Journal for Geometry and Graphics
Volume 8 (2004), No. 1, 81–90.

Keyframes Extraction Method for Motion Capture Data

Koichi Matsuda[1], Kunio Kondo[2]

[1] *Iwate Prefectural University*
152-52, Sugo, Takizawa-mura, Iwate-ken, Japan
email: matsuda@soft.iwate-pu.ac.jp

[2] *Saitama University*
255, Shimo-ohkubo, Sakura-ku, Saitama-ken, Japan
email: kondo@ke.ics.saitama-u.ac.jp

Abstract. In this paper, we propose a keyframes extraction method of motion capture data for reducing data quantity. To choose keyframes from motion capture data, we apply a refinement technique for handwritten curves to the data.

As a result of comparing original data with reduced data, it is confirmed that the path of each data is nearly equal. Features of our method are: (1) data reduction that utilizes characteristics of the measured motion is possible, and (2) the user can control the ratio of reduction. It became possible that the data quantity is kept to about 50%, while features of the original motion are retained.

Key Words: computer animation, motion capture data, key-frame

MSC 2000: 68U05

1. Introduction

Animation techniques have been utilized in various fields because animation has advantages in general. The advantages are that it can express various pictures that can't be expressed in the photographic pictures. Therefore, computer animation is of great demand in the field of picture production because of the progress of computer graphics and the realization of labor saving in a production work with a computer. Especially, animation is being employed in movies, TV programs, commercials and other entertainment works.

In the field of 3D animation, data of actual movements are taken into a computer with input device such as a motion capture system. The volume of the data is a large quantity. Many researches, therefore, propose physics-based techniques that generate realistic movement [6, 1, 8]. On the other hand, knowledge and techniques based on the experience of animators are necessary to make emphasized expressions. There is a technique of movement

ISSN 1433-8157/$ 2.50

generation of characters such as human and animal, a process that produces computer animations [2, 3, 7, 5]. For emphasizing the motion of characters by using motion capture data, users should control a large quantity of frame data. So, only keyframes should be shown to the users for reducing their labor.

The purpose of this study is

(1) to reduce lengthy information by extracting keyframes that show features of the movement from motion capture data, and

(2) to make handling the data easy for users.

In this paper, we propose a keyframes extraction method of the data by using a refinement method for handwritten curves. Our method has three processes for reducing the data:

- First, the *segmentation* of a motion graph that is the angle sequence on each axis (x, y, z).
- Second, finding *characteristic frames* from curve sections of the graph.
- And third, *overlapping* each reduced motion graph.

We evaluate the proposed method by applying it to many motion capture data.

As a result of comparing original data with reduced data, it is confirmed that the path of each data is nearly equal. Features of our method are

(1) data reduction that utilizes characteristics of the measured motion, and

(2) users can control the ratio of reduction.

Thus it became possible that the data volume is kept to about 50%, while features of the original motion are retained.

2. Keyframes extraction method

In this section we explain a technique to extract keyframes from motion data by finding feature frames of the data.

Motion data is recorded as rotation angles of each joint in time series, and a joint has angles with each axis, x-axis, y-axis and z-axis. These data can be drawn as a graph of the angle and the time. We call this graph *"motion graph"* (see Fig. 1).

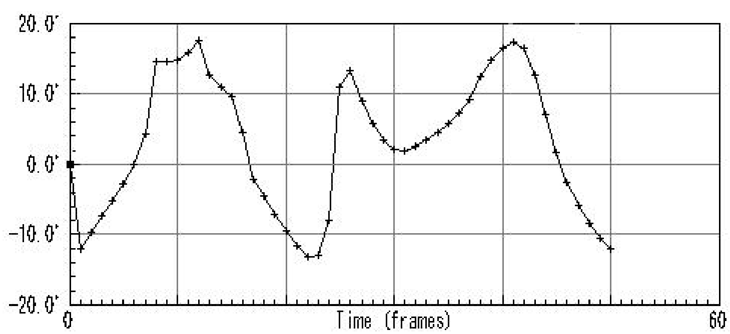

Figure 1: An example of a motion graph.

After figuring out the keyframes on a motion graph, a minimum set of frames can be left. To find keyframes of the motion graph, we apply a technique that enables the user to get clean copies of handwritten curves; this is called *"interactive sequential drawing method" (ISD)* [4]. When we modify ISD for motion graphs, the motion data is reduced automatically.

2.1. Interactive sequential drawing

ISD can retrieve control points of a natural spline curve from handwritten curves. So, necessary points that become characteristics from a motion graph can be extracted by using our technique.

ISD can remove hand deflection or distortion from handwritten curves, select control points for the curves from a local and global point of view. As a result, the user can get clean copies of handwritten curves.

The technique consist of two algorithms.

- *Connection* (Fig. 2(1))
  Some handwritten curves consisting of some segments are changed into one curve.
- *Refinement* (Fig. 2(2))
  Deflection or distortion points are removed.

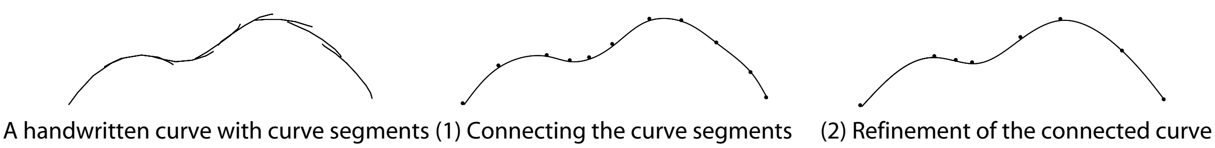

Figure 2: Processing a handwritten curve

2.2. Segmentation of a motion graph

When there is a section that has only a little angle differ on a motion graph, this section is replaced by a horizontal straight line. It needs to divide a motion graph into curved and straight line sections, because ISD is a curve refinement method.

A straight line section is recognized when the amplitude width of a certain section is under 1/100 of the maximum amplitude width (see Fig. 3).

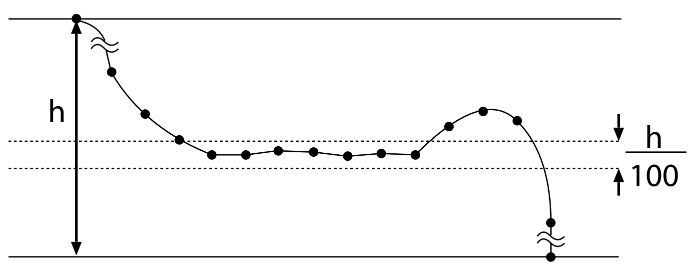

Figure 3: Definition of a straight line section

Motion graphs are interpolated with a natural spline curve; so for a straight line section one must choose two points at each endpoint. If only one point is chosen at each end of the section, the section will be interpolated not by a straight line but by a curve (see Fig. 4).

2.3. Reduction of keyframes

Keyframes of motion capture data are defined by all keyframes on each motion graph. If a frame is required at some graphs, the frame is adopted as a keyframe. The keyframes in the above method are called *"full-keyframes"*.

Fig. 5 shows the definition of full-keyframes. In Fig. 5(b), the numbers of keyframes for Heading, Pitch, Bank are 4, 4, 7, respectively. In the worst case in Fig. 5, 11 keyframes are

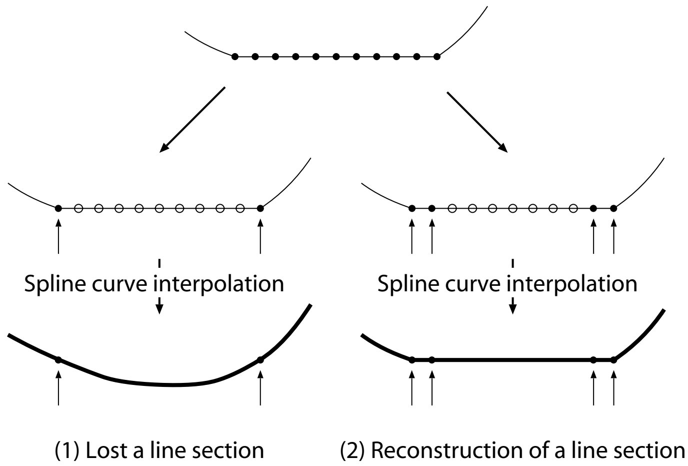

Figure 4: Keeping a straight line section

necessary to reconstruct the motion graphs of three axes completely, because these graphs change independently. On the worst case in Fig. 5 the frame number of "full-keyframes" is about three times of the number of Heading's keyframes. So, we suppose a reduction method for "full-keyframes":

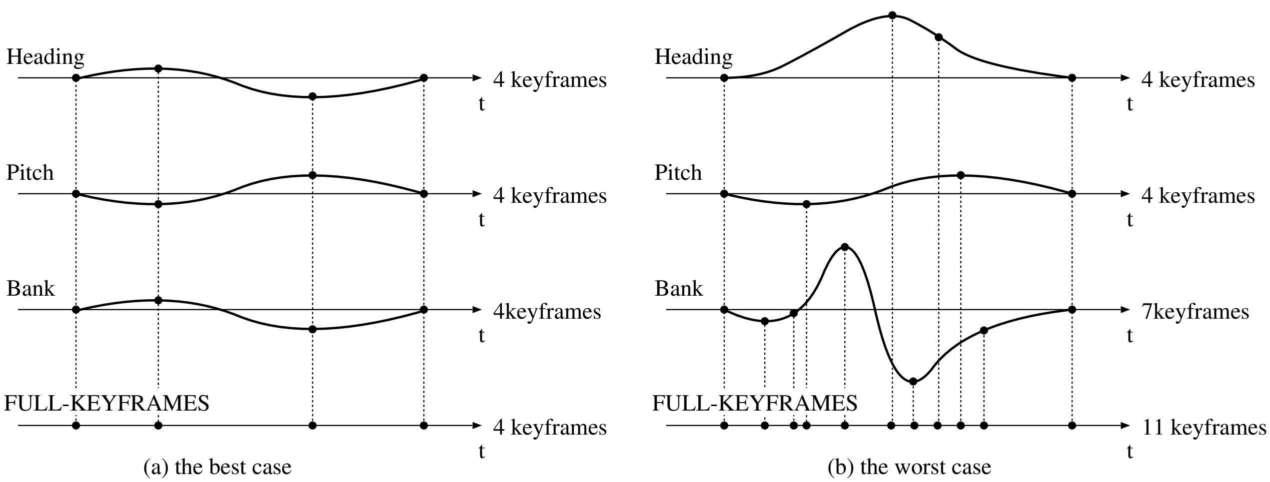

Figure 5: Definition of *full-keyframes*

First, a frame that satisfies one of the following conditions is defined a *"fixed frame"* (see Fig. 6). Fixed frames are defined as points which can't be deleted.

(1) A point that has a local maximal or minimum value;

(2) Both ends point of a straight line section;

(3) A point that has "large angle difference" on both sides, and the both sides' points. Here a large angle difference means that the difference exceeds 50% of the maximum amplitude width in the motion graph.

Second, a frame that satisfies one of the following conditions and is not a fixed frame, is defined a *"reduction frame"* that can be removed.

(1) When there is no frame data between characteristic frames, the frames are replaced by a new characteristic frame that is set at the center of the frames (see Fig. 7(1)).

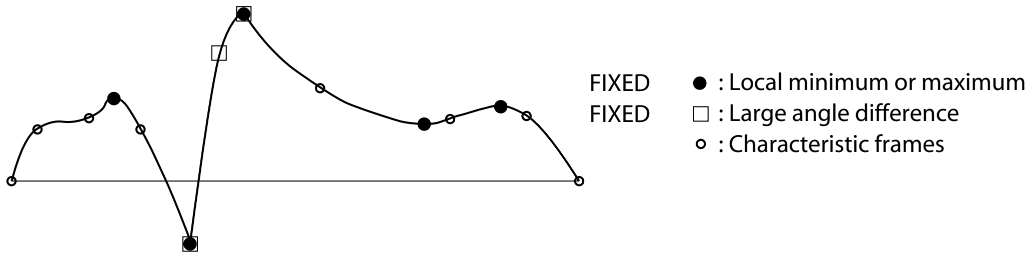

Figure 6: *Fixed* frames

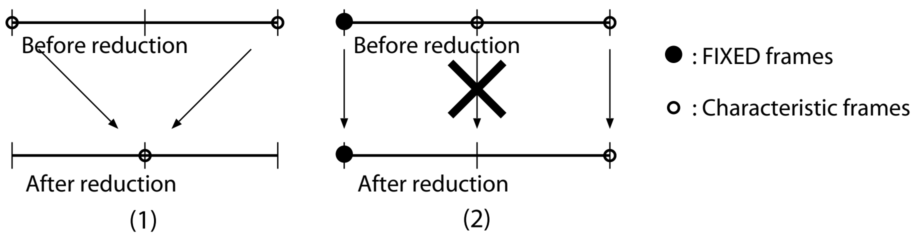

Figure 7: *Reduction* frames

(2) If three frame data have continued and the central point is not a fixed frame, the central point is removed (see Fig. 7(2)).

3. Experiment

We show examples of motion pictures using keyframes with the proposed method. The 3D human body used on this experiment is divided into 18 parts. The motion pictures are created on NewTek LightWave3D (LW). We programmed our method for the motion capture data file of LW on UNIX. All the setups of LW are used with default, and threshold on ISD is 0.13 that became experimentally good value.

The Figures 8–9 show the original motion data ("both hands thrust" and "Swing 1") that are rendered with LW.

The Tables 1–2 show results of keyframes extraction. The meaning of items in the tables is as follows:

Parts means the name of the 3D human body.

F is the number of frames in original motion data.

H/P/B Heading / Pitch / Bank. These are the angles of the x-, y-, and z-axis.

Lap The number of keyframes by overlapping three motion graphs.

Del The number of keyframes after reduction.

Ratio "Del : Lap $\times$ 100".

The Tables 1–2 show that the keyframe extraction of motion data from an axis can decrease the number of frames to 50%–70%, but "Overlapping" can decrease the frames about 30%. So, by applying the "reduction" process to the keyframes after the "Overlapping" process, the frames finally decrease about 50%.

The Figs. 10–12 show motion graphs that were created by the proposed technique on LW. These graphs concern the left foot of Table 1 and the right shoulder of Table 2. At each figure, left graph is an original one and the right graph is obtained after "reduction".

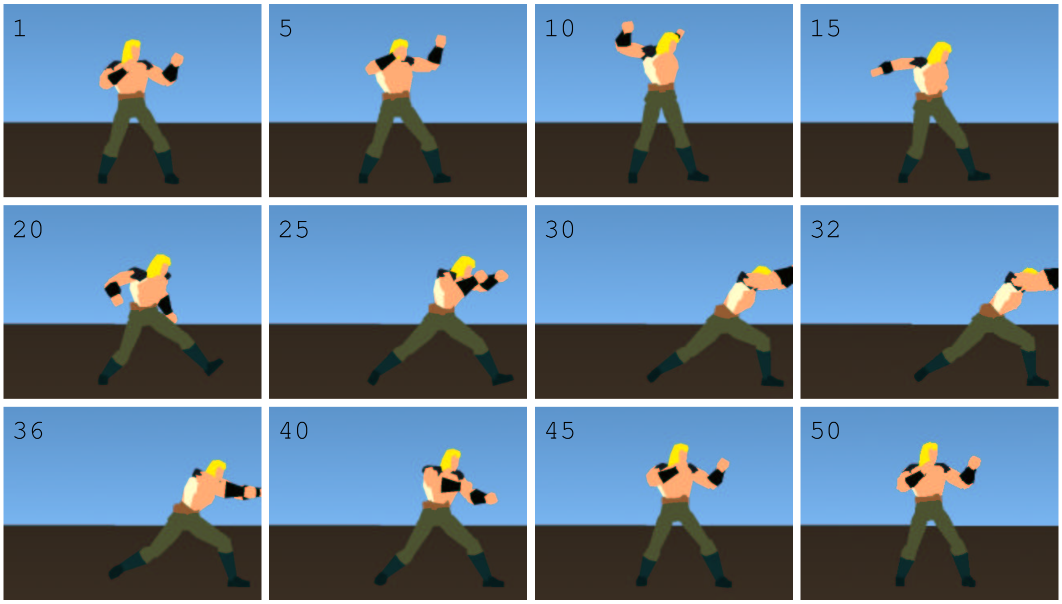

Figure 8: Both hands thrust: 51 frames

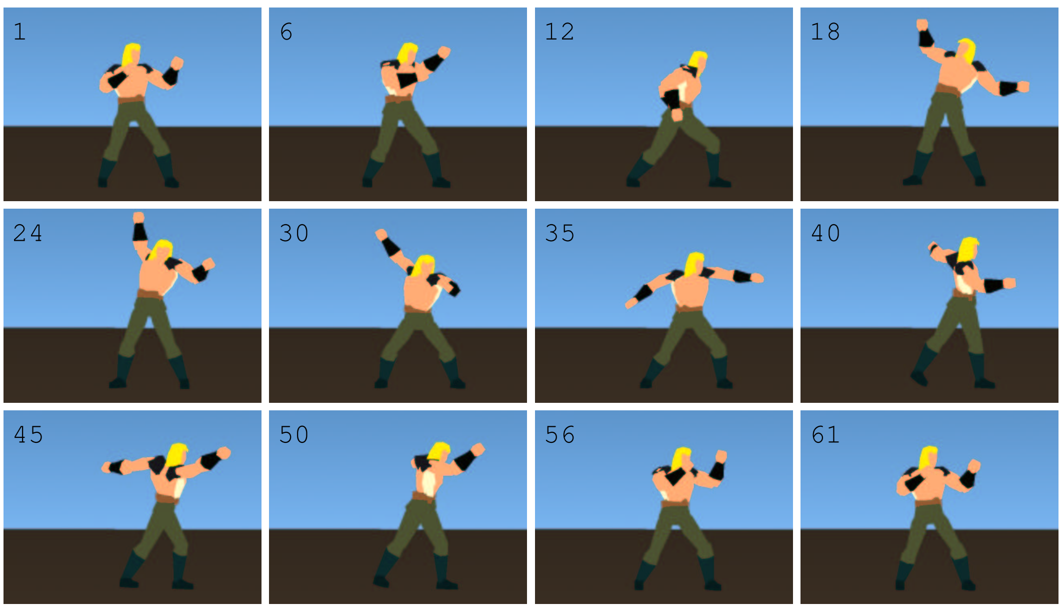

Figure 9: Swing 1: 62 frames

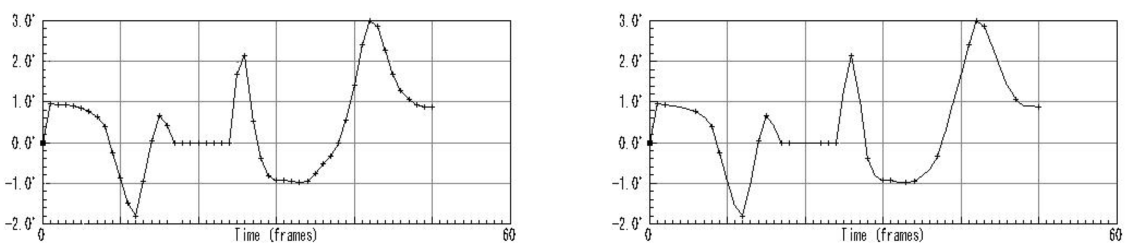

Figure 10: Left foot (Table 1) frame data (Heading)

Table 1: Reduction example (Fig. 8)

Parts	*F*	*H/P/B*	*Lap*	*Del*	*Ratio*
hip	51	15/ 9/20	33	25	51.0%
chest	51	21/18/21	41	26	49.0%
neck	51	21/22/21	40	28	45.1%
head	51	19/27/21	38	29	45.1%
shoulder (R)	51	24/13/16	35	21	58.8%
upper arm (R)	51	18/20/20	34	29	43.1%
fore arm (R)	51	23/13/29	36	25	51.0%
hand (R)	25	9/11/ 9	12	10	60.0%
shoulder (L)	51	24/32/16	44	31	39.2%
upper arm (L)	51	20/25/22	39	28	45.1%
fore arm (L)	51	23/23/19	37	27	47.1%
hand (L)	25	10/12/10	14	13	48.0%
thigh (R)	51	32/20/13	39	28	45.1%
calf (R)	51	18/23/14	33	27	47.1%
foot (R)	51	32/20/20	41	30	41.2%
thigh (L)	51	26/16/16	35	26	49.0%
calf (L)	51	17/21/20	36	25	51.0%
foot (L)	51	25/19/16	37	26	49.0%
Average	\	20.9/19.1/17.9	34.7	25.2	47.6%

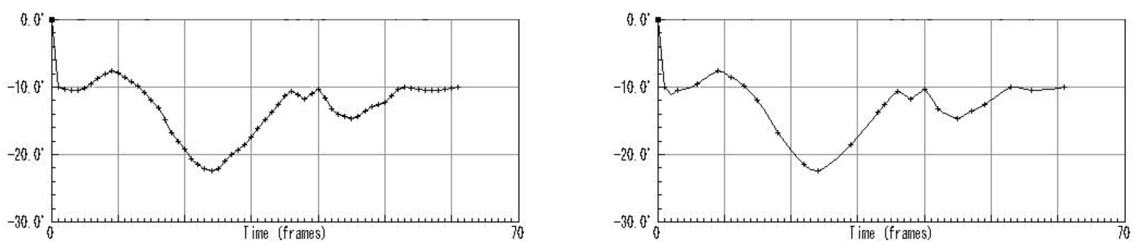

Figure 11: Right shoulder (Table 2) frame data (Pitch)

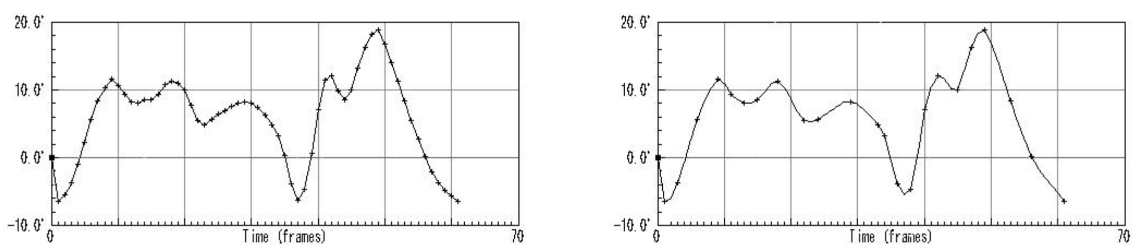

Figure 12: Right shoulder (Table 2) frame data (Bank)

The Figures 13–14 show a motion path that is part of a 3D human model. Fig. 13 shows that a similar path can be drawn with keyframes that decrease 60% frames from the original motion data, and Fig. 14 shows that we can get a similar path with the proposed method.

About 45%–50% were reduced as a result of applying our method to 18 kinds of motion data (see Table 3).

Table 2: Reduction example (Fig. 9)

PARTS	*F*	*H/P/B*	*Lap*	*Del*	*Ratio*
hip	62	13/23/24	38	27	56.5%
chest	62	23/24/32	51	31	50.0%
neck	62	19/32/18	42	28	54.8%
head	62	18/33/22	44	30	51.6%
shoulder (R)	62	21/18/27	41	24	61.3%
upper arm (R)	62	27/24/24	46	31	50.0%
fore arm (R)	62	29/23/29	44	35	43.5%
hand (R)	25	9/11/ 9	12	10	60.0%
shoulder (L)	62	24/30/26	48	34	45.2%
upper arm (L)	62	25/27/26	47	30	51.6%
fore arm (L)	62	15/29/13	38	30	51.6%
hand (L)	25	10/12/10	14	13	48.0%
thigh (R)	62	47/27/27	56	30	51.6%
calf (R)	62	25/23/27	45	35	43.5%
foot (R)	62	35/30/21	53	36	41.9%
thigh (L)	62	37/21/21	51	28	54.8%
calf (L)	62	26/31/32	50	30	51.6%
foot (L)	62	28/35/18	54	30	51.6%
Average	\	23.9/25.2/22.6	43.0	28.4	50.9%

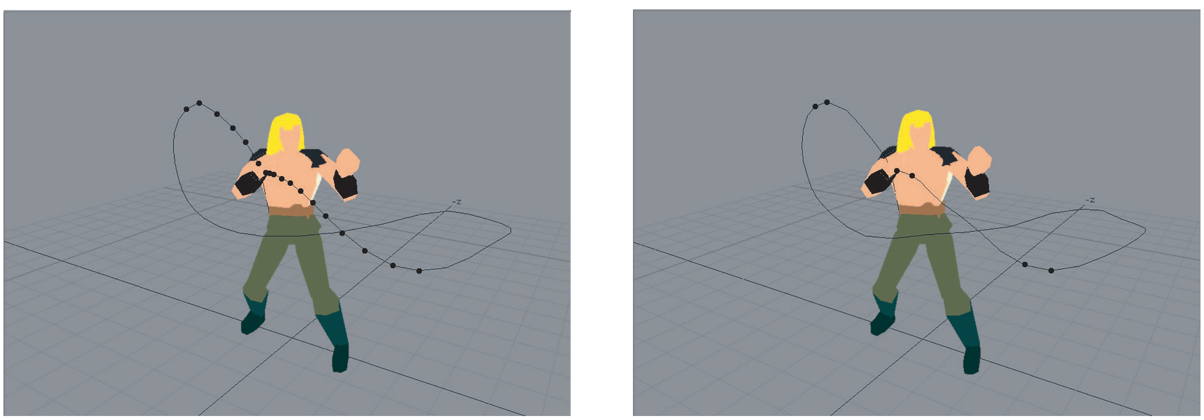

Figure 13: Right hand path, left: original, right: proposed method

4. Conclusion

This paper proposes a method of extracting keyframes that show the characteristics of motion data. The extraction of keyframes from the data are executed with a curve refinement method for handwritten drawings. As a result, with the proposed method the number of frames of motion data is reduced about 50%.

Furthermore, the reduction performance will still be better when applied to motion data that have more than 1000–2000 frames.

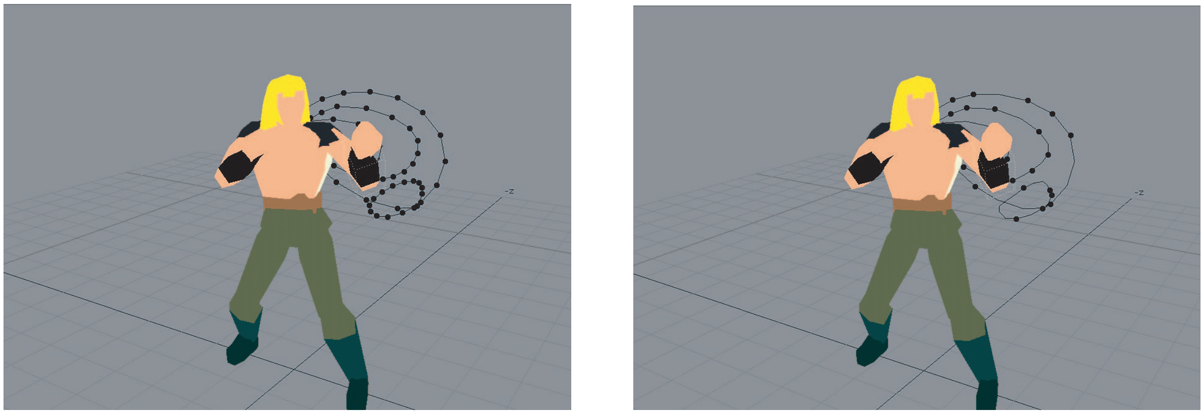

Figure 14: Left arm path, left: original, right: proposed method

Table 3: Examples of keyframe reduction percentage

Action	*Frame Number*	*Ratio*
Swing 2	52	54.0%
Hook	62	47.5%
Jab	17	43.8%
High kick	36	46.1%
High punch	24	46.3%
Back fist	56	48.5%
Sliding foot sweep	48	48.9%
Rolling sobat	25	43.3%
Low kick	48	47.9%
Low punch1	32	43.8%
Low punch2	27	47.5%
Rolling step	58	49.1%
Fore step	30	52.6%
Back step	25	47.6%
Combo kick	55	47.3%
Combo punch	53	49.8%
Dash	46	46.3%
Striking	50	51.7%

References

[1] P. Faloutsos, M. van de Panne, D. Terzopoulos: *Composable Controllers for Physics-Based Character Animation*. SIGGRAPH 2001 Proc., 251–260.

[2] J.K. Hodgins, N.S. Pollard: *Adapting Simulated Behaviors For New Characters*. SIGGRAPH 97 Proc., 153–162.

[3] M. Kobayashi, K. Kondo, H. Sato: *Emphasized Expressions Using Motion Filter in Creating Animation*. Proc. 8th ICECGDG, vol. 2, 451–454 (1998).

[4] K. Matsuda, K. Kondo: *Interactive Sequential Drawing for Sketch Inputting* [in Japanese]. Journal of Information Processing Society of Japan **40**, no. 2, 593–601 (1999).

[5] S. Sato, K. Kondo, H. Sato, S. Shimada, M. Kaneko: *Motion Filter: Emphasized Expressions on Computer Animation* [in Japanese]. The Institute of Image Information and Television Engineers, vol. 49, no. 10, 1280–1287 (1995).

[6] H.C. Sun, D.N. Metaxas: *Automating gait generation.* SIGGRAPH 2001 Proc., 261–269.

[7] M. Unuma, K. Anjyo, R. Takeuchi: *Fourier Principles for Emotion-based Human Figure Animation.* SIGGRAPH 95 Proc., 91–96.

[8] V.B. Zordan, J.K. Hodgins: *Tracking and modifying upper-body human motion data with dynamic simulation.* Eurographics Animation Workshop, Computer Animation and Simulation'99 (1999).

Received August 1, 2002; final form May 11, 2004