

מבוא לעיבוד תמונות – דוח מכין חלק א'

מגישים

	1
שם:	ליאור וובצ'וק
מספר סטודנט:	207584715
קורס מעבדה:	מעבדה בחשמל 2

	2
שם:	אלמוג אדטו
מספר סטודנט:	318782976
קורס מעבדה:	מעבדה בחשמל 2

תאריך הגשה:	26-Mar-2023
סמסטר:	אביב תשפ"ג

תרגיל 1

1. הסבירו את סוגי התמונות:

– binary

– grayscale

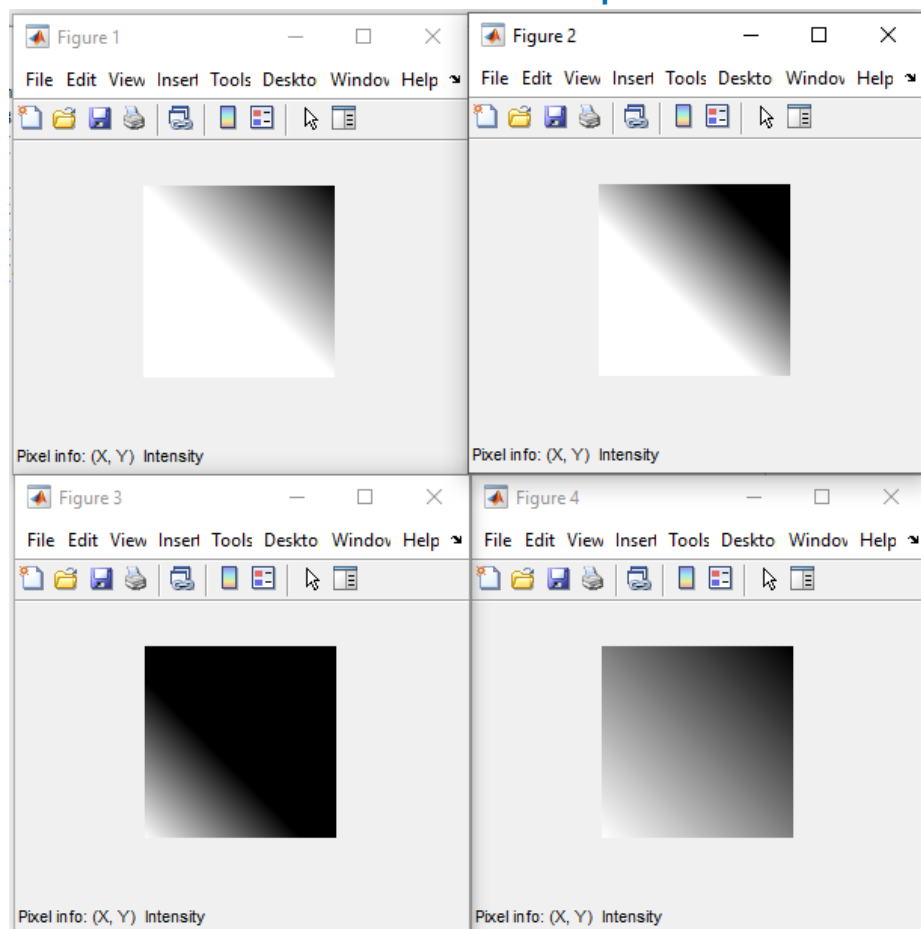
– colored

– multispectral

– Indexed

תמונה זו היא תמונה (למעשה מטריצה $m \times n$) בה לכל פיקסל משוייך מס' בינארי – 0 (המייצג שחור) או 1 (המייצג לבן).
תמונה זו היא תמונה (למעשה מטריצה $m \times n$), בה לכל פיקסל משוייך מספר בתחום מסויים. ככל שהמספר גדול יותר, כך עוצמת הפיקסל גדלה – כאשר המס' הנמוך ביותר בתחום מייצג שחור והמספר הגבוה ביותר מייצג לבן. ב-toolbox ב-matlab, כל פיקסל מקבל ערכים בטווח הרציף $[0,1]$.
תמונה זו היא תמונה (למעשה מטריצה $m \times n$), בה לכל פיקסל משוייכים 3 מספרים בתחום מסויים – אחד עבור כל ערך RGB (אדום, ירוק, כחול). ככל שהמספר גדול יותר, כך עוצמת הפיקסל גדלה. כך, על ידי שקלול ערכי הצבעים, ניתן לייצר צבעים רבים בתמונה. ב-toolbox ב-matlab, כל פיקסל מקבל ערכים בטווח הרציף $[0,1]$.
בדומה ל-colored רק שכאן בנוסף ל-3 ערכי RGB לכל פיקסל נשמרים 3 ערכי ערוצים נוספים – ערוצי InfraRed.
תמונה זו היא תמונה (למעשה מטריצה $m \times n$) בה לכל פיקסל משוייך אינדקס, המייצג שורה בטבלת מיפוי. כל שורה בטבלה מכילה 3 ערכי RGB, וכך ניתן בדומה ל-colored אך ייצוג מטריצי שונה) להגיע לצבעים רבים.

2. הציגו את התמונות שהתקבלו.



הסבירו מה מהות השוני בין התמונות (התייחסו למושג "תחום דינאמי").

התחום הדינאמי הוא תחום ערכי עוצמת הפיקסלים בו מצויים הפיקסלים בתמונה. ניתן לראות כי עבור התמונות השונות ישנם תחומים דינמיים שונים בפקודה המייצרת אותם, ולכן נראות שונה. ערכים במטריצה ממנה ייצרנו את התמונה, שהם מעל המקסימום, יקבעו לערך המקסימלי (כלומר לבן), וערכים שהם מתחת למינימום התחום הדינמי המוגדר יקבלו את הערך המינימלי (כלומר שחור).

האם התמונות שונות בערכי הפיקסלים?

התמונות זהות בערכי הפיקבלים בחיתוך התחומים הדינמיים המשותף להן. מחוץ לחיתוך, הפיקסלים הם בעלי הערך המקסימלי או המינימלי כפי שהוסבר קודם.

3. מהו התחום הדינאמי עבור התמונה כאשר הסוגריים ריקות [] ?

כאשר הסוגריים ריקות, התחום הדינאמי נקבע בין המינימום במטריצה המתקבלת למקסימום במטריצה המתקבלת – כלומר כל התחום. במקרה זה – 0 עד 255.

4. הסבירו מה מבצעת הפונקציה `impixelinfo` ומדוע היא שימושית.

פונקציה זו מאפשרת לקבל מידע על כל פיקסל במעבר העכבר עליו – הקואורדינטות שלו והעוצמה שלו. פונקציה זו שימושית כי ניתן לקבל בעזרתה ערכים מדויקים של עוצמות, כמו גם להבדיל בין עוצמות שונות שלא בהכרח ניתן להבדיל ביניהן בעזרת העין האנושית.

תרגיל 2

1. מהי היסטוגרמה (Histogram) של תמונה?

היסטוגרמה של תמונה היא כמות הפיקסלים הצבעים בצבע מסויים או עוצמה מסויימת, עבור כל צבע/עוצמה.

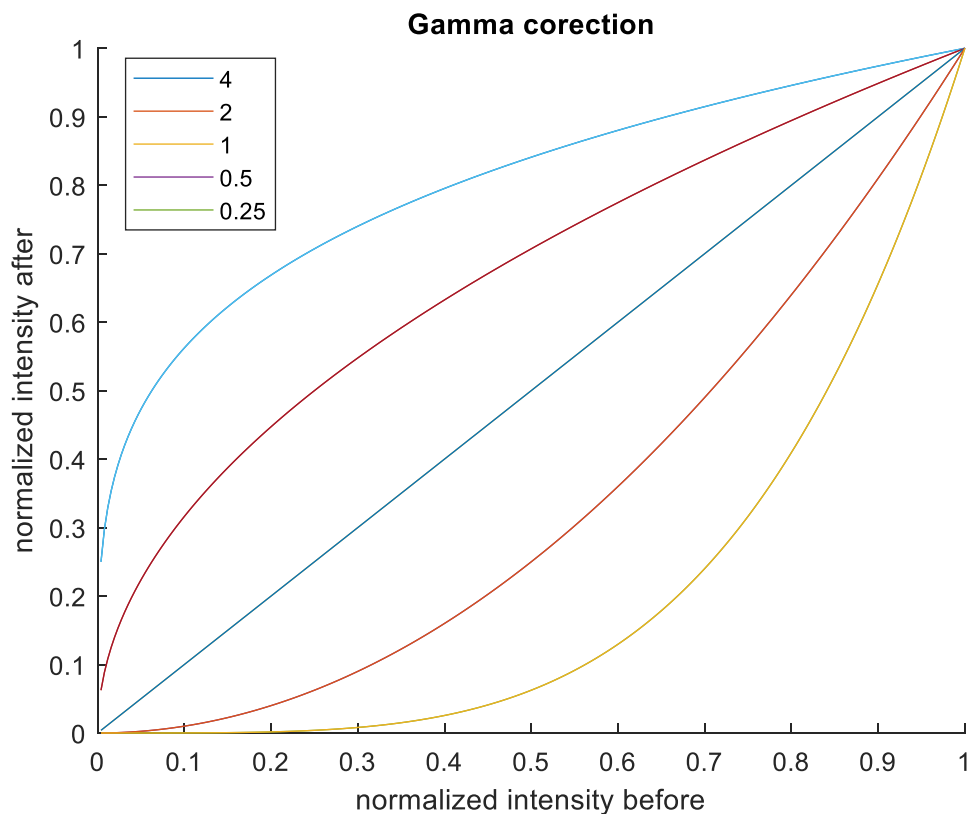
מה ניתן ללמוד מההיסטוגרמה על התמונה?

ניתן ללמוד מהי השכיחות של כל צבע. ניתן להשתמש בכך כדי לזהות תחומים בהם צבעים חוזרים יותר מהמצופה, ועל סמך כך לתקן את התמונה בדרכים שונות.

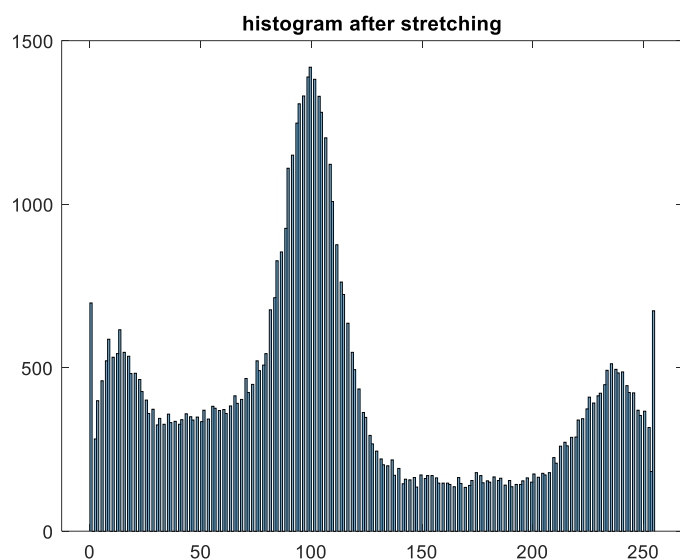
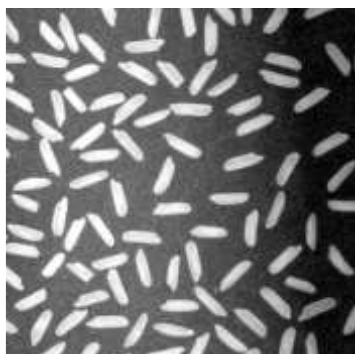
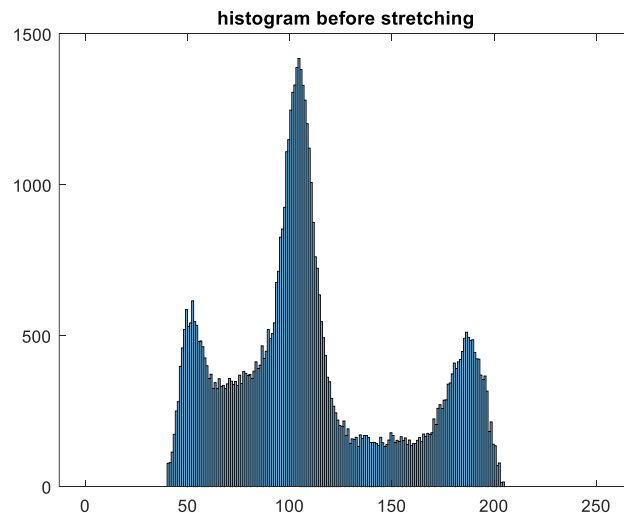
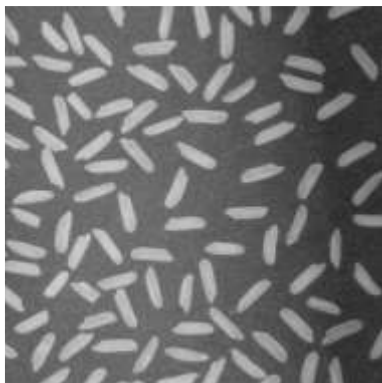
2. מדוע מבצעים מתיחת היסטוגרמה (נקרא גם Intensity Adjustment, Contrast Stretching)?

מבצעים פעולה זו בתמונות בהן התחום הדינמי אינו מנוצל במלואו. בתמונות כאלו, בהן מרבית הפיקסלים מרוכזים בתחום עוצמות מסויים, ניתן למתוח תחום זה ובכך לנצל את כל התחום הדינאמי. מכיוון שמרבית הפיקסלים מרוכזים בתחום שמתחנו, רק פיקסלים מעטים (שאיים בתחום) יקבלו את ערכי הקצה, ובאופן כללי נקבל תוצאה טובה יותר.

3. שרטטו את פונקציית ההמרה שמבצע תיקון גאמה – מוצא לעומת כניסה. ציר x יתאר את ערך הפיקסל לפני התיקון, וציר y את הערך אחריו.



4. שרטטו שתי היסטוגרמות. האחת של תמונה לפני שעברה מתיחת היסטוגרמה, והשנייה – אחריה. תארו במילים את הקשר בין ההיסטוגרמות.



ניתן לראות כי פעולת מתיחת היסטוגרמה מבצעת מתיחה של התחום הדינמי של התמונה לטווח המלא (במקרה זה 0-255) שהוא התחום הדינמי המלא האפשרי. הדבר מביא לניצול מלא של רמות האפור. ניתן לראות כי צורת ההיסטוגרמה נשמרה זהה, אך מתוחה.

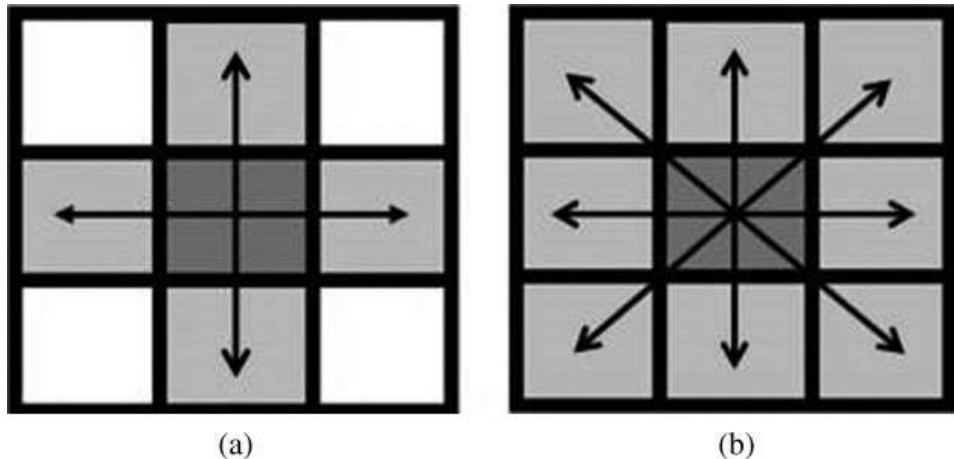
5. הסבירו מהי פעולת שוויון היסטוגרמה (Histogram Equalization), ובמה היא שונה ממתחת היסטוגרמה.

פעולת שוויון היסטוגרמה הינה פעולה המותחת את ערכי ההיסטוגרמה ומגדילה את התחום הדינמי לטווח המלא של ערכי האפור, אך בנוסף לכך מבצעת חלוקה מחדש של הפיקסלים בצורה יותר אחידה – כלומר שלכל עוצמה תהיה דומה של פיקסלים בצבע זה.

תרגיל 3

1. מהם סוגי הקישוריות (Connectivity) בתמונות? הדגמו באיור.

קישוריות בתמונות באופן כללי היא מצב בו מתייחסים לשכניו של פיקסל, ולא רק לכל פיקסל כעומד בפני עצמו. הסוגים השונים של הקישוריות נובעים מכמות וסוג השכנים אליהם מתייחסים בניתוח ועיבוד התמונה. סוגי קישוריות שונים הם למשל קישוריות-4-שכנים (תמונה a), המתייחסת רק לשכנים שמלמעלה, למטה, שמאל וימין, או קישוריות-8-שכנים (תמונה b), המתייחסת לכל 8 השכנים של פיקסל כלשהו.



כיצד משתמשים בקישוריות לגילוי עצמים?

ניתן להשתמש בקישוריות לגילוי עצמים, על ידי חיפוש שפות. כאשר בין פיקסלים מקושרים ישנו הפרש גדול (ניגודיות גדולה), ניתן להסיק כי קיימת שפה במקום זה – ובכך לגלות עצמים. כאשר ישנה קישוריות רבה באיזור מסוים (פיקסלים בעלי עוצמה דומה), ניתן להסיק כי מדובר באותו עצם.

2. מה מבצעת הפונקציה bwconncomp?

הפונקציה מקבלת כקלט תמונה בינארית (שחור ולבן), ומחזירה כפלט מבנה נתונים המכיל את קבוצות הקשירות שבתמונה. על ידי קישוריות, התמונה מסווגת את קבוצות הקשירות, ובמבנה הנתונים קיים מידע כמו מספר קבוצות הקשירות וכן הפיקסלים השייכים לכל קבוצת קשירות).

3. מה מבצעת הפונקציה labelmatrix ומה מתקבל ממנה?

פונקציה זו מקבלת כקלט את מבנה הנתונים שהוחזר כפלט מהפונקציה הקודמת, ומחזיר מטריצה בה לכל פיקסל משייך אינדקס של קבוצת קשירות מסוימת. הדבר מאפשר לקרוא ולהבין את התוצאות של הפונקציה bwconncomp בצורה יותר ברורה.

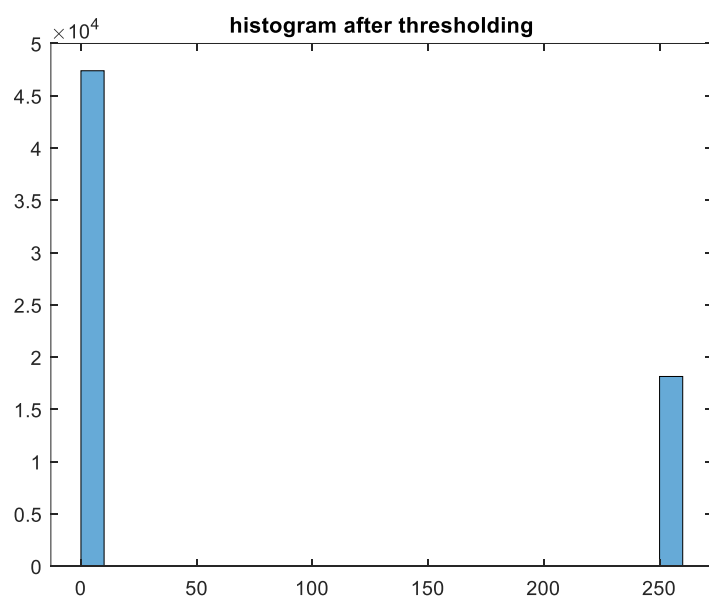
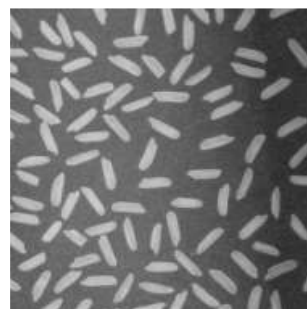
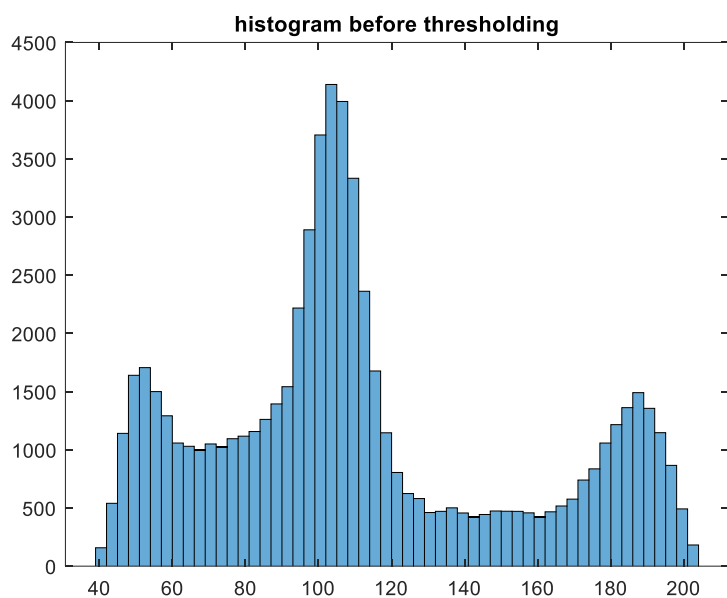
מה חשיבותן של שתי פונקציות אלו?

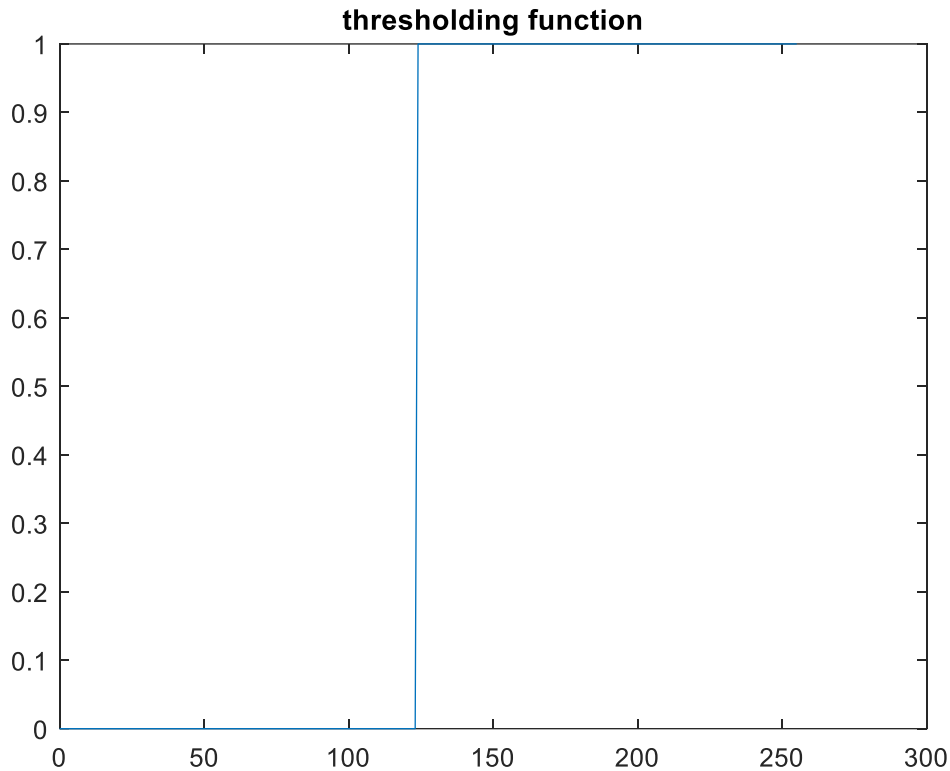
4. חשיבותן היא בכך שהן מאפשרות לסווג קבוצות קשירות בתמונה – ולמעשה כך ניתן להפריד ולגלות את העצמים השונים בתמונה.

5. מהו חיתוך בסף (Thresholding) ובאיזה מקרים הוא שימושי?

חיתוך בסף הוא פעולה שעיקרה קביעת סף עוצמה מסוים – וצביעת הפיקסלים שמעליו בלבן והפיקסלים שמתחתיו בשחור. סף העוצמה נקבע בצורה דטרמיניסטית או על ידי אלגוריתם שקובע אותו על סמך פרמטרי התמונה (למשל היסטוגרמה וכו'), וכן יכול להיות גלובאלי (עבור תמונה) או מקומי (ספים שונים לאזורים שונים בתמונה). הדבר שימושי כדי לאחד עוצמות דומות לכדי אותה עוצמה בינארית, כדי לסווג אותן מאוחר יותר לקבוצות קשירות (למשל ע"י הפונקציות שתוארו לעיל). כמו כן שימושי כדי להבליט את קבוצות הקשירות על פני הרקע.

6. תארו את השינוי ברמות ההארה של תמונה העוברת חיתוך בסף ע"י גרפים: הגרף הראשון יכיל את ההארה ההתחלתית (ציר x) לעומת הסופית (ציר y). הגרפים השני והשלישי יכילו את ההיסטוגרמה של התמונה המקורית, ושל התמונה הסופית (שרטוטים איכותיים).





7. נתונה תמונה כללית אשר בה מספר עצמים שכל אחד מהם בגוון שונה ולא ידוע, ורקע בגוון אחר לא ידוע. בידכם רכיב המבצע פעולת חיתוך בסף עם סף שערכו אפס. הסבירו כיצד ניתן לקבל תמונה בינארית יחידה המכילה את כל העצמים בלבן בעזרת הפעלת רכיב זה בלבד פעם אחת.

מכיוון שלא ניתן לקבוע מי מהצבעים הוא צבע הרקע נניח שאנחנו יודעים מהו צבע הרקע. כעת נחסיר מכל פיקסל בתמונה את ערך הצבע של הרקע ונבצע על כל הפיקסלים ערך מוחלט. כעת נבצע פעולת חיתוך בסף עם ערך 0 ורק הרקע שערכו היה 0 ישאר 0 וכל שאר הפיקסלים יהפכו ל256.

8. רשמו שתי דוגמאות לתהליכים מציאותיים הדורשים ביצוע סגמנטציה.

1. הבחנה בין גופים שונים בכביש ברכב אוטונומי – למשל בין הולכי רגל או רכבים חולפים לבין הכביש עצמו.

2. רפואי – זיהוי גידולים והבחנה ביניהם לרקמות בריאות בגוף האדם.

תרגיל 4

- 1. פרטו על סוגי הרעש הבאים והשפעתם על התמונה. ציינו את הפרמטרים לכל אחד מהם:**
 - Poisson (Shot noise) –** רעש פואסוני הוא רעש המצטרף לכל פיקסל ומפולג לפי התפלגות פואסונית. הפרמטרים הם תוחלת ושונות.
 - Salt & Pepper –** מדובר ברעש בו פיקסלים מסויימים מחליפים את עוצמתם ללבן או שחור באופן רנדומלי. הפרמטרים הם ההסתברויות של פיקסל להפוך לשחור, להפוך ללבן, או להישאר ללא שינוי.
 - Speckle -** רעש אדיטיבי שמצטרף, המפולג התפלגות אחידה בקטע מסויים. הפרמטרים הם גבולות הקטע בו מפולג רעש זה.
- 2. אילו רכיבים או פעולות במצלמה עלולים להכניס רעש (לא טשטוש) לתהליך הצילום וכיצד?**

רעש תרמי מחלקים בעלי התנגדות במצלמה, פגמים/לכלוך/שריטות על העדשה, וכו'.
- 3. פרטו על פעולתן של הפונקציות special ו-imfilter וכיצד הן קשורות זו לזו.**

הפונקציה imfilter מקבלת כקלט מסנן רב מימדי (המיוצג כמטריצה), ותמונה (מטריצה), ומחזיר כפלט את התמונה לאחר הפעלת המסנן עליה. הפונקציה special מקבלת כקלט פרמטרים של מסנן רצוי, ומחזירה כפלט את המסנן. הקשר ביניהן הוא שהפונקציה imfilter מקבלת כקלט את המסנן ש-special מייצרת.
- 4. מהן כל אפשרויות הריפוד (padding) בפונקציה imfilter? הסבירו עליהן בקצרה. מהי האפשרות הטובה ביותר לדעתכם עבור תמונות? נמקו.**

האפשרויות הן:

 - Symmetric –** הריפוד מחוץ לתמונה יהיה על ידי שיקוף הפיקסלים הקרובים לקצוות התמונה – על גבי קצה התמונה.
 - Replicate –** הריפוד מחוץ לתמונה יהיה על ידי שכפול פיקסלי הקצה של התמונה לתחומים שמחוץ לה.
 - Circular –** הריפוד מחוץ לתמונה מניח כי התמונה הינה מחזורית – כלומר ריפוד על ידי הפיקסלים הנמצאים בצד הנגדי לקצה בו אנו מרפדים.

לדעתינו האפשרות הכי טובה היא replicate, כי סביר להניח שגופים הנמצאים בקצוות התמונה ימשיכו גם מעבר לה. אין בסיס להנחת מחזוריות של תמונה או שיקוף הפיקסלים הנמצאים בקצה.
- 5. הסבירו על דרך פעולתו, יתרונו וחסרונותיו של מסנן חציון (Median).**

מסנן חציון מחזיר את ערך החציון בין כל ערכי האזור הדרוש (למשל מטריצה 3x3) עבור פיקסל מסויים. היתרון של מסנן זה הוא שניתן להתעלם מערכים גדולים שלא בהכרח קשורים לאזור התמונה (למשל רעש salt&pepper), כיוון שהם יעלמו בביצוע פעולת החציון. החיסרון הוא שהוא יכול לגרום לטשטוש התמונה.
- 6. הסבירו על דרך פעולתו, יתרונו וחסרונותיו של מסנן גאوسی (Gaussian).**

מסנן גאوسی הוא מסנן אשר ממצע את ערכי האזור הדרוש סביב כל פיקסל, כאשר המשקל של כל פיקסל שכן נקבע על פי ההתפלגות הגאוסית הנקבעת על פי מרחק. היתרון הוא שמדובר במסנן לינארי, דבר המפשט מאוד את החישובים איתו. החסרון, גם כאן, הוא שהוא גורם לטשטוש והתמונה והשפות בה.

7. טשטשו תמונת grayscale לבחירתכם בעזרת מסנן טשטוש כלשהו לבחירתכם והציגו את התוצאות.



השתמשנו במסנן גאוס'י.

תרגיל 5

1. מהי קוונטיזציה (Quantization)?

קוונטיזציה היא פעולה בה ערכי הפיקסלים השונים בתמונה מומרים לטווח בדיד, סופי ומוגדר מראש של ערכי עוצמה.

מדוע אנו צריכים לבצע קוונטיזציה?

אנו צריכים לבצע קוונטיזציה על מנת להוריד את מימד הנתונים בבעיה – כלומר את מימד ערכי הפיקסלים שהמחשב ידרש לטפל בהם. הדבר יוכל לפשט מאוד את החישובים, ויכול להיות שימושי גם לדחיסת תמונות.

2. מהו תפקידם של הפרמטרים $values$ - i levels בתהליך הקוונטיזציה?

הפרמטר $values$ מייצג את רמות העוצמה השונות אליהן נקוונטט את ערכי הפיקסלים בתמונה הדגומה.

הפרמטר $levels$ מייצג את עוצמות הפיקסלים ביניהן נבצע קוונטוט – כלומר מהן רמות העוצמה ביניהן מקוונטטים הפיקסלים.

3. הסבירו מהי שגיאה ריבועית ממוצעת (MSE).

שגיאה ריבועית ממוצעת, היא מדד מקובל למדידת "מרחק" בין שתי תמונות. המרחק נקבע על ידי סכימת ריבוע הפרש ערכי אותו פיקסל בין שתי התמונות.

4. היעזרו בפונקציה $immse$ ומדדו MSE בין שתי התמונות שקיבלתם בסעיף 7 בתרגיל הקודם:

הערך שחושב על ידי `matlab` הינו 168.8079.

מדדו MSE בין התמונה המקורית לעצמה:

0

מה המסקנה משתי המדידות?

ניתן לראות כי מדד השגיאה הריבועית הממוצעת הוא מדד טוב למדידת שוני בין תמונות. נסיק כי מזעור השגיאה הריבועית הממוצעת הוא קריטריון טוב לטיבו של מסנן, שכן אנו נשאף שפלט הסיון יהיה כמה שיותר זהה לתמונה המקורית – כלומר $MSE = 0$ כפי שקילבנו בסעיף זה.

5. מהו ההבדל העיקרי בין התמרת DCT (DCT Transform) להתמרת הפורייה, ומדוע הוא חשוב?

התמרת DCT היא התמרה הכוללת הכפלה בקוסינוסים ממשיים במקום באקספוננטים מרוכבים את איברי הווקטור. ההבדל העיקרי הוא שפלט ה- DCT הוא ממשי טהור, בעוד התמרת פוריה היא מרוכבת. הבדל זה חשוב מכיוון שיותר קל לעבוד עם ערכים ממשיים, ויש צורך בפחות זיכרון על מנת לאחסן את פלט ההתמרה (שכן משתנה מרוכב כולל שני חלקים שיש לשמור – חלק ממשי וחלק מדומה, ואילו כאן קיים החלק הממשי בלבד).

6. מה לדעתכם מקבלים כאשר מבצעים התמרת פורייה של ריבוע לבן (255) מלא על פני רקע

שחור (0)? נמקו. רמז: חשבו תחילה על המקרה החד-ממדי. מומלץ לא להיעזר ב-Matlab.

במקרה החד מימדי, מדובר בהתמרת פוריה של חלון, שהיא פונקציית ה-Sync. לכן, בדו מימד נקבל Sync דו מימדי, כלומר מעגלים מחזוריים באמפליטודה שהולכת וקטנה ככל שמתרחקים מהמרכז.

תרגיל 6

1. איזה סוג של תמונה היא תמונת שפות? מדוע?

תמונת שפות היא תמונה המכילה את השפות בלבד מתמונת קלט כלשהי. השפות הן פיקסלים בעוצמה מלאה (לבן), ושאר התמונה הינה בעוצמה 0 (שחור). המטרה בכך היא שהפיקסלים המעניינים בתמונה (השפות) יהיו בעוצמה מלאה, במקום שהפיקסלים הלא מעניינים (שהם מרבית הפיקסלים בתמונה) יהיו בעוצמה מלאה.

2. פרטו על דרכי הפעולה של האופרטורים הבאים לגילוי שפות:

Sobel – אופרטור זה מחשב את הנגזרת הראשונה (הגרדיאנט) עבור כל פיקסל (על ידי סביבתו כמובן), והפיקסלים להם גרדיאנט גדול יסווגו כקצה.

Laplacian of Gaussian – אופרטור זה מעביר את התמונה דרך מסנן גאוס, ולאחר מכן מוצא את הנגזרת השנייה (לפלסיאן) לכל פיקסל. הנקודות שיווגו כקצה, הן הנקודות בהן הלפלסיאן מתאפס – כלומר נק' קיצון של הנגזרת הראשונה (שיפוע גדול בערכו המוחלט – כלומר שפה בין 2 תווים).

Canny – אופרטור זה מחשב גם הוא את הנגזרת הראשונה (גרדיאנט) עבור כל פיקסל, ומסווג נקודות מעל סף מסויים כנקודות שפה. כמו כן, עבור פיקסלים שנגזרתם הינה בערך ביניים כיוון הגרדיאנט נבדק ואזורים בהם יש רצף שמצביע לכיוון מסויים יסווגו גם הם כשפה.

3. רשמו לפחות שני גורמים טבעיים או סיבות טבעיות שונות שבגללן קיימות שפות בתמונות.

קיימות שפות בתמונות עקב הימצאות חפצים שונים/תווים שונים בתמונה. סיבה נוספת יכולה להיות תאורה שונה של אזורים שונים בתמונה.

4. ציינו שתי בעיות שונות היכולות להפריע לתהליך גילוי השפות.

בעיה שיכולה להפריע היא תאורה שונה לאורך אותו עצם. כמובן שנרצה לסווג את קצה העצם כשפה, אך תאורה שונה על פניו יכולה לשבש תהליך זה. בעיה אפשרית נוספת היא תמונה ברזולוציה נמוכה, המקשה על זיהוי מיקום מדויק של השפה או זיהויה בכלל.