Motion Retrieval Using Motion Strokes over Character Sketch

Mr. Raju G. Masand[#1], Mr. A.N. Bhute[#2]

[#1]*Student, Department of information technology*
Sinhgad college of engineering ,University of pune,India
[#2]*Associate Professor, Department of information technology*
Sinhgad college of engineering ,University of pune, India

Abstract -- **The growth in the motion capture data has increased the importance of motion retrieval Systems. The idea of motion capturing originated two decades ago, motion capture systems had been developed that allow to track and record human motions at high spatial and temporal resolutions. The resulting motion capture data is used to analyze human motions in fields such as sports sciences and biometrics (person identification), and to synthesize realistic motion sequences in data driven computer animation. There are various techniques for the creation of new, realistic motions from pre-recorded motion clips that have become an active field of research. Such techniques depend on motion capture databases covering a broad spectrum of motions in various characteristics. The 2D sketch figures can express a wide range of human motion, and they can be easily drawn by people without any professional training. The paper is implemented as the user will draw the motion strokes over a sketch like walk, jump, punch, kick, etc. The relevant motion clips based on the drawn motion will be retrieved from the database. This paper gives the description of the proposed system. In which there will be one interface through which the user will specify the query by drawing the sketch. Then proving the motion to it, based on the motion specified the relevant motion video clips will be fetched from the database. Database already contains the motion clips of some motion type. This approach actually employs motion pattern discovery and matching scheme that breaks human motion into parts or hierarchical motion representation. It actually divides the human body into parts and then try to find the motion in those parts.**

Keywords - motion capture, indexing, retrieval, sketch recognition, motion Strokes, trajectory.

I.   INTRODUCTION

Motion capture is the process of sampling the posture and location information of a subject over time. The subject is usually a person, an animal or a machine. In case the subject is a person or animal, it is sometimes referred to as an 'actor'. Especially in the entertainment industry, motion capture is also frequently abbreviated as 'mocap'.

The technical goal of motion capture is to get the motion data of certain points of interest on the subject, so that either some parameters of the motion (e.g., speed, angle, distance, etc.) can be calculated or the data can be used to control or drive something else. In case parameters of the motion are calculated, the application may be motion analysis, sports analysis, biomechanics, biodynamic, etc. In case the data is used to drive a computer generated (CG) character or a scenery to mimic the motion, it is referred to as animation or visual special effects (VFX), etc. In case the data is used to control a machine, the application may be tele-surgery, tele-robotics, motion feedback control, etc. In case the data is used to control some displays or something else, the application may be virtual reality, interactive games, virtual training, virtual rehabilitation, motion directed music, etc.

Using motion capture, a system will be developed capable of capturing the real" time motion of a user's hand drawn motion. The motion data captured will be mapped to a control interface, which will control the movement of a object, The system will be designed for a generic user therefore allowing more design focus on the interactive process of motion capture. It will conform to general usability standards allowing the majority of users to correctly operate the system with little. The announced film production of Tintin has recently made news due to its ground breaking technology, which allows the director to look at a monitor showing a real time fully rendered 3D virtual set, with the actors animated as their characters.

The second, and less common until recently, is advanced user interfaces. The keyboard, mouse and games controller have dominated the area of Human Computer Interaction (HCI) since the late 1970s, but recently there has been movement away from these towards more innovative and natural methods of interaction. Mocap systems are increasingly being used in vision based interfaces such as gesture recognition, for sign language applications. The Nintendo Wii, launched in 2007, gives users the ability to interact with games similar to that of a real world experience. It is supplied with a controller which uses motion capture to track its position in accordance with the television allowing users to use it in a variety of different ways e.g. golf club, tennis racket. Since human motion data is spatio-temporal, looking into this data to understand the contents or finding a specific part of them requires some effort from the users watching the animation from first to last frame while changing the viewing angles. As the size of the database increases, the

problem becomes more serious. Therefore, most public motion libraries use keywords as the medium for browsing and searching through motion data.

The user can look through the database by reading the keywords and directly retrieve the motion data by typing one of a keyword as an input query. However, the keywords have obvious limitations as the medium for motion data. The keywords may be subject to the person who decided on them. In addition, a few words are not sufficient to explain and identify many parts of human motion. For example, when there are two different free styled dances in the database, it would be hard to decide on a proper keyword for each, which will uniquely describe the dances well and will be easy for the user to think of when he or she tries to directly retrieve it with a keyword. If we value the success rate of the retrieval, a choice of compromise would be one general word such as "Dance" or "Free-Style Dance" for both. In terms of motion browsing, displaying static images of well-selected key postures would be a good alternative. The key postures can be selected in procedurally, so we do not need to rely on a subjective decision. Anonymous motion and even meaningless movements can be represented in this way, and the user can understand the contents of the motion data quickly and intuitively by looking through the key postures in time sequence. However, with this method it is difficult to support a direct retrieval interface with any input query. When the database size is large, searching for a large volume of the images would again be a burden for the users. In this paper, stick figures are suggested as a unified medium for both browsing and searching for human motion data. The stick figure is a simple drawing style, which is used to depict a brief human motion. It is very easy to draw, even as children can do it without any training, and can expressive a wide range of human motion well. In our interface, each motion data file in the database is visualized as a consecutive stick figure image, each of which describes a selected key posture. The challenge in the motion retrieval is to reduce the difference between the users' sketches and generated images of the same motion as much as possible. Visualizing the motion as the same style as the user's sketches is the key idea to achieve this goal. Not only the drawing style, but also the procedures generating the image imitates people's drawing. The preliminary survey was conducted, in which the participants were asked to explain given motion data through their hand sketched stick figures. There were number of observations which are being implemented in this paper.

II. RELATED WORK

Visualizing human motion data has been interesting subject because of its spatio-temporal domain and high dimensionality. Bouvier -Zappa et al. [6] added cartoon-like signs such as speed lines and noise waves to the character

body to show a short past motion of the current moment of the motion. Yasuda et al. [4] displayed densely sampling key postures on the horizontal time-bar. The user can see the change of the postures by looking through it from left to right. Assa and his colleagues introduced a systematic method selecting representative key postures from motion data. To visualize motion data by displaying the selected key posture with speed lines in a raw, but the main difference is that our images will be compared to hand-sketched images for the retrieval. In this method it is tried to make the images as similar to people's hand drawing of human motion as possible. The complexity of motion data also makes difficult its retrieval using a simple input query. For example, Müller et al. [8] defined and extracted the simple features from motion data and then used it for an efficient comparison between different motion data. Deng etal. suggested the partial matching of the motion, in which the query can be a single part of body motion (e.g. left leg) or a mixture of body parts in different motion clips. In these methods, however, if the user did not have the query motion which is similar to the desired one, the only way to get it would be to capture it directly. In the interface which uses hand-sketched stick figures as the input queries, and they can be sketched by using a general mouse or tablet. Sketch input, in the character animation, are mostly used to define a trajectory constraint, which the character or its specific joint point should move alone. Thorne et al. used predefined gestures each of which corresponds to a specific style of locomotion. By sketching multiple gestures continuously on the 3D environment, the user can define both constraints, the travelling path and walking styles. The kind of locomotion is limited by the number of defined gestures. Since stick figures are as simple as people can draw without training, they have been used as the user input for posture constraints. Pan et al. developed a system in which the user can set the key-postures of a 2D character animation by sketching its skeletal structure. There also have been studied of reconstructing a 3D key-posture from a 2D sketched stick figure without the knowledge of the captured database. On the other hand, we use the input stick figures as queries to retrieve similar motion segments from the large database. Wei end Chai developed a likelihood model of 3D character poses given stick figure-like 2D strokes. Because the model was learned from a large motion data, this system also could be used an interactive static posture retrieval system. Our definition of stick figures includes the moving trajectories of

III. SYSTEM OVERVIEW

The proposed system, consists of interface through .which the user will be able to draw the motion stroke over the predefined sketch. User can draw different motion strokes like of walking, jumping, punching kicking, The main idea is to encode the motion trajectories of the query and the clips in the

database by using a small set of gradients. This allows an efficient indexing of the motions and a fast retrieval, by matching the x-gradients and y-gradients coefficients of the motions. To support complex motion clips which contain several actions, it begins by splitting such clips into sub clips. Each sub clip contains only one action. In our implementation, the segmentation is done by locating the key poses. Initially input sketch is partitioned into different parts like upper body portion , lower body portion. Then determine which body limb is selected then based on the limb selection videos will be fetched from the database. The same sectioning of the body into different limbs is also done in the database videos the user draws the trajectories onto the sketch. This trajectory along with the lib is processed and based on these videos from the database are retrieved. The hierarchy of the is organized in four levels: full body, upper/lower body, all the body main limbs (leg, arm, etc), and the body joints. This description supports queries describing movement in several scales—both full body motion and up to movement of the various joints, and let our system has the flexibility and ability of retrieving logically and numerically similar motions. To retrieve motions, the user sketches the desired motion.

IV.   FRAMEWORK OVERVIEW :

Fig no.1 presents an overview of system framework. The method takes as input raw motion data and extracts meaningful features from the motion to provide a compact, representative space to index into the database. End users specify queries as combinations of sequences of a variety of motion features and the framework returns motion sub sequences that satisfy these properties. Motion Feature Extraction and Indexing (Offline). Given a large motion database it first compresses the motion data by extracting only relevant key frames. The motion data is represented as a set of curves for each joint angle.

A set of candidate key frames are selected which map to the extreme points of these curves to produce a simplified representation which approximates the original curve within a reasonable error threshold. Next, it define an extensible set of motion keys which characterize the different structural, dynamic, and geometric properties of the motion over a time window . Keys can be computed directly from the motion, or computed using other keys. During an offline process, we compute all key values for all motions in the database. Key values may be of different data-types and may have arbitrary ranges. For intuitive query specification and efficient indexing, we first define a minimal alphabet for each key. Key values are converted into this language to populate a tire data structure which facilitates efficient motion subsequence matching which is independent of the number of motions in the database.
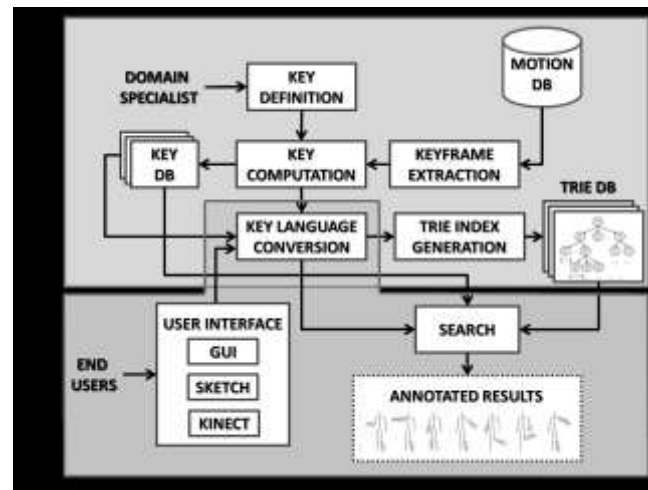

Fig no.1 framework

V.   HOW PEOPLE DRAW STICK FIGURES :

To observe how people draw stick figures when they explain human motion, The participants were asked to draw stick figures that explain the given motion clips. The motion clips were bowing, slow/fast walking, slow/fast kicking, and tuning kick. Figure 2 shows input interface for drawing the sketch, the common rules can be formed for basic components such as selected key postures, added path lines, and viewing angle. Key Postures The key postures are mostly selected at the moment when the motion changes the direction or stops. For example, the key postures of the kicking motion were at the moment when the kicking foot(fig no. 2.3) is at the peak and the kicking is finished. Moving Path In some motion data, the points of changing the moving direction are not clear. For example, while the hand(fig. no. 2.1) is moving along a circle, the direction changes continuously. In these cases, participants usually added a curved line to depict the moving trajectory. In addition, when the motion is relatively fast, multiple lines were drawn along the trajectory like speed lines in comics. The sketching interface allows users to define the motion by simply drawing a motion line (called "motion stroke") describing the character motion. The motion strokes are drawn on a selected character pose and a selected view. The main goal of the method is to infer the corresponding 2D trajectories from the hand-drawn strokes. System begins by allowing the user to select a desired character pose Next, the user specifies the motion by adding strokes to the character. Using a single key posture for motion retrieval is insufficient to distinguish between motions, In this method, the proposed way of using motion lines, on top of the selected posture, to better define the desired motion. Each of the hand-drawn motion strokes is approximated to an ellipsoid. This type of quadric surface is a good candidate for trajectory fitting of many hand-drawn sketches, and is relatively robust to the camera viewpoint differences. It assigns each stroke to the joint nearest to its starting point. The following page shows the different motion strokes that are being implemented in this paper such as punching(fig. no. 2.1), walking(fig. no. 2.2), kicking(fig. no.2.3), jumping(fig. no.2.4), two motion strokes simultaneously(fig. no.2.5)
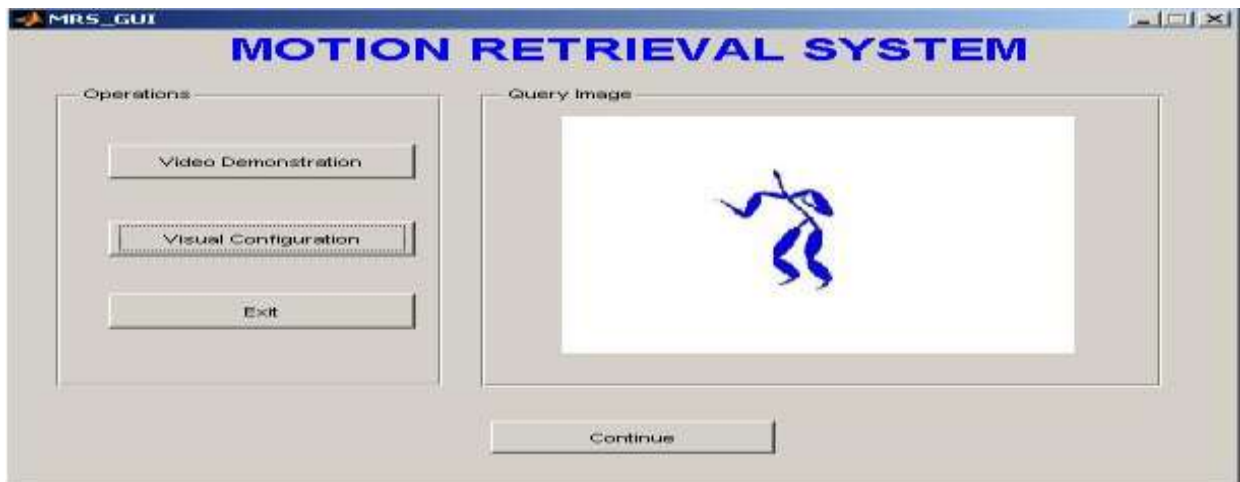
.

System Snapshots

Fig no.2 input to the system
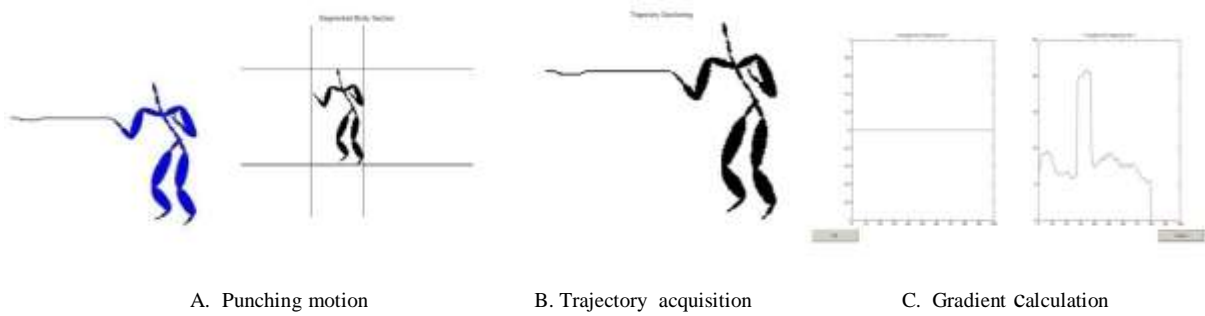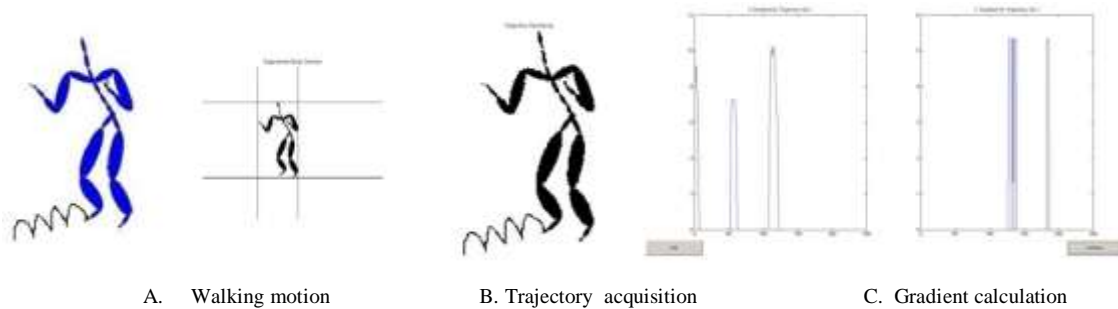
Input : Punching motion Stroke

A. Punching motion   B. Trajectory acquisition   C. Gradient calculation

Fig. no 2.1 Output screen of punching video retrieval

Input : Walking motion Stroke

A.   Walking motion   B. Trajectory acquisition   C. Gradient calculation
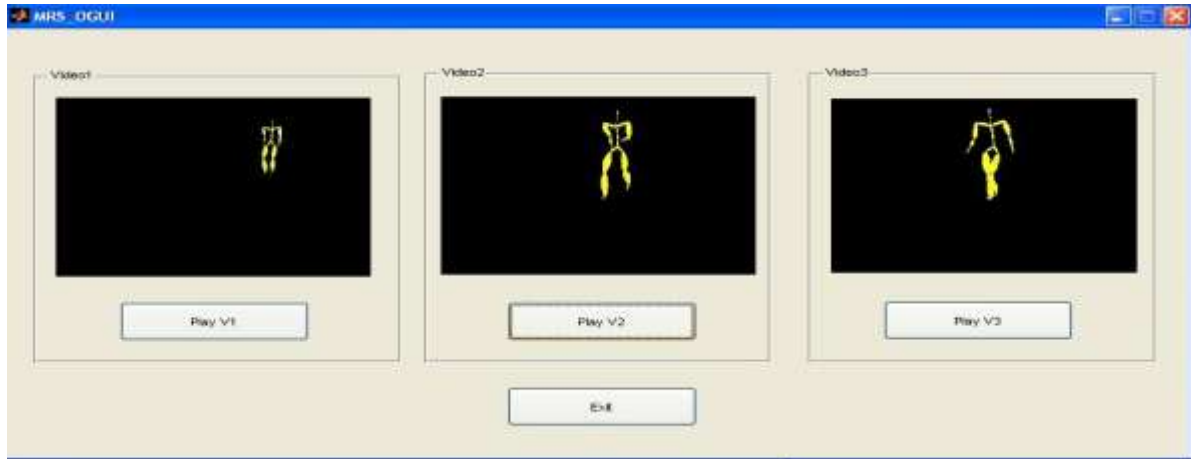
Fig. no 2.2 Output screen of Walking video retrieval

Input : Kicking motion Stroke

A.   Kicking motion   B. Trajectory acquisition   C. Gradient calculation
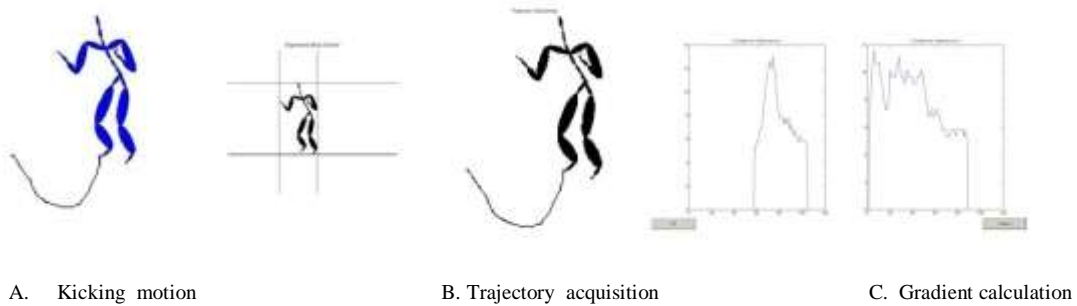
Fig. no 2.3 Output screen of Kicking video retrieval

Input : Jumping motion Stroke

A.   Jumping motion   B. Trajectory acquisition   C. Gradient calculation
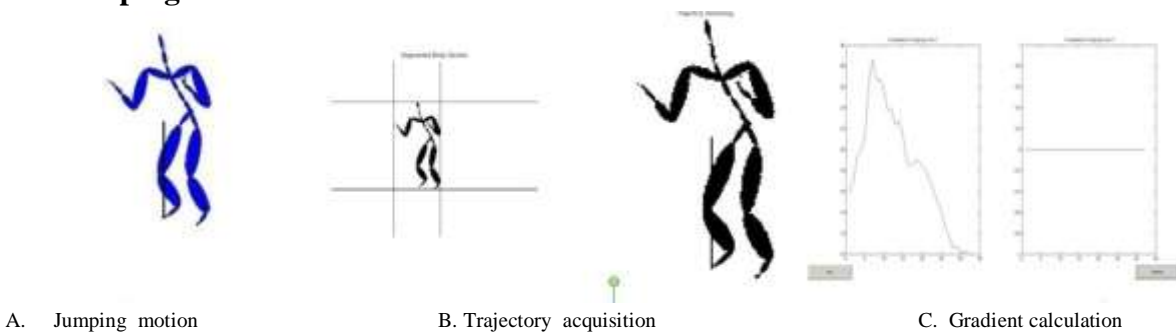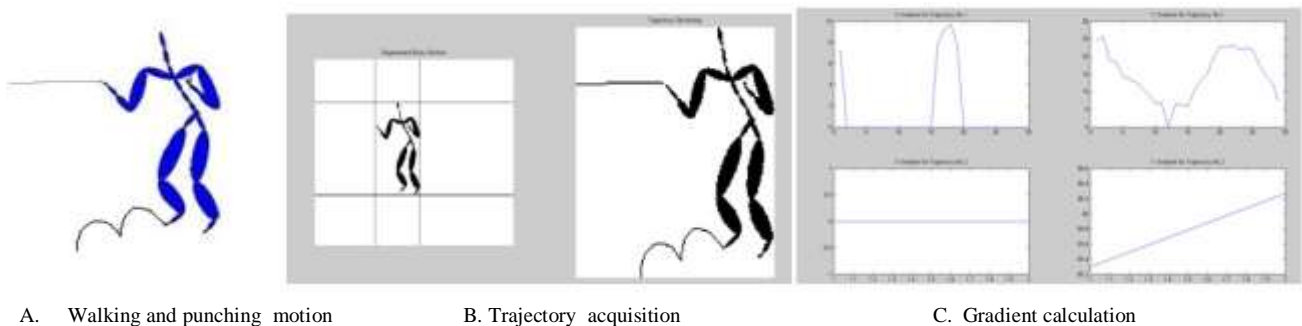
Fig. no 2.4 Output screen of Jumping video retrieval

Input : Walking and punching motion Stroke

| A. Walking and punching motion | B. Trajectory acquisition | C. Gradient calculation |

Fig. no 2.5 Output screen of Walking and punching video retrieval

VI. PROCESSING , INDEXING AND RETRIEVAL

The motion data pre-processing step consists of three main steps: body hierarchy construction that partitions the human body into a number of parts in the spatial domain, motion segmentation and normalization that segment part-based human motions and then group them into a set of basic motion prototypes (called motion patterns), which essentially partitions human motions in the temporal domain, and motion pattern extraction that detects and discovers patterns by grouping similar motion segments.
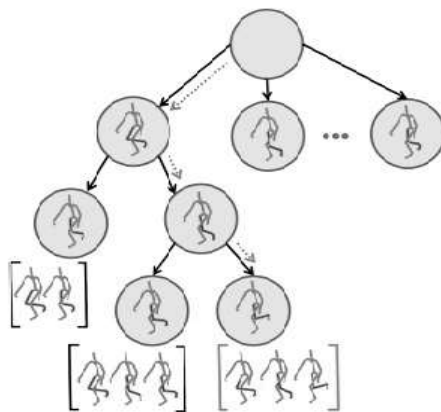
Fig. no. 3 database indexing

Fig no. 4 Figure motion data pre-processing

A hierarchical human structure illustrated in Fig is constructed based on the spatial connectivity of the human body. The whole human body is first divided into ten meaningful basic parts, e.g., head, torso, left arm, left hand, etc. and then a hierarchy with four layers is built accordingly. The hierarchy includes a total of eighteen nodes: ten leaf nodes stand for basic body parts, the parent nodes in the middle layers correspond to meaningful combinations of child nodes, and the root node represents the entire human body. We choose a human hierarchy representation, because it provides a logical control granularity. Furthermore, a multilayer hierarchy naturally take care of the correlations among several human parts that will be exploited for follow-up motion similarity computation. In this work, we use joint angles rather than 3D marker positions for representing human motion data, due to the fact that the joint angle representation is convenient for unifying the motions

Typically, raw data from motion capture systems are long motion sequences with high variances, e.g., tens of minutes. However, the motions of interest in many applications are shorter motion clips that satisfy users' specific requirements. Therefore, an automated motion segmentation process that adaptively chops the long motion into short clips is necessary for the later motion retrieval algorithm.
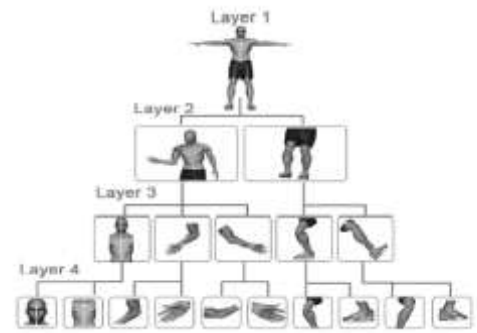

Figure no.5 hierarchy level construction of human body

VII. SYSTEM ALGORITHM

Input : Sketch with motion trajectory
Output : videos consisting of the same motion stroke given in input.

Pseudo-code :

1. Sketch selection
2. Sketch pre-processing
3. Acquire the drawn trajectory and save its co-ordinate vectors.
4. To find min. Distance for limbs
5. Body sectioning along with the trajectory
6. Determine the limb index for the trajectory.
7. Trajectory mapping onto body
8. Gradient calculations :
   Calculate the directional gradient of the trajectory along the co-ordinate axes.

   i) The rate of change in the spatial dimension is calculated along the axes.

   ii) Above value is converted into discrete form for the comparison purpose.

9. Gradient mapping and correlation(binarisation and Rate of change of gradient) i.e. The co-relation between the indexed gradient and the current trajectory is calculated.
10. Based on the match between the above two components, rank is assigned to the videos.
11. Video retrieval using indexes.

VIII. EXPERIMENT AND RESULTS

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 in the graph.

For the different motion strokes there will be change in no. of iterations and no. of pixels.

System is able to implement the five motions like punching, kicking, jumping, running and walking. Other different motions we can implement like dance moves, other sports activities. These new motions can be implemented by training the algorithm.

The database consists of 100 motion clips which are of 2 seconds to 10 seconds each. Database is taken from Mocap[11].

Result analysis using Precision and Recall

Based on the experiments performed to retrieve the videos
- **Precision**: fraction of retrieved docs that are relevant
  = P(relevant|retrieved)
- **Recall**: fraction of relevant docs that are retrieved
  = P(retrieved|relevant)

	Relevant	Non relevant
Retrieved	tp	Fp
Not Retrieved	fn	tn

Fig No. 6 (Table no. 1)

Precision P = tp/(tp + fp)
- Recall   R = tp/(tp + fn)

The following precision and recall is obtained on the different Motion strokes :

Motion	Precision(in %)	Recall(in %)
Jump	66.66	75
Walk	66.66	75
Kick	66.66	50
Punch	100	75
Run	66.66	60

Fig. no. 7 (table no. 2)

IX.   CONCLUSION AND DISCUSSION

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 report 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 motion trajectory of walking, jumping, kicking, punching, etc. in the respective limb such as punching motion stroke would be provided at the hands or kicking motion stroke would be provided in response, it should get the motion clips in which these motions are present. So the proposed system is able to retrieve the five king of motions proposed with around 60-70 % accuracy. For all actions precision and recall results are good.

REFERENCES

[1] M. S. Lew, N. Sebe, C. Djeraba, And R. Jain. Contentbased Multimedia Information Retrieval: State Of The Art And Challenges. Acm Transactions On Multimedia Computing,Communications, And Applications, Vol. 2, Pp. 1-19, 2006.

[2] Marchand-Maillet, S.: Content-Based Video Retrieval: An Overview. Technical Report 00.06, Cui - University Of Geneva, Geneva (2000)

[3] Taskiran, C., Chen, J.Y., Albiol, A., Torres, L., Bouman, C., Delp, E.: Vibe: A Compressed Video Database Structured For Active Browsing And Search. Ieee Transactions On Multimedia 6, 103–118 (2004)

[4] 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)

[5] A System For Analyzing And Indexing Human-Motion Databases.In Proceedings Of The 2005 ACM SIGMOD International Conference On Management Of Data, ACM, New York, NY, USA,SIGMOD '05, 924–926.

[6] Efficient Motion Retrieval In Large Motion Databases Mubbasir Kapadia_ I-Kao Chiangy Tiju Thomasz Norman I. Badlerx Joseph T Kider Jr University Of Pennsylvania

[7] M. Thorne And D. Burke, "Motion Doodles: An Interface For Sketching Character Motion," ACM Trans. Graphics, Vol. 23, No. 3, Pp. 424-431, 2004.

[8] Y.-Y. Tsai, W.-C. Lin, K.-B. Cheng, J. Lee, And T.-Y. Lee, "Real- Time Physics-Based Third Biped Character Animation Using An Inverted Pendulum Model," IEEE Trans. Visualization And Computer Graphics, Vol. 16, No. 2, Pp. 325-337, Mar. 2010.

[9] CHIENG FENG T., GUNAWARDANE P., DAVIS J., JIANG B.: Motion Capture Data Retrieval Using An Artist's Doll. In Pattern Recognition, 2008. ICPR 2008. 19th International Conference(2008), Pp. 1–4

[10] DAVIS J., AGRAWALA M., CHUANG E., POPOVI´C Z., SALESIN D.: A Sketching Interface For Articulated Figure Animation. In Proceedings Of The ACM SIGGRAPH/Eurographics Symposium On Computer Animation (2003), Eurographics Association, Pp. 320–328

[11] CMU Graphics Lab Motion Capture Database (http://mocap.cs.cmu.edu/).

[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