הטכניון – מכון טכנולוגי לישראל
הפקולטה למדעי המחשב

עיבוד תמונות דיגיטלי 2360860
חורף תשפ"ו

<u>תרגיל בית מס׳ 3</u>

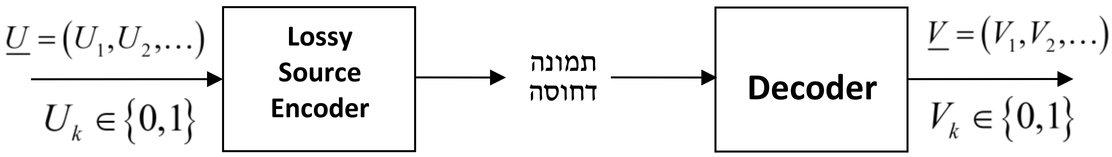
מועד הגשה: עד 1.1.25 בשעה 23:59. הגשה אלקטרונית דרך Moodle.

<u>שאלה מס׳ 1</u>

נתון מקור בינארי סימטרי חסר-זיכרון U (כלומר, $p(U=0)=p(U=1)=\frac{1}{2}$). נרצה לדחוס מקור זה בקצב R עם עיוות ממוצע D, כאשר מדד העיוות הוא עיוות Hamming:

$$d_H(u,v)=\begin{cases}0, & u=v\\ 1, & u\neq v\end{cases}$$

הדחיסה מתוארת באיור הבא:



כאשר $U\in\{0,1\}$ הוא המקור, ו- $V\in\{0,1\}$ הוא מוצא המפענח והעיוות הממוצע נתון ע"י $D=E\{d_H(U,V)\}$.

א. רשמו את פונקציית הקצב-עיוות של המקודד האופטימלי של המקור כפונקציה של D.

לצורך קידוד המקור מוצעות שתי סכמות דחיסה:

I. כאשר המקור מפיק '0' הופכים אותו ל-'1' בהסתברות q שערכה נבחר כך שהעיוות הממוצע לא יחרוג מ- D. המוצא נדחס ללא עיוות נוסף בקצב R נמוך ככל שניתן.

II. בהינתן וקטור מקור באורך N, נשדר רק את NR הסיביות הראשונות, ואת השאר המקלט ישלים באפסים (עם עיוות D).

ב. מצאו ביטוי ל- q כפונקציה של D עבור סכמה I.

ג. חשבו את הפונקציה $R_I(D)$ של סכמה I, ושרטטו אותה.

ד. חשבו את הפונקציה $R_{II}(D)$ של סכמה II, ושרטטו אותה.

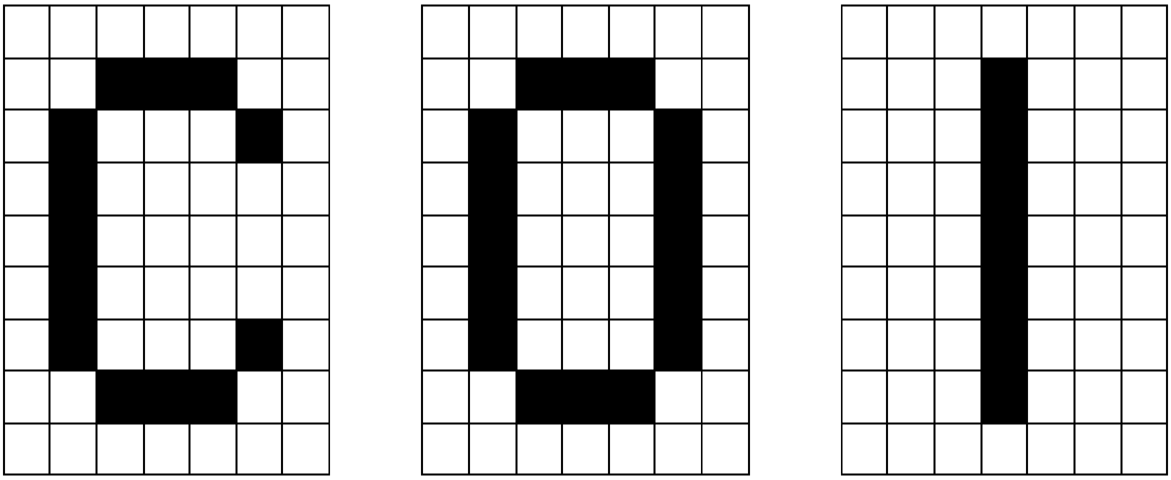
ה. איזו מבין שתי הסכמות טובה יותר? הסבירו.

<u>שאלה מס׳ 2</u>

חלקי שאלה זו אינם תלויים זה בזה.

<u>חלק א׳</u>

נתון מקור אותיות המייצר תמונות בינאריות של שלוש האותיות הבאות באנגלית 'l', 'o' ו-'c' בלבד, בגודל 9×7, המתוארות בתמונות הבאות **<u>(פיקסל לבן = 0)</u>**:



א. לתמונות מתווסף רעש אדיטיבי W, אשר מוסיף לכל פיקסל אחד משלושת הערכים [-a,0,a] באופן בלתי תלוי ביתר הפיקסלים ובהסתברות שווה. הציעו שיטה לשחזור התמונה כאשר a ידוע, ורשמו את כל ערכי a שעבורם לא ניתן לשחזר.

<u>חלק ב׳</u>

בחברת המשקאות הקלים "קוקו-לוקו" נפתח פס ייצור חדש עבור המשקה הפופולרי "פאמטה". עקב תקלות במילוי הבקבוקים הוחלט להציב מצלמה מול פס הייצור כדי לגלות בקבוקים שלא מולאו עד הסוף. פס הייצור עוצר כל מספר שניות והמצלמה מצלמת תמונת רמות אפור בתחום [0,255] של מספר בקבוקים במיקום אופקי לא ידוע על גבי רקע שחור **(רמת אפור 0)**. אפשר להניח שאין חצאי בקבוקים בתמונה. תמונה לדוגמה:

ב. הבקבוקים בלבד מקבלים רעש שיאים כך ש-40% מהפיקסלים שלהם נפגעים מהרעש. הציעו שיטה טובה ככל הניתן לתיקון הבקבוקים בתמונות תוך שמירה מרבית על איכותן.

<u>חלק ג'</u>

נתונה התמונה $U[m,n]$, המכילה אך ורק P העתקים של התמונה $\psi[m,n]$, כאשר קואורדינטות המרכזים של העתקי $\psi[m,n]$ בתמונה $U[m,n]$, הינם $\{(m_i,n_i)\}_{i=1}^{P}$. ניתן להניח כי ההעתקים של $\psi[m,n]$ לא נחתכים בקצוות התמונה $U[m,n]$. במקרה של חפיפות בין העתקים של $\psi[m,n]$, הערך המתאים ב$U[m,n]$ הוא סכום ההעתקים. ערכם של שאר הפיקסלים בתמונה $U[m,n]$ הוא 0. נסמן ב-H את הפעולה היוצרת את $U[m,n]$ מתוך $\psi[m,n]$, כלומר: $U[m,n]=H\{\psi[m,n]\}$.

התמונה $\psi[m,n]$, בגודל $M\times M$, מקיימת את המשוואה הבאה:

$$\psi[m,n]=\begin{cases}1, & m=\dfrac{M+1}{2}\\ 0, & m\neq\dfrac{M+1}{2}\end{cases}$$

התמונה $U[m,n]$ נחשפה לרעש שיאים (Salt & Pepper), כך שרק 3% מהפיקסלים מכילים רעש.

ג. המהנדס הנודע חואן טייזר הציע להשתמש במסנן חציון 3×3 (עם סביבה 8) לניקוי הרעש. תארו באופן <u>איכותי</u> ו<u>בקצרה</u> את התמונה במוצא המסנן. התייחסו בתשובתכם לתמונה המקורית ולרעש.

<u>שימו לב</u>: בסעיף הבא (סעיף ד'), הקואורדינטות $\{(m_i,n_i)\}_{i=1}^{P}$ <u>אינן</u> נתונות.

ד. בהתבסס על מסננים מוכרים, הציעו סכמה המבוססת על סינון לשחזור התמונה $U[m,n]$ באופן כזה שהשגיאה $\left\|U[m,n]-\hat{U}[m,n]\right\|$ תהיה מינימלית, כאשר $\hat{U}[m,n]$ הינה התמונה המשוחזרת. אין צורך לפתור אנליטית, אלא רק להציע ולהסביר את הסכמה.

<u>שאלה מס' 3</u>

במערה נטושה בבקעת רוהאן התגלו מגילות קלף הכתובות בכתב הוביט עתיק, המכיל אותיות מארבעת הסוגים הבאים: Γ, ⅂, L, □. על מנת לנתח באופן אוטומטי את המגילות בשפת ההוביט, הן נסרקות ומנותחות במחשב. לאחר הסריקה מתקבלות תמונות בגודל $N\times N$ (N גדול מאוד, כך שניתן להזניח אפקטי קצוות). תמונה כזו לדוגמה היא התמונה הבינארית I (פיקסל ריק = אפס):

		1	1						
		1					1		
							1	1	
	1	1				1	1	1	
		1				1		1	
						1	1	1	

א. בהינתן תבנית לאות ד, $T_{\text{ד}} =$

1	1
	1

מבצעים קורלציה מרחבית דו-ממדית (התאמת תבניות) בין I ל- $T_{\text{ד}}$. נסמן את תוצאת התאמת התבניות הזו ב- $I_{\text{ד}}$. איירו את התמונה $I_{\text{ד}}$.

ב. לאחר מכן מבצעים פעולת סף על $I_{\text{ד}}$ כדי לגלות את מיקום התבניות $T_{\text{ד}}$ בתמונה.
1. באיזה סף כדאי לבחור? ציינו במפורש איזה פיקסלים יעברו את הסף שבחרתם.
2. האם ישנן אותיות המזוהות באופן שגוי כאות ד? הסבירו.

ג. באופן דומה לסעיף א' מגדירים את $T_{\square} =$

1	1	1
1		1
1	1	1

ומבצעים קורלציה מרחבית דו-ממדית בין I ל- $T_{\square}$. נסמן את תוצאת התאמת התבניות הזו ב- $I_{\square}$. איירו את התמונה $I_{\square}$.

ד. באופן דומה לסעיף ב' מבצעים פעולת סף על $I_{\square}$ כדי לגלות את מיקום התבניות $T_{\square}$ בתמונה.
באיזה סף כדאי לבחור? ציינו במפורש איזה פיקסלים יעברו את הסף שבחרתם.

ה. מעוניינים בשיטה ליצירת תמונה בינארית Z שבה הפיקסלים היחידים בעלי ערך 1 הם אלו שבהם קיימת האות ד <u>בלבד ולא אף אות מסוג אחר</u>. הציעו שתי שיטות פשוטות שונות, וכתבו עבורן אלגוריתם בפסאודו-קוד, ליצירת Z. ניתן (אך לא חובה) להשתמש במשתנים הבאים: I, $T_{\text{ד}}$, $T_{\square}$, $I_{\text{ד}}$ ו- $I_{\square}$.

<u>שאלה מס' 4</u>

התמונה הבדידה X בגודל $K \times K$ עוברת עיוות והרעשה לפי המודל $\underline{Y} = f\left(\underline{X}\right) + \alpha \underline{W}$ ומתקבלות המדידות $\underline{Y}$. $\underline{W}$ הוא רעש אדיטיבי המפולג לפי $\mathrm{N}\left(\underline{\mu}, \sigma_W^2 I\right)$ בת"ס ב- $\underline{X}$, ו- $f\left(\underline{X}\right)$ היא פונקציה וקטורית. נניח בתחילה כי $f\left(\underline{X}\right) = H\left(\underline{X} + \underline{\mathbf{1}}\right)$ כאשר H היא מטריצת פעולה ידועה של גרעין טשטוש כלשהו ו- $\underline{\mathbf{1}}$ הוא וקטור של אחדות. <u>הערה:</u> כל הווקטורים בשאלה זו הם בייצוג עמודה (CS).

א. מצאו משערך של התמונה $\underline{X}$ מתוך תמונת המדידות $\underline{Y}$ כאשר $\alpha = 0$.

כעת נתון כי $\alpha = 2$ ועל התמונה X נתון ידע מוקדם בדמות הפילוג הבא:

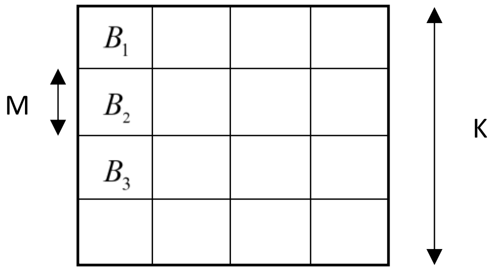
$$p_X\left(\underline{x}\right) \propto \exp\left\{-\frac{\left\|D_x\underline{x}\right\|_2^2 - \left\|D_y\underline{x}\right\|_2^2}{2\sigma_S^2}\right\}$$

כאשר D_x ו- D_y הם אופרטורים של נגזרות <u>אחוריות</u>.

ב. מצאו את משערך ה-MAP של התמונה $\underline{X}$ מתוך המדידות $\underline{Y}$.

ג. כיצד משפיע σ_W על אופי השערוך של $\underline{X}$? כיצד משפיע σ_S על אופי השערוך? נמקו.

נגדיר מטריצת פעולה G_k אשר מטרתה לחלץ את הבלוק ה- k, B_k, בגודל $M \times M$ מהתמונה לפי $\underline{B}_k = G_k\underline{X}$. ניתן להניח כי M מחלק את K וכי הבלוקים לא חופפים. תמונה להמחשה:



ד. האם G_2 פעולה ספרבילית? האם היא קבועה במקום? נמקו בקצרה (אין צורך להוכיח).

נתונים שני המשערכים הבאים עבור התמונה $\underline{X}$:

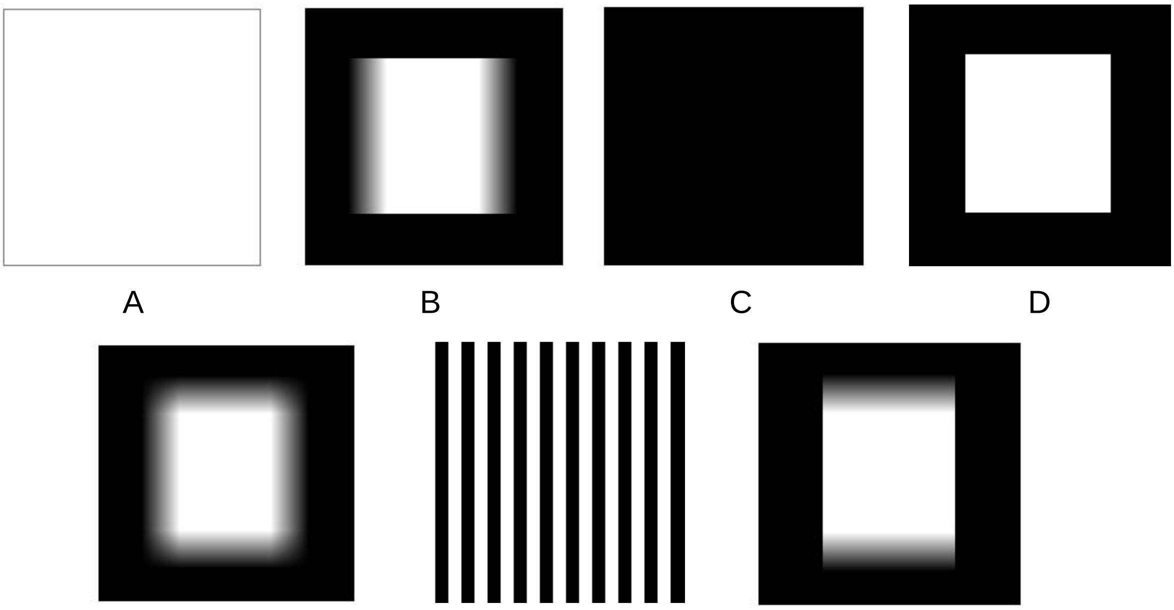
$$\hat{X} = \underset{X}{argmin}\frac{1}{2}\left\|\underline{Y} - f\left(\underline{X}\right)\right\|_2^2 + \lambda\left\|L_K\underline{X}\right\|_2^2$$

$$\hat{X} = \underset{X}{argmin}\frac{1}{2}\left\|\underline{Y} - f\left(\underline{X}\right)\right\|_2^2 + \lambda\sum_{k=1}^{n}\left\|L_M G_k\underline{X}\right\|_2^2$$

כאשר L_K ו- L_M הן פעולות לפלסיאן על וקטורים בגודל $K^2 \times 1$ ו-$M^2 \times 1$ בהתאמה.

ה. הסבירו באופן איכותי מה יהיו ההבדלים בשערוך התמונה $\underline{X}$ בין שני המשערכים הנ"ל.

נתונות תמונות רמות האפור A-G הבאות בתחום [0,1] **<u>(לבן = 1)</u>**. כל התמונות זהות בגודלן:



A B C D

E F G

נתונות בנוסף הפונקציות 1-4 המהוות הסתברויות א-פריוריות (priors):

$$\boxed{1}\ \exp\left\{-\left\|D_x\underline{X}\right\|_2^2\right\} \qquad \boxed{2}\ \exp\left\{-\left\|\underline{X}-D_x\underline{X}-D_y\underline{X}\right\|_2^2\right\}$$

$$\boxed{3}\ \exp\left\{-\left\|D_y\underline{X}\right\|_2^2\right\} \qquad \boxed{4}\ \exp\left\{-\frac{1}{\left\|\underline{X}+\varepsilon\mathbf{1}\right\|_2^2}\right\}, \quad (\varepsilon<<1)$$

ו. <u>לכל אחת</u> מהפונקציות 1-4, מצאו את התמונה (או התמונות) מתוך A-G שהיא בעלת הסתברות א-פריורית מירבית, ואת התמונה (או התמונות) שהיא בעלת הסתברות א-פריורית הכי נמוכה. נמקו.

כעת שוב $\alpha=0$, לא נתון ידע מוקדם על $\underline{X}$ והפונקציה $f\left(\underline{X}\right)$ מקיימת $f\left(\underline{X}\right)=N\underline{X}$, כאשר N היא מטריצה בת"ס ב- $\underline{X}$ ואיבריה מפולגים .i.i.d לפי $\mathrm{N}\left(0,\sigma^2\right)$.

ז. מצאו את התוחלת והשונות של $\underline{Y}$ בהינתן $\underline{X}$. הסבירו.

ח. רשמו ביטוי עבור משערך ה-ML של התמונה $\underline{X}$ מתוך המדידות $\underline{Y}$. פתחו את המשערך ומצאו תנאי על $\underline{X}$ לצורך השערוך. הסבירו את התוצאה שקיבלתם ומשמעותה.

<u>שאלה מס׳ 5</u>

נסמן ב- H מטריצה שאיננה ריבועית אך H^TH מטריצה הפיכה, ו- $\underline{N}$ הינו וקטור רעש המפולג $\underline{N}\sim\mathrm{N}\left(0,\sigma^2I\right)$, כאשר σ פרמטר ידוע. תמונה כלשהי $\underline{X}$, הנתונה בסידור עמודה (CS) ובלתי תלויה

ב- $\underline{N}$, עוברת עיוות הנתון ע"י המשוואה $\underline{Y} = H\underline{X} + \underline{N}$. הניחו כי המודל המקדים (prior) המתאים לתמונה הינו $P\left(\underline{X}\right) = C_1 \exp\left\{-\alpha \left\|L\underline{X}\right\|^2\right\}$.

א. <u>בסעיף זה בלבד</u> נתון כי L הינה מטריצת הפעולה של הגרעין הבא: $l = \frac{1}{9}\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$.

הסבירו איכותית אילו תמונות מועדפות על-ידי המודל עבור $\alpha \gg 1$.

מעתה הניחו כי L הינה מטריצה ידועה <u>כלשהי</u>.
<u>שימו לב</u>: בסעיפים ב', ג' ו-ו' יש להגיע למשערכים מהצורה $\hat{X} = A\underline{Y}$ כאשר A מטריצה מתאימה.

ב. חשבו ורשמו משערך ML ל- $\underline{X}$ על סמך המדידה $\underline{Y}$. נסמן משערך זה ב- $\hat{\underline{X}}_1$.

ג. חשבו ורשמו משערך MAP ל- $\underline{X}$ על סמך המדידה, $\underline{Y}$. נסמן משערך זה ב- $\hat{\underline{X}}_2$.

ד. נניח כי $\sigma^2 \cong 0$. מהו ההבדל בין ביצועי המשערכים $\hat{\underline{X}}_1$ ו- $\hat{\underline{X}}_2$ (האם משערך אחד טוב מהשני, האם ביצועיהם זהים, או שלא ניתן לדעת)? נמקו.

מעתה הניחו כי מודל מקדים (prior) לתמונה <u>אינו ידוע</u>, וכי H הינה מטריצה ריבועית שאיננה הפיכה $\left(\det\left(H\right) = 0\right)$.

מהנדס א' הציע, באופן שרירותי, את המודל הבא לתמונה: $P\left(\underline{X}\right) = C_1 \exp\left\{-\frac{1}{2s^2}\left\|\underline{X}\right\|^2\right\}$, כאשר הפרמטרים $C_1, s > 0$ ידועים. נסמן ב- $\hat{\underline{X}}_3$ את משערך ה- MAP המתקבל בשערוך $\underline{X}$ על-סמך $\underline{Y}$ לפי המודל של מהנדס א'.

ה. הסבירו אילו תמונות מועדפות על-ידי המודל עבור $s \cong 0$ ועבור $s \gg 1$.

מהנדס ב' טוען שמכיוון שאין לתמונה מודל מקדים ידוע, לא ניתן לבצע שערוך MAP, ובמקום זאת יש לשערך את $\underline{X}$ על-סמך $\underline{Y}$ ע"י משערך ML בלבד.

ו. חשבו ורשמו משערך ML יציב נומרית המתאים לדרישותיו של מהנדס ב'.

נסמן את המשערך שקיבלתם בסעיף ו' ב- $\hat{\underline{X}}_4$. מהנדס ג' טוען שעבור אוסף פרמטרים מסוים $\left(C_1, s\right)$ המשערך $\hat{\underline{X}}_3$, שהינו משערך MAP, זהה למשערך $\hat{\underline{X}}_4$, שהינו משערך ML.

ז. מצאו פרמטרים $\left(C_1, s\right)$ עבורם טענתו של מהנדס ג' נכונה, או הסבירו מדוע היא איננה נכונה.