

Image Processing - Exercise 1

Shir Hana Stern, shirste, 325945236

מבוא

מטרת התרגיל היא לזהות את המעבר בין שתי סצנות שונות בקטע וידאו, באמצעות ניתוח שינויים בתוכן הוויזואלי בין פריימים עוקבים. המעבר בין הסצנות מתרחש כאשר יש שינוי חד במידע הוויזואלי מפריים אחד למשנהו, המעיד על מעבר לסצנה חדשה.

הטכניקה המרכזית שבה השתמשתי כדי לפתור את התרגיל כללה השוואה בין היסטוגרמות של פריימים עוקבים. ההיסטוגרמות מייצגות את התפלגות הצבעים בפריים, והשוואת היסטוגרמות של שני פריימים מגלה עד כמה הם דומים או שונים זה מזה.

קטגוריה 1 מתאפיינת בכך שכל אחת משתי הסצנות בסרטון כמעט במצב סטטי, ואז יש שינוי בולט מאוד בין ההיסטוגרמות של שתי הסצנות.

קטגוריה 2, לעומת זאת, מכילה שתי סצנות שלפחות אחת מהן משנה באמצע את רמת הטשטוש של תמונת הסצנה, מה שהופך את זיהוי המעבר בין הסצנות לקשה יותר. שינוי החדות בין פריימים של אותה סצנה עשוי להיות הבדל משמעותי בין היסטוגרמות הפריימים ולכן יכול להקשות על מציאת המעבר הנכון בין שתי הסצנות השונות בקטע הוידאו.

אלגוריתם

1) קריאת פריימים מהסרטון: השלב הראשון הוא קריאת קובץ הוידאו והפקת הפריימים שלו. כל פריים נשמר כמערך דו-ממדי של ערכי פיקסלים.

2) לולאה על פני כל הפריימים:

- עבור כל פריים בסרטון, אם הוא מקטגוריה 1 מתבצעת המרה לגווני אפור ואם הוא מקטגוריה 2 מתבצעת המרה לשחור-לבן.
- חישוב של היסטוגרמה כך שנקבל היסטוגרמה אחת לכל פריים.
- השוואת פריימים עוקבים: עבור כל פריים (למעט הראשון), מתבצע חישוב של ההבדל בין היסטוגרמת הפריים הנוכחי להיסטוגרמת הפריים הקודם. ההבדל מחושב על ידי חיבור ההבדלים המוחלטים בין הערכים של הפיקסלים (מדד פשוט לשינוי בין פריימים).
- מעקב אחרי ההבדל המקסימלי: אם ההבדל של היסטוגרמות בין הפריים הנוכחי לפריים הקודם גדול מההבדל המקסימלי שנמצא עד כה, מתבצע עדכון של ההבדל המקסימלי ומספרי הפריימים נשמרים כנקודות חיתוך סצנה פוטנציאליות (הפריים לפני השינוי והפריים הנוכחי).
- החזרת התוצאה: הפונקציה מחזירה זוג מספרים, שהם מספרי הפריימים המייצגים את נקודות חיתוך הסצנה, כלומר את הפריימים שבהם היה שינוי בסצנה.

השינוי היחיד שעשיתי באלגוריתם בין הקטגוריה הראשונה לשנייה, היה שעבור קטגוריה 2 ביצעתי המרה של כל פריים לבינארי. המרת כל פריים לתמונה בינארית של שחור ולבן עזר בסינון של השפעת

השינויים הקטנים והלא משמעותיים (כמו טשטוש) ופישוט ההיסטוגרמות על ידי שמירה רק על אינפורמציה חשובה. בשיטה הזו, השפעת שינויי ההיסטוגרמות בין פריימים המשנים לפתע את החדות אך השייכים לאותה סצנה קטנה.

פרטי מימוש

(1) קריאת נתוני הווידאו – בשלב הראשון, קובץ הווידאו נקרא ומחולק למערך פריימים של `numpy` בעזרת הפונקציה `read_video()` השייכת לספרייה `mediapy`. המערך מכיל את כל פריימי הווידאו, מה שמאפשר גישה לכל פריים בצורה של מערך דו-ממדי עבור כל ערך צבע (RGB).
(2) לולאה על פני כל הפריימים:

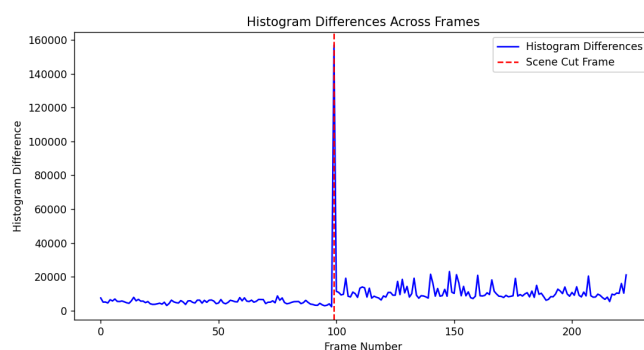
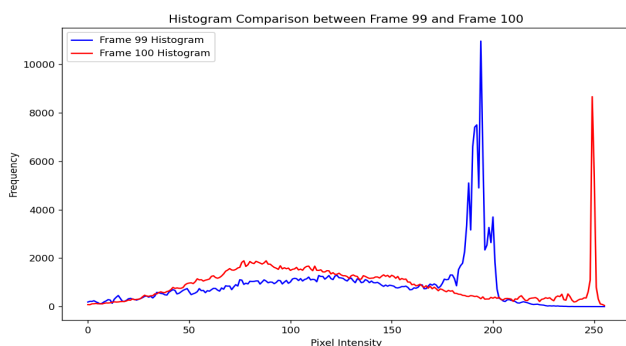
- המרת הפריימים לשחור לבן/ גווני אפור – כל פריים מומר לתמונה בגווני אפור על ידי שימוש בספריית `numpy` שמבצעת חישוב ממוצע הערכים לאורך ציר הצבעים (RGB) עבור כל פיקסל, ואם הווידאו הוא מקטגוריה 2 יוצרים תמונה בינארית של שחור ולבן.
- חישוב ההיסטוגרמה של כל פריים – לכל פריים מחושבת ההיסטוגרמה (באמצעות פונקציה של ספריית `numpy`). ההיסטוגרמה מאפשרת להשוות את הפריימים ביניהם במונחים של השינויים בבהירות ובהתפלגות הצבעים המרכיבים כל פריים.
- השוואת פריימים עוקבים – הפריימים העוקבים נבדקים על ידי חישוב ההבדל בין ההיסטוגרמות שלהם. ההבדל מחושב על ידי שימוש בפונקציות של ספריית `numpy` שמבצעות חיבור של ההבדלים המוחלטים בין כל ערך בהיסטוגרמות, מה שמעיד על השינוי בתמונה בין שני פריימים עוקבים.
- מעקב אחרי ההבדל המקסימלי (יישום שלי) – כל פעם שמחושב הבדל חדש בין שני פריימים, הוא מושווה להבדל הקודם שנמצא. אם ההבדל הנוכחי גדול יותר מההבדל המקסימלי הקודם, נשמרים הפריימים הנוכחיים. בסופו של תהליך, נשמרים זוג הפריימים עם ההבדל המקסימלי.
- החזרת זוג הפריימים עם ההבדל הגדול ביותר (יישום שלי) – בסיום האלגוריתם, מוחזר זוג הפריימים שבו נמצא ההבדל הגדול ביותר בהיסטוגרמות שלהם. זוג זה מצביע על חיתוך סצנה שבו התרחש השינוי המובהק ביותר מבחינת הבהירות.

האלגוריתם שלי עבור שתי הקטגוריות לא מחזיר את זוג הפריימים שבו ההבדל בין ההיסטוגרמות עובר סף מסוים, אלא מחזיר את זוג הסצנות שבהן ההבדל בין ההיסטוגרמות הוא **המקסימלי**. נשים לב שעבור קטגוריה 2 האלגוריתם מבצע המרה לשחור-לבן (בינארי). המרה זו מפחיתה את הרגישות לשינויים בבהירות או בחדות בתוך אותה סצנה, כיוון שהיא שומרת רק על מידע ניגודיות בסיסי. לעומת זאת, בגווני אפור, שינויים קטנים בבהירות או בחדות יכולים לגרום להבדלים משמעותיים יותר בהיסטוגרמה, גם אם הסצנה לא השתנתה. לכן, המעבר לשחור-לבן מתאים במיוחד לזיהוי מעבר בין סצנות, תוך צמצום "רעש" הנובע משינויים מינוריים בתוך אותה סצנה, ולכן הוא היה נדרש עבור סרטונים מקטגוריה 2.

השתמשתי ב-Mediapy לקריאת קובץ הווידאו והמרתו לפריימים וב-Numpy לחישוב ההיסטוגרמות וביצוע חישובים מתמטיים יעילים.

תוצאות קטגוריה 1

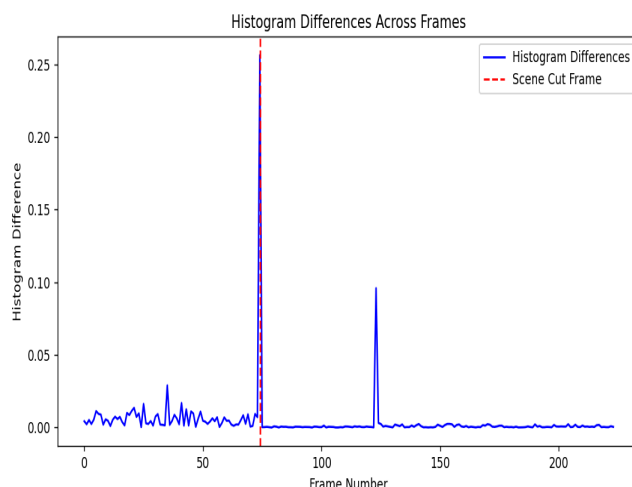
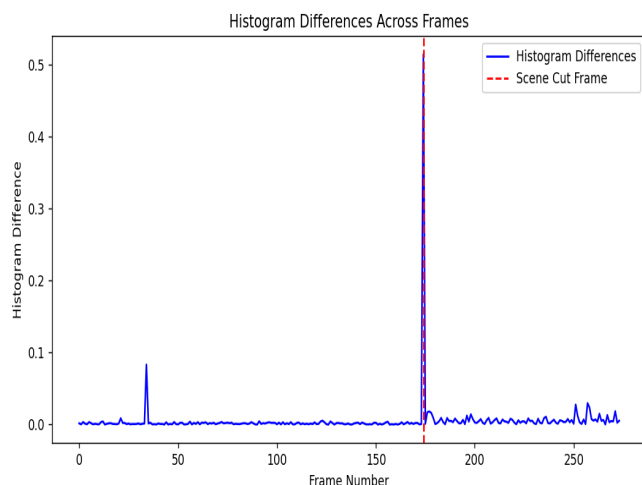
בסרטונים מקטגוריה 1, זוהו חיתוכי סצנות במדויק בהתבסס על ההבדל המקסימלי בין ההיסטוגרמות של פריימים עוקבים. בסרטון הראשון חיתוך הסצנה זוהה בין פריימים 99 ל-100, ובסרטון השני בין פריימים 149 ל-150. גרפים המציגים את ההבדלים בין ההיסטוגרמות לאורך הפריימים מראים קפיצה משמעותית בנקודות החיתוך, ותמונות מוקטנות של הפריימים לפני ואחרי החיתוך מדגישות את השינוי הוויזואלי בבירור.



החלטה על מיקום חיתוך הסצנה בוצעה על פי השינוי המשמעותי בהיסטוגרמות בין פריימים עוקבים. גרף ראשון מציג עבור הסרטון הראשון מקטגוריה 1 את **נקודת השיא של הבדל ההיסטוגרמות** שהתרחש בין פריימים 99 ו-100. גרף זה ממחיש את החיתוך בצורה ברורה, כיוון שבנקודת השיא יש שינוי מאוד בולט בין שני הפריימים. נתון זה מחזק את ההשערה שמדובר בחיתוך סצנה. בנוסף, גרף שני מציג את **השינוי בהיסטוגרמות** בין פריימים 99 ו-100, שבו ניתן לראות שינוי חד ודרמטי, דבר המעיד על חיתוך סצנה. גרף זה מראה באופן ברור את השינוי המשמעותי בין שני הפריימים, מה שמעיד על המעבר בין סצנה אחת לשנייה.

תוצאות קטגוריה 2

בסרטונים מקטגוריה 2 זוהו חיתוכי סצנות באופן מדויק, בהתבסס על ההבדל המקסימלי בין ההיסטוגרמות של פריימים עוקבים (לאחר המרה של כל פריים לשחור ולבן). בסרטון השלישי זוהה חיתוך סצנה בין פריימים 174 ל-175, ובסרטון השני בין פריימים 74 ל-75. הסרטונים בקטגוריה 2 שונים מאלו בקטגוריה 1 בכך שבקטגוריה 1 הסצנות לרוב סטטיות והמעברים בין הסצנות מלווים במעברים חדים, בעוד שבקטגוריה 2 יש סצנות דינמיות, עם תנועה ושינוי חדות בין פריימים השייכים לאותה הסצנה. בקטגוריה 1, ההבדלים בהיסטוגרמות בין פריימים עוקבים בולטים יותר, שכן יש הבדל משמעותי אחד בין הסצנות, מה שמקל על זיהוי חיתוך הסצנה. לעומת זאת, בקטגוריה 2 השינויים בהיסטוגרמות רבים יותר לאורך הסרטון, מה שמקשה על זיהוי החיתוך בצורה ברורה.



הגרפים המצורפים מציגים את נקודות השיא בהבדלי ההיסטוגרמות עבור הסרטונים בקטגוריה 2. בניגוד לגרפים שהתקבלו עבור הסרטונים בקטגוריה 1, ניתן להבחין שקיימות שתי נקודות שיא בהבדל ההיסטוגרמות. פתרון המתבסס על סף (threshold) בודד אינו מתאים במקרה זה, שכן הוא מזהה רק נקודת שיא מקומית שמספיק שעברה איזשהו סף מסוים, מבלי להבחין בשינויים השונים והמשמעותיים המתרחשים במהלך הסרטון. לכן, בחרתי להשתמש באלגוריתם המחפש את המקסימום הגלובלי של ההבדלים בהיסטוגרמות, מה שמאפשר זיהוי מדויק יותר של חיתוך הסצנות, גם כאשר יש יותר מנקודת שיא אחת.

בנוסף, עבור סרטונים מקטגוריה 2 בחרתי לבצע המרה של כל פריים לשחור ולבן בינארי על מנת להפחית את הבדל ההיסטוגרמות של פריימים מאותה סצנה כאשר מתבצע שינוי בחדות/איכות הפריים תוך כדי הסצנה. המרה זו עזרה לי להפחית "רעשים" שהיו עלולים להפריע לאלגוריתם שלי למצוא את חיתוך הסצנות המתאים, שכן ללא המרה זו, האלגוריתם היה מזהה את שני הפריימים העוקבים שבהם חדות התמונה משתנה כנקודת החיתוך בין שתי הסצנות אף על פי שהן שייכות לאותה הסצנה.

מסקנות

* **חיתוכי סצנות בקטגוריה 1:** בסרטונים מסוג 1, עם סצנות סטטיות ומעברים חדים, ההבדלים המשמעותיים בהיסטוגרמות בין פריימים עוקבים אפשרו זיהוי מדויק של החיתוכים.

* **אתגרים בקטגוריה 2:** בסרטונים מסוג 2, עם שינויים בחדות באותה סצנה, ההבדלים המשמעותיים בין פריימים עוקבים יצרו מספר היסטוגרמות מקסימליות לוקאליות, מה שהקשה על זיהוי חיתוכים.

* **יתרון של המרה לבינארי:** בקטגוריה 2, המרת כל פריים לתמונה בינארית של שחור ולבן עזר בסינון של השפעת השינויים הקטנים והלא משמעותיים (כמו טשטוש) ופישוט ההיסטוגרמות על ידי שמירה רק על אינפורמציה חשובה. בשיטה הזו, השפעת השינויים בפריים חד ומטושטש של אותה הסצנה קטן.

* **ביצועי האלגוריתם:** הגישה של זיהוי ההבדל המקסימלי הגלובלי בהיסטוגרמות הוכיחה את עצמה כיעילה עבור שתי הקטגוריות.

לסיכום, המסקנה העיקרית היא שזיהוי חיתוכי סצנות בסרטונים סטטיים, באמצעות השוואת היסטוגרמות, הוא תהליך פשוט יחסית. לעומת זאת, בסרטונים דינמיים, בהם יש שינויים תכופים ועדינים בהיסטוגרמות בין פריימים, יש צורך בשיטות מתקדמות יותר על מנת לזהות חיתוכים בצורה מדויקת.