

תרגיל Matlab מספר 2

הוראות כלליות

- יש לשמור ולטעון תמונות איתם תעבדו ב**ספרייה אחת מעל** הספרייה בה נמצאים קבצי הקוד. לדוגמה:
`>> I = imread('..\Lena.jpg');`
- יש להימנע ככל האפשר משימוש בלולאות בקוד, ולעשות שימוש בכתיב מטריצי.
- יש להגיש את כל התרגיל כסקריפט יחיד. ניתן ורצוי לממש שאלות או חלקים מהשאלות כפונקציות (לא לשכוח לצרף אותם כקבצים נפרדים!!).
- יש להקפיד לתת כותרת מתאימה לכל תמונה או גרף, וגם לצירים כאשר זה רלוונטי. יש לדאוג שגרפים ותמונות יוצגו בגודל מספק להבנת תוכנם.
- יש לכתוב כל סעיף ב-cell שונה. יוצרים cell חדש ע"י כתיבת התווים '%%' בתחילת שורה ריקה. (תווים אלו מבדילים בין cells). התווים הופכים להיות מודגשים ואז אפשר להפעיל בנפרד את ה-cell הנוכחי, הצבוע בצהוב, בעזרת Ctrl+Enter. לכל cell יש לרשום את מספר הסעיף ככותרת.
- יש לתעד את הקוד באופן סביר ובמיוחד במקומות בהן מבוצעות פעולות לא טריוויאליות.
- את התרגיל יש להגיש אלקטרונית דרך **Moodle**, בחלק של Matlab Assignments בצורה הבאה:
יש לשמור את קבצי התרגיל שלכם ללא תמונות קלט/פלט כלשהן, בקובץ ZIP כך:
ex2_<ID1>_<ID2>.zip, כאשר <ID1> ו-<ID2> הינם מספרי הסטודנט של שני המגישים.
לדוגמא: **ex2_012345678_987654321.ZIP** עבור תרגיל של מגישים בעלי מספרי הסטודנט 012345678, 987654321.
- ההגשה ב-moodle תעשה רק ע"י אחד מבני הזוג. יש לוודא כי גם שם קובץ ה-zip וכן גם העמוד הראשון של קובץ ה-pdf יכלול את שני מספרי ת.ז. של שני המגישים.
קובץ ה-ZIP הנ"ל יכלול את הקבצים הבאים:
א. סקריפט הרצת התרגיל וכל יתר הקוד הרלוונטי הנדרש להשגת כל התוצאות.
ב. קובץ PDF יחיד עם תשובות לשאלות התיאורטיות יחד עם כל תמונה וגרף שהתקבלו כפלט מהקוד. יש לרשום את פרטי המגישים במסמך.
ג. אין צורך להגיש את התמונות או את קבצי העזר שניתנו לכם.
- את התרגיל יש להגיש (מומלץ בזוגות) עד ליום א' בתאריך **26.5.2019** בשעה **23:55**.
על כל יום איחור בהגשה (שלא אושר מראש ע"י סגל הקורס) יורדו 4 נקודות אוטומטית. הגשה מאוחרת תעשה ישירות לאחראי התרגילים ולא דרך Moodle.
- התייעצות עם חברים מותרת ואף מומלצת, אולם את הקוד עליכם לכתוב בצורה עצמאית. הסגל יתייחס בחומרה המקסימלית להעתקות.
- מספר קבצי עזר ב-Matlab מצורפים באתר. כמו כן ניתן להיעזר כמובן ב-help של Matlab.
- שאלות לגבי התרגיל ניתן להפנות לאחראי תרגילי ה-Matlab דרך הפורום המתאים ב-Moodle או במייל.

בהצלחה!

1. סינון מרחבי ופעולות מורפולוגיות

הצעות לפונקציות שימושיות בשאלה זו:

strel, imerode, imdilate, imopen, imclose, imtophat, imbothat, imsharpen, medfilt2, im2bw

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, המתקבלים ע"י:

$$SE = strel(shape, parameters)$$

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 לבודד 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. גילוי וחידוד שפות בתמונה רועשת

הצעות לפונקציות שימושיות בתרגיל זה:

edge, imnoise, fspecial, imfilter, medfilt2, subplot, suptitle

- א. צלמו תמונת "סלפי" באמצעות הסמארטפון וטענו אותה למשתנה בשם selfie.
- שנו את גודלה ל-400X300 בערך באמצעות imresize מבלי לשנות את היחס בין אורך לרוחבה.
 - המירו אותה לתמונת רמות אפור באמצעות rgb2gray.
 - לבסוף, הפעילו על התמונה מסנן גאوسی בעל סטיית תקן 0.5 באמצעות imgaussfilt.
- ב. ישנן מספר שיטות לגילוי שפות בתמונה. בסעיף זה נבחן שלוש מהן על התמונה selfie. לשם כך נשתמש בפונקציה edge עם פרמטר המציין את שם האופרטור ועוד פרמטרים רלוונטיים. הציגו את מפת השפות של שלושת השיטות הבאות באותו figure באמצעות subplot:
1. אופרטור Sobel
 2. אופרטור Roberts
 3. אופרטור Canny
- מצאו ערכי סף וסיגמא טובים לטעמכם לכל אחד מהאופרטורים בנפרד. הסבירו באופן איכותי איזה אופרטור עדיף לדעתכם, ולמה?
- הערה: לשם השוואה, ניתן לבצע zoom על שפה מובהקת.

המסנן הבא מבצע חידוד שפות לתמונה, כאשר a הינו פרמטר בתחום $[1,0]$.

$$\delta - a\nabla^2 = \begin{pmatrix} 0 & -a & 0 \\ -a & 1+4a & -a \\ 0 & -a & 0 \end{pmatrix}$$

- ג. הפעילו על התמונה selfie את המסנן עבור מספר ערכי a . במוצא המסנן יש לבצע Clipping לערכים החורגים מ- $[255,0]$.
- מצאו את ערך a הנותן את התוצאה האופטימלית עבור התמונה לדעתכם. הציגו את התוצאה. זכרו לבצע מתיחה לטווח הדינמי של התמונה.
 - הסבירו מדוע מסנן זה מבצע פעולה של חידוד שפות, ומה תפקיד הפרמטר a .
- ד. כעת הוסיפו לתמונה רעש Salt and Pepper בצפיפות 0.06 (היעזרו ב-imnoise), וחזרו על פעולת חידוד השפות. הציגו את התוצאה עבור $a=[0.3,0.5,0.7]$ והסבירו מדוע התקבל הגבר של הרעש.
- ה. 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 $a=[0.3,0.5,0.7]$.

3. לכידת תמונה

הצעות לפונקציות שימושיות בחלק זה: VideoReader, readFrame, hasFrame.

בשאלה זו נתנסה בשיטת **Rolling Shutter**, אשר משמשת ללכידת תמונות ומיושמת ברוב מכשירי הצילום כיום (למעשה, בכל המכשירים בהם מותקן חיישן CMOS). בשיטה זו, התמונה נרכשת ע"י סריקה של הפריים בצורה אופקית ואנכית. כתוצאה מכך, לא כל החלקים של סצנת התמונה מוקלטים באותו זמן, אך הם מוצגים יחד באותו זמן. בעקבות שיטת הדגימה המיושמת ב-Rolling Shutter, עצמים אשר נעים במהירות גבוהה עלולים להופיע עם עיוותים צפויים בתמונה. בשאלה זו נתוודע לחלק מעיוותים אלו. מומלץ לצפות בסרטון המעניין הבא המדגים את אופן פעולת השיטה:

<https://www.youtube.com/watch?v=dNVtMmLlnoE>

- א. צפו בסרטון הצבע GallopingHorse.mp4 באמצעות תוכנה ייעודית כלשהי, ולאחר מכן טענו אותו ל-Matlab בעזרת הפקודה VideoReader. הסרטון ישמר במשתנה מסוג VideoReader, כאשר קריאה של מסגרת מהסרטון תעשה באמצעות הפקודה readFrame. על-מנת ללמוד כיצד לקרוא ואף לשנות מאפיינים שונים של הסרטון (כגון גודל כל מסגרת), קראו על הפקודה VideoReader ב-Help של Matlab.
- ב. כעת נבצע סימולציה של לכידת תמונת צבע בשיטת Rolling Shutter. בצעו סריקה אנכית של הסצנה, כך שמכל מסגרת בסרטון תילקח שורה אחת לתמונה הסופית. צפו בתמונה שהתקבלה ורשמו אילו עיוותים ניתן לצפות בה.
- ג. חזרו על סעיף ב', אך הפעם בצעו סריקה אופקית משמאל לימין.
- ד. חזרו על סעיף ג', אך הפעם בצעו סריקה אופקית מימין לשמאל. הסבירו את ההבדל בין תוצאות סעיפים ג' ו-ד'.

4. עקיבה בוידאו

הצעות לפונקציות שימושיות בחלק זה: `imshow`, `normxcorr2`, `ind2sub`.

א. פתחו את הקובץ `VideoTracking.mat`. הקובץ מכיל את המטריצות `Template` ו-`Frames`, כאשר `Template` הינו האובייקט שאחריו נרצה לעקוב בסרטון השמור ב-`Frames`. הציגו את `Template` וצפו בסרטון (`Frames`) באמצעות `imshow`.

ב. ממשו אלגוריתם לזיהוי ה-`Template` בכל אחת מהתמונות ב-`Frames`, בכל תמונה סמנו את המיקום המזוהה בעזרת ריבוע שחור בגודל של `Template` המקיף את האובייקט המזוהה. לשם זיהוי ה-`Template` נכיר מדד דמיון נוסף, מדד הקרוס-קורלציה המנורמלת:

$$\rho_{TI}[x, y] = \frac{R_{TI}[x, y]}{\sqrt{R_{TT}\left[\frac{M}{2}, \frac{N}{2}\right] R_{II}[x, y]}}$$

$$R_{TI}[x, y] = \sum_{i=1}^M \sum_{j=1}^N T[i, j] \cdot I\left[x + i - \frac{M}{2} - 1, y + j - \frac{N}{2} - 1\right] \quad \text{כאשר}$$

היעזרו בפונקציית `Matlab` בשם `normxcorr2`.

כל אחת מהתמונות ב-`Frames` חפשו את ההתאמה המיטבית ל-`Template`. ההתאמה המיטבית מתוארת על ידי ערך מקסימלי של הקורלציה המנורמלת.

ג. צפו בסרטון הכולל את סימון ה-`Template`, האם האלגוריתם הצליח לזהות את ה-`Template` בכל אחת מהתמונות ב-`Frames`? הציגו מספר הצלחות וכישלונות (במידה וקיימים) יחד עם סימון ה-`Template`. הסבירו את הצלחות/כישלונות הזיהוי.

ד. תקנו את האלגוריתם כך שיבצע זיהוי מוצלח במסגרות בהם נכשל קודם לכן. הציגו מסגרות אלו לאחר התיקון.

5. שחזור תמונות

הצעות לפונקציות שימושיות בחלק זה:

conv2, reshape, imnoise, meshgrid, semilogy, sparse, imgradientxy, divergence.

נתונה התמונה Y אשר מהווה גרסה רועשת של התמונה X . לצורך שחזור X מתוך Y נרצה למזער את הביטוי הבא:

$$\varepsilon^2\{\underline{X}\} = (\underline{X} - \underline{Y})^T (\underline{X} - \underline{Y}) + \lambda (D\underline{X})^T (D\underline{X})$$

כאשר \underline{X} הנו וקטור המורכב מהעמודות של התמונה X , \underline{Y} וקטור המורכב מהעמודות של התמונה הרועשת Y ו- D הוא אופרטור הלפליסיאן אשר נתון ע"י הגרעין:

$$D_{\text{kernel}} = \begin{bmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

λ הינו פרמטר הרגולריזציה.

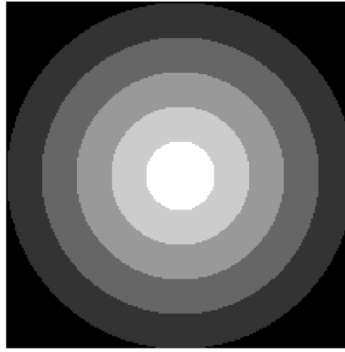
על מנת לשחזר את תמונת המקור נפעיל תהליך איטרטיבי המבוסס על אלגוריתם Steepest Descent:

$$\hat{\underline{X}}_{k+1} = \hat{\underline{X}}_k - \mu_k \underline{G}_k = \hat{\underline{X}}_k - \mu_k \left((I + \lambda D^T D) \hat{\underline{X}}_k - \underline{Y} \right)$$

נאתחל תהליך זה ב- $\underline{X}_0 = \underline{Y}$, ונקבע את גודל הצעד μ_k באופן הבא:

$$\mu_k = \frac{\underline{G}_k^T \underline{G}_k}{\underline{G}_k^T (I + \lambda D^T D) \underline{G}_k}$$

א. בנו תמונה שתראה באופן הבא:



לתמונה שש רמות אפור מ-0 עד 250 בקפיצות של 50. העיגול הפנימי ברדיוס 20 פיקסלים, וכל מסגרת עגולה היא ברוחב של 20 פיקסלים. גודל התמונה הוא 200X200.

נגדיר ב- a את מספר האלקטרוני הנקלטים במצלמה אשר מתורגמים לרמת אפור אחת. נניח כי $a = 5$. כעת הכניסו לתמונה רעש פואסוני באופן המדמה רעש אמיתי המתקבל במערכות אופטיות:

1. הכפילו את ערכי רמות האפור בתמונה ב- a כדי להמיר את התמונה ליחידות של אלקטרוני.
2. המירו את התמונה המתקבלת ל-uint16. המרה זו הכרחית עבור שימוש נכון ב-imnoise בשלב הבא.
3. הפעילו על התמונה את הפונקציה imnoise עם ארגומנט 'poisson'. בפעולה זו למעשה ערך כל פיקסל (כמות האלקטרוני) שחושב לעיל מומר למשתנה אקראי פואסוני. משתנה אקראי פואסוני הוא בעל שונות השווה לתוחלת. התוחלת במקרה שלנו הוא מספר האלקטרוני שהתקבל בשלב 2 לעיל.
4. חלקו את התמונה ב- a כדי להחזיר את התמונה לרמות אפור רגילות.
5. הפעילו round על התמונה ובצעו clipping לתחום [0 255] על-מנת לממש קוונטיזציה.
6. המירו את התמונה ל-double לטובת המשך השאלה.

התמונה הרועשת המתקבלת תהיה התמונה Y .

ב. ממשו את האלגוריתם אשר תואר לעיל כפונקציה הבאה:

`[Xout, Err1, Err2] = DenoiseByL2(Y, X, numIter, lambda)`

Y – מטריצת $M \times N$ המתארת את התמונה הרועשת.

X – התמונה המקורית.

numIter – מספר האיטרציות שעל אלגוריתם לבצע.

lambda – פרמטר הרגולריזציה.

Xout – התמונה המשוחזרת.

Err1, Err2 – וקטורים באורך של מספר האיטרציות אשר מכילים את השגיאות הבאות:

$$Err1\{\hat{\underline{X}}_k\} \triangleq \left(\hat{\underline{X}}_k - \underline{Y}\right)^T \left(\hat{\underline{X}}_k - \underline{Y}\right) + \lambda (D\hat{\underline{X}}_k)^T (D\hat{\underline{X}}_k)$$

$$Err2\{\hat{\underline{X}}_k\} \triangleq \left(\hat{\underline{X}}_k - \underline{X}\right)^T \left(\hat{\underline{X}}_k - \underline{X}\right)$$

הפונקציה תציג את התמונה המתקבלת לאחר כל 10 איטרציות.

הערה: ניתן לממש את הפונקציות ע"י כפל מטריצות דלילות או ע"י קונבולוציות עם גרעינים.

ג. הפעילו את הפונקציה שכתבתם בסעיף ב' על התמונה שיצרתם. השתמשו ב- $\lambda = 0.5$ וב- $\text{numIter}=50$. התבוננו בתמונה לאחר כל 10 איטרציות והציגו את התמונה הטובה ביותר בעיניכם. הציגו גרף לוגריתמי של השגיאות ($\text{Err1}, \text{Err2}$) כתלות במספר האיטרציות (במערכת צירים אחת), והסבירו.

ד. כעת ננסה שחזור עם prior של Total Variation, כלומר נעבוד עם פונקציית העלות:

$$\mathcal{E}^2\{\underline{X}\} = (\underline{X} - \underline{Y})^T (\underline{X} - \underline{Y}) + \lambda \cdot TV\{\underline{X}\}$$

כאשר TV הוא פונקציית ה-Total Variation.

על מנת לשחזר נפעיל תהליך איטרטיבי כמו בסעיף ב', רק עבור צעד גרדיאנט המתאים ל-TV:

$$\hat{\underline{X}}_{k+1} = \hat{\underline{X}}_k + \frac{\mu_k}{2} \underline{U}_k = \hat{\underline{X}}_k + \frac{\mu_k}{2} \left(2(\underline{Y} - \hat{\underline{X}}_k) + \lambda \nabla \cdot \left(\frac{\nabla \hat{\underline{X}}_k}{\sqrt{|\nabla \hat{\underline{X}}_k|^2 + \varepsilon_0^2}} \right) \right)$$

נאתחל תהליך זה ב- $\underline{X}_0 = \underline{Y}$, ונשתמש בגודל צעד קבוע $\mu_k = 100\varepsilon_0$.

ממשו את האלגוריתם המתאים כפונקציה הבאה:

`[Xout, Err1, Err2]=DenoiseByTV(Y, X, numIter, lambda)`

כאשר הפרמטרים הם כמו בסעיף ב', למעט Err1 , שזאת יוגדר כך:

$$\text{Err1}\{\hat{\underline{X}}_k\} \triangleq (\hat{\underline{X}}_k - \underline{Y})^T (\hat{\underline{X}}_k - \underline{Y}) + \lambda \cdot TV\{\hat{\underline{X}}_k\}$$

הפונקציה תציג את התמונה המתקבלת לאחר כל 50 איטרציות.

ה. הפעילו את הפונקציה שכתבתם בסעיף ד' על התמונה שיצרתם. השתמשו ב- $\lambda = 20$ וב- $\text{numIter}=200$. התבוננו בתמונה לאחר כל 10 איטרציות והציגו את התמונה הטובה ביותר בעיניכם. הציגו גרף לוגריתמי של השגיאות ($\text{Err1}, \text{Err2}$) כתלות במספר האיטרציות (במערכת צירים אחת), והסבירו.

ו. חוו את דעתכם על התוצאות. האם הן מתאימות לתאוריה?

ז. צלמו תמונה חיצונית מאזור מוכר ברחבי קמפוס הטכניון. המירו אותה לרמות אפור והקטינו אותה לגודל סביר. נסו את התהליך כולו עליה והציגו את התוצאות.