

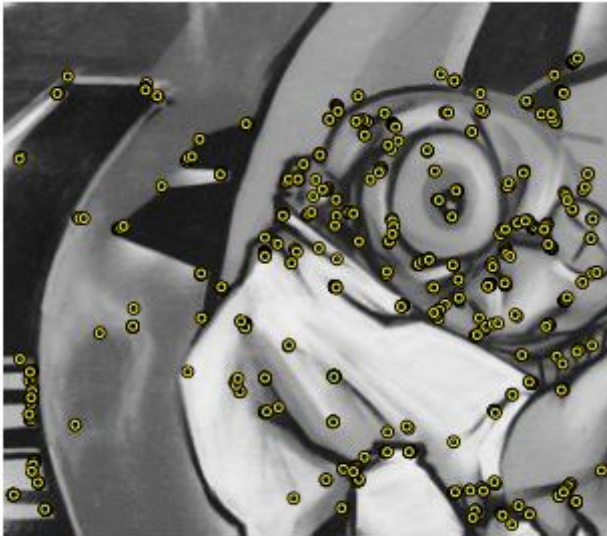
Local Features And Matching For Object Instances

Presented by:

Guy Katz

שיטת זיהוי מבוסס עצמים

- למידת העצם
- יצירת פיצורים לעצמים שבתמונה
- יצירת מסד נתונים לכול עצם יחד עם הפיצורים שמתאימים לכל עצם.
- מציאת עצמים בתמונה.
- ברגע שמוצג עצם בתמונה נוספת גילוי הפיצורים שלו.
- התאמה של הפיצורים של העצמים בתמונה הנוספת לפיצורים של העצמים במסד הנתונים



יתרונות בפיצרים מקומיים לא ברי-שינוי.

- מקומיות

- בגלל שהפיצרים הם מקומיים אז הם חסינים לאי-סדר (clutter), הסתרה (occlusion)

- יחודיות

- כל פיצרים בנפרד יכול להיות מותאם למאגר מידע גדול.

- כמות

- לכל אוביקט יכולים להוצר הרבה פיצרים.

- יעילות

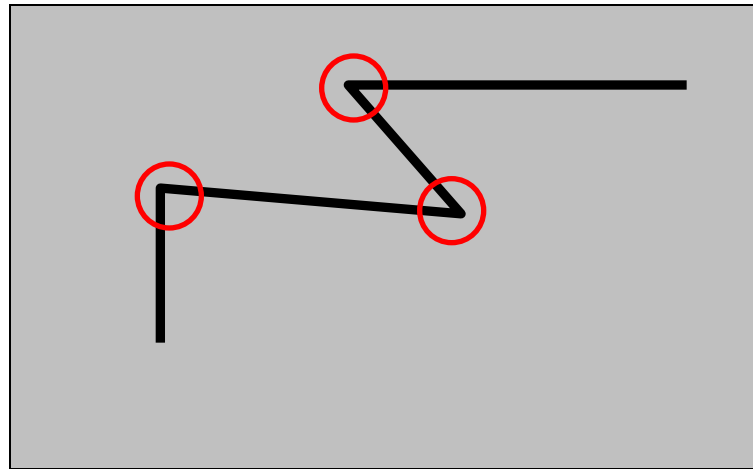
- זמני הריצה מאד קרובים לזמן ריצה ב"זמן-אמת".

יתרונות (המשך...)



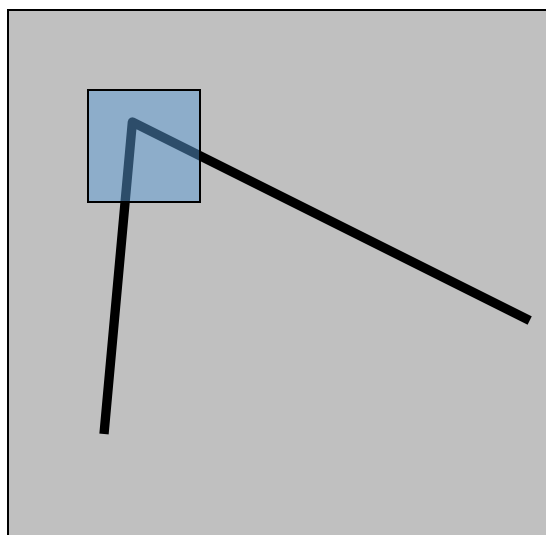
דוגמא למציאת נקודות פוטנציאליות

Harris corner detector

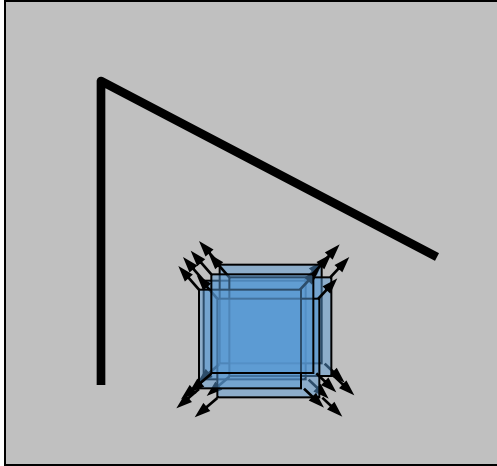


הרעיון:

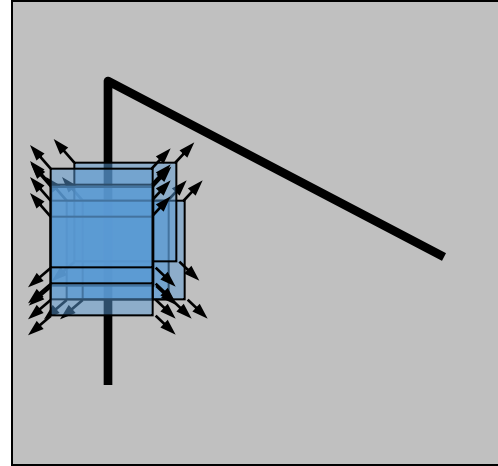
- זיהוי פשוט של נקודת עניין בהסתכלות דרך חלון קטן.
- אם החלון יוזז אז נקבל שינוי גדול ברמת העצימות.



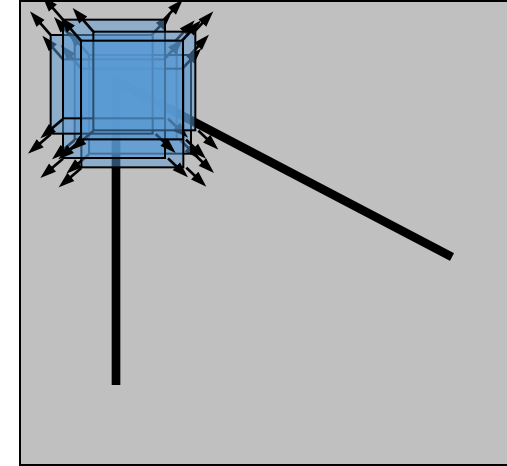
Harris Detector: Basic Idea



איזור "שטוח" :
אין שינוי בכל הכיוונים



איזור "שפה" :
אין שינוי בכיוון השפה



איזור "פינה" :
שינוי משמעותי בכל
הכיוונים

Harris Detector: מתמטיקה

שינוי בעוצמה עבור הזזה $[u, v]$:

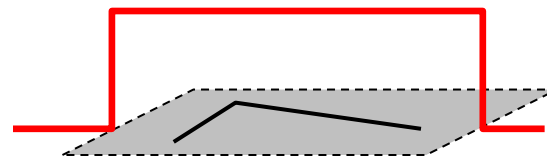
$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

Window
function

Shifted
intensity

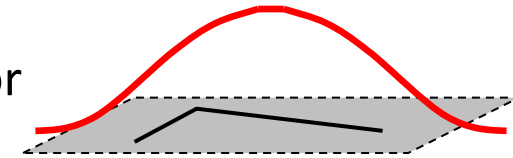
Intensity

Window function $w(x, y) =$



1 in window, 0 outside

or



Gaussian

Harris Detector: מתמטיקה

$$E(u, v) = \sum_{x, y} w(x, y) [I(x+u, y+v) - I(x, y)]^2$$

עבור קטנים $[u, v]$: $I(x+u, y+v) = I(x, y) + uI_x + vI_y$

נקבל:

$$E(u, v) = \sum_{x, y} w(x, y) \left\| \begin{bmatrix} I_x(x, y) & I_y(x, y) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \right\|^2 =$$

$$\begin{bmatrix} u & v \end{bmatrix} \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} u & v \end{bmatrix} M \begin{bmatrix} u \\ v \end{bmatrix}$$

מתמתיקה: Harris Detector

עבור הזזה קטנה של $[u, v]$ נקבל הערכה בילינארית:

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

כאשר המטריצה M היא מטריצה 2×2 מחושבת מהנגזרות של העצם בתמונה:

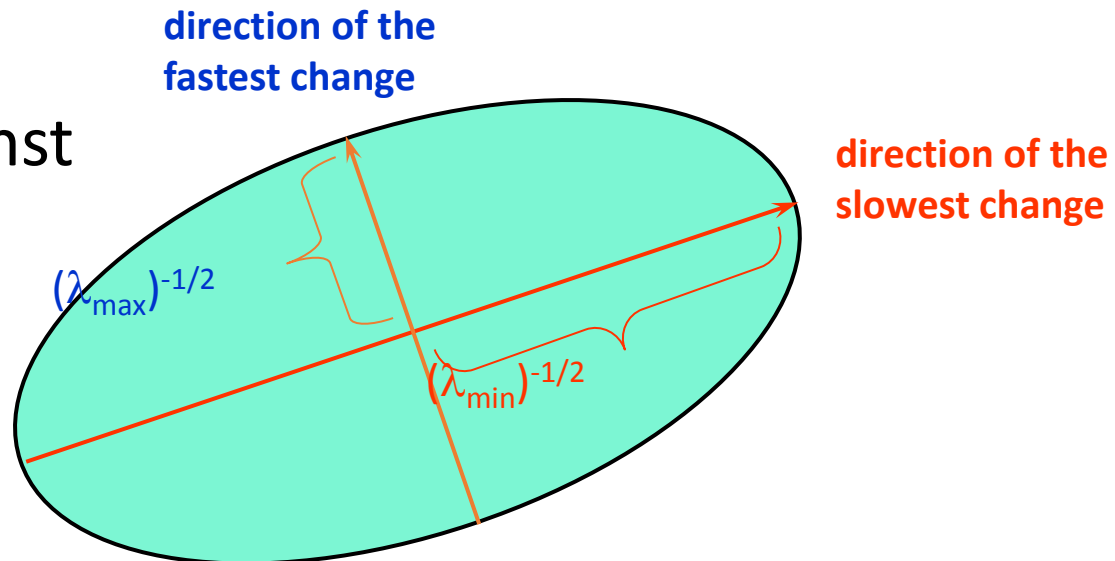
$$M = \sum_{x, y} w(x, y) \begin{bmatrix} I_x^2 & I_x I_y \\ I_x I_y & I_y^2 \end{bmatrix}$$

Harris Detector: מתמטיקה

שינוי בעוצמה בהזזה בחלון: ערכים עצמים

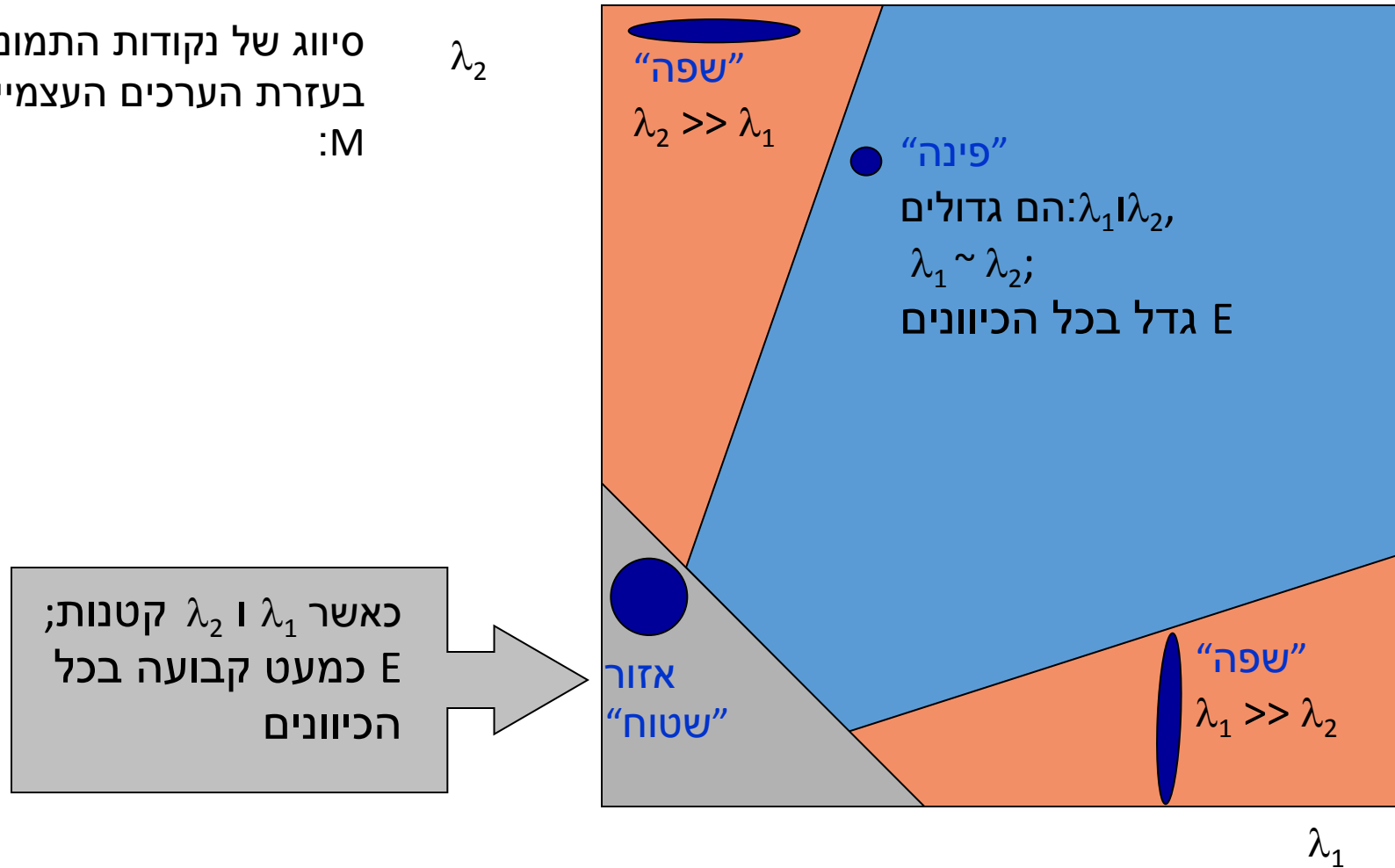
$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix} \quad \lambda_1, \lambda_2 - \text{eigenvalues of } M$$

$E(u, v) = \text{const}$ אליפסה



Harris Detector: מתמתיקהב

סיווג של נקודות התמונה
בעזרת הערכים העצמיים של
M:



Harris Detector: מתמטיקה

קביעת מדד עבור הפינות:

$$R = \det M - k (\text{trace } M)^2$$

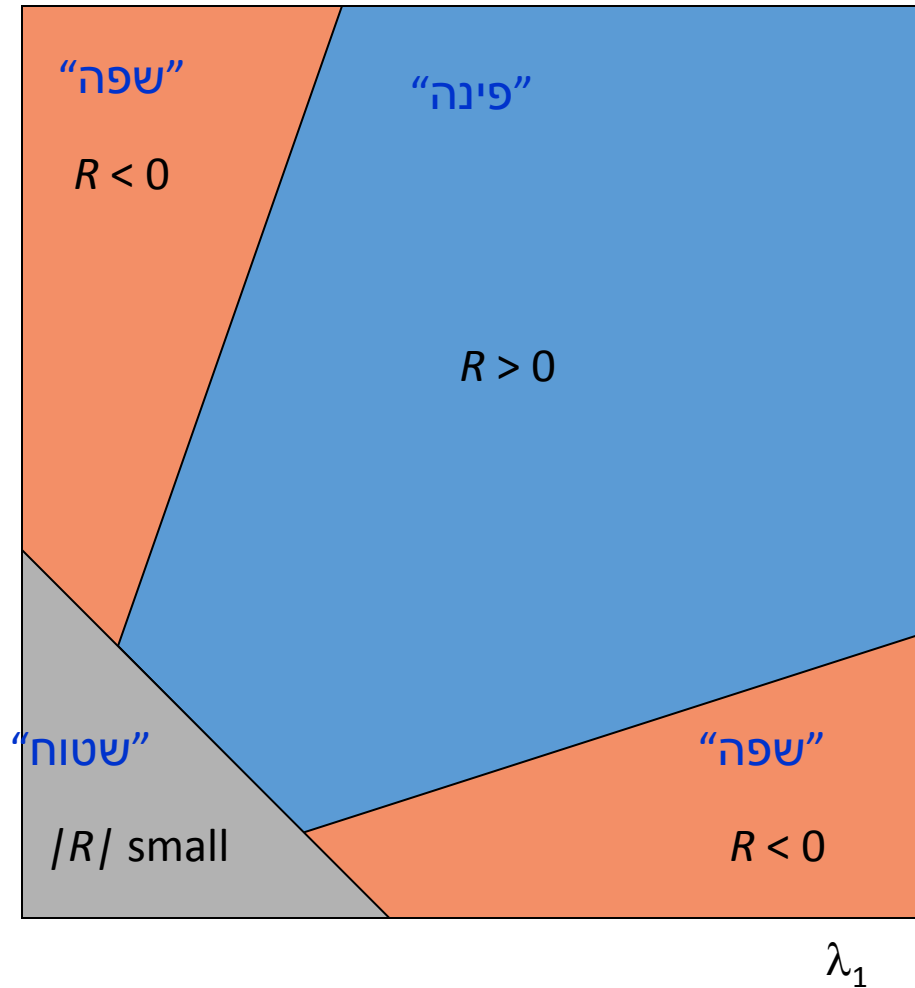
$$\det M = \lambda_1 \lambda_2$$

$$\text{trace } M = \lambda_1 + \lambda_2$$

(k – empirical constant, $k = 0.04-0.06$)

Harris Detector: מתמטיקה

- λ_2
- R תלוי רק בערכים העצמיים של M
 - R גדול – פינה
 - R הוא שלילי עם אבל גדול – שפה
 - $|R|$ הוא קטן – שטוח



Harris Detector

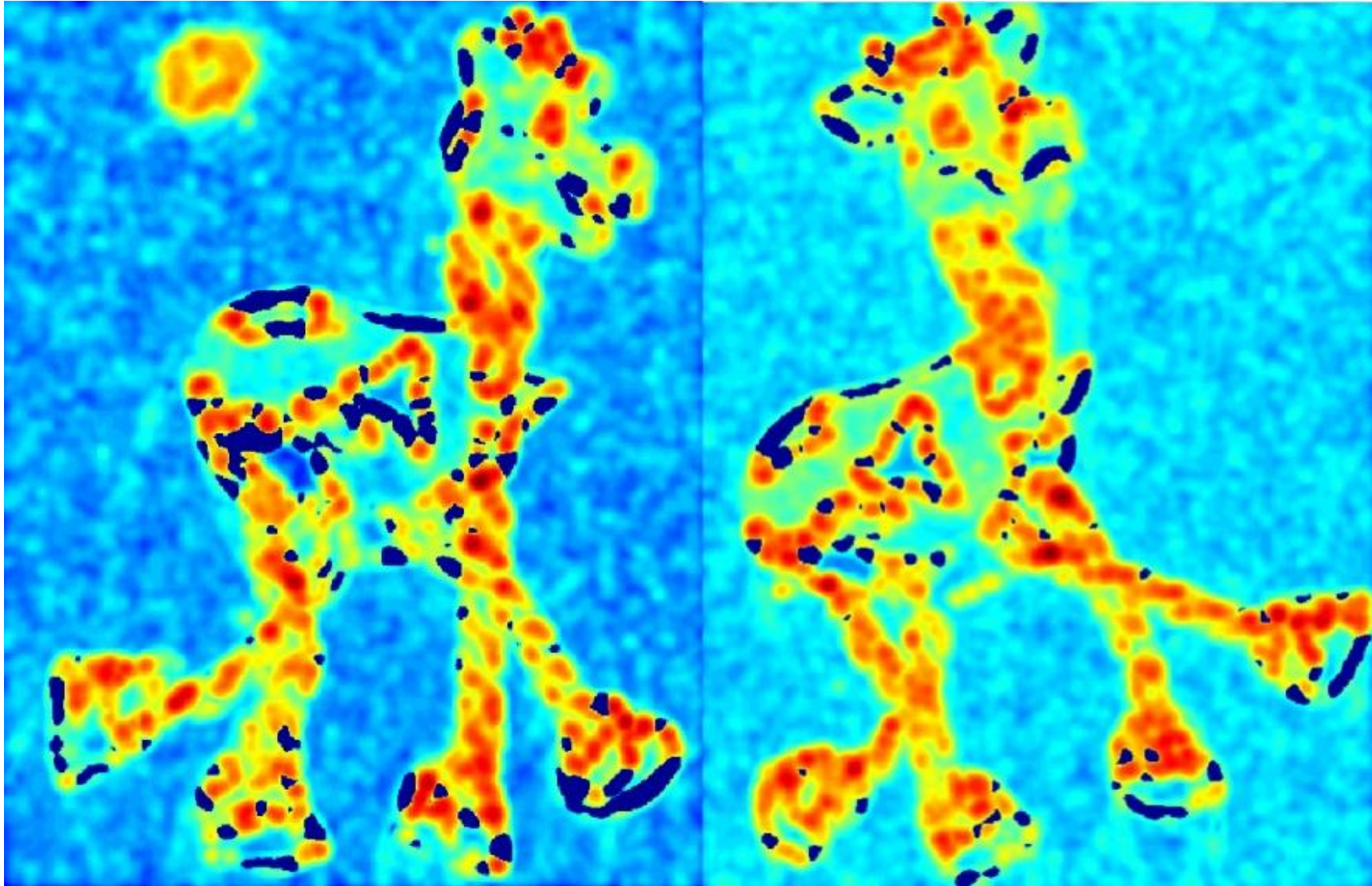
- אלגוריתם:
- מציאת פינות אשר ינפקו R גדול ($\text{threshold} < R$)
- בחירת הנקודות אשר ה R שלהם הוא מקסימום מקומי.

Harris Detector: תרשים עבודה



Harris Detector: תרשים עבודה

חישוב פינות בהתאם ל R



Harris Detector: תרשים עבודה

מציאת נקודות עם פינה גדולה בהתאם $R > \tau$



Harris Detector: Workflow

בחירת נקודות אשר להם R מקסימום מקומי



Harris Detector: Workflow



Harris Detector: סיכום

- שינוי העוצמה הממוצע בכיוון $[u, v]$ ניתן לביטוי על ידי במשוואה ביליניארית

$$E(u, v) \cong [u, v] M \begin{bmatrix} u \\ v \end{bmatrix}$$

- תאור של הנקודה במונחים של ערכים עצמים של מטריצה M :
קביעת ערך מדיד לפינה

$$R = \lambda_1 \lambda_2 - k(\lambda_1 + \lambda_2)^2$$

- נקודה שהיא פינה צריכה להיות בעלת שינוי עוצמה גדול בכל הכיוונים, R צריך להיות גדול וחיובי

יצירת SIFT

- שלב ראשון – יצירת פרמידת DOG:

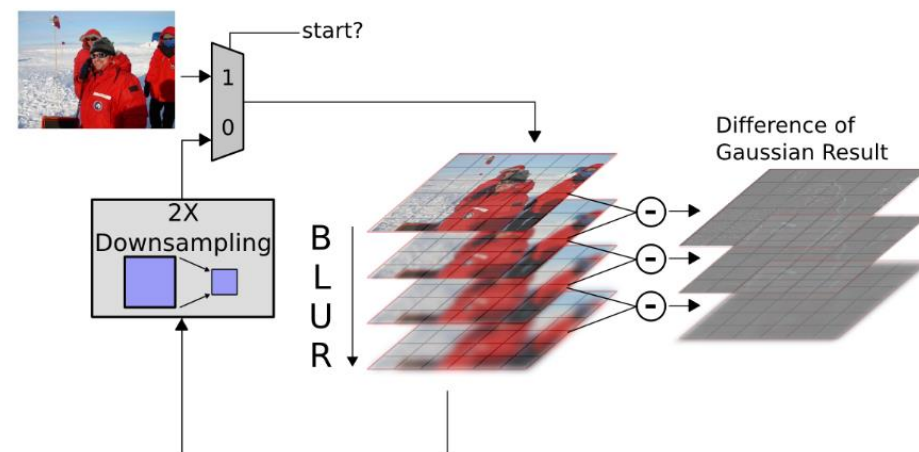


Figure 1: Difference of Gaussian pyramid generation

Difference of Gaussian ScaleSpace Pyramids *

for SIFT Feature Detection

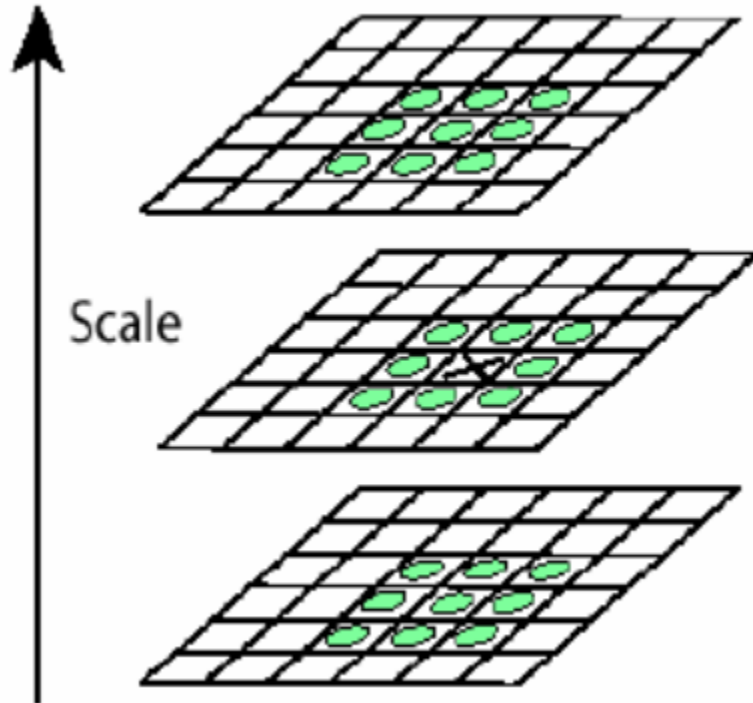
Ballard Blair and Chris Murphy

- יצירה של 3 רמות ב3 אוקטבות שונות – שינוי הס

יצירת SIFT

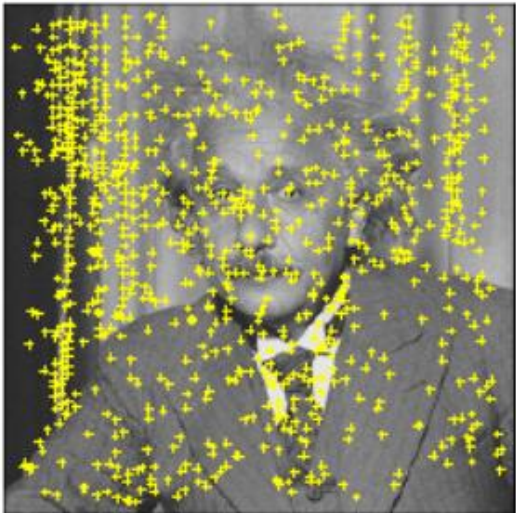
שלב שני – מציאת מקסימום/מינימום מקומיים:

- בודקים אם פיקסל הוא מקסימום/מינימום ברמה הראשונה
- אם כן, ממשיכים לרמה הבאה ובודקים בשאר הרמות. אם לא, עוברים לפיקסל הבא.
- מתקדימים בשאר הרמות לבדיקת מקסימום/מינימום
- אם הפיקסל הוא מקסימום/מינימום בכל הרמות אז הוא נבחר בתור מקסימום/מינימום

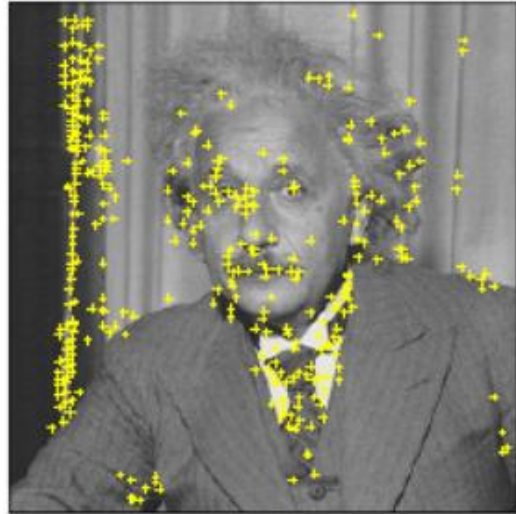


יצירת SIFT

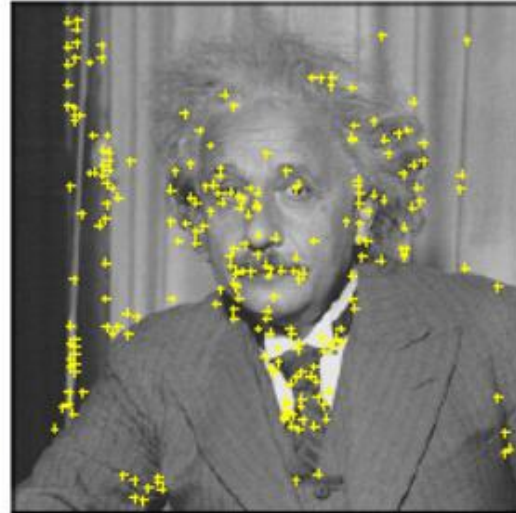
- שלב שלישי – הסרת נקודות לא יציבות.
- הסרת נקודות בעלות ניגודיות נמוכה
- כאשר ההבדל אינו מספיק מובהק.
- נקודות לאורך הקצוות
- עוזר בשימור כנגד שינוי בסיבוב וקנה מידע.



Scale-space extremas



After low contrast removal



After removal of edge responses

יצירת SIFT

יצירת דסקריפטור

- עבור כל נקודת מפתח מחושב גראדיאנט בכל הכיוונים:

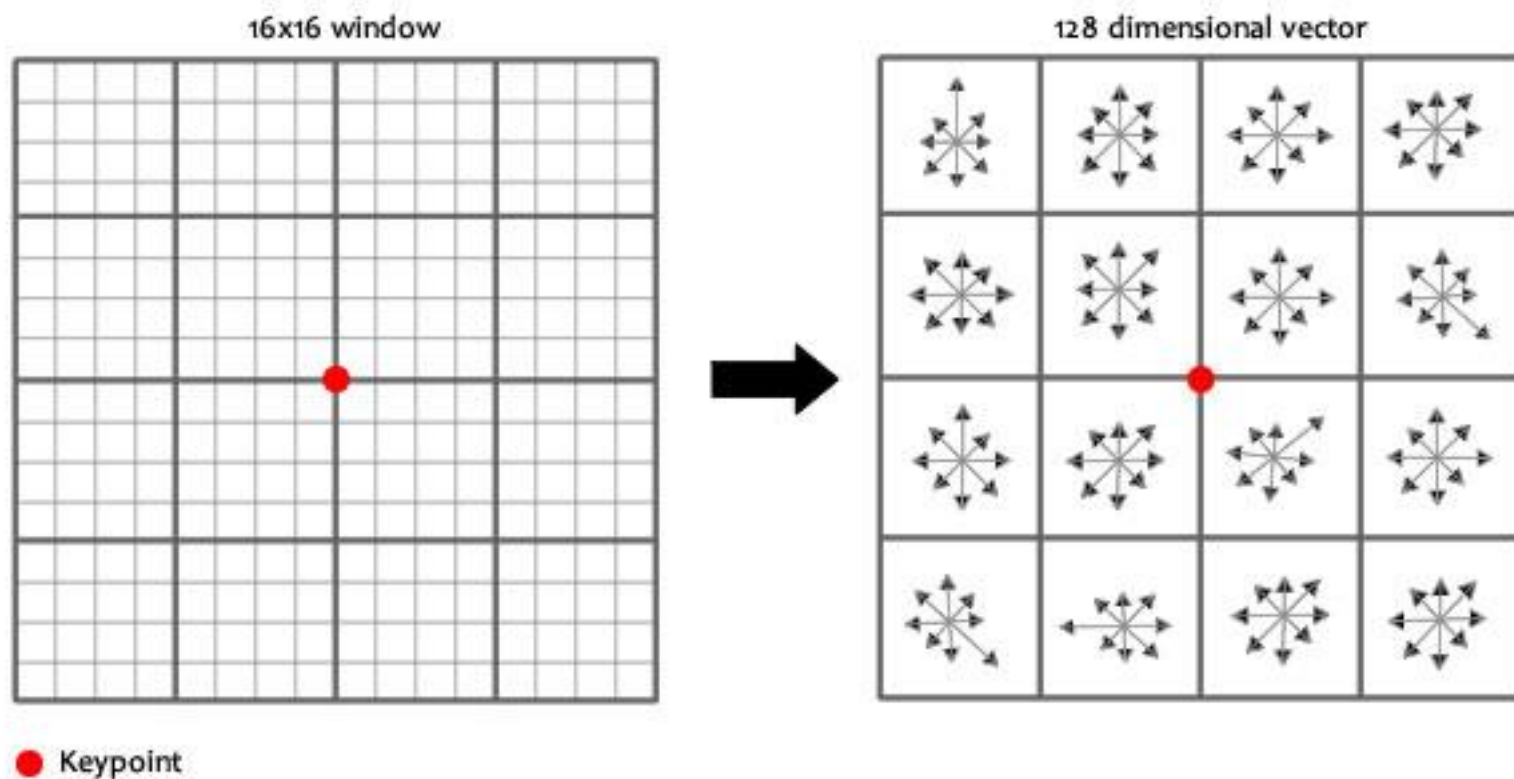
- גודל $M_{ij} = \sqrt{(A_{ij} - A_{i+1,j})^2 + (A_{ij} - A_{i,j+1})^2}$

- כיוון $R_{ij} = \text{atan2}(A_{ij} - A_{i+1,j}, A_{i,j+1} - A_{ij})$

- מחלקים את התמונה לחלונות בגודל 128 פיקסלים.

יצירת SIFT

יצירת דסקריפטור (המשך..)



יצירת SIFT

יצירת דסקריפטור (המשך..)

- יוצרים היסטוגרמה עבור עם מספר bin-ים קבוע.
- מחפשים מהו הכיוון הדומיננטי ואותו רושמים ראשון.
- רושמים את שאר הכיוונים לפי סדר הכיוונים.
- לבסוף נוצר וקטור עם 128 ערכים – דסקריפטור.

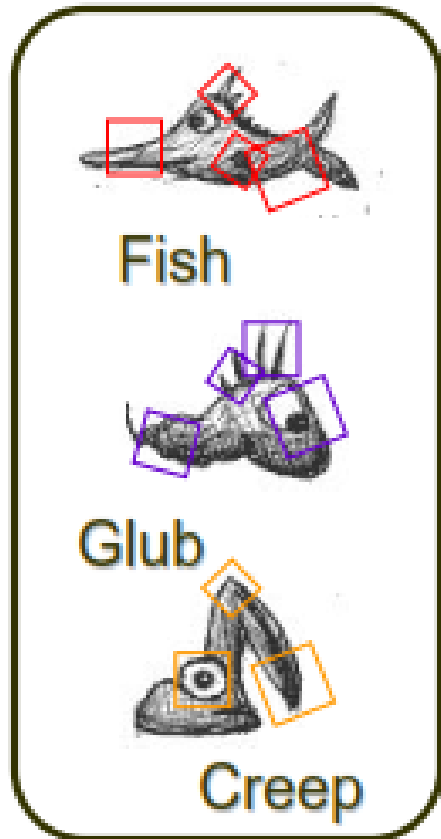
זיהוי עצמים - SIFT

• אימון המערכת לזיהוי עצמים – יצירת פיצרים

ואיחסונם במאדר המיד

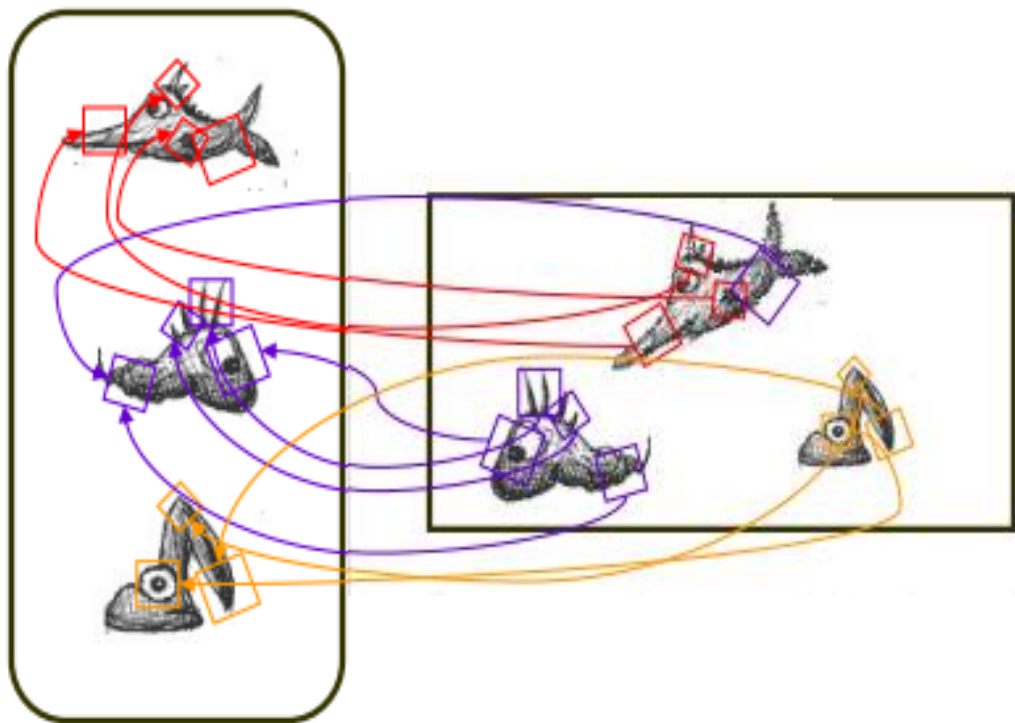
• הפיצרים מכילים קורדינאטות של הדסקריפטור של העצם.

• ידוע לאיזה עצם כל פיצר שייך.



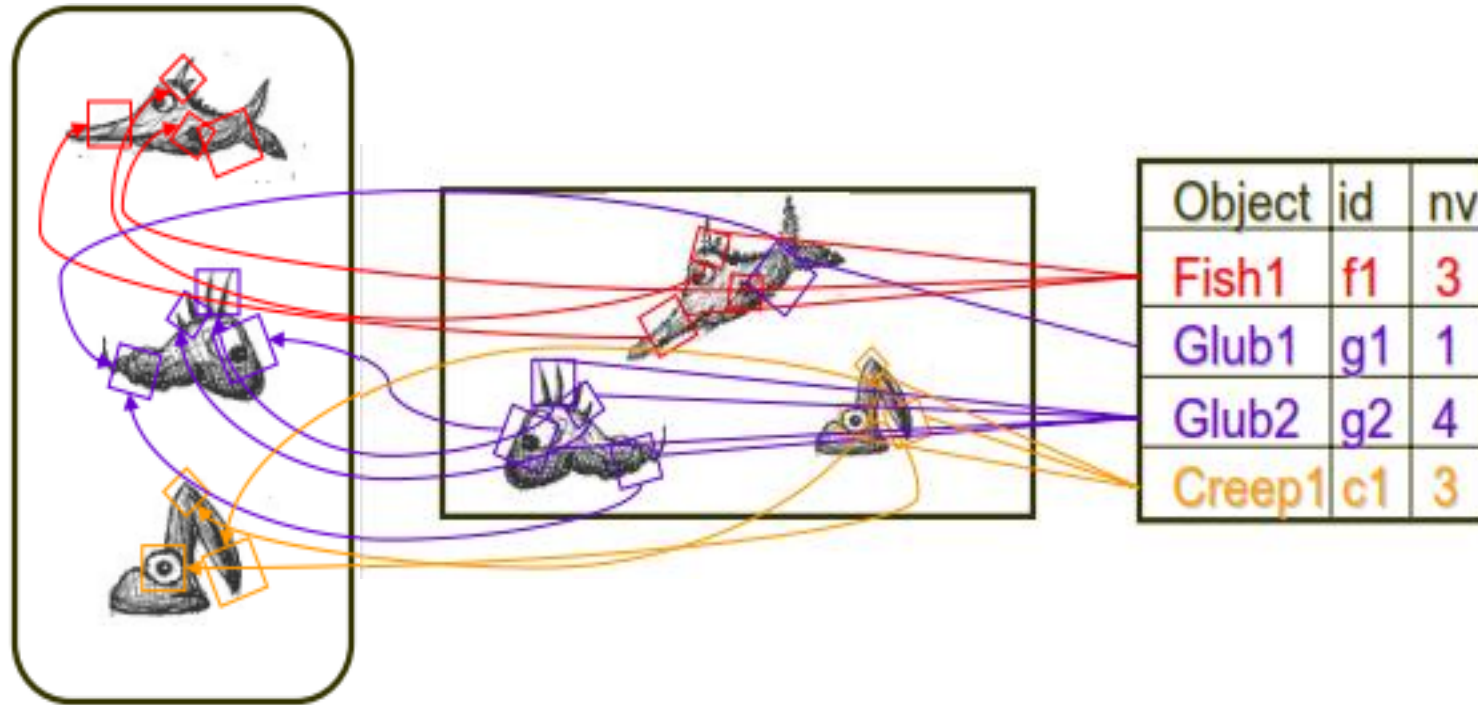
זיהוי עצמים - SIFT

- בשלב בא אנחנו רוצים לזהות את העצמים בתמונה החדשה.
- נניח כי אלו התוצאות ויש שגיאה אחת.



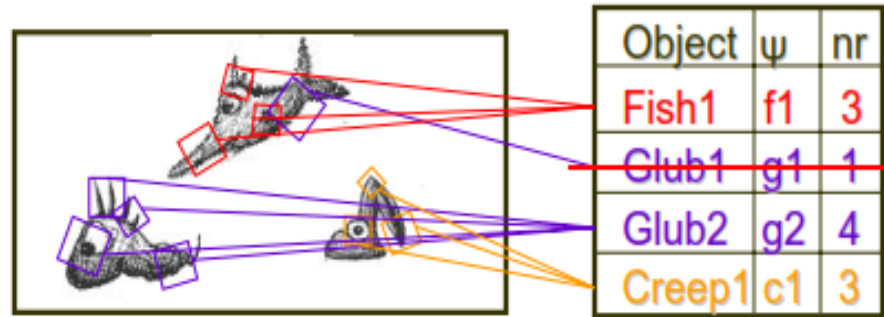
זיהוי עצמים - SIFT

- כל מפתח מצביע עבור עצם ידוע בתמונת המקור – בהתאמה למיקום, קנה מידע ואוריינטציה.



זיהוי עצמים - SIFT

- כל עצם שקיבל פחות מ-3 הצבעות נמחק.
- לבסוף מנסים להתאים גאומטריה.

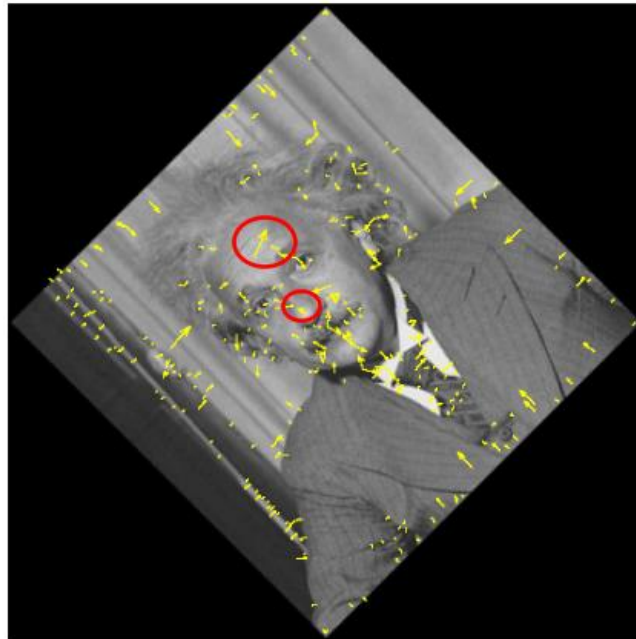
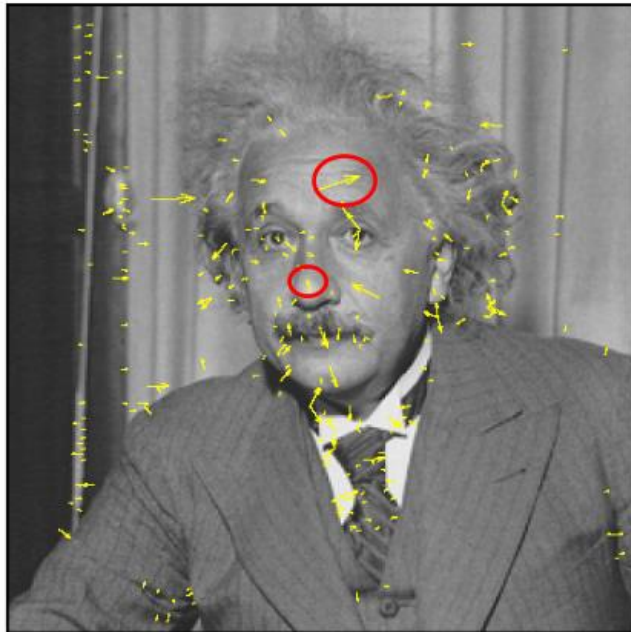


תכונות האי-שינוי של הדסקריפטור

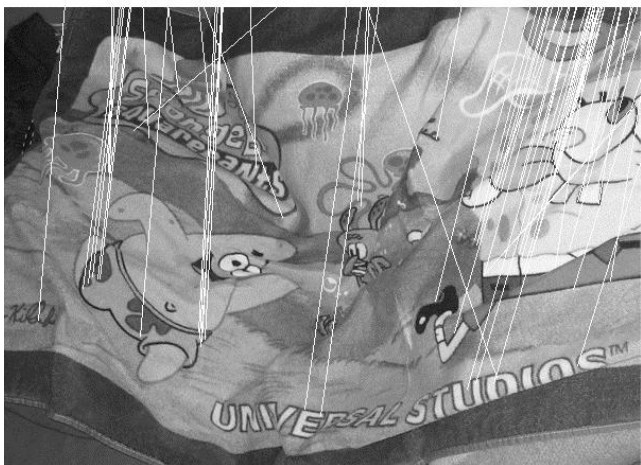
- אינו משתנה בסיבוב.
- אינו משתנה בקנה מידה.
- אינו משתנה בשינויים תאורה.
- אינו משתנה בטרנספורמציות אפינניות.

אי-שינוי בסיבוב + קנה מידה

- עמידות הדסקריפטור בפני סיבוב נובעת מכך שהדסקריפטור בנוי מכיוונים קונוניים – תמיד הכיוון הדומיננטי יוצמד לכיוון הקונוני הראשון.
- עמידות הדסקריפטור נובעת מכך שכאשר בחרנו את הנקודות הנקודות הנבחרות היו מקסימום/מינימום בכמה קנה מידה שונים.



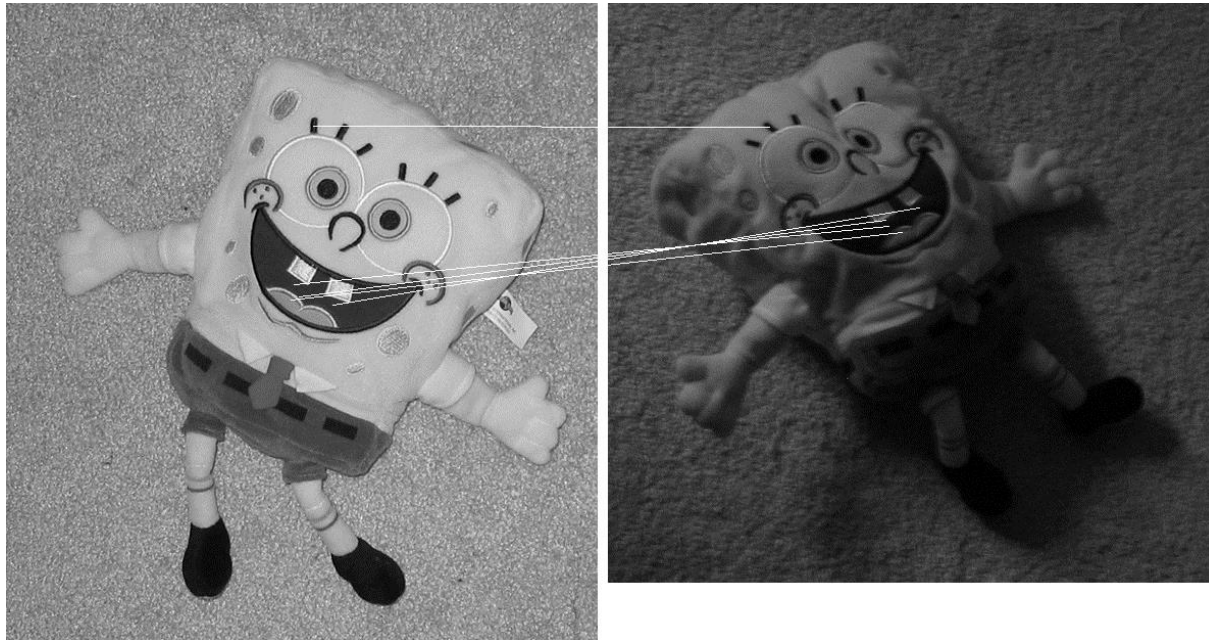
אי-שינוי בטרנספורמציה אפינית



- עמידות הדסקריפטור לטרנספורמציה אפינית
נובע מכך שהנקודות הן מקסימום/מינימום מקומי
של אזור קטן יחסית וההשפעה של שינויים גדולים
היא די קטנה.

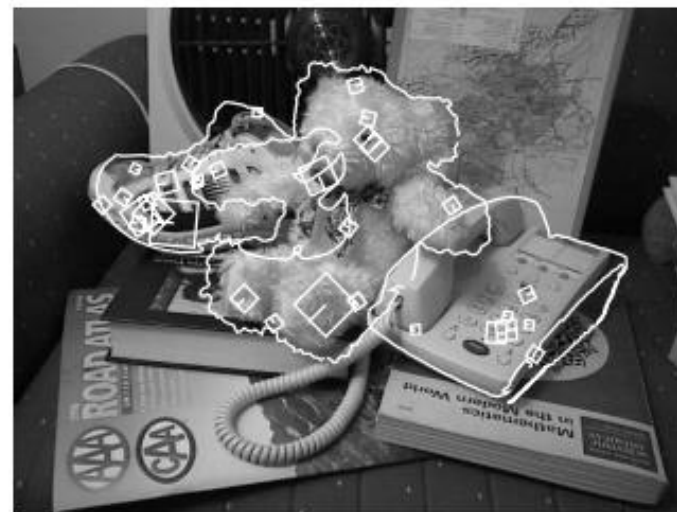
אי-שינוי בתאורה

- מכיוון שהדסקריפטור בנוי מהגראדיאנט השינוי בתאורה אינו משפיע על גודל הגראדיאנט.



ניסויים ותוצאות

זיהוי עצמים והסתרה



ניסויים ותוצאות

זיהוי עצמים ותאורה



ניסויים ותוצאות

עמידות / חסינות

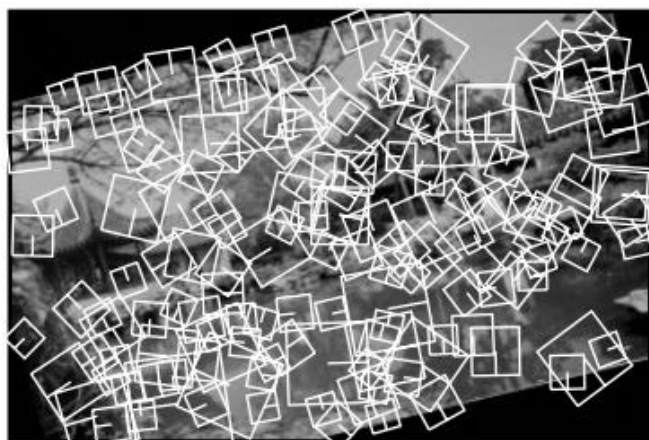
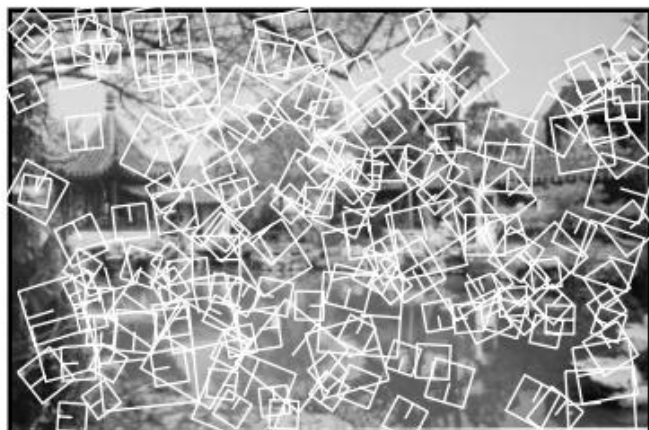


Image transformation	Match %	Ori %
A. Increase contrast by 1.2	89.0	86.6
B. Decrease intensity by 0.2	88.5	85.9
C. Rotate by 20 degrees	85.4	81.0
D. Scale by 0.7	85.1	80.3
E. Stretch by 1.2	83.5	76.1
F. Stretch by 1.5	77.7	65.0
G. Add 10% pixel noise	90.3	88.4
H. All of A,B,C,D,E,G.	78.6	71.8

מקורות מידע:

- Object Recognition from Local Scale-Invariant Features, Lowe, ICCV 1999.
- Local Invariant Feature Detectors: A Survey, Tuytelaars and Mikolajczyk. Foundations and Trends in Computer Graphics and Vision, 2008.
- Matching with Invariant Features, Darya Frolova, Denis Simakov, The Weizmann Institute of Science, March 2004.
- **“Distinctive Image Features from Scale Invariant Features”**, *International Journal of Computer Vision*, 2004, by David Lowe.