Fast Color Matching Using Weighted Subspace on Medicine Package Recognition

Kenjiro Sugimoto and Sei-ichiro Kamata
Graduate School of Information, Production and Systems, Waseda University
2-7 Hibikino Wakamatsu-ku Kitakyushu-shi Fukuoka, 808-0135 JAPAN
sugimoto@asagi.waseda.jp, kam@waseda.jp

Abstract

This paper presents a color matching technique using weighted subspace on medicine package recognition. The proposed method is more compact and lower-complex than scalable color descriptor and dominant color descriptor, which are employed by MPEG-7. Our method is based on subspace matching: A color object is treated as a subspace derived from its color distribution. Unlike mutual subspace method, it is specially designed for color matching. Specifically, weighted subspace and a distance-based dissimilarity are employed instead of normalized subspace and similarity based on canonical angles of MSM. Experiments show that the proposed method outperforms the conventional methods in terms of description size, building/matching speed, and recognition rate.

1 Introduction

Medicine package recognition is a significant technique for preventing dispensing errors, namely incorrect prescription of a medicine or its dosage. Fast and accurate recognition algorithms are required in order to actualize smooth and reliable dispensing operation. This research targets Press-Through Package (simply called a *package* in the paper), which is one of the most popular packages for pills/tablets/capsules, as shown in Figure 1. Color information is an important features of the packages. A package contains several pills/tablets/capsules and some characters/symbols are regularly printed on its surface. Every package is uniquely color-designed, normally showing in a few distinct colors. Thus, color matching is a reasonable solution for medicine package recognition. Although symbol recognition is also effective, color matching is nevertheless useful to fast prune candidates prior to symbol recognition. This paper focuses on low-level color descriptors.

Many descriptors including MPEG-7[1] have been proposed in the past. The conventional descriptors can be classified into two groups: description based on histogram[2, 3] and dominant color[4, 5]. MPEG-7 employs scalable color descriptor (SCD) as the former and dominant color descriptor (DCD) as the latter. Histogram description is easy to build and match, whereas the size is relatively large, e.g. RGB space divided into 8×8×8 requires 512 bins. Dominant color description is much more compact, whereas clustering process to find dominant color consumes much computer resources. In order to resolve the drawbacks, a color descriptor using eigenvectors and eigenvalues was proposed in [6]. The method treats a color distribution as eigenvectors weighted by their corresponding contribution ratio. However, three unconvincing points remain: (1) the reason why the weighted eigenvectors can efficiently characterize color distributions, (2) concrete difference from mutual subspace method (MSM)[7], a successful technique based on subspace matching on pattern recognition, (3) effect of dimension reduction. Therefore this paper specifies the three points and presents a modified color matching method based on the analysis.

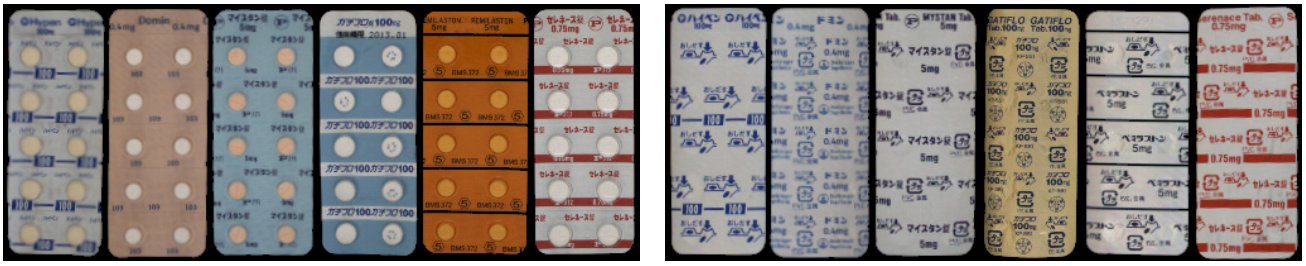
The rest of the paper is organized as follows: Section 2 analyzes subspace description of color distributions and outlines drawbacks of MSM for color matching. Section 3 proposes a modified subspace method specialized for color matching. Section 4 evaluates its performance by comparing with MSM, SCD and DCD. Finally, Section 5 provides our conclusions.

2 Color Matching Using Subspace Method

2.1 Analysis on subspace description

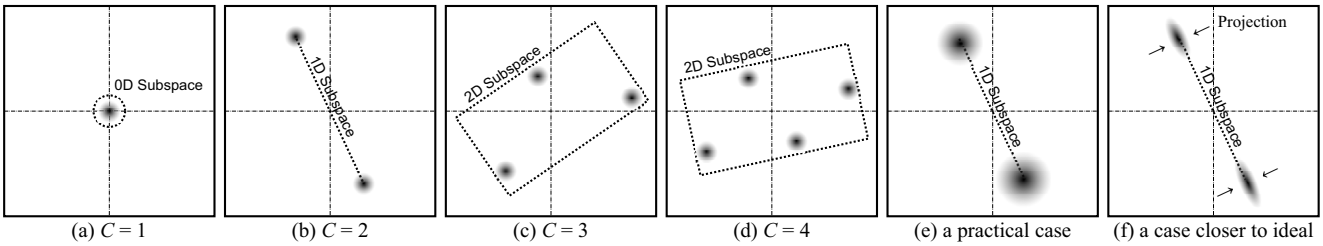
Subspace description requires the following condition: A color distribution in K-dimensional color space consists of C color clusters where C is unknown but satisfies $C \leq K + 1$, e.g. $K = 3$ for RGB space and $K = 4$ for CMYK space. Each cluster is assumed to be a point. Figure 2 illustrates some examples of distributions with $C = 1, 2, 3, 4$ under $K = 2$. If $C = 1$ (see Figure 2a), the cluster locates on a point. If $C = 2$ (see Figure 2b), the two clusters can produce a line. If $C = 3$ (see Figure 2c), the three clusters can produce a plane. The point, line, and plane are zero-, one-, and two-dimensional subspaces respectively. By contrast, if $C = 4$ (see Figure 2d), the four clusters still produce a plane because of limitation of $K = 2$. Thus, as long as $C \leq K + 1$, eigenvalues are effective to represent C and power relationship between clusters. Next, another helpful information is the direction of eigenvectors. Even if C is the same, eigenvectors can distinguish distributions with distinct clusters. Accordingly, a subspace in K-dimensional space can characterize up to $(K + 1)$ clusters.

Another positive effect of subspace description is noise reduction for more practical cases. Practically, each cluster is a dense manifold, not a point. Consider an ideal distribution (see Figure 2b) and its more practical one (see Figure 2e) composed of two clusters. The two clusters in the ideal case are exactly located on a one-dimensional subspace and its orthocomplement contains no energy. By contrast, the orthocomplement in the practical case will contain some energy. Projection to the subspace can moves the practical distribution closer to the ideal one (see Figure 2f). Therefore, subspace description behaves also as noise reduction if $C \leq K$.



(a) Obverse side (pill-observable side) (b) Reverse side (pill-unobservable side)

Figure 1: Medicine packages called PTP.



(a) $C = 1$ (b) $C = 2$ (c) $C = 3$ (d) $C = 4$ (e) a practical case (f) a case closer to ideal

Figure 2: Illustrations of distributions consisting of few clusters.

This subspace approach is more friendly than dominant color one with clustering process. Generally, clustering process requires to accurately estimate C. If two of three clusters are fairly close to each other, clustering process may incorrectly suppose $C = 2$. This is a sensitive problem to noise and fatal for matching phase. On the contrary, subspace building requires no estimation of C. Even if this situation incorrectly builds a one-dimensional subspace in spite of $C = 3$, the difference from the correct subspace is just a small second eigenvalue. Hence, the subspace approach can more easily describe correct cluster features than dominant color one.

2.2 Mutual subspace method

MSM[7], a successful technique for pattern matching on character recognition and face recognition, treats a pattern distribution as a subspace and recognizes anonymous patterns by subspace matching. Consider a distribution consisting of N vectors in K-dimensional space. Let $\boldsymbol{x}_n$ $(n = 1, 2, \ldots, N)$ denote one of the K-dimensional vectors composing the distribution. Its covariance matrix can be expressed as

$$\boldsymbol{S} = \frac{1}{N}\sum_{n=1}^{N}(\boldsymbol{x}_n - \bar{\boldsymbol{x}})(\boldsymbol{x}_n - \bar{\boldsymbol{x}})^{\mathsf{T}},$$

where $^{\mathsf{T}}$ indicates transposition and $\bar{\boldsymbol{x}}$ denotes the mean vector of $\boldsymbol{x}_n$. PCA to S derives eigenvalues λ_k $(k = 1, 2, \ldots, K)$ and their corresponding eigenvectors $\boldsymbol{\phi}_k$. Note that $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_K \geq 0$ and $(\boldsymbol{\phi}_i, \boldsymbol{\phi}_j) = \delta_{ij}$ $(i, j = 1, 2, \ldots, K)$ where $(\cdot, \cdot)$ indicates inner product and δ_{ij} denotes Kronecker delta. An r-dimensional subspace $(r \leq K)$ is spanned by the r eigenvectors $\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_r$.

Subspace matching requires a metric between subspaces. Consider two r-dimensional subspaces $\mathcal{P}$ and $\mathcal{Q}$. Let $\boldsymbol{\Phi} = [\boldsymbol{\phi}_1, \boldsymbol{\phi}_2, \ldots, \boldsymbol{\phi}_r]$ and $\boldsymbol{\Psi} = [\boldsymbol{\psi}_1, \boldsymbol{\psi}_2, \ldots, \boldsymbol{\psi}_r]$ denote their $K \times r$ orthonormal bases respectively. Also let λ_k and ν_k denote their eigenvalues respectively. MSM employs the following similarity

$$s(\mathcal{P}, \mathcal{Q}) = \sup_{\boldsymbol{u}\in\mathcal{P}, \boldsymbol{v}\in\mathcal{Q}} \frac{(\boldsymbol{u}, \boldsymbol{v})^2}{\|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2} = \cos^2\theta_1, \qquad (1)$$

where $\|\boldsymbol{u}\| \neq \boldsymbol{0}$ and $\|\boldsymbol{v}\| \neq \boldsymbol{0}$. θ_k $(k = 1, 2, \ldots, r)$ is called the k-th canonical angle between $\mathcal{P}$ and $\mathcal{Q}$. Eq. (1) is extended to the following summation in [8].

$$s(\mathcal{P}, \mathcal{Q}) = \sum_{k=1}^{r}\cos^2\theta_k. \qquad (2)$$

According to [9], $\cos^2\theta_k$ correspond to the k-th eigenvalues of $\boldsymbol{A} = \boldsymbol{\Psi}^{\mathsf{T}}\boldsymbol{\Phi}\boldsymbol{\Phi}^{\mathsf{T}}\boldsymbol{\Psi}$. Hence, MSM calculates Eq. (2) by applying PCA to $\boldsymbol{A}$ for each subspace matching.

2.3 Drawbacks of mutual subspace method

MSM is inefficient to color matching because it is just designed for character recognition or face recognition, not for color matching. This subsection specifies two drawbacks of MSM by comparing color matching with character recognition.

Firstly, MSM assumes $\lambda_k = \nu_k = 1$. This is equivalent to normalization of $(\boldsymbol{x}_n - \bar{\boldsymbol{x}})$. This operation is inadequate since eigenvalues are helpful information for color matching as stated above. In color matching, eigenvalues are related to the distance between clusters and seem interpretable as vividness. Vividness is an important feature for categories and should be utilized for matching. On the other hand, in character recognition, eigenvalues signify the contrast of images. A character category probably contains different contrast images, e.g. both of dark and light characters. MSM ignores eigenvalues in order to be invariant to contrast. However, this is unsuitable for color matching.

Secondly, the similarity with canonical angles is ineffective for color matching. Eq. (1) or Eq. (2) basically target sparse space. Normally, character recognition manages $K \geq 100$ whereas color matching manages $K \leq 16$ at most even for multispectral imaging.

Table 1: The test sets used in the experiments

Test Set	#Image	#Category	Note
OBV	900	84	Obverse side
REV	936	84	Reverse side
ALL	1836	168	OBV+REV

Low-dimensional space is more difficult to be sparse than high-dimensional one. Additionally, if $r = K$, $s(\mathcal{P}, \mathcal{Q}) = K$ anywise. Because $\boldsymbol{A} = \boldsymbol{I}$, that is projection to the original space. As a result, functionable options are only $r = 1, 2$ for RGB color space. Hence, Eq. (2) is inflexible for low-dimensional space such as color space.

3 Color Matching Using Weighted Subspace

The proposed method is designed based on the analysis in Section 2. In general, a subspace is described as eigenvectors weighted by their corresponding eigenvalues.

$$\boldsymbol{p}_k = \lambda_k^{\frac{1}{2}} \boldsymbol{\phi}_k, \qquad \boldsymbol{q}_k = \nu_k^{\frac{1}{2}} \boldsymbol{\psi}_k.$$

Note that the weight coefficients are equivalent to standard deviation on the direction of principal axis. Their orthogonal bases can be expresses as

$$\boldsymbol{P} = [\boldsymbol{p}_1, \boldsymbol{p}_2, \ldots, \boldsymbol{p}_r], \qquad \boldsymbol{Q} = [\boldsymbol{q}_1, \boldsymbol{q}_2, \ldots, \boldsymbol{q}_r].$$

The proposed method defines a dissimilarity between $\mathcal{P}$ and $\mathcal{Q}$ by using distance between $\boldsymbol{p}_k$ and $\boldsymbol{q}_k$. In order to be invariant to sign indefiniteness of eigenvectors, the following function is introduced as their distance.

$$l^2(\boldsymbol{p}_k, \boldsymbol{q}_k) = \min\left(\|\boldsymbol{p}_k - \boldsymbol{q}_k\|^2, \|\boldsymbol{p}_k + \boldsymbol{q}_k\|^2\right),$$

where $\min(\cdot, \cdot)$ indicates a function which returns the minimum value of the arguments. The proposed dissimilarity is defined as

$$d^2(\mathcal{P}, \mathcal{Q}) = \sum_{k=1}^{r} l^2(\boldsymbol{p}_k, \boldsymbol{q}_k). \tag{3}$$

If $\boldsymbol{p}_k$ or $\boldsymbol{q}_k$ is sign-reversed in advance such that $l^2(\boldsymbol{p}_k, \boldsymbol{q}_k) = \|\boldsymbol{p}_k - \boldsymbol{q}_k\|^2$, Eq. (3) can be rewritten as

$$d^2(\mathcal{P}, \mathcal{Q}) = \|\boldsymbol{P} - \boldsymbol{Q}\|_{\mathrm{F}}^2 = \mathrm{tr}\left((\boldsymbol{P} - \boldsymbol{Q})^{\mathsf{T}}(\boldsymbol{P} - \boldsymbol{Q})\right),$$

where $\|\cdot\|_{\mathrm{F}}$ indicates Frobenius norm.

4 Experiments and Discussion

This section evaluates performance of the proposed descriptor by comparing with MSM, SCD and DCD. Table 1 lists the three test sets used in the experiments. OBV contains obverse side images (see Figure 1a), REV contains reverse ones(see Figure 1b), and ALL is simply the mix of OBV and REV. All the images are 24-bits RGB and each image captures one whole package, i.e. no offcuts. The image size is roughly from QVGA to VGA. Each test set is divided into two subsets for training and matching phases. The training phase determines the prototype subspace of each category by applying PCA to the averaged covariance matrix of all the color distributions which belongs to the category. The test environment is Intel Core2 Quad 2.33GHz CPU with 2GB memory. The used implementation of SCD and DCD is available from [10].

Table 2: Size and calculation time[ms/package].

Descriptor	Size	Training	Matching
SCD	256	6.85	7.04
DCD	KC at least	374.05	245.21
MSM	K^2 at most	**1.70**	2.72
Proposed	K^2 at most	**1.70**	**1.79**

4.1 Compactness, complexity, and accuracy

Table 2 indicates the description size of the competitors. The proposed descriptor is one of the most compact methods. SCD requires 256 elements to represent a quantized HSV color space into $8 \times 8 \times 4$ bins. The other methods require far fewer elements than SCD. MSM and the proposed descriptor require at most nine elements in RGB case. Note that their degree-of-freedom is smaller than the size due to $(\boldsymbol{\phi}_i, \boldsymbol{\phi}_j) = \delta_{ij}$.

Table 2 also indicates their training and matching time. The proposed descriptor shows the shortest time in them. DCD is severely slower than the others and rather insufficient for smooth dispensing operation, since clustering process runs heavy iterative computation. The others are sufficiently fast. Matching of MSM is nearly 1[ms/package] slower than that of the proposed descriptor because of PCA to $\boldsymbol{A}$ for each matching.

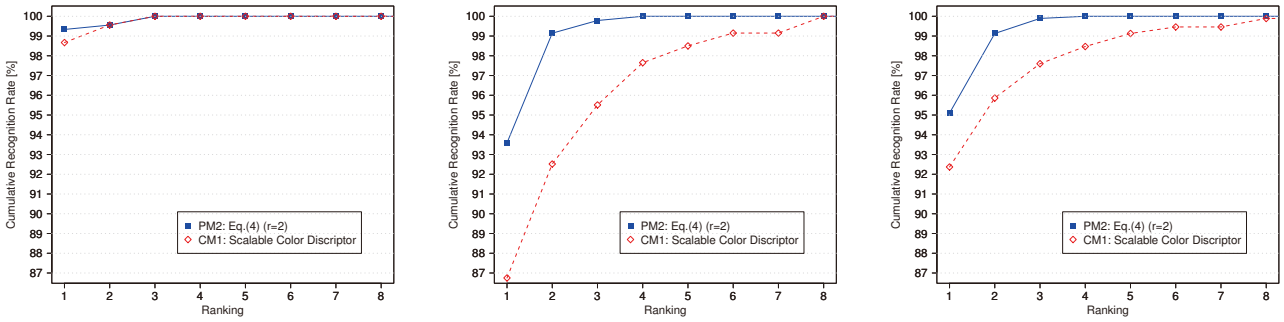
Table 3 lists the recognition rate and underlines the best rate in $r = 1, 2, 3$ for each test set of each method. Comparison between CM2 and PM1 reveals efficiency of the distance-based dissimilarity against canonical angles. MSM are inefficient in $r = 2, 3$. Next, comparison between PM1 and PM2 shows that weighting approach is effective for color matching. PM2 achieves the highest rate in the competitors. Figure 3 shows the cumulative recognition rate of SCD and the proposed descriptor (PM2). The proposed descriptor dominantly outperforms SCD in REV and ALL. Table 4 lists incorrect answers of some queries. In view of human eye sense, both packages of each pair share their dominant colors.

4.2 The number of clusters and dimensions

The relationship between C and r, as mentioned in Section 2.1, can be examined also in Table 3. First, OBV shows higher rate than REV in every case. Because obverse side shows in more colors than reverse side in most cases on account of a pill color. More colors cause easier category separation. Most obverse and reverse side images in the test sets show in two and three colors respectively. Under the fact that an r-dimensional subspace can characterize up to $C = (r + 1)$ clusters, $r = 2$ and $r = 1$ are appropriate to OBV and REV respectively. Actually, PM1 shows the best rate at the parameters in OBV and REV. Furthermore, rate difference between $r = 1, 2$ also supports our hypothesis. The diminution in OBV (6.44[%]) is clearly bigger than that in REV (1.28[%]). Second eigenvector is more meaningful for OBV than for REV. The reason is attributed to the relationship between C and r.

Table 3: Recognition rate[%]. ('×' indicates invalid parameter setting)

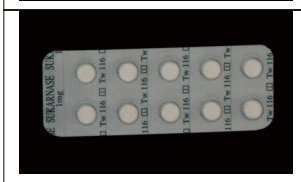
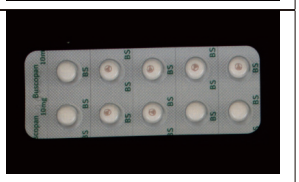
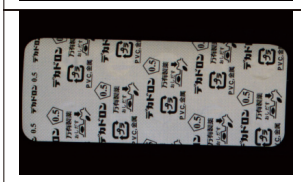
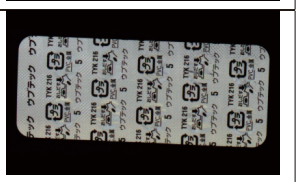
		OBV			REV			ALL		
		$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$	$r=1$	$r=2$	$r=3$
CM1	SCD[1]	98.67			86.75			92.37		
CM2	MSM[7], Eq. (2)	82.00	59.78	×	77.78	15.60	×	68.63	29.85	×
CM3	SIKK2010[6]	×	×	98.67	×	×	83.33	×	×	88.24
PM1	Eq. (3) with $\lambda_k = 1$	82.00	84.44	83.11	77.78	36.75	35.04	68.63	54.58	53.16
PM2	Eq. (3)	92.89	99.33	99.11	92.31	93.59	92.52	89.00	95.10	94.77



(a) OBV (b) REV (c) ALL

Figure 3: Cumulative recognition rate.

Table 4: Queries and their incorrect answers.

Query Image	Incorrect Answer
	
	
	

5 Conclusions

This paper presented a color matching technique using weighted subspace on medicine package recognition. Our experiments revealed that the proposed descriptor outperformed MSM and MPEG-7 low-level color descriptors in terms of training/matching time, description size and recognition rate. Thus, the proposed descriptor is efficient as a low-level color descriptor. However, in view of medicine package recognition, color-based approach is insufficient to achieve reliable checking, because color approach cannot distinguish similar color categories. Combination of another view point algorithms, e.g. symbol recognition, is desired for more accurate recognition.

Acknowledgment

This work was supported by the Grant-in-Aid for Japan Society for the Promotion of Science (JSPS).

References

[1] T. Sikora, "The MPEG-7 visual standard for content description — an overview," *IEEE Trans. on CSVT*, vol. 11, pp. 696–702, June 2001.

[2] J. Hafner, H. S. Sawhney, W. Equitz, M. Flickner, and W. Niblack, "Efficient color histogram indexing for quadratic form distance functions," *IEEE Trans. on PAMI*, vol. 17, pp. 729–736, July 1995.

[3] L. V. Tran and R. Lenz, "PCA-based representation of color distributions for color-based image retrieval," *Proc. of ICIP*, vol. 2, pp. 697–700, Oct. 2001.

[4] Y. Deng, B. S. Manjunath, C. Kenney, M. Moore, and H. Shin, "An efficient color representation for image retrieval," *IEEE Trans. on IP*, vol. 10, pp. 140–147, Jan. 2001.

[5] N.-C. Yang, W.-H. Chang, C.-M. Kuo, and T.-H. Li, "A fast MPEG-7 dominant color extraction with new similarity," *J. of VCIR*, vol. 19, pp. 92–105, Feb. 2008.

[6] K. Sugimoto, K. Inoue, Y. Kuroki, and S. Kamata, "A color distribution descriptor for medicine package recognition," *Proc. of CJKPR*, pp. 64–69, Nov. 2010.

[7] K. Maeda and S. Watanabe, "A pattern matching method with local structure," *IEICE Trans.*, vol. J68-D, pp. 345–352, March 1985. (in Japanese).

[8] K. Fukui, O. Yamaguchi, K. Suzuki, and K. Maeda, "Face recognition under variable lighting condition with constrained mutual subspace method: learning of constraint subspace to reduce influence of lighting changes," *IEICE Trans.*, vol. J82-D-II, pp. 613–620, Apr. 1999. (in Japanese).

[9] F. Chatelin, *Eigenvalues of matrices*. John Wiley & Sons, 1 ed., August 1993.

[10] http://savvash.blogspot.com/.