On the Usefulness of Color for Kinship Verification from Face Images

Xiaoting Wu†, Elhocine Boutellaa†, Miguel Bordallo López†, Xiaoyi Feng*, Abdenour Hadid†‡
*Northwestern Polytechnical University, School of Electronics and Information, Xi’an, China
†Division Architecture des Systemes et Multimedia, Centre de Developpement des Technologies Avancees CDTA, Algeria
‡Center for Machine Vision and Signal Analysis, University of Oulu, Finland

Abstract—Automatic kinship verification from faces aims to determine whether two persons have a biological kin relation or not by comparing their facial attributes. This is a challenging research problem that has recently received lots of attention from the research community. However, most of the proposed methods have mainly focused on analyzing only the luminance (i.e. gray-scale) of the face images, hence discarding the chrominence (i.e. color) information which can be a useful additional cue for verifying kin relationships. This paper investigates for the first time the usefulness of color information in the verification of kinship relationships from facial images. For this purpose, we extract joint color-texture features to encode both the luminance and the chrominance information in the color images. The kinship verification performance using joint color-texture analysis is then compared against counterpart approaches using only gray-scale information. Extensive experiments using different color spaces and texture features are conducted on two benchmark databases. Our results indicate that classifying color images consistently shows superior performance in three different color spaces.

I. INTRODUCTION

Automatic facial kinship verification aims at checking whether two persons have a biological kin (i.e. family) relation or not by comparing their face appearance. This has a number of applications such as finding missing children, family album organization, creation of family trees or image annotation.

Kinship verification is a relatively new research topic in computer vision that is progressively gaining attention from the research community. It is a very challenging problem as it has to deal with images captured in the wild under adverse pose, expression, illumination and occlusion conditions. In addition, kinship verification shows wide intra-class and inter-class variations, where the pair of input images may be from persons of different sex and with a large age difference. Furthermore, persons from the same family may exhibit a large difference in face attributes while faces of persons with no kin relation may look similar.

To tackle the kinship verification problem, the research community has collected and made publicly available several face databases in both constrained and unconstrained conditions. These databases contain face images depicting pairs with different kin relations (and without kin relations). The most studied kin relations include Father-Son (F-S), Father-Daughter (F-D), Mother-Son (M-S) and Mother-Daughter (M-D).

So far, most of the proposed methods in the literature for automatic kinship verification have mainly focused on analyzing only gray-scale face images, hence discarding color information which can be a useful additional cue for verifying kin relationships. From a biological point of view, the chromaticity of the face is tied to genetically expressed features, such as eye color or skin tone. This hereditary features are many times present in kin-related persons in a similar manner. This paper investigates for the first time the usefulness of color information in the verification of kinship relationships from facial images. We aim at answering the question of whether color does help to improve kinship verification or not. For this purpose, we compare the performance of several baseline methods used in their traditional gray-scale variants against their counterparts utilizing color information. More specifically, we extract joint color-texture features to encode both the luminance and the chrominance information in the color images. The kinship verification performance using joint color-texture analysis is then compared against counterpart approaches using only gray-scale information. Extensive experiments using different color spaces and baseline methods are conducted on two benchmark kinship databases. The experimental results indicate that classifying color images consistently shows superior performance, in three different color spaces.

The main contributions of this work can be summarized as follows:

- We investigate for the first time the usefulness of color information for automatic kinship verification from face images;
- We propose the extraction of joint color-texture features from the luminance and chrominance channels of color images;
- We perform extensive experiments using different color spaces and baseline methods on two benchmark kinship databases;
- We show that consistent performance improvement can be obtained when using color images compared with their gray-scale counterparts, especially when using the HSV color space.

The rest of this paper is organized as follows. Section II reviews the related work on kinship verification and on the use of color in face analysis. Section III describes our methodology for studying the role of color in kinship verification. Section IV presents the experiments and discusses the results. Finally, concluding remarks and future directions are drawn in Section V.
II. RELATED WORK

Kinship verification has been traditionally studied in psychology and neuroscience aiming at understanding the resemblance of faces of the same family [1, 2, 3]. To the best of our knowledge, the first attempt to conduct automatic kinship verification based on face images is described in the work of Fang et al. [4]. The authors considered both local and global features and used forward selection methodology to find the most discriminative inherited facial features. A classification accuracy of 70.67% was reported on a small test set of image pairs using K-Nearest-Neighbors. Since this first work, several attempts have been made to solve the problem of kinship verification.

Among these contributions, the use of metric learning has resulted in very good performance as can be seen in the popular work of Lu et al. [5], which was later extended by Yan et al. [6] and Hu et al. [7]. More recently, methods based on deep learning have shown high performance. For example, Zhang et al. [8] applied CNN structure for kinship verification achieving excellent results, while Boutellaa et al. [9] showed that the combination of shallow and deep features can further improve the results. Deep learning approaches, however, require large amounts of training data and might not be applicable in every situation.

It appears that most of the proposed methods for kinship verification are mainly based on analyzing only gray-scale face images, hence discarding the color information. The question which arises then is whether this color cue is useful for verifying kin relationships or not. The use of color information in the recognition of nonface objects has been the subject of much debate [10]. However, there has only been a small amount of work which examines its contribution to face analysis. Color information has for example been applied to face analysis problems such as face detection, face recognition [11, 12] and spoofing detection [13]. These color based methods typically aim at learning discriminating color spaces where the classification becomes more affordable compared to the gray-scale space.

To the best of our knowledge, no previous work has yet explicitly explored the role of color information for kinship verification. Our present work aims at answering the question of whether color does help to improve kinship verification or not. To gain insight into the role of color information for kinship verification, we compare the performance of several baseline methods used in their traditional gray-scale variants against their counterparts utilizing color information. More specifically, we extract joint color-texture features to encode both the luminance and the chrominance information in the color images. The kinship verification performance using joint color-texture analysis is then compared against counterpart approaches using only gray-scale information. Our work is partially inspired by the work of Boulkenafet et al. [13] which proposed a joint texture-chrominance feature extraction scheme using color Local Binary Patterns (LBP) and applied it to better discriminate fake faces from genuine ones for biometric anti-spoofing.

III. PROPOSED METHODOLOGY

To study the usefulness of color in automatic kinship verification:

- We considered three baseline methods, successfully used in automatic kinship verification [14], namely Local Phase Quantization (LPQ) [15], Binarized Statistical Image Features (BSIF) [16] and the Neighborhood Repulsed Metric Learning (NRML) [5].
- We considered three color spaces namely RGB, HSV and YCrCb.
- As the baseline methods were originally designed for gray-scale images, we extended the methods to include color information by considering a joint texture-color analysis and combining the features extracted from different color channels (e.g. R, G and B).
- We compared the performance of these three baseline methods (LPQ, BSIF and NRML) in the three color spaces (RGB, HSV and YCrCb) versus their performance in the gray-scale space (Gray).
- We experimented with two different benchmark kinship face databases namely The Tri-subject Kinship Face Database (TSKinFace) [17] and The Kinship Face in the Wild dataset (KinFaceW-II) [5].

A. Neighborhood repulsed metric learning

The first baseline method employs the popular local binary patterns (LBP) features with metric learning. The neighborhood repulsed metric learning (NRML) is an approach for kinship verification proposed by Lu et al. [5]. NRML learns a distance where the face pairs with a kin relation are pulled close to each other and those without a kin relation are pushed away. This is achieved by seeking a metric d:

$$d(x_i, y_i) = \sqrt{(x_i - y_i)^T A(x_i - y_i)}$$  (1)

where $(x_i, y_i)$ is the pair of parent-child features (LBP features in our case); $i = j$ means a positive pair and $d$ should be as small as possible, and $(x_i \neq y_i)$ means a negative pair and $d$ should be as large as possible. $A$ is a symmetric and positive semidefinite square matrix obtained by solving the following optimization problem:

$$\max_A J(A) = J_1(A) + J_2(A) + J_3(A)$$

$$= \frac{1}{NK} \sum_{i=1}^{N} \sum_{t=1}^{k} d^2(x_i, y_{t_i}) + \frac{1}{NK} \sum_{i=1}^{N} \sum_{t=2}^{k} d^2(x_{it_2}, y_{t_1}) - \frac{1}{N} \sum_{i=1}^{N} d^2(x_i, y_i)$$  (2)

$J_1$ ensures that $t_1$th $k$-nearest neighbors of child $y_i$ are separated as far as possible with parent $x_i$ in the learned distance metric space; and $J_2$ ensures that $t_2$th $k$-nearest neighbors of parent $x_i$ are separated as far as possible with child $y_i$ in the learned distance metric space; whereas $J_3$ guarantees
that the positive pairs \((x_i, y_i)\) are pulled as close as possible in the learned distance metric space.

**B. Local phase quantization**

Local Phase Quantization (LPQ) [15] is a blur robust image descriptor. The LPQ descriptor is based on the insensitivity of the low-frequency phase components to centrally symmetric blur. Therefore, LPQ employs the phase information of short-term Fourier transform, which is locally computed on a window around each pixel of the image. LPQ is computed using four complex low frequencies: \(u_0 = (\alpha, 0), u_1 = (\alpha, \alpha), u_2 = (0, \alpha), u_3 = (-\alpha, -\alpha)\) where \(\alpha\) is a small scalar frequency (\(\alpha << 1\)) ensuring the blur is centrally symmetric. Each pixel \(x\) of the image is characterized by a vector \(F_x\) of complex frequencies:

\[
F_x = \{\text{Re}\{F(x, u_0), F(x, u_1), F(x, u_2), F(x, u_3)\},
\text{Im}\{F(x, u_0), F(x, u_1), F(x, u_2), F(x, u_3)\}\},
\]

(3)

\(\text{Re}\{\cdot\}\) and \(\text{Im}\{\cdot\}\) denotes the real and the imaginary part of a complex number.

To maximize the information preservation by the quantization process, the coefficients should be statistically independent. Therefore, a decorrelation step based on a whitening transform is applied in LPQ before the quantization process. Subsequently, the vector of whitened coefficients is quantized via a simple shareholding scheme:

\[
q_i = \begin{cases} 
0 & \text{if } f'_i < 0 \\
1 & \text{otherwise}
\end{cases}
\]

(4)

where \(f'_i\) is the \(i\)th whitened coefficient. Finally, the resulting binary quantized coefficients are represented as integer value in \([0-255]\) as follows:

\[
LPQ(x) = \sum_{i=0}^{8} q_i 2^{i-1}
\]

(5)

**C. Binarized statistical image feature**

Binarized statistical image feature (BSIF) [16] is a binary texture descriptor belonging the same family of LBP and LPQ. However, instead of using hand-crafted filters, BSIF utilizes a small set of natural images to automatically learn a fixed set of filters. This set of filters is learned based on the statistics of small set of natural images to automatically learn a fixed set of filters. Where, instead of using hand-crafted filters, BSIF utilizes a texture descriptor belonging the same family of LBP and LPQ.

In order to learn a useful set of \(n\) filters \(W_r\), the statistical independence of the responses \(s_i\) should be maximized. To achieve that, a sufficient numbers of patches are randomly sampled from the training images. The patches are then normalized to zero mean and principal component analysis (PCA) is applied to minimize their dimension to \(n\). Finally, the filters are derived by applying the independent component analysis (ICA) algorithm. Once the filter matrix \(W\) is computed, it can be directly utilized for calculating BSIF features from any image.

**D. Classification scheme**

Our kinship verification scheme is depicted in Fig. 1. In a nutshell, a target pair of face images is given as an input. These two images are first converted into different color spaces (e.g. HSV). Then, the features (e.g. BSIF) are separately extracted from each channel of the considered color space (e.g. \(H\), \(S\) and \(V\)). The features are then concatenated to form one enhanced feature vector. Finally, we apply cosine similarity between the feature vectors of the pair of the two face images. Cosine similarity between two vectors \(X\) and \(Y\) is defined as following:

\[
sim(X, Y) = \frac{x \cdot y}{\|x\| \cdot \|y\|}
\]

(8)

where \(\| \cdot \|\) is the Euclidean norm, and \(X, Y\) refer to parent and child feature vectors. A small value of \(\sim(X, Y)\) means a high probability that the parent \(X\) and the child \(Y\) have a kin relation.

**IV. EXPERIMENTAL RESULTS AND ANALYSIS**

To gain insight into whether color does help to improve kinship verification or not, we performed extensive experiments on two different benchmark kinship face databases namely The Tri-subject Kinship Face Database (TSKinFace) [17] and The Kinship Face in the Wild dataset (KinFaceW-II) [5]. We compared the performance of the three baseline methods (LPQ [15], BSIF [16] and NRML [5]) in the three color spaces (RGB, HSV and YCbCr) versus their performance in the gray-scale space (Gray). In this section, we describe the experimental data and setup, and then discuss the obtained results.

**A. Experimental Data**

1) TSKinFace: The Tri-subject Kinship Face Database, (TSkinFace)[17] is a recently published dataset that contains two groups of family-based kinship relations: Father-Mother-Daughter (FM-D), Father-Mother-Son (FM-S). FM-D and FM-S relations have 513 and 502 tri-subject groups respectively, which are collected from wild circumstances from public figures from the Internet. The face images are cropped into a resolution of 64×64 pixels. For comparative analysis, we reorganized the dataset by dividing the Father-Mother-Daughter group into Father-Daughter and Mother-Daughter relationships, and the Father-Mother-Son group into Father-Son and Mother-Son relationships. Thus, we obtained four kin relations. Fig. 2 shows samples of positive pairs and negative pairs from the Tri-subject Kinship Face Database (TSkinFace).
2) Kinship Face in the Wild dataset: The Kinship Face in the Wild dataset (KinFaceW)[5] consists of two different subsets: KinFaceW-I and KinFaceW-II. Both subsets include four kinds of parent-child relationships: Father-Daughter, Father-Son, Mother-Daughter, Mother-Son. The dataset provides a description of the image collection process, depicting images under uncontrolled environments with no restrictions in terms of pose, lighting, background, expression, age, ethnicity and partial occlusion. Images from KinFaceW-II are acquired from the same photograph. Images from KinFaceW-I are acquired from different images, although many times they also show similar capturing conditions [18]. These images are collected from the internet including public figures and their parents or children. Examples from the dataset KinFaceW-II can be seen in Fig.3.

The face images are cropped into a resolution of is 64×64 pixels and aligned using the position of the eyes. KinFaceW-I includes respectively 156, 134, 116, and 127 pairs of each category. KinFaceW-II includes 250 pairs for each relationship type.

B. Experimental Setup

In our experiments, we followed the defined protocol of the KinFaceW dataset and adopted a similar protocol with the recently introduced TSKinFace dataset. In this context, all pairs with kin relation are used as positive pairs. Negative pairs are generated randomly by parents and children which do not have a kin relation. Moreover, parents and children in negative part appear only once. Thus, the experimental protocol for training and testing contains always 50% positive pairs and 50% negative pairs.

To evaluate the performance of the different baseline methods, we applied five-fold validation in both datasets, TSKinFace and KinFaceW-II. To ensure a fair comparison between methods and databases, the results are reported with the accuracy measured on the Receiver Operating Characteristic (ROC) curves. To validate the statistical significance of our results, we have performed statistical tests (ANOVA Single Factor) across the results of the five different folds and four different relationships, for a total of 20 data points.
C. Experimental Results

Tables I - Table IV show the obtained results on the TSKinFace database, comparing the performance of the three baseline methods (LPQ [15], BSIF [16] and NRML [5]) in the three color spaces ($RGB$, $HSV$ and $YC_bC_r$) versus the performance of these baseline methods in the gray-scale space ($Gray$). From these results, we can clearly see that the use of color consistently improves the kinship performance compared with the corresponding gray-scale approaches. When comparing the different baseline methods, we can see that BSIF yields in the best overall performance, outperforming the NRML based metric learning method. When comparing the color spaces, although the improvement can be seen in all of them, $HSV$ seems to give the best overall performance. This superiority can be noted for all individual folds. A statistical ANOVA test performed across all five folds and all four relationships showed that the better performance of $HSV$ is statistically significant. Hence, we can conclude that color does provide some discriminative information that can help in boosting the kinship verification performance.

Regarding the performance in different kinship groups, the results show that relationships composed by members of different sex (F-D and M-S) obtain lower overall accuracy when compared with same gender pairs (F-S and M-D). This is in line with previous observations [9]. Kinship verification is indeed a complex research problem with many factors affecting the performance of the verification methods, such as different gender or large age gaps.

In addition to the results given in the tables (Tables I - Table IV), Fig. 4 show the results in terms of ROC curves on the TSKinFace dataset clearly demonstrating that color improves the performance in all scenarios.

To validate our findings, we have performed experiments on a secondary dataset, KinFaceW-II. One of the most used kinship verification datasets, KinFaceW-II is comprised of image pairs captured in similar conditions [18], but has been used as a benchmark in many previous works and its use facilitates the comparison against the state-of-the-art [19].

Fig. 5 shows the obtained results, validating our findings and pointing out the usefulness of color information for Kinship verification. The ROC curves are shown for the best performing feature (BSIF) for grayscale and three different color spaces.

V. CONCLUSION

Chromaticity-related features such as eye-color or skin tone play a role in kinship verification. In this paper, we presented the first study in the literature exploring the role of color information in kinship verification. We performed extensive experiments on two different benchmark kinship face databases namely The Tri-subject Kinship Face Database (TSKinFace) and The Kinship Face in the Wild dataset (KinFaceW-II). We compared the performance of the three baseline methods (LPQ, BSIF and NRML) in the three color spaces ($RGB$, $HSV$ and $YC_bC_r$) against the performance of these baseline methods in the gray-scale space ($Gray$). The obtained results indicated that color seems to improve the performance of kinship verification in all configurations.

The choice of the baseline methods, the color spaces and the databases is mainly motivated by the popularity and the availability of the chosen material. This is important for reproducible research and fair comparison with state of the art. The work is by no mean complete as our results and findings should be considered as preliminary and hence should be further validated on other larger databases and other baseline methods. It is worth noting that in this experimental setup, the $HSV$ color space seems to outperform other color spaces. This intriguing result might be due to the uncorrelated nature of its components and should be further investigated.

The databases selected for the experimental validation have negative classes that are randomly generated. The inclusion of pairs in the negative class presenting similar characteristics in terms of age, ethnicity or appearance (look-alike) could have an impact in the verification rates, and should be further studied.

However, tackling the problem of kinship verification is not the main focus of the paper. The focus is on comparing the performance of color-based methods versus gray-scale counterparts. The results could be further improved by exploring the complementarity of the features or the color spaces.

Among other open questions are: does ethnicity affect the importance of color for kinship verification? How the different spectral channels of face images should be combined to take
full advantages of the color information? How the environment (illumination conditions, camera calibration, etc.) affects the choice of optimal color space? These suggest that color cue has not yet shown its full potential and need further investigation.

REFERENCES