמונות בודדות. בדרך כלל מדובר בכ- 30 תמונות בשניה לשם קבלת אשליה של רצף שימושיים על הסדרה בקצב זה. שימושים על צמד תמונות עוקבות אופייניים בסרט כלשהו, ההתרשות הראשונה היא שמדובר בתמונות זהות. כאשר מידת הפעילות (כמויות התנועה) בסרט מועטה, התמונות תהיו באמת כמעט זהות. השוני בין התמונות בד"כ מייחס לתנועה ביניהן. תכוונה זו מהווה את המוטיב העיקרי האלגוריתמים הדנים בסדרות של תמונות.

מהם הטיפולים הדוריים לסדרות של תמונות? ובכן, כל שאמרנו עד כה לתמונות בודדות רלוונטי גם לסדרות – נרצה לדגום ולבצע קוונטייזציה, נרצה לשפר במגוון אופנים, נרצה לשזר מקלוקלים ידועים, ונרצה לדחוס. במרבית המקרים ניתן לקחת אלגוריתמים שנעדו לתמונה בודדה ולאמץ לסדרה של תמונות עיי הפעלתם על כל התמונה בנפרד. בד"כ גישה זו לא מוצחת בשל שתי סיבות עיקריות:

1. משימה המבוצעת על סידרה של תמונות צריכה בד"כ לתת תוצאה טובה יותר בהשוואה לגישה בה כל תמונה מטופלת בנפרד. זאת כיון שבסדרה קיים יותר מידע עליון ניתן להתבסס. לכן למשל, ניתן לנוקות רעש בסדרה טוב יותר מאשר מהתמונה בודדה, וכן לדחוס סדרה עמוק יותר עבור אותה שגיאה מאותה סיבה בדיק.

2. משימה המבוצעת על סידרה צריכה בד"כ להיות קלה יותר בהשוואה למשימה המבוצעת על תמונה בודדה. זאת כיון שם כבר סיימנו את המשימה עבור התמונה הקודמת, הרי שהטיפול לתמונה הנוכחית כמעט עשוי עוד לפני התחילו בשל הדמיון הרב בין התמונות. ברור כי לכל אלגוריתם יש למצוא את המכינזם לניצול תכוונה זו.

במסגרת פרק זה אנו נתרכו בשלוש שימושות עיקריות - הראשונה מציאת וקטור תנוצה בין תמונות. משימה זו מהויה מרכיב מרכזי בובאנו לטפל בסדרות של תמונות, ולרוב אינה מהויה מטרה בפני עצמה. המשימה השנייה בהណון היא ניקוי רעש, והמשימה השלישייה היא דחיסת סרטים. אנו נראה כי שני היישומים האחרונים משתמשים בשערוך תנוצה (ב דרכים שונות) על מנת למצות את הפוטנציאל הטמון במידע הנתון.

1.9 תנוצה בין תמונות

נציג את הבעיה באופן הבא: נתונות שתי תמונות עוקבות בסרט וידאו - $X[k, j, t-1]$ ו- $X[k, j, t]$. השאלה הראותה ששאל היא - מהו הקשר בין תמונות אלה? נציג מספר צמדים מתוך סדרות של תמונות, ועל-פיהם ננסה להציג כלל כללי. ציור 1.9 מראה שלושה צמדים עוקבים של תמונות שלושה סרטים שונים. לכל צמד, מוצגת גם תמונת ההפרש בערך מוחלט, וכפי שניתן לראות, תמונה זו בעירה שחורה - דבר המעיד על הפרש נמוך. האזוריים הנעים הם החלקים עם השגיאה הגדולה. על בסיס זה מוצע המודל הבא לתאר את הקשר שבין שתי התמונות:

$$X[k, j, t] = X[k - dx(k, j, t), j - dy(k, j, t), t - 1]$$

כאשר $(dx(k, j, t), dy(k, j, t))$ הם מרכיבי התנועה בציר האופקי והאנכי של הפיקסל ה- (k, j) בזמן t . משווהה זו טוענת בפועל כי כל פיקסל בתמונה בזמן t מוקשו בפיקסל כלשהו מהתמונה הקודמת, כאשר יתכן שפיקסל זה גם עבר תמונה. לכן, רמות אפור נותרות קבועות בזמן, אך זוות ימנה שמאלה ומעלה-מטה. ציור 9.2 ממחיש רעיון זה לוקטוריים חד-ממדניים. הרעיון המרכזי הוא שאותות שונות קשורות זה לזה דרך עיוות של הקואורדינטות.

שאלה ראשונית בובאנו לטפל בסדרות של תמונות היא כיצד מוצאים את שדה התנועה dz/dt ? שדה תנוצה זה יכול לשמש אותנו לשם ייחוס נוכן של הפיקסלים בין התמונות ובשל כך לייעל לאין ערוך פעולות כמו שיפור או דחיסה. סוגיה זו קרויה בשם הכלול "שערוך תנוצה". כמות הנעלמים בבעיה כללית זו עצומה - פעמיים מגודל התמונה. במסגרת פרק זה נציג שתי אסכולות מרכזיות בנושא זה של שערוך תנוצה - שערוך התנועה לכל פיקסל בנפרד, ושערוך לבולוקים של פיקסלים. לשתי גישות אלו יש יישומים מגוונים ושוניים, והן משלימות זו את זו.

לעתים קרובות התנועה בין התמונות ניתנת לתיאור פרמטרי פשוט. נסקור בקצרה מספר מודלים כאלה, ונראה את תכונותיהם. אנו נניח כי הקואורדינטות נתונות במישור ברכף:

1. הזזה גלובלית - במקרה זה הפיקסל המצו依 במישור בנקודה (y, x) עובר למיקום חדש באופן הבא:



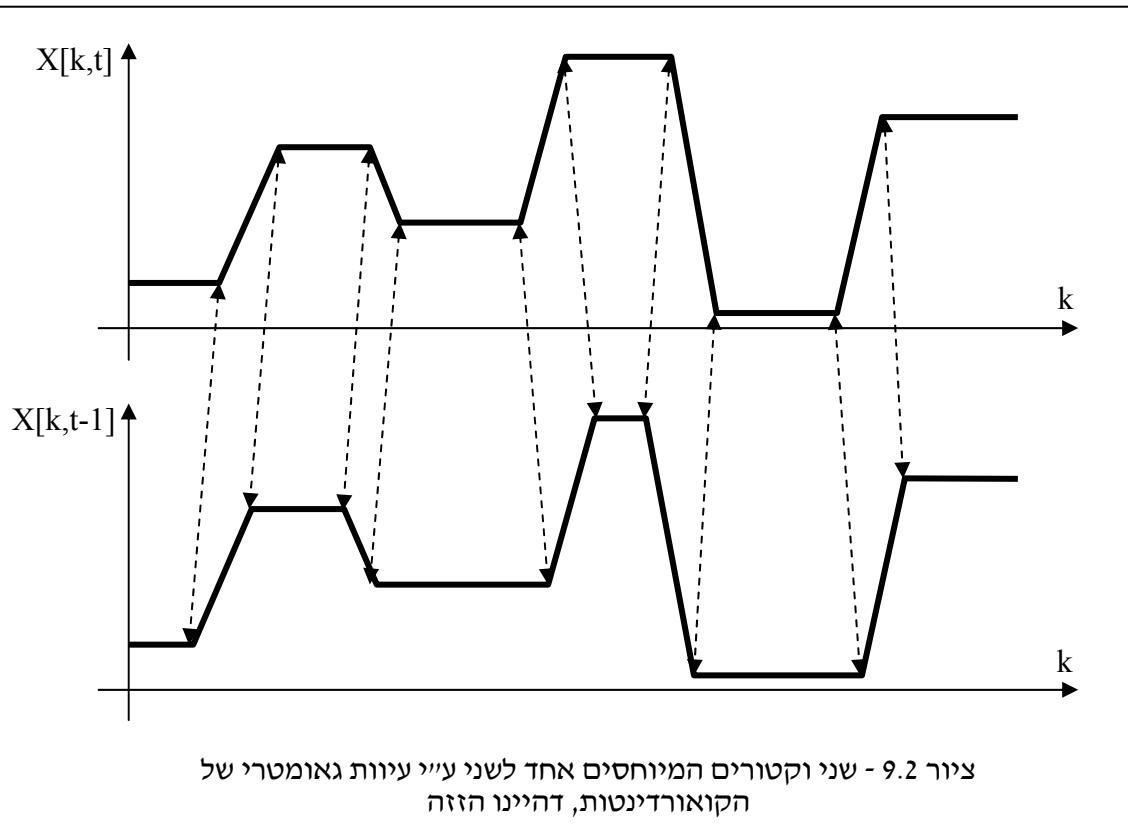
ציור 9.1 - צמדים תמונות עוקבים בשלושה סרטיים שונים, ותמונות ההפרש שלהם בערך מוחלט (לאחר הכפלת בערך קבוע לשם ניצול מלא התchrom הדינامي

$$\begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = \begin{bmatrix} x \\ y \end{bmatrix}_t + \begin{bmatrix} e \\ f \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \end{bmatrix}_t - \begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = -\begin{bmatrix} e \\ f \end{bmatrix}$$

במקרה זה, שדה התנועה בניו מחייבים שגודלם קבוע ושווה $\left[\begin{smallmatrix} -e \\ -f \end{smallmatrix} \right]$.

2. הזזה וסיבוב גלובליים (עיוות אוקלידי) - במקרה זה המיקום החדש נתון ע"י הקשר:

$$\begin{aligned} \begin{bmatrix} x \\ y \end{bmatrix}_{t-1} &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t + \begin{bmatrix} e \\ f \end{bmatrix} \Rightarrow \\ \begin{bmatrix} dx(x, y, t) \\ dy(x, y, t) \end{bmatrix} &= \begin{bmatrix} x \\ y \end{bmatrix}_t - \begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = \begin{bmatrix} 1-\cos(\theta) & -\sin(\theta) \\ \sin(\theta) & 1-\cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t - \begin{bmatrix} e \\ f \end{bmatrix} \end{aligned}$$



המטריצה מסובבת בזווית θ ואז מתבצעת פעולה הזזה כמפורט. הפעם שלושה פרמטרים מאפיינים את מלאה התנועה. שדה התנועה הפעם הם חיצים תלויים מקום, כפי שניכר מהמשוואות

3. עיוות דמיון - במקרה זה המיקום החדש נתון ע"י הקשר:

$$\begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t + \begin{bmatrix} e \\ f \end{bmatrix}$$

הסיבוב וההזזה מתבצעים כמפורט. בנוסף, יש הגדלה/הקטנה שונה בכל ציר, ובסה"כ 5 פרמטרים. ברור כי גם הפעם יתקבל שדה חיצים משתנה במקומות.

4. עיוות אפיני - הקשר הפעם נתון ע"י:

$$\begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t + \begin{bmatrix} e \\ f \end{bmatrix} \Rightarrow$$

$$\begin{bmatrix} dx(x, y, t) \\ dy(x, y, t) \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix}_t - \begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = \begin{bmatrix} 1-a & -b \\ -c & 1-d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t + \begin{bmatrix} e \\ f \end{bmatrix}$$

כך שישנו שישה פרמטרים המאפיינים את התנועה. תנועה אפינית מכילה בתוכה את עיוות הדמיון עם אפשרות של מרכיב הקורי shear - גזירה - המבצע מתיחה אלכסונית של הצירים.

5. **עיוות פרויקטיבי** - הקשר בעט מסובך יותר ונתנו ע"י :

$$\begin{bmatrix} x \\ y \end{bmatrix}_{t-1} = \frac{\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t + \begin{bmatrix} e \\ f \end{bmatrix}}{\begin{bmatrix} g & h \\ i & j \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}_t + 1}$$

ובעת שמונה פרמטרים נתונים את הקשר בין הקואורדינטות. בעיות זה ובכל קודמי, קווים ישרים מומרים לקווים ישרים.

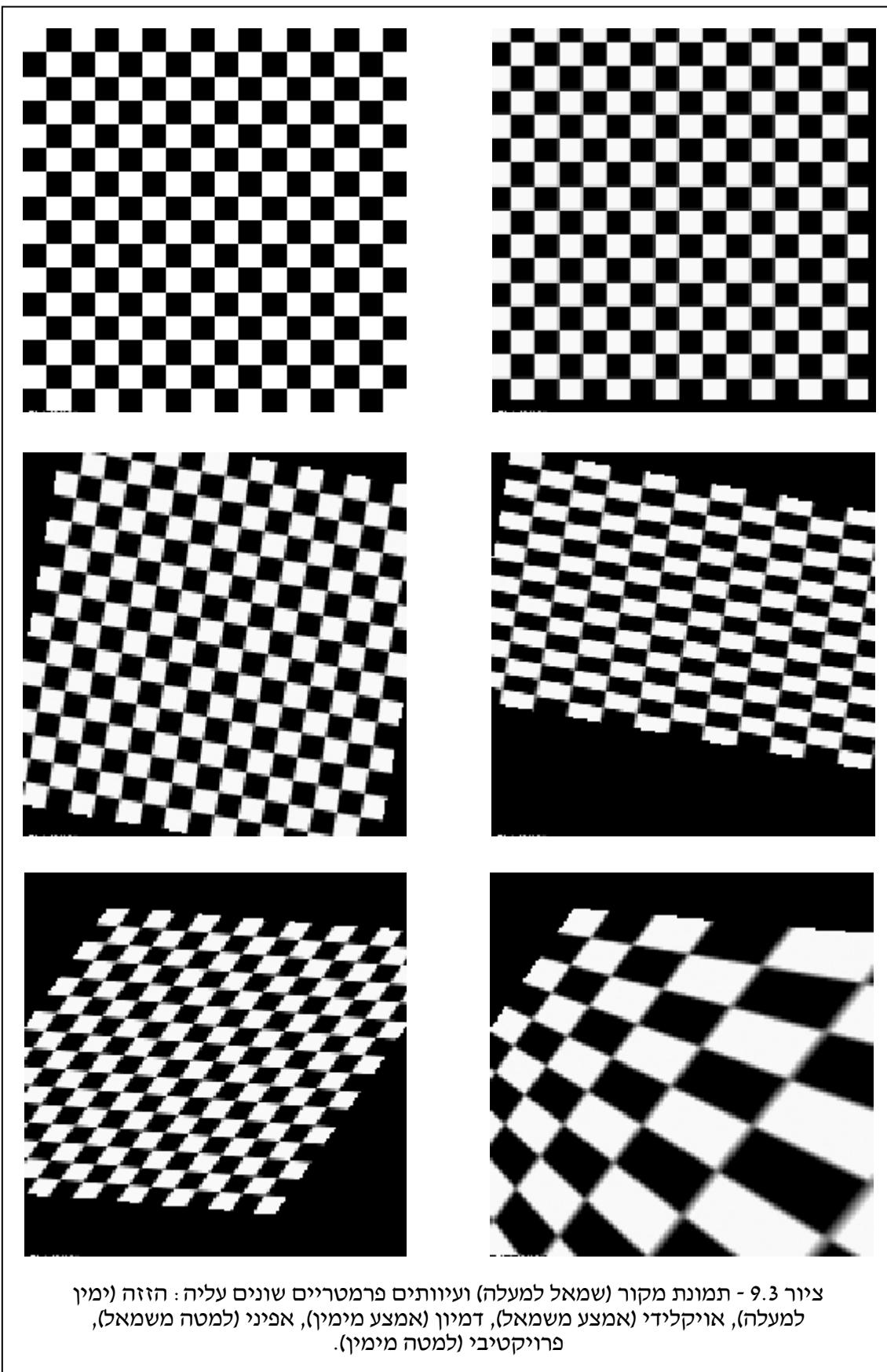
6. **עיוות פולינומיAli** - בניגוד לעיוותים הקודמים אשר להם קשר פיזיקלי ישיר בדרך היעוצרן של תמונות, בהיותם הטלות מגוונות מתלת מימד לדו-מימד, ניתן להציג קשרים שרירתיים תלויים בין הקואורדינטות. קשר אפשרי מסווג זה הוא הקשר הבא :

$$x(t-1) = a_1 + b_1 \cdot x(t) + c_1 \cdot y(t) + d_1 \cdot x(t)y(t) + e_1 \cdot x(t)^2 + \dots$$

$$y(t-1) = a_2 + b_2 \cdot x(t) + c_2 \cdot y(t) + d_2 \cdot x(t)y(t) + e_2 \cdot x(t)^2 + \dots$$

בציור 9.3 מובאת תמונה לוח שחמט ועיוותים עליה על מנת להמחיש עיוותים אלו. ההזזה (שורה עליונה מוצעה עם וקטור תנוצה של [2.1,4.2], העיוות האוקלידי (אמצע בצד שמאל) בוצע עם הזזה של [-15,15] וזווית של 10°. בעיות הדמיון נעשו שימוש באותו זווית, עם הזזה של [-5,5] ועם הגדלות של 0.7 ו- 1.5 בהתאם לשני היצירום. העיוות האפיני התקבל עם הזזה של [-45,-15] וארבעת איברי המטריצה הם [1, 0.7, -0.1, 1.3]. לקבלת העיוות הפרויקטיבי עשוינו שימוש באותו ערכי התמורה כבקרה האפיני עם שימוש בשני הפרמטרים הנוספים [0.01, 0.0062]. בדוגמאות אלו, כאשר מתקבלת נקודה חדשה במקום שאינו שלם, בחרנו לבצע אינטרפולציה ביליניארית ע"י שימוש באربعة השכנים של הנקודה. כמו כן, נקודות שפולשות אל מחוץ לשטח התמונה נחשבות כאפסים.

אם נחזיר לצמד תמונות עוקבות בסדרה אופיינית, תנועה גלובלית ניתנת לתיאור ע"י מספר פרמטרים מצומצם ניתנת ליחסו לתנועת מצלמה כאשר העצם המצלום נייח ורחוק יחסית מהמצלמה, או אם העצם מישורי. כאשר לעומת זאת מצלמת מושך מטרידית עם עומק משמעותי ביחס למטריקת התנועה מרכיבת יותר כיון שנקודות קרובות נעות מהר יותר מנקודות רחוקות. תנועה גלובלית ניתנת גם לקבלה כאשר המצלמה נייחת ואילו עצם קשיח הממלא את כל שטח התמונה נעה בתלת-מימד. בכל מקרה לאחר של עצמים תלת-מימדיים קרוביים ומצלמה נעה, או תנועה מקומית של עצמים (כגון מכונית בכביש), כשהתמונה נייחת והמכונית מוצאה חלק הראריה), מתקובל שדה תנועה מורכב יותר בו עיקרונו לכל פיקסל צפוי להיות חזק שונה המעיד על כיון התנועתו. במסגרת הסעיפים הבאים נדון בשאלת כיצד משערכים את שדה התנועה הזה. אנו נබיל את הדיוון לשערוך תנועה כללי מבלי להניח מודל פרמטרי - ישם אלגוריתמים המסוגלים לפעול הרבה יותר מדויק ומהר אם מניחים מראש מבנה מוגבל של תנועה.



נקודה אחרונה בה נדוע במסגרת סעיף זה היא ייצוג פועלות עיוות גיאומטרי כלשהו כמכפלה במטריצה. נסטכל על המשוואה

$$X[k, j, t-1] = X[k + dx(k, j, t), j + dy(k, j, t), t]$$

משווהה זו קובעת כי אופרטור העיוות הגיאומטרי המסויים על התמונה $[k, j, t-1] X$ מניב את התמונה $X[k, j, t-1]$. אופרטור זה ליניארי כיון שם נפעיל אותו על צירוף ליניארי של תמונות נקבל צירוף ליניארי של תוצאות העיוות. لكن ניתן ליצgo כמטריצה אשר מסומן $F[t, 1]$, כאשר סימון זה קובע כי אופרטור זה מעביר תמונה זמן t לנוקודת הזמן $t-1$. במקרה הכללי, $F[t, k]$ מעבירה את התמונה בזמן t בזמן k .
ברישום לקסיקוגרפי נקבל את הקשר הבא:

$$\underline{X}(t-1) = F(t, 1) \underline{X}(t)$$

כאשר נניח כי אורך הוקטורים המעורבים הוא N . אם נסטכל על שורה אחת במערכת משווהות זו נוכל להבין כיצד בנזיה המטריצה F . שורה מסויימת תיראה כך:

$$X_{n_0}(t-1) = X[k_0, j_0, t-1] = \sum_{n=1}^N F_{[n_0, n]}[t, 1] X_n(t)$$

משווהה זו אומرت כי הפיקסל ה- $t-1$ בתמונה הקודמת (אינדקס זה הוא האינדקס בסידור לקסיקוגרפי, והוא שקול למיקום ה- $[j_0, k_0]$ בדו-מימד) מתקיים כצירוף ליניארי של פיקסלים מהתמונה הנוכחית. מקדמי הצירוף הליניארי הזה יהיו יסתכמו ל-1. למשל, אם הפיקסל ה- $t+3$ מתקיים על ידי הזהה של הפיקסל ה- $t+3$, הרי שכל השורה תהיה אפסים ורק האיבר $F_{[n_0, t+3]}[t, 1]$ יהיה 1.

בעיקרו, כאשר אנו נתונים ביצוג של עיוות גיאומטרי כמטריצה, אנו מניחים ייצוג דיסקרטי של התמונה. לכן, בבנייה המטריצה ליישום התנועה הנדרשת מסתתרת גם תפיסת האינטראופולציה שתימושם. למשל, אם נרצה אינטראופולציה השכן הקרוב (Nearest Neighbor), כל שורה ב- F תכיל 1 במקום כלשהו ואפס בלבד. בשימוש באינטראופולציה בי-ליניארית נקבל כי כל שורה מכילה 4 איברים אי-שליליים סוכומים 1, ואלו מייצגים את המשקלות של רבייעיות הפיקסלים הסמוכים לנוקודה אליה עברנו. במקרה, המטריצה המתבקשת דיליה מאוד, דהיינו, מרבית איבריה הם אפסים, וכך קל לאחסנה ולתמן. איתה אם נכון.

9.2 שערוך תנועה בין תמונות - גישת בלוקים

נניח כי אנו עוסקים בתמונות בגודל של 512 על 512 פיקסלים. מציאת וקטור התנועה לכל פיקסל בנפרד מחייבת בפנינו מספר קשיים אובייקטיביים:

1. כמות הנעלמים גדולה ושווה לפעמיים מספר הפיקסלים. גם אם יוצע תהליך מהיר לחישוב הוקטור לכל פיקסל, ביצוע תהליך זה כל כך הרבה פעמים הופך לשימושה כבזה מבחינה חשובה.
2. גם אם איינו חשובים מכמות החישובים, כמות הזיכרון הנדרשת מהוות גורם מפריע. לא די בכך שעליינו לשמור את שתי התמונות הנדונות, גם علينا לשמור את התוצאה כשני מערכיים בגודל התמונה שבמקרה הכללי צפויים להחזיק ערכיהם ממשיים. במקביל, אם علينا לשדר לגורם כלשהו שדה וקטורים זה, הרי שכמויות המידע הנדרש לשידור עצמה - למעשה - יותר מהתמונה עצמה. לכן, באлогוריتمי דחיסה אין מקום לבניית שדה תנועה מלא המיחס וקטור לכל פיקסל.
3. אם איינו חשובים מחישובים או זיכרנו עדיין קיימת בעיה. אם נתבונן בפיקסל $[k,j,t]X$ ונצא לחפש את מקורו בתמונה הקודמת, קרוב לוודאי שיתקבלו מספר פתרונות אפשריים. כמו כן, רוש בתמונה ישפייע באופן קיצוני על איקות ההתאמנה. איך נברור מtones את הפתרון הנכון? בחירה טבעית היא להניח סוג של חלקות על שדה התנועה המתkeletal. אם לפיקסל כלשהו יש וקטור תנועה נתון, צפוי כי שכניו ינוו באופן דומה, וזאת מתוך הנחה שהם שייכים לאותו עצם בצלום.

בשל שלושת הסיבות, נוכל להציג את התפיסה הבאה - במקומות ליחס וקטור תנועה שונה לכל פיקסל בתמונה, נחלק את התמונה לבלוקים (לדוגמה בגודל של 8 על 8 פיקסלים), ולכל בלוק כזו ניחס וקטור תנועה אחד. באופן זה נצטרך לחשב ולאגורר הרבה פחות וקטורים (לדוגמה שלנו - פי 64 פחות). מעבר לכך, גם נפתר גם את הבעיה השלישית, כיון שאם נחשף את מיקומו של הבלוק כמקרה אחד בתמונה הקודמת, אנו צפויים לקבל התאמה חסינה יותר לרוש. החסרונו של תפיסה זו ברור - אם בבלוק נתון קיימות שתי אוכולוסיות של פיקסלים שנעו באופן שונה - אנו לא נדע מכך וקרוב לוודאי שניחס לכל הבלוק וקטור לא מוצלח.

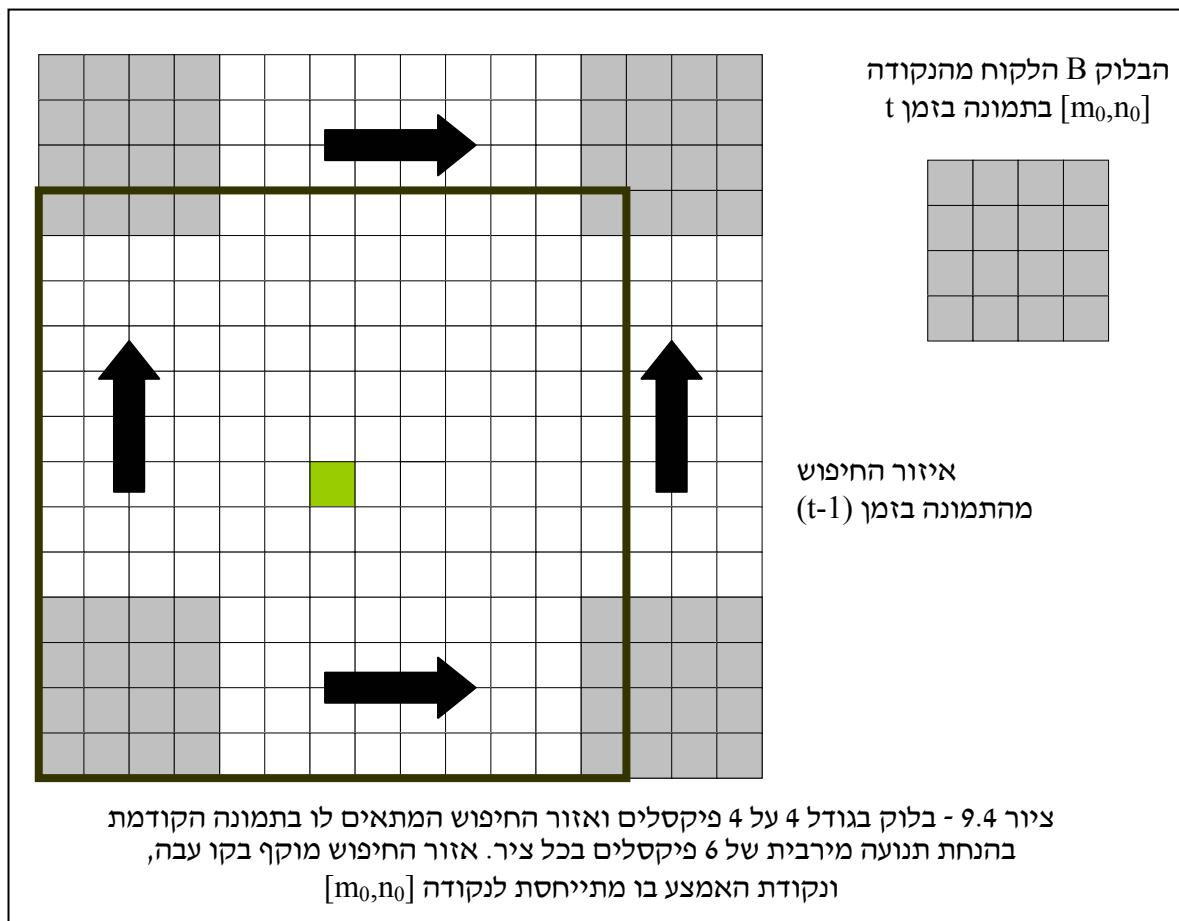
נניח כי בחרנו לעבוד עם בגודל של 512 על 512, וחילקנו אותה ל- 64 על 64 בלוקים סמוכים שככל אחד בגודל של 8 על 8 פיקסלים. לכל בלוק כזו علينا למצוא וקטור תנועה, והשאלה בה נטפל כעת היא כיצד נעשה זאת. השיטה בה ננקוט היא הבאה - לבלוק נתון מהתמונה $[k,j,t]X$ נבדוק עבור כל מיקום אפשרי בתמונה הקודמת $[j,k,t-1]X$ את מידת ההתאמתו. מידת ההתאמנה תיקבע ע"י החסרונו מהפיקסלים במיקום המוצע, העלאת ההפרשנים בריבוע וסכום. ערך נמוך יעיד על התאמה טובה וערך גבוה על התאמה גרועה. אם נסמן את הבלוק המטופל ב- B, ונניח כי בלוק זה נלקח ממהיקום $[m_0,n_0]$ בתמונה $[k,j,t]X$, כלומר:

$$B[k,j] = X[m_0 + k, n_0 + j, t] \quad 1 \leq k, j \leq 8$$

אנו ננסה להביא למינימום את הפונקציה הבאה:

$$E[m,n] = \sum_{k=1}^8 \sum_{j=1}^8 (B[k,j] - X[m+m_0+k, n+n_0+j, t-1])^2$$

הערכאים $[n_0, m+n_0]$ מטארים את המיקום ממנו עקנו את הבלוק המועמד מתוך התמונה הקודמת, ופונקציה זו מודדת בעצם את המרחק בין B לבלוק זה. בפועל, אין צורך לבדוק את שיעור ההתאמה לכל $[n, m]$ בשטח התמונה - יש להתמקד באזור חיפוש סביר, תוך התייחסות להסתה המקסימלית האפשרית. למשל, אם ידוע לנו כי בלוק לא ינוע יותר מ- 6 פיקסלים לכל כיוון, הדבר מגדיר אזור חיפוש סביר הנקודה $[n_0, m_0]$ שגודלו 13 על 13 פיקסלים. לכן, נתוניינו בערכי $E[m, n]$ עבור ערכי m ו- n בתחום -6 ועד +6. ציור 9.4 ממחיש רעיון זה עבור בלוקים בגודל 4 על 4.



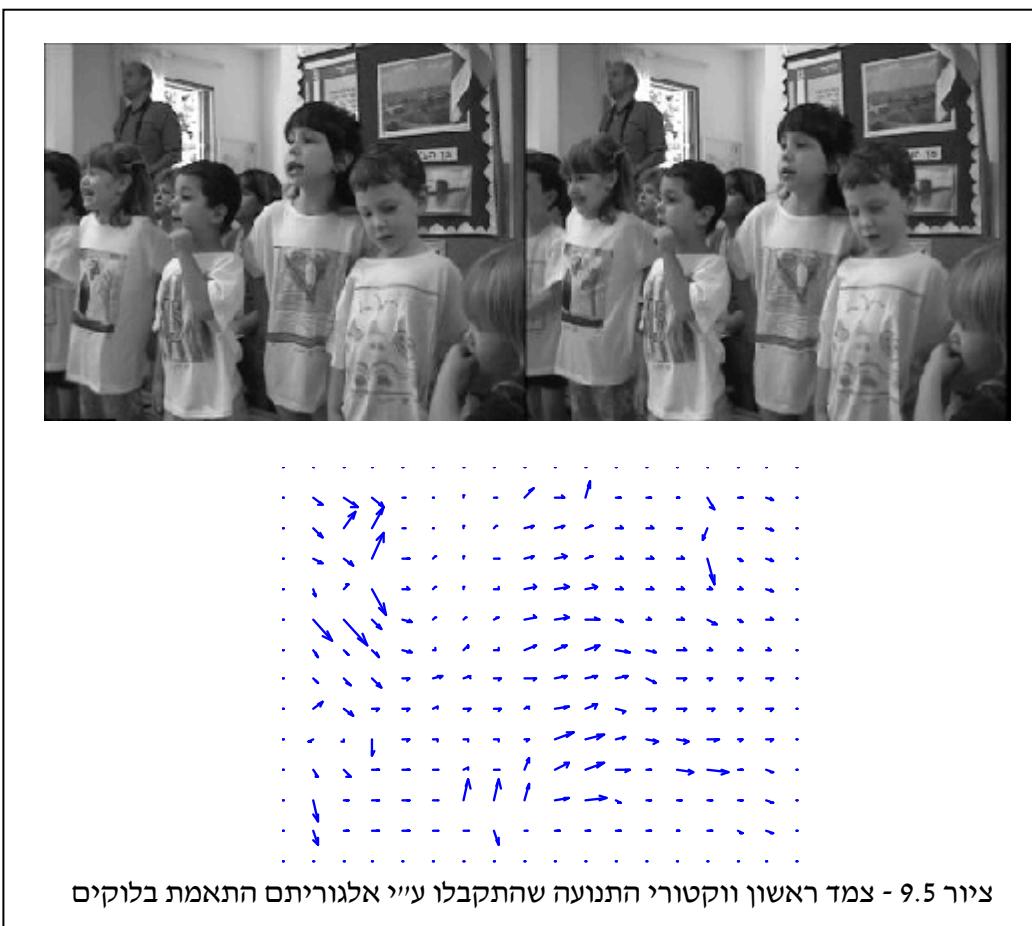
תהליך המינימיזציה דורש חיפוש מייגע - מעבר על כל המיקומים האפשריים (מעבר תנוצה מקסימלית k יש $(2p+1)^2$ מיקומים כאלה), חישוב הערך E , ומיציאת הנמק שבנייתם. מסתבר כי לישומים רבים גם תהליך זה אינו נחטיב כזול מבחינה חישובית, ונדרש אלגוריתם מהיר ממנו. דרך אחת להקל היא לשנות את E לפונקציה הבאה:

$$E[m,n] = \sum_{k=1}^8 \sum_{j=1}^8 |B[k,j] - X[m+m_0+k, n+n_0+j, t-1]|$$

באופן זה המרנו הعلاה בריבוע לערך מוחלט - פועלה קלה יותר.

בנוסף, בהנחה שהפונקציה $[n,m]E$ הינה פונקציה קמורה עם מינימה יחידה (כפי שסביר שבדרך כלל מתקיים), ניתן להציג מגוון אלגוריתמים מהירים למציאת נקודת המינימום. אנו לא נציג אלגוריתמים אלו במסגרת זו, אך נאמר כי בrama עקרונית הם מנגעים לחשב את $[n,m]E$ באופן סלקטיבי ולפי התוצאות הקודמות לקבוע את מיקומן של נקודות הבדיקה הבאות.

בצירורים 9.5 ו- 9.6 אנו מראים שוב את צמדי התמונות מצירור 9.1 והפעם עם תוצאה שעורך התנועה בגישה זו לבлокים של 16 על 16 ועם הנחת תנוצה מירבית של 10 פיקסלים לכל צד. נשים לב כי בצדדים השני והשלישי הייתה תנוצה גלובלית של המצלמה, וחיצי התנועה נראים כהזהה בצד השני וכזום פנימה בצד השלישי. בצד הראשון התנועה היא מקומית וכך יוצא שמרבית וקטורי התנועה קרובים לאפס.





ציור 9.6 - צמד שני ווקטוריו התנועה שהתקבלו ע"י אלגוריתם התאמת בлокים



ציור 9.7 - צמד שלישי ווקטוריו התנועה שהתקבלו ע"י אלגוריתם התאמת בлокים

9.3 שערוך תנועה בין תמונות - גישה מלאה

החסרון המרכזי בגישה שערוך התנועה שתוארה בסעיף קודם היא רמת הדיקוק שניתן לצפות ממנה. ברור כי חלוקה לבlokים מקילה על החישובים ועל צרכית הזיכרון - אך הדבר בא על חשבון דיקוק, כיון שההנחה שפיקסלים בבלוק נאים כמקשה אחת אינה נכונה. התחליף הוא למצוא וקטור תנועה לכל פיקסל, ובזה נ逋וק בסעיף זה. המשוואה אותה צריכים וקטורי התנועהקיימים היא המשוואת הבאה:

$$I(x + dx(x, y, t), y + dy(x, y, t), t) = I(x, y, t-1)$$

אותה כבר ראיינו בתחלת פרק זה בגרסה שונה מעט (דיסקרטית, עם תנועה הפוכה). שתי בעיות מתעוררות בבואה לעשות שימוש במשוואת זו:

1. הביטוי המתקבל הינו פונקציה מורכבת של הנעלמים.
2. הביטוי כופה אילוץ אחד על שני נעלמים - דבר המעיד כי חסר מידע בבואה לקבוע את וקטורי התנועה.

על מנת לפתור את הבעיה הראשונה מוצעת הגישה הדיפרנציאלית, אשר מהוות כיום את הרעיון המרכזי לשירות אלגוריתמי שערוך תנועה. הרעיון פשוט בעיקרו - נניח כי התנועה $[dx(x, y, t), dy(x, y, t)]$ קטנה יחסית (כלומר, מצויה בתחום $[-1, +1]$ פיקסלים). כיון לכך, נוכל לפתח את אגף שמאל של המשוואת לטור טיילור סביב הנקודה (x, y, t) ולהזניח איברים מסדר גדול מ-1. מקבל:

$$\begin{aligned} I(x, y, t-1) &= I(x + dx(x, y, t), y + dy(x, y, t), t) = \\ &= I(x, y, t) + \frac{\partial I(x, y, t)}{\partial x} dx(x, y, t) + \frac{\partial I(x, y, t)}{\partial y} dy(x, y, t) = \\ &= I(x, y, t) + I_x(x, y, t)dx(x, y, t) + I_y(x, y, t)dy(x, y, t) \end{aligned}$$

קיבלו משוואת ליניארית בנעלמים שלנו $[dx(x, y, t), dy(x, y, t)]$, ומקדמי משווה זו ידועים לנו. אנו מקבלים את שתי התמונות ויכולים לחשב גם נגורות אופקיות ואנכית של התמונה (x, y, t) . נשים לב כי הגורם החופשי הוא ההפרש בין התמונות העוקבות, וניתן להתייחס אליו כנגורת זמנית. פתרנו אם כך את הבעיה הראשונה אותה הצגנו. המשוואת שהתקבלה קוריה "משוואת אילוץ הבהירות" - (Brightness Constrained Equation - BCE)

במקרה הדיסקרטי علينا להשתמש בקירוב הנגורות כפי שהוצע בפרק 5. עברו המקרה הדיסקרטי רצוי מאד לבצע החלקה מקדימה על שתי תמונות הכניסה, על מנת לדוחות רוש עשוי להזיק בחישובי הנגורות, ועל מנת להפוך את ה-BCE למדוקיקת יותר. כאשר האוזר חלק יותר, הנגורות הגבוהות של התמונה חלשות יותר ואז ההזנחה שלנו בטור טיילור הקוצץ נכונה יותר.

נציג כעת שני אלגוריתמים לשערוך תנועה המשתמשים ב- BCE בובאים לשערוך את התנועה. האלגוריתם הראשון הוצע עיי Lucas ו- Kanade, והוא מציע מוטיב הדומה לתחליק השערוך בבלוקים אותו ראיינו בסעיף קודם. נניח כי רצוננו לשערוך את התנועה לפיקסל $[k_0, j_0]$ (והתחליק שיтвор יבוצע לכל פיקסל באופן דומה). נפתח חלון של 5 על 5 פיקסלים סביב פיקסל זה ונניח כי לכל הפיקסלים יש אותה תנועה, אותה נסמן $[dx[k_0, j_0, t], dy[k_0, j_0, t]]$. בכל פיקסל יש לנו משווה ליניארית אחת בה שני ערכים אלו הם הנעלמים - נרכז את 25 המשוואות שמתקבילות ונמצא את הפתרון הטוב ביותר במובן LS, דהיינו, אם נרכז את כל המשוואות באופן הבא :

$$A = \begin{bmatrix} I_x[k_0 - 2, j_0 - 2, t] & I_y[k_0 - 2, j_0 - 2, t] \\ \vdots & \vdots \\ I_x[k_0, j_0, t] & I_y[k_0, j_0, t] \\ \vdots & \vdots \\ I_x[k_0 + 2, j_0 + 2, t] & I_y[k_0 + 2, j_0 + 2, t] \end{bmatrix} \quad V = \begin{bmatrix} dx[k_0, j_0, t] \\ dy[k_0, j_0, t] \end{bmatrix}$$

$$B = \begin{bmatrix} I[k_0 - 2, j_0 - 2, t - 1] - I[k_0 - 2, j_0 - 2, t] \\ \vdots \\ I[k_0, j_0, t - 1] - I[k_0, j_0, t] \\ \vdots \\ I[k_0 + 2, j_0 + 2, t - 1] - I[k_0 + 2, j_0 + 2, t] \end{bmatrix}$$

$$\Rightarrow AV = B$$

אנו נגדיר את וקטור התנועה V להיות :

$$V_{\text{opt}} = \underset{V}{\text{ArgMin}} \|AV - B\|^2 = (A^T A)^{-1} A^T B$$

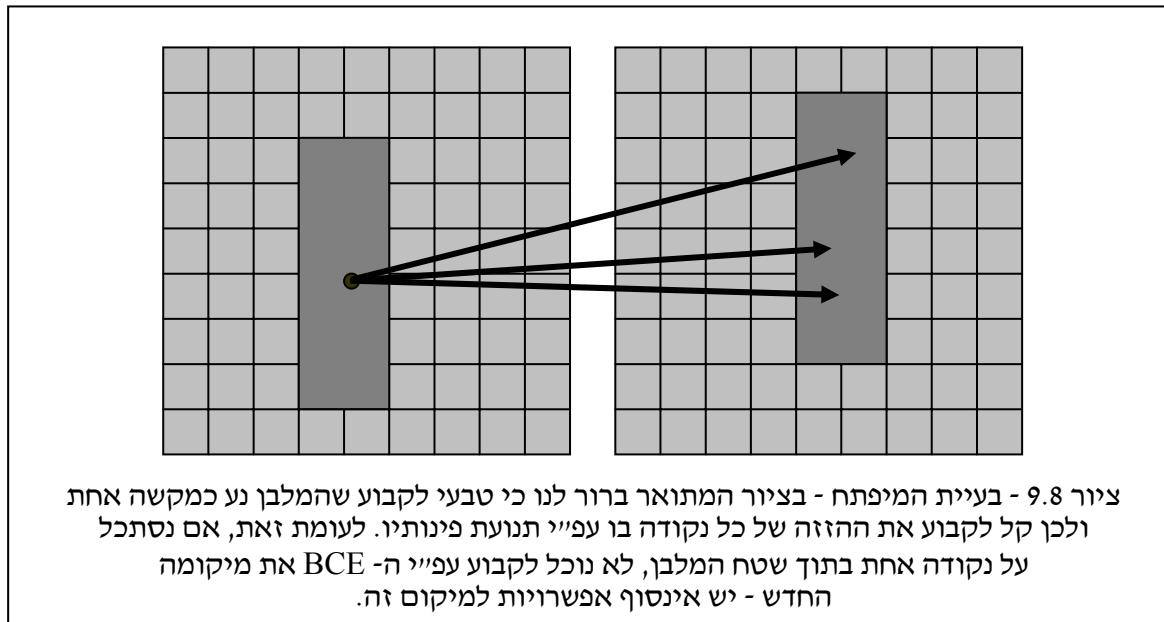
וכך, על-ידי ערכי התמונות ונגורותיהם קיבל את וקטור התנועה. המטריצה אותה יש להפוך כאן היא בגודל של 2 על 2, ולכן מרכיבות חישובית אינה מטרידת כאן. שיפור מסויים לאלגוריתם יכול להתקבל אם נכניס משקל לכל משווה לפי מרחקה מהפיקסל המרכז.

מסתבר כי אלגוריתם זה מניב תוצאות טובות מאוד בדרך כלל, כל עוד התנועה קטנה כדי שהנחנו. כאשר היא אינה קטנה, יש להפעיל את האלגוריתם, לקבל שדה תנועה, לעוות את התמונה $I(x, y, t)$ אל $I(x, y, t-1)$ לפי התנועה שהתקבלה ולהזור על התחליק. באופן זה התמונות מתקרבות זו אל זו בכל איטרציה והתנועה היא התנועה המctrברת.

תופעה מעניינת המתבלט באלגוריתם זה כמודר-לוואי היא שלכל פיקסל אנו יכולים לקבל מידע לאמינותה הוקטור שעבורו. לצורך חישוב הוקטור علينا להפוך את המטריצה $A^T A$, ואם מטריצה זו לא הפיכה ברור כי אין פתרון. מצב זה מתקבל כאשר כל השורות ב- A תלויות ליניארית זו בזו - למשל - אם כל השורות זהות. מצב זה מוכר בשם "תופעת המפתח" (Aperture Problem). ציור 9.8 ממחיש את

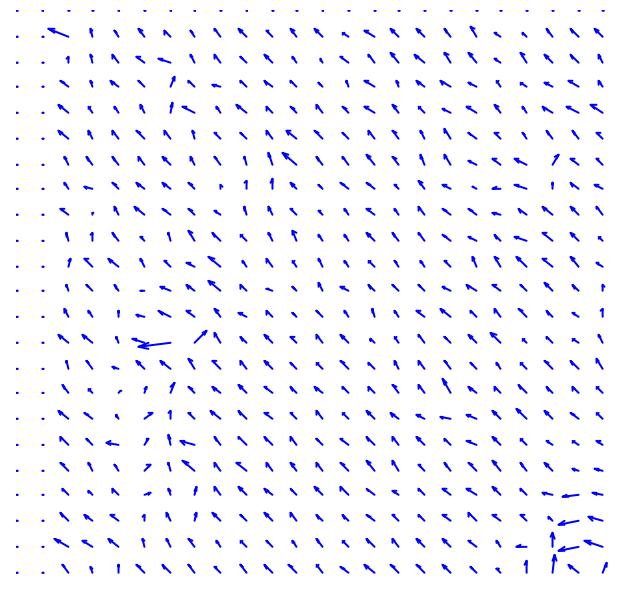
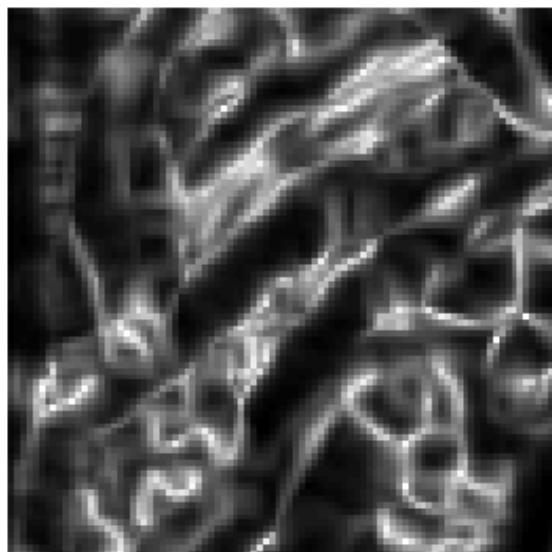
פרק 9 - עיבוד סדרות של תמונות

התופעה, ורואים כי כמשמעותם דרך "חור המנעל" על נתח תמונה, קשה לקבוע את מיקומה החדש אם איןנו עושים שימוש בקונטקט הסביבתי שלו. באופן זה תיארנו מצב קיצוני בו וקטור התנועה לא ניתן לחישוב. במקרים פחות קיצוניים, המטריצה הנחफכת אמנס אינה סינגולרית, אך קרובה לכך. מdad יפה למידת בריאותה הוא ה- Condition Number - הקטן חלקי הגדל. אם יחס זה קטן מס' מסוים נוכל לקבוע כי הוקטור המתקבל אינו כל כך אמין.



ציור 9.10 מראה צמד תמונות עם הזויה מלאכותית מדוייקת של פיקסל אחד שמאלה ולמעלה. ציור זה גם מראה כיצד נראה תמונות הנקודות האופקית, האנכית והזמןית. ציור 9.11 מראה את תוצאת השערוך של התנועה לפי אלגוריתם Lukas & Kanade. בתמונה האמיןויות ורואים כי אורי שפה מניבים שערוך טוב יותר.





ציור 9.11 - תוצאות שערוך תנועה עפ"י אלגוריתם Lukas & Kanade. למעלה - מפת האמיןויות של הוקטורים בהירים יותר מעידים על אמינות גדולה יותר. למטה - שדה התנועה המשוערך

בדוגמה זו בה התנועה היא $[+1, -1]$, הוקטור המומוצע בשדה הוקטורים המשוערך הוא $[0.937, 0.919]$. ע"י שינויים קליים כגון הגדלת חלון העבודה, הכנסת משקלות לאיברי החלון, ושינוי דרך ביצוע הנגזרת לגישה מדוייקת יותר, ניתן להגיע לדיקט טוב יותר.

אלגוריתם שני לביצוע שערוך התנועה מרכיב יותר אך יפה יותר, והוא פותח על ידי Horn ו- Schunck. אלגוריתם זה מציע להתייחס לשדה התנועה כולה בבעיות אופטימיזציה ולמוצא את כלו במקביל. נסמן בוקטור \underline{U} את כל קוואורדינטות ה- x של שדה התנועה שאינו ידוע, וב- \underline{V} את קוואורדינטות ה- y. אלה וקטוריים המתאימים מSIDEOR לקסיקוגרפיה של תמונות התנועה האופקית והאנכית, ממש כשם שאנו עושים זאת לתמונות רגילים. משוואות ה- BCE מצביעות את הקשר הבא:

$$\begin{aligned} & \begin{bmatrix} I_x[1,1,t] & 0 & I_y[1,1,t] & 0 \\ \vdots & & \vdots & \\ 0 & I_x[M,N,t] & 0 & I_y[M,N,t] \end{bmatrix} \begin{bmatrix} \underline{U} \\ \underline{V} \end{bmatrix} = \begin{bmatrix} I_t[1,1,t] \\ \vdots \\ I_x[M,N,t] \end{bmatrix} \\ \Rightarrow & [H_x, H_y] \begin{bmatrix} \underline{U} \\ \underline{V} \end{bmatrix} = \underline{Y} \end{aligned}$$

כאשר הנחנו כי התמונות הן בגודל N על M. למעשה הקשר הנ"ל הוא פשוט ריכוז של כל משווה אילוץ הבחרות לכל הפיקסלים לביטוי אחד. יש כאן MN משוואות עם MN^2 גלמים. השתמש כעת בתוכנת החלוקות הצפואה לשדה התנועה. אם נפעיל על "התמונה" \underline{U} גזירה כגון לפלייאן, נרצה כי נורמות התוצאה תהיה נמוכה. באופן זה נקבע חלקות על \underline{U} ועל \underline{V} , ממש כשם שעשינו זאת בדיאון בשחזר תמונות. לכן, פונקציית המחריר אותה נזעך היא:

$$\begin{aligned} \{\hat{\underline{U}}_{\text{opt}}, \hat{\underline{V}}_{\text{opt}}\} &= \underset{\underline{U}, \underline{V}}{\text{ArgMin}} \quad \left\| H_x \underline{U} + H_y \underline{V} - \underline{Y} \right\|^2 + \lambda \left[\|D\underline{U}\|^2 + \|D\underline{V}\|^2 \right] = \\ \Rightarrow \begin{bmatrix} \hat{\underline{U}}_{\text{opt}} \\ \hat{\underline{V}}_{\text{opt}} \end{bmatrix} &= \begin{bmatrix} H_x H_x + \lambda D^T D & H_x H_y \\ H_y H_x & H_y H_y + \lambda D^T D \end{bmatrix}^{-1} \begin{bmatrix} H_x \underline{Y} \\ H_y \underline{Y} \end{bmatrix} \end{aligned}$$

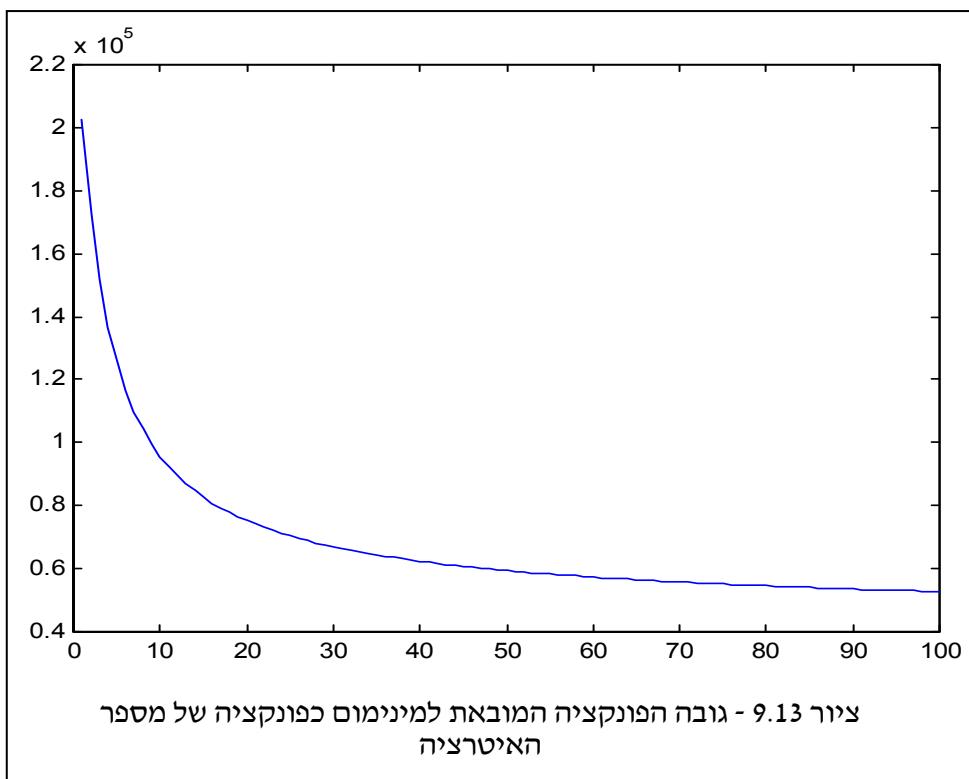
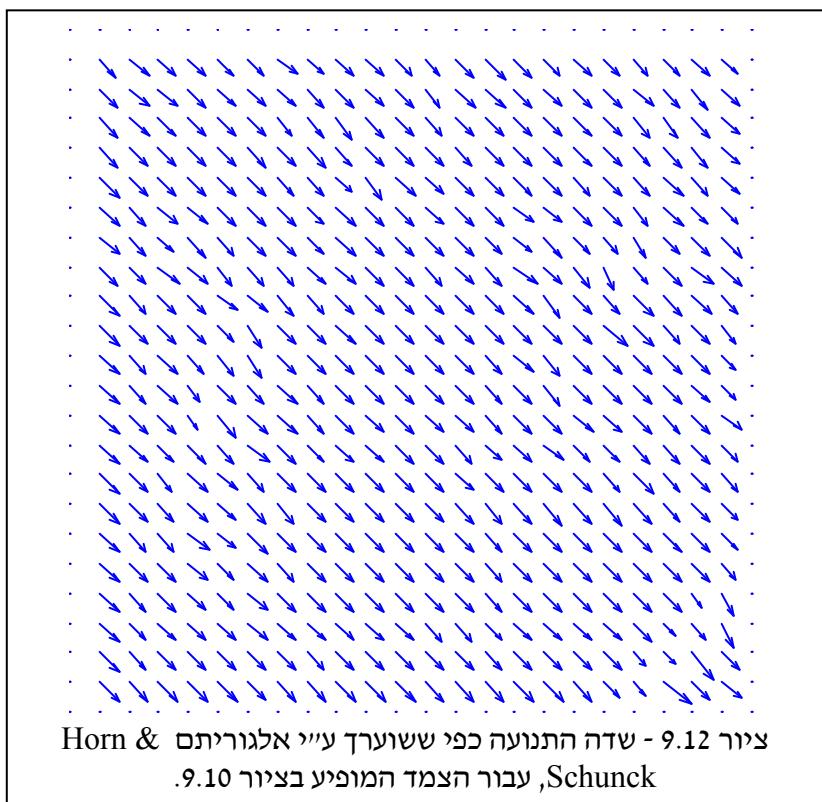
הגודל λ שולט במידת החלוקות הנכפיה על הפתרון - עבור ערך גבוה מאוד נקבל שדה התנועה מאד חלק, ועבור ערך נמוך נרצה שינויים מרוחביים בוקטוריים. כמו בשחזר תמונות, היפוך המטריצה המתתקבל אינו בא בחשבון, וכתחליף מוצעת גישה איטרטיבית לפי ה- SD, בה מתחילה בניחוש (נניח אפסים), ומעדכנים הפתרון לפי:

$$\begin{bmatrix} \hat{\underline{U}} \\ \hat{\underline{V}} \end{bmatrix}_{k+1} = \begin{bmatrix} \hat{\underline{U}} \\ \hat{\underline{V}} \end{bmatrix}_k - \mu \left\{ \begin{bmatrix} H_x H_x + \lambda D^T D & H_x H_y \\ H_y H_x & H_y H_y + \lambda D^T D \end{bmatrix} \begin{bmatrix} \hat{\underline{U}} \\ \hat{\underline{V}} \end{bmatrix}_k - \begin{bmatrix} H_x \underline{Y} \\ H_y \underline{Y} \end{bmatrix} \right\}$$

גם באלגוריתם זה, אין צורך לבנות מטריצות או לייצג את הנעלמים בפועל בוקטוריים. הפעולות המתבצעות במשואה זו על תמונות הנעלמים \underline{U} ו- \underline{V} פשוטות - הפעלת לפלייאן כקונבולוציה, וחכלה במטריצות אלכסוניות כפירשו הכפלת התמונות \underline{U} ו- \underline{V} איבר-איבר באיברי האלכסון הראשי.

פרק 9 - עיבוד סדרות של תמונות -

בצירור 9.12 מובאות תוצאות של אלגוריתם זה לאוטו צמד כבתמונות 9.10 ו- 9.11. בחרנו $\lambda = 100$ וכן נקבע חלוקה חזה על הוקטורים. הוקטור הממוצע הוא $[0.9952, -0.9921]$. בצייר 9.13 מובא הגרף המראה את גובה הפונקציה שהובאה למינימום כפונקציה של מספר האיטרציה.



4.9 ניקוי רעש בסדרות של תמונות

ניבור כתה לדון באפליקציות העושות שימוש בשערוך תנועה. במסגרת סעיף זה נראה כיצד מנקים סידרת התמונות מרעש אדיטיבי גאוסי ולבן. לצורך יישום זה נדרש להשתמש בשערוך תנועה מדויק ככל האפשר, וכך אנו נניח כי אלגוריתם כדוגמת האלגוריתמים הדיפרנציאליים משמש אותנו.

נתחיל את הדיון בקביעת סימונים. סידרת התמונות המקורית (לא רעש מסומנת ב-) $\underline{X}(t)$, בה כל איבר הוא וקטור המתאר תמונה בגודל M על N פיקסלים. הشرط הפגום הוא $\underline{Y}(t)$, אשר גם הוא בניו מתמונות בגודל M על N , אך קיים הקשר הבא:

$$\underline{Y}(t) = \underline{X}(t) + \underline{N}(t)$$

כלומר, התמונה $\underline{Y}(t)$ היא גרסה רועשת של $\underline{X}(t)$. הרעש כאמור לבן, וזה אומר כי ערכיו במקום ובזמן מפולגים באופן זהה (פיילוג נאוסי) וחסר תלות. בהגעה לתמונה $\underline{Y}(t)$, רצוננו להפיק שערוך מוצלח ונקי מרעש שיסמן $\underline{\hat{X}}(t)$. אנו נניח מערכת סיבתית, הכוללת מערכת שבה $\underline{\hat{X}}(t)$ נבנה על-פי מידע שכבר התקבל ולא עפ"י תנועות עתידיות. הנחה אחרתה היא שאנו יכולים לשערך את התנועה בין כל צמד תמונות (ואפילו רועשות) בבדיקה מעוללה. לכן, כנסנו שערך את התנועה בין הצמד $\{\underline{X}(t-k), \underline{X}(t)\}$ קיבל אותה תוצאה כמו עבור הצמד $\{\underline{Y}(t-k), \underline{Y}(t)\}$. פועלות העיונות הגיאומטרי הזו מסומן ב- $G(t, k)$ - זהה מטריצה המשמעותת התנועה שיש להסביר את $\underline{X}(t-k)$ על מנת לקבל את $\underline{X}(t)$, וכבר פגשנו את הפעולה ההופוכה לה F בתחילת פרק זה. לכן קיים הקשר:

$$\underline{X}(t) = G(t, k) \underline{X}(t-k)$$

השוויון במשווה זו אינו מדויק כיון שככל ש- k גדול יותר כך יותר מידע חדש מוזן לתמונה עם הזמן וכן זה שווין מוקרב, ואיכותו נופלת עם הגדלת k . לו עבדנו עם תמונה בודדה, ורצינו לנוקות רעש לפי המתודולוגיה של פרק 6 - שחזור תמונות - היינו מציגים את המינימיזציה של הפונקציה הבאה:

$$\hat{\underline{X}}_{\text{Intra}}(t) = \underset{\underline{X}(t)}{\operatorname{ArgMin}} \left\| \underline{Y}(t) - \underline{X}(t) \right\|^2 + \lambda \left\| D\underline{X}(t) \right\|^2$$

כאשר האיבר הראשון בפונקציה המחייב כופה קירבה ל- \underline{Y} , והאיבר השני כופה חלקות מרחבית ע"י פעולה גזירה כלשהי. הפתרון קריוי Intra מכיוון שהוא אינו פועל בין תמונות (Inter-frame) אלא בתוכן (Intra-frame). בשל כך גם אין כל שימוש בתנועה ביןטיים. נניח כי אנו רוצים לצרף למשחק את התמונה הקודמת. נוכל להציג את הפונקציה הבאה:

$$\hat{\underline{X}}_{In}(t) = \underset{\underline{X}(t)}{\operatorname{ArgMin}} \left\| \underline{Y}(t) - \underline{X}(t) \right\|^2 + \alpha \left\| G(t,1) \underline{Y}(t-1) - \underline{X}(t) \right\|^2 + \lambda \left\| D \underline{X}(t) \right\|^2$$

כעת כפינו על הפתרון $\hat{\underline{X}}(t)$, לאחר עיות נאות, להיות קרוב לתמונה הקודמת $\underline{Y}(t-1)$. המקדם α יהיה מספר כלשהו קטן במקצת מ-1 הקובע את חשיבות הקרבה החדשה ביחס לקרבה לתמונה הנוכחית. נוכל אז להציג ביטוי מורכב יותר אשר יביא בחשבון את כל התמונות בסדרה מזמן $t=0$. ביטוי זה הוא :

$$\hat{\underline{X}}(t) = \underset{\underline{X}(t)}{\operatorname{ArgMin}} \sum_{k=0}^t \alpha^k \left\| G(t,k) \underline{Y}(t-k) - \underline{X}(t) \right\|^2 + \lambda \left\| D \underline{X}(t) \right\|^2$$

נשים לב כי כל איבר הוכפל במשקל α^k כך שהתמונות הקרובותתקבלנה משקל גדול, והרחוקות משקל הדועך לאפס. נפתרו בעית אופטימיזציה זו ונראה כי פתרונה אינו מורכב. גזירה והשוואה לאפס נתונים :

$$2 \sum_{k=0}^t \alpha^k [\underline{X}(t) - G(t,k) \underline{Y}(t-k)] + 2\lambda D^T D \underline{X}(t) = 0$$

$$\left[\sum_{k=0}^t \alpha^k I + \lambda D^T D \right] \hat{\underline{X}}(t) = \left[\frac{1-\alpha^{t+1}}{1-\alpha} I + \lambda D^T D \right] \hat{\underline{X}}(t) = \sum_{k=0}^t \alpha^k G(t,k) \underline{Y}(t-k)$$

לכוארה נראה כי עליינו לזכור את כל התמונות בסדרה, מנקודת הזמן 0 ועד זמן t , לשם מציאת התמונה הנוכחית במוצא. בפועל נוכל להשתמש בקשר

$$G(t,k) = G(t,1)G(t-1,1)\cdots G(t-k+2,1)G(t-k+1,1)$$

ולקבל

$$\begin{aligned} \underline{P}(t) &= \sum_{k=0}^t \alpha^k G(t,k) \underline{Y}(t-k) = \sum_{k=0}^t \alpha^k G(t,1) \cdots G(t-k+1,1) \underline{Y}(t-k) = \\ &= \underline{Y}(t) + \alpha G(t,1) \sum_{k=1}^t \alpha^k G(t-1,1) \cdots G^T(t-k+1,1) \underline{Y}(t-k) = \\ &= \underline{Y}(t) + \alpha G(t,1) \sum_{j=0}^{t-1} \alpha^k G(t-1,1) \cdots G^T(t-j,1) \underline{Y}(t-j+1) = \\ &= \underline{Y}(t) + \alpha G(t,1) \underline{P}(t-1) \end{aligned}$$

כלומר, על מנת לחשב את $\hat{\underline{X}}(t)$ علينا לחת את התמונה $\underline{P}(t-1)$ (שנניח שמצויה בידינו), לעדכנה לפי המשוואה הרקורסיבית הניל', ולפתור את מערכת המשוואות

$$\left[\frac{1}{1-\alpha} I + \lambda D^T D \right] \hat{\underline{X}}(t) \cong \underline{P}(t)$$

השוויון בערך נובע מכך שהנחנו כי α גדול דיו ולכן $0 \leq \alpha$. מערכת משווהות זו יכולה להיפטר באופן איטרטיבי ע"י אלגוריתם ה- SD הצעוד בכיוון הגרדיאנט :

$$\hat{\underline{X}}_{k+1}(t) = \hat{\underline{X}}_k(t) - \frac{\mu}{1-\alpha} \hat{\underline{X}}_k(t) - \lambda D^T D \hat{\underline{X}}_k(t) + \mu \underline{P}(t)$$

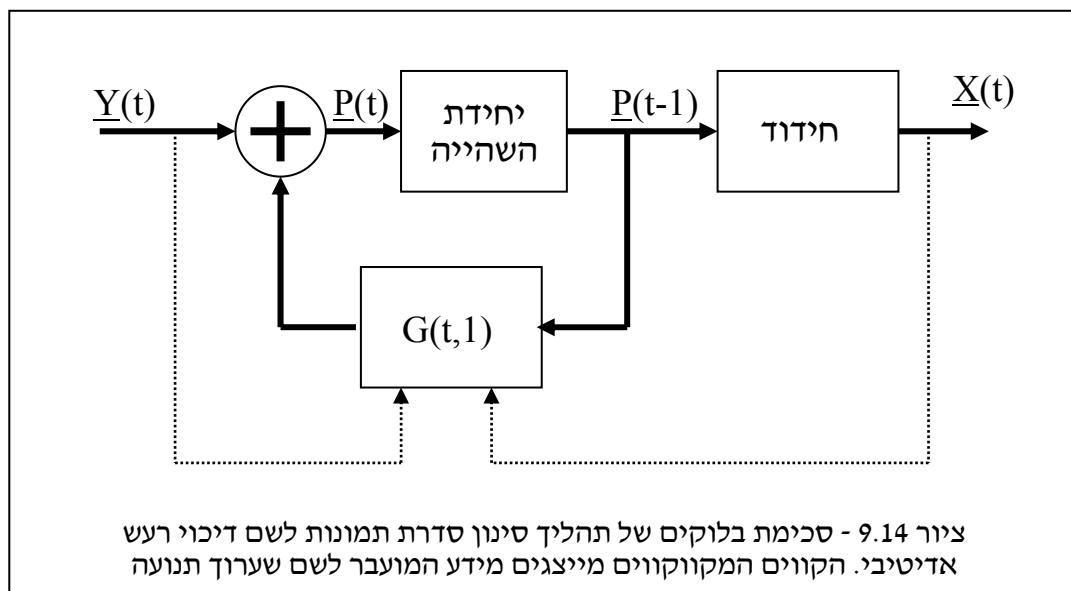
כתחליף, כיוון שהמטריצה הנחפה הינה מטריצה במבנה בלוק-סיבובי, ניתן להפעיל בהיפוך כMSN ליניארי קבוע במקום.

הוקטור $\underline{P}(t)$ אינו אלא ממוצע זמן של התמונות בסדרה, כאשר בממוצע זה ישן שתי תופעות חשובות - (I) קיים משקל דועץ אקספוננציאלית לתמונות רחוקות, ו- (II) התמונות מקוזזות תנועה בטרם תיכנסנה לממוצע. התהליך האיטרטיבי נועד לבצע חידוד על תמונה ממוצע זה. אם נקבע $\lambda = 0$, כלומר לא מרכיב של חלקות מרוחبات בפתרון, יתקבל כי :

$$\begin{aligned}\hat{\underline{X}}(t) &= (1-\alpha)\underline{P}(t) = (1-\alpha)\underline{Y}(t) + (1-\alpha)\alpha G(t,1)\underline{P}(t-1) \\ &\Rightarrow \hat{\underline{X}}(t) \approx (1-\alpha)\underline{Y}(t) + \alpha G(t,1)\hat{\underline{X}}(t-1)\end{aligned}$$

קיבלנו כי $\hat{\underline{X}}(t)$ מתקבל ע"י משואה רקורסיבית מסדר ראשון -MSN בעל קווטב יחיד בזמן.

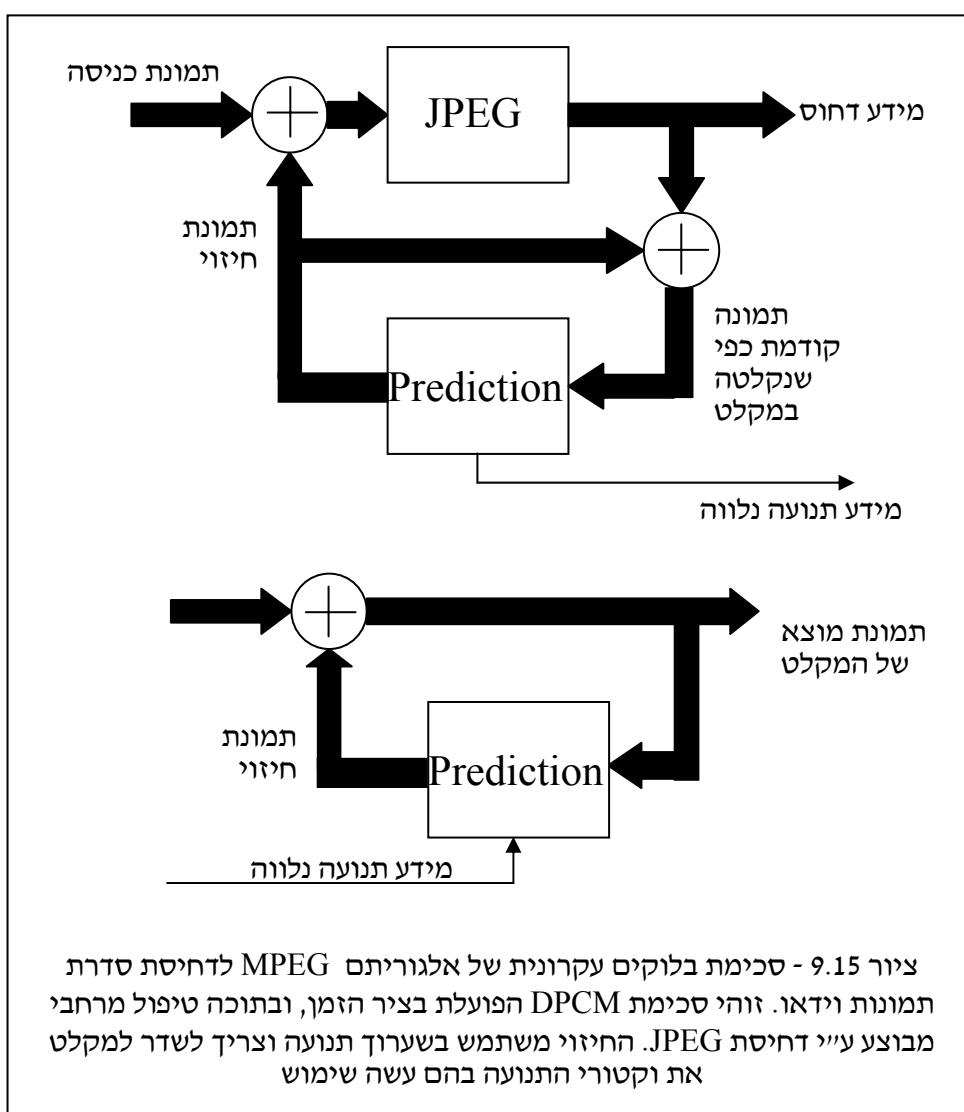
נקודה חשובה בתהליך זה היא דיקוק התנועה בה נעשה שימוש. על המטריצה G לייצג תנועה מדויקת בדיקוק תת-פייסלי, אם נרצה תוצאה באיכות טוביה. לכן רצוי לבצע את שערוך התנועה בין $\hat{\underline{X}}(t-1)$ ו- $\underline{Y}(t)$, ולא בין $\underline{Y}(t-1)$ ל- $\underline{Y}(t)$. ציור 9.14 מראה סכימת בלוקים של פעולה הניתני לסדרות של תמונות.



9.5 דחיסת סדרות של תמונות ומבוא ל- MPEG

בסעיף זה נדנו בדחיפסה של סדרות של תמונות כאשר הדחיפסה נעשית *On-Line*, דהיינו, על כל תמונה בהגעה. זאת בניגוד לדחיפסה של סדרת תמונות כמידע תלת-מימדי השמור על הדיסק כמקרה אחד. כזכור, דחיסת מידע מצריכה ניצול של היתירותיות באותן הנדחס. יתרות בסדרת תמונות הינה יתרות הסתברותי, יתרות מרחבית, ולראשונה גן - יתרות זמנית, שהרי כבר אמרנו שתמונות עוקבות כמעט זהות למעט תנועה.

במסגרת סעיף זה נציג אלגוריתם דחיסה אחד מיני רבים אפשריים. ייחודה של אלגוריתם דחיסה זה בפשטותו היחסית וביצועיו הטובים. למעשה, אלגוריתם זה הוא הגרעין המרכזי באלגוריתם ה- *MPEG* המהווה כיום את הסטנדרט לדחיסת וידאו. הרעיון המרכזי באלגוריתם זה הוא ניצול התהיליך שהוגדר באלגוריתם ה- *JPEG*, והכללו לו לכך שימוש על סדרות של תמונות. הרעיון הוא להשתמש בחיזוי לצורך מציאו היתירות הזמנית, והתמורה לצורך מציאו היתירות המרחבית. סכימת בלוקים של אלגוריתם זה מתוארת בציור 9.15.



כפי שניתן לראות, הסכימה היא סכימת DPCM קלאסית בה העורכים מעבירים תמונות שלמות. בהגיעה התמונה העכשווית $(t)X$, נבנית עבורה תמונה חייזי אותה נסמך ב- $(t)\tilde{X}$. תמונה זו נבנית ע"י ליקיחת התמונה הקודמת כפי שנקלטה במקלט, $(t)\hat{X}$, וביצוע קיזוז תנועה עליה לפי תנועה לשוערכה בין שתי התמונות האחרוניות. וקטורי התנועה לא מוצאים במקלט, ועל מנת שהמקלט ישמש באותו חייזי בדיק עלינו לשדר את וקטורי התנועה. לשם כך רצוי לעשות שימוש בשערוך תנועה אחד לבlok של פיקסלים, בו יש מעט וקטוריים לשדר. בשל כך, תמונה החיזי נבנית באופן הבא:

1. חלק את התמונה הנוכחי $(t)X$ לבוקים בגודל של 16 על 16 פיקסלים הממלאים את שטח התמונה.
2. בנה תמונה חייזי שתסמן $(t)\tilde{X}$ - נאתחל תמונה זו כתמונה אפסים בגודל של $(t)X$.
3. לכל בлок מהבלוקים הנ"ל חפש את מיקומו בתמונה הקודמת $(t-1)\hat{X}$. בהנחה כי נמצא המיקום, קח את הבלוק המתאים מהתמונה $(t-1)\hat{X}$ והעתק אותו למיקום הבלוק ב- $(t)\tilde{X}$. הסיבה לכך שאנו עובדים עם $(t-1)\hat{X}$ ולא עם $(t-1)X$ (התמונה המקורית) היא שרצוננו לקבל אותה תמונה חייזי במקלט.
4. באופן זה אנו מטפלים בכל בлок וממלאים את כל שטח התמונה $(t)\tilde{X}$. התוצאה היא תמונה חייזי המנסה להיות קרובה ככל האפשר לתמונה $(t)X$ - זהה בדיק מהותה של תמונה חייזי.

לאחר בניית החיזי $(t)\tilde{X}$, מחוסרת תמונה זו מהתמונה המקורית ומקבלים את תמונה שגיאת החיזי:

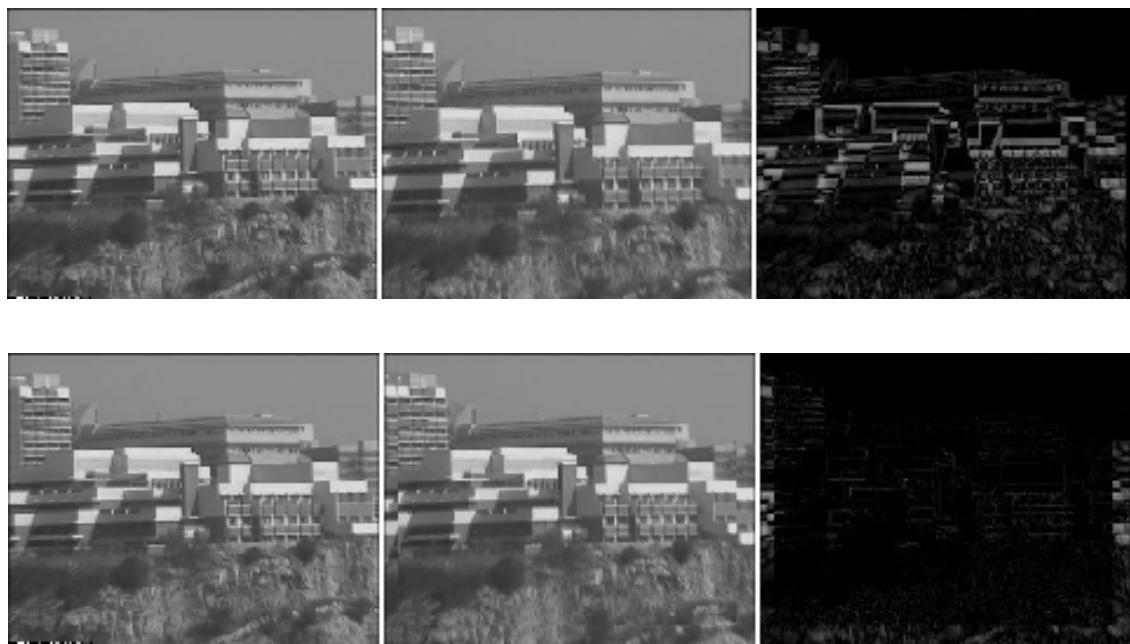
$$E(t) = X(t) - \tilde{X}(t)$$

תמונה הפרש זו אמורה להיות חלה ומרוכזת סביב האפס. אם בסכימת ה- DPCM המקורית שגיאת החיזי הייתה סקלאר ועbara קוונטייזציה, הרי שגם אנו נזקקים לקוונטייזציה, אך כיוון ש- $E(t)$ תמונה ויש בה עדין יתרות מרחבית, רצוי לנצלה. זאת עושים ע"י דחיסה המזכירה את אלגוריתם ה- JPEG - התמונה $E(t)$ מחלוקת לבוקים, כל בлок עבר התמרת DCT דו-מימדית, ומקדמי ההתרמה עוברים קוונטייזציה לפי חשיבותם. התוצאה בשלב זה היא התמונה $\hat{E}(t)$ - אשר אמורה להיות דומה לתמונה $E(t)$ למעט המידע שנזרק.

במקלט מתקבלת התמונה $(t)\hat{E}$ ואליה מחוברת תמונה החיזי $(t)\tilde{X}$ - זהה אותה תמונה חייזי כבמשדר, כיוון שהיא נבנית עפ"י אותם וקטורי תנועה ועפ"י התמונה $(t-1)\hat{X}$. התוצאה היא:

$$\hat{X}(t) = \hat{E}(t) + \tilde{X}(t) = E(t) + \tilde{X}(t) - [E(t) - \hat{E}(t)] = X(t) - [E(t) - \hat{E}(t)]$$

ושוב קיבלנו כי שגיאת סכימת ה- DPCM אינה אלא זו שמוכנסת ע"י הקוונטייזציה בבלוק ה- JPEG. צייר 9.16 מראה צמד תמונות, ההפרש ביניהן, תמונה החיזי הנבנית ושגיאת החיזי.

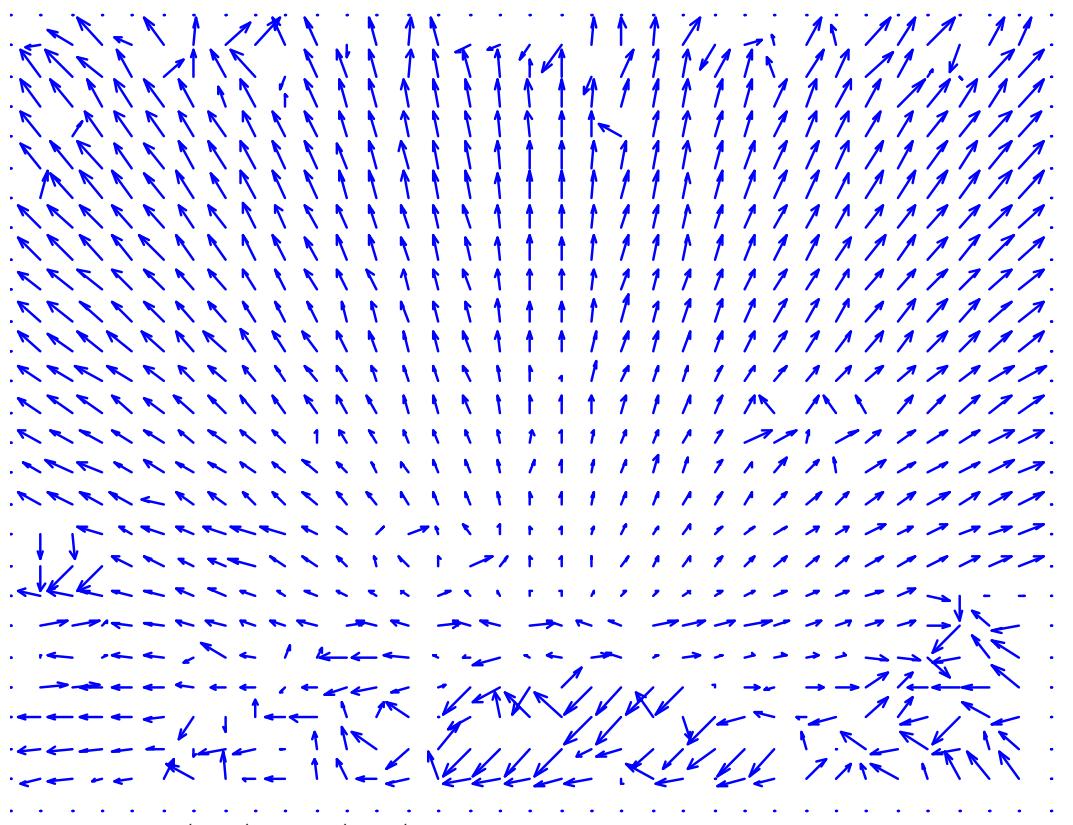


צירור 9.16 - למעלה : צמד תמונות בעקבות סדרת וידאו (משמאלי) ותמונות ההפרש ביןיהם בערך מוחלט (מיימן). למטה : התמונה העכשווית (שמאל), תמונה החיזוי הנבנית מתוך התמונה הקודמת ע"י קיזוז תנעה בבלוקים של 8 על 8 (אמצע), ותמונה שגיאת החיזוי (ימין).

בתמונה זו, אנרגיית השגיאה בין שתי התמונות (הפרש הפיקסלים ברייבור המוחולק במספר הפיקסלים הוא 1172, בעוד שאנרגיית שגיאת החיזוי היא 235 - יחס של כ- 4. באלגוריתם שיעוש בחרכנו לא לטפל בבלוקים המצוים בקצוות התמונה על מנת להימנע מהתופעה קצה המציגות טיפול ייחודי. בשל כך השגיאה בקצוות גדולה יותר מאשר במרכז. כאשר מסירים את השולטים, האנרגיה של השגיאה המקורית היא 1265 ואילו זו של החיזוי היא 65 - יחס של קרוב ל- 20.

התמונה הנבנית כתמונה חיזוי היא תמונה הבניה כקולאז' של בלוקים המועתקים ממיקומים שונים של התמונה הקודמת בסרט הנקלט. בשל כך, לא מפתיע לגלוות כל מיני תופעות של גבולות צורמים של בלוקים בתמונה זו. למרות שבצמוד לתמונה 9.16 קיבלנו שגיאה נמוכה, גם בה קיימת אונה תופעה. תמונה 9.17 מראה את תמונה החיזוי, וניכר כי בתחום בצד ימין מורגשת התופעה ביותר אחרת. ביתר התמונה לא רואים את הבעה, בשל אופיו של היעד המוצלים הבניי במלוקים כמעט אחידים עם שפות אנקיות ואופקיות, ובשל התנועה פשוטה בין שתי התמונות. בציור זה גם מובאת מפת התנועה שע-ל-פיה נבנה החיזוי. מפה זו ניכר כי ברוב האזורים התנועה חלקה ולכנן לא צפוי לקבל קווי תפראים בין הבלוקים. בתחום התמונה החיצים מבולבלים ולכנן שם יש סיכוי סביר לגלוות את התופעה בחrifיות יתר.

אם סרט וידאו דורש זיכרון של 30 תמונות בגודל של 500 על 500 פיקסלים בשניה, הרי שמדובר בקצב של 7.5 MB לשניה. אלגוריתם ה- MPEG יכול להוריד קצב זה ביחס של 10-200 (תלו依 באיכות הנחוצה). הפטונציאלי הטמען ב יתרות הזמן גדול והוא מKENה לטיפול בסדרות יחס דחיסה גדול יותר עבור אותה שגיאה ריבועית לתמונה נתונה.



ציור 9.17 - תמונה החיזוי ומפת התנועה לפי בלוקים של 8 על 8

פרק 10 – שחזור תמונה מהיטלים

219	פרק 10 – שחזור תמונה מהיטלים
220	10.1 בעיות שחזור מהיטלים וניסוח המתמטי
221	10.2 התמרת ראדו
223	10.3 משפט החתק המרכזי והיפוך התמרת ראדו
226	10.4 שחזור מהיטלים ע"י גישות תדר
228	10.5 שחזור מהיטלים ע"י קונבולוציה
229	10.6 שחזור מהיטלים ע"י גישות אלגברית

בפרק 6 דנו בשחזור (Restoration) תמונה מקלקלול בעל מודל ידוע, וראינו כלים מובססי שערוך סטטיסטי נרთמיים לשם פתרון בעיות מסווג זה. למעשה, בפרק 6 הנחנו את התשתיות המתמטיות לפרק מרכזי בעיבוד אות הקרווי בעיות היפוך (Inverse Problems). בעית היפוך היא בעיה בה נתון סט מדידות, נתון קשר מתמטי מוגדר היטב (ולאו דווקא דטרמיניסטי – ראה עניין הרוש בשחזור תמונה) בין מדידות אלו ובין איזשהו אות אידיאלי רצוי, ועלינו לבצע היפוך לשם הפקת האות הרצוי מהמדידות המצוויות. בשחזור תמונה המדידות הן התמונה המקולקלת, האות האידיאלי היא התמונה המקורית והנקיה, והקשר בין השתיים ידוע כנובע מאופרטור ליניארי (טשטוש) ותוספת של רעש בעל מבנה סטטיסטי ידוע.

בפרק זה נדון בשחזור (Reconstruction) תמונה מתוך סט מדידות שלה אשר מתקבל כהטלוות. מסתבר כי מבנה מעין זה נפוץ בצלום רפואי, מדידות עקיפות של קרקע, חישה מרוחק ע"י מכ"ם, ועוד. גישת צילום עקיפה זו ידועה בשם טומוגרפיה ממוחשבת (Computerized Tomography) (CT). אנו ראה כי מבנה בעיה זו נופל אף הוא לקטגורית בעיות ההיפוך, וככזה, כלים שכבר פגשנו הם הכלים הנאותים לטיפול בעיות השחזור, אלא שנדרשות התאמות בשל השוני הקיים. גם כאן נראה כי הקשר בין המדידות לאות האידיאלי ניתן לביטוי כמערכת ליניארית והוספה רעש, אלא שהפעם לממדידות אין הקשר ויזואלי – אלה הן הטלות חד-ממדיות שיחד מכילות את המידע המאפשר חזרה לתמונה המוטלת. וזהו הסיבה לאבחנה בין המושג Reconstruction ובין Restoration.

לצד הצגת נושא שחזור מהיטלות כבעיתת היפוך ושימוש בכלים שכבר הוצגו לשם בניית התמונה מהיטളותיה, נציג גם יסודות מתמטיים של אופטור ההטלה, התמרת ראנון הנובעת ממנה, ויצוג תדרי של כלל הבעיה. ליצוג זה פוטנציאלי חשוב בביטוי שחזור מהיר של תמונה, אך, אם להודות על האמת, שחזור זה סובל מאותן חולשות מוכרות שגרמו לנו לתהוות החמצה בדיון בשחזור תמונה מטשטוש.

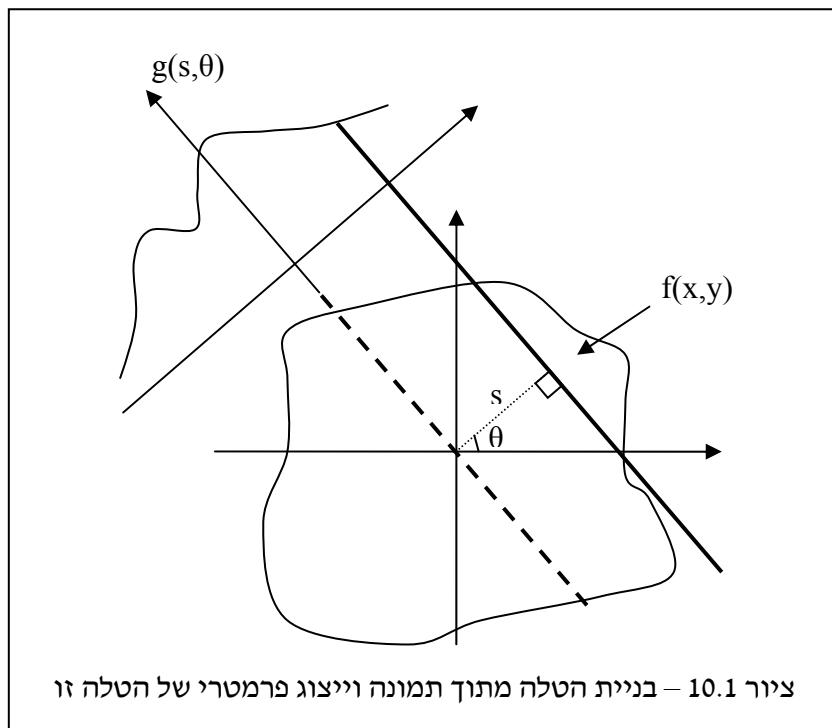
הצגת נושא זה מובאת לשם השלמות של ההצעה, והיכרות עם מושגי היסוד בתחום.

10.1 בעיות שחזור מהיטלים וניסוחו המתמטי

נתחיל את הדיון בפונקציה דו-מימדית המוגדרת מעל הרץ – $f(x,y)$. פונקציה זו מייצגת את התמונה אותה רצוננו לשחזר. אנו מניחים שתמונה זו מייצגת צפיפות חומר דו-מימד כפונקציה של המקום (x,y) . נניח כי קרון בעוצמה התחלתית I_0 עוברת דרך משטח החומר $f(x,y)$, נעה לאורך הישר L , ומכלט מתעד את עוצמת הקרון בموقع החומר. בכל אינטראול אינפיטיסימלי באורך L_p בו עוברת הקרון לאורכו של הישר היא מונחת בעוצמה יחסית $\exp\{-f(x,y)dL\}$ (בהנחה שביחידת אורך קרצה זו (x,y) קבוע בערכו). לכן, לאחר מעבר לכל אורכו של הישר נקבל כי העוצמה הנקלטה בموقع היא מכפלת הניחותים. מכפלה זו מוגרת לסכום המעריכים באקספוננט,

$$I = I_0 \cdot \exp\left\{-\int_L f(x,y) dL\right\}$$

אם נניח כעת כי למעשה מסרק מקבילי של קרניים חלף דרך החומר, נקבל פונקציה של ערכי מוצא במקומות ערך בודד כלשהו. מבנה זה מתואר בציור 10.1.



אנו ניצג כל קרן ישרה L ע"י שני פרמטרים – זוויתו ביחס לאופק θ ומרחקו מהראשית s כמו שמתואר בציור. תחת יצוג זה, הפונקציה לאחר החלטה היא

$$I(s, \theta) = I_0 \cdot \exp \left\{ - \int \int_{x, y} f(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy \right\}$$

בhinaten משפחת פונקציות $I(s, \theta)$ המתאפיינת לזרויות הטלה שונות בתחום $\pi < \theta \leq 0$ ולהסתה מהראשית בתחום $-\infty < s < \infty$, רצוננו לבנות את $f(x, y)$. נגדיר את הפונקציה $g(s, \theta) =$

$$g(s, \theta) = -\log \left\{ \frac{I(s, \theta)}{I_0} \right\} = \int \int_{x, y} f(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy$$

אנו רואים כי בין החלטות המודדות ובין פונקציית היעד לשחזור יש קשר ליניארי פשוט. קשר זה מתרחש את המודל הנוכחי בעיית ההיפוך – אנו נוכל קשר זה בובאו לפטור את בעיית ההיפוך זו.

נתמיכת בקצרה לסוגי הדמיה הניטנים לתיאור הניל': כאשר התמונה $f(x, y)$ הינה פונקציה צפיפות הריקמה בחalk גוף כלשהו, והקרן הנורית הינה קרן רנטגן, אנו מקבלים סידרת הטלות רנטגן והשחזור הנוצר ידוע בשם CT-X-Ray. הטלות יכולות גם להתקבל בדרכים אחרות כגון החזרה מטוויה לא ידוע, או מקורות המצויים בתחום התווך. למשל, בהשא סיסמית בה המטרה היא לקבל מפת צפיפות הקרקע מתחת לפני השטח, נוצר הלם קול (בדיב' ע"י מוט חומר נפץ שהוכנס לקדח בקרקע), והחישה נעשית ע"י סידרת חיישני רעד או קול הפרוסים בשטח. כדוגמה אחרת, בצלום PET (Positron Emission Tomography) חומר רדיואקטיבי המצוי בתחום הרקמות בשיעורים שונים מתפרק ומשחרר אגב כך קרינת גמא. במקרה זה המקורות נמצאים בתחום הגוף, ומיקומם כמותם ביחידת שטח היא הנעלם אותו נשוחר. המידע הנtentoon הוא תיעוד אנרגטי הקרינה מחוץ לגוף. כדוגמה שלישית נזכיר את צילום ה- MRI בו שדה מגנטי מבוקר משנה את הספינים (כיוון המגנט) של גרעיני אטומי החנקן, ועם מדידה של שינויים אלו מתקבלת תמונה הטלה של ריכוזי הנזולים בריקמה. דוגמאות אלו, וישן גם אחרות, ממחישות דרכים שונות להשיג מידע על גודל בלתי ניתן למדידה ע"י קבלת הטלות או מידע המתייחס להטלות בדרך כלשהי.

10.2 התמרת רצון

למעשה הוגדרה התמרת רצון לעלה בהגדרת משפחת הטלות. בהינתן פונקציה המוגדרת מעל הקואורדינטות הקרטזיות, $f(x, y)$, התמרת רצון שלה שתסמן $\mathfrak{R}\{f(x, y)\}$ נתונה ע"י

$$g(s, \theta) = \mathfrak{R}\{f(x, y)\} = \int \int_{x, y} f(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy$$

כאשר תחום הגדרת הפונקציה g הוא $\pi < \theta \leq 0$ ו- $s < \infty$. כל ערך (s, θ) מהו אינטגרציה על פני ישר אחד על הפונקציה $f(x, y)$. מעניין לציין כי אם נחיל את התמורה עם תמונה f המוגדרת מעל קוואורדינטות פולריות (r, φ) , נקבל כי התמרת רצון תהיה

$$\begin{aligned}
 g(s, \theta) = \Re\{f(x, y)\} &= \iint_{x, y} f(x, y) \delta(x \cos \theta + y \sin \theta - s) dx dy = \\
 &= \int_{r=0}^{\infty} \int_{\varphi=0}^{2\pi} r f(r, \varphi) \delta(r \cos \varphi \cos \theta + r \sin \varphi \sin \theta - s) dx dy = \\
 &= \int_{r=0}^{\infty} \int_{\varphi=0}^{2\pi} r f(r, \varphi) \delta(r \cos(\varphi - \theta) - s) dx dy
 \end{aligned}$$

משמעות ביטוי זה היא שבאינטגרציה על-פni הפונקציה $f(r, \varphi)$ ליצירת ערך התמורה $f(s, \theta)$ נאספות כל הנקודות $\{(r, \varphi) \mid r \cos(\varphi - \theta) = s\}$. זאת בinguוד לישר בקואורדינטות (x, y) .

נדון כעת במספר תכונות בסיסיות של התמורה רадון $\{g(s, \theta) = \Re\{f(x, y)\}\}$: מההגדרה ניכר כי התמורה רадון היא התמורה ליניארית, דהיינו, עבור שתי תמונות $f_1(x, y)$ ו- $f_2(x, y)$ מתקבל כי $\Re\{\alpha f_1(x, y) + \beta f_2(x, y)\} = \alpha \Re\{f_1(x, y)\} + \beta \Re\{f_2(x, y)\}$. כמו כן ברור כי אם התמונה $f(x, y)$ בעלת תמן סופי, דהיינו, $\forall (x, y) \mid x^2 + y^2 \geq D$, $f(x, y) \equiv 0$, אז גם התמורה רадון בעלת תמן סופי $-D \leq s \leq D$. התמורה רадון מחזוריות בזווית $-g(s, \theta) = g(-s, \theta + \pi)$. סימטריות זו היא הסיבה להגבלת ההגדרה לאינטראול $[0, \pi]$.

כאשר מזוזת התמונה (x_0, y_0) בשיירור $f(x, y)$ נקבל התמורה חדשה המזוזת רק בציר s . קל לראות:

$$\begin{aligned}
 \Re\{f(x - x_0, y - y_0)\} &= \iint_{x, y} f(x - x_0, y - y_0) \delta(x \cos \theta + y \sin \theta - s) dx dy = \\
 &= \iint_{x, y} f(x, y) \delta(x \cos \theta + y \sin \theta + x_0 \cos \theta + y_0 \sin \theta - s) dx dy = \\
 &= g(s - x_0 \cos \theta - y_0 \sin \theta, \theta).
 \end{aligned}$$

באופן דומה ניתן להראות תכונת הגדלה $\Re\{f(ax, ay)\} = g(as, \theta) / |a|$. תכונה אחרתה שנזכיר היא שימור הממוצע אשר קלה להוכיחו ישירות מההגדרה –

$$\iint_{x, y} f(x, y) dx dy = \int_s g(s, \theta) ds$$

תכונה זו טריויאלית אם נחשב על משמעותה הגיאומטרית. בבחירה זווית הטלת שרירותית, אם נתיחס למקרה בצייר 10, כל ערך בהטלת הוא אינטגרציה לאורך ישר כלשהו (המוכתב ע"י s). ע"י אינטגרציה לאורכה של כל הטלת, אנו אוספים את כל הישרים המקבילים ובכך מקבלים את סך נפח הפונקציה $f(x, y)$.

כיוון שאנו חושבים במונחים של התמורה, שאלת מיידית שנשאלת היא האם התמורה המוגדרת הפיכה, דהיינו, האם בהינתן $g(s, \theta)$ בתחום $\pi < \theta < \infty$ ו- $-\infty < s < \infty$ ניתן לחזור לתמונה $f(x, y)$. מיד נראה כי היפוך התמורה אפשרי, וכייך לעשותו.

10.3 משפט החתך המרכזי והיפוך התמרת רадון

דבר מעניין ובעל חשיבות קורה כאשר מתבצעת התמרת פוריה על וקטורי הטללה. בהינתן הפונקציה הדו-מימדית $g(s, \theta)$ נבצע התמרת פוריה ביחס למשתנה s ונקבל פונקציה חדשה

$$\mathcal{J}_1\{g(s, \theta)\} = \mathcal{J}_1\{\mathfrak{F}\{f(x, y)\}\} = G(u, \theta)$$

נפתח פעולה זו ונקבל

$$\begin{aligned} G(u, \theta) &= \mathcal{J}_1\{g(s, \theta)\} = \int_s g(s, \theta) \exp\{-j2\pi s u\} ds = \\ &= \mathcal{J}_1\{\mathfrak{F}\{f(x, y)\}\} = \iint_{x,y} f(x, y) \delta(x \cos \theta + y \sin \theta - s) \exp\{-j2\pi s u\} ds dx dy = \\ &= \iint_{x,y} f(x, y) \left[\int_s \delta(x \cos \theta + y \sin \theta - s) \exp\{-j2\pi s u\} ds \right] dx dy = \\ &= \iint_{x,y} f(x, y) \exp\{-j2\pi(u x \cos \theta + u y \sin \theta)\} dx dy \end{aligned}$$

ביטוי זה נראה סתום. אך נזכיר כי התמרת פוריה דו-מימדית על הפונקציה $f(x, y)$ נתונה ע"י

$$F(\omega_x, \omega_y) = \iint_{x,y} f(x, y) \exp\{-j2\pi(\omega_x x + \omega_y y)\} dx dy$$

אם נתייחס לנקודות התדר בקואורדינטות פולריות $\omega_x = u \cos \theta, \omega_y = u \sin \theta$ נקבל כי התמרת הפוריה

העוברת מ- $f(x, y)$ ל- $F(u \cos \theta, u \sin \theta) = \tilde{F}(u, \theta)$ ניתנת לרשום כ:

$$\tilde{F}(u, \theta) = \iint_{x,y} f(x, y) \exp\{-j2\pi(u x \cos \theta + u y \sin \theta)\} dx dy$$

וזהו בדיק הפונקציה $G(u, \theta)$ שקיבלו בעקבות התמרת פוריה חד-מימדית על התמרת רадון. תוצאה זו ידועה בשם "משפט החתך המרכזי" (Projection Slice Theorem). פירוש הדבר הוא שההינתן התמרת רадון של $f(x, y)$ אותה סימנו כ- (s, θ, g) , ביצוע ההתרמה הפוכה ייעשה ע"י השלבים הבאים המתוארים גם בסכימת הבלוקים בצייר 10.2 :

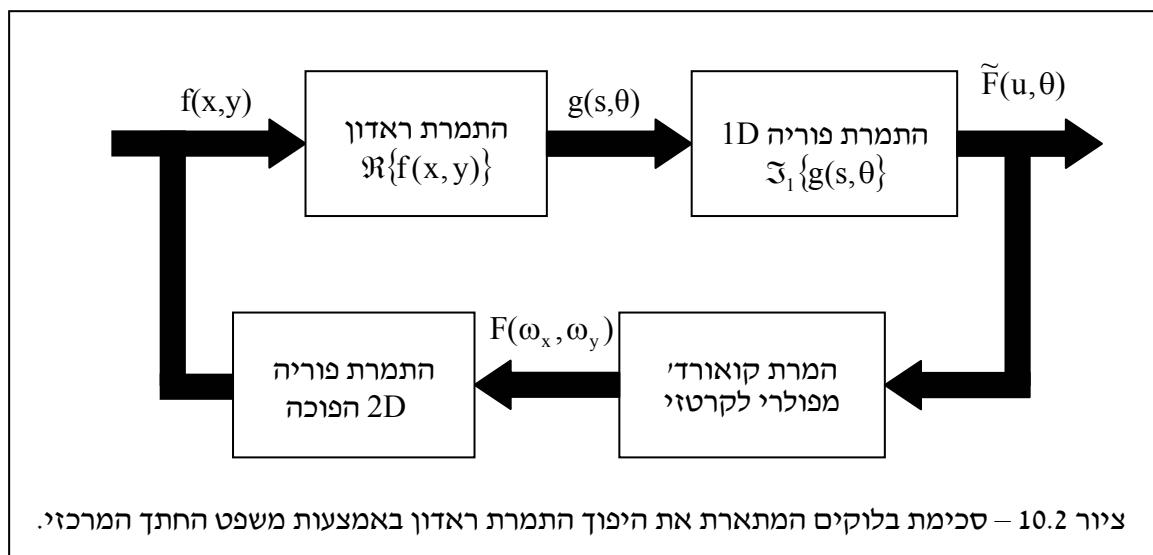
1. בצעי התמרת פוריה חד-מימדית על הפונקציה - $g(s, \theta)$ ביחס למשתנה s . התוצאה היא הפונקציה הדו-מימדית החדשה $G(u, \theta)$. אנו ראיינו כי זהה גם התמרת הפוריה הפולרית

$$\tilde{F}(u, \theta)$$

2. המריי קואורדינטות $\{u, \theta\}$ המהוות קואורדינטות פולריות לקואורדינטות קרטזיות מהצורה $\omega_x = u \cos \theta, \omega_y = u \sin \theta$. ברצף פעולה זו היא פעולה באנאליטית – רישום שונה ולא יותר

(כשנעביר לתיאור דיסקרטי נגלה שהדברים מסתבכים). כך תתקבל הפונקציה $F(\omega_x, \omega_y)$

3. בצעי התמרת פוריה הפוכה דו-מימדית על $F(\omega_x, \omega_y)$ לקבלת $f(x, y)$.



למעשה, דרך משפט החתך המרכזי הצגנו דרך אפשרית נוספת לביצוע היפוך לתמרת רадון. כל צעדי התהליך מוגדרים היטב לפיכך התמרת הפוכה נראה כתהליך יציב ואמין המוביל לקשר חד-חד-ערכי בין תמונה והתמertaה.

ישנן גם דרכי אחרות לביצוע התמרת הפוכה – נציג שתיים כאלה – גישה ישירה וגישה ה הטלות לאחר. נתחיל עם הגישה הישירה אותה נציג ללא הוכחה. בהינתן $g(s, \theta)$ – התמרת רадון של הפונקציה קיימם הקשר הבא:

$$f(x, y) = \frac{1}{2\pi^2} \int_0^\infty \int_{-\infty}^\infty \frac{\frac{\partial}{\partial s} g(s, \theta)}{x \cos \theta + y \sin \theta - s} ds d\theta$$

וקשר זה אינו אלא התמרת הפוכה בדרך ישירה. מהביטוי הנ"ל ניכר כי הפונקציה $g(s, \theta)$ עוברת שלב ראשון בו היא נוצרת לפי s – פועלה הניתנת להתייחסות כמספר מגביר גובהים בציר אחד. לאחר מכן מתבצעת פועלה המוכרת בשם "התמרת הילברט". אנו לא ניכנס למשמעות פועלות אלו מושם שחשיבותן תיאורטית גרידא.

קיימות דרך אחרת לבניית התמרת רадון הפוכה. ברור כי דרך זו קשורה בעבותות לשתי הגישות הקודמות. עם זאת, חשיבותה בזווית ראייה אחרת שהיא מספקת לעניין היפוך התמרת – זווית שתוביל לאלגוריתמים מעשיים בשחזור תמונה מהיטליה. נתחיל בהגדרה של אופרטור מעניין הקורוי "אופרטור ההטלה לאחר". אופרטור זה המוגדר ע"י

$$b(x, y) = \int_0^\pi g(x \cos \theta + y \sin \theta, \theta) d\theta$$

מתחילה בפונקציה ה הטלות $g(s, \theta)$ ובפעולת אינטגרציה על-פניה מקבל פונקציה במקום (y, x) . אם נתיחס לנקודת יעד (y, x) קבועה, הרי שהאינטגרציה נעה על-פנוי (s, θ) לאורץ קו קבוע ב- θ וסינוסואידית ב- s . מה כה מיוחד באופרטור זה? ובכן – מספר דברים אותם נציין ללא הוכחה.

אופרטור ההטלה לאחרו הוא למשה ה- Adjoint של התמרת ראנון. במקרים אחרים ופשוטות, לו היינו חושבים על התמורה דיסקרטית ונוטנים להטמרת ראנון יציג כמטריצה R המכפילה את התמונה f (בסיידור לקסיקוגרפי), הרי שהאופרטור שהגדנו כאן אינו אלא פועלת R^T (של המטריצה). שנגייע לדיוון בדרכים אלגבריות לשחזר מהיטלים נראה כי אופרטור זה חיוני בהיפוך התמורה.

מסתבר כי קיים הקשר הבא המצביע כי תהליך ההטלה לאחרו עשוה כיברת דרך קראת היפוך התמורה. קשר זה אומר שאם נתחיל בפונקציה (x,y,f) , נבצע התמרת ראנון לקבלת (θ, ϕ, g) , נפעיל את אופרטור ההטלה לאחרו לקבלת (x,y,b) , נקבל כי :

$$b(x,y) = f(x,y) \otimes \frac{1}{\sqrt{x^2 + y^2}}$$

כלומר, בפעולה נוספת של דה-קונבולוציה ע"י מסנן שייצגו התידי הינו כגור רדילי נגיע חוזרת לתמונה המקורי. וכך אנו רואים דרך שלישית לביצוע התמורה ההפוכה.

אנו נעצור את הדיוון כאן ונעבור לגישות מעשיות לשחזר מהיטלים. פירוש הדבר שאנו מפסיקים לחושב במונחי הרץ שליוו אותנו עד כה, ונתחיל לחושב במונחים דיסקרטיים - ידוע לנו סט היטלים **באוסף הזוויות ובערכי הסטת הישר**

$$\theta_n = \pi / N, n = 0, 1, \dots, N-1$$

$$s_m = mS / M, -M/2 \leq m < M/2$$

יש אם כך N היטלים הפרוסות ברוחוי זווית קבועים, וכל היטלה כוללת M דגימות בתחום $[S, +S]$.

גודל צעד הדגימה S/M יוכתב ישירות מהתדר הרדילי המירבי אותו רצוננו לשחזר בתמונה $[f, j]$. בהינתן תמונה דו-מימדית בעלת התמרת פורייה המתאפשרת מחוץ לדיסקה ברדיוס $\omega_0/2$, כל היטלה שלה אף היא חסומת תדר עם אותו סף. לכן, כל היטלה צריכה להידגים במרווה שהיא קטן מ- $\omega_0/2$. בדרך זו נוכל לקבוע את M .

כמויות היטלים N קשה יותר לקביעה ובחירה צריכה להיות קשורה ל- $\omega_0/2$ אף היא, אם כי בדרך עקיפה. במונחים גסים של חשבונות המדע, אם יש לנו MN ערכי מדידה, סביר כי התמונה המשוחזרת $[f, j]$ תכיל כמות דומה או נמוכה של דגימות, שאם לא כן – בעיית השחזר אינה מוגדרת היטב.

בידיעה שהתמונה חסומת תדר לרדיוס $\omega_0/2$, פירוש הדבר שמדובר הדגימה הkartizi צריך להיות לכל היוטר $D = 1/2\omega_0$. אנו גם מניחים שהתמונה f בעלת תמק מרוחבי בצורת דיסקה שרדיאוסה S כך הרי קבענו שהגבנו את S לאינטראול $[S, +S]$. לכן, בניתו גס, ביריבע בגודל $2S$ על S ייכנסו $(2S/2\omega_0)^2$ דגימות. כל היטלה מכילה לפחות $(2S/2\omega_0)^2$ דגימות, ולכן תידרשנה $(2S/2\omega_0)^2$ היטלים

לפחות על מנת לספק את הצורך לאוזן בין המידע הנוכחי והמידע המבוקש. בפועל, רצוי לתת יותר מידע מנעלמים לשם ייצוג השחזור.

עניין אחר עליו יש לתת את הדעת בעבר לאותו דגומים במקומות ובמרחב רадון הוא עניין דרך היוצרן של ההטלות מהמידע הדגום. דרך אחת לחשוב על הקשר היא במונחי אותן חסומי תדר שנדגמו. אפשרות אחרת היא לחשב על התמונה המשוחזרת כתמונה הבנויה מפיקסלים בעלי תמק ריבועי אשר על-פניהם ערך קבוע של (x,y) .

10.4 שחזור מהיטלים ע"י גישות תדר

אחת הדרכים הפופולריות לשחזור תמונה מהטלות היא גישת התדר המתבססת על משפט החתק המרכזי. משפט זה טוען כי ביצוע התמרת פורייה חד-מימדית על כל הטלה מביא אותנו לפונקציה חדשה אשר ניתנת להתייחסות כתמונה הרצואה בתדר כשהיא דוגמה בקואורדינטות פולריות. בעת נתאר תהליך מעשי שנועד לשחזור תמונה בגישה זו, תוכן התיחסות לעובדה שככל הגדים המעורבים דוגמים.

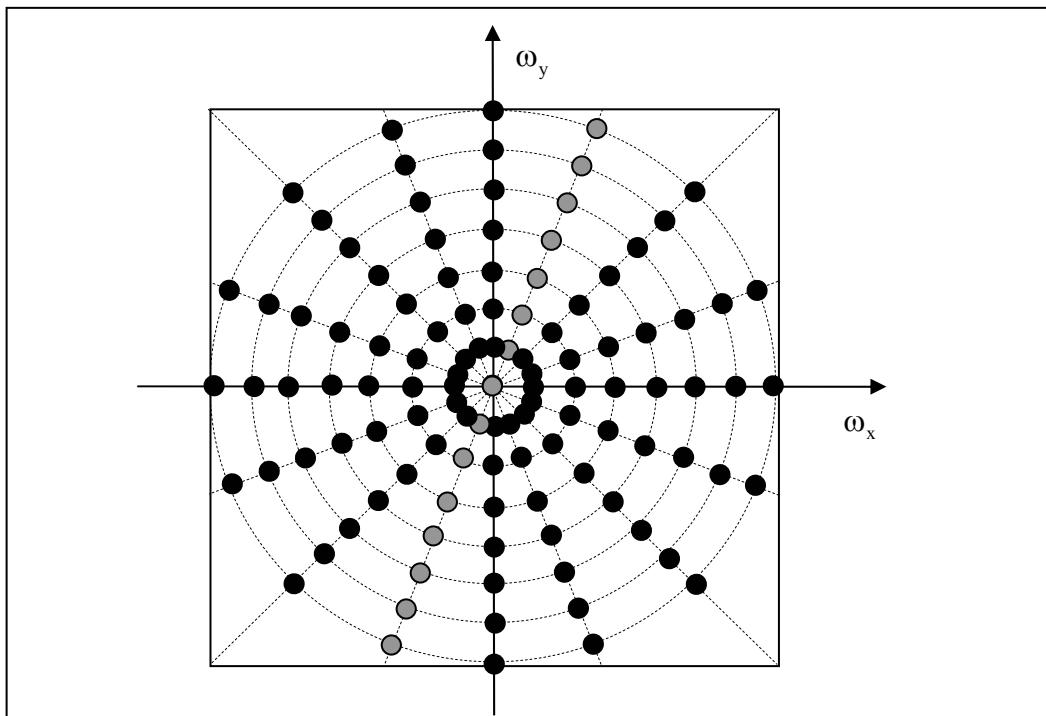
נקודת המוצא היא הפונקציה $[g_{m,n}]$ המוגדרת בתמך $N < n \leq M/2, 0 \leq m < M/2$. כל שורה (m קבוע) – $[g_{:,m}]$ במנוחה Matlab) של מערך זה מייצגת זווית קבועה, וכל עמודה (n קבוע) – $[g_{m,:}]$ במנוחה Matlab) מייצגת הטלה אחת.

בצעד ראשון ניקח כל הטלה ונבצע עליה 1D-FFT. למרות שככל הטלה כזו מכילה M ערכים, אין סיבה להגביל את ה-FFT ל- M ערכי מוצא. ע"י ריפוד באפסים של וקטור הכניסה ניתן לייצר סיידרת מוצאת באורך גדול יותר. אלגוריתמית, ניקח את המערך הדו-מימי בגודל M שורות על N עמודות, נסיף לכל עמודה N אפסים לקבלת מערך חדש בגודל $(N+1) \times M$, אז לכל עמודה נבצע FFT חד-מימי. המוצא יהיה מערך בגודל זהה $(N+1) \times M$, בו כל עמודה מהווה חתק מרכזי בתדר, כמפורט בציור 10.3. התיכים אלו פורים במרוחבי זווית קבועים ונוטנים בפועל דוגימה פולרית של מרחב התדר.

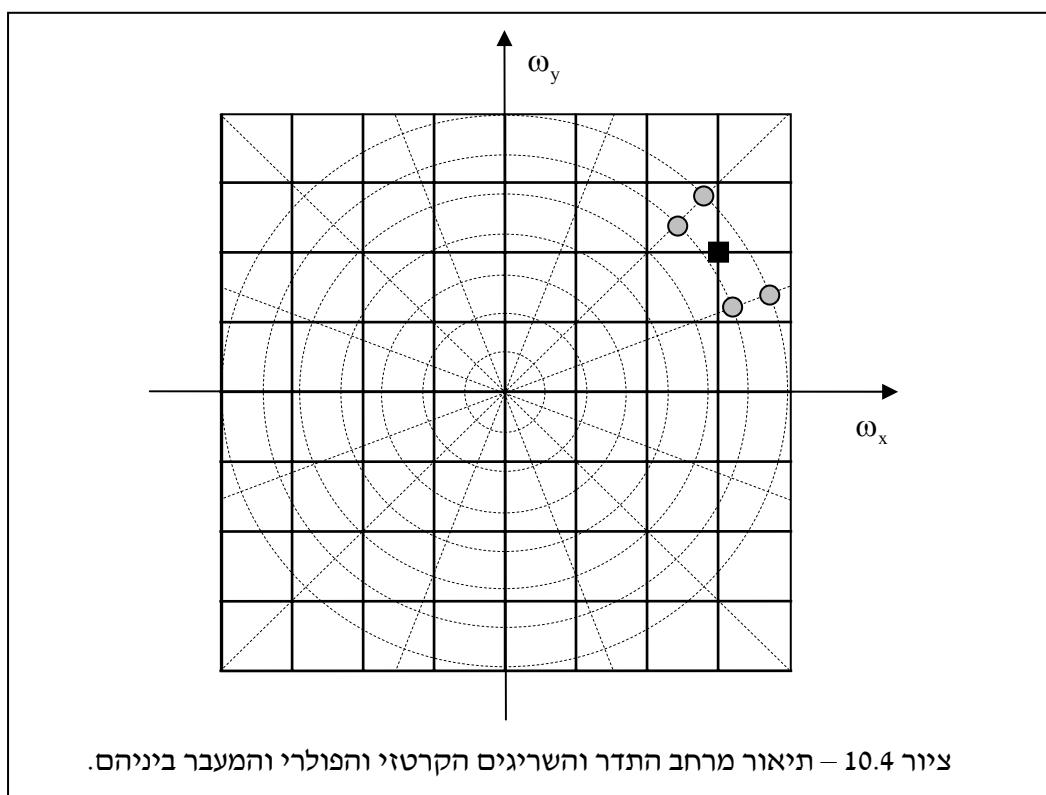
כל שעליינו לעשות עכשו הוא התמרת פורייה הפוכה מהדוגימות הפולריות כמפורט בציור 10.3 אל התמונה במקומות $[f_{j,k}]$ כשהיא בקואורדינטות פולריות. לروع המזל, התמרה כזו לא ידועה לנו, ואם נבחר לעבוד לפי הגדרת התמרת פורייה נאלץ לבצע תהליך נומרי עתיק חישובים. האלטרנטיבה היא המרת אוסף הדוגימות הפולריות באוסף אחר קרטזוי. עבור אוסף כזה, ביצוע 2D-FFT ייתן בצורה אמינה ומהירה את התוצאה הרצואה. המרת הקואורדינטות תיעשה ע"י אינטראפולציה – כל נקודת יעד בשרגיג הקרטזוי תחושב כממוצע משוקלל של שכנותיה מהשריג הפולרי עבורן יש לנו ערכים. ציור 10.4 ממחיש המקרה זו. זו גם הסיבה לריפוד באפסים שביצענו בצעד הראשון – ריפוד זה יוצר לכל קרן החולפת דרך הראשית יותר נקודות דוגימה, ולכך האינטראפולציה שתיעשה תהיה מדויקת יותר. ברור כי בכך יש לדאוג ליותר קרניים כאלו – זה קשר ישיר למספר ההטלות הנוכחיות.

פרק 10 - שחזור תמונה מהיטלים

לאחר ביצוע האינטראפולציה וקיבלת N נקודות תדר בשציג קרטזי, התמרת 2D-FFT הפוכה תוביל לתמונה הרצiosa. בדרך כלל, עד האינטראפולציה מהוות אבן גוף בתהליך שתואר וביצועיו יჩידים.



ציור 10.3 – תיאור מרחב התדר והדgesיות הנთונות לאחר התמרת FFT לכל הטלה. קו הנקודות האפור מתאר אוסף דgesיות הנובע מהטליה אחת



ציור 10.4 – תיאור מרחב התדר והשריגים הקרטזי והפולרי והמעבר ביניהם.

10.5 שחזור מהיטלים ע"י קובולוציה

שחזור מהיטלים בגישה התדר הישירה נובעת ממשפט החतך המרכזי אשר הוכיח להיות אחת הדריכים לביצוע התמרת ראנון הפוכה. באופן דומה, נוכל להציג אלגוריתם לשחזור מהיטלים המבוסס על אופרטור ההטלה לאחרו.

שוב נקודת המוצא היא הפונקציה $[m,n] \rightarrow g$ המוגדרת בתמך N . כל שורה (m קבוע) של מערך זה מייצגת זווית קבועה, וכל עמודה (n קבוע) מייצגת הטלה אחת שלמה. לפי התכונה המרכזית של אופרטור ההטלה לאחרו, אם ניקח את סט ההטלות $[n,m] \rightarrow g$ וنبצע אינטגרציה מהצורה (דיסקרטיזציה פשוטה של נוסחת האופרטור)

$$b(x,y) = \sum_{n=0}^{N-1} g \left[x \cos\left(\frac{\pi n}{N}\right) + y \sin\left(\frac{\pi n}{N}\right), n \right]$$

תתקבל תמונה הקשורה לתמונה הרצiosa עד-כדי פעולה ליניארית וקבועה במקום. נוכל לדגום את מישור המקום ע"י $x=kd$, $y=jd$ מרוח הדגימה ולכתוב

$$\cdot b(kd, jd) = b[k, j] = \sum_{n=0}^{N-1} g \left[kd \cos\left(\frac{\pi n}{N}\right) + jd \sin\left(\frac{\pi n}{N}\right), n \right]$$

ברור כי לצורך לעגל את הגודל שמחליף את האינדקס m בפונקציה g בסכימה על מנת לקבל שימוש בפונקציה הדיסקרטית ממנה התחנו.

הצעד השני והאחרון שנדרש לשם השלמת השחזור הוא ביצוע הסינון. הקשר עליו אנו משתמשים

הוא :

$$\begin{aligned} b(x,y) &= f(x,y) \otimes \frac{1}{\sqrt{x^2 + y^2}} \\ \Rightarrow B(\omega_x, \omega_y) &= F(\omega_x, \omega_y) \cdot \tilde{\mathcal{Z}} \left\{ \frac{1}{\sqrt{x^2 + y^2}} \right\} \\ \Rightarrow F(\omega_x, \omega_y) &= B(\omega_x, \omega_y) \cdot \tilde{\mathcal{Z}}^{-1} \left\{ \frac{1}{\sqrt{x^2 + y^2}} \right\} \end{aligned}$$

המשן הנדרש לשם קבלת F הוא גוזר רדילי אידאלי, ומסנן זה אינו מעשי כיון שתהיה לו נטייה להגבר רעשים. לכן, בספרות מציעים מגוון מסנני שחזור שנראים כגוזר סביב הראשית, ודועכים כשמתרחכים מהראשית על מנת לדכא הגבר רעש. רעיון הירירטי זה לאزر לנו ואנו ראיינו כמוותו שדיםרו על שחזור תМОנות מטשטוש ורעש, שם ראיינו כי רגולרייזציה המדעית את המשן המשוחרר רצואה לשם קבלת תוצאות טובות. עם זאת, גם ראיינו כי יש אלטרנטיבה בעלת רקע תיאורתי נאות לשם תכנון שחזור זה המבוסס על שערוך סטטיסטי. אנו ניגש למשפחה שיטות אלו בסעיף הבא.

חשוב להבהיר כי בדומה לגישות התדר, גם גישה זו סובלת מצרות לרוב בשל הדגימות השונות, קירובי האינטגרלים לסכימות סופיות, ולבסוף, ביצוע סינון הפוך לא יציב.

6.10 שחזור מהיטלים ע"י גישות אלגבריות

אם נחזור להתחלה, נוכל להציג מודל הקשור את ההטלות הנתונות לתמונה הרצוייה כscalar הגדלים המעורבים מיוצגים כדגומים. מודל זה יתייחס לתמונה היעד $[k,j] f$ כוקטור עמודה בסידור לקסיקוגרפי f . מטריצה R_0 תציג את תהליכי ההטלה של f לקבלת וקטור ההטלה \underline{g}_0 , ונקבל סידרת משוואות מהצורה

$$\left\{ \underline{g}_{\theta_n} = R_{\theta_n} f + \underline{v}_n \right\}_{n=0}^{N-1}$$

לכל זווית הטלה יש משווה כזו, המטריצה המתארת את ההטלה מוגדרת היטב, ועלינו להתייחס באופן נאות לרוש אדיטיבי שמוספי בכל הטלה – אנו נניח כי רוש זה גאוסי, בעל ממוצע אפס, לבן, וחסר קורלציה בין ההטלות השונות. עפ"י כלים שכבר ראיינו בפרק 6, יישום ה- Maximum Likelihood לבעיית האופטימיזציה

$$\underset{f}{\text{Min.}} \quad \epsilon\{f\} = \sum_{n=0}^{N-1} \left\| \underline{g}_{\theta_n} - R_{\theta_n} f \right\|_2^2$$

אשר פתרונה נתון ע"י

$$f_{\text{opt}} = \left[\sum_{n=0}^{N-1} R_{\theta_n}^T R_{\theta_n} \right]^{-1} \left[\sum_{n=0}^{N-1} R_{\theta_n}^T \underline{g}_{\theta_n} \right]$$

ונכל גם לרכז את כל המשוואות למערכת אחת ע"י:

$$\underline{g} = \begin{bmatrix} \underline{g}_{\theta_0} \\ \underline{g}_{\theta_1} \\ \vdots \\ \underline{g}_{\theta_{N-1}} \end{bmatrix} = \begin{bmatrix} R_{\theta_0} \\ R_{\theta_1} \\ \vdots \\ R_{\theta_{N-1}} \end{bmatrix} f + \begin{bmatrix} \underline{v}_0 \\ \underline{v}_1 \\ \vdots \\ \underline{v}_{N-1} \end{bmatrix} = Rf + \underline{v}$$

ואז הפתרון הניל אין אלא

$$f_{\text{opt}} = (R^T R)^{-1} R^T \underline{g}$$

עיוון קצר יגלה כי הביטוי $R^T \underline{g}$ אינו אלא הפעלת אופרטור ההטלה לאחרו אותו הצגנו קודם. ההכפלה בהיפוך המטריצה $R^T R$ נותנת את הקונבולוציה לה נדרשו בשיטה זו. המסקנה הנגזרת היא שיטת השחזור מבוססת קונבולוציה אינה אלא שחזור מבוסס ML. כczzo, אנו מכירים את חסרוןותיה הטוב ויודעים כי זו אינה שיטה חזקה דיה להתמודד עם רוש ועם אופרטור קלקלול בעל Condition-Number גרווע (כלומר, מטריצת קלקלול קרובה לסינגולרית). ראיינו כי הפתרון הוא רגולרייזציה, ובפרט כזו מבוססת MAP בה אנו מניחים מבנה לתמונה הנעלמת. לשם דוגמה, עברו רגולרייזציה פשוטית מהסוג

$$\underset{f}{\text{Min.}} \quad \epsilon\{f\} = \sum_{n=0}^{N-1} \left\| \underline{g}_{\theta_n} - R_{\theta_n} f \right\|_2^2 + \lambda \|Df\|_2^2$$

$$\underline{f}_{\text{opt}} = \left[\sum_{n=0}^{N-1} R_{\theta_n}^T R_{\theta_n} + \lambda D^T D \right]^{-1} \left[\sum_{n=0}^{N-1} R_{\theta_n}^T \underline{g}_{\theta_n} \right] = [R^T R + \lambda D^T D]^{-1} R^T \underline{g}$$

אם קודם סינו ע"י מסנן גוזר תיאורתי המיצג ע"י ההיפוך של $R^T R$, הרי שהמסנן שמוסע כאן אינו אלא גירסה מרווחנת שלו כפי שגישות היחסיות בספרות בחרו להפעיל. ברור כי ניתן להציג דרכים מוצלחות יותר כרוח הדברים בהם נגענו בפרק 6, כגון שימוש בפונקציות רובסיטיות ועוד.

מרכיב חדש ואשר מסבך את העניינים בשחזרים מבוססי ייצוג מטריצי כגון אלו הוא בעית אי-دلילותן של המטריצות המעורבות. בשחזר תמונות מטשטוש, מטריצת הקלוקול דليلת מאוד וכך כמובן גם השחלוף שלה. כאן לעומת זאת כל שורה במטריצה R מכילה N איברים (כל אינטגרציה על ישר חוצה את התמונה ווסףת N ערכי פיקסלים) ששוכמו. יתרה מכך, מבנה המטריצה אינו פשוט במו במקורה של טשטוש מקומי – מיקום האיברים השונים מאפס דורש עבודה חישוב מורכבת יותר. אלו נושאים בעלי השפעה ישירה על מרכיבות הפתרון הנומרי לשחזר. מיסיבות אלו מתחשי אלגוריתמים בתחום זה קיצורי דרך והאוצאות של שחזר איטרטיבי \underline{f} . אנו נראה כאן שתי גישות אלו – אחת פשוטה להבנה אך איטית להחריד, ואחרת בה ישנה האצה משמעותית אך אי-דיאוקים שנובעים מהאצה זו. נקדמים ונאמר כי יש שיטות רבות אחרות.

ראינו כי בסיס הדברים הפעלת ML יחד עם רגולרייזציה יוצרים את פונקציית המחיר הבאה בבעית האופטימיזציה אותה עליינו לפתור:

$$\epsilon\{\underline{f}\} = \|\underline{g} - R\underline{f}\|_2^2 + \lambda \|D\underline{f}\|_2^2$$

גרדיאנט פונקציה זו נתון ע"י

$$\frac{\partial \epsilon\{\underline{f}\}}{\partial \underline{f}} = 2R^T(R\underline{f} - \underline{g}) + 2\lambda D^T D\underline{f}$$

ולכן איטרציות של אלגוריתם השיפוע המירבי (Steepest Descent) ייעשו ע"י המשוואה האיטרטיבית

$$\underline{f}_{k+1} = \underline{f}_k - \mu (R^T(R\underline{f}_k - \underline{g}) + 2\lambda D^T D\underline{f}_k)$$

ברור כי אנו לא מעוניינים לבנות מטריצות לשם יישום האלגוריתם, ולכן ניתן למשווה זו הסבר במונחים של פעולות על התמונה f – הכפלת D אינה אלא הפעלת מסנן לפלייאן (או גוזר אחר שimdooz) חלקות, הפעלת D^T זהה בשל סימטריות המטריצה. עד כה זה החלק הקל. הפעלה R על התמונה פירושה בניית הטלות מהתמונה – זהה פעולה טובעניות בחישובים אך היא חונית. הפעלה R^T היא פעולה ההטלה לאחרו אותה פגשנו כבר. שתי הפעולות אלו (R ו- R^T) יכתיבו את סיבוכיות האלגוריתם באיטרציה אחת. יש לזכור כי לשם התכנסות קרוב לוודאי שנדקק לעשרות איטרציות ופירוש הדבר אלגוריתם איטי למדי.

אם נקבע על הנקודה הכוabit מכל/algoritm שהוצע – יהיה זה חישוב ההטלה וההטלה לאחרו – אם אין מנוס מחשבים, רצוי שנידרש לכך פעמיים ספרות בלבד. זה הרעיון מאחורי ההאצה הבאה

המודרנת לאנשי עיבוד אותות 알고ריתם ה- LMS. נציג 알고ריתם זה ללא איזכור של נושא הרגולרייזציה לשם פשוטות. פונקציית היעד שלנו ניתנת לתיאור כ:

$$\epsilon\{\underline{f}\} = \left\| \underline{g} - R\underline{f} \right\|_2^2 = \sum_{j=1}^L (g_j - r_j f_j)^2$$

בתיאור זה פירקנו את נורמת השגיאה ב- L המשוואות לסכום של L ביטויים ריבועיים פשוטים, כאשר L הוא מספר המדידות הסקלאריות (מספר ההצלחות כפול מספר הדגימות בכל הטליה). 알고ריתם ה- SD מציין הצגנו מציע את העדכון

$$\underline{f}_{k+1} = \underline{f}_k - \mu R^T (R\underline{f}_k - \underline{g}) = \underline{f}_k - \mu \sum_{j=1}^L r_j^T (r_j \underline{f}_k - g_j)$$

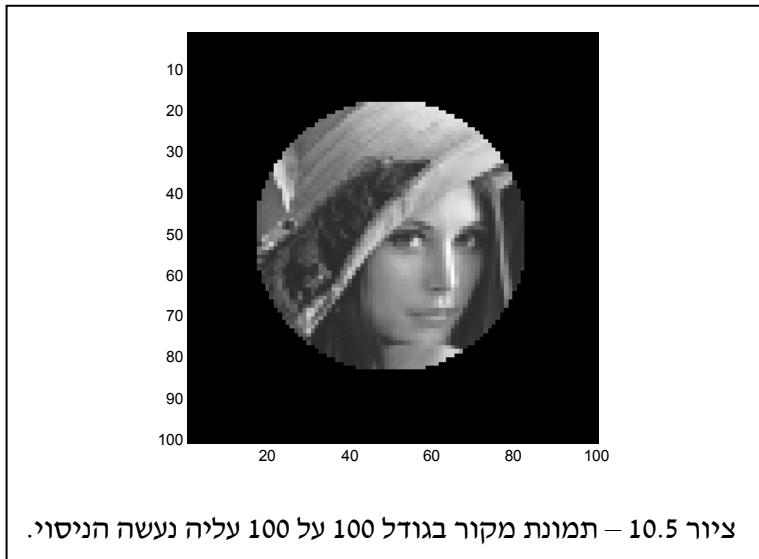
אנו רואים כי העדכון ניתנו אף הוא להצגה כפירוק של L עדכנים. באlgorigitms הניל, L עדכנים מחושבים עפ"י תוכאת האיטרציה הקודמת, ורק אז נעשה עדכון מסיבי לפתרון. 알고ריתם ה- LMS מציין כתחליף לעדכן את f אחרי כל אחד מאיברי הגרדיאנט. לכן האלגוריתם יהיה הבא

$$\begin{cases} \underline{f}_1^k = \underline{f}_L^{k-1} - \mu r_1^T (r_1 \underline{f}_L^{k-1} - g_1) \\ \underline{f}_2^k = \underline{f}_1^k - \mu r_2^T (r_2 \underline{f}_1^k - g_2) \\ \vdots \\ \underline{f}_{j+1}^k = \underline{f}_j^k - \mu r_{j+1}^T (r_{j+1} \underline{f}_j^k - g_{j+1}) \\ \vdots \\ \underline{f}_L^k = \underline{f}_{L-1}^k - \mu r_L^T (r_L \underline{f}_{L-1}^k - g_L) \end{cases}$$

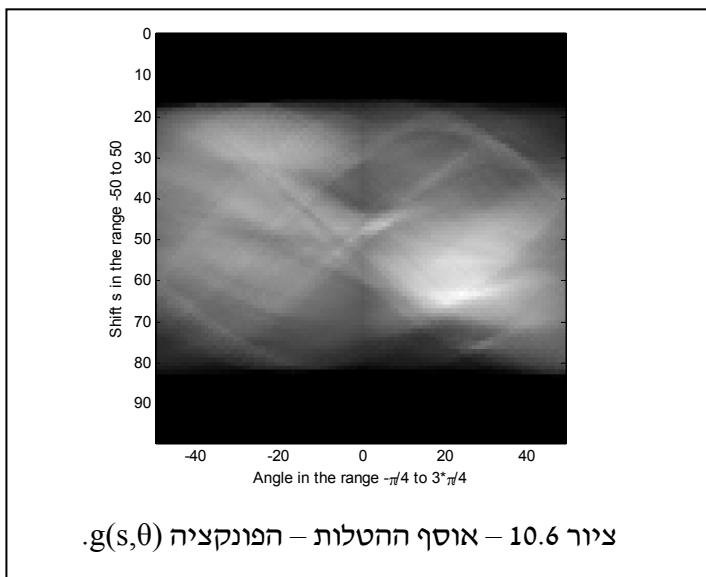
באופן זה, כל מדידה סקלארית שמנעה יוצרת עדכון לפתרון – דבר זהו למדי במצבים בהם חישוב יכול להיעשות במהלך רכישת מדידות. האלגוריתם הניל מציין L עדכנים שחזור במהלך של כמה חישובים ששකולה לאיטרציה בודדה של ה- SD. עם זאת, ניתן להמשיך ולעדכן ע"י סיבוב נוסף של עדכנים על המדידות.

מסתבר כי אלגוריתם זה בעל תכונות התכנסות טובות למדי – אנו נימנע מלaiceness לפני התיאורטי של ניתוח זה ונציג כי בנסיבות חישובים השקולה לאיטרציה אחת של ה- SD (דהיינו – פעמי אחדות חישוב הצלות ופעם שימוש באופרטור ההטלה לאחר מכן) מגיע לאלגוריתם ה- LMS לתמונה בעלת תוכן רב יותר ומשמעותי. יתרון משמעותי של אלגוריתם ה- LMS עליו לא דיברנו הוא יכולת העקבה שלו – במקרה בו האובייקט המצלום נע תוך כדי רכישת הצלות, ניתן כי נרצה לשחזר סרט ולא תמונה בודדה. אלגוריתם ה- LMS מייצר סידרת תמונות בהן להצלות האחרונות יותר משקל בקביעת התמונה וכן נקבל סידרת תמונות נאמנה יותר למתרחש.

לשומות תיאור נושא השחזור מהיטלים, נציג כאן תוכאות ניסוי קצר ומלאכוטי במקצת. אנו מתחילה בתמונה מקור – תמונה בגודל של 100 על 100 המתוארת בציור 10.5.



מתמונה זו אנו יוצרים סידרת היטלים הכללת 100 היטלות בפיזור זווית אחד באינטראול $(\pi/4, \pi/4)$, כשבכל היטלה יש 100 דגימות המתייחסות להסתה s בתחום $-50 \leq s \leq 49$. לצורך יצירת ההיטלות מהמידע הדיסקרטית נעשית הפרדה בטיפול בין ישרים בזווית אופקיות $[-\pi/4, \pi/4]$ וישרים בזווית אנכיות $[\pi/4, 3\pi/4]$. אם נתיחס לישרים בזווית אופקית, עבור היטלה בזווית מסויימת θ והסתה s , וערבי הקואורדינטה x בתחום $[-50, 49]$ (אנו מניחים כי מרכזו התמונה הוא הראשית) נקבע את הקואורדינטות המתאימות y להיות אלו המקיים את משוואת הישר $0 = s - x \cos \theta + y \sin \theta$. בהינתן אוסף של הקואורדינטות (y, x) הכללת 100 צמדים, חלקם ייפול מחוץ לתמונה ואלו צריכים להיזרק. הנוטרים מתארים מקום גאותרי בו נופל הישר על שטח התמונה. נעלם את קואורדינטות y וננסכים (שקלול לאינטגרציה) את הפיקסלים בנקודות המתקבלות לשם קבלת $g(s, \theta)$. תהליך זה חוזר על עצמו 100 פעמים נגד כל ערך של s ו-100 פעמים נגד כל ערך של θ . תמונה ההיטלות מתוארת בצייר 10.6.

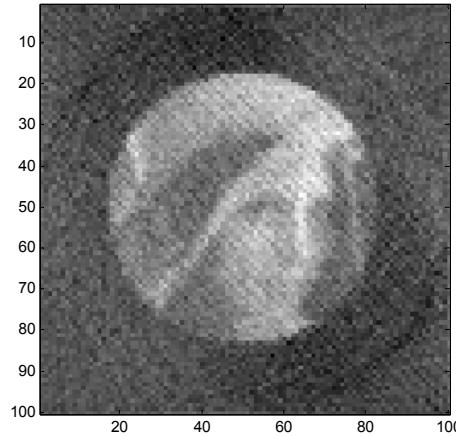
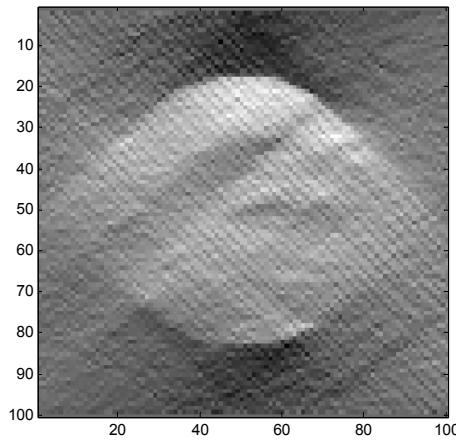


פרק 10 - שחזור תמונה מהיטלים

הטיפול הישרים בעלי זווית אנכית דומה כשהפעם ע נקבע להיות 100 ערכים בתחום [49,50] ו- x נקבע מトーך המשווה $0 = s \cos \theta + y \sin \theta$ ועובר העגלה. כתחליף להעגלה (שפירושה שיטת nearest-Neighbor), ניתן לבצע הטלות מדויקות יותר ע"י אינטראפלציה נאמנה יותר לשיש, אך כפי שנראה, בשל מלאכותיות הניסוי, אין עניין זה יוצר הפרעה משמעותית.

כל ערכי ההטלה זההמו ברעש גaussi ולבן בעל ממוצע אפס ושונות $\sigma = 200$ יש לזכור כי ההטלות צוברות כ- 100 ערכי רמות אפור בתחום [0,255], ולכן עצמת רעש זו מקבילה לרעש גaussi בעל שנות $\sigma = 20$ בכל פיקסל.

ציור 10.7 מתאר את תוצאות האלגוריתם ה- LMS עבור חצי סריקה מלאה וסריקה מלאה של מערכת המשוואות. אנו רואים כי כבר לאחר סריקה אחת, שימושה ביצוע $R^T R$ עם אחת כל אחד, התוצאה כבר "טובה". בסימולציה זו בחרנו שלא לערוב את הרגולרייזציה. لكن התוצאה סובלת מותפעת "צלצולים" האופיינית כל-כך לפתרון ML ישיר.

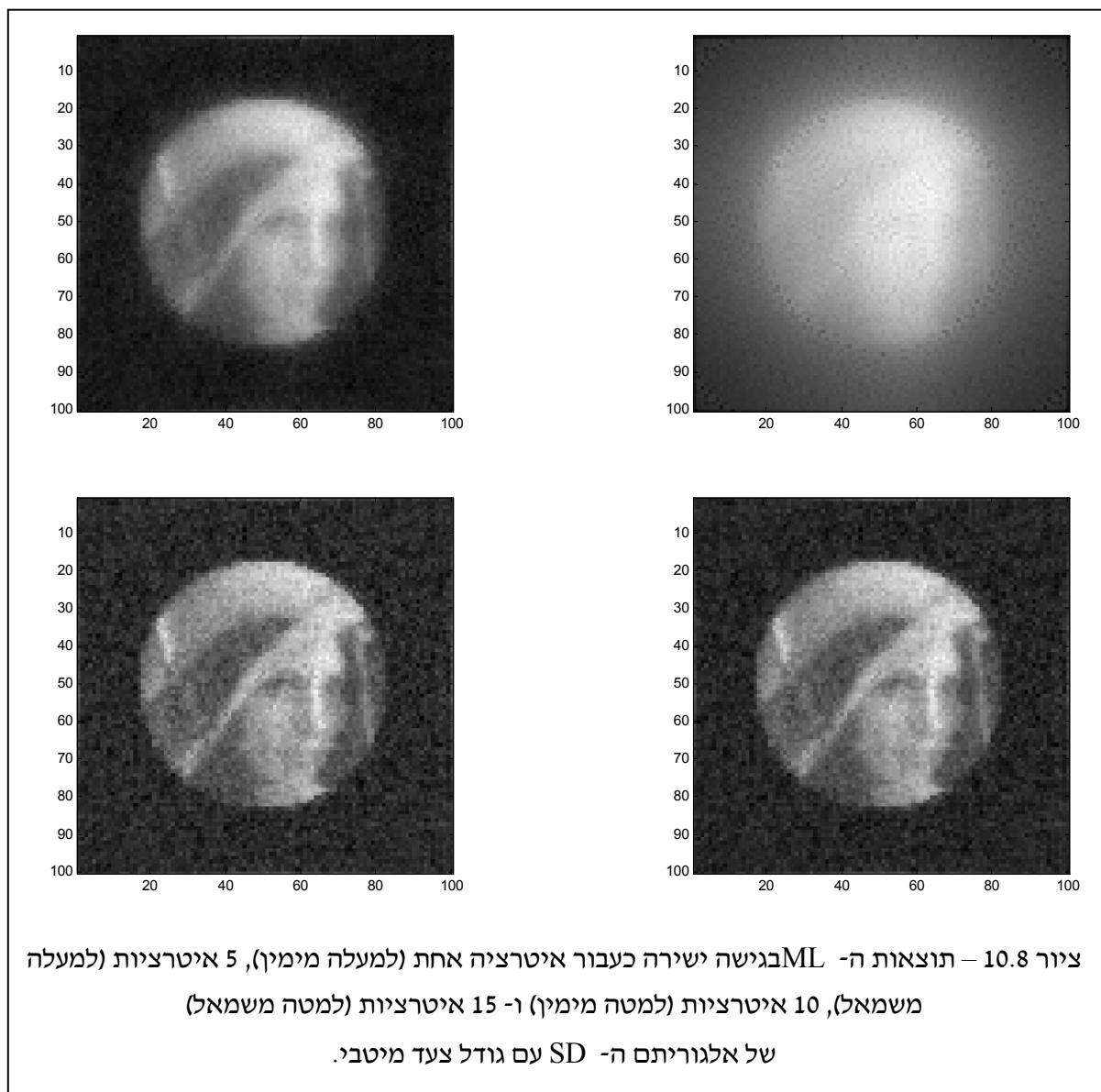


ציור 10.7 – תוצאת ה- LMS לאחר חצי סריקה וסריקה מלאה אחת.

פרק 10 - שחזור תמונה מהיטלים

בביצוע ה- LMS יצרנו את הוקטורים \vec{z} ממש כשם שאלו נבנו בעת הפקת החטולות. לכן, גם אם יצרנו פגיעה בתהליך הדגימה של הישרים, פגעה זו הופיעה גם בעת השחזור ואינה מהוות בעיה. בישומים מעשיים בהם אין זה כך, יכול הפער בין המודל המיציר את החטולות, ומהודל בו משתמשים לשחזר להיות מקור לשגיאות גדולות. לעומת זאת נוכל לטעון כי סטיות בין הייצור האמיתי של החטולות לבין המודל שישמש לשחזר נספג לתוך הרעש האדיטיבי שבמודל.

כיוון שהוקטורים \vec{z} נאספו, וכיוון שמיידי מערכת המשוואות מאפשרים זאת, צברנו גם את המטריצה R במלואה – זו היא מטריצה בגודל 10000 על 10000 שבה כל שורה כוללת עד 100 איברים שהם '1' (שפירושו – אלו נצטברו בהטלה). בהינתן מטריצה זו אנו יכולים להפעיל את הגישה הישירה כפי שתואר קודם. ציור 10.8 מראה תוצאות גישה זו לאחר {1,5,10,15} איטרציות ה- SD עם גודל צעד אופטימלי לקבלת מהירות התכנסות מירבית.



ציור 10.8 – תוצאות ה- LMS בגישה ישירה בעבר איטרציה אחת (למעלה מימין), 5 איטרציות (למעלה משמאל), 10 איטרציות (למטה מימין) ו- 15 איטרציות (למטה משמאל)
של אלגוריתם ה- SD עם גודל צעד מיטבי.

פרק 11 – מבנים פירמידליים ושימושיהם בתמונות

פרק 11 – מבנים פירמידליים ושימושיהם בתמונות
235	11.1 הפירמידה הגאוסית
236	11.2 הפירמידה הפלסיאנית
240	11.3 התמרת Wavelet ע"י ה-
249	Lifting Scheme

מושג ששב וועלہ במהלך הדיוון בתמונות הוא מושג ה-"ריזולוציה". ריזולוציה של תמונה הינה שם נרדף למידת חדותה, אן לגודל הפרט הקטן ביותר הניתן בה לאבחנה. לא תמיד יותר פיקסלים משמשם יותר ריזולוציה. בהקשר זה, הדיוון במשפט הדגימה מציע מدد ברור הקשר בין מרוחק הדגימה ובין הריזולוציה של התמונה, כזו נמדדת במונחים תדריים. החולשה של תפיסה זו בהיותה גלובלית וחד-משמעותית. באשר לגלובליות, משפט הדגימה מציע מרוחק דגימה אחד התקף לכל התמונה – אין אלו ידועים לאפין אותן בעלי ריזולוציה תלויות מקום ומרוחק הדגימה הנאות לייצוגם בכל מקום. באשר לח-משמעותית, מדובר בשניהם כי במקום מסוים בתמונה יש רק ריזולוציה אחת? אם נפעל כך, בהכרח נציג את הריזולוציה הגבוהה ביותר במקומות, ונתעלם מפוגון תופעות הקורנות באזורה נקודת ואשר להן ריזולוציה טבעיות אחרות. שתי חולשות מהוות את התשתיות לצורך במבנה פירמידליים בהם ריזולוציה יכולה להשתנות ממקום למקום, וכל מקום יכול לקבל ייצוג נאות במגוון ריזולוציות.

תפיסה מולטי-ריזולוציונית כבר הוצאה כסדרה בתתרמת Wavelet, אלא שם בחנו את ההתרמה כהתמורה ולא ייצוג החושב "ריזולוציה". בפרק זה נחזר על עיקרי הדברים העומדים מאחוריה התמורה זו ודוגמאות לה, אלא שההדגש יהיה הפעם על היות שיטות אלה שיטות ייצוג מולטי-ריזולוציוניות המKENOT גמישות בעיבודים שונים. אין גישה מאוחדת לתחום זה והדרך בה נבחר להציג את הדברים היא הצגת האפשרויות השונות ודרבן את היישומים הנובעים מהן.

11.1 הפירמידה הגאוסיאנית

ככל, פירמידה הינה אוסף תמונות המאורגנות בסדר כלשהו (נציין סידור אנכי המתחילה למטה ועולה למעלה) ואשר כולל נובעות מהתמונה אחת שעבירה עיבוד. נניח כי נתונה לנו תמונה $(t,m)X$ בגודל N . נגיד תמונה זו כתתיתתה של הפירמידה $P_G(0)$ ונסמן אותה כ- $P_G(0)$. נפעיל על תמונה זו מסנן LPF מסווג כלשהו ונקבל תמונה בה הפרטים מטושטשים קמעה. התמונה החדשה, $P_G(1)$, תהיה השכבה השנייה בפירמידה. נמשיך ונטשטוש את התמונה $P_G(1)$ ונקבל את $P_G(2)$ וכך הלאה. אוסף התמונות שתתקבל יהיה הפירמידה הרצiosa, כאשר כל תמונה מייצגת רמת רזולוציה מסוימת.

נניח כי מסנן ה- LPF הוא כזה שمدכה בצורה נאותה תדרים גבוהים מעבר ל- $2/\pi$ (מחצית טווח התדרים). פירוש הדבר הוא שהתמונה לאחר סינון יכולה לעבור דילול ביחס 2:1 ללא אובדן. לכן, יחד עם פעולות הסינון תבואנה פעולות דילול והtapes תקינה. ציור 11.1 ממחיש את המבנה המתkeletal.



מהו המסנן בו יש לעשות שימוש? ישן מגוון גישות לבחירת המסנן. לדוגמה, נוכל לחפש בחירהה שתבטיח דיכוי תדרים גבוהים Ciوات על-מנת שהדילול יהיה מותר ללא אובדן בקרוב טוב. בעבודתם הציעו Burt ו- Adelson גישה שונה – הם הציעו את הכללים הבאים אוטם מסנן זה יקיים, בהנחה שזו

מסנן ספרטילי באורך $2N+1$ שtagobet ההלם שלו היא $\{w_k\}_k$

$$1. \sum_{k=-N}^N w_k = 1,$$

$$2. \forall k > 0, w_k = w_{-k}$$

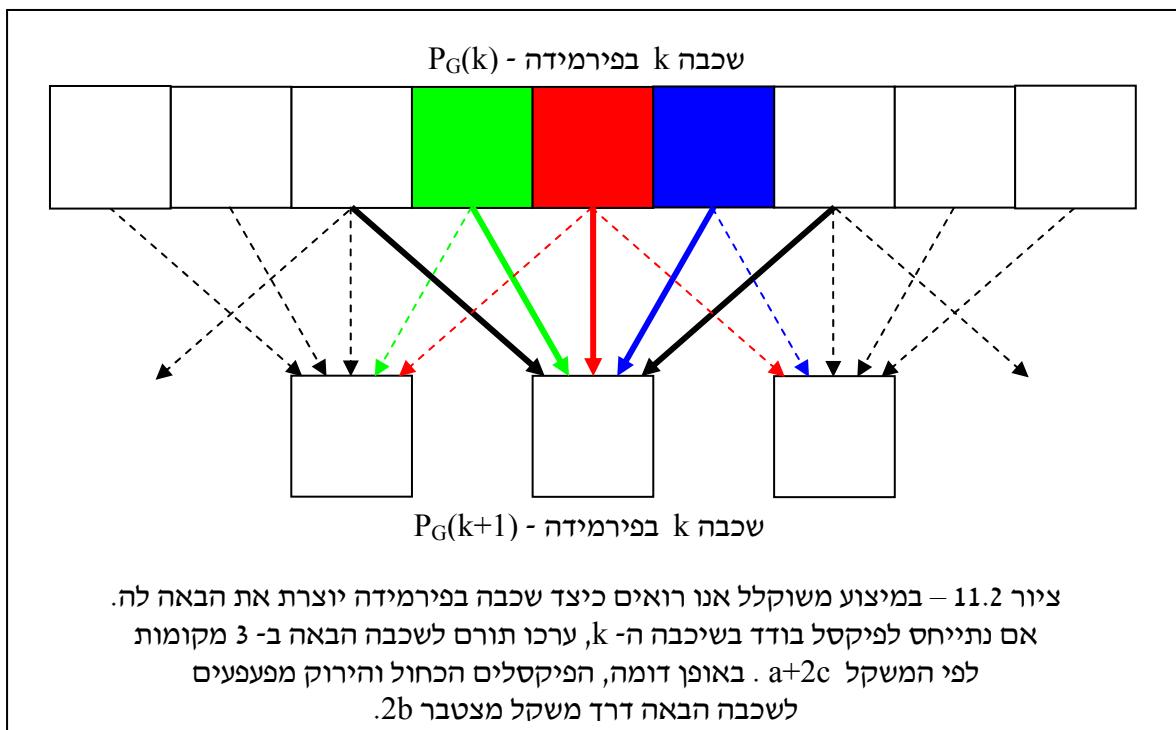
$$3. \forall k > 0, w_k \leq w_{k-1},$$

$$4. \forall k, \sum_{j=-\infty}^{\infty} w_{k+2j} = \text{const}$$

פרק 11 - מבנים פירמידליים ו שימושיהם בתמונות

הדרישה הראשונה מבטיחה כי הפעלת מסנן זה על התמונה תשמר את הממווצע המקומיי. הדרישת השנייה יוצרת מסנן סימטרי בו אין הטיה לצד אחד על פני الآخر. הדרישת השלישית באה לוודא דעיכת מונוטונית של השפעת השכנים על הפיקסל המרכזי בסינון. הדרישת הרביעית והמורכבת מכולן פירושה שככל פיקסל תורם בצורה זהה ביצירת שכבות הפירמידה הבאה המדוללת. ציור 11.2 ממחיש עניין זה ומצדיק את ניסוח הדרישת כפי שנקבע במשווהה.

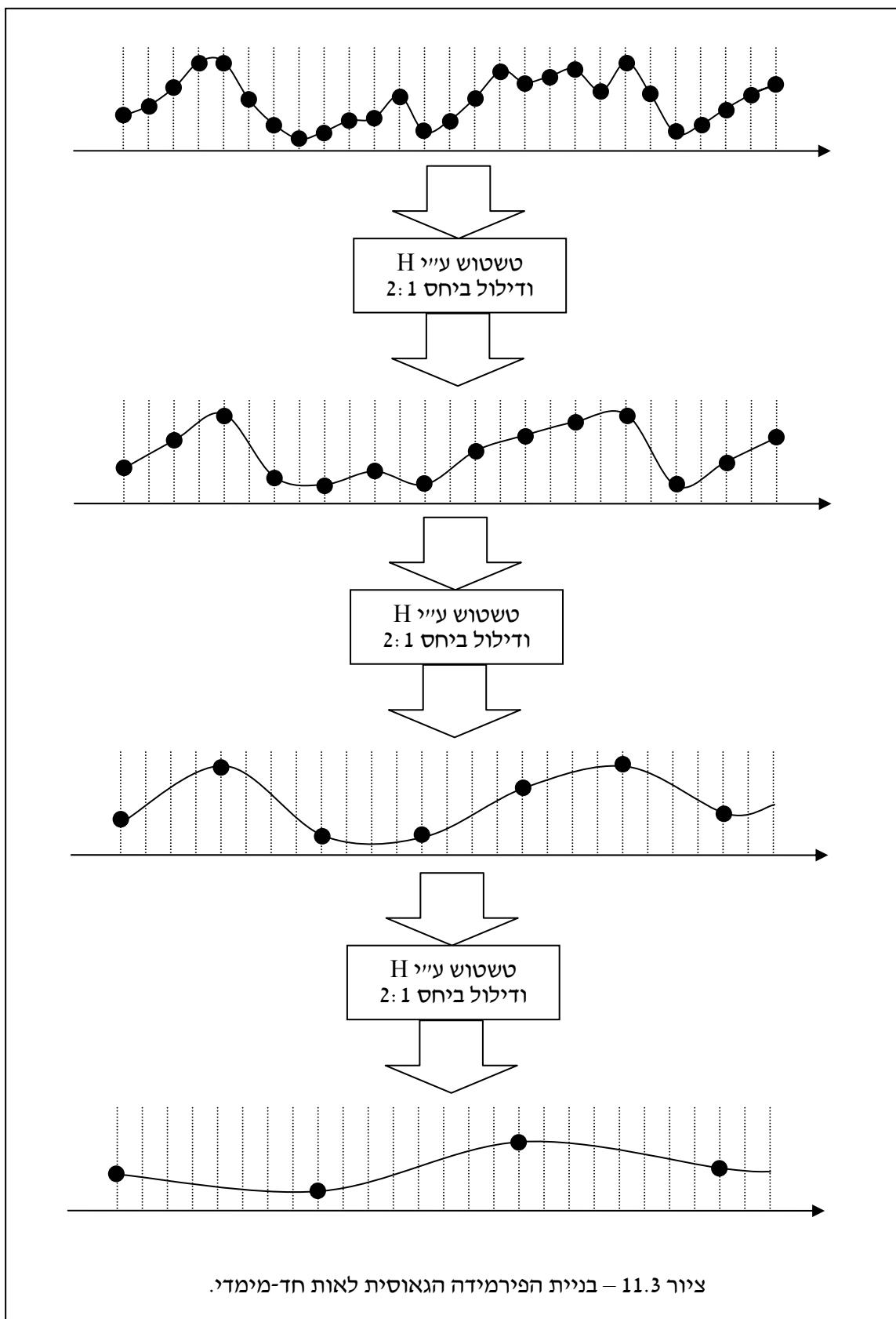
לשם דוגמה, אם נניח כי מסנן בניו $c = \{c, b, a, b, c\}$, הדרישת הראשונה תהיה $a \leq b \leq c$. הדרישת השנייה מوطמעת כבר בייצוג הפרמטרי שבחרנו, והדרישת השלישית תהיה $b = 0.25$ (צרוּף דרישת ראשונה ורביעית), a צריך להיות גדול מ- 0.25 (דרישה שלישית), ו- c יהיה $c = 0.5(a - 0.25)$. עבור הבחירה $a = 0.4$ מקבל מסנן דומה לפעמן גאוסי ומכאן השם "הפירמידה הגaussית". ציור 11.3 ממחיש את דרך בניית הפירמידה הגaussית לאותות חד-מימדיים, והרי הדבר זהה בדו-מימד.



ציור 11.2 – במציאות משוקל לנו לראות כיצד שכבה בפירמידה יוצרת את הבאה לה.
אם נתיחס לפיקסל בודד בשכבה ה- k , ערכו תורם לשכבה הבאה ב- 3 מקומות
לפי המשקל $c + 2a$. באופן דומה, הפיקסלים הכהולים והירוק מפעפים
לשכבה הבאה דרך משקל מצטבר $2b$.

מה אפשר לעשות בעזרת עזרת פירמידה גaussית? לא הרבה. זהו בראש ובראשונה אמצעי לייצוג התמונה במגוון רזולוציות. בעוד השכבה התחתונה עשויה מאמץ להראות את כל הפרטים, שכבות עליונות יותר זורקות את-אט את הפרטים הקטנים ומתרכזות בסיפור הכללי של התמונה. לדוגמה, בציור 11.1, נוצרת הcovariance המפורטת בשכבה הראשונה הופכת לחסרת פרטים מהשכבה השלישית ואילך.

נזכיר מספר שימושים בהם עושים שימוש בפירמידה Gaussית. מבנה פירמידלי זה יכול לעזור בחיפוש תבניות בתמונה. אחת הדרכים פשוטות ביותר והישירות ביותר לחפש תבנית בתמונה היא שיטות הקורלציה שדומות לה כבר פגשו בשערוך תנואה של בלוקים (ראה פרק 9). בהינתן תמונה X בגודל



N על N, ובהינתן תבנית Y בגודל L על L ($N < L$) אותה אנו מחותים (נניח פרצוף), החיפוש יכול להיעשות ע"י החישוב

$$C[n, m] = \frac{\left(\sum_{k=1}^L \sum_{j=1}^L X[n+k, m+j] Y[k, j] \right)}{\sqrt{\left(\sum_{k=1}^L \sum_{j=1}^L X[n+k, m+j]^2 \right) \left(\sum_{k=1}^L \sum_{j=1}^L Y[k, j]^2 \right)}}$$

בכל מקום $[m, n]$ בתמונה, וקביעת מציאותו במקומות בו ערך זה גדול מס' מסוים. זהה כמות עצומה של חישובים ונitin להקטינה בצורה משמעותית אם נבנה פירמידה גאומטרית ל- X ול- Y. במקרה זה, חיפושים מקדים בשכבות גבהות יותר של הפירמידה יכולים לגלוות איזורים חדשניים בהם יעשה חיפוש גם ברזולוציות גבהות (דהיינו, שכבות תחתונות של הפירמידה). שימוש אחר הוא כאשר הפירמידה נעשית רק על X והחיפוש נעשה בכל השכבות עם Y המקורי. גישה זו נוחה כאשר מחותים את התבנית בידיעה כי היא צפופה להופיע במנוגון גדלים אפשריים.

למעשה, ניתן להשתמש בפירמידה הגאומטרית לצורך שערוך תנועה בבלוקים והדבר יכול להביא להאצה ולהגדלת הדיקט בתהליך השערוך. לדוגמה, נניח שעורך התנועה צריך להיעשות בתמונות בגודל N על N בבלוקים של 32 על 32, וגדלי וקטורי התנועה הם 16 פיקסלים לכל כיוון. ניתן להמיר תהליך זה בשערוך תנועה על צמד תמונות שיכבה שנייה בפירמידה הגאומטרית ואשר גודלו הוא $N/2$. נניח שהפעם גDAL הפלים גDAL 16, ולכן וקטורי התנועה לא קטן. אבל עם השינוי שבוצע איזור החיפוש של הוקטור קטן בפקטור 4 (כי וקטורי התנועה המירביים הם 8 לכל כיוון כתעט) וכן גם התמוך לחישוב ה- MSE (בחישוב ההתאמה נצבור 256 איברים ולא 1024). פירוש הדבר חסכו בפקטור 16 רק במעבר לשכבת הפירמידה השנייה. למעשה, אין סיבה לעוזר כאן, ונitin לעלות עוד ועוד כשכל שכבה הבלוק ואיזור החיפוש קטנים. ברור כי יש לעוזר לפני שהבלוק מגיע לגודל קטן מיד שיביא לשערוך לאאמין, וכך ראוי לעוזר בגודל 4 על 4 (שכבה רביעית במקרה זה).

בහינתן שעורך שכבה $k+1$, יש לעדכנו ע"י ירידה חוזרת לשכבה ה- k. עדכון זה צריך לשכבה ה- k. בחיפוש מקומי סביר הוקטור שנמצא בשלב קודם (כשהוא מוכפל ב- 2). באופן כזה הקטנו באופן משמעותי את כמות החישובים הדרישה למציאות וקטור התנועה. לכן, אם נדרשו במקור לאיזור חיפוש בגודל 32 על 32 שימושו 1024 נקודות בדיקה של מיקום הבלוק, הרי שהפעם בשכבה הרביעית נבצע חיפוש ב- 4 על 4 מקומות ועם בלוק קטן משמעותית, כך שכמות החישובים זינחה. עם הפעוע לשכבה הבאה, נחפש בכל פעם רק ב- 8 מקומות סביר הפתרון הקודם, וכך נחסוך באופן גס ביחס של כ- 100 בזמנים החישובים הכוללת, כיוון שرك החישובים של השכבה התחתונה בפירמידה מהותיים בהשוואה.

בנוסף, הדיקט של השערוך צפוי לגודל משתני סיבות - אנו נוכל בלבצע חיפוש באיזור חיפוש גדול יותר, ומכיון שהוקטור הראשוני מחושב ע"פ מותווה כלל בתמונה תוך התעלמות מפרטים קטנים שיכולים להשחוף השערוך לערך לא רלוונטי. עם הפעוע מטה בפירמידה, אנו מגבלים את העדכון לשכיבת הפתרון הקודם, וכך מכונים את החיפוש למקום הרואוי. הדבר דומה למינימיזציה של פונקציה מאוד לא חלקה וחושש לנפילה למינימה לוקאלית. המעבר לשכבות הרזולוציה הבאות בפירמידה מושל להחלהקת פונקציה זו ומחיקת של מינימיות לוקליות כתוצאה לכך.

11.2 הפירמידה הפלטיאנית

החסרון המרכזי של הפירמידה הגאוסית הוא שכבות הפירמידה לא כוללות רזולוציות זרות המצטרפות זו לזו, אלא כל שכבה כוללת הפקת מסך הרזולוציות. כך יצא שהשכבה הראשונה בפירמידה כוללת את **כל מרביי התמונה** (זו התמונה ממנה אנו יוצאים בדרך!), וכל שכבה מעלה כוללת תוכן רזולוציוני המוכל בתמונות שמתחנה. חסרונו זה קל לתיקון ע"י החסרת תמונות – דבר המביא לבנייתה של הפירמידה הפלטיאנית.

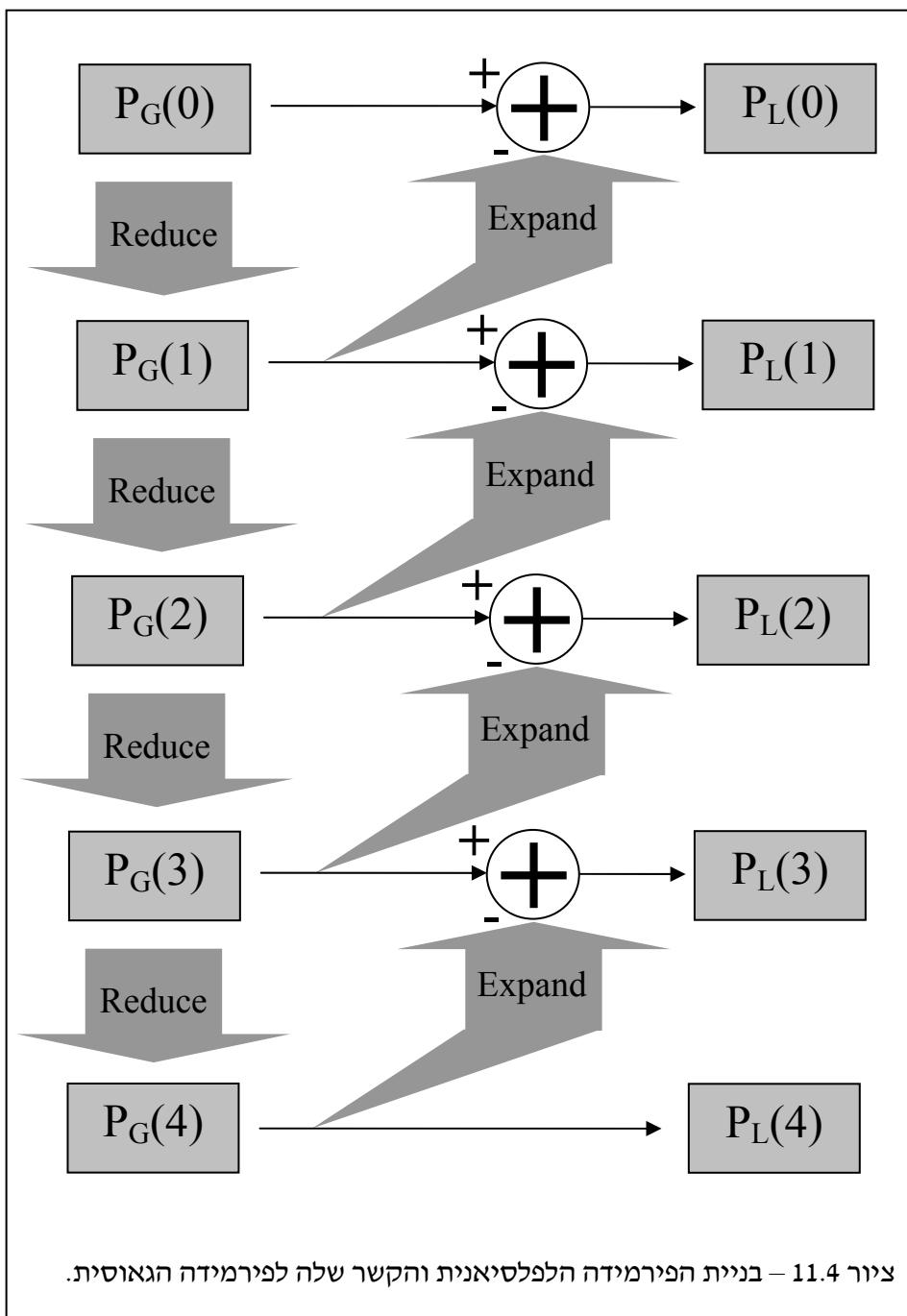
על רקע הנאמר לעיל, נציג לבנות את הפירמידה הפלטיאנית P_L באופן הבא – את התמונה X בגודל N על N מושטש ונקלטת התמונה HX בגודל זהה. נחסר ונגדיר את השכבה הראשונה בפירמידה ע"י $HX - X = P_L(0)$. זהה תמונה הפרשים בגודל המקורי N על N המכילה את הפרטים הקטנים כשהם מתוארים כഫזרים של הערך האמתי במקומות פחות ממוצע מוקומי שלו – כברرأינו בעבר כי זהה פועלות הפלטיאן על תמונה. את התמונה HX נוכל לדלך בשם שעשינו בפירמידה הגאוסית P_G . על התמונה המתבקשת שגודלה $N/2$ על $N/2$ נוכל לחזור על התהליך – טשטוש, החסרה שייצרת את השכבה הבאה בפירמידה, ודילול. נקבל סידרת תמונות תמונות הפרש ובראש הפירמידה תמונה קטינה המתלבצת עם התמונה בראש הפירמידה הגאוסית.

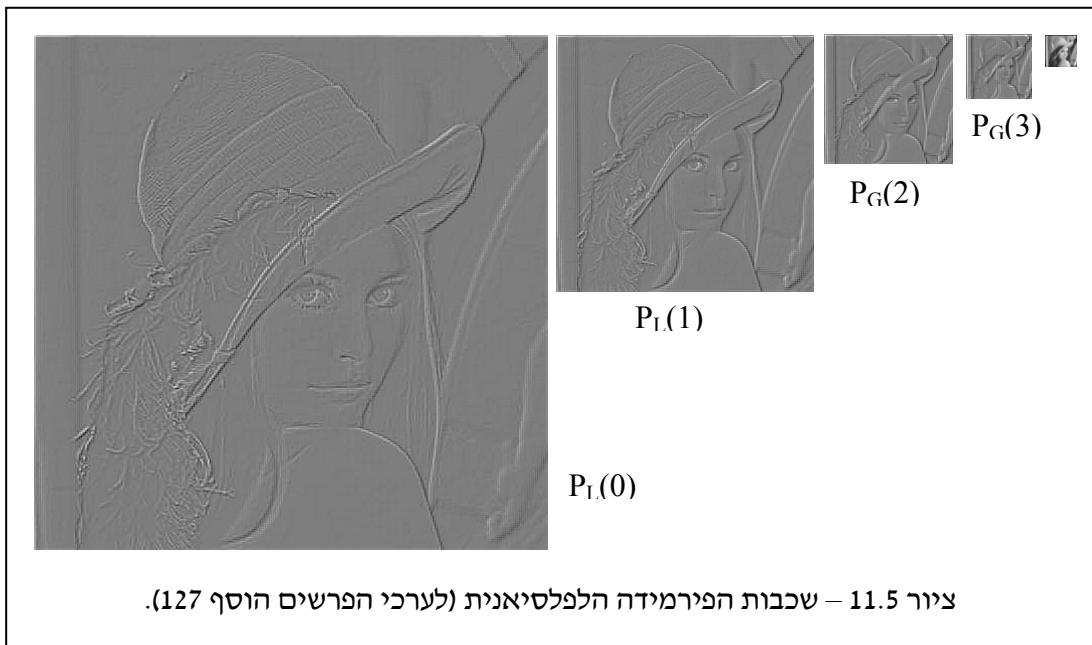
הגינוי ומפתחו ככל שהוא, תהליך פשוט זה של בניית סובל מחולשה מהותית והוא חוסר יכולת לחזור לתמונה המקורית. זאת כיוון שבפעולות הדילול אנו זורקים מידע שאינו מצוי בעת השחזור. אם נתיחס לבסיס הפירמידה, התמונה $P_L(0)$ מכילה בצורה נאותה את התדרים הגבוהים של התמונה. אבל, תדריה הנמוכים אשר אמורים להישמר בתמונה HX נפגמו בשל דילול תמונה זו לשם בניית צעדי הפירמידה הבאים.

תיקון הבעיה יכול להיעשות באופן הבא. במקום להגדיר את $P_L(0)$ כהפרש בין X ל- HX , נגדיר תמונה זו כהפרש בין X ל- HX לאחר דילול והרחבה. כלומר, התמונה HX תעבור דילול ממוקדם, ואז תנופה חוזרת לגודל המלא באינטראפלציה (כלשהו!) כל אינטראפלציה תספק את הצורך בשחזור התמונה מהפירמידה!). מתמונה מוגדלת זו יעשה החיסור. במונחים אלגבריים, נגדיר את האופרטור $Reduce$ כאופרטור מטשטש ומדיל, והאופרטור $Expand$ כאופרטור המבצע הגדלה חוזרת. אז הפירמידה תיבנה בסידרת הצעדים הבאים, כאשרנו משתמשים בבנייה הפירמידה הגאוסית:

1. התחלי מהתמונה X . זהה גם $P_G(0)$ – שכבת הפירמידה הגאוסית ההתחלתית.
2. קבע את k ל- 1.
3. בצעד ה- k , בצעוי פעולה $Reduce$ על $P_G(k-1)$ וקבלו את התמונה $P_G(k)$ – שכבת הפירמידה הגאוסית ה- k .
4. בצע פעולה $Expand$ על $P_G(k)$ והגדיר $P_L(k) = P_G(k) - Expand\{P_G(k)\}$
5. הגדל את k . אם הגיענו לשכבה $M = k$ קבע $P_L(k) = P_G(k)$ ועצור. אחרת חוזר לשלב 3.

בבנייה זו של הפירמידה, בניית התמונה המקורית פשוטה ותיעשה בצורה רקורסיבית ע"י הפעלת **Expand** על התמונה בראש הפירמידה ($P_L(M)$), הוספת השכבה ה- $M-1$, שוב **Expand** וחזור עד להגעה לתחתית הפירמידה הלפלסיאנית, כשההתמונה הנוצרת זהה במדויק לו ממנה התחלו. ציור 11.4 ממחיש את דרך בניית שתי הפירמות והקשר ביניהן, וציור 11.5 מראה את התמונות המתקבלות בפועל בהקלה לצירור 11.1. תהליך בניית התמונה מהפירמידה הלפלסיאנית מתואר בציור 11.6 - פעוף מעלה בציור 11.1 כמשמעותיים בשכבה (4) $P_L(4)$ מוביל לתמונה הרצוייה, וכל לראות את הדמיון לתהליכי הבניה בציור 11.4.





ציור 11.5 – שכבות הפירמידה הלפלסיאנית (לערכי הפרשים הוסף 127).

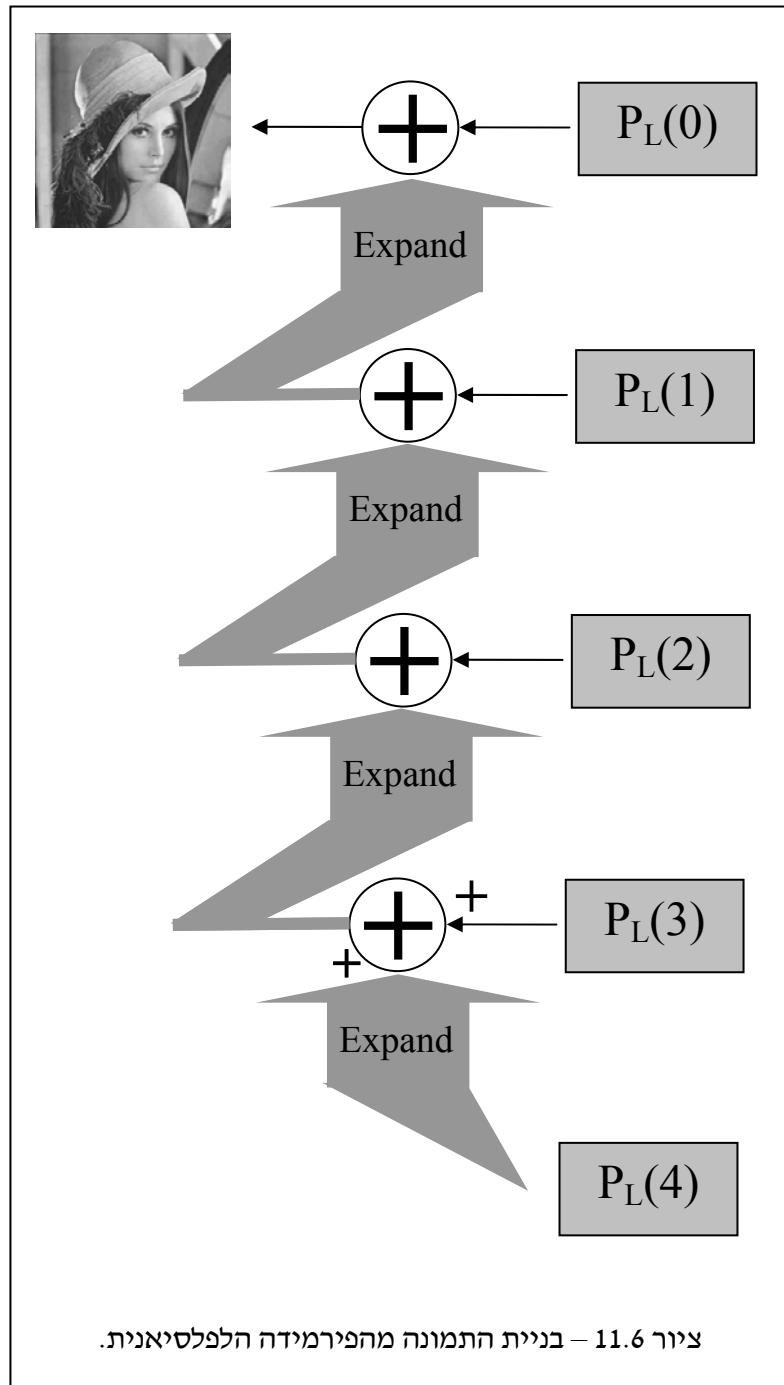
כיצד נבחר את פעולה Expand ? אמרנו קודם כי לא חשוב כיצד תוגדר, בנית התמונה מהפירמידה תביא לשחזר מדויק כל עוד התהליך עקבי בموון שאותה Expand משמשת בבנייה הפירמידה ובשחזור. בציור 11.5 בנו את הפירמידה על ידי פעולה Pixel Replication בו בעליה ברזולוציה כל פיקסל מייצר רביעית ערכיים זהים לו. זהה קרוב לוודאי הגישה פשוטה ביותר האפשרית וגם הגורעה ביותר.

מה ההבדל בין בחירות שונות? נעה על שאלת זו דרך מתן אינטראפרטציה מעט שונה לפירמידה הלפלסיאנית. כאשר בונים את התמונה $P_L(k-1)$ אנו עושים זאת ע"י המשוואة
$$P_L(k-1) = P_G(k-1) - \text{Expand}\{P_G(k)\} = P_G(k-1) - \text{Expand}\{\text{Reduce}\{P_G(k-1)\}\}.$$

כלומר, התמונה $P_L(k-1)$ בפירמידה הגאוסית היא תמונה היסוד לבניה זו. אנו מחשבים את התמונה $P_G(k-1)$ $\text{Expand}\{\text{Reduce}\{P_G(k-1)\}\}$ ומשתמשים בה לחיזוי של התמונה המקורית $P_G(k-1)$. תמונה החיזוי היא תמונה נטולת פרטים והיא מנסה לנבא את התמונה שמתחתיה (המפורטת יותר) בפירמידה הגאוסית. אם אין פרטים חדים בתמונה $P_G(k-1)$, החיזוי יביא לשאריות קטנה. חיזוי מוצלח הוא חיזוי בו התמונה $P_L(k-1)$ כוללת אנרגיה דלה. אם החיזוי מוצלח זה יכול להיות בגלל שתי סיבות:

1. ישכן כי תוכן התמונה הוא כזה שאין רמת פרטים גבוהה (רזולוציה גבוהה). במקרה כזה החיזוי ייתן ערכיים מאד לאפס. במקרה כזה הפירמידה מצליח לומר לנו את מידת הרזולוציה האמיתית של התמונה (ואפילו כתלות במקום) לפי העוצמה של תמונות הפירמידה בכל מקום. אם לדוגמה מהשכבה השנייה ומטה מתקבלים ערכי אפס בחצי העליון של התמונה בעוד שערכי שכבות אלה בחלק התחתון שונים מהוותית ממאפס, נדע כי החצי העליון מוגבל ברזולוציה וכי היה בפחות פיקסלים בפקטור 4 לתיאור אזור זה.

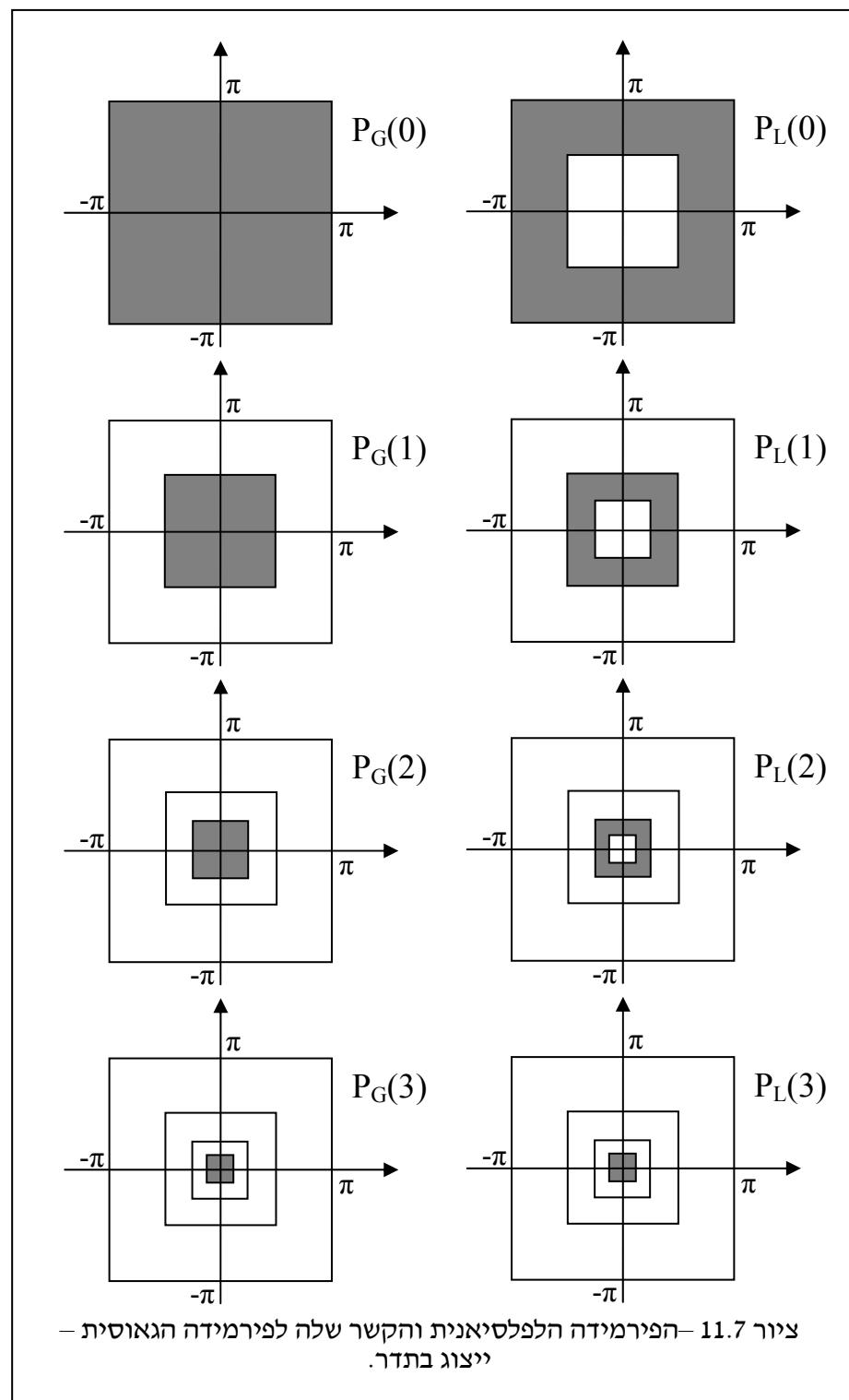
ללא קשר לנקודת הקודמת, ישן דרכים שונים להגדיר את הפעולה Expand וכל דרך תיתן שגיאת חיזוי שונה. רצוי לבחור בהגדרת Expand שתיתן חיזוי מוצלח – חיזוי שיקטין את אנרגיית השארית ב- $P_L(k-1)$. עם זאת, יש גם עניין בבחירה של פעולה פשוטה שקללה לביצוע מבחינה חישובית.



צייר 11.7 ממחיש את משמעות הפירמידה הלפלסיאנית (והגאוסיאנית) בתדר. אנו רואים את התמך התדרי π^2 – בו התמונה מלאת מידע. הפירמידה הגאוסיאת לוקחת רבע פנימי המתקבל ע"י מסנן מעביר נומכיס הולך וצר בכל שלב וכך מתרכזות בעליית השכבות בתדרים הולכים וקטנים. הפירמידה הלפלסיאנית בונה תМОנות הנבנות כמסנן מעביר פס כשהפס נע לעבר הראשית ונهاיה צר יותר

פרק 11 - מבנים פירמידליים ו שימושיהם בתמונות

ו יותר עם התקרובותו בראשית. מצור זה גם ברור כי כל שעליינו לשות הוא לסכם את רצועות התדר הלו בפירמידה הפלסיאנית על מנת לקבל את התמונה המלאה. עניין מטה מעט הוא נושא הדילול שלא ניכר מהציגו בתדר, כיוון שמדובר בהתרמה מנורמלת.



במקור הוצאה הפירמידה הפלסיאנית כפירוק של תמונה לשם דחיסתה. הרעיון פשוט – בניית הפירמידה לתמונה, ואז ביצוע קוונטייזציה לשכבותיה כך שערכאים קרובים לאפס מאופסים, וכל ערך אחר מומך לערכי ייצוג קלים לדחיסה. ניתן אף לחשב על קוונטיזור שונה לכל שכבה. בפועל ביצועי אלגוריתם

זה נופלים מלאה של JPEG. יש לזכור כי תמונה בגודל N על N מומרת בתהליך בניית הפירמידה ל- N^2 פיקסלים - N^2 לתמונה $(0, P_L)$, וכן הלאה, לקבלת הטו

$$\cdot \sum_{k=0}^{\log N} \left[\frac{N}{2^k} \right]^2 \approx N^2 \cdot \sum_{k=0}^{\infty} 4^{-k} = \frac{N^2}{1 - \frac{1}{4}} = \frac{4N^2}{3}$$

לכן, כבר לאחר בניית הפירמידה יש לנו ניפור של כמות המידע ב- 33%. לעומת זאת, המידע במבנה החדש הינו נמוך אנטרופיה כיון שככלו מידע הפרושים שעבורו כברرأינו כי הפילוג בעל צורה לפלסיינית (הנה עוד סיבה לשמה של הפירמידה).

אם רצוננו בדחיפסה, נוכל גם לבחור את הפעולה **Expand** באופן מיטבי להקטנת שגיאות החיזוי, ועימן את האנתרופיה. ניתן להציג גישות שונות לבחירת **Expand**. דרך אטרקטיבית היא התייחסות לתהליך הניפור כבעיית **Up-Scaling** של תמונה ועשית שימוש בכלים שלמדנו בפרק 6 על שחזור תמונות תוך שימוש ב- **Prior** לבניית חיזוי מוצלח יותר.

מעניין לציין כי אלגוריתם דחיסה זה מצטייני תיאורטית באמצעות החיזוי וההתמורה שנדונו בהרחבה בפרק 8 כדריכים לדחיסת תמונות. עד כה נתנו אינטראפטציה של חיזוי לפירמידה הפלסיינית, אם כי זהו חיזוי מעניין בהיותו לא סיבתי כמו החיזויים בהם דנו בפרק 8, ומכאן כוחם פוטנציאלי לביצועים טובים יותר מגישות הסורקנות את התמונה באופן סיבתי חד-מימדי.

ניתן גם לחשב על בניית הפירמידה כהתמורה ואפיו התמורה ליניארית. לצורך זה נניח כי פעולות **the-Reduce** וה- **the-Expand** הין פעולות ליניאריות וככלאלה הן ניתנות לייצוג כמטריצות R_k ו- E_k . המטריצה R_k מקבלת תמונה בגודל $(N/2^k) \times (N/2^k)$ לאחר סיור לקסיקוגרפי ומונבה את תוצאתה הטעטווש במסנן ליניארי כלשהו (ואפיו תלוי מקום) כשלאחריו דילול. לכן מטריצה זו תהיה מלכנית עם פירות שורות מעמודות, בגודל של $(N^2/2^{2k}) \times (N^2/2^{2k+2})$. באופן דומה, המטריצה E_k מקבלת תמונה בגודל $(N/2^{k-1}) \times (N/2^{k-1})$, ובמצעת הגדלה ואינטראפטציה לקבלת תמונה גדולה פי 2 בכל ציר (שוב – כוקטור בייצוג לקסיקוגרפי). לכן מטריצה זו תהיה בגודל $(N^2/2^{2k}) \times (N^2/2^{2k+2})$.

כדוגמה לצמד מעניין של מטריצות **the-Reduce** וה- **the-Expand**, ניתן לבחור את R_k כך שיבצע טשטוש כבפירמידה הגאוסית עם $\{w_k\}_k = \{0.05, 0.25, 0.4, 0.25, 0.05\}$ ואחריו דילול ביחס 1:2. את המטריצה E_k נוכל לבחור כפעולות ה- **Adjoint** דהיינו $-R_k^T = E_k$. בשיל דרך בניית ערכי המסנן לפירמידה הגאוסית ייצא כי אינטראפטציה זו מתאימה (אם כי נדרש לכפול ב- 4 את תוצאתה, כפי שנראה). אם נכתב מספר שורות אופייניות ב- R_k כפעולה חד-מימדית, הוא יראו כך

$$R_k = \begin{bmatrix} & & & & & \vdots \\ 0.05 & 0.25 & 0.4 & 0.25 & 0.05 & & \\ \dots & 0 & 0.05 & 0.25 & 0.4 & 0.25 & 0.05 & \dots \\ & & & 0.05 & 0.25 & 0.4 & 0.25 & 0.05 \\ & & & & \vdots & & & \end{bmatrix}$$

פעולת השחלוף תיתן מטריצה בעלת המבנה

$$E_k = \begin{bmatrix} & & & & & \vdots \\ & 0.05 & & & & \\ & 0.25 & & & & \\ & 0.4 & 0.05 & & & \\ & 0.25 & 0.25 & & & \\ \dots & 0.05 & 0.4 & 0.05 & \dots & \\ & 0.25 & 0.25 & & & \\ & 0.05 & 0.4 & & & \\ & 0.25 & & & & \\ & 0.05 & & & & \vdots \\ & & & & & \end{bmatrix}$$

אנו רואים כי שורות מטריצה זו מסתכמות כולם ל- 0.5 – דבר המבטיח כי צירופים ליניאריים שנעשים ע"י מטריצה זו יכולים להיחשב כאינטראפולציה. לכן יש לכפול ב- 2 את התוצאה לשם תיקון הערך המומוצע. תוכנות סכומי השורות הקבועים נובעת מהדרישה הריבועית של תרומה שווה בתכנון המסנו לפירמידה הגאומטרית. במעבר לדו-מימד, כיוון שככל הפעולות פרידוט, מכפלת קרונקר של מטריצות אלה יתנו את הפעולה הדו-מימדית הכלולית, ולכן נדרש גורם תיקון של 4.

בעזרת מטריצות אלו, בניית הפירמידה הגאומטרית תיראה כך, כאשרו מתחילה בתמונה המקור X

בגודל N על N (ניקח דוגמא בה עומק הפירמידה הוא 5) :

$$\begin{aligned} \underline{P}_G(0) &= \underline{X} \\ \underline{P}_G(1) &= R_0 \underline{X} \\ \underline{P}_G(2) &= R_1 R_0 \underline{X} \\ \underline{P}_G(3) &= R_2 R_1 R_0 \underline{X} \\ \underline{P}_G(4) &= R_3 R_2 R_1 R_0 \underline{X} \end{aligned}$$

מתוך קשרים אלו, בניית הפירמידה הלפלסיאנית תיעשה ע"י

$$\begin{aligned} \underline{P}_L(0) &= \underline{P}_G(0) - E_0 \underline{P}_G(1) = (I - E_0 R_0) \underline{X} \\ \underline{P}_L(1) &= \underline{P}_G(1) - E_1 \underline{P}_G(2) = (R_0 - E_1 R_1 R_0) \underline{X} \\ \underline{P}_L(2) &= \underline{P}_G(2) - E_2 \underline{P}_G(3) = (R_1 R_0 - E_2 R_2 R_1 R_0) \underline{X} \\ \underline{P}_L(3) &= \underline{P}_G(3) - E_3 \underline{P}_G(4) = (R_2 R_1 R_0 - E_3 R_3 R_2 R_1 R_0) \underline{X} \\ \underline{P}_L(4) &= \underline{P}_G(4) = R_3 R_2 R_1 R_0 \underline{X}. \end{aligned}$$

לכן, הפירמידה הלפלסיאנית אינה אלא המכפלה

$$\begin{bmatrix} \underline{P}_L(0) \\ \underline{P}_L(1) \\ \underline{P}_L(2) \\ \underline{P}_L(3) \\ \underline{P}_L(4) \end{bmatrix} = \begin{bmatrix} I - E_0 R_0 \\ (I - E_1 R_1) R_0 \\ (I - E_2 R_2) R_1 R_0 \\ (I - E_3 R_3) R_2 R_1 R_0 \\ R_3 R_2 R_1 R_0 \end{bmatrix} \underline{X}$$

במונחי התמרה בהם עסקנו בפרק 7, זהה התמרה ליניארית (גם ההתמרה וגם היפוכה). בשל יצירת מסדרה של $\frac{3}{N^2} 4$ מקדמי ייצוג, התמרה זו כוללת יתרות (Over-Complete).

יש לנו גם את ההתמרה ההפוכה – תהליך קritisת הפירמידה הפלסיאנית לתמונה המקור.

המשווה לשימוש תהיה

$$\underline{X} = E_0 [E_1 [E_2 [E_3 \underline{P}_L(4) + \underline{P}_L(3)] + \underline{P}_L(2)] + \underline{P}_L(1)] + \underline{P}_L(0)$$

ביצוג מטריצי, קשר זה יראה כך:

$$\underline{X} = \begin{bmatrix} I & E_0 & E_0 E_1 & E_0 E_1 E_2 & E_0 E_1 E_2 E_3 \end{bmatrix} \begin{bmatrix} \underline{P}_L(0) \\ \underline{P}_L(1) \\ \underline{P}_L(2) \\ \underline{P}_L(3) \\ \underline{P}_L(4) \end{bmatrix}$$

ומכפלת מטריצת ההתמרה ההפוכה במטריצת ההתמרה תהיה

$$\begin{aligned} & \begin{bmatrix} I & E_0 & E_0 E_1 & E_0 E_1 E_2 & E_0 E_1 E_2 E_3 \end{bmatrix} \begin{bmatrix} I - E_0 R_0 \\ (I - E_1 R_1) R_0 \\ (I - E_2 R_2) R_1 R_0 \\ (I - E_3 R_3) R_2 R_1 R_0 \\ R_3 R_2 R_1 R_0 \end{bmatrix} = \\ &= I - E_0 R_0 + E_0 (I - E_1 R_1) R_0 + E_0 E_1 (I - E_2 R_2) R_1 R_0 + \\ &+ E_0 E_1 E_2 (I - E_3 R_3) R_2 R_1 R_0 + E_0 E_1 E_2 E_3 R_3 R_2 R_1 R_0 = I \end{aligned}$$

אנו רואים כי בדרך הגדרת ההתמרה ההפוכה בינוי בפועל את מטריצת הפסאודו-היפוך של מטריצת ההתמרה, ובכך הבטחנו חוזרת לתמונה ממנה יצאו. מכאן גם ברור כי בבחירה צמד הפעולות – התמרה קדימה וההתמרה ההפוכה – בחרכנו בהתמרה שmbיאה למינימום את מקדמי התמרה – דבר שמשמעותו עס הרצון לשימוש בתוצאה לדחיסה.

ישום מעניין אחר בו עושים שימוש בפירמידה הפלסיאנית הוא מיזוג תמונות ברענון שזכה לשם Multi-Resolution Spline. נניח כי נתונות שתי תמונות האמורות להתחבר ביניהן (לשם פסיפס). לרוב חיבור ישיר של התמונות מניב תפְר צורם בשל אי התאמה של תנאי ההארה, סטיות קלות בתיאום ביניהן ועוד. באמצעות פירמידה הפלסיאנית ניתן למיזוג תמונות תוך כדי טשטוש נעים של התפר. לשם פשוטות נניח כי שתי תמונות שוות גודל צרכות להתחבר כך שאחת תורמת את החצי הימני, והאחרת את החצי השמאלי. חיבור שני חצאים אלו ישירות יגרום כאמור לתפר בולט ולא טבעי. האלטרנטיבה המוצעת היא פירוק שני התמונות לפירמידות הפלסיאניות, ובנית פירמידה משולבת בה כל שכבה בנוייה כמיזוג ישיר

פרק 11 - מבנים פירמידליים ושימושיהם בתמונות

של חצאי התמונה מהפירמידות של שתי התמונות. כאשר משוחרים את התמונה מהפירמידה החדשה, ראיינו כבר כי תהליך הבניה מתואר ע"י

$$\begin{aligned}\underline{X} &= E_0 [E_1 [E_2 [E_3 P_L(4) + P_L(3)] + P_L(2)] + P_L(1)] + P_L(0) = \\ &= P_L(0) + \\ &\quad + E_0 P_L(1) + \\ &\quad + E_0 E_1 P_L(2) + \\ &\quad + E_0 E_1 E_2 P_L(3) + \\ &\quad + E_0 E_1 E_2 E_3 P_L(4).\end{aligned}$$

לכן, השכבה העליונה בעלת הרזולוציה הנמוכה תעבור סידרה ארוכה של צעדי Expand שבהם היא מטושטשת. לכן, אם שני החצאים אינם توאימים בשכבה זו, תבוא מריחה ותיצור תפיר נקי בין השתיים. כשןרד עם שכבות הפירמידה, כל שכבה עברת סידרה הולכת וקצרה של פעולות Expand ולכן הטשטוש הולך ופוחת. בשכבה האחרונה נמזג את הפרטים ללא כל טשטוש. בוצרה זו מידע נמוך תדר יטושטש חזק יותר ממידע גובה תדר, ממש כפי שאנו רוצים שיקרת. זאת כיון שכשמיידע בתמונה חלק מאדן, נחוץ לטשטוש רחב תמק לשם חיבור שני המתחים, ולעומת זאת, כשהמידע פחות חלק, החלקה רחבה تعالים פרטניים ותזיק. ציור 11.8 מראה תוצאות תהליכי זה על מיזוג שני תמונות זהות שכוללות כל אחת "תנאי הארה אחרים" (ובפועל, תוספת של שתי שדות אקריאים וחלקים). פרטים נוספים על יישום זה והפירמידה הלפלסיאנית מצויים במאמריהם של Burt ו- Adelson.



ציור 11.8 – מיזוג תמונות בעזרת פירמידה לפלסיאנית. לתמונה Lena הוספו פעמיים שדה אקראי שונה שהוחלק. בלקיחת חצי מכל התמונה קו התפר ניכר וצורם. במיזוג של שתי הפירמידות הלפלסיאניות קו זה טושטש ללא אובדן פרטים בתמונה. בסימולציה זו עשינו שימוש בפעולת Expand בתמונה. המהווה Adjoint לפעולת ה- Reduce כפי שתואר קודם.

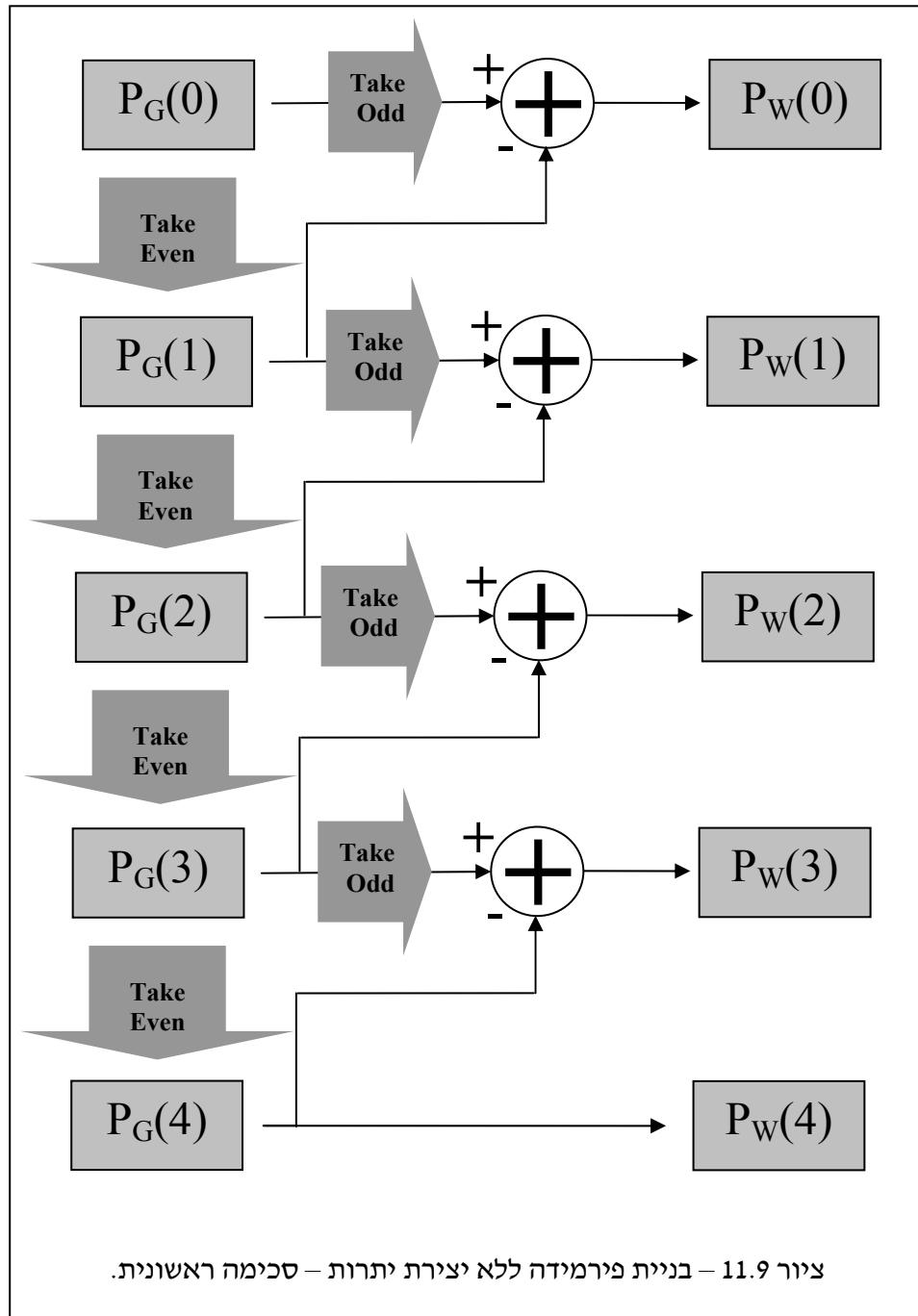
11.3 התמרת ה- Wavelet ע"י Lifting Scheme

הפירמידה הלפלסיאנית נתנה לנו ייצוג נאות של מגוון הרזולוציות הקיימות בתמונה, אך ראיינו כי בפирוק זה יצרנו יותר מידע ביחס של 4/3. דבר זה פעל נגדנו, למשל, כאשר קשנו לדוחס בצורה עילית את התמונה. האם ניפוח מידע זה חיוני? בפרק זה נראה כי ניתן להציג סכימה דומה מאוד לפירמידה הלפלסיאנית בה אין ניפוח מידע. הדבר המעניין שיצוץ מתוך הסכימה המוצעת אינו אלא התמרת ה-Wavelet בה כבר פגשנו בפרק 7. חשוב להבהיר כי פרק זה אינו בא להציג את התמרת ה-Wavelet מהזוויות הקלאסית ולא יכול להקיים את כל המידע בתחום – נקודת מבטנו צרה ותתמקד ביצוג מולטי-רזולוציוני וייעיל של התמונה הנתונה, תוך בניית פירמידה הדומה לפירמידה הלפלסיאנית.

נזכיר את התהליך הבסיסי הבא שיתואר לאות חד-מימדי \underline{X} - וקטור באורך $N = 2^n$. כזכור, בבניית פירמידה גאוסית אותן זה הוא גם השכבה התחתונה - $P_G(0)$. נפעיל פעולה Reduce על $P_G(0)$ ונקבל אותן באורך $N/2 = 2^{n-1}$ לו קראנו $P_G(1)$. בשלב זה נניח כי פעולה Reduce פשוט מבצעת דילול ובוחרת את האיברים הזוגיים בסידרה. בעזרתה אותן נבצע חיזוי של איבורי האי-זוגיים של הווקטור המקורי $P_G(0)$, ונחשב את שגיאות החיזוי. נגיד וקטור שגיאות אלו $C = P_G(0) - P_G(1)$. וקטור זה יהיה באורך $2^{n-1} = N/2$ והוא ישמר כתיאור מרכיבי המידע העדין ביותר באות המטופל וזויה שכבת האפס בפירמידה שאנו בונים כעת.

בשלב זה, בהתייחס ל- $P_G(1)$, נוכל לחזור על התהליך – פעולה Reduce לקבלת $P_G(2)$, חיזוי החלק האי-זוגי השני $P_G(1) = P_G^{\text{Odd}}(1) - P_G(2)$, קבלת וקטור שגיאות חיזוי $P_G^{\text{Odd}}(1) - P_G(2)$ (שנסמן $P_W(1)$) – ו然后再 חזר חלילה. תהליך זה יכול להימשך עד לשלב ה-($n-1$) – יי בו $P_G(n)$ הוא סקלר. השכבה الأخيرة בפירמידה שנוצרה, $P_W(n)$, תהיה פשוט $P_G(n)$. אוסף כל הדגימות שייצרו יכלול רק N דגימות, וכי אין ניפוח בכמות המידע, כל השכבות למעט האחרונה כוללות ערכי הפרשים. פסי התדר המכוסים בשכבות אלו זהים למתחואר בציור 11.7. ציור 11.9 ממחיש את דרך בניית הפירמידה החדשה המוצעת.

בהתנתו הפירמידה $P_W(n)$, בניית אותן המקורי תיעשה ע"י היפוך סדר הפעולות. נתחיל ב- $P_W(n)$ אשר הינו האיבר השני $P_G(n-1)$ – זה סקלר. נוסיף את ערכו ל- $P_W(n-1)$ (סקלר אף הוא), ונקבל את האיבר הראשון של $P_G(n-1)$. כך השלמנו את בניית $P_G(n-1)$. באופן דומה, הוספה $P_W(n-2)$ ל- $P_G(n-2)$ תיתן את החלק האי-זוגי ב- $P_G(n-1)$, והחלק הזוגי אינו אלא $P_G(n-1)$. תהליך זה יעלה כל הדרך עד לבניית $P_G(0)$ הוא אותן המקורי ממנו התחילנו. ציור 11.10 מראה תהליך זה.



דוגמה 11.1: ניקח את הסדרה הבאה ונפעיל עליה את התהליך שתואר.

$$\underline{X} = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14 \ 15 \ 16] = P_G(0)$$

ה הפרדה לאיברים הזוגיים והאי-זוגיים תיתן

$$P_G(1) = [2 \ 4 \ 6 \ 8 \ 10 \ 12 \ 14 \ 16] = \underline{X}_E$$

$$\underline{X}_O = [1 \ 3 \ 5 \ 7 \ 9 \ 11 \ 13 \ 15]$$

לאחר החיזוי קיבל

$$P_W(0) = [-1 \ -1 \ -1 \ -1 \ -1 \ -1 \ -1 \ -1]$$

הפרדת $P_G(1)$ לזוגיים ואי-זוגיים תיתן באופן דומה

$$\underline{P}_G(2) = [4 \quad 8 \quad 12 \quad 16] \Rightarrow \underline{P}_W(1) = [-2 \quad -2 \quad -2 \quad -2]$$

ובאופן דומה

$$\underline{P}_G(3) = [8 \quad 16] \Rightarrow \underline{P}_W(2) = [-4 \quad -4]$$

$$\underline{P}_G(4) = [8] \Rightarrow \underline{P}_W(3) = [-8]$$

$$\underline{P}_W(4) = [8]$$

ולכן פירמידת המוצא תהיה

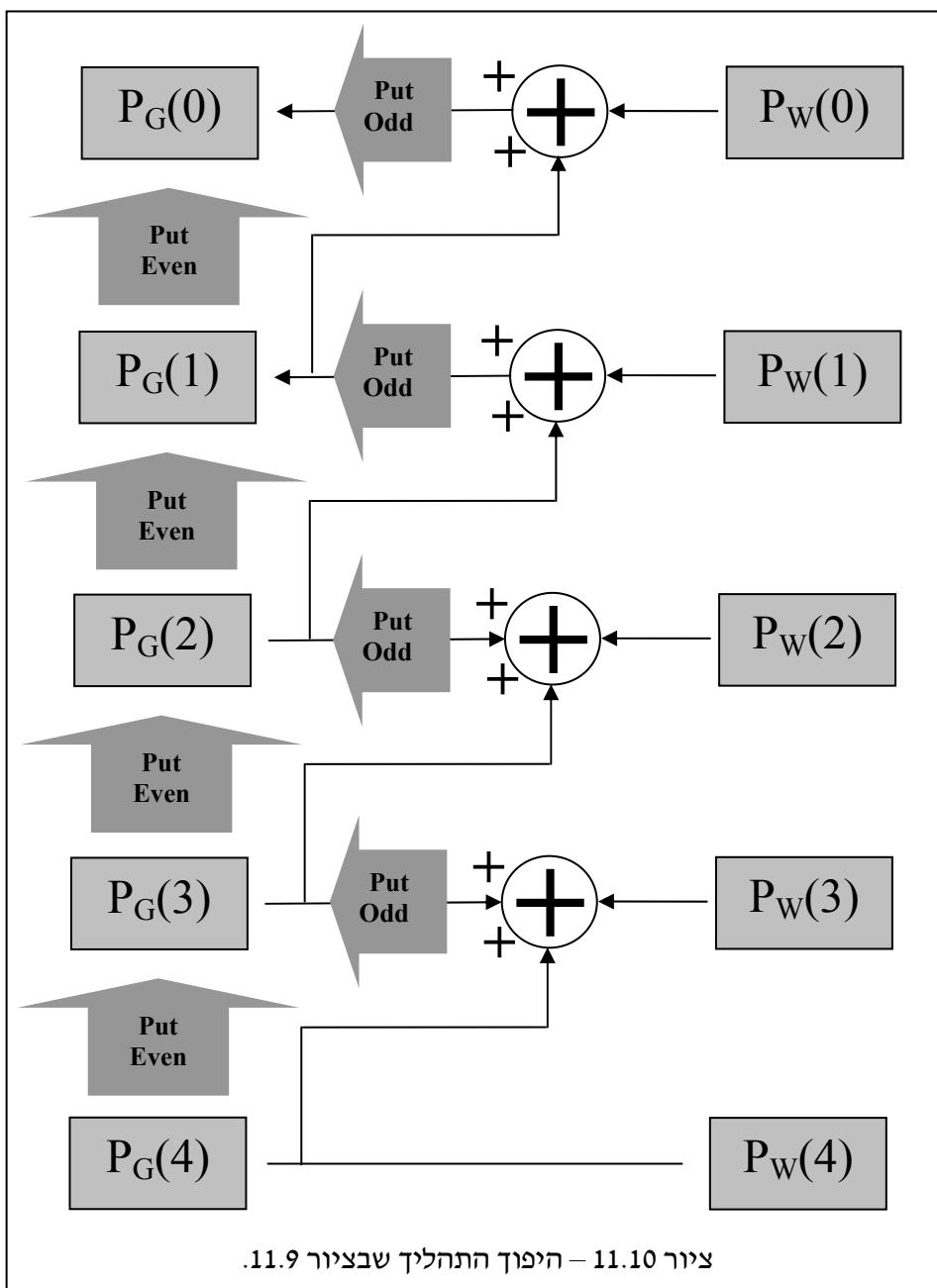
$$\underline{P}_W(0) = [-1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1 \quad -1]$$

$$\underline{P}_W(1) = [-2 \quad -2 \quad -2 \quad -2]$$

$$\underline{P}_W(2) = [-4 \quad -4]$$

$$\underline{P}_W(3) = [-8]$$

$$\underline{P}_W(4) = [8]$$



פרק 11 - מבנים פירמידליים ו שימושיהם בתמונות

נשים לב לכך שבחרית פעולת Reduce מורכבת יותר מאשר דילול פשוט תיצור בעיה בהיפוך, כיוון שביצועה בכיוון ההפוך אינה בהכרח אפשרית. עובדה מצערת זו מובילת לשאלת: האם נוכל להכליל סכימה זו להשתמש בפעולות Reduce ו-Expand כליליות יותר מאשר בחירת זוגיים ואי-זוגיים? הכללה כזו הוצעה תחת השם Sweldens Lifting Scheme ע"י באמצעות שנות ה-90' והוצאה כדריך אלטרנטיבית מפנה לבניית התמרת Wavelet-ה. הרעיון הוא הבא – בהינתן סידרת המקור \underline{X} מוצע לפצלה לאיברים זוגיים \underline{X}_E ו-אי-זוגיים \underline{X}_O מקודם. בעזרת סידרת הזוגיים \underline{X}_E יש לבנות חיזוי לאי-זוגיים ולהחליפים בערכי שניאת החיזוי, דהיינו החיזוי יהיה פונקציה מהצורה $\text{Pred}\{\underline{X}_E\}$, וערך האי-זוגיים יתפכו להיות $\underline{X}_O^{\text{New}} = \underline{X}_O - \text{Pred}\{\underline{X}_E\}$. למעשה יכולנו לעצור כאן ואז צירורים 11.9 ו- 11.10 היו משתנים לכלול אבן בניין חדשה Pred באופן המתוואר בציור 11.11, שבו זה אנו מתרכזים בשלב הראשון. ברור כי בהינתן $P_G(1)$ יכול כל התהליך לחזור על עצמו שוב ושוב עד להגעה בראש הפירמידה עם $P_G(n)$ סקלר.

בסכימה חדשה זו הבלוק Pred ייתן שיפור בביטויים כיון שהニアת החיזוי יכולה להיות קטנה יותר עברו בחירה אותה. למשל, ניתן לבחור למצע שני שכנים מימיינו ומשמאלו של האיבר האי-זוגי לשם חיזוי טוב. עברו הבחירה הכללית יותר

$$\text{Pred}\{\underline{X}_E\}[j] = \sum_{j=-L}^L \alpha_j \underline{X}_E[k-j] \Rightarrow \underline{X}_O^{\text{New}}[k] = \underline{X}_O[k] - \sum_{j=-L}^L \alpha_j \underline{X}_E[k-j]$$

ונכל לבחור מקדים לבניית חיזוי אופטימלי ע"י מינימיזציה מהצורה

$$\underset{\alpha_{-L}, \dots, \alpha_L}{\text{Minimize}} \sum_{k=1}^{N/2} \left(\underline{X}_O[k] - \sum_{j=-L}^L \alpha_j \underline{X}_E[k-j] \right)^2$$

דוגמה 11.2: ניקח את הסדרה הבאה ונפעיל עליה את התהליך שתואר, כשנnia כי החיזוי ממצע שכנים (ובקצוות חיסוי פשוט של שכן יחיד).

$$\underline{X} = [1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8 \ 9 \ 10 \ 11 \ 12 \ 13 \ 14 \ 15 \ 16] = P_G(0)$$

ה הפרדה לאיברים הזוגיים והאי-זוגיים תיתן

$$P_G(1) = [2 \ 4 \ 6 \ 8 \ 10 \ 12 \ 14 \ 16] = \underline{X}_E$$

$$\underline{X}_O = [1 \ 3 \ 5 \ 7 \ 9 \ 11 \ 13 \ 15]$$

לאחר החיזוי נקבל

$$\text{Pred}\{\underline{X}_E\} = [2 \ 3 \ 5 \ 7 \ 9 \ 11 \ 13 \ 15]$$

$$\underline{P}_W(0) = [-1 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0]$$

הפרדת $P_G(1)$ לזוגיים ואי-זוגיים תיתן באופן דומה

$$P_G(2) = [4 \ 8 \ 12 \ 16] \Rightarrow \underline{P}_W(1) = [-2 \ 0 \ 0 \ 0]$$

ובאופן דומה

$$\begin{aligned}\underline{P}_G(3) &= [8 \quad 16] \Rightarrow \underline{P}_W(2) = [-4 \quad 0] \\ \underline{P}_G(4) &= [8] \Rightarrow \underline{P}_W(3) = [-8] \\ \underline{P}_W(4) &= [8]\end{aligned}$$

ולכן פירמידת המוצא תהיה

$$\begin{aligned}\underline{P}_W(0) &= [-1 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0 \quad 0] \\ \underline{P}_W(1) &= [-2 \quad 0 \quad 0 \quad 0] \\ \underline{P}_W(2) &= [-4 \quad 0] \\ \underline{P}_W(3) &= [-8] \\ \underline{P}_W(4) &= [8]\end{aligned}$$

וזהו תוצאה טובה יותר מאשר שמספרית האיברים הם אפסים. למעשה, גם את בעיית הקצה השמאלי בוקטוריו שגיאת החיזוי ניתן לפחות ע"י שימוש בשני שכנים וביצוע אקסטרופולציה, ושל פשوطה הוקטור ממנו יצאו, ניתן לקבל כי כל השכבות כוללות אפסים למעט שתי השכבות האחרונות. בוקטורוי מידע בעלי תוכן מורכב יותרanno צפויים לקבל תוכן שונה מאפס.

כל שתואר עד כה עדין אינו סכימת Lifting Scheme המלאה. לצורך השלמתה מוצע שלב

$$\begin{aligned}\text{חישוב נוסף בו לאחר חישוב העדכון } \underline{X}_O^{\text{New}} &= \underline{X}_O - \text{Pred}\{\underline{X}_E\} \\ \underline{X}_E^{\text{New}} &= \underline{X}_E + \text{Update}\{\underline{X}_O^{\text{New}}\}\end{aligned}$$

הרעיון בסיסי צעד נוסף זה הוא להטיב את צורת האות \underline{X}_E לפני סיום החישוב כך שייהי חלק יותר וככזה בעל פוטנציאל דחיסות טוב יותר לפחות הבא. זהו בעצם המקור לשם Lifting כי פעולה זו היא שהופכת את המערכת להתרמת Wavelet. ציור 11.12 מראה את הסכימה המלאה.

אם נניח כי פעולות Predict ו-Expand הין ליניארית, נוכל לתאר את פעולה כמטריצות P ו- U המכפילות וקטוריים. במקרה זה נקבל כי הסכימה המוצעת ניתנת לרישום ע"י

$$\begin{aligned}\underline{X}_O^{\text{New}} &= \underline{X}_O - P\underline{X}_E \\ \underline{X}_E^{\text{New}} &= \underline{X}_E + U\underline{X}_O^{\text{New}} = \underline{X}_E + U\underline{X}_O - UP\underline{X}_E\end{aligned}$$

או ברישום נוח יותר

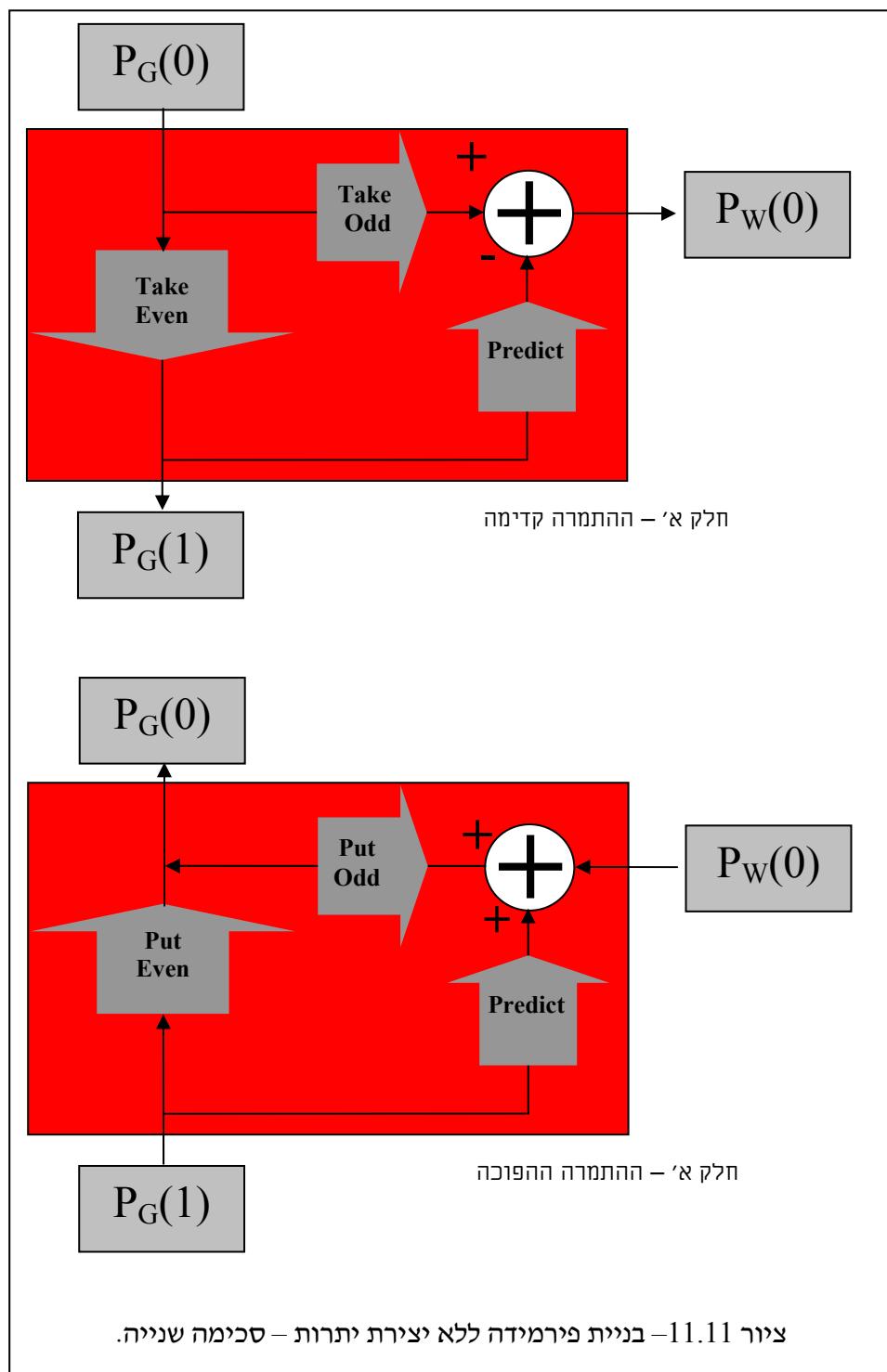
$$\cdot \begin{bmatrix} \underline{X}_O^{\text{New}} \\ \underline{X}_E^{\text{New}} \end{bmatrix} = \begin{bmatrix} I & -P \\ U & I - UP \end{bmatrix} \begin{bmatrix} \underline{X}_O \\ \underline{X}_E \end{bmatrix}$$

דוגמה 11.3: נחבר $I = 0.5I$ ו- $P = 0.5I$ ונקבל

$$\cdot \begin{bmatrix} \underline{X}_O^{\text{New}} \\ \underline{X}_E^{\text{New}} \end{bmatrix} = \begin{bmatrix} I & -P \\ U & I - UP \end{bmatrix} \begin{bmatrix} \underline{X}_O \\ \underline{X}_E \end{bmatrix} = \begin{bmatrix} I & -I \\ 0.5I & 0.5I \end{bmatrix} \begin{bmatrix} \underline{X}_O \\ \underline{X}_E \end{bmatrix}$$

פירוש הדבר הוא שモץ האיזוגים יהיה הפרשים של צמדים סמוכים, ומוצא הזוגים יהיה ממוצעים.

אם נכפול את $\underline{X}_E^{\text{New}}$ ב- 2 נקבל כי זהו בדיקת השלב הראשון בהתרמת Haar.



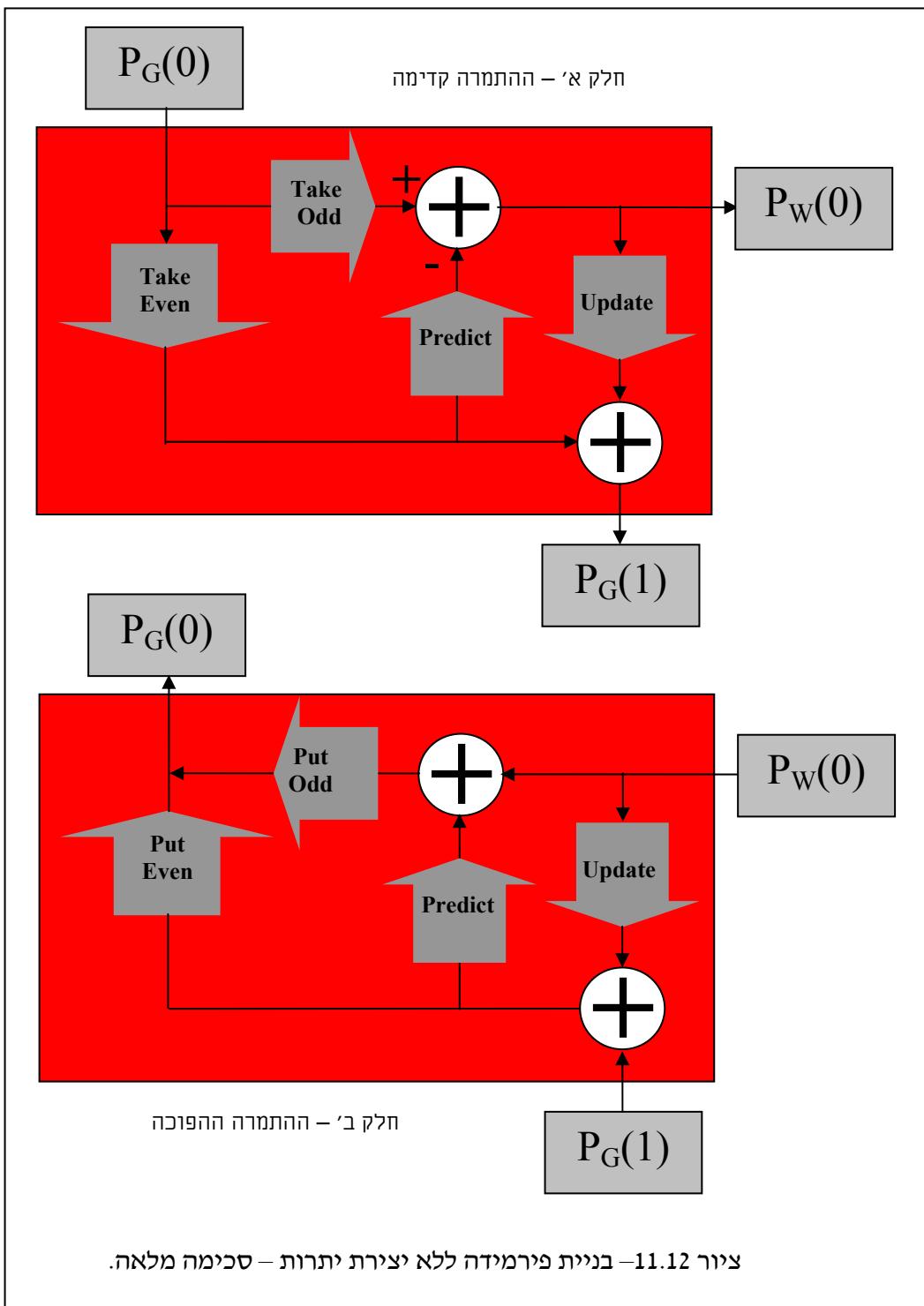
ברור כי נוכל לבחר פעולות מורכבות יותר ל- P ו- U וכן ליצור התמורות Wavelet שונות. יש לשים לב כי אם היעד הוא התהמלה אורתונורמלית, נוצר קשר בין בחירת P ו- U אותו יש לקיים, ומסתבר כי אז תכונו הצמד אינו קל. אבל, חשוב להבהיר כי לכל בחירה שרירותית של P ו- U , אף כי תוביל להטהמלה שאינה יוניטרית, עדין נוכל לטעון כי יש התהמלה הפוכה פשוטה שנייתנת לרישום כ:

פרק 11 - מבנים פירמידליים ו שימושיהם בתמונות

$$\begin{aligned}\underline{X}_O^{\text{New}} &= \underline{X}_O - P \underline{X}_E \Rightarrow \underline{X}_O = \underline{X}_O^{\text{New}} + P \underline{X}_E \\ \underline{X}_E^{\text{New}} &= \underline{X}_E + U \underline{X}_O - U P \underline{X}_E \\ \Rightarrow \underline{X}_E^{\text{New}} &= \underline{X}_E - U P \underline{X}_E + U \underline{X}_O^{\text{New}} + U P \underline{X}_E \Rightarrow \underline{X}_E = \underline{X}_E^{\text{New}} - U \underline{X}_O^{\text{New}} \\ \Rightarrow \underline{X}_O &= \underline{X}_O^{\text{New}} + P \underline{X}_E^{\text{New}} - U \underline{X}_O^{\text{New}}\end{aligned}$$

או ברישום נוח יותר

$$\cdot \begin{bmatrix} \underline{X}_O \\ \underline{X}_E \end{bmatrix} = \begin{bmatrix} I - PU & P \\ -U & I \end{bmatrix} \begin{bmatrix} \underline{X}_O^{\text{New}} \\ \underline{X}_E^{\text{New}} \end{bmatrix}$$



פרק 11 - מבנים פירמידליים ו שימושיהם בתמונות

מכיוון שפעולות Update נועדה להחליק את וקטור הזוגיים הנותר, נוכל להציג מציאות מסנן ליניארי אופטימלי להשגת יעד זה. מהמשווה $\underline{X}_E^{\text{New}} = \underline{X}_E + U \underline{X}_O^{\text{New}}$ ניכר כי אף פעולה Update ה-
ניתנת לרשום כצירוף ליניארי

$$\text{Update} \left\{ \underline{X}_O^{\text{New}} \right\}_k = \sum_{j=-L}^L \beta_j \underline{x}_O^{\text{New}} [k-j]$$

הרי שניתנו לאמץ גם רישום מטריצי בו S היא מטריצת הזזה בדוגימה אחת ימינה, ולכתוב

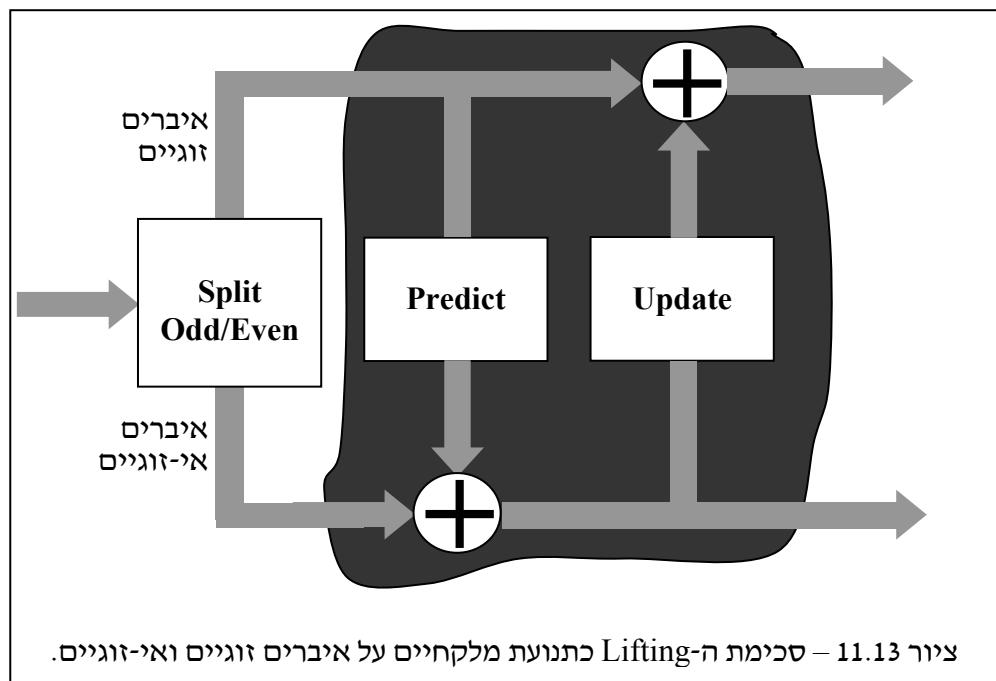
$$U \underline{X}_O^{\text{New}} = \begin{bmatrix} \vdots & \vdots & \vdots \\ S^{-L} \underline{X}_O^{\text{New}} & \dots & \underline{X}_O^{\text{New}} & \dots & S^L \underline{X}_O^{\text{New}} \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} \beta_{-L} \\ \beta_0 \\ \beta_L \end{bmatrix} = M_X \underline{\beta}$$

אז, נוכל להציג קביעת מקדמי המسان $\underline{\beta}$ ע"י קביעת חלקות מרבית על התוצאה של פעולה Update-
תוֹץ שימוש באופרטור גזירה D ולקבל

$$\underset{\underline{\beta}}{\text{Minimize}} \| D(M_X \underline{\beta} + \underline{X}_E) \|_2^2 \Rightarrow \underline{\beta}_{\text{opt}} = [D^T M_X^T M_X D]^{-1} D^T M_X^T \underline{X}_E$$

מסתבר כי סכימת ה- Lifting ניתנת לשכלול עוד ועוד ע"י הוספת פעולות Predict ו-
Update כמפורט בציור 11.13. כל פעולה כזו הפיכה, ולכן קיימת ההתרמה ההפוכה. עם זאת,
לא ברור כיצד יש לתכנן פעולות אלו לביצועים מיטביים. בהינתן התמרת Wavelet ניתן לגנות את
פעולות ה-P וה-U אשר תבנהו לאוטו יעד, ומימוש ההתרמה בפועל ע"י Lifting נחשב יעיל יותר.

כנקודה נוספת נזכיר כי פעולות ה- Predict וה- Update לא חייבות להיות ליניאריות – דבר זה
פותח פתח להתרמות חדשות ומעניינות – נשא זה עדין בחיתוליו ויש בו פוטנציאל עצום.



פרק 12 – מבוא לתורת הצלב

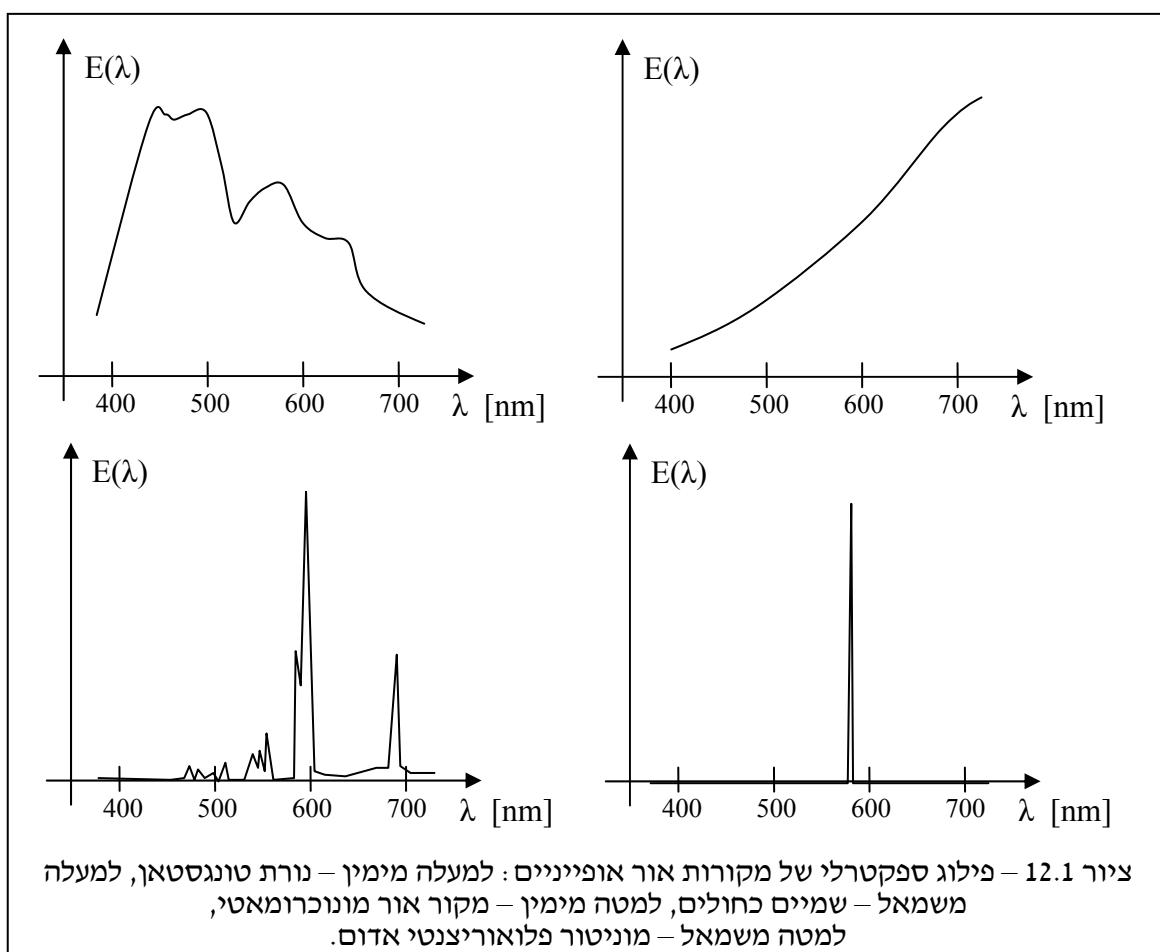
פרק 12 – מבוא לתורת הצלב	
257	12.1 צבע - מהפיזיקה לעין האנושית
258	12.2 התאמות צבע
261	12.3 שיטות ייצוג צבע בתמונות
265	12.4 ייצוג תמונות צבע
271	12.5 שיפור ושחזור תמונות צבע
276	12.6 דחיסת תמונות צבע
283	

רוב הדיוון בספר זה מתרכו בטיפול בתמונות Gray-Scale בהן כל פיקסל כולל ערך סקלרי. מן הבדיקה המתמטית, הכללת הנאמר לתמונות וקטוריות בהן כל פיקסל כולל וקטור ערכאים אינה פשוטה ואנייה חד-משמעות. ברור שניתן לטפל במידע תומני זה כאוסף בלתי תלוי של תמונות ולבצע פעולות על כל תמונה כפי שלמדנו – האם זה מספק? ברוב המקרים לא! אם כך, איך ניקח בחשבון את הקשרים בין הערכאים השונים בוקטורי הפיקסלים?

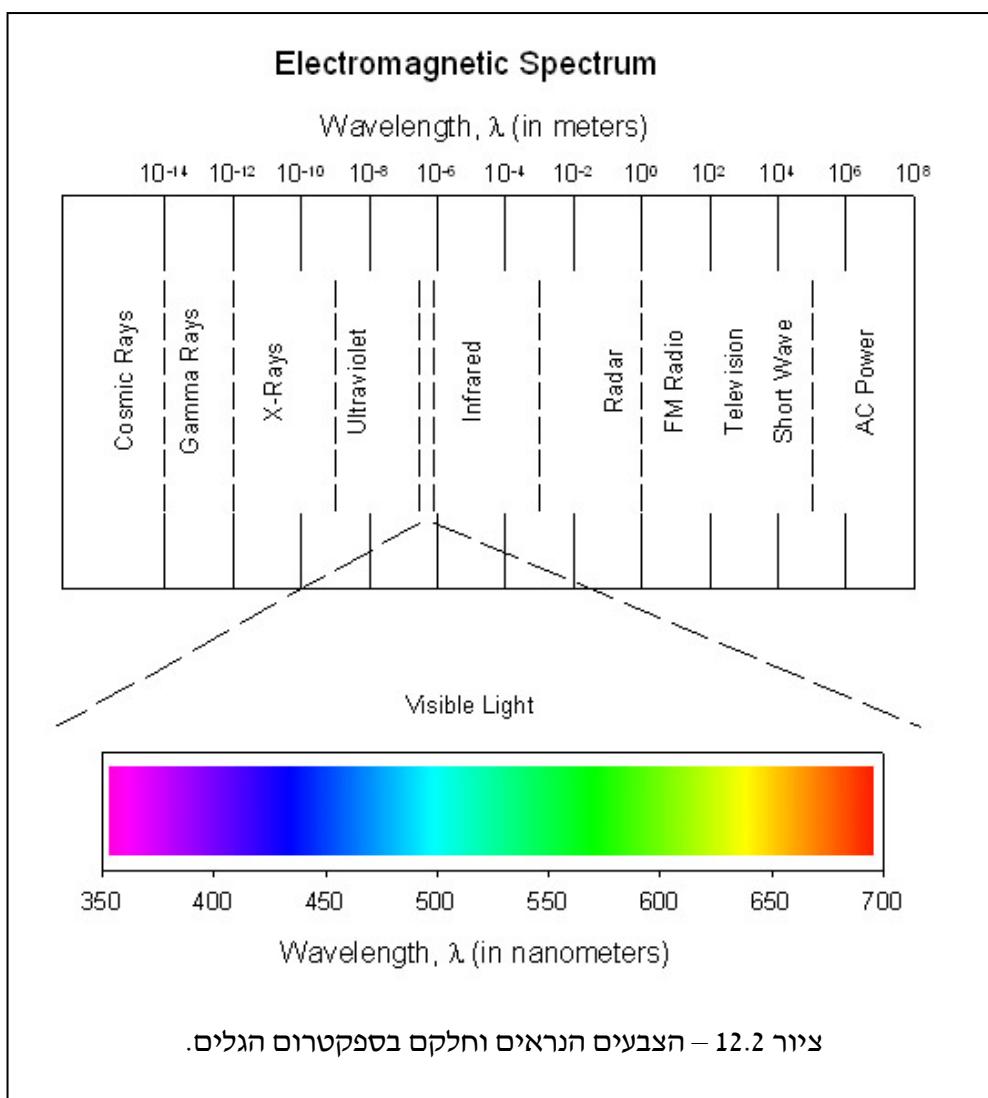
כאשר וקטור ערכי הפיקסל הינו שכבות הצלב R, G, ו-B, מעבר למורכבות שבטיפול בוקטורי ערכאים, נכנס מרכיב חדש למשחק והוא מערכת הראייה האנושית. השלשה RGB אינה מקבץ זמן של ערכאים – יש קשר ביניהם וקשר זה יש להגדיר לפי הדרך בה מערכת הראייה האנושית ממזגת שכבות אלו. ואmens, הטיפול בצלב בתמונות מורכב, ורבית העוסקים בעיבוד תמונה משתמשים בגישות אד-הוקיות ליישומים המופעלים על תמונות צבע. תורה הצלב היא עניין מורכב מכיוון שהיא דורשת עירוב של ידע בפיזיקה (לשם תיאור היוצרות הצלב מגלים אלקטромגנטיים), הבנה במערכת הראייה האנושית (על מנת להבין כיצד הצלב נמדד עין), ויכולת בעיבוד תמונות הבאה מעל כל אלה. בשל כך, שאלות רבות בעניין הטיפול בתמונות צבע עומדות עדיין פתוחות. אנו נציג בפרק זה את היסודות הבסיסיים בתחום, ובעזרת אלו נציג הכללה לפרקי היסוד בעיבוד תמונה – ייצוג, שיפור, שחזור, ודחיסת של תמונות צבע.

12.1 צבע – מהפיזיקה לעין האנושית

היווצרה של תמונה בעין האנושית מתחילה במקור אוור המאיר את היעד הנצפה. אוור זה הינו גל אלקטרומגנטי שנitin לטיור כצירוף לניארי של אורך גל שונים ובעוצמות שונות. אורך הגל מחושב כמוינה בין מהירות האור ותדר הגל - $f/c = \lambda$. ניתן של גל זה ייתן לנו את פילוגו הספקטרלי ($E(\lambda)$) – מעין גרע של העוצמות בכל אורך גל. הפילוג הספקטרלי מלמד הרבה (לא הכל!) ישנו עניין הפaza שהה לא עסוק) על אופי הגל. לשם דוגמה, בציור 12.1 מוצגים פילוגים ספקטרליים (מקורבים) של מספר מקורות אוור אופייניים.



גל מונוכרומטי הוא גל בעל אנרגיה באורך גל יחיד (כמו פונקציה הלם כפונקציה של λ). אם ניצור מקור אור כזה עם אורך גל יחיד ונitin לשיליטה ונציג אותו ישירות לצופה אנושי, יראה הצופה צבעים שונים כפונקציה של אורך הגל. ציור 12.2 מראה את הצבעים המתקבלים מניסוי זה. מעניין לציין שגלי בעלי אורך גל של פחות מ- 350nm או מעל 750nm אינם כלל. התוחום [350nm, 750nm] הוא התחום הנראה. בציור 12.2 אנו גם רואים מיהם הגלים שמחוץ לתחום זה, כשהשכנים הם התוחום האינפרא-אדום בಗלים ארוכים והאולטרה סגול בಗלים קצריים. לדוגמה, גל באורך של 480nm יראה כחול, גל באורך של 570nm יראה צהוב, וגל באורך של 610nm יראה כתום.



ציור 12.2 – הצבעים הנראים וחלוקם בספקטרום הגלים.

בשלב שני של הייצורות התמונה פוגע גל האור באובייקטים שונים, כשהלכל חפש ולכל חומר יש תכונות בליעה שונות של הגל. הבליעה תהיה פונקציה של אורך הגל, ופונקציה של הזווית בין כיוון מקור האור ובין המשטח עליו מאיריים. כתוצאה מהמפגש בין גל האור ונקודה על פני האובייקט המצלום ייבלו חלק מהגל וחלק יוחזר בפלוג ספקטורי חדש שנთנו ע"י

$$E_{\text{out}}(\lambda) = E_{\text{in}}(\lambda)R(\lambda)\cos\theta$$

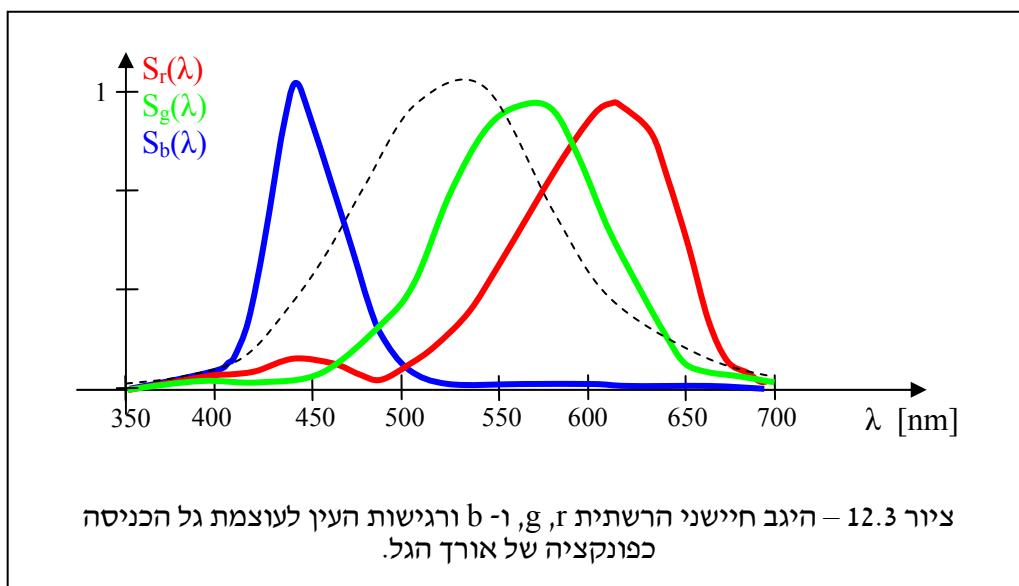
כשהפונקציה $(\lambda) R$ נובעת מתכונות הבליעה של החומר ו- θ היא הזווית בין מקור האור לפני שטח האובייקט. כך למשל, אם הזווית היא 90° , מוחזרת אנרגיה מירבית משטח האובייקט. הכנסת הגורם $\cos\theta$ נובעת מהנחה כי האובייקט בעל פני שטח למברטאים (בניגוד למראה למשול).

בשלב שלישי ואחרו בשלבי הייצור של התמונה משתמשת מערכת הראייה האנושית - אל העין מגיע הגל $(\lambda) E_{\text{out}}$ לאחר הפגעה בחפצים הנצפים. היה זה נפלא לו בראשית העין היו חישנים המבצעים ניתוח ספקטורי מלא של הגל הנכנס ורואים אל מולא המידע המתאר את הגל, אך בפועל אין זה כך. כבר באמצע המאה ה- 19 הסיקו מדענים וביניהם Maxwell ו- Helmholtz כי חישת הצבע נעשית ע"י שלושה

חישנים החסים צבעי יסוד שונים. נסמן חישנים אלו ב- R, G, ו- B. לכל אחד מחישנים אלו עוקמת היגב לאורכי הגל S_r, S_g, S_b , ומוצא חישנים אלו נתון לפי

$$R = \int_{\lambda} S_r(\lambda) E(\lambda) d\lambda \quad G = \int_{\lambda} S_g(\lambda) E(\lambda) d\lambda \quad B = \int_{\lambda} S_b(\lambda) E(\lambda) d\lambda$$

כאשר $E(\lambda)$ הוא הגל הנופל על הרשנית. פעולה זו מבצעת שלושה ממוצעים ממושקלים של הפילוג הספקטורי, כשפונקציות היגב כפי שנמדדו בניסויים נראים בציור 12.3. יש לזכור כי בהתיחסותנו לגל זה אנו חושבים על נקודה בודדה בתמונה – בפועל יש להתייחס לגל זה עם כפונקציה של המקום (y,x), כך שבכל נקודה ברשנית פועל המודל הנ"ל עם גל $E(\lambda)$ פעל פילוג שונה.



התפיסה המוצעת לפיה שלוש אינטגרציות ממוצעות יוצרות שלושה צבעי יסוד קרויות (תיאורית שלושת הצבעים) והיא רווחת במגוון תחומים בהםណן צבע וחישתו. TriChromatic Theory

גרף נוסף שמקובל להציג בדיון על תכונות מערכת הראייה הוא גרף רגישות העין לעוצמת גל האור כפונקציה של אורך הגל. גם גרף זה מוצג בציור 12.3 וכך מוקוו. גרף זה פירשו שams גל מונוכרומטי בעל אנרגיה אחת יוזן ישירות לעין, העין תבחן בו חזק יותר סביב הירוקים וחלש עד כדי אי-ראייה בקצוות התחום הנראה.

מתוך התיאור שניתן ברור כי ניתן אינטגרט פילוגים ספקטוריים עבור אותה שלשה של R, G, ו- B שנמדדים ע"י החישנים, כמובן, ככלומר, מידע אובד מעבר למפלוג ספקטורי מלא לשולחה ערכים מייצגים. הבחירה בשם Blue, Green, Red היא בגל שלושת מרכיבים אלו הם צבעים מונוכרומטיים, אלא בכלל שהגון הכללי אותו הם חשים (מרכז הבודד של פונקציית ההיגב) נופלת באופן כללי על צבעים אלו.

נוח להציג את האינטגרציה הנ"ל בייצוג דיסקרטי ע"י דגימה מספיק צפופה של ציר ה- λ בה יש N דגימות. וקטור העמודה E הוא וקטור ערכי העוצמה כפונקציה של אורך הגל באורך N, והוקטורים

הם רגישיות התאים השונים בחישת ה- R, G, ו- B. אנו מנהים כי וקטוריים אלו בעלי סכום מינורמל,

$$\underline{S}_r \underline{1} = \underline{S}_g \underline{1} = \underline{S}_b \underline{1} = 1$$

נגיד את המטריצה \underline{S} כמיוזג שלושת וקטוריים אלה למטריצה בגודל 3 על N (מספר נקודות הדגימה על הציר λ). אז

$$R = \underline{S}_r \underline{E}, \quad G = \underline{S}_g \underline{E}, \quad B = \underline{S}_b \underline{E} \Rightarrow \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} - & \underline{S}_r & - \\ - & \underline{S}_g & - \\ - & \underline{S}_b & - \end{bmatrix} \underline{E} = \underline{S} \underline{E} = \underline{C}$$

מהנאמר לעלה ברור כי בהינתן \underline{C} הכוללת שלושת ערכי ה- RGB, ישנו אינסוף אפשרויות ל- \underline{E} שיניבו צבע זה, כזו מערכת של 3 משוואות ב- N גורמים.

על רקע התיאור הנ"ל ברור מדוע טבעי לייצג תמונה ספרטית שלושת RGB בכל פיקסל. גם מרבית המצלמות הדיגיטליות מפעילות שלושה סוגים חישניים באופן דומה אם כי עם עוקומות היגב שונות מעט. דרך השגת חישניים שונים היא פשוטה – משתמשים בחישניים בעלי אפיון היגב דומה, אך יוצקים על שטח פנים מסיכת צבע המסתננת חלק מאורכי הגל, וכך מעצבים את ההיגב השונה שלהם ל- R, G, ו- B.

אין כל מניעה ליצור מצלמות בהן היגב לאורכי הגל שונה בתכלית מזה של מערכת הראייה האנושית. בפרט, ניתן לבנות חישניים ש- "יראו" גלים אורך מ- 700nm ואלו הן מצלמות האינפרא-אדום. ישן גם מצלמות בהן במקום שלושה סוגים חישניים יש יותר (עד כמה מאות), כשמטרתן להתקדם לחישת הפילוג הספקטורי במלואו, ומסתבר כי חישה זו חשובה ביישומים של חישה מרוחק (תצלום לווני) ובצלום רפואי של רקמות ועוד. תחום זה קרי לא במקרה "חישה מולטי-ספקטRELית".

12.2 התאמות צבע

המרות צבע נחוצות אם ישנו אי-תיאום בין חישת הצבע במערכת אחת והציגתה במערכת אחרת. נתחיל את הדיון בחישת צבע ע"י מצלמה וביצוע התאמאה לשם הזנת הצבע הנמדד לעין אנושית.

נניח כי בצלמה דיגיטלית נתונים לנו שלושה סוגים חישניים בעלי מטריצת חישה \underline{P} (מקביל למטריצה \underline{S} אותה הגדרנו עבור מערכת הראייה האנושית אך עם עוקומות היגב שונות), ואת מוצא המצלמה אנו מזינים למשך כך שהינתן שלושת RGB המשך מציגה בצורה נאמנה לנכפה במערכת הראייה האנושית (כמילימטרות – בהינתן שלושת RGB פולט המשך אור בעל פילוג ספקטורי כך שכשוויכפל ב- \underline{S} ניתן שוב את אותה שלשה).

נרצה כמובן לבצע תיקון על ערכי ה- RGB הנמדדים במצלמה לפני העברתם למסך, לשם מעבר מערכי החישה שהתקבלו ב- \underline{P} (במצלמה) לאלו שיתקבלו ב- \underline{S} (בעין). עבור פילוג ספקטורי \underline{E} שהוצע למצלמה, כל שיש לנו הוא סט של שלושה ערכים שייקראו $\underline{C}_{\text{camera}} = \underline{P}\underline{E}$ ורצונו בסט אחר שייקראו $\underline{C}_{\text{Eye}}$. נגידר מטריצה M בגודל 3 על 3 שתמיר את האחד לאחר רנץ' לקבוע את זהותה. לו היה נמדד ישירות, הצביע בעין היה $\underline{C}_{\text{Eye}} = S\underline{E}$. בפועל צבע זה יקורב ע"י $\underline{C}_{\text{Eye}} \approx M\underline{P}\underline{E}$. לכן, מטריצה M שתביא למינימום את הביטוי

$$f\{M\} = \int_{\underline{E} \in \Omega} \|\underline{M}\underline{P}\underline{E} - \underline{S}\underline{E}\|_2^2 d\underline{E}$$

הרעיון הוא למצוא M שתיהיה נאמנה לכל המקבץ האפשרי של פילוג הספקטורים הכניטה \underline{E} . גזירת ביטוי זה ליחס ל- M נותנת

$$\frac{\partial f\{M\}}{\partial M} = \int_{\underline{E} \in \Omega} (\underline{M}\underline{P} - \underline{S}) \underline{E} \underline{E}^T \mathbf{P}^T d\underline{E} = 0 = \underline{M}\underline{P} \left[\int_{\underline{E} \in \Omega} \underline{E} \underline{E}^T d\underline{E} \right] \mathbf{P}^T - \underline{S} \left[\int_{\underline{E} \in \Omega} \underline{E} \underline{E}^T d\underline{E} \right] \mathbf{P}^T$$

נגידר את האינטגרל **בביטוי הנילכ** – $\int_{\underline{E} \in \Omega} \underline{E} \underline{E}^T d\underline{E} = \mathbf{R}$ ונקבל

$$M_{\text{Opt}} = \mathbf{S} \mathbf{R} \mathbf{P}^T [\mathbf{P} \mathbf{R} \mathbf{P}^T]^{-1}$$

המטריצה \mathbf{R} היא מטריצת האוטוקורלציה של \underline{E} . עבור פילוג ערכים אקראי בתחום $[0,1]$ בהסתברות אחידה, ולא קורלציה בין ערכים סמוכים (כמו ברעש לבן!), מטריצה זו תהיה

$$\int_{\underline{E} \in \Omega} \underline{E} \underline{E}^T d\underline{E} = \mathbf{R} = \frac{1}{12} \mathbf{I} + \frac{1}{4} \mathbf{1}$$

דרך החישוב של מטריצה זו פשוט – כל איברי האלכסון הראשי זהים ושוויים ל-

$$\int_0^1 x^2 dx = \frac{x^3}{3} \Big|_0^1 = \frac{1}{3}$$

כל האיברים מחוץ לאלכסון הראשי גם הם זהים ושוויים ל –

$$\int_0^1 \int_0^1 xy dx dy = \int_0^1 x dx \cdot \int_0^1 y dy = \frac{x^2}{2} \Big|_0^1 \cdot \frac{y^2}{2} \Big|_0^1 = \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$$

ניתן גם לחשב את \mathbf{R} בהנחה פילוג אחרית על שדה ההארה \underline{E} , ולקחת בחשבון למשל משקל שונה לפילוגים בהתאם לתמונות בהן אנו עוסקים. ניתן גם לבצע את בניית \mathbf{R} מתוך סט דיסקרטי של הארות צפויות.

בביטוי של M האופטימי-Ano הופכים מטריצה בגודל 3 על 3 אשר תהיה לא סינגולרית אם $\text{rank}\{\mathbf{P}\} = 3$ ואם \mathbf{R} מטריצה חיובית מוגדרת. שתי דרישות אלו מתקיימות ($\text{rank}\{\mathbf{P}\} = 3$ משמעו שלושת סוגי החישנים במצלמה פורסם מרחב תלת ממדי – אם לא – זהו בזבוז של חישנים). במקרה פרטי מעניין, אם $\mathbf{P} = \mathbf{S}$ נקבל כי $I = M$ – כלומר לא נדרש המרה, וזה כמובן צפוי.

בניגוד התאמה אחרת היא בעית התאמת צבע כך שפילוג ספקטורי נתון יוצג בכך בעל מאפיינים מסוימים כך שהעין האנושית תראה את הגוונים הנכונים. עניין זה דומה מעט בניתוחו והוא מוביל בנית עקומות ה- CIE שנזכיר בהמשך.

נניח כי בידינו התקן פלט (למשל, מסך מחשב טלוויזיה) המציג תמונה. דרך הצגת התמונה היא ע"י בחירת מספר (אולוי שלוש ואולוי יותר) צבעי יסוד, בניית צירוף ליניארי שלהם והצגתם. ברור כי התקן פלט זה אינו יכול להציג כל פילוג ספקטורי אפשרי של אור בשל השימוש כמעט完全 צבעי יסוד. עם זאת, אין כל חשיבות להפקת הפילוג הספקטורי המדוקיק – די בהפקת פילוג ספקטורי אשרCSI שיאפשר לעין אනושית הוא ייתן אותה שלשת RGB, ובכך יידאו כי הפקת הצבעים מהתקן הפלט תואמת את ציפויותה של מערכת הראייה האנושית.

ב- 1930 קבעה ועדת בשם CIE (Commission Internationale de L'Eclairage) סטנדרטיזציה של הפקת צבע בהתקני פלט לפי האופן הבא. נבחרו שלושה צבעים מונוכרומטיים כצבעי יסוד מהן מורכב כל צבע צירוף ליניארי. ועדת ה- CIE קבעה צבעי יסוד אלה לאורכי הגל 700nm (אדום), 546.1nm (ירוק) ו- 435.8nm (כחול). נכל ליצץ סידרת פילוגים ספקטוריים אלו כמטריצה P בגודל 3 על N בה בכל שורה יש איבר אחד השווה 1 והוא המיקום אורץ הגל הנאות. לכן, בדומה להנחה שנעשתה קודם, סכום איברי כל שורה מנורמל להיות 1.

בהינתן פילוג צבע אמיתי כלשהו \underline{E} (ווקטור באורך N), נציג צירוף ליניארי של שלושת צבעי ה- P לשם הפקת הצבע בעין – צירוף ליניארי זה יהיה $\underline{\beta}^T P$ – וזהו פילוג ספקטורי חדש. בעוד ששירות העין תחשש את שלשת ה- RGB הנתונה ע"י $\underline{S}\underline{E}$, הרי שע"י הצבעים היסודיים נקבל כי שלשת ה- RGB שהעין תחשש היא $\underline{\beta}^T S P$. נרצה כמובן שתתי שלשות ה- RGB תלכדנה (או תהינה קרובות) לכל צבע \underline{E} אפשרי ע"י בחירת הווקטור $\underline{\beta}$ בדומה נאותה כפונקציה של \underline{E} .

הדרך בה בחרה ועדת ה- CIE לתקן את הבעה היא הבאה: במקומות קבוע פונקציה כללית מהצורה $\underline{E}(\underline{\beta})$, נניח כי את קבועה ה- $\underline{\beta}$ נעשה לצבעים מונוכרומטיים בלבד. לצבע מונוכרומטי \underline{E}_k (דהיינו, '1' במקומות ה- k בווקטור ואפס בכל מקום אחר) נקבל את המשוואה

$$S P^T \underline{\beta}_k = S \underline{E}_k \Rightarrow \underline{\beta}_k = [S P^T]^{-1} S \underline{E}_k$$

וממשוואה זו נחלץ את שלשות המשקלות $\underline{\beta}_k$. ב巡视ת ה- k מ- 1 ועד N נקבל שלוש פונקציות משקל הקróיות Tristimulus Values. פונקציות אלו מייצגות את המשקל היחסית של כל צבע יסוד בנית צבע עד מונוכרומטי בעין. ציור 12.4 מראה את הפונקציות הללו, וכפי שניכר, לצבעים מסוימים נדרש משקל שלילי לצבע היסוד האדום לקבל האפקט הרצוי.

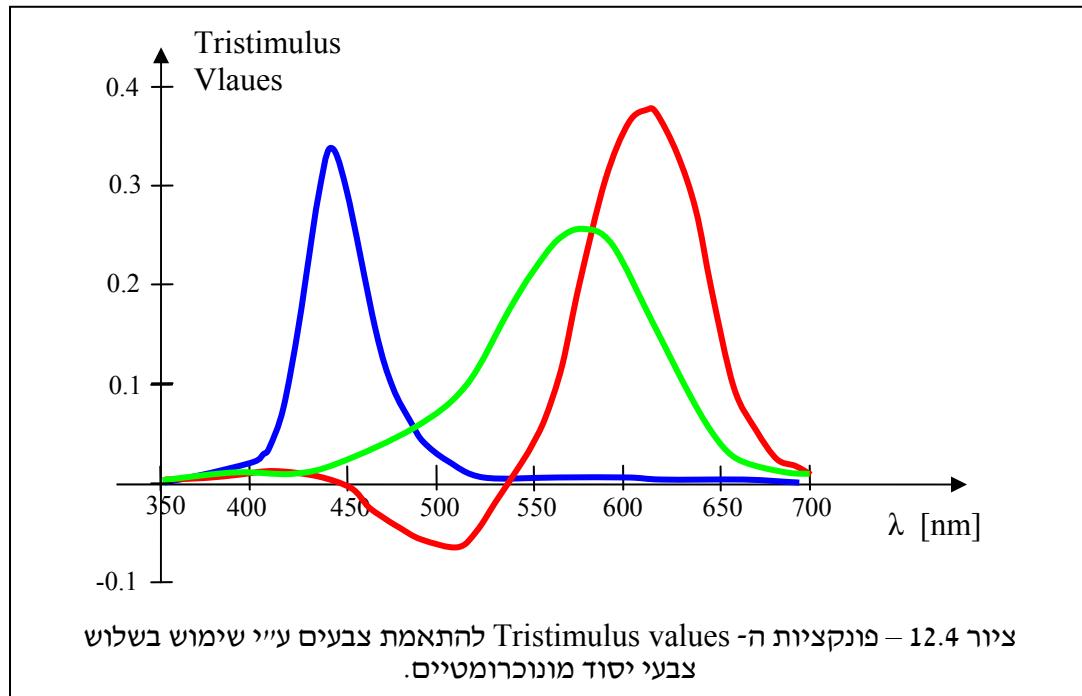
ברישום אחר, ניתן לצרף את כל המשוואות לשווה מטריצית אחת בה הצבעים \underline{E} יملאו עמודותיה של מטריצה E בגודל N על N וזו כמובן תהיה מטריצת היחידה בשל היות כל צבעים אלו מונוכרומטיים עם סריקה מלאה של אורך הגל. אז נקבל כי פונקציות המשקל הנ"ל אינן אלא

$$\left. \begin{array}{l} \mathbf{S}\mathbf{P}^T \underline{\beta}_1 = \mathbf{S}\underline{E}_1 \\ \vdots \\ \mathbf{S}\mathbf{P}^T \underline{\beta}_k = \mathbf{S}\underline{E}_k \\ \vdots \\ \mathbf{S}\mathbf{P}^T \underline{\beta}_N = \mathbf{S}\underline{E}_N \end{array} \right\} \Rightarrow \mathbf{S}\mathbf{P}^T \mathbf{B} = \mathbf{S}\mathbf{E} \Rightarrow \mathbf{B} = [\mathbf{S}\mathbf{P}^T]^{-1} \mathbf{S}\mathbf{E} = [\mathbf{S}\mathbf{P}^T]^{-1} \mathbf{S}$$

בגיעה צבע כללי \underline{E} , הוצע ע"י ה-CIE להתייחס אליו כצירוף ליניארי של אותן מונוכרומטיים, לכל צבע מונוכרומטי להשתמש במשקלות הנאותים, וכך לקבל את סט המשקלות הכללי. בתיאור אלגברי זה אינו אלא אומר כי

$$\begin{aligned} \underline{E}_{\text{New}} &= \sum_{k=1}^N w_k \underline{E}_k \\ \Rightarrow \underline{\beta}_{\text{New}} &= \sum_{k=1}^N w_k \underline{\beta}_k = \sum_{k=1}^N w_k [\mathbf{S}\mathbf{P}^T]^{-1} \mathbf{S}\underline{E}_k = [\mathbf{S}\mathbf{P}^T]^{-1} \mathbf{S} \sum_{k=1}^N w_k \underline{E}_k = [\mathbf{S}\mathbf{P}^T]^{-1} \mathbf{S}\underline{E}_{\text{New}} = \mathbf{B}\underline{E}_{\text{New}} \end{aligned}$$

כלומר, בפשטות, בהינתן צבע חדש יחושבו מקדמי הצירוף הליניארי כממוצע משוקל של משקלות ה- Tristimulus Values.



אם זהה מטריצת ה- \mathbf{B} הטובה ביותר? נציג גישה כללית יותר למציאת \mathbf{B} המיטבית באופן הבא – נניח כי המטריצה \mathbf{B} קבועה ולכל צבע המגיע משתמש בשלוש המשקלות $\underline{\beta}_{\text{New}} = \mathbf{B}\underline{E}_{\text{New}}$. בהגעה הצבע $\underline{E}_{\text{New}}$, העין תחשוש ישירות את השלשה $\mathbf{S}\underline{E}_{\text{New}}$, בעוד שבגישה שלילוב צבעי היסוד המוצע נתקבל

את השלשה $\mathbf{SP}^T \underline{\beta}_{\text{New}} = \mathbf{SP}^T \mathbf{B} \underline{E}_{\text{New}}$. מطبع הדברים, נרצה כי שתי שלשות RGB אלה יתלכדו לכל צבע. לכן נדרש

$$\forall \underline{C}, \quad \mathbf{SP}^T \mathbf{B} \underline{E} = \underline{S} \underline{E} \quad \Rightarrow \quad \mathbf{SP}^T \mathbf{B} = \mathbf{S} \quad \Rightarrow \quad \mathbf{B} = [\mathbf{SP}^T]^{-1} \mathbf{S}$$

ותוצאה זו זהה לוצאה שקיבנו קודם.

אם נחזר לרגע להתחלה, משמעות תוצאה זו היא שבгинון מסך מחשב המיצג צבע ע"י מיזוג של שלושה צבעים יסוד ע"י צירוף ליניארי, ניתן "לرمות" את העין ולספק בכך פילוג ספקטרלי אשר יתאים לצבעו אוטה העין הייתה רואה לו צפחה בגל האור המקורי. אבל, תוצאה זו בעלת ערך תיאורתי כיון שמשקל שלילי יוצר בעיה במימוש. שאלת מיידית שעולה מכך היא השאלה הבאה – האם קיימים סט של צבעים יסוד פיזיקליים (כלומר בעלי ספקטרום אי-שלילי) שיתנו כי \mathbf{B} אי-שלילית אף היא? ננסח זאת כاإוסף הדרישות

$$\forall i, j, \quad \mathbf{B}[i, j] \geq 0, \quad \mathbf{P}[i, j] \geq 0, \quad \mathbf{B} = [\mathbf{SP}^T]^{-1} \mathbf{S}$$

מסתבר כי על מנת לתת תשובה חיובית לשאלת זו נדרשים מספר גדול משלוש של צבעים יסוד (יש לשים לב שזו נוסחת החישוב של \mathbf{B} אינה כמתואר כי המטריצה \mathbf{SP}^T אינה ריבועית, וישן דרגות חופש שבמיוחד לקיומו של פתרון). לעומת זאת, אם נתיר שימוש בצבעים יסוד לא פיזיקליים (כolumbia שיכולים בספקטרום שלהם ערכי אנרגיה שליליים), הדבר אפשרי, ויוצג זה מוביל למפת ה- XYZ-CIE שתשוכר בהמשך. ברור כי שימוש בפתרון מעין זה לא יינתן לשלים בניה בפועל של צבע (שהרי צבע יסוד אינםMRI מימוש), אך חשיבותו מפה זו גודלה לשם הצגת פריסת הצבעים האפשרית.

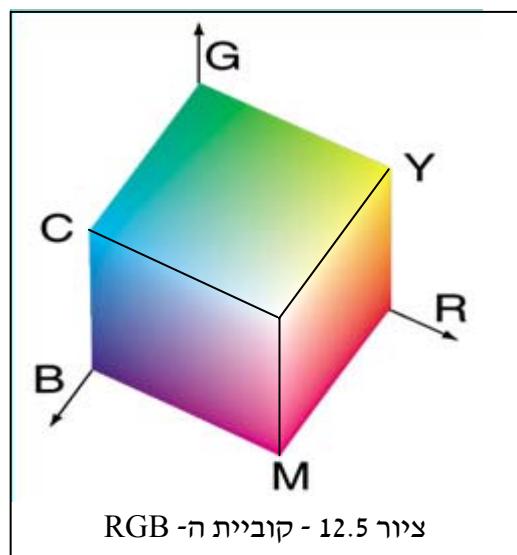
12.3 שיטות ייצוג צבע בתמונות

כאמור, שיטה ייצוג צבע טבעיות לתמונות כשהן במדיה דיגיטלית היא שיטת ה- RGB. אם נניח כי כל שכבת צבע זוכה לייצוג ב- 8 סיביות, ניתן לחשב על שלשה זו כ��דרה קובייה של צבעים, וכל צירוף של 24 סיביות יוצר אחד מתוך כ- 16 מיליון צבעים אפשריים בתוך חלל הקובייה הזה. קובייה ה-RGB מוצגת בציור 11.5. נשים לב כי הצבע [R=255,G=255,B=0] נותן את הצבע צהוב המתואר כפינה אחת של הקובייה. פינות מעניינות אחרות הן [R=255,G=0,B=255] – Magenta, ו-[R=0,G=255,B=255] – Cyan. השילוב [R=0,G=255,B=0] שנוטן את ה- (תכלת). השילוב [R=255,G=255,B=255] נותן לבן. צבעי שילוב אלו חשובים בהקשר של החסרות צבעים אשר ידוע בהמשך בקצרה.

הערה: יש מקום להיות יותר מדויקים, ולציין את הקובייה הניל' כשהיא מעוותת. הרווחים שנבע מהציוויל הניל' הוא צבעי היסוד R, G, ו- B אורתונורמלים זה לזה – ובכך חבוייה הנחיה על חוסר תלות ביניהם, ואין זה כלל כך. אם נניח למשל של שלשה זו היא כאלו המתקבלים במערכות הראיה האנושית, הרי שמצויר 12.3 ברור לנו כי מכפלה פנימית בין פונקציות המשקל שתוארו תיתן קורלציה כלשהי בין היצירם שיצרנו. דרך הציור הנכונה יותר תביא לפוליטופ (מצולע תלת-ממדי) שהינו עיוות של הקובייה כך שהזווית בין הצבעים תהיה חדה יותר מ- 90°. פוליטופ זה יהיה במצב פירמידה (או יותר נכון – קונו – קונו)

שראשה בראשית, כיוון שם שלושת RGB קיימת ונראית, הרי שכל הכפלת שלה בקבוע ייתן צבע נראה גם כן.

ניתן גם להסתכל על קובייה זו בדרך שונה, כפי שמציע ציור 11.5, ולמהות קו דמיוני בין נקודות הראשית [0,0,0] ובין [255,255,255]. לכל אורך ישר זה נקבל צבע אפור בהירות משתנית משוחרר מוחלט לבן מוחלט. הצבע הלבן בגל מתיחס לפילוג קרוב לאחד של כל אורכי הגל (להזיכרכם, כשהאנו משתמשים במושג "לבן" באפיון רעש למשל, אנו מתכוונים כי פילוגו בתדר אחד), ועבור פילוג זה R=G=B (כלומר, השטח מתחת הגראפים בציור 12.3 שווה בשלושת העוקמים, כפי שכבר הנחנו).



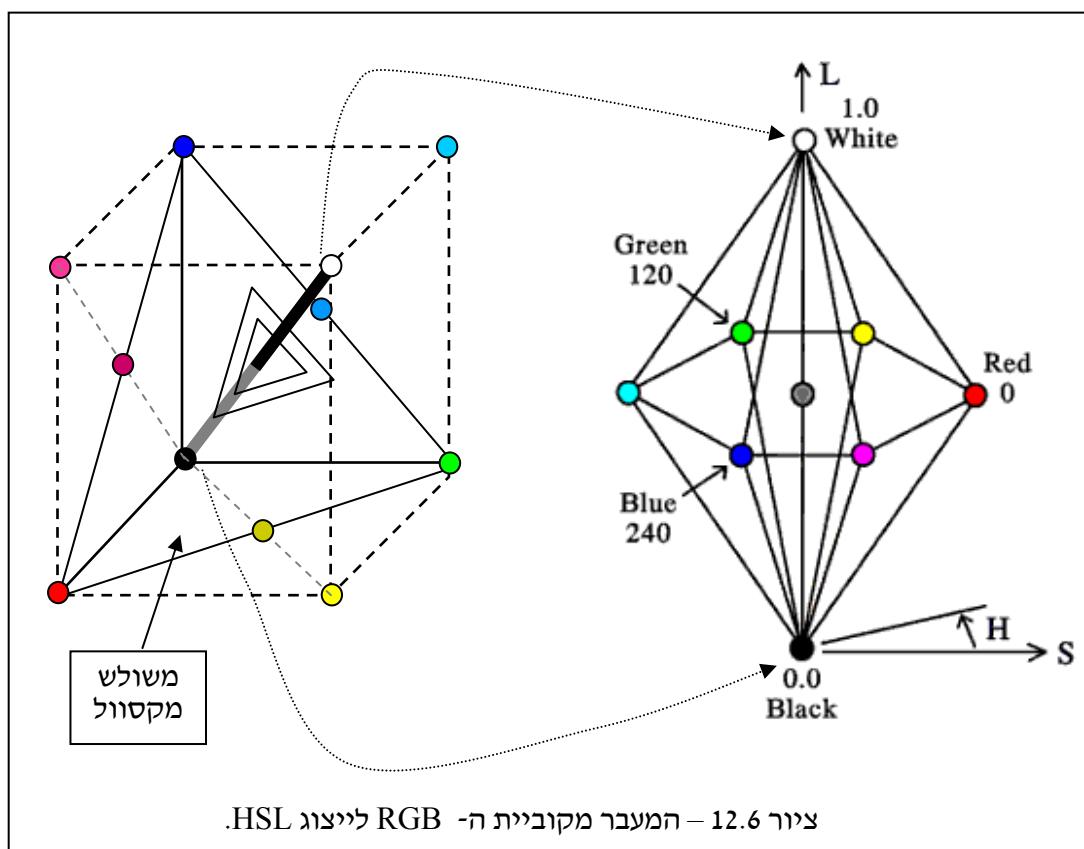
משתמע לכך כי אם נרצה להמיר תמונה צבע לתרומות גווני אפור, נחליף כל שלושת ערכים [R,G,B] בנקודה הקרויה ביוטר על האלכסון שהוצאה את הקובייה משוחרר לבן, וברור כי זהו פשוט מיצוע $(R+G+B)/3$. אבל נוכל לעשות יותר – במקרה שימוש ב-[R,G,B] ניתן לייצג את הצבע ע"י שלושת ערכים חדשות שתיתן הפרדה טבעיות בין בהירות לתוכנות הצבע. כפי שהוצע קודם, המיקום לאורך ציר הבבירות (נקרא Brightness או Intensity) יהיה הערך הראשון בשלשה החדש. בקביעת הבבירות התמקמו על החतך המתואר כמשולש אשר הינו אכן לקו הציר שחזור-לבן – משולש שנקרא משולש מקסול.

על פני משולש מקסול, מיקום ייצג את הצבע. דרך אחת לעשות זאת היא קביעת המרחק בערך מוחלט מקו הציר (נקודת המרכז של המשולש) עדות לעומקו של הצבע (נקרא saturation). בנוסף, הכוון אליו פנינו – או במילים אחרות – אורך הגל הדומיננטי – יעד על הגוון (קרי Hue). כך נקבל ייצוג חדש לצבע הקרי HSL (או HSI) והוא מתואר בציור 12.6. ישנן מספר דרכים לבצע את החישוב האנליטי העובר משלשת RGB ל-HSL, אך אנו נוטר על התענג שבhzגת המשוואות שמיישמות מעבר זה בפועל.

הרעיון להפריד בין בהירות לפירוטי הצבע מפתחה וחוזר על עצמו במגוון ייצוגים אלטרנטיביים כגון YIQ, CIE-XYZ, La*b*. בחרה פופולרית מבין ייצוגי הצבע קרויה YIQ. יציג זה הוצע ע"י ועדת שעסקה בייצוג צבע לשידורי טלוויזיה (National Television Systems Committee - NTSC) וייצוג זה הוא המישם באותות טלוויזיוניים. מתוך רצון להציג המירה פשוטה מ- RGB (צבעי היסוד של CIE ולא אלו של מערכת הראייה) ל-YIQ הוצאה ההמרה היליניארית הבאה:

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.523 & 0.312 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

ביצוג זה, מרכיב ה- Y הוא המקביל בהירות והוא מוצע כממוצע משוקל של R, G, ו- B עם דגש על הירוק – זהו האות לפיו תיראה התמונה בטלוויזיה שחורה לבן. שני המרכיבים האחרים מאפיינים צבע שהבחנה אינה כמקודם – גוון ועומק, אלא שני צירים כמעט אונקיים זה לזה שפורסם את התנועה על פני שטח משולש מקסול שוויי בהירות שהגדנו בציור 12.6. מטריצה זו הפיכה ולכון ההתמרה חוזרת פשוטה אף היא. הערכים במטריצה באים לשקף מחד את שלושת צבעי היסוד שנתונים במסך הטלוויזיה, ומайдך את הרצון לשובב את צירוי קוביית ה- RGB בזווית מסוימת לקבלת הפרדה טובה בין צבע ובין בהירות.

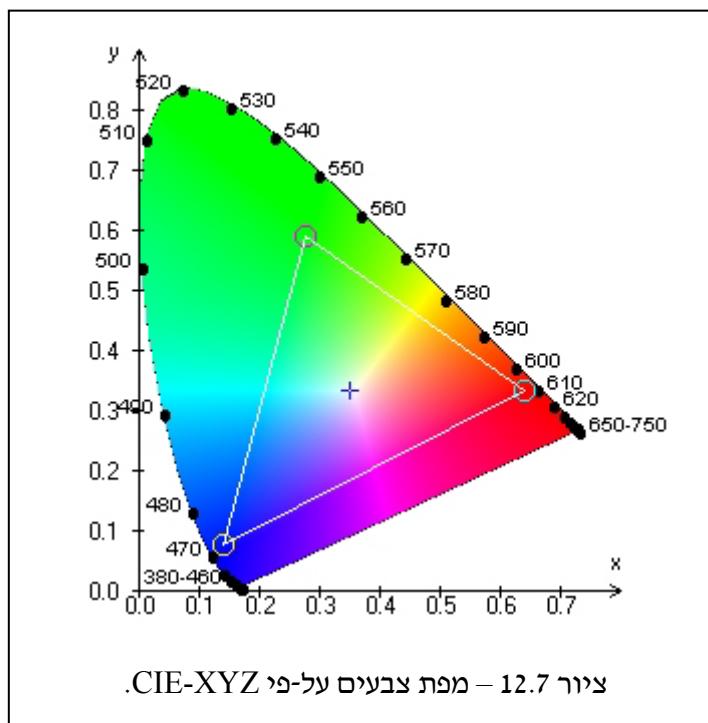


ייצוג אחר שהזכיר הוא ה- CIE-XYZ והוא כבר הוזכר קודם. הרעיון הוא הבא: אם אנו מותרים על העניין בהירות, הרי שנרצה להתרמז בנקודות על פני המשולש לו קראננו משולש מקסול.

משולש זה מאופיין במשוואה $R+G+B=Const$. בהינתן שלושת RGB, די בשני ערכים על מנת לתאר את המיקום של הצבע,

$$X = \frac{R}{R+G+B}; Y = \frac{G}{R+G+B}; Z = \frac{B}{R+G+B} \Rightarrow X+Y+Z=1$$

כנגד כל צמד ערכי (X,Y), Z ידוע ומיותר, ולכל צמד כזה במישור נוכל ליחס צבע על כל מאפיינו (גונו בסיסי ועומק הצבע). נוכל לייצר מפת צבעים באופן זה, ומפה מעין זו מתוארת בציור 12.7. חשוב לציין כי בניית מפה זו תלויה בבחירה צבעי היסוד R, G, ו- B שמהם התחלנו, ובבנייהתה מעורבת המטריצה B שהציגנו קודם על מנת לעבור לחישות צבע כפי שהעיננו תיתן. כפי שכבר רأינו, שימוש צבעי היסוד של CIE (המוניוכרומטיים) לשם השגת צבעים מסוימים מוביל למשקל שלילי. המפה שמתוארת כאן מציעה שילובים של ערכי X ו- Y חיוביים לשם קבלת כל צבע בספקטרום הצבעים, ועובדת זו נשענת על שימוש צבעי יסוד (B,G,R) שאינם פיזיקליים, כפי שהבהירנו קודם.



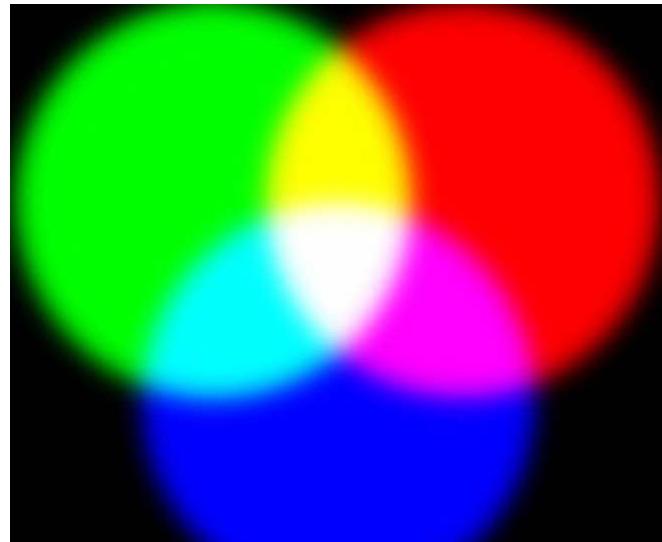
בעיה שמתעוררת עם מפה זו היא שמרחקים שוים על-פניה אינם משקפים מרחקי צבע שוים ביחס למערכת הראייה האנושית. דרך אחת להציג סוגיה זו היא הצגת אליפסות שווות "חישה" סביבן נקודות הפזרות על פני מפה זו. מסתבר כי בחלק הימני עיון של המפה נדרש גדול יותר על פני המפה לשם קבלת שינוי שווה בעיני האנושית. עניין זה חשוב למדדימערכות בהן מרחקי צבעים ממשימים (וכמעט כל נושא שיותג בהמשך נופל לקטגוריה זו). ייצוג b^*La הוצע כיעות המפה הנילן כך שמרחקים אוקלידיים פשוטים על-פניה יהיו מרחקים הוגנים.

כושא אחרון במסגרת הדיון בייצוגי צבע, ניגע בנושא חיבור וחיסור של צבעים. מבלתי לומר זאת במפורש הנחנו עד כה כי צביעת נקודה אחת בשני צבעים מהווה את חיבורם בייצוג RGB, כלומר חיבור מרכיביהם באופן אלגברי פשוט. כמו כן, כשמייצגנו שני צבעי יסוד התקבל צבע ביניים כקומבינציה ליניארית פשוטה. תפיסה זו מתאימה למסך בו נקודת המוצא היא צבע שחור, והכנסת אור מקרבת את

גונו לבן. בגישה זו, חיבור צבעים כמתואר בציור 12.8 טבעי ונכון. לשם דוגמה, אנו רואים כי כאשר מצרפים את הירוק לאדום מתקיים צהוב, ועם הוספת הכחול יתקבל לבן.

לעומת זאת, כאשרנו מדפיסים על דף נייר, נקודת המוצא הפוכה – זהו הצבע הלבן – וכל הוספת צבע מקרבתו לשחור. אם כך, בהינתן שלושת RGB אוטה רצוננו להפיק על גבי נייר, באיזה צבע יסוד נשתמש וכי怎 נזגם? מוצע להשתמש בצבעים המשלימים אותם כבר פגשנו בתיאור קוביית ה-RGB בציור 12.4. הצבעים המשלימים הם Cyan, Magenta, ו-Yellow. ואכן השם CMY. השם 'משלימים' נובע מהעובדת שהוספת צבע משלים לבן יוצר בפועל את הצבע המקורי – למשל, בהינתן דף לבן ועליו מודפס גוון ה-Magenta, אם נקבע מעליו בצבע שני Cyan נקבל צבע כחול (ראה ציור 12.9).

פירוש הדבר הוא שבעובדת עם קוביית ה-RGB علينا להתייחס לראשית צבע הלבן וממנו אנו פועלים לבניית השחור, בעוד שבגישה המקורית ראיינו את השחור בראשית וממנו בשילובי צבעים התקרבו לבן.

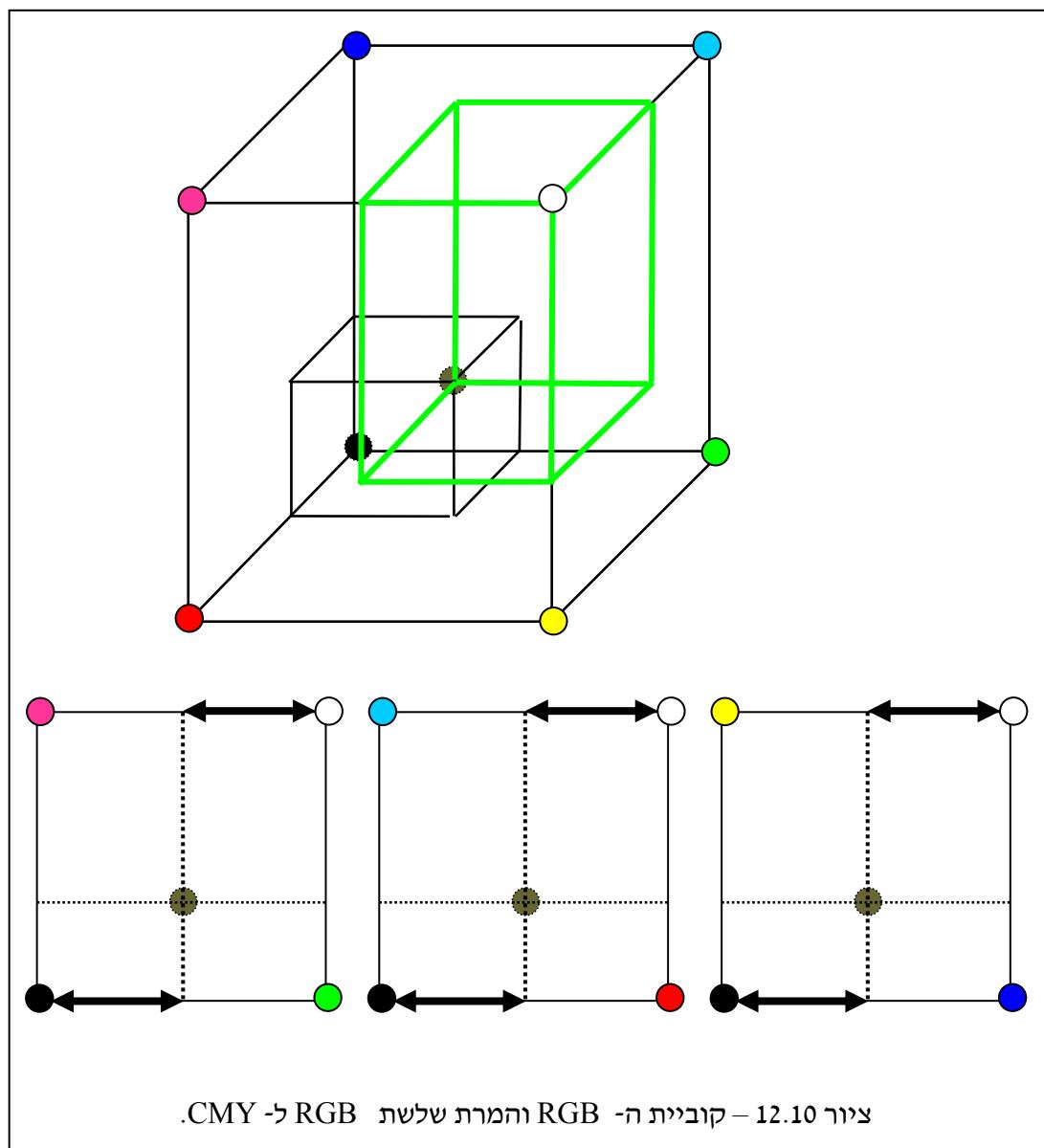
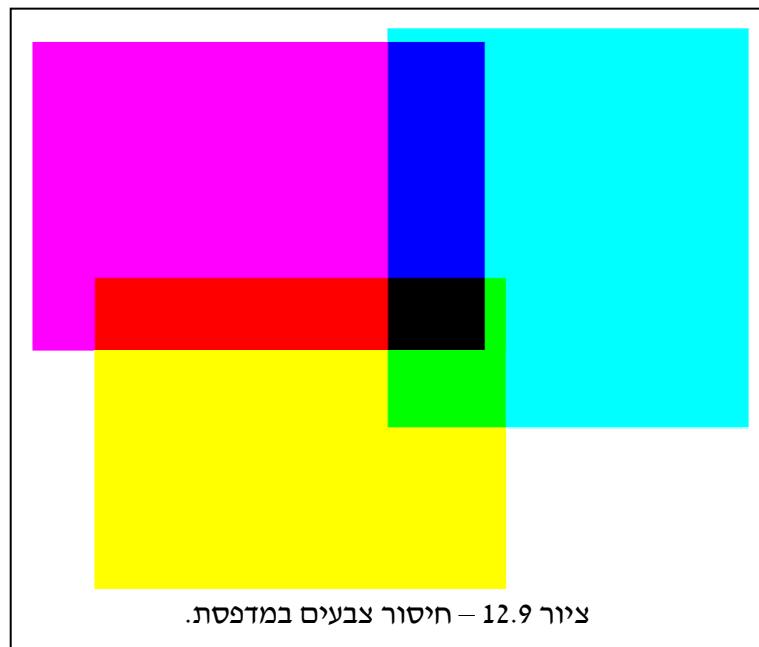


ציור 12.8 – חיבור צבעים במסך.

נניח כי נתונה לנו השלשה [R,G,B] המיוצגים ב- 8 סיביות, ורצוננו להדפסה על דף נייר לבן. לו היה זה מוניטור, היינו מצרפים את הגוונים R באדום, G בירוק ו-B בכחול שבמסך והגון הרצוי היה מתUEL. בהדפסה הגישה תהיה שונה. אם נציג את נקודת היעד בתוך קוביית ה-RGB, علينا למזוג את הצבעים המשלימים CMY לפיה

$$, C = 255 - B, \quad M = 255 - G, \quad Y = 255 - R$$

ותהlik זה מוסבר בציור 12.10. לדוגמה, כאשרנו נעים מנקודת הלבן לעבר נקודת ה-Magenta לשם קביעת הנקודות הנחוצה של צבע זה, אנו נעים בפועל בקו מקביל לקו הנע מהשחור לירוק. لكن מידת ה-Magenta הנחוצה היא מלאה הלבן פחות הירוק שבצבע הנתון.

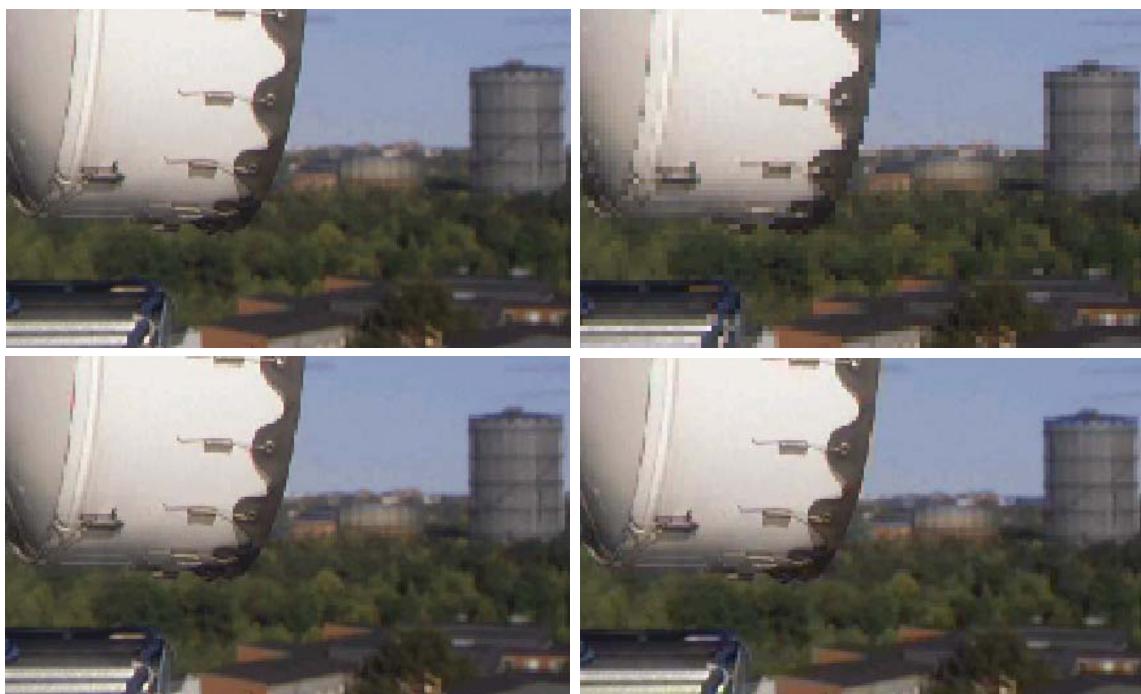


4.12.4 ייצוג תМОונות צבע

בדיון ביצוג תМОונות עוסקו בשני נושאים מרכזיים – דגימה וקונטיזציה. נחזר לשני נושאים אלו בקצרה ונראה מה השוני הנובע מהטיפול בצבע.

דגימה של תМОונת צבע יכולה לקבל ניתוח זהה לניתוח שנטנו לתМОונת שחור-לבן עם הבדיקה כי יש לבצע את תהליך הדגימה לכל סקלר ביצוג הוקטורי של הצבע. אם התМОונה מיוצגת ב-RGB, לרוב נגלה כי מרוח הדגימה הנחוץ קרוב מאוד ואף זהה בשלוש שכבות התМОונה.

לעומת זאת, אם הייצוג כולל הפרדה של בהירות וגוון נקלט לרוב כי מרוח הדגימה בהירות נדרש להיות קטן יותר, בעוד שבשכבות הגוון המרוח יכול להיות גם יותר. תופעה זו נובעת מתכונות מערכת הראייה האנושית בה אין רזולוציה טובה לצבע בעוד שנדרשת רזולוציה גבוהה בהירות. ציור 12.11 ממחיש תופעה זו ע"י הציגת תМОונה (כל הczęgot הימן בסופו של דבר ב-RGB) שיוצגה ב-YIQ, שבהachat התМОונות כל השכבות נתונות באותה רמת פירוט, באחת התМОונות ישנו דילול או אינטראפולדציה ב-Y (פעולת הדילול והאינטראפולדציה יוצרות בלוק של 2 על 2 או 4 על 4 פיקסלים בעלי אותו ערך) עם השארת I ו-Q ללא קלוקול, ותМОונה שלישית נתונה כשבכבה ה-Y שומרה כבמוך בעוד שכבות ה-I וה-Q עברו דילול או אינטראפולדציה.



ציור 12.11 – השפעת דילול או אינטראפולדציה על חdot תМОונת צבע.

למעלה משמאל – תМОונת המקור.

למטה מימין – תМОונת המקור עם דילול או אינטראפולדציה ביחס 1:2 בכל ציר לשכבה ה-Y.
למטה משמאל – תМОונת המקור עם דילול או אינטראפולדציה ביחס 1:2 בכל ציר לשכבות ה-I וה-Q.
למטה מימין – תМОונת המקור עם דילול או אינטראפולדציה ביחס 1:4 בכל ציר לשכבות ה-I וה-Q.

אנו רואים כי פגיעה (ואפילו קלה) בשכבות הביהירות ניכרת מיד. לעומת זאת, אבדן $\frac{3}{4}$ משכבות ה- I וה- Q נותרת אבדן זעום שקשה לאבחנה (בסק-הכל חצי ממידע התמונה הושלך כמעט ללא נזק). כאשר הדילול בשכבות הכרומה הינו 1:4 בכל ציר (דהיינו, יותר 1/16 מכלל המידע בהן) הנזק מתחילה להיות מוגרש, אך עדין לא באותה עצמה בה נראה הנזק שהסבירו - Y עם דילול פחות אגרסיבי. בדיקת תוכנה זו היא שameda לנגד עיניהם של חברי ועדת ה- NTSC כשבחרו בייצוג QI. מטרתם העיקרית הייתה חסכו ברווח פס לשידור, ועם המעבר ל- QI השידור כולל פס רגיל לשכבות ה- Y ופסי תזרר זרים הרבה יותר לשתי שכבות הכרומה. אנו נראה כי תוכנה זו תשוב ותשמש אותנו בבואנו לבצע שחזור או דחיסה של תמונות צבע.

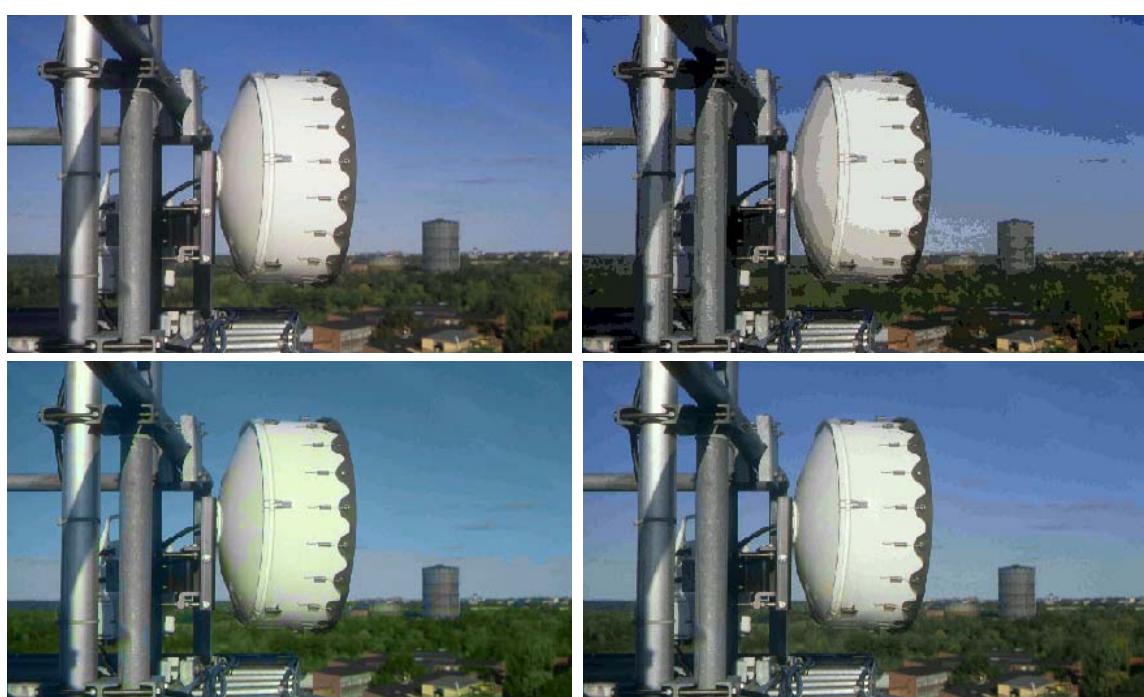
קוונטיזציה של תמונות צבע יכולה להיעשות ע"י קוונטיזציה של כל אחת מהשכבות הסקלריות בנפרד. בפרט, כבר הזכרנו כי בייצוג RGB מקובל לתת 8 סיביות לכל שכבה – דבר המוביל להיווצרן של 16 מיליון צבעים אפשריים בייצוג. אם נתבונן בקוביות ה- RGB נראה קוונטיזציה זו בקוביות שוות נפח אשר גבולותיהן מקבילות לצירם כשריג קרטזי אחד. בגישה ספרטילית זו, הורדת כמהות הסיביות באחת השכבות תביא לרובה לפגיעה מוגשת, וכך אנו תקווים על 24 סיביות לפחות.

האם באמת לא נוכל לרדת מ- 16 מיליון צבעים? התשובה היא כמובן לא אם הגישה ממשיכה להיות ספרטילית ב- RGB ואם נדרש התאמה לכל כל התמונות. לעומת זאת, אם נרצה עירוב הצירים או בניית קוונטיזציה תלויות תמונה, נוכל בהחלט לחסוך.

עירוב הצירים משמעו שתאי הקוונטיזציה בקוביות ה- RGB יהיו בעלי צורה כללית יותר, ואת זאת ניתן להשיג למשל ע"י עיות הקובייה ע"י מעבר לייצוג אחר (כגון QI) וביצוע של קוונטיזציה איחידה על הצירים החדשניים עם בחירת פחות סיביות לאזורים פחות רגילים. למשל, בייצוג QI נידרש ל- 8 סיביות ל- Y, אך ברוח הדיון הקודם נוכל לפגוע בשכבות ה- I וה- Q ביצוגן ב- 6 סיביות בלבד. ציר 12.12 מראה תוצאות ניסוי מעין זה בו רואים כי אם יש להפחית כמהות הסיביות, הרבה פחות נזק נגרם אם הפחתה זו נעשית על שכבות הצלב.

הערה: בקונטיזציה של שכבות ה- QI בניסוי הנ"ל נעשה שימוש בקונטיזר יונייפורי ממורכב. לשם יישום קוונטיזר כזה יש לקבוע את ערך המינימום והמקסימום. בהנחה ששכבות ה- RGB נתונות בתחום $[0,1]$, לשכבת ה- Y הטווח יהיה $[0,1]$, לשכבת ה- I הטווח יהיה $[-0.6,0.6]$ ולשכבת ה- Q הטווח יהיה $[-0.53,0.53]$. ערכי מינימום ומקסימום אלה מתקיים בפשטות מתוך מטריצת ההמרה.

למעשה, ניתן ללבת הרבה יותר ורחוק אם מוכנים להכניס את המרכיב השני שהוזכר לעלה, הוא עניין התלות בתמונה. בגישה הקודמת מוצעת קוונטיזציה איחידה, ואם ניתן לפילוג הערכים בתמונה להשפיע, הרי שנוכל לתקן קוונטיזר שאינו איחידי, וכבר כזאת ניצור אדפטציה של תהליך הקונטיזציה לתמונה אותה אנו מעבדים. ברור כי בשמירת תמונה זו יש לשמור מידע נלווה על מדרגו הקונטיזרים השונים לשם בניית התמונה מערבי האינדקסים שלה.



צ'ור 12.12 – השפעת קוונטיזציה שונה בשכבות YIQ.

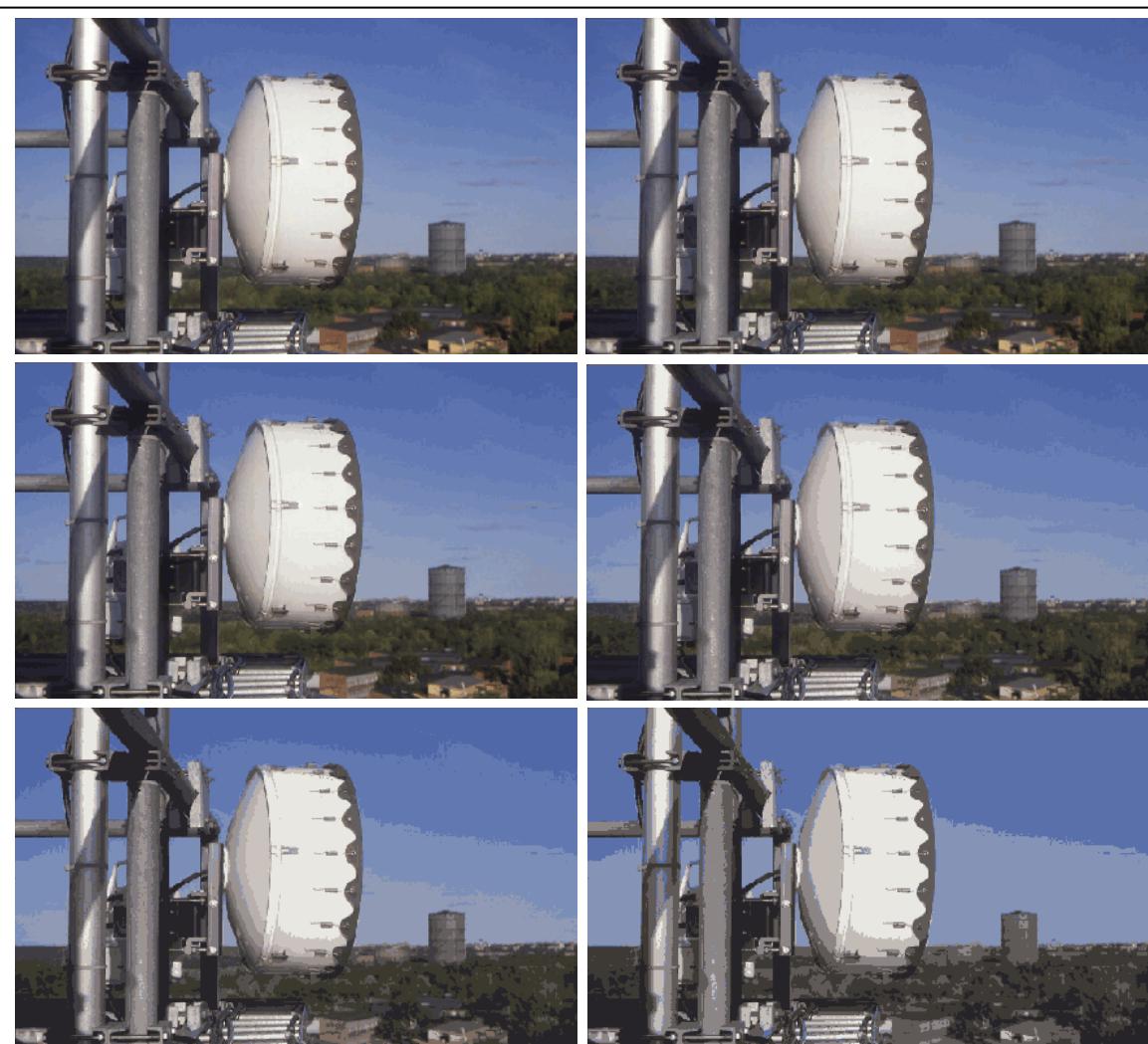
- למעלה משמאל – הקצתה 6 סיביות לכל שכבה ב- YIQ (סה"כ – 18 סיביות לפיקסל).
- למעלה מימין – הקצתה 6 סיביות ל- I ו- Q ו- 3 סיביות ל- Y (סה"כ – 15 סיביות לפיקסל).
- למטה משמאל – הקצתה 6 סיביות ל- Y ו- 4 סיביות ל- I ו- Q (סה"כ – 14 סיביות לפיקסל).
- למטה מימין – הקצתה 5 סיביות לכל שכבה (סה"כ – 15 סיביות לפיקסל).

גישה כללית יותר להורדת כמות הצלבים היא גישה שתיפרד מהתפיסה הספרטנית גם בייצוג האלטרנטיבי ותפעל ע"י אישכול (Clustering). אם נחזור לרגע לייצוג התמונה ב- RGB, נוכל לבנות היסטוגרמאות תלת-מימדיות בה לכל תא [R,G,B] יש מונה הסופר את מספר הופעות של שלשה זו. לתמונה נתונה ניתן לצפות כי יהיו צבירים – קבוצה גדולה של הופעות של אותו צבע או צבעים קרובים למדוי, ועבורם נוכל לבחור צבע מייצג יחיד.

זו בדיקת הגישה המוביילה לקוונטיזציה וקטוריית שנדונה בפרק 4. אנו נתיחס לאוסף הפיקסלים בקבוצות וקטורים בתלת-מימד, ונחפש מילון ייצוג בעל מספר צבעי ייצוג שייקבע מראש. פירוש הדבר שמבנה התמונה יהפוך למבנה חדש בו יש טבלת צבעים לבניית התמונה שתכלול N צבעים (זהו המילון שנמצא), וכל פיקסל יוכל אינדקס אל טבלת הצבעים על מנת לומר מהו הצבע המאכלס פיקסל זה.

ב- Matlab מבנה זה נקרא `Indexed` והפקודה `rgb2ind` ממירת מתמונה RGB למבנה החדש `Indexed`. הפקודה `image` וסט צבעים `Palette`. הצגת תמונת `Indexed` נעשית ע"י הצגת מערך האינדקסים כתמונה ע"י הפקודה `colormap`, ואז מייפוי הערכים לצבעיהם המתאימים ע"י הפקודה `colormap`. בציור 12.13 מוצגת התמונה שכבר הצינו בציורים קודמים כשהיא מומרת למבנה `Indexed` עם מספר משתנה של צבעים ב- `Palette`. בעוד שבגישה הקודמת נדרשו לכ-14 סיביות לשמר את מראה התמונה, הרי שכآن

אנו רואים כי די ב- 6 סיביות והתוצאה קרובה מאוד. אנו נראה כי גישה זו מובילה לאלטרנטיבת מעניינת בדוחיסת תМОנות צבע.



ציור 12.13 – ייצוג תמונה במבנה Indexed במספר צבעים משתנה. למעלה משמאל – 256, למעלה מימין – 128, אמצע משמאל – 64, אמצע מימין – 32, למטה משמאל – 16, ולמטה מימין – 8 צבעים.

כיצד תיעשה פעולה תכנון המילון במעבר מהתמונות RGB לתמונה Indexed ? גישה אפשרית אחת היא האלגוריתם שהוזג בפרק 4 – זהו האלגוריתם המכפיל את אלגוריתם Max-Lloyd ונקרא - K-Means או בתחום של דחיסת מידע – LBG על שם שלושת יוצרים (Linde, Buso, Gray). עם זאת, ישן אלטרנטיבות פשוטות יותר אשר לא פעם נוטנות תוצאה וייזואלית נעימה יותר. אלגוריתמים כמו Octree, Median-Cut ואחרים מוכרים בספרות כאלטרנטיבות טובות, אך אנו נעצר את הדיוון בנושא זה כאן.

מרחב צבע לבניית ה- PalettePalette מקובל להיות מרחב ה- RGB. הסיבה לכך היא הרצון למדוד מרחקים בין צבעים בצורה הוגנת. כזכור (ראה פרק 4), בניית המילון נעשה שימוש איטרטיבי בצד חישוב מרכז כובד הדורש מיצוע של צבעים, וכך מיפוי הדורש חישוב מרחקים בין צבעים. במרחב RGB שתי פועלות אלה קלות – בהינתן שתי שלשות RGB, מרחק אוקלידי ביןיהן (סכום ריבועי הפרשיים)

מהוות הערכה לא רעה למידת המרחק הויזואלי בין צבעים אלו, ولكن שימוש ב- RGB נראה מתאים. לעומת זאת, למשל, במעבר ל- YIQ, מרחב הצבע מתעוזות ויחסוב מרחק אוקלידי אינו הוגן. מסתבר כי נושא זה של מרחקים אוקלידיים ונאמנותם לפער צבעים בעין מהותי ומושך עניין רב. בספרות ישנו דיוון עמוק בסוגיה זו והוצעו מרחב צבע כגון $*La^*$ שנועדו לפתור בעיה זו בדיק.

אם כבר נבנה ה-Palette, נוכל לשפר את התמונה ע"י בחירת אינדקסים בצורה שונה מעט – דבר בו נפגשנו בעבר כשדיברנו על Half-Toning. נניח כי בສירה שורה לאחר שורה קבענו את האינדקסים עד לפיקסל ה-(j,k), וכעת עליינו לקבוע את האינדקס באותה מקום זה. בגישה רגילה נקבע את האינדקס להצבע על הצבע ב- Palette אשר יהיה הקרוב ביותר לצבע בפיקסל זה. בבחירה זו אנו מתעלמים מהבחירה שנעשתה לסייעת הפיקסל, ובכך יכול להיות שאנו מניצחים שגיאת מצטברת. תופעה זו בדיק היא העומדת מאחוריה מראה השמיים בציור 12.12 בבחירה 8 צבעי Palette. נוצר מצב בו נוצר קו תפיר בין שתי רמות צבע – קו שאינו שם במציאות. האלטרנטיבה היא גישה הקרויה Error Diffusion, בה אנו בוחרים אינדקסים מתוך כדי צבירת שגיאה וחתיכחות אליה.

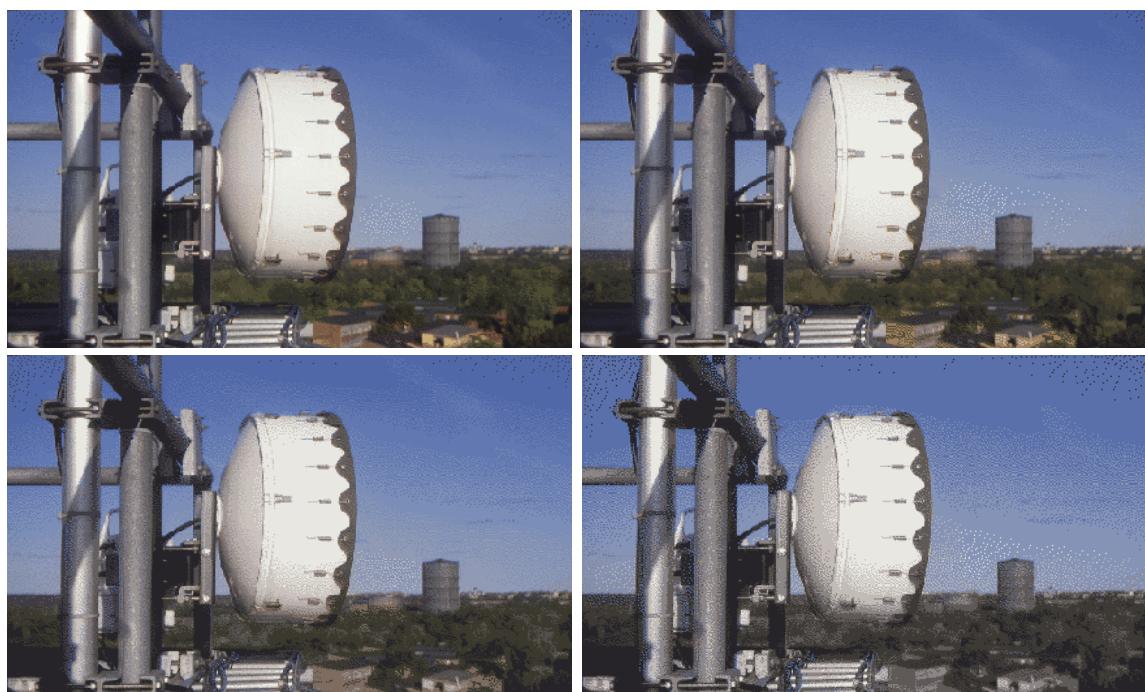
בשיטת ה- Error-Diffusion אנוначיל בכך שנדיר שכנות סיבתיות לפיקסל ה-(j,k) – קבוצת פיקסלים עבורה כבר נעשתה החלטה לגבי האינדקס – נסמן שכנות זו ב- Ψ . נוכל להניח לשם פשוטות כי שכנות זו כוללת את שלושת הפיקסלים מעלה ומשמאלו לפיקסל הנוכחי. בכל פיקסל k בשכנות זו יש שלשה של RGB מקורית שתסומן $\{R_p, G_p, B_p\}$. לכל פיקסל k יש גם את הצבע שנבחר מתוך ה-Palette אשר יסומן כ- $\{\hat{R}_p, \hat{G}_p, \hat{B}_p\}$. השגיאה המומוצעת על-פני שכנות זו נתונה כ- :

$$\Delta R = \sum_{p \in \Psi} R_p - \hat{R}_p \quad \Delta G = \sum_{p \in \Psi} G_p - \hat{G}_p \quad \Delta B = \sum_{p \in \Psi} B_p - \hat{B}_p$$

כעת, בהינתן הפיקסל החדש x בעל השלשה $\{R, G, B\}$, נרצה לבחור את הצבע המתאים לפיקסל זה. אנו נניח לפיקסל זה את הצבע הקרוב ביותר ל- $\{R - \Delta R, G - \Delta G, B - \Delta B\}$ מתוך צבעי ה-Palette. בזורה זו, אם השגיאה היא אפס זהותית, ניתן מפעלים את האלגוריתם הפשוט של בחירת השכן הקרוב מתוך המילון.

כאשר למשל השגיאה באדום שלילית, פירוש הדבר שהגון האדום הכליל בסביבה כהה מהדרוש יש להבהירו. לכן, החסרת שגיאה זו תיתן ערך יעד חדש בהיר יותר מולו נחפש שכן קרוב ביותר ב밀ון. ציור 12.14 מראה 4 תמנונות בהן נעשתה פעולה ה- Error-Diffusion בבחירה האינדקסים. בתמונות אלו יש 16, 32, 64 ו- 8 צבעים, כמו רבייעית התמונות התחתונות בציור 12.13, והשיפור ביחס לגישה הרגילה ניכר.

שיטת ה- Error Diffusion רבת פנים, וניתן להכללה בדרכים שונות. ניתן, למשל, לדרש כי בחירת צבע שאינו השכן הקרוב המיידי, לכפות רק תת-קבוצה של צבעים אפשריים על מנת למנוע קפיצות צבע עצבניות מיידי. שיטה זו רווחת במדפסות צבע בהן יש מספר מחסניות צבע (יותר מ- 3) בבואר להדפיס תמונה (בניגוד לגרפיקה או טקסט, בהם מגנון זה מכובה).



ציור 12.14 – ייצוג תמונה במבנה Indexed במספר צבעים משתנה עם error-diffusion. משמאל – 8 צבעים, מימין – 16, למעלה מימין – 32, ומטה משמאל – 64, למקרה מינימלי.

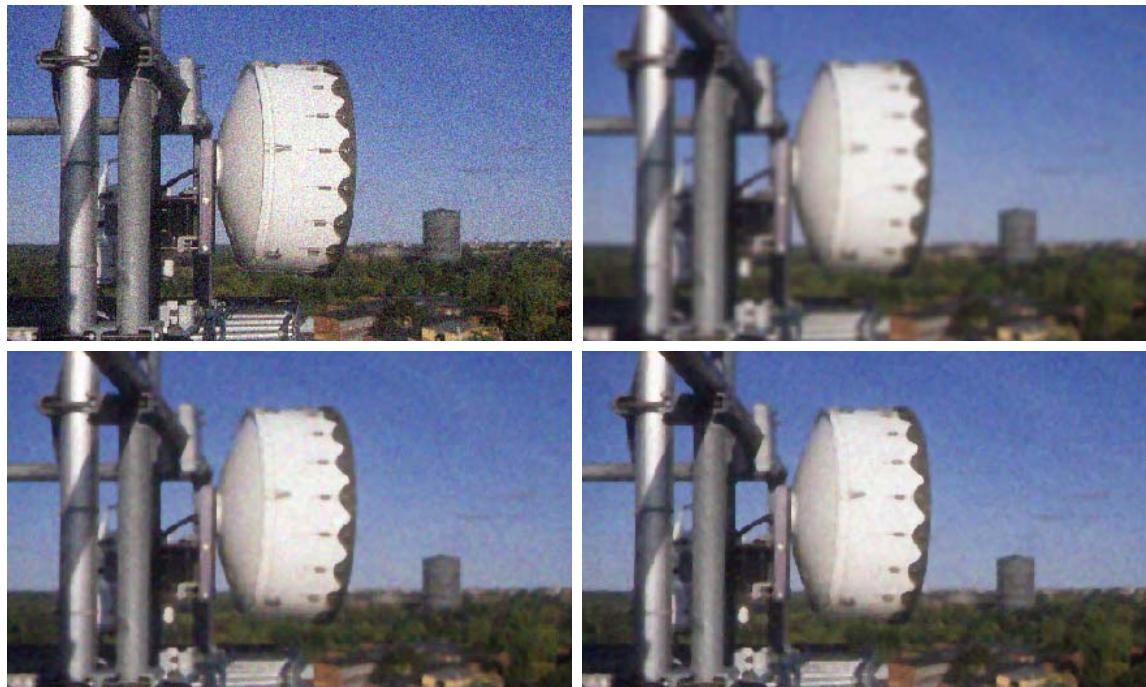
בקשר של תמונות בעלות מבנה Indexed מקובל לעתים קרובות השימוש בפסאודו-צבע. אם נתונה לנו תמונה שאינה מתארת צלום אלא גודל כלשהו (טמפרטורה כפונקציה של המיקום, עומק מים, שדות אלקטромגנטיים ועוד), ניתן ליחס לערכיה צבעים לשם קבלת אפקט ויזואלי חזק יותר. אין כל חוקיות בכך הבחירה, וזאת יתר או מנות מאשר מודע (כמו הרבה דברים בויזואלייזציה). אנו סתפק רק באיזור המושג.

12.5 שיפור ו恢復 תמונה צבע

פעולות היסטוגרמה וסינון של תמונה צבע יכולים להישנות על כל אחת משכבות ה- RGB בנפרד, ובהרבה מקרים כך זה אمنם מופעל. בגישה זו יתכונו עיותני צבע מטורידים. לדוגמה, אם רצוננו לנוקות ורעש מהתמונה ע"י החקיקה פשוטה, ונפעיל מסנן 5 על 5 אחד על התמונה, במערכות המחשבים ימוגו צבעים לייצור צבע חדש שאינו בהכרח קיים בתמונה המקורית. גישה אלטרנטיבית מקובלת היא הפעלת הסינון על תמונה לאחר המרת לייצוג המפריד בהירות ממופיעני צבע.

אם נתייחס לרגע לישום הבסיסי של הורדת רעש מתמונה צבע תוך כדי שמירה על שפות, יוכל להציג שלוש אלטרנטיבות – פועלות סינון ליניארי על שכבות ה- RGB באופן בלתי תלוי, פועלות חציון על שכבות ה- RGB באופן בלתי תלוי, או המרת ל- YIQ, ביצוע חציון על שכבות ה- Y, וטשטוש פשוט על שכבות ה- I וה- Q. תוכנות גישות אלה מוצגות בציור 12.15, ואנו רואים שוב כי הפרדה של בהירות

מצבע משרתת את המטרה של פישוט הפתרון ע"י הקצאת התהיליך המורכב (חץין) לבתיות, והסתפקות בפתרון פשוט (סינון ליניארי) בשכבות הצבע.



ציור 12.15 – תמונה מקור רועשת (למעלה משמאל) ותיקונה ע"י סינון ליניארי בלתי תלוי בשכבות ה- RGB (למעלה מימין), סינון חציון בלתי תלוי לשכבותיה ב- YIQ (למטה משמאל), וסינון בייצוג YIQ בו Y מסונן ע"י חציון ו- IQ ע"י מסנן ליניארי.

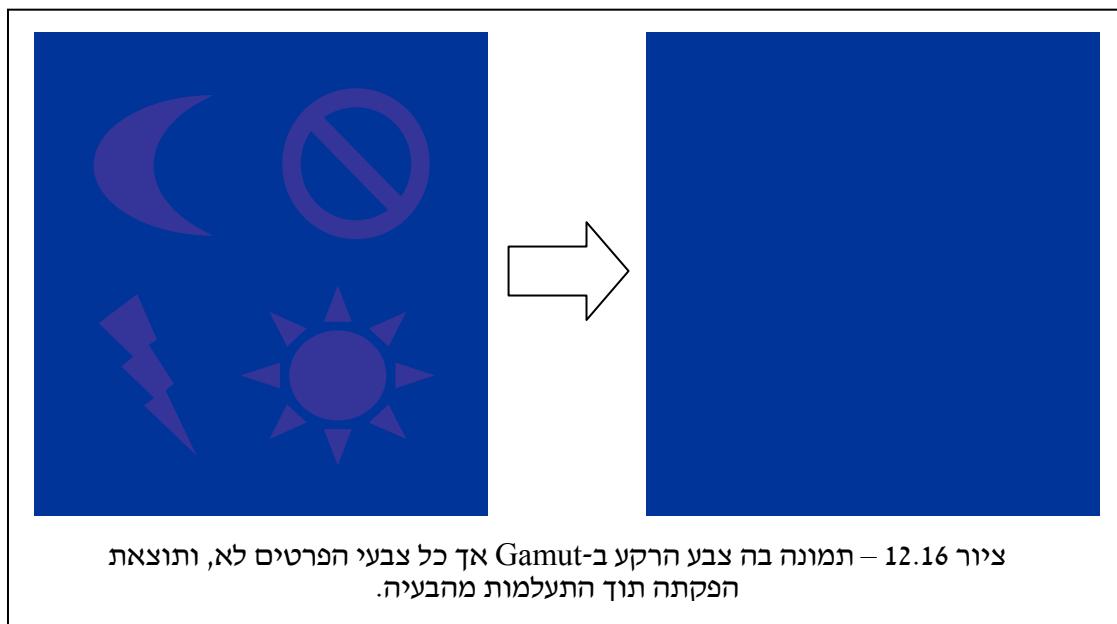
באופן דומה נוכל לטפל בתמונות עם מגוון הכלים שהוצעו בפרק 5 – פעולות היסטוגרמה, סינון וסינון אדפטיבי – כל אלה יכולים להיעשות ע"י שימוש בייצוג נוח המפריד את הבתיות אשר עליה יעשה עיקר העבודה. גם שחזור תמונות מטשטוש ידוע יכול להיעשות בצורה זו – בייצוג השחזור על שדה הבתיות והשארת שכבות הצבע כמות שחן, שהרי ממילא לא רואים את חdonן.

במסגרת פרק זה רצוננו להתבונן בבעיות מיוחדות הקשורות לשיפורה של תמונה הנובעת ממערכות הצבע, וכאלה הדורשות פתרונות חדשים. אנו נתמקד בשלושה נושאים מסווג זה – מגבלות Gamut והתייחסות אליהן, פונקציות הסתברות לתמונות צבע – הקשר בין שכבות ה- RGB, ושהazor מתת-דגימה – Demosaicing –

המושג Gamut פירושו טווח הצבעים הניתנים לשימוש בהתקן Plat. לפיכך, Gamut של מדפסת (או מסך) הוא אוסף כל שלשות ה- RGB אותן יכולה המדפסת לייצר. לרוב ה- Gamut מוגבל ואני כולל את מלאה קוביית ה- RGB. בהינתן תמונה אותה רצוננו להפיק בהדפסה, מה علينا לעשות לתמונה בטרם הפנויותה למדפסת, על מנת לקחת בחשבון את מגבלותה? אם נתעלם והתמונה אמנים מכילה צבעים מחוץ ל- Gamut, התוצאה תהיה קיצוץ של צבעים אלה לצבע הקרוב ביותר אשר ניתן להדפסה – זהה הטלה

פשותה של צבוי התמונה על טווח ה-Gamut. במקרים רבים זה מספיק, ובעיקר אם התמונה לא כוללת פיקסלים רבים עם גוונים מחוץ ל-Gamut.

במקרים אחרים, פעולה קיצוץ זו מעליימה פרטים חשובים. אם באזור שמרביתו מחוץ ל-Gamut ישן שפota המתארות פרטים חשובים, מרבית שפota אלה תימחקנה לחלוtin בעקבות פעולה ההטלה הפשטota. דוגמה סינטיטית לתופעה זו נתונה בציור 12.16.



אם ננסה את תהליך ההטלה כפעולות מינימיזציה, ישפוק הדבר אוור על דרך פתרון אפשרית. נסמן את תמונה המקור כ- \underline{Y} , ואת תמונה היעד כ- \underline{X} . הוא יתואר כקובוצה Θ . לכן, פעולה ההטלה פירושה שבהינתן התמונה \underline{Y} , علينا למצוא את \underline{X} הקרוב אליה ביותר (פיקסל מול פיקסל) כך שהתמונה \underline{X} יכולה בתוך הקובוצה Θ . בניסוח מתמטי,

$$\underline{X}_{\text{Opt}} = \underset{\underline{X}}{\text{ArgMin}} \left\| \underline{X} - \underline{Y} \right\|_2^2 \quad \text{Subject to } \underline{X} \in \Theta$$

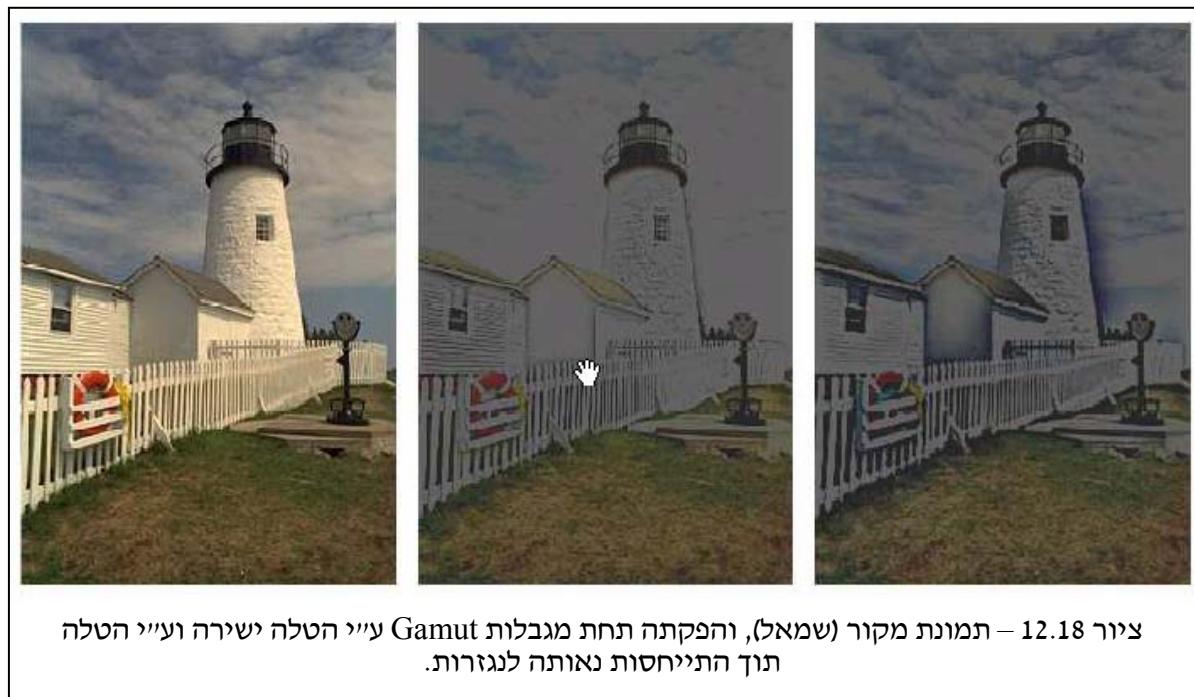
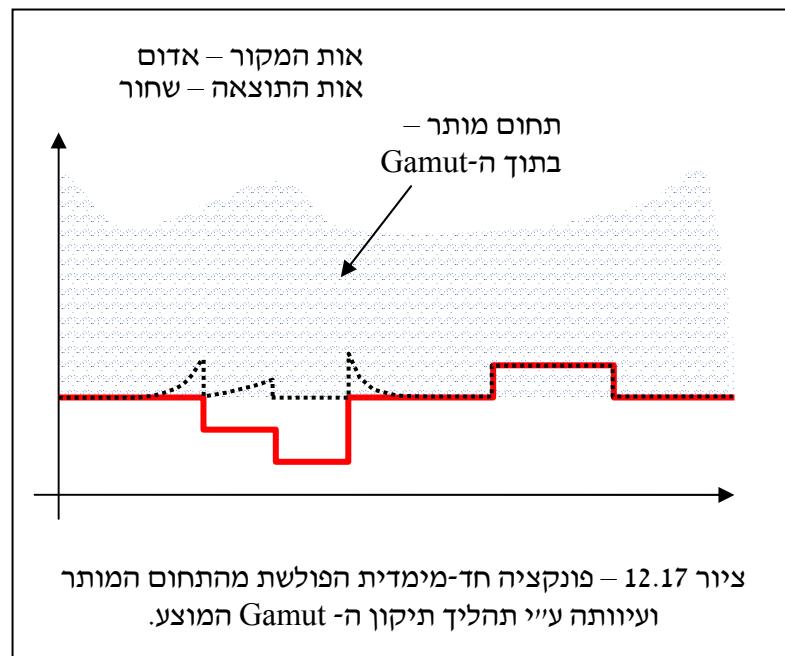
וכאמור, הפתרון פשוט למדי ופועל על כל פיקסל באופן בלתי תלוי בשכניו.

שיעור של תהליך זה יבוא אם ניתן לשכנים כוח להתערב בתוצאת הקביעה של הפיקסל הנוכחי. זאת נוכל להשיג עיי מידת מרחק בין התמונות \underline{X} ל- \underline{Y} אשר תראה לא רק הפרשי רמות RGB בין הפיקסלים בשתי התמונות, אלא גם את ההפרשים בנגזרות. נוכל להציג ביטוי מהצורה

$$\underline{X}_{\text{Opt}} = \underset{\underline{X}}{\text{ArgMin}} \left\| \underline{X} - \underline{Y} \right\|_2^2 + \lambda \left\| \mathbf{D}(\underline{X} - \underline{Y}) \right\|_2^2 \quad \text{Subject to } \underline{X} \in \Theta$$

הגורם החדש מחשב נגזרות מרחביות על שכבות ה-RGB (כל אחת בנפרד, ויתכן אף שכמה כיוונים), והדרישה היא שהקירבה שתהייה בין \underline{X} ל- \underline{Y} תהיה גם בין נגזרותיהם המרחביות. לכן, אם יש שפה בתמונה אחת, מן הראוי שתהייה שפה גם באחרות. עבור הדוגמה שבציור 12.15 המשמעות היא שהרקע ישנה את ערכו קרוב לשפות, על מנת לאפשר קפיצה שתיראה במסגרת ה-Gamut המוגבל. ציור 12.17

ממחיש גרפית את השפעת השימוש בngezrot על פונקציה חד מימדית (למשל, שורה אחת בתמונה שבצior). ציור 12.19 מראה דוגמה מעשית עם התוצאה הצפואה.



הפתרון הנומירי של בעיית האופטימיזציה שניתנה יכול להיות עיי איטרציות השיפוע המרבי,
תוך כדי הטלה על מרחב האילוצים. אך

$$\hat{\underline{X}}_{k+1} = \operatorname{Proj}_{\Theta} \left\{ \hat{\underline{X}}_k - \mu \left[\hat{\underline{X}}_k - \underline{Y} + \lambda \mathbf{D}^T \mathbf{D} (\hat{\underline{X}}_k - \underline{Y}) \right] \right\}$$

במהותה, פועלת ההטלה היא אותה פעולה שעשינו קודם – חישוב הצלב הקרוב ביותר מ-*Gamut* כזו נעשה לכל פיקסל בנפרד. השוני הוא במערכות של נזירות מרחביות בעדכון הפתרון. ישן כMOVEDן דרכים להציג התכונות התהיליך, אך אין לא נכנס לנושא זה. אם ה-*Gamut* יוצר קבוצה קמורה, אלגוריתם זה מובטח לתוכנס.

אנו עוברים כעט לנושא של פונקציות הסטברות לIALIZEDות כשהפעם נרצה לקחת בחשבון צבע. זהו נושא הנמצא עדין בחיתוליו, והידע המצו依 בידנו כיים, לרבות מה שיוצג כאן לראשונה בלבד. עבורIALIZEDות

שchorolben כבר ניסחנו פונקציית הסטברות זו כפיגול Gibbs מהצורה

$$P(\underline{X}) = C \cdot \exp\left\{-\alpha^2 \rho\{\underline{DX}\}\right\}$$

(ראה פרק 6) כאשר *C* נועד לנרגול פונקציה זו, *α* הוא פרמטר שקובע את מידת פיזור הסטברות, ו-
(x) ק היא פונקציה רובסטית כגון פונקציית הערך המוחלט, על מנת להרשות שפות בתמונה. אנו רוצים להכליל פונקציה זו לIALIZEDות צבע וההשערה הראשונה שבאה בחשבון היא המבנה

$$\begin{aligned} P(\underline{X}) &= C \cdot \exp\left\{-\alpha^2 [\rho\{\underline{DX}_R\} + \rho\{\underline{DX}_G\} + \rho\{\underline{DX}_B\}]\right\} = \\ &= C \cdot \exp\left\{-\alpha^2 \rho\{\underline{DX}_R\}\right\} \cdot \exp\left\{-\alpha^2 \rho\{\underline{DX}_G\}\right\} \cdot \exp\left\{-\alpha^2 \rho\{\underline{DX}_B\}\right\} \end{aligned}$$

כלומר, בבחירה זו הצענו כי שכבות ה- RGB כל אחת לגופה חלקה למקוטיעין. הפגם בבחירה זו הוא שאנו לא מצאים קשר בין השכבות השונות. למעשה, מהעובדת שההסטברות נכתבת כמכפלה של שלוש פונקציות הסטברות לכל שכבה נובע כי אנו מניחים אי-תלות. האמם שכבות ה- RGB בלתי תלויות זו בזו?

בחינה של סוגיה זו מגלת כי ישנו קשר הדוק הקובל עלי מיקומן של שפות וכיווניהן צריך להתאים בין שכבות ה- RGB. ישן דרכים מגוונות לנחס זאת מתמטית. אחת הדרכים פשוטות היא הרישום

$$\begin{aligned} P(\underline{X}) &= C \cdot \exp\left\{-\alpha^2 [\rho\{\underline{DX}_R\} + \rho\{\underline{DX}_G\} + \rho\{\underline{DX}_B\}]\right\} \cdot \\ &\quad \cdot \exp\left\{-\beta^2 [\eta\{\underline{D}(\underline{X}_R - \underline{X}_G)\} + \eta\{\underline{D}(\underline{X}_G - \underline{X}_B)\} + \eta\{\underline{D}(\underline{X}_B - \underline{X}_R)\}]\right\} \end{aligned}$$

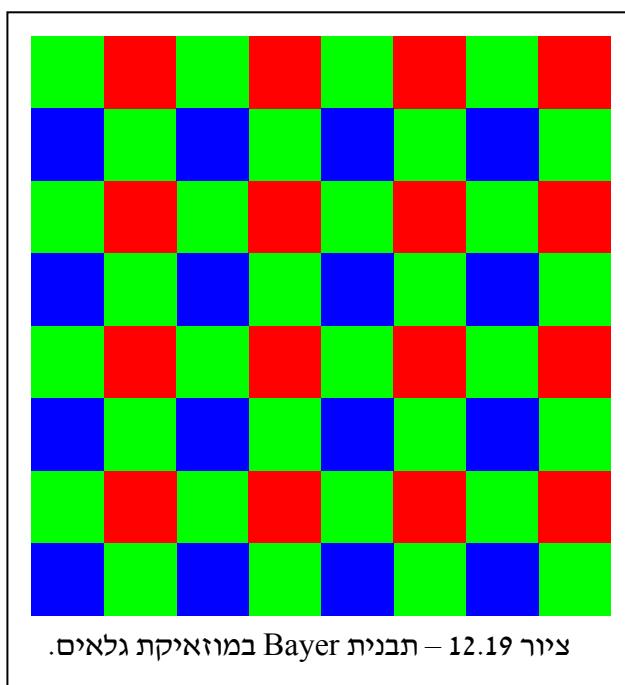
המרכיב הראשון דורש כאמור חלוקות למקוטיעין של כל שכבה בנפרד, והמרכיב השני דורש התאמת שפות בין צמדי שכבות. הפונקציה η מתמחרת את שינויי השפות בין השכבות, וגם היא יכולה להיבחר כפונקציה רובסטית כלשהי. דרך אחרת להסתכל על הביטוי השני היא הדרך הבאה – גם(TM)נות ההפרש בין שתי שכבות צריכה להיות חלקה למקוטיעין ועם מעט שפות ככל האפשר. ברור כי אם השפות תתלבבד נקבל כי אנרגיה זו תוקטן, כפי שהוא רצוי. ניתן להמשיך ולשפר פונקציה זו לקחת בחשבון כי עוצמת הנזורת בשפה אינה זהה בין השכבות ולכן גורם תיקון שיופיע חלק מהנעלים.

בاهינתן פונקציית הסטברות מוצלחת לIALIZEDות צבע, אנו חמושים לפתרון בעיות היפוך בהן הנעלם היא TMIZEDות צבע. נוכל בדרך זו לפתח סינון מבוסס MAP לשם שחזור TMIZEDה מטשטוש ורעש, ובכך לעקוף את ההצעה האמפירית להמיר התמונה ל- YIQ בטרם השחזור.

בURITY שחוֹר מענינית המחייבת פונקציית Prior מהצורה הניל' היא בURITY ה-Demosaicing.

מצלמות דיגיטליות תמונה צבע צרכות בפועל שלושה מערכים שווים גודל אשר יחוֹש את שכבות ה-R, G וה-B בנפרד. אם לא די בכך, הרי שלושת מערכיו גלאים אלה צרכיים "לראות" את היעד עם תיאום גיאומטרי מלא – ככל שלוש פיקסלים מתואמת RGBattiichס למקומם זהה באובייקט המצלום. דבר זה אפשרי ע"י הכנסת האור למצלמה ופיזולו בפריזמה לשכבות הגלאים.

כל הגישה זו יקרה למדי אך יש מצלמות שפועלות כך והוא נחטיבת מצלמות מיקצועיות יותר וכן כמובן גם יקרים יותר. האלטרנטיבה היא שימוש במרקם גלאים אחד שבו חלק מהגלאים רואים יrox, חלק אדום וחלק כחול, כשהבחירה נעשית ע"י יציקת מסיכה קטנה על פני כל גלאי לבחירת הצבע המתאים. אחת הדרכים ליצור מזאיקה זו היא ה-Bayer Pattern אשר נראה בציור 12.19.



בURITY השחוֹר אם כך היא: בהינתן מזאיקת הערכים (שפירוש תת-דגימה של תמונה הצבע המלא – RGB), שחזור את הערכים החסרים. שיטות לשחזור התמונה המלאה קרויות אלגוריתמי Demosaicing, וכאליה קיימים רבים. יש לזכור כי **כל** מצלמה דיגיטלית ביתית כיום כוללת בתוכה Demosaicing ייחודי, ולא פעם איקות התמונה בצילום נקבע על-פי האלגוריתם שיוושם.

ניתן להציג גישה פשוטית שמבצע שחזור של הירוקים מהירוקים, האדומים מהאדומים והכחולים מהכחולים. שחזור זה יכול להיעשות ע"י שיטות אינטראופולציה שלוקחות ערכים ידועים בסביבה, מקרים פולינום דו-מימדי מסדר כלשהו, ומຕוך קירוב זה קובעים את הערך החסר. גישה זו תיתן לרוב תוצאה גרועה בסביבת שפות, בשל אי-תיאום של מיקום השפה בשכבות הצבע – דבר היוצר הדרדר.

שיטת מוביסטט על שחזור MAP אשר תעשה שימוש בפונקציית ההסתברות שהוצעה קודם
תציג את בעיית המינימיזציה הבאה :

$$\text{Min}_{\underline{X}_R, \underline{X}_G, \underline{X}_B} \left\{ \begin{array}{l} \|\mathbf{S}_R \underline{X}_R - \underline{Y}_R\|_2^2 + \|\mathbf{S}_G \underline{X}_G - \underline{Y}_G\|_2^2 + \|\mathbf{S}_B \underline{X}_B - \underline{Y}_B\|_2^2 + \\ + \alpha^2 [\rho \{\mathbf{D}\underline{X}_R\} + \rho \{\mathbf{D}\underline{X}_G\} + \rho \{\mathbf{D}\underline{X}_B\}] + \\ + \beta^2 [\eta \{\mathbf{D}(\underline{X}_R - \underline{X}_G)\} + \eta \{\mathbf{D}(\underline{X}_G - \underline{X}_B)\} + \eta \{\mathbf{D}(\underline{X}_B - \underline{X}_R)\}] \end{array} \right\}$$

בביטוי זה, הסימן \underline{Y}_R ניתן למדידות בPixelsים בהם יש ערך אדום, ובאופן דומה אנו מגדירים את \underline{Y}_G ו- \underline{Y}_B . האופרטור \mathbf{S}_R הוא האופרטור המקבל תמונה אדום מלאה ובוחר ממנה את הערכים שקיים
במזיאקה. גזירת ביטוי זה ביחס ל- $\underline{X}_R, \underline{X}_G, \underline{X}_B$ נותנת

$$\begin{aligned} \frac{\partial f}{\partial \underline{X}_R} &= \mathbf{S}_R^T (\mathbf{S}_R \underline{X}_R - \underline{Y}_R) + \alpha^2 \mathbf{D}^T \rho' \{\mathbf{D}\underline{X}_R\} + \beta^2 \mathbf{D}^T [\eta' \{\mathbf{D}(\underline{X}_R - \underline{X}_G)\} + \eta' \{\mathbf{D}(\underline{X}_R - \underline{X}_B)\}] \\ \frac{\partial f}{\partial \underline{X}_G} &= \mathbf{S}_G^T (\mathbf{S}_G \underline{X}_G - \underline{Y}_G) + \alpha^2 \mathbf{D}^T \rho' \{\mathbf{D}\underline{X}_G\} + \beta^2 \mathbf{D}^T [\eta' \{\mathbf{D}(\underline{X}_G - \underline{X}_R)\} + \eta' \{\mathbf{D}(\underline{X}_G - \underline{X}_B)\}] \\ \frac{\partial f}{\partial \underline{X}_B} &= \mathbf{S}_B^T (\mathbf{S}_B \underline{X}_B - \underline{Y}_B) + \alpha^2 \mathbf{D}^T \rho' \{\mathbf{D}\underline{X}_B\} + \beta^2 \mathbf{D}^T [\eta' \{\mathbf{D}(\underline{X}_B - \underline{X}_R)\} + \eta' \{\mathbf{D}(\underline{X}_B - \underline{X}_G)\}] \end{aligned}$$

ובתהליך איטרטיבי ניתן להפעיל עדכון מבוסס SD שיתכנס למינימום של הפונקציה שתוארה. אם נתבונן בביטויי הגזירה נגלה, בדומה לשחזור תמונות, כי אין כל צורך בייצוג וקטורי מטריצי ובניסוח אופרטורים כמטריצות. אבני הבניין הנחוצים פשוטים למדי :

- א. החסרת תמונות – כגון ביצוע $(\mathbf{S}_R \underline{X}_R - \underline{Y}_R) - \underline{X}_R$ או $(\mathbf{S}_B \underline{X}_B - \underline{Y}_B) - \underline{X}_B$. פעולה זו תיעשה כהכרה פיקסל מפיקסל כששתי התמונות שמורות כמערך דו-מימדי.
- ב. הפעלת אופרטור גזירה כלפלסיאן – ה嚢פלה ב- D או D^T . פעולה זו תיעשה כקובולויזיה על התמונה, עם הכנסת שיקולי קצוטות כגון המשכה מחזורתית או תמונה מראה.
- ג. הפעלת פונקציה סקלרית על Pixelsים בתמונה – כגון $\{\mathbf{D}(\underline{X}_R - \underline{X}_B)\}$ או $\{\rho'(\mathbf{D}\underline{X}_R)\}$.
- הפעולות אלו אין אלא פועלות LUT על כל פיקסל. למשל, עבור הבחירה $x(m, n) = |x|$, הפעולה הנדרשת היא נגזרת פונקציה זו – דהיינו פועלות סימנו. אם ערך הפיקסל שלילי, -1 ואם חיובי $+1$.
- ד. הפעלת האופרטור \mathbf{S}_R פירושה בחירות Pixelsים שישם במזיאקה והתעלמות מאחרים. הפעולה המשוחלפת פירושה הכנסת אפסים לערכים החסרים.

קיימות כאמור גישות שונות ומגוונות לבניית תמונה ה- Demosaic. המשותף למצלחות שבוחן הוא ניסוח נאות של הקשר בין שכבות הצבע. בציור 12.20 מוצגות שתי תוצאות של תהליך Demosaicing מבוסס במידה חישובית¹ (כלומר, מראים למערכת מספר דוגמאות של תמונה מזיאקה והתמונה שהיתה צריכה להתקבל, והמערכת לומדת על הקשר בין שכבות הצבע מתוך דוגמאות אלה). התוצאות מתייחסות לטיפול בכל שכבה בנפרד וטיפול משולב שבו רואים הרבה פחות פגעי צבע.

¹ תוצאות אלו נוצרו בעבודת המחקר של דרי חגי הל-אר מאוניברסיטת חיפה.



ציור 12.20 – תוצאה Demosaicing – גישה המטפלת בשכבות בנפרד (מיימין) וגישה המערבת שכבות (משמאל) – באדיבות דרי' חגיית הל-אור.

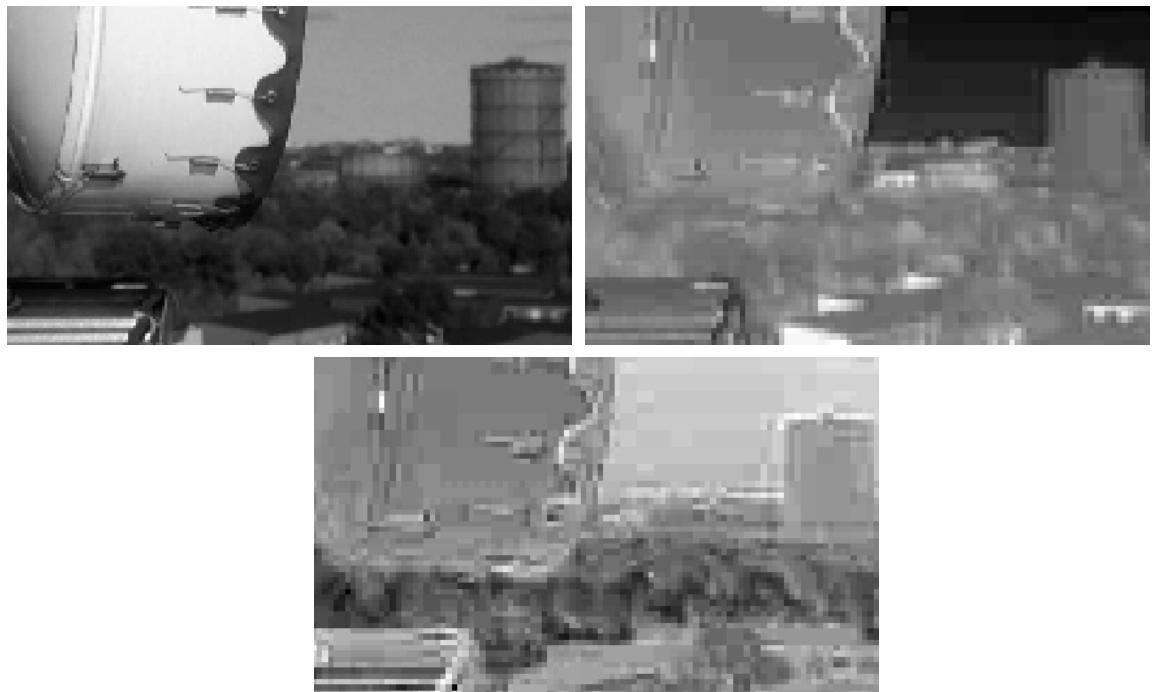
12.6 דחיסת תמונות צבע

כמו בכל סוגיה אחרת בעיבוד תמונות, כשהדבר נוגע לתוכנת צבע, מוצע הפתרון האד-הוקי של הפרדת הבתיירות ממאפייני הצבע. כבר רأינו כי ניתן להחזיק תמונה צבע בייצוג QIY בן אנו זורקים 50% מסך המידע ע"י דילול ביחס 1:2 בכל ציר בשכבות ה- I ו- Q ללא נזק ניכר. הגישה מכאן יכולה להמשיך ע"י דחיסת כל אחת מהשכבות בנפרד בכל אחת מהגישות שנדונו קודם, כגון גישת חיזוי או התמרה, כשהם מלאוים בקונטיזציה מתאימה. בcontra זו אנו מנצלים במידה מסוימת את הקשר בין השכבות ע"י ההמרה והדילול של I ו- Q. עם זאת, אם נסתכל על התמונות הללו, כמו שמתוארות בציור 12.21, ברור כי עדין קיים קשר ביןיהם, גם אם חלש, ודחיסתן בנפרד מבזבזת יתרות זו.

נוכל גם להציג אלגוריתמי דחיסה שרואים את הקשר בין השכבות בצורה ישירה. כשהתחלנו את הדיוון בדחיסת תמונה הזכרנו את שיטת ה- PCM שאינה אלא קוונטיזציה סקלרית ישירה על רמות האפור. בהכללה טبيعית, נציג קוונטיזציה וקטורית על שלשות ה- RGB במספר מצומצם של רמות ייצוג ובכך לדחוס את התמונה – זהו בדיקת התהילה שעשינו מילא לשם ייצוג נוח במסגרת הדיוון בתמונות Indexed. מסתבר כי ניתן אף לעשות יותר – אם נתיחס למשפחת התמונות הסינטטיות כגון ציורים, גרפיקה, ועוד, הרי שמקבץ הצבעים שיידרש יהיה קטן. דבר זה יביא לדחיסות גבוהה גם בשל העובדה הקטן יותר, וגם (וכנראה בעיקר) בשל העובדה שסט האינדקסים בעל אלף-בית קטן Palett.

בحينו סט האינדקסים עבור תמונה סינטטית, יתקבל כי אלה יוצרים תמונה חלקה למוקוטען. כמובן, יהיו אזורים רבים בהם האינדקס לא משתנה בין שכנים, ופירוש הדבר פוטנציאלי דחיסה נוספת. שימוש בזמנים כמו Run-Length-Encoding או וריאציות שונות של אלגוריתם Lempel-Ziv יכולים

לשם דחיסה ללא עיות של האינדקסים, וזוהי בדיקת הגישה המובילה לפורטט ה- GIF וה-PNG. למעשה שני פורטטים אלה וווחים מידי באינטרנט בשל פשוטותם ויעילותם.



צייר 12.21 – שכבות התמונה ב- Y (למעלה ממשمال), I (למעלה מימין) ו- Q (למטה). רואים כי יש קשר ביניהן – זהה יתרות שאותה ניתן לנצל בדחיסה משולבת.

אם נמשיך עם אותו קו מחשבה שננקט בפרק 8, הרי שקבענו כי דחיסת PCM (קונטיזציה ישירות לערכי התמונה) אינה מוצלת כמו דחיסת הפרשים הסטטיסטיים – דבר שהוביל אותנו לגישות חיזוי. נוכל להציג סכימת DPCM בה נזהה את שלשת ה- RGB בפיקסל על פי השכנים בדרך כלל. נגידר סט שכנים סיבתי בສריקה בדרך כללו של הפיקסלים, השתמש בשכנים הסיבתיים לבנות חיזוי של שלושת ערכי ה- RGB, ונחסיר החיזוי מהערכים האמתיים שישנם בפיקסל. את אותן השגיאה נוכל להזין לקונטיזר וקטורי שאותו נתקן ע"י בניית סט גדול של הפרשי חיזוי צפויים. מעניין לציין שאפילו אם החיזוי של כל שכבה יעשה בנפרד (דבר שוביל לשגיאות חיזוי סקלריות שניתנות לחישוב בנפרד שלוש סכימות DPCM נפרדות), עדין קידוד וקטורי תלת-מימדי עדיף על פני קונטיזציה סקלרית. זאת כיון שבקונטיזציה סקלרית אנו מוצאים קשר בין השכבות, ומעבר לכך, מרצפים את המרחב בצורה יעילה יותר.

נספחים

נספחים :	
285	נספח א' – הרחבות לקוונטייזר אופטימלי
286	נספח ב' – התמרת KLT כהתמורה אופטימלית והרחבותיה
294	נספח ג' – הקצאת סיביות אופטימלית בדחסית
299	נספח ד' – נושאים מתקדמים בשחזור תמונות
302	

בפרק זה נסיף דיון במספר נושאים מגוונים שמעשוריים את הדיון שניתן בפרקם קודמים. בחירת הנושאים הנוספים נעשתה תוך חיפוש של סוגיות בעלות יופי תיאורטי וכן תרומה מעשית חשובה לתהום. מלכתחילה הכוונה הייתה להכניס נושאים אלו לפרק המתאים, וכנראה שכז' יעשה לבסוף.

נספח א' – הרחבות לקוונטייזר אופטימלי

בפרק 4 למדנו על תהליך הקוונטייזציה ועל דרכיים לעילו. ראיינו כי ניתן לתכנן קוונטייזר אופטימלי להבאת השגיאה הריבועית למינימום. במסגרת נספח זהណון בשתי וריאציות של קוונטייזר אופטימלי זה. בחלק הראשון ננסה לפשט את תהליך תכנון הקוונטייזר למקרים בהם פישוט כזה אפשרי. אנו נראה כי כאשר מספר המדרגות בפונקציית הקוונטייזציה גדול דיו, קירוב מפשט כזה אפשרי ואז תכנון הקוונטייזר היא משימה קלה יותר. חלק שני בנספח זה יידען בהחלפת קרייטריון האופטימליות למינימיזציה של האנטרופיה. אנו נראה כי שילוב של קוונטייזר של אחריו קידוד אנטרופיה יוצר פתרונות שונים בתכלית.

1. קוונטייזר Max-Lloyd

נתחיל באזכור מהיר של קוונטייזר Max-Lloyd. נתון לנו אות (נניח כי הוא חד-מימדי כיון שבכל הדיוון שיובהה כאן נתיחס למקרה של הקוונטייזציה הסקלרית) $[a]_x$. במקורו ישן שתי גישות לאפיין מקור זה – האחת, דטרמיניסטיבית, תאמר כי $[a]_x$ נתון לנו עם אינסוף דגימות (כלומר $\infty, \dots, 0, 1, 2, \dots, a$), ואנו נוכל לייצר היסטוגרמה של ערכיו $[a]_x$ אשר מסומן $(x)_k$ והיא תיתן לנו את שכיחותו של הערך x . גישה סטוכסטית המוביילה לאותו מקום היא האמירה שערכיו f מפולגים לפי הפלוג הסקלרי $(x)_k = p$ – כלומר k היא פונקציה ההסתברות להגשה של ערך x .

הקוונטייזר אותו אנו עומדים לתכנן הינו פונקציית מדרגות כמתואר בציור A.1. הפונקציה Q מבצעת קפיצה בערכי x השווים $\{r_k\}_{k=0}^N$ (אלו הם רמות החלוקת), וערכה באינטרוול $[r_{k-1}, r_k]$ הוא $r_N = +\infty$ ו- $r_0 = -\infty$ (אלו הם רמות הייצוג). אנו נניח כי ערכי הפונקציה מקבלים כל ערך ממשי, ולכן $r_N = +\infty$ ו- $r_0 = -\infty$ קבועים וידועים מראש.

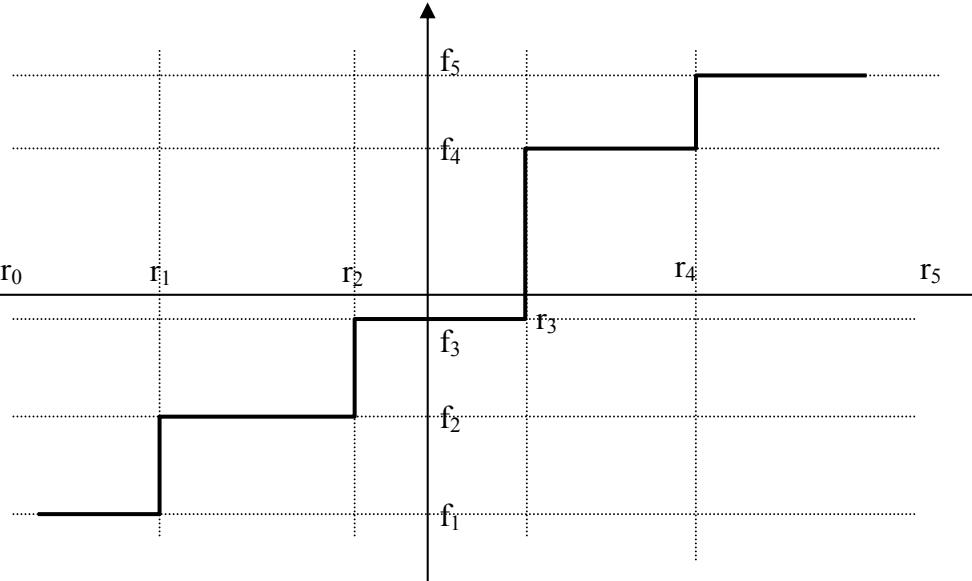
בתפיסה הדטרמיניסטיבית נגדיר את השגיאה הממוצעת בין האות x ובין הקוונטייזציה שלו $\{x\}_k$

ע"י

$$\bar{E} = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N (x[n] - Q\{x[n]\})^2$$

כיוון שהפונקציה Q היא פונקציית הקוונטייזציה וזו סקלרית וחסרת זכרון, ניתן לשנות את דרך הסכימה בסכימה על פני הערכים האפשריים של x . המבנה החדש יהיה

$$\bar{E} = \lim_{L \rightarrow \infty} \frac{1}{L} \sum_{n=1}^L (x[n] - Q\{x[n]\})^2 = \int_{-\infty}^{\infty} p(x) (x - Q\{x\})^2 dx$$

ציור 1.A – תיאור גרפי של הפונקציה $(x)Q$ – פועלות הקוונטיזציה.

בשינוי זה קבענו כי כל ערך של x מופיע $(x)Q$ פעמים (לאחר נירמול), ולכן במקומות לסכום על ציר הזמן אנו סוכמים על ציר ערכי x . עם שינוי זה קיבלנו כי ייצוג הבעה חוזר למבנה הסטטיסטי, כיוון שהשגיאה הממוצעת לפי גישה סטטיסטית נתונה ע"י

$$\bar{E} = E\{(x - Q\{x\})^2\} = \int_{-\infty}^{\infty} p(x)(x - Q\{x\})^2 dx$$

וזהו כמוון אותו ביטוי בדיקוק.

כיוון שהפונקציה Q היא פונקציה מדרגות המבצעת קפיצה בערכי x השווים $\{r_k\}_{k=0}^N$, וערכה קבוע ושווה f_k , נוכל לפרק את האינטגרל הניל לנתחים ולקבל

$$\bar{E} = \int_{-\infty}^{\infty} p(x)(x - Q\{x\})^2 dx = \sum_{k=1}^N \int_{r_{k-1}}^{r_k} p(x)(x - f_k)^2 dx$$

מטרתנו היא מציאת $\{f_k\}_{k=1}^N$ שיביאו שגיאה זו למינימום. גזירת השגיאה לפי הנעלם f_k מניבה

$$\begin{aligned} \frac{\partial}{\partial f_k} \sum_{k=1}^N \int_{r_{k-1}}^{r_k} p(x)(x - f_k)^2 dx &= \int_{r_{k-1}}^{r_k} p(x) \frac{\partial}{\partial f_k} (x - f_k)^2 dx = 2 \int_{r_{k-1}}^{r_k} p(x)(f_k - x) dx = 0 \\ \Rightarrow f_k &= \frac{\int_{r_{k-1}}^{r_k} x \cdot p(x) dx}{\int_{r_{k-1}}^{r_k} p(x) dx} \end{aligned}$$

כלומר, בידיעת הסידרה $\{r_k\}_{k=0}^N$, רמות הייצוג יימצאו כמרכזי כובד של האינטראולים הקבועים. באופן דומה, גזירה לפי x נותנת

$$\begin{aligned} \frac{\partial}{\partial r_k} \sum_{k=1}^N \int_{r_{k-1}}^{r_k} p(x)(x - f_k)^2 dx &= \frac{\partial}{\partial r_k} \left[\int_{r_{k-1}}^{r_k} p(x)(x - f_k)^2 dx + \int_{r_k}^{r_{k+1}} p(x)(x - f_{k+1})^2 dx \right] = \\ &= p(r_k)(r_k - f_k)^2 - p(r_k)(r_k - f_{k+1})^2 = 0 \\ \Rightarrow r_k &= \frac{f_k + f_{k+1}}{2} \end{aligned}$$

כלומר, הפעם קיבלנו תוצאה דואלית לפיה בידיעת $\{f_k\}_{k=1}^N$ נוכל לקבוע את $\{r_k\}_{k=0}^N$ כערכי המוצע שבין רמות הייצוג.

בשל העובדה המשוואות הכלולות סטומה, הוצע פתרון איטרטיבי, בו מתחילה בನיחוש מוצלח של רמות ההחלטה (או רמות הייצוג – זה לא משנה), וمعدכנים את רמות הייצוג לפי המשוואות הניל. עם הערכים המעודכנים חוזרים ומשנים את רמות ההחלטה, וחוזר חלילה.

2. קירוב קוונטייזר Max-Loyd Companding

אחת הביעות המרכזיות בתכנון קוונטייזר Max-Lloyd היא בעיית ההינעלות על מינימה לוקלית. להצלחת התהילה האיטרטיבי הניל יש קשר הדוק לבחירת האיתחול, ולשימוש בשיטות החלקה שונות. מסתבר שעבור קוונטייזרים בהם מספר המדרגות גדול דיו ניתן להשיג פתרון מקרוב לבעה הניל תוך הימנעות ממינימות לוקליות. גישה זו אותה נציג מיד ידועה בשם Companding בשל סיבה אותה נזכיר בהמשך. ההנחה היסודית אותה נבקש לקיים היא שגודל המדרגות כה קטן ביחס לשינויים בפונקציית הפילוג שנוכל לקבוע כי בקירוב טוב

$$\bar{E} = \int_{-\infty}^{\infty} p(x)(x - Q\{x\})^2 dx = \sum_{k=1}^N \int_{r_{k-1}}^{r_k} p(x)(x - f_k)^2 dx = \sum_{k=1}^N \int_{r_{k-1}}^{r_k} p(r_k)(x - f_k)^2 dx$$

כלומר, באינטראול $[r_{k-1}, r_k]$ אנו מבקשים להניח כי פונקציית הפילוג בקירוב טוב קבועה ושווה לפיכך ל- $p(r_k)$ – ערך פונקציה זו בקצת אינטראול זה (נוכל במידה דומה להניח כי פונקציית הפילוג מתийיחס

לנקודת המצע של האינטראול, דהיינו $p(0.5[r_k + r_{k-1}])$, ודבר זה נוח יותר). תחת הנחה זו קיבל כי השגיאה נתונה כ:

$$\bar{E} = \sum_{k=1}^N \int_{r_{k-1}}^{r_k} p(r_k)(x - f_k)^2 dx = \frac{1}{3} \sum_{k=1}^N p(r_k) [(r_k - f_k)^3 - (r_{k-1} - f_k)^3]$$

גזרה של ביטוי זה לפי f_k תיתן

$$\frac{\partial}{\partial f_k} \sum_{k=1}^N p(r_k) [(r_k - f_k)^3 - (r_{k-1} - f_k)^3] = p(r_k) [-(f_k - r_k)^2 + (f_k - r_{k-1})^2] = 0$$

$$\Rightarrow f_k = \frac{r_{k-1} + r_k}{2}$$

זו לא תוצאה זהה לתוצאה אחרת ראיינו קודם קודם כיון שהפעם נקבע שרטת **היצוג מתකבלת** באמצעות האינטראול $[r_{k-1}, r_k]$ (דבר מעורר תהיה אם חשוב על קצוות הגף! אנו נראה כי למרות צורה זו, הביטויים אליהם נגעו לא סובלים מבעה). נוכל להציב ביטוי זה חוזה לשגיאה ולקבל כי

$$\bar{E} = \frac{1}{3} \sum_{k=1}^N p(r_k) [(r_k - f_k)^3 - (r_{k-1} - f_k)^3] = \frac{1}{12} \sum_{k=1}^N p(r_k) (r_k - r_{k-1})^3$$

וכעת עליינו להביא למינימום ביטוי זה כפונקציה של רמות החלטה. במקום לנוקוט בגישה ישירה למינימיזציה, נציג החלטת הנעלמים. נסמן את הגודל

$$\mu_k = p(r_k)^{1/3} \cdot (r_k - r_{k-1})$$

ואז השגיאה שאויה יש להביא למינימום היא

$$\bar{E} = \frac{1}{12} \sum_{k=1}^N p(r_k) (r_k - r_{k-1})^3 = \frac{1}{12} \sum_{k=1}^N \mu_k^3$$

נשאל מהם המספרים μ_k שיביאו לשגיאה זו לפחות? אלא שישנו אילוץ שערכיכם אלו חייביםקיימים:

$$\sum_{k=1}^N \mu_k = \sum_{k=1}^N p(r_k)^{1/3} (r_k - r_{k-1}) \cong \int_{r_0}^{r_N} p(x)^{1/3} dx = K$$

כיוון שחוק הפילוג (x) ידוע, הגודל K ניתן לקביעה מראש והוא תלוי רק בחוק הפילוג. לכן, ניתן את בעיית האופטימיזציה

$$\text{Minimize}_{\mu_k} \sum_{k=1}^N \mu_k^3 \quad \text{Subject to} \quad \sum_{k=1}^N \mu_k = K$$

ע"י קופלי לגרנזי ונקבל

$$\begin{aligned} L\{\mu_1, \mu_2, \dots, \mu_N\} &= \sum_{k=1}^N \mu_k^3 - \lambda \left(\sum_{k=1}^N \mu_k - K \right) \\ \frac{\partial L}{\partial \mu_j} &= 3\mu_j^2 - \lambda = 0 \Rightarrow \mu_j = \sqrt{\frac{\lambda}{3}} = \frac{K}{N} \end{aligned}$$

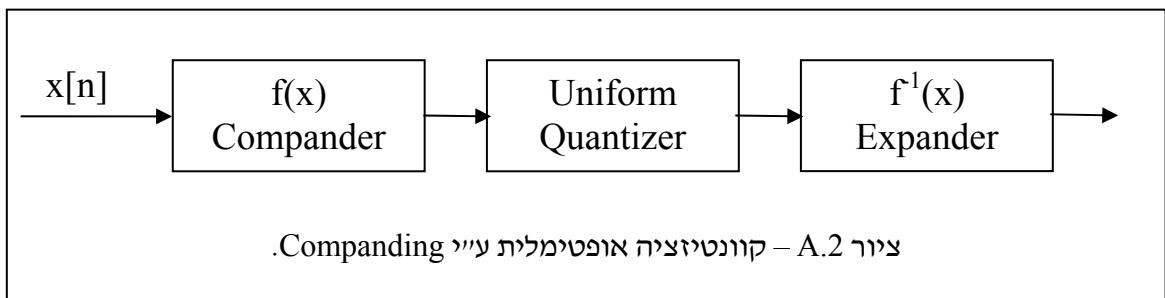
התוצאה מעניינת מאוד וקובעת כי להשגת אופטימום על כל μ_j להיות שווים. כיוון שה- μ הוא בקירוב טוב השטח מתחת פונקציית ההסתברות בחזקת שליש באינטראול $[r_{k-1}, r_k]$, פירוש הדבר הוא

שים לחת את הפונקציה $f(x)^{1/3}$ ולחילקה ל- N נתחים שווים שטח, ואלה יגדירו את מיקום רמות החלטה האופטימליות. בצורה סידורית נמצא את רמות ההחלטה -начילה עם $z_0 = \infty$, ונקבע את z_1 כנקודה בה צבירה הפונקציה $f(x)^{1/3}$ שטח השווה ל- $N/1$ משטחה הכלול (הלא הוא K). משם נחפש את z_2 בצורה דומה וחזרה כך עד z_{N-1} , כאשרור כי $\infty = z_N$. השגיאה הנובעת מקוונטייזר זה נתונה ע"י

$$\bar{E} = \frac{1}{12} \sum_{k=1}^N \mu_k^3 = \frac{1}{12} \left(\frac{K}{N} \right)^3 \cdot N = \frac{K^3}{12N^2} = \frac{K^3}{12} 2^{-2J}$$

ביטוי זה חשוב ונעשה בו שימוש תכוף – הוא נותן הערכה לשגיאת הקוונטייזציה המומוצעת כפונקציה של פונקציית התפלגות (חבוי ב- K) וכפונקציה של מספר הסיביות המוקצות לייצוג (ל- J סיביות יש 2^J רמות). מביטוי זה נובע כי כל הוספת סיבית מקטינה את השגיאה הריבועית בפקטור 4.

از' למה Companding? מכיוון שישנה גישה אחרת שモבילה אותה תוצאה בדיק. בשנות ה-60' בהם שיקולי חומרה היו מהותניים נשאלת השאלה הבאה – האם ניתן יהיה להחילף את הקוונטייזר האופטימי (שסביר כי יהיה בעל רוחב מדרגות משתנה) בתהיליך שקול כמתואר בציור A.2, בו האות הכניסה עובר עיות חסר זיכרון (פولات LUT), אחריו קוונטייזר אחד, ולבסוף עיות הפוך?



על סמך הנition שהוצע לעיל למעשה קיבלנו את התשובה, לפיה ניתן בבחירה להקרב לביצועי קוונטייזר Max-Lloyd עבור קוונטייזיה צפופה מספיק, וזאת נשיג ע"י הפונקציה

$$w = f(x) = \frac{\int_{-\infty}^x p(z)^{1/3} dz}{\int_{-\infty}^{\infty} p(z)^{1/3} dz}$$

פונקציה זו ממירה את המשטנה האקראי x (נשברנו – אין יותר ראייה דטרמיניסטיבית!) הקיים על פני כל ציר הממשיים במשטנה חדש w הקיים באינטראול $[0,1]$. קוונטייזיה אחת על פני w הוא בדיק המירושם לקירוב קוונטייזר Max-Lloyd שקיבלנו קודם (בשל הדרישת שעיל כל ה- z_i להיות שווים).

באשר לעיון החזר (קראנו לו Expander), כל תכilitו היא לשנות את רמות הייצוג שיקבעו כפוזרים אחדים בתחום $[0,1]$ לערךים שמקנים את ערכי x ולבן העיונות לאחר. דרך אגב, קל לטועות ולחשוב כי העיונות הנחוץ f צריך להיות זה שימיר את המשטנה האקראי x לבעל פילוג אחד, וכפי שראויים אין זה כך כלל!

דוגמה A.1 – נניח כי x בעל פילוג אחד בתחום $[0,1]$. במקרה זה פונקציית העיונות המוצעת ($f(x)$) היא פונקציה היחידה, דהיינו, $x=w$. لكن פילוג אחד קיבל בשיטה זו קוונטייזציה איחודית, שהיא גם האופטימלית.

דוגמה A.2 – נניח כי x בעל פילוג מהצורה

$$\cdot p(x) = \begin{cases} 0 & x < 0 \\ 2x & x \in [0,1] \\ 0 & x > 1 \end{cases} \Rightarrow f(x) = \begin{cases} 0 & x < 0 \\ x^{4/3} & x \in [0,1] \\ 1 & x > 1 \end{cases}$$

לעומת זאת, לו רצינו להפוך את x לבעל פילוג אחד היה נדרשת פונקציית העיונות x^2 (חויה על נושא יישור היסטוגרמות תוכל לרענן את זיכרונכם בדרך לקבלת תוצאה זו). כפי שהוא רואים, אלה שתי פונקציות שונות בתכלית.

3. קוונטייזר אופטימלי במונחי D-R

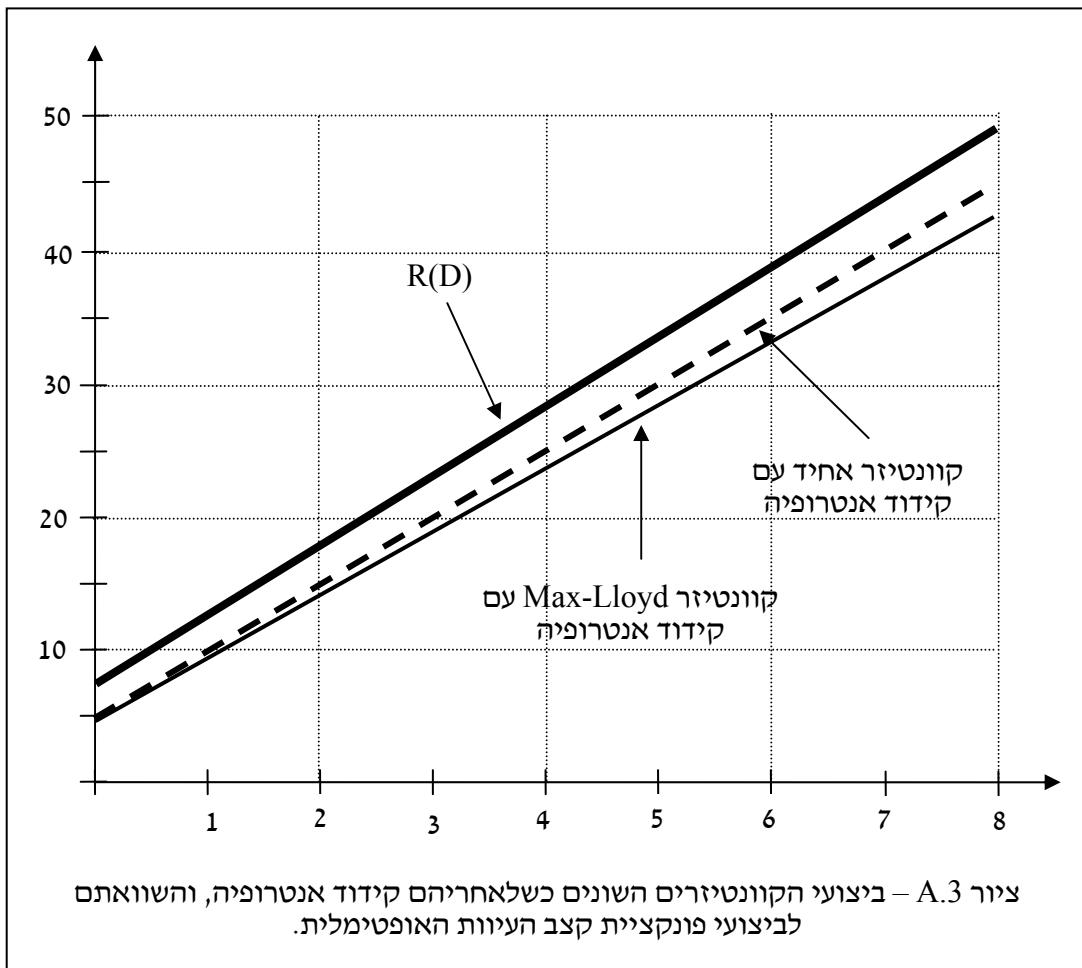
נניח כי נתנו לנו משטנה אקראי x לבן (אין תלות בין דגימותיו) בעל חוק פילוג גausי $\{N(0, \sigma^2)\}$ והוא רצונו לקודד. תורת האינפורמציה מלמדת אותנו כי פונקציית קצב העיונות עבור משטנה זה נתונה ע"י

$$R(D) = \frac{1}{2} \log_2 \left(\frac{\sigma^2}{D} \right) \Leftrightarrow D(R) = \sigma^2 \cdot 2^{-2R}$$

כש- D הוא השגיאה הריבועית הממוצעת. חשוב להבהיר – חישוב פונקציית קצב העיונות בדרך כלל מרכיב למדוי וועל-פי רוב לא ניתן להגעה לביטוי סגור כבמקרה זה. כזכור, פונקציה זו מייצגת את הקידוד הטוב ביותר האפשרי למשטנה זה, ובביטוי זה קרוי גם חסם שנון (Shannon).

נוכל כמובן לגשת לעניין קידוד המשטנה x באמצעות מעשיים – הפעלת קוונטייזציה כלשהי. כיוון שאנו חושבים במונחים של דחיסה, טבעי שנססה לשחות הרבה ככל האפשר ביצוג המשטנה x ולבן נפעיל קידוד אנטרופיה על מוצא הקוונטייזר. כמובן, ניקח את סט רמות הייצוג שלכל אחת סבירות הופעה, ונחשב את האנטרופיה הסקלרית מתוך הנחה שבידינו אלגוריתם שימושי אנטרופיה זו. בפועל, כזכור, מקודד אנטרופיה כמו Huffman מתקרב כדי סיבית אחת אנטרופיה. קוונטייזר ראשון אותו נפעיל הוא הקוונטייזר האחד, כנסנה את גודל המדרגה עד לקבלת שגיאה ריבועית מינימאלית. על התוצאה נחשב אנטרופיה (כגンド כל מספר של מדרגות) ונקבל מקום חדש של קצב-夷וות מעשי. נוכל לחזור עת התהילה כשהפעם הקוונטייזר הוא קוונטייזר Max-Lloyd, וכך נקבל תופעה מוזרה – קוונטייזר Max-Lloyd חלש

יוטר מהאחד! תוצאות אלו מוצגות בציור 3.A. השאלה המרכזית בה נרצה להתמקד היא - איך זה יתכן?



ובכן, בישומו של קונטיזר Max-Lloyd ישנו כוח המתנגד לkidוד אנטרופיה. כאשר פונקציית הפילוג גבוהה יבחר בצד הקונטיזר Max-Lloyd לצוף מדרגות כדי להקטין את שגיאת הייצוג, בעוד שבאזורים נמוכים המדרגות תרווינה. משמעותו תוצאה זו היא שסבירות הופעתן של רמות הייצוג בקונטיזר Max-Lloyd יהיה קרובה לאחד, וזה אומר שהקידוד אנטרופיה יהיה חסר תרומה. בኒיגוד לזה, קונטיזר אחד ייתן שוני משמעותי בין סבירות רמות הייצוג השונות, בכך יפתח פתח לביצועים טובים לkidוד האנטרופיה.

כל זה מוביל לשאלת הבסיסית – מהו הקונטיזר הטוב ביותר במונחי $R(D)$ בהנחה כי מבוצע kidod antropfia על רמות הייצוג? מסתבר כי ניתן לתכנן קונטיזר אופטימלי במונחי $D-R$, אך זה אינו פשוט. עבור המקרה הגאוסי, בחרית קונטיזר אחד אינה רחוכה כלל מהקונטיזר האופטימלי. יתרה מזה, לא ניתן להשיג קונטיזר סקלרי שיגיע לחסם שנון כיוון לכך שהוא השגת החסם נחוץ kidod וקטורי (על אף שהיא לבן!). קיימות עבודות מתחילה שנות ה-70 שעשו בעיה זו וספקו דרכי לתכנן את הקונטיזר האופטימלי, אך אלו נעזרו את הדיוון בנושא זה בשלב זה.

הערה 1: נשים לב כי חסם שנון מציע את הביטוי $D(R) = 2^{-2R} \cdot \sigma^2$ ומתואצאה שקיבלונו עבור קוונטייזר מבוסס Companding קיבלונו $D = K^3 2^{-2R} / 12$, כאשר נלקח בחשבון קידוד אנטרופיה על מוצא הקוונטייזר. יוצא שפקטור השיפור בהוספת סיבית בשתי העקומות שווה ל-4, דבר שמעיד כי שתי עקומות אלו בעלות שיפוע זהה.

הערה 2: ממש כשם שיש מקום לשלב מקודד אנטרופיה בМОצא קוונטייזר סקלרי ולתכנן את הצמד לקבלת ביצועים מיטביים במונחי R-D, יש אפשרות לדבר באופן דומה על קידוד וקטורי. לא יעלה על הדעת שימוש באלגוריתם ה-K-Means ייתן את הביצועים הטובים ביותר מכיוון שבדומה לקוונטייזר Max-Lloyd החלוקה שהוא מציע פועלת נגד קידוד אנטרופיה יעיל.

נספח ב' – התמרת KLT כהתמורה אופטימלית והרחבותיה

בפרק 7 הצגנו את התמרת ה- KLT כהתמורה אופטימלית, והציגו שני מבנים לפיה נבחנת אופטימליות זו. בספח זה נציג בפירוט את ניתוח האופטימליות של התمرة זו, והרחבות מעניינות הנובעות מכך.

1. אופטימליות התמרת ה- KLT

יש מגוון דרכי לבנות את התמרת ה- KLT – אנו מציע דרך סידורית קלה יותר להבנה. נניח כי נתון לנו אוסף וקטורים $\{v_k\}_{k=1}^N$ ורצוננו לתקן עבורם התمرة ליניארית אשר תייצג מטריצה בגודל $N \times N$ הkoplat וקטוריים אלה. לשם קביעת מידת הצלחת ההתمرة, נניח את התהליך הבא – בהינתן וקטור v zusätzlich לתוצאה Ax , ישארו M ראשונים ללא שינוי, וכל היתר יאפסו. וקטור $v = A_x + M$ איברי וקטור התוצאה x , ישארו M ראשונים ללא שינוי, וכל היתר יאפסו. מכיון ש Ax מתקבל ע"י המכפלה $v = I_M A_x = A_x I_M$, כשהמטריצה I_M היא מטריצה אלכסונית עם $M = 1$. חדש זה מתקבל ע"י המכפלה $v = I_M A_x = A_x I_M$, תבחר למשה את השורה הראשונה $M - A$ אותה נסמן כ- \underline{b}_1^T כשל יתר השורות אינן רלוונטיות. הווקטור \underline{b}_1 יהיה בעל איבר ראשון שונה מאשר מ-0 וכל האחרים אפסים בהגדרה. לכן, בשזרור מותק מקדים בכך זה נקבל כי רק העמודה הראשונה במטריצה B משמשת. לכן, נרצה להביא למינימום את הביטוי

$$f(\underline{a}_1, \underline{b}_1) = \sum_k \left\| \underline{v}_k - \underline{b}_1 \underline{a}_1^T \underline{v}_k \right\|_2^2$$

ביטוי זה מבטא את רצוננו לייצג בעזרת המקדם היחיד שנותר את המרב האפשרי מוקטור המקור. ניכר כי לא יכול להיות לביטוי זה פתרון יחיד כי אם צמד וקטורים מסוימים $[\underline{a}, \underline{b}]$ מהווים פתרון אופטימלי, אז כך גם כל צמד $[\underline{a}, \underline{b}]$. לכן, לשם יצירת ייחדות פתרון נפתרו הבעיה הנ"ל תוך כפיה האילוץ $\|\underline{a}\|_2 = 1$. נניח לרגע כי מצאנו את \underline{a} , ונשאל מהו \underline{b} האופטמלי. התוצאה המתבקשת היא

$$\begin{aligned} \frac{\partial}{\partial \underline{b}_1} f(\underline{a}_1, \underline{b}_1) &= \frac{\partial}{\partial \underline{b}_1} \sum_k \left\| \underline{v}_k - \underline{b}_1 \underline{a}_1^T \underline{v}_k \right\|_2^2 = \\ &= \sum_k (\underline{v}_k - \underline{b}_1 \underline{a}_1^T \underline{v}_k) \underline{v}_k^T \underline{a}_1 = (I - \underline{b}_1 \underline{a}_1^T) \left(\sum_k \underline{v}_k \underline{v}_k^T \right) \underline{a}_1 = 0 \end{aligned}$$

נגידר את המטריצה R ע"י

$$\cdot R = \sum_k v_k u_k^T$$

מטריצה זו היא מטריצת הקוריאנס של אוסף הוקטוריים. עבור מקבץ אינסופי של וקטורים אשר להם ניתן להתייחס כוקטורים אקראיים, מטריצה זו מתארת את התלותות החדדיות בין איברי הוקטור.
מתקובל,

$$\cdot (I - b_1 a_1^T) R a_1 = 0 \Rightarrow b_1 = R a_1 [a_1^T R a_1]^{-1} = \frac{1}{\lambda} R a_1$$

כאן סימנו את הסקלר $a_1^T R a_1 = \lambda$. נציב תוצאה זו חזרה לפונקציית המחיר המקורי ונקבל כי עליינו להביא למינימום את הפונקציה

$$\cdot f(a_1) = \sum_k \left\| v_k - \frac{1}{\lambda} R a_1 a_1^T v_k \right\|_2^2$$

נניח כי הפתרון האופטימלי הוא וקטור עצמי של המטריצה R , דהיינו קיים כי $a_1^T a_1 = \lambda$. בהכפלה משמאלי ב- a_1^T נקבל $\lambda a_1^T a_1 = \lambda \|a_1\|_2 = \lambda$, כך שהסקלר שהגדכנו קודם אינו אלא הערך העצמי שמתיחס לפתרון שנבחר. תחת הנחה זו גם נקבל כי

$$\cdot b_1 = R a_1 [a_1^T R a_1]^{-1} = \frac{1}{\lambda} R a_1 = \frac{1}{\lambda} \lambda a_1 = a_1$$

נוכל להציב את הקשר של היהת הפתרון וקטור עצמי של R ולקיים כי פונקציית המחיר אותה יש להביא למינימום במבנה פשוט יותר,

$$\begin{aligned} f(a_1) &= \sum_k \left\| v_k - \frac{1}{\lambda} R a_1 a_1^T v_k \right\|_2^2 \stackrel{\substack{R a_1 = \lambda a_1}}{=} \sum_k \left\| v_k - a_1 a_1^T v_k \right\|_2^2 = \\ &= \sum_k v_k^T (I - a_1 a_1^T) (I - a_1 a_1^T) v_k = \\ &= \sum_k \left[\|v_k\|_2^2 - 2 v_k^T a_1 a_1^T v_k + v_k^T a_1 a_1^T a_1 a_1^T v_k \right] = \sum_k \left[\|v_k\|_2^2 - |a_1^T v_k|^2 \right] \end{aligned}$$

בפיתוח הניל השתמשנו בעובדה ש- $|a_1^T v_k|^2 = \|a_1\|_2^2$. מבוטוי זה נובע כי מציאת מינימום לפונקציית המחיר זו שköלה למציאת מקסימום לפונקציית המחיר

$$\cdot g(a_1) = \sum_k |a_1^T v_k|^2 = a_1^T \left[\sum_k v_k v_k^T \right] a_1 = a_1^T R a_1$$

המקסימום של ביטוי זה (בכפוף לאילוץ $\|a_1\|_2 = 1$) הינו הווקטור העצמי המתיחס לערך העצמי הגדול ביותר של המטריצה R , וכך הנחה שעשינו לפני מספר שלבים על אופיו של הפתרון מתאמת.

נניח כי עשינו כל זאת וקיבלו את השורה הראשונה ב- A (והעמודה הראשונה ב-B). נוכל לחזור על התהליך ולהפוך את השורה השנייה ע"י הבאה למינימום של פונקציית המחיר

$$\cdot f(a_2, b_2) = \sum_k \left\| v_k - (b_2 a_2^T + b_1 a_1^T) v_k \right\|_2^2 = \sum_k \left\| v_k - b_2 a_2^T v_k - b_1 a_1^T v_k \right\|_2^2$$

כשהפעם אנו מנהים כי הCMD [ב,ג] נתון. תחילה אלגברי דומה לזו שתואר לעיל יביא לבחירת \underline{a}_2 כוקטור העצמי המתייחס לערך העצמי השני של המטריצה R , וכך תחילה זה ממש. נקודה מעניינת היא שנדמה כי הפתרון שנקראם כאן יהיה תת-אופטימלי בשל העובדה התהיליך סדרתי – למעשה אין זה כך – אם תיעשה האופטימיזציה בצורה מקבילת על כל המטריצה A (ו- B) נגיע לאותו פתרון ממש.

נחזיר לרגע לתוצאה שקיבלנו במהלך הניתוח ושנראתה כפרט טכני. מצאנו כי כתחליף לביעית מינימיזציה מורכבת יש לנו בעיית מקסימיזציה בה פונקציית המחיר היא

$$g\{\underline{a}_1\} = \sum_k \left| \underline{a}_1^T \underline{v}_k \right|^2 = \underline{a}_1^T \left[\sum_k \underline{v}_k \underline{v}_k^T \right] \underline{a}_1 = \underline{a}_1^T R \underline{a}_1$$

לביטוי זה משמעות גיאומטרית מעניינת – בהינתן מקבץ הדוגמאות $\{\underline{v}_k\}_{k=1}^N$ אנו מחפשים כיוון הטלה \underline{a}_1 עליו יתקבל כי הפיזור הוא מרבי! ציור 1. B. ממחיש בחירה זו בדו-מימד.



תכונה נוספת מעניינת היא התכונה הבאה: בהינתן וקטור כלשהו \underline{u} , נוכל לייצג אותו כצירוף ליניארי של וקטוריים עצמאיים אלה ולכתוב

$$\underline{v} = \sum_{n=1}^N v_n \underline{a}_n$$

ההטמרה ע"י A פירושה המכלה במטריצה ששורותיה הן \underline{a}_n^T , ובשל העובדה הווקטורים העצמיים אורתונורמליליים, המכלה זו תיתן את אוסף המקדמים v_n , וזה תהיה תוצאה ההטמרה $\underline{u} = A \underline{v}$. אם נחשב את את מטריצת הקוריאנס של מוצא ההטמרה נקבל

$$R_U = \sum_k u_k u_k^T = A \sum_k v_k v_k^T A^T = A R A^T = \Lambda$$

בשחמטריצה Λ היא מטריצה אלכסונית המאוכסלת בערכים העצמיים של R . יצא כי פועלות ההתרמה הלבינה את המקדים – יקרה וקטורי ייצוג חדשים בהם אין קורלציה בין איברים שונים.

ICA ל- PCA 2.

התמרת ה- KLT ידועה לטטיטיסטיקים בשם PCA, כשם זה בא לתאר בדיקת שראיינו קודם – אנו מוצאים את הכוונים העיקריים של הנתונים הוקטורים לבניית התמורה זו. יש חסרונות רבים להתרמה זו, אך יתרונה הגדל של התמורה זו בפשטותה החישובית ובኒקיוו המתמטי המוביל לפתרון סגור – עובדות המסבירות את הפופולריות הרבה של שיטה זו. בשורש השנים האחרונות נעשו ניסיונות שונים להכליל את ה- PCA. המשותף לכל הכללות אלו הוא אובדן הפשטות!

הכללה ראשונה שנציג היא ניסיון למנוע מחריגים להשפיע על בניית התמורה. אם נתיחס לבניית כיוון הטעלה הראשוני, הרי שהביטוי המקורי המובא למקסימום הוא

$$g\{\underline{a}\} = \sum_k |\underline{a}^T \underline{v}_k|^2$$

ועבור וקטור בעייתי הנתון ערך חריג יוסט כיוון הטעלה. כתחליף, שימוש בפונקציה

$$g\{\underline{a}\} = \sum_k [h(\underline{a}^T \underline{v}_k)]^2$$

כאשר (x) מרכז ערכים גדולים תוכל בהחלט לפטור ולוחקית בעיה זו. לדוגמה, שימוש בפונקציה

$$h(x) = \begin{cases} T & x > T \\ x & |x| \leq T \\ -T & x < -T \end{cases}$$

תשרת מטרה זו. כמובן שאז צפה ועולה בעיית המקסימיזציה של הפונקציה החדשה, שהרי אין זו עדין בעיית ערכים עצמים. הפתרון המסתמן בעבודות רבות היא גישת הגראדיינט הסטוכסטי בה למעשה

מעדכנים את \underline{a} לפי כל וקטור \underline{v}_k בנפרד ע"י הגרדייאנט של הפונקציה $h(\underline{a}^T \underline{v}_k)^2$. מתקיים

$$\cdot \underline{a}_{k+1} = \underline{a}_k + \alpha \frac{\partial h(\underline{a}^T \underline{v}_k)^2}{\partial \underline{a}} \Big|_{\underline{a}_k} = \underline{a}_k + \alpha h(\underline{a}_k \underline{v}_k) \underline{v}_k$$

השיטה נקראית גראדיינט סטוכסטי כיוון שההنجichות לאוסף הוקטורים היא כסט מדידות מצטבר לפונקציית מחיר שבסכום שלב נתון אינה ידועה, ורק בצבירת אינסוף וקטורים יש לה משמעות דטרמיניסטית.

גישה דומה מאוד ניתן להציג בה תתקבל אינטראפטציה אחרת למגמי – אם נניח כי במקום שימוש בפונקציה h ייחידה נציג פונקציית מחיר מהצורה

$$g\{\underline{a}\} = \sum_k [h_1(\underline{a}^T \underline{y}_k)]^2 + \sum_k [h_2(\underline{a}^T \underline{y}_k)]^2 + \dots$$

נקבל תוצאה מעניינת. לדוגמה, עבור $x = h_1$ המופעלת לבדה ראיינו כי קיבלנו הלבנת מוצא ההתרמה. הלבנה זו פירושה ביטול הקורלציה בין מקדמי ההתרמה. בוקטורים אקראיים גאוסיים חוסר קורלציה שקול לחוסר תלות סטטיסטית (שני משתנים בלתי תלויים סטטיסטית אם $(x,y) = p(k)$). בכל משתנה אקראי אחר אין זה כך. לכן, אם היחיד הוא הורדת תלות סטטיסטית ולא רק ביטול קורלציה, הרי שההתהיליך מורכב יותר לאין ערוך ושימוש מספר פונקציות h יכול להיות חיקוי לא רע כלל להשגת יעד חדש זה. קיימת התוכונה הבאה לשני משתנים בלתי תלויים סטטיסטית:

$$\forall h, E\{h(x)h(y)\} = E\{h(x)\}E\{h(y)\}$$

עד כה התרכזנו בביטול הקורלציה ע"י שימוש ב- $x = h_1$ (שהרי משתנים חסרי קורלציה מאופיינים בתכונה $E\{xy\} = E\{x\}E\{y\}$). לכן, אם נבנה פונקציית מחיר כמפורט, נתקרב יותר לחוסר תלות סטטיסטית. האלגוריתם במקרה זה גם הוא ישבן על גרדיאנט סטוכסטי, אך יכול פשוט יותר איברים בעדכון התוצאה.

אם אמנס רצוננו בביטול תלות סטטיסטית, הגישה הישירה לעד זה היא שימוש במושג האנתרופיה המשותפת. אנו לא נציג נושא זה כאן, והספרות בתחום זה ענפה.

נספח ג' – הקצאת סיביות אופטימלית בדחיסה

נושא יפה שעולה בדחיסת אוטות הוא נושא הקצאת סיביות אופטימלית. הרעיון הבסיסי הוא חלוקה חכמה של תקציב סיביות נתון לשם קבלת שגיאה מינימלית. מעניין לציין כי למטרות יופיו של רעיון זה, הוא אינו מיושם בסטנדרט ה- JPEG שם יכול היה להתאים. בפועל מנגנונים אחרים בסטנדרט זה נותנים אפקט דומה. בנספח זה נתאר את הבעיה ואת דרך פתרונה.

1. הצגת הבעיה – דחיסה מבוססת התמරה

שיטת דחיסה עילית בתמונות מציעה חלוקת התמונה לבלוקים בגודל קבוע (נניח 8 על 8), וביצוע התמരה על כל בלוק. רצוייה התמരה קרובה ל- KLT על מנת לאפשר קיצוץ מוקדים עם נזק מינימלי – למידע הבלוק. ראיינו בפרק 7 כי שימוש ב- DCT מקרוב היטב את ה- KLT, ומביא עימיו יתרון חשוב – קיומו של אלגוריתם מהיר וקבוע (כלומר ללא צורך לשדר את תוכנות התמരה) ליישום התמരה. אנו נניח כי בטרם בוצעה התמരה ה- DCT לכל בלוק, הוחסר ממוצע הבלוק וקודד בנפרד (כפי שבפועל יש לעשות). פירוש הדבר שבכל בלוק יש 63 מוקדים (כי מוקדם ה- DC נוטרל).

כעת בידנו N בלוקים כשבכל אחד 63 ערci התמരה, ורצוננו בהקצאה חכמה של תקציב הסיביות בין מוקדי התמരה לקבלת שגיאה מינימלית. אנו מניחים כי חוקיות הקצאת הסיביות תהיה זהה מבлок לבלוק, דהיינו, מוקדם [3,4] בכל הבלוקים קיבל כמות סיביות זהה. הנחה זו הגיונית למדי כיון שמקדם במיקום זהה בבלוק הינו בעל חשיבות דומה. טיפול חכם יותר יניח כי תמונה מכילה בלוקים מקורות שונים לגמרי (בלוקים הלkopים ממוקם, אזורי שפה, אזורי חלקים, אזורי רעש) ונכון יהיה לבצע הקצאת סיביות שונה למוקדי התמരה לכל סוג בלוק בנפרד. במקרה זה רצוי לבצע אישכול (Clustering) על בלוקי התמരה, ואז לתכנן את הקצאת הסיביות לכל קבוצה בנפרד.

נחזיר למקרה בו בידנו קבוצת בלוקים עליה מתכנן הקצאת סיביות קבועה. כאמור, בידנו תקציב סיביות, ועל כל קבוצה של בלוקים בין המוקדים יש לדרש שסקס סיביות מותאים לתקציב ממנו יצאו – דבר זה יהווה אילוץ בעיית האופטימיזציה אותה אנו עומדים להגדיר. בהינתן כמות סיביות מסוימת למוקדם התמരה, ובידעה של פילוגו (שהרי יש לנו N מוקדים כאלו ונitin להתייחס להיסטורמה שלהם, או לפיזורם), נוכל לקבוע את השגיאה הצפואה בקונטיזציה סקלרית של מספרים אלו. לצורך זה נוכל לעשות שימוש בקשרים אותם קיבלנו בנספח א' למשל. לכן, נוכל לחשב את השגיאה הכלולות ולהביאה למינימום בכפוף לאילוץ שהזכרנו. מטרת תהליך זה היא לקבוע 63 מספרי קסט של כמות סיביות לכל מוקדם התמരה, ואלו ינבעו כפתרון בעיית האופטימיזציה. נציג כעת את הרכיב הבעייה ופתרונה בצורה מתמטית מסודרת.

2. הציגת הבעיה – ייצוג מתמטי

נתון לנו סט של קטגורים אקראיים - $\{\underline{x}_k\}_{k=1}^L$. כל וקטור כולל N איברים חסרי קורלציה ביןיהם המפולגים גאוסיאנית עם ממוצע אפס ושונות σ_k (במילים אחרות, בדוגמה שהזיכרון קודם של קידוד תמונה, יש לנו $L=63$ קבוצות ובכל קבוצה N מספרים סקלריים בעלי שונות ידועה אשר מייצגים את אותו מקדם התמורה בבלוקים שונים). נתונים לנו B סיביות לקידוד כלל כל NL המספרים הנתונים. נסמן ב- $\{n_k\}_{k=1}^L$ את כמות הסיביות שתוקצה למקדם ה- j . לכן, האילוץ שיש על אוסף הנעלמים לקיים הוא

$$N \cdot \sum_{k=1}^L n_k = B$$

כלומר אילוץ ליניארי פשוט שדורש שסך הסיביות יתאים לתקציב הסיביות הכלול.

אנו נניח כי שגיאת הקוונטייזציה (שgiaה ריבועית ממוצעת) של משתנה אקראי גaussiy בעל ממוצע אפס ושונות σ אשר קודד ב- k מ סיביות נתונה ע"י הביטוי

$$E^2 = \text{Const} \cdot \sigma_k^2 \cdot 2^{-2n_k}$$

לכן, השגיאה הכלולת בבלוק תהיה

$$E_T^2 = \sum_{k=1}^L \text{Const} \cdot \sigma_k^2 \cdot 2^{-2n_k}$$

הבעיה שנרצה לפתור היא לפיך

$$\underset{\{n_k\}_k}{\text{Min}} \quad \sum_{k=1}^L \text{Const} \cdot \sigma_k^2 \cdot 2^{-2n_k} \quad \text{Subject to:} \quad N \cdot \sum_{k=1}^L n_k = B$$

פתרון בעיה זו בעזרת כופלי לגרנזי' נתן

$$\begin{aligned} L\{n_1, n_2, \dots, n_L\} &= \sum_{k=1}^L \text{Const} \cdot \sigma_k^2 \cdot 2^{-2n_k} + \lambda \left(\sum_{k=1}^L n_k - \frac{B}{N} \right) \\ \Rightarrow \quad \frac{\partial L}{\partial n_k} &= -2 \ln 2 \cdot \text{Const} \cdot \sigma_k^2 \cdot 2^{-2n_k} + \lambda = 0 \\ \Rightarrow \quad n_k &= \frac{1}{2} \cdot \log_2 \left(\frac{2 \ln 2 \cdot \text{Const} \cdot \sigma_k^2}{\lambda} \right) \end{aligned}$$

קיבילנו כי הקצאת הסיביות תתגדל עם ההגדלת השונות – דבר צפוי למדי. הצבהchorה לאילוץ נתן

$$\begin{aligned} \frac{B}{N} &= \frac{1}{2} \cdot \sum_{k=1}^L \log_2 \left(\frac{2 \ln 2 \cdot \text{Const} \cdot \sigma_k^2}{\lambda} \right) = L \cdot \log_2 \left(\frac{2 \ln 2 \cdot \text{Const}}{\lambda} \right) + \log_2 (\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2) \\ \Rightarrow \lambda &= 2 \ln 2 \cdot \text{Const} \cdot 2^{\frac{1}{L} \log_2 (\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2)} \cdot \frac{B}{NL} \end{aligned}$$

והצבת תוצאה זו חוזרת לביטוי של הקצאת הסיביות נתנת

$$\begin{aligned}
n_k &= \frac{1}{2} \cdot \log_2 \left(\frac{2 \ln 2 \cdot \text{Const} \cdot \sigma_k^2}{\lambda} \right) = \frac{1}{2} \cdot \log_2 \left(\frac{\sigma_k^2}{2^L \log_2(\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2) - \frac{B}{NL}} \right) = \\
&= \frac{1}{2} \cdot \log_2(\sigma_k^2) - \frac{1}{2L} \log_2(\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2) + \frac{B}{NL} = \\
&= \frac{1}{2} \cdot \log_2 \left(\frac{\sigma_k^2}{\sqrt[L]{\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2}} \right) + \frac{B}{NL}.
\end{aligned}$$

התוצאה שהתקבלה מעניינת ובעל הרכב אינטואיטיבי. הגורם NL/B אינו אלא כמות הסיביות המוצעת לכל מקדם – יש B סיביות שמתחלקות בין NL מקדים. והיה וכל השונוויות שוות, נῆפה כי חלוקה אחידה זו היא שתתקבל. ואמנם, הגורם הראשון קובע את משקלת היחסית של השונוות ה- k -ית מול המוצע הגיאומטרי של כל השונוויות. אם כל השונוויות זהות גורם זה יהיה 0, וההקצאה תהיה אחידה. לכל מבנה אחר נקבל ההקצאה דינמית המעדיפה מקדים גבוהים שונות, כפיו.

בעיה טורנית אחרת עם תוצאה זו היא העובדה שלא כפינו על הסידירה לנתן מספרים אי-שליליים. בהחלטת יתכן מצב בו מתקיים הקצאה שלילית עבור מקדם התמורה בעל פיזור קטן מאוד ביחס לאחרים. במקרה כזה, במקומות לשלב אילוץ חיוביות בתוך כופלי הלגרנז' ניתן פשוט לפתרו פעמיים וראשונה כמתואר, ופעמי שנייה כשהאנו משתמשים את המקדים שזכו להקצאה שלילית, וקובעים את הקצאותם לאפס. על היותר נבצע הקצאה אופטימלית כבנוסחה הניל עם L קטן יותר (כי חלק מהמקדים הוצאו). יתכן כי נדרש למספר שבבי פתרון בצורה זו. לב העניין הינו פשוט קביעת סף על השונוויות לפיו כל מקדם בעל שונוות קטנה מס' מסוימים מקבל 0 סיביות ומושפט מהתהילה החישובי. בעיה שנייה בתחילת שהוצע כאן היא הצורך בערכי הקצאה שלמים. בפועל, תוצאה ההקצאה עברת העגלה, ואז נבדק האילוץ. במקרה של סטייה גדולה מדי ניתן לחזור על התהילה החישובי עם B שונה לשם קבלת תוצאה מתוקנת.

מה תהיה השגיאה המינימלית שתושג? נציב את נוסחתה ההקצאה האופטימלית לתוך ביתוי השגיאה ונקבל

$$\begin{aligned}
E_T^2 &= \sum_{k=1}^L \text{Const} \cdot \sigma_k^2 \cdot 2^{-2n_k} = \text{Const} \cdot 2^{-2B/NL} \cdot \sum_{k=1}^L \sqrt[L]{\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2} \\
&= \sqrt[L]{\sigma_1^2 \cdot \sigma_2^2 \cdot \dots \cdot \sigma_L^2} \cdot L \cdot \text{Const} \cdot 2^{-2B/NL}
\end{aligned}$$

אנו רואים מספר תופעות בביתי זה :

- א. סך השגיאה פרופורציוני למוצע הגיאומטרי של השונוויות המקדים.
- ב. סך השגיאה פרופורציוני באופן אקספוננציאלי הפוך לתקציב הסיביות למקדם בודד.
- ג. אם נניח כי כל השונוויות שוות נקבל כי השגיאה מתלכדת עם הנוסחה המקורית של שגיאת קוונטיוזציה לסקלר מוקוד.

כנגדה אחרתה חשוב להבהיר כי קוונטיוזציה שתיעשה תפעל בהנחה ממוצע אפס על המשתנים האקראים שם לא כן, נדרש להعبر את המוצע לצורך הפענוח. כמו כן, סידירת השונוויות חייבה להיות מוקודדת אף היא על מנת לאפשר קביעת רמות הייצוג במקלט.

נספח ד' – נושאים מתקדמים בשחזור תמונות

בנספח זה נזכיר מספר נושאים מעניינים בשחזור תמונות שלא נדונו בפרק 6. נתחיל עם שחזור מובוס אנטרופיה מרבית אשר מהוות אלטרנטיבה לגישת ה- ML וה- MAP. גישה זו עיילה ופופולרית בטיפול בתמונות בחקר אסטרונומי בהן הרקע כהה מאוד ואחד. נושא שני שייסкар הוא נושא השיחזור העיור בו פונקציית הטשטוש אינה ידועה. התוצאות בתחום זה רבות ורובן מורכבות למדי (וקשורות באורח פלא ל- ICA שהוזכר בנספח ב'). אנו נתמקד ביסודות בלבד ונראה גישות שונות לתקיפת הבעיה. נושא שלישי שנזכיר הוא בעיית denoising – הורדת רעש מתמונה. אנו לכאורה יודעים הכל על נושא זה בהיותו מקרה פרטי של בעיית השיחזור, אך נראה כי עם פישוט הבעיה ושימוש בהתמרת Wavelet נוצר מרווח מעניין לתוצאות מرتתקות. נושא אחרון בו ניגע בקשר הוא שילוב מעניין בין שחזור תמונות (או למעשה פתרון בעיות היפוך) ובין למידה חישובית. שילוב זה כומח מבניית Prior מובוס דוגמאות.

1. שחזור מובוס אנטרופיה מינימלית

אנו ראיינו כבר דרכים לפתרון בעיית השיחזור של התמונה \underline{X} ממדידתה \underline{Y} הנתונה ע"י $\underline{Y} = \underline{H}\underline{X} + \underline{N}$, כאשרנו מניחים כי H מטריצה הפועלת כטשטוש התמונה ו- N הוא רעש אקראי מפולג גausianית עם ממוצע אפס ומטריצת קורוריанс $I_n^2 \sigma$ (זהיינו, רעש לבן). ראיינו בפרט כי ה策חה במשימת השיחזור תלואה בבחירה נאותה של רגולרייזציה ובלעדיה (שימוש ב- ML) יניב תוצאות עלבות.

בצלום אסטרונומי של גורמי שמיים התמונות הנרכשות בטלסקופ סובלות מהפרעות טשטוש ורעש ולפיכך הקשר $\underline{Y} = \underline{H}\underline{X} + \underline{N}$ מתאים היבט. לעומת זאת, טיפול בתמונות אלו ע"י כלי הרגולרייזציה שהוצעו בהם כפינו חלקות אינן מוצלח כיון שבתמונות כאלו המבנה בדרך כלל כולל רקע כהה מאוד אשר מאלל את רוב שיטתה של התמונה (רמות אפור קרובות לאפס), ושיאים בהירים קטני שטח המיצגים גורמי שמיים. ישנה סבירות רבה שטיפול קונונציוני שמתאים לתמונות בעלות מבנה-Piece wise-smooth מהסוג שמקובל בצלום דיגיטלי ביתי יביא למחיקת כוכבים בתמונה, וזה כמובן אינו רצוי. לכן עלינו להחליף את אסטרטגיית הרגולרייזציה לתמונות אלו.

אם ניקח את הוקטור הנעלם \underline{X} ונתיחס אליו לרוגע כפונקציית הסתברות (אחרי נירמול ע"י חלוקה בסכום איבריו), נוכל לטעון שתי טענות חשובות אשר להן קשר להצלחת תהליך השיחזור:
א. כמו כל פונקציית הסתברות, כל ערכי הפיקסלים חייבים להיות אי-שליליים! תcona זו מתאימה לרצונו לשחזר תמונה עם רמות אפור אי-שליליים. נוכל כמובן לחזור לכלי ה- ML וה- MAP ולהוסיף תcona זו כאילוץ בעיית האופטימיזציה שמובילה לשחזר. כפיטת אילוץ זה קלה – בתהליך איטרטיבי מקובל לאחר כל עדכון לאפס את הערכים השליליים. בהתייחסות לתמונה כפונקציית הסתברות אנו פטורים מאילוץ זה – הוא משולב באופן טבעי כפי שנראה. במקרה של תמונות בהן הרקע קרוב לאפס, כפיטת אילוץ זה תורמת הרבה לאיכות התוצאה כיון שאילוץ ML אשר ייתן תופעות צלולים יאלץ להעלימן אם מדובר בשפה בין רקע כהה מאוד שאינו אפס ובין شيئا בהיר.

ב. ההתייחסות ל- \underline{X} כפונקציית הסתברות כמוות כמו לומר – מה ההסתברות למצוא בפיקסל מסוימים כוכב – ערך בהיר פירושו הסתברות גבוהה, ורוב הרקע מתייחס להסתברות קלושה לקיום כוכב. על חוק פילוג זה שהוגדר נוכל לומר כי אנו מצפים שיהיה בעל אנטרופיה נמוכה מאוד. כזכור מפרק 8, אנטרופיה תגיא לשיאו עבור פילוג אחד ותקנון ככל שחוק הפילוג מתקרב להתנהגות דטרמיניסטית. קיומם של מספר شيئاים בודד בתמונה ייתן אנטרופיה מאוד נמוכה וזו תועדף על מספר شيئاים זהה אך עם מריחה.

נניח אם כך את בעיית השחזר שbideno כבעית מינימיזציה של אנטרופיה בכפוף לאילוץ הנובע מהמידדות.

$$\text{Min}_{\underline{x}} - \sum_{k,j} x[k,j] \cdot \ln x[k,j] \quad \text{Subject to: } \|\underline{H}\underline{x} - \underline{Y}\|_2^2 = \sigma^2$$

או ביטוי מקביל בו האילוץ מוצג כתוספת לפונקציית המחיר עם מקדם 5 שהינו כופל לגרניי של האילוץ)

$$\text{Min}_{\underline{x}} - \sum_{k,j} x[k,j] \cdot \ln x[k,j] + \frac{\lambda}{2} \|\underline{H}\underline{x} - \underline{Y}\|_2^2$$

נשים לב לתכונה מעניינת שתסייע לנו – עניין נרמול \underline{X} . מקובל להניח כי אופרטור הטרטוש משמר ממוצע, כלומר סכום איברי הגרעין המטבש הוא 1. לכן, בהינתן ממוצע התמונה \underline{X} יהיה שווה בקירוב טוב לממוצע התמונה \underline{Y} והרי זה ערך ידוע. לכן, אם נתחל את התהליך ע"י חלוקת \underline{X} בסכום איבריה נביא אותה גם כן לבנייה של חוק הסתברות, ומכאן ואילך נוכל לכפות על \underline{X} להיות גם היא מנורמלת. מציאת המינימום תיעשה ע"י גזירת הפונקציה. גזירת ביטוי האנטרופיה לפי איבר יחיד ייתן $\log x + 1$.

$$\frac{\partial}{\partial \underline{x}} \left[\sum_{k,j} -x[k,j] \cdot \ln x[k,j] + \frac{\lambda}{2} \|\underline{H}\underline{x} - \underline{Y}\|_2^2 \right] = -1 - \log \underline{X} + \lambda \underline{H}^T (\underline{H}\underline{x} - \underline{Y}) = 0$$

משוואה זו סטומה וקשה לפתורן בדרך ישירה, אך נוכל הציע תהליכי איטרטיביים שונים לבניית הפתרון. לדוגמה, שימוש באלגוריתם ה- SD יביא לאיטרציה מהצורה

$$\underline{X}_{n+1} = \underline{X}_n - \mu \left[-1 - \log \underline{X}_n + \lambda \underline{H}^T (\underline{H}\underline{x}_n - \underline{Y}) \right]$$

גישה אחרת קרויה שיטת החלפה. שיטה נומրית קלאסית לפתור משוואות סטומות מהצורה $\underline{X} - \underline{X} = \Phi(\underline{X}_n)$.

$$-1 - \log \underline{X}_{n+1} + \lambda \underline{H}^T (\underline{H}\underline{x}_n - \underline{Y}) = 0 \Rightarrow \underline{X}_{n+1} = \exp \left\{ \lambda \underline{H}^T (\underline{H}\underline{x}_n - \underline{Y}) - 1 \right\}$$

הרעיוון במשוואה זו הוא שאנו נוכל להבטיח שתתקבל התכונות של סידרת הפתרונות, נקבל עבור מספר רב של איטרציות כי $\underline{X}_n = \underline{X}_{n+1}$ וזו יתקבל פתרון שמקיים את איפוס הגרדיאנט כרצוי.

משמעותו לציין כי אי-שליליות הפתרון תנבע ישירות מפעולת האיטרציה, ולכן לא נדרש לכפottaה באמצעות נוספם. לעומת זאת, ניתן כי סכום איברי \underline{X} לאחר האיטרציה לא יהיה 1, ולכן לאחר כל איטרציה יש לנורמל את הוקטור.

2. שחזור עיוור

לכל אורךו של פרק 6 הנחנו כי במודל הקלקול $\underline{N} + H\underline{X} = \underline{Y}$ אנו יודעים במדויק את פונקציית הטשטוש. באופן מעשי זה לרוב לא נכון, עם קביעה שרירוטית של המטריצה H אנו יכולים לגרום לשגיאות שחזור קשות. קיימים היקף גדול של עבודות תחת השם "שחזור עיוור" שבאים להוריד הנחה זו, וモוצעים דרכים שונים (ומשנות, חיברים להודות) לשחזור בתנאים קשים אלו.

עבודות ראשונות בתחום בחרו לנשות ולקבע את מידת הטשטוש ע"י מציאת מקומות בתמונה בהן יש נקודה בהירה על רקע כהה (או להיפך). במקרה של קיום "פונקציית דلتא" כזו, ההנחהות באזורה תהיה המשנן המטשטש. בגין נקודה בודדת ניתן לחפש שפה חדשה ולעשות זאת במספר שפות ובזויות שונות וכך לשזר את הטשטוש לפני תחילת תהליך השחזור. שיטות אלו חולשות למדי בשל תלותן במאפייני תמונה שאינם זמינים בהכרח, בשל הפרעה שהריעש יוצר, ובשל הגבר השגיאות שנבע מהפרדת הבעה לשתי בעיות – קביעת הטשטוש ולאחריה שחזור התמונה.

גישות מודרניות יותר רואו שיש לשלב את קביעת הטשטוש כחלק בתהליך השחזור. הוצעו מגוון

כלים כאלה, כשהמובילים בהם הן הגישות הבאות:

1. אלגוריתם ה- EM – זהה הרחבת של משערק הסבירות המירבית (ML) לתנאים בהם חסר מידע בעיית השערוק. המידע החסר אצלו הוא הטשטוש ולכן מבנה זה טבעי לשימוש. אלגוריתם ה- EM מציע תהליך איטרטיבי של פתרון ה- ML שבו כל איטרציה בנוייה משני נתחים – הראשון בו נקבע המידע החסר ע"י מוצע (Expectation) של פתרון ה- ML, והשני בו מוצע ה- ML עם הנחת ידיעת המידע החסר משלב קודם.
2. שיטות מבוססות ICA – ניתן להציג את בעיית השחזור העיוור כפירוק של עירוב ליניארי של אותות בלתי-תלויים-סטטיסטיות. זהה בעיתת ה- ICA, ופתרונות עבורה קיימים אם כי אינם פשוטים.
3. שיטות מבוססות הנחות מודל על הטשטוש – ממש שם שיכלנו לקבוע פונקציית הסתברות על התמונה הלא נודעת ולהיעזר בה בبنית הפתרון, ניתן להציג גישה דומה לטובת הטשטוש. אם ידוע כי גרעין הטשטוש מהוות פונקציה חלקה, או אם ידוע כי הוא קיים לאורך קו (בשל טשטוש Motion Blur), מידע זה יכול לשמש בبنית השחזור.

גישה פשוטה למדדי מהמשפחה השלישי (הנחה מודל על הטשטוש) שהוצאה ע"י Kave ו-You/notant שילוב של מציאת הטשטוש תוך כדי מציאת התמונה המשוחזרת. אם השתמש במונחים מפרק 6, לו ידענו את H היינו אמורים להביא למינימום את הפונקציה

$$\epsilon\{\underline{X}\} = \|\underline{Y} - H\underline{X}\|_2^2 + \lambda \rho\{\underline{DX}\}$$

אם נניח כי אופרטור הטשטוש קבוע במקומות ובעל מבנה חלק (כגון גאוסיאן), נוכל להתיחס לפונקציה הבאה כפונקציית המחיר למינימיזציה

$$\tilde{\epsilon}\{\underline{X}, h\} = \|\underline{Y} - H(h) \cdot \underline{X}\|_2^2 + \lambda \rho\{\underline{DX}\}$$

בביטוי זה אנו כופים על הפתרון (h) להיות חלק ממש כמו תמונה, כשהמתקדם למרכיב מחיר זה צריך להיבחר בקפידה כדי למשקל נכוון את חשיבות חלקות זו. הסימון (h) H אינו אלא בנייה מטריצת

הטשטוש מתוק גרעין הטשטוש בהנחה של אופרטור קבוע במקומות. גזירה של פונקציה זו ביחס לנעלמים תיתן

$$\frac{\partial \tilde{\epsilon}\{\underline{X}, \underline{h}\}}{\partial \underline{X}} = H^T [H\underline{X} - \underline{Y}] + \lambda D^T \rho' \{D\underline{X}\}$$

$$\frac{\partial \tilde{\epsilon}\{\underline{X}, \underline{h}\}}{\partial \underline{h}} = X_M^T [X_M \underline{h} - \underline{Y}] + \gamma D^T \rho' \{D\underline{h}\}$$

כיוון שישנה קונבולוציה בין \underline{X} ל- \underline{h} , ניתן להחליף את תפקידם במכפלה $H\underline{X}$ ולבנות מטריצה מ- \underline{X} שתכפול את הוקטור \underline{h} – זה בדיקת מה שעומד מאחורי הסימן X_M .

פתרון נומרי של בעיה זו ייעשה ע"י קביעת איתחול ל- \underline{h} ול- \underline{X} , והפעלת ירידה לפי הגרדיאנט (SD) לשירוגין בעזרת שתי המשוואות הניל עד להתקנסות. כיוון שפונקציית היעד אינה קמורה, אין כל הבטחה למציאת הפתרון הטוב ביותר, ובמידה רבה בחירת האיתחול קובעת את תוצאות הסיום.

3. הורדת רעש (Denoising)

הורדת רעש מתמונה אינה אלא מקרה פרטי של בעיית השיחזור הכללית. לכן, בשימוש בשערוך MAP עם טיפול נאות בשפות, תוצאה התהליך תהיה נקודת המינימום של הפונקציה

$$\epsilon\{\underline{X}\} = \|\underline{Y} - \underline{X}\|_2^2 + \lambda \rho\{D\underline{X}\}$$

שאלה מעניינת שנבחנה בספרות היא שאלת בחירת האופרטור D . מסתבר כי עבור בחירת D כמטריצה אורתונורמלית מתתקבל פתרון סגור ונוח אשר קל במיוחד לישום. עד כה فعلנו תחת ההנחה כי D מיישם גזירה כלשהי. נוכל להציג התמרת Wavelet אורתונורמלית כאופרטור זה – הרכיב המטריצה כולל נגירות שוונוט (מסני HPF במרוחבי תדר שונים), ולכן ראייה זו נראית סבירה. נניח אם כך כי D היא מטריצה אורתונורמלית ונסמן $\underline{X}_D = \underline{X}$. מתקיים כי מינימיזציה של פונקציית המchiaר הניל שköלה למינימיזציה של

$$\epsilon\{\underline{X}\} = \|\underline{Y} - \underline{X}\|_2^2 + \lambda \rho\{D\underline{X}\} = \|D^T (\underline{Y}_D - \underline{X}_D)\|_2^2 + \lambda \rho\{\underline{X}_D\} =$$

$$= \|\underline{Y}_D - \underline{X}_D\|_2^2 + \lambda \rho\{\underline{X}_D\} = \sum_{k=1}^N (y_d[k] - x_d[k])^2 + \lambda \rho\{x_d[k]\}$$

בפיתוח הניל עשינו שימוש בכך שnormata L_2 אינה מושפעת מאופרטור אורתונורמלי, דהיינו $\|\underline{DX}\|_2 = \|\underline{X}\|_2$. כמו כן השתמשנו בעובדה שהפונקציה מהויה סכום פונקציות סקלריות על ערכי הוקטור המוזן –

$$\rho\{\underline{V}\} = \sum_{k=1}^N \rho\{v_k\}$$

ברישום הניל אנו מניחים כי התמונה כוללת N פיקסלים וכך גם מספרם של מקדמי התמara.

יופיו של הביטוי שהתקבל נובע מכך שחייב של כל מקדם $[k]_d$ יכול להיעשות בנפרד כפתרון של בעית אופטימיזציה סקלרית פשוטה. علينا להביא למינימום את הפונקציה

$$\begin{aligned}
 f(x) &= (x - y)^2 + \lambda \rho(x) \\
 \Rightarrow f'(x) &= 2(x - y) + \lambda \rho'(x) = 0 \\
 \Rightarrow x &= y - \frac{\lambda}{2} \rho'(x)
 \end{aligned}$$

ובינתיים נראה כי יידרש פתרון נומרי לפתרון בעיה זו (אם כי הוא פשוט! בשל היות הנעלם חד-ממדי).

. Achot haubudot haomcorot biyotra b'tachos ha-Donoho-Denoising. Hayav ubodatam shel Johnston, nukodat haomzah shlahem hiiyata abchana sheam masteklim ul tmonot roushet dror makdimi ha-Wavelet shlahem, roa'im ci yishenim hraba makdimim katanim shnobuim az'or ruk merush. Le'an, diycio alu ybi'a laafeket nikui. B'milim achrot, kabu ha'shniim ci tmonah nikkia merush zricha lehiiot be'ulat makdimi Wavelet shroobim apsimim. Batragom l'shpato, hem b'horo b'fonkzia robustiy mahzora

$$\rho(x) = |x|^p \quad p \in (0,1]$$

uber $0 = k$ anu soperim b'po'el at m'shar haibrim batmara shosha mafas, vekh hotrim liyud shel diliot makdimi batmara. Afilo zo'u b'chirah la-kmora, nitnu l'ptor b'zorra sugora be'utit opfimiyyatza zo. B'makom l'hatiyach lengozot (shatu'urana be'utit b'shel ai-chalkot ha-fonkzia) nazor l'fonkziyah m'hair hamlaa shkubut b'makraha zo ci

$$\varepsilon_{\underline{X}} = \sum_{k=1}^N (y_d[k] - x_d[k])^2 + \sum_{k=1}^N \lambda |x_d[k]|^0$$

hgorem ha'shni soper m'shar ai'rim shosha mafas boktor \underline{X}_D vhaibar ha'rasun ro'ach le'ravot kibba b'in voktor zo' lo kotor ha'mididot \underline{Y}_D (sheino ala \underline{Y} la'achor batmara Wavelet). Nnich ci makdimi batmara b- \underline{Y}_D sozru b'szor yord be'dek molchel tz' shakruba'im shbehem lafas matziim batmara boktor. Am ulaino lafas ai'rim, bror ci yish le'shotot zot b'zvab ha'oktorim sm shogiaha ribouyah negromat ktnah. Le'an, ntaril um hachirah $\underline{X}_D = \underline{Y}_D$ vna'af ot haibar ha'achron, zo shlefni v'chazar chilah. Um ai'pos haibrim nban' at urk ha-fonkzia \underline{X} vnu'zor at tahlik ai'pos b'makom bo batmila ulia b'sgiahah. Am nban'ro lehishar M ai'rim r'ashonim shonim mafas yihya m'hair b'sgiahah

$$\varepsilon_M \{\underline{X}\} = \sum_{k=M+1}^N y_d[k]^2 + \lambda M$$

b'muber l-1-M ai'rim shonim mafas (b'shel ai'pos ai'ber nosf) b'sgiahah t'hia

$$\varepsilon_{M-1} \{\underline{X}\} = \sum_{k=M}^N y_d[k]^2 + \lambda(M-1)$$

am $\varepsilon_{M-1} \{\underline{X}\} > \varepsilon_M \{\underline{X}\}$ yish l'hspik at ai'pos haibrim. Drisha zo mi'tרגמת l-

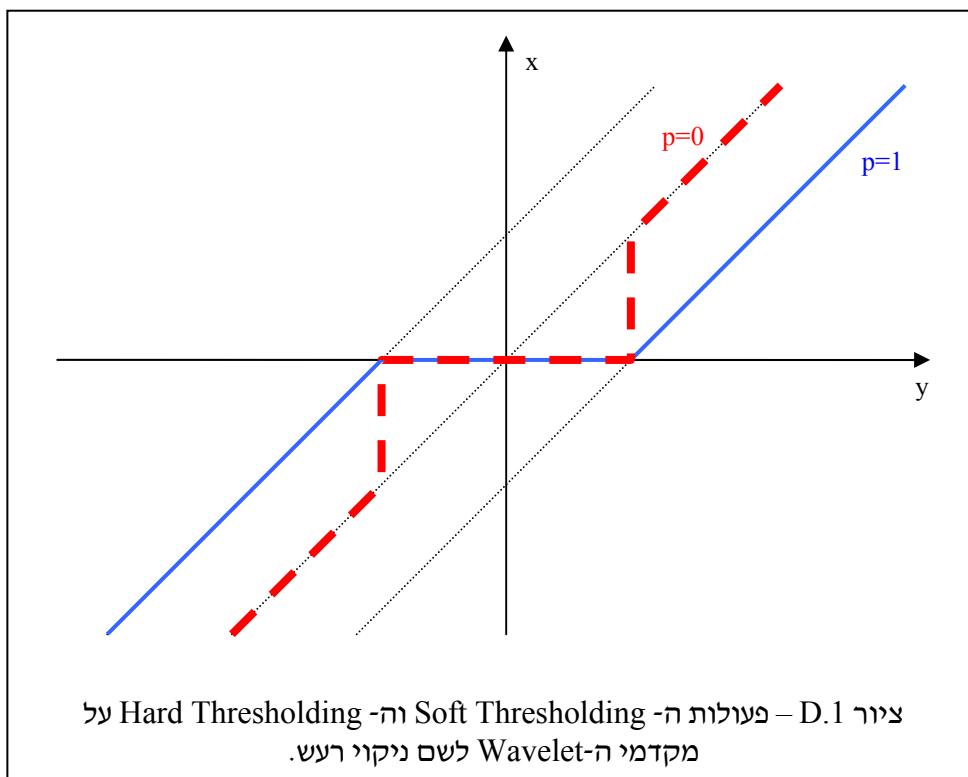
$$\begin{aligned}
 \varepsilon_{M-1} \{\underline{X}\} &= \sum_{k=M}^N y_d[k]^2 + \lambda(M-1) > \sum_{k=M+1}^N y_d[k]^2 + \lambda M = \varepsilon_M \{\underline{X}\} \\
 \Rightarrow y_d[M]^2 &> \lambda \quad \Rightarrow \quad |y_d[M]| > \frac{\sqrt{\lambda}}{2}
 \end{aligned}$$

התקבל אלגוריתם לניקוי רעש שמציע את השלבים הבאים – יש להתחיל בלקיחת וקטור המדידות \underline{y} וביצע עליו התמרת Wavelet. את כל מקדמי שערכם המוחלט מתחת $\sqrt{2}/\lambda$ יש לאפס. את הוקטור הנוצר יש להמיר חוזה ע"י התמרת Wavelet הפוכה ומתקבלת תוצאה ניקוי הרעש.

החלפת ק-ב- 1 ייתן קירוב לדليلות הרצואה ויביא לרכיב בדרכ פועלות האלגוריתם. יש לשים לב כך שבבחירה $k=2$ תעבור נגד דليلות ותעדיף ערכים רבים קטנים על פני ערך אחד גדול בהתרמה, כפי שכבר ראיינו בפרק 6, ולכן אינה רצואה. שימושים בבחירה הפורנקציה הרובשתית עם $k=1$ קיבל שפתרון בעיית הרעש בתמונה יצא כפתרון של N בעיות בלתי תלויות מהצורה

$$x = y - \frac{\lambda}{2} \text{sign}(x) = y - \frac{\lambda}{2} \rho'(x)$$

אם y גדול מ- $-\lambda/2$ אז פתרון המשווה זו הוא $y - \lambda/2 = x$. באופן דומה, אם y קטן מ- $-\lambda/2$, הפתרון יהיה $y + \lambda/2 = x$. לכל ערך באינטראול שבין שני ספינים אלה לא קיים פתרון המשווה זו בשל אי-הרציפות של פועלות הסימן, ואז נציגו כפתרון אפס לשם קבלת קרבה מרבית בין שני האגפים, ולשם קבלת עקבות עם הפתרון עבור $y=0$. ציור 1 מראה את פועלות השינוי על מקדמי Wavelet- p בשני המקרים שנוחתו.



4. שחזור מבוסס דוגמאות

נושא חשוב שצובר תאוצה ונמצא ברובו בזירה מחקרית עדין הוא נושא השימוש בדוגמאות כתחליף לפונקציית הסתברות אנליטית לתמונה. אם נחשוב על כך לרגע, יש נאיביות גדולה ברישום מהצורה

$$P(\underline{X}) = \text{Const} \cdot \exp\{-\lambda \rho \{\underline{D} \underline{X}\}\}$$

כיוון שהבחירה בביטוי אנליטי שונה לגזירה מחייב פשרות אוצריות עם הממציאות – תמונה לרוב נוטה להיות חלקה (חוץ מקרים שבהיא לא כל כך חלקה, כמו במרקמים), שפות נוטות להיות חדות (חוץ מקרים שבהם שפה שאינה חזה קורית, כמו במקרה של תנאי האורה לא נוחים), ועוד ועוד. כך יוצאה של מזבח הנוחות המתמטית אלו מקרים את הצלחת פתרון בעיית ההיפוך בה נעשה שימוש ב-*Prior* מעין זה.

ומהו האלטרנטיבה? דוגמאות! הרבה מאוד דוגמאות! מאגר עצום של תמונות שתישמרנה כדי ליצג את מרבית התופעות שאנו צופים לראות בתמונה. בהינתן דוגמאות אלה נוכל תיאורתית לנשות ולבנות את פונקציית הסתברות, אז לבצע את השחזור, אלא שגישה זו מאוד לא יעילה. נסיוון מעין זה הוצע ע"י Zhu Mumford ו-*Zhu* – הם בחרו פונקציית הסתברות מעט יותר כללית מהצורה

$$P(\underline{X}) = \text{Const} \cdot \exp\left\{-\sum_k \lambda_k \rho_k \{D_k \underline{X}\}\right\}$$

בפונקציה זו לא פועל אופרטור גזירה יחד אלא מספר אופרטורים כאלה, ועל כל אחת מתוצאות הגזירה מופעלת פונקציה רובטית אחרת. עיקר עבודתם של Zhu Mumford ו-*Zhu* הייתה פתרון השאלה – בהינתן בנק של תמונות דוגמה, האם נוכל למצוא את הבחירה הטובה ביותר ליותר לאופרטורים D_k ו- הבחירה הפונקציות הרובטיות שנלויות להן? ניתנת הפתרון אותה בחרו מרכיבת למדוי, וקרוב לוודאי שנייה פשוטה לא אובדן ביצועים. הנקודה היא שהם יצרו את הקשר בין פונקציית ה-*Prior* ובין אוסף דוגמאות. ניתנים היבטים שעורך הפרמטרי ואין זו בהכרח הגישה המוצלחת כיון שהיא כופה סדר של מבנה הפונקציה ממנו מתחילה.

אם נוכל לעשות שימוש בדוגמאות בצורה ישירה יותר? תשובה שהולכת ונركמת בימים אלה. פאן מסויים אותו נזכיר נבע מעבודות על סינטז מרקמים של Effros ואחרים. עבודה של העת الأخيرة (פורסמה חיליקית ב-2003) היא עבדותם של Weissman ו-Sarousi (בעלי עבר מהטכניון) והיא נראית כמו שביא לההフィכה בתחום זה של שימוש בדוגמאות לשם פתרון בעיות היפוך, כאשרם הנחות היסוד מורכבות עוד יותר – הדוגמאות הבן נעשו שימושו נבעות מהמידע הפגום הנוכחי לשחזור.

נציע את הגישה הבאה לשימוש במאגר התמונות, ונתיחס לצורך התיאור של השימוש במאגר בעיית ניקוי הרעש מתמונה. נבחר חלון של $(2L+1) \times (2L+1)$ פיקסלים סביב כל פיקסל במאגר, ובננה בצורה כזו מיליון עצום של וקטורים באורך $(2L+1)^2$. לכל פיקסל בתמונה המוקללת הנתונה נתיחס לחלון בגודל $(2L+1) \times (2L+1)$ סביבו, נתעלם מערך הפיקסל עצמו המצויב במרכזו ובניא מהמילון שנבנה את כל הוקטורים בהם ישנו דמיון גדול לתוכן הבלוק (למעט איבר האמצע). דמיון זה יכול להימدد כ-

MSE, או בדרכים אחרות שלוקחות בחשבון שפות, כיוונים ועוד. בעת שבידנו K דוגמאות של בלוקים בעלי רקע דומה נסתכל על פילוג הערך המצווי במרכזו וזה יאמר לנו את חוק פילוגו של הפיקסל המטופל. נניח כי מצאנו פילוג גאוסי בעל ממוצע m ושונות s . נניח כי הערך שנדד בפועל הוא y ואנו שואלים במא להחליפו. שימוש במשערץ MAP יאמר כי את x יש לקבוע לפי

$$\hat{x}_{\text{opt}} = \operatorname{Arg} \min_x \frac{1}{\sigma_n^2} (x - y)^2 + \frac{1}{s^2} (x - m)^2 = \frac{\frac{y}{\sigma_n^2} + \frac{m}{s^2}}{\frac{1}{\sigma_n^2} + \frac{1}{s^2}}$$

(אנו מניחים ידיעת שונות הרעש בתמונה הנתונה). ברור כי אין לנו מוחייבים לייצוג ע"י פילוג גאוסי וכל פילוג אחר יאפשר מציאת x אופטימלי, כשהמה שנדרש באמת הוא היסטוגרמת ערכי הפיקסל האמצאי בבלוקים שהובאו מהמילון כבלוקים קרובים.

אם נחזור על תהליך זה לכל פיקסל ונעשה זאת במספר סיבובים על פני כל הפיקסלים, תתקבל תמונה ששואבת את ערכיה מتوزع מאגר הנתונים.

לדוגמה, נניח כי בתמונה המדודה בחלון מסוים שכל תוכנו בעל רמת אפור קרובה ל- 100 הפיקסל המרכזיות הוא 255. נניח כי לקיחת הקונטקט (כל הבלוק למעטמרכזו) והבאתו בבלוקים דומים מהמאגר הניבה 10000 בבלוקים רלוונטיים וביהם פילוג המרכזיות מצבע בסבירות גבוהה על 100, יצא כי s (פייזור הערכים) קטן מאוד ולכן הערך במרכזה הבלוק יחולף ב- 100, וכך ניקינו רעש נקודתי. לרעש זה אין זכות להתקיים כי במאגר התמונות הוא נדייר עד כדי בלתי אפשרי. לא די בסיבוב אחד על הפיקסלים כיון שבסיבוב ראשון הזנו קונטקט פגום בשל הרעש שנכלל בו, ומטבע הדברים ההחלה שנבעה שימוש בקונטקט זה שגויה במקצת. סיבובים נוספים בהם הופחת הרעש יסייעו את ניקוי התמונה.

הגישה שהוצעה כאן בנויה בצורה אינטואיטיבית, והא דומה מאוד לגישה בה מילא Effros חורים המركמים באופן מוצלח למדי. ניתן לאמץ ראייה הסתברותית יותר תיאורטיבית לתהליך זה, אך אנו נעצור את הדיוון כאן.