Human Motion Simulation and Action Corpus

Gang Zheng[1], Wanqing Li[1], Philip Ogunbona[1],
Liju Dong[1,2], and Igor Kharitonenko[1]

[1] School of Computer Science and Software Engineering, University of Wollongong, Australia
[2] College of Information Engineering, Shenyang University, P.R. of China
`{gz207,wanqing,philipo,liju,igor}@uow.edu.au`

Abstract. Acquisition of large scale good quality training samples is becoming a major issue in machine learning based human motion analysis. This paper presents a method to simulate continuous gross human body motion with the intention to establish a human motion corpus for learning and recognition. The simulation is achieved by a temporal-spatial-temporal decomposition of human motion into actions, joint actions and actionlets based on the human kinematic model. The actionlet models the primitive moving phase of a joint and represents the muscle movement governed by kinesiological principles. Joint actions and body actions are constructed from actionlets through constrained concatenation and synchronization. Methods for concatenation and synchronization are proposed in this paper. An action corpus with small number of action vocabularies is created to verify the feasibility of the proposed method.

Keywords: Human Motion, Actions, Simulation, Motion Editing.

1 Introduction

Automatic recognition of human motion has been one of the most active research topics in computer vision in the past decades [1,2]. Recent research [1,2,3,4,5] has demonstrated that machine learning approach has the potential to lead to a generic solution. However, the success of machine learning approach hinges on the quality and amount of training samples. Although limited individual action samples maybe obtained using motion capture devices [6,7,8,9,10,11], acquisition of large amount of training samples remains challenging considering the cost and amount of work required to obtaining such training samples. In this paper, we propose to generate large amount of quality motion data through simulation for the purpose of learning and recognition.

Simulation of human motion has been previously studied mainly from the perspective of graphical animation [6,7,8] and qualitative and quantitative biomechanical studies [12,13] with a focus on individual actions and small part of the body. For instance, the commercial software *Anybody* [14] and *Jack Software* [15] are used for biomechanical and ergonomic study and design of products. The software enables users to position biomechanically accurate digital humans of

V.G. Duffy (Ed.): Digital Human Modeling, HCII 2007, LNCS 4561, pp. 314–322, 2007.

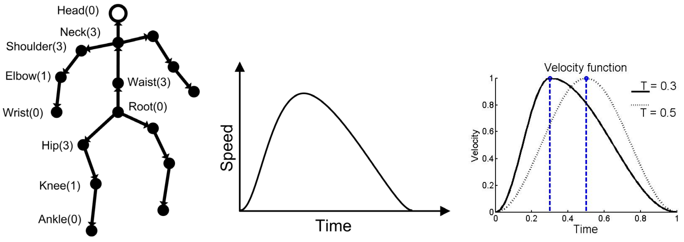
various sizes in virtual environments, assign them tasks and analyze their performance. Both softwares employ complex kinetics analysis and empirical models [16,17,18] and are able to simulate individual actions in a natural and realistic manner. However, they are not designed for generating large amount of motion data, especially the variations of individual actions.

This paper presents a method to simulate gross human body motion with the intention to establish a human motion corpus for learning and recognition. We propose to simulate human motion by hierarchically decomposing it into actions, joint actions and primitive actions, referred to as *actionlets*. The actionlet models the primitive moving phase of a joint and represents the muscle movement governed by kinesiological principles. Joint actions and body actions are constructed from actionlets through constrained concatenation and synchronization.

The rest of the paper is organized as follow. Section 2 presents the kinematic model and the hierarchical decomposition of human motion into the primitive movement, actionlets. Section 3 describes how a complex human motion can be generally simulated through hierarchical concatenation and synchronization of the actionlets. An approach to establish a motion corpus based on the proposed simulation scheme is given in Section 4. Section 5 demonstrates some simulation reults and conluding remarks are made in Section 6.

2 Kinematic Model and Motion Decomposition

Human body is often viewed as an articulated system of rigid links or segments connected by joints. Fig. 1(a) shows a kinematic model with 15-joint, 22-DOFs (Degree Of Freedom). The model is tailored to represent the postures of human body, which is descriptive enough for gross human motion. A detailed model may be adopted with more DOFs and segments such as hands and feet.



(a) (b) (c)

Fig. 1. The adopted kinematic model and velocity curves. (a) Kinematic model and main joints, (b) A typical velocity curve based on kinesiological study [19], (c) Two simulated velocity curves using a concatenation of two half cycle *sine* function (Eq. 1).

Human motion can be considered as a continuous change of his/her body posture and, therefore, represented as a sequence of temporally sampled postures. A posture at a given time is decided by the spatial configuration of the segments of the human body. The co-existance of spatial and temporal nature has made human motion much more complicated to describe and analyse than speech. In [20], we proposed to decompose *human motion* into a series of temporally concatenated actions (temporal decomposition). Each action is formed by a set of coordinated *joint actions* (spatial decomposition). A joint action is further decomposed into a sequence/sequences of *actionlets*(temporal decomposition), where each actionlet represents a cycle of acceleration and deceleration of a joint according to kinesiological study [19,21]. Fig. 1(b) shows a typical velocity curve [19] of an actionlet. This curve may be emulated by the following function

$$v(t) = \begin{cases} \frac{1}{2}(1 + \sin(\frac{\pi}{T}(t - \frac{T}{2}))); & 0 \le t \le T \\ \frac{1}{2}(1 + \sin(\frac{\pi}{1-T}(t - (T - \frac{1-T}{2})))); & T \le t \le 1, \end{cases} \tag{1}$$

where T is a normalized variable that controls the shape of the function. Fig. 1(c) shows the emulated curves with different $T = 0.3$ and $T = 0.5$ respectively.

In order to represent the posture in a way that is independent of the body size [20], Euler angle is adopted in our research to represent the spatial configuration of the joints of the kenematic model. Let $\boldsymbol{Y} = (y_0, ..., y_{21})$, where $y_0, ..., y_{21}$ are all Euler angles corresponding to 22 DOFs. An action that happens between the period, $t \in [T_{begin}, T_{end}]$, can be represented as

$$\boldsymbol{Y(t)} = (y_0(t), ..., y_{21}(t)); t \in [T_{begin},\ T_{end}]. \tag{2}$$

where $y_i(t), t \in [T_{begin}, T_{end}]$ represents the time series of the Euler angles of the $i'th$ DOF. $y_i(t)$ may consists of N_i cycle of accelarion and deceleration or silence (static periods). In other words, $y_i(t)$ is a concatenation of N_i actionlets. Therefore,

$$Y(t) = \begin{cases} concat(f_{[0,0]}(\tau_{0,0}), f_{[0,1]}(\tau_{0,1}), \cdots, f_{[0,N_0-1]}(\tau_{0,N_0-1})) \\ \vdots \\ concat(f_{[21,0]}(\tau_{21,0}), f_{[21,1]}(\tau_{21,1}), \cdots, f_{[21,N_{21}-1]}(\tau_{21,N_{21}-1})) \end{cases} \tag{3}$$

where $concat(\cdot)$ represents concatenation, and N_i represents the number of actionlet in the $i'th$ DOF. $f_{[i,j]}(.)$ is the Euler angle series of the $j'th$ actionlets in the $i'th$ DOF, which can be specified by users or obtained from motion capture data. $\tau_{i,j}$ is the time variable corresponding to each $f_{[i,j]}(.)$. The velocity, temporal changing rate of the Eulaer angle, of any actionlet follows the basic principle of muscle movement as shown in Fig. 1(b,c).

Let $v(t)$ be a velocity function of any $y_i(t)$ that consists of one actionlet. The condition of actionlet is

$$v(t_{begin}) = v(t_{end}) = 0 \tag{4}$$

$$v(t) \neq 0, t_{begin} \le t < t_{end}, \tag{5}$$

where t_{end} is the exact next zero velocity point to the t_{begin}, so that any kind of action concatenation can be abstracted into concatenation between actionlets.

According to [22], any muscle acts in cycles of contracting, keeping still, strectching and keeping still. In the angular velocity space, it means a cycle of acceleration and deceleration. Soderberg [19] gave a typical angular velocity curve cooresponding to one joint action (Fig. 1(b)) that can be approximated by half cycle of *sine* function as shown in Fig. 1(c).

3 Simulation

With the proposed decomposition, simulation of human motion amounts to the composition of a sequence of actions using actionlets. Any actionlet has to satisfy all possible physical constraints which the corresponding joint may be subject to. The two major composition operations that are required to build a specific type of actions (e.g. running) from actionlets are *concatenation* and *synchronization*. The reality of the simulated motion depends on how well these operations perform.

Concatenation is employed to build a joint action from actionlets or to form a motion from a sequence of actions. We apply the continuity of the Euler angles and their first order derivatives (velocity) to all DOFs. When two actionlets or actions do not meet the concatenation criteria, either the two actionlets or actions will be blended or a transition actionlet or action is inserted between them.

Considering two actionlet sequences SA and TA, where SA has two actionlets $sa1$ and $sa2$, TA has one actoinlet $ta1$. Suppose $[t_{sa1}, t_{sa2}]$ is the time scope of $sa1$, $[t_{sa2}, t_{sa3}]$ is the time scope of $sa2$, and $[t_{ta1}, t_{ta2}]$ is the time scope of $ta1$. The concatenation between SA and TA can be generally represented as an insert operation, i.e. to 'cut' a piece of action $[t_{begin}, t_{end}]$ from source action SA and 'paste' it into target action TA at the time t_{target}, where $t_{sa1} < t_{begin} < t_{sa2} < t_{end} < t_{sa3}$ and $t_{ta1} < t_{target} < t_{ta2}$. The result action has to be maintained smooth.

Suppose v_a is the velocity function corresponding to each actionlet, a. The process of the insertion operation is summaried in table 1 and table 2, where the target actionlet is changed to two actionlets $ta1'$ and $ta2'$, and

$$t'_{ta1} = t_{ta1},$$

$$t'_{ta2} = t_{target} + (t_{sa2} - t_{begin}),$$

$$t'_{ta3} = t_{ta2} + (t_{end} - t_{begin}).$$

Here, the source actionlet is 'pasted' into the target actionlet. Similar process is employed when there are more than one actionlet in the time scope $[t_{begin}, t_{end}]$. This algorithm can maintain the smoothness of target's anguler displacement because the velocity values at the ends of target actionlets keep zero.

Synchronization is required when an action involves a number of joints. The joints have to act in a coordinated manner in order for the action to be realistic and meaningful. A typical example is "walking" where the arms and legs and the two joints (keen and hip) on the legs have to be synchronised properly.

Table 1. The source and target actionlets before concatenation

Actionlet	Begin time	End time	Velocity function
$sa1$	t_{sa1}	t_{sa2}	v_{sa1}
$sa2$	t_{sa2}	t_{sa3}	v_{sa2}
$ta1$	t_{ta1}	t_{ta2}	v_{ta1}

Table 2. The source and target actionlets after concatenation

Actionlet	Begin time	End time	Velocity function
$ta1'$	t'_{ta1}	t'_{ta2}	v_{ta1}
$ta2'$	t'_{ta2}	t'_{ta3}	v_{sa2}

Let $\wp_g(t)$ be a joint action group, then

$$\wp_g(t) \equiv \{y_k(t)\,|k \in \Im\} \tag{6}$$

where $\Im$ is a set of valid joints, and $y_k(t)$ corresponding to the $k'th$ DOF showed in Eq. 2.

Assume $\wp_p(t)$ and $\wp_q(t)$ are two groups of joint actions, we consider both $\wp_p(t)$ and $\wp_q$ as periodic functions, the synchronization between them can be represented as

$$\wp_q(t) = \wp_p(sh * t + \phi) \tag{7}$$

where sh denotes the temporal scale coefficient and ϕ is a phase variable.

4 Action Corpus

The major objective of the proposed simulation is to provide an easy and low-cost scheme to generate enough quality motion data, action corpus, for learning and recognition. With the hierarical decomposition of the human motion into the primitive actionlets, the process to establish such an action corpus becomes to define a set of resuable actionlets and construct actions and sequence of actions from the actionlets. This is a reverse process of decomposition as illustrated in Fig. 4. To make the construcion process easy, we introduce a number of intermediate reusable components: *joint actions*, *limb actions* and *joint group action.*

1. *Joint Action.* In the kinemati model as showed in Fig. 1(a), joints may have more than 1 DOF. *Joint action* is defined as a group of coupled DOF functions $y_i(t)$ within a specified time scope for a particular joint. For example, the joint action on the left shoulder joint includes $y_3(t)$, $y_4(t)$ and $y_5(t)$ and its dimentionality is 3.

2. *Limb Action.* We define a *limb action* as a group of two or more phisically connected joint actions in a specified time scope. For example, the left arm includes left elbow and left shoulder and the limb action of the left arm includes $y_3(t)$, $y_4(t)$, $y_5(t)$ and $y_7(t)$, where $y_7(t)$ corresponds to the left elbow's DOF function. This limb action is 4 dimension.
3. *Joint Group Action.* A joint group action is defined as a group of any coombination of joint actions and limb actions. In other words, it could be any combination of DOF functions within $y_i(t)$, where $0 \leq i \leq 21$, for there are 22 DOF functions.

The definition of *joint actions*, *limb actions* and *joint group actions* connects the real world's description of human actions and the underliying mathematical approximation models, the action models and the actionlet models. They provide us with a set of flexible and resuable *action units* to describe and simulate a variety of human motions. With a small number of resualble action units together with the concatenation and synchrononization methods described above, virtually, unlimited types of human motion can be simulated.

Due to the demand for large number of samples of a particular action for learning, we introduce a *mutation* process in the simulation to generate variations of a particular action. The mutation process eliminates the need for specifying an action for each sample and allows us to easily generate samples with variations for the same action. There are two ways to mutate the samples. One is to choose same type of actionlets but with different velocity functions from the actionlet database. Another is to perturb the parameters of the velocity functions as shown in Fig 1(c).

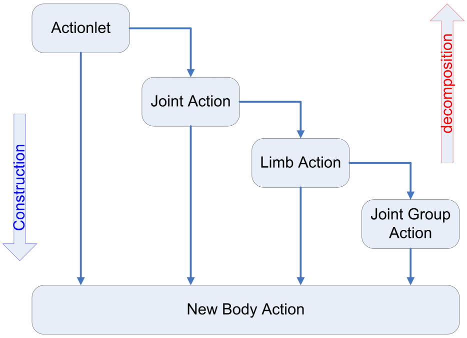


Fig. 2. Decomposition and Construction

5 Results and Visualization

A simulation system has been implemenetd with a number of essential motion editing functions and a set of parametrized actionlets. Humanoid animation (H-Anim), part of W3D/VRML ISO standard, is adoped for visualization to

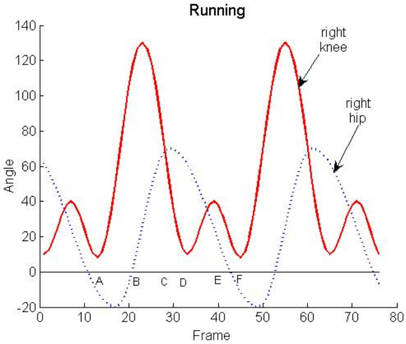


Fig. 3. Angle curve of two joints for running

check the reality of the simulated actions. Fig. 4 shows a simulated 'running' action. Fig. 5 shows the angle changes of the two joints: right knee and right hip. Corresponding to the six indicated positions, six static postures were cut from a 3-D browser as shown in Fig. 4.

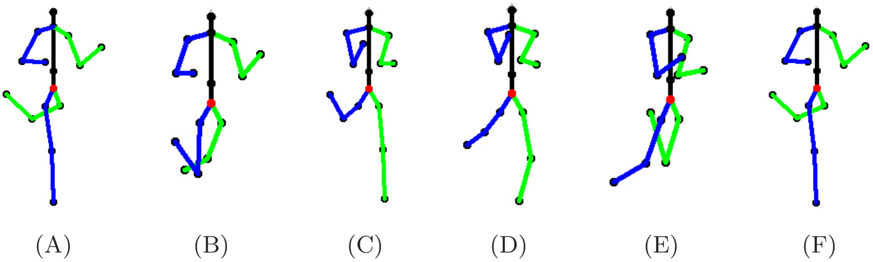
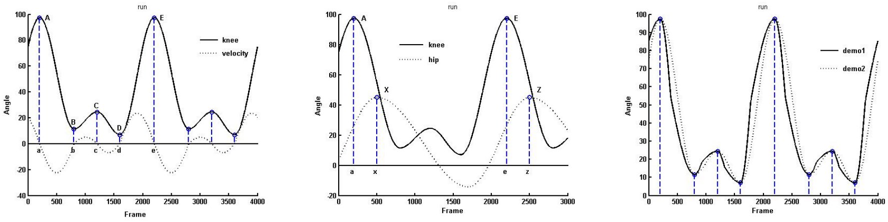


Fig. 4. Running action. A, B, C, D, E, and F corresponding to the six positions marked in Fig. 5

Fig. 5(a) shows one *joint action* (knee) while running at the temporal resolution of 3000 frames per second. The action was composed of two types of actionlets AB (or CD) and BC (or DE). Fig. 5(b) shows the synchronisation between joint "knee" and "hip" for running. *Actionlets* AB, BC, CD, DE construct the *joint action* AE for "knee" and *joint action* XZ for hip. There is a phase $|A - X|$ difference between these two *joint actions*, but the two *joint actions* have the same period, that is $|E - A| = |Z - X|$. All *action unit* mentioned above are stored in one database that can be reused to generate other actions. After the mutation process, a large variety of the 'running' action were generated. Fig. 5(c) shows the mutation of the two joints in action 'running'.



(a) (b) (c)

Fig. 5. Simulation of action ”running”. (a) Two types of actionlets for joint ”knee” (AB and BC). (b) Synchronization of ”knee”’ and ”hip” (c) Mutation result of ”knee”, actionlets were modelled by Eq. 1 with $T = 0.3$ (solid line) and $T = 0.5$ (dotted line).

6 Conclusion

In this paper, we improved the action model proposed in [20], and resolved the concatenation issue in the simulation process. Using the method, two or more actions can be concatenated together without complicated kinematic analysis, and keep good nature-looking. The system is based on novel scheme to decompose human motion into actions, joint actions and actionlets. This decomposition makes feasible a large scale simulation of human body motion in which not only a virtually unlimited types of actions can be simulated from a finite set of actionlets, but also an unlimited variations of one action can be generated.

References

1. Moeslund, T.B., Granum, E.: A survey of computer vision-based human motion capture. Computer Vision and Image Understanding 81, 231–268 (2001)
2. Moeslund, T.B., Hilton, A., Kruger, V.: A survey of advanced in vision-based human motion capture and analysis. Computer Vision and Image Understanding 104, 90–126 (2006)
3. Hu, W., Tan, T., Wang, L., Maybank, S.: A survey on visual surveillance of object motion and behaviors. IEEE Transactions on Systems, Man. and Cybernetics, Part. C 34(3), 334–352 (2004)
4. Zhang, X., Liu, Y., Huang, T.S.: Motion analysis of articulated objects from monocular images. IEEE Transactions on Pattern Analysis and Machine Intelligence 28(4), 625–636 (2006)
5. Green, R.D., Guan, L.: Quantifying and recognizing human movement patterns from monocular video images–Part I: A new framework for modelling human motion. IEEE Trans Circuits and Systems for Video Technology 14(2), 179–190 (2004)
6. Boulic, R., Magnenat-Thalmann, N., Thalmann, D.: A global human walking model with real-time kinematic personification. International Journal of Computer Graphics 6(6), 344–358 (1990)
7. Fang, A.C., Pollard, N.S.: Efficient synthesis of physically valid human motion. Proceedings of ACM Transactions on Graphics 22(3), 417–426 (2003)

8. Xiao, Z., Yang, X., Zhang, J.J.: Motion data correction and extrapolation using physical constraints. IEEE Proceedings of the 9th International Conference on Information Visualisation (IV 05) 00, 571–576 (2005)
9. Faraway, J.: Human animation using nonparametric regression. Department of Statistics University of Michigan, Tech. Rep. 382 (2002)
10. Ijspeert, A., Nakanishi, J., Schaal, S.: Movement imitation with nonlinear dynamical systems in humanoid robots. In: Proceedings of the IEEE International Conference on Robotics and Automation, Washington, DC, May 2002, vol. 2 , pp. 1398–1403 (2002)
11. Pettre, J., Simeon, T., Laumond, J.: Planning human walk in virtual environments. IEEE / RSJ International Conference on Intelligent Robots and Systems (IROS'02) vol. 3, pp. 3048–3053 (September 2002)
12. Chang, C.-C., Brownb, D.R., Bloswickb, D.S., Hsiang, S.M.: Biomechanical simulation of manual lifting using spacetime optimization. Journal of Biomechanics 34(4), 527–532 (2001)
13. Park, W., Chaffin, D.B., Martin, B.J., Faraway, J.J.: A computer algorithm for representing spatial-temporal structure of human motion and a motion generalization method. Journal of Biomechanics 33, 2321–2329 (2005)
14. [Online]. Available: `http://www.ugs.com/products/tecnomatix/human_performance/jack/`
15. [Online] Available: `http://www.anybodytech.com`
16. Wrotek, P., Jenkins, O.C., McGuire, M.: Dynamo: dynamic, data-driven character control with adjustable balance. In: sandbox '06: Proceedings of the 2006 ACM SIGGRAPH symposium on Videogames, pp. 61–70. ACM Press, New York, USA (2006)
17. Kry, P.G., Pai, D.K.: Interaction capture and synthesis. In: SIGGRAPH '06: ACM SIGGRAPH 2006 Papers, pp. 872–880. ACM Press, New York, USA (2006)
18. Xia, S., Qiu, X., Wang, Z.: A novel framework for athlete training based on interactive motion editing and silhouette analysis. In: VRST '05: Proceedings of the ACM symposium on Virtual reality software and technology, pp. 56–58. ACM Press, New York, USA (2005)
19. Soderberg, G.L.: Kinesiology application to pathological motion, 1st edn. Williams & Wilkins (1986)
20. Zheng, G., Li, W., Ogunbona, P., Dong, L., Kharitonenko, I.: Simulation of human motion for learning and recognition. In: AI 2006. LNCS (LNAI), vol. 4304, pp. 1168–1172. Springer, Heidelberg (2006)
21. Tözeren, A.: Human Body Dynamics. Springer, Heidelberg (1999)
22. Zatsiorsky, V.M.: Kinematics of Human Motion, 1st edn. Human Kinetics (1998)