Journal of Electronic Imaging

JElectronicImaging.org

On the usefulness of hyperspectral imaging for face recognition

Simone Bianco

SPIE. IS&T imaging.org

Simone Bianco, "On the usefulness of hyperspectral imaging for face recognition," *J. Electron. Imaging* **25**(6), 063020 (2016), doi: 10.1117/1.JEI.25.6.063020.

On the usefulness of hyperspectral imaging for face recognition

Simone Bianco*
University of Milano-Bicocca, DISCo, Dipartimento di Informatica, Sistemistica e Comunicazione, Viale Sarca 336, Milan 20126, Italy

Abstract. Hyperspectral cameras provide additional information in terms of multiple sampling of the visible spectrum, holding information that could be potentially useful for biometric applications. This paper investigates whether the performance of hyperspectral face recognition algorithms can be improved by considering single and multiple one-dimensional (1-D) projections of the whole spectral data along the spectral dimension. Three different projections are investigated and found by optimization: single-spectral band selection, nonnegative spectral band combination, and unbounded spectral band combination. Since 1-D projections can be performed directly on the imaging device with color filters, projections are also restricted to be physically plausible. The experiments are performed on a standard hyperspectral dataset and the obtained results outperform eight existing hyperspectral face recognition algorithms. *© 2016 SPIE and IS&T* [DOI: 10.1117/1.JEI.25.6.063020]

Keywords: hyperspectral imaging; biometrics; face recognition.

Paper 16497 received Jun. 17, 2016; accepted for publication Nov. 21, 2016; published online Dec. 15, 2016.

Since intraperson differences are often larger than interpersonal ones in the presence of variations in viewing point and illumination conditions, face recognition is still a challenging problem.

Most of the current research is based on features extracted from grayscale or RGB images, which are usually acquired in the visible spectrum.[1,2]

With the aim of increasing the dimensions in face images, many researchers have considered the use of hyperspectral imaging.[3–7] Hyperspectral imaging can increase facial discrimination by capturing more biometric measurements such as the spectral response of faces. A hyperspectral image is a data cube with two spatial dimensions and one spectral dimension. It is captured by a hyperspectral camera which operates in multiple narrow bands and densely samples the radiance information in both space and wavelength, producing a radiance spectra at every pixel.

In addition to face appearance, spectral measurements in multiple wavelengths can also measure subsurface tissue features[4] which may be significantly different for each person.

Although the high dimensionality of hyperspectral data is a desirable feature for separating the different identities, at the same time it poses new challenges such as interband misalignments and low signal-to-noise ratio (SNR) in certain spectral bands.

Due to the high dimensionality of hyperspectral data, discriminative feature extraction for face recognition is more challenging than two-dimensional (2-D) images. The different approaches for dimensionality reduction and feature extraction range from the subsampling of the hyperspectral data[4,5,7] to the more promising approaches which use whole-band features.[3,8]

Starting from the best hyperspectral method in the state-of-the-art,[8] this paper investigates whether the use of linear projections along the spectral dimension can improve face recognition performance with respect to the use of the full hyperspectral data. Although applied in a different way, the use of linear projections has already been shown to improve classification accuracy in the 2-D domain when applied to PCA,[9,10] eigenfaces,[11] and fisherfaces.[11,12] The experiments are performed on the PolyU hyperspectral[3,13] standard hyperspectral face database. The results are compared with eight existing hyperspectral face recognition algorithms.

Preliminary findings reported in this paper appeared in Ref. 14, where no constraints were given on the investigated projections. This paper extends[14] in several ways

- Since one-dimensional (1-D) projections can be performed directly in the imaging device with the use of color filters, additional experiments with physically plausible filters (i.e., projections) are added.
- Two different face alignments are considered in the experiments: a manual single-band alignment and an automatic full-band one.
- Results are compared with a larger number of methods in the state of the art.

1 Baseline Method

The proposed method builds on the method of Uzair et al.,[8] which has three main steps that are respectively devoted to the normalization of the variations in illumination, to the extraction of the features, and to the classification.

In the first step, the individual bands are filtered with a circular (8,1) neighborhood illumination variations step (LBP)[15] filter to normalize for the illumination variations.

The second step is the feature extraction step which is based on a three-dimensional discrete cosine transform

*Address all correspondence to: Simone Bianco, E-mail: simone.bianco@disco.unimib.it

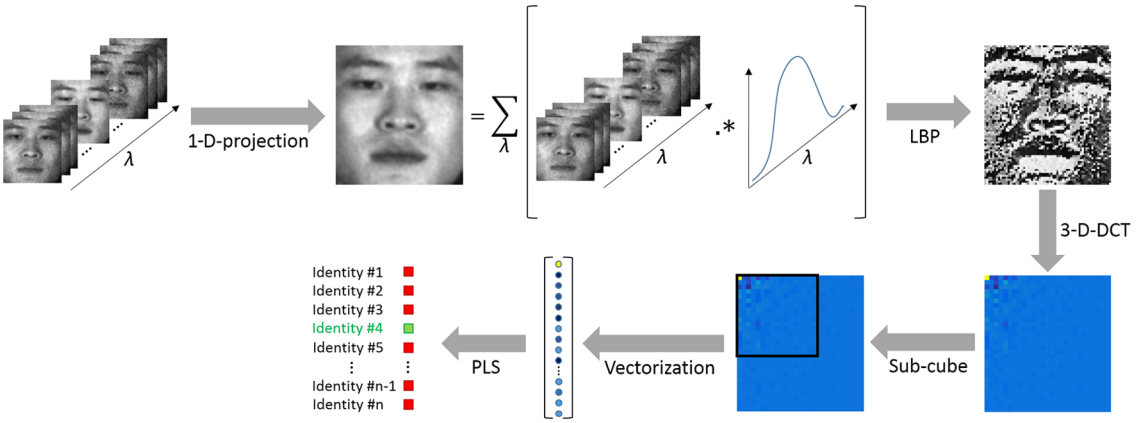

Fig. 1 Flowchart of the entire processing chain of the proposed method.

(3-D-DCT). The discrete cosine transform (DCT)[16] decomposes a discrete signal into a linear combination of independent cosine basis functions. DCT tends to generate a representation in which most of the signal information is encoded in the low-frequency coefficients. By selecting as features only the low-frequency coefficients, a compact representation can be obtained. The 3-D-DCT of a hyperspectral cube $H(x,y,\lambda)$ with size $N_1 \times N_2 \times N_3$ is given by

$$F(u,v,w) = \Omega_1(u)\Omega_2(v)\Omega_3(w) \sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} \sum_{\lambda=0}^{N_3-1} H(x,y,\lambda) \times \cos\frac{\pi(2x+1)u}{2N_1} \cos\frac{\pi(2y+1)v}{2N_2} \cos\frac{\pi(2\lambda+1)w}{2N_3}, \tag{1}$$

with $u = \{0,\ldots,N_1-1\}$, $v = \{0,\ldots,N_2-1\}$, $w = \{0,\ldots,N_3-1\}$, and $\Omega_i(\cdot)$ is defined $\sqrt{1/N_i}$ if its argument is zero, and $\sqrt{2/N_i}$ otherwise.

The low frequency coefficients near the origin of $F(u,v,w)$ represent most of the energy of the hyperspectral cube, therefore, the high-frequency coefficients can be discarded. The feature vector is constructed by sampling a frequency subcube $\Gamma(u,v,w)$ of dimensions $(\alpha \times \beta \times \gamma)$ keeping only the low-frequency elements around the origin of $F(u,v,w)$. The subcube $\Gamma(u,v,w)$ is then vectorized and normalized to unit magnitude to obtain the final feature vector $f \in \mathbb{R}^d$, where $d = \alpha\beta\gamma$, which is then used for classification.

The third and last steps consist of the use of the partial least squares (PLS) regression[17] for the classification. PLS models the relations between sets of observed variables by means of latent variables. In its general form, PLS creates orthogonal score vectors by maximizing the covariance among different variable sets. The only parameter to be set in PLS is the number of latent variables to use.

2 Linear Spectral Projection

The proposed method builds on top of the method of Uzair et al.[8] to understand whether the full hyperspectral information is actually needed to improve face recognition accuracy or if a projection of it suffices. The projection is applied directly to the hyperspectral cube $H(x,y,\lambda)$ (i.e., the radiance data), before any step of the method in Ref. 8. The projection is performed along the spectral dimension and depends on the set of weights $W(\lambda_i) = w_i$, $i = 1,\ldots,N_3$

$$P(x,y) = \sum_{i=1,\ldots,N_3} H(x,y,\lambda_i) W(\lambda_i). \tag{2}$$

The projection $P(x,y)$ is thus a 2-D image, forcing $\gamma = 1$ for the subcube size. The entire processing chain of the proposed method is reported in the flowchart in Fig. 1.

In this work, three different projections are considered. The first one is

$$W_1(\lambda) = \delta_{\lambda_0}(\lambda) = \begin{cases} 1 & \text{if } \lambda = \lambda_0 \\ 0 & \text{otherwise} \end{cases}, \tag{3}$$

and can be seen as a band selection operator or a pass-band optical filter.

The second projection is

$$W_2(\lambda_i) = w_i, i = 1,\ldots,L \quad \text{s.t. } \forall\, w_i : w_i \in \mathbb{R}, 0 \le w_i \le 1, \tag{4}$$

which can be seen as a nonnegative linear combination of the different hyperspectral bands. This is an operation analog to what optical filters do in traditional imaging and could be done using a monochrome digital camera coupled with a custom designed filter.

The third projection is an unbounded linear combination of the hyperspectral bands and can be defined as in Eq. (4) removing the lower and upper bounds on the filter coefficients w_i, i.e.,

$$W_3(\lambda_i) = w_i, i = 1,\ldots,L \quad \text{s.t. } \forall\, w_i : w_i \in \mathbb{R}. \tag{5}$$

This is a generalization of the second one and is the only one that cannot be realized through an optical filter since it could have negative coefficients as well as $|w_i| > 1$. The optimal $W_1(\lambda)$ projection is obtained by exhaustive search, while for both $W_2(\lambda)$ and $W_3(\lambda)$ a particle swarm optimization (PSO)[18,19] is used. PSO is a population-based stochastic optimization technique. A population of individuals is initialized as random guesses to the problem solutions and a communication structure is also defined, assigning neighbors for each individual to interact with. These individuals are candidate solutions. The particles iteratively evaluate the fitness of the candidate solutions and remember the location where

Fig. 2 A hyperspectral face cube from the PolyU-HSFD dataset.

they had their best success. The best solution of each individual is called the particle best or the local best. Each particle makes this information available to its neighbors. Movements through the search space are guided by these successes. The swarm is typically modeled by particles in multidimensional space that have a position and a velocity. These particles move into the search space and have two essential reasoning capabilities: the memory of their own best position and the knowledge of the global best position (or the best position of their neighbors). Members of a swarm communicate good positions to each other and adjust their own position and velocity based on these good positions. The reason for choosing the PSO algorithm is that PSO makes no assumptions about the problem being optimized and can search very large spaces of candidate solutions. Furthermore, PSO is a derivative-free method and its population-based nature permits an easy parallelization to speed up optimization time.

Experiments are conducted using a 10-fold cross validation. For each split, let us call $\mathbf{t} = \{t_1, \ldots, t_n\}$ the random indices of subjects used for training and $\mathbf{T} = \{T_1, \ldots, T_m\}$ the random indices of subjects used for testing. For each subject j in $\mathbf{t}$ let $(g_{j,1}, \ldots, g_{j,G})$ be the corresponding gallery cubes and $(p_{j,1}, \ldots, p_{j,P})$ the corresponding probes ones. The fitness function corresponds to the average recognition rate evaluated on the subjects in $\mathbf{t}$ using a leave-one-out cross-validation procedure: in turn, each single gallery cube $H_{t_j}^{g_{j,i}}$, $i = 1, \ldots, G$, with $t_j \in \mathbf{t}$ is excluded from training and its identity is predicted. Once the optimization of weights is ended, the subjects in $\mathbf{T}$ are added to the gallery and the identities of their probe cubes are predicted. In this way, performances are always evaluated on identities not used for weights optimization.

3 Experiments

3.1 *Dataset*

The hyperspectral face database used is the Hong Kong Polytechnic University hyperspectral face database (PolyU-HSFD).[3,13] It consists of hyperspectral image cubes acquired using a CRI's VariSpec Liquid Crystal Tuneable Filter. Each cube contains 33 bands acquired in the 400- to 720-nm spectral range in 10-nm steps. The database has been collected over a long period of time and shows significant appearance variations of the subjects (e.g., changes of hair style and skin conditions). SNR in bands near the blue wavelength is very low, and the database contains interband misalignments due to subject movements during the acquisition at the different wavelengths.

The database contains a total of 48 different subjects, of which 13 are females and 35 are males. For each of the first 25 subjects from four to seven cubes are available, while the remaining 23 subjects have only one cube each. Following the experimental protocol of Refs. 3 and 7, only the first 25 subjects are used in the experiments. For each subject, two cubes are randomly selected for the gallery and the remaining cubes are used as probes. The random selection is repeated 10 times and the results are reported in terms of average recognition rate over the 10 random splits. For face alignment, as in Ref. 3, the eye, nose tip, and mouth corners' coordinates were manually located, and a subregion containing the face was cropped from each band, normalized, and scaled to one quarter size.

An example of the hyperspectral face cubes used is reported in Fig. 2, while examples of appearance variations are reported in Fig. 3.

3.2 *Compared Hyperspectral Face Recognition Algorithms*

The eight existing hyperspectral face recognition algorithms used for comparisons include spectral signature matching,[4] spectral angle measurement,[6] spectral eigenface,[5] 2-D PCA,[3] 3-D Gabor wavelets,[7] 2-D and 3-D-DCT with PLS regression,[8] and spatiospectral information fusion with PLS regression.[20] The parameters of these algorithms are set as follows. For the spectral signature matching algorithm,[4] five adjacent square regions of size 17×17 pixels arranged in a cross pattern are used to represent hair, forehead, and cheeks. For the lips, square regions of size 9×9 pixels are used. For spectral eigenface,[5] 99% of the energy is preserved by retaining 48 PCA basis vectors. For 2-D PCA,[3] 99% of the energy is preserved by retaining 27 PCA basis vectors. For the 3-D Gabor method, 52 Gabor wavelets are used for feature generation as recommended by Ref. 7. For the 2-D and 3-D-DCT[8] methods, the parameters are taken as suggested by the authors: $\alpha = \beta = \gamma = 10$ for the subcube size to extract the features and 45 PLS basis.

3.3 *Results*

The results of the hyperspectral face recognition algorithms compared are reported in terms of average recognition rate in

Fig. 3 Examples of appearance variations. The same hyperspectral band corresponding to $\lambda_{15} = 540$ nm is selected for all subjects.

Table 1. The results of spectral signature matching,[4] spectral angle measurement,[6] spectral eigenface,[5] 2-D PCA,[3] 3-D Gabor wavelets,[7] and 2-D and 3-D-DCT with PLS regression[8] are all taken from Ref. 8 with the only exception being the 3-D-DCT method for which the results using our implementation are also reported.

It is possible to notice that the proposed method outperforms the best algorithm in the state of the art by 2% (for the single-band selection) up to 4% (for the unbounded band combination). The best projections found for $W_1(\lambda)$, $W_2(\lambda)$, and $W_3(\lambda)$ are shown in Fig. 4. Interestingly, the band selected by $W_1(\lambda)$ and the bands receiving higher weights by $W_2(\lambda)$ and $W_3(\lambda)$ are localized at the oxyhemoglobin peak absorption valley.[3,21]

Table 1 Average recognition rates and standard deviations (%) for 10-fold experiments on the PolyU-HSFD database.

Algorithm	Average recognition rate ± std
Spectral signature[4]	24.63 ± 3.87
Spectral angle[6]	25.49 ± 4.36
Spectral eigenface[5]	70.30 ± 3.61
D PCA[3]	71.11 ± 3.16
D Gabor wavelets[7]	90.19 ± 2.09
D-DCT + PLS[8]	91.43 ± 2.10
D-DCT + PLS[8]	93.00 ± 2.27
D-DCT + PLS (author's implementation)	93.32 ± 3.13
Spatiospectral information fusion + PLS[20]	95.20 ± 1.60
Proposed [$W_1(\lambda)$, single-band selection]	97.20 ± 1.66
Proposed [$W_2(\lambda)$, nonnegative band combination]	98.34 ± 1.83
Proposed [$W_3(\lambda)$, unbounded band combination]	99.11 ± 1.21

As already stated in Sec. 2, the projections $W_1(\lambda)$ and $W_2(\lambda)$ could both be realized through an optical filter since they do not have negative coefficients. The projection $W_3(\lambda)$, instead can not be realized through a single-optical filter, but by exploiting the linearity of Eq. (2) it could be realized by subtracting two different optical filters $W_3^+(\lambda_i)$ and $W_3^-(\lambda_i)$

$$P(x,y) = \sum_{i=1,\ldots,N_3} H(x,y,\lambda_i) W_3^+(\lambda_i) - \sum_{i=1,\ldots,N_3} H(x,y,\lambda_i) W_3^-(\lambda_i), \tag{6}$$

where

$$W_3^+(\lambda_i) = \begin{cases} W_3(\lambda_i) & \text{if } w_i > 0 \\ 0 & \text{otherwise} \end{cases}, \tag{7}$$

and

$$W_3^-(\lambda_i) = \begin{cases} -W_3(\lambda_i) & \text{if } w_i < 0 \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

Some examples of the projected output given by applying Eq. (2) with the optimal $W_1(\lambda)$, $W_2(\lambda)$, and $W_3(\lambda)$ projections found, are shown in Fig. 5.

From the images reported, it is possible to see that using the $W_1(\lambda)$ projection results in sharper images, due to the fact that only one spectral band is used. On the contrary, since $W_2(\lambda)$ and $W_3(\lambda)$ use the whole spectra, they make interband misalignments evident resulting in more blurred images.

In Fig. 6, some examples of errors across the 10-fold experiments when using the $W_3(\lambda)$ projection are reported. The two gallery images are reported for each example together with the probe image and the gallery images of the incorrectly assigned identity.

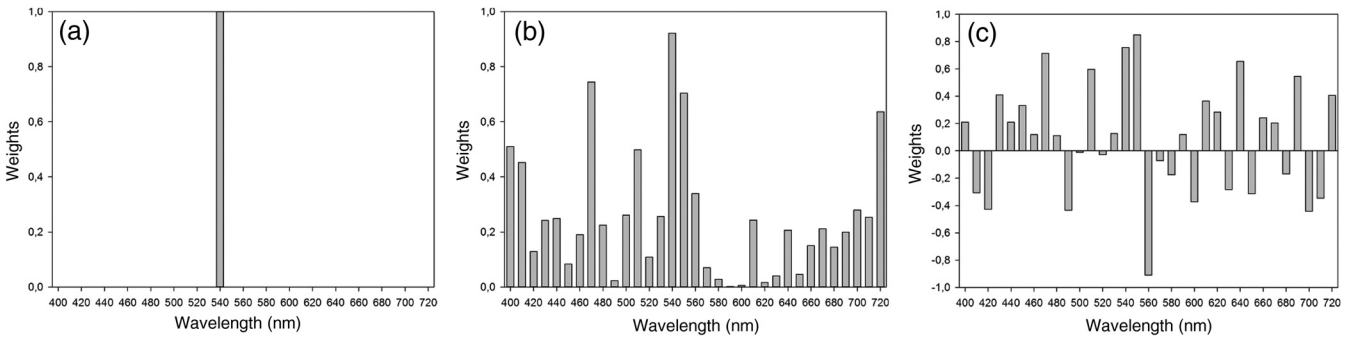

Fig. 4 Best projections found: (a) $W_1(\lambda)$, (b) $W_2(\lambda)$, and (c) $W_3(\lambda)$.

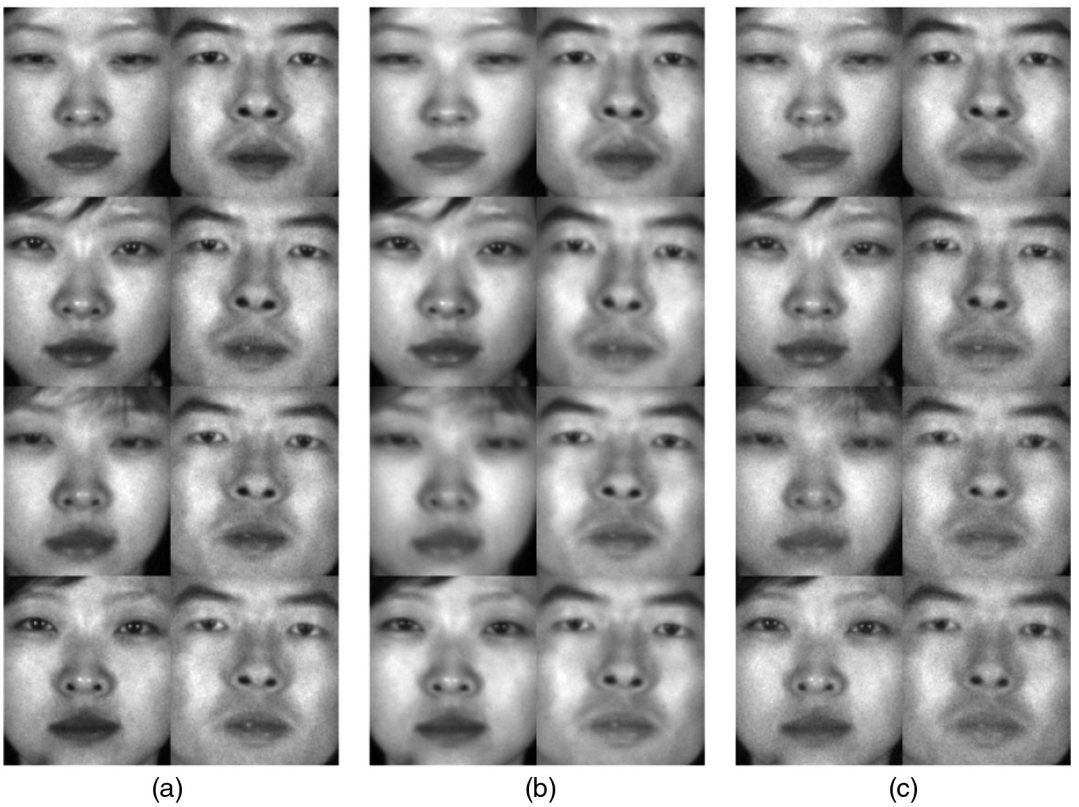

Fig. 5 Examples of the projections obtained by applying the optimal projections found. (a) $W_1(\lambda)$, single-band selection; (b) $W_2(\lambda)$, nonnegative linear combination; (c) $W_3(\lambda)$ and unbounded linear combination.

The sensitivity of the proposed method is analyzed in Fig. 7 by plotting equal recognition rate curves as a function of the number of PLS basis and subcube size ($\alpha = \beta$ and $\gamma = 1$, due to the effect of the projection).

3.4 *Physically Plausible Projections*

All the experiments in the previous section were performed without considering the physical plausibility of the projections. Since the projections can be performed with optical filters, in this section we perform additional experiments considering real filters. To this end, a set of 500 color filters have been downloaded from the Rosco website.[22] In addition to these filters, the spectral sensitivity functions of 28 different cameras, including professional DSLRs, point-and-shoot, industrial, and mobile cameras have been obtained.[23] The spectral sensitivity functions of the cameras and the filters are reported in Fig. 8. For visualization purposes, a k-means with $k = 25$ has been run on the 500 color filters and only the centroids obtained are plotted. Three different experiments have been carried out, obtained by coupling the color filters with the camera sensitivities:

I. Selection of the best color filter without considering the camera sensitivities. This is equivalent to placing the color filter on a gray-level camera.

II. Selection of the best color filter coupled with the camera sensitivity of one-color channel. This is equivalent to placing the color filter on a color camera and considering just the recording of a single-color channel.

III. Selection of the best color filter coupled with the full camera sensitivities. This is equivalent to placing the color filter on a color camera and considering the recording of all the color channels.

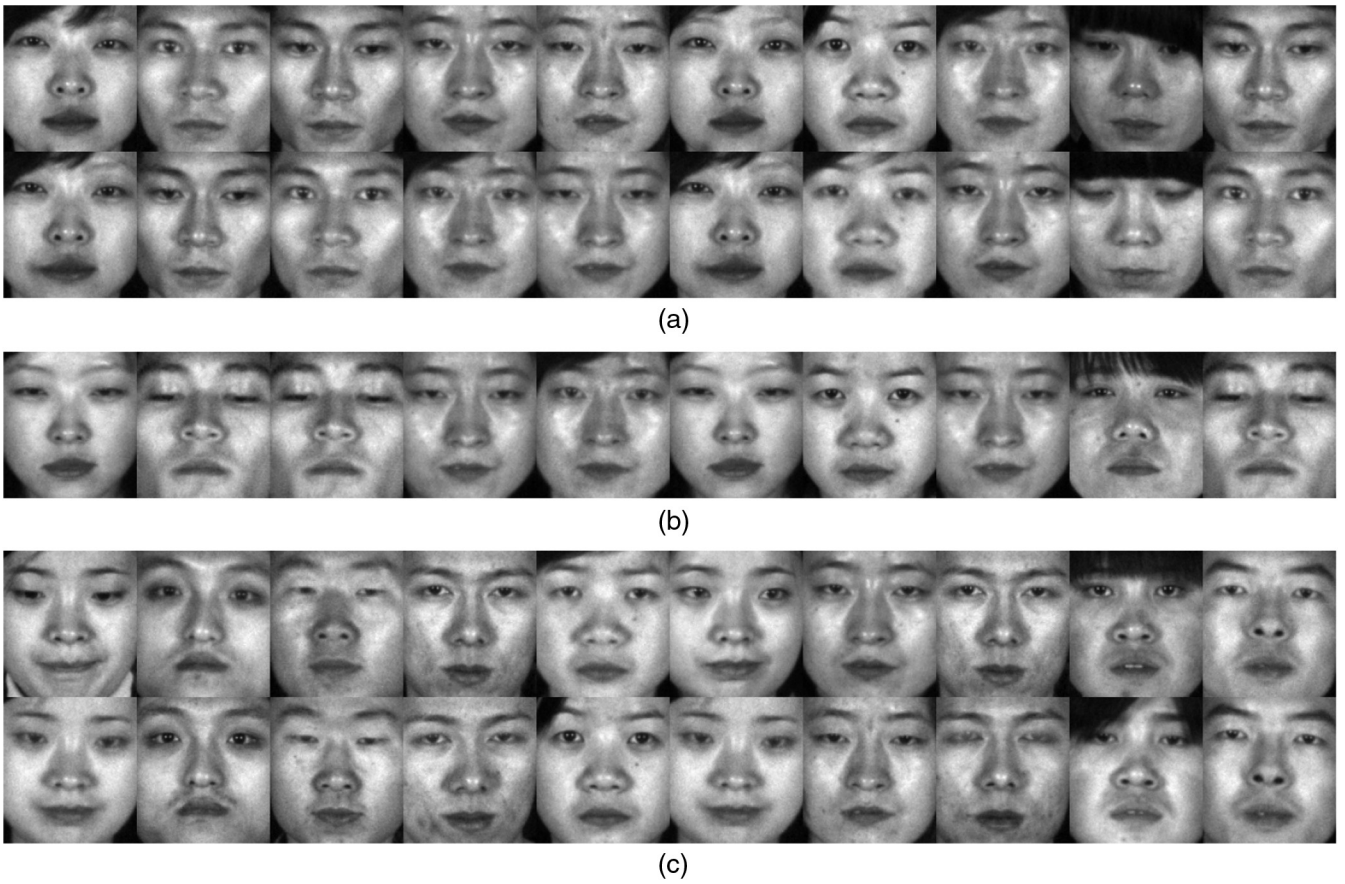

Fig. 6 Examples of errors for the 10-fold experiment using the $W_3(\lambda)$ projection: (a) gallery cubes, (b) probes, (c) gallery cubes for the predicted identity.

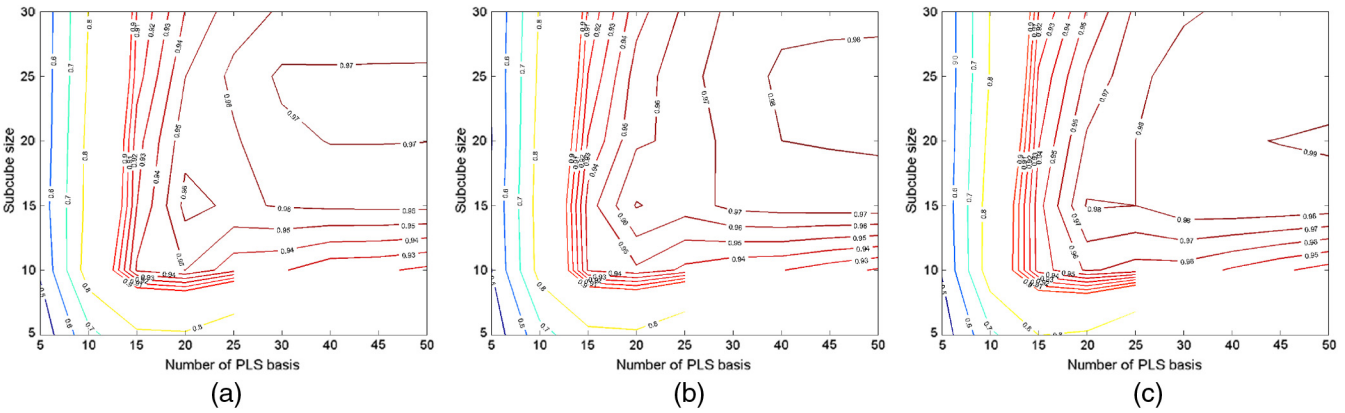

Fig. 7 Equal recognition rate curves as a function of subcube size (y-axis) and number of PLS basis (x-axis): (a) single-band selection, (b) nonnegative linear combination, and (c) unbounded linear combination.

Furthermore, in addition to the single-band manual alignment performed in Sec. 3.1, we perform a fully automatic alignment using congealing.[24] This permits to align all the spectral bands, removing the interband misalignments visible in Fig. 5. A comparison of the average face across all bands and all the dataset using the manual alignment and congealing is reported in Fig. 9.

For all the experiments, the number of PLS basis and subcube sizes are kept fixed and equal to the optimal ones found for $W_3(\lambda)$ in the previous section. For all the experiments, the best filter or combination of filter and camera sensitivities are obtained by exhaustive search.

The results of the three experiments with both alignments are reported in Table 2. From the results, it is possible to see that using the same face alignment used in the previous section, the recognition accuracy using a physically plausible 1-D projection (i.e., experiments I and II) is worse than that obtained by the $W_2(\lambda)$. Using a 3-D projection (i.e.,

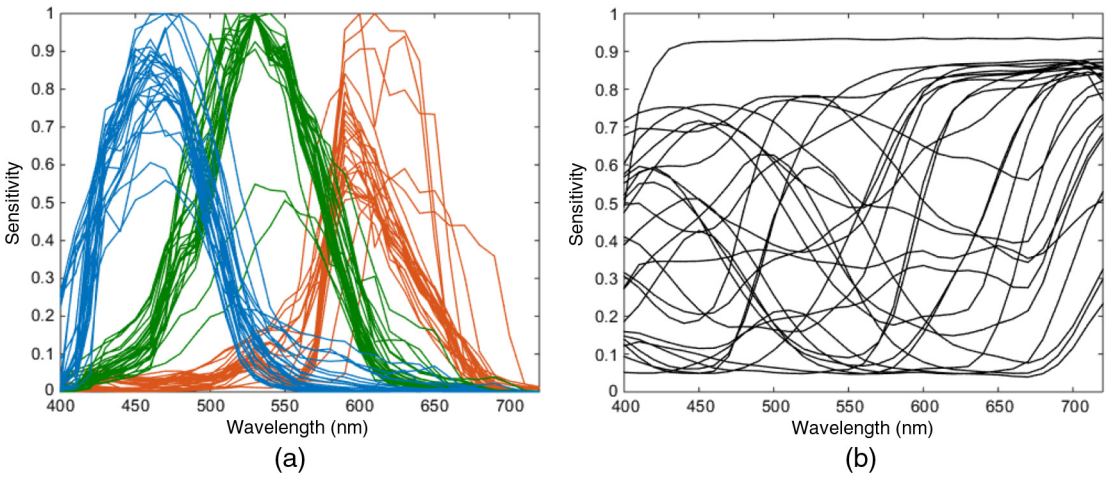

Fig. 8 Spectral sensitivity functions of (a) the 28 cameras and (b) of the 25 centroids obtained by *k*-means from the 500 filters.

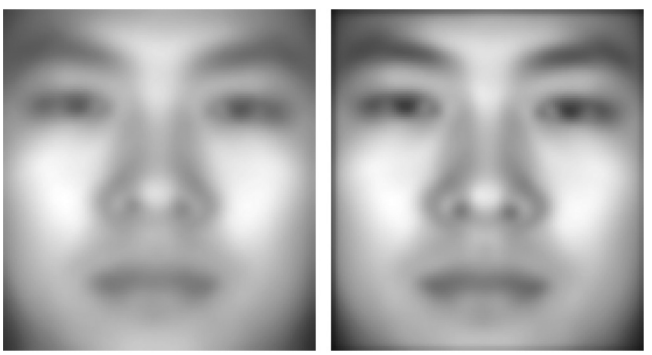

Fig. 9 Average face across all bands and all the dataset using (a) the single-band manual alignment and (b) congealing.

experiment III) permits to obtain identical results with respect to $W_2(\lambda)$, but still lower than those of $W_3(\lambda)$. Using congealing instead, the accuracy obtained in experiments I and II is already higher than that obtained by $W_3(\lambda)$. Furthermore, all the results obtained are outperformed by those obtained in experiment III.

The plots of the selected filters and camera sensitivities selected for each experiment are shown in Fig. 10. Using the single-band manual alignment, the best color filter selected in experiment I is 263, while experiments II and III selected the same color filter (251) coupled with the spectral sensitivities of the same camera (Point Grey Grasshopper 50S5C), with experiment II using just the green channel. A similar behavior can also be seen using the congealed images. In this case, the best color filter selected in experiment I is 201, while experiments II and III selected the same color filter (194) coupled with the spectral sensitivities of the same camera (Canon 5D Mark II), with experiment II using just the blue channel.

As a further analysis of the performance, the receiver operating characteristic (ROC) curve is shown in Fig. 11. Four different curves are reported that are relative to four different algorithms: 3-D-DCT with PLS regression,[8] spatio-spectral information fusion with PLS regression,[20] the proposed method using the $W_3(\lambda)$ projection (i.e., last line in Table 1), and the proposed method using a color filter on multiple color channels (i.e., experiment III, last line in Table 2). From the plots, it is possible to see that the curves relative to the proposed method have a larger area, thus further confirming the results reported in Tables 1 and 2.

Finally, in Fig. 12 the execution time of the proposed method is reported. Two different instances are considered: in the first one the spectral projection is mathematically

Table 2 Average recognition rates and standard deviations (%) for 10-fold experiments on the PolyU-HSFD database.

Experiment	Alignment type	Average recognition rate ± std
I (color filter on gray-level camera)	Single-band (manual)	98.12 ± 0.02
II (color filter on single-color channel)	Single-band (manual)	98.25 ± 0.01
III (color filter on multiple color channels)	Single-band (manual)	98.34 ± 0.02
I (color filter on gray-level camera)	Congealing	99.66 ± 0.01
II (color filter on single-color channel)	Congealing	99.69 ± 0.01
III (color filter on multiple color channels)	Congealing	99.79 ± 0.01

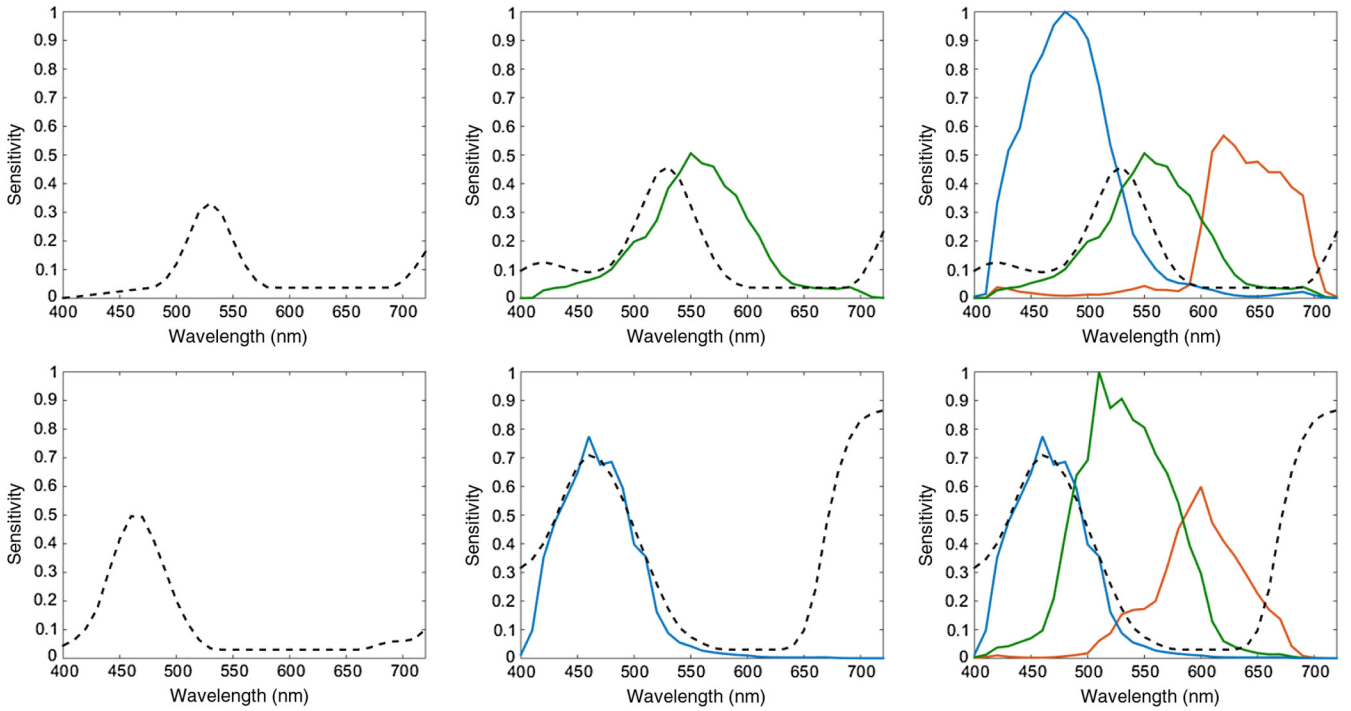

Fig. 10 Optimal filters and camera sensitivities found for experiments (a) I, (b) II, and (c) III varying the alignment method: single-band manual alignment (top) and congealing (bottom).

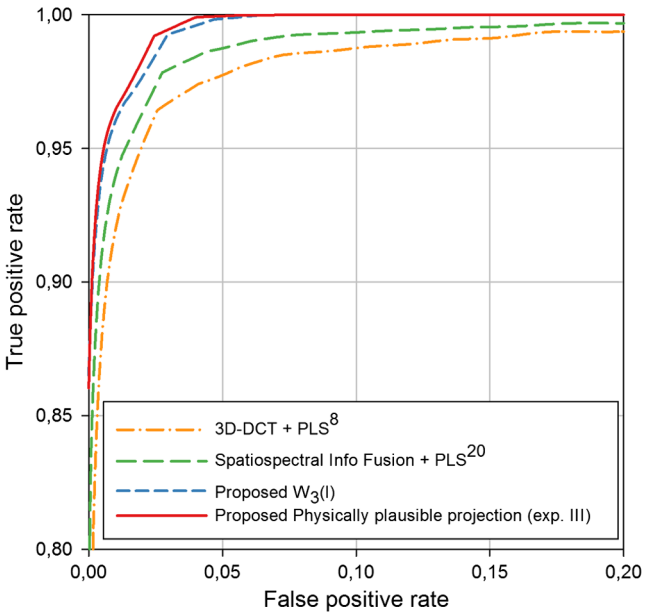

Fig. 11 ROC curve of the best proposed methods in Tables 1 and 2 with those of the best methods in the state of the art.

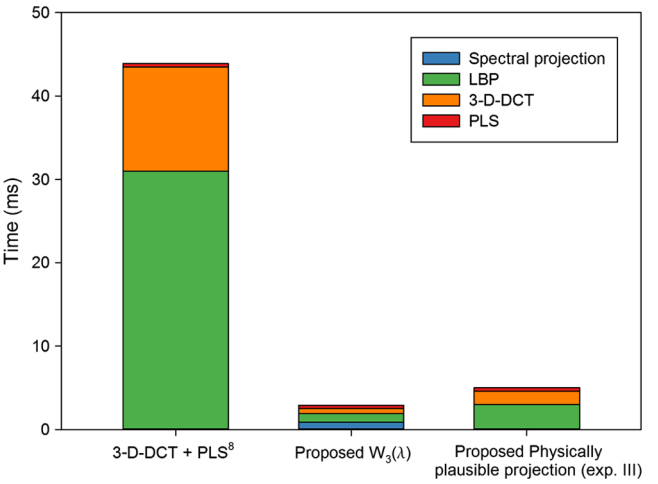

Fig. 12 Execution time of two variants of the proposed method and the baseline method[8] divided into the different common modules they are composed of.

performed using the optimized $W_3(\lambda)$ projection (and thus impacts on execution time), while in the second one it is performed by placing a color filter on a color camera (and thus does not impact on execution time). Since the proposed method builds on top of Ref. 8, its execution time is reported as a comparison. Execution times are subdivided to see the time-impact of the different modules that the considered algorithms share. Times are measured on a machine with Windows 7, 16 GB of RAM, i5-2500K processor, using pure MATLAB® code (MATLAB® version R2014b). From the plots, it can be noticed that the most expensive operations in Ref. 8 are the normalization for the LBP and the feature extraction step (3-D-DCT). These operations are expensive mostly because they are performed on all the spectral bands. The proposed method, by reducing the spectral dimension through linear spectral projections, is able to reduce the computational time from 43.4 ms to 3.9 to 5.9 ms.

3.5 *Cross-Database Experiments*

In order to investigate the generalization ability of the proposed method, experiments on two further standard datasets

Table 3 Average recognition rates and standard deviations (%) for 10-fold experiments on the CMU-HSFD and UWA-HSFD databases.

Algorithm	Average recognition rate ± std	
	CMU-HSFD	UWA-HSFD
3-D-DCT + PLS[8]	97.21 ± 1.54	96.26 ± 1.96
Spatiospectral information fusion + PLS[20]	99.10 ± 0.57	98.20 ± 1.21
Proposed $W_3(\lambda)$ (unbounded band combination with congealing)	99.47 ± 0.31	99.32 ± 0.87
Proposed physically plausible projection (Exp. III with congealing)	99.94 ± 0.01	99.86 ± 0.01

are carried out. The first dataset is the CMU hyperspectral face database (CMU-HSFD),[25] that consists of 4 to 20 cubes of 48 subjects with each cube containing 65 bands covering the spectral range of 450 to 1090 nm in steps of 10 nm. The second dataset is the UWA-hyperspectral face database (UWA-HSFD)[20] that consists of 120 hyperspectral image cubes of 70 subjects with each cube containing 33 bands covering the spectral range of 400 to 720 nm in steps of 10 nm. Following Ref. 20, faces are cropped and resized to 40×40 and 30×30 pixels for CMU and UWA-HSFD, respectively. For both the databases, 10-fold cross validation is performed by randomly selecting gallery/probe combinations in each fold. Two variants of the proposed method are evaluated: the proposed method using the $W_3(\lambda)$ projection and the proposed method using a color filter on multiple color channels. In both cases, the same parameters used on the PolyU-HSFD dataset are used. To this end, the spectral dimensions of the cubes in the CMU dataset have been downsampled to match that of both PolyU-HSFD and UWA-HSFD datasets. The two variants are compared with the best methods in the state of the art, i.e., 3-D-DCT with PLS regression[8] and spatiospectral information fusion with PLS regression.[20] Numerical results in terms of recognition accuracy are reported in Table 3, while ROCs are shown in Fig. 13. From both the numerical results and the plots, it is possible to see that the two variants of the proposed method are able to outperform the best methods in the state on both datasets, even if not specifically optimized on these unseen datasets.

4 Conclusion

In this paper, it is shown that the performance of hyperspectral face recognition algorithms can be improved by just considering single and multiple 1-D projections along the spectral dimension of the full spectral cube. Three different projections have been investigated: single-spectral band selection, nonnegative spectral band combination, and unbounded spectral band combination.

Experiments were performed on a standard hyperspectral dataset and the results of the proposed algorithm were compared with eight existing hyperspectral face recognition algorithms. Experimental results showed that the application of the optimal linear projections can improve the performance of the best hyperspectral face recognition algorithm in the state of the art by almost 4%, reaching an average recognition rate on a 10-fold experiment of 99.11%. Since the linear projections considered in this paper can be directly performed on the acquisition device with a proper transmission filter, we also investigated the use of linear projections compatible with physically plausible optical filters. Experimental results showed that the recognition rate using physically plausible filters is less than 1% lower than that of using optimized projections. The use of an automatic full-band face alignment algorithm permits to further improve the recognition accuracy to 99.79% using physically plausible filters. Cross-database experiments further confirmed the effectiveness and robustness of the proposed method.

Possible future works involve the evaluation of the performance of the proposed method at the different stages of a typical imaging processing pipeline[26] with particular interest on illuminant estimation and removal.[27]

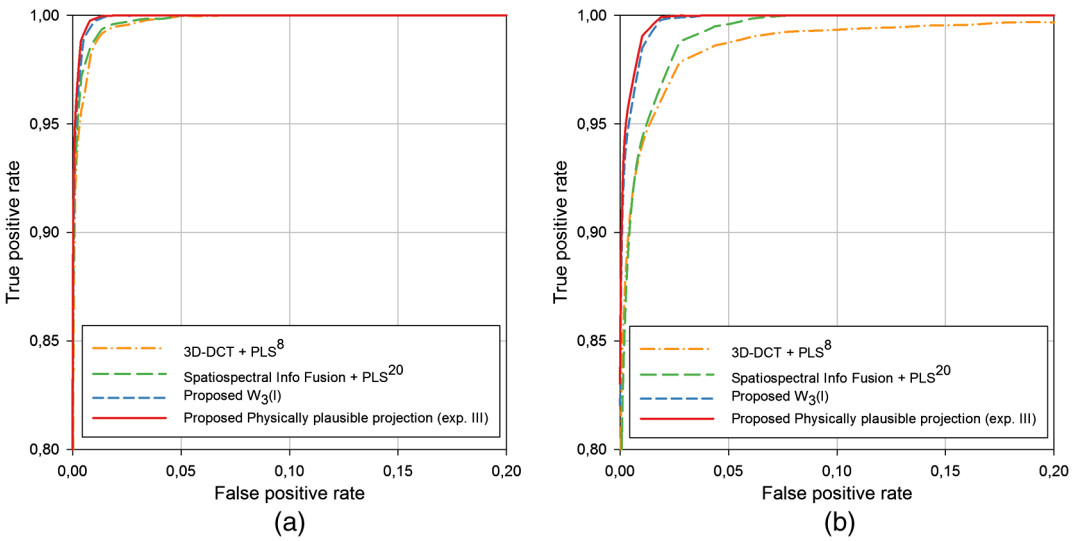

Fig. 13 ROC curves of the best proposed methods in Tables 1 and 2 with those of the best methods in the state of the art (a) on the CMU-HSFD database and (b) on the UWA-HSFD database.

Acknowledgments

The work has been partially done during a visiting period at Stanford University. The author would like to thank Professor J. E. Farrell who provided insight and expertise that greatly assisted the research.

References

1. W. Zhao et al., "Face recognition: a literature survey," *ACM Comput. Surv.* **35**(4), 399–458 (2003).
2. S. Z. Li and A. K. Jain, *Handbook of Face Recognition*, 2nd ed., Springer-Verlag, London (2011).
3. W. Di et al., "Studies on hyperspectral face recognition in visible spectrum with feature band selection," *IEEE Trans. Syst. Man Cybern. Part A* **40**(6), 1354–1361 (2010).
4. Z. Pan et al., "Face recognition in hyperspectral images," *IEEE Trans. Pattern Anal. Mach. Intell.* **25**(12), 1552–1560 (2003).
5. Z. Pan, G. Healey, and B. Tromberg, "Comparison of spectral-only and spectral/spatial face recognition for personal identity verification," *EURASIP J. Adv. Signal Process.* **2009**, 1–7 (2009).
6. S. A. Robila, "Toward hyperspectral face recognition," *Proc. SPIE* **6812**, 68120X (2008).
7. L. Shen and S. Zheng, "Hyperspectral face recognition using 3D Gabor wavelets," in *21st Int. Conf. on Pattern Recognition (ICPR 2012)*, pp. 1574–1577, IEEE (2012).
8. M. Uzair, A. Mahmood, and A. S. Mian, "Hyperspectral face recognition using 3d-dct and partial least squares," in *British Machine Vision Conf. (BMVC)* (2013).
9. D. Skočaj, A. Leonardis, and H. Bischof, "Weighted and robust learning of subspace representations," *Pattern Recognit.* **40**(5), 1556–1569 (2007).
10. P. Y. Han, A. T. B. Jin, and L. H. Siong, "Eigenvector weighting function in face recognition," *Discrete Dyn. Nat. Soc.* **2011**, 1–19 (2011).
11. M. Kawulok, J. Wu, and E. R. Hancock, "Supervised relevance maps for increasing the distinctiveness of facial images," *Pattern Recognit.* **44**(4), 929–939 (2011).
12. D.-L. Li et al., "Face recognition using nonparametric-weighted fisherfaces," *EURASIP J. Adv. Signal Process.* **2012**(1), 1–11 (2012).
13. "The Hong Kong Polytechnic University hyperspectral face database (PolyU-HSFD)," http://www4.comp.polyu.edu.hk/~biometrics/hsi/hyper_face.htm (02 December 2016)
14. S. Bianco, "Can linear data projection improve hyperspectral face recognition?" in *Computational Color Imaging*, pp. 161–170, Springer (2015).
15. T. Ahonen, A. Hadid, and M. Pietikäinen, "Face recognition with local binary patterns," in *Computer Vision (ECCV 2004)*, pp. 469–481, Springer (2004).
16. N. Ahmed, T. Natarajan, and K. R. Rao, "Discrete cosine transfom," *IEEE Trans. Comput.* **100**(1), 90–93 (1974).
17. R. Rosipal and N. Krämer, "Overview and recent advances in partial least squares," in *Subspace, Latent Structure and Feature Selection*, pp. 34–51, Springer (2006).
18. J. Kennedy, "Particle swarm optimization," in *Encyclopedia of Machine Learning*, pp. 760–766, Springer (2011).
19. S. Bianco and R. Schettini, "Two new von kries based chromatic adaptation transforms found by numerical optimization," *Color Res. Appl.* **35**(3), 184–192 (2010).
20. M. Uzair, A. Mahmood, and A. Mian, "Hyperspectral face recognition with spatiospectral information fusion and pls regression," *IEEE Trans. Image Process.* **24**(3), 1127–1137 (2015).
21. W. Zijlstra, A. Buursma, and W. M. Van der Roest, "Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin, and methemoglobin," *Clin. Chem.* **37**(9), 1633–1638 (1991).
22. Rosco, www.rosco.com (02 December 2016).
23. J. Jiang et al., "What is the space of spectral sensitivity functions for digital color cameras?" in *2013 IEEE Workshop on Applications of Computer Vision (WACV)*, pp. 168–179, IEEE (2013).
24. G. B. Huang, V. Jain, and E. Learned-Miller, "Unsupervised joint alignment of complex images," in *IEEE 11th Int. Conf. on Computer Vision (ICCV 2007)*, pp. 1–8, IEEE (2007).
25. L. J. Denes, P. Metes, and Y. Liu, *Hyperspectral Face Database*, Carnegie Mellon University, The Robotics Institute (2002).
26. S. Bianco et al., "Color correction pipeline optimization for digital cameras," *J. Electron. Imaging* **22**(2), 023014 (2013).
27. S. Bianco and R. Schettini, "Adaptive color constancy using faces," *IEEE Trans. Pattern Anal. Mach. Intell.* **36**(8), 1505–1518 (2014).

Simone Bianco obtained his PhD in computer science at DISCo (Dipartimento di Informatica, Sistemistica e Comunicazione) of the University of Milano-Bicocca, Italy, in 2010. He obtained his BSc and the MSc degrees in mathematics from the University of Milano-Bicocca, Italy, in 2003 and 2006, respectively. He is currently assistant professor and his research interests include computer vision, machine learning, optimization algorithms, and color imaging.