Unidimensional Multiscale Local Features for Object Detection under Rotation and Mild Occlusions

Michael Villamizar, Alberto Sanfeliu and Juan Andrade Cetto
{mvillami, sanfeliu, cetto}@iri.upc.edu

Institut de Robòtica i Informàtica Industrial, CSIC-UPC
Llorens Artigas 4-6, 08028 Barcelona, Spain

Abstract. In this article, scale and orientation invariant object detection is performed by matching intensity level histograms. Unlike other global measurement methods, the present one uses a local feature description that allows small changes in the histogram signature, giving robustness to partial occlusions. Local features over the object histogram are extracted during a *Boosting* learning phase, selecting the most discriminant features within a training histogram image set. The *Integral Histogram* has been used to compute local histograms in constant time.

1 Introduction

Color histograms are often used as local features for object identification and tracking [1, 2], specially, given its invariance to pose change. However its main drawback is its sensitivity to illumination conditions. Schiele and Crowley [3] have extended the idea of representing the object by histograms, incorporating other local image features like the gradient magnitude, orientation and laplacian, resulting in a multidimensional histogram representation. This approach performs robust object recognition under different viewing conditions, such as, orientation, scale and view points changes.

With the propose of attaining object detection for real time applications, many methods have arisen that tackle the feature computation cost. One simple and effective method is based on the use of integral images. Viola and Jones [4] presented their integral image based on accumulation of pixel intensities over the image axes. Other extensions have been proposed to calculate other local properties efficiently. Villamizar *et al* [5] and Porikli [6] developed the Integral Histogram, with which is possible to compute rapidly any local histogram independently of its size and location.

2 Proposed Method

In this paper we combine the benefits of speed from the integral image computation with the invariant properties that color histograms give, and build on top
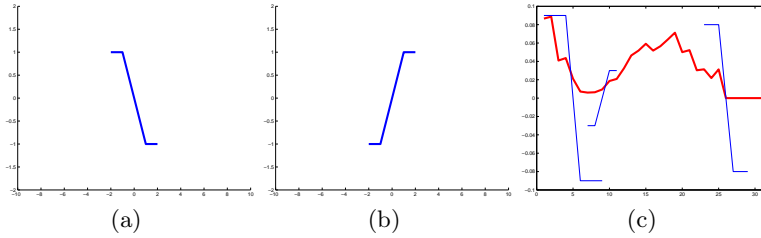
Fig. 1. Haar features a) sign + b) sign − c) Haar description

of our previous work on Boosting algorithms to produce a fast and robust object recognition system. The major benefit of the proposed method relies on its robustness to partial occlusions, since the histogram matching is performed locally, using a robust boosted classifier based on the combination of local features.

The work reported here introduces a novel multiscale unidimensional histogram representation based on a linear combination of Haar features, that follows the spirit of other typical feature sets learned via Boosting. These histograms are efficiently computed using our previously reported integral histogram image [5] and we compare on its use for object detection against the Swain and Ballard histogram intersection metric.

2.1   Local Features

We propose to describe objects by means of intensity level histograms, in order to achieve viewpoint invariance [1, 3]. However, our similarity measurement relies on a linear boosted classifier that uses Haar local features over the histogram signature. Histogram matching is carried out locally.

Those local features that are more discriminant during the Boosting learning phase are selected as weak hypothesis or classifiers, and their linear combination gives a strong hypothesis, called strong classifier. The Haar local features showed in Figure 1 represent a simple and suitable form to describe a histogram signal. They encode the inflexions in the histogram at any location, width, and sign. Consequently, the object can be modelled as a Haar decomposition of its histogram signature using the more relevant coefficients (see Figure 1c). Intensity level histograms are computed from both the patch training images (30x30 pixels) for the Boosting stage, and patches extracted from test images.

2.2   Boosting Classifier

Feature selection is performed via AdaBoost [7]. AdaBoost extracts in each iteration the weak classifier (feature width, location and sign) that best discriminates objects from background training histogram images. A weak classifier can be expressed as

$$h(s) = \begin{cases} 1 & : & s * f > t \\ 0 & : & \text{otherwise} \end{cases} \,, \tag{1}$$

where s is a training sample histogram, f is the feature being tested, with all its parameters (width, location and sign), $*$ indicates the convolution operation, and t is the response threshold. The algorithm selects the most discriminant weak classifier h, as well as its contribution α in classifying the entire training set, as a function of the classification error ϵ.

$$\alpha = \frac{1}{2} \ln \frac{1 - \epsilon}{\epsilon} \qquad (2)$$

At each iteration, the algorithm also updates a set of weights over the training set. Initially, all weights are set equally, but on each round, the weights of missclassified samples are increased so that the algorithm is forced to focus on such hard samples in the training set the previously chosen classifiers missed. In a certain way, the technique is similar to a Support Vector Machine, in that both search for a class separability hyperplane, although using different distance norms, l_2 for SVMs, and l_1 for boosting [8]. The dimensionality of the separating hyperplane in AdaBoost is given by the number N of weak classifiers that form the strong classifier:

$$H(s) = \begin{cases} 1 & : \quad \sum_{i=1}^{N} \alpha_i h_i(s) \geq \frac{1}{2} \sum_{i=1}^{N} \alpha_i \quad \text{object} \\ 0 & : \quad \text{otherwise} \qquad\qquad\qquad\quad \text{no-object} \end{cases} . \qquad (3)$$

2.3 Integral Histogram

An integral image is a representation of the image that allows fast computation of features because it does not work directly with the original image intensities. Instead, it works over an incrementally built image that adds feature values along rows and columns. Once computed this image representation, any one of the local features can be computed at any location and scale in constant time [4].

Extending the idea of having cumulative data at each pixel in the integral image, we have proposed to store on it the histogram data instead of intensity sums [5]. The integral histogram stores intensity level histograms which, once constructed, allow for the computation of histogram within a rectangular area in constant time.

The value of the integral histogram s at coordinates u, v contains the intensity histogram of the region above and to the left of u, v, inclusive,

$$s(u, v) = \sum_{i \leq u, j \leq v} s(i, j) . \qquad (4)$$

then, it is possible to compute for example, the intensity histogram in a rectangular region, called *Area*, simply by adding and subtracting the cumulative histograms at its four corners in the integral histogram representation (see figure 2),

$$histogram(Area) = s(A) + s(D) - s(B) - s(C) \qquad (5)$$
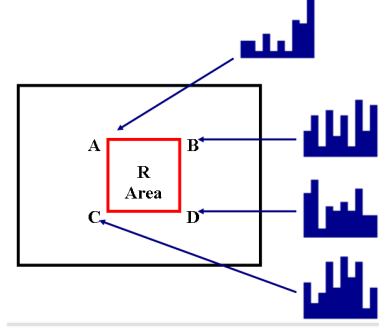
Fig. 2. The integral histogram.

Furthermore, the construction of the integral histogram is computed iteratively with

$$s(u,v,bin) = bin(I(u,v)) + s(u-1,v,bin) + s(u,v-1,bin) - s(u-1,v-1,bin) \tag{6}$$

where

$$bin(I(u,v)) = \begin{cases} 1 & : & I(u,v) \in bin \\ 0 & : & \text{otherwise} \end{cases} \tag{7}$$

3   Object Detection

We have decided to compare our proposed method with the known Swain and Ballard method [1] in terms of classification. The tests are based on patch images of both object and background outdoor scenes. The Swain and Ballard color intersection metric is defined as

$$\bigcap(H,T) = \sum_{i=1}^{m} \min(H(i), T(i)) \tag{8}$$

where $H(i)$ is the new class test histogram, $T(i)$ the reference histogram associated to the object image and m the number of bins. This method makes the comparison with such specific object histogram $H(i)$, that is, it only uses one canonical image to perform object detection.

Conversely, our method performs a object detection in a local manner, taking into account possible changes in the histograms, due to small object translation, non uniform illumination, scale and partial occlusions. As the learning process is carried out over a set of training histogram images, the selected weak classifiers become robust to small image transformations present in the training set. Some of the training histogram images that have been used for our proposed method are shown in the Figure 3a-l.
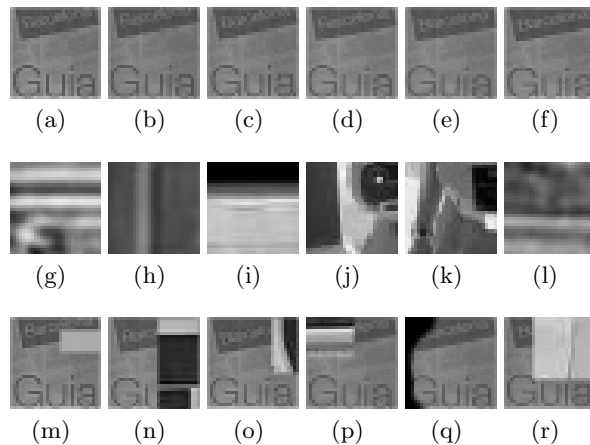
Fig. 3. Training set images. a-f) Object g-l) Background m-r) Occluded object

The first experiment consists in applying the measurement over a validation set of 30 object patch images. In the proposed method, we required a training set of 50 object images and 100 background images. The number of bins selected was of 12 as tradeoff between reliability and computation burden. The results appear in table.

The second experiment is carried over 300 background patch images, extracted from outdoor and indoor scenes, (none with the object). This test is performed to show the method performance to background scenes and its discrimination. One false positive is detected for our method (figure 4).

The third experiment was aimed at evaluating the descriptor robustness to mild occlusions. The table shows the results for the occluded object shown in Figure 3(m-r). Thanks to the local matching property of the proposed method, the correct detection is high.

Method	Correct	False negatives	False positives	Test
Proposed method	96%	4%	0%	Validation
Swain & Ballard	93%	6%	0%	
Proposed method	99%	0%	1%	Background
Swain & Ballard	85%	0%	15%	
Proposed method	90%	10%	0%	Occlusions
Swain & Ballard	73%	27%	0%	

The detection is performed by applying the strong classifier $H(s)$ over the entire test image, at every location and scale. Therefore the use of the integral histogram is of utmost importance in this hard task.
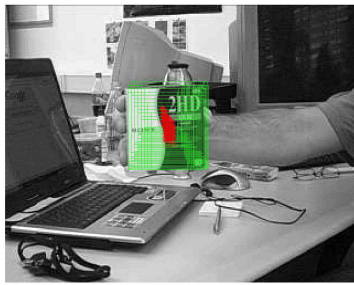
Some detection results are shown in Figure 5. We can appreciate that the detection is achieved even when the object presents several scales, locations and rotations in the plane.
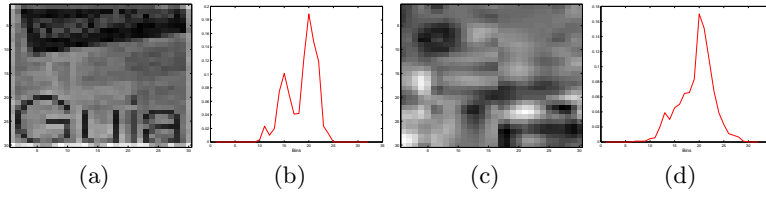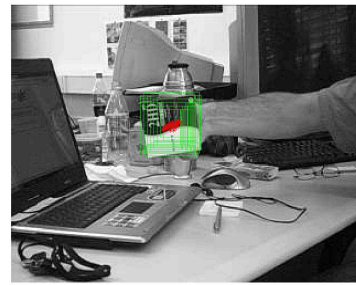
Fig. 4. False positive a) Object image b) Object histogram c) False positive patch d) False positive histogram
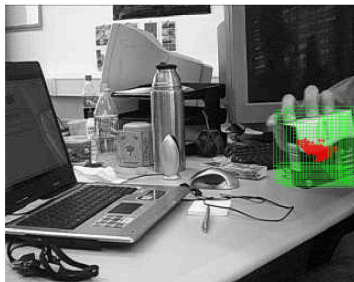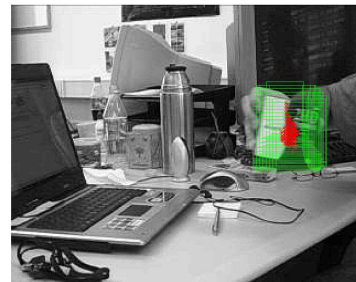

(a)

(b)

(c)

(d)

(e)

Fig. 5. Object detection a) Box object b-e) Object detection

4   Conclusions

We have presented an appearance method to perform object detection invariant to object scale and orientation changes, and robust under partial occlusions. The classification rule is based on Haar local features extracted during a Boosting training phase, giving a local measurement that accounts for small changes in the object histogram signature. In order to reduce the computational cost of performing object detection, the integral histogram has been incorporated.

References

1. M. Swain and D. Ballard. Color indexing. *Int. J. Comput. Vision*, 7(1):11–32, 1991.
2. F. Moreno, J. Andrade-Cetto, and A. Sanfeliu. Fusion of color and shape for object tracking under varying illumination. In F. J. Perales, A. J. C. Campilho, N. P. de la Blanca, and A. Sanfeliu, editors, *Pattern Recognition and Image Analysis*, volume 2652 of *Lect. Notes Comput. Sci.*, pages 580–588, Puerto de Andratx, Jun. 2003. Springer-Verlag.
3. B. Schiele and J.L. Crowley. Object recognition using multidimensional receptive field histograms. In B. Buxton and R. Cipolla, editors, *Proc. 4th European Conf. Comput. Vision*, volume 1065 of *Lect. Notes Comput. Sci.*, pages 610–619, Cambridge, Apr. 1996. Springer-Verlag.
4. P. Viola and M. Jones. Rapid object detection using a boosted cascade of simple features. In *Proc. 15th IEEE Conf. Comput. Vision Pattern Recog.*, pages 511–518, Kauai, Dec. 2001.
5. M. Villamizar, A. Sanfeliu, and J. Andrade-Cetto. Computation of rotation local invariant features using the integral image for real time object detection. In *Proc. 18th IAPR Int. Conf. Pattern Recog.*, volume 4, pages 81–85, Hong Kong, Aug. 2006.
6. F. Porikli. Integral histogram: a fast way to extract histograms in cartesian spaces. In *Proc. 19th IEEE Conf. Comput. Vision Pattern Recog.*, volume 1, pages 829–836, San Diego, Jun. 2005.
7. Y. Freund and R. E. Schapire. A decision-theoretic generalization of on-line learning and an application to boosting. *J. Comput. Syst. Sci.*, 55(1):119–139, Aug. 1997.
8. G. Rätsch, B. Schölkopf, S. Mika, and K.-R. Müller. SVM and Boosting: One class. Technical report, GMD First, Nov. 2000.