Classifying a New Descriptor Based on Marr's Visual Theory

J.M. Pérez-Lorenzo[1], S. García Galán[1], A. Bandera[2], R. Vázquez-Martín[2], and R. Marfil[2]

[1] Dept. Ing. Telecomunicación, Universidad de Jaén, Linares-23700, Spain
jmperez@ujaen.es
[2] Dept. Tecnología Electrónica, Universidad de Málaga, Málaga-29071, Spain

Abstract. Descriptors are a powerful tool in digital image analysis. Performance of tasks such as image matching and object recognition is strongly dependent on the visual descriptors that are used. The dimension of the descriptor has a direct impact on the time the analysis take, and less dimensions are desirable for fast matching. In this paper we use a type of region called curvilinear region. This approach is based on Marr's visual theory. Marr supposed that every object can be divided in its constituent parts, being this parts cylinders. So, we suppose also that in every image there must be curvilinear regions that are easy to detect. We propose a very short descriptor to use with these curvilinear regions in order to classify these regions for higher visual tasks.

1 Introduction

Image matching is a fundamental aspect of many problems in computer vision. Many approaches have been presented in the literature for achieving this goal [1] [3] [5] [8] . In a general way, an image matching method can be divided into three steps. In a first step a technique for searching visual regions with a high repeatability is employed. In a second one these regions are characterized with descriptors, and in the third step the descriptor vectors are matched between different images.

In [9] an approach to visual scenes matching was presented. In biological object recognition no much work has solved the question of specifity, that is, which properties of an object are exactly encoded in the neural representations used to recognize that object. Also the problem of feature selection is a fundamental question [10], although these features seem to be very sensitive to particular combinations of local shape, colour, and texture properties. The approach proposed in [9] was based in Marrs "Generalizad Cilinders" visual theory [7]. Marrs suggests that objects can be formed with an alphabet of simple geometrical shapes, being these shapes cylinders along the main axes of the object. By this way, objects can be represented in terms of their constituent parts.

So the scheme of image matching proposed in [9] defends that a particular type of regions, called curvilinear regions, can be easily detected in digital images, and those regions could be compared in their complexity to those features

J. Mira et al. (Eds.): IWINAC 2009, Part I, LNCS 5601, pp. 205–212, 2009.

analysed by the IT cortex for achieving objects recognition. The algorithm can be divided in several steps. Firstly the curvilinear regions are obtained thanks to a colour segmentation and the verification of some previously established geometrical properties, called curvilinear properties. As the result of this analysis step the method obtains the set of detected curvilinear regions. In a second step the normalized regions are employed as the input of a shape descriptor, which is a contour-based approach to object representation that uses a curvature function [2]. In the third step the recognition stage matches the obtained individual features to a database of features from known scenes using a nearest-neighbour algorithm. This algorithm uses the curvature description, colour description and position of the region in the image. The length of the descriptor used in the matching step was of 260 parameters.

In a general way, the dimension of the descriptor has a direct impact on the time the matches take, and less dimensions are desirable for fast interest point matching [1]. However, lower dimensional feature vectors are in general less distinctive than their high-dimensional counterparts. So, a good goal is to develop both a detector and descriptor that are fast to compute while not sacrificing performance [1].

In this paper a new descriptor is employed. The descriptor uses the colour parameters of the segmentation detection and the parameters that are computed for each region to decide whether it is a curvilinear region or not. So, this descriptor is a vector formed by colour and geometric properties. In the matching step a classification based on standard statistical pattern recognition has been employed. Experiments show that some visual objects can be correctly matched by using a descriptor with just a few parameters. The main difference of the work presented in this paper with the work in [9] is that the descriptor and matching steps have been replaced by new steps that use a much shorter descriptor, resulting in a much more efficient performance of the system. Although only a few experiments have been achieved we believe that they are promising for a further and deeper research in this visual matching method.

The rest of the paper is organized as follows: Section 2 briefly defines a curvilinear region and its properties. Section 3 describes the implementation of the curvilinear region detector. Section 4 presents the descriptor we use in this work for characterizing the curvilinear regions. Section 5 presents the methods we have used for the classification stage. Experimental results for the whole strategy is presented in Section 6. The paper concludes along with conclusions and future work in Section 7.

2 Curvilinear Regions Definition

We define a curvilinear region as a parameter vector $\{a(l), w_l(l), w_r(l)\}_{l=0..L}$, being L the length of the region, $a(l)$ a vector defining the axis between the right and left borders ($b_r(l)$ and $b_l(l)$), and $w_r(l)$ and $w_l(l)$ the widths of the curvilinear region (see Fig. 1).The curvilinear conditions to be satisfied by these regions are:

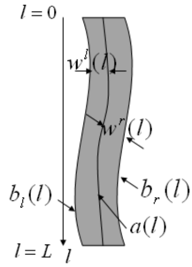

Fig. 1. Definition of curvilinear region. $b_l(l)$,$b_r(l)$: left and right borders, $a(l)$: medial axis, $w_l(l)$,$w_r(l)$: left and right widths.

i) Geometric similarity around the region axis
ii) The ratio between its average width and its total length must be less than a predefined threshold
iii) Left and right borders must be locally parallel
iv) Colour along this axis should be homogeneous.

3 Curvilinear Regions Detector

The implementation of the curvilinear region detector uses in a first step a segmentation in the HSI colour space, based on the Bounded Irregular Pyramid (BIP) [6]. With this segmentation the method ensures that the obtained regions comply with the homogeneous colour property of the curvilinear regions. Next, properties i), ii) and iii) are checked based on some geometrical parameters of the regions. For the estimation of the parameters, the skeleton of the region is extracted with the algorithm based on the d8-distance described in [4], which can approximate the distance transform inside the region in only two steps. The skeletons are generally not connected, so a post-process is needed before estimating further parameters. The skeleton obtained is defined as the set of connected pixels $p_s = (i_s, j_s), 0 \leq s \leq N-1$, being N the number of pixels being evaluated of the skeleton. In Fig. 2.a, 2.b and 2.c an original image, the segmentation image and the estimated skeletons are represented.

i) Symmetry around the skeleton. The method looks for those pixels which comply with the requirement of symmetry around the axis. The normal vector is calculated for each pixel ps in the skeleton, and the cross-points between the normal and the left and right borders of the region are estimated. If we define p_s^l and p_s^r as these cross-points, then we obtain the triplets $(p_s, p_s^l, p_s^r), 0 \leq s \leq N-1$. The symmetry condition can be defined as:

$$E_{\Delta w} \leq U_{\Delta w}(1 - e^{-\frac{N^2}{2\sigma^2_{\Delta w}}}) \tag{1}$$

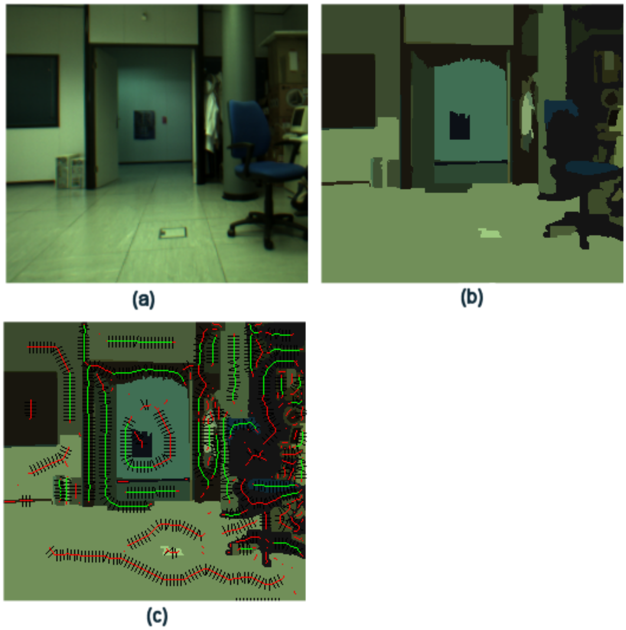

Fig. 2. Estimated skeletons over the segmented image. a) Original image. b) Segmented image. c) Estimated skeletons with the estimated normal vectors.

with

$$E_{\Delta w} = \frac{1}{N}\sum_{s=0}^{N-1}(\Delta w_s - \overline{\Delta w})^2 \tag{2}$$

$$\Delta w_s = |w_s^l - w_s^r| \tag{3}$$

$$\overline{\Delta w} = \frac{1}{N}\sum_{s=0}^{N-1}\Delta w_s \tag{4}$$

being w_s^l the Euclidean distance between pixels p_s and p_s^l and w_s^r the Euclidean distance between pixels p_s and p_s^r. $U_{\Delta w}$ is a parameter of the method.

ii) Ratio. Given a position s in the skeleton, the width w_s of the region is estimated as the Euclidean distance between pixels p_s^l and p_s^r. The following condition must be satisfied:

$$L_{max} \geq U_w \frac{1}{N}\sum_{s=0}^{N-1} w_s \tag{5}$$

being L_{max} the maximum length that the curvilinear region could have, which is estimated with all the connected pixels of the skeleton. U_w is a parameter of the method.

iii) Borders Parallelism. To check the borders parallelism requirement we estimate the tangential vectors on the borders at pixels p^l_s and p^r_s, and then we calculate the angle between those vectors and the normal vector given a position s, obtaining α^l_s and α^r_s. The following condition must be satisfied:

$$\overline{\Delta\alpha} \leq U_{\Delta\alpha} \quad (6)$$

being

$$\overline{\Delta\alpha} = \frac{1}{N}\sum_{s=0}^{N-1} |\alpha^l_s - \alpha^r_s| \quad (7)$$

and $U_{\Delta\alpha}$ is a parameter of the method.

The detector tries to join as many pixels as possible to form a curvilinear skeleton. For doing that, the algorithm starts in an endpoint of the skeleton and it looks for adding the connected pixels checking if Ec. 1, Ec. 5 and Ec. 6 are true with the new added pixel. When all the pixels have been evaluated inside a region, the curvilinear skeletons with close endpoints are linked. Those parts of the objects with a skeleton evaluated as a curvilinear skeleton are considered curvilinear region by the detector. In our experiments, we demand that these regions must have a minimum length of 10 pixels.

4 Curvilinear Regions Descriptor

The purpose of this work is trying to use a descriptor as shorter as possible and being able to match visual regions in a proper way. In [9] we decided to use a 260-dimensional descriptor, but in this new approach we suggest to use some of the previously estimated parameters as part of the regions descriptor.

In [9] we used a curvature function to characterize the region boundaries. In this new approach we apply this curvature function to the skeleton of the region and calculate:

$$C_T = \sum_{s=0}^{N-1} fc_s \quad (8)$$

$$\overline{C_T} = \frac{C_T}{N} \quad (9)$$

being fc_s the curvature function of the object shape. An angle-based curvature estimator is used, where the curve orientation is estimated at each point with respect to a reference direction. The approach is based on a k-slope algorithm which estimates the curvature using a k value which is adaptively changed [2]. C_T and $\overline{C_T}$ are part of the region descriptor.

The same geometrical parameters that are used to decide if the region complies with the curvilinear properties, plus some of the colour values of the segmented region and the curvature function mean values, form the region descriptor. Specifically, the descriptor has got the next parameters:

i) H (Hue value of the colour of the region),
ii) S (Saturation value of the colour of the region),
iii) $E_{\Delta w}$ (Estimated value in Ec. 2),
iv) $\overline{\Delta w}$ (Estimated value in Ec. 4),
v) $\overline{\Delta \alpha}$ (Estimated value in Ec. 7),
vi) C_T (Estimated value in Ec. 8) and
vii) $\overline{C_T}$ (Estimated value in Ec. 9)

5 Classification Stage: Gaussian Mixture Model-Based Classifier

For classifications purposes, a number of standard statistical pattern recognition (SPR) exists. The basic idea behind SPR is to estimate the probability density function (pdf) for the feature vectors of each class. In the Gaussian Mixture Model (GMM) classifier, each class pdf is assumed to consist of a mixture of a specific number K of multidimensional Gaussian distributions. The GMM classifier only needs to store the set of estimated parameters for each class. The Expectation-Maximization (EM) algorithm is used to estimate the parameters of each Gaussian component and the mixture weights.

We have used a three-component GMM classifier with diagonal covariance matrices because it has shown a slightly better performance than other SPR classifiers. The performance of the system does not improve when using a higher number of components in the GMM classifier. The GMM classifier is initialized using the K-means algorithm with multiple random starting points.

6 Experimental Results

For the experiment results an indoor environment has been captured in 30 images. Four different curvilinear regions were detected in each of the 30 images. In Fig. 3, we can see the four regions of interest detected in 4 of the 30 images. So, the experiments were focused to match those regions and see the classification success.

The classifications results were calculated using a ten-fold cross-validation evaluation where the dataset to be evaluated was randomly partitioned, so that 10% was used for testing and 90% was used for training. The process was iterated with different random partitions and the results were averaged.

The algorithm to classify one region uses a nearest neighbour-based matching strategy. Once the classifier has been trained, the algorithm estimates the membership of the region to be tested with every region class. If the probability index is above a threshold, the region is considered to be a member of the class. If the region to be tested belongs to more than one class, then the algorithm decides that the correct class is the one with the highest probability.

In our experiment the average ratio of success of the ten-fold cross-validation evaluation has been 94.46%.

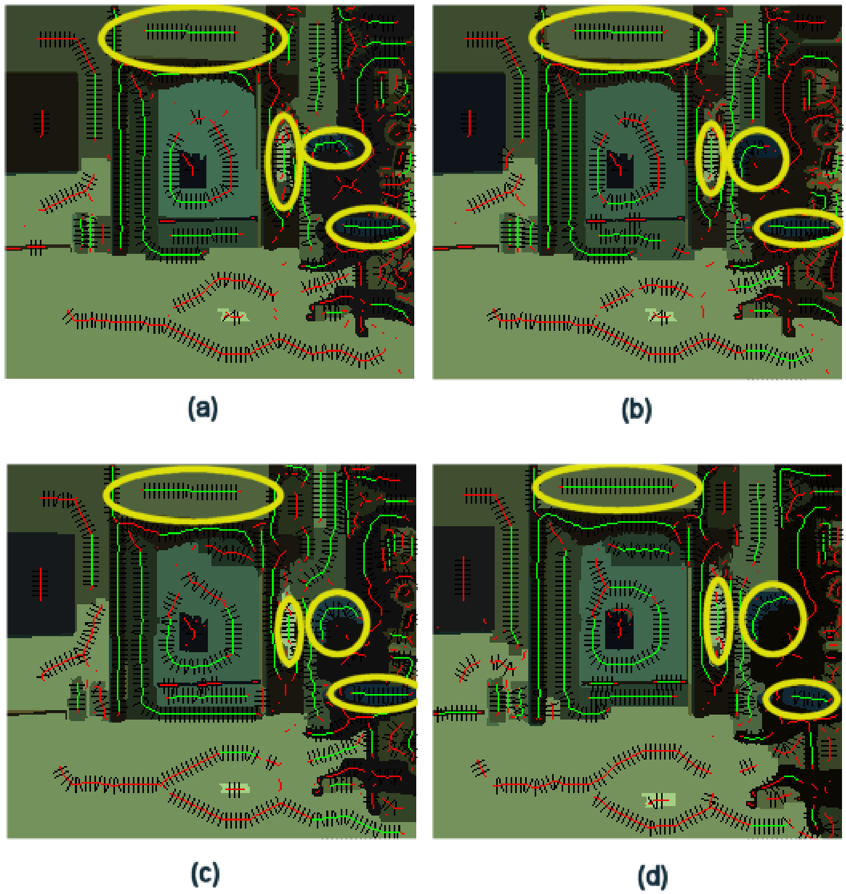

Fig. 3. Images captured from the same scene with slightly different viewpoints. The four curvilinear regions were detected in every image. These four regions were used to test the classifier.

7 Conclusions and Future Work

We have continued the work presented in [9]. Our work propose a scheme of image matching based on the detection of curvilinear regions in the images. In this paper we have used a very short curvilinear region descriptor in order to get a better performance in the matching stage. This descriptor is formed by the values of a colour segmentation and a few geometrical features of the regions. We have tested this descriptor in an indoor environment, where several regions have been matched with a high success percentage in a set of 30 images.

However, the results presented still belong to a preliminary stage of the experiments. More experiments should be taken in more diverse scenes in order to test the repeatability of the detection and the performance of the classifier, and a comparative study with other techniques would be also desirable. So, these issues should be our future work.

Acknowledgements

This work has been partially granted by the Spanish Junta de Andalucía project P07-TIC-02713 and P07-TIC-03106.

References

1. Bay, H., Tuytelaars, H., Van Gool, L.: SURF: Speeded up robust features. In: Leonardis, A., Bischof, H., Pinz, A. (eds.) ECCV 2006. LNCS, vol. 3951, pp. 404–417. Springer, Heidelberg (2006)
2. Bandera, A., Urdiales, C., Arrebola, F., Sandoval, F.: Corner detection by means of adaptively estimated curvature function. Electronics Letters 36(2), 124–126 (2000)
3. Ke, Y., Sukthankar, R.: PCA-SIFT: A more distinctive representation for local image descriptors. In: 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2004), vol. 2, pp. 506–513 (2004)
4. Klette, G.: A comparative discussion of distance transformation and simple deformations in digital image processing. Machine Graphics & Vision 12(2), 235–256 (2003)
5. Lowe, D.: Towards a computational model for object recognition in IT Cortex. In: First IEEE International Workshop on Biologically Motivated Computer Vision, Seoul, Korea, pp. 20–31 (2000)
6. Marfil, R., Rodriguez, J.A., Bandera, A., Sandoval, F.: Bounded irregular pyramid: a new structure for color image segmentation. Pattern Recognition 37(3), 623–626 (2004)
7. Marr, D., Nishihara, H.K.: Representation and recognition of the spatial organization of three-dimensional shapes. Proceedings of the Royal Society of London B: Biological Sciences 200, 269–294 (1978)
8. Matas, J., Chum, O., Urban, M., Pajdla, T.: Robust wide baseline stereo from maximally stable extremal regions. In: Proceedings of the British Machine Vision Conference, vol. 1, pp. 384–393 (2002)
9. Pérez-Lorenzo, J.M., Bandera, A., Reche-López, P., Marfil, R., Vázquez-Martín, R.: An approach to visual scenes matching with curvilinear regions. In: Mira, J., Álvarez, J.R. (eds.) IWINAC 2007. LNCS, vol. 4528, pp. 409–418. Springer, Heidelberg (2007)
10. Ullman, S., Vidal-Naquet, M., Sali, E.: Visual features of intermediate complexity and their use in classification. Nature Neuroscience 5, 682–687 (2002)