A Signer-Independent Sign Language Recognition System Based on the weighted KNN/HMM

Siqi Liu
Department of Electronics Information Engineering
Xi'an Technological University
Xi'an, China
837928697@qq.com

Qinkun Xiao
Department of Electronics Information Engineering
Xi'an Technological University
Xi'an, China
xiaoqinkun10000@163.com

***Abstract*—Aiming at the feature of signer-independent sign language recognition the training data complexity caused by mass data and noticeable distinctions between different people data, the weighted KNN/HMM model is presented in this paper. This model is made of two blocks, which is part of sign language classification and recognition. In classing part, the KNN (K-Nearest Neighbor, KNN) is used to learn the training samples. Considering the different contributions of sign language features to pattern classification we give different weight to different characteristics. And the category of test sample is decided by the sum of weighted distance. In recognition part, weighted KNN classification result is taken as the state-input of HMM (Hidden Markov models, HMM) to implement sign language recognition, combine with the ability to temporal data modeling and fuzzy inference of HMM model. Experiment results show that weighted KNN/HMM sign language recognition model are efficient on either recognition speed or recognition rate.**

Keywords-Signer-Independent Sign Language Recognition; Hidden Markov Models; weighted K-Nearest Neighbor

I. INTRODUCTION

The aim of Sign Language Recognition is to provide efficient and accurate mechanisms to translate sign language into text or speech. Most studies in the sign language recognition field always employ HMM used in recognition and the SVM (Support Vector Machine, SVM) or KNN used in classification. The [2] proposed approach is to find the optimal weights via Artificial Bee Colony (ABC) algorithm [3], in order to improves the correct classification performance in Iris, Hagerman, and Breast Cancer data set. The [4] add the depth information to effectively locate the 3D position of the hands in the sign language recognition system, then using the hidden Markov models to recognize the variety of sign language movement changing on the time domain. The [5] proposed algorithm represents the first real attempt in the Italian Sign Language case. The automatic recognizer is based on Hidden Markov Models (HMMs) and video features have been extracted using the OpenCV open source library, the rate of recognition is 45.23%. The [6] proposed a continuous sign language recognition system for analytical reasoning by Product-HMMs method of sign language signal, test accuracy rate reaches 89%. In the [7] frequency cepstral coefficients is used to extract speech features. K nearest neighbor classifier is used for the first time to author knowledge in Pashto language to classify the features of speech and compare its accuracy with linear discriminate analysis. The overall average recognition exactitude of 76.8% is obtained.

But the study on the sign language recognition are based on unspecific individuals are only less compared with specific individuals. Therefore, proposed the weighted KNN/HMM model solve the following problems:(1) It is too complexity to analysis sign language signal cased by the limitation topological structure of HMM,(2) HMM require each gesture vector corresponding to each state of gesture is independent and lacks classification characteristics. In this paper, a sign language database is established through collection different gesture of people, we can obtain related sequences according to classification functional of KNN algorithm, to increase the classification characteristic of HMM recognition method and reduce the number of estimated parameters of the model, improving sign language recognition efficiency.

II. HMM MODULE

HMM (Hidden Markov Models) is a statistical method which is well-known and use widely. It uses Markov chain to simulate the change of statistical properties and this change is indirectly described by observing sequence.

A. HMM Model mathematical description

HMM [8][9] is a typical stochastic finite automata, each observation node corresponds to a specific state of the node, the HMM graphics mode is as shown in figure 1, in which the gray nodes represent observations nodes and the white nodes represent implied node.

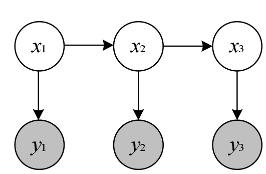

Figure 1. The HMM Graphics Mode

The implicit sequence and observation sequence of 1: t is represented by $\mathbf{X_t}$, $\mathbf{Y_t}$. The x_t is used to express the

978-1-4799-8646-0/15 $31.00 © 2015 IEEE
DOI 10.1109/IHMSC.2015.71

implied variable value of t moment and the y_t is used to express the observed variables value of t moment. X_t represents a implicit variable that presence n states, i.e. $x_t \in \{1,2,3,\ldots,n\}$. Y_t represents a discrete variable having m kind of possible values and $y_t \in \{c_1,c_2,\ldots,c_m\}$, which c represents any kind of possible values and also represents an m-dimensional feature vector V^m.

There are three main model parameters of HMM when the model output parameter is defined as $P(y_t \mid x_t)$:

(1)The initial hidden state distribution matrix: $\boldsymbol{\pi}=(\pi_i)_{1\times n}$ and $\pi_i=P(x_1=i)$.

(2)The implicit state transition matrix: $\mathbf{A}=(a_{ij})_{n\times n}$ and $a_{ij}=P(x_t=i|x_{t-1}=j)$.

(3) The observation matrix: $\mathbf{B}=(b_i)_{1\times n}$, we can use matrix B to represent the output parameters if the observation value is a finite discrete signal to obtained HMM third argument, which $b_i(k)=P(y_t=k|x_t=i)$.

Above all, HMM parameters are simplified for：$\boldsymbol{\lambda} = (\mathbf{A}, \mathbf{B}, \boldsymbol{\pi})$

If the observed sequence Y_T is known and the HMM parameters $\boldsymbol{\lambda}$ is unknown, the observing sequence training and learning based on some algorithms when the sign of an unknown language data for testing. Finally the most consistent system parameters of the observed sequence are obtained, which making the probability *of* $P(\mathbf{Y}_T|\boldsymbol{\lambda})$ maximum. To find the best parameter $\boldsymbol{\lambda}^*$:

$$\boldsymbol{\lambda}^* = \arg\max_{\lambda} P(\mathbf{Y}, \mathbf{X} \mid \boldsymbol{\lambda}) \qquad (1)$$

B. HMM Reasoningaaa

We could use the decoding algorithm of Viterbi to find out the most likely implied sequence $\mathbf{X}_T$ if the parameter λ and the observation sequence $y_1,y_2,\cdots,y_T$ is known. To find the implied sequence of X meets the following formula:

$$\hat{\mathbf{X}} = \arg\max_{\mathbf{x}} P(\mathbf{X}, \mathbf{Y} \mid \boldsymbol{\lambda}) = \arg\max_{\mathbf{x}} P(\mathbf{X} \mid \mathbf{Y}, \boldsymbol{\lambda}) \qquad (2)$$

The variables of the Viterbi is defined as: $\delta_t(i)=\max_{x(1),\cdots x(t-1)} P(x_1,\cdots x_{t-1},x_t=i,\mathbf{Y}_T|\lambda)$ and the $\triangle_t(i)$ is defined as the implicit sequence state of t-1 moment ago.

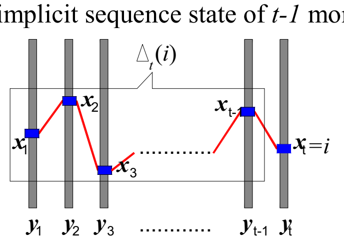

Figure 2. the Meaning Schematic Of $\triangle_t(i)$

The process of Viterbi decoding algorithm is shown as follows:

initialization: $\delta_1(i)=\pi_i b_i(y_1) \qquad 1 \le i \le n$

$\Delta_1(i)=0$

recursion: $\delta_t(j)=[\max_{1\le i\le n} \delta_{t-1}(i)a_{ij}]b_j(y_t) \qquad 2\le t\le T, 1\le j\le n$

$\Delta_t(j)=\arg\max_{1\le i\le n}[\delta_{t-1}(i)a_{ij}]b_j(y_t)$

calculation: $P(\hat{X},Y\mid\lambda)=\max_{1\le i\le n}[\delta_T(i)]$

traceback: $\hat{X}_T=\arg\max_{1\le i\le n}[\delta_T(i)]$

$\hat{X}_t=\Delta_{t+1}(\hat{X}_{t+1}) \qquad t=T-1,T-2,\ldots,1$

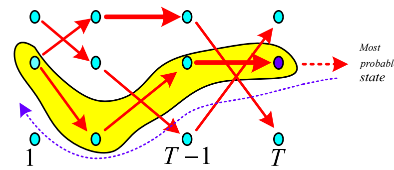

Figure 3. Reasoning Process Diagram

III. The Weighted KNN Module

KNN (K-Nearest Neighbor, KNN) [10] is a very famous statistical algorithm in pattern recognition field and a kind of concise and effective nonparametric classification method, which also can be used in the field of sign language recognition.

A. KNN module introduction

The KNN algorithm is to find the K nearest neighbor of test sample in the training sample, which decides the category of test sample according to the K nearest neighbor. The method as following shown: given a test sample z and find out its K neighbor samples $\{z_{1NN},\ldots,z_{KNN}\}$ in the inside library of training sample, then calculated Euclidean distance between z and neighbor samples $\{d^2_{z,z_{1NN}},\ldots,d^2_{z,z_{KNN}}\}$,

$$d^2_{Z,Z_{iNN}} = \| z - z_{iNN} \|_2^2 \;, \; i=1,2,\ldots,K \qquad (3)$$

Determining the adjacent values by calculate Euclidean distance of the test samples, the smaller distance shows that the greater similarity, otherwise shows the smaller similarity.

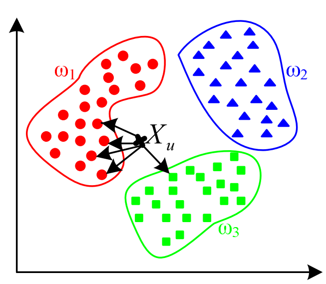

Figure 4. KNN Classification Schematic Diagram

B. The Weighted KNN module

Because the feature vector property of KNN classification method is irrelevant that resulting from the classification bias, so we gives different weight[11] to different characteristics which have different influence to classification process. Selecting the weights corresponding to the K neighbor samples { $w_{z,z_{1NN}}, \ldots, w_{z,z_{KNN}}$ }.

$$w_{z,z_{iNN}} = \exp(-d^2_{z,z_{iNN}} / 2\sigma^2) \; i=1,\ldots,K \qquad (4)$$

σ is nuclear parameter and its size determines the change speed of weight. To calculate the weighted average of K neighbor samples,

$$z_{mean} = \frac{\sum_{i=1}^{K} w_{Z,Z_{iNN}} \bullet z_{iNN}}{\sum_{i=1}^{k} w_{Z,Z_{iNN}}} \qquad (5)$$

the Euclidean distance of K neighbor samples and z_{mean}.

$$d^2_{Z,Z_{mean}} = \| z - z_{mean} \|_2^2 \qquad (6)$$

Compared with the distance and threshold t, if $d^2_{z,z_{mean}} \le t$ and z is considered as the inside library of sample, otherwise the z samples belong to outside library samples.

Given different weights to K neighbor sample of test sample library through the weighted KNN classifier. These weights changed with the Euclidean distance in a decay exponential curve, making neighbors samples that close to test samples obtain larger weights. Compared with the traditional KNN, the weighted KNN is a kind of locality classifier has stronger classification function, which given different weights to K neighbor samples through the distance and emphasize the contribution of local information.

IV. THE WEIGHTED KNN/HMM SYSTEM MODEL

A. Input Device: This paper uses a single camera to capture single hand movements, implemented the sign language gesture recognition system based on MATLAB.

B. Training: The recognition experiment consists in a database that we collecting 30 isolated-sign images for 10 signer, namely, the inside library of samples. Training and identify the parameters of test sample by the weighted KNN/HMM module to get their semantics.

When input a test sample, detecting whether it belong to the inside sample of library or not. We could select a KNN classifier by pretreatment and classification training to obtain classification results if it was. The classification results are taken as the state-input of HMM model to establishment HMM template library after learning, and then use HMM recognizer to identify the semantic. To find the sequence meet the conditions when an unknown sign language data for testing, which the semantics corresponding to the implicit sequence is the recognition result. The weighted KNN model flow chart as show in figure 5, HMM model flow chart as show in figure 6.

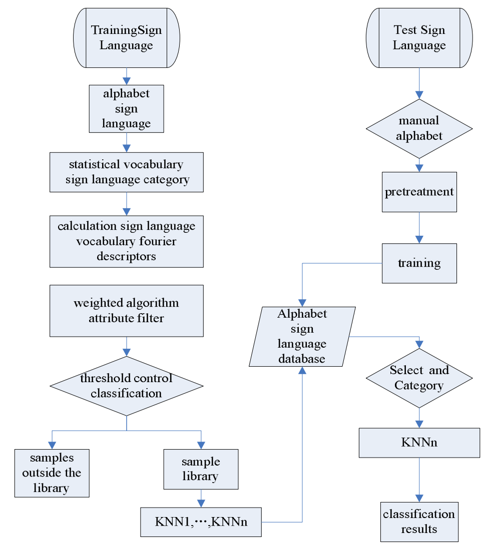

Figure 5. The Weighted KNN Model Flow Chart

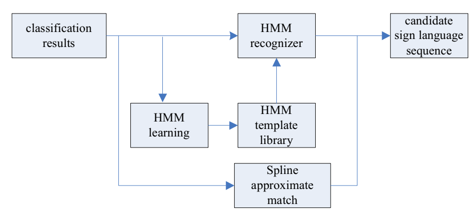

Figure 6. The HMM Model Flow Chart

V. THE EXPERIMENTAL RESULTS AND CONCLUSIONS

The 30 alphabet sign language are A, B, C, D, E, F, G, H, J, K, L, M, N, O, P, Q, S, T, U, W, X, Y, Z, CH, SH, ZH, and NG. Alphabet V, for example, figure 7 shows different alphabet sign language of 10 different people.

Figure 7. Collection Of Alphabet Sign Language To 10 Different People

In this experiment, we take the 300 samples as training samples and randomly input 150 gestures for recognition. In the classifier, the result changed in classification corresponding to the value of k. The classification precision rate and recall rate is shown in figure 8, which we can see the precision increased while the K value increased. At the same time, we compared with the recognition rate of traditional HMM model and the weighted KNN/HMM when the K value is changed. The result is shown in table I.

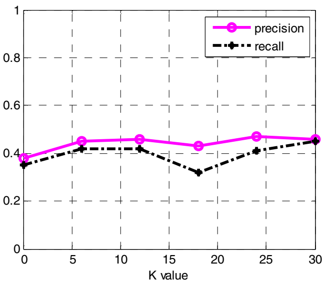

Figure 8. KNN Classification Rate

TABLE I. MODEL IDENTIFICATION

Recognition Model	Recognition Model					
	K=5	*K=10*	*K=15*	*K=20*	*K=25*	*K=30*
Weighted KNN/HMM	41.41	48,.83	48.44	44.93	51.17	46.48
HMM	23.44	26.56	33.21	34.32	28.52	29.33

The weighted KNN/HMM hybrid method combined with the classification function of Weighted KNN and the statistical properties of HMM. The test shows that system based on the weighted KNN/HMM is feasible and practical. The training and recognition speed of this model improve significantly compared with the HMM-based sign language recognition system, but the identification rate needs to be improved because of the feature vector expression inaccurate.

ACKNOWLEDGMENT

This work is partly supported by the National Basic Research Project of China (No. 2010CB731800) and the China National Foundation (No. 60972095, 61271362).

REFERENCES

[1] Jiang Li.gesture recognition technology based on CAS-GLOVE data glove [D].Beijing Jiaotong University,2006

[2] Halil Yigit.A weighting approach for KNN classifier[J].Electronics,Computer and Computation (ICECCO), 2013,228 – 231

[3] D.Karaboga,B. Gorkemli,C. Ozturk, and N. Karaboga,"A Comprehensive Survey:Artificial Bee Colony (ABC) Algorithm and Applications,"Artificial Intelligence Review,doi:10.1007/s10462-012-9328-0, 2012.

[4] Yeh-Kuang Wu,Hui-Chun Wang,Liung-Chun Chang,and Ke-Chun Li.S. Ramanna.Using HMMs and Depth Information for Signer-Independent Sign Language Recognition.(Eds.): MIWAI 2013, LNCS 8271, pp. 79–86, 2013.

[5] Chang CC, Lin CJ (2011) LIBSVM:A library for support vector machines. ACM Trans Intell Syst Technol 2:1–27

[6] Shin-Han Yu.Chung-Lin Huang;Shih-Chung Hsu;Hung-Wei Lin,Hau-Wei Wang Vision-Based Continuous Sign Language Recognition using Product HMM Pattern Recognition (ACPR), 2011 First Asian Conference on 28-28 Nov.2011 pp510 – 514

[7] Zakir Ali·Arbab Waseem Abbas,·T.M. Thasleema·Burhan Uddin, Tanzeela Raaz·Sahibzada Abdur Rehman Abid, Database development and automatic speech recognition of isolated Pashto spoken digits using MFCC and K-NN,Int J Speech Technol (2015) 18:271–275DOI 10.1007/s10772-014-9267-z

[8] Xiaoqin Kun, Gao Song Bayesian Networks and Intelligent Information Processing [M],Beijing National Defense Industry Press,

[9] McGuire R M, Hernandez-Rebollar J,Starner T. Towards a one-way American sign language translator[C]//IEEE International Conference on Automatic Face and Gesture Recognition, Seoul, Korea.IEEE Computer Society Press, 2004: 620- 625.

[10] R. Duda,P .Hart .Pattern Classifieation and Seene Analysis[M].New York :John and Wiley&Sons,1973.

[11] Wasito I,Mirkin B. K Nearest Neighbour approach in the least-squares data imputation algorithms[J].Information Sciences,2005,169(l-2):1-25.