A Novel Inheritable Color Space with Application to Kinship Verification

Qingfeng Liu Ajit Puthenputhussery Chengjun Liu
New Jersey Institute of Technology
ql69, avp38, cliu@njit.edu

Abstract

Anthropology studies discover that some genetic related facial features, which are inherited by children from their parents, can be used for kinship verification. This paper investigates an important inheritable feature — color and presents a novel inheritable color space (InCS) and a generalized InCS (GInCS) framework with application to kinship verification. Specifically, a novel color similarity measure (CSM) is first defined. Second, based on this similarity measure, a new inheritable color space (InCS) is derived by balancing the criterion of minimizing the distance between kinship pairs and the criterion of maximizing the distance between non-kinship pairs. Unlike conventional color spaces, e.g. the RGB color space, the proposed InCS, which is learned automatically from the data, captures the inheritable information between parent and child. Third, theoretical and empirical analysis show that the proposed InCS exhibits the decorrelation property, which is positively related to the performance of kinship verification. Robustness to the illumination variation is also discussed. Fourth, a generalized InCS framework is presented to extend the InCS from the pixel level to the feature level for improving the performance and the robustness to illumination variation. The proposed InCS is evaluated on several popular datasets, namely the KinFaceW-I dataset, the KinFaceW-II dataset, the UB KinFace dataset, and the Cornell KinFace dataset. Experimental results show that the proposed InCS is able to (i) improve the conventional color spaces such as RGB, YUV, YIQ color spaces by a large margin, (ii) achieve robustness to the illumination variation, and (iii) outperforms other popular methods.

1. Introduction

Kinship verification from facial images, which is an emerging research area in computer vision, has gained increasing attention in recent years [4], [29], [19], [3], [32]. Pioneer works in anthropology [1], [2] believe that there are some genetic related features which are inherited by children from their parents that can be used to determine the kinship relations.

One important inheritable feature is color. Our observation shows that color information, which is complementary to shape and texture, usually leads to better performance for kinship verification (the cross-ethnicity parents and children are not considered in this paper since such cases are minority). Conventional color spaces such as RGB, YUV, YIQ, YCbCr have shown their superiority for face recognition [24]. Recently many color methods [9, 10, 11, 34, 26] are proposed for object recognition, object detection and action recognition. These methods, which focus on illumination invariant properties of color, are not specifically designed for capturing the inheritable information between kinship images for recognizing the kinship relations.

This paper thus presents a novel inheritable color space and a generalized InCS framework for kinship verification. Specifically, a new color similarity measure (CSM), which represents the accumulation of the similarity measures of corresponding color component between two images, is first defined. A novel inheritable color space (InCS) is then automatically derived by trading off the criterion of minimizing the distance between the kinship pairs and the criterion of maximizing the distance between the non-kinship pairs. Some properties of InCS are further discussed. Theoretical analysis shows that the proposed InCS possesses the decorrelation property, which decorrelates the color components of the color difference matrix between two images. Experimental analysis also discovers that the decorrelation property, which is measured in terms of average statistical correlation (ASC), often leads to better performance. The robustness to the illumination variation is also discussed. Finally, a generalized InCS (GInCS) framework is proposed by extending the InCS from the pixel level to the feature level for improving the performance and the illumination invariance. An example is presented by applying the Fisher vector [8] to the RGB component images prior to the computation of the component images in the InCS.

Experimental results on four representative datasets: the KinFaceW-I dataset, the KinFaceW-II dataset [19], the UB KinFace dataset [29], and the Cornell KinFace dataset [4] show the feasibility of the proposed method.

2. Related Work

Kinship verification. The pioneer works of kinship analysis originate from anthropology and psychology community. Bressnan et al. [2] evaluated the phenotype matching on facial features and claimed that parents have correlated visual resemblance with their offspring. Studies [1] in anthropology have confirmed that children resemble their parent more than other people and they may resemble a particular parent more at different ages. Later work [4] by Fang et al. shows the feasibility of applying computer vision techniques for kinship verification. Xia et al. [29] proposed a transfer subspace learning based algorithm by using the young parents set as an intermediate set to reduce the significant divergence in the appearance distributions between children and old parents facial images. Lu et al. [19] proposed neighborhood repulsed metric learning (NRML) in which the intraclass samples within a kinship relation are pulled as close as possible and interclass samples are pushed as far as possible for kinship verification. Dehghan et al. [3] proposed to apply the generative and the discriminative gated autoencoders to learn the genetic features and metrics together for kinship verification. Yan et al. [31] proposed a multimetric learning method to combine different complementary feature descriptors for kinship verification and later [32] proposed to learn the discriminative mid-level features by constructing a reference dataset instead of using handcrafted descriptors.

Color space. Color information contributes significantly to the discriminative power of image representation. Conventional color spaces such as RGB, YUV, YIQ, YCbCr have shown their ability for improving the performance of face recognition [24, 16, 12, 13]. For a detail comparison among different color spaces, please refer [24]. Van de Sande et al. [26] show that color information along with shape features yield excellent results in image classification system. Khan et al. [10] proposed the use of color attributes as an explicit color representation for object detection. Zhang et al. [34] proposed a new biologically inspired color image descriptor that uses a hierarchical nonlinear spatio-chromatic operator yielding spatial and chromatic opponent channels. Khan et al. [22] show that better results can be obtained for object recognition by explicitly separating the color cue to guide attention by means of a top-down category-specific attention map. Yang et al. [33] proposed a new color model - the $g_1g_2g_3$ model based on the log chromacity color space, which preserves the relationship between R, G and B in the model. Rahat Khan et al. [11] cluster color values together based on their discriminative power such that the drop of mutual information of the final representation is minimized.

3. A Novel Inheritable Color Space (InCS)

Conventional color spaces such as YUV, YIQ, YCbCr, Opponent color space etc., have shown their ability for recognition problems by considering the illumination invariant properties. However, they are not deliberately designed to capture the inheritable information between kinship images for recognizing the kinship relations. We therefore present a novel inheritable color space (InCS) for kinship verification by deriving a transformation $\mathbf{W} \in \mathbb{R}^{3\times3}$ from the original RGB color space.

In detail, given a pair of images I_{p_i} and I_{c_i} $(i = 1, 2, ..., m)$ with a size of $h \times w$, these two images can be represented as two matrices $\mathbf{m}_{p_i}$ and $\mathbf{m}_{c_i} \in \mathbb{R}^{3\times n}$ $(n = h \times w)$ respectively, where each row vector is the concatenation of column pixels in each color component (red, green and blue) of the image. Note that in Section 5, each row vector will be a feature vector (e.g. Fisher vector) for each color component. Formally, $\mathbf{m}_{p_i}$ and $\mathbf{m}_{c_i}$ are defined as $\mathbf{m}_{p_i} = [\mathbf{p}_{i1}, \mathbf{p}_{i2}, ..., \mathbf{p}_{in}]$ and $\mathbf{m}_{c_i} = [\mathbf{c}_{i1}, \mathbf{c}_{i2}, ..., \mathbf{c}_{in}]$ respectively, where $\mathbf{p}_{ij}$ and $\mathbf{c}_{ij} \in \mathbb{R}^{3\times1}(j = 1, 2, ..., n)$ are vectors that consists of the color values of three color components at each pixel for two compared images respectively. Then the new representation $\mathbf{x}_i$ and $\mathbf{y}_i \in \mathbb{R}^{3\times n}(i = 1, 2, ..., m)$ in InCS is computed from $\mathbf{m}_{p_i}$ and $\mathbf{m}_{c_i}$ by applying the transformation $\mathbf{W}$ as follows:

$$\begin{aligned} \mathbf{x}_i &= \mathbf{W}^T\mathbf{m}_{p_i} = [\mathbf{x}_{i1}, \mathbf{x}_{i2}, ..., \mathbf{x}_{in}] \\ \mathbf{y}_i &= \mathbf{W}^T\mathbf{m}_{c_i} = [\mathbf{y}_{i1}, \mathbf{y}_{i2}, ..., \mathbf{y}_{in}] \end{aligned} \tag{1}$$

where $\mathbf{x}_{ij}, \mathbf{y}_{ij} \in \mathbb{R}^{3\times1}(j = 1, 2, ..., n)$.

To incorporate the information of all the three color components, we first define a new color similarity measure (CSM). Let $\mathbf{x}_i(c) \in \mathbb{R}^{n\times1}(c = 1, 2, 3)$ be the c-th color component vector of the InCS in image I_{p_i}, whose values are identical to the c-th row of $\mathbf{x}_i$. And $\mathbf{y}_i(c)$ is defined similarly. Then the CSM has the following mathematical form:

$$\begin{aligned} d(I_{p_i}, I_{c_i}) &= \mathbf{Tr}\{(\mathbf{x}_i - \mathbf{y}_i)(\mathbf{x}_i - \mathbf{y}_i)^T\} \\ &= \sum_{j=1}^{n}(\mathbf{p}_{ij} - \mathbf{c}_{ij})^T\mathbf{A}(\mathbf{p}_{ij} - \mathbf{c}_{ij}) \end{aligned} \tag{2}$$

where $\mathbf{Tr}\{\cdot\}$ denotes the trace of a matrix and $\mathbf{A} = \mathbf{W}\mathbf{W}^T \in \mathbb{R}^{3\times3}$. Note that since $\mathbf{W}$ is a transformation applied to the RGB color space, it is reasonable to assume it is full rank. As a result, $\mathbf{A}$ is a positive semidefinite matrix, which guarantees CSM is a metric — satisfying the non-negativity and the triangle inequality [30].

From equation 2, the CSM may be interpreted in two ways. On the one hand, the CSM may be interpreted as the accumulation of the distance metrics between corresponding pixel color values of two images. On the other hand, the CSM may also be interpreted as a summation of the

Euclidean similarity measures between each corresponding color component, which is denoted as $\sum_{c=1}^{3}(\mathbf{x}_i(c) - \mathbf{y}_i(c))^T(\mathbf{x}_i(c) - \mathbf{y}_i(c))$.

Then based on such a similarity measure, our inheritable color space (InCS) is derived with the goal of pushing away the non-kinship samples as far as possible while keeping the kinship ones as close as possible by optimizing the following objective function:

$$\min_{\mathbf{W}} \sum_{(I_{p_i}, I_{c_i}) \in T} d(I_{p_i}, I_{c_i}) - \sum_{(I_{p_i}, I_{c_i}) \in F} d(I_{p_i}, I_{c_i}) \quad (3)$$

where T and F are the sets of kinship pairs and non-kinship pairs respectively. Note that the ratio of two terms can also be applied as an objective function.

To optimize equation 3, we further define two matrices, namely the true verification matrix $\mathbf{V}_t = \sum_{(I_{p_i}, I_{c_i}) \in T} \sum_{j=1}^{n} (\mathbf{p}_{ij} - \mathbf{c}_{ij})(\mathbf{p}_{ij} - \mathbf{c}_{ij})^T \in \mathbb{R}^{3\times 3}$ that characterizes the color variations among the kinship image pairs, and the false verification matrix $\mathbf{V}_f = \sum_{(I_{p_i}, I_{c_i}) \in F} \sum_{j=1}^{n} (\mathbf{p}_{ij} - \mathbf{c}_{ij})(\mathbf{p}_{ij} - \mathbf{c}_{ij})^T \in \mathbb{R}^{3\times 3}$ that captures the color variations among all the non-kinship image pairs. Note that the definitions of $\mathbf{V}_t$ and $\mathbf{V}_f$ are similar to the within-scatter matrix and the between-scatter matrix defined in discriminant analysis [6].

As a result, the objective function in equation 3 can be rewritten as equation 4 by introducing a constraint $\mathbf{W}^T\mathbf{V}_t\mathbf{W} = \mathbf{I}$ on $\mathbf{W}$ to exclude the trivial solution. The solution $\mathbf{W}$ of optimizing the objective function in equation 4 consists of the eigenvectors of matrix $\mathbf{V}_t^{-1}(\mathbf{V}_t - \mathbf{V}_f)$. Finally, the InCS is derived from the original RGB color space by applying the learned transformation matrix $\mathbf{W}$.

$$\begin{aligned} &\min_{\mathbf{W}} \mathbf{Tr}(\mathbf{W}^T(\mathbf{V}_t - \mathbf{V}_f)\mathbf{W}) \\ &s.t. \mathbf{W}^T\mathbf{V}_t\mathbf{W} = \mathbf{I} \end{aligned} \quad (4)$$

It can be discovered from equation 4 that the proposed InCS seeks to balance the criterion of minimizing the color variations among the kinship image pairs and the criterion of maximizing the color variations among the non-kinship image pairs. Therefore, InCS is capable of capturing the color variations between the kinship images, which is a favorable property for kinship verification.

4. Properties of the InCS

This section presents the properties of the proposed InCS. Specifically, we first show that our proposed InCS exhibits the decorrelation property and establish its connection to the performance. Second, we analyze the robustness of our InCS to illumination variations at both pixel level and feature level based on the diagonal illumination model [5, 26].

4.1. The Decorrelation Property

Given the derived transformation matrix $\mathbf{W} = [\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3]$ where $\mathbf{w}_i \in \mathbb{R}^{3\times 1}(i = 1, 2, 3)$, the pixel color values in vector $\mathbf{x}_{ij} = [x_{ij1}, x_{ij2}, x_{ij3}]^T$ and $\mathbf{y}_{ij} = [y_{ij1}, y_{ij2}, y_{ij3}]^T$ can be represented as $x_{iju} = \mathbf{w}_u^T\mathbf{p}_{ij}$, $y_{iju} = \mathbf{w}_u^T\mathbf{c}_{ij}$, $x_{ijv} = \mathbf{w}_v^T\mathbf{p}_{ij}$ and $y_{ijv} = \mathbf{w}_v^T\mathbf{c}_{ij}$ where $u, v = 1, 2, 3$ and $u \neq v$. Note that x_{iju} and x_{ijv} are the pixel values of the u-th component and the v-th component of $\mathbf{x}_{ij}$ in the InCS respectively. y_{iju} and y_{ijv} are defined in a similar fashion.

For kinship verification, an important variable is the difference matrix $\mathbf{d}_i$ for the corresponding pixels between two images, which is computed as $\mathbf{d}_i = \mathbf{x}_i - \mathbf{y}_i$, because the similarity between two images can be represented in terms of $\mathbf{d}_i$ if some similarity measures, such as the CSM or the Euclidean distance, are applied. As a result, the decorrelation of the components of $\mathbf{d}_i$ can reduce the information redundancy for measuring the similarity and enhance the kinship verification performance.

We therefore present the decorrelation property of the proposed InCS as follows. It states that the u-th component (row) $\mathbf{d}_i(u)$ and the v-th component $\mathbf{d}_i(v)$ $(u, v = 1, 2, 3$ and $u \neq v)$ of the difference matrix $\mathbf{d}_i$ are decorrelated.

Property 4.1. ***The Decorrelation Property***. *If each component of the color difference matrix $\boldsymbol{d}_i$ is centered, we have the following statistical correlation $S(u, v)$ between the u-th component (row) $\boldsymbol{d}_i(u)$ and the v-th component $\boldsymbol{d}_i(v)$ $(u, v = 1, 2, 3$ and $u \neq v)$ of the difference matrix $\boldsymbol{d}_i$.*

$$\begin{aligned} S(u, v) &= \mathcal{E}\left\{(\boldsymbol{d}_i(u) - \mathcal{E}\left\{\boldsymbol{d}_i(u)\right\})(\boldsymbol{d}_i(v) - \mathcal{E}\left\{\boldsymbol{d}_i(v)\right\})^T\right\} \\ &= 0 \end{aligned} \quad (5)$$

Proof. First, $\mathcal{E}\left\{\mathbf{d}_i(u)\right\}$ and $\mathcal{E}\left\{\mathbf{d}_i(v)\right\}$ are zero since each component of the color difference matrix $\mathbf{d}_i$ is centered. Second, the solution of optimizing equation 4 shows that $\mathbf{W}$ consists of the eigenvectors of matrix $\mathbf{V}_t^{-1}(\mathbf{V}_t - \mathbf{V}_f)$, which can be further proved to be the same as the eigenvectors of matrix $(\mathbf{V}_t + \mathbf{V}_f)^{-1}\mathbf{V}_f$. In other words, we have the following results

$$\begin{aligned} &(\mathbf{V}_t + \mathbf{V}_f)^{-1}\mathbf{V}_f\mathbf{W} = \mathbf{W}\Lambda \\ \Rightarrow &\mathbf{W}^T(\mathbf{V}_t + \mathbf{V}_f)\mathbf{W}\Lambda = \mathbf{W}^T\mathbf{V}_t\mathbf{W} \\ \Rightarrow &\mathbf{W}^T(\mathbf{V}_t + \mathbf{V}_f)\mathbf{W} = \Lambda^{-1} \end{aligned} \quad (6)$$

where Λ is a diagonal matrix that is composed of the eigenvalues of $(\mathbf{V}_t + \mathbf{V}_f)^{-1}\mathbf{V}_f$.

Then we have the following statistical correlation $S(u, v)$ between the u-th and v-th color components $(u, v =$

$1, 2, 3$ and $u \neq v$) for $\mathbf{x}_i$ and $\mathbf{y}_i$ as equation 7:

$$\begin{aligned} S(u,v) &= \mathcal{E}(\mathbf{d}_i(u)\mathbf{d}_i(v)^T) \\ &= \frac{1}{m}\sum_{i=1}^{m}\mathbf{d}_i(u)\mathbf{d}_i(v)^T \\ &= \frac{1}{m}\sum_{i=1}^{m}(\mathbf{x}_i(u)-\mathbf{y}_i(u))(\mathbf{x}_i(v)-\mathbf{y}_i(v))^T \\ &= \frac{1}{m}\sum_{i=1}^{m}\sum_{j=1}^{n}(x_{iju}-y_{iju})(x_{ijv}-y_{ijv})^T \\ &= \frac{1}{m}\mathbf{w}_u^T(V_t+V_f)\mathbf{w}_v \\ &= 0 \end{aligned} \tag{7}$$

□

To further reveal the advantage of the decorrelation property, we conduct experiments by using the average statistical correlation (ASC) defined below as an indicator of the degree of decorrelation for different color spaces.

$$ASC = \frac{1}{3}\sum_{u\neq v}|S(u,v)| \tag{8}$$

where $|\cdot|$ is the absolute value. Experimental results in Section 7.2 show that in general, the smaller the ASC is, the better the performance the corresponding color space achieves.

4.2. Robustness to Illumination Variations

Another important property of the proposed InCS is the robustness to illumination variations. The illumination variations of an image can be modeled by the diagonal model [5, 26], which corresponds to the Lambertian reflectance model under the assumption of narrow band filters. Specifically, The diagonal model is defined as a diagonal transformation $\mathbf{L} \in \mathbb{R}^{3\times3}$ on the RGB values of each pixel and a shift $\mathbf{s} \in \mathbb{R}^{3\times1}$ as follows:

$$\begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}\begin{pmatrix} R \\ G \\ B \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \\ s_3 \end{pmatrix} \tag{9}$$

where a, b and c are the diagonal elements of matrix $\mathbf{L}$, and s_1, s_2 and s_3 are the elements of vector $\mathbf{s}$.

Based on this diagonal model, five types of common illumination variations, namely light intensity change, light intensity shift, light intensity change and shift, light color change as well as light color change and shift, can be identified [26]. First, the light intensity change assumes $a = b = c > 0$ and $s_1 = s_2 = s_3 = 0$, which means the pixel values change by a constant factor in all color components. The light intensity change is often caused by the differences of the intensity of light, shadows and shading. Second, the light intensity shift assumes $a = b = c = 1$ and $s_1 = s_2 = s_3 > 0$. The light intensity shift is mainly due to the diffuse lighting. Third, the light intensity change and shift combines the above two changes and assumes $a = b = c > 0$ and $s_1 = s_2 = s_3 > 0$. Fourth, the light color change assumes $a \neq b \neq c > 0$ and $s_1 = s_2 = s_3 = 0$, which means each color component scales independently. Finally, the light color change and shift assumes $a \neq b \neq c > 0$ and $s_1 \neq s_2 \neq s_3 > 0$.

Now we present theoretical analysis to show the following conclusion under the illumination condition that the diagonal transformation $\mathbf{L}$ applied is the same within each image pair (parent and child image in this image pair have the same $\mathbf{L}$) but different across image pairs and the shift $\mathbf{s}$ is applied similarly.

- Our InCS is robust to light intensity change and light color change.
- Our InCS becomes robust to light intensity shift, light intensity change and shift as well as light color change and shift with the help of the proposed CSM or the Euclidean distance to cancel out the shift.

To prove that, let $\mathbf{V}_t^u$ and $\mathbf{V}_f^u$ be the new true verification matrix and the new false verification matrix, respectively, after the illumination variations are applied to each original image pair (I_{p_i}, I_{c_i}) by the diagonal transformation $\mathbf{L}_i$ and the shift $\mathbf{s}_i$ $(i = 1, 2, ..., m)$. Note that $\mathbf{L}_i$ may be different across image pairs but are the same within each image pair, and so is $\mathbf{s}_i$. Then we have $\mathbf{V}_t^u = \mathbf{L}_i^t\mathbf{V}_t\mathbf{L}_i$ and $\mathbf{V}_f^u = \mathbf{L}_i^t\mathbf{V}_f\mathbf{L}_i$. Then the new transformation $\mathbf{W}^u$ has the relation to the original transformation $\mathbf{W}$ as $\mathbf{W} = \mathbf{L}_i\mathbf{W}^u$. As a result, the derived image pixel vectors for InCS with the illumination variations are as follows:

$$\begin{aligned} (\mathbf{W}^u)^T(\mathbf{L}_i^t\mathbf{p}_{ij}+\mathbf{s}_i) &= \mathbf{W}^T\mathbf{p}_{ij} + (\mathbf{W}^u)^T\mathbf{s}_i \\ (\mathbf{W}^u)^T(\mathbf{L}_i^t\mathbf{c}_{ij}+\mathbf{s}_i) &= \mathbf{W}^T\mathbf{c}_{ij} + (\mathbf{W}^u)^T\mathbf{s}_i \end{aligned} \tag{10}$$

First, if the shift $\mathbf{s}_i$ is not considered ($s_1 = s_2 = s_3 = 0$), then the image pixel vectors for InCS with the illumination variations are the same as those without illumination variations, which means our InCS is robust to light intensity change and light color change regardless of the value of $\mathbf{L}_i$.

Second, if the shift $\mathbf{s}_i$ is considered ($s_1 \neq s_2 \neq s_3 > 0$), when our CSM or the Euclidean distance is applied, it is easy to see that the effect of the offset term $(\mathbf{W}^u)^T\mathbf{s}_i$ is canceled out when comparing the parent and child image. Then our InCS is also robust to light intensity shift, light intensity change and shift as well as light color change and shift with the help of the proposed CSM.

However, the pixel level inheritable color space cannot guarantee the robustness to illumination variations under

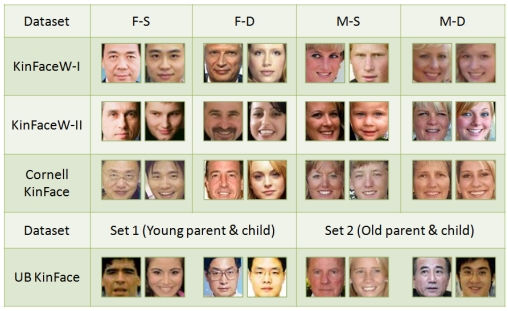

Figure 1. Example images from the KinFaceW-I, the KinFaceW-II dataset, the Cornell KinFace dataset and the UB KinFace dataset.

a more general illumination condition: both the diagonal transformation $\mathbf{L}_i$ and shift $\mathbf{s}_i$ are different for parent and child image within the kinship image pair respectively. In this case, we will appeal to the feature level inheritable color space, namely the generalized InCS (GInCS) described in Section 5. We will show our GInCS is also robust to illumination variations in Section 5 under a more general illumination condition.

For empirical evaluation of the robustness of the InCS, please refer to Section 7.3, where we evaluate the kinship verification performance under different illumination conditions based on the five types of illumination variations described above.

5. The Generalized InCS (GInCS) Framework

Our novel InCS can be further generalized from the pixel level to the feature level to enhance the generalization performance and improve the robustness to image variabilities, such as illumination variations. One example is to apply the Fisher vector method [8] to the red, green, and blue component images prior to the computation of the component images in the InCS. Note that other feature extraction methods, such as Gabor wavelets [14], LBP [28], Feature LBP [7] may be applied as well to define new generalized InCS for kinship verification.

We first briefly review the Fisher vector method [8], which has been widely applied for visual recognition problems such as face recognition [25], object recognition [8]. Specifically, let $\mathbf{X} = \{\mathbf{d}_t, t = 1, 2, ..., T\}$ be the set of T local descriptors (e.g. SIFT descriptors) extracted from the image. Let $\mu_{\boldsymbol{\lambda}}$ be the probability density function of $\mathbf{X}$ with a set of parameters $\boldsymbol{\lambda}$, then the Fisher kernel [8] is defined as $K(\mathbf{X}, \mathbf{Y}) = (\mathbf{G}^{\mathbf{X}}_{\boldsymbol{\lambda}})^T \mathbf{F}^{-1}_{\boldsymbol{\lambda}} \mathbf{G}^{\mathbf{Y}}_{\boldsymbol{\lambda}}$, where $\mathbf{G}^{\mathbf{X}}_{\boldsymbol{\lambda}} = \frac{1}{T} \bigtriangledown_{\boldsymbol{\lambda}} \log[\mu_{\boldsymbol{\lambda}}(\mathbf{X})]$, which is the gradient vector of the log-likelihood that describes the contribution of the parameters to the generation process. And $\mathbf{F}_{\boldsymbol{\lambda}}$ is the Fisher information matrix of $\mu_{\boldsymbol{\lambda}}$. Essentially, the Fisher vector is derived from the explicit decomposition of the Fisher kernel according to the fact that the symmetric and positive definite Fisher information matrix $\mathbf{F}_{\boldsymbol{\lambda}}$ has a Cholesky decomposition as $\mathbf{F}^{-1}_{\boldsymbol{\lambda}} = \mathbf{L}^T_{\boldsymbol{\lambda}} \mathbf{L}_{\boldsymbol{\lambda}}$. Therefore, the Fisher kernel $K(\mathbf{X}, \mathbf{Y})$ can be written as a dot product between two vectors $\mathbf{L}_{\boldsymbol{\lambda}} \mathbf{G}^{\mathbf{X}}_{\boldsymbol{\lambda}}$ and $\mathbf{L}_{\boldsymbol{\lambda}} \mathbf{G}^{\mathbf{Y}}_{\boldsymbol{\lambda}}$ which are defined as the **Fisher vectors** of $\mathbf{X}$ and $\mathbf{Y}$ respectively. A diagonal closed-form approximation of $\mathbf{F}_{\boldsymbol{\lambda}}$ [8] is often used where $\mathbf{L}_{\boldsymbol{\lambda}}$ is just a whitening transformation for each dimension of $\mathbf{G}^{\mathbf{X}}_{\boldsymbol{\lambda}}$ and $\mathbf{G}^{\mathbf{Y}}_{\boldsymbol{\lambda}}$. The dimensionality of Fisher vector depends only on the number of parameters in the parameter set $\boldsymbol{\lambda}$.

We now present the generalized InCS framework by using the Fisher vector as an example. First, the Fisher vector is computed for each color components in the RGB color space. Second, these Fisher vectors maybe taken as the input of the learning process by optimizing the objective function in equation 4 to derive the Fisher vector based generalized InCS. Finally, all the derived GInCS component vectors are normalized and concatenated as one augmented component vector for kinship verification.

It is easy to see that our GInCS preserves the decorrelation property since the learning process remains the same. Besides, our GInCS is also robust to illumination variations under a more general illumination condition: the diagonal transformation $\mathbf{L}$ and the shift $\mathbf{s}$ are different within and between the image pairs, respectively. First, our implementation is based on the Fisher vector, which depends on the SIFT descriptors extracted from each color component image. The SIFT descriptor is intensity shift-invariant [26] since it is based on the gradient of image which takes the derivative to cancel out the intensity shift. Therefore, our GInCS is robust to intensity shift operations of the diagonal model. Second, the SIFT descriptor is often normalized so that it is robust to light intensity change under any diagonal transformation. Third, the SIFT descriptor is computed and normalized for each color component image independently so that it is also robust to light color change [26]. As a result, our Fisher vector based generalized InCS is robust to light intensity change, light color change, light intensity shift, light intensity change and shift as well as light color change and shift under a more general illumination condition. The generalized InCS framework with Fisher vector can significantly improve the results and outperform other popular methods (see Section 6.1).

6. Experiments

Our proposed method is assessed on four representative kinship verification databases: the KinFaceW-I dataset, the KinFaceW-II dataset [19], the UB KinFace dataset [29], and the Cornell KinFace dataset [4]. Example images are shown in figure 1.

The pattern vector we implement for InCS is the concatenation of column pixels of the RGB color space. The pattern vector for the generalized InCS is the Fisher vector. First, in each color component, the dense SIFT feature is

Methods	F-S	F-D	M-S	M-D	Mean
CSML [21]	61.10	58.10	60.90	70.00	62.50
LMNN [27]	63.10	58.10	62.90	70.00	63.30
NRML [19]	64.10	59.10	63.90	71.00	64.30
MNRML [19]	72.50	66.50	66.20	72.00	69.90
GGA [3]	70.50	70.00	67.20	74.30	70.50
ANTH [3]	72.50	71.50	70.80	75.60	72.60
DGA [3]	76.40	72.50	71.90	77.30	74.50
DMML [31]	74.50	69.50	69.50	75.50	72.25
MPDFL [32]	73.50	67.50	66.10	73.10	70.10
GInCS	**77.25**	**76.90**	**75.82**	**81.44**	**77.85**

Table 1. Comparison between the GInCS and other popular methods on the KinFaceW-I dataset.

derived with a step size of 1 and five scale patch sizes as 2, 4, 6, 8, 10. The dimensionality 128 of the SIFT feature is further reduced to 64 by PCA. And the spatial information [25] is added to the SIFT feature with 2 more dimensions which means the final dimensionality of the SIFT feature is 66. Then, a Gaussian mixture model with 256 components is derived. As a result, Fisher vector is derived as 33792 (2*256*66) dimension vector. Power transformation [8] is applied to the extracted Fisher vector. Euclidean distance or our color similarity measure can be further applied to compute the similarity between two images. Finally, a two class linear support vector machine is used to determine the kinship relations between images.

6.1. KinFaceW-I and KinFaceW-II dataset

The KinFaceW-I and the KinfaceW-II dataset contain four kinship relations: father-son (F-S), father-daughter (F-D), mother-son (M-S), and mother-daughter (M-D). In the KinFaceW-I dataset, there are 156, 134, 116, and 127 image pairs for each relation. In the KinFaceW-II dataset, there are 250 pairs of the images for each relation. In our experiments, we conduct 5-fold cross validation where both the KinFaceW-I dataset and the KinFaceW-II dataset are divided into five folds having the same number of image pairs. The SIFT flow [15] is extracted to pre-process the images so that the corresponding parts between two images are enhanced. Our novel color similarity measure is applied to compute the distance between two images. Experimental results on table 1, table 2, show that our method is able to outperform other popular methods on both datasets. Note that although better performances are reported in [17], they use multiple features while our method uses only one single feature to achieve good results.

6.2. UB KinFace dataset

The UB KinFace dataset consists of 600 images of 400 persons from 200 families. For each family, there are three images, which correspond to child, young parent and old parent. In our experiments, images are aligned according to the eye position and cropped to 64×64 such that the background information is removed and only the facial region is used for kinship verification. Two subsets of images are constructed, where set 1 consists of 200 child and young parent pairs and set 2 consists of 200 child and old parent pairs. The 5-fold cross validation is also conducted. Experimental results on table 3 show that our method is able to outperform other popular methods.

Methods	F-S	F-D	M-S	M-D	Mean
CSML [21]	71.80	68.10	73.80	74.00	71.90
LMNN [27]	74.80	71.10	75.80	76.00	74.50
NRML [19]	76.80	73.10	76.80	77.00	75.70
MNRML [19]	76.90	74.30	77.40	77.60	76.50
DMML [31]	78.50	76.50	78.50	79.50	78.25
MPDFL [32]	77.30	74.70	77.80	78.00	77.00
GInCS	**85.40**	**77.00**	**81.60**	**81.60**	**81.40**

Table 2. Comparison between the GInCS and other popular methods on the KinFaceW-II dataset.

Methods	Set 1	Set 2
MCCA [23]	65.50	64.00
MMFA [23]	65.00	64.00
LDDM [20]	66.50	66.00
DMMA [18]	65.50	63.50
MNRML [19]	66.50	65.50
DMML [31]	74.50	70.00
GInCS	**75.80**	**72.20**

Table 3. Comparison between the GInCS and other popular methods on the UB KinFace dataset.

6.3. Cornell KinFace dataset

The Cornell KinFace dataset contains 143 pairs of kinship images where 40%, 22%, 13% and 26% are the father-son (F-S), father-daughter (F-D), mother-son (M-S), and mother-daughter (M-D) relations. In our experiments, the images are preprocessed in the same fashion as the UB KinFace dataset and 5-fold cross validation is conducted for each relation respectively. Experimental results on table 4 show that our method is able to outperform other popular methods.

7. Comprehensive Analysis

This section presents a comprehensive analysis of our proposed InCS method including (i) the comparison with other color spaces, (ii) analysis of decorrelation property and (iii) robustness to the illumination variations. All the experiments are conducted on the the KinFaceW-I dataset and the KinFaceW-II dataset. Without other specification,

Methods	F-S	F-D	M-S	M-D	Mean
MCCA [23]	71.50	65.80	73.50	63.50	68.57
MMFA [23]	71.50	66.40	73.50	64.50	68.97
LDDM [20]	73.00	66.90	74.50	67.50	70.47
DMMA [18]	71.00	65.50	73.00	65.50	68.75
MNRML [19]	74.50	68.80	77.20	65.80	71.57
DMML [31]	76.00	70.50	77.50	71.00	73.75
GInCS	**78.20**	**73.00**	**78.80**	**73.50**	**75.87**

Table 4. Comparison between the GInCS and other popular methods on the Cornell KinFace dataset.

Color Spaces	F-S	F-D	M-S	M-D	Mean
Grey	58.01	51.88	49.93	52.83	53.16
RGB	66.60	56.74	55.16	60.25	59.69
rgb	59.96	57.81	62.90	65.70	61.59
CIE-XYZ	65.32	55.60	54.73	61.05	59.18
YUV	66.65	61.52	58.64	63.85	62.67
YIQ	65.70	60.03	59.93	61.76	61.86
YCbCr	66.03	60.77	57.77	64.56	62.28
Opponent	66.98	61.58	58.24	65.42	63.06
InCS	66.35	60.06	60.76	**69.67**	**64.21**

Table 5. Comparison between the InCS and other color spaces on the KinFaceW-I dataset.

the process of the experiments is as follows. The concatenation of column pixels is applied first in the InCS to derive the pattern vector for each color component. Then the pattern vectors in each component are normalized and concatenated as an augmented pattern vector. The euclidean distance is further applied to calculate the similarity between different images.

7.1. Performance of Different Color Spaces

This section presents the performance of different color spaces including the RGB, rgb (normalized RGB), YUV, YIQ, YCbCr, CIE-XYZ and Opponent color space. In order to conduct a fair comparison, the pattern vector is extracted by concatenating the column pixels in each color components of each image for different color spaces. Experimental results demonstrated in table 5 and table 6 shows that the proposed InCS is able to achieve better performance than the original RGB color space by a large margin especially on the KinFaceW-II dataset which has a larger size of training images.

7.2. The Decorrelation Property

This section investigates the decorrelation property of our proposed InCS in terms of the average statistical correlation (ASC) for all the image pairs (including the training and testing pairs). It can be concluded from the empirical results in table 7 and table 8: (i) our proposed InCS is able to achieve the lowest value of ASC since it can capture the color variations between two images and decorrelates the color components of the difference matrices between two images; (ii) the final performances of different color spaces are quite related to the value of ASC and lower value of ASC usually will result in better performance.

Color Spaces	F-S	F-D	M-S	M-D	Mean
Grey	56.60	54.60	55.20	54.80	55.30
RGB	68.20	60.80	59.60	57.80	61.60
rgb	69.20	57.20	61.40	61.60	62.35
CIE-XYZ	67.00	60.00	60.20	58.20	61.35
YUV	71.40	62.40	65.60	64.00	65.85
YIQ	73.60	62.00	66.20	64.80	66.65
YCbCr	71.60	59.80	66.80	64.60	65.70
Opponent	71.80	59.80	69.00	64.40	66.25
InCS	**73.60**	**62.40**	**71.40**	**71.80**	**69.80**

Table 6. Comparison between the InCS and other color spaces on the KinFaceW-II dataset.

Color Spaces	ASC	Performance
RGB	0.93	59.69
rgb	0.57	61.59
CIE-XYZ	0.97	59.18
YUV	0.63	62.67
YIQ	0.58	61.86
YCbCr	0.59	62.28
Opponent	0.55	63.06
InCS	**0.26**	**64.21**

Table 7. The values of average statistical correlation of different color spaces on the KinFaceW-I dataset.

Color Spaces	ASC	Performance
RGB	0.91	61.60
rgb	0.60	62.35
CIE-XYZ	0.96	61.35
YUV	0.63	65.85
YIQ	0.56	66.65
YCbCr	0.60	65.70
Opponent	0.59	66.25
InCS	**0.30**	**69.80**

Table 8. The values of average statistical correlation of different color spaces on the KinFaceW-II dataset.

7.3. The Robustness to Illumination Variations

To show the robustness to illumination variation for our proposed method, two groups of experiments are conducted for InCS and GInCS respectively.

In particular, the first group of experiments evaluates the robustness of InCS to illumination variations under a re-

KinFaceW-I	F-S	F-D	M-S	M-D	Mean
InCS	66.35	60.06	60.76	69.67	64.21
InCS + LI	66.35	60.06	60.76	69.67	64.21
InCS + LC	66.35	60.06	60.76	69.67	64.21
KinFaceW-II	F-S	F-D	M-S	M-D	Mean
InCS	73.60	62.40	71.40	71.80	69.80
InCS + LI	73.60	62.40	71.40	71.80	69.80
InCS + LC	73.60	62.40	71.40	71.80	69.80

Table 9. Performance on the KinFaceW-I and the KinFaceW-II datasets when illumination changes. "LI" means light intensity change and "LC" means light color change.

stricted illumination condition: for each pair of images, the same diagonal transformation **L** and the same shift **s** are applied, while for different pairs, the diagonal transformation **L** and the shift **s** may be different. We conduct the following experiments for this group: (i) First, the diagonal transformation **L**, whose elements fall into (0, 1], is randomly generated and the elements of **s** are set to zero for evaluating the robustness to light intensity change (LI) and light color change (LC). As shown in table 9, our InCS is indeed invariance to light intensity change and light color change under the restricted illumination condition. (ii) Second, the diagonal transformation **L** is randomly generated and so is the shift **s** for evaluating the robustness to light intensity shift (LS), light intensity change and shift (LIs) as well as light color change and shift (LCs). The Euclidean distance is applied. Transformed image pixel values outside the range [0, 255] are still kept but normalized later. As shown in table 10, our InCS is indeed invariant to light intensity shift, light intensity change and shift as well as light color change and shift under the restricted illumination condition with the help of our CSM or the Euclidean distance. Note that all the experiments are conducted for five times and the average results are reported. The results are exactly the same for all five iterations with our CSM or the Euclidean distance. But cosine similarity produces slightly different results for each iteration.

The second group of experiments assesses the robustness of GInCS to illumination variations under a more general illumination condition: the diagonal transformation **L** and the shift **s** are different for all the images (no matter within or between the image pairs). Similarly, transformed image pixel values outside the range [0, 255] are still kept but normalized later. The results in table 11 show that our GInCS method achieves illumination robustness to light intensity change, light color change, light intensity shift, light intensity change and shift as well as light color change and shift. Note that all the experiments are also conducted for five times and the average results are reported. Due to the randomness of computing the Fisher vector (random sampling of SIFT descriptors to estimate the GMM model), sometimes slightly better results are produced.

KinFaceW-I	F-S	F-D	M-S	M-D	Mean
InCS	66.35	60.06	60.76	69.67	64.21
InCS + LS	66.35	60.06	60.76	69.67	64.21
InCS + LIs	66.35	60.06	60.76	69.67	64.21
InCS + LCs	66.35	60.06	60.76	69.67	64.21
KinFaceW-II	F-S	F-D	M-S	M-D	Mean
InCS	73.60	62.40	71.40	71.80	69.80
InCS + LS	73.60	62.40	71.40	71.80	69.80
InCS + LIs	73.60	62.40	71.40	71.80	69.80
InCS + LCs	73.60	62.40	71.40	71.80	69.80

Table 10. Performance on the KinFaceW-I and the KinFaceW-II datasets when illumination changes. "LS" means light intensity shift, "LIs" means light intensity change and shift, and "LCs" means light color change and shift.

KinFaceW-I	F-S	F-D	M-S	M-D	Mean
GInCS	77.25	76.90	75.82	81.44	77.85
GInCS + LI	77.25	76.54	77.99	80.98	78.19
GInCS + LC	76.60	76.90	76.70	80.20	77.60
GInCS + LS	77.25	76.51	77.99	81.07	78.20
GInCS + LIs	77.25	75.41	77.55	80.61	77.71
GInCS + LCs	77.25	76.15	77.12	79.81	77.58
KinFaceW-II	F-S	F-D	M-S	M-D	Mean
GInCS	85.40	77.00	81.60	81.60	81.40
GInCS + LI	85.29	77.60	81.80	81.60	81.57
GInCS + LC	85.93	77.00	82.00	81.40	81.58
GInCS + LS	84.97	77.00	82.20	80.40	81.14
GInCS + LIs	84.97	76.80	82.40	81.20	81.34
GInCS + LCs	84.97	77.00	82.80	80.80	81.39

Table 11. Performance of GInCS on the KinFaceW-I and the KinFaceW-II dataset for different illumination variations

8. Conclusion

This paper presents a novel inheritable color space for kinship verification. A novel color similarity measure is first defined. The InCS is then learned by pulling close kinship pairs and pushing away non-kinship pairs. Further analysis show that the InCS possesses the decorrelation property and the robustness to the illumination variation. Then a generalized InCS framework is presented to extend the InCS from the pixel level to the feature level. The evaluations show the feasibility of the proposed method.

References

[1] A. Alvergne, C. Faurie, and M. Raymond. Differential facial resemblance of young children to their parents: who do

children look like more? *Evolution and Human Behavior*, 28(2):135 – 144, 2007.

[2] P. Bressan and M. Dal Martello. Talis pater, talis filius: Perceived resemblance and the belief in genetic relatedness. *Psychological Science*, 13(3):213–218, 2002.

[3] A. Dehghan, E. G. Ortiz, R. Villegas, and M. Shah. Who do i look like? determining parent-offspring resemblance via gated autoencoders. In *CVPR*, pages 1757–1764. IEEE, 2014.

[4] R. Fang, K. D. Tang, N. Snavely, and T. Chen. Towards computational models of kinship verification. In *ICIP*, pages 1577–1580, 2010.

[5] G. D. Finlayson, M. S. Drew, and B. V. Funt. Spectral sharpening: sensor transformations for improved color constancy. *J. Opt. Soc. Am. A*, 11(5):1553–1563, May 1994.

[6] K. Fukunaga. *Introduction to Statistical Pattern Recognition (2Nd Ed.)*. Academic Press Professional, Inc., 1990.

[7] J. Gu and C. Liu. Feature local binary patterns with application to eye detection. In *Neurocomputing*, pages 138–152, 2013.

[8] H. Jegou, F. Perronnin, M. Douze, J. Sanchez, P. Perez, and C. Schmid. Aggregating local image descriptors into compact codes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 34(9):1704–1716, Sept 2012.

[9] F. S. Khan, M. A. Rao, J. van de Weijer, A. D. Bagdanov, A. Lopez, and M. Felsberg. Coloring action recognition in still images. *International Journal of Computer Vision (IJCV)*, 105(3):205–221, 2013.

[10] F. S. Khan, M. A. Rao, J. van de Weijer, A. D. Bagdanov, M. Vanrell, and A. Lopez. Color attributes for object detection. In *CVPR 2012*, 2012.

[11] R. Khan, J. van de Weijer, F. Shahbaz Khan, D. Muselet, C. Ducottet, and C. Barat. Discriminative color descriptors. In *Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on*, pages 2866–2873, June 2013.

[12] C. Liu. Extracting discriminative color features for face recognition. *Pattern Recognition Letters*, 32(14):1796–1804, 2011.

[13] C. Liu. Effective use of color information for large scale face verification. *Neurocomputing*, 101:43–51, 2013.

[14] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. In *IEEE Transactions on Image Processing*, pages 467–476, 2002.

[15] C. Liu, J. Yuen, and A. Torralba. SIFT flow: Dense correspondence across scenes and its applications. *IEEE Trans. Pattern Anal. Mach. Intell.*, 33(5):978–994, 2011.

[16] Z. Liu and C. Liu. Fusion of color, local spatial and global frequency information for face recognition. *Pattern Recognition*, 43(8):2882–2890, 2010.

[17] J. Lu, J. Hu, V. E. Liong, X. Zhou, A. Bottino, I. U. Islam, T. F. Vieira, X. Qin, X. Tan, S. Chen, Y. Keller, S. Mahpod, L. Zheng, K. Idrissi, C. Garcia, S. Duffner, A. Baskurt, M. Castrillon-Santana, and J. Lorenzo-Navarro. The FG 2015 Kinship Verification in the Wild Evaluation. In *FG 2015*, pages 1–7, May 2015.

[18] J. Lu, Y.-P. Tan, and G. Wang. Discriminative multimanifold analysis for face recognition from a single training sample per person. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(1):39–51, Jan 2013.

[19] J. Lu, X. Zhou, Y.-P. Tan, Y. Shang, and J. Zhou. Neighborhood repulsed metric learning for kinship verification. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 36(2):331–345, 2014.

[20] Y. Mu, W. Ding, and D. Tao. Local discriminative distance metrics ensemble learning. *Pattern Recognition*, 46(8):2337 – 2349, 2013.

[21] H. Nguyen and L. Bai. Cosine similarity metric learning for face verification. In *ACCV*, volume 6493, pages 709–720, 2011.

[22] F. Shahbaz Khan, J. van de Weijer, and M. Vanrell. Top-down color attention for object recognition. In *ICCV*, pages 979–986, Sept 2009.

[23] A. Sharma, A. Kumar, H. Daume, and D. Jacobs. Generalized multiview analysis: A discriminative latent space. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on*, pages 2160–2167, June 2012.

[24] P. Shih and C. Liu. Comparative assessment of content-based face image retrieval in different color spaces. *IJPRAI*, 19(7):873–893, 2005.

[25] K. Simonyan, O. M. Parkhi, A. Vedaldi, and A. Zisserman. Fisher Vector Faces in the Wild. In *BMVC*, 2013.

[26] K. E. A. van de Sande, T. Gevers, and C. G. M. Snoek. Evaluating color descriptors for object and scene recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32(9):1582–1596, 2010.

[27] K. Q. Weinberger and L. K. Saul. Distance metric learning for large margin nearest neighbor classification. *J. Mach. Learn. Res.*, 10:207–244, 2009.

[28] L. Wolf, T. Hassner, and Y. Taigman. Descriptor based methods in the wild. In *Real-Life Images workshop at the European Conference on Computer Vision (ECCV)*, October 2008.

[29] S. Xia, M. Shao, J. Luo, and Y. Fu. Understanding kin relationships in a photo. *Multimedia, IEEE Transactions on*, 14(4):1046–1056, Aug 2012.

[30] E. P. Xing, M. I. Jordan, S. Russell, and A. Y. Ng. Distance metric learning with application to clustering with side-information. In *NIPS*, pages 505–512, 2002.

[31] H. Yan, J. Lu, W. Deng, and X. Zhou. Discriminative multimetric learning for kinship verification. *IEEE Transactions on Information Forensics and Security*, 9(7):1169–1178, 2014.

[32] H. Yan, J. Lu, and X. Zhou. Prototype-based discriminative feature learning for kinship verification. *IEEE Transactions on Cybernetics*, 2015.

[33] Y. Yang, S. Liao, Z. Lei, D. Yi, and S. Li. Color models and weighted covariance estimation for person re-identification. In *ICPR 2014*, pages 1874–1879, Aug 2014.

[34] J. Zhang, Y. Barhomi, and T. Serre. A new biologically inspired color image descriptor. In *Computer Vision ECCV 2012*, pages 312–324. Springer Berlin Heidelberg, 2012.