A Fuzzy Kernel Motion Classifier for Autonomous Stroke Rehabilitation

Zhe Zhang, *Student Member, IEEE*, Luca Liparulo, *Student Member, IEEE*, Massimo Panella, *Member, IEEE*, Xudong Gu, and Qiang Fang, *Member, IEEE*

***Abstract*—Autonomous poststroke rehabilitation systems which can be deployed outside hospital with no or reduced supervision have attracted increasing amount of research attentions due to the high expenditure associated with the current inpatient stroke rehabilitation systems. To realize an autonomous systems, a reliable patient monitoring technique which can automatically record and classify patient's motion during training sessions is essential. In order to minimize the cost and operational complexity, the combination of nonvisual-based inertia sensing devices and pattern recognition algorithms are often considered more suitable in such applications. However, the high motion irregularity due to stroke patients' body function impairment has significantly increased the classification difficulty. A novel fuzzy kernel motion classifier specifically designed for stroke patient's rehabilitation training motion classification is presented in this paper. The proposed classifier utilizes geometrically unconstrained fuzzy membership functions to address the motion class overlapping issue, and thus, it can achieve highly accurate motion classification even with poorly performed motion samples. In order to validate the performance of the classifier, experiments have been conducted using real motion data sampled from stroke patients with a wide range of impairment level and the results have demonstrated that the proposed classifier is superior in terms of error rate compared to other popular algorithms.**

***Index Terms*—Autonomous motion classification, fuzzy pattern recognition, inertia sensing devices, stroke rehabilitation, unconstrained membership function (MF).**

I. Introduction

STROKE is an acute cerebrovascular disease caused by haemorrhage or blockage in brain blood vessels. The cerebral ischemia as a direct consequence of stroke can lead to severe neurological damage and depending on the region affected, a number of impairments including muscle weakness, sensory loss, aphasia, cognitive problem and visuospatial dysfunction [1]. Stroke is known as the leading cause of death and ongoing disability in the world [2]. According to statistics, approximately 17 million people worldwide had a new or recurrent stroke incident in 2010 and 5.9 million deaths during the year were stroke related [3]. By 2010, there were also 33 million stroke survivors in the world and many of them are still suffering from stroke impairments and unable to live independently which has created a great challenge for health care and rehabilitation institutes around the globe.

Poststroke rehabilitation has been proven to be essential and effective in helping stroke patients to gradually regain some of their body functionality in numerous researches[4]. However, as the incidence of stroke is continuously increasing with the aging of the population, the amount of health care expenditure contributes to stroke-related programs is also rising rapidly. In Australia, the average first year cost per case for first time Ischemic Stroke (IS) and Intracerebral Haemorrhage (ICH) stroke patients in 2004 were A$6022 and A$3977, respectively. The lifetime cost per case for IS and ICH were estimated to be A$57.106 and A$49.995, which add up to A$1.7 billion and A$232 million for the entire nation. It is also worth noting that the inpatient rehabilitation expenses have contributed 28.7% and 27.9% to the overall cost for each type of stroke [5].

In recent years, unsupervised and home-based rehabilitation training programs have attracted substantial amount of research attentions not only for cost reduction but also to improve rehabilitation outcome. Early supported discharge (ESD) is a widely recognized program which aims to accelerate discharge by providing comprehensive support for the patients to continue rehabilitation training in community or home setting [6]–[8]. ESD allows patients to train in a more familiar environment where they are going to perform activities of daily living and create a smoother transition. However, how to maintain the intensity and quality of rehabilitation training in an environment with no or reduced supervision is one of the primary issues that must be addressed to ensure the rehabilitation outcome.

One viable solution, which has been proposed in various researches, is to introduce an automatic management system which is capable of unsupervised motion recording and classification during a rehabilitation training program, and thus, the doctors will be able to track patients' training performance without having to attend every training session [9]–[14]. Inertial measurement unit (IMU)-based motion tracking system is commonly considered as most suitable for home-based rehabilitation application due to its advantages of being compact, cost effective, and relatively easy to operate compared to its counterparts, such as visual-based tracking systems [15]–[17]. Because IMU-based systems are prone to output drifting problem caused by measurement noise accumulated over time, it is common to apply pattern recognition techniques on the sensor

Manuscript received October 25, 2014; revised March 20, 2015; accepted April 16, 2015. Date of publication May 6, 2015; date of current version May 9, 2016.

Z. Zhang and Q. Fang are with the School of Electrical and Computer Engineering, RMIT University, Melbourne, Vic. 3000, Australia (e-mail: zhe.zhang@rmit.edu.au; john.fang@rmit.edu.au).

L. Liparulo and M. Panella are with the Department of Information Engineering, Electronics and Telecommunications, University of Rome "La Sapienza," 00184 Rome, Italy (e-mail: luca.liparulo@uniroma1.it; massimo.panella@uniroma1.it).

X. Gu is with the Rehabilitation Medical Centre, Jiaxing 2nd Hospital, Zhejiang 314000, China (e-mail: jxgxd@hotmail.com).

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/JBHI.2015.2430524

output directly without motion reconstruction. Therefore, to achieve high accuracy and optimal efficiency, the selection of data processing methods and classification algorithms is crucial.

The objective of pattern recognition is to solve machine learning problems by grouping or classifying objects into a number of categories automatically and efficiently [18]. When the recognition process assigns the objects a label associated with a finite number of discrete categories, using the priori knowledge from sample targets, the process is known as supervised classification [19]. A number of studies have presented regarding the implementation of various classification techniques for solving human motion classification problems.

In [20], the authors have hypothesized that two electromyographic (EMG) channel recordings could provide useful information for evaluating the outcome of rehabilitation determining the spatial characteristics of motor activity. The analysis has been carried out defining 14 different movements and using support vector machines (SVMs) with the implementation of radial basis function (RBF) kernels. In a similar way, Sazonov *et al.* [21] carried out a study about the automatic recognition of postures and activities in stroke patients with the goal of applying the findings in rehabilitation programs. The automatic recognition approach, which was based on a SVM classification of the sensor data sampled by a wearable shoe-based device, provided good performance in terms of classification error rate. Nevertheless, in both articles, the small number of subjects and the manual signal segmentation method cannot guarantee the optimal generalization of the proposed method.

Instance-based learning algorithm such as K-nearest neighbor has been previously proposed to address the motion classification problem [13], [14]. Unlike many other pattern recognition techniques, instance-based method retains the original input instances for classification without a learning process to generalized data into a set of inference rules. This type of the learning strategy is generally referred as lazy learning, as the most of the work is delayed until the evaluation stage when the query is made and as a result, the classification process can become too cumbersome and impractical for many applications. When template matching-based approach is implemented with multinomial motion classification, the problem is amplified as the heavy querying process may have to be repeated multiple times to locate the optimal match. In order to be integrated into regular rehabilitation training, the classification process must be computationally inexpensive to perform even with large number of motion types. Therefore, despite the classification performance, instance-based learning cannot be considered as the most suitable candidate for this application.

For classification algorithms, one of the most challenging problem is the overlapping of classes. Generally, a parametric model is built on the basis of the statistical or data-driven assumptions that are developed from a set of labeled observations. The model is used for assigning any point in the data space with one or more class labels and in some cases, a specific label can be determined for each point using a certain decision making process [22], [23]. The classical crisp approaches do not allow the classes to overlap since the estimated model only outputs a binary decision that associates each input pattern with a single class [24]. Contrarily, a statistical classifier estimates *a posteriori* probability with a value between 0 and 1 for each point in the data space to be associated with each class [25]. A suited decision theory or, more commonly, a threshold can be applied to make decisions through a winner-takes-all (WTA) strategy.

In a similar way, the fuzzy approaches consider the uncertainties within the data and define fuzzy labels using a degree of membership to each class with a range between zero and one. A set of fuzzy values can be determined for each pattern by evaluating a suited membership function (MF) [26]–[31]. In the application of motion classification during rehabilitation training, the complex movements performed by the poststroke patients are very likely to create overlapping data structures. In this case, several patterns may have partial membership to different classes, each representing a particular motion. Fuzzy logic can address the problem by assigning each pattern a label after evaluating the MFs for all the classes, and a WTA strategy, which is explained in greater detail in Section III, can be applied to make decisions. Several works in the literature demonstrated that fuzzy systems are useful in biomedical signal processing. OMalley *et al.* [32] invoked the well-known fuzzy c-means algorithm to find natural groupings among gait variables. It subsequently categorizes new subjects according to the discovered groups using data collected from 156 subjects with stroke. Chan *et al.* [33] proposed a fuzzy approach to classify single-site EMG signals for multifunctional prosthesis control.

In this paper, a fuzzy kernel motion classifier for autonomous poststroke rehabilitation is proposed. The system utilizes the geometrically unconstrained fuzzy MFs previously discussed in [34]–[36]. The flexible shape of the MF can effectively manage the motion class overlapping problem which is aggravated by stroke patients' irregular performance. It is shown in this study that the proposed model is capable of differentiating complex rehabilitation exercise movements efficiently using the accelerometer and gyroscope measurements. The classifier is designed to work with stroke patients suffering from various degrees of motion impairment in rehabilitation training environments with no or reduced supervision such as patient's home or community center. Preliminary experiments with motion samples from actual stroke patients were conducted to validate the performance of the system.

The rest of this paper is organized as follows. In Section II, an overview is presented as well as the detailed explanation of the proposed classification system. In Section III, the inner fuzzy classifier and the used fuzzy MFs are illustrated. The experiment setup and the numerical results are explained in Sections IV and V, respectively. Finally, in Section VI, the conclusion and the future works are discussed.

II. System Overview

The overview of the proposed rehabilitation motion classification system is shown in Fig. 1. The motion data were sampled at 400 Hz using the accelerometer and gyroscope from an XSens IMU module (MTi-300). The sensor node was attached to patient's forearm with the positive direction of x-axis pointing the elbow, as illustrated in Fig. 2. The node was aligned

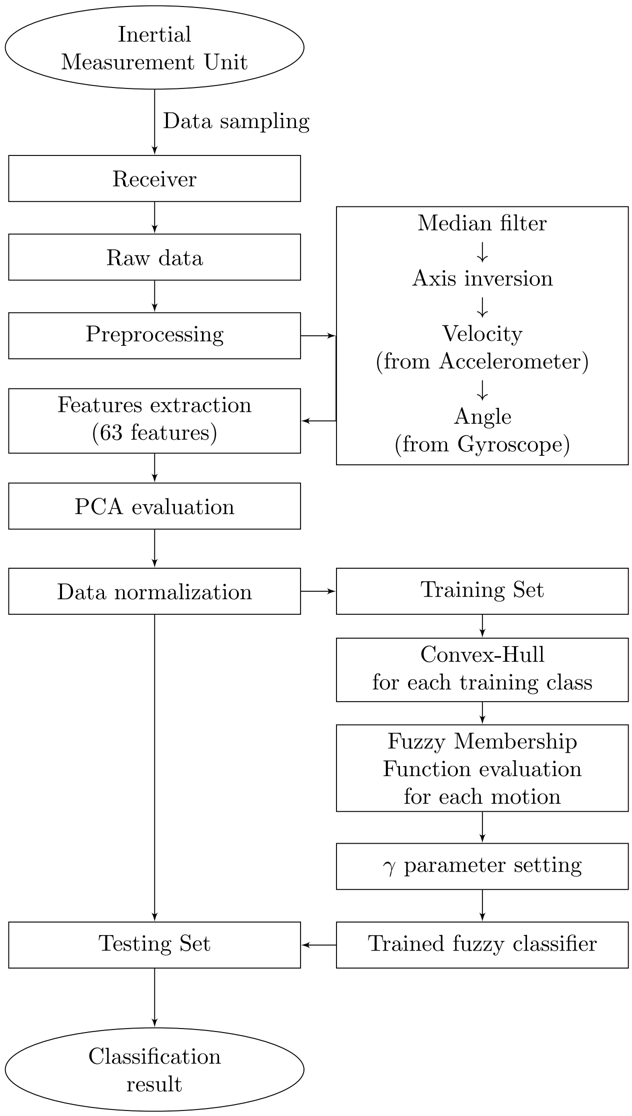

Fig. 1. Flowchart of the proposed system for upper limb motions classification in poststroke rehabilitation.

with the centerline of the back of the hand and mounted as close to the wrist joint as possible while making sure it will not be affected by any wrist movement. The raw data collected were then fed through a sequence of preprocessing procedures. Median filter were first applied to remove spikes in the signal caused by electrical noise or signal drop-off without distortion to the waveform. In practical situations, it needs to be taken into account that the patients may have different sides of the body affected by stroke. Therefore, during the preprocessing stage, data collected from left-hand impaired patients had the *Y*-axis of accelerometer data and *X*- and *Z*-axis of gyroscope data inverted in order to rectify the difference. After the axis inversion, the three-axis accelerometer reading was purified by removing the static offset caused by gravity and then integrated into velocity. However, the raw accelerometer measurement was kept in separated data sequences in order to retain the information of static acceleration. The three-axis angular velocity measured by gyroscope was also integrated into orientation; therefore, after the

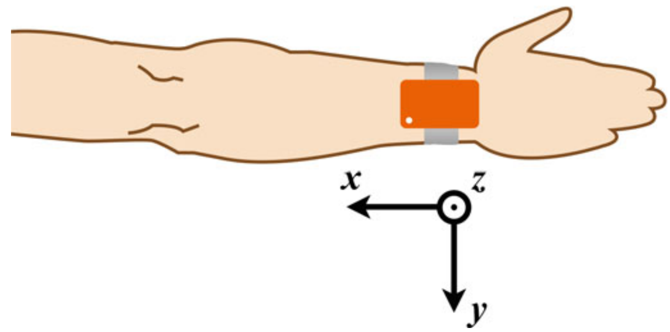

Fig. 2. Position and orientation of the sensor node.

preprocessing, nine data sequences (velocity, raw acceleration, and orientation) were ready for feature extraction. In total, 63 features were used as shown in Table I and a principal component analysis (PCA) process is applied in order to reduce the data complexity before the training classifier. The extracted features were then normalized before being fed into the proposed fuzzy classifier described in the following.

A. *Principal Component Analysis*

PCA is a widely adopted statistical tool which is capable of performing spatial redundancy reduction through orthogonal linear transformation and explaining the variances in a large number of variables using only a few principal components. It has been demonstrated in a number of researches that by properly implementing the PCA feature reduction process, the performance of a classifier can be boosted significantly [37] and PCA can also be applied as signal decomposition and dimension reduction tool in solving various engineering problems [38]–[40].

The essence of PCA is to apply orthogonal transformation on the covariance matrix of the input feature data and to produce uncorrelated components with sequentially maximum possible variances. After an initial preprocessing step of data normalization, which ensures that the feature values are scaled to the range between 0 and 1 and the PCA can work with absolute reference values, the principal components are sorted in descending order of significance based on the eigenvalues, and the percentage of contribution is calculated to indicate the amount of information carried by each component. By carefully choosing the number of principal components based on the total contribution or information covered, it is possible to greatly reduce the large number of features without losing much information. To a classifier, this means large amount of input information can be processed using significantly less time without sacrificing accuracy, which is the reason why PCA is considered as an important part of the proposed classification system.

B. *Data Normalization*

A second preprocessing step of data normalization is evaluated for scaling the PCA results to the range between 0 and 1. Let M be the number of patterns of the dataset $\mathcal{D} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_M\}$ and N be the number of data features; that is,

TABLE I
LIST OF THE FEATURES

Feature type	Feature names	Feature description
Mean	MeanACCX, MeanACCY, MeanACCZ, MeanANGX, MeanANGY, MeanANGZ, MeanVELX, MeanVELY, MeanVELZ	Mean magnitude calculated for acceleration, angular position, and velocity data sequence
Standard Deviation	StdACCX, StdACCY, StdACCZ, StdANGX, StdANGY, StdANGZ, StdVELX, StdVELY, StdVELZ	Standard deviation calculated for acceleration, angular position and velocity data sequence
Duration	DurACCX, DurACCY, DurACCZ, DurANGX, DurANGY, DurANGZ, DurVELX, DurVELY, DurVELZ	Effective duration calculated by counting the number of continues samples with absolute magnitude greater than the 20th percentile for each data sequence
Energy	EneACCX, EneACCY, EneACCZ, EneANGX, EneANGY, EneANGZ, EneVELX, EneVELY, EneVELZ	Signal energy calculated by taking the sum of squared magnitude for each data sequence
Dominant Frequency Power	PowACCX, PowACCY, PowACCZ, PowANGX, PowANGY, PowANGZ, PowVELX, PowVELY, PowVELZ	Power Spectral Density (PSD) at dominant frequency calculated for each data sequence after time-frequency domain conversion using Fast-Fourier Transform (FFT)
Dominant Frequency	FreqACCX, FreqACCY, FreqACCZ, FreqANGX, FreqANGY, FreqANGZ, FreqVELX, FreqVELY, FreqVELZ	Dominant frequency calculated by locating the peak PSD for each data sequence after FFT
Mean Power	AvePowACCX, AvePowACCY, AvePowACCZ, AvePowANGX, AvePowANGY, AvePowANGZ, AvePowVELX, AvePowVELY, AvePowVELZ	Average power of the power spectrum calculated for each data sequence after FFT

each pattern of the dataset is represented by a N-tuple of real numbers

$$\mathbf{x}_m = [x_{m1},\ x_{m2},\ \ldots, x_{mN}]\ ,\ m = 1, \ldots, M\,. \qquad (1)$$

Since the data features are completely heterogeneous, the patterns cannot be normalized in a global way, but on each column independently

$$x_{mj} \longleftarrow \frac{x_{mj} - b_j}{a_j - b_j}\ ,\ m = 1, \ldots, M\ ,\ j = 1, \ldots, N \qquad (2)$$

where $a_j = \underset{m=1,\ldots,M}{\arg\max}\ \{x_{mj}\}$ and $b_j = \underset{m=1,\ldots,M}{\arg\min}\ \{x_{mj}\}$.

In cross-validation test, a real implementation condition is simulated where new subjects must be classified with the knowledge obtained only from the training set. Even if the outcome of classification is unknown for the subject to be classified, the real measures associated with the patients exercise are known before the training and classification process. Therefore, the above preprocessing procedures, including the PCA and normalization, can be repeated for any new patient, and the trained classifier can be applied to any held out subjects in the test.

III. PROPOSED FUZZY MOTION CLASSIFIER

The upper limb motions classification problem in poststroke patients has been solved using the proposed model shown in Fig. 3, which will be referred to in the following as fuzzy kernel classifier (FKC). After data normalization, the classifier evaluates a label vector $\mathbf{L}$, for any pattern $\mathbf{x}$

$$\mathbf{L} = \left[\mu_{(1)}(\mathbf{x})\ \ \mu_{(2)}(\mathbf{x}), \ldots, \mu_{(K)}(\mathbf{x})\right] \qquad (3)$$

where the kth element of $\mathbf{L}$ represents the fuzzy MF of the pattern to the kth class. We adopt the WTA criterium for decision

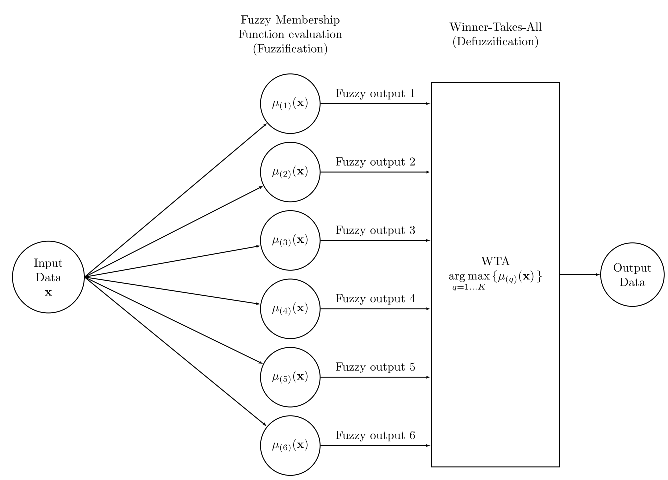

Fig. 3. Structure of the proposed classifier.

making; therefore, we chose the maximum value in the vector $\mathbf{L}$ and assign the final crisp label representing the estimated motion class for pattern $\mathbf{x}$.

The first step is to determine the training set which is a matrix where each row represents a vector of features to be used for the model learning. The proposed classifier establishes a set of fuzzy MFs to associate the patterns of each motion to the corresponding class. The used MFs, based on the geometrical representation of the data and point-to-polygon distance evaluation, have been presented in [34]. These MFs are constructed by taking regular polygons which cover all the patterns of each class and H kernel functions on both the vertices and the centroid of the corresponding polygon. The mathematical

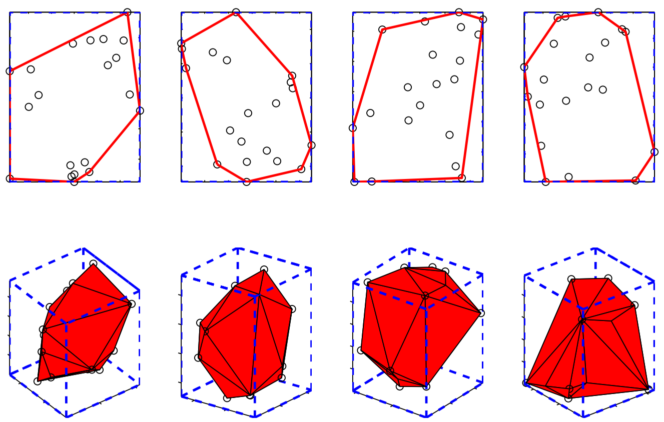

Fig. 4. Example in 2-D and 3-D of the convex hull to evaluate the points for placing the kernel of the MF.

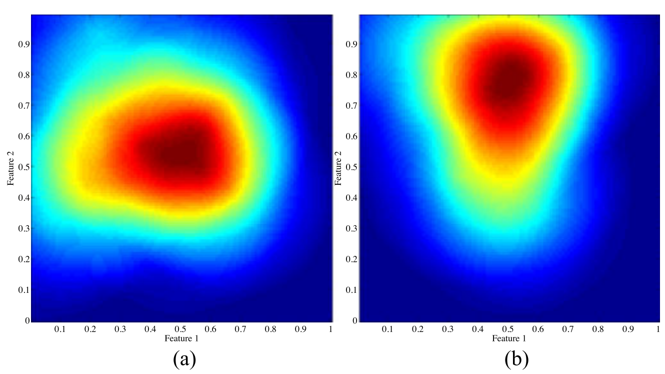

Fig. 5. Examples of the adopted MF visualized in 2-D plane with two different sets of randomly selected samples.

representation of the process is shown as follows:

$$\mu^{(\mathrm{cone})}(\mathbf{x}) = \max\left[0, 1 - \frac{\gamma}{\sqrt{N}} d_2(\mathbf{x}, \mathbf{c})\right] + \sum_{i=1}^{H} \max\left[0, 1 - \frac{\gamma}{\sqrt{N}} d_2(\mathbf{x}, \mathbf{v}_i)\right] \quad (4)$$

where N is the number of dimensions, $d_2(\mathbf{x}, \mathbf{c})$ is the point-to-centroid Euclidean distance, $d_2(\mathbf{x}, \mathbf{v}_i)$ is the point-to-ith-vertex Euclidean distance, and γ is the parameter that define the slope of the MF.

During data analysis if the number of dimensions is high, the identification of all the vertices of a polytope will lead to very high computational cost. The convex hull of a set of points is the smallest convex set that contains the points [41] and it is considered as an effective solution for real-time polygon boundary evaluation in many scientific disciplines [42]–[44]. For these reasons, as demonstrated in Fig. 4, we used this mathematical algorithm for the evaluation of the vertices in which it is possible to place the linear function during the MFs construction [45]. The unconstrained convex shape of the MF can be clearly observed in Fig. 5, where two MF examples, computed using ten randomly selected points from the $[0, 1]$ space, are visualized in 2-D planes.

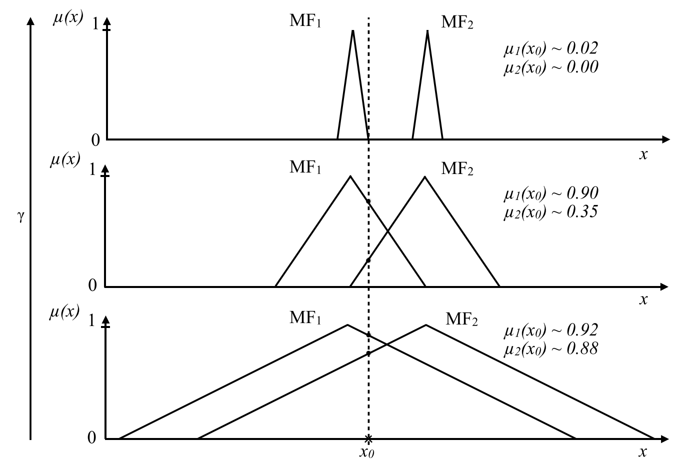

Fig. 6. Fuzzy MF evaluation with the γ parameter varying.

The γ parameter represents the sharpness of the MF: The higher the value of γ, the faster the function goes to zero, as depicted in Fig. 6. The optimization of this parameter helps the correct estimation of the pattern membership to each class. An excessively small value might result a critical class overlapping, while a large value may also cause indetermination as the area covered by MFs will be too small and the degree of membership to all the known classes could be very close to zero.

IV. Experiment Protocol

The experiments were conducted in collaboration with the Rehabilitation Medical Center of the 2nd Hospital of Jiaxing in China. All the experiment procedures and data access were approved by the ethics committees of the hospital and the university. After the informed consent and selection process, 14 stroke patients from the center including ten males and four females with an average age of 60.3 (ranged from 32 to 78) had participated in this research and in total 531 motion data were collected. The patient subjects were selected with a wide range of impairment level (Brunnstrom stage I–V) to test if the system was suitable for various stages of rehabilitation training. During the selection process, the participants were examined by the experienced doctors and the ones with severe cognitive, perceptual or communication problem, or any other health conditions that could have been not suitable for the experiment has been carefully excluded.

The motion sampling process utilized six common classical upper extremity rehabilitation exercises which are Bobath handshake, straight arm palm press, shoulder horizontal flexion and extension, forehead reaching with elbow, shoulder touching, and wrist turn. The motions are selected as they are the most frequently performed exercises in the hospital which are familiar to the patients and they are also able to cover the different perspectives of patients' motion impairments including multiple joint flexibility, muscle strength, and spasticity. The patients were asked to follow a video demo which repeats ten times for each exercise and no additional assistance was provided except for safety reasons. The participants were encouraged to attempt all six motions; however, due to the difference in impairment severity, most of the patients at low Brunnstrom stage (I–III) were not able to complete every exercise. Despite some motions were badly performed, every complete motion samples from all the 14 patients were included in the dataset to ensure the validity of the test result. However, six patients who had better performance during the exercise were selected to form a separate dataset in order to demonstrate the influence of motion quality on classification result.

V. Results and Discussion

The classification performance is evaluated in a tenfold cross-validation process, where ten rounds of classification are performed with different pairs of data subsets for training and testing. For each round, the classifier is trained with the patterns from the training set only, and the number of incorrectly classified motions is evaluated using the held out patterns from the test set. The original data are randomly partitioned into ten subsets with equal sizes and it is guaranteed that all classes are covered in each subset. Each fold uses a different subset for validation and the final result is determined using the average of the classification errors over the ten folds.

Two datasets with different groups of patients involved are used in the performance evaluation in order to demonstrate the influence of motion quality on classification accuracy:

1) Group with six patients (6P): The dataset consists of motion samples from six patients at higher Brunnstrom stages who are able to perform all the six motions with relatively high quality. The dataset includes 360 patterns with 60 in each of the six motion classes.
2) Group with 14 patients (14P): The dataset consists of motion samples from all 14 patients with large variation in motion quality due to the difference in impairment severity. The dataset includes 531 patterns with different number of samples in each of the six motion classes as some patients at low Brunnstrom stages have difficulties in completing certain movements.

The PCA analysis transforms the 63 features into principal components, which are ranked based on their percentage of contribution in terms of describing data variance. Theoretically, the motion classifier is able to work with any number of principal components as the more components are included, the more information from the original dataset will be covered. However, most of the variance is explained in the first few principal components and by adding extra components, the computation cost will rise significantly.

In order to demonstrate the sensitivity of the results of the proposed algorithm to its main parameters, we performed a grid search procedure by varying the number of principal components from 2 to 7 and the value of γ in a practically computable range between 0.2 and 20 with a 0.2 step. Any further increase above this range would result in the separation of the kernels within the MFs, and therefore, it would impair the performance of the classifier. The error rate versus the value of γ and the number of principal components are plotted in Figs. 7 and 8 for 6P and 14P datasets, respectively. It can be seen that the error rate drops significantly, while more information are recovered by adding extra principal component for both dataset.

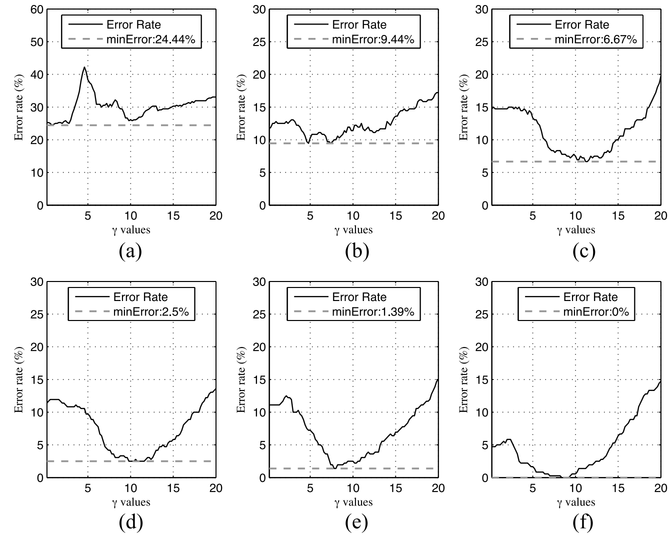

Fig. 7. Six patients classification. Error rate obtained with varying γ from 0.2 to 20 with a step of 0.2 and varying the number of the principal components used: (a) (2, cumulative variance (CV) = 21.61%), (b) (3, CV = 36.46%), (c) (4, CV = 48.30%), (d) (5, CV = 57.77%), (e) (6, CV = 68.86%), and (f) (7, CV = 72.33%).

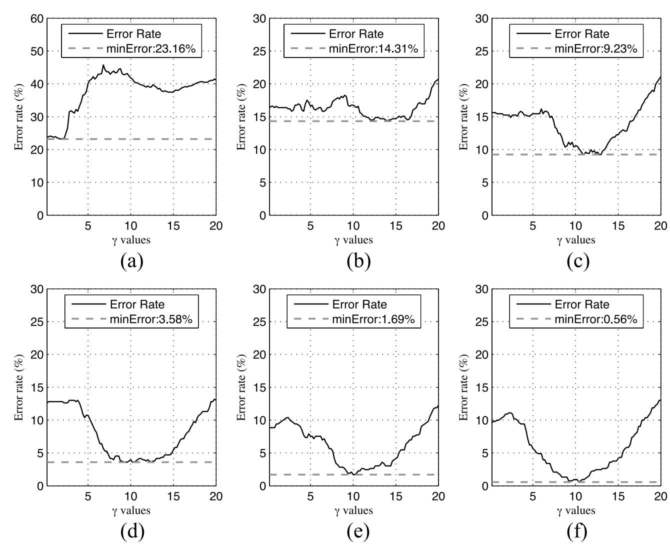

Fig. 8. Fourteen patients classification. Error rate obtained with varying γ from 0.2 to 20 with a step of 0.2 and varying the number of the principal components used: (a) 2, (b) 3, (c) 4, (d) 5, (e) 6, and (f) 7.

A $0\,\%$ of error rate was finally achieved using seven principal components and when $\gamma^* = 8.2$. In the case with 14P dataset as shown in Fig. 8, where the low quality motion samples are introduced, some performance deterioration can be observed compared to 6P group, although less than 1% error rate was also achieved with the same number of principal components. The γ value for the best result was slightly shifted to 10.2, which indicates that the shape of the MF boundary are optimized to be steeper to accommodate the additional class overlapping due to the increase of uncertainties. The detailed classification results for 14P dataset using $\gamma^* = 10.2$ can be found in the confusion matrix listed in Table II, which shows for each row the number

TABLE II
CONFUSION MATRIX FOR THE CLASSIFICATION TEST USING 14 PATIENTS

	Estimated outcome					
Actual value	M1*	M2*	M3*	M4*	M5*	M6*
M1	141	0	0	0	0	0
M2	0	80	0	0	0	0
M3	0	0	90	0	0	0
M4	1	0	1	58	0	0
M5	0	0	1	0	79	0
M6	0	0	0	0	0	80

TABLE III
ERROR RATE COMPARISON BETWEEN DIFFERENT CLASSIFIERS WITH 95% CONFIDENCE ERROR BOUNDS

Algorithm used	6 Patients	14 Patients
FKC	0.00	0.56 ± 0.64
Neurofuzzy classifier	1.67 ± 1.32	1.70 ± 1.10
Classification tree (CART)	5.28 ± 2.31	4.90 ± 1.84
SVM	1.32 ± 1.18	1.11 ± 0.89
LDA	1.67 ± 1.32	8.35 ± 2.35
QDA	0.56 ± 0.77	1.32 ± 0.97
Naive Bayes	3.06 ± 1.78	6.03 ± 2.02
PNN	11.11 ± 3.25	67.23 ± 4.00

All the values are expressed in (%).

of patterns of every real motion $(\mathrm{M1}, \mathrm{M2}, \ldots, \mathrm{M6})$ that are assigned to the estimated outcome $(\mathrm{M1}^*, \mathrm{M2}^*, \ldots, \mathrm{M6}^*)$.

A number of widely used classification algorithms, trained in the MATLAB software (version R2013a), have also been tested for comparison in order to confirm that we can obtain a smaller error rate using our fuzzy approach. The first one is a neurofuzzy classifier whose parameters are adapted by the scaled conjugate gradient method [46]. A decision classification tree (CART) [47] is used as a classic approach, together with the famous SVM by using a RBF kernel with the sequential minimal optimization method. Moreover, the following probabilistic classifiers have been tested: Linear discriminant analysis (LDA) and quadratic discriminant analysis (QDA) [48]; naive Bayes classifier [49]; the feed-forward probabilistic neural network (PNN) [50].

The same datasets, preprocessing techniques, and experimental procedures are used for comparing the different classification algorithms. A grid search is applied by varying the inner parameters of each training algorithm and the number of principal components in a suitable range. The best tenfold cross-validation result, obtained for each algorithm within the grid search, is reported in Table III. QDA provides a natural benchmark comparison to the fuzzy classifier. Therefore, in Fig. 9, the QDA performance is also depicted as a function of number of principal components.

It can be seen from the result that the proposed FKC has a clear advantage over other popular methods in terms of accuracy for poststroke rehabilitation motion classification. SVM which has been widely adopted in motion tracking applications can be treated as a benchmark [51], [52] and close to 1% error rate was achieved after tuning a RBF kernel SVM. Despite the accuracy, the implementation of SVM in this application can

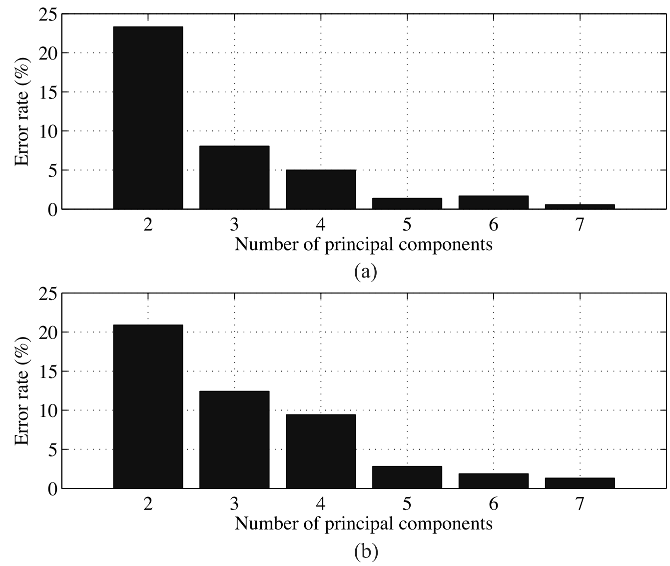

Fig. 9. QDA performance as a function of number of principal components for both (a) 6P and (b) 14P.

be limited as the choice of kernel can dramatically affect the performance. The computation time, which multiplies rapidly as the size of the dataset increases, is another factor which hinders its implementation [53]. On the other hand, the proposed FKC requires only a single γ parameter to be tuned in the validation phase and the computation time is relatively quicker especially in testing phase once the convex set is determined.

VI. CONCLUSION

This paper has presented a novel fuzzy kernel motion classifier which is specifically designed for autonomous stroke rehabilitation applications. The system is capable of accurately identifying common rehabilitation exercise movements during a stroke patient's routine training session using the kinetic data collected through a IMU attached to patient's wrist. By implementing the proposed motion classifier in a telerehabilitation training system, the doctors will be able to track patients training performance remotely without having to follow through the entire session. In addition to monitor patients performance and adherence, the motion classification can also provide significant support for other telerehabilitation features. For example, many pattern-recognition-based automatic impairment level and body function evaluation system relies on the accurate identification of certain predefined movements in order to realize unsupervised implementation. Template matching-based motion evaluation approaches can also benefit from the proposed motion classifier, which can potentially reduce the indexing effort of the template database.

The advantage of the proposed classification approach is twofold: to obtain a multivalued classification label representing the different degrees of reliability in classifying every type of movements for a patient; to improve the decision making process by WTA strategy in order to obtain better experimental results in assigning a single outcome.The optimized geometrically unconstrained MFs adopted in the classifier can effectively manage the motion class overlapping issue which is one of the major obstacles in classification problems especially when dealing with irregular motion samples performed by stroke patients with various degree of body functionality impairment.

The proposed classifier has also undergone a series of validation experiments and the results have demonstrated superior performances compared to other popular pattern recognition algorithms. When including no less than seven features as input, extracted by means of PCA, the fuzzy kernel motion classifier is able to achieve 0% error rate for low impairment level patient group and 0.56% for all patients. The future work will focus on further improving the model by including more motion types and samples, and integrating the motion classifier into an autonomous rehabilitation training system which can provide high quality stroke rehabilitation training outside hospital with no or reduced supervision.

References

[1] WHO, "Stroke—1989. Recommendations on stroke prevention, diagnosis, and therapy. Report of the WHO task force on stroke and other cerebrovascular disorders," *Stroke*, vol. 20, pp. 1407–1431, Oct. 1989.

[2] National Stroke Foundation, "National stroke audit—Rehabilitation services report 2012," *National Stroke Foundation*, Melbourne, Australia, 2012.

[3] M. H. Feigin, V. R. Forouzanfar, G. A. Krishnamurthi, M. Mensah, D. A. Connor, A. E. Bennett, R. L. Moran, L. Sacco, T. Anderson, M. Truelsen, N. O'Donnell, S. Venketasubramanian, C. M. BarkerCollo, W. Lawes, Y. Wang, E. Shinohara, M. Witt, M. Ezzati, C. Naghavi, and Murray, "Global and regional burden of stroke during 1990–2010: Findings from the global burden of disease study 2010," *Lancet*, vol. 383, no. 9913, pp. 245–255, 2014.

[4] National Stroke Foundation, "Clinical guidelines for stroke management 2010," *National Stroke Foundation*, Melbourne, Australia, 2010, pp. 78–95.

[5] D. A. Cadilhac, R. Carter, A. G. Thrift, and H. M. Dewey, "Estimating the long-term costs of ischemic and hemorrhagic stroke for Australia: New evidence derived from the north east Melbourne stroke incidence study (NEMESIS)," *Stroke*, vol. 40, no. 9913, pp. 915–921, 2009.

[6] G. E. Gresham, P. W. Duncan, W. B. Stason, "Post-stroke Rehabilitation: Clinical Practice Guideline, 16," Rockville, MD: U.S. Department of Health and Human Services. Public Health Service, Agency for Health Care Policy and Research, AHCPR Publication no. 95-0662, 1995.

[7] Intercollegiate Stroke Working Party, "National Clinical Guidelines for Stroke (2nd edition.)," Royal College of Physicians of London, London, 2000.

[8] R. Teasell, M. Meyer, A. McClure, C. Pan, M. Murie-Fernandez, N. Foley, and K. Salter, "Stroke rehabilitation: An international perspective," *Topics Stroke Rehabil.*, vol. 16, no. 1, pp. 44–56, 2009.

[9] Y. Chen, Y. Zhou, X. Cheng, and Mi, "Upper limb motion recognition based on two-step SVM classification method of surface EMG," *Int. J. Control Autom.*, vol. 6, no. 3, pp. 249–266, 2013.

[10] Y. Geng, L. Zhang, D. Tang, X. Zhang, and G. Li, "Pattern recognition based forearm motion classification for patients with chronic hemiparesis," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, 2013, pp. 5918–5921.

[11] K. Kaczmarczyk, A. Wit, M. Krawczyk, and J. Zaborski, "Gait classification in post-stroke patients using artificial neural networks," *Gait Posture*, vol. 30, no. 2, pp. 207–210, 2009.

[12] N. Yang, M. Zhang, C. Huang, and D. Jin, "Motion quality evaluation of upper limb target-reaching movements," *Med. Eng. Phys.*, vol. 24, no. 2, pp. 115–120, 2002.

[13] Z. Zhang, Q. Fang, and F. Ferry, "Upper limb motion capturing and classification for unsupervised stroke rehabilitation," in *Proc. IEEE Ind. Electron. Soc. Conf.*, 2011, pp. 3832–3836.

[14] Z. Zhang, Q. Fang, L. Wang, and P. Barrett, "Template matching based motion classification for unsupervised post-stroke rehabilitation," in *Proc. Int. Symp. Bioelectron. Bioinformat.*, 2011, pp. 199–202.

[15] H. Ghasemzadeh, R. Jafari, and B. Prabhakaran, "A body sensor network with electromyogram and inertial sensors: Multimodal interpretation of muscular activities," *IEEE Trans. Inform. Technol. Biomed.*, vol. 14, no. 2, pp. 198–206, Mar. 2010.

[16] H. Zhou and H. Hu, "Human motion tracking for rehabilitation—A survey," *Biomed. Signal Process. Control*, vol. 3, no. 1, pp. 1–18, 2008.

[17] H. Zhou, O. Hu, and N. Harris, "Application of wearable inertial sensors in stroke rehabilitation," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, 2005, vol. 7, pp. 6825–6828.

[18] S. Theodoridis and K. Koutroumbas, *Pattern Recognition: Theory and Applications*, 4th ed. New York, NY, USA: Academic, 2009.

[19] C. M. Bishop, *Pattern Recognition and Machine Learning (Information Science and Statistics)*. Secaucus, NJ, USA: Springer-Verlag, 2006.

[20] O. Urra, A. Casals, and R. Jana, "Evaluating spatial characteristics of upper-limb movements from EMG signals," in *XIII Mediterranean Conference on Medical and Biological Engineering and Computing*, L. M. Roa Romero, Ed. New York, NY, USA: Springer-Verlag, 2014, vol. 41, pp. 1795–1798.

[21] E. Sazonov, G. Fulk, N. Sazonova, and S. Schuckers, "Automatic recognition of postures and activities in stroke patients," in *Proc. IEEE Eng. Med. Biol. Soc. Conf.*, 2009, pp. 2200–2203.

[22] Z. Yan, Z. Wang, and H. Xie, "The application of mutual information-based feature selection and fuzzy LS-SVM-based classifier in motion classification," *Comput. Methods Prog. Biomed.*, vol. 90, no. 3, pp. 275–284, 2008.

[23] D. Tsujinishi and S. Abe, "Fuzzy least squares support vector machines for multiclass problems," *Neural Netw.*, vol. 16, nos. 5/6, pp. 785–792, 2003.

[24] R. Duda, P. Hart, and D. Stork, *Pattern Classification*. New York, NY, USA: Wiley, 2001.

[25] R. Parisi, A. Cirillo, M. Panella, and A. Uncini, "Source localization in reverberant environments by consistent peak selection," in *Proc. IEEE Int. Conf. Acoust., Speech, Signal Process.*, Apr. 2007, vol. 1, pp. I-37–I-40.

[26] L. Zadeh, "Fuzzy sets," *Inform. Control*, vol. 8, pp. 338–353, 1965.

[27] L. A. Zadeh, "Similarity relations and fuzzy orderings," *Inform. Sci.*, vol. 3, no. 2, pp. 177–200, 1971.

[28] D. Nauck and R. Kruse, "A neuro-fuzzy method to learn fuzzy classification rules from data," *Fuzzy Sets Syst.*, vol. 89, no. 3, pp. 277–288, 1997.

[29] H. Ishibuchi, T. Nakashima, and T. Murata, "Performance evaluation of fuzzy classifier systems for multidimensional pattern classification problems," *IEEE Trans. Syst., Man, Cybern., B, Cybern.*, vol. 29, no. 5, pp. 601–618, Oct. 1999.

[30] M. Panella and G. Martinelli, "Neurofuzzy networks with nonlinear quantum learning," *IEEE Trans. Fuzzy Syst.*, vol. 17, no. 3, pp. 698–710, May 2009.

[31] M. Panella, "A hierarchical procedure for the synthesis of anfis networks," *Adv. Fuzzy Syst.*, vol. 2012, pp. 1–12, 2012.

[32] M. O'Malley, M. Abel, D. Damiano, and C. Vaughan, "Fuzzy clustering of children with cerebral palsy based on temporal- distance gait parameters," *IEEE Trans. Rehabil. Eng.*, vol. 5, no. 4, pp. 300–309, Dec. 1997.

[33] F. Chan, Y.-S. Yang, F. Lam, Y.-T. Zhang, and P. Parker, "Fuzzy EMG classification for prosthesis control," *IEEE Trans. Rehabil. Eng.*, vol. 8, no. 3, pp. 305–311, Sep. 2000.

[34] L. Liparulo, A. Proietti, and M. Panella, "Fuzzy membership function based on point-to-polygon distance evaluation," in *Proc. IEEE Conf. Fuzzy Syst.*, Hyderabad, India, 2013, pp. 1–8.

[35] L. Liparulo, A. Proietti, and M. Panella, "Fuzzy clustering using the convex hull as geometrical model," *Adv. Fuzzy Sys.*, vol. 2015, art. no. 265135 (13 pages), 2015

[36] L. Liparulo, A. Proietti, and M. Panella, "Improved online fuzzy clustering based on unconstrained kernels," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, Aug. 2015, pp. 1–8.

[37] L. Cao, K. Chua, W. Chong, H. Lee, and Q. Gu, "A comparison of PCA, KPCA and ICA for dimensionality reduction in support vector machine," *Neurocomputing*, vol. 55, nos. 1/2, pp. 321–336, 2003.

[38] R. Panerai, A. Ferreira, and O. Brum, "Principal component analysis of multiple noninvasive blood flow derived signals," *IEEE Trans. Biomed. Eng.*, vol. 35, no. 7, pp. 533–538, Jul. 1988.

[39] H. Sadeghi, F. Prince, S. Sadeghi, and H. Labelle, "Principal component analysis of the power developed in the flexion/extension muscles of the hip in able-bodied gait," *Med. Eng. Phys.*, vol. 22, no. 10, pp. 703–710, 2000.

[40] A. Subasi and M. Gursoy, "EEG signal classification using PCA, ICA, LDA and support vector machines," *Expert Syst. Appl.*, vol. 37, no. 12, pp. 8659–8666, 2010.

[41] D. T. Lee and F. P. Preparata, "Computational geometry—A survey," *IEEE Trans. Comput.*, vol. C-33, no. 12, pp. 1072–1101, Dec. 1984.

[42] J. Sklansky, “Measuring concavity on a rectangular mosaic,” *IEEE Trans. Comput.*, vol. C-21, no. 12, pp. 1355–1364, Dec. 1972.
[43] M. Maisto, M. Panella, L. Liparulo, and A. Proietti, “An accurate algorithm for the identification of fingertips using an RGB-D camera,” *IEEE Trans. Circuits Syst.*, vol. 3, no. 2, pp. 272–283, Jun. 2013.
[44] W. Guiliang, D. Xiangqian, L. Bo, and X. Peiyong, “A new semi-supervised clustering method based on convex hull,” in *Proc. Int. Conf. Intell. Comput. Technol. Autom.*, 2011, vol. 2, pp. 1036–1038.
[45] C. B. Barber, D. P. Dobkin, and H. Huhdanpaa, “The quickhull algorithm for convex hulls,” *ACM Trans. Math. Softw.*, vol. 22, no. 4, pp. 469–483, Dec. 1996.
[46] M. F. Moller, “A scaled conjugate gradient algorithm for fast supervised learning,” *Neural Netw.*, vol. 6, no. 4, pp. 525–533, 1993.
[47] L. Rokach and O. Maimon, *Data Mining with Decision Trees: Theory and Applications*. River Edge, NJ, USA: World Scientific, 2008.
[48] W. J. Krzanowski, Ed., *Principles of Multivariate Analysis: A User’s Perspective*. New York, NY, USA: Oxford Univ. Press, 1988.
[49] I. Rish, “An empirical study of the naive Bayes classifier,” *American Association for Artificial Intelligence*, 2001, pp. 41–46.
[50] D. F. Specht, “Probabilistic neural networks,” *Neural Netw.*, vol. 3, no. 1, pp. 109–118, Jan. 1990.
[51] R. Begg and J. Kamruzzaman, “A machine learning approach for automated recognition of movement patterns using basic, kinetic and kinematic gait data,” *J. Biomech.*, vol. 38, no. 3, pp. 401–408, 2005.
[52] C. Schldt, I. Laptev, and B. Caputo, “Recognizing human actions: A local SVM approach,” in *Proc. Int. Conf. Pattern Recog.*, 2004, vol. 3, pp. 32–36.
[53] C. Burges, “A tutorial on support vector machines for pattern recognition,” *Data Mining Knowl. Discovery*, vol. 2, no. 2, pp. 121–167, 1998.

Zhe Zhang (S’10) was born in Tianjin, China, in 1988. He received the B.E. degree (Hons., 1st class) in electrical and electronic engineering from RMIT University, Melbourne, Australia, in 2010, where he is currently working toward the Ph.D. degree with the School of Electrical and Computer Engineering.

His major research interests include wearable sensor network, signal processing, and pattern recognition applied in the field of rehabilitation and biomedical engineering.

Luca Liparulo (S’13) was born in Campobasso, Italy, in 1983. He received the B.Sc. degree in telecommunications, the M.Sc. degree in communications engineering, and the Ph.D. degree in information and communication engineering all from University “La Sapienza,” Rome, Italy, in 2008, 2011, and 2015, respectively.

He is currently a Postdoctoral Research Fellow at the University of Rome “La Sapienza,” Rome. His research interests include the application of fuzzy logic, neural networks, and computational intelligence techniques in complex systems optimization.

Massimo Panella (M’01) was born in Rome, Italy, in 1971. He received the five-year Laurea degree (*summa cum laude*) in electronic engineering and the Ph.D. degree in information and communication engineering, both from University “La Sapienza,” Rome, in 1998 and 2002, respectively.

He is an Assistant Professor of circuit theory and learning systems for pervasive computing at the University of Rome “La Sapienza.” His research interests include circuit theory, computational intelligence, and pervasive computing for modeling, optimization, and control of complex systems.

Xudong Gu received the MBChB degree from Nanjing Medical University, Jiangsu, China, in 1989, and the M.R.M (Specialist) and M.R.M degrees from Zhejiang Provincial Peoples Hospital, Zhejiang, China, and the University of Wollongong, Wollongong, Australia, in 1997 and 2004, respectively.

He is currently the Chairman of the Rehabilitation Medical Centre, Jiaxing 2nd Hospital, Zhejiang, and a Professor at Jiaxing University, Zhejiang. He is a Standing Committee Member of Chinese Medical Doctor Association, Rehabilitation Doctor Association, and an Editor of the Chinese Journal of Physical Medicine and Rehabilitation.

Qiang Fang (M’99) received the B.S. degree in applied physics form Tsinghua University, Beijing, China, in 1991, and the Ph.D degree in biomedical engineering from Monash University, Melbourne, Australia, in 2000.

He joined RMIT University, Melbourne, Australia, in 2000, as a Permanent Academic Staff, and is currently leading a research team in wearable and implantable technologies applicable to rehabilitation and neurology. His major research interests include intelligent and miniaturized medical instrumentation, wearable and implantable body sensor network, and pervasive computing technologies.