Proceedings of the International Workshop on Intelligent Information Systems
IIS'2011, September 13-14, 2011, Chisinau, Republic of Moldova

Novel Structural Elements for Image Descriptor

Andrej Agarkov

Abstract: It is proposed to describe the local features neighborhood by additional structural elements. These structural elements are grouped in the graph. This one captures the image structure and to facilitate its analysis. Such graph allows to restore the original image. It opens the possibility for conversion of natural gray and color images in vector form.

Keywords: local invariant features, descriptors, image recognition, objects detection.

1 Introduction

In recent years, recognition of images have become increasingly popular methods for applying the local invariant features. These ones are derived using such filters as the DoG and LoG [1-4]. This approach has been successfully used for objects tracking, mobile robots navigation, objects detection, image description and recognition.

To define each of local invariant features are used a variety of descriptors. They describe the brightness characteristics of the neighborhood around its center [1, 5, 6]. As a rule, used neighborhood equal to three features scale. This makes it impossible to use such descriptors to define the homogeneous objects with a weak texture. Also such descriptors do not allow to group local features as belonging to same homogeneous area or object.

In this paper, we propose to describe the local features neighborhood by additional structural elements. The nature of these elements is the same as that of the local features. It allows not to complicate the basic method greatly. These structural elements are grouped in the form of a graph. This one captures the image structure and to facilitate its analysis. These structural elements allow restoring the original image. It opens the possibility for conversion of natural gray and color images in vector form.

2 The allocation of additional structural elements

Local invariant features, which form the set E, correspond to the extremums of the function $D(x, y, \sigma)$. This function is an original image

multiscale representation produced by filters DoG or LoG. Position of the extremums determines the center coordinates and scale of the corresponding local features.

Additional structural elements correspond to the extremums in the direction of multiscale representation planes, i.e. functions $D(x, y, \sigma = const)$. These extremums correspond to the lengthy detail. As follows from the properties of DoG and LoG such details can be either border or lengthy homogeneous regions.

It is proposed the following procedure for the additional structural elements retrieval. In each multiscale representation plane $D(x, y, \sigma_i)$ extremums in the direction are searched. These ones form the set $T = \{p_j = p(x_j, y_j, \sigma_{i,j})\}$. This set is partitioned into disjoint subsets $\overline{T} = \{t_j : \bigcup_j t_j = T; \forall j, k,, j \neq k, t_j \cup t_k = \varnothing\}$. Each one consists of extremums with the same coordinates in the image plane and is arranged in a sequence of adjacent planes. These planes are corresponding to a sequence of smoothing scales σ:

$$t = \{p(x_j, y_j, \sigma_{i,j}) : \forall j, k, j \neq k, (x_j, y_j) = (x_k, y_k),;$$
$$\{\sigma_{i,j} : i = i_1, ..., i_{N_t}, i_{m+1} = i_m + 1\}\}$$

where N_t - number of extremums p in the set t. Each set $t_i \subset T$ forms an additional structural element. This structural element is characterized by the coordinates (x_i, y_i) in the image plane, scale $\sigma_{extr} = \arg\min_{p(x,y,\sigma) \in t_i} \max\left(D(x, y, \sigma)\right)$ (further - the character scale) and average brightness value (further - the character brightness).

It should be noted that $E \subset T$, since the extremum of the function $D(x, y, \sigma)$ is also an extremum in the direction for $D(x, y, \sigma = const)$. It allows to define the local features by additional structural elements located in the vicinity of their centers.

3 Descriptor

Local features descriptor is formed by additional structural elements. It should reflect their relative positions. For this purpose it is used graph. Each vertex of the graph is associated with a structural element. The edges

of the graph represent the relative positions of structural elements, including the neighboring relations between them. Because $E \subset T$, some of the vertices correspond to local features. Thus, formed graph $G_I^T(V^T, E^T)$ defines the whole of the original image. A separate local features descriptors correspond to the subgraphs formed by structural elements located in their vicinity.

Connected subgraphs of $G_I^T(V^T, E^T)$, formed through the neighborhood relationships, correspond to the lengthy homogeneous regions. It allows to find the structural relationship between the local features (accessory one homogeneous area, and other relationships, reflecting the structure of images and objects).

4 Image restoration

Because each structural element is a region with character brightness and center at some point in the image plane, it is possible to restore the original image. However, such an image would be devoided of small details that are smaller than the initial smoothing scale.

Since the aim is to investigate the possibility of restoring, it is used the most simple way to make it. It draws circles with character scale radius and character brightness in the center of each structural element. If the circles overlap, the pixels brightness determined by the character brightness of a structural element with smaller scale. If the scales are equal - the pixel brightness value is averaged.

Fig. 1 shows the original and restored images. The size of the original image is 384x288. The initial value of the scale of smoothing - $\sigma_0 = 1.4$.

a b

Figure 1. Image restoration. a) – original image, b) – restored image.

As seen from Fig. 1, the restored image, as expected, is devoided of small details that are smaller than the minimum scale of smoothing.

However, visual assessment shows that the result of restoration is not much different from the original.

5 Conclusion

In this paper it is proposed a method to extract the image structural elements. It is based on using the multiscale image representation. This representation is constructed by applying DoG or LoG filters. It is the same filter as for local invariant features extracting. Thus these structural elements define local invariant features neighborhood and relationships. It allows to use local features to define a homogeneous objects, i.e. expand their capabilities and applicability. Also, it allows to restore the original image by applying these elements. It makes possible the conversion of natural images into vector form.

References

[1] David G. Lowe. *Distinctive Image Features from Scale-Invariant Keypoints*. International Journal of Computer Vision, 2004, vol. 2, no 60, pp.91-110.

[2] David G. Lowe. *Local Feature View Clustering for3D Object Recognition*. In Proc. CVPR, Springer, 2001, pp.682-688.

[3] Hongli Deng,Wei Zhang,EricMortensen,Thomas Dietterich Linda Shapiro, *Principal Curvature-Based Region Detector for Object Recognition*. IEEE Conference on Computer Vision and Pattern Recognition, CVPR '07, 2007, pp.1-8.

[4] Wei Zhang, Hongli Deng, Thomas G. Dietterich and Eric N. Mortensen. *A Hierarchical Object Recognition System Based on Multi-scale Principal Curvature Regions*. 18th International Conference on Pattern Recognition, 2006, vol. 1, pp.778-782.

[5] Yan Ke, Rahul Sukthankar. *PCA-SIFT: A More Distinctive Representation for Local Image Descriptors*. IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004, vol. 2, pp.506-513.

[6] Lazebnik S., Schmid C., and Ponce J. *A Sparse Texture Representation Using Affine-Invariant Regions*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Madison, Wisconsin, USA, 2003, vol. II, pp.319-324.

Andrej Agarkov[1]

[1] Institute of Artificial Intelligence, Donetsk, Ukraine
E-mail: aav@iai.donetsk.ua