

ראייה אנושית גישה חישובית

תרגיל 3

מגשים: רז דרזי ורן שחם

11 ביוני 2017

חלק I

תיאורטי

נגדיר את הקונבולוציה של שתי מטריצות X, Y כך:

$$Z[m, n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X[i, j] Y[m-i, n-j]$$

ונסמן:

$$Z = X * Y$$

שאלה 1

נראה שקונבולוציה היא אסוציאטיבית. יהיו X, Y, Z מטריצות. אזי:

$$\begin{aligned} ((X * Y) * Z)[l, k] &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} (X * Y)[m, n] Z[l-m, k-n] \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \left(\sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X[i, j] Y[m-i, n-j] \right) Z[l-m, k-n] \\ &= \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (X[i, j] Y[m-i, n-j] Z[l-m, k-n]) \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} X[i, j] Y[m-i, n-j] Z[l-m, k-n] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X[i, j] \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} Y[m-i, n-j] Z[l-m, k-n] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X[i, j] \sum_{m=-\infty}^{\infty} \sum_{n=-\infty}^{\infty} Y[m, n] Z[(l-i)-m, (k-j)-n] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X[i, j] (Y * Z)[l-i, k-j] = (X * (Y * Z))[l, k] \end{aligned}$$

לכל k, l , ולכן שוויון מתקיים:

$$(X * Y) * Z = X * (Y * Z)$$

שאלה 2

יהיו a, b סקלרים ו- X, Y, Z מטריצות. אזי:

$$\begin{aligned} (X * (aY + bZ)) [m, n] &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X [i, j] (aY + bZ) [m - i, n - j] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X [i, j] aY [m - i, n - j] + \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} X [i, j] bZ [m - i, n - j] \\ &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (X [i, j] aY [m - i, n - j]) + \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (X [i, j] bZ [m - i, n - j]) \\ &= a \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (X [i, j] Y [m - i, n - j]) + b \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} (X [i, j] Z [m - i, n - j]) \\ &= aX * Y [m, n] + bX * Z [m, n] \end{aligned}$$

לכל m, n לכן מתקיים $X * (aY + bZ) = aX * Y + bX * Z$.

שאלה 3

עבור K_x, K_y נגדיר את הלפלסיאן של I באופן הבא:

$$\nabla^2 I = (I * K_x) * K_x + (I * K_y) * K_y$$

$$\text{נניח ש-} K_x = \frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, K_y = \frac{1}{2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \text{ ונגדיר } K_l = \frac{1}{4} \begin{pmatrix} 1 & -4 & 1 \\ 1 & 1 & 1 \end{pmatrix} \text{ מתקיים:}$$

$$\begin{aligned} \nabla^2 I &\stackrel{1}{=} (I * K_x) * K_x + (I * K_y) * K_y \stackrel{2}{=} I * (K_x * K_x) + I * (K_y * K_y) \\ &\stackrel{3}{=} I * (K_x * K_x + K_y * K_y) \stackrel{4}{=} I * \left(\left(\frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \right) * \left(\frac{1}{2} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \right) + \left(\frac{1}{2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \right) * \left(\frac{1}{2} \begin{pmatrix} -1 & 1 \\ 1 & 1 \end{pmatrix} \right) \right) \\ &\stackrel{5}{=} I * \left(\frac{1}{4} \begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \frac{1}{4} \begin{pmatrix} 1 & -2 & 1 \\ 1 & 1 & 1 \end{pmatrix} \right) \stackrel{6}{=} I * \frac{1}{4} \begin{pmatrix} 1 & -4 & 1 \\ 1 & 1 & 1 \end{pmatrix} \stackrel{7}{=} I * K_l \end{aligned}$$

כאשר:

1. מהגדרת הלפלסיאן

2. שאלה 1 (אסוציאטיביות הקונבולוציה)

3. שאלה 2

4. הגדרת K_x, K_y

5. הגדרת הקונבולוציה. למשל, $(K'_x * K'_x)[0, 0]$ כאשר $K'_x = \begin{pmatrix} 1 & -1 \end{pmatrix}$ היא:

$$\begin{aligned} \left(\begin{pmatrix} 1 & -1 \end{pmatrix} * \begin{pmatrix} 1 & -1 \end{pmatrix} \right) [0, 0] &= \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} \begin{pmatrix} 1 & -1 \end{pmatrix} [i, j] \begin{pmatrix} 1 & -1 \end{pmatrix} [0-i, 0-j] \\ &= \sum_{j=-\infty}^{\infty} \begin{pmatrix} 1 & -1 \end{pmatrix} [0, j] \begin{pmatrix} 1 & -1 \end{pmatrix} [0, -j] \\ &= \sum_{j=-1}^1 \begin{pmatrix} 1 & -1 \end{pmatrix} [0, j] \begin{pmatrix} 1 & -1 \end{pmatrix} [0, -j] \\ &= 0 \cdot (-1) + 1 \cdot 1 + (-1) \cdot 0 = 1 \end{aligned}$$

ובאופן דומה $(K'_x * K'_x)[0, 1] = -2$ ו- $(K'_x * K'_x)[0, 2] = 1$ ולכן $K_x * K_x = \frac{1}{4} \begin{pmatrix} 1 & -2 & 1 \end{pmatrix}$.

6. הגדרת חיבור מטריצות (בהתאם להנחות בתרגיל שחלק שאינו מוגדר במטריצה הוא 0 ושהאיבר האמצעי במטריצה הוא האינדקס ה-0 בכל מימד)

7. הגדרת K_l

שאלה 4

נניח שקיים K_α כך ש- $K_l * K_\alpha = \bar{\delta}$ כאשר:

$$\bar{\delta}[i, j] = \begin{cases} 1 & i, j = (0, 0) \\ 0 & \text{otherwise} \end{cases}$$

נשים לב:

$$(\nabla^2 I) * K_\alpha = (I * K_l) * K_\alpha = I * (K_l * K_\alpha) = I * \bar{\delta}$$

מהגדרת הקונבולוציה, לכל m, n :

$$(I * \bar{\delta})[m, n] = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} I[i, j] \bar{\delta}[m-i, n-j] = I[m, n]$$

כאשר המעבר האחרון נובע מהגדרת $\bar{\delta}$:

$$\bar{\delta}[m-i, n-j] = 1 \iff m-i = n-j = 0 \iff i = m, j = n$$

ולכן:

$$(\nabla^2 I) * K_\alpha = I$$

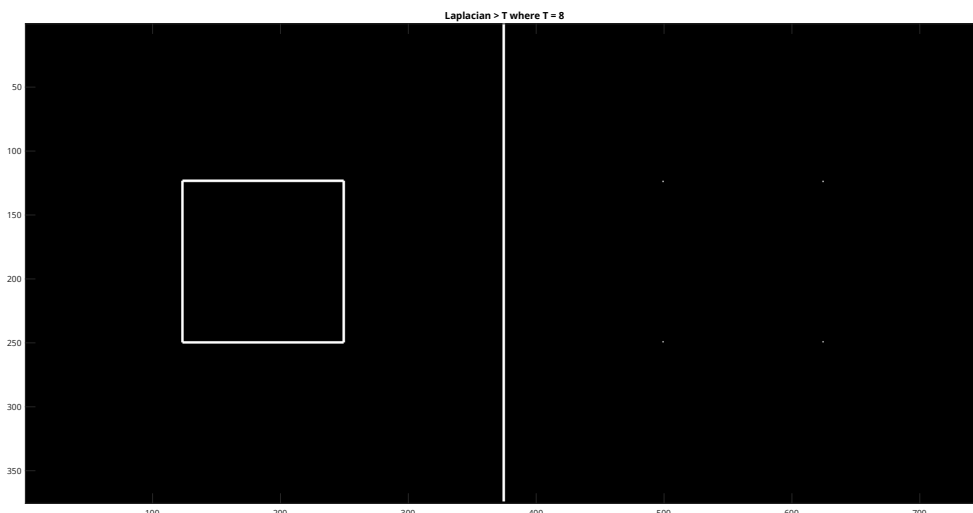
כנדרש.

חלק II

תכנותי

שאלה 4

הריבוע השמאלי נראה בהיר יותר. לאחר חיפוש אחר ערכי T אפשריים, מצאנו כי $T = 8$ מקיים את התכונה הבאה: עבור החצי השמאלי של התמונה ישנם $1,000 <$ פיקסלים שערכם בערך מוחלט גדול מ- T ; עבור החצי הימני ישנם 4 פיקסלים כאלה. באיור המצורף אפשר



איור 1: שאלה 4 - הלפלסיאן בריבוע השמאלי גדול מ- T ובריבוע הימני קטן מ- T (למעט ב-4 פינות הריבוע)

לראות תכונה זאת.

שאלה 5

כמובן שניתן להבחין בצלב הבהיר על אלכסוני התמונה (פינות הריבועים). גם כאן, לאחר חיפוש ממצא על ערכי T אפשריים התקבל ש- $T = 5$ מקיים את התכונה: ישנם פיקסלים גדולים מ- T על אלכסון התמונה (52 כאלה) וכן אין פיקסלים גדולים מ- T מחוץ לאלכסון. כמו בשאלה הקודמת, האיור המצורף מתאר את הממצא.

שאלה 6

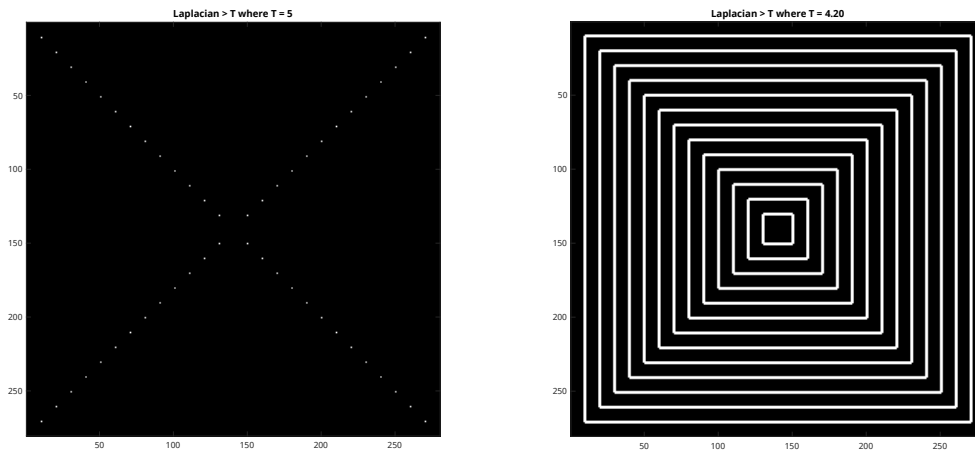
לאחר חיפוש ממצא על ערכי T אפשריים, לא מצאנו ערך שיכול להסביר את התופעה - כלומר לא קיים ערך T עבורו הלפלסיאן בחצי הטבעת העליון גדול מ- T ובחצי התחתון קטן מ- T . נסיק שלא קיים T כזה - כי כמות הפיקסלים בשני חציי הטבעת יורדת באופן הדרגתי ל-0 עבור ערכי T גדלים, ובקצב דומה בשניהם (ראו איור).

שאלה 7

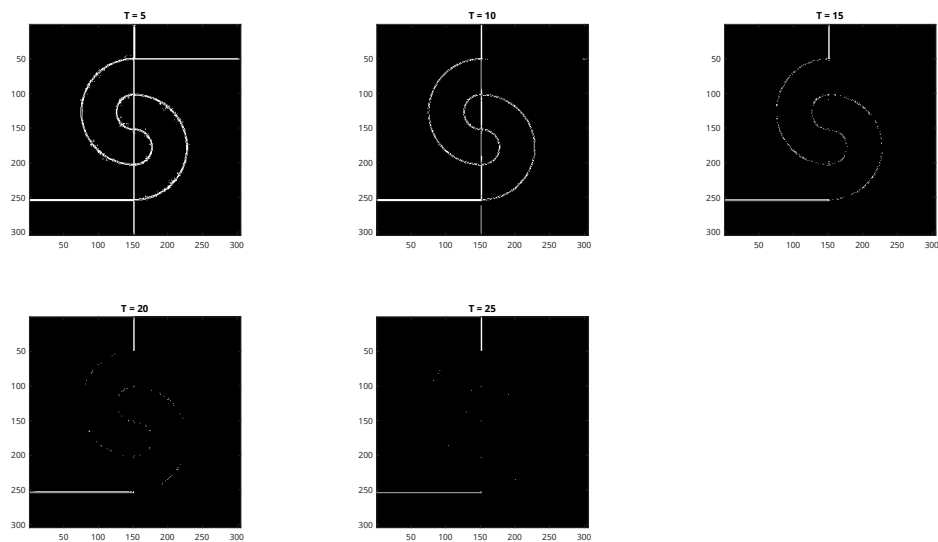
כפי שראינו בשאלות 4-6, חלק מהאשליות ניתן להסביר באמצעות ה-Laplacian וחלק לא. בפרט, ראינו שאת האשליות שהופיעו בשאלות 4, 5 ניתן להסביר באופן זה וזו שב-6 לא.

היפותזה אלטרנטיבית היא שהצבע הנתפס הוא פונקצייה כלשהי של מאפיינים בסיסיים בתמונה ושל ידע מוקדם (prior) לגבי התמונה (או לגבי תמונות טבעיות, באופן כללי). כלומר, ישנה רגולריות בתמונות אותה אנחנו לומדים ומנסים להכליל לתמונות חדשות. דוגמה ל-prior כזה הוא שלכל חומר יש גוון אחיד. כלל זה יסביר את האפקט ב-kofka-ring באופן הבא: אנחנו תופסים את החצי התחתון של הטבעת כעשוי מאותו "חומר" של הרקע השמאלי, ואת החצי העליון כמאותו חומר של הרקע הימני.¹ תחת ההנחה שאובייקט מאותו חומר הוא בעל גוון אחיד, נתפוס את שני חציי הטבעת כבעלי צבע שונה (בגלל ההבדל הגדול בצבעים בין הרקע הימני לשמאלי). מאפייני התמונה במקרה הזה יתפסו תפקיד משני בהבנת הצבע של הטבעת. במילים אחרות, ידע סמנטי ("גבוה") משפיע על תהליכי תפיסה "נמוכים".

¹גם כאן נכנס prior לפיו אובייקטים הם בעלי צורה "חלקה" יחסית, כלומר הם לא נחתכים ומשתנים במהירות גדולה מדי.



איור 2: שאלה 5 - משמאל: הלפלסיאן בפינות הריבוע (על האלכסון) גדול מ- T ובשאר התמונה קטן מ- T ; מימין: תמונה דומה עבור ערך סף $T - \varepsilon$ - כלומר מראה שהסף T שמצאנו הדוק



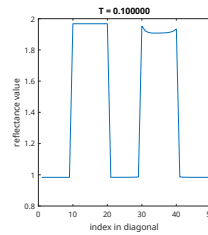
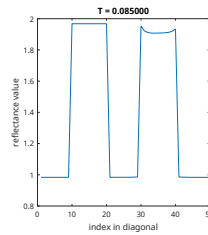
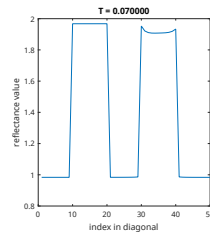
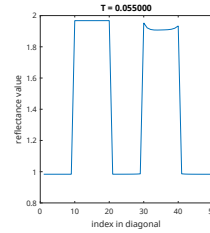
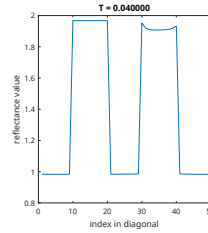
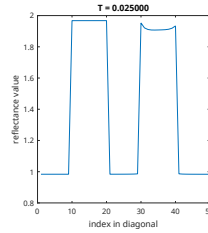
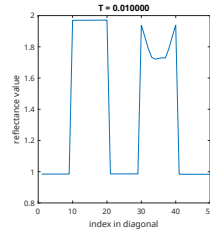
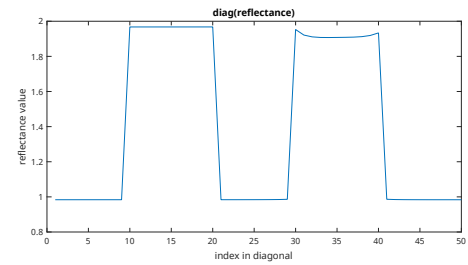
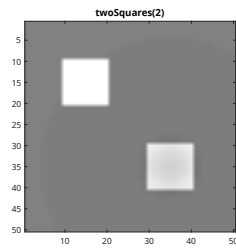
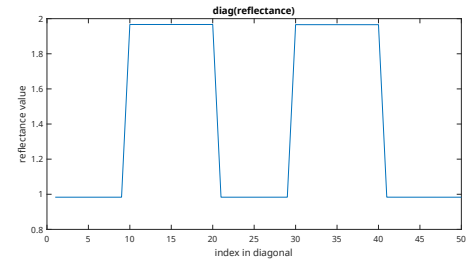
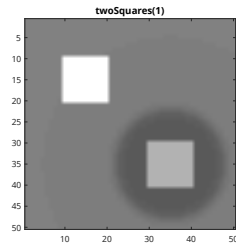
איור 3: שאלה 6 - ערכי הלפלסיאן הגדולים מ- T עבור ערכי T שונים. אפשר לראות שאין ערך בו החלק העליון הוא בעל יותר פיקסלים הגדולים מ- T גדול מבחלק התחתון.

שאלה 9

עבור הגירוי הנוצר עם $\text{twoSquares}(1)$ וערך סף $T = 0.07$ האלגוריתם מחזיר בדיוק את אותו ערך reflectance . עבור הגירוי השני האלגוריתם מחזיר ערך החזרה משתנה בתוך הריבוע המוצל - כלומר מחזיר שהריבוע המוצל כהה יותר במרכזו ובהיר יותר בפריפריה - בהתאמה לצל העמום המוטל עליו.

נזכור כי ההנחות עליהן האלג' מתבסס הן:

- שינוי בסיגנל הוא שינוי במקדם ההחזרה או בתאורה, אך לא בשניהם
 - שינוי חד בתמונה הוא שינוי במקדם ההחזרה ואילו השינוי הדרגתי הוא שינוי בתאורה
- ההנחה המופרת בגירוי השני היא ההנחה לפיה השינוי בסיגנל הוא רק באחד מהגורמים, ולא בשניהם. במקרה זה, הריבוע התחתון מורכב מסיגנל משתנה הן בגלל הצל המוטל עליו והן בגלל הפרדה שלו מהרקע. שני הגורמים משתנים במקביל ואילו האלגוריתם שמתבסס על הנחה (1) משווה את ה(שינוי ב)סיגנל לערך הסף כדי להפריד כל שינוי לאחד משני הגורמים האפשריים. כאשר שני הגורמים לא ברי הפרדה, השוואה כזו לסף נאיבי אינה יעילה ובטח שלא יכולה לחלץ את מידת השינוי בשני הגורמים ביחד. שינוי ערכי T לא עוזר, כפי שאפשר לראות באיור ובהתאם לנדון למעלה.



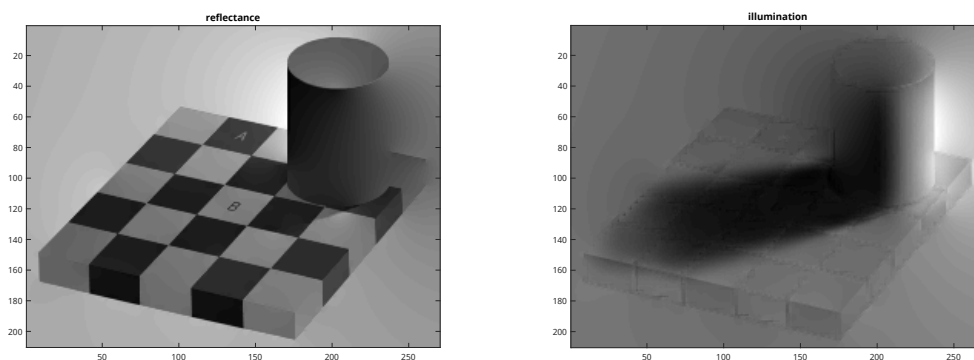
איור 4: שאלה 9 - על גירוי המקיים את ההנחות האלגוריתם מבצע בצורה מושלמת (עם ערך הסף הנבחר). עבור הגירוי השני, אף ערך סף לא נותן תוצאה מושלמת (כפי שאפשר לראות בגרפים התחתונים)

שאלה 10

הנקודות אכן בגוון זהה: $A = B = 0.4196$. אחרי הרצת האלגוריתם, בתמונת מקדמי ההחזרה המתקבלת מתקיים $R(A) = 0.566 <$ $R(B) = 0.876$, בהתאמה לחוויית האשליה. הנקודה A נתפסת כאפורה כהה יחסית, בעוד ש- B נתפסת כלבנה.

על לוח השחמט האלגוריתם מחזיר תוצאות לא רעות, אם כי לא מושלמות. ניתן לראות שהריבועים השחורים בצל כהים יותר מאלה מחוץ לצל ובאופן דומה עבור ריבועים בהירים, אע"פ שהאנחנו תופסים אותם כבעלי צבע (מקדם החזרה) זהה. גם הגליל מקבל ערכי מקדם החזרה שונים בחלקיו השונים בשונה מהתפיסה.

גם כאן מופרת הנחת השינוי הבודד - הצל מוטל על אזורים בהם ישנו גם שינוי בצבע המשבצת, ולכן שינוי ערך הסף לא יפתור את טעות האלגוריתם בחלק זה. בנוסף, מופרת גם ההנחה השנייה לפיה שינוי חד בסיגנל הוא כתוצאה משינוי במקדם ההחזרה - למשל בחלקו העליון של הגליל קיים שינוי חד שנובע משינוי בתאורה ולא ב- ρ . באופן דומה גם בחלקו הימני-תחתון של הגליל יש שינוי קטן בסיגנל שנובע משינוי ב- ρ ולא משינוי בתאורה. גם כאן, שינוי ערך הסף לא צפוי לשפר את התוצאה מסיבות דומות לאלו שהוצגו בשאלה הקודמת.

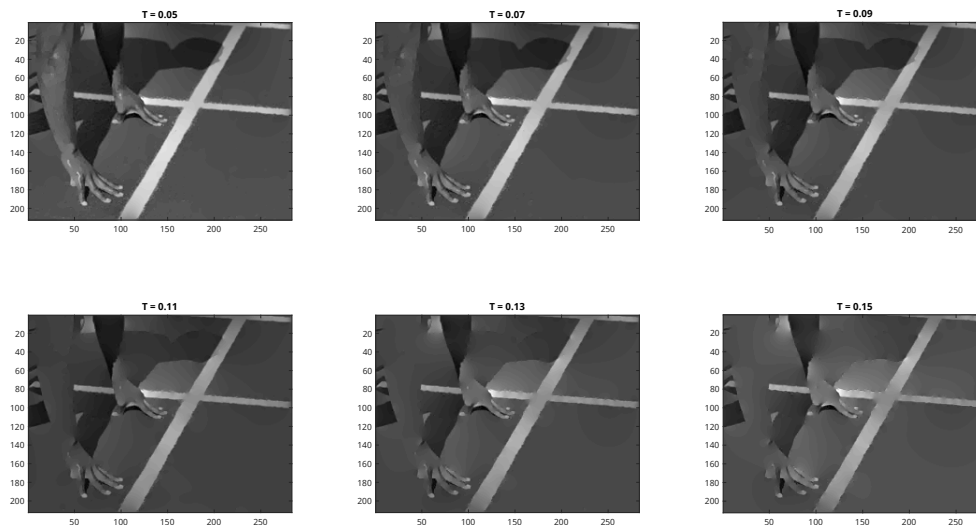


איור 5: שאלה 10 - על לוח השחמט הנחות האלגוריתם מתקיימות בצורה לא רעה ולכן גם ביצועי האלגוריתם בהתאם. על הגליל התוצאות בינוניות בגלל הפרת הנחות.

שאלה 11

האלגוריתם נכשל בהסרת הצל. כפי שנדון בסעיפים הקודמים, ההנחה לפי שינוי חזק בסיגנל הוא תוצאה של שינוי מקדם ההחזרה אינה מתקיימת. לפיכך, הגדלת T תוביל לסיווג של שינויים חדים יותר כנובעים משינויי תאורה "תעלים" את הצל. עם זאת, מכיוון שהשינוי בסיגנל באזור קצוות הצל דומה לשינוי בקצוות היד, ולכן גם ההבדל בין היד למשטח (הרקורטן?) תסווג כשינוי בתאורה, וגם זו תוצאה שגויה.

רמזים שאפשר להוסיף לאלגוריתם הם "מסכה" במימדים זהים לתמונה המקורית, עם ערכים $\{1, \dots, k\}$ עבור קבוע $k \in \mathbb{N}$. תמונה זו תציין לאלגוריתם שבמקומות בהם הערך במסכה זהה, החומר בתמונה המקורית זהה. כך, כאשר נבדק תנאי ה-threshold בקורדינטה כלשהי, נפסול סיווג שינוי בסיגנל כשינוי בתאורה אם הערכים במסכה בקורדינטה זו זהים. באופן דומה, אפשר להזין מסכה דומה המתארת אזורים עם תאורה דומה ושינוי אנלוגי לזה שהוצג קודם בתנאי הסף.



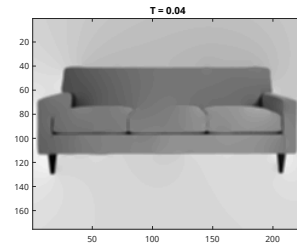
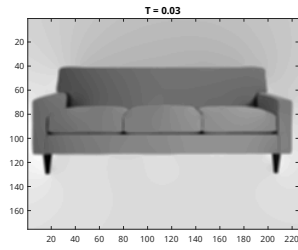
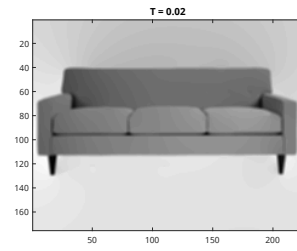
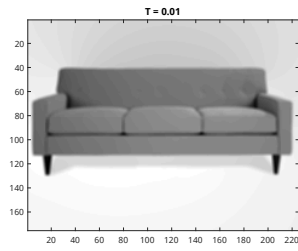
איור 6: שאלה 11 - כישלון בהסרת הצל - כאשר מגדילים את ערך הסף הצל נבלע במשטח, אך גם יד האצן.

שאלה 12

קל לראות שהאלגוריתם מסיר את הצל בהצלחה, עבור ערכי T לא קטנים מדי. עבור ערכי הסף הגדולים יותר, גם השינויים בתאורה בגב הספה (איפה שהכפתורים) נעלמים, ואנחנו מקבלים תמונה כמעט מושלמת של מקדמי ההחזרה כפי שהם נתפסים.

בתמונה זו ההנחות של האלגוריתם מתקיימות בצורה טובה: שינויים חדים בסיגנל נובעים משינוי בחומר ושינויים חלשים נובעים משינוי בתאורה.

ההבדל האיכותי בין תמונה זו לבין התמונות הקודמות, לדעתנו, הוא הפשטות של התמונה - הגיאומטריה של האובייקט המוצג היא פשוטה (משטחים ארוכים וישרים, אין רקע, אין עצמים עגולים המחזירים אור בצורה משתנה לאורך פני שטחם וכו') וכן התאורה היא אחידה ונראית כמוטלת ממקור יחיד מגבוה. לעומתה, בתמונה הקודמת יש הרבה פרטים ושינויים חזקים בסיגנל שנובעים מתאורה חזקה ומשטחים מורכבים (הידיים) המחזירים את האור בצורה שונה באופן דרמטי בחלקיהם השונים.



איור 7: שאלה 12 - אין הרבה הבדל בין ערכי T השונים (למעט אולי הקטן ביותר בו האלגוריתם לא ממש עובד) - כולם מצליחים במידה סבירה.

חלק III

קוד