מקליד: עידן אריה
מתרגל: ד"ר אורן קאפח
תאריך: 16-01-2019

עיבוד תמונות
89-560-03

טרנספורמציות

יש לנו תמונה, ואנחנו רוצים להזיז אותה. או לסובב אותה. או להגדיל/להקטין אותה. אלו פעולות שלא משנות את הערכים של הפיקסלים, אלא רק את המיקום שלהם:

$$(x,y) \to (x',y') = (f(x,y), g(x,y))$$

- הזזה - Translation

$$\mathrm{Tr} : (x,y) \overset{(a,b)}{\to} (x+a, y+b)$$

- הגדלה - Scaling

$$\mathrm{S} : (x,y) \overset{(a,b)}{\to} (ax, by)$$

- סיבוב - Rotation

$$\mathrm{R} : (x,y) \overset{\theta}{\to} (x\cos\theta - y\sin\theta, x\sin\theta + y\cos\theta)$$

תזכורת: איך זוכרים את הנוסחה של סיבוב? אם יש לנו שני וקטורים בגדלים v_1, v_2 וזויות θ_1, θ_2, ומייצגים אותם בתור מספרים מרוכבים, הכפלה שלהם היא:

$$v_1 \measuredangle\theta_1 \cdot v_2 \measuredangle\theta_2 = v_1 v_2 \measuredangle\theta_1 + \theta_2$$

נזכור גם ש:

$$\measuredangle\theta = \cos\theta + i\sin\theta$$

ואז, כדי לסובב את (x,y) בθ, עושים:

$$(x+iy)(\cos\theta + i\sin\theta) = (x\cos\theta - y\sin\theta) + i(x\sin\theta + y\cos\theta)$$

לכאורה אפשר היה פשוט לעבור פיקסל פיקסל ולמצוא לכל אחד את המיקום החדש שלו. אבל אז, אם למשל נגדיל וקטור פי 3:

$$\begin{bmatrix}1 & 2 & 3\end{bmatrix} \to \begin{bmatrix}? & ? & 1 & ? & ? & 2 & ? & ? & 3\end{bmatrix}$$

יהיו לנו פיקסלים ריקים. לכן, במקום זה, נלך על הפיקסלים של התוצאה - ולכל מיקום נמצא את הפיקסל בתמונה המקורית שהולך אליו.
מה עושים כשלא מגיעים לאינדקס שלם?

- אפשר לקחת את השכן הכי קרוב.
- אפשר לעשות ממוצע משוקלל לפי האינדקס שיצא.

מקליד: עידן אריה
מתרגל: ד"ר אורן קאפח
תאריך: 16-01-2019

האלגוריתם

עבור הזזה

```
[rows, cols] = size(i);
ni = zeros(rows, cols);
for r = 1:rows
    for c = 1:cols
        o_r = r - b
        o_c = c - a
        if 1 ≤ o_r ≤ rows and 1 ≤ o_c ≤ cols
            ni(r, c) = i(o_r, o_c)
        end
    end
end
```

שימו לב שאם הערכים יוצאים מהתמונה - אנחנו לא לוקחים כלום, ומשאירים אותם בתור 0. זה מה שקורה כשמזיזים תמונה.

עבור הגדלה

```
[rows, cols] = size(i);
ni = zeros(rows, cols);
for r = 1:rows
    for c = 1:cols
        o_r = round(r / b)
        o_c = round(c - a)
        if 1 ≤ o_r ≤ rows and 1 ≤ o_c ≤ cols
            ni(r, c) = i(o_r, o_c)
        end
    end
end
```

עבור סיבוב

```
[rows, cols] = size(i);
ni = zeros(rows, cols);
for r = 1:rows
    for c = 1:cols
        o_r = round(−c sin(−θ) + r cos(−θ))
        o_c = round(c cos(−θ) + r sin(−θ))
        if 1 ≤ o_r ≤ rows and 1 ≤ o_c ≤ cols
            ni(r, c) = i(o_r, o_c)
        end
    end
end
```

מקליד: עידן אריה
מתרגל: ד"ר אורן קאפח
תאריך: 16-01-2019

עיבוד תמונות
89-560-03

נשים לב - לגבי הגדלה וסיבוב

ההגדלה והסיבוב נעשים יחסית לראשית הצירים. אם רוצים להגדיל או לסובב סביב נקודה אחרת - צריך לשלב גם הזזה.

טרנספורמציות באמצעות קונבולוציה

כדי לעשות הזזה באמצעות קונבולוציה, עושים $i*\delta(a,b)$ כאשר $\delta(a,b)$ הוא פילטר שבו יש רק 1 אחד, לא במרכז, וכל השאר 0. המיקום של הδ נקבע לפי a,b - למשל:

$$\delta(2,0) = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

מציאת הזזה

אם יש לנו שתי תמונות, שאחת מהן היא הזזה של השנייה, איך מחשבים בכמה הן זזו?
לפי משפט הקונובולוציה, אם $f*\delta(a,b)=h$ אז $F\cdot G=H$ (כאשר F,G,H התמרות פוריה של $f,\delta(a,b),h$ בהתאמה). ולכן:

$$G=\frac{H}{F}$$

ואז עושים ifft על G בשביל לקבל את δ, ומוצאים איפה הpeak שלו.

הערה: אם בF יש אפסים, אפשר להחליף אותם בערך מאוד קטן.

טרנספורמציות באמצעות התמרת פוריה

יש לנו תמונה בגודל $n\times n$. עושים לה fft ומקבלים מטריצת מקדמים F בגודל $n\times n$. יוצרים ממנה מטריצה בגודל $2n\times 2n$ הנראית ככה:

$$\begin{bmatrix} F_{00} & 0 & F_{01} & 0 & F_{02} & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ F_{10} & 0 & F_{11} & 0 & F_{12} & \cdots \\ 0 & 0 & 0 & 0 & 0 & \cdots \\ F_{20} & 0 & F_{21} & 0 & F_{22} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

מה יקרה כשנעשה ifft לתמונה הזאת?
התדר $(0,0)$ נשאר אותו דבר - כלומר הבהירות נשארת. כל תדר אחר גדל פי 2 (בתדירות שלו) - כלומר פי 2 יותר מהר. לכן התמונה שנקבל מifft תהיה חזרה 4 פעמים (פעמיים לאורך ופעמיים לרוחב) של התמונה המקורית.
עכשיו, מה אם נעשה shift לF (כך שF_{00} יהיה באמצע) ונעתיק את F למרכז של מטריצה בגודל $2n\times 2n$ (כאשר מסביב אפסים)? כשמחזירים למישור התמונה, מה נקבל?
התדרים ישארו בדיוק אותו דבר, אבל הגודל עליו הם מתפרסים גדל - לכן נקבל גירסה גדולה (מרוחה) של התמונה המקורית.