Hyperspectral face recognition with a spatial information fusion for local dynamic texture patterns and collaborative representation classifier

Min Hao1,2 | Guangyuan Liu1,2 | Desheng Xie3

1 College of Electronic and Information Engineering, Southwest University, Chongqing, China
2 Chongqing Key Laboratory of Non-linear Circuit and Intelligent Information Processing, Southwest University, Chongqing, China
3 Suson Co., Ltd., Sichuan, China

Correspondence
Guangyuan Liu, Lab for Signal and Information Processing, Southwest University, Chongqing, China. Email: liugy@swu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Number: 61872301; Chongqing Key Laboratory of Non-linear Circuit and Intelligent Information Processing; Fundamental Research Funds for the Central Universities, Grant/Award Number: XDJK2019C021

Abstract
Hyperspectral face recognition provides improved classification rates due to its abundant information in the face cubes of every subject in hyperspectral face databases. However, while offering excellent opportunities, it also brings new challenges, such as low signal-to-noise ratio, interband misalignment, and high data dimensionality. Based on these ad hoc problems, literature has already proposed some optimisation methods including dimensionality reduction, image denoising, and alignment to perform face recognition, yet lacking comprehensive evaluation. This paper proposes a novel hyperspectral face recognition algorithm that is based on spatial information fusion for feature extraction (histogram of local dynamic texture patterns) and collaborative representation classifier for classification. Meanwhile, the algorithm is applied to three popular hyperspectral face databases, Carnegie Mellon University (CMU)-hyperspectral face database (HSFD), University of Western Australia (UWA)-HSFD, and Hong Kong Polytechnic University (PolyU)-HSFD databases. Experimental results demonstrate that CMU-HSFD and UWA-HSFD databases achieve very competitive classification results. PolyU-HSFD database also achieves rather good classification rates. The best recognition results are 98.5% ± 0.95, 96.6% ± 0.98, and 94.0% ± 2.86 for CMU-HSFD, UWA-HSFD and PolyU-HSFD, respectively. It demonstrates experimentally that this algorithm can be used to recognise faces. Moreover, we compared eight existing state-of-the-art face recognition techniques with our proposed method in performing hyperspectral face recognition. In this research, we formulate hyperspectral face recognition as an image-set classification problem and evaluate the performances compared with other kinds of algorithms. Comparisons with the eight existing hyperspectral face recognition techniques on three standard datasets show that the proposed algorithm outperforms most other state-of-the-art algorithms, indicating that it is a promising approach for hyperspectral face recognition.

1 | INTRODUCTION

With popularity over the use of hyperspectral cameras, hyperspectral imaging (HSI) technique has been explored to conduct research on many fields such as remote sensing and computer vision [5, 6]. HSI provides both spatial and spectral information, where more useful measurements can be produced. Therefore, it brings new opportunities to discriminate objects through the spatiotemporal modalities.

Hyperspectral face recognition is a very important yet challenging problem [7–9]. By acquiring more spectral information (usually, the visible spectrum and beyond), face recognition can be possibly improved. The advantages of adding the spectral properties lie in the following aspects such as enhancing the inter-person differentiation [9] (referring to not only the surface appearance but also the underlying subsurface features) and offering the discriminative perception between a real and synthetic face [10]. Essentially, it is because that HSI brings more biometric signatures. In the literature, there have already been some existing research methods exploring either spatial/spectral features alone or spatiotemporal features jointly. The most popular indicators applied in face recognition research

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.
© 2021 The Authors. IET Image Processing published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology
contain local binary pattern (LBP) \([7, 11, 12]\), histogram of oriented gradients (HOG) \([14, 15]\), log-polar transform \([14]\), Gabor filter bank descriptor \([1, 16]\), sparse tensor embedding (STE) \([17]\), scale invariant feature transform (SIFT) \([18, 19]\), and spatiotemporal corner feature \([20]\), and so forth. The above indicators demonstrate very robust and competitive achievement alone or by combination in many experimental validations.

Meanwhile, HSI can also bring some drawbacks including inter-band misalignment, intra- and inter-person differences, low signal-to-noise ratio (SNR), and high data dimensionality \([21–23]\). For example, because of some casual movements of a subject such as eye blinking, head rotation, and body shaking, these actions may necessarily lead to the undesired misalignment. The misalignment would induce intra- and inter-person variations. Both natural and synthetic lights demonstrate lower spectral intensity near the blue part of the whole spectrum. Therefore, these images acquired by HSI demonstrate lower SNRs. Given the abundant spectral information, it will introduce the curse of dimensionality problems, increasing the computational complexity and information redundancy rate.

It is due to these challenges that they motivate the researchers to indulge in hyperspectral face recognition. But existing research efforts do not consider the collaborative fusion of the variance-based (or standard deviation-based, for convenience these two terms will be used interchangeably) spatial local dynamic texture patterns (LDTP) information (also known as sensitivity) and oriented gradients. Therefore, a novel biometric measurement (histogram of LDTP (HoLDTP)) is formally proposed to perform hyperspectral face recognition.

Besides the feature extraction, another dominant process is to choose effective face classification methods in terms of different constraint conditions. Considering the structure of HSI face cube, image-set classification algorithms can be chosen to do the classification (in fact, it is to solve the convex optimisation problem) \([8, 24]\). Classification methods such as sparse representation classifier \([8, 25, 26]\), collaborative representation classifier (CRC) \([8, 14]\), partial least squares (PLS) \([13]\), affine hull-based image-set distance (AHISD) \([2]\), convex hull-based image-set distance (CHISD) \([2]\), discriminative canonical correlation \([27]\), manifold-manifold distance \([28]\), manifold discriminant analysis \([29]\), and sparse approximated nearest point (SANP) \([24]\) have been already developed for face recognition. Recently, there has been growing attention in the research of representation theory. Among all the representation methods, CRC is a classical sparse representation, proven to be an effective way for dealing with the classification process. Therefore, in this study, we introduce our proposed hyperspectral face recognition algorithm that combines HoLDTP operator with CRC classifier.

2 RELATED RESEARCH WORK

For hyperspectral face recognition, there is a gradually increasing number of research addressing these issues. Many researches achieve well under active investigation on several databases acquired under controlled or unconstrained conditions. Pan et al. \([10]\) demonstrated the useful discriminants for human face recognition over a database, which contains 31 bands over the near-infrared (0.7–1.0 \(\mu\)m) for 200 subjects. The Carnegie Mellon University constructed a database (lasting two months at multiple sessions), which recruited 18 subjects. Fifty-four diverse faces were collected in the spectral range from 0.45 to 1.1 \(\mu\)m. And they also proved the spectral asymmetry for personal identification \([30]\). Di et al. \([31]\) established hyperspectral face database (HSFD) containing 25 subjects and 33 bands over the visible light spectrum (0.4–0.72 \(\mu\)m), demonstrating that the recognition result with the selective bands outperforms those by using a single band, the whole bands, or even the canonical red, green, blue (RGB) bands. The literature represented the effectiveness of face recognition in multispectral ranges.

Most of the existing recognition methods perform feature extraction in their own and special ways. Ning et al. \([32]\) developed a new method to reduce dimensionality, named biomimetic uncorrelated locality discriminant projection (BULDP). Experimental findings showed that the proposed BULDP method obtained better discriminating performance for face recognition. Uzair et al. \([13]\) investigated hyperspectral face recognition using spatiotemporal covariance-based band fusion strategy and a PLS regression classifier. The fusion strategy was more effective than both simple spatial averaging and spectral integration since it captured the spatial plus the spectral feature variances. They achieved the state-of-the-art recognition performances for three public HSFDs. Zhao et al. \([17]\) proposed an STE approach where an image was viewed as a third-order tensor. Also, they defined sparse neighbourhoods and their tensor weights in STE. The experiments showed that the recognition performance outperformed that of the state-of-the-art approaches. Xie et al. The authors in \([11]\) performed hyperspectral face recognition using LBP and simplified Weber local descriptor, which were used to extract the effective local patterns from spatiotemporal information. Finally, it demonstrated an outstanding performance (92.8\%) for the UWA-HSFD. Also, Fan et al. \([33]\) investigated the existing limitations of using LBP to recognise the face. Then the gradient-weighted strategy was introduced to adjust regional histograms, developing the smoothness effect in terms of the discriminative contributions of the local positions. Also, Gaussian fuzzy region membership was present to enhance the registration robustness. Experimental results from the notoriously challenging the extended Yale B face corpus showed that the recognition ability was significantly great even in the most challenging conditions such as dimmer illumination and worse body distortion. Jabid et al. \([34]\) proposed a new local feature pattern (local directional pattern (LDP)) to recognise the face. This descriptor mainly encoded the different edge responses in all eight directions around each pixel. LDP histograms from several blocks were generated and then regularly concatenated into a unified feature vector. The results demonstrated that the proposed pattern had superior robustness compared with other existing popular descriptors. Tong et al.\([35]\) described a pattern descriptor (referred to as local dominant directional
symmetrical coding patterns (LDDSCPs)). The dominant directional pattern of expression can be constructed through the comparative computation between the convolution score of each pixel and the average of its neighbouring points. Then, considering the directional symmetry of the two groups, the authors stacked their corresponding histogram codes into a unified LDDSCP vector. Experimental results demonstrated that the LDDSCP descriptor, compared with other operators such as LBP and Gabor, achieved prominent recognition accuracy and computational complexity. Ghias et al. [36, 37] explored to describe the high-frequency detail in facial expressions and pose changes. They proposed a representation based on reliable anatomical features (such as vesselness features). This made the proposed representation much more robust both to pose and scale changes, significantly outperforming previously described methods. Through their research focusing only on the infrared spectrum (not extended to the visible spectrum and other ranges), they provided a new view of face recognition based on the anatomical features. Li et al. [20] proposed a novel corner feature for HSI defined as an extreme-constrained spatiotemporal corner. The experimental results represented that this novel approach can discriminate sufficient corner patterns from HSI, and its recognition performance was relatively better than those of other popular methods.

Meanwhile, there are also other advanced research in face recognition algorithms such as image-set-based recognition algorithms and composite RGB image-based recognition algorithms. Chang et al. [38] introduced a distribution separation measure for face recognition. Through ranking the separation measurement values, the optimal spectral range can be determined automatically. Based on a variety of experimental processes, the composite images formed by the selective spectrums can have more robust performance than the conventional fusion images. Chen et al. [14] performed comparative analysis for five existing descriptors that are popular in handling 2D face recognition. With a CRC as the classification choose, the best recognition accuracy for the PolyU-HSFD was 96.4% ± 2.3 and for the CMU-HSFD was 98.0%±0.7. Fan et al. [25] proposed a novel representation scheme named kernel sparse representation (or coding) codes a signal over a dictionary such that $y\approx\alpha$, and $\alpha$ is a sparse vector.

3 | CONTRIBUTION OF THE STUDY

Our contribution is to construct a novel feature operator (i.e. HoLDTP) for face recognition. Through the fusion of variance-based mapping (LDTP) and the oriented gradient information, HoLDTP can capture dynamic local textures that are more robust. To the best of our knowledge, this is the first attempt to incorporate LDTP with the oriented gradient information. The operator can substantially produce more abundant texture information, improving the dynamic representation of facial local texture. Based on the proposed HoLDTP, this study is to explore the feasibility/effectiveness of our proposed algorithm for hyperspectral face recognition. Meanwhile, it also provides an alternative promising algorithm for real-time hyperspectral face recognition.

4 | RESEARCH INTERPRETATION

4.1 | Research hypotheses

This study is designed to perform hyperspectral face recognition using our proposed method. The evaluation strategy is based on the image-set classification. Hence, the proposed operator in this study is assessed on its feature-characterising performance. According to the empirical analysis, this study hypothesises the following:

(1) Subject-specific image set lies in a unified linear subspace with carrying discriminative clustering properties between any two subspaces and close to each other inside the same subspace. That is, any spectral probe face image belonging to one subject can be represented by the linear combination of all entry vectors inside the image set. Further, it means that any probe can be coded as a sparse linear combination of all vectors inside all subspaces. For example, sparse representation (or coding) codes a signal $y$ over a dictionary $\emptyset$ such that $y\approx\alpha$, and $\alpha$ is a sparse vector.

(2) Wavelength-specific image set is also regarded as a manifold subspace. Their clustering properties between two subspaces are subject to the specific wavelength. It is known empirically that the subspace to subspace distance between bandwidths where there is a larger difference is significantly larger than one between those being a smaller difference.

(3) Certain linear combination of sample vectors can be represented as a probe image, intensively demonstrating that enough training samples for every image set do exist. On the other hand, it assumes that each subspace is approximating to a completed set. Figure 1 shows the discriminative clustering distributions under different mapping conditions (subject- or wavelength-specific). These structural hierarchies intuitively comply with our empirical assumptions for the intrinsic distributions of image sets.
FIGURE 1 Scatter diagram of image sets based on (a) subject, and (b) wavelength. Note that we select sets as a comparison plot whose wavelengths are 450, 460, 720 and 730 nm. And the illustration data are chosen from CMU-hyperspectral face database (HSFD). The axes represent the first three largest principal components.

TABLE 1 Hyperspectral face databases details

| Database | Subjects | hyperspectral cubes | Bands | Spectral range |
|----------|----------|---------------------|-------|----------------|
| CMU      | 48       | 150                 | 65    | 450–1090 nm    |
| UWA      | 80       | 147                 | 33    | 400–720 nm     |
| PolyU    | 25       | 113                 | 33    | 400–720 nm     |

FIGURE 2 Hyperspectral face cube from the CMU-HSFD database. Only even-numbered bands are shown.

FIGURE 3 Hyperspectral face cube from the UWA-HSFD

4.2 | Experimental data

We performed experiments on three standard databases. They are all popularly referred in many research fields and also available in public repositories. Table 1 gives a summary of the databases. The relevant database descriptions and processing strategies implemented in this study are explained as follows.

For the CMU-HSFD [30] (as illustrated in Figure 2), it contains sequential face images in 65 spectral bands covering the spectrum range of 450–1090 nm at 10 nm step. Forty-eight subjects participate in this collection period. Meanwhile, for each subject, 4–20 cubes are captured at different sessions by controlling different lighting conditions. But note that only the face cubes obtained with all the lights on are kept in our experiment. Finally, after the above exclusion, the experimental data only contains 147 HSI cubes. The gallery is formed through randomly taking only one cube for each subject, and then the remaining cubes are left to be probes. The spatial resolution of the images is 640 × 480. In the experiments, for each face image (spectral band), we automatically detect the face and eye regions and subsequently normalise for rotation and scale variations. Following the eye coordinates, a normalised fixed image of size 40 × 40 is then cropped.

For the UWA-HSFD [13, 39], the face cube contains 33 bands covering the spectral range of 400–720 nm with a 10 nm step. There are 80 subjects in this database with each subject consisting of 1–4 cubes. Figure 3 shows a hyperspectral face cube from the UWA-HSFD database. Subsequently, we randomly select one cube for each subject as a gallery with the remaining cubes as probes. In this dataset, each face image has a size of 420 × 420, cropped from the original image. Through the unified normalisation, faces are finally resized to 60 × 60 pixels.

For the PolyU-HSFD [23, 40] (Figure 4), each cube consists of 33 bands. The database also covers the spectral range of 400–
Hyperspectral imagery has been adopted recently to enhance the spectral dimensionality information of the facial regions and can boost the predictive performance of face recognition due to capturing more biometric characteristics, especially introducing much more spectrum information of faces. Capturing spatial characteristics alone on every 2D facial image inside the 3D hyperspectral facial cubes, this work proposes a novel information fusion, considering the collaborative binding of the LDTP information related to its corresponding variance (sensitivity) and the oriented gradient information. After fusion with these features, every 2D image can be coded as a new feature vector. As a feature input, the CRC-based voting strategy is used to classify the faces. The proposed scheme in this study for hyperspectral face recognition is organised as follows. First of all, facial region detection is performed automatically in HSFD; subsequently, the input facial image is relevantly cropped. Second, features are extracted through the cropped facial image. Eventually, we classify the probe face into a reasonable label. After obtaining all classified labels for each probe face in a cube, a voting evaluation will be performed to determine the final class label for the whole cube. The process of face recognition can be given below.

5.1 Automatic facial region detection in HSFD

A hyperspectral face cube, in general, consists of face images captured at multiple wavelengths of the electromagnetic spectrum. Hence, face recognition researches have been increasingly extended from the visible spectrum towards hyperspectral ranges. However, it is known that some spectrum ranges (e.g., bands near the blue wavelength region) have a very low SNR. It directly conducts a fact that some facial regions, in one way, may not be detected accurately. Thus, we assume here that the locations of geometric centres of the faces over all bands are aligned almost or have less displacement. In this case, an alternative calibration method is used to solve the problem of miss detections. That is, all the centre positions of all faces in the image cube can be estimated first. Then the weight values (voting around the detected centre position using Gaussian distribution function) are multiplied into the face matrix, respectively. After the similar weighted process for all faces in the same image cube, the accumulative collective face matrix has been achieved. Note that if a probable facial region has not been detected, the weights can be set to zeros. Finally, the average of the coordinates in a local neighbouring vicinity achieving maximum scores over all spectral bands is then identified as the face centre of the HSI cube. Meanwhile, the average weight and height of eigenface are obtained by calculating the mean weight and height of all detected faces. With the achieved values such as centre, weight, and height, all the facial regions in the image cube (especially those face images where facial regions have not been truly detected) can be determined. It is noteworthy that before face detection, noise reduction is performed for every hyperspectral face band image separately using block-matching and 3D filtering method [41]. Then, the facial region is automatically detected according to Masayuki Tanaka’s Matlab code [42]. Besides, this work applies the same operation as above for eye region detection. After two pupil centre positions are determined separately, rotation and scale normalisation can be conducted to all face images.

5.2 Facial feature extraction

This study innovatively introduces a variance-based spatial LDTP information fusion method to perform the facial feature extraction. Explicitly in the hyperspectral face, local texture information, known as an important intrinsic characteristic, can demonstrate its specificity and sensitivity. Specificity denotes the variable individual differences between different persons, while sensitivity represents the variable spatial reflectance discriminations for the specific subject. Hereby, we construct a new feature fusion of spatial LDTP information. Figure 5 illustrates the fusion process in detail.

5.2.1 Histogram of local dynamic texture patterns

As illustrated in the construction diagram in Figure 5, the spatial HoLDTP feature is introduced for characterising the local dynamic variations, for which we get the inspiration from other local texture patterns such as LBP, and HOGs. Accordingly, we substantially consider the discriminative oriented gradient of the
local variance or sensitivity. Note that we are not considering the pixel value but taking the local variance as the basic processing unit. This choice can rather introduce the local variable structure. Thereby, the second-order descriptor further strengthens the ability of HoLDTP to distinguish local texture information.

In this illustration, the first step is converting the original image from a face cube into a local variance image matrix. For example, we can statistically capture all the local variance values in terms of a fine-grained unit cell (UC; shown as the red rectangle in the above pipeline and set to 8 × 8 pixels in this study) across the whole face image. The way of UC stepping across the face can be divided into two ways (non-overlapping and overlapping). The overlapping method is chosen in this study to compute the local dynamic variance of each pixel, considering the diversity of variance distributions. It means that variance values will be statistically calculated around the centre of every pixel in the face. Finally, each face can be mapped into a new face image as related to its local dynamic texture pattern.

Second, we perform another feature mapping to extract the LBP value [7, 12]. Unsurprisingly, the generated variance matrix almost shows similar variable values, making them vulnerable to small disturbances. To address the prominent thresholding problem, a local ternary pattern is introduced to partially alleviate this problem, denoted in the following form.

\[
LTP_3^{(P,R)} = \sum_{i=1}^{P} \chi(g_i)3^{-1}, \quad \chi(g_i) = \begin{cases} 
0, & \mu - g_i > \tau \sigma \\
1, & |g_i - \mu| \leq \tau \sigma \\
2, & g_i - \mu > \tau \sigma 
\end{cases}
\]

where \(g_i\) is the variable value of the \(i\)th neighbouring point around each locus, \(\mu\) represents the mean value of its neighbouring points, \(\sigma\) is the corresponding standard deviation, and \(\tau\) is defined as the control parameter for adjusting the similarity between the neighbours. For example, 1 can be viewed as denoting neighbourhood variables sufficiently similar to \(\mu\). \(P, R\) mean the number of the neighbouring points around a locus and radius of the vicinity, respectively.

Further, in order to highlight the local texture, an adaptive weighting strategy has also been incorporated into the above computation. We employ the following weighting function:

\[
\psi_g(\|\nabla I\|_1) = \exp\left\{ \frac{\|\nabla I\|_1}{k^2} \right\} - 1
\]

where \(\|\nabla I\|_1\) denotes the average gradient magnitude (i.e. averaged \(L1\) norm) in the square neighbourhood of the pixel point, \(k\) is a free parameter trained from the gallery, and \(\psi_g(\cdot)\) the weight concerning a specified locus. Finally, the LBP value has the following formulation:

\[
\tilde{LTP}_3^{(P,R)} = LTP_3^{(P,R)} \cdot \psi_g(\|\nabla I\|_1)
\]

After the above processing, the local ternary pattern (LTP) operator can sufficiently characterise the local texture information. Figure 6 illustrates the comparative performance of different modifications. From these observations, we note that LBP has brought more redundant noises than LTP, and similarly, the weighted LBP also produces more singular perturbations.
in some local fine-grained loci than the weighted LTP, especially during the periorbital areas. The significant differences demonstrate the excellent ability of weighted LTP to capture local texture. Explicitly, as illustrated in the construction process, a similar fixed UC is selected as a mask template to compute the LTP value for each local region. Then, each LTP value will be calculated into its corresponding decimal type. Hereeto, LDTP mapping is obtained through the iterative LTP operation, demonstrating the generic distribution of local dynamic textures.

After that, the HOG operator is introduced to characterise the oriented gradients [15, 43]. As known in the literature, HOG is applied to describe the normalised histograms of local gradient patterns. Moreover, this study also introduces another coordinate pair beside the conventional horizontal-vertical pair as shown in Figure 7. Regarding the block segmentation discussed in many partition methods, we can also attempt to segment the whole face region into a grid of \( N \times N \) unit blocks (UBs). For example, an illustration of a face mapping segmented into \( 3 \times 3 \) blocks is presented in Figure 5. Certainly, the number of block segmentation is dependent on the specific application field.

Next, a face image is represented as a new feature vector (HoLDTP). Accordingly, the face cube will be constructed and grouped to form a HoLDTP Matrix. In summary, the proposed facial feature extraction method is explained in detail as shown in Algorithm 1.

5.3 Face representation

Through the above operation, each face image could be represented as a HoLDTP vector. Accordingly, each face cube can be coded as a HoLDTP matrix, each column that denotes a face image acquired under a specified wavelength. Finally, all cubes can be rearranged as a feature dictionary matrix \( \mathbf{H} = [H_1, ..., H_i, ..., H_L] \in \mathbb{R}^{d \times L} \), where \( d \) is the dimensionality of face feature, \( L \) is the number of subjects, and \( H_i = [H_{ij_1}, ..., H_{ij_k}, ..., H_{ij_m}] \in \mathbb{R}^{r \times \kappa \times \lambda} \) denotes the \( m_i \) set of face cubes of the \( i \)th subject, such that \( L = N \cdot \sum m_i \), where \( H_{ij} = [H_{ij1}, ..., H_{ijk}, ..., H_{ij\lambda}] \in \mathbb{R}^{r \times \kappa \times \lambda} \) represents the HoLDTP features for each spectral face in the image cube, and \( \{\lambda_k\}_{k=1}^N \) is a wavelength set.

5.4 CRC-based voting classification

Sparse representation and low-rank approximation have recently gained more attention for recognising human face [44]. Zhang et al. [8] found that the propelling influence for developing the performance of face recognition was the collaborative representation rather than the sparse representation. Chen et al. [14, 45] applied the voting techniques-based CRC classifier for hyperspectral face recognition. Li et al. [46] proposed a novel collaborative representation-based nearest neighbour algorithm for classifying face images. The method determined the label
of a probe face by majority voting of those with $k$ largest representation weights, also considering the local within-class collaborative representation (CR). Akhtar et al. [47] showed that sparseness should not be completely ignored for computational gains. They proposed a dense collaborative representation with a sparse representation that achieved higher accuracy and lower computational time in the resulting experiments. In this study, a sparse representation is applied to hyperspectral face recognition. The voting technique, as in [14], is applied to hyperspectral face recognition.

Let $X = [X_1, X_2, ..., X_C] \in R^{d \times n}$ denote the face dictionary whose columns are the training image vectors of the face cubes, where $C$ is the number of classes. Then,

$$X_i = [X_{i1}, X_{i2}, ..., X_{in}] \in R^{d \times n}$$

where $X_i$ is the training samples of class $i \in [1, C]$, and $n_i$ is the number of face cubes in the $i$th class, each of which is captured under different conditions. Further,

$$X_i' = [(X_{i1})', (X_{i2})', ..., (X_{in})'] \in R^{d \times N}$$

where $N$ is the number of spectral bands.

Let $Y = [y_1, y_2, ..., y_N] \in R^{d \times N}$ be a testing cube with $N$ face images. Based on the assumptions in Section 4, we assume that the image set is approximating to a complete set that requires containing enough training samples as much as possible. Then, once a query face image $y_k \in R^d$ comes, it can be coded as $y_k = X \tilde{\alpha} \in R^d$, where $\tilde{\alpha} = [\alpha_1, ..., \alpha_i, ..., \alpha_C]$ and $\alpha_i$ is the coding vector associated with the class $i$. In CRC, one has to solve the following optimisation problem:

$$\tilde{\alpha} = \arg \min_{\alpha} \|y_k - X \alpha\|^2_2 + \lambda \|\alpha\|^2_2$$

where $\lambda$ is a regularisation parameter and $\tilde{\alpha} = [\tilde{\alpha}_1, ..., \tilde{\alpha}_i, ..., \tilde{\alpha}_C]$. This optimisation problem can be solved through the following closed-form solution:

$$\tilde{\alpha} = (X^T X + \lambda I)^{-1} X^T y_k$$

While the calculation of $(X^T X + \lambda I)^{-1} X^T$ offline prior to classification can develop the computational complexity effectively. For a face $y_k$, it will be classified based on the optimal minimum of the least square error $e_i = \|y_k - X_i \tilde{\alpha}\|_2$, that is, identity($y_k$) = $\arg \min_i e_i$

Regarding the decision scheme of the majority-winning policy, there are a series of classification criteria for all face images $y_k \in Y$. Then we collect its statistic sorting results in descending order, denoting the form by $c = (c_1, c_2, ..., c_L)$ with $c_i \geq c_{i+1}$. Whichever method we choose (either full voting or partial voting), the identity of $Y$ is recognised to be the class that gets the maximum voting score (its number of appearances) within the chosen set of bands. In this study, according to the literature [14, 23, 45], we have chosen the bands 8—15, 7—43, and 7—30 as the selected partial subset of face bands for UWA-HSFD, CMU-HSFD, and PolyU-HSFD, respectively.

6 | EXPERIMENTS AND RESULTS

For validating the performance of our algorithm proposed in this work in hyperspectral face recognition, we compare the proposed algorithm with eight existing state-of-the-art face recognition algorithms using three standard hyperspectral face databases (CMU, UWA, and PolyU HSFD).

For all the databases, 10-fold cross-validation is performed in our experiments by randomly choosing different gallery/probe pairs in each fold. Besides, all experiments are iteratively run 100 times. As recommended in [8], for our proposed CRC-based voting classifiers, the regularisation parameters $\lambda$ are set to 0.005, 0.005, and 0.003 for CMU, UWA, and PolyU databases, respectively. A UC of size $3 \times 3$ is chosen as the standard mask template. And stepping depth applied in this study is set to 1 for obtaining the local variance around each pixel. Besides, the control parameter $\tau$ is set to 0.7, and $k$ is designated as 2 in all datasets. For other algorithms, their parameters are optimised as recommended by the original authors, respectively. The detailed specifications are shown as follows.

For 3D Gabor wavelets, each face cube is convolved with 52 wavelets for feature extraction similar to [1]. For the AHISD method, the bound is chosen between 1 and 5. For CHISD, the error penalty parameter is set to be the same as in [2] $(C = 100$ for gray-scale features and $C = 50$ for LBP in linear support vector machine (SVM)). Both methods apply principal component analysis (PCA) to preserve 90% of energy as before. For SANP, the regularisation parameter $\lambda$ is by default set to 0.01 like in [3]. In Fisherfaces setting [4], the gallery is constructed by selecting two samples per class at least. For LBP algorithm [7], the combination of $LBP_{9,1}^{p^2}$ operator and window size of $10 \times 10$ is determined to achieve the best recognition performance. In CRC algorithm [8], the regularisation parameters $\lambda$ are set to 0.005, 0.005, and 0.003 for CMU, UWA and, PolyU databases, respectively. In [13], normalisation is done to alleviate the illumination variations with a circular $(8,1)$ neighbourhood LBP filter. A cuberoot of size $3 \times 3 \times 3 \lambda$ is formed to slide iteratively with the step of one pixel. Table 2 shows the average recognition rates of these algorithms on the three public databases.

In view of the comparative results from the tabulation, our proposed method can almost outperform all of those state-of-the-art algorithms for the three public databases. Specifically, four methods including 3D Gabor wavelets, AHISD, CHISD, and SANP all have lower accuracy rates than the five remaining ones for CMU and UWA HSFD, but for PolyU HSFD, other three methods (Fisherfaces, LBP, and CRC) generate lower effects than other six algorithms. Only two methods (band fusion+PLS and our proposed algorithm) could demonstrate the best evaluation qualities on all three HSFD. It is worth noting that our algorithm only considers the local spatial information, not yet extracting the spectral characteristics from the wide spectrum. Although the spatiotemporal spectral fusion brings more discriminative features in the band fusion algorithm, there is no significant improvement in recognising face images. And these two methods coincidently show a similar validation for face recognition. Particularly, regarded as the modified version of LBP, our algorithm has actually improved the recognition accuracies for
the three HSFD, especially developing the discriminative ability in PolyU HSFD. Figure 8 shows the average recognition rate versus eigenface dimension for these databases. We validate the number of eigenface dimensions from 20 to 200 (300) in all our comparative analyses. Therefore, in our experiments, we fixed the number of eigenface dimensions to 270, 120, and 150 for CMU, UWA, and PolyU databases, respectively. In the recognition analysis, we performed a comparison between person and face recognition (see Figure 8). Note that person recognition denotes the subject classification as the whole face cube is regarded as a subject to recognise different persons, and face recognition is to classify each face image. Observing from the experimental results, we find that in the three databases, the average accuracy rates in face recognition achieve poorer performances than in-person recognition across all range of eigenface dimensions. For person recognition, they can obtain the best recognition performances when the number of eigenface dimensions reaches 270, 120, and 150 for CMU, UWA, and PolyU databases, respectively. After arriving at some levels, the performances will be consistently maintained. Furthermore, the same variable regularities also clearly appear in face recognition. In view of actual variations such as pose, facial expression, and illumination, the performances achieved in PolyU HSFD are a bit worse than in both CUM and UWA HSFD for all these algorithms. This may be mainly because that comparison with other databases PolyU contains more appearance variations of the subjects including hairstyle changes and skin conditions, and therefore has a much lower SNR due to the abundant mixture of noises such as inter-band misalignment, illumination, and pose, and so forth. Although these variations can influence the recognition performance, our proposed method still outperforms other algorithms in the three available public databases.

7 | DISCUSSION AND CONCLUSION

In this study, a novel feature formulation is proposed to combine the image set-based CRC classification method for hyperspectral face recognition. A joint representation of face image is defined by the fusion of variance-based LDTP into the oriented gradient information. Unlike the feature coding of a single image, our method for facial feature extraction choose the local variance as a variable indicator to characterise the spatial structures, initially combining the properties of LDTP and oriented gradient. Meanwhile, CRC has the advantage of sparsely representing the probe image from the set samples of the respective set. It has achieved a promising performance in face recognition. Therefore, our method is further introduced to optimise the problem of hyperspectral face recognition.

A thorough experimental evaluation is conducted on three benchmark datasets for face recognition, and the results are compared to eight existing state-of-the-art algorithms. Experimental results demonstrate that our proposed method can achieve a more robust recognition performance for the three public HSFD, almost outperforming all those algorithms. Especially for PolyU HSFD that contains more appearance variations of the subjects and also has much lower SNR than the other two databases, our method can mostly improve the recognition performances, showing great robustness to subject movements or expression variations that are unavoidable. It shows a more robust capability in recognising the noisy datasets. Meanwhile, our method also demonstrates better classification effects for CMU and UMA-HSFD. Especially, two methods (i.e. band fusion+PLS and our proposed method) can achieve approximately equal recognition performances on each of the HSFD. But it is worth noting that our method only considers the local spatial information, not yet exploiting the spectral characteristics. It manifests that for spatiotemporal fusion, the discriminative feature with PLS cannot bring more significant improvement in recognising face than our method. That is, our method can use fewer feature information to obtain more excellent effects. Furthermore, compared with LBP, our algorithm has actually also improved the recognition accuracies for the three HSFD, especially developing the discriminative ability in PolyU HSFD. To recap, although the process is influenced by the existing abundant mixture of noises such as inter-band misalignment, illumination, and pose, our proposed method can still consistently achieve the promising performance across all experiments only using the spatial information, while the performances of other methods fluctuate even with tuned parameters on different datasets/features. Therefore, the experimental results reveal that our method has the advantages of better characterising the LDTP and noise reduction due to the introduction of the variance-based spatial gradient patterns. Simple LBP does not have these advantages.

To sum up, our main contribution is proposing a novel hyperspectral face recognition algorithm based on LDTP for feature formulation and CRC for classification. The proposed algorithm was tested on three standard databases and compared
FIGURE 8  Recognition analysis: Average accuracy rate versus eigenface dimension respectively on (a) CMU, (b) UWA, and (c) PolyU HSFD. Each illustration is evaluated by performing person and face recognition, respectively.
with eight existing state-of-the-art algorithms. This is the first time such a comprehensive formulation of combining LDTP and oriented gradient for hyperspectral face recognition has been performed. Our algorithm outperformed existing state-of-the-art algorithms on the three standard databases indicating that our algorithm is a promising approach for hyperspectral face recognition.

Hyperspectral face recognition is a promising but challenging classification problem. The technique has penetrated deep into many fields such as remote sensing and computer vision. These applications motivate the researchers to develop a more robust evaluation system. Our proposed method can mostly improve the recognition performances by introducing a new LDTP indicator, showing its promising capability. Further research is extended to bind the spectral property into the spatial local information, considering the comprehensive spatiotemporal information fusion. Another consideration is to perform band selection. It aims at representing a face with even sparser approximation where a probe face can be the linear combination of the sample images with the principal component bands. The authors may continue to work on other classification methods such as deep convolution neural networks for hyperspectral face recognition.

ACKNOWLEDGEMENTS

This research is supported by the National Natural Science Foundation of China (Grant No.61872301), Chongqing Key Laboratory of Non-linear Circuit and Intelligent Information Processing and Fundamental Research Funds for the Central Universities (XDJK2019C021). The authors are grateful to their colleagues at the Institute of Signal and Information Processing at Southwest University for their thoughtful suggestions. M. Hao would like to appreciate the helpful comments and reviews of Professor G. Liu in particular. M. Hao also acknowledges the help of D. Xie for his technical suggestions and feedback in the early versions of this manuscript. Last but not least, we would like to thank the anonymous reviewers for their constructive comments and suggestions.

ORCID

Min Hao
https://orcid.org/0000-0001-5696-7745

REFERENCES

1. Shen, L., Zheng, S.: Hyperspectral face recognition using 3D Gabor wavelets. In: Proceedings of the 21st International Conference on Pattern Recognition (ICPR2012), Tsukuba, Japan, pp. 1574–1577 (2012)
2. Cevikalp, H., Triggs, B: Face recognition based on image sets. In: CVPR 2010–IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, USA, pp. 2567–2573 (2010)
3. Hu, Y., Mian, A.S., Owens, R.: Face recognition using sparse approximated nearest points between image sets. IEEE Trans. Pattern Anal. Mach. Intell. 34(10), 1992–2004 (2012)
4. Belhumeur, P.N., Hespanha, J., Kriegman, D.J.: Eigenfaces vs. Fisherafaces: Recognition using class specific linear projection. In: Proceedings of the 4th European Conference on Computer Vision, Cambridge, UK, pp. 45–58 (1996)
5. Ren, J., et al.: Effective feature extraction and data reduction with hyperspectral imaging in remote sensing. IEEE Signal Process Mag. 31(4), 149–154 (2014)
6. Kwon, H., Tai, Y-W.: RGB-guided hyperspectral image upsampling. In: 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, pp. 307–315 (2015)
7. Ahonen, T., Hadid, A., Pietikainen, M.: Face recognition with local binary patterns. In: 8th European Conference on Computer Vision, Prague, Czech Republic, pp. 460–481 (2004)
8. Zhang, L., Yang, M., Feng, X.: Sparse representation or collaborative representation: Which helps face recognition?. In: IEEE International Conference on Computer Vision, ICCV 2011, Barcelona, Spain, pp. 471–478 (2011)
9. Pan, Z., et al.: Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Mach. Intell. 25(12), 1552–1560 (2003)
10. Pan, Z.H., et al.: Face recognition in hyperspectral images. IEEE Trans. Pattern Anal. Mach. Intell. 25(12), 1552–1560 (2003)
11. Xie, Z.H., et al.: Fusion of LBP and SWLD using spatio-spectral information for hyperspectral face recognition. Proc. SPIE 10620, 1062016.1–1062016.6 (2018)
12. Zhao, G.Y., Pietikainen, M.: Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans. Pattern Anal. Mach. Intell. 29(6), 915–928 (2007)
13. Uzair, M., Mahmood, A., Mian, A.: Hyperspectral face recognition with spatiotemporal information fusion and PLS regression. IEEE Trans. Image Process. 24(3), 1127–1137 (2015)
14. Chen, G., Li, C., Sun, W.: Hyperspectral face recognition via feature extraction and CRC-based classifier. IET Image Proc. 11(4), 266–272 (2016)
15. Déniz, O., et al.: Face recognition using histograms of oriented gradients. Pattern Recognit. Lett. 32(12), 1598–1603 (2011)
16. Ren, C., et al.: Band-reweighed Gabor kernel embedding for face image representation and recognition. IEEE Trans. Image Process. 23(2), 725–740 (2014)
17. Zhao, H., Sun, S.: Sparse tensor embedding based multispectral face recognition. Neurocomputing 133, 427–436 (2014)
18. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (2004)
19. Cheung, W., Hamarneh, G.: N-SIFT: N-dimensional scale invariant feature transform. IEEE Trans. Image Process. 18(9), 2012–2021 (2009)
20. Li, Y.S., et al.: Extreme-constrained spatial-spectral corner detector for image-level hyperspectral image classification. Pattern Recognit. Lett. 109, 110–119 (2018)
21. Cho, W., et al.: Hyperspectral face recognition using improved inter-channel alignment based on qualitative prediction models. Opt. Express 24(24), 27637 (2016)
22. Wang, J., Chang, C.I.: Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis. IEEE Trans. Geosci. Remote Sens. 44(6), 1586–1600 (2006)
23. Wei, D., et al.: Studies on hyperspectral face recognition in visible spectrum with feature band selection. IEEE Trans. Syst. Man Cybern. Part A Syst. Humans 40(6), 1354–1361 (2010)
24. Hu, Y., Mian, A.S., Owens, R.: Sparse approximated nearest points for image set classification. In: CVPR 2011, Colorado Springs, CO, USA, pp. 121–128 (2011)
25. Fan, Z.Z., et al.: Virtual dictionary based kernel sparse representation for face recognition. Pattern Recognit. 76, 1–13 (2018)
26. Qiao, T., et al.: Joint bilateral filtering and spectral similarity-based sparse representation: A generic framework for effective feature extraction and data classification in hyperspectral imaging. Pattern Recognit. 77, 316–328 (2018)
27. Kim, T.K., Kittler, J., Cipolla, R.: Discriminative learning and recognition of image set classes using canonical correlations. IEEE Trans. Pattern Anal. Mach. Intell. 29(6), 1005–1018 (2007)
28. Wang, R., et al.: Manifold-manifold distance with application to face recognition based on image set. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition, Anchorage, AK, USA, pp. 1–8 (2008)
29. Nguyen, C.H., Artemiadis, P.: EEG feature descriptors and discriminant analysis under Riemannian Manifold perspective. Neurocomputing 275, 1871–1883 (2018)
30. Denes, L.J., Metes, P., Liu, Y.: Hyperspectral Face Database. Carnegie Mellon University, The Robotics Institute, Pittsburgh, Pennsylvania (2002)
31. Di, W., et al.: Studies on hyperspectral face recognition in visible spectrum with feature band selection. IEEE Trans. Syst. Man Cybern. Part A Syst. Humans 40(6), 1354–1361 (2010)
32. Ning, X., et al.: BULDSP: Biomimetic uncorrelated locality discriminant projection for feature extraction in face recognition. IEEE Trans. Image Process. 27(5), 2575–2586 (2018)
33. Fan, J., Arandjelović, O.: Employing domain specific discriminative information to address inherent limitations of the LBP descriptor in face recognition. In: 2018 International Joint Conference on Neural Networks (IJCNN), Rio de Janeiro, Brazil, pp. 1–7 (2018)
34. Jabid, T., Kabir, M.H., Chae, O.: Local directional pattern (LDP) for face recognition. Int. J. Innovative Comput. Inf. Control 8(4), 2423–2437 (2012)
35. Tong, Y., Chen, R.: Local dominant directional symmetrical coding patterns for facial expression recognition. Comput. Intell. Neurosci. 2019, 1–13 (2019)
36. Ghiass, R.S., et al.: Vesselness features and the inverse compositional AAM for robust face recognition in thermal IR. In: Proceedings of the Twenty-Seventh AAAI Conference on Artificial Intelligence, Bellevue, Washington, USA, pp. 357–364 (2013)
37. Ghiass, R.S., et al.: Illumination-invariant face recognition from a single image across extreme pose using a dual dimension AAM ensemble in the thermal infrared spectrum. In: The 2013 International Joint Conference on Neural Networks (IJCNN), Dallas, TX, USA (2013)
38. Chang, H., et al.: Improving face recognition via narrowband spectral range selection using Jeffrey divergence. IEEE Trans. Inf. Forensics Secur. 4(1), 111–122 (2009)
39. Uzair, M., Mahmood, A., Mian, A.S.: Hyperspectral face recognition using 3D-DCT and partial least squares. In: Proceedings of the British Machine Vision Conference 2013, Bristol, UK (2013)
40. PolyU-HSFD. http://www4.comp.polyu.edu.hk/~biometrics/hsi/hyper_face.htm (2015). Accessed 22 Jan. 2015
41. Dabov, K., et al.: Image denoising by sparse 3-D transform-domain collaborative filtering. IEEE Trans. Image Process. 16(8), 2080–2095 (2007)
42. Tanaka, M.: Matlab code for face parts detection. https://ww2.mathworks.cn/matlabcentral/fileexchange/36855-face-parts-detection (2016)
43. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: International Conference on Computer Vision & Pattern Recognit. (CVPR ’05) San Diego, CA, USA (2005)
44. Wright, J., et al.: Robust face recognition via sparse representation. IEEE Trans. Pattern Anal. Mach. Intell. 31(2), 210–227 (2009)
45. Chen, G., Sun, W., Xie, W.: Hyperspectral face recognition with log-polar Fourier features and collaborative representation based voting classifiers. IET Biom. 6(1), 36–42 (2017)
46. Li, W., et al.: Collaborative-representation-based nearest neighbor classifier for hyperspectral imagery. IEEE Geosci. Remote Sens. Lett. 12(2), 389–393 (2015)
47. Akhtar, N., Shafait, F., Mian, A., Efficient classification with sparsity augmented collaborative representation. Pattern Recognit. 65, 136–145 (2017)

How to cite this article: Hao M, Liu G, Xie D. Hyperspectral face recognition with a spatial information fusion for local dynamic texture patterns and collaborative representation classifier. IET Image Process. 2021;15:1617–1628.
https://doi.org/10.1049/ipr2.12131