Chinese Sign Language Recognition Based on An Optimized Tree-structure Framework

Xidong Yang, Xiang Chen*, *Member, IEEE*, Xiang Cao, Shengjing Wei, Xu Zhang, *Member, IEEE*

***Abstract*—Chinese sign language (CSL) subword recognition based on surface electromyography (sEMG), accelerometer (ACC) and gyroscope (GYRO) sensors was explored in this paper. In order to fuse effectively the information of these three kinds of sensors, the classification abilities of sEMG, ACC, GYRO and their combinations in three common sign components (one or two handed, hand orientation and hand amplitude) were evaluated firstly and then an optimized tree-structure classification framework was proposed for CSL subword recognition. Eight subjects participated in this study and recognition experiments under different testing conditions were implemented on a target set consisting of 150 CSL subwords. The proposed optimized tree-structure classification framework based on sEMG, ACC and GYRO obtained the best performance among seven different testing conditions with single sensor, paired-sensor fusion and three-sensor fusion, and the overall recognition accuracies of 94.31% and 87.02% were obtained for 150 CSL subwords in user-specific test and user-independent test respectively. Our study could lay a basis for the implementation of large-vocabulary sign language recognition system based on sEMG, ACC and GYRO sensors.**

***Index Terms*—Accelerometer, electromyography, gyroscope, sensor fusion, sign language recognition**

I. Introduction

Sign language (SL) is a kind of structured gesture and has been commonly used among the deaf and hearing-impaired people. However, there appear to be communication barriers exist between the deaf and the hearers especially those who are incompetent in SL [1], [2], [3]. Sign language recognition (SLR), which can translate SL into text or speech automatically and serve as one of the most suitable way of human-computer interaction (HCI), attracts a large number of researchers' attention [1], [4], [5]. Meanwhile, it is urgent to build assistive systems like SL interpreters in order to eliminate the communication barriers between the deaf and the hearing society as well as to provide better user experience in HCI.

In general, SL gestures contain abundant information, such as hand shape, orientation, location, facial expression and the movement of body, arms or hands [2], [3], [6]. There are two main categories of SLR approaches in term of the sensing technologies. The first one is computer vision-based approach. This technique utilizes image devices such as camera to capture SL information and image processing technique to finish SLR [1], [3], [4]. For instance, Kishore and Kumar [7] developed a video-based Indian SLR system and achieved 96% average classification rate over 80 signs. The main drawback of vision-based technology is that its performance is vulnerable to the influence of environment such as background and illumination [1], [4], [8]. The other SLR technique is based on data glove which uses several subtle sensors such as strain gauges and hand trackers to detect the movement pattern of hand and fingers. This method can obtain high recognition accuracy in large-scale SLR [4], [8], [9]. Gao et al. [10] firstly reported a continuous SLR work over 5113 signs with two sensory gloves and three trackers, and achieved 90.8% average classification rate. Li et al. [11] collected real-world data using a pair of low-cost digital gloves and achieved 87.4% word accuracy in the evaluation of 1024 testing sentences involving 510 Chinese sign language (CSL) words. Nevertheless, wearing the cumbersome data glove to collect hand and finger movements will potentially go against the convenient, efficient and natural intention of HCI [1], [4], [9].

Compared with the traditional SLR approaches mentioned above, alternative SL sensing techniques based on low-cost, wearable and high-portable devices such as surface electromyography (sEMG) and inertial sensors have been recently introduced [12]. SEMG can reflect behavior and movement patterns by measuring electrical activities from the corresponding muscles. It is suitable for the recognition of fine motions such as wrist and finger movements with wide applications in the field of active control [13], [14] and medical rehabilitation [15], [16], [17]. For example, Yun et al. [18] collected 5-channel sEMG signals from the forearm to classify 26 single characters of American Sign Language and acquired an approximate accuracy of 94%. Inertial sensors such as accelerometer (ACC) and gyroscope (GYRO) also have great advantages in capturing information of hand gesture. Accelerometer is usually applied to measure accelerations related to vibrations and the gravity. ACC-based gesture recognition systems are extremely suitable for the recognition of large-scale space trajectory [13], [8]. Wang et al. [19] developed a digital pen input device based on a 3-axis accelerometer for the recognition of 10 handwritten digits and 8

Manuscript received September 15, 2015. This work was supported by the National Nature Science Foundation of China under Grants 61271138 and 61431017.

X. Chen* is with the Institute of Biomedical Engineering, University of Science and Technology of China (USTC), Hefei 230027, China (corresponding author: +86-0551-6360-1175; e-mail: xch@ustc.edu.cn).

X. Yang, X. Cao, S. Wei and X. Zhang are with the Institute of Biomedical Engineering, University of Science and Technology of China (USTC), Hefei 230027, China (e-mail: yangxid@mail.ustc.edu.cn; flander@mail.ustc.edu.cn; wt901218@mail.ustc.edu.cn; xuzhang90@ustc.edu.cn).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

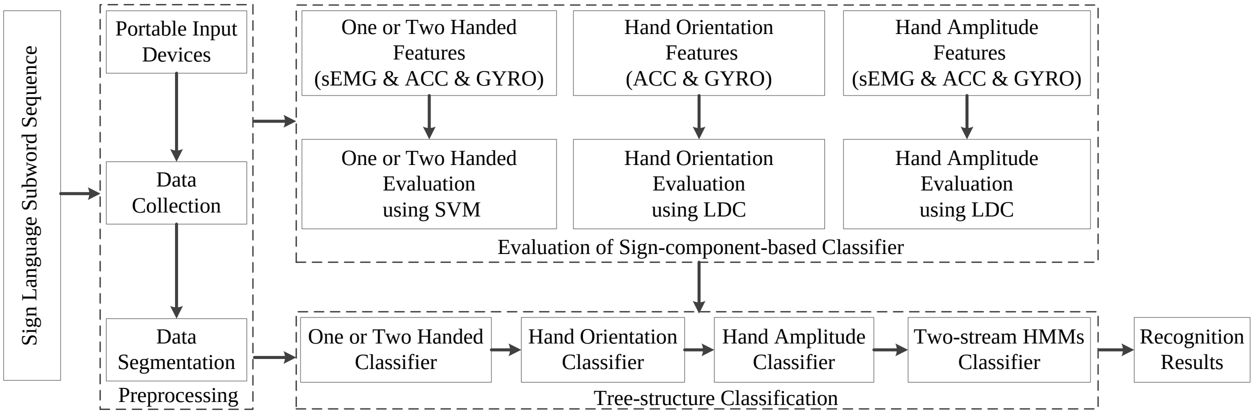

Fig. 1. The block diagram of the proposed SLR approach.

gesture trajectories with the overall average accuracy of 98% and 98.75%, respectively. Gyroscope is helpful in judging gesture space position and can accurately collect the arm and hand rotation information. Preferable recognition can be realized based on gyroscope especially for small-scale rotation [20], [21]. Dermitzakis et al. [15] explored gesture recognition in the context of upper-limb prosthetics using gyroscope sensors and an optimal classification rate of 97.53% over 22 gestures in total was obtained.

Due to their small-size, light-weight, low-cost and wearable properties, both sEMG and inertial sensors can be easily designed as an armband [21], a watch-like wristband [8], [9], or be integrated into electronic devices like mobile phones [12], [13], which can facilitate the implementation of practical SLR systems compared with conventional vision-based or data glove-based approaches. Many previous researches indicated that the fusion of sEMG and inertial sensors could enhance the performance of hand gesture recognition significantly and enlarge the size of recognizable gesture set. Chen et al. [22] reported that the fusion of sEMG and ACC achieved 5%-10% improvement in the recognition accuracies for 24 kinds of hand gestures. Boschmann et al. [23] reported that the combination of sEMG, ACC and GYRO significantly outperformed sEMG or the fusion of sEMG and ACC. Besides, Li et al. [9] developed a portable SLR system based on sEMG and ACC sensors and achieved an impressive accuracy of 95.78% over 121 frequently used CSL subwords in user-specific condition. Zhang et al. [8] built a real-time interactive system based on sEMG and ACC sensors, realized the control of a virtual Rubik's cube using 18 gestures, and achieved 97.6% recognition accuracies in user-specific condition and 90.2% in user-independent condition, respectively.

Although relatively good results have been obtained in SLR and HCI implementation based on solo sEMG, ACC, GYRO sensors and their different combinations, the number of the recognizable gestures is limited and still can't meet the needs of daily communication between the deaf and hearing society. So far, according to the literature, the maximum recognizable gesture vocabulary based on the wearable and portable sEMG and inertial sensors is 121 subwords in user-specific condition [9] and 40 words in user-independent condition [24], respectively. The relatively high identification accuracy is commonly achieved only in user-specific condition in which an individual user has to attend data acquisition to train his own SLR models prior to actual SLR implementation. Consequently, taking the practicability into account, large vocabulary sign recognition with high accuracy is extremely important for the realization of practical SLR systems.

In order to enlarge the size of recognizable gesture set and improve the SLR performance, this study presents an optimized tree-structure classification framework for recognizing 150 frequently-used CSL subwords. The gesture information was recorded via a home-made portable and wearable wristband integrating sEMG, ACC and GYRO sensors. The classification performance was evaluated by using sEMG, ACC, GYRO and their combinations to demonstrate the feasibility of the proposed method.

II. Methodology

Fig. 1 shows the block diagram of the proposed SLR approach combining multichannel sEMG, 3-D ACC and 3-D GYRO data. The preprocessing module includes data collection procedure and data segmentation. The tree-structure classification and evaluation of sign-component-based classifier are the major work of this study. As mentioned above, sEMG can effectively reflect the muscle activity intensity, and inertial sensors can capture the information of movement trajectory. Therefore, three common components, namely one or two handed, hand orientation and hand amplitude, were selected as the target components of the proposed tree-structure classification framework to improve the recognition performance and to cut down the recognition time consumption. Based on the performance evaluation of the three kinds of sensors and their combinations in the classification of these three components, an optimized tree-structure classification method was proposed and successfully applied in the final recognition of 150 frequently used CSL subwords. The details of all individual modules are described below.

A. *Target CSL Subwords and Subjects*

There are more than 5500 sign gestures in all according to the Chinese sign language dictionary and the target gesture set

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

TABLE I
THE TYPICAL FEATURES OF THE TARGET GESTURE SET

Typical components	Category	Example (Chinese character)	Number
One or Two Handed	One-handed	you(你)	81
	Two-handed	home(家)	69
Hand Orientation	Horizontal	road(路)	56
	Vertical	happy(高兴)	55
	Longitude	admit(承认)	39
Hand Amplitude	Low Amplitude	have(有)	34
	High Amplitude	movie(电影)	46
	Middle Amplitude	think(思考)	70

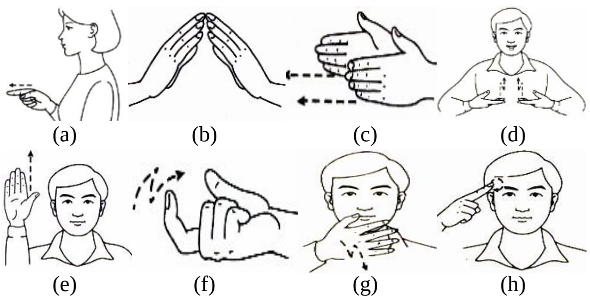

Fig. 2. Examples of the typical components. (a) you(你). (b) home(家). (c) road(路). (d) happy(高兴). (e) admit(承认). (f) have(有). (g) movie(电影). (h) think(思考).

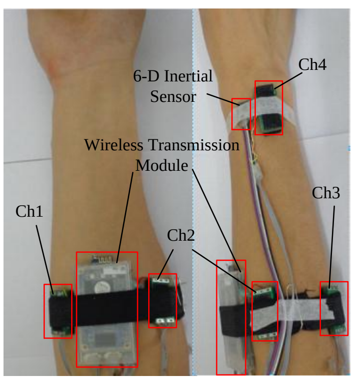

Fig. 3. Placement of 4-channle sEMG and 6-D inertial sensors on right forearm. Here 'Ch' denotes the channel of sEMG sensor.

was comprised of 150 frequently-used nouns and verbs about food, clothing, housing, transportation etc., which was determined according to the basic necessities of life. The typical features of the target gesture set including one or two handed, hand orientation and hand amplitude are summarized in Table I. Concretely, the 150 subwords includes 81 one-handed subwords and 69 two-handed subwords, three typical hand orientations and three different hand amplitude levels. The examples of the typical components are illustrated in Fig. 2.

Eight right-handed male volunteers (age: 25 ± 0.7 yr) participated in this study. All subjects were healthy without any neuromuscular or joint diseases. The experiments were carried out under the condition that all subjects were informed of the objective and procedures of this study. Before data collection, each subject was required to learn all the subwords from a teaching video transcribed by a graduate student who was considerably familiar with these subwords.

B. *Data Collection*

In data measurement, two flexible and wireless acquisition modules were worn on both of the right and left forearms symmetrically. For each module, a 6-D inertial sensor (including a 3-D ACC and a 3-D GYRO, MPU6050, InvenSense, Inc., Sunnyvale, CA, USA) and four sEMG sensors were placed in a band formation around the forearm (see Fig. 3). The 6-D inertial sensor was placed on the back of the forearm near the wrist. Each sEMG sensor has two parallel Ag-AgCl bars with a 10 mm×1 mm physical dimension and a 10-mm center-to-center spacing respectively. A two-stage amplifier with a total gain of 54 dB and a band-pass filter of 20-500 Hz bandwidth were used as the preprocessing steps for the differential sEMG signals in each channel. Then the collected analog sEMG signals were digitalized by a 12-bit AD converter (ADS8638, Texas Instruments, Inc., Dallas, TX). The four sEMG sensors were placed over four proper positions to target the forearm muscles: *extensor digiti minimi*, *palmaris longus*, *extensor carpi ulnaris*, and *extensor carpi radialis*, respectively, as demonstrated in Fig. 3. The locations of the target muscles were determined based on the understanding of the anatomy of human body. SEMG signals were recorded at 1000-Hz sampling rate, the inertial signals at 100-Hz sampling rate, and the data was recorded to a computer via Bluetooth for further processing.

The recorded signals were monitored via a computer screen which was placed in front of the subjects during the experiment. For each subject, the data collection included 5 experiment sessions scheduled in different days. In each session, the subject was asked to perform all the 150 subwords, each with 5 repetitions. During the performance, the subject's arm needed to get back to the relaxing state of drooped naturally when a subword was finished and then to the beginning position of the next subword. Consequently, 3750 CSL subword samples in total were collected from each subject for experimental analysis.

C. *Data Segmentation*

Data segmentation was conducted to automatically detect the active segment corresponding to each subword execution from the continuous signal. Because there is no apparent muscle contraction during the non-active segment between two consecutive subwords, the sEMG signal was effectively used for data segmentation as it directly represents muscle activities. Before the data segmentation, each sEMG channel was filtered by a band-pass filter of 20-500 Hz bandwidth in the range of dominent energy [25] and the inertial signal was filtered through a smooth filter. In fact, all one-handed CSL subwords selected in this study were performed by the right hand, so only the sEMG signals collected from right hands were considered as the references for data segmentation. Considering the synchronism of inertial signals with the sEMG signals, a 10-fold up-sampling was conducted for the inertial signals as the sampling rate of sEMG signals is ten times of inertial signals.

The active segment detection method applied in this study is based on sample entropy (SampEn) analysis of sEMG. The

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

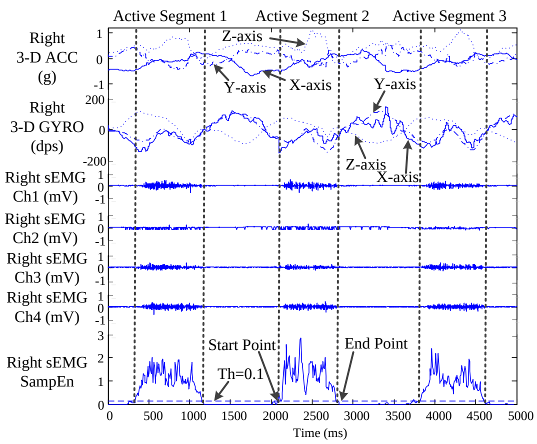

Fig. 4. The example of data segmentation results for CSL subword “full(满)”.

SampEn is derived from approximate entropy (ApEn) [26], [27] and has achieved successful application in physiological time series analysis [27]. The calculation procedure of SampEn is defined as (1):

$$SampEn(m,r,N) = -\ln\left[\frac{B^{m+1}(r)}{B^{m}(r)}\right] \tag{1}$$

where m is the dimension and N is the sequence length, r is the tolerance for accepting matrices and $B^{m}(r)$ is the probability that two sequence match for m points. It is reported that r can be chosen as $(0.15-0.25)\times SD$ empirically [26], [27], where SD refers to the standard deviation of the original time sequence. We set $m=2$ and $r=0.25\times SD$ in this study. Considering that users were asked to relax after each sign in the data collection experiment, the data segmentation based on SampEn analysis could be conducted based on the following five steps:

1) Compute averaged sEMG time series across multiple sEMG channels.
2) Apply a series of overlapping analysis windows to the averaged sEMG time series with a window length of 32 ms and an increment of 8 ms.
3) Apply the SampEn analysis on the analysis windows to represent the muscle activities complexity where the SampEn may increase with burst of muscle activity and decrease with the muscle loosen.
4) Compare the SampEn with a predefined threshold. The active segment of a subword begins when the SampEn of the signal exceeds the threshold and continues until the SampEns of consecutive N frames below the threshold. The active segments for both sEMG and inertial signals are determined by the same boundaries.
5) Abandon the segments whose lengths are less than N as noise and N is set as 15 frames (equivalent to 120ms), which is consistent with our pilot studies [8], [9], [28].

Fig. 4 shows the example of data segmentation results for CSL subword “full (满).” Considering the example sign is a one-handed subword, only the signals of right hand including 3-D ACC, 3-D GYRO and 4-channel sEMG signals are shown. The SampEn stream rising above the threshold denotes the beginning of the active segment, while SampEn falling below the threshold denotes its end. The three active segments corresponding to the three repetitions of the example subword are successfully marked with dashed line. In addition, it has been reported that the background noise is the dominant factor which should be considered in choosing the proper threshold [8], [9]. Fortunately, the data collection experiments of this study were all conducted in a favorable environment, and all active segments can be successfully and automatically detected when the threshold is chosen as 0.1 according to Fig. 4.

D. *Evaluation of Sign-component-based Classifiers*

In general, each subword can be simply divided into a few basic sign components, such as hand shape, trajectory, location, facial expression and body posture and so on [2], [3], [4], [28]. Because different subwords may have same sign components and numerous gestures can be composed of several basic components, a large-scale SLR should be achieved based on the small-scale recognition of sign components. In our pilot study [28], a sign-component-based framework was successfully proposed in the recognition of 121 CSL subwords based on the sEMG and ACC sensors. This study will firstly evaluate the classification abilities of sEMG, ACC, GYRO and their combinations in three common sign-component-based classifiers including one or two handed classifier, hand orientation classifier and hand amplitude classifier, and then the final tree-structure classification method is proposed based on the best identification result of each classifier in order to enlarge the gesture vocabulary and improve the recognition performance.

1) *One or Two Handed Classifier*

The CSL subwords can be categorized as either one-handed signs or two-handed signs. In this study, only the right hand is applied in performing the one-handed signs while both hands are involved for two-handed signs. For one-handed subword, there is no apparent muscle contraction or hand movement of left hand, thus the mean absolute value (MAV) of the sEMG signal and the standard deviation (Std) of the ACC and GYRO signal measured from the left forearm are comparatively low. MAV and Std are defined as (2) and (3) respectively, where x represents the signal, u and N are the mean value and length of the signal respectively. A classifier of support vector machine (SVM), which is originally proposed to solve the dichotomy problem, is employed as the one or two handed classifier in this study. The mechanism of SVM is to find an optimal separating hyperplane that could satisfy the classification requirement as well as to exhibit good performance by maximizing the blank space on either side of the hyperplane [29].

$$MAV = \frac{1}{N}\sum_{i=1}^{N}|x_i| \tag{2}$$

$$Std = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x(i)-u\right)^{2}} \tag{3}$$

The MAVs of 4-channel sEMG, Stds of both ACC and GYRO measured from the left hand were employed to form a feature vector in one or two handed classifier based on SVM. In order to evaluate the contribution of sEMG, ACC and GYRO

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

sensors in one or two handed classifier, seven different testing conditions with single sensor, paired-sensor fusion and three-sensor fusion on the recognition of 150 CSL subwords across eight subjects were conducted in user-specific manner and user-independent manner, respectively. In the training phase, the feature vector extracted from the included sensors was sent into the SVM to train a hyperplane. As for the testing phase, all the unknown samples were labeled as either one-handed sign or two-handed sign based on the distance to the corresponding hyperplane. Finally, the sensor combination which achieved the best performance in one or two handed classifier will be utilized in the final tree-structure classification.

2) Hand Orientation Classifier

There are various hand orientations according to different palm facings when performing SL. As shown in Table I, three typical hand orientations including vertical, horizontal and longitude are involved in the target gesture set. It is reported that the mean value (Mean) of 3-D ACC active segments can efficiently represent different hand orientations no matter whether the hand gesture is a static subword or a dynamic subword [8], [9]. The calculation formula of Mean is shown in (4) and the parameters have the same meaning with (2) and (3). Besides, the GYRO can provide complementary angular information for ACC as previously stated, which can characterize the spatial position of gestures more accurately. Therefore, the Mean of ACC and GYRO measured from both hands are used as feature vector to classify hand orientations.

$$Mean = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{4}$$

The fuzzy K-means clustering and linear discriminant classifier (LDC) are employed in the hand orientation classifier. A fuzzy K-means algorithm [30] is proposed to partition the training samples into K clusters where each cluster denotes a kind of similar hand orientation. We set the number of clusters K = 3 for the hand orientation classifier to implement a rough hand orientation classification in this study. Subsequently, a LDC [30] is trained to determine the pattern branch of input data for hand orientation classification based on the distance to the corresponding clusters of typical hand orientation according to (5).

$$P\left(\varphi \middle| \phi_j\right) = G\left(\varphi, \mu_j, \Sigma_j\right) \tag{5}$$

where μ_j and Σ_j are the mean vector and covariance matrix of class ϕ_j, φ is the feature vector , and the multivariate Gaussian distribution is denotes as $G\left(\cdot, \mu_j, \Sigma_j\right)$.

In this approach, the left and right hand orientation classifier are trained independently. In the testing phase, by calculating the maximum likelihood of a sample φ belonging to the hand orientation class ϕ_j, each testing sample is directly classified into one of three orientation classes according to the minimum distance to the three cluster centers of the typical orientation classes. For one-handed subwords, the testing samples are directly fed into the right hand orientation classifier. As for the two-handed subwords, the testing samples are first sent into the left classifier, and then into the right orientation classifier. The contribution of ACC and GYRO sensors in hand orientation classifier is evaluated in three conditions: ACC-only, GYRO-only and fusion of ACC and GYRO, respectively. The feature combination which obtained the best identification result is applied in the final subword recognition method for hand orientation classification.

3) Hand Amplitude Classifier

Another typical classifier of sign component is based on hand amplitude during SL execution. Three hand amplitude levels are defined as presented in Table I and the typical examples are shown in Fig. 2. Actually, MAV of sEMG and Std of inertial signals can reflect hand intensity or hand movements instantaneously. Therefore, for each subword segment, the MAV of sEMG, Std of the 3-D ACC and 3-D GYRO measured from both hands formed the feature vector for hand amplitude classifier.

Similar to the hand orientation classifier, the left and right hand amplitude classifier are trained independently. The fuzzy K-means clustering and LDC are proposed for the training and classification of the hand amplitude classifier. The number of clusters is chosen as K = 3 to match the three kinds of typical hand amplitudes involved in this study. In the training phase, all the training samples are assigned to the cluster of subwords with approximate arm amplitude. Then, a LDC is trained for the hand amplitude classifier based on the clustering results. In the testing phase, the unknown input gesture is evaluated by calculating the likelihood via the LDC. For one-handed subwords, the testing samples are directly classified according to the right hand amplitude classifier. As for two-handed subwords, the testing samples are sent into the left and right hand amplitude classifier in sequence. Seven testing conditions including single-sensor, paired-sensor fusion and three-sensor fusion in user-specific and user-independent manner are involved in the evaluation of hand amplitude classifier. Additionally, the sensor combination of the best classification result in hand amplitude classifier will be employed in the proposed subword recognition approach.

E. Tree-structure Classification

The tree-structure classification is capable of simplifying the complex multi-category classification greatly by dividing it into multiple dichotomous issues and then solves it hierarchically. In general, a decision tree consists of a root node, a set of internal nodes and some leaf nodes, where each leaf node indicates a class [31]. The input of a decision tree is a group of features that represent different attributes. Each non-leaf node indicates a classifier that can split the samples into different branches according to different attributes. At last, each sample is assigned at a certain leaf node with a final decision. Tree-structure classification idea was constantly employed in SLR [8], [9] and could achieve robust performance particularly with large vocabulary SLR [8], [31]. In our pilot study [9], the tree-structure classification was successfully applied in the recognition of 121 CSL subwords

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

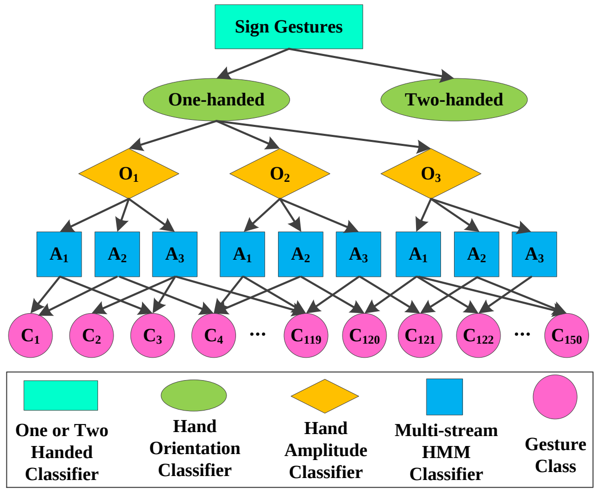

Fig. 5. The structure of decision tree for SLR.

based on the sEMG and ACC sensors and promising recognition accuracy was achieved. This study will follow the same research strategy as [9] and extend the tree-structure classification with the fusion of sEMG, ACC and GYRO sensors, and enlarge the gesture vocabulary further.

Based on the comprehensive analysis of the SL subwords characteristics, an extended decision tree was proposed in this study for SLR, with its structure shown in Fig. 5. The root node is a representation of the whole gesture candidates which cover all the possible subwords classes. The classification procedure of one-handed subwords and two-handed subwords is the same and only the classification of one-handed subwords is illustrated in Fig.5. In the recognition phase, each unknown test sample is fed into one or two handed classifier, hand orientation classifier, hand amplitude classifier and the final multi-stream HMM classifier in sequence to obtain the eventual identification result.

One or two handed classifier, hand orientation classifier, hand amplitude classifier in the decision tree could be considered as rough classifiers, as they only assign the testing samples into several rough classes based on some different attributes. As mentioned above, the best identification result of each rough classifier will be applied in the final tree-structure classification framework. The multi-stream HMM (MSHMM) classifier is the special part of the tree-structure classifier that it gives the final classification result of the testing samples. In other words, the preliminary three classifiers are utilized to reduce the searching range of MSHMM classifier which could be helpful to improve the recognition performance and to cut down the recognition time consumption. The feature vector of MSHMM classifier contains MAV and $[f_1, f_2, f_3, f_4, f_5]$ of sEMG signals, and Mean and Std of ACC and GYRO signals. The feature vector of $[f_1, f_2, f_3, f_4, f_5]$ can describe the frequency-domain contents of sEMG from the time-domain perspective and the concrete calculation procedure can be found in [32]. Additionally, in order to eliminate the individual difference in the execution speed of each subword as far as possible, the active segments are divided into 50 frames equally before the feature extraction.

It is reported that MSHMM is capable of integrating multi-sensor information to recognize a certain number of CSL subwords[8], [9]. Here the implementation of MSHMM is based on two standard single-stream HMMs, that is sEMG and inertial stream models, denoted as λ_c^E and λ_c^I separately. The stream fusion of λ_c^E and λ_c^I can be realized according to a linear weighted combination function of log-likelihood as (6) shows.

$$\log P(\mathrm{O}|\lambda_c) = \omega \log P(\mathrm{O}_E|\lambda_c^E) + (1-\omega)\log P(\mathrm{O}_I|\lambda_c^I), 1 \le c \le \mathrm{C} \quad (6)$$

where O_E and O_I are observed feature sequence from both sEMG and inertial streams, λ_c is an MSHMM corresponding to the c-th subword class, and C is the total number of subword classes of this node. Considering the trade-off between high performance and low complexity, the stream weight factor ω was set to a constant value of 0.5 in this study. Similar to [8], [9], the hidden states of the MSHMM and the mixture components of modeling each observation are empirically set to 5 and 3, respectively. In this study, one-handed subwords and two-handed subwords are modeled respectively. For one-handed subwords, only the features derived from sEMG and inertial sensors on the right hand are used to train MSHMM. As for the two-handed subwords, two separate MSHMMs are trained using the same training procedure with one-handed subwords to model the left and right hand subword with the feature set of the corresponding hand, respectively.

In the testing phase, the unknown samples are sent into the MSHMM classifier with the feature vector sequences of sEMG, ACC and GYRO to get the recognition results. For one-handed subwords, the recognition result is determined by (7).

$$c^* = \arg\max_{1\le c\le C} P\left(\mathrm{O}_R|\lambda_{Rc}\right) \quad (7)$$

where O_R is the observation sequence of right hand and λ_{Rc} is a MSHMM corresponding to the c-th subword class, c^* represents the final class which maximizes the likelihood $P(\mathrm{O}_R|\lambda_{Rc})$ for the right hand.

For two-handed subwords, the likelihood $P(\mathrm{O}_R|\lambda_{Rc})$ and $P(\mathrm{O}_L|\lambda_{Lc})$ are computed respectively to get two overall likelihood scores with the same approach. The final recognition result is assigned as the two-handed subword c^* with the highest overall combined likelihood score.

$$c^* = \arg\max_{1\le c\le C} \left(P\left(\mathrm{O}_R|\lambda_{Rc}\right) + P\left(\mathrm{O}_L|\lambda_{Lc}\right)\right) \quad (8)$$

where $P\left(\mathrm{O}_L|\lambda_{Lc}\right)$ represents the likelihood of an observed feature sequence O_L belonging to the left hand class model λ_{Lc}.

III. Results and Analysis

In order to evaluate the performance of the proposed method, pattern recognition of all 150 CSL subwords was conducted in both user-specific manner and user-independent manner, respectively. In the data collection experiments, 5 groups of data samples were obtained for each subject. Therefore, we implemented a fivefold cross validation scheme that four of the five groups of data samples were used for training while the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

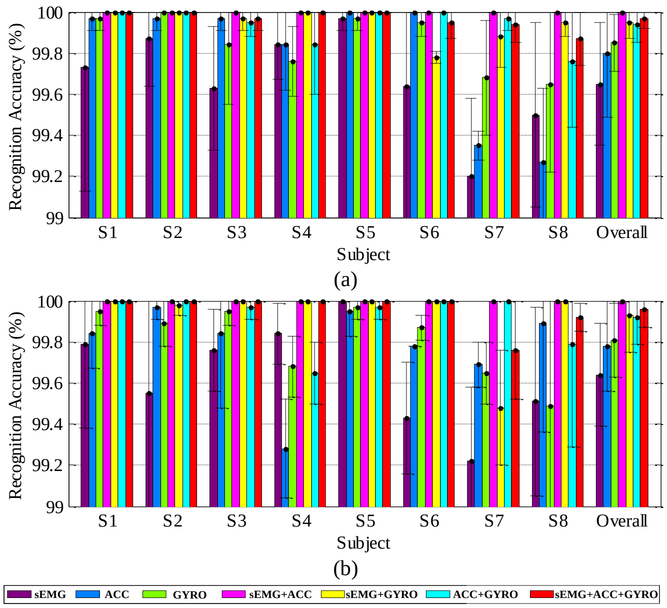

Fig. 6. The classification results of one or two handed classifier under seven different sensor combinations for eight subjects (S1-S8). (a) User-specific test. (b) User-independent test.

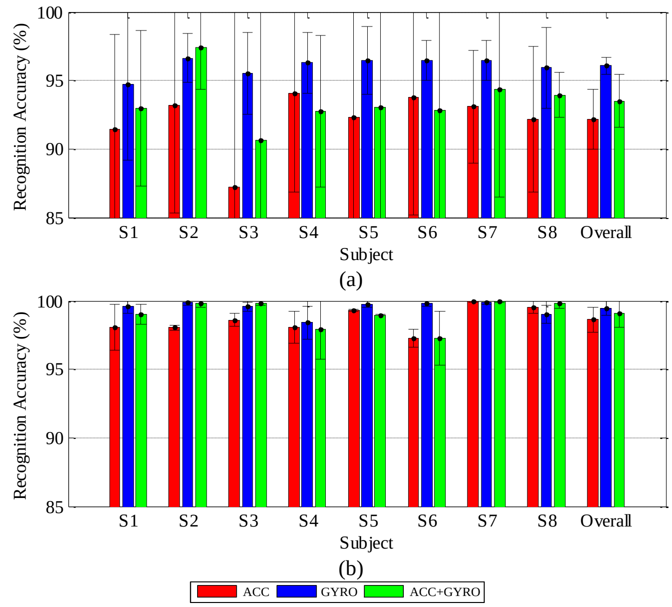

Fig. 7. The classification results of hand orientation classifier under three different sensor combinations for eight subjects (S1-S8). (a) User-specific test. (b) User-independent test.

remaining one for testing in the user-specific classification experiments. For the user-independent classification, the data collected from seven subjects were used for training, and the data from the remaining subject for testing. As mentioned above, different feature vectors extracted from training data and testing data were used for the training and testing of three component classifiers and MSHMM classifier. In other words, although the original data used for training and testing for each classifier are the same, the training feature vectors and testing feature vectors for different classifiers are different.

A. Evaluation Results of Sign-component-based Classifier

According to the classification method presented in Section II, the performance of each sign component classifier on 150 CSL subwords classification was examined in user-specific and user-independent manner, respectively. Fig. 6 shows the classification results of one or two handed classifier under different sensor combinations for eight subjects. The accuracy was always higher than 99% under each testing condition no matter in user-specific or user-independent manner, which demonstrated that the sEMG, ACC and GYRO sensor was able to be applied successfully in one or two handed classifier. Besides, paired-sensor fusion and three-sensor fusion achieved a little higher average classification accuracy than single-sensor condition. A perfect 100% recognition accuracy was achieved with the fusion of sEMG and ACC sensor in both user-specific and user-independent manner, which proved that the fusion of sEMG and ACC sensor was the best combination among the different sensor fusion conditions in one or two handed classifier.

The classification results of hand orientation classifier under three different testing conditions for eight subjects are presented in Fig. 7. It can be observed that the best performance of user-specific and user-independent classification were all obtained under GYRO-only condition, with the overall identification result of 96.07% (SD: 0.65%) in user-specific test

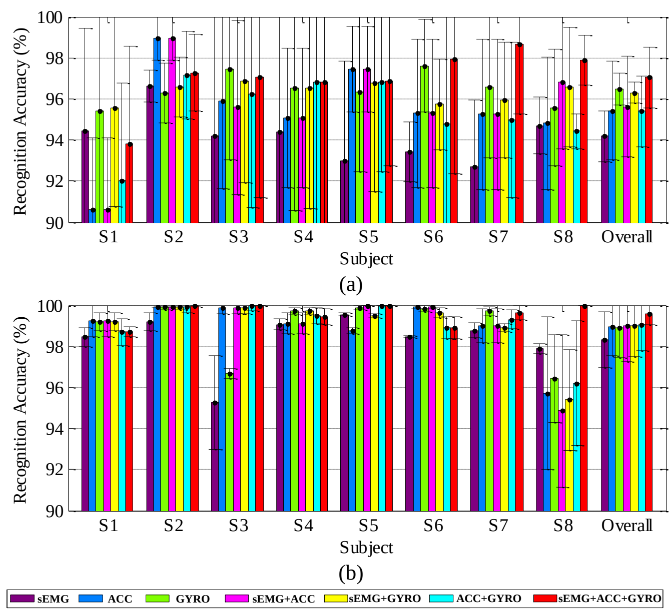

Fig. 8. The classification results of hand amplitude classifier under seven different sensor combinations for eight subjects (S1-S8). (a) User-specific test. (b) User-independent test.

and 99.49% (SD: 0.52%) in user-independent test, respectively. Actually, the results achieved in ACC-only and fusion of ACC and GYRO were also comparatively promising and acceptable with the overall identification rate ranging from 92.18% (SD: 2.17%) to 93.51% (1.93%) in user-specific condition and 98.62% (SD: 0.91%) to 99.08% (SD: 1.01%) in user-independent condition separately. However, the standard deviation of the recognition results obtained with GYRO-only was smaller than that of ACC-only and fusion of ACC and GYRO sensor, which indicated that GYRO sensor achieved much more robust results. Therefore, the comprehensive results illustrated that both ACC and GYRO sensors were well used in hand orientation classifier while relatively higher recognition performance was achieved

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

TABLE II
THE RECOGNITION RESULTS OF 150 CSL SUBWORDS ACROSS EIGHT SUBJECTS (S1-S8) IN BOTH USER-SPECIFIC AND USER-INDEPENDENT CONDITION.

Conditions		1st Test (%)		2nd Test (%)		3rd Test (%)		4th Test (%)		5th Test (%)		Overall (%)	
		Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std	Mean	Std
User-specific	S1	81.6	34.8	97.1	12.4	97.7	7.50	85.9	27.5	97.1	9.90	91.9	7.58
	S2	97.5	9.60	93.8	20.6	95.2	14.0	95.1	14.5	97.7	7.20	95.9	1.68
	S3	93.2	18.1	97.9	8.70	95.7	12.2	94.4	16.4	95.2	15.3	95.3	1.73
	S4	84.9	28.8	95.3	14.6	96.5	12.4	96.4	12.1	94.8	16.0	93.6	4.90
	S5	94.7	17.3	97.1	10.5	95.6	13.7	90.9	19.9	95.7	13.8	94.8	2.32
	S6	91.1	18.0	90.6	18.6	94.3	17.3	93.5	21.6	95.9	17.7	93.1	2.22
	S7	92.4	6.40	95.9	16.1	96.4	17.3	93.9	14.8	92.6	12.2	94.2	1.85
	S8	95.5	15.3	97.9	10.8	96.3	8.40	92.7	14.2	96.8	14.9	95.8	1.94
User-independent	S1	90.2	22.2	87.4	28.8	78.4	37.2	70.7	42.3	83.4	31.8	82.0	7.72
	S2	91.0	22.3	92.6	19.8	85.7	28.5	88.1	24.8	88.9	23.4	89.3	2.66
	S3	90.6	23.9	92.1	22.8	93.1	20.1	92.1	22.9	88.3	24.8	91.2	1.85
	S4	85.8	26.3	82.5	32.5	88.6	24.8	86.5	27.7	84.1	31.1	85.5	2.33
	S5	92.8	23.7	90.9	25.5	93.1	24.5	94.0	22.3	93.3	25.1	92.8	1.16
	S6	81.2	40.5	88.6	38.2	81.6	35.0	80.5	34.0	78.0	40.4	82.0	3.97
	S7	89.6	31.8	88.3	27.3	79.7	32.4	87.1	32.7	85.1	31.5	86.0	3.86
	S8	90.7	26.4	93.7	37.8	85.3	35.1	80.1	25.4	87.2	33.4	87.4	5.23

in GYRO-only condition.

For the hand amplitude classifier, the evaluation results of sEMG, ACC and GYRO sensor in the classification of 150 CSL subwords are illustrated in Fig. 8. The overall recognition results were satisfactory in both user-specific and user-independent condition. As shown in Fig. 8, the sEMG-only condition achieved the lowest overall classification accuracy of 94.18% (SD: 1.24%) in the user-specific manner and 98.34% (SD: 1.35%) in the user-independent manner, respectively. The relatively higher recognition accuracy was obtained in GYRO-only (Mean: 96.48%, SD: 0.78%), fusion of sEMG and GYRO (Mean: 96.31%, SD: 0.50) and three-sensor fusion (Mean: 97.04, SD: 1.46%) conditions in user-specific classification. The average identification rates were approximately close to each other under different sensor combinations in user-independent test, with the best performance (Mean: 99.58%, SD: 0.52%) achieved under three-sensor combination condition. Consequently, the total classification results demonstrated that the fusion of sEMG, ACC and GYRO sensor achieved the best performance of hand amplitude classifier in both user-specific and user-independent condition.

It can be observed from Fig. 7 and Fig. 8 that the classification results of user-independent test in both hand orientation and hand amplitude were superior to the results of user-specific test. In actual, it has been well known that for the fuzzy K-means cluster, more valid training samples are used, higher performance might be achieved. According to the data collection introduced in Section II, the training samples of the user-independent test were much more than those of user-specific test, which might be the key factor that led to the higher classification accuracy of hand orientation classifier and hand amplitude classifier in the user-independent test.

B. Subword Recognition Results

The user-specific and user-independent classification rates of 150 CSL subwords achieved by the proposed tree-structure method are reported in Table II. The 5 tests presented in Table II including 1st Test, 2nd Test, 3rd Test, 4th Test and 5th Test are corresponding to the recognition results of each experiment session respectively. The overall recognition accuracies of 94.31% (SD: 1.41%) and 87.02% (SD: 3.96%) were obtained in user-specific and user-independent condition, respectively. It was not unexpected that the recognition rates in user-specific classification were higher than that in user-independent classification. Actually, there exist great challenges to implement user-independent SLR due to the individual differences of sEMG and difference of inertial signals caused by different performing habits among subjects in the data collection experiments. Therefore, the recognition results achieved in this study were relatively outstanding and satisfactory in both user-specific and user-independent condition.

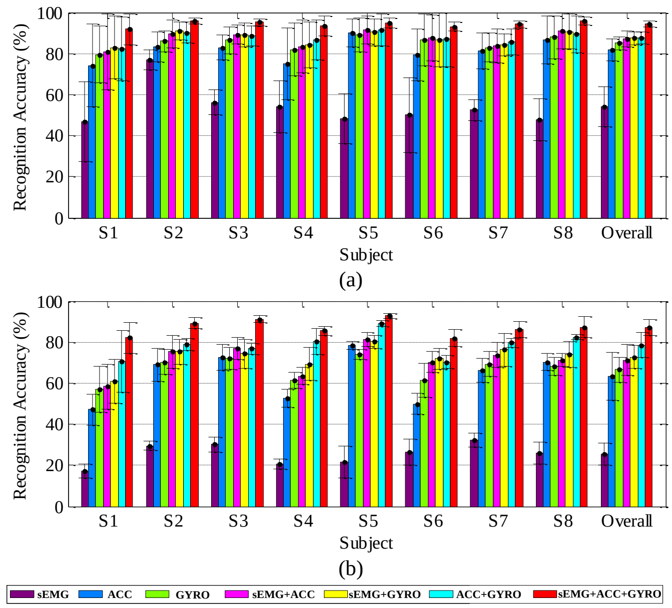

Fig. 9. The recognition results of 150 CSL subwords under seven different sensor fusion strategies for eight subjects (S1-S8). (a) User-specific test. (b) User-independent test.

C. Results Comparison between Different Sensor Fusion Strategies

In order to examine the performance of the proposed sensor fusion strategy on automatic SLR based on sEMG, ACC and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/JBHI.2016.2560907, IEEE Journal of Biomedical and Health Informatics

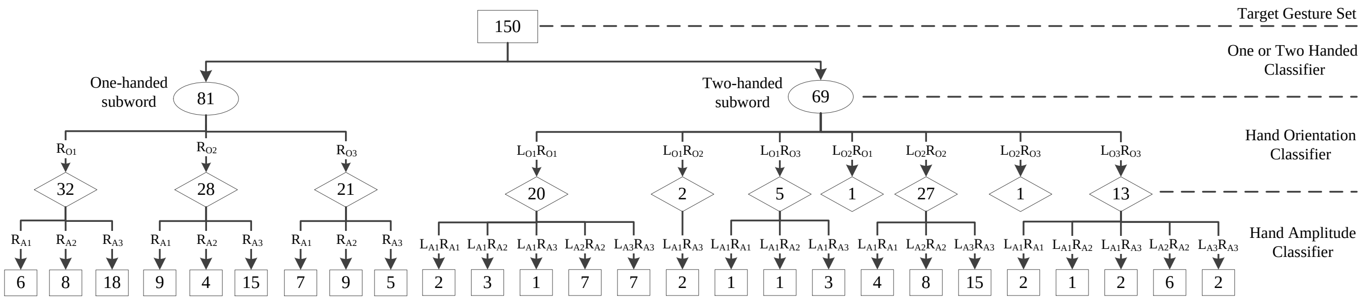

Fig. 10. The variation of candidate number through the rough component-based classifiers of the proposed tree-structure classification framework. L and R mean the left hand and right hand respectively. O_1, O_2 and O_3 indicate three typical hand orientations respectively. A_1, A_2 and A_3 indicate three different hand amplitudes respectively.

GYRO sensors, recognition results with single sensor, paired-sensor and three-sensor fusion on the recognition of 150 CSL subwords across eight subjects are given in Fig. 9. For the recognition tests using single sensor or paired-sensor fusion, the features of the excluded sensors were removed from the tree-structure classification, the branches of the tree-structure classification were only determined by the features from the testing sensors and then the recognition result was directly given at the final level of the tree-structure classification according to the corresponding HMM streams of the remaining sensors. It can be seen from Fig. 9 (a) that the average recognition accuracy achieved by the sensor fusion of sEMG, ACC and GYRO is 94.31% (SD: 1.41%) for the user-specific test, which significantly outperformed the best performance of single sensor (Mean: 85.07%; SD: 3.25%) and paired-sensor fusion (Mean: 87.64%; SD: 2.93%). Similar results were also obtained from user-independent classification as shown in Fig. 9 (b). The average identification rate of 87.02% (SD: 3.96%) was achieved over 150 CSL subwords recognition combining sEMG, ACC and GYRO sensors, with an impressive 20% and approximate 9% improvement than the greatest classification results of single sensor (Mean: 66.69%; SD: 6.02%) and paired-sensor fusion (Mean: 78.52%; SD: 6.16%) respectively. In the single sensor condition, the recognition results of ACC-only or GYRO-only significantly outperformed those of sEMG-only in both user-specific and user-independent classification. The possible reason was that the dynamic gestures involved were much more than the static gestures in the target vocabulary, while the former were performed with large-scale space trajectory which was captured by the inertial sensors easily, but the latter were almost fine motions which might be distinguished effectively using sEMG sensors. For the paired-sensor fusion condition, the identification rates were all higher than those of single sensor conditions whether in user-specific or user-independent test. Therefore, it is clear that our proposed fusion strategy for SLR system based on sEMG, ACC and GYRO sensors was effective in enhancing the recognition performance of 150 CSL subwords in both user-specific and user-independent classification.

IV. Discussion and Future Work

In order to implement reliable large-vocabulary CSL subword recognition, different from related researches, the main contributions of this study can be summarized as follows: 1) expanding the target gesture set to 150 frequently used CSL subwords; 2) sensing gesture information with sEMG, ACC and GYRO sensors (integrated in a form of home-made portable and wearable wristband) and evaluating the contribution of each kind of sensor in three common sign component classifiers; 3) improving the tree-structure classification method for CSL subwords based on the fusion of three kinds of sensors and demonstrating the effectiveness of sensor fusion strategy with sEMG, ACC and GYRO sensors.

It is well known that the component-based SLR is a promising and effective approach to expand the sign vocabulary. Compared with our pilot studies [8], [9], [28], this study introduced the GYRO sensor into the tree-structure classification and evaluated the classification abilities of sEMG, ACC, GYRO and their combinations in three common sign components and subwords. The feature of the optimized tree-structure classification framework is that three preliminary component classifiers were adopted to reduce the searching range of MSHMM classifier, which could be helpful to improve the recognition performance and to cut down the recognition time consumption. Fig. 10 gives an example of the candidate number variation in an ideal classification condition using the optimized tree-structure classification. As shown in Fig.10, the candidate number was reduced level by level in the tree structure classification, from 150 at the top of the tree to 1~18 for the final MSHMM classification. Based on the comparison between the proposed tree-structure classification method with the MSHMM-only method, where only multi-stream HMMs were utilized to recognize CSL subwords, we found that the proposed tree-structure classification framework could not only increase approximate 3% recognition accuracy, but also significantly reduce the time consumptions from 1.79 second per word (s/w) to about 0.40 s/w (testing platform: Intel Xeon E5-2650 at a 2.6-GHz CPU with a 16-GB RAM), which demonstrated the effectiveness of tree-structure classification in enhancing the performance of SLR.

The combination of sEMG and inertial sensors is a promising sensing technique which can facilitate the implementation of portable and wearable SLR systems a lot. Many efforts have been made toward this direction and the main differences of several recent studies associated with the fusion of multi-sensor information for SLR are summarized in Table III for

TABLE III
RELATIVE STUDIES WITH DIFFERENT COMBINATIONS OF sEMG AND INERTIAL SENSORS FOR SLR

Studies	Sensor Type of Each Hand	Handed	Subject Numbers	Language	Vocabulary	Recognition Accuracies	
Kosmidou [3]	1 3-D ACC 5 bipolar sEMG	Single	3	Greek SL	60 words	User-specific	> 93%
Zhang [8]	1 3-D ACC 5 bipolar sEMG	Single	10	Natural Gesture	18 gestures	User-specific User-independent	97.6% 90.2%
Boschmann [23]	1 3-D ACC 1 3-D GYRO 4 bipolar sEMG	Single	2	Natural Gesture	8 gestures	User-independent	≈ 90%
Wu [24]	1 3-D ACC 1 3-D GYRO 4 bipolar sEMG	Single	4	American SL	40 words	User-specific User-independent	95.94% ≈ 40%
Li [28]	1 3-D ACC 4 bipolar sEMG	Double	5	Chinese SL	121 subwords	User-specific	96.5%
The proposed	1 3-D ACC 1 3-D GYRO 4 bipolar sEMG	Double	8	Chinese SL	150 subwords	User-specific User-independent	94.3% 87.0%

comparison. As observed from Table III, the overall recognition accuracy in user-specific classification achieved in this study is a bit lower than the results presented in [8], [24], [28], while the target gesture set is extended further. For the user-independent classification, the current study indicates a great progress on portable and wearable SLR systems considering the relatively large vocabulary of signs and the promising recognition performance. Additionally, only user-specific classification has been targeted in numerous studies such as the work described in [3], [9], [19], [28]. Although the user-specific SLR system could be flexible to develop with relatively high recognition accuracy, the burden of data collection for training models before using the SLR system and the weak universality will hinder the convenient and natural properties of HCI. Therefore, it is of great importance to explore the user-independent SLR. As far as we know, this study is the first attempt toward large-vocabulary user-independent SLR based on a portable and wearable wristband-like sensing system. Considering the challenge and complexity of user-independent classification due to the individual differences of execution speed and movement scale, the recognition performance achieved in this study is still satisfactory, which will facilitate the implementation of actual SL interpreter.

In addition, the significance of sensor fusion with sEMG and ACC sensors in enhancing the performance of hand gesture recognition has been investigated in our pilot studies [8], [9], [22]. In this study, the good recognition results indicate the potential of the proposed fusion technique of sEMG, ACC and GYRO for the realization of large-vocabulary SLR system. The overall recognition accuracy in three-sensor fusion condition is higher than that of fusion with sEMG and ACC sensor with the improvement of 7.2% in user-specific test and 15.7% in user-independent test, respectively, which demonstrated the poential of GYRO sensor in large vocabulary SLR. The complementary functionality of sEMG, ACC and GYRO signals has also been explored by Boschmann et al. [23] in the recognition of 8 hand and wrist motions. This paper further expands the same finding for CSL subword recognition with a larger vocabulary according to our proposed fusion strategy.

This study presents an automatic SLR method based on the combination of wearable sEMG and inertial sensors. The promising performances especially on user-independent recognition lay a foundation for the implementation of universal SLR system. Portable SLR systems with the proposed recognition framework will be explored in the future. In addition, it is reported that the recognition rate depends a lot on the feature selection [33]. Therefore, a thorough and comprehensive exploration on the feature extraction and selection will become another part of our future effort to further improve the recognition performance.

V. CONCLUSION

This paper explored the application of wearable sEMG, ACC and GYRO sensors in Chinese sign language subword recognition. After the evaluation of sEMG, ACC, GYRO and their combinations in the classification of three common sign components, an optimized tree-structure CSL subword classification framework was proposed. Experimental results on the identification of 150 CSL subwords across eight subjects validated the effectiveness of the proposed tree-structure classification method with the overall accuracies of 94.31% in user-specific test and 87.02% in user-independent test, respectively. Our study has laid a good foundation for the realization of large-vocabulary SLR systems based on portable and wearable sensors.

ACKNOWLEDGMENT

The authors are grateful to all the subjects for their participation in the data collection experiments in this study. Our special thanks also go to Dr. Z. Lu and R. Su for their constructive suggestions on the experimental scheme.

REFERENCES

[1] J. Ma, W. Gao, J. Wu, and C. Wang, "A continuous Chinese Sign Language recognition system," in *Proc. 4th Int. Conf. IEEE Automatic Face and Gesture Recognition* , 2000, pp. 428-433.
[2] V. E. Kosmidou, L. J. Hadjileontiadis, and S. M. Panas, "Evaluation of surface EMG features for the recognition of American Sign Language

gestures," in *Proc. 28th Ann. Int. IEEE EMBS Conf.*, 2006, pp. 6197-6200.

[3] V. E. Kosmidou and L. J. Hadjileontiadis, "Sign language recognition using intrinsic-mode sample entropy on sEMG and accelerometer data," *IEEE Transactions. Biomedical Engineering*, vol. 56, pp. 2879-2890, Dec. 2009.

[4] M. E. Al-Ahdal and N. M. Tahir, "Review in sign language recognition systems," in *Proc. IEEE ISCI*, 2012, pp. 52-57.

[5] S. C. Ong and S. Ranganath, "Automatic sign language analysis: A survey and the future beyond lexical meaning," *IEEE Transactions. Pattern Analysis and Machine Intelligence*, vol. 27, pp. 873-891, Jun. 2005.

[6] Y. Guo, Q. Wang, S. Huang, and A. Abraham, "Hand gesture recognition system using single-mixture source separation and flexible neural trees," *Journal of Vibration and Control*, vol. 20, no. 9, pp. 1333-1342, 2014.

[7] P. Kishore and P. R. Kumar, "A video based Indian sign language recognition system (INSLR) using wavelet transform and fuzzy logic," *IACSIT International Journal of Engineering and Technology*, vol. 4, no. 5, pp. 537-542, Oct. 2012.

[8] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A framework for hand gesture recognition based on accelerometer and EMG sensors," *IEEE Transactions. Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 41, no. 6, pp. 1064-1076, Nov. 2011.

[9] Y. Li, X. Chen, J. Tian, X. Zhang, K. Wang, and J. Yang, "Automatic recognition of sign language subwords based on portable accelerometer and EMG sensors," in *Int. Conf. Multimodal Interfaces and the Workshop on Machine Learning for Multimodal Interaction*, 2010, pp. 1-7.

[10] W. Gao, G. Fang, D. Zhao, and Y. Chen, "Transition movement models for large vocabulary continuous sign language recognition," in *Proc. 6th Int. Conf. IEEE Automatic Face and Gesture Recognition*, 2004, pp. 553-558.

[11] K. Li, Z. Zhou, and C.-H. Lee, "Sign Transition Modeling and a Scalable Solution to Continuous Sign Language Recognition for Real-World Applications," *ACM Transactions. Accessible Computing*, vol. 8, no. 2, pp. 1-23, Jan. 2016.

[12] M. Zhang and A. A. Sawchuk, "Human daily activity recognition with sparse representation using wearable sensors," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, pp. 553-560, May. 2013.

[13] Z. Lu, X. Chen, Q. Li, X. Zhang, and P. Zhou, "A hand gesture recognition framework and wearable gesture-based interaction prototype for mobile devices," *IEEE Transactions. Human-Machine Systems*, vol. 44, no. 2, pp. 293-299, Apr. 2014.

[14] A. H. Al-Timemy, G. Bugmann, J. Escudero, and N. Outram, "Classification of finger movements for the dexterous hand prosthesis control with surface electromyography," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, pp. 608-618, May. 2013.

[15] K. Dermitzakis, A. H. Arieta, and R. Pfeifer, "Gesture recognition in upper-limb prosthetics: A viability study using dynamic time warping and gyroscopes," in *Proc. Ann. Int. Conf. IEEE Engineering in Medicine and Biology Society*, 2011, pp. 4530-4533.

[16] S. H. Roy, M. Cheng, S.-S. Chang, J. Moore, G. De Luca, S. Nawab, and C. J. De Luca, "A combined sEMG and accelerometer system for monitoring functional activity in stroke," *IEEE Transactions. Neural Systems and Rehabilitation Engineering*, vol. 17, no. 6, pp. 585-594, Dec. 2009.

[17] S. Amsuss, P. M. Goebel, N. Jiang, B. Graimann, L. Paredes, and D. Farina, "Self-correcting pattern recognition system of surface EMG signals for upper limb prosthesis control," *IEEE Transactions. Biomedical Engineering*, vol. 61, no. 4, pp. 1167-1176, Apr. 2014.

[18] L. K. Yun, T. T. Swee, R. Anuar, Z. Yahya, A. Yahya, and M. R. A. Kadir, "Sign Language Recognition System using SEMG and Hidden Markov Model," in *Recent Advances in Mathematical Methods, Intelligent Systems and Materials*, 2013, pp. 50-53.

[19] J.-S. Wang and F.-C. Chuang, "An accelerometer-based digital pen with a trajectory recognition algorithm for handwritten digit and gesture recognition," *IEEE Transactions. Industrial Electronics*, vol. 59, no. 7, pp. 2998-3007, Jul. 2012.

[20] S. Chernbumroong, S. Cang, and H. Yu, "Genetic Algorithm-Based Classifiers Fusion for Multisensor Activity Recognition of Elderly People," *IEEE Journal of Biomedical and Health Informatics*, vol. 19, pp. 282-289, Jan. 2015.

[21] J. Cannan and H. Hu, "Feasibility of Using Gyro and EMG Fusion as a Multi-Position Computer Interface for Amputees," in *Proc. 4th Int. Conf. Emerging Security Technologies*, 2013, pp. 75-78.

[22] X. Chen, X. Zhang, Z.-Y. Zhao, J.-H. Yang, V. Lantz, and K.-Q. Wang, "Hand gesture recognition research based on surface EMG sensors and 2D-accelerometers," in *Proc. 11th IEEE ISWC*, 2007, pp. 11-14.

[23] A. Boschmann, B. Nofen, and M. Platzner, "Improving transient state myoelectric signal recognition in hand movement classification using gyroscopes," in *Proc. 35th Ann. Int. IEEE Engineering in Medicine and Biology Society*, 2013, pp. 6035-6038.

[24] J. Wu, Z. Tian, L. Sun, L. Estevez, and R. Jafari, "Real-time American sign language recognition using wrist-worn motion and surface EMG sensors," in *Proc. 12th Int. Conf. IEEE Wearable and Implantable Body Sensor Networks*, 2015, pp. 1-6.

[25] J.-U. Chu, I. Moon, and M.-S. Mun, "A real-time EMG pattern recognition system based on linear-nonlinear feature projection for a multifunction myoelectric hand," *IEEE Transactions. Biomedical Engineering*, vol. 53, pp. 2232-2239, Nov. 2006.

[26] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, pp. H2039-H2049, 2000.

[27] X. Zhang and P. Zhou, "Sample entropy analysis of surface EMG for improved muscle activity onset detection against spurious background spikes," *Journal of Electromyography and Kinesiology*, vol. 22, pp. 901-907, Jun. 2012.

[28] Y. Li, X. Chen, X. Zhang, K. Wang, and Z. J. Wang, "A sign-component-based framework for chinese sign language recognition using accelerometer and sEMG data," *IEEE Transactions. Biomedical Engineering*, vol. 59, no. 10, pp. 2695-2704, Oct. 2012.

[29] A. Wang, W. Yuan, J. Liu, Z. Yu, and H. Li, "A novel pattern recognition algorithm: Combining ART network with SVM to reconstruct a multi-class classifier," *Computers & Mathematics with Applications*, vol. 57, pp. 1908-1914, 2009.

[30] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification*, 2nd ed. New York: Wiley, 2001, Section 10.4.4.

[31] G. Fang, W. Gao, and D. Zhao, "Large vocabulary sign language recognition based on fuzzy decision trees," *IEEE Transactions. Systems, Man and Cybernetics, Part A: Systems and Humans*, vol. 34, no. 3, pp. 305-314, May. 2004.

[32] R. N. Khushaba, L. Shi, and S. Kodagoda, "Time-dependent spectral features for limb position invariant myoelectric pattern recognition," in *Communications and Information Technologies (ISCIT), 2012 International Symposium on*, 2012, pp. 1015-1020.

[33] F. Al Omari, J. Hui, C. Mei, and G. Liu, "Pattern Recognition of Eight Hand Motions Using Feature Extraction of Forearm EMG Signal," *Proceedings of the National Academy of Sciences, India Section A: Physical Sciences*, vol. 84, pp. 473-480, 2014.