

תרגיל מחשב מספר 2

הוראות כלליות

1. יש לשמור ולטעון את התמונות שקשורות לתרגיל **בספרייה אחת מעל** הספרייה בה נמצאים קבצי הקוד.
2. יש להימנע ככל האפשר משימוש בלולאות בקוד, ולעשות שימוש בכתוביב מטריצי.
3. יש להגיש את כל התרגיל כסקריפט יחיד. ניתן ורצוי לממש שאלות או חלקים מהשאלות כפונקציות (לא לשכוח לצרף אותם!!).
4. יש להקפיד לתת כותרת מתאימה לכל תמונה או גרף, וגם לצירים כאשר זה רלוונטי. יש לדאוג שגרפים ותמונות יוצגו בגודל מספק להבנת תוכנם.
5. יש להפריד ויזואלית בין סעיפים שונים בקוד בעזרת הערות.
6. יש לתעד את הקוד באופן סביר ובמיוחד במקומות בהן מבוצעות פעולות לא טריוויאליות.
7. את התרגיל יש להגיש אלקטרונית **דרך Moodle**, בחלק של "תרגילי מחשב" בצורה הבאה:
יש לשמור את קבצי התרגיל שלכם ללא תמונות קלט/פלט כלשהן, בקובץ ZIP כך:
code2_<ID>.zip
כאשר <ID> הוא מספר ת.ז. של המגיש. לדוגמה: **code2_012345678.ZIP** עבור תרגיל של מגיש בעלי מספר ת.ז. 012345678.
קובץ ה-ZIP הנ"ל יכול את הקבצים הבאים:
א. מחברת הרצת התרגיל וכל יתר הקוד הרלוונטי הנדרש להשגת כל התוצאות.
ב. קובץ PDF יחיד עם תשובות לשאלות התיאורטיות. יש לרשום את פרטי המגיש במסמך. לחילופין ניתן לרשום את התשובות המילוליות במחברת בצורה מודגשת וברורה, ולהימנע ממסמך PDF.
ג. אין צורך להגיש את התמונות או את קבצי העזר שניתנו לכם.
8. **את התרגיל יש להגיש עד ליום שישי בתאריך 30.5.2023 בשעה 23:55.**
על כל יום איחור בהגשה (שלא אושר מראש ע"י סגל הקורס) יורדו 4 נקודות אוטומטית. הגשה מאוחרת תעשה ישירות למייל ולא דרך Moodle.
9. התייעצות עם חברים מותרת ואף מומלצת, אולם את הקוד עליכם לכתוב בצורה עצמאית. הסגל יתייחס בחומרה המקסימלית להעתקות.
10. שאלות לגבי התרגיל ניתן לרשום בפורום המתאים ב-Piazza או במייל.

בהצלחה!

1. סינון מרחבי ופעולות מורפולוגיות

חלק א' – פעולות מורפולוגיות

בשאלה זו נכיר מספר פעולות מורפולוגיות בסיסיות ונעזר בהן כדי לחלץ אותיות מתמונה של מקלדת. בנוסף לארבעת הפעולות הבסיסיות שראינו בכיתה, נגדיר עוד שתי פעולות:

Top Hat – מחסירה מהתמונה המקורית את תוצאת הפעולה Open על התמונה: $A - A \circ B$. הפעולה משמרת שיאים בתמונה ומשפרת ניגודיות.

Bottom Hat – מחסירה את התמונה המקורית מתוצאת הפעולה Close על התמונה: $A \bullet B - A$. הפעולה משמרת שפלים ומשפרת ניגודיות.

לקריאה נוספת: http://utam.gg.utah.edu/tomo03/03_mid/HTML/node120.html

א. כדי להתרשם מששת הפעולות הנ"ל טענו את התמונה crazyBioComp.jpg והפעילו עליה את הפעולות בעזרת מספר אלמנטי בניה (Structure element) לבחירתכם.

ישנם מספר אלמנטים מובנים וכן גם ניתן להגדיר אלמנטים נוספים. הציגו את התוצאות.

ב. טענו את התמונה keyboard.jpg והציגו אותה.

ג. יצרו כעת שלושה אלמנטי בניה חדשים. האלמנט הראשון יהיה קו אנכי, השני יהיה קו אופקי והשלישי ריבוע, כולם עם צלע בגודל 8.

הפעילו שחיקה (erode) עם שני האלמנטים הראשונים (הקווים) הנ"ל על התמונה והציגו אותן. מהם המבנים הגיאומטריים הנשמרים בתמונה?

ד. כעת סכמו את שתי התמונות מהסעיף הקודם והציגו את תמונת הסכום. בחרו בערך הסף 0.2 והפכו את התמונה לבינארית. הציגו את התוצאה המתקבלת והסבירו.

חלק ב' – סינון

ה. בשלב זה נרצה להיעזר במסנן חציון כדי לבדוד את הקלידים. הקלידים הינם רכיבי הקשירות המכילים את הטקסט. תחילה בצעו היפוך לוגי (NOT) לתמונה מסעיף ד' ובצעו סינון חציון באמצעות סביבה בגודל 8x8. הסבירו מדוע דווקא סינון חציון מתאים (מדוע למשל לא בחרנו במסנן מיצוע?).

חלק ג' – בחזרה לפעולות מורפולוגיות

ו. כעת נרצה למקם טוב יותר את הקלידים שכן יתכן שתמונת הקשירות מסעיף ה' מכילה גם את שפות הקלידים. הפעילו שחיקה על התמונה מסעיף ה' באמצעות האלמנט השלישי (הריבוע). הציגו את התוצאה המתקבלת.

ז. בצעו פעולת חיתוך בין התמונה מסעיף ו' לבין התמונה המקורית (למשל ע"י כפל התמונות). שימו לב שכפל מוגדר מעל טיפוסים זהים לכן יש צורך להמיר את התמונה מסעיף ו' לטיפוס uint8. חדדו את התמונה באמצעות imsharpen והציגו את התמונה המתקבלת.

ח. בכדי להיפטר מהרקע המיותר בכל קליד הפעילו שוב פעולת סף עם ערך סף מתאים לדעתכם. הציגו את התוצאה הסופית והסבירו את התהליך הכולל שבצעתם.

2. התאמת תבניות

לאורך תרגיל זה נשתמש במדד ה-SSD (Sum of Squared Differences) למדידת הדמיון בין תמונות ה-Template (התבנית שנרצה למצוא) ובין התמונה המרכזית. מדד זה נתון על ידי:

$$S[x, y] = \sum_{i=1}^M \sum_{j=1}^N \left(T[i, j] - I \left[x + i - \frac{M}{2} - 1, y + j - \frac{N}{2} - 1 \right] \right)^2$$

כאשר T היא התבנית, I היא התמונה ו- (M, N) הם ממדי התבנית. ההתאמה המיטבית מתקבלת עבור ערך SSD מינימלי.

הערה: על מנת לבצע שימוש נכון במדד ה-SSD הפכו את כל התמונות בתרגיל זה ל-double.

א. כתבו פונקציה המקבלת תמונה ותבנית ומחשבת את מדד ה-SSD בכל הפיקסלים בתמונה. בקצוות התמונה, היכן שהתבנית ממילא אינה נכנסת, אפשר להחזיר ערך גדול מאוד.

ב. טענו את התמונה Text.jpg. תמונה זו מכילה טקסט קצר של Edgar Allan Poe.

בסעיף זה עליכם לגלות מהו גודל הגופן בו כתוב הטקסט. לשם כך השתמשו בתמונות

E10.jpg, E11.jpg, E12.jpg, E14.jpg, E16.jpg התואמות לגדלי גופן 10, 11, 12, 14, 16.

- הציגו את חמש התמונות יחדיו.

- ממשו אלגוריתם המשתמש במדד ה-SSD ומוצא לאיזו תמונת Template התאמה מיטבית לטקסט. מהי מסקנתכם לגבי גודל הגופן בטקסט?

ג. בסעיף זה עליכם לגלות כמה פעמים מופיעה האות a בטקסט וכמה פעמים מופיעה האות t בטקסט (כולל capitals). חתכו מתוך הטקסט את התבניות הדרושות לחישוב וקחו מסגרת הצמודה לאות ככל הניתן (הסתכלו על E10,...,E16 כדוגמה). השתמשו במדד ה-SSD להתאמה, קבעו סף לערכי ה-SSD על מנת להפריד בין התאמה וחוסר התאמה וספרו כמה פעמים מופיעה כל אות. יש לקבוע סף אחיד לשתי האותיות וה- $capitals$ שלהן (תנאי או פונקציה כלשהי הפועלים על ערכי ה-SSD ולא מספר קבוע); בחנו את וקטורי ה-SSD על מנת למצוא סף בצורה מתאימה.

- הציגו את התבניות שבהן השתמשתם.

- הסבירו את הסף הנבחר.

- רשמו את מספר המופעים של כל אות בשני הגדלים. האם הצלחתם לזהות את כל מופעי האותיות?

ד. עזרו ל-Mr. Smoke-too-much שאינו מסוגל להגיד את האות c , להקריא את הטקסט. עשו זאת ע"י זיהוי כל מופעי האות c בטקסט והחלפתם באות k . לשם כך השתמשו בתבניות c.jpg, k.jpg. הציגו את הטקסט החדש. האם הוא נראה טבעי?

3. חידוד שפות בתמונה רועשת

נטען כי המסנן $\delta - a\nabla^2 = \begin{pmatrix} 0 & -a & 0 \\ -a & 1+4a & -a \\ 0 & -a & 0 \end{pmatrix}$ מבצע חידוד שפות לתמונה. a הינו פרמטר הנע בתחום $[0,1]$. כאשר –

$$Laplacian = \nabla^2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \delta = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

- א. טענו את התמונה Bananas.jpg והפעילו עליה את המסנן עבור מספר ערכי a . במוצא המסנן יש לבצע Clipping לערכים החורגים מ- $[0,255]$.
- מצאו את ערך a הנותן את התוצאה האופטימלית עבור התמונה לדעתכם. הציגו את התוצאה.
- הסבירו מדוע מסנן זה מבצע פעולה של חידוד שפות, ומה תפקיד הפרמטר a .
- ב. כעת הוסיפו לתמונה רעש Salt and Pepper בצפיפות 0.03 (העזרו ב-imnoise), וחזרו על פעולת חידוד השפות. הציגו את התוצאה עבור $a = 0.2, 0.7$ והסבירו מדוע התקבל הגבר של הרעש.
- ג. הציגו דרך לשיפור התהליך של חידוד שפות בתמונה מורעשת, כאשר ידוע כי הרעש הינו מסוג Salt and Pepper. ממשו את הצעתכם והציגו את התוצאות עבור $a = 0.2, 0.7$.
- ד. חזרו על סעיפים ב'-ג' עבור רעש גאوسی עם תוחלת 0 ושונות 0.003.
- ה. חזרו על סעיפים א' ו-ד' (אין צורך לחזור על ב' + ג') עבור הגרעין הבא (לפליסיאן בסביבה 8 מלאה):

$$\delta - a\nabla^2 = \begin{pmatrix} -a & -a & -a \\ -a & 1+8a & -a \\ -a & -a & -a \end{pmatrix}$$

- מהם ההבדלים הצפויים להתקבל מההגדרה השונה של הגרעין? האם ניתן לראות אותן בתמונה המחודדת? הסבירו.
- הערה: כדי להבחין בהבדלים מומלץ להתמקד (Zoom) על שפה של אובייקט ולהתבונן בה בכיוונים שונים ולבצע השוואה עם סעיף ד'.

4. פירמידות

- א. כתבו פונקציות ליצירה של פירמידה גאוסית ופירמידה לפלסיאנית בעלות 5 שכבות כל אחת מתוך תמונת מקור שניתנת כפרמטר. ניתן להשתמש בפונקציות ספריה.
- ב. כתבו פונקציה המשחזרת תמונה מתוך פירמידה לפלסיאנית בת 5 שכבות. ניתן להשתמש בפונקציות ספריה.
- ג. הציגו את שכבות הפירמידות הנ"ל עבור התמונה Mandril והתמונה Toucan. השווה את שכבות הפירמידה הלפלסיאנית של שתי התמונות. מהם ההבדלים בניהן?
- ד. הראו כי ביצוע פירמידה לפלסיאנית לתמונה ואז שיחזורה מהפירמידה נותן את התמונה המקורית.
- ה. כעת נממש אלגוריתם פשוט לשיפור איכות של תמונה. התמונות focus1 ו-focus2 הן של אותה הסצנה רק עם מיקוד שונה. בעזרת פירמידה לפלסיאנית בת 4 שכבות נרצה לקבל תמונה אחת שכולה בפוקוס מתוך שתי התמונות בהן הפוקוס הוא חלקי. ממשו את האלגוריתם, הראו את התוצאות והסבירו מדוע זה מצליח.

5. התמרות

א. כתבו פונקציה המבצעת התמרת DCT דו-ממדית לבלוקים בגודל $M \times M$ (התמונה מחולקת לבלוקים שאינם חופפים, וגודלם הוא $M \times M$. לכל אחד מהבלוקים, בנפרד, מופעלת התמרת DCT דו-ממדית). הפונקציה תיראה כדלקמן:

$$[C] = \text{DCTcoeff}(\text{Pic}, K)$$

Pic – תמונה בגודל $N \times N$.

K – מספר בלוקים בשורה או עמודה ($K = N/M$).

C – מטריצת מקדמי ה-DCT. גודלה כגודל Pic.

• ניתן להניח כי N מתחלק ב-K.

ב. הפעילו את הפונקציה על התמונה Tiffany.jpg, עבור $K=8,64,512$. הציגו את תמונות מטריצות המקדמים המתקבלות. מה הסיבה להבדל בין המטריצות? מה הסיבה להבדלים בין הבלוקים השונים במטריצה? (מומלץ להסתכל על התמונות המתקבלות, פעם תוך שימוש ב-scaling לוגריתמי ופעם ללא שימוש ב-scaling)

נרצה כעת לייצג את התמונה עם $N^2/4$ איברים בלבד.

ג. כתבו פונקציה המבצעת התמרת DCT דו-ממדית לבלוקים בגודל $M \times M$, ומחזירה בכל בלוק רק את L המקדמים בעלי הערך המוחלט הגבוה ביותר. שאר המקדמים מאופסים. הפונקציה תיראה כדלקמן:

$$[C] = \text{DCT_L_coeff}(\text{Pic}, K, L)$$

Pic – תמונה בגודל $N \times N$.

K – מספר בלוקים בשורה או עמודה ($K = N/M$).

L – שבר (fraction) המייצג את מספר המקדמים הנשמרים בכל בלוק.

C – מטריצת מקדמי ה-DCT. גודלה כגודל Pic.

ד. כתבו פונקציה המבצעת התמרת DCT הופכית דו-ממדית לבלוקים בגודל $M \times M$. הפונקציה תיראה כדלקמן:

$$[\text{Pic}] = \text{iDCTcoeff}(C, K)$$

Pic – תמונה בגודל $N \times N$.

K – מספר בלוקים בשורה או עמודה.

C – מטריצת מקדמי ה-DCT. גודלה כגודל Pic.

ה. הפעילו את הפונקציה מסעיף ג' על התמונה Tiffany.jpg, כך שיישמרו רק $N^2/4$ מקדמים (כלומר, בכל בלוק יישמרו רק $M^2/4$ איברים). לאחר מכן, שחזרו את התמונה באמצעות הפונקציה מסעיף ד'. בצעו זאת עבור $K = 1, 8, 64$. הציגו את התמונה המקורית ואת כל תמונות השחזור המתקבלות ב-figure אחד (בעזרת subplot). האם ניתן להבחין בהבדלים בין התמונות?

ו. ננסה לבחון את ההבדלים בצורות נוספות:

1. הציגו את תמונת ההפרש בין תמונת השחזור המתקבלת בסעיף ה' והתמונה המקורית (עבור $K=1, 8, 64$). הקפידו להציג colorbar בכל התמונות. ניתן לבצע scaling או מתיחה בשביל לראות

טוב יותר את ההפרשים.

2. חשבו את ה-MSE בין תוצאת השחזור והתמונה המקורית. בצעו זאת עבור $K=1, 2, 4, 8, 16, 64, 128, 256$, והציגו גרף המתאר את ה-MSE כפונקציה של K .

• תזכורת:
$$MSE = \frac{1}{NM} \sum_{n=1}^N \sum_{m=1}^M (NewPic[n, m] - Pic[n, m])^2$$

מהי המסקנה? איזה K נותן את התוצאה הטובה ביותר? הסבירו.