ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 7, Issue 10 , October 2020

Algorithm for Searching Objects in Images Using a Descriptor

Mukhammadi Shamiev Olimovich

Teacher of the Department of Software Engineering, Samarkand Branch of the Tashkent University of Information Technologies named after Muhammad al-Khwarizmi, Uzbekistan

ABSTRACT: This paper discusses the development of a descriptor for searching for objects in images. The interest in the issue of pattern recognition is great. Already today, you can meet the fruits of scientific discoveries in everyday life: software for text recognition, systems for recognizing faces, license plates of cars passing on the highway, as well as systems for recognizing a person's emotional state by facial expression. Due to the complexity of the tasks set, highly specialized recognition systems are widely used. Such systems are much easier to make: smaller subject area and simpler system requirements.

KEY WORDS: image analysis, object search, structural descriptor, graph nesting, computer vision

I.INTRODUCTION

Searching for objects in images is an urgent task in various areas of computer vision: human-machine interfaces, biometrics, Earth remote sensing, image databases, technological process control systems, etc.

Objects are often searched based on matching using descriptors. A descriptor is a method that identifies a certain area of an image based on a set of features. The following groups of two-dimensional image descriptors are distinguished [1, 2]: local binary descriptors, descriptors based on spectral representation, descriptors based on basic functions, shape descriptors. It should be noted that some methods, according to their characteristics, can be attributed to different groups at the same time.

Local binary descriptors represent a description of a small area of the image in the form of binary vectors. The best known local descriptors are local binary patterns and their modifications (Local Binary Patterns - LBP) [3]. This group also includes BRIEF (Binary Robust Independent Elementary Features) [2], ORB (Oriented BRIEF), BRISK (Binary Robust Invariant Scalable Keypoints) [1], etc.

Spectral representation-based descriptors use different quantities to identify regions: intensity, color, gradients, statistical characteristics, etc. Spectral-based descriptors are more computationally intensive than binary descriptors. These methods include the following descriptors and their modifications: SIFT (Scale Invariant Feature Transform), SURF (Speeded Up Robust Features) [1], DAISY, HoG (Histogram of Gradients), correlation templates (Sum of Absolute Differences (SAD), Sum of Squared Differences (SSD), Normalized Cross Correlation (NCC)), Local Gradient Pattern (LGP) [2], Freeman's code (Chain Code Histograms (CCH)) [3], Haar features (HAAR Features) [1], etc.

Basis function descriptors represent the description of an image in given spaces. The most famous descriptor based on basis functions is the Fourier descriptor. This group of descriptors can also include methods of sparse coding (Sparse Coding) [1]. An example of a descriptor in this group is the Bag of Words method. Sparse coding descriptors use a set of codes to identify objects instead of basic functions.

The disadvantage of many approaches is that in computer vision systems it is often difficult to distinguish the boundaries of the analyzed object due to the effects of a sharp change in lighting, mutual overlap, a complex textured background, a change in perspective, etc.

The paper proposes a structural descriptor that identifies an area consisting of hundreds and thousands of pixels. This area is comparable to images of objects searched for in computer vision systems. To construct a descriptor, it is proposed to use a graph whose vertices are the features of the image. In previous works, the issues of matching, clustering of features and object detection were considered on the basis of spectral graph theory. In the presented work, the novelty is the following: the use of a descriptor for detecting faces in images is considered, and the accuracy of detecting faces from the angle of rotation of a person's head in space is studied.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 7, Issue 10 , October 2020

II.DEVELOPMENT OF A GRAPH-BASED STRUCTURAL DESCRIPTOR

The advantage of structural approaches is that they allow you to analyze a large variety of elements based on a small number of simple components and rules for the formation of a graphical model. Also, structural methods allow you to describe those characteristics of an object that exclude its assignment to another class, which increases the reliability of recognition.

The proposed structural descriptor includes the construction of graphs of the reference image and the current image, the embedding of graphs in the vector space and classification.

III. BUILDING A GRAPH FROM AN IMAGE

The input of the descriptor receives the current image and the reference image obtained at the initialization stage. It should be noted that when objects rotate in space, some features in different images will disappear. Therefore, to improve the reliability of the search, it is proposed to track the centers of mass of the feature segments. When the object is rotated, the set of special points in such areas may change, but the area itself will be present in the image (Fig. 1). Features are highlighted using the SURF detector [1].

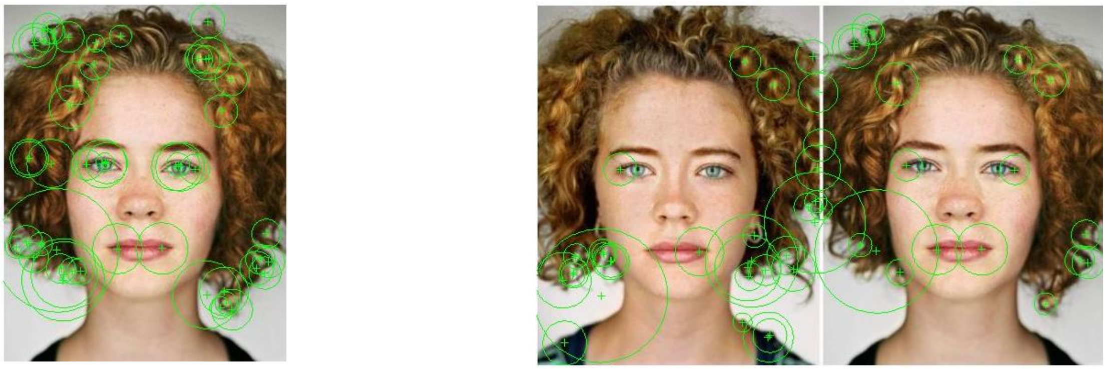

Figure: 1. Clustering of features highlighted in the image of a human face by the SURF detector

Clustering of singular points is carried out. During this process, all the special points selected in the image are grouped based on the method of connected components. In this case, a graph is constructed based on the Delaunay triangulation. The W parameter is set, and all edges in the graph whose length is greater than W. To select the W parameter, a histogram of distributions of pairwise distances between vertices is constructed. For images with a well-pronounced cluster structure, the histogram will have two peaks, one of which corresponds to intra-cluster distances, the second - to inter-cluster distances. The W parameter is selected from the minimum zone between these peaks. Only the closest peaks remain connected. If in any received segment there are few features (less than 3), then it is not considered further. Thus, segments are formed that correspond to the most characteristic areas of the image.
After clustering, the image graph is built. The centers of mass of the obtained segments are selected as the vertices. The graph is constructed based on the Delaunay triangulation.
The normalized Laplace matrix of the graph is calculated based on the expression (1):

$$L_n=\begin{cases} 1, if\ u = v\ and d_v \neq 0; \\ -\frac{1}{\sqrt{d_u d_v}}, if\ A(u,v) = 1 \\ 0, inothercases \end{cases} \quad (1)$$

where d_u , $d_v$are the degrees of the vertices u and v, respectively.

ISSN: 2350-0328
International Journal of Advanced Research in Science, Engineering and Technology
Vol. 7, Issue 10 , October 2020

IV.EMBEDDING OF GRAPHS IN VECTOR SPACE AND CLASSIFICATION OF OBJECTS IN IMAGES.

To search for objects, it is proposed to embed graphs into a vector space. This will make it possible to represent images in the form of vectors of numerical characteristics, which will enable an approximate comparison of structures that does not require exact comparison of graphs. The advantage of the developed approach is the invariance to rotation of the image on the plane, since the spectral characteristics of the graph do not depend on the marking of its vertices.
To compare the candidate image with the reference, it is proposed to use the spectral graph theory. To solve this problem, the graphs constructed from the compared images are embedded into the vector space. In this case, the graphs of the compared images are converted into a vector of numerical characteristics, on the basis of which the comparison is performed.
Based on the decomposition of the normalized Laplace matrix, the spectral characteristics of the graph are calculated:

$$L_n=\Phi\Lambda\Phi^T, \qquad (2)$$

where Λ is the dioganal eigenvalue matrix $\lambda_1, \lambda_2, \lambda_3 ... \lambda_{|v|}$;
Φ - matrix of eigenvectors $\varphi_1, \varphi_2, \varphi_3 ... \varphi_{|v|}$.
The heat core is a solution to the equation and is calculated using eigenvalues and eigenvectors:

$$H_t(u,v)=\sum_{i=1}^{|v|} e^{-\lambda t}\, \varphi_i(u)\varphi_i(v) \qquad (3)$$

Therefore, for the vertex u, the coordinate vector is calculated as follows:

$$y_u=(e^{\frac{\lambda_1 t}{2}}\varphi_1(u), e^{\frac{\lambda_2 t}{2}}\varphi_2(u), \dots, e^{\frac{\lambda_{|v|} t}{2}}\varphi_{|v|}(u)\ ) \qquad (4)$$

we get the following curvature value for points *u* and *v*:

$$k(u,v)=\sqrt{24(1-d_E)} \qquad (5)$$

Thus, the distances between the graphs can be described using the Hausdorff metric:

$$HD(G_1,G_2)=\max_{i\epsilon V_1} \max_{j\epsilon V_1} \min_{I\epsilon V_2} \min_{J\epsilon V_2} \|k_2(I,J)-K_1(i,j)\| \qquad (6)$$

V.ALGORITHM FOR THE IMPLEMENTATION OF A STRUCTURAL DESCRIPTOR BASED ON GRAPHS

The graph-based structural descriptor implementation algorithm consists of the following steps.

Step 1. The features of the candidate image and the reference image are highlighted.
Step 2. Clustering of features of two images is carried out.
Step 3. Delaunay graphs are constructed based on the centers of mass of the segments. The normalized Laplace matrices (1) and heat cores (3) are calculated.
Step 4. Using the Young-House-Holder transformation, the coordinates of the graph vertices are projected into the vector space (4).
Step 5. The values of the curves connecting the points are calculated. Curvature matrices are constructed (5).
Step 6. The matrix of distances between the characteristics of the reference image graph and the analyzed image graph is calculated (6).
Step 7. To visualize the data, the multidimensional scaling method is used, which makes it possible to display the degree of closeness of the candidate image to any object from the reference database.
Step 8. Clustering of points occurs in a space of lower dimension based on the method of connected components. The inclusion of points in the cluster indicates the assignment of images of objects to a certain image-reference.

ISSN: 2350-0328

International Journal of Advanced Research in Science, Engineering and Technology

Vol. 7, Issue 10 , October 2020

VI. CONCLUSION

When studying the developed approach, it was revealed that the descriptor has invariance to image rotation on a plane, as well as the ability to detect objects with a rotation angle in space of up to 50 °. It was also revealed that the use of the centers of mass of the feature segments as the vertices of the graph significantly increases the stability of the approach when changing the camera angle of the object.

REFERENCES

[1]. T. Ojala, M. Pietikäinen, D. Hardwood, Comparative study of texture measures with classification based on feature distributions / Pattern Recognition. 1996. – Vol. 29, Issue 1. – P. 51-59. – DOI: 10.1016/0031-3203(95)00067-4.

[2]. E. Rublee, V. Rabaud, K. Konolige, G. Bradski, An efficient alternative to SIFT or SURF / // IEEE International Conference on Computer Vision (ICCV). – 2011. – P. 2564-2571. – DOI: 10.1109/ICCV.2011.6126544.

[3]. Barinov, A.E. Constrained Spectral Clustering Algorithm for Selecting a Human Face in Images / Dynamics of systems, mechanisms and machines. – 2016. – V. 2, № 1. – P. 222-228.