Vision and SLAM

Speaker: Ricardo Vázquez Martín

Acción Integrada –'Visual-based interface for robots and intelligent environments'. Coimbra (Portugal) - 05/03/2009

INSTITUTO DE SISTEMAS E ROBÓTICA
UNIVERSIDADE DE COIMBRA

Ingeniería de Sistemas Integrados
Departamento de Tecnología Electrónica
Universidad de Málaga (Spain)

Contents

- Introduction
- Distinctive Features
- Hierarchical grouping algorithm
- Comparative study
- Problems in Visual SLAM

Contents

- Introduction
- Distinctive Features
- Hierarchical grouping algorithm
- Comparative study
- Problems in Visual SLAM

Motivation: Why vision?

- Vision systems are passive and of high resolution
- A huge amount of information (colour, texture or shape)
- Problems: a large amount of information, lighting, dynamic backgrounds and view-invariant matching

Contents

Distinctive Features

Interest Points

- Moravec (1977): intensity in a local window
- Harris (1988): local moment matrix
- Shi and Tomasi (1994): affine image transformation
- SUSAN (1997): no assumption about the local image structure
- FAST (2006): machine learning for fast corner detection

Distinctive Features

Interest Points

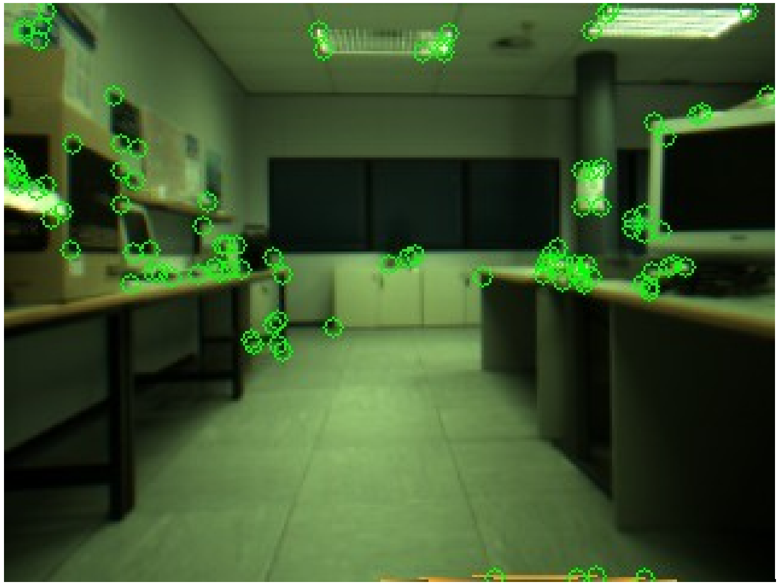
In the Harris street corner

Harris corners

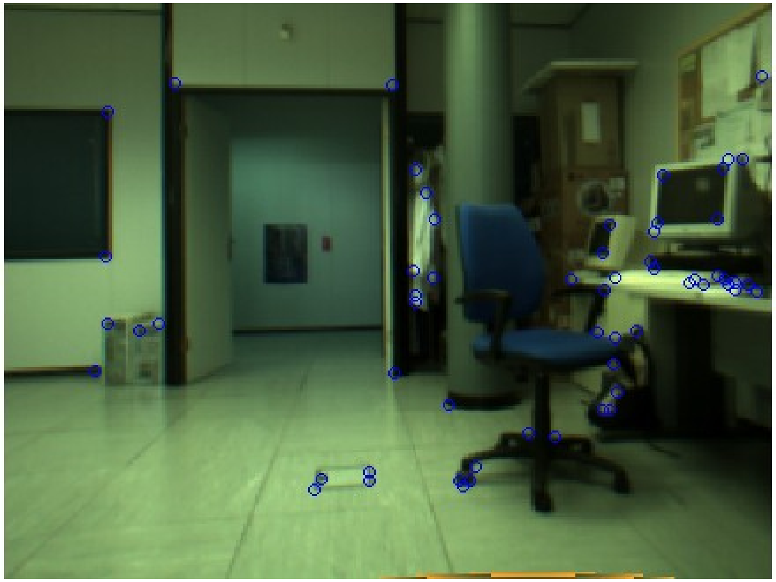
Distinctive Features

Interest Points

Harris corners



FAST



Distinctive Features

Interest Points - Advantages

- Salient in images
- Good invariance properties
- Low computation cost and very numerous

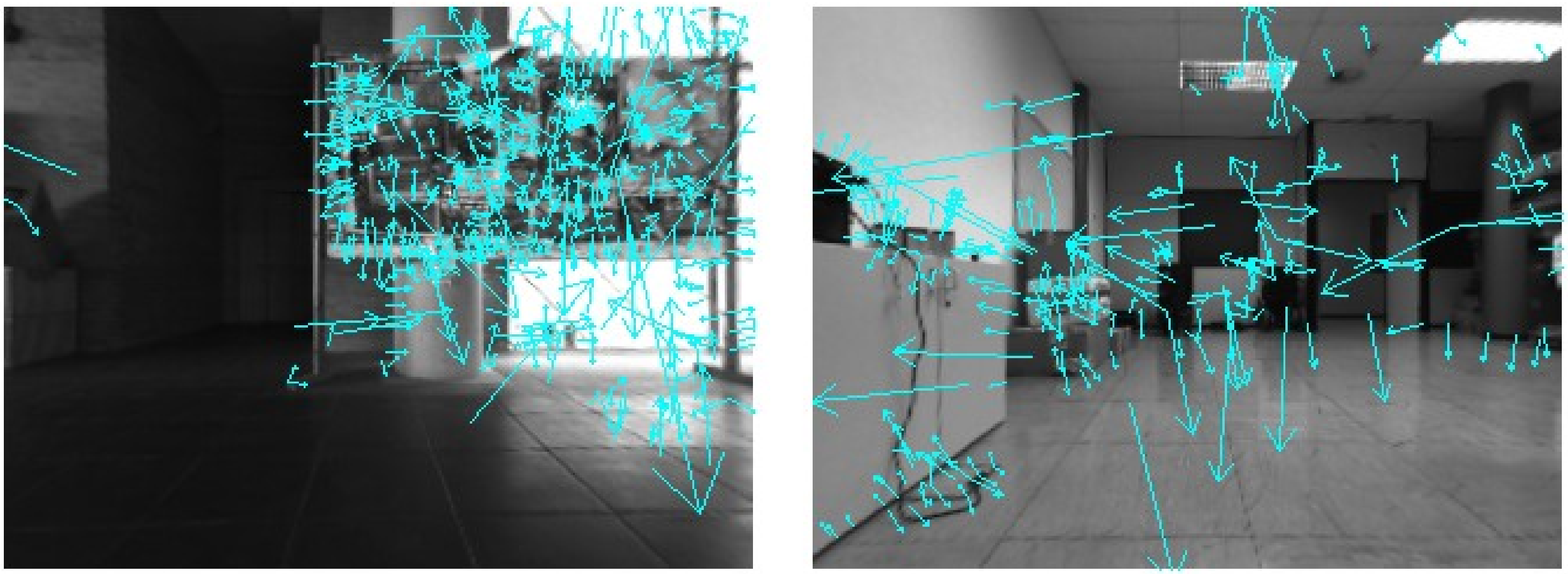
Interest Points - Drawbacks

- Repeatability steeply decrease with significant scale changes

Scale Invariant Features

- Harris-Laplacian (2001): Harris Corner + normalized laplacian to select scale
- Scale Invariant Feature Transform - SIFT (1999): Difference of Gaussians
- Speeded Up Robust Features - SURF (2008): Determinant of the Hessian matrix

Scale Invariant Features - SIFT



Distinctive Features

Affine Invariant Features

- Invariance under arbitrary viewing conditions: affinity
- Region shape must be adapted: covariant
- Pattern should be normalized to use an invariant feature descriptor
- Affine invariant construction method: second moment matrix or autocorrelation matrix

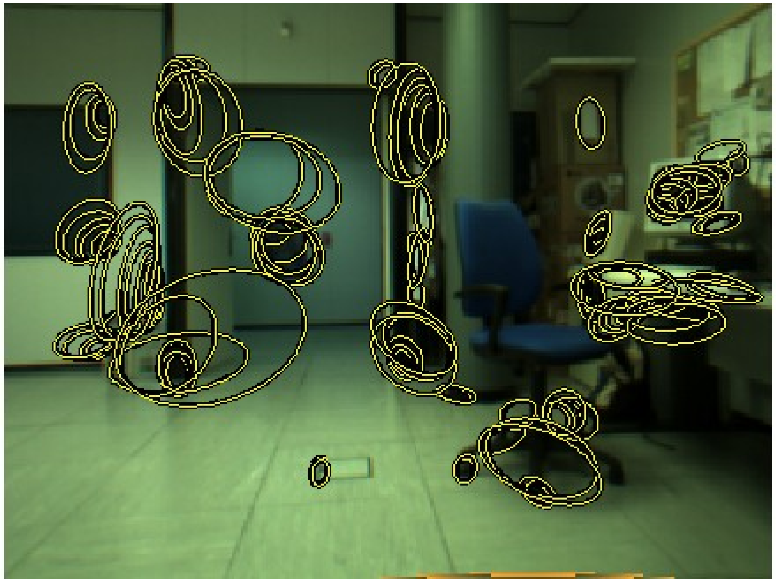
Distinctive Features

Affine Invariant Features

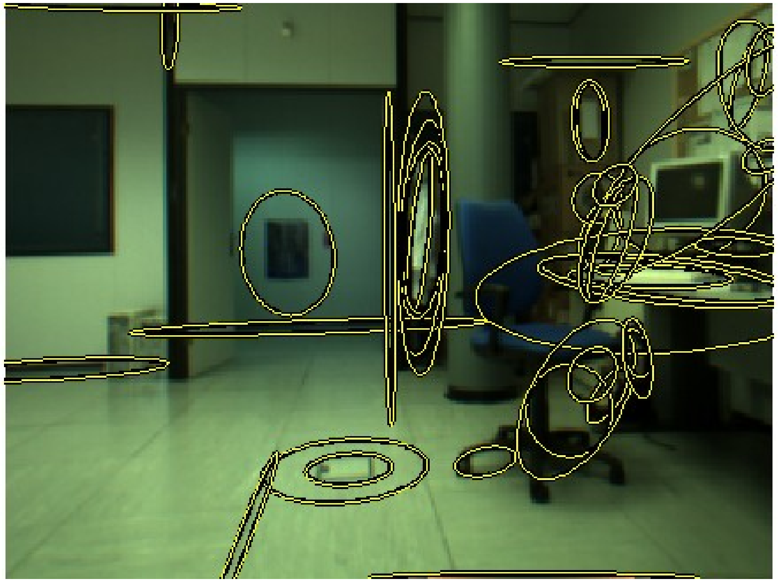
- Harris-Affine and Hessian-Affine (2006)
- Max. Stable Extremal Regions – MSER (2002)
- Intensity Extrema Based Region – IBR (2004)
- Edge Based Region detector – EBR (2004)
- Entropy Based Region detector – salient regions (2004)

Affine Invariant Features

Hessian Affine



MSER



Feature description - definition

- Detected features must be characterized to solve the correspondence problem

Feature description - descriptors

- Correlation window
- Invariant to scale and rotation: SIFT, PCA-SIFT, SURF, GLOH

Contents

- Introduction
- Distinctive Features
- Hierarchical grouping algorithm
- Comparative study
- Problems in Visual SLAM

Hierarchical grouping mechanism

Region detection in three stages

- Pre-segmentation: image is segmented in blobs of uniform colour
- Perceptual grouping: a smaller partition of the image is obtained merging blobs
- Visual feature detection and normalization: some constraints are imposed the set of obtained regions

Hierarchical grouping mechanism

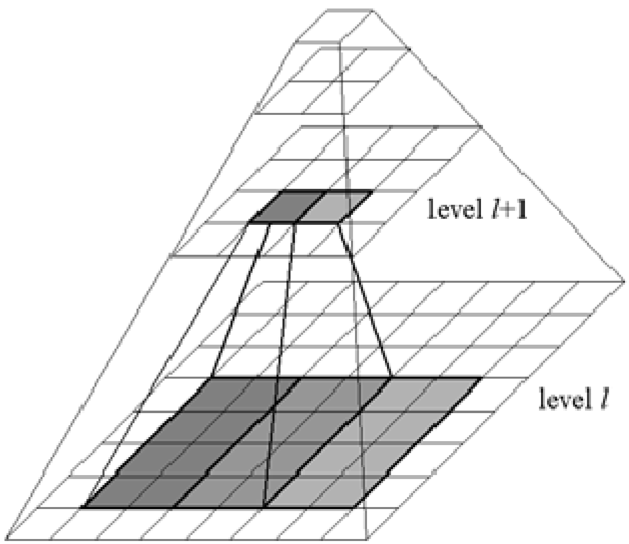
Pre-segmentation stage

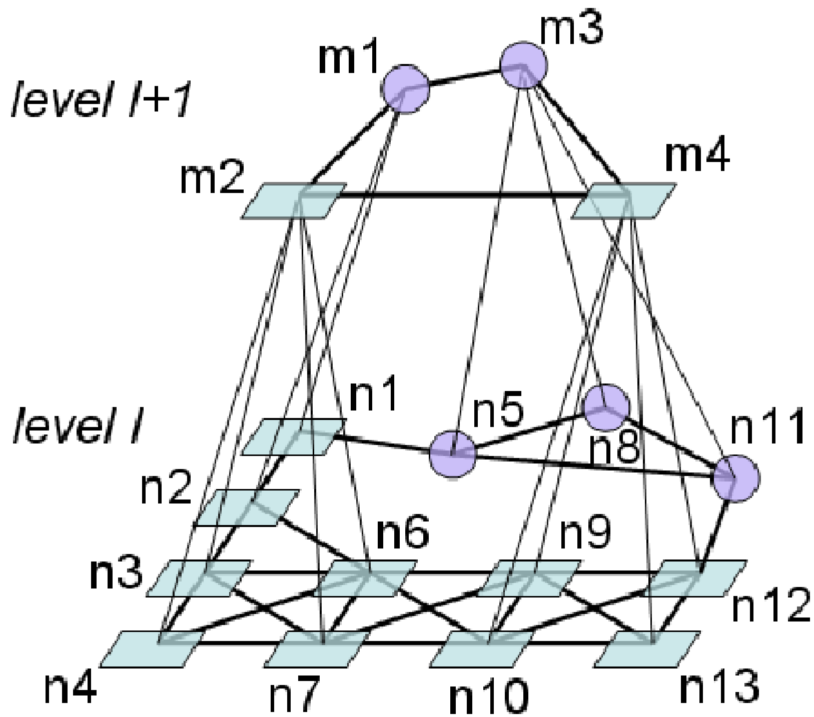
- Homogeneous regions: color
- Bounded Irregular Pyramid
- New decimation process to avoid the shift variance problem (uBIP)

•Marfil, R. et al, “Perception-based Image Segmentation Using the Bounded Irregular Pyramid”, Pattern Recognition Symposium 2007

Hierarchical grouping mechanism

Pre-segmentation stage





Hierarchical grouping mechanism

Perceptual grouping stage

- Pre-segmented blobs are grouped following a distance criterion
- Colour contrast and the shared boundary are used to simplify the image partition (and disparity in the stereo vision version)

$$\Upsilon(n_i, n_j) = \sqrt{w_1 \cdot \left(\frac{d(n_i, n_j) \cdot min(b_i, b_j)}{\alpha \cdot c_{ij} + \beta(b_{ij} - c_{ij})}\right)^2 + w_2 \cdot (disp(n_i) - disp(n_j))^2}$$

Hierarchical grouping mechanism

Feature detection and normalization

- Regions that fulfil some conditions are selected as features
- Area of a region: a percentage of the whole image
- Regions must not be in an image border
- High color contrast between a feature and its surrounding regions

Hierarchical grouping mechanism

Feature description – colour histogram

- Masked with a kernel to take into account spatial information

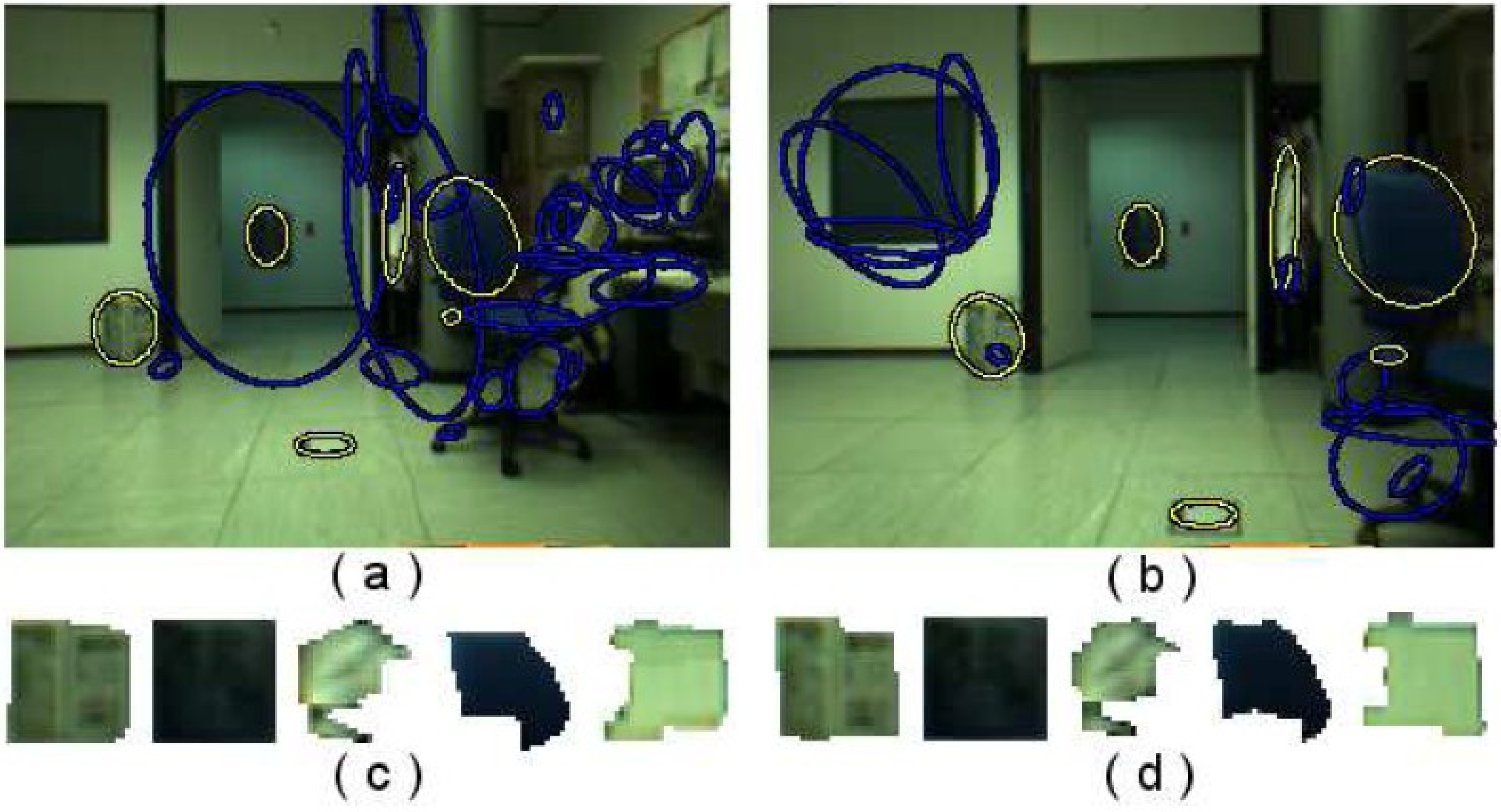
$$s_n = \frac{1}{\mathcal{C}} \sum_{(x,y)_i \in \zeta_v} \mathcal{N}((x,y)_i)\delta(\gamma[I((x,y)_i] - n)$$

- Similarity between regions using a metric derived from the Bhattacharyya coefficient

$$d(\mathbf{p},\mathbf{q}) = \sqrt{1 - \rho[\hat{\mathbf{p}},\hat{\mathbf{q}}]} \qquad \rho[\hat{\mathbf{p}},\hat{\mathbf{q}}] = \sum_{i=1}^{m} \sqrt{\hat{p}_i \cdot \hat{q}_i}$$

Hierarchical grouping mechanism

Experimental results



Hierarchical grouping mechanism

Experimental results

Hierarchical grouping mechanism

Experimental results

(c)

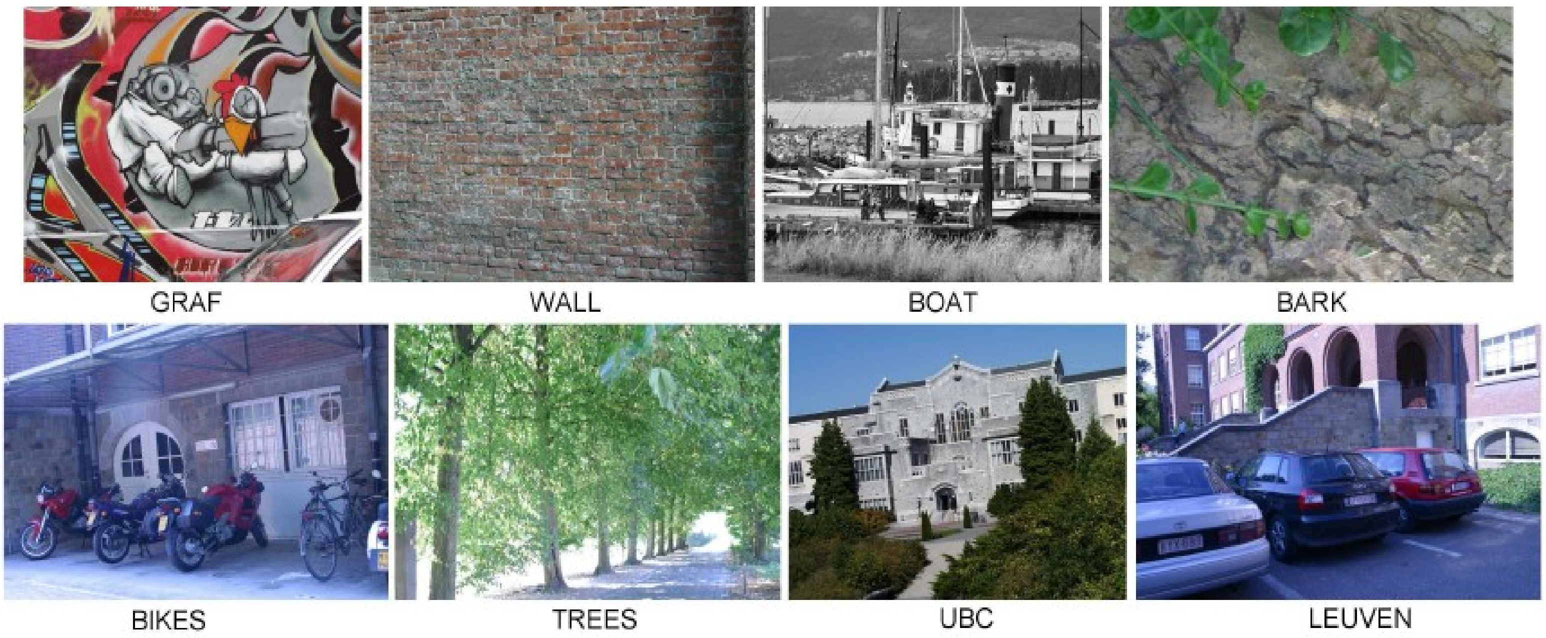
Hierarchical grouping mechanism

Conclusions

- Does not rely on the extraction of interest points features and on differential methods
- Affine region detector based in image intensity (colour) analysis
- mid-level segmentation coherent with the human-based image decomposition
- Features detected in scale-space with an underlying semantic significance

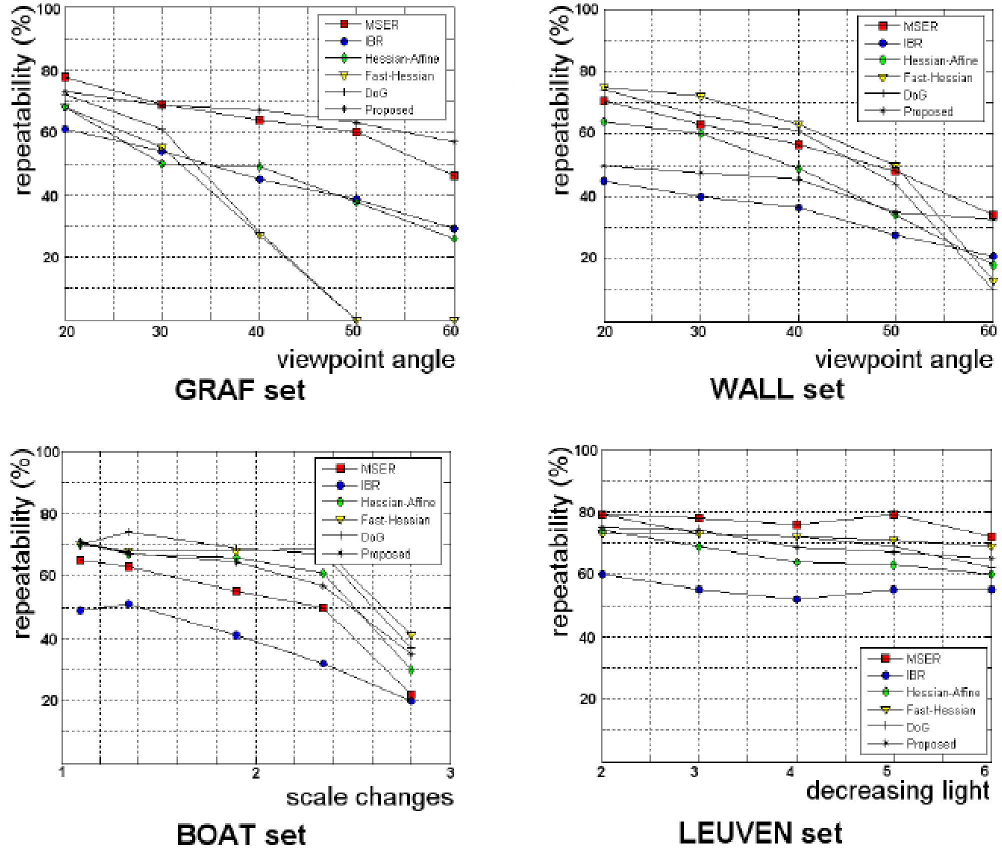
Contents

A comparative study: database



- http://www.robots.ox.ac.uk/~vgg/research/affine

A comparative study: Region detector

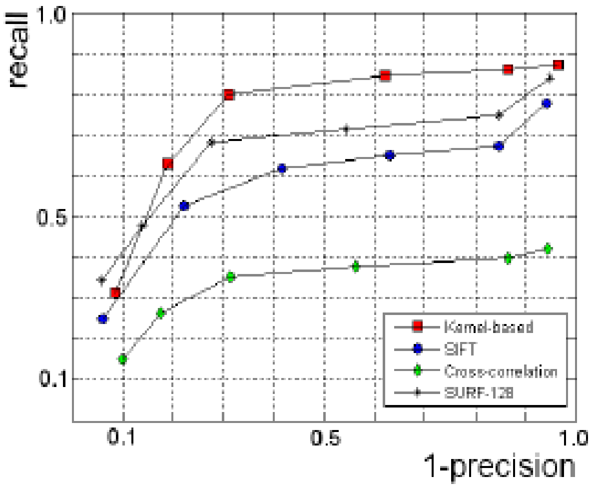


A comparative study: Computing times

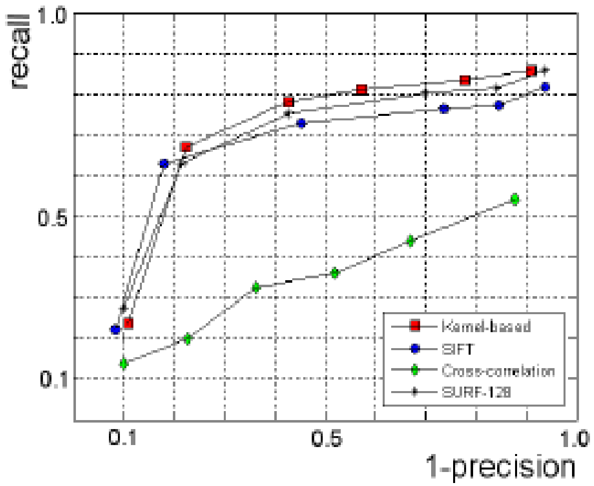
	Height	Regions	F	Q	SV	Time (sec)
LP	25.5	73.7	743.2	1011.5	30.2	2.75
MP	32.9	107.6	650.1	818.5	29.3	3.42
HP	11.4	76.1	670.3	955.1	28.4	4.23
CoP	74.2	91.2	630.7	870.2	30.5	2.85
BIP	8.7	83.6	720.2	1090.1	44.1	0.20
uBIP	9.3	60.5	700.1	950.3	24.3	0.23

detector	Number of regions	Run time (sec)
DoG	1520	0.39
Hessian-affine	1649	2.43
Fast-Hessian	1418	0.12
MSER	533	0.56
IBR	679	9.77
Proposed	147	0.32

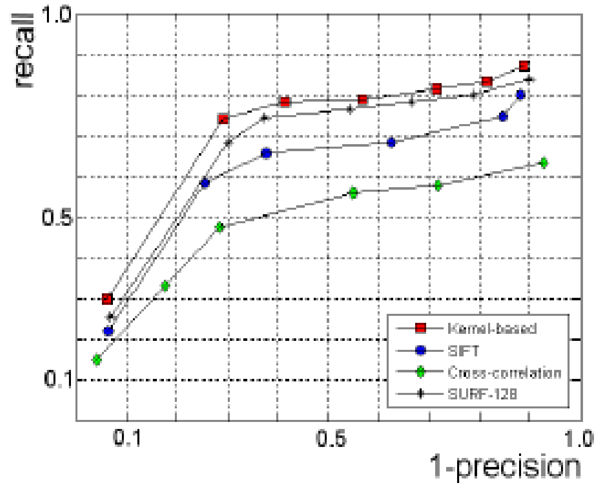
A comparative study: Descriptor



GRAF set



BIKES set



LEUVEN set

$$recall = \frac{\#\, correct\ matches}{\#\, correspondences} \qquad 1 - precision = \frac{\#\, false\ matches}{\#\, correct\ matches + \#\, false\ matches}$$

Contents

- Introduction
- Distinctive Features
- Hierarchical grouping algorithm
- Comparative study
- Problems in Visual SLAM

Problems in Visual SLAM

Monocular SLAM

- Javier Civera, Andrew J. Davison and J.M.M. Montiel " Inverse Depth Parametrization for Monocular SLAM“, IEEE Transactions on Robotics Vol 24(5) pp 932-945. October 2008

Problems in Visual SLAM

Monocular SLAM – using affine regions

Vision and SLAM

Speaker: Ricardo Vázquez Martín

Acci *s and*

- **Thanks for your attention!!**
- **Any questions/advise?**

Ingeniería de Sistemas Integrados
Departamento de Tecnología Electrónica
Universidad de Málaga (Spain)