IMPROVING OBJECT DETECTION AND RECOGNITION FOR
SEMANTIC MAPPING WITH AN EXTENDED INTENSITY AND
SHAPE BASED DESCRIPTOR

ERICKSON R. NASCIMENTO1,∗, GABRIEL L. OLIVEIRA1, MARIO F. M. CAMPOS1, ANTONÔO W. VIEIRA1,2
1Departamento de Ciência da Computação - UFMG - Belo Horizonte, Brazil
2Departamento de Ciências Exatas - CCET - UNIMONTES - Montes Claros, Brazil
∗Corresponding author: erickson@dcc.ufmg.br

I. Introduction
In this work, we propose a descriptor composed of Binary Appearance and Shape Elements (BASE) based on BRIEF that efficiently combines texture and geometrical shape information to improve discriminative power. We used this new descriptor to recognize objects under different illumination conditions with an adaptive boost classification framework that provides semantic information for a mapping task.

II. BASE Descriptor
Our descriptor is inspired by the work of [1]. We propose an extension to their work by embedding geometrical cues during descriptor construction. The center of the image patch p is positioned at the location of each keypoint found by the FAST detector. For all positions in a set of (x, y)-locations we evaluate the function:

\[ f(p) = \begin{cases} 1 & \text{if } p(x) < p(y) \lor (N(x), N(y)) \geq \rho \\ 0 & \text{otherwise} \end{cases} \]

where \( p(x) \) is the pixel intensity at position \( x = (u, v)^T \) and \( N(x) \) is the normal vector of the projection of pixel \( p(x) \) in the point cloud. The final descriptor is encoded as a binary string computed by:

\[ b(p) = \sum_{i=1}^{256} 2^{i-1} f(p, x_i, y_i) \]

III. Semantic Mapping
Our classifier is composed of binary weak classifiers (set of descriptors) integrated by the AdaBoost algorithm.

Algorithm 1 Semantic Map(c):

1: while true do
2: \( p = \) Particles\( \{ \text{leafPosition()} \}
3: \( f = \) getHOGFeature()
4: \( K = \) FAST(I)
5: \( T = \) \( \{ \text{p} | p \in K \}\)
6: Find label class \( c^* \) solving:
7: \( c^* = \arg \max_{c \in T} \sum_{p \in T} h_p(T) \)
8: if \( c^* \neq \text{“rest”} \) then
9: \( \text{max}(g) = c^* \)
10: end if
11: end while

IV. Matching Performance

Figure 1. Objects used for classification and detection experiments. The last image is an example of negative sample used in the training and test steps.

Figure 2. a) ROC curve of matching using BASE descriptor. We note high true positive rate with low false positive rate for all objects in the dataset. b) CPU time for Learning and Classification steps. While the use of BASE in learning step is approximately 2 times of SURF and it is closer to BRIEF, in the classification performance our descriptor was almost 2 times faster than BRIEF and 4 times faster than SURF.

V. Semantic Mapping Experiments

Figure 3. Confusion matrices (rows-normalized) between the nine classes. We observe a much better classification for BASE descriptor justified by the clear diagonal on its confusion matrix even consuming less CPU time. We also note that the classifiers built with BRIEF and SURF descriptors present a strong bias toward Toolbox class.

Figure 4. The use of BRIEF or SURF results in a large number of false positive detection (Red stars). While was detected (Cyan squares) only one object with SURF and BRIEF detected two, BASE found six without generating too much false positive detections. Green circles mean correct detection and classification.

VI. Conclusions
We have proposed a new lightweight descriptor that efficiently combines intensity and shape information to construct fast and low memory consumption signatures for keypoints. We compared this descriptor against standard descriptors in the literature. This descriptor improved robotics mapping and object recognition in several experiments.

References

Acknowledgements

[Image: CNPq, FAPESP, Unimontes]