Can Linear Data Projection Improve Hyperspectral Face Recognition?

Simone Bianco(✉)

University of Milano-Bicocca, 20126 Milano, Italy
simone.bianco@disco.unimib.it

Abstract. This paper investigates if the performance of hyperspectral face recognition algorithms can be improved by considering 1D projections of the whole spectral data along the spectral dimension. Three different projections are investigated: single spectral band selection, non-negative spectral band combination, and unbounded spectral band combination. Experiments are performed on a standard hyperspectral dataset and the obtained results outperform seven existing hyperspectral face recognition algorithms.

1 Introduction

Since intra-person differences are often larger than inter-person ones in 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 inter-band 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 2D images. The different approaches for dimensionality reduction and feature extraction range from the sub-sampling of the hyperspectral data [4,5,7] to the more promising approaches which use whole-band features [3,8].

A. Trémeau et al. (Eds.): CCIW 2015, LNCS 9016, pp. 161–170, 2015.
DOI: 10.1007/978-3-319-15979-9_16

Starting from the best hyperspectral method in the state of the art [8], this paper investigates if the use of linear projections along the spectral dimension can improve face recognition performance with respect to the use of the full hyperspectral data.

The experiments are performed on the PolyU Hyperspectral [3,9] standard hyperspectral face database. The results are compared with seven existing hyperspectral face recognition algorithms.

2 Baseline Method

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

The first step consists in filtering the individual bands with a circular (8,1) neighborhood LBP [12] 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 (3D-DCT). The Discrete Cosine Transform (DCT) [13] decomposes a discrete signal into linear combination of independent cosine basis functions. DCT tends to generate a representation in which the low-frequency coefficients encode most of the signal information. A compact representation can be obtained by selecting as features only the low-frequency coefficients. The 3D-DCT of a hyperspectral cube $H(x, y, \lambda)$ with size $N_1 \times N_2 \times N_3$ is given by

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

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, and therefore the high-frequency coefficients can be discarded. In order to construct the feature vector, in [8] a frequency sub-cube $\Gamma(u, v, w)$ of dimensions $(\alpha \times \beta \times \gamma)$ is sampled by retaining only the low-frequency elements around the origin of $F(u, v, w)$. The sub-cube $\Gamma(u, v, w)$ 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 last step consists in the use of the Partial Least Squares (PLS) regression [14] 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 between different variable sets. The only parameter to be set in PLS is the number of latent variables to use.

3 The Proposed Method

Building on top of the method of Uzair et al. [8], this work wants to understand if 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 [8], 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 2D image, forcing $\gamma = 1$ for the sub-cube size. 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

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

which can be seen as a non-negative linear combination of the different hyperspectral bands. This is an operation analogue 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 equation 4 removing the lower and upper bounds on the filter coefficients w_i, i.e.:

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

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) [10,11] 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 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

Fig. 1. A hyperspectral face cube from the PolyU-HSFD dataset

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.

4 Experiments

4.1 Dataset

The hyperspectral face database used is the Hong Kong Polytechnic University Hyperspectral Face Database(PolyU-HSFD) [3,9]. It consists of hyperspectral image cubes acquired using a CRIs VariSpec Liquid Crystal Tuneable Filter. Each cube contains 33 bands acquired in the 400-720nm spectral range in 10nm 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, skin conditionss, etc.). Signal to noise ratio (SNR) in bands near the blue wavelength is very low, and the database contains inter-band misalignments due to subject movements during the acquisition at the different wavelengths.

The database contains a total of 48 subjects (13 females and 35 males). For each of the first 25 subjects four to seven cubes are available, while the remaining 23 subjects only have one cube each. Following the experimental protocol of [3,7], 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 ten times and the results are averaged. As in [3] the eye, nose tip, and mouth corners coordinates were located manually for image registration, 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 Figure 1, while examples of appearance variations are reported in Figure 2.

4.2 Compared Hyperspectral Face Recognition Algorithms

The seven existing hyperspectral face recognition algorithms used for comparisons include Spectral Signature Matching [4], Spectral Angle Measurement [6],

Fig. 2. Examples of appearance variations. The same hyperspectral band corresponding to $\lambda_{15} = 540$nm is selected for all subjects.

Spectral Eigenface [5], 2D PCA [3], 3D Gabor Wavelets [7], and 2D and 3D-DCT with PLS regression [8]. The parameters of these algorithms are set as follows. For spectral signature matching algorithm [4], five adjacent square regions of size 17x17 pixels arranged in a cross pattern are used to represent hair, forehead and cheeks. For the lips, square regions of size 9x9 pixels are used . For Spectral Eigenface [5], 99% energy is preserved by retaining 48 PCA basis vectors. For 2D PCA [3], 99% energy is preserved by retaining 27 PCA basis vectors. For the 3D Gabor method, 52 Gabor wavelets are used for feature generation as recommended by [7]. For the 2D and 3D-DCT [8] method the parameters are taken as suggested by the authors: $\alpha = \beta = \gamma = 10$ for the sub-cube size to extract the features, and 45 PLS basis.

4.3 Results

The results of the hyperspectral face recognition algorithms compared are reported in terms of average recognition rate in Table 1. The results of Spectral Signature Matching [4], Spectral Angle Measurement [6], Spectral Eigenface [5], 2D PCA [3], 3D Gabor Wavelets [7], and 2D and 3D-DCT with PLS regression [8] are all taken from [8], with the only exception of the 3D-DCT method for which the results using an 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 4.3% up to 6.11%. The best projections found for $W_1(\lambda)$, $W_2(\lambda)$, and $W_3(\lambda)$ are reported in Figure 3. 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 [3,15].

As already said in Section 3 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 exploiting the linearity of equation 2 it could be realized by subtracting two

Table 1. Average recognition rates and standard deviations (%) for ten-fold experiments on the 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)
2D PCA [3]	71.11 (3.16)
3D Gabor Wavelets [7]	90.19 (2.09)
2D-DCT + PLS [8]	91.43 (2.10)
3D-DCT + PLS [8]	93.00 (2.27)
3D-DCT + PLS (author's implementation)	93.32 (3.13)
Proposed ($W_1(\lambda)$, single band selection)	97.20 (1.66)
Proposed ($W_2(\lambda)$, non-negative band combination)	98.34 (1.83)
Proposed ($W_3(\lambda)$, unbounded band combination)	**99.11 (1.21)**

different optical filters $W_3^+(\lambda_i)$ and $W_3^-(\lambda_i)$:

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

where

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

and

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

Some examples of the projected output given by applying equation 2 with the optimal $W_1(\lambda)$, $W_2(\lambda)$, and $W_3(\lambda)$ projections are reported in Figure 4.

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 inter-band misalignments evident resulting in more blurred images.

In Figure 5 some examples of errors across the ten-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.

The sensitivity of the proposed method is analyzed in Figure 6 by plotting equal recognition rate curves as a function of number of PLS basis and sub-cube size ($\alpha = \beta$ and $\gamma = 1$, due to the effect of the projection).

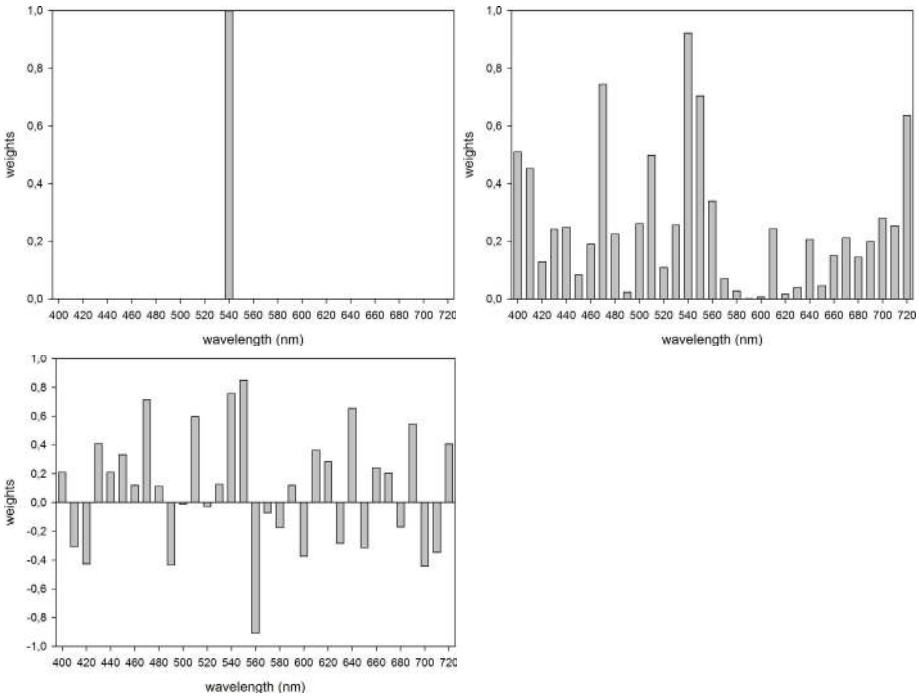

Fig. 3. Best projections found: W_1 (top left), W_2 (top right), and W_3 (bottom left)

Fig. 4. Examples of the projections obtained by applying the optimal projections found. $W_1(\lambda)$, single band selection (left); $W_2(\lambda)$, non-negative linear combination (middle); $W_3(\lambda)$ and unbounded linear combination (right).

Fig. 5. Examples of errors for the ten-fold experiment using the $W_3(\lambda)$ projection: gallery cubes (top), probes (middle), gallery cubes for the predicted identity (bottom)

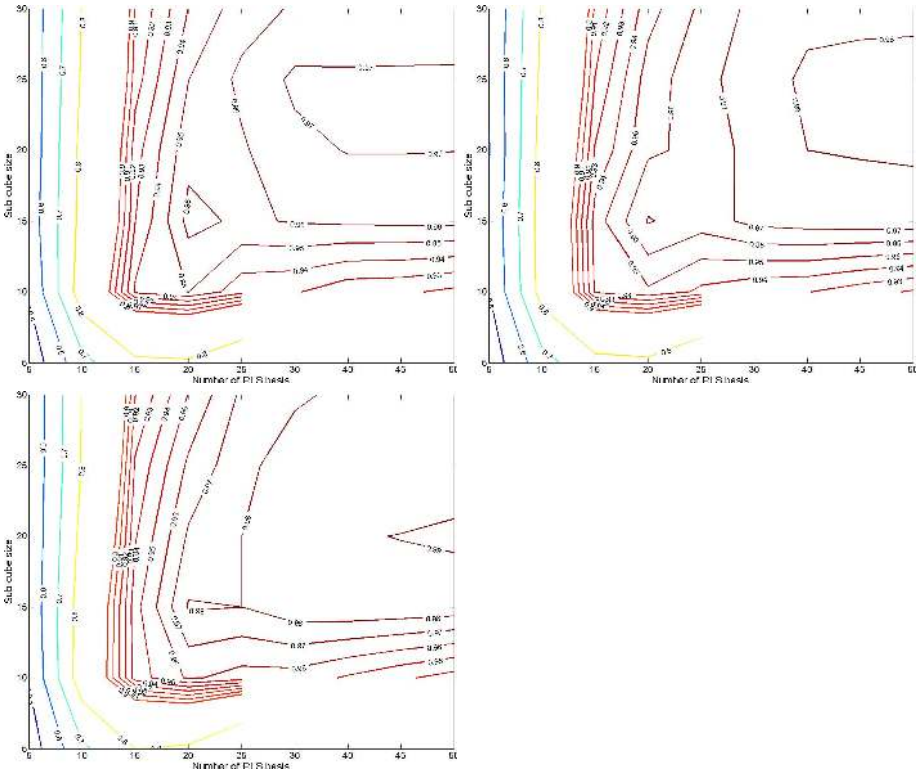

Fig. 6. Equal recognition rate curves as a function of sub-cube size (y-axis) and number of PLS basis (x-axis): single band selection (top left), non-negative linear combination (top right), and unbounded linear combination (bottom left)

5 Conclusion

In this paper it is shown that the performance of hyperspectral face recognition algorithms can be improved by just considering 1D projections along the spectral dimension of the full spectral cube. Three different projections have been investigated: single spectral band selection, non-negative 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 seven 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 more than 6%, reaching an average recognition rate on a ten-fold experiment of more than 99%.

As future work it will be investigated the use of linear projections compatible with physically plausible optical filters, by adding smoothness constraint on the projection weights. It will be also studied if multiple linear projections can further improve the recognition rate and which are the best fusion strategies. Furthermore, it will be investigated if the approach proposed in this work can be applied to hyperspectral images recovered from traditional RGB images using spectral recovery techniques [16].

References

1. Zhao, W., Chellappa, R., Phillips, P.J., Rosenfeld, A.: Face recognition: A literature survey. ACM Computing Surveys (CSUR) **35**(4), 399–458 (2003)
2. Li, S.Z., Jain, A.K.: Handbook of face recognition, 2nd edn. Springer (2011)
3. Di, W., Zhang, D., Pan, Q.: Studies on hyperspectral face recognition in visible spectrum with feature band selection. IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans **40**(6), 1354–1361 (2010)
4. Pan, Z., Healey, G., Prasad, M., Tromberg, B.: Face recognition in hyperspectral images. IEEE Transactions on Pattern Analysis and Machine Intelligence **25**(12), 1552–1560 (2003)
5. Pan, Z., Healey, G., Tromberg, B.: Comparison of spectral-only and spectral/spatial face recognition for personal identity verification. EURASIP Journal on Advances in Signal Processing **8**, 943602 (2009)
6. Robila, S.A.: Toward hyperspectral face recognition. In: Electronic Imaging, International Society for Optics and Photonics, pp. 68120X–68120X (2008)
7. Shen, L., Zheng, S.: Hyperspectral face recognition using 3D gabor wavelets. In: 21st IEEE International Conference on Pattern Recognition (ICPR), pp. 1574–1577 (2012)
8. Uzair, M., Mahmood, A., Mian, A.: Hyperspectral face recognition using 3D-DCT and partial least squares. In: British Machine Vision Conference, pp. 57.1–5-7.10 (2013)
9. PolyU-HSFD. www4.comp.polyu.edu.hk/~biometrics/
10. Kennedy, J., Eberhart, R.: Particle swarm optimization. In: Proceedings of the IEEE International Conference on Neural Networks, vol. 4, pp. 1942–1948 (1995)

11. Bianco, S., Schettini, R.: Two new von Kries based chromatic adaptation transforms found by numerical optimization. Color Research & Application **35**(3), 184–192 (2010)
12. Ahonen, T., Hadid, A., Pietikäinen, M.: Face recognition with local binary patterns. In: Pajdla, T., Matas, J.G. (eds.) ECCV 2004. LNCS, vol. 3021, pp. 469–481. Springer, Heidelberg (2004)
13. Ahmed, N., Natarajan, T., Rao, K.R.: Discrete cosine transform. IEEE Transactions on Computers **100**(1), 90–93 (1974)
14. Rosipal, R., Krämer, N.: Overview and recent advances in partial least squares. In: Subspace, Latent Structure and Feature Selection, pp. 34–51 (2006)
15. Zijlstra, W.G., Buursma, A., Meeuwsen-Van der Roest, W.P.: Absorption spectra of human fetal and adult oxyhemoglobin, de-oxyhemoglobin, carboxyhemoglobin, and methemoglobin. Clinical Chemistry **37**(9), 1633–1638 (1991)
16. Bianco, S.: Reflectance spectra recovery from tristimulus values by adaptive estimation with metameric shape correction. JOSA A **27**(8), 1868–1877 (2010)