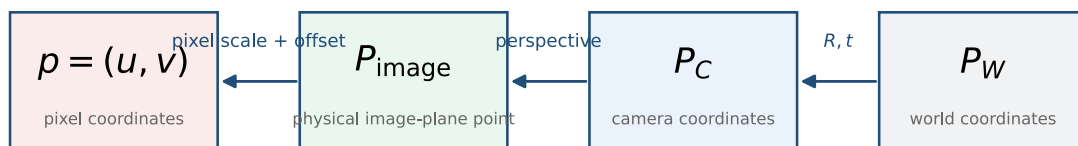
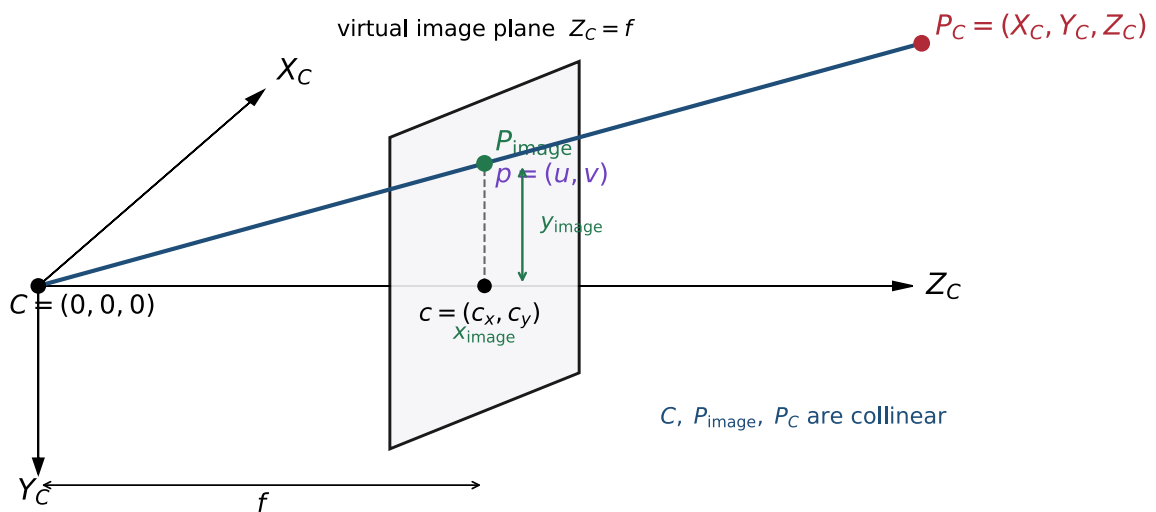


מן הנקודה במרחב אל הפיקסל בתמונה

מודל מצלמת נקב, דמיון משולשים, הקואורדינטות (u, v) והקשר לכיול ולמציאת תנוחת מצלמה

3D camera geometry



מטרת המסמך: להבין בהדרגה כיצד נקודה תלת-ממדית $P_C = (X_C, Y_C, Z_C)$ מוקרנת לנקודה דו-ממדית בתמונה $p = (u, v)$, ומדוע מופיעים בנוסחה אורך המוקד והעומק.

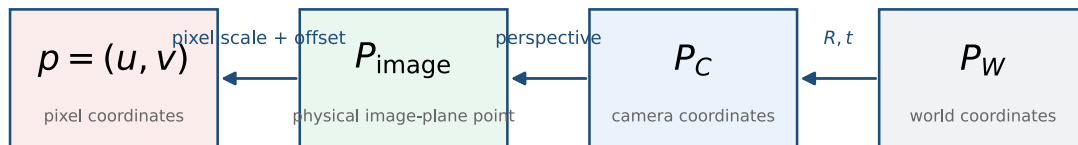
$$P_C = (X_C, Y_C, Z_C) \mapsto p = \pi(P_C) = (u, v)$$

מפת המסמך

1. מן הנקודה בעולם אל נקודת ההיטל והפיקסל
2. מערכת הצירים של המצלמה
3. קרן הראייה במודל מצלמת נקב
4. דמיון המשולשים ודוגמת עומק מפורטת
5. הנקודה הראשית (c_x, c_y)
6. מעבר ממילימטרים לפיקסלים
7. הניסוח המדויק $p = \pi(P_C)$
8. מטריצת הכיול K
9. דוגמה מספרית
10. ההיטל ההפוך וקרן במרחב
11. עיוות עדשה
12. כיול באמצעות Chessboard או ChArUco
13. מיקום וכיוון המצלמה בתלת־ממד
14. חיבור למסלול רובוט ולשדה ראייה

1. מן הנקודה בעולם אל נקודת ההיטל והפיקסל

לפני שמתחילים בחישוב, חשוב להפריד בין ארבעה תיאורים שונים. הם קשורים זה לזה, אך אינם אותה מערכת קואורדינטות.



- P_W - נקודה תלת ממדית במערכת הצירים של העולם או של לוח הכיול.
- $P_C = (X_C, Y_C, Z_C)$ - אותה נקודה לאחר העברה למערכת הצירים של המצלמה.
- $P_{image} = (x_{image}, y_{image}, f)$ - נקודת החיתוך הגאומטרית של קרן הראייה עם מישור התמונה הווירטואלי. הקואורדינטות x_{image}, y_{image} נמדדות ביחידות אורך, למשל מילימטרים.
- $p = (u, v)$ - הקואורדינטות בפיקסלים של אותה נקודת היטל בתמונה הדיגיטלית.

$$P_{image} = (x_{image}, y_{image}, f)$$

$$P_{image} \leftrightarrow p = (u, v)$$

הבחנה מרכזית:

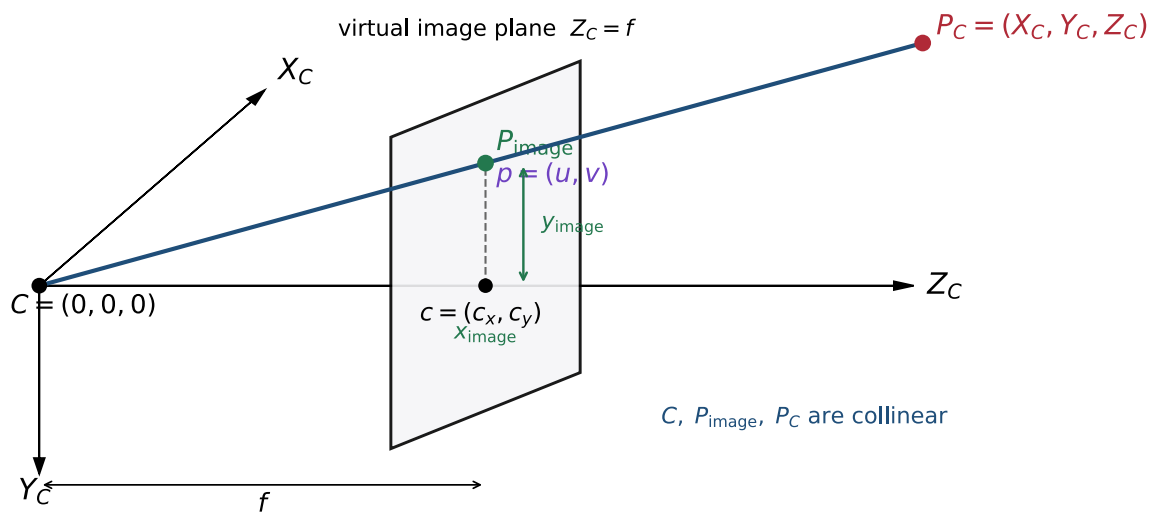
- P_{image} הוא המיקום הפיזיקלי של נקודת החיתוך על מישור התמונה.
 - $p = (u, v)$ הן קואורדינטות הפיקסלים של אותה נקודה בתמונה הדיגיטלית.
- הערת סימון:** אין פרמטר נוסף בשם m . נשתמש לאורך המסמך רק בסימונים המפורשים המופיעים בשתי הנוסחאות שלמעלה.

2. מערכת הצירים של המצלמה והציור התלת ממדי

במוסכמה הנפוצה של OpenCV:

- ציר X_C פונה ימינה ביחס למצלמה.
- ציר Y_C פונה למטה בתמונה.
- ציר Z_C הוא הציר האופטי ופונה קדימה מן המצלמה אל הסצנה.

3D camera geometry



- מרכז המצלמה הוא הראשית $C=(0,0,0)$.
- P_C היא הנקודה התלת ממדית במערכת המצלמה.
- הקרן הישרה ממרכז המצלמה אל הנקודה חותכת את מישור התמונה ב- P_{image} .
- הקואורדינטות הדיגיטליות של אותה נקודת חיתוך הן $p=(u,v)$.

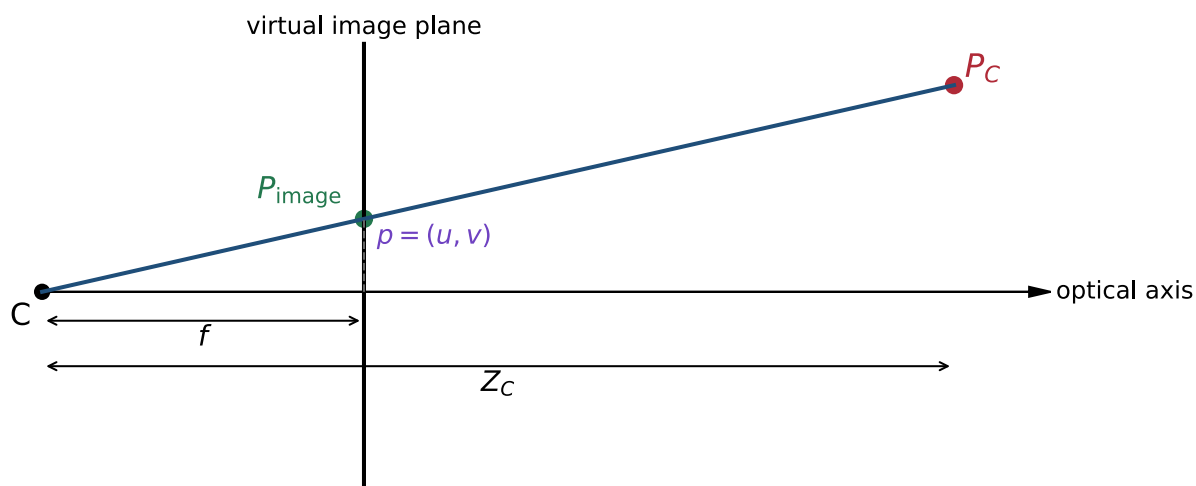
פירוש העומק: Z_C הוא הרכיב לאורך הציר האופטי. הוא אינו בהכרח המרחק האוקלידי המלא של הנקודה מן המצלמה.

3. מודל מצלמת הנקב: הקרן היא קו ישר

במודל מצלמת נקב מניחים שכל הקרניים עוברות דרך מרכז אופטי יחיד C . הקרן ממרכז המצלמה אל P_C היא:

$$\ell(\lambda) = C + \lambda P_C = \lambda(X_C, Y_C, Z_C), \quad \lambda \geq 0$$

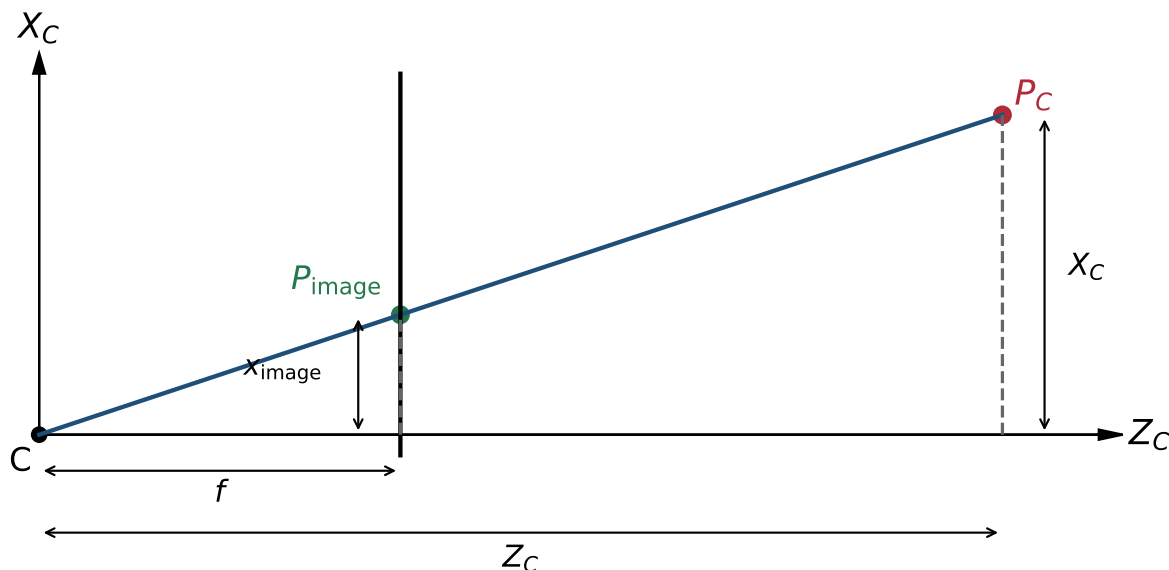
נשתמש במישור תמונה וירטואלי שנמצא לפני המצלמה במרחק מוקד f , כלומר במישור $Z=f$. הקרן חותכת את המישור בנקודה P_{image} .



שלוש הנקודות C , P_{image} ו- P_C נמצאות על אותו קו ישר. הנקודה P_{image} היא נקודת החיתוך של הקרן עם מישור התמונה.

4. דמיון המשולשים: מדוע מחלקים בעומק?

נתחיל בחתך דו ממדי במישור X_C-Z_C . שני המשולשים בצירוף הם ישרי זווית וחולקים את אותה זווית ליד מרכז המצלמה. לכן הם דומים.



במשולש הקטן הבסיס הוא f והגובה הוא x_{image} . במשולש הגדול הבסיס הוא Z_C והגובה הוא X_C . מדמיון המשולשים:

$$\frac{x_{\text{image}}}{f} = \frac{X_C}{Z_C} \implies x_{\text{image}} = f \frac{X_C}{Z_C}$$

באותו אופן, בחתך Y_C-Z_C :

$$y_{\text{image}} = f \frac{Y_C}{Z_C}$$

למה מחלקים בעומק? כאשר Z_C גדל, היחסים X_C/Z_C ו- Y_C/Z_C קטנים. לכן הנקודה הרחוקה יותר מוקרנת קרוב יותר לציר האופטי. זהו מקור הפרספקטיבה.

דוגמה מפורטת - ללא סימון מקוצר

נניח:

$$f=20 \text{ mm}, \quad X_C=0.10 \text{ m}=100 \text{ mm}$$

הנקודה נמצאת 10 סנטימטרים הצידה מן הציר האופטי. נשווה שני עומקים, ונשתמש בכל החישוב במילימטרים.

מקרה ב: עומק של שני מטרים

$$Z_C=2 \text{ m}=2000 \text{ mm}$$

$$x_{\text{image}}=20 \cdot (100/2000)=1 \text{ mm}$$

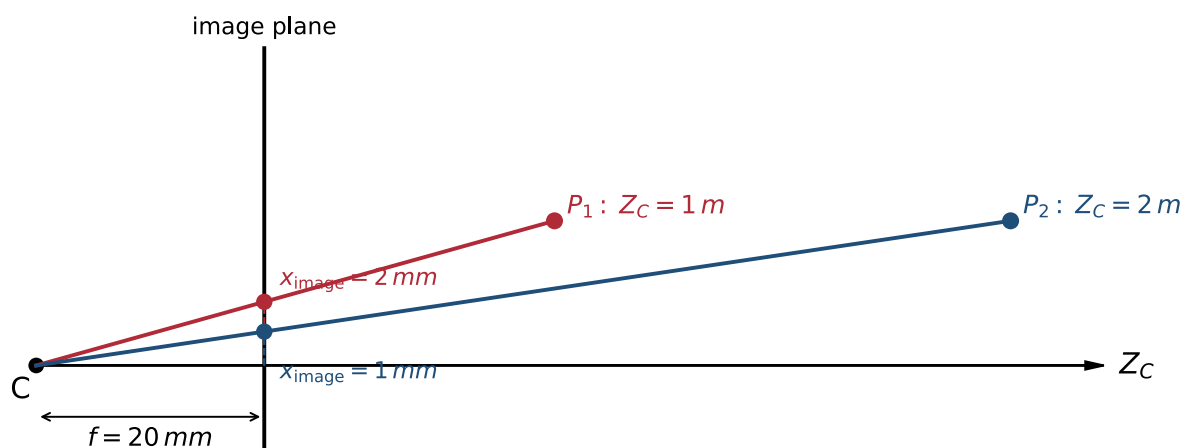
ההיטל נמצא רק מילימטר אחד מן הציר האופטי.

מקרה א: עומק של מטר אחד

$$Z_C=1 \text{ m}=1000 \text{ mm}$$

$$x_{\text{image}}=20 \cdot (100/1000)=2 \text{ mm}$$

ההיטל נמצא 2 מילימטרים מן הציר האופטי על מישור התמונה.



Same lateral offset $X_C = 0.10 \text{ m}$; doubling depth halves the image displacement.

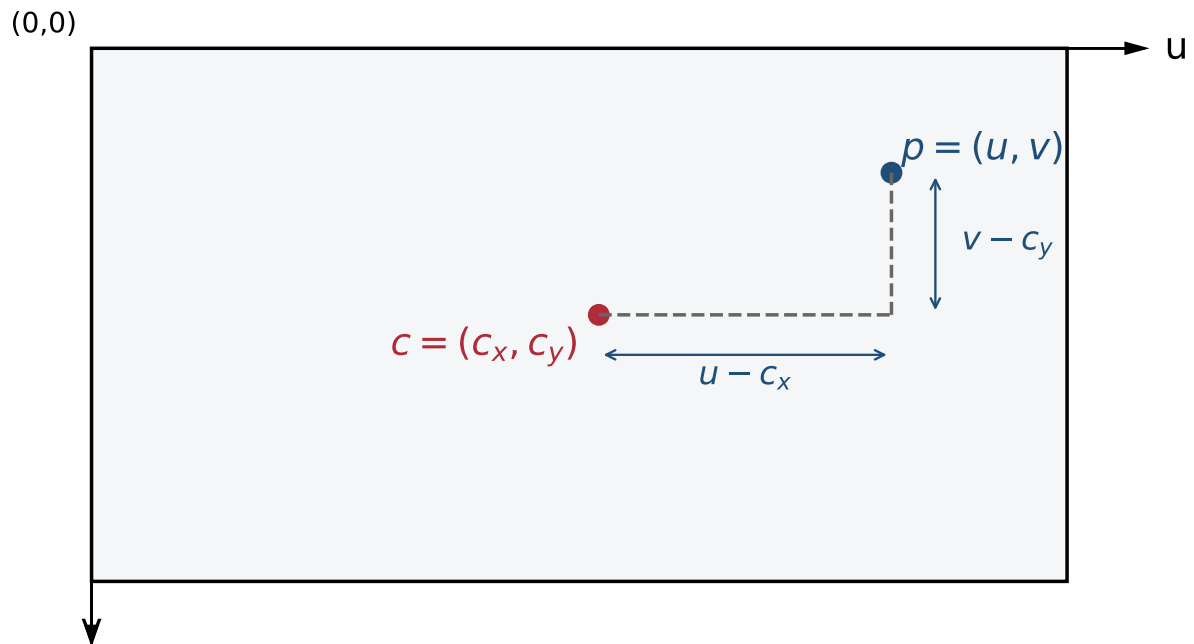
מסקנה: הכפלת העומק מ-1 מטר ל-2 מטר הקטינה את מרחק ההיטל מן הציר האופטי

מ-2 מילימטרים ל-1 מילימטר. לא נוסף שום פרמטר חדש; x_{image} הוא פשוט

הקואורדינטה הפיזיקלית של נקודת ההיטל על מישור התמונה.

5. הנקודה הראשית (c_x, c_y)

הציר האופטי חותך את מישור התמונה בנקודה הנקראת **הנקודה הראשית** - principal point. בקואורדינטות פיקסלים נסמן אותה $c=(c_x, c_y)$.



הנקודה $p=(u,v)$ היא נקודת ההיטל של P_C . המרחקים שלה מן הנקודה הראשית הם:

- $u - c_x$ בכיוון האופקי;
- $v - c_y$ בכיוון האנכי.

אם הנקודה במרחב נמצאת בדיוק על הציר האופטי, אז $X_C=Y_C=0$, ולכן נקודת ההיטל שלה היא בדיוק $(u,v)=(c_x, c_y)$.

6. ממידות פיזיקליות לפיקסלים

מן המשולשים הדומים קיבלנו את $x_{\text{image}}, y_{\text{image}}$ ביחידות אורך, למשל מילימטרים על החיישן. תמונה דיגיטלית נמדדת בפיקסלים.

נסמן ב- s_x וב- s_y את מספר הפיקסלים ליחידת אורך בשני הכיוונים. אז:

$$u - c_x = s_x x_{\text{image}}, \quad v - c_y = s_y y_{\text{image}}$$

נציב את תוצאות דמיון המשולשים ונגדיר אורכי מוקד ביחידות פיקסלים:

$$f_x = s_x f, \quad f_y = s_y f$$

ולכן מתקבלות נוסחאות ההיטל בפיקסלים:

$$u = f_x \frac{x_c}{z_c} + c_x, \quad v = f_y \frac{y_c}{z_c} + c_y$$

מדוע מופיע f בנוסחת u ? אורך המוקד קובע את קנה המידה של ההיטל על מישור התמונה. **מדוע מופיע z_c ?** העומק קובע כמה הפרספקטיבה מקטינה את הסטייה הצידית.

7. הניסוח המדויק של פונקציית ההיטל

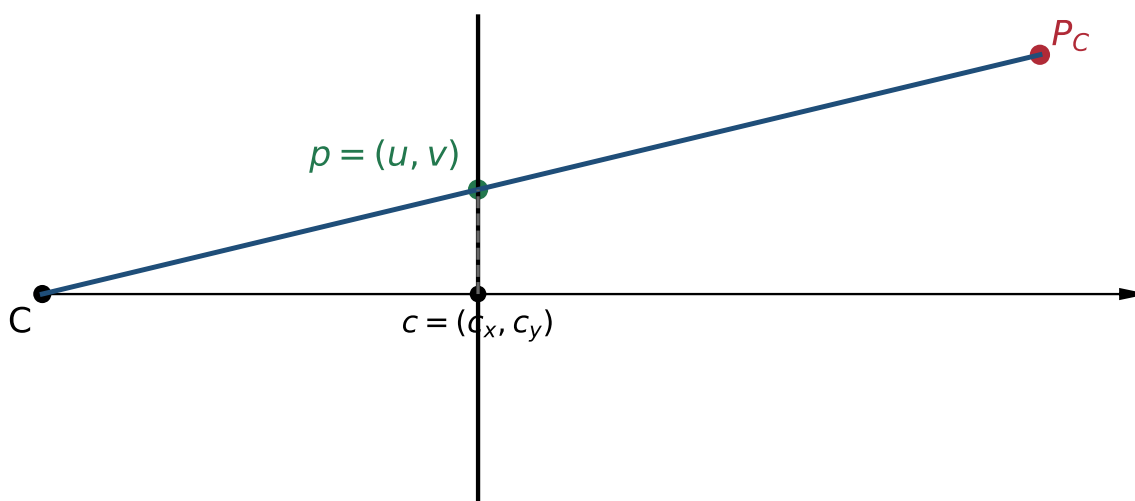
כדי להדגיש את התלות בנקודה התלת־ממדית, נגדיר פונקציית היטל π . הכתיבה המדויקת היא:

$$p = \pi(P_C) = \left(f_{x\overline{Z}_C} \frac{X_C}{Z_C} + c_x, f_{y\overline{Z}_C} \frac{Y_C}{Z_C} + c_y \right)$$

כלומר, שתי קואורדינטות הפיקסל תלויות בנקודה התלת־ממדית ובפרמטרי המצלמה. אפשר להדגיש זאת בכתיבה הפונקציונלית הבאה:

$$u(P_C) = f_{x\overline{Z}_C} \frac{X_C}{Z_C} + c_x, \quad v(P_C) = f_{y\overline{Z}_C} \frac{Y_C}{Z_C} + c_y$$

למה מופיעים f ו- Z_C בנוסחה של u ? משום ש- u אינו נתון שרירותי; הוא הקואורדינטה של ההיטל של P_C . אורך המוקד קובע את קנה המידה, והעומק קובע את ההקטנה הפרספקטיבית.



p היא הנקודה הגיאומטרית על מישור התמונה; (u, v) הן הקואורדינטות שלה. הכתיבה $p = (u, v)$ אינה מציבה נקודה חדשה מחוץ לקרן. C, p, P_C קולינאריות.

8. מטריצת הכיול K

את ארבעת הפרמטרים הפנימיים המרכזיים מרכזים במטריצה:

$$K = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix}$$

בקואורדינטות הומוגניות אפשר לכתוב:

$$Z_C [u, v, 1]^T = K [X_C, Y_C, Z_C]^T$$

לאחר הכפל מתקבלים $Z_C u = f_x X_C + c_x Z_C$ ו- $Z_C v = f_y Y_C + c_y Z_C$. חלוקה ב- Z_C מחזירה את שתי נוסחאות ההיטל.

מטריצת K מתארת את הפרמטרים **הפנימיים** של המצלמה: אורך מוקד בפיקסלים והנקודה הראשית. היא אינה מתארת את מיקום המצלמה בחדר.

9. דוגמה מספרית מלאה

נניח כיול שבו:

$$f_x=1000, \quad f_y=980, \quad c_x=640, \quad c_y=360$$

ונקודה במערכת המצלמה:

$$P_C = (0.20, 0.10, 2.00) \text{ m}$$

נחשב:

$$u = 1000 \cdot (0.20/2.00) + 640 = 740$$

$$v = 980 \cdot (0.10/2.00) + 360 = 409$$

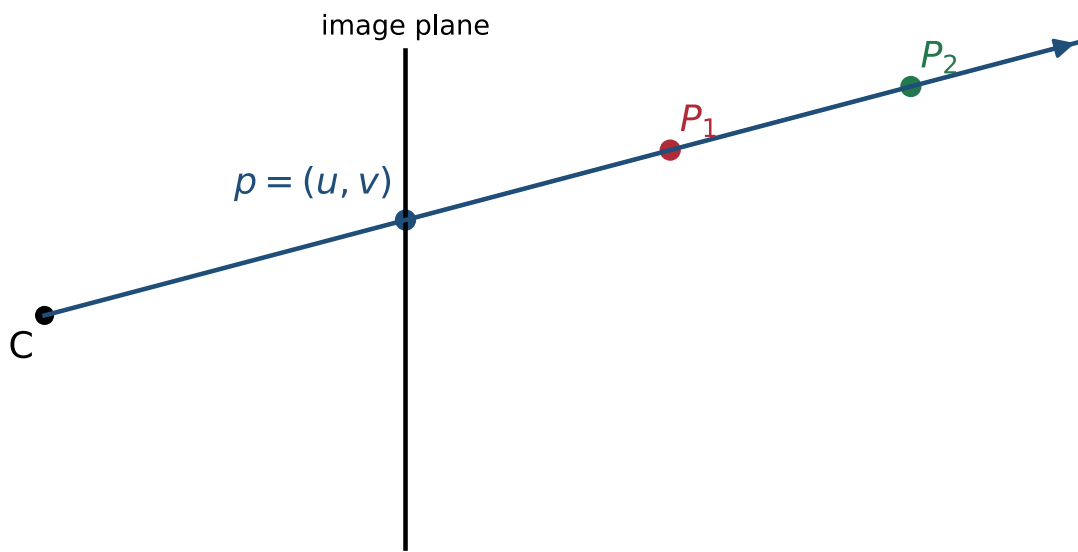
לכן $p=(740,409)$. ביחס לנקודה הראשית $(640,360)$, הנקודה נמצאת 100 פיקסלים ימינה ו-49 פיקסלים למטה.

אם נכפיל את העומק ונשמור את X_C, Y_C ללא שינוי, נקבל נקודה קרובה יותר למרכז התמונה. זו השפעת הפרספקטיבה.

10. הכיוון ההפוך: פיקסל אחד מגדיר קרן, לא נקודה יחידה

הנוסחה הקדמית לוקחת נקודה תלת-ממדית ומחזירה פיקסל. בכיוון ההפוך, פיקסל יחיד אינו קובע עומק. הוא קובע כיוון של קרן.

$$P_C(\lambda) = \lambda K^{-1}[u, v, 1]^T, \quad \lambda > 0$$



הנקודות P_1 ו- P_2 מוקרנות לאותו פיקסל משום שהן על אותה קרן. כדי לקבוע עומק צריך מידע נוסף: מצלמה שנייה, חיישן עומק, מישור ידוע, גודל ידוע, או תנועה בין כמה תמונות.

11. עיוות עדשה

המודל עד כה הוא מודל אידיאלי. עדשה אמיתית מוסיפה עיוות רדיאלי ועיוות משיקי. לכן בתהליך כיול מעריכים גם וקטור עיוות D , הכולל בדרך כלל מקדמים כגון:

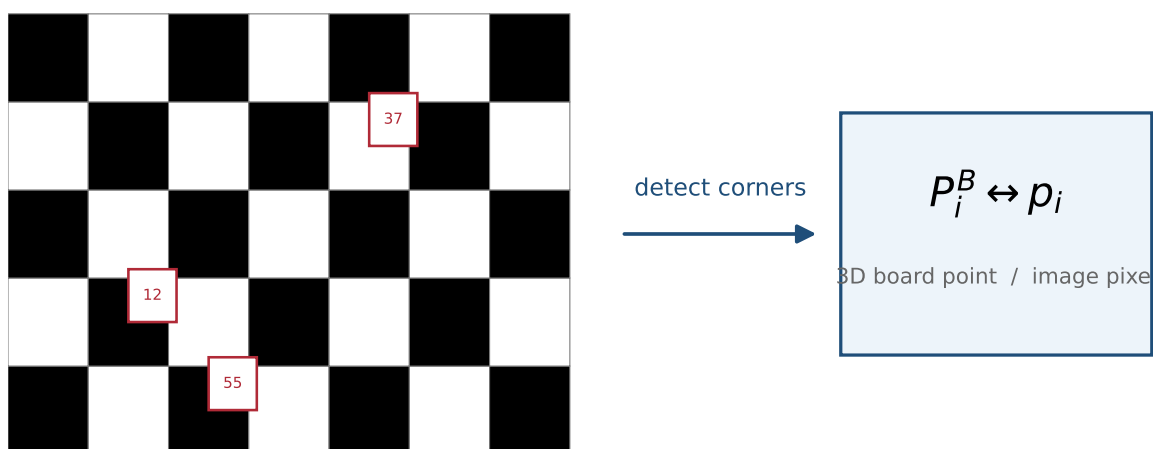
• k_1, k_2, k_3 - עיוות רדיאלי;

• p_1, p_2 - עיוות משיקי.

בפועל ההיטל נכתב בקיצור $p = \pi(K, D, P_C)$. לאחר תיקון העיוות אפשר להשתמש במודל מצלמת הנקב שתואר לעיל.

12. כיול באמצעות Chessboard או ChArUco

בלוח כיול ידועות מראש הקואורדינטות הגיאומטריות של הפינות. אם הלוח מישורי, אפשר להגדיר כל פינה כ- $P_i^B = (X_i, Y_i, 0)$. בתמונה מזהים את הפיקסל המתאים $p_i = (u_i, v_i)$.



מתמונות רבות, בזוויות ובמרחקים שונים, פותרים בעיית מינימום של שגיאת ההקרנה מחדש ומעריכים את K ואת D .

הליך מעשי מומלץ

1. להדפיס לוח מדויק ולהדביקו למשטח קשיח ושטוח.
2. לשמור על אותה עדשה, רזולוציה, zoom ו-focus שבהם תשתמש לאחר הכיול.
3. לצלם בערך 20-40 תמונות מגוונות: מרכז, קצוות, הטיות ומרחקים שונים.
4. לזהות את הפינות ולבצע כיול פנימי.
5. לבדוק שגיאת הקרנה מחדש ולפסול תמונות גרועות.

13. מציאת מיקום וכיוון המצלמה בתלת-ממד

לאחר הכיול הפנימי, נקודה במערכת העולם מועברת למערכת המצלמה באמצעות סיבוב והזזה:

$$P_C = R_{CW}P_W + t_{CW}$$

אחר כך מפעילים את ההיטל הפרספקטיבי. הנוסחה המאוחדת היא:

$$\tilde{p} \sim K[R_{CW} | t_{CW}] \tilde{P}_W$$

ב-OpenCV הפונקציה `solvePnP` מחזירה טרנספורמציה מן העולם או הלוח אל המצלמה. כדי לקבל את מרכז המצלמה במערכת העולם הופכים את הטרנספורמציה:

$$C_W = -R_{CW}^T t_{CW}$$

הפרדה חשובה: הכיול הפנימי מוצא את K, D . הערכת התנוחה מוצאת את R, t . רק שילוב שניהם מאפשר להקריין נקודות עולם לפיקסלים ולמצוא את מיקום וכיוון המצלמה.

14. חיבור למסלול רובוט ולשדה ראייה

אם המצלמה מותקנת על רובוט, בכל זמן מתקבלים מיקום המצלמה $C_W(t)$ וכיוונה $R_{WC}(t)$. ציר הראייה הקדמי של המצלמה הוא $e_z = (0, 0, 1)^T$, ולכן כיוון המבט בעולם הוא:

$$d_W(t) = R_{WC}(t)e_z$$

עבור נקודת מטרה Q_W , וקטור קו הראייה הוא $\ell(t) = Q_W - C_W(t)$. אם חצי מפתח שדה הראייה הוא α , תנאי הראייה הוא:

$$\frac{d_W(t) \cdot \ell(t)}{\|\ell(t)\|} \geq \cos \alpha$$

כך מתחברים יחד מודל מצלמת הנקב, כיול המצלמה, תנוחת המצלמה, מסלול הרובוט וקרני הראייה שעסקנו בהן בתרגילים הקודמים.

סיכום: שרשרת החישוב המלאה

שלב 1 - מן העולם למצלמה

$$P_C = R_{CW}P_W + t_{CW}$$

שלב 2 - חלוקה בעומק

$$x_n = \frac{x_c}{Z_c}, \quad y_n = \frac{y_c}{Z_c}$$

שלב 3 - מעבר לפיקסלים

$$u = f_x \frac{x_c}{Z_c} + c_x, \quad v = f_y \frac{y_c}{Z_c} + c_y$$

שלב 4 - ניסוח פונקציונלי

$$p = \pi(P_C) = \left(f_x \frac{x_c}{Z_c} + c_x, f_y \frac{y_c}{Z_c} + c_y \right)$$

המשפט שכדאי לזכור: P_C היא נקודה במרחב, p היא נקודת ההיטל שלה בתמונה, ו- (u, v) הן הקואורדינטות בפיקסלים של אותה נקודה. הופעת f_x, f_y ו- Z_C נובעת מדמיון משולשים ומהמרת יחידות לפיקסלים.