

## מציאת ספים - Edge Detection

המטרה: למצוא את גבולות הגזרה של כל אזור.

יש 3 תכונות ש-edge detection צריך לענות עליהם:

1. זיהוי טוב - כמה שפחות false positive<sup>1</sup> וכמה שפחות false negative<sup>2</sup>.

2. good localization - כמה שפחות לפספס במיקום של השפה.

3. single response - אנחנו רוצים שהעובי של השפה יהיה פיקסל אחד.

מה שמאפיין שפה זה גרדיינט<sup>3</sup> גדול - שינוי גדול ברמות האפור. בשביל למצוא גרדיינט צריך נגזרת - איך זה עובד בדו-מימד?

$$s_x = \frac{f(x + \Delta x, y) - f(x, y)}{\Delta x} = f(x + 1, y) - f(x, y) = \text{conv2}\left(i, \begin{bmatrix} 1 & -1 \end{bmatrix}\right)$$

$$s_y = \frac{f(x, y + \Delta y) - f(x, y)}{\Delta y} = f(x, y + 1) - f(x, y) = \text{conv2}\left(i, \begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \text{conv2}\left(i, \begin{bmatrix} 1 & -1 \end{bmatrix}'\right)$$

ואז

$$s = \sqrt{s_x^2 + s_y^2}$$

ואז בודקים אם זה עובר סף מסויים.

האם זה עובד? לא מספיק טוב. אם למשל יש לנו כדור, יש אזור שהוא מתחיל להיות כהה, ויהיה לנו טווח רחב שהשיפוע בו משמעותי. במקום לקבל single response תהיה לנו שפה מאוד רחבה. למעשה, הפתרון הזה לא עונה מספיק טוב על אף אחד משלושת הדרישות.

במקום זה יש אלגוריתם יותר טוב - Canny Edge Detector.

## Canny Edge Detector

זהו שיפור לאלגוריתם. קודם כל אנחנו רוצים single response בכיוון הגרדיינט. איך מוצאים את כיוון הגרדיינט? יש לנו  $s_x$  ו  $s_y$ , ולכן את הכיוון  $\theta$  אפשר למצוא באמצעות:

$$\tan \theta = \frac{s_y}{s_x}$$

במעבר בין האובייקט לרקע צריך להיות הגרדיינט הכי חזק. לכן יש לנו שלב בשם non-maximum supression שבו אנחנו מסתכלים על הפיקסל המרכזי. אנחנו רוצים לעגל את הזווית לכפולה של  $45^\circ$ , ובהתאם לזווית אנחנו יודעים מי שני השכנים שצריך לבדוק מולם. לא צריך לעשות arctan - אפשר להשתמש במונוטוניות של  $\tan$  בשביל העיגול:

$$\begin{cases} 0 & -0.4142 \leq \tan \theta \leq 0.4142 \\ 1 & 0.4142 \leq \tan \theta \leq 2.4142 \\ 2 & 2.4142 \leq |\tan \theta| \\ 3 & -2.4142 \leq \tan \theta \leq -0.4142 \end{cases}$$

לפי הכיוון מוצאים את שני השכנים, ומשווים להם. זה נותן לנו גם single response (כי בד"כ יש רק מקסימום אחד) וגם good localization (כי לוקחים את המקסימלי). איך משיגים זיהוי טוב?

<sup>1</sup>אמרנו שזה פיקסל שפה וזה לא פיקסל שפה.  
<sup>2</sup>אמרנו שזה לא פיקסל שפה וזה כן פיקסל שפה.  
<sup>3</sup>שיפוע

בד"כ יש לנו סף  $T$ , ואנחנו משווים מולו. אם  $T$  שלנו גבוה מדי יהיו לנו הרבה false negatives, ואם הוא נמוך מדי יהיה לנו הרבה false positives.  
במקום סף אחד נבחר שניים - סף גבוה  $T_H$  וסף נמוך  $T_L$ . את פיקסל השפה הראשון נמצא באמצעות  $T_H$ , ועבור פיקסלים סמוכים לפיקסלי שפה שכבר מצאנו נשתמש ב- $T_L$ . למנגנון הזה של קביעת שני ספים קוראים Hysteresis, והוא משמש למשל בטרמוסטטים.  
איך עושים את זה בפועל? עבור כל פיקסל:

```
if  $T_H < s(\text{neighbors})$   
     $T = T_H$   
else  
     $T = T_L$   
end
```

הבעיה בשיטה הזאת היא שאין פיעפוע. אבל אם נבחר  $T$  לפי השאלה אם השכנים הם שפה או לא, נגלה רק את השפות שנמצאות אחרי פיקסלי השפה שמצאנו. הפתרון - לעשות מעבר כפול. במעבר השני עוברים הפוך, וככה מוצאים שכנים לפי הכיוון השני.