Human Motion Retrieval Using Video or Drawn Sketch

Sumit S. Rathi
PG Student, Department of Computer Engineering
G. H. Raisoni College of Engineering and Management,
Wagholi, Pune-412207
Email: sumitrathi50@gmail.com

Mr. Sachin Patil
Faculty, Department of Computer Engineering
G. H. Raisoni College of Engineering and Management,
Wagholi, Pune-412207
Email: sachin.3400@gamil.com

***Abstract*—**The importance of motion retrieval is increasing now a days. The majority of existing motion retrieval labor intensive, there has been a recent paradigm move in the animation industry with an increasing use of pre-recorded movement of animating exclusive figures. An essential need to use motion catch data is an efficient method for listing and accessing movements. In this work, a novel sketching interface for interpreting the problem is provided. This simple strategy allows the user to determine the necessary movement by drawing several movement swings over a attracted personality, which needs less effort and extends the users expressiveness. To support the real-time interface, a specific development of the movements and the hand-drawn question is needed. Here we are implementing the Conjugate Gradient method for retrieving motion from hand drawn sketch and video. It is an optimization and prominent iterative method. It is fast and uses a small amount of storage.

Keywords*-Conjugate Gradient, Indexing, Motion Retrieval, Motion Strokes, Sketching, Trajectory.*

*****

I. INTRODUCTION

Motion capture is the process of sampling the location and position information of an object over time. The object is usually an animal, a machine or a person. In case the object is an animal or a person, it is generally referred to as an 'actor'. Especially, motion capture is also abbreviated as 'mocap' in the entertainment industry. Over the last few years, motion capture information has developed and grabbed a main part in animated graphics, games and 3D surroundings. The genuine individual activity is an integral part of media like video games and movies. More natural figures make for more immersive environments and more credible effects. At the same time, realistic animation of individual activity is a challenging task, as people have confirmed to be skilled at critical the intricacies of human movement and determining discrepancies.

One common solution for this problem is motion capture and recovery. However, while activity capture and retrieval is an efficient way of acquiring realistic human motion, by itself it is a technique for recreating activity. Motion capture information has confirmed to be difficult to alter, and editing techniques are efficient only for small changes to an activity. The purpose of motion capture is to get the motion data of certain points of interest in the object, so that either some parameters of the motion (e.g., distance, speed, etc.) can be calculated or the information can be utilized to control or drive something else. Motion parameters are calculated and used in the application like sports analysis, motion analysis, biomechanics, etc. The data is used to control a machine; the application can be tele-robotics, tele-surgery, motion feedback control, etc. In case the data is used to control some displays or something else, the application may be virtual rehabilitation, interactive games, virtual training, virtual reality, motion directed music, etc.

A system will be developed capable of capturing the real time motion of the user's hand drawn sketch or motion video. The motion data captured will be recorded to a control interface, which will manage the movement of an object, The system will be intended for a common user therefore permitting more design focus on the interactive procedure of motion capture. It will confirm to general usability values permitting the majority of users to appropriately operate the system. Thorne and Burke present a more convenient approach where an activity is created by drawing a continuous series of lines, arcs, and circles[7]. They are parsed and then planned to a parameterized set of movements. A sketch program for creating Kung-fu movements is presented in the work of Li et al. [8]. They use pictures to explain the first and last positions. Next, they recover the similar postures as well as the in-between movements by projecting the supports and the 3D data source activity trajectories to 2D plane, and related them in that space. The controller uses motion capture to track, its position in accordance with the television, allowing users to use it in a variety of various ways e.g. cricket, tennis racket, golf. Since human motion data are spatial temporal, looking into this data to understand the contents or finding a specific part of them requires some effort. In animation, the users are watching the animation from initial to last frame with different the view position. The problem causes more serious as the size of the database increases.

Our proposed system permits the iterative improvement of the selections, limiting the motion to fit a more accurate pose description. By combining with the encoding of the query and motion repository, the system can be used in interactive scenarios. Here we are implementing the Conjugate Gradient method for retrieving motion from hand drawn sketch and video. The Conjugate Gradient method is an optimization method, also one of the most prominent iterative methods. It is fast and uses a small amount of storage.

II. RELATED WORK

Author [1] has suggested a new technique to index positions of two figures carefully getting each other. The technique is based on the concept of logical tangles, and it is proven that we can classify various positions of two figures twisted with each other. They have also proven that a platform line technique using low-level features such as the position of the joint parts can experience from categorizing such positions. They suggested scribing the troubles created between the worldwide ways linking the end effectors; another strategy to scribe the troubles is to calculate the regional GLI between smaller parts such as those created by the divisions. Such a strategy might be more effective as we will only need to encode the regional community where the troubles are consisting. However, a disadvantage is that another strategy to calculate the resemblances of positions which are consisting of different sections will be required. The system [2] makes 3D types for a wide type of form drawings; certain restrictions avoid it from working globally. One is that the contour-completion strategy is localthe finished from shape relies on the geometry of the beginning and finishing factors, but disregards the rest of the feedback shape; it will need a much further knowing of form finalization to deal with this. Our inflation criteria currently need adjusting constants; the constants that generate the most satisfactory looking outcomes actually generate self-intersecting areas, especially in places like armpits (i.e., between a branch and a body).

Qinkun Xiao et. al [13] have suggested a content-based motion taken information recovery criteria. The suggested WGM criteria first select associate frames and the weight values for every associate frame. A calculated chart is developed for evaluating two motions, and to evaluate the likeness between the two motions, the matching on this calculated chart is used. The suggested WGM-based motion recovery criterion has been tried on the CMU [6] data source. Experimental outcomes and comparison with pre-existing matching methods illustrate that the suggested technique outperforms different methods for motion recovery. In this document [4], authors have suggested a two-level statistical model, known as movement structure, to catch complicated movement characteristics. Motion structure is shown by a set of movement textons and their submission. Regional characteristics are taken by movement textons using straight line powerful techniques, while international characteristics are modeled by changing between the textons. With the learned movement structure, they can synthesize and modify movement easily. Limitations: Although our strategy has been effective in generating realistic and powerful movement, there stay several places for improvement.

Zhigang Deng, Qin GU and Qing Li [5] showed a perceptually steady, example centered human motion recovery system focused around a modern pattern removal and matching technique. Given a query motion, their technique can efficiently recover wisely relative motions from an extensive motion’s information data source. The expertise of their technique specifically profits from the quick performance of the established KMP sequence matching computation and the KDtree structure. To evaluate the perfection and convenience of their technique, they instructed similar client studies to evaluate its query perfection by distinct our technique and three the condition of the arts, example centered motion look for methods. By analyzing user study outcomes, they found that their technique was measurably powerful regarding look for perfection, and its look for motion outcomes are consequently organized in an approximately perceptually consistent order. High dimensionality and spatio-temporal domain is an interesting subject in Visualizing human motion data. To show Bouvier -Zappa et al. [9] added cartoon-like signs such as noise waves and speed lines to show a short past motion of the character body with the current moment of the motion. On the basis of horizontal time-bar Yasuda et al. [10] displayed densely sampling key postures. The user can view the change of in postures by observing through it from left to right.

A systematic method in motion data for selecting key postures is introduced by Assa and his colleagues. This method is used to make the images as similar to hand drawing of motion as possible. By using a simple input query its difficult to retrieve the complex motion data.

ISSN: 2321-8169
5478 - 5483

TABLE I. COMPARISON BETWEEN MOTION CAPTURE TECHNIQUE

Method name	Feature	Concept / Algorithm	Limitations	Advantages
Pixel based	Color, pixel	Comparison based on RGB/gray values	The objects are away in the frame are ignored	Fast, simple
Region based	Motion field, shape, color	Try to extract two fields stationary and motion fields are in two different region	Problem in extracting the region when the region are overlapping	Less memory required
Inter frame difference	Color, motion, pixel	Comparison based on block by block	Slow moving object cannot be tracked	Fast
Kernel density estimation	Pixel	Background is given by a histogram of most recent pixel values	More memory is required	Accurate
Median filter	Pixel, speed	Based on averaging the previous background to get new background	More computational time is required	Used to remove the noise
Conjugate gradient method	Pixel, angle	Based on calculation of x gradient and y gradient	More calculation is required	Fast and used small storage space

For example, Mller et al. [11] suggested to extract the simple features of motion data and further used it for an efficient comparison between various motion data.

Deng etal defined the partial matching of the motion, in which the query can be a mixture of body parts or a single part of body motion (e.g. left leg) in different motion clips. Here, if the user did not get desired output the only way to get it would be to capture it directly. To define the trajectory constrains sketch input is used, in the character animation, which its specific joint point or character should move alone. To a specific style of locomotion Thorne et al. used predefined gestures. The user can define the constraints, the Walking style and travelling path by sketching multiple gestures continuously in the 3D environment. The kind of locomotion is restricted by the number of defined gestures. As stick figures can draw without training by the user, they have been used as the input for posture constraints. A system developed by Pan et al. In which the key-postures of a 2D character animation set by the user by sketching its skeletal structure. There also have been studied from 2D sketched stick figure to reconstruct 3D key-posture. Wei end Chai defined a likelihood model of 3D character poses given stick figure-like 2D strokes because the model was used to retrieve static posture from a large motion database. In this paper sketch or video containing the emotion strokes is taken as input query motion trajectories are encoded by the conjugate gradient method.

III. PREPROCESSING, INDEXING, RETRIEVAL

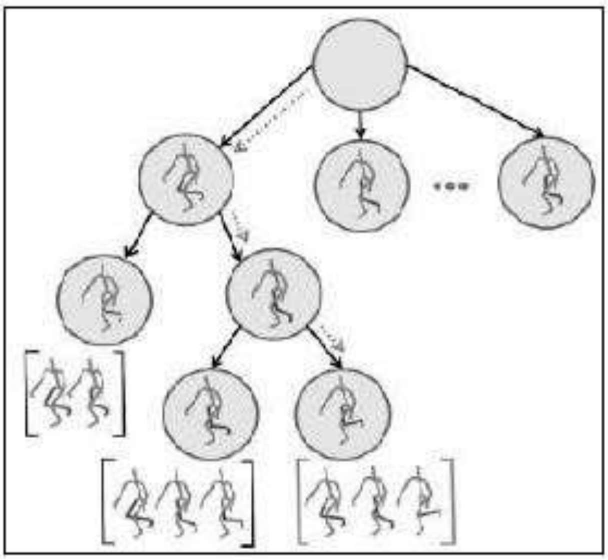

Fig 1: Database indexing.

The pre-processing of motion data consists of three main steps: hierarchy construction of body that divide the human body into a number of parts in the motion segmentation (fig 2), spatial domain and normalization that segment part based motions and then group them into a set of motion patterns (called motion prototype), which basically partitions human motions in the temporal domain, and motion pattern extraction that identify and determine patterns by grouping similar motion segments.

Based on the spatial connectivity of the human body a hierarchical human structure is constructed in (fig 3). The human body is separated into ten meaningful basic parts, e.g. left arm, head, left hand, torso, etc. and then a hierarchy with four iterative layers are constructed accordingly. It includes ten leaf nodes stand for basic body parts, eighteen nodes: the parent nodes in the middle layers

 ISSN: 2321-8169
5478 - 5483

correspond to meaningful grouping of child nodes and the root node represents the complete human body.

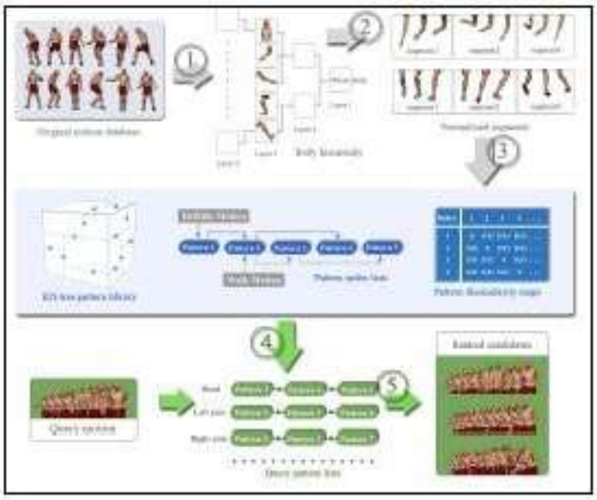

Fig.2: Motion data preprocessing.

We prefer a human hierarchical representation, since it gives a logical control granularity. In addition, a multilayer hierarchy usually takes care of the correlations among various human parts that will be exploited for motion matching computation. In our system, we use joint angles instead of 3D marker positions on behalf of human motion data, due to the fact that the joint angle representation is suitable for unifying the motions.

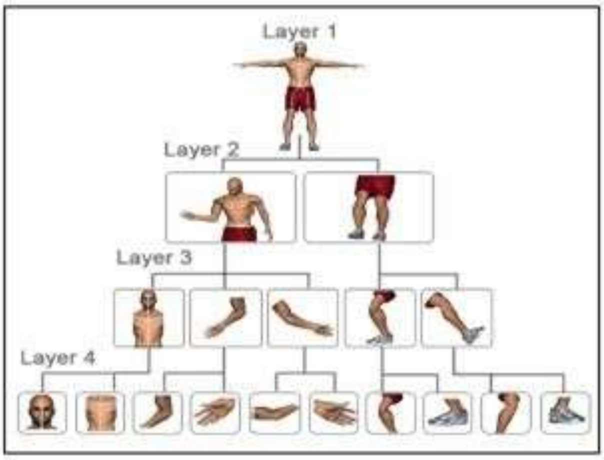

Fig.3: Hierarchy level construction of human body.

Normally, raw motion capture systems having long motion sequences with variances, e.g., a couple of minutes. However, desire motions in many applications are shorter motion clips that satisfy the users particular necessities. Thus, automated motion segmentation processes that adaptively cut the long motion into short clips are essential for the later motion retrieval algorithm.

IV. IMPLEMENTATION DETAILS

A. System Overview

The proposed system, consist of an interface by which the user will be able to give motion stroke input as a skech or motion video. A user can give various motion strokes like jumping, walking, kicking, punching, through video or sketch. The proposed system contains two major parts; motion encoding and motion retrieval. The main idea is to encode the motion trajectories of the input and the videos in the database by using a small set of conjugate gradients. This permits an efficient indexing of the motions and a fast retrieval by matching conjugate gradient coefficients of the motions. Motion videos are divided into the sub clips to support complex motion videos which contain numerous actions. Each sub clip associated with only single action. In our implementation, by locating the key poses the segmentation is done.

Initially input sketch/sub clip is divided into different parts like lower body portion, upper body portion. Then find out which body limb is chosen, then based on the limb selection videos will be obtained from the database. Database video divided into the different limbs with the same section of the body. The user gives the input trajectories in the form of sketch or video. Videos from the database are retrieved on the basis of the trajectory of limb. The hierarchy of the structured in four levels: lower/upper body, full body, all the body main limbs (leg, head, etc.) and the body joints. This description supports input describing movement in various scalesboth body motion and up to the movement of the several joints, and let our system has the flexibility and capability of retrieving numerically and logically similar motions. To retrieve motions, the user draws the sketch or gives input video with desired motion.

B. System Architecture

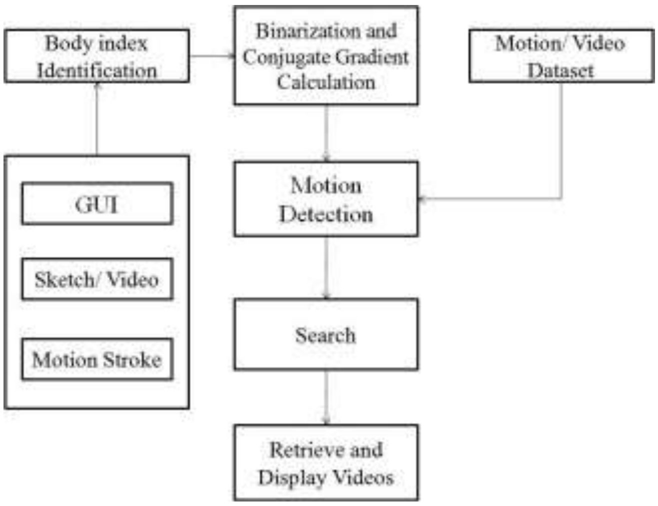

Fig.4: System Architecture.

C. Pseudo-code

Input: Sketch and video with motion trajectory

Output: Retrieve videos having same motion strokes in input,

Step1: Input Image, sketch or video

Step2: Input pre processing

Step 3: Obtain the trajectory and save its co-ordinate vectors. Step4: Calculate min distance for limbs by using Euclidian Step5: Selecting body with trajectory.

Step6: Calculate limb index for trajectories.

Step7: Mapping trajectory on body with minimum error threshold.
Step8: Conjugate Gradient calculations: (Calculate the conjugate
gradient by conjugate gradient method of the trajectory with the co- ordinate axes.)
Step9: Conjugate Gradient mapping and correlation i.e. The co-relation between the database and query trajectory is calculated.
Step 10: Based on the similarity matching between the above two Components videos are ranked and retrieve.

D. Experimental Setup

The system is built using Java frameworks (version jdk 6) on Windows platform. The Netbeans (version 6.9) is used as a development tool. The system doesnt require any specific hardware to run; any standard machine is capable of running the application.

E. Dataset

Here we are using public CMU motion database [6], which provide us about 6000 video clips and more than lacks of frames.

V. EXPECTED RESULTS

A. Result Analysis

The system uses gradient algorithm to calculate the coordinates of the pixels. Then based on the various motion strokes provided by the user these trajectories are stored in spatial domain. The graph below shows the x-gradient and y gradient of the single stroke provided by the user. X-gradient will map the x coordinates of the input sketch to the no. Of iterations and no. of pixels of the x axis in the graph and y gradient will map the y coordinates of the input sketch to the no. Of iterations and no. of pixels of the y axis of the graph. For the different motion strokes there will be change in no. Of iterations and no. of pixels. The system is able to implement the five motions like walk, running ,jumping, running and walking. Other different motions we can implement like dance moves, other sports activities. These new motions can be implemented by the training algorithm.

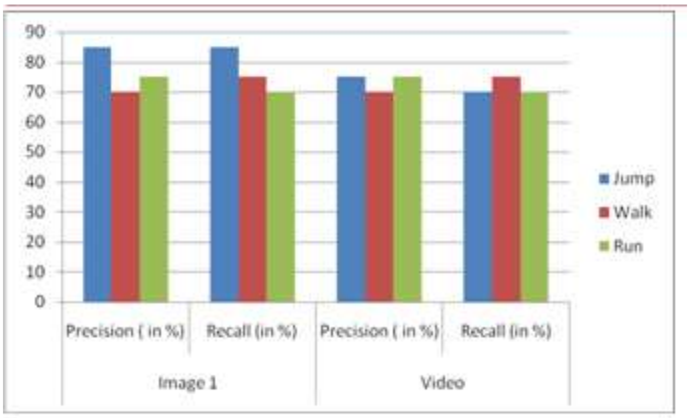

Fig.5: Accuracy graph.

VI. CONCLUSION

The video content has been increasing day by day so there is a strong requirement to retrieve these videos in a very quick time.For the same purpose there should be some efficient method to retrieve it quickly. In this paper the new way of retrieving the videos is proposed and implemented. The implemented algorithm first determines to which limb the motion is provided, based on that it retrieves the videos. In this user specifies the required motion stroke such as the motion trajectory of walking, jumping,running etc. In the respective limb such as punching motion stroke would be provided at the foot,it should get the motion clips in which these motions are present. So the proposed system is able to retrieve the three kinds of motions proposed with around 70-80% accuracy.

ACKNOWLEDGMENT

We would like to thank the researchers as well as publishers for making resources available and teachers for their guidance. We also thank the college authority for providing the required infrastructure and support. Finally, we would like to extend a heartfelt gratitude to friends and family members.

REFERENCES

[1] E.S.L. Ho and T. Komura, Indexing and Retrieving Motions of Characters in Close Contact, IEEE Trans. Computer Graphics and Visualization, vol. 15, no. 3, pp. 481-492, May/June 2009.J. Clerk Maxwell, A Treatise on Electricity and Magnetism, 3rd ed., vol. 2. Oxford: Clarendon, 1892, pp.68–73.
[2] O. Karpenko and J. Hughes, Smoothsketch: Third Free-Form Shapes from Complex Sketches, ACM Trans. Graphics, vol. 25, no. 3, pp. 589-598, 2006.
[3] E. Keogh, Exact Indexing of Dynamic Time Warping,VLDB 02: Proc. 28th Int'l Conf. Very Large Data Bases, pp. 406-417, 2002.
[4] Y. Li, Y.L.T. Wang, and H.-Y. Shum, Motion Texture: A Two- Level Statistical Model for Character Motion Synthesis,ACM Trans. Graphics, vol. 21, no. 3, pp. 465-472, 2002.
[5] Z. Deng, Q. Gu, and Q. Li, Perceptually Consistent Example Based Human Motion Retrieval I3D 09: Proc. Symp. Interactive 3D Graphics and Games, pp. 191-198, 2009.
[6] CMU 2003. Motion Capture Database, http:// mocap. cs.cmu.edu/, 2011.
[7] C.-Y. Chiu, S.-P. Chao, M.-Y. Wu, S.-N. Yang, and H.-C. Lin, Content-Based Retrieval for Human Motion Data,J. Visual Comm. and Image Representation, vol. 15, no. 3, pp. 446-466, 2004.
[8] Q.L. Li, W.D. Geng, T. Yu, X.J. Shen, N. Lau, and G. Yu, Motionmaster: Authoring and Choreographing Kung-Fu Motions by Sketch Drawings, SCA 06: Proc. ACM SIGGRAPH/ Eurographics Symp. Computer Animation, pp. 233-241, 2006.
[9] Efficient Motion Retrieval In Large Motion Databases Mubbasir Kapadia I-Kao Chiangy Tiju

ISSN: 2321-8169
5478 - 5483

Thomasz Norman I. Badlerx Joseph T Kider Jr University Of Pennsylvania.

[10] Bouvier-Zappa S., Ostromoukhov V., Poulin P.: Motion Cues For Illustration Of Skeletal Motion Capture Data.In Proceedings Of The 5th International Symposium On Nonphotorealisticanimation And Rendering (2007)

[11] M. Muller, T. Roder, and M. Clausen, Efficient Content-Based Retrieval of Motion Capture Data, ACM Trans. Graphics, vol. 24, no. 3, pp. 677- 685, 2005.

[12] Min-Wen Chao, Chao-Hung Lin, Member,Human Motion Retrieval from Hand-Drawn Sketch IEEE, Jackie Assa, and Tong-Yee Lee, Senior Member IEEE, IEEE Transactions On Visualization And Computer Graphics, Vol. 18, NO. 5, May 2012

[13] Qinkun Xiao and Luo Yichuang, Lv Zhongkai, Motion Retrieval Based on Graph Matching And Revised Kuhn-Munkres Algorithm published in Signal Processing, Communication and Computing (ICSPCC), 2013 IEEE International Conference.