Novel EFM-KNN Classifier and a New Color Descriptor for Image Classification

Abhishek Verma[1] and Chengjun Liu

Abstract—**We propose a new CGSF+PHOG descriptor and perform image classification using a novel EFM-KNN classifier, which combines the Enhanced Fisher Model (EFM) and the K Nearest Neighbor (KNN) decision rule. We integrate the oRGB-SIFT descriptor with other color SIFT features to produce the Color SIFT Fusion (CSF) and the Color Grayscale SIFT Fusion (CGSF) descriptors. The CGSF is integrated to the PHOG to obtain the novel CGSF+PHOG descriptor. The effectiveness of the proposed new descriptor and the classification method is evaluated using two grand challenge datasets: the Oxford flower database and the MIT scene database. The classification results using the EFM-KNN classifier show that (i) the CGSF+PHOG descriptor improves recognition performance upon other descriptors; and (ii) the oRGB-SIFT, the CSF, and the CGSF perform better than the other color SIFT descriptors.**

Index Terms— **EFM-KNN classifier, CGSF+PHOG descriptor, oRGB-SIFT descriptor, Color SIFT Fusion (CSF), Color Grayscale SIFT Fusion (CGSF), image classification.**

I. Introduction

IMAGE classification using color features has been shown to achieve high success rate due to the fact that color features contain significant discriminative information [1]-[4]. Color based image classification can be very useful in the identification of object and natural scene categories [4]. Color features derived from different color spaces display different properties. Two important criteria for color feature detectors are that they should be stable under varying viewing conditions, such as changes in illumination, shading, highlights, and they should have high discriminative power. Color features such as the color histogram, color texture and local invariant features provide varying degrees of success against image variations such as viewpoint and lighting changes, clutter and occlusions [5]-[7].

Lately, the detection and recognition of locally affine invariant regions has gained prominence [8]-[10]. Affine region detectors when combined with the intensity Scale-Invariant Feature Transform (SIFT) descriptor [9] are able to outperform many alternatives [8]. In this paper, we extend the SIFT descriptor to different color spaces including the oRGB color space [11]. We integrate the oRGB-SIFT descriptor with other color SIFT features to produce the Color SIFT Fusion (CSF) and the Color Grayscale SIFT Fusion (CGSF) [4]. Furthermore, we combine the CGSF with the Pyramid of Histograms of Orientation Gradients (PHOG) to obtain a new CGSF+PHOG descriptor for image category classification. Image classification is implemented using the novel EFM-KNN classifier, which combines the Enhanced Fisher Model (EFM) [12], [13] and the K Nearest Neighbor (KNN) decision rule. The effectiveness of the proposed classification method and descriptors will be evaluated using two grand challenge datasets: the Oxford flower database and the MIT scene database.

[1] corresponding author

A. Verma and C. Liu are with the Department of Computer Science, New Jersey Institute of Technology, University Heights, Newark, NJ 07102 USA (e-mail: {av56, chengjun.liu}@njit.edu).

II. Related Work

This section briefly surveys the recent work on color image retrieval. In recent years, use of color as a means to face recognition [1], [3], [14] and object and scene classification [4] has gained popularity. Color features can capture discriminative information by means of the color invariants, color histogram, color texture, etc. One of the earlier works is the color indexing system designed by Swain and Ballard, which uses the color histogram for image inquiry from a large image database [15]. More recent work on color based image classification appears in [1], [2], [4], [16] that propose several new color spaces and methods for face, object and scene category recognition. Evaluation of local color invariant descriptors is performed in [5]. Fusion of color models, color region detection and color edge detection have been investigated for representation of color images [7]. Key contributions in color, texture, and shape abstraction have been discussed in Datta et al. [6].

Efficient retrieval requires a robust feature extraction method that has the ability to learn meaningful low-dimensional patterns in spaces of very high dimensionality [10], [17], [18]. Low-dimensional representations are also important when one considers the intrinsic computational aspect. PCA has been widely used to perform dimensionality reduction for image indexing and retrieval [12], [19]. Recently, Support Vector Machine (SVM) classifier for multiple category recognition has gained popularity [20] though it suffers from the drawback of being computationally too expensive on large scale image classification tasks. The EFM classifier has achieved good success for the task of image based recognition [13], [21], [22].

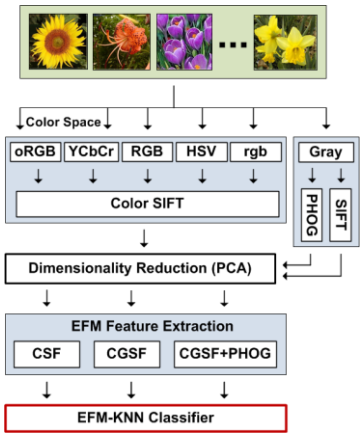

Fig. 1. Multiple feature fusion methodology using the EFM-KNN classifier.

III. Color Spaces and the New Color Descriptor

We first review in this section five color spaces in which our new color descriptor is defined, and then discuss five conventional SIFT descriptors: the RGB-SIFT, the rgb-SIFT, the HSV-SIFT, the YCbCr-SIFT, and the grayscale-SIFT descriptors. Next we present the oRGB-SIFT descriptor and two combined color SIFT descriptors: the Color SIFT Fusion (CSF), and the Color Grayscale SIFT Fusion (CGSF) [4]. Finally, we present a new CGSF+PHOG descriptor obtained from combining the CGSF with the Pyramid of Histograms of Orientation Gradients.

A color image contains three component images, and each pixel of a color image is specified in a color space, which serves as a color coordinate system. The commonly used color space is the RGB color space. Other color spaces are usually calculated from the RGB color space by means of either linear or nonlinear transformations.

To reduce the sensitivity of the RGB images to luminance, surface orientation, and other photographic conditions, the rgb color space is defined by normalizing the R, G, and B components. The HSV color space is motivated by human vision system because human describes color by means of hue, saturation, and brightness. Hue and saturation define chrominance, while intensity or value specifies luminance [23]. The YCbCr color space is developed for digital video standard and television transmissions. In YCbCr, the RGB components are separated into luminance, chrominance blue, and chrominance red.

The oRGB color space [11] has three channels L, C_1 and C_2. The primaries of this model are based on the three fundamental psychological opponent axes: white-black, red-green, and yellow-blue. The color information is contained in C_1 and C_2. The value of C_1 lies within [-1, 1] and the value of C_2 lies within [-0.8660, 0.8660]. The L channel contains the luminance information and its values range between [0, 1].

$$\begin{bmatrix} L \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.2990 & 0.5870 & 0.1140 \\ 0.5000 & 0.5000 & -1.0000 \\ 0.8660 & -0.8660 & 0.0000 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (1)$$

The SIFT descriptor proposed by Lowe transforms an image into a large collection of feature vectors, each of which is invariant to image translation, scaling, and rotation, partially invariant to the illumination changes, and robust to local geometric distortion [9]. The key locations used to specify the SIFT descriptor are defined as maxima and minima of the result of the difference of Gaussian function applied in the scale-space to a series of smoothed and resampled images. SIFT descriptors robust to local affine distortions are then obtained by considering pixels around a radius of the key location.

The grayscale-SIFT descriptor is defined as the SIFT descriptor applied to the grayscale image. A color SIFT descriptor in a given color space is derived by individually computing the SIFT descriptor on each of the three component images in the specific color space. This produces a 384 dimensional descriptor that is formed from concatenating the 128 dimensional vectors from the three channels. As a result, the four color SIFT descriptors are defined: the RGB-SIFT, the YCbCr-SIFT, the HSV-SIFT, and the rgb-SIFT descriptors.

The Pyramid of Histograms of Orientation Gradients (PHOG) descriptor [24] is able to represent an image by its local shape and the spatial layout of the shape. The local shape is captured by the distribution over edge orientations within a region and the spatial layout by tiling the image into regions at multiple resolutions. The distance between two PHOG image descriptors then reflects the extent to which the images contain similar shapes and correspond in their spatial layout [24].

The four color SIFT descriptors are defined in the oRGB color space and the fusion in different color spaces. In particular, the oRGB-SIFT descriptor is constructed by concatenating the SIFT descriptors of the three component images in the oRGB color space. The Color SIFT Fusion (CSF) descriptor is formed by fusing the RGB-SIFT, the YCbCr-SIFT, the HSV-SIFT, the oRGB-SIFT, and the rgb-SIFT descriptors. The Color Grayscale SIFT Fusion (CGSF) descriptor is obtained by fusing further the CSF descriptor and

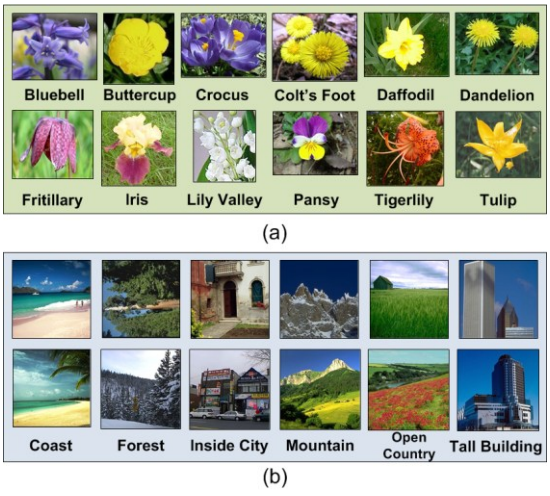

Fig. 2. Example images from the (a) Oxford flower dataset and (b) the MIT scene dataset.

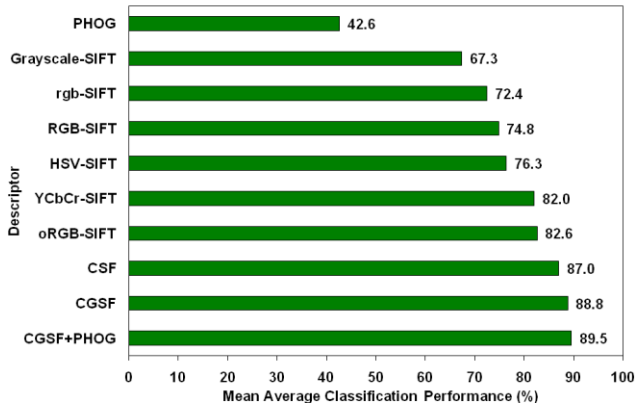

Fig. 3. The mean average classification performance of the ten descriptors fusing the EFM-KNN classifier on the Oxford flower dataset: the oRGB-SIFT, the YCbCr-SIFT, the RGB-SIFT, the HSV-SIFT, the rgb-SIFT, the grayscale-SIFT, the PHOG, the CSF, the CGSF, and the CGSF+PHOG descriptors.

the grayscale-SIFT descriptor. The CGSF is combined with the PHOG to obtain the new CGSF+PHOG color descriptor.

IV. The Novel EFM-KNN Classifier

Image classification using the new descriptor introduced in the preceding section is implemented using the novel EFM-KNN classifier [12], [13], which combines the Enhanced Fisher Model (EFM) and the K Nearest Neighbor (KNN) decision rule [25]. Let $X \in \mathrm{R}^N$ be a random vector whose covariance matrix is Σ_X:

$$\Sigma_X = \varepsilon\{[X - \varepsilon(X)][X - \varepsilon(X)]^t\} \tag{2}$$

where $\varepsilon(.)$ is the expectation operator and t denotes the transpose operation. The eigenvectors of the covariance matrix Σ_X can be derived by PCA:

$$\Sigma_X = \Phi \Lambda \Phi^t \tag{3}$$

where $\Phi = [\phi_1 \phi_2 \phi_3 .. \phi_N]$ is an orthogonal eigenvector matrix and $\Lambda = diag\{\lambda_1, \lambda_2, ..., \lambda_N\}$ a diagonal eigenvalue matrix with diagonal elements in decreasing order. An important application of PCA is dimensionality reduction:

$$Y = P^t X \tag{4}$$

where $P = [\phi_1 \phi_2 \phi_3 .. \phi_K]$, and $K < N$. $X \in \mathrm{R}^K$ thus is composed of the most significant principal components. PCA, which is derived based on an optimal representation criterion, usually does not lead to good image classification performance. To improve upon PCA, the Fisher Linear Discriminant (FLD) analysis [25] is introduced to extract the most discriminating features.

The FLD method optimizes a criterion defined on the within-class and between-class scatter matrices, S_w and S_b [25]:

$$S_w = \sum_{i=1}^{L} P(\omega_i) \varepsilon\{(Y - M_i)(Y - M_i)^t \mid \omega_i\} \tag{5}$$

$$S_b = \sum_{i=1}^{L} P(\omega_i)(M_i - M_0)(M_i - M_0)^t \tag{6}$$

where $P(\omega_i)$ is a priori probability, ω_i represent the classes, and M_i and M are the means of the classes and the grand mean, respectively. The criterion the FLD method optimizes is $J_1 = tr(S_w^{-1} S_b)$, which is maximized when Ψ contains the eigenvectors of the matrix $S_w^{-1} S_b$ [25]:

$$S_w^{-1} S_b \Psi = \Psi \Delta \tag{7}$$

where Ψ, Δ are the eigenvector and eigenvalue matrices of $S_w^{-1} S_b$, respectively. The FLD discriminating features are defined by projecting the pattern vector Y onto the eigenvectors of Ψ:

$$Z = \Psi^t Y \tag{8}$$

Z thus is more effective than the feature vector Y derived by PCA for image classification.

The FLD method, however, often leads to overfitting when implemented in an inappropriate PCA space. To improve the generalization performance of the FLD method, a proper balance between two criteria should be maintained: the energy criterion for adequate image representation and the magnitude criterion for eliminating the small-valued trailing eigenvalues of the within-class scatter matrix [12]. A new method, the Enhanced Fisher Model (EFM), is capable of improving the generalization performance of the FLD method [12]. Specifically, the EFM method improves the generalization capability of the FLD method by decomposing the FLD procedure into a simultaneous diagonalization of the within-

TABLE I
Category Wise Descriptor Performance (%) Split-out With the EFM-KNN Classifier on the Oxford Flower Dataset. Note that the Categories are Sorted on the CGSF+PHOG Results

Category	CGSF+ PHOG	CGSF	CSF	oRGB SIFT	YCbCr SIFT	Gray SIFT
Daisy	**100**	98	98	**100**	98	93
Sunflower	**100**	**100**	**100**	**100**	**100**	95
Windflower	**100**	98	92	92	92	90
Tigerlily	**98**	**98**	97	**98**	95	78
Dandelion	**95**	93	92	92	92	82
Bluebell	**93**	**93**	90	85	79	49
Colt's Foot	93	93	**95**	90	93	83
Lily Valley	**93**	91	90	82	80	78
Pansy	91	**93**	89	76	78	75
Cowslip	87	**90**	88	84	88	46
Iris	**87**	85	80	75	72	78
Buttercup	84	**85**	84	83	82	49
Fritillary	84	**85**	83	80	83	75
Snowdrop	**83**	81	78	62	60	63
Daffodil	82	80	**83**	78	73	45
Crocus	**75**	69	68	63	59	25
Tulip	**75**	74	73	64	70	37
Mean	**89.5**	**88.8**	**87.0**	**82.6**	**82.0**	**67.3**

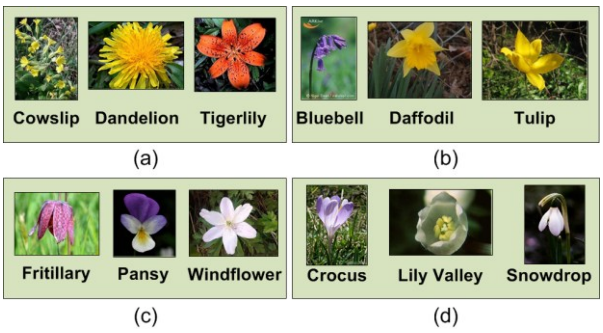

Fig. 4. Image recognition using the EFM-KNN classifier on the Oxford flower dataset: (a) example images unrecognized using the grayscale-SIFT descriptor but recognized using the oRGB-SIFT descriptor; (b) example images unrecognized using the oRGB-SIFT descriptor but recognized using the CSF descriptor; (c) images unrecognized using the CSF but recognized using the CGSF descriptor; (d) images unrecognized using the CGSF but recognized using the CGSF+PHOG descriptor.

class and between-class scatter matrices [12]. The simultaneous diagonalization is stepwise equivalent to two operations as pointed out by Fukunaga [25]: whitening the within-class scatter matrix and applying PCA to the between-class scatter matrix using the transformed data. The stepwise operation shows that during whitening the eigenvalues of the within-class scatter matrix appear in the denominator. Since the small (trailing) eigenvalues tend to capture noise [12], they cause the whitening step to fit for misleading variations, which leads to poor generalization performance. To achieve enhanced performance, the EFM method preserves a proper balance between the need that the selected eigenvalues account for most of the spectral energy of the raw data (for representational adequacy), and the requirement that the eigenvalues of the within-class scatter matrix (in the reduced PCA space) are not too small (for better generalization performance) [12].

Image classification is implemented using the EFM-KNN classifier, and Fig. 1 shows the fusion methodology of multiple descriptors using the EFM-KNN classifier.

V. Experiments

A. Dataset and Experimental Methodology

We apply the following two publicly accessible datasets to evaluate our proposed descriptors and classification method: the Oxford Flower dataset [26] and the MIT Scene dataset [27]. The Oxford flower dataset [26] consists of 17 species of flowers with 80 images per category. All the images are in color, JPEG format and the mean image size is 560x560 pixels. There are species that have a very unique visual appearance, e.g. Fritillaries and Tigerlilies, as well as species with very similar appearance, for example Dandelions and Coltsfoot. The large intra-class variability and the small inter-class variability make this dataset very challenging. See Fig. 2(a) for some sample images. The MIT scene dataset [27] has 2,688 images classified as eight categories: 360 coast, 328 forest, 374 mountain, 410 open country, 260 highway, 308 inside of cities, 356 tall buildings, and 292 streets. All of the images are in color, in JPEG format, and the average size of each image is 256x256 pixels. There is a large variation in light, pose and angles, along with a high intra-class variation. The sources of the images vary (from commercial databases, websites, and digital cameras) [27]. See Fig. 2(b) for some sample images from this dataset.

TABLE II
Comparison of the Classification Performance (%) with Other Methods on the Oxford Flower Dataset

Our Method		[26]		[28]	
RGB-SIFT	74.8	Color	73.7	Shape	68.88
HSV-SIFT	76.3	Shape	71.8	Color	59.71
YCbCr-SIFT	82.0	Texture	56.0*	Texture	59.00
oRGB-SIFT	**82.6**				
CSF	**87.0**				
CGSF	**88.8**				
CGSF+PHOG	**89.5**	Fusion	81.3	Fusion	82.55

* Approximate value inferred from Fig. 12 in [26]

The classification task is to assign each test image to one of a number of categories. The performance is measured using a confusion matrix, and the overall performance rates are measured by the average value of the diagonal entries of the confusion matrix. For the Oxford flower dataset we select three sets of 40 training images per class and 20 test images per class (same data splits as used in [26]). For the MIT scene dataset we randomly select five sets and each set consists of 2000 images for training (250 images per class) and the rest 688 images for testing. Within each set there is no overlap in the images selected for training and testing. The classification scheme on the datasets compares the overall and category wise performance of ten different descriptors: the oRGB-SIFT, the YCbCr-SIFT, the RGB-SIFT, the HSV-SIFT, the rgb-SIFT, the grayscale-SIFT, the PHOG, the CSF, the CGSF, and the CGSF+PHOG descriptors. Classification is implemented using the EFM-KNN classifier.

B. Evaluation of Color Descriptors and EFM-KNN Classifier on the Oxford Flower Dataset

The first set of experiments assesses the overall

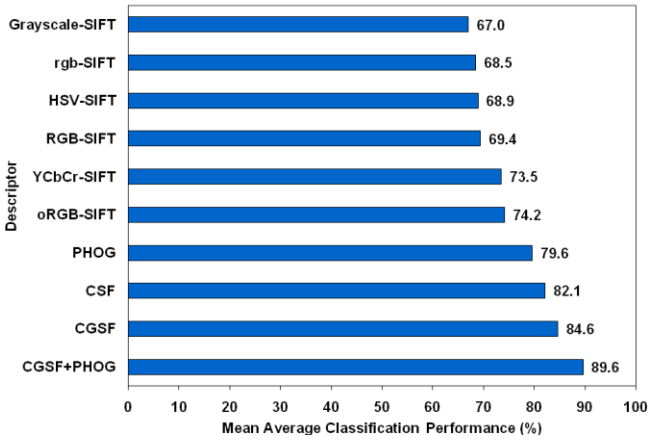

Fig. 5. The mean average classification performance of the ten descriptors using the EFM-KNN classifier on the MIT scene dataset: the oRGB-SIFT, the YCbCr-SIFT, the RGB-SIFT, the HSV-SIFT, the rgb-SIFT, the grayscale-SIFT, the PHOG, the CSF, the CGSF, and the CGSF+PHOG descriptors.

TABLE III
CATEGORY WISE DESCRIPTOR PERFORMANCE (%) SPLIT-OUT WITH THE EFM-KNN CLASSIFIER ON THE MIT SCENE DATASET. NOTE THAT THE CATEGORIES ARE SORTED ON THE CGSF+PHOG RESULTS

Category	CGSF+ PHOG	CGSF	CSF	oRGB SIFT	Gray SIFT	PHOG
Highway	**100**	**100**	**100**	**100**	82	91
Forest	**99**	97	97	88	90	95
Inside City	**97**	90	88	75	86	80
Coast	**91**	77	66	63	65	84
Street	90	**93**	89	90	60	86
Mountain	**88**	82	79	69	60	75
Tall Building	**85**	81	80	65	62	71
Open Country	**67**	58	57	43	32	56
Mean	**89.6**	**84.6**	**82.1**	**74.2**	**67.0**	**79.6**

classification performance of the ten descriptors on the Oxford flower dataset. Note that for each category we implement three-fold cross validation for each descriptor using the EFM-KNN classifier to derive the average classification performance. As a result, each descriptor yields 17 average classification rates corresponding to the 17 image categories. The mean value of these 17 average classification rates is defined as the mean average classification performance for the descriptor. Fig. 3 shows the mean average classification performance of various descriptors.

The best recognition rate that we obtain is 89.5% from the CGSF+PHOG, which is a very respectable value for a dataset of this size and complexity. The oRGB-SIFT achieves the classification rate of 82.6%. It outperforms the other color SIFT descriptors. It is noted that fusion of the color SIFT descriptors (CSF) improves upon the grayscale-SIFT by a huge 19.7% margin. The grayscale-SIFT descriptor improves the fusion (CGSF) result by a good 1.8% margin upon the CSF descriptor.

The second set of experiments assesses the five best descriptors and the grayscale-SIFT using the EFM-KNN classifier on individual image categories. From Table I it can be seen that three categories achieve 100% success rate and over 50% of the categories achieve a success rate of more than 90% with the CGSF+PHOG descriptor. Sunflower achieves 100% success rate across the best five descriptors. Crocus and Tulip have a success rate of 75% indicating high intra-class variability and low inter-class variability. The average success rate for the top 10 categories with the CGSF+PHOG descriptor is a respectable 95%. Individual color SIFT features improve upon the grayscale-SIFT on most of the categories. The CSF almost always improves upon the grayscale-SIFT, this indicates that various color descriptors are not redundant. The CGSF improves upon the CSF; furthermore most categories perform at their best when we combine the PHOG with the CGSF.

The final set of experiments further assesses the performance of the descriptors based on the correctly recognized images. See Fig. 4(a) for some example images that are not recognized by the EFM-KNN classifier using the grayscale-SIFT descriptor but are correctly recognized using the oRGB-SIFT descriptor. This reaffirms the importance of color and the distinctiveness of the oRGB-SIFT descriptor for

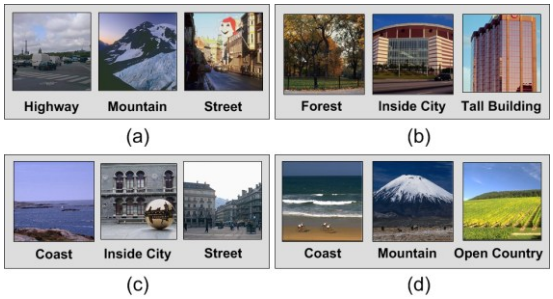

Fig. 6. Image recognition using the EFM-KNN classifier on the MIT scene dataset: (a) example images unrecognized using the grayscale-SIFT descriptor but recognized using the oRGB-SIFT descriptor; (b) example images unrecognized using the oRGB-SIFT descriptor but recognized using the CSF descriptor; (c) images unrecognized using the CSF but recognized using the CGSF descriptor; (d) images unrecognized using the CGSF but recognized using the CGSF+PHOG descriptor.

image category recognition. Fig. 4(b) shows images unrecognized using the oRGB-SIFT descriptor but recognized using the CSF descriptor, Fig. 4(c) shows images unrecognized using the CSF but recognized using the CGSF descriptor and Fig. 4(d) shows images unrecognized using the CGSF but recognized using the CGSF+PHOG descriptor.

Table II shows a comparison of our results with those obtained by Nilsback [26] and Varma [28]. Our technique outperforms the state of the art on this dataset even without combining color descriptors or considering texture and shape features independently. Each of the four color SIFT descriptors outperform descriptors in [26], [28]. Combined SIFT descriptors (CSF, CGSF and CGSF+PHOG) improve over the fusion result in [26] and SVM 1-vs-All fusion result in [28], previously the best result on this dataset.

C. Evaluation of Color Descriptors and EFM-KNN Classifier on the MIT Scene Dataset

The first set of experiments on this dataset assesses the overall classification performance of the ten descriptors. Note that for each category we implement five-fold cross validation for each descriptor using the EFM-KNN classifier to derive the average classification performance. Fig. 5 shows the mean average classification performance of various descriptors. The best recognition rate that we obtain is 89.6% from the CGSF+PHOG, which is a very respectable value for a dataset of this size and complexity. The oRGB-SIFT achieves the classification rate of 74.2%. It outperforms the other color SIFT descriptors. It is noted that fusion of the color SIFT descriptors (CSF) improves upon the grayscale-SIFT by a huge 15.1% margin. The grayscale-SIFT descriptor improves

TABLE IV
COMPARISON OF THE CLASSIFICATION PERFORMANCE (%) WITH OTHER METHOD ON THE MIT SCENE DATASET

# train	# test	Our Method		[27]
2000	688	CSF	**82.1**	-
		CGSF	**84.6**	
		CGSF+PHOG	**89.6**	
800	1888	CSF	79.4	
		CGSF	81.9	
		CGSF+PHOG	**86.7**	83.7

the fusion (CGSF) result by a good 2.5% margin upon the CSF descriptor.

The second set of experiments assesses the five best descriptors and the grayscale-SIFT using the EFM-KNN classifier on individual image categories. From Table III it can be seen that the top five categories achieve a success rate of over 90%. The Highway category achieves a success rate of 100% across the best four descriptors. Individual color SIFT features improve upon the grayscale-SIFT on most of the categories. The CSF results on each of the eight categories show improvement upon the grayscale-SIFT and the CGSF improves upon the CSF. Integration of PHOG with the CGSF to obtain the CGSF+PHOG highly benefits most categories and in particular there is a significant increase in the classification performance upon the CGSF results for the Coast, Inside City and Open Country categories where the increment is in the range of 7% to 14%.

The final set of experiments further assesses the performance of the descriptors based on the correctly recognized images. See Fig. 6(a) for some example images that are not recognized by the EFM-KNN classifier using the grayscale-SIFT descriptor but are correctly recognized using the oRGB-SIFT descriptor. Fig. 6(b) shows images unrecognized using the oRGB-SIFT descriptor but recognized using the CSF descriptor, Fig. 6(c) shows images unrecognized using the CSF but recognized using the CGSF descriptor and Fig. 6(d) shows images unrecognized using the CGSF but recognized using the CGSF+PHOG descriptor.

From Table IV it can be seen that on the 800 training images (100 images per class) and 1688 testing images we achieve 86.7% success rate with the CGSF+PHOG descriptor. This improves over the result of authors in [27] by a good 3%.

VI. Conclusion

We have proposed a new CGSF+PHOG descriptor and performed image classification using a novel EFM-KNN classifier, which combines the Enhanced Fisher Model (EFM) and the K Nearest Neighbor (KNN) decision rule. Results of the experiments using two grand challenge datasets show that our oRGB-SIFT descriptor improves recognition performance upon other color SIFT descriptors. The fusion of Color SIFT descriptors (CSF), Color Grayscale SIFT descriptor (CGSF), and the new CGSF+PHOG descriptor show significant improvement in the classification performance, which indicates that various color-SIFT descriptors and grayscale-SIFT descriptor are not redundant for image classification.

References

[1] C. Liu and J. Yang, "ICA color space for pattern recognition," *IEEE Trans. on Neural Networks*, vol. 20, no. 2, pp. 248-257, 2009.

[2] J. Yang and C. Liu, "Color image discriminant models and algorithms for face recognition," *IEEE Transactions on Neural Networks*, vol. 19, no. 12, pp. 2088-2098, 2008.

[3] P. Shih and C. Liu, "Comparative assessment of content-based face image retrieval in different color spaces," *Int. Journal of Pattern Recognition and Artificial Intelligence*, vol. 19, no. 7, pp. 873-893, 2005.

[4] A. Verma, S. Banerji, and C. Liu, "A new color SIFT descriptor and methods for image category classification," *in the 2010 Int. Congress on Computer Applications and Computational Science*, pp. 819-822, Singapore.

[5] G. Burghouts and J.M. Geusebroek, "Performance evaluation of local color invariants," *Computer Vision and Image Understanding*, vol. 113, pp. 48-62, 2009.

[6] R. Datta, D. Joshi, J. Li, and J. Wang, "Image retrieval: Ideas, influences, and trends of the new age," *ACM Computing Surveys*, vol. 40, no. 2, pp. 509-522, 2008.

[7] H. Stokman and T. Gevers, "Selection and fusion of color models for image feature detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 29, no. 3, pp. 371-381, 2007.

[8] K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L. Van Gool, "A comparison of affine region detectors," *Int. Journal of Computer Vision*, vol. 65, nos. 1-2, pp. 43-72, 2005.

[9] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. Journal of Computer Vision*, vol. 60, no. 2, pp. 91-110, 2004.

[10] C. Liu, "A bayesian discriminating features method for face detection," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 25, no. 6, pp. 725-740, 2003.

[11] M. Bratkova, S. Boulos, and P. Shirley, "oRGB: a practical opponent color space for computer graphics," *IEEE Computer Graphics and Applications*, vol. 29, no. 1, pp. 42-55, 2009.

[12] C. Liu and H. Wechsler, "Robust coding schemes for indexing and retrieval from large face databases," *IEEE Trans. on Image Processing*, vol. 9, no. 1, pp. 132-137, 2000.

[13] C. Liu and H. Wechsler, "Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition," *IEEE Trans. on Image Processing*, vol. 11, no. 4, pp. 467-476, 2002.

[14] C. Liu, "Capitalize on dimensionality increasing techniques for improving face recognition grand challenge performance," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 28, no. 5, pp. 725-737, 2006.

[15] M.J. Swain and D.H. Ballard, "Color indexing," *Int. Journal of Computer Vision*, vol. 7, no. 1, pp. 11-32, 1991.

[16] C. Liu, "Learning the uncorrelated, independent, and discriminating color spaces for face recognition," *IEEE Trans. on Information Forensics and Security*, vol. 3, no. 2, pp. 213-222, 2008.

[17] C. Liu and H. Wechsler, "Evolutionary pursuit and its application to face recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 6, pp. 570-582, 2000.

[18] C. Liu and H. Wechsler, "Independent component analysis of gabor features for face recognition," *IEEE Trans. on Neural Networks*, vol. 14, no. 4, pp. 919-928, 2003.

[19] C. Liu, "Gabor-based kernel with fractional power polynomial models for face recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 26, no. 5, pp. 572-581, 2004.

[20] J. Zhang, M. Marszalek, S. Lazebnik, and C. Schmid, "Local features and kernels for classification of texture and object categories a comprehensive study," *Int. Journal of Computer Vision*, vol. 73, no. 2, pp. 213-238, 2007.

[21] C. Liu and H. Wechsler, "A shape and texture based enhanced fisher classifier for face recognition," *IEEE Trans. on Image Processing*, vol. 10, no. 4, pp. 598-608, 2001.

[22] C. Liu, "Enhanced independent component analysis and its application to content based face image retrieval," *IEEE Trans. Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 2, pp. 1117-1127, 2004.

[23] C.G. Gonzalez and R.E. Woods, *Digital Image Processing*, Upper Saddle River, NJ: Prentice-Hall, 2001.

[24] A. Bosch, A. Zisserman, and X. Munoz, "Representing shape with a spatial pyramid kernel," *in the 2007 Int. Conf. on Image and Video Retrieval*, pp. 401-408, New York, NY.

[25] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, San Diego, CA: Academic Press, 2nd ed., 1990.

[26] M. E. Nilsback and A. Zisserman, "A visual vocabulary for flower classification," *in the 2006 Int. Conf. on Computer Vision and Pattern Recognition*, vol. 2, pp. 1447-1454, Washington, DC.

[27] A. Oliva and A. Torralba, "Modeling the shape of the scene: a holistic representation of the spatial envelope," *Int. Journal of Computer Vision*, vol. 42, no. 3, pp. 145-175, 2001.

[28] M. Varma and D. Ray, "Learning the discriminative power-invariance trade-off," *in the 2007 IEEE Int. Conf. on Computer Vision*, pp. 1-8, Rio de Janeiro.