Sign Language Recognition Based on Trajectory Modeling with HMMs

Junfu Pu, Wengang Zhou, Jihai Zhang, Houqiang Li

University of Science and Technology of China

Introduction

Sign language is one of the most important ways for communication between the deaf and the normal people. With broad social impact, this problem has attracted considerable attention from many researchers around the world. Sign language recognition (SLR) targets on interpreting and understanding the sign language for convenience of communication between the deaf and the normal people. The problem is challenging due to the large variations for different signers and the subtle difference between sign words. In this paper, we propose a new method for isolated sign language recognition based on trajectory modeling with hidden Markov models (HMMs). The method is based on trajectories of sign words. The data captured by Kinect consists of a set of 3D points, which are the axis locations of joints in each temporal stamp. We use the trajectories of both hands for recognition.

In our approach, we first normalize and re-sample the raw trajectory data and partition the trajectory into multiple segments. To represent each trajectory segment, we proposed a new curve feature descriptor based on shape context. After that, hidden Markov model is used to model each isolated sign word for recognition.

Methods

Framework

The framework of our method is shown in Fig. 1. The steps in the left column consist the stage of training. For a testing trajectory of sign word, the HMMs that we've trained are used for recognition.

Feature Extraction

Fig. 3 shows the procedure for curve feature extraction.

1. **Shape context**
   As is shown in Fig. 2, shape context describes the distribution of the other points in the neighborhood of a reference point.
2. **Codebook Training**
   K-means algorithm is used for training the codebook.
3. **Quantization**

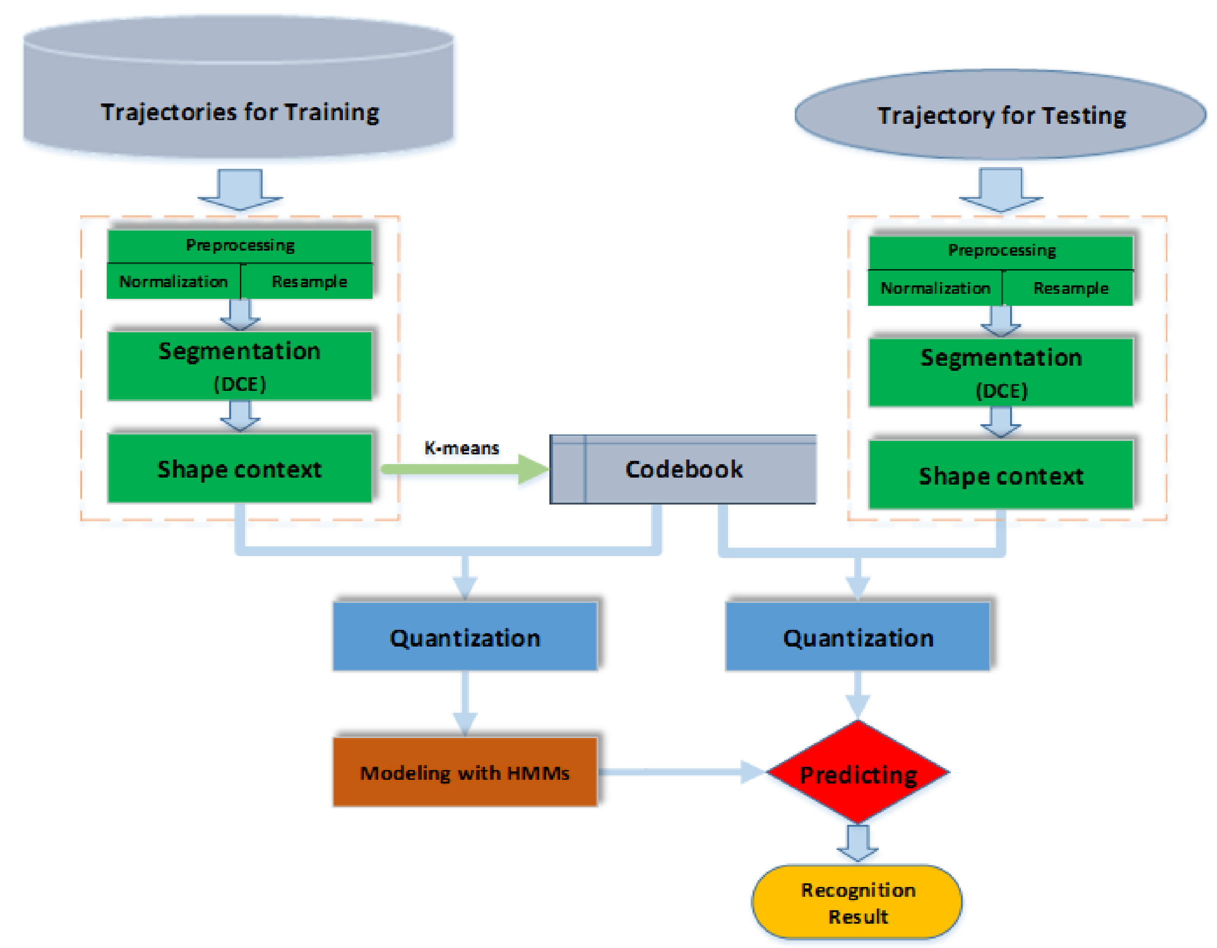

Fig. 1 The illustration of framework

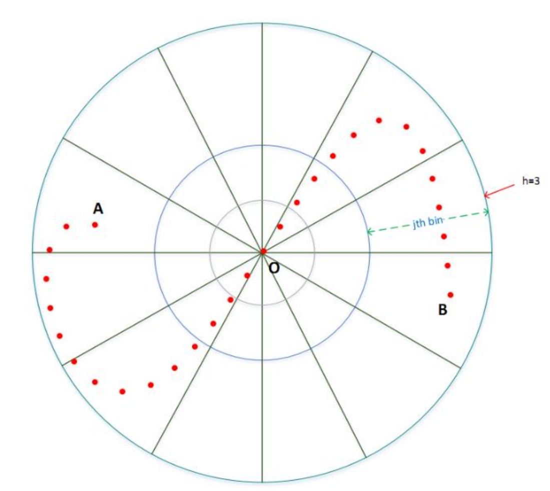

Fig. 2 Shape context

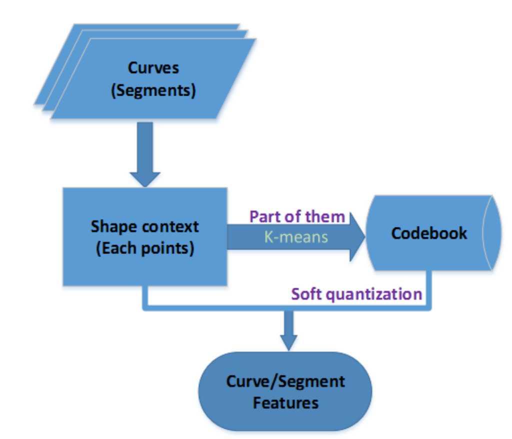

Fig. 3 Curve feature extraction

Character Modeling by HMMs

1. **Preprocessing**
2. **DCE Algorithm**
   We use DCE algorithm for trajectory segmentation.
3. **HMM Modeling**
   The HMM is modeled by the parameter vector $\lambda = (A, B, \pi)$, where A is the transition probability matrix, B is the emission probability matrix. Given the unknown testing sequence, we classify it to class C_p with the following rule:

$$C_p = \max_{C_i} \log p(O|\lambda_{C_i}), i = 1, 2, \ldots, N$$

Results

Table 1. The recognition rates for different methods on Subset A

Subset A	Top1	Top5	Top10
Normal HMM	0.322	0.548	0.653
CM_VoM	0.576	0.824	0.894
Our method	0.673	0.866	0.898

Table 2. The recognition rates for different methods on Subset B with unseen signers

Subset B	Top1	Top5	Top10
Normal HMM	0.125	0.271	0.386
CM_VoM	0.451	0.719	0.819
Our method	0.544	0.773	0.827

Dataset

Subset A

100 sign word, performed by 14 signers for 5 times

Subset B

100 sign word, performed by 36 signers for 5 times

Table 1 and Table 2 show the recognition rates on both datasets for different methods.

Conclusions

In this paper, we propose a new approach for Chinese sign language recognition based on trajectory modeling. The method is inspired from the shape recognition with shape context. We partition the projected curve of sign word trajectory into multiple segments and represent each segment into histogram feature by shape context quantization. With these features, the HMMs are applied for modeling the sign words. The experiments show that our method outperforms the comparison methods by a large margin on a large dataset containing 100 sign words with over 25,000 samples.

For the future work, we will integrate both trajectory of sign word and hand shapes for more accurate SLR recognition.