



Sketch-based human motion retrieval via selected 2D geometric posture descriptor



Jun Xiao^a, Zhangpeng Tang^{a,*}, Yinfu Feng^a, Zhidong Xiao^b

^a School of Computer Science, Zhejiang University, Hangzhou 310027, PR China

^b Bournemouth University, Poole Dorset BH12, United Kingdom

ARTICLE INFO

Article history:

Received 7 September 2014

Received in revised form

5 December 2014

Accepted 5 January 2015

Available online 12 January 2015

Keywords:

Motion retrieval

Sketch-based

Feature selection

Computer animation

ABSTRACT

Sketch-based human motion retrieval is a hot topic in computer animation in recent years. In this paper, we present a novel sketch-based human motion retrieval method via selected 2-dimensional (2D) Geometric Posture Descriptor (2GPD). Specially, we firstly propose a rich 2D pose feature call 2D Geometric Posture Descriptor (2GPD), which is effective in encoding the 2D posture similarity by exploiting the geometric relationships among different human body parts. Since the original 2GPD is of high dimension and redundant, a semi-supervised feature selection algorithm derived from Laplacian Score is then adopted to select the most discriminative feature component of 2GPD as feature representation, and we call it as selected 2GPD. Finally, a posture-by-posture motion retrieval algorithm is used to retrieve a motion sequence by sketching several key postures. Experimental results on CMU human motion database demonstrate the effectiveness of our proposed approach.

© 2015 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

1. Introduction

In the last few years, motion capture technique has been widely used in various applications, such as featured films, computer games and virtual reality [1–4]. However, the process of human motion capture is very time consuming and expensive [5]. So, it is important to reuse the pre-captured human motion in database. To this end, the first and important step is to retrieve similar human motion sequences from motion database.

To retrieve a desired motion sequence from a large-scale human motion database, the user usually has to describe the character of the human motion clips to be retrieved. There are several ways to achieve this goal. The simplest way is using textual description like key-words: “kicking”

and “fighting” or a more detailed sentence like “a kick of the left foot followed by a punch” [6]. The textual description is intuitive and efficient, while it requires a lot of manual work for annotating all of the motion sequences in database. To overcome this shortcoming, the content-based human motion retrieval has attracted much attention, which retrieves the motion clips via submitting a similar and short motion clip [5,7,8] as query. However, sometimes it is hard to acquire the appropriate motion clips as query. For example, if we want to retrieve a back flip and a handstand walk, it is a little difficult for the user to perform such actions. Although Numaguchi et al. [9] proposed to use a puppet interface to generate the query motion clip, such a puppet consists with several sensors and potentiometers, which is not easy to obtain for a general user.

Inspired by traditional 2D animation wherein the experienced animator sketches several key character postures, which are then used as the seeds to generate a new motion, the sketch-based human motion retrieval becomes popular in computer animation in recent years [10–13]. Different

* Corresponding author.

E-mail addresses: junx@zjuem.zju.edu.cn (J. Xiao), zhangpengtang@gmail.com (Z. Tang), fyf200502@hotmail.com (Y. Feng), zxiao@bournemouth.ac.uk (Z. Xiao).

from the other kind of input, using a 2D sketch brings in the following benefits: intuitiveness, fast posture definition, quick production of sample and first-pass animated scenes [14]. In light of this, we also utilize the sketch interface to retrieve human motion from database in this paper.

As we known, there are many factors which may affect the motion retrieval performance including motion representation, matching algorithm, indexing structure and so on. Motion representation is a the core-stone in constructing an efficient motion retrieval system. It is also the bridge to connect the 2D sketch and the 3D human motion. Traditional geometric features such as joints location and joint-joint direction are effective to character 3D human posture but not for the hand-drawn sketch. The hand-drawn sketch is usually changeable because different user has diverse sketch style and the body proportion is hard to be unified for the same posture. Thus, we has to find a proper feature to represent both 2D sketch and 3D posture.

In this paper, we propose a 2D Geometric Posture Descriptor (2GPD) including four kinds of geometric features. Similar geometric feature has been applied in graphics and computer vision [15], while here we propose a new rich posture feature set to exploit the geometric properties and relationships between different human body parts. 2GPD emphasizes the relational body part configuration, which is consistent with human perception.

Taken the flexibility of human-drawn sketch into account, the features involved distance metric are normalized with respect to the human body height and all of the features are discretized into several intervals. To ensure the retrieval efficiency, the feature dimension should not be too high. In our work, the total feature dimension of 2GPD is 738 while not all of them are useful and discriminative to describe a 2D posture. Therefore, we design a semi-supervised feature selection algorithm derived from Laplacian Score algorithm to select a compact and discriminative feature subset from the original 2GPD.

The remainder part of this paper is organized as follows. We first review some related work in Section 2. Then, a detailed description of 2DGP and our proposed semi-supervised feature selection algorithm are introduced in Sections 4 and 5, respectively. Section 5 presents the retrieval framework and algorithm. Finally, we give the experimental results and conclusions in Sections 6 and 7.

2. Related work

Sketch has been widely used in computer animation and computer graphics. Igarashi et al. [16] and Zeleznik et al. [17] proposed a sketch-based modeling system, which utilizes input strokes for mesh creation. Lee et al. [18] presented a method which can derive improved shadow image from single posture sketch. Sketch is also used for facial expressions by Lau [19] and Seol [20]. In computer animation, sketch also has been successfully used in motion synthesis, motion retrieval and posture design [21–27,14].

In order to establish a connection between a 2D sketch and a 3D posture, 3D posture reconstruction from 2D sketch is widely used in some research. Davis et al. [10] proposed a method which reconstructs several possible 3D postures from a hand-drawn key posture stick figure.

Meanwhile, a set of constraints and assumptions are applied to return the most likely 3D postures to the user. Lin et al. [23] presented an intuitive sketch interface that allows the user to creating a 3D human character in a sitting position on a chair. They reduced the reconstruction solution space by considering the interaction between the character and environment and adding physics constraints. The reconstruction problem was formulated into a nonlinear optimization problem. Li et al. [27] utilized the human figure sketches corresponding to the initial and closing posture of a Kung-fu form to retrieval and refinement the Kung-fu motions. And sketch trajectory on specific moving joints helps the retrieval process. However, reconstructing the 3D posture from a 2D sketch posture is an ill-posed problem due to many possible solutions.

A few researchers [22,21] used sketch to define the trajectory constraints which the character or its specific joints move along. But some trajectories are hard to draw in 2D sketch interface. For example, if the user sketches a posture in the view plane XOY but hand joint moves along the Z-axis.

Another kind of method is also used to compare the 2D sketch with 3D posture. First the 3D posture is projected into 2D plain, and then the same geometric feature is extracted on both the projected posture and the 2D sketch. As far as geometric feature concerned, some information like limb length and joint coordinate are removed directly because each user has different sketch style and the body proportion is hard to be unified for the same posture. The direction of the limbs, an intuitional features, is usually used [21,28] on sketch. However, only the direction of the limbs cannot character a motion sufficiently. Just as the 3D posture representation [15], the relative geometry relation of the joints and the limbs is necessary. In the field of computer vision, multi-feature [29–31] is widely used and some method such as LDA is applied to select the features or information with high discrimination [32,33]. Inspired by those work, we propose a novel feature in this paper called 2GPD combining four kinds of geometry feature, which is effective to character the hand-drawn sketch.

The 2GPD contains four different types of geometry feature, which results in high dimensionality. To reduce the dimension of feature representation, a semi-supervised feature selection method derived from Laplacian Score [34] is used in our work. Generally, feature selection approaches can be roughly classified into wrapper and filter. The wrapper approaches evaluate the features using the learning algorithm. Thus, they wrap the selection process around the learning algorithm [35–37]. The filter approaches examine intrinsic properties of the data to evaluate the features prior to the learning tasks, which always rely on the class labels, most commonly evaluating correlations between features and the class label. Classic filter approaches include Pearson correlation coefficients, data variance [38], Fisher score [39], and Kolmogorov–Smirnov test [40]. He et al. [34] proposed a filter method for unsupervised feature selection, which evaluates a feature by Laplacian Score. In this paper, a semi-supervised feature selection [41] method derived from Laplacian Score is utilized, which seeks features with high score in a carefully designed nearest neighbor graph [42].

3. 2D geometric posture descriptor

Usually, a 2D posture is obtained by projecting a 3D posture or the user's sketch, which is encoded as a collection of 2D joint coordinates. In this paper, we use a 12-joint skeleton model as shown in Fig. 1(a). The skeleton is a tree structure, where the joints are nodes.

We define a set of joints and lines on the human skeleton as shown in Fig. 1. There are 12 joints and 26 lines in total, which are explained as below.

- **Joint:** Each joint J is encoded with its coordinate (J_x, J_y) . There are 12 joints in total.
- **Line:** If one of the following three constraints is satisfied, $L_{J_a \rightarrow J_b}$ is the line from joint J_a to J_b .
 1. J_a and J_b are directly adjacent in the skeleton model. This produces 11 lines.
 2. If one of J_a and J_b is the end site (such as head, hand and foot joint), and the other one is the joint which is two steps away from that end joint (it means that one joint is an ancestor of the other and the difference of between their depths in the skeleton tree is two). For example, $L_{LHand \rightarrow Neck}$, $L_{LFoot \rightarrow Pelvis}$ and

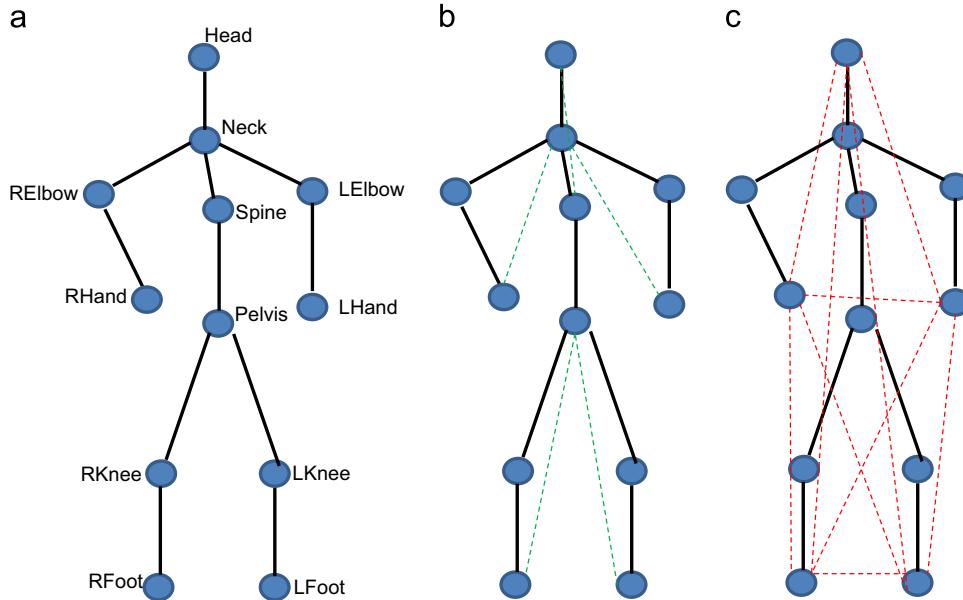


Fig. 1. (a) 2D skeleton model, blue spheres are joints and black lines are limbs. (b) Lines generated by the end site joints and the joints two steps away from end site joints. (c) Lines generated by end site joints.

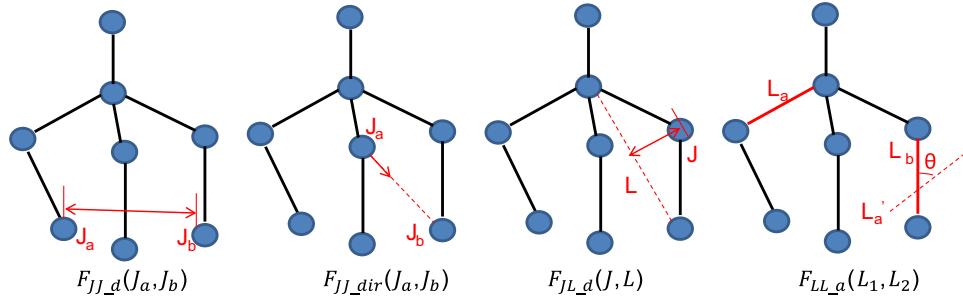


Fig. 2. Four feature types.

$L_{Head \rightarrow Spine}$ are three valid lines. This produces 5 lines, which is presented in Fig. 1(b). The reason why considering these lines is skeleton tree towards end sites are important in pose perception.

3. $L_{J_a \rightarrow J_b}$ is a line if both J_a and J_b are end sites. This produces 10 lines, which is presented in Fig. 1(c). These lines are considered because the relation between end sites is of great importance in pose identification.

J_a to J_b is the normalized vector $\overrightarrow{J_a J_b}$:

$$F_{J\|_d}(J_a J_b) = \frac{\overrightarrow{J_a J_b}}{\|\overrightarrow{J_a J_b}\|} \quad (2)$$

- **Joint–Line Distance** $F_{JL_d(J, J_a \rightarrow J_b)}$: This is the distance from joint J to line $L_{J_a \rightarrow J_b}$:

$$F_{JL_d(J, J_a \rightarrow J_b)} = \frac{2S_{\Delta J_a J_b}}{F_{J\|_d}(J_a J_b)} \quad (3)$$

where $S_{\Delta J_a J_b}$ is the area of the triangle. The calculation of area can be accelerated by employing Helen formula because we have already calculated all the edge length in the triangle as feature $F_{J\|_d}$ in Formula (1).

- **Line–Line Angle** $F_{LL_a(L_{J_a \rightarrow J_b}, L_{J_c \rightarrow J_d})}$: This is the angle from line $L_{J_a \rightarrow J_b}$ to $L_{J_c \rightarrow J_d}$, which ranges from 0 to π :

$$F_{LL_a(L_{J_a \rightarrow J_b}, L_{J_c \rightarrow J_d})} = \arccos(F_{J\|_d}(J_a J_b) \odot F_{J\|_d}(J_c J_d)) \quad (4)$$

where \odot is the dot product operator on two vectors.

As shown in Table 1, considering 12 joints, 26 lines and four kinds of 2D geometric posture descriptors, 717 kinds of feature term are measured and the total feature dimension is 783.

3.2. Feature normalization and discretization

In the process of motion retrieval, features are first extracted from the input sketch and the projected posture. Then the distance between them is computed. The feature has to describe the similarity between the users sketch and the projected posture accurately; meanwhile, it cannot be affected by the users changeable input. Each user has different sketch style, so the body proportion is hard to be unified. To avoid this problem, the distance related feature is normalized by the height of the body:

$$F_{J\|_d}(J_a J_b) = \frac{F_{J\|_d}(J_a J_b)}{H_{body}}$$

$$F_{JL_d(J, J_a \rightarrow J_b)} = \frac{F_{JL_d(J, J_a \rightarrow J_b)}}{H_{body}}$$

where H_{body} equals the distance from joint hand to joint foot.

After feature normalization, the feature is discretized. The distance feature (Joint–Joint Distance and Joint–Line Distance), which ranges from 0 to 1, is discretized into n value in the interval $[0, 1]$. For example, the interval $[0, 1]$ is divided into n parts averagely $[[0, v_1], [v_1, v_2], [v_2, v_3], \dots, [v_{n-1}, 1]]$, if a feature $f_k \in (v_i, v_j]$, then set $f_k = v_j$. In a similar way, the angle feature (Line–Line Angle) is discretized into n value in the interval $[0, 180]$. For the Joint–Joint Direction

feature, the direction vector interval is divided as in Fig. 3, every direction feature is located in one subinterval.

The normalization and discretization processes allow the user to sketch postures not so accurately, thus increasing the robustness of the retrieval approach.

4. Semi-supervised feature selection

In this paper, a semi-supervised feature selection algorithm derived from Laplacian Score [34] is utilized to reduce feature dimension. Laplacian Score is fundamentally based on Laplacian Eigenmaps [43] and Locality Preserving Projection [44]. The basic idea of Laplacian Score is to assess the features according to their locality preserving power and it is based on the observation that, two data points are probably related to the same topic if they are close to each other. In many learning problems such as classification, the local structure of the data space is very important. A nearest neighbor graph is constructed to model the local geometric structure in this method and Laplacian score seeks those features which respect this graph structure [34,45,46].

Let LS_r denotes the Laplacian Score of the r -th feature. Let f_{ri} denotes the i -th sample of the r -th feature, $i = 1, \dots, n$. The semi-supervised feature selection algorithm can be described as follows:

- Construct a nearest neighbor graph G with n nodes. The i -th node corresponds to X_i . We put an edge between nodes i and j if X_i and X_j are close, i.e. X_i is among k nearest neighbors of X_j or X_j is among k nearest neighbors of X_i . In our work, there are some labeled training data, which means we have some 2D postures and the similarity between them is known by manual observation. An edge is added between two nodes known as similar and the edge is tagged.
- If nodes i and j are connected, set

$$S_{ij} = \begin{cases} e^{-(\|x_i - x_j\|^2/t)} & \text{Edge}_{ij} \text{ is not tagged} \\ 1 & \text{Edge}_{ij} \text{ is tagged} \end{cases}$$

where t is a suitable constant. Otherwise, set $S_{ij} = 0$. The weight matrix S of the graph models the local structure of the data space.

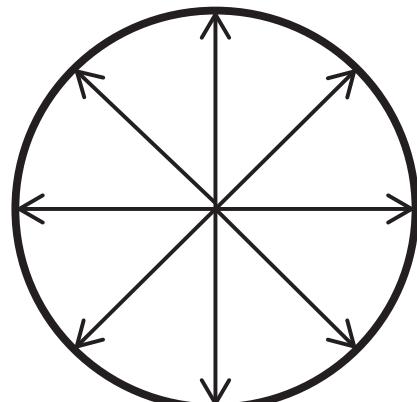


Fig. 3. The direction feature is discretized into one suninterval.

Table 1
2D pose feature count.

Type	$F_{J\ _d}$	$F_{J\ _dir}$	F_{JL_d}	F_{LL_a}	Total
Number	66	66	260	325	717
Dimension	66	132	260	325	783

- For the r -th feature, we define $f_r = [f_{r1}, f_{r2}, \dots, f_{rm}]$ $D = \text{diag}(S\mathbf{1})$, $\mathbf{1} = [1, \dots, 1]^T$ $LS = D - S$ where the matrix LS is often called graph Laplacian. Let

$$\tilde{f}_r = f_r - \frac{f_r^T D \mathbf{1}}{\mathbf{1}^T D \mathbf{1}} \mathbf{1}.$$

- Compute the Laplacian Score of the r -th feature as

follows:

$$LS_r = \frac{\tilde{f}_r^T LS \tilde{f}_r}{\tilde{f}_r^T D \tilde{f}_r}.$$

- According to the scores of all the features, the features with highest k scores are selected.

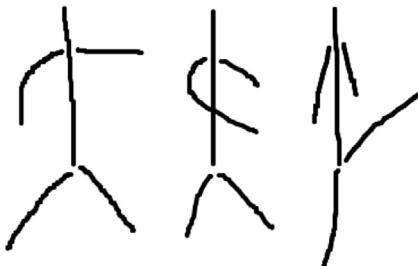


Fig. 4. Sketching example.

5. Motion retrieval

5.1. Framework

Inspired by Tang [28], we built a retrieval system which allows the user to define a motion by sketch several key postures. The posture sketch has some simple limits such as (a) one sketch posture includes five strokes; (b) the first stroke is body; (c) the view tilt and elevation angles are set to zero and a azimuth angle that makes the motion easily distinguished is selected. Once the user's sketch is obtained,

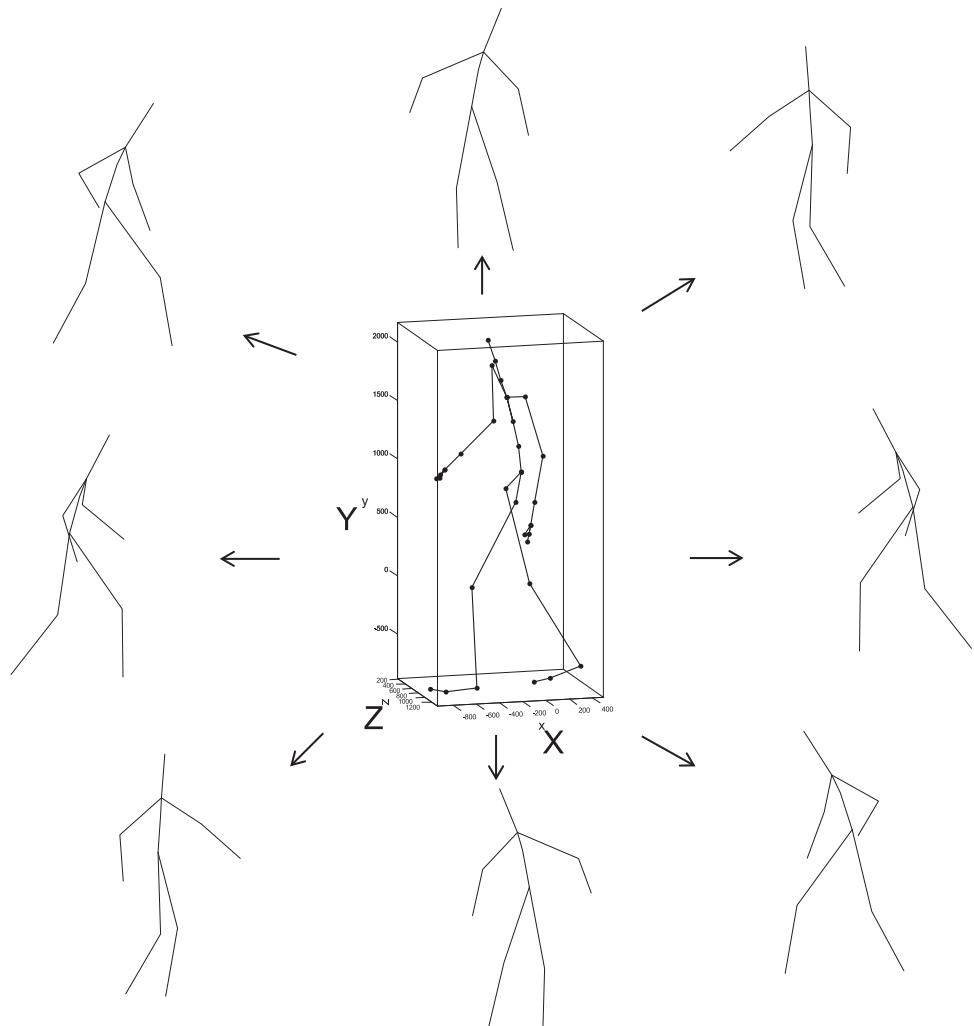


Fig. 5. Posture projection. A 3D posture is projected into 8 2D planes in average. The project plains are parallel to XOZ plain.

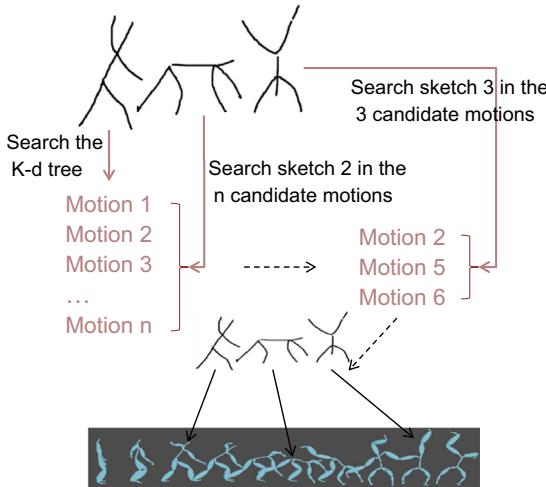


Fig. 6. The overflow of our retrieval algorithm with an example.

the joints are labeled automatically (the rules above make this process become possible). [Fig. 4](#) shows a sketch example.

In the training stage, the key postures are firstly extracted for all the motions in the database. Here the key posture extraction algorithm proposed by Tang [28] is used. Then the key postures are projected into eight 2D plains, orthographic projection is applied and eight projection directions are denoted as $\{D_k | k = 1, 2, \dots, 8\}$, where

$$D_k = \left(\sin\left(\frac{k-1}{8} \cdot 2\pi\right), 0, \cos\left(\frac{k-1}{8} \cdot 2\pi\right) \right). \quad (5)$$

[Fig. 5](#) shows an example of posture projection.

After all the projected postures are created, some similar 2D postures are labeled manually. Then the 2GPD feature is detected and the feature selection algorithm in [Section 4](#) is used to reduced the feature dimension. What is more, some index structure such as K-D-tree or R-tree can applied to index the feature space, which can speed up the k-NN retrieval process. In this work, k-d tree is used as Tang [28] introduced.

5.2. Retrieval algorithm

For a sequence of input sketch postures, what the user desires is a consecutive motion, in which the similar postures should appear in the same order as input postures. When given n sketch postures $\{S_1, S_2, \dots, S_n\}$, our posture by posture retrieval algorithm is presented in [Algorithm 1](#).

Algorithm 1. Retrieval algorithm.

Input: n sketches: $\{S_1, S_2, \dots, S_n\}$

Output The motions which include all the sketches.

- 1: **Initialize** $i=1$.
- 2: If $i > n$, the algorithm is terminated. Otherwise, search the k-d tree and return m_c candidate postures for S_i . After selection by the user, m postures $\{p_i^1, p_i^2 \dots p_i^m\}$ are left.
- 3: For each $p_i^j (1 \leq j \leq m)$, search all of the other sketches $S_k (k \neq i)$ orderly in the motion m_i^j which p_i^j lies in. If all S_k is found in m_i^j , then m_i^j is added into result sets.
- 4: If none of $m_i^j (1 \leq j \leq m)$ is chosen in step 3, **Set** $i=i+1$, **Goto** Step 2. Otherwise, the algorithm is terminated.

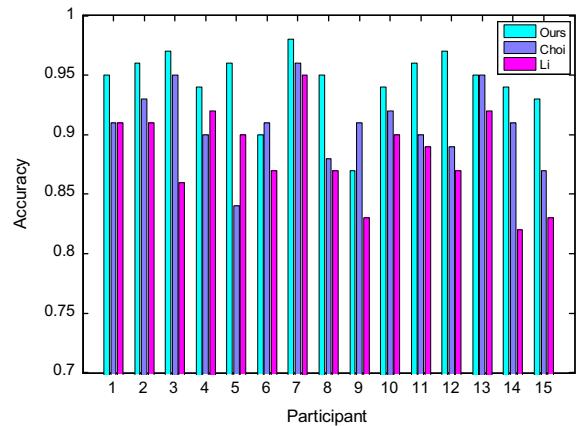


Fig. 7. Single posture retrieval result. The average accuracy of three methods are 94.47%, 90.87%, 88.33%.

[Fig. 6](#) shows an example of the retrieval algorithm.

6. Experimental result

Our experiments run on a desktop computer with a 3.10 GHz Intel-Core-i5 cpu, 4 GB RAM and Windows7 operating system. To demonstrate the efficiency of our approach, we use a subset of CMU motion database [47,48] as experimental data, which contains 1272 motion data files. Since the pure sketch-based motion retrieval method is rare, our results are compared with Choi's [21] method and Li's method [27]. A similar sketch like ours was used in Choi's method and a complex human figure sketch with trajectory constraints was used in Li's method.

The result of sketch based motion retrieval cannot be evaluate objectively. It is a good choice that the 3D motions returned by retrieval system are evaluated by the users themselves. To this end, a user study is conducted to show the effectiveness of our method. Fifteen graduate student participants are randomly chosen in our experiments.

First the effectiveness of single posture retrieval is evaluated. The participants sketching a posture and four similar candidate 3D posture results are given. If one of the results is accepted by the participant, we treat it as a successful retrieval. Each participant draw 20 different sketches in our experiment. The retrieval accuracy is shown in [Fig. 7](#). [Fig. 8](#) presents some single posture retrieval examples.

Next, we evaluate the accuracy of motion sequence retrieval. First, the participants are required to browse the whole database motions. Then the participants are required to retrieve ten motions. For each motion, the participants retrieve them by sketching several key postures and our system will return several candidate motion results. It is regarded as a successful retrieval if one of the candidate motions meets the participant's requirement. Since the trajectory constraints is utilized in Li's method [27] but not in ours, we evaluate their method separately: with trajectory and with no trajectory. The experiment earlier is repeated for ten times. As shown in [Fig. 9](#), the

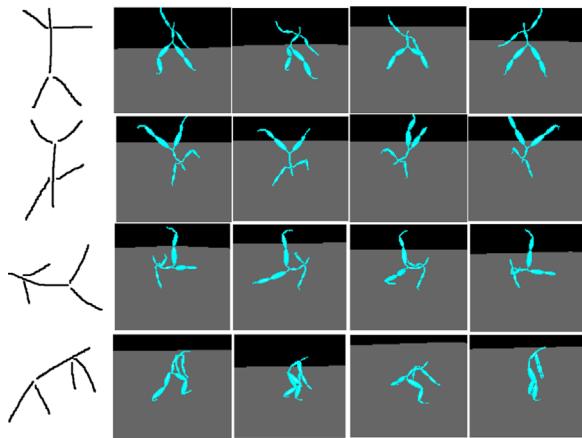


Fig. 8. Four examples of single posture retrieval in our user study.

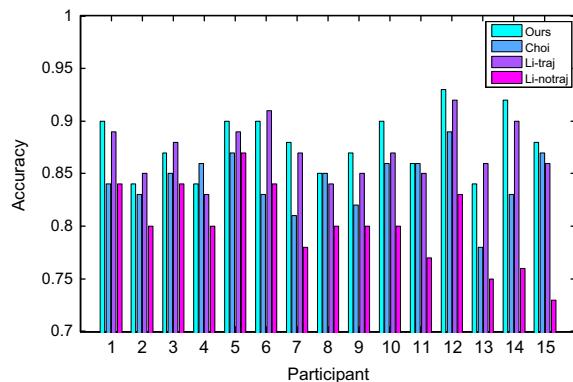


Fig. 9. Motion sequence retrieval result. The average accuracy of the methods are 87.87%, 84.33%, 87.13%, 80.07%.

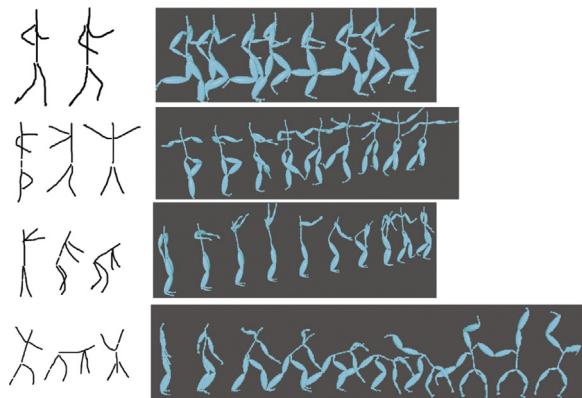


Fig. 10. Four examples of motion sequence retrieval in our user study.

accuracy of Li's method with trajectory is close to ours. But the retrieval time is doubled when the trajectory constraints is added in their method. Fig. 10 presents some motion sequence retrieval examples.

The selected feature dimension in the experiment earlier is set to 81, Fig. 11 shows how does the selected feature dimension affects the result accuracy.

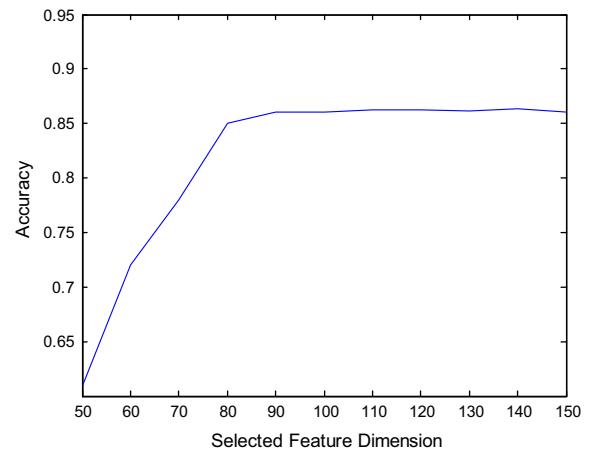


Fig. 11. Results with different feature dimension. When the feature dimension reach about 80, the accuracy grow barely.

7. Conclusion

In this paper, we present a 2D posture feature called 2D geometric pose descriptor, which describes a 2D sketch posture or projected posture by combine four kinds of geometric relation feature. The normalization and discretization to the feature make our method more robust to the different style of sketches. The semi-supervised feature selection method is designed to select a more compact and effective feature subset as feature representation, which makes the retrieval performance be more efficient.

In the future, we would like to study how to combine multiple different features like our 2GPD and spatial-temporal feature as feature representation, which would boost the algorithm's performance. Besides, our posture by posture retrieval algorithm can be extended to clip based retrieval algorithm.

Acknowledgements

This research is supported by the National High Technology Research and Development Program (2012AA011502), the National Key Technology R&D Program (2013BAH59F00), the Zhejiang Provincial Natural Science Foundation of China (LY13F020001), the Zhejiang Province Public Technology Applied Research Projects (No. 2014C33090) and partially supported by the grant of the "Sino-UK Higher Education Research Partnership for Ph.D. Students" Project funded by the Department of Business, Innovation and Skills of the British Government and Ministry of Education of P.R. China.

Appendix A. Supplementary data

Supplementary data associated with this paper can be found in the online version at <http://dx.doi.org/10.1016/j.sigpro.2015.01.004>.

References

- [1] C. Böhm, S. Berchtold, D. Keim, Searching in high-dimensional spaces: index structures for improving the performance of multimedia databases, *ACM Comput. Surv.* 33 (3) (2001) 322–373.
- [2] T. Huang, H. Liu, G. Ding, Motion retrieval based on kinetic features in large motion database, in: Proceedings of the 14th ACM International Conference on Multimodal Interaction, ACM, Santa Monica California, USA, 2012, pp. 209–216.
- [3] Y. Feng, J. Xiao, Y. Zhuang, X. Yang, J.J. Zhang, R. Song, Exploiting temporal stability and low-rank structure for motion capture data refinement, *Inf. Sci.* 277 (2014) 777–793.
- [4] J. Xiao, Y. Feng, M. Ji, X. Yang, J.J. Zhang, Y. Zhuang, Sparse motion bases selection for human motion denoising, *Signal Process.*, <http://dx.doi.org/10.1016/j.sigpro.2014.08.017>.
- [5] M. Müller, T. Röder, M. Clausen, Efficient content-based retrieval of motion capture data, in: *ACM Transactions on Graphics (TOG)*, vol. 24, ACM, Los Angeles, USA, 2005, pp. 677–685.
- [6] A. Yoshitaka, T. Ichikawa, A survey on content-based retrieval for multimedia databases, *IEEE Trans. Knowl. Data Eng.* 11 (1) (1999) 81–93.
- [7] C. Sun, I. Junejo, H. Foroosh, Motion retrieval using low-rank subspace decomposition of motion volume, in: Computer Graphics Forum, vol. 30, Wiley Online Library, Asheville, USA, 2011, pp. 1953–1962.
- [8] T. Qi, Y.-F. Feng, J. Xiao, Y.-T. Zhuang, X.-S. Yang, J.-J. Zhang, A semantic feature for human motion retrieval, *Comput. Anim. Virtual Worlds 24 (3–4) (2013) 399–407*.
- [9] N. Numaguchi, A. Nakazawa, T. Shiratori, J.-K. Hodgins, A puppet interface for retrieval of motion capture data, in: Proceedings of the 2011 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, ACM, Vancouver, Canada, 2011, pp. 157–166.
- [10] J. Davis, M. Agrawala, E. Chuang, Z. Popović, D. Salesin, A sketching interface for articulated figure animation, in: Proceedings of the 2003 ACM SIGGRAPH/Eurographics Symposium on Computer Animation, Eurographics Association, 2003, pp. 320–328.
- [11] Y. Lin, 3D character animation synthesis from 2D sketches, in: Proceedings of the 4th International Conference on Computer Graphics and Interactive Techniques in Australasia and Southeast Asia, ACM, Kuala Lumpur, Malaysia, 93–96, pp. 2006.
- [12] C. Mao, S. Qin, D. Wright, A sketch-based gesture interface for rough 3D stick figure animation, in: Eurographics Workshop on Sketch-Based Interfaces and Modeling, The Eurographics Association, 2005, pp. 175–183.
- [13] X.K. Wei, J. Chai, Intuitive interactive human-character posing with millions of example poses, *IEEE Comput. Graph. Appl.* 31 (4) (2011) 78–88.
- [14] I. Yoo, J. Vanek, M. Nizovtseva, N. Adamo-Villani, B. Benes, Sketching human character animations by composing sequences from large motion database, *Vis. Comput.* 30 (2) (2014) 213–227.
- [15] C. Chen, Y. Zhuang, F. Nie, Y. Yang, F. Wu, J. Xiao, Learning a 3D human pose distance metric from geometric pose descriptor, *IEEE Trans. Vis. Comput. Graph.* 17 (11) (2011) 1676–1689.
- [16] T. Igarashi, S. Matsuo, H. Tanaka, Teddy: a sketching interface for 3D freeform design, in: ACM Siggraph 2007 Courses, vol. 21, ACM, San Diego, USA, 2007.
- [17] R.C. Zeleznik, K.P. Herndon, J.F. Hughes, SKETCH: an interface for sketching 3D scenes, in: ACM SIGGRAPH 2007 Courses, ACM, San Diego, USA, 2007, p. 19.
- [18] V.J. Lee, C.L. Zitnick, M.F. Cohen, Shadowdraw: real-time user guidance for freehand drawing, in: *ACM Transactions on Graphics (TOG)*, vol. 30, ACM, Vancouver, Canada, 2011, p. 27.
- [19] M. Lau, J. Chai, Y.-Q. Xu, H.-Y. Shum, Face poser: interactive modeling of 3d facial expressions using facial priors, *ACM Trans. Graph.* 29 (1) (2009) 3.
- [20] Y. Seol, J. Seo, P.H. Kim, J. Lewis, J. Noh, Artist friendly facial animation retargeting, *ACM Trans. Graph.* 30 (6) (2011) 162.
- [21] M.-G. Choi, K. Yang, T. Igarashi, J. Mitani, J. Lee, Retrieval and visualization of human motion data via stick figures, in: Computer Graphics Forum, vol. 31, Wiley Online Library, 2012, pp. 2057–2065.
- [22] M.-W. Chao, C.-H. Lin, J. Assa, T.-Y. Lee, Human motion retrieval from hand-drawn sketch, *IEEE Trans. Vis. Comput. Graph.* 18 (5) (2012) 729–740.
- [23] J. Lin, T. Igarashi, J. Mitani, M. Liao, Y. He, A sketching interface for sitting pose design in the virtual environment, *IEEE Trans. Vis. Comput. Graph.* 18 (11) (2012) 1979–1991.
- [24] J. Pan, J.J. Zhang, Sketch-based skeleton-driven 2D animation and motion capture, in: *Transactions on Edutainment VI*, Springer, Chengdu, China, 2011, pp. 164–181.
- [25] M. Thorne, D. Burke, M. van de Panne, Motion doodles: an interface for sketching character motion, in: *ACM Transactions on Graphics (TOG)*, vol. 23, ACM, Los Angeles, USA, 2004, pp. 424–431.
- [26] E. Jain, Y. Sheikh, M. Mahler, J. Hodgins, Three-dimensional proxies for hand-drawn characters, *ACM Trans. Graph.* 31 (1) (2012) 8.
- [27] Q.-L. Li, W.-D. Geng, T. Yu, X.-J. Shen, N. Lau, G. Yu, MotionMaster: authoring and choreographing Kung-fu motions by sketch drawings, in: Proceedings of the 2006 ACM SIGGRAPH/Eurographics symposium on Computer animation, Eurographics Association, 2006, pp. 233–241.
- [28] Z. Tang, J. Xiao, Y. Feng, X. Yang, J. Zhang, Human motion retrieval based on freehand sketch, *Comput. Anim. Virtual Worlds 25 (3–4) (2014) 273–281*.
- [29] L. Zhang, Y. Han, Y. Yang, M. Song, S. Yan, Q. Tian, Discovering discriminative graphlets for aerial image categories recognition, *IEEE Trans. Image Process.: Publication of the IEEE Signal Processing Society* 22 (12) (2013) 5071–5084.
- [30] L. Zhang, Y. Gao, C. Hong, Y. Feng, J. Zhu, D. Cai, Feature correlation hypergraph: exploiting high-order potentials for multimodal recognition, *IEEE Trans. Cybern.* 44 (8) (2013) 1408–1419.
- [31] Y. Feng, J. Xiao, Y. Zhuang, X. Liu, Adaptive unsupervised multi-view feature selection for visual concept recognition, in: *The 11th Asian conference on Computer Vision (ACCV)*, Lecture Notes in Computer Science, Springer, Daejeon, Korea, 2013, pp. 343–357.
- [32] L. Zhang, Y. Gao, Y. Xia, Q. Dai, X. Li, A fine-grained image categorization system by cellet-encoded spatial pyramid modeling, *IEEE Trans. Ind. Electron.* PP (99) (2014) 1.
- [33] L. Zhang, M. Song, X. Liu, L. Sun, C. Chen, J. Bu, Recognizing architecture styles by hierarchical sparse coding of blocklets, *Inf. Sci.* 254 (2014) 141–154.
- [34] X. He, D. Cai, P. Niyogi, Laplacian score for feature selection, in: *Advances in Neural Information Processing Systems*, 2005, pp. 507–514.
- [35] R. Kohavi, G.H. John, Wrappers for feature subset selection, *Artif. Intell.* 97 (1) (1997) 273–324.
- [36] Y. Han, Y. Yang, X. Zhou, Co-regularized ensemble for feature selection, in: *Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence*, AAAI Press, Beijing, China, 2013, pp. 1380–1386.
- [37] Y. Han, J. Zhang, Z. Xu, S.-I. Yu, Discriminative multi-task feature selection, in: *AAAI (Late-Breaking Developments)*, 2013.
- [38] H. Liu, H. Motoda, *Feature Selection for Knowledge Discovery and Data Mining*, Springer, 1998.
- [39] Q. Gu, Z. Li, J. Han, Generalized Fisher Score for Feature Selection, *arXiv:<http://arXiv.org/abs/arXiv:1202.3725>*.
- [40] H.W. Lilliefors, On the Kolmogorov-Smirnov test for normality with mean and variance unknown, *J. Am. Stat. Assoc.* 62 (318) (1967) 399–402.
- [41] Y. Han, Y. Yang, Y. Yan, Z. Ma, N. Sebe, X. Zhou, Semisupervised feature selection via spline regression for video semantic recognition, *IEEE Trans. Neural Netw. Learn. Syst.* PP (99) (2014) 1.
- [42] L. Zhang, M. Song, X. Liu, J. Bu, C. Chen, Fast multi-view segment graph kernel for object classification, *Signal Process.* 93 (6) (2013) 1597–1607.
- [43] M. Belkin, P. Niyogi, Laplacian eigenmaps and spectral techniques for embedding and clustering, in: *NIPS*, vol. 14, 2001, pp. 585–591.
- [44] X. Niyogi, Locality preserving projections, in: *Neural information processing systems*, vol. 16, 2004, p. 153.
- [45] L. Zhang, Y. Yang, Y. Gao, Y. Yu, C. Wang, X. Li, A probabilistic associative model for segmenting weakly-supervised images, *IEEE Trans. Image Process.* 23 (9) (2014) 4150–4159.
- [46] L. Zhang, Y. Gao, Z.R., Q. Tian, X. Li, Fusion of multichannel local and global structural cues for photo aesthetics evaluation, *IEEE Trans. Image Process.* 23(3) (2014) 1419–1429.
- [47] Carnegie-Mellon MoCap Database, (<Http://mocap.cs.cmu.edu/>).
- [48] J. Xiao, Y. Feng, W. Hu, Predicting missing markers in human motion capture using l1-sparse representation, *Comput. Anim. Virtual Worlds 22 (2–3) (2011) 221–228*.