15-381
ARTIFICIAL INTELLIGENCE

LECTURE 15:
VISION III: FEATURES

FALL 2010

Several slides taken from: Kristen Grauman, Alexei Efros, David Lowe

SESSION 1: OCTOBER 21, 2010

Tim Wilson	1.6-Bit Pattern Databases
Jinyoung Park	A Neural Implementation of the Kalman Filter
William Devanny	A New Algorithm for Weighted Partial MaxSAT
Zachary Sparks	A Temporal Proof System for General Game Playing
Vivek Pai	Adapting to the Shifting Intent of Search Queries
Derek Basehore	Algorithms for Finding Approximate Formations in Games
Madhav Bhagat	An Online Algorithm for Large Scale Image Similarity Learning
Justin Zhu	An Online Algorithm for Large Scale Image Similarity Learning
Dustin Hellstern	An Online Algorithm for Large Scale Image Similarity Learning
Keith Ainsworth	Argumentation System with Changes of an Agent's Knowledge Base
Steven Williams	Asynchronous Multi-Robot Patrolling against Intrusions in Arbitrary Topologies
Nicole Carter	Asynchronous Multi-Robot Patrolling Against Intrusions in Arbitrary Topologies
Eric Wu	Automatic Attribution of Quoted Speech in Literary Narrative
Daniel Chen	Canadian Traveler Problem with Remote Sensing
Anthony Hugh	CAO: A Fully Automatic Emoticon Analysis System
Vikram Rajkumar	CAO: A Fully Automatic Emoticon Analysis System
Travis Mandel	CAO: A Fully Automatic Emoticon Analysis System
Karl Hellstern	Constructing Topological Maps using Markov Random Fields and Loop-Closure Detection
Laura Glenndenning	Creating Dynamic Story Plots with Continual Multiagent Planning
Michael Wang	Decentralised Coordination of Mobile Sensors Using the Max-Sum Algorithm
Jon Boerner	Efficient Online Learning and Prediction of Users' Desktop Actions
Ben Parr	Finding Optimal Solutions to Cooperative Pathfinding Problems
Amos Yuen	Finding Optimal Solutions to Cooperative Pathfinding Problems
Malcolm Greaves	From Generic Knowledge to Specific Reasoning for Medical Image Interpretation Using Graph based Representations
Dev Doshi	Functional Network Reorganization in motor cortex can be explained by reward-modulated Hebbian Learning
Aaron Hsu	Integrating Constraint Satisfaction and Spatial Reasoning
James Moffatt	Learning Models of Object Structure
Anurag Mengle	Learning Simulation Control in General Game-Playing Agents
Phil Brown	Multi-task Gaussian Process Learning of Robot Inverse Dynamics
Haw-Shiuan Chang	Neural Implementation of Hierarchical Bayesian Inference by Importance Sampling
Nathan Herzing	New Improvements in Optimal Rectangle Packing

Q&A (5min)

Q&A (5min)

Q&A + BUFFER (15 min)

Q&A (5min)

SESSION II: OCTOBER 26, 2010

Ryan Cahoon	Nonparametric Curve Extraction Based on Any Colony System
Yuzhou Xin	On Stochastic and Worst-case Models for Investing
Nick Sidawy	Optimal Rectangle Packing on Non-Square Benchmarks
Zhiwei Huang	Parallel Depth First Proof Number Search
Arjun Sinha	Predicting the Importance of Newsfeed Posts and Social Network Friends
Sarun Savetsila	Probabilistic Collision State Checker for Crowded Environments
Timothy Carson	Reasoning with Lines in Euclidean Space
Tarush Aggarwal	Search Space Reduction Using Swamp Hierarchies
Torrey Brenner	Searching Without a Heuristic: Efficient Use of Abstraction
Rohit Madhu	Sensing and Predicting Shared Bicycling Usage in the City
Kapil Easwar	Single-Frontier Bidirectional Search
Alexandre Rebert	Single-Frontier Bidirectional Search
William Mitchell	Solving Stochastic Games
Todd Eisenberger	TBA*: Time Bounded A*
Adam Mihalcin	The Genetic Algorithm as a General Diffusion Model for Social Networks
Apaorn Suveepattananont	Topological Relations between Convex Regions
Sanil Shah	Towards an Intelligent Code Search Engine
Akshay Dave	Training a Multilingual Sportscaster: Using Perceptual Context to Learn Language
Faye Han	Truth, Justice, and Cake Cutting
Franklin Ta	Truth, Justice, and Cake Cutting
Joe Appel	UCT for Tactical Assault Planning in Real-Time Strategy Games
Katherine Brady	Urban Security: Game-Theoretic Resource Allocation in Networked Physical Domains
Te Gao	Using Closed Captions as Supervision for Video Activity Recognition
Jason McDonald	Using Reasoning Patterns to Help Humans Solve Complex Games
Christopher Tomaszewski	Using Stereo for Object Recognition
Mo Fahmy	Visual Contextual Advertising: Bringing Textual Advertisements to Images
Ian Clanton-Thuon	Visual Contextual Advertising: Bringing Textual Advertisements to Images
Flavia Grosan	Visual Contextual Advertising: Bringing Textual Advertisements to Images
Melissa Wagner	What if the Irresponsible Teachers Are Dominating?
Huaishu Peng	What Is an Opinion About Exploring Political Standpoints Using Opinion Scoring Model
Dustin Haffner	Whose Vote Should Count More: Optimal Integration of Labels from Labelers of Unknown Expertise

Q&A (5min)

Q&A (5min)

Q&A + BUFFER (15 min)

Q&A (5min)

LIGHTNING ROUND

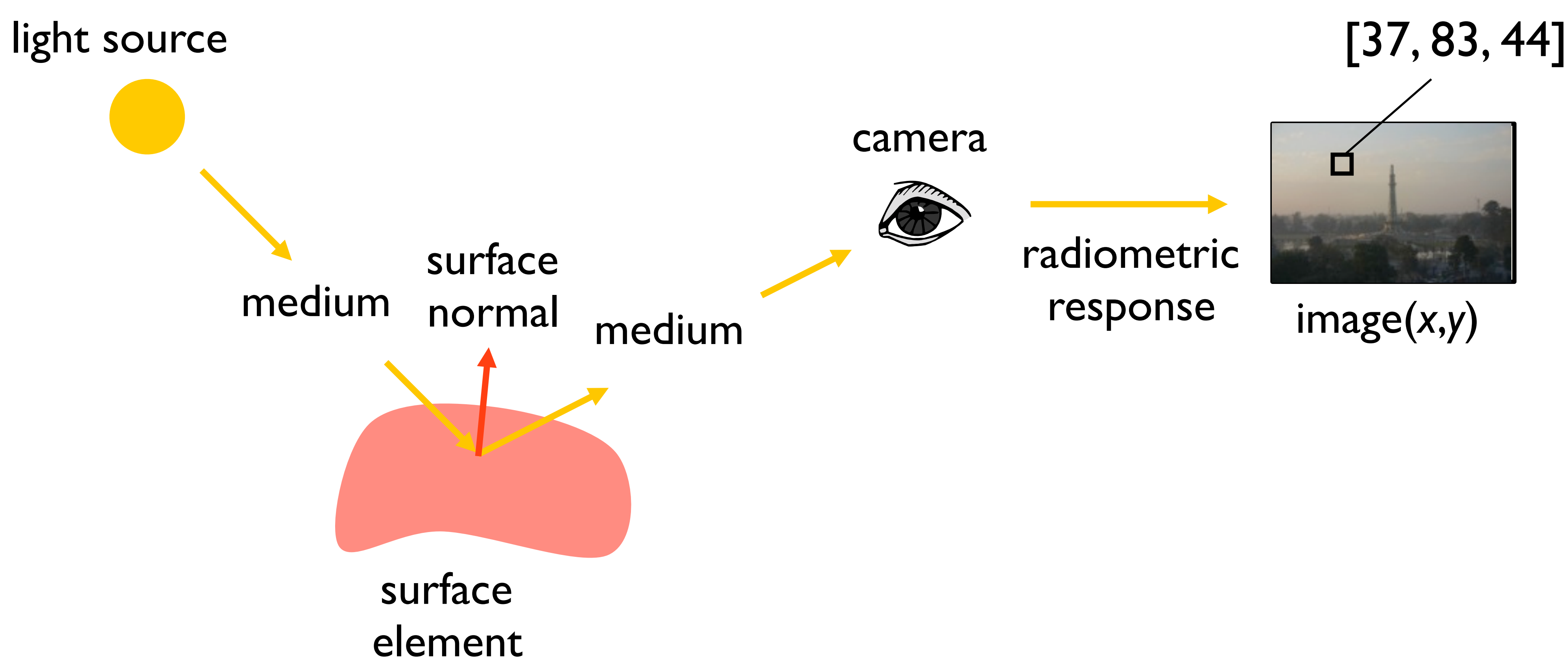
- REHEARSE!
- SALIENCY: DON'T TRY TO INCLUDE EVERYTHING
- SUNSHINE: BE POSITIVE
- RAIN: BE CRITICAL
- REMEMBER GRANDMA: 10 WORDS
- **REHEARSE!**

LIGHTNING ROUND

- 1.5 MINUTE OVERVIEW OF PAPER
- CRITERIA FOR GRADING:
  - ALGORITHMIC EXPLANATION: INPUT, OUTPUT, PROCESS
  - WHY IS THIS PAPER INTERESTING: APPLICATION OR ELEGANCE
  - HOW CAN I IMPROVE THIS?
  - HOW INTERESTING THE SELECTED PAPER IS

VISION REVIEW

RADIOMETRY



Scene Irradiance → Scene Radiance → Image Irradiance → Intensity

Light Source → BRDF → Radiometric Image Formation → Camera Response Function →

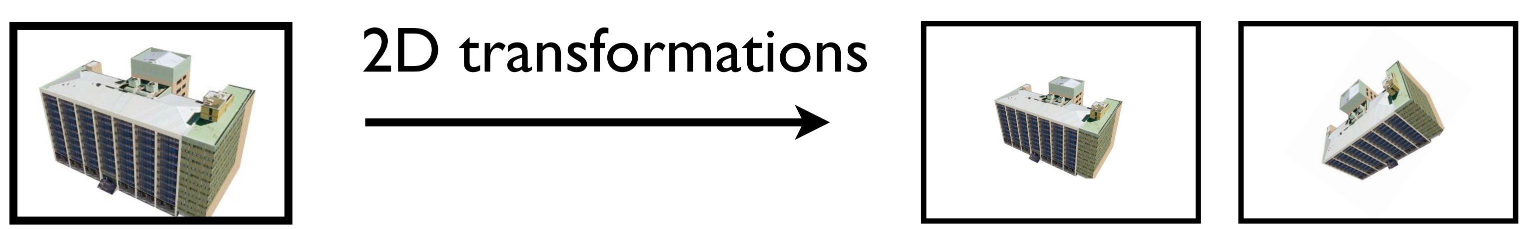
VISION REVIEW

GEOMETRY

$$\mathbf{X} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

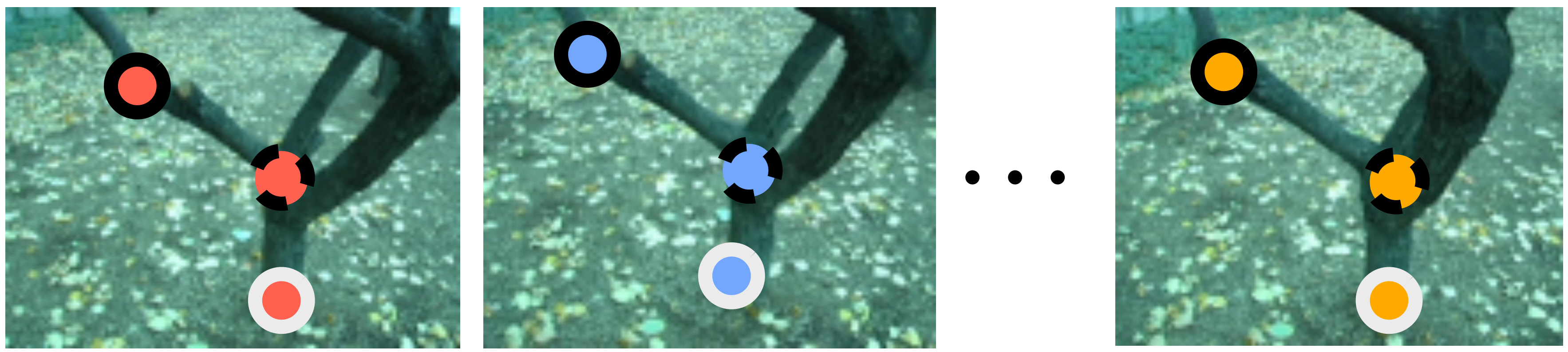
P Projection matrix

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix}$$



VISION REVIEW

SCENE RECONSTRUCTION



1 2 … *F*

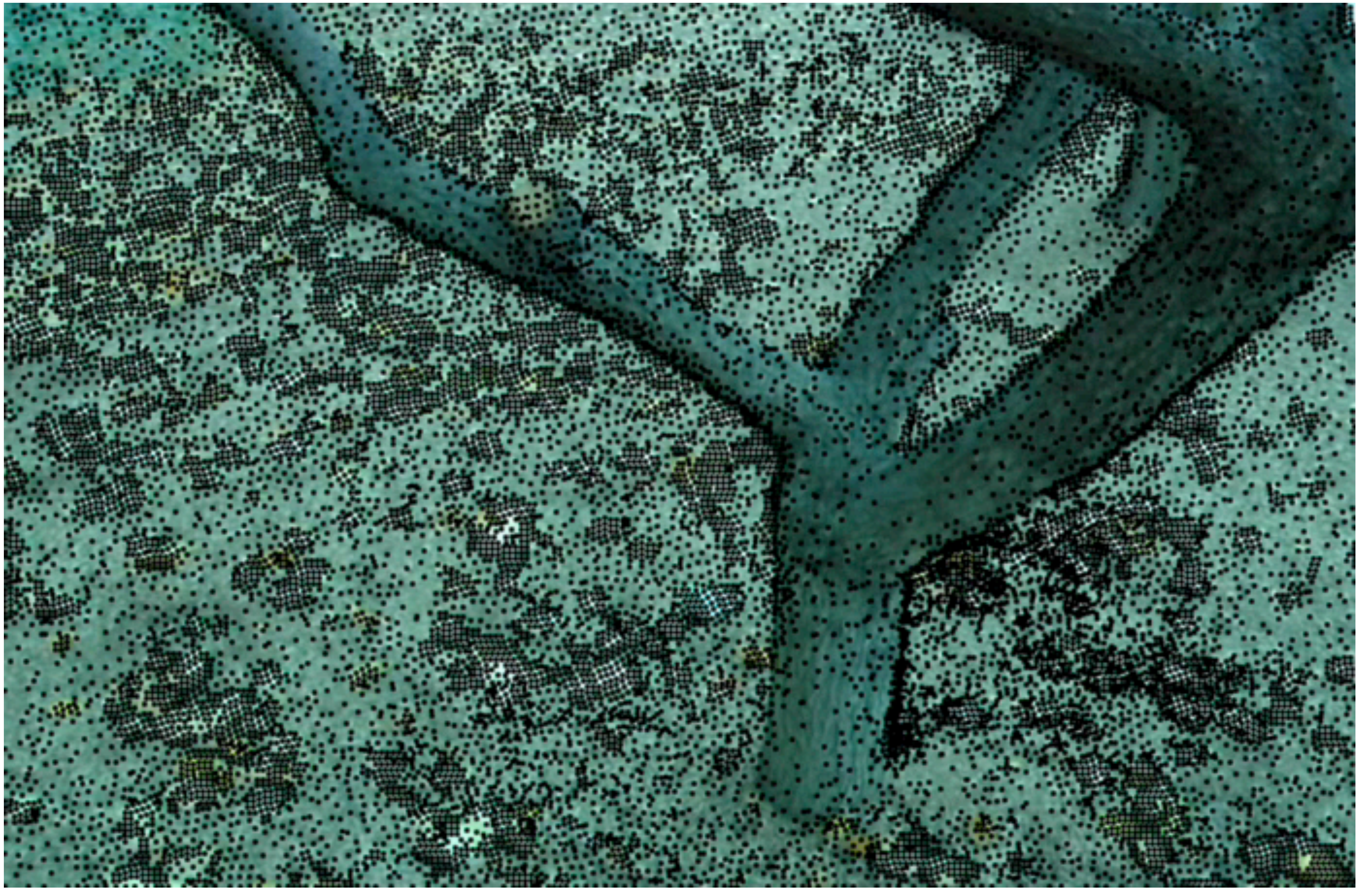
rank 4 ⟶ $$\mathbf{W} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1P} \\ y_{11} & y_{12} & \cdots & y_{1P} \\ x_{21} & x_{22} & \cdots & x_{2P} \\ y_{21} & y_{22} & \cdots & y_{2P} \\ \vdots & & & \vdots \\ x_{F1} & x_{F2} & \cdots & x_{FP} \\ y_{F1} & y_{F2} & \cdots & y_{FP} \end{bmatrix}$$

(rank 3 if you subtract mean)

2x*F* (*F* frames)

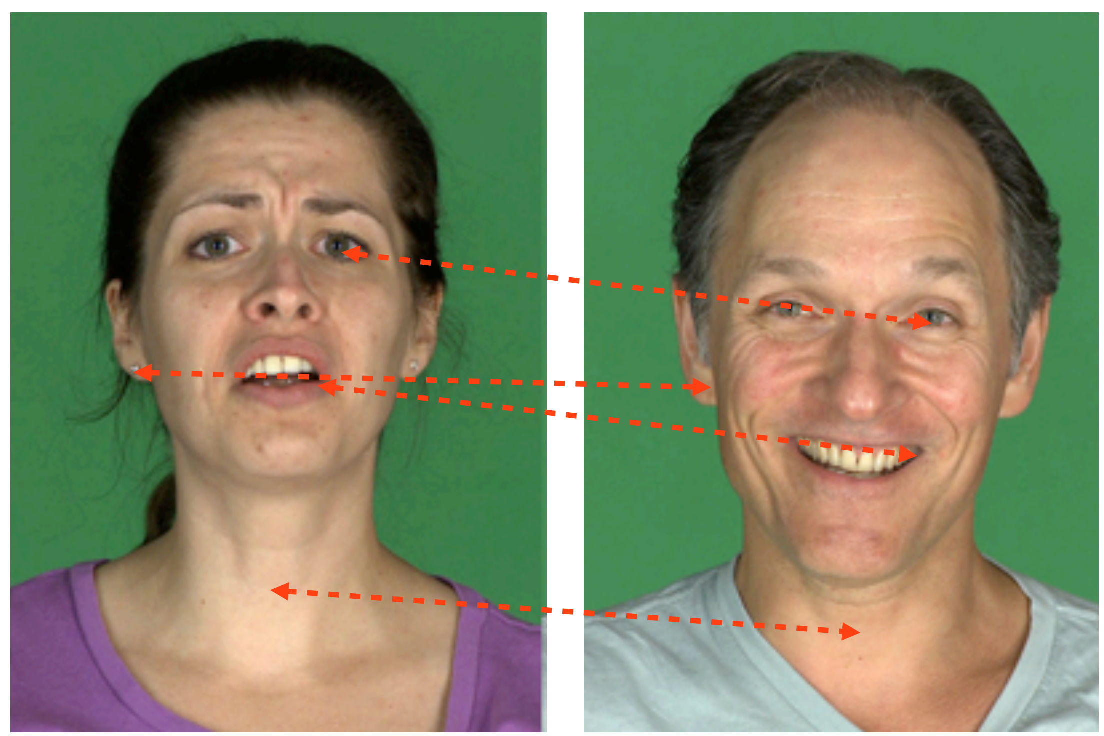
P (*P* points)

FEATURE TRACKING

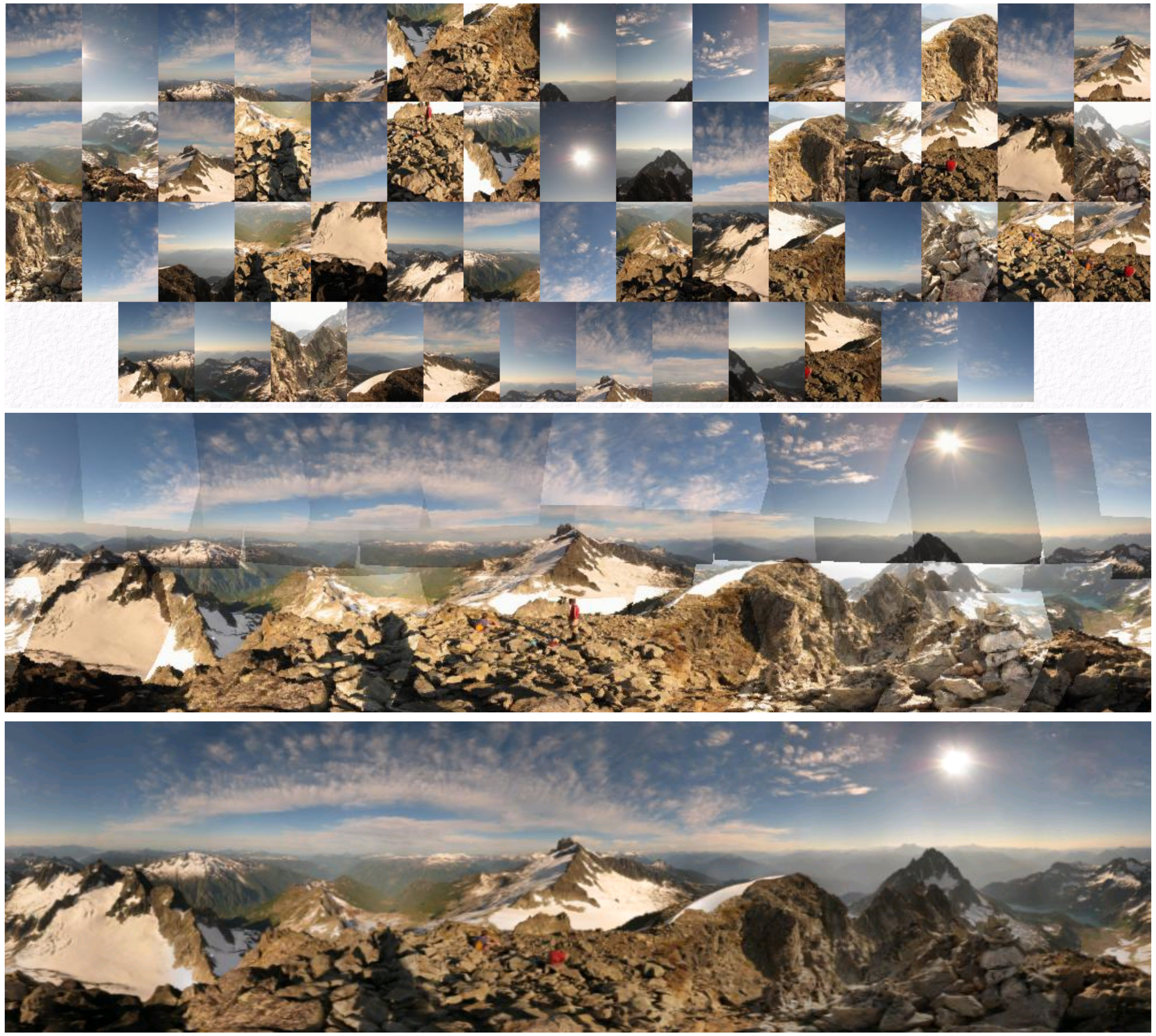


“THE THREE MOST IMPORTANT PROBLEMS IN COMPUTER VISION ARE:

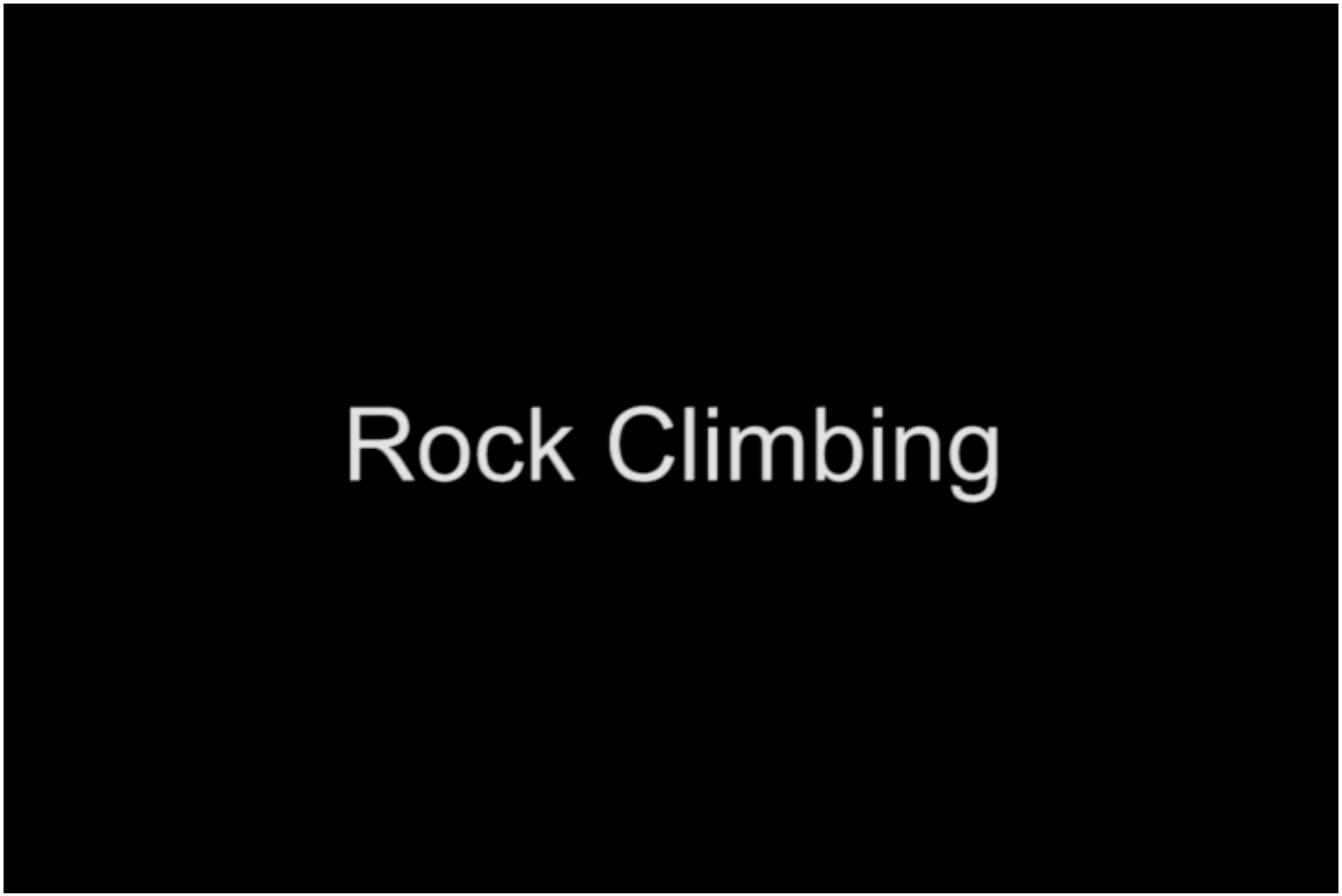
REGISTRATION/MATCHING/CORRESPONDENCE



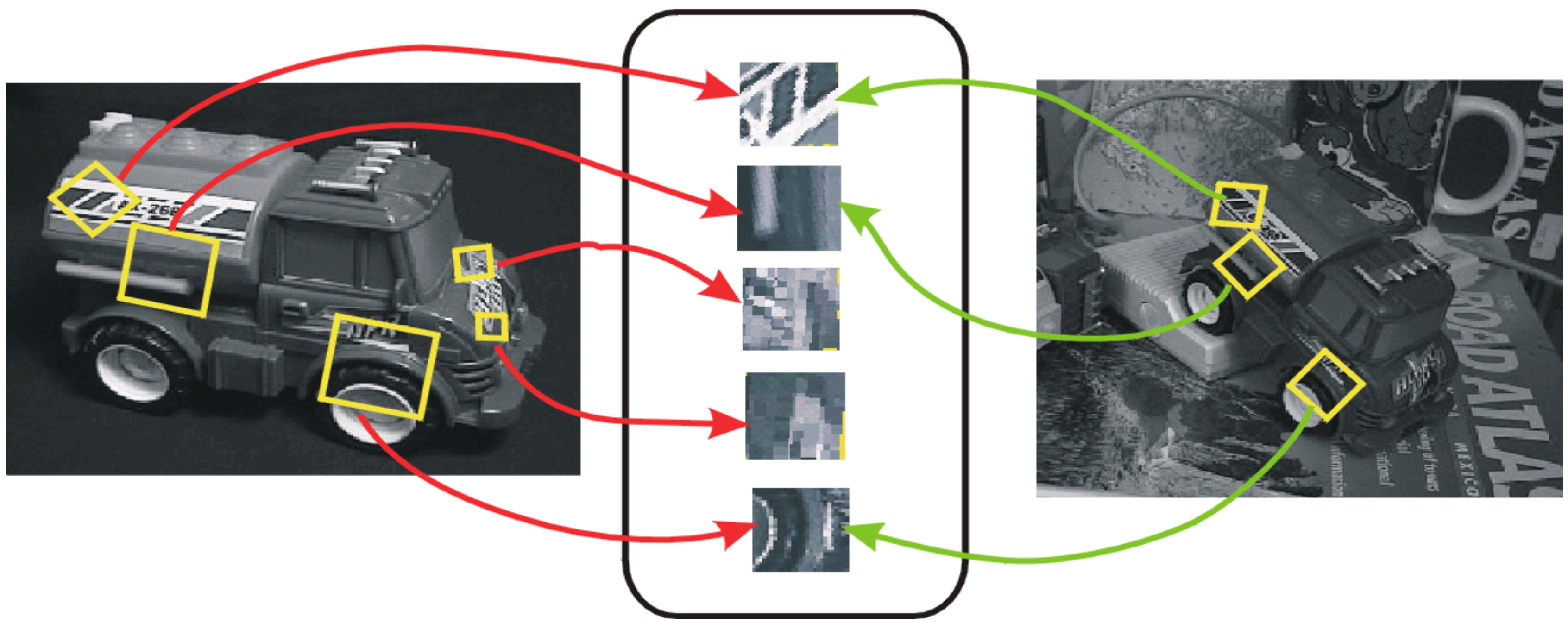
PANORAMAS



3D RECONSTRUCTION



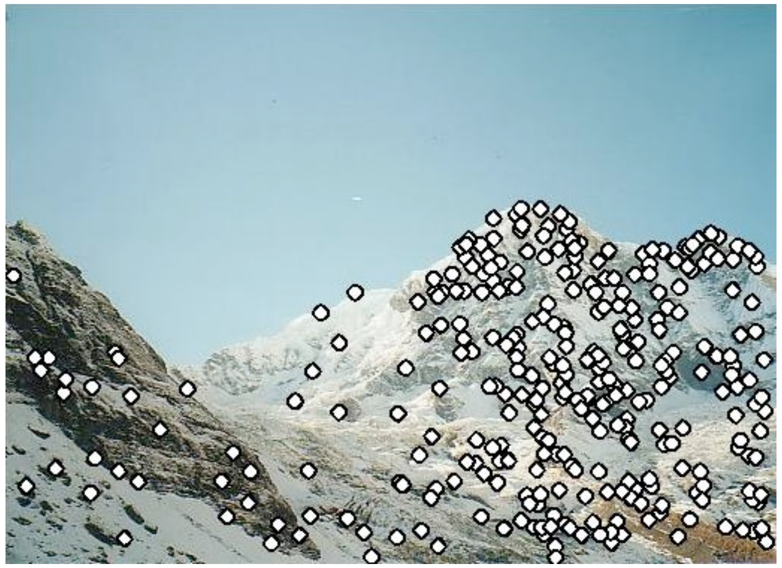
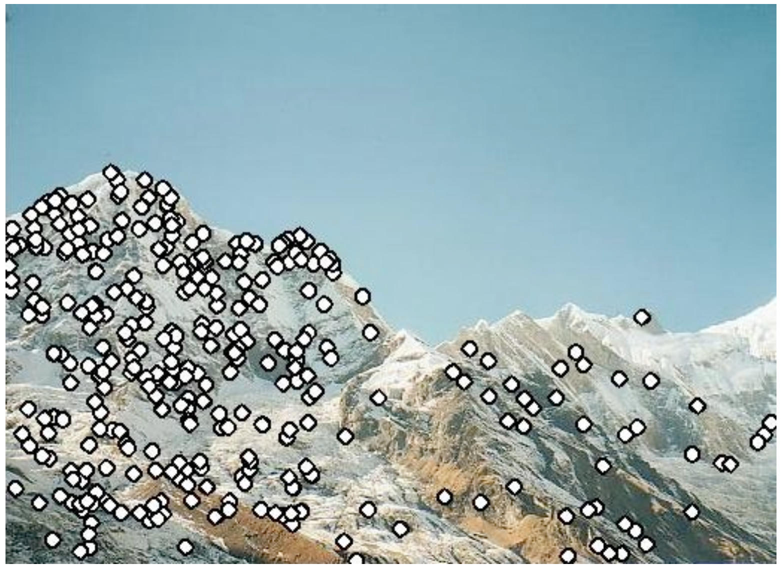
OBJECT RECOGNITION



REGISTRATION

IN THIS CLASS...

- **DETECTION**: IDENTIFY INTEREST POINT
- DESCRIPTION: EXTRACT FEATURE VECTORS FOR EACH POINT
- REGISTRATION: DETERMINE CORRESPONDENCE BETWEEN INTEREST POINTS

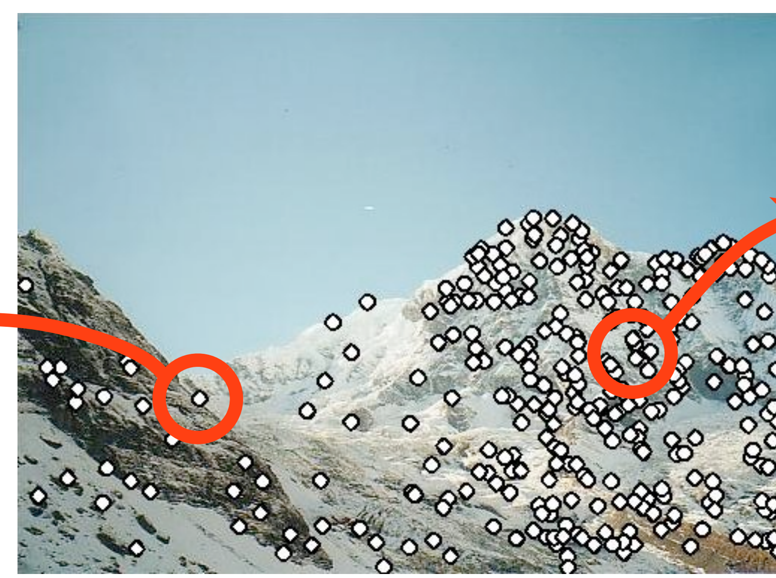



INTEREST POINTS
= SALIENT POINTS
= IMAGE FEATURES
= KEYPOINTS
= CORNER POINTS

LOCAL FEATURES

- DETECTION: IDENTIFY INTEREST POINT
- **DESCRIPTION**: EXTRACT FEATURE VECTORS FOR EACH POINT
- REGISTRATION: DETERMINE CORRESPONDENCE BETWEEN INTEREST POINTS

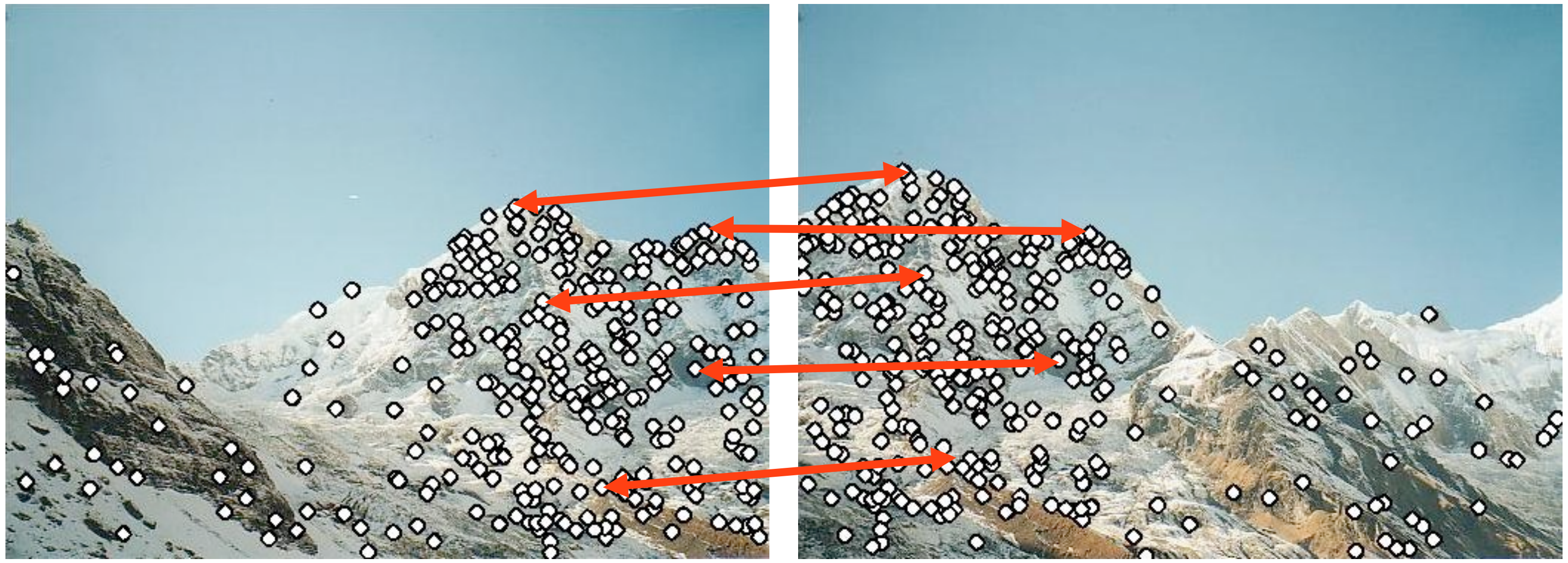
$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{12} \\ \vdots \\ \mathbf{x}_{1N} \end{bmatrix}$$



$$\mathbf{x}_2 = \begin{bmatrix} \mathbf{x}_{21} \\ \mathbf{x}_{22} \\ \vdots \\ \mathbf{x}_{2N} \end{bmatrix}$$

LOCAL FEATURES

- DETECTION: IDENTIFY INTEREST POINT
- DESCRIPTION: EXTRACT FEATURE VECTORS FOR EACH POINT
- **REGISTRATION**: DETERMINE CORRESPONDENCE BETWEEN INTEREST POINTS

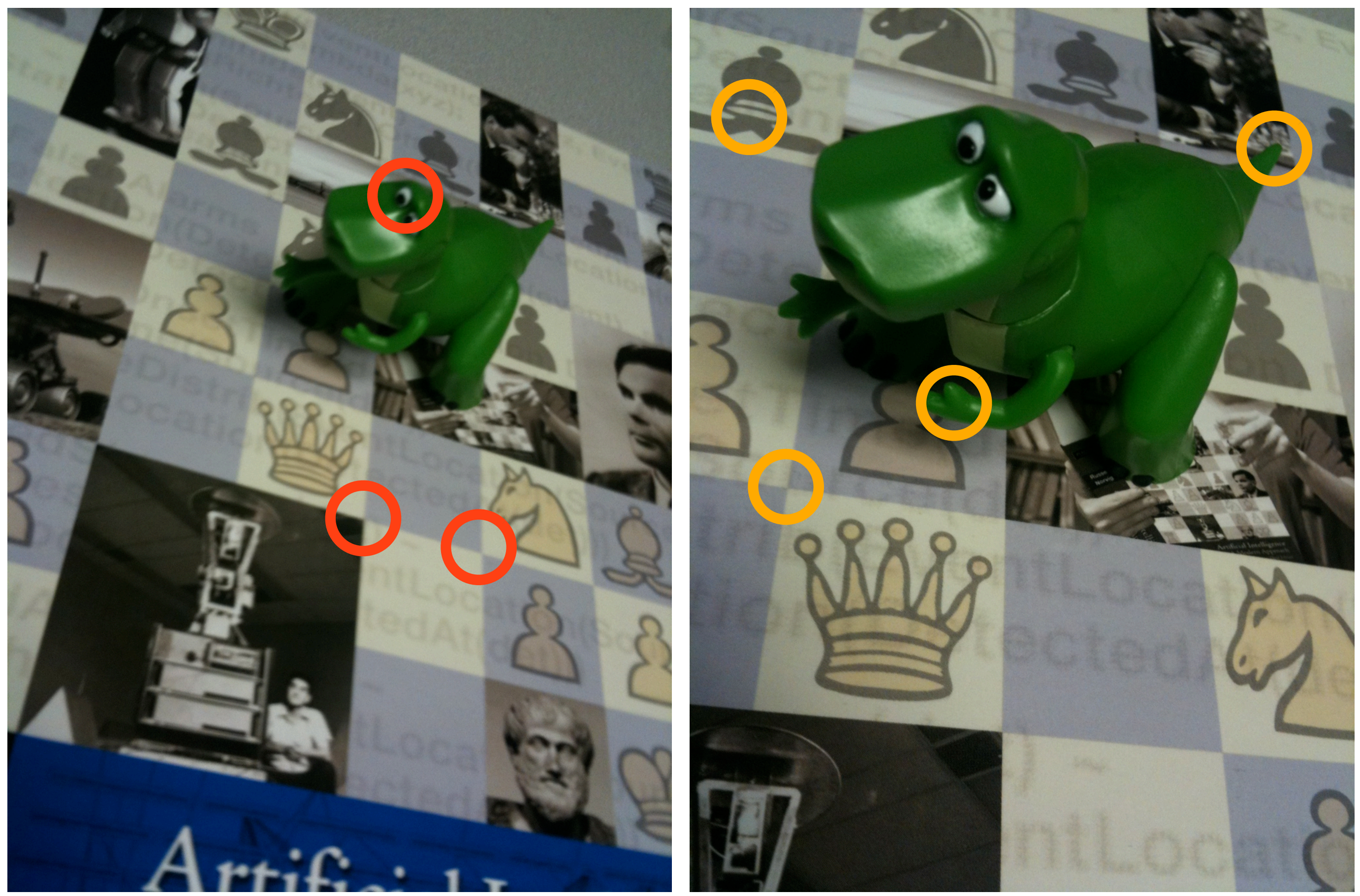


HOW DO WE EVALUATE GOOD INTEREST POINTS?

DESIRABLE PROPERTIES

- **REPEATABILITY**: INVARIANCE TO GEOMETRIC AND RADIOMETRIC TRANSFORMATIONS
- SALIENCY
- COMPRESSION
- LOCALITY

REPEATABILITY



If two sets are disjoint, matching is impossible

DESIRABLE PROPERTIES

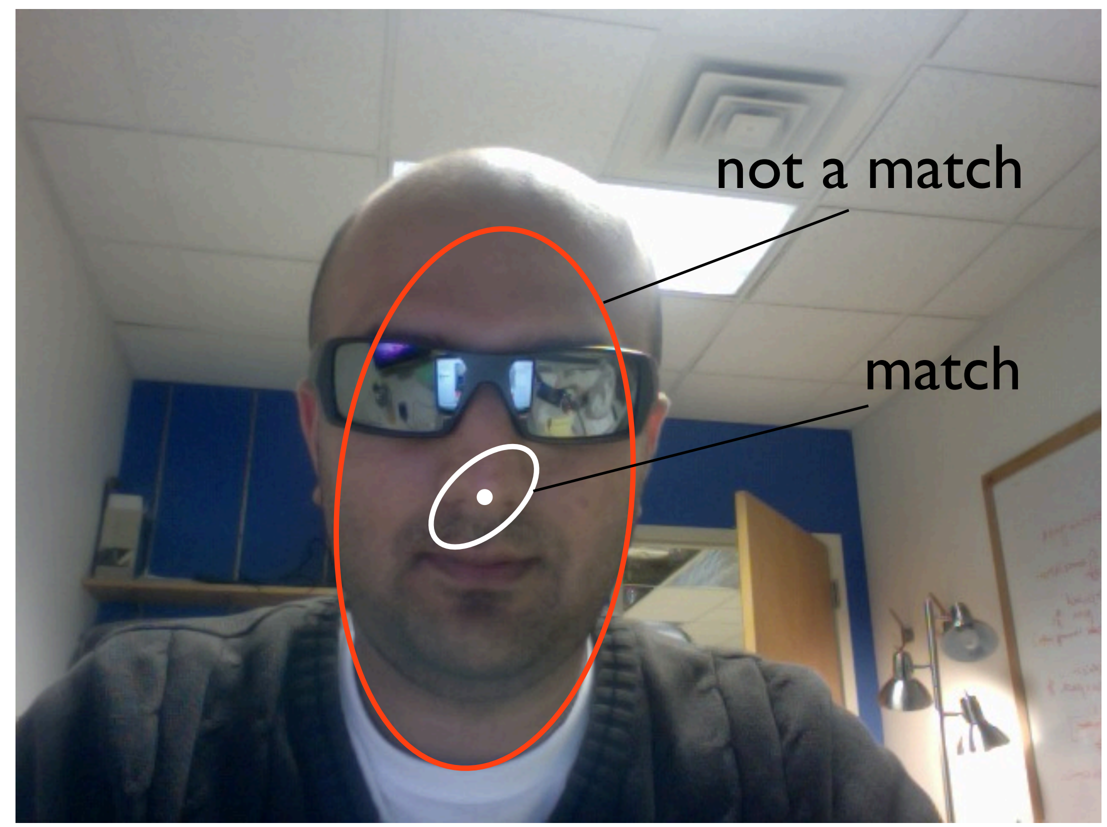
- REPEATABILITY
- **SALIENCY:** EACH POINT IS DISTINCTIVE
- COMPRESSION
- LOCALITY

SALIENCY

DESIRABLE PROPERTIES

- REPEATABILITY
- SALIENCY:
- COMPACTNESS
- **LOCALITY:** FEATURE ‘SUPPORT’ IS LOCAL; REMAINS ROBUST TO OCCLUSION

LOCALITY

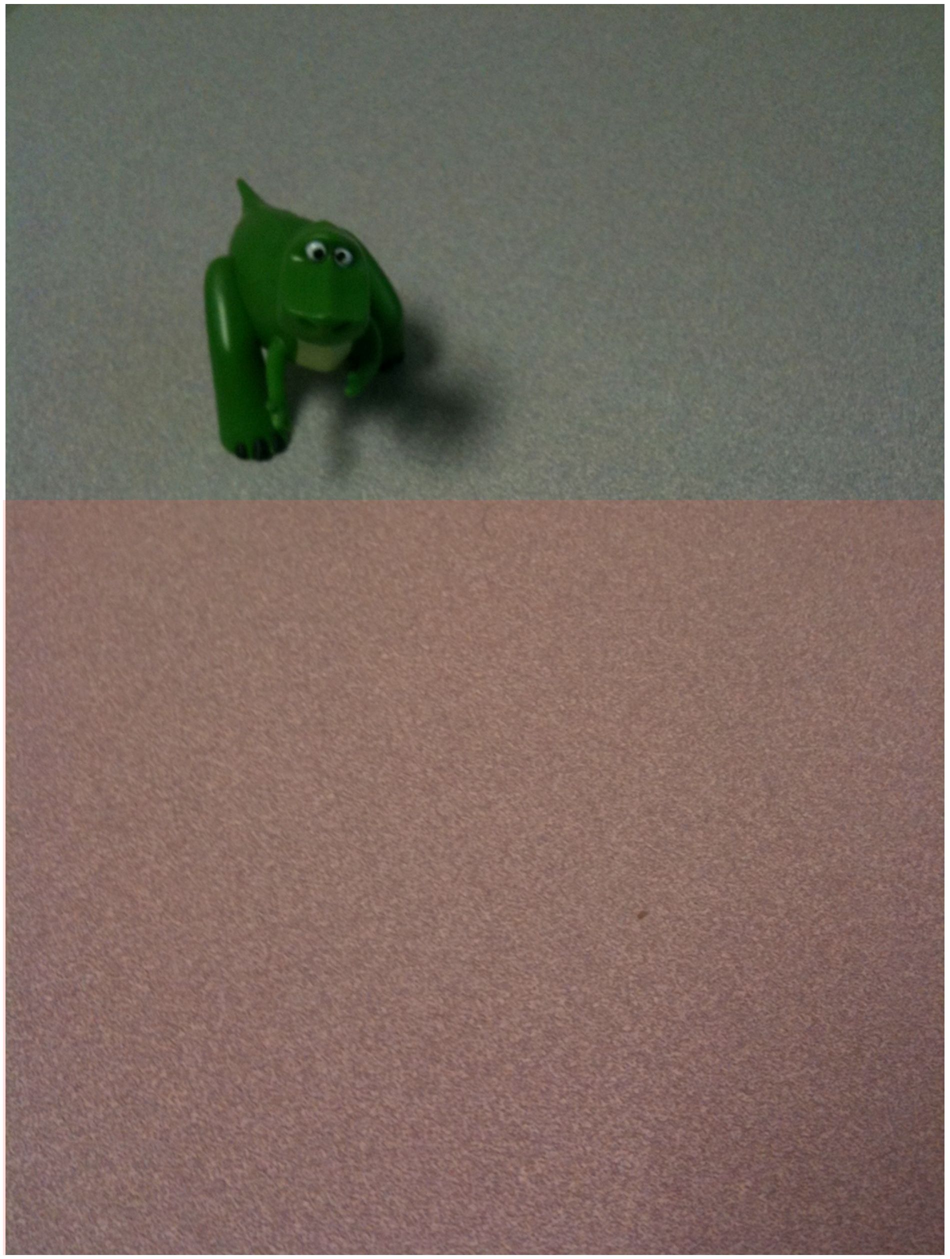


trade-off between saliency and locality

DESIRABLE PROPERTIES

- REPEATABILITY
- SALIENCY
- **COMPRESSION:** THERE SHOULD BE FEWER FEATURES THAN PIXELS
- LOCALITY

COMPRESSION



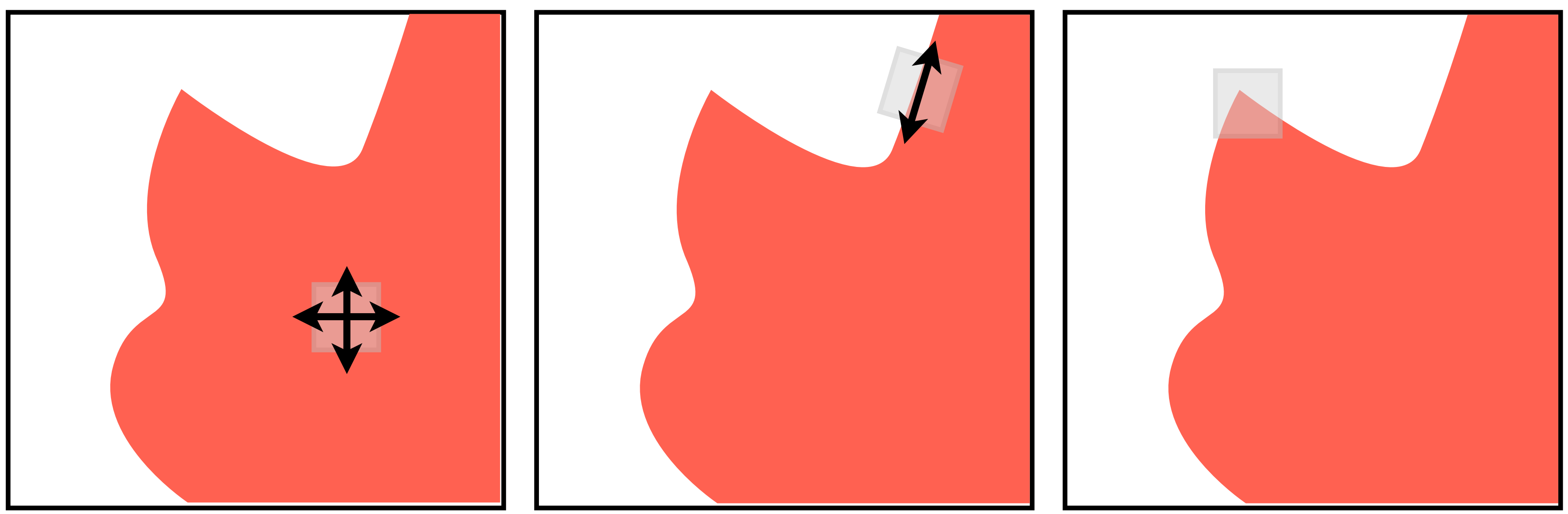
DESIRABLE PROPERTIES

- REPEATABILITY
- SALIENCY
- COMPRESSION: THERE SHOULD BE FEWER FEATURES THAN PIXELS
- LOCALITY
- **EFFICIENCY**

HOW SHOULD WE SELECT INTEREST POINTS?

CHANGE WE CAN BELIEVE IN

- **IDEA**: SHIFTING PATCH IN ANY DIRECTION SHOULD PRODUCE LARGE CHANGE



Flat areas
No change in
all directions

Edges
No change
along edge

Corners
Change in
all direction

CONVOLUTION

Kernel (e.g. Gaussian)

image

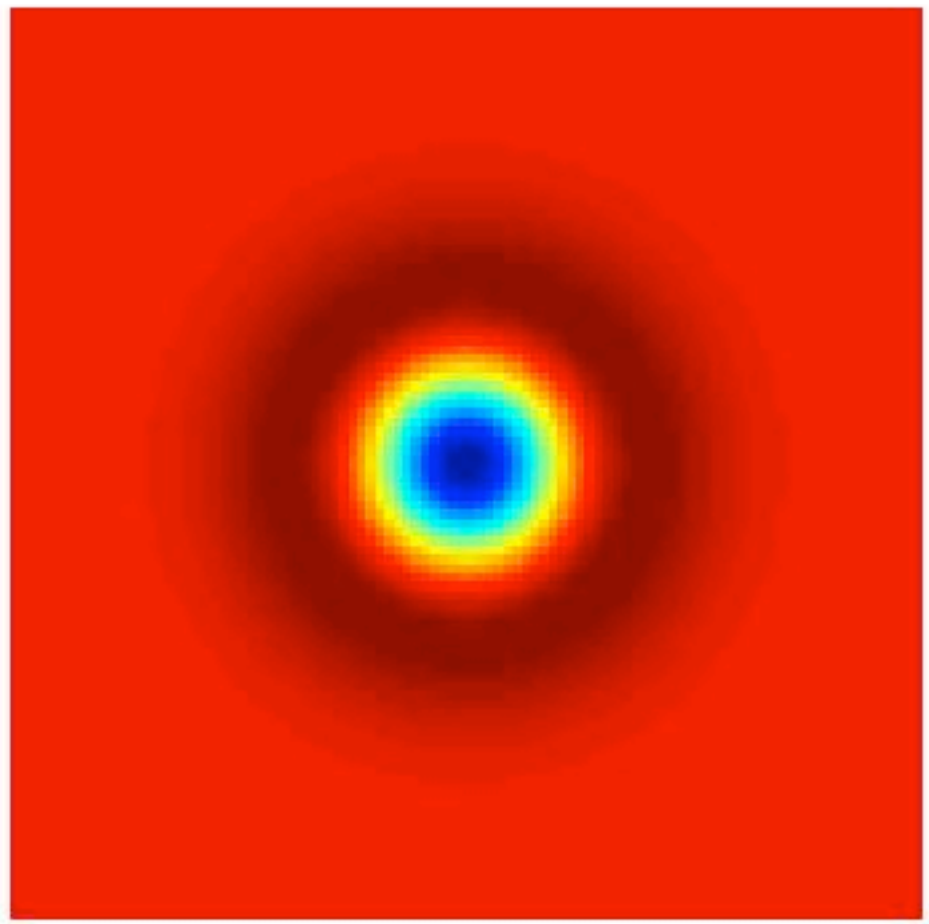
$$L(x, y, \sigma) = G(x, y, \sigma) * I(x, y)$$

convolution operator

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$$

LAPLACIAN OF GAUSSIAN

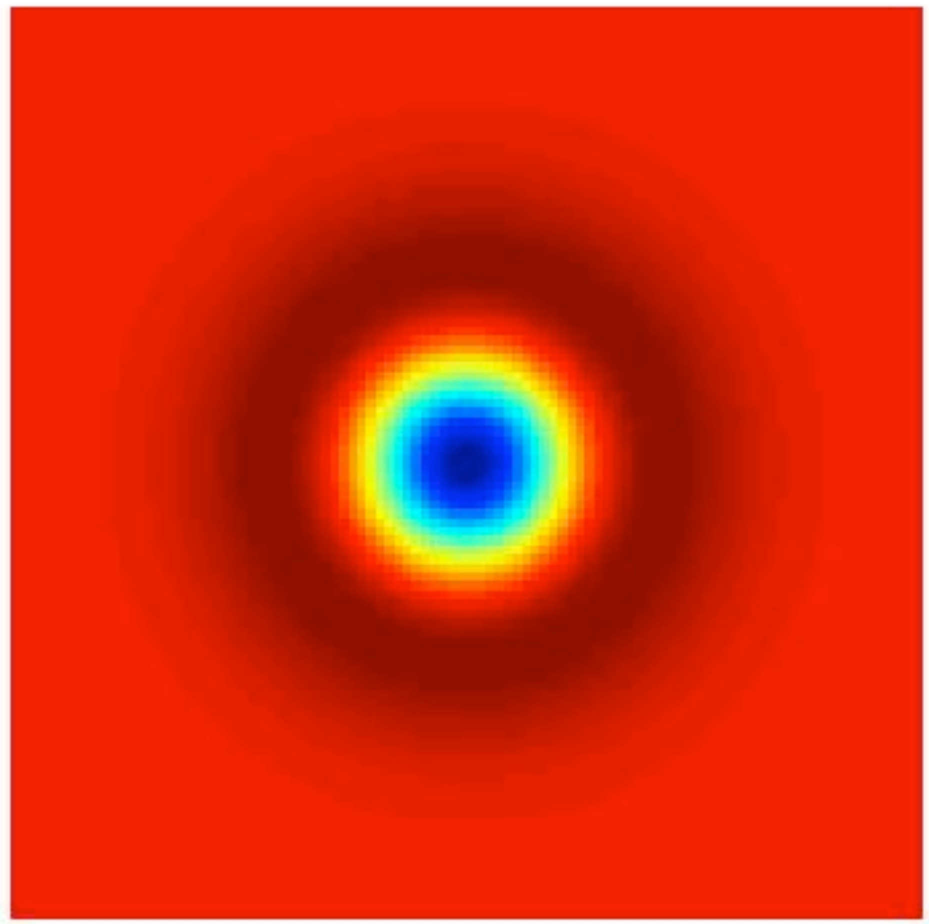
MARR-HILDRETH OPERATOR



$$\nabla^2 f(x, y) = \frac{\delta^2 f}{\delta x^2} + \frac{\delta^2 f}{\delta y^2}$$

IDEA: SHIFTING PATCH IN ANY DIRECTION SHOULD PRODUCE LARGE CHANGE?

DIFFERENCE OF GAUSSIAN



$$D(x, y, \sigma) = (G(x, y, k\sigma) - G(x, y, \sigma)) * I(x, y)$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

FAST APPROXIMATION TO THE LAPLACIAN OF GAUSSIAN

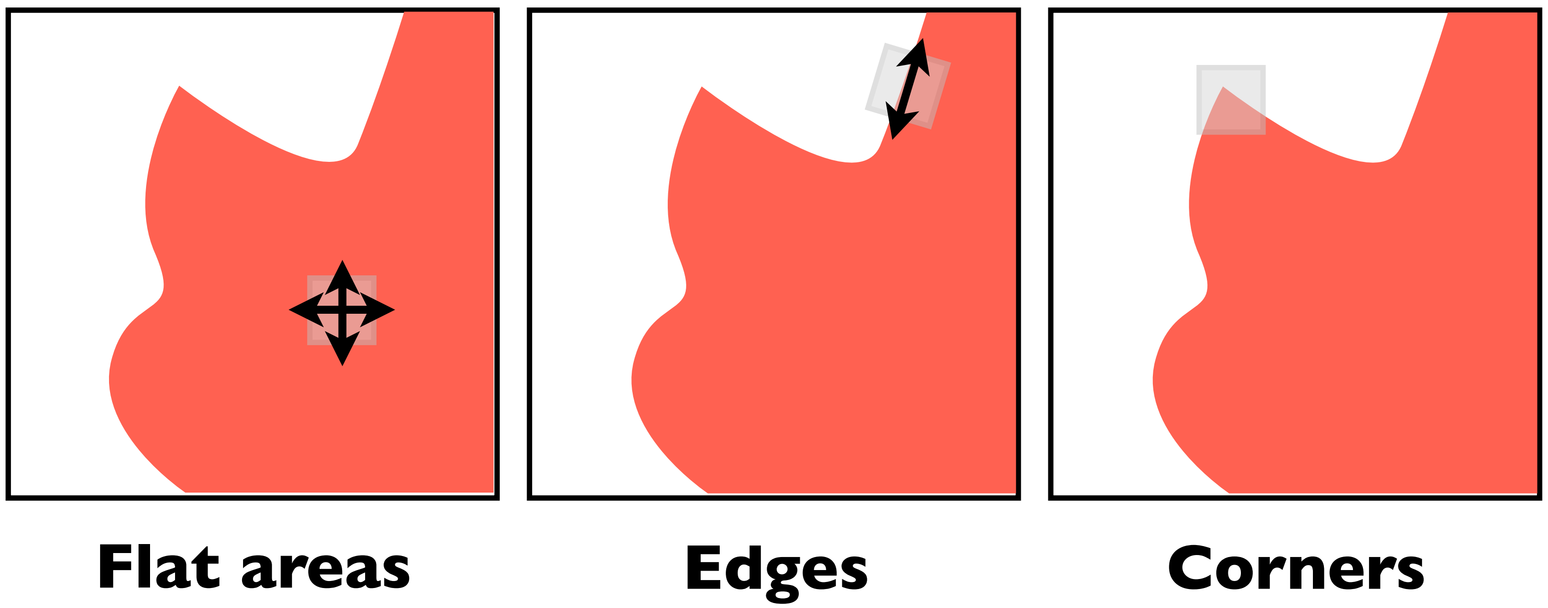
EDGE ELIMINATION

Change in *D*
(computer by taking adjacent pixel differences)

Hessian Matrix

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

The eigenvalues of H are proportional to the principal curvature of D



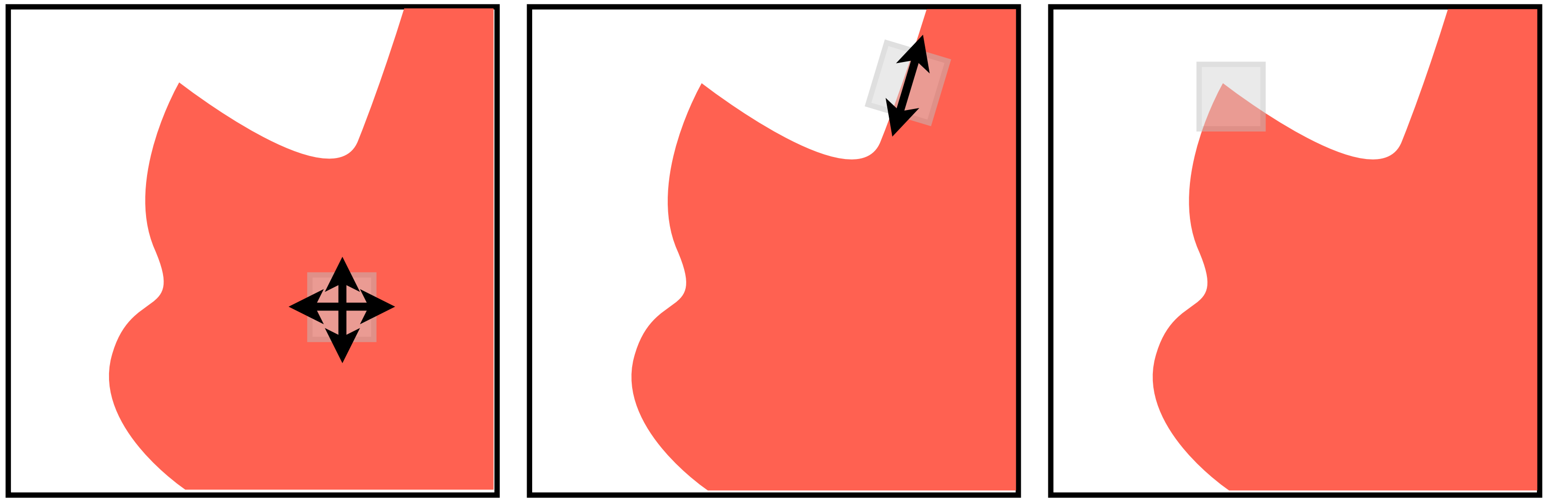
CORNERNESS

HARRIS-STEPHENS CORNER

Hessian Matrix

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

Eigen values of **H**: α, β



two small eigenvalues

one large one small eigenvalues

two large eigenvalues

CORNERNESS

HARRIS-STEPHENS CORNER

Hessian Matrix

$$\mathbf{H} = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix}$$

Eigen values of **H**: α, β

$$\mathrm{Tr}(\mathbf{H}) = D_{xx} + D_{yy} = \alpha + \beta$$

$$\mathrm{Det}(\mathbf{H}) = D_{xx}D_{yy} - D_{xy}^2 = \alpha\beta$$

$$M_c = \alpha\beta - \kappa(\alpha + \beta)^2 = \det(\mathbf{H}) - \kappa \mathrm{trace}^2(\mathbf{H})$$

$$M_c > \mathrm{threshold}$$

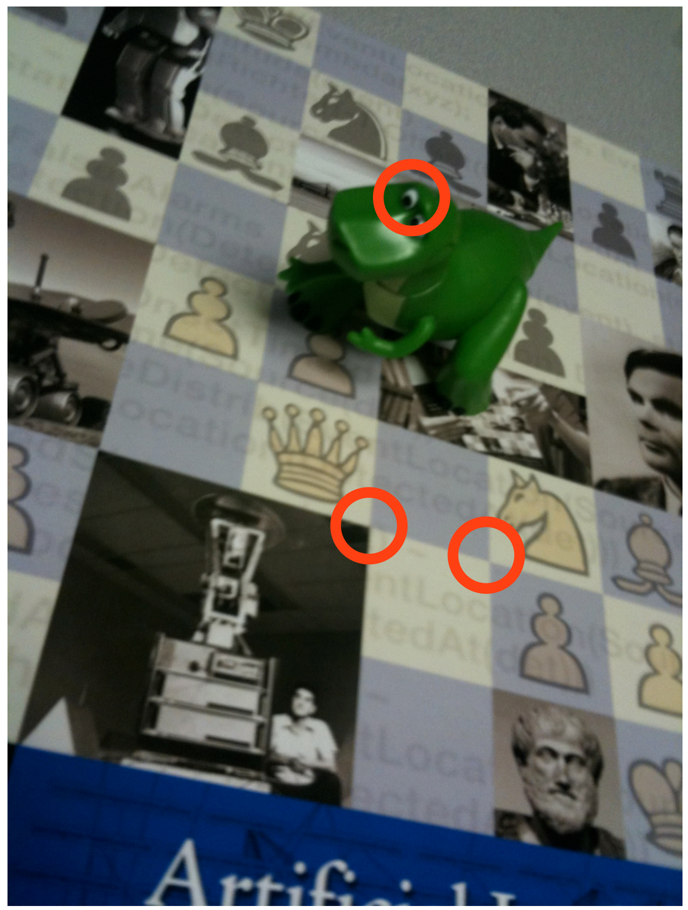
HARRIS CORNERS

- ESTIMATE DIFFERENCE OF GAUSSIAN AT APPROPRIATE SCALE
- CONSTRUCT HESSIAN MATRIX FOR EACH PIXEL
- ESTIMATE M FOR EACH PIXEL
- PERFORM NON-MAXIMA SUPPRESSION FOR EACH PIXEL

HARRIS CORNERS

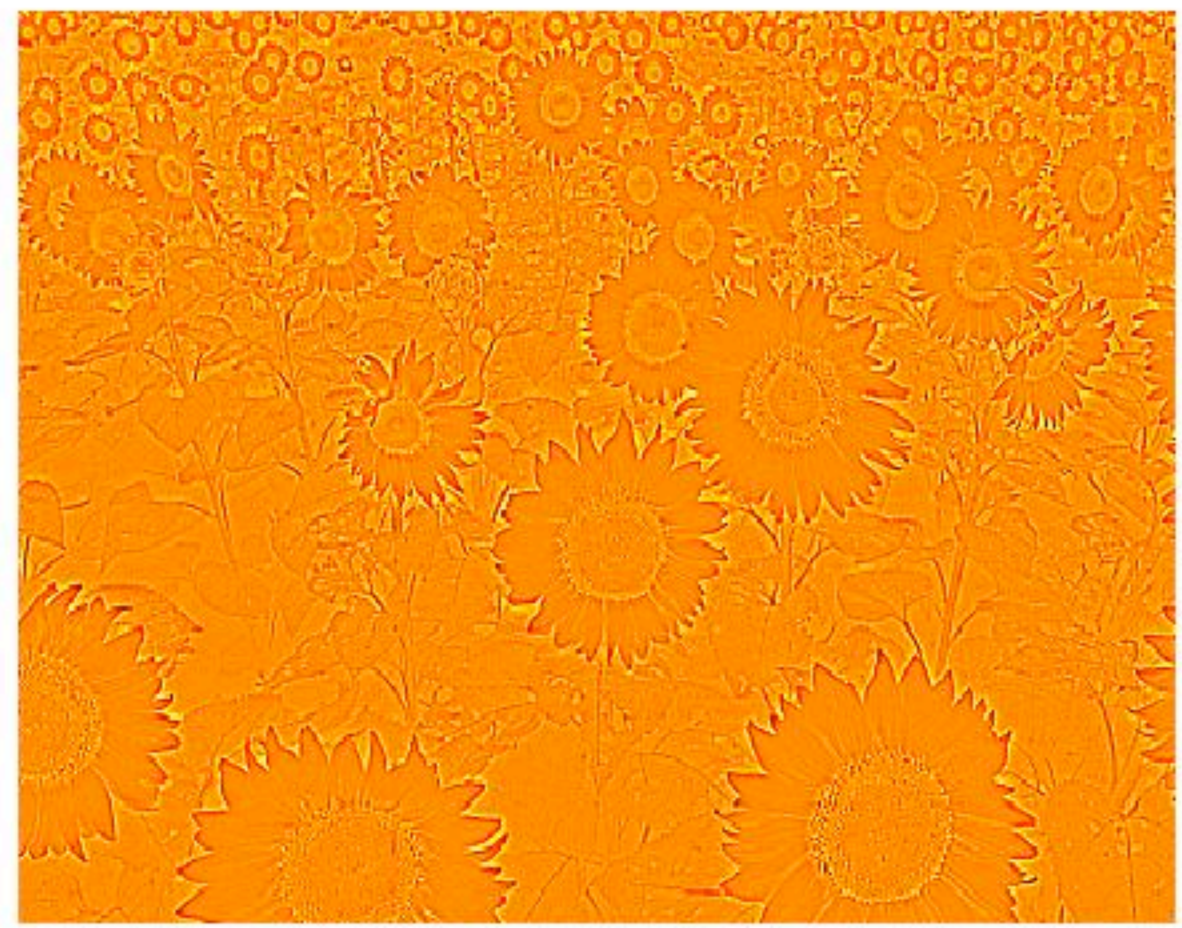
- TRANSLATION INVARIANT
- ROTATION INVARIANT
- ROBUST TO ILLUMINATION CHANGES

HOW DO WE SELECT INTEREST POINTS THAT ARE REPEATABLE OVER DIFFERENT SCALES?

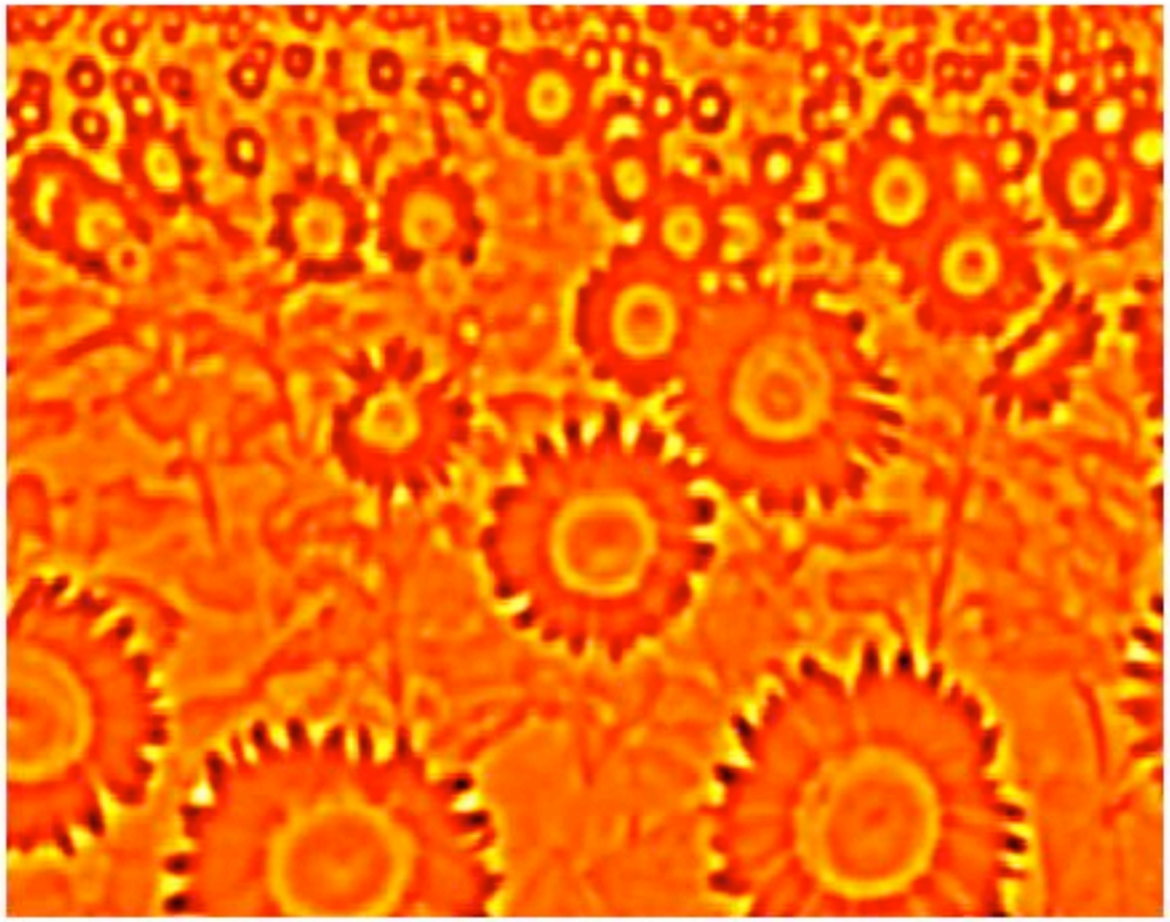


SCALE SPACE

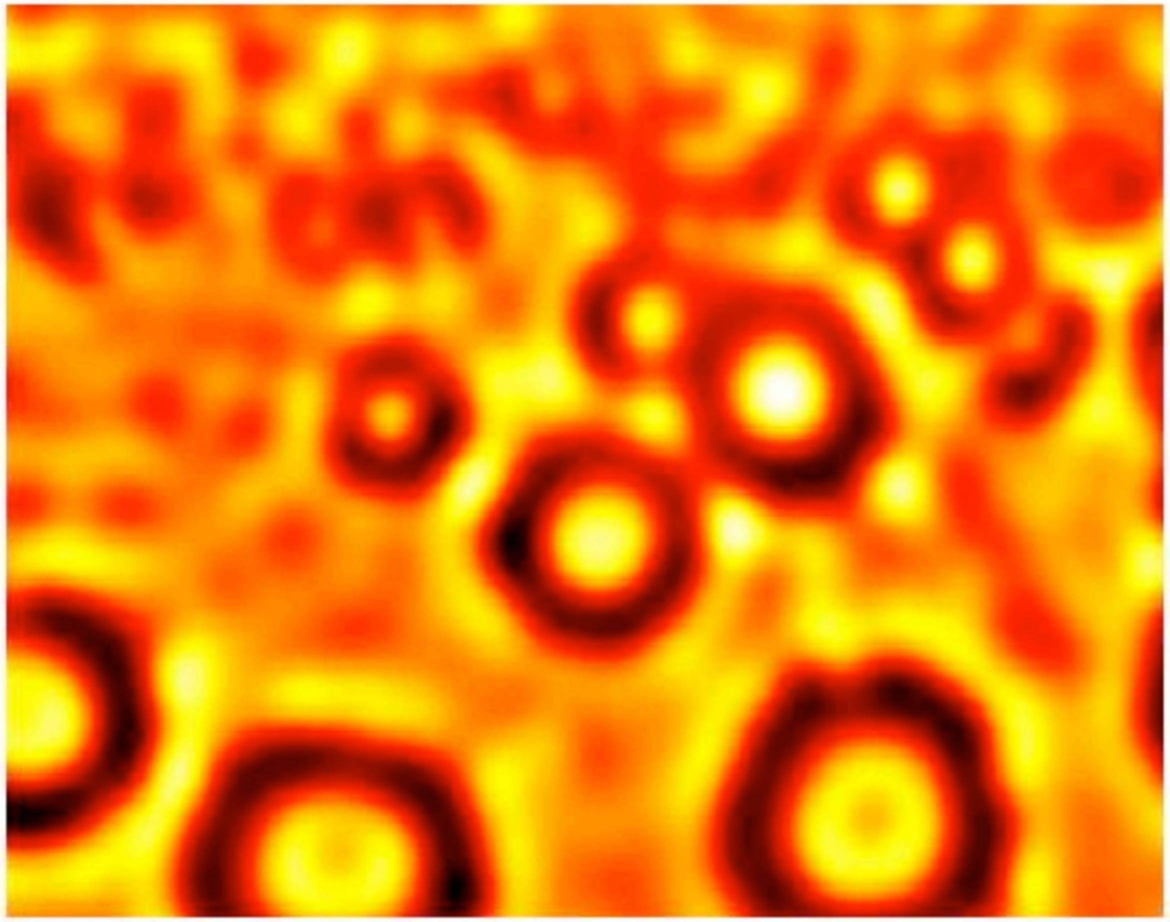
SCALE SPACE



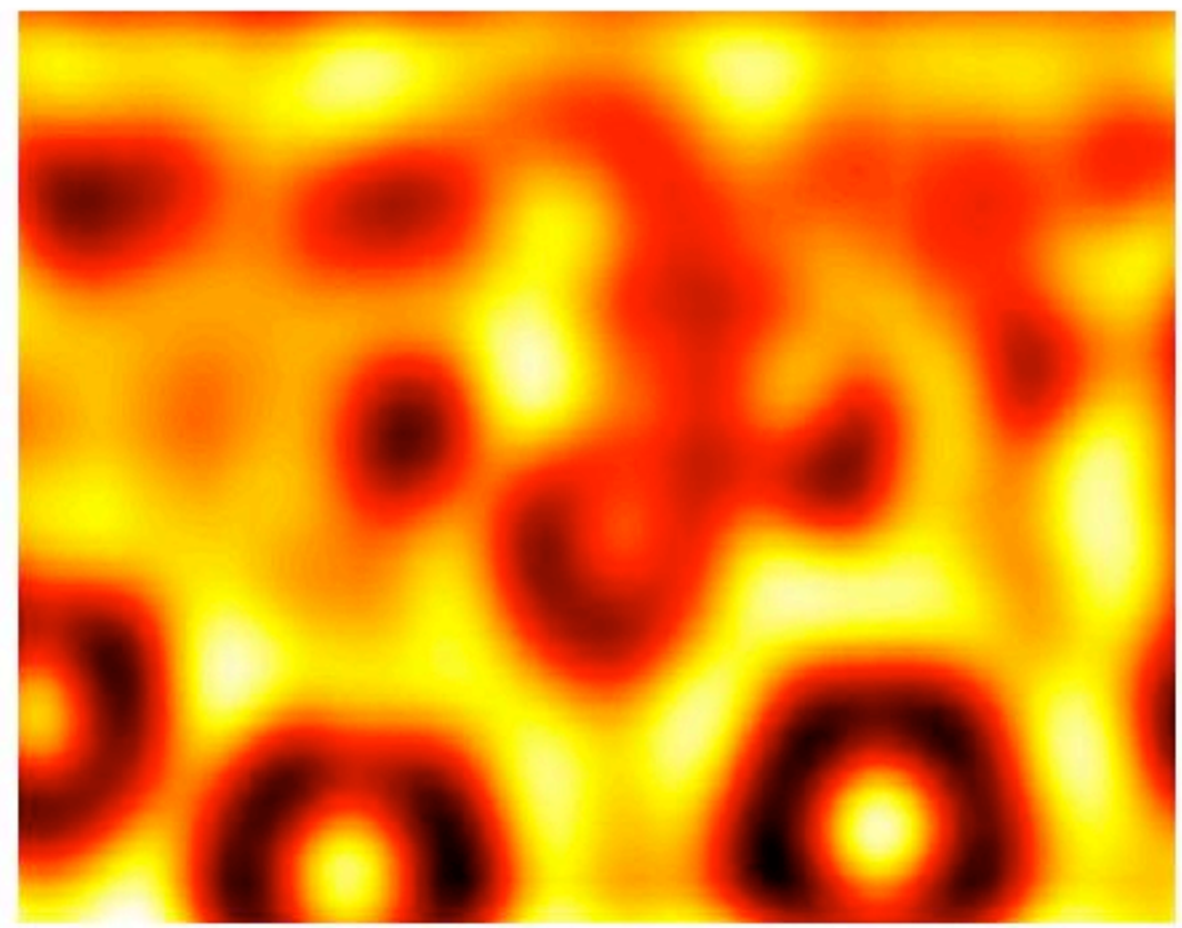
SCALE SPACE



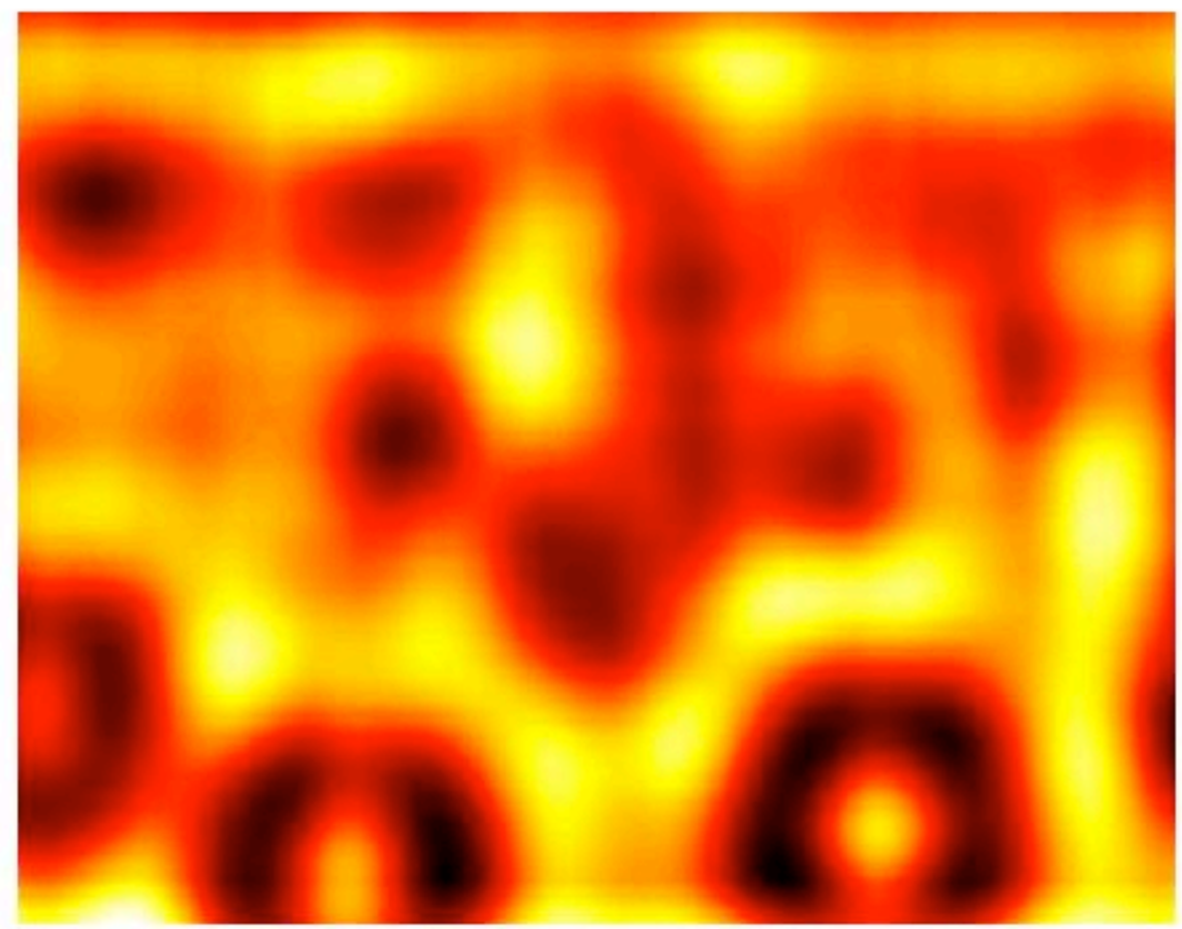
SCALE SPACE



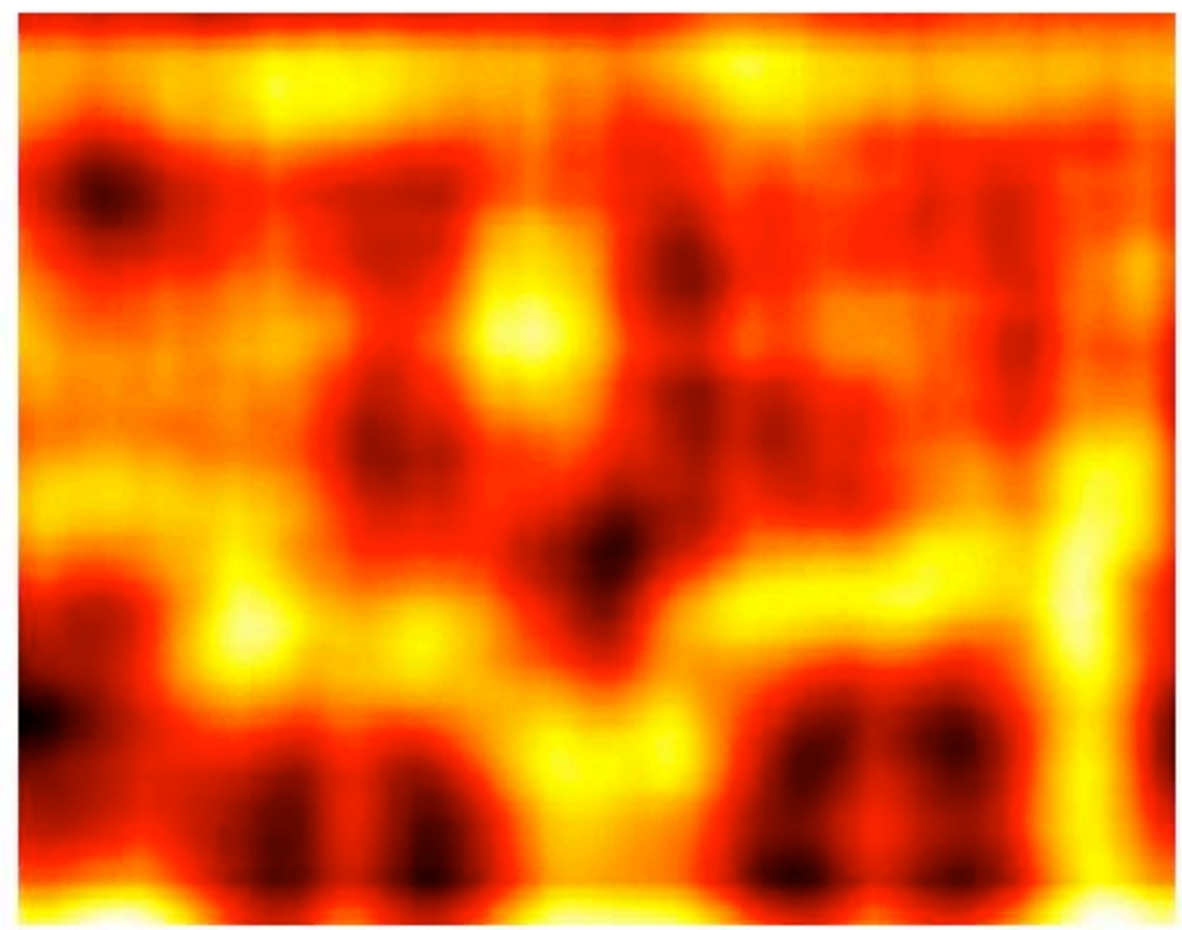
SCALE SPACE



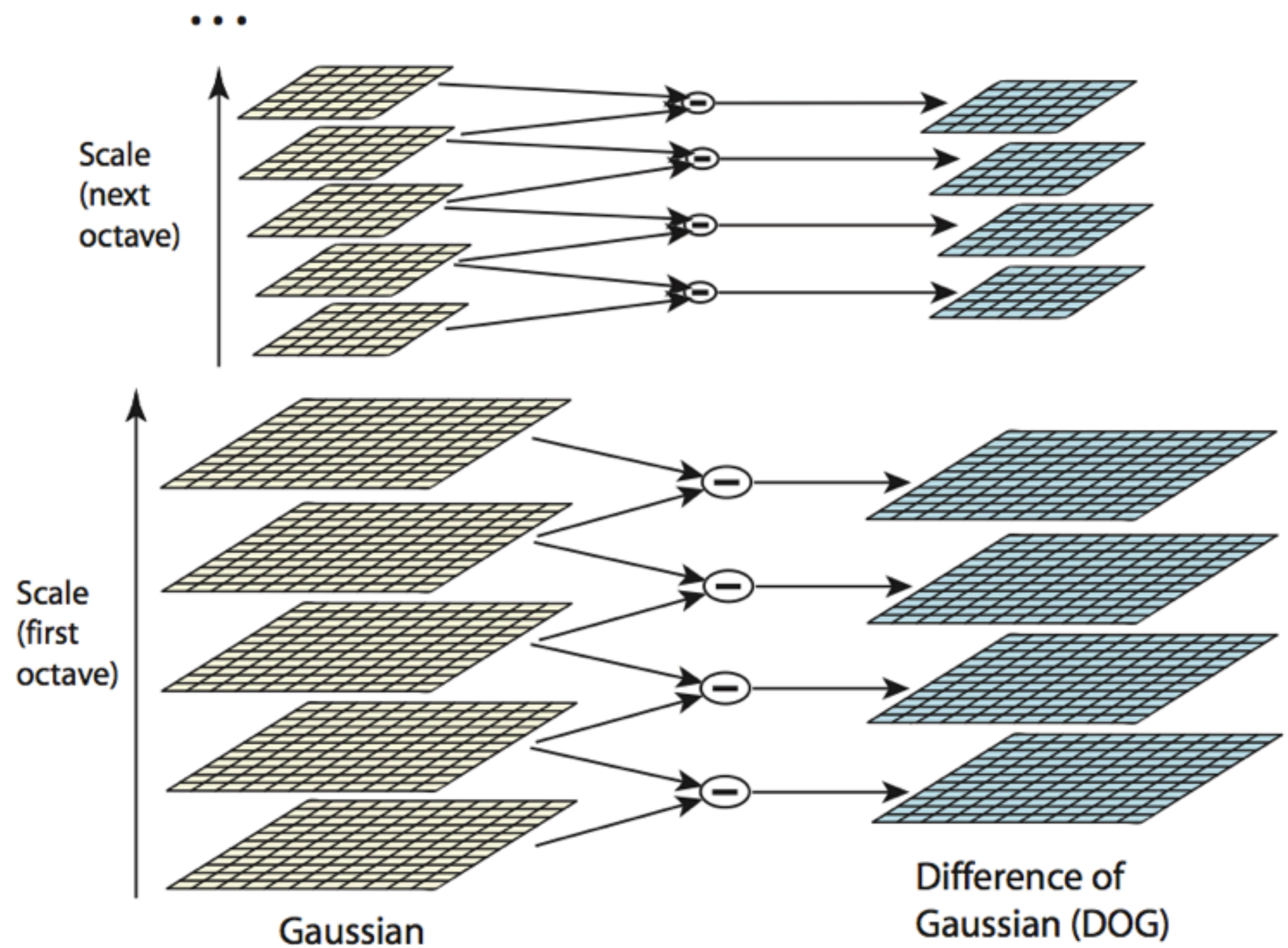
SCALE SPACE



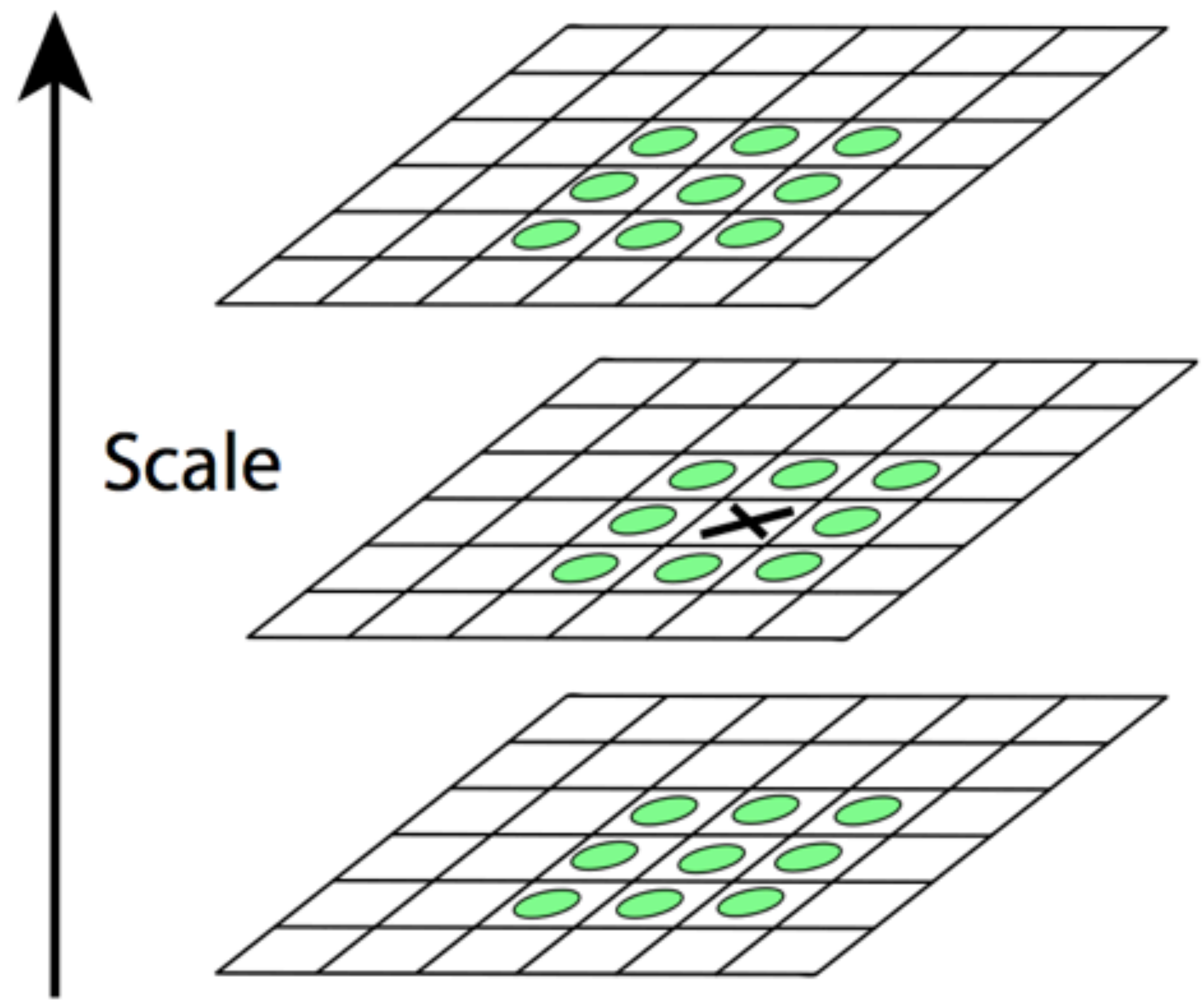
SCALE SPACE



SCALE SPACE


...
Scale
(next
octave)
Scale
(first
octave)
Gaussian
Difference of
Gaussian (DOG)

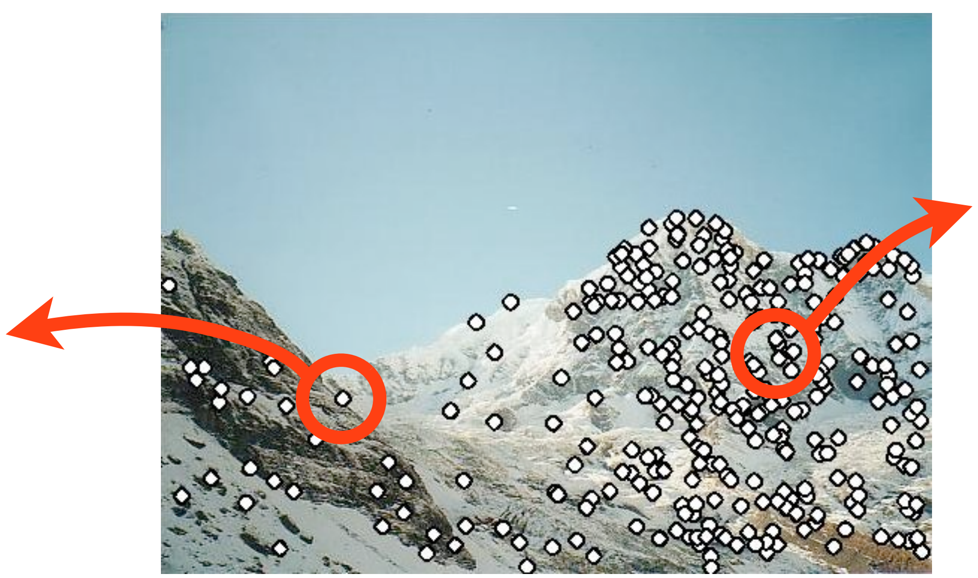
LOCAL EXTREMA DETECTION



LOCAL FEATURES

- DETECTION: IDENTIFY INTEREST POINT
- **DESCRIPTION**: EXTRACT FEATURE VECTORS FOR EACH POINT
- REGISTRATION: DETERMINE CORRESPONDENCE BETWEEN INTEREST POINTS

$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{12} \\ \vdots \\ \mathbf{x}_{1N} \end{bmatrix}$$

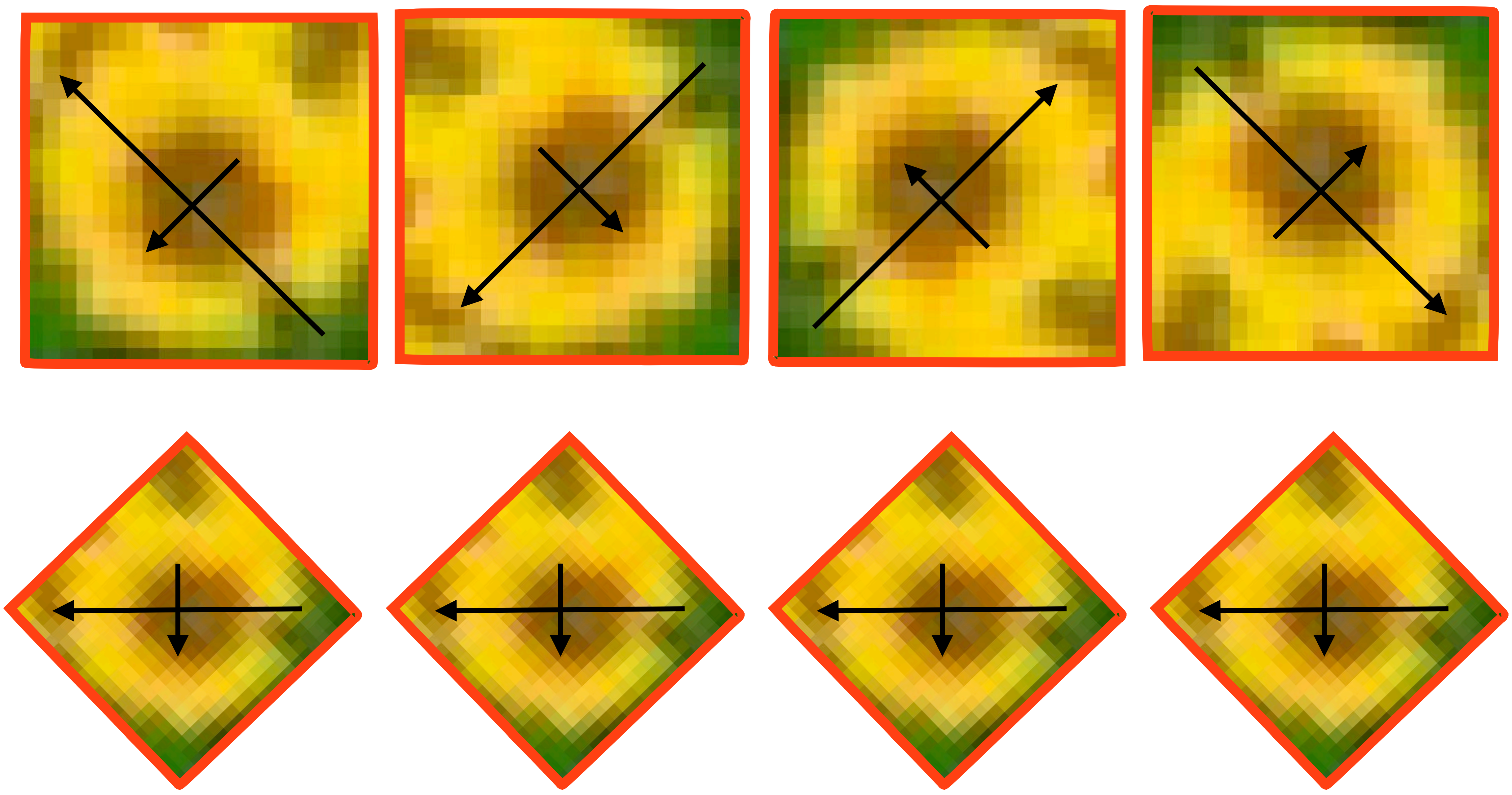


$$\mathbf{x}_2 = \begin{bmatrix} \mathbf{x}_{21} \\ \mathbf{x}_{22} \\ \vdots \\ \mathbf{x}_{2N} \end{bmatrix}$$

VECTORIZATION

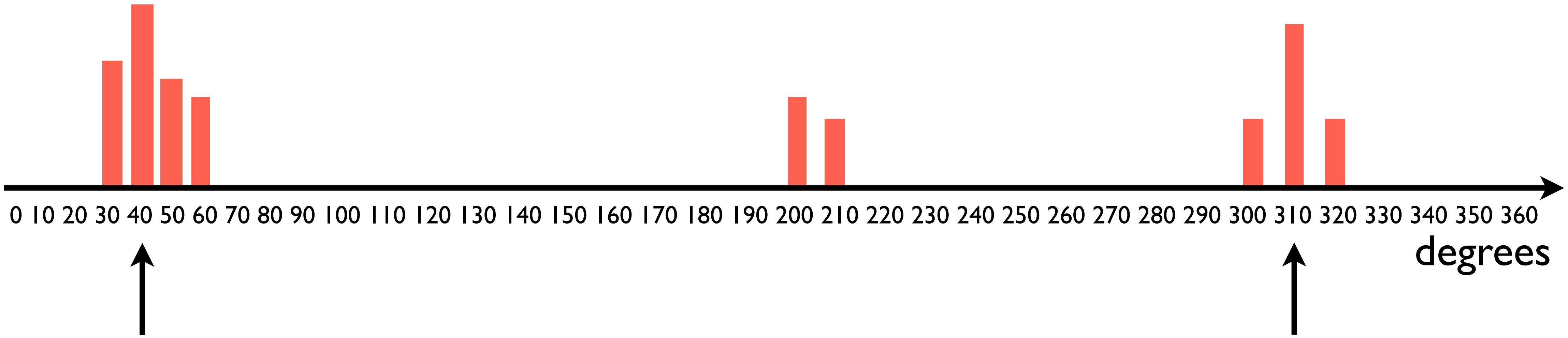
Invariance?

HOW DO WE ACHIEVE ROTATION INVARIANCE?

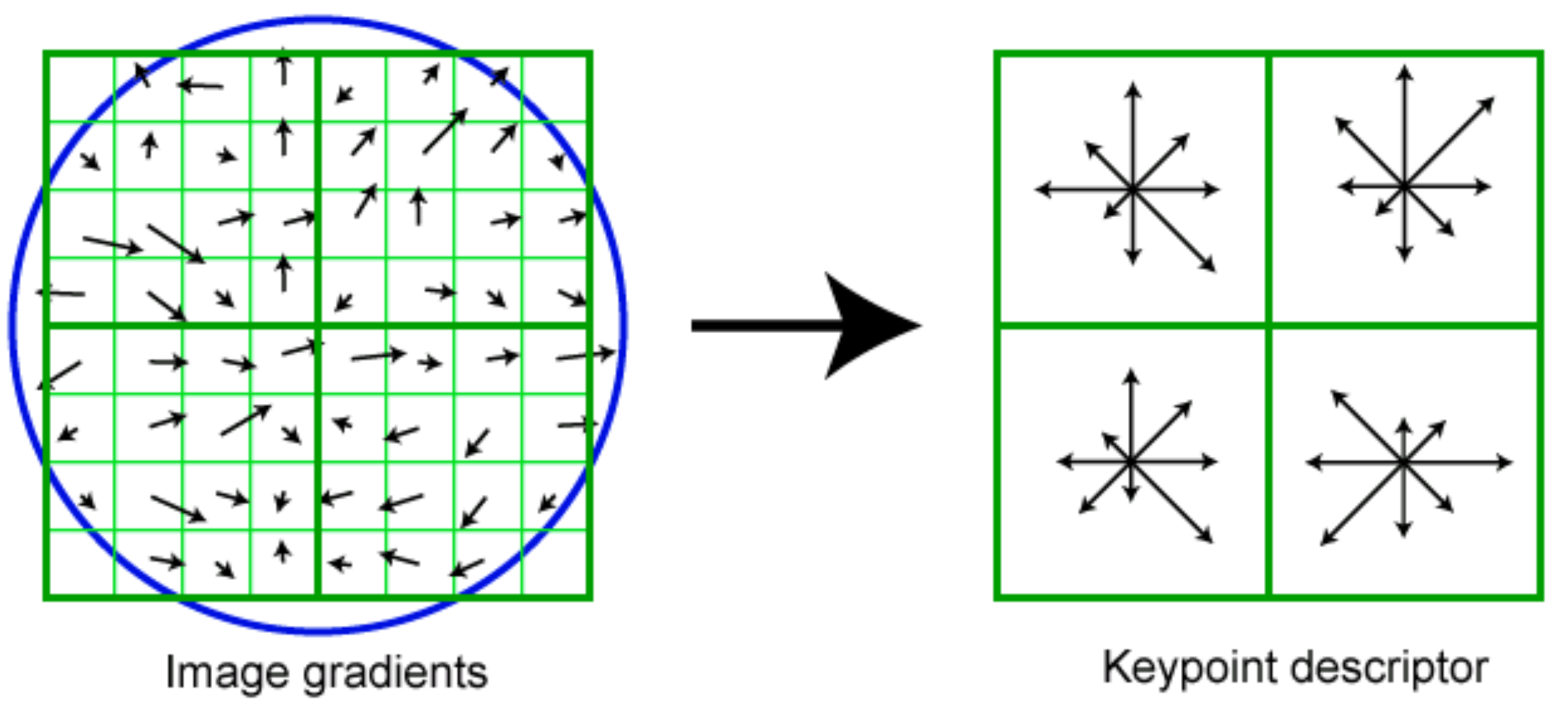


ORIENTATION ASSIGNMENT

- COMPUTE **GRADIENT MAGNITUDE** AND **GRADIENT DIRECTION**
- CREATE AN **ORIENTATION HISTOGRAM**
- SELECT PEAK ORIENTATION
  - ALSO SELECT ANY PEAK THAT IS WITHIN 80% OF PEAK
- ROTATE PATCH ACCORDING TO PEAK ORIENTATION TO 'UP'



TRANSLATION, ROTATION, AND SCALE INVARIANT DESCRIPTOR



	8x8	2x2
In practice:	16x16	4x4

(4 x 4) cells x 8 orientations = 128 length descriptor

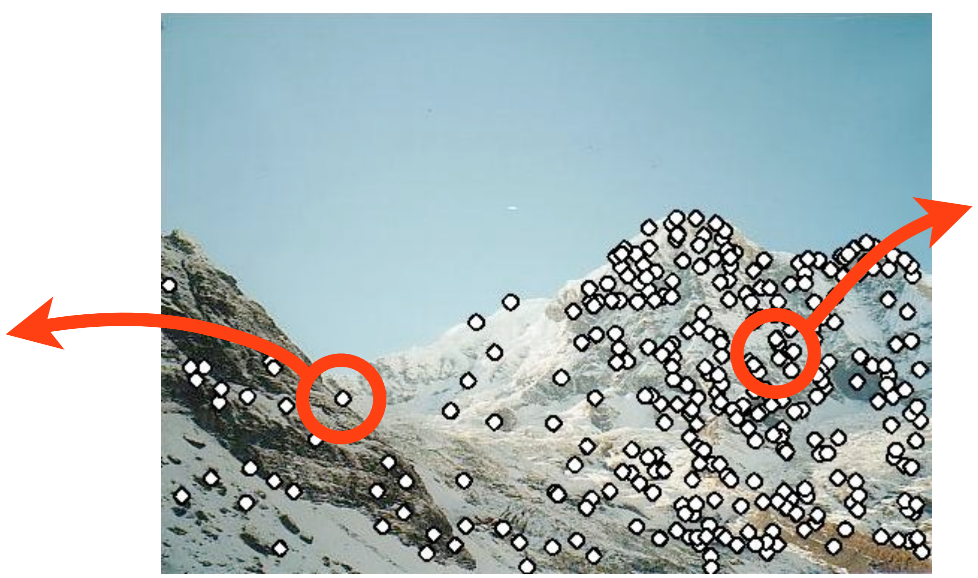
SCALE INVARIANT FEATURE TRANSFORM

- DAVID LOWE, DISTINCTIVE IMAGE FEATURES FROM SCALE INVARIANT KEYPOINTS, *INTERNATIONAL JOURNAL OF COMPUTER VISION*, 2004

LOCAL FEATURES

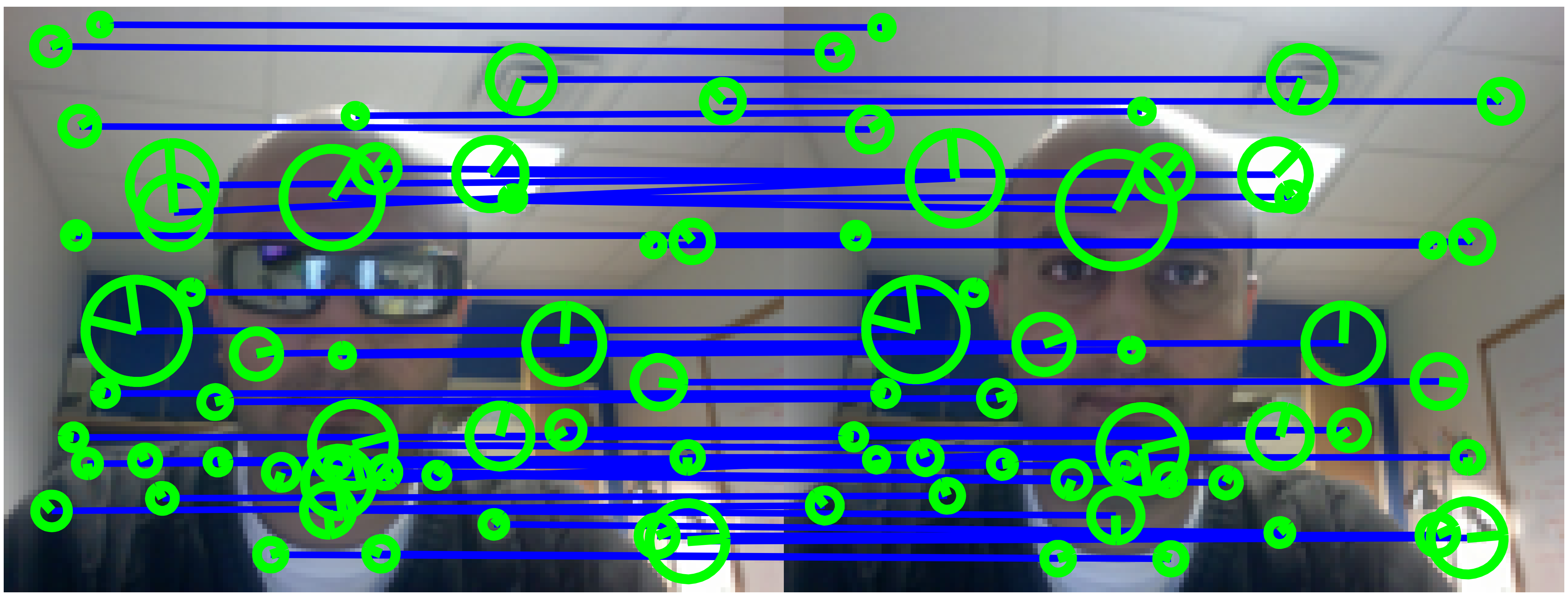
- DETECTION: IDENTIFY INTEREST POINT
- DESCRIPTION: EXTRACT FEATURE VECTORS FOR EACH POINT
- **REGISTRATION**: DETERMINE CORRESPONDENCE BETWEEN INTEREST POINTS

$$\mathbf{x}_1 = \begin{bmatrix} \mathbf{x}_{11} \\ \mathbf{x}_{12} \\ \vdots \\ \mathbf{x}_{1N} \end{bmatrix}$$



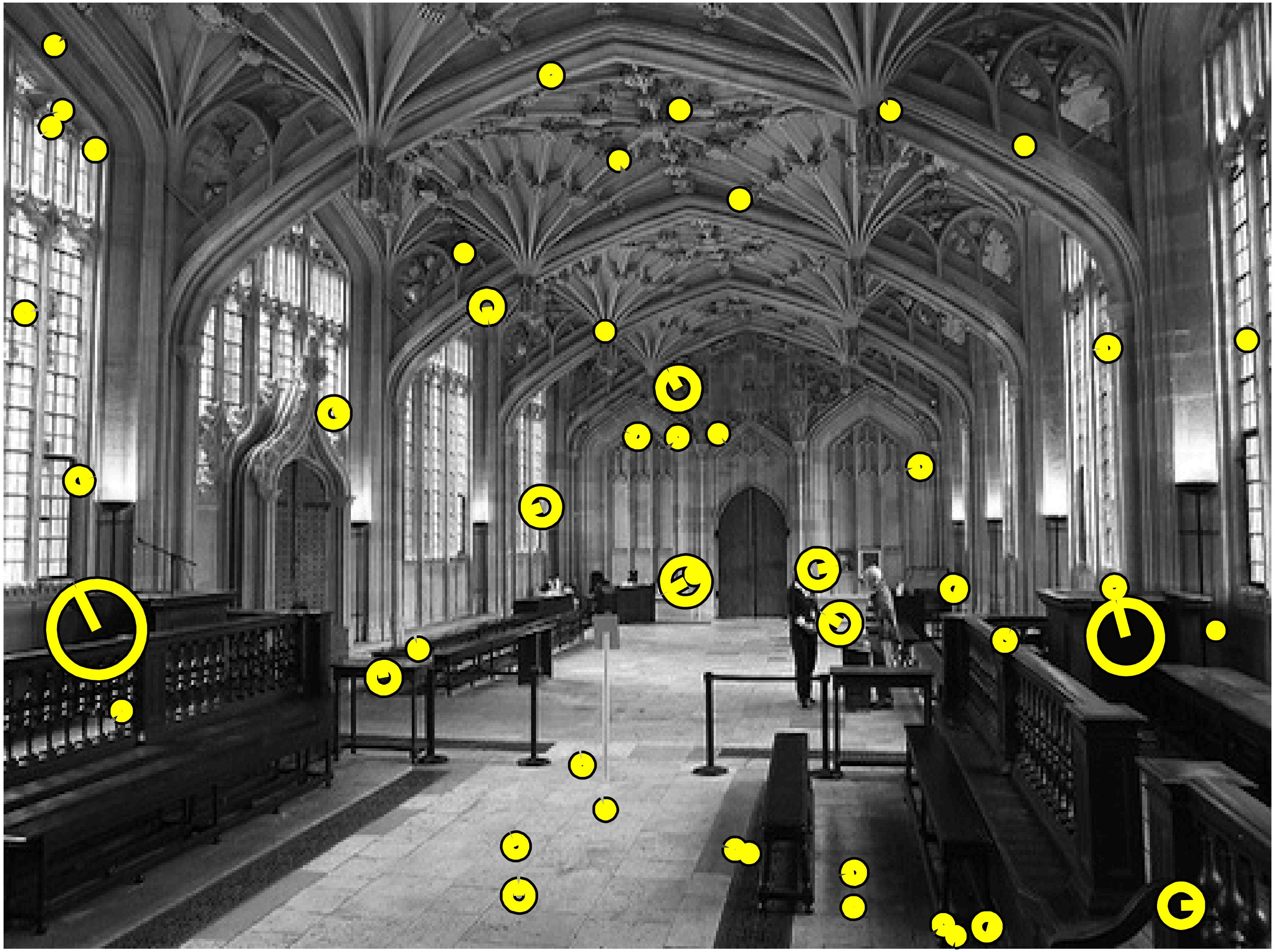
$$\mathbf{x}_2 = \begin{bmatrix} \mathbf{x}_{21} \\ \mathbf{x}_{22} \\ \vdots \\ \mathbf{x}_{2N} \end{bmatrix}$$

SIFT MATCHING

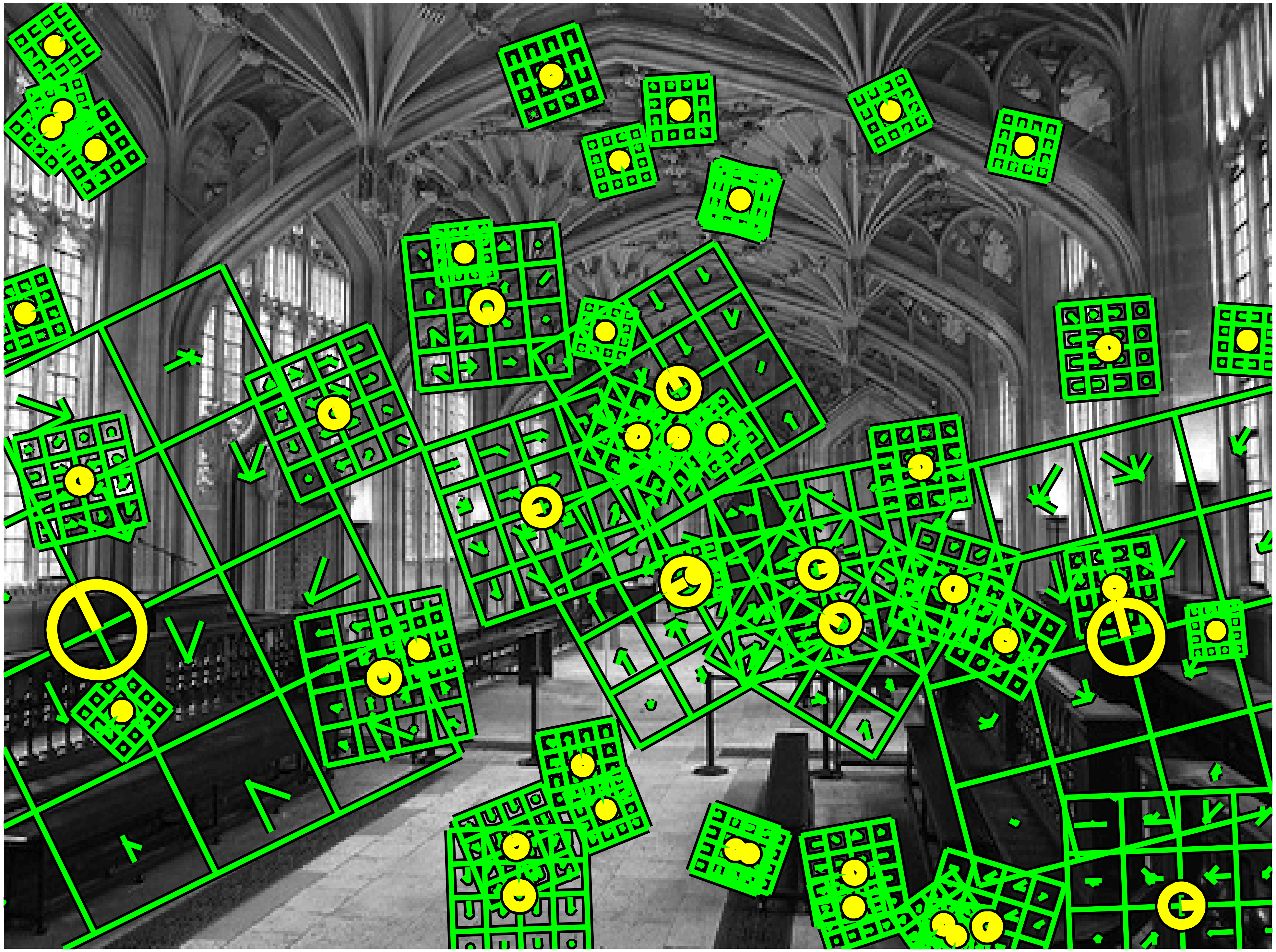


SIFT

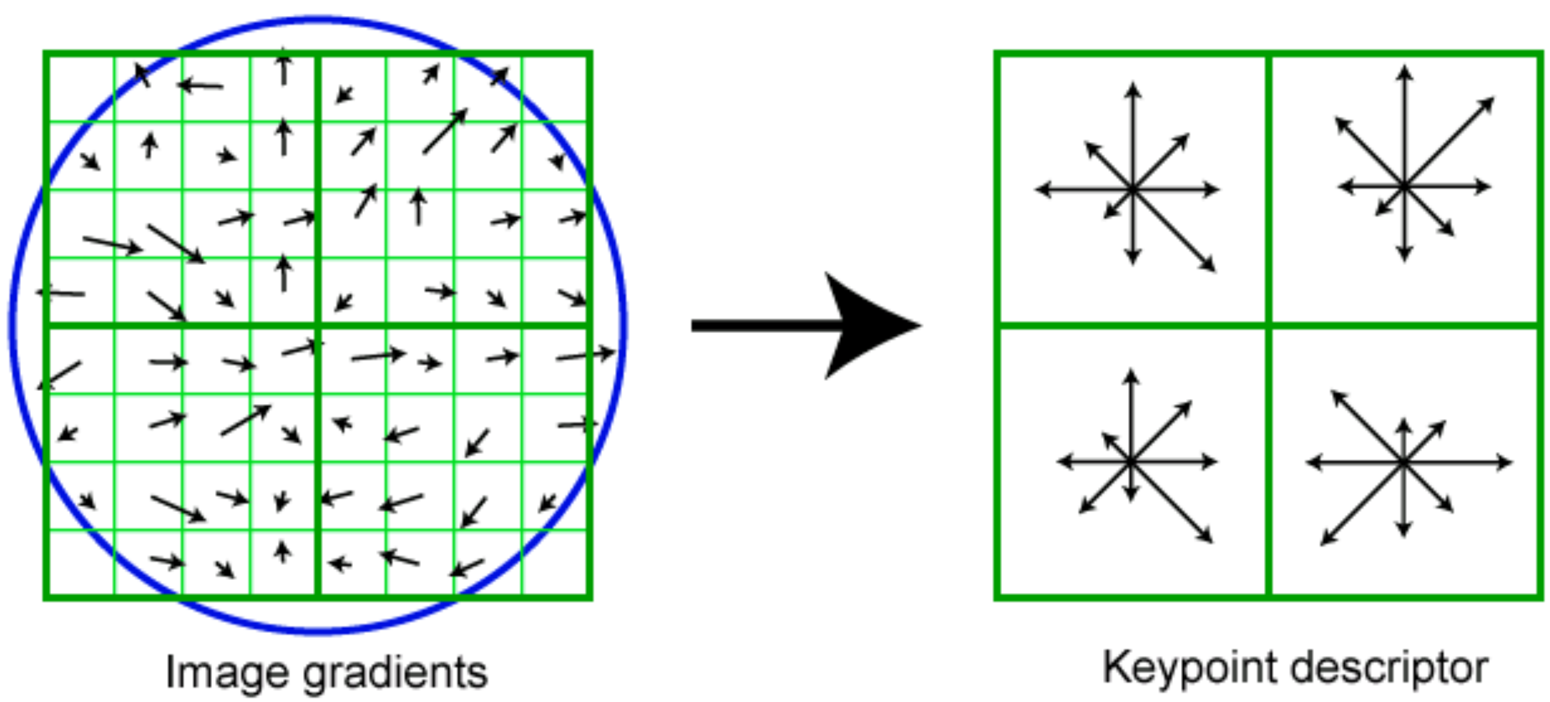
SIFT



SIFT



TRANSLATION, ROTATION, AND SCALE INVARIANT DESCRIPTOR



8x8 2x2

In practice: 16x16 4x4

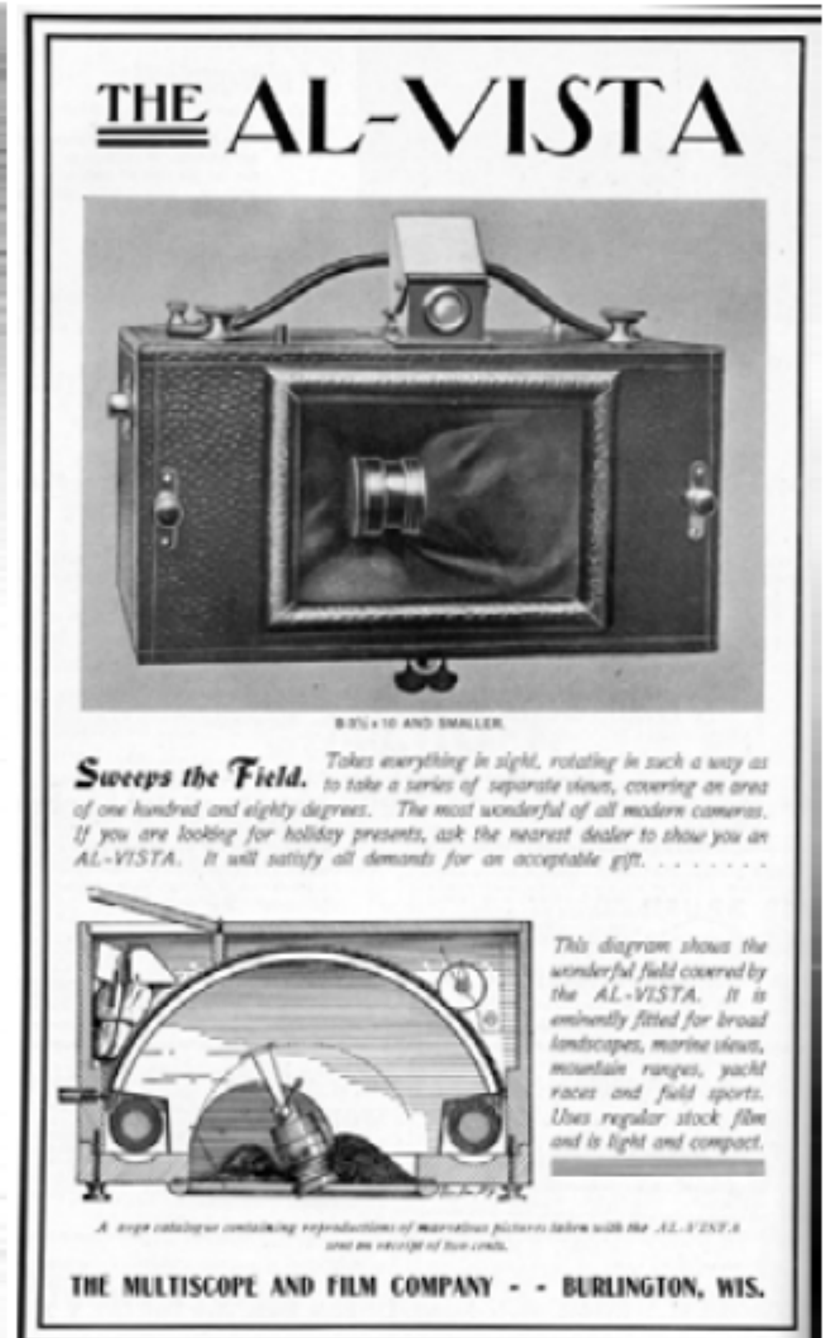
(4 x 4) cells x 8 orientations = 128 length descriptor

SIFT

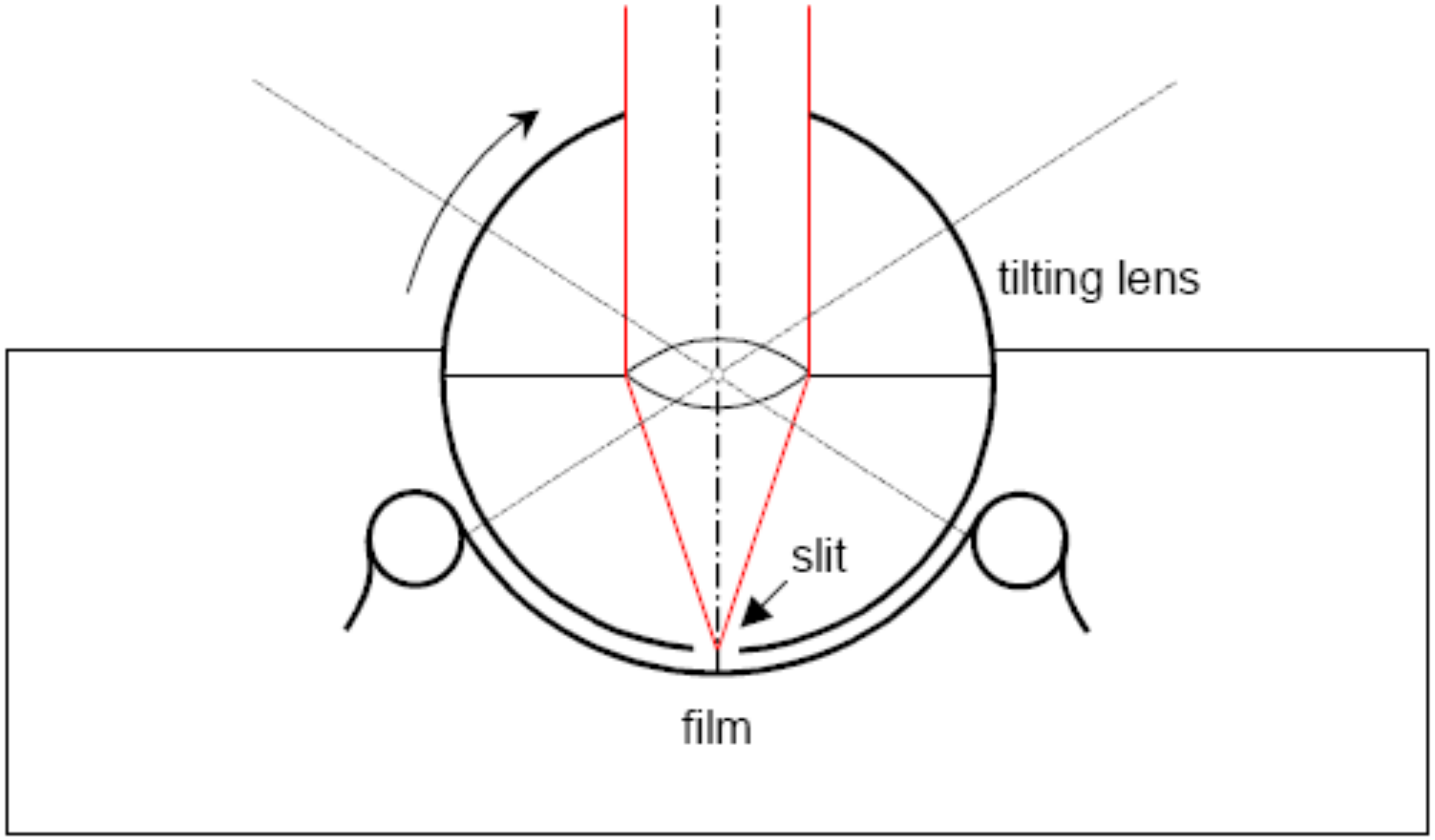
IN THIS CLASS...

HOW TO DETECT AND MATCH POINTS THAT ARE ROBUST TO:

- TRANSLATION,
- ROTATION,
- SCALE,
- ILLUMINATION CHANGES



Al-Vista, 1899 ($20)



Swing lens (1843 – 1980s)