

חלק ב' – רקע ודו"ח מכין

בחלקה השני של המעבדה אנו נעסוק בעיבוד של וידאו - דהיינו סדרה של תמונות, עוקבות כרונולוגית, וקרובות אחת לשנייה בזמן (ולרוב גם במקום).

- מומלץ לקרוא בויקיפדיה תחת הערך "וידאו":

<http://he.wikipedia.org/wiki/%D7%95%D7%99%D7%93%D7%90%D7%95>

כיוון שאנו נעסוק בוידאו, הרי שמלבד הפרטים הנתונים בכל תמונה בנפרד, נרצה גם לחלץ מידע המתפתח בזמן, כגון תנועה, שינויי תאורה וכו'.

1. מרכז המסה

על-מנת לעקוב אחרי תנועה של עצמים (קשיחים), נרצה לעקוב אחרי מרכז המסה שלהם.

$$\mu_x = \left(\frac{\iint xf(x,y)dxdy}{\iint f(x,y)dxdy} \right) \quad \blacksquare \text{ מרכז מסה בציר ה-} x :$$

1.1 מהו עצם קשיח? מהו עצם שאינו קשיח? תנו דוגמאות.

1.2 ממשו פונקציות MATLAB: `[x_com,y_com]=calc_com(bw_im)`, המקבלת בתור קלט תמונת שחור לבן (מסכה), ומחשבת את מרכז המסה של התמונה.

1.3 ממשו פונקציות MATLAB: `[x_com,y_com]=calc_com_vec(bw_im)`, המקבלת בתור קלט תמונת שחור לבן (מסכה), ומחשבת את מרכזי המסה של כל העצמים השונים בתמונה. הפונקציה תחזיר את וקטורי מרכזי המסה של העצמים השונים. (רמז: השתמשו בפונקציה "bwlabel")

1.4 טענו את התמונה "pieces_bw.bmp" והפעילו את הפונקציה הנ"ל על התמונה המתקבלת. סמנו את נקודות מרכז המסה השונות בתמונה (ע"י הצבה של הערך 2 בנקודות אלו למשל). הציגו את התוצאה.

2. למידה

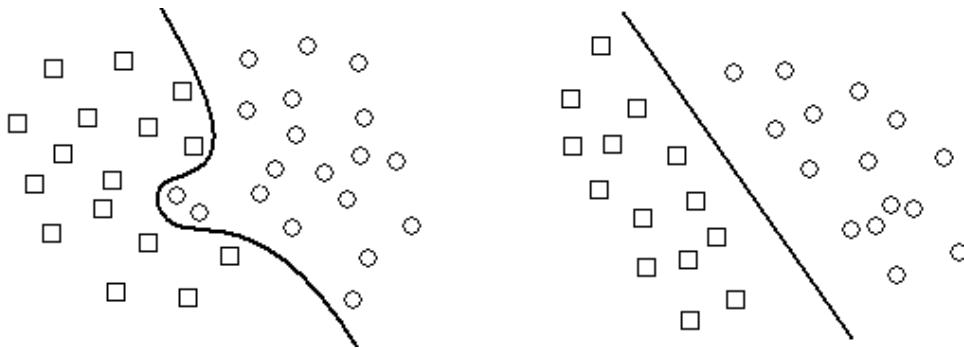
אחד מן הכלים החשובים ביותר בסיווג אוטומטי הוא השימוש במערכות לומדות. אנו נרצה להשתמש במערכת לומדת מסוג SVM (Support Vector Machine) על-מנת ללמוד ובהמשך למצוא מאפיינים שבעזרתם נרצה לסווג.

- ניתן לקרוא בויקיפדיה תחת הערך "support vector machine" :

http://en.wikipedia.org/wiki/Support_vector_machine

- רשות: קראו את המאמר "A Tutorial on v-Support Vector Machines" של: "Pai-Hsuen Chen, Chih-Jen Lin, Bernharn Scholkopf"

המערכת הלומדת מקבלת כקלט סט של נקודות לימוד ותיוגם הנכון, ומניבה כפלט חלוקה של מרחב הדגימות לאזורי החלטה :



בשלב הסיווג, המערכת תבצע החלטה לגבי זהותו של האובייקט המתויג לפי אזור ההחלטה שאליו הוא שייך. הערה: קיימות גם מערכות לומדות המניבות את הסבירות שהאובייקט אכן שייך לכל אחד מאזורי ההחלטה. וכמו כן את שסבירות שהוא אינו שייך לאף אחד מהמקבצים (clusters).

2.1) הסבירו מהו מסווג חלש ומהו מסווג חזק. הסבירו כיצד ניתן ליצור מסווג מורכב ממספר מסווגים לינאריים. הסבירו מהי "התאמת יתר", כיצד מקבלים אותה, מתי היא שימושית ולחילופין מהן הסכנות הטמונות בה.

2.2) הורידו מאתר הקורס את הקובץ "SVM.zip"

(2.3)

- צרו סט לימוד של 4 קבוצות בגודל 100 של ו"א גאוסיים דו מימדיים בת"ס (השתמשו ב- "randn(100,2)", עם תוחלות [2,-2],[-2,-2],[-2,2],[2,2], בהתאמה, ושונות של 0.5. הקצו ערכי סיווג של 1,2,3,4 לקבוצות.
- הציגו את ארבע הקבוצות ב-plot אחד, הקצו צבע שונה לכל קבוצה לפי השיוך הראשוני שלה. למדו את המערכת ע"י שימוש בפקודה: `model = svm_train(y_train,X_train)`
- צרו סט בדיקה של 400 ו"א גאוסיים דו מימדיים בת"ס עם תוחלת אפס ושונות 1.

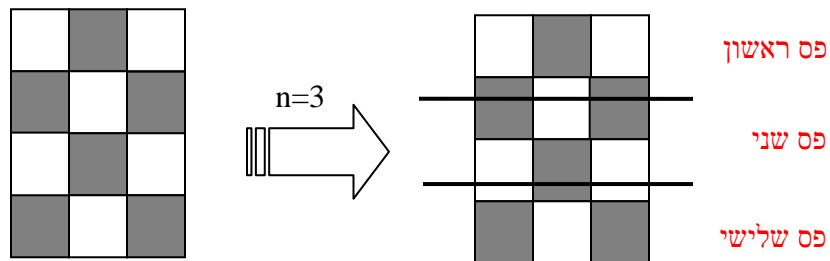
- סווגו את נתוני הבדיקה ע"י שימוש בפקודה:

`Labels_Test = svm_test(rand(size(X_test,1)), X_test, model)`

- הציגו את תוצאת הסיווג ב-plot אחד, הקצו צבע שונה לכל קבוצה לפי הסיווג שלה. שימו לב לגבולות איזורי ההחלטה של המסווג, מה אתם יכולים להגיד עליהם?

(2.4)

- ממשו פונקציות MATLAB: `partial_mass=find_partial_mass(mask,n)` המקבלת כקלט תמונת שחור-לבן. הפונקציה מחלקת את התמונה ל- n פסים שווים. שימו לב כי אם גובה התמונה אינו מתחלק ב- n , הפונקציה תחלק את השורה המתאימה באמצע השורה (sub-pixel division).
- לדוגמא: תמונה שגובהה ארבע פיקסלים, מחולקת ל-3 פסי גובה.



- הפונקציה תחזיר וקטור המכיל את אחוז המסה היחסי בכל פס גובה (שימו לב כי סכום האברים של וקטור הפלט יהיה 1). בדוגמא שלנו הפלט יהיה: $(5/18, 1/3, 7/18)$.

(2.5) הורידו מן האתר את התמונה "Yin_Yang.jpg", הריצו עליה את הפונקציה עם $n=2,3,4,5,6,7$ הציגו את התוצאות.

(2.6) הריצו את הפונקציה עם $n=2,3,4,5,6,7$ על תמונה אקראית: `mask=rand(17,1000)`. הציגו את התוצאות.

חלק ב' – ניסוי מעבדה

הורידו מאתר הניסוי את הקובץ "data2" המכיל את התמונות ופונקציות ה-MATLAB עבור החלק השני של הניסוי.

במהלך חלק זה של הניסוי ננסה לממש מערכת המזהה ומאפיינת תנועה של כלי רכב (או עצמים אחרים) בסרט וידאו, ובמקרה הצורך יודעת לזהות את כלי הרכב לפי מספרו. את המספרים אנו נזהה ע"י שימוש במערכת לומדת היברידית (משולבת), של עץ החלטה ולמידה.

1. טעינת וידאו

טענו את הקובץ `blurred_car_with_digits` המכיל רצף של תמונות של מכונית בתנועה. קובץ זה מכיל `Cell_Array`, המכיל סט של מסגרות (פריימים) של סרטון וידאו. ניתן לגשת לתמונה ה-`n` ע"י הפקודה:
`image_n=blurred_car_with_digits{n}`

2. כיוול

בתהליך רכישת התמונה קיים רעש. אנו נרצה למצוא ערך סף הגיוני לעוצמת רעש זה. (1) חלצו את תמונת ההפרשים (בערך מוחלט) בין המסגרת הראשונה והאחרונה בסרטון (תמונה זו תהיה בשחור-לבן). הציגו את ההיסטוגרמה שלה. (2) בחרו ערך סף הגיוני, על פי ההיסטוגרמה, כך שמעליו כל פיקסל הפרש יהיה חשוד כעצם חדש שנכנס לתמונה. הסבירו איך בחרתם את הסף ואיך בחרתם את טווחי ההיסטוגרמה.

3. ניתוח הוידאו

- צרו `Cell_Array` חדש המכיל את המסכה של "הרכב" בלבד בכל מסגרת מתאימה.
 - חשבו את ההפרש (בערך מוחלט) בין המסגרת המתאימה והמסגרת הראשונה.
 - סננו פיקסלים אשר ערכם מתחת לסף הנ"ל.
 - בצעו פעולות סינון/חיזוק/מילוי, בדומה לחלק א' (הניחו כי מימדיו של הרכב הינם לפחות כמה מאות פיקסלים). צרפו את הפרמטרים הרלוונטיים לדו"ח המסכם.
- השתמשו בסרטון זה על-מנת לחלץ את הפרמטרים הבאים (בצורה אוטומטית):
 - מספר המסגרת בה המכונית נכנסה לראשונה לתמונה, ומספר המסגרת האחרונה בה המכונית עדיין בתמונה.
 - מספר המסגרת בה המכונית נכנסה לראשונה לתמונה במלואה, ומספר המסגרת האחרונה בה המכונית עדיין בתמונה במלואה.

פרטו כיצד בצעתם משימות אלו.

- עבור כל אחת מן המסגרות בהן המכונית בתוך התמונה במלואה חשבו את מרכז המסה והמהירות (ביחידות של פיקסל למסגרת) של הרכב.
 - חשבו את המהירות הממוצעת (וקטורית), והמקסימלית של הרכב במהלך תנועתו.
 - חשבו את הזווית (הממוצעת) בה נע הרכב (שימו לב כי ראשית הצירים היא הפינה השמאלית העליונה).
- 3) הציגו בגרף (דו-מימדי) את נקודות מרכז המסה של הרכב לאורך תנועתו. תארו את תנועת הרכב בזמן בקצרה.

4. סיווג ספרות ע"י למידה

- 1) טענו את הקובץ "digit_learning.m" המממש מערכת המנסה ללמוד לזהות ספרות 0-9 ע"י חילוף מאפייני הספרות.
- 2) כפי שניתן להיווכח המערכת אינה עושה עבודה טובה מספיק בשלב זה. התבוננו בספרות, ונסו לחשוב על מאפיינים נוספים שיכולים לשפר את הסיווג. אלו מאפיינים הינם חשובים בספרות, לדעתכם?
- 3) הביאו את המערכת לאפס אחוז שגיאת סיווג אימון, ופחות מ-10 אחוז שגיאת סיווג מבחן. פרטו את המאפיינים בהם השתמשתם. מדוע לא כדאי להביא את שגיאת סיווג מבחן לאפס אחוז גם כן?

5. זיהוי מספר רכב

- 1) טענו שוב את הקובץ blurred_car_with_digits המכילה רצף של מסגרות של מכונית בתנועה. התמונות מטושטשות ומקשות על תהליך הזיהוי. עתה ננסה לחדד את התמונה על-מנת לקבל את הספרות בצורה טובה יותר. באופן כללי ניתן לתאר את תהליך הרכישה של תמונה של תנועה, כקונבולוציה בין האובייקט הנייח ומסנן תנועה מתאים. לכן ניתן להפוך חלק מתופעת הטשטוש ע"י הפעלת פעולת קונבולוציה הפוכה (deconvolution).
- 2) התבוננו באחת מן התמונות המטושטשת ונסו לשערך את אורך המריחה שנוצרה בעת תהליך רכישת התמונה (השתמשו ב-zoom ובפקודה-pixel on אם יש צורך). העזרו גם בנתוני מהירות הרכב אותם חיבתם מקודם. צרו מסנן היוצר עיוות בכיוון התנועה של הרכב ע"י פקודת MATLAB:
- $$FPS = fspecial('motion',length,teta)$$
- 3) חדדו את התמונה ע"י הפקודה (השתמשו בערך ב-10-15 איטרציות):
- $$Sharp_image = deconvblind(image,FPS,NUM_ITERATIONS)$$
- נסו אורכי מסנן שונים סביב האורך שאותו צפיתם. מדוע יש צורך בתהליך איטרטיבי לצורך ביצוע פעולה זו?

- 4) השתמשו בפונקציה "rotate_im" על-מנת לסובב את התמונה בזווית $teta$ כך שהספרות יהיו מאונכות.
- 5) חזרו על הפעולה (סיבוב וסינון) עם אורך המסנן שנראה כי פועל "הכי טוב" עבור כל המסגרות של תנועת הרכב. האם האיכות זהה בכל המסגרות?
- 6) צרו מסכה של הספרות בלבד עבור כל מסגרת. השתמשו בכלים שנלמדו בחלק א'.
- 7) חלצו עבור הספרות בכל מסגרת את המאפיינים שהשתמשם בהם בשלב סיווג הספרות, וסווגו אותם אוטומטית.
- 8) בצעו סינון מסוג ML (Maximum Likelihood) על הסיווגים השונים של כל ספרה (יתכן והם יהיו זהים).
- 9) רשמו את התוצאה שקיבלתם. האם קיבלתם את המספר הנכון (נסו לשפר את המסכה של הספרות אם לא)?
- 10) הסבירו את הקשיים שנתקלתם בהם בסעיפים הקודמים, מה מבין הפעולות שקלות לנו כאנשים, קשה עבור מערכת ממוחשבת? אלו מבין השלבים הנ"ל יהיו זהים לכל סרט המכיל מכונית בתנועה, ואילו יצטרכו להיות מותאמים לכל סרט ספציפי?