

Confidence-based Cue Integration for Visual Place Recognition

Andrzej Pronobis

CAS/CVAP, KTH

Stockholm, Sweden

`pronobis@csc.kth.se`



Barbara Caputo

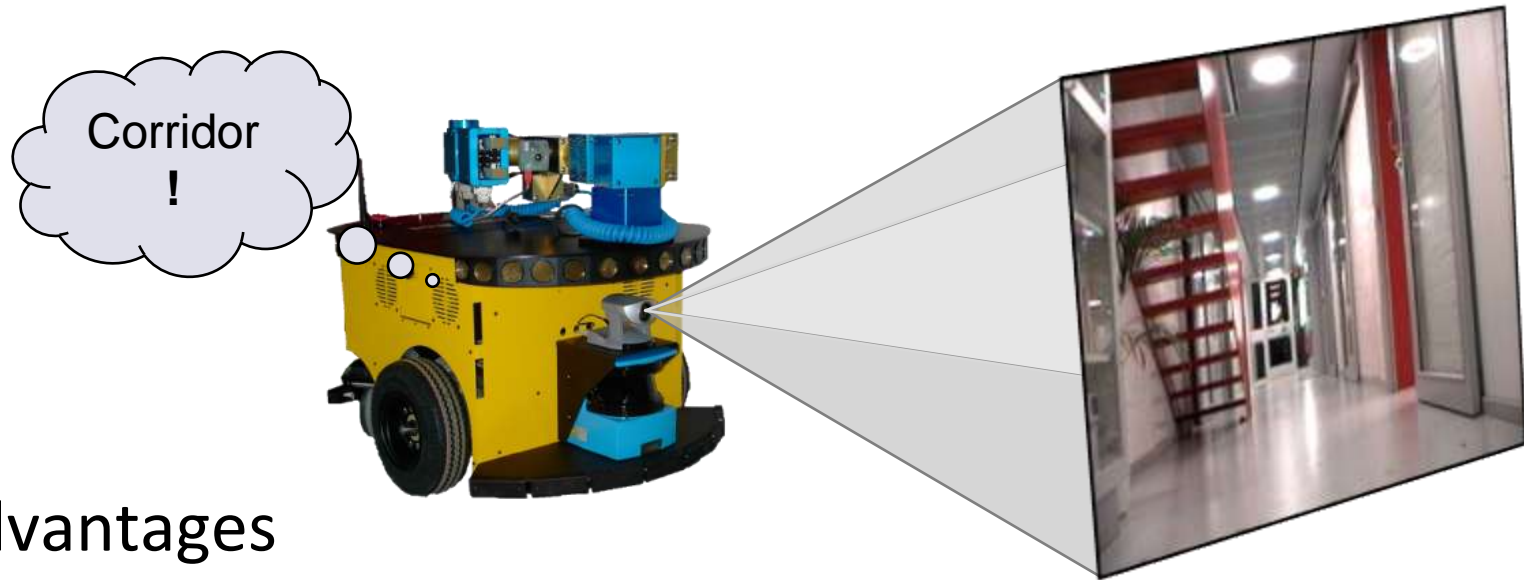
IDIAP, Martigny, Switzerland

EPFL, Lausanne, Switzerland

`bcaputo@idiap.ch`



Visual Place Recognition for Topological Localization



Advantages

- Rich sensor
 - Loop closing
 - Solving the kidnapped robot problem
- Source of contextual information
 - Semantics encoded in visual appearance
- Portable and cost-effective



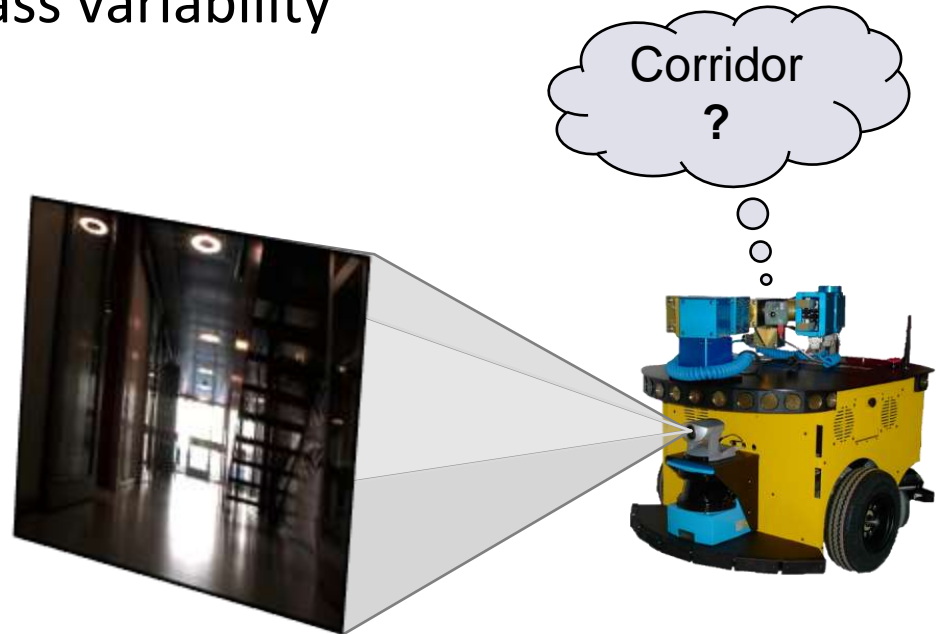
Challenges

□ Challenges

- Large amounts of information
- Noise and large within class variability
 - Illumination conditions
 - Human activity
 - Viewpoints
- Temporal variations

□ Possible solutions

- Information fusion
- Ignorance detection
- Adaptive algorithms





Contribution

- Visual place recognition system employing local and global features
- Confidence estimation and hypotheses ranking method for discriminative classifiers
- Confidence-based cue integration scheme
- Thorough experimental evaluation in the domain of mobile robot topological localization
 - Robustness to illumination changes, human activity, viewpoint variations
 - Real-time efficiency



Visual Place Recognition

Definition of the Problem

- Fully supervised approach [Pronobis et al. '06]
- Training:



- Recognition:





Experimental Setup

The IDOL Database





Experimental Setup

The IDOL Database

- Five rooms of different functionality



One-person office



Corridor



Two-persons office



Kitchen



Printer area

- Two different robot platforms

- Three illumination settings over three weeks



Cloudy



Sunny



Night





Experimental Procedure

- ❑ Experiment 1 – Stable illumination conditions
- ❑ Experiment 2 – Varying illumination conditions
- ❑ Experiment 3 – Recognition across platforms

Training



Testing





Experimental Procedure

- Experiment 1 – Stable illumination conditions
- Experiment 2 – Varying illumination conditions
- Experiment 3 – Recognition across platforms

Training



Testing

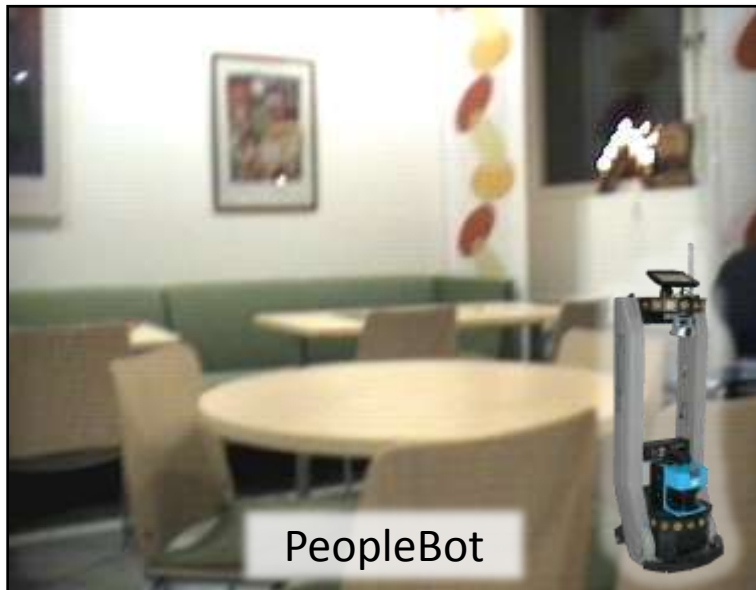




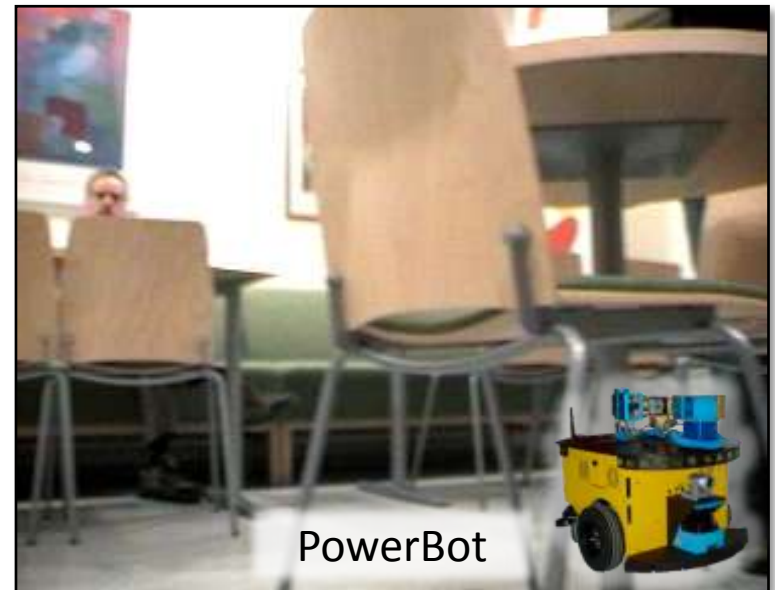
Experimental Procedure

- Experiment 1 – Stable illumination conditions
- Experiment 2 – Varying illumination conditions
- Experiment 3 – Recognition across platforms

Training

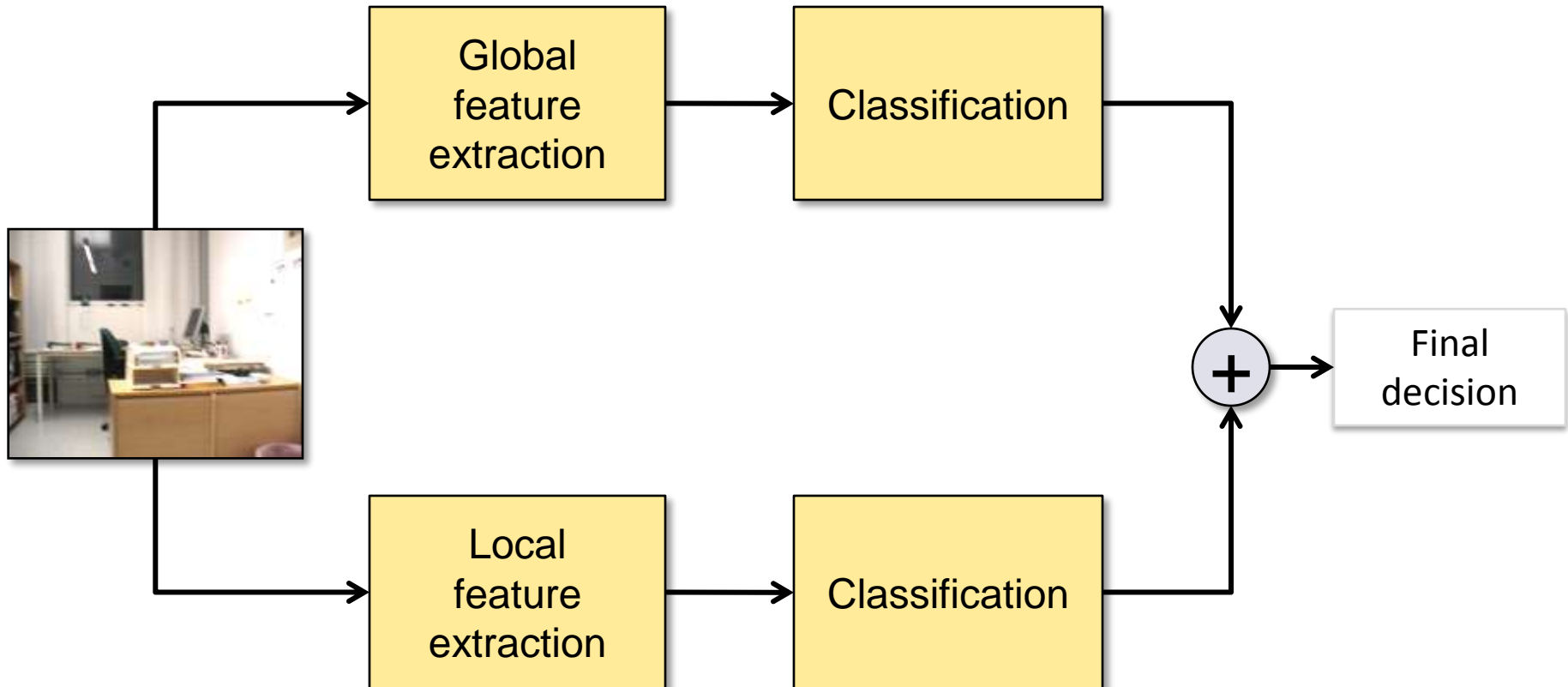


Testing





The Visual Place Recognition System Overview

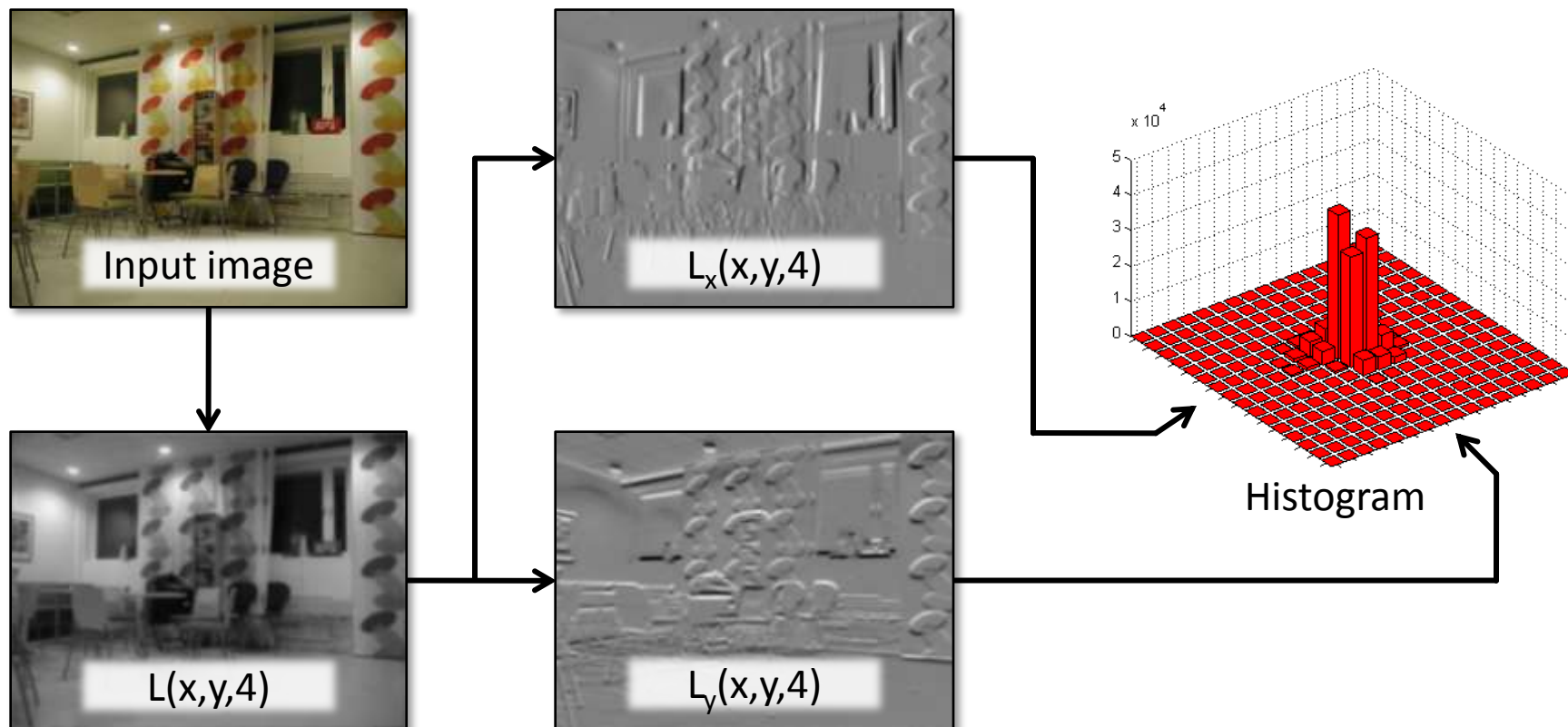




The Visual Place Recognition System

Feature Extraction – Global Features

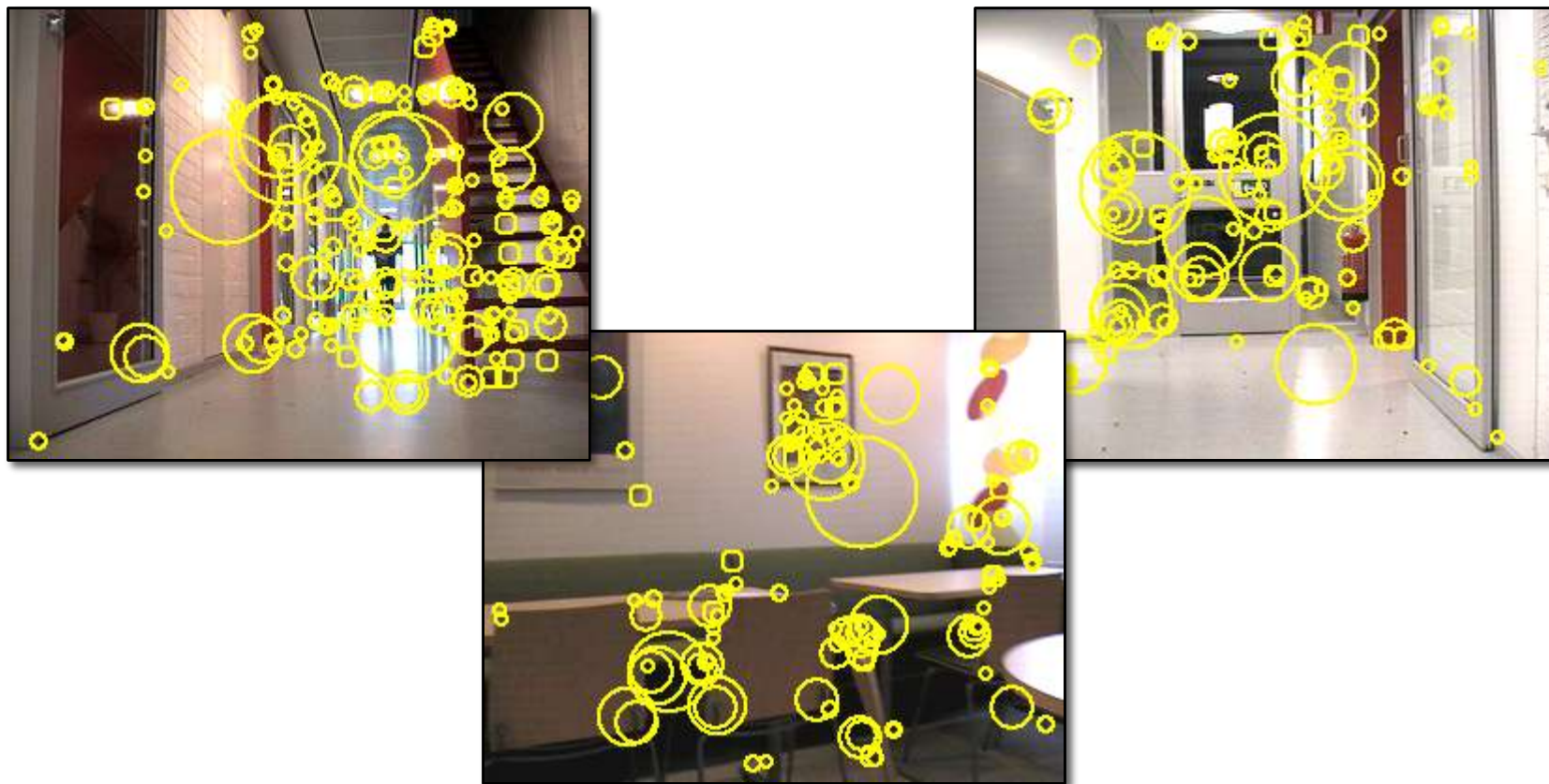
- High dimensional Composed Receptive Field Histograms (CRFH) [Linde & Lideberg '04]



The Visual Place Recognition System

Feature Extraction – Local Features

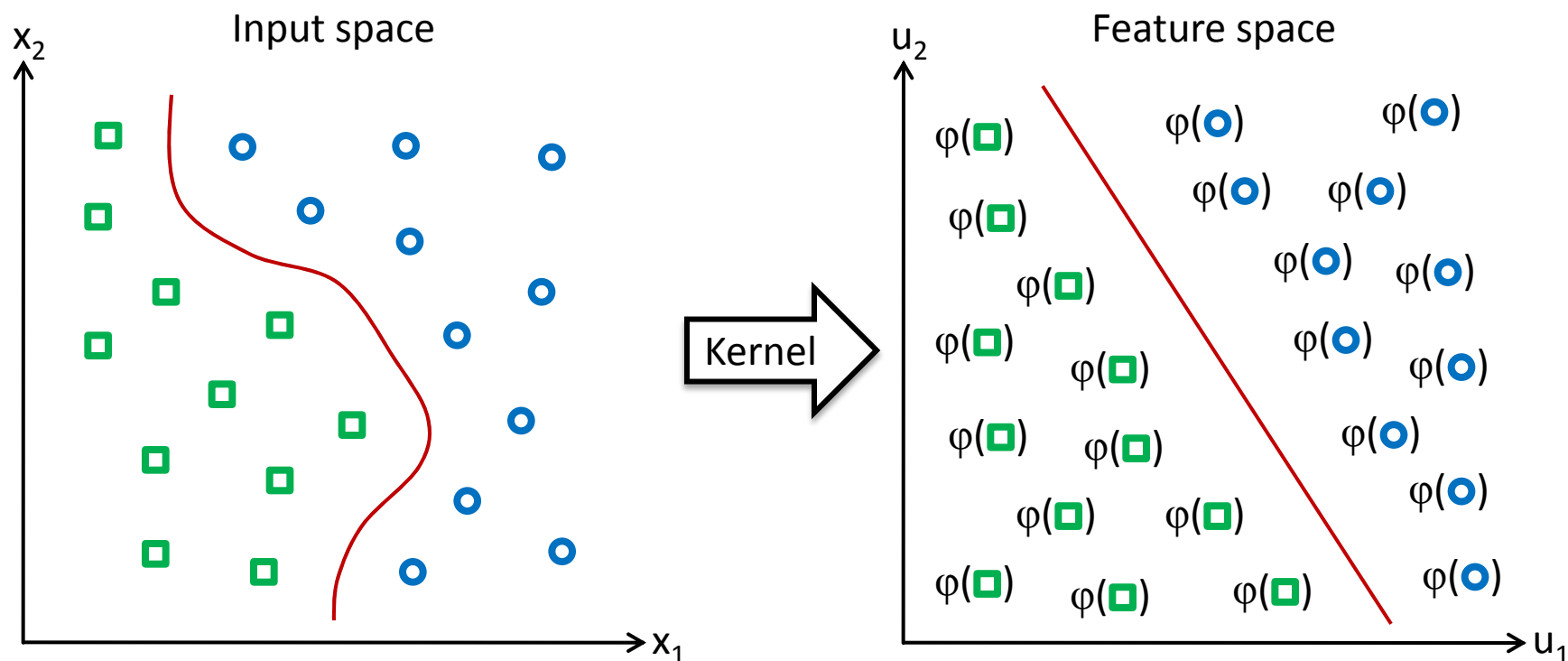
- Harris-Laplace interest-point detector [Mikolajczyk & Schmid '01] and SIFT descriptor [Lowe '04]





The Visual Place Recognition System Classification

Support Vector Machines [Cristianini & Taylor '99]



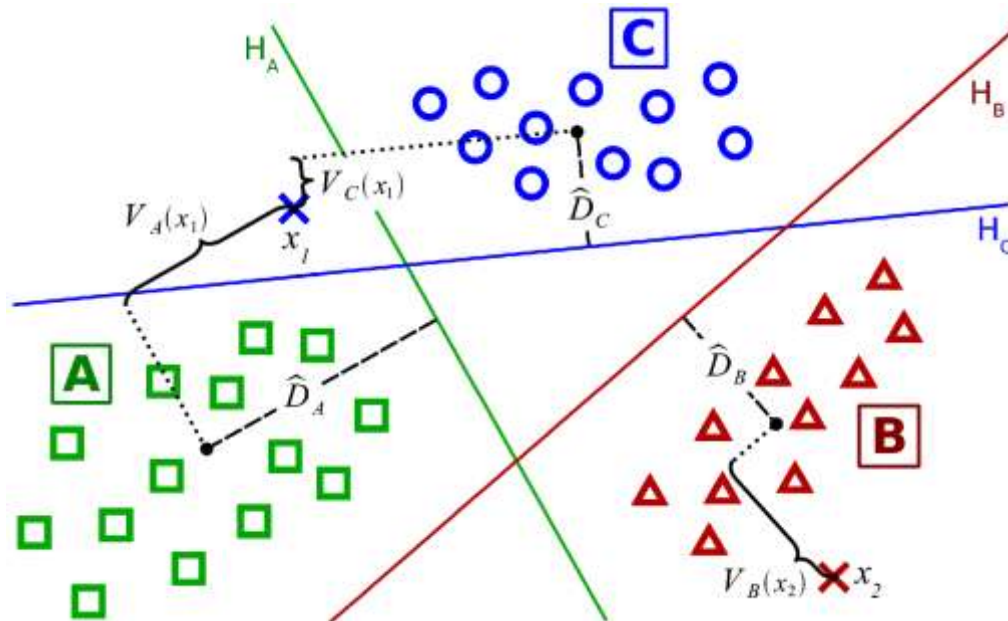
Two standard multi-class extensions: O-a-A, O-a-O

Superior generalization, no confidence information



Confidence Estimation and Hypotheses Ranking

- Confidence information and hypotheses ranking derived from distances between samples and hyperplanes
- Solution based on the one-against-all principle



Scores:

$$V_j(x) = \left| \hat{D}_j - D_j(x) \right|$$

Best hypothesis:

$$j^* = \arg \min_{j=1 \dots M} \{V_j(x)\}$$

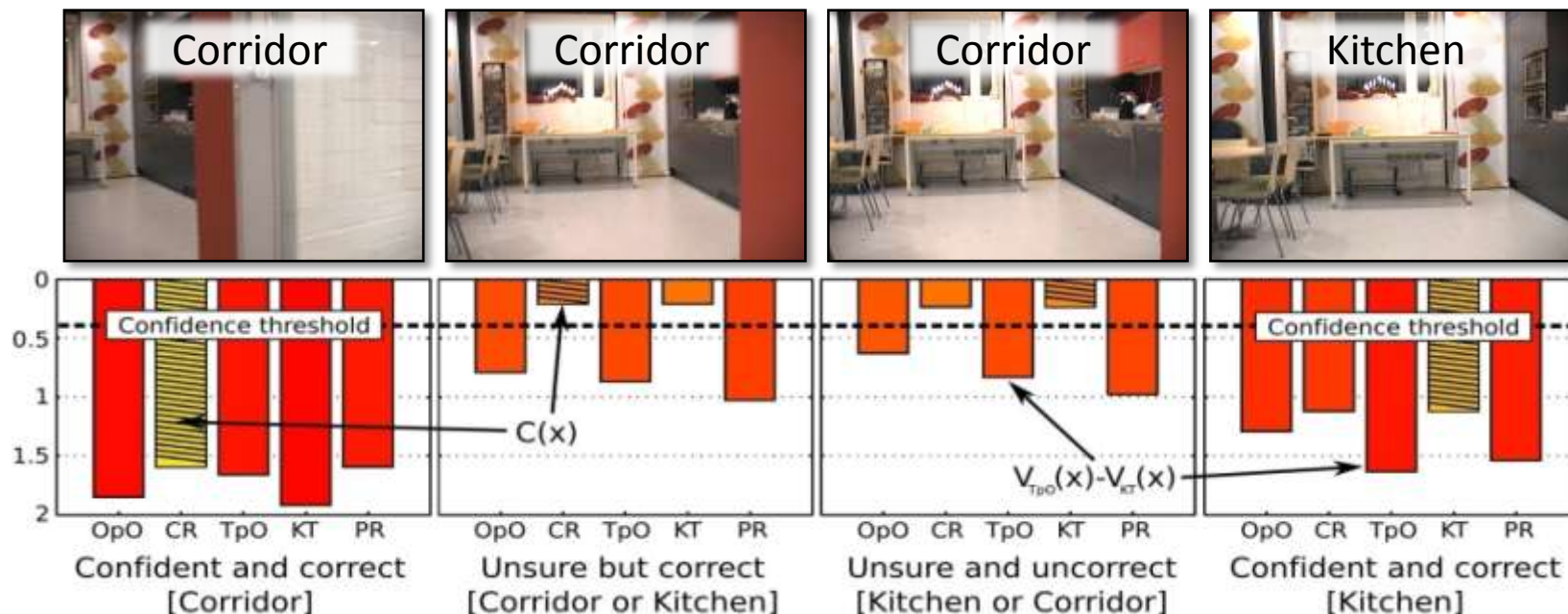
- Confidence and order of hypotheses is derived from V_j

$$C(x) = \min_{j=1 \dots M, j \neq j^*} \{V_j(x)\} - V_{j^*}(x)$$



Confidence Estimation Results

More informative decision

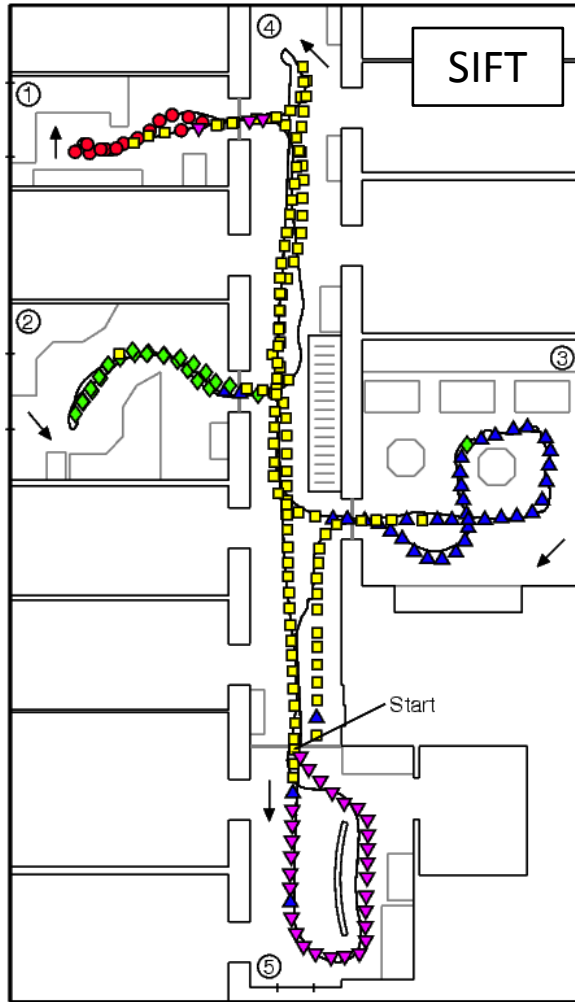


Modified one-against-all approach:

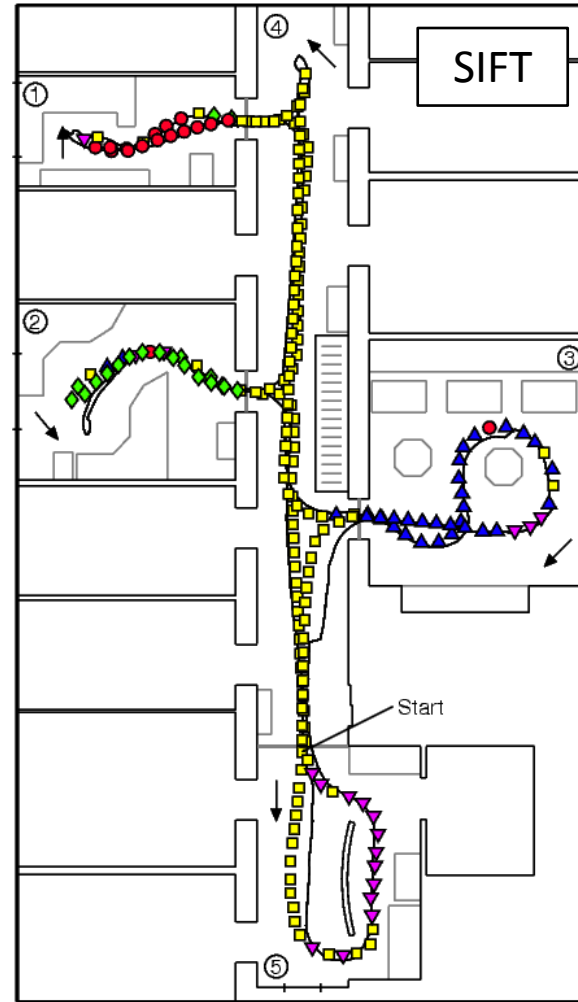
- Eliminates misclassified samples
- More reliable confidence estimates and more accurate hard decisions than standard one-against-one and one-against-all



Cue Integration Motivation



PeopleBot Cloudy -> PeopleBot Night



PeopleBot Cloudy -> PowerBot Night



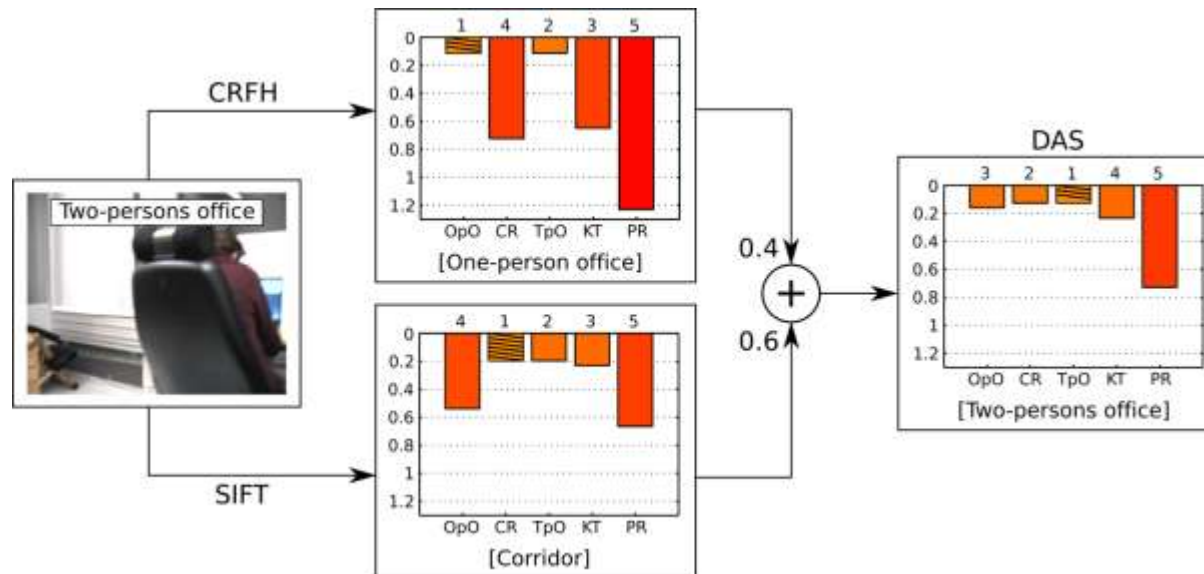
Cue Integration

G-DAS

- Generalization of DAS [Nilsback & Caputo '04]
- Scores generated by several classifiers (P) accumulated through weighted summation:

$$a_1 * \boxed{V_1, V_2, \dots, V_M}^{(1)} + a_2 * \boxed{V_1, V_2, \dots, V_M}^{(2)} + \dots + a_P * \boxed{V_1, V_2, \dots, V_M}^{(P)} = \boxed{V_1, V_2, \dots, V_M}^{(\Sigma)}$$

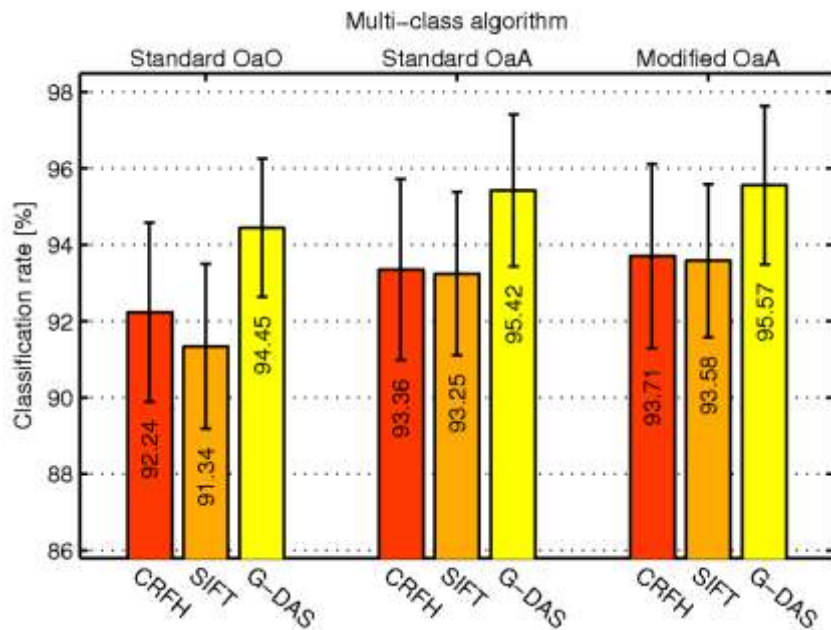
- Can give correct results even if all classifiers are wrong



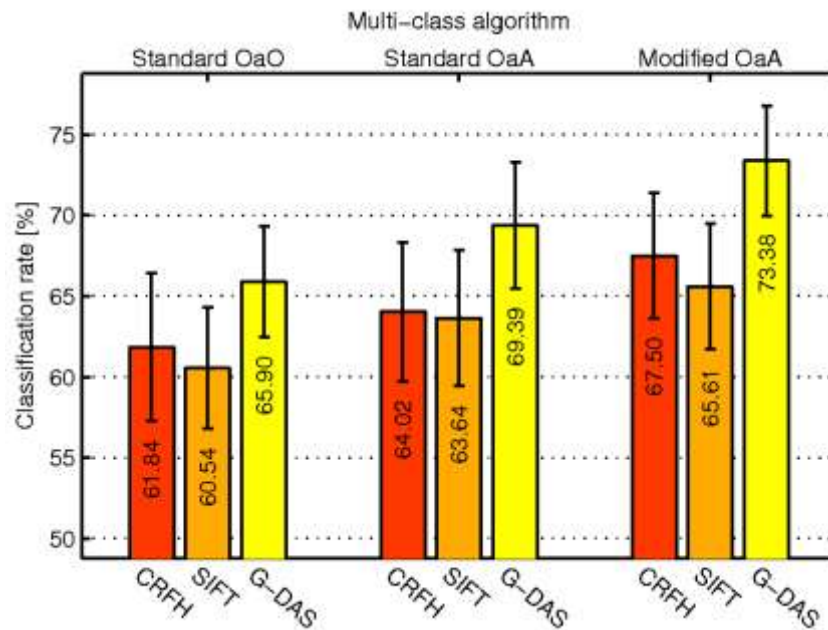


Cue Integration Results

Experimental results



Exp. 1 – Stable illumination conditions



Exp. 3 – Recognition across platforms

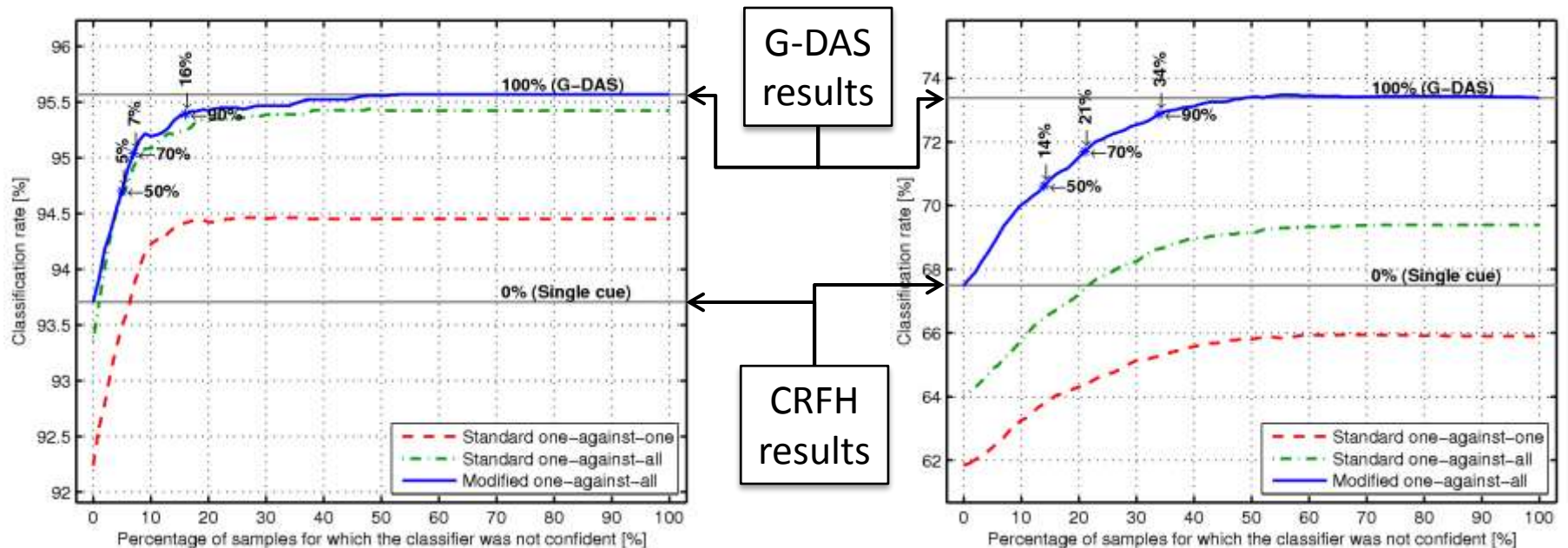
Consistent gain in accuracy (+1.9% to +7.8%)

Modified OaA approach outperforms standard methods



Confidence-based Cue Integration

- G-DAS most effective for cases of low confidence
- Using all cues can be expensive and unnecessary
- Solution: **extract and use additional cues only when confidence is not satisfactory**



Exp. 1 – Stable illumination conditions

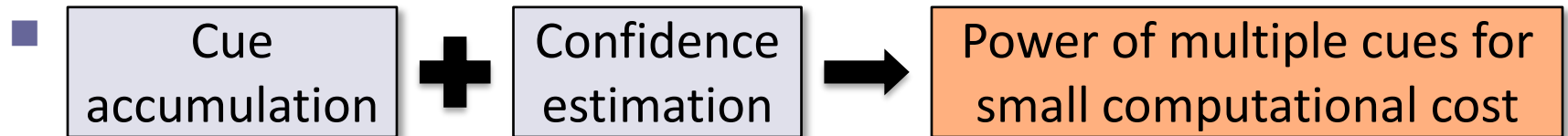
Exp. 3 – Recognition across platforms



Summary

□ Conclusions

- SVMs + local and global visual features = effective approach to visual place recognition
- Integration of different visual cues increases robustness
- Useful confidence information encoded in classifier's outputs



- Applicable for sensor integration and all large-margin classifiers

□ Ongoing and future work

- Additional visual/non-visual cues (texture, color, range info)
- Place categorization, semantic labeling
- Adaptive semi-supervised alg. for dynamic environments

Thank you



Contact:

`pronobis@csc.kth.se`

`http://www.csc.kth.se/~pronobis`

The IDOL database:

`http://cogvis.nada.kth.se/IDOL`





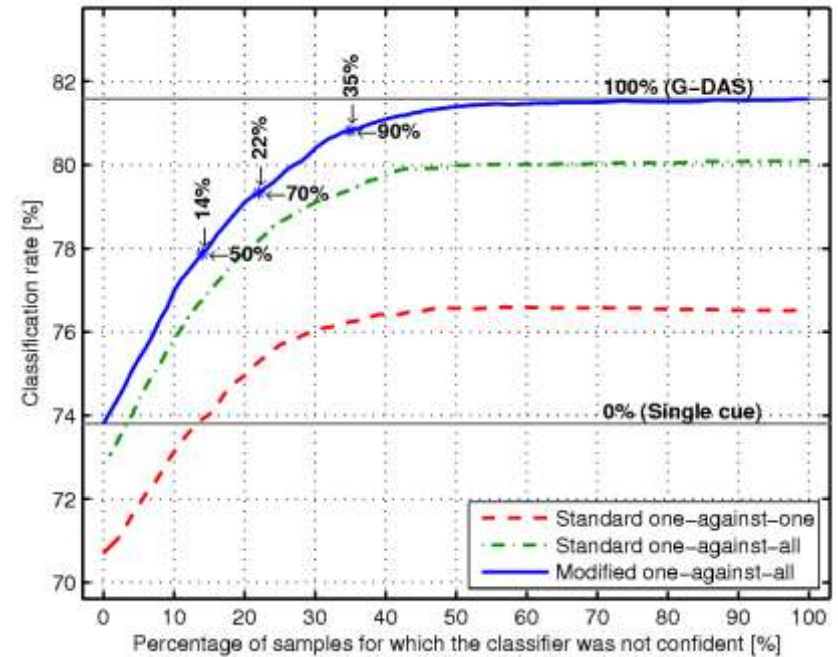
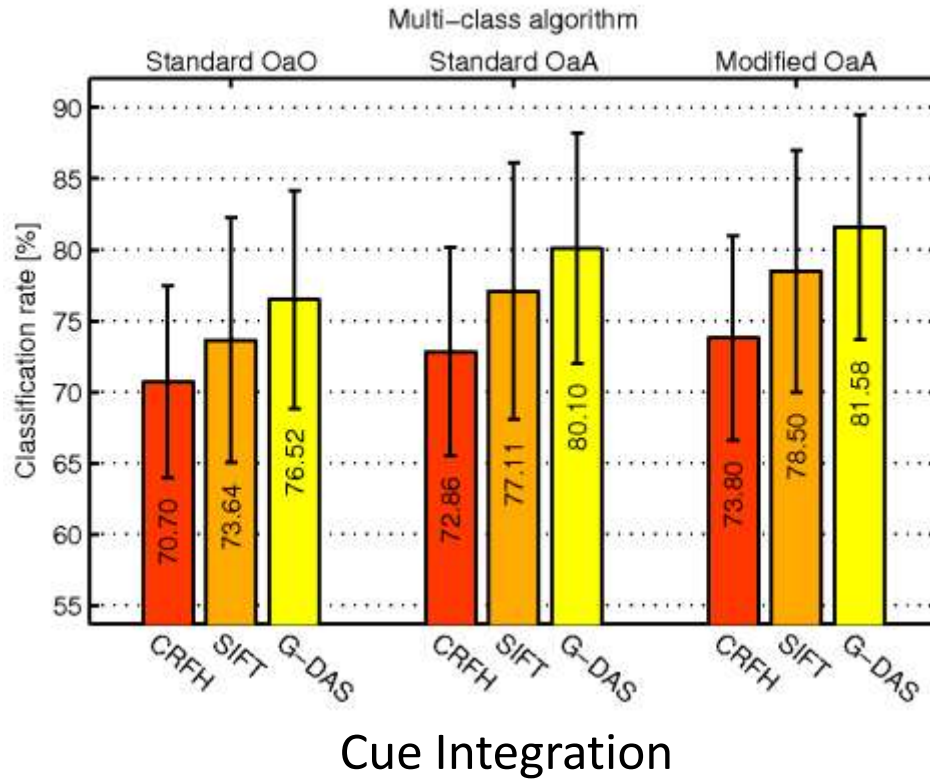
A Typical Experiment

CRFH only, trained when Cloudy, testing at Night



Experimental Results

Varying Illumination Conditions

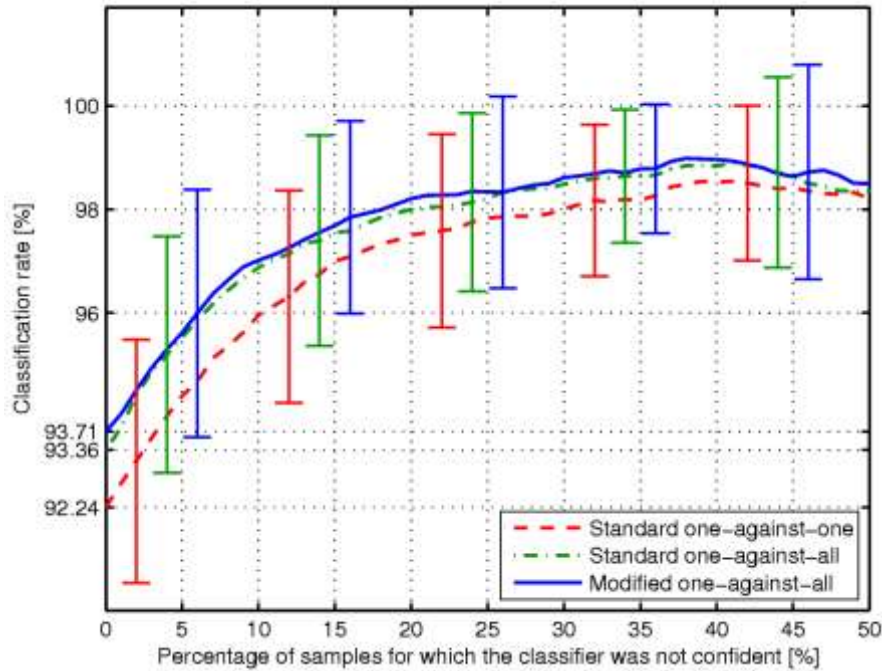


Confidence-based Cue Integration

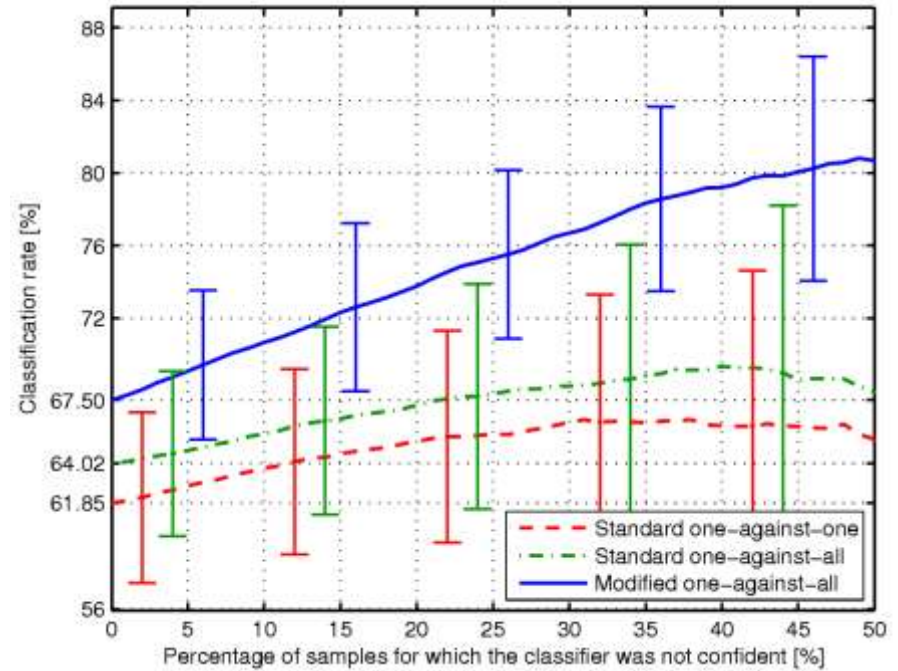


Experimental Results

Confidence Estimation



Exp. 1 – Stable illumination conditions



Exp. 3 – Recognition across platforms