

ראייה ממוחשבת – תשפ"ב

תרגיל בית 2

9/5/2023

הנחיות כלליות:

- ההגשה היא בזוגות או ביחידים בלבד.
- בתרגיל זה יש לממש כל שאלה בשפת מטלב או פייתון, כאשר ניתן להשתמש בספריות קיימות לפעולות בסיסיות כמו קריאה ושמירה של תמונות, ציור על גבי תמונה ופונקציות יותר ספציפיות שמותר להשתמש בהן לפי הפירוט בכל שאלה. ספריות מומלצות במטלב הן: Image-Processing and Computer Vision Toolboxes כמו גם ספריית vl_feat. עבור פייתון אפשר לציין את Pillow, OpenCV, scikit-image.
- יש להגיש את כל הקוד שאתם כתבתם (לא כולל קוד מספריות אחרות שנעזרתם בהן) בקובץ יחיד בשם code.zip.
- יש להגיש את כל התוצאות שתתבקשו להציג כאשר יש תיקייה נפרדת לכל בעיית סטריאו, כשהכל ארוז בקובץ יחיד בשם results.zip.
- יש להגיש מסמך אחד בפורמט pdf, שבו יש לציין שם ותעודת זהות. הסברי הבנוס צריכים להיות שם (אם עונים על הסעיף הזה).
- את התרגיל יש להגיש דרך אתר המודל. אפשר להגיש עדכונים של ההגשה (כל עוד זה לפני מועד ההגשה) כמו גם ערעורים לבדיקה במקרה הצורך במייל חוזר באותו שרשור.
- התרגיל להגשה עד ליום רביעי ה-31 במאי בשעה 23:59. כל יום איחור יגרור הורדה של 4 נק', עד לתאריך אחרון של יום שישי ה-2 ביוני. לא תינתן דחייה כלשהי (פרט לבקשות מוצדקות שיגיעו עד ה-27 במאי).

שאלות:

1. (100 נק') Photometric Stereo based View Synthesis

המטרה היא ליצור סדרת זוויות צפיה חדשות של סצנה, המקשרת בין זוג תמונות במצב rectified. ניתן לעשות זאת ע"י שינוי ה-pose של המצלמה וסינתוז של התמונה שהייתה נרכשת מזווית הצילום החדשה. בכל דוגמא יש זוג תמונות במצב rectified שצולמו על ידי זוג מצלמות זהות (עם אותם intrinsics) כאשר המרחק בין המצלמות (baseline) הוא כ-10 ס"מ (0.1 מטר).

שלבים:

א. חשבו את מפות ה-disparity (משני הכיוונים) על ידי אלגוריתם הסטריאו הלוקאלי הבא. יש לממש את שיטת winner-takes-all אחרי aggregation על cost-volume שמחושב על בסיס ה-census-transform, עם סיון consistency:

- חשבו census-transform על כל אחת מהתמונות
- חשבו את ה cost-volume (שימו לב שנתון ה-disparity המקסימלי בקובץ txt).
- בצעו אגרגציה לוקאלית
- מצאו את המינימום לכל מיקום
- סננו התאמות על סמך consistency-test

ב. חשבו מתוכן את שתי מפות העומקים.

ג. עבור התמונה השמאלית: חשבו reprojection של כל קואורדינטות התמונה (פיקסלים) אל המרחב התלת-ממדי. ניתן לעשות זאת בשימוש במטריצה K של ה intrinsics (ההופכית) ובמפת העומקים D. הנקודות בתלת-ממד יהיו במערכת הקואורדינטות של המצלמה, כלומר שאנחנו משתמשים במטריצת המצלמה $P = K[R|T]$ כאשר $R = I$, $T = 0$.
ד. הטילו בחזרה את הנקודות התלת-מימדיות אל מישור המצלמה השמאלית, במקומה המקורי. כעת, סנתזו את התמונה המתקבלת ע"י העתקה של ערכי הפיקסלים ב RGB מהתמונה המקורית. התוצאה אמורה להיות זהה לתמונה המקורית (עד כדי חורים) - זוהי בסה"כ בדיקת נכונות של תהליך ה-2D-3D-2D.

ה. חזרו על הפעולה עבור סדרה של 11 מיקומי מצלמה, במרווחים אחידים של 1 ס"מ על ה-baseline שבין שתי התמונות. ניתן לעשות זאת ע"י מיפולציה של T במטריצת ה extrinsics.

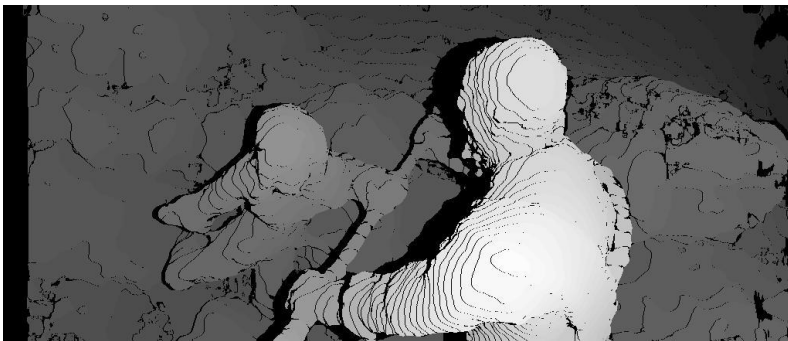
הערה: סיפקנו לכם דוגמא פתורה, שממנה נלקחו הויזואליזציות שלמטה. בין היתר, יש שם קבצי txt שניתן לטעון של מטריצות ה-disparities וה-depth (בנוסף לויזואליזציות שלהן). ניתן להתחיל לפתור ישירות את החלק השני של התרגיל (שלבים ג והלאה) בהינתן מפות העומקים (לפחות עד שנלמד על חישובי סטריאו בשיעור הקרוב).

להגשה (עבור כל סצנה, בתיקייה נפרדת):

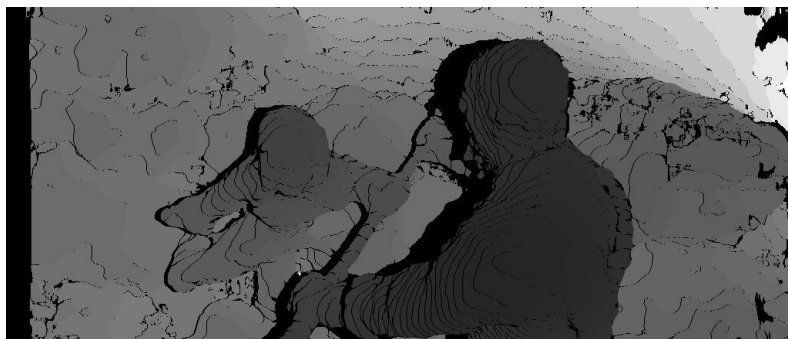
1. זוג התמונות המקוריות עם השמות `im_left.jpg`, `im_right.jpg`:



2. זוג תמונות ה-disparity עם השמות `disp_left.jpg`, `disp_right.jpg` (באיזורים ללא הדאטה צריך לשים אפס ובסוף לחלק את כל התמונה בערך המקסימלי כדי לקבל 'תמונה' ב- $[0,1]$):



3. זוג תמונות העומק עם השמות `depth_left.jpg`, `depth_right.jpg` (גם כאן, באיזורים ללא הדאטה צריך לשים אפס ובסוף לחלק את כל התמונה בערך המקסימלי כדי לקבל 'תמונה' ב- $[0,1]$):



4. סדרת התמונות המסונתזות עם השמות synth_i.jpg $i=1, \dots, 11$ כשאר i עולה משמאל לימין):



- ובנוסף** (10 נק'): שימו לב ל"חורים" (אזורים מושחרים) בתוצאות. ישנם שני סוגים של חורים: (I) קיום דקים; (II) חורים משמעותיים יותר (שנראים כמו צלליות).
- הסבירו את המקור של כל אחד משני סוגי החורים האלו
 - הציעו והסבירו פתרון שבו אין חורים בכלל. כדאי לזכור שניתן להשתמש בשתי תמונות המקור במקרה זה.
 - צרו תיקייה חדשה נוספת עם כל הפלטים הנ"ל, גם למקרה הזה.