Adaptive Color Independent Components based SIFT Descriptors for Image Classification

Danni Ai[1], Xianhua Han[1], Xiang Ruan[2] and Yen-Wei Chen[1]
[1] *Graduate School of Science and Enginnering, Ritsumeikan University, Japan*
[2] *Omron corporation, Japan*

Abstract

This paper proposes an adaptive color independent components based SIFT descriptor (termed CIC-SIFT) for image classification. Our motivation is to seek an adaptive and efficient color space for color SIFT feature extraction. Our work has two key contributions. First, based on independent component analysis (ICA), an adaptive and efficient color space is proposed for color image representation. Second, in this ICA-based color space, a discriminative CIC-SIFT descriptor is calculated for image classification. The experiment results indicate that (1) contrast between objects and background can be enhanced on the ICA-based color space and (2) the CIC-SIFT descriptor outperforms other conventional color SIFT descriptors on image classification.

1. Introduction

Feature extraction, as one of the most important steps for image classification, can capture a certain visual property of an image. Recently, local SIFT descriptors (Scale-Invariant Feature Transform) have been widely used for image classification [1]. An inherent problem with SIFT is that its description for the local shape is by only gray-scale intensities [2]. Since color provides valuable information, several color SIFT descriptors have been proposed by combining color information with local SIFT descriptors, such as RGB-SIFT, Transformed color SIFT[2], HSV-SIFT[3], CSIFT[4] and Opponent-SIFT[2]. What they are in common is that: they firstly transform an original color image into other color space to obtain a new color representation, and then, extract features in the transformed color space. However, the physical model based transformation, which is used to obtain the new color representation, is defined uniquely and is not adaptive to input images. Moreover, it is a fact that different image categories have specific color distributions. Thus, for efficient color representation, it is necessary to find an adaptive transformation by analyzing each image category rather than using a unique physical model.

Interest is growing in developing image coding by statistical analysis. In our previous work [5], we have shown that among the three color independent components obtained from RGB color space, two are in an opposing-color model, by which the responses of R, G, B cones are combined in opposing fashions. Moreover, Buchsbaum et al.[6] found that the use of opposing fashions is the most efficient way to encode human photoreceptors. This coincides with the idea of finding an efficient color representation.

In this paper, we employ Independent Component Analysis (ICA)to find an adaptive and efficient color space (color independent components) for each image category and extract SIFT features in the ICA-based color space, which is termed as CIC-SIFT (Color Independent Component based SIFT). We show that the CIC-SIFT descriptor is more discriminative for image classification than other exiting color SIFT descriptors.

2. Color independent components based SIFT descriptors (CIC-SIFT)

2.1. Independent Component Analysis

Independent Component Analysis (ICA) is a linear transformation that transforms a set of random data to be statistically independent of each other [7].

Considering $\mathbf{x}$ is a combination of independent source components $\mathbf{s}$ with the mixing matrix $\mathbf{A}$:

$$\mathbf{x} = \mathbf{A}\mathbf{s} \tag{1}$$

The goal of ICA is to find the transformation matrix $\mathbf{W}$, so that each element of the resulting vector $\mathbf{y}$ becomes as independent as possible:

1051-4651/10 $26.00 © 2010 IEEE
DOI 10.1109/ICPR.2010.596

$$\mathbf{y} = \mathbf{W}\mathbf{x} \quad (2)$$

Bell and Sejnowski [8] have proposed a neural learning algorithm for ICA. The approach is to maximize joint entropy by stochastic gradient ascent. The updating formula for **W** is:

$$\mathbf{W}^{(t+1)} = \mathbf{W}^{(t)} + \mu\left[\mathbf{I} - \varphi\left(\mathbf{y}^{(t)}\right)\left(\mathbf{y}^{(t)}\right)^T\right]\mathbf{W} \quad (3)$$

where μ is a learning coefficient, t is a iteration number, and $\phi(\mathbf{y}) = 1 - 2/\left(1+e^{-\mathbf{y}}\right)$ is calculated for each component of **y**. A whitening technique was used to accelerate this computation. If the iteration converges, **y** is considered to be equivalent to **s** except in scale and permutation.

In this paper, we consider that each R, G, B component is a linear combination of three independent components, which can be found by ICA, and propose a adaptive color independent component based SIFT descriptor (termed CIC-SIFT) for efficient image classification.

2.2. Color independent components based SIFT descriptors (CIC-SIFT)

The proposed CIC-SIFT can be summarized in three steps: (1) Learning step: a transformation matrix $\mathbf{W}_{ICA}$ is learned for each category by using ICA; (2) Color transformation: original images components are transformed into three independent components by using the adaptive transformation matrix $\mathbf{W}_{ICA}$; (3) Feature extraction: the color independent components are utilized to calculate CIC-SIFT descriptors.

Assuming that there are n categories in a database and the information of each image are carried by RGB components. In order to find an efficient ICA-based color space for color representation, we learn an ICA transformation matrix from each category. The color intensity of each pixel $\mathbf{x}_{RGB} = \left[x_R, x_G, x_B\right]^T$ is used as a sample.

As we already know, the goal of ICA is to seek a transformation matrix $\mathbf{W}_{ICA}$ to linearly transform input data so that outputs are as statistically independent of each other as possible. For each category, n transformation matrix $\mathbf{W}_{ICA_i}$ $\left(i \in [1,2,\cdots,n]\right)$ (3×3) can be calculated by using the learning algorithm proposed by Bell and Sejnowski [4], which can transform the original components ($x_R \;\; x_G \;\; x_B$) of images into three independent components $\left(x_{IC1} \;\; x_{IC2} \;\; x_{IC3}\right)$. The independent components are used as the new color representation (ICA-based color space) to calculate CIC-SIFT descriptors.

The standard SIFT descriptor has a 128-dimension feature for a keypoint in a gray-scale image and can be described as $\mathbf{f}_{Gray} = [f_1, f_2, \cdots, f_{128}]^T$. Thus the CIC-SIFT descriptor for a keypoint of three independent components is represented as

$$\mathbf{f}_{CIC-SIFT} = [f_1^{(IC1)} \cdots f_{128}^{(IC1)}, f_1^{(IC2)} \cdots f_{128}^{(IC2)}, f_1^{(IC3)} \cdots f_{128}^{(IC3)}]^T$$

which has 384 dimension.

3. Image classification Experiment

In order to demonstrate the effectiveness, the CIC-SIFT is applied to image classification experiments.

3.1. Image feature

Since the main purpose of this paper is to discuss the efficient color representation and discriminative color SIFT descriptors, we only used color SIFT features for image classification experiments though color SIFT features can be used together with other features for achieving higher classification rate.

In each transformed color channel, overlapped regular patches, with the size of 32×32 pixels, are extracted to compute their local image gradient. Seven color SIFT features are compared in our experiments: standard SIFT, RGB-SIFT, Transformed color SIFT, HSV-SIFT, CSIFT, Opponent-SIFT and CIC-SIFT descriptors.

3.2. Image classification

In our method, bag-of-feature is used as feature representation [9]. We employed SVMs as a classifier, which has been widely used and shown to be efficient for image classification [10].

What we should mention is that the test images need to be transformed by all learned transformation matrices since we do not really know which category they belong to in advance. Therefore, for each test image (T), we can obtain n^2 classification probabilities (n categories × n transformation matrices). The test image (T) is allocated to the category (C) that has the highest value of probabilities, as shown in Eq. 4:

$$C(\mathrm{T}) = \arg\max_i \left(P_{\mathbf{W}_1}(\mathrm{T}, i), \cdots, P_{\mathbf{W}_n}(\mathrm{T}, i)\right) \quad (4)$$

where $i = 1,2,...,n$ is the category label; $\mathbf{W}_j$ ($j = 1,2,...,n$) is transformation matrices learned from each category; $P(x)$ denotes classification probabilities of a test image allocated into different categories.

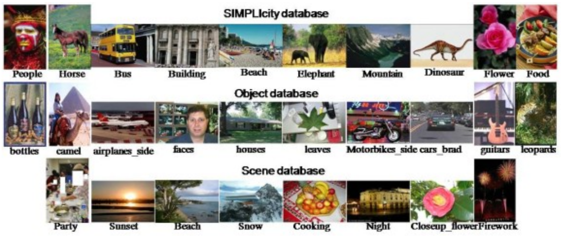

Figure 1: Image databases.

3.3. Image database

We evaluate our CIC-SIFT descriptors and other six ones on three diverse databases: SIMPLICITY image database [11], object database composed of ten categories collected from Caltech and ref. [12] (all the datasets is available at ref. [13] and [14]) and our own scene database collected from internet. Typical image of each category are shown in Fig. 1.

4. Experimental results

4.1. Color space transformation

In order to discuss the efficiency of ICA-based color representation, we transformed original images into seven different color spaces (Gray-scale space, RGB color space, Transformed color space, HSV color space, Color invariant space [4], Opponent color space and ICA-based color space).

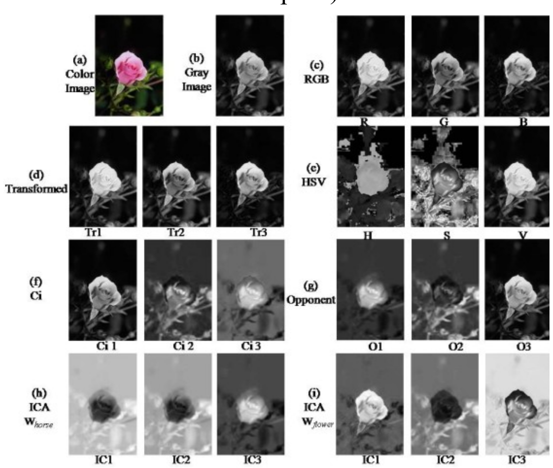

Figure 2: Transformed Color Space

A sample image is shown in Fig. 2. As we can see from Fig.2(c), both the object (flower) and the background (leaves) appear in original R, G and B channels clearly. They seem to be similar compared to those obtained by transformed color transformation (Fig.2(d)) which is normalized from RGB color space. In HSV color space, the object is difficult to distinguish in the hue-channel and not boosted significantly in the saturation and value channel (Fig.2(e)). We can not distinguish the object from the background in Color invariant space clearly (Fig.2(f)). In Opponent color space, summation component O3 that corresponds to illumination can not enhance the contrast between the object and the background, while the contrast can be enhanced in opponent components O1 and O2 (Fig.2(g)). However, since the transformation matrix, which is defined uniquely, is not a suitable one for the flower category, the boundary between the object and the background is not very clear.

Our method is to seek an adaptive transformation matrix for each category based on ICA and present efficient color representation for corresponding color image. If we look at flower category for example, the transformation matrix is calculated by ICA as:

$$\mathbf{W}_{ICA_{flower}} = \begin{bmatrix} -0.0446 & 0.8608 & -0.7229 \\ 0.2014 & -1.1898 & 0.9565 \\ -0.2962 & -0.4821 & 1.3432 \end{bmatrix} \quad (5)$$

Two rows of $\mathbf{W}_{ICA_{flower}}$ reflect an opposing-color model by combining R, G and B into opposing color components. The interesting point is that there is no summation component, which usually appears in many transforms.

With this transformation matrix $\mathbf{W}_{ICA_{flower}}$, we can transform the flower image into ICA-based color space (show in Fig. 2(i)), where the flower (object) is boosted while the leaves (background) are suppressed significantly. Note that more clear boundaries between the object and the background appear in Fig. 2(i) than those appear in Fig. 2(h), where the flower image is transformed by an uncorresponding transformation matrix ($\mathbf{W}_{horse}$), which is learned from the horse category.

4.2. Classification Results

4.2.1. SIMPLICITY image database. In SIMPLIcity

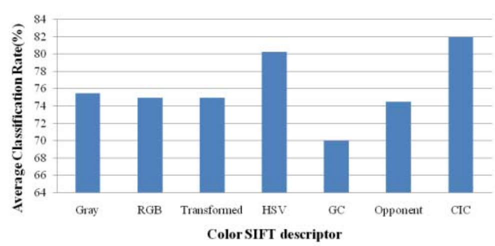

Figure 3: Average classification rate in the SIMIPLICITY image database.

database, 1000 images included in ten different categories (Fig. 1) are divided into two parts: 600 color images in the training database (60 for each category) and 400 color images in the test database (40 for each category). As shown in Fig. 3, the average classification rate of our proposed CIC-SIFT descriptor is estimated as 82%, which is the highest among all other SIFT descriptors. Thus, it denotes that CIC-SIFT descriptors can improve image classification.

4.2.2. Object database. Object database is composed of ten object categories summarized in Table 1.

Table 1. Image number of each category

Category	leopards	airplanes	bottles	camel	cars_brad
Number	200	1074	247	356	1155
Category	faces	guitars	house	leaves	motorbikes
Number	450	1030	1000	186	826

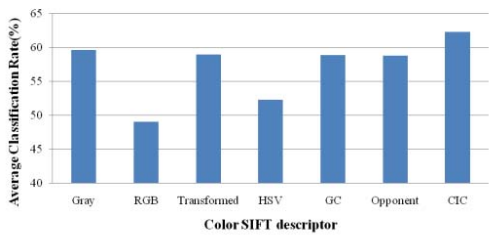

Figure 4: Average classification rate in the object database.

Figure 4 shows results of classification experiments using 40 images per category for training and the rest for testing. Our proposed CIC-SIFT descriptors based classification rate is 62.3%, which is the highest one among all other SIFT descriptors. Thus, we can see that CIC-SIFT descriptors are efficient and discriminative for the object classification.

4.2.3. Scene database. The scene database is composed of eight scene categories, including beach, party, cooking, sunset, firework, snow, flower and night (Fig. 1). Each category has 400 images. 40 images per category are used as training data, and the rest are test data. We can clearly see from Fig. 5 that the average classification rate is 69.1% using our CIC-SIFT descriptors, which is much higher than other best results of 62.1%, achieved using Opponent-SIFT.

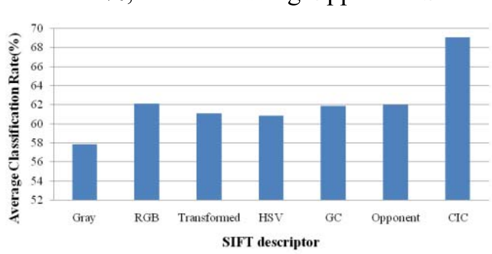

Figure 5: Average classification rate in the scene database.

Thus, we can conclude that our proposed CIC-SIFT descriptors are efficient for both object and scene image calssification.

5. Conclusion

In the paper, we proposed a novel color independent components based SIFT descriptor (CIC-SIFT) for image classification. The proposed ICA-based color space is adaptive and efficient to each category, which can be used for enhancing the contrast between the object and the background. The proposed CIC-SIFT descriptor can extract more discriminative local features for image classification. The experimental results show that our proposed CIC-SIFT descriptor is superior to other conventional color SIFT descriptors for both object and scene images.

6. References

[1] David G. Lowe. Distinctive Image Features from Scale-Invariant Keypoints. The International Journal of Computer Vision, 2004.

[2] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluation of Color Descriptors for Object and Scene Recognition. IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2008

[3] A. Bosch, A. Zisserman, and X. Muoz. Scene classification using a hybrid generative/discriminative approach. IEEE Trans. Pattern Analysis and Machine Intell, 30(04):712- 727, 2008

[4] A. E. Abdel-Hakim and A. A. Farag. CSIFT: A SIFT Descriptor with Color Invariant Characteristics. Computer Vision and Pattern Recognition (CVPR'06). 2, 1978-1983, 2006

[5] X. Y. Zeng, Y. W. Chen, and Z. Nakao. Classification of remotely sensed images using independent component analysis and spatial consistency. Journal of Advanced Computational Intelligence and Intelligent Informatics, 8(2):216-222, 2004

[6]G.Buchsbaum and A. Gottschalk. Trichromacy opponent colours coding and optimum colour information transmission in the retina. Proceedings of the Royal Society Lodon, B, 220:89-113, 1983

[7] P. Comon, Independent component analysis-a new concept? Signal Processing 36, 287-314, 1994.

[8] A.J.Bell and T.J.Sejnowski. An information maximization approach to blind separation and blind deconvolution. Neural Computation, 1129-1159, 1995.

[9] 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.

[10] B. Scholkopf and A. Smola. Learning with Kernels: Support Vector Machines, Regularization, Optimization and Beyond. MIT Press, Cambridge, MA, 2002.

[11] J. Z. Wang, J. Li, and G. Wiederhold. SIMPLIcity: Semantics sensitive Integrated Matching for Picture Libraries. IEEE Trans. On Pattern Analysis and Machine Intelligence, 23(9):947-963, 2001.

[12]R. Fergus, P. Perona, and A. Zisserman. A Sparse Object Category Model for Efficient Learning and Exhaustive Recognition. Proc. of the IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2005.

[13] http://www.robots.ox.ac.uk/~vgg/data.html

[14] http://www.vision.caltech.edu/html-files/archive.html