Eurographics/ ACM SIGGRAPH Symposium on Computer Animation (2006)
M.-P. Cani, J. O'Brien (Editors)

MotionMaster: Authoring and Choreographing Kung-fu Motions by Sketch Drawings

Q. L. Li [1], W. D. Geng[1], T. Yu[1], X. J. Shen[1], N. Lau[2] and G. Yu[2]

[1]State Key Lab of CAD & CG, College of Computer Science, Zhejiang University, PR China
[2]Multimedia Innovation Center, School of Design, The Hong Kong Polytechnic University, HK SAR

Abstract

Sketch-drawings is an intuitive and comprehensive means of conveying movement ideas in character animation. We proposed a novel sketch-based approach to assisting the authoring and choreographing of Kungfu motions at the early stage of animation creation. Given two human figure sketches corresponding to the initial and closing posture of a Kungfu form, and the trajectory drawings on specific moving joints, MotionMaster can directly rapid-prototype the realistic 3D motion sequence by sketch-based motion retrieval and refinement based on a motion database. The animators can then preview and evaluate the recovered motion sequence from any viewing angles. After the 3D motion sequence has been associated with the 2D sketch drawing, the animator can also interactively and iteratively make changes on the 2D sketch drawing, and the system will automatically transfer the 2D changes to the 3D motion data of current interests. It greatly helps the animator focus on the movement idea development during the evolutionary process of building motion data for articulated characters.

Categories and Subject Descriptors (according to ACM CCS): I.3.7 [Computer Graphics]: Three-Dimensional Graphics and Realism:Animation

1. Introduction

The process of creating an animation generally requires many steps going from abstract movement idea conceptualization to completion of concrete animation sequence. While there has been much work in tools for specifying the motion of animated character, less work has gone into the important preliminary stages such as movement design and planning. From the point of view of motion choreography, authoring motions is a compositional design process that involves creating, structuring, and forming body movements. Building computer-based motion authoring tools in terms of a choreographer's design skills and conventions requires an understanding of mental model of choreographer's design process, and how our creative process operates when we interact with computer systems. Authoring and visualizing body movements in an immediate and responsive way will help provide a more intuitive, direct and transparent relationship with the conception and development of movement ideas [CBM*93] [SCLG90].

When planning a character animation, it is useful to start with a simple 2D drawing of the 3D character. These sketching thumbnails, through its images and annotations, describes how various joints move in relation to one another in the overall figures. The earlier work on computer-based motion choreography can be traced back to the end of 1970. Brown and Smoliar developed a graphical tool to assist the editing and creation of Labanotation sequence for human choreographer [BS76]. In light of the iterative and interactive nature of the choreographic process, Calvert et al investigated how to support the choreographer in compositional process, and developed a visual idea generator with intuitive interface of Stage-, Body-, and Stance- Editors [SCLG90] [CBM*93]. Most of these existing efforts focus on movement notating, and intuitive interfaces to devise and create a notation-based choreography. They often require the end user to newly learn a notating system for character movement choreographing.

However, sketch-drawings are a natural choice for the creation of graphical media such as character movement, as the artists can quickly and easily sketch 2D figures, and traditional animators often begin work by quickly sketching thumbnails of a character in key poses to capture the charac-

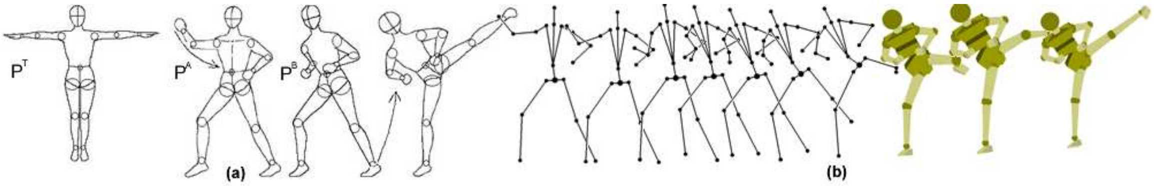

Figure 1: *(a) The input 2D posture sketches and motion lines of joint trajectory. (b)The resulting 3D motion sequence matched to 2D sketches*

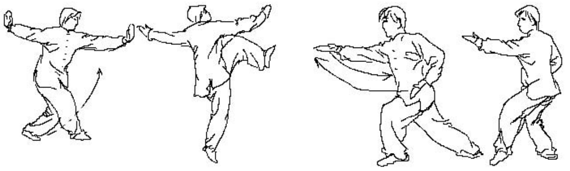

Figure 2: *Sketch drawing examples of Chinese Kungfu forms.*

ter's overall motion [DAC*03]. Our goal here is to provide design support such as movement exploring, reconciling, juxtaposing and synthesizing during the process of creating character motions. We choose Chinese Kungfu, a popular theme in feature film and video game industry, as motion domain. Chinese Kungfu motions are natural and aesthetic movements of attack and defense, such as the punch, strike, hook, block, chop and various kicks, and these martial art techniques have been handed down by drawings, or pictures with text explanation for hundreds of years. As shown in Figure 2, a typical martial art form is made of an initial posture, intermediate movements (attack and/or defense) and a closing posture. The initial and closing postures are often represented as hand-drawn human figures. The trajectories of hit joints within a Kung-fu form are often indicated by cursive arrowed lines.

In order to assist the conception and development of Kungfu forms, a sketch-based motion authoring and choreographing system, MotionMaster, is proposed. Its input is solely based on sketch-drawings. The principle behind it is that MotionMaster should be able to rough out realistic character motions at the early stage of motion design, as the rapid-prototyped 3D movements helps the motion composer not only to evaluate the choreography work, but also to follow and develop the intent of movement idea in an animation sequence. Such a Kung-fu motion authoring problem can be formalized as follows: P^T is T-pose skeleton of a specific character, P^A and P^B are the initial and closing postures of a Kungfu forms respectively(see Figure 1), S is a set of cursive lines that indicate the trajectory of joints during the intermediate movements in current Kung-fu form. Based on the reference posture P^T, the user interactively specifies the joint correspondence relationship of P^A and P^B accordingly. The system will automatically infer the desired 3D motion sequence that matches the input P^A, P^B and S.

The remainder of this paper is organized as follows. We first review the related work in Section 2, and explain the pipeline and scheme of sketch-based motion authoring and choreographing in Section 3. Then we describe how to generate the best-fit motion segments from a motion database in Section 4. Section 5 presents how to iteratively optimize and refine the motion segments in question. We show the experimental results and conclude with a discussion of our approach in Section 6 and 7.

2. Related work

The understanding of Kung-fu motions via sketch drawings involves the recognition of initial and closing postures, and the trajectory recovery of moving joints in the intermediate movements within the Kungfu form. The relevant work in computer vision and animation community can be described by three main categories as follows.

2.1. 3D posture recovery from single view

3D posture recovery from a single view is ill-defined due to insufficient spatial information. Its primary challenge is that many 3D poses may be consistent with a given 2D stick figure. In computer vision community, the early work on the problem of reconstructing 3D poses from a monocular video sequence can be traced back to 1980s [LC85]. Several recent surveys on these techniques have been explored in [Gav99] [MG01]. In general, there are two main approaches, namely, model-based and learning-based, to the 3D pose reconstruction problem. The model-based approach assumes that a 3D skeleton is known a priori, and uses an articulated human model to generate possible 3D postures that match the 2D human figure. In order to obtain the best 3D posture, a set of physical, environmental, or dynamic constraints is then applied to cull invalid 3D postures generated initially [Tay00] [DCR01]. In learning-based approaches, they automatically learn a mapping between features in the 2D image and that in the 3D posture through stochastic learning processes [RSR01]. It requires a large set of training data to learn prior correspondence knowledge of posture features in 2D image and specific 3D motion.

By taking the guiding data set in the learning-based approach and a priori knowledge of human model and constraints in the model-based approach, Chiu et al exploits a posture library and constraints to guide the reconstruction work. This hybrid approach will retrieve an appropriate candidate from the library in terms of the similarity of the 3D posture's 2D projection and the human figure in the image [CWW*04]. The 3D pose recovery techniques used here is in spirit similar to the one used in our system. However, our system goes significantly beyond theirs. we rapidly reconstruct the entire 3D motion sequence expressed in the 2D sketches, in stead of the 3D postures.

2.2. Sketch-based interface for animation

A useful and common approach to interfacing 2D-3D animations is to provide an intuitive sketch-based interface, as hand drawn animation sketches provides a flexible and simple means of "roughing out" expressive animated characters. In the sketch-based motion generation, Sabiston attempted to pose stick skeleton figure from a series of 2D thumbnail sketches using a simple "flipbook" analogy. Assuming the animator has a reasonable knowledge of perspective and foreshortening, the system is able to determine the joint angles for the 3D figure by calculating the amount of foreshortening on hand-drawn limbs [SAB91].

Davis et al provides a simple sketching interface for articulated figure animation. The user draws the skeleton on top of the 2D sketch, and then the system constructs the set of poses that exactly match the drawing, It also allows the user to guide the system to the desired character pose [DAC*03]. Thorne et al focused on the high-level motions, and developed cursive motion notations that can be used to draw motions. A desired motion is created for the character by sketching a gesture such as continuous sequence of lines, arcs, and loops [TBP04]. Oshita presents a pen-based intuitive interface to generate the human motion interactively. He employs the pen as a metaphor of sketch, and directly maps pen movements to the figure motions [Osh04].

Regarding the sketch-based character modelling, Igarashi et al present a sketching interface, Teddy, for quickly and easily designing freeform models such as stuffed animals and other rotund objects. The user draws several 2D freeform strokes interactively on the screen and the system automatically constructs plausible 3D polygonal surfaces [IMT99]. Rademacher proposed the view dependent geometry that the specified model can change with the view point. The inputs to the system are a 3D model of a character (the base model) and a set of drawings of the character from various viewpoints. The user then manually matches the viewing direction and the shape of the base model with each drawing thus creating a key viewpoint and key deformation pair for every drawing [RAD99]. Chaudhuri et al further proposed a technique of injecting view-dependent stylization into the animation. They allow the animator to make the character respond automatically to changes in the view direction. This sketch-based input mechanism interfaces the rich stylization of traditional 2D animation with 3D character animation [CKB04].

2.3. Motion editing and synthesis

The task of specifying the motion of animated object to the computer is surprisingly difficult. The animator must be able to specify subtle details of motion to convey the personality of a character or the mood of an animation in compelling fashion. All available tools require a tradeoff between automation and control. Two types of semiautomatic techniques have been developed for creating human motions: model-based and data-driven [YKH04]. Model-based approaches use simulation, search, and optimization to generate character motion by restricting the space of possible motions via kinematic, dynamic, or biomechanical models. This approach provides a flexible and compact representation of motion, but can sometimes be difficult to construct and control, or may fail to generate natural-looking motions if the models do not sufficiently constrain the motion. Data-driven approaches utilize captured motion to create animations that contain the subtle movement details recorded from a human actor. These do not provide a perfect solution either. However, intensive use of motion capture data has become a common strategy for producing plausible movements. A recent survey was given in [GY03].

Regarding the synthesis of Kungfu motions, the most valuable work was recently presented in [CCYL04]. They describe a novel framework for synthesizing new motions based on given motion manuals and corresponding capture examples. At first, a set of basic motion text is extracted from exercise manuals using text analysis method. Then an annotation method to build the mapping table between basic motion texts and their corresponding motion clips in the given examples. The new motion with given textual description can be synthesized by converting the description sentences into a sequence of text manual forms and then concatenating the corresponding motion clips to form the desired motion.

3. Overview of motion authoring and choreographing in MotionMaster

In character animation, thumbnail sketching is the initial and improvisational stage in authoring and choreographing motions. The professional computer animators often draw key poses on paper before building them in the computer. The characters are often hand-drawn as simple ellipsoidal volumes, or as mannequin figure sequences which will from a coarse animated motion. It then provides a starting point for the animators to iteratively optimize and refine the rough motions in their minds, and the desired 3D motions are finally created with this coarse-to-fine strategy. The rapid-prototyping of motion during motion planning stage will not only relieve the animator of tedious and detailed work for evaluation, but more importantly, it will play the role of a visual idea generator, which can greatly assist the animator to develop the movement idea intuitively.

The true evaluation of motion-be-designed is to view the physical realization of the final motion sequence, therefore the rapid-prototyped motion should be as realistic as possible. From the point of view of computer aided conceptual design, a certain degree of transparency in automating the rapid-prototype of motion is also required, as it will let the animator focus more on the movement idea, minimizing the interruption to the animator's mental thinking process. In addition, the alternate view of 3D motion is also important in the juxtaposition of different frames of reference movements. The studying on the same object, idea and entity from different points of view allows the choreographer to think in various (perhaps unconventional) ways about the motion of current interests [SCLG90].

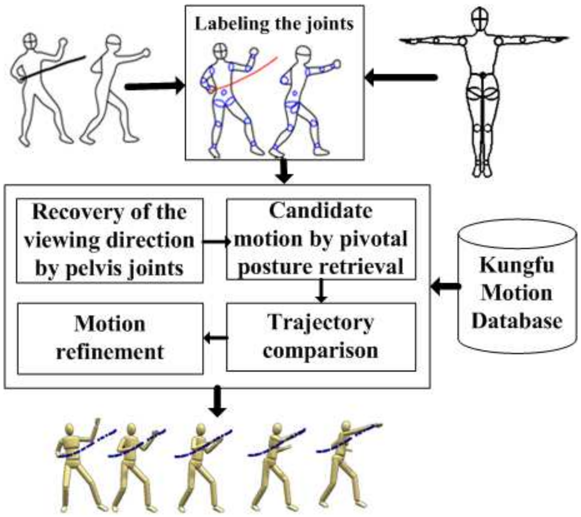

Figure 3: *The framework and pipeline of MotionMaster.*

A framework and pipeline of sketch-based motion authoring and choreographing is proposed in terms of the aforementioned principles. As shown in Figure 3, there are four major components of MotionMaster: recovery of viewing direction, generating candidate motion by posture retrieval, distilling the best-fit motion segments by trajectory comparison, and refining and optimizing the motion in question by 2D sketches. The major algorithmic steps of the motion authoring and choreographing pipeline are given below:

1. Character specification. The user first draws a 2D T-pose mannequin figure with joint points from scratch, or loads one T-pose sketch figure drawn before. The T-pose sketch drawing can also be generated from the 3D skeleton in the motion library. The segment/bone ratios and joint configuration in the T-pose figure are employed when drawing the other sketch figures afterwards.
2. Input of sketches. The user inputs the 2D sketch figures of the initial and closing pose of Kung-fu form respectively. The trajectory of moving joints are marked by cursive motion lines. Once the user confirms that the current sketch drawings are accomplished, the system will ask the user to explicitly mark the 2D position of all the joints on the input sketch drawings in terms of user-defined T-pose character skeleton.
3. Recovery of root orientation relative to viewing direction. Assuming that the viewing direction is fixed along the Z axis, the system reconstructs the root orientation by the 2D coordinate of left-hip, right hip and pivotal joints using scaled orthographic projection and segments' lengths/ratios in the 2D T-pose mannequin figure.
4. Locating motion clips by pivotal posture retrieval. For each frame in the motion capture database, the system will re-project it into the 2D image plane in terms of the recovered root orientations, and then compare it with the sketch figures of initial pose. If they are similar, the current motion clips will be picked out as candidate 3D motion sequences.
5. Candidate motion segments generation. For each motion clip contains the candidate 3D initial posture, we also compare its sequential frames with the closing postures in the 2D image plane. If they are matched, the motion segment between these two frames will be extracted as candidate 3D motion segments.
6. Trajectory comparison. For each candidate motion segment, we re-project its 3D trajectory of the specified joints to the 2D plane in terms of the recovered viewing direction, and then compare it with the trajectories drawn by the user. If they are matched, the current motion segment will be returned for further refinement.
7. Motion refinement by 2D sketches. The user can evaluate the recovered motion from any angle, and further make the 3D motion segments consistent with the input 2D sketch. The user can also rework or refine it by editing 2D sketch drawings. The changes on the joint trajectory and postures on 2D sketches will be transferred to 3D motion by inverse kinematics under the constraints of keeping movement continuity.

4. Generating motion segmentations matched to 2D sketches

2D sketches are in general insufficient to interpret the 3D motion data thoroughly, extra information and domain knowledge are needed to infer the desired 3D motion sequence. We employed the similar idea used for 3D posture recovery in [CWW*04] , and a motion database is exploited to guide the recovery of 3D motion sequence. Suppose that two sketches of a human figure are given, the viewing direction of each sketch will be recovered in terms of the reference T-pose. The proposed approach will first retrieve a set of candidate motion clips from the motion database, and extract the motion segments in which the 2D projection of the first and last frames are similar to the human figure in the two 2D sketches respectively. Secondly it will compare the projected 2D trajectories of the specified joints with the annotated trajectory lines of the corresponding joints in the input 2D sketches, and return the current motion segment as the resulting candidate motion if they are matched.

4.1. The alignment of viewing direction and pivotal posture retrieval

The animators often employ the convention of scaled orthographic projection to draw the 2D sketches, therefore we also align the viewing direction under the assumption of orthographic projection. The pelvis part of human body has a constant spatial relationship while performing various movements, and we employ it as a reference pivot to align the 3D posture with 2D sketch. As shown in Figure 4, the viewing direction is pointing along the -Z axis, and there are thee joints, the pivotal(root) joint at the waist, the left hip joint, and the right hip joint, in the pelvis of human torso. The distances between the pelvis joints in the T-pose skeleton are denoted as

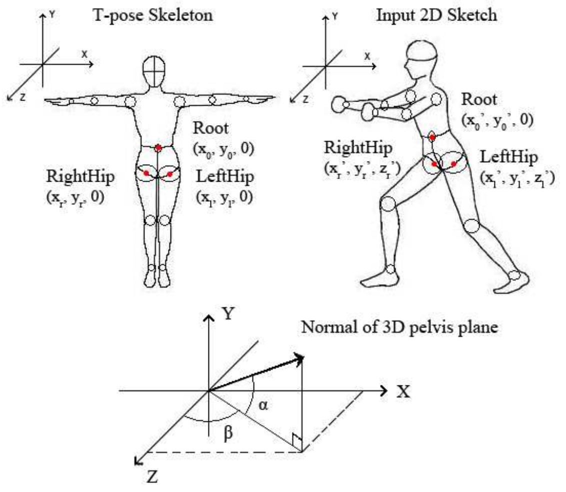

Figure 4: *The normal recovery of 3D pelvis plane.*

$$d_{left}=\sqrt{(x_l-x_0)^2+(y_l-y_0)^2},$$
$$d_{right}=\sqrt{(x_r-x_0)^2+(y_r-y_0)^2},$$
$$d_{leftright}=\sqrt{(x_l-x_r)^2+(y_l-y_r)^2},$$

Assuming that the root joint of the input 2D sketch is positioned on the Z=0 plane, and the user-defined T-pose skeleton has the same projection scale ratio with the 2D sketches drawn later on, the depths of lefthip and righthip joints in the input 2D sketch in Figure 4 are as follows:

$$z_l'=\pm\sqrt{d_{left}^2-(x_l'-x_0')^2-(y_l'-y_0')^2}$$
$$z_r'=\pm\sqrt{d_{right}^2-(x_r'-x_0')^2-(y_r'-y_0')^2}$$

By further applying the distance constraint $d_{leftright}$, we can easily get the two possible normals of 3D pelvis plane in the input 2D sketch. In order to make the alignment process transparent to the user, MotionMaster does not ask the user to manually remove the incorrect human body orientation, and each 3D frame in the motion database will be projected twice when performing the posture and trajectory comparison.

In MotionMaster, we choose the +Z axis as the intermediate reference direction for the alignment. With the two angles of the recovered normal of 3D pelvis plane, α and β, shown in Figure 4, we similarly define the angles, α' and β', for the pelvis normal of the 3D posture in motion database, and the final alignment transformation matrix is

$$R_y(-\beta')R_x(\alpha')R_x(-\alpha)R_y(\beta),$$

where $R_y()$ and $R_x()$ are the rotational transformation matrix along the Y and X axis respectively.

The next step is to compare the projection of a 3D frame with the 2D sketch. In order to minimize the positional imprecision problem in sketch drawings, the similarity metric is defined by the relative spatial relationship among the joints. Suppose that $P(R_0, j_{11}, j_{12}, ..., j_{1n}, b_{11}, b_{12},......,b_{1m}, l_{11}, l_{12}, ..., l_{1k})$ is the projection of a 3D posture from the motion database, , and $S(R_1,j_{21}, j_{22}, ..., j_{2n}, b_{11}, b_{22},......,b_{2m}, l_{21}, l_{22}, ..., l_{2k})$ is the 2D hand-drawn sketch, where R_0 and R_1 are the 2D positions of respective pivotal joint in P and S, j_{1i} and j_{2i} are the i^{th} joint's relative position within the bounding-box of the entire skeleton in P and S, and l_{1i} and l_{2i} are the angle between i^{th} bone direction and the positive Y axis in P and S respectively. b_{1i} is the segment ratio of the i_{th} projected bone length against its corresponding 3D length, b_{2i} is the segment ratio of the i_{th} sketched bone length against its corresponding length in the reference T-pose sketch. The similarity metric between P and S is defined as

$$Dist(P,S)=\sum_{i=1}^{k}w_1|l_{2i}-l_{1i}|+\sum_{i=1}^{m}w_2|b_{2i}-b_{1i}| + \sum_{i=1}^{n}w_3|(j_{2i}-R_1)-(j_{1i}-R_0)|$$

where w_1, w_2, w_3 are the weights specified by the user.

Ideally, we can directly search for the matched frames with the given threshold of Dist(P,S). However, we find that if the best-fit frame is located in a specific motion sequence, its neighboring frames are also picked out as top similar postures. In order to solve this redundancy problem, we further classify the retrieved frames from the same motion sequence into clusters. In each frame cluster, only one representative frame will be employed to locate the candidate motion clip in the motion database. It effectively helps to locate more potential matched motion clips from the motion database.

4.2. Distilling best-fit motions by trajectory comparison filter

The pivotal posture retrieval will often generate multiple candidate motion segments, as the resulting motion segments matched to the sketch drawings is only ensured in a sense of posture similarity. As shown in Figure 5(a), the initial and closing frames of two resulting motion segments are similar to the input sketches, but their trajectories of the hit-joint are obviously different. therefore it is necessary to distill the desired motion segments by trajectory comparison filter such that the trajectory of specified joints are also matched to that in the input sketch.

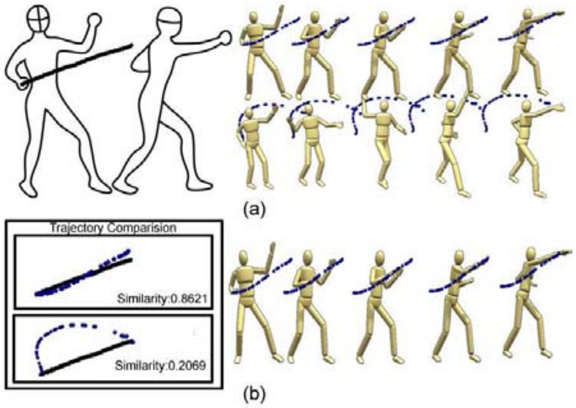

Figure 5: *(a) The matched motion segments by pivotal posture retrieval; (b) The trajectory comparison filter and the survived motion segment.*

A trajectory S is defined as S=[(s_1,t_1),(s_2,t_2),......,(s_N,t_N)] , where N is the number of sample points in S, and s_i =($s_{i,x}$,$s_{i,y}$) is the i^{th}2D point on S sampled at t_i. The basic standpoint to compare the trajectories is to convert the 2D trajectory into 1-D array. We employ the turning angle of trajectory as the metrics [JM04].

Denoting L=$\sqrt{(s_{i+1,x}-s_{i,x})^2+(s_{i+1,y}-s_{i,y})^2}$ as the distance of between two consecutive sample points in S, the turning angle θ_i is defined as $\sin\theta_i = (s_{i+1,y}-s_{i,y})/L$, and $\cos\theta_i = (s_{i+1,x}-s_{i,x})/L$. Given two trajectories S_1=[(s_1^1,t_1^1),(s_2^1,t_2^1),......,(s_n^1,t_n^1)] and S_2=[(s_1^2,t_1^2),(s_2^2,t_2^2),......,(s_m^2,t_m^2)], The trajectories can be converted into the turning angle representations.

$$S_1=[(t_1^1,(\sin\theta_1^1,\cos\theta_1^1)),......,(t_n^1,(\sin\theta_n^1,\cos\theta_n^1))]$$

$$S_2=[(t_1^2,(\sin\theta_1^2,\cos\theta_1^2)),......,(t_m^2,(\sin\theta_m^2,\cos\theta_m^2))]$$

The distance of turning angle, TAD, is recursively defined as follows:

$$TAD(S_1,S_2)=\begin{cases} 0, & \text{m=n=0;}\\ m, & \text{n=0;}\\ n, & \text{m=0;}\\ cost(s_1^1,s_1^2)+MinRest(S_1,S_2), & \text{otherwise.}\end{cases}$$

where cost(s_i^1, s_i^2) is defined as

$$cost(s_i^1,s_i^2)=\{\begin{matrix} 0 & |\sin\theta_i^1-\sin\theta_i^2|+|\cos\theta_i^1-\cos\theta_i^2|<\varepsilon\\ 1 & otherwise\end{matrix}$$

ε is the tolerance given by the user. MinRest() is defined as

$$MinRest(S_1,S_2)=\min\{TAD(S_1,Rest(S_2)),TAD(Rest(S_1),S_2),\\ TAD(Rest(S_1),Rest(S_2))\}$$

Where the Rest() function for the trajectory is defined as

Rest([(s_1,t_1),(s_2,t_2),......,(s_n,t_n)])=[(s_2,t_2),......,(s_n,t_n)].

By the similar DTW algorithmic principles in [VGG02] [COO05], the similarity between two trajectories is finalized as:

$$similarity(S_1,S_2)=1-\frac{TAD(S_1,S_2)}{Max(m,n)}$$

Due to the noise in hand-drawn sketches, the median filter is applied to smooth the trajectory for preprocessing purpose. In MotionMaster, the default threshold of acceptable matching of two trajectories is set as 0.7. An instance to test the effectiveness of trajectory comparison algorithm is given in Figure 6. A sample about how to remove the unfit motion segment by trajectory is given in Figure 5 (b).

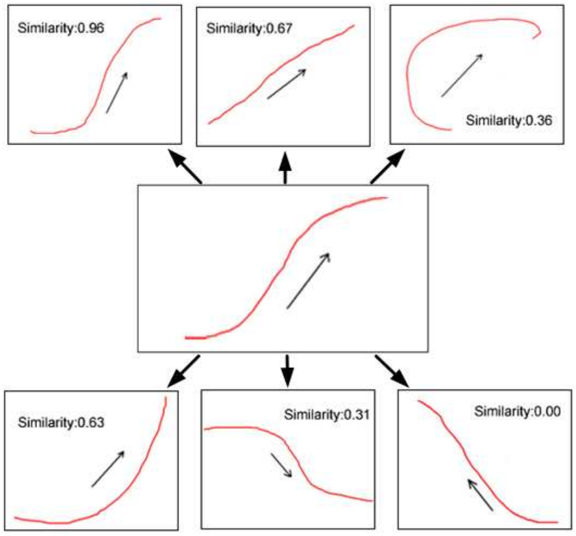

Figure 6: *Testing instance of trajectory comparison filter, the source trajectory is at the center.*

5. Sketch-based motion refinement

After the trajectory comparison process, the survived motion segments have been very close to the desired motion in the sketch drawings. However, it does not mean that they are the same to each other. In order to enforce the constraints in the input sketches, the sketch-based motion refinement is introduced to automatically make changes on the 3D candidate motion segments such that its 2D projection can be consistent with the input 2D sketches.

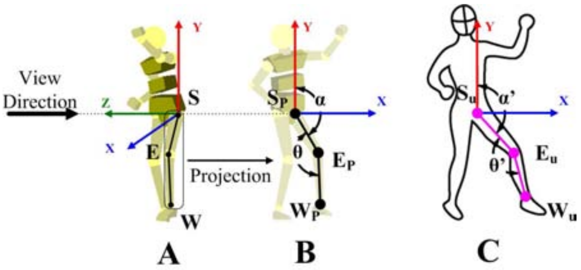

Figure 7: *Example of adjusting posture. A is the 3D frame, B is the projected 2D posture, C is the input 2D sketch.*

The inconsistency between them are mainly embodied in the 2D posture and trajectory of Kungfu forms. Figure 7 shows an example of posture difference in the lower part body, the relevant joints angles are denoted as θ, θ', α, α', and the relevant joint positions are denoted as W_p, S_p, E_p, W_u, S_u, E_u. Let $S_1 = |W_p - E_p|/|S_p - E_p|$, $S_2 = |W_u - E_u|/|S_u - E_u|$, then the posture adjustment is converted into an optimization problem. The objective function is defined as

$$\Delta_{posture} = w_1|\theta - \theta'| + w_2|\alpha - \alpha'| + w_3|S_1/S_2 - 1|$$

Where w_1, w_2, w_3 are the weights set by the user.

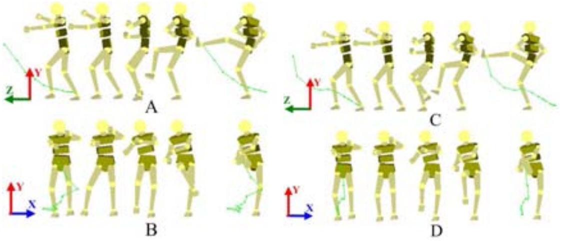

Figure 8: *A and B are the side view and front view of a resulting motion by directly applying inverse kinematic technique for 2D positions of joints. The apparent flickering of right ankle joint is observed in the front view. C and D are the side view and front view of an improved resulting motion using inverse kinematic with movement continuity constraints*

After the initial and closing posture are refined, we first blend the remaining 3D motion sequence in terms of the changes in the first and last frames, and then further refine the trajectory of joints in the remaining motion frames by the inverse kinematic technique. Due to the missing depth information in the sketches, the direct application of positional inverse kinematic techniques will cause the serious joint flickering problem(see Figure 8 B). Therefore we add more constraints to maintain the movement continuity of 3D joints. The final optimal function for motion refinement is given

$$\Delta_{motion} = w_1 \times Diff(TD) + w_2 \times \Delta_{posture} + w_3 \times Diff(Z) + w_4 \times Diff(V)$$

Where w_1, w_2, w_3, w_4 are the weights for different constraints. Diff(TD), Diff(Z) and Diff(V) are the function to compute the difference between trajectories, depths, and velocities respectively. The testing results show that it facilitates the flickering problem obviously (see Figure 8 C and D).

Figure 9: *Example of motion refinement via 2D hand-drawn sketches.*

In order to better assist the evolution process of motion authoring and choreography, we open this interface to the end user in MotionMaster. After the 3D motion data has been associated with the 2D sketch, the user can iteratively make changes on it and the system automatically transfer the 2D changes to the 3D motion sequence. Figure 9 is an instance to manually refine the motion by 2D sketches.

6. Experimental results and discussion

In MotionMaster,the sketch drawings of the new character are guided by the user-defined T-pose skeleton. While drawing the character poses, no special constraints are imposed on the content of drawing, except that the overall size of the drawn character should be in the same/similar projection scale as that in the user-defined T-pose skeleton. The success with them partly depends on the user's ability to draw figures in motion. In the hands of skilled animator, sketch-based authoring and choreographing is a quite flexible, ideal for the communication of complex jointed figures. The current volume of the motion database is about 24000 frames (around 400 Kungfu forms) captured from professional Kungfu actors. In order to maintain the flexibility of database and perform the testing for further development, no index mechanism is implemented yet. Each 3D frame will be reprojected to 2D plane on the fly during the pivotal posture retrieval. MotionMaster can return the resultant 3D motion sequence around 30 seconds when running on Pentium4 2.4GHZ windows platform. The performance is acceptable, as motion authoring and choreography task often proceeds a period of days or weeks. In order to testify the effectiveness of the system, we made a series of experiments to recover the 3D motions by sketches. In most cases, MotionMaster can effectively generate the desired motion from the library. Sometimes, MotionMaster will also return multiple resulting motions (see Figure 10).

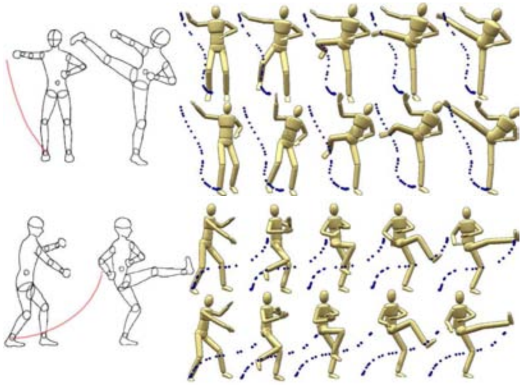

Figure 10: *Multiple resulting motion segments matched to the input 2D sketches.*

The major concern of this approach is the reliability and robustness due to the imprecision of sketch drawings. In MotionMaster, we proposed a joint labeling/marking mechanism as a post-processing step to semi-automatically cleanup the sketch drawings. These labelled joints will automatically form the neat 2D skeleton data in terms of user-defined T-pose skeleton. In order to analyze the reliability of 3D recovery due to imprecision, we stochastically choose 2000 posture frames from the motion database, and add the perturbation to the root orientation. The tolerance in distance metric caused by erroneous torso orientation is then calculated and evaluated. Figure 11 shows the average distance magnitude from the viewing direction with 0, 45, 90, 135 degrees respectively. Based on this analysis and intuitive visual evaluation of resultant motions, we set the default threshold of pivotal motion retrieval as 10 in MotionMaster. That is to say, the tolerance of recovered root orientation error in MotionMaster is around 30 degrees. A skilled animator can easily achieve that in their sketch drawings.

The major limitation of MotionMaster is that it will cause failure if the system can not find any motion segments matched to the input 2D sketches. One of the solutions to overcome it is to enlarge the content coverage of motion database. However, with the increased volume of motion database, an index mechanism should be designed for the pivotal posture retrieval. Based on the current similarity metric, the top 10 similar postures from the sketch-based motion retrieval are mostly in their neighboring frames. Therefore the representative frames of initial and closing posture for each Kungfu form will be the first choice as indexing postures . To the best of our knowledge, 10000 Kungfu forms will cover more than 90% of the martial arts techniques in Chinese Kungfu. It needs 20000 indexing postures in total. Therefore, by the indexing mechanism of representative frames, we are still able to make MotionMaster return the rapid-prototyped motion sequence within a certain amount of seconds.

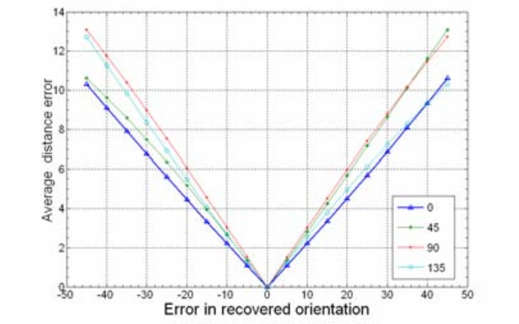

Figure 11: *Tolerance analysis of root orientation in pivotal posture retrieval.*

7. Conclusion and future work

The primary contribution of this paper is that it can directly rough out the desired 3D motion sequence merely from two input sketch drawings, while most existing work focus on the 3D posture recovery from single view sketches. Moreover, this technique can also be extended into the other rhythmic motion domain such as dancing movement.

From the point of view of human computer interaction, MotionMaster significantly goes beyond the sketching interface in [DAC*03]. Their system can only reconstruct the 3D keyframe postures, you have to redraw the in-between frames even if you merely want to change the trajectory of the hit-joint in a Kungfu-form, while MotionMaster builds the overall 3D motion sequence from sketch drawings, and the motion editing have been integrated into the system. The user can easily change the hit-joint trajectory solely by redrawing the annotation motion lines. It is flexible and convenient for the animator to choreograph the Kungfu motions, especially on the combat of multiple persons. Compared with Motion Doodles in [TBP04], MotionMaster focuses more on sketching low-level motions that involve the trajectory specification of moving joints of a figure. These annotated trajectory drawings are consistent with the user's cognitive intuition, and no special training is needed for the animators. Motion Doodles aims at the high level motions such as specification of motion paths of the entire articulated figure, and the end user needs to learn the special annotation vocabulary. In spite of these differences, MotionMaster and Motion Doodles could complement each other in the choreography task.

In the future, we will not only expand the data volume to 5000-10000 Kungfu forms with an indexing mechanism, but also plan to integrate this technique with the sketch-based stylization in [CKB04] to form a solution that can effectively assist the production of hand-drawn animations.

8. Acknowledgements

We wish to thank Shun Zhang for making the demo videos. This work was supported by NSF of China(60373032), by the research funds of Science and Technology Department of Zhejiang Province, and by Bureau of Science and Technology of Ningbo City.

References

[BS76] BROWN M. D., SMOLIAR S. W.: A graphics editor for labanotation. In *Proc. SIGGRAPH '76* (1976), pp. 60–65.

[CBM*93] CALVERT T. W., BRUDERLIN A., MAH S., SCHIPHORST T., WELMAN C.: The evolution of an interface for choreographers. In *Proceedings of SIGCHI Conference on Human Factors in Computing Systems* (1993), pp. 115–122.

[CCYL04] CHAO S. P., CHIU C. Y., YANG S. N., LIN T. G.: Taichi synthesizer: a motion synthesis framework based on key postures and motion instructions. *Computer Animation and Virtual World 15* (2004), 259–268.

[CKB04] CHAUDHURI P., KALRA P., BANERJEE S.: A system for view-dependent animation. *Computer Graphics Forum 23*, 3 (2004), C411–C419.

[COO05] CHEN L., OZSU M., ORIA V.: Robust and fast similarity search for moving object trajectories. In *Proceedings of SIGMOD* (2005), pp. 491–502.

[CWW*04] CHIU C. Y., WU M. Y., WU Y. C., WU C. C., YANG S. N.: Retrieval and constraint-based human posture reconstruction from a single image. In *Workshop on Modelling and Motion Capture Techniques for Virtual Environments* (2004).

[DAC*03] DAVIS J., AGRAWALA M., CHUANG E., POPOVIC Z., SALESIN D.: A sketching interface for articulated figure animation. In *ACM SIGGRAPH/Eurographics Symposium on Computer Animation* (2003), pp. 320–328.

[DCR01] DIFRANCO D., CHAM T.-J., REHG J.: Reconstruction of 3d figure motion from 2d correspondences. In *IEEE Proceedings of Computer Vision and Pattern Recognition* (2001), pp. 1–307–1–314.

[Gav99] GAVRILA D. M.: The visual analysis of human movement: a survey. computer vision and image understanding. *Computer Vision and Image Understanding 73*, 1 (1999), 82–98.

[GY03] GENG W., YU G.: Reuse of motion capture data in animation: a review. In *Lecture notes in computer science 2669* (2003), pp. 620–629.

[IMT99] IGARASHI T., MATSUOKA S., TANAKA H.: Teddy: A sketching interface for 3d freeform design. In *Proceedings of SIGGRAPH 99* (1999), pp. 409–416.

[JM04] JIN Y., MOKHTARIAN F.: Efficient video retrieval by motion trajectory. In *Proceedings of British Machine Vision Conference* (2004), pp. 667–676.

[LC85] LEE H.-J., CHEN Z.: Determination of 3d human body postures from a single view. *Computer Vision, Graphics, and Image Processing 30* (1985), 148–169.

[MG01] MOESLUND T., GRANUM E.: A survey of computer vision-based human motion capture. *Computer Vision and Image Understanding 81*, 3 (2001), 231–268.

[Osh04] OSHITA M.: Pen-to-mime: A pen-based interface for interactive control of a human figure. In *EUROGRAPHICS workshop on sketch-based interfaces and modeling* (2004), pp. 43–52.

[RAD99] RADEMACHER P.: View-dependent geometry. In *Proceedings of SIGGRAPH 99* (1999), pp. 439–446.

[RSR01] ROSALES R., SIDDIQUI M., REHG J.: Estimating 3d body pose using uncalibrated cameras. In *IEEE Proceedings of Computer Vision and Pattern Recognition* (2001), pp. 821–827.

[SAB91] SABISTON W. R.: Extracting 3d motion from hand-drawn animated figures. In *Master Thesis OF Massachusetts Institute of Technology* (1991).

[SCLG90] SCHIPHORST T., CALVERT C., LEE C. W., GAUDET S.: Tools for interaction with the creative process of composition. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (1990), pp. 167–174.

[Tay00] TAYLOR C. J.: Reconstruction of articulated objects from point correspondences in a single uncalibrated image. *Computer Vision and Image Understanding 80*, 3 (Dec. 2000), 349–363.

[TBP04] THORNE M., BURKE D., PANNE V. D.: Motion doodles: An interface for sketching character motion. In *Proceedings of SIGGRAPH'04* (2004), pp. 424–431.

[VGG02] VLACHOS M., G.KOLLIOS, GUNOPULOS D.: Discovering similar multidimensional trajectories. In *Proceedings of International Conference on Data Engineering* (2002), pp. 673–684.

[YKH04] YAMANE K., KUFFNER J. J., HODGINS J. K.: Synthesizing animations of human manipulation tasks. In *Proc. SIGGRAPH '04* (2004), pp. 532–539.