מקליד: עידן אריה
מרצה: פרופ' נתן נתניהו
תאריך: 2018-12-05

עיבוד תמונות
89-560-01

בהינתן תמונה, ראינו איך אפשר לייצר מהפיקסלים של התמונה את פיקסלי השפה - אוסף של פיקסלים שהאופרטורים שהשתמשנו בהם תייגו אותם כפיקסלי שפה.

בתמונה מציאותית פיקסלי השפה הם לא סתם - יש סיבה שתהליך הedge detection תייג אותם כשפה. הם נמצאים כנראה על מעבר בין אזור בתמונה לאזור אחר, שיש שם גרדיינט - שינוי משמעותי ברמות האפור או בגוון. קרוב לוודאי שהוא נמצא על פני גבול של אובייקט. אם זה אובייקט מלאכותי אז כנראה שהם יהיו בקווים ישרים.

מה המטרה עכשיו? במקום אוסף פיקסלי השפה, אנחנו רוצים להתאים צורות גאומטריות - לאו דווקא קווים ישרים, אולי איזו קשת מעגלית. יש לנו איזה מודל מונח - מושג על הצורה של התמונה שמנסים לרכוש, ורוצים להתאים את פיקסלי השפה למודל הזה.

ייצוג קונטור

קריטריונים:
1. **יעילות**: ייצוג פשוט וקומפקטי.
2. **דיוק**: התאמה מדויקת לנקודות.
   זו לא גיאומטריה נקייה - גם אם יש לנו קו ישר לא בהכרח נוכל לייצג אותו באמצעות שתי נקודות, כי יש אילוצים (למשל הדיסקרטיזציה)
   מבחינה מתמטית אפשר לבדוק את זה באמצעות רגרסיה סטטיסטית.
3. **אפקטיביות**: מתאים לצרכים יותר מתקדמים של עיבוד תמונה או של ראייה ממוחשבת.

התאמת עקומה Curve Fitting

אפשר לייחס מרחק מנקודות השפה למודל המונח. נרצה איזה מדד להתאמה:

- maximum absolute error (MAE):

$$\mathrm{MAE} = \max_i |d_i|$$

- mean squared error (MSE):

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n} d_i^2$$

- normalized maximum error (NME):

$$\epsilon = \frac{\mathrm{MAE}}{S} = \max_i \frac{d_i}{S}$$

אם היה לנו אינסוף זמן היה אפשר להסתכל על כל המודלים בעולם ולבחור את זה שנותן את ההתאמה הכי טובה (בהתאם לקריטריון שבחרנו).

שיטות אינטרפולציה

דוגמה: מציאת polyline - ייצוג רב-קווי. יש כמה קטעים ישרים שצריך להתאים לנתונים, ויש אלגוריתם איטרטיבי/רקורסיבי שמוצא בכל פעם את הסטייה הגבוהה ביותר ומפצל אותו.

שיטות אפרוקסימציה

להבדיל מאינטרפולציה - שם ההתאמה חייבת לעבור דרך נקודות השפה - באפרוקסימציה הקירוב לא חייב לעבור דרך השפה.

שיטה אחת היא ordinary least squares:

Ordinary Least Squares

יש n פיקסלי שפה, ומסתכלים על סטיות אנכיות (על ציר הy) r_i. מנסים למצוא את הישר האופטימלי, שממזער את סכום ריבועי הסטיות:

$$\sum_{i=1}^{n} r_i^2 \to \min$$

הנעלמים הם a,b ממשוואת הישר $y=ax+b$, והנתונים הם x_i, y_i:

$$r_i = y_i - y\left(x_i\right) = y_i - \left(ax_i + b\right)$$

ולכן המשוואות הן:

$$0 = \frac{\partial}{\partial a}\left(\sum_{i=1}^{n} r_i^2\right) \qquad 0 = \frac{\partial}{\partial b}\left(\sum_{i=1}^{n} r_i^2\right)$$

ומבצעים את הגזירה (בלי לפתוח את הריבוע):

$$0 = \frac{\partial}{\partial a}\left(\sum_{i=1}^{n} r_i^2\right) = \cancel{2}\sum_{i=1}^{n}\left(y_i - ax_i - b\right)\left(-x_i\right)$$

$$0 = \frac{\partial}{\partial b}\left(\sum_{i=1}^{n} r_i^2\right) = \cancel{2}\sum_{i=1}^{n}\left(y_i - ax_i - b\right)\left(-1\right)$$

ומקבלים שתי משוואות בשני נעלמים.

ייצוג בהתאם לפרמטרים אלטרנטיביים

ייצוג באמצעות a וb הוא בעייתי, כי אי אפשר לייצג ככה קוים אנכיים. לכן נרצה להשתמש בייצוג אלטרנטיבי:

$$\rho = x\cos\theta + y\sin\theta$$

כאשר ρ הוא המרחק מראשית הצירים (בד"כ מרכז התמונה) וθ היא הזווית של הגובה מציר הx. בעצם (x,y) היא נקודת המפגש עם הגובה מראשית הצירים.

לכן הנוסחה שצריך למזער היא:

$$\sum_{i=1}^{n}\left(x_i\cos\theta + y_i\sin\theta - \rho\right)^2 \to \min$$

רגרסיה (Regression)

באופן כללי יש לנו משוואה סתומה:

$$f\left(x, y, a_1, a_2, \ldots, a_p\right) = 0$$

וצריך למצוא רגרסיה סטטיסטית ולפתור p משוואות עם p נעלמים. אם אפשר לבטא בצורה $y = f\left(x, a_1, a_2, \ldots, a_p\right)$ אז יותר קל.

Least Median of Squares

הממוצע הוא לא תמיד מדד טוב - לפעמים נרצה LME - התאמה שבה החציון של הסטיות הוא המינימלי.

Hough Transform

לכל נקודה (כלומר לכל פיקסל שפה) נחשוב על הישר הדואלי - כל נקודה תהפוך במרחב הפרמטרים (מרחב C, M) לישר:

$$(x_i, y_i) \rightarrow c = -x_i m + y_i$$

עבור x, y מסוימים יש הרבה ערכי c, m שמקיימים אותם - וכולם נמצאים על ישר אחד. אם חוזרים למישור התמונה - כל (c, m) מתאים לישר במישור התמונה, וכל הישרים האלה עוברים דרך (x_i, y_i).

אם כל ה(x_i, y_i) נמצאים על ישר אחד, אז צריך להיות ישר אחד (c, m) שעובר דרך כולן - ואז זה אומר שבמישור הפרמטרים כל הישרים (x_i, y_i) עוברים דרך הנקודה (c, m).

נשים ♡: במישור התמונה (x, y) הם נקודות ו(c, m) הם ישרים. במישור הפרמטרים (x, y) הם ישרים ו(c, m) הם נקודות.

מרחב הפרמטרים הוא דיסקרטי, אז אפשר לעבור תא תא על כל ערכי (m, c) האפשריים, ולספור כמה מישרי (x_i, y_i) עוברים דרך אותה נקודה. עכשיו מחפשים במרחב הפרמטרים נקודות בהן יש peak - מקסימום מקומי. אלו כנראה נקודות שמייצגות קווים משמעותיים.

נשים ♡: אנחנו לא מוגבלים כאן למציאת מופע בודד - יכול להיות שנמצא כמה peaks.

אלגוריתם:
1. בצע קוונטיזציה של מרחב הפרמטרים
2. אתחל ב0 מונה לכל תא
3. לכל נקודת שפה (x, y) במרחב התמונה, הגדל ב1 את כל התאים במרחב הפרמטרים המקיימים את משוואת המודל.
4. מצא נקודות במרחב השפה שהמונה שלהן מהווה מקסימום מקומי.

אבל נזכור שאי אפשר לייצג ככה קווים אנכיים! אז מה קורה אם נייצג את הקווים עם ρ, θ? הפונקציה המייצגת כל ישר הופכת להיות יותר מסובכת:

$$(x, y) \rightarrow \rho = x \cos\theta + y \sin\theta$$

אבל האלגוריתם נשאר אותו דבר.

קוונטיזציה בייצוג (ρ, θ)

אם הקו נמצא מתחת לראשית הצירים, אז הρ שלילי. לכן הθ צריך להיות בטווח $[-90°, 90°)$. הρ מחושב לפי $|\rho| \leq \frac{\sqrt{h^2+w^2}}{2}$.

מה לגבי גודל תאים?ככל שתא יותר קטן אז הוא ייתן יותר דיוק - אבל זה יכול לפגוע בקוונטיזציה. יש כאן tradeoff.

קשתות מעגליות

איך מתאימים קשת מעגלית? Hough Transform תקף לכל מודל. במקרה של מעגל - הנעלמים הם a, b, r:

$$(x_i - a)^2 + (y_i - b)^2 \approx r^2$$

זה נותן לנו "משטח" (למרות שהוא לא מישורי). ואת אותו אלגוריתם שעשינו עבור קווים ישרים אפשר לעשות עבור המשוואות האלו.

המשוואות הפרמטריות הן:

$$x = a + r\cos\theta \qquad y = b + r\sin\theta$$

או:

$$a = x - r\cos\theta \qquad b = y - r\sin\theta$$

אבל הזווית θ ידועה לנו - כי אנחנו יודעים מה הגדריינט של השפה, והוא אמור להיות מאונך למשיק באותה נקודת שפה. לכן ה-θ היא האוריינטציה של הגרדיינט.

לכן אפשר לסלק את r:

$$b = a\tan\theta - x\tan\theta + y$$

לכן אפשר להשתמש באינפורמציה הנוספת הזאת כדי לחסוך מימד אחד - הפרמטרים הם x_i, y_i, θ_i והנעלמים הם רק a, b.

אם מוצאים ככה peakים $\left(\hat{a}, \hat{b}\right)$, עדיין צריך למצוא את הרדיוסים של הקשתות. איך עושים את זה? עושים עוד Hough Transform, אבל בחד-מימד. הפעם הפרמטרים הם $x_i, y_i, \hat{a}, \hat{b}$ והנעלם היחיד הוא r, ומשתמשים במשוואה:

$$\left(x_i - \hat{a}\right)^2 + \left(y_i - \hat{b}\right)^2 \approx r^2$$

וככה עושים לכל $\left(\hat{a}, \hat{b}\right)$ שמוצאים, כדי למצוא את כל הקשתות (המעגלים, בעצם) המעניינות.