

Rank Preserving Discriminant Analysis for Human Behavior Recognition on Wireless Sensor Networks

Dapeng Tao, Lianwen Jin, *Member, IEEE*, Yongfei Wang, and Xuelong Li, *Fellow, IEEE*

Abstract—With the rapid development of the intelligent sensing and the prompt growing industrial safety demands, human behavior recognition has received a great deal of attentions in industrial informatics. To deploy an utmost scalable, flexible, and robust human behavior recognition system, we need both innovative sensing electronics and suitable intelligence algorithms. Wireless sensor networks (WSNs) open a novel way for human behavior recognition, because the heavy computation can be immediately transferred to a network server. In this paper, a new scheme for human behavior recognition on WSNs is proposed, which transmits activities' signals compressed by Hamming compressed sensing to the network server and conducts behavior recognition through a collaboration between a new dimension reduction algorithm termed rank preserving discriminant analysis (RPDA) and a nearest neighbor classifier. RPDA encodes local rank information of within-class samples and discriminative information of the between-class under the framework of Patch Alignment Framework. Experiments are conducted on the SCUT Naturalistic 3D Acceleration-based Activity (SCUT NAA) dataset and demonstrate the effectiveness of RPDA for human behavior recognition.

Index Terms—Discriminant analysis, human behavior recognition, rank preserving, wireless sensor networks (WSNs).

I. INTRODUCTION

HUMAN behavior recognition is a complex issue, spans many disciplines, and receives intensive attentions in industrial informatics. The basic steps involve sensing signal acquisition, information processing and pattern classification. The

Manuscript received June 14, 2012; revised September 09, 2012, December 05, 2012, and February 17, 2013; accepted March 15, 2013. Date of publication March 28, 2013; date of current version December 12, 2013. This work was supported in part by the National Basic Research Program of China (973 Program) under Grant 2012CB316400, the National Natural Science Foundation of China under Grants 61075021, Grant 61201348, Grant 61125106, Grant 91120302, and Grant 61072093, the National Science and Technology Support Plan under Grant 2013BAH65F01 and Grant 2013BAH65F04, the Guangdong Natural Science Funds S2011020000541, the Guangdong Scientific and Technology Research Plan 2012A010701001, the Fundamental Research Funds for the Central Universities of China under Grants 2012ZP0002 and Grant D2116320, and Shaanxi Key Innovation Team of Science and Technology under Grant 2012KCT-04. Paper no. TII-12-0412.

D. Tao, L. Jin, and Y. Wang are with the School of Electronic and Information Engineering, South China University of Technology, GuangZhou 510640, China (e-mail: dapeng.tao@gmail.com; lianwen.jin@gmail.com; yongfei.wang.ee@gmail.com).

X. Li is with the Center for OPTical IMagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an 710119, China (e-mail: xuelong_li@opt.ac.cn).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TII.2013.2255061

development of sensing electronics and intelligent algorithms results in innovations, efficiencies, and cost savings in many areas [80]. In recent years, a dozen of effective methods have been proposed to automatically recognize human behavior and benefits the industrial informatics. We can simply group these methods into two categories: computer vision-based and accelerometer-based systems.

Many computer vision-based human behavior analysis systems have been developed over the past decades [13], [27], [28], [42], [46], [52], [55], [56], [70], [75]. This kind of systems can be accomplished by the several important steps [3], [18], which are object detection, object segmentation, feature extraction and classification. Representative works are listed below. Thureau [62] utilized histogram of oriented gradient (HOG) descriptors to represent actions. Kellokumpu *et al.* [29] introduced the dynamic local binary pattern (LBP) descriptor. Shao *et al.* [56] compared representative feature descriptors, such as HOG-HOF [41] and LBP-TOP [76]. Niebles *et al.* [51] proposed an unsupervised learning based topic model for human activity classification. In contrast to unsupervised learning algorithms, supervised learning, such as support vector machines (SVM) [21], [40], [64], conditional random fields (CRFs) [58] and Adaboost [19], can effectively exploit the label information to improve accuracy of recognition. Bian *et al.* [7] proposed the transfer topic model (TTM) to solve the problem of the amount of training samples is insufficient. Vision-based human behavior analysis systems cannot perform well for industrial environment, because these systems are sensitive to lighting conditions.

Accelerometer-based physical activity recognition [24], [32], [43], [53] is an important and exciting alternative, which has been receiving increasing attention in recent years in industrial informatics. It investigates the use of acceleration signals from the accelerometer attached to the human body so as to analyze and classify daily physical activities such as walking, running, and standing. The progress in physical activity recognition is also meaningful to many disciplines and applications such as health monitoring, context-awareness, and smart surveillance. In terms of accelerometer-based physical activity recognition, it has to be decided how many accelerometers should be placed on a human body.

In earlier studies, several accelerometers were attached to a human body, which made the tester feel uncomfortable [32]. To improve this, researchers tried to recognize the physical activities with only one accelerometer [24], [53]. It is worth noting that there are three commonly used features in accelerometer-based human behavior recognition, including fast Fourier transform (FFT) coefficients [6], DCT coefficients [72], and time-domain feature [6]. It has been demonstrated that, with FFT

features, a satisfactory recognition rate can be achieved in accelerometer-based physical activity recognition. This result may derive from the fact that people are performing some regular movements when they are walking, running, and jumping. FFT coefficients can capture the frequency information of the cyclic movements, and thus the discriminative information of the activities is contained in the FFT coefficients.

Multifunctional wireless sensors [2], [9], [31], [36], [39], [47], [48], [60], [65], which have achieved a huge success, boost the concept of wireless sensor networks (WSNs) and attract researchers who work on human behavior recognition for industrial informatics. For example, Ghasemzadeh *et al.* [22] presented a scheme of classification model to recognize human activities by utilizing body-worn inertial sensor networks. A large-scale WSNs based on ZigBee protocol, is widely used in industrial monitoring [50]. It is a low-consumption device and has the capability to sense the variation information (e.g., temperature, pressure, and revolution speed) in a reliable way. WSN-based ZigBee protocol can be used in the application of human behavior recognition, because of the following reasons: 1) it has low power consumption, and thus we can control the size of body-worn sensing module; 2) the sensing modules build network flexibly and rapidly, since the protocol brings about the nature of self-organization and self-configuration; and 3) it has a relatively low cost of the sensing module.

However, increasing network size causes a wide range of issues, e.g., data transmission cost and network lifetime [9], [12]. Furthermore, given limited computational resource in WSNs, it is impossible to accomplish human behavior recognition directly. The success of compressed sensing [10], [17] opens a door to effectively and efficiently compress physical information, and thus we consider transmitting compressed data to the network server where the behavior features will be extracted and the activities will be recognized. However, classical compressive sensing algorithms require polynomial time [11] for signal reconstruction. It is substantially expensive for WSNs. Hamming compressed sensing (HCS) [79] is more suitable for WSNs, it can fast recover a digit signal from the quantization of its few measurements, because it has linear recovery time.

After obtaining the digital signals, we need intelligent algorithms for classification [26], [71]. Since the digital signal is of high dimensionality, we prefer dimension reduction for signal preprocessing to avoid the over-fitting problem. The dimension reduction results in a succinct yet effective representation of a sample in the original high-dimensional space. Many effective dimension reduction methods have been proposed over the past few decades [5], [20], [23], [59], [63]. In this paper, we introduce the rank order information to improve discriminant learning for human behavior recognition and present a new dimension reduction scheme termed rank preserving discriminant analysis (RPDA).

Based on the above descriptions, precisely recognizing human behaviors captured by one 3-D accelerometer becomes to be reality through the following steps: 1) utilizing HCS to compress the accelerometer signal, and then transmitting the compressed data to the network server via WSNs; 2) utilizing HCS to decode the compressed accelerometer signal in the net-

work server; 3) training the RPDA projection matrix by using a small number of labeled samples; and 4) classifying the RPDA projected samples by using the nearest neighbor classifier, and 5) returning the recognition results to the wireless sensor.

Fig. 1 shows the architecture of the proposed human behavior recognition on wireless sensor networks. To collect activity data, the client terminal can be comprised primarily by an accelerometer, microprocessor and ZigBee transceiver module. The bandwidth of ZigBee is suit for the data of 3-D acceleration-based activity compressed by Hamming compressed sensing (HCS). The main computing process of the human behavior recognition is actually provided by Network server, including HCS decoding, training RPDA projection matrix by using a small number of labeled samples stored in Network server, and classifying the RPDA projected samples by using the nearest neighbor classifier.

The main contribution of this paper is the newly developed RPDA for human behavior recognition. HCS is a general information compression technique, which helps to reduce the time delay of transmission and is not tied to the proposed RPDA. Given the limited page length, we will not detail the other parts which are easy to implement based on the references cited in this paper.

The remainder of this paper is organized as follows. In Section II, related works on dimension reduction are reviewed which are important for human behavior recognition and experiment section. We detail the newly proposed Rank Preserving Discriminant Analysis in Section III. Section IV shows the experimental results on the SCUT Naturalistic 3-D Acceleration-based Activity (SCUT NAA) dataset [72]. Finally, concluding remarks and suggestions for future work are presented in Section V.

II. RELATED WORK

Dimension reduction algorithms can be simply grouped into two categories: unsupervised and supervised learning algorithms. In unsupervised learning, the class label information is unavailable. Principal component analysis (PCA) [25], [35] is the most representative unsupervised dimension reduction algorithm. It provides a roadmap for efficient relevant information extracting from the high-dimensional data space by reconstructing Gaussian distributed data through maximizing the trace of the total scatter. This simple, unsupervised and parametric method has helped the researches in various fields from neuroscience to computer vision, such as, image analysis, facial expression recognition [57], [73], and data compression [69], because of its simplicity.

Laplacian eigenmaps (LEs) [5], which is a geometrically motivated algorithm, not only avoid the limitation of the number of projection vectors naturally, but also consider the nonlinearity of the data distribution. LE preserves the local geometry information in order to uncover the intrinsic geometrical structure of the original high-dimensional data by building a graph model which encodes the neighborhood information. Therefore, LE is an efficient nonlinear dimension reduction method. Locality preserving projections (LPP) [23] is the linear approximation of LE. Though LPP is a linear dimension reduction algorithm,

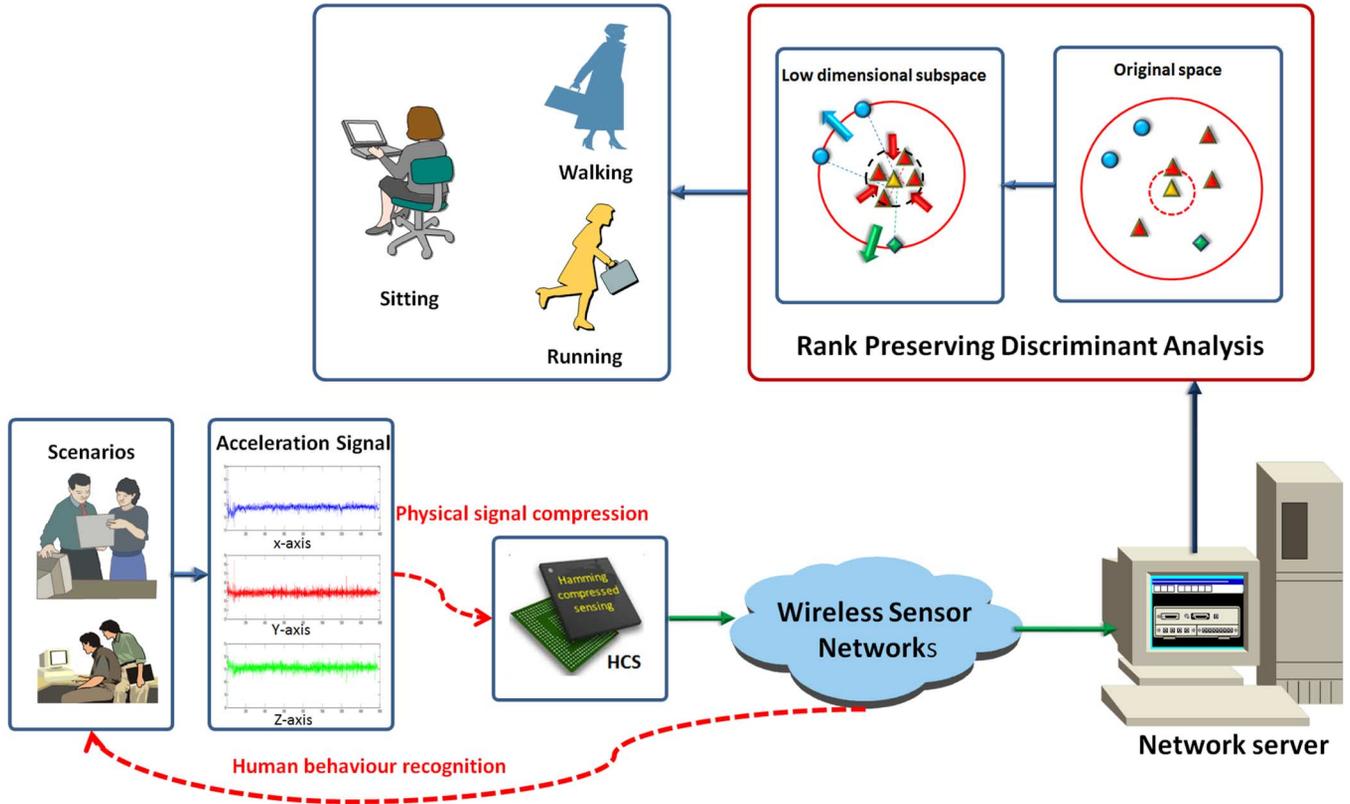


Fig. 1. Human behavior recognition on WSNs. This scheme contains the following five components: 1) using HCS to compress the acceleration signal and transmit the compressed signal to network sever via WSNs; 2) using HCS to decode the compressed signal on the server; 3) training Rank Preserving Discriminant Analysis projection matrix by using a small number of labeled samples; 4) classifying the RPDA projected samples by using the nearest neighbor classifier; and 5) returning the recognition results to the wireless sensor.

it approximates the nonlinear problem properly by partially preserving the local geometry information.

In supervised learning, a training sample consists of an input instance and the associated class label. Linear discriminant analysis (LDA) [20] is the most representative supervised dimension reduction algorithm. It seeks to find a projection direction which minimizes the trace of the within-class scatter matrix and maximizes the trace of the between-class scatter matrix, simultaneously. In general, LDA performs excellently under the circumstances that different classes have an equal within-class scatter. Nevertheless, LDA suffers from the two main drawbacks. First, it is a globally linear dimension reduction algorithm and fails to discover the nonlinear structure hidden in the high-dimensional space. Second, to use LDA, small sample size (SSS) is a big problem [61]. LDA needs a large number of training samples to acquire a good model. A dozen of algorithms have been proposed to deal with the SSS problem of LDA, such as PCA plus LDA [4] and direct LDA (DLDA) [74]. However, they fail to consider the local geometry of within-class samples and the discriminative information in selected subspace.

Although existing dimension reduction algorithms [49] have been applied to human behavior recognition, there is still room to improve the classification precision. Recently, it has been observed that the Euclidean metric suffers from the concentration of measure phenomenon [16], since the difference of distances between pairs of high-dimensional samples are fairly indistinguishable. Extensive numerical experiments [15], [30] verified

the ranking of neighbors is important. Transferring distance information to rank orders benefits to recover the intrinsic data properties.

Nonmetric multidimensional scaling (MDS) [78] aims to preserve rank order information by matching distances among all data in the low-dimensional space with distances among all data in the original high-dimensional space. In addition, data-driven high-dimensional scaling (DD-HDS) [45] was presented to improve the performance of the representation of high-dimensional data. Recently, Lespinats *et al.* [44] proposed RankVisu to preserve small dissimilarities as possible, since small rank orders are more important.

In contrast to the classical spectral analysis-based dimension reduction, we introduce rank order information to human behavior recognition and present a new dimension reduction algorithm, RPDA. It differs from the aforementioned dimension reduction in considering the influence of rank order information in within-class and between-class. By introducing a penalized factor of distances that takes the concentration of measure phenomenon [16] into account, it preserves as much as possible the local rank order information of the within-class formed local patch. In addition, it is remarkable that the process of dimension reduction always companies variations in the original distribution. Therefore, we model the process which extracts the local discriminative information of the between-class by intentionally ignoring the rank order information. In order to understand RPDA, we fabricate the local rank information of within-class

samples and the discriminative information of the between-class under the framework of patch alignment framework (PAF) [78].

III. RPDA

Here, we present a new supervised dimension reduction algorithm, RPDA. The concentration of measure phenomenon [16] has significant impact on the performance of dimension reduction tools, because the differences of distances between pairs of high-dimensional samples are fairly indistinguishable [1]. A direct solution is to preserve the rank order information of within-class samples [15] in the process of transforming samples from the high-dimensional space to a low-dimensional subspace.

It is insufficient to consider only the rank order information, because the process of dimension reduction always companies variations in the original distribution [14]. Therefore, we design a discriminant information extracted way that ignores the between-class samples rank order information. This strategy is feasible for high-dimensional data to selectively shrink or stretch a suitable manifold. Similar to other spectral analysis abased dimension reduction algorithms, it can be built under Patch Alignment Framework (PAF) [78], because PAF offers a platform to manipulate the local rank information of within-class samples and the discriminative information of the between-class samples. Under PAF, all these algorithms can be reasonably divided into part optimization and whole alignment two stages.

The training samples of the discriminative dimension reduction possess class labels. Given a training set in the high-dimensional space R^D , i.e., $X = [x_1, x_2, \dots, x_N] \in R^{D \times N}$, and each sample has the corresponding class label $C_i \in Z^n$. The objective is to find a projection matrix $U \in R^{D \times d}$ to linearly map samples from the high-dimensional space R^D to a low-dimensional subspace R^d , with $d < D$, i.e., $Y = U^T X = [y_1, y_2, \dots, y_N] \in R^{d \times N}$.

A. Part Optimization for RPDA

For a given labeled sample x_i , we can find its k_1 closest within-class samples $x_{i^1}, \dots, x_{i^{k_1}}$ and k_2 closest between-class samples $x_{i_1}, \dots, x_{i_{k_2}}$ to form a local patch $X_i = [x_i, x_{i^1}, \dots, x_{i^{k_1}}, x_{i_1}, \dots, x_{i_{k_2}}] \in R^{D \times (k_1 + k_2 + 1)}$. In addition, we define that R_{ij} is the rank of the sample j for the sample i . By using RPDA, our objective is to achieve a new low-dimensional representation $Y_i = [y_i, y_{i^1}, \dots, y_{i^{k_1}}, y_{i_1}, \dots, y_{i_{k_2}}] \in R^{d \times (k_1 + k_2 + 1)}$ for each local patch, where the between-class distances will be as large as possible and the within-class rank order information will be preserved as much as possible.

Fig. 2 illustrates the process of part optimization in the situation when $k_1 = 2$ and $k_2 = 3$. It shows that in the projected subspace (right-hand side of Fig. 2), the intra-class rank order information of sample y_i (i.e., the rank of red triangles from the yellow triangle) is preserved whereas the distances between y_i and the samples from other classes (blue circle and green square) are large.

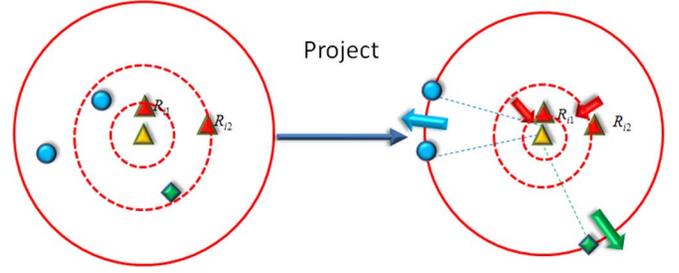


Fig. 2. Process of part optimization. The yellow triangle is a labeled sample. The red triangles are $k_1 = 2$ closest within-class samples. Blue circles and green squares are the $k_2 = 3$ closest between-class samples.

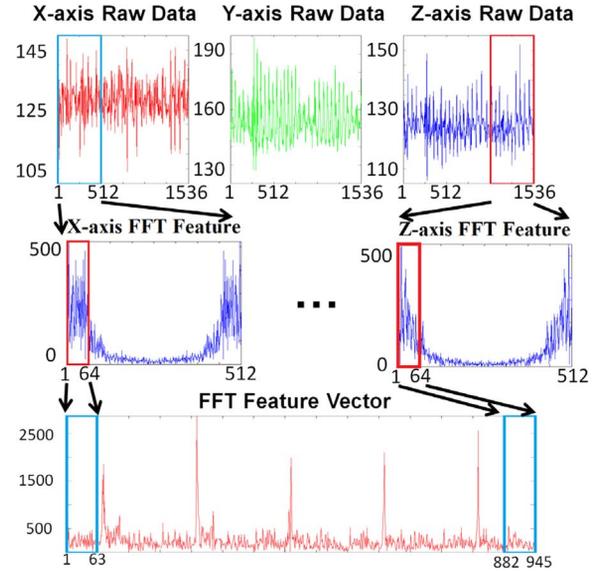


Fig. 3. Procedure of features extraction. The first row represents the raw data collected from the accelerometer. The x -axis contains values described sampling points and the sampling interval is 0.01 s. The y -axis contains values described acceleration attribute. For example, $y = 125$ means gravity = 0 g, and each 26 units means 1 g, where $g = 9.8 \text{ m/s}^2$. The second row is FFT feature of 512 points from the raw data. The third row is 945 dimensional FFT feature.

For each local patch in the low-dimensional subspace, we maximize the margin that is the sum of the distances between y_i and k_2 samples between-class as

$$M(y_i) = \sum_{p=1}^{k_2} \|y_i - y_{i_p}\|^2. \quad (1)$$

By using a rank matrix, we can easily obtain the rank order information. However, the rank matrix R , in which the entry R_{ij} is the rank of the sample j with respect to the sample i , is generally not symmetric. This is because the sample x_j is the k th nearest neighbor of x_i , but not vice versa. Therefore, it cannot be directly used in PAF. In contrast, the distance matrix of the local patch is symmetric. Therefore, we introduce a penalized factor in the distance matrix to well solve the problem arose by the concentration of measure phenomenon, and we have

$$R(y_i) = \sum_{j=1}^{k_1} \|y_i - y_{i_j}\|^2 (w_i)_j \quad (2)$$

where $(w_i)_j$ is the penalized factor to emphasize the distinction between small and large distances in the original distribution with a large and a small weighting, respectively. As expected, the small distances in original space will cause more heavily penalization in low-dimensional subspace.

Motivated by the success of LE [5] and the effective neighborhood relations preservation by heat kernel, we propose a penalized factor as

$$(w_i)_j = \begin{cases} \exp\left(\frac{-\|x_i - x_{ij}\|^2}{t}\right), & \text{if } x_{ij} \in N_{k_1}(x_i) \\ 0, & \text{otherwise} \end{cases} \quad (3)$$

Since the local patch X_i can be regarded as approximately linear [54], we can encode both the rank order information and the discriminative information to achieve the entire objective function of part optimization

$$\arg \min_{y_i} \left(\sum_{j=1}^{k_1} \|y_i - y_{ij}\|^2 (w_i)_j - \gamma \sum_{p=1}^{k_2} \|y_i - y_{ip}\|^2 \right) \quad (4)$$

by combining (1) and (2) via a tradeoff parameter γ where $\gamma \in [0, 1]$ is a tradeoff parameter to integrate the contributions of intra-class samples and those of the between-class samples in the part optimization stage.

We further deduce (4) to

$$\arg \min_{y_i} \sum_{j=1}^{k_1} \|y_i - y_{ij}\|^2 (w_i)_j - \gamma \sum_{p=1}^{k_2} \|y_i - y_{ip}\|^2 = \arg \min_{Y_i} \text{tr}(Y_i L_i Y_i^T) \quad (5)$$

where $\text{tr}(\cdot)$ is the trace operator, $L_i = \begin{bmatrix} -e_{k_1+k_2}^T \\ I_{k_1+k_2} \end{bmatrix} \text{diag}(v_i) \begin{bmatrix} -e_{k_1+k_2} \\ I_{k_1+k_2} \end{bmatrix}$,

$e_{k_1+k_2} = [1, \dots, 1]^T \in R^{k_1+k_2}$, $I_{k_1+k_2} = \text{diag}(1, \dots, 1)$, $Y_i = \begin{bmatrix} y_i, y_{i^1}, \dots, y_{i^{k_1}}, y_{i_1}, \dots, y_{i_{k_2}} \end{bmatrix}$, and

$v_i = \underbrace{[(w_i)_1, \dots, (w_i)_{k_1}]}_{k_1}, \underbrace{[-\gamma_{k_1+1}, \dots, -\gamma_{k_2}]}_{k_2}$.

B. Whole Alignment for RPDA

By using a sample selection matrix, the coordinate of the low-dimensional representation Y_i is selected from the globe coordinates $Y = U^T X = [y_1, y_2, \dots, y_N] \in R^{d \times N}$, i.e.,

$$Y_i = Y S_i \quad (6)$$

where $S_i \in R^{N \times (K+1)}$ is the selection matrix. Let $F_i = \{i, i_1, \dots, i_K\}$ be the index set, and then each entry of the selection matrix is defined by

$$(S_i)_{pq} = \begin{cases} 1, & \text{if } p = F_i\{q\} \\ 0, & \text{else.} \end{cases} \quad (7)$$

According to (6), the part optimization (5) can be rewritten as

$$\arg \min_Y \text{tr}(Y S_i L_i S_i^T Y^T). \quad (8)$$

We have the whole alignment by summing over all the part optimizations defined in (8) to achieve

$$\arg \min_Y \sum_{i=1}^N \text{tr}(Y S_i L_i S_i^T Y^T) = \arg \min_Y \text{tr}(Y L Y^T). \quad (9)$$

Considering simplicity, $U^T U = I_d$ is imposed on (9) to uniquely determine the orthogonal projection matrix U according to $Y = U^T X$, where I_d is an identity matrix of size $d \times d$. Thus, (9) is transformed to

$$\begin{aligned} & \arg \min_Y \text{tr}(U^T X L X^T U) \\ \text{s.t. } & U^T U = I_d. \end{aligned} \quad (10)$$

By utilizing the Lagrange's multiplier method [78], we transform (10) to a generalized eigenvalue problem. The projection matrix U is given by d eigenvectors associated with d smallest eigenvalues of $X L X^T$. In addition, PCA is recommended to be applied to the original high-dimensional data for removing the noise.

C. Alternative Penalized Factors

As a penalized factor, (3) is very flexible. In general, the selection of parameter t is in directly proportional to the average distances of pairwise within-class samples. To avoid the selection of parameter t , we can construct a penalized factor as

$$(w_i)_j = \begin{cases} \frac{x_i^T x_{ij}}{\|x_i\| \cdot \|x_{ij}\|}, & \text{if } x_{ij} \in N_{k_1}(x_i) \\ 0, & \text{otherwise.} \end{cases} \quad (11)$$

This form is the cosine similarity of the two samples and very classical in information retrieval [33].

In addition, we utilize a sigmoid-like weighting function [45] as an alternative to improve the performance of rank preserving:

$$(w_i)_j = \begin{cases} 1 - \int_{-\infty}^{\|x_i - x_{ij}\|} f(u|\mu, \sigma) du, & \text{if } x_{ij} \in N_{k_1}(x_i) \\ 0, & \text{otherwise} \end{cases} \quad (12)$$

where $f(u|\mu, \sigma)$ is a Gaussian variable probability density function parametrized in terms of a mean μ and a standard deviation σ .

Similar to DD-HDS [45], the mean μ and the standard deviation σ can be estimated by

$$\mu = \frac{\text{mean}_{1 \leq i < j \leq N}(d_{ij}) - 2(1 - \lambda) \text{std}_{1 \leq i < j \leq N}(d_{ij})}{2} \quad (13)$$

$$\sigma = 2\lambda \text{std}_{1 \leq i < j \leq N}(d_{ij}) \quad (14)$$

where we have imposed the mean and the std operation upon the distribution of distances between all pairwise samples in the high-dimensional space (d_{ij}). Thus, by modeling the difference of distances between small and large distances, the above weighting function preserves the intra-class rank order information. The parameter λ is a user-defined parameter and can be chosen in the range of $[0, 1]$. To be clear, the larger λ is selected and the effect of rank preserving will be more significant that

affects the performance of dimension reduction. In addition, the parameter λ can be selected by cross-validation. In real applications, the penalized factor defined in (11) is the most convenient scheme. We can avoid the annoying step of parameter selection. However, penalized factors defined in (3) and (12) are more flexible for rank preserving. Based on the above discussions, we summarize RPDA in Algorithm 1.

In RPDA, we first build part optimizations for all training samples by calculating the matrix L_i according to (5). Then, the L_i matrixes are summed into the whole alignment matrix L according to (9). In addition, to avoid the selection of parameter t , penalized factor equation (3) can be replaced by (11) or (12).

Algorithm 1: Rank Preserving Discriminant Analysis

Input: Training set $X = [x_1, x_2, \dots, x_N] \in R^{D \times N}$; Class label $C_i \in Z^n$ d : dimension of the reduced space.

Output: Orthogonal projection matrix $U = [u_1, u_2, \dots, u_d] \in R^{D \times d}$

- Step 1) (optional) Use PCA projection matrix U_{PCA} to reconstruct the original training set X ;
- Step 2) Part optimization: Construct N patches for the training set according to (1) and (2), and calculate the matrix L_i for each patch by using (5);
- Step 3) Whole alignment: Sum all the patches in a global coordinate over all samples, and compute the whole alignment objective function (9);
- Step 4) Compute the projection matrix U_{RPDA} whose column vectors are the d eigenvectors of XLX^T associated with d smallest eigenvalues.
- Step 5) Return the final projection matrix
 $U = U_{PCA}U_{RPDA}$.
-

D. Time Complexity Analysis

Given N training samples in a D dimensional space, the time complexity of RPDA consists of two parts, including the whole alignment matrix L calculation and the eigenvalue calculation. The first part is $O((D + k_1 + k_2) \times N^2)$, where k_1 is the number of closest with-class samples and k_2 is the number of closest between-class samples. When $k_1 + k_2 \ll D$, we have $O(D \times N^2)$. The second part is $O(d \times N^2)$, where d is dimension after dimension reduction. Thus, the whole time complexity of RPDA is $O((D + d) \times N^2)$.

IV. EXPERIMENTAL RESULTS

Here, the experiments of human behavior recognition are conducted on the SCUT Naturalistic 3-D Acceleration-based Activity (SCUT NAA) dataset [72] to demonstrate the performance of the proposed RPDA. It contains 1278 samples of ten object categories. Since HCS affects the subsequent dimension reduction algorithm, we set $m/n = 4$, where m is the number of 1-bit measurements, and n is the size of a sequence of sample. This setting performs well in our system. In this paper, one commonly used feature in accelerometer based activity

TABLE I
DEFINITION OF THE TEN KINDS OF BEHAVIOR

Classes	Definition
Relaxing	Sitting & doing nothing
Walking	Walking 50 m at normal speed
Walking quickly	Walking 50 m faster than normal speed
Walking backwards	Walking backwards for 50 m
Running	Jogging 100 meters
Step walking	Moving the feet alternately in the rhythm of a marching step without advancing
Jumping	Jumping for 45s without advancing
Upstairs	Ascending stairs
Downstairs	Descending stairs
Cycling	Cycling with a real bike

recognition, the FFT coefficients [6], was extracted. The performance is measured by using the average accuracy for each class and the confusion matrix between the ground truth label information and the most likely inferred label information. The confusion matrix can contribute to get a better understanding of where the approach is going wrong. Details of the experimental setup and baseline models are given below.

A. SCUT NAA Dataset

The SCUT NAA dataset [72] is the first publicly available 3-D acceleration-based human behavior dataset. By utilizing one tri-axial accelerometer located on a fixed position (three alternatives include the waist belt, the trousers pocket, and the shirt pocket), we collected 1278 samples from 44 subjects (34 males and 10 females) naturally. All 44 subjects were students at South China University of Technology. Note that these college students were enrolled from different cities in China, the average age and the variance are 21.2 and 0.7, respectively. In addition, ten types of human behaviors were selected. Table I shows that these human behaviors include a range of common activities involving industrial production, e.g., light intensity activities such as sitting, moderate intensity activities such as step walking, and vigorous activities such as jumping and running. Therefore, the SCUT NAA dataset is suitable for human behavior research in industrial informatics.

Note that the data acquisition subjects to some conditions and the numbers of samples in different class are different, e.g., the class of "cycling" has only 30 samples, since only 30 peoples can ride. Inspired by the leave-one-out cross validation, observations of a single subject from the original dataset were using as the test data, and the remaining samples are used for model training. The training set was used to learn the orthogonal projection matrix. The test set was used for performance evaluation. Therefore, the SCUT NAA dataset can be divided into 44 splits.

B. FFT Descriptor

FFT features were extracted from the raw acceleration data corresponding to each axis. The size of the window used in the process of feature extraction was 512 points, with 256 sample points overlapped between consecutive windows. For each sliding window, the first 64 FFT coefficients were retained, respectively, while the first coefficient corresponding to the current component was abandoned. The raw data corresponding

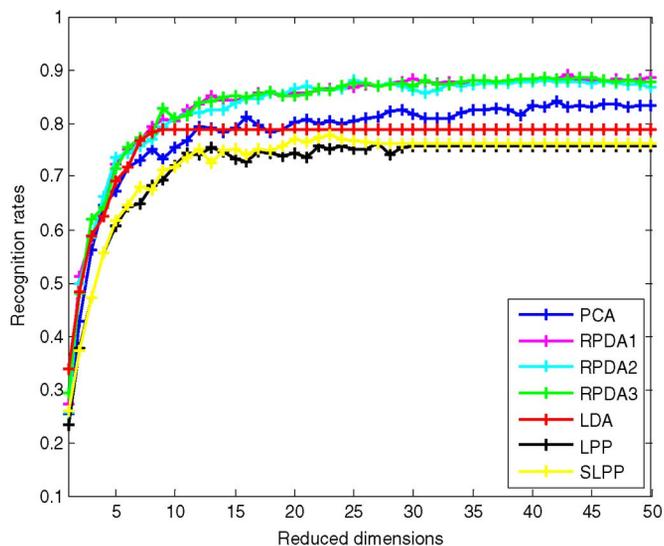


Fig. 4. Average recognition rate versus dimension reduction on the test sets.

to each activity contain 4096 sample points, which leads to a 945-dimensional FFT feature.

C. Baselines and Performance Measures

Here, the performance of RPDA was evaluated by comparing with four representative algorithms, including PCA, LDA, LPP and SLPP [8]. According the penalized factors defined in (3), (11), and (12), we named the proposed algorithm RPDA1, RPDA2, and RPDA3, respectively. These algorithms have their own merits. PCA and LPP are unsupervised algorithms. LDA and SLPP are supervised algorithms. Before we conduct LDA, LPP, SLPP and RPDA, the first stage is the PCA projection. In the PCA stage, $N - C$ dimensions are retained to ensure that within-scatter matrix S_w in LDA [38] is nonsingular, because the number of the original features is much larger than the number of training samples. In order to accelerate the learning process, we also conduct PCA step to retain $N - 1$ dimensions in LPP, SLPP, and RPDA.

The Nearest Neighbor (NN) rule was used in classification in the test stage. The performance is measured by the average accuracy for each class on the FFT features. To better understand different approaches, the confusion matrices between the ground truth class label and the most likely inferred label information are reported.

D. Experimental Results and Analysis

The average recognition rate is computed, which is varied with the number of dimensionalities. The result is shown in Fig. 4. Table II reports the best average accuracy and the corresponding dimensionalities of all of the algorithms on the test sets of all splits. It can be observed that RPDA outperforms the others in terms of recognition rate. In addition, RPDA1 classification confusion matrix was shown in Fig. 7. It shows: 1) “walking” is confused with “walking quickly” and “walking backwards” frequently; 2) “downstairs” is similar to “upstairs”; and 3) “cycling” is easy to be classified as “step walking.”

TABLE II
BEST AVERAGE RECOGNITION RATES OF SEVEN ALGORITHMS
ON THE FFT FEATURES

Classes	PCA	LDA	LPP	SLPP	RPDA1	RPDA2	RPDA3
Relaxing	0.7907	0.7209	0.8140	0.8372	0.9302	0.9302	0.8372
Walking	0.9000	0.7333	0.8333	0.7667	0.8000	0.8333	0.8000
Walking quickly	0.6818	0.8409	0.4545	0.7500	0.7955	0.7500	0.7955
Walking backwards	0.7273	0.6818	0.5455	0.6364	0.7955	0.8182	0.8636
Running	0.8636	0.8636	0.5455	0.6591	0.9091	0.8864	0.9091
Step walking	1.0000	0.9545	1.0000	1.0000	1.0000	1.0000	1.0000
Jumping	0.9773	0.9773	0.9318	0.8864	0.9773	0.9773	0.9773
Upstairs	0.9091	0.5909	0.7045	0.8182	0.8182	0.8636	0.8409
Downstairs	0.7273	0.6364	0.6364	0.5682	0.9091	0.8409	0.8864
Cycling	0.8409	0.8864	0.7727	0.8636	0.9545	0.8864	0.9318
Average accuracy	0.842 (42)	0.787 (9)	0.724 (30)	0.779 (23)	0.889 (43)	0.879 (41)	0.884 (44)

Fig. 5 shows the confusion matrices between neighborhood ranks in the original space and the dimensionality reduced subspaces (obtained by different dimension reduction algorithms) on all training sets. Fig. 6 shows the average percentage of perfectly preserved the top six within-class samples neighborhood ranks and the top nine within-class samples neighborhood ranks on all training sets. A higher value of a matrix entry corresponds to a darker square in the figure. From the viewpoint of rank preserving, the perfect dimension reduction algorithm generates a black diagonal matrix. PCA outperforms the other algorithms in terms of rank order information preserving. LDA cannot capture the nearest neighborhood rank order information. It is worth noting that RPDA algorithm outperforms the other algorithms in terms of rank order information preserving, except for PCA. The experimental results discovered the following two points.

- 1) PCA and LDA are the global linear dimension reduction algorithms. In our experiment, PCA outperforms the other algorithms in terms of rank order information preserving. This is because PCA considers preserving rank orders information globally and manifold learning algorithms consider preserving rank order information locally, the distribution of training samples benefits PCA. However, PCA preserves the within-class rank order information as well as the between-class rank. Therefore, it is not suitable for classification tasks. LDA preserves the discriminative information but ignores the rank order information. In addition, the between-class scatter matrix S_B is of rank $C - 1$, where C is the number of class, and we can obtain no more than $C - 1$ project vectors. These two issues caused that the performance of LDA performs more poorly than PCA in terms of classification accuracy when the subspace dimensionality increases.
- 2) LPP and SLPP are manifold learning algorithms and preserve the rank order information in a local patch. They are not as good as RPDA and PCA in terms of rank order information preserving shown in Fig. 6. In addition, they tend to preserve the rank order information of between-class intentionally, and thus they cannot model the discriminative information precisely. Therefore, they performed not as well as other algorithms.

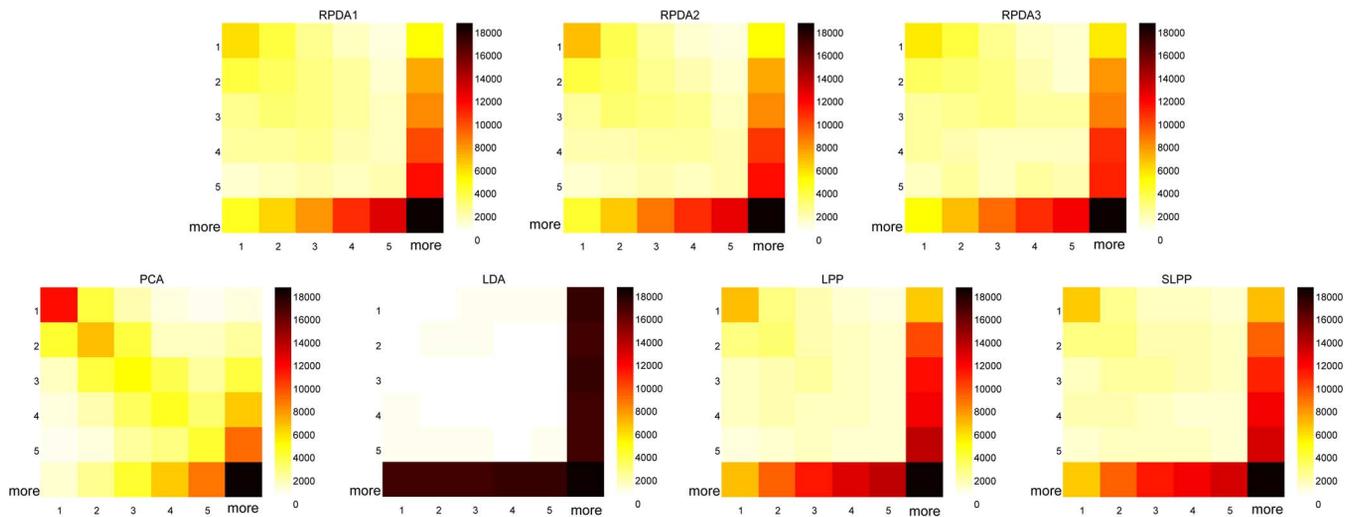


Fig. 5. Confusion matrix between neighborhood ranks in the original high-dimensional space and the subspace learned by a particular dimension reduction algorithm on all training sets. The x -axis contains values described ranks in original space. The y -axis contains values described ranks in output space. A higher value of a matrix entry corresponds to a darker square in the figure and a perfect dimension reduction from the viewpoint of rank order information preserving results in a black diagonal.

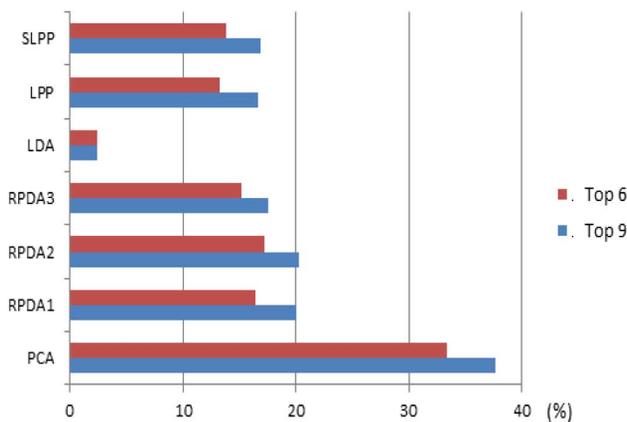


Fig. 6. Percentage of perfectly preserved the within-class samples neighborhood ranks on all training sets.

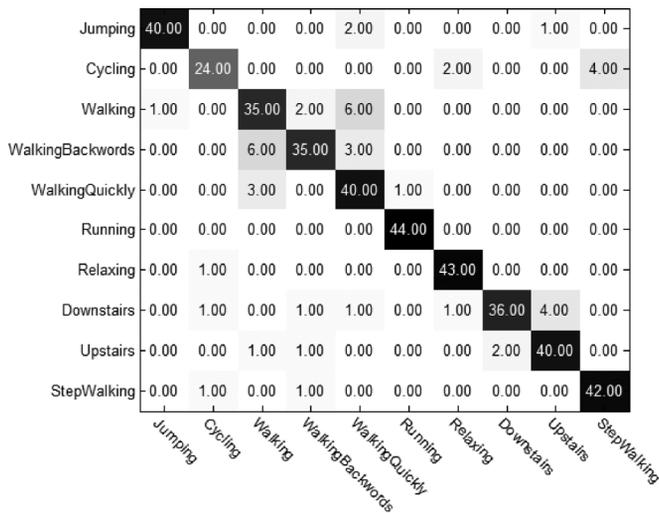


Fig. 7. RPDA1 classification confusion matrix for all testing sets.

TABLE III
BEST AVERAGE RECOGNITION RATES OF SEVEN ALGORITHMS
ON THE FFT FEATURES

Classes	PCA	LDA	LPP	SLPP	RPDA1	RPDA2	RPDA3
Average accuracy	0.872	0.790	0.835	0.847	0.891	0.891	0.889

(SVM classifier)

TABLE IV
BEST AVERAGE TRAINING TIME OF FIVE ALGORITHMS
ON THE FFT FEATURES

Classes	PCA	LDA	RPDA1	RPDA2	RPDA3
Average training time(second)	0.15	0.25	1.81	2.3	0.96

These results show the balance of the within-class nearest neighborhood ranks and the between-class nearest neighborhood ranks is important for improving the human behavior recognition accuracy.

In addition, we have tested the SVM classifier with the RBF kernel under the same experimental setting. It can be observed that the best accuracy achieved by RPDA1 and RPDA2 in Table III. Comparing with Table II, the SVM classifier improved the accuracy of recognition compared with the NN classifier. In addition, Table IV reports average training time of dimension reduction algorithms, including PCA, LDA and RPDA, as well as RPDA's variants. We conduct experiments on an Intel Core 2 Duo T9600 2.8GHz computer with a 4-Gbyte memory.

V. CONCLUDING REMARKS

In this paper, a new manifold learning based dimension reduction algorithm, Rank Preserving Discriminant Analysis (RPDA), was presented under the framework of patch alignment. Technically, rank orders do not encounter the concentration of measure phenomenon. Therefore, three type penalized factors were developed to encode the rank order information of the within-class samples. The between-class samples rank order information was ignored while the discriminative model of the between-class samples was constructed. Comparing to the classical unsupervised dimension reduction algorithms (e.g. PCA and LPP) and representative supervised dimension reduction algorithms (e.g. LDA and SLPP), RPDA has shown many competitive and attractive properties. In human behavior recognition for industrial informatics, RPDA is superior to the above algorithms in terms of recognition rate.

RPDA has some important parameters that affect the performance of the subsequent classification. Automatic selection of the parameters is important and should be investigated in a very careful way. Unlabeled samples are helpful for enhancing the classification performance. In application of human behavior recognition, it is practical to collect a large number of unlabeled samples. The sample distribution should be reckoned as a prior to improve the classification accuracy. Therefore, a semisupervised extension RPDA should be carefully considered. These two issues are important for industrial informatics.

In the future, we will apply the proposed RPDA to other applications, e.g., scene classification [37], [77] and multimedia tagging [34], [66]–[68]. Since robust visual features will cause the dimension of the image representation to be high, dimension reduction results in a succinct and effective representation for subsequent steps.

In addition, WSNs can be well applied to human behavior recognition, this utilization has some limitations. First, it is limited to power supply, because our WSNs are powered by batteries in considering mobility. Second, it cannot monitor a large area, because Zigbee is a short-range wireless transfer protocol.

ACKNOWLEDGMENT

The authors would like to thank the anonymous reviewers whose comments have improved the presentation of this paper significantly.

REFERENCES

- [1] C. C. Aggarwal, A. Hinneburg, and D. A. Keim, *On the Surprising Behavior of Distance Metrics in High-Dimensional Space*, J. V. Bussche and V. Vianu, Eds. Berlin, Germany: Springer, 2001, vol. 1973, Lecture Notes in Computer Science, pp. 420–434.
- [2] G. Anastasi, M. Conti, and M. Di Francesco, “A comprehensive analysis of the MAC unreliability problem in IEEE 802.15.4 wireless sensor networks,” *IEEE Trans. Ind. Inf.*, vol. 7, no. 1, pp. 52–65, Feb. 2011.
- [3] J. Antonino-Daviu, S. Aviyente, E. G. Strangas, and M. Riera-Guasp, “Scale invariant feature extraction algorithm for the automatic diagnosis of rotor asymmetries in induction motors,” *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 100–108, Feb. 2013.
- [4] P. Belhumeur, J. Hespanha, and D. Kriegman, “Eigenfaces vs. Fisherfaces: Recognition using class specific linear projection,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 19, no. 7, pp. 711–720, Jul. 1997.
- [5] M. Belkin and P. Niyogi, “Laplacian eigenmaps and spectral techniques for embedding and clustering,” *Adv. Neural Inf. Process. Syst.*, vol. 14, pp. 585–591, 2002.
- [6] L. Bao and S. S. Intille, “Activity recognition from user-annotated acceleration data,” in *Proc. 2nd Int. Conf. Pervasive Computing*, 2004, pp. 1–17.
- [7] W. Bian, D. Tao, and Y. Rui, “Cross-domain human action recognition,” *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 298–307, Apr. 2012.
- [8] D. Cai, X. He, and J. Han, Using Graph Model for Face Analysis Dept. Comput. Sci., Univ. of Illinois at Urbana-Champaign, Tech. Rep. 2636, 2005.
- [9] C. Caione, D. Brunelli, and L. Benini, “Distributed compressive sampling for lifetime optimization in dense wireless sensor networks,” *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 30–40, Feb. 2012.
- [10] E. J. Candès and T. Tao, “Near-optimal signal recovery from random projections: Universal encoding strategies?,” *IEEE Trans. Inf. Theory*, vol. 52, no. 12, pp. 5406–5425, Dec. 2006.
- [11] E. J. Candès and T. Tao, “The dantzig selector: Statistical estimation when p is much larger than n ,” *Ann. Statistics*, vol. 35, no. 6, pp. 2313–2351, 2007.
- [12] T. M. Chivewe and G. P. Hancke, “A distributed topology control technique for low interference and energy efficiency in wireless sensor networks,” *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 11–19, Feb. 2012.
- [13] T. Cuong and M. M. Trivedi, “3-D posture and gesture recognition for interactivity in smart spaces,” *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 178–187, Feb. 2012.
- [14] P. Demartines and J. Hérault, “Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets,” *IEEE Trans. Neural Netw.*, vol. 8, no. 1, pp. 148–154, Feb. 1997.
- [15] P. J. Deschavanne, A. Giron, J. Vilain, G. Fagot, and B. Fertil, “Genomic signature: Characterization and classification of species assessed by chaos game representation of sequences,” *Mol. Biol. Evol.*, vol. 16, pp. 1391–1399, 1999.
- [16] D. L. Donoho, “High-dimensional data analysis: The curses and blessings of dimensionality,” in *Proc. Amer. Math. Soc. Lecture*, Los Angeles, CA, USA, 2000.
- [17] D. L. Donoho, “Compressed sensing,” *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.
- [18] S. M. Elsayed, R. A. Sarker, and D. L. Essam, “An improved self-adaptive differential evolution algorithm for optimization problems,” *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 88–99, Feb. 2013.
- [19] A. Fathi and G. Mori, “Action recognition by learning mid-level motion features,” in *Proc. Int. Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
- [20] R. A. Fisher, “The use of multiple measurements in taxonomic problems,” *Ann. Eugenics*, vol. 7, pp. 179–188, 1936.
- [21] B. Geng, D. Tao, C. Xu, L. Yang, and X. Hua, “Ensemble manifold regularization,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 6, pp. 1227–1233, Jun. 2012.
- [22] H. Ghasemzadeh and R. Jafari, “Physical movement monitoring using body sensor networks: A phonological approach to construct spatial decision trees,” *IEEE Trans. Ind. Inf.*, vol. 7, no. 1, pp. 66–77, Feb. 2011.
- [23] X. He and P. Niyogi, “Locality preserving projections,” *Adv. Neural Inf. Process. Syst.*, vol. 16, pp. 153–160, 2004.
- [24] Z. He and L. Jin, “Weightlessness feature—a novel feature for single tri-axial accelerometer based activity recognition,” in *Proc. Int. Conf. Pattern Recognit.*, 2008, pp. 1–4.
- [25] H. Hotelling, “Analysis of a complex of statistical variables into principal components,” *J. Educational Psychol.*, vol. 24, pp. 417–441, 1933.
- [26] D. K. Hunter, H. Yu, M. S. Pukish III, J. Kolbusz, and B. M. Wilamowski, “Selection of proper neural network sizes and architectures—A comparative study,” *IEEE Trans. Ind. Inf.*, vol. 8, no. 2, pp. 228–240, May 2012.
- [27] Z. Ju and H. Liu, “A unified fuzzy framework for human hand motion recognition,” *IEEE Trans. Fuzzy Syst.*, vol. 19, no. 5, pp. 901–903, Oct. 2011.
- [28] S. Katsura, W. Yamanouchi, and Y. Yokokura, “Real-world haptics: Reproduction of human motion,” *IEEE Ind. Electron. Mag.*, vol. 6, no. 1, pp. 25–31, 2012.
- [29] V. Kellokumpu, G. Zhao, and M. Pietikäinen, “Human activity recognition using a dynamic texture based method,” in *Proc. Brit. Mach. Vis. Conf.*, 2008, pp. 885–894.
- [30] J. B. Kruskal, “Non-metric multidimensional scaling: A numerical method,” *Psychometrika*, vol. 29, pp. 115–129, 1964.

- [31] R. Kyusakov, J. Eliasson, J. Delsing, J. Deventer, and J. Gustafsson, "Integration of wireless sensor and actuator nodes with IT infrastructure using service-oriented architecture," *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 43–51, Feb. 2013.
- [32] K. V. Laerhoven, A. Schmidt, and H.-W. Gellersen, "Multi-sensor context aware clothing," in *Proc. 6th Int. Symp. Wearable Comput.*, 2002, pp. 49–56.
- [33] K. Lang, "Newsweeder: Learning to filter netnews," in *Proc. 12th Int. Conf. Mach. Learning*, 1995, pp. 331–339.
- [34] G. Li, M. Wang, Z. Lu, R. Hong, and T. S. Chua, "In-video product annotation with web information mining," *ACM Trans. Multimedia Computing, Commun., Appl.*, vol. 8, no. 4, p. 55, 2012.
- [35] J. Li and D. Tao, "Simple exponential family PCA," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 24, no. 3, pp. 485–497, 2013.
- [36] H. Li, A. Monti, and F. Ponci, "A fuzzy-based sensor validation strategy for AC motor drives," *IEEE Trans. Ind. Inf.*, vol. 8, no. 4, pp. 839–848, Nov. 2012.
- [37] F. Lian, Y. Lin, C. Kuo, and J. Jean, "Voting-based motion estimation for real-time video transmission in networked mobile camera systems," *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 172–180, Feb. 2013.
- [38] J. Lu, K. Plataniotis, and A. Venetsanopoulos, "Face recognition using LDA-based algorithms," *IEEE Trans. Neural Netw.*, vol. 14, no. 1, pp. 195–200, Feb. 2003.
- [39] N. Kubota, T. Obo, and H. Liu, "Human behavior measurement based on sensor network and robot partners," *J. Adv. Computational Intell. Intelligent Inf.*, vol. 14, no. 3, pp. 309–315, 2010.
- [40] I. Laptev, B. Caputo, C. Schüldt, and T. Lindeberg, "Local velocity-adapted motion events for spatio-temporal recognition," *Comput. Vis. Image Understanding*, vol. 108, no. 3, pp. 207–229, 2007.
- [41] I. Laptev, M. Marszalek, C. Schmid, and B. Rozenfeld, "Learning realistic human actions from movies," in *Proc. Int. Conf. Comput. Vis. Pattern Recognit.*, 2008, pp. 1–8.
- [42] H. Liu, S. Chen, and N. Kubota, "Guest editorial special section on intelligent video systems and analytics," *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, p. 90, Feb. 2012.
- [43] R. C. Luo and C.-C. Chang, "Multisensor fusion and integration: A review on approaches and its applications in mechatronics," *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 49–60, Feb. 2012.
- [44] S. Lespinats, B. Fertil, P. Villemain, and J. Hérault, "RankVisu: Mapping from the neighborhood network," *Neurocomputing*, vol. 72, no. 13–15, pp. 2964–2978, 2009.
- [45] S. Lespinats, M. Verleysen, A. Giron, and B. Fertil, "DD-HDS: A tool for visualization and exploration of high dimensional data," *IEEE Trans. Neural Netw.*, vol. 18, no. 5, pp. 1265–1279, Oct. 2007.
- [46] S. Livatino, F. Banno, and G. Muscato, "3-D integration of robot vision and laser data with semiautomatic calibration in augmented reality stereoscopic visual interface," *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 69–77, Feb. 2012.
- [47] H. Mahboubi, J. Habibi, A. G. Aghdam, and K. Sayrafian-Pour, "Distributed deployment strategies for improved coverage in a network of mobile sensors with prioritized sensing field," *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 451–461, Feb. 2013.
- [48] S. Maiti, V. Verma, C. Chakraborty, and Y. Hori, "An adaptive speed sensorless induction motor drive with artificial neural network for stability enhancement," *IEEE Trans. Ind. Inf.*, vol. 8, no. 4, pp. 757–766, Nov. 2012.
- [49] J. Mantyjarvi, J. Himberg, and T. Seppanen, "Recognizing human motion with multiple acceleration sensors," in *Proc. IEEE Int. Conf. Syst., Man, Cybern.*, 2001, pp. 747–752.
- [50] D. Miorandi, E. Uhlemann, S. Vitturi, and A. Willig, "Guest editorial: Special section on wireless technologies in factory and industrial automation, Part I," *IEEE Trans. Ind. Inf.*, vol. 3, no. 2, pp. 95–98, May 2007.
- [51] J. C. Niebles, H. Wang, and F.-F. Li, "Unsupervised learning of human action categories using spatial-temporal words," *Int. J. Comput. Vis.*, vol. 79, no. 3, pp. 299–318, 2008.
- [52] N. Nguyen, D. Phung, S. Venkatesh, and H. Bui, "Learning and detecting activities from movement trajectories using the hierarchical hidden Markov model," in *Proc. IEEE Int. Conf. Comput. Vis. Pattern Recognit.*, 2005, pp. 955–960.
- [53] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proc. Nat. Conf. Artif. Intell.*, 2005, pp. 1541–1546.
- [54] L. K. Saul and S. T. Roweis, "Think globally, fit locally: Unsupervised learning of low dimensional manifold," *J. Mach. Learning Res.*, vol. 4, pp. 119–155, 2003.
- [55] L. Shao, L. Ji, Y. Liu, and J. Zhang, "Human action segmentation and recognition via motion and shape analysis," *Pattern Recognit. Lett.*, vol. 33, no. 4, pp. 438–445, Mar. 2012.
- [56] L. Shao and R. Mattivi, "Feature detector and descriptor evaluation in human action recognition," in *Proc. ACM Int. Conf. Image Video Retrieval*, 2010, pp. 477–484.
- [57] M. Song, D. Tao, X. Huang, C. Chen, and J. Bu, "Three-dimensional face reconstruction from a single image by a coupled RBF network," *IEEE Trans. Image Process.*, vol. 21, no. 5, pp. 2887–2897, May 2012.
- [58] C. Sminchisescu, A. Kanaujia, and D. N. Metaxas, "Conditional models for contextual human motion recognition," *Comput. Vis. Image Understanding*, vol. 104, no. 2–3, pp. 210–220, 2006.
- [59] D. Tao and L. Jin, "Discriminative information preservation for face recognition," *Neurocomputing*, vol. 91, pp. 11–20, 2012.
- [60] L. Tang, K. Wang, and Y. Huang, "Study of speed-dependent packet error rate for wireless sensor on rotating mechanical structures," *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 72–80, Feb. 2013.
- [61] D. Tao, X. Li, X. Wu, and S. J. Maybank, "Geometric mean for subspace selection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 260–274, Feb. 2009.
- [62] C. Thureau, "Behavior histograms for action recognition and human detection," *Human Motion-Understanding, Modeling, Capture and Animation*, pp. 299–312, 2007.
- [63] X. Tian, D. Tao, and Y. Rui, "Sparse transfer learning for interactive video search reranking," *ACM Trans. Multimedia Comput., Commun., Applicat.*, vol. 8, no. 3, p. 26, 2012.
- [64] V. N. Vapnik, *Statistical Learning Theory*. New York, NY, USA: Wiley, 1998.
- [65] G. Wang, L. Tao, H. Di, X. Ye, and Y. Shi, "A scalable distributed architecture for intelligent vision system," *IEEE Trans. Ind. Inf.*, vol. 8, no. 1, pp. 91–99, Feb. 2012.
- [66] M. Wang, X. S. Hua, J. Tang, and R. Hong, "Beyond distance measurement: Constructing neighborhood similarity for video annotation," *IEEE Trans. Multimedia*, vol. 11, no. 3, pp. 465–476, Apr. 2009.
- [67] M. Wang, R. Hong, G. Li, Z. J. Zha, S. Yan, and T. S. Chua, "Event driven web video summarization by tag localization and key-shot identification," *IEEE Trans. Multimedia*, vol. 14, no. 4, pp. 975–985, Aug. 2012.
- [68] M. Wang, B. Ni, X. Hua, and T. S. Chua, "Assistive tagging: A survey of multimedia tagging with human-computer joint exploration," *ACM Computing Surveys*, vol. 44, no. 4, p. 25, 2012.
- [69] Z. Wang, C.-S. Leung, Y.-S. Zhu, and T.-T. Wong, "Data compression on the illumination adjustable images by PCA and ICA," *Signal Process.: Image Commun.*, vol. 19, no. 10, pp. 939–954, 2004.
- [70] D. Wu and L. Shao, "Silhouette analysis based action recognition via exploiting human poses," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 23, no. 2, pp. 236–243, Feb. 2012.
- [71] T. Xie, H. Yu, J. Hewlett, P. Rózycki, and B. M. Wilamowski, "Fast and efficient second-order method for training radial basis function networks," *IEEE Trans. Neural Netw. Learning Syst.*, vol. 23, no. 4, pp. 609–619, 2012.
- [72] Y. Xue and L. Jin, "A naturalistic 3D acceleration-based activity dataset & benchmark evaluations," in *Proc. IEEE Int. Conf. Syst., Man Cybern.*, 2010, pp. 4081–4085.
- [73] J. Yang, D. Zhang, A. F. Frangi, and J.-Y. Yang, "Two-dimensional PCA: A new approach to appearance-based face representation and recognition," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 26, no. 1, pp. 131–137, 2004.
- [74] H. Yu and J. Yang, "A direct LDA algorithm for high-dimensional data with application to face recognition," *Pattern Recognit.*, vol. 34, pp. 2067–2070, 2001.
- [75] Z. Zhang and D. Tao, "Slow feature analysis for human action recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 3, pp. 436–450, Mar. 2012.
- [76] G. Zhao and M. Pietikäinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 29, no. 6, pp. 915–928, Jun. 2007.
- [77] T. Zhang, S. Liu, C. Xu, and H. Lu, "Mining semantic context information for intelligent video surveillance of traffic scenes," *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 149–160, Feb. 2013.
- [78] T. Zhang, D. Tao, X. Li, and J. Yang, "Patch alignment for dimensionality reduction," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1299–1313, Sep. 2009.
- [79] T. Zhou and D. Tao, "1-bit Hamming compressed sensing," in *Proc. IEEE Int. Symp. Inf. Theory*, 2012, pp. 1862–1866.

- [80] T. Zhang, S. Liu, C. Xu, and H. Lu, "Mining semantic context information for intelligent video surveillance of traffic scenes," *IEEE Trans. Ind. Inf.*, vol. 9, no. 1, pp. 149–160, Feb. 2013.



Dapeng Tao received the B.S. degree in electronics and information engineering from Northwestern Polytechnical University, Xi'an, China. He is currently working toward the Ph.D. degree in information and communication engineering with South China University of Technology, Guangzhou, China.

His research interests include machine learning, computer vision, and cloud computing.



Lianwen Jin (M'09) received the B.S. degree from the University of Science and Technology of China, Hefei, China, in 1991, and the Ph.D. degree from South China University of Technology, Guangzhou, China, in 1996.

He is currently a Professor with the School of Electronic and Information Engineering, South China University of Technology, Guangzhou, China. He is the author or coauthor of more than 100 scientific papers. His research interests include image processing, handwriting analysis and recognition,

machine learning, cloud computing, and intelligent systems.

Prof. Jin is member of the IEEE Signal Processing Society, the IEEE Communication Society, the IEEE Computer Society, the China Image and Graphics Society, and the Cloud Computing Experts Committee of China Institute of Communications. He received the award of New Century Excellent Talent Program of MOE in 2006, the Guangdong Pearl River Distinguished Professor award in 2011. He served as a Program Committee member for a number of international conferences, including ICMLC 2007–2011, ICFHR 2008–2012, ICDAR 2009, ICDAR 2013, ICPR 2010, ICPR 2012, and ICMLA 2012.



Yongfei Wang received the B.S. degree in information engineering from Guangdong University of Technology, Guangzhou, China. He is currently working toward the M.S. degree in information and communication engineering at South China University of Technology, Guangzhou, China.

His research interests include machine learning and computer vision.

Xuelong Li (M'02–SM'07–F'12) is a Full Professor with the Center for Optical Imagery Analysis and Learning (OPTIMAL), State Key Laboratory of Transient Optics and Photonics, Xi'an Institute of Optics and Precision Mechanics, Chinese Academy of Sciences, Xi'an, China.