מבוא לראייה ממוחשבת (22928) – ממ"ן 13

מגיש: דותן אסלמן

<u>שאלה 1</u>

מה תהיה ההשפעה של פילטר מקסימום, פילטר שמחשב את המקסימום בסביבה של NxN פיקסלים? איזה סוג של רעש הפילטר יסיר בצורה יעילה ומה יהיו תופעות הלוואי על התמונה כתוצאה מהפעלת פילטר זה?

פילטר "מקסימום" הינו פילטר לא לינארי שמבצע בחירה, בדומה לפילטר "חציון" שנלמד, רק במוקם לבצע בחירה של הערך החציוני הוא יבחר את הערך המקסימלי.

הפילטר יעיל בהסרת רעשי "פילפל" –פיקסלים שרופים שמתבטאים בערך כהה (שחור) וכן בהחלקת רעש אדטיבי מאזורים 'חלקים' בתמונה (ע"י יישור שלהם למקסימום הסביבתי) למרות שבמקרה של הרעש האדטיבי יכול להיווצר אפקט של הגברת הרעש במידה והערך של הפיקסל יהיה גבוה מהערך ללא הרעש.

הפגיעה העיקרית תהיה ברזולוציה המרחבית של התמונה, כלומר פגיעה בתדרים הגבוהים עד כדי העלמה של פריטים קטנים (קטנים ביחס לחלון הפילטר הנבחר) במידה ואותם פריטים קטנים הם כהים ביחס לסביבה שלהם. בנוסף, הפילטר יזיז קצוות בתמונה לכיוון ש'יגדיל' את הצד הבהיר יותר.

<u>שאלה 2</u>

א. רשום 2 יתרונות ו2 חסרונות של hough transform

יתרונות:

- מתאים גם במידה ויש יותר ממופע אחד של הצורה שמחפשים (לדוג' במקרה של קוים – יותר מקו אחד בתמונה)
- מתמודד טוב עם הסתרות

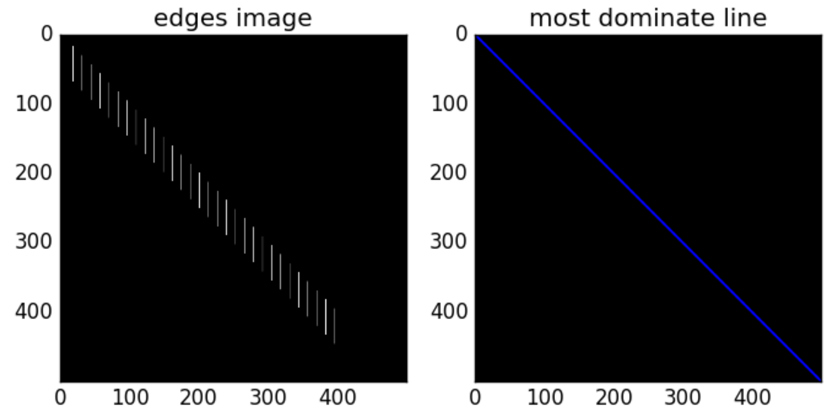
חסרונות:

- חלוקה שונה של המרחב הפרמטרי(יותר/פחות חלוקה של זוויות ומרחקים) משפיעה משמעותית על התוצאות ועל הרגישות לרעש וקשה לבחור את הפרמטרים הנכונים.
- במידה ומחפשים צורות מורכבות המרחב הפרמטרי גדל אקספוננציאלית עם מורכבות המודל ואיתו גם משך החיפוש.

ב. האם גילוי בhough יכול להיות false positive? אם כן- תן דוגמה, אם לא – הסבר למה זה לא אפשרי.

כן, כיוון שכל נק' בתמונת הedges תורמת לכל הקווים האפשריים שעוברים דרכה עלול להיווצר מצב שנצברים votes על קו שלא באמת קיים בתמונה. מצב זה יכול לקרות הן מרעש בתמונה שנצבר להצבעות או מקרים חריגים בתצורת התמונה ש'מרמים' את ההתמרה.

לדוג' יצרתי תמונה בעלת מספר קווים אנכיים שכל נק' בהם תורמת לקו אלכסוני שלא באמת קיים בתמונה, אך הוא הקו הדומיננטי ביותר שhough מוצא מסיבות ברורות



ג. בהנחה שצריך לגלות אוביקט בכל הסיבובים, ובכל הסקאלות האם אפשר להשיג זאת בעזרת GHT?

לא ניתן להשתמש בGHT על מנת למצוא צורות בסיבובים וסקאלות שונים כיוון שהvoting יאספו במרחב הפרמטרי היא לפי הזווית והמרחק של הנגזרת הכיוונית לכיוון מרכז המסה של הצורה, ואם נראה את הצורה בזוויות שונות נזהה כיווני נגזרת שונים שיגזרו 'הצבעה' במרכז מסה לא נכון. כנ"ל לגבי שינויי סקאלה שיגזור הצבעות למרכזי מסה 'על אותו הקו' של המרכז הנכון אך לא במקום הנכון בגלל הבדלי הסקאלה. מהסיבה הזו לא ניתן להשתמב בGHT בצורה שהיא אדישה לסיבוב וscale.

<u>שאלה 3</u>

הגדר בקצרה את המושגים הבאים

א. Camera intrinsic calibration parameters - הפרמטרים הפנימיים של המצלמה : 5 רמות חופש הכוללים את : אורך מוקד ביחידות של פיקסלים לכל ציר (2 ערכים, fx, fy במטריצה), זווית הskew של הפיקסל במידה ואינו מלבן מושלם(ערך אחד, s במטריצה) וההיסט של הפיקסלים ל(0,0) של התמונה (2 ערכים, x0,y0 במטריצה)
הערכים מיוצגים ע"י מטריצה 3x3 ונהוג לסמנה באות K:

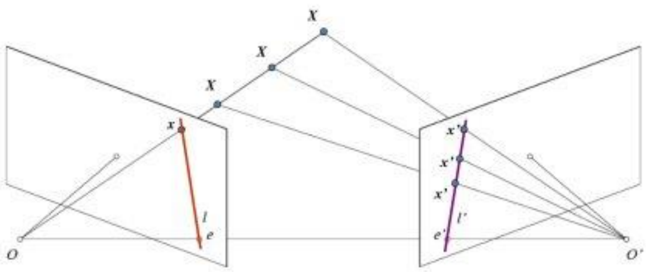
$$K = \begin{bmatrix} fx & s & x_0 \\ 0 & fy & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

ב. Camera extrinsic calibration parameters – הפרמטרים החיצוניים של המצלמה : 6 רמות חופש הכוללים את מיקום המצלמה ביחס לעולם התלת מימדי (x,y,z) והpose של המצלמה (זווית) ביחס ל3 הצירים.
נהוג לסמן :

[R | t] ,   t = -RC

כאשר R – מטריצת הסיבוב ב3 הצירים ו C וקטור שמכיל את הקורדינטות של מיקום המצלמה ביחס למע' הצירים של העולם.

ג. Epipoles and epipolar lines –
בסטריאו – בהינתן נק' בתמונה א' הישר האפיפולרי שלו הינו הישר עליו תהיה הנק' התואמת לה בתמונה ב'. הנקודה האפיפולית הינה הנקודה דרכה עוברים כל הישרים האפיפולרים של תמונה מסויימת.



את נק' האפי-פול של כל מצלמה בצילום סטראו ע"י נתוני מיקומי המצלמות ללא תלות בסצינה בתור נק' החיתוך של הbaseline (הקו העובר בין מרכז ההטלה של מצלמה א' למרכז ההטלה של מצלמה ב') עם אזור התמונה (למעשה אפשר להסתכל על האפיפול בתור המיקום בתמונה בו מופיע מרכז ההטלה של המצלמה השניה) יש לציין שהאפיפול יכול ליפול גם מחוץ לתמונה עצמה אך יופיע על המישור האינסופי שלה.

ד. Svd – singular value decomposition שיטה לפירוק מטריצה מלבנית **A** ל3 מטריצות :

$$A = U\Sigma V$$

כאשר U ו V הינם אורטונורמליות וΣ מטריצה אלכסונית אשר מכילה את הערכים הסינגולריים של A.

מספר הערכים הסינגולריים השונים מ0 מבטא את ה rank של המטריצה

השיטה דומה לחישוב eigenvalues ו eigenvectors (eigen decomposition) רק ללא הצורך של המטריצה להיות ריבועית ו SDP אך ניתן באמצעותה לחשב PCA.

ה. Bayes rule – חוק בייס, משוואה שמאפשרת לנו לחשב הסתברות מותנית של θ בהינתן D על בסיס ההסתברות המותנית ההפוכה (D בהינתן θ) וההסתברות בנפרד של כ"א מהמשתנים, או באופן מלא:

$$\boldsymbol{P}(\theta|\mathrm{D}) = \frac{\boldsymbol{P}(\mathrm{D}|\theta)\mathrm{P}(\theta)}{P(D)}$$

ניתן גם להסתכל על זה באופן סמנטי יותר כך:

$$\boldsymbol{posterior} = \frac{\boldsymbol{likelihood} * \boldsymbol{prior}}{\boldsymbol{evidance}(\boldsymbol{normalizer})}$$

ו. Essential matrix – המטריצה ההכרחית, מטריצה 3x3 שמבטאת את הקשר בין נקודה תלת מימדית y הנצפית במצלמה א' לנקודה התואמת לה 'y הנצפית במצלמה ב', הקשר מבטא את האילוצים האפיפולרים ומוגדר באופן הבא:

$$(\boldsymbol{y}')^{\boldsymbol{T}}\boldsymbol{E}\boldsymbol{y} = \boldsymbol{0}$$

כאשר E – המטריצה ההכרחית.

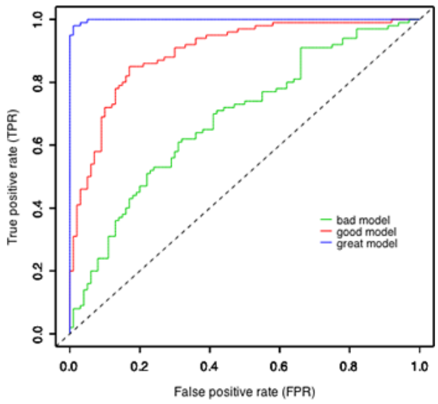
ז. Fundamental matrix – המטריצה היסודית, מטריצת 3x3 בעלת תפקוד דומה למטריצה ההכרחית רק שעובדת על יחידות של פיקסלים בדו מימד בין 2 התמונות בסטראו, למעשה מדובר בשילוב של המטריצה ההכרחית עם מטריצות הקליברציה של כל אחת מהמצלמות.

$$(\boldsymbol{x}')^{\boldsymbol{T}}\boldsymbol{F}\boldsymbol{x} = (\boldsymbol{x}')\boldsymbol{M}_{\boldsymbol{l}}^{-\boldsymbol{1}}\boldsymbol{E}\boldsymbol{M}_{\boldsymbol{r}}\boldsymbol{x} = \boldsymbol{0}$$

כאשר $\boldsymbol{M}_{\boldsymbol{r}}$ ו $\boldsymbol{M}_{\boldsymbol{l}}$- הינן מטריצות הקליברציה של המצלמה הימנית והשמאלית בהתאמה. x ו 'x הינן הנק' בדו מימד(בתמונה) בניגוד לנק' בתלת מימד שהופיעו במטריצה ההכרחית.

באמצעות המטריצה היסודית ניתן לחשב לכל נקודה בתמונה א' את הישר האפיפולרי התואם לה בתמונה ב'

ח. ROC curve - Receiver operation characteristic curve גרף להצגה/מדידת איכות של classifier בינארי. ציר X מכיל את היחס בין FP (התראות שווא) לבין כמות הדוגמאות השליליות שיש (מה שלמעשה נותן את ה False positive rate) וציר Y מכיל את היחס בין ה TP (התראות אמת) לבין כמות הדוגמאות החיוביות, מה שנותן את ה True positive rate. המטרה כמובן של קלאסיפייר טוב היא TPR גבוה תוך שמירה על FPR מינימאלי וניתן לחשב את הAUC – area under the curve בתור מדד לטיב של מדד זה.
לדוג :



<u>שאלה 4</u>

מה הקשר בין ה essential matrix לבין הfundamental matrix?

שתי המטריצות מתייחסות לקשר בין נק' שצולמה במצלמה א' למיקום שלה בתמונה ב'. ה essential matrix מתייחסת למיקום של הנק' שצולמה ע"י מצלמה א' לפי המערכת של מצלמה ב' **אך עדיין בתלת מימד** ועוד הfundamental matrix מתייחסת ליחס בין המיקומים בדו מימד וביחידות של פיקסלים, ובכך משלבת בתוכה גם את נתוני הכיול הפנימיים של כ"א מהמצלמות.

$$(\boldsymbol{x'})^{\boldsymbol{T}}\boldsymbol{Fx} = (\boldsymbol{x'})\boldsymbol{M_l^{-1}EM_r x} = \ \boldsymbol{0}$$

כאשר $\boldsymbol{M_r}$ ו $\boldsymbol{M_l}$- הינן מטריצות הקליברציה של המצלמה הימנית והשמאלית בהתאמה. x ו x' הינן הנק' בדו מימד(בתמונה) בניגוד לנק' בתלת מימד שהופיעו במטריצה ההכרחית.

כיצד מאפשרת ה fundamental matrix לשפר את מציאת ההתאמות בין נקודות עניין בין 2 תמונות?

בהינתן נק' x בתמונה שצולמה ע"י מצלמה א' הכפלה שלה עם המטריצה היסודית F תיתן לנו את משוואת הישר האפיפולרי עליו נמצאת הנקודה התואמת לx בתמונה שצולמה ע"י מצלמה ב', שיוויון זה נובע מהמשוואה הבאה:

$$(\boldsymbol{x'})^{\boldsymbol{T}}\boldsymbol{Fx} = 0$$

כאשר *F* – המטריצה היסודית. כלומר הנק' החדשה נמצאת על הישר שנוצר מהמכפלה של המטריצה עם הנקודה ולכן תוצאת המכפלה היא 0.

שאלה 5

א. מהי טרנספורמציה פרוייקטיבית?

תיאור כללי להתמרה הפיכה שממפה ישרים לישרים.

בפועל מדובר במטריצה בגודל 3x3 שמאפשרת לעשות טרנזלציה, סיבוב, scale, שיקוף ושינויי פרספקטיבה לתמונה.

מקרים פרטיים של התמרה פרוייקטיבית הם התמרה אפינית, similarity , טרנסלציה וכו'

נקרא גם הומוגרפיה.

ב. נתונות 2 קבוצות של נקודות מתאימות במישור, תאר אלגוריתם לשיערוך הטרנספורמציה הפרויקטיבית בין 2 הקבוצות ע"פ עיקרון lease squares?

כדי למצוא את מטריצת הטרנספורמציה H (בגודל 3 על 3) נשטח אותה לוקטור h ונפתור את מערכת המשוואות הלינארית

$$Ah = 0$$

כאשר כל צמד נקודות תואמות יוצרות 2 משוואות שאפשר לבטא באמצעות 2 שורות במטריצה A

$$\begin{aligned} \mathbf{a}_x &= (-x_1, -y_1, -1, 0, 0, 0, x_2'x_1, x_2'y_1, x_2')^T \\ \mathbf{a}_y &= (0, 0, 0, -x_1, -y_1, -1, y_2'x_1, y_2'y_1, y_2')^T . \end{aligned}$$

ו Aכולו יבנה באופן הבא:

$$A = \begin{pmatrix} \mathbf{a}_{x1}^T \\ \mathbf{a}_{y1}^T \\ \vdots \\ \mathbf{a}_{xN}^T \\ \mathbf{a}_{yN}^T \end{pmatrix} .$$

נפתור את בעיית הleast square ההומוגנית, כלומר

$$\min|Ah|^2$$

נבצע פירוק לערכים סינגולריים למטריצה A וניקח את הוקטור הסינגולרי אשר תואם לערך הסינגולרי הנמוך ביותר וזה יהיה הוקטור h, נעשה לו reshape חזרה ל3 על 3 על מנת לקבל מטריצה.

מה מספר ההתאמות המינימלי לצורך שיערוך?

כיוון שהומוגרפיה מלאה מכילה 8 פרמטרים נצטרך 4 התאמות(כלומר סהכ 8 נק') כיוון שכל התאמה מייצרת 2 משוואות וכך בפועל יהיו לנו 8 משוואות ו8 נעלמים.

ג. בהנחה שחלק מההתאמות אינו נכון, תאר אלגוריתם לשערוך הטרנספורמציה במצב זה

ניתן במקרה הזה להשתמש ב**RANSAC**, ע"פ שיטה זו נגריל 4 התאמות, נחשב עבורן הומוגרפיה באמצעות least square ונבדוק את ההתאמה של ההומוגרפיה לשאר הנקודות (נטיל אותן ונראה אם הן קרובות לנקודות התואמות להן), נעשה זאת עד T פעמים או עד אשר קיבלנו "קונצנזוס" (סף מסויים של נקודות מסכימות עם השיערוך), ברגע שהגענו לעצירה ניקח את ההגרלה הטובה ביותר

(זו שצברה קונצנזוס גדול ביותר) ונחשב על כל הנקודות בקונצנזוס least square וזו תהיה התוצאה הסופית

שאלה 6

א. האם מטריצת ההטלה P יכולה להכיל פרמטרים שמפצים על עיוותי העדשה?

לצערינו לא, עיוותי העדשה (radial distortion) הינם עיוותים לא לינארים ולכן אי אפשר לפצות עליהם באמצעות המודל הלינארי של מטריצת ההטלה. דרך פשוטה להבין שהעיוותים לא לינארים הם שעיוותי עדשה יכולים 'לעקם' קווים ישרים בתמונה וזה סותר את ההגדרה של טרנספורמציות לינאריות בהן קוים ישרים נשארים קווים ישרים(ראה הגדרת הומוגרפיה בשאלה 5)

ב. כתוב 2 רשימות, אחת מתארת את הפרמטרים הפנימיים של המצלמה והשניה את החיצוניים

פנימיים (סהכ 5):

- אורך המוקד ביחידות של פיקסלים לכל ציר – 2 ערכים
- ההסתה ממרכז ההטלה ל(0,0) בתמונה – 2 ערכים
- ה*skew* של פיקסל על הגלאי (מבטא את הזווית של הפיקסל במקרה והוא לא בדיוק מלבני) – ערך 1

וכך נראת מטריצת הפרמטרים הפנימיים:

$$K = \begin{bmatrix} fx & s & x_0 \\ 0 & fy & y_0 \\ 0 & 0 & 1 \end{bmatrix}$$

חיצוניים (סהכ 6):

- המיקום *C=(x,y,z)* - מיקום מרכז ההטלה של המצלמה ביחס לעולם – 3 ערכים
- מטריצת *R* – מטריצת סיבוב שמבטאת את זווית ההסתכלות של המצלמה ביחס למע' הצירים של העולם, משלבת בתוכה סיבוב ב3 הצירים ולכן מכילה 3 ערכים

בהרכבה המטריצה נראית כך :

[R | t] , t = -RC

ג. הראה כיצד אפשר להציג את מטריצת ההטלה P ע"י מכפלה של 2 מטריצות, אחת מייצגת את הפרמטרים הפנימיים ואחת את החיצוניים :

מטריצת הקליברציה המלאה *P* (שילוב של פנימיים וחיצוניים) מתוארת באופן הבא:

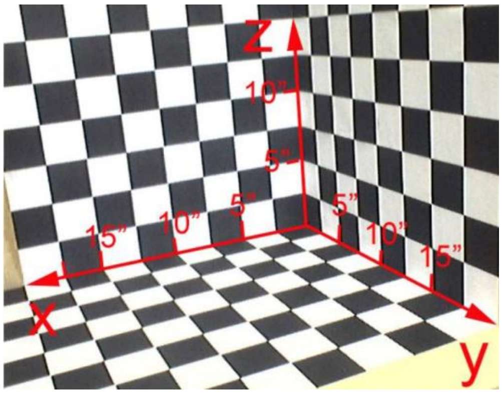
$$P = K[R|t]\,, t = -RC$$

ד. הראה את השלבים של אלגוריתם לשערוך המטריצה P מתוך תמונה אחת של "אובייקט קליברציה", מה הנתונים שמוזנים לאלגוריתם?

הקלט של אלגורתם זה הוא נקודות בתלת מימד והמיקום שלהם בתמונה בדו מימד (ראשית הצירים בעולם התלת מימדי נבחרת 'אקראית' וכל השאר ביחס אליה) ועכשיו יש לנו צמדים של

(u,v) <-> P

כיוון שיש לנו 11 רמות חופש במטריצת הקליברציה יש צורך ב6 התאמות (כל התאמה 2 משוואות סך הכל 12 משוואות ו11 נעלמים)



ניתן מכל צמד כזה לייצר את 2 המשוואות הבאות :

$$(m_1 - u_i m_3)P_i = 0$$

$$(m_2 - v_i m_3)P_i = 0$$

לכל צמד i, כאשר m_j הינה השורה ה j במטריצת הקליברציה M (אותה אנו רוצים לחשב)

ניתן לצבור משוואות אלו למטריצה גדולה ולבצע שוב least square כפי שמתואר בשאלה 5 (נבצע svd וניקח את הוקטור הסינגולרי לו יש את הערך הסינגולרי הנמוך ביותר ואז נבצע reshape ל3 על4)

<u>שאלה 7</u>

נתון מצלמת *pinhole* בעלת הנתונים : $\text{center} = (320,240)$, $f = 500, S_x = S_y = 1$. ההתמרה לקורדינטות ה"עולם" נתונה ע"י $X_{cam} = RX_{world} + T$.

כאשר $T = [70,95,120]^T$, $R = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

א. ***כתוב את מטריצת ההטלה של המצלמה***:

מטריצת הפרמטרים הפנימיים:

$$K = \begin{bmatrix} 500 & 0 & 320 \\ 0 & 500 & 240 \\ 0 & 0 & 1 \end{bmatrix}$$

מטריצת הפרמטרים החיצוניים הינה מטריצת 4 על 4 מהצורה:

$$\begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix}$$

נחשב באמצעותם את P

$$P = [K|0]\begin{bmatrix} R & T \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 500 & 320 & 73400 \\ -500 & 0 & 240 & 76300 \\ 0 & 0 & 1 & 120 \end{bmatrix}$$

ב. ***מה תהיה הנקודה בתמונה שמתאימה לנק' Xw = [150, 200, 400]***

נחשב זאת ע"י הכפלה במטריצה P

$$x = PX_w = [301400,97300,520]^T = [\mathbf{579.61}, \mathbf{187.11}, \mathbf{1}]^T$$

שאלה 8

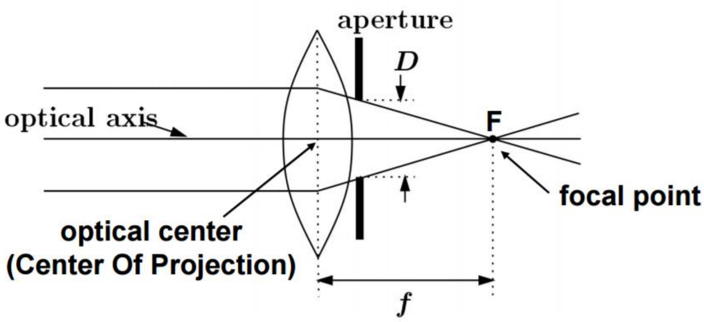
א. תחת אילו תנאים ה vanishing point של קו ישר בתמונה תופיע במישור התמונה באינסוף?

אם הקו הישר יהיה מקביל לצירי התמונה (כלומר- על מישור התמונה לאורך ציר X או ציר Y) הvanishing point שלו יהיה ב'אינסוף'

ב. ***תחת אילו תנאים תוטלנה יותר מנקודה אחת של אותו קו ישר בעולם לאותה vanishing point במישור התמונה?***

אם הקו הישר הוא אנכי למישור התמונה (כלומר מקביל לציר Z) הנקודות ה'רחוקות' שלו יופיעו על אותה נקודה.

ג. ***נתונה נק' בעולם P = (400,600,1200) מוטלת ל p = (24,36) במערכת קורדינטות ניצבת לציר Z, ומוקד המצלמה נמצא ב (0,0,f) , מהו f ?***



במצב אידאלי (המצלמה מיושרת לZ ומרכז ההטלה בראשית הצירים) זה ניתן לחשב את p ע"י צמד המשוואות הפשוט :

$$p(x) = f * X/Z$$

$$p(y) = f * Y/Z$$

במקרה הספציפי שלנו (ומספיק להסתכל על משוואה אחת)

$$f * \frac{600}{1200} = \frac{f}{2} = 36 => \boldsymbol{f = 72}$$

נוודא מול המשוואה השניה:

$$\frac{f}{3} = \frac{72}{3} = 24$$