PCA based modification of SIFT-like methods for object class recognition

Agnieszka Owczarek

Institute of Electronics, Technical University of Lodz,
211/215 Wolczanska str., 90-924 Lodz, e-mail: agnieszka.jadwiga.owczarek@p.lodz.pl

This paper discusses a novel PCA based modification of standard SIFT and PCA-SIFT algorithms for the purpose of object class recognition. New descriptors intended to be simultaneously distinctive enough to describe the difference between features belonging to separate categories and general enough to capture the variations among features from the same class are proposed.

The experimental results, gained for a test database, showing the reliability of introduced approach are presented.

Keywords and phrases: object recognition, principle component analysis, SIFT.

Introduction

Object recognition is a fundamental aspect of majority of visual systems, starting from autonomous mobile platform control systems thorough navigation systems aiding disabled people and ending on variety of monitoring applications. In such systems detection of group of objects belonging to the same category very often is more crucial than recognition of one strictly defined representative of them. For instance, in automatic control systems for mobile platforms extraction of all objects belonging for example to a class of 'tree', 'bench' or 'human' may let easily to create map of obstacles and consequently give great contribution to algorithms of safe path planning.

The main challenge in recognizing object categories comes up with reliable model that can described common, for all object belonging to the given category, features and simultaneously be flexible enough to cover the single object variability (caused e.g. by different size, pose, changing appearance due to lighting conditions or viewpoint transformations). Several approaches that deal with this problems can be found in literature [1–4]. As all these solutions are different in the sense e.g. of number of properties that should described class of object or their types, all of them put great impact on the representation. It is said that representation must contains the complex information concerning the features collection with respect to their description, distinctive appearance and special mutual position.

In this paper focus is put mainly on the aspect of single feature description. Thus, the spatial relationships between different features belonging to objects of the same category as well as the way the interests points are localized in the image are not taken into consideration. In order to find key points location SIFT algorithm proposed by D. Low in [5] is used. To find the best description of features for object class recognition two PCA based modification of standard SIFT descriptors are proposed and their effectiveness is compared. The first one introduces the direct SIFT descriptor length reduction using PCA analysis. The second one was inspired by Ke in [6]. In this method instead of using SIFT`s smoothed weighted histograms, the normalized gradient patches are used. In both cases PCA analysis is used to transform high dimensional space of correlated properties into low dimensional space of features that are common for all objects from analyzed category.

The remainder of this paper is organized as follows. Section 1 reviews the relevant aspects of modified algorithms: SIFT and PCA-SIFT. Sections 2 presents Principle Components Analysis technique. Section 3 details proposed modification. In section 4 experimental results are provided. Finally, Section 5 summarizes the contributions of this paper.

Review of modified algorithms

SIFT algorithm is said to be the most reliable method in the task of automatic object recognition. Using it the best matching results can be obtained. No wonder that since it has been introduced by Low many its modifications were proposed. Below standard SIFT descriptor and one of its alternatives is presented.

SIFT Algorithm

As described in [5] SIFT algorithms includes four stages of computation: scale-space peak selection, keypoint localization, orientation assignment and keypoint descriptor. In the first stage potential points of interest are identified by searching for local picks in a series of difference-of-Gaussian images (DoG). In the second stage, previously found points are localized with sub-pixel accuracy and reduced if found to be not stable. In third one dominant orientations for each keypoint is identified ensuring the invariance to transformations. In final stage, local image descriptor based upon the image gradients histograms in its local neighborhood is build. A 4 × 4 array of gradient histograms, each with 8 orientation bins, created by sampling the magnitudes and orientations of the image gradient in the patch around the keypoint, captures the rough spatial structure of the feature patch. Such 128-element vector after normalization to unit length becomes the final descriptor.

PCA-SIFT Algorithm

The PCA-SIFT descriptor introduced in [6] is based on the same input data as standard SIFT descriptor. It means that outputs from the first three stages (the sub-pixel location, scale, and dominant orientations) are common for both algorithms. The difference lies in the local patches descriptions. In case of PCA-SIFT method, patches of the size 41 × 41 centered over the keypoint localization and rotated to align its dominant orientation to a canonical direction are extracted from the given scale. Then, for every patch horizontal and vertical gradient maps are computed. Such maps are subsequently concatenated into vector with 2 × 39 × 39 = 3042 elements and projected onto low-dimensional space using eigenspace. In order to build the eigenspace PCA algorithm is used. Top n eigenvectors (experimentally set to 20) of covariance matrix of numerous collection of 3042-element vectors derived form gradient patches forms the eigenspace. According to authors, PCA-SIFT algorithm is both significantly more accurate and much faster than the standard SIFT.

Principle Component Analysis

Principle Component Analysis (PCA) is a standard technique used for dimensionality reduction [7]. Since it was invented in 1901 by Karl Pearson it has found many applications in computer vision problems. The most popular one is a concept of eigenfaces described in [8].

The main idea of PCA is to transform a number of possible correlated variables into a smaller number of uncorrelated variables. Mathematically it means decomposition of the covariance matrix, created for a set of observations, according to following formula:

$$C \cdot V_e = D \cdot V_e \tag{1}$$

where: C — covariance matrix,
V_e — eigenvectors matrix of C,
D — diagonal matrix of eigenvalues of C.

This operation satisfies the rule that the greatest variance among the observations corresponds to the first principle component (the eigenvector with the largest eigenvalue), the second one to the second principle component and so on. Therefore, taking only few first components and projecting analyzed data into the new space formed by them, the dimensionality reduction with the high information retaining can be achieved.

Proposed approach

Proposed solution is derived from the algorithms shortly described in section 1 of this paper with a combination of a novel, in the sense of the object categories recognition task, approach to PCA technique. Opposite to the object recognition, for object class recognition the descriptors that the best described what is common for all objects belonging to the category and make it distinctive among other classes, rather than what is unique for a single representative are demand. Therefore, in proposed method, using PCA technique, eigenvectors corresponding to the highest eigenvalues are discarded in favour of the ones characterized by smaller eigenvalues. Thanks to it features with the highest variation among the class, deciding about the single object specific properties, are not taken into account. Below, analyzed algorithms with introduced modification are described in more details.

The first stage is to create a pattern eigenspace. In order to do it, a set of test images containing different representatives of given category are analyzed. Using standard SIFT or PCA-SIFT algorithms corresponding local features descriptors are computed. Afterwards, on the modified covariance matrix of these descriptors (introduce in [8] for a dataset containing much more features that observations) PCA is applied. Then, as mention in the first paragraph of this section, eigenvectors with the highest associated eigenvalues are discarded (experimentally set that eigenvectors containing ca. 90–95% of information should be got rid of) and the fixed amount of top remanding ones (marked as 'N') builds the pattern eigenspace.

Fig. 1. The exemplary face images from test set with marked point of interest.

Fig. 2. Final results of class of object detection for proposed modification of PCA-SIFT algorithm (four images on the left) and SIFT algorithm (four images on the right).

In the next stage, from the previously computed eigenspace, the pattern descriptor is obtain. To do it, firstly the collection of standard input descriptors is projected into a new space and then the mean values of every feature in the new space forms the new low-dimensional descriptor (the length of the descriptor is determined by the number of eigenvectors 'N' on which eigenspace is based on). These two stages can be pre-computed once for every feature and their finals results (reduced eigenspace and pattern descriptor) stored for future calculation.

Once the pattern eigenspaces and descriptors are created, searched features can be identified in random images containing demand class of objects. Hence, in a new image (not belonging to the test set of objects) standard local features descriptors (SIFT or PCA-SIFT) are computed and projected into our eigenspace. As a result, a new descriptors are obtained that in next stage are compared with pattern one.

To determine whether two descriptors belong to the same feature, presenting the same class of objects in different images, the Euclidean distance between descriptors is used. Two vectors are said to represent the same feature if the difference between the first and the second smallest distance is higher than 80% of the second one (like in a standard SIFT algorithm).

Experimental results

To assess the viability of proposed approached to recognition of class of object, the experiment with features located on the human faces was performed. The database containing images of 153 individuals (both man and female) was used (faces94 database — see acknowledgment). From this database a test set of images was selected (34 faces) and corresponding to the same feature (in my experiment located in the nose hole) standard SIFT and PCA-SIFT descriptors were computed and transformed into class-of-object related ones using the methodology presented in section 3 of this paper. The exemplary face images with marked interested feature are presented in Fig. 1.

To reduce a dimensionality and make descriptor more general for representation of the hole class of objects, in PCA analyzes sixteen top eigenvectors containing 95% of variability were discarded. As a result a new pattern eigenspace was created using ten top eigenvectors from remaining set. Hence, the size of final pattern descriptor was reduced to 10 elements (e.g. from 128 of standard SIFT descriptor).

Figure 2 presents final result of class of object detection (four images on the left corresponds to PCA-SIFT method, and four on the right to the SIFT algorithm). Five the best matchings to the pattern are marked with appropriate numbers. In Table1 comparison of reliability for both modification is shown.

Table 1. Comparison of reliability of proposed modification of SIFT and PCA-SIFT algorithms.

Algorithm	Found in 5 the best	Found as the best one (1st one)	Not found (among 5 the best)
PCA_SIFT	45%	17%	55%
SIFT	66%	35%	34%

Conclusion and future work

The modification, introduced for the recognition of class of object, of algorithms generally used for object detection was presented in former sections. Gain results show that reliability of proposed modification for both analyzed algorithms is too poor to be considered as the best solution in unchanged form (better results were achieved for standard SIFT algorithms with the greatest recognition on the level of 66% for images from the same database as the test images), but probably can be taken into account as a part of more complex algorithms, in which class of objects is described by more features with respect to their special relationships. In such approaches proposed modification can be useful as a compact and timely effective way of describing a single characteristic point (but it must be verified).

Acknowledgments

This research has made use of the faces94 face database provided by Dr Libor Spacek from University of Essex (http://cswww.essex.ac.uk/mv/allfaces/faces94.html).

References

[1] Amit, Y., and D. Geman. ”A computational model for visual selection”. *Neural Computation* 11(7), 1999: 1691--1715.

[2] Borenstein, E., and S. Ullman. “Class-specific, top--down segmentation”. *Proceedings of ECCV* 2002: 109--124.

[3] Burl, M., M. Weber, and P. Perona. “A probabilistic approach to object recognition using local photometry and global geometry”. *Proceedings of ECCV* 1998: 628--641.

[4] Schneiderman, H., and T. Kanade. “A statistical method for 3d object detection applied to faces and cars”. *Proceedings of CVPR* 2000.

[5] Lowe, D.G. “Distinctive Image Features from Scale-Invariant Keypoints”. *International Journal of Computer Vision* 2004: 91–110.

[6] Ke, Y., and R. Sukthankar. “PCA-SIFT: A More Distinctive Representation for Local Image Descriptors”. *2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04)* 2 (2004): 506--513.

[7] Joliffe, I.T. *Principal Component Analysis.* Springer-Verlag: 1986.

[8] Turk, M., and A. Pentland. “Face recognition using eigenfaces”. *Proceedings of Computer Vision and Pattern Recognition* 1991: 586–591.