

## עיבוד וניתוח וידאו – תרגיל 2

### חלק 1

#### שאלה 1

משמעות האילוץ של בהירות קבועה הינו:

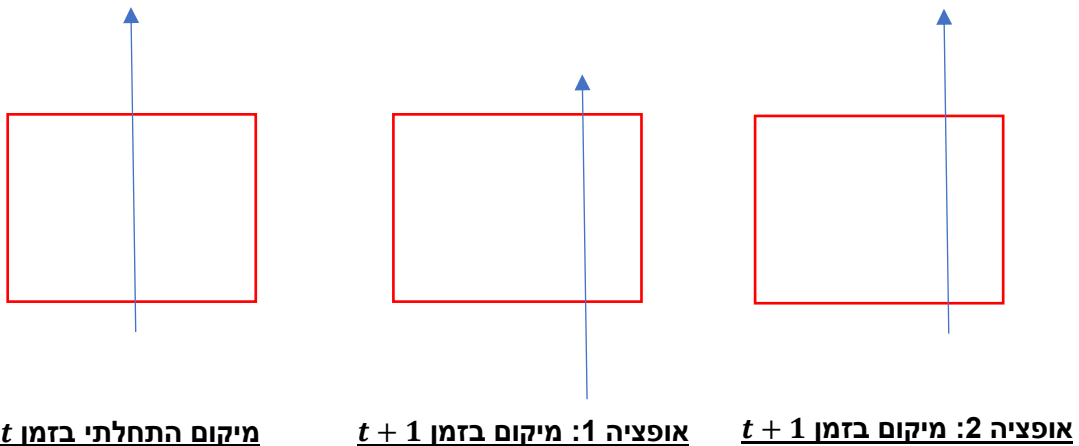
$$I_1(x, y) = I_2(x + u, y + v)$$

כלומר כל פיקסל שזז ממקום מסוים למקום אחר בתמונה ישמור על רמת בהירות זהה. הנחה זאת מקלה מאוד לפתרונות בעיית optic flow מכיוון שבצורה זאת ניתן לבנות מודל מתמטי המבוסס על קירוב טיילור מסדר ראשון, המניחה שעבור פיקסלים בסביבה מסוימת וקטור תזוזה זהה, ולכן פתרון המשוואה הופך משוואה אחת עבור כל פיקסל למערכת משוואות עבור סביבת פיקסלים.

בעולם האמיתי ההנחה כי הבהירות לא משתנה היא לא נכונה, עצמים רפלקטיביים הנעים מאזור מוצל לאזור מואר יכולים לשנות את רמת הבהירות באופן קיצוני, כלומר יתכנו מצבים בהם ההנחות הבסיסיות של המודל יהיו לא נכונות.

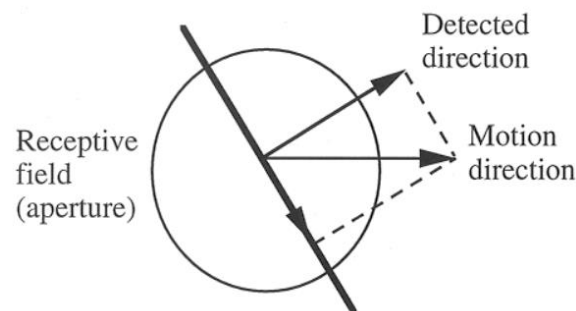
#### שאלה 2

בעיית האפרטורה אומרת כי יתכנו מצבים בהם לא נדע את הכיוון האמיתי בו עצם בתמונה זז בגלל שהחלון דרכו אנו מביטים הוא סופי, וייתכן והאובייקט לא יכנס בצורה מלאה לחלון זה. נראה דוגמא:



ניתן לראות כי בגלל גודל האפרטורה הסופי שלנו, יכולים להיות מצבים בהם לא נדע את התנועה המדויקת אלא נדע רק את כיוון התנועה הניצבת לגרדיאנט בתמונה.

דוגמא נוספת:



פיתרון אחד שניתן לעשות על מנת למזער את הבעיה הוא הגדלת החלון ומעבר למודל 2D אפיני במקום טרנסלציה, כלומר הגדלת כמות המשתנים.

### שאלה 3

בבעיית ה-optic flow קיימת משוואה אחת ושני נעלמים.

$$0 = I_t + I_x u + I_y v$$

כאשר  $I_t = I_2 - I_1$ . והנעלמים שלנו הם  $u, v$ .

לכן ב-Lucas-Kanade פתחו סביב כל פיקסל חלון, למשל בגודל  $5 \times 5$  והניחו כי בסביבה הקטנה הזו הפיקסלים ינועו ביחד לאותו כיוון. לכן מייצרים עוד ועוד משוואות, אשר כל אחת ניתנת לייצוג כישר במישור  $u, v$  כאשר הפתרון הוא בנקודת החיתוך של הקווים (25 משוואות עבור 2 נעלמים).

### שאלה 4

ההנחה הזאת אינה בהכרח נכונה עובר שפות של אובייקטים, מכיוון שהתזוזה במקרים כאלה של פיקסלים שכנים השייכים לאובייקטים שונים עשויים לנוע לכיוונים שונים.

### שאלה 5

האלגוריתם שאנו מציעים יעשה שימוש ב-depth map. בגלל העובדה שאנו פותרים בעיה המכילה 2 נעלמים באמצעות 25 משוואות נציע את הפתרון הבא:

נפתח חלון  $5 \times 5$  ונשלח אותו ליחידה המחשבת את ה-depth map, הפלט של היחידה תהיה מטריצה בגודל  $5 \times 5$  המשייכת כל פיקסל לאובייקט באמצעות מתן ערך integer שונה לפיקסלים השייכים לאובייקט שונה (אובייקט 1 יקבל את הערך 1, אובייקט 2 את הערך 2...). מכיוון שהחלון של  $5 \times 5$  יחסית קטן, לא נצפה להופעה של מספר אובייקטים רב (בוודאות חד ספרתי, ובביטחון רב ברוב המקרים עד גג 3-4 אובייקטים)

עבור כל אובייקט נמצא את שני הנעלמים  $u_i, v_i$  היחודיים לו באמצעות סט המשוואות הנובע מפיתוח טיילור שביצענו, וזאת רק עם המידע השייך לפיקסלים המשייכים לאובייקט זה. כלומר כעת לא יהיו לנו 25 משוואות עבור 2 נעלמים, אלא יותר נעלמים עם פחות משוואות (בגלל ריבוי האובייקטים).

למשל עבור depth map שיגילה 2 אובייקטים, אובייקט 1 בוגדל 14 פיקסלים, ואובייקט 2 עם 11 פיקסלים נקבל:

עבור אובייקט 1 – 14 משוואות עם 2 נעלמים  $u_1, v_1$ .

עבור אובייקט 2 – 11 משוואות עם 2 נעלמים  $u_2, v_2$ .

נציין בזאת כי ככל שיהיו יותר אובייקטים שיזהו על ידי ה-depth map כך הדיוק של פתרון המשוואות יהיה פחות טוב. בנוסף קיימת מגבלה של 12 אובייקטים עם 12 נעלמים עבור חלון של  $5 \times 5$ , בפועל מקרה זה יניב תוצאה מאוד לא מדויקת, וכמו כן ניתן להניח כי הוא לא ישים ברוב המקרים.