עבודה בראיה ממחושבת

בנושא רגיסטרציה של תמונות (Registration)

מגיש: נועם קציר

מרצה: דר' קורנל לוסטינג

הקדמה

אמא שלי, שקוראת בכפות ידיים, טוענת שקיימת גישה לפיה ניתן להיעזר בקריאה בכף היד כדי לנתח מרכיבים מאישיותו של אדם, רגשותיו, מה הוא חווה בהווה ואיזה פוטנציאל יש לו לעתיד. הרעיון של קריאה בכף היד נשמע קסום, ומיד אנו עלולים לחשוב שמדובר בשרלטנות בגלל הדימוי שנוצר לתחום, אך מדובר בגישה שמבוססת על הפסיכותרפיה היונגיאנית (דיסציפלינה בפסיכולוגיה). כמו כן, כאשר מתעמקים, מסתבר שזהו הליך בו אנו מחפשים קווים מסוימים בכף היד ולפיהם ניתן להבחין בתכונות מוגדרות.

מקובל לבצע הטבעה של כף היד באמצעות דיו, פשוט לוקחים גלגלת עם דיו, מעברים אותה על היד ואז מניחים את היד על דף נייר ומטביעים הדפס של כף היד. על ההדפס נהוג לרשום איזו יד זו, שם האדם, תאריך, ולעיתים גם הטבעה נוספת של האגודל של אותה יד באזור מרוחק של הדף.

אמי מנהלת זה שנים, מאז התחילה ללמוד קריאה בכף היד, מאגר של כפות הידיים של בני המשפחה לאורך השנים וניתן להבחין באופן ההתפתחות של האדם לאחר צפייה בסדרת ההדפסים.

לכן, עלתה לי השאלה, מדוע שלא אשתמש בגישה של ראיה ממחושבת לחלץ את אותן התכונות שמחפש אדם שקורא בכף היד, באופן אוטומטי? הנחתי שבאמצעות ראיה ממוחשבת, ניתן להבחין בדברים אשר קשה להבחין בהם בעין אנושית, ולכן ניתן למצוא את השינויים שקיימים בין שני הדפסים אשר ביניהם פער זמן מסוים.

כדי שאוכל לעשות זיהוי מדויק של שינויים חשוב לבצעה התאמה בין התמונות, למקם את כפות הידיים בכל הדפס באותו מיקום כך שנוכל לבצע פעולות כגון חיסור תמונות כדי לזהות את השינויים. וכך למעשה הצלחתי להגדיר בעבודה זו אלגוריתם של רגיסטרציה (Registration) על תמונות כף יד, כשלב מקדים לאלגוריתם עתידי שיוכל לבצע קריאה לכף היד.

כעט ניגש לתיאור האלגוריתם.

האלגוריתם

לשם הפרויקט השתמשתי בתמונות של כפות הידיים שלי בשני הדפסים משנת 2008 אחד לכל יד, ומשנת 2015 אחד לכל יד, תוך שימוש בתמונות בסריקה ברזולוציה גבוהה (600X600DPI).

לשם ביצוע הרגיסטרציה רציתי למצוא תחילה את הידיים, למפות מהו מרכז כף היד והיכן ממוקמות האצבעות, ואז חשבתי שאסובב את הידיים לזווית סיבוב האחידה ולבסוף אבצע גרירה של מרכז כף היד של אחת לעבר השנייה.

אחלק את שלבי כתיבת האלגוריתם ל-12 שלבים:

1. בשלב הראשון שאלתי, כיצד אוכל להבדיל אילו מהפיקסלים בתמונה הם אכן של כף היד?

ראה תמונה של יד שמאל משנת 2008:



כפי שניתן לראות, בתמונה זו יש מספר בעיות נוספות מכיוון ויש כאן גם את היד אך גם אגודלים נוספים וכן תיעוד של ההדפס שעלול להפריע למהלך יישום של האלגוריתם.

לכן, לאחר מספר ניסיונות למצוא את היד תוך התעלמות מרעשי הרקע כמו כתב היד והאגודלים הנוספים, הגעתי למסקנה כי יש להגדיר מספר אילוצים לאלגוריתם כדי לעבוד עם התמונות:

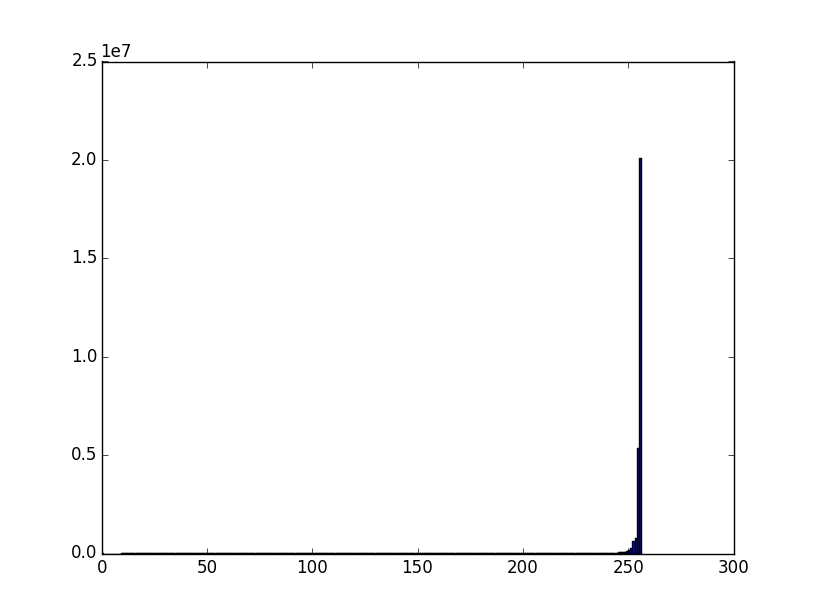
* בתמונה יש רק יד אחת
* ידוע לי לגבי כל תמונה של איזו יד זו ומאיזו שנה
* אין הערות בכתב יד ואין אגודלים נוספים

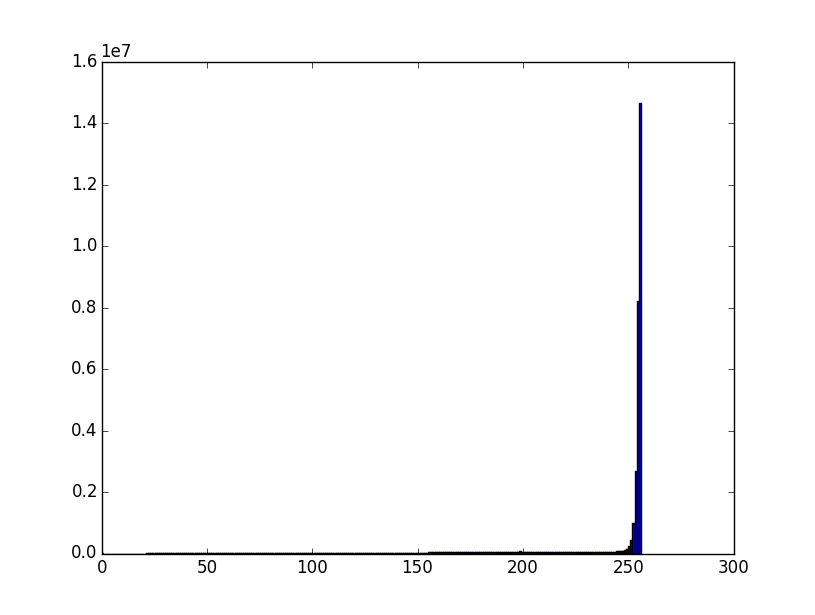
אז לצורך כך ביצעתי עריכה לתמונות אלו כדי לעמוד בדרישות, להלן דוגמה לתמונה לאחר העריכה:



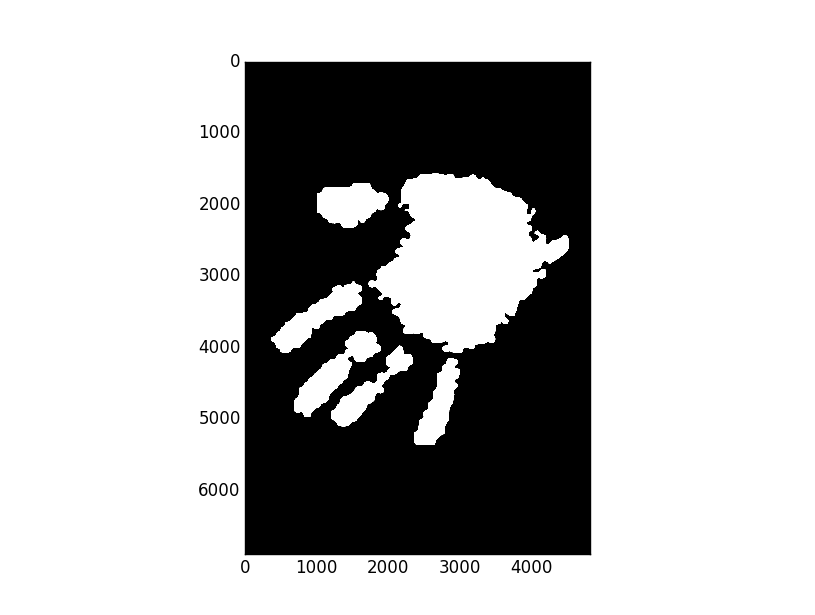
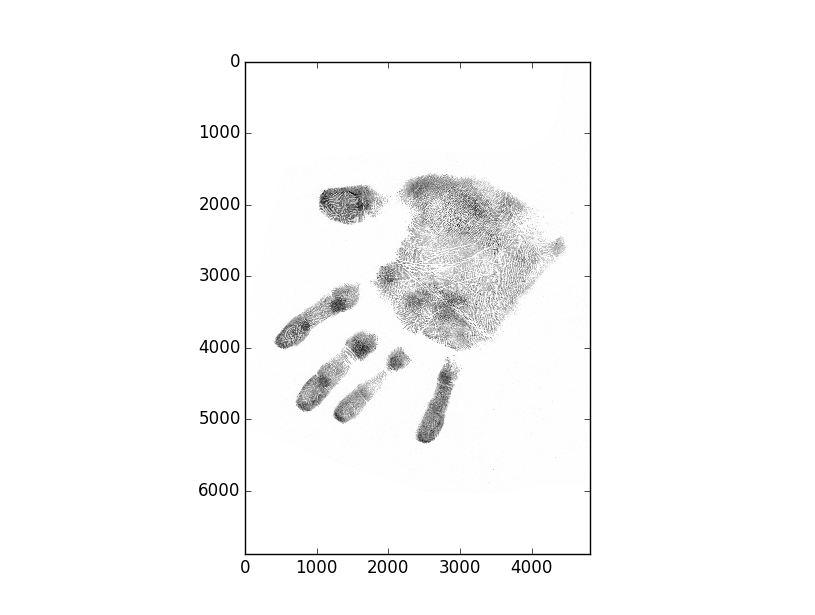
1. כעת, כאשר קיבלתי תמונה המכילה רק יד, השאלה השנייה הייתה כיצד אמצא את הפיקסלים של היד ואבצע הפרדה של הפיקסלים של כף היד מהפיקסלים אשר שייכים לרקע?

לאחר שציירתי היסטוגרמה של מספר תמונות הגעתי למסקנה כי כל פיקסל שהצבע שלו ב grayscale נמצא בין ציון סף 0 ל 245, מכיל בעיקר את היד, ולא את הצבע של הרקע שהוא כמובן לבן (255)





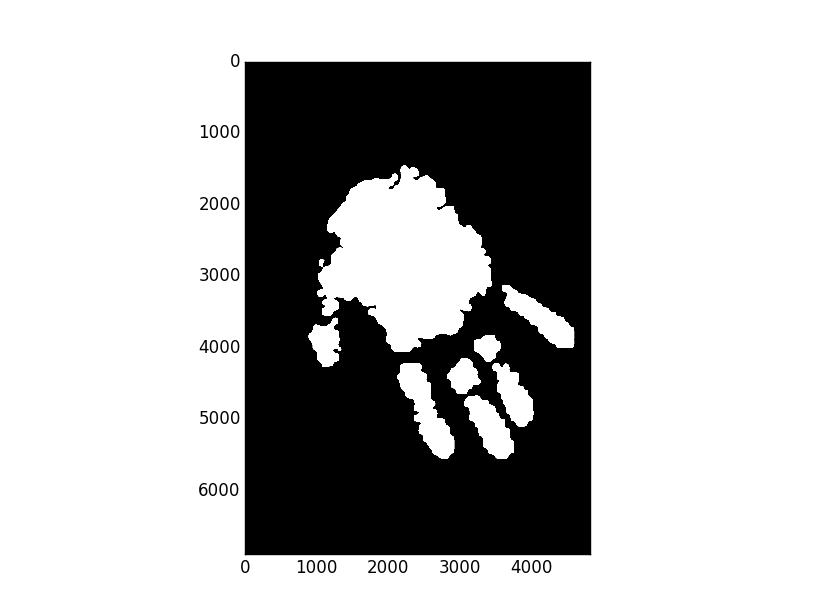
1. השאלה השלישית הייתה, כיצד למצוא את קווי המתאר שמקיפים את כף היד כדי שבעזרתם נוכל לבנות תמונה בינארית של כף היד, לשם כך נעזרתי בפונקציה cv2.findContours(), בעזרתה ניתן למצוא את כל קווי המתאר, (הקונטורים) בתמונה, אך אם נבצע מילוי של כל הקונטורים בתמונה אנו עלולים לקבל אזורים המכילים פיקסלים בודדים. אם כך, שאלה נוספת היא כיצד נבדיל בין פיקסלים טועים לבין האזורים הגדולים? מסתבר כי עבור כל קונטור יש מספר פעולות שניתן לקבל אודותיו אחת מהן מאפשרת לנו לחשב את השטח הכלוא בתוך הקונטור, ובעזרת מציאת ציון סף מתאים נוכל להתעניין רק בקונטורים החשובים. מכיוון ומדובר בתמונות גדולות יחסית הציון שנקבע היה 300000. כלומר כל מה שגדול משטח זה הוא כנראה אזור עניין חשוב. כך הצלחנו להגיע לתמונה הבינארית הבאה עבור תמונה של יד שמאל מ 2015:



1. לאחר מכן, חישבתי את קונטורים של התמונה הבינארית המפורשת, ויצרתי רשימת אובייקטים עבור כל קונטור ושמרתי עבורו מספר תכונות מעניינות:

* השטח שלו
* האם הוא מרכז כף היד (קונטור בעל השטח הגדול ביותר)
* את מרכז הכובד שלו (תוך שימוש במומנטים של קונטור)
* האם כבר עברתי עליו וקבעתי לאיזו אצבע הוא שייך
* את המרחק של מרכז הכובד של כל אובייקט ממרכז הכובד של מרכז כף היד

1. מכאן נדמה היה שאני קרוב מאוד לסיום, אלא שבחלק מההדפסים האצבע לא התקבלה כמקשה אחת והיית מורכבת ממספר חתיכות, ראה תמונה:



ניתן לראות באופן ברור כי יש שתי אצבעות עבורן יש הפרדה בתמונה הבינארית במפרק.

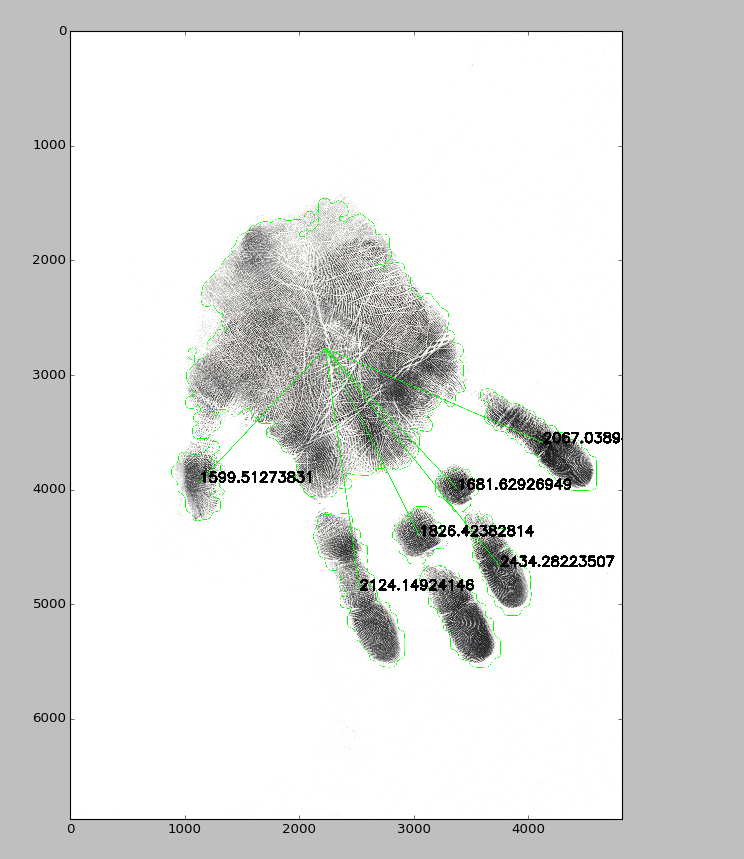
התלבטתי כיצד אמצא כי חלק מסוים שייך לאצבע מסוימת.

חשבתי שאם אצור תמונה בינארית אשר מכילה קו, שעובר דרך הקונטור של קצה האצבע, ואז עבור כל קונטור של אובייקט מבין האובייקטים האחרים אצור תמונה שמכילה מילוי של הקונטור. אז אוכל לבצע כפל איבר-איבר בין התמונה הבינארית של הקונטור המלא לבין התמונה הבינארית שמכילה את הקו. אם תוצאת ההכפלה תהיה גדולה מאפס, סימן שאלמנט בינארי זה שייך לקצה הזה של האצבע. כך אעשה עבור כל קצות האצבעות, ואוכל למפות את הרכיבים של כף היד במדויק.

1. אך נתקלתי בבעיה בניתוח הזה, כלומר, מה הן באמת קצות האצבעות, ומה הם באמת האלמנטים הרחוקים ביותר ממרכז כף היד, ומה מבטיח לי שחלק מהמפרק של האצבע השלישית לא רחוק יותר ממרכז כף היד מאשר האגודל?

הגעתי למסקנה שרק לקבוע שקצות האצבעות הם האיברים הרחוקים ביותר זה לא מספיק.

לכן, החלטתי שאעבור על כל האובייקטים כאשר הם מסודרים לפי המרחק ממרכז כף היד בסדר יורד ואבצע את הבדיקה על כל אחד מהם אם עוד לא מיפו אותו לאיזו אצבע, ואז אבצע את האלגוריתם לשיוך לאצבעות הנ"ל וכך מכיוון שחלקים מסוימים עברו מיפויי אין מה לדאוג משום שבוודאות קצה של אצבע רחוק מחלק במפרק של אותה אצבע מהמרכז, להלן ראה תמונה לדוגמא:

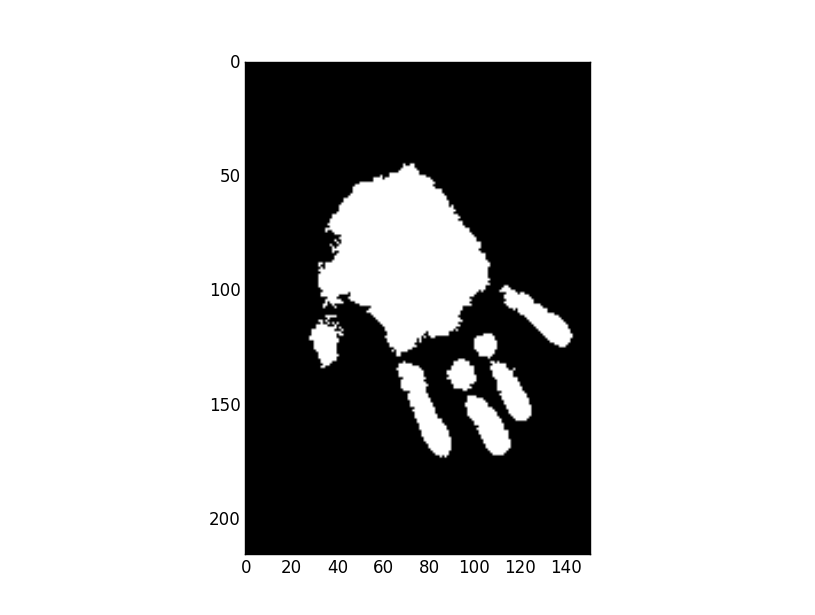


1. בשלב זה, השאלה הבאה הייתה איך באמת אמצא קו שמתאר את הקונטור של קצה האצבע?

ראיתי כי ל OpenCV יש יכול התאים אליפסה שמקיפה את הקונטור בצורה אידיאלית או להעביר דרכו את הקו העובר דרך ציר הסיבוב שלו.

לאחר מספר ניסיונות הבנתי שאין הבדל מהותי בין הכיוון של הרדיוס הגדול של האליפסה לכיוון הישר העובר דרך הקונטור. אך הייתה בעיה אחת עם הרעיון הנ"ל, מדובר בתמונות גדולות מאוד, ולכן חישוב של כפל איבר-איבר יכול לקחת זמן רב.

1. כדי להתגבר על הבעיה יצרתי תוך שימוש בפירמידות תמונה קטנה של התמונה וגם של התמונה הבינארית. על התמונות הקטנות (האיבר החמישי בפירמידה שמידותיו קטנות ביחס של ביחס למקור), יכולתי לבצע את החישוב הנ"ל במהירות וביעילות רבה ויתר, כך יכולתי למפות את האצבעות בעזרת האלגוריתם הנ"ל עם עליות חישוב מזעריות. ראה דוגמא לתמונה בינארית מוקטנת:



מכאן הסקתי שאין כל חשיבות לביצוע החישוב על התמונות הגדולות וניתן לבצע את כל החישובים על התמונות הקטנות (ראה דוגמא לתמונה מוקטנת בינארית), ורק לבסוף לחזור לתמונה המקורית.

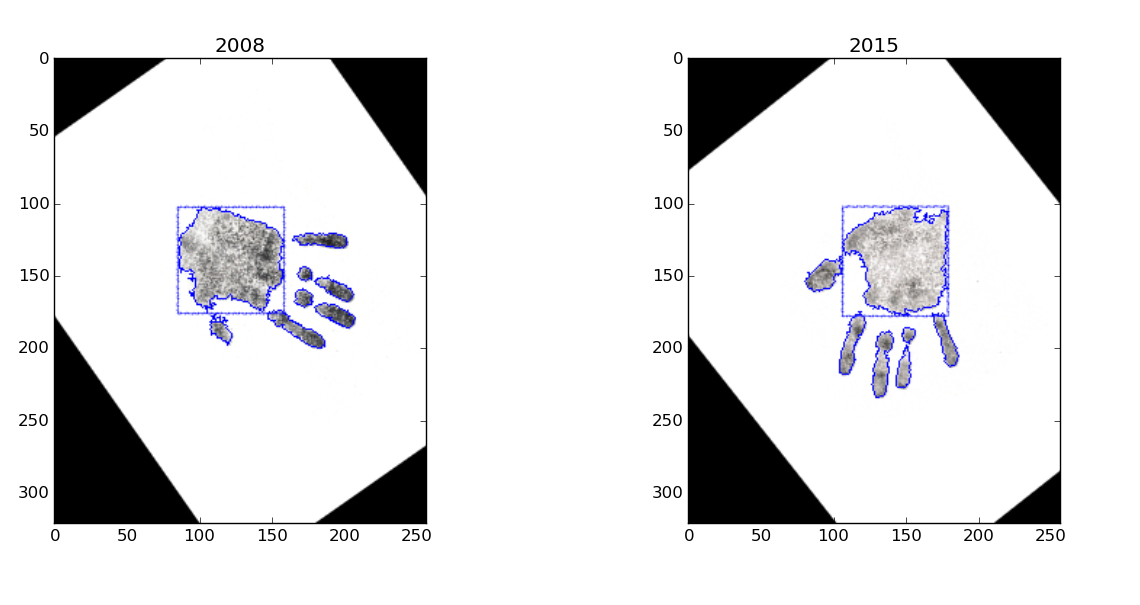
1. מכאן עלתה השאלה, כיצד חוזרים לאחר החישובים לתמונה המקורית?

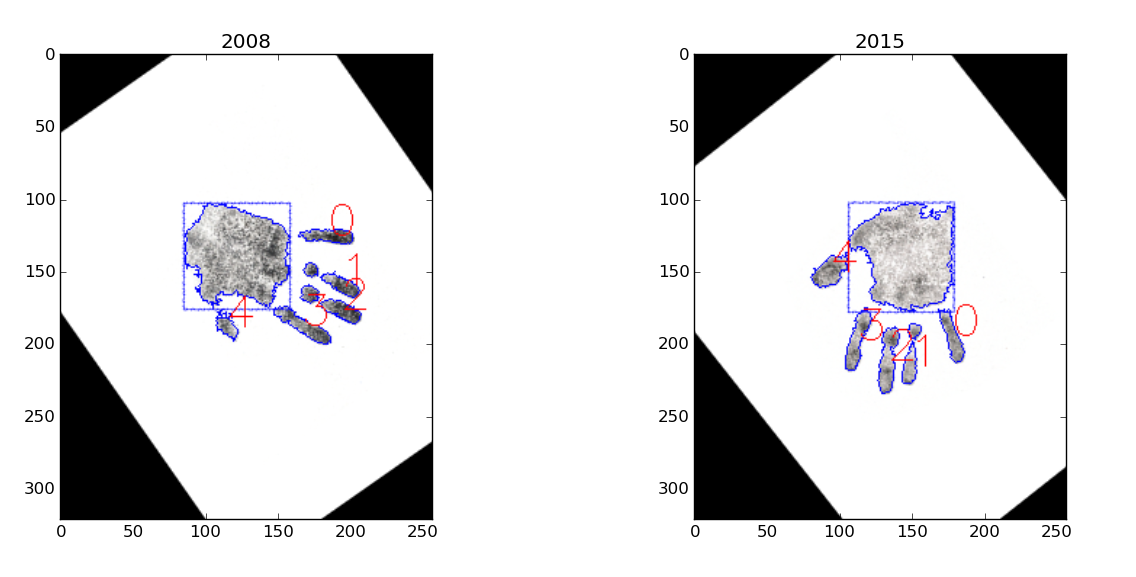
לכך יש שתי תשובות אפשריות:

* ניתן לצייר את הקונטור הקטן לאחר החישובים, ואז לבצע הגדלה על התמונה הבינארית ולאחר מכן למפות מחדש את הקונטורים והמרכזים של כל חלק. לאחר חיפוש ברשת האינטרנט [1] מצאתי כי לפתרון זה יש חסרון בולט משום שעלולים להיווצר עיוותים מסוג artifacts.
* מסתבר שהקונטור שמתקבל מ OpenCV הוא למעשה רצף של נקודות שמרכיבות את הקונטור, ולכן ניתן פשוט להכפיל את כל הנקודות של הקונטור ביחס ההגדלה בין התמונה הקטנה לגדולה. כמו כן, להכפיל באופן דומה את מרכזי הכובד של האובייקטים לאחר החישוב. פתרון זה עדיף על הפתרון הראשון כיוון שהוא מדויק יותר וחסכוני יותר מבחינת זמן הריצה שלו.

1. כעת, לאחר המיפוי של האצבעות רק נותר למספר כל אצבע, איך עושים את זה?

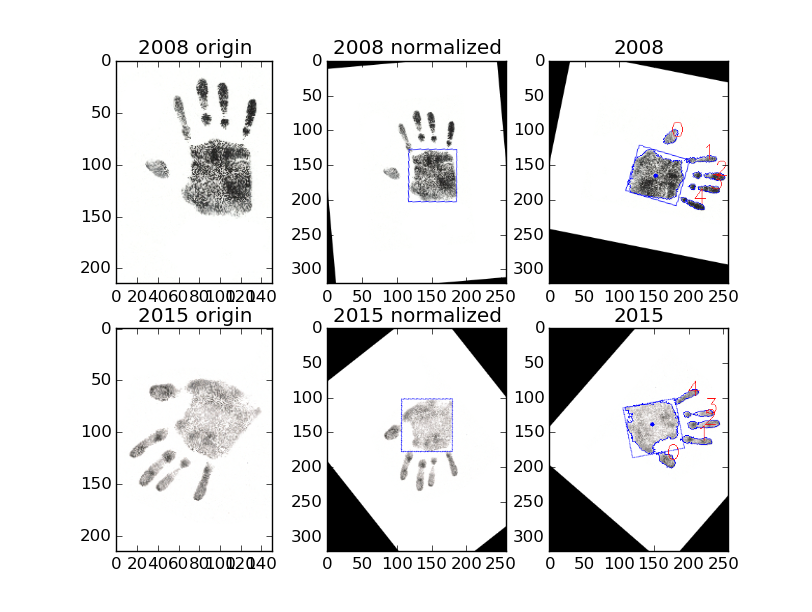
אילו יכולתי לנרמל את מקום היד בצורה ברורה כך שהיד תפנה לאחד מהכיוונים: ימינה, שמאלה, למעלה או למטה, יהיה קל למפות את האצבעות בידיעה באיזו יד מדובר. כך למשל, אם היד היא יד שמאל והיא מכוונת כלפי מעלה אז ארבע קצות אצבעות מתוך 5 יהיו בעלות מרכזי כובד מעל מרכז הכובד של מרכז כף היד, וההצבעות יהיו מסודרות מאגודל ועד זרת מימין לשמאל.

גיליתי שניתן באמצעות OpenCV למצוא את המלבן המסובב שמתאר בצורה מיטבית את הקונטור. חישבתי את המלבן הנ"ל עבור הקונטור של מרכז כף היד ואז סובבתי את הקונטורים של האצבעות ואת מרכזי הכובד שלהם ביחס לציר סיבוב של מרכז כף היד בעזרת טרנספורמציית סיבוב בזווית בה נוטה המלבן, ולאחר מכן מיינתי את האצבעות לפי השיטה הנ"ל כך שהתקבל: 



1. בשלב זה, השאלה הייתה, כיצד לסובב יד אחת כלפי השנייה ולמקם אותן אחת תחת השנייה, חשבתי שאוכל לקבוע את הזווית של היד ביחס לראשית הצירים ולסובב את שתי הידיים כך שיהיו בזווית אפס מעלות ביחס לצירים. לשם כך קבעתי שלפי שלב הנרמול ניתן לדעת את הזווית במונחים של 0,90,180,270 מעלות. אם כך, כיצד ניתן למצוא את הזווית של היד ביחס למלבן?

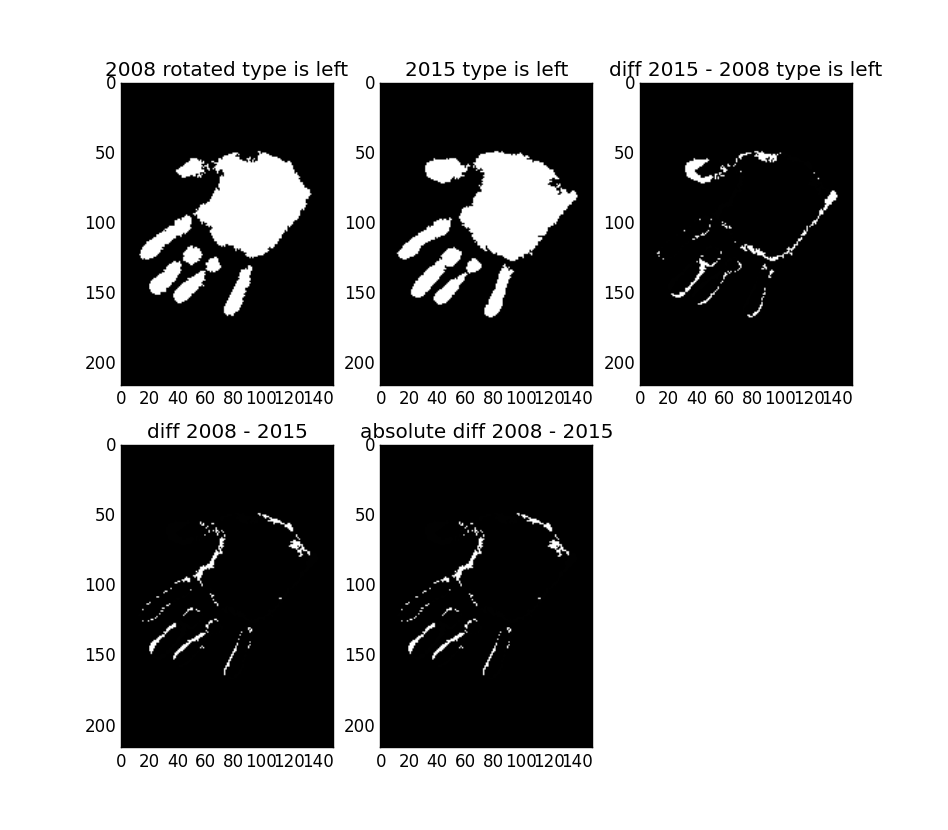
הנחתי שמספיק להשתמש בזווית של האצבע השלישית שנמצאת יחסית במרכז ולחשב את הזווית יחסית לנטייה הכללית. כך הגעתי למצב הבא:



נדמה היה שהגעתי כמעט להתאמה מלאה, אך כאשר ביצעתי חיסור בין התמונות גם לאחר הזזה ביחס להפרש בין המרכזים, הגעתי לדיוק לא מספיק טוב.

1. בעזרת חומר שלמדתי בקורס "ראיה ממוחשבת רבת מבטים", למדתי שניתן לשערך את הטרנספורמציה של הסיבוב וההזזה שבין שתי התמונות, על ידי רגרסיה על נקודות עניין בשתי התמונות בעזרת קירוב הריבועים הפחותים. כמובן ל OpenCV יש פונקציה מובנית המאפשרת חישוב שכזה cv2.estimateRigidTransform()

וכמובן יש לנו בעצם 5-6 נקודת עניין ברורות. קצות האצבעות ומרכז כף היד (לא בכל התמונות יש מפרקים נוספים). ובחלקן האגודל מתחבר למרכז כף היד ולכן בחרתי למפות את האצבעות מהזרת לאגודל כדי שאף אם יהיו פחות נקודות, הזיהוי עדיין יצליח. כך בעזרת הטרנספורמציה הנ"ל הצלחתי לבצע התאמה על סמך התמונות הבינאריות בצורה מיטבית (ראה תמונה)



לאן ניתן להתקדם מכאן?

מפאת חוסר הזמן, את ההתאמה ביצעתי רק על התמונות הקטנות ומכאן יש לבצע חזרה לתמונות המקור, כפי שהסברתי קודם לכן.

כמו כן, ההתאמה כאן לא לקחה בחשבון את הקווים המסומנים על כף היד דבר אשר דורש אופטימיזציה נוספת.

הניסיון לבצע אופטימיזציה על ידי פונקציה של ההפרש בין התמונות בערך מוחלט שלאחר ההתאמה, כאשר הפרמטרים של הפונקציה הם הזזה וסיבוב של אחת התמונות כלפי השנייה, תוך שימוש בפונקציה scipy.optimize.minimize() לא הצליח לי בצורה טובה.

אני מאמין כי קיימות שיטות טובות יותר בשביל השלב העדין של הרגיסטרציה לאחר שביצעתי את ההתאמה הנוכחית כשלב מקדים.

וכמובן ניתן להתקדם ולהוסיף אלגוריתם שמנתח את כף היד ומבצע קריאה בכף היד.

מיקום הקוד שמופיע בדו"ח זה

עבודה זו נכתבה כמובן בעזרת Python(2.7.x) תוך שימוש ב OpenCV(2.4.11), וכל הקוד של הפרויקט הנ"ל זמין להורדה ישירות מ GitHub, בכתובת הבאה:

<https://github.com/noamkatzir/palm-hand-reading>

ביבליוגרפיה

[1] <http://answers.opencv.org/question/44580/can-i-resize-a-contour/> פוסט שדן בדרכים לביצוע resize לקונטור

[2] <http://www.pyimagesearch.com/category/building-a-pokedex/> סדרת מאמרים בבלוג של "Adrian Rosebrock" שמדברת על בניה של "פוקדע" שמשתמש בשיטות של ראיה ממוחשבת בעזרת python. מצאתי בדיעבד החומרים שם נתנו לי אישור רשמי שאני בכיוון הנכון