מקליד: עידן אריה
מרצה: פרופ' נתן נתניהו
תאריך: 2018-10-31

עיבוד תמונות
89-560-01

טרנפורמציות לינאריות (Linear Transformation/Mapping)

הגדרה: מיפוי של ערך אפור a לערך אפור b $b = T(a)$ בהתאם לפונקציה לינארית

אם בשורה התחתונה התוצאה של טרנפורמציה (כלשהי - לאו דווקא לינארית) היא קטנה מערך מינימלי (0) או גדולה מערך מקסימלי (בד"כ 255) - "קוטעים" אותה.

Linear Scaling

$$b = \frac{b_{\max} - b_{\min}}{a_{\max} - a_{\min}} \cdot (a - a_{\min}) + b_{\min}$$

$a_{\min}$ ממופה ל$b_{\min}$, $a_{\max}$ ממופה ל$b_{\max}$, והערכים ביניהם ממופים בצורה לינארית.
כאשר $a_{\max} - a_{\min} > 1$ הקונטרסט יגדל. כאשר $0 < a_{\max} - a_{\min} < 1$ הקונטרסט יקטן.

תשליל

$$b = a_{\max} - a$$

כהה הופך לבהיר ובהיר הופך לכהה.

אופרטורים בוליאנים

OR $b = a + k$

AND $b = a \cdot k$

XOR $b = a \oplus k$

NEG $b = \overline{a}$

טרנספורמציות לא-לינאריות

הנפוצות הן:

לוגריתמיות: הקונטרסט יגדל באזורים הנמוכים/חשוכים, ויקטן באזורים הגבוהים/בהירים:

$$b = k \cdot \ln(1 + a)$$

(עבור $k = 46$ $a = 255$ ימופה ל$b = 255$)

אקפוננציאליות: הקונטרסט יגדל באזורים הגבוהים/בהירים, ויקטן באזורים הנמוכים/חשוכים:

$$b = e^{\frac{a}{k}} - 1$$

אופרציות בין תמונות

עד עכשיו דיברנו על אופרציות על תמונה בודדת, אבל אפשר להגדיר אופרציות על שתי תמונות f ו־g (תמונת היעד נסמן o):

$$o\left(i,j\right)=f\left(i,j\right)+g\left(i,j\right)$$

$$o\left(i,j\right)=f\left(i,j\right)-g\left(i,j\right)$$

$$o\left(i,j\right)=f\left(i,j\right)\cdot g\left(i,j\right)$$

$$o\left(i,j\right)=\frac{f\left(i,j\right)}{g\left(i,j\right)}$$

ממוצע תמונות - ניקוי רעש

רכישת תמונה מלווה ברעש n:

$$g\left(x,y\right)=f\left(x,y\right)+r\left(x,y\right)$$

הרעש הוא ערך אקראי - אין פונקציה אנליטית לחשב אותו, ולכן נשתמש באמצעים הסתברותיים כדי לנטרל אותו. כאשר התוחלת של הרעש היא 0 ואין קורולציה בינו לסיגנל המקורי f, אם ניתן לרכוש מספר תמונות עבור מערכת אופטית נתונה אפשר לעשות ממוצע כדי למזער את הרעש:

$$o\left(x,y\right)=\frac{1}{k}\sum_{i=1}^{K}g_{i}\left(x,y\right)=f\left(x,y\right)+\frac{1}{K}\sum_{i=1}^{K}r_{i}\left(x,y\right)$$

כאשר: $f\left(x,y\right)$ התמונה האמיתית
$g_i\left(x,y\right)$ הרפליקה ה־i
$r_i\left(x,y\right)$ רמת הרעש עבור הרפליקה ה־i

הרעש הוא משתנה אקראי בכל פיקסל, בלתי תלוי בפיקסלים הסמוכים (למרות שהם נובעים מאותה פונקציית צפיפות הסתברות). התוחלת של הרעש היא 0, ולכן המיצוע מקטין את השונות של הרעש פי K (או מקטין את סטיית התקן פי $\sqrt{K}$), שכן עבור משתנים אקראיים X,Y כך ש־$Y=aX$ עבור a קבוע:

$$\mathrm{VAR}\left[Y\right]=\mathrm{E}\left[\left(Y-\overline{Y}\right)^{2}\right]=\mathrm{E}\left[\left(aX-a\overline{X}\right)^{2}\right]=\mathrm{E}\left[a^{2}\left(X-\overline{X}\right)^{2}\right]=a^{2}\mathrm{E}\left[\left(X-\overline{X}\right)^{2}\right]=a^{2}\mathrm{VAR}\left[X\right]$$

ולכן:

$$\mathrm{VAR}\left[\frac{1}{K}\sum_{i=1}^{K}r_{i}\left(x,y\right)\right]=\frac{1}{K^{2}}\sum\mathrm{VAR}\left[\sum_{i=1}^{K}r_{i}\left(x,y\right)\right]$$

מה השונות של סכום משתנים אקראיים בלתי תלויים?

למה: $\mathrm{E}\left[X_1\cdot X_2\right]=\mathrm{E}\left[X_1\right]\cdot\mathrm{E}\left[X_2\right]$

מסקנה: הקורלציה בין X_1 ו־X_2 היא 0

הוכחה:

$$\mathrm{E}\left[\left(X_1-\overline{X_1}\right)\left(X_2-\overline{X_2}\right)\right]=\mathrm{E}\left[X_1X_2\right]-\overline{X_2}\mathrm{E}\left[X_1\right]-\overline{X_1}\mathrm{E}\left[X_2\right]+\overline{X_1}\cdot\overline{X_2}$$

ומכיוון ש־$\overline{X_1}=\mathrm{E}\left[X_1\right]$ ו־$\overline{X_2}=\mathrm{E}\left[X_2\right]$ מקבלים:

$$\mathrm{E}\left[X_1X_2\right]-\overline{X_2}\mathrm{E}\left[X_1\right]-\cancel{\overline{X_1}\mathrm{E}\left[X_2\right]}+\cancel{\overline{X_1}\cdot\overline{X_2}}=\mathrm{E}\left[X_1X_2\right]-\overline{X_1}\cdot\overline{X_2}$$

לכן:

$$\mathrm{VAR}\left[X_1+X_2\right]=\mathrm{E}\left[\left(\left(X_1+X_2\right)-\left(\overline{X_1}+\overline{X_2}\right)\right)^2\right]=E\left[\left(\left(X_1-\overline{X_1}\right)+\left(X_2-\overline{X_2}\right)\right)^2\right]=$$

$$=E\left[\left(X_1-\overline{X_1}\right)^2\right]+E\left[\left(X_2-\overline{X_2}\right)^2\right]=\mathrm{VAR}\left[X_1\right]+\mathrm{VAR}\left[X_2\right]$$

ועבור $K>2$ משתנים r_i עם אותה שונות:

$$\mathrm{VAR}\left[\sum_{i=1}^{K} r_i\left(x,y\right)\right]=\sum_{i=1}^{K}\mathrm{VAR}\left[r_i\left(x,y\right)\right]=K\cdot\mathrm{VAR}\left[r\left(x,y\right)\right]$$

כלומר:

$$\mathrm{VAR}\left[\frac{1}{K}\sum_{i=1}^{K} r_i\left(x,y\right)\right]=\frac{1}{K^2}K\cdot\mathrm{VAR}\left[r\left(x,y\right)\right]=\frac{1}{K}\mathrm{VAR}\left[r\left(x,y\right)\right]$$

כלומר השונות קטנה פי K ולכן סטיית התקן קטנה פי $\sqrt{K}$.

חיסור תמונות

$$o\left(x,y\right)=\left|f\left(x,y\right)-g\left(x,y\right)\right|$$

בעלת חשיבות לגילוי שינויים - ממלא תפקיד חשוב באפליקציות רבות.
פעולת החיסור עצמה פשוטה, אבל בשביל שתהיה אפקטיבית צריך להגיע למצב של חפיפה מושלמת בין התמונות, ולהתחשב בהבדלים ברמות אפור.

הגדרת אות ורעש בתמונה

גודל אות המידע בתמונה (נקייה מרעש) שווה לשונות התמונה (הקונטרסט). עבור תמונה f בגודל $N\times M$ נגדיר:

$$\sigma_f^2=\frac{1}{NM}\sum_i\sum_j\left(f_{i,j}-b_f\right)^2$$

$$\sigma_n^2=\frac{1}{NM}\sum_i\sum_j\left(n_{i,j}-E_n\right)^2$$

כאשר:

σ_f	גודל האות	
σ_n	גודל הרעש	
b_f	בהירות התמונה	
E_n	תוחלת הרעש	

יס אות לרעש (signal to noise ratio) הוא המנה :

$$\mathrm{SNR}=\frac{\sigma_f^2}{\sigma_n^2}=\left(\frac{\sigma_f}{\sigma_n}\right)^2$$

בשביל שהתמונה תהיה ברורה מה שחשוב זה לא עצמת הרעש עצמו אלא הSNR. אם $\mathrm{SNR}<1$ אז $\log\mathrm{SNR}<0$ ולכן אומרים שהSNR שלילי ושה"סיגנל טבול ברעש) - ולמרות זאת יודעים לעשות שחזור של הסיגנל המקורי.

סוגי רעש

למה שבכלל נרצה להוסיף רעש? בד"כ כדי לבדוק אלגוריתם של עיבוד תמונה או ראייה ממוחשבת נרצה להוסיף רעש לתמונות כדי לראות שהאלגוריתם עדיין מצליח לבצע את המטלה בצורה נכונה.

רעש גאוסי (נורמלי)

$$h_i\left(a\right)=\frac{e^{\frac{-(a-\mu)^2}{2\sigma^2}}}{\sigma\sqrt{2\pi}}$$

כאשר: σ סטיית התקן

μ תוחלת הרעש

עובדה: 99.7% מערכי רמות האפור שוכנים באינטרבל $[\mu-3\sigma,\mu+3\sigma]$.

רעש מלח־פלפל

רעש שמתבטא בערכים קיצוניים בלי קשר לכלום. אם יש הסתברות ל$2p$ שזה יקרה בכל פיקסל אז:

$$n\left(a\right)=\begin{cases}s & a=a_s\\ p & a=a_p\\ 0 & \text{otherwise}\end{cases}$$

הערכים הקיצוניים קרויים outliers.

הערות כלליות על רעש

- הרעש אינוי תלוי ברמות האפור.
- הרעש בפיקסל מסויים אינו תלוי ברעש בפיקסלים השכנים.

מקליד: עידן אריה
מרצה: פרופ' נתן נתניהו
תאריך: 2018-10-31

עיבוד תמונות
89-560-01

אופרטור הקונבולוציה

$$g(x,y) = f * h(x,y) = \sum_{(i,j)\in H} f(x-i, y-j)\, h(i,j)$$

(ה* זה אופרטור הקונבולוציה)
כאשר $0 \le x-i \le n-1$ ו $0 \le y-j \le m-1$ וH הינו תחום ערכי האינדקסים עבורם מוגדר המסנן $h(i,j)$
בעצם כל פיקסל $g(x,y)$ הוא קומבינציה לינארית של **<u>סביבת</u>** הפיקסל במקור $f(x,y)$.

נשים ♡: לא נוגעים בתמונה המקורית - התוצאות מחושבות לפי תמונת הקלט ונרשמות בתמונת הפלט.

מה קורה בקצוות? יש קונבנציות שונות:

- להשאיר את הערכים כמו שהם
- לאפס את הערכים
- להשתמש בתמונה ציקלית

מקובלות מסכות בגודל 3×3 (כלומר $i,j \in \{-1,0,1\}$) אבל אפשר להשתמש גם ב5×5 או 7×7. משתמשים בד"כ בגדלים אי־זוגיים כדי

מסנן מיצוע (לינארי)

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

$$g(x,y) = \sum_{i=-1}^{1}\sum_{j=-1}^{1} \frac{f(x-i, y-j)}{9}$$

אפשר גם להשתמש בפילטרים יותר גדולים.

משמעות הממוצע

$$\sum_{i=1}^{N} (x_i - \hat{x})^2 \to \min$$

$$\frac{\partial}{\partial \hat{x}}(\mathrm{LHS}) = 2 \cdot \sum_{i=1}^{N} (x_i - \hat{x}) \equiv 0$$

ומכאן

$$\hat{x} = \frac{1}{N}\sum_{i=1}^{N} x_i$$

כלומר הממוצע הוא אומדן אופטימלי ביחד לקריטריון מזעור סכום ריבועי השגיאה ביחס לאוסף דגימות נתון. לכן זה טוב לטיפול ברעשים גאוסיאנים.

מסנן חציון (לא לינארי)

זה מסנן לא לינארי - אין כאן פעולת קונבולוציה. המסכה מגדילה רק אלו פיקסלים משתתפים בחישוב, ולוקחים את החציון של הערכים האלה. למשל במסכה של 3×3 מקבלים 9 ערכים, ואם $f_1 \leq f_2 \leq \cdots \leq f_8 \leq f_9$ אז לוקחים את f_5.

החציון מטפל בעיקר ברעשי מלח־פלפל, שכן חציון מסלק ערכים קיצוניים (כל עוד יש פחות מ50% outliers)

מה צריך להיות גודל הפילטר?

1×1? לא באמת עושה כלום (אלא אם כן המקדם שונה מ1)

זוגי? יש בעיה כי אין פיקסל באמצע.

ככל שהפילטר יותר גדול, השונות של הרעש קטנה, אבל מצד שני החדות של התמונה יורדת.