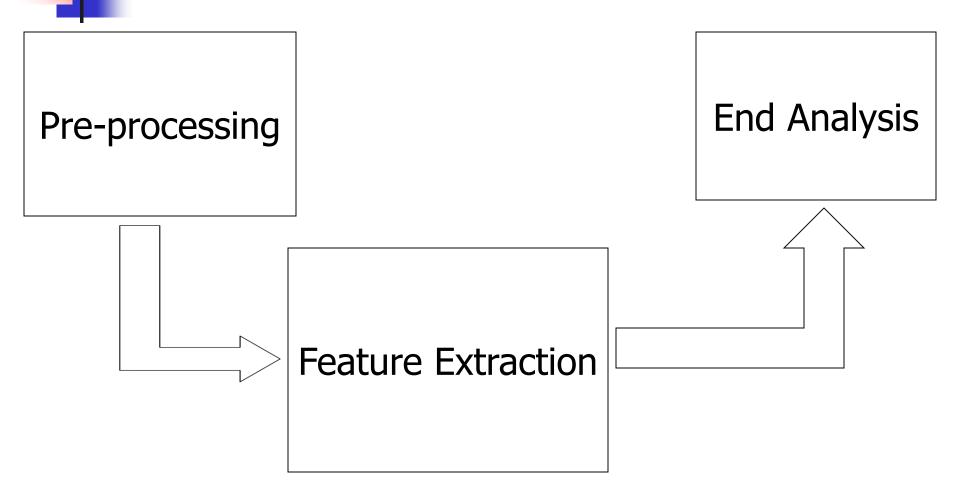
Feature detection and description

Jayanta Mukhopadhyay
Dept. of Computer Science and Engg.

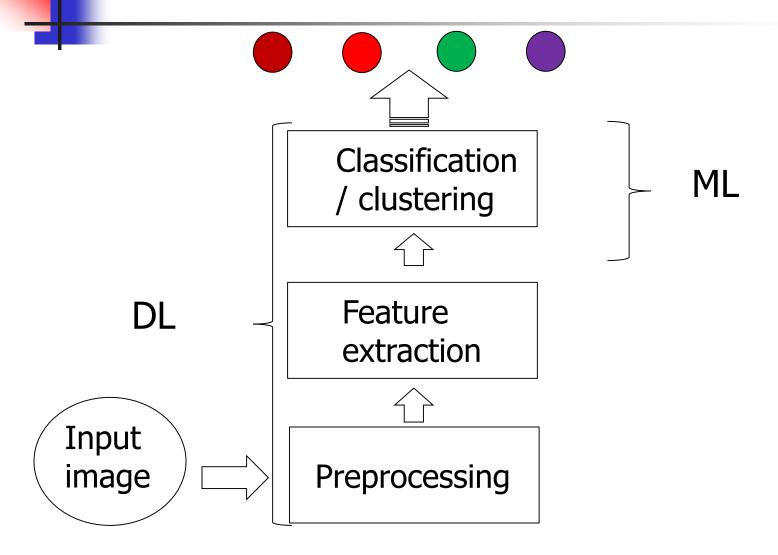
A feature of an image?

- Characterizes its visual content.
 - A part of an organization on describing a region
 - A higher level organization than a pixel
- A point in a multidimensional space
 - feature vector: n-Dimensional vector: x ∈ Rⁿ
- Represents
 - a point within a neighborhood
 - a pattern
 - a patch
 - an object
 - the whole image





Role of features in learning visual content



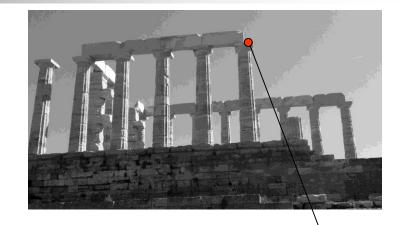


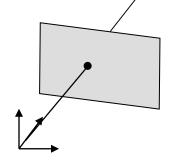
A few typical problems

- To determine structure / geometry of objects.
- To determine material properties of objects.
- To determine an object's position, category, and its role / interaction with other objects.
- To find similar images from a library given a query image

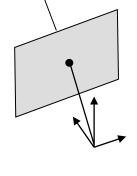
Example: Structure Recovery







Given two views of the scene recover the unknown camera displacement and 3D scene structure



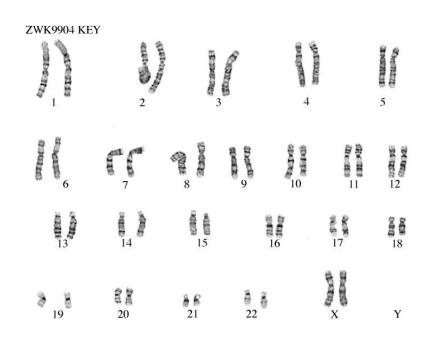
Adapted from www.cs.gmu.edu/~kosecka/lect4.ppt



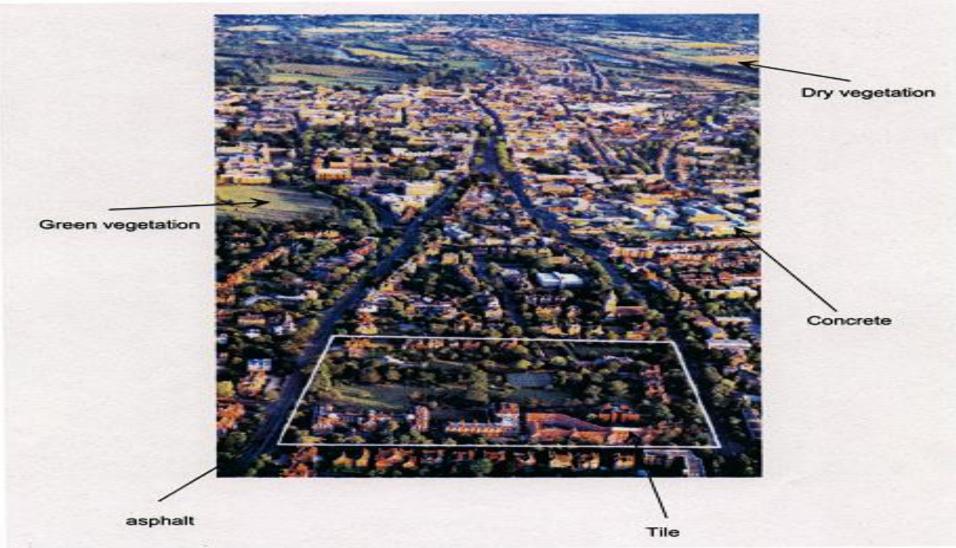
Example: Karyotyping

Pairing homologous chromosomes.





Example: Material Property from Remote Sensing Images



Courtesy: www.sal.ocean.washington.edu/teaching/lectures/classification.ppt

Example: Document Analysis

most after the wint own Papers; nowed allowers: the burley are until one. stope and any distressed a THE OTHER PARTY WAS I ARTHURSTED SHOWS NOT A THE WAY AND ADDRESS OF THE PARTY AND ADDRESS O WHITE WAY AND A PARK A P. I. soften largest contract fallence. White day you will a my over his a WHEN HER MINERAL PROPERTY. MARKATON STORY ASSESSED. MET PINE STREET ON APPEAL WORLDS THE WAY THE REST PROPERTY. Series wheels who plant SHORT AT PART IN THE LOCATION I THE RESERVE OF SERVICE Pitting is not cell more forms. ROSS arms drop appell sibratalest etcs ches fights volker a Personal state of the state of मेरे किम ज़र्त कर करता बाँड वर्कि । Responsible to application (IN CHECK FOR SPRING NO. WORK WAS REPORT MANUFACTURE. NAMES OF STREET AND ADDRESS OF PROBLEM OUR COLUMNS proton syrue custon warms where a THE RESERVE WE SHEET IS NOT MAN SAME AS A STRUCTURE OF THE PERSON NAMED IN COLUMN 2 AND RESCOULTED, STATES OF A PARTY enter you gill so the service of

WHEN PRIVATE AND COMP. INCOME.

करिया का सम अनेत अंदर्

AN ANY THE PERSON WHEN

er or elign at either her i

THE REAL PROPERTY WAS NOT AN A

WITCH

JAMES THE RESIDENCE AND ADDRESS. क्या नाम के गाँच अब अभी महान STREET ATTENDED NOT THE OWN. WE SET TO AN ARRANGE OF THE A 4940 SHAD IN TURNS 45 BESTER SERVERS THE RESERVED AND ADDRESS. THE DIE WE HAVE ADDRESS OFFICE A there were not send so hit. MICH WITH SUPERING ARE ROOMED OF THE WHEN A STREET THE SEC. WITH ACRES THE OTHER AND A WHEN MER WER PERSON THE WHEN PERSON AND A PERSON AS PROPERTY. wet whom all is written poor? AND AN YOR ACCOUNT OF ACCOUNT AN I WITH LOUIS THROWING YOU DECK OF WHITE SELECTION WHEN I AT RES YOUR WITH WITHIN WRITE I PARTIES NAME AND ADDRESS OF Religion on the second activity street a report or size will a AND NOT THE REAL PROPERTY. नामा जिल्ला कर्ते करते क्षेत्र प्रतिकात the Report Will will unlike to WHITE PERSONS AND PROPERTY. stilet effet der efferet utv i print store with the case tools. affects for one services with a medication was weard with a (fewers of the action to Seen) CHES WING AND ADDRESS OF A PRINTER PRINTER WITH WAY WAY WAY AF SHADEN (FOR 10% NIT) stitm from our wifes which NOT AN AND THE REAL PROPERTY. seed more registered with the register. Marchael and the Senior of the

Mark 33 46 NO NO 101 MCV 4TH STATES NO. 27 WHILE A WELL PERCE WORKER DESIGNATION MEDICAN STATE OF THE PARTY AND PERSONS ASSESSED.

REPRESENTED FOR

MARKET BY STREET WITH

man with will be so with MITTER NO. STATE SUCK SEED. SERVICE AND AND RESTREEN 254 911 22 010 6764 762 266 MR STEE SERVICE AND LAND TO THE PARTY OF THE PARTY A Major eries permits white-year my that the six tier on Maxithe one trailing lie tell. saw street work free laws all a man or we form as files prime HE SIC O'C ACCESS ACCES. HE STATE AND TOWN THE PARTY. SERVICE AND PROPERTY. are for elect apper urbi-WHEN THE PERSON WELL MILE TOTALS LISTED THE JULY WELL MEETS STORT AND AN ADDRESS. also one on on all enwithin williams and notes accom-Mary's ween at other wall. MEETS STOPP AND ARMS AND ADDRESS. with a liver spec after wors o and there are no personal to WERE THE REPORTS OF THE BE BE 175, SK-4C TOTAL TOTAL with process out 40 less who lively free next white i BRICE INVOCA, FROM NOW WITH MALE Mills cow will be for many

model with a real or stops. STREET, THE REAL PROPERTY. Name round whose or a WE WIN STATE OF LIST VICES then berfiet am uffler men : WHEN THE RESERVE ON were my my frem were we are at an other own o MAN COMMITTER SERVICE AND ADDRESS. WHICH DANKED HE REPORT ! ANY OF RESIDENCE PROPERTY AND A OF TRECIPE OF STATE OF teres by any of the course of the SHEET AND WHEN THE REAL PROPERTY. STREET, SON GLOSS ST. Bell all an ery attenues of a or feet speed togo like Constant. and the free rate more reports a A RESIDENCE STOP STOP AND AND AND May 62 House Property of HOUSE, SCHOOL BOOK BOOK OF THE WOOD OF THE REAL PROPERTY. CHIEF CO. PROP SPEC THE STREET STREET, STREET, STREET, METERIORISTANCE ON ANCHE with a linear open side. NOT ON FAMILY DISCRETE NOTE IN A SHE'S JULY ON HOME HERE'S limits reportify your allies a THE REPORT OF ALL PROPERTY. THE RESIDENCE OF STREET AN OF AN ASSESSMENT A ADD NOT THE REAL PROPERTY. the address of the later was a FEW WALL DIT CON TIME. PRINT AT WHITE BETWEEN THE REAL PRINTERS. HEAD AND STATEMENT SPECIA With the the the the witter READ AND CHIEF TYPES WERE A SHOWN WHEN THE REAL PROPERTY. THE REPORT OF THE PARTY OF MOREOVER AND ADDRESS AND ADDRESS. AND PROCESS OF THE WAY WITH THE A CHANGE OF SECRETARY PROPERTY. Named Stiller or helf-less size in मिक्न रहत नहां का ताल करते पर । and elegated that influence has a ER Penner will structure west was explose specified with a and the cell fell self-actions: NAME OF TAXABLE PARTY. PERSONAL PROPERTY AND A PERSON Reservoire degrees off was WHEN REPORT WHE WINDOW YOU I WHEN DAMES OFFICE WHEN IN WHICH COUNTY AND THE WORLD STORY DATE MAKEN HIS ARREST OF THE WHEN HE WIS STONE WITH A THE DWG W VIV YOURS सामाने कवित्रकात्रक प्राप्तानी साहर र A PERSONAL PROPERTY AND AND ASSESSED. Water of the add with from a and Bridger, etc., \$50 days not a ATTENDED BY THE PARTY STATE APPROPRIEST OF THE NOW AND DE THE WHITE HE BE NOT BE WHY DRIVE BE WITH LEWIS .. all washes when with left a THE RESERVE WHEN THE PARTY AND PARTY. WATER FORCES AND SOTHIN WICHOW विकास एक मान्य भारत कराई है। STREET AND WAY BE SEEN. WAR STREET STREET, ASK SEE MANUFACTURE AND MALE AND ADDRESS OF THE PARTY ADDRESS OF THE PARTY ADDRESS OF THE PARTY AND ADDR

THE PART AND REAL PROPERTY AND A

on his office an expecting a

Steer It on the last the HOTELS WITH SOUTHING WITH PRINCIPLE AND DESIGNATION OF BETWEEN AND ACTION STORES IN MARKETTAL

PERSONAL PROPERTY AND PERSONS tipe with CHG to be until MORE WHITE MICE OF THE BAT I MODEL HOW HE'VE ADVISE VALUE OF THE REST PRINTS WHEN YOU AND A MIN. WITH THAT HE WAS DRIVEN. MARKS STREET, ST. TO. WHILE A MINISTER AND RESIDENCE. was first for 20 for 20 five . Water Street Street, Treet, rathly AND HOSP WITH THE STREET, MARKET SPICE STREET THE TOTAL WITE SHE MORE WOMEN. IN WHICH FOR WHICH WHICH IS MINISTRAL STREET, SERVING are the above name of the MINISTRAÇÃO NIÑOS PROCESTOS. may active seem specially by BOTH ROOM WITH BY DOWN A WIRE KINN COLUMN TO WAS sections or been seen or security of WHEN THE PERSON AND A STATE AND IN MED STATE AND ADDRESS. well allow ages alter were a MALESCAN SCHOOL SCHOOL STATE A WEST HIS WENNINGSHIPS BE SERVICED, FOR MITTER STREET, with other visc trivials lies i will live him now who y Built larges, fairs, we all ma-Management and the factors

Example: Scene matching









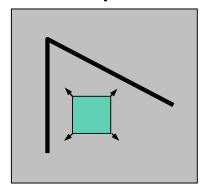
- Detection
- Description
- Matching

Translation, rotation, scale ...

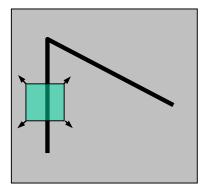
Feature detection

A local measure for uniquely identifying the feature.

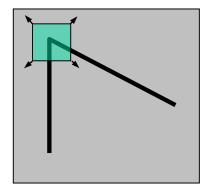
How does the measure change on shifting windows at different points?



"flat" region: no change in all directions



"edge": no change along the edge direction

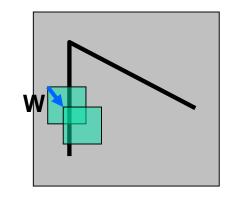


"corner":
significant
change in all
directions

Feature detection

Consider shifting the window **W** by (u,v).

- how do the pixels in W change?
- compare each pixel before and after by summing up the squared differences (SSD).
- this defines an SSD "error" of E(u,v).



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

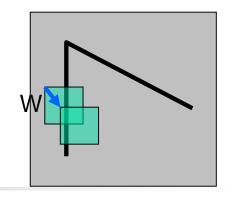
\$mall motion assumption

$$I(x+u,y+v) = I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v + \text{higher order terms}$$

For small u and v

$$I(x+u,y+v) \approx I(x,y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v$$
 $\approx I(x,y) + [I_x \ I_y] \begin{bmatrix} u \\ v \end{bmatrix}$ shorthand: $I_x = \frac{\partial I}{\partial x}$

Feature detection



$$I(x + u, y + v) \approx I(x, y) + \frac{\partial I}{\partial x}u + \frac{\partial I}{\partial y}v \approx I(x, y) + [I_x \ I_y] \begin{bmatrix} u \\ v \end{bmatrix}$$

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

$$pprox \sum_{(x,y)\in W} \left[I(x,y) + \left[I_x \ I_y\right] \left[\begin{array}{c} u \\ v \end{array}\right] - I(x,y)\right]$$

$$\approx \sum_{(x,y)\in W} \left[[I_x \ I_y] \left[\begin{array}{c} u \\ v \end{array} \right] \right]^2$$

$$E(u,v) = \sum_{(x,y) \in W} [u \ v] \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

 $\begin{bmatrix}
I_{x} & I_{y} \begin{bmatrix} u \\ v \end{bmatrix}
\end{bmatrix}^{I} \begin{bmatrix}
I_{x} & I_{y} \begin{bmatrix} u \\ v \end{bmatrix}
\end{bmatrix}$ $= \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} I_{x} \\ I_{y} \end{bmatrix} \begin{bmatrix} I_{x} & I_{y} \begin{bmatrix} u \\ v \end{bmatrix}
\end{bmatrix}$ $= \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} I_{x}^{2} & I_{x}I_{y} \\ I_{y}I_{x} & I_{y}^{2} \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$

$$E(u,v) = \sum_{(x,y)\in W} [u\ v] \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$

Feature detection

$$E(u,v) = \sum_{(x,y) \in W} [u \ v] \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$
 For the example above

- Suppose the center of the green window moved to anywhere on the blue unit circle.
- Directions for the largest and the smallest E values?
 - Eigenvectors of *H*.

Quick eigenvalue/eigenvector review

$$Ax = \lambda x$$

The **eigenvectors** of a matrix A are the vectors x that satisfy:

$$det(A - \lambda I) = 0 \ det \begin{bmatrix} h_{11} - \lambda & h_{12} \\ h_{21} & h_{22} - \lambda \end{bmatrix} = 0$$

The scalar λ is the **eigenvalue** corresponding to x

The eigenvalues are found by solving:

$$\lambda_{\pm} = \frac{1}{2} \left[(h_{11} + h_{22}) \pm \sqrt{4h_{12}h_{21} + (h_{11} - h_{22})^2} \right]$$

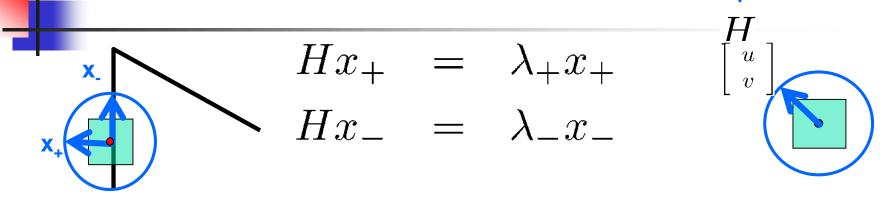
For eigen vector solve the following:

$$\begin{bmatrix} h_{11} - \lambda & h_{12} \\ h_{21} & h_{22} - \lambda \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0$$

Once you know λ , you find eigen vector $[x\ y]^T$ by solving the above.

Slide adapted from Darya Frolova, Denis Simakov, Weizmann Institute.

$$E(u,v) = \sum_{(x,y) \in W} [u \ v] \begin{bmatrix} I_x^2 & I_x I_y \\ I_y I_x & I_y^2 \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix}$$
 Feature detection



Eigenvalues and eigenvectors of H

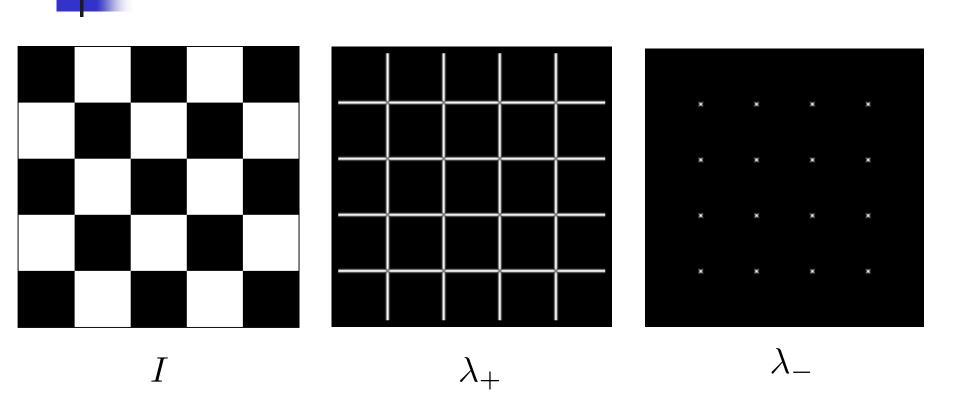
- Define shifts with the smallest and largest change (E value)
- x_{+} = direction of **largest** increase in E.
- λ_{+} = amount of increase in direction x_{+}
- x_{\perp} = direction of **smallest** increase in E.
- λ = amount of increase in direction x_{\perp}

Feature scoring function

E(u,v) to be **large** for small shifts in **all** directions

- the *minimum* of E(u,v) should be large, over all unit vectors $[u \ v]$.
- this minimum is given by the smaller eigenvalue (λ₋) of H

Feature detection



Feature detection: Algorithm

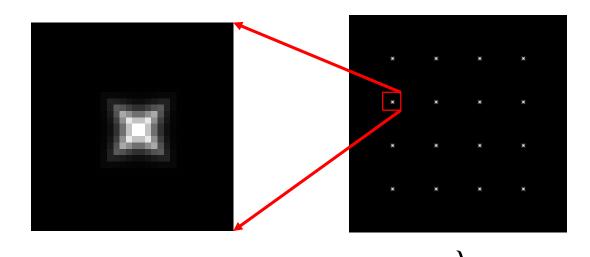
- Compute the gradient at each point in the image.
- Obtain the H matrix from the entries in the gradient.

$$H = \begin{bmatrix} \overline{I_{\chi}^2} & \overline{I_{\chi}}\overline{I_{y}} \\ \overline{I_{y}}\overline{I_{\chi}} & \overline{I_{y}^2} \end{bmatrix}$$

- Compute eigenvalues of H.
- Locate points with large response (λ_{-} > threshold)
- Select points where λ₋ is a local maximum as features.

Local maximum: illustration

• Points with $\lambda_{\underline{}}$ as a local maximum.



The Harris operator

 $\lambda_{\underline{l}}$ is a variant of the "Harris operator" for feature detection.

$$f = \frac{\lambda_1 \lambda_2}{\lambda_1 + \lambda_2}$$

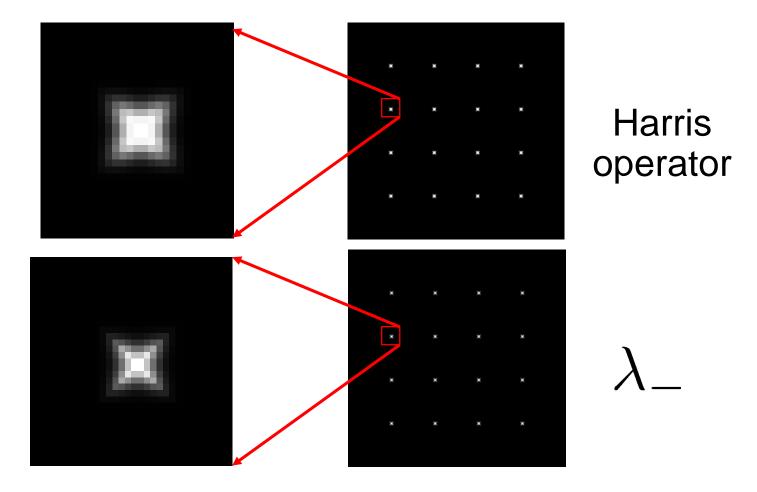
$$= \frac{determinant(H)}{trace(H)}$$

• The trace is the sum of the diagonals, i.e.,

$$trace(H) = h_{11} + h_{22}$$

- Very similar to λ but less expensive (no square root)
- Called the "Harris Corner Detector" or "Harris Operator"
- Lots of other detectors, this is one of the most popular.

The Harris operator



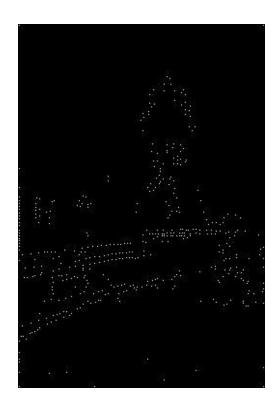
Harris detector example



Input image

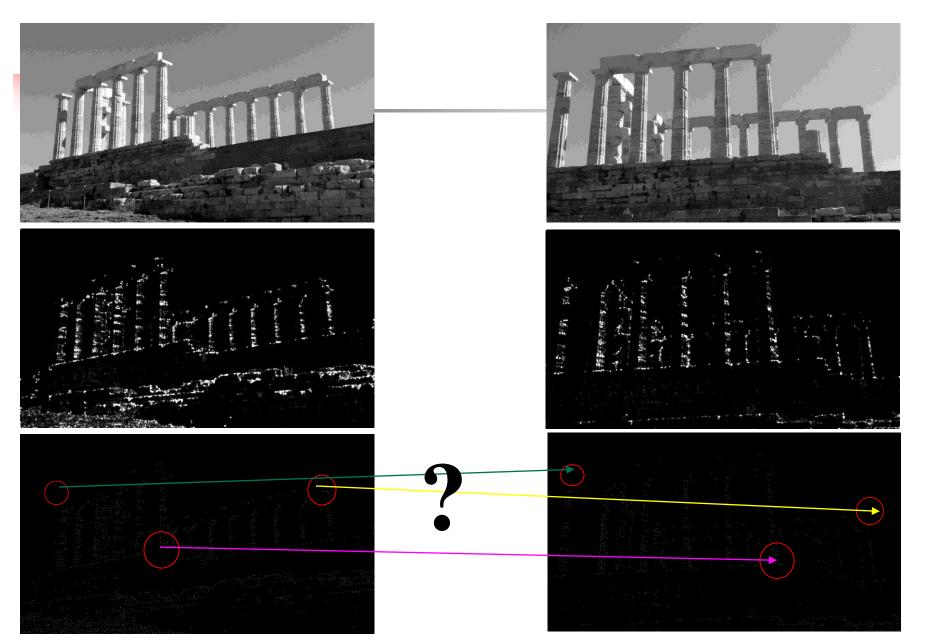


f-Value



Local maxima

Feature matching





Matching with Features

- Detect feature points in both images.
- Describe them by local statistics.
- Find corresponding pairs (Matching).



Invariance

- Is it possible to select same features under various transformations?
 - Rotation
 - Change of illumination
 - Scale

O .

Scale invariant detection

Key idea for detecting corners Find scale that gives local maximum of f.

- f is a local maximum in both position and scale
- Common definition of f: Laplacian (or difference between two Gaussian filtered images with different s.d.'s.).



Invariance

Consider two images I_1 and I_2 so that I_2 is a transformed version of I_1

- For example,
 - Translation
 - Rotation
 - Scale
 - Reflection
 - Non-uniform scaling
 - Illumination
 - View ...

Tranformational Invariance: Detection of the same features regardless of the transformation.

Detection and description: to be invariant

Both should be ensured.

- 1. Detector to be transformational invariant.
 - Harris measure is invariant to translation and rotation.
 - Scale requires handling of multi-resolution representation
 - Usual approach: Computing features at multiple scales using a Gaussian pyramid.
 - More precise computation locates features at "the best scale" (e.g., SIFT)
- 2. Feature descriptor to be transformational invariant.
 - captures the information in a region around the detected feature point.
 - e.g. histogram of gradient directions in a square window centering a feature point.

Finding Keypoints – Scale, Location

Information content on varying scale?









Scale Invariant Detection

Functions for determining scale Convolve with image and observe extrema in 3-D.

Kernels:

Laplacian

$$L = \sigma^2 (G_{xx}(x, y, \sigma) + G_{yy}(x, y, \sigma))$$

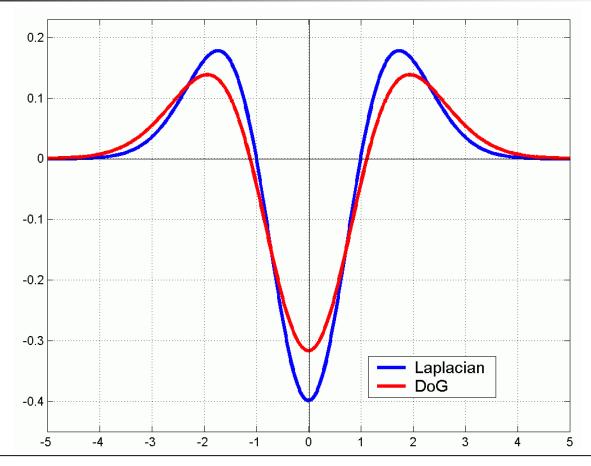
Difference of Gaussians

$$DoG = G(x, y, k\sigma) - G(x, y, \sigma)$$

where Gaussian

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





Both kernels are invariant to scale and rotation.



Relationship between LoG and DoG operator

$$\frac{\partial G}{\partial \sigma} = \sigma \nabla^2 G$$

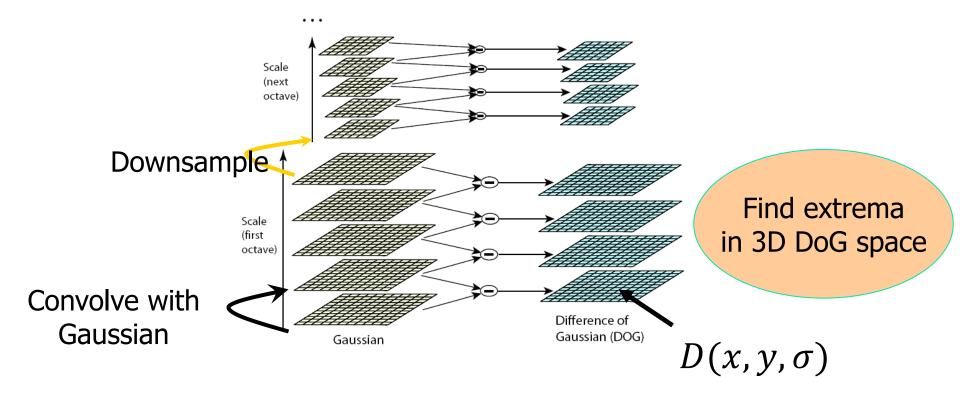
$$\sigma \nabla^2 G = \frac{G(x, y, k\sigma) - G(x, y, \sigma)}{k\sigma - \sigma}$$

$$G(x, y, k\sigma) - G(x, y, \sigma) = (k - 1)\sigma^2 \nabla^2 G$$

The factor (k-1) is kept constant across scales.

→ does not influence extrema locations.

Finding Keypoints - Scale, Location



Typical example DOGs at different scales)





Octave:2



 $\sigma = 2^{2/3}$



 $\sigma = 2^{5/6}$



 $\sigma=2$



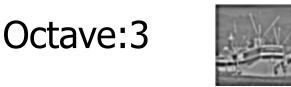
$$\sigma = 2^{7/6}$$

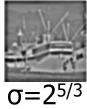


 $\sigma = 2^{4/3}$



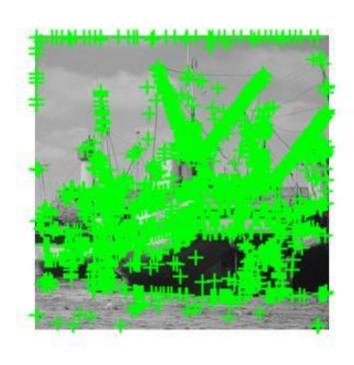
 $\sigma = 2^{3/2}$







Extrema of DOG stack: Example)





Keypoint localization

- A lot of points, some of them are not good enough.
 - Edge points
- Locations may be not accurate.

Eliminating edge points

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

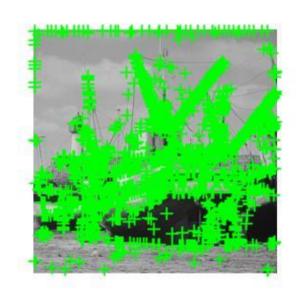
D(.)=DoG of the function.

- Edge point: large principal curvature across the edge but a small one in the perpendicular direction.
- The eigenvalues of H are proportional to the principal curvatures, so two eigenvalues shouldn't differ too much.

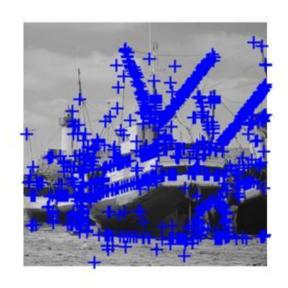
Eliminate keypoint if the ratio greater than
$$\frac{Tr(H)^2}{Det(H)} < \frac{(r+1)^2}{r}$$
 for high value of r (say 10)

Slide adapted from Darya Frolova, Denis Simakov, Weizmann Institute.

Key-points filtering: An example

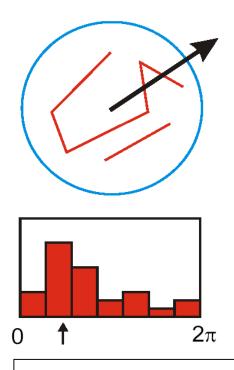


Initial set



After ratio threshold

Orientation assignment



- Create histogram of local gradient directions at selected scale.
- Assign canonical orientation at peak of smoothed histogram.
- Each key specifies stable 2D coordinates (x, y, scale, orientation).

If there are two major orientations, use both.

Slide adapted from Darya Frolova, Denis Simakov, Weizmann Institute.

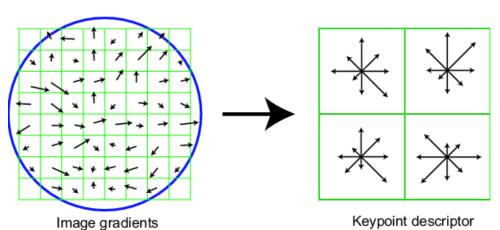


Keypoint Descriptors

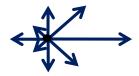
- At this point, each keypoint has
 - location
 - scale
 - orientation
- Next is to compute a descriptor for the local image region about each keypoint that is
 - highly distinctive
 - invariant as possible to variations such as changes in viewpoint and illumination

Scale Invariant Feature Transform (Lowe'99, ICCV)

- Take 16x16 square window (orientation corrected) around detected feature
- Compute edge orientation (angle of the gradient 90°) for each pixel
- Throw out weak edges (threshold gradient magnitude)
- Create histogram of surviving edge orientations



 $\begin{array}{cccc}
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & & \\
 & & \\
 & & & \\
 & & \\
 & & & \\
 & & \\
 & & & \\
 & & \\
 & & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 & & \\
 &$



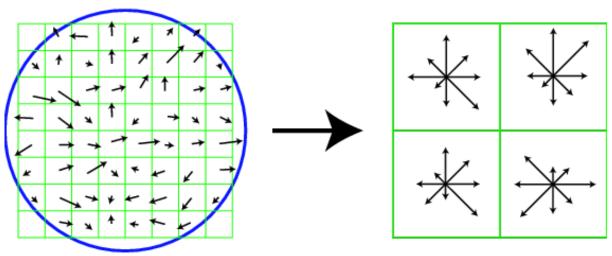
Adapted from slide by David Lowe

SIFT descriptor

- Divide the 16x16 window into a 4x4 grid of cells (2x2 case shown below).
- Compute an orientation histogram for each cell

16 cells * 8 orientations = 128 dimensional

descriptor.



Keypoint descriptor

Robustness

Capable of handling

- changes in viewpoint
 - Up to about 60 degree out of plane rotation
- significant changes in illumination
 - Sometimes even day vs. night

Speeded-Up Robust Features (SURF): Another descriptor

- Speeded-Up Robust Features (SURF)
 - (Bay et al. ECCV, 2006)
 - Box-type convolution filters and use of integral images to speed up the computation.
 - Use of Hessian operator for key point detection → Local maxima of det(H).
 - Accumulate orientation corrected Haar wavelet responses.



Hessian Operator

Convolution with the Gaussian second order derivative with image.

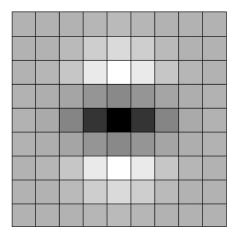
$$\mathcal{H}(\mathbf{x}, \sigma) = \begin{bmatrix} D_{\chi\chi}(\mathbf{x}, \sigma) & D_{\chi\chi}(\mathbf{x}, \sigma) \\ D_{\chi\chi}(\mathbf{x}, \sigma) & D_{\chi\chi}(\mathbf{x}, \sigma) \end{bmatrix}$$

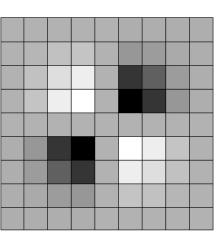
Keypoint: Maximum of $det(\mathcal{H}(.))$ over space and scale.

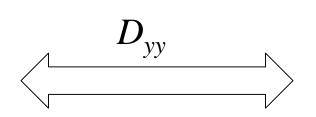
Approximation of Gaussian by



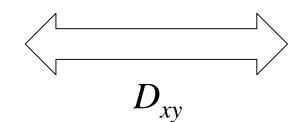
Bay et al, Speeded up robust features (SURF), CVIU, 2008

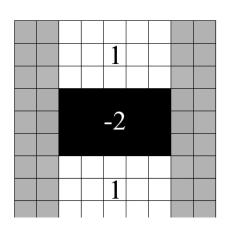


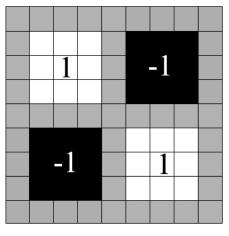




9x9 Box-filters are approximation of Gaussian width 1.2.







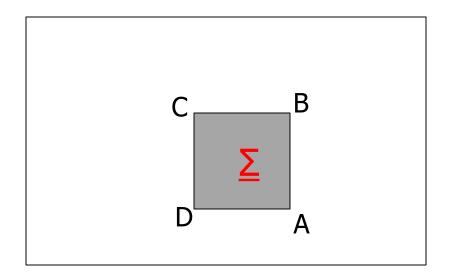
$$\det(H_{approx}) = D_{xx}D_{yy} - (wD_{xy})^2$$

w∼0.9



Fast computation using the integral image

$$I_{\Sigma}(\mathbf{x}) = \sum_{i=0}^{i \le x} \sum_{j=0}^{j \le y} I(i,j)$$

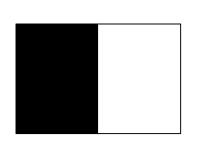


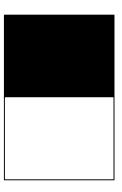
Only 3 additions and four memory access.

$$\Sigma = A - B - C + D$$



- Dominant orientation by accumulating Haar horizontal and vertical responses in a rotated sliding window (of width 60°) at the scale of key point.
- The longest vector provides the dominant direction.



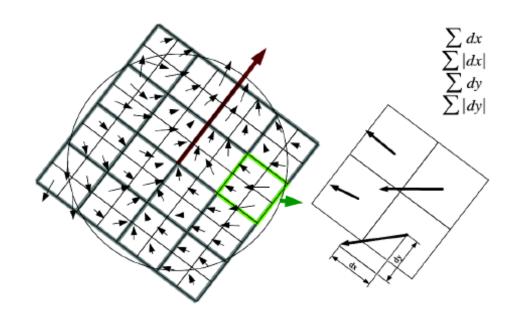


6 operations needed for computing each filter response using integral image.

Haar Filters
Box filter implementation.

SURF: Sum of Haar Wavelet responses

- Partitioned into 4x4 square sub-regions.
- Haar wavelet
 responses at
 regularly spaced
 5x5 sample patches
 in each sub-region.
- Each sub-region has 4D vector.
- Concatenate them to 64 D vector.



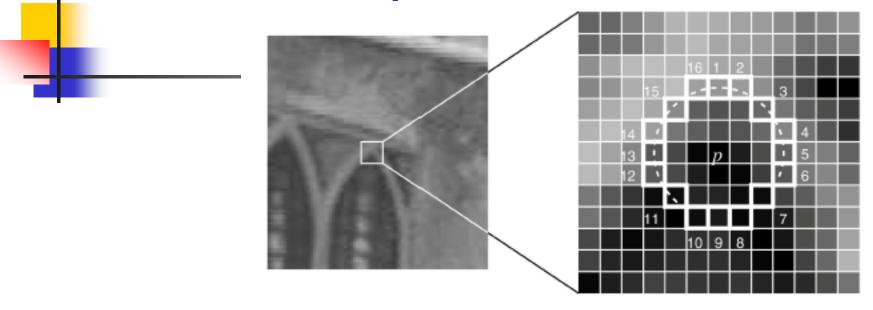
Size of the window=20 x scale

Bay et al, Speeded up robust features (SURF), CVIU, 2008



- FAST: Features from Accelerated Segment Test (Rosten et al, PAMI, 2010).
- BRIEF: Binary Robust Independent Elementary Features (Colonder et al, ECCV, 2010).
- ORB: Oriented FAST and Rotated BRIEF (Rublee et al, ICCV, 2011)

FAST: Principle



- 12 point test: If there exists 12 consecutive points in the set of 16 points brighter than the central pixel (Rosten et al ICCV'05).
- Modified strategy: Train decision tree on the boolean conditions given labelled data.

Rosten et al, Faster and Better: A Machine Learning Approach to Corner Detection, PAMI, 2010

FAST: Partitioning of points on the circle and Training a DT

$$S_{p \to x} = \begin{cases} d, & I_{p \to x} \le I_p - t & (darker) \\ s, & I_p - t < I_{p \to x} < I_p + t & (similar) \\ b, & I_p + t \le I_{p \to x} & (brighter) \end{cases}$$

- Classification of points in three classes and create three partitions for a point.
- Train a decision tree.

Rosten et al, Faster and Better: A Machine Learning Approach to Corner Detection, PAMI, 2010

4

BRIEF: Principle

- O Generate randomly a set of n_d pairs of locations $(\mathbf{x_i}, \mathbf{y_i})$ in a patch centered around a point.
- Perform the following boolean test.

$$\tau(\mathbf{p}; \mathbf{x}, \mathbf{y}) = \begin{cases} 1 & if \ \mathbf{p}(\mathbf{x}) < \mathbf{p}(\mathbf{y}) \\ 0 & otherwise \end{cases}$$

BRIEF descriptor: n_d Dimensional binary string

$$f_{n_d}(\mathbf{p}) = \sum_i 2^{i-1} \tau(\mathbf{p}; \mathbf{x}_i, \mathbf{y}_i)$$

128, 256, 512

Colonder et al, BRIEF: Binary Robust Independent Elementary Features, ECCV, 2010

ORB: Principle

- FAST does not operate across scale.
- Apply FAST detector on pyramid of smoothened images.
- Orientation by intensity centroid: The vector from the center of the patch and to a centroid considering the intensity distribution (intensity weighted center of patch).
- Rotate the patch by the angle and compute BRIEF called steered BRIEF.

Matching

- 'Representation of a key-point by a feature vector.
 - e.g. $[f_0 f_1 ... f_n]^T$
- Use distance functions / similarity measures.
 - L₁ norm

$$L_1(\vec{f}, \vec{g}) = \sum_{i=0}^{n} |f_i - g_i|$$

L₂ norm

$$L_2(\vec{f}, \vec{g}) = \left(\sum_{i=0}^n |f_i - g_i|^2\right)^{\overline{2}}$$

L_n norm

$$L_p(\vec{f}, \vec{g}) = \left(\sum_{i=0}^n |f_i - g_i|^p\right)^{\overline{p}}$$

Region descriptors

- Patch descriptors
- Texture descriptors
- Image / Sub-Image global descriptors

Patch Descriptor: Histogram of Gradients (HoG)

- Compute centered horizontal and vertical gradients with no smoothing.
- Compute gradient orientation and magnitudes,
- For color image, pick the color channel with the highest gradient magnitude for each pixel.
- For a 64x128 image, divide the image into 16x16 blocks of 50% overlap. →7x15=105 blocks in total.

N.Dalal and B. Triggs, Histograms of oriented gradients for human detection, CVPR-2005

4

Histogram of Gradients (HoG)

- Each block: 2x2 cells with size 8x8.
- Quantize the gradient orientation into 9 bins.
- The vote is the gradient magnitude.
- Interpolate votes between neighboring bin center.
- The vote can also be weighted with Gaussian to down-weight the pixels near the edges of the block.
- Concatenate histograms.
 - Feature dimension: 105x4x9 = 3,780

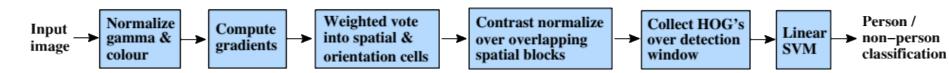
N.Dalal and B. Triggs, Histograms of oriented gradients for human detection, CVPR-2005

Object detection with patch descriptors.

- Typical examples:
 - Pedestrian detection
 - Character recognition
- Detection as a classification task.
 - Generate labeled sample feature descriptors.
 - Train a classifier.
 - NN, SVM, Decision Tree, Random Forest
 - Label an unknown patch using its descriptor.

Applications

Pedestrian Detection



N.Dalal and B. Triggs, Histograms of oriented gradients for human detection, CVPR-2005



Non-maximal suppression

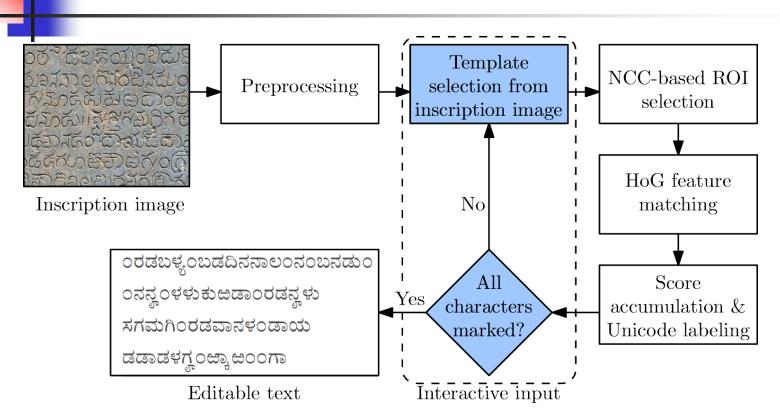
- Expected to get a high detection score with neighboring overlapping patches.
 - Select the patch with locally maximal score.
- A greedy approach:
 - Select the best scoring window
 - It is expected to cover the target object.
 - Suppress the windows that are too close to the selected window.
 - Search next top-scoring windows out of the rest.

Character Spotting (E-PURALEKHAK)

- A tool for digital paleography
 - Aids analysis of inscription
 - Converts inscribed substrate to editable text
- Processing pipeline
 - Preprocessing denoising and normalization ([0 255])
 - Search possible locations of user indicated character by cross correlation
 - Prune the search results by HoG feature matching
 - Parse the Unicode file editable text

Shashaank. M. Aswatha et al., A Method for Extracting Text from Stone Inscriptions using Character spotting, ACCV, 2014.

Character Spotting (E-PURALEKHAK)



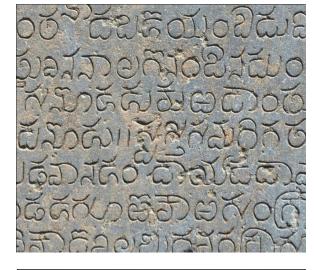
Character spotting – process flow

Shashaank. M. Aswatha et al., A Method for Extracting Text from Stone Inscriptions using Character spotting, ACCV, 2014.



E-PURALEKHAK

Original Inscription Image



Intermediate Result

Unicode generation



ಂರಡಬಳ್ಯಂಬಡದಿನನಾಲಂನಂಬನಡುಂ

ಂನನ್ಯಂಳಳುಕುಱಡಾಂರಡನ್ಯಳು

ಸಗಮಗಿಂರಡವಾನಳಂಡಾಯ

ಡಡಾಡಳಗ್ಯಂಱ್ಕಾಱ೦೦ಗಾ

Editable text



Texture descriptor



- Texture: spatial arrangement of the colors or intensities in an image
 - A quantitative measure of the arrangement of intensities in the region.

Texture descriptors

- Edge density and direction
- Local Binary Pattern (LBP).
- Co-occurrence Matrix.
- Laws' texture energy features.

Edge density and direction

- Compute gradient at each pixel.
- The descriptor: normalized histograms of magnitudes and directions of gradients over a region.
 (II (rescr) II (dir))
 - $(H_R(m'ag), H_R(dir))$

Normalized histogram of directions.

- Numbers of bins in histograms kept small (e.g. 10).
 - Use L1 norm between the feature vectors as a distance measure.
 Normalized histogram → Area = 1;



Local Binary Pattern (LBP).

$$b(i) = \begin{cases} 1 & if \ (I(i) > I(c)) \\ 0 & Otherwise \end{cases}$$

$$LBP(c) = \sum_{i=0}^{7} b(i)2^{i}$$
 You may have different ordering of neighbors

neighbors.

- Values range from 0 to 255.
- Obtain normalized histogram over a region.
- Not rotational invariant.
- Invariant to illumination and contrast.

T. Ojala, M. Pietikainen, and D. Harwood, A Comparative Study of Texture Measures with Classification Based on Feature Distributions, Pattern Recognition, vol. 29, pp. 51-59, 1996.

Variations of LBP

- Making it rotational invariant.
 - A circular neighborhood of radius R, with P pixels at equal intervals of angles.
 - Use interpolation if does not belong to the discrete grid.

$$LBP_{P,R}(c) = \sum_{i=0}^{P-1} b(i)2^{i}$$
 $LBP_{8,1} \leftarrow \rightarrow LBP$

$$LBP_{P,R}^{ri}(c) = \min\{ROR(LBP_{P,R}(c), i) \mid i = 0,1,2,...P - 1\}$$

where ROR(x, i) performs a circular bit-wise right shift on the P-bit number x, i times. 36 distinct values for $LBP^{ri}_{8,1}$.

T. Ojala, M. Pietikainen, and T.i Maenpaa, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, TPAMI, VOL. 24, NO. 7, JULY 2002, pp. 971-987

Variations of LBP

- Uniform pattern
 - Not more than 2 spatial transitions in the bit sequence.

Use rotation invariant value

• U(111<mark>01</mark>111)=2

by computing minimum

- U(10001001)=4

applying ROR operator.

$$LBP_{P,R}^{riu2}(c) = \begin{cases} \sum_{i=0}^{P-1} b(i)2^{i} & if \ U(LBP_{p,R}(c)) \leq 2\\ P+1 & Otherwise \end{cases}$$

Exactly P+1
uniform
patterns, hence
P+2 distinct
values.

T. Ojala, M. Pietikainen, and T.i Maenpaa, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, TPAMI, VOL. 24, NO. 7, JULY 2002, pp. 971-987

Nine uniform patterns of $LBP_{8,R}^{riu2}$

 The nine uniform patterns and the numbers inside them correspond to their unique codes.



Variance Measures of the Contrast

- Local variance of intensities of uniform pattern.
- Normalized histogram of local variances.

$$var_{P,R}(c) = \frac{1}{P} \sum_{p=0}^{P-1} (g_p - \mu)^2 \qquad \mu = \frac{1}{P} \sum_{p=0}^{P-1} g_p$$

Another robust representation: $\frac{LBP_{P,R}^{CBP}}{var_{P,R}}$

T. Ojala, M. Pietikainen, and T.i Maenpaa, Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns, TPAMI, VOL. 24, NO. 7, JULY 2002, pp. 971-987

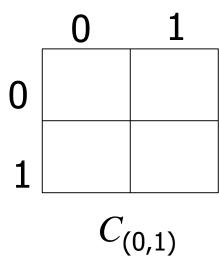
Co-occurrence Matrix (C_r)

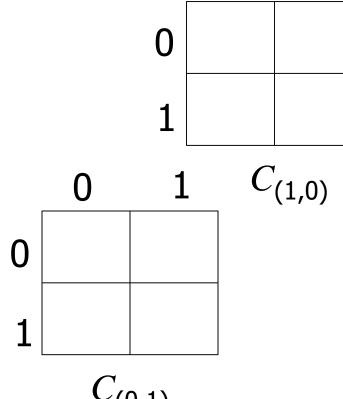
- $C_r(x,y)$: How many times elements x and y occur at a pair of pixels related spatially (designated by r in the notation).
 - e.g. \mathbf{p} r \mathbf{q} denotes \mathbf{q} is shifted from \mathbf{p} by a translation of $\mathbf{t}=(a,b)$, i.e. $\mathbf{q}=\mathbf{p}+\mathbf{t}$.
 - $C_{(a,b)}(x,y)$: Number of cases in an image where $I(\mathbf{p})=x$ and $I(\mathbf{p}+\mathbf{t})=y$.



Co-occurrence Matrix (C_r)

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0

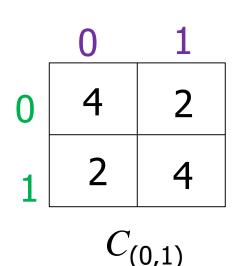


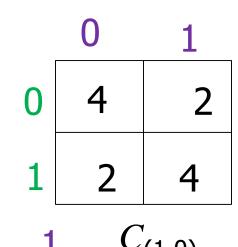


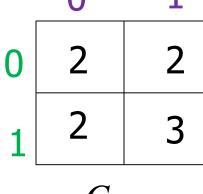


Co-occurrence Matrix (C_r)

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0





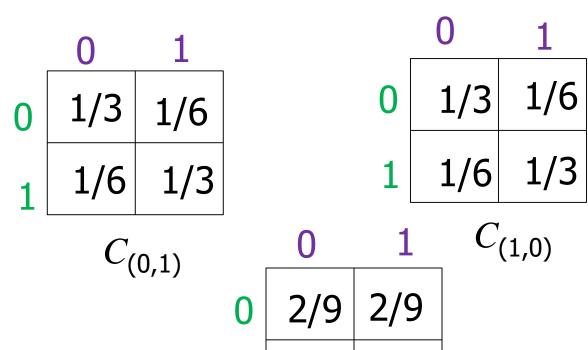




Normalized Co-occurrence Matrix (N_r)

Divide by the sum of frequencies in a matrix.

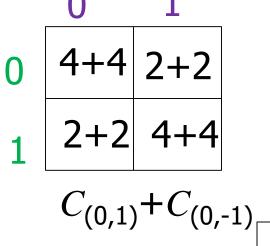
0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0

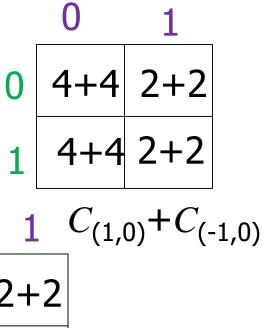


Symmetric Co-occurrence Matrix (S_r)

$$S_r(x,y) = C_r(x,y) + C_{-r}(x,y)$$

0	0	1	1
0	0	1	1
1	1	0	0
1	1	0	0





$$C_{(1,1)}+C_{(-1,-1)}$$

Features from Normalized Cooccurrence Matrix

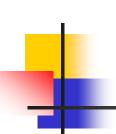
$$Energy = \sum_{x} \sum_{y} N_r^2(x, y)$$

$$Entropy = -\sum_{x} \sum_{y} N_r(x, y) \log_2 N_r(x, y)$$

$$Contrast = \sum_{x} \sum_{y} (x - y)^2 N_r(x, y)$$

$$Homogeneity = \sum_{x} \sum_{y} \frac{N_r(x, y)}{1 + |x - y|}$$

$$Correlation = \frac{\sum_{x} \sum_{y} N_r(x, y) xy - \mu_x \mu_y}{\sigma_x \sigma_y}$$



Features from Normalized Cooccurrence Matrix

$$Correlation = \frac{\sum_{x} \sum_{y} (x - \mu_{x}) (y - \mu_{y}) N_{r}(x, y)}{\sigma_{x} \sigma_{y}}$$

Mean and s.d. of row sums

$$f(x) = \sum_{v} N_r(x, y) \qquad g(y) = \sum_{x} N_r(x, y)$$

Mean and s.d. of column sums

$$g(y) = \sum_{x} N_r(x, y)$$

4

Laws' texture energy features

A set of 9 5x5 masks used to compute texture energy.

L5 (Level): [1 4 6 4 1]^T

E5 (Edge): $[-1 -2 \ 0 \ 2 \ 1]^T$

S5 (Spot): $[-1 \ 0 \ 2 \ 0 \ -1]^T$

R5 (ripple): [1 -4 6 -4 1]^T

Computation with mask: Convolution

A mask: Outer product of any pair. e.g. E5L5: E5.L5^T

$$\begin{bmatrix} -1 \\ -2 \\ 0 \\ 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

K. Laws, "Rapid Texture Identification", in SPIE Vol. 238: Image Processing for Missile Guidance, 1980, pp. 376-380.

Laws' texture energy features

Take

average of

responses

of two

masks.

A set of 9 5x5 masks used to compute texture energy.

L5 (Level): [1 4 6 4 1]^T

E5 (Edge): [-1 -2 0 2 1]^T

S5 (Spot): [-1 0 2 0 -1]^T

R5 (ripple): $[1 -4 6 -4 1]^T$

16 such masks possible.

Combine a few pairs to make 9 masks.

L5E5 and E5L5

L5R5 and R5L5

E5S5 and S5E5

L5S5 and S5L5

E5R5 and R5E5

S5R5 and R5S5

S5S5

R5R5

E5E5

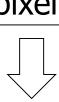
K. Laws, "Rapid Texture Identification", in SPIE Vol. 238: Image Processing for Missile Guidance, 1980, pp. 376-380.

Laws' texture energy

Input image

15x15 window

Subtract local average at each pixel.



Convolve with 16 masks.

L5E5 and E5L5 L5R5 and R5L5 E5S5 and S5E5 L5S5 and S5L5 E5R5 and R5E5 S5R5 and R5S5 S5S5 R5R5 E5E5

9 dimensional feature space.



Combine a few symmetric pairs to 9 channels.



Compute energy map for each channel.

Sum of absolute values in a 15x15 window.



Use of texture descriptors

- Detection of object patches represented by textured patterns.
- Segmentation of images.
- Classification / Matching
 - Generate a Library of labelled feature descriptors.
 - Detection of classes (class labels).
 - Matching to the nearest texture descriptor.

Image / Object Descriptor

- Bag of visual words
 - Compute key-point based feature descriptors over a library of images.

 K-means clustering
 - Quantize them (clustering) to form a finite set of representative descriptors (visual words).
 - For an image assign the nearest visual word corresponding to the feature descriptor of a key point.
 - Represent by each image by a histogram of visual words.

Sivic et. al., Discovering objects and their location in images, ICCV' 2005.

Vector of locally aggregated descriptors (VLAD)

- Form the codebook of visual words as in BoVW representation.
 - $C_1, C_2, ..., C_k$ Cluster centers of dimension D (say).
- Each local descriptor x in an image is associated to one of these visual words.
- Accumulate the differences w.r.t. the corresponding cluster center.

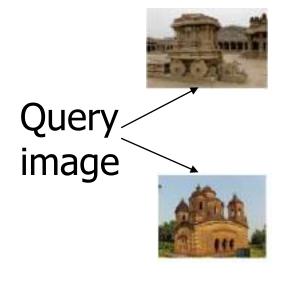
$$v_i = \sum_{x \text{ assigned to } C_i} (x - C_i)$$

 \circ VLAD descriptor=V/||V|| Dimension: k.D

Jegou et al., Aggregating local descriptors into a compact image representation, CVPR, 2010.



- Content based image retrieval
 - Image search based on visual content















Retrieved images from a database.















Summary of Techniques

- Scale and transformation invariant feature detection:
 - Harris corner-Laplacian Maximum.
 - DOG Maximum.
 - Intensity weighted FAST.

- Scale and transformation invariant Feature descriptor:
 - SIFT, SURF, ORB

Summary of Techniques

- Region and texture descriptors.
 - HoG
 - Edge density
 - LBP
 - Co-occurrence matrix
 - Laws' texture energy

- A few applications
 - Key point descriptor: Matching corresponding points of a scene.
 - Region descriptor:Object detection
 - Global descriptor :Image retrieval.
- Image global descriptor
 - BoVW
 - VLAD



