

מבוא, הגדרות ומושגי יסוד

תמונה מנגנון שמעביר אותנו מהעולם התלת-מימדי לאיזושהי ישות דו-מימדית. יש את מישור התמונה, שם נמדדים ערכי התמונה, ועוצמת האור הנרשמת בנקודה (x, y) במישור התמונה נסמן ב- $I(x, y)$.

יש אובדן אינפורמציה בגלל שיצאנו מעולם תלת-מימדי ועכשיו האינפורמציה שאנחנו אמורים לעבד היא דו-מימדית.

ראייה אנושית	ראייה ממוחשבת
תמונה או וידאו	
חישה דרך העין	חישה דרך התקן
ניתוח באמצעות המוח	ניתוח באמצעות מערכת ממוחשבת
זיהוי דברים	

עיבוד תמונה זה בניית אינפורמציה ברמה יותר מתקדמת מהמידע הגולמי, כאשר המטרה הסופית היא ראייה ממוחשבת.

מבחינה אופטית

האור היוצא ממקור האור פוגע במשטחים וחוזר אל אמצעי החישה שמזהה את עצמת האור בכל נקודה.

מקורות לקבלת תמונה

- סריקה אופטית באמצעות מערך לינארי של קולטים הרגישים לאור.
- תמונות רנטגן (X-ray)
- עד היום רופאים ניתחו את זה, אבל מחשבים יכולים לסייע. למשל להצליב ממקורות שונים.
- תמונות טווח (*range images*) - בכל נקודה ונקודה מה שנקלט זה העצמה היחסית לאמצעי החישה.
- תמונת צבע - שלושה ערוצים שונים (אדום, ירוק, כחול - RGB). מדידת עצמת האור בכל צבע בנפרד.
- תמונות לוויין - בתחום מאוד גדול של אורכי גל. יכולה להכיל הרבה מאוד ערוצים ספקטרליים (לא רק RGB)

דגימה וכימוי (Sampling and Quantization)

דגימה אי אפשר לייצג כל נקודה ונקודה, כי יש אינסוף מידע (אפילו לא בר-מניה!) של נקודות. לכן מחליטים על חלוקה לתאים (סריג - grid) כאשר לכל תא (picture element - פיקסל) רושמים את עצמת האור. ככל שהתאים יותר קטנים הרזולוציה יותר גדולה. ככל שהתאים יותר גדולים ככה מאבדים יותר מידע.

$I(i, j)$ הערך של עצמת האור בכל תא במטריצה. זה ערך אפור - gray level value

כימוי הערך שנרשם יכול להיות מתחום מאוד גדול (אינסופי!) של ערכים - לכן מגבילים את עצמנו ועושים quantization - הקונבנציה המספקת היא להשתמש בבייט (byte) אחד - זה נותן 256 רמות שונות של אפור - מס עד 255. 00000000 זה שחור ו11111111 זה לבן.

- בתמונה בינארית יש לכל פיקסל ערך של 0 או 1.

יש לנו בעצם הפסד של מידע (*Degradation*) גם מבחינה מרחבית וגם מבחינת טווח ערכים.

עיבוד תמונה (Image Processing)

לאחר שיש לנו ייצוג דיגיטלי של התמונה, אפשר לעבד אותה לצרכים שונים כאשר המוטיבציה המרכזית היא לספק מידע לצרכים של ראייה ממוחשבת.

- **תצוגה (display)** - להציג את התמונה באופן שבני אדם יוכל להבין. תהליך הפוך לרכישת התמונה - מעביר אותה חזרה לעולם התלת מימדי.

- **שיפור איכות (enhancement)**

- **שחזור תמונה (restoration)** - לפעמים תמונה מתקלקלת בגלל תופעות פיזיקליות ואחרות, ויש צורך לשחזר את התמונה המקורית.

- למשל - הייתה בעיה בטלסקופ האבל, והיו צריכים לתקן את התמונות.

- **דחיסה (compression)** - צריך הרבה זיכרון כדי לאחסן תמונות, והרבה תעבורה כדי להעביר אותה. צריך להקטין את הגודל בלי לפגוע (יותר מדי) באיכות.

יש שתי פעולות מאוד חשובות לראייה ממוחשבת:

- **Edge Detection** - בשפה יש שינוי חד בצבע, שמלמד על שינוי פיסי באובייקט (מה שמבדיל אותו מאובייקטים שכנים)

- **Segmentation** - חלוקת התמונה לאזורים הומוגניים, במטרה לזהות את האובייקט מתוך אותם אזורים.

יש קשר בין הפעולות - edge הוא גבול של segment, אבל יש עוד דרכים להגיע לסגמנטים.

יישומים

- **זיהוי תווים - Optical Character Recognition (OCR)**. למשל מערכת חניונים שמתוך תמונה של לוחית רישוי מוצאת את מספר הרישוי. לוחיות רישוי זה יחסית קל, כי הכל מאוד סטנדרטי, אבל עדיין צריך לבצע זיהוי ספרות. אבל המטרה המהותית של OCR זה להפוך מסמכים סרוקים לקבצים שאפשר לבצע עליהם חיפוש ועריכה.

- **זיהוי פנים - Facial Image Processing**. תפס תאוצה אחרי 9/11, במטרה לזהות אנשים שלא אמורים להיות במקום ומהווים סיכון ביטחוני.

- זיהוי הבעות פנים - Facial Expression Recognition

- **מעקב אחרי אובייקטים - Object Tracking**. למשל בשביל לפתור מחלוקות באירועי ספורט. זיהוי המסלול הרציף של האובייקט מתוך תמונות בדידות.

- **ניתוח מידע רפואי - Medical Image Analysis**.

- **רכבים אוטונומיים - Autonomous Vehicles**.