Global Selection vs Local Ordering of Color SIFT Independent Components for Object/Scene Classification

Dan-ni AI†, Student Member, Xian-hua HAN††, Nonmember, Guifang DUAN†, Member, Xiang RUAN†††, Nonmember, and Yen-wei CHEN†††, Member

SUMMARY This paper addresses the problem of ordering the color SIFT descriptors in the independent component analysis for image classification. Component ordering is of great importance for image classification, since it is the foundation of feature selection. To select distinctive and compact independent components (IC) of the color SIFT descriptors, we propose two ordering approaches based on local variation, named as the localization-based IC ordering and the sparseness-based IC ordering. We evaluate the performance of proposed methods, the conventional IC selection method (global variation based components selection) and original color SIFT descriptors on object and scene databases, and obtain the following two main results. First, the proposed methods are able to obtain acceptable classification results in comparison with original color SIFT descriptors. Second, the highest classification rate can be obtained by using the global selection method in the scene database, while the local ordering methods give the best performance for the object database.

key words: independent component analysis, localization, sparseness, SIFT, color, object/scene classification

1. Introduction

Image classification is one of the most important problems in the fields of computer vision and pattern recognition. As a fundamental technique for image analysis and visual searching, it has received considerable attention [1]. To achieve high accuracy, image classification requires distinctive image features for effective image representation.

Recently, many features have been proposed for image classification [11]. Among them, the scale-invariant feature transform (SIFT) descriptor which was proposed by Lowe [10], has significant impact on representing the distinctive characteristics of images. It extracts distinctive features from a gray-scale image and describes the local shape using edge-gradient histograms. Y. Ke et al. [12] proposed PCA-SIFT, similar to SIFT, which represents the support region by a vector of gradient patch and then reduces the dimension by principle component analysis (PCA). It was showed that PCA-SIFT was both more distinctive and compact than the original SIFT descriptor. Independent component analysis (ICA), as an alternative statistical algorithm, was used to represent the interest points in [19].

Koen E.A. van de Sande et al. extended the standard SIFT descriptor to color SIFT descriptors by adding color information, since color plays an important role in distinguishing different types of objects [9]. They presented a comparative study of several global and local descriptors including the histograms of RGB, opponent, hue, rg, transformed color, color moments, moment invariants, standard SIFT and the color SIFT descriptors. Their experiments demonstrated that the accuracy of ranked category classification results using color SIFT descriptors was much better than those of other descriptors. Vijay Chandrasekhar et al. [2] argued that transmission and storage of local features are of critical importance for mobile visual search applications, and presented a comprehensive survey of SIFT compression schemes. In [22], we have utilized ICA to seek an adaptive and efficient color subspace for color SIFT feature extraction and showed the classification accuracy can be significantly improved. Following the previous work, this study attempts to select both more distinctive and compact color SIFT independent components for image classification.

Compared to PCA, ICA can produce statistically independent non-Gaussian components by de-correlating the higher order moments in addition to the first and second order moments of the statistical distribution. Meanwhile, it is able to describe local variations [14]. In PCA-based feature representation, the criterion of feature selection is to keep the variation of the original data as much as possible. However, ICA-based feature representation is devoid of such kind of criterions. Conventionally, independent components (IC) selection is based on PCA [19], [20]. In our previous study, PCA is first applied to color SIFT descriptors, and then followed by ICA [23]. Thus, in fact the selection results are decided by PCA, i.e. the global variation. However, for some kinds of images, such as object images, the local variation is more important for classification. In this paper, we propose two ordering methods for sorting the color SIFT independent components, which are based on the local variation. The first, called the localization-based IC ordering, orders the IC according to the localized variation [14]; the other, called sparseness-based IC ordering, employs the criteria of sparseness [5] for the IC ordering. Differing from the conventional method, the proposed methods calculate all the IC first and then sort them according to the proposed criteria for the IC selection. We evaluate the proposed methods and the conventional method [23] by conducting experiments with object image classification and

Copyright © 2011 The Institute of Electronics, Information and Communication Engineers
scene image classification. Meanwhile, we try to find the inclined power of each method and to demonstrate their different applicability.

This paper is organized as follows. The definition of independent component analysis is reviewed in Sect. 2. Section 3 presents the motivation of this study by discussing the ordering of the independent components. In Sect. 4, the experimental design is presented and the results are discussed. Finally, conclusions are drawn in Sect. 5.

2. Independent Component Analysis

Assume \( \mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_N] \) denotes \( N \) training sample sets that have a zero mean. Here \( \mathbf{x}_i, i = 1, 2, \cdots, N \), is the \( i \) th sample feature with \( L \) dimensions. It is considered to be a combination of unknown independent source components \( \mathbf{s}_i = [s_{i1}, s_{i2}, \cdots, s_{iL}]^T \) with an unknown mixing matrix \( \mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \cdots, \mathbf{a}_L] \)

\[
\mathbf{x}_i = \sum_{j=1}^{L} \mathbf{a}_j s_{ij} = \mathbf{A} s_i
\]

The objective of ICA is to seek an unknown demixing matrix \( \mathbf{W} \). The resulting vector,

\[
\hat{\mathbf{s}}_i = \mathbf{Wx}_i
\]

recovers the source vector \( \mathbf{s}_i \), except for possible permutation and rescaling.

In order to simplify estimation of the demixing matrix \( \mathbf{W} \), the training sample \( \mathbf{x}_i \) is pre-whitened by a whitening matrix \( \mathbf{V} \), namely:

\[
\mathbf{z}_i = \mathbf{Vx}_i
\]

where \( \mathbf{z}_i \) is the whitened training sample. The whitening matrix \( \mathbf{V} \) can be obtained by classical PCA:

\[
\mathbf{V} = \Sigma^{-1} \mathbf{U}^T
\]

given that \( \mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_L] \) is the orthogonal matrix of eigenvectors and \( \Sigma^2 = \text{diag} \{\lambda_1, \lambda_2, \cdots, \lambda_L\} \), \( \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_L \) is the diagonal matrix of eigenvalues. Therefore, instead of estimating \( \mathbf{W} \), the whitened \( \mathbf{z}_i \) is used to calculate a new matrix \( \mathbf{Q} \), so that

\[
\hat{\mathbf{s}}_i = \mathbf{Qz}_i
\]

is a good estimation of the source vector \( \mathbf{s} \) with possible permutation and rescaling.

Since the matrix \( \mathbf{Q} \) can not be calculated analytically, several ICA algorithms have been proposed [16], [26], [27]. The Fast-ICA method [16], which does not have a learning rate or other adjustable parameters, is used in our method. The row vector of \( \mathbf{Q} \) is denoted by \( \mathbf{q}_j^T, j = 1, 2, \cdots, L \). Using the kurtosis of the projection of \( \mathbf{z} \) into \( \mathbf{q}_j^T \), the cost function of FastICA is expressed as follows:

\[
\text{kurt}(\mathbf{q}_j^T \mathbf{z}) = E \left\{ (\mathbf{q}_j^T \mathbf{z})^4 \right\} - 3 \| \mathbf{q}_j \|_4^4
\]

The objective of ICA is to seek an unknown demixing matrix \( \mathbf{W} \). The resulting vector,

\[
\hat{\mathbf{s}}_i = \mathbf{Wx}_i
\]

subject to \( \| \mathbf{q}_j \|_2^2 = 1 \)

A Lagrangian multiplier \( \alpha \) is introduced to solve the above optimization problem:

\[
J(\mathbf{q}_j) = E \left\{ (\mathbf{q}_j^T \mathbf{z})^4 \right\} - 3 \| \mathbf{q}_j \|_4^4 + \alpha \left( 1 - \| \mathbf{q}_j \|_2^2 \right)
\]

As we know the kurtosis is greater or less than zero for a non-Gaussian random variable, and zero for a Gaussian random variable. The densities of the mixed variables are closer to a Gaussian density than the densities of the independent components. We therefore seek either the maximal or minimal kurtosis value prior to searching for the independent components. After differentiating Eq. (8) with respect to \( \mathbf{q}_j \) and setting it to 0, we obtain:

\[
\mathbf{q}_j = \frac{2}{\alpha} \left\{ E \left\{ \mathbf{z} (\mathbf{q}_j^T \mathbf{z})^3 \right\} - 3 \mathbf{q}_j \right\}
\]

According to the constraint (Eq. 7), the effect of the coefficient \( 2/\alpha \) can be removed by projecting \( \mathbf{q}_j \) onto the unit sphere. Therefore, the fixed-point algorithm for ICA is obtained as:

\[
\mathbf{q}_j = \frac{E \left\{ \mathbf{z} (\mathbf{q}_j^T \mathbf{z})^3 \right\} - 3 \mathbf{q}_j}{\| \mathbf{q}_j \|_2}
\]

The fixed-point algorithm can be used to estimate several independent components. In order to prevent the different components from converging to the same maxima, we de-correlate the outputs \( \mathbf{q}_1^T \mathbf{z}, \mathbf{q}_2^T \mathbf{z}, \cdots, \mathbf{q}_L^T \mathbf{z} \) after every iteration. It means that we need to add an orthogonализing projection after Eq. (9): when \( k \) vectors \( \mathbf{q}_1, \mathbf{q}_2, \cdots, \mathbf{q}_k \) have been estimated, the ‘projections’ \( \{\mathbf{q}_{k+1, r}\} \) \( r = 1, \cdots, k \) of the previously estimated \( k \) vectors are subtracted from \( \mathbf{q}_{k+1} \) [16]:

\[
\mathbf{q}_{k+1} = \mathbf{q}_{k+1} - \sum_{r=1}^{k} (\mathbf{q}_{k+1, r}) \mathbf{q}_r
\]

After all of the \( \mathbf{q}_j \) have been calculated, the demixing matrix \( \mathbf{W} \) is obtained as:

\[
\mathbf{W} = \mathbf{QV}
\]

3. Ordering of the Independent Component

3.1 A Conventional Independent Components Selection Method (PCA-Based Global IC Selection Method)

A PCA-based global IC selection method is a conventional components selection method. Initially, PCA is used to select the components, retaining as much variation as possible. Then, ICA is performed to obtain the independent components from the low-dimensional PCA subspace. Eigenvectors \( \mathbf{u}_i \) (obtained via PCA and introduced in Sect. 2) associated with the first \( D \) largest eigenvalues are used to form the projection subspace:
The original sample vector \( \mathbf{x} \) is projected onto the PCA subspace and represented by a low-dimensional vector \( \mathbf{y} \):

\[
\mathbf{y} = \mathbf{U}_D^T \mathbf{x}
\]

According to the pre-whitening processing of ICA (Eq. (4) and Eq. (13)), the low-dimensional vector \( \mathbf{y} \) can be used to seek independent components:

\[
\tilde{\mathbf{s}} = \mathbf{Wx} = \mathbf{Q}\Sigma_D^{-1}\mathbf{U}_D^T \mathbf{x} = \mathbf{Q}\Sigma_D^{-1} \mathbf{y}
\]

where \( \Sigma_D = \text{diag}[\lambda_1, \lambda_2, \cdots, \lambda_D] \), \( (\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_D) \) are the diagonal matrices of the eigenvalues.

Since PCA requires a Gaussian distribution of the input data, and is able to describe global feature variations [14], this method is called a global IC selection method, and is shown in Fig. 4 (a).

### 3.2 Proposed Independent Components Ordering Methods (Local IC Ordering Methods)

Unlike the PCA-based global IC selection method, the proposed IC ordering methods first directly apply ICA to the color SIFT descriptors and obtain independent components (shown in Eq. (2)). A localization-based IC method and a sparseness-based IC method [7] are then used to order and select components in order of information significance. ICA does not have the restriction that the input data must be drawn from a Gaussian distribution, and is able to describe local variations. Therefore, these proposed methods can be called local IC ordering methods. In the next two subsections, we focus on these two ordering methods in detail.

#### 3.2.1 Localization-Based Independent Components Ordering

The localization-based IC method orders the independent components according to the locality and the amplitude of the variation [14]. Columns of \( \mathbf{A} \) in the ICA model are called basis functions, and the \( s_{ij} \) component of the \( \mathbf{s}_i \) vector becomes a coefficient of the \( i \)-th basis vector in the observed vector \( \mathbf{x} \) (see Eq. (1)). When \( \mathbf{s}_i \) is varied in a range, the components of \( \mathbf{x} \) show a variation with a certain amplitude. For noisy modes, these variations are relatively small and are not localized, while for sample considering modes, these variations are localized and have a large amplitude.

The localization-based IC method orders the independent components based on the following criterion [14]:

\[
C = V \cdot M \cdot \frac{O}{F}
\]

where \( V \) is the width of the histogram of the projections, that are obtained by projecting observed descriptors into basis functions. \( M \) is the maximum variation amplitude, which is determined by weighing the basis functions \( \mathbf{A} \) with \( s_{ij} \) changing in the range of corresponding \( V \). \( O \) is the total surface of the variations whose amplitude is larger than 50% of \( M \).

\( F \) is the ratio of the number of components owning greater than \( M/2 \) amplitude and the total number of components in the descriptors (see Fig. 1). Since the noisy modes show variations with small amplitudes that occur along the entire descriptor, the \( O/F \) will be small for these modes and the value of \( C \) will be lower. The localization-based IC ordering method leads to a selection of independent components according to the value of \( C \). Components with the highest \( C \) are placed first and those with a low \( C \) are placed last:

\[
C_{\hat{w}_1} > C_{\hat{w}_2} > \cdots > C_{\hat{w}_L}
\]

where, \( \hat{w}_j (j = 1, 2, \cdots, L) \) denotes the row vectors (the basis functions) of the demixing matrix \( \mathbf{W} \).

#### 3.2.2 Sparseness-Based Independent Components Ordering

It is proved that most independent components have a super-Gaussian distribution and the corresponding basis functions are very sparse. They are similar to localized and oriented receptive fields in image decomposition obtained by using ICA [3], [5]. The sparseness-based IC ordering method determines independent components based on their sparseness. The \( p \) norm [6] of a basis function is used as a measure of the degree of sparseness [7]:

\[
\|\mathbf{A}_i\| = \left( \sum_{j=0}^{L-1} |A_{ij}|^p \right)^{1/p}, \quad \text{with} \quad p < 1
\]

where \( A_{ij} \) is \( j \)-th component of \( i \)-th basis function \( \mathbf{A}_i \), and \( L \) is the dimension of the descriptors. The smaller \( \|\mathbf{A}_i\| \) is, the sparser the basis function is. We rearrange the basis functions according to their degree of sparseness. Components with the lowest \( \|\mathbf{A}_i\| \) are placed first and those with a high value of \( \|\mathbf{A}_i\| \) are placed last:

\[
\|\mathbf{A}_{\hat{w}_1}\| < \|\mathbf{A}_{\hat{w}_2}\| < \cdots < \|\mathbf{A}_{\hat{w}_L}\|
\]

where, \( \hat{w}_j (j = 1, 2, \cdots, L) \) denotes the row vector of demixing matrix \( \mathbf{W} \).
4. Experiments

4.1 Object/Scene Database

Our experiment uses two diverse databases to evaluate the performance of three ICA-based feature dimension reduction methods. These were derived analytically in the previous section:

1. Figure 2 displays the Object database used by R. Fergus and others [8]. It is composed of the nine object categories and is summarized in Table 1. Because the image sizes vary from 16,000 to 530,432 pixels, we have resized all of the images to approximately 120,000 pixels retaining aspect ratio. The experiment is performed with 150 images per individual category for training, and the rest of the database for testing.

2. We have compiled a database of scenes to use for training test cases. Figure 3 displays the eight scene categories: parties, beaches, cooking, night, firework, sunset, snow and a generated closeup of a flower. Each category has 400 images. The resized images are 500 × 375 or 375 × 500 pixels in size. 150 images of each category are used as training data. The remainder are used as test data.

| Category | Number |
|----------|--------|
| leopards | 200    |
| airplanes| 1074   |
| bottles  | 247    |
| camel    | 356    |
| cars_brad| 1155   |
| faces    | 450    |
| guitars  | 1030   |
| house    | 1000   |
| motorbikes| 826   |

4.2 Experimental Setup

The RGB color space is a basic color space for color image representation. It can be transformed into other color spaces as required [9], [13], [22]. The proposed methods are explored by applying them to RGB-SIFT descriptors in the RGB color space and Opponent-SIFT descriptors in the Opponent color space, as well in case where the classification performance would be better if no prior knowledge about object and scene categories is available [9]. SIFT descriptors with 128-dimension (in each channel) are computed over the 32 × 32 regular grids that were used in our previous study and worked well for the same databases applied to object and scene classification [22]. The grids overlap by 50% in both horizontal and vertical positions. We combine the SIFT descriptors of the corresponding position in each channel into a 384 dimension feature vector (128 × 3). Based on localized variation, we used the proposed methods to select 128 independent components from Color-SIFT IC (with the same dimensions as the dimension of gray-scale SIFT descriptors). Naturally, any reasonable number of IC can be selected. Since the main purpose of this study is to discuss the effectiveness of local IC ordering methods for independent components selection, we only use the ICA-based color SIFT descriptors for image classification experiments. They may be used together with other features to achieve higher classification rate.

Subsequently, the orderless bag-of-features [28] are used as a features representation. We build a visual vocabulary by clustering the feature vectors (descriptors) from the training set, and then represent each image in the data set as a histogram of visual words drawn from the vocabulary. In our experiments k-means [17] is used as a clustering method. The number of clusters (visual words) is 500 for both the object database and the scene database. We employed SVMs as a classifier. This has been widely used and has been shown to be efficient for object/scene classification [4]. In our experiments, the LIBSVM package has been employed [18]. The classification process is summarized in Fig. 4.

4.3 Evaluation Criterion for Image Classification

The performance of the classification is measured using the average classification rate $A_{CR}$ and an $F_1$-measure. Assuming that there are $R$ categories in the database for each category $i$ ($i = 1, 2, \cdots, R$), the $F_1$-measure is defined as follows:

$$F_1 = \frac{2 \times \text{recall}_i \times \text{precision}_i}{\text{recall}_i + \text{precision}_i}$$

where

$$\text{recall}_i = \frac{tp_i}{tp_i + fn_i}$$

$$\text{precision}_i = \frac{tp_i}{tp_i + fp_i}$$
Fig. 4  Selection of color SIFT independent components for object/scene classification. (a) Conventional global IC selection method (b) Proposed local IC ordering methods.

where $t_{pi}$ is the number of images correctly classified; $t_{pi} + f_{ni}$ is the total number of relevant images; $t_{pi} + f_{pi}$ is the total number of retrievals. $F_1$-measure is the harmonic mean of recall and precision.

The average classification rate $ACR$ is shown as follows:

$$ACR = \frac{\sum_{i=1}^{R} recall_i}{R}$$

(24)

4.4 Experimental Results

In this subsection, local IC ordering methods and global IC selection method are evaluated using comparative experiments. These are done in terms of the behaviour of the average classification rate $ACR$ and $F_1$-measure in object and scene databases.

4.4.1 Comparison among Variations Using the Global IC Selection and Local IC Ordering Methods

As described in Sect. 3, “Localization-based IC Ordering” is used to order the IC according to their locality and the amplitude of the variation; “Sparseness-based IC Ordering” rearranges the IC according to their sparseness; and “PCA-based Selection” selects components preserving as much of the variance as possible. For nice visualization, we have drawn the variation in the RGB color space for the above three methods in Fig. 7 (The left, middle, and right columns express: the variation of IC obtained with “Localization-based IC Ordering”, the variation of IC obtained with “Sparseness-based IC Ordering” and the variation of eigenvectors obtained with “Global IC Selection” when the corresponding weight factors are varied between $-\sqrt{\frac{V}{2}}$ and $\sqrt{\frac{V}{2}}$, and $-3\sigma$ and $3\sigma$). It is clear that the vectors calculated from ICA show localized variations, whereas those yielded by PCA show global variations. Meanwhile,

by using local IC ordering methods, IC, which are kept at the top position, are of either localization and have large amplitude or sparseness.

4.4.2 Selection for Local Independent Components

By using local IC ordering methods, we re-arrange the IC according to the values of localization or sparseness (Eqs. (18) and (20)), and select the first 128 IC (more localization and sparseness) and the last 128 IC (less localization and sparseness) according to Eq. (2)). In Fig. 5, the first 128 IC are denoted as “localization128max” and the last 128 IC are labeled as “localization128min”. The aver-
Fig. 7  Left column: the variation of IC obtained with “Localization-based IC Ordering”; middle column: the variation of IC obtained with “Sparseness-based IC Ordering” and right column: the variation of eigenvectors obtained with “PCA-based Selection” when the corresponding weight factors are varied between $-V/2$ and $V/2$, and $-3\sigma$ and $3\sigma$. Rows (1), (2), ..., (6) denote the variations of the $1^{st}$, $10^{th}$, $80^{th}$, $128^{th}$, $374^{th}$ and $384^{th}$ components, respectively.

Fig. 8  Average classification rates of the scene database using eleven SIFT descriptors.

Image classification results clearly demonstrate that the components with large localization (the first 128 IC) are more efficient than those with less localization (the last 128 IC). In addition, the difference of the classification rates between the first 128 IC and the last 128 IC in the object database is larger than that in the scene database. The same results are obtained by using sparseness-based IC ordering method, which is the other local ordering method. In Fig. 6, the first 128 IC with more sparseness are denoted as “sparseness128min” and the last 128 IC with less sparseness are labeled as “sparseness128max”. This means that the classification results in the object database are more sensitive to local information than that in the scene database. This agrees with the factor that for object database the local changes in images are important clues, while for a scene database, focusing global changes in whole images are required. It is possible that local IC ordering methods may be better than the global IC selection method for the object database, and vice versa for the scene database. This hypothesis is supported by the experimental evidence discussed in Sect. 4.

4.4.3 Image Classification Results for Scene/Object Database

In this subsection, we investigate the efficiency of the proposed local IC ordering methods for Scene/Object images in RGB color space and opponent color space. In Fig. 8 and Fig. 9, 128 components are selected for each descrip-
Here, “PCAcolorSIFT128” means components selection carried out only by PCA; “ICAcolorSIFT128” denotes that components are first selected by PCA and then transformed into independent components by ICA. “Localization128max” and “Sparseness128min” mean that the independent components are first found by ICA, and then the first 128 components are retained according to the ordering methods (mentioned in Sect. 3.2). “Gray128” and “ColorSIFT384” denotes the original SIFT descriptors extracted from gray-scale image and RGB or opponent color space, respectively. Thus, a total of 11 kinds of descriptors are compared: five in the RGB color space, five in the opponent color space, and one in the gray-scale space.

Figure 8 shows that the highest average classification rate is obtained by using a global IC selection method (ICAcolorSIFT128 descriptors), which not only reduces the high dimension of the original color SIFT descriptor, but also improves the scene classification results. In Fig. 9 “Localization128max” and “Sparseness128min” obtained by the local IC ordering methods are superior to all other methods in the object database case.

We now taking a closer look at the classification results obtained using each of the two databases. For sake of perspicuity, only the results obtained in RGB color space are illustrated by comparing the F1-measure for six types of SIFT descriptors. As shown in Fig. 10, F1-measure displays the highest values by using the global ordering method (“ICAcolorSIFT128”) in five categories: beach, closeup_flower, cooking, night and sunset. Meanwhile, it gives a higher value when using the global ordering method than when using local ordering methods (“Localization128max” and “Sparseness128min”) in the night category. In the object database, shown in Fig. 11, the local ordering methods are superior to the global ordering method for seven categories. These results indicate that the global IC selection method is more effective for scene classification while the local IC
ordering methods are more efficient for object classification.

4.4.4 Comparison among Computational Time

The classification experiments are divided into three phases: Color SIFT extraction (including dimension reduction processing for descriptors), Bag-of-Feature (calculate the histogram of each color image) and Classification. The computational time for each phase are shown in Fig. 12. As shown in Fig. 12, we just calculate SIFT descriptors in the RGB color space over regular grids and do not need to detect keypoints, so feature extraction is not computationally expensive. In the phase of “Color SIFT extraction”, though the proposed method should require a multiplication of a demixing matrix with a size of 128 \times 384 for dimension reduction, the computational cost for this operation is trivial (only 24 ms) compared to the time of ColorSIFT384 (242 ms). Contrarily, descriptors assigned to a visual word (Bag of features) spend more computation time for ColorSIFT384. Since selecting limited number of features is able to efficiently reduce the computational time for computing distances to each visual word, the whole image classification time is efficiently reduced.

5. Conclusion and Future Works

In the paper, we proposed a strategy of using local variation to order the color SIFT independent components for image classification. Two local IC ordering methods, localization-based IC ordering and sparseness-based IC ordering were presented. We investigated 11 kinds of descriptors in three color spaces for object image classification and scene image classification. In comparison with standard color SIFT descriptors, the proposed local IC ordering methods are more distinctive and compact inspiring improvements in both object and scene classification. Additionally, local IC ordering methods are efficient feature selection for object images case, while global variation based IC selection method works well for scene images case.

Acknowledgements

This work was supported in part by the Research Matching Fund for Private Universities from MEXT (Ministry of Education, Culture, Sports, Science, and Technology) and in part by the Research Fund from Ritsumeikan University.

We acknowledge the kind assistance and comments of Prof. Lakhmi C. Jain and Dr. Jeff Tweedale at the University of South Australia, which led to important improvements in this study.

References

[1] D. Lu and Q. Weng, “A survey of image classification methods and techniques for improving classification performance,” Int. J. Remote Sensing, vol.28, pp.823–870, 2007.
[2] V. Chandrasekhar, M. Makar, G. Takacs, D.M. Chen, S.S. Tsai, R. Grzeszczuk, and B. Girod, “Survey of SIFT compression schemes,” Proc. International Conference on Pattern Recognition (ICPR), Istanbul, Turkey, Aug. 2010.
[3] A.J. Bell and T.J. Sejnowski, “The “independent components” of natural scenes are edge filters,” Vision Research, vol.37, no.23, pp.3327–3338, 1997.
[4] B. Scholkopf and A. Smola, Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond, MIT Press, Cambridge, MA, 2002.
[5] X.H. Han, Y.W. Chen, and Z. Nakao, “Robust edge detection by independent component analysis in noisy image,” IEICE Trans. Inf. & Syst., vol.E87-D, no.9, pp.2204–2211, Sept. 2004.
[6] J.O. Smith III, Mathematics of the Discrete Fourier Transform (DFT): With music and audio applications, pp.1–264, W3K Publishing, 2003.
[7] J. Karvanen and A. Cichocki, “Measuring sparseness in noisy signals,” Proc. ICA 2003, pp.125–130, 2003.
[8] R. Fergus, P. Perona, and A. Zisserman, “A sparse object category model for efficient learning and exhaustive recognition,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2005.
[9] K.E.A. van de Sande, T. Gevers, and C.G.M. Snoek, “Evaluating color descriptors for object and scene recognition,” IEEE Trans. Pattern Anal. Mach. Intell., vol.32, no.9, pp.1582–1596, 2010.
[10] D.G. Lowe, “Distinctive image features from scale-invariant keypoints,” Int. J. Comput. Vis., vol.60, no.2, pp.91–110, 2004.
[11] R. Datta, D. Joshi, J. Li, and J.Z. Wang, “Image retrieval: Ideas, influences, and trends of the new age,” ACM Comput. Surv., vol.40, no.2, pp.1–60, 2008.
[12] Y. Ke and R. Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors,” CVPR 04 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004.
[13] M. Tkalci and J.F. Tasic, “Colour spaces: Perceptual, historical and applicational background,” EUROCON 2003. Computer as a Tool, The IEEE Region 8, vol.1. 22-24, pp.304–308, 2003.
[14] M. Uzumcu, A.F. Frangi, J. Reiber, and B. Lelieveldt, “Independent component analysis in statistical shape models, medical imaging 2003,” Image Processing, vol.5032, pp.375–383, 2003.
[15] P. Comon, “Independent component analysis—a new concept?,” Signal Process., vol.36, pp.287–314, 1994.
[16] A. Hyvärinen, “Fast and robust fixed-point algorithms for independent component analysis,” IEEE Trans. Neural Netw., vol.10, no.3, pp.626–634, 1999.
[17] J.B. MacQueen, “Some methods for classification and analysis of multivariate observations,” Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability, pp.281–297, University of California Press, 1967.
[18] C.C. Chang and C.J. Lin, “LIBSVM: A library for support vector machines,” Software available at: http://www.csie.ntu.edu.tw, 2001.
[19] T. Han, W. Li, T. Wang, and Y. Wang, “Independent components analysis for representation interest point descriptors,” Lect. Notes Comput. Sci., vol.4113, pp.1219–1223, 2006.
[20] M. Yamazaki and S. Fels, “Local image descriptors using supervised kernel ICA,” Pacific-Rim Symposium on Image and Video Technology (PSIVT), pp.94–105, 2009.

[21] K. Chakrabarti and G. Minos, “Approximate query processing using wavelets,” The VLDB Journal, pp.199–223, 2001.

[22] D. Ai, X. Han, X. Ruan, and Y. Chen, “Color independent components based SIFT descriptors for object/scene classification,” IEICE Trans. Inf. & Syst., vol.E93-D, no.9, pp.2577–2586, Sept. 2010.

[23] D. Ai, X. Han, G. Duan, X. Ruan, and Y. Chen, “ICA vs PCA color SIFT descriptors for scene classification,” Meeting on Image Recognition Understanding, IS2-74, 2010.

[24] V. Ferrari, T. Tuytelaars, and L. Van Gool, “Simultaneous object recognition and segmentation by image exploration,” ECCV, pp.40–54, 2004.

[25] A. Opelt, M. Fussenegger, A. Pinz, and P. Auer, “Weak hypotheses and boosting for generic object detection and recognition,” ECCV, pp.71–84, 2004.

[26] A.J. Bell and T.J. Sejnowski, “An information maximization approach to blind separation and blind deconvolution,” Neural Comput., vol.7, pp.1129–1159, 1995.

[27] J.F. Cardoso, “Higher order contrasts for independent component analysis,” Neural Comput., vol.11, no.1, pp.157–192, 1999.

[28] J.G. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid, “Local features and kernels for classification of texture and object categories: A comprehensive study,” Beyond Patches Workshop, Computer Vision and Pattern Recognition (CVPR), 2006.

Dan-ni Ai received a B.E. degree in 2005 from Xi’an Jiaotong University, China, an M.E. degree in 2008 from Xi’an Jiaotong University, China. She is currently a PhD candidate at the Graduate School of Engineering and Science, Ritsumeikan University. Her current research interests include image processing and analysis, feature extraction and image annotation.

Guifang Duan received the Doctor degree in integrated science and engineering from Ritsumeikan University, Japan, in 2009. He is currently a Postdoc fellow at the Institute of Science and Technology, Ritsumeikan University. His research interests include computer vision, machine learning, and visual inspection. He is the member of IEEE.

Xian-hua Han received a B.E. degree in 1999 from ChongQing University, China, an M.E. degree in 2002 from ShanDong University, JiNan, China, and a D.E. degree in 2005 from University of the Ryukyus, Okinawa, Japan. From 2006 to 2007, she was a lecture with Electronics and Information Engineering School Central South University of Forestry and Technology, China. She is currently a post-doctoral fellow at college of Information Science and Engineering, Ritsumeikan University. Her current research interests include image processing and analysis, feature extraction, image annotation and pattern recognition.

Xiang Ruan received the BS degree in Shanghai Jiaotong University in 1997 and the MS and the PhD degrees in Osaka City University in 2001 and 2004. He is currently a researcher of electronic components company of OMRON Corp. His current research interests include computer vision, pattern recognition and image processing. He is a member of the IEEE and the IEEE Computer Society.

Yen-wei Chen received a B.E. degree in 1985 from Kobe Univ., Kobe, Japan, a M.E. degree in 1987, and a D.E. degree in 1990, both from Osaka Univ., Osaka, Japan. From 1991 to 1994, he was a research fellow with the Institute of Laser Technology, Osaka. From Oct. 1994 to Mar. 2004, he was an associate Professor and a professor with the Department of Electrical and Electronic Engineering, Univ. of the Ryukyus, Okinawa, Japan. He is currently a professor with the college of Information Science and Engineering, Ritsumeikan University, Kyoto, Japan. He was a visiting professor with the Oxford University, Oxford, UK in 2003. He is an Overseas Assessor of Chinese Academy of Science and Technology. He is an associate Editor of International Journal of Image and Graphics (IJIG) and editorial board members of the International Journal of Knowledge-based and Intelligent Engineering Systems and the International Journal of Information. His research interests include pattern recognition, image processing and machine learning. He has published more than 200 research papers in these fields.