



## חלק א' – עיבוד תמונה מבוסס משוואת דיפוזיה

### הקדמה

#### *משימות מרכזיות של עיבוד תמונה*

ראשית, נסקור מספר משימות בסיסיות של עיבוד תמונה. משימות אלה עומדות בבסיס של אפליקציות רבות בתחומים כמו טלוויזיה דיגיטלית, הדמיה רפואית, ראייה ממוחשבת, מצלמות דיגיטליות ועוד.

– המשימה הראשונה היא סגמנטציה, קרי חלוקה של התמונה לסגמנטים איזורים קשירים בעלי רמת אפור אחידה. זהו שלב מקדים הכרחי באפליקציות רבות של הדמיה רפואית וראיה ממוחשבת. שלב זה יכול לסייע בגילוי שפות ומציאת איזורי עניין בתמונות.

המשימה השנייה קשורה לבעיה הבאה - כל מערכת חישה או תקשורת סובלים מהתווספות רעש לסיגנל הרצוי. השאיפה היא לקבל תמונה שמצד אחד היא נקיה ככל הניתן מרעש ומצד שני שומרת על רוב המאפיינים הבולטים שלה, כמו שפות, מרקמים וצבע. הרעיון הבסיסי של ניקוי מרעש משתמש בעובדה שרעש מאופיין על ידי תדרים גבוהים ולכן עושה שימוש ב-LPF כדי להפחית את הרעש. כל שימוש ברעיון כזה בהכרח סובל מ-trade-off בין הפחתת רעש לאיבוד חדות ופרטים קטנים. אומנם קיימים אלגוריתמים מבוססי מודל סטטיסטי לניקוי מרעש שמניחים התפלגות מסוימת של הרעש ומשיגים שחזור טוב של התמונה הרצויה, אך במקרה שהנחה זו לא תקפה כבר לא ניתן להבטיח ניקוי טוב. על כן, יש העדפה לאלגוריתם שידע לטפל בסוגים רבים של רעש ללא הנחה מוקדמת לגבי אופי הרעש.

סוג אחר של פגיעה באיכות התמונה מתרחש כאשר המצלמה לא בפוקוס או שהעצם/המצלמה בתנועה, ואז התמונה המתקבלת היא מטושטשת ומרוחה. לרוב מקובל להחזיר לתמונה את חדותה על ידי הנחת מודל לגבי סוג הטשטוש (למשל תנועה), שערך הפרמטרים שלו ולבסוף הפעלת מסנן להסרת הטשטוש. יש תרחישים שבהם קשה לבנות מודל שכזה או שהטשטוש נובע ממספר מקורות (ואז המודל יהיה הרבה יותר מורכב). על כן, יש עניין באלגוריתם שידע להגביר את החדות בתמונה מבלי להניח שום הנחות לגבי המקורות לטשטוש.

בעיה מורכבת יותר היא בעיה של שיפור תמונה, ובפרט תמונה שעברה אינטרפולציה מרזולוציה נמוכה לגבוהה יותר. תמונה כזו סובלת ממספר בעיות שעיקרן רעשי אינטרפולציה (בצורת "טבעות" מסביב לשפות), אובדן חדות ופיקסליזציה. כלומר יש כאן שילוב של מספר עיוותים שפוגעים באיכות התמונה ומכאן הקושי בבניית מודל לאפיון עיוותים אלה.



ניתן להתייחס לבעיה כאל בעיית חיזוי ועל ידי הנחת מודל לגבי ספקטרום התמונה ניתן לבנות מסנן "אופטימלי" מתאים. אך גישות אלה עדיין לא מצליחות להתמודד בצורה מספקת עם כל העיוותים ש תוארו, בשל חוסר התאמה של איזורים שונים בתמונה למודל הגלובלי שהונח לגבי הספקטרום. עבור כל המשימות שתוארו לעיל פותחו לאורך השנים אלגוריתמים ספציפיים המנסים לתקוף את כל אחת מהבעיות בנפרד תוך שימוש בהנחות מסוימות ובניית כלים המתאימים לבעיה הספציפית. בחלק הראשון של ניסוי זה נציג שיטה בעלת ביסוס מתמטי ופיזיקלי נרחב, אשר מנסה לטפל בכל המשימות תחת קורת גג אחת – משוואות דיפרנציאליות חלקיות. ההתאמה לכל משימה תעשה על ידי שינוי פרמטרי המשוואה. גישה זו תוקפת את בעיות אלו על ידי התמקדות במבנה בסיסי של תמונות (שפות, זרימה, חלקות וכו') וניסיון להדגיש תכונות מסוימות בתמונות שמערכת הראייה האנושית מעדיפה.

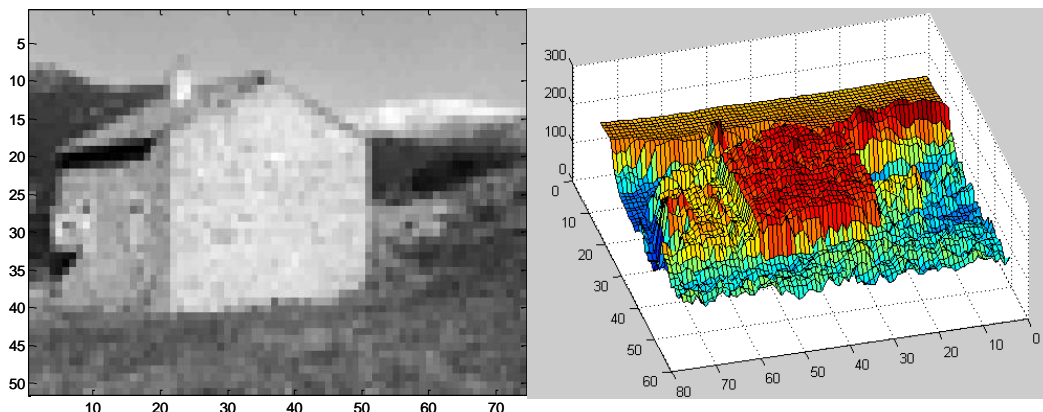
אנחנו נתמקד בשתי משימות של עיבוד תמונה:

1. ניקוי מרעש

2. סופר-רזולוציה מתמונה בודדת

### הקשר בין משוואות דיפרנציאליות חלקיות לעיבוד תמונה

בתחום של עיבוד תמונה נהוג לחשוב על תמונה כאל אות דו- מימדי. השינוי המחשבתי הראשון הנדרש הוא לחשוב על תמונת רמת אפור כעל משטח תלת מימדי, שבו ציר  $z$  מציין את ערך רמת האפור (ראה איור למטה). תמונה דיגיטלית היא דגימה של משטח זה בכל אחד מהצירים (סריג דגימה במישור  $x-y$  וקוונטיזציה בציר  $z$ ). ניתן לייחס למשטח זה משמעויות פיזיקליות שונות, כאשר הדוגמא הבולטת ביותר היא מפת ריכוזי חלקיקים, ובה נתרכז בחלק הראשון של הניסוי. כעת ניתן להפעיל על המשטח הנ"ל (המהווה את התמונה שלנו) תהליכים פיזיקליים. תהליכים אלה (שינויים זמניים ומרחביים של משטח) מתוארים בצורה טבעית על ידי משוואה דיפרנציאלית חלקית. במקרה של מפת ריכוזי חלקיקים המד"ח המתאימה היא משוואת דיפוזיה.





הרעיון הבסיסי הוא לעשות שימוש במד "ח בעל משמעות פיזיקלית על מנת להניע תהליכים של עיבוד תמונה. בתהליך של עיבוד תמונה אנו מעוניינים להביא לשינוי בתמונה בכיוון "רצוי" כלשהו, למשל הדגשת שפות או החלקה של איזורים חלקים יחסית. ניתן לחשוב על יצירת שינוי זה כאל תהליך שבו מקשרים בין ערכי רמות אפור של פיקסלים בסביבה קרובה. מכאן המוטיבציה להשתמש במשוואה המתארת דיפוזיה של חלקיקים. תהליך של דיפוזיה המושך לכיוון של השוואת ריכוזי חלקיקים אנלוגי לתהליך של החלקה בתמונה.

## 1. סקירת תהליכי דיפוזיה שונים

### מדיפוזיה לינארית לדיפוזיה לא לינארית

משוואת הדיפוזיה הבסיסית ביותר היא משוואת דיפוזיה לינארית הנתונה על ידי  $I_t = c\nabla^2 I$ ,  $I(t=0) = I_0$ , כאשר  $c$  הוא מקדם הדיפוזיה והוא הומוגני וקבוע בזמן. נשים לב כי זוהי משוואה פרבולית. למשוואה זו קיים פתרון אנליטי והוא נתון על ידי:

$$(1.1) I(x, y, t) = I_0(x, y) * G(x, y, t), \text{ where } G(x, y, t) = \frac{1}{\sqrt{4\pi ct}} e^{-\frac{x^2+y^2}{4ct}}$$

כלומר התמונה ברגע  $t$  מתקבלת על ידי קונובולוציה של התמונה המקורית עם גאוסיאן שרוחבו גדל עם  $t$ . מכאן ניתן לראות שלאחר זמן ארוך מספיק התמונה תהיה בקירוב בעל רמת אפור אחידה השווה לממוצע רמות האפור של התמונה המקורית.

המסקנה מהקביעה האחרונה היא שדיפוזיה לינארית מבצעת החלקה גלובלית. החסרון בכך הוא באיבוד השפות והפרטים הקטנים. מכאן ניתן להבין שרצוי לתת טיפול שונה לאיזורים חלקים יחסית ולשפות. כלומר אנו מעוניינים להכניס לוקליות לתהליך הדיפוזיה וזאת על ידי מקדם דיפוזיה לוקלי שתלוי בתמונה (ליתר דיוק בגרדיאנטים שלה). כמו כן, להשגת דרגת חופש נוספת נאפשר למקדם הדיפוזיה להשתנות גם בזמן.

באופן זה מתקבלת משוואת דיפוזיה לא לינארית מהצורה:

$$(1.2) I_t = \nabla \cdot (c(x, y, t)\nabla I)$$

כדי להפריד את הטיפול בין איזורים חלקים לשפות נקבע את מקדם הדיפוזיה להיות פונקציה מונוטונית יורדת של עוצמת הגרדיאנט, כלומר פונקציה



מהצורה  $c(x, y, t) = g(\|\nabla I(x, y, t)\|)$  . באופן זה מקדם הדיפוזיה יהיה גבוה עבור איזורים חלקים יחסית, כלומר יעודד החלקה נוספת שלהם, ויהיה נמוך עבור שפות, כלומר ישמר אותם.

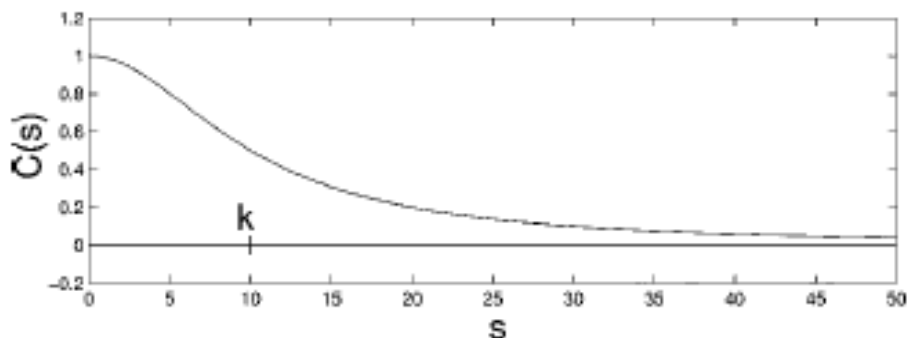
סוג כזה של משוואה יביא כעבור מספיק זמן לתמונה שהיא בעל אופי חלק למקוטעין. כלומר היא מתאימה בצורה טובה לביצוע משימה של סגמנטציה וזיהוי שפות בצורה אמינה.

### הצעות לקביעת מקדם הדיפוזיה

הגישה שתוארה לעיל הוצגה ב- 1990 במאמרם של Perona & Malik (להלן PM).

ההצעה של PM עבור מקדם הדיפוזיה היתה: 
$$(1.3) \quad g(s) = \frac{1}{1 + (s/K_1)^2}$$

התנהגות מקדם הדיפוזיה כתלות בעוצמת הגרדיאנט מתוארת בגרף הבא:



ראוי לציין ששימוש במקדם כזה משמר את המינימום והמקסימום הגלובלי בתמונה ומביא ליציבות של התהליך.

כפי שצוין קודם לכן הצעה זו מתאימה לביצוע משימה של סגמנטציה . עם זאת, היא איננה מתאימה לבעיה של ניקוי מרעש שכן רעש מאופיין אף הוא על ידי עוצמה גבוהה של גרדיאנטים. ההצעה של PM תביא לשימור של חלק ניכר מהרעש כיוון שהוא יקבל טיפול דומה לזה של השפות.



Catte ושות' הציעו ב-1992 רעיון כיצד לפתור בעיה זו. לצורך חישוב מקדם הדיפוזיה הם הציעו להשתמש בגרסה מוחלקת של התמונה. כלומר המשוואה מקבלת עתה את הצורה:

$$(1.4) I_t = \nabla \cdot (g(|\nabla I_v|) \nabla I), I_v = I * \text{guass\_filter}$$

כאשר רוחב הפילטר הגאוסית הוא  $\sigma$ . באופן זה, רעש ושפות מקבלים כעת טיפול שונה – שפות יהיו בעלות גרדיאנטים גבוהים יחסית גם בתמונה המוחלקת, בעוד שלרעש יתאימו גרדיאנטים נמוכים יחסית לאחר ההחלקה (בגלל שלרעש יש ממוצע אפס).

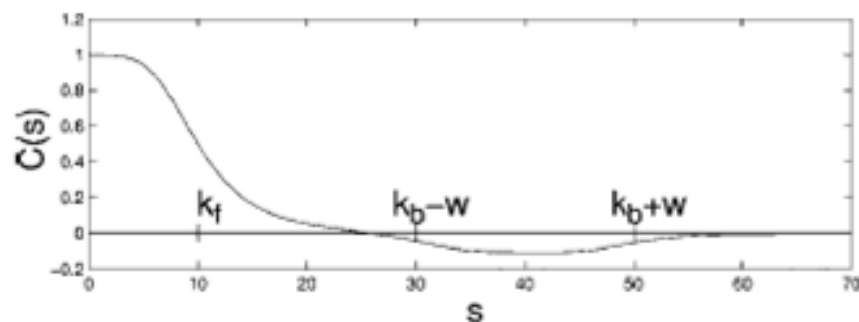
משוואות הדיפוזיה שהוצעו עד כה היו בעלות מאפיינים פיזיקליים, קרי זרימה של חלקיקים עם מפל הריכוזים בלבד. השליטה על התהליך נעשתה על ידי שינוי מקדם הדיפוזיה באופן לוקלי. מקדם זה קיבל ערכים בין 0 ל-1 בלבד והיה מונוטוני יורד בעוצמת הגרדיאנט, וכך ניתן היה לשמר את השפות המשמעותיות בתמונה.

ב-2004 גילבוע, סוכן וזאבי הציעו להסיר את המגבלה הפיזיקלית ולאפשר למקדם הדיפוזיה לקבל ערכים שליליים קטנים עבור תחום גרדיאנטים מסוים. תחום זה הוא תחום של גרדיאנטים גבוהים יחסית, כלומר תחום שאמור לייצג את השפות המשמעותיות בתמונה. המשמעות של שינוי זה הוא שכעת חלקיקים יכולים לזרום כנגד מפל הריכוזים ובכך להגדיל את הגרדיאנט. מטרת ההצעה היתה לאפשר הדגשה של שפות, שמהווה נדבך חשוב במשימה של שיפור תמונה.

חשוב להדגיש שהצעה זו היא "לא פיסיקלית" מהגדרתה, אך תחת בחירה נכונה של הפרמטרים בתהליך ניתן למנוע התבדרות ולקבל את התוצאה הרצויה – הדגשת שפות. הצעה זו זכתה לשם **FAB** (Forward and Backward diffusion) ומקדם הדיפוזיה שלה הוא מהצורה:

$$(1.5) g(s) = \frac{1}{1 + \left(\frac{s}{K_f}\right)^4} - \frac{\beta}{1 + \left(\frac{s - K_b}{W}\right)^4}$$

התנהגות מקדם הדיפוזיה כתלות בעוצמת הגרדיאנט מתוארת בגרף הבא:





הסיבה לצורה שנבחרה עבור הגרדיאנטים הגבוהים היא כדי לאפשר הדגשת שפות מבוקרת ולמנוע הדגשת יתר.

הצעות לקביעת ערכי  $K_f, K_b, W$  :

א. חישוב עוצמת הגרדיאנט הממוצע בכל חלון , שמסומן ב-MAG (Mean Absolute Gradient), ולקחת  $[K_f, K_b, W] = [2, 4, 1] * MAG$  (1.6).

ב. חישוב היסטוגרמה של עוצמות הגרדיאנטים בכל חלון שגדולות מ-MAG.  $K_f$  נקבע להיות  $\gamma * MAG$ , ואילו קצוות תחום הגרדיאנטים שאותו אנו מעוניינים להגביר ( $K_b - W, K_b + W$ ) נקבעים על ידי העברת ספים (מסומנים ב-per1 וב-per2) על פונקציית ההתפלגות המצטברת של עוצמות הגרדיאנטים שגדולות מ-MAG.

הערה - מכיוון שתמונות טבעיות הן ברובן חלקות, מירב הגרדיאנטים מרוכזים בפיק גדול ב-0. על כן עדיף להסתכל רק על החלק בהיסטוגרמה שמתאים לגרדיאנטים גבוהים יחסית, לצורך קביעת תחום הגרדיאנטים שאותו אנו מעוניינים להגביר.

תיאור מדויק של ההצעה השניה מובא כאן :

$$g[n] = \sum_{k=1}^n h[k]$$

$h[n]$  is a normalized histogram of the gradient magnitudes that are larger than MAG (Mean Absolute Gradient)

$$a = x[\text{argmin}(|g[n] - \text{per1}|)]$$

$$b = x[\text{argmin}(|g[n] - \text{per2}|)]$$

where  $x$  is a vector containing the bins of the histogram

$$(1.7) K_f = \gamma * MAG, K_b = \frac{a+b}{2}, W = K_b - a$$

## 2. סקירת שיטות נומריות לפתרון מד"ח

המשוואות שהוצגו בפרק הקודם הן משוואות רציפות בזמן ובמקום. מכיוון שמפעילים אותן על תמונות דיגטליות יש להחליף את אופרטורי הגרדיאנט



והדיברגנס בגירסאות דיסקרטיות. כמו כן, המשוואות הן "מאוד" לא לינאריות ולכן אין להן פתרון אנליטי. לכן, גישה שבה ניתן לנקוט כדי לפתור משוואות אלו היא לפתור אותן בצורה איטרטיבית וזוהי מהות הגישה הנומריית. גישה זו מכניסה אלמנט דיסקרטי גם לציר הזמן של התהליך. כלומר, בכל צעד זמן מקבלים תמונה חדשה מתוך תמונה בצעד הזמן הקודם. ככל שגודל הצעד גדול יותר ניתן לבצע שינוי משמעותי יותר בין תמונות עוקבות, אך עלולים לסבול מאי-יציבות נומרית (ואף התבדרות) אם צעד הזמן גדול מידי.

מטרת פרק זה היא לא להעביר קורס מזורז בשיטות נומריות, אלא להדגיש את הקשיים העומדים בפנינו כשאנו באים לעשות דיסקרטיזציה של מד"ח רציפות ולפתור מד"ח לא לינאריות. בפרק זה נרצה להבין את חשיבות בחירת גודל הצעד בתהליך הדיפוזיה הדיסקרטי ואת המקור לבעיות של אי יציבות נומרית.

### דיסקרטיזציה של אופרטורי הנגזרת במקום ובזמן עבור אות חד מימדי

$$\frac{\partial f}{\partial t} \rightarrow \frac{f^{k+1} - f^k}{\tau}, \text{ כאשר } \tau \text{ הוא גודל צעד הזמן}$$

$$\left(\frac{\partial f}{\partial x}\right)_i \rightarrow \frac{f_{i+1}^k - f_{i-1}^k}{2} \text{ (נגזרת מרכזית)}$$

### פתרון נומרי של משוואת הדיפוזיה החד-מימדית

משוואת הדיפוזיה הלא לינארית החד-מימדית בזמן רציף מקבלת את הצורה:

$$u_t = \frac{\partial}{\partial x} \left( g \left( \left| \frac{\partial u}{\partial x} \right| \right) \frac{\partial u}{\partial x} \right)$$

הגישה הפשוטה ביותר היא גישה של **explicit** – לפי גישה זו הדיסקרטיזציה של אגף ימין של המשוואה נעשית רק לפי האות באותו רגע ( $u^k$ ). לאחר דיסקרטיזציה מקבלים את המשוואה הבאה:

$$\frac{u^{k+1} - u^k}{\tau} = A(u^k)u^k \quad \Rightarrow \quad (2.1) \quad u^{k+1} = (I + \tau A(u^k))u^k$$

כאשר  $A(u^k)$  היא מטריצה שמהווה גרסה דיסקרטית של אופרטור הגזירה המרחבית, כולל הפונקציה הלא לינארית  $g(\cdot)$ . קיימים ביטויים מפורשים למטריצה  $A$ . השוני בין ביטויים אלו נובע מקירובים שונים לנגזרות המרחביות. נביא לדוגמה אחד מהם שעושה שימוש בנגזרת מרכזית (אין צורך להתעמק בפיתוח שלו):



$$A_{ij}(u^k) = \begin{cases} \frac{g_j^k + g_i^k}{2h^2}, & j \in N(i) \\ -\sum_{l \in N(i)} \frac{g_l^k + g_i^k}{2h^2}, & j = i, \\ 0, & \text{else} \end{cases} \quad \text{where} \quad g_i^k = g \left( \frac{u_{i+1}^k - u_{i-1}^k}{2h} \right)$$

גישה זו מניבה סכימה איטרטיבית. כלומר האות ברגע הבא  $u^{k+1}$  מתקבל על ידי הכפלה של מטריצה, שחושבה מתוך האות הנוכחי, באות הנוכחי. הסיבוכיות החישובית של איטרציה אחת היא נמוכה, אך מסתבר שיש הגבלה תיאורטית על גודל הצעד להבטחת יציבות נומרית -  $\tau \leq 0.5$ . לכן, אם ברצוננו לקבל את התמונה לאחר זמן  $t$  מסוים יש צורך לבצע מספר איטרציות בהתאם לגודל הצעד שנבחר. כלומר אנו משלמים בסיבוכיות זמן חישוב מבחינת מספר הצעדים הדרוש להגעה לזמן  $t$ .

דרך להתגבר על מגבלה זו של גודל צעד מקסימלי היא להשתמש בסכימה שהיא

**semi-implicit**. בסכימה זו המטריצה  $A$  עדיין מחושבת מתוך האות ברגע הנוכחי, אולם כעת היא מוכפלת באות ברגע הבא. סכימה זו מניבה מערכת משוואות לינארית מהצורה הבאה:

$$\frac{u_i^{k+1} - u_i^k}{\tau} = A(u^k)u^{k+1} \Rightarrow (2.2) \quad (I - \tau A(u^k))u^{k+1} = u^k$$

מספר הנעלמים  $N$  הוא כגודל האות ברגע הבא ומספר המשוואות הוא  $N$ . באופן כללי, מדובר במטריצה מאוד גדולה ( $N \times N$ ) והמחיר של הפיכתה הוא רב מבחינה חישובית ( $O(N^2)$ ).

עם זאת, במקרה שלנו מטריצת המערכת  $A$  היא תלת-אלכסונית (יש לה איברים שונים מאפס רק באלכסון הראשי ובשני האלכסונים המסמוכים אליו). עבור מטריצה כזו קיים אלגוריתם מוכח, שהוא אלגוריתם Thomas, לפתרון יעיל ומדויק של מערכת המשוואות (חישוביות של  $O(N)$ ). מבחינה תיאורטית ניתן להשתמש בסכימה זו בכל צעד זמן שיידרש, גדול כרצוננו.

**פתרון נומרי של משוואת הדיפוזיה עבור המקרה הדו-מימדי (תמונות)**

בגישה של explicit ההכללה היא מיידית, כאשר עבור כל אחת מהנגזרות ב-  $x$  וב-  $y$  מבצעים המרה לאות חד-מימדי על ידי RS ו-CS בהתאמה לפני חישובן ואז לאחר כפל במטריצה  $A$  המתאימה חוזרים לגודל התמונה המקורי.





ניתן לסכם זאת באופן הבא :

$$(2.3) \quad u^{k+1} = u^k + \tau(A_x(u^k)u_{rs}^k)_{\text{reshape to } u} + \tau(A_y(u^k)u_{cs}^k)_{\text{reshape to } u}$$

כאן המגבלה התיאורטית על גודל הצעד המקסימלי להבטחת יציבות נומרית היא  $\tau \leq 0.25$ .

הערה – הסכימה האיטרטיבית לא מחייבת עבודה עם מטריצה . השימוש במטריצה בעיקר מאפשר הצגה נוחה של השיטות. ניתן להתייחס למטריצה כאל אופרטור מרחבי.

בשורה התחתונה, ניתן פשוט לחשב מפורשות לכל פיקסל את הנגזרות המרחביות בכל אחד מהכיוונים שנבחרו בתמונה ולכפול אותן במקדמי הדיפוזיה המתאימים להן. בשיטות PM ו-FAB, שבהן נשתמש בניסוי, נעבוד עם 4 כיוונים בתמונה (צפון, דרום, מערב ומזרח), במקום 2 כיוונים (אנכי ואופקי) שבהם נעשה שימוש כאשר מחשבים נגזרות מרכזיות. בהתאם לכך, יש גם 4 מקדמי דיפוזיה מתאימים לכל פיקסל. סכימה איטרטיבית זו נראית כך :

$$(2.4) \quad u_{i,j}^{k+1} = u_{i,j}^k + \tau(c_N \nabla_N u + c_S \nabla_S u + c_E \nabla_E u + c_W \nabla_W u)_{i,j}^k$$

$$c_m = g(|\nabla_m u|), \quad m \text{ is the direction of the gradient : N,S,E,W}$$

בגישה של semi-implicit ההכללה לדו-מימד היא :

$$(I - \tau(A_x(u^k) + A_y(u^k)))u^{k+1} = u^k$$

כאשר כעת  $A_x$  ו- $A_y$  הם אופרטורים כלליים ולא מטריצות כיוון ש- $u$  הוא תמונה ולא וקטור. אם היינו רוצים להפוך את  $A_x$  ו- $A_y$  למטריצות היה צורך לסדר את  $u$  ב- RS או CS, ואז חיבור המטריצות  $A_x(u^k) + A_y(u^k)$  היה בהכרח מביא למטריצה שהיא לא תלת-אלכסונית כי רק אחת ממטריצות אלו היא תלת אלכסונית ( $A_x$  במקרה של RS ו- $A_y$  במקרה של CS). כלומר, לא ניתן להשתמש באלגוריתם Thomas ואין דרך יעילה לפתור במדויק את מערכת המשוואות. לכן מוצעת דרך אחרת, מקורבת, לפתרון הבעיה.

### AOS – Additive Operator Splitting

ראשית, נציג את הבעיה כבעיה של הפיכת אופרטור:

$$(2.5) \quad u^{k+1} = (I - \tau(A_x(u^k) + A_y(u^k)))^{-1} u^k$$

לצורך פתרון בעיה זו נשתמש בקירוב AOS:

$$(2.6) \quad u^{k+1} = \frac{1}{2} [(I - 2\tau A_x(u^k))^{-1} + (I - 2\tau A_y(u^k))^{-1}] u^k$$

לקירוב זה יש את אותו פיתוח טיילור מסדר ראשון כמו לבעיה המקורית.

כעת, לאחר שימוש בקירוב זה ניתן להציג את הבעיה בתור:



$$(2.7) \quad u^{k+1} = \frac{1}{2} (u_1^{k+1} + u_2^{k+1})$$

$u_1$  is the solution of  $(I - 2\tau A_x(u_{rs}^k)) u_{rs}^{k+1} = u_{rs}^k$

$u_2$  is the solution of  $(I - 2\tau A_y(u_{cs}^k)) u_{cs}^{k+1} = u_{cs}^k$

כלומר, עברנו מבעיה אחת שאנו לא יודעים לפתור בצורה יעילה לשתי בעיות שניתנות לפתרון מדויק ויעיל על ידי אלגוריתם Thomas. המחיר ששילמנו הוא קירוב, אך ניתן להאמין שלא יהיה לדבר השפעה מבחינה ויזואלית

במהלך הניסוי אנחנו לא נשתמש בשיטות שהן semi-implicit, ובפרט בסכימת AOS. עם זאת, חשוב לדעת שניתן בשיטה זו להגיע לאותם ביצועים כמו בשיטת explicit, רק שכעת נזדקק רק למספר בודד של צעדים ולא למספר רב של צעדים כמו בשיטת explicit.

### 3. ניקוי מרעש – שימוש במדד PSNR כמדד לטיב הניקוי

בפרק הראשון הוצגו שתי הצעות למקדמי דיפוזיה יכולים לשמש אותנו במשימה של ניקוי מרעש. הצעות אלו אמורות לנקות את הרעש, תוך שימור השפות (PM) ואף הדגשתן (FAB). אנו מעוניינים להשוות בין ביצועי ההצעות השונות.

קריטריון הטיב המתמטי המקובל במשימות של שחזור תמונה מקורית

$$(3.1) \quad PSNR = 10 \log_{10} \left( \frac{255^2}{E[(I_{org} - I)^2]} \right)$$

מתוך תמונה שעברה שינוי הוא:

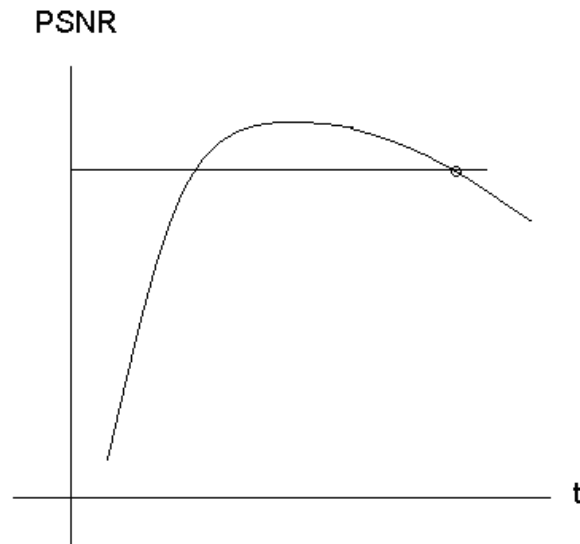
נשים לב שה-PSNR מחושב ביחידות של dB (דציבלים).

חשוב לציין כי תמונות בעלות ערך דומה של PSNR יכולות להיות בעלות שוני רב מבחינה ויזואלית, כי ה-PSNR הוא קריטריון גלובלי שמחושב על כל התמונה. קיימים מאפיינים שמערכת הראיה האנושית מאוד רגישה לכל שינוי בהם, כמו צורות גיאומטריות וחדות של שפות. יש מצבים שבהם עיוות מסוים במאפיינים אלו יתבטא בירידה קלה בלבד ב-PSNR, אך מבחינה ויזואלית הפגיעה תהיה משמעותית. לכן לא נוכל להסתפק רק בהשוואה המתמטית, ויש צורך גם בהשוואה ויזואלית. עבור כל תמונה נוכל להתרשם מאיכותה, הן מבחינת כמות הרעש שנוקה והן מבחינת הפגיעה בחדות ובפרטים הקטנים של התמונה.



בנוסף, מכיוון שאנו עוסקים בתהליכים והתוצאה הסופית של התהליך תלויה מאוד בזמן ש בו נמשך התהליך (מספר צעדי הזמן בסכימה האיטרטיבית), נשאלת השאלה באיזו נקודת זמן צריך לבצע את ההשוואה המתמטית והויזואלית בין ההצעות השונות.

כפי שנראה בניסוי, גרף אופייני של ה-PSNR כתלות בזמן התהליך נראה כך:



יש קודם כל עליה מונוטונית (בדרך כלל חדה) עד לערך של מקסימום גלובלי ולאחריו ירידה מונוטונית, שהיא על פי רוב מתונה יותר מהעליה. נקודת המקסימום הגלובלי היא הנקודה שבה מבחינה תיאורטית אמורים לקבל את השחזור הטוב ביותר של התמונה המקורית מתוך התמונה הרועשת. לכן נרצה להשוות בין השיטות השונות, הן מבחינת ערך המקסימום של ה-PSNR והן מבחינה ויזואלית בנקודת המקסימום.

בנוסף, כדי לבצע השוואה ויזואלית "הוגנת" ולהתרשם איכותית מה כל שיטה מעדיפה בעת ביצוע משימה של ניקוי מרעש, נבצע השוואה ויזואלית בין התמונות שהתקבלו מכל אחת מהשיטות כאשר הן מגיעות בקירוב לערך מסוים של ה-PSNR בשלב הירידה מהמקסימום (הנקודה שמסומנת באיור למעלה).

#### 4. דיפוזיה כיוונית מגבירת קוהרנטיות (CeD)

##### מבוא

עד כה בכל משוואות הדיפוזיה שהצגנו נעשה שימוש רק בעוצמת הגרדיאנט על מנת לחשב את מקדם הדיפוזיה בכל פיקסל. בחלק השני של הספר נרצה להתמקד בתכונה נוספת מלבד שפות (עוצמת גרדיאנטים) שיש לה חשיבות במערכת הראיה אנושית והיא זרימה וכיווניות בתמונה. למעשה, לא ניצלנו עד



כה את המידע שיש בידינו לגבי כיווני הגרדיאנטים . בחלק זה של הספר נראה כיצד ניתן לנצל מידע זה על מנת ליצור דיפוזיה כיוונית

ניתן לשאוב מוטיבציה לחקר הזרימה בתמונה מתחום האומנות . לטכניקה של משיכת המכחול של הצייר יש השפעה רבה על התחושה שהצופה מקבל מהציור . התנועה האימפרסיוניסטית במאה ה-19 חקרה לעומק את נושא זה וציוריה התאפיינו במשיכות מכחול עזות, שיצרו תחושה של זרימה והתרחשות בתמונה. לכן, ניתן לקוות שהגברה של זרימה בתמונה תוכל לתרום לשיפורה מנקודת הראות של הצופה.

בהתאם לכך בחינת טיב התמונות בחלק זה של הספר לא יעשה בעזרת קריטריון טיב מתמטי (כמו למשל ה-PSNR שבו נעשה שימוש בפרק של ניקוי מרעש), אלא יהיה יותר ויזואלי ויהיה קשור למאפיינים שחשובים למערכת הראיה האנושית.

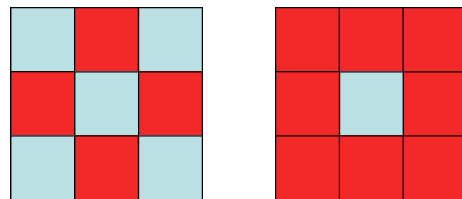
### הצגת אופרטור גזירה חדש

Scharr ו-Weickert הציעו ב-2002 להחליף את אופרטור הנגזרת המרכזית,

$$(4.1) \frac{\partial}{\partial x} \rightarrow \frac{1}{32} \begin{pmatrix} -3 & 0 & 3 \\ -10 & 0 & 10 \\ -3 & 0 & 3 \end{pmatrix} \text{ באופרטור גזירה חדש, } \frac{\partial}{\partial x} \rightarrow \frac{1}{2} \begin{pmatrix} -1 & 0 & 1 \end{pmatrix}$$

אופרטור זה מביא לשיפור היציבות הנומרית ומסייע לחישוב מדויק יותר של גרדיאנטים שכיוונם לא ניצב לצירים x ו-y.

כאשר משתמשים בנגזרת מרכזית במשוואת דיפוזיה, הדיפוזיה בין פיקסלים סמוכים נעשית רק בצורה אופקית או בצורה אנכית (כלומר רק 4 שכנים משפיעים על העידכון של כל פיקסל בכל איטרציה). אופרטור הנגזרת החדש מאפשר דיפוזיה ב-2 כיוונים אלכסוניים בנוסף לכיוון האנכי והאופקי (כלומר כעת 8 שכנים משפיעים על העידכון). לכן הוא נחשב אינווריאנטי לסיבובים.



בנוסף, כפי שנראה בהמשך, לצורך ביצוע דיפוזיה כיוונית חשוב לנו לשערך בצורה כמה שיותר מדויקת את כיווני הגרדיאנטים, ועל כן ברורה חשיבות האופרטור החדש.

### הגדרת הטנזור המבני

לצורך דיפוזיה כיוונית, אנו מעוניינים לעבוד עם אוריינטציות (זוויות בין 0 ל-180) במקום כיוונים (זוויות בין 0 ל-360). מעתה ואילך כאשר נשתמש במונח "כיוון" נתכוון בעצם ל"אוריינטציה". לכן במקום לעבוד עם וקטור גרדיאנט



עבור כל פיקסל נעבוד עם טנזור שמתקבל ממכפלה טנזורית של וקטור זה בעצמו:

$$\nabla I \otimes \nabla I = \nabla I \nabla I^T = \begin{pmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{pmatrix} \nabla I = \left( \frac{\partial I}{\partial x}, \frac{\partial I}{\partial y} \right) \rightarrow$$

כלומר, כעת כל פיקסל מאופיין על ידי 4 ערכים (4 איברי המטריצה הנ"ל). במאמר נטען שכדי למנוע רגישות לרעש עדיף לעבוד עם גירסא מוחלקת של התמונה  $I_\sigma$  במקום התמונה המקורית  $I$ .

הטנזור המבני הממוצע (להלן "הטנזור המבני") מתקבל על ידי קונבולוציה של התמונה המתאימה לכל אחד מאיברי המטריצה עם פילטר מיצוע גאוסי ברוחב  $\rho \gg \sigma$  כך ש:

$$(4.2) J_\rho = K_\rho * (\nabla I_\sigma \nabla I_\sigma^T)$$

הטנזור המבני משקף את האופי של סביבה ברדיוס  $\rho$  סביב הפיקסל. **הגדרת מושג הקוהרנטיות**

נסמן את הערכים העצמיים של הטנזור המבני ב-  $\mu_{1,2}$ , כאשר  $\mu_1 > \mu_2$ , ואת הוקטורים העצמיים האורתונורמליים המתאימים להם ב-  $v_{1,2}$ .

נגדיר את כיוון הוקטור  $v_2$ , המתאים לערך העצמי הקטן, להיות כיוון הקוהרנטיות. זהו הכיוון שבו יש את הפלקטואציות הקטנות ביותר בתוך הסביבה, כלומר כיוון הזרימה הדומיננטי. הוקטור  $v_1$  אורתוגונאלי לוקטור  $v_2$  ובמקרה שבו הפיקסל נמצא באיזור שפה הוא מכוון בכיוון שניצב לשפה. עוצמת הקוהרנטיות מוגדרת על ידי:

$$(4.3) \text{coherence magnitude} = (\mu_1 - \mu_2)^2$$

קוהרנטיות גבוהה משמעה זרימה חזקה בכיוון דומיננטי בסביבה של הפיקסל. כדי לשערך בצורה אמינה את כיוון זרימה זה השתמשנו באופרטור הנגזרת החדש בעת חישוב הטנזור המבני. למעשה נרצה לבחור את גודל הסביבה  $\rho$  בטנזור המבני להיות מתאים לעובי של קו אופייני בתמונה.

### הכללה לצבע

לצורך עבודה עם תמונת צבע נגדיר את הטנזור המבני הממוצע להיות:

$$\begin{pmatrix} 1 + \tilde{R}_x^2 + \tilde{G}_x^2 + \tilde{B}_x^2 & \tilde{R}_x \tilde{R}_y + \tilde{G}_x \tilde{G}_y + \tilde{B}_x \tilde{B}_y \\ \tilde{R}_x \tilde{R}_y + \tilde{G}_x \tilde{G}_y + \tilde{B}_x \tilde{B}_y & 1 + \tilde{R}_y^2 + \tilde{G}_y^2 + \tilde{B}_y^2 \end{pmatrix}$$

באופן זה כיווני הזרימה בתמונת צבע יקבעו לפי מטריקה שמשלבת את כל הערוצים, מה שימנע יצירת עיוותים בתמונה. עיוותים כאלה עלולים להיווצר אם חישוב הטנזור המבני היה מתבצע בכל ערוץ בנפרד. למעשה, אנו יוצרים בצורה זו סנכרון בעבודה על הערוצים השונים.



**שילוב הטנזור המבני והקוהרנטיות בתהליך הדיפוזיה**

נרצה עתה לנצל את הכלים החדשים שהוצגו בפרק הקודם ליצירת דיפוזיה כיוונית מגבירת קוהרנטיות. ראשית, ניתן להציג את תהליך הדיפוזיה הלא-לינארית שהוצג בחלק הראשון של הספר באופן הבא:

$$I_t = \nabla \cdot (c(x, y, t) \nabla I) = \nabla \cdot \begin{pmatrix} g(|I_x|) & 0 \\ 0 & g(|I_y|) \end{pmatrix} \begin{pmatrix} I_x \\ I_y \end{pmatrix}$$

כלומר, למעשה וקטור הגרדיאנט מוכפל משמאל במטריצה אלכסונית שבאלכסונה הראשי מופיעים מקדמי הדיפוזיה בכיוון האנכי ובכיוון האופקי, כאשר אותה פונקציה משמשת לחישוב שני המקדמים. מטריצה זו היא שקובעת את אופי הדיפוזיה הלא לינארית בכל פיקסל. למעשה ניתן ליצור סוגים חדשים של דיפוזיה על ידי שימוש בכל ארבעת איבריה.

לכן, נציג את משוואת הדיפוזיה באופן הבא:

$$(4.4) \quad I_t = \nabla \cdot \left( D(x, y, t) \begin{pmatrix} I_x \\ I_y \end{pmatrix} \right)$$

כאשר D תהיה מטריצת מקדמי הדיפוזיה.

$$(4.5) \quad D = \begin{bmatrix} v_1 & v_2 \\ \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$$

מוצע לקבוע את המטריצה D באופן הבא:

כאשר  $v_{1,2}$  הם הוקטורים העצמיים של הטנזור המבני ו- $\lambda_{1,2}$  ישלטו על הדיפוזיה בכל אחד מהכיוונים של הוקטורים הנ"ל (ערכיהם ינועו בין 0 ל-1 באופן דומה למקדמי הדיפוזיה שהוצגו בחלק הראשון של הספר). כלומר  $\lambda_2$  ישלוט בדיפוזיה בכיוון הקוהרנטיות ו- $\lambda_1$  ישלוט בדיפוזיה בכיוון הניצב לו (במקרה של פיקסל הנמצא באיזור שפה, זהו הכיוון הניצב לשפה). ניתן להגדיר פונקציות שונות ל- $\lambda_1$  ול- $\lambda_2$  כדי לקבל התנה גות שונה של הדיפוזיה בכיוון הקוהרנטיות ובכיוון הניצב לו. פונקציות אלה יקבעו לפי המשימה שאנו מעוניינים לבצע.

כדי להבהיר נקודה זו נציג את הבעיה באופן יותר מתמטי:

כל וקטור גרדיאנט ניתן להצגה בתור  $\nabla I = \omega_1 v_1 + \omega_2 v_2$  מכיוון ש- $v_{1,2}$  פורשים את המרחב הוקטורי  $R^2$ . הם ההיטלים של וקטור הגרדיאנט על כל אחד מהכיוונים. נגדיר את השטף להיות  $D \nabla I = \lambda_1 \omega_1 v_1 + \lambda_2 \omega_2 v_2$ . אזי אם  $\vec{j} = \begin{pmatrix} j_1 \\ j_2 \end{pmatrix}$



וקטור הגרדיאנט מכוון בכיוון שקרוב לכיוון הקוהרנטיות (כלומר  $\omega_2 \gg \omega_1$ ) נקבל שהשטף נשלט בעיקר על ידי  $\lambda_2$ . באופן דומה אם וקטור הגרדיאנט מכוון בכיוון שניצב לכיוון הקוהרנטיות השטף נשלט בעיקר על ידי  $\lambda_1$ .

סכימת העבודה של תהליך ה-CeD היא כדלקמן:

- (1) Calculate  $\nabla I, \nabla I_\sigma$
- (2)  $J_\rho = K_\rho * (\nabla I_\sigma \nabla I_\sigma^T)$
- (3) Calculate eigen vectors of  $J_\rho - v_{1,2}$
- (4) Calculate  $\lambda_{1,2}$  according to the chosen functions
- (5) 
$$D = [v_1 \ v_2] \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$
- (6) 
$$\vec{j} = \begin{pmatrix} j_1 \\ j_2 \end{pmatrix} = D \nabla I = \begin{pmatrix} aI_x + bI_y \\ bI_x + cI_y \end{pmatrix}$$
- (7) 
$$I_t = \nabla \cdot (D(x, y, t) \nabla I) = \frac{\partial j_1}{\partial x} + \frac{\partial j_2}{\partial y}$$

לגבי תמונות צבע, נשים לב כי השטף מחושב עבור כל ערוץ בנפרד, תוך הסתמכות על הטנזור המבני שחושב משילוב בין כל הערוצים. על מנת להתרשם באופן איכותי מתהליך הדיפוזיה הכיוונית שבנינו ניתן לבחון כשלב ראשון שתי הצעות "פשוטות" לערכים של  $\lambda_{1,2}$ :

1. הגברת הזרימה בתמונה, תוך שימור שפות: להגברת הזרימה ניקח  $\lambda_2 \equiv 1$ , כלומר נבצע החלקה מקסימלית בכיוון הקוהרנטיות. לצורך שימור שפות, נאפשר רק החלקה מינימלית בכיוון זה, כלומר ניקח  $\lambda_1 \equiv \alpha$ , כאשר  $\alpha \ll 1$ .

$$(4.6) \lambda_1 \equiv \alpha, \lambda_2 \equiv 1$$

2. הגברת הזרימה בתמונה, תוך הדגשת שפות: להגברת זרימה ניקח  $\lambda_2 \equiv 1$ , כמו בהצעה הראשונה. כדי להדגיש שפות ולא רק לשמרן ניקח  $\lambda_1 \equiv -\beta$ .

$$(4.7) \lambda_1 \equiv -\beta, \lambda_2 \equiv 1$$

בשתי ההצעות הנ"ל נעשה שימוש בערכים גלובליים עבור  $\lambda_{1,2}$ , כלומר אין תלות בעוצמת הקוהרנטיות או בעוצמת הגרדיאנט. בניסוי נבחן מספר תרחישים שבהם ההצעה הראשונה של דיפוזיה כיוונית "גלובלית" יכולה להיות שימושית:

1. השלמת צורות גיאומטריות ובפרט קווים מקוטעים.

2. הפיכת תמונה לדמוית ציור (משיכות מכחול) – "אומנות בעזרת מחשב".



3. שיפור ביצועים במשימה של ניקוי מרעש עבור תמונה שיש בה כיווניות גבוהה – נראה דוגמא של ניקוי מרעש עבור סינוס קוצנטרי.

השליטה על הדיפוזיה הכיוונית נעשה דרך הפרמטרים  $\lambda_{1,2}$ . כמובן שפרמטרים אלו לא חייבים להיות גלובליים וניתן לקבוע אותם להיות פונקציה כלשהי, ובפרט פונקציה של עוצמת הגרדיאנט או של עוצמת הקוהרנטיות. יש להתאים את הפונקציות הנבחרות לכל אחת מהפרמטרים למשימה שאותה רוצים לבצע בכיוון המתאים. הצעה לפונקציות כאלה תינתן בפרק הבא.

### 5. סופר-רזולוציה מתמונה בודדת (Single Image Super-Resolution)

בפרק זה נרצה להשתמש בכלים שהוצגו בפרקים הקודמים לשם בניית סכימות לביצוע משימה של סופר-רזולוציה מתמונה בודדת.

ראשית, נבחן את הפגמים שמאפיינים תמונה שעברה אינטרפולציה מרזולוציה נמוכה. לשם כך נתבונן בתמונה בגודל 150X113 שעברה הגדלה פי 3 בכל ציר:



הפגמים הבולטים הם פיקסליזציה, טשטוש, חוסר זרימה ורעשים "טבעתיים" מסביב לשפות (Ringing effects).

נבחן הצעות שונות לסופר-רזולוציה:

1. הפעלת דיפוזיה לא-לינארית שאינה כיוונית, שעושה שימוש ברכיב אחורי בלבד של





## שיטת FAB.

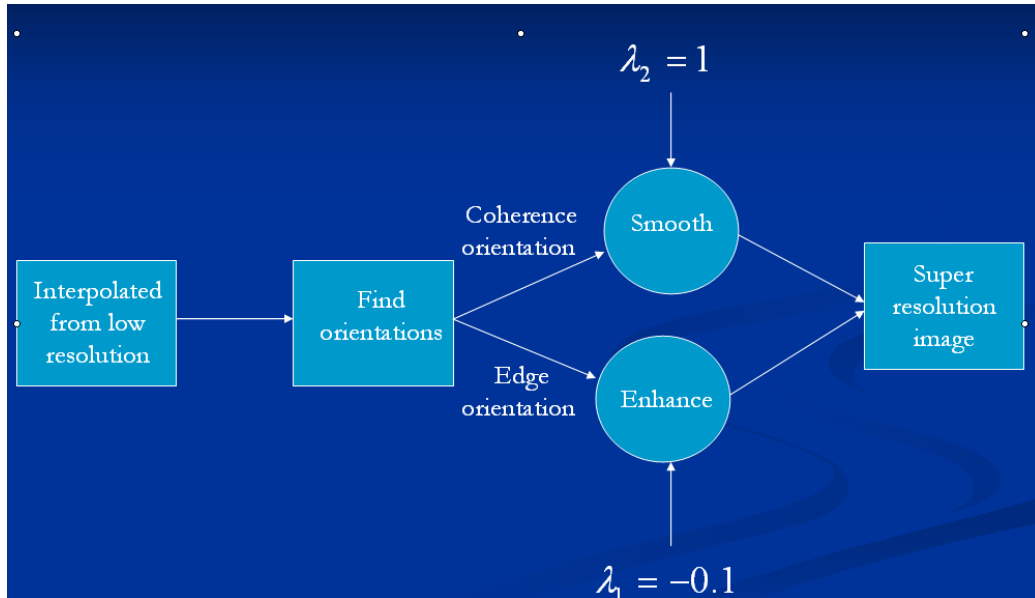
הצעה זו יכולה לתת טיפול חלקי בלבד לבעיות שהוצגו לעיל. השפות אומנם יודגשו אך ישארו "מפוקסלות". ניתן לנסות לטפל ברעשים ה"טבעתיים" על ידי הוספת רכיב קדמי קטן ב-FAB, אך נשלם בפגיעה במרקמים.

2. סכימה לשיפור תמונה שמבוססת על דיפוזיה כיוונית. בניסוי נבחן תחילה סכימה גלובלית, שמבוססת על ההצעה הגלובלית השניה מהפרק הקודם. לאחר שנאבחן את הבעיות בהצעה זו, נעבור מפרמטרים גלובליים לפרמטרים שהם תלויי עוצמת גרדיאנט ועוצמת קוהרנטיות, שאמורים לרסן ולעדן את הדיפוזיה הכיוונית, כך שתיווצר תמונה עם תחושה פחות "מלאכותית". ניתן לצפות שדיפוזיה כיוונית מגבירת שפות וקוהרנטיות תדע להתמודד עם הבעיות שפורטו לעיל. היא תחזיר את החדות והזרימה בתמונה כך שהשפות יהיו מודגשות יותר וקווים יהיו רציפים יותר. כמו כן, היא תחליק את האיזורים מסביב לשפות בצורה כיוונית להורדת הרעשים "הטבעתיים".

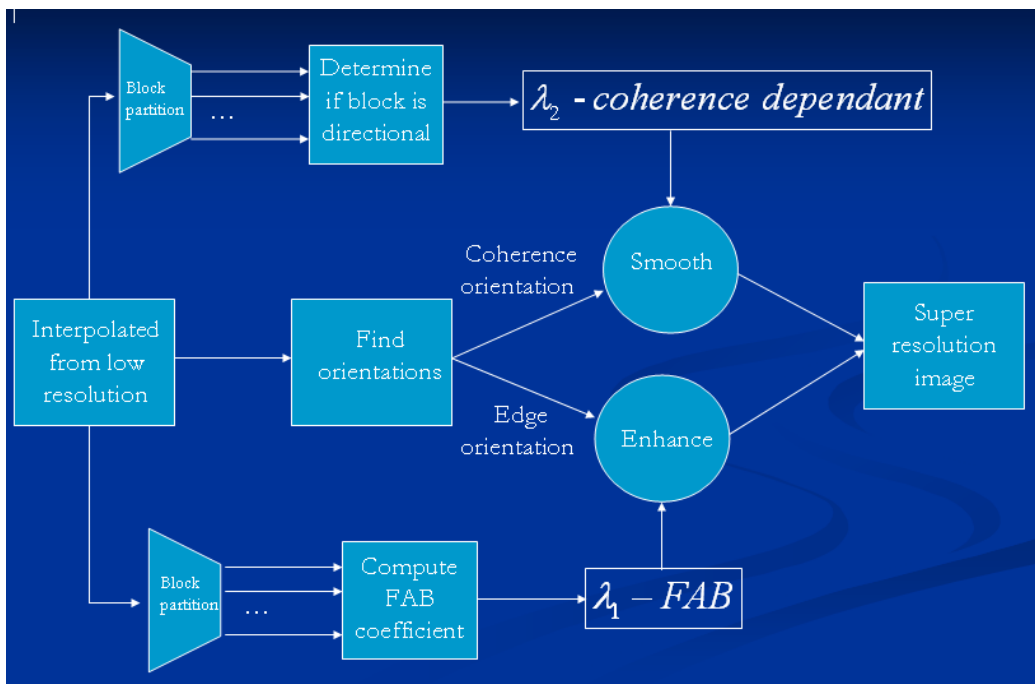


סכימות לשיפור תמונה שמבוססות על דיפוזיה כיוונית

א. סכימה גלובלית



ב. סכימה אדפטיבית

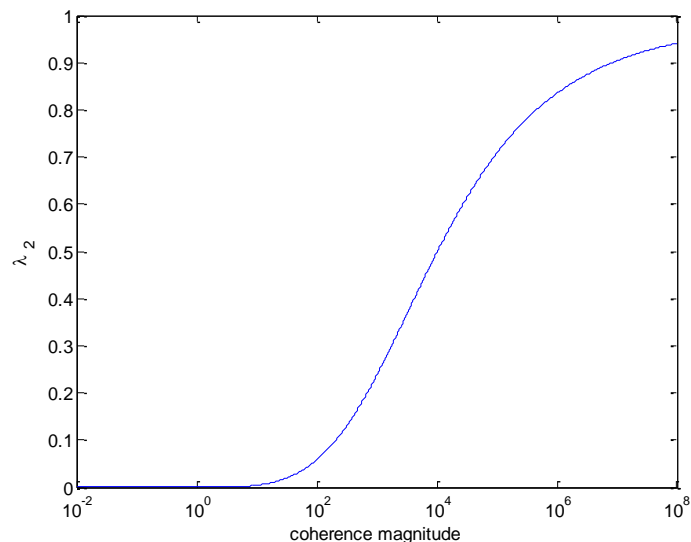


בסכימה זו לוקחים את  $\lambda_1$  להיות תלוי עוצמת גרדיאנט, כאשר התלות היא כמו במקדם FAB עם רכיב אחורי בלבד (מקדמי ה-FAB מחושבים לכל בלוק כפי שהוסבר בפרק הראשון), ואילו את  $\lambda_2$  לוקחים להיות פונקציה מונוטונית עולה



של עוצמת הקוהרנטיות . בנוסף, בבלוקים שהוחלט לגביהם שהם אינם כיווניים, לא מתבצעת החלקה בכיוון הקוהרנטיות ( $\lambda_2$  מאופס). ראוי לציין כי איסוף סטטיסטיקה על תמונות מראה שערכי הקוהרנטיות נעים על פני סדרי גודל רבים . תחום אופייני של ערכים הוא:  $[10^{-4}, 10^6]$ . אנו מעוניינים לקבל ערך גבוה (קרוב ל-1) של  $\lambda_2$  עבור ערכי קוהרנטיות גבוהים וערך נמוך (קרוב ל-0) עבור ערכי קוהרנטיות נמוכים, אבל נעדיף להימנע מהתנהגות "מדרגתית" ולעבוד עם פונקציה שנפרשת בצורה טובה על פני טווח ערכי הקוהרנטיות. לשם כך ניתן למשל ליצור תלות של  $\lambda_2$  בלוגריתם של עוצמת הקוהרנטיות ולא בעוצמת הקוהרנטיות עצמה.

צורה או פיינית של פונקציה שתלויה בעוצמת הקוהרנטיות ושמקיימת את הדרישה הנ"ל היא:



### ביבליוגרפיה

1. P. Perona, J. Malik, "Scale-space and edge detection using anisotropic diffusion", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, No. 7, pp. 629 - 639, July 1990.
2. F. Catte, P. L. Lions, J. M. Morel and T. Coll, "Image Selective Smoothing and Edge Detetction by Nonlinear Diffusion", SIAM J. Numer. Anal., Vol. 29, No. 1, pp. 182-193, Feb. 1992.
3. G. Gilboa, N. Sochen, Y. Y. Zeevi, "Forward-and-Backward Diffusion Processes for Adaptive Image Enhancement and Denoising", IEEE Transactions on Image Processing, Vol. 11, No. 7, pp. 689-703, July 2002
4. J. Weickert, B. M. ter Haar Romeny, M. A. Viergever, "Efficient and Reliable Schemes for Nonlinear Diffusion Filtering", IEEE Transactions on Image Processing, Vol. 7, No. 3, pp. 398-410, March 1998.



5. J. Weickert, "Coherence-Enhancing Diffusion of Colour Images", Proc. VII National Symposium on Pattern Recognition and Image Analysis (VII NSPRIA), Vol. 1, pp. 239-244, April 1997.

6. J. Weickert, H. Schar, "A scheme for Coherence-Enhancing Diffusion Filtering with Optimized Rotation Invariance", Journal of Visual Communication and Image Representation, Vol. 13, No. 1, pp. 103-118, March 2002.

---



### הערות טכניות לקראת הניסוי:

1. כל פונקציות ה-Matlab והתמונות המוזכרות בשאלות ההכנה ובמהלך הניסוי ניתנות להורדה מהאתר.
2. בשאלות ההכנה נרצה לעבוד לשם פשטות עם תמונות רמות אפור. לכן לאחר טעינת התמונה בעזרת הפקודה `imread` יש להמיר את התמונה לרמות אפור על ידי הפקודה `rgb2gray`.
3. במהלך הניסוי נעבוד עם תמונות צבע ולכן שם אין לבצע המרה לרמות אפור. ראוי לציין כי ברקע התיאורטי לא הוסברה ההכללה לצבע עבור דיפוזיה לא-לינארית. בקוד ה-Matlab הקיים נעשה שימוש בהכללה טריוויאלית (למרות שקיימות גישות סדורות יותר לטיפול בצבע כמו עבודה במסגרת Beltrami Flow). עבור דיפוזיה כיוונית השתמשנו בהכללה לצבע שהופיעה בפרק 4 של התיאוריה.
4. לפני שמבצעים תהליך דיפוזיה על התמונה יש להמיר אותה לפורמט של `double` כדי שיהיה ניתן לבצע עליה פעולות כמו סינון.

### דוח הכנה למפגש הראשון

יש לענות על השאלות הבאות בצורה מפורטת ומנומקת. בשאלות שבהן נדרשת כתיבת קוד Matlab, יש לצרף לדוח ההכנה את קוד ה-Matlab.

מומלץ להשתמש בפונקציות העזר הבאות ב-Matlab:

**calc\_image\_derivatives** – מקבלת תמונה ואת סוג המסנן שבו משתמשים לצורך חישוב הנגזרות, ומחשבת את הנגזרות בכיוון האופקי והאנכי לכל פיקסל.

**compute\_structural\_tensor** – מקבלת את הנגזרות האופקית והאנכית של התמונה ואת הפרמטר  $\rho$  ומחשבת את הטנזור המבני. הטנזור המבני מאורגן ב-3 מטריצות, כל אחת בגודל של התמונה. לא צריך מטריצה רביעית כמו שהופיע ברקע התיאורטי כי האיברים באלכסון המשני של הטנזור המבני תמיד שווים זה לזה.

**compute\_coherence\_and\_orientations** – מקבלת את הטנזור המבני ומחשבת את עוצמת הקוהרנטיות וכן את כיווני הקוהרנטיות והשפות בכל פיקסל.

### שאלות הכנה



1. כפי שראינו במשוואה 1.1 הפעלת דיפוזיה לינארית על תמונה שקולה לקונבולוציה של התמונה עם גאוסיאן . רוחב הגאוסיאן גדל ככל שתהליך הדיפוזיה נמשך זמן רב יותר – הוא שווה  $2ct$  . כדי לקבל אנלוגיה לתהליך הדיפוזיה הדיסקרטי נרשום את זמן התהליך באופן הבא :  $t = N \cdot dt$  , כאשר  $N$  מספר הצעדים ו- $dt$  גודל צעד הזמן. נקח  $c = 1$  ונקבל שרוחב הגאוסיאן שווה  $2N \cdot dt$  .

עבור התמונה house.jpg :

- בדוק מה קורה כתוצאה מהפעלת דיפוזיה עבור  $N=10,100,500$
- ודא עבור כל אחד מהמקרים שהערך הממוצע של התמונה לא משתנה במידה משמעותית בעקבות תהליך הדיפוזיה.
- מהן מסקנותיך לגבי דיפוזיה לינארית?

הערה - ניתן לצורך כך להשתמש בפקודות fspecial עם האופציה 'gaussian' ולאחריה imfilter עם האופציה 'replicate'. יש לקחת את סטיית התקן של הפילטר הגאוסיה להיות  $\sqrt{2N \cdot dt}$  ואת גודל הפילטר להיות  $2N + 1$  (כאשר מספר זה גדול מגודל התמונה, יש להגביל את הפילטר להיות בגודל התמונה).

2. חשוב לשים לב כי במשוואה 1.4 מוכפל מקדם הדיפוזיה בגרדיאנט שמחושב מתוך התמונה הלא מוחלקת. תן הסבר להגיון שעומד מאחורי העבודה בשיטה זו.

רמז : ההצעה הנאיבית יותר היתה פשוט להחליף את התמונה בגרסה המוחלקת שלה וכל להיפטר מהרעש אחת ולתמיד . מה החסרון של הצעה כזו ואיך מנסים להתגבר עליו בשיטה שהוצעה?

3. עבור התמונה house.jpg :

א. בחר שלושה חלונות – אחד שנמצא באזור השמים, שני שנמצא באזור הבנין ושלישי שנמצא באזור הדשא . לכל אחד מהחלונות צייר הסטוגרמה מנורמלת והיסטוגרמה מנורמלת מצטברת (הנרמול הוא לערכים בין 0 ל-1) של עוצמת הגרדיאנט.

ב. חשב לכל חלון את ה-MAG. סמן על ההיסטוגרמה המנורמלת את ערכי  $K_f, K_b, W$  לפי ההצעה שמופיעה במשוואה 1.6. מה מסקנתך לגבי הצעה זו?

הערה – לצורך חישוב עוצמת הגרדיאנט ניתן להשתמש בפונקציה `calc_image_derivatives` עם הפרמטר `filter_type='central_derivative'`.

4. הוכח שלשני הביטויים במשוואות 2.5 ו-2.6 יש אותו פיתוח טיילור מסדר ראשון.

רמז : קירוב טיילור מסדר ראשון של  $\frac{1}{1-x}$  הוא  $1+x$ .



5. תן דוגמא פשוטה לבלוק של  $3 \times 3$  פיקסלים שעבורו נקבל את המצב הבא:  
 א. צעד דיפוזיה אחד שבו נעשה שימוש באופרטור הנגזרת המרכזית ישאיר את הפיקסל המרכזי ללא שינוי.  
 ב. צעד דיפוזיה אחד שבו נעשה שימוש באופרטור הנגזרת החדש (משוואה 4.1) ישנה את הפיקסל המרכזי.

6. עבור התמונה fingerprint.jpg :

א. חשב לכל פיקסל את הטנזור המבני.

ב. חשב את עוצמת הקוהרנטיות בכל פיקסל ואת הערך הממוצע של עוצמת הקוהרנטיות בתמונה.

ג. שרטט את כיוון הקוהרנטיות (נוח להשתמש בפקודה quiver) עבור הפיקסלים שעוצמת הקוהרנטיות בהם גדולה מעוצמת הקוהרנטיות הממוצעת.

**הערות:** 1. בשאלה זו יש להשתמש בפונקציות `compute_structural_tensor`

ו-`compute_coherence_and_orientations`.

2. כדי שיהיה נוח להתבונן בחצים המכוונים בכיוון הקוהרנטיות, מומלץ לשרטט חצים בסקלה פחות צפופה (נדגום כל פיקסל רביעי בכל אחד מהצירים).



## מהלך הניסוי

### ניסוי 1 - ניקוי מרעש על ידי דיפוזיה לא-לינארית

בחלק זה יש להשתמש בפונקציה ה-matlab הבאה: `non_linear_diffusion`

א. טען את התמונה `Paignton_Shops.jpg` והוסף לה רעש גאוסני לבן עם שונות 0.005. ניתן להשתמש לשם כך בפקודה הבאה:  
`I=255*imnoise(I/255,'gaussian',0,sigma_noise);`

ב. הפעל תהליך דיפוזיה לא-לינארית עם מקדם דיפוזיה PM ועם הפרמטרים הבאים: גודל צעד זמן – 0.125  
 מספר צעדי זמן מקסימלי – 100  
 $\beta = 0.1$ ,  $K=10$ ,  $\sigma^2 = 0.5$

אל תשלח את הפרמטרים `optimal`, `PSNR_thr`. הם יקבלו את ערכי ה-default שלהם, כך שיתבצעו 100 איטרציות.

הערה – בניסוי 1 אל תשלח את הפרמטר `is_SR`. הוא יקבל את ערך ה-default שלו שהוא אפס. זהו הערך הרצוי כאן כי אנו לא מבצעים משימה של סופר-רזולוציה.

ג. הצג את התמונה המתקבלת לאחר 100 איטרציות. מהי מסקנתך לגבי המשימה שמבצע תהליך דיפוזיה עם מקדם PM לאחר זמן רב?

ד. שרטט גרף של ה-PSNR כפונקציה של הזמן עבור תהליך הדיפוזיה עם מקדם PM. לאחר כמה איטרציות ה-PSNR הגיע לערכו המקסימלי? זוהי הנקודה שבה יעצר התהליך עבור האופציה `optimal=1`.

ה. חזור על הסעיף הקודם עבור תהליך דיפוזיה עם מקדם FAB (אפשר להסתפק ב-50 צעדי זמן). כעת שיש בידך את הגרפים של PSNR כפונקציה של הזמן עבור שתי ההצעות, קבע ערך סף ל-PSNR שבו נרצה לעצור את שני התהליכים (העצירה צריכה להעשות בשלב של ירידת ה-PSNR ויש לוודא ששני התהליכים יכולים להגיע בירידה לערך זה). ערך סף זה ישמש אותנו כ-`PSNR_thr` עבור האופציה `optimal=2`.

ו. חשב את התמונות המתקבלות בכל אחת משתי ההצעות (FAB, PM) כאשר התהליך נעצר ב-PSNR המקסימלי (`optimal=1`). השווה בין ערכי ה-PSNR המקסימליים בשתי ההצעות. בנוסף, הצג את התמונות וערוך השוואה ויזואלית איכותית בין התמונות.





ז. כדי לעשות השוואה ויזואלית יותר "הוגנת" בין התמונות, חשב את התמונות המתקבלות בכל אחת משתי ההצעות כאשר עוצרים את התהליך באותו ה-PSNR, לפי ערך הסף שקבעת בסעיף ה' (optimal=2). הצג את התמונות וענה על השאלה הבאה – מה מעדיפה כל אחת מההצעות?

ח. נרצה להתרשם מחשיבות בחירת גודל צעד הזמן. עד עכשיו עבדנו עם גודל צעד זמן  $\tau = 0.125$ , שעומד בדרישה התיאורטית  $\tau \leq 0.25$ . בדוק מה קורה בכל אחת משתי ההצעות (PM, FAB) כאשר לוקחים גודל צעד זמן מעבר לערך הסף – למשל  $\tau = 0.3$ . תן דעתך גם לירידה בערך ה-PSNR וגם לפגיעה הויזואלית.

## ניסוי 2 – דיפוזיה כיוונית

בחלק זה יש להשתמש בפונקציה ה-matlab הבאה:

coherence\_enhancing\_diffusion

א. טען את התמונה fingerprint.jpg והפעל עליה תהליך של דיפוזיה כיוונית מגבירת קוהרנטיות עם הפרמטרים הבאים:

גודל צעד זמן – 0.125

מספר צעדי זמן – 72

$$\beta = 0.1, \rho = 6, \alpha = 0.001, \nu^2 = 0.5$$

הפונקציה עבור  $\lambda_{1,2}$  לפי משוואה 4.6 (יש לשלוח 'global' עבור כל אחת מהן).

סוג הפילטר עבור חישוב הנגזרת – אינווריאנטי לסיבובים ('rotation\_invariant').

הערה – אל תשלח את שאר הפרמטרים, כך שהם יקבלו את ערכי ה-default שלהם.

הצג את התמונה המקורית ואת התמונה לאחר תהליך הדיפוזיה הכיוונית איזו משימה מבצעת כאן הדיפוזיה הכיוונית?

ב. טען את התמונה tree.jpg. זוהי דוגמא אופיינית לתמונה "טבעית". הפעל עליה תהליך של דיפוזיה כיוונית עם אותם פרמטרים כמו בסעיף הקודם. בנוסף, שלח את הפרמטר draw\_coh\_directions=1, כך שבתחילת התהליך ובסופו יוצגו כיווני הקוהרנטיות בתמונה.

מהתבוננות ב"תמונות החצים", מה מסקנתך לגבי האופן שבו הדיפוזיה הכיוונית פועלת על הוקטורים המצביעים בכיוון הקוהרנטיות?



הצג את התמונה המקורית ואת התמונה לאחר תהליך הדיפוזיה הכיוונית. מה מסקנתך לגבי האופן שבו הדיפוזיה הכיוונית פועלת על תמונה טבעית?

ג. השתמש בקוד הבא כדי ליצור תמונה של סינוס קוצנטרי:

```
[x,y]=meshgrid(1:0.5:401,1:0.5:401);
R=(x-201).^2+(y-201).^2;
I=255*abs(sin(R/100));
I=I(401:601,201:401);
```

הוסף לסינוס הקוצנטרי רעש גאוסני לבן עם שונות 0.005. נסה לנקות את הרעש ב-3 אופנים שונים:

(1) הפעלת תהליך דיפוזיה לא-לינארית עם מקדם דיפוזיה PM עם הפרמטרים שהופיעו בניסוי 1 – סעיף ב' ועם  $\text{optimal}=1$  (עצירת התהליך עם הגעה לערך PSNR מקסימלי).

(2,3) הפעלת תהליך דיפוזיה כיוונית עם הפרמטרים שהופיעו בסעיף א' (פרט לשינוי אחד -  $\rho = 4$ ), פעם אחת עם פילטר של נגזרת מרכזית ('central\_derivative') ופעם שניה עם הפילטר האינוריאנטי לסיבובים ('rotation invariant'). בשני המקרים עצור את התהליך עם הגעה ל-PSNR מקסימלי ( $\text{optimal}=1$ ).

הצג את התמונות המתקבלות בכל אחד מ-3 השיטות והשווה ביניהן, הן מבחינת ה-PSNR והן ויזואלית. מה מסקנתך לגבי שימוש בדיפוזיה כיוונית במשימה של ניקוי מרעש? מתי דיפוזיה כיוונית יכולה להשיג ביצועים טובים יותר מדיפוזיה לא-לינארית "רגילה" (שאינה כיוונית)? נסה להסביר את ההבדל שהתקבל בין תהליכים (2), (3) כשעבדנו עם פילטרים שונים לחישוב הנגזרות.

ד. טען את התמונה wall.jpg הפעל עליה תהליך של דיפוזיה כיוונית ב-4 אופנים שונים. כל תהליכי הדיפוזיה ימשכו 48 צעדי זמן והם יבדלו זה מזה בפרמטר  $\rho$  ובפונקציות שיבחרו עבור  $\lambda_{1,2}$ :

בשני התהליכים הראשונים הפונקציות עבור  $\lambda_{1,2}$  יהיו לפי משוואה 4.6 ('global' עבור שתיהן) ואילו הפרמטר  $\rho$  יהיה פעם אחת 2 ופעם שניה 10. בשני התהליכים האחרונים הפונקציות עבור  $\lambda_{1,2}$  יהיו לפי משוואה 4.7 ('global' עבור  $\lambda_2$  ו-'negative' עבור  $\lambda_1$ ) ואילו הפרמטר  $\rho$  יהיה פעם אחת 2 ופעם שניה 10. הערה – שאר הפרמטרים כמו בסעיף א'.

הצג את התמונות שהתקבלו בכל אחד מ-4 התהליכים ורשום את מסקנותיך לגבי



ההבדל בין שתי ההצעות שהופיעו בסוף פרק 4 (משוואות 4.6-4.7) ולגבי המשמעות של הפרמטר  $\rho$ .

### ניסוי 3 – סופר-רזולוציה מתמונה בודדת

בחלק זה אין להציג את התוצאות בגרפים, אלא לשמור את התמונות בפורמט '.jpg'

לשם כך יש להשתמש בפקודה `imwrite`. חשוב לשים לב שלפני השמירה צריך להמיר בחזרה מפורמט `double` ל-`uint8`.

א. טען את התמונה `wizard.jpg`. ראשית, ניצור בצורה מלאכותית תמונה ברזולוציה נמוכה (מספר פיקסלים קטן). הקטן את התמונה פי 4 בכל ציר בעזרת הפקודה `imresize`. השתמש באפשרות '`bicubic`' בתור שיטת האינטרפולציה.

אנחנו מעוניינים לבצע סופר-רזולוציה לתמונה בעזרת מספר הצעות שנסקרו בעמודים 15-17. לשם כך, נבצע קודם כל אינטרפולציה סטנדרטית לתמונה לרזולוציה הרצויה.

הגדל את התמונה פי 3 בכל ציר בעזרת הפקודה `imresize`. השתמש באפשרות '`bicubic`' בתור שיטת האינטרפולציה. שמור תמונה זו בשם `wizard_LQ.jpg`.

הצג את התמונה שהתקבלה לאחר האינטרפולציה. באילו תופעות ניתן להבחין בה? מה ההשפעות השליליות של האינטרפולציה ומדוע יש צורך בשיפור התמונה?

ב. בסעיף זה נבחן את ההצעה לבצע סופר-רזולוציה בעזרת דיפוזיה לא-לינארית שאינה כיוונית, ובפרט שימוש ברכיב אחורי בלבד של שיטת FAB. לשם כך יש להשתמש בפונקציה `non_linear_diffusion` עם הפרמטרים הבאים:

גודל צעד זמן – 0.125

מספר צעדי זמן מקסימלי – 48

$$\beta = 0.1, K=10, v^2 = 0.5$$

יש לשלוח לפונקציה גם `optimal=0` כדי שיתבצעו כל 48 האיטרציות ו-`is_SR=1` כי אנו במשימה של סופר-רזולוציה ורוצים להשתמש ברכיב אחורי בלבד של FAB.



שמור את התמונה שהתקבלה לאחר הסופר- רזולוציה בשיטה זו בשם wizard\_SR1.jpg. הצג את תמונה זו ותאר את התוצאות. אילו שיפורים שיטה זו הכניסה לתמונה? באילו בעיות השיטה לא הצליחה לטפל כראוי?

ג. לאור התוצאות של הסעיף הקודם, נעבור לבחון הצעה לבצע סופר-רזולוציה בעזרת דיפוזיה כיוונית. בסעיף זה נבחן את הסכימה הגלובלית שתוארה בעמוד 16 – א'. לשם כך יש להשתמש בפונקציה coherence\_enhancing\_diffusion עם הפרמטרים הבאים:

גודל צעד זמן – 0.125

מספר צעדי זמן – 48

$$\beta = 0.1, \rho = 6, \alpha = 0.001, \nu^2 = 0.5$$

הפונקציות עבור  $\lambda_{1,2}$  - לפי משוואה 4.7 ('global' עבור  $\lambda_2$  ו-'negative' עבור  $\lambda_1$ ).

– אינווריאנטי לסיבובים ('rotation\_invariant'). סוג הפילטר עבור חישוב הנגזרת

שמור את התמונה שהתקבלה לאחר הסופר- רזולוציה בשיטה זו בשם wizard\_SR2.jpg. הצג את תמונה זו ותאר את התוצאות. מה מסקנותיך לגבי הצעה זו?

ג. לאור התוצאות של הסעיף הקודם, נעבור מסכימה גלובלית לסכימה אדפטיבית. יש לעבוד באותו האופן כמו בסעיף הקודם, רק שהפעם יש להשתמש בפונקציות הבאות עבור  $\lambda_{1,2}$ : 'FAB-only\_backward' עבור  $\lambda_1$  ו-'log\_coh\_cond' עבור  $\lambda_2$ .

שמור את התמונה שהתקבלה לאחר הסופר- רזולוציה בשיטה זו בשם wizard\_SR3.jpg. הצג את תמונה זו ותאר את התוצאות. אילו שיפורים שיטה זו הכניסה לתמונה? השווה את שיטה זו לשתי השיטות הקודמות.

### כתיבת דוח מסכם

הדוח המסכם צריך לכלול תשובות לכל השאלות שהופיעו ב חלק "מהלך הניסוי". בדוח יש לכלול גם את כל הגרפים שנתבקשתם לשרטט.



לגבי התמונות שנתבקשתם להציג – כדי לחסוך הדפסה צבעונית שלהן , יש לשלוח באימייל מצגת שמכילה עבור כל סעיף את התמונות בלבד (ללא הסברים).

לגבי קוד המטלב – יש לשלוח באימייל קובץ zip שמכיל את כל פונקציות המטלב ששימשו אתכם בחלק "מהלך הניסוי".