

ירודות באיכותן) אשר גורמת לטשטוש H . אנו נניח כי טשטוש זה הוא ליניארי וקבוע במקום. לכן, גרעין הטשטוש $H[m,n]$ מספר את סיפורה המלא של פעולה זו.

ישנם מקרים בהם מקור הטשטוש אינה האופטיקה אלא תנועה של המצלמה. כאשר המצלמה זהה בזמן החשיפה, תתקבל מריחה כיוונית הקרויה Motion Blur. מתוך התמונה המקולקלת ניתן לרוב לחשב את גרעין הטשטוש אשר הינו ליניארי וקבוע במקום.

נחזור לסיפור על המצלמה: בנוסף, מטריצת הגלאים בה המצלמה עושה שימוש סובלת מרעשים, ולכן התמונה לאחר הטשטוש מקבלת מרכיב אדיטיבי $N[m,n]$ של רעש אקראי בעל ממוצע אפס. אנו נניח לכל אורכו של פרק זה כי רעש זה הינו רעש לבן והומוגני בעל פילוג גאوسي. פירוש הדבר שבכל פיקסל מצטרף מספר אקראי גאوسي בעל שונות σ_N קבוע במקום, וממוצע אפס.

לכן, הקשר בין התמונה הרצויה $x[m,n]$ לתמונה המתקבלת $Y[m,n]$ נתון ע"י הביטוי:

$$Y[m,n] = X[m,n] \otimes H[m,n] + N[m,n]$$

לשם נוחות מקובל לסדר את התמונות בסידור לקסיקוגרפי לפי עמודות (בחירה שרירותית), ואז הקשר הנ"ל הוא:

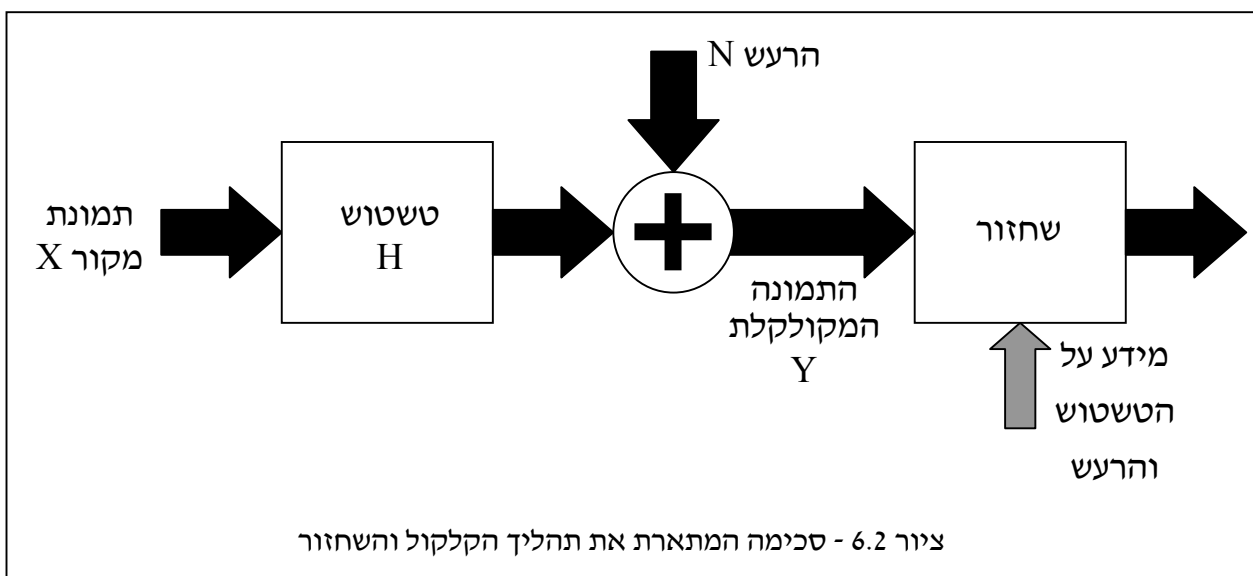
$$\underline{Y} = \underline{H}\underline{X} + \underline{N}$$

המטריצה H היא אופרטור הטשטוש בגודל $[L^2 \times L^2]$. בהמשך טבעי לדיון מפרק 2, נוכל לומר כי כיוון ש- H מייצגת אופרטור קבוע במקום, הרי שמטריצה זו בעלת מבנה בלוק טופליץ. אם מטריצה זו גדולה מאוד ביחס לגודל הגרעין המטשטש, נוכל אף להניח בקירוב טוב כי יש לה מבנה בלוק סיבובי. אנו נדקק להנחות אלו בהמשך.

ציור 6.1 מראה תמונת מקור X וגרסה מקולקלת שלה ע"י טשטוש אופייני לאופטיקה (טשטוש גאوسي בעל חתך רדיאלי), וגרסה של מריחה הנובעת מתנועה. בשני המקרים הוסף רעש אדיטיבי. כאמור, המטרה היא לקחת את התמונה המקולקלת, את המידע על גורמי הקלקול, ולשחזר את תמונת המקור. ציור 6.2 מראה סכימת בלוקים של התהליך במלואו (קלקול ותיקון). המידע המוזן לתהליך השחזור הוא התמונה המקולקלת עצמה, גרעין הטשטוש, ואופיו הסטטיסטי של הרעש (במקרה שלנו רק הפרמטר σ_N).

נשים לב כי לפי המודל הנ"ל יכולנו בקלות לטפל בטשטוש ליניארי תלוי מקום. גם אז אופרטור הטשטוש ניתן לתיאור כמטריצה H , אלא שאז אין לה מבנה בלוק טופליץ. הבעיה עם טשטוש תלוי מקום היא הצורך לדעת אותו, ולבנות ממנו את המטריצה H . אם בשל הבנה עמוקה של התהליך הפיסיקלי של יצירת התמונה טשטוש זה ידוע, הרי שרוב הדברים שייאמרו בהמשך רלוונטיים גם לטשטוש תלוי מקום.

ציור 6.1 - תמונת מקור וגרסאות מקולקלות שלה ע"י טשטוש גאוסי (למעלה בצד ימין) וטשטוש כתוצאה מתנועה (שמאל למטה). בשני המקרים הוסף רעש גאוסי לבן בעל שונות $\sigma_N = 5$



אבל, מתעוררות שתי בעיות: ראשית, כלל לא ברור כי ל- \mathbf{H} קיים היפוך. למעשה, במרבית המקרים ניתן לקבוע בודאות כי לא קיים היפוך! קל יותר לראות זאת דווקא בייצוג בתדר. אם \mathbf{H} לא הפיכה (כלומר סינגולרית), אזי גם \mathbf{H}_T סינגולרית. לכן, באיברי אלכסונה הראשי יש אפסים. מטריצה כזו לא ניתנת להיפוך! מבחינה אינטואיטיבית ניתן להסביר זאת באופן הבא: נניח כי קיים אפס כזה בשורה K כלשהי. אפס זה כפל את האיבר ה- K של התמרת התדר של תמונת הכניסה X . לכן, ערך זה אבד ולא ניתן יהיה לשחזרו.

תיקון לבעיה זו יהיה (ושוב, אנו "חושבים" בתדר) לקחת כל ערך אלכסון ראשי שהוא אפס, ולהמירו בערך אחר קרוב לאפס - נניח $1e-10$. גישה דומה אך מעט שונה היא להשתמש במסנן המשחזר הבא (הנתון בתדר):

$$\mathbf{G}_T[k] = \frac{\mathbf{H}_T^*[k]}{|\mathbf{H}_T[k]|^2 + \text{Const}}$$

כאשר הקבוע קטן משמעותית מ- 1 (סדר גודל מומלץ - 0.001). לכן, בתדרים נמוכים בהם $\mathbf{H}_T^*[k]$ בערכו המוחלט קרוב ל- 1, הקבוע זניח ואנו מקבלים התנהגות של מסנן הפוך. עבור תדרים גבוהים בהם $\mathbf{H}_T^*[k]$ מתקרב לאפס, הקבוע משמש כהגנה מחלוקה באפס.

עד כה בחרנו להתעלם מקיומו של הרעש. מסתבר כי אפילו אם הרעש חלש, השפעתו בתהליך של סינון הפוך היא דרמטית. נשתמש שוב בתיאור אינטואיטיבי של הבעיה ע"י ייצוג הבעיה בתדר: אם נניח כי בשורה ה- k של \mathbf{H}_T יש ערך נמוך יחסית (נניח $1e-3$). נניח כמו כן כי בנקודת תדר זו, עוצמת האות המקורי ועוצמת הרעש נמוכות מאוד אך הרעש חזק יותר מהאות. תופעה זו מתרחשת כמעט תמיד בתדרים גבוהים בהם כמעט אין נוכחות לאות המקור ואז יוצא שהרעש דומיננטי. נניח לכן - $X_T[k] = 1e-3$ ואילו הרעש חזק יותר - $N_T[k] = 1e-2$. הפעלת המסנן ההפוך פירושה הכפלת שני ערכים אלו בערך $1e3$. גרמנו באופן זה להכנסה חזקה מאוד של רעש לתמונת התוצאה.

המסנן המשחזר שהוצע קודם יתנהג יותר טוב. במקרה כזה, נקבל שהן הרעש והן האות מועברים כשהם מוכפלים בגודל:

$$\frac{1 \times 10^{-3}}{1 \times 10^{-6} + \text{Const}} \stackrel{\cong}{\uparrow} 1$$

Const=0.001

תוצאה זו לא נוראית כמו הכפלה ב- 1000, אך אינה טובה לכשעצמה. בנקודת תדר זו היה ראוי שנכפול את הכניסה בערך נמוך מ- 1 לשם דיכוי הרעש. ממילא האות טובל שם ברעש גדול ממנו, ולכן אין

מה להפסיד. נציע אם כך את הגישה הבאה: בכל נקודת תדר נברר מהו היחס בין עוצמת האות לעוצמת הרעש, ולבצע מסנן הפוך רק בתדרים בהם יחס זה גבוה. במקומות בעלי יחס נמוך, בהם הרעש דומיננטי, כתחליף למסנן הפוך, יש להפעיל דיכוי (כלומר הכפלה בערך נמוך מאוד מ-1). דיכוי זה יקטין את האות שם, אך גם את הרעש שממילא דומיננטי יותר.

הבעיה עם גישה זו היא הצורך בשני גדלים שאת שניהם אין לנו - עוצמת האות המקורי ועוצמת הרעש בכל נקודת תדר. כתחליף, ניתן להניח כי עוצמת הרעש קבועה וזהה בכל נקודת תדר, ושווה σ_N . הנחה זו נכונה כאשר הרעש לבן. באשר לאות המקורי, הנחה סבירה היא שעוצמתו יחסית הפוך למרחק מהראשית בתדר. כלומר, עבור תדרים נמוכים עוצמתו צפויה להיות גבוהה, ועם העלייה בתדר עוצמתו יורדת. ביצוע התמרות פוריה למגוון גדול של תמונות מגלה כי כלל זה נכון ברוב מוחלט של המקרים. ההבדל בין התמונות מתבטא בדרך כלל כמעט רק בפאזה, ומזו אנו מתעלמים פה.

לסיכום, התהליך המוצע הוא ביצוע התמרת פוריה לאופרטור הטשטוש, ולקבל את H_T . אם נסמן את ערכי האלכסון הראשי של מטריצה זו ב- $H_T[k]$, מסנן השחזור בתדר, G_T יתקבל ע"י:

$$G_T[k] = \frac{H_T^*[k]}{|H_T[k]|^2 + \alpha \sigma_N^2 (r[k])^2}$$

כאשר * מייצג צמוד קומפלקסי, ו- $r[k]$ הוא המרחק אל הראשית במישור התדר. כך למשל, עבור $k=1$ זהו איבר ה-DC ולכן $r[1]=0$. גישה זו דומה להצעתנו הקודמת, למעט העובדה שכעת ה-"קבוע" שלנו גדל עם התדר, ולכן הופך דומיננטי מהר יותר.

מדוע ביטוי זה? בשל היות המכנה גדול מאפס לכל k , ברור כי לא נקבל היפוך של 0, וכך נפתרה הבעיה הראשונה. שנית, עבור k נמוכים, כאשר האות חזק מאוד ביחס לרעש, הגודל $r[k]$ קטן מאוד, ונקבל כי המסנן הוא בקירוב טוב היפוך מדויק כיוון שהאיבר הימני במכנה יהיה זניח. כאשר הרעש דומיננטי מאוד, איבר זה (הימני במכנה) יהיה הדומיננטי במכנה ולכן המסנן המשחזר יהיה קרוב לאפס, כנדרש.

לאחר שהתקבל G_T (מטריצה אלכסונית) ניתן לבצע התמרת תדר הפוכה לקבלת G . הפעלת G על התמונה \underline{Y} תניב את השחזור. כתחליף, ניתן לבצע את השחזור בתדר ע"י הכפלת G_T ב- \underline{Y}_T וביצוע התמרת תדר הפוכה על התוצאה - שתי גישות אלו זהות לחלוטין.

המסנן המוצע מוכר מאוד וקרוי מסנן Wiener. בנייתו באופן מדויק וקבלת הביטוי שהוצע נעשית מתוך שיקולים הסתברותיים, שלא הצגנו כאן. ציור 6.3 מראה תמונת מקור מטושטשת ורועשת, ושחזור שלה ע"י מסנן היפוך. בעוד שהמרחק (במובן MSE) בין תמונת המקור לתמונה המקולקלת הוא 103, הרי שהמרחק בין תמונת המקור לשחזור כאן הוא $7e6$! התוצאה ללכ כל ספק גרועה!

בציור 6.4 אנו מראים לאותה תמונה בדיוק את תוצאת מסנן Wiener, עבור שלושה ערכי α שונים. עבור $\alpha = 5e-6$ מתקבלת שגיאה של 87.1, עבור $\alpha = 1e-6$ השגיאה היא 78.8, ועבור $\alpha = 5e-7$ השגיאה עולה ל-94.8. אנו רואים כי חל שיפור בתהליך השחזור, הן באופן ויזואלי, והן באופן כמותי. עם זאת, קשה להתנתק מתחושה של החמצה כשמסתכלים בתוצאות - השיזור אינו מתקרב כלל לאיכות של תמונת המקור, ונשאלת השאלה האם זה הכי טוב שניתן להשיג? למזלנו, התשובה על כך היא לא! ניתן להציע פתרונות איכותיים יותר, אך הדבר מחייב שיטות מורכבות יותר, אשר יידונו בסעיף הבא.



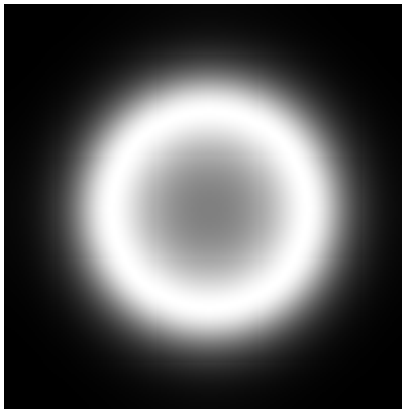

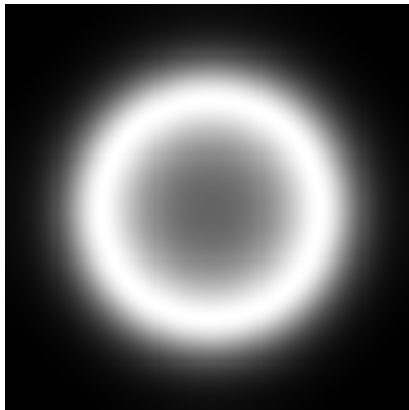



דוגמה 6.1: נניח כי אנו עוסקים באותות חד מימדיים, הטשטוש הוא הפונקציה הבאה (בתדר):

$$H_T[k] = \frac{10}{k^2 + 10} \quad k = 0, 1, \dots, 100.$$

המסנן המשחזר שהוצע ראשון בסעיף זה יהיה:

$$G_T[k] = \frac{\frac{10}{k^2 + 10}}{\left[\frac{10}{k^2 + 10}\right]^2 + 0.001} = \frac{10(k^2 + 10)}{100 + 0.001(k^2 + 10)^2}$$

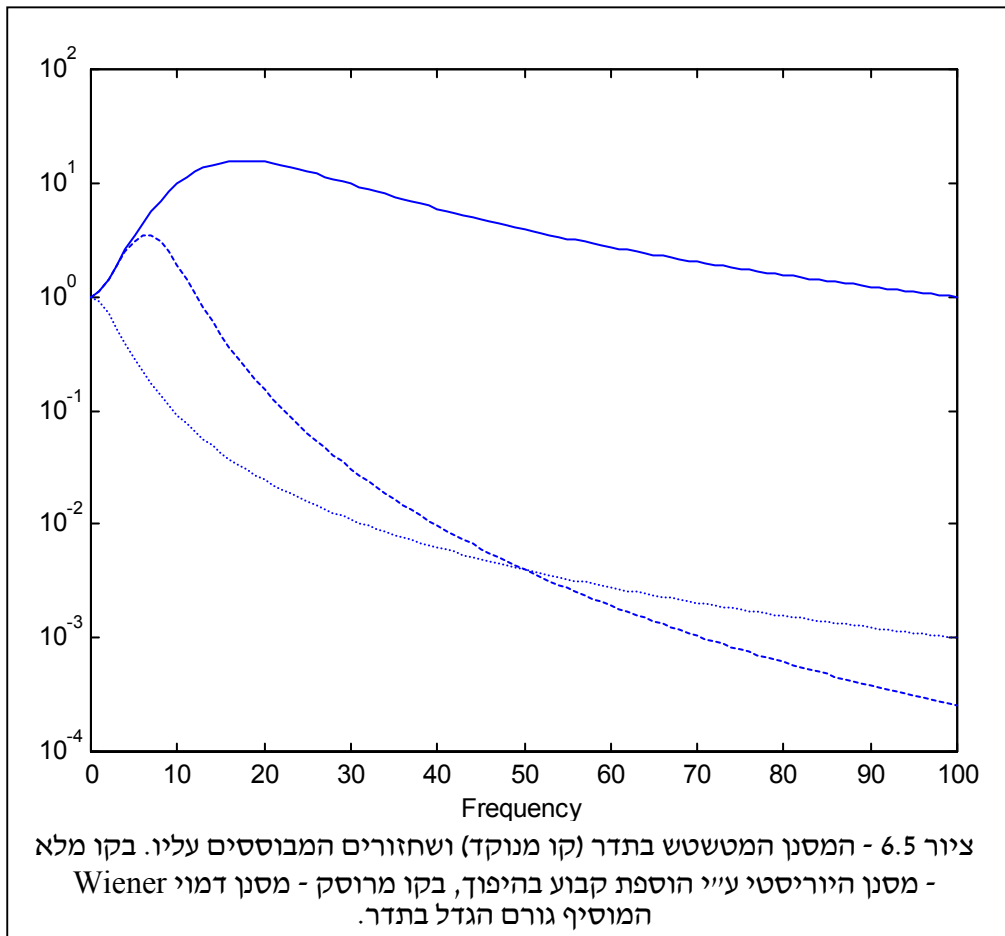
	
	
	

ציור 6.4 - שחזור עי"י מסנן Wiener. שמאל - המסננים (בערך מוחלט), ימין - תוצאות השחזורים. שלושת התוצאות מתייחסות לשלושה ערכים שונים של הפרמטר α

ואילו המסנן המשחזר השני יהיה (נבחר כאן קבוע באופן שרירותי):

$$G_T[k] = \frac{\frac{10}{k^2 + 10}}{\left[\frac{10}{k^2 + 10}\right]^2 + 0.001[0.4k^2]} = \frac{10(k^2 + 10)}{100 + 0.0004(k^2 + 10)^2 k^2}$$

ציור 6.5 מראה את שלושת מסננים אלו (המסנן המטשטש, ושני המשחזרים).



אנו רואים כי המסנן הבנוי על הוספת קבוע נותן אמנם דעיכה בהגבר אחרי תדר מסוים, אך בכל מקרה הוא גבוה. מסנן ה-Wiener מצליח לרדת חזק יותר ואף לתת ניחות חזק יותר משל המסנן המטשטש - תוצאה רצויה כיוון שאנו מניחים שבאזור זה חזק הרעש מהאות. בשני המקרים, ההתנהגות באיזור התדרים הנמוכים דומה - התנהגות של מסנן הפוך אשר אמורה לשחזר יפה מרכיבי תדר אלה.

6.3 משערכים סטטיסטיים – ML ו- MAP

בסעיף זה נציג באופן תיאורטי שני משערכים סטטיסטיים, בהם נעשה שימוש לטובת בניית אלגוריתמי שחזור תמונה טובים יותר. משערכים אלו הם משערכך הסבירות המירבית (ML - Maximum Likelihood), ומשערכך ההסתברות הסופית המקסימלית (Maximum A-posteriori) (MAP - Probability). בשני המקרים הסיפור מתחיל די דומה - נתון לנו סט נתונים שיסומן Y , ורצוננו לשערכך ערך לא ידוע X . ידוע לנו קשר בין X ל- Y באופן כלשהו (במונחים של פונקציות הסתברות), ורצוננו להציע דרך "חכמה" לחישוב של הנעלם X . נמחיש את הסיפור הנ"ל דרך דוגמה פשוטה.

דוגמה 6.1: נניח נתונה לנו סידרה בת 100 מספרים y_k אשר ידוע כי הוגרלו עפ"י חוק פילוג גאוסית חסר זיכרון עם ממוצע לא ידוע ושונות 1. הוקטור Y יהיה סט מספרים אלו. הנעלם יהיה M - הממוצע. ברמה אינטואיטיבית ידוע לנו כי שערכך ממוצע במצב כזה יכול להיעשות פשוט ע"י ממוצע המדידות, דהיינו:

$$\hat{M} = \frac{1}{100} \sum_{k=1}^{100} y_k$$

משערכך הסבירות המרבית דורש מאתנו לבנות את פונקצית ההסתברות $P(Y/X)$, כלומר פונקציה שתקבע מהי הסבירות שהמדידות תהינה Y עבור ערך הנעלם X . פונקציה זו היא פונקציה של המשתנה X . אם הצלחנו לנסח פונקציה כזו, המשערכך מציע כי נמצא את ערכו של X המביא את הסתברות זו למקסימום:

$$\hat{X}_{ML} = \underset{X}{\text{ArgMax}} P(Y/X)$$

נחזור לדוגמה הקודמת ונראה כיצד פונקציה זו מתקבלת.

דוגמה 6.1 (המשך): חוק ההסתברות של הוקטור Y הוא פשוט מכפלה של 100 הסתברויות לכל אחד מאיבריו, כאשר כל איבר כזה הוא הסתברות גאוסית. אם נניח כי אנו יודעים שהממוצע הוא M נקבל:

$$\begin{aligned} P(Y/M) &= \prod_{k=1}^{100} P(y_k/M) = \prod_{k=1}^{100} C \cdot \exp\left\{-\frac{(y_k - M)^2}{2}\right\} = \\ &= C^{100} \cdot \exp\left\{-\frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2\right\} = L(M) \end{aligned}$$

הפונקציה שהתקבלה נקראת פונקצית הסבירות L , והיא מעל המשתנה M - הממוצע הלא נודע. כזכור, רצוננו למצוא את M שיביא פונקציה זו למקסימום, ולכן:

$$\text{ArgMax}_M \left\{ C^{100} \cdot \exp \left\{ -\frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2 \right\} \right\} = \text{ArgMin}_M \left\{ \sum_{k=1}^{100} (y_k - M)^2 \right\} = \frac{1}{100} \sum_{k=1}^{100} y_k$$

ואנו רואים כי ה-ML נתן את התוצאה שהצענו קודם מתוך אינטואיציה.

דוגמה 6.2: ניח מקרה דומה, אלא שהפעם סט המדידות נובע מתוך התפלגות אקפוננציאלית מהצורה:

$$P(y_k / M) = C \cdot \exp\{-\lambda|y_k - M|\}$$

מהו משעריך ה-ML לממוצע M במקרה זה? הפתרון בדומה לקודם יהיה לכתוב את ההסתברות המותנית:

$$\begin{aligned} P(Y/M) &= \prod_{k=1}^{100} P(y_k / M) = \prod_{k=1}^{100} C \cdot \exp\{-\lambda|y_k - M|\} = \\ &= C^{100} \cdot \exp\left\{-\lambda \sum_{k=1}^{100} |y_k - M|\right\} = L(M) \end{aligned}$$

מינימיזציה הפעם תניב:

$$\text{ArgMax}_M \left\{ C^{100} \cdot \exp \left\{ -\lambda \sum_{k=1}^{100} |y_k - M| \right\} \right\} = \text{ArgMin}_M \left\{ \sum_{k=1}^{100} |y_k - M| \right\} = \text{MED}\{y_1, y_2, \dots, y_{100}\}$$

התוצאה הפעם פחות צפויה. אנו רואים כי חישובו של ממוצע הפילוג תלוי למעשה בחוק הפילוג עצמו. יכולנו לשערך את הממוצע ע"י מיצוע אלגברי, אך במקרה זה שערך זה צפוי להיות פחות טוב (לפחות במובן ML).

עד כה דיברנו על משעריך ה-ML. לפי גישה זו הנחנו כי הנעלם הוא פרמטר בחוק הפילוג ולא

משתנה אקראי בעצמו. מה אם ידוע לנו כי הנעלם X הוא אכן משתנה אקראי, וידוע לנו חוק הפילוג שלו? למשל, עבור הדוגמה הקודמת, מה אם ידוע לנו כי הממוצע M הוא משתנה אקראי גאוסי בעצמו עם ממוצע 5 ושונות 2. כיצד נרתום מידע זה על מנת לשערך טוב יותר את הנעלם?

התשובה היא משערך ה-MAP, המציע את השימוש בחוק Bayes. משערך ה-MAP מציע להביא למקסימום את ההסתברות $P(X/Y)$, שהינה ההסתברות לקבלת הערך X כשנתונות המדידות Y (בדיוק הפוך ל-ML). נוסחת Bayes קובעת כי:

$$P(X/Y) = \frac{P(X, Y)}{P(Y)} = \frac{P(X)P(Y/X)}{P(Y)}$$

ולכן משערך ה-MAP ינסה להשיג מקסימום להסתברות זו ביחס לערכו של X . מקסימיזציה ביחס ל- X פירושה שהמכנה $P(Y)$ אינו משפיע כלל, ולכן:

$$\hat{X}_{MAP} = \underset{X}{\text{ArgMax}} P(X/Y) = \underset{X}{\text{ArgMax}} \{P(X)P(Y/X)\}$$

ישנו הרבה דמיון לשערוך ה-ML. ההבדל הוא בכך שפונקציית הסבירות $P(Y/X)$ מוכפלת בפונקציית ההסתברות של X (הקרויה לעיתים **Prior-Knowledge**) כקיצור ל-**Prior-Knowledge** מידע מקדים. (מעניין לציין שכאשר אין לנו מידע מקדים על X , אנו נניח כי הפילוג שלו מתפזר אחיד על פני כל הערכים האפשריים, ואז נקבל כי משערך ה-MAP הופך להיות משערך ה-ML).

דוגמה 6.3: בהמשך לדוגמה 6.1, נניח כי ידוע כי הממוצע הלא נודע מפולג גאוסית לפי:

$$P(M) = C_0 \exp\left\{-\frac{(M-5)^2}{2\sigma_M^2}\right\}$$

במקרה זה, שערוך ה-ML לא משתנה. לעומת זאת שערוך ה-MAP יהיה:

$$\begin{aligned} \underset{M}{\text{ArgMax}} \left\{ C^{100} \cdot \exp\left\{-\frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2\right\} \cdot C_0 \cdot \exp\left\{-\frac{(M-5)^2}{2\sigma_M^2}\right\} \right\} = \\ = \underset{M}{\text{ArgMin}} \left\{ \frac{1}{2} \sum_{k=1}^{100} (y_k - M)^2 + \frac{(M-5)^2}{2\sigma_M^2} \right\} = \frac{5/\sigma_M^2 + \sum_{k=1}^{100} y_k}{100 + 1/\sigma_M^2} \end{aligned}$$

עבור $\sigma_M \rightarrow 0$ נקבל כי התוצאה שואפת ל-5. זה צפוי כי כאשר $\sigma_M \rightarrow 0$ פירוש הדבר שבודאות מוחלטת אנו יודעים כי הממוצע שווה 5 (עוד בטרם התקבלו מדידות). עבור $\sigma_M \rightarrow \infty$ נקבל כי הממוצע

האלגברי הוא השערך, כמו שהציע ה-ML. גם זו תוצאה צפויה כיוון שעבור שונות גדולה מאוד, פירוש הדבר הוא שאין לנו בעצם מידע מקדים על הנעלם, ולכן יש לדלותו מהמדידות בלבד.

6.4 שחזור תמונות מבוסס ML ו- MAP

נחזור כעת לבעיית השחזור שלנו, כשאנו מנסים להשתמש במשערך ה-ML. וקטור המדידות הוא

\underline{Y} , ווקטור הנעלמים שלנו הוא \underline{X} . הקשר ביניהם נתון ע"י $\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$, כאשר \mathbf{H} מטריצת טשטוש ו- \underline{N} וקטור רעש אקראי גאוסי בעל חוק פילוג:

$$P_N(\underline{N}) = C \cdot \exp\left\{-\frac{\underline{N}^T \underline{N}}{2\sigma_N^2}\right\}$$

עפ"י ביטוי זה, ההסתברות $P(\underline{Y}/\underline{X})$ אינה אלא:

$$P_N(\underline{Y}/\underline{X}) = C \cdot \exp\left\{-\frac{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]}{2\sigma_N^2}\right\}$$

זאת כיוון שאם מתייחסים ל- \underline{X} כוקטור פרמטרים נתון, אזי הגורם $\mathbf{H}\underline{X}$ מהווה תוספת דטרמיניסטית, והוקטור החדש $\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$ הינו גאוסי גם כן, בעל אותה שונות, ועם ממוצע חדש שהינו הגודל שהוסף. משערך ה-ML יהיה הוקטור \underline{X} אשר יביא את ההסתברות הנ"ל למקסימום. לכן:

$$\begin{aligned} \hat{\underline{X}}_{ML} &= \underset{\underline{X}}{\text{ArgMax}}\{P_N(\underline{Y}/\underline{X})\} = \underset{\underline{X}}{\text{ArgMax}}\left\{\exp\left\{-\frac{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]}{2\sigma_N^2}\right\}\right\} = \\ &= \underset{\underline{X}}{\text{ArgMin}}\left\{\frac{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]}{2\sigma_N^2}\right\} = \underset{\underline{X}}{\text{ArgMin}}\{[\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}]\} = \\ &= [\mathbf{H}^T \mathbf{H}]^{-1} \mathbf{H}^T \underline{Y} \end{aligned}$$

התוצאה שהתקבלה נכונה רק אם קיימת מטריצה הפוכה ל- $\mathbf{H}^T \mathbf{H}$, שפירושה שעל \mathbf{H} עצמה להיות לא סינגולרית. למעשה, משערך ה-ML מציע לנו להביא למינימום את הפונקציה

$f(\underline{X}) = [\mathbf{H}\underline{X} - \underline{Y}]^T [\mathbf{H}\underline{X} - \underline{Y}]$ - הצעה בהחלט אינטואיטיבית שפירושה, מצא את התמונה שאם תטושטש ע"י \mathbf{H} היא תהיה קרובה ככל האפשר ל- \underline{Y} .

התוצאה שהתקבלה אינה אלא ביצוע של מסנן הפוך כיוון שאם \mathbf{H} ריבועית מתקבל:

$$\hat{\underline{\mathbf{X}}}_{\text{ML}} = [\mathbf{H}^T \mathbf{H}]^{-1} \mathbf{H}^T \underline{\mathbf{Y}} = \mathbf{H}^{-1} \mathbf{H}^{-T} \mathbf{H}^T \underline{\mathbf{Y}} = \mathbf{H}^{-1} \underline{\mathbf{Y}}$$

כאשר \mathbf{H} אינה ריבועית) בניגוד למקרה שלנו), התוצאה הנ"ל הינה היפוך מוכלל (Pseudo-Inverse) המוכר לנו מאלגברה ליניארית. תוצאה מעניינת שהתקבלה כאן היא שלעוצמת הרעש אין משקל בקביעת המסנן המשחזר.

בכל מקרה, ההנחה שקיים היפוך ל- \mathbf{H} היא הנחה מגבילה. על מנת להיפטר מהצורך בהנחה זו,

מקובל לעשות שימוש ברגולריזציה. הרעיון די פשוט - המטריצה $\mathbf{H}^T \mathbf{H}$ חיובית חצי מוגדרת - עובדה שפירושה שכל ערכיה העצמיים ממשיים ואי-שליליים. לו היו כל הערכים העצמיים מעל האפס ממש, הרי ש- $\mathbf{H}^T \mathbf{H}$ הייתה הפיכה. לכן, נוסיף ל- $\mathbf{H}^T \mathbf{H}$ מטריצת יחידה (\mathbf{I}) מוכפלת בקבוע חיובי קטן מאוד ε^2 . תוספת זו תזיז את כל הערכים העצמיים למעלה, ותביא למצב בו $\mathbf{H}^T \mathbf{H} + \varepsilon^2 \mathbf{I}$ הפיכה. לכן, השחזור יהיה:

$$\hat{\underline{\mathbf{X}}}_{\text{ML}} = [\mathbf{H}^T \mathbf{H} + \varepsilon^2 \mathbf{I}]^{-1} \mathbf{H}^T \underline{\mathbf{Y}}$$

ההנחה היא שאם ε^2 קטן דיו לא להשפיע על ההתנהגות הכללית, ועם זאת, בעל ערך גבוה דיו למניעת בעיות יציבות נומריות, נקבל שחזור לא רע. בעקרון, יכולנו להוסיף למטריצה $\mathbf{H}^T \mathbf{H}$ כל מטריצה חיובית מוגדרת \mathbf{M} ולהשיג את אותו אפקט, אך בעצם איננו יודעים איזו מטריצה לבחור. גילינו, אם כך, כי משיקולי ML יוצא שיש להפעיל מסנן הפוך לשם שחזור, וניתן לייצבו ע"י בחירה שרירותית של מטריצת רגולריזציה.

נשתמש כעת במשעך ה- MAP. לצורך השימוש בו עלינו לעשות שימוש בפונקציה הסתברות

$P(\underline{\mathbf{X}})$. פונקציה הסתברות זו פירושה שבהינתן תמונה כלשהי $\underline{\mathbf{X}}$, נוכל לומר מהי מידת סבירותה. מה עושה תמונה לסבירה יותר? שאלה קשה אשר עד היום אין תשובה מוצלחת עליה. אבל, ישנן בהחלט תשובות חלקיות בהן נוכל לעשות שימוש.

הצעה מאוד פופולרית לאומדן הסתברות של תמונה היא ההצעה לעשות שימוש במידת חלקותה

של התמונה. ניתן לטעון כי תמונה חלקה (כמשטח בתלת-מימד) יותר סבירה מתמונה שאינה חלקה. ברור כי באזורים חלקים בתמונה הנחה זו נכונה, ובאזורים שפות הנחה זו אינה נכונה. נתעלם בשלב זה מפגם זה (מאוחר יותר נראה דרכים להתגבר על סטייה זו). את חלקותה של התמונה נוכל למדוד ע"י הפעלת נגזרת כלשהי (גרדיאנט, לפלסיאן...) וחשוב סכום ריבועי ערכי התוצאה. אם התמונה חלקה, הרי שתוצאת הנגזרת תהיה נמוכה בכל מקום ולכן הערך הכולל יהיה נמוך. אם בתמונה יש הרבה שפות נקבל כי

הנגזרות גבוהות לכן הערך הכולל יהיה גבוה. אם נניח כי המטריצה \mathbf{D} מייצגת פעולת נגזרת מסדר שנבחר, פונקצית ההסתברות המוצעת היא:

$$P(\underline{X}) = C \cdot \exp\left\{-\frac{[\underline{\mathbf{D}}\underline{X}]^T [\underline{\mathbf{D}}\underline{X}]}{2\sigma_D^2}\right\}$$

הביטוי $\underline{\mathbf{D}}\underline{X}$ מקבל את התמונה X ומבצע עליה פעולת גזירה, והביטוי $[\underline{\mathbf{D}}\underline{X}]^T [\underline{\mathbf{D}}\underline{X}]$ לוקח את תוצאת הגזירה מעלה בריבוע כל איבר ומסכם. לכן, ערך זה יהיה נמוך לתמונות חלקות, וגבוה לתמונות שאינן חלקות. השימוש במבנה של פונקצית אקספוננט נובע מנוחות שתתבהר בהמשך. הפרמטר σ_D קובע את שונותה של הסתברות זו. בכל אופן, רואים שעבור תמונה חלה ההסתברות גבוהה, ועבור תמונה עם ריבוי שפות, ההסתברות נמוכה, כנדרש.

משעריך ה- MAP מביא את מכפלת פונקצית הסבירות (הלוקחה מה- ML) ופונקצית ה- Prior הנ"ל למקסימום, ולכן:

$$\begin{aligned} \hat{X}_{MAP} &= \underset{\underline{X}}{\text{ArgMax}} \{P_N(\underline{Y}/\underline{X})P_X(\underline{X})\} = \\ &= \underset{\underline{X}}{\text{ArgMax}} \left\{ \exp\left\{-\frac{[\underline{Y} - \underline{\mathbf{H}}\underline{X}]^T [\underline{Y} - \underline{\mathbf{H}}\underline{X}]}{2\sigma_N^2}\right\} \cdot \exp\left\{-\frac{[\underline{\mathbf{D}}\underline{X}]^T [\underline{\mathbf{D}}\underline{X}]}{2\sigma_D^2}\right\} \right\} = \\ &= \underset{\underline{X}}{\text{ArgMin}} \left\{ \frac{[\underline{Y} - \underline{\mathbf{H}}\underline{X}]^T [\underline{Y} - \underline{\mathbf{H}}\underline{X}]}{2\sigma_N^2} + \frac{[\underline{\mathbf{D}}\underline{X}]^T [\underline{\mathbf{D}}\underline{X}]}{2\sigma_D^2} \right\} = \\ &= \underset{\underline{X}}{\text{ArgMin}} \left\{ [\underline{Y} - \underline{\mathbf{H}}\underline{X}]^T [\underline{Y} - \underline{\mathbf{H}}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} [\underline{\mathbf{D}}\underline{X}]^T [\underline{\mathbf{D}}\underline{X}] \right\} = \\ &= \left[\underline{\mathbf{H}}^T \underline{\mathbf{H}} + \frac{\sigma_N^2}{\sigma_D^2} \underline{\mathbf{D}}^T \underline{\mathbf{D}} \right]^{-1} \underline{\mathbf{H}}^T \underline{Y} \end{aligned}$$

אנו רואים כי הפעם, המסנן התקבל עם רגולריזציה שנבעה מתוך שיקולי ה- MAP. כלומר, אם רצוננו בשחזור מוצלח, מוטב שנבחר לעשות שימוש מטריצת גזירה כלשהי. כמו כן, אנו רואים כי מנת שתי השונויות פועלת במשקל לגורם הרגולריזציה הנוסף.

בהשוואה ל- ML, משעריך ה- MAP מציע מינימיזציה של הפונקציה

$$f(\underline{X}) = [\underline{Y} - \underline{\mathbf{H}}\underline{X}]^T [\underline{Y} - \underline{\mathbf{H}}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} [\underline{\mathbf{D}}\underline{X}]^T [\underline{\mathbf{D}}\underline{X}]$$

בפונקציה זו המרכיב הראשון זהה לזה של ה-ML, והמרכיב השני מושך את הפתרון לעבר תמונות חלקות יותר. אם תמונה מסוימת מצליחה להביא לערך מאוד נמוך את הגורם הראשון אין בכך די לשם קבלתה כפתרון - המחיר השני גם הוא משחק תפקיד בקביעת הפתרון. כאשר σ_D נמוך מאוד פירוש הדבר שאנו דורשים שהתוצאה תהיה חלקה מאוד, ואמנם בפונקציה f הגורם השני נהיה דומיננטי, ולכן הפתרון יהיה חלק. עבור $\sigma_D \rightarrow 0$ הבעיה דורשת חלקות מוחלטת, ולכן התמונה תהיה קבוע עבור גזירה ראשונה ומישור מוטה עבור גזירה שניה. כאשר $\sigma_D \rightarrow \infty$ אנו מחלישים את דרישת החלקות ואז יתנקז הפתרון לתוצאת ה-ML.

כנקודה אחרונה בסעיף זה נתייחס לפתרון מעשי של בעיית השחזור ע"י הכלים שהוצגו לעיל. ניקח את תוצאת ה- MAP כיוון שהיא מכילה את ה-ML, ונראה כיצד מעשית משיגים שחזור באמצעותה. נניח כי לשם רגולריזציה אנו משתמשים במסנן 3 על 3 לפלסיאן. אם נניח כי התמונה בגודל 200×200 , הרי שהמטריצות המעורבות (\mathbf{D} , \mathbf{H}) תהיינה בגודל של $40,000 \times 40,000$. מטריצות אלו דלילות (מרבית איברייהן אפסים), אך היפוך מטריצה בגודל כזה אינו בא בחשבון. שתי אלטרנטיבות קיימות הן פתרון במישור התדר, ושיטות איטרטיביות.

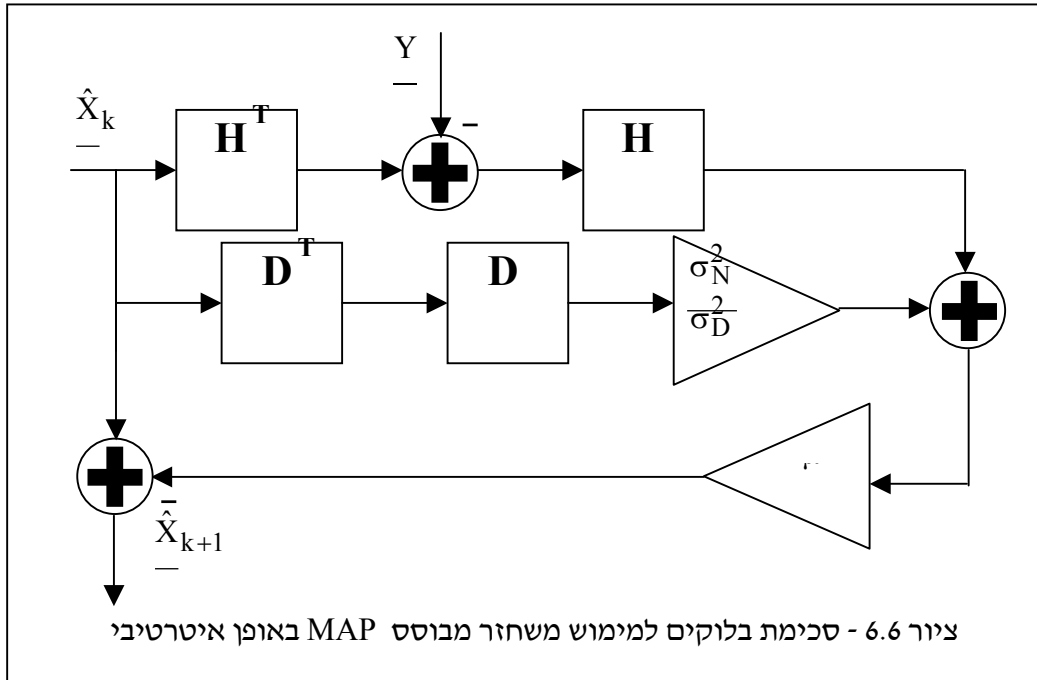
באשר לפתרון איטרטיבי, הרעיון הבסיסי ביותר הוא אלגוריתם השיפוע המירבי (Steepest Descent). הרעיון הוא להתחיל מניחוש מוצלח לפתרון (למשל $\hat{\mathbf{X}}_0 = \mathbf{Y}$) ולשפרו שוב ושוב ע"י ירידה בכיוון הפוך לכיוון הגרדיאנט של הפונקציה $f(\mathbf{X})$. הגרדיאנט מצביע על הכיוון בו חלה העלייה הגדולה ביותר, ולכן הליכה בכיוון הפוך פירושה ירידה מירבית. לכן, משוואת האיטרציה תהיה:

$$\hat{\mathbf{X}}_{k+1} = \hat{\mathbf{X}}_k - \mu \left. \frac{\partial f(\mathbf{X})}{\partial \mathbf{X}} \right|_{\hat{\mathbf{X}}_k} = \hat{\mathbf{X}}_k - \mu \left[\mathbf{H}^T (\mathbf{H} \hat{\mathbf{X}}_k - \mathbf{Y}) + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{D} \hat{\mathbf{X}}_k \right]$$

הגודל μ הוא גודל הצעד לעדכון - ערך גדול מידי ייתן התבדרות, ערך קטן מידי ייתן התכנסות איטית מידי, ויש לבחור אותו בחוכמה (קיימות דרכים לקביעתו באופן אוטומטי, שלא יידונו כאן). התהליך המתואר מובטח להתכנס לפתרון במדויק של הבעיה בשל היותה של הפונקציה f קמורה, ובשל תכונות האלגוריתם המוצע.

כיוון ש- \mathbf{H} ו- \mathbf{D} מסננים ליניאריים וקבועים במקום, הרי שגם לאחר שחלוף פעולות אלו הינן ליניאריות וקבועות במקום. יתרה מזו, כאשר הטשטוש בעלת גרעין סימטרי, מתקבל כי $\mathbf{H}^T = \mathbf{H}$ (גם הלפלסיאן מקיימת תכונה זו). לכן, אין כלל צורך לעבור עם תמונות בסידור לקסיקוגרפי או מטריצות כה גדולות. ניתן פשוט להתייחס למשוואה הנ"ל כסדרת צעדי סינון שעל הפתרון הקודם לעבור לשם עדכון. ציור 6.6 מראה את סכימת הבלוקים של תהליך העדכון.

ציור 6.7 מראה גרף של ערך הפונקציה $f(\underline{X})$ שמובאת למינימום בתהליך האיטרטיבי כפונקציה של מספר האיטרציה. בהרצה זו נעשה שימוש בלפליסיאן לצורך רגולריזציה, ו- $\sigma_D^2 = 25$. רואים כי חלה ירידה עקבית עד להתכנסות. הגרפים השונים מתייחסים לערכי μ שונים. כפי שכבר נאמר, עבור ערך נמוך מידי ההתכנסות איטית, ועבור ערך גדול מידי חלה התבדרות.

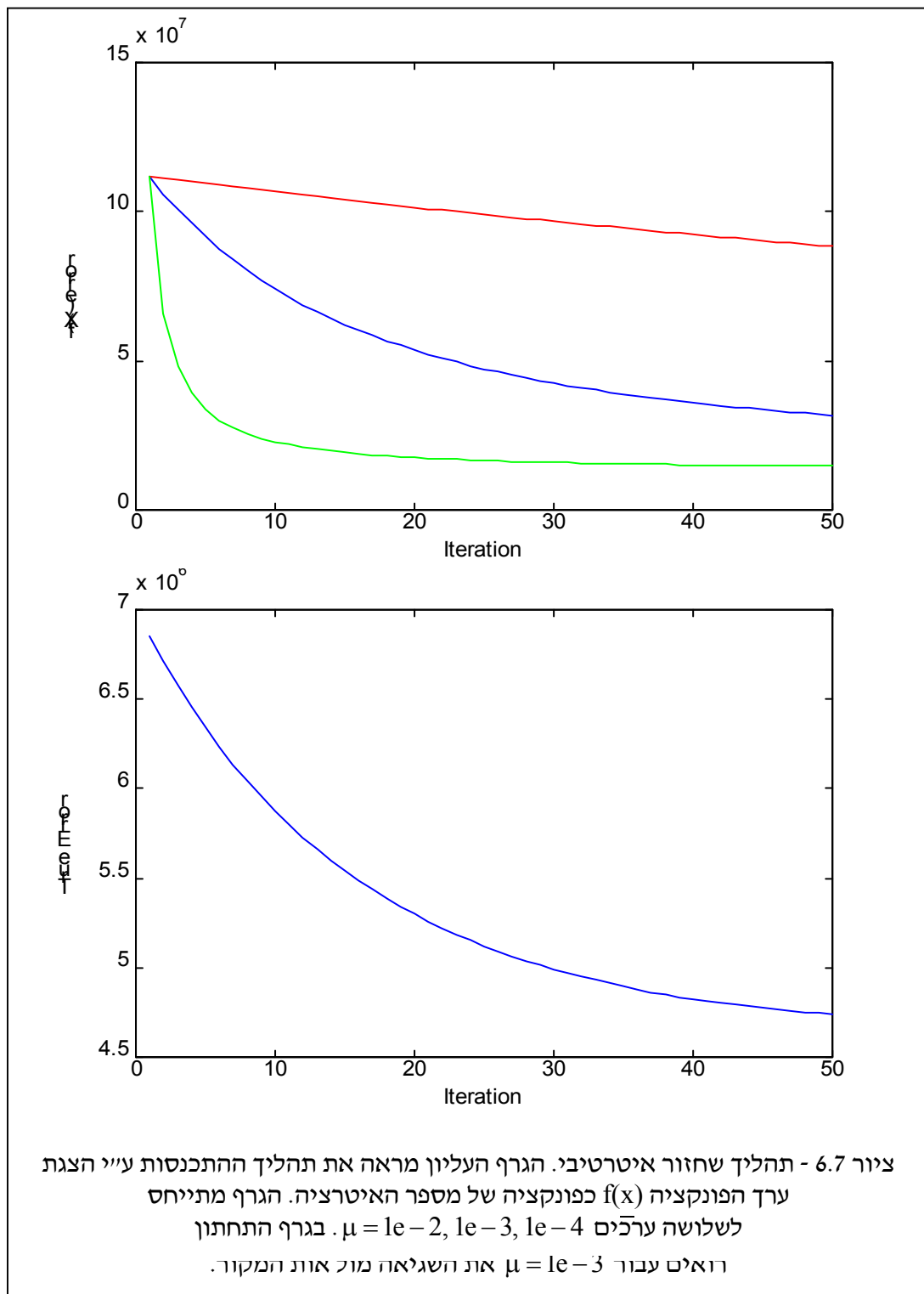


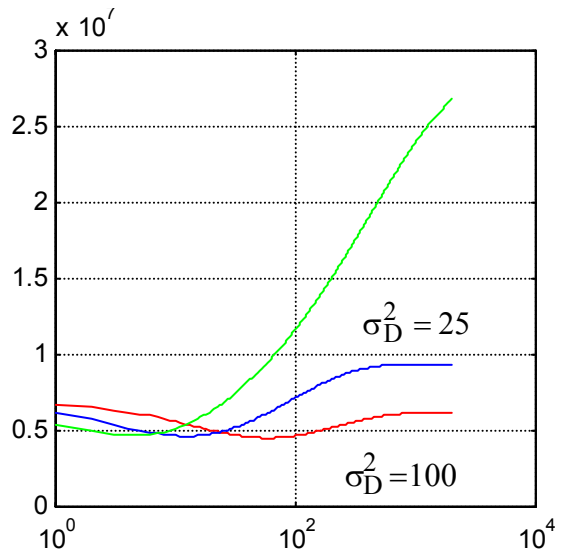
הגרף השני בציור זה מראה את השגיאה $\|\underline{X} - \hat{\underline{X}}\|^2$. כפי שרואים, גם כאן מושגת ירידה עקבית, אם כי אין זו סתירה לתיאוריה לקבל גם עליה - נזכור כי רצוננו האמיתי הוא בהשגת מינימום לפונקציה זו, אך התמונה המקורית אינה קיימת בידינו. לכן, הפונקציה אותה אנו מביאים למינימום $f(\underline{X})$ אינה חיקוי לשגיאה האמיתית.

ציור 6.8 מראה תוצאות שחזור MAP ע"י הגישה האיטרטיבית למגוון ערכי σ_D . אנו רואים כי עבור בחירת ערך גבוה התוצאה רועשת מידי (תופעת ה- "צלצולים" - Ringing) בשל הגבר הרעש. עבור ערך נמוך התוצאה חלקה. בציור זה גם מובאת השגיאה האמיתית (מול התמונה האידיאלית) לשם השוואה. רואים כי לשלושת ערכי σ_D מתקבל אפקט דומה בו השגיאה יורדת באיטרציות הראשונות ואז עולה ומתייצבת על ערכה הסופי (כעבור 2000 איטרציות!).

נציג כעת את הפתרון המבוסס על התמרת תדר כששוב אנו מתייחסים לפתרון MAP. כזכור מפרקים קודמים ומתחילת פרק זה, המטריצה T המייצגת את התמרת הפורייה הדיסקרטית הזו

מימדית, מלכסנת מטריצות בלוק סיבוביות במקרה שלנו, הן H והן D בעלות מבנה כזה בשל היותן ייצוג לפעולה ליניארית וקבועה במקום. לכן :





ציור 6.8 - תמונת מקור (למעלה בצד שמאל), תמונה מקולקלת (למעלה בצד ימין), תוצאות שחזור MAP עם ערכים שונים של σ_D^2 (100 - שורה שניה בצד שמאל, 25 - שורה שניה בצד ימין, 1 - שורה תחתונה בצד שמאל). כן מצורף גרף המתאר את השגיאה האמיתית כפונקציה של מספר האיטרציה

$$\begin{aligned} \left[\mathbf{H}^T \mathbf{H} + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{D} \right] \hat{\mathbf{X}} &= \mathbf{H}^T \mathbf{Y} \\ \Rightarrow \left[\mathbf{T} \mathbf{H}^T \mathbf{T}^H \mathbf{T} \mathbf{H} \mathbf{T}^H + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{T} \mathbf{D}^T \mathbf{T}^H \mathbf{T} \mathbf{D} \mathbf{T}^H \right] \mathbf{T} \hat{\mathbf{X}} &= \mathbf{T} \mathbf{H}^T \mathbf{T}^H \mathbf{T} \mathbf{Y} \\ \Rightarrow \left[\mathbf{H}_T^H \mathbf{H}_T + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}_T^H \mathbf{D}_T \right] \hat{\mathbf{X}}_T &= \mathbf{H}_T^H \mathbf{Y}_T \end{aligned}$$

הוקטור $\hat{\mathbf{X}}_T$ הוא התמונה הנדרשת במישור התדר. באופן דומה, \mathbf{Y}_T היא התמונה הנמדדת, גם בתדר. הקשר בין שתי תמונות אלו נתון ע"י המשוואה הנ"ל, כשהמטריצות הכופלות (בשני האגפים) אלכסוניות! לכן, נוכל לקבל קשר סקלרי פשוט בין האיבר ה- k של \mathbf{Y}_T לאיבר ה- k של $\hat{\mathbf{X}}_T$:

$$\left[\mathbf{H}_T^H \mathbf{H}_T + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}_T^H \mathbf{D}_T \right] \hat{\mathbf{X}}_T = \mathbf{H}_T^H \mathbf{Y}_T \Rightarrow \hat{\mathbf{X}}_T[k] = \frac{(\mathbf{H}_T[k])^*}{\left(|\mathbf{H}_T[k]|^2 + \frac{\sigma_N^2}{\sigma_D^2} |\mathbf{D}_T[k]|^2 \right)} \mathbf{Y}_T[k]$$

קשר זה הוא בדיוק הקשר שהוצג כמסנן Wiener, כך שהפעם בנינו ממש את המסנן מתוך שיקולים תיאורטיים מבוססי שערך MAP. עם זאת לגישה אשר נדונה בסעיף זה יש יתרון חשוב- הכללתה לשחזור ע"י סינון לא ליניארי תהיה קלה ופש וטה. נשים לב כי פעולת הלפלטיאן היא מסנן מעביר גבוהים (HPF), ולכן התמרת התדר של פעולת הלפלטיאן הינה גורם שעולה עם התדר. לכן, במכנה בביטוי הנ"ל הגורם של הרגולריזציה יהיה זניח בתדרים נמוכים (בנדרש) ומשמעותי בתדרים גבוהים.

לסיכום ביניים - גישת ה- MAP הביאה לנו יכולת לבנות את מסנן ה- Wiener מתוך שיקולים עמוקים יותר. אבל - לא קיבלנו בינתיים משהו חדש, והביצועים הצפויים (וזה ניכר מצויר 6.8) אינם מלבבים. היתרון של שיטת ה- MAP הוא בכך שהיא ניתנת להכללה לסינון לא ליניארי, כפי שנראה בסעיף הבא.

6.5 שיפור פונקצית ההסתברות לתמונות

גילינו כי פעולת הרגולריזציה המבוססת על חלקות גורמת להחלקת יתר של התמונה. כאשר מקטינים את השפעת ה-Prior, אנו מתקרבים לביצועי מסנן מהפך, וזה כמובן גרוע. כך יוצא שעלינו למצוא ערך ביניים אשר מנסה ליהנות מכל העולמות, ובפועל סובל מכל המחלות, והתוצאות מאכזבות. הסיבה לאי-הצלחה כאן היא העובדה שהנחנו הנחה לא נכונה במודע - שכל סביבה בתמונה תעדיף להיות חלקה. ברור שבאזורים חלקים הדבר נכון, אך אזורי שפות הם העיקר בתמונה, ובהם פגענו. הפתרון לפיכך הוא לשנות את חוק ההסתברות שהנחנו על תמונות. נזכור כי הנחתנו המקורית היא:

$$P(\underline{X}) = C \cdot \exp \left\{ - \frac{[\underline{DX}]^T [\underline{DX}]}{2\sigma_D^2} \right\}$$

גישה אפשרית אחת לשפר את חוק ההסתברות זה היא להישאר במסגרת פילוג גאואסי, אך להרשות שונות תלוית מקום. אפשרות אחרת היא לשנות את חוק הפילוג מגאואסי למשהו אחר נכון יותר. אנו נכיר שתי גישות אלו, ונראה כי קיים דמיון ביניהן.

נתחיל באפשרות של הרחבת ההסתברות הגאוסית לשונות תלוית מקום. ההסתברות הגאוסית במקרה הכללי ביותר לוקטור \underline{X} היא:

$$P(\underline{X}) = C \cdot \exp \left\{ - \frac{1}{2} [\underline{X} - \underline{M}]^T \mathbf{Q}^{-1} [\underline{X} - \underline{M}] \right\}$$

כאשר \underline{M} הוא הממוצע והמטריצה \mathbf{Q} היא מטריצת הקווריאנס. לפי הגירסה הקודמת בחרנו ממוצע אפס, ומטריצת שונות

$$\mathbf{Q}^{-1} = \mathbf{D}^T \mathbf{D}$$

באשר לממוצע, לא נציע שינוי. לעומת זאת, חסרון בולט במטריצת הקווריאנס היא היותה בעלת מבנה בלוק טופליץ (כיוון ש- \mathbf{D} בעלת מבנה כזה). במילים אחרות, מטריצת הקווריאנס מניחה כי בכל פיקסל יש אותה התנהגות (מקביל לתכונת ה-Linear Space Invariance). על מנת לגוון מעט את הבחירה, נציע הכנסתה של מטריצת משקל אלכסונית \mathbf{W} , ונקבל

$$\mathbf{Q}^{-1} = \mathbf{D}^T \mathbf{W} \mathbf{D}$$

כאשר מטריצה זו תכיל ערכים זהים על אלכסונה, נתלכד עם הבחירה הקודמת, וכאשר הערכים ישתנו לאורך האלכסון, נקבל התנהגות תלויית מקום. משערך ה-MAP הפעם יתקבל כנקודת המינימום של הביטוי

$$f(\underline{X}) = [\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} [\mathbf{D}\underline{X}]^T \mathbf{W} [\mathbf{D}\underline{X}]$$

ומביטוי זה נוכל להסיק כיצד עלינו לפעול בבחירת ערכי \mathbf{W} . הוקטור $\mathbf{D}\underline{X}$ היא תמונת נגזרת (כלשהי) של תמונת המקור \underline{X} . באזורים שאמורים להיות חלקים נכפה חלוקת ע"י בחירת משקל 1. במקומות בהם צפויה שפה, נבחר ערך הקרוב לאפס. באופן זה נכפה חלקות באופן תלוי מקום. באשר לבנייתה של \mathbf{W} , נוכל לקחת את התמונה המקולקלת \underline{Y} , ולהשתמש בה כדי לקבוע בכל פיקסל האם ישנה שפה אם לאו.

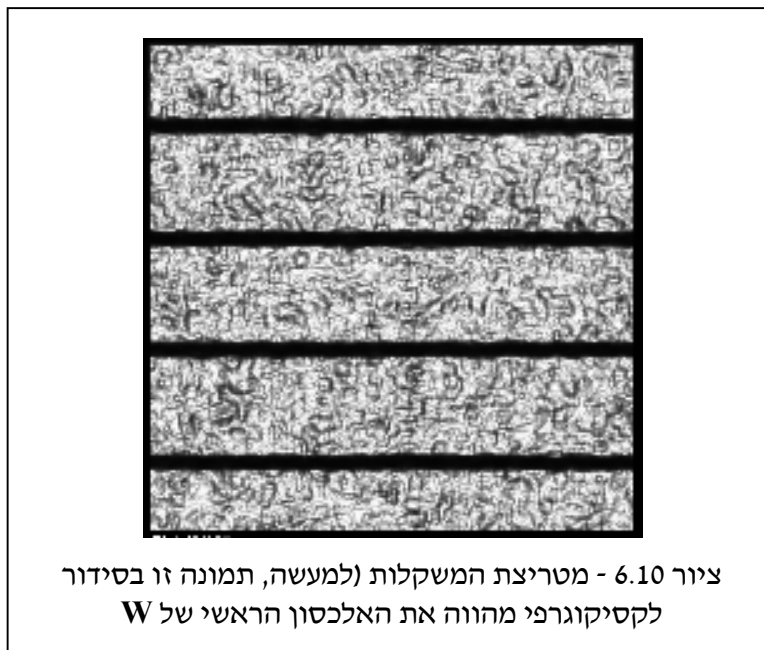
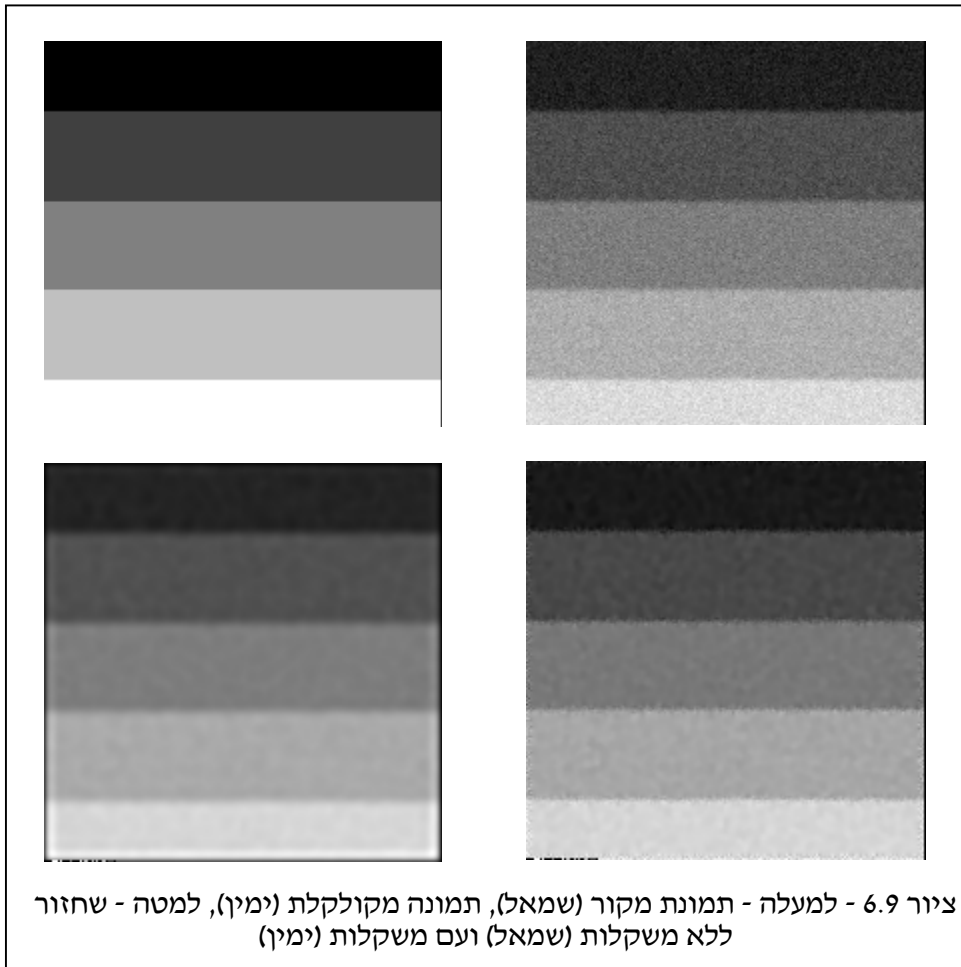
בניגוד למקרה הקודם, הפעם לא ניתן לעבוד בתדר כיוון שחלק מהפעולות אינן קבועות במקום. גישת האיטרציות פועלת באופן הבא:

$$\hat{\underline{X}}_{k+1} = \hat{\underline{X}}_k - \mu \left. \frac{\partial f(\underline{X})}{\partial \underline{X}} \right|_{\hat{\underline{X}}_k} = \hat{\underline{X}}_k - \mu \left[\mathbf{H}^T (\mathbf{H}\hat{\underline{X}}_k - \underline{Y}) + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \mathbf{W} \mathbf{D} \hat{\underline{X}}_k \right]$$

כך שהסכימה דומה מאוד לקודמת, עם שינוי אחד - לאחר הפעלת מסנן גוזר על התמונה, יש לכפול את התמונה בתמונת מסיכה שהינה בדיוק המשקלות שנבחרו.

לעיתים ניתן אף לבחור ערך משקל שלילי במקום בו צפויה שפה לשם השגת פעולה הפוכה של חידוד. בחירת הערך השלילי מסוכנת משום שאז יתכן שהפונקציה הנ"ל כבר אינה קמורה, ובעלת נקודות מינימה רבות. כמו כן הדבר מסכן את יציבות התהליך האיטרטיבי שתואר לעיל.

בציור 6.9 נראה תהליך השחזור עבור תמונה סינתטית. בציור מובאת התמונה המקורית, התמונה המקולקלת, שחזור עם מישקול קבוע (הגישה הישנה) ושחזור עם משיקול תלוי מקום. שגיאת השחזור בני המקרים היא - ללא משקלות - MSE=27.06, ועם משקלות - MSE=21.001. ערכים אלו התקבלו ע"י החסרת התמונה המקורית מתוצאת השחזור וסיכום ריבועי ההפרשים. לצורך זה הושמטה רצועה של 5 פיקסלים מקצוות התמונה, מכיון שממילא שם פעולותינו לא מוצלחות בכל שיטה. רואים שמושג שיפור משמעותי ע"י הכנסת המשקלות לשחזור. חשוב להבהיר כי השיפור הפוטנציאלי טוב אף יותר כיוון שניתן היה להגדיל את המקדם של הרגולריזציה ולכפות חלקות חזקה יותר באזורים החלקים, כך שפוטנציאל גישה זו לא מוצא עד תום. ציור 6.10 מראה את מטריצת המשקל שנבחרה.



המשקלות נבנו ע"י סידרת הצעדים הבאה :

1. לקיחת התמונה המקולקלת והחלקתה ע"י מסנן 5 על 5. צעד זה נועד לדכא את הרעש בתמונה לפני חישוב נגזרות.

2. חישוב הנגזרות האופקית והאנכית ע"י מסנן מהצורה $[1 \ 0 \ -1]$ (בשני הצירים), וחישוב נורמת הגרדיאנט (סכום ריבועי שני הנגזרות). כעת יש לנו תמונה בה ערך גבוה מעיד על שפה, וערך נמוך מתייחס לאזור חלק.

בניית משקל בכל נקודה ע"י הנוסחה $-0.01 + 10/[Grad + 10]$. כך קיבלנו כי כאשר Grad הוא 0, המשקל הוא 0.99, וכאשר הוא נוטה לאינסוף המשקל הוא -0.01.

עד כה תיארונו כיצד נכליל את הנחת הגאוסיות שלנו ע"י הכנסת מומנט סדר שני מורכב יותר. נראה כעת גישה אחרת שמשמעותה בריחה מהנחת הגאוסיות אל חוקי פילוג אחרים. נחזור לבסיס - אנו לוקחים את התמונה X ומבצעים עליה פעולת גזירה כלשהי D . נסתכל על פיקסל מסויים ונניח כי ערך הנגזרת הוא g . בשיטה הקודמת, g עולה בריבוע והתוצאה היא "מחירו" של פיקסל זה. מדוע לעלות בריבוע ולא ברביעית? מדוע לא ערך מוחלט? וניתן כמובן להציע מגוון אפשרויות אחרות.

ננסה להבין מה המשמעות של בחירות שונות של החזקה בדרך הבאה. נניח שרצוננו לייצר פונקציה $f[k]$ (חד-מימדית) חלקה ודיסקרטית אשר תעבור מ-0 ל-1 ידוע לנו כי הפונקציה ב- $k=0$ היא 0, וב- $k=10$ היא 1. נציע שתי חלופות - לפונקציה f ונבחן את מחירן. חלופה ראשונה תהיה ש- f_1 תעלה בהדרגה 10 מדרגות בגובה 0.1 כל אחת, וחלופה שניה בה f_2 עולה ב- $k=5$ בבת אחת מ-0 ל-1. נניח כי פעולת הנגזרת היא פשוט הפרשים של דגימות סמוכות. לפי השיטה המקורית נקבל

$$\sum_{k=1}^{10} (f_1[k] - f_1[k-1])^2 = 0.1^2 + 0.1^2 + \dots + 0.1^2 = 10 \cdot 0.1^2 = 0.1$$

$$\sum_{k=1}^{10} (f_2[k] - f_2[k-1])^2 = 0^2 + \dots + 1^2 + \dots + 0^2 = 1$$

לכן ברור כי תמחור כזה יעדיף עליה הדרגתית. נניח כעת כי החלפנו את החזקה בערך מוחלט. נקבל:

$$\sum_{k=1}^{10} |f_1[k] - f_1[k-1]| = |0.1| + |0.1| + \dots + |0.1| = 10 \cdot 0.1 = 1$$

$$\sum_{k=1}^{10} |f_2[k] - f_2[k-1]| = 0 + \dots + |1| + \dots + 0 = 1$$

והפעם קיבלנו כי מחיר שני הגישות זהה. נמשיך במגמה זו ונציע שורש על הערך המוחלט ונקבל:

$$\sum_{k=1}^{10} |f_1[k] - f_1[k-1]|^{0.5} = |0.1|^{0.5} + |0.1|^{0.5} + \dots + |0.1|^{0.5} = 10 \cdot 0.1^{0.5} = 3.16$$

$$\sum_{k=1}^{10} |f_2[k] - f_2[k-1]|^{0.5} = 0 + \dots + |1|^{0.5} + \dots + 0 = 1$$

ובמקרה זה הפונקציה תעדיף עליה חדה על-פני עליה מתונה.

מדוע כל זה חשוב? ובכן, אנו מניחים שכאשר מבוצעת הפעולה \mathbf{DX} , ערכים (בערך מוחלט) בסביבת האפס מתייחסים לאזור נמוך, ולכן הגישה הריבועית יפה להם, כלומר נרצה שמעבר בין רמות יקרה בהדרגה. לעומת זאת, מערך מסוים ואילך כנראה שמדובר בשפות, ולכן נרצה שמעבר בין רמות יקרה באופן חד. לכן, נציע באופן כללי כי פונקצית המחיר לחלקותה של התמונה תהיה:

$$\sum_k \rho(\mathbf{DX}|_k)$$

כלומר, נבצע כמקודם פעולת גזירה \mathbf{D} , וכל ערך בתמונת התוצאה יוכנס לפונקציה $\rho(x)$. אם $\rho(x) = x^2$ חזרנו לגרסה המקורית. אם נבחר $\rho(x)$ שיתנהג כפונקציה ריבועית סביב הראשית, וימתן את עלייתו מנקודה מסוימת, הרי שנשיג את שרצינו. ישנן מספר אפשרויות - נציג שתיים מהן:

$$\rho(x) = \frac{x^2}{x^2 + \alpha} \quad \text{or} \quad \rho(x) = \begin{cases} -T(2x + T) & x < -T \\ x^2 & -T \leq x \leq T \\ T(2x - T) & x > T \end{cases}$$

הבחירה הראשונה ידועה בשם פונקצית קושי, והשנייה בשם פונקצית Huber-Markov. שימוש

בפונקציות $\rho(x)$ מוכרת בסטטיסטיקה בשם שערך רובסטי (Robust Estimation), והפונקציות זכו לשם M-functions.

הבעיה עם גישת השערך הרובסטי היא שעבור בחירה של פונקציות שעולות בקצב נמוך מליניארי (כמו הדוגמה הראשונה מבין השתיים במשוואה הקודמת), הפונקציה שמובאת למינימום כבר אינה קמורה, ואז מתקבלות תופעות של נקודות מינימה רבות, ובעיות התכנסות למינימה הגלובלית בהתאם לכך. בשל כך, בחירת פונקצית ערך מוחלט $\rho(x) = |x|$ זוכה לפופולריות רבה.

עבור שימוש בפונקצית המחיר החדשה, הפונקציה הכוללת שישי להביא למינימום היא:

$$f(\mathbf{X}) = [\mathbf{Y} - \mathbf{HX}]^T [\mathbf{Y} - \mathbf{HX}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{DX}]$$

כאשר לפי הסימון שלנו, ρ על וקטור הוא פעולת $\rho(x)$ על כל אחד מאיבריו, וסיכום התוצאות. מתוך ביטוי זה ניתן ללכת אחורה ולקבוע כי בעצם הנחנו כי פונקציה ההסתברות של מרחב התמונות נתונה ע"י:

$$P(\underline{X}) = C \cdot \exp \left\{ -\frac{\rho[\underline{DX}]}{2\sigma_D^2} \right\}$$

פונקציות פילוג בעלות מבנה זה קרויות פילוג Gibbs. המשותף לכולן הוא ההסתברות ל- \underline{X} יחסית הפוך דרך אקספוננט למידת האנרגיה שלו, והמעריך הוא סוג של אנרגיה.

האם יש קשר בין גישה זו לגישה הקודמת של בחירת מטריצת משקל? לשם גילוי קשר כזה, נכתוב את משוואת האיטרציה שיש להפעיל לשם מינימיזציה של הביטוי שנכתב למעלה:

$$\hat{X}_{k+1} = \hat{X}_k - \mu \frac{\partial f(\underline{X})}{\partial \underline{X}} \Big|_{\hat{X}_k} = \hat{X}_k - \mu \left[\mathbf{H}^T (\mathbf{H}\hat{X}_k - \underline{Y}) + \frac{\sigma_N^2}{\sigma_D^2} \mathbf{D}^T \rho'[\underline{D}\hat{X}_k] \right]$$

משמעות משוואה זו היא שיש לגזור את \underline{X} כמקודם, ואז להעביר כל פיקסל דרך נגזרת הפונקציה $\rho(x)$, ואת התוצאה שוב לגזור. לשם דוגמה - עבור $\rho(x) = 0.5x^2$ ברור כי $\rho'(x) = x$, ואז נקבל כי פעולת ה- ρ' משאירה את הערכים כמות שהם. עבור $\rho(x) = |x|$ הנגזרת תהיה $\rho'(x) = \text{sign}(x)$, כך שכל פיקסל לאחר נגזרת יושווה לאפס - שליליים יהפכו ל-1, וחיוביים ל+1. בהשוואה למשוואה האיטרטיבית במקרה של השימוש במשקל, שני הגורמים הנ"ל מתחלפים:

$$\text{(גישת } \rho(x) \text{)} \quad \rho'[\underline{D}\hat{X}_k] \Leftrightarrow \mathbf{W}\underline{D}\hat{X}_k \quad \text{(משקל)}$$

בשני המקרים התהליך מתחיל זהה - התמונה \hat{X}_k עוברת דרך פעולת הגזירה \mathbf{D} . בשני המקרים הפעולה הבאה היא פעולה על כל פיקסל בנפרד. האחת כופלת במשקל כלשהו, והאחרת מעבירה דרך פונקציה כלשהי. לכן, אם בגישת המשקלות לא היינו משתמשים במטריצת משקל קבועה לאורך כל התהליך האיטרטיבי, אלא משנים אותה בכל איטרציה להיות

$$\mathbf{W}|_{\text{pos}} = \frac{\rho'[\underline{D}\hat{X}_k]_{\text{pos}}}{[\underline{D}\hat{X}_k]_{\text{pos}}}$$

היינו מקבלים אותה תוצאה בדיוק. מה משמעות בחירה זו? לדוגמה עבור $\rho(x) = 0.5x^2$ יתקבל כי \mathbf{W} אינה אלא 1 לכל הפיקסלים. עבור $\rho(x) = |x|$ נקבל כי המשקל יחסי הפוך לגודל הנגזרת בערכה המוחלט, כיוון ש:

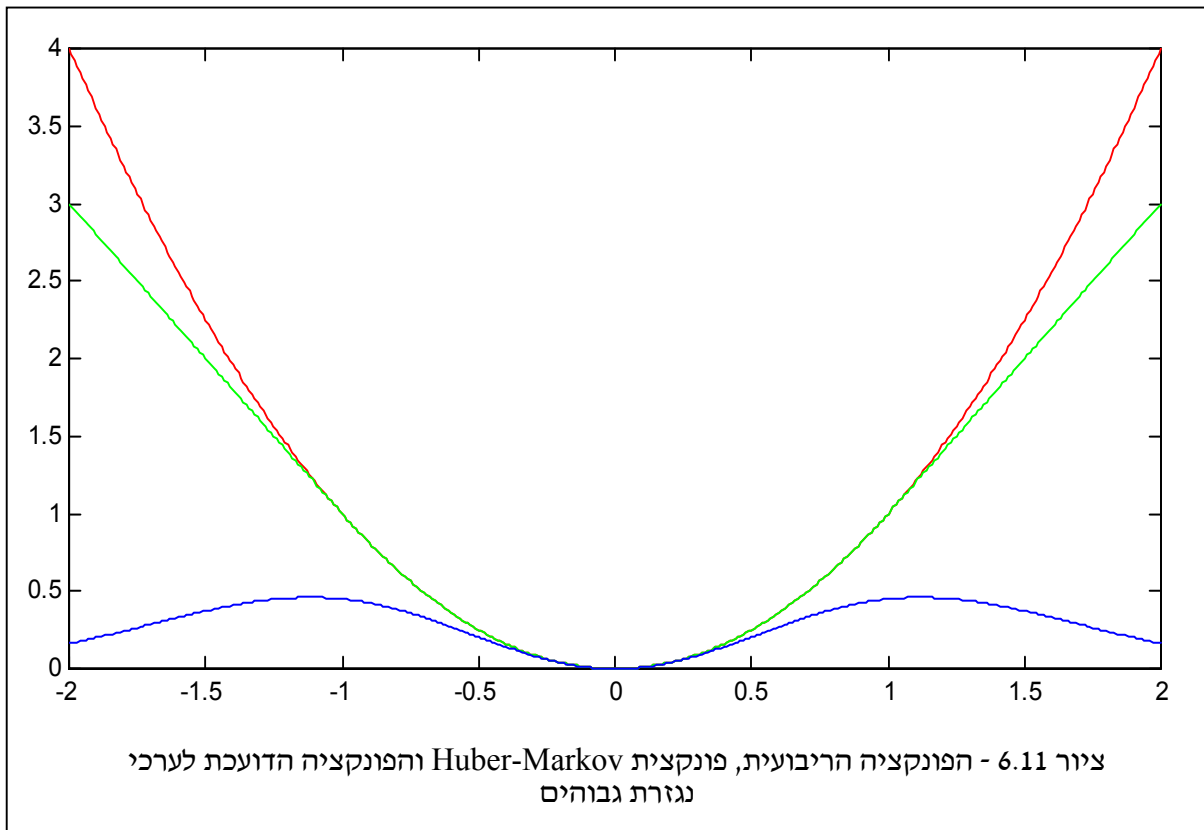
$$W|_{\text{pos}} = \frac{\rho[\mathbf{D}\hat{X}_k]_{\text{pos}}}{[\mathbf{D}\hat{X}_k]_{\text{pos}}} = \frac{\text{sign}(\mathbf{D}\hat{X}_k|_{\text{pos}})}{|\mathbf{D}\hat{X}_k|_{\text{pos}}} = \frac{1}{|\mathbf{D}\hat{X}_k|_{\text{pos}}}$$

אם נחשוב על מה שקיבלנו, הרי שלו יכולנו לקבוע את מטריצת המשקלות לא לפי \underline{Y} אלא לפי \underline{X} היינו צפויים לקבל תוצאה טובה יותר. לכן, כתחליף לחישוב \mathbf{W} מבוסס \mathbf{Y} , מציעה התוצאה שהתקבלה שנחשב את \mathbf{W} בכל איטרציה על סמך תוצאת האיטרציה הקודמת \hat{X}_k , וזה קורב יותר לשימוש ב- \underline{X} המקורית.

רעיון מעניין הוא הרעיון הבא: מדוע לא לעשות שימוש בפונקציות שאפילו יורדות עבור ערכים גבוהים. למשל, נציע את הפונקציה הבאה:

$$\rho(x) = x^2 \cdot \exp\{-\alpha x^2\}$$

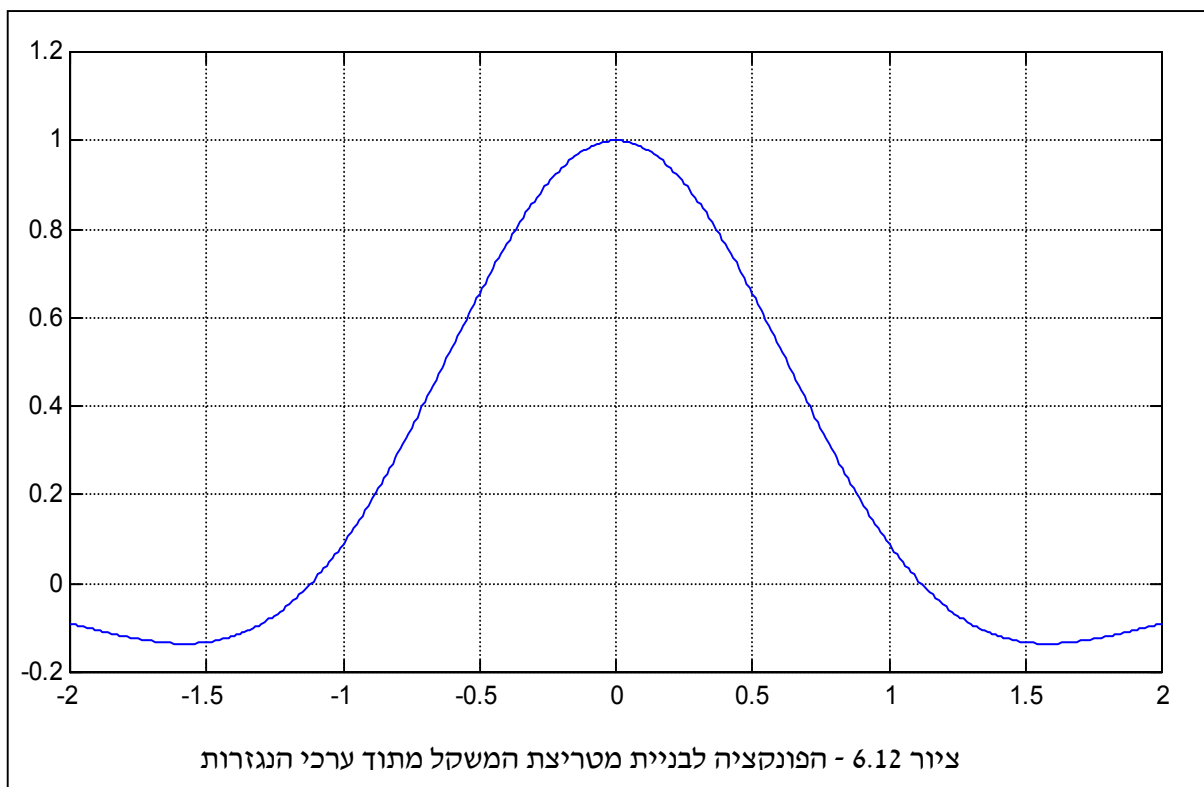
בחירה זו תתנהג כפונקציה ריבועית לערכים נמוכים, אך לאחר עליה מסוימת תתקבל ירידה לאפס יעבור ערכים גבוהים. ציור 6.11 מראה את הפונקציה הריבועית, פונקציית Huber-Markov, ובחירת פונקציה חדשה זו. בכל פונקציה שעולה מונוטונית אנו גובים מחיר שולי (שולי - מתייחס לנגזרת הפונקציה) חיובי.



עבור פונקציות יורדות נעודד היווצרם של שפות אם נגזרתם גדולה. עבור הבחירה הנ"ל, המשקל שמקביל לשימוש יהיה:

$$w(x) = \frac{\rho'(x)}{x} = 2 \exp(-\alpha x^2) [1 - \alpha x^2]$$

עבור הגרף הנ"ל מתקבל כי חישוב במשקל נעשה לפי הגרף הבא, המסגיר כי רעיון בחירת ערכים שליליים אינו נטול בסיס:



6.6 בעיות היפוך אחרות והקשר לבעיית השחזור הקלאסית

אם נסכם את שראינו עד כה, הרי שעבור בעיות בעלות המבנה $\underline{Y} = \mathbf{H}\underline{X} + \underline{N}$, בהן \mathbf{H} מערכת ליניארית כלשהי ידועה, \underline{X} תמונה רצויה שלא ניתנת למדידה, \underline{Y} תמונה אחרת שנמדד במקומה, ו- N - וקטור רעש אקראי גאוסני לבן בעל עוצמה σ_N , הצעת שחזור לא רעה ל- \underline{X} היא מינימיזציה של הביטוי:

$$f(\underline{X}) = [\underline{Y} - \mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{H}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{D}\underline{X}]$$

כאשר D אופרטור גזירה ליניארי כלשהו, הפונקציה $\rho(x)$ היא פונקצית שערך רובסטית שנבחרת מראש, ו- σ_D פרמטר השולט בעוצמת החלקות שנרצה לכפות.

עד כה השתמשנו בתוצאה הנ"ל לשם פתרונה של בעיית השחזור הקלאסית של בניית שחזור לתמונה מטושטשת ורועשת. נשאלת השאלה - האם ניתן לעשות שימוש בתוצאה הנ"ל לבעיות שחזור אחרות? התשובה היא כן, ואנו נציג מספר בעיות כאלה.

האפשרות הראשונה היא ניקוי תמונה מרעש. במקרה זה, המודל נותר כמות שהוא למעט העובדה ש- $\mathbf{H} = \mathbf{I}$. לכן, במקרה זה, עלינו להביא למינימום את

$$f(\underline{X}) = [\underline{Y} - \underline{X}]^T [\underline{Y} - \underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\mathbf{D}\underline{X}]$$

וכמובן התהליך שמתקבל פשוט יותר בהשוואה לזה המקורי (כיוון שאנו פטורים מהפעלת \mathbf{H} בתהליך האיטרטיבי).

אפשרות שניה היא הגדלת תמונה. בהינתן תמונה \underline{Y} , נשאל כיצד להגדילה פי r בכל ציר. אפשרות אחת טריוויאלית היא שכפול פיקסלים, כך שכל פיקסל יוחף בבלוק של r על r פיקסלים זהים לו. גישה זו מניבה תמונה באיכות ירודה למדי. בפרק 3 בו דנו בשחזור מדגימה תיארונו הגדלה ע"י אינטרפולציה ביליניארית. התוצאה בגישה זו טובה יותר, אך עדיין, באזור שפה בתמונה נקבל מריחה לא נעימה של המעבר, ושוב תתקבל תמונה באיכות ירודה.

אפשרות חדשה היא להתייחס לבעיית ההגדלה כבעיית שחזור. אנו נניח כי הייתה תמונה גדולה ויפה \underline{X} אשר עברה תהליך של טשטוש \mathbf{H} ודילול S לקבלת התוצאה הירודה \underline{Y} . לכן, הבעיה אותה נרצה לפתור תהיה מינימיזציה של:

$$f(\underline{X}) = [\underline{Y} - \mathbf{S}\mathbf{H}\underline{X}]^T [\underline{Y} - \mathbf{S}\mathbf{H}\underline{X}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\underline{D}\underline{X}]$$

באופן זה עשינו שימוש בעובדה שדילול היא פעולה ליניארית הניתנת לתיאור כמכפלת מטריצה בוקטור. מטריצה זו מלבנית ולא ריבועית - אם התמונה בכניסתה היא 100 על 100, ובמוצאה 25 על 25, גודל המטריצה יהיה 625 שורות על 10,000 עמודות. פעולת הדילול היא פעולה ספרבילית וניתן לפיכך לכותבה כמכפלת קרונקר בין שתי פעולות חד-מימדיות.

באשר לדרך בחירת האופרטורים \mathbf{H} ו- \mathbf{S} להגדלת תמונה, \mathbf{S} ייקבע ישירות לפי יחס ההגדלה אותו רצוננו להשיג. \mathbf{H} יכול להיבחר כ- \mathbf{I} אם אין רצוננו באפקט של חידוד. בחירת קיומו של טשטוש תגרום לתוצאה להיות חדה יותר. במקרים רבים ניתן לקבוע מדוע התמונה קטנה מידי מתוך שיקולים פיזיקליים, ואז \mathbf{H} יכולה להיבחר בהתאמה לדרך היווצרה של התמונה.

בעיה אחרת לגמרי אשר גם היא נופלת לקטגוריה של משפחת הבעיות הנ"ל היא בעיית הסופר-רזולוציה. נתונה לנו מצלמה ירודה באיכותה ורצוננו לצלם יעד כלשהו. היעד מיוצג כתמונה אידיאלית \underline{X} , ואילו מה שהמצלמה מניבה היא התמונה \underline{X} לאחר סידרת פעולות מזיקות - הזזה, טשטוש, דילול, ולבסוף רעש. במקום לצלם פעם אחת אנו נצלם M פעמים, בהזזות שונות ולכן נקבל:

$$\begin{aligned} \underline{Y}_1 &= \mathbf{S}\mathbf{H}\mathbf{F}_1 \underline{X} + \underline{N}_1 \\ \underline{Y}_2 &= \mathbf{S}\mathbf{H}\mathbf{F}_2 \underline{X} + \underline{N}_2 \\ &\vdots \\ \underline{Y}_M &= \mathbf{S}\mathbf{H}\mathbf{F}_M \underline{X} + \underline{N}_M \end{aligned}$$

הזזת תמונה (כל עיוות גיאומטרי שהוא!) ניתנת לתיאור כפעולה ליניארית ולכן כמטריצה. זאת כיוון שאם ניקח צירוף ליניארי של שתי תמונות ונזיז את שילובן, יהיה זה כמו להזיז כל אחת מהן ואז לצרפן ליניארית.

ב- M הפעמים בהם צילמנו עשינו שימוש באותה מצלמה ירודה באיכותה, ולכן \mathbf{H} ו- \mathbf{S} זהים לכל התמונות שנוצרו. לעומת זאת, תוך כדי צילום זזנו ולכן האוריינטציה הגיאומטרית בין התמונות השתנתה. זו מיוצגת כאן ע"י האופרטורים הליניאריים \mathbf{F} . כך התקבל סט תמונות \underline{Y} . נוכל לקחת את כל השוואות הנ"ל ולצרפן לאחת לקבלת:

$$\begin{bmatrix} \underline{Y}_1 \\ \underline{Y}_2 \\ \vdots \\ \underline{Y}_M \end{bmatrix} = \begin{bmatrix} \mathbf{S}\mathbf{H}\mathbf{F}_1 \\ \mathbf{S}\mathbf{H}\mathbf{F}_2 \\ \vdots \\ \mathbf{S}\mathbf{H}\mathbf{F}_M \end{bmatrix} \underline{X} + \begin{bmatrix} \underline{N}_1 \\ \underline{N}_2 \\ \vdots \\ \underline{N}_M \end{bmatrix} = \tilde{\mathbf{H}} \underline{X} + \tilde{\underline{N}}$$

ושוב, אנו רואים אותו מבנה. לכן, שחזור תמונת הסופר-רזולוציה ייעשה ע"י מינימיזציה של הביטוי

$$f(\mathbf{X}) = [\underline{\mathbf{Y}} - \tilde{\mathbf{H}}\underline{\mathbf{X}}]^T [\underline{\mathbf{Y}} - \tilde{\mathbf{H}}\underline{\mathbf{X}}] + \frac{\sigma_N^2}{\sigma_D^2} \rho[\underline{\mathbf{D}}\underline{\mathbf{X}}] = \sum_{k=1}^M \|\underline{\mathbf{Y}} - \mathbf{SHF}_k \underline{\mathbf{X}}\|^2 + \frac{\sigma_N^2}{\sigma_D^2} \rho[\underline{\mathbf{D}}\underline{\mathbf{X}}]$$

בדומה לנאמר עבור הגדלת תמונה, על מנת להפעיל תהליך סופר-רזולוציה, עלינו לדעת את האופרטורים המעורבים. ידיעת אופרטורי התנועה כרוכה שתהליך שערך תנועה בין התמונות הנתונות - נושא שיידון בפרקים מתקדמים יותר. כמקודם, \mathbf{S} יכול להיקבע ישירות מתוך יחס ההגדלה הרצוי. \mathbf{H} ייקבע ע"י הבנת מאפייני המצלמה בה נעשה שימוש, או פשוט בבחירה שרירותית לפי מידת החידוד שרצוננו להשיג בתוצאה הסופית.