

Image Processing - Exercise 1

Shachar Cohen



מבוא

המטרה העיקרית של תרגיל זה היא לזהות מעברים בין סצנות בקטעי וידאו נתונים, על ידי ניתוח ההיסטוגרמות של הפריימים בסרטון. כפי שלמדנו בשיעור, היסטוגרמת צבעים היא ייצוג של התפלגות הצבעים בתמונה. היסטוגרמת הצבעים מייצגת את מספר הפיקסלים הצבועים בכל צבע. לפיכך, היינו מצפים שההיסטוגרמות של תמונות דומות יהיו דומות גם כן. בפרט נצפה שפריימים עוקבים בסרטון בהם אין מעבר סצנות, יהיו בעלי היסטוגרמות דומות. לפיכך נצטרך להגדיר מטריקה כלשהי להשוואת ההיסטוגרמות בין כל שני פריימים עוקבים (בהמשך הדו"ח).

ההבדל העיקרי בין הסרטונים בקטגוריה הראשונה לבין הסרטונים בקטגוריה השנייה הוא באופי ההיסטוגרמות שלהם. בקטגוריה הראשונה ניתן לראות כי ההיסטוגרמות של פריימים עוקבים בתוך אותה סצנה הם בעלי התפלגות צבעים דומה (כלומר היסטוגרמה לא סוכמת דומה). לכן עבור קטגוריה הזו נוכל להשתמש במרחק בין שתי ההיסטוגרמות של פריימים עוקבים כדי להעריך האם היה מעבר סצנה בין הפריימים האלו. בקטגוריה השנייה ישנם פריימים עוקבים באותה סצנה בעלי עיוות קל בintensity של הצבעים בתמונה, כך שהמרחק בין ההיסטוגרמות של שני הפריימים גדל באופן משמעותי. במקרה הזה, מדידת המרחקים בין ההיסטוגרמות הסוכמות תסייע לזהות מעברי סצנות, משום שההיסטוגרמות סוכמות פחות רגישות לשינויים קלים בצבעי התמונה.

אלגוריתם

בשל מאפייני התרגיל, ההנחה המוקדמת שלי בשני האלגוריתמים היא שקיים מעבר סצנה (יחיד) בסרטון הנתון. תחת ההנחה הזו, ציפיתי שהמרחק בין שני פריימים בהם יש מעבר סצנה יהיה מקסימלי תחת מטריקה מסוימת. כלומר, בהינתן $d: F \times F \rightarrow \mathbb{R}$ פונקציה למדידת מרחקים בין שני פריימים, f_1, \dots, f_n הפריימים הנתונים בסרטון (לפי הסדר), נצפה שמעבר הסצנה בסרטון יהיה:

$$(i, i + 1) = \operatorname{argmax}_{(j, j + 1), j \in [n-1]} \{d(j, j + 1)\}$$

המטריקה שהגדרתי למדידת המרחק בקטגוריה הראשונה היא המרחק האוקלידי (נורמה 2) בין שתי ההיסטוגרמות הרגילות של הפריימים. המטריקה שהגדרתי למדידת המרחק בקטגוריה השנייה היא המרחק האוקלידי בין שתי ההיסטוגרמות הסוכמות של שני הפריימים. המטריקות האלו הן פשוטות יחסית וסביר להניח שבמקרים מסובכים יותר (שינוי קיצוני בתאורה באמצע הפריים, אפקטים כלשהם וכד') הן לא יעבדו, אבל הן טובות מספיק לתרגיל הזה.

שלבי האלגוריתם:

- א. טעינת הפריימים בסרטון ב-grayscale.
- ב. נאתחל מרחק מקסימלי ל-0, ערך החזרה ל-null, והיסטוגרמה קודמת ל-null.
- ג. איטרציה על כל i מ-1 עד n , כאשר n הוא מספר הפריימים בסרטון. בכל איטרציה:
 - 1) נחשב את ההיסטוגרמה (הסוכמת/ הרגילה) של הפריים ה- i .
 - 2) נחשב את המרחק האוקלידי בין ההיסטוגרמה של הפריים ה- i להיסטוגרמה של הפריים ה- $i-1$.
 - 3) נבצע השמה של ההיסטוגרמה הנוכחית לתוך ההיסטוגרמה הקודמת.
 - 4) אם המרחק הנוכחי גדול מהמרחק המקסימלי, נשמור את האינדקס של הפריים הנוכחי והפריים הקודם בתור ערך ההחזרה.
- ד. לאחר סיום הריצה, נחזיר את ערך ההחזרה ששמרנו.

זמן ריצה: נסמן w, h מימדי התמונה בכל פריים ו- n מספר הפריימים בסרטון. טעינת הסרטון והמרתו לשחור לבן לוקחת $O(h \cdot w \cdot n)$, וכיוון שבכל איטרציה אנחנו מחשבים היסטוגרמה פעם אחת ב- $O(w \cdot h)$ ואז משווים בין שתי ההיסטוגרמות, סיבוכיות ההשוואה תהיה $O(n \cdot w \cdot h)$ כלומר סה"כ $O(n \cdot w \cdot h)$.

פרטי מימוש

שימוש בספרייה: השתמשתי בספרייה mediapy איתה עבדנו בCoding bootcamp בשל הפשטות והנוחות שלה, וכמובן בספרייה numpy לצורך חישובים ועבודה עם הדאטא.

יעילות בזכרון: כדי לחסוך בזכרון השתמשתי במבנה הנתונים VideoReader של mediapy, שבעצם משמש כאיטרטור ומאפשר לקרוא פריים אחד בכל פעם. טיפוס הנתונים הזה מקבל את סוג סקלת הצבעים שבה אנחנו

רוצים להשתמש ומעביר את הפריים בכל פעם לסקלת הצבעים הזו (במקרי שלנו – גווני אפור). כדי להשוות בין ההיסטוגרמה של הפריים הנוכחי לפריים הקודם, נמנעתי מלשמור את כל הפריים ושמרתי רק את ההיסטוגרמה שכבר חישבנו. בנוסף, כדי להימנע מכפילות קוד השתמשתי באותה פונקציה עבור שני סוגי הסרטונים והתאמתי את סוג ההיסטוגרמה לסוג הסרטון. יש כאן איזון עדין בין יעילות בזכרון לבין יעילות בזמן ריצה – כביכול יכולתי לחשב את המרחקים בין כל שני פריימים במקביל ב-numpy ולא ע"י איטרטור, אבל תחת ההנחה שאורך הסרטון יכול להיות אפילו כמה דקות (אלפי/מאות אלפי פריימים), היעילות בזכרון חשובה ויכולה לקצר בפועל גם את זמן הריצה.

אתגרים: שיווי ההיסטוגרמות: כשהתחלתי לעבוד על התרגיל חשבתי להשתמש באלגוריתם לשיווי ההיסטוגרמות על כל פריים, לפני ההשוואה בין הפריימים. ציפיתי ששיווי ההיסטוגרמות יעזור לאלגוריתם לזהות פריימים עוקבים באותה סצנה שיש בהם שינוי בהירות. בפועל שיווי ההיסטוגרמות הגדיל את הרעש בתמונות מסוימות (למדנו על כך בכיתה) והקשה (!) על האלגוריתם לזהות תמונות באותה סצנה.



דוגמה לניסוי שעשיתי בשיווי ההיסטוגרמות

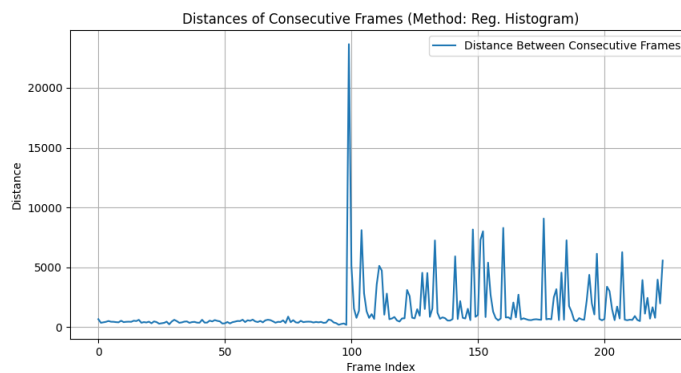
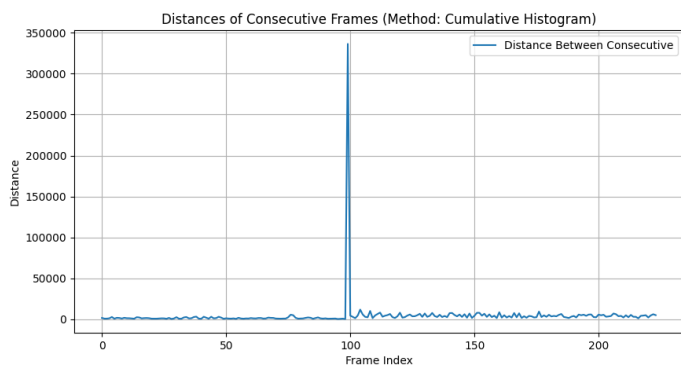
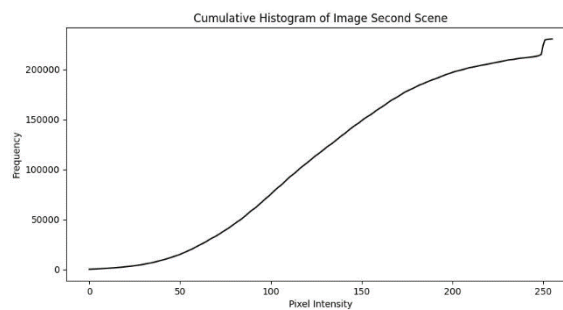
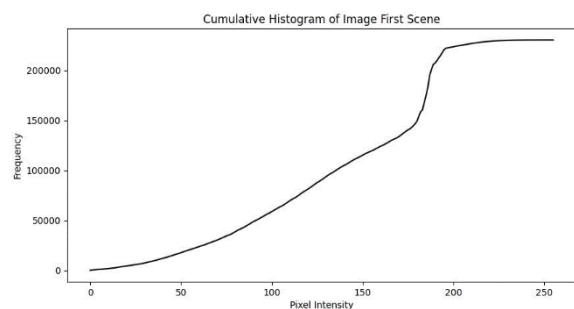
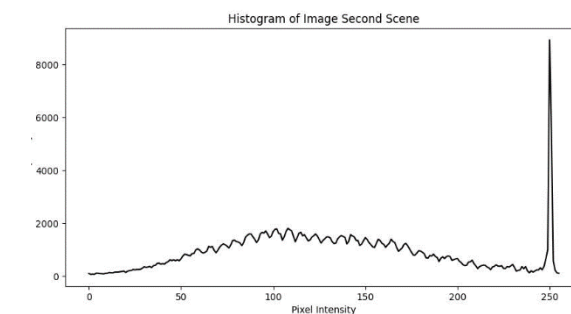
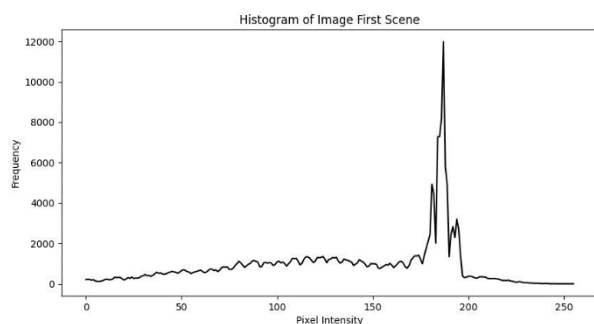
זיהוי אובייקטים בתמונה: אחרי הכשלון המוחץ של שיווי ההיסטוגרמות חיפשתי ספריות שמזהות עצמים בתמונה, כדי שבכל זאת אוכל לזהות את אותו עצם בין שני פריימים בסצנה מסוימת גם אם הבהירות משתנה באמצע הסצנה. גיליתי שבספרייה OpenCV יש ממשק שמאפשר לזהות עצמים בתמונה, אבל בסוף החלטתי שזו קצת רמאות ולא השתמשתי בממשק הזה.

תוצאות קטגוריה 1

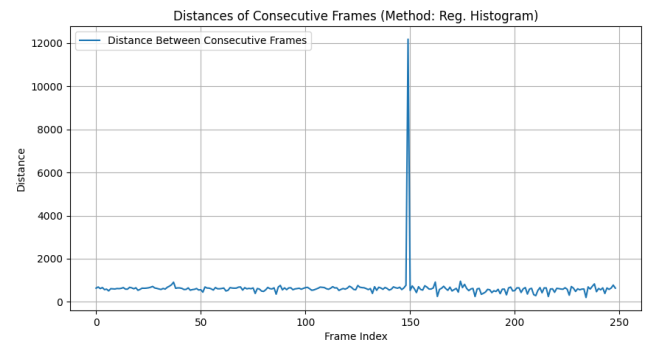
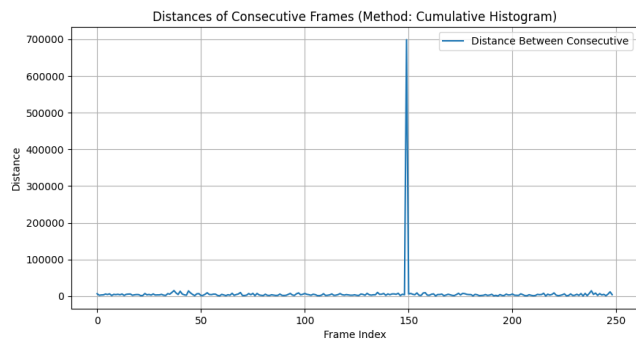
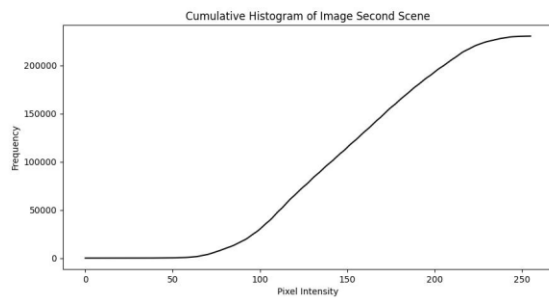
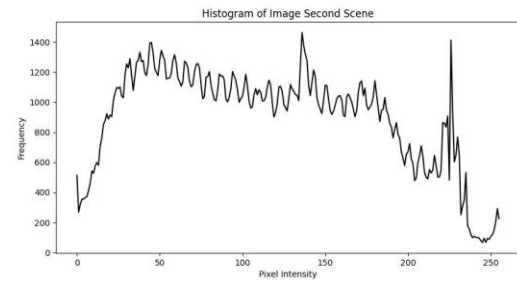
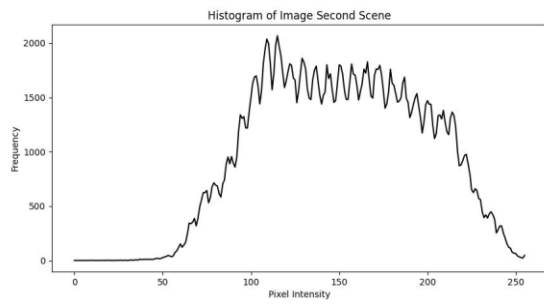
בשרטוטים מטה ניתן לראות את שתי התמונות שזוהו בתור מעבר פריים ע"י האלגוריתם בכל אחד מהסרטונים – האלגוריתם באמת זיהה את מעבר הסצנה (פריימים 99 ו-100, פריימים 149 ו-150). מתחת לתמונות ניתן לראות את ההיסטוגרמה הרגילה ואת ההיסטוגרמה הסוכמת של כל אחת מהתמונות, וכן את המרחקים בין פריימים עוקבים בשתי השיטות לאורך הסרטון. במקרה הזה, כשהפעלנו את שני האלגוריתמים (השוואה של ההיסטוגרמה סוכמת והשוואה של ההיסטוגרמה רגילה) קיבלנו את אותה התוצאה, אך ניתן לראות כי בהיסטוגרמה הסוכמת התוצאה מובהקת יותר. ניתן לראות כי בשתי ההיסטוגרמות הפיקסלים "מתקבצים" סביב ערך גוון אפור מסוים ששונה זה מזה, כלומר פיזור הערכים בתמונה נמוך יחסית ומתרכז בנקודה שונה. התמונות בפריימים השונים באותה סצנה הם פעלי התפלגות צבעים דומה, כלומר ההתקבצות סביב צבע מסוים היא זו שעזרה להבחין

במעבר הפריימים. זה מסביר מדוע אלגוריתם שיווי היסטוגרמות לא יסייע לזיהוי מעבר פריימים, אלא רק יבלבל את האלגוריתם.

סרטון ראשון



סרטון שני

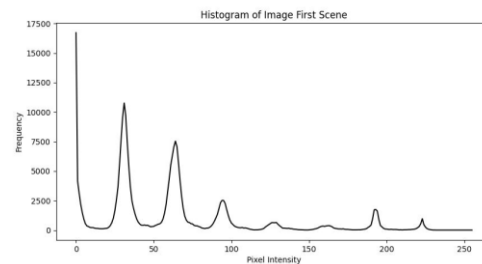
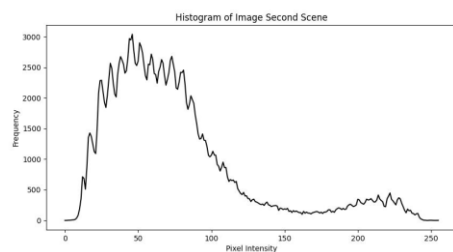


תוצאות קטגוריה 2

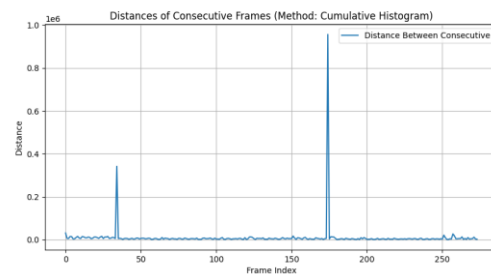
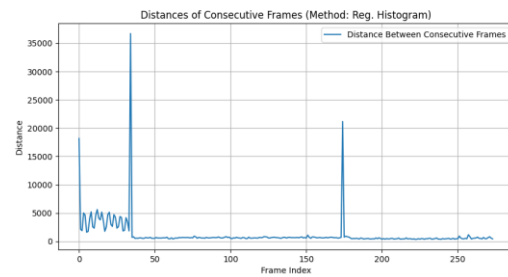
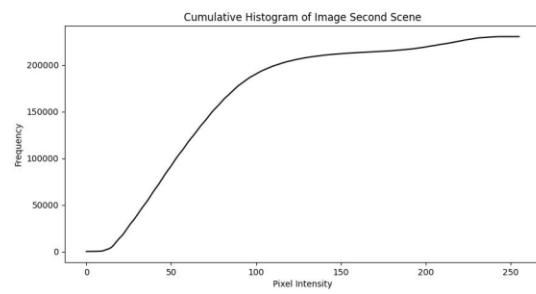
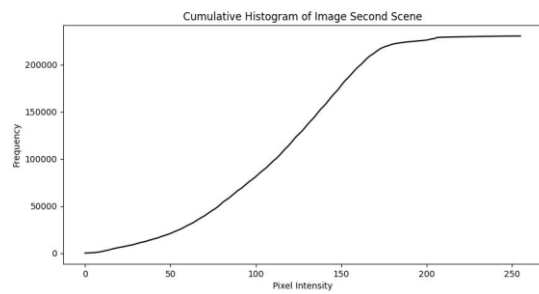
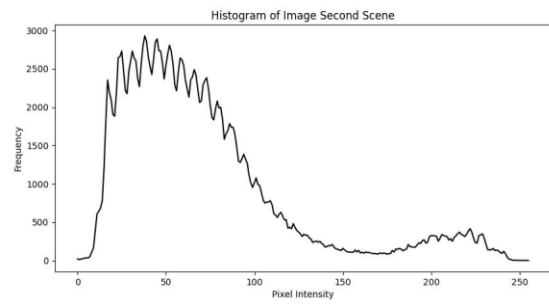
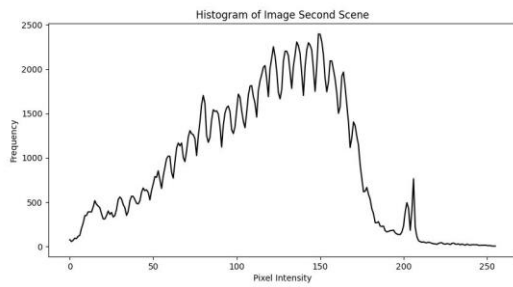
מצורפות התוצאות של שני הסרטונים. ניתן לראות כי האלגוריתם הסוכם אכן זיהה את מעבר הפריים נכון. בניגוד למקרה הראשון, בו ההיסטוגרמה הסוכמת סייעה להפוך את התוצאה למובהקת יותר אך לא שינתה את ההחלטה, ניתן לראות כי בקטגוריה הזו התקבלו שתי תוצאות שונות עבור ההיסטוגרמה הרגילה וההיסטוגרמה הסוכמת. כפי שהסברתי במבוא, ההיסטוגרמה הרגילה מושפעת יותר מעיוותים קטנים בתמונה, וניתן לראות כי גם בפועל ההיסטוגרמה הרגילה זיהתה שינוי משמעותי בין שני פריימים באותה סצנה עם שינוי קל בצבעים ולכן סיווגה אותם כמעבר סצנה. למשל בסרטון הראשון, האלגוריתם של קטגוריה 1 סיווג את המעבר בין הפריימים הנ"ל כמעבר סצנה.



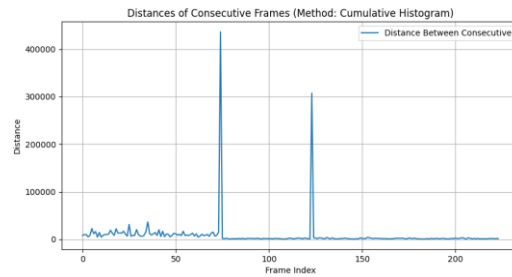
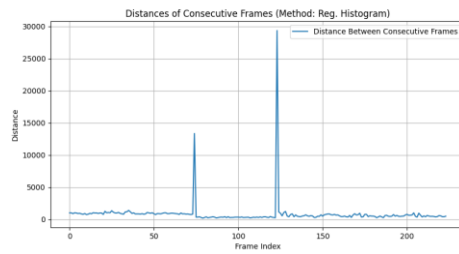
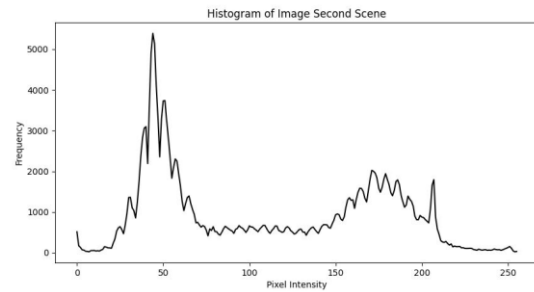
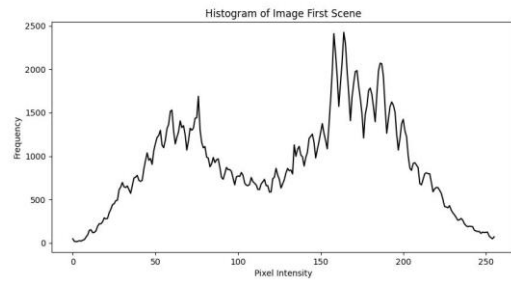
ניתן לראות כי ההיסטוגרמה של התמונה הימנית היא יותר "דיסקרטית", כלומר יש פחות פיזור של צבעים, לעומת התמונה השמאלית, שהיא יותר "חלקה".



סרטון ראשון



סרטון שני



מסקנות

לסיכום, בתרגיל זה ראינו כי על אף שקיימים היום אלגוריתמים מתקדמים לזיהוי עצמים, ניתן לזהות שינויים משמעותיים בתמונה גם בשיטות מסורתיות יותר כמו השוואת היסטוגרמות. ראינו כי היסטוגרמה שאינה סוכמת רגישה לשינויים קטנים בתמונה, כמו דיסקרטיזציה של ערכי הצבע בתמונה, ולכן אינה יעילה במיוחד למדידת הדמיון בין תמונות. לעומת זאת, היסטוגרמה סוכמת פחות רגישה לשינויים כאלו ומתאימה יותר למדידת הדמיון בין תמונות ולהבנת התפלגות הצבעים בתוך התמונה. בנוסף, למדתי לעבוד עם ספריות בסיסיות בעיבוד תמונה ועיבוד וידאו ולכתוב תוכנית בסיסית ל-histogram equalization, שזה מגניב ממש :)