



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



# החלפת רקע בצלילום וידאו

מאת :

**שלו דוד ינון נוסבאום**

בנהנחיתת :

**יעקב צירני**

chorf-אביב 2007



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



# Video background substitution

By:

**Inon Nussbaum      Shalev David**

Supervised by:

**Jacob Chernoi**

Winter-Spring 2007



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



## תודות

ברצוננו להודות למנהל הפרויקט, יעקב צירני, על הדרכתו במהלך הפרויקט.  
תודה רבה גם לצוות המעבדה לחקור הראייה ומדעי התמונה, בראשות יוחנן ארז, על  
 נכונותם לעוזר ולסייע בכל אשר היה דרוש לנו לביצוע הפרויקט.



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



## תוכן העניינים

6	תקציר
7	ABSTRACT
8	מבוא
8	רקע
9	הגדרות
10	מבנה העבודה
10	סביר כללי
10	תיאור סכמטי
11	מציאות המסלח
12	סביבת העבודה
13	רקע תיאורתי
13	"התאמת בלוקים" (BLOCK MATCHING)
13	הקדמה ומוטיבציה
13	שערוך תנועה
16	אלגוריתם לשערוך התנועה
21	אלגוריתם INTELLIGENT SCISSORS
21	רקע ומוטיבציה
22	תיאור דאלגוריתם
28	החסרת רקע (BACKGROUND SUBTRACTION)
28	הקדמה
28	מייצוע הרקע
29	GMM
34	שיטת הרקע הרץ (Running Background)
36	עיבוד תמונה
36	מרחבי צבעים
36	מרחב RGB
37	מרחב HSV
42	מרחב YCbCr
43	ב unify ה-INTERLACE
45	פעולות על תמונת מסכה
45	פועלות לצורכי ניקוי רעים ומידע מיתר
45	פועלות לצוריך השלמת מידע חסר
54	עבודה מעשית וניתוח תוצאות
54	עקביה תוך שימוש בהתאם בלוקים
54	בסיס האלגוריתם
62	פיקורת וסינון הבלוקים הנעקבים
66	SHIPOR המסכה באמצעות אלגוריתם ה-INTelligent SCISSORS
66	סביר הרעיון
66	האלגוריתם שפותח



**הטכניון - מכון טכנולוגי לישראל**  
**הפקולטה להנדסת חשמל**  
**המעבדה לחקור הראייה ומדעי התמונה**



76.....	מציאת המסכה תוך שימוש בשיטת הרקע הרץ .....
87.....	אגישה שנבחנה בפרויקט .....
87.....	תרשים בלוקים .....
89.....	מיצאות .....
94.....	הנחות יסוד .....
<b>95.....</b>	<b>סיכום ומסקנות .....</b>
<b>96.....</b>	<b>הצעות שיפור ומחקר עתידי .....</b>
<b>97.....</b>	<b>רשימת מקורות .....</b>



## תקציר

מטרת הפרויקט היא להחליף רקע בצילומים וידיאו,(Clipper) כולם לבזבז את האובייקט אותו מעוניינים להשאיר בצילומים, ולהמיר את סביבתו בתמונה.

כדי לבזבז את אותו אובייקט יש לסוג כל פיקסל כרקע אחורי או קדמי, כאשר הרקע האחורי יוחלף בתמונה, והרקע הקדמי הוא האובייקט הנדון.

האלגוריתם המוצע בפרויקט משלב בין למידת רקע בשיטת "רקע רץ" (Running background), שערוך לא סיבתי המבוסס על התנהלות האובייקט סביבה מרכז המסה שלו, וביצוע פעולות מורפולוגיות. ביצוע האלגוריתם מתבצע בשלושה מרחבי צבעים שונים – RGB, HSV, YCBCR.

האלגוריתם מומש בשפת התכנות MATLAB. הוא קולט סרט וידאו כלשהו מסוג AVI ותמונה. פلت התוכנית הוא סרט וידאו מסוג AVI אשר בו התמונה מהויה תחליפן רקע של האובייקט הסרט הוידיאו הנקלט.



## Abstract

The purpose of this project is to substitute a video background, meaning to isolate the object which one would like to keep, and replace its environment with a picture.

In order to isolate this object, there is a need to classify every pixel to one of two sets: background or foreground. Background pixels will be substituted by the picture, while the foreground pixels are kept.

The proposed algorithm in the project combines background learning using the “Running Background” method, a non casual prediction based on the behavior of the object around its center of mass, and performing morphological operations. The execution of the algorithm occurs in three different color dimensions – RGB, HSV and YCBCR.

The algorithm was implemented in MATLAB. It receives an AVI type video stream and a picture. The output of the program is an AVI video stream file in which the picture substitutes the object’s original background.



## מבוא

### רקע

"החלפת רקע" (Background Substitution) בסרט וידאו הינה בעיה מורכבת שטרם נמצא לה פתרון מושלם אשר לוקח בחשבון את תנאי הסביבה (גשם, שלג, רוחות), איזות הצילום (רעים אדיטיביים למיניהם), סיבוכיות הרקע, שינוי תוארה בסרט וצדומה. קיימות אפליקציות רבות שפותחו למטרות גילוי עצמים נעים תוך שימוש בשיטות של "החסרת רקע" (Background Subtraction). חוסר הסטציונריות של הרקע מהוות את הבעיה המרכזיית ברוב האפליקציות שפותחו עד כה בנושא. שיטות של החסרת רקע משערכות מודל של רקע ומבצעות עדכון שלו בכל פריים כך שינויים סיבתיים ברקע נלקחים בחשבון. גילוי אובייקטים נעים בפריים מסוימים מתבצע ע"י החסירה בין הפריים לבין הרקע המשוערך על מנת לקבל תמונה הפרשים אשר ממנה ניתן לחלק את העצמים אשר נמצאים בתנועה ואינם מהווים חלק מהרקע. באפליקציות אלו נעשה שימוש לצורך עקיבה אזרחית (בעיקר עקיבה אחר בני-אדם ומכוניות בכבישים ובמעבר חצייה) ממצלמות נייחות שהוצבו באזור. שיטות אלו מניחות כי העצמים נמצאים בתנועה מתמדת וחולפים על פני המצלמה, וכי ברקע חלים שינויים קלים וקראיים בלבד כגון תנעות עליים הנידפים ברוח וגלים שקטים. במרבית הטכניקות הללו כאשר הדמות משתנה במקומות ומפסיקת לנוע היא "נבלעת" והאפליקציה מחשיבה אותה כרקע.

תכלית פרויקט זה היא לזהות את הדמות גם כאשר היא איננה בתנועה מתמדת ולמנוע ממנה "להיבלו" ברקע כאשר היא נייחת, ולכן טכניקה זאת לבדה אינה פתרון הולם. ישנן אפליקציות שימושה בטכניקות אחרות אולם אפליקציות אלו שואבות את המידע על העולם משתי מצלמות וכן יוצרות ראיית סטריאו (ראייה תלת ממדית) המאפשרת להן להבחן יותר קלות בשכבות שונות ולהבדיל בין הדמיות לבין הרקע.



בפרויקט זה המידע על העולם נשאב מצלמה אחת בלבד, ומכיון שכך התמונה המתקבלת היא תמונה דו-ממדית ולא ניתן להשליך אליה טכניקות של ראיית סטריאו.

### הגדירות

תמונה הפרשים – הערך המוחלט של תמונה המתקבלת מחיסור בין שתי תמונות כלשהן.

מסכה – תמונה בינהרית של שחור לבן (מיוצגת ע"י מטריצת אפסים ואחדים) אשר בה פיקסלים שחורים (בעל ערך '0') מייצגים "רקע אחורי" (background) ופיקסלים לבנים (בעל ערך '1') מייצגים "רקע קדמי" (foreground), דהיינו העצם אשר את רקעו מעוניינים להחלף.

מסכה סופית – מסכה המשמשת לייצרת סרט הפלט.

נתיב מסכה – תמונה בינהרית של שחור לבן (מיוצגת ע"י מטריצת אפסים ואחדים) המשלימה את תמונה המסכה (כלומר סכום הוא '1').

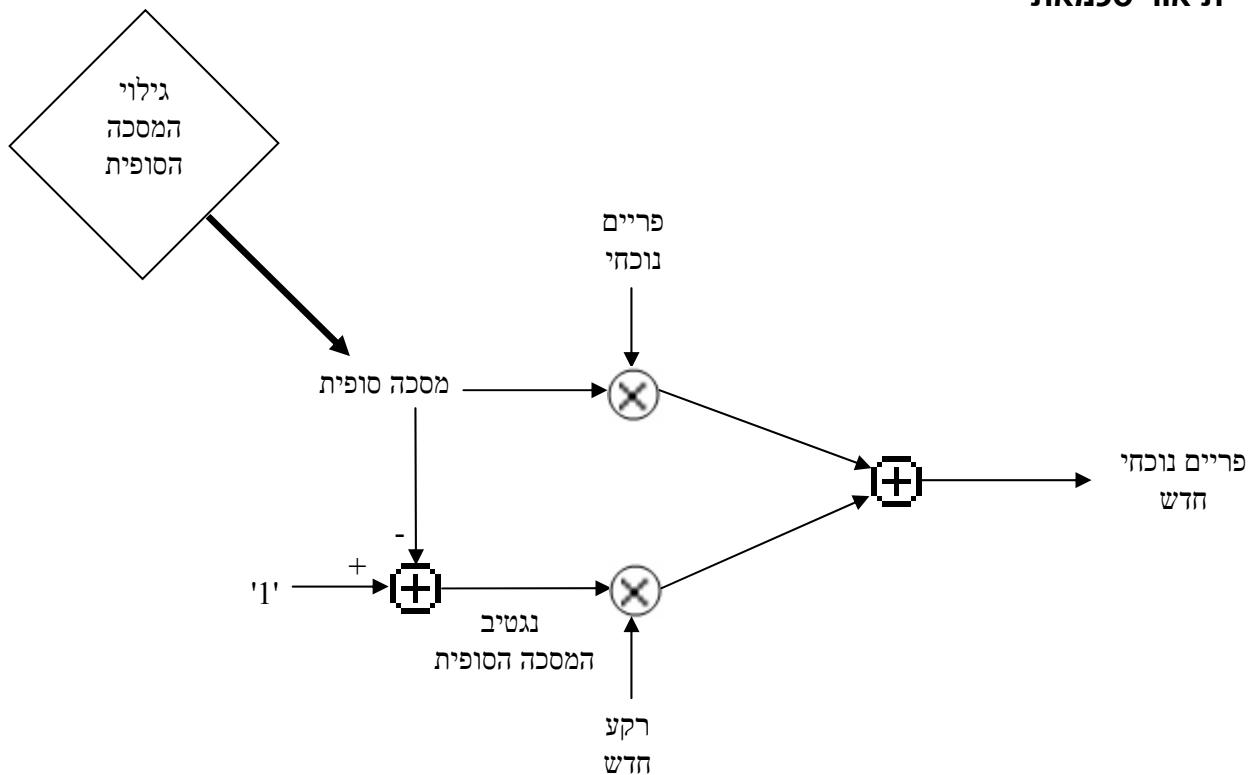
רקע חדש – תמונה שנקלטה מהמשתמש באפליקציה, במטרה ליוצר סרט וידיאו חדש בו תמונה זו תהווה תחליף לרקע המקורי של העצם, אשר את רקעו מעוניינים לשנות.

## מבנה העבודה

### הסבר כללי

הפרויקט מטרכו ברובו במציאת מסכה אשר תבודד את הדמויות מהסביבה שלה. לאחר שנמצאה מסכה כזו כל שנותר לעשות הוא להכפיל בין המסכה לבין הפריים הנוכחיים, להכפיל בין גטיב המסכה לבין תמונה הרקע החדש, ולהחבר ביניהם. ההכפלה בין שתי תמונות אליה מתיחסים כמו הינה הכפלה איבר-איבר בין שתי המטריצות המתארות את התמונה, ולא הכפלה רגילה בין שתי מטריצות.

### תיאור סכמטי



איור 1 – תיאור סכמטי



### מציאת המסכה

במהלך הפרויקט מומשו ושולבו כמה דרכי בניסיון לפתור את הבעיה בצורה אופטימאלית:

- ביצוע פעולות על תמונה בינהית נתונה למטרת הורדת הורדות רעים והשלמת דמות מתקבלת רצiosa.
- עקיבה המבוססת על טכניקת "התאמת בלוקים" (Block Matching).
- שימוש באלגוריתם Intelligent Scissors המשמש למציאה חכמה של שפות.
- למידת רקע בשיטת "רקע רץ" (Running Background).
- שערוך מיקום הדמות בפריים מסוימים בהסתמך על מידע מיקומו מפריים אחרים.

עקיבה המבוססת על טכניקת "התאמת בלוקים" מומשה בכמה דרכים שיפורטו בהמשך, אולם לא ניתן לבסס עליה אפליקציה שלמה מפאת ריגושתה לרעש והעובדה כי היא גוררת טוויות מצטברות כאשר יש תנואה – יודגס בהמשך.

המודיבציה לשימוש באלגוריתם Intelligent Scissors הייתה לקחת מסכה נתונה וליצור מסכה אחרת שתיצמד לשפות החיצונית של הדמות, תהה הדקה יותר עליה ותבודד אותה לחלוتين מסביבתה.

כאשר ישנים שינויים סביבתיים ברקע המקורי (כאשר הבולטים ביניהם שינויים תאורה) יש צורך בלמידה הרקע על מנת לאפשר הסתגלות ועמידות לשינויים. השיטה בה נעשה שימוש בפרויקט זה היא שיטת "רקע רץ" (Running Background).

על מנת לנצל את היתירות הזמן הקיימת בין פריימים עוקבים, יתרונות הנובעת מותדר הדגימה הגובה של מצלמת וידיאו (לרוב בין 15-30 פריימים בשנייה), ניתן לשערך את המסכה בפריים מסוימים מתוך הסתכבותם על פריימים קודמים ופריימים עוקב. ראוי



רק לציין כי היתירות הזמנית מתבססת על ההנחה כי העצם המבוקש איננו מסוגל לשנות בתזוזיות את כיוון תנועתו פעמיים בשלושה פרימיוםים עוקבים.

#### **סביבה העבודה**

התוכנות נעשו בסביבת עבודה Matlab 2006Ra גרסה 7.2 וההרצות בוצעו על מחשב בעל מעבד מסוג 64 bit AMD athlon(tm) 1.81 GHz ו- 1 GB זיכרון RAM. מערכת הפעלה של במחשב הייתה Windows XP. הסרטים שהורצו היו בעלי מסגרות בגודל 720x480 פיקסלים, אלא אם כן מצוין אחרת.



## רקע תיאורטי

### "התאמת בלוקים" (*Block Matching*)

#### הקדמה ומוטיבציה

בධיסת וידיאו, כאשר מעוניינים לנצל את היתירות הזמנית בין פרייםים עוקבים המשמשים לעיתים בחיזוי ובקייזו תנועה על מנת לשפר את יחס הדחיסה. יחס הדחיסה צפוי להשתפר ככל שהתנועה בסרט הוידיאו רבה יותר.

**חיזוי:** ייצרת חיזוי של המסגרת הנוכחית המבוסס על מסגרות קודמות (אחד או יותר).  
**קייזו תנועה:** ייצרת תמונה שארית על ידי החסרת תמונה החיזוי מהמסגרת הנוכחית.

בධיסת וידיאו מתבצע קייזו תנועה על מנת ליצור תמונה הפרשים בעלת אנרגיה נמוכה יותר כדי להפחית את המידע הנשלח.

בפרויקט זה אין צורך בקייזו תנועה אלא רק בשעורכה למטרת עקיבה. שעורך תנועה מותבצע על ידי השוואת איזור במסגרת הנוכחית לאזוריים שכנים במסגרת הקודמת בניסיון למצוא את ההתאמה 'הטובה ביותר' (כזו המביאה לשגיאת חיזוי קטנה ככל האפשר). אין קריטריון אבסולוטי להתאמה 'הטובה ביותר', ישנו כמה קריטריונים שונים שלכל אחד מהם יתרונות אחרים.

#### شعורך תנועה

אות וידיאו מורכב מרצף של פרייםים, רצף המהווה בפועל סדרת תמונות. סדרת תמונות תסומן ע"י  $\{X_t\}_{t=0}^T$ , כאשר  $X_t \in \{J, K\}$ , הינו גודל התמונה והפרמטר  $t$  מייצג את מימד הזמן.

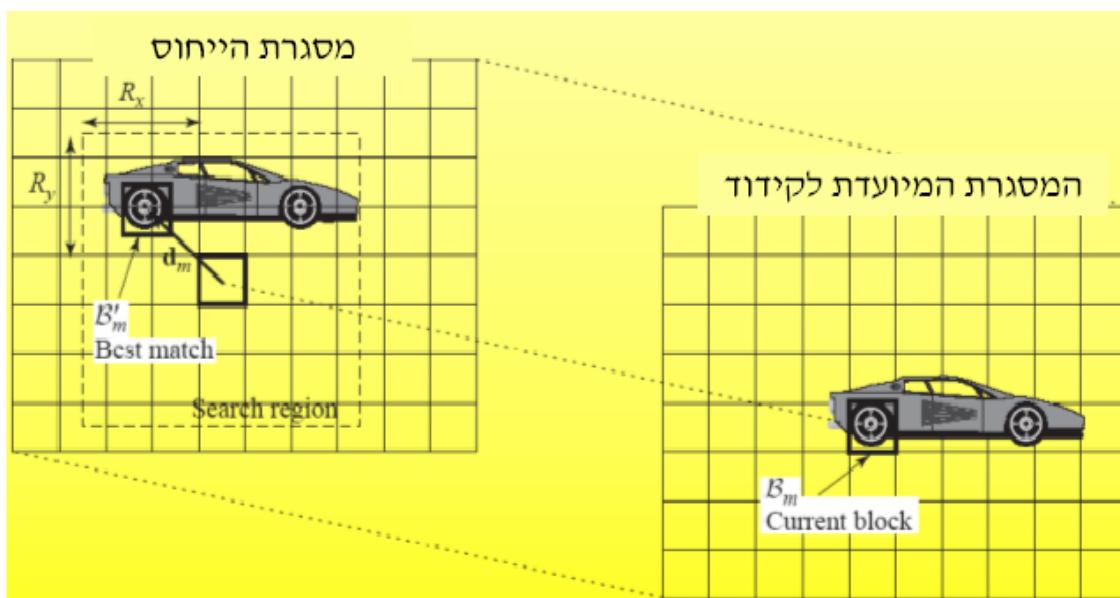


cut רוצים לשערך תנועה בשיטת "התאמת בלוקים" עבור תמונה בזמן  $t$  כלשהו. מסגרת נוכחית תוגדר בתור התמונה  $[k,j,t]_X$  ומסגרת הייחוס תוגדר בתור התמונה  $[k,j,t-1]_X$ . בתחילת שיעריך התנועה תחיליה יש לחלק את המסגרת הנוכחית ואת מסגרת הייחוס לבлокים בגודל של  $N \times M$ . לאחר מכן עברו כל בלוק במסגרת הנוכחית שרוצים לשעריך (בלוק מטרה) מתבצעת השוואה בין לבוקים בהזותות שונות במסגרת הייחוס.

בלוק מטרה שנלקח מהמיקום  $[m_0, n_0]$  בתמונה הנוכחית יסומן ב- B ועבورو מתקיים :

$$B[k,j] = X[m_0 + k, n_0 + j, t]; \quad 1 \leq k \leq M, 1 \leq j \leq N$$

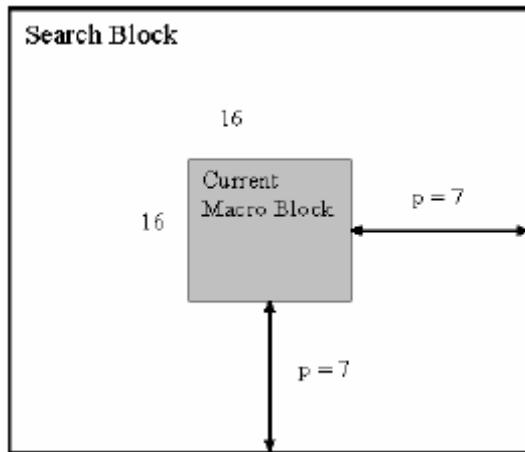
רוצים לחת קבוצת בלוקים מסגרת הייחוס ולמצוא מתוך קבוצת בלוקים זו את הבלוק אשר עבورو מידת ההתאמה בין הבלוק המטרה היא הגבוהה ביותר. קבוצת הבלוקים הנלקחת מסגרת הייחוס נקראת איזור החיפוש, כפי שמצוח גם באירור 2 .(search region)



איור 2 - הממחשה של איזור החיפוש במסגרת הייחוס



גודל איזור החיפוש מוגדר לרוב לפि פרמטר חיפוש  $p$  אשר מצין את מספר הpixels הנלקחים מכל כיוון ביחס למיקומו של בלוק המטרה (במסגרת הייחוס). הדבר מודגם באIOR 3.



AIOR 3 – איזור החיפוש מוגדר לפি פרמטר  $p$

קיימות מספר אפשרויות להגדלת ממד הטיב להתאמת בלוקים. בפרויקט יהיה שימוש בקריטריון ה-SAD ובקריטריון הקורלציה המנורמלת.  
 תהיו  $[m, n]$  קואורדינאטה כלשהי המתארת את המיקום היחסי בין בלוק כלשהו במסגרת הייחוס (בלוק באיזור החיפוש) לבין בלוק המטרה.  
 עבור קריטריון ה – SAD – (Sum of Absolute Differences) השגיאה מוגדרת לפि

$$Err[m, n] = \sum_{k=1}^N \sum_{j=1}^N |B[k, j] - X[m + m_0 + k, n + n_0 + j, t - 1]|$$

התאמת טובה לפি קריטריון זה תהיה שגיאה נמוכה, והבלוק המתאים ביותר באיזור החיפוש יהיה זה שմביא את הביטוי הניל למינימום.



עבור קרייטריוון הקורלציה המנורמלת השגיאה מוגדרת לפי

$$Corr[m, n] = \frac{\sum_{k=1}^N \sum_{j=1}^N B[k, j] X[m + m_0 + k, n + n_0 + j, t-1]}{\left( \sum_{k=1}^N \sum_{j=1}^N (X[m + m_0 + k, n + n_0 + j, t-1])^2 \right)^{0.5} \left( \sum_{k=1}^N \sum_{j=1}^N (B[k, j])^2 \right)^{0.5}}$$

קורלציה מנורמלת מקבלת ערכים חיוביים בין אפס לאחד וההתאמה טוביה לפי קרייטריוון זה תהיה קבלת ערך גבוה. הבלוק המתאים ביותר באזור החיפוש יהיה זה שմביא את הביטוי הניל' למקרים מסוימים.

בפרויקט נעשה שימוש בבלוקים ריבועיים בלבד, כלומר ככלו שעבורם  $N=M$ , ומהשך הדיוון יתיחס רק לבлокים כאלה.

### אלגוריתמים לשערוך התנועה

ישנם אלגוריתמים רבים לשערוך תנועה. השניים בהם נעשה שימוש בפרויקט לצורך עקיבה הם :

– אלגוריתם חיפוש בו מוחשבת ההתאמה הטובה ביותר ביותר עבור כל אחד מהבלוקים באזור החיפוש. באלגוריתם זה מובטח כי יתקבל הבלוק בעל מידת ההתאמה הגבוהה ביותר באזור החיפוש, אולם הוא הכביד ביותר מבחינה חישובית להיות ולא מתבצעת סלקציה כלשהי. Full-Search



עבור אלגוריתם ה- Full-Search גודל איזור החיפוש הוא  $2S \times 2S = 4S^2$  כאשר  $M = p + S$ . כמו כן, לפי הגדרת אלגוריתם חיפוש זה מספר הבדיקות זהה לגודל איזור החיפוש (כי מוחפשים את כל האפשרויות), כלומר גם מספר ההתאמות הנבדקות הוא  $2S \times 2S = 4S^2$ .

N – אלגוריתם חיפוש איטרטיבי בו מוחשבת ההתאמה הטובה ביותר על ידי הקטנת איזור החיפוש בכל צעדים, כפי שמתואר להלן :

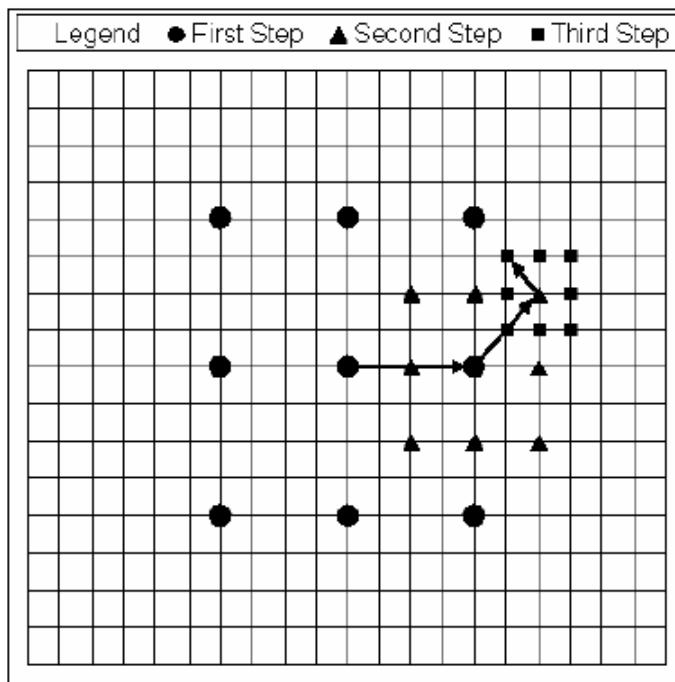
- (i) סימון מיקום בлок המטרה במסגרת הייחוס (לא תזוזה) ב  $(0,0)$ .
- (ii) קביעת גודל צעד של  $S = 2^{N-1}$  (N מגדיר את גודל איזור החיפוש).
- (iii) בניית איזור חיפוש המורכב מהפיקסל במיקום  $(0,0)$  ומשמונה הפיקסלים המתקברים מלהזות של  $S \pm$  בכל כיוון.
- (iv) בחירת המיקום בעל מידת ההתאמה הטובה ביותר בין תשעה הפיקסלים שלעיל כמרכז החיפוש החדש וסימונו ב  $(0,0)$ .
- (v) עצירת האלגוריתם אם  $S = 1$ , אחרת הקטנת גודל הצעד פי 2 ( $S = \frac{S}{2}$ ) ו חוזרת על התהליך החל משלב (iii).

גודל הצעד  $S$  בכל איטרציה באלגוריתם הינו חזקה שלמה של 2. עקרונית, בדומה לתיאור הכללי של אלגוריתמי שערוך התנועה, ניתן להגדיר פרמטר חיפוש  $k$  כלשהו שאיננו חזקה שלמה של 2, אלא שאו יש להגדיר את N באופן הבא :

$$N = \lfloor \log_2(p+1) \rfloor ; \quad S = 2^{N-1}$$



להלן איור בו הדוגמה של אלגוריתם ה-N-Step :



איור 4 – הדוגמה של אלגוריתם ה - N-Step . וקטור התנועה בדוגמה זו הוא (3,-3).

עבור אלגוריתם ה - N-Step גודל איזור החיפוש בכל צעדים הוא  $2S \times 2S$  כאשר  $S = 2^{N-1}$ . התוצאה יכולה להיות גם מחוץ לאזור זה, האזור המלא שבו היא יכולה להיות הוא :

$$(2\sum S_n) \times (2\sum S_n) = \left(2 \sum_{n=1}^N 2^{n-1}\right) \times \left(2 \sum_{n=1}^N 2^{n-1}\right) = \left(2(2^N - 1)\right) \times \left(2(2^N - 1)\right) = 4(2^N - 1)^2$$



כמו כן, באלגוריתם ה - N-Step תחילת בודקים 9 אפשרויות, לאחר מכן את ה - 8 הנוספות סביבב הceği סבירה מביניהן, לאחר מכן שוב את ה - 8 הנוספות סביבב הceği סבירה מבין האחרונות וחוזר חלילה עד עצירת האלגוריתם. בסך הכל מקבלים  $8N+1 = 8(N-1)+9$  בדיקות.

לסיכום, עבור בלוק באלגוריתם ה - N-Step גודל איזור החיפוש הוא  $4(2^N - 1)^2$  ומספר ההתאמות הנבדקות הוא  $8N+1$ .

באיור 5 מוצגות מסגרת ייחוס, מסגרת נוכחות, ומפת וקטורי התנועה כפי שימושגיים ע"י אלגוריתם התאמת בלוקים.



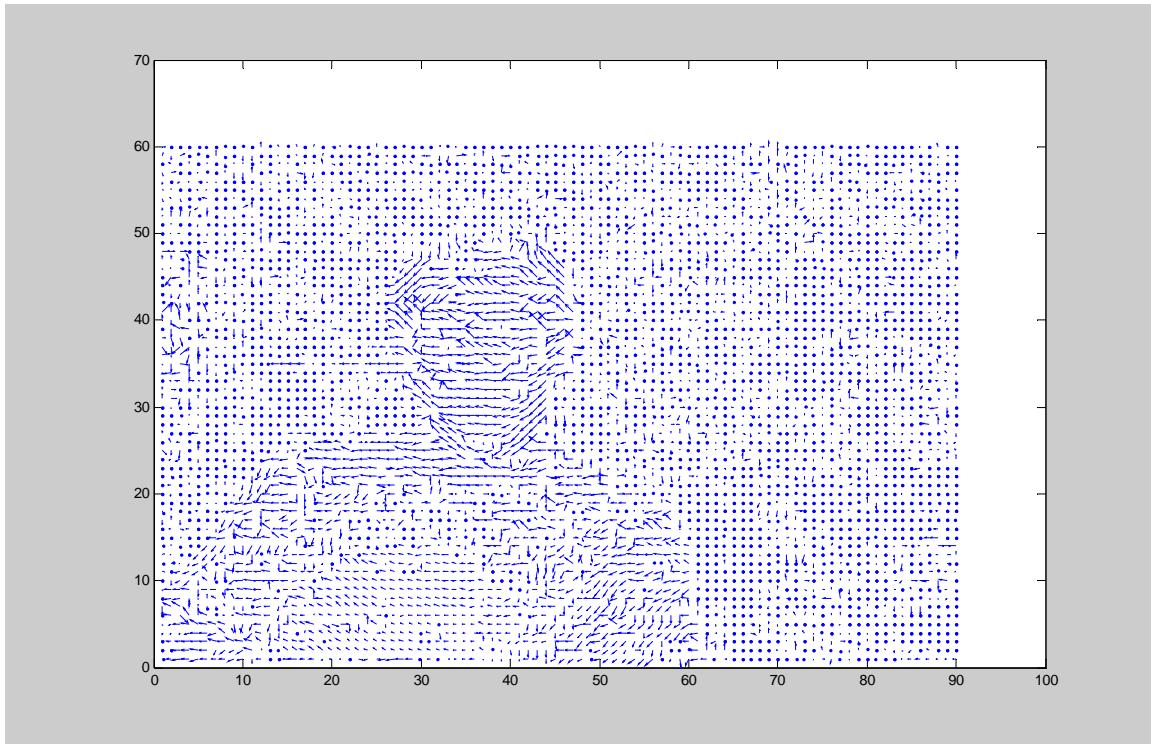
איור 5 – מסגרת הייחוס



איור 5 – מסגרת נוכחת



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקירת הראייה ומדעי התמונה



איור 5 – מפת וקטורי התנועה המתתקבלת משיעורך התנועה עבור שתי המסגרות.



## אלגוריתם *Intelligent Scissors*

### רקע ומוטיבציה

אחד הנושאים החשובים ברשותת תקשורת הוא בעית ניתוב בין מקור לעד. רשת תקשורת מיוצגת לרוב על ידי גרף שהוא אוסף של צמתים שחלקים מחוברים זה לזה בקשרות. המקור והיעד הם צמתים קשירים כלשהם והמטרה היא למצוא מסלול (אוסף של קשרות) המחבר ביניהם כך שהניתוב ביניהם יהיה אופטימלי ("מחייב מינימלי"). כל קשת מקבלת ערך מסומי המכונה "משקל הקשת" וקריטריון נפוץ לאופטימליות ניתוב הוא מסלול בו סכום משקלים קשווים הוא מינימלי.

הרעיון אפוא הוא לייצג את התמונה כgraf בו כל פיקסל מהווה צומת המחבר בקשרות לשומות שכנו. לכל קשת בgraf יינתן משקל, כאשר מעוניינים להעניק משקלים נמוכים לקשווים אשר שימושו שפות בתמונה המקורית. ככל שהshape חזקה יותר רוצים משקל הקשת המתאימה לה יהיה נמוך יותר.

Intelligent Scissors הינו כלי אינטראקטיבי אשר משמש לחילוץ מדויק של אובייקטים מתוך דוגמאות ורכבתם בתמונות אחרות. הכלי הניל מאפשר חופש למשתמש וمتקשר אליו בעזרת העכבר. כל שעל המשתמש לעשות הוא להקיף את האובייקט אותו הוא מעוניין לחלץ באמצעות העכבר ולסמן בתוכו נקודת קלשי (Seed Point). לאחר שמקבל הכלי intelligent Scissors את המידע מהמשתמש, הוא מפעיל אלגוריתם תכנות דינامي דיסקרטי (DP) אשר ממיר את הבעה בבעיה של מציאת מסלולים אופטימליים בgraf דו מימדי, מתוך שאיפה שהמסלולים יתארו במדויק את קו השפה של האובייקט הרצוי.

בפרויקט זה האפליקציה הממומשת אינה מאפשרת בשום שלב שימוש ישיר בעכבר ולכן היה צורך בפיתוח אלגוריתם אשר מדמה את פעולה העכבר, ואשר יפורט עליו בהמשך.



### תיאור האלגוריתם

האלגוריתם משלב מספר שיטות לגילוי שפנות במטרה לביצוע בצורה אופטימלית גילוי מדויק של שפנות אובייקטיבים בתמונה ומסתמך על "תבניות משקלים" המגדירות את "מחירו" של כל פיקסל בתמונה יחסית לשכנו. "מחירו" של כל פיקסל בתמונה לצורך זיהוי גבולות של אובייקטיבים יקבע בסכום משוקלל של שלושת המרכיבים הבאים:

$f_z$  - פונקציה המבוססת על אופרטור הלפלסיאן. זוהי פונקציה חד ממינית מסדר שני אשר המשטנה הבלתי תלוי שלה הוא פיקסל היעד.

чисוב  $f_z$ :

את הלפלסיאן בנקודה מסוימת מחשבים ע"י פעולה מסכחה של התמונה עם אופרטור הלפלסיאן המוגדר ע"י המטריצה:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & -4 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

את התוצאה המתקבלת מפעולת מסכחה זו נסמן ע"י  $I_L(q)$  ואז  $f_z(q)$  ייחסב ע"י:

$$f_z(q) = \begin{cases} 0; & \text{if } I_L(q) = 0 \\ 1; & \text{if } I_L(q) \neq 0 \end{cases}$$



- פונקציה המבוססת על גרדיאנט האmplיטודה. זהה פונקציה חד ממדית מסדר ראשון אשר המשתנה הבלתי תלוי שלה הוא פיקסל היעד.

чисוב :  $f_G$

תחילה יש לחשב את הנגורות החלקיות בתמונה. עבור תמונה דיגיטלית המוצגת באופן דיסקרטי חישוב הנגורות החלקיות יבוצע לפי :

$$I_X(i,j) = F(i,j+1) - F(i,j)$$
$$I_Y(i,j) = F(i+1,j) - F(i,j)$$

כאשר X הציר האופקי, ו-Y הינו הציר האנכי.

גודלו של וקטור הגרדיאנט יחשב לפי  $G(i,j) = \sqrt{I_X^2 + I_Y^2}$  ונקבל לבסוף :

$$f_G(q) = f_G(i,j) = \frac{\max(G) - G(i,j)}{\max(G)} = 1 - \frac{G(i,j)}{\max(G)}$$

לאחר הגדרת תבנית המשקלים  $f_G$ , מציאת המסלול האופטימלי בשלב הבא יבוצע כך שיחסוב המרחק מפיקסל אחד לשכנו יהיה תלוי האם הוא ממוקם באלכסון אליו או אנכי/אופקי אליו : אם הפיקסל השכן נמצא באלכסון אזי מרחקו מוכפל ב-  $\sqrt{2}$ , ואם הוא אנכי/אופקי אליו אזי מרחקו מוכפל ב- 1.

- פונקציה המבוססת על גרדיאנט הcyonיות. זהה פונקציה דו ממדית מסדר שני אשר המשתנים הבלתיים תלויים שלה הם פיקסל היעד ופיקסל המקור.



חישוב :  $f_D$

מגדירים תחילת את הגדים :

$$D(p) = (I_Y(p), -I_X(p))$$
$$L(p, q) = \begin{cases} q - p; & \text{if } D(p) \cdot (q - p) \geq 0 \\ p - q; & \text{if } D(p) \cdot (q - p) < 0 \end{cases}$$
$$d_p(p, q) = D(p) \cdot L(p, q)$$
$$d_q(p, q) = L(p, q) \cdot D(q)$$

ולקבלת גרדיאנט הכווניות עבור הקשת  $(q, p)$  מתקבל על ידי הביטוי :

$$f_D(p, q) = \frac{2}{3\pi} \left( a \cos(d_p(p, q)) + a \cos(d_q(p, q)) \right)$$

שכלול תבניות המשקלים של שלוש הפונקציות מספק גילי שפה אמין ומדויק למדי.  
השכלול מתבצע ע"י הגדרת משקל לכל פונקציה באופן הבא :

$$l(p, q) = \omega_Z \cdot f_Z(q) + \omega_G \cdot f_G(q) + \omega_D \cdot f_D(p, q)$$

$l(p, q)$  היא הפונקציה המייצגת את המחיר המקומי של הקשת כאשר  $p$  ו-  $q$  הם שני פיקסלים סמוכים, ו-  $\omega_Z, \omega_G, \omega_D$  הם המשקלים של הפונקציות  $f_Z, f_G, f_D$  בהתאם.

לאחר שיכלול זה מתקבלת מטריצה המכילה ערך עבור כל פיקסל. ערך זה מהוווה "מחיר" מדויק יותר עבור כל פיקסל אשר שימוש בפונקציה בודדת בלבד. המטריצה



הסופית שמתאפשרת לאחר חישוב משקלים כל הפיקסלים בתמונה, מהוות את תבנית המשקלים עליה מופעל אלגוריתם ה- Dijkstra שמטרתו למצוא מסלולים אופטימליים שבזורתם מאותרת לבסוף השפה.

המסלול האופטימלי, שהינו בעצם קו השפה של האובייקט הרצוי, מאותר רק לאחר הגדרת תחום סביבה האובייקט המבוקש. הגדרת התחום מתבצע ע"י המשמש: סימון נקודה אחת מדוייקת על אובייקט ומסלול רציף של נקודות בסביבתו.

סימון הנקודות באלגוריתם ה- Intelligent Scissors מתבצע באמצעות העבר וכל נקודה במסלול אותו מקיף המשמש כניסה למגר הנזונים עימיו עובד האלגוריתם. על- מנת לאפשר לאלגוריתם לעבוד בצורה נכונה, הנקודות שנקלטו במגר הנזונים עוברות דילול. בפרויקט זה נכתב אלגוריתם אשר מסמן את נקודות המסלול התוחם באופן אוטומטי ללא צורך בשימוש בעכבר.

מציאת המסלול המתאר את קו השפה מתבצעה בשלבים, כאשר בכל שלב נמצא חלק קטן של המסלול והוא מהוות את המשכו של הקטע שהתגלה בשלב אחד לפניו על מסלול השפה. האלגוריתם משתמשים כאשר הוא מגיע חזרה לנקודת ההתחלה ואז נוצר מסלול סגור המתאר את קו השפה.

בכל שלב מוגדרת נקודה הנבחרת כנקודת ההתחלה של הקטע, היא נקודת ה- seed point. בשלב הראשון של אלגוריתם ה- Intelligent Scissors נקודה זו מוגדרת על ידי המשמש באמצעות העכבר. בפרויקט זה אין שימוש בעכבר ולכן נקודת ה- seed point הראשונה נבחרה כנקודת מרכז המסה של האובייקט, מתוך הנחה שנקודת מרכז המסה נמצאת בתחום האובייקט.

בכל שלב, מציאת קטע מהשפה מתבצע על ידי חיפוש המסלולים האופטימליים בין כמה נקודות עוקבות, שסמננו על ידי המשמש בסביבת האובייקט, לבין הנקודה הנבחרת כ- seed point. איתור המסלולים האופטימליים בכל שלב מתבצע על ידי הפעלת אלגוריתם ה- Dijkstra. מבצעים חיתוך בין המסלולים האופטימליים שנמצאו,

וקטע זה משוויך כחלק מקו השפה של האובייקט. נקודות הסיום של הקטע מהוות את נקודות התחילה של הקטע הבא והוא ה- seed point החדש של השלב הבא. בפרויקט זה לא מומש אלגוריתם ה- Intelligent Scissors במלואו, אלא בסיס בחלקו על קוד קיים תוך התאמת הקוד לדרישות. הקוד עלייו בוצעה התבבשות נלקח מפרויקט עבר [12]. איורים 7-6 נלקחו מפרויקט העבר המזכיר.

באIOR 6 ניתן לראות ארבעה שלבים של איתור נקודות seed point. בכל שלב נמצאים שלושה מסלולים (שלוש נקודות שסומנו ע"י המשמש אל הנקודה התחלה), המתוארת כ- seed point של אותו שלב. בסיום כל שלב, לאחר מציאת המסלול משותף לששת המסלולים, עוברים לשלב הבא שבו נקבעת ה- seed point כנקודה الأخيرة של המסלול המשותף שאוטר בשלב הקודם. בקורס שלבים זו מתקדם האלגוריתם עד אשר נגלה השפה במלואה.



איור 6 – הדגמת שלבים באיתור נקודות seed point.

כפי שכבר צוין, הנקודות המקיפות את האובייקט עוברות דילול. דילול מועט מדי של נקודות עשוי להכיל קטע משותף שאינו מתאר את קו השפה בשל ציפויות הנקודות וקרבתם זו לזו. מצד שני דילול רב מדי של נקודות גם הוא אינו נותן תוצאות טובות, משום שאם נקודת התחלה והסיום שנבחרו רוחקota זו מזו המסלול עלול לסתות למסלול לא רצוי.

בפרויקט העבר נמצא כי יחס דילול של 5:1 הינו בחירה טובה ובו נעשה שימוש. כמו דוגמאות למציאת מסלולים אופטימליים עבור יחס דילול זה מופיעות באירור 7.



איור 7 – מסלולים אופטימליים שנמצאו עבור יחס הדילול שנבחר 5:1.



### החסרת רקע (*Background Subtraction*)

#### הקדמה

שיטה שימושית לחילוץ אובייקטים נעים מרקע רצף פריימרים היא שיטת "החסרת רקע". שיטה זו כוללת שערוך תמונה רקע ללא עצמים נעים, הפחחת כל פריים חדש מתמונה זו, ויצירת תמונה ביןראית לפי סך. התוצאה המתקבלת היא תמונה ביןראית אשר מבדילה בין הרקע המשוערך לבין העצמים הנעים.

קיים קושי רב בשערוך הרקע כאשר ישנים שינויים סביבתיים ברקע כגון שינוי תאורה ותזוזות מכלמה.

הגישות הרבות לפתרון בעיה זו נבדלות אלה מלאה בסוג מודל הרקע בו משתמשים ובתהליך עדכונו המודל.

גישה עיקריות הן מיצוע הרקע, מודל GMM ושיטת הרקע הרץ, כאשר שיטת ה-GMM מאפשרת ישות של רקע מרובה והאחרות של רקע יחיד.

#### מיצוע הרקע

הנתת היסוד בשיטת לימוד זו היא כי קיים אתחול ראשוני של תמונה רקע ללא עצמים נעים לפני תחילת הרצת האלגוריתם. בזמן הרצת האלגוריתם מותר לרקע להשתנות והמטרה היא ללמידה את השינויים.

בשיטת זו הרקע מוגדר כממוצע (או חציון) מהפריימרים הקודמים, כאשר זה הוא מספר ידוע וקבוע מראש.

שיטת זו מאד מהירה, אך צורכת זיכרון רב ככל ש- $n$  גדול יותר וככל שתמונה סרט הויידיאו גדולה יותר (סיבוכיות זיכרון  $O(n \cdot \text{size}(\text{frame}))$ ).



מציאת הרקע B בזמן t תבוצע לפि :

$$B_t = \frac{\sum_{i=t-n}^{i=t-1} Frame_i}{n}$$

על פי ביטוי זה הרקע מחושב בכל פעם מחדש לפחות לפி  $t$  הפריים האחרונים, כך שאם אובייקט כלשהו מופיע בכל  $t$  הפריים הללו, הוא יילקח בחשבון בעת חישוב הרקע והרקע לא יהיה נקי ממנו. הדבר יפגע כמובן בביטויים היות ותמונה הרקע המוחשבת לא תהיה נקייה ממחיצים שאוטם מעוניינים לסוג מסווג בתוך רקע קדמי בין אם הם נעים ובין אם הם נייחים. השיטה הנילא לא משיגה את המטרות הרצויות בפרויקט זה.

### מודל GMM

לעתים רבות כאשר מעוניינים לחשב פילוג של וקטור אקראי כלשהו  $x$  עושים זאת בגישה הפרמטרית. בגישה זו מניחים כי הפילוג הנדרש  $p_x(x)$  הינו בעל צורה ידועה, המוגדרת עד כדי וקטור פרמטרים  $\theta$ , כלומר

$$p_x(x) = p_x(x | \theta)$$

לרוב מדובר בוקטור פרמטרים ממשיים, דהיינו  $\theta = (\theta_1, \dots, \theta_P)^T \in \mathbb{R}^P$ .

קיימות שתי גישות: גישה בייסיאנית וגישה לא בייסיאנית. בגישה הביאיסיאנית מניחים כי הפרמטר (הוקטור)  $\theta$  הינו משתנה מקרי, בעל פילוג הסתברות אפrioרי ידוע  $p(\theta)$ , ואילו בגישה הלא בייסיאנית רואים את הפרמטר  $\theta$  כגודל קבוע וטרמיניסטי, ובפרט מנעים מהනחות סטטיסטיות כלשהן לגביו.



במקרים רבים הפילוג המשוערך ( $p_x(x)$  הינו מולטי-מודלי, כלומר מרוכז במספר אזורים למרחב. מודל מקובל עבור פונקציית הצפיפות במקרה זה הינו מודל תערובת Gaussian Mixture (mixtures), ובמיוחד התערובת הגאוסית, הידוע בתור מודל ה-GMM (Model).

זהו מודל לא בייסיאני שצורתו הכללית היא:

$$p_x(x|\theta) = \sum_{j=1}^J w_j \eta(x; \mu_j, \Sigma_j)$$

כאשר

$$\eta(x; \mu, \Sigma) = \frac{1}{(2\pi)^{\frac{d}{2}} |\Sigma|^{\frac{1}{2}}} e^{-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)}$$

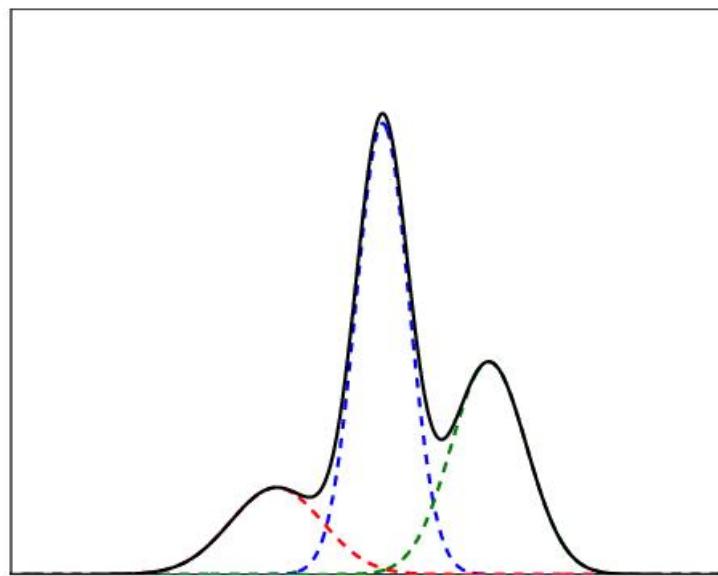
על מנת לקבל את צפיפות ההסתברות לפי מודל זה, הפרמטרים אותם יש לשערך הם

$$\theta = \left\{ w_j, \mu_j, \Sigma_j \right\}_{j=1}^J, \text{ ובעצם וקטור הפרמטרים הינו}$$

על מנת ש  $p_x(x|\theta)$  אכן תהיה צפיפות הסתברות נדרש להתקיים התנאי

$$w_j \geq 0, \quad \sum_{j=1}^J w_j = 1, \quad \text{�כל וקטור פרמטרים } \theta, \text{ ומכאן ש} \int_{-\infty}^{\infty} p_x(x|\theta) dx = 1$$

באיור 8 ניתן לראות את צורתו הכללית של פילוג הסתברות כלשהו כפי שהוא משוערך באמצעות מודל ה – GMM עבור תערובת של שלושה משתנים אקראיים גאוסיים.



איור 8 – מודל ה – GMM עבור  $J=3$ .

מודל ה-GMM למטרת למידת רקע מיושם על ידי שעורך של צפיפות ההסתברות לצפיה בערך (ווקטור) מסוים של פיקסל. את המודל מפעילים על כל פיקסל בתמונה בנפרד ובזמנים מחזיקים עבור כל פיקסל כזה בפרמטרים  $\{w_j, \mu_j, \Sigma_j\}_{j=1}^J$  ובכל פרויים חדש ובעצם מוחזקים אותם לפי הערכיהם הנצפים בפרויים החדש והערכיהם שליהם בפרויים הקודמים. ככלمر בזמן מסוים  $t$  מעדכנים את הפרמטרים  $\{w_{j,t}, \mu_{j,t}, \Sigma_{j,t}\}_{j=1}^J$  לפי פרמטר למידה  $\alpha$  ובהסתמך על הפרמטרים שנלמדו בזמן  $t-1$ .



ערך (ווקטור) של פיקסל מסוים בזמן  $t$  יסומן כ  $X_t$  ולפי מודל ה-GMM צפיפות ההסתברות שלו תשוערך לפי:

$$p_{x_t}(X_t) = \sum_{j=1}^J w_{j,t} \eta(X_t; \mu_{j,t}, \Sigma_{j,t})$$

מספר הגאוסיאנים בתערובת,  $J$ , נקבע בהתאם לזכרון הפניו ולכוח העבודה של המחשב, כאשר ערך מקובל עבורו הוא בין 3 ל-5. כמו כן נהוג להניח כי בין האלמנטים שמהם בנוי הווקטור  $X_t$  לא קיימת תלות סטטיסטיות וכי השונות שלהם זהה, ואז

$$\text{לוקחים מטריצת קוואריאנס מהצורה } I = \sum_{j,t} \sigma_{j,t}^2.$$

הנחה הנ"ל נותנת הקליה חישובית על חישוב דיקום מסוים, מפני שבין האלמנטים לרוב כן קיימת איזושהי תלות סטטיסטית. נוח לחתה למשל את הווקטור  $(R_t, G_t, B_t)$  שבו בעצם כן קיימת תלות מסוימת בין האלמנטים. באופן כללי ניתן גם לעבור מרחב RGB למרחבים אחרים כגון מרחב HSV או מרחב YCbCr, אשר בהם הקורלציה בין המרכיבים השונים קטנה יותר.

כפי שכבר צוין קודם לכן, ישנו כללי עדכון רבים עבור  $\{w_{j,t}, \mu_{j,t}, \Sigma_{j,t}\}_{j=1}^J$  לפי פרמטר  $\alpha$ . האלגוריתמים המרכזיים מופיעים במקורות [1][2].



לאחר שעודכנו הפרמטרים הניל ממיינים את הגאוסיאנים של התערובת לפי סדר עולה של היחס  $\frac{w_{j,t}}{\sigma_{j,t}}$ ,  $1 \leq j \leq J$ . לאחר המיוון B הגאוסיאנים הראשונים בזמן  $t$  נבחרים כמודל הרקע כאשר B נקבע לפי סף מסוים T (בין 0 ל-1) על ידי הביטוי:

$$B_t = \arg \min_b \left( \sum_{j=1}^b w_{j,t} > T \right)$$

כך בעצם אנו מאפשרים מודל רב של הרקע (רקע מולטי-מודלי) להיות ובעצם B יכול להיות גדול מ-1. מספר המודלים שיתארו את הרקע תלוי בסף T. ככל ש T קטן יותר, הדבר יגרור B קטן יותר והנטיה תהייה להשתמש במודל בעל רקע יחיד. מאידך, ככל ש T גדול יותר, הדבר יגרור B גדול יותר והנטיה תהייה להשתמש ביוטר מודלים של הרקע.

ישנן מספר שיטות שונות לצורך שינון ומיוון נקודות (במרחב הוקטורית המייצג את ערך הפיקסל) לגאוסיאנים השונים. לרוב מקובל להשתמש במרחב של עד 2.5 מהותולת כאיינדיקציה לשיעיותו לאוטו גאוסיאן (ניתן לשיעך ערך מסוים של פיקסל ליותר מגאוסיאן אחד). המיוון והשינון נחוץ הן לצורך עדכון הפרמטרים והן לצורך סיוג נקודות רקע או חלק מהאובייקט.

בשיטת זו בכל פריים כל הפיקסלים ללא יוצא מן הכלל מתעדכנים לפי הפריים החדש, ככלומר אין סלקטיביות. היות והרקע הינו מולטי-מודלי תהיה הבחנה בין חפצים נעיםLivni, 2013. בין הרקע. הבחנה הניל תהיה טובה יותר מאשר בשיטה של "מיוצע רקע" מפני שהרקע הינו מולטי-מודלי. לכן חפץ מסוים ייטמע ברקע רק אם הוא יהיה במצב סטטי



לפרק זמן ארוך מספיק כך שיוצר גאוסיאן שיופיע באותו בעל יחס  $\frac{w_{j,t}}{\sigma_{j,t}}$  דומיננטי ביחס

לגאוסיאנים האחרים. שיטה זו צפופה לתות תוצאות טובות עבור חפצים נעים, אולם תתקל בקשישים רבים עבור אובייקטים שאינם בתנועה מתמדת. שיטה זו מושגה רק חלק מהמטרות הרצויות בפרויקט זה אולם לא את כולם.

### שיטת הרקע הרץ (Running Background)

בדומה להנחה היסוד בשיטת הלימוד על פי מיצוע הרקע גם בשיטה זו מניחים כי קיימים אתחול ראשוני של תמונה רקע ללא עצמים נעים לפני תחילת הרצת האלגוריתם. בזמן הרצת האלגוריתם מותר לרקע להשנות והמטרה היא ללמידה את השינויים בשיטה זו, בニיגוד לשיטת "מיצוע הרקע", נשמרת תמונה רקע ייחידה אשר מתעדכנת בכל פריים חדש כמשמעות משוקלל בין הפריים החדש לבין תמונה הרקע הקודמת, לפי פرمטר למידה  $\alpha$ .

באופן כללי, ללא סלקטיביות, הרקע בזמן  $t$  מחושב לפי הביטוי:

$$B_t = \alpha \cdot Frame_t + (1-\alpha) \cdot B_{t-1}$$

כפי שהוזכר קודם לכן, ללא סלקטיביות מתקבל מודל רקע שמטמייע בתוכו גם חפצים נעים. אם מעוניינים למנוע הטמעה זו ולהפריד בין החפצים הנעים לרקע וללמידה רק את הרקע יש לבצע סלקציה שבה הלמידה מתבצעת רק לפי הפיקסלים אשר מסווגים בתור רקע אחריו. הסלקציה מתבצעת על ידי חישוב הרקע בזמן  $t$  לפי הביטוי:

$$B_t(x, y) = \begin{cases} \alpha \cdot Frame_t(x, y) + (1-\alpha) \cdot B_{t-1}(x, y) & \text{if } Frame_t(x, y) \text{ is background} \\ B_{t-1}(x, y) & \text{if } Frame_t(x, y) \text{ is foreground} \end{cases}$$



**הטכניון - מכון טכנולוגי לישראל**  
**הפקולטה להנדסת חשמל**  
**המעבדה לחקור הראייה ומדעי התמונה**



כאשר שיטה זו כוללת סלקטיביות של הפיקסלים השונים היא משיגה את מרבית המטרות בפרויקט זה: היא מצליחה לעקוב بصورة טובה אחר אובייקטים נאים גם אם הם אינם בתנועה מתמדת.

הבעיה בלמידה זו היא שהרקע המוסתר על ידי האובייקט אינו נלמד, ומודל הרקע לא מटעדכו עבورو. הדבר בא לידי ביטוי בעיקר כאשר חפץ נע מסויים עוצר תנועתו ולאחר מכן חוזרשוב לנוע.

בפרויקט הוחלט להשתמש במודל זה לצורכי שערוך הרקע.



## עיבוד תמונה

### מרחבי צבעים

בניגוד לעולם האמתי הרציף, במחשב תמונות הן דיגיטליות ומיצגות על ידי ערכים דיסקרטיים. הסיבה לכך היא שהמחשב הינו בעל זיכרון סופי, ולכן עליו לבצע כימות (קוונטייזציה).

תמונות רמות אפור מוצגות על ידי ערך ממשי סקלרי אשר מצין את רמת הבاهירות עבור כל פיקסל בתמונה. לעומת זאת, תמונות צבעוניות, בהן נעשה שימוש בפרויקט זה, יש לייצג על ידי ערך ממשי וקטורי על מנת לדעת גם את הצבע של הפיקסל ולא רק את רמת הבاهירות שלו.

קיימות אפשרויות שונות לייצוג ערכי הוקטוריים, זאת בהתאם למרחב הצבעים אשר בוחרים לשיך אליו את הוקטור.

ישנם מרחבי צבעים רבים. מרחבי הצבעים שבהם נעשה שימוש במהלך הפרויקט הם HSV, YCbCr ו- RGB, כאשר כל אחד מהמרחבים הניל פורש את כל הצבעים הקיימים בטבע.

### מרחב RGB

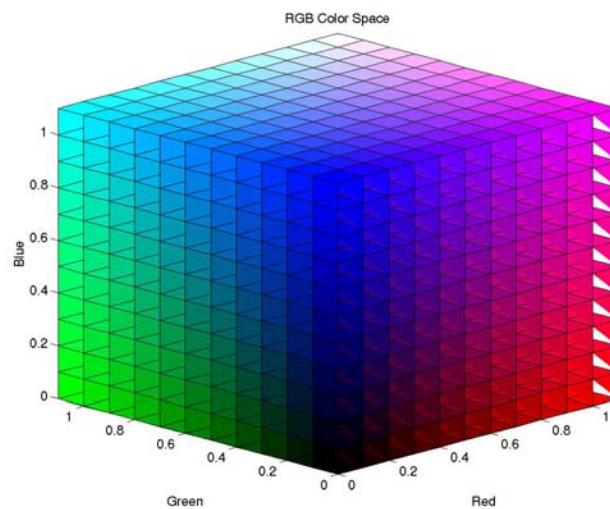
בתמונות שמיוצגות בפורמט RGB כל פיקסל מיוצג על ידי וקטור ממימד 3 לפי :

$$x = \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

כאשר הקואורדינאטה R מייצגת את עוצמת הצבע האדום, הקואורדינאטה G מייצגת את עוצמת הצבע הירוק והקואורדינאטה B מייצגת את עוצמת הצבע הכחול.



את המרחב RGB נוח לייצג באמצעות צירים קרטזיות ולהסתכל על המרחב כקובייה תלת ממדית, כפי שמודגם באיוור 9.



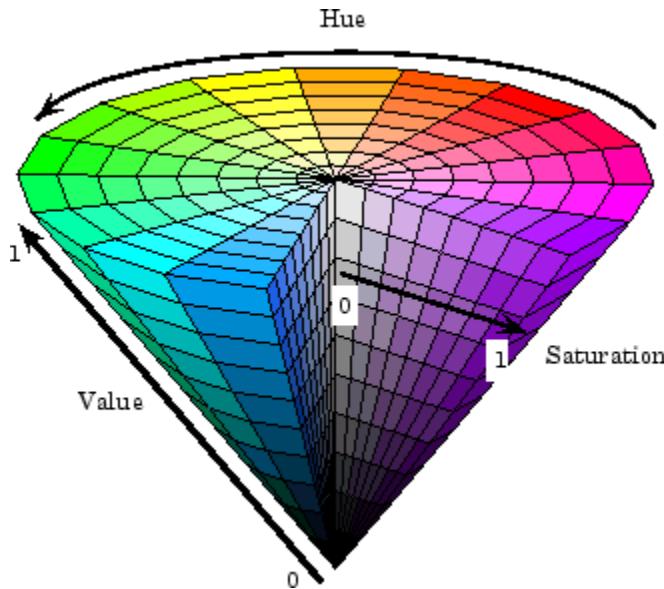
.איור 9 – מרחב RGB.

### מרחב HSV

ביצוג HSV כל פיקסל מיוצג על ידי וקטור ממימד 3 לפי :

$$x = \begin{pmatrix} H \\ S \\ V \end{pmatrix}$$

כאשר הקואורדינאטה H מייצגת את אורך הגל דומיננטי (גוון - hue), הקואורדינאטה S מייצגת את הטוher (רווייה - saturation) והקואורדינאטה V (value) מייצגת את רמת הבחרות. מבט על מרחב HSV מוצג באיוור 10.



איור 10 – מרחב HSV.

על מנת להגדיר את H ו-S יש לדון תחיליה בדיאגרמת הצבעים כפי שנקבעה על ידי וועדת ה-CIE.

דיאגרמת הצבעים הינה כלי שמצבע מיפוי של כל הצבעים על משטח דו-מימדי. הגישה מותיחסת לשלוות עקומי ההיינוט של קולטני אור היום של האדם:  $\bar{x}(\lambda), \bar{y}(\lambda), \bar{z}(\lambda)$ .



בහינתן פילוג הספק ( $\lambda f$ ) מחשבים את שלישיות ה- TRISTIMULUS ( $X, Y, Z$ ) לפי האינטגרלים הבאים:

$$X = \int_{-\infty}^{\infty} f(\lambda) \bar{x}(\lambda) d\lambda$$

$$Y = \int_{-\infty}^{\infty} f(\lambda) \bar{y}(\lambda) d\lambda$$

$$Z = \int_{-\infty}^{\infty} f(\lambda) \bar{z}(\lambda) d\lambda$$

שלישייה זו מחסרים מימד אחד ויוצרים את שלישייה ה- TRICHIROMATIC :  $(x, y, z)$

$$x = \frac{X}{X + Y + Z}$$

$$y = \frac{Y}{X + Y + Z}$$

$$z = \frac{Z}{X + Y + Z}$$

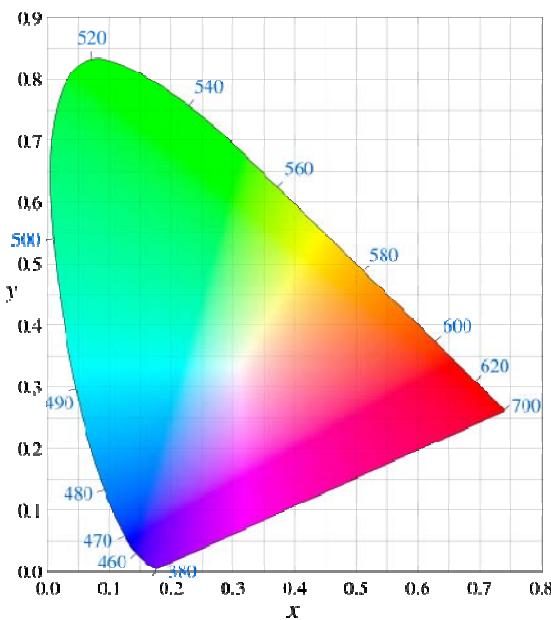
קיימת תלות בין השלישייה זו להיות ומתקיים  $1 = z + y + x$  ולכן מאבדים מימד אחד ונשארים רק שני משתנים בלתי תלויים אשר כל אחד מהם מהווה דרגת חופש אחת. המימד אותו מאבדים הוא בעצם מימד הבחרות, אשר עליו ניתן לוותר להיות מעוניינים לבטא את הצבע בלבד (בחירה לא מעוניינת).



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



דיגרמת ה – CIE מייצגת את שלישית  $(z, y, x)$  בדו-מימד לפי המשתנים  $(y, x)$  היות והשלישיה הניל פורשת כאמור מימד 2 ולא 3. הצבעים בדיגרמה מוצגים באירור 11.



איור 11 - דיגרמת ה – CIE עם החלוקה לצבעים.

תחילה יש לציין כי המעטפת העליונה של הדיגרמה בעלת הצורה הפרבולית שיצכת לדיגרמה ומהווה את כל הצבעים המונוכרומטיים, דהיינו צבעים שפילוג ההספק שלהם מרוכז כולם באורך גל מסוים (פילוג בעל צורה מתמטית  $(\lambda_1 - \lambda) = P\delta(\lambda)$ ). על המעטפת מצוינים מספרים המייצגים את אורך הגל של הצבע המונוכרומטי באותו נקודה. לעומת זאת, המעטפת התחתונה בעלת הצורה היליניארית אינה כללת בדיגרמה והנקודות עליה אינן צבעים. כמו כן, ניתן לראות כי הצבע הלבן מופיע במרכז הדיגרמה סביבה הנקודת  $(x, y) = \left(\frac{1}{3}, \frac{1}{3}\right)$  צפוי מפני שפילוג ההספק של האור בלבן הינו אחיד ( $f(\lambda) = const$ ).

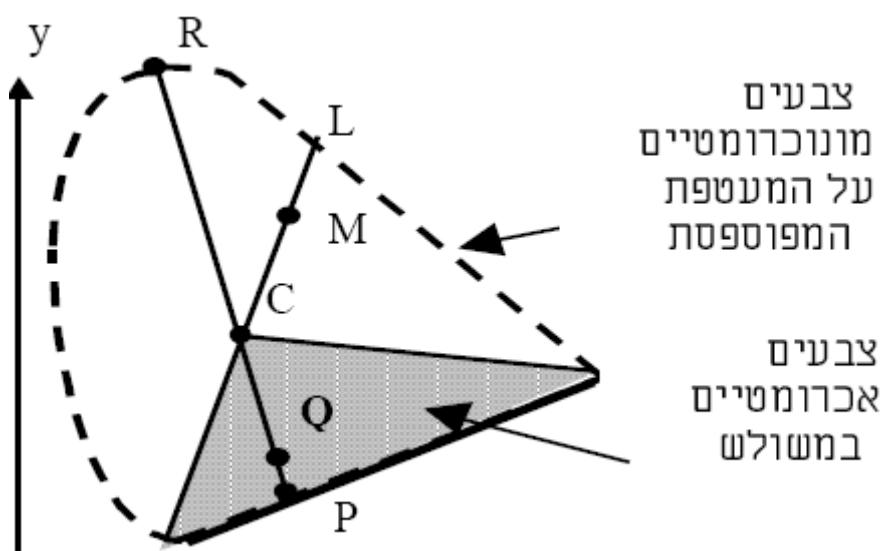


צבעים אקרומטיים הם הצבעים שאלייהם לא ניתן להגיע על ידי תהליך חיבור צבעים מונוכרומטיים עם הצבע הלבן.

גונו או אורך גל דומיננטי של צבע מסוים בנקודה כלשהי בדיאגראמה הוא אורך הגל של הצבע המונוכרומטי הנדרש לחיבור עם הצבע הלבן על מנת להגיע אליו הצבע. עבור צבעים אקרומטיים אורך הגל הדומיננטי יהיה מינוס אורך הגל של הקו המחבר עם הצבע המונוכרומטי, זאת מפני שהנקודות על המעתפת התחתונה אינן מתארות צבעים מונוכרומטיים ואינן שייכות לדיאגראמה.

רוויה היא מושג המבטא מרחק בין צבע כלשהו לבין הצבע הלבן, כך שככל שהצבע רחוק יותר מהלבן הוא יותר רווי. המדיידה היא כמותית ומונומלת בין הערכים 0 ו-1. רוויה גדולה מעידה כי הצבע קרוב יותר לצבעים המונוכרומטיים.

המחשה של המושגים מוצגת באירור הבא :



איור 12 – המוחשת המושגים המדוברים על גבי דיאגרמת ה – CIE.



באיור 12 ערך הרווחה של הצבע M יבוטא על ידי היחס  $\frac{|CM|}{|CL|}$  וערך הרווחה של הצבע Q

יבוטא על ידי היחס  $\frac{|CQ|}{|CP|}$ . הגוון של הצבע M הינו אורך הגל של הצבע L על המעטפת

המפוספסת ואורך הגל של הצבע האקרומטי Q הינו מינוס אורך הגל של הצבע R.

ב – MATLAB קיימת הפונקציה `rgb2HSV(I)` הממירת תמונה מייצוג במרחב RGB  
לייצוג במרחב HSV ונעשה בה שימוש בפרויקט.

#### מרחב YCbCr

מרחב זה מתקיים לאחר הפעלה של טרנספורמציה ליניארית על מרחב RGB, לפי :

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{pmatrix} 16 \\ 128 \\ 128 \end{pmatrix} + \frac{1}{256} \begin{pmatrix} 65.738 & 129.057 & 25.06 \\ -37.945 & -74.494 & 112.439 \\ 112.439 & -94.154 & -18.285 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

שימוש נפוץ במרחב זה נעשה בתחום של דחיסת תונות, זאת מפני שבמרחב YCbCr  
הקורלציה בין הקואורדינאות קטנה יותר מאשר הקורלציה במרחב RGB. רכיב Y  
מכיל את רוב המידע (מידע על בהירות), בעוד Cb ו – Cr – טומנים בחובם פחות מידע  
(מידע על מאפייני הצבע).

ב – MATLAB קיימת הפונקציה `rgb2YCbCr(I)` הממירת תמונה מייצוג במרחב RGB  
לייצוג במרחב YCbCr ונעשה בה שימוש בפרויקט.



## בעיית ה-INTERLACE

מסיבות ההיסטוריות של שימוש בטכנולוגיה ישנה, מצלמות הוידיאו המיעדות לפורמט טלוויזיה מצלמות 50 שניות בשניתה. כל שני שנות ממוגנים לפריים בודד, כאשר השורות הזוגיות שייכות לשדה אחד והאי זוגיות לאחר, כך שלבסוף מתקבל וידיאו בקצב של 25 פריימים בשניתה. מיזוג שני שנות עוקבים בדרך כלל לא נראה לעין האנושית בעת צפייה בסרט, אך כאשר מתבוננים בפריים בודד ניתן להבחין בשני השדות השונים. תופעה זו נקראת Interlace והיא פוגעת באיכות הסרט במיוחד כאשר האובייקט המצלום נע במהירות, מכיוון שבמקרה זה השוני בין כל שני פריימים עוקבים בולט יותר לצופה.

קיים מספר פתרונות לבאג' Interlace, כמו למשל פתרונות המבוססים על ביצוע אינטראפלציה או דצימציה לסרט הוידיאו. בפרויקט זה הוחלט לפתור את הבעיה על ידי הקטנת הרזולוציה של הפריים פי 4, תוך שימוש בשדה אחד בלבד, ככלומר ויתור על שורות ועמודות זוגיות/אי-זוגיות.

ויתור מעין זה אمنם מקטין את הרזולוציה אך עדין שומר על יחס הממדים המקורי של התמונה, בנוסף לכך הוא מקצר את זמן הרכיצה ומכל חישובית. יש לציין כי קיימות מצלמות שבאג' Interlace לא קיימת בהן ועל כן בפרויקט ניתן לחפש בחירה האם להשתמש בפתרון זה.

באירור 13 רואים את בעיית ה- Interlace ואת הפתרון המתתקבל לאחר הקטנת הרזולוציה פי 4.



איור 13 – תמונה עליונה לפניו תיקו בעיות ה – Interlace, תמונה תחתונה לאחר הקטנת רזולוציה.



### פעולות על תמונה מסכה

לאחר שמספרים תמונה הפרשים נתונה לתמונה בינהית של '0' ו-'1', מתקבלת תמונה מסכה אשר יש צורך בהעליה עיבוד נוסף על מנת לנחות רעשים, להפטר ממידע מיותר ולהשלים מידע חסר. בצהורה זו התקבלת תמונה מסכה שתתאר את מיקום האובייקט באופןמדויק יותר.  
לצורך עיבוד זה בוצעו על תמונות המסכות מספר פעולות.

### פעולות לצורך ניקוי רעשים ומידע מיותר

סרטוי וידיאו לרוב אינס נקיים מרעשים, כמו למשל רעש גאושי אדטיבי, רעש "פלפל ומלח" וכדומה. בנוסף, לעיתים הסף שנקבע למיפוי תמונה השחור לבן נזוק מדי ומtooסף מידע מיותר.

על מנת להתגבר על הבעיה הלו נכתב הפונקציה `clean_noise_param`. הפונקציה זו מקבלת כפרמטרים את תמונה המסכה אותה רוצים לעבד ופרמטר סף השולט על עוצמת הניקוי. הפונקציה מוצאת את כל קבוצות הקשרות בתמונה המסכה (על ידי פונקציית `bwlabel` Matlab), מתייגת כל אחת מהן, ומסירה את אלו שגודלה איננו עולה על גודל הסף שנקבע.

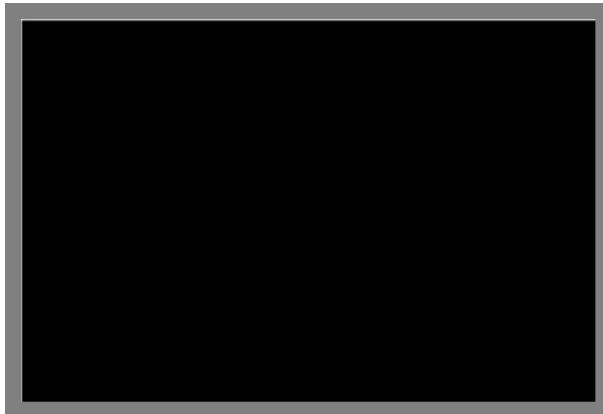
במקרים מסוימים בתמונה מסכה של האובייקט מתקבלת שפה "משווננת", כלומר כביכול שפת האובייקט מוקפת "שפיצים" לא טבעיות אשר בולטים לעין האנושית ויוצרים תחושה לא נעימה של צורה שבורה. על מנת להחליק את השפה ולהפכה ליותר טבעיות ואמיתית נעשה שימוש בפעולת החץון הדו-ممדיית (`medfilt2`) ובפונקציה המורפולוגית `imclose`.



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



באיור 14 מודגמת הורדת רעש מפריים שאין בו אובייקט להדגשת פועלות ניקוי הרעש.



איור 14 – מימין התמונה לפניה פועלות ניקוי הרעש, ומשמאלו אחראית ללא אובייקט.

באיור 15 מודגמת הורדת רעש מפריים שיש בו אובייקט להמחשת פועלות ניקוי הרעש.



איור 15 – מימין התמונה לפניה פועלות ניקוי הרעש, ומשמאלו אחראית, הפעם עם אובייקט



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



באיור 16 מופיעה המחזה של פועלות החציון ושל פועלות *imclose* לצורך החלקת שפת האובייקט המתתקבל במסכה. החלקת השפות מתבצעת על התמונה השמאלית המופיעה באיור 15.



איור 16 - הפעלת הפועלות על התמונה השמאלית המופיעה באיור 15 נותנת את תמונה זו, אשר בה ניתן לראות כי השפות הוחלקו למטען מראה טבעי יותר.



### פעולות לצורך השלמת מידע חסר

בשתי הדוגמאות שראנו בפרק קודם, ניתן לראות כי במקרים רבים ניתן לתקן את החסרים באמצעות פונקציית `imreconstruct`. במקרה הנוכחי, לעומת זאת, יש לנו רק דוחה (background) ונקודות אובייקט (object points), ללא מידע על צורה או גודלה של הנקודות. במקרה כזה, ניתן להשתמש בפונקציית `imfill`, שמיושמת ב MATLAB כ`imfill(bw, 'holes')`.

הפעולה שמבצעים היא פשוטה: מילוי כל חור בדוחה (background) באמצעות הנקודות האובייקט. בפועל, הנקודות יוכנסו למשתנה `holes`, ופונקציית `imfill` תמליל את החור. בסיום הפעולה, נקבל דוחה מילויים מלאים.

הפעולה מוגדרת ב MATLAB כ`filled = imfill(bw, holes)`.

#### 1.

לעתים הבעיה מתבטאת בקבלת אובייקט שיש בו חורים. בתוכנת Matlab קיימת פונקציה מורפולוגית בשם `imfill`, אשר מקבלת את הפרמטר `'holes'` וממלאת כל קבוצת פיקסלים שחורים המוקפת בפיקסלים לבנים. שימוש בפונקציה זו אינו מהוות פתרון מספק, זאת משומש שבנוסח להשלמת המידע על האובייקט, היא עלולה להוסיף מידע נוסף, להיות ולא כל קבוצת פיקסלים שחורים המוקפים בפיקסלים לבנים היא קבוצה נוספת מעוניינית לסוג חלק מהאובייקט. למשל אדם שם ידים על המותניים. מעבר לכך, כאשר אובייקט כלשהו נחתך על ידי גבולות המסגרת כך שהוא אינו מופיע בשלהותו בפריים, והחורים שנוצרים כוללים בתוכם את הפיקסלים של הגבולות, הפונקציה לא ממלאת את החורים הללו.

לכן, על מנת לספק פתרון הולם יותר להשלמת המידע החסר נכתבת פונקציה בשם `fill_holes_param`. פונקציה זו מקבלת כפרמטרים את תמונה המסכה אותה רוצים לעבד ופרמטר סף השולט על עצמת מילוי החורים. הפונקציה מוצאת את כל קבוצות הקשיירות של האפסים בתמונה המסכה, מתייגת כל אחת מהן, וממלאת ב'1' את אלו שגודלה אינו עולה על גודל הסף שנקבע. צורת הסתכלות מקבילה על פונקציה זו היא כלל ביצוע פעולה `clean_noise_param` על נגטיב המסכה, וביצוע נגטיב על התוצאה.



יתרונותיה של פונקציה זו על פני פונקציית `imfill` הוא בכך שהיא משתמש יכולה לשנות בגודל החורים שימולאו. חורים גדולים בדרך כלל לא נגרמים כתוצאה ממיידע חסר אלא הם חורים טבעיים ש"שייכים" לאובייקט. מסיבה זו גודל החור הוא מدد טוב להבחנה בין חורים שיש למלא בין-Calotte שלא. בנוסף, פונקציה זו מסוגלת למלא חורים גם אם הם "נחתכו" על ידי גבולות המסגרת.

הדבר מומחש באיור 17.



איור 17 - תמונה מחרורת של אדם, כאשר אחד החורים הוא בגבולות המסגרת.



איור 17 – מימין התמונה לאחר שהופעלה עלייה פעולה Imfill ומשמאל לאחר שהופעלה עלייה פעולה fill\_holes\_param. ניתן לראות איך פעולת ה fill\_holes\_param הtagbraה גם על חורים בגבולות המסגרת, היכן שפעולת ה fill\_imfill נכשלה.

## .2.

לעתים הבעיה מתבטאת ב"חצאי חורים", כולם חורים שנמצאים גם על שפת האובייקט ולא מוקפים כולם בפיקסלים בעלי ערך '1'. המחשה של "חצאי חורים" מופיעה באיור 18.

היות ו"חצאי חורים" אלו אינם מוקפים בערכי '1' וAINS נחתכים על ידי גבולות המסגרת, יש צורך לסגור אותם בדרך אחרת על מנת שייה ניתן למלא אותם בעזרת הפונקציה fill\_holes\_param.

ניתן להשתמש בפונקציות מורפולוגיות כגון: imerode, imdilate, imclose. בפרויקט נעשו שימוש בפונקציות אלה אולם הן סיפקו פתרון חלקי ולא מלא וכן היה צורך ביצוע פעולות עריכה נוספת.



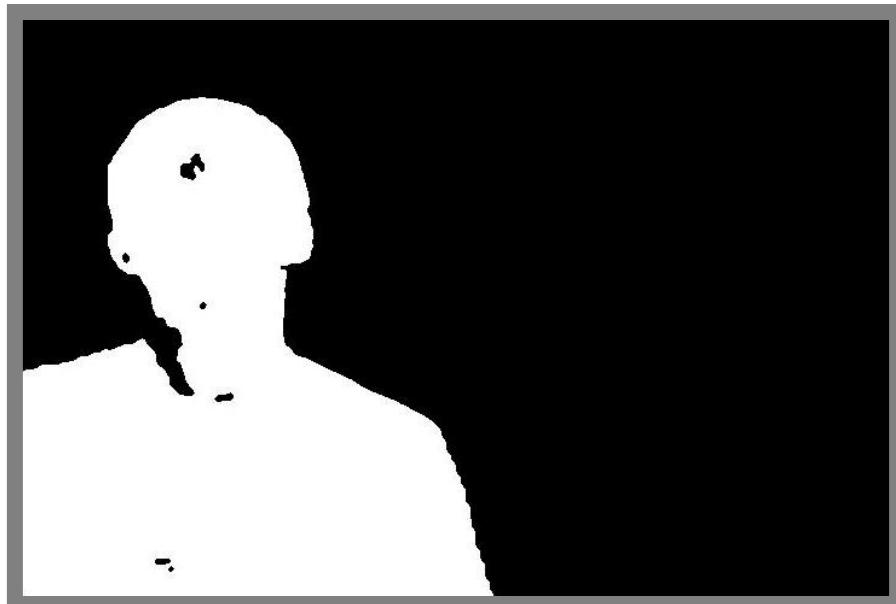
קייםית יתרונות זמניים בין פרוייקטים עוקבים, יתרונות זו נובעת מtàדר הדגימה הגבוה של מצלמת הוידיאו ויכולת לשמש כלי עזר להשלמת המידע החסר. הוחלט לנצל את היתירות הזמנית לצורך שערוך המסכה בפריים מסויים, תוך הסתמכות על המסכות מהפריים הקודם ומהפריים העוקב לו. ההנחה היא כי האובייקט לא משנה את כיוונו תנועתו פעמיים בשלושת הפרוייקטים הללו, אחרת יתוסף מידע כוזב.

שערוך המסכה מתבצע באופן הבא:

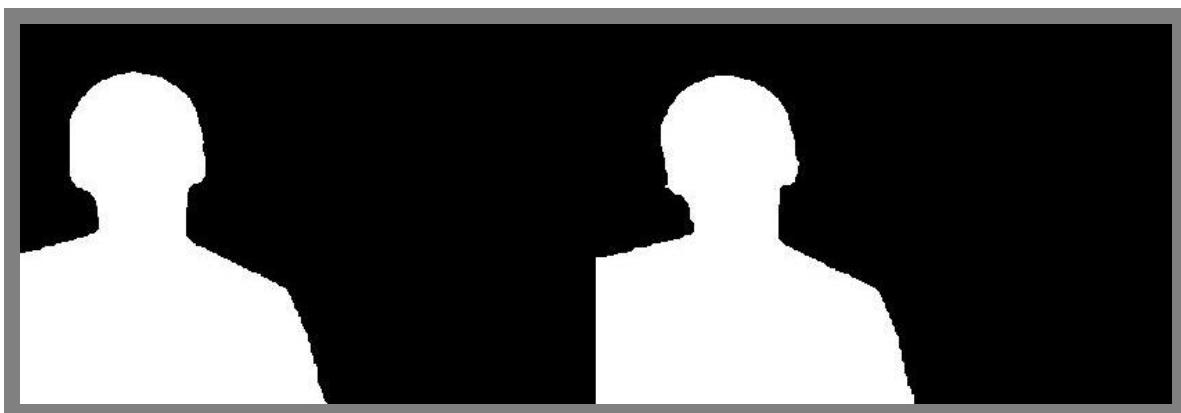
תחילה מכפילים איבר-איבר בין תמונה המסכה של הפריים הקודם לבין תמונה המסכה של הפריים העוקב (ביחס לפריים הנוכחי). לאחר מכן מוצאים את מרכזה המסכה עבור כל אחת מתמונות המסכה, ממרכזים אותן סביב נקודת מרכז המסכה של המסכה הנוכחית "המחוררת" ומכפילים ביניהן איבר-איבר.

לבסוף, מוסיפים את שתי המכפלות הללו למסכה הנוכחית וכך ניתן להשלים חלק מהמידע החסר, אם לא את כולו.

המחשה של שימוש שערוך המסכה לפי שלבים מופיעה בסט התמונות באIOR 18 – התמונות מוצגות כך שנitin יהיה לראות את כל התהליך ולהבין כיצד בעצם מתווסף מידע וnewlineים "חצאי החורים" המדוברים.



איור 18 - מסכה ראשונית בה ניתן להבחן בחצי חור מצד השמאלי של הצוואר.



איור 18 - שתי המסכות, מימין המסכה העוקבת ומשמאל הקודמת, ללא תקלות.



איור 18 – מימין ניתן לראות את תוצאת השعروץ לפי קורלציה ישירה ומשמאלו את תוצאת השعروץ לפי קורלציה סביבי מרכזו המסה.



איור 18 – תמונה המסכה הסופית המתקבלת לאחר תוספות השعروוכים למסכה הראשונית וביצוע פעולות העריכה שתואורו קודם לכן.



## עבודה מעשית וניתוח תוצאות

### עקביה תוך שימוש בהתקנת בלוקים

#### בבסיס האלגוריתם

תזכורת : מסגרת נוכחית מוגדרת בתוור התמונה  $X[k,j,t] \in [1,1] \times [J,K]$  ומסגרת הייחוס מוגדרת בתוור התמונה  $X[k,j,t-1] \in [1,1] \times [J,K]$ . כמו כן, בלוק מטרה הינו בלוק אשר שייך למסגרת הנוכחית ואשר רוצים למצוא עבورو את ההתקנה הטובה ביותר במסגרת הייחוס.

שימוש העקביה בפרויקט דומה מאוד למימוש התקנת הבלוקים עבור דחיסת תמונה. בדחיסת תמונה מתבצע חיפוש של בלוק מטרה במסגרת הייחוס. כאן לעומת זאת החיפוש יתבצע בכיוון הפוך, דהיינו יתבצע ניסיון למצוא בלוק מסוים במסגרת הייחוס במסגרת הנוכחית. בלוק זה יוגדר כ"בלוק נעקב" ואיזור החיפוש הפעם יהיה במסגרת הנוכחית ולא במסגרת הייחוס.

המטרה היא לחתך שתי מסגרות - מסגרת ייחוס ומסגרת נוכחית, וMSCA של מסגרת הייחוס - אשר תוגדר בתוור "MSCA של ייחוס", ולבנות MSCA עבור המסגרת הנוכחית. העקביה מתבצע אך ורק על קבוצת בלוקים השויכת לMSCA של ייחוס ולא על כל הבלוקים במסגרת הייחוס. קבוצת הבלוקים יכולה להיות למשל רק שפת MSCA (ואז מתבצעת עקביה רק על שפת MSCA) וכיולה להיות למשל כל MSCA (ואז העקביה היא אחר כל בלוק במסגרת הייחוס).

локחים בלוק כלשהו מקבוצת הבלוקים - זה יהיה הבלוק הנעקב. מכפילים בלוק זה איבר-איבר בבלוק המקביל לו במסגרת הייחוס על מנת לחתך בחשבונו אך ורק פיקסלים המכילים מידע נכון ולהתעלם מפיקסלים שבהם יש מידע של רקע אחריו אשר איננו נכון. כל בלוק באזורי החיפוש (אשר שייך למסגרת הנוכחית הפעם) גם כן יוכפל באותו בלוק MSCA ורק לאחר מכן יתבצע החיפוש.

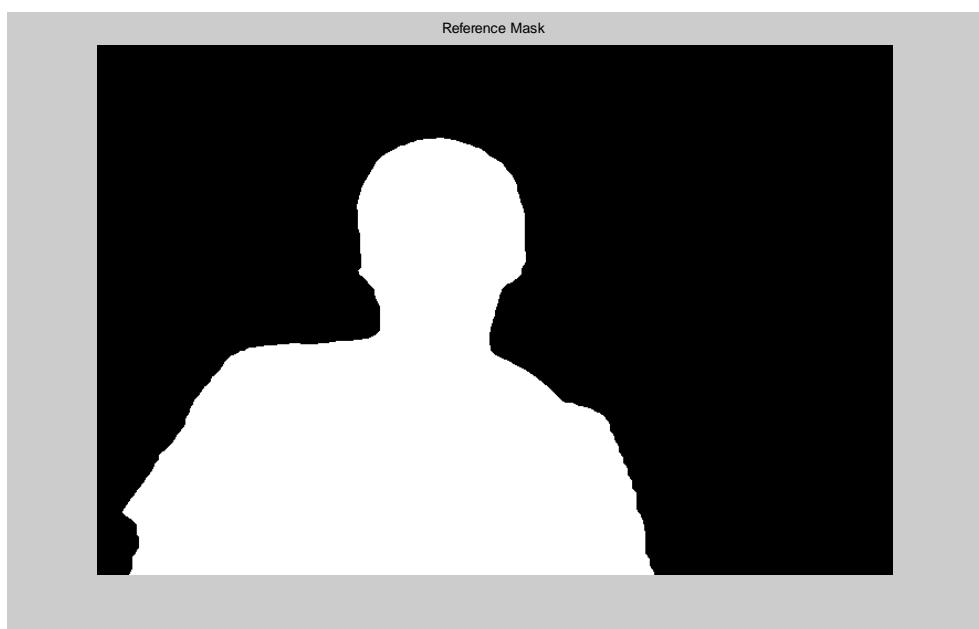
קובוצת הבלוקים נבחרה כך :



לוקחים קבוצת פיקסלים אשר שייכים למסכה ולכל פיקסל מגדריים בлок בגודל הרצוי (באמצעות בחירת גודל פרמטר  $m$ ) סביבו כך שהנקודה משתמש מרכזו המסה של הבלוק הנבחר (למעט מקרים קצה כאשר חורגים מסגרת התמונה). קבוצת הבלוקים תהיה כל הבלוקים אשר נבחרו באופן זהה. בדומה לזה ניתן כי חלק מהבלוקים יהיו בעלי תחום חופף אם הנקודות, שבסביבן הן מוגדרות, מСПיק קרובות זו לזו.

לצורך ביצוע העקבה נכתבת הפונקציה `TraceObject` בכמה גרסאות שונות, כאשר השוני בין הגרסאות הוא שיטת התאמת הבלוקים ( `Full-Search` או `N-Step` ) וקריטריון מידת ההתאמה ( מקסימום קורלציה מנורמלת או מינימום סכום הערכים המוחלטים של ההפרשיות ). משיקולי סיבוכיות הוחלט להשתמש בשיטת `N-Step` וכל התוצאות שיוצגו עבור עקבה יהיו עברו שיטה זו.

להלן תמונות של מסכת ייחוס, מסגרת ייחוס ומסגרת נוכחת:



איור 19 – מסכת ייחוס



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



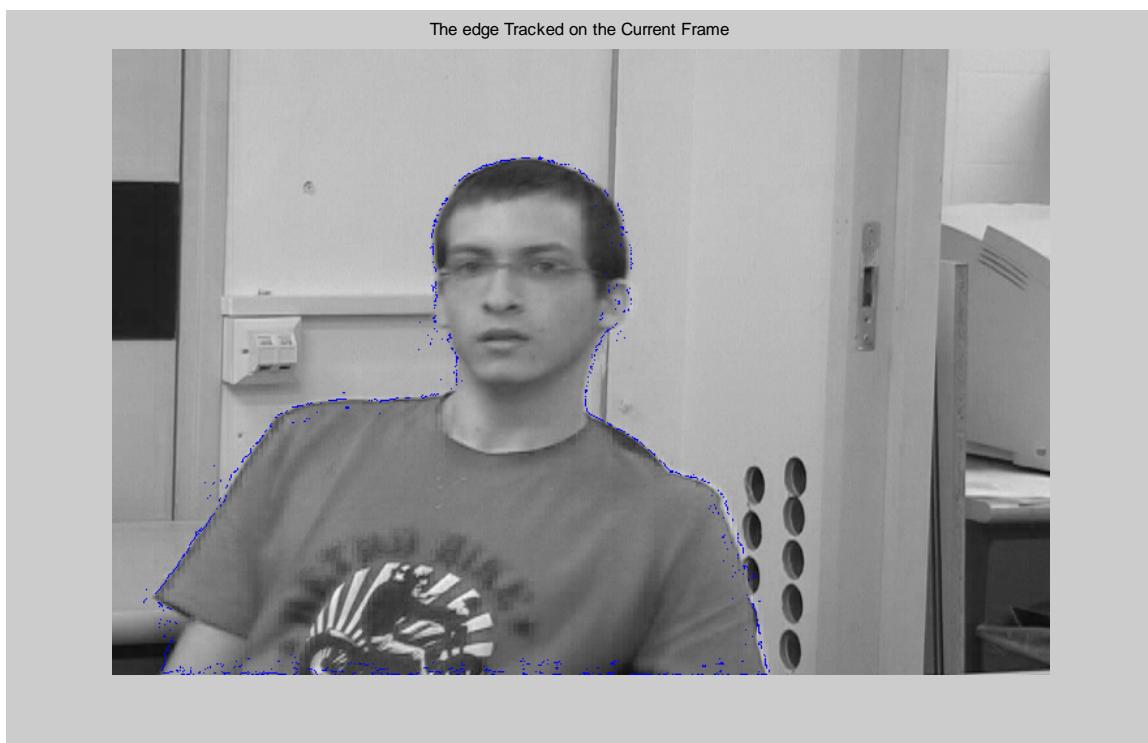
איור 19 – מסגרת ייחוס



איור 19 – מסגרת נוכחית



לאחר איטרציה אחת של עקייבת, המסכה הנוכחית המתקבלת עברו בлок בגודל 7 פיקסלים ואיזור חיפוש בגודל 16 פיקסלים, עברו קבוצת בלוקים שנלקחה מנקודות השפה של המסכה, היא :



איור 20 - הנקודות בכחול הינן נקודות השפה לאחר שיירוץ תנועתן ממסכת הייחוס, ממסגרת הייחוס ומהמסגרת הנוכחית. כאן מוצגות הנקודות על גבי התמונה הנוכחית וניתן לראות שאכן הן מתyiישבות ברובן על השפה.

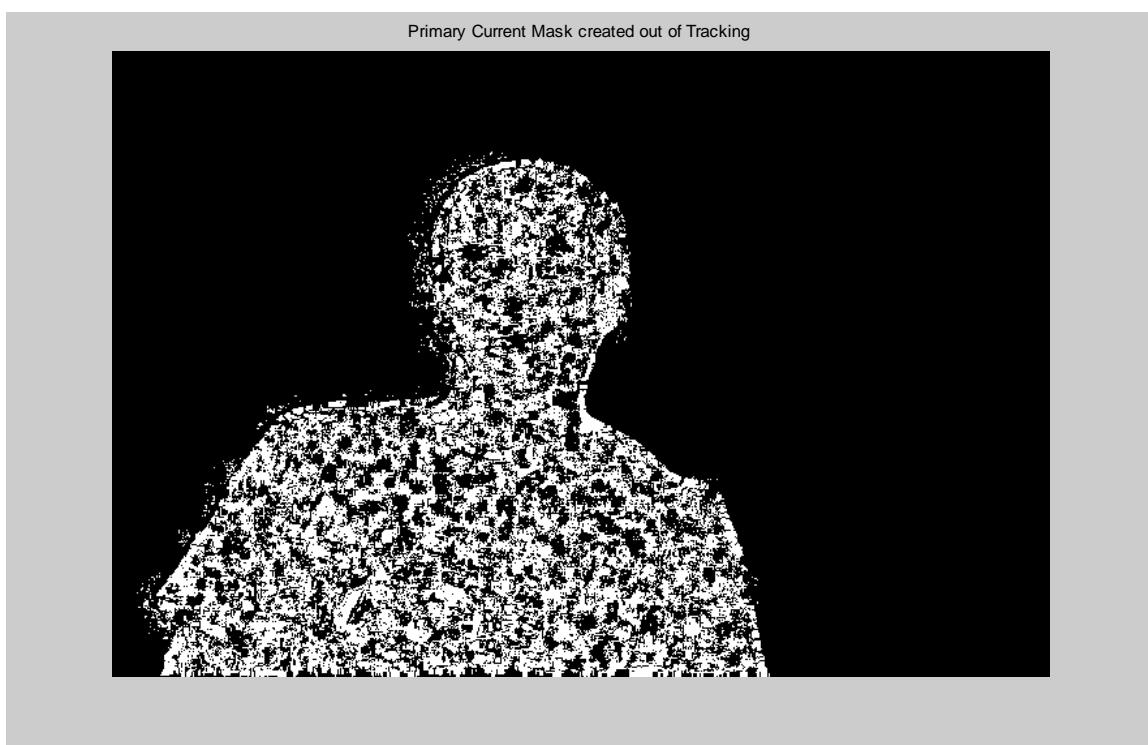
ניתן לראות שהעקביה אחר השפה היא טובה, אולם אין זה מספיק כדי למלא את תזוז האובייקט וליצור מסכה שלמה, זאת מפני שהשפה שהתקבלה היא מדוילת מדי. לכן בשלב הבא בוצע ניסיון לעקוב אחר כל הנקודות במסכה.



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



לאחר איטרציה אחת של עקיבה, המסכה הנוכחיית המתקבלת עברו אוטם פרמטרים (גודל בלוק וגודל איזור חיפוש), אבל עברו קבוצת בלוקים שנלקחת הפעם מכל הנקודות במסכה, היא:



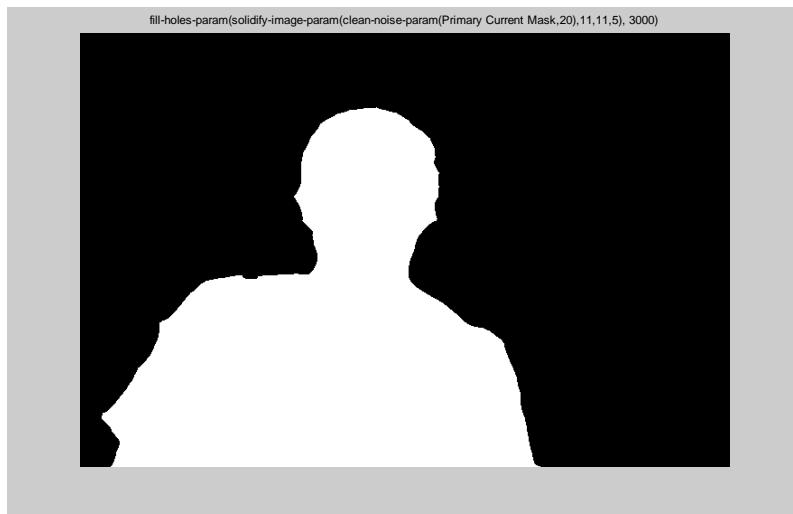
איור 21 – עקיבה אחר כל הנקודות במסכה



**הטכניון - מכון טכנולוגי לישראל**  
**הפקולטה להנדסת חשמל**  
**המעבדה לחקור הראייה ומדעי התמונה**



לאחר ביצוע של ערכות על התמונה, כפי שהובր בפרק עיבוד תמונה מתתקבלת המסכה הנוכחית הבאה :



איור 22 – עריכה על עקיבה אחר כל הנקודות במסכה

אם נבודד את הרקע הקדמי (האדם באירועים) מהמסגרת הנוכחית וממסגרת הייחוס להשוואה נקבל את התמונות הבאות :



איור 23 – חוטכים את הדמות ממסגרת הייחוס



איור 23 – חוטכים את הדמות מהמסגרת הנוכחית

נלקחה מסכת ייחוס טובה, אולם לא מושלמת ( ישנה גליישה קלה אל מעבר לשרוול החולצה הימני ). מעט קשה להבחן ויזואלית בהבדלים אולם אם מתבוננים בקצת התחthon השמאלי של המסגרות ניתן להבחן בתנועה.

המסכה החדשה שנוצרה הינה טובה מאוד ושרורץ התנועה עובד יפה ברוב חלקיו הגוף (ראש והחלק השמאלי של הגוף למשל) אולם במקומות בהם מסכת הייחוס לא הייתה מושלמת וככללה מידע כוזב של רקע אחורי (סיווג שגוי) המסכה הנוכחית שנוצרה כללה את המידע השגוי שוב ואף הרחיבתה אותו. בנוסף ישנו מקומות שבהם המסכה שנוצרה לא הייתה צמודה לחלוتين לאובייקט.



המסקנה היא שהשימוש בעקביבה הניל לאורך מספר איטרציות רצופות יגרור שגיאה מצטברת שתהייה בולטת לעין המתבונן כפי שמודגם באיוור 24.



. איור 24 – המחתה השגיאה המצטברת לאחר חמיישה פרויימרים עוקבים.

המסקנה הבלתי נמנעת אם כן היא שהאלגוריתם הבסיסי בתווך עקביבה איננו מספיק לבדו, להיות ושגיאה מצטברת כה משמעותית לאחר חמיישה פרויימרים בלבד היא בלתי נסבלת ויש צורך לבצע תיקון לאלגוריתם.

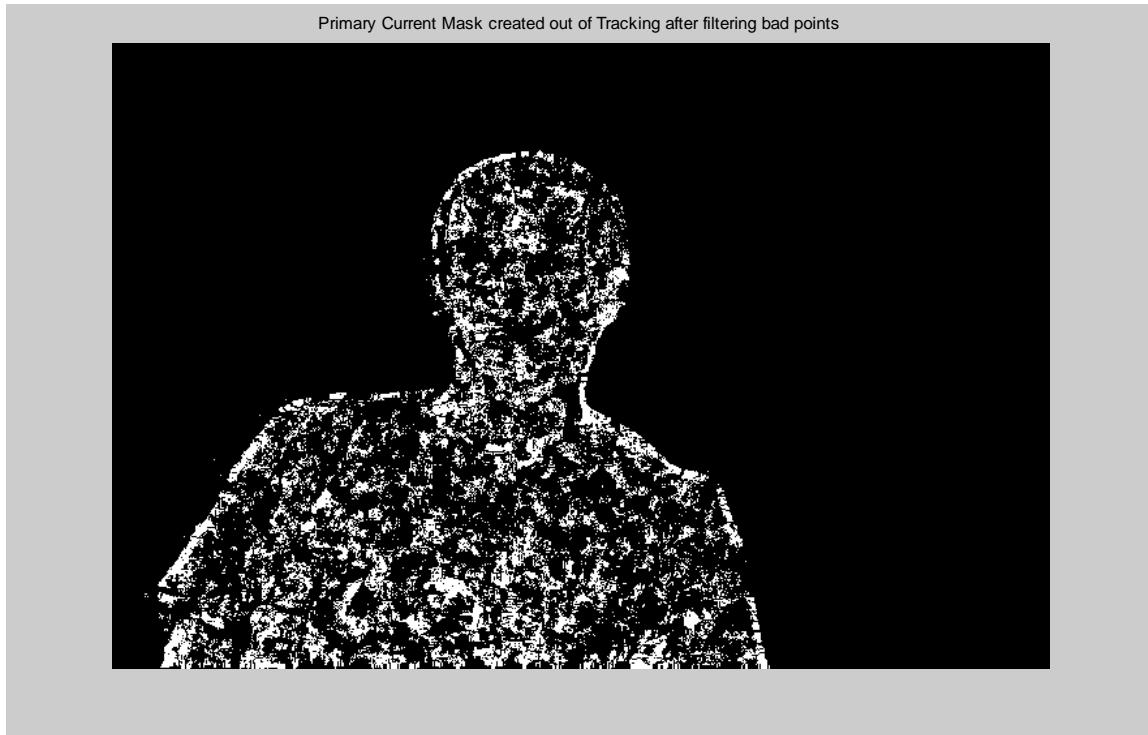


### ביקורת וסינון הנקודות הנעקבים

בגישה הבסיסית שהראיינו נוצרה בעיה של שגיאה מצטברת והמודל לא עבד בצורה טובה. על מנת לשפר את המודל ולנסות ולהפחית את השגיאה המצטברת, הוחלט להעביר את הנקודות דרך מסנן אשר נותן ביקורת על הנקודות ומוריד את הנקודות אשר הביקורת עליהם היא שלילית.

קריטריון הביקורת שנבחר בפרויקט הוא לקיחת נקודה, הגדרת סביבה עבורה מתחן נקודות שייכות למסכת הייחוס בלבד, חישוב ממוצע וקטורי התנועה המתקבלים עבור כל הנקודות בסביבה, חישוב ממוצע המרחקים של וקטורי התנועה מהממוצע חשוב, ולבסוף החלטה על הסרת הנקודה המקורית במידה וממוצע המרחקים עולה על סף מסוים. בפועל, מעיפים נקודות שהסיבה הקרובה שלן לא נעה בקירוב כמקשה אחת. להיות ובפרויקט מדובר בתמונות טבעיות שרוב המידע שלן מרוכז בתדרים נמוכים, נקודות שסועגו לא נconaה בתור רקע קדמי, הסטייה שלן מהממוצע צפואה להיות גדולה יותר (בגלל האקרואיות) וכך ניתן להבדילן מנקודות שסועגו נconaה ולהקטין את השגיאה המצטברת.

תוצאת העקיבה אחר כל הנקודות במסכה, לאחר סינון נקודות עליהם הביקורת שלילית, מופיעה באIOR 25.

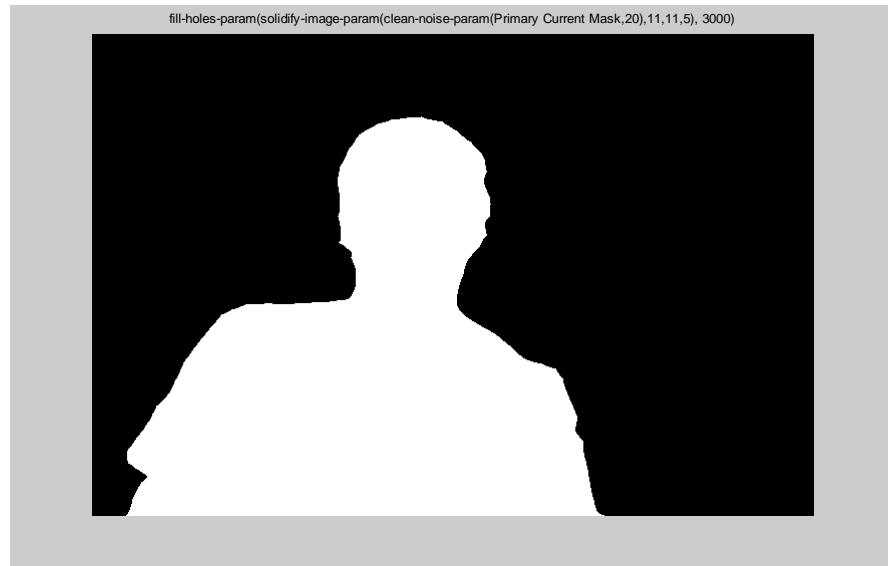


איור 25 - המסכה הראשונית המתקבלת לאחר עקיבה ולאחר סינון הנקודות. ניתן לראות בהשוואה לאיור 21 ש מרבית הנקודות מסביב לאובייקט, אשר גולשות אל מעבר לשפה ושיכוכות לרקע האחורי מסוננות לאחר הביקורת ומתקבלת מסכה שבצורתה הכללית טובת בהרבה.

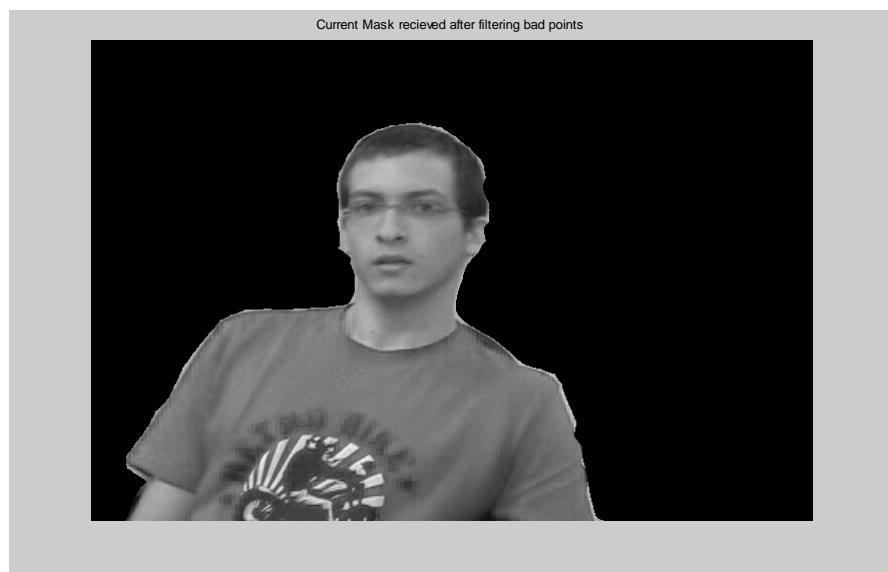
ניתן לראות שהמסכה הראשונית שהתקבלה השתפרה משמעותית לאחר סינון הנקודות הלא אמינות.-cut ניתן לבצע ערכיות על המסכה ולקבל מסכה נוכחות סופית טובת יותר. המסכה הסופית שהתקבלה וחילוץ האובייקט מהמסגרת הנוכחית מוצגים באיור .26



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



איור 26 – המסכה הסופית שהתקבלת לאחר מתן ביקורת על נקודות וסינון נקודות עליהן הביקורת איננה טובה.



איור 26 – חילוץ האובייקט מהמסגרת הנוכחית השתפר ביחס לזה שהוצע לאיור 23.



המסכה המתקבלת לאחר סינון הנקודות טוביה יותר ומדויקת יותר. ניתן לראות למשל את השיפור בשרוול חולצת יד ימין, כאשר בעצם הרקע האחורי שסוג לא נכון נעלם. ניתן להקיש מכך כי הסינון עוזר במידה מה להתגבר על הגורמים לשגיאת המצברת ולשפר את התוצאות. מצד שני, המסכה מעט "נאכלת" במקומות מסוימים על שפת האובייקט, דבר הפוגם במידה מה בתמונה המתקבלת ואשר עלול לגרום בעצםו לשגיאת מצברת שכן חלקים שנאכלו באובייקט לא יעקבו בעתיד.

הчисIRON הגדל בתהליך זה הוא שזמן ההרצאה הפך בלתי נסבל עד כדי כך שזה לא ריאלי להריץ את האלגוריתם הנ"ל על סרטו קלשו. היה אפוא צורך באלגוריתם אחר, מהיר יותר, אשר יאפשר לשפר מסכה נוכחת.



## שיפור המasca באמצעות אלגוריתם ה- *Intelligent Scissors*

### הסבר הרעיון

כפי שהוסבר בפרק התיאורטי, אלגוריתם ה- Intelligent Scissors מאפשר מציאה חכמה של שפות בתחום נתון של התמונה. כפי שהוזכר, המasca איננה תמיד צמודה לחלווטין לשפט האובייקט. הסיבה נעוצה בכך שהספ, אשר על פיו ממירים את תמונה ההפזרים לתמונה מסכה בינהית, אינו מייצג בעקבות זאת כל התמונה. המטרה היא לשפר את המasca כך שתהייה צמודה לשפט האובייקט ולת��ן אותה במקרים בהם הצבירה שגיאה.

אלגוריתם ה- Intelligent Scissors הינו כלי אינטראקטיבי אשר מקבל אינפורמציה מהמשתמש באפליקציה באמצעות העכבר. האינפורמציה שהוא מקבל היא עקום סגור (במרחב הדו-מימדי של התמונה) אשר מכיל בתוכו את האובייקט במלואו, ובנוסף גם קואורדינאות של נקודה כלשהי בתוך האובייקט, היא נקודת ה- seed point. האלגוריתם מקבל את המידע הניל ומחזיר למשתמש מסכה המכילה את האובייקט בלבד מבודד מסביבתו.

בפרויקט זה אינטראקציה כזו בין המשתמש לבין האפליקציה מנוגדת תחילה למטרת הפרויקט להיות והיא מונעת ממנו לעבוד בצורה אוטומטית. היה אם כך צורך בפיתוח אלגוריתם המדמה את פעולה העכבר ובכךאפשר למצוא במצבה עצמאית עקום המקיים את האובייקט, ונקודות point seed.

### האלגוריתם שפותח

תחילה, נמצאת מסכה ראשונית פשוטה אשר מכילה את האובייקט בצורה סבירה. את המסכה שנמצאה מנפחים מעט ע"י פעולות מורפולוגיות של Matlab על מנת להבטיח שהיא תכיל את כל האובייקט, גם אם זה עולה בהוספה של מידע מיותר למסכה (רקע).

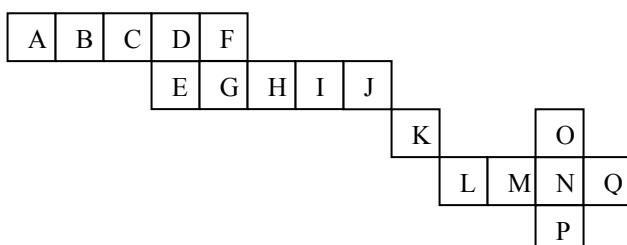


מוצאים את שפת המסכה ומתקבל עוקם אשר מכיל בתוכו את האובייקט. כמו כן, ניתן ללקחת בטור נקודת seed point את נקודת מרכז המסכה של המסכה שנמצאה, או ללקחת נקודת כלשהי אקראיית מתוך המסכה הראשונית (לפנוי הניפוח).

כאשר מקיפים את הדמות באמצעות העכבר, הנקודות נכנסות למאגר הנתונים של כלי ה - Intelligent Scissors לפי סדרן האוקלידי, כלומר הנקודות מסודרות ברצף. לכן, על מנת להשלים את פועלות הדמיית העכבר יש צורך בסידור נקודות שפת המסכה ברצף. לאחר שבוצעו כל אלה, מאגר הנתונים של כלי ה - Intelligent Scissors הושלם ומוחזרת כפלט מסכה צמודה יותר לאובייקט.

לצורך פועלות הדמיית העכבר אשר ממיינת את הנקודות לפי הסדר נכתבה הפונקציה PathArrange אשר פועלת באופן הבא :

תחילה, הפונקציה מקבלת כקלט את קבוצת נקודות השפה (המקום שלחן למרחב המיקום הדו-מימדי של התמונה) של האובייקט אשר להן שתי נקודות שכנות בדיזוק, מתוך שמונה שכניו, גם כן שייכות לשפת האובייקט. קבוצת נקודות זו תكون קבוצת הקולט. למשל באירור 27 יילקו רק הנקודות L,J,I,B. נקודות אלו אינן בהכרח מופיעות בסדר מסוין, כלומר רשימת נקודות הקולט הללו יכולה להיות מיוצגת למשל בסדר I,J,K,L,B.



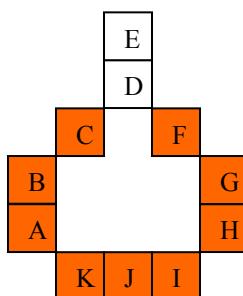
איור 27



מתוך קבוצת נקודות הקלט בוחרים נקודה שרירותית כלשהי. נקודה זו מהוות תחילתה של שרשרת נקודות אשר תהיינה בסופה של דבר מסודרות לפי מיקומן. CUT מתחילה בתחילת מיוון הנקודות שבמסגרתו בכל שלב מתווספת נקודה לשרשראת באופן הבא :

- מוצאים את השכנים של הנקודה שהתווספה לשרשראת בשלב הקודם.
- בוחרים שכן אחד באופן שרירותי, במידה והוא שייך לקבוצת הקלט והוא עוד לא מופיע בשרשראת מוסיפים אותו לשרשראת וועברים לשלב הבא.
- במידה ואותו שכן לא מופיע בקבוצת הקלט או שכבר נמצא בשרשראת בוחרים את השכן השני. אם השכן השני שייך לקבוצת הקלט והוא עוד לא מופיע בשרשראת מוסיפים אותו לשרשראת וועברים לשלב הבא.
- אם שני השכנים כבר מופיעים בשרשראת אזי השרשרת כבר מכילה את כל הנקודות בצורה ממויינת. זהו תנאי העצירה.
- במידה גם השכן השני לא מופיע בקבוצת הקלט או שכבר נמצא בשרשראת, מחפשים את הנקודה הקרובה ביותר שכן מופיעה בקבוצת הקלט ולא מופיעה עדין בשרשראת, מוסיפים אותה לשרשראת וועברים לשלב הבא.

למשל עבור קבוצת הנקודות המופיעעה באיוור 28 שרשרת הנקודות שיווחזו, לפי הסדר לאחר המיוון עבור נקודת התחלה שרירותית A תהיה : A,B,C,F,G,H,I,J,K.



איור 28 – בלבן נקודות שפה שאין שייכות לקבוצת הקלט, בכתום נקודות הקלט.



באיור 31 אשר מופיע בהמשך ניתן לראות כיצד האלגוריתם שתואר מאיר את השפה ומקיף את האובייקט.

ליבת האלגוריתם מומשא בפונקציה DrawPathNew אשר נוערת בכמה פונקציות עוזר בעצמה. הפונקציה מקבלת כקלט את המסכה הראשונית לאחר ניפוחה, ומחזירה את כל מאגר הנתונים הדרושים לפתרון הבעיה באמצעות הכליל האינטראקטיבי של Intelligent Scissors. **כלי ה- Scissors** בו נעשה שימוש נלקח מפרויקט עבר [12].

להלן תמונה המסכה הראשונית המנופחת :



איור 29 - המסכה לאחר ניפוח. את מסכה זו שולחים כקלט לפונקציה DrawPathNew במטרה לבנות את מאגר הנתונים עבור **כלי ה- Intelligent Scissors**.



ולאחר הכפלת במסגרת הנוכחית ניתן לראות אכן את ההרחבה ביחס לאובייקט ושאכן האובייקט מוכל במלואו בمسכה - איור 30.



איור 30 – הדמות אכן מוכלת במסכה המנופחת.



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה

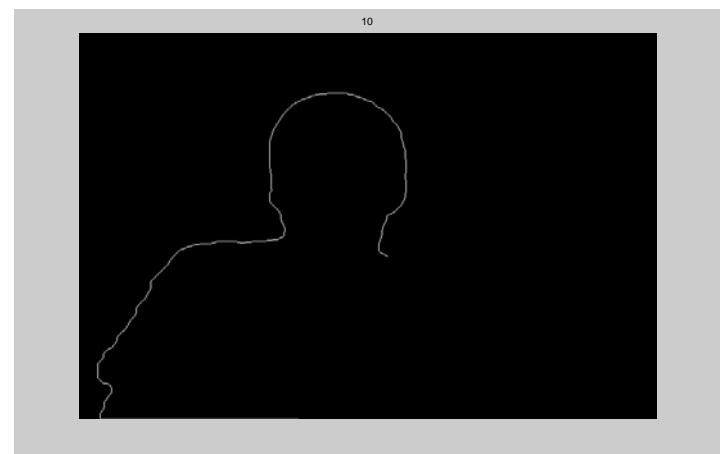
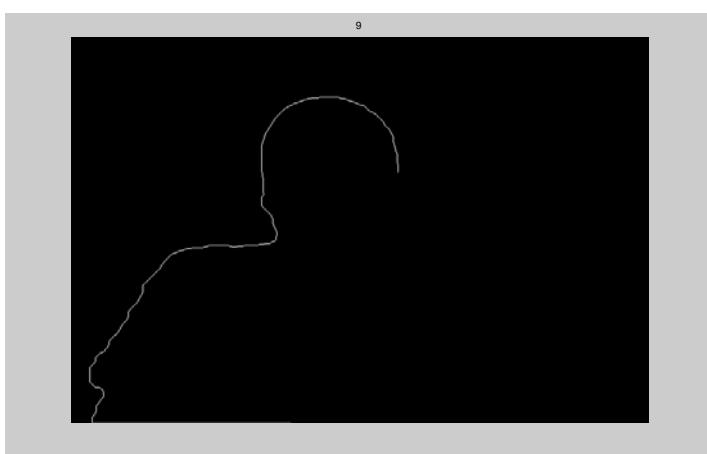
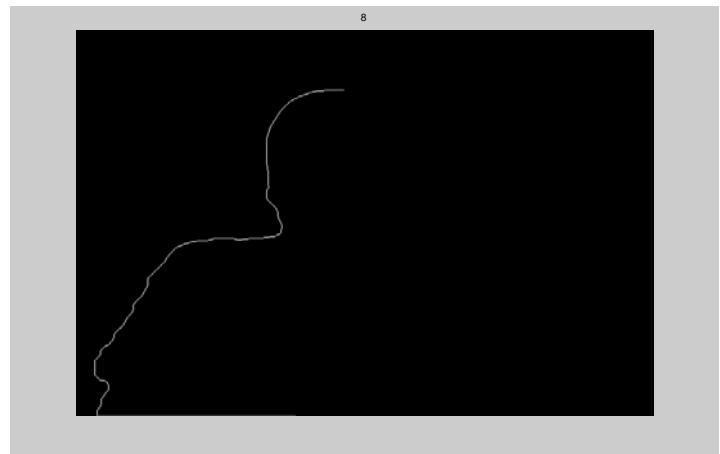
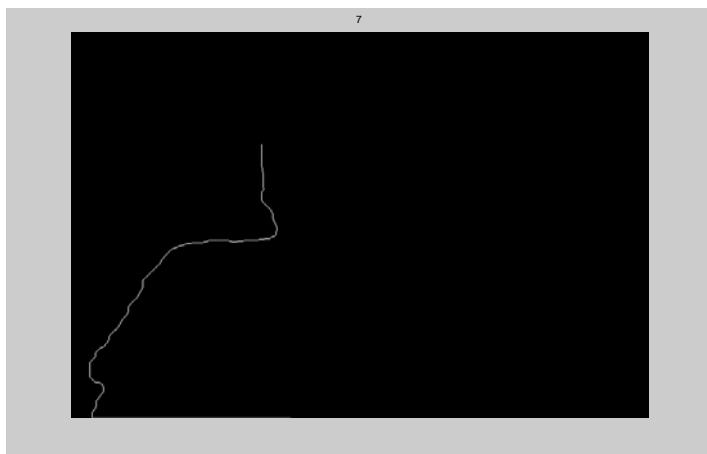


פעולת הדמיית העכבר – איור 31 ניתן לראות איך השפה מצוירת מסביב לאובייקט באופן עצמאי ורציף.



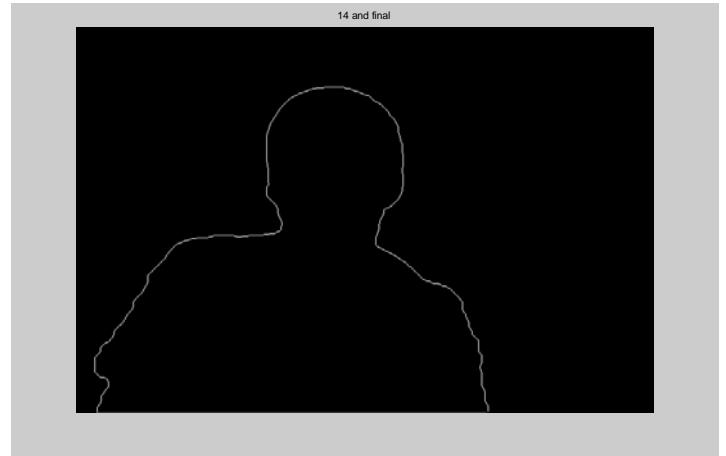
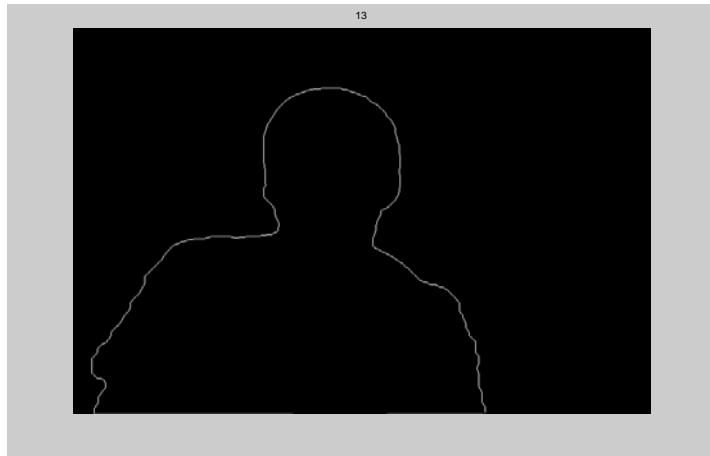
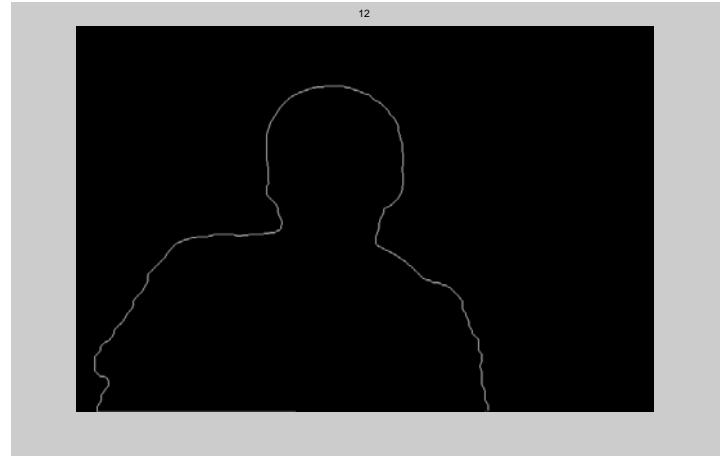
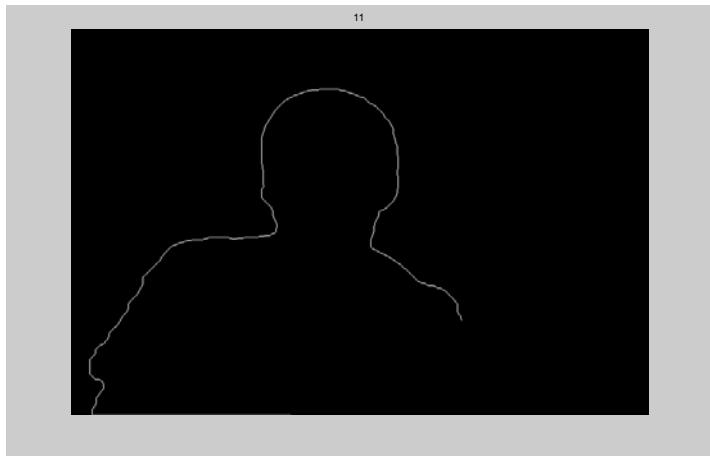


הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקירת הראייה ומדעי התמונה





הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



איור 31 – רצף של 14 תמונות מוצגות בסדר כרונולוגי להדגמת פעולה הדמיית העכבר.



באיור 32 ניתן לראות את קווי המתאר של הדמות בכחול כפי שモצא אותן אלגוריתם h - Intelligent Scissors.



איור 32 - קווי המתאר מסומנים בכחול עבור פריים מסוימים

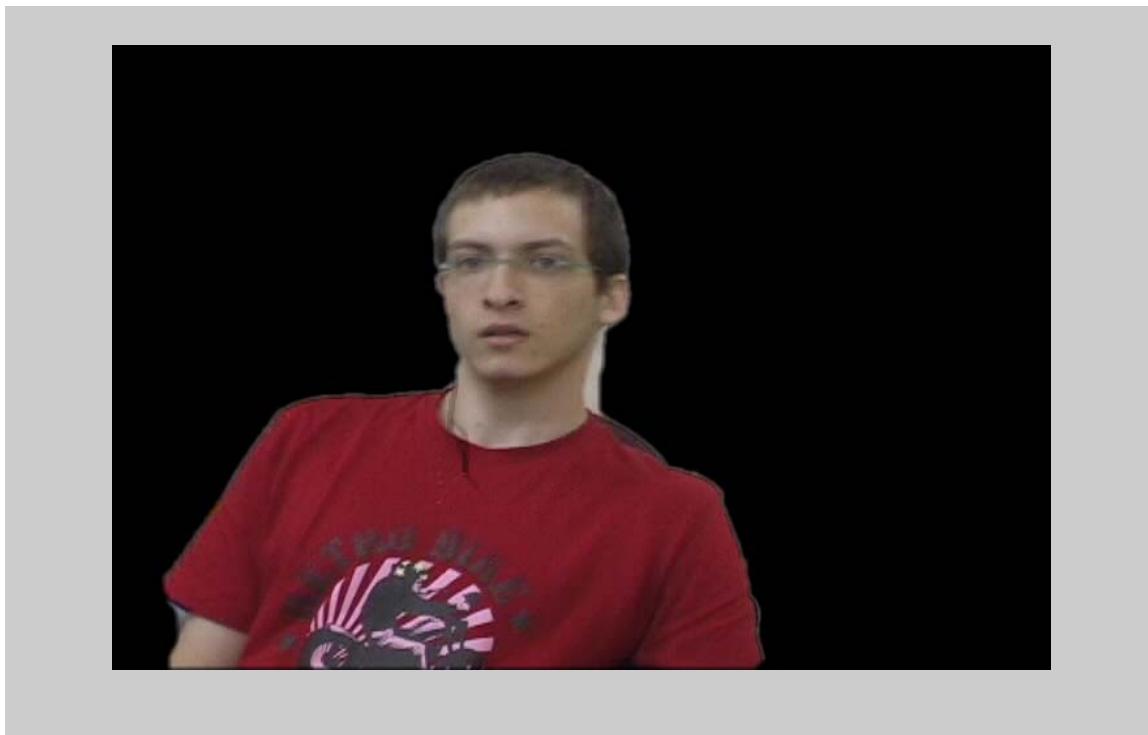
ניתן לראות כי האלגוריתם זיהה את שפות האובייקט בצורה טובה, אך במקום בו הייתה ברקע שפה חזקה הוא זיהה אותה במקום את שפת האובייקט ולכן קווי המתאר אינם מתאימים במדוק אט הדמות ועלולים להיות שיבושים כאשר הרקע האחורי מורכב ומכיל שפות חזקות בקרבת האובייקט.

בסק הכל התוצאה המתקבלת מושיטה זו הינה תוצאה טובה אך בעייתית במידה מסוימת להיות ובפרויקט זה אחת השאייפות היא להשיג ביצועים טובים ללא תלות



במורכבות הרקע האחורי, ובאלגוריתם זה לא מושגת מטרה זו. בנוסף, זמן הריצה ארוך במעט צען כדי לאייטרציה והיה רצוי במצב מודל עילית יותר מבחינת סיבוכיות הזמן.

כאשר חותכים את הדמויות מהמסגרת הקיימת מתקבלת תמונה שהיא פلت של אלגוריתם ה- Intelligent Scissors או אשר מוצגת באיוור 33



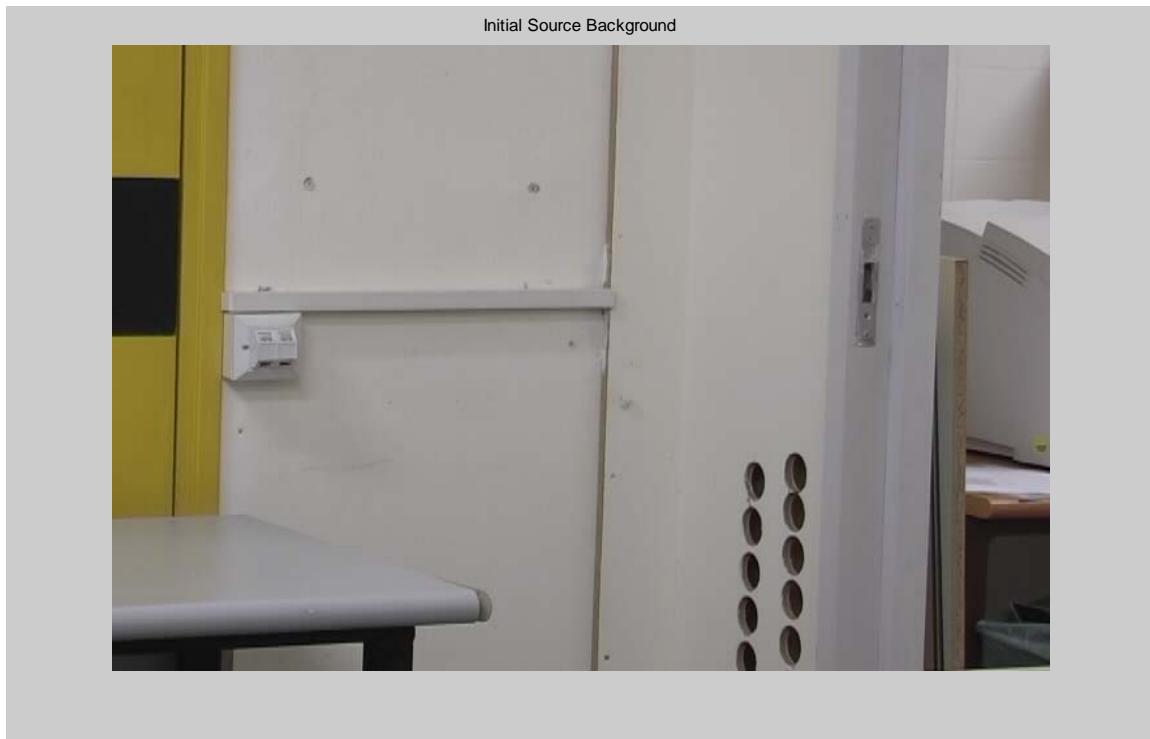
איור 33 - חילוץ האובייקט. כפי שהווטר, אכן האובייקט לא נלקח בצורה מושלמת וניתן לראות גם חלק מהרקע האחורי.



### מציאת המסכה תוך שימוש בשיטת הרקע הרץ

כפי שהוסבר בחלק התיאורטי, בגלל שינויים שתרחשים ברקע ניתן לשפר את הביצועים על ידי למידת הרקע עם הזמן. שימוש בטכניקה זו מאפשר להתגבר על בעיות שנובעות משינויים בתנאי הסביבה ובעיקר מפני שינויי בעוצמות ההארה שהמצלמה מאלצת.

תחילה, יש לאותחל תמונה רקע ראשונית. על כן יש לחת מספר פרימיום אשר בהם לא מופיע אובייקט וברקע לא יופיעו תנודות קיזזוניות ויהיה ניתן להתייחס אליו בקשר כסטטי. הפרמטר  $\alpha$  מייצג את מספר הפרימומים הראשונים הניל הנקחים לשם אתחול תמונה רקע. ניתן לראות רקע לדוגמה באירור 34.



איור 34 - תמונה שמרת את הרקע הראשוני לפי  $\alpha=9$ .



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



קצב הלימוד של הרקע נקבע לפי פרמטר  $\alpha$ , המכונה פרמטר הלימוד. פרמטר זה נקבע אמפירית לערך 0.125 והתוצאות עבورو מופיעות באירועים 35-39.



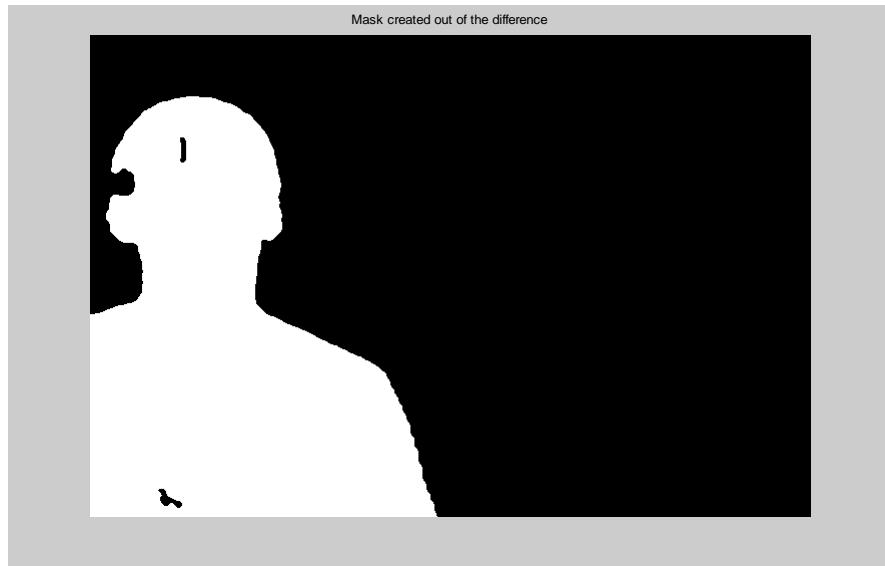
איור 35 – התמונה הנוכחית, האובייקט נכנס לתמונה.



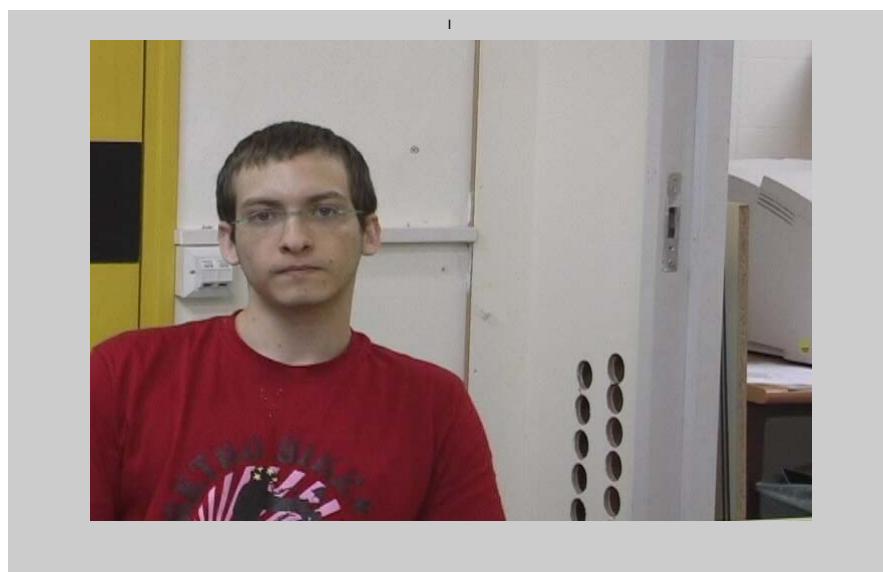
איור 35 – הרקע הנלמד כאשר האובייקט רק נכנס לתמונה



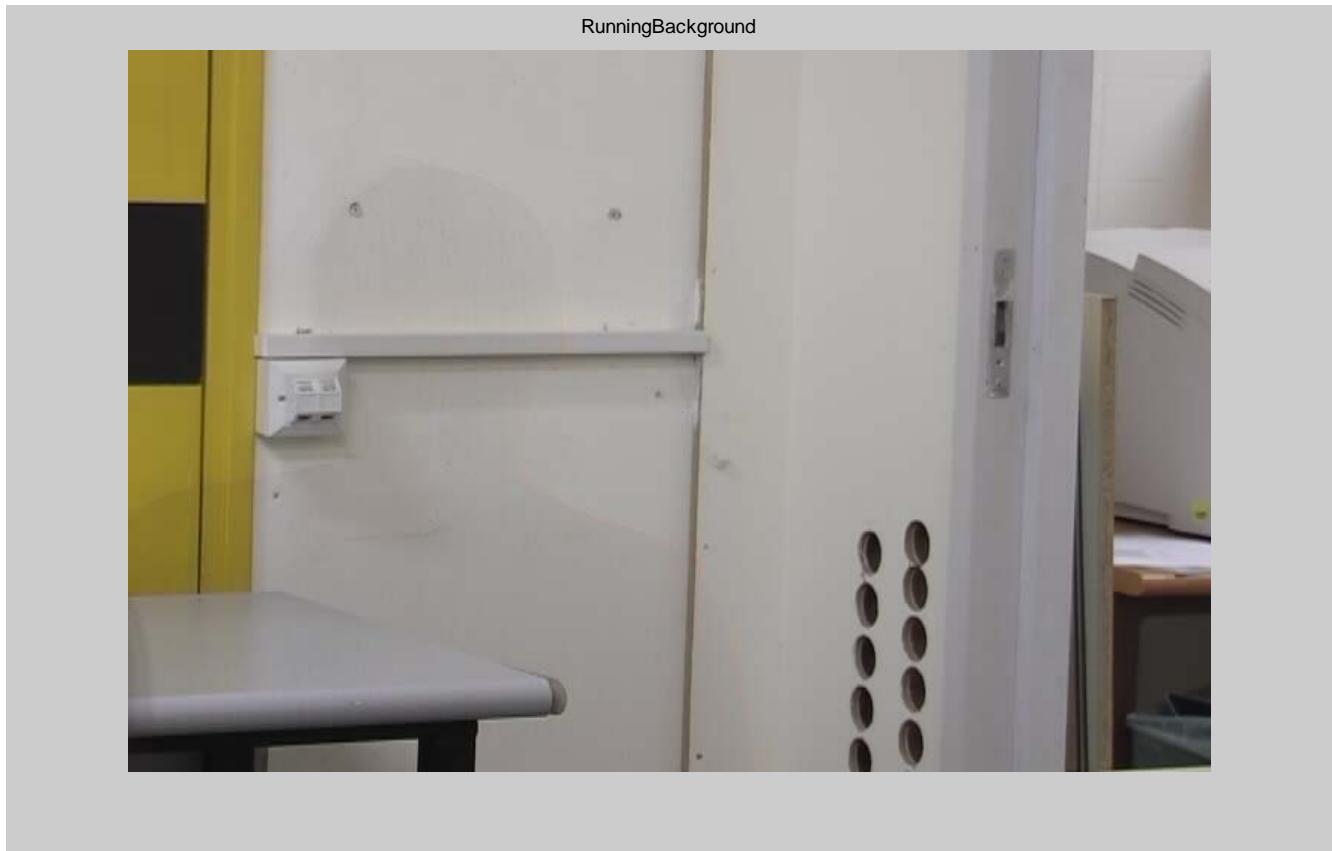
הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



איור 35 - המסכה שהתקבלה מחיסור בין התמונה לבין הרקע הנלמד. ניתן לשפר את המסכה על ידי מילוי חורים לקבלת מסכה טובה יותר.



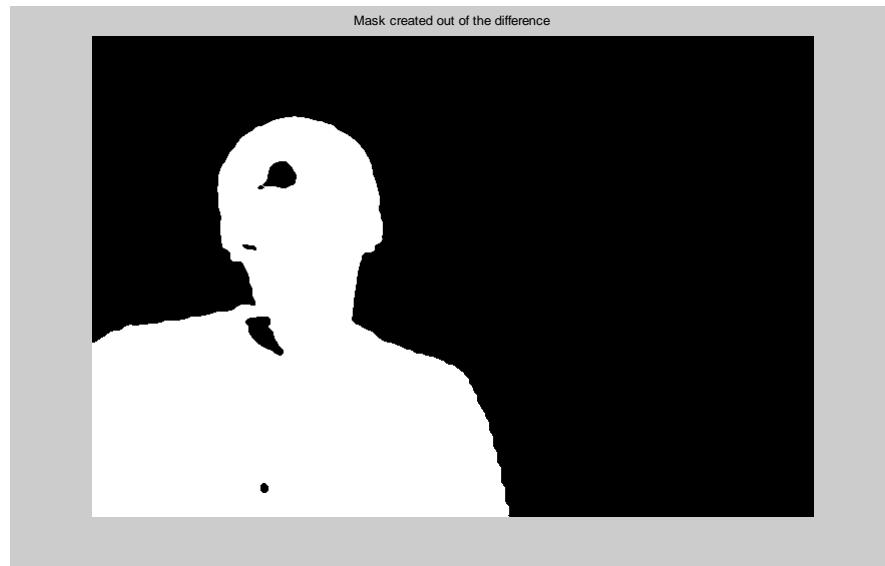
איור 36 - 20 פרויימים לאחר מכון



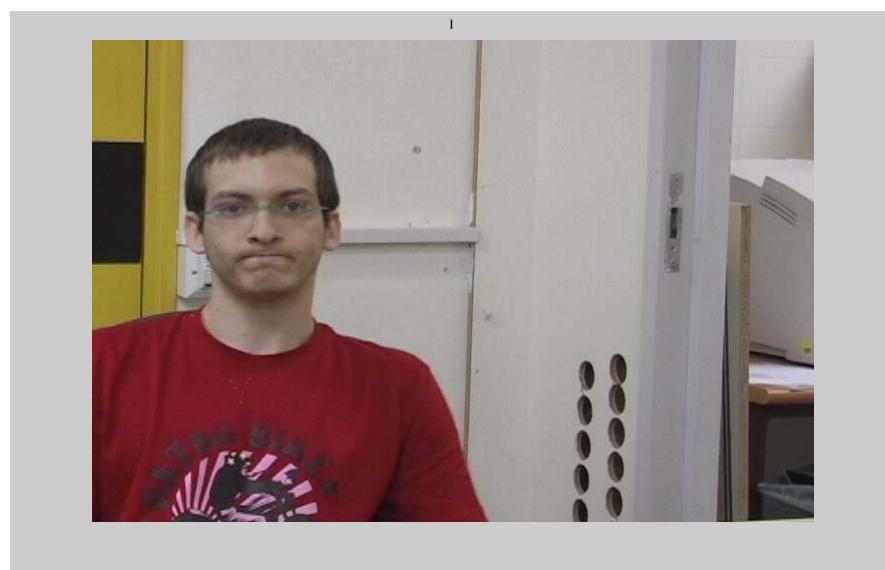
איור 36 - הרקע הנלמד לאחר 20 פרימים אלו. ניתן לראות כי אחרי פרק הזמן שעבר האזורים שנלמדו הסתגלו לשינוי התאורה, בעוד שהמקומות שסווגו כרקע קדמי ואשר בהם לא התבכעה למידה לא הסתגלו לשינוי התאורה ולכן מתkowski בתמונה "צללית" באזור בו הדמות נמצאת.



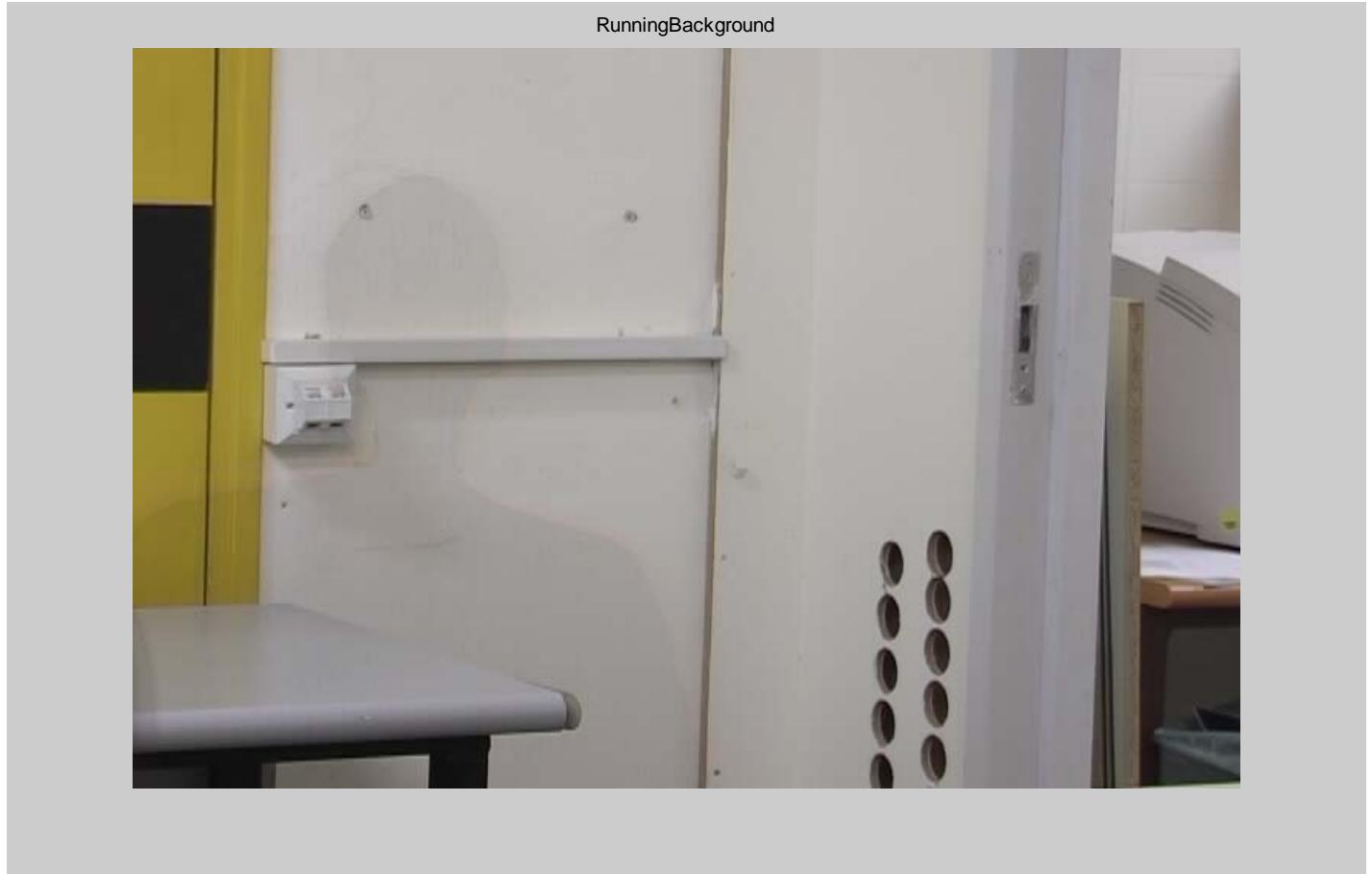
הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



איור 36 – המסכה 20 פריימים מאוחר יותר. מסכה טובה, ניתן על ידי מילוי חורים להפוך אותה למצוינת.



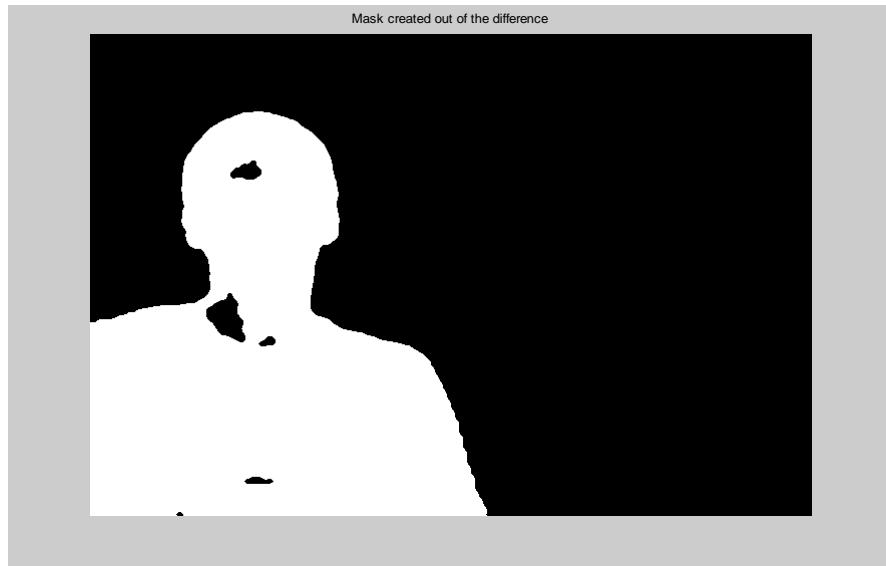
איור 37 – התמונה לאחר 20 פריימים נוספים



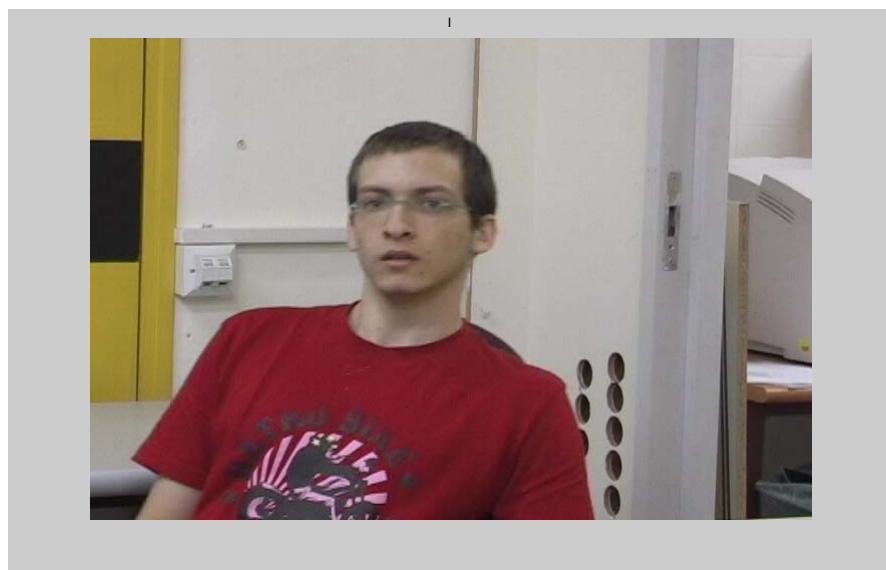
אייר 37 - הרקע הנלמד 20 פרייםים לאחר מכן. הפיקסלים שהוסתרו על ידי הדמות לפיקי זמן ארוכים יותר נלמדו פחות. ניתן לראות בתמונה זו חיזוק של הדברים שנאמרו באյור בקודם וניתן להבחן בצללית בצורה ברורה יותר.



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



איור 37 - המסכה טוביה, לאחר מילוי חורים תתקבל מסכה מדויקת.



איור 38 - התמונה לאחר 20 פריימים נוספים



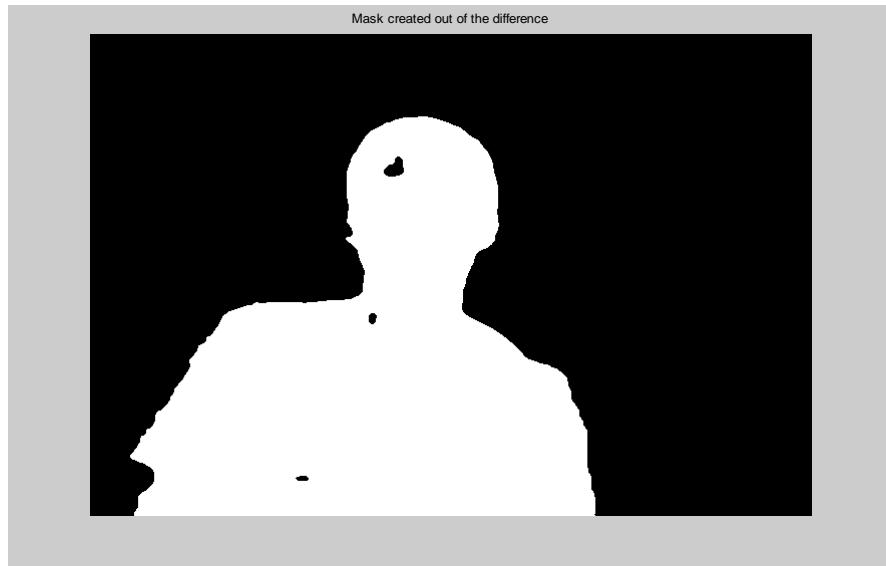
הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקירת הראייה ומדעי התמונה



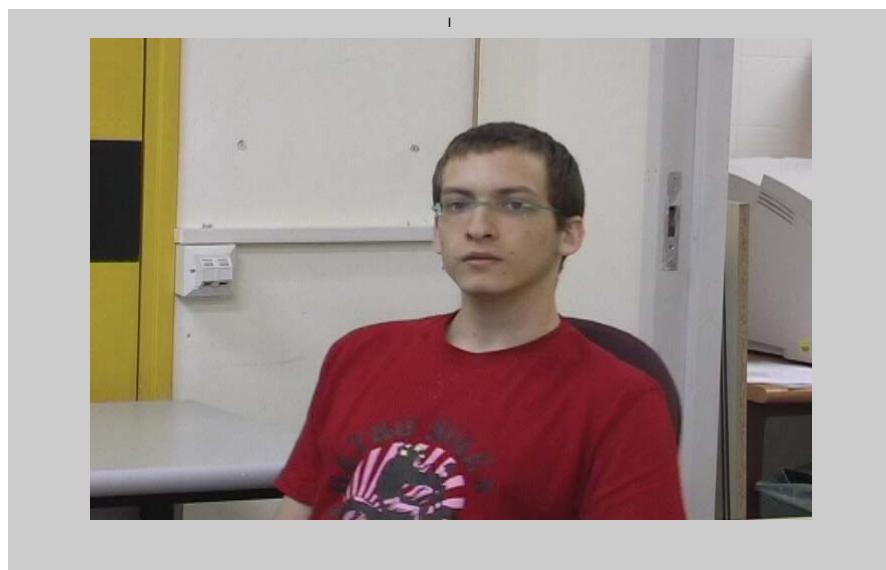
איור 38 - הראש הוזז למקום אחר ולכן האзор שהווסתר על ידיו לפניו כן נלמד ותמונה הרקע הרץ הסתגלה בחזרה לשינויי התאורה שנוצרו בפרק הזמן שחלף.



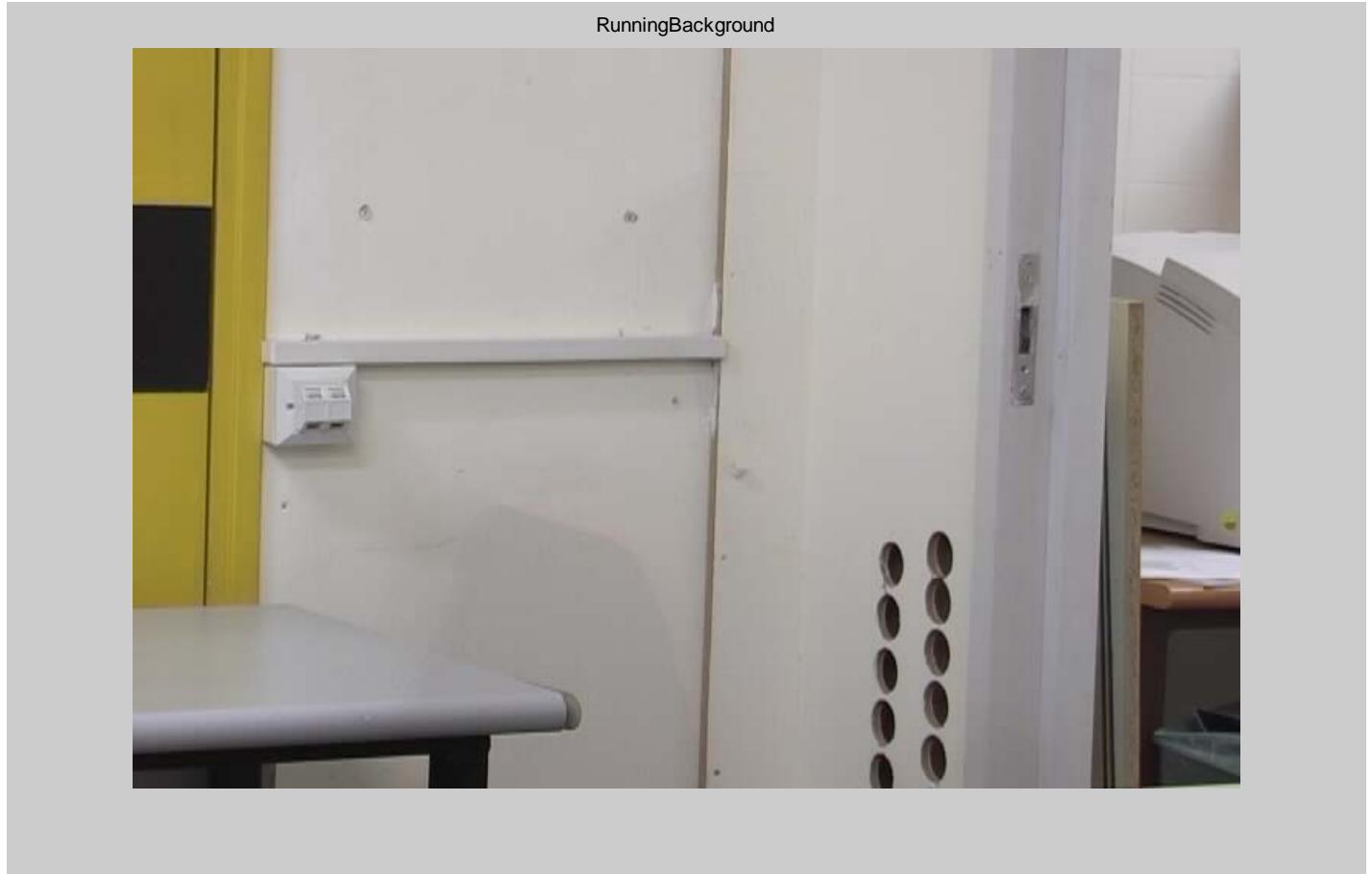
הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



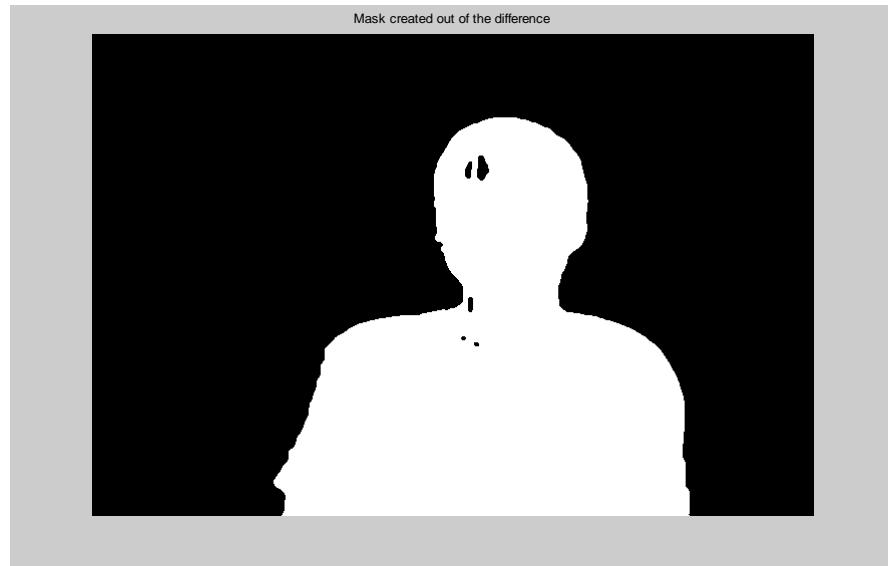
איור 38 - המסכה כפי שהתקבלה על ידי חישור בין התמונה לבין הרקע הנלמד.



איור 39 - התמונה לאחר 30 פריימרים נוספים



איור 39 - הרקע הנלמד, הפעם חלק ניכר מהגוף זו וניתן לראות בצורה בולטת את התעדכנות הרקע הרץ. באזוריים שנחשפו כתוצאה מתזוזת הגוף הצללית נעלמה, ואילו באזוריים שלא נחשפו הצללית נשארה. הדבר מדגיש את הלמידה באזוריים החשופים ואת אי-הלמידה באזוריים שאינם חשופים, כפי שצוין בפרק התיאורטי על גישת הלמידה הנ"ל.



איור 39 - המסכה שנוצרה על ידי הפחטה בין הרקע הארץ.

בז' הכל התוצאות המתתקבלות משימוש ברקע הלומד הן טובות ומשמש ניתן לראות כיצד הרקע הלומד מסתגל לשינויים ברקע האחורי ומאפשר מציאת מסכה טובה. בנוסף, זמן הריצה עבור איטרציה בודדת הינו מהיר מאוד ביחס לגישות הקודומות שהוצעו.

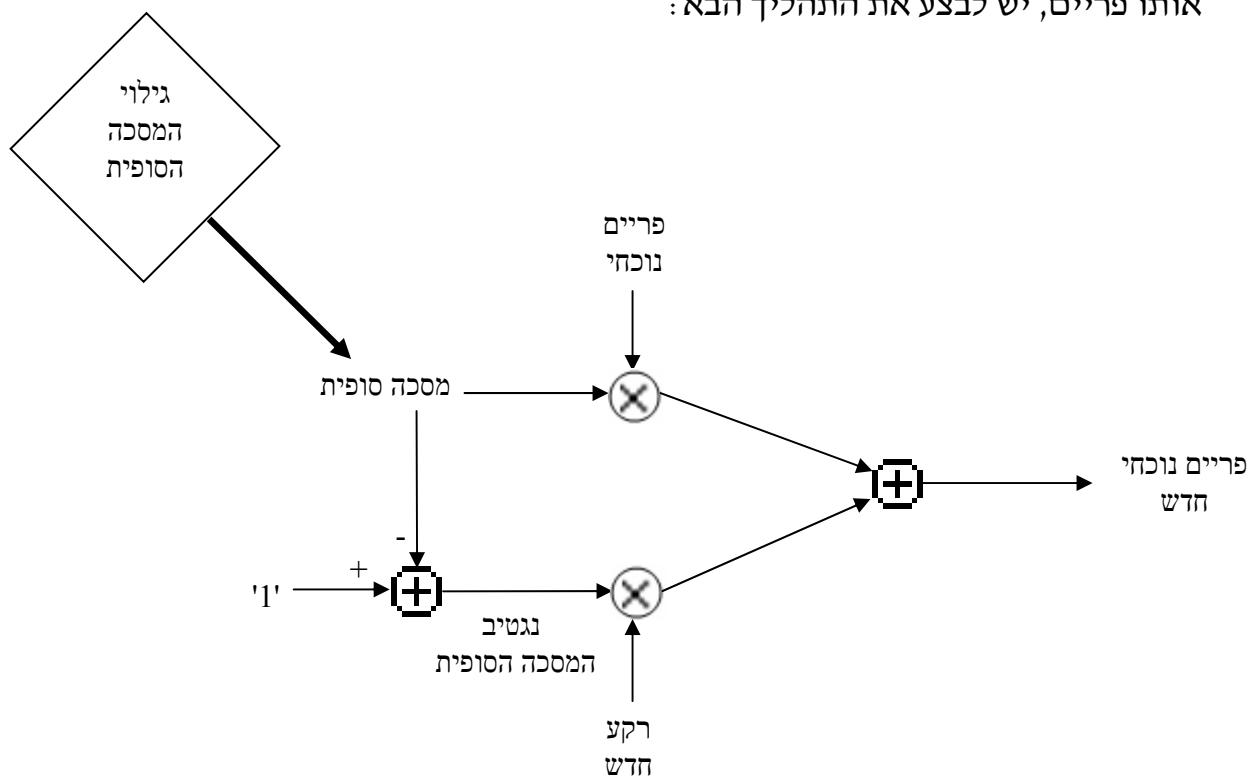


### הגישה שנבחרה בפרויקט

מהתוצאות שהתקבלו ניתן להסיק כי כדאי יהיה לשלב את הרקע הלומד עם תוספת של פעולות להשלמת מידע חסר ולהסרת מידע כוזב כפי שהסביר והציג בפרק עיבוד התמונה.

### תרשים בלוקים

כפי שהציג בפרק המבוא, לצורך החלפת הרקע עבור פריים מסוים, בהינתן המסכה של אותו פריים, יש לבצע את התהליך הבא :



איור 40 – סכימת בלוקים המתארת את קבלת התוצר בתוצר בפריים מסוים בהינתן המסכה

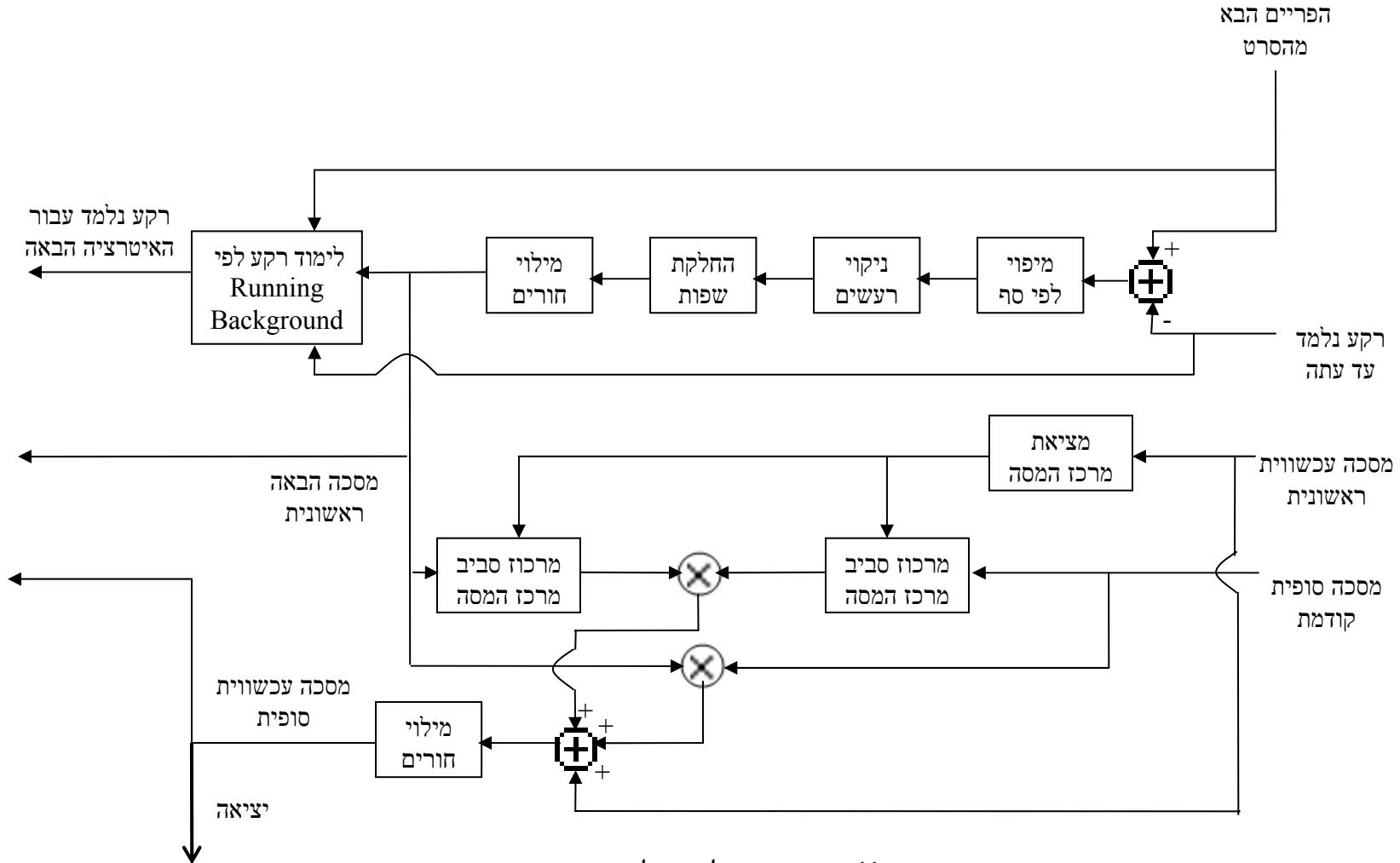
בלוק גילוי המסכה הסופית מומש בפרויקט זה באופן הבא :



# הטכניון - מכון טכנולוגי לישראל

## הפקולטה להנדסת חשמל

### המעבדה למחקר הראייה ומדעי התמונה



**איור 41 – מימוש בлок גילוי המסכה הפנימית**



#### תוצאות

האלגוריתם המבוסס על הגישה הניל הורץ לדוגמה על סרט בגודל 720x480 פיקסלים, ובאורך של כ-35 שניות בקצב דגימה של 30 פריימרים לשניה. התקבלו התוצאות הבאות:

תמונה הרקע אליה הוועבר האובייקט:



איור 42 – הרקע אליו הוועבר האובייקט

מספר פריימרים מהסרט שנוצר, כאשר משמאלי סרט הכניסה ומימין סרט היציאה -  
פלט המודל, מוצגים בעמודים הבאים באյור 43 :



הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה





הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת החשמל  
המעבדה לחקור הראייה ומדעי התמונה





הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה





הטכניון - מכון טכנולוגי לישראל  
הפקולטה להנדסת חשמל  
המעבדה לחקור הראייה ומדעי התמונה



ניתן לראות כי התוצאות משביעות רצון, ואכן האובייקט הועבר מהרקע המקורי שלו לרקע החדש בצורה משכנעת. חילוץ האובייקט לשם החלפת הרקע בוצע בצורה טובה הן כאשר האובייקט היה בתנועה והן כאשר הוא היה במנוחה.

מבחינת סיבוכיות, המודל הנ"ל מבצע בmäßig איטרציה כ- 5 שניות עבור פרימים שנלקחו במלואם (ללא דצימציה שנوعדה לפטור את בעיית ה – Interlace) כאשר גודלם, כפי שצוין, הוא 720x480 פיקסלים.



### הנחות יסוד

על מנת לאפשר את פועלות המודל יש להניח את הנחות הבאות:

- מעוניינים לחוץ את כל העצמים הדומיננטיים הנעים בסרט הווידיאו ללא אבחנה. לא ניתן למשל סרט בו מופיעים שני אנשים לחוץ רק אחד מהם, אלא אך ורק את שנייהם.
- הרקע האחורי רשאי להיות מורכב ככל שייהי
- הצלום נעשה על ידי מצלמה נייחת
- יחס אות לרעש גבוה. עבור רעש חזק מדי ביחס לתמונה עלולים להיות שיבושים רבים באיכות הסרט הסופי שנוצר כפלט המודל.
- קיימת תמונה רקע ראשונית שאוначלה ללא עצמים נעים, דהיינו את המודל יש להריץ מעט לפני כניסה האובייקטים אותם מעוניינים לחוץ בעtid.
- הרקע אשר אליו מעבירים את האובייקט חייב להיות לפחות בגודל מסגרתבודדת של סרט הכניסה.
- לא ניתן להניח מודל כללי של רעש



## סיכום ומסקנות

האלגוריתם הורץ על מספר סרטים שונים וניתן היה להבחן במאפיינים שונים. הסרט, שהוצגה עבورو החלפת הרקע, היה סרט בעל SNR (יחס אות לרעש) גובה ועבورو קיבלו ביצועים טובים. ניתן לראות כי חילוף הרקע התבצע בהצלחה ובאיכות טובה.

כאשר נלקחו סרטים בעלי SNR נמוך איכות המסקה נפגמה. אמנם ניתן היה להתגבר ביכולת על הרעשים באמצעות סינון פשוט כפי שהוצעו בפרק עיצוב התמונה, אולם ביחד עם הרעש מסוון לרוב גם חלק מהאובייקט וכותזאה מכך נוצרת מסקה אסולה מעט שאינה יפה לעין המתבונן. בנוסף, הסרטים מסויימים היו בעיות של השתקפות וצלליות דומיננטיות של האובייקט אשר מהו היה יותר קשה להפטר באמצעות הסימונו שהוצג מפהת גודלן.

כמו כן יש לציין כי הספרים למיפוי תמונה רמות אפור לתמונה בינהית נקבעו בצורה ידנית לכל סרט בנפרד כך שהתקבלו תוצאות טובות. לא היה סף גלובלי אחד שעבד טוב לכל הסרטים.

לסיכומו של עניין, המודל שמומש עובד באיכות טובה עבור סרטים ללא רעש גובה, ובאיכות פחותה טובה עבור סרטים בעלי רעשים גבוהים וואו השתקפות וצלליות דומיננטיות. בנוסף היה צורך לקבוע ידנית את הסף בכל סרט בנפרד.



### הצעות שיפור ומחקר עתידי

כפי שצוין בפרק הסיכום והמסקנות, בתנאים מסוימים נפגמת איקות התוצר הסופי. להלן כמה הצעות שעשוויות לפתור את הבעיה ולשפר את איקות התוצאה:

#### סינון רעים מקדים –

כפי שצוין איקות התוצר הסופי נפגמת כאשר סרט הכניסה, שעליו מבצעים את העיבוד, רועש מדי. בפרויקט זה התבצע סינון רעים כחלק מהעיבוד לאחר שהתקבלו פרימרים מהסרט המקורי, אולם לא התבצע ניקוי רעים על תמונה הסרט לפי העיבוד – וזאת מפני שלא ניתן היה להניח מודל רעש מסוים. סביר להניח שסינון רעים כללי שאינו מניח הנחות סטטיסטיות על הרעש יגרום לכך שהעיבוד יהיה מוצלח יותר על סרטים רועשים.

#### הורדת צלליות והשתקפות –

צלליות והשתקפות דומיננטיות בסרטים מפריון לפעלת האלגוריתם בפרויקט. עיבוד מקדים שיוריד את הצלליות והשתקפות יכול לעזור לקבל תוצאות טובות יותר.

#### מציאת ספים אוטומטית –

הספרים על פיהם מופיעו תמונות האפור לתמונות ביןaries נקבעו באופן ידני לכל סרט וסרט. מציאת אוטומטית של הספרים אלה תשפר את כלליות הקוד ואולי אף את איקות התוצאה.



## רשימת מקורות

- [1] Stauffer C, Grimson W. E. L. “*Adaptive background mixture models for real-time tracking*”. in *Proceedings. 1999 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No PR00149)*. IEEE Comput. Soc. Part Vol. 2, 1999.
- [2] KaewTraKulPong, P., Bowden, R.. “*An improved adaptive background mixture model for realtime tracking with shadow detection*”. In: *Proc. 2nd European Workshop on Advanced Video Based Surveillance Systems*, 2001
- [3] Eric N. Mortensen and William A. Barrett, “*Intelligent Scissors for Image Composition*”, *Computer Graphics Proceedings, Annual Conference Series, 1995*, pp. 191-198.
- [4] Barjatya A., “*Block matching algorithms for motion estimation*”; DIP 6620 Spring 2004 Final Project Parper.
- [5] P. L. Rosin, “*Thresholding for change detection*”, *Computer Vision and Image Understanding*, vol. 86, pp. 79–95, 2002.
- [6] Tian, Y.L.[Ying-Li], Hampapur, A.[Arun], “*Robust Salient Motion Detection with Complex Background for Real-Time Video Surveillance*”, *Motion05(II: 30-35)*.
- [7] V.Kolmogorov, A.Criminisi, A.Blake, G.Cross, C.Rother, *Microsoft Research Ltd*, “*Bi-Layer Segmentation of Binocular Stereo Video*”, *CVPR05(II: 407-414)*.
- [8] ר. מair, *הרצאות במערכות לומדות*, סמסטר אביב תשס"ז, פרק 4 : הערכת פילוגי הסתריות, הפקולטה להנדסת חשמל, טכניון, 2007,
- [9] מ.פורת, “*קורס במערכות דאייה ושמייה*”, ערך גיל בן עמי, טכניון.



[10] מ.אלעד, "עיבוד וניתוח תמונות", מהדורה שנייה, אוקטובר 2000, הטכניון.

[11] א.סוצקובר, נ.חייט, י.אברגיל, "דוחיתת תמונות ואוטומטיות ויזואו, תדריך ניסוי", המעבדה לעיבוד אוטומטיות ותמונה, הפקולטה להנדסת חשמל, טכניון, מרץ 2006.

[12] S.Cohen, K.Tsimbelman, "*Intelligent Scissor for edge detection*", Instructor: Idan Shatz *Vision & Image Since Lab, Technion, 2003*

[13] איור 8 נלקח מ -

<http://www.robots.ox.ac.uk/~parg/projects/ica/riz/Thesis/Figs/var/MoG.jpeg>

[14] איור 9 נלקח מ -

<http://www.owlnet.rice.edu/~elec301/Projects02/artSpy/patmac/RGB.jpg>

[15] איור 10 נלקח מ -

<http://www.mathworks.com/access/helpdesk/help/toolbox/images/hsvcone.gif>

[16] איור 11 נלקח מ -

[http://en.wikipedia.org/wiki/CIE\\_1931\\_color\\_space](http://en.wikipedia.org/wiki/CIE_1931_color_space)

[17] איור 42 נלקח מ -

[http://www.tertullian.org/sabratha\\_temple\\_of\\_isis.jpg](http://www.tertullian.org/sabratha_temple_of_isis.jpg)