A Discriminative Approach to Robust Visual Place Recognition

A. Pronobis■, B. Caputo◇, P. Jensfelt■, H.I. Christensen■

■CAS/KTH, Stockholm, Sweden
{pronobis, patric, hic}@nada.kth.se

◇IDIAP, Martigny, Switzerland
bcaputo@idiap.ch

Contribution

- Vision-based indoor place recognition system robust to visual variability introduced by
  - Varying illumination (natural/artificial light)
  - Human activity
  - Small view-point variations

Contribution

- Vision-based indoor place recognition system robust to visual variability introduced by
  - Varying illumination (natural/artificial light)
  - Human activity
  - Small view-point variations
- Thorough experimental evaluation in the domain of mobile robot topological localization
  - Using three different platforms
  - Under varying illumination
  - Over a significant time span

Motivation

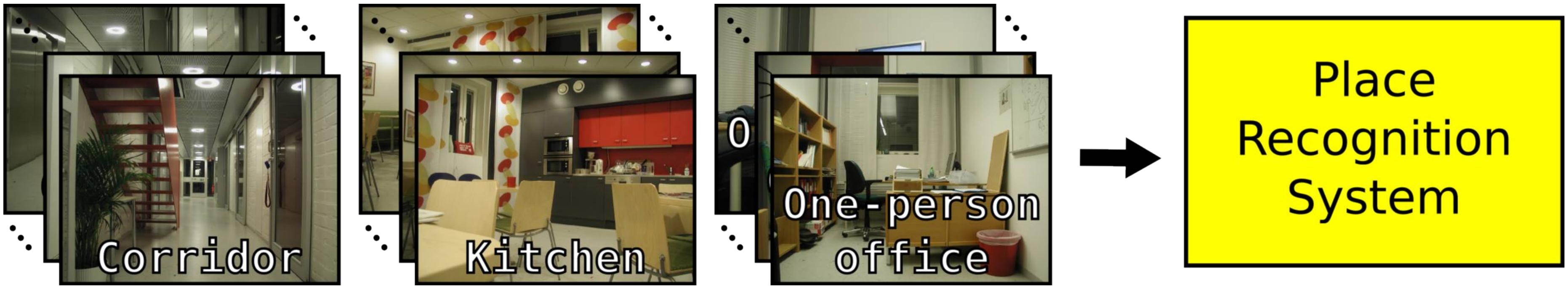
- Localization – a fundamental competence for mobile autonomous systems
- Place recognition for topological localization
  - Method for loop closing
  - Recovery from the kidnapped robot problem
  - Source of contextual information
  - Possible solution for scalability issues
- Vision-based solutions
  - Provide cues unavailable for other sensors
  - Portable and cost-effective

Our Approach

- Fully supervised, appearance-based method

Our Approach

- Fully supervised, appearance-based method
- Training:



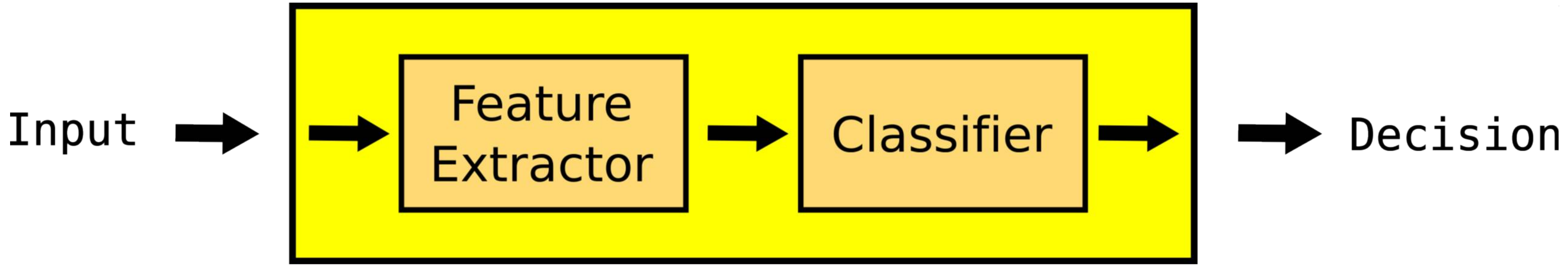
Our Approach

- Fully supervised, appearance-based method
- Training:

- Recognition:

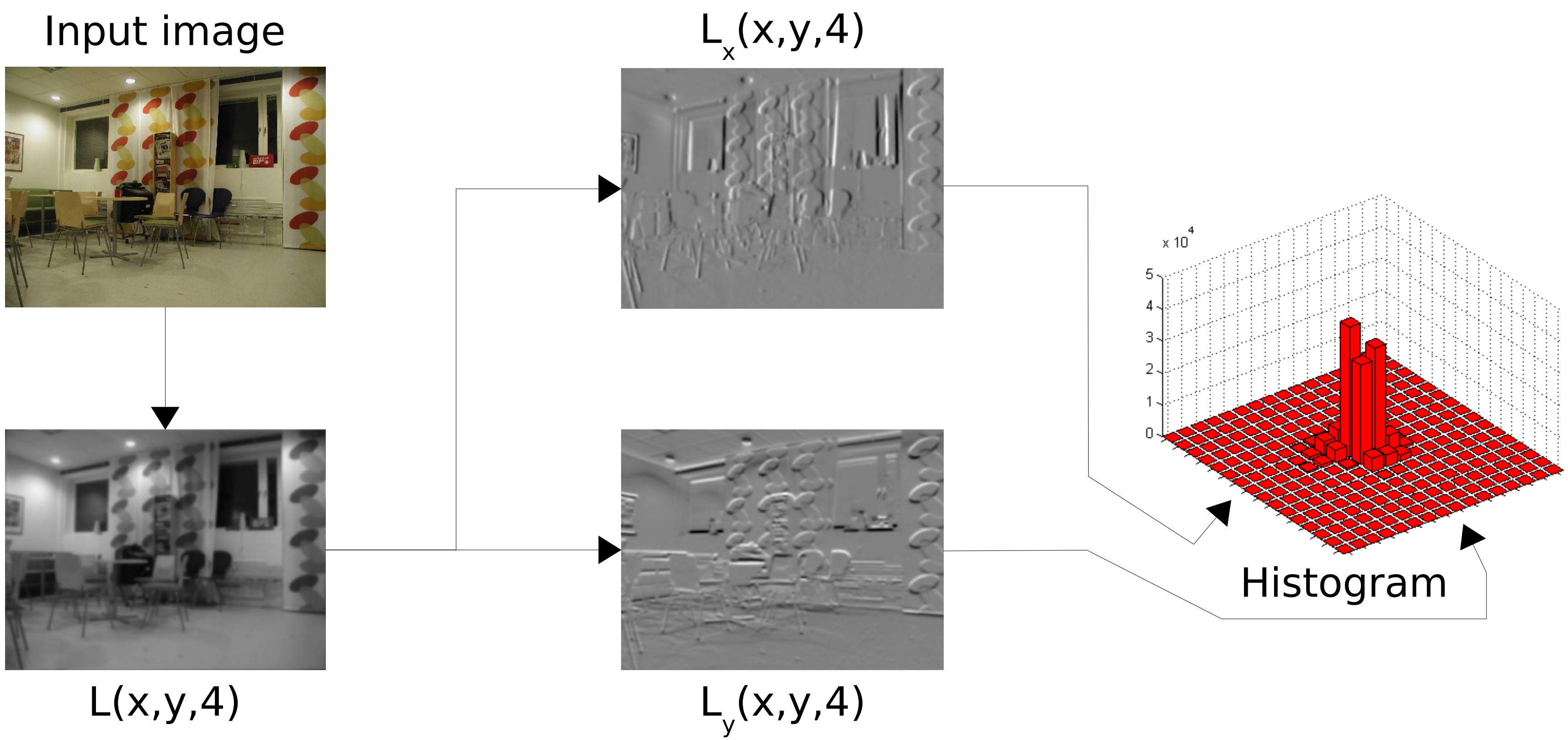
Our Approach

- Assumption: encoding of the global configuration of a scene is informative enough for recognition
- The system consists of two parts:
  - Feature extractor
    High Dimensional Composed Receptive Field Histograms (CRFH) [Linde and Lindeberg '04]
  - Classifier
    Support Vector Machines [Vapnik '98]



Our Approach
Feature extraction – High dimensional CRFH

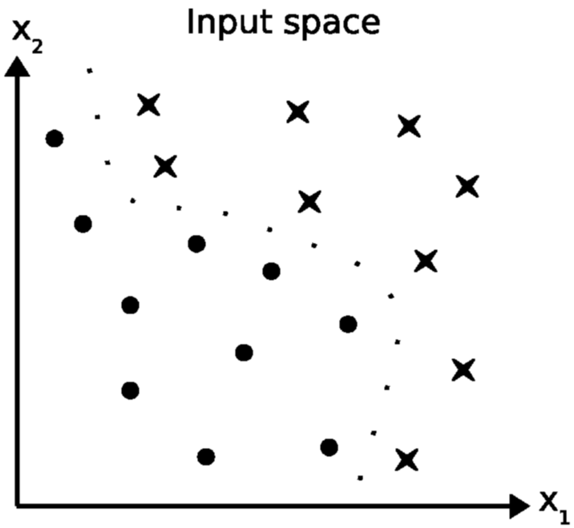
- Sparse multi-dimensional statistical representation of responses of several descriptors applied to the input image



- Our case: 6 dimensional histograms, 2nd order Gaussian derivative filters applied to the intensity channel

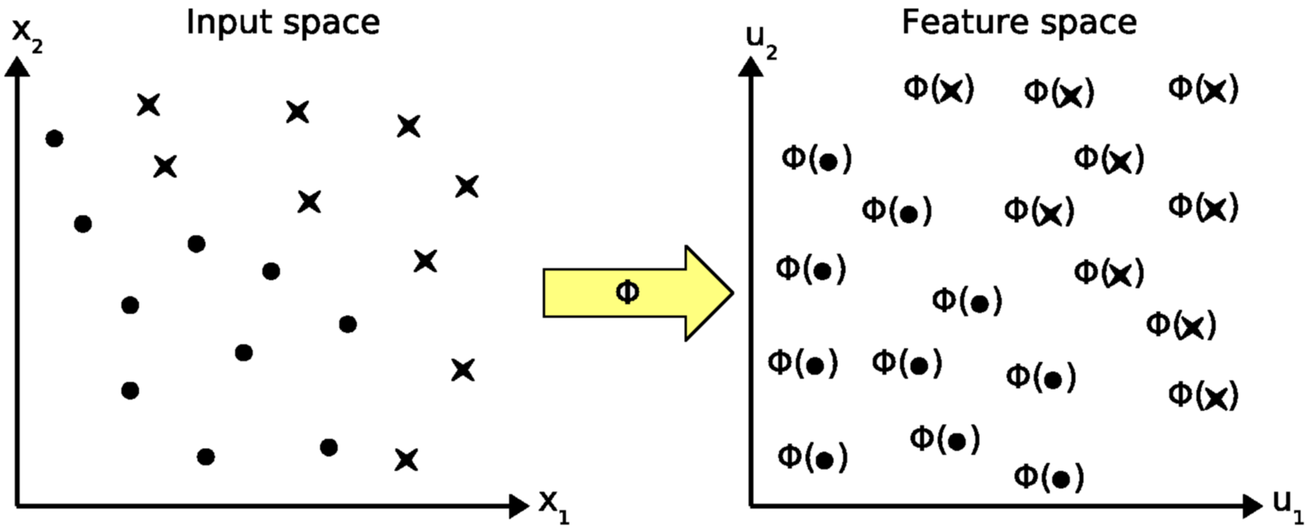
Our Approach
Classification – Support Vector Machines

- Large margin, discriminative classifier separating training data by a hyperplane in a high dimensional feature space



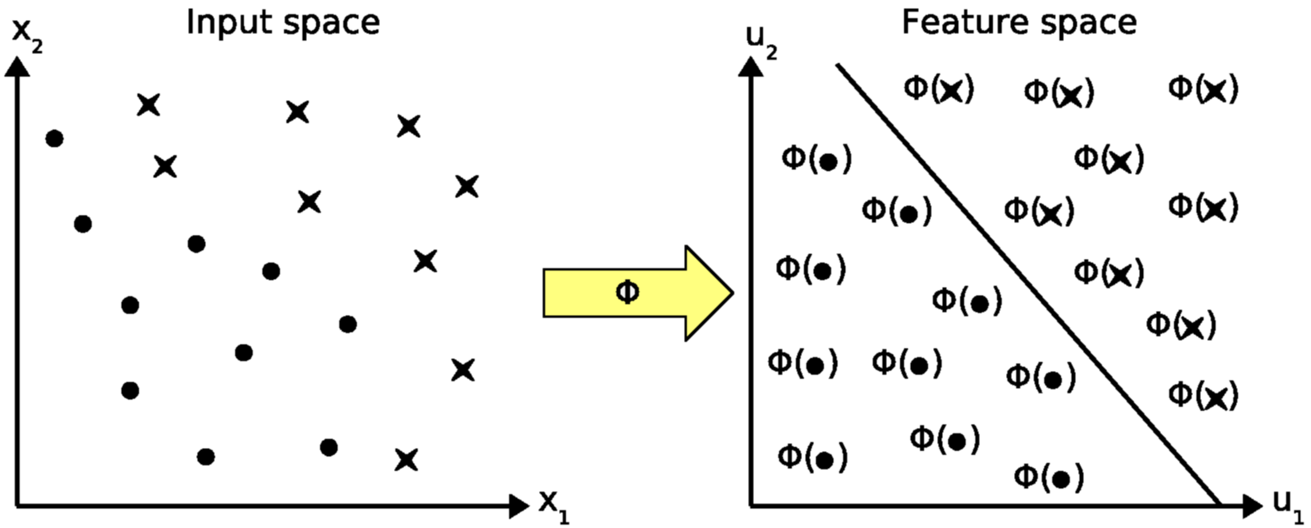
Our Approach
Classification – Support Vector Machines

- Large margin, discriminative classifier separating training data by a hyperplane in a high dimensional feature space



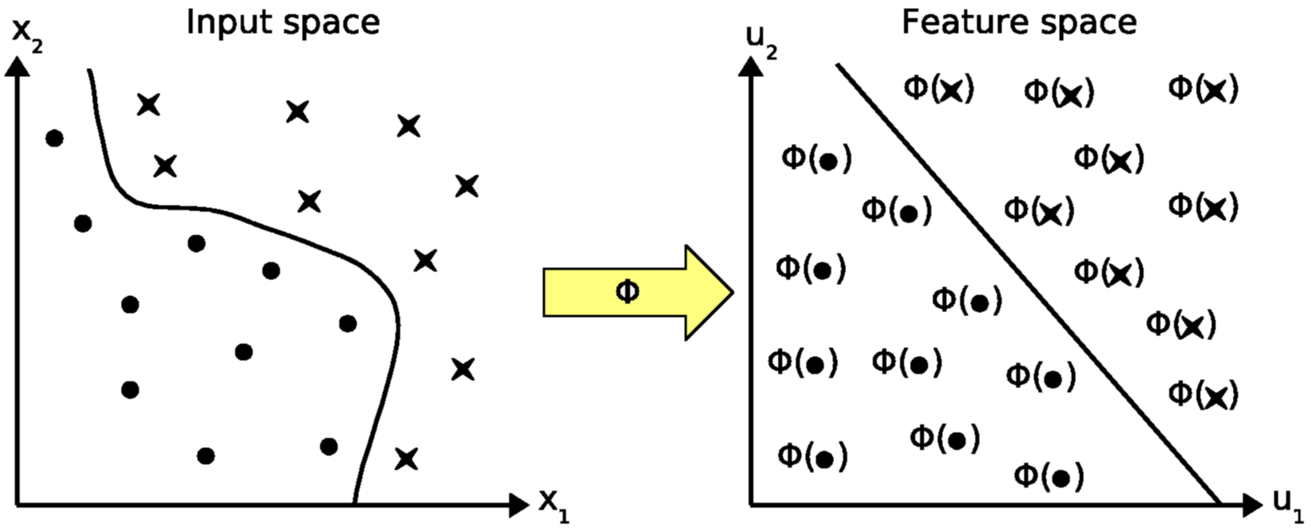
Our Approach
Classification – Support Vector Machines

- Large margin, discriminative classifier separating training data by a hyperplane in a high dimensional feature space



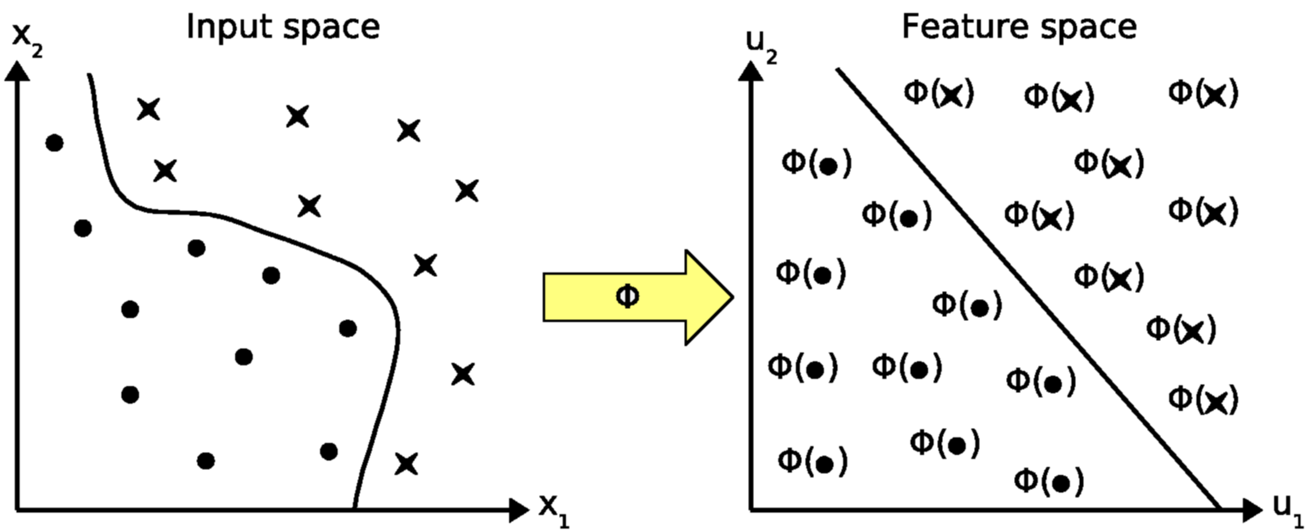
Our Approach
Classification – Support Vector Machines

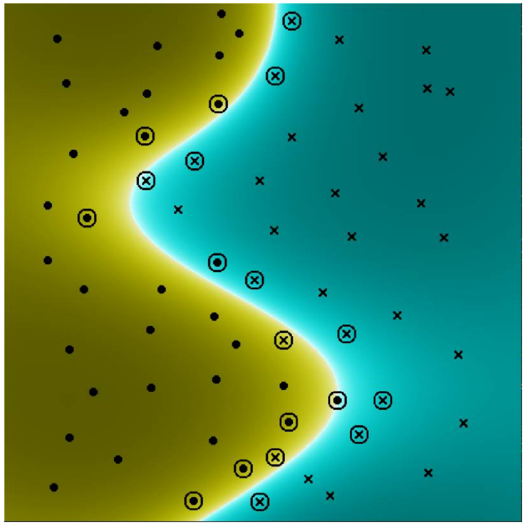
- Large margin, discriminative classifier separating training data by a hyperplane in a high dimensional feature space



Our Approach
Classification – Support Vector Machines

- Large margin, discriminative classifier separating training data by a hyperplane in a high dimensional feature space

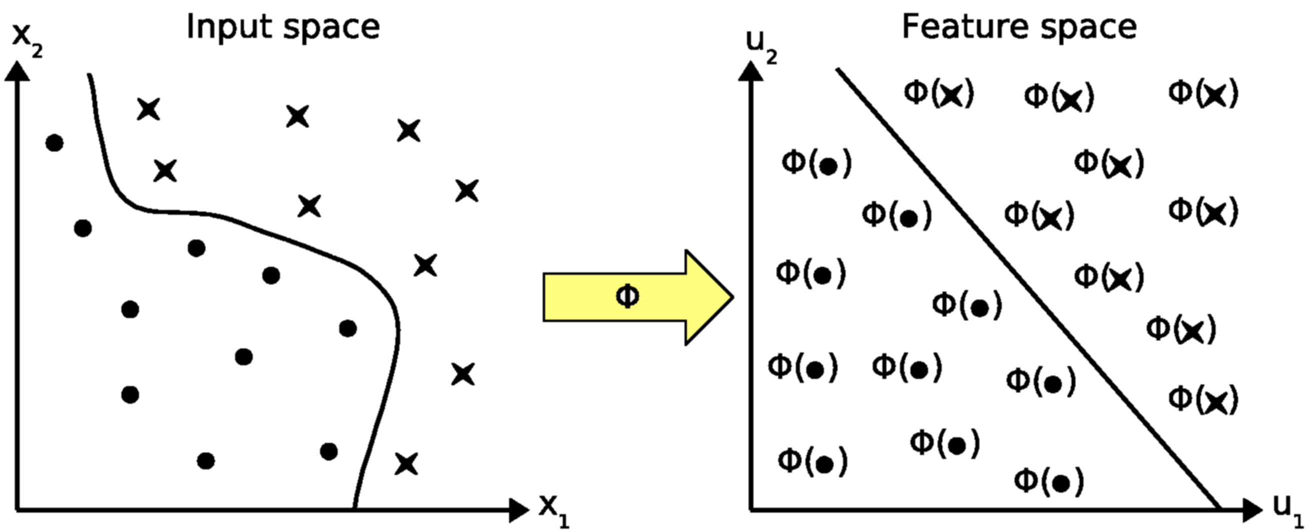


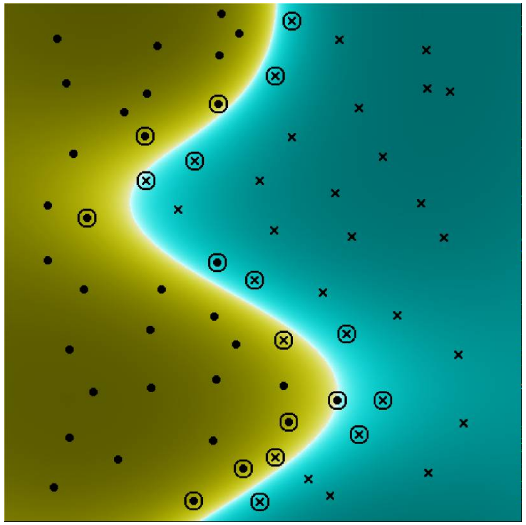


- The discriminant function is parametrized by a subset of training vectors

Our Approach
Classification – Support Vector Machines

- Large margin, discriminative classifier separating training data by a hyperplane in a high dimensional feature space





- The discriminant function is parametrized by a subset of training vectors
- Very good generalization performance
- Can be performed efficiently - kernel trick

Pronobis, Caputo, Jensfelt, Christensen. A Discriminative Approach to Robust Visual Place Recognition.

Experimental Setup
The Environment

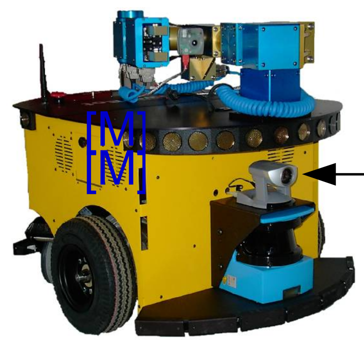
- Real office environment
- Each room represents a different functional area

Experimental Setup
Image Acquisition

- Three different platforms

Camera
(98cm)

PeopleBot Minnie



Camera
(36cm)

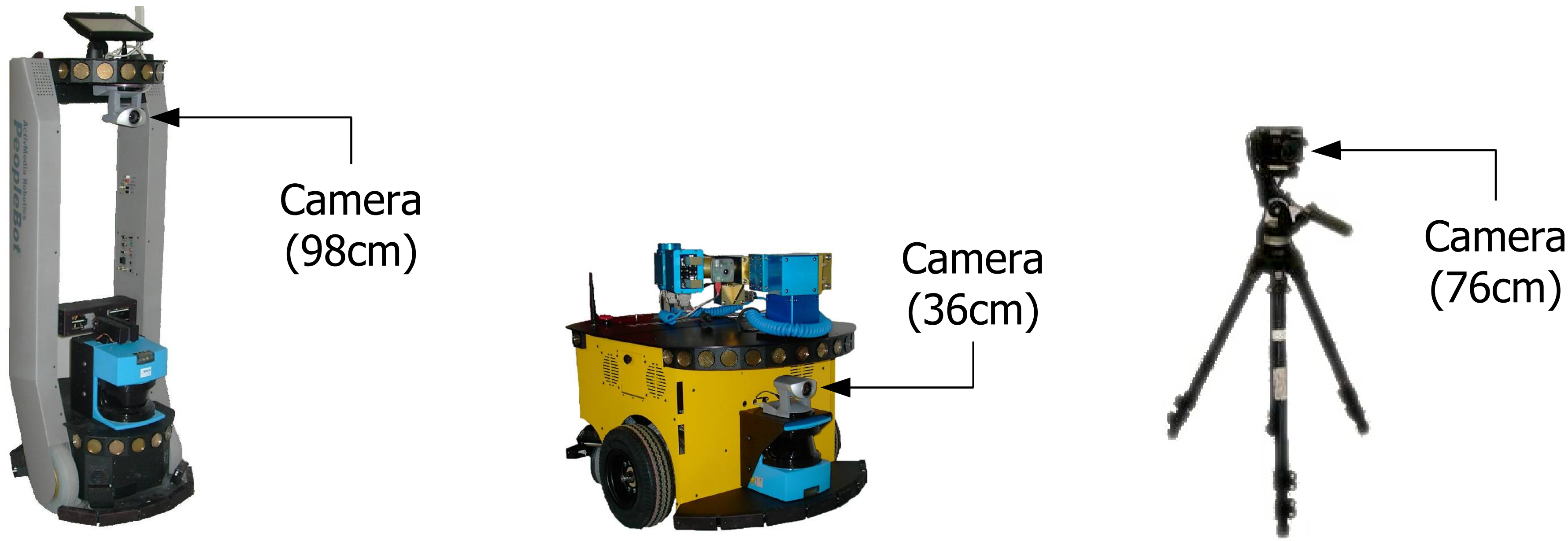
PowerBot Dumbo

Camera
(76cm)

Camera on a tripod

Experimental Setup
Image Acquisition

- Three different platforms



PeopleBot Minnie   PowerBot Dumbo   Camera on a tripod

- Three illumination conditions
- Significant span of time

Several types of variability captured

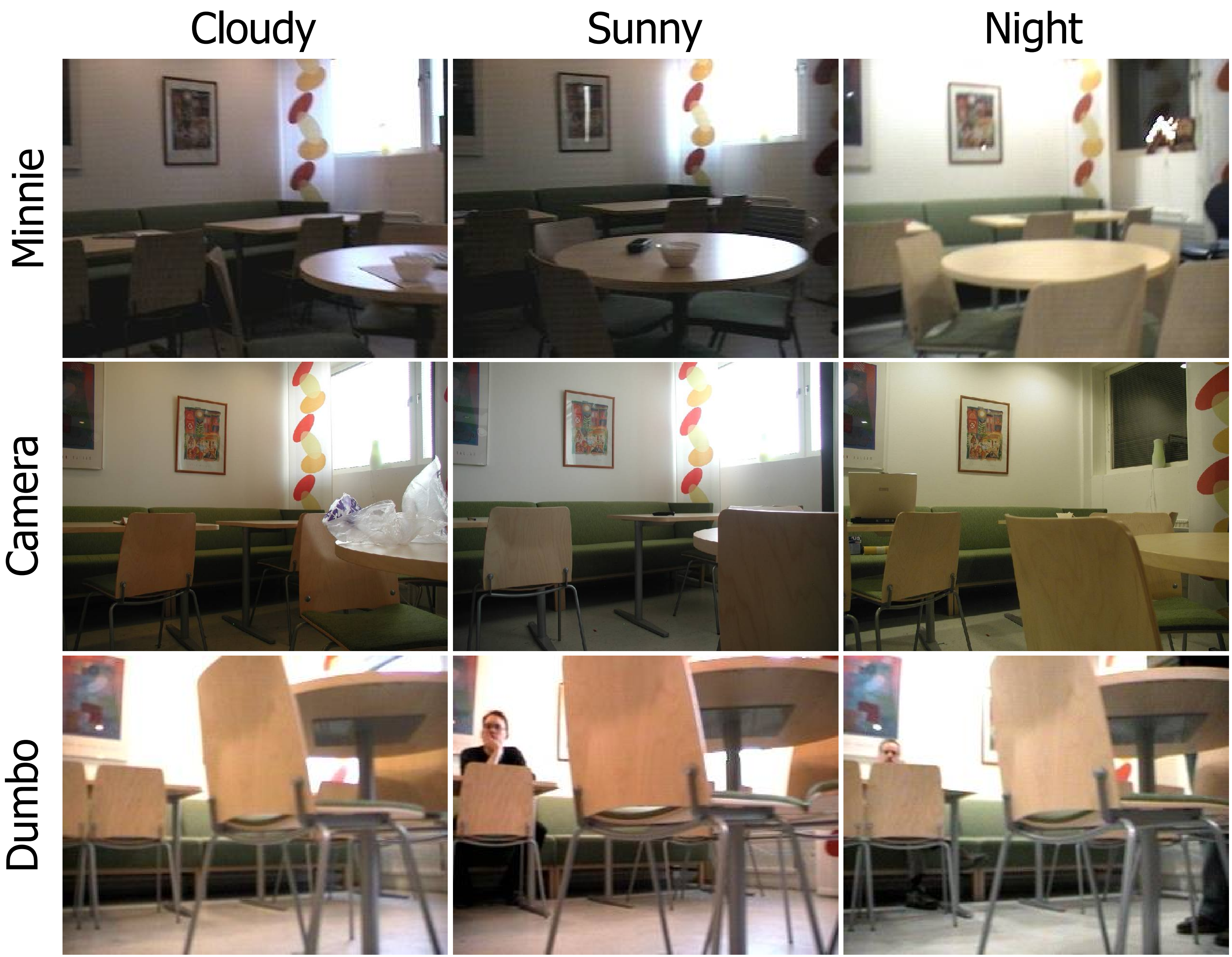
Exemplary Pictures – The Corridor

Cloudy Sunny Night

Minnie
Camera
Dumbo

Pronobis, Caputo, Jensfelt, Christensen. A Discriminative Approach to Robust Visual Place Recognition.

Exemplary Pictures – The Kitchen



Pronobis, Caputo, Jensfelt, Christensen. A Discriminative Approach to Robust Visual Place Recognition.

Experimental Evaluation

- The system was evaluated in three sets of experiments

Experimental Evaluation

- The system was evaluated in three sets of experiments

- Experiment 1 – Stable illumination conditions
  - Reference experiment

Experimental Evaluation

- The system was evaluated in three sets of experiments

- Experiment 1 – Stable illumination conditions
  - Reference experiment
- Experiment 2 – Varying illumination conditions
  - Evaluating robustness

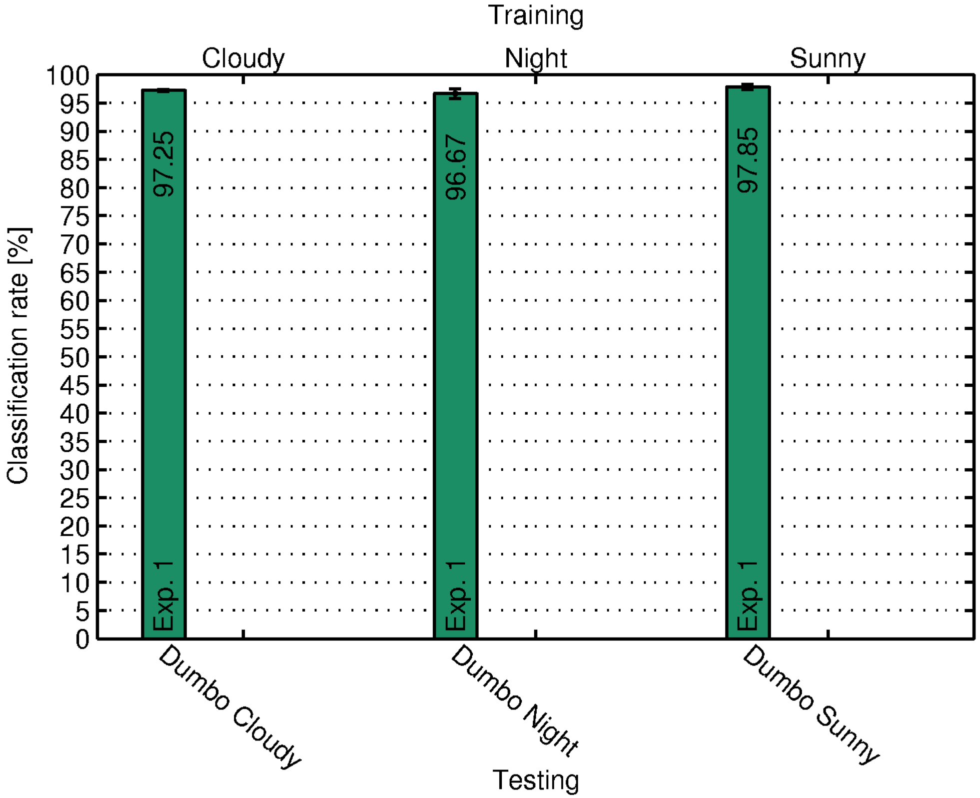
Experimental Evaluation

- The system was evaluated in three sets of experiments

- Experiment 1 – Stable illumination conditions
  - Reference experiment
- Experiment 2 – Varying illumination conditions
  - Evaluating robustness
- Experiment 3 – Recognition across platforms
  - Can a model trained on images acquired using one device be useful for localization of another device?

Experimental Results
Exp. 1 - Stable Illumination Conditions

Training on images acquired with Dumbo



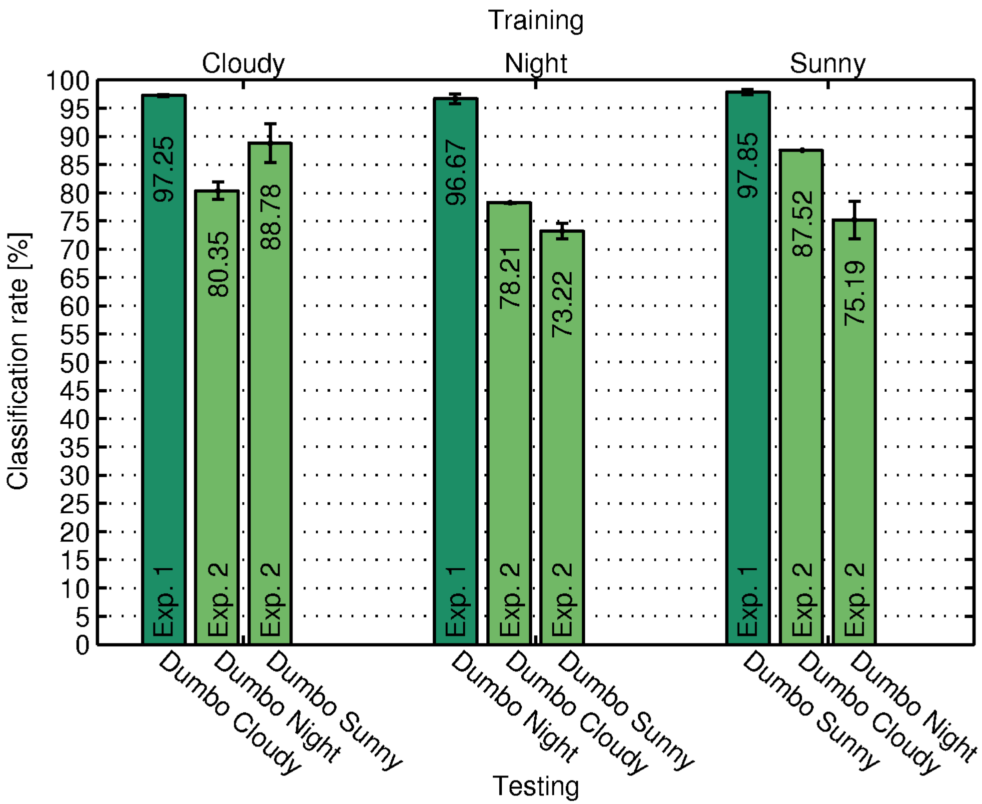
Training on the INDECS database

Not available – only one set of data for each illumination condition

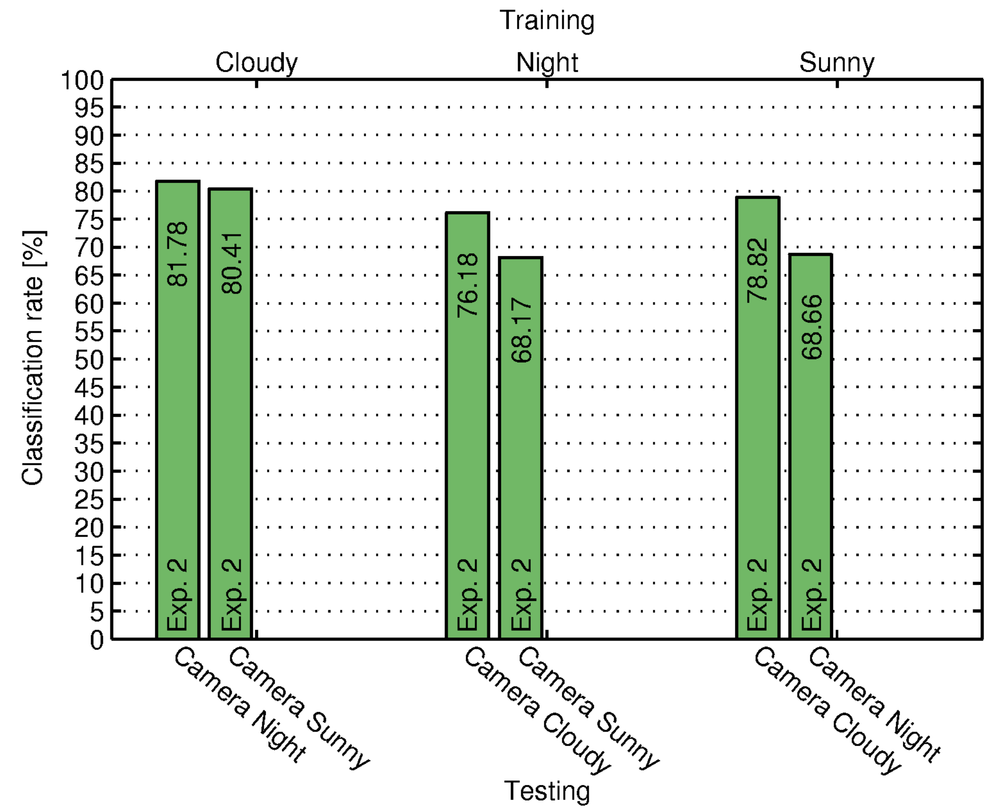
- Average class. rate: 97.2% for Dumbo and 95.5% for Minnie
- Each room treated equally, chance 20%

Experimental Results
Exp. 2 – Varying Illumination Conditions

Training on images acquired with Dumbo



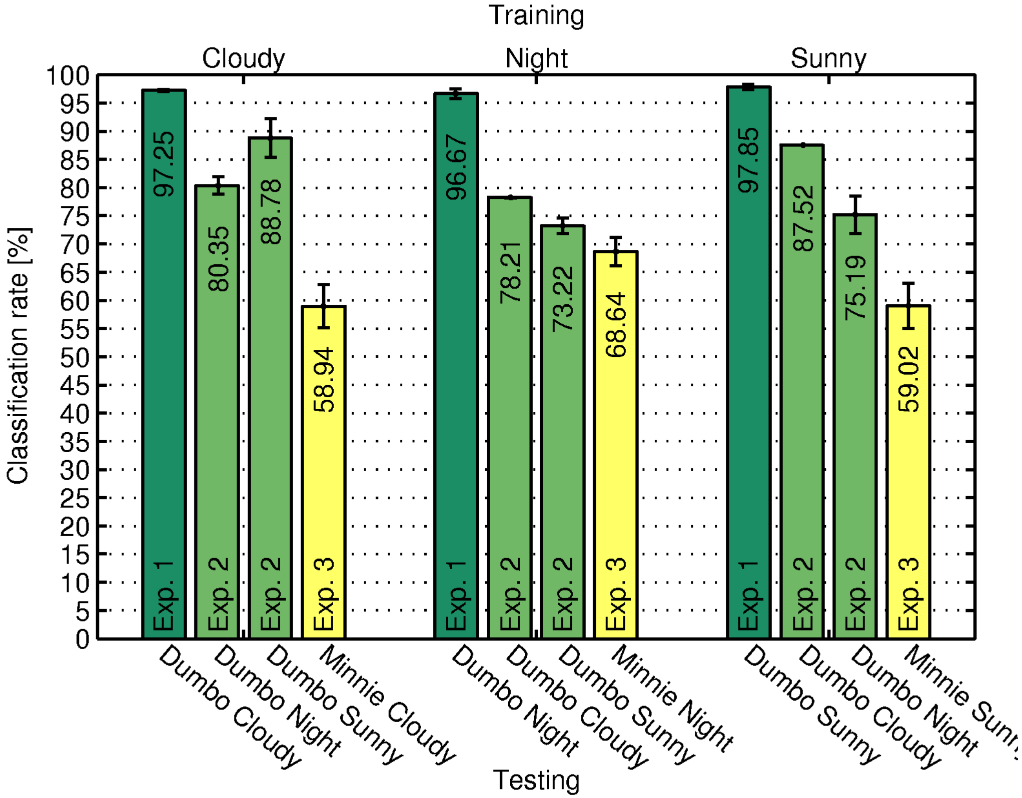
Training on the INDECS database



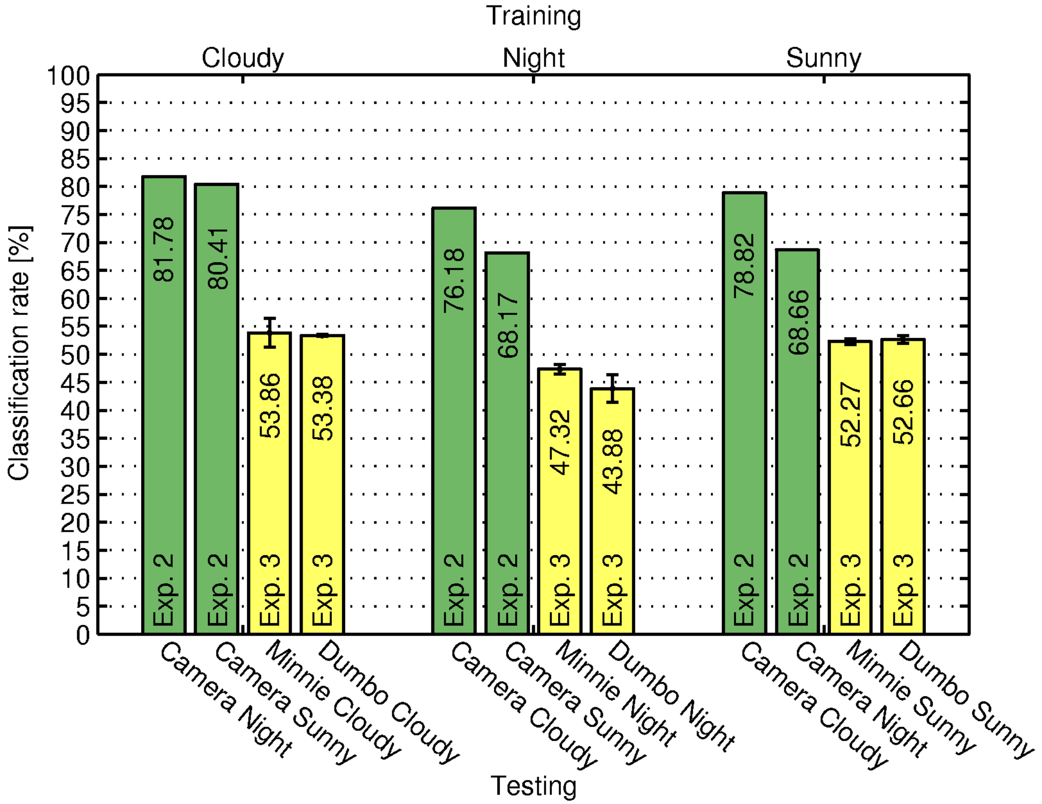
- Best performance when trained in cloudy weather (average class. rate: 84.6% for Dumbo, 81.0% for INDECS)
- 100% impossible to achieve with such approach

Experimental Results
Exp. 3 – Recognition Across Platforms

Training on images acquired with Dumbo



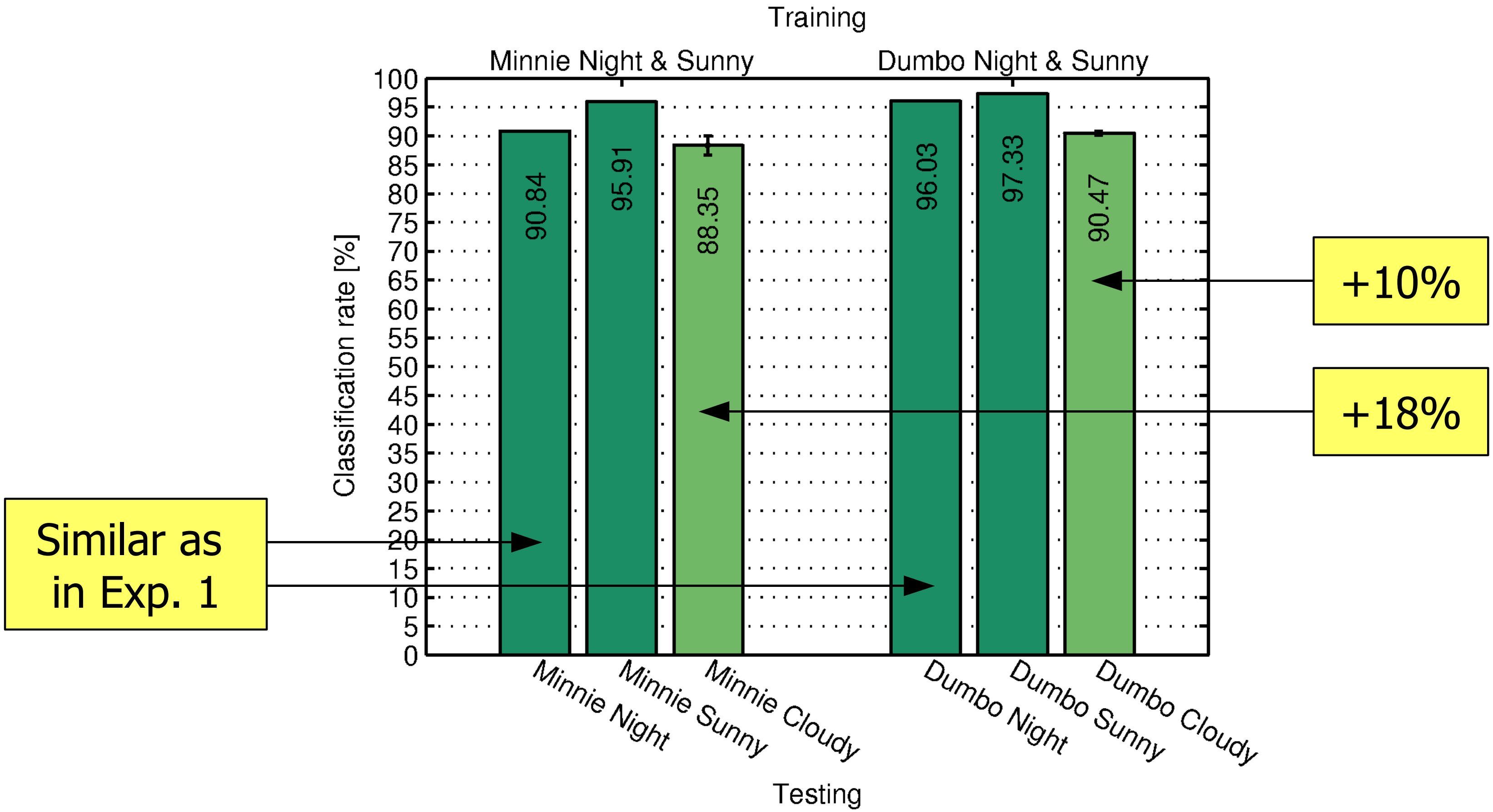
Training on the INDECS database



- Still up to about 70% of images classified correctly
- INDECS: only about 50%, but very large variability

Experimental Results
Robustness and Efficiency

- Our goal: robust and efficient solution (little training data)
- Robustness can be improved – more data required



Training on two sequences acquired under different illumination conditions

Summary and Future Work

- Rich global descriptor + SVM = robust and efficient approach to visual-based place recognition
- Successful in handling significant changes in illumination and other variations that occur in real-world environments
- Recognition time 350ms
- Future work:
  - Incorporating illumination invariance into image descriptor
  - Fusing information from more than one image
  - Handling very long-time variations through incremental learning and adaptability (ongoing)
  - Place categorization

Thank you

Contact: `pronobis@nada.kth.se`

The data: `http://cogvis.nada.kth.se/INDECS`
`http://cogvis.nada.kth.se/IDOL`

Full video: `http://www.nada.kth.se/~pronobis`