# Simultaneous Recognition and Assessment of Post-Stroke Hemiparetic Gait by Fusing Kinematic, Kinetic, and Electrophysiological Data

Chengkun Cui, Gui-Bin Bian, Zeng-Guang Hou, *Senior Member, IEEE*, Jun Zhao, Guodong Su, Hao Zhou, Liang Peng, and Weiqun Wang

***Abstract*— Gait analysis for the patients with lower limb motor dysfunction is a useful tool in assisting clinicians for diagnosis, assessment, and rehabilitation strategy making. Implementing accurate automatic gait analysis for the hemiparetic patients after stroke is a great challenge in clinical practice. This study is to develop a new automatic gait analysis system for qualitatively recognizing and quantitatively assessing the gait abnormality of the post-stroke hemiparetic patients. Twenty-one post-stroke patients and twenty-one healthy volunteers participated in the walking trials. Three of the most representative gait data, i.e., marker trajectory (MT), ground reaction force (GRF), and electromyogram, were simultaneously acquired from these subjects during their walking. A multimodal fusion architecture is established by using these different modal data to qualitatively distinguish the hemiparetic gait from normal gait by different pattern recognition techniques and to quantitatively estimate the patient's lower limb motor function by a novel probability-based gait score. Seven decision fusion algorithms have been tested in this architecture, and extensive data analysis experiments have been conducted. The results indicate that the recognition performance and estimation performance of the system become better when more modal gait data are fused. For the recognition performance, the random forest classifier based on the GRF data achieves an accuracy of 92.26% outperformed other single-modal schemes. When combining two modal data, the accuracy can be enhanced to 95.83% by using the support vector machine (SVM) fusion algorithm to fuse the MT and GRF data. When integrating all the three modal data, the accuracy can be further improved to 98.21% by using the SVM fusion algorithm. For the estimation performance, the absolute values of the correlation coefficients between the estimation results of the above three schemes and the Wisconsin gait scale scores for the post-stroke patients are 0.63, 0.75, and 0.84, respectively, which means the clinical relevance becomes more obvious when using more modalities. These promising results demonstrate that the proposed method has considerable potential to promote the future design of automatic gait analysis systems for clinical practice.**

***Index Terms*— Post-stroke hemiparesis, gait analysis, multimodal fusion, marker trajectory (MT), ground reaction force (GRF), electromyogram (EMG).**

Manuscript received November 8, 2017; revised February 11, 2018; accepted February 12, 2018. Date of publication March 2, 2018; date of current version April 6, 2018. This work was supported in part by the National Natural Science Foundation of China under Grant 61720106012, Grant U1713220, Grant 61533016, Grant 61603386, Grant 91648208, and Grant 61421004, and in part by the Beijing Natural Science Foundation under Grant L172050. *(Corresponding author: Zeng-Guang Hou.)*

C. Cui is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, and also with the University of Chinese Academy of Sciences, Beijing 100049, China (e-mail: cuichengkun2014@ia.ac.cn).

G.-B. Bian, L. Peng, and W. Wang are with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China (e-mail: guibin.bian@ia.ac.cn; liang.peng@ia.ac.cn; weiqun.wang@ia.ac.cn).

Z.-G. Hou is with the State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China, also with the University of Chinese Academy of Sciences, Beijing 100049, China, and also with the CAS Center for Excellence in Brain Science and Intelligence Technology, Beijing 100190, China (e-mail: zengguang.hou@ia.ac.cn).

J. Zhao, G. Su, and H. Zhou are with the Beijing Bo'ai Hospital, China Rehabilitation Research Center, Beijing 100068, China (e-mail: zaojun@aliyun.com; ZKPTSGD@126.com; halfsunny@foxmail.com).

Digital Object Identifier 10.1109/TNSRE.2018.2811415

## I. Introduction

POST-STROKE hemiparesis is a condition usually caused by cerebrovascular blockage or rupture that affects motion control of the entire left or right side of the body. Gait impairment is a common problem in post-stroke hemiparetic patients, which greatly reduces their quality of life, because walking ability is very important in conducting many daily activities. Hemiparetic gait is characterized by shorter stride length, slower velocity, slower cadence, spatial and temporal left-right asymmetry, and more energy consumption in comparison with normal gait [1]. The motor rehabilitation treatment after hemiparesis is essential to help the patients recover their motor function. In order to make appropriate treatment strategies for the patients with stroke, it is very crucial to accurately recognize and assess their gait abnormality. The automatic gait analysis can become an effective solution, which has already caught much attention in clinical diagnosis and assessment [2]–[7].

In recent years, advanced measuring techniques have made it easier to collect the kinematic, kinetic or electrophysiological data during walking [8]–[10], which greatly facilitates the development of automatic gait analysis. Many researchers focused on the recognition problem of gait abnormality, such as [11]–[19]. Different pattern recognition algorithms were adopted to construct classification systems to automatically identify the pathological or normal gait patterns based on the recorded gait data. The trained classification systems had some generalization ability to unknown gait data, which could provide some effective information for clinical diagnosis. However, classifying different gait patterns into two or more categories is essentially a qualitative analysis method, which is not subtle enough to produce valuable information for the assessment of patients' lower limb motor function.

Several studies have explored and designed different gait scores (e.g., the GGI [20], GDI [21], GPS [22], GVI [23] and COGS [24]) to implement the quantitative analysis of the pathological gait patterns. These gait scores essentially measured the similarity between a subject's gait pattern and the average gait pattern of a reference group. However, most gait scores were established on some mathematical expressions with pre-set fixed parameters, which might cause poor generalization performance on unknown gait data. Furthermore, these gait scores were usually calculated based on only a single modal data source (e.g., kinematic data), which could only summarize a subject's gait characteristics from a particular aspect and might lead to a one-sided assessment result.

According to the above exposition, the qualitative and quantitative analysis methods may be combined to develop a more powerful automatic gait analysis system. Meanwhile, different modal gait data, including kinematic, kinetic and electrophysiological data, need to be collected and integrated to achieve a more comprehensive analysis result. In this way, it is possible to embed the outcome of the automatic gait analysis into the conventional clinical decisions.

This paper further extends and develops our previous work [25] to a new field: automatic gait analysis of post-stroke hemiparetic patients. A data-driven-based multimodal fusion architecture is proposed to qualitatively recognize the hemiparetic gait, and quantitatively estimate the patients' walking ability. Three different modal time-varying dynamic data, i.e., marker trajectory (MT), ground reaction force (GRF) and electromyogram (EMG), were synchronously recorded from the subjects from both pathological and normal groups. These multimodal data are served as the inputs of the fusion architecture. After preprocessing and feature extraction, different modal data are fed to their individual classifiers, respectively. Next, the outputs of these classifiers are combined by a fusion algorithm to achieve the final analysis results, which contain the qualitative predicted class and quantitative estimated score.

The main contributions of this study include:

1) A multimodal fusion architecture based on integrating kinematic, kinetic and electrophysiological data is designed to qualitatively and quantitatively analyze the post-stroke hemiparetic gait abnormality.
2) A novel probability-based gait score is defined to quantify the differences between the post-stroke hemiparetic gait and normal gait, and can be used to implement the quantitative assessment of the post-stroke patients' walking ability.
3) Seven decision fusion algorithms are adopted and tested in the proposed architecture, respectively. To the best of our knowledge, most of them have never been studied in the field of automatic gait analysis.

TABLE I
DEMOGRAPHIC DATA OF THE SUBJECTS

| Variables | | Pathological group | Normal group | P |
|---|---|---|---|---|
| Age (years) | | 47.88±12.32 | 47.38±11.09 | 0.89 |
| Height (cm) | | 169.05±6.10 | 170.95±6.50 | 0.33 |
| Weight (kg) | | 71.86±10.83 | 70.57±11.21 | 0.71 |
| Gender | Male | 16 | 16 | 1.00 |
| | Female | 5 | 5 | |

## II. METHODS

### A. Subjects

This study recruited twenty-one post-stroke hemiparetic patients from the China Rehabilitation Research Center, Beijing, China, as the pathological group. At the same time, twenty-one matched healthy volunteers were enrolled as the normal reference group.

For the pathological group, the inclusion criteria included: 1) a first stroke with unilateral hemiparesis, 2) unilateral hemispheric lesions confirmed by computed tomography or magnetic resonance imaging, 3) age 20-75 years, 4) ability to walk at least 10 m without any help or assistive devices, 5) ability to understand verbal commands and to cooperate with the experimental procedures, 6) no other diagnosed diseases known to affect walking performances. For the normal group, the inclusion criteria were no history of any neurological or musculoskeletal disorders. Exclusion criteria for both groups were previous history of severe diseases about heart, lung, liver, kidney, etc.

Table I lists the demographic data of the two group subjects, including age, height, weight and gender. According to the different types of statistical data, the corresponding statistical test methods are adopted to analyze the significance of differences. For numerical data, such as age, height and weight, the unpaired t-tests are used to compare the differences in these data between post-stroke patients and healthy subjects, respectively. For categorical data, like gender, the chi-squared test is employed to examine the differences between groups. All the above statistical test methods are conducted by using SPSS (IBM Corp., USA) at a significant level of 0.05. From the rightmost column of Table I, we can clearly see that there are no significant differences ($P > 0.05$) in the demographic data between the pathological group and normal group.

This research was approved by the Ethics Committee of China Rehabilitation Research Center (approval number: 2016-059-1), and an informed consent form was signed by each subject.

### B. Measurements

All the experiments were carried out at the Gait Analysis Laboratory, China Rehabilitation Research Center. Three different types of gait data during human walking, including the kinematic, kinetic and electrophysiological data, were simultaneously recorded by a motion capture system and multiple sensors.

Fig. 1. Marker placement for static calibration: (a) anterior and (b) posterior views.

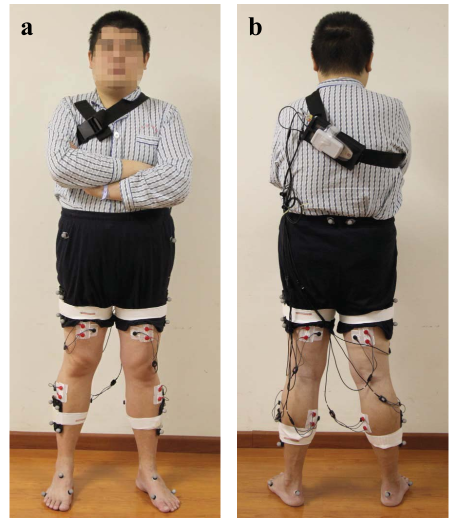

Fig. 2. Marker placement after static calibration and EMG electrode placement: (a) anterior and (b) posterior views.

The motion trajectories of the reflective markers placed on each subject's lower body were recorded by a six-camera Qualisys motion capture system (Qualisys AB, Sweden) at a sampling frequency of 200 Hz. Before the walking trials, a static calibration of the Qualisys system was conducted with a subject upright standing at the center of the floor for about 3 seconds to ensure that all the cameras could accurately record the positions of the markers. The purpose of this calibration was to locate the subject's bone segments and joint center positions in the system. During the static calibration, there were 38 markers glued to the subject's body (see Fig. 1) based on the Calibrated Anatomical System Technique (CAST) [26]. Specifically, these markers were placed on the anterior superior iliac spine, posterior superior iliac spine, greater trochanter, lateral thigh, medial and lateral femoral epicondyles, lateral shank, medial and lateral malleoli, calcaneus, first and fifth metatarsal heads, and second metatarsal base of both lower limbs. After the calibration, 10 markers were removed (greater trochanter, medial and lateral femoral epicondyles, medial and lateral malleoli on both sides of the body), and the other 28 markers were remained on the subject (see Fig. 2) to record dynamic limb motion data during the walking trials.

The GRF data were captured at 1000 Hz by using two Bertec force plates (Bertec Corp., USA) embedded in the floor (see Fig. 3). The foot strike and toe off events were determined by using a vertical GRF threshold of 10 N. In order to keep the subjects from deliberately changing their strides to strike the force plates, we used a black cloth to cover the force plates and surrounding area.

Eight channels of surface EMG signals were collected at 1000 Hz by a Biomonitor ME6000 (Mega Electronics Ltd., Finland). The monitored muscles involved the rectus femoris, long head of biceps femoris, tibialis anterior, medial head of gastrocnemius of both lower limbs. The EMG electrode placement was based on the human anatomical locations [27] (see Fig. 2). The acquisition processes of the above three different modal data were controlled and synchronized by the Qualisys Track Manager (Qualisys AB, Sweden).

Fig. 3. Force plates embedded in the floor.

The subjects were required to walk barefoot at their natural velocity. The starting point was adjusted for each subject to guarantee that he or she could make full contact with the force plates. Moreover, the subjects were guided to watch a big black dot on the opposite wall to transfer their attention away from the force plates covered by a black cloth, and to keep an upright body position during walking. In order to prevent large physical consumption and muscle fatigue, the subjects were given as much rest as they required between different trials. It is worth noting that all the walking trials were carried out under the guidance of an experienced physiotherapist. A walking trial is considered as a successful trial when a single full contact of a subject's each foot on the corresponding force plate is achieved. Full contact with the force plate can be confirmed visually in the Qualisys Track Manager. Each subject was required to accomplish 9 successful walking trials.

The subsequent analysis is conducted on the gait pattern from one full gait cycle. Since the two force plates are placed as shown in Fig. 3, a full gait cycle in this study is defined from left foot strike to the next left foot strike, during which the subject should make full contact with the two force plates in turns. Therefore, from each walking trial, only one full gait cycle can be extracted. The beginning and ending events

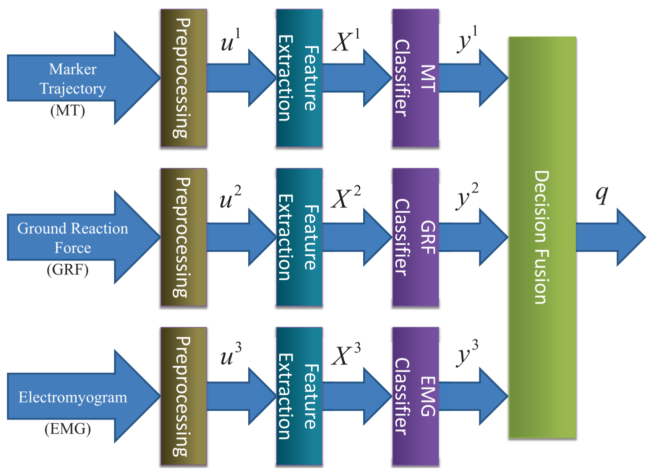


Fig. 4. Multimodal fusion architecture for automatic gait analysis.

of the gait cycle can be automatically detected and marked by using Visual 3D (C-Motion Inc., USA). In this work, an experimental sample is defined as the data collected during a full gait cycle. According to the subject's group, the ground truth label of each sample is labeled as 0 (pathological group) or 1 (normal group).

### C. Multimodal Fusion Architecture

The motivation of designing multimodal fusion architecture is to obtain more comprehensive and accurate analysis results by exploiting the complementarity between different modal gait data. In Fig. 4, we show the diagram of the proposed fusion architecture for automatic gait analysis. This architecture is established mainly based on the integration of semantic information from different modal gait data. As the figure presents, multimodal data are processed separately to yield individual predictions which can represent the semantic information of the corresponding modalities. Then, these individual predictions are combined later to derive the final prediction. The above description can be formulated as follows.

Let us define $u^m$ as an experimental sample of the $m$th modality after pre-processing, and define the feature vector extracted from $u^m$ as

$$X^m = [\xi_1^m, \xi_2^m, \ldots, \xi_n^m, \ldots, \xi_{N_m}^m]^T, \quad m = 1, 2, \ldots, M \qquad (1)$$

where $\xi_n^m$ is the $n$th feature of $X^m$. The feature vector from each modality is separately fed to the corresponding individual classifier to produce the individual predicted result, which can be denoted as

$$y^m = h^m(X^m) \qquad (2)$$

where $y^m$ represents the individual predicted probability achieved by the individual classifier $h^m$ of the $m$th modality. The individual predicted probability is a value containing semantic information of the corresponding modality, which describes the possibility that a subject's gait pattern belongs to a certain class. Of course, for different modal gait data, the obtained probabilities may be dissimilar. We hope to make the best use of the complementarity between the semantic information from different modal data. Therefore, these individual predicted probabilities are concatenated together to construct a semantic information vector. At the fusion stage, this semantic information vector is processed by a decision fusion algorithm to obtain the final predicted result, which can be expressed as

$$q = f(y^1, y^2, \ldots, y^M) \qquad (3)$$

where $q$ is the final predicted probability and $f$ is the corresponding fusion algorithm. A subject's gait pattern will be automatically assigned to the pathological group when the predicted probability is less than 0.5. It is worthwhile pointing out that this architecture can be easily extended to the situations in which more than three modal data are used in parallel.

The training process of the fusion architecture can be summarized as a four-step procedure. To begin with, the training dataset needs to be divided into two non-overlapping parts: individual-training dataset and development-training dataset. Then, the individual-training dataset and ground truth labels are utilized to build some candidate individual classifiers. Subsequently, the development-training dataset is fed to these candidate classifiers to find the best individual classifier for each modality by comparing the recognition performance. The outputs of these best individual classifiers on the development-training dataset are concatenated to form semantic information vectors. Finally, these semantic information vectors and ground truth labels are used to train the tunable parameters of the fusion algorithm. During the training procedure, we employ 3-fold cross validation for the determination of the hyper-parameters in some classification or fusion models.

As it is mentioned in Section II-B, for each subject, there are 9 full gait cycles extracted from 9 successful walking trials. Therefore, the data recorded in 378 gait cycles from 42 subjects are regarded as the full dataset. Since it is usually very difficult to get enough multimodal data for training the automatic gait analysis system in clinical practice, the size of the testing dataset is arranged larger than that of the individual-training or development-training dataset. For convenience, the data in the first three gait cycles of each subject are served as the individual-training dataset, and the data in the middle two gait cycles of each subject are used as the development-training dataset. The data in the remaining four gait cycles of each subject are employed to construct the testing dataset. Namely, the data division ratio is 3:2:4 for the individual-training, development-training and testing dataset.

### D. Data Analysis

*1) Preprocessing:* The data analysis is performed by using MATLAB (MathWorks Inc., USA). The MT data are low-pass filtered at 6 Hz to remove noise. The amplitudes of the MT data are normalized to body height to reduce the influence of anthropometric differences between the subjects. Similar to the processing of MT, the GRF data are low-pass filtered at 20 Hz, and their amplitudes are normalized to body weight. The EMG data are band-pass filtered from 10 Hz to 100 Hz, notch filtered at 50 Hz, and full-wave rectified. Then, the data of each channel are normalized by the maximum EMG amplitude of the corresponding muscle. In this study, the maximum EMG amplitude for each muscle is calculated as the average value of the EMG peak amplitudes across multiple gait cycles from one subject.

Due to the differences between the recorded gait cycles, the equal sample length requirement in pattern classification problem is not satisfied, which will result in the failure of training for a classifier. To this end, each modal data are separately time-normalized to 101 points (representing the gait cycle from 0% to 100%) channel by channel, using the cubic spline interpolation and resampling methods. Next, for each modality, the time-normalized data from different channels are concatenated to form the respective sample vector. For MT, each sample vector has 8484 dimensions (28 markers $\times$ 3 spatial directions $\times$ 101 points in time). For GRF, each sample vector has 606 dimensions (2 force plates $\times$ 3 spatial directions $\times$ 101 points in time). For EMG, each sample vector has 808 dimensions (8 pairs of electrodes $\times$ 101 points in time). As the three modal data will be severally fed to the respective classifiers, the equal sample length requirement only needs to be satisfied among the samples of each single modality rather than the samples between different modalities.

*2) Feature Extraction:* The high dimensional data probably deteriorate the generalization performance of the pattern classification algorithms when the number of experimental samples is not very large [28]. To this end, the principal component analysis (PCA) [29] is used for each modal data to reduce redundant information and extract effective gait features. The number of retained principal components is determined by using a criterion of 80% of total data variance.

*3) Individual Classifiers:* The features extracted from each modal data are served as the input of the corresponding individual classifier. In the present study, five popular classification algorithms [30] are considered as the candidate individual classifiers for each modality, which are the support vector machine (SVM), neural network (NN), random forest (RF), naive Bayes (NB) and k-nearest neighbor (KNN). Among them, the SVM adopts a radial basis kernel function, and its output scores are transformed to posterior probabilities by using Platt's method [31]. The NN is a feedforward network with a hidden layer, and the activation functions are sigmoid functions.

According to the training procedure presented in Section II-C, the candidate individual classifiers are established by using the individual-training dataset to optimize their adjustable parameters. Then, these candidate models are utilized to process the development-training dataset, and the best individual model for each modality is selected by comparing the recognition performance. The selection results are RF for MT, RF for GRF, and NN for EMG.

*4) Fusion Algorithms:* At the fusion stage, the outputs of the selected classifiers on the development-training dataset are used to train the decision fusion model. In general, the decision fusion model can be established by two ways: classification-based fusion algorithms and rule-based fusion algorithms [32]. The former integrates different modal information by using data-driven-based classification techniques, while the latter is based on pre-defined calculation rules. In the present paper, we adopt five classification-based fusion algorithms (e.g., SVM, NN, etc.) that have been outlined in Section II-D-3. At the same time, we consider two rule-based fusion methods, which are the average rule (AR) and max rule (MR) fusion algorithms [33]. The above seven fusion algorithms here are employed to build the fusion model rather than the previous individual model. The main difference lies that the inputs of the fusion model are the predicted probabilities which contain semantic information of different modalities. We hypothesize that by using an appropriate fusion algorithm, different modal semantic information can be effectively integrated, which has the potential to fully mine the complementarity between multimodal gait data. It is worth mentioning that except to the NN-based decision fusion algorithm whose performance has been reported in the relevant literature, other decision fusion algorithms adopted in this work have not yet been investigated in the field of automatic gait analysis.

*5) Walking Ability Score:* Based on the above exposition, the proposed architecture can be used to qualitatively recognize different gait patterns. However, how to simultaneously implement quantitative estimation of the gait differences between pathological and normal groups is still an open problem. In this study, a novel custom-defined gait score is proposed to embed the quantitative analysis method into this architecture.

As a value containing semantic information of gait data, the predicted probability of an individual classifier or fusion model can quantitatively reflect the possibility that a subject's gait pattern belongs to a certain class. Accordingly, we define this probability as a new gait score named as walking ability score (WAS) to quantify the differences between pathological and normal gait patterns. The WAS represents a subject's gait pattern as a real number with a range from 0 to 1. When the WAS is closer to 0, it indicates that the gait abnormality is more severe. Therefore, the WAS can be used to quantitatively assess the subjects' walking ability. Furthermore, the WAS of a fusion model is a further integration of the local WASs from different modal individual classifiers. Thus, the WAS after fusion represents more comprehensive semantic information, which may effectively exploit the complementarity between different modalities. Moreover, since the WAS is obtained from the objective dynamic gait data, the proposed architecture can be applied to objectively support clinical decisions. According to the above analysis, when compared with the traditional clinical scales, the WAS can provide a more objective, quantitative and comprehensive assessment result.

Although the WAS can be used to quantitatively assess the walking ability of both post-stroke patients and healthy subjects, we focus on the assessment for the patients with the consideration of clinical practice. By considering the influence of measurement errors, we use the walking ability mean score (WAMS) as the final estimation result, which is the mean value of the WASs across multiple experimental samples from a patient.

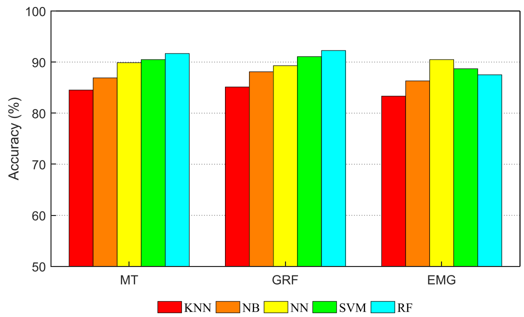


Fig. 5. Recognition accuracies of single-modal schemes.

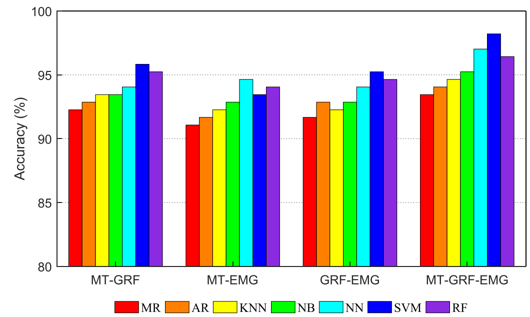


Fig. 6. Recognition accuracies of multimodal fusion schemes.

## III. Results

### A. Performance Metric

At first, we consider the evaluation of the qualitative recognition performance of the proposed architecture. In this work, we select the accuracy as the performance metric, which can be mathematically expressed as [30]

$$Accuracy = \frac{1}{N}\sum_{i=1}^{N} I(c_i = l_i) \tag{4}$$

where $N$ is the number of samples, $c_i$ and $l_i$ are the predicted class and ground truth label of the $i$th sample, respectively, and $I(\cdot)$ is an indicator function.

Since there are no standard performance metrics for the custom-defined gait scores, it is a great challenge for the evaluation of the quantitative estimation performance. In this study, we intend to explore an evaluation method from the perspective of clinical relevance. As the Wisconsin Gait Scale (WGS) [34] is widely used for the clinical assessment of the post-stroke hemiparetic gait, the clinical relevance of the proposed WAMS is evaluated by comparing with the WGS score.

Based on the above exposition, we take the recognition performance metric as the criterion to determine the best individual classification scheme and the best fusion scheme on the testing dataset. After that, the relationship between the WGS scores and the WAMSs of these best schemes is further explored and discussed.

### B. Qualitative Recognition Performance

*1) Single-Modal Schemes:* As a first step, we report the recognition performance of the five candidate classifiers on the testing dataset, respectively. Figure 5 shows the recognition accuracies of different individual classifiers by processing each single modal data. By comparing the recognition accuracies, we can find that not all the candidate individual classifiers show the same recognition performance for each modal data. For instance, the RF and SVM classification models achieve accuracies over 90% when applied to MT or GRF data, while only the NN classification model obtains an accuracy over 90% when applied to EMG data. The above inconsistent results can probably be interpreted by the “no free lunch” theorem [35], which indicates that no algorithms can work very well over all data due to the cost functions and sample characteristics. For the best recognition performance of each modality on the testing dataset, the results are consistent with that on the development-training dataset, i.e., RF for MT, RF for GRF, and NN for EMG, which means that the selection of individual classifiers is reasonable. Furthermore, the RF model using GRF data yields the highest accuracy of 92.26% among these schemes, which is chosen as the best single-modal scheme (BSS).

*2) Multimodal Fusion Schemes:* In the following, we explore the fusion forms of multimodal gait data, and test the recognition performance of different decision fusion algorithms. The fusion forms can be the combinations of any two modalities or three modalities. For full validation, all kinds of combinations are considered, i.e., MT-GRF, MT-EMG, GRF-EMG and MT-GRF-EMG. For each combination, we evaluate seven decision fusion algorithms mentioned in Section II-D-4 on the testing dataset, respectively. Figure 6 shows the experimental results. In most cases, the accuracies of multimodal fusion schemes are higher than those of the single-modal schemes. Furthermore, similar to the findings in the single-modal schemes, not all the fusion algorithms demonstrate the same performance for each combination form. Among these algorithms, the SVM, RF and NN fusion models show relatively better identification capability on all these modal combinations. Specifically, for the combinations of any two modal data, the highest accuracy of 95.83% is achieved by the SVM fusion model integrating MT and GRF data, which is selected as the the best two-modal fusion scheme (BTWFS). For the combination of three modalities, the highest accuracy of 98.21% is obtained by the SVM fusion model combining the three modal data, which is determined as the best three-modal fusion scheme (BTHFS).

In order to directly compare the best recognition performance in the cases of using different numbers of modalities, the recognition results of the best single-modal, two-modal and three-modal schemes are duplicated in Fig. 7. We can clearly see that the BTWFS outperforms the BSS by 3.57%, while the BTHFS outperforms the BSS and BTWFS by 5.95% and 2.38%, respectively. These results indicate that the recognition performance can be obviously improved by appropriately integrating more modal gait data. It can support our hypothesis that the complementarity between different modal data can be better mined by using multimodal fusion techniques. Besides, the highest accuracy among all the

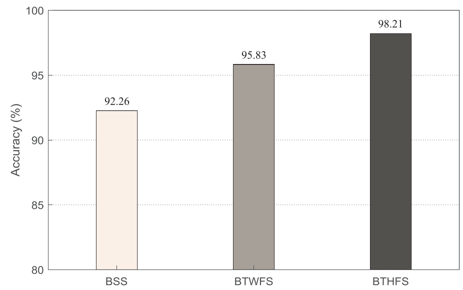


Fig. 7. Comparison on the recognition performance of the three best schemes.

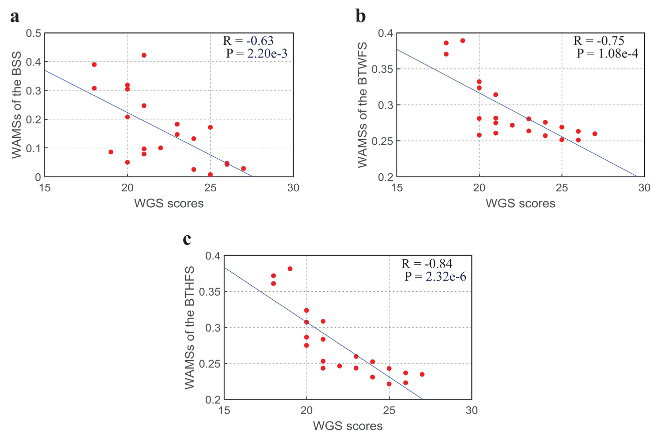


Fig. 8. Comparison on the estimation performance of the three best schemes: (a) BSS, (b) BTWFS and (c) BTHFS.

different schemes is more than 98% on the testing dataset. Since the size of the testing dataset is larger than that of the individual-training or development-training dataset, it is a relatively difficult classification problem. Therefore, this highest accuracy indicates a certain potential application value of the proposed fusion architecture in automatic clinical diagnosis of the hemiparetic patients after stroke.

### C. Quantitative Estimation Performance

As it is mentioned in Section III-A, the WGS score is used to compare with the proposed WAMS. All the post-stroke patients' walking performances have been assessed by an experienced physician using the WGS. The WGS consists of 14 items, whose score ranges from 14 to 45. In contrast to the WAMS, the lager a WGS score is, the more abnormal the gait performs.

It is worth noting that the traditional scale scoring depends on the clinician's personal experience. Different clinicians may give different scale scores on the same patient. Therefore, the comparison results are affected by uncertain subjective factors. The purpose of this comparison is just to explore a evaluation method for the quantitative estimation performance of the proposed architecture. By considering that the three best schemes have been determined, we take these best schemes as representatives to study their estimation performance on the testing dataset.

The relationship between the WGS scores and WAMSs of the three best schemes is shown in Fig. 8. The Pearson correlation coefficient and the result of correlation test at a significance level of 0.05 are provided on the top left corner of each subgraph, respectively. We can clearly see that there is a significant negative correlation between the WGS scores and WAMSs of each best scheme ($P < 0.05$). Furthermore, the absolute value of the correlation coefficient increases as the number of modalities increases. The absolute value of the correlation coefficient between the WGS scores and the WAMSs of the BTHFS is as high as 0.84, which indicates a strong linear relationship. Consequently, the proposed WAMS has potential for clinical use in quantitatively assessing the gait abnormality of the post-stroke patients.

## IV. Discussions

Most of the existing automatic gait analysis systems cannot achieve simultaneous qualitative and quantitative analysis of patients' gait abnormality. Furthermore, these systems usually use only a single data source, which is difficult to obtain adequate comprehensive and accurate analysis results in clinical practice. The present study gives a preliminary exploration of combining multimodal gait data to qualitatively and quantitatively analyze the gait abnormality in patients with stroke. Under the proposed fusion architecture, the MT, GRF and EMG data can be effectively integrated to implement the qualitative identification of the hemiparetic gait and the quantitative assessment of the patients' lower limb motor function.

Perhaps the most exciting findings are that with our measured data, the recognition accuracy of different gait patterns is enhanced as the number of used modalities increases in most cases. Meanwhile, for the best schemes of different modalities, the correlation between the probability estimations of the patients' walking ability and the clinical scale ratings becomes more obvious when using more modal data, which enable the estimations to be more comprehensible to the clinicians. In addition, other findings of this study are discussed as follows.

For the inputs of the automatic gait analysis systems, most researches adopt the time-invariant static gait data, such as kinematic or kinetic variables at some specific time points of the gait cycle, statistical characteristics derived from multiple gait cycles, etc. In contrast, three different time-variant dynamic gait data are served as the inputs of the system in this study. The static gait data do have some advantages, e.g., clear physical meaning. However, their defects are also very obvious. To begin with, the selection of different static gait data is normally dependent on the prior knowledge of the clinicians or researchers. Then, the calculation of some static gait data is time consuming and may cause error accumulation. Finally, and most importantly, only using the static gait data will lose a lot of time-dependent information during walking. The use of dynamic gait data can be a good way to avoid the above problems.

The structures of the existing automatic gait analysis systems are usually designed as a single classifier or a single calculation expression, whereas we design a complex architecture including the individual classifiers and fusion model.

These conventional structures can only process a single modal gait data source. As each single modal data can only reflect one aspect of human gait, the results processed by these structures may be one-sided and inaccurate. In contrast, the proposed fusion architecture can implement the integration of different modal gait data. The complementarity between different modal gait data can be effectively utilized in this architecture to obtain more comprehensive and accurate results.

The WAMS can quantitatively express a subject's walking ability as a real number between 0 and 1. Compared with the other gait scores reviewed in Section I, the main advantage lies in that the WAMS of the fusion model is a synthetic score which contains the semantic information of different modal gait data, while other gait scores are obtained based on the gait characteristics from only a single modal data source. Furthermore, the comparison with the WGS scores has validated the clinical relevance of the WAMS to some extent.

It is worth noting that there are some limitations in this work. Firstly, due to the time normalization of the multi-modal gait data, the differences between the subjects' walking velocities are greatly weakened. This gait characteristic will be given attention to and used in the future work. Secondly, due to the lack of standard evaluation metrics for the custom-defined gait scores, the quantitative estimation performance of the proposed architecture needs to be further validated on more clinical experiments. Thirdly, the inclusion criteria in this study limit that the fusion architecture is specifically designed for analyzing the abnormal gait after stroke. Therefore, this study needs to be further extended to the analysis of gait abnormality caused by other diseases. Finally, the designed experiments do not involve the assessment of the effect of a rehabilitation process or therapeutic treatment. This task will be explored in depth in our future research.

## V. Conclusion

Accurate recognition and assessment of gait abnormality play an important role in making effective rehabilitation treatment strategies for the post-stroke hemiparetic patients. This research proposes a multimodal fusion-based automatic gait analysis architecture which can qualitatively identify the patients' altered gait patterns after stroke, and quantitatively estimate their walking ability. The kinematic, kinetic and electrophysiological data recorded during walking can be processed and integrated in this architecture. The results show that both the recognition performance and estimation performance become better as more different modal gait data are used. The best three-modal fusion scheme produces over 98% recognition accuracy, and its estimation results indicate a strong significant negative correlation with the WGS scores. These results may facilitate the automatic gait analysis system to better support the traditional clinical decisions, and thereby enhance the efficiency of rehabilitation treatments.

## Acknowledgment

The authors would like to thank all the subjects who participated in the walking trials.

## References


[1] G. Chen, C. Patten, D. H. Kothari, and F. E. Zajac, "Gait differences between individuals with post-stroke hemiparesis and non-disabled controls at matched speeds," *Gait Posture*, vol. 22, no. 1, pp. 51–56, 2005.

[2] J. Y. Goulermas, A. H. Findlow, C. J. Nester, D. Howard, and P. Bowker, "Automated design of robust discriminant analysis classifier for foot pressure lesions using kinematic data," *IEEE Trans. Biomed. Eng.*, vol. 52, no. 9, pp. 1549–1562, Sep. 2005.

[3] H. Dejnabadi, B. M. Jolles, and K. Aminian, "A new approach for quantitative analysis of inter-joint coordination during gait," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 2, pp. 755–764, Feb. 2008.

[4] B. Mariani, S. Rochat, C. J. Büla, and K. Aminian, "Heel and toe clearance estimation for gait analysis using wireless inertial sensors," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 11, pp. 3162–3168, Nov. 2012.

[5] M. Karg, W. Seiberl, F. Kreuzpointner, J.-P. Haas, and D. Kulić, "Clinical gait analysis: Comparing explicit state duration HMMs using a reference-based index," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 23, no. 2, pp. 319–331, Mar. 2015.

[6] M. U. B. Altaf, T. Butko, and B. H. Juang, "Acoustic gaits: Gait analysis with footstep sounds," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 8, pp. 2001–2011, Aug. 2015.

[7] Y. Qi, C.-B. Soh, E. Gunawan, K.-S. Low, and R. Thomas, "Assessment of foot trajectory for human gait phase detection using wireless ultrasonic sensor network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 88–97, Jan. 2016.

[8] P. Parijat, T. E. Lockhart, and J. Liu, "EMG and kinematic responses to unexpected slips after slip training in virtual reality," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 2, pp. 593–599, Feb. 2015.

[9] M. W. Whitmore, L. J. Hargrove, and E. J. Perreault, "Gait characteristics when walking on different slippery walkways," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 1, pp. 228–239, Jan. 2016.

[10] P. Terrier, J. Carré, M.-L. Connaissa, B. Léger, and F. Luthi, "Monitoring of gait quality in patients with chronic pain of lower limbs," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1843–1852, Oct. 2017.

[11] W.-L. Wu and F.-C. Su, "Potential of the back propagation neural network in the assessment of gait patterns in ankle arthrodesis," *Clin. Biomech.*, vol. 15, no. 2, pp. 143–145, 2000.

[12] J. Kamruzzaman and R. K. Begg, "Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 12, pp. 2479–2490, Dec. 2006.

[13] N. Mezghani *et al.*, "Automatic classification of asymptomatic and osteoarthritis knee gait patterns using kinematic data features and the nearest neighbor classifier," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 3, pp. 1230–1232, Mar. 2008.

[14] B. Zhang, Y. Zhang, and R. K. Begg, "Gait classification in children with cerebral palsy by Bayesian approach," *Pattern Recognit.*, vol. 42, no. 4, pp. 581–586, 2009.

[15] C. Scheffer and T. Cloete, "Inertial motion capture in conjunction with an artificial neural network can differentiate the gait patterns of hemiparetic stroke patients compared with able-bodied counterparts," *Comput. Methods Biomech. Med. Eng.*, vol. 15, no. 3, pp. 285–294, 2012.

[16] D. Laroche *et al.*, "A classification study of kinematic gait trajectories in hip osteoarthritis," *Comput. Biol. Med.*, vol. 55, pp. 42–48, Dec. 2014.

[17] F. Wahid, R. K. Begg, C. J. Hass, S. Halgamuge, and D. C. Ackland, "Classification of Parkinson's disease gait using spatial-temporal gait features," *IEEE J. Biomed. Health Informat.*, vol. 19, no. 6, pp. 1794–1802, Nov. 2015.

[18] W. Zeng, F. Liu, Q. Wang, Y. Wang, L. Ma, and Y. Zhang, "Parkinson's disease classification using gait analysis via deterministic learning," *Neurosci. Lett.*, vol. 633, pp. 268–278, Oct. 2016.

[19] E. Dolatabadi, B. Taati, and A. Mihailidis, "An automated classification of pathological gait using unobtrusive sensing technology," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 12, pp. 2336–2346, Dec. 2017.

[20] L. M. Schutte, U. Narayanan, J. L. Stout, P. Selber, J. R. Gage, and M. H. Schwartz, "An index for quantifying deviations from normal gait," *Gait Posture*, vol. 11, no. 1, pp. 25–31, 2000.

[21] M. H. Schwartz and A. Rozumalski, "The gait deviation index: A new comprehensive index of gait pathology," *Gait Posture*, vol. 28, no. 3, pp. 351–357, 2008.

[22] R. Baker *et al.*, "The gait profile score and movement analysis profile," *Gait Posture*, vol. 30, no. 3, pp. 265–269, 2009.

[23] A. Gouelle, F. Mégrot, A. Presedo, I. Husson, A. Yelnik, and G. F. Penneçot, "The gait variability index: A new way to quantify fluctuation magnitude of spatiotemporal parameters during gait," *Gait Posture*, vol. 38, no. 3, pp. 461–465, 2013.

[24] J. Christian, J. Kröll, G. Strutzenberger, N. Alexander, M. Ofner, and H. Schwameder, "Computer aided analysis of gait patterns in patients with acute anterior cruciate ligament injury," *Clin. Biomech.*, vol. 33, pp. 55–60, Mar. 2016.

[25] C. Cui, G.-B. Bian, Z.-G. Hou, J. Zhao, and H. Zhou, "A multimodal framework based on integration of cortical and muscular activities for decoding human intentions about lower limb motions," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 4, pp. 889–899, Aug. 2017.

[26] A. Cappozzo, F. Catani, U. D. Croce, and A. Leardini, "Position and orientation in space of bones during movement: Anatomical frame definition and determination," *Clin. Biomech.*, vol. 10, no. 4, pp. 171–178, 1995.

[27] A. O. Perotto, *Anatomical Guide for the Electromyographer: The Limbs and Trunk*, 4th ed. Springfield, IL, USA: Charles C Thomas, 2005.

[28] S. J. Raudys and A. K. Jain, "Small sample size effects in statistical pattern recognition: Recommendations for practitioners," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 13, no. 3, pp. 252–264, Mar. 1991.

[29] K. Fukunaga, *Introduction to Statistical Pattern Recognition*, 2nd ed. San Diego, CA, USA: Academic, 1990.

[30] T. Hastie, R. Tibshirani, and J. Friedman, *The Elements of Statistical Learning Data Mining, Inference, and Prediction*, 2nd ed. New York, NY, USA: Springer, 2009.

[31] J. C. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," in *Advances in Large Margin Classifiers*. Cambridge, MA, USA: MIT Press, 2000, pp. 61–74.

[32] P. K. Atrey, M. A. Hossain, A. El Saddik, and M. S. Kankanhalli, "Multimodal fusion for multimedia analysis: A survey," *Multimedia Syst.*, vol. 16, no. 6, pp. 345–379, 2010.

[33] J. Kittler, M. Hatef, R. P. W. Duin, and J. Matas, "On combining classifiers," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 20, no. 3, pp. 226–239, Mar. 1998.

[34] A. A. Rodriquez *et al.*, "Gait training efficacy using a home-based practice model in chronic hemiplegia," *Arch. Phys. Med. Rehabil.*, vol. 77, no. 8, pp. 801–805, 1996.

[35] D. H. Wolpert and W. G. Macready, "No free lunch theorems for search," Santa Fe Inst., Sante Fe, NM, USA, Tech. Rep. SFI-TR-05-010, 1995.