

Manifold Learning and Recognition of Human Activity Using Body-Area Sensors

Mi Zhang

*Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089 USA
Email: mizhang@usc.edu*

Alexander A. Sawchuk

*Ming Hsieh Department of Electrical Engineering
University of Southern California
Los Angeles, CA 90089 USA
Email: sawchuk@sipi.usc.edu*

Abstract—Manifold learning is an important technique for effective nonlinear dimensionality reduction in machine learning. In this paper, we present a manifold-based framework for human activity recognition using wearable motion sensors. In our framework, we use locally linear embedding (LLE) to capture the intrinsic structure and build nonlinear manifolds for each activity. A nearest-neighbor interpolation technique is then applied to learn the mapping function from the input space to the manifold space. Finally, activity recognition is performed by comparing trajectories of different activity manifolds in the manifold space. Experimental results validate the effectiveness of our framework and demonstrate that manifold learning is promising for the task of human activity recognition using wearable motion sensors.

Keywords—assisted living technologies; human activity recognition; manifold learning; body-area sensors

I. INTRODUCTION

As life expectancy generally increases among the world population, assisted living technologies that aid the elderly in their daily lives are particularly significant. Among these technologies, human activity recognition has received increasing attention in both academia and industry in recent years. Advances in semiconductor and MEMS technologies have enabled the development of miniaturized sensing equipments in multiple modalities such that people can wear these devices unobtrusively. Therefore, it is possible to use these wearable sensors to build systems that recognize human activities and understand human behaviors [1].

Existing wearable sensor-based activity models fall into two broad categories based on the level of signal granularity. The first category represents activities using a “whole-motion” model in which continuous sensor streams are divided into fixed length windows whose length is chosen such that all relevant information can be extracted from each single window. This global information is then transformed into a feature vector used as input to the classifier [2]. The second category represents activities using motion primitives that capture the local information. Activity models are then built on top of these motion primitives [3]. Although both models have been shown to be very effective in existing studies, they can achieve good performance only in relatively high dimensional feature spaces [4] [5]. This computational overhead limits their real world applications.

To lower the dimensionality and thus reduce the computational overhead, in this work, we propose a framework based on manifold learning techniques that embeds the high-dimensional human activity signals into a low-dimensional space for compact representation and recognition. The idea of this manifold-based framework stems from the observation that the sensor signals of a subject performing certain activity are constrained by the physical body kinematics and the temporal constraints posed by the activity being performed. Given these constraints, it is expected that the sensor signals vary smoothly and lie on a low-dimensional manifold embedded in the high-dimensional input space. Moreover, these manifolds capture the intrinsic activity structures and act as trajectories to characterize different types of activities. These motivate the analysis of human activities in the low-dimensional manifold space rather than the high-dimensional input space.

The keys to the success of the manifold-based framework are: (1) extracting meaningful activity manifolds that preserve the intrinsic structure of the human activity signals; and (2) constructing effective recognition algorithms to perform activity classification in the low-dimensional manifold space. For the first point, the main challenge is that activity manifolds are in nature nonlinear and even twisted. Because of such nonlinearity, linear models such as principal component analysis (PCA), and linear discriminant analysis (LDA) are not able to discover the underlying manifold structures. For the second point, the activity manifolds may have different shapes and lengths for different activities and even the same activity because different subjects may perform the same activity in different styles. The classification algorithm should be robust enough to handle these inter-class and intra-class variations.

Based on the considerations mentioned above, we focus on developing a human activity recognition framework based on nonlinear manifold learning techniques. Techniques such as isometric feature mapping (Isomap) [6], local linear embedding (LLE) [7], and Laplacian Eigenmap [8] are able to capture the low-dimensional nonlinear manifolds embedded in the high-dimensional input spaces for synthetic examples as well as real world applications, such as face recognition [9], visual speech recognition [10], visual object

tracking [11], vision-based body pose identification [12] and human movement analysis [13] [14]. However, there have been relatively fewer studies on practical applications of manifold learning for wearable sensor-based human activity recognition. Thus, there are two goals for this work. The first goal is to investigate whether there exists a compact low-dimensional manifold representation for the activity signals sampled from the wearable motion sensors. The second goal is to explore the feasibility of applying manifold learning techniques for human activity recognition in the low-dimensional manifold space.

The rest of this paper is organized as follows. Section II briefly reviews some existing work on human activity recognition. Section III introduces the sensing platform and dataset used for this study. Section IV describes the details of the manifold-based framework for human activity representation and recognition. Section V presents the evaluation results of this framework. Finally, section VI concludes this paper and establishes directions for future work.

II. RELATED WORK

Based on the granularity level human activities are modeled, existing activity recognition methods based on wearable sensors can be broadly classified into two categories: “whole-motion”-based approaches, and motion primitive-based approaches.

In the case of “whole-motion” model, different combinations of features and classifiers have been extensively studied on different sets of activities. In [2], Bao *et al.* studied statistical and frequency domain features in conjunction with four classifiers including decision trees (C4.5), decision tables, naive Bayes and nearest-neighbor. Among them, the decision tree achieved the best performance. Zhang *et al.* in [4] aimed to identify the most important features to recognize various human activities. They compared statistical features with physical features extracted based on the physical parameters of human motion. They showed physical features could make significant contributions to the performance of the recognition system.

In the case of motion primitive-based methods, in [3], motion primitives were constructed by dividing the activity trajectory into fixed-length windows with identical spatial duration, where each window was mapped to a motion primitive based on its trajectory direction in the Cartesian space. The problem of activity recognition was then formulated as a standard string-matching problem. As a further extension, Fihl *et al.* in [15] replaced the standard deterministic string-matching algorithm with a probabilistic-based string-matching strategy by using probabilistic edit distance instead of the standard edit distance. However, string-matching-based approaches are sensitive to noise and perform poorly in the presence of high intra-class variances. To overcome this problem, Zhang *et al.* in [5] developed

a statistical framework based on the Bag-of-Features (BoF) model and demonstrated superior performance.

More recently, research in human motion analysis has shifted toward using nonlinear manifolds to capture the structure of activity signals in the low-dimensional spaces, especially in the computer vision community. The goal here is to find a compact low-dimensional representation for the activity signals. Wang *et al.* in [11] used Isomap to learn the low-dimensional intrinsic object structure for the task of visual tracking, and have obtained significantly improved performance. In [12], Elgammal *et al.* applied LLE to construct human gait manifolds from silhouettes extracted from videos, and designed a mapping function based on generalized radial basis functions (GRBF) to infer 3D body poses from the constructed manifolds. Blackburn *et al.* in [13] followed the same idea as in [12], but used Isomap for manifold construction and dynamic time warping technique for manifold recognition.

In this work, we follow the basic principles of the nonlinear manifold learning techniques and apply them to build and recognize activity manifolds from the signals sampled from the wearable body-area motion sensors.

III. SENSING PLATFORM AND DATASET

For this work, data is recorded using an off-the-shelf multimodal sensing platform called MotionNode [16]. MotionNode is a 6-DOF inertial measurement unit (IMU) specifically designed for human motion sensing applications. It integrates a 3-axis accelerometer, a 3-axis gyroscope, and a 3-axis magnetometer. In this work, only the data sampled from the accelerometer and gyroscope is considered. The measurement range for each axis of accelerometer and gyroscope is $\pm 6g$ and $\pm 500dps$ respectively. The sampling rates for both accelerometer and gyroscope are set to 100 Hz. This setting is high enough to capture all the details of normal human motion.

Six participants with different gender, age, height, and weight are recruited to perform nine types of activities: walk forward, walk left, walk right, go upstairs, go downstairs, jump up, run, stand, and sit. These activities correspond to the most common activities in people’s daily life and are useful for both elder care and personal fitness applications. During data collection, a MotionNode is attached firmly onto the participant’s right front hip. Each participant performs five trials for each activity on different days at various indoor and outdoor locations without supervision.

IV. MANIFOLD-BASED FRAMEWORK

Figure 1 shows the block diagram of our manifold-based framework. The proposed framework consists of two stages. In the training stage, the streaming sensor data of each activity is first divided into a sequence of fixed-length window cells whose length is much smaller than the duration of the activity itself (in this work, we use a window cell size

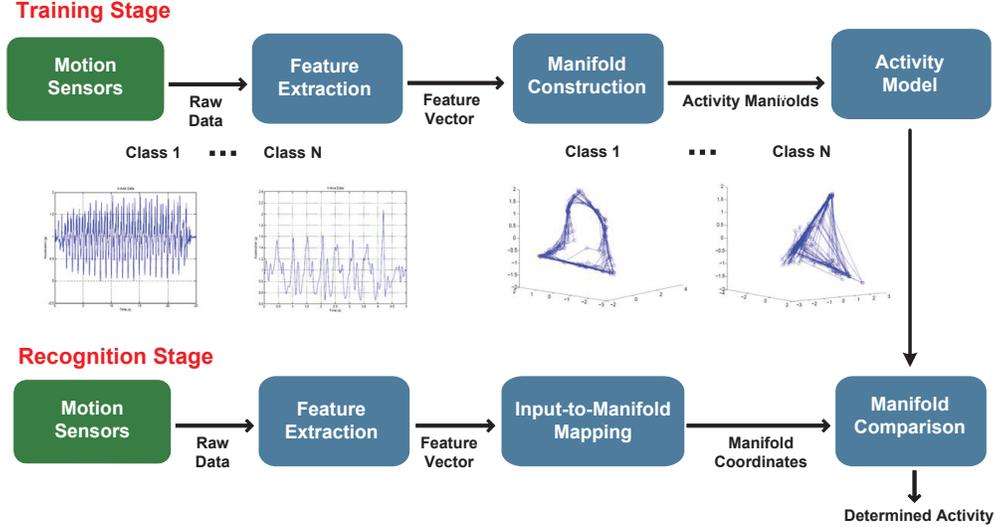


Figure 1. The block diagram of the manifold-based human activity recognition framework

of 0.1 second). Features are extracted from each window cell to form a local feature vector. The supervised LLE algorithm (as described in this section later) is then applied to map each high-dimensional local feature vector onto a low-dimensional manifold to construct activity manifold for each activity class. In the recognition stage, the unknown stream of sensor data is first transformed into a sequence of local feature vectors. These feature vectors are then mapped into the low-dimensional manifold space by the manifold projection mapping function learned in the training stage by means of the nearest-neighbor interpolation technique (as described in this section later). To classify the unknown sensor data, its newly constructed manifold is compared to the manifolds of the known activity classes. Finally, the manifold is classified as the activity class that has the most similar manifold. In the remainder of this section, we present the details of all the components of this framework.

A. Feature Extraction

For wearable sensor-based human activity recognition, a variety of features both in time and frequency domains have been investigated within the framework of the “whole-motion” model. Popular examples are mean, variance, entropy, correlation, FFT coefficients etc. However, since the total number of samples within each window cell is small, complex features such as entropy and FFT coefficients may not be reliably calculated. Therefore, we only consider features that can be reliably calculated with a small number of samples. Table I lists the features we include in this work. These features are extracted from each axis of both accelerometer and gyroscope. Therefore, the dimensionality of the input feature space is 30.

Feature	Description
Mean	The DC component (average value) of the signal over the window
Standard Deviation	Measure of the spreadness of the signal over the window
Root Mean Square	The quadratic mean value of the signal over the window
Averaged derivatives	The mean value of the first order derivatives of the signal over the window
Mean Crossing Rate	The total number of times the signal changes from below average to above average or vice versa normalized by the window length

Table I
FEATURES AND THEIR BRIEF DESCRIPTIONS

B. Learning Activity Manifolds

In this work, we adapt a LLE framework [7] to capture the intrinsic structures of the activity signals and construct the corresponding low-dimensional activity manifolds. We choose LLE over other manifold learning techniques such as Isomap and Laplacian Eigenmap because LLE makes fewer assumptions about the activity signals and runs much faster [17]. Although LLE was initially proposed as an unsupervised manifold learning algorithm, here, we utilize the class label information and construct manifolds for each activity class separately in a supervised manner.

Let $\mathbf{X} = \{\mathbf{x}_i \in R^D, i = 1, \dots, N\}$ be the input activity signal segment with length of N in the D -dimensional input space after feature extraction, where \mathbf{x}_i represents the local feature vector associated with the i^{th} window cell within the segment and acts as a single point in R^D . LLE takes \mathbf{X} as input and computes the corresponding coordinate vectors $\mathbf{Y} = \{\mathbf{y}_i \in R^d, i = 1, \dots, N\}$ in the d -dimensional manifold space ($d \ll D$). The procedure of LLE algorithm consists of three steps and is described as follows.

1) **Find neighborhood:** Find K nearest neighbors for each point $\mathbf{x}_i, i = 1, \dots, N$ in the D -dimensional input space. In this work, the Euclidean distance is used to

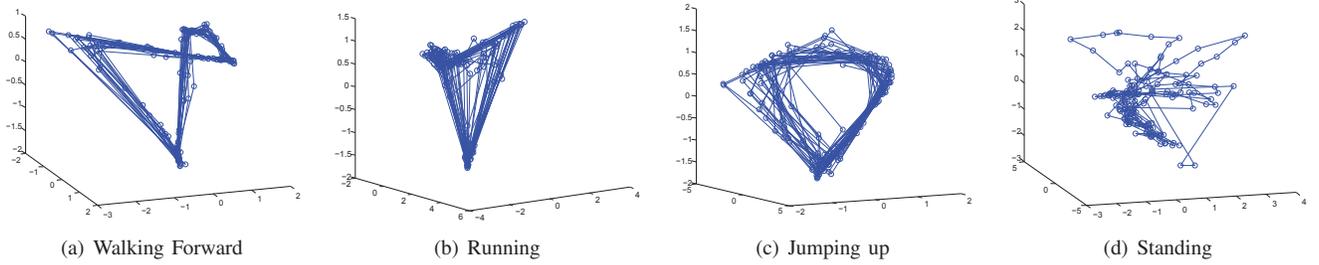


Figure 2. Manifolds of four different types of activities visualized in 3D spaces

measure the similarity between points after each feature dimension is normalized to zero mean and unit variance. The value of K is determined empirically.

2) **Compute reconstruction weights:** Assuming that each point and its neighbors lie on a locally linear patch of the underlying manifold, each point can be reconstructed as a linear combination of its K nearest neighbors found in the first step. The objective of this step is to compute the reconstruction weights that minimize the global reconstruction error measured by the cost function:

$$\epsilon(\mathbf{W}) = \sum_{i=1}^N \left\| \mathbf{x}_i - \sum_{j=1}^N W_{ij} \mathbf{x}_j \right\|^2 \quad (1)$$

where W_{ij} represents the contribution (weight) of \mathbf{x}_j to the reconstruction of \mathbf{x}_i .

To compute the weights W_{ij} , the cost function is minimized subject to two constraints: (1) $W_{ij} = 0$, if \mathbf{x}_j is not one of \mathbf{x}_i 's K nearest neighbors, and (2) $\sum_{j=1}^K W_{ij} = 1$, if \mathbf{x}_j is among \mathbf{x}_i 's K nearest neighbors. The solution (optimal weights W_{ij}) of this optimization problem can be found by solving a least-square problem [18].

3) **Construct d -dimensional embedding:** The constrained weights W_{ij} derived from step 2 characterize the intrinsic geometric properties of each point and its neighbors, and by design, they are invariant to transformations from D -dimensional input space to d -dimensional manifold space. Therefore, the same weights W_{ij} that reconstruct \mathbf{x}_i in D -dimensional input space can also reconstruct its embedded manifold coordinates \mathbf{y}_i in d -dimensional manifold space. Based on this characteristic, the manifold coordinates \mathbf{y}_i can be computed by minimizing the embedding cost function:

$$\Phi(\mathbf{Y}) = \sum_{i=1}^N \left\| \mathbf{y}_i - \sum_{j=1}^N W_{ij} \mathbf{y}_j \right\|^2 \quad (2)$$

Similar to step 2, to compute the manifold coordinates \mathbf{y}_i , the embedding cost function is minimized subject to two constraints: (1) $\sum_{i=1}^N \mathbf{y}_i = 0$, and (2) $\frac{1}{N} \sum_{i=1}^N \mathbf{y}_i \mathbf{y}_i^T = \mathbf{I}$. These two constraints make the problem well-posed, and the optimization problem is transformed into an eigenvalue

problem, in which we select the d non-zero eigenvectors corresponding to the d smallest eigenvalues to provide the desired d manifold coordinates [7].

Figure 2 illustrates the resulting manifolds in 3D spaces for four different activities: *walk forward*, *run*, *jump up*, and *stand*. As illustrated, the manifolds (trajectories) of *walk forward* (Figure 2(a)), *run* (Figure 2(b)), and *jump up* (Figure 2(c)) evolve along a closed nonlinear curve in the embedded space respectively. This is because these three activities are either periodic or semi-periodic, causing the trajectories of different cycles to overlap each other. More importantly, these results indicate that there exists a compact low-dimensional manifold representation for these activities. However, for the activity *stand* in Figure 2(d), there does not exist a clear trajectory representing the activity itself. This result is expected since *stand* is aperiodic and static such that it is difficult to extract a consistent trajectory. Therefore, it is not useful to recognize *stand* and other similar aperiodic and static activities using this manifold-based framework. Based on this observation, we do not take *stand* and *sit* into consideration from now on.

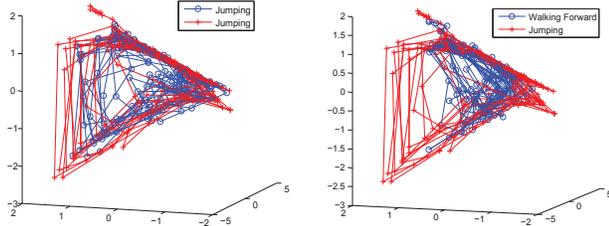
C. Learning Input-to-Manifold Mapping

As shown in the previous subsection, given the input coordinates in D dimensions, LLE provides the embedding coordinates in d dimensions directly. In other words, the mapping function $f : R^D \rightarrow R^d$ is not explicitly given by LLE. For the task of activity recognition, however, we need to compute the embedding coordinates corresponding to new test activity segments. In principle, we could rerun the entire LLE algorithm with the original training dataset augmented by the test activity segment. For large datasets of high dimensionality, however, this approach is prohibitively expensive. Thus, it is necessary to derive an explicit mapping function between the high and low dimensional spaces.

In this work, we use the non-parametric mapping function proposed in [18]. The mapping function is inspired from the LLE algorithm described in the previous subsection and learned by means of the nearest-neighbor interpolation technique. Specifically, to compute the embedding coordinates $\hat{\mathbf{y}}$ for a new input $\hat{\mathbf{x}}$, we perform the following three steps: (1) identify the K nearest neighbors of $\hat{\mathbf{x}}$ among the

training set (denoted as $\hat{x}_i, i = 1, \dots, K$); (2) compute the linear weights W_i that best reconstruct \hat{x} from its K nearest neighbors, subject to the constraint $\sum_{i=1}^K W_i = 1$; (3) since the neighbors of \hat{x} have known corresponding embedding coordinates (denoted as $\hat{y}_i, i = 1, \dots, K$), \hat{y} is then obtained by linearly combining these embedding coordinates with the recovered weights W_i . That is, $\hat{y} = \sum_{i=1}^K W_i \hat{y}_i$.

As an example, Figure 3 shows the resulting manifolds after mapping a test segment of activity *jump up* (Figure 3(a)) and a test segment of activity *walk forward* (Figure 3(b)) into the same activity *jump up* manifold space respectively. In both figures, the points in blue represent the mapped test segment. It is obvious to see that the manifolds of activity segments belonging to the same activity class have similar shapes and highly overlapped trajectories while the shapes of manifolds of different activity segments are quite distinct. This observation indicates that activity recognition can be performed by comparing the shapes of manifolds.



(a) The result after mapping a test segment of activity *jump up* to the activity *jump up* manifold space (b) The result after mapping a test segment of activity *walk forward* to the activity *jump up* manifold space

Figure 3. Mapping results of the non-parametric mapping function

D. Recognizing Activity Manifolds

Based on the observations in Figure 2 and 3, activity recognition is performed by comparing trajectories of manifolds in the low-dimensional space. One issue of trajectory comparison is that trajectories of manifolds from different activity classes or the same activity class but from different segments may be misaligned and have different lengths. Therefore, a distance measure that can handle misalignment and variations in trajectory lengths is desired. In this work, we use a variant of the Hausdorff metric, that is, the “mean value of the minimums”, to measure the distance between different manifolds:

$$Dist(M_1, M_2) = \frac{1}{T_{M_1}} \sum_{i=1}^{T_{M_1}} \min_{1 \leq j \leq T_{M_2}} \|M_1(i) - M_2(j)\| \quad (3)$$

where M_1 and M_2 are two manifolds under comparison, T_{M_1} and T_{M_2} are the lengths of M_1 and M_2 respectively, and $M_1(i)$ is the i^{th} point on the manifold M_1 [14]. Since the Hausdorff metric is directional, the distance measure is thus modified to ensure symmetry:

$$D(M_1, M_2) = Dist(M_1, M_2) + Dist(M_2, M_1) \quad (4)$$

Based on this distance measure, the recognition procedure is as follows. The test activity segment is first mapped into the manifold space of each known activity class to construct its manifold. This newly constructed manifold is then compared to the manifolds of each known activity class. The test activity segment is classified as the activity class that has the most similar manifold.

V. EVALUATION

In this section, we evaluate the effectiveness of our manifold-based framework. We divide the dataset into training set and test set. Both sets cover segments from all activity trials performed by all participants. Activity manifolds and the corresponding parameters are learned from the training set. A confusion table is built from the test set to illustrate the performance of the framework.

A. Estimating the Intrinsic Dimensionality

As our first experiment, since there exist compact low-dimensional manifold structures for human activity signals, it is important to estimate the manifolds’ intrinsic dimensionality. In this work, we use residual variance proposed in [6] for the estimation. The residual variance is defined as

$$\text{residual variance} = R^2(D_I, D_M) \quad (5)$$

where D_I , and D_M are the Euclidean distance matrices in the input space and low-dimensional embedding space, respectively, and R is the standard linear correlation coefficient, taken over all entries of D_I , and D_M . The lower the residual variance is, the better the high-dimensional input data are represented in the low-dimensional manifold space.

Figure 4 illustrates the values of residual variance as a function of the dimensionality of the manifold space for different activities. To avoid overfitting, the intrinsic dimensionality of the manifold d is estimated by looking for the “elbow” at which the curve ceases to decrease significantly with added dimensions [6]. As expected, the intrinsic dimensionality for different activity manifolds are different. Specifically, for activity *walk forward* (Figure 4(a)), *walk left* (Figure 4(b)), *walk right* (Figure 4(c)), and *run* (Figure 4(f)), the estimated intrinsic dimensionality is 3. For activity *go upstairs* (Figure 4(d)) and *go downstairs* (Figure 4(e)), the estimated intrinsic dimensionality is 4. For activity *jump up* (Figure 4(g)), the estimated intrinsic dimensionality is 2. It should be noted that the higher the intrinsic dimensionality is, the more dimensions of variation and complicated structure the activity has. Therefore, our result indicates that *go upstairs* and *go downstairs* contain the most complicated structures while *jump up* has the simplest structure among all the activities, which to some extent matches our intuition. Finally, since the intrinsic dimensionalities are different for different activities, activity manifolds are constructed and classified in their own intrinsic dimensionality spaces respectively.

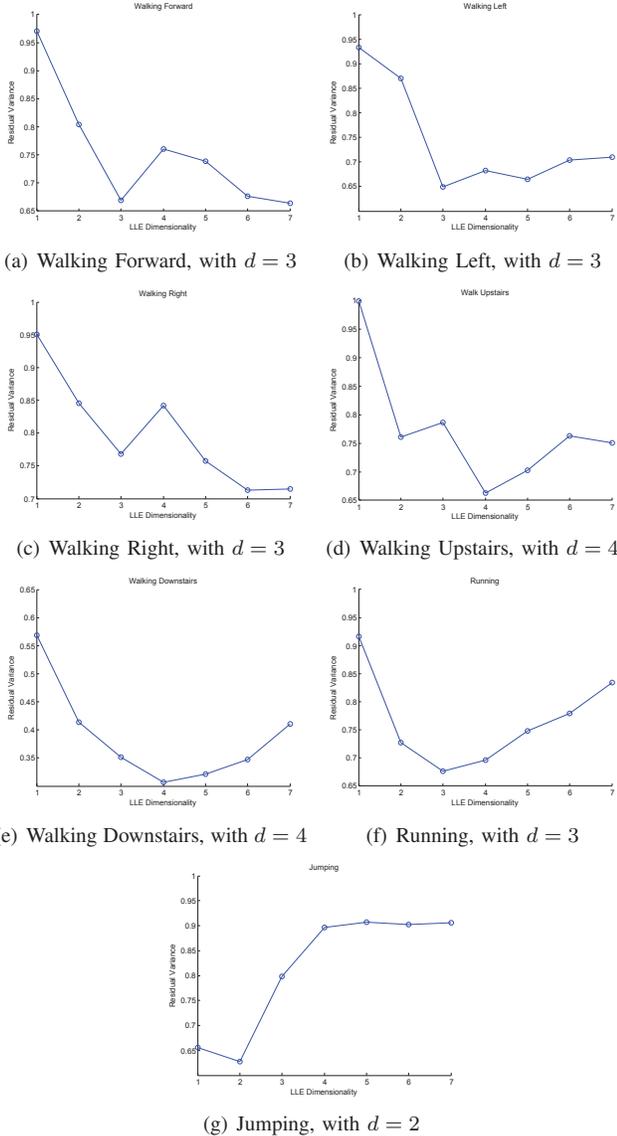


Figure 4. Estimation of the intrinsic dimensionality of different activity manifolds based on residual variance

B. Impact of the Number of Nearest Neighbors

The other key parameter of the framework is the number of nearest neighbors (K) defined in the first step of LLE. It is obvious to see that a small K may falsely divide a continuous manifold into disjoint sub-manifolds. To the extreme, the LLE algorithm can only recover embeddings whose intrinsic dimensionality is strictly less than K [18]. In contrast, a large K may violate the basic assumption of local linearity. Furthermore, if K is larger than the dimensionality of the input space (in our case, $D = 30$), the local reconstruction weights in the second step of LLE are no longer uniquely defined [18]. Given these constraints, in this study, we experiment with K ranging from 5 to 25. Five-fold cross validation is used to evaluate the performance. The

best K is determined as the one at which the classification accuracy reaches the maximum.

Figure 5 shows the average misclassification rates as a function of K at 5, 10, 15, 20, and 25. The error bars represent the standard deviation across five folds in cross validation. As illustrated in the figure, the misclassification rate drops significantly from $K = 5$ and reaches the minimum at $K = 10$. When K is larger than 10, the misclassification rate increases. This observation demonstrates that using 10 nearest neighbors is the best to construct activity manifolds for our dataset.

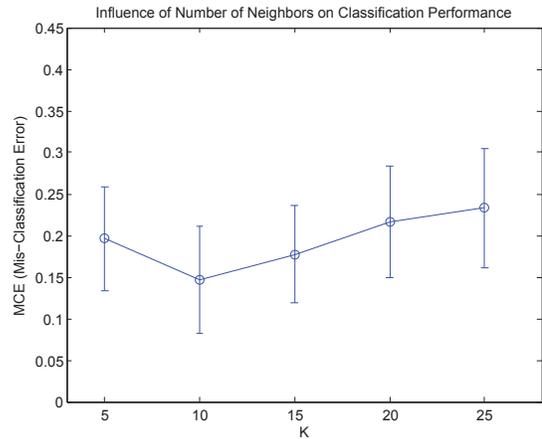


Figure 5. Impact of the number of nearest neighbors (K) on the classification performance of the manifold-based framework

C. Confusion Table

Finally, the confusion table for the test set with $K = 10$ is shown in Table II. The overall recognition accuracy across all activities is 80.3%. If we examine the recognition performance for each activity individually, *jump up* (with a 94.1% precision and 90.9% recall) and *run* (with a 98.2% precision and 88.9% recall) are the two easiest activities to recognize. Compared to other activities, *go upstairs* and *go downstairs* have relatively low precision values (75.5% and 74.5% respectively). This is because these two activities can be confused with each other and other walking-related activities. Finally, *walk forward* and *walk right* perform the worst in the sense that they have the lowest recall value (71.9%) and precision value (74.0%) respectively. As illustrated in the table, *walk forward*, *walk left* and *walk right* are commonly confused. This indicates that the manifolds of these three activities have similar shapes and trajectories such that our manifold-based framework has difficulties in differentiating them from each other.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented a human activity recognition framework using wearable motion sensors based on manifold learning. As a conclusion, our framework is able to

		Classified Activity						Total	Recall	
		Walk forward	Walk left	Walk right	Go up stairs	Go down stairs	Run forward			Jump up
Ground Truth	1 Walk forward	146	23	27	2	2	1	2	203	71.9%
	2 Walk left	12	178	28	3	1	0	0	222	80.2%
	3 Walk right	19	27	182	1	3	0	0	232	78.4%
	4 Go up stairs	2	1	2	40	3	0	1	49	81.6%
	5 Go down stairs	1	1	3	2	35	1	0	43	81.4%
	6 Run forward	3	2	3	3	1	112	2	126	88.9%
	7 Jump up	1	2	1	2	2	0	80	88	90.9%
Total		184	234	246	53	47	114	85		
Precision		79.3%	76.1%	74.0%	75.5%	74.5%	98.2%	94.1%		

Table II

CONFUSION TABLE WHEN USING 10 NEAREST NEIGHBORS. THE ENTRY IN THE i^{th} ROW AND j^{th} COLUMN IS THE COUNT OF ACTIVITY INSTANCES FROM CLASS i BUT CLASSIFIED AS CLASS j . THE OVERALL RECOGNITION ACCURACY IS 80.3%.

capture the intrinsic low-dimensional manifold structures for activities that are either periodic or semi-periodic. Furthermore, we demonstrate that activity recognition can be performed on top of this compact representation and achieves promising results. For future work, we aim to improve the classification performance by investigating the use of statistical manifold comparison approaches (for instance, Hidden Markov Model (HMM)) to replace the current non-statistical distance measure-based approach.

REFERENCES

- [1] Mi Zhang and Alexander A. Sawchuk. A customizable framework of body area sensor network for rehabilitation. In *International Symposium on Applied Sciences in Biomedical and Communication Technologies (ISABEL)*, pages 1–6, Bratislava, Slovak Republic, November 2009.
- [2] Ling Bao and Stephen S. Intille. Activity recognition from user-annotated acceleration data. In *International Conference on Pervasive Computing*, pages 1–17, Linz/Vienna, Austria, 2004.
- [3] Thomas Stiefmeier, Daniel Roggen, and Gerhard Tröster. Gestures are strings: efficient online gesture spotting and classification using string matching. In *International Conference on Body Area Networks (BodyNets)*, pages 16:1–16:8, Florence, Italy, 2007.
- [4] Mi Zhang and Alexander A. Sawchuk. A feature selection-based framework for human activity recognition using wearable multimodal sensors. In *International Conference on Body Area Networks (BodyNets)*, Beijing, China, November 2011.
- [5] Mi Zhang and Alexander A. Sawchuk. A bag-of-feature-based framework for human activity representation and recognition. In *ACM International Conference on Ubiquitous Computing (Ubicomp) Workshop on Situation, Activity and Goal Awareness*, Beijing, China, September 2011.
- [6] Joshua B. Tenenbaum, Vin de Silva, and John C. Langford. A Global Geometric Framework for Nonlinear Dimensionality Reduction. *Science*, 290(5500):2319–2323, 2000.
- [7] Sam T. Roweis and Lawrence K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323, 2000.
- [8] Mikhail Belkin and Partha Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Computation*, 15:1373–1396, 2002.
- [9] Ming-Hsuan Yang. Extended isomap for pattern classification. In *Eighteenth National Conference on Artificial Intelligence (AAAI)*, pages 224–229, Edmonton, Alberta, Canada, 2002.
- [10] Christoph Bregler and Stephen M. Omohundro. Nonlinear manifold learning for visual speech recognition. In *International Conference on Computer Vision (ICCV)*, pages 494–499, Cambridge, Massachusetts, USA, 1995.
- [11] Qiang Wang, Guangyou Xu, and Haizhou Ai. Learning object intrinsic structure for robust visual tracking. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 227–233, Madison, Wisconsin, USA, 2003.
- [12] Ahmed M. Elgammal and Chan-Su Lee. Inferring 3D body pose from silhouettes using activity manifold learning. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 681–688, Washington, DC, USA, 2004.
- [13] Jaron Blackburn and Eraldo Ribeiro. Human motion recognition using isomap and dynamic time warping. In *Workshop on Human Motion: Understanding, Modeling, Capture and Animation*, pages 285–298, Rio de Janeiro, Brazil, 2007.
- [14] Liang Wang and David Suter. Analyzing human movements from silhouettes using manifold learning. In *IEEE International Conference on Video and Signal Based Surveillance (AVSS)*, Sydney, New South Wales, Australia, 2006.
- [15] Preben Fihl, Michael B. Holte, Thomas B. Moeslund, and Lars Reng. Action recognition using motion primitives and probabilistic edit distance. In *International Conference on Articulated Motion and Deformable Objects (AMDO)*, pages 375–384, Mallorca, Spain, 2006.
- [16] <http://www.motionnode.com>.
- [17] Lawrence K. Saul, Kilian Q. Weinberger, Fei Sha, Jihun Ham, and Daniel D. Lee. Spectral methods for dimensionality reduction. *Semisupervised Learning*. MIT Press: Cambridge, MA, USA, 2006.
- [18] Lawrence K. Saul and Sam T. Roweis. Think globally, fit locally: unsupervised learning of low dimensional manifolds. *Journal of Machine Learning Research*, 4:119–155, December 2003.