Scale & Affine Invariant Interest Point Detectors

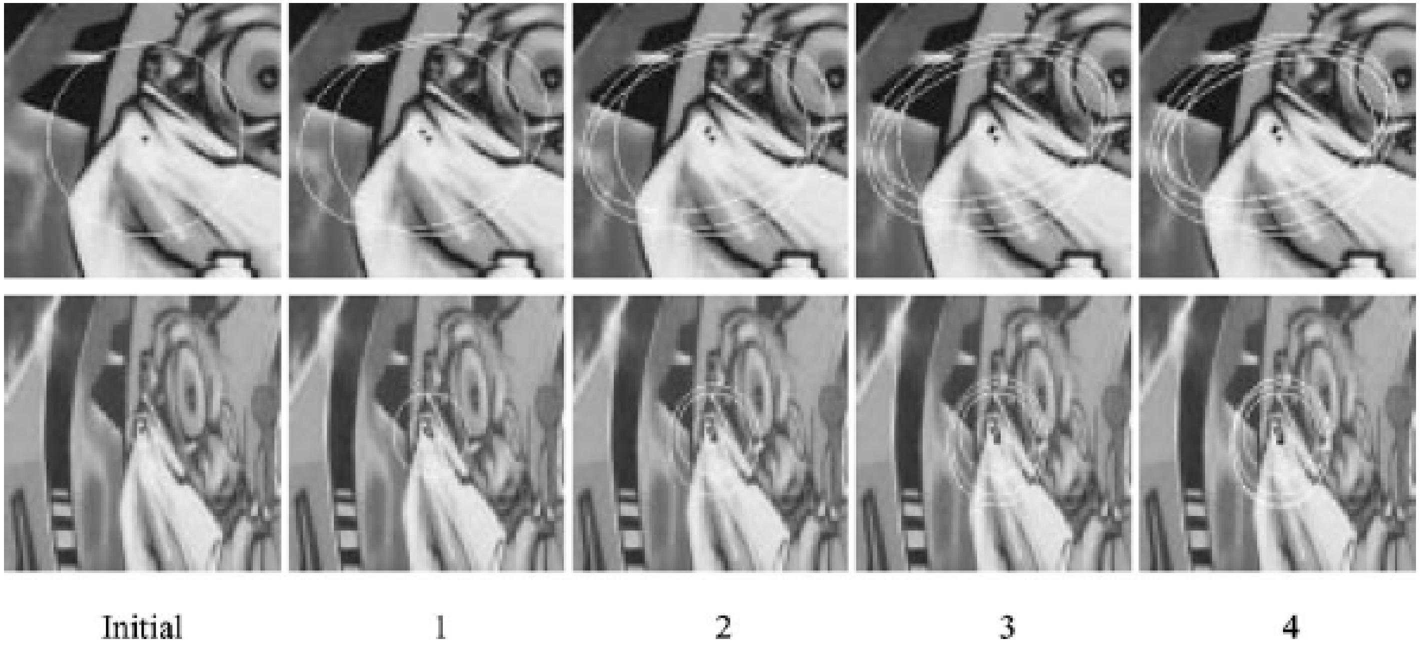
Mikolajczyk & Schmid

presented by Dustin Lennon

Paper Goal

- Combine Harris detector with Laplacian
  - Generate multi-scale Harris interest points
  - Maximize Laplacian measure over scale
  - Yields scale invariant detector
- Extend to affine invariant
  - Estimate affine shape of a point neighborhood via iterative algorithm

Visual Goal



Background/Introduction

- Basic idea #1:
  - scale invariance is equivalent to selecting points at characteristic scales
    - Laplacian measure is maximized over scale parameter
- Basic idea #2:
  - Affine shape comes from second moment matrix (Hessian)
    - Describes the curvature in the principle components

Background/Introduction

- Laplacian of Gaussian
  - Smoothing before differentiating
  - Both linear filters, order of application doesn’t matter
  - Kernel looks like a 3D mexican hat filter
  - Detects blob like structures
  - Why LoG: A second derivative is zero when the first derivative is maximized
- Difference of Gaussian
  - Subtract two successive smoothed images
  - Approximates the LoG

Background/Introduction

- But drawbacks because of detections along edges
  - unstable

- More sophisticated approach using penalized LoG and Hessian
  - Det, Tr are similarity invariant
  - Reduces to a consideration of the eigenvalues

Background/Introduction

- Affine Invariance
  - We allow a linear transform that scales along each principle direction
  - Earlier approaches (Alvarez & Morales) weren't so general
    - Connect the edge points, construct the perpendicular bisector
      - Assumes qualities about the corners
  - Claim is that previous affine invariant detectors are fundamentally flawed or generate spurious detected points

Scale Invariant Interest Points

- Scale Adapted Harris Detector

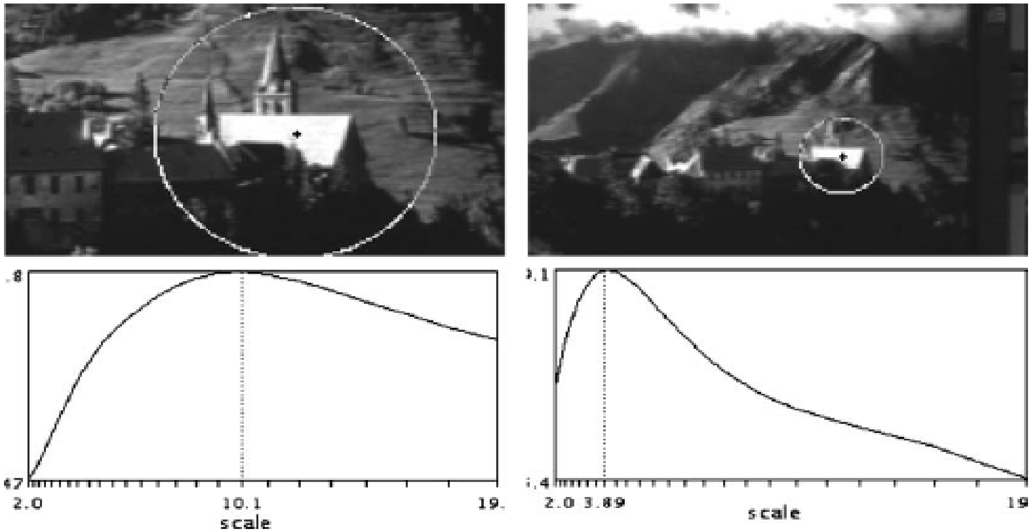
$$\mu(\mathbf{x}, \sigma_I, \sigma_D) = \begin{bmatrix} \mu_{11} & \mu_{12} \\ \mu_{21} & \mu_{22} \end{bmatrix}$$
$$= \sigma_D^2 \, g(\sigma_I) * \begin{bmatrix} L_x^2(\mathbf{x}, \sigma_D) & L_x L_y(\mathbf{x}, \sigma_D) \\ L_x L_y(\mathbf{x}, \sigma_D) & L_y^2(\mathbf{x}, \sigma_D) \end{bmatrix} \quad (1)$$

- Harris Measure

$$cornerness = \det(\mu(\mathbf{x}, \sigma_{\mathbf{I}}, \sigma_{\mathbf{D}})) - \alpha \operatorname{trace}^2(\mu(\mathbf{x}, \sigma_{\mathbf{I}}, \sigma_{\mathbf{D}})) \quad (2)$$

Characteristic Scale

- Sigma parameters
  - Associated with width of smoothing windows
  - At each spatial location, maximize LoG measure over scale
    - Characteristic scale
  - Ratio of scales corresponds to a scale factor between two images



Harris-Laplace Detector

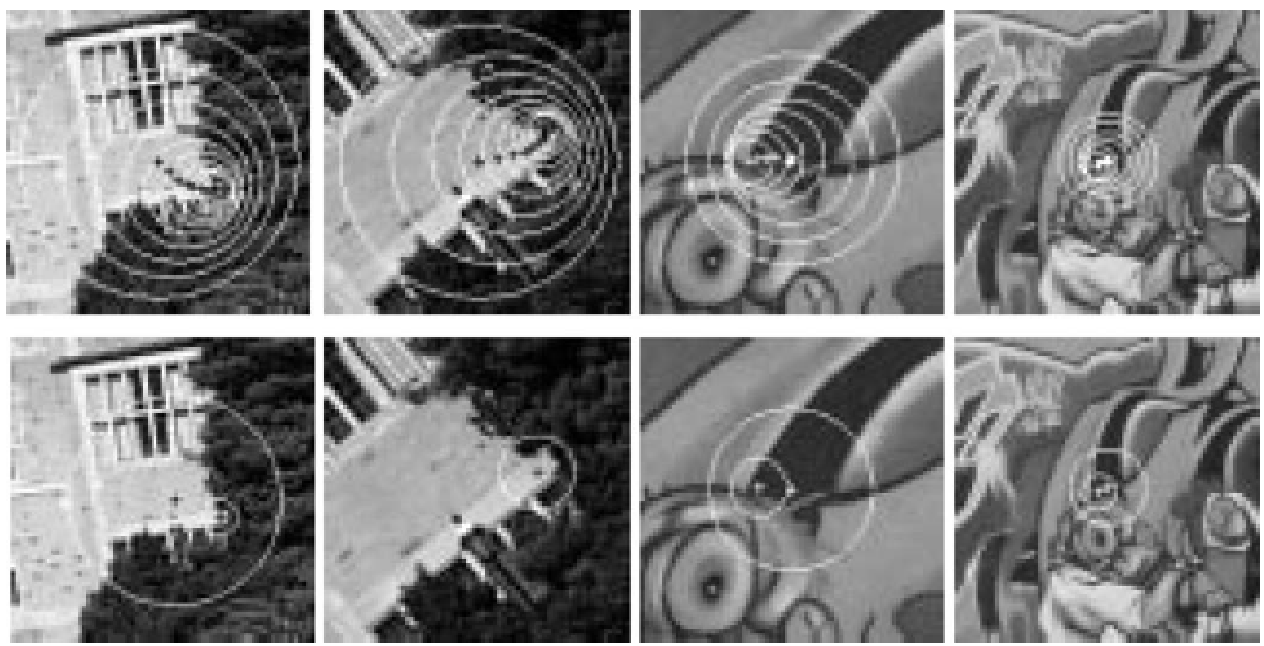
- Algorithm
  - Pre-select scales, sigma_n
  - Calculate (Harris) maxima about the point
    - threshold for small cornerness
  - Compute the matrix mu, for sigma_I = sigma_n
  - Iterate

1. Find the local extremum over scale of the LoG for the point $\mathbf{x}^{(k)}$, otherwise reject the point. The investigated range of scales is limited to $\sigma_I^{(k+1)} = t\sigma_I^{(k)}$ with $t \in [0.7, \ldots, 1.4]$.
2. Detect the spatial location $\mathbf{x}^{(k+1)}$ of a maximum of the Harris measure nearest to $\mathbf{x}^{(k)}$ for the selected σ_I^{k+1}.
3. Go to Step 1 if $\sigma_I^{(k+1)} \neq \sigma_I^{(k)}$ or $\mathbf{x}^{(k+1)} \neq \mathbf{x}^{(k)}$.

Harris-Laplace Detector

The authors claim that both scale and location converge. An example is shown below.

1. Find the local extremum over scale of the LoG for the point $\mathbf{x}^{(k)}$, otherwise reject the point. The investigated range of scales is limited to $\sigma_I^{(k+1)} = t\sigma_I^{(k)}$ with $t \in [0.7, \ldots, 1.4]$.
2. Detect the spatial location $\mathbf{x}^{(k+1)}$ of a maximum of the Harris measure nearest to $\mathbf{x}^{(k)}$ for the selected σ_I^{k+1}.
3. Go to Step 1 if $\sigma_I^{(k+1)} \neq \sigma_I^{(k)}$ or $\mathbf{x}^{(k+1)} \neq \mathbf{x}^{(k)}$.



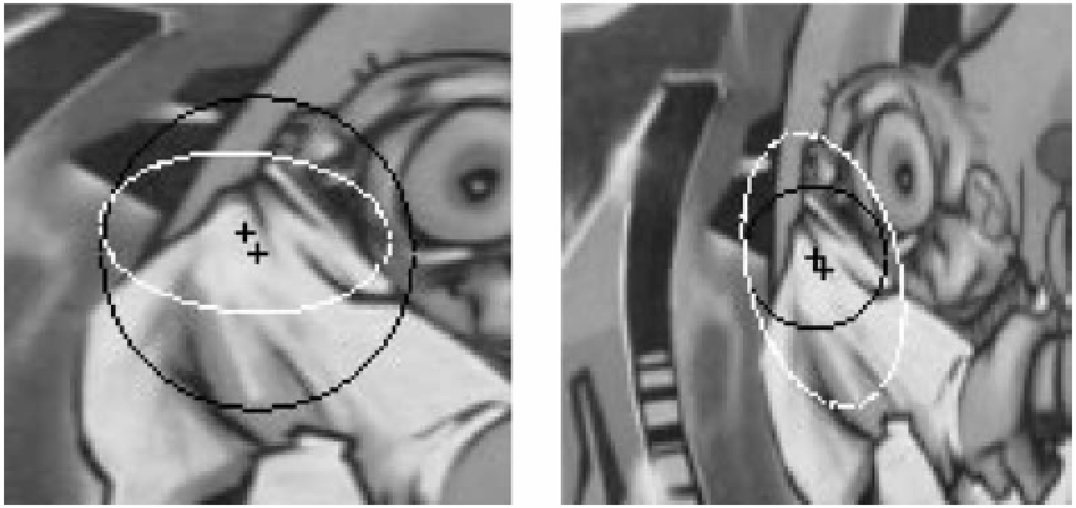
Harris Laplace

- A faster, but less accurate algorithm is also available.

- Questions about Harris Laplace
  - What about textured/fractal areas?
    - Kadir’s entropy based method
  - Local structures over a wide range of scales?
    - In contrast to Kadir?

Affine Invariance

- Need to generalize uniform scale changes
- Fig 3 exhibits this problem



Affine Invariance

The authors develop an affine invariant version of mu:

$$\mu(\mathbf{x}, \Sigma_I, \Sigma_D) = \det(\Sigma_D)\, g(\Sigma_I) * ((\nabla L)(\mathbf{x}, \Sigma_D)(\nabla L)(\mathbf{x}, \Sigma_D)^T) \quad (4)$$

Here Sigma represents covariance matrix for integration/differentiation Gaussian kernels

The matrix is a Hermitian operator.

To restrict search space, let Sigma_I, Sigma_D be proportional.

Affine Transformation

- Mu is transformed by an affine transformation of x: $\mathbf{x}_R = A\mathbf{x}_L.$

$$\mu(\mathbf{x}_L, \Sigma_{I,L}, \Sigma_{D,L}) = A^T \mu(\mathbf{x}_R, \Sigma_{I,R}, \Sigma_{D,R}) A$$
$$= A^T \mu(A\mathbf{x}_L, A\Sigma_{I,L}A^T, A\Sigma_{D,L}A^T) A \quad (5)$$

If we denote the corresponding matrices by:

$$\mu(\mathbf{x}_L, \Sigma_{I,L}, \Sigma_{D,L}) = M_L \quad \mu(\mathbf{x}_R, \Sigma_{I,R}, \Sigma_{D,R}) = M_R$$

these matrices are then related by:

$$M_L = A^T M_R A \quad M_R = A^{-T} M_L A^{-1} \quad (6)$$

In this case the differentiation and integration kernels are transformed by:

$$\Sigma_R = A \Sigma_L A^T$$

Let us suppose that the matrix M_L is computed in such a way that:

$$\Sigma_{I,L} = \sigma_I M_L^{-1} \quad \Sigma_{D,L} = \sigma_D M_L^{-1} \quad (7)$$

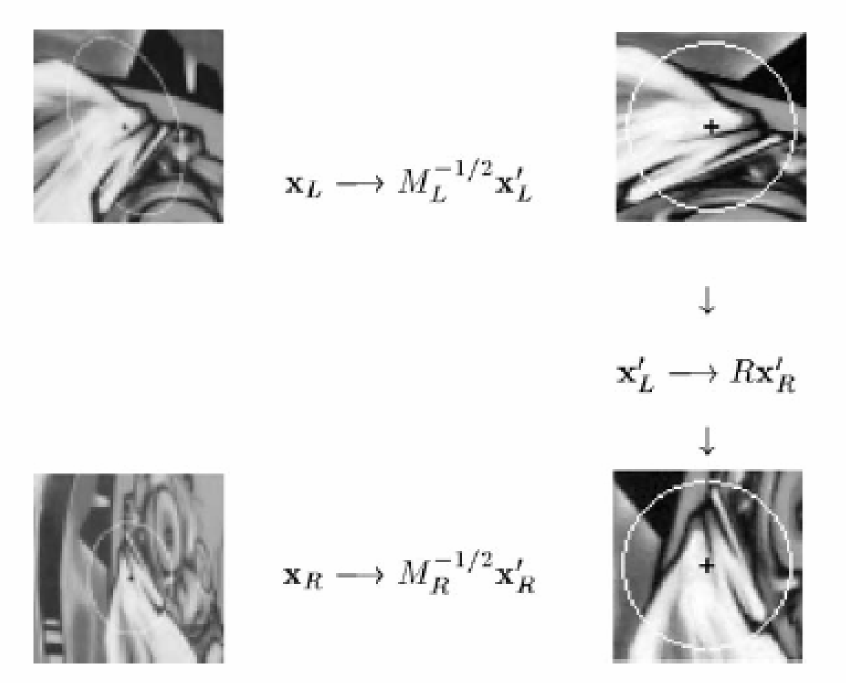
$$\begin{aligned} \Sigma_{I,R} &= A\Sigma_{I,L}A^T = \sigma_I \left(A M_L^{-1} A^T\right) \\ &= \sigma_I (A^{-T} M_L A^{-1})^{-1} = \sigma_I M_R^{-1} \\ \Sigma_{D,R} &= A\Sigma_{D,L}A^T = \sigma_D \left(A M_L^{-1} A^T\right) \\ &= \sigma_D (A^{-T} M_L A^{-1})^{-1} = \sigma_D M_R^{-1} \end{aligned} \quad (8)$$

$$A = M_R^{-1/2} R M_L^{1/2}$$

$$\mathbf{x}_R = A\mathbf{x}_L = M_R^{-1/2} R M_L^{1/2} \mathbf{x}_L,$$
$$M_R^{1/2} \mathbf{x}_R = R M_L^{1/2} \mathbf{x}_L$$

Affine Invariance

- Lots of math, simple idea
- We just estimate the Sigma covariance matrices, and the problem reduces to a rotation only
  - Recovered by gradient orientation



Isotropy

- If we consider mu as a Hessian, its eigenvalues are related to the curvature

$$Q = \frac{\lambda_{\min}(\mu)}{\lambda_{\max}(\mu)} \qquad (10)$$

- We choose sigma_D to maximize this isotropy measure.
- Iteratively approach a situation where Harris-Laplace (not affine) will work

Harris Affine Detector

- Spatial Localization
  - Local maximum of the Harris function
- Integration scale
  - Selected at extremum over scale of Laplacian
- Differentiation scale
  - Selected at maximum of isotropy measure
- Shape Adaptation Matrix
  - Estimated by the second moment matrix

Shape Adaptation Matrix

- Iteratively update the mu matrix by successive square roots
  - Keep max eigenvalue = 1
  - Square root operation forces min eigenvalue to converge to 1
  - Image is enlarged in direction corresponding to minimum eigenvalue at each iteration

Integration/Differentiation Scale

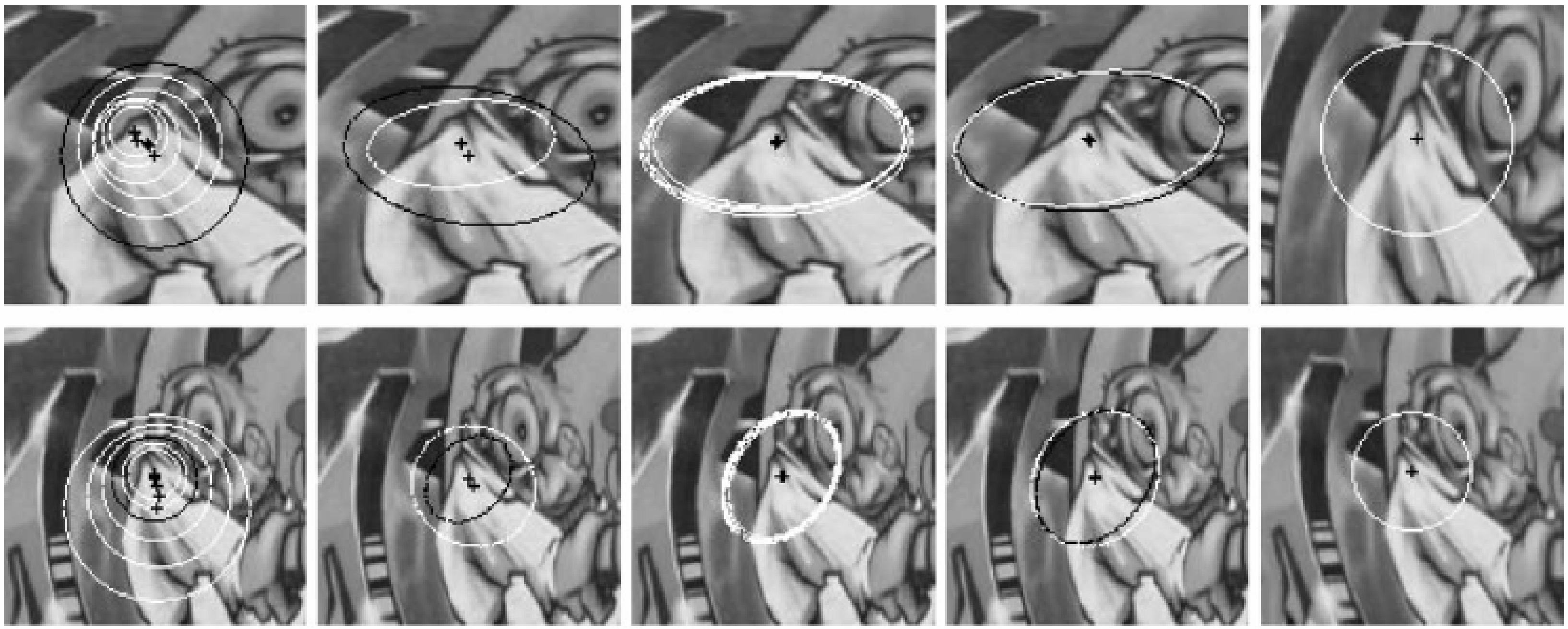
- Shape Adaptation means
  - only need sigmas corresponding to the Harris-Laplace (non affine) case.
    - Use LoG and Isotropy measure

- Well defined convergence criterion in terms of eigenvalues

Detection Algorithm

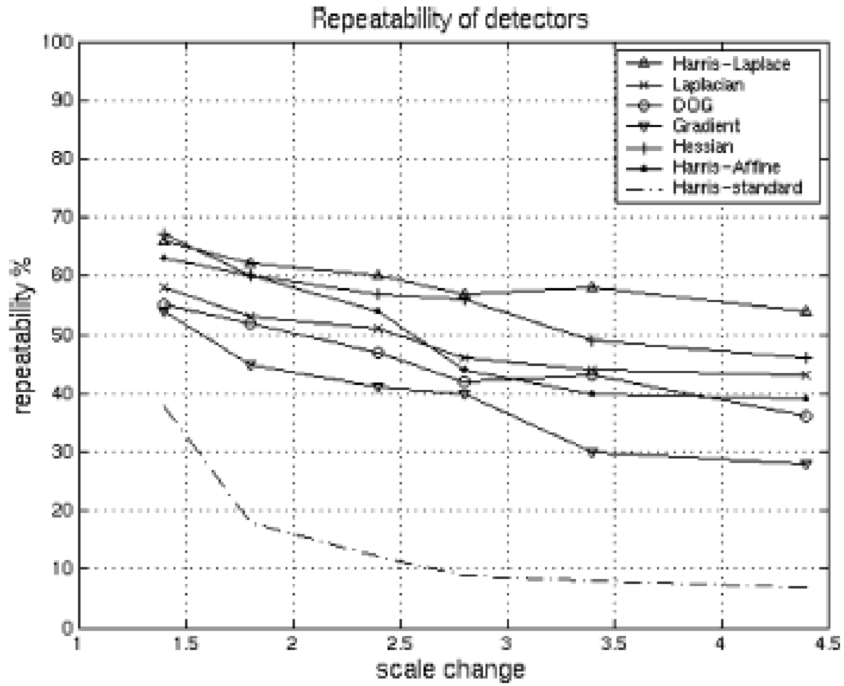
1. initialize $U^{(0)}$ to the identity matrix
2. normalize window $W(\mathbf{x}_w) = I(\mathbf{x})$ centered on $U^{(k-1)}\mathbf{x}_w^{(k-1)} = \mathbf{x}^{(k-1)}$
3. select *integration scale* σ_I at point $\mathbf{x}_w^{(k-1)}$
4. select *differentiation scale* $\sigma_D = s\sigma_I$, which maximizes $\frac{\lambda_{\min}(\mu)}{\lambda_{\max}(\mu)}$, with $s \in [0.5, \ldots, 0.75]$ and $\mu = \mu(\mathbf{x}_w{}^{(k-1)}, \sigma_I, \sigma_D)$
5. detect *spatial localization* $\mathbf{x}_w^{(k)}$ of a maximum of the Harris measure (Eq. (2)) nearest to $\mathbf{x}_w^{(k-1)}$ and compute the location of the interest point $\mathbf{x}^{(k)}$
6. compute $\mu_i^{(k)} = \mu^{-\frac{1}{2}}(\mathbf{x}_w^{(k)}, \sigma_I, \sigma_D)$
7. concatenate transformation $U^{(k)} = \mu_i^{(k)} \cdot U^{(k-1)}$ and normalize $U^{(k)}$ to $\lambda_{\max}(U^{(k)}) = 1$
8. go to Step 2 if $1 - \lambda_{\min}(\mu_i^{(k)})/\lambda_{\max}(\mu_i^{(k)}) \geq \epsilon_C$

Detection of Affine Invariant Points

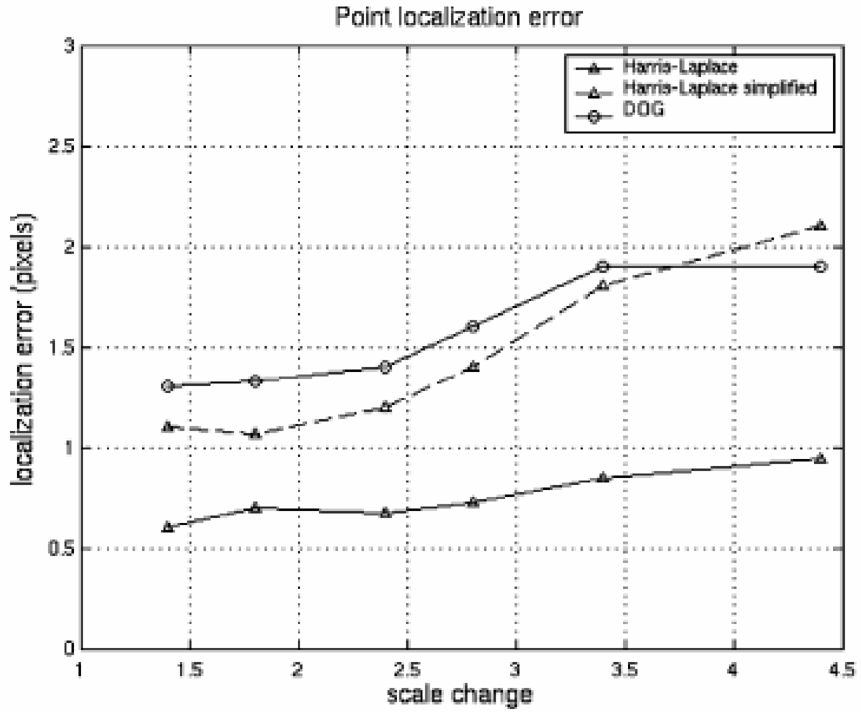


(a) (b) (c) (d) (e)

Results/Repeatability

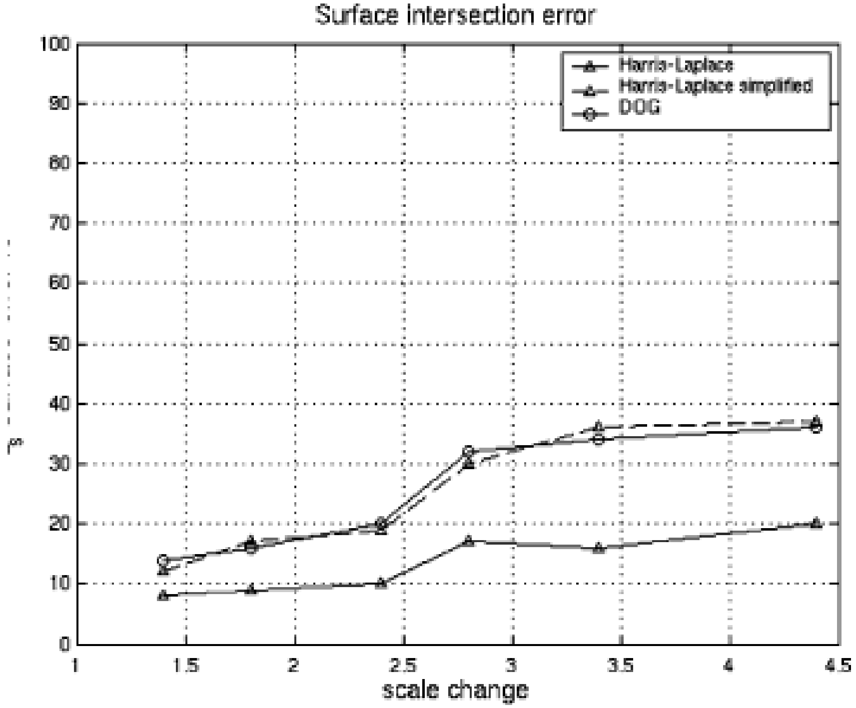


Results/Point Localization Error



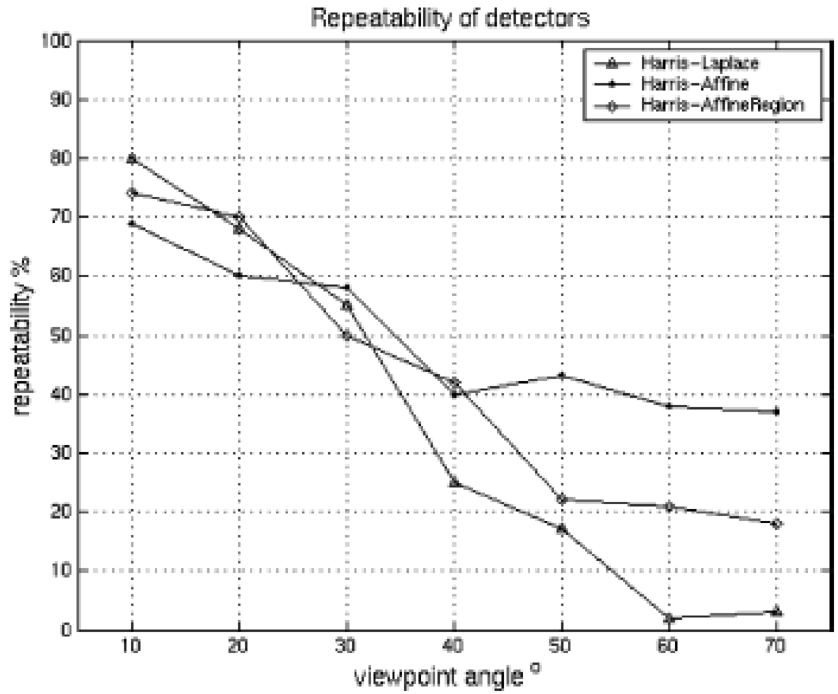
(a)

Results/Surface Intersection Error

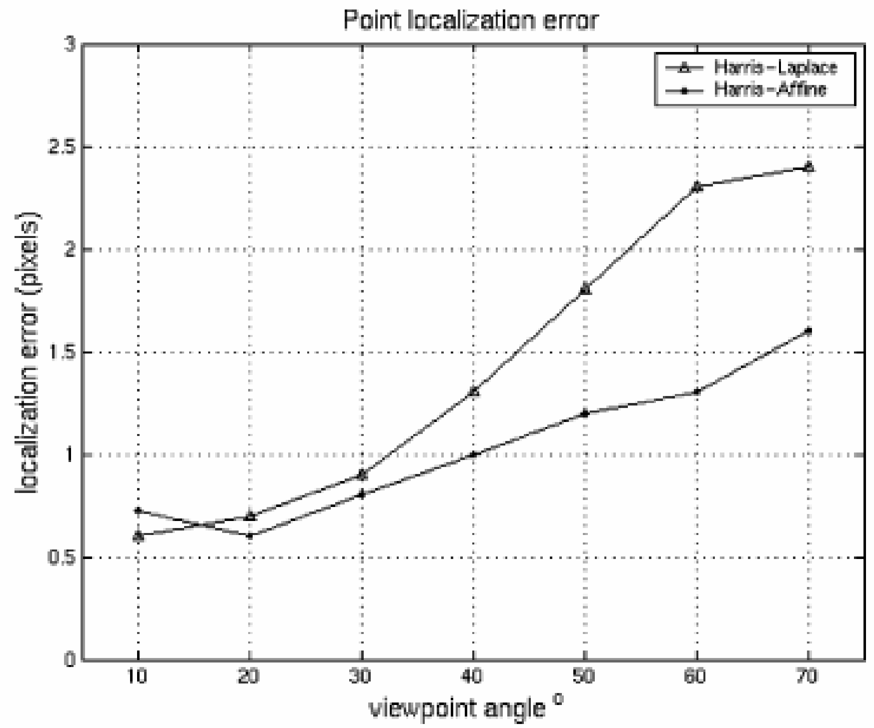


(b)

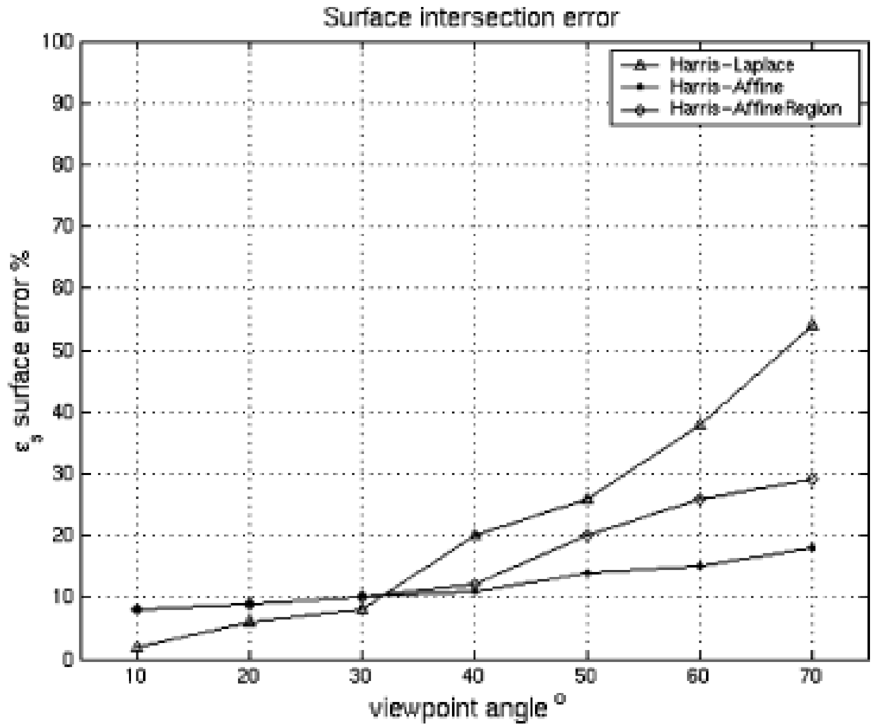
Results/Repeatability



Point Localization Error

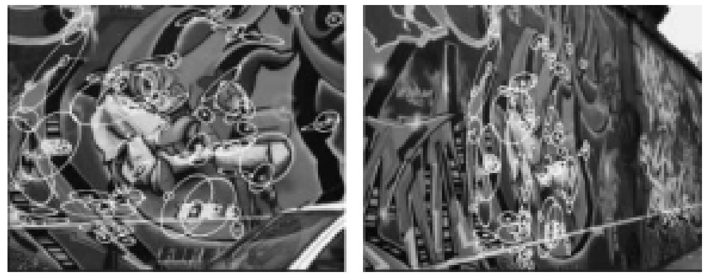


Surface Intersection Error

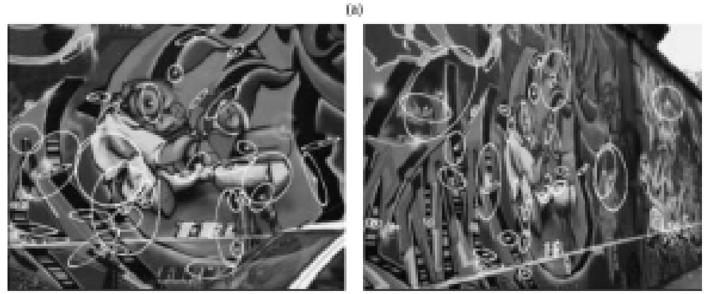


Applications

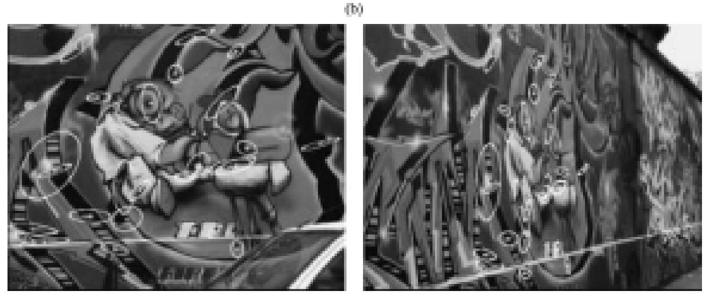
Applications



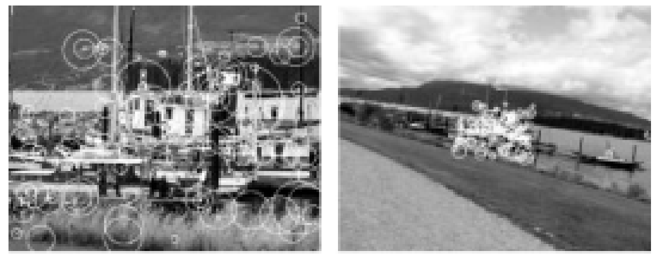
(a)



(b)



Applications



(a) Scale change of 3.9 and rotation of 17°.

(b) Scale change of 1.8 and viewpoint change of 30°

(c) Scale change of 1.7 and viewpoint change of 50°

Conclusions

- Results – impressive
- Methodology – reasonably well-justified
- Possible drawbacks?