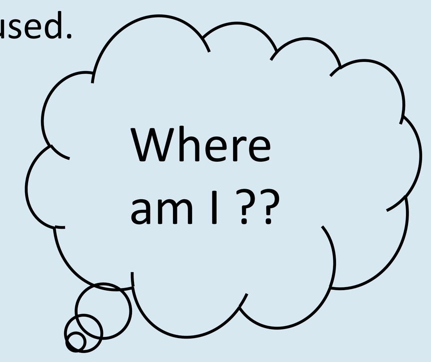
Place Recognition System for Localization of Mobile Robots

Raghavender Sahdev, Asheer Bachoo and John K. Tsotsos
Dept. of Electrical Engineering and Computer Science and Center for Vision Research
York University , Toronto, Canada
{sahdev, abachoo, tsotsos}@cse.yorku.ca

Introduction

- The robot learns from experience and recognizes previously seen / unseen topological places in known / unknown environments.
- Our system has been practically tested with a novel dataset developed by us.
- A HOUP (Histogram of Oriented Uniform Patterns) descriptor is used to represent an image.
- Classifiers (1 Nearest Neighbor for Place recognition and SVM for place categorization) have been used.



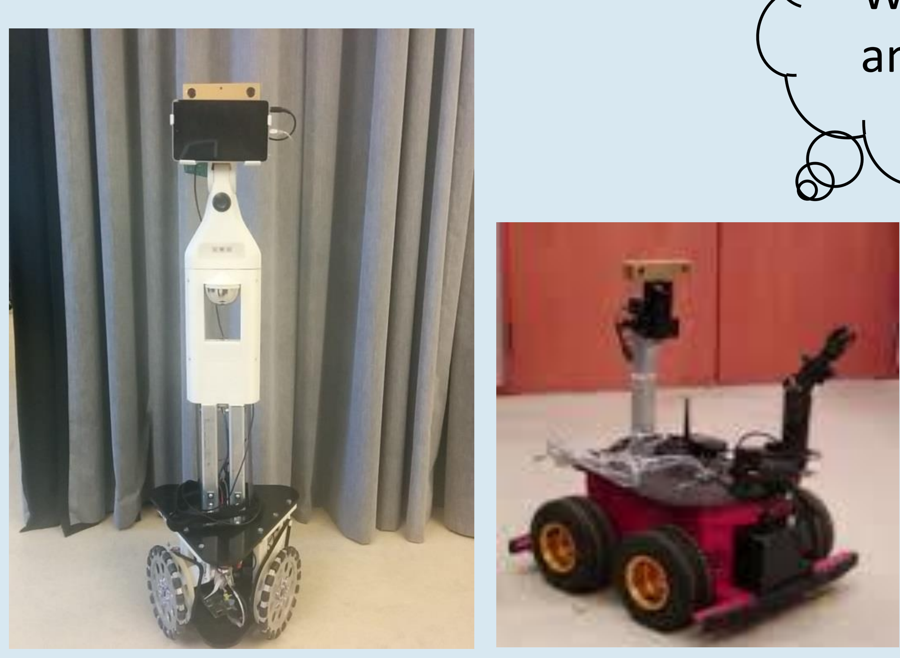


Fig. 1. Robots being used for place recognition Virtual Me (left) and Pioneer (right)

The HOUP Descriptor

- The Histogram of Oriented Uniform Patterns (HOUP) is a distribution based descriptor used to build the histogram which describes the frequency content of the image;

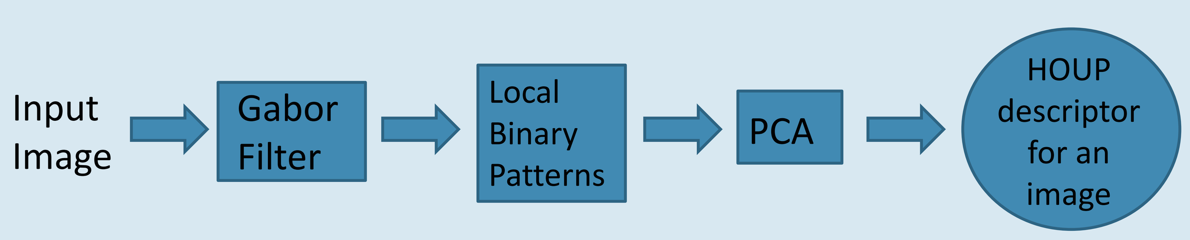


Fig. 2. Generation of the HOUP descriptor

- Initially the image is convolved with a Gabor filter tuned to 6 different orientations from 0 to 5π/6.

$$v_k(x) = \left| \sum_{x'} i(x') g_k(x - x') \right|$$

- Here $v_k(x)$ is the output of the convolved image with the Gabor filter $g_k(x-x')$ at a specific frequency and orientation. $i(x')$ is the input image to the Gabor filter.
- The output $v_k(x)$ of the Gabor filter is passed to generate the local binary patterns (LBPs) of the image as in [1].
- We use a 3x3 neighborhood to generate the LBPs; we get 58 uniform patterns out of the 256 total patterns. 59 dimensions used with one dimension to represent non uniform patterns
- Gabor filter at 6 different orientations gives a 59 * 6 = 354 dimensional representation of an image sub block.
- 354 is brought down to 70 using the Principal Component Analysis.

Sub Division Scheme

- Each image is divided into sub blocks to generate different features which would provide an informative representation of the image.
- We divide the given image into 1x1, 2x2, 3x3, 4x4 and 5x5 blocks. So in total we have 55 sub blocks * 3 frequencies = 165 candidate features.
- A HOUP descriptor for each image sub block at a specific Gabor frequency is computed.
- The 3x3 sub division scheme gives the best representation.
- 9 features each of 70 dimensionality (630 dimensions for an image) is used at a particular frequency.

Our Dataset

- Dataset built at 3rd floor of Lassonde Building at York University.
- Our Dataset has 11 places (as shown) each scene has 3 representations – the left, right and depth image.
- The dataset was built in 2 different lighting conditions day and night using 2 robots *(Virtual Me* and *Pioneer).* Robots were manually driven.
- The Camera (Point Grey Bumblebee 2 stereo vision camera). is mounted at heights of 117cms and 88cms for *Virtual Me* and *Pioneer* respectively.
- Each image has a resolution 640 x 480 at a frame rate of approximately 3 frames per second.
- Each place has 60 – 200 images and in total we have 1800 – 2000 images which are used for training and a similar image sequence for testing.

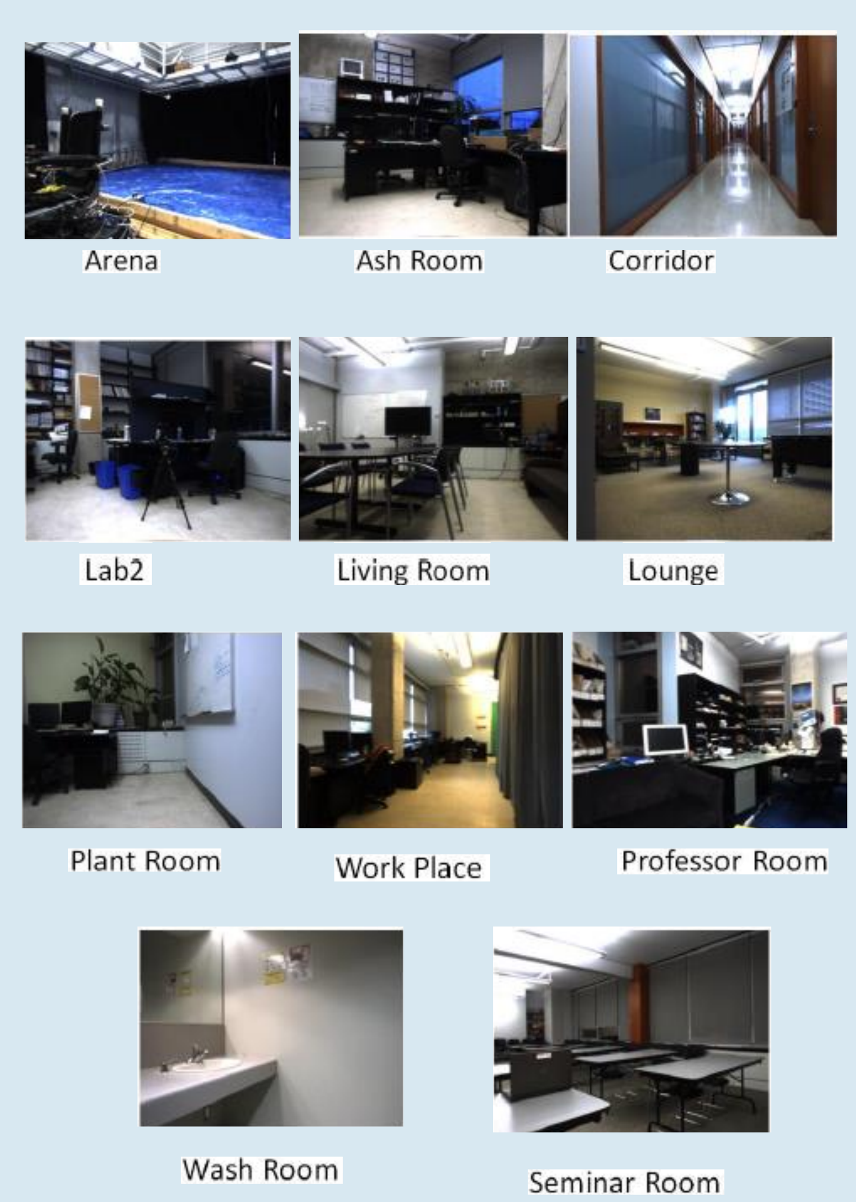


Fig. 3. Eleven different places using which the dataset was generated

Experimental Results

- Four sets of experiments were conducted to evaluate the performance of our system
  - Same Robot Same Lighting Conditions.
  - Same Robot Different Lighting Conditions
  - Different Robot Same Lighting Conditions
  - Different Robot Different Lighting Conditions

Experiment	Training Set	Testing Set	Lighting Conditions	Accuracy
1	Pioneer	Pioneer	Same	98
	Virtual ME	Virtual ME	Same	98
2	Pioneer	Pioneer	Different	93
	Virtual ME	Virtual ME	Different	93
3	Pioneer	Virtual ME	Same	92
	Virtual ME	Pioneer	Same	92
4	Pioneer	Virtual ME	Different	82
	Virtual ME	Pioneer	Different	85

Fig. 4. Accuracies reported by our system on the dataset generated by us.

Generalizability of our Method

- Our proposed method was also tested with the KTH Idol dataset [2]; it performs very well giving accuracies comparable to those reported by the highest on this dataset by Fazl-Ersi and Tsotsos [3].
- We did place categorization by implementing an SVM classifier; we got an accuracy of 75 % on the UIUC dataset built by Lazebnik et al. (2006) [4].
- The robot was driven manually in a new environment through different places; places were labeled for the robot during training.
- The robot could then successfully identify each place with high accuracy during the testing phase.
- Our method has proven to generalize over new environments by training the robot once and testing it in that environment.

References

[1] Ojala T, Pietikainen M and Maenpaa T (2002) Multi resolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24(7): 971-987

[2] Pronobis A, Caputo B, Jesfelt P, et al. (2006) A discriminative approach to robust visual place recognition. In: *proceedings of International conference on robots and systems,* pp. 3829-3836

[3] Fazl-Ersi and Tsotsos (2012) Histogram of Oriented Uniform Patterns for Robust place recognition and categorization. In: The International Journal for Robotics Research.

[4] Lazebnik S, Schmid C and Ponce J (2006) Beyond bag of features: Spatial Pyramid Matching for recognizing natural scene categories. In *proceedings of the IEEE international conference on computer vision and pattern recognition,* pp 2169 – 2178