

דו"ח סיכום פרויקט: א'

עקיבת עין ממצלמה פשוטה לצורך  
שליטה בתוכנת קריאה לילדים  
Webcam Eye Tracking for  
Controlling a Reading Software for  
Children

מבצעים:

טום-אבי שפירא Tom-Avi Shapira  
ליטל יעקובוב Lital Yakobov

מנחה:

אורי בריט Ori Bryt  
סמסטר רישום: אביב תשע"ט  
תאריך הגשה: מרץ, 2020

בשיתוף עם:

הפקולטה לחינוך למדע וטכנולוגיה,  
הפקולטה להנדסה ביו-רפואית.

P 4773-2-19

## תודות

אנו מודים לפרופ' ציפי הורוביץ-קראוס, מהפקולטה לחינוך, מדע וטכנולוגיה והפקולטה להנדסה ביו-רפואית שהגתה את הרעיון לפרויקט, ולאשת הקשר שלנו ד"ר מיכל זיוון, מהפקולטה לחינוך, מדע וטכנולוגיה שסייעו לנו במסירות במהלך הפרויקט.

אנו מודים מקרב לב לאורי בריט, המנחה שלנו, על התמיכה הרבה והעזרה ללא רבב לכל אורכו של הפרויקט. כמו כן, אנו מודים ליאיר משה, לנמרוד פלג ולשאר חברי מעבדת SIPL על הסיוע הרב.

# תוכן עניינים

|       |   |    |
|-------|---|----|
| 1     | מבוא  | 1  |
| 1.1   | מוטיבציה                                    | 1  |
| 1.2   | הנחות הפרויקט                               | 2  |
| 2     | סקירת ספרות                                 | 3  |
| 2.1   | VIOLA & JONS                                | 3  |
| 2.2   | HEAD ORIENTATION AND GAZE DETECTION         | 6  |
| 2.3   | GAUSSIAN PYRAMID                            | 7  |
| 2.4   | CIRCULAR HOUGH TRANSFORM                    | 8  |
| 3     | תיאור האלגוריתם                             | 11 |
| 3.1   | זיהוי פנים, עיניים ואף                      | 12 |
| 3.2   | מיקוד המסגרת התוחמת את העיניים              | 13 |
| 3.3   | זיהוי האישונים                              | 15 |
| 3.4   | ביצוע עקיבה אחר תנועת האישונים              | 17 |
| 3.4.1 | קביעת נקודת ייחוס                           | 17 |
| 3.4.2 | כיול המערכת                                 | 20 |
| 3.4.3 | ניתוח וקביעת כיוון ההסתכלות                 | 21 |
| 3.5   | ניתוח וידאו רציף                            | 22 |
| 3.5.1 | סנכרון וידאו                                | 22 |
| 3.5.2 | ניתוח מסגרות לצורך שלב כיול המערכת          | 23 |
| 3.5.3 | ניתוח מסגרות לצורך שלב קביעת כיוון ההסתכלות | 25 |
| 3.5.4 | תיקון קביעת כיוון ההסתכלות                  | 28 |
| 4     | סימולציות                                   | 31 |
| 4.1   | מאגר המידע                                  | 31 |
| 4.2   | תיאור הסימולציות                            | 34 |
| 4.3   | סיבוכיות זמן                                | 36 |
| 5     | תוצאות                                      | 37 |
| 5.1   | ניתוח התוצאות                               | 37 |
| 5.2   | תוצאות סרטוני העקיבה הרציפים                | 40 |
| 5.3   | תוצאות סרטוני העקיבה האקראיים               | 46 |

|                |                        |
|----------------|------------------------|
| 50.....        | 5.4. תוצאות כלליות     |
| <b>51.....</b> | <b>6. סיכום</b>        |
| 51.....        | 6.1. מסקנות            |
| 52.....        | 6.2. עבודה עתידית      |
| <b>53.....</b> | <b>7. רשימת מקורות</b> |

# רשימת איורים

- 1 ..... איור 1 - מימין רזולוציה נמוכה. משמאל רזולוציה גבוהה.
- 2 ..... איור 2 - דוגמה לתמונת משתמש ממאגר המידע שאספנו.
- 3 ..... איור 3 - זיהוי פנים לפי Viola & Jones.
- 4 ..... איור 4 - Haar Features.
- 4 ..... איור 5 - שימוש בתמונה אינטגרלית לחישוב סכום במלבן ABCD.
- 5 ..... איור 6 - Haar Features עבור אזור העיניים.
- 5 ..... איור 7 - Haar Features עבור אזור האף.
- 6 ..... איור 8 - מודל גיאומטרי לפנים.
- 7 ..... איור 9 - זיהוי המשולש: מרכזי העיניים ותחתית האף יחד עם קו המבט.
- 7 ..... איור 10 - פירמידה גאוסיאנית.
- 8 ..... איור 11 - מישור  $(m,b)$ .
- 8 ..... איור 12 - מישור  $(x,y)$ .
- 9 ..... איור 13 - תהליך ההצבעה עבור התמרת Hough.
- 9 ..... איור 14 - התמונה המקורית עבור.
- 9 ..... איור 15 - תמונת ההתמרה עבור.
- 11 ..... איור 16 - תרשים מלבנים של הפרויקט.
- 12 ..... איור 17 - מציאת פנים עיניים ואף לפי Viola & Jones.
- 13 ..... איור 18 - פירמידה גאוסיאנית של מסגרות העיניים.
- 14 ..... איור 19 - בחירת התבנית המתאימה ברמה 2.
- 14 ..... איור 20 - בחירת התבנית המתאימה ברמה 1.
- 14 ..... איור 21 - מסגרת העין לפני (משמאל) ואחרי (מימין) תהליך ה-TM.
- 15 ..... איור 22 - התאמה יחידה בתמונה המקורית לאחר CHT.
- 16 ..... איור 23 - יותר מהתאמה אחת ב-CHT לאחר עיבוד.
- 16 ..... איור 24 - בחירת המעגל המתאים להיות אישון לפי הפרמטרים.
- 18 ..... איור 25 - מסגרות פנים עם הסטה שונה המתקבלות מהפעלת אלגוריתם V&J.
- 18 ..... איור 26 - מסגרות עיניים שונות עם הסטה שונה.
- 19 ..... איור 27 - קביעת האף כנקודת היחוס.
- 20 ..... איור 28 - מבט לקצה שמאל של המסך.
- 20 ..... איור 29 - מבט לקצה ימין של המסך.
- 20 ..... איור 30 - גבולות הגזרה של האישון בחלל העין.
- 21 ..... איור 31 - גבולות הגזרה של האישון בחלל העין ומיקום האישון הנוכחי ביחס אליהן.

|    |   |
|----|---|
| 21 | איור 32 - חלוקת המסך ל-3 מקטעים שווים.....  |
| 23 | איור 33 - זיהוי הצפצוף בתחילת סרטון המשתמש.....   |
| 27 | איור 34 - וקטורי הפלט של האלגוריתם המייצגים את כיוון ההסתכלות הנקבע לכל עין בנפרד.....                          |
| 27 | איור 35 - הסכמה בין שתי העיניים.....  |
| 27 | איור 36 - חוסר הסכמה בין שתי העיניים.....   |
| 28 | איור 37 - תיקון כיוון ההסתכלות במקרה בו הזיהוי באחת העיניים נמצא מחוץ לגבולות הגזרה.....                        |
| 29 | איור 38 - תיקון כיוון ההסתכלות במקרה בו הזיהוי באחת העיניים לא התבצע.....                                       |
| 29 | איור 39 - תיקון כיוון ההסתכלות במקרה בו הזיהוי בשתי העיניים נמצא מחוץ לגבולות הגזרה.....                        |
| 30 | איור 40 - תיקון כיוון ההסתכלות במקרה בו הזיהוי בשתי העיניים לא התבצע.....                                       |
| 31 | איור 41 - מבט לקצה ימין של המסך.....  |
| 31 | איור 42 - מבט לקצה שמאל של המסך.....  |
| 32 | איור 43 - חלוקת המסך ל-5 מקטעים.....  |
| 34 | איור 44 - קביעת כיוון הסתכלות (מספר מקטע) שונה כתוצאה מהזזת החוצץ על פני הסרטון.....                            |
| 35 | איור 45 - קביעת כיוון הסתכלות (מספר מקטע) כתוצאה מקטימת קצוות החוצץ.....  |
| 38 | איור 46 - פלט הכלי הוויזואלי לניתוח איכותי של האלגוריתם.....  |
| 41 | איור 47 - הדגמה לסרטון עקיבה רציף: המילה הראשונה.....   |
| 41 | איור 48 - הדגמה לסרטון עקיבה רציף: המילה השנייה.....  |
| 41 | איור 49 - הדגמה לסרטון עקיבה רציף: המילה השלישית.....   |
| 44 | איור 50 - הדגמה לתופעת הריצוד בה המשתמש אינו מסתכל אל המילה הרצויה והאלגוריתם קובע נכונה את כיוון ההסתכלות..... |
| 45 | איור 51 - הדגמה לתזוזת ראש הפוגעת בשלב הכיול.....   |
| 47 | איור 52 - הדגמה לסרטון עקיבה אקראי: המילה הראשונה.....  |
| 47 | איור 53 - הדגמה לסרטון עקיבה אקראי: המילה השנייה.....   |
| 47 | איור 54 - הדגמה לסרטון עקיבה אקראי: המילה השלישית.....  |

## רשימת טבלאות

|    |  |
|----|--|
| 33 | טבלה 1 - התפלגות מספר המילים בסרטוני העקיבה..... |
| 40 | טבלה 2 - תוצאות סרטוני העקיבה הרציפים.....       |
| 46 | טבלה 3 - תוצאות סרטוני העקיבה האקראיים.....      |

# תקציר

קריאה היא יכולת אנושית חשובה, החל מהגילאים הצעירים. כדי לבצע משימה זו, על המוח לסנכרן ולשלב כמה מערכות קוגניטיביות, כגון: שליטה, שפה ועיבוד התמונות המתקבלות מהעיניים. למעשה, העיניים הן המפתח לביצוע מטרה זו. עבור אנשים הלוקים בדיסלקציה משימת הקריאה היא מורכבת יותר מכיוון שענייהם מרצדות בין חלקי הטקסט.

מטרת פרויקט זה היא זיהוי עיניי הקוראים תוך עקיבה וניתוח תנועת האישונים על ידי שימוש במצלמת רשת. במסגרת הפרויקט, פיתחנו אלגוריתם זה בסביבת MATLAB. ביצועי האלגוריתם נבדקו על פני מאגר נתונים רחב הכולל תמונות וסרטונים של נבדקים בגילאים שונים שצולמו בתנאים מגוונים. נמצא כי האלגוריתם מצליח לזהות ולעקוב אחר עיניי הקוראים בצורה טובה.

## Abstract

Reading is a relatively important human ability. In order to perform this task, our brain must engage and synchronize several cognitive systems such as cognitive control, language, and visual recessing. The eyes can be used in this case as the “key” to the level of difficulty in synchronization between these cognitive systems. This can be extremely helpful for dyslexic children struggling with their reading.

In this project, we aim to develop reading program for children using their eye movement patterns. Our goal is to identify the readers' eyes while tracking and analyzing their eye movement. In this project we developed an algorithm in MATLAB workspace. The algorithm's performance was tested using a dataset with images and videos of different readers in wide range of ages and photo conditions using webcam. The algorithm exhibited good identification and tracking analysis abilities.

# 1. מבוא

## 1.1. מוטיבציה

קריאה היא משימה מורכבת, המוח שלנו צריך לבצע במקביל כמה משימות קוגניטיביות, ביניהן: זיהוי המילים בעזרת העיניים, ניתוח מידע חזותי במוח וחיבור המילים לכדי משפט לוגי תקין. הקריאה מבוססת על תהליך סריקה חזותי המורכב מנעיצות המבט ותנועות הקפצה קצרות של העיניים להמשך שורה.

עבור אנשים המאובחנים בדיסלקציה, אשר בעלי יכולות קוגניטיביות נמוכות, קיים קושי במיקוד המבט בנקודה הרצויה ובסריקה ראייתית וכן, בלבול בין האותיות.

מטרת הפרויקט היא ראשית לזהות את עיני המשתמשים, תוך התמקדות באישונים. לאחר מכן, לעקוב אחר תנועתם וזאת כדי לנתח את תזוזת העיניים ואת כיוון ההסתכלות על הטקסט המוצג על המסך. כך נוכל להבין וללמוד בצורה איכותית על תהליך הקריאה של המשתמש, ובפרט, של משתמשים המתקשים בקריאה.

האתגר המרכזי בפרויקט זה, בניגוד לפתרונות הקיימים בספרות המשתמשים במצלמות איכותיות, הוא השימוש במצלמת רשת וזאת במטרה להנגיש את אלגוריתם העקיבה לכל משתמש.



איור 1 - מימין רזולוציה נמוכה. משמאל רזולוציה גבוהה



## 1.2. הנחות הפרויקט

במסגרת הפרויקט פיתחנו אלגוריתם המזהה עיניים ואישונים ומבצע עקיבה אחרי תנועתם תוך שימוש במצלמת רשת. פיתחנו כלי ליצירת מאגר מידע הכולל תמונות וסרטונים, ובדקנו עליהם את האלגוריתם שיצרנו.



איור 2 - דוגמה לתמונת משתמש ממאגר המידע שאספנו

במימוש האלגוריתם הנחנו מספר הנחות :

- הקריאה היא בציר האופקי בלבד.
- בעת הצילום המשתמש נמצא במרחק סביר מהמצלמה, כך שראשו ישר ודומיננטי בתמונה.
- בעת הצילום המשתמש אינו מזיז את ראשו כלל, כך שבמהלך הקריאה רק אישוניו זזים וראשו אינו מוטת.
- החדר מואר, בתהליך העקיבה בשלב הוידאו דרשנו בנוסף תנאי תאורה אחידים.

## 2. סקירת ספרות

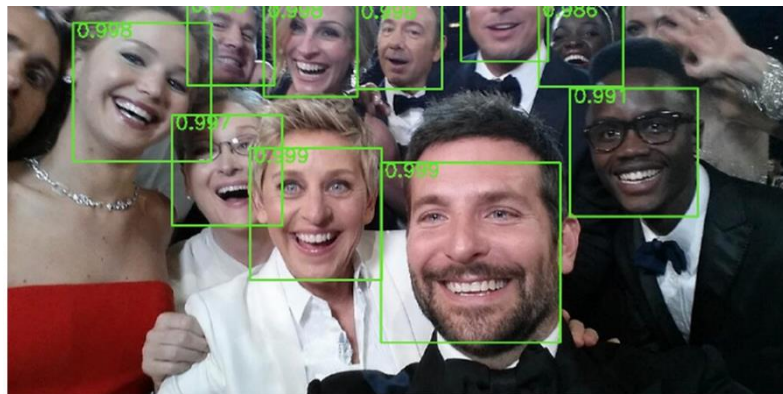
במהלך הפרויקט בוצעה סקירת ספרות שמטרתה הכרות עם אלגוריתמי עיבוד תמונה שונים ובחינה של אלגוריתמים קיימים לזיהוי העיניים ועקיבה אחריהן. נסקור את התרומות העיקריות מהספרות הרלוונטיות בפרויקט.

### 2.1 Viola & Jones

זהו אלגוריתם הזיהוי הראשון של אלמנטים בreal time אשר פותח על ידי Paul Viola וMichael Jones ב2001. המוטיבציה הראשונית בפיתוח האלגוריתם הייתה זיהוי פנים. ניתן להתאים אלגוריתם זה לזיהוי אובייקטים ממחלקות שונות. בפרויקט אנו משתמשים באלגוריתם לזיהוי פנים, עיניים ואף, בהתבסס על מאמר [1]. הנחות העבודה בעת השימוש באלגוריתם הן שכל הפנים מוכלות בתמונה, הן אינן מוטות לאף צד והמבט הוא ישר.

יתרונות האלגוריתם הם :

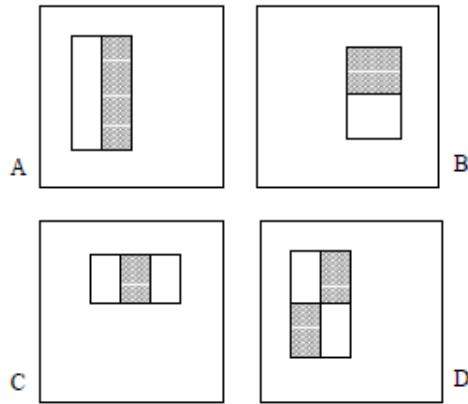
- Robust – אחוז זיהוי גבוה וקצב מאוד נמוך של false-positive.
- Real time – קצב מינימלי של 2[fps].



איור 3 - זיהוי פנים לפי Viola & Jones

לאלגוריתם יש ארבעה שלבים עיקריים :

- שימוש ב Haar features .
- יצירת תמונה אינטגרלית.
- Adaboost training .
- Cascading Classifiers .

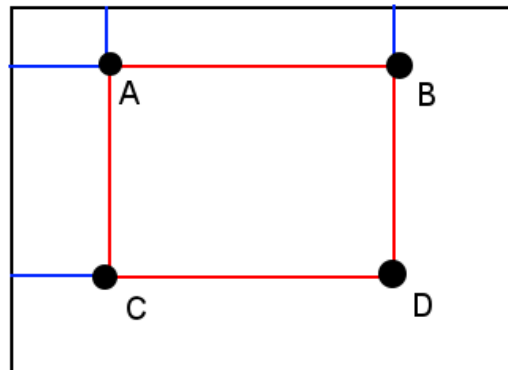


איור 4 - Haar Features

Haar Features משמשים לזיהוי אובייקטים בתמונה. הרעיון המרכזי הוא ביצוע סכום של הפיקסלים בתמונה בתוך אזור מלבני מוגדר.

למעשה זיהוי פנים היא משימה מעט יותר מורכבת ונדרש שימוש במלבנים (לדוגמא הארבעה המתוארים באיור משמאל) באלגוריתם של Viola & Jones. הערך המתקבל עבור כל מלבן כזה הוא ההפרש שבין סכום הפיקסלים בתמונה שבתוך המלבן הכהה לבין סכום הפיקסלים בתמונה שבתוך המלבן הבהיר.

על ידי שימוש במבנה נתונים של תמונה אינטגרלית, הסכימה באזור מלבני מתבצעת בזמן שהוא קבוע. זה כמובן מקנה יתרון מהירות גדול על פני אלטרנטיבות אחרות. בעצם כל נקודה בתמונה האינטגרלית מייצגת סכום של כל הפיקסלים מעל ומשמאל לפיקסל הנ"ל בתמונה המקורית. לאחר יצירת התמונה האינטגרלית, חישוב הסכום של הפיקסלים בתמונה בתוך כל אזור מלבני הוא פשוט, ומצריך שימוש ב-4 נקודות בלבד בתמונה האינטגרלית.



איור 5 - שימוש בתמונה אינטגרלית לחישוב סכום במלבן ABCD

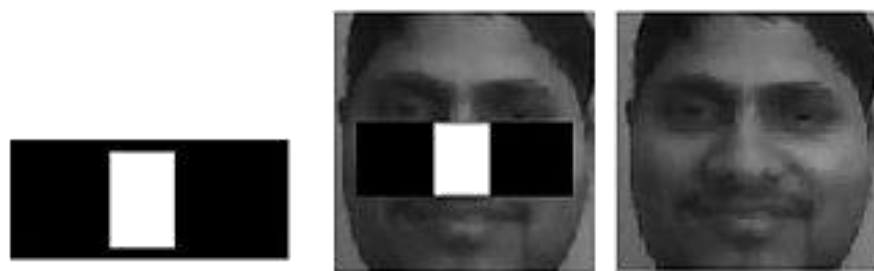
$$\sum_{\substack{x_0 < x \leq x_1 \\ y_0 < y \leq y_1}} i(x, y) = I(D) + I(A) - I(B) - I(C)$$

הרעיון המרכזי הוא שהפנים האנושיות כולן חולקות כמה מאפיינים משותפים, ונוכל לבצע התאמה שלהם על ידי שימוש ב-Haar Features. למשל:  
 -אזור העיניים כהה יותר מאשר האזור העליון של הלחיים.



איור 6 - Haar Features עבור אזור העיניים

-אזור הגשר של האף הוא בהיר יותר לעומת העיניים.



איור 7 - Haar Features עבור אזור האף

הרכב המאפיינים היוצרים תווי פנים מתאימים:  
 -מיקום וגודל: עיניים, שפתיים, גשר האף.  
 -הערך המתקבל עבור כל מלבן: מייצג את ההפרש.

בשלב הבא נעשה שימוש בגרסה של האלגוריתם הלומד Adaboost כדי למצוא את המאפיינים הטובים ביותר וכן לאמן את המסווגים להשתמש בהם. בכך נוצר מסווג "חזק" המורכב בצורה לינארית ממסווגים "חלשים".

$$h(\mathbf{x}) = \text{sgn} \left( \sum_{j=1}^M \alpha_j h_j(\mathbf{x}) \right)$$

כל מסווג "חלש" הוא פונקציית סף פשוטה, המבוססת על feature מסוים. למשל עבור feature  $j$ :

$$h_j(\mathbf{x}) = \begin{cases} -s_j & \text{if } f_j < \theta_j \\ s_j & \text{otherwise} \end{cases}$$

ערך הסף, הקוטביות והקבועים נקבעים במהלך האימון.

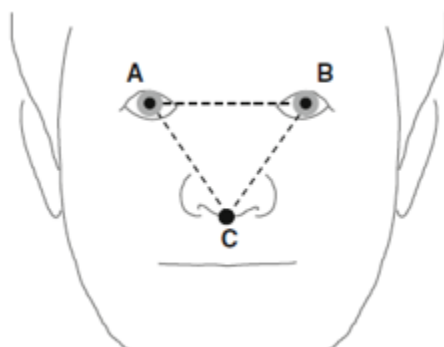
בשלב האחרון נעשה שרשור של מספר מסווגים, כך שכל המידע בפלט של מסווג אחד ניתן כקלט נוסף למסווג הבא בשרשור. למעשה, כל מסווג בודק את התוצאה עבור חלק מהמאפיינים, בסדר הולך ועולה בכמות ובמקביל והולך ויורד בחשיבות. כך שהמסווג הבא הוא מדויק יותר. הדרישה המרכזית הייתה שהמסווג יהיה מהיר מספיק כדי להיות מיושם במערכת עם מעבד שעובד בהספק נמוך כמו פלאפונים ומצלמות. למעשה, החיפוש של האובייקט במסגרת כולה נעשה על ידי הזזת החלון ברחבי התמונה ובדיקה של כל מיקום אפשרי של המסווג.

## 2.2 Head Orientation And Gaze Detection

זהו אלגוריתם זיהוי כיוון הסתכלות אשר פותח על ידי Jeremy Yirmeyahu Kaminski ועמיתיו בשנת 2006, לפי מאמר [2].

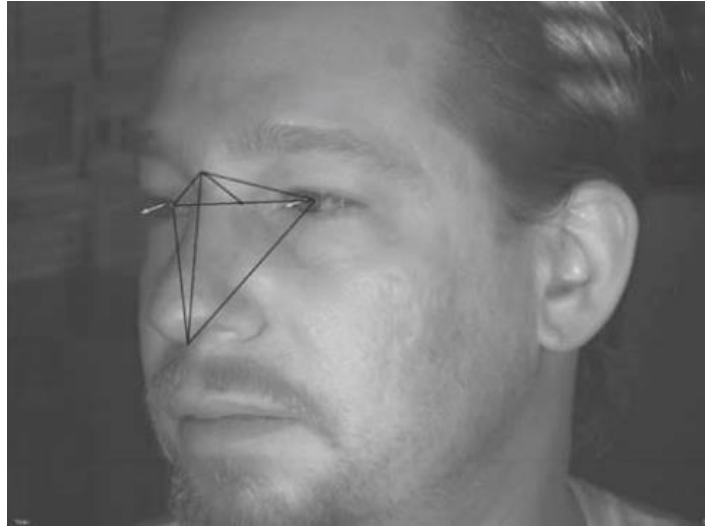
הקלט שלו הוא תמונה יחידה של פנים המתבוננות לכיוון כלשהו, והפלט שלו הוא מנח הראש וכיוון המבט. יתרונות האלגוריתם הם שהוא רובסטי ולא דורש קליברציה מוקדמת.

האלגוריתם מבוסס על מודל גיאומטרי של הפנים והעיניים. נשים לב כי שתי העיניים ותחתית האף יוצרים יחד משולש תלת-ממדי.



איור 8 - מודל גיאומטרי לפנים

האלגוריתם מזהה את כיוון הסתכלות באופן נפרד עבור כל עין, כשלרוב מתקבלות תוצאות כמעט זהות עבור שתי העיניים. בכך ניתן לראות את רובסטיות האלגוריתם.

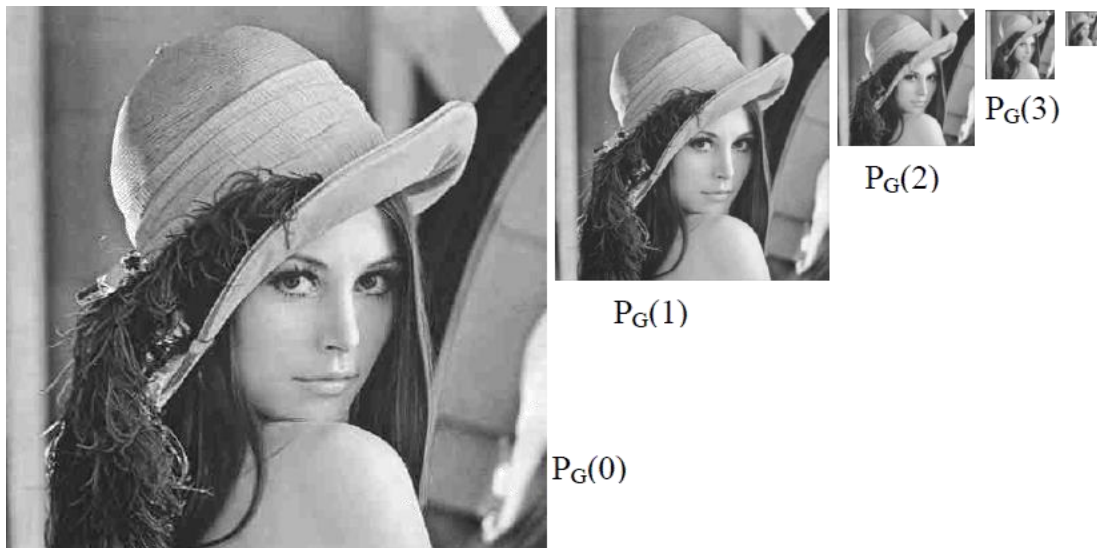


איור 9 - זיהוי המשולש : מרכזי העיניים ותחתית האף יחד עם קו המבט

נציין כי לא השתמשנו באלגוריתם זה, מפני שעסקנו בפתרון בעיית העקיבה בדו-ממד.

## 2.3. Gaussian Pyramid

פירמידה גאוסיאנית היא אוסף של תמונות הנובעות מתמונה אחת מקורית (שנמצאת בתחתית הפירמידה) שעברה החלקה (העברה במסגן מעביר נמוכים) ותת-דגימה בפקטור 2 בכל ממד (הקטנת ממד התמונה). כך שלמעשה כל תמונה בפירמידה נוצרת מטשטוש ודילול של קודמתה.



איור 10 - פירמידה גאוסיאנית

כך קיבלנו ייצוג של התמונה המקורית במגוון רזולוציות, כשהשכבות העליונות מתמקדות בפרטיים העיקריים בתמונה וזורקות את הפרטים הקטנים.

במקור [3] נטען כי המבנה הפירמידלי עוזר בין היתר בחיפוש תבניות בתמונה ומקל משמעותית על התהליך. הרעיון הוא לבצע חיפוש מקדים בשכבות הגבוהות יותר של הפירמידה ובכך לגלות אזורים חשודים בהם ייעשה חיפוש גם ברזולוציות הגבוהות שהן השכבות הנמוכות יותר של הפירמידה. אנו עושים שימוש בכך בפרויקט במציאת מסגרת תוחמת ממוקדת יותר של עיניי המשתמש. נציין בנוסף כי שימוש זה מקל גם במציאת ההתאמה, שכן פשוט יותר להתאים תבנית קטנה ובעלת פחות פרטים, וכן מצמצם את הזמן הנדרש עבור חיפוש התבנית.

## 2.4 Circular Hough Transform

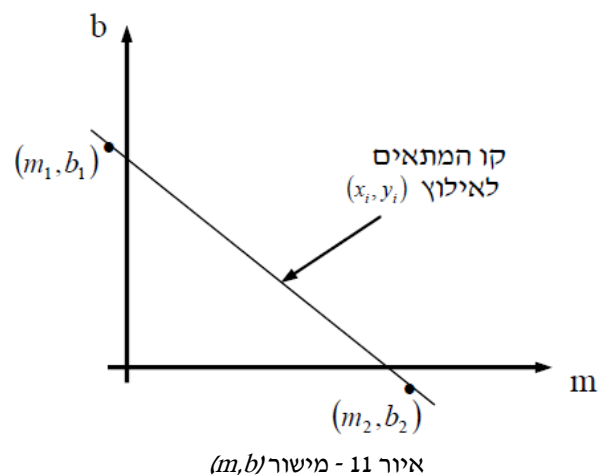
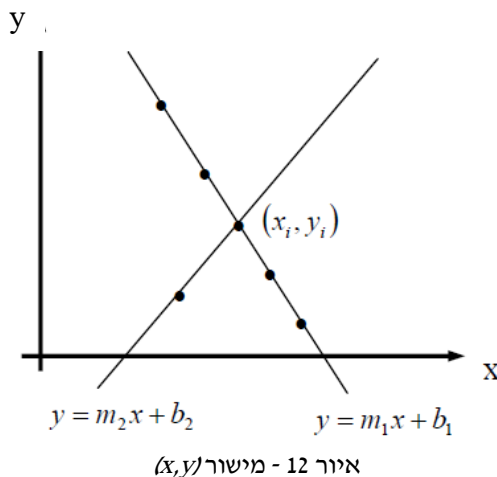
התמרת Hough היא כלי לזיהוי קווים ומעגלים, נרחיב על כך לפי מאמר [4]. באופן הגיוני ניתן להניח כי קווים ישרים בתמונה הם למעשה שפות, על כן האלגוריתם מסייע בעצם גם במציאת שפות בתמונה. עבור זיהוי קווים: ההתמרה מבצעת מעבר מהמישור הרגיל של  $(x,y)$  למישור הפרמטרים  $(m,b)$  המייצגים שיפוע וחיתוך עם הציר האנכי של ישר.

דרך כל נקודה ניתן להעביר אינסוף קווים שונים.

נתייחס לכל ישר כזה כנקודה במישור  $(m,b)$  בהתאם.

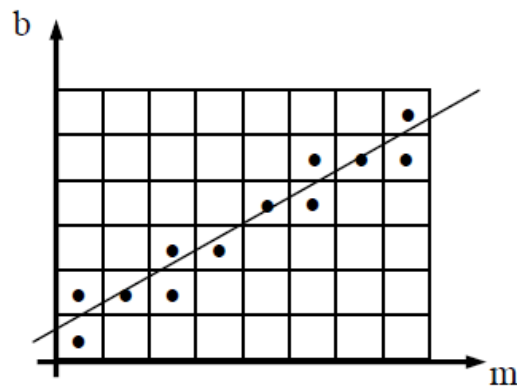
למעשה, כל הישרים העוברים דרך נקודה כלשהי במישור  $(x,y)$  מהווים ישר במישור  $(m,b)$  לפי :

$$b = -x_i m + y_i$$



ומכאן שעבור כל נקודה במישור  $(x,y)$  נקבל ישר במישור  $(m,b)$  המאפיין את כל הקווים שעוברים דרך הנקודה במישור  $(x,y)$ . נחזור על הפעולה עבור כל הנקודות בתמונה.

המשמעות של חיתוך הרבה קווים במישור  $(m, b)$  בנקודה מסוימת, היא שהפרמטרים  $(m, b)$  של נקודת החיתוך מתאימים לקו שעובר דרך הרבה נקודות במישור  $(x, y)$ . מספר הקווים הנחתכים מייצג את מספר הנקודות במישור  $(x, y)$  דרכן עובר הקו. כך הופכת בעיית מציאת הקווים הישרים לבעיית מציאת נקודות חיתוך. האלגוריתם ממומש בפועל על ידי בניית מטריצה בגודל של  $Nb \times Nm$ , וביצוע הצבעה - כל קו העובר דרך תא במטריצה מעלה את ערך התא ב1. לבסוף, מחפשים מקסימום על פני כלל התאים.



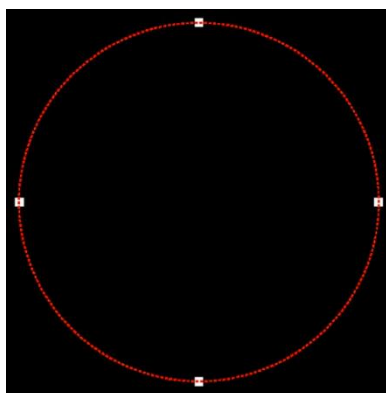
איור 13 - תהליך ההצבעה עבור התמרת Hough

עבור זיהוי מעגלים: ההתמרה מבצעת מעבר מהמישור הרגיל של  $(x, y)$  למרחב הפרמטרים  $(r, x_c, y_c)$  המייצגים את רדיוס המעגל ומרכזו.

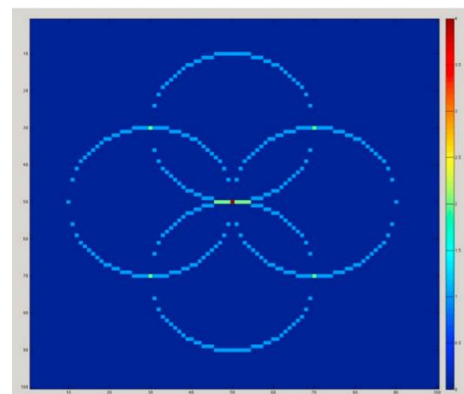
דרך כל נקודה ניתן להעביר אינסוף מעגלים שונים.

נתייחס לכל מעגל כזה כנקודה במרחב הפרמטרים  $(r, x_c, y_c)$ .

למעשה, כל המעגלים העוברים דרך נקודה כלשהי במישור  $(x, y)$  מהווים מעין קונוס במרחב הפרמטרים  $(r, x_c, y_c)$ . ניתן למצוא את המעגלים על ידי מציאת חיתוכים בין הקונוסים (תלת ממד) או על ידי פירוק לשני שלבים. בשלב הראשון, מקבעים רדיוס ומוצאים את מרכז המעגל האידיאלי (דו-ממד), ובשלב השני, מוצאים את אורך הרדיוס האידיאלי (חד-ממד). למשל עבור אורך רדיוס ידוע:



איור 14 – התמונה המקורית עבור Circular Hough Transform



איור 15 – תמונת ההתמרה עבור Circular Hough Transform

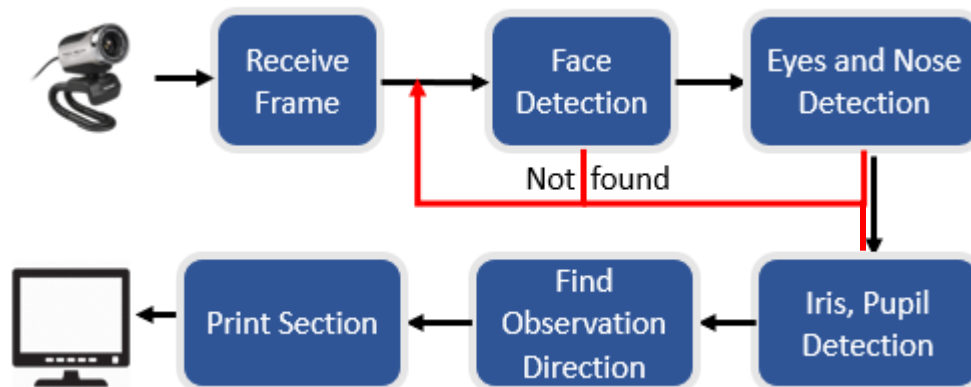


בהתמרת Hough המעגלית בתמונה מימין, ניתן לראות שנוצר מעגל עבור כל אחת מ-4 הנקודות בתמונה המקורית, שמרכזו הוא באותה הנקודה שמתאימה לו בתמונה המקורית. באופן דומה להתמרת Hough עבור קווים, בונים מטריצה ומבצעים תהליך הצבעה כך שערכי התאים במטריצה עולים ב-1 עם כל מעגל שעובר דרכם. לבסוף, מחפשים מקסימום על פני כלל התאים למציאת מרכז המעגל שבו עוברות כל הנקודות שבתמונה המקורית. בפרויקט אנו משתמשים בהתמרת Hough המעגלית בשלב מציאת הקשתית והאישון כמרכזה.

### 3. תיאור האלגוריתם

במהלך תיאור האלגוריתם נתייחס לניתוח עבור עין אחת. הניתוח עבור העין השנייה סימטרי ובלתי תלוי בראשונה. לשם פשטות ראשית נתייחס לאלגוריתם ושלביו עבור תמונות בודדות. לאחר מכן, תוך המשך תיאורו, נבצע את המעבר לניתוח וידאו רציף. לאלגוריתם מספר שלבים :

- זיהוי פנים, עיניים ואף.
- מיקוד המסגרת התוחמת את העיניים.
- זיהוי אישונים.
- ביצוע עקיבה אחר תנועת האישונים.



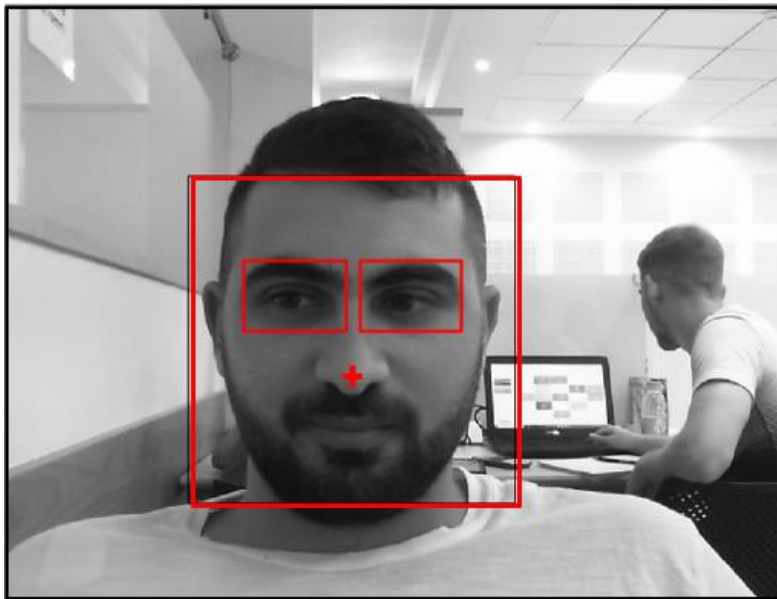
איור 16 - תרשים מלבנים של הפרויקט

### 3.1. זיהוי פנים, עיניים ואף

השתמשנו באלגוריתם של Viola & Jones על מנת לזהות את הפנים, העיניים והאף של המשתמש. טרם השימוש באלגוריתם, נרמלנו את ממוצע ערכי רמות האפור בתמונה להיות 147, בחרנו ערך זה בצורה אמפירית בהתבסס על מאגר התמונות שלנו, תמונות עם ממוצע רמות אפור זה הניבו תוצאות זיהוי אישון טובות יותר בשלבים מאוחרים יותר של האלגוריתם.

פלט האלגוריתם הוא מסגרות התחמות את האלמנטים המבוקשים, במקרה של האף סימנו את מרכז המסגרת מפני שזהו המידע הנחוץ לפעולת האלגוריתם. במידה ואחד או יותר ומהאלמנטים לא עבר זיהוי כנדרש, דילגנו על המסגרת ועברנו למסגרת הבאה. בחרנו להשתמש אך ורק במסגרות המכילות את שתי העיניים ולא לנתח עבור עין בודדת כדי להקל על מימוש האלגוריתם שמתבסס על ניתוח זהה עבור שתי העיניים שזוהו. נציין שבחירה זו נעשתה תחת ההנחה שאחוזי הצלחת אלגוריתם Viola & Jones מאוד גבוהים לפי הספרות, אנו נאשש זאת בנוסף בפרק התוצאות.

בחלק מהתמונות בוצע זיהוי לא יחיד של פנים, עיניים או אף. עבור תמונות אלו ביצענו סינון לבחירת האלמנט בצורה נכונה. זאת על ידי הסתמכות על המיקום הצפוי של האלמנטים בתוך התמונה על פי המבנה של הפנים.



איור 17 - מציאת פנים עיניים ואף לפי Viola & Jones

## 3.2. מיקוד המסגרת התוחמת את העיניים

לפני שלב מציאת האישון עצמו, מיקדנו את המסגרת התוחמת סביבו, זאת במטרה להקל על זיהוי ולסנן מראש זיהויים לא נכונים שיתכן והיו מתקבלים בחלקים של התמונה שנחתכו לאחר שלב זה. נציין כי מיקוד המסגרת סביב האישון התבצע רק בציר  $y$ , זאת מפני שבציר  $x$  קווי המתאר של העין ברזולוציה הנוכחית לא היו ברורים מספיק. לכן המיקוד לא היה איכותי דיו ולעיתים אף חתך חלקים חשובים בעין.

את תהליך המיקוד עשינו באמצעות Template Matching (TM). כדי לשפר את הביצועים השתמשנו בפירמידה גאוסיאנית של מסגרות העיניים שקיבלנו כפלט מ Viola & Jones בשלב הקודם. תחילה ביצענו TM על הרמות הגבוהות יותר בפירמידה, ולאחר מכן השתמשנו בתוצאה כדי למקד את החיפוש ברמות הנמוכות יותר.

השתמשנו בשלוש רמות של הפירמידה: תמונת הבסיס (רמה 0), 21 רמות נוספות.

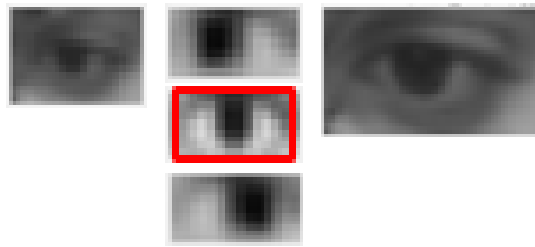


איור 18 - פירמידה גאוסיאנית של מסגרות העיניים

עבור שלב ה-TM יצרנו בסך הכל 24 תבניות (לכל עין). 8 תבניות לכל כיוון מבט: ימין, מרכז ושמאל. נציין כי התבניות נבנו על סמך עין ימין ושוקפו עבור עין שמאל. נדון על ההשפעה של כך בפרק התוצאות. מתוך 8 התבניות, תבנית אחת מיועדת לזיהוי הראשוני עבור רמה 2 (הגבוהה ביותר) והשאר מיועדות עבור רמה 1 והן בגדלים שונים, זאת במטרה למצוא התאמה מירבית ומצומצמת ככל הניתן סביב הקשתית. נציין כי יצרנו תבניות בשלושת כיווני המבט העיקריים כדי להקל על מציאת ההתאמה המירבית עבור עיני המשתמש המביטות בחלקים שונים במסך, כך שבסופו של דבר תבחר התבנית שהניבה את ההתאמה הטובה ביותר (בכל רמה של הפירמידה הגאוסיאנית).

כמתואר, להתאמה ראשונית ביצענו TM ברמה 2. זאת ההתאמה הגסה ביותר שנותנת אינדקס בסיסית לצמצום ראשון של מסגרת העין. שלב זה חשוב מפני שהוא מהווה בסיס להמשך ההתאמה ברמה 1, בלעדיו בהרבה מקרים לא נמצאה התאמה טובה ברמה 1. לאחר שמצאנו התאמה נחתך את התמונה בהתאם ונבצע הרחבה של התמונה לרמה 1 בפירמידה הגאוסיאנית.

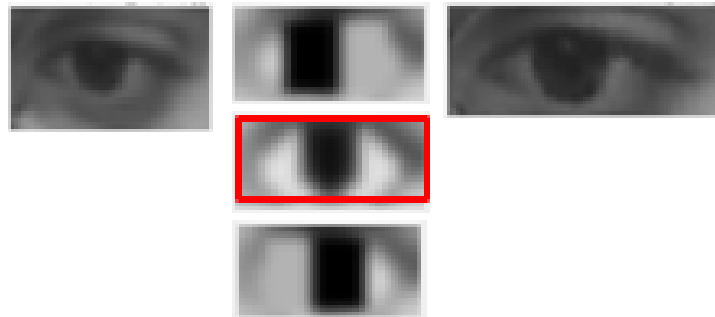
בפועל ההרחבה מטשטשת את התמונה ברמה 1 בצורה משמעותית ולכן במקום זאת בחרנו לחתוך את התמונה ברמה 1 לפי המסגרת המתאימה שנמצאה ברמה 2, מוגדלת בפקטור 2, זהו פקטור הגודל בין הרמות.



איור 19 - בחירת התבנית המתאימה ברמה 2

ניתן לראות באיור לעיל כי מבין 3 כיווני המבט ההתאמה המרבית לתמונה השמאלית באיור מתקבלת עבור כיוון המבט למרכז. התבנית המתאימה מסומנת במסגרת האדומה. לבסוף, לאחר ביצוע החיתוך המתואר לעיל קיבלנו את התמונה הימנית באיור זה. נציין כי בחרנו להציג 3 דוגמאות לתבניות המייצגות כיוון מבט כל אחת. שאר התבניות מייצגות אף הן את אותם כיווני המבט פרט לכך שהן בגדלים שונים (כדי לשפר את תהליך ה-TM). לכן לצורך ההמחשה הסתפקנו בהצגת 3 תבניות בגודל כלשהו.

באותו אופן, ביצענו TM עבור רמה 1 וחתכנו את התמונה ברמה 0 לפי המסגרת המתאימה שנמצאה מוגדלת בפקטור 2.



איור 20 - בחירת התבנית המתאימה ברמה 1

פלט האלגוריתם הוא מסגרת קטנה ומדוייקת יותר של העיניים. ניתן לראות זאת בצורה טובה באיור הבא:



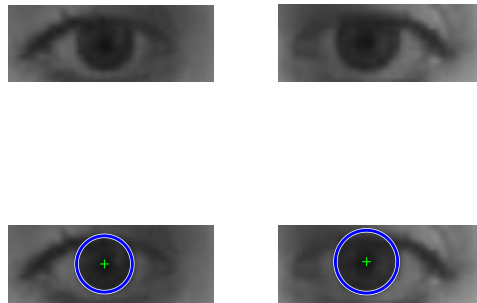
איור 21 - מסגרת העין לפני (משמאל) ואחרי (מימין) תהליך ה-TM

### 3.3. זיהוי האישונים

לאחר קבלת המסגרת המצומצמת מהשלב הקודם, מצאנו את האישון. ביצענו זאת על ידי מציאת הקשתית, ולאחר מכן את מרכזו שהוא האישון. תהליך מציאת הקשתית בוצע על ידי הפעלת  $CHT$  Circle Hough Transform על המסגרת המצומצמת. כפי שכבר ציינו, חשוב להפעיל את ההתמרה רק בשלב זה. זאת כיוון שהיא לא יציבה והפעלתה מסווגת הרבה מסיחים שהתקבלו במסגרת העין המקורית כמעגלים.

הוספנו פרמטרים שונים שמייחדים את הקשתית אשר עזרו לנו בזיהוי המעגל שאנו רוצים למצוא: טווח רדיוסים צפוי, ערך רמות אפור נמוך בעיגול המתקבל, טווח מיקומים צפוי של מרכז המעגל. כל אלו נקבעו על סמך מאגר התמונות שאספנו באופן אמפירי.

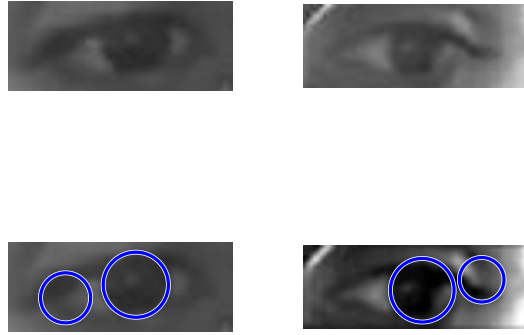
כאשר נמצאה התאמה יחידה שעונה על פרמטרי הקשתית שהגדרנו לעיל, חילצנו לפלט את מרכז המעגל כאישון.



איור 22 - התאמה יחידה בתמונה המקורית לאחר  $CHT$

כאשר לא נמצאה התאמה או שנמצאה התאמה יחידה שאינה מתאימה לנו לפי הפרמטרים שהוגדרו, מסגרת העין עוברת עיבוד הכולל מתיחת קונטרסט, חידוד  $LPF$ . זאת במטרה להבליט את הקשתית מיתר חלקי העין כך שאלגוריתם  $CHT$  יצליח לזהות אותה. תהליך זה נקבע על סמך העובדה שלרוב הקשתית היא חלק בולט וכה יותר ביחס לסביבתה.

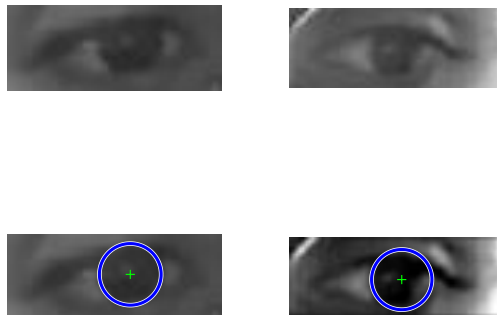
לאחר מכן, ביצענו מחדש את התמרת Hough וברוב המקרים אכן נמצא המעגל הרצוי. במידה וגם לאחר שלב עיבוד זה לא נמצאה הקשתית, ויתרנו על ניתוח מסגרת העין הנוכחית.



איור 23 - יותר מהתאמה אחת ב-CHT לאחר עיבוד

עבור האיור לעיל למשל, ניתן לראות תמונה של זוג עיניים לפני (הזוג העליון) ואחרי (הזוג התחתון) עיבוד, כך שמימין יש את עין ימין ומשמאל את עין שמאל. באיור זה לא נמצא תחילה מעגל מתאים עבור העין הימנית. על כן, העברנו אותה דרך אלגוריתמי עיבוד שונים (כמתואר). ניתן לראות כי מסגרת העין הימנית המקורית למעלה, שונה ממסגרת העין הימנית לאחר עיבוד שנמצאת למטה. כתוצאה מהעיבוד, נמצאו שתי התאמות פוטנציאליות. לעומת זאת, עבור העין השמאלית נמצאו מלכתחילה שתי התאמות פוטנציאליות.

כאשר נמצאה יותר מהתאמה אחת בחרנו את זו בעלת רמת האפור הממוצעת הנמוכה ביותר וזאת בתנאי שמיקום האישון שנמצא אכן בטווח שהוגדר לעיל.



איור 24 - בחירת המעגל המתאים להיות אישון לפי הפרמטרים

לדוגמא באיור 23, בחרנו במעגלים שערך רמות האפור הממוצע שלהם הוא נמוך יותר, כמו כן גם מרכז האישון ממוקם בטווח שהוגדר.

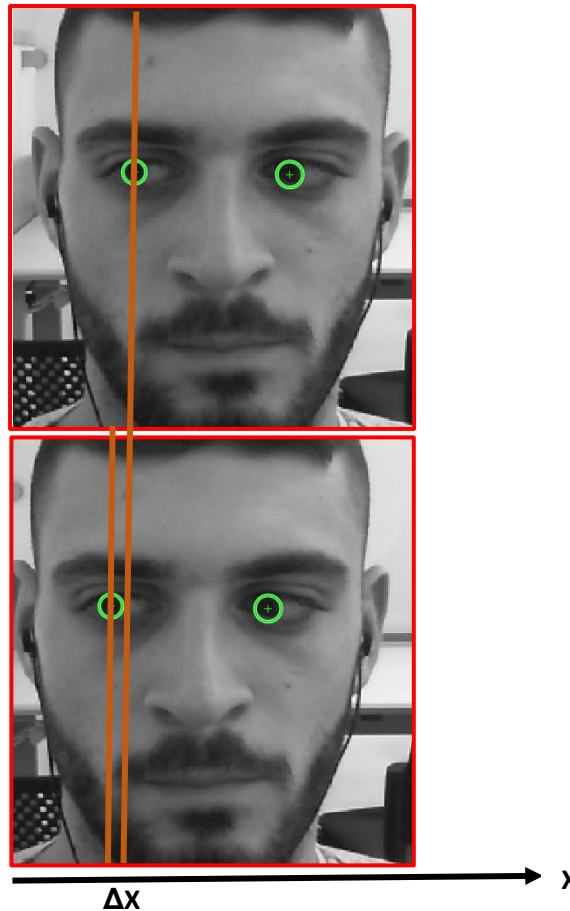
## **3.4. ביצוע עקיבה אחר תנועת האישונים**

### **3.4.1. קביעת נקודת ייחוס**

ראשית על מנת לבצע את העקיבה יש לקבוע נקודת ייחוס יציבה בתמונת הפנים שהתקבלה בתת הפרק [3.1](#). זאת כדי שנוכל להשוות מסגרות שונות. הצורך להשוות מסגרות שונות נובע תחילה מהצורך לקבוע את כיוון הסתכלות. לשם כך נצטרך להשוות את כיוון ההסתכלות הנוכחי המתקבל בכל פעם, לכיוון הסתכלות שקיבלנו מראש בעת כיול המערכת. למעשה יש צורך לקבוע גבולות גזרה לאישון כדי להבין מהו התחום שהוא יכול לנוע בו, ובהתאם למיקומו בתוך התחום, נוכל לקבוע את כיוון ההסתכלות הנוכחי. על נושא כיול המערכת נרחיב בתת הפרק הבא.

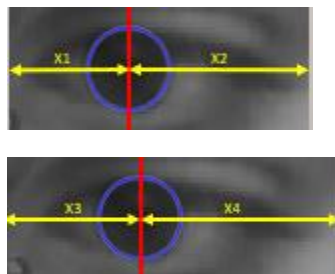
סיבה נוספת בגללה נרצה לקבוע נקודת ייחוס היא העובדה שאם נצלם את המשתמש בשתי תמונות עוקבות, כאשר הוא מסתכל לאותו הכיוון בדיוק ולא זו, נקבל כי פלט האלגוריתם יכול להיות שונה. הסיבה לכך היא שאלגוריתם Viola & Jones חותך את מסגרות תמונות הפנים והעיניים בצורה שונה כל פעם (כפי שמוצג באיורים 24, 25 בהמשך) ואם נמדוד את מיקום האישון לפי מיקומו בתוך המסגרת שהתקבלה למעשה נקבל את מרחקו מהפינה השמאלית העליונה שלה שמשתנה לפי אופן החיתוך הנעשה על-ידי Viola & Jones. הדרך להתגבר על כך היא הגדרת נקודת ייחוס יציבה ללא קשר למסגרות המתקבלות מ Viola & Jones, כך שנקבע את מיקום האישון ביחס אליה.





איור 25 - מסגרות פנים עם הסטה שונה המתקבלות מהפעלת אלגוריתם  $V&J$

ניתן לראות באיור 24 כי על אף שהמשתמש מתבונן לאותו הכיוון בשתי המסגרות העוקבות, מתקבל ערך  $x$  שונה עבור האישון. זאת עקב מסגרת פנים שונה שמניב האלגוריתם של Jones ו Viola. מכאן נובע הצורך בנקודת ייחוס, שביחס אליה יתבצעו המדידות.

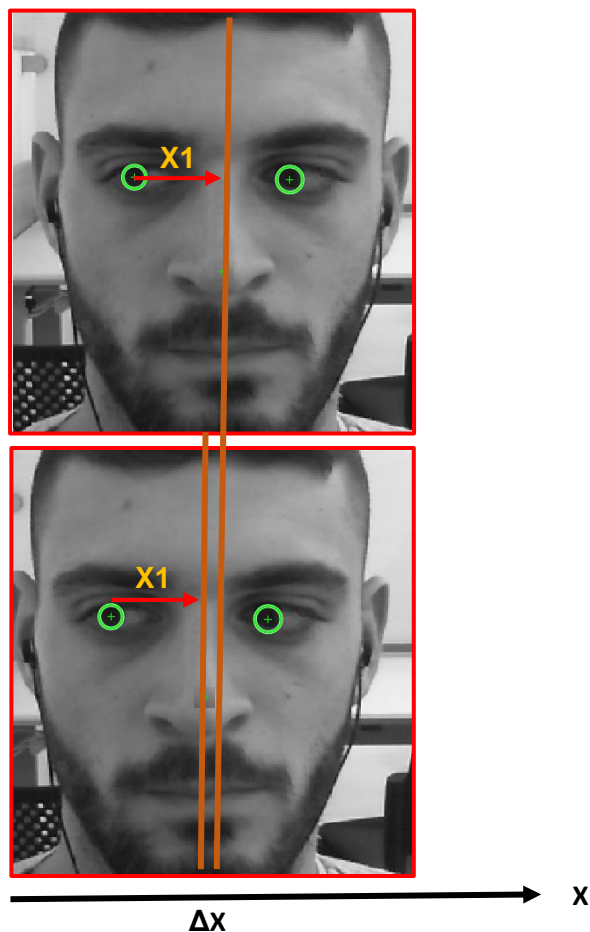


איור 26 - מסגרות עיניים שונות עם הסטה שונה

כמו כן, ניתן לראות באיור 25 שתי מסגרות עיניים שונות עבור כיוון הסתכלות זהה. כך שגם בנסיון זה לקבוע את כיוון ההסתכלות ביחס אליהן הבעיה שתוארה לעיל קיימת. לא נוכל לבצע מדידה מהימנה של ערך  $x$  של האישון מאף קצה, זאת כיוון שנקבל שוני משמעותי בין הערכים בשתי המסגרות.

כדי לפתור בעיה זו נשתמש בזיהוי האף באמצעות Viola & Jones ונקבע אותו להיות נקודת הייחוס, במקום הפינה השמאלית העליונה של המסגרת שהתקבלה. מבדיקות על מאגר התמונות שלנו מצאנו כי נקודה זו אכן יחסית יציבה.

נוכל להסביר את יציבות זיהוי האף בכך שהוא איבר סימטרי בפנים וכי אינו רגיש להבעות ולא זו בחלל הפנים. בשל כך, גם במקרה בו המסגרת התוחמת את האף המתקבלת מאלגוריתם Viola & Jones שונה במעט מפעם לפעם, מרכזה עדיין יחסית יציב וקבוע בתמונת הפנים הכוללת כפי שבדקנו באופן אמפירי.

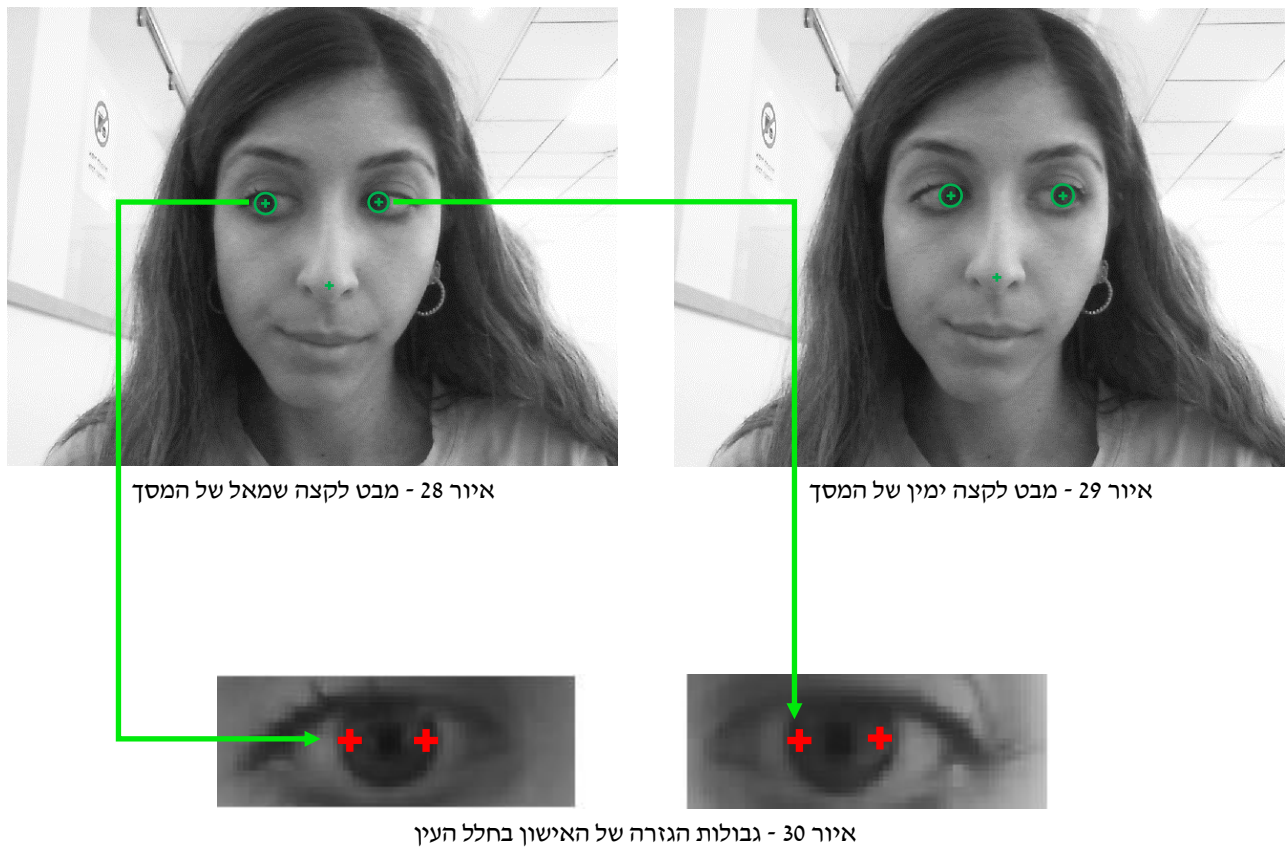


איור 27 - קביעת האף כנקודת הייחוס

### 3.4.2. כיול המערכת

כדי לקבוע את כיוון ההסתכלות של המשתמש, ניעזר בשתי תמונות קליברציה, אשר מגדירות את מיקום האישון בעת התבוננות המשתמש לשני קצוות המסך. את מיקום האישון בכל תמונה נמצא לפי האלגוריתם שתואר עד כה.

לאחר מציאת נקודת הייחוס בשלב הקודם, נשתמש בה כדי לקבוע את מיקום האישון בתמונות הקליברציה ביחס אליה, עבור כל משתמש.



כפי שניתן לראות בתמונות לעיל, שני המיקומים של האישונים (לכל עין) שמצאנו מהווים גבולות גזרה למיקום האישון בעת התבוננות המשתמש על הטקסט המוצג על המסך. למעשה מרגע הגדרת גבולות הגזרה של תנועת האישון נצפה שהיא תוגבל בציר ה-x בין שני ה '+' האדומים בכל עין ללא תלות בשניה.

### 3.4.3. ניתוח וקביעת כיוון ההסתכלות

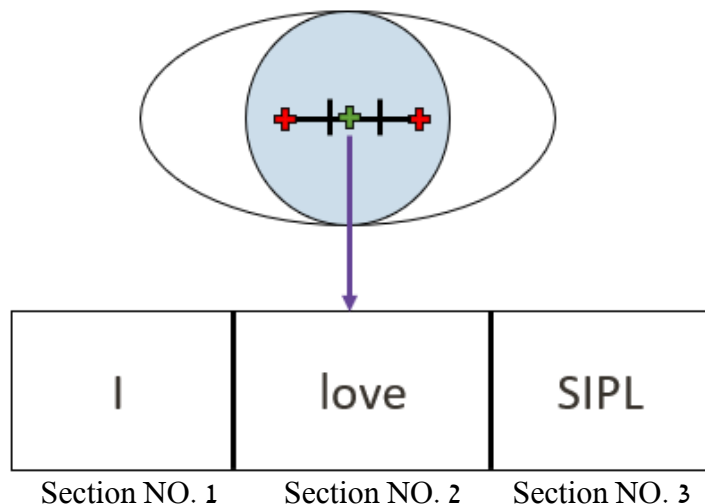
בהנתן תמונה בה המשתמש מסתכל למקום מסויים במסך, נמצא את מיקום האישון שלו בכל עין ללא תלות בשניה. גם במקרה זה, לאחר מציאת נקודת הייחוס, נשתמש בה כדי לקבוע את מיקום האישון ביחס אליה. כעת, לאחר מציאת הן המיקום הנוכחי והן מיקום האישון בתמונות הקליברציה נוכל לקבוע מיקום יחסי ביניהן. (זאת מפני שכל המיקומים הנ"ל הם ביחס לנקודת הייחוס היציבה שקבענו).



איור 31 - גבולות הגזרה של האישון בחלל העין ומיקום האישון הנוכחי ביחס אליהן

נשים לב שאכן כפי שציפינו, מיקום האישון שקובע את כיוון ההסתכלות המשתמש (אשר מסומן ב '+' הירוק) תחום בין גבולות הגזרה שהוגדרו בשלב הקליברציה (אשר מסומנים ב '+' האדומים).

כעת, לאחר קביעת המיקום היחסי נוכל להסיק את כיוון ההסתכלות המשתמש על המסך. כדי לקבוע את כיוון ההסתכלות נחלק בצורה לינארית למספר זהה של חלקים את המסך ואת התחום שבין שני מיקומי האישונים בתמונות הקליברציה. אנו בחרנו 10 חלקים סך הכל כאשר כל חלק נקרא מקטע. ההתאמה נעשית בצורה חד ערכית כך שמספר המקטע שמכיל את האישון בתמונת העין שווה למספר המקטע עליו מסתכל המשתמש במסך. נציין שהמרחק הממוצע בין מיקומי האישון המיוחסים לגבולות הגזרה לפי מאגר הסרטונים אותו ניתחנו הינו 10 פיקסלים, כלומר בחלוקת המסך ל-10 מקטעים, בממוצע לכל מקטע מותאם פיקסל בודד. נרחיב על כך בפרקי התוצאות והמסקנות.



איור 32 - חלוקת המסך ל-3 מקטעים שווים

באיור לעיל ניתן לראות הדגמה לחלוקת המסך ל-3 מקטעים. המקטע הראשון הוא בין הפלוס האדום השמאלי עד לקו השחור השמאלי, המקטע השני הוא בין שני הקווים השחורים והשלישי הוא בין הקו השחור הימני לפלוס האדום הימני. הקביעה לאן מסתכל המשתמש נעשית לפי התחום בו נמצא ה-"+" הירוק.

## 3.5. ניתוח וידאו רציף

כדי להקל על המעבר מניתוח תמונה בודדת לניתוח וידאו שלם נגדיר כעת frame buffer. זהו למעשה חוצץ המכיל מספר מסגרות כאשר בכל מסגרת זוהה אישון. נציין שוב כי התהליך סימטרי לשתי העיניים, כלומר קיים חוצץ לכל עין בנפרד, לכן לשם פשטות נתאר אותו רק עבור עין אחת. מעתה נשתמש במונח "חוצץ" בכדי לתאר frame buffer.

העיקרון עליו נתבסס בעת ניתוח הווידאו הוא הסקה כוללת על סמך מספר מסגרות עוקבות, זאת כדי לקבוע את כיוון הסתכלות המשתמש. כלומר קבלת החלטה על סמך חוצץ. זאת בניגוד לניתוח בתחילת פרק זה שהסתמך בכל פעם על תמונה בודדת.

בדרך זו נדייק את ניתוח האלגוריתם תוך ניצול המסגרות הרבות שמספק לנו סרטון וידאו. נציין כי כל מסגרת שמכיל החוצץ עוברת את כל תהליך העיבוד שהוצג [באיור 15](#) ולמעשה מפורט בתת סעיפים [3.4 – 3.1](#) בפרק זה.

כעת, יתרון המסגרות הרבות שמספק הווידאו טומן בחובו גם חיסרון עיקרי. חשוב להבין שלמרות שאחוזי הצלחת האלגוריתם בזיהוי אישון המשתמש היו גבוהים על מאגר התמונות שהצגנו, עדיין היה אחוז מועט של מסגרות בהן אלגוריתם Viloa & Jones כשל או שלא הצלחנו לזהות נכונה את האישון. נציין שהסיבות לכך נובעות מאי עמידה בהנחות הפרויקט שהגדרנו ורזולוצייה ירודה של המצלמה איתה עבדנו. מפני שלווידאו יש אופי שונה מאשר תמונות, בעת המעבר לווידאו ישנן מסגרות שעוברות עיבוד לא טוב והאישון לא מזוהה בהן נכונה. לכן הדבר מאלץ את סינון. נרחיב על כך בהמשך תת פרק זה.

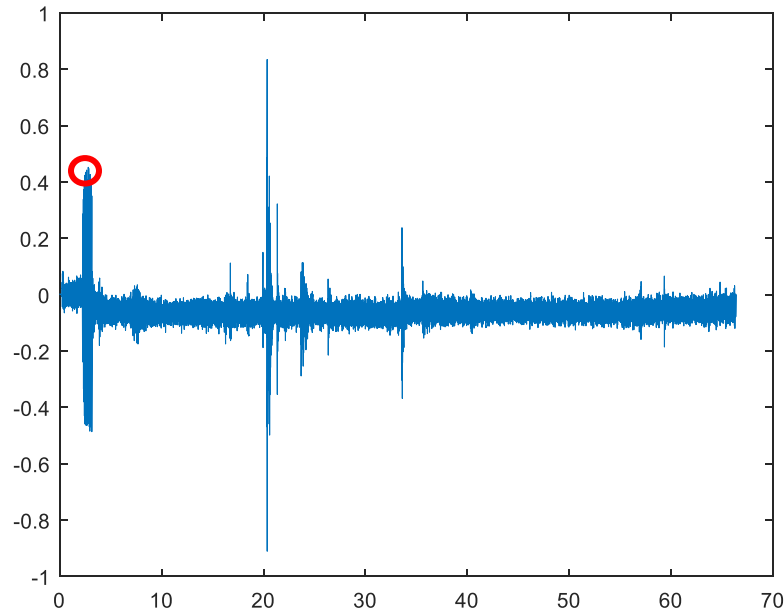
### 3.5.1. סנכרון וידאו

בעת המעבר מניתוח תמונות לניתוח סרטוני וידאו נתקלנו בבעיית סנכרון. למעשה, השניות הבודדות שלקחו לנו כעורכי הניסוי מהרגע שבו התחלנו לצלם את המשתמש ועד לרגע שבו הפעלנו את הסרטון המופיע על המסך עליו הסתכל דרשו מאיתנו לבצע סנכרון, כך שלא נכלול אותן בניתוח סרט הווידאו של המשתמש.

בסיס בעיית הסנכרון נבע מהצורך שלנו ככותבי האלגוריתם לנתח את נכונותו, כלומר לבדוק עבור כל חוצץ האם תוצאת כיוון ההסתכלות שקבע אכן נכונה. לשם כך נדרשנו להשוות תוצאת כל חוצץ בסרטון המשתמש לכיוון ההסתכלות ידוע שהוגדר מראש לפי הסרטון עליו הסתכל, כלומר לבצע סנכרון בין הסרטונים הנ"ל.

כך, במידה והייתה התאמה הכרזנו על "הצלחה", אחרת "כישלון" עבור החוצץ.

הפתרון שהצענו למטרה זו הוא שימוש בשמע. ברגע תחילת הסרטון עליו הסתכל המשתמש נשמע צפצוף. מעצם זיהוי הצפצוף בעת ניתוח סרטון המשתמש הסקנו מאיזו שניה בסרטון עלינו להתחיל את העקיבה אחרי עיניו.



איור 33 - זיהוי הצפצוף בתחילת סרטון המשתמש

ניתן לראות בגרף המוצג לעיל את האמפליטודה הגבוהה יחסית בתחילתו, הדבר מסמן את הישמע הצפצוף למטרת הסנכרון.

### 3.5.2. ניתוח מסגרות לצורך שלב כיוול המערכת

נכליל את שלב כיוול המערכת המתואר בתת פרק [3.4.2](#) עבור וידאו רציף. כאמור, בתחילת הסרטון עליו הסתכל המשתמש הוא נדרש להסתכל לשני קצוות המסך וכך דגמנו את מיקום אישורו והגדרנו את גבולות הגזרה של תנועתו. בפועל הדבר נעשה בוידאו רציף בו נדרש המשתמש להסתכל לכל כיוון למשך 3 שניות. האלגוריתם דוגם מתוך 3 השניות הללו את השניה המרכזית, זאת כדי לדגום את האישור במצב יציב כאשר הוא ממוקד. במקרה שלנו מפני שאנו דוגמים בקצב של 30 [Fps] למעשה נקבל חוצץ המכיל 30 מסגרות שונות שבהתבסס עליהן נוכל לקבוע את גבולות הגזרה של תנועת האישור. את הנתונים הטכניים ניתן לראות בפרק [4.1](#). כפי שכבר תיארו בהקדמת תת פרק זה, בחלק מהמסגרות הללו האישור לא מזוהה נכונה ולכן יש צורך בסינון.

לצורך כך נעביר את מסגרות הכיול בשני סוגי מסננים :

1. מסנן רמות אפור: מסנן זה מתבסס על העיקרון שהאישון הוא האובייקט הכי כהה בתמונת העין הסופית שקיבלנו מהאלגוריתם (לאחר תת פרק 3.2). למעשה, זהו מסנן סף שקובע על סמך ממוצע רמות האפור של האובייקט האם הוא יכול להיות ייצוג לאישון או לא, זאת בייחס לרמות האפור של שאר האובייקטים שזוהו כאישונים בשאר המסגרות שנכנסו לחוצץ. את הסף נקבע באופן הבא:

$$TH = \min + \frac{\max - \min}{3}$$

If gray\_level > TH  
remove

כלומר, לאחר שנחשב את רמות האפור הממוצעות של האישונים בכל 30 המסגרות שהתקבלו, נשמור רק את אלו הנמצאות בשליש התחתון של התחום בין הערך המקסימלי (max) למינימלי (min) שמצאנו. זאת בהתבסס על כך שנצפה שרמת האפור הממוצעות של האישון תהיה נמוכה. חשוב להבין שאין הדבר אומר שרק שליש מסך המסגרות, כלומר 10, יעברו את הסף. אלא שכל המסגרות הנמצאות בשליש התחתון של תחום רמות האפור שהוגדר יעברו אותו. נציין שבחרנו את השליש התחתון בצורה אמפירית על סמך ניתוח מאגר הסרטונים שלנו.

לאחר המעבר במסנן זה אכן רוב המסגרות שהכילו זיהוי אישון שגוי, כלומר מסיח, סוננו. המסיחים שנשארו היו לרוב עיגולים במקומות כהים יותר באזור העין כגון שפות חלל העין או אזור הגבה במידה ולא נחתך בשלבי האלגוריתם המוקדמים יותר. כדי להתגבר על כך הגדרנו מסנן נוסף המתבסס על מיקום.

2. מסנן מיקום: כפי שצויין, לאחר המעבר במסנן רמות האפור, כמות המסיחים קטנה משמעותית ולכן החלק היחסי שלהם מסך כל המסגרות שנותרו קטן מאוד. בהתבסס על כך מסנן המיקום שבנינו מסנן על סמך מיקום חריג שנקבע במספר מסגרות בודדות ביחס לכל שאר המסגרות. כך למעשה אנו מסננים מסגרות הקובעות מיקום אישון חריג הן בציר X והן בציר Y ביחס לכל שאר המסגרות.

מסנן זה בנוי על עיקרון של מסנן הפרש ומסנן ישר הפרש גדול מדי בין קביעת מיקום האישון של שתי מסגרות עוקבות (כלומר נגזרת גבוהה ולא צפויה). על סמך מאגר הסרטונים שלנו וברזולוציה הנתונה מצאנו בצורה אמפירית את ערך הסף לשינוי בנגזרת להיות 5 פיקסלים.

לאחר המעבר בשני סוגי המסננים שהוגדרו לעיל ובכדי לקבל החלטה אודות מיקום האישון הסופי לצורך קביעת גבולות הגזרה של תנועתו, ביצענו הן מיצוע והן פעולת חציון על המסגרות שנותרו בחוצץ. שתי הפעולות בוצעו במקביל ובאופן בלתי תלוי אחת בשנייה, זאת במטרה לבדוק מי מהן עדיפה תוך נסיון למקסם את ביצועי האלגוריתם. נרחיב על כך בפרקי הסימולציות והתוצאות.

### 3.5.3. ניתוח מסגרות לצורך שלב קביעת כיוון ההסתכלות

נכליל את שלב קביעת כיוון ההסתכלות המתואר בתת פרק [3.4.3](#) עבור וידאו רציף. כאמור, בהנתן תמונה בה המשתמש הסתכל למקום מסויים במסך, מצאנו את מיקום האישון שלו בכל עין ללא תלות בשניה. בפועל העקיבה נעשית בווידאו בו המשתמש עוקב אחרי המילים בצורה רציפה. המילים מופיעות במהירויות שונות, חצי שניה ושניה. בחרנו מהירויות אלה כדי לדמות תהליך קריאה רגיל ואיטי של המשתמש בהתאם. מפני שאנו דוגמים בקצב של 30 [Fps] וכדי לתמוך במהירות של חצי שניה, למעשה נקבל חוצץ המכיל 15 מסגרות שונות שבהתבסס עליהן נוכל לקבוע את כיוון ההסתכלות המשתמש. חשוב להבין שבבחירת גודל החוצץ ישנו trade-off. ככול שהוא יכול יותר מסגרות, קביעת כיוון ההסתכלות תהיה איכותית ומדויקת יותר מפני שהיא תתבסס על יותר מידע, אך מנגד נוכל לתמוך במהירות קריאה איטית יותר. כפי שכבר תיארנו בהקדמת תת פרק זה, בחלק מהמסגרות הללו האישון לא מזוהה נכונה ולכן יש צורך בסינון.

לצורך כך נעביר את מסגרות העקיבה בשני סוגי מסננים, בדומה לנאמר בנוגע לסינון מסגרות הכיול בתת הפרק הקודם:

1. מסנן רמות אפור: מסנן זה מתבסס על העיקרון שהאישון הוא האובייקט הכי כהה בתמונת העין הסופית שקיבלנו מהאלגוריתם (לאחר תת פרק [3.2](#)). למעשה, זהו מסנן סף שקובע על סמך ממוצע רמות האפור של האובייקט האם הוא יכול להיות אישון או לא, זאת בייחס לרמות האפור של שאר האובייקטים שזוהו כאישונים בשאר המסגרות שנכנסו לחוצץ. קביעת הסף עבור מסגרות העקיבה שונה מעט מקביעת סף מסגרות הכיול כפי שתואר בתת הפרק הקודם. למעשה, במקרה זה נחזיק בנוסף סף גלובלי שיקבע האם החוצץ תקף או לא. עבור ניתוח העקיבה הראשוני הסף הגלובלי יוגדר להיות הסף שחושב בשלב הכיול. לאחר מכן בכל איטרציה של העקיבה, כאשר כל איטרציה מיוחסת לחוצץ חדש, סף זה יתעדכן בהתאם לקראתה. ראשית נבין את המוטיבציה לקביעת הסף הגלובלי. כפי שהוסבר בתת פרק [3.1](#) ממוצע רמות האפור נורמלו כך שיהיו אחידים בכל התמונות, זאת לצד ההנחה שתוארה בתחילת פרק זה שתנאי התאורה אחידים בעת צילום הוידאו מביאים להבנה שאנו מצפים לקביעת סף יחסית אחיד לכל החוצצים. מטרת הסף הגלובלי היא לזהות מתי לחוצץ חושב סף חריג ביחס לאחרים, זה נעשה באופן הבא:

$$\begin{aligned} & \text{If } min > global\_TH \\ & \quad TH = global\_TH \\ & \text{Else } TH = min + \frac{max - min}{3} \\ & \text{If } gray\_level > TH \\ & \quad \text{remove} \end{aligned}$$



כלומר, לאחר שנחשב את רמות האפור הממוצעות של האישונים בכל 15 המסגרות שהתקבלו, נבדוק תחילה האם הערך המינימלי (min) שחושב גדול מהסף הגלובלי. במידה וכן, ותחת ההנחות שתוארו לעיל, מדובר בחוצץ בעל סף חריג ולכן לא נרצה להשתמש בו. אחרת, באותו אופן כמו בשלב הכיול, נגדיר את הסף כך שנשמור רק את המסגרות הנמצאות בשליש התחתון של התחום בין הערך המקסימלי למינימלי שמצאנו. זאת בהתבסס על כך שנצפה שרמת האפור הממוצעות של האישון תהיה נמוכה. גם במקרה זה חשוב להבין שאין הדבר אומר שרק שליש מסך המסגרות, כלומר 5, יעברו את הסף. אלא שכל המסגרות הנמצאות בשליש התחתון של תחום רמות האפור שהוגדר יעברו אותו. נציין שוב שבחרנו את השליש התחתון בצורה אמפירית על סמך ניתוח מאגר הסרטונים שלנו. לסיום כדי לעדכן את הסף הגלובלי לקראת איטרצית העקיבה הבאה, נמצע בינו לבין הסף שחושב עבור החוצץ הנוכחי:

$$global\_TH = avg(global\_TH, TH)$$

נציין שבמידה והחוצץ לא נפסל אזי הסף שמחושב לו בלתי תלוי בסף של קודמו. כלומר אם יהיו שינויים קטנים בתנאי התאורה על פניו של המשתמש כתוצאה מתזוזה קלה שלא יובילו לפסילת החוצץ, האלגוריתם ימשיך לעבוד כנדרש. תכונה זו של אי תלות החוצצים האחד בשני (עד כדי מקרה הקיצון של פסילת חוצץ) מקנה לאלגוריתם את היכולת להסתגל לשינויים קטנים בתנאי התאורה לאורך זמן, תוך עדכון הסף הגלובלי בהתאם.

בדומה לשלב הכיול, לאחר המעבר במסנן זה אכן רוב המסגרות שהכילו זיהוי אישון שגוי, כלומר מסיח, סוננו. המסיחים שנשארו היו לרוב עיגולים במקומות כהים יותר באזור העין כגון קצוות חלל העין או אזור הגבה במידה ולא נחתך בשלבי האלגוריתם המוקדמים יותר.

כדי להתגבר על כך, באופן זהה לחלוטין לנאמר בשלב הכיול, הגדרנו מסנן נוסף המתבסס על מיקום. **2. מסנן מיקום:** כפי שצויין, לאחר המעבר במסנן רמות האפור, כמות המסיחים קטנה משמעותית ולכן החלק היחסי שלהם מסך כל המסגרות שנותרו קטן מאוד. בהתבסס על כך מסנן המיקום שבנינו מסנן על סמך מיקום חריג שנקבע במספר מסגרות בודדות ביחס לכל שאר המסגרות. כך למעשה אנו מסננים מסגרות הקובעות מיקום אישון חריג הן בציר X והן בציר Y ביחס לכל שאר המסגרות.

מסנן זה בנוי על עיקרון של מסנן הפרש ומסנן כאשר יש הפרש גדול מדיי בין קביעת מיקום האישון של שתי מסגרות עוקבות (כלומר נגזרת גבוהה ולא צפויה). על סמך מאגר הסרטונים שלנו וברזולוציה הנתונה מצאנו בצורה אמפירית את ערך הסף לשינוי בנגזרת להיות 5 פיקסלים.

לאחר המעבר בשני סוגי המסננים שהוגדרו לעיל ובכדי לקבל החלטה אודות מיקום האישון הסופי, ביצענו הן מיצוע והן פעולת חציון על המסגרות שנותרו בחוצץ. שתי הפעולות בוצעו במקביל ובאופן בלתי תלוי אחת בשנייה, זאת במטרה לבדוק מי מהן עדיפה תוך נסיון למקסם את ביצועי האלגוריתם. נרחיב על כך בפרקי הסימולציות והתוצאות.

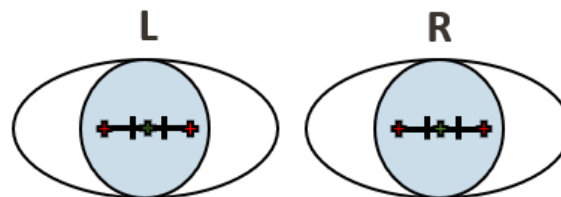
כעת, לאחר שקבענו מהו כיוון הסתכלות המשתמש על סמך החלטת החוץ, נייצג את ההחלטות בשני וקטורים כפונקציה של הזמן. וקטור לכל עין, שכן ההחלטות מתקבלות באופן בלתי תלוי לשתייהן. כל איבר בווקטור מכיל מספר מקטע עבור החוץ הרלוונטי, לפי הסדר הכרונולוגי של החוצצים ברזולוציה זמנית של 0.5 שניה.

משמעות מספר המקטע מודגמת באיור 31 בתת פרק 3.4.3.

| Left Eye | Right Eye |                 |
|----------|-----------|-----------------|
| 1        | 1         | $T = 0$ [sec]   |
| 1        | 2         | $T = 0.5$ [sec] |
| 2        | 2         | $T = 1$ [sec]   |
| 3        | 3         | .               |
| 4        | 5         | .               |
| 5        | 6         | .               |
| 8        | 8         | .               |
| .        | .         |                 |
| .        | .         |                 |
| .        | .         |                 |

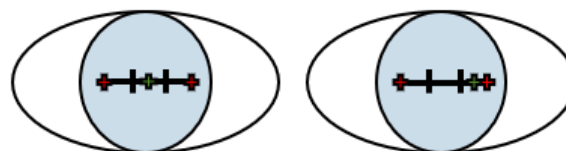
איור 34 - וקטורי הפלט של האלגוריתם המייצגים את כיוון ההסתכלות הנקבע לכל עין בנפרד

כעת, בחלק מניתוחי העקיבה אחר אישוני המשתמש התקבלה הסכמה בין שתי העיניים לגבי מספר המקטע עליו הוא מתבונן, ובחלק לא. ניתן להבחין בזאת גם באיור 33 לעיל. יחד עם זאת נציין כי גם כאשר לא הייתה הסכמה בין שתי העיניים, בממוצע ההבדל המקסימלי בתוצאות היה 3 מקטעים. למעשה, כפי שכבר ציינו, האלגוריתם פועל על כל עין בנפרד ולכן גם תוצאת קביעת שתי העיניים אודות כיוון ההסתכלות בלתי תלויה בניהן. לשם הדגמה נציג 3 מקטעים.



איור 35 - הסכמה בין שתי העיניים

למשל באיור 34, הן הניתוח עבור העין הימנית והן הניתוח עבור העין השמאלית קבע כי מספר המקטע שבו מתבונן הצופה הוא 2. זאת לעומת איור 35 בו הניתוח עבור העין הימנית קבע כי הצופה מתבונן במקטע מספר 3, ואילו הניתוח עבור העין השמאלית קבע כי הצופה מתבונן במקטע מספר 2.



איור 36 - חוסר הסכמה בין שתי העיניים

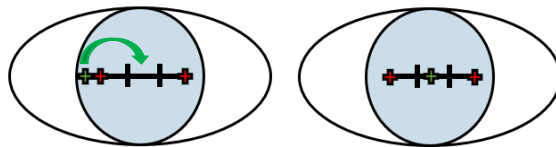
במסגרת פרויקט זה לא עסקנו בקביעת כיוון ההסתכלות סופי במקרה בו ישנו חוסר הסכמה בין שתי העיניים. עם זאת, נציע שתי דרכים לקביעת כיוון ההסתכלות סופי במקרה זה:

1. מיצוע בין החלטת שתי העיניים.
2. קביעת כיוון ההסתכלות על סמך העין שהניבה תוצאות מדויקות יותר. נציין כי ניתן לפתח כיוון זה ולבדוק האם ישנם מקרים בהם עין ימין מניבה תוצאות טובות יותר לעומת מקרים בהם עין שמאל מניבה תוצאות טובות יותר, זאת כתלות במספר המקטע עליו מסתכל המשתמש. למעשה נצפה לתוצאות שונות בין שתי העיניים עקב הסקת כיוון ההסתכלות בצורה לינארית. לדוגמה: נשער כי כאשר המשתמש מסתכל לחלקו הימני של המסך, עין ימין תניב תוצאות טובות יותר. זאת בשל העובדה שהזווית בין מיקום העין למיקום הטקסט עליו מסתכל המשתמש על המסך קטנה יותר, כלומר ההנחה הלינארית תקפה יותר.

### 3.5.4. תיקון קביעת כיוון ההסתכלות

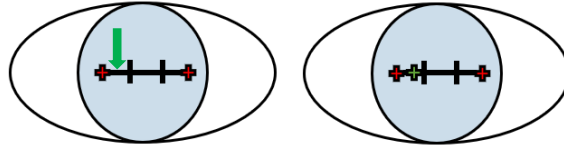
במקרה של וידאו, בניגוד לתמונה בודדת, נוכל להסתמך על ממד הזמן ורציפות הוידאו כדי לבצע תיקונים אודות קביעת כיוון ההסתכלות המשתמש במקרי קיצון. לשם הדגמה נציג 3 מקטעים המקרים אותם נתקן:

1. הזיהוי באחת העיניים נמצא מחוץ לגבולות הגזרה שהוגדרו לתנועת האישון. כפי שכבר ציינו בתת פרק [3.4.3](#), תיאורטית מקרה זה לא אפשרי ולא נצפה לקבלו. מעשית בשל תזוזות ראש של המשתמש, גם אם היו קטנות, מקרה זה יכול להתקבל. בכדי להתגבר על כך נקבע את כיוון ההסתכלות בעין שחרגה מגבולות הגזרה להיות זהה לזה שנקבע בעין המקבילה. ניתן היה לתקן כך שהעין שחרגה תקע את כיוון ההסתכלות להיות מספר המקטע הקרוב ביותר לזיהוי שהתקבל, כך שמספר המקטע יהיה חוקי.



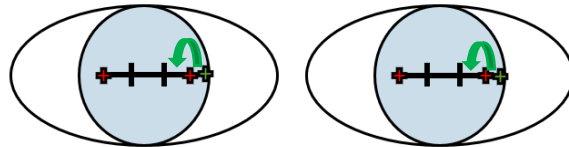
איור 37 - תיקון כיוון ההסתכלות במקרה בו הזיהוי באחת העיניים נמצא מחוץ לגבולות הגזרה

2. הזיהוי באחת העיניים לא התבצע. מקרה זה יכול להתקבל כאשר היה רצף של 15 מסגרות עוקבות שהאלגוריתם של Viola&Jones נכשל בהן, כאשר כל 15 המסגרות שנכנסו לחוצץ באיטרציה הנוכחית סוננו על-ידי המסננים שהגדרנו בתת הפרק הקודם או כאשר החוצץ כולו נפסל בהתאם לסף הגלובלי הנוכחי. כדי להתגבר על כך נקבע את כיוון ההסתכלות בעין שלא התבצע בה זיהוי להיות זהה לזה שנקבע בעין המקבילה.



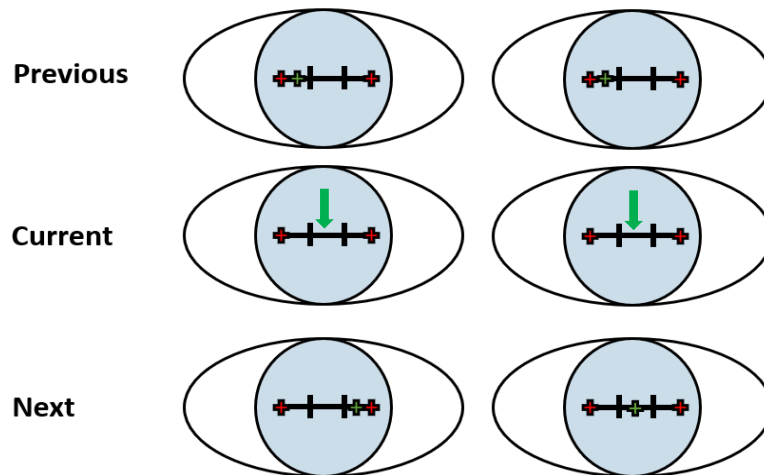
איור 38 - תיקון כיוון ההסתכלות במקרה בו הזיהוי באחת העיניים לא התבצע

3. הזיהוי בשתי העיניים נמצא מחוץ לגבולות הגזרה שהוגדרו לתנועת האישון. במקרה זה בניגוד למקרה שתואר בסעיף 1 לא נוכל להסתמך על אחת העיניים בתיקון. כדי להתגבר על כך נקבע את כיוון ההסתכלות בשתי העיניים להיות מספר המקטע הקרוב ביותר לזיהוי שהתקבל, כך שמספר המקטע יהיה חוקי. כלומר, אם הזיהוי קרוב יותר למקטע מספר 1 נקבע כי המשתמש מסתכל אליו, ובהתאם למקטע מספר 10. נציין שאלו שני מספרי המקטעים היחידים שניתן לתקן אליהם במקרה זה. בנוסף נציין שלא התקבל מקרה בו עין אחת נדרשה לתיקון שיצביע למקטע מספר 1 ואילו העין השנייה נדרשה לתיקון שיצביע למקטע מספר 10. זאת בהתאם לנאמר בתת הפרק הקודם שגם כאשר לא הייתה הסכמה בין שתי העיניים, ההבדל בתוצאות עבורן היה עד כדי 3 מקטעים.



איור 39 - תיקון כיוון ההסתכלות במקרה בו הזיהוי בשתי העיניים נמצא מחוץ לגבולות הגזרה

4. הזיהוי בשתי העיניים לא התבצע. במקרה זה בניגוד למקרה שתואר בסעיף 2 לא נוכל להסתמך על אחת העיניים בתיקון. בכדי להתגבר על כך נקבע את כיוון ההסתכלות בשתי העיניים להיות ממוצע כיוון ההסתכלות שזוהה בחוצץ הקודם והחוצץ הבא. אם מדובר בחוצץ הראשון או האחרון כך שלא ניתן לעשות מיצוע כפי שתואר לעיל, מספר המקטע נבחר להיות 1 או 10 בהתאם. זאת על סמך הידע המקדים שהמשתמש מתחיל לקרוא ממקטע מספר 1 ומסיים לקרוא במקטע מספר 10. במידה ולא נרצה להסתמך על ידע מקדים זה, נוכל לקבוע את כיוון ההסתכלות לפי החוצץ העוקב או הקודם בהתאם. הצעת תיקון זו מסתמכת על העובדה שאין חוסר זיהוי בשתי העיניים למשך שני חוצצים עוקבים, ואכן לפי תוצאות האלגוריתם חוסר זיהוי בשתי העיניים במקביל קורה לעיתים רחוקות למדי. נציין כי המיצוע הנ"ל הינו חשבוני. כלומר, לא מתחשב בעובדה שתהליך הקריאה הוא ציקלי. לכן נוכל לשפר את תיקון מקרה זה על-ידי ביצוע מיצוע ציקלי עבור ניתוח סרטון הקריאה הרציפה.



איור 40 - תיקון כיוון ההסתכלות במקרה בו הזיהוי בשתי העיניים לא התבצע

בנוסף, חשוב להבין שהצעת תיקון זו מבטלת את אפשרות האלגוריתם להיות אלגוריתם real-time. זאת מפני שהיא משתמשת בחישוב שמתבצע בחוצץ העוקב לחוצץ הנוכחי, כלומר במילים אחרות משתמשת במידע עתידי.

על סיבוכיות זמן האלגוריתם נרחיב בפרק הסימולציות, לעת עתה נציין שהיא לא תומכת באלגוריתם real-time. אך במידה והיא הייתה תומכת בכך, נציע שתי דרכים כדי לאפשר תכונה זו:

1. התבססות רק על סמך החוצץ הקודם לנוכחי, ללא מיצוע עם החוצץ העתידי.
2. בהנחה וזמן קביעת כיוון הסתכלות עבור חוצץ שואף לאפס, ובהנחה שאנו מאפשרים השהייה של 0.5 שניה, כן נוכל לחשב את כיוון ההסתכלות גם על סמך החוצץ העתידי ולמצע. זאת מפני שאנו זקוקים למידע עתידי מחוצץ עוקב אחד בלבד, כלומר מידע עתידי מנקודת זמן של 0.5 שניה קדימה.

## 4. סימולציות

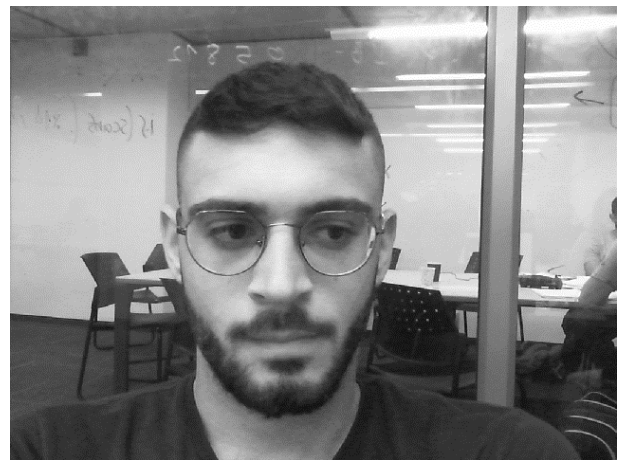
### 4.1. מאגר המידע

לצורך בחינת הביצועים של האלגוריתם, בנינו מאגר תמונות וסרטונים. המאגר צולם בעזרת מצלמת רשת של חברת Microsoft בעלת רזולוציה 640X480 אשר מצלמת בקצב של 30[fps]. כדי לבנות את המאגר צילמנו אנשים בטווחי גילאים 20-50, ויצרנו מאגר הכולל 100 תמונות (200 עיניים) וכן 46 סרטוני עקיבה.

מאגר התמונות מורכב מתמונות צבע אשר צולמו במקומות שונים (מעבדות, ספריות, מרחבי למידה, בתים) ובתנאי תאורה מגוונים. טרם תחילת ביצוע האלגוריתם התמונות הומרו לרמות אפור, כל תמונה סיפקה ניתוח זוג עיניים באופן בלתי תלוי אחת בשנייה. ראשית, התמונות שימשו לפיתוח החלק הראשון של האלגוריתם בו זיהינו את עיני המשתמש והאישונים בלבד (ללא עקיבה). תחילה צילמנו עבור כל נבדק שתי תמונות בהן הוא מסתכל לשני קצוות המסך.



איור 42 - מבט לקצה שמאל של המסך



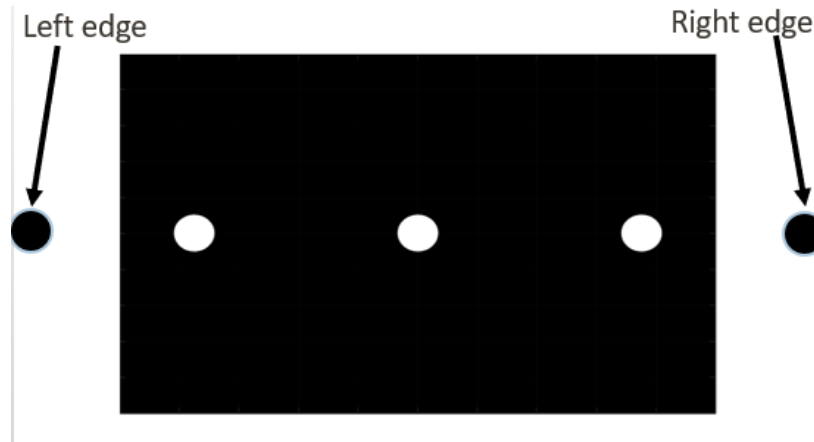
איור 41 - מבט לקצה ימין של המסך

לאחר מכן, הרחבנו את המאגר עם נבדקים נוספים כך שכל נבדק סיפק 5 תמונות:

- שתי תמונות כיול עם מבט לקצוות המסך (כמתואר מקודם).
- שלוש תמונות נוספות בהן הסתכל למקומות שונים בתחום המסך.

בסך הכול, המסך עליו הסתכל המשתמש חולק ל-5 מקטעים, כך שכל תמונה נלקחה בעת הסתכלותו על מקטע אחר.

על תמונות אלו בדקנו את אלגוריתם זיהוי העיניים שפיתחנו, אחוזי הצלחת שלב זה עמדו על כ 98%. בשלב הבא ובהסתמך עליהן, פיתחנו את אלגוריתם העקיבה וזיהוי כיוון ההסתכלות תוך בדיקת ביצועיו. בסך הכול הגענו למאגר הכולל 100 זוגות עיניים.



איור 43 - חלוקת המסך ל-5 מקטעים

לאחר מכן עברנו לניתוח של וידאו. צילמנו 24 נבדקים שעקבו אחרי 2 סוגי סרטונים. תהליך צילום הסרטונים בוצע במעבדת SIPL בטכניון ובתנאי תאורה אחידים. הראשון מבניהם מדמה קריאה רציפה של טקסט בציר האופקי ומכיל מילים באורכים שונים אשר מופיעות על גבי המסך, כל פעם מילה בודדת, במהירויות שונות. המילים מרכיבות משפטים לוגיים. אורכו 50 שניות. צילמנו סה"כ 24 סרטונים כאלו. השני מבניהם זהה לראשון פרט לכך שהוא מציג מילים בצורה אקראית על המסך. אורכו 30 שניות. צילמנו סה"כ 22 סרטונים כאלו. מטרתו היא לנטרל את יכולת המשתמש לנבא היכן המילה הבאה תופיע על המסך. נציין כי גם מאגר הווידאו הומר לגווני אפור טרם תחילת ביצוע האלגוריתם, שכן ניתוח וידאו זהה בבסיסו לניתוח תמונה בודדת, מסגרת (frame) אחר מסגרת.



|          |  |  |  |  |  |  |  |  |  |
|----------|--|--|--|--|--|--|--|--|--|
| Football |  |  |  |  |  |  |  |  |  |
|----------|--|--|--|--|--|--|--|--|--|

איור 5 – הדגמה לווידאו (הוידאו מוצג בצורה מתוחה)

נציין כי האיור לעיל מוצג ביחס ממדים מתוח ולא טבעי במטרה להרחיב ויזואלית גודל כל מקטע וזאת כדי שיהיה ניתן לראות בצורה טובה יותר את החלוקה למקטעים.  
מענה כל האיוורים מסוג זה יוצגו ביחס הממדים הנ"ל.  
את התפלגות מספר המילים בסרטוני העקיבה נציג בטבלה הבאה :

טבלה 1 - התפלגות מספר המילים בסרטוני העקיבה

| סרטון עקיבה רציף    |                               |                               |            |
|---------------------|-------------------------------|-------------------------------|------------|
| מספר המילים סך-הכול | מספר המילים עבור מהירות גבוהה | מספר המילים עבור מהירות נמוכה | אורך המילה |
| 58                  | 28                            | 30                            | 1          |
| 25                  | 15                            | 10                            | 2          |
| 15                  | 1                             | 14                            | 3          |
| 2                   | 0                             | 2                             | 4          |
| סרטון עקיבה אקראי   |                               |                               |            |
| מספר המילים סך-הכול | מספר המילים עבור מהירות גבוהה | מספר המילים עבור מהירות נמוכה | אורך המילה |
| 36                  | 14                            | 22                            | 1          |
| 13                  | 3                             | 10                            | 2          |
| 10                  | 0                             | 10                            | 3          |
| 2                   | 0                             | 2                             | 4          |



## 4.2. תיאור הסימולציות

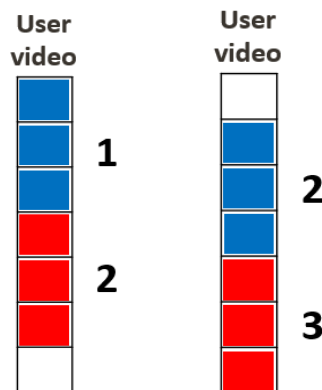
ראשית, נציין שבכל הסימולציות ניתחנו את ביצועי האלגוריתם ב offline, כלומר ניתחנו סרטון עקיבה קיים ולא בזמן אמת.

כעת, את ביצועי האלגוריתם מדדנו ביחס לתוצאות עקיבה ידועות מראש שציפינו לקבל לפי הסרטון עליו הסתכל המשתמש. במטרה למקסם את ביצועי האלגוריתם, בדקנו אותו בתנאים שונים בעזרת בניית מספר רחב של סימולציות.

נרחיב על הפרמטרים המבדילים את הסימולציות אחת מהשנייה:  
תחילה נזכיר, כפי שכבר ציינו בתת הפרקים 3.5.2 ו 3.5.3, ההחלטה אודות כיוון ההסתכלות בעת ניתוח וידאו מתבצעת על סמך חוצץ המכיל מסגרות עוקבות.

1. מיצוע לעומת פעולת חציון על חוצץ: למעשה כל מסגרת מספקת לנו מידע על מיקום האישון של המשתמש. בכדי להסיק כיוון כללי על סמך כל המסגרות בחוצץ נוכל למצע את מיקומי האישון בכולן או להפעיל פעולת חציון. למעשה פרמטר זה בודק כיצד עדיף לקבל את ההחלטה אודות כיוון ההסתכלות. על ההשפעה של כך על תוצאות האלגוריתם נרחיב בפרק התוצאות.

2. הזזת החוצצים על פני סרטון העקיבה בעת הניתוח:  
הרעיון בפרמטר זה הוא לבצע הזזה על פני ניתוח כל הסרטון במטרה להבין האם ישנה בעיית סנכרון בעת השוואת תוצאות האלגוריתם לתוצאות הצפויות.  
כל חוצץ מכיל 15 מסגרות, פעולת ההזזה קודמת לפעולת הסינון ולכן מתבצעת על פני כל מסגרות הסרטון. כעת, ייתכן וההחלטה אודות כיוון ההסתכלות עבור חוצץ מסוים תשתנה כתלות במסגרות אותן הוא מכיל, גם אם מדובר רק במסגרת אחת שונה. הדבר מודגם באיור 43 מטה.

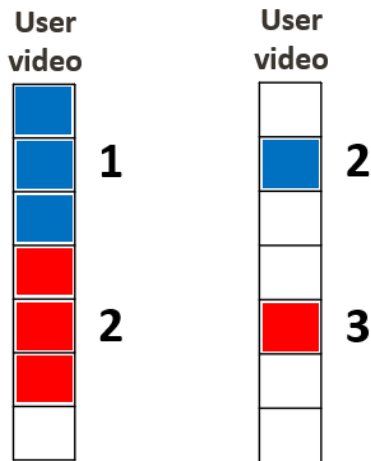


איור 44 - קביעת כיוון הסתכלות (מספר מקטע) שונה כתוצאה מהזזת החוצץ על פני הסרטון

באיור 43 ניתן לראות הדגמה להזזה המתוארת. נניח כי כל ריבוע מייצג מסגרת בסרטון וכי כל חוצץ מכיל 3 מסגרות. באיור מוצגים 2 חוצצים בצבעים שונים, כחול ואדום.  
בצד השמאלי ניתן לראות את החלטת כל חוצץ, לעומת החלטת כל חוצץ בצד הימני לאחר ההזזה.

כל שנותר לעשות הוא לבחון את תוצאות האלגוריתם לפני ולאחר ההזזה ולקבוע האם ישנה בעיית סנכרון. נרחיב על כך בפרק התוצאות.

3. קטימת החוצץ משני קצותיו: הרעיון בפרמטר זה הוא לנתח רק את המסגרות בהן האישון יציב ומקובע על מילה ספציפית. כלומר ללא המסגרות שצולמו בעת תנועת האישון אל ומהמילה עליה מסתכל המשתמש, שכן הן יכולות להוות מסיח בעת חישוב כיוון ההסתכלות. כל חוצץ מכיל 15 מסגרות, פעולת הקטימה קודמת לפעולת הסינון ולכן מתבצעת על פני כל מסגרות הסרטון. כעת, ייתכן וההחלטה אודות כיוון ההסתכלות עבור חוצץ מסוים תשתנה כתלות בכך שנקטום את המסגרות בקצותיו. הדבר מודגם באיור 44 מטה.



איור 45 - קביעת כיוון הסתכלות (מספר מקטע) כתוצאה מקטימת קצוות החוצץ

באיור 44 ניתן לראות הדגמה לקטימה המתוארת. נניח כי כל ריבוע מייצג מסגרת בסרטון וכי כל חוצץ מכיל 3 מסגרות. באיור מוצגים 2 חוצצים בצבעים שונים, כחול ואדום. בצד השמאלי ניתן לראות את החלטת כל חוצץ, לעומת החלטת כל חוצץ בצד הימני לאחר הקטימה. כל שנותר לעשות הוא לבחון את תוצאות האלגוריתם לפני ולאחר הקטימה ולקבוע האם היא תורמת. נרחיב על כך בפרק [התוצאות](#).

לאחר שהסברנו על הפרמטרים המבדילים בין הסימולציות, נציין אילו סימולציות ביצענו משילובם :

- החלטה לפי מסנן מיצוע :
  - ללא הזזה.
  - הזזה של 5 מסגרות.
  - הזזה של 10 מסגרות.
  - קטימת 4 מסגרות מתחילת ומסוף החוצץ. (מתוך 15 מסגרות סך-הכול).
- החלטה לפי מסנן חציון :
  - ללא הזזה.
  - הזזה של 5 מסגרות.
  - הזזה של 10 מסגרות.
  - קטימת 4 מסגרות מתחילת ומסוף החוצץ. (מתוך 15 מסגרות סך-הכול).

סך-הכול ביצענו לכל סרטון וידאו 8 סימולציות כמתואר לעיל.

### 4.3. סיבוכיות זמן

ראשית, נציין שבמסגרת מטרות פרויקט זה לא התעסקנו בייעול סיבוכיות זמן, לכן נציג את נתוני סיבוכיות הזמן שהתקבלה ולא נרחיב על כך מעבר.

זמן ניתוח מסגרת בודדת ארך במוצא כחצי שניה על מחשבי מעבדת SIPL בטכניון המכילים GPU. כתוצאה מכך זמן סימולציה בודדת על סרטון שלם המכיל כ 1800 מסגרות לקח כרבע שעה. סך-הכול ניתחנו 46 סרטוני וידאו, כשכול סרטון מנותח 8 פעמים, כל פעם בתנאי סימולציה אחרים. לכן נעריך את זמן ניתוח כל מאגר הסרטונים קרוב ל 4 ימים. נציין שעל מחשב סטנדרטי ללא GPU זמן ניתוח מסגרת בודדת ערך במוצא כשניה. כתוצאה מכך וכפי שכבר ציינו בתת הפרק [3.5.4](#), סיבוכיות הזמן שהתקבלה לא מתאימה לתנאי real-time.

## 5. תוצאות

### 5.1. ניתוח התוצאות

כדי לנתח את ביצועי האלגוריתם בצורה יעילה ואיכותית, פיתחנו כלים לניתוח התוצאות. פיתוחם היה בלתי נפרד משלבי פיתוח האלגוריתם במסגרת פרויקט זה. הראשון מבניהם עוסק בניתוח כמותי מדויק והשני בניתוח ויזואלי איכותי. נפרט על שני הכלים בהרחבה:

#### 1. ניתוח כמותי מדויק:

כלי זה מתבסס על השוואה וקטורית בין שני וקטורים כפונקציה של הזמן, אשר מכילים מספרי מקטעים המעידים על כיוון הסתכלות ברזולוציה זמנית של חצי שניה. (כפי שמוצג באיור 33 בתת פרק [3.5.3](#)).

ההשוואה הווקטורית מתבצעת בין וקטור הפלט של האלגוריתם (עבור כל עין בנפרד) לבין וקטור קבוע וידוע, אותו יצרנו מראש, המכיל את מספרי המקטעים הנכונים כפונקציה של הזמן בהתאם לסרטון העקיבה עליו הסתכל המשתמש. נציין שההשוואה אפשרית מעצם סנכרון שני הסרטונים כמתואר בתת פרק [3.5.1](#). כמו כן נציין ששלב הניתוח הותאם הן לסרטוני העקיבה המדמים קריאה רציפה והן לסרטוני העקיבה האקראיים.

כעת, כתוצאה מההשוואה הווקטורית קיבלנו וקטור בינארי המעיד על "הצלחה" או "כישלון" ברזולוציה זמנית של חצי שניה. נציין שעבור מילה הנפרסת על מספר מקטעים נדרש מהאלגוריתם לקבוע את כיוון ההסתכלות להיות אחד מהם כדי להכריז על "הצלחה", שכן האלגוריתם מספק כפלט רק מספר מקטע אחד בכל פעם.

במטרה לנתח את האלגוריתם בצורה איכותית ומעמיקה, ניתחנו עבור כל סימולציה את התוצאות לפי מספר פרמטרים:

- ניתוח כל סרטון הוידאו.
- ניתוח הוידאו לפי מהירות הצגת המילים.
- ניתוח הוידאו לפי אורך מילה ביחידות של מקטע.

נציין שהניתוחים הללו בוצעו פעמיים. בראשונה "הצלחה" הוגדרה כאשר האלגוריתם קבע כיוון הסתכלות מדויק ונכון. בשנייה "הצלחה" הוגדרה כאשר האלגוריתם קבע כיוון הסתכלות עד כדי קודמו ועוקבו של מספר המקטע הנכון, זאת במטרה להבין איכותית האם כיוון ההסתכלות שנקבע הוא נכון. נציין כי סימון הניתוח השני בטבלאות התוצאות שיוצגו בהמשך הינו "+1-".

בנוסף לאלו, עבור כל סימולציה שנותחה, חילצנו מתוך מספר המסגרות הכולל (totalFrameNum) שנותח את מספר המסגרות שעל סמכן נקבע כיוון ההסתכלות לאורך כל הניתוח (goodFrameNum). זאת במטרה לבחון כמה היא אמינה לצד בחינת המסננים שבנינו כמתואר בתת הפרק [3.5.3](#).

## 2. ניתוח ויזואלי איכותי:

כלי זה מאפשר לצפות בסרטון המשתמש שצולם, במקביל לסרטון המילים שהוצג לפניו וקביעת האלגוריתם אודות כיוון ההסתכלות.

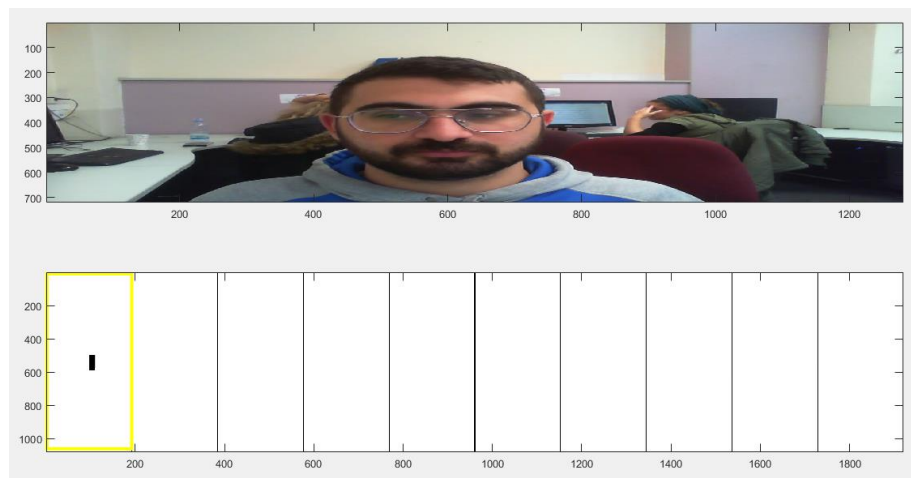
בדומה לניתוח הכמותי, נציין שההצגה הוויזואלית אפשרית מעצם סנכרון שני הסרטונים כמתואר בתת פרק [3.5.1](#).

המוטיבציה לניתוח זה היא לקבל מושג איכותי האם האלגוריתם זיהה את האזור עליו הסתכל המשתמש, גם אם לא צדק במדויק.

כמו כן, כלי זה מאפשר לעקוב אחר המשתמש ולבחון האם שמע להוראות ושמר על הנחות הפרויקט, בדגש על ההוראה לא להטות ולהזיז את הראש כלל.

נוסף לכך, כלי זה מאפשר לבדוק האם המשתמש אכן הסתכל על המילה שנדרש להסתכל עליה בכל נקודת זמן.

לאחר בחינת מאגר הסרטונים בעזרת כלי זה גילינו כי בחלק מן המקרים המשתמש כלל לא הסתכל לאן שנדרש. לכן, למרות שהאלגוריתם קבע נכונה את כיוון ההסתכלות בהתאם למקטע המסך עליו הסתכל המשתמש, בעת הניתוח הכמותי המדויק הדבר נחשב לכישלון במקום הצלחה. נרחיב על נקודה זו בהמשך הדיון בפרק זה.



איור 46 - פלט הכלי הוויזואלי לניתוח איכותי של האלגוריתם

באיור לעיל ניתן לראות הדגמה לפלט הכלי הוויזואלי לניתוח האלגוריתם. ניתן להבחין בעיני המשתמש אשר מסתכל על המקטע הנכון בו מופיעה המילה. כמו כן ניתן לראות את קביעת האלגוריתם אודות כיוון ההסתכלות המסומנת בריבוע צהוב סביב המקטע המתאים.

בתת הפרקים הבאים נציג תוצאות כמותיות מדויקות המשקללות את כל ניתוחי הסימולציות על כל הסרטונים. לאחר מכן נציג תוצאות איכותיות וויזואליות. כל זאת עבור כל אחד מסוגי הסרטונים, רציף ואקראי.

נציין שבפועל התוצאות המשוקללות חושבו על פני 18 סרטוני עקיבה רציפים ו 16 אקראיים. הסיבה לכך היא שטרם ניתוח תוצאות סרטוני העקיבה ביצענו עליהם סינון במטרה לנפות סרטוני עקיבה לא מייצגים שלא עמדו בהנחות הפרויקט.

עבור כל סימולציה משוקללת נציג טבלת תוצאות המתארת את אחוזי ההצלחה שלה. כותרת כל טבלה תעיד על הפרמטרים שנבחרו עבור הסימולציה שהיא מתארת באופן הבא:

Filter\_FrameBufferSize\_Shift0/5/10 (Number)

כאשר:

- Filter: באיזה מסנן השתמשנו כדי לקבל החלטה אודות כיוון ההסתכלות, ממוצע או חציון.
- FrameBufferSize: גודל החוצץ. אם לא נקטם, גודלו המקורי הוא 15 מסגרות והדבר מתואר בסימון "1-15". אם נקטם בקצותיו, לפי המתואר בתת הפרק 4.1 הוא מכיל את מסגרות 5-11 ולכן הדבר מתואר בסימון "5-11".
- Shift0/5/10: הזזת החוצץ על פני סרטון העקיבה בעת הניתוח, כמתואר בתת הפרק 4.1. גודל זה מתאר את גודל ההזזה ביחידות של מסגרות, {0,5,10}.
- Number: מספר המיוחס לסימולציה אשר מזהה אותה באופן יחיד למטרות התייחסות בהמשך פרק זה.

נסביר לגבי מספר פרמטרים נוספים המופיעים בטבלאות:

- totalFrameNum: מספר המסגרות הכולל שעברו ניתוח בסימולציה.
- goodFrameNum: מספר המסגרות שעברו את המסננים ועל סמכן נקבעה תוצאת העקיבה.
- Left/right eye: אחוז ההצלחה המדויק עבור עין שמאל/ימין בהתאם.
- left/right eye +-1: אחוז ההצלחה האיכותי עבור עין שמאל/ימין בהתאם.
- all video: אחוז ההצלחה על גבי כל הסרטון.
- part 1 (velocity 1 sec): אחוז ההצלחה בחלק הסרטון בו המילים נעות במהירות של שניה על גבי המסך.
- part 2 (velocity 0.5 sec): אחוז ההצלחה בחלק הסרטון בו המילים נעות במהירות של חצי שניה על גבי המסך.
- Word length i: אחוז ההצלחה עבור מילים באורך i מקטעים.

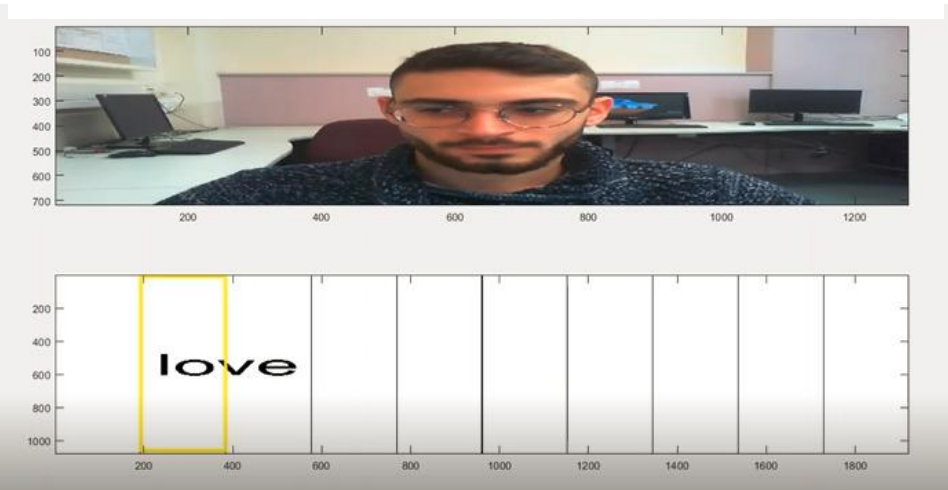
## 5.2. תוצאות סרטוני העקיבה הרציפים

| mean_1-15_shift0(1)       |                           |          |             |           |              | median_1-15_shift0(4)     |                           |          |             |           |              |
|---------------------------|---------------------------|----------|-------------|-----------|--------------|---------------------------|---------------------------|----------|-------------|-----------|--------------|
| totalFrameNum             | 1500.0                    | left eye | left eye +1 | right eye | right eye +1 | totalFrameNum             | 1500.0                    | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum              | 1073.1                    |          |             | 1079.1    |              | goodFrameNum              | 1073.2                    |          |             | 1079.9    |              |
| all video                 |                           | 41.8     | 71.0        | 43.3      | 70.6         | all video                 |                           | 43.2     | 73.3        | 45.5      | 73.8         |
| part 1 (velocity 1 sec)   |                           | 45.2     | 73.0        | 48.6      | 75.1         | part 1 (velocity 1 sec)   |                           | 46.4     | 73.2        | 52.0      | 77.5         |
| part 2 (velocity 0.5 sec) |                           | 37.6     | 68.4        | 36.5      | 64.8         | part 2 (velocity 0.5 sec) |                           | 39.0     | 73.4        | 37.2      | 69.1         |
| word length 1             | all video                 | 31.5     | 65.9        | 32.0      | 65.7         | word length 1             | all video                 | 32.4     | 68.9        | 33.3      | 69.0         |
|                           | part 1 (velocity 1 sec)   | 31.4     | 67.1        | 33.7      | 69.6         |                           | part 1 (velocity 1 sec)   | 32.7     | 67.3        | 36.7      | 72.2         |
|                           | part 2 (velocity 0.5 sec) | 31.7     | 64.7        | 30.3      | 61.6         |                           | part 2 (velocity 0.5 sec) | 31.9     | 70.6        | 29.6      | 65.5         |
| word length 2             | all video                 | 49.4     | 75.3        | 51.1      | 73.2         | word length 2             | all video                 | 52.0     | 77.4        | 54.1      | 76.7         |
|                           | part 1 (velocity 1 sec)   | 54.1     | 77.6        | 57.6      | 78.2         |                           | part 1 (velocity 1 sec)   | 55.3     | 77.6        | 62.4      | 80.6         |
|                           | part 2 (velocity 0.5 sec) | 46.3     | 73.7        | 46.7      | 69.8         |                           | part 2 (velocity 0.5 sec) | 49.8     | 77.3        | 48.6      | 74.1         |
| word length 3             | all video                 | 63.5     | 81.6        | 67.8      | 82.0         | word length 3             | all video                 | 65.5     | 82.0        | 72.5      | 85.1         |
|                           | part 1 (velocity 1 sec)   | 63.0     | 80.7        | 68.5      | 81.9         |                           | part 1 (velocity 1 sec)   | 64.7     | 81.1        | 72.3      | 84.5         |
|                           | part 2 (velocity 0.5 sec) | 35.3     | 47.1        | 29.4      | 41.2         |                           | part 2 (velocity 0.5 sec) | 38.2     | 47.1        | 38.2      | 47.1         |
| word length 4             | all video                 | 82.4     | 85.3        | 88.2      | 94.1         | word length 4             | all video                 | 79.4     | 85.3        | 88.2      | 94.1         |
|                           | part 1 (velocity 1 sec)   | 82.4     | 85.3        | 88.2      | 94.1         |                           | part 1 (velocity 1 sec)   | 79.4     | 85.3        | 88.2      | 94.1         |
| אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          | אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| mean_1-15_shifts(2)       |                           |          |             |           |              | median_1-15_shifts(5)     |                           |          |             |           |              |
| totalFrameNum             | 1500.0                    | left eye | left eye +1 | right eye | right eye +1 | totalFrameNum             | 1500.0                    | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum              | 1081.9                    |          |             | 1061.8    |              | goodFrameNum              | 1081.9                    |          |             | 1061.8    |              |
| all video                 |                           | 40.4     | 68.9        | 40.9      | 68.7         | all video                 |                           | 41.6     | 70.4        | 43.6      | 70.9         |
| part 1 (velocity 1 sec)   |                           | 44.4     | 70.9        | 44.7      | 73.3         | part 1 (velocity 1 sec)   |                           | 44.5     | 70.1        | 47.7      | 74.2         |
| part 2 (velocity 0.5 sec) |                           | 35.3     | 66.4        | 36.1      | 62.8         | part 2 (velocity 0.5 sec) |                           | 37.8     | 70.9        | 38.5      | 66.8         |
| word length 1             | all video                 | 32.7     | 65.8        | 33.1      | 66.3         | word length 1             | all video                 | 33.5     | 67.1        | 35.0      | 68.2         |
|                           | part 1 (velocity 1 sec)   | 34.1     | 66.9        | 32.7      | 70.4         |                           | part 1 (velocity 1 sec)   | 33.1     | 64.9        | 35.5      | 70.6         |
|                           | part 2 (velocity 0.5 sec) | 31.1     | 64.7        | 33.4      | 62.0         |                           | part 2 (velocity 0.5 sec) | 33.8     | 69.5        | 34.5      | 65.5         |
| word length 2             | all video                 | 42.8     | 69.2        | 43.8      | 66.6         | word length 2             | all video                 | 44.2     | 71.1        | 48.0      | 70.1         |
|                           | part 1 (velocity 1 sec)   | 47.6     | 70.6        | 51.8      | 71.8         |                           | part 1 (velocity 1 sec)   | 47.6     | 70.0        | 54.7      | 73.5         |
|                           | part 2 (velocity 0.5 sec) | 39.6     | 68.2        | 38.4      | 63.1         |                           | part 2 (velocity 0.5 sec) | 42.0     | 71.8        | 43.5      | 67.8         |
| word length 3             | all video                 | 62.4     | 80.0        | 63.5      | 80.0         | word length 3             | all video                 | 64.7     | 80.8        | 67.1      | 82.0         |
|                           | part 1 (velocity 1 sec)   | 60.5     | 79.4        | 62.6      | 79.8         |                           | part 1 (velocity 1 sec)   | 63.0     | 79.8        | 66.4      | 81.5         |
|                           | part 2 (velocity 0.5 sec) | 44.1     | 44.1        | 38.2      | 41.2         |                           | part 2 (velocity 0.5 sec) | 44.1     | 47.1        | 38.2      | 44.1         |
| word length 4             | all video                 | 70.6     | 73.5        | 64.7      | 79.4         | word length 4             | all video                 | 70.6     | 79.4        | 64.7      | 79.4         |
|                           | part 1 (velocity 1 sec)   | 70.6     | 73.5        | 64.7      | 79.4         |                           | part 1 (velocity 1 sec)   | 70.6     | 79.4        | 64.7      | 79.4         |
| אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          | אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| mean_1-15_shift10(3)      |                           |          |             |           |              | median_1-15_shift10(6)    |                           |          |             |           |              |
| totalFrameNum             | 1500.0                    | left eye | left eye +1 | right eye | right eye +1 | totalFrameNum             | 1500.0                    | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum              | 1079.3                    |          |             | 1078.3    |              | goodFrameNum              | 1059.3                    |          |             | 1065.4    |              |
| all video                 |                           | 34.1     | 65.3        | 32.6      | 63.6         | all video                 |                           | 33.4     | 64.5        | 34.9      | 66.0         |
| part 1 (velocity 1 sec)   |                           | 36.8     | 66.3        | 35.8      | 65.7         | part 1 (velocity 1 sec)   |                           | 38.0     | 64.9        | 39.7      | 68.8         |
| part 2 (velocity 0.5 sec) |                           | 30.6     | 64.0        | 28.5      | 61.1         | part 2 (velocity 0.5 sec) |                           | 27.4     | 63.9        | 28.7      | 62.4         |
| word length 1             | all video                 | 26.8     | 62.8        | 24.9      | 61.7         | word length 1             | all video                 | 26.6     | 61.0        | 27.7      | 62.8         |
|                           | part 1 (velocity 1 sec)   | 27.3     | 63.5        | 26.7      | 64.7         |                           | part 1 (velocity 1 sec)   | 28.6     | 60.8        | 30.2      | 65.7         |
|                           | part 2 (velocity 0.5 sec) | 26.3     | 62.0        | 23.1      | 58.4         |                           | part 2 (velocity 0.5 sec) | 24.4     | 61.1        | 25.0      | 59.7         |
| word length 2             | all video                 | 36.2     | 64.5        | 36.9      | 61.6         | word length 2             | all video                 | 35.5     | 66.6        | 36.9      | 65.4         |
|                           | part 1 (velocity 1 sec)   | 37.6     | 62.4        | 38.8      | 57.6         |                           | part 1 (velocity 1 sec)   | 41.8     | 65.3        | 42.4      | 64.1         |
|                           | part 2 (velocity 0.5 sec) | 35.3     | 65.9        | 35.7      | 64.3         |                           | part 2 (velocity 0.5 sec) | 31.4     | 67.5        | 33.3      | 66.3         |
| word length 3             | all video                 | 53.7     | 74.1        | 51.8      | 73.3         | word length 3             | all video                 | 50.6     | 72.9        | 56.1      | 77.6         |
|                           | part 1 (velocity 1 sec)   | 51.7     | 72.7        | 50.4      | 72.3         |                           | part 1 (velocity 1 sec)   | 50.4     | 71.8        | 55.5      | 77.3         |
|                           | part 2 (velocity 0.5 sec) | 41.2     | 47.1        | 35.3      | 44.1         |                           | part 2 (velocity 0.5 sec) | 26.5     | 44.1        | 32.4      | 41.2         |
| word length 4             | all video                 | 70.6     | 82.4        | 55.9      | 73.5         | word length 4             | all video                 | 73.5     | 76.5        | 58.8      | 79.4         |
|                           | part 1 (velocity 1 sec)   | 70.6     | 82.4        | 55.9      | 73.5         |                           | part 1 (velocity 1 sec)   | 73.5     | 76.5        | 58.8      | 79.4         |
| אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          | אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| mean_5-11_shift0(7)       |                           |          |             |           |              | median_5-11_shift0(8)     |                           |          |             |           |              |
| totalFrameNum             | 700.0                     | left eye | left eye +1 | right eye | right eye +1 | totalFrameNum             | 700.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum              | 534.6                     |          |             | 520.1     |              | goodFrameNum              | 506.6                     |          |             | 520.1     |              |
| all video                 |                           | 43.2     | 71.8        | 44.3      | 72.2         | all video                 |                           | 43.2     | 72.8        | 46.0      | 74.6         |
| part 1 (velocity 1 sec)   |                           | 46.6     | 73.4        | 50.1      | 75.7         | part 1 (velocity 1 sec)   |                           | 46.1     | 72.9        | 52.5      | 76.9         |
| part 2 (velocity 0.5 sec) |                           | 38.8     | 69.7        | 36.9      | 67.8         | part 2 (velocity 0.5 sec) |                           | 39.4     | 72.6        | 37.7      | 71.8         |
| word length 1             | all video                 | 33.0     | 67.3        | 33.4      | 69.1         | word length 1             | all video                 | 32.9     | 68.4        | 35.0      | 71.5         |
|                           | part 1 (velocity 1 sec)   | 33.7     | 67.8        | 36.1      | 72.0         |                           | part 1 (velocity 1 sec)   | 33.3     | 67.1        | 39.4      | 73.5         |
|                           | part 2 (velocity 0.5 sec) | 32.1     | 66.8        | 30.5      | 66.0         |                           | part 2 (velocity 0.5 sec) | 32.4     | 69.7        | 30.3      | 69.3         |
| word length 2             | all video                 | 50.1     | 73.6        | 51.5      | 71.5         | word length 2             | all video                 | 50.1     | 75.3        | 52.7      | 75.3         |
|                           | part 1 (velocity 1 sec)   | 52.9     | 74.1        | 60.0      | 74.1         |                           | part 1 (velocity 1 sec)   | 50.6     | 73.5        | 59.4      | 75.9         |
|                           | part 2 (velocity 0.5 sec) | 48.2     | 73.3        | 45.9      | 69.8         |                           | part 2 (velocity 0.5 sec) | 49.8     | 76.5        | 48.2      | 74.9         |
| word length 3             | all video                 | 66.3     | 83.5        | 69.0      | 83.5         | word length 3             | all video                 | 66.7     | 83.5        | 71.8      | 83.9         |
|                           | part 1 (velocity 1 sec)   | 65.1     | 82.8        | 68.1      | 83.2         |                           | part 1 (velocity 1 sec)   | 65.5     | 82.8        | 70.6      | 83.2         |
|                           | part 2 (velocity 0.5 sec) | 41.2     | 47.1        | 41.2      | 44.1         |                           | part 2 (velocity 0.5 sec) | 41.2     | 47.1        | 44.1      | 47.1         |
| word length 4             | all video                 | 79.4     | 88.2        | 85.3      | 88.2         | word length 4             | all video                 | 79.4     | 88.2        | 88.2      | 88.2         |
|                           | part 1 (velocity 1 sec)   | 79.4     | 88.2        | 85.3      | 88.2         |                           | part 1 (velocity 1 sec)   | 79.4     | 88.2        | 88.2      | 88.2         |
| אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          | אין מילים באורך 4 בחלק 2  | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |

טבלה 2 - תוצאות סרטוני העקיבה הרציפים



איור 47 - הדגמה לסרטון עקיבה רציף: המילה הראשונה



איור 48 – הדגמה לסרטון עקיבה רציף: המילה השנייה



איור 49 – הדגמה לסרטון עקיבה רציף: המילה השלישית



באירורים לעיל ניתן לראות המחשה לסרטון עקיבה רציף בו המשתמש עוקב אחר המילים המופיעות על גבי המסך. ניתן להבחין בכך שעיניו של המשתמש מסתכלות אל המילה הרצויה ואכן תוצאות האלגוריתם המסומנת בריבוע צהוב קובעת נכונה את כיוון ההסתכלות.

נבחן את תוצאות הסרטון הרציף במספר היבטים :

#### 1. מהירות הצגת המילים על המסך :

ניתן לראות כי עבור כל הסימולציות, אחוזי דיוק נמוכים יותר עבור המהירות הגבוהה (חצי שניה) לעומת הנמוכה (שניה). לאחר מיצוע על גבי התוצאות שהוצגו בטבלה 2 התקבל כי אחוז ההצלחה המדויק עבור מילים שהוצגו במהירות גבוהה עומד על 35.4% בעוד שאחוז ההצלחה המדויק עבור מילים שהוצגו במהירות נמוכה עומד על 45%. ככל הנראה אישון המשתמש יציב פחות במקרה בו מהירות הצגת המילים גבוהה יותר, כלומר ישנן פחות מסגרות איכותיות שבהסתמך עליהן נקבע כיוון ההסתכלות.

#### 2. אורך המילה ביחידות של מקטע :

באופן כללי, ככול שאורך המילה גדל כך תוצאות האלגוריתם טובות יותר. הדבר הגיוני מפני שיש יותר מרחב בקביעת כיוון ההסתכלות, כלומר הוא יכול להתפרס על פני מספר מקטעים. לאחר מיצוע על גבי התוצאות שהוצגו בטבלה 2 התקבל כי אחוז ההצלחה המדויק עבור מילים שהתפרסו על גבי מקטע אחד עומד על 31.5%, אחוז ההצלחה המדויק עבור מילים שהתפרסו על גבי 2 מקטעים עומד על 46% ואחוז ההצלחה המדויק עבור מילים שהתפרסו על גבי 3 מקטעים עומד על 63.3%.

נרצה לציין לעומת האמור לעיל, כי ניתן לראות לדוגמה בסימולציה מספר 5 שהצלחת האלגוריתם בזיהוי מילים באורך 3 טובה יותר מזיהוי מילים באורך 4. ככל הנראה זו אנומליה, כלומר מספר המילים בסרטון העקיבה באורך 4 אינו מאפשר ממוצע אמין ולכן גם אם לא התבצע זיהוי נכון רק על חלק קטן מהם, הדבר התבטא בירידת אחוזים גדולה. בנוסף, לכל אורך מילה ניתחנו את תוצאות האלגוריתם עבור ההצגה במהירויות השונות. נשים לב כי באופן כללי גם במקרה זה התוצאות טובות פחות עבור המהירות הגבוהה, גם כאן בחלק מהסימולציות ישנן אנומליות אך הן לא משמעותיות. נשים לב כי ככול שאורך המילה גדל כך ההפרש בין אחוזי ההצלחה ב-2 המהירויות השונות גדל גם הוא. נשער שככול שאורך המילה היה גדול יותר כך היו פחות מילים באורך זה בסרטון העקיבה ולכן כל זיהוי לא נכון של העקיבה פגע משמעותית יותר באחוזי ההצלחה.

#### 3. קביעת כיוון הסתכלות מדויק לעומת איכותי :

במקרה האיכותי "הצלחה" הוגדרה כאשר האלגוריתם קבע כיוון הסתכלות עד כדי קודמו ועוקבו של מספר המקטע הנכון. במקרה זה אחוזי ההצלחה עלו משמעותית בכלל הסימולציות. כלומר איכותיות האלגוריתם מצליח לקבוע כיוון הסתכלות נכון.

נרחיב כעת על הנאמר בתת פרק [3.4.3](#) ונאמר שמפני שבממוצע המרחק בין גבולות הגזרה שנקבעו בשלב הכיול היה 10 פיקסלים, לכל מקטע יוחס פיקסל בודד. כלומר המשמעות של קביעת כיוון

הסתכלות איכותי במקרה זה היא עד כדי הפרש של פיקסל בודד לכל כיוון. מדובר ברזולוציה שקשה מאוד לקבוע לפיה כיוון הסתכלות מדויק. בפועל תחום המרחקים בין גבולות הגזרה היה בין 5 ל 15 פיקסלים ומצאנו יחס ישר בין גודל המרחק הנ"ל לבין הצלחת האלגוריתם. כלומר ככול שהמרחק היה גדול יותר, כך תוצאות האלגוריתם היו טובות יותר מפני שלכל מקטע יוחסו יותר פיקסלים במוצע.

#### 4. השוואת מסנן מיצוע לעומת מסנן חציון:

ניתן לראות כי התוצאות פחות או יותר זהות עבור שני המסננים ולא נוכל לקבוע מי מהם עדיף. עובדה זו מאששת את איכות המסננים שבנינו בתת הפרק [3.5.3](#), שכן אם אין שוני גדול בין פעולת מיצוע לחציון הדבר מעיד על כך שלא היו ערכים קיצוניים בעת ביצוע הפעולות. כלומר המסיחים אכן סוננו כפי שציפינו.

#### 5. השפעת הזזת החוצצים על פני סרטון העקיבה בעת הניתוח:

באופן כללי ניתן לראות ירידה בתוצאות האלגוריתם כאשר מבצעים את פעולת ההזזה, בנוסף הירידה גדלה ככול שההזזה גדולה יותר. נסיק מכך שלא הייתה בעיית סנכרון משמעותית בעת ניתוח תוצאות האלגוריתם עבור הסרטון הרציף. כלומר הסנכרון שתואר בתת הפרק [3.5.1](#) הינו טוב.

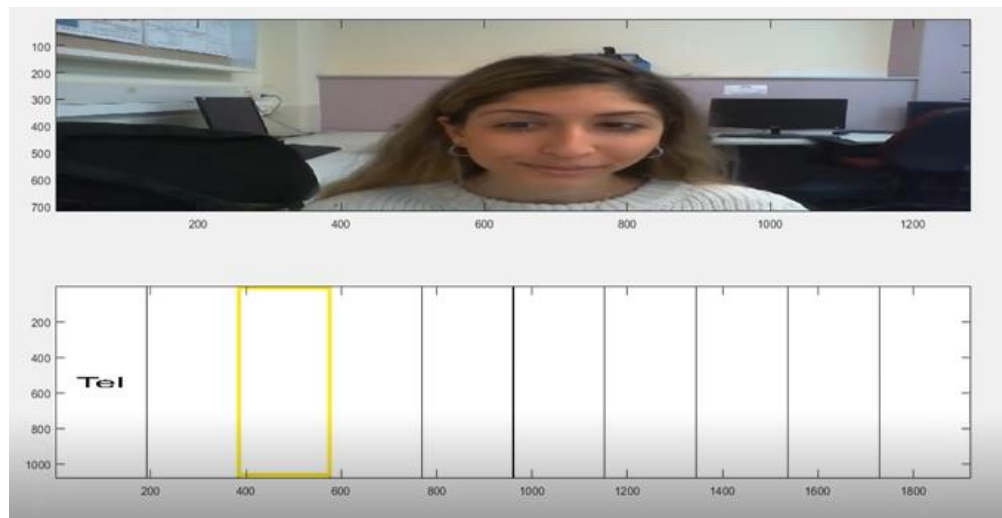
#### 6. השפעת קטימת החוצץ משני קצותיו:

ניתן לראות כי התוצאות פחות או יותר זהות כאשר מבצעים את פעולת הקטימה. בדומה לנקודה מספר 4, הדבר מאשש את כך שהמסיחים אכן סוננו כפי שציפינו.

#### 7. השוואת תוצאות עין ימין לעומת שמאל:

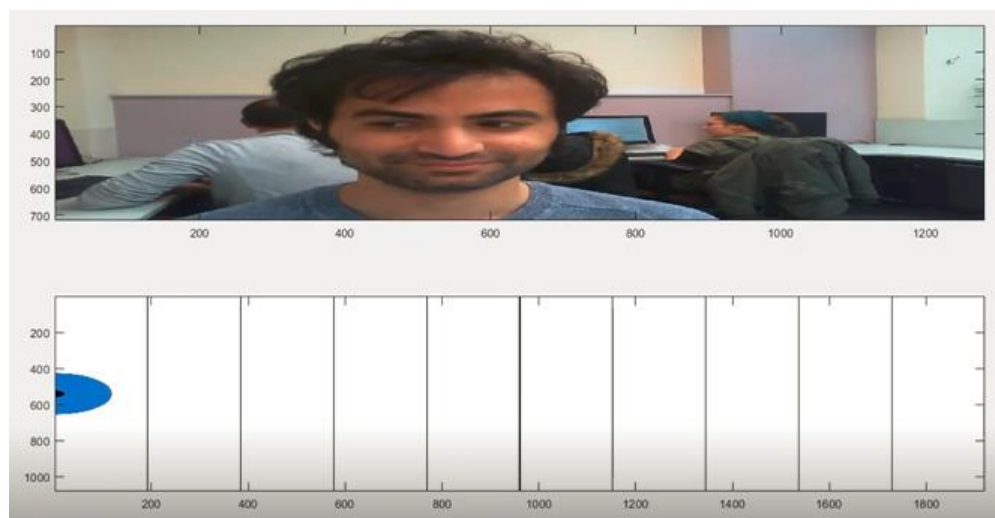
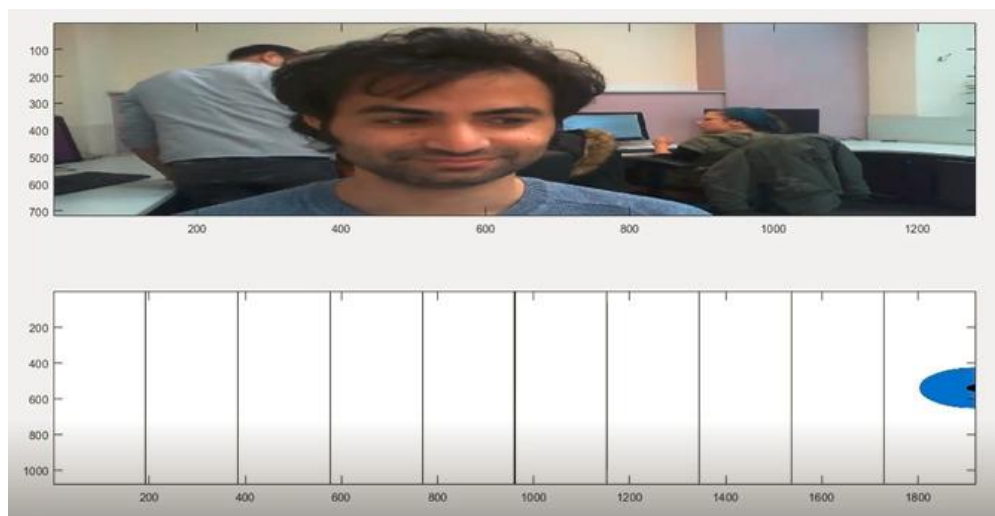
פרט לסימולציה מספר 3, עין ימין מובילה במעט על פני עין שמאל. נסביר זאת בשל העובדה ששלב ה TM היה מעט יותר מוצלח עבור עין ימין (נרחיב על אחוזי ההצלחה בהמשך). כמתואר בתת פרק [3.2](#), זאת מפני שהתבנית שהשתמשנו בה לצורך שלב זה נבנתה במקור בהתאם לעין ימין ושוקפה עבור עין שמאל.

יחד עם כל האמור לעיל, נציין כי בשלב ניתוח התוצאות הבחנו בתופעת הריצוד שבה המשתמש לא הסתכל לאן שנדרש ובעצם לא עקב אחר ההוראות. בפרק [5.1](#) הסברנו על הכלי הוויזואלי שפיתחנו בעזרתו ניתן לראות את סרטון המשתמש לצד הסרטון עליו הסתכל ובמקביל לאבחנה לאן האלגוריתם קבע שהסתכל. כאשר בחנו ויזואלית את הסרטונים שצילמנו בעזרת כלי זה, מצאנו כי בהרבה מהמקרים בהם לכאורה האלגוריתם נכשל למעשה הוא קבע כיוון הסתכלות נכון בהתאם למיקום האמיתי עליו הסתכל המשתמש, מיקום שלא תאם את מיקום המילה עליה היה אמור להסתכל. לכן אחוזי הדיוק הנמוכים התקבלו בחלק מהמקרים בשל אי עקיבה של המשתמש אחר ההוראות. ניתן לראות הדגמה באיור הבא, בו המשתמש לא מסתכלת אל עבר המילה שמופיעה על גבי המסך והאלגוריתם אכן מסמן נכונה בריבוע צהוב לאיזה מקטע היא מסתכלת בפועל.



איור 50 - הדגמה לתופעת הריצוד בה המשתמש אינו מסתכל אל המילה הרצויה והאלגוריתם קובע נכונה את כיוון ההסתכלות

תופעה נוספת בה הבחנו כאשר בדקנו ויזואלית את הסרטונים שצילמנו בעזרת הכלי, היא תזוזות ראש משמעותית בשלב הכיול ובשלבי בעקיבה אחר הטקסט המוצג על המסך. אי עמידה זו של המשתמש אחר ההוראות הובילה לפגיעה באחוזי הדיוק. הדגמה לכך ניתן לראות באיורים למטה:

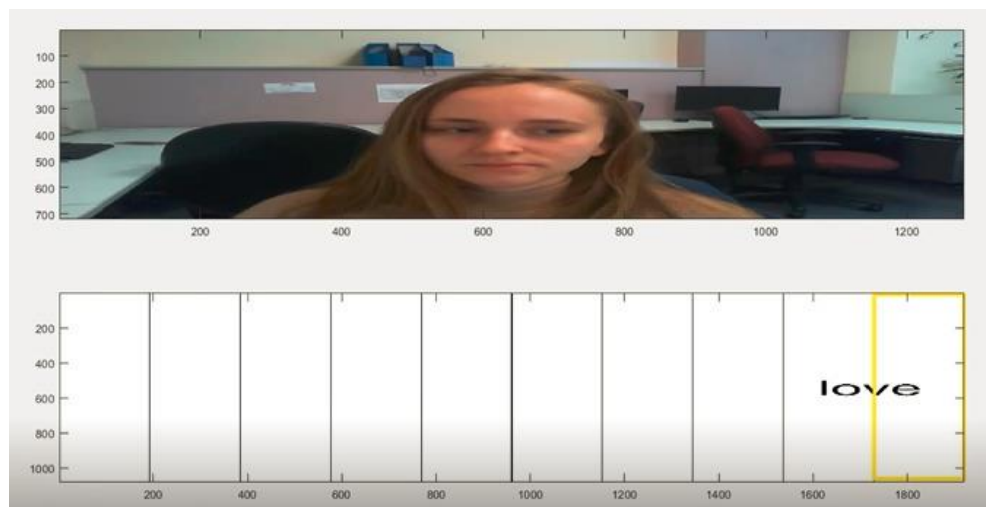


איור 51 - הדגמה לתזוזת ראש הפוגעת בשלב הכיול

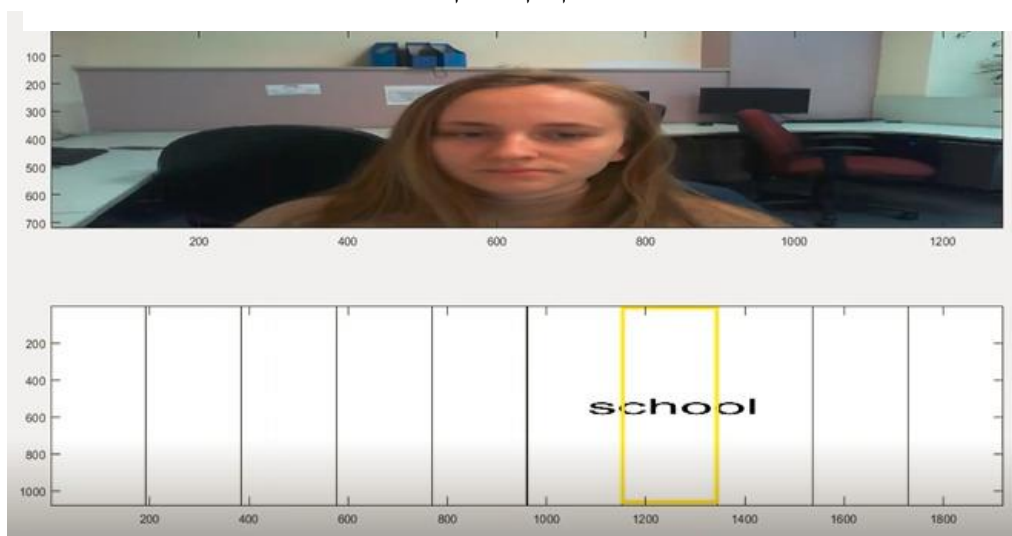
## 5.3. תוצאות סרטוני העקיבה האקראיים

| th5_mean_1-15_shift0 (1)    |                           |          |             |           |              |
|-----------------------------|---------------------------|----------|-------------|-----------|--------------|
| totalFrameNum               | 900.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 730.8    |             | 622.6     |              |
| all video                   |                           | 26.2     | 45.3        | 25.2      | 44.0         |
| part 1 (velocity 1 sec)     |                           | 32.8     | 50.7        | 30.7      | 49.5         |
| part 2 (velocity 0.5 sec)   |                           | 9.3      | 31.5        | 10.7      | 29.8         |
| word length 1               | all video                 | 18.6     | 42.5        | 19.3      | 42.5         |
|                             | part 1 (velocity 1 sec)   | 24.1     | 47.9        | 24.3      | 50.0         |
|                             | part 2 (velocity 0.5 sec) | 10.1     | 34.0        | 11.3      | 30.7         |
| word length 2               | all video                 | 32.6     | 44.3        | 26.7      | 40.7         |
|                             | part 1 (velocity 1 sec)   | 40.6     | 51.8        | 32.4      | 45.3         |
|                             | part 2 (velocity 0.5 sec) | 5.9      | 19.6        | 7.8       | 25.5         |
| word length 3               | all video                 | 41.2     | 54.7        | 40.6      | 52.4         |
|                             | part 1 (velocity 1 sec)   | 41.2     | 54.7        | 40.6      | 52.4         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 47.1     | 55.9        | 44.1      | 50.0         |
|                             | part 1 (velocity 1 sec)   | 47.1     | 55.9        | 44.1      | 50.0         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_mean_1-15_shift5 (2)    |                           |          |             |           |              |
| totalFrameNum               | 900.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 698.1    |             | 609.8     |              |
| all video                   |                           | 32.4     | 53.8        | 31.2      | 53.7         |
| part 1 (velocity 1 sec)     |                           | 39.3     | 60.2        | 36.9      | 59.6         |
| part 2 (velocity 0.5 sec)   |                           | 14.5     | 37.4        | 16.6      | 38.4         |
| word length 1               | all video                 | 23.9     | 49.5        | 23.0      | 51.3         |
|                             | part 1 (velocity 1 sec)   | 29.1     | 56.1        | 27.5      | 59.1         |
|                             | part 2 (velocity 0.5 sec) | 15.5     | 39.1        | 16.0      | 39.1         |
| word length 2               | all video                 | 38.9     | 54.3        | 36.7      | 52.0         |
|                             | part 1 (velocity 1 sec)   | 47.6     | 61.8        | 41.8      | 57.1         |
|                             | part 2 (velocity 0.5 sec) | 9.8      | 29.4        | 19.6      | 35.3         |
| word length 3               | all video                 | 49.4     | 65.3        | 51.2      | 63.5         |
|                             | part 1 (velocity 1 sec)   | 49.4     | 65.3        | 51.2      | 63.5         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 58.8     | 70.6        | 44.1      | 58.8         |
|                             | part 1 (velocity 1 sec)   | 58.8     | 70.6        | 44.1      | 58.8         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_mean_1-15_shift10 (3)   |                           |          |             |           |              |
| totalFrameNum               | 900.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 702.9    |             | 615.5     |              |
| all video                   |                           | 40.0     | 66.1        | 39.2      | 63.3         |
| part 1 (velocity 1 sec)     |                           | 45.9     | 70.3        | 44.7      | 67.8         |
| part 2 (velocity 0.5 sec)   |                           | 24.9     | 55.0        | 25.3      | 51.6         |
| word length 1               | all video                 | 27.6     | 59.5        | 28.9      | 60.0         |
|                             | part 1 (velocity 1 sec)   | 29.9     | 63.4        | 31.8      | 64.7         |
|                             | part 2 (velocity 0.5 sec) | 23.9     | 53.4        | 24.4      | 52.5         |
| word length 2               | all video                 | 49.3     | 71.5        | 46.6      | 61.1         |
|                             | part 1 (velocity 1 sec)   | 55.3     | 74.1        | 51.8      | 65.3         |
|                             | part 2 (velocity 0.5 sec) | 29.4     | 62.7        | 29.4      | 47.1         |
| word length 3               | all video                 | 64.1     | 78.8        | 61.8      | 75.3         |
|                             | part 1 (velocity 1 sec)   | 64.1     | 78.8        | 61.8      | 75.3         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 82.4     | 85.3        | 64.7      | 76.5         |
|                             | part 1 (velocity 1 sec)   | 82.4     | 85.3        | 64.7      | 76.5         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_mean_5-11_shift0 (7)    |                           |          |             |           |              |
| totalFrameNum               | 420.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 339.8    |             | 286.9     |              |
| all video                   |                           | 27.7     | 47.7        | 26.7      | 47.1         |
| part 1 (velocity 1 sec)     |                           | 34.2     | 54.4        | 32.9      | 52.8         |
| part 2 (velocity 0.5 sec)   |                           | 10.7     | 30.4        | 10.7      | 32.2         |
| word length 1               | all video                 | 20.9     | 46.1        | 20.4      | 46.2         |
|                             | part 1 (velocity 1 sec)   | 27.0     | 54.0        | 25.9      | 53.7         |
|                             | part 2 (velocity 0.5 sec) | 11.3     | 33.6        | 11.8      | 34.5         |
| word length 2               | all video                 | 32.6     | 47.5        | 31.2      | 45.7         |
|                             | part 1 (velocity 1 sec)   | 40.0     | 57.1        | 38.8      | 52.9         |
|                             | part 2 (velocity 0.5 sec) | 7.8      | 15.7        | 5.9       | 21.6         |
| word length 3               | all video                 | 42.9     | 52.4        | 40.6      | 50.6         |
|                             | part 1 (velocity 1 sec)   | 42.9     | 52.4        | 40.6      | 50.6         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 41.2     | 55.9        | 41.2      | 52.9         |
|                             | part 1 (velocity 1 sec)   | 41.2     | 55.9        | 41.2      | 52.9         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_median_1-15_shift0 (4)  |                           |          |             |           |              |
| totalFrameNum               | 900.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 730.8    |             | 622.4     |              |
| all video                   |                           | 26.2     | 46.5        | 26.1      | 44.9         |
| part 1 (velocity 1 sec)     |                           | 32.6     | 52.7        | 32.4      | 50.5         |
| part 2 (velocity 0.5 sec)   |                           | 9.7      | 30.4        | 10.0      | 30.4         |
| word length 1               | all video                 | 19.0     | 44.0        | 20.1      | 44.4         |
|                             | part 1 (velocity 1 sec)   | 24.1     | 50.8        | 25.4      | 51.9         |
|                             | part 2 (velocity 0.5 sec) | 10.9     | 33.2        | 11.8      | 32.8         |
| word length 2               | all video                 | 32.1     | 47.1        | 28.1      | 40.7         |
|                             | part 1 (velocity 1 sec)   | 40.6     | 55.9        | 35.9      | 47.1         |
|                             | part 2 (velocity 0.5 sec) | 3.9      | 17.6        | 2.0       | 19.6         |
| word length 3               | all video                 | 40.6     | 53.5        | 41.8      | 51.8         |
|                             | part 1 (velocity 1 sec)   | 40.6     | 53.5        | 41.8      | 51.8         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 47.1     | 52.9        | 44.1      | 47.1         |
|                             | part 1 (velocity 1 sec)   | 47.1     | 52.9        | 44.1      | 47.1         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_median_1-15_shift5 (5)  |                           |          |             |           |              |
| totalFrameNum               | 900.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 698.1    |             | 609.8     |              |
| all video                   |                           | 32.8     | 53.2        | 30.9      | 53.0         |
| part 1 (velocity 1 sec)     |                           | 38.9     | 59.2        | 36.6      | 59.2         |
| part 2 (velocity 0.5 sec)   |                           | 17.0     | 37.7        | 15.9      | 37.0         |
| word length 1               | all video                 | 24.3     | 50.0        | 23.0      | 51.6         |
|                             | part 1 (velocity 1 sec)   | 28.6     | 56.1        | 26.7      | 59.4         |
|                             | part 2 (velocity 0.5 sec) | 17.6     | 40.3        | 17.2      | 39.5         |
| word length 2               | all video                 | 38.9     | 53.4        | 36.2      | 50.2         |
|                             | part 1 (velocity 1 sec)   | 46.5     | 61.8        | 44.1      | 57.6         |
|                             | part 2 (velocity 0.5 sec) | 13.7     | 25.5        | 9.8       | 25.5         |
| word length 3               | all video                 | 50.0     | 62.4        | 48.8      | 61.2         |
|                             | part 1 (velocity 1 sec)   | 50.0     | 62.4        | 48.8      | 61.2         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 58.8     | 64.7        | 47.1      | 55.9         |
|                             | part 1 (velocity 1 sec)   | 58.8     | 64.7        | 47.1      | 55.9         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_median_1-15_shift10 (6) |                           |          |             |           |              |
| totalFrameNum               | 900.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 702.9    |             | 615.4     |              |
| all video                   |                           | 43.2     | 67.6        | 43.2      | 66.2         |
| part 1 (velocity 1 sec)     |                           | 48.5     | 71.0        | 49.3      | 69.7         |
| part 2 (velocity 0.5 sec)   |                           | 29.4     | 58.8        | 27.3      | 57.1         |
| word length 1               | all video                 | 30.6     | 60.8        | 33.2      | 61.4         |
|                             | part 1 (velocity 1 sec)   | 33.7     | 63.4        | 37.4      | 65.2         |
|                             | part 2 (velocity 0.5 sec) | 25.6     | 56.7        | 26.5      | 55.5         |
| word length 2               | all video                 | 58.4     | 75.1        | 50.7      | 69.7         |
|                             | part 1 (velocity 1 sec)   | 61.8     | 77.1        | 56.5      | 71.2         |
|                             | part 2 (velocity 0.5 sec) | 47.1     | 68.6        | 31.4      | 64.7         |
| word length 3               | all video                 | 61.8     | 78.8        | 64.1      | 76.5         |
|                             | part 1 (velocity 1 sec)   | 61.8     | 78.8        | 64.1      | 76.5         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 79.4     | 85.3        | 70.6      | 76.5         |
|                             | part 1 (velocity 1 sec)   | 79.4     | 85.3        | 70.6      | 76.5         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| th5_median_5-11_shift0 (8)  |                           |          |             |           |              |
| totalFrameNum               | 420.0                     | left eye | left eye +1 | right eye | right eye +1 |
| goodFrameNum                |                           | 339.8    |             | 286.9     |              |
| all video                   |                           | 28.4     | 48.2        | 27.6      | 47.4         |
| part 1 (velocity 1 sec)     |                           | 34.5     | 55.1        | 33.4      | 53.3         |
| part 2 (velocity 0.5 sec)   |                           | 12.5     | 30.4        | 12.5      | 32.2         |
| word length 1               | all video                 | 22.4     | 46.1        | 21.2      | 46.4         |
|                             | part 1 (velocity 1 sec)   | 27.5     | 53.5        | 25.7      | 53.5         |
|                             | part 2 (velocity 0.5 sec) | 14.3     | 34.5        | 14.3      | 35.3         |
| word length 2               | all video                 | 31.7     | 49.3        | 32.6      | 46.2         |
|                             | part 1 (velocity 1 sec)   | 40.0     | 60.6        | 41.2      | 54.7         |
|                             | part 2 (velocity 0.5 sec) | 3.9      | 11.8        | 3.9       | 17.6         |
| word length 3               | all video                 | 42.4     | 53.5        | 40.6      | 52.4         |
|                             | part 1 (velocity 1 sec)   | 42.4     | 53.5        | 40.6      | 52.4         |
| אין מילים באורך 3 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |
| word length 4               | all video                 | 44.1     | 52.9        | 44.1      | 50.0         |
|                             | part 1 (velocity 1 sec)   | 44.1     | 52.9        | 44.1      | 50.0         |
| אין מילים באורך 4 בחלק 2    | part 2 (velocity 0.5 sec) | NaN      | NaN         | NaN       | NaN          |

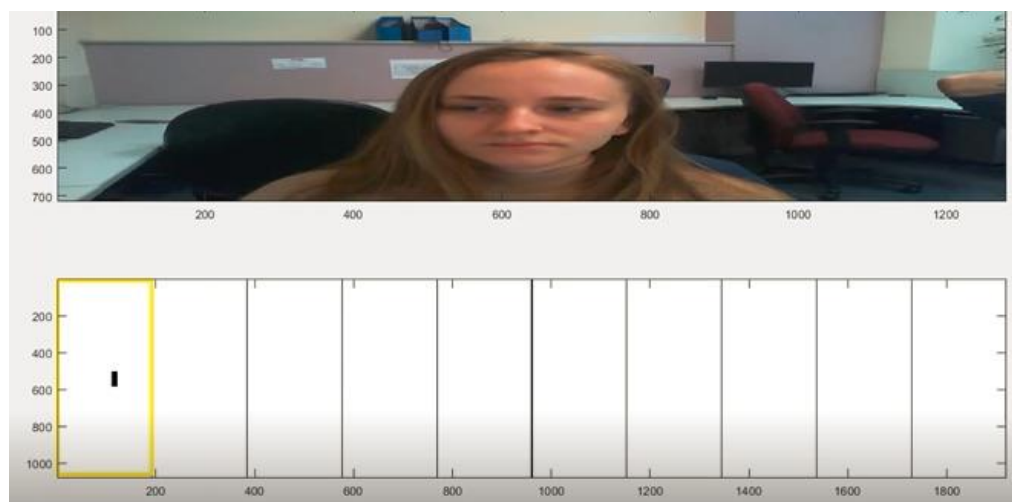
טבלה 3 - תוצאות סרטוני העקיבה האקראיים



איור 52 - הדגמה לסרטון עקיבה אקראי: המילה הראשונה



איור 53 - הדגמה לסרטון עקיבה אקראי: המילה השנייה



איור 54 - הדגמה לסרטון עקיבה אקראי: המילה השלישית

באיורים לעיל ניתן לראות המחשה לסרטון עקיבה אקראי בו המשתמשת עוקבת אחר המילים המופיעות על גבי המסך. ניתן להבחין בכך שעיניה של המשתמשת מסתכלות אל המילה הרצויה ואכן תוצאות האלגוריתם המסומנת בריבוע צהוב קובעת נכונה את כיוון ההסתכלות.

נבחן את תוצאות הסרטון האקראי במספר היבטים:

ראשית, נשים לב כי ללא שימוש בפרמטר ההזזה, באופן כולל תוצאות האלגוריתם נמוכות משמעותית עבור העקיבה אחר הסרטון האקראי לעומת הרציף. כאשר נשווה את תנאי סימולציה ממספר 4 עבור שני המקרים נקבל כי במקרה של הסרטון הרציף האלגוריתם הגיע ל 45% הצלחה, לעומת המקרה של הסרטון האקראי בו השיג 26% הצלחה. נסביר זאת בכך שלמשתמשים היה קשה לבצע עקיבה אקראית ולעקוב אחר המילים המוצגות במהירויות הנתונות. נשער כי במקרה זה זמן התגובה של המשתמש שלא היה מודע למיקום המילה הבאה (בניגוד למקרה הקודם בו התבצעה קריאה רציפה) התבטא בתזוזות אישון מהירות בין חלקי המסך שהובילו לאי יציבותו. בנוסף המרחק הממוצע שהאישון היה צריך לעבור ממילה למילה היה גדול יותר, זה מתבטא ביותר מסגרות "מבוזבזות" בהן האישון נמצא בתנועה ולא יציב ולכן ירידה באמינות ובביצועים. ניתן אף להגיד כי בחלק מהמקרים האישון ריצד בחיפוש אחר המילה.

לעומת זאת, כאשר עשינו שימוש בפרמטר ההזזה נכחנו לגלות כי תוצאות האלגוריתם במקרה העקיבה אחר הסרטון האקראי דומות לתוצאותיו במקרה העקיבה אחר סרטון הרציף. כאשר נבחן את תנאי סימולציה מספר 4 עבור מקרה העקיבה הרציף נקבל כאמור 45% הצלחה, כאשר נבחן את תנאי סימולציה מספר 6 עבור מקרה העקיבה האקראי נקבל 43% הצלחה.

#### 1. מהירות הצגת המילים על המסך:

גם במקרה זה, ניתן לראות עבור כל הסימולציות, שהתוצאות נמוכות יותר עבור המהירות הגבוהה לעומת הנמוכה. במקרה זה הדבר אף מתבטא בצורה משמעותית יותר.

#### 2. אורך המילה ביחידות של מקטע:

גם במקרה זה, באופן כללי ובהפרדה למהירויות שונות, ככול שאורך המילה גדל כך תוצאות האלגוריתם טובות יותר. הדבר הגיוני מפני שיש יותר מרחב בקביעת כיוון ההסתכלות, כלומר הוא יכול להתפרס על פני מספר מקטעים.

#### 3. קביעת כיוון הסתכלות מדויק לעומת איכותי:

מעבר לנאמר על נושא זה בניתוח תוצאות הסרטונים הרציפים, ובהתאם לכך שהסקנו עקב התוצאות הנמוכות משמעותית שאישון המשתמש לא יציב כיוון שהוא לא מספיק לעקוב אחר המילים המוצגות, נוכל לומר כי הגדלת מרחב התוצאות הנכונות עבור קביעת כיוון ההסתכלות מפצה על העקיבה האיטית של המשתמש ותורם לשיפור התוצאות.

#### 4. השוואת מסנן מיצוע לעומת מסנן חציון:

ניתן לראות כי התוצאות פחות או יותר זהות עבור שני המסננים ולא נוכל לקבוע מי מהם עדיף. עובדה זו מאששת את איכות המסננים שבנינו בתת הפרק [3.5.3](#), שכן אם אין שוני גדול בין פעולת

מיצוע לחציון הדבר מעיד על כך שלא היו ערכים קיצוניים בעת ביצוע הפעולות. כלומר המסיחים אכן סוננו כפי שציפינו.

5. השפעת הזזת החוצצים על פני סרטון העקיבה בעת הניתוח:

באופן כללי ניתן לראות עלייה בתוצאות האלגוריתם כאשר מבצעים את פעולת ההזזה, בנוסף העלייה גדלה ככול שההזזה גדולה יותר. בהתבסס על תוצאות ניתוח הסרטון הרציף נסיק שגם כעת אין בעיית סנכרון, שכן שיטת הסנכרון זהה בשני המקרים. נוכל לשער שאכן זמן התגובה של המשתמש היה גדול יותר, כלומר המשתמש "פיגר" בהזזת עיניו לעומת הופעת המילה, בגלל שהיא הופיעה במקום לא צפוי, ולכן הסטה של החוצצים קדימה מפצה על הפיגור הזה.

6. השפעת קטימת החוצץ משני קצותיו:

ניתן לראות כי התוצאות פחות או יותר זהות כאשר מבצעים את פעולת הקטימה. בדומה לנקודה מספר 4, הדבר מאשש את כך שהמסיחים אכן סוננו כפי שציפינו.

7. השוואת תוצאות עין ימין לעומת שמאל:

ניתן לראות שתוצאות שתי העיניים יחסית זהות. במקרה זה לא יכולנו להסיק האם אחת מהעניים טובה במעט מהשנייה.



## 5.4. תוצאות כלליות

נציין מספר אחוזי הצלחה על ניתוח מאגר הסרטונים שלנו המכיל 46 סרטונים. (55,800 מסגרות המוכלות ב 3720 חוצצים לכל עין).

נציין את אחוזי ההצלחה של שלבי האלגוריתם, יחד ולחוד :

1. אלגוריתם Viola & Jones : 97.65% מכלל המסגרות.
2. TM עבור עין ימין : 98.5% מהמסגרות שעברו את שלב 1.
3. TM עבור עין שמאל : 97.7% מהמסגרות שעברו את שלב 1.
4. אחוז המסגרות שעברו את שלבים 1-3 בהצלחה ולאחר מכן עברו את המסננים שתוארו בתת פרק 3.5.3 : 79.3% מהמסגרות שעברו את שלבים 1-3.
5. אחוז המסגרות שעברו את שלבים 1-3 בהצלחה ולאחר מכן נפסלו בשל הסף הגלובלי : 5.2% מהמסגרות שעברו את שלבים 1-3.
6. לפי 4,5 נקבל כי אחוז המסגרות שעברו את שלבים 1-3 בהצלחה, לא נפסלו על ידי הסף הגלובלי וסוננו בהמשך התהליך : 15.5% מהמסגרות שעברו את שלבים 1-3.

נשים לב כי בסך הכול כל אחד מהשלבים 1-4 בפני עצמו מצליח להניב תוצאות טובות. בסופו של דבר, כאשר נתחשב בכל ארבעת השלבים, נקבל כי בממוצע עבור כל סרטון 76% מהמסגרות שימשו בפועל לקביעת כיוון ההסתכלות.

כעת, כפי שנאמר בתת הפרק 3.1, ההנחה שאחוזי הצלחת אלגוריתם Viola & Jones גבוהים אכן הייתה נכונה. בנוסף, כפי שכבר ציינו, ההבדל הקטן בין אחוזי הצלחת שלב ה TM בין שתי העיניים נובע מכך שהתבנית שהשתמשנו בה לצורך שלב זה נבנתה במקור בהתאם לעין ימין ושוקפה עבור עין שמאל. כמו כן, נציין כי האחוז היחסית גדול של המסגרות שעברו את המסננים, בנוסף לעובדה שהתקבלו תוצאות דומות עבור מסנן מיצוע וחציון (כפי שכבר נאמר בתת הפרק הקודם), מעידים על אמינות האלגוריתם. עם זאת, נאמר כי לא מצאנו קשר סטטיסטי בין אחוז המסגרות שבהסתמך עליהן נקבע כיוון ההסתכלות לבין טיב הקביעה בכל סימולציה בנפרד.

נציין בנוסף מספר תוצאות שחילצנו בשלב ניתוח האלגוריתם :

1. ההבדל בין תוצאת קביעת כיוון ההסתכלות של שתי העיניים היה 0.53 מקטעים בממוצע.
2. ההבדל המקסימלי בין קביעת כיוון ההסתכלות של שתי העיניים היה 3 מקטעים בממוצע.
3. אחוז הפעמים בו עין אחת קבעה כיוון הסתכלות כלשהו והשניה נכשלה עומד על 28% מתוך 3720 פעמים (מספר החוצצים לכל עין).
4. אחוז הפעמים בו שתי העיניים נכשלו בקביעת כיוון ההסתכלות עומד על 3% מתוך 3720 פעמים (מספר החוצצים לכל עין).

## 6. סיכום

### 6.1. מסקנות

ראשית, נשמח לציין כי הצלחנו לפתח אלגוריתם זיהוי ועקיבה מוצלח המבוסס על אלגוריתמי ניתוח ועיבוד תמונה קלאסיים. כל שלב בפני עצמו הניב תוצאות טובות כמו גם חיבורם לכדי אלגוריתם שלם.

אנו סבורים כי נקודת התורפה של האלגוריתם היא רזולוציית המסגרות איתה עבדנו. כפי שכבר ציינו בתת פרק [3.4.3](#), בממוצע לכל מקטע במסך יוחס פיקסל בודד. כלומר בכדי להגיע לתוצאות מדויקות וטובות נדרש להגיע לדיוק של פיקסל בודד במסגרות אותן עיבדנו. אנו סבורים כי במידה ורזולוציית המסגרות איתה עבדנו הייתה טובה יותר התוצאות היו משתפרות משמעותית. כלומר ככול שנצליח לייחס יותר פיקסלים למקטע כך התוצאות ישתפרו. זאת בעיקר על סמך מדד ה"הצלחה" שהוגדר עד כדי מקטע קודם או עוקב לזה הרצוי, ובהסתמך על שמצאנו שכול שאורך המילה גדל כך גם תוצאות העקיבה טובות יותר. (שכן ככול שאורך המילה גדל הדבר מדומה לשיפור הרזולוציה מפני שיותר פיקסלים מיוחסים אליה). בנוסף מתוך האמור לעיל, נסיק כי תמיכה בפחות מ-10 מקטעים, למשל רק ב-5 תסייע להשיג אחוזי דיוק טובים יותר. נציין כי יתכן ובעבודה עם רזולוציה טובה יותר נוכל להסיר את ההנחה שראש המשתמש צריך להיות דומיננטי בתמונה.

נוסף לכך, בשל האופטיקה של מצלמת הרשת, נאלצנו לקבוע תנאי תאורה אחידים, שכן הדבר מאוד משפיע כאשר משתמשים במצלמות רשת. אנו מסיקים כי במידה והיינו משתמשים במצלמה איכותית יותר היינו יכולים להקל בדרישת תנאי התאורה האחידים.

כעת, כפי שניתן לראות בפרק [5](#), מהירות הצגת המילים הינה פקטור משמעותי. ככול שהיא גדלה אחוזי דיוק האלגוריתם היו נמוכים יותר. נשער שהסיבה לכך היא אי יציבות האישון על המילה עליה המשתמש נדרש להסתכל שלעיתים אף גורמת לריצוד אצלו. ניתן לראות דמיון בין מקרה זה לבין סרטוני העקיבה האקראיים גם בהם קיבלנו את תופעת ריצוד האישון (בשתי המהירויות). כלומר בשני המקרים תופעת הריצוד גרמה לפגיעה משמעותית בתוצאות.

בהקשר לנקודה זו, חשוב לנו לציין כי תופעת הריצוד פוגעת ביכולת שלנו לנתח את ביצועי האלגוריתם נכונה, שכן במידה והמשתמש לא הסתכל לאן שנדרש נקבל בוודאות "כישלון" בעת ניתוח התוצאות. נדגיש כי במקרה זה של ריצוד אין הדבר מעיד על כשלון אמיתי של האלגוריתם שכן הוא קבע כיוון הסתכלות שנחשב לשגוי אך ורק בשל תופעת הריצוד שגרמה לכך שהמשתמש לא הסתכל לאן שנדרש. מקרים אלו גרמו לירידה לא מוצדקת בביצועי האלגוריתם כפי שנמדדו בפרק [5](#) וניתן לראות לכך הדגמה באיור 49 בפרק זה.

בכל הנוגע להנחות האלגוריתם שתוארו בתחילת פרק 3, מצאנו את כולן כהכרחיות. אכן כאשר ראש המשתמש לא היה דומיננטי, ישר ויציב במהלך כל העקיבה קיבלנו תוצאות ירודות. בפרט כאשר התבצעו תזוזות ראש משמעותיות בעת שלב הכיול, ניתוח העקיבה הניב תוצאות ירודות משמעותיות. ניתן לראות הדגמה לכך באיור 50 בתת פרק 5.1.

## 6.2. עבודה עתידית

כחלק מדרישות הפרויקט ביצענו ניוון הבעיה לדו-ממד. על מנת לבצע עקיבה איכותית ומדויקת יותר נדרש להתייחס לבעיה התלת-ממדית המלאה המביאה בחשבון גם את מנח הראש. נציין כי בפועל מעטים המשתמשים שאכן לא הזיזו את ראשם כלל בעת העקיבה (גם אם תזוזות ראש מעטות), דבר זה שגרם לירידה באחוזי הדיוק של האלגוריתם. על כן, נוכל לשער כי התחשבות בהטיית הראש תסייע בשיפור התוצאות. כמו כן, התייחסות לבעיה המלאה תעזור גם בהנגשת האלגוריתם לילדים כפי שתוכנן במקור.

בעת קביעת כיוון ההסתכלות התבססנו על קירוב לינארי כפי שהסברנו בתת פרק 3.4.3. ביצענו התאמה בצורה חד ערכית בין מספר המקטע שמכיל את האישון בתמונת העין שווה לבין מספר המקטע עליו מסתכל המשתמש במסך.

בפועל העין היא לא משטח דו-ממדי כפי שניסינו לדמות במטרה לפשוט הבעיה. על מנת להגיע לתוצאות מדויקות יותר יש ליצור מודל גיאומטרי המתחשב בכך אשר לא מבצע חלוקה לינארית.

כמו כן, בכל הנוגע להתפלגות המילים בסרטוני העקיבה לפי מהירות ואורך, נציע כשיפור עתידי לקבוע את התפלגות זו להיות אחידה כך שיהיו מספיק מילים מכל סוג.

לסיום, לצד כל הנאמר לעיל, נציע כמחשבה לשיפור עתידי נוסף שימוש בטכניקות למידה עמוקה לצד כלי ניתוח ועיבוד התמונה הקלאסיים בהם השתמשנו. זאת במטרה לשפר הן את תוצאות האלגוריתם והן את סיבוכיות הזמן שלו.

## 7. רשימת מקורות

1. P. Viola and M.J. Jones, "Robust Real-Time Face Detection", Int. J. Comput. Vision 57, 2 (May 2004), 137-154.
2. J. Y. Kaminski, D. Knaan, and A. Shavit, "Single image face orientation and gaze detection", Mach. Vision Appl. 21, 1 (October 2009), 85-98
3. Michael Elad, "Introduction to Image Processing", Technion, 1999.
4. [https://en.wikipedia.org/wiki/Hough\\_transform](https://en.wikipedia.org/wiki/Hough_transform)