Curve Matching from the View of Manifold for Sign Language Recognition

Yushun Lin[1,2], Xiujuan Chai[1], Yu Zhou[3], and Xilin Chen[1]

[1]Key Lab of Intelligent Information Processing of Chinese Academy of Sciences
(CAS), Institute of Computing Technology, CAS, Beijing 100190, China
[2]University of Chinese Academy of Sciences, Beijing 100049, China
[3]Institute of Information Engineering, CAS, Beijing 100093, China

Abstract. Sign language recognition is a challenging task due to the
complex action variations and the large vocabulary set. Generally, sign
language conveys meaning through multichannel information like tra-
jectory, hand posture and facial expression simultaneously. Obviously,
trajectories of sign words play an important role for sign language recog-
nition. Although the multichannel features are helpful for sign represen-
tation, this paper only focuses on the trajectory aspect. A method of
curve matching based on manifold analysis is proposed to recognize iso-
lated sign language word with 3D trajectory captured by Kinect. From
the view of manifold, the main structure of the curve is found by the
intrinsic linear segments, which are characterized by some geometric fea-
tures.Then the matching between curves is transformed into the match-
ing between two sets of sequential linear segments. The performance of
the proposed curve matching strategy is evaluated on two different sign
language datasets. Our method achieves a top-1 recognition rate of 78.3%
and 61.4% in a 370 daily words dataset and a large dataset containing
1000 vocabularies.

1 Introduction

Sign language is one of the most common communication means for hearing-
impaired community. With the rising public concern on deaf-mutes, sign lan-
guage recognition has become a hot research topic nowadays. At the same time,
it's also a complex and challenging problem for dealing with multichannel infor-
mation, including trajectory, hand posture and even facial expression.

There exist several common approaches for sign language recognition. The
early approaches often applied artificial neural networks (ANN). A method for
Japanese sign language recognition using recurrent neural network was proposed
by Murakami et al. (1991) [1]. They developed a system to recognize a finger
alphabet of 42 symbols, with the input of data gloves. Huang et al. (1995) [2]
presented an isolated sign recognition system using a Hopfield ANN, which can
recognize 15 different gestures accurately. After that, Kim et al. (1996) [3] trained
a Fuzzy Min Max ANN with x, y, z coordinates and angles provided by data
gloves to recognize 25 isolated gestures with a success rate of 85%. Another

popular approach is Hidden Markov Model, as sign language recognition is a kind of temporal pattern recognition, just like speech recognition. Grobel et al. (1997) [4] presented an HMM based isolated sign (gesture) recognition system, which achieved a recognition rate of 94% for 262 signs. They used cotton gloves with several colour-markings in fingers, palm and back of the hand to get both trajectory and hand shape features. Starner et al .(1998) [5] proposed an HMM based system for recognizing sentence-level continuous American Sign Language (ASL) in a lexicon of 40 words and demonstrated that HMMs presents a strong technique for recognizing sign language.

However, in this paper we tried to focus on the problem of sign language recognition with only trajectory information. Then the problem of sign language recognition can be transformed to curve matching in 3D space. In this field, many work has been done on 2D curve matching. Mokhtarian et al. (1986) [6] introduced a scale-based description at varying levels of details to recognize planer curves. An affine-invariant method was proposed by Zuliani et al. (2004) [7]. Efrat et al. (2007) [8] introduced continuous dynamic time warping measures for calculating the curves' similarity. As to 3D curve matching, Pajdla et al. (1995)[9] presented a method for this problem using semi-differential invariants, without computation of high order derivatives. Kishon et al. (1990) [10] proposed a method, which transformed the 3D curves into 1D numerical strings of rotation and translation invariant to match. Most of work above focuses on finding out the longest corresponding parts of two curves, while in sign language recognition application we concern more about the feature representation and the matching, i.e. the quantization of the distance between two curves.

To realize the curve matching, our solution is inspired by the view of manifold. Each sign curve is regarded as a manifold, then the curve matching problem is reformulated as the distance calculating between two corresponding manifolds. Maximum Linear Segment (MLS) by hierarchical divisive clustering is introduced to explore the structure information of curves. To well describe the sign, a hand-elbow co-occurrence feature is proposed, in which, the hand is dominant and the elbow is subordinative. With a novel designed distance measurement, the final matching between two sets of sequential linear segments is achieved with dynamic time warping (DTW) matching.

The remaining part of this paper is organized as follows. Section 2 gives an overview on the manifold analysis based curve matching method. Section 3 is the detail for each key module, including preprocessing, Maximum Linear Segment, the hand-elbow co-occurrence feature and the matching of sequential linear segments. Section 4 is about the experiments and the analysis. Section 5 is the conclusion.

2   Method Overview

To solve the problem of sign language recognition with only trajectory information, we propose the manifold analysis based curve matching method. The proposed method mainly consists of three parts, i.e., preprocessing, Maximum Lin-

ear segment (MLS), hand-elbow co-occurrence feature extraction and matching of sequential linear segments. Fig. 1 gives the illustration for the main flowchart of the proposed method.

The preprocessing step attempts to normalize the data and remove the noise. In consideration of various scales from different signers, the trajectory curves should be normalized by signers' size at the first step of preprocessing. Then it must be re-sampled in order to remove the effect of noise, especially those brought by inconsistent velocities.

Maximum Linear Segments are further designed to model the locally linear structure of the curves by using hierarchical divisive clustering (HDC). Then the intrinsic structure of the curves can be discovered. Some geometric features are extracted to characterize each linear segment and to well describe the sign curve, a novel hand-elbow co-occurrence feature is proposed. The final matching between sequential linear segments is obtained by DTW matching. It is worth mentioning that in the case of two-hands words, the hands are dealt separately and their DTW distances are combined with an equal weight for the final decision.
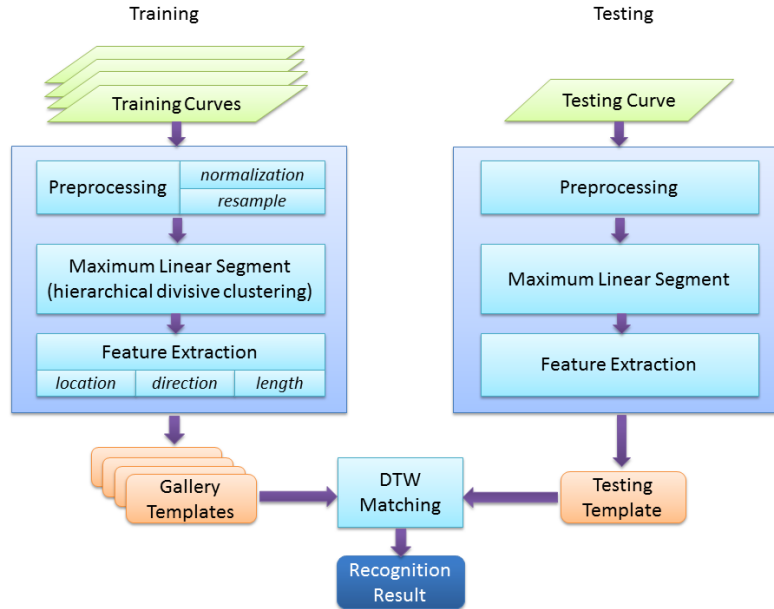

Fig. 1. Flowchart of the proposed method.

3   Manifold Analysis based Curve Matching

3.1   Preprocessing

The target of preprocessing consists of two aspects, scale normalization and data resampling. On one hand, the raw data of trajectory obtained from the device are 3D locations of joints, which should not be used directly. Since different signers have different body sizes, the scales of trajectory curves played by them may also be different. In order to avoid the effect of various scales, data of trajectory should be normalized from the very beginning. In this work, the signers are required to stand in front of the device with their hands down at the beginning of recording. The locations of both hands and the head in the first frame are selected as reference points for normalization.
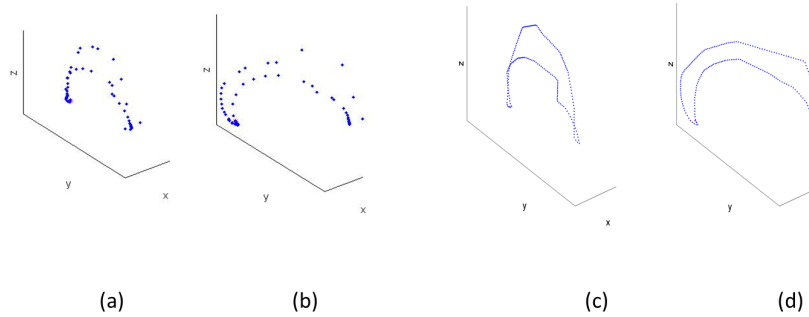

(a)   (b)   (c)   (d)

Fig. 2. Some exemplars of resampling. (a) and (b) are raw curves. (c) and (d) are re-sampled ones corresponding to (a) and (b) respectively.

On the other hand, resampling is essential for two reasons. First of all, it helps to remove the noise caused by inconsistent velocities. Because there is no explicit rules about velocity in standard Chinese Sign Language, velocity of the trajectory seems to be a random factor. It means that different signers may play a same word in different velocities, and even one signer may play the same word in different velocities in two times. Resampling operation can effectively remove the difference of velocities and keep the curve shape information well. The other reason is for the convenience of division of Maximum Linear Segment, which requires sufficient sample points. There exists many algorithms for resampling [11], and the $1 algorithm proposed by Wobbrock, J. O., et al(2007) [12] is used in our method because of its less time cost and simpleness. The main idea of this algorithm is the equidistant linear interpolation along the curve. The

interval length of interpolation is the length of the curve divided by the number of re-sampled points. Some exemplars of resampling are shown in Fig. 2 , from which, it can be seen that the curve is re-sampled evenly while maintaining the original shape.

3.2   Maximum Linear Segment

After preprocessing, normalized curves that consist of a certain number of even points are obtained. Each curve can be regarded as a simple manifold of one dimension. In mathematics, manifold is a topological space that resembles Euclidean space near each point locally. Similar to Maximum Linear Patch (MLP) introduced by [13], we introduce a concept of Maximum Linear Segment (MLS), to represent the structure information of curves. The maximum here means that the partition of segments made by MLS has maximum linearity for current step of division. In order to quantify the linear degree of curves, the definition of non-linearity degree is introduced. The non-linearity degree of a curve is defined as the geodesic distance along the curve divided by Euclidean distance between the starting point and end point. An illustration of geodesic distance and Euclidean distance between two points on a curve is given in Fig. 3. Denote $d(p)$ as the non-linearity degree of segment $S(p)$ with n points from p_1 to p_n. Then $d(p) = \sum_{i=1}^{n-1} ||p_{i+1} - p_i||/||p_n - p_1||$. Just as the name implies, the larger non-linearity degree, the less linear it is.
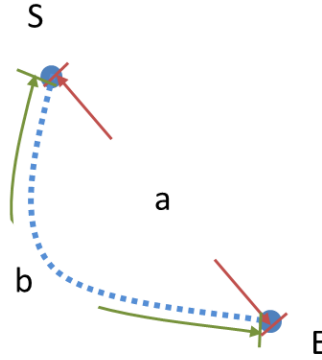

Fig. 3. The illustration of Euclidean distance a and geodesic distance b between point S and point E on the blue curve.

The main idea of MLS is to regard the whole curve as a manifold, then split it into certain segments according to the non-linearity degree. Different partitions

represent different distributions of local linear structure of curves. MLS is able to keep the local linear property of segments as much as possible.

Let's denote X as the whole curve with N points, and M is the number of segments that the curve will be divided into. Then problem can be formulated as looking for a set of segments $X^{(i)}(i = 1, 2, ...M)$ that the non-linearity degree of each division step is minimum and satisfy the following equations:

$$X = \bigcup_{i=1}^{M} X^{(i)}(i = 1, 2, ...M). \tag{1}$$

$$X^{(i)} \cap X^{(j)} = \emptyset(i \neq j, i, j = 1, 2, ...M). \tag{2}$$

$$X^{(i)} = \left\{ x_1^{(i)}, x_2^{(i)}, ..., x_{Ni}^{(i)} \right\} (\sum_{i=1}^{M} Ni = N). \tag{3}$$

To achieve the target, we use hierarchical divisive clustering (HDC) to get the segments. The algorithm for MLS by HDC is described in Algorithm 1.

---

1 Given a curve C and a fixed number M ;
2 Define the set of segments as S and the number of segments in S as n ;
3 Set S empty ;
4 Add C to S ;
5 n=1 ;
6 **while** *n is not equal to M* **do**
7   Find the segment S_m with maximum non-linearity degree in S ;
8   **for** *Each point p in S_m* **do**
9   Divide S_m from p into two segments: $S_1(p)$, $S_2(p)$ and calculate their non-linearity degree $d_1(p)$, $d_2(p)$;
10   **end**
11   Find the point p_m that minimize $max(d_1(p), d_2(p))$;
12   Add $S_1(p_m)$, $S_2(p_m)$ to S ;
13   Remove S_m from S ;
14   n=n+1 ;
15 **end**
16 Output the final set of segments S ;

Algorithm 1: MLS by hierarchical divisive clustering (HDC).

The specific value of M is a tuning parameter. In our experiment, we find that 72 is a proper value. Finally, we can get a set of Maximum Linear Segments divided from a curve, which are the elementary units for the following matching steps. An example of HDC with 4 steps is shown in Fig. 4. Fig. 5 gives the MLS result of a real sign trajectory curve.
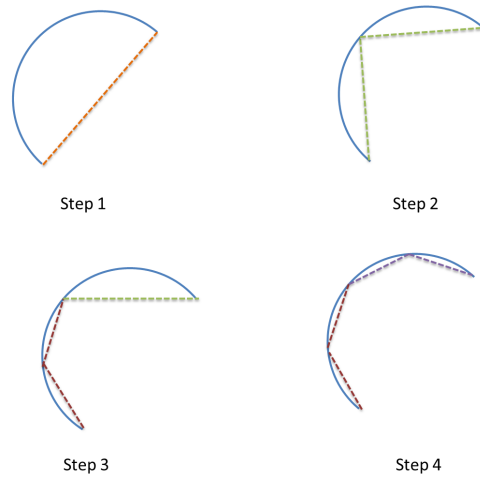
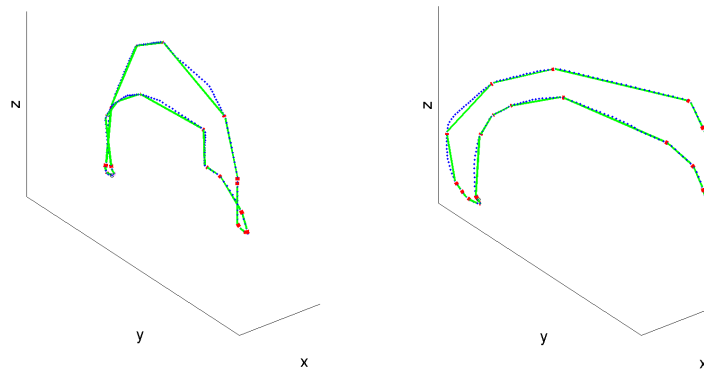**Fig. 4.** An example of a HDC with 4 steps. The segments of the curve are represented by dash lines.

Fig. 5. The MLS of a real sign trajectory. The green lines represent the Maximum Linear Segments. In this example, M is set to 20 for easy understanding.

3.3   Matching of Sequential Segments

Since we have gotten the Maximum Linear Segments for each curve, the problem of calculating the distance between two curves is transformed to the calculation of distance between two sequences of segments. In order to accomplish the target, there are two problems needed to be solved.

Measurement of Two Matched Segments.   The first problem is feature extraction to measure two matched segments. Considering that the geometric information is most important in trajectory data, the location, direction and length are extracted to characterize each segment. Then based on this kind of feature representation, the distance between two matched segments should be defined.

From the view of manifold, the segment is a linear subspace, which can be abstracted to the line segment. The location of a segment is defined as the midpoint of the line segment. Just as its name implies, the direction of a segment is defined by the difference between the end point and the starting point. The length of segment is defined as the Euclidean distance between the starting point and end point of the line segment. Denote the two matched segments as S_1, S_2, the distance between them is defined as follows. (see Equ. 4)

$$d(S_1, S_2) = (d_e + d_{len})/\left[(d_p + 1)/2\right].\qquad(4)$$

where d_e is the Euclidean distance between the locations of two segments, d_{len} is the length difference between two segments, that is $d_{len} = |len(S_1) - len(S_2)|$ where $len(S)$ represents the length of segment S. $d_p = \cos\theta$ is the cosine of the angle between two segments. (see Fig. 6) Actually, the plane in Fig. 6 is just simplified for visualization. Two segments S_1 and S_2 are the vectors in 3D space. Thus the angle between these two 3D vectors can be calculated.
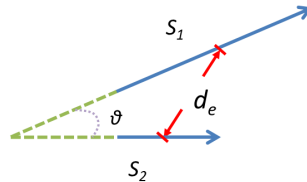

Fig. 6. Illustrations of features. S_1 and S_2 are two matched segments (blue line segments).

Measurement of Segment Sequences. The other problem is the matching between two sets of sequential linear segments. Although the amount of segments we get from each curve is set to be the same, it doesn't means that the segments should be matched simply just by their indexes. In spite that the sign curves belonging to the same vocabulary will have roughly similar structure, corresponding segments may appear with different indexes due to the random variance when playing a sign. The signer cannot repeat the same word with exactly the same trajectory.

However, it's true that the matching of segment sequences for two curves should obey the rules of sequential matching. Let's denote $X_1^{(i)}$ as the ith segment of curve X_1, and denote $X_2^{(j)}$ in a similar way. The sequential matching means that there will not exist a match between $X_1^{(i)}$ and $X_2^{(j)}$ if there already exists a match between $X_1^{(m)}$ and $X_2^{(n)}$, where $(i - m) * (j - n) < 0$. To meet the requirement of order preservation, DTW is applied in our solution. In sequential analysis, DTW is used for measuring similarity between two temporal sequences which may vary in time. It has been widely used in many applications, including handwriting recognition [14] and signature verification [15]. The most well-known one should be automatic speech recognition [16], in which DTW is used to cope with different speaking speeds. Generally speaking, DTW can be applied to the analysis for any temporal data which can be turned into a linear sequence, including the trajectory data in our problem. It can determine the matches that minimize the distance between two segment sequences under the condition of sequential matching.

Since we have measurement for two segments, the DTW distance between two segment sequences can be calculated to evaluate the similarity between two curves. The formulation of DTW distance between two segment sequences Seq_1 and Seq_2 is given in Equ. 5.

$$D_{dtw}(Seq_1, Seq_2) = \frac{1}{Card(I)} \sum_{(i,j) \in I} d(S_{1i}, S_{2j}). \tag{5}$$

where S_{1i} denotes the ith segment in Seq_1 and S_{2j} is the jth segment in Seq_2. I is the matching set in the warping path, which minimize the DTW distance between Seq_1 and Seq_2. $Card(I)$ denotes the cardinality of set I.

3.4   Hand-elbow Co-occurrence Feature

It is obvious that the hands' location features describe the majority of trajectory information in a sign word. At the same time, the elbow can also provide some discriminative information for trajectory curve recognition. The elbow movement can be regarded as a subordinative motion to hand. In order to avoid the noise effect, a novel hand-elbow co-occurrence feature is proposed.

Considering that the hand movement is dominant, the Maximum Linear Segments are obtained from hand curve and set as the reference to get the corresponding elbow segments. In other words, there is a one-to-one correspondence

between raw data points of hands and elbows. Unfortunately, this one-to-one correspondence will disappear after resampling and MLS. Our solution is mapping the hand segment points of re-sampled curves back to the nearest raw hand curves. Then we can get corresponding segment points of the raw elbow curves because of the one-to-one correlation. In the end, mapping the segment points of the raw elbow curves to the re-sampled elbow curves. The correspondence between hands and elbows' segments is built. The procedure for building the mapping relations is illustrated in Fig. 7.
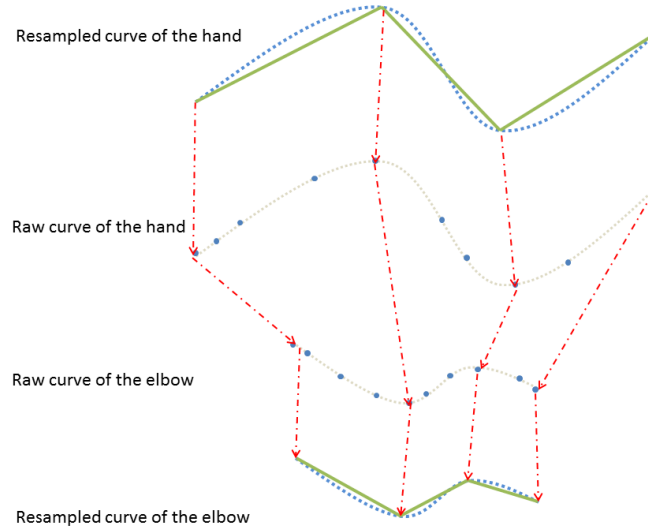

Fig. 7. An illustration for the procedure of building the hand and elbow mapping relations. The blue points represent the raw data points on raw curves while the green lines represent the segments on re-sampled curves. The red dash lines indicates the mapping relations.

4   Experiments

4.1   Datasets and Experimental Setting

To evaluate the performance of the proposed method, our experiments are conducted on two Chinese Sign Language datasets collected by ourselves.

Dataset A. Dataset A has 370 vocabularies that are widely used in daily life. Each word is played 5 times. So there are 5×370 sign videos totally. All of the data is played by one female deaf student.

Dataset B. Dataset B has 1000 different vocabularies. Each word is played 3 times, in other words there are 3×1000 words in all. All the words are played

by a male deaf student. Fig. 8 shows an example of the key frames for a word in this dataset.

Fig. 8. An example of key frames from one of the words in Dataset B.

The data is captured by Kinect, which can provide color image and depth map simultaneously. With the public windows SDK, the joint locations can be obtained in real-time. For the convenience to get abundant data, Kinect has been used in gesture recognition [17], 3D body scanning [18] and sign language recognition [19, 20]. Among all 20 joints provided, 5 joints including head, both hands, and both elbows are used in our implementation. It is obvious that the locations of both hands are essential for our problem of sign language recognition. The locations of both elbows are good supplement for hands. The location of the head is used for normalization in preprocessing.

Since we have 5 groups of data in Dataset A and 3 groups of data in Dataset B, the leave-one-out cross-validation strategy is adopted in our experiments.

4.2   Baseline Method

Hidden Markov Model (HMM) is a statistical learning method for modeling a system, that is assumed to be a Markov process with hidden states. It is a classic method in temporal pattern recognition, such as speech, gesture and handwriting recognition [21], which is similar to our problem of trajectory recognition in sign language recognition. Therefore HMM is regarded as a baseline method in this paper. On the other hand, we also compare to the curve matching method

presented by Chai, X., et al(2013) [22], which matches the location of trajectory points (denoted as PLM), for a sanity check.

To give a concrete comparison with our method, we carried out the HMM experiments in two versions with different features. One is based on the simple normalized skeleton points with head reference, which is denoted as HMM_N. And the other adopts the pairwise features as the input (denoted as HMM_P). The pairwise feature is also generated by hands, elbows and the head, which is introduced by Wang, J., et al. (2012) [23].

4.3   Experimental Results and Analysis

Location feature of elbows is considered to be a supplement to the location feature of hands. We have shown the comparison among the recognition rates of elbows' feature, hands' feature and both hands and elbows' feature on Dataset B in Fig. 9. From this figure, it can be seen that the hand-elbow co-occurrence feature can enhance the recognition rate significantly by considering the elbow feature simultaneously.
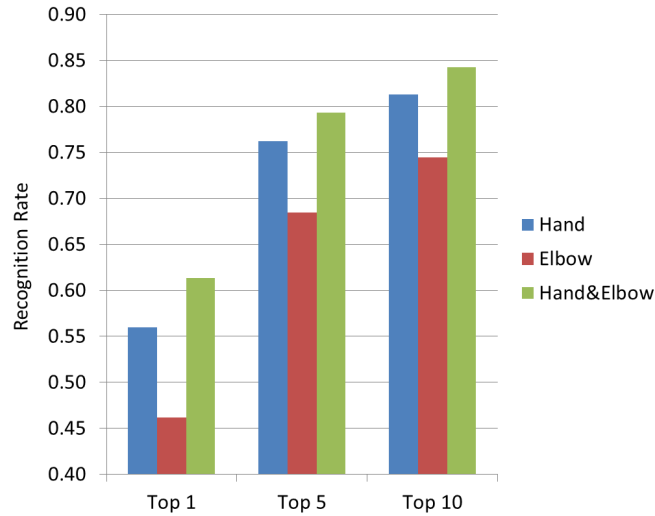

Fig. 9. The comparison among the recognition rates by using elbows' feature, hands' feature and hand-elbow co-occurrence feature on Dataset B.

The recognition rates of HMM with normalized feature (HMM_N), HMM with pairwise feature (HMM_P), method proposed in [22] (PLM) and our method on both Dataset A (see Table 1) and Dataset B (see Table 2) are shown in the

Table 1. The recognition rates for all four methods in Dataset A.

Dataset A	HMM_N	HMM_P	PLM	Our method
Top 1	0.792	0.753	0.810	0.783
Top 5	0.901	0.879	0.919	0.902
Top 10	0.928	0.914	0.946	0.929

Table 2. The recognition rates for all four methods in Dataset B.

Dataset B	HMM_N	HMM_P	PLM	Our method
Top 1	0.336	0.597	0.591	0.614
Top 5	0.563	0.813	0.770	0.793
Top 10	0.655	0.868	0.819	0.843

following tables. Fig. 10 gives the top-1 recognition rates for all three methods on the two datasets.

Judging from these results, in Dataset A, a dataset of smaller scale, our method is comparable with the other methods. In Dataset B, a more persuasive dataset with large vocabulary, our method achieves much higher recognition rate than the other methods. The HMM_N performs well on Dataset A, which is of smaller scale. But its performance decreases significantly on challenging Dataset B. However, our method of manifold analysis based curve matching performs much better on Dataset B, for the reason that it characterizes the local linear structure of curves better.

5   Conclusion

In this paper, we propose a novel 3D curve matching method for sign language recognition. The method is inspired from the view of manifold and divides the trajectory curve into Maximum Linear Segment (MLS), which describes the intrinsic structure of the curve. Based on the definition of distance between two matched segments, the similarity between two curves is also evaluated through the matching between sequential linear segments. Furthermore, a novel hand-elbow co-occurrence feature is proposed to enhance the representation ability for sign curve by considering the hand and elbow movement simultaneously. The experiments show that our method performs better than the baseline methods on a large dataset containing 1000 vocabularies. However, it can be seen that the trajectory information may be not enough for sign language recognition, especially when confronted with the large vocabularies. A directly future work is to explore the fusion feature by combining the motion trajectory and the hand shape for robust sign language recognition.
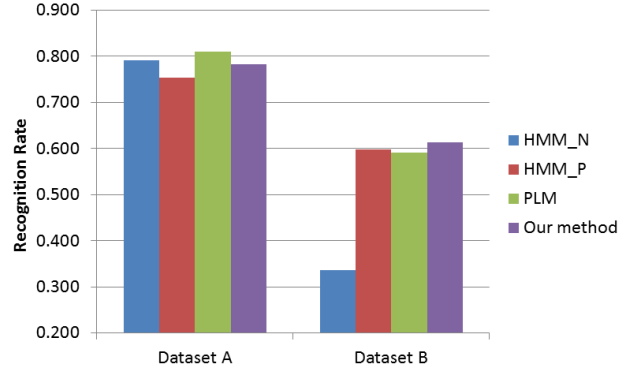
Fig. 10. The comparison of top-1 recognition rates among HMM_N, HMM_P, PLM and the proposed method.

References

1. Murakami, K., Taguchi, H.: Gesture recognition using recurrent neural networks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. CHI '91, New York, NY, USA, ACM (1991) 237–242
2. Huang, C.L., Huang, W.Y., Lien, C.C.: Sign language recognition using 3-d hopfield neural network. In: Image Processing, 1995. Proceedings., International Conference on. Volume 2. (1995) 611–614 vol.2
3. Kim, J.S., Jang, W., Bien, Z.: A dynamic gesture recognition system for the korean sign language (ksl). Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on **26** (1996) 354–359
4. Grobel, K., Assan, M.: Isolated sign language recognition using hidden markov models. In: Systems, Man, and Cybernetics, 1997. Computational Cybernetics and Simulation., 1997 IEEE International Conference on. Volume 1. (1997) 162–167 vol.1
5. Starner, T., Weaver, J., Pentland, A.: Real-time american sign language recognition using desk and wearable computer based video. Pattern Analysis and Machine Intelligence, IEEE Transactions on **20** (1998) 1371–1375
6. Mokhtarian, F., Mackworth, A.: Scale-based description and recognition of planar curves and two-dimensional shapes. Pattern Analysis and Machine Intelligence, IEEE Transactions on **PAMI-8** (1986) 34–43

7. Zuliani, M., Bhagavathy, S., Manjunath, B., Kenney, C.: Affine-invariant curve matching. In: Image Processing, 2004. ICIP '04. 2004 International Conference on. Volume 5. (2004) 3041–3044 Vol. 5

8. Efrat, A., Fan, Q., Venkatasubramanian, S.: Curve matching, time warping, and light fields: New algorithms for computing similarity between curves. J. Math. Imaging Vis. **27** (2007) 203–216

9. Pajdla, T., Gool, L.V.: Matching of 3-d curves using semi-differential invariants. In: Computer Vision, 1995. Proceedings., Fifth International Conference on. (1995) 390–395

10. Kishon, E., Hastie, T., Wolfson, H.: 3-d curve matching using splines. In Faugeras, O., ed.: Computer Vision ECCV 90. Volume 427 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (1990) 589–591

11. Shahraray, B., Anderson, D.: Uniform resampling of digitized contours. Pattern Analysis and Machine Intelligence, IEEE Transactions on **PAMI-7** (1985) 674–681

12. Wobbrock, J.O., Wilson, A.D., Li, Y.: Gestures without libraries, toolkits or training: A \$1 recognizer for user interface prototypes. In: Proceedings of the 20th Annual ACM Symposium on User Interface Software and Technology. UIST '07, New York, NY, USA, ACM (2007) 159–168

13. Wang, R., Shan, S., Chen, X., Chen, J., Gao, W.: Maximal linear embedding for dimensionality reduction. Pattern Analysis and Machine Intelligence, IEEE Transactions on **33** (2011) 1776–1792

14. Bahlmann, C., Burkhardt, H.: The writer independent online handwriting recognition system frog on hand and cluster generative statistical dynamic time warping. Pattern Analysis and Machine Intelligence, IEEE Transactions on **26** (2004) 299–310

15. Martens, R., Claesen, L.: On-line signature verification by dynamic time-warping. In: Pattern Recognition, 1996., Proceedings of the 13th International Conference on. Volume 3. (1996) 38–42 vol.3

16. Sakoe, H., Chiba, S.: Dynamic programming algorithm optimization for spoken word recognition. Acoustics, Speech and Signal Processing, IEEE Transactions on **26** (1978) 43–49

17. Ren, Z., Meng, J., Yuan, J., Zhang, Z.: Robust hand gesture recognition with kinect sensor. In: Proceedings of the 19th ACM International Conference on Multimedia. MM '11, New York, NY, USA, ACM (2011) 759–760

18. Tong, J., Zhou, J., Liu, L., Pan, Z., Yan, H.: Scanning 3d full human bodies using kinects. Visualization and Computer Graphics, IEEE Transactions on **18** (2012) 643–650

19. Zafrulla, Z., Brashear, H., Starner, T., Hamilton, H., Presti, P.: American sign language recognition with the kinect. In: Proceedings of the 13th International Conference on Multimodal Interfaces. ICMI '11, New York, NY, USA, ACM (2011) 279–286

20. Sun, C., Zhang, T., Bao, B.K., Xu, C., Mei, T.: Discriminative exemplar coding for sign language recognition with kinect. Cybernetics, IEEE Transactions on **43** (2013) 1418–1428

21. Al-Hajj Mohamad, R., Likforman-Sulem, L., Mokbel, C.: Combining slanted-frame classifiers for improved hmm-based arabic handwriting recognition. Pattern Analysis and Machine Intelligence, IEEE Transactions on **31** (2009) 1165–1177

22. Chai, X., Li, G., Lin, Y., Xu, Z., Tang, Y., Chen, X., Zhou, M.: Sign language recognition and translation with kinect. In: IEEE Conf. on AFGR. (2013)

23. Wang, J., Liu, Z., Wu, Y., Yuan, J.: Mining actionlet ensemble for action recognition with depth cameras. In: Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. (2012) 1290–1297