www.hapetek.co.il

עיבוד וניתוח תמונות

046200

סיכום הקורס

עדכון אחרון : 7/2/2009

תוכן עניינים

1. אותות ומערכות בדו-מימד

1.1. הקדמה

אנחנו נדבר על תמונות, שהן פונקציות דו-מימדיות. תמונות אלו יהוו קלט למערכות מתאימות, המבצעות פעולה על פונקציה דו מימדית $f(x,y)$, ופולטות (תמונה) פונקציה דו מימדית אחרת, $g(x,y)$.

$$f(x,y) \longrightarrow \boxed{H} \longrightarrow g(x,y)$$

1.2. לינאריות

מערכת דו-מימדית H תיקרא לינארית כאשר מתקיימות שתי התכונות הבאות לכל $\alpha, f(x,y), g(x,y)$:

1. כפל בסקלר : $H\{\alpha f(x,y)\} = \alpha H\{f(x,y)\}$
2. אדיטיביות : $H\{f(x,y) + g(x,y)\} = H\{f(x,y)\} + H\{g(x,y)\}$

דוגמא לפונקציה לינארית : הזזה.
דוגמא לפונקציה לא לינארית : פעולת סף (Threshold).

1.3. זכרון

במערכת חסרות זיכרון (מערכות פעולת נקודה), הערך של g תלוי רק בערך של f באותה נקודה (x,y).
דוגמה למערכות שאינן חסרות זיכרון : מיצוע, גרדיאנט.

1.4. קביעות במקום

מערכות קבועות במקום (Space-Invariant)
זוהי מערכת שהיא קומוטטיבית עם הזזה, כלומר :

$$H\{f(x-\alpha, y-\beta)\} = g(x-\alpha, y-\beta)$$

לכל x, y, α, β, לכל $f(x,y)$.
דוגמא למערכת קבועה במקום : פעולת הנגזרת.

דוגמא למערכת שאינה קבועה במקום : מערכת שמאפסת את כל הפיקסלים מחוץ לחלון קבוע. כמובן שהזזה של תמונה שעברה במערכת כזו לא תראה כמו התמונה המקורית שעברה באותה המערכת (ניתן לחשוב על החלון כהתבוננות מתוך מכונית נוסעת, כאשר הנוף הוא התמונה – בנסיעה, מבחינתנו הנוף נע, ואנחנו רואים רק מה שעובר דרך החלון. במשך הנסיעה התמונה משתנה).

1.5. תגובה להלם

ראשית, נגדיר פונקצית דלתא דו מימדית של דירק. זוהי פונקציה, אשר לכל $f(x,y)$ רציפות, מקיימת :

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,\delta(x-\alpha, y-\beta)\,dxdy = f(\alpha,\beta)$$

נביט על פונקצית ההלם כתמונה, ולכן פונקצית הדלתא הדו-מימדית $\delta(x-\alpha, y-\beta)$ היא בעצם פיקסל אחד לבן בנקודה (α,β), וזאת על-גבי תמונה שחורה.

תגובה להלם של מערכת לינארית בדו-מימד :

$$\delta(x-\alpha, y-\beta) \longrightarrow \boxed{H} \longrightarrow h(x, y; \alpha, \beta)$$

הפונקציה $h(x,y;\alpha,\beta)$ תלויה בארבעה משתנים, שהם הנקודה במרחב והנקודה בה פעלה פונקצית הדלתא. ניתן להסתכל על התגובה להלם h כתמונה להלם, $h(x,y)$ כאשר ההלם ניתן בנקודה (α,β). הם ההלם יינתן בנקודה אחרת, נאמר (α_2,β_2), המערכת לא בהכרח תוציא לנו כפלט את אותה התמונה $h(x,y)$.

1.6. אינטגרל הסופרפוזיציה

אם המערכת H היא לינארית, אזי תגובת המערכת $g(x,y)$ לקלט $f(x,y)$ הוא

$$H\{f(x,y)\}=g(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\alpha,\beta)h(x,y;\alpha,\beta)\,d\alpha\, d\beta$$

זהו אינטגרל הסופרפוזיציה.

1.7. מערכות לינאריות וקבועות בזמן (LSI)

במקרה שמערכת לינארית היא גם קבועה בזמן, אזי מתקיים:

$$h(x,y;\alpha,\beta)=h(x-\alpha,y-\beta)$$

כלומר התגובה להלם תהיה תמונה שערכיה תלויים במרחקם מהמקום בו ניתן ההלם.
מערכות כאלו נקראות Linear Space Invariant (באנלוגיה לאותות שמע חד מימדיים, LTI)

במקרה זה, אינטרל הסופרפוזיציה ייראה כך:

$$H\{f(x,y)\}=g(x,y)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\alpha,\beta)h(x-\alpha,y-\beta)\,d\alpha\, d\beta$$

וזהו אינטגרל הקונבולוציה.

דוגמא למערכת לינארית, שאינה קבועה במקום: פעולת התמקדות (Zoom). בפעולה זו, נקודות שונות עוברות טיפול שונה, בהתאם למיקומם בתמונה המקורית. לדוגמא – פיקסל במרכז לא עובר שינוי, אבל פיקסל בקצה התמונה יכול להעלם, או להשתנות בצורה ניכרת.
כאשר נכניס הלם תמונה שנמצא קצת ימינה מהאמצע, הוא יוזז ימינה, כאשר נכניס הלם תמונה שנמצא קצת שמאלה מהאמצע, הוא יוזז שמאלה – כל נקודה עוברת טיפול שונה.

1.8. פרידות (Separability)

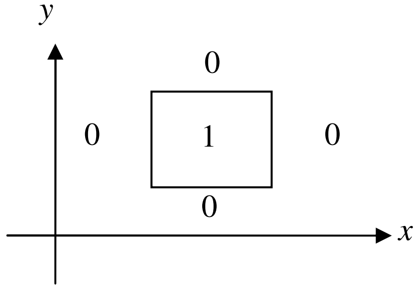
פונקציה ספרבילית היא פונקציה דו-מימדית המקיימת:

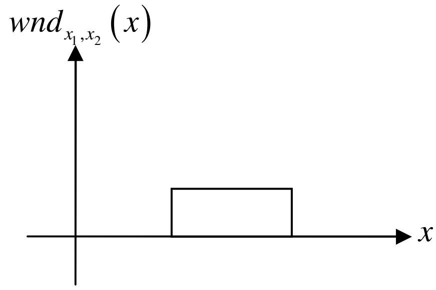
$$f(x,y)=f_1(x)\cdot f_2(y)$$

לדוגמה, פונקצית חלון דו-מימדית היא פונקצית ספרבילית המורכבת ממכפלת שתי פונקציות חלון חד-מימדיות:

$$wnd(x,y)=wnd_{x_1,x_2}(x)\cdot wnd_{y_1,y_2}(y)=\left(u(x-x_1)-u(x-x_2)\right)\left(u(y-y_1)-u(y-y_2)\right)$$

כאשר כאן wnd הוא קיצור לפונקצית חלון.





הגדרה: מערכת H תקרא ספרבילית אם תגובת ההלם שלה, h, ספרבילית.
במקרה זה, מתקיים:

$$h(x,y;\alpha,\beta)=h_1(x;\alpha)\cdot h_2(y;\beta)$$

במקרה כזה, אינטגרל הסופרפוזיציה ייראה כך:

$$H\{f(x,y)\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\alpha,\beta)h(x,y;\alpha,\beta)\,d\alpha\, d\beta=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(\alpha,\beta)h_1(x;\alpha)h_2(y;\beta)\,d\alpha\, d\beta$$

$$=\int_{-\infty}^{\infty}h_2(y;\beta)\left(\int_{-\infty}^{\infty}f(\alpha,\beta)h_1(x;\alpha)\,d\alpha\right)d\beta$$

כלומר, ניתן לבצע את האינטגרל בשני חלקים, כל פעם על משתנה אינטגרציה אחד (ניתן קודם לפעול על שורות, ואז לפעול על עמודות, או להפך).

1.9. התמרות פורייה דו-מימדיות

1.9.1. התמרת פורייה רציפה לאותות רציפים

הגדרת ההתמרה:

$$F\left(u,v\right)=\mathcal{F}\left\{f\left(x,y\right)\right\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f\left(x,y\right)e^{-2\pi j\left(ux+vy\right)}dxdy$$

וההתמרה ההפוכה:

$$f\left(x,y\right)=\mathcal{F}^{-1}\left\{F\left(u,v\right)\right\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}F\left(u,v\right)e^{2\pi j\left(ux+vy\right)}dudv$$

מהגדרת ההתמרה ההפוכה, ומזכרון נושן עבור התמרות חד-מימדיות, ניתן לומר שכל תמונה היא לא יותר מאשר סופרפוזיציה של אינסוף תמונות בסיס מהצורה $e^{2\pi j\left(ux+vy\right)}$, כאשר לכל תמונה כזו יש משקל $F\left(u,v\right)$.

1.9.2. תגובת התדר

תגובת התדר של מערכת לינארית וקבועה בזמן H, אשר תגובת ההלם שלה היא $h\left(x,y\right)$ היא

$$H\left(u,v\right)=\mathcal{F}\left\{h\left(x,y\right)\right\}$$

1.9.3. תכונות ההתמרה

- יחידות;
- לינאריות;
- היפוך צירים:

$$\mathcal{F}\left\{f^{*}\left(x,y\right)\right\}=F^{*}\left(-u,-v\right)$$

$$\mathcal{F}\left\{f\left(-x,-y\right)\right\}=F\left(-u,-v\right)$$

- ספרביליות ההתמרה: לכל אות $f\left(x,y\right)$, ניתן לבצע אינטגרל על מימד אחד, ואז על המימד השני:

$$F\left(u,v\right)=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f\left(x,y\right)e^{-2\pi j\left(ux+vy\right)}dxdy=\int_{-\infty}^{\infty}\left(\int_{-\infty}^{\infty}f\left(x,y\right)e^{-2\pi jux}dx\right)e^{-2\pi jvy}dy$$

- מתיחה וכיווץ (Scaling)

$$\mathcal{F}\left\{f\left(ax,by\right)\right\}=\frac{1}{\left|ab\right|}F\left(\frac{u}{a},\frac{v}{b}\right)$$

אינטואיטיבית, כאשר מכווצים תמונה, אנו מקטינים את אורך הגל, ולכן התדר של התמונה גדל – כלומר בתחום התדר נראה אנרגיות בתדרים גבוהים יותר. רעיון זה דומה לעיקרון אי הוודאות – ככל שמכווצים את המידע במקום, הוא מתפרס על תחום גדול יותר בתדר.

- הזזה במקום והזזה בתדר (אפנון):

$$\mathcal{F}\left\{f\left(x-x_0,y-y_0\right)\right\}=e^{-2\pi j\left(ux_0+vy_0\right)}F\left(u,v\right)$$

$$\mathcal{F}^{-1}\left\{F\left(u-u_0,v-v_0\right)\right\}=e^{2\pi j\left(xu_0+yv_0\right)}f\left(x,y\right)$$

- משפט הקונבולוציה

$$\mathcal{F}\left\{f_1\left(x,y\right)*f_2\left(x,y\right)\right\}=F_1\left(u,v\right)\cdot F_2\left(u,v\right)$$

$$\mathcal{F}\left\{f_1\left(x,y\right)\cdot f_2\left(x,y\right)\right\}=F_1\left(u,v\right)*F_2\left(u,v\right)$$

כאשר כאן פעולה הקונבולוציה היא פעולה דו-מימדית.

- משפט פרסבל

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|f\left(x,y\right)\right|^2dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left|F\left(u,v\right)\right|^2dudv$$

התמרה לאות ספרבילי

כאשר מבצעים התמרה לאות ספרבילי, ניתן לבצע התמרה לכל מימד בנפרד. כלומר, כאשר

$$f(x,y)=f_1(x)\cdot f_2(y)$$

אזי

$$\mathcal{F}^{2D}\left\{f(x,y)\right\}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f(x,y)e^{-2\pi j(ux+vy)}dxdy=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}f_1(x)f_2(y)e^{-2\pi jux}e^{-2\pi jvy}dxdy$$

$$=\int_{-\infty}^{\infty}f_1(x)e^{-2\pi jux}dx\int_{-\infty}^{\infty}f_2(y)e^{-2\pi jvy}dy=\mathcal{F}^{1D}\left\{f_1(x)\right\}\cdot\mathcal{F}^{1D}\left\{f_2(y)\right\}$$

1.10. אותות ומערכות ליניאריות דיסקרטיות בדו-מימד

תמונה דו מימדית דיסקרטית תסומן $f[m,n]$, כאשר $m,n\in\mathbb{Z}$. m הוא אינדקס השורות; n הוא אינדקס העמודות.

את האותות הדיסקרטים בחד-מימד ראינו כוקטורים אינסופיים (סדרות של מספרים). אותות דו-מימדיים (תמונות) ניתנים להתסכלות כמטריצות בגודל אינסופי של מספרים.

1.10.1. דלתא של קרוניקר בדו-מימד

הגדרה:

$$\delta[m,n]=\delta[m]\cdot\delta[n]=\begin{cases}1, & m=n=0\\ 0, & else\end{cases}$$

התגובה להלם של מערכת לינארית דיסקרטית:

$$\delta[m-k,n-r]\longrightarrow\boxed{H}\longrightarrow h[m,n;k,r]$$

1.11. סכום הסופרפוזיציה

כאשר מדברים על מערכת דיסקרטית H לינארית, באופן מקביל לאינטגרל הסופרפוזיציה, המוצא הוא סופרפוזיציה של התגובה להלם עם הכניסה:

$$g[m,n]=H\left\{f[m,n]\right\}=\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty}f[k,r]h[m,n;k,r]$$

כמו במקרה הרציף (אינטגרל), ניתן להתסכל על המוצא $g[m,n]$ כעל סכום (הסכימה של אברי המטריצה f) משוקלל (ההכפלה באברי מטריצה h). לכל נקודה (פיקסל) בפלט $g[m,n]$ יש מטריצת משקלים מתאימה.

כאשר H היא גם מערכת קבועה במקום, מתקיים

$$h[m,n;k,r]=h[m-k,n-r]$$

ואז, כמו במקרה הרציף, סכום הסופרפוזיציה מתנוון לסכום הקונבולוציה:

$$g[m,n]=H\left\{f[m,n]\right\}=\sum_{-\infty}^{\infty}\sum_{-\infty}^{\infty}f[k,r]h[m-k,n-r]$$

מערכת (פעולה) לינארית ספרבילית היא מערכת מהצורה

$$h[m,n;k,r]=h_1[m;k]\cdot h_2[n;r]$$

1.12. התמרות פורייה לאותות דיסקרטיים

הגדרה: התמרת פורייה של תמונה דו-מימדית דיסקרטית $f[m,n]$ (ההתמרה היא רציפה!):

$$F_D(\theta_1,\theta_2)=\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}f[m,n]e^{-j(\theta_1 m+\theta_2 n)}$$

כאשר ההתמרה היא מחזורית- 2π ב θ_1 וב θ_2, ולכן מספיק להתסכל על התחום הדו-מימדי

$$\theta_1 \in [-\pi,\pi] \times \theta_2 \in [-\pi,\pi]$$

להתמרת פורייה זו יש כמובן אינסוף ערכים (רציפים), בהתאמה לאינסוף הערכים שבמקום (אינסוף איברי $f[m,n]$).

כעת נסתכל כל תמונה המוגברת בגודלה: תמונה סופית עם M שורות ו N עמודות. במקרה זה כניראה לא צריך אינסוף איברים במישור התדר כדי לתאר את התמונה, כי יש רק $M \cdot N$ דרגות חופש לתמונה, הניתנים כקלט להתמרת הפורייה.
אנחנו נדגום את התמרת הפורייה בתדרים הבאים:

$$\begin{cases} \theta_1 = \frac{2\pi}{M}k, & k = 0,1,2,...,M-1 \\ \theta_2 = \frac{2\pi}{N}r, & r = 0,1,2,...,N-1 \end{cases}$$

כלומר, אנחנו יכולים להתסכל על תמונה (מטריצה) בגודל סופי $M \times N$, ועל התמרת הפורייה שלה, שגם תהיה מטריצה סופית, באותו גודל.
למי שעדיין לא שם לב, זוהי התמרת DFT בדו-מימד. הנה ההגדרה:

$$2D-DFT\{f[m,n]\} = F[k,r] = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} f[m,n] e^{-2\pi j\left(\frac{mk}{M}+\frac{nr}{N}\right)}$$

וההתמרה ההפוכה:

$$2D-IDFT\{F[k,r]\} = f[m,n] = \frac{1}{MN}\sum_{k=-\infty}^{\infty}\sum_{r=-\infty}^{\infty} F[k,r] e^{2\pi j\left(\frac{mk}{M}+\frac{nr}{N}\right)}$$

בדומה למקרה של התמרת פורייה רציפה, אנו רואים כי מהגדרת התמרה ההפוכה, ניתן לומר שכל תמונה היא לא יותר מאשר סופרפוזיציה של אינסוף תמונות בסיס מהצורה $e^{2\pi j\left(\frac{mk}{M}+\frac{nr}{N}\right)}$ $e^{2\pi j(ux+vy)}$, כאשר לכל תמונה כזו יש משקל $F[k,r]$.

נשים לב כי פעולת ההתמרה היא ספרבילית:

$$F[k,r] = \sum_{n=-\infty}^{\infty}\left(\sum_{m=-\infty}^{\infty} f[m,n] e^{-2\pi j\frac{mk}{M}}\right) e^{-2\pi j\frac{nr}{N}}$$

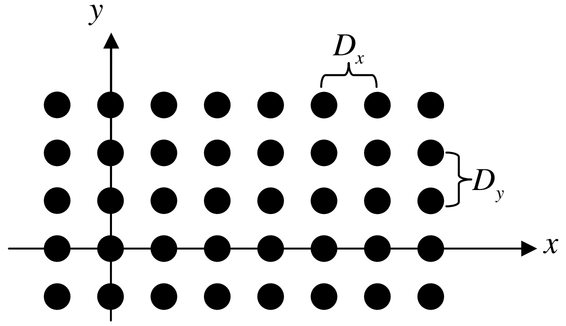
ולכן, ניתן לבצע התמרת DFT (חד-מימדית) על השורות (עמודות), ואז לעשות התמרת DFT על העמדות (שורות).
כמובן שאת התמרות ה DFT החד-מימדיות ניתן לבצע בעזרת אלגוריתם FFT היעיל.

1.13. דגימה ושחזור

אנחנו רוצים לקחת אות רציף $f(x,y)$, $x,y \in \mathbb{R}$ (ל x,y יחידות של מקום, כלומר מטר, מילי-מטר וכו'), ואת האות הרציף לדגום כדי לקבל אות בדיד $f[m,n]$, $m,n \in \mathbb{N}$ הם אינדקים, נטולי יחידות, של שורה ועמודה.

נגדיר מסרק הלמים רציף דו-מימדי:

$$\text{comb}(x,y;D_x,D_y) = \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\delta(x-mD_x, y-nD_y)$$



זהו סריג אינסופי של פונקציות דלתא. נכפיל את האות $f(x,y)$ בפונקציית המסרק, וזהו תהליך הדגימה שלנו:

$$\tilde{f}(x,y) = f(x,y)\cdot \mathrm{comb}(x,y;D_x,D_y) = f(x,y)\sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty}\delta(x-mD_x, y-nD_y)$$
$$= \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} f(x,y)\delta(x-mD_x, y-nD_y)$$
$$= \sum_{n=-\infty}^{\infty}\sum_{m=-\infty}^{\infty} f[m,n]\delta(x-mD_x, y-nD_y)$$

קיבלנו תמונה רציפה $\tilde{f}(x,y)$, שהיא סכום של פונקציות דלתא, המוכפלות במשקלים שונים $f[m,n]$ - משקלים אלו הם הפונקציה הבדידה. כלומר:

$$f[m,n] = f(mD_x, nD_y)$$

התמרת הפורייה (הרציפה) של מסרק ההלמים (הרציף):

$$\mathcal{F}\left\{\mathrm{comb}(x,y;D_x,D_y)\right\} = \frac{1}{D_xD_y}\mathrm{comb}\left(u,v;\frac{1}{D_x},\frac{1}{D_y}\right)$$

התמרת פורייה של האות הדגום $\tilde{f}(x,y)$:

$$\tilde{F}(u,v) = \mathcal{F}\left\{\tilde{f}(x,y)\right\} = \mathcal{F}\left\{f(x,y)\cdot\mathrm{comb}(x,y;D_x,D_y)\right\}$$
$$= F(u,v) * \frac{1}{D_xD_y}\mathrm{comb}\left(u,v;\frac{1}{D_x},\frac{1}{D_y}\right)$$

כלומר אינסוף שכפולים של התמרת האות המקורי $f(x,y)$, גם בתדרים מאוד גבוהים (תופעה לא מפתיעה, מכיוון שכפלנו את $f(x,y)$ במסרק הלמים, מה שגורר שינויים מאוד חדים (דלתות), וכשיש שינויים מהירים כאלו במישור המקום, יש תדרים מאוד גבוהים במישור התדר).

שחזור הדגימה הוא לא יותר מאשר הצגת המטריצה ($f[m,n]$ - המספרים אותם שמרנו במחשב) על מסך.

1.13.1.תופעת ההתחזות

תופעת ההתחזות (Aliasing) יכולה להתקיים גם כאן, כמו שראינו בדגימה חד-מימדית. כדי למנוע תופעה זו, נרצה שלתמונה יהיה תמך סופי במישור התדר, כלומר נרצה להעבירה דרך מסנן מעביר נמוכים (LPF) בעולם הרציף (כך שהתמונה לא תהיה חדה לגמרי). בצורה מעשית, האופטיקה של המצלמה המצלמת את התמונה מהווה LPF, ולכן התמונה המתקבלת היא בעלת תמך סופי בתדר.

עלינו לדאוג שהמרחק בין השכפולים, $\frac{1}{D_x}$ ו $\frac{1}{D_y}$, הנוצרים במישור התדר יהיה מספיק גדול, כך שלא תהיה חפיפה בין השכפולים (רפליקות). אם u_0 ו v_0 הם הערכים המקסימליים של התמונה בתדר, יש לעמוד בתנאי נייקויסט:

$$\begin{cases} \dfrac{1}{D_x} > 2u_0 \\ \dfrac{1}{D_y} > 2v_0 \end{cases}$$

שחזור אידיאלי של האות המקורי:

$$\hat{f}(x,y) = \mathcal{F}^{-1}\left\{H(u,v)\cdot\tilde{F}(u,v)\right\}$$

ובמישור המקום, הפעולה המתאימה היא קונבולוציה:

$$\hat{f}(x,y) = \tilde{f}(x,y) * h(x,y)$$

אם $H(u,v)$ היא פונקצית חלון דו-מימדית, אזי תגובת ההלם התמאימה היא

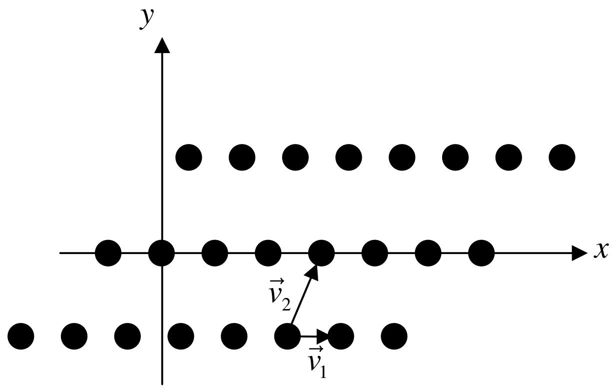
$$h(x,y) = \mathrm{sinc}\left(\frac{x}{D_x},\frac{y}{D_y}\right)$$

כמובן ששיחזור זה אינו מעשי, כי השחזור ע״י תצוגה על גבי מסך הוא תצוגה של מלבן אחד לכל פיקסל – לא פונקצית sinc אינסופית. כלומר, בפועל $h(x,y)$ הוא מלבן, מה שמקביל לשחזור ZOH-Zero Order Hold המוכר.

נשים לב שפעולות השחזור והדגימה המדוברות כאן לא רלוונטי רק לתמונות רציפות הנדגמות, אלא גם לדגימה של תמונה בדידי קיימת לתמונה בעלת מימדים יותר קטנים. לדוגמא, אם יש לנו תמונה בגודל 1000×1000, נוכל להקטין אותה ע״י דגימה של כל פיקסל עשירי, ולקבל תמונה חדשה בגודל 100×100. גם במקרה זה עלולים להתקל בתופעת ההתחזות (Aliasing), ולכן יש לטשטש את התמונה לפני התהליך – כלומר להעבירה במסנן LPF.
באופן אנלוגי – ניתן לשחזר תמונה בגודל 100×100 לתמונה בגודל 1000×1000. ישנן שיטות שונותלבצע זאת, לדוגמה אינטרפולציה בי-לינארית (Bilinear Interpolation) או אינטרפולצית השכן הקרוב (Nearest Neighbour Interpolation).

1.1.13.2 דגימה מוכללת

עד עכשיו דגמנו בעזרת סריג מלבני, כלומר פונקצית המסרק שלנו היתה שכפולים לאורך ציר x ולאורך ציר y במרווחים שווים. ניתן לדגום אות בסריג כללי יותר.



סריג זה עדיין מחזורי, אך לא באופן ספרבילי ב x וב y, כמו במקרה הפרטי של סריג מלבני.
בסריג זה, נקודות הדגימה הן קומבינציה לינארית של שני הוקטורים המגדירים את הסריג:

$$\underline{v}_1=\begin{bmatrix}v_{1,x}\\ v_{1,y}\end{bmatrix},\underline{v}_2=\begin{bmatrix}v_{2,x}\\ v_{2,y}\end{bmatrix}$$

קואורדיטה של נקודת דגימה מחושבת כך:

$$\begin{bmatrix}s_1\\ s_2\end{bmatrix}=\begin{bmatrix}\underline{v}_1 & \underline{v}_2\end{bmatrix}\begin{bmatrix}m\\ n\end{bmatrix}=\begin{bmatrix}v_{1,x} & v_{2,x}\\ v_{1,y} & v_{2,y}\end{bmatrix}\begin{bmatrix}m\\ n\end{bmatrix}=\underline{\underline{V}}\cdot\underline{n}$$

במקרה הפרטי של סריג מלבני, ניקח

$$\underline{v}_1=\begin{bmatrix}D_x\\ 0\end{bmatrix},\underline{v}_2=\begin{bmatrix}0\\ D_y\end{bmatrix}$$

ונקבל שנקודות הדגימה הן, כצפוי:

$$\begin{bmatrix}s_1\\ s_2\end{bmatrix}=\begin{bmatrix}\underline{v}_1 & \underline{v}_2\end{bmatrix}\begin{bmatrix}m\\ n\end{bmatrix}=\begin{bmatrix}D_x & 0\\ 0 & D_y\end{bmatrix}\begin{bmatrix}m\\ n\end{bmatrix}=\begin{bmatrix}mD_x\\ nD_y\end{bmatrix}$$

יש לשים לב שצריך לקחת את $\underline{v}_1,\underline{v}_2$ בלתי תלויים לינארית (בת״ל), כי אחרת אנחנו נדגום על קו ישר, ולא על המישור.
מכיוון ש $\underline{v}_1,\underline{v}_2$ בת״ל, הרי ש $\det\underline{\underline{V}}\neq0$.
מסרק ההלמים הכללי המתאים לדגימה כללית נרשם בצורה הבאה:

$$\text{comb}\left(\underline{x};\underline{\underline{V}}\right)=\sum_{\underline{n}}\delta\left(\underline{x}-\underline{\underline{V}}\underline{n}\right)$$

כאשר $\underline{x}=\left(x,y\right)$ קואורדינטה במישור, $\underline{n}=\left(n,m\right)\in\mathbb{Z}^2$ וקטור של מספרים שלמים, ו $\underline{\underline{V}}$ המטריצה המייצגת את וקטורי הדגימה, כפי שראינו קודם.

התמרת פורייה של פונקצית מסרק ההלמים היא

$$\mathcal{F}\left\{\text{comb}\left(\underline{x};\underline{\underline{V}}\right)\right\}=\frac{1}{\left|\det\underline{\underline{V}}\right|}\text{comb}\left(\underline{u};\underline{\underline{V}}^{-T}\right)$$

כאשר הסימון A^{-T} מציין את $\left(\underline{\underline{A}}^{-1}\right)^T$, כלומר פעולת Transpose על המטריצה ההופכית של $\underline{\underline{A}}$.
בנוסף, הסימון $\underline{u}$ הוא וקטור הקואורדינאטות במישור התדר, כלומר $\underline{u}=\left(u,v\right)$.

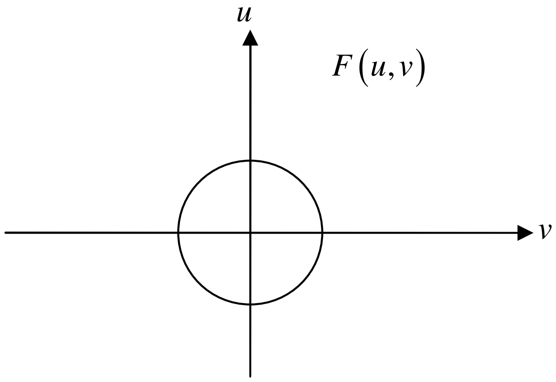
נמשיך ונשתמש במסרק הכללי. הפונקציה $\tilde{f}\left(\underline{x}\right)$ הדוגמת בעזרת המסרק הכללי את הפונקציה $f\left(\underline{x}\right)$ היא:

$$\tilde{f}\left(\underline{x}\right)=f\left(\underline{x}\right)\cdot\operatorname{comb}\left(\underline{x};\underline{\underline{V}}\right)$$

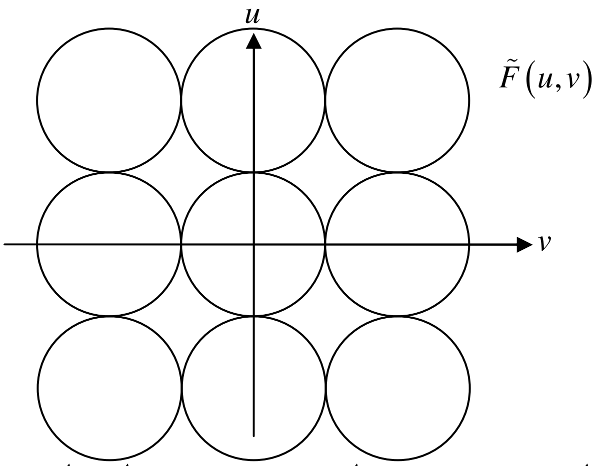
ובמישור התדר:

$$\tilde{F}\left(u,v\right)=\mathcal{F}\left\{\tilde{f}\left(\underline{x}\right)\right\}=F\left(u,v\right)*\frac{1}{\left|\det\underline{\underline{V}}\right|}\operatorname{comb}\left(\underline{u};\underline{\underline{V}}^{-T}\right)$$

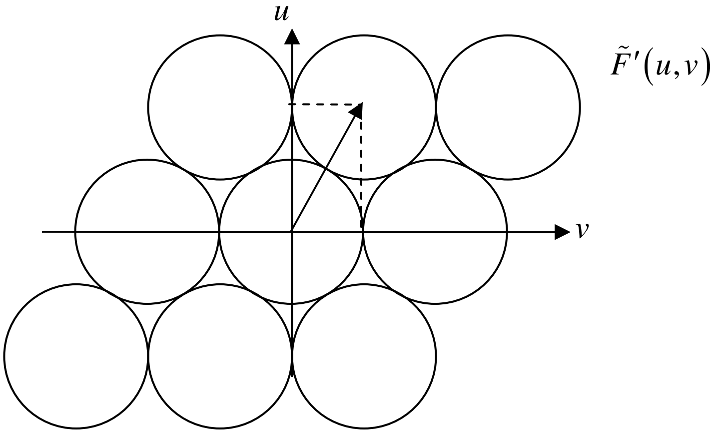
לעיתים נרצה לדגום תמונות אשר אין להם כיווניות מסויימת (איזוטרופיות), לדוגמא צילום של דשא, או חול. כלומר, אנו מסתכלים במישור התדר על התחום הבא:



נרצה לדגום כמה שפחות, ועדיין לא לקבל Aliasing. בדגימה המלבנית הרגילה, נקבל שכפולים של התמך המעגלי, משיקים זה לזה באופן הבא:



ומסתבר שניתן לצופף את השכפולים בתדר יותר ביעילות, וזאת ע"י סריג לא מלבני, אלא בעזרת סריג משושה:



כאשר התמונה בתדר יותר צפופה, המשמעות במישור המקום היא פחות דגימות. הוקטור המסומן מייצג את סכום הוקטורים בהם השתמשנו לצורך הדגימה.

שטח תא היחידה של סריג הדגימות מוגדר להיות

$$S_{lattice} = \left|\det \underline{\underline{V}}^{-T}\right| = \frac{1}{\left|\det \underline{\underline{V}}\right|}$$

ולכן צפיפות הדגימות הינה

$$D = \frac{1}{S_{lattice}} = \left|\det \underline{\underline{V}}\right|$$

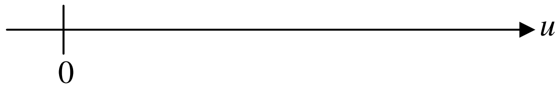
ניתן גם להתסכל על נצילות הסריג, כאשר נתון לנו השטח הסופי של תמך ההתמרה, $S_{support}$:

$$\eta = \frac{S_{support}}{S_{lattice}} = \frac{S_{support}}{\left|\det \underline{\underline{V}}^{-T}\right|}$$

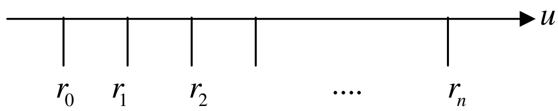
2. כימוי (Quantization)

ערך של פיקסל בתמונה מייצג כמות אור מסויימת שנפלה על הגלאי של המצלמה. ערכים אלו הם רציפים מבחינתנו (בעיקרון ערכים אלו אינם רציפים, מכיוון שלכל פוטון יש כמות מסויימת של אנרגיה, ולכן האנרגיה הפוגעת בגלאי כלשהו חייבת להיות כפולה של כמות האנרגיה של הפוטון – כלומר מספר שאינו רציף), אבל לא ניתן לאכסן בזיכרון מעשה ידי אדם אינפורמציה רציפה. לכן, יש לבצע עיגול של מספרים ממשיים למספרים בדידים. זוהי פעולת הקוונטיזציה – עיגול למספרים בדידים. למשל, נהוג לייצג פיקסלים בתמונת שחור/לבן (כאשר כל פיקסל מיוצג ע"י 8 ביטים) בעזרת הערכים הבדידים $0,1,2,3,...255$, כאשר כל מספר מייצג רמת בהירות, כאשר 0 מייצג בהירות מינימלית (שחור) ו 255 מייצג בהירות מקסימאלית (לבן).

נביט בפיקסל המקבל באופן התחלתי ערך רציף u כלשהו :

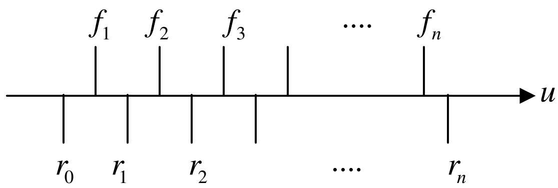


כעת, נחלק את תחום הערכים האפשריים לפיקסל לתחומים ע"י הצבת גבולות בערכים $\{r_i\}_{i=0}^{n}$. לגבולות אלו נקרא "רמות החלטה" :



כל ערך של פיקסל חדש יכול ליפול רק בתחום אחד בין שתי רמות החלטה.

כעת נוסיף "רמות ייצוג" – אילו הם מספרים $\{f_i\}_{i=1}^{n}$ אשר ייצגו פיקסלים הנופלים בכל תחום המוגדר ע"י שתי רמות החלטה :



ערכי הפיקסלים u הם הקלט של תהליך הכימוי, ואילו רמות הייצוג הם הפלט של תהליך הכימוי. לכל ערך רציף של פיקסל יותאם ערך בדיד אחד מתוך רמות הייצוג $\{f_i\}_{i=1}^{n}$. לדוגמה, אם נרצה לעגל למספרים שלמים, $\{f_i\}_{i=1}^{n}$ יהיו סדרת המספרים השלמים. רמות ההחלטה הם הספים שכנגדם נמדד ערך הכניסה (ערך הפיקסל הרציף) כדי להתאימו לתחום החלטה מסויים. לדוגמה, אם נרצה לעגל למספרים שלמים,רמות ההחלטה יהיו המספרים החצי שלמים

$$\{r_i\}_{i=0}^{n}:\ r_i = i + \frac{1}{2}, \qquad i \in \mathbb{Z}$$

וכך, כל המספרים הנופלים, לדוגמא, בין 6.5 לבין 7.5, ייקבלו את הערך 7.

2.1. היסטוגרמה

נביט בתמונת גווני אפור (שחור/לבן), בה כל פיקסל u מיוצג ע"י מספר בין 0 ל 255 (ישנם 256 גווני אפור).
נבצע סריקה על כל התמונה, ונשמור וקטור $\underline{v}$, בעל 256 רכיבים, באופן הבא : רכיב i של הוקטור יכיל את מספר המופעים של של הגוון i. השרטוט של רכיבי $\underline{v}$ תתן את פילוג הגוונים בתמונה. פילוג זה נקרא היסטורגרמה.

2.2. קוונטייזר אחיד

זהו קוונטייזר אשר המרווחים בין רמות ההחלטה ובין רמות הייצוג קבועים ושווים :

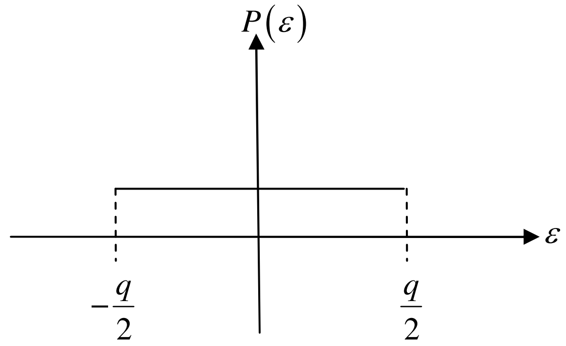
$$r_k - r_{k-1} = f_k - f_{k-1} = q$$

כאן q נקרא "צעד הקוונטיזציה" (או "מרווח הקוונטיזציה"). לדוגמה, כאשר עיגלנו למספרים שלמים, מרווח הקוונטיזציה הוא 1. אם לדוגמה מעגלים סכומי כסף לאלפי שקלים, המרווח יהיה 1000.

2.3. שגיאת תהליך הכימוי

כמובן שתהליך הכימוי גורם לאובדן של אינפורמציה, ולכן ישנה בד"כ שגיאה.

לדוגמה, נביט בקוונטייזר אחיד. גודל השגיאה בקוונטייזר אחיד יהיה לכל היותר $\frac{q}{2}$.
בהנחה שרמות האפור בכניסה לקוונטייזר יכולות לקבל כל ערך בהסתברות אחידה, נקבל צפיפות הסתברות אחידה גם לערך השגיאה של הקוונטייזר האחיד:



ואז, השגיאה הממוצעת ($E[\varepsilon]$) היא אפס, כלומר בממוצע ערכי הפיקסלים לא משתנים, כלומר התמונה במוצא הקוונטייזר לא תהיה יותר בהירה או יותר כהה מאשר התמונה המקורית בכניסת הקוונטייזר.

כעת נביט על השונות של שגיאת הקוונטייזר האחיד:

$$E\left[\varepsilon^2\right]=\int_{-\frac{q}{2}}^{\frac{q}{2}}\varepsilon^2\cdot\frac{1}{q}d\varepsilon=\left[\frac{1}{q}\frac{\varepsilon^3}{2}\right]_{-\frac{q}{2}}^{\frac{q}{2}}=\frac{1}{q}\frac{\frac{q^3}{8}+\frac{q^3}{8}}{2}=\frac{q^2}{8}$$

למדד זה קוראים גם "השגיאה הריבועית הממוצעת" (MSE – Mean Square Error). בד"כ נניח כי זהו מדד סביר לאיכות של תמונה, אבל האם זהו באמת מדד טוב?
דוגמא:
נביט בתמונה בגודל 1000×1000 שחורה לגמרי, כאשר בשורה 28 יש קו דק (בגובה פיקסל אחד) בצבע לבן.
כעת נניח שהתמונה עברה שני עיבודים. לאחר העיבוד הראשון, התמונה שנוצרה היתה זהה, מלבד הקו הדק, שכעת נמצא בשורה 29 במקום בשורה 28. לפי מדד השגיאה שפגשנו, MSE, התרומה לשגיאה תהיה בשורה 28, שם כל הפיקסלים הלבנים של הקו הפכו לשחורים, ובנוסף בשורה 29, שם הפיקסלים השחורים המקוריים הם עתה בצבע לבן. לאחר העיבוד השני, התמונה שהתקבלה שחורה לגמרי. לפי MSE, התרומה לשגיאה תהיה רק משורה 28, שם הקו הלבן נעלם וכעת הכל שחור. לסיכום – מבחינת מדד ה MSE, התמונה השחורה יותר קרובה למציאות מאשר התמונה שבה הפס הלבן זז מטה. זאת למרות שלצפופה האנושי מצב זה ממש לא הגיוני – התמונה עם הפס הלבן בשורה 29 נראה מאוד דומה לתמונה עם הפס המקורי בשורה 28 (די קשה להבחין, מתוך 1000 השורות היכן הקו הלבן...), והתמונה השחורה לגמרי, היא התמונה עם השגיאה הגדולה – שם כל האינפורמציה אבדה.
בכל זאת, משתמשים במדד זה בגלל פשטות החישוב שלו.

נגדיר מדד לשגיאה בצורה מדוייקת.

סימון	הסבר
u	ערך פיקסל בכניסת הקוונטייזר
u'	ערך פיקסל ביציאה (אחת מרמות הייצוג)
$d(u,u')$	מדד השגיאה – פונקציה של הכניסה והיציאה
$P_U(u)$	פילוג רמות האפור בתמונה

לדוגמא, נוכל לבחור $d(u,u')=(u-u')^2$.

נבחר מדד לשגיאה הגלובלית $D=E\left[d(u,u')\right]$ (כלומר תוחלת). לא היינו חייבים לבחור מדד כזה – יכולנו גם לבחור את השגיאה המקסימלית מבין כל הפיקסלים שבתמונה, כלומר $D=\max\left\{d(u,u')\right\}$.

כעת, מעניינים אותנו פיקסלים רק ברמות (בערכים) שקיימות בתמונה – אלו רמות ההחלטה שלנו:

$$D=E\left[d(u,u')\right]=\int_{r_0}^{r_n}d(u,u')P_U(u)\,du$$

מכיוון שהערך שמתקבל עבור כניסה בין שתי רמות החלטה $\left[r_{k-1},r_k\right]$ הוא f_k, נוכל לכתוב:

$$D=E\left[d\left(u,u'\right)\right]=\sum_{k=1}^{n}\int_{r_{k-1}}^{r_k}d\left(u,f_k\right)P_U\left(u\right)du$$

במקרה הפרטי, שבו $d\left(u,f_k\right)=\left(u-f_k\right)^2$, נקבל:

$$D=E\left[d\left(u,u'\right)\right]=\sum_{k=1}^{n}\int_{r_{k-1}}^{r_k}\left(u-f_k\right)^2P_U\left(u\right)du$$

נשים לב שהפונקציה $D\left(r_0,r_1,...,r_n,f_1,f_2,...,f_n\right)$ רציפה בכל משתניה, מכיוון שאם נזיז את אחד מהמשתנים ב ε_0 קטן, לא נקבל קפיצות חריגות של D לאינסוף.
כדי להביא למינימום את השגיאה הגלובלית, נגזור ונשווה ל 0:

$$\begin{cases}\dfrac{\partial D}{\partial r_k}=0, & k=1,2,3,...,n-1\\[2mm] \dfrac{\partial D}{\partial r_f}=0, & k=1,2,3,...,n\end{cases}$$

נזכר בכלל הבא:

$$\frac{d}{d\alpha}\int_{u_0(\alpha)}^{u_1(\alpha)}g\left(x\right)dx=g\left(u_1\right)\cdot\frac{du_1}{d\alpha}-g\left(u_0\right)\frac{du_0}{d\alpha}$$

ראשית נניח כי ידועות רמות הייצוג f_k, ונגזור לפי רמות ההחלטה r_k. נשים לב שרמה r_k תופיע בשני אינטגרלים:

$$\int_{r_{k-1}}^{r_k}d\left(u,f_k\right)P_U\left(u\right)du+\int_{r_k}^{r_{k+1}}d\left(u,f_{k+1}\right)P_U\left(u\right)du$$

ואז:

$$0=\frac{\partial D}{\partial r_k}=d\left(r_k,f_k\right)P_U\left(r_k\right)-d\left(r_k,f_{k+1}\right)P_U\left(r_k\right)$$

נניח ש $P_U\left(r_k\right)\neq 0$, ונקבל

$$d\left(r_k,f_k\right)=d\left(r_k,f_{k+1}\right)$$

תנאי זה נקרא תנאי השכן הקרוב (Nearest Neighbour).
ניתן לראות שיש דרגות חופש בבחירת הפתרון לבעיה זו.

עבור מדד MSE, נניח שידועים שתי רמות הייצוג f_1,f_2, ואנחנו צריכים להחליט איפה לשים את r_1. לפי התנאי שלעיל, יש להגדיר את r_1 בדיוק באמצע: $r_1=\dfrac{f_2+f_1}{2}$.
נראה זאת באופן כללי. התוצאה שקיבלנו אומרת שנרצה שהשגיאה מימין תהיה שווה לשגיאה משמאל, ועבור מדד MSE, נרצה שהמרחק בין r_k ל f_k (מהביטוי $\left(r_k-f_k\right)^2$) יהיה שווה למרחק בין r_k ל f_{k+1} (מהביטוי $\left(r_k-f_{k+1}\right)^2$).
כלומר, עבור בחירה של $D=E\left[\left(u-u'\right)^2\right]$, כלומר MSE, נקבל

$$r_k=\frac{f_k+f_{k+1}}{2}$$

כעת נניח שידועות רמות ההחלטה, ונגזור לפי רמות הייצוג f_k:

$$0=\frac{\partial D}{\partial f_k}=\frac{\partial}{\partial f_k}\left(\int_{r_{k-1}}^{r_k}d\left(u,f_k\right)P_U\left(u\right)du\right)$$

כאשר לא לקחנו את כל האינטגרלים כי רמת הייצוג f_k תופיע רק באחד מהם. נמשיך:

$$0=\int_{r_{k-1}}^{r_k}\frac{\partial}{\partial f_k}\left(d\left(u,f_k\right)\right)P_U\left(u\right)du$$

ועבור מדד השגיאה הריבועית:

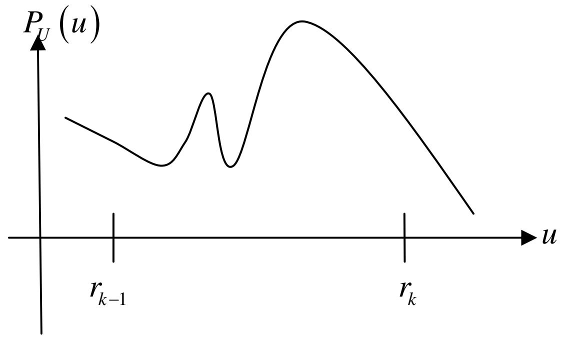
$$0=\int_{r_{k-1}}^{r_k}\frac{\partial}{\partial f_k}\left(\left(u-f_k\right)^2\right)P_U\left(u\right)du=-2\int_{r_{k-1}}^{r_k}\left(u-f_k\right)P_U\left(u\right)du=-2\int_{r_{k-1}}^{r_k}uP_U\left(u\right)du-2\int_{r_{k-1}}^{r_k}f_kP_U\left(u\right)du$$

כלומר

$$f_k=\frac{\int_{r_{k-1}}^{r_k}uP_U\left(u\right)du}{\int_{r_{k-1}}^{r_k}P_U\left(u\right)du}$$

כלומר, בהינתן רמות החלטה סמוכות r_{k-1}, r_k, ניתן לחשב היכן להגדיר את רמת הייצוג f_k שביניהן.

לדוגמה, נניח כי בין שתי רמות החלטה סמוכות r_{k-1}, r_k התפלגות הכניסה היא:



אזי רמת הייצוג f_k תתקבל קרוב יותר ל r_k בגלל השקלול שמבצעים האינטגרלים, שהו בעצם מרכז הכובד.

לסיכום נאמר שנהוג להשתמש בשגיאה הריבועית הממוצעת כדי להעריך את טיב הקוונטיזציה. ניתן להשתמש ב

$$MSE=\sum_{k=1}^{N}\int_{r_{k-1}}^{r_k}\left(u-f_k\right)^2P_U\left(u\right)du$$

כדי לחשב את התוחלת של השגיאה הריבועית, בהנתן היסטוגרמת תמונת המקור u, או ב

$$MSE=\frac{1}{MN}\sum_{m=1}^{M}\sum_{n=1}^{N}\left(u\left[m,n\right]-u'\left[m,n\right]\right)^2$$

כדי לחשב את הסטייה המעשית בין תמונת המקור u לתמונה שאחרי הקוונטיזציה u'.

2.4. קוונטייזר מקס-לויד (Max-Loyd)

זהו הקוונטייזר שממזער את השגיאה הריבועית הממוצעת, MSE, כפי שראינו בדוגמה שלעיל.
כדי למצוא את רמות הייצוג ורמות ההחלטה של קוונטייזר שכזה, עלינו לפתור את מערכות המשוואות:

$$\begin{cases} r_k=\dfrac{f_k+f_{k+1}}{2} \\ f_k=\dfrac{\int_{r_{k-1}}^{r_k}uP_U\left(u\right)du}{\int_{r_{k-1}}^{r_k}P_U\left(u\right)du} \end{cases}$$

נוכל לפתור באופן איטרטיבי:

1. נבחר סט התחלתי $\left\{r_k\right\}$;
2. נחשב את $\left\{f_k\right\}$ ע"י הנוסחה לחישוב מרכז הכובד שקיבלנו;
3. נעדכן את $\left\{r_k\right\}$ ע"י הנוסחה לחישוב השכן הקרוב שקילבנו;
4. תנאי עצירה אפשריים: האם רמת הדיוק מספיקה לנו? האם עשינו מספיק איטרציות? האם קצב השינוי של $\left\{r_k\right\}$ קטן עם האיטרציות? האם ה MSE הנוכחי מספק אותנו?
נוכל להחליט שסיימנו, ואם לא החלטנו שסיימנו, נמשיך לשלב 2.

חשוב לציין שפתרונות איטרטיביים מסוג זה לא מבטיחים להגיא לפתרון האופטימלי, אלא לפתרון נכון בצורה מקומית.

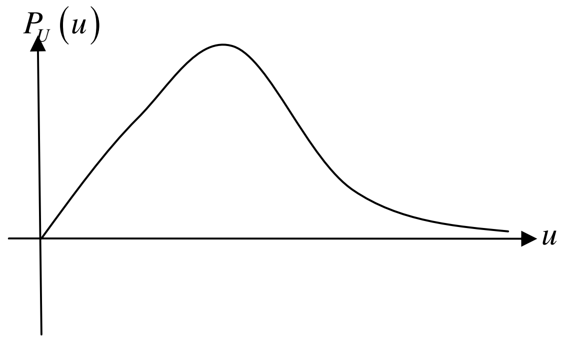
2.5. קוונטייזר אחיד

בקוונטייזר אחיד מתקיים :

$$\begin{cases} r_k = r_0 + kq \\ f_k = r_0 + \left(k - \frac{1}{2}\right)q \end{cases}, \quad k = 1,2,3,...n$$

נראה שיש כאן שתי דרגות חופש – בחירת r_0 ובחירת צעד הקוונטיזציה q, אך ראינו כי השגיאה תקטן ככל ש q יקטן, לכן בחירת q איזה באמת דרגת חופש.

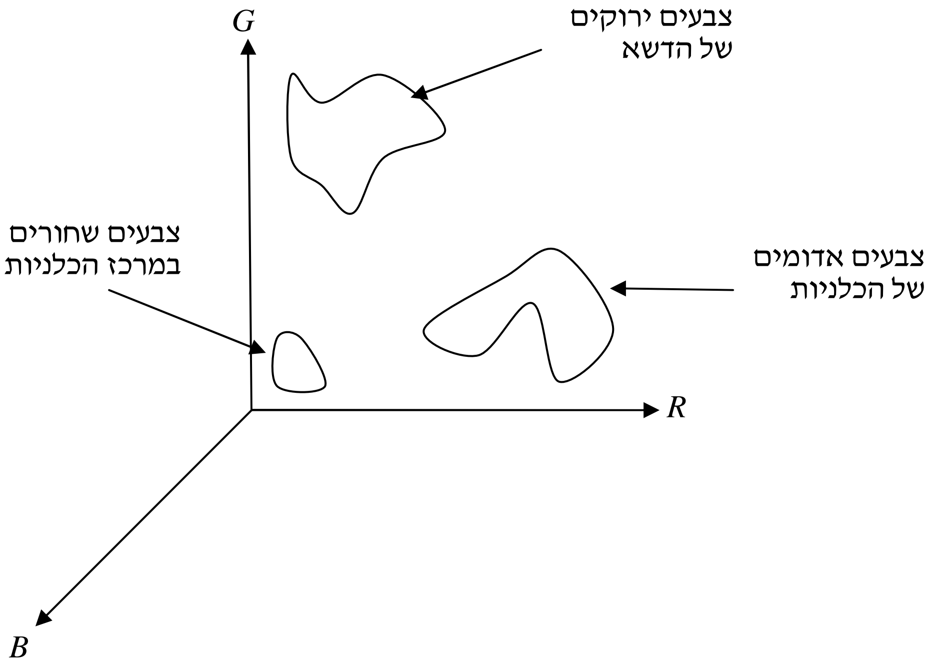
בעצם, קוונטייזר אחיד לא כל כך טוב לנו. נביט בפילוג ההתסברות של הכניסה הבא (היסטוגרמה) :



זוהי תמונה יחסית כהה, מכיוון שרוב הפיקסלים בעלי גוון כהה. קוונטייזר אחיד היה מפזר את רמות ההחלטה באופן אחיד ע״ג ההיסטוגרמה, מה שהיה גורם לשגיאה גבוהה עבור פיקסלים כהים.

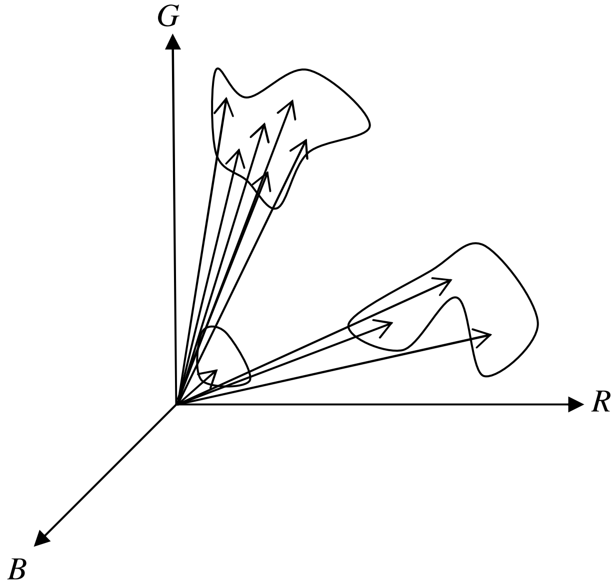
2.6. כימוי של תמונות צבעוניות

תמונה צבעונית ניתנת לייצוג על-ידי 3 מטריצות, לכל אחד מהגוונים אדום (R), ירוק (G) וכחול (B).
בצורה נאיבית, ניתן לבצע כימוי לכל אחד מהערוצים (צבעים) בנפרד. בד״כ פעולה זו לא יעילה.
נביט, לדוגמא, בתמונה של דשא וכלניות, ונציג Scatter-Plot של הגוונים :



את ה Scatter-Plot יוצרים ע״י התבוננות בערכי ה (R,G,B) של כל פיקסל, ושרטוט נקודה מתאימה לכל וקטור שכזה. ניתן לראות בבירור שבתמונה זו, אם היינו משתמשים בכימוי אחיד לכל ערוץ, היה בזבוז – יש תחומים בכל ערוץ שכלל אינם בשימוש.

נוכל לבצע קוונטיזציה וקטורית: נחלק את ה Scatter-Plot לאיזורים בעלי חשיבות (אלו יגדירו איזורי החלטה S_i, במקביל למונח רמות ההחלטה r_i במקרה של גווני האפור), ולכל איזור החלטה כזה זה נגדיר וקטור המייצג את הצבעים באותו איזור (וקטור ייצוג $\underline{f_i}$, במקביל למונח רמת ייצוג f_i במקרה של גווני האפור):



כך, עבור כל פיקסל במוצא, נשמור את מספר הוקטור המתאים לו, המייצג את צבעו. בדוגמה הזו בחרנו צבע אחד מייצג לצבעי השחור, 3 צבעים מייצגים לגווני האדום ו 6 גווני ירוק.

לכל וקטור i יש לשמור את הצבע (R,G,B) המתאים לו. נשמור נתונים אלו בטבלה, התקרא Color Map:

B	G	R	וקטור
0	12	3	0
65	74	15	1
200	111	21	2
			.
			.
			10

נבחן כעת איך יכול הקוונטייזר לחשב את וקטור הייצוג המתאים לפיקסל כלשהו $\underline{x}=(R,G,B)$. ע"פ תנאי השכן הקרוב, אם לכל j, $i \neq j$, מתקיים

$$d\left(\underline{x},\underline{f_i}\right) \leq d\left(\underline{x},\underline{f_j}\right)$$

אזי $x \in S_i$, ולכין ייוצג ע"י וקטור הייצוג $\underline{f_i}$.

כעת, נאמר שנתונים איזורי ההחלטה S_i, והשאלה היא איך נקבע את וקטורי הייצוג. נחשב את מרכז הכובד של מקבץ הפיקסלים באזור ההחלטה הזה, כלומר:

$$\underline{f_i} = \frac{1}{N_i}\sum_{\underline{x}\in S_i} \underline{x}$$

כאשר N_i הוא מספר הפיקסלים בתוך האזור S_i.

נבחן כעת את אופן ייצוג הצבעים בצורה אופטימלית מבחינת מערכת הראייה האנושית.

כידוע, ניתן לייצג פיקסלים ע"י ערכי (R,G,B) שלהם. עבור העין האנושית, עצמת האור של כל פיקסל היא נתון יותר חשוב מאשר צבע הפיקסל (לכן תמונות שחור/לבן יותר ברורות לעין האנושית).

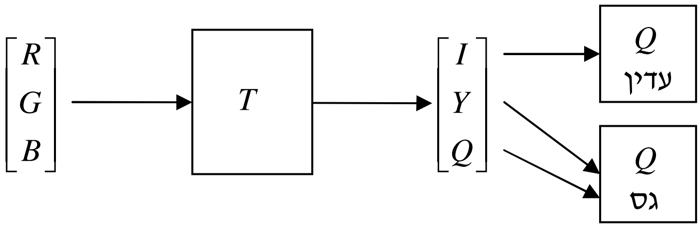
את עצמות האור של פיקסל (Luminance), בהינתן ערכי (R,G,B) שלו, ניתן לחשב, לדוגמא, ע"י הסכום $R+G+B$. ניתן לסכום ערכים אלו גם עם משקלים מסויימים. בכל מקרה, ניתן לבצע טרנספורמציה לינארית על השלשה (R,G,B), ולקבל מספר אחד המייצג את עצמת האור של הפיקסל.

בנוסף לעצמת האור, יש לשמור עוד אינפורמציה, הנקראת Chromaticity, המייצגת את צבע הפיקסל. אינפורמציה זו נשמרת בשני מספרים נוספים, הנסמן כרגע ב (Y,Q). כאמור, העין האנושית פחות רגישה לפרמטרים אלו.

בסופו של עניין, נוכל למצוא טרנספורמציה T אשר תתן לנו את ערכי (I,Y,Q) לכל פיקסל:

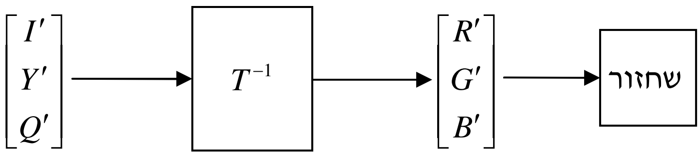
$$\begin{bmatrix} I \\ Y \\ Q \end{bmatrix} = \underline{\underline{T}}_{3\times 3} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

מכיוון שהעין האנושית רגישה יותר לעוצמת האור, ופחות לצבע, נרצה לבצע קוונטיזציה עדינה לפרמטר I, וקוונטיזציה גסה יותר לשני הפרמטרים הנוספים:



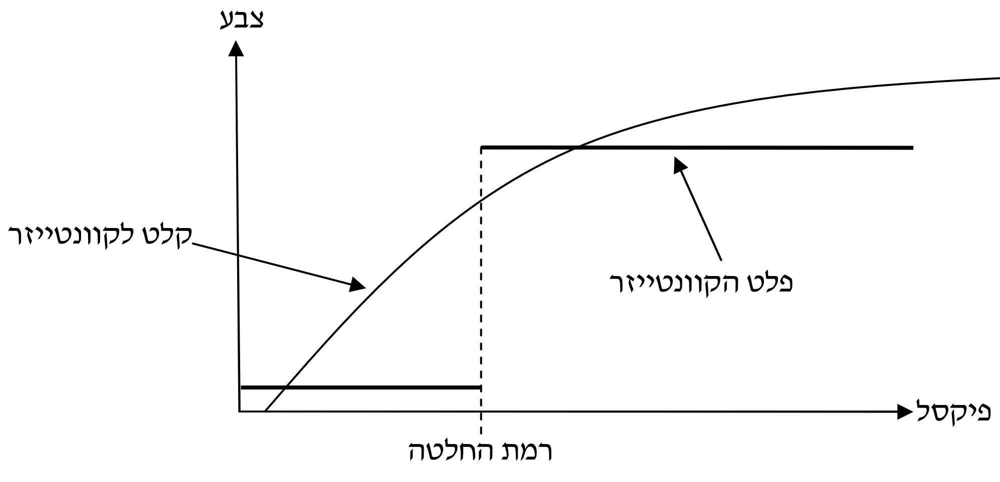
כך, ניתן לחסוך בכמות האינפורמציה שיש לשמור עבור תמונה צבעונית, מכיוון שהתחשבנו בעובדה שקוונטיזציה גסה של האינפורמציה לגבי צבעי הפיקסלים תתן לנו תוצאה שהצופה האנושי יהיה מרוצה ממנה.

שחזור ערכי (R,G,B) של כל פיקסל יתבצעו ע״י הטרנספורמציה ההפוכה. כמובן שיש לשים לב שערכי (I,Y,Q) עבור השחזור שונים מערכי (I,Y,Q) שקיבלנו לאחר הטרנספורמציה הראשונה, מכיוון שאלו עברו קוונטיזציה. כמובן שגם ערכי (R,G,B) המשוחזרים יהיו, מאותה סיבה, שונים.



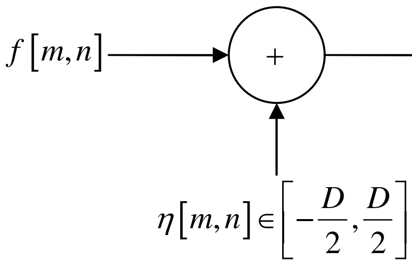
2.7. תופעת Contouring

כשמבצעים קוונטיזציה, ניתן להבחין בתופעה שבה מופיעים הבדלים ניראים לעין בין אזורים בצבעים שונים – נראה כי יש קו מפריד בין שני אזורים אלו, ומכאן השם Contour. נקרא גם Banding.
ניתן לראות את הסיבה לתופעה זו, בגרף הבא:



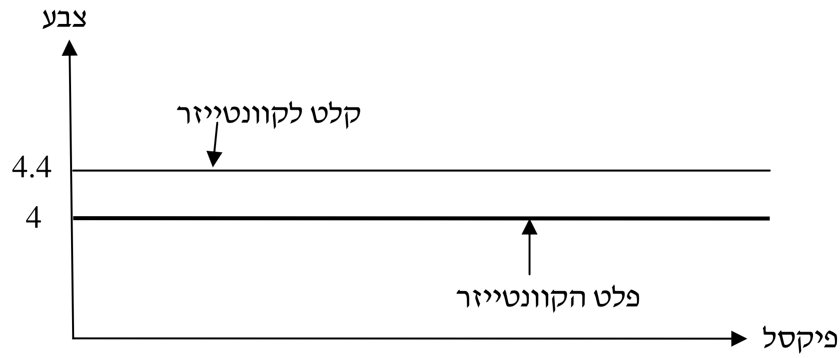
2.8. פעולת Dithering

פעולה הנועדה להעלים את עקבות תופעת ה Contouring.
זוהי פעולה של הכנסה של רעש η הנע בטווח ערכים ידועים, בכוונה תחילה, לפני פעולת הקוונטיזציה Q, וזאת כדי לעדן את המעברים החדים הקורים בגלל המעברים בין רמות הייצוג.

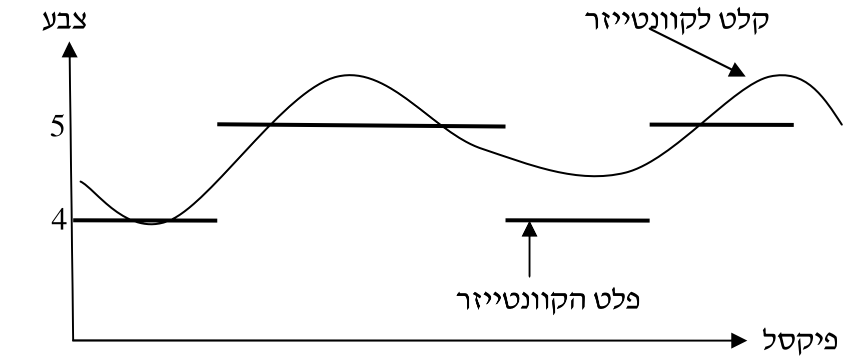


דוגמא

נביט באזור בתמונה, אשר ערכי הפיקסלים בו (גוון אפור) הם קבועים, בערך 4.4 . נאמר שהקוונטייזר מעגל למספרים שלמים, ולכן נקבל:



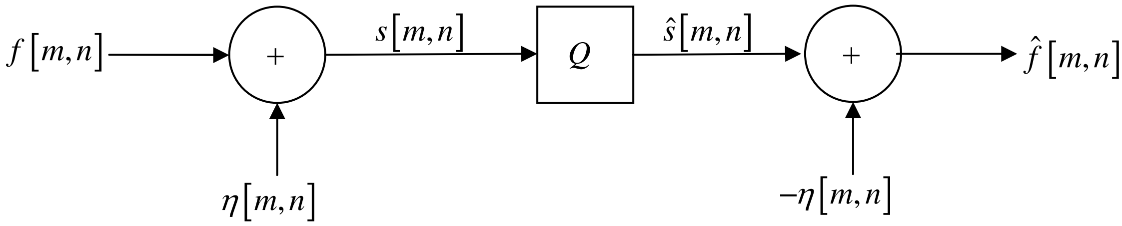
כלומר, התמונה העברה כימוי תראה בהירה יותר מאשר המקורית. לאחר שנוסיף את הרעש, נקבל:



ואנו רואים שבאותו איזור שהיה בהיר יותר, יהיו עכשיו צבעים משתנים, כך שלעין האנושית (שבין היתר מבצעת מיצוע על קבוצות של פיקסלים, ולא רואה פיקסלים בודדים) תהיה תחושה של צבע יותר כהה מהצבע שערכו 4 .

את התמונה שעברה כימוי ($\hat{s}[m,n]$) נשמור במחשב.

כעת נרצה להציג את התמונה (לשחזר את $f[m,n]$). נניח כי שמרנו במחשב, בנוסף לתוצאת הקוונטיזציה $\hat{s}[m.n]$, את שדה הרעש $\eta[m.n]$, ולפני הצגת התמונה, נחסיר את הרעש:



נחשב את שגיאת התהליך:

$$d = f - \hat{f} = (s-\eta) - (\hat{s}-\eta) = s - \hat{s}$$

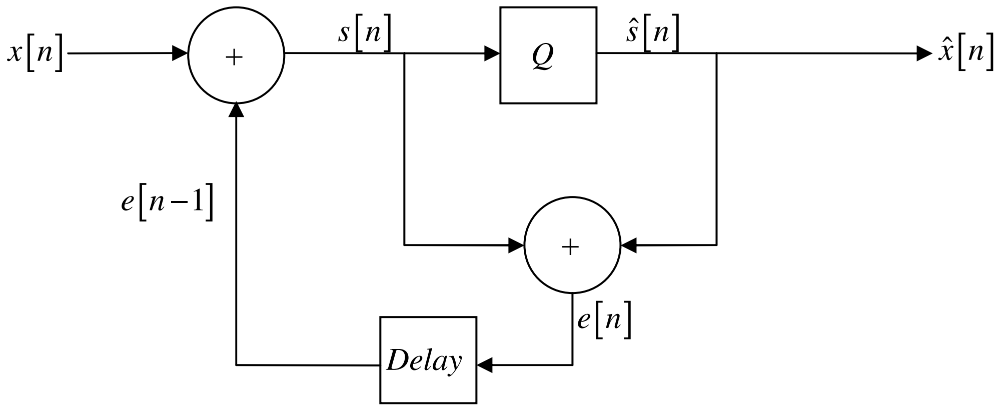
כלומר השגיאה הכוללת שווה לשגיאה של פעולת הקוונטיזציה – לא הכנסנו שגיאה נוספת בגלל ה Dithering.

שימו לב כי הנחנו כאן שאנחנו שומרים את מטריצת הרעש $\eta[m,n]$. פעולה זו יקרה מבחינת זכרון, ולכן נשתמש בפסאודו-רעשים חסכוניים יותר, כגון סדרה הנראת אקראית, המורכבת מהספרות המרכיבות את המספר $\pi = 3.1415...$.

2.9. Error Diffusion

נמשיך כעת ונבחן שיטה טובה יותר להתגבר על תופעת ה Contouring של פעולת הקוונטיזציה. ראינו שפעולת ה Dithering עזרה לנו להציג את התמונה בצורה טובה יותר כאשר היו מעברים בין רמות ייצוג. אבל יחד עם זאת, תוספת הרעש תפגע בתמונה באזורים בהם הערך האמיתי תואם לרמת הייצוג – אם, בהתאם לדוגמה שלעיל כאשר לא מוסיפים רעש, ערכי הפיקסלים הם 4, אזי הקוונטייזר ייצג אותם כראוי – ערך המוצא יהיה 4. הוספת הרעש תפגע בייצוג זה, שכן הרעש מתווסף לכל הפיקסלים בתמונה.

שיטת Error Diffusion (פעפוע השגיאה) באה להוסיף רעש בצורה חכמה יותר מאשר שיטת Dithering. בשיטה זו אנחנו נבדוק אם שגיאת הקוונטייזר, בזמן פעולתו, מצריכה הכנסה של "רעש" (לא באמת נכניס רעש, אלא משהו אחר).



כאן אנו רואים הצגה וקטורית $x[n]$ של כלל הפיקסלים של התמונה (כלומר יש לקחת את כל $M \cdot N$ הפיקסלים של התמונה $f[n,m]$ לפי הסדר, נאמר משמאל למעלה עד צד ימין למטה, ולהניח אותם בוקטור אחד ארוך).

בשגיאת הקוונטיזציה, $e[n] = \hat{s}[n] - s[n]$, נשתמש כדי להוסיפה לערך הפיקסל הבא בתמונה, כלומר נבצע

$$s[n] = x[n] + e[n-1]$$

דוגמא:

נביט באזור מהדוגמא הקודמת, בו ערכי התמונה הם 4.4 והקוונטייזר מעגל למספרים שלמים:

n	$x[n]$	$e[n-1]$	$s[n]$	$\hat{x}[n]$
1	4.4	0	4.4	4
2	4.4	0.4	4.8	5
3	4.4	-0.2	4.2	4
4	4.4	0.2	4.6	5
5	4.4	-0.4	4	4
6	4.4	0	4.4	4
7	4.4	0.4	4.8	5

לשימחתנו, קיבלנו אפקט רצוי – הערך 4.4 יקבל במוצע ערכים קוונטיזציה משתנים, וכך, לאחר השחזור ומיצוע מערכת הראייה, הצופה האנושי יראו ערך שבין 4 ל 5, כמו בתמונה המקורית.

היתרון הנוסף, כפי שרמזנו בתחילה, הוא שערכים שלא צריכים להשתנות, אכן לא משתנים. לדוגמא:

n	$x[n]$	$e[n-1]$	$s[n]$	$\hat{x}[n]$
1	4	0	4	4
2	4	0	4	4
3	4	0	4	4

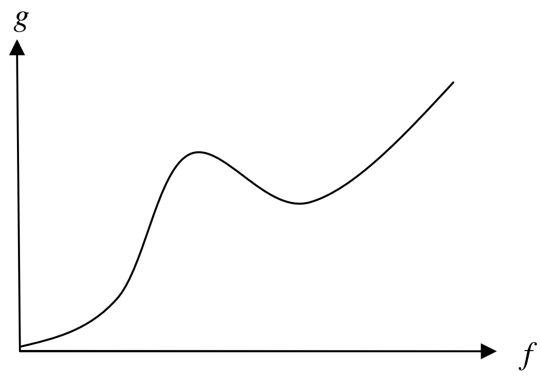
לסיכום, כשאנחנו משתמשים במנגנון אדפטיבי זה, אין צורך להכניס רעש סינטזי לתמונה.

3. פעולות נקודה

3.1. הקדמה

פעולות נקודה הן פעולות שבהן המוצא בנקודה $[n,m]$ תלוי רק בכניסה בנקודה $[n,m]$. באופן אנלוגי למערכות בחד-מימד, סוג פעולות אלו נקראו "חסרות זיכרון".

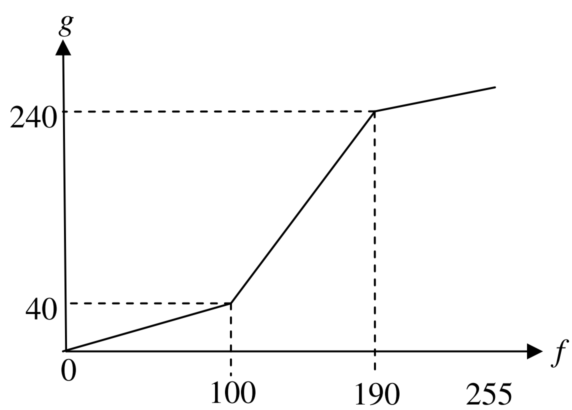
לדוגמא, ניתן להביט בפונקציה הבאה:



כאשר f ערך פיקסל בכניסה (הערך המציין את גוון האפור שלו), ו g ערך הפיקסל (גוון האפור) במוצא. כאמור g תלוי רק ב f, ולא בשכניו.

3.2. מתיחת ניגודיות (Contrast)

הפעולה הבאה:



מבצעת "מתיחת Contrast", מכיוון שהתחום הנמוך והגבוה של גווני האפור עוברים לתחומים צרים, בעוד התחום הצר במרכז גווני האפור עובר לתחום רחב במוצא. ע"י כך שהתחום הצר יחסית באמצע, $[100,190]$, נמתח לתחום רחב יותר $[40,240]$, ישנה כאן הגברה של התחום, במובן שהצופה האנושי יוכל להבחין טוב יותר בפרטים הנמצאים בתחום זה. מאידך, התחום $[0,100]$ מתכווץ לתחום $[0,40]$, ולכן פרטים כהים (שגווני האפור שלהם במקור היו בין 0 ל 100) יהיו פחות ניתנים להבחנה לאחר מתיחת הקונטרסט.

מתי נראה שיפור בתמונה?

כאשר תמונה היא יחסית אפרורית, כלומר אין הרבה גוונים קרובים מאוד לקצוות (שחור ולבן), הפעולה לא תגרום להפסד ניכר בקצוות, ומצד שני, התחום שבו רוב הפיקסלים מופיעים יעבור מתיחה, וייוצג ע"י תחום רחב יותר של גוונים, ולכן הצופה האנושי יראה שיפור משמעותי בהבחנה בפרטים.

דוגמא לפונקציה למתיחת קונטרסט

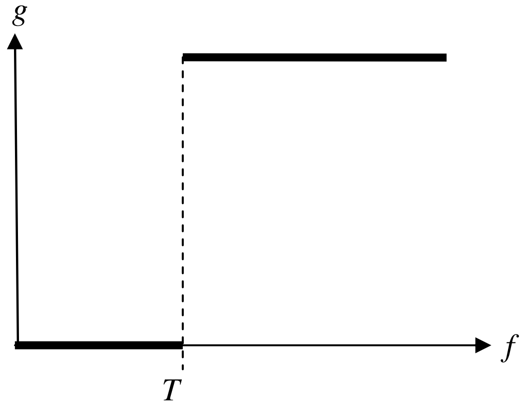
כאשר תמונה מיוצגת ב 255 גווני אפור, נחשב את $f_{\max}$, גוון האפור המקסימלי ו $f_{\min}$, גוון האפור המינימלי, ונבצע:

$$g = 255 \cdot \frac{f - f_{\min}}{f - f_{\max}}$$

באופן זה, נשתמש באופן אפקטיבי יותר בתחום בדימני שבין 0 ל 255, כך שנייצד את כל רמות האפור הקיימות בתמונה בעזרת התחום הזה (בתחום המקסימלי).

3.3. פעולת סף (Thresholding)

פעולה זו הופכת תמונת גווני לאפור לתמונה בעלת שני צבעים בלבד, וזאת ע״י הבחנה בסף מוגדר $T \in [0, 255]$:



סף זה לא חייב להיות באמצע, כלומר בערך 128, אלא יכול להקבע על-ידנו.
פעולה זו נקראת גם ״בינאריזציה״.

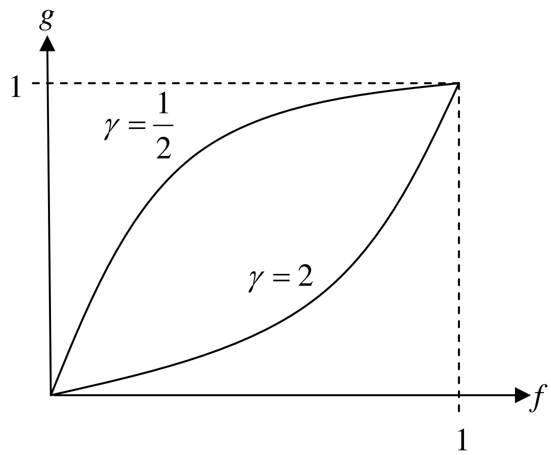
3.4. תיקון גאמא

תיקון גאמא (γ Correction) מבצע מתיחת קונטרסט ע״י משפחה של קווים מעוקלים מהצורה

$$g = f^{\gamma}$$

נביט על גווני האפור כמנורמלים לערכים בתחום $[0,1]$.

בפעולה זו, חוץ מהערכים 0 ו 1, הנשארים בערכם המקורי, ניתן למתוח את תחום הגוונים הבהירים או את תחום הגוונים הכהים, בהתאם ל γ. המתיחה בתחום אחד באה על חשבון הכיוון בתחום האחר, כפי שניתן לראות בגרף הבא.



ניתן לשלב ולבצע תיקון גאמא לאחר ביצוע מתיחת קונטרסט:

$$g = \left(\frac{f - f_{\max}}{f_{\max} - f_{\min}} \right)^{\gamma}$$

מתיחת הקונטרסט מביאה את ערכי הגוונים האפורים לתחום $[0,1]$, ולאחר מכן מתבצע תיקון גאמא.

תיקון גאמא הוא תיקון שנכנס בעקבות עיוות שנוצר בעת שידורי הטלביזיה. מסיבות של תאימות לאחור, תיקונים כאלה מוכנסים היום למשדרים וגם למקלטים, למרות שכבר כלל אין צורך באלו.

3.5. עיצוב היסטוגרמה

את מונח ההיסטוגרמה כבר פגשנו בפרקים הקודמים.
נניח כי היסטוגרמות (צפיפות ההסתברות של גווני האפור) תמונת המקור, $P_f(x)$, ותמונת הפלט $P_g(x)$ הן רציפות.
ניתן לראות את השפעת מתיחות הקונטרסט שביצענו בסעיפים הקודמים על ההיסטוגרמה של התמונה.
נרצה כעת לעצב את ההיסטוגרמה של תמונה כלשהי, כלומר להביא את ההתפלגות הידועה של תמונת המקור $P_f(x)$ להתפלגות רצוייה $P_g(x)$. נרצה לדעת מה היא הטרנספורמציה T (נניח מונוטונית) שתשנה את ההיסטרגמה באופן הזה.
נביט בפונקציות ההסתברויות המצטברות:

$$F_f(x) = \Pr\{f \le x\} = \int_{-\infty}^{x} P_f(x)\,dx$$

$$F_g(x) = \Pr\{g \le x\} = \int_{-\infty}^{x} P_g(x)\,dx$$

נניח ש T מונוטונית עולה:

$$F_f(x) = \Pr\{f \le x\} = \Pr\{T(f) \le T(x)\}$$

וכמובן $g = T(f)$, ולכן

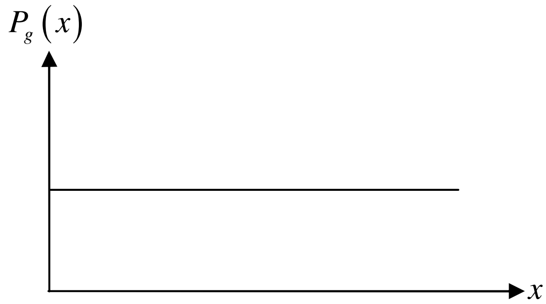
$$F_f(x) = \Pr\{Tf \le Tx\} = \Pr\{g \le T(x)\} = F_g(T(x))$$

ומכיוון ש $F_f(x)$ וגם $F_g(x)$ מונוטוניות עולות, יש להם פונקציות הופכיות, ואז מהשוויון שלעיל ניתן לקבל את הטרנספורמציה שחיפשנו:

$$T(x) = F_g^{-1}(F_f(x))$$

3.6. קיזוז היסטוגרמה (Histogram Equialization)

נביט כעת במקרה הפרטי בו נרצה שההיסטוגרמה במוצע תהיה אחידה (קבועה לכל גוון אפור):



לכן במקרה זה:

$$F_g(x) = \Pr\{g \le x\} = \int_{-\infty}^{x} P_g(x)\,dx = x$$

וההופכית:

$$F_g^{-1} = x$$

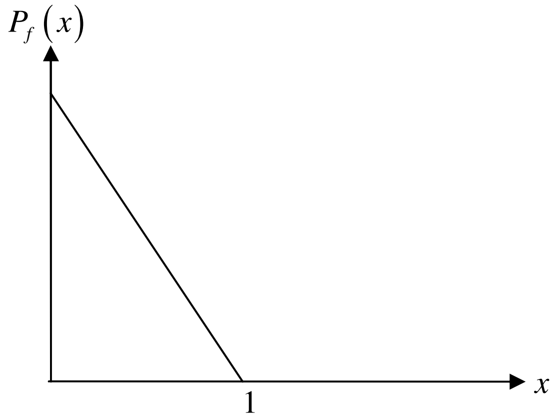
ולכן קיבלנו כי

$$T(x) = F_g^{-1}(F_f(x)) = F_f(x)$$

דוגמא

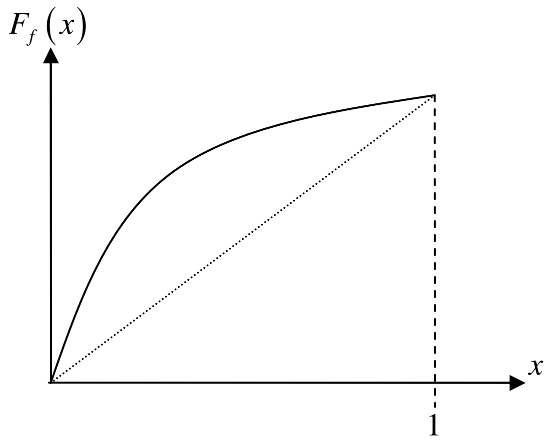
נביט בתמונה בעלת ההיסטורגמה הבאה:

$$P_f(x) = \begin{cases} 2(1-x), & 0 \le x \le 1 \\ 0, & else \end{cases}$$



נקבל:

$$F_f\left(x\right)=\int_{-\infty}^{x}P_f\left(x\right)dx=2x-x^2$$



בעזרת ההשוואה לקו האלכסוני (שלא משנה את ההיסטוגרמה – זוהי פונקצית היחידה), אנחנו רואים שהטרנספורמציה שלנו, $T\left(x\right)=F_f\left(x\right)$, תבהיר את התמונה. דבר זה מתאים לעובדה כי ראינו שהתמונה המקורית בעלת צבעים מאוד כהים.

כעת נרצה שההיסטורגמה של תמונת הקלט תהפוך להיסטוגרמה הבאה:

$$P_g\left(x\right)=\begin{cases}2x, & 0\le x\le 1\\ 0, & else\end{cases}$$

ע״י חישוב נקבל

$$F_g\left(x\right)=x^2$$

ולכן הטרנספורמציה המתאימה היא

$$T\left(x\right)=F_g^{-1}\left(F_f\left(x\right)\right)=F_g^{-1}\left(2x-x^2\right)=\sqrt{2x-x^2}$$

ניתן לראות שיש כאן בעצם הרכבה של פוקציות, כאשר הפונקציה הראשונה, $F_f\left(x\right)$, מבצעת את הקיזוז (Equalization) שראינו בהתחלה, והפעולה השנייה, $F_g^{-1}\left(x\right)$, מביאה את ההיסטוגרמה האחידה לצורה המבוקשת.

3.7. טבלאת איתור (Look-up Table)

עד כה, התייחסנו להסיטוגרמות כאל פונקציות רציפות. אך כמובן שהדבר לא כך.

בפועל, פעולות נקודה מתבצעות בעזרת טבלאות איתור, כאשר הטרנספורמציה מעבירה ערך של פיקסל לערך חדש בהתאם לטבלה. לדוגמא, הטבלה הבאה מייצגת טרנספורמציה שפועלת על תמונה בעלת 4 גווני אפור אפשריים. באופן כללי, הטרנספורמציה תבהיר את התמונה:

סידורי	גוון בכניסה, f	גוון במוצא, g
1	12	200
2	7	180
3	100	255
4	2	80

דוגמא

בתמונה f יש 20 פיקסלים, וסה"כ 8 רמות אפור. להלן וקטור ההיסטוגרמה של התמונה:

$$h_f = [1 \quad 2 \quad 3 \quad 4 \quad 4 \quad 3 \quad 2 \quad 1]$$

ניתן לראות שגווני הפיקסלים מרוכזים ברמות האמצעיות.

נרצה לקבל את ההיסטוגרמה הבאה במוצא:

$$h_g = [4 \quad 3 \quad 2 \quad 1 \quad 1 \quad 2 \quad 3 \quad 4]$$

באופן אנלוגי לפונקצית ההסתברות המצטברת, נחשב את פונקצית ההיסטוגרמה המצטברת של ההתפלגות המקורית f ושל ההתפלגות הרצויה g:

$$H_f = [1 \quad 3 \quad 6 \quad 10 \quad 14 \quad 17 \quad 19 \quad 20]$$

$$H_g = [4 \quad 7 \quad 9 \quad 10 \quad 11 \quad 13 \quad 16 \quad 20]$$

במקרה הרציף, דרשנו ש $F_g(T(x)) = F_f(x)$, כלומר באופן אנלוגי, נרצה ש $H_g(T(x)) = H_f(x)$. כלומר, הטרנספורמציה צריכה לקחת גוון x כלשהו (בעל ערך בהיסטוגרמה המצטברת $H_f(x)$) ולהמיר אותו לגוון בעל אותו ערך בהיסטוגרמה המצטברת $H_g(x)$. בגלל האופי הבדיד של ההיסטוגרמות, כמובן שלא ניתן לעשות זאת. לכן, נבחר את הטרנספורמציה באופן שהערכים יהיו קרובים מספיק. לדוגמה, פיקסלים עם גוון 2 (בתמונה המקורית, מתאימים 3 פיקסלים בעלי גוון 2 או פחות) יעברו להיות פיקסלים בעלי גוון 1 (כי בהיסטוגרמה המצטברת הרצויה יש 4 פיקסלים בעלי גוון זה או פחות, וזה מספר הפיקסלים הקרוב ביותר ל 3). לסיכום, נקבל את ה Look-Up Table הבא:

גוון בכניסה, f	גוון במוצא, g
1	1
2	1
3	2
4	4
5	6
6	7
7	8
8	8

נוכל לחשב את ההיסטוגרמה הסופית שנקבל, בעקבות הטרנספורמציה שהגדרנו:

$$\tilde{h}_g = [3 \quad 3 \quad 0 \quad 4 \quad 0 \quad 4 \quad 3 \quad 3]$$

איך חושבה ההיסטוגרמה? לדוגמה – גוון 1: ע"פ הטרנספורמציה, כל הפיקסלים שמקור היו בגוון 1 (יש פיקסל אחד כזה, לפי h_f) וגם כל הפיקסלים שבמקור היו בגוון 2 (יש שני פיקסלים כאלו, לפי h_f) יהיו במוצא בגוון 1 סה"כ, $\tilde{h}_g(1) = 3$. גוון 5: ע"פ הטבלה, אף פיקסל לא יהיה בגוון כזה, לכן $\tilde{h}_g(5) = 0$.

4. פעולות מרחביות

4.1. מבוא

פעולות מרחבית הן פעולות אשר, בניגוד לפעולות נקודה, מתחשבות ביותר מנקודה אחת כדי לייצר את הפלט בנקודה מסויינת.

4.2. סינון לינארי

סינון לינארי מתבצע ע״י העברת האות (התמונה) דרך מערכת לינארית עם תגובה להלם $h[n,m]$:

$$f[m,n] \longrightarrow \boxed{H} \longrightarrow g[m,n]$$

אות המוצא מתקבל ע״י קונבולוציה :

$$g[m,n]=f[m,n]*h[m,n]=\sum_{k=-\infty}^{\infty}\sum_{r=-\infty}^{\infty}h[k,r]f[m-k,n-r]$$

ראינו שביצוע קונבולוציה, באופן גרפי ובמקביל לקונבולוציה בחד-מימד, יש לבצע היפוך (שתי פעולות ראי) של מטריצת גרעין הקונבולוציה, ביצוע מעבר על התמונה, הכפלה וסכימה איבר איבר.

לעיתים נוח להגדיר ״פעולת מסכה״, אשר די דומה לפעולת הקונבולוציה. אם המסכה w היא תמונת ראי (היפוך האיברים מימין-לשמאל ומעלה-מטה) של גרעין הקונבולוציה h :

$$w[m,n]=h[-m,-n]$$

פעולת המסכה בעזרת w , המבצעת את אותה הפעולה כמו קונבולוציה עם h , מוגדרת כך :

$$g[m,n]=\sum_{k=-\infty}^{\infty}\sum_{r=-\infty}^{\infty}w[k,r]f[m+k,n+r]$$

מבחינה גרפית, הפעולה זהה לפעולת הקונבולוציה, רק שאין צורך לבצע את פעולות הראי על גרעין הפעולה – פשוט לוקחים את מטריצת המסכה, ומעבירים אותה על תמונת המקור.
יש לשים לב שפעולת מסכה אינס אסוציאטיבית כמו פעולת הקונבולוציה, כלומר, בעוד ש

$$(f*g)*h=f*(g*h)$$

מתקיים

$$(f\odot g)\odot h\neq f\odot(g\odot h)$$

כאשר $\odot$ מסמל כאן את פעולת המסכה.

4.3. מיצוע מקומי

זוהי פעולת סינון לינארי עם המסנן (לדוגמא) :

$$h=\frac{1}{9}\begin{bmatrix}1&1&1\\1&1&1\\1&1&1\end{bmatrix}$$

פעולה זו תפיק במוצע את הערך הממוצע של פיקסל ושמונת שכניו. כאשר מסתכלים על איזור אחיד (גוון אפור זהה), המסנן לא ישנה אותו, אך כאשר יש מעברים בין גוונים, מסנן זה ימצע את המעברים, ולכן יווצר טשטוש.

פעולת מיצוע היא בעתם סינון-העברת-נמוכים, כלומר LPF – Low Pass Filtering.

מסנן זה טוב בסינון רעשים לבנים. זאת מכיוון שתכולת התדר של רעש לבן היא די אחידה על-גבי כלל התדרים (כלומר, יש לרעש הלבן אנרגיה בכל התדרים), לעומת תמונות, אשר סטטיסטית יש להן אנרגיה בעיקר בתדרים הנמוכים.

באופן כללי, שימוש במסנן לינארי כדי לסנן רעשים הוא בעייתי, כי מסנן כזה נוטה לטשטש שפות בתמונה (שפה היא איזור שבו יש שינויים חדים בגווני האפור).

4.4. מסנן חציון (Median)

עבור סביבה מסויימת (ניקח דוגמה של 3×3 פיקסלים), הערך במוצא יהיה החציון של 9 ערכי הפיקסלים השכנים.

נראה כי מסנן זה אינו ליניארי:

א. החציון של המטריצה הבאה הוא 0:

0	0	0
1	0	0
1	1	1

ב. החציון של מטריצה זו הוא גם 0:

1	1	1
0	0	1
0	0	0

ג. אבל, החציון של סכום המטריצות אינו $0=0+0$, אלא 1:

1	1	1
1	0	1
1	1	1

לדוגמא, נביט בפעולת המסנן, בגודל 3×3, על המטריצה הנ"ל:

		5	5	6	
		6	5	6	
		7	4	6	

הערך במוצא, בפיקסל $[3,3]$, יהיה החציון של 9 הערכים $4,5,5,5,6,6,6,6,7$, כלומר 6.

יתרונו הגדול של מסנן חציון הוא בכל שהוא עמיד לשינויים חדים בערכים. למשל, גם אם אחד הפיקסלים בערך 6 היה מקוול עיוות לערך גבוה כלשהו, נאמר 150, עדיין המוצא היה 6. המסנן הממצע שפגשנו קודם היה משופע משינוי כזה, אבל מסנן זה עמיד לשינוי זה.

לכן, פילטר זה אפקטיבי מאוד בסינון רעשים תוך כדי שמירה על שפות.

4.5. גזירת תמונה

נחשב את הגרדיאנט של התמונה:

$$\vec{\nabla} f(x,y)=\frac{\partial f}{\partial x}\hat{x}+\frac{\partial f}{\partial y}\hat{y} \qquad \Rightarrow \qquad \left|\vec{\nabla} f(x,y)\right|=\sqrt{\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2}$$

הגרדיאנט של התמונה הוא מדד טוב לגבי שינוי העוצמה בתמונה, בכל מרחב התמונה.

כמובן שהתמונות שלנו הן דיסקרטיות, ולכן נוכל רק לקרב את ערך הנגזרת. ביטוי לקירוב הנגזרת בפיקסל $[m,n]$, בכיוון העמודות הוא

$$f[m,n]-f[m-1,n]$$

והוא נקרא נגזרת אחורית. ניתן גם לחשב את הנגזרת הקדמית:

$$f[m+1,n]-f[m,n]$$

נוכל לקחת את ממוצע בין שתי אלו, לקבל:

$$\frac{\partial}{\partial y}f[m,n]\approx\frac{f[m+1,n]-f[m-1,n]}{2}$$

ובצורה דומה, עבור נגזרת בכיוון השורות :

$$\frac{\partial}{\partial x} f[m,n] \approx \frac{f[m,n+1]-f[m,n-1]}{2}$$

חישוב המטריצה המייצגת את $\frac{\partial f}{\partial x}$ שקול לביצוע קונבולוציה של התמונה עם הגרעין $h=\frac{1}{2}\begin{bmatrix}1 & 0 & -1\end{bmatrix}$, כלומר

$$\frac{\partial}{\partial x} f[m,n] = \frac{1}{2} f * \begin{bmatrix}1 & 0 & -1\end{bmatrix}$$

הנגזרת היא סוג של מסנן-מעביר-גבוהים, HPF – High Pass Filter. אם ניקח תמונה מורעשת, הרעשים יוגברו בתדרים הגבוהים. לכן, בד״כ נרצה לסנן את התמונה לפני גזירתה, כלומר להעבירה במסנן LPF. נאמר שאנו מסנים את התמונה במסנן הממצע שראינו קודם :

$$h_1 = \frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

לכן, בסיכומו של דבר, כדי לגזור נרצה להעביר את התמונה בשני המסננים הללו :

$$\frac{\partial}{\partial x} f[m,n] = (f * h_1) * h = \frac{1}{18} f * \begin{bmatrix}1 & 0 & -1\end{bmatrix} * \begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix}$$

$$= \frac{1}{18} f \begin{bmatrix}1 & 1 & 0 & -1 & -1\\ 1 & 1 & 0 & -1 & -1\\ 1 & 1 & 0 & -1 & -1\end{bmatrix}$$

ואז המסנן האחד שקיבלנו הוא BPF – Band Pass Filter, מכיוון שהוא מסנן תדרים נמוכים (התרומה של המסנן הממצע) ומסנן תדרים גבוהים (התרומה של המסנן הגוזר).

לצורך גילוי שפות, נקרב כעת את הנגזרת השנייה ע״י חישוב ההפרש בין הנגזרת הקדמית לנגזרת האחורית (שקול לביצוע נגזרת קדמית על תמונת הנגזרת עצמה) :

$$\frac{\partial^2}{\partial y^2} f[m,n] \approx \left(f[m+1,n]-f[m,n]\right)-\left(f[m,n]-f[m-1,n]\right)$$

$$= f[m+1,n]-2f[m,n]+f[m-1,n]$$

ובכיוון השורות נקבל :

$$\frac{\partial^2}{\partial x^2} f[m,n] \approx f[m,n+1]-2f[m,n]+f[m,n-1]$$

וכמו בנגזרת הראשונה, הפעולה האחרונה שקולה לביצוע קונבולוציה עם המסנן $h=\begin{bmatrix}1 & -2 & 1\end{bmatrix}$.

ניתן לקבל את אופרטור הלפלסיאן

$$\nabla^2 f(x,y) = \frac{\partial^2 f(x,y)}{\partial x^2} + \frac{\partial^2 f(x,y)}{\partial y^2}$$

בעזרת קונבולוציה עם המטריצה :

$$h_{lap} = \begin{bmatrix}0 & 1 & 0\\ 1 & -4 & 1\\ 0 & 1 & 0\end{bmatrix}$$

הלפלסיאן הינו מסנן HPF, מהיותו גוזר (ניתן לראות בקלות כי איזורים בעלי ערכים אחידים (תדר DC – ללא שינויים) נעלמים.

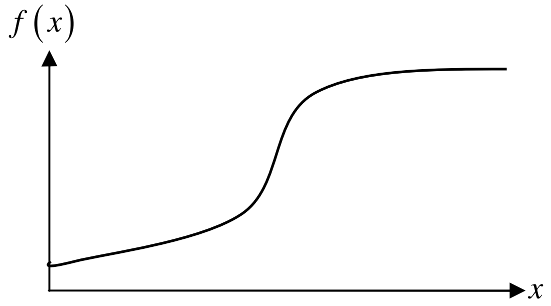
ניתן לקבל מסנן HPF גם בעזרת הפעולה הבאה :

$$h_{HPF} = \delta[m,n] - h_{LPF} = \begin{bmatrix}0 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 0\end{bmatrix} - \frac{1}{9}\begin{bmatrix}1 & 1 & 1\\ 1 & 1 & 1\\ 1 & 1 & 1\end{bmatrix} = \frac{1}{9}\begin{bmatrix}-1 & -1 & -1\\ -1 & 8 & -1\\ -1 & -1 & -1\end{bmatrix}$$

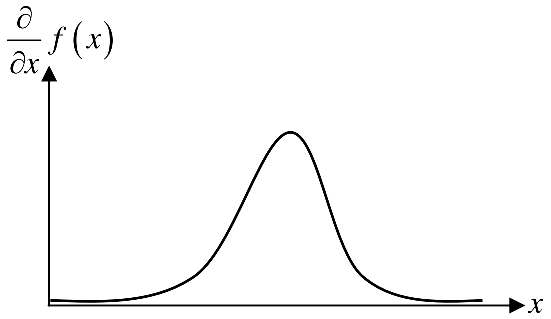
פעולה זו מסירה מהתמונה המקורית (קולבולוציה עם מסנן היחידה, δ) את הסינון הנמוך, ולכן מה שנותר הוא תמונת עם התדרים הגבוהים.

4.6. חידוד תמונה

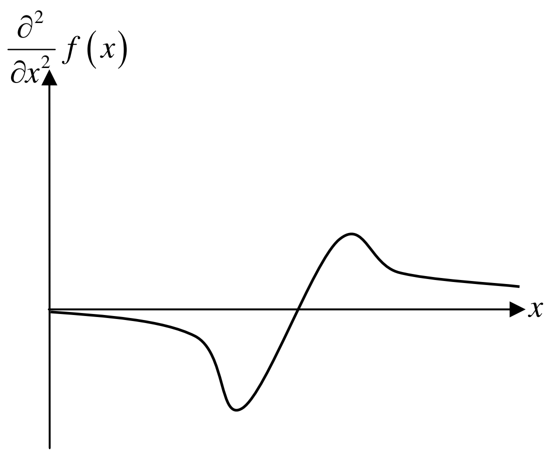
נשתמש בנגזרת כדי לבצע חידוד של תמונות.
נביט בפרופיל של גווני אפור f באזור שפה:



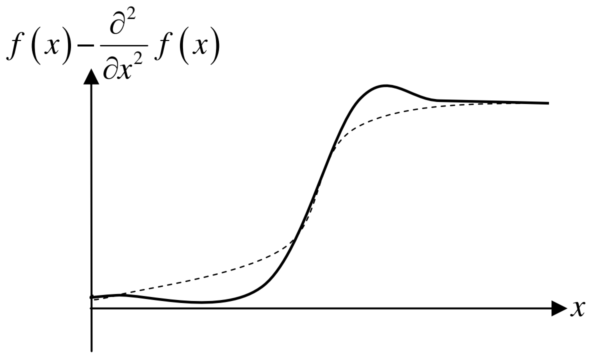
פרופיל הנגזרת באותו אזור:



והנגזרת השנייה:



אם נחסיר מהתמונה המקורית את הנגזרת השנייה, נקבל אפקט של הדגשת שפה (במקוקו הפונקציה f):



לפעולה זו קוראים Unsharp-Masking, ומתמטית ניתן לקבל אותה כך:

$$g[m,n] = f[m,n] - \alpha \cdot \nabla^2 f[m,n]$$

וכבר ראינו שניתן לממש פעולה זו בעזרת קונבולוציה:

$$g[m,n]=f[m,n]-\alpha\cdot\nabla^2 f[m,n]=f[m,n]*\left(\delta[m,n]-\alpha h_{lap}\right)$$
$$=f[m,n]*\left(\begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix}-\alpha\begin{bmatrix}0&1&0\\1&-4&1\\0&1&0\end{bmatrix}\right)=$$
$$=f[m,n]*\begin{bmatrix}0&-\alpha&0\\-\alpha&1+\alpha 4&-\alpha\\0&-\alpha&0\end{bmatrix}$$

4.7. גילוי שפות

כדי לגלות שפות בתמונה f , ראשית נחשב את נגזרתה השנייה:

$$g[m,n]=\left|\nabla^2 f[m,n]\right|$$

לאחר מכן נגדיר גוון סף T ונבצע פעולת סף.

ככל שהסף יהיה גבוה יותר, נקבל פחות איזורים הנחשבים שפות.

4.8. מסנן Wallis

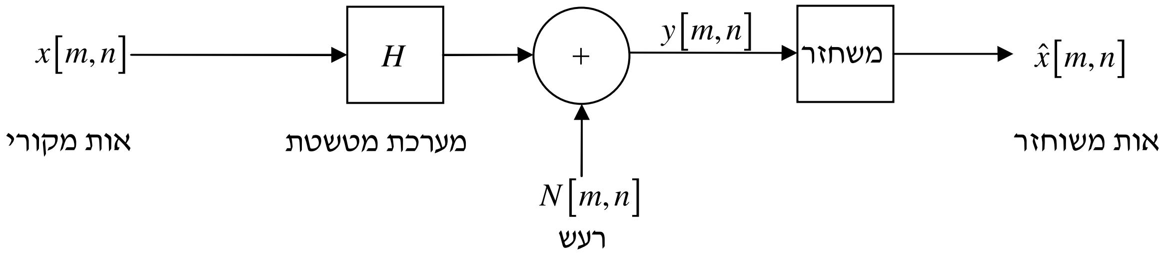
זהו מסנן אדפטיבי לתוכן התמונה: הוא עובד על איזורים קטנים ולכל איזור פועל בצורה המתאימה לו.

<להשלים אם צריך>

5. שחזור תמונות

5.1. בעית השחזור

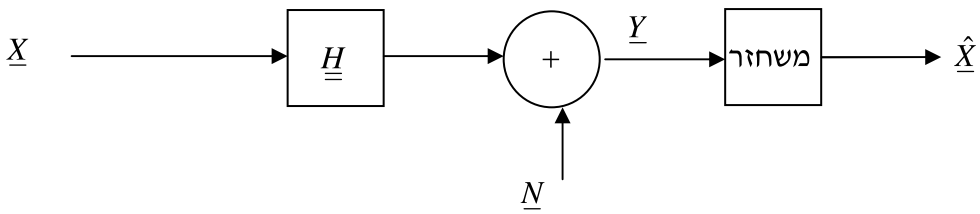
נרצה לשחזר תמונה שעברה הפרעה, ע"פ המודל הבא:



כבר ראינו בעבר מערכת שנועדה לקזז הפרעה משודרת – תיקון גאמא. במקרה זה יודעים את העיוות הנוצר, ולכן לא מסובך לתקנו.

כאשר H היא מערכת LSI, כלומר מתבצעת ע"י קונבולוציה, אזי לפעולת השחזור קוראים דה-קובולוציה (De-convolution).

נציג את הסכמה שלעיל כאשר המערכת הלינארית היא לא יותר מאשר מטריצה, והאותות (התמונות) מסודרים בוקטור, בסידור עמודה או בסידור שורה:



כאשר $\underline{X}$ היא התמונה המקורית ו $\underline{Y}$ היא התמונה שנשמרה בזיכרון המחשב:

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

במשוואה זו, $\underline{Y}$ ו $\underline{\underline{H}}$ נתונים, יכול להיות לנו ידע סטטיסטי על הרעש $\underline{N}$, והיינו רוצים לשחזר את $\underline{X}$.

על $\underline{X}$ אנו יודעים כמה פרטים, כמו למשל העובדה שערכיה ממשיים חיוביים, ואת מימדיה.

5.2. שערוך סטטיסטי

זהו שערון כאשר נתונים לנו נתונים סטטיסטיים לגבי התמונות $\underline{X}$ או $\underline{Y}$.

נסמן ב $P_{\underline{X}}\left(\underline{x}\right)$ את פונקצית צפיפות ההסתברות (PDF – Probability Density Function) של התמונה $\underline{X}$. מרחב המדגם שלנו הוא בעצם כל התמונות האפשריות. לדוגמה, ההסתברות לקיומה של תמונה מרוכבת היא 0.
באותו אופן, נסמן ב $P_{\underline{Y}}\left(\underline{y}\right)$ את פונקצית צפיפות ההסתברות של התמונה $\underline{Y}$.

בהנתן תמונת המקור $\underline{X}$, ניתן לשאול מה היא ההסתברות לקבלת תמונה $\underline{Y}$. פונקצית צפיפות ההסתברותה המתאימה לכך היא $P_{\underline{Y}|\underline{X}}\left(\underline{y} \mid \underline{x}\right)$. באופן מקביל, כאשר נתונת תמונת התוצאה $\underline{Y}$, נוכל לבדוק את ההסתברות לקבלת תמונה מקור $\underline{X}$: $P_{\underline{X}|\underline{Y}}\left(\underline{x} \mid \underline{y}\right)$.

5.3. משערך סבירות מירבית (Maximum Likelihood)

זהו משערך המביא למקסימום את ההסתברות לקבלת התמונה $\underline{Y}$, בהנתן תמונת מקור $\underline{X}$:

$$\hat{\underline{X}}_{ML} = \underset{\underline{X}}{\operatorname{argmax}}\, P_{\underline{Y}|\underline{X}}\left(\underline{y}\,|\,\underline{x}\right)$$

אם אין רעש, אזי

$$P_{\underline{Y}|\underline{X}}\left(\underline{y}\,|\,\underline{x}\right) = \delta\left(\underline{y} - \underline{\underline{H}}\underline{x}\right)$$

נניח כי הרעש הוא רל״ג (רעש לבן גאוסי) בעל ממוצע 0, כאשר c מקדם נרמול:

$$P_{\underline{N}}\left(\underline{n}\right) = c\exp\left\{-\frac{\left|\underline{n}\right|^2}{2\sigma_{\underline{N}}^2}\right\} = c\exp\left\{-\frac{\underline{n}^T\underline{n}}{2\sigma_{\underline{N}}^2}\right\}$$

ואז

$$P_{\underline{Y}|\underline{X}}\left(\underline{y}\,|\,\underline{x}\right) = P_{\underline{N}}\left(\underline{y} - \underline{\underline{H}}\underline{x}\right) = c\exp\left\{-\frac{\left(\underline{y} - \underline{\underline{H}}\underline{x}\right)^T\left(\underline{y} - \underline{\underline{H}}\underline{x}\right)}{2\sigma_{\underline{N}}^2}\right\} = c\exp\left\{-\frac{\left|\underline{y} - \underline{\underline{H}}\underline{x}\right|^2}{2\sigma_{\underline{N}}^2}\right\}$$

כאשר $\underline{y} - \underline{\underline{H}}\underline{x}$ הוא ההפרש בין האות הנתון לשערוך הדטרמיניסטי – והוא התפלגות הרעש (כשיש רעש, ל $\underline{y}$ יש סטייה מ $\underline{\underline{H}}\underline{x}$ שהיא בדיוק הסטייה של הרעש). נקבל

$$\hat{\underline{X}}_{ML} = \underset{\underline{X}}{\operatorname{argmax}}\, \exp\left\{-\frac{\left|\underline{y} - \underline{\underline{H}}\underline{x}\right|^2}{2\sigma_{\underline{N}}^2}\right\}$$

והאקספוננט מקסימלי כאשר החזקה מינימלית, לכן

$$\hat{\underline{X}}_{ML} = \underset{\underline{X}}{\operatorname{argmin}}\left\{\left|\underline{y} - \underline{\underline{H}}\underline{x}\right|^2\right\}$$

כלומר, אנחנו מחפשים $\underline{X}$ כך שאם נטשטש אותו ע״י מסנן $\underline{\underline{H}}$, נקבל משהו שמאוד דומה לנתון - $\underline{Y}$. נחפש את המינימום ע״י גזירה והשוואה לאפס:

$$\begin{aligned}
0 = \frac{\partial}{\partial x}\left|\underline{y} - \underline{\underline{H}}\underline{x}\right|^2 &= \frac{\partial}{\partial x}\left(\left(\underline{y} - \underline{\underline{H}}\underline{x}\right)^T\left(\underline{y} - \underline{\underline{H}}\underline{x}\right)\right) = -2\underline{\underline{H}}^T\left(\underline{y} - \underline{\underline{H}}\underline{x}\right)\\
\Rightarrow \quad & -\underline{\underline{H}}^T\underline{y} + \underline{\underline{H}}^T\underline{\underline{H}}\underline{x} = 0\\
\Rightarrow \quad & \underline{\underline{H}}^T\underline{\underline{H}}\underline{x} = \underline{\underline{H}}^T\underline{y}\\
\Rightarrow \quad & \hat{\underline{X}}_{ML} = \left(\underline{\underline{H}}^T\underline{\underline{H}}\right)^{-1}\underline{\underline{H}}^T\underline{y}
\end{aligned}$$

כאשר $\left(\underline{\underline{H}}^T\underline{\underline{H}}\right)^{-1}\underline{\underline{H}}^T$ היא הפסאודו-הפיכה של $\underline{\underline{H}}$, ואם $\underline{\underline{H}}$ מטריצה ריבועית והפיכה, אז מתנוונים למקרה של מטריצה הופכית

$$\hat{\underline{X}}_{ML} = \left(\underline{\underline{H}}^T\underline{\underline{H}}\right)^{-1}\underline{\underline{H}}^T\underline{y} = \underline{\underline{H}}^{-1}\left(\underline{\underline{H}}^T\right)^{-1}\underline{\underline{H}}^T\underline{y} = \underline{\underline{H}}^{-1}\underline{\underline{I}}\underline{y} = \underline{\underline{H}}^{-1}\underline{y}$$

אם $\underline{\underline{H}}^T\underline{\underline{H}}$ אינה הפיכה (ואז ל $\underline{\underline{H}}$ אין פסאודו-הפיכה), נבצע רגולריזציה: נגדיר מטריצה

$$\underline{\underline{H}}' = \left(\underline{\underline{H}}^T\underline{\underline{H}} + \varepsilon^2\underline{\underline{I}}\right)$$

כאשר ε קטן מאוד. מטריצה $\underline{\underline{H}}^T\underline{\underline{H}} + \varepsilon^2\underline{\underline{I}}$ תהיה הפיכה מכיוון ש $\underline{\underline{H}}^T\underline{\underline{H}}$ היא מטריצה שהערכים העצמיים שלה הם אי-שליליים, ואז הערכים העצמיים של $\underline{\underline{H}}^T\underline{\underline{H}} + \varepsilon^2\underline{\underline{I}}$ גדולים ממש מאפס עבור $\varepsilon \neq 0$.

5.4. משערך MAP (Maximum a-posteriori Probability)

הרעיון במשערך סטטיסטי זה הוא להתחשב במידע סטטיסטי מוקדם שיש לגבי האות שאותו מנסים לשחזר, $\underline{X}$, כלומר ידוע $P_{\underline{X}}\left(\underline{x}\right)$.

לפי חוק Bayes:

$$P_{\underline{X}|\underline{Y}}\left(\underline{x}\mid\underline{y}\right)=\frac{P_{\underline{Y}|\underline{X}}\left(\underline{y}\mid\underline{x}\right)P_{\underline{X}}\left(\underline{x}\right)}{P_{\underline{Y}}\left(\underline{y}\right)}=\frac{P_{\underline{X},\underline{Y}}\left(\underline{x},\underline{y}\right)}{P_{\underline{Y}}\left(\underline{y}\right)}$$

מכיוון ש $P_{\underline{Y}}\left(\underline{y}\right)$ לא תלוי בערכי $\underline{x}$, נקבל כי

$$\underline{\hat{X}}_{MAP}=\underset{\underline{X}}{\operatorname{argmax}}\,P_{\underline{X}|\underline{Y}}\left(\underline{x}\mid\underline{y}\right)=\underset{\underline{X}}{\operatorname{argmax}}\,P_{\underline{Y}|\underline{X}}\left(\underline{y}\mid\underline{x}\right)P_{\underline{X}}\left(\underline{x}\right)$$

ניתן לראות שאם $P_{\underline{X}}\left(\underline{x}\right)$ לא ידוע, כלומר אין ידע פריורי (ידע מוקדם) על $\underline{X}$, אנחנו מתנוונים חזרה לשערוך .Maximum Likelihood

לבעיה זו אין פתרון מתמטי, אלא פתרון הנדסי. לדוגמא, ניתן לומר שיש לנו ידע מקדים על $\underline{x}$, כלומר $P_{\underline{X}}\left(\underline{x}\right)$, בכך שהתמונה לא תקבל ערכים שליליים (כי הפיקסלים הם מספרים ממשיים חיוביים), ושאין בתמונה פיל אדום (או לפחות בסבירות מאוד נמוכה), או שהתמונה חלקה.
בד"כ תמונות הן חלקות, ולכן ניתן להשתמש בלפלסיאן. את פעולת הלפלסיאן ניתן לייצג ע"י אופרטור לינארי, כלומר מטריצת טופליץ, נסמנה $\underline{\underline{D}}$. כלומר, נשתמש ב

$$P_{\underline{X}}\left(\underline{x}\right)=c\exp\left\{-\frac{\left|\nabla^2\underline{x}\right|^2}{2\sigma_x^2}\right\}=c\exp\left\{-\frac{\left|\underline{\underline{D}}\underline{x}\right|^2}{2\sigma_D^2}\right\}$$

נוכל להעביר את התמונה באופרטור הלפלסיאן, ולראות עד כמה התוצאה קרובה לאפס, מכיוון שעבור תמונות חלקות, הלפלסיאן די מתאפס – וע"פ ההגדרה שלעיל, ההסתברות תהיה גבוהה לתמונה זו. אם יהיו הרבה שינויים בתמונה המוצעת בתור התמונה המשוחזרת, נקבל ערכים גבוהים ללפלסיאן, מה שיקטין את ההסתברות לתמונה מוצעת זו.
אם נניח כי הרעש הוא רל"ג, אזי כפי שראינו בפיתוחים הקודמים, נקבל כי

$$\underline{\hat{X}}_{MAP}=\underset{\underline{X}}{\operatorname{argmax}}\,P_{\underline{Y}|\underline{X}}\left(\underline{y}\mid\underline{x}\right)=\underset{\underline{X}}{\operatorname{argmax}}\left\{\exp\left\{-\frac{\left|\underline{y}-\underline{\underline{H}}\underline{x}\right|^2}{2\sigma_N^2}\right\}\exp\left\{-\frac{\left|\underline{\underline{D}}\underline{x}\right|^2}{2\sigma_D^2}\right\}\right\}$$

ואז יש למזער את הארגומנט של האקספוננטים, כלומר

$$\underline{\hat{X}}_{MAP}=\underset{\underline{X}}{\operatorname{argmin}}\left\{\frac{\left|\underline{y}-\underline{\underline{H}}\underline{x}\right|^2}{2\sigma_N^2}+\frac{\left|\underline{\underline{D}}\underline{x}\right|^2}{2\sigma_D^2}\right\}$$

נשים לב שהפתרון במודל Maximum Likelihood לא היה תלוי בשונות הרעש σ_N. במודל זה, σ_D הוא גודל שבשליטתנו: ככל שנרצה להרשות תמונות משוחזרות עם נגזרות גבוהות (שינויים חדים), נגדיל את σ_D.

במודל ה Maximum Likelihood, רצינו למצוא את ה $\underline{x}$ שיסביר הכי טוב את $\underline{y}$, כלומר את $\underline{x}$ שעבר טשטוש $\underline{\underline{H}}$. כלומר, אם התוצאה לא קרובה ל $\underline{y}$, נקבל "קנס". במודל MAP, ניתן לרשום, בהמשך לפיתוח שלעיל, כי

$$\underline{\hat{X}}_{MAP}=\underset{\underline{X}}{\operatorname{argmin}}\left\{\left|\underline{y}-\underline{\underline{H}}\underline{x}\right|^2+\frac{\sigma_N^2}{\sigma_D^2}\left|\underline{\underline{D}}\underline{x}\right|^2\right\}$$

ואז רואים כי נקבל "קנס" גם על היות התמונה לא חלקה מספיק (לפי הקריטריון σ_D שבחרנו). בשיטת MAP, ניתן להביא לידי ביטוי את ה trade-off בין דרישת ההתאמה לנתונים לבין הדרישה לחלקות ע"י הביטוי $\frac{\sigma_N^2}{\sigma_D^2}$.

נחפש את המינימום ע"י גזירה והשוואה לאפס:

$$0=\frac{\partial}{\partial x}\left[\left|\underline{y}-\underline{\underline{H}}\underline{x}\right|^2+\frac{\sigma_N^2}{\sigma_D^2}\left|\underline{\underline{D}}\underline{x}\right|^2\right]=-2\underline{\underline{H}}^T\left(\underline{y}-\underline{\underline{H}}\underline{x}\right)+2\frac{\sigma_N^2}{\sigma_D^2}\underline{\underline{D}}^T\underline{\underline{D}}\underline{x}$$

$$\Rightarrow\qquad -\underline{\underline{H}}^T\left(\underline{y}-\underline{\underline{H}}\underline{x}\right)+\frac{\sigma_N^2}{\sigma_D^2}\underline{\underline{D}}^T\underline{\underline{D}}\underline{x}=0$$

$$\Rightarrow\qquad -\underline{\underline{H}}^T\underline{y}+\underline{\underline{H}}^T\underline{\underline{H}}\underline{x}+\frac{\sigma_N^2}{\sigma_D^2}\underline{\underline{D}}^T\underline{\underline{D}}\underline{x}=0$$

$$\Rightarrow\qquad \underline{\hat{X}}_{MAP}=\left(\underline{\underline{H}}^T\underline{\underline{H}}+\frac{\sigma_N^2}{\sigma_D^2}\underline{\underline{D}}^T\underline{\underline{D}}\right)^{-1}\underline{\underline{H}}^T\underline{y}$$

המטריציה שיש להפוך כאן מזכירה את המטריצה שקיבלנו בשיטת Maximum Likelihood לאחר הרגולציה. במקרה הזה, המטריצה $\underline{\underline{D}}^T\underline{\underline{D}}$ מבצעת את פעולת הרגולציה. הרעיון הוא שאם $\underline{\underline{H}}^T\underline{\underline{H}}$ אינה הפיכה, אזי יש מספר תמונות מקור $\underline{x}$ שיכולות לתת את התמונה הנצפית $\underline{y}$, אבל כאן אנחנו בוחרים מכל האפשרויות האלו את התמונה הכי חלקה.

מעשית, לא משחזרים תמונות באופן שתארנו, בגלל הבעייתיות הטכנית בהיפוך של מטריצות גדולות.

5.4.1. שימוש ב Steepest Descent

אנחנו מחפשים את נקודת המינימום של הביטוי

$$J\left(\underline{x}\right)=\left|\underline{y}-\underline{\underline{H}}\underline{x}\right|^2+\frac{\sigma_N^2}{\sigma_D^2}\left|\underline{\underline{D}}\underline{x}\right|^2$$

מכיוון שאנחנו יודעים שקיימת נקודמת מינימום, ושהיא היחידה, נוכל להתסכל על הפונקציה בנקודה מסויימת, ואז לזוז לאורכה, בכיוון היוביל אותנו למינימום, כלומר בכיוון הפוך לגרדיאנט, בצעדים. באופן איטרטיבי, נבצע

$$\underline{\hat{x}}_{k+1}=\underline{\hat{x}}_k-\mu\left.\frac{\partial J\left(\underline{x}\right)}{\partial\underline{x}}\right|_{\underline{x}=\hat{x}_k}$$

כאשר μ הוא גודל הצעד אותו עושים בכל פעם.

6. דחיסת תמונות

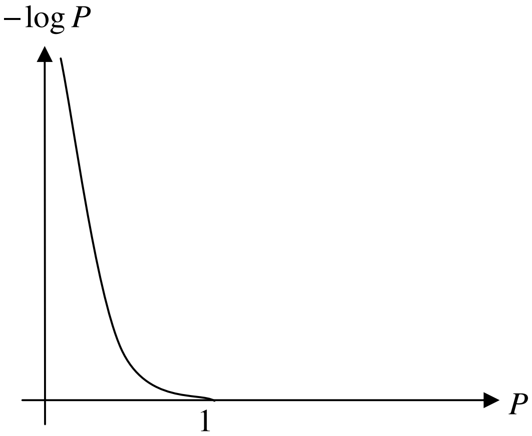
6.1. תורת האינפורמציה

6.1.1. אינפורמציה

ההסתברות לאירוע E היא $0 \le P(E) \le 1$. מידת האינפורמציה I באירוע, ביחידות של ביטים, היא

$$I(E)=\log_2 \frac{1}{P(E)}[bits]=-\log_2 P(E)[bits]$$

האינפורמציה היא מדד לאי-הוודאות של האירוע. ככל שהאירוע וודאי יותר, הוא אינו נושא אינפורמציה.



עבור שני אירועים בלתי-תלויים E_1 ו E_2, האינפורמציה היא אדיטיבית (מתווספת):

$$I(E_1,E_2)=\log_2 \frac{1}{P(E_1,E_2)}=\log_2 \frac{1}{P(E_1)P(E_2)}=$$

$$=\log_2 \frac{1}{P(E_1)}+\log_2 \frac{1}{P(E_2)}=I(E_1)+I(E_1)$$

6.1.2. אנטרופיה

נסתכל במקור הפולט k תווים, כאשר לכל תו אלף-בית סופי. כלומר הערכים שכל תו מקבל הם בדידים וסופיים:

$$A_0=\{a_i\}_{i=1}^{k}$$

התווים הם בלתי תלויים ביניהם. האינפורמציה הממוצעת של המקור הפולט את התויים הבלתי תלויים:

$$H(u)=-\sum_{u\in A_0} \log_2 (P_U(u))P_U(u)$$

דוגמא

מקור בינארי חסר זיכרון (הטלת מטבע):

$$P_U(u=1)=p$$

$$P_U(u=2)=1-p$$

האנטרופיה:

$$H(u)=-\sum_{u\in A_0} \log_2 (P_U(u))P_U(u)=-p\log_2 p-(1-p)\log_2(1-p)$$

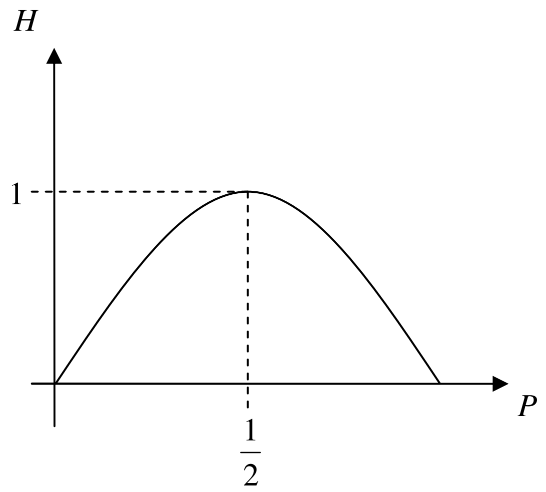
יחידות האנטרופיה הם $\left[\frac{bits}{symbol}\right]$.

במקרה של וודאות מלאה, כלומר $p=0$ או $p=1$, לא תהיה אינפורמציה מהמאורעות, ולכן גם האנטרופיה שנקבל היא אפס:

$$H(u)=-1\log_2 1-0\log_2 0=0$$

במקרה של מטבע הוגן, $p=\frac{1}{2}$, נקבל אינפורמציה מלאה מהאירועים, ולכן אנטרופיה מקסימלית:

$$H(u)=-\frac{1}{2}\log_2\frac{1}{2}-\frac{1}{2}\log_2\frac{1}{2}=-\log_2\frac{1}{2}=1$$



באופן כללי:

$$0\le H(u)\le \log_2 k$$

כאשר ההתפלגות היא אחידה, אי הוודאות מקסימלית, ואז מקבלים אנטרופיה מקסימלית:

$$H(u)=-\sum_{u\in A_0}\log_2\left(P_U(u)\right)P_U(u)=-\sum_{u\in A_0}\frac{1}{k}\log_2 k=\log_2 k$$

6.2. משפט ההצפנה של Shannon

המשפט:

1. מקור u בעל אנטרופיה $H(u)$ לא ניתן לייצוג במספר סיביות ממוצע לסימבול הקטן מ $H(u)$.
2. לכל $\varepsilon>0$ קטן כרצוננו, קיים עקרונית אלגוריתם דחיסה לקצב נמוך מ $H(u)+\varepsilon$.

משפט Shannon מציב חסם תחתון לביצועים של דחיסה משמרת. כלומר כדי לדחוס אות (מקור לביטים) בצורה שתשמור על כל המידע, האנטרופיה שלו, $H(u)$, היא חסם תחתון למספר הביטים לסימבול שנצטרך להשתמש בהם. המשפט גם אומר שניתן להתקרב לאותו חסם כרצוננו, אך אינו מפרט את האלגוריתם.

מדובר על הקצב הממוצע – מספר הביטים הממוצע לייצוג סימבול (פיקסל). לא הכרחי כי לכל פיקסל נשתמש מספר ביטים אחיד (לדוגמה, קוד הופמן), אבל הממוצע המשוקלל של מספר הביטים הנדרשים לייצוג כל פיקסל – הוא החסם המדובר.

6.2.1. חוסר תלות בין תווים
כשדיברנו על אנטרופיה, אמרנו שהמקור (התמונה שלנו) פולט סימבולים (הפיקסלים בתמונה) באופן בלתי תלוי. אבל, כשמדובר בתמונה, בד"כ יש קשר בין פיקסלים שכנים: אם פיקסל היה כהה, יש סבירות גבוהה שגם הפיקסל שלידו כהה.

6.2.2. תלות בין תווים
נביט במקור הפולט סדרה אקראית של וקטורים

$$\underline{X}=\left[x_1 \quad x_2\right]$$

כאשר x_1 ו x_2 הם חלק מהאלף-בית של המקור, כלומר $x_1,x_2\in\left\{a_i\right\}_{i=1}^{k}$.

נאמר שזוגות המספרים הם זוגות של פיקסלים עוקבים בתמונה, ולכן קיימת תלות בין רכיבי הוקטור. עדיין בין כל שני וקטורים, יש אי-תלות. כלומר סדרת הוקטורים היא IID = Independent Indentically Distributed.

לכל צמד מספרים יש הסתברות, ולכן האנטרופיה במקרה זה:

$$H\left(\underline{X}\right)=H\left(\left[x_1 \quad x_2\right]\right)=-\sum_{i,j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\log_2 P_{x_1,x_2}\left(a_i,a_j\right)$$

אם הפיקסלים הצמודים x_1 ו x_2 היו בלתי-תלויים, ההסתברות $P_{x_1,x_2}\left(a_i,a_j\right)$ היתה מכפלת ההסתברויות $P_{x_1}\left(a_i\right)P_{x_2}\left(a_j\right)$, אך לא כך הדבר. ע"י שימוש בחק Bayes נקבל

$$\begin{aligned}H\left(\underline{X}\right)&=-\sum_{i,j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\log_2 P_{x_1}\left(a_i\right)P_{x_2|x_1}\left(a_j\,|\,a_i\right)\\&=-\sum_{i,j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\left(\log_2 P_{x_1}\left(a_i\right)+\log_2 P_{x_2|x_1}\left(a_j\,|\,a_i\right)\right)\\&=-\sum_{i,j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\log_2 P_{x_1}\left(a_i\right)-\sum_{i,j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\log_2 P_{x_2|x_1}\left(a_j\,|\,a_i\right)\end{aligned}$$

נפתח את הסכום הראשון:

$$\begin{aligned}S_1&=-\sum_{i=1}^{k}\sum_{j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\log_2 P_{x_1}\left(a_i\right)=-\sum_{i=1}^{k}\sum_{j=1}^{k}P_{x_1}\left(a_i\right)P_{x_2|x_1}\left(a_j\,|\,a_i\right)\log_2 P_{x_1}\left(a_i\right)\\&=-\sum_{i=1}^{k}P_{x_1}\left(a_i\right)\log_2 P_{x_1}\left(a_i\right)\underbrace{\sum_{j=1}^{k}P_{x_2|x_1}\left(a_j\,|\,a_i\right)}_{1}=-\sum_{i=1}^{k}P_{x_1}\left(a_i\right)\log_2 P_{x_1}\left(a_i\right)=H\left(x_1\right)\end{aligned}$$

נפתח את הסכום השני:

$$\begin{aligned}S_2&=-\sum_{i=1}^{k}\sum_{j=1}^{k}P_{x_1,x_2}\left(a_i,a_j\right)\log_2 P_{x_2|x_1}\left(a_j\,|\,a_i\right)=-\sum_{i=1}^{k}\sum_{j=1}^{k}P_{x_1}\left(a_i\right)P_{x_2|x_1}\left(a_j\,|\,a_i\right)\log_2 P_{x_2|x_1}\left(a_j\,|\,a_i\right)\\&=\sum_{i=1}^{k}P_{x_1}\left(a_i\right)\left(-\sum_{j=1}^{k}P_{x_2|x_1}\left(a_j\,|\,a_i\right)\log_2 P_{x_2|x_1}\left(a_j\,|\,a_i\right)\right)=\sum_{i=1}^{k}P_{x_1}\left(a_i\right)H\left(x_2\,|\,x_1=a_i\right)=H\left(x_2\,|\,x_1\right)\end{aligned}$$

כלומר

$$H\left(\underline{X}\right)=H\left(\left[x_1 \quad x_2\right]\right)=H\left(x_1\right)+H\left(x_2\,|\,x_1\right)$$

כאשר ידוע כי

$$H\left(x_2\right)\geq H\left(x_2\,|\,x_1\right)$$

אינטאיטיבית, מכיוון ש $H\left(x_2\right)$ היא מדד לחוסר הוודאות של הפיקסל השני x_2, אז $H\left(x_2\right)\geq H\left(x_2\,|\,x_1\right)$ מכיוון שבהינתן הפיקסל הראשון, חוסר הוודאות לגבי הפיקסל השני קטן. כאשר כל הפיקסלים בלתי תלויים, יתקבל השוויון.

מכיוון שלכל הפיקסלים אותה סטטיסטירה, אזי $H\left(x_1\right)=H\left(x_2\right)$, ולכן קיבלנו

$$H\left(\underline{X}\right)=H\left(x_1\right)+H\left(x_2\,|\,x_1\right)\leq H\left(x_1\right)+H\left(x_2\right)=2H\left(x_1\right)$$

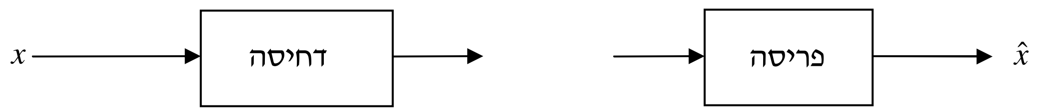
לכן, אם נקודד זוגות פיקסלים, נטפל במידע עם פחות אנטרופיה, וע"פ משפט הצפינה, נוכל להשתמש בפחות ביטים.

נכליל את התוצאה שלעיל עבור מקור הפולט N תווים, ונקבל

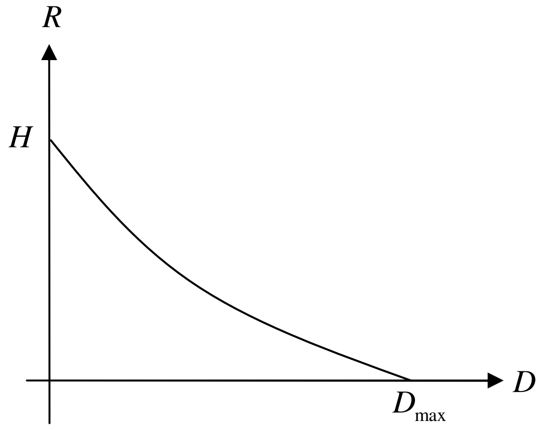
$$H\left(\underline{X}\right)\leq\sum_{n=1}^{N}H\left(x_n\right)$$

6.3. דחיסה לא משמרת

מערכת כללית לדחיסה ופריסה מחדש של המידע הדחוס נראית כך :



כאשר הדחיסה משמרת, המידע הנפרס לאחר הדחיסה זהה למידע המקורי.
אחרת, נוכל להגדיר מדד לעיוות המתרחש $d\left(x,\hat{x}\right)$, ונרצה שהעיוות הממוצע, כלומר $E\left[d\left(x,\hat{x}\right)\right]$ לא יעבור חסם מסויים. נציג את פונקציית קצב עיוות :



פונקציה זו (תמיד קמורה ממש) מתארת את העיוות במידע כפונקציה של קצב האינפורמציה (כמות הביטים לסימבול). כאמור, אם לא נרצה עיוות כלל, נצטרך להתשמש בקצב אינפורמציה השווה לאנטרופיה של המקור.

ככל שנבצע קוונטיזציה גסה יותר, נוריד את מספר הביטים לייצוג הסימבולים, ונוסיף לעיוות. נראה כי ככל שנבצע קוונטיזציה גסה יותר ויותר, העיוות יגדל, אבל בעצם יש חסם תחתון לכמות הביטים שצריך כדי לייצג סימבול.
לדוגמא, נבצע קוונטיזציה לתמונה בעלת 128 רמות אפור, כאשר לקוונטייזר רמת ייצוג אחת. נבחר רמת אפור 64 (כך העיוות הממוצע המחושב יהיה קטן יותר מאשר אם היינו בוחרים את רמת האפור 0). קיבלנו תמונה עם רמת אפור אחת, ולכן איבדנו את כל האינפורמציה. כלומר, יש רמה של קוונטיזציה שמעבר לה אין טעם להסתכל על העיוות, כל אל האינפורמציה כבר אבדה לנו.

6.4. קידוד תמונות

דרכים להורדת אנטרופיה :

6.4.1. קידוד Huffman

אלגוריתם הקידוד :

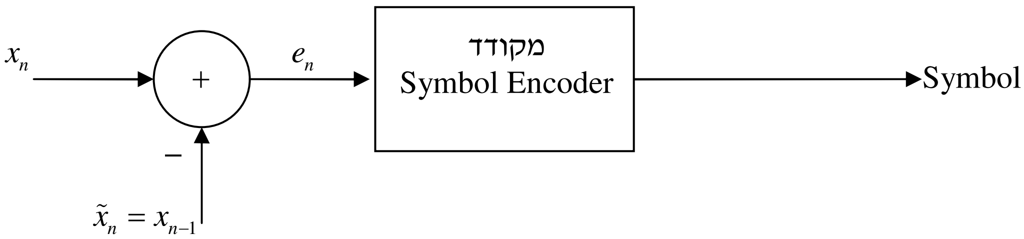
1. סדר את ההסברויות לקבלת כל סימבול בסדר יורד. כל סימבול יהיה עלה בעץ שאנחנו נבנה כעת.
2. כל עוד יש יותר מצומת אחד :
   - חבר את שני הצמתים בעלי ההסתברויות הנמוכות ביותר לצומת חדש אשר ההסברות שלו הוא סכום ההסברויות של שני הצמתים שחיברנו.
   - הקצה סיבית '1' לענף התחתון החדש וסיבית '0' לענף העליון החדש.
3. הקוד לכל סימבול הוא הסיביות הרשומות על כל ענף בדרך לעלה שלו, כאשר מתחילים משורש העץ.

6.4.2. ניצול פילוג לא אחיד

ראינו כי ככל שהפילוג אחיד יותר, האנטרופיה גדולה יותר.

6.4.3. ניצול תלות בין פיקסלים סמוכים

נשתמש במערכת הבאה ליצירת הסימבולים :



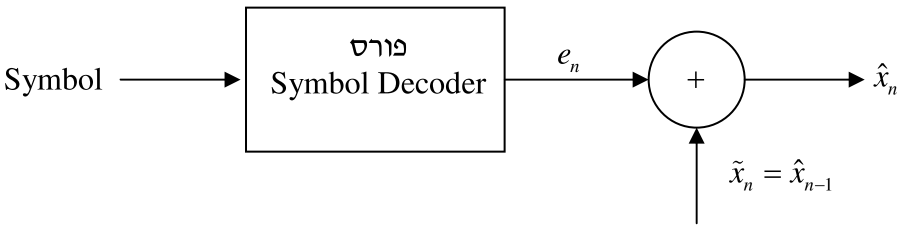
את הפיקסל הראשון לא נדחוס, ונשמור את ערכו (כלומר החזאי לפיקסל הראשון הוא אפס, $\tilde{x}_1 = 0$). נשתמש בערך הפיקסל הראשון כדי לחזות את הערך של הפיקסל השני. כך נמשיך, וכל פיקסל ישמש חזאי לפיקסל הבא בתור.
הפילוג שנקבל זהה לנגזרת של הפילוג של התמונה המקורית. נקבל פילוג צר יחסית סביב ה 0. את ערכי ההפרשים, e_n, נעביר למקודד (הקידוד יכול להתבצע, למשל, ע״י קידוד הופמן). מכיוון שיצרנו פילוג צר יותר בכניסה למקודד הסימבולים, האנטרופיה (חוסר הוודאות) קטנה, וכך ניתן לשמור את הנתונים, עדיין במצב Lossless (דחיסה משמרת – ללא איבוד אינפורמציה), בעזרת פחות ביטים לפיקסל.

כך יראו ערכי הפיקסלים בתמונה המקורית, כאשר אנחנו כרגע שומרים את ערכו של הפיקסל השביעי, בתמונה בגודל 4×5:

x_{n-6}	x_{n-5}	x_{n-4}	x_{n-3}	x_{n-2}
x_{n-1}	x_n			

נוכל לחזות באופן אולי יותר יעיל, אם במקום להסתמך רק על הפיקסל הקודם בתור חזאי, ניקח ממוצע בין הפיקסלים הקודמים שסמוכים לפיקסל החדש שעליו מסתכלים, וניקח את הממוצע הזה בתור חזאי. עבור המטריצה שלעיל, ניקח את הממוצע $\frac{x_{n-6} + x_{n-5} + x_{n-1}}{3}$ כדי לשערך את ערכו של x_n.

מכיוון שקידדנו את הפיקסלים בצורת הפרשים, יש לבנות מערכת פריסה מתאימה:



הפיקסל הראשון נתון לנו ללא הדחיסה. כדי לחשב את הפיקסל השני, יש להוסיף לשגיאה (השגיאה היא בעצם הסימבול שקודדנו בתהליך הקידוד) את ערך הפיקסל הראשון. כך הלאה – לחישוב כל פיקסל, נסכום את ערך הפיקסל המשוחזר הקודם עם השגיאה של הפיקסל הנוכחי.

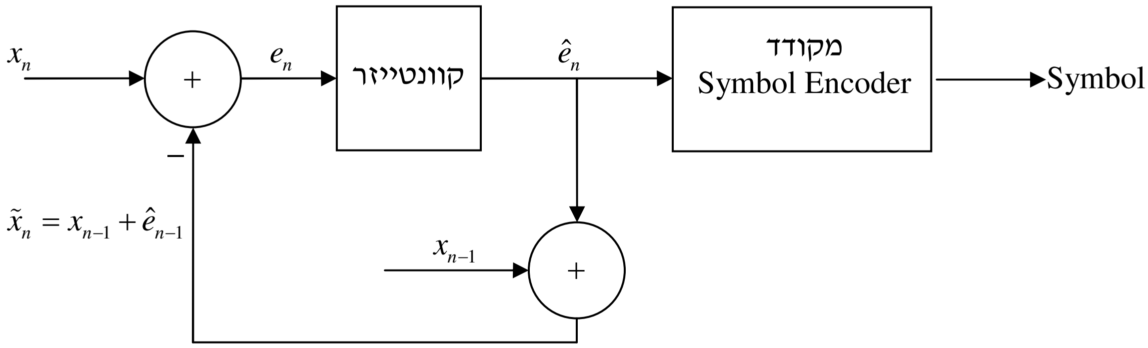
עד כאן הקטנו את האנטרופיה (את המספר הממוצע של ביטים שיש לקודד לכל סימבול) ע״י ציפוף ההיסטוגרמה של המקור. כאמור, פעולות אלו לא גרמו לאיבוד מידע, ולכן זוהי דחיסת Lossless.

כעת אנחנו רוצים להוריד עוד את קצב האינפורמציה, וזאת במחיר איבוד מידע, ע״י קוונטיזציה של רמות האפור המתקבלות לאחר ציפוף ההיסטוגרמה ע״י ביצוע ההפרשים.

כאשר לקחנו את הפיקסל הקודם להיות הערך החזוי של הפיקסל הנוכחי, הנחנו בעצם שהפיקסל הקודם שוחזר במדויק, אך זה לא המקרה כאשר הערכים המשוחזרים של הפיקסלים הם רמות הייצוג – כלומר הם כוללים את שגיאת הקוונטיזציה. אם נפעיל את אותו השחזור שהגדרנו מקודם על האות שעבר קוונטיזציה, שגיאות הקוונטיזציה יתווספו לתהליך השחזור, ומהר מאוד (לאחר שחזור של מספר פיקסלים לא גדול) התמונה המשוחזרת כבר לא תהיה דומה למקור.

6.4.4. חיזוי כולל קוונטיזציה DPCM – Differencial Pulse Code Modulation

דחיסה:



את הפיקסל הראשון לא נדחוס, ולא נבצע עליו קוונטיזציה. כדי לחשב את הפיקסל השני, יש להוסיף לשגיאת הקוונטיזציה, $\hat{e}_n$, את ערך הפיקסל הראשון. כך הלאה – לחישוב כל פיקסל, נחשב את ההפרש בין הפיקסל הנוכחי לבין החזאי לפיקסל הנוכחי. החזאי לפיקסל הנוכחי, $\tilde{x}_n$, הוא סכום של ערך הפיקסל הקודם עם שגיאת הקוונטיזציה של הפיקסל הקודם

.

תהליך הפריסה נותר זהה לתהליך בו לא ביצענו קוונטיזציה.

6.5. ניצול מגבלות מערכת הראייה האנושית

ניתן, בעזרת התמרה לינארית מתאימה, לקבל וקטור מקדמים אשר ניתן בעזרתו לשחזר את התמונה המקורית.
כדי לנצל את מגבלות מערכת הראייה האנושית, נבצע קוונטיזציה לוקטור המקדמים, כך שתבוצע קוונטיזציה גסה למקדמים שמערכת הראייה פחות רגישה אליהם, וקוונטיזציה עדינה למקדמים שהעין רגישה אליהם יותר.

7. התמרות

7.1. התמרה ספרבילית

מערכת לינארית תקרא ספרבילית כאשר

$$g[m,n]=\sum_{k=-\infty}^{\infty}\sum_{r=-\infty}^{\infty}h_1[m,k]h_2[n,r]f[k,r]$$
$$=\sum_{k=-\infty}^{\infty}h_1[m,k]\left(\sum_{r=-\infty}^{\infty}h_2[n,r]f[k,r]\right)$$

כלומר, ניתן לסכום בנפרד על העמודות ועל השורות.

כשהמערכת ספרבילית, ניתן לכתוב את הקשר בין תמונת הכניסה לתמונת היציאה כך

$$\underline{\underline{g}}_{M\times N}=\underline{\underline{H}}_{1_{M\times M}}\cdot\underline{\underline{f}}_{M\times N}\cdot\underline{\underline{H}}^T_{2_{N\times N}}$$

ובסידור לקסיקורגפי:

$$\underline{g}^{rs}=\underline{\underline{H}}_{MN\times MN}\underline{f}^{rs}$$

ואם המערכת ספרבילית, אזי

$$\underline{\underline{H}}=\underline{\underline{H}}_1\otimes\underline{\underline{H}}_2$$

או בצורה

$$\underline{g}^{cs}=\tilde{\underline{\underline{H}}}\underline{f}^{cs},\qquad \tilde{\underline{\underline{H}}}=\underline{\underline{H}}_2\otimes\underline{\underline{H}}_1$$

כאשר $\underline{\underline{H}}_1=\underline{\underline{H}}_2$, נקבל $\tilde{\underline{\underline{H}}}=\underline{\underline{H}}$.

7.1.1. מספר פעולות כפל

כדי לכפול שורה אחת ב $\underline{\underline{f}}_{M\times N}$ עם עמודה אחת ב $\underline{\underline{H}}_{2_{N\times N}}$, יש לבצע N^2 כפלים, לכן יש לבצע סה"כ MN^2 כפלים.

עבור ההכפלה ב $\underline{\underline{H}}_{1_{M\times M}}$ נקבל עוד M^2N פעולות, לסיכום יש $MN^2+M^2N=MN(M+N)$.

אם נתעלם מהספרביליות, נצטרך להכפיל את $\underline{\underline{H}}_{MN\times MN}$ ב $\underline{f}^{rs}$, יש צורך ב $(MN)^2$.

ואז נעדיף לנצל את הספרביליות, ואז נכפול פי $\frac{MN(M+N)}{(MN)^2}=\frac{M+N}{MN}$ פחות פעמים.

7.2. התמרות יוניטריות

זוהי התמרה ע"י מטריצה $\underline{\underline{A}}$ יוניטרית, כלומר מטריצה המקיימת את התכונה הבאה:

$$\underline{\underline{A}}^{-1}=\underline{\underline{A}}^H\triangleq\left(\underline{\underline{A}}^*\right)^T$$

הפעלת ההתמרה ע"י ההכפלה במטריצה. כלומר עבור תמונה המיוצגת בסידור לקסיקוגרפי $\underline{v}$, וקטור מקדמי הייצוג שלה, במרחב ההתמרה הוא

$$\underline{u}=\underline{\underline{A}}\underline{v}$$

וההתמרה ההפוכה:

$$\underline{v}=\underline{\underline{A}}^H\underline{u}$$

נכתוב את ההתמרה ההפוכה בצורה שונה:

$$\begin{bmatrix}v_1\\v_2\\\vdots\\\vdots\\v_{MN}\end{bmatrix}_{MN\times1}=\begin{bmatrix}\begin{pmatrix}\vdots\\\underline{a}_1\\\vdots\end{pmatrix}&\begin{pmatrix}\vdots\\\underline{a}_2\\\vdots\end{pmatrix}&\cdots&\begin{pmatrix}\vdots\\\underline{a}_{MN}\\\vdots\end{pmatrix}\end{bmatrix}_{MN\times MN}\begin{bmatrix}u_1\\u_2\\\vdots\\\vdots\\u_{MN}\end{bmatrix}_{MN\times1}$$

כלומר, $\underline{a}_k$ הן עמודות מטריצת ההתמרה $\underline{\underline{A}}$. מקבלים כי

$$\underline{v}=\sum_{k=1}^{MN}u_k\cdot\underline{a}_k$$

כלומר, ניתן לכתוב את התמונה המקורית $\underline{v}$ כסכום של תמונות בסיס $\left\{\underline{a}_k\right\}_{k=1}^{MN}$.

ניתן לראות שמתקיים משפט פרסבל:

$$\left|\underline{u}\right|^2=\left|\underline{\underline{A}}\,\underline{v}\right|^2=\left(\underline{\underline{A}}\,\underline{v}\right)^H\left(\underline{\underline{A}}\,\underline{v}\right)=\underline{v}^H\underline{\underline{A}}^H\underline{\underline{A}}\,\underline{v}=\underline{v}^H\underline{\underline{I}}\,\underline{v}=\underline{v}^H\underline{v}=\left|\underline{v}\right|^2$$

7.3. התמרה יוניטרית ספרבילית

ראינו שמערכת היא ספרבילית כאשר ניתן לרשום:

$$\underline{\underline{g}}_{M\times N}=C_{M\times M}\cdot\underline{\underline{f}}_{M\times N}\cdot D_{M\times M}^T$$

כאשר $\underline{\underline{C}}$ ו $\underline{\underline{D}}$ יוניטריות, ניתן למצוא את ההתמרה ההפוכה:

$$\underline{\underline{f}}=\underline{\underline{C}}^H\underline{\underline{g}}\,\underline{\underline{D}}^*$$

ואז ניתן לכתוב גם

$$\underline{f}^{rs}=\left(\underline{\underline{C}}^H\otimes\underline{\underline{D}}^*\right)\underline{g}^{rs}$$

(וכמובן במימוש הפעולה נעדיף לנצל את הספרביליות ולבצע את $\underline{\underline{f}}=\underline{\underline{C}}^H\underline{\underline{g}}\,\underline{\underline{D}}^*$).

7.3.1. תאור 2D-DFD כמכפלת קרוניקר

הגדרת התמרת DFD בדו-מימד:

$$F\left[k,r\right]=2D-DFT\left\{f\left[m,n\right]\right\}=\sum_{m=0}^{M-1}\sum_{n=0}^{N-1}f\left[m,n\right]W_M^{-mk}W_N^{-nr}$$

כאשר

$$W_N=e^{j\frac{2\pi}{N}}$$

קל לראות שזאת פעולה ספרבילית.

נגדיר שתי מטריצות בגודל $M\times N$:

$$\left\{\underline{\underline{W}}_M\right\}_{m,k}=W_M^{-mk}$$
$$\left\{\underline{\underline{W}}_N\right\}_{m,k}=W_N^{-mk}$$

המטריצות סימטריות מכיוון ש

$$\left\{\underline{\underline{W}}_M\right\}_{m,k}=W_M^{-mk}=W_M^{-km}=\left\{\underline{\underline{W}}_M\right\}_{k,m}$$
$$\left\{\underline{\underline{W}}_N\right\}_{m,k}=W_N^{-mk}=W_N^{-km}=\left\{\underline{\underline{W}}_N\right\}_{k,m}$$

כלומר

$$\underline{\underline{W}}_M^T=\underline{\underline{W}}_M$$
$$\underline{\underline{W}}_N^T=\underline{\underline{W}}_N$$

ואז ניתן לכתוב

$$\underline{\underline{F}}=\underline{\underline{W}}_M\underline{\underline{f}}\,\underline{\underline{W}}_N$$

כאשר $\underline{\underline{F}}$ היא התמונה בתחום התדר, ואז

$$\underline{F}^{rs}=\left(\underline{\underline{W}}_M\otimes\underline{\underline{W}}_N\right)\underline{f}^{rs}$$

וההתמרה ההפוכה:

$$\underline{\underline{f}}=\frac{1}{MN}\underline{\underline{W}}_M^*\underline{\underline{F}}\,\underline{\underline{W}}_N^*$$

ובסידור לקסיקוגרפי:

$$\underline{f}^{rs} = \frac{1}{MN}\left(\underline{\underline{W}}_M^* \otimes \underline{\underline{W}}_N^*\right)\underline{F}^{rs}$$

כאשר הפקטור $\frac{1}{MN}$ קיים כדי "לתקן" את היחס באנרגיות של התמונה בין מישור המקום למישור התדר. כלומר – התמרה זו אינה יוניטרית. נוכל להגדיר התמרה דומה, יוניטרית:

$$\tilde{\underline{\underline{W}}}_M = \frac{1}{\sqrt{M}}\underline{\underline{W}}_M$$

$$\tilde{\underline{\underline{W}}}_N = \frac{1}{\sqrt{N}}\underline{\underline{W}}_N$$

ונקבל התמרה יוניטרית:

$$\tilde{\underline{\underline{F}}} = \tilde{\underline{\underline{W}}}_M \underline{\underline{f}} \tilde{\underline{\underline{W}}}_N \quad \Rightarrow \quad \underline{\underline{f}} = \tilde{\underline{\underline{W}}}_M^* \underline{\underline{F}} \tilde{\underline{\underline{W}}}_N^*$$

ראינו כי מספר פעולות הכפל כאשר מנצלים את ספרביליות ההתמרה הוא $MN\left(M+N\right)$, אבל במקרה של התמרת DFT דו מימדית, ניתן להשתמש באלגוריתם FFT הידוע, ובכך לקבל סיבוכיות של $MN\log_2 MN$.

7.4. התמרת DCT – Discrete Cosine Transform

זוהי התמרה יוניטרית ספרבילית, המוגדרת ע"י:

$$\left\{\underline{\underline{C}}_N\right\}_{k,n} = \begin{cases} \frac{1}{\sqrt{N}}, & k=0 \\ \sqrt{\frac{2}{N}}\cos\left(\frac{\pi\left(2n+1\right)k}{2N}\right), & else \end{cases}$$

הפעלת ההתמרה ע"י:

$$\underline{\underline{F}}_{(M\times N)} = \underline{\underline{C}}_{M_{(M\times M)}} \underline{\underline{f}}_{(M\times N)} \underline{\underline{C}}^T_{N_{(N\times N)}}$$

וההתמרה ההפוכה:

$$\underline{\underline{f}} = \underline{\underline{C}}_M^T \underline{\underline{F}}\underline{\underline{C}}_N$$

משורה לשורה, תדר הקוסינוס גדל (השורה הראשונה, עבור $k=0$, קבועה). כמו כן, יש תזוזה בפאזה של $\frac{\pi k}{2N}$ משורה לשורה.
התמרה זו מזכירה את התמרה פורייה, בה מטילים את התמונה על פונקציות בסיס שהינן מעריכיות מרוכבות.
השוואה בין DCT ל DFT:

- גם ב DFT התדר גדל משורה לשורה, אך אין שינוי בפאזה
- DFT מרוכבת, בעוד DCT ממשית
- בשתי ההתמרות המטריצות הן אורתונורמליות, ולכן יוניטריות

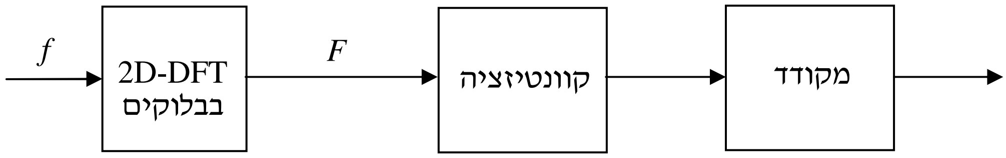
נביט בהתמרה ההפוכה:

$$\underline{\underline{f}} = \underline{\underline{C}}_M^T \underline{\underline{F}}\underline{\underline{C}}_N$$

אם נרצה לקבל את אחת מ MN תמונות הבסיס, יש לשים פיקסל אחד בעל ערך 1 בהתמרה $\underline{\underline{F}}$, ובשאר הפיקסלים 0 – כלומר הלם.

פונקציות קוסינוס אינן פונקציות עצמיות של פעולת הקונבולוציה, כמו אקספוננטים שראינו בהתמרת פורייה – אז מדוע כדי בכל זאת להשתמש בפונקציות אלו כדי לייצג תמונות? מסתבר שכשמבצעים התמרת DCT על תמונות טבעיות, אז רוב האנרגיה נדחסת במעט איברים בפינה השמאלית עליונה של תמונת ההתמרה. כך ניתן לבצע קוונטיזציה עדינה על אותו חלק חיוני, ולוותר על רוב תחום התדרים (ע"י קוונטיזציה גסה בחלקים בתחום זה).
מסתבר שהעיוות שמקבלים כתוצאה מהקוונטיזציה קטן יותר מאשר אם היינו עושים קוונטיזציה על התמרת DFT של אותה תמונה.

תהליך הדחיסה בתקן JPG:



פעולת ה 2D-DFT מתבצעת על בלוקים לא חופפים בגודל של 8×8. ע"י חלוקה לבלוקים מקבלים תוצאה טובה יותר, מכיוון שיש אזורים שהם שפות ואזורים שהם חלקים, וכך ניתן לטפל בכל בלוק בהתאם לאנרגיות שבו.

7.5. התמרת Hadamard

זוהי התמרה ספרבילית, לאותות (וקטורים) באורך 2^n.

עבור וקטור באורך $2^1 = 2$, מטריצת ההתמרה הינה

$$\underline{\underline{H}}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

ניתן לבנות את מטריצות Hadamard היותר גדולות ע"פ כלל רקורסיבי:

$$\underline{\underline{H}}_n = \underline{\underline{H}}_1 \otimes \underline{\underline{H}}_{n-1}$$

למשל

$$\underline{\underline{H}}_2 = \underline{\underline{H}}_1 \otimes \underline{\underline{H}}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} & \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\\ \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix} & -\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1 & 1 & 1 & 1\\ 1 & -1 & 1 & -1\\ 1 & 1 & -1 & -1\\ 1 & -1 & -1 & 1\end{bmatrix}$$

מכיוון שזוהי פעולה ספרבילית, שתי השיטות לחישוב ההתמרה הן (לדוגמא עבור $\underline{\underline{H}}_2$):

$$\underline{U}_{rs} = \underline{\underline{H}}_2 \underline{u}^{rs}$$

$$\underline{\underline{U}} = \underline{\underline{H}}_1 \underline{\underline{u}} \underline{\underline{H}}_1^T$$

היתרון הראשון הוא שאין הכפלה במספרים ממשיים כמו ב DFT, והשני הוא שפעולות ההכפלה הן מאוד פשוטות – כפל ב 1 או ב -1.

7.6. התמרת Haar

יש עולם שלם של התמרות הנקראות התמרות Wavelets, או בשמן המלא DWT – Discrete Wavelets Transform.

התמרת Haar היא מקרה פרטי של התמרת DWT.

נתמיר זוג איברים ע"י ההתמרה:

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}$$

ונקבל

$$\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix}a+b\\ a-b\end{bmatrix}$$

כאשר האיבר הראשון שקיבלנו, $a+b$, מייצג מסנן LPF (ביצוע של ממוצע), והאיבר השני, $a-b$, מייצג HPF (הפרש).

התמרת Haar פועלה על שורות בנפרד, ואז על עמודות, כלומר זוהי התמרה ספרבילית.

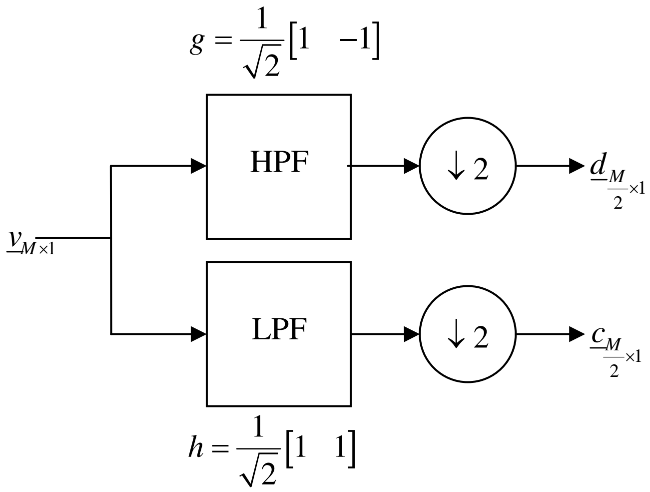
נפעיל את ההתמרה $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix}$ בבלוקים על שורה בתמונה, כלומר:

$$\begin{matrix} \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix} & \\ & \\ & \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix} \\ & \\ & & \frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix} \end{matrix} \begin{bmatrix} a \\ b \\ c \\ d \\ e \\ f \end{bmatrix}$$

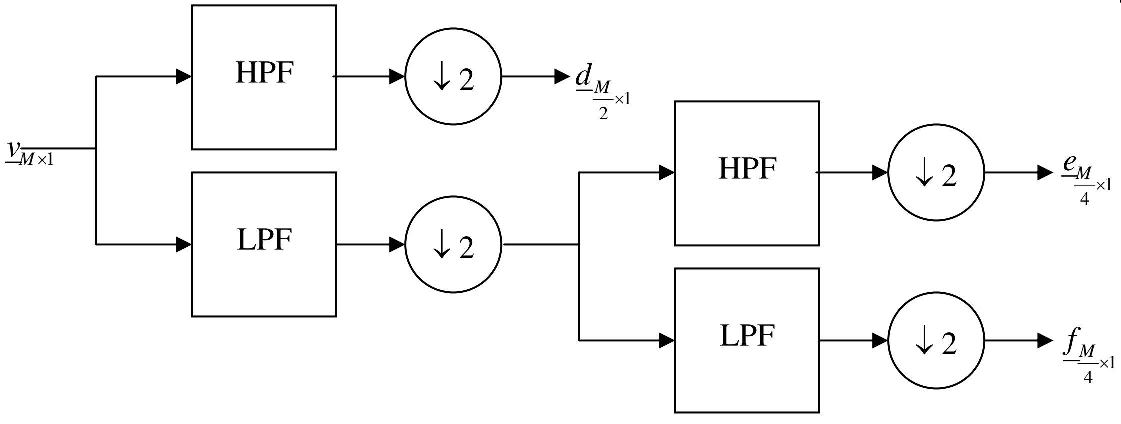
כלומר, מהוקטור $\begin{bmatrix}a & b & c & d & e & f\end{bmatrix}^T$ נייצר את האיברים $\frac{1}{\sqrt{2}}\begin{bmatrix}a+b & c+d & e+f\end{bmatrix}^T$. הפעולה הזו שקולה לביצוע קונבולוציה עם המסנן (המטשטש) $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\end{bmatrix}$, ואז "זריקת" כל איבר שני, כלומר ביצוע Downsampling (דילול) בפקטור 2.

כעת נבצע אותה פעולה, עם המסנן שני, מסנן ה HPF המתאים שראינו מקודם: $\frac{1}{\sqrt{2}}\begin{bmatrix}1 & -1\end{bmatrix}$.

לסיכום, קיבלנו:



וקטור $\underline{d}$ הוא וקטור ההפרשים המקומיים, ווקטור $\underline{c}$ הוא וקטור הממוצעים המקומיים.
נשים לב שפעולה זו הפיכה – כלומר ניתן לשחזר את a,b מידיעת הממוצעים וההפרשים שלהם.
המסננים h,g אותם אנחנו בוחנים הם אלו הספציפיים להתמרת Haar, התמרות DWT אחרות מתעסקות עם מסננים שונים.
נמשיך לבצע את המהלך המתואר באופן רקורסיבי לתחום התדרים הנמוך (כלומר לוקטור המתקבל מסינון בלוקים LPF), ונקבל



בחלוקה הראשונה, קיבלנו חלוקה לתדרים גבוהים ותדרים נמוכים, ובחלוקה השנייה מבצעים את אותה החלוקה לתדרים הנמוכים שקיבלנו מהחלוקה הראשונה.

יתרון ההתמרה הוא אפשרות לייצוג Multiscale של התמונה, כלומר ניתן על-סמך וקטור התדרים הכי נמוכים לראות תמונה קטנה מייצגת של התמונה המקורית. כשמוסיפים מידע לגבי התדרים היותר גבוהים, התמונה תהיה יותר חדה.

8. נספח מתמטי

8.1. התמרות פורייה דו-מימדיות

8.1.1. פונקציות שימושיות

מס	פונקציה
1	$\mathrm{rect}(x)=\begin{cases}1, & \lvert x\rvert\le\frac{1}{2}\\ 0, & else\end{cases}$
2	$\mathrm{sgn}(x)=\begin{cases}1, & x>0\\ 0, & x=0\\ -1, & x<0\end{cases}$
3	$\mathrm{sinc}(x)=\frac{\sin\pi x}{\pi x}$
4	$\mathrm{comb}(x)=\sum_{n=-\infty}^{\infty}\delta(x-n)$
5	$\mathrm{tri}(x)=\begin{cases}1-\lvert x\rvert, & \lvert x\rvert\le 1\\ 0, & else\end{cases}$

8.1.2. תכונות ההתמרה

מס	תכונה	נוסחה
		$f(x,y)\xrightarrow{F}F(u,v)$
1	דואליות	$F(x,y)\xrightarrow{F}g(-u,-v)$
2		
3		
4		
5		
6		
7		
8		
9		
10		
11		
12		
13		
14		

8.1.3. התמרות ידועות

מס	אות בזמן רציף	התמרת פוריה רציפה
1	$\delta(x,y)$	1
2	$\delta(x-x_0,y-y_0)$	$\exp\{-2\pi j(x_0u+y_0v)\}$
3	$\exp\{2\pi j(xu_0+yv_0)\}$	$\delta(u-u_0,u-u_0)$
4	$\exp\{-\pi(x^2+y^2)\}$	$\exp\{-\pi(u^2+v^2)\}$
5	$\mathrm{rect}(x,y)$	$\mathrm{sinc}(u,v)$
6	$\mathrm{tri}(x,y)$	$\mathrm{sinc}^2(u,v)$
7	$\mathrm{comb}(x,y)$	$\mathrm{comb}(u,v)$

		8
		9
		10
		11
		12
		13
		14

8.2. זהויות טריגונומטריות

$$\cos\left(\alpha+\beta\right)=\cos\alpha\cos\beta-\sin\alpha\sin\beta$$

8.3. פונקצית דלתא

$$\delta\left(x-x_0\right)\delta\left(y-y_0\right)=\delta\left(x-x_0,y-y_0\right)$$

$$\delta\left(\frac{x}{\alpha}-x_0\right)=\delta\left(\frac{1}{\alpha}\left(x-\alpha x_0\right)\right)=\left|\alpha\right|\delta\left(x-\alpha x_0\right)$$

8.4. חשבון מטריצות

$$\underline{\underline{A}}=\begin{bmatrix} a & b \\ c & d \end{bmatrix} \quad \Rightarrow \quad \underline{\underline{A}}^{-1}=\frac{1}{ad-bc}\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$$\det\left(k\underline{\underline{A}}\right)=k^2\det\left(\underline{\underline{A}}\right)$$

מטריצה סימטרית היא זו המקיימת:

$$\underline{\underline{A}}^T=\underline{\underline{A}}$$

מטריצה הרמיטית היא מטריצה סימטרית וממשית:

$$\left(\underline{\underline{A}}^*\right)^T=\underline{\underline{A}}$$

למטריצה זו ערכים עצמיים ממשיים.

למת $ABCD$:

$$\left(\underline{\underline{A}}-\underline{\underline{B}}\,\underline{\underline{C}}\,\underline{\underline{D}}\right)^{-1}=\underline{\underline{A}}^{-1}+\underline{\underline{A}}^{-1}\underline{\underline{B}}\left(\underline{\underline{C}}^{-1}-\underline{\underline{D}}\,\underline{\underline{A}}^{-1}\underline{\underline{B}}\right)^{-1}\underline{\underline{D}}\,\underline{\underline{A}}^{-1}$$

8.5. הסתברות

8.5.1. הגדרת הסתברות מותנית

$$\Pr\left(A\mid B\right)\triangleq\frac{\Pr\left(A\cap B\right)}{\Pr\left(B\right)}$$

8.5.2. חוק Bayes

$$\Pr\left(A\mid B\right)=\frac{\Pr\left(B\mid A\right)\Pr\left(A\right)}{\Pr\left(B\right)}$$

8.5.3. משתנה אקראי גאוסי

$$X\sim N\left(\mu,\sigma\right)\quad\Rightarrow\quad P_X\left(x\right)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left\{-\frac{\left(x-\mu\right)^2}{2\sigma^2}\right\}$$

כאשר

$$E\left[X\right]=\mu,\qquad E\left[\left(X-\mu\right)^2\right]=\sigma^2$$

8.5.4. וקטור אקראי גאוסי

$$\underline{X}_{n\times1}\sim N\left(\underline{\mu},\underline{\underline{\Lambda}}\right)\quad\Rightarrow\quad P_{\underline{X}}\left(\underline{x}\right)=\frac{1}{\sqrt{\left(2\pi\right)^n\det\underline{\underline{\Lambda}}}}\exp\left\{-\frac{1}{2}\left(\underline{x}-\underline{\mu}\right)^T\underline{\underline{\Lambda}}\left(\underline{x}-\underline{\mu}\right)\right\}$$

כאשר

$$\underline{\mu} = E\left[\underline{X}\right], \qquad \underline{\underline{\Lambda}} = E\left[\left(\underline{X}-\underline{\mu}\right)^T\left(\underline{X}-\underline{\mu}\right)\right]$$

8.5.5. טרנספורמציה של משתנים אקראיים

נתון מ״א X בעל פונקצית PDF $P_X\left(x\right)$. כמו כן, נתונה טרנספורמציה $Y=T\left(X\right)$. כאשר T מונוטונית, מתקיים

$$P_Y\left(y\right)=\left.\frac{P_X\left(x\right)}{T'\left(x\right)}\right|_{x=T^{-1}\left(y\right)}$$

אם T אינה מונוטונית, כלומר קיימים מספר x-ים שנותנים ערך y, נסמן את אותם ה x-ים ב $x_1,x_2,...,x_n$, ואז:

$$P_Y\left(y\right)=\sum_{i=1}^{n}P_X\left(x_i\right)\left.\frac{1}{T'\left(x_i\right)}\right|_{x_i=T^{-1}\left(y\right)}$$

8.6. סידור לקסיקורפי (Lexicographic) של מטריצות

למען פשטות הטיפול בתמונות, מאוד שימושי למפות את המטריצה $\underline{\underline{A}}$ הדו-מימדיות (בעלות M שורות ו N עמודות, כלומר במימד $M\times N$) המייצגת את התמונה שלנו לוקטור עמודה אחד ארוך (באורך MN, כלומר מימדו יהיה $MN\times 1$).

8.6.1. סידור עמודות (Column Stack)

בסידור זה, אנחנו לוקחים את כל עמודות $\underline{\underline{A}}$ ומניחים אותן זו מתחת לזו בוקטור $\underline{A}^{cs}$:

$$\underline{\underline{A}}=\begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{M,1} & \cdots & \cdots & a_{M,N} \end{bmatrix} \quad\Rightarrow\quad \underline{A}^{cs}=\begin{bmatrix} a_{1,1} \\ a_{2,1} \\ \vdots \\ a_{M,1} \\ a_{1,2} \\ a_{2,2} \\ \vdots \\ a_{M,2} \\ \vdots \\ a_{1,N} \\ a_{2,N} \\ \vdots \\ a_{M,N} \end{bmatrix}$$

8.6.2. סידור שורות (Row Stack)

בסידור זה, אנחנו לוקחים את כל שורות $\underline{\underline{A}}$, מבצעים על כל אחד מהן פעולת Transpose, כדי שיהיו בצורת עמודה, ומניחים אותן זו מתחת לזו בוקטור $\underline{A}^{rs}$:

$$\underline{\underline{A}} = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,N} \\ a_{2,1} & \ddots & & \vdots \\ \vdots & & \ddots & \vdots \\ a_{M,1} & \cdots & \cdots & a_{M,N} \end{bmatrix} \quad \Rightarrow \quad \underline{A}^{rs} = \begin{bmatrix} a_{1,1} \\ a_{1,2} \\ \vdots \\ a_{1,N} \\ a_{2,1} \\ a_{2,2} \\ \vdots \\ a_{2,N} \\ \vdots \\ a_{M,1} \\ a_{M,2} \\ \vdots \\ a_{M,N} \end{bmatrix}$$

8.7. מטריצות מיוחדות

8.7.1. מטריצת טופליץ (Toeplitz)

זוהי מטריצה אשר כל איבריה, אשר לאורך אלכסון כלשהו, שווים. עבור מטריצה $\underline{\underline{T}}_{M\times N}$, ניתן לנסח זאת כך:

$$\underline{\underline{T}}\left[m+1,n+1\right] = \underline{\underline{T}}\left[m,n\right], \qquad \begin{cases} 0 \le m \le M-1 \\ 0 \le n \le N-1 \end{cases}$$

במטריצת טופליץ $\underline{\underline{T}}_{M\times N}$ יכולים להיות $N+M-1$ ערכים שונים.

לדוגמא:

$$\underline{\underline{T}}_{M\times N} = \begin{bmatrix} t_0 & \cdots & & t_{-N+1} \\ t_1 & t_0 & & \vdots \\ \vdots & & \ddots & t_{-1} \\ t_{M-1} & \cdots & \cdots & t_0 \end{bmatrix}$$

ובניסוח אחר, המציג כי איברי המטריצה תלויים רק בהפרש בין אינדקס השורה לעמודה:

$$\underline{\underline{T}}\left[m,n\right] = t_{m-n}, \qquad t_i \in \mathbb{C}$$

דוגמא

עבור שני האותות:

$$x\left[n\right], \quad n = 0,...,M-1$$

$$h\left[n\right], \quad n = 0,...,N-1$$

נביט בקונבולוציה לינארית (בשונה מקונבולוציה ציקלית) חד-מימדית שלהם:

$$y\left[n\right] = h\left[n\right] * x\left[n\right] = \sum_{k=0}^{N-1} h\left[k\right] x\left[n-k\right] = \sum_{k=0}^{M-1} x\left[k\right] h\left[n-k\right]$$

נרצה לבטא את פקולת הקונבולוציה בעזרת כפל מטריצה:

$$\underline{y} = \underline{\underline{T}}\underline{x}$$

האותות באורך N, לכן אין לנו בעיה להשתמש באידקסים מחוץ לתחום $\left[0,N-1\right]$ - פשוט נאפס את האיברים באינדקסים אלו. נבחין בכך ש

$$
\begin{aligned}
&y[0]=x[0]h[0]\\
&y[1]=x[0]h[1]+x[1]h[0]\\
&y[2]=x[0]h[2]+x[1]h[1]+x[2]h[0]\\
&\vdots
\end{aligned}
$$

נגדיר שני וקטורי עמודה :

$$
\begin{aligned}
&y=\begin{bmatrix} y[0] & y[1] & \cdots & y[N+M-1]\end{bmatrix}^T\\
&x=\begin{bmatrix} x[0] & x[1] & \cdots & x[M-1]\end{bmatrix}^T
\end{aligned}
$$

ואז ניתן לכתוב את פעולת הקונבלוציה בעזרת המשוואה המטריצית :

$$\underline{y}_{(M+N-1)\times 1}=\underline{\underline{T}}_{(M+N-1)\times M}\,\underline{x}_{M\times 1}$$

כאשר המטריצה המתאימה לייצג את פעולת הקונבולוציה היא מטריצת טופליץ :

$$
\underline{\underline{T}}_{(M+N-1)\times M}=\begin{bmatrix}
h[0] & 0 & 0 & \cdots & \cdots & 0\\
h[1] & h[0] & 0 & & & 0\\
h[2] & & h[0] & 0 & & 0\\
\vdots & & & \ddots & \ddots & \vdots\\
h[N-1] & & & & \ddots & 0\\
0 & \ddots & & & & h[0]\\
0 & \ddots & \ddots & & & h[1]\\
\vdots & & \ddots & \ddots & & h[2]\\
\vdots & & & 0 & \ddots & \vdots\\
0 & \cdots & 0 & 0 & 0 & h[N-1]
\end{bmatrix}
$$

כלומר

$$
\begin{bmatrix}
y[0]\\ y[1]\\ y[2]\\ y[3]\\ y[4]\\ \\ \\ \vdots\\ \\ x[N+M-1]
\end{bmatrix}
=\begin{bmatrix}
h[0] & 0 & 0 & \cdots & \cdots & 0\\
h[1] & h[0] & 0 & & & 0\\
h[2] & & h[0] & 0 & & 0\\
\vdots & & & \ddots & \ddots & \vdots\\
h[N-1] & & & & \ddots & 0\\
0 & \ddots & & & & h[0]\\
0 & \ddots & \ddots & & & h[1]\\
\vdots & & \ddots & \ddots & & h[2]\\
\vdots & & & 0 & \ddots & \vdots\\
0 & \cdots & 0 & 0 & 0 & h[N-1]
\end{bmatrix}
\begin{bmatrix}
x[0]\\ x[1]\\ x[2]\\ x[3]\\ \vdots\\ x[M-1]
\end{bmatrix}
$$

8.7.2. מטריצה סיבובית

מטריצה זו היא מקרה פרטי של מטריצת טופליץ. זוהי מטריצה ריבועית $\underline{\underline{C}}_{N\times N}$ אשר כל שורה בה היא הזזה ב 1 של השורה שמעליה (נאמר שהשורה הראשונה היא העוגן). ניתן לנסח את חוקיות ערכי האיברים כך :

$$\underline{\underline{C}}[m,n]=c_{(m-n)\bmod N}$$

מטריצה זו משמשת בייצוג של פעולת קונבולוציה ציקלית, באופן מקביל לייצוג פעולת הקונבולוציה הלינארית ע״י מטריצת טופליץ.

עבור שני האותות :

$$
\begin{aligned}
&x[n], \quad n=0,...,N-1\\
&h[n], \quad n=0,...,N-1
\end{aligned}
$$

נביט בקונבולוציה ציקלית חד-מימדית שלהם :

$$y[n]=h[n]*x[n]=\sum_{k=0}^{M-1}x[k]h\big[(n-k)\bmod N\big]$$

קונבולוציה זאת ניתנת לייצוג ע"י מטריצה, באופן הבא :

$$\begin{bmatrix} y[0] \\ y[1] \\ \vdots \\ y[N-1] \end{bmatrix}_{N\times 1} = \begin{bmatrix} h[0] & h[1] & & h[N-1] \\ h[N-1] & h[0] & h[1] & \\ & & \ddots & \\ h[1] & & h[N-1] & h[0] \end{bmatrix}_{N\times N} \cdot \begin{bmatrix} x[0] \\ x[1] \\ \vdots \\ x[N-1] \end{bmatrix}_{N\times 1}$$

8.7.3. מטריצת בלוקים

זוהי מטריצה אשר מייצגים כל קבוצת איברים שלה בתור מטריצה בפני עצמה, קטנה יותר, כאשר כל המטריצות הקטנות (בלוקים) הן בגודל זהה $P\times Q$ (מימד הבסיס). אם מטריצת הבלוקים $\underline{\underline{B}}$ בנוייה מסה"כ M שורות של בלוקים ו N עמודות של בלוקים, סה"כ מימד מטריצת הבלוקים יהיה $MP\times NQ$:

$$\underline{\underline{B}}=\begin{bmatrix} \underline{\underline{A}}_{1,1} & \underline{\underline{A}}_{1,2} & \cdots & \underline{\underline{A}}_{1,N} \\ \underline{\underline{A}}_{2,1} & & & \vdots \\ \vdots & & \ddots & \vdots \\ \underline{\underline{A}}_{M,1} & \cdots & \cdots & \underline{\underline{A}}_{M,N} \end{bmatrix}_{MP\times NQ}$$

8.7.4. מטריצת בלוק-טופליץ

אם יש שיוון בין הבלוקים של מטריצת טופליץ $\underline{\underline{B}}$, המוגדר ע"י הקשר של מטריצת טופליץ, כלומר אם עבור כל הבלוקים $\underline{\underline{A}}_{i,j}$ של $\underline{\underline{B}}$ מתקיים

$$\underline{\underline{A}}_{i,j}=\underline{\underline{A}}_{i-j}$$

אזי המטריצה $\underline{\underline{B}}$ תקרא מטריצת בלוק-טופליץ

8.7.5. מטריצת בלוק-טופליץ כפולה

מטריצת בלוק-טופליץ $\underline{\underline{B}}$ תיקרא מטריצת בלוק-טופליץ כפולה אם כל בלוק שלה, $\underline{\underline{A}}_{i,j}$, הוא מטריצת טופליץ יש לשים לב שמטריצה כזו אינה בהכרח מטריצת טופליץ, כלומר $\underline{\underline{B}}$ לא חייבת להיות גם היא מטריצת טופליץ.

ניקח תמונה $\underline{\underline{X}}_{M\times N}$ ונבצע איתה קונבולוציה (דו-מימדית) עם גרעין $\underline{\underline{H}}_{P\times Q}$:

$$\underline{\underline{Y}}^{cs}_{(M+P-1)\times(N+Q-1)}=\underline{\underline{H}}_{P\times Q}*X^{cs}_{M\times N}$$

נמפה את התמונה לוקטור עמודות $\underline{X}^{cs}$. את פעולת הקונבולוציה הדו-מימדית של התמונה $\underline{\underline{X}}^{cs}$ עם גרעין $\underline{\underline{H}}$, לקבלת התמונה $\underline{\underline{Y}}^{cs}$, ניתן לייצג ע"י פעולה לינארית של מטריצה מסוג בלוק-טופליץ כפולה :

$$\underline{Y}^{cs}=\begin{bmatrix} \underline{\underline{T}}_0 & 0 & \cdots & \cdots & 0 & 0 \\ \underline{\underline{T}}_1 & \underline{\underline{T}}_0 & 0 & & & 0 \\ \vdots & \underline{\underline{T}}_1 & \ddots & & & \vdots \\ \underline{\underline{T}}_{N+Q-1} & & \ddots & \ddots & & \vdots \\ 0 & \ddots & & \ddots & \ddots & \vdots \\ 0 & 0 & \underline{\underline{T}}_{N+Q-1} & \cdots & \underline{\underline{T}}_1 & \underline{\underline{T}}_0 \end{bmatrix}\overline{X}^{cs}$$

גדלי הבלוקים $\underline{\underline{T}}_i$ הוא $(M+P-1)\times M$, וגודל מטריצת הבלוק-טופליץ כולה, $\underline{\underline{H}}$, הוא

$$\big((M+P-1)(N+Q-1)\big)\times(M\cdot N)$$

8.7.6. מטריצה בלוק-סיבובית

זוהי מטריצת בלוקים $\underline{\underline{B}}$, אשר הבלוקים שלה, $\underline{\underline{A}}_{i,j}$ מקיימים את תכונת הסיבוביות:

$$\underline{\underline{A}}_{i,j}=\underline{\underline{A}}_{(i-j)\bmod N}$$

8.7.7. מטריצה בלוק-סיבובית כפולה

זוהי מטריצת בלוק-סיבובית $\underline{\underline{B}}$, אשר הבלוקים שלה, $\underline{\underline{A}}_{i,j}$ הם מטריצות סיבוביות.
מטריצות מסוג זה משמשות בייצוג של פעולות קונבולוציה ציקלית דו-מימדיות.

8.7.8. מטריצה פסאודו-הופכית

למטריצה ריבועית $\underline{\underline{A}}_{N\times N}$ יכולה להיות מטריצה הופכית, אשר מסומנת $\underline{\underline{A}}_{N\times N}^{-1}$.
למטריצות $\underline{\underline{B}}_{M\times N}$ שאינן ריבועיות, מגדירים מטריצה פסאודו-הופכית (Pseudo-Inverse):

$$\underline{\underline{B}}_{M\times N}^{-1}\triangleq\left(\underline{\underline{B}}^T\underline{\underline{B}}\right)^{-1}\underline{\underline{B}}^T$$

אשר כמובן לא תמיד קיימת.

8.8. מכפלת קרוניקר

מכפלת קרוניקר $\underline{\underline{C}}$ בין מטריצה $\underline{\underline{A}}_{M_1\times M_2}$ למטריצה $\underline{\underline{B}}_{N_1\times N_2}$ היא יצירת של מטריצת בלוקים $\underline{\underline{C}}$, כאשר הבלוקים הוא מטריצות $\underline{\underline{B}}_{N_1\times N_2}$ מוכפלות באיברי מטריצה $\underline{\underline{A}}_{M_1\times M_2}$, ומסומנת $\underline{\underline{C}}=\underline{\underline{A}}\otimes\underline{\underline{B}}$:

$$\underline{\underline{C}}=\underline{\underline{A}}\otimes\underline{\underline{B}}=\left\{a[m,n]\cdot\underline{\underline{B}}\right\}=\begin{bmatrix} a[1,1]\cdot B & a[1,1]\cdot B & a[1,M_2]\cdot B \\ & & \\ a[M_1,1]\cdot B & & a[M_1,M_2]\cdot B \end{bmatrix}$$

תכונות:

מכפלת קרוניקר אינה קומוטטיבית, כלומר $\underline{\underline{A}}\otimes\underline{\underline{B}}\neq\underline{\underline{B}}\otimes\underline{\underline{A}}$
$\left(\underline{\underline{A}}+\underline{\underline{B}}\right)\otimes\underline{\underline{C}}=\underline{\underline{A}}\otimes\underline{\underline{C}}+\underline{\underline{B}}\otimes\underline{\underline{C}}$
$\left(\underline{\underline{A}}\otimes\underline{\underline{B}}\right)\otimes\underline{\underline{C}}=\underline{\underline{A}}\otimes\left(\underline{\underline{B}}\otimes\underline{\underline{C}}\right)$
$\forall\alpha\in\mathbb{R}:\qquad \alpha\left(\underline{\underline{A}}\otimes\underline{\underline{B}}\right)=\left(\alpha\underline{\underline{A}}\right)\otimes\underline{\underline{B}}=\underline{\underline{A}}\otimes\left(\alpha\underline{\underline{B}}\right)$
$\left(\underline{\underline{A}}\otimes\underline{\underline{B}}\right)^T=\underline{\underline{A}}^T\otimes\underline{\underline{B}}^T$
$\left(\underline{\underline{A}}\otimes\underline{\underline{B}}\right)\left(\underline{\underline{C}}\otimes\underline{\underline{D}}\right)=\underline{\underline{A}}\underline{\underline{C}}\otimes\underline{\underline{B}}\underline{\underline{D}}$
$\left(\underline{\underline{A}}\otimes\underline{\underline{B}}\right)^{-1}=\underline{\underline{A}}^{-1}\otimes\underline{\underline{B}}^{-1}$
$\left(\underline{\underline{A}}\otimes\underline{\underline{B}}\right)=\left(\underline{\underline{A}}\otimes\underline{\underline{I}}\right)\left(\underline{\underline{I}}\otimes\underline{\underline{B}}\right)$

8.9. חדו"א של וקטורים ומטריצות

8.9.1. הגדרות

הוקטורים הם וקטורי עמודה.

1. נגזרת של וקטור $\underline{y}$ לפי סקלר x, כלומר $\frac{\partial\underline{y}}{\partial x}$, הוא וקטור אשר האיבר ה n שלו הוא נגזרת האיבר ה n בוקטור $\underline{y}$, לפי הסקלר x, כלומר $\frac{\partial\underline{y}[n]}{\partial x}$.

2. נגזרת של סקלר y לפי וקטור $\underline{x}$, כלומר $\frac{\partial y}{\partial \underline{x}}$, הוא וקטור אשר האיבר ה n שלו הוא נגזרת של הסקלר y לפי האיבר ה n של הוקטור $\underline{x}$, כלומר $\frac{\partial y}{\partial \underline{x}[n]}$.
3. נגזרת של מטריצה $\underline{\underline{y}}$ לפי סקלר x, כלומר $\frac{\partial \underline{\underline{y}}}{\partial x}$, היא מטריצה אשר האיבר ה $[m,n]$ שלה הוא נגזרת האיבר ה $[m,n]$ במטריצה $\underline{\underline{y}}$, לפי הסקלר x, כלומר $\frac{\partial \underline{\underline{y}}[m,n]}{\partial x}$.
4. נגזרת של סקלר y לפי מטריצה $\underline{\underline{x}}$, כלומר $\frac{\partial y}{\partial \underline{\underline{x}}}$, היא מטריצה אשר האיבר ה $[m,n]$ שלה הוא נגזרת הסקלר y לפי האיבר ה $[m,n]$ במטריצה $\underline{\underline{x}}$, כלומר $\frac{\partial y}{\partial \underline{\underline{x}}[m,n]}$.
5. הנגזרת $\frac{\partial \underline{y}^T}{\partial \underline{x}}$, היא מטריצה אשר האיבר ה $[m,n]$ שלה הוא נגזרת האיבר ה n בוקטור $\underline{y}$, לפי האיבר ה m בוקטור $\underline{x}$, כלומר $\frac{\partial \underline{y}[m]}{\partial \underline{x}[n]}$.

8.9.2. תוצאות

$\frac{\partial}{\partial \underline{x}}\left(\underline{a}^T \underline{x}\right) = \frac{\partial}{\partial \underline{x}}\left(\underline{x}^T \underline{a}\right) = \underline{a}$
$\frac{\partial}{\partial \underline{x}}\left(\underline{x}^T \underline{\underline{A}}\underline{x}\right) = \left(\underline{\underline{A}}^T + \underline{\underline{A}}\right)\underline{x}$
$\frac{\partial z}{\partial \underline{x}} = \frac{\partial \underline{y}}{\partial \underline{x}}\frac{\partial z}{\partial \underline{y}}$
בדרך כלל $\frac{\partial z}{\partial \underline{x}} \neq \frac{\partial z}{\partial \underline{y}}\frac{\partial \underline{y}}{\partial \underline{x}}$
$\frac{\partial}{\partial \underline{x}}\underline{\underline{A}}^T \underline{x} = \frac{\partial}{\partial \underline{x}}\underline{x}^T \underline{\underline{A}} = \underline{\underline{A}}$
$\frac{\partial}{\partial \underline{x}}\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)^T\left(\underline{\underline{A}}\underline{x}+\underline{b}\right) = 2\underline{\underline{A}}^T\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)$
$\frac{\partial}{\partial \underline{x}}\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)^T\underline{\underline{C}}\left(\underline{\underline{A}}\underline{x}+\underline{b}\right) = \underline{\underline{A}}^T\left(\underline{\underline{C}}+\underline{\underline{C}}^T\right)\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)$
$\frac{\partial}{\partial \underline{x}}\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)^T\left(\underline{\underline{D}}\underline{x}+\underline{e}\right) = \underline{\underline{A}}^T\left(\underline{\underline{D}}\underline{x}+\underline{e}\right)+\underline{\underline{D}}^T\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)$
$\frac{\partial}{\partial \underline{x}}\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)^T\underline{\underline{C}}\left(\underline{\underline{D}}\underline{x}+\underline{e}\right) = \underline{\underline{A}}^T\underline{\underline{C}}\left(\underline{\underline{D}}\underline{x}+\underline{e}\right)+\underline{\underline{D}}^T\underline{\underline{C}}^T\left(\underline{\underline{A}}\underline{x}+\underline{b}\right)$