MVA2005 IAPR Conference on Machine VIsion Applications, May 16-18, 2005 Tsukuba Science City, Japan
13-24

Automatic Construction of the Motion Database which Allows to Search Contents by a Motion Name

Takashi YUKAWA
Akita Keizaihoka University
Junior College
46-1 Morisawa, Sakura, Shimokitate,
Akita-shi, Akita, 010-8515 Japan
yukawa@akeihou-u.ac.jp

Naoko OBARA and Hideo TAMAMOTO
Faculty of Engineering and Resource Science,
Akita University
1-1 Tagata Gakuen-machi,
Akita-shi, Akita, 010-8502 Japan
{obara, tamamoto}@ie.akita-u.ac.jp

Abstract

In order to reuse the human motion data recorded by the motion capture, we have proposed a motion database construction method named BUYOFU. In this method, captured motion data are divided into a basic motion which is easy to reuse, and the word explaining the motion is used as a search key. Since we have been performing these processes manually at present, if the amount of basic motion data increases registration and maintenance of data will difficult. In order to reuse the captured motion data efficiently, we propose the method of assigning a search key automatically to the divided motion data, and the method of retrieving the target basic motion data from the database by specifying a motion name.

1 Introduction

Character animation that expresses human motion has been used in a wide field from such amusements as movies, televisions, and video games to human engineering and welfare engineering. The motion capture system is used to generate a realistic human motion. This system can obtain the 3-dimensional time series of position and angle data of a sensor that is attached to an actor's body using optics or magnetism.

Though a motion capture system is effective to obtain correct human motion, there is a problem that we cannot use it easily. This is because the system is very expensive and the actor who performs demanded motion is needed. Therefore, the method that makes it possible to reuse the obtained motion data is desired [1].

There are some products to reuse the motion data which are acquired by a motion capture system [2, 3]. In order to retrieve required motion data with these products, we have to choose the data from a category list consisting of a lot of motion. Therefore, if the number of motion data increases, it will become difficult to search required data.

We have proposed the recording and reusing method of human motion data named BUYOFU aiming at recording human motion and at reusing the 3-dimensional human motion data [4]. In this method, the obtained motion data are divided into some pieces that are called basic motion. The basic motion is easy to reuse, and the name expressing its contents is assigned as a search key. Since the registration work to databases, such as division into basic motion of motion data and search key assignment, is conducted manually at present, if the amount of data increases, a problem will arise that registration and maintenance of data become difficult.

In this paper, we propose the method of assigning automatically the search key to basic motion data when basic motion data are registered to a database, and the method of generating a search key from a motion name in order to solve above mentioned problem. We show in an experiment that the proposed method can reduce the burden of the work that basic motion data are registered to a database, and also can enhance the efficiency of reuse of captured motion data.

2 Automatic Assignment to the Basic Motion Data of a Search Key

2.1 Obtaining basic motion data

We used magnetic-type motion capture system to obtain human motion data. Fig. 1(a) shows an actor whose motion is being obtained. Fifteen sensors are attached to the actor at the position shown in Fig. 1(b), and motion data are recorded at the sampling rate of 30 frames per 1 second. The data obtained by one sensor of the motion capture system are time series of data that consist of six elements of the positions (X, Y, Z) and angles (R_x, R_y, R_z) in 3-dimensional space.

Since the directions where an actor performs while obtaining motion or the size of an actor's body varies, even

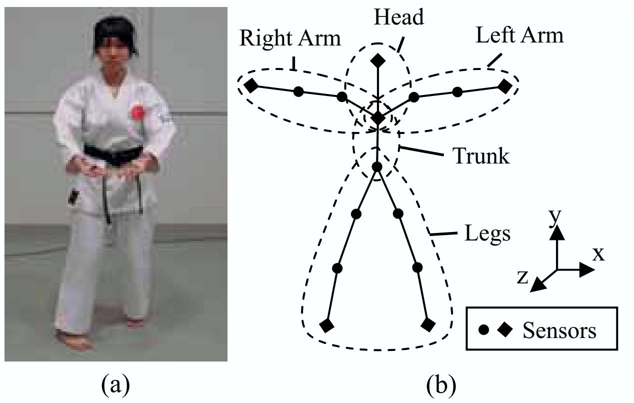

Figure 1. Obtaining motion (a) an actor (b) sensor positions and body parts

if the actor performs the same motion, the motion will be recorded as different motion.

To remove the influence of the direction from the obtained motion, the data recorded by the sensors attached at a head part, a left and right arm parts, and a trunk part are converted into the position and the angle on the basis of a position the waist. The data of the legs are converted so that the front of the body may always turn to the front.

Moreover, the obtained motion data are converted so that the length of each part of the body may become equal to the length of a standard model's one to remove the influence of the difference in the length of an actor's body size. We call this process standardization [4, 5].

In BUYOFU, the basic motion expresses a motion of a certain part of the body. Therefore, the recorded motion data are divided into five parts, that is, a head part, a right and left arm parts, a trunk part, and a leg part shown in Fig. 1 (b). There are two or more sensors in each part. We use only one sensor among them that is considered to detect the best feature in each part (Fig. 1 (b) ♦marks).

In order to make reusable the acquired motion data, the data are divided into some pieces of basic motion. Since at present motion data are examined by viewing and are divided into basic motion by the method by the method discussed in [4], this division is very tedious. By using the automatic division method proposed in [5], the burden of division can be reduced. In this paper, we assume that the acquired motion data are appropriately divided into basic motion in advance.

2.2 Search key assignment to basic motion data

Since basic motion data express simple motion for a short time, change of the value of each element (X, Y, Z, R_x, R_y, R_z) in basic motion is also simple. Hence, we consider we can express the feature of basic motion by using the method that the basic motion data of each element are sampled by n points at an equal interval and that m-valued linear quantization is carried out for each sampled data. In order to make the actor's performing speed be independent of motion data, every basic motion data are sampled by n points irrespective of the data length. In order to regard rotational maximum (2π) as the minimum value (0) with respect to rotation data, the value of angle is linearly quantized by m+1 values and the maximum value m+1 is set to 0.

Since the feature of one basic motion element consists of n pieces of m-value in this method, one basic motion is expressed with 6n length feature vector. By replacing the value acquired as mentioned above with one of m kinds of alphabet which begins from 'A', basic motion can be written in the sequence of 6n characters. We call this character sequence basic motion feature vector, and use it as a search key when registering basic motion data to a database.

As an example, the feature vector obtained with n=5 and m=5 is shown in Fig. 2. In this example, the value of five points sampled at an equal interval is quantized by five values, and the value is replaced with the alphabet of 'A-E'. In the example of Fig. 2, the X element for basic motion feature vector is DDDEE, Y element for basic motion feature vector is BCCCB, and so on. Consequently, the feature vector DDDEEBCCCCBCCDDDEEEDCBCDDCCDEE is assigned to basic motion shown in Fig. 2 as a search key.

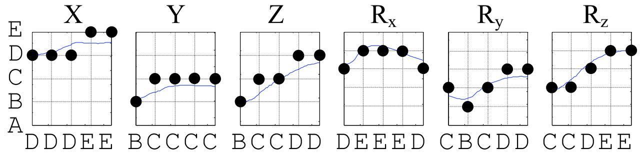

Figure 2. Generation of the basic motion feature vector

Motion Name	X Y Z R_x R_y R_z
$W_1W_2W_3$	DDDDCBCCCCDCBBBEEEEECCDDDAEEED
$W_4W_2W_3$	CCDDDBBBBBCCCBBDEEEEBBCCCDCCCC
$W_5W_2W_3$	DDDDDBCCBBDCCCCEEEEECCCCCAEDCC
W_2W_3	--DD-B-----C----EEEE----------

Figure 3. Partial feature vector

3 Retrieving Basic Motion Data by Given Motion Name

3.1 Motion name

In this paper, we assume that the basic motion name consists of one or more concatenation of the word showing the part name of the body used for motion, a direction of motion, and a position of the part of the body, and the kind of motion. We call each word a motion word.

Basic motion can be considered to be the change from one posture to another posture. The posture when basic motion is starts is called a start posture, and the posture when the basic motion ends is called an end posture. A basic motion name is described as follows.

Start Posture Name > End Posture Name

An end posture name can also be specified to be the start posture name of the next basic motion. When there is no suitable name of an end posture, the basic motion name is used as an end posture name. For example, the name of "sitting on a chair" is used as a posture name after the motion "sitting on a chair."

3.2 Motion word dictionary

We consider that there is a certain relation between the motion word that constitutes a motion name and the feature of basic motion data. For example, the motion name of "sitting on a chair" consists of two motion words of "sitting on" and "a chair." In this example, a motion word "a chair" can be related with the height of y coordinate of the waist after sitting on a chair, and the motion word "sitting on" means that the value of y coordinate of the waist decreases.

Thus, a motion word has a relation with the value or the change of the value of the specific element of the basic motion feature vector. First, the teacher motion data which are assigned to the basic motion name are prepared. Next, the element is found out which is common to the basic motion feature vector (partial feature vector) containing a certain motion word. Finally, a table is made which corresponds basic motion feature vector with a motion word.

We call this table a motion word dictionary.

There is a motion word which has strong correlation with another motion word. There is a motion word which is used in a different meaning depending on the kind of motion name, for example, the word “put” which is common among motion names “put on” and “put off.” Since the motion word used in such a different meaning has few elements common to the feature vector, the combined two motion words existing in a basic motion name is added to a motion word dictionary.

The example of derivation of the partial feature vector is shown in Fig. 3. In this example, in order to obtain the partial feature vector of motion word “W_2W_3” common to the three basic motion names “$W_1W_2W_3$”, “$W_4W_2W_3$”, and “$W_5W_2W_3$”, the mean and distribution of each feature vector element are calculated. Next, the mean value of an element with the value of distribution less than a threshold is used as the feature of the element common to motion word “W_2W_3.” The element whose value of distribution is more than a threshold is decided not to have a common feature, and the value of the element is replaced with the mark of ‘−’. We calculate partial feature vectors for all the target motion words, and create the motion word dictionary.

3.3 Automatic search key generation

The following procedure is performed for generating a suitable basic motion search key from the specified motion name.

1. The combination of two motion words is created for each of the start posture name and the end posture name which constitutes a motion name.
2. The partial basic motion feature vector corresponding to the combination of two words created by 1. is retrieved from a motion word dictionary.
3. The mean value of each element of the partial basic motion feature vector obtained by 2. is calculated. In case the values of all elements are ‘−’, the ‘−’ is used as a mean value.

The feature vector of basic motion generated using this procedure may contain the element to which the value is not assigned (i.e. ‘−’). In this case, the value ‘−’ is not used as the feature of the basic motion. When the generated key is compared with a search key in a database to retrieve basic motion, the value ‘−’ is excluded from the generated key.

4 Experiment

In order to verify the effectiveness of the proposed method, we made an experiment. Three kinds of the form (kata) data of the karate-do each of which is performed by three actors ware used for human motion data.

In order to use these data as teacher data, obtained data were divided into the basic motion, and basic motion name was assigned to a basic motion data before the experiment. We created 96 kinds of 288 teacher basic motion data.

4.1 Experiment system

The overview of an experiment system is shown in Fig.

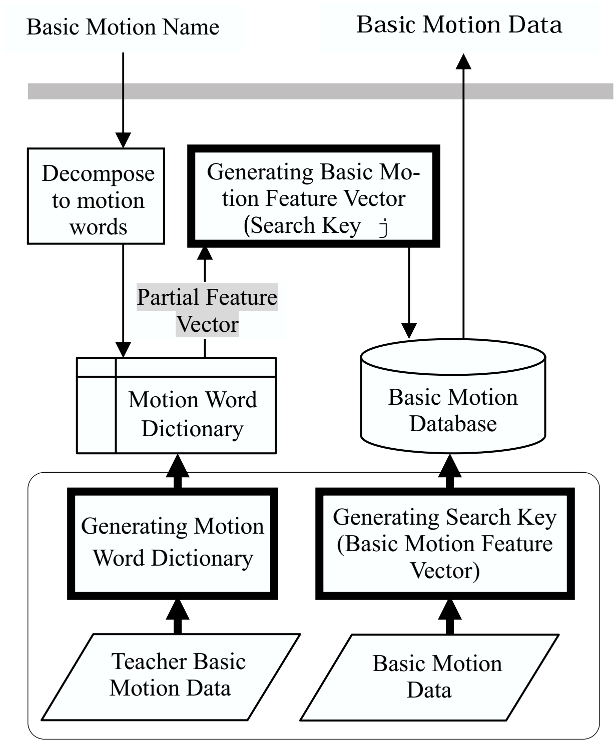

Figure 4. Overview of experiment system

.

4. The feature vector is generated from basic motion data and a motion word dictionary is made using teacher motion data. A basic motion database is constructed. In order to investigate whether correct basic motion data are retrieved, we applied some basic motion names to the constructed basic motion database.

4.2 Generation of the basic motion feature vector

First, search key was assigned to the basic motion data. In this experiment, the value of each basic motion element is sampled by five points and the sampled value was quantized by five values. As a result, one basic motion was expressed with the feature vector of basic motion of 30 characters.

The basic motion database was constructed which used the feature vector of basic motion as a search key.

4.3 Creation of a motion word dictionary

Next, the motion name of the teacher data was decomposed into the motion word, and correspondence with the feature basic motion vector was examined for the combination of two motion words. As a result, 96 kinds of basic motion names were generated, and 265 kinds of combined motion words were generated from 96 kinds of motion names. The feature basic motion vectors for every combination of two motion words were created.

4.4 Retrieving basic motion data

In order to verify the effectiveness of the method of retrieving the target basic motion data from a basic motion database by specifying the motion name, the experiment was conducted using five kinds of basic motion names shown in Table 1.

For example, using a basic motion name “Tsukami Hikiyose > Nukite Chudan Tsuki”, the “Tsukami Hikiyose” is derived from the combination of two words of

Table 1. Performance Data

Basic Motion Name	# of appearances	# of retrieves	# of correct answers	# of faults
SeikenChudanTsuki>ChudanYokoUke	12	12	12 (100.0%)	0
Hikite>SeikenChudanTsuki	12	8	8 (66.7%)	0
TsukamiHikiyose>NukiteChudanTsuki	9	12	9 (100.0%)	3
NukiteChudanTsuki>TsukamiHikiyose	9	8	8 (88.9%)	0
ChudanYokoUke>Hikite	6	7	6 (100.0%)	1

"Tsukami" and "Hikiyose" in a start posture name. "Nukite Chudan", "Nukite Tsuki", and "Chudan Tsuki" are derived from the combination of three words of "Nukite", "Chudan" and "Tsuki" in an end posture name. The following partial feature vector is obtained for each combination of two motion words by searching a motion word dictionary.

Tsukami Hikiyose
`CCDDDCCBBBCCCCBEEE---CCCC-CCCC`
Nukite Chudan
`CCDDDCC-BB-CCC--EEE--CCCC--CCC`
Nukite Tsuki
`CCDDDCC-BB-CCC--EEE--CCCC--CCC`
Chudan Tsuki
`C-DDD-C---CC----EEE--CCCC-----`

By compounding these partial vectors, the following vector was obtained as a feature vector of basic motion name "Tsukami Hikiyose > Nukite Chudan Tsuki."

`CCDDDCCBBBCCCCBEEEE--CCCC-CCCC`

The target basic motion data can be retrieved from a basic motion database by using this feature vector as a search key.

The experimental result is shown in Table 1. Each column of the table indicates from the left the basic motion name (basic motion name specified to search), the number of times of an appearance (number with which the basic motion name is included in the database), the number of detection (the number of basic motion retrieved from the database by the generated search key), the number of correct answers, and the number of incorrect detection (the number of the basic motion name specified among the number of detection).

4.5 Discussion

As shown in Table 1, the ratio of the number of correct answers to the number of times of appearance is 90.0% on average. However, there is the case where the number of the incorrect detection is high for some basic motion names.

The basic motion with high incorrect detection includes many motion words which are used in the other basic motion name. Because the other basic motion name with the same motion word has the same feature vector, incorrect basic motion is detected. To solve this problem, a motion words need to be selected whose feature vector is different from each other.

The experimental result shows that the proposed method must be an effective candidate when a basic motion database is constructed.

5 Conclusion and future work

In this paper, we propose a method of 1) automatic assignment of search key based on the feature of basic motion, 2) generation of a search key from the word expressing the basic motion, in order to enable efficient reuse of the human motion data acquired by the motion capture system. The experiment was conducted using some basic motion names, and the effectiveness of proposed method was shown.

Since it is possible to register basic motion data to a database automatically by using this system, effective reuse of motion capture data is expectable.

Future work includes 1) application of this method to the data of other kinds of motion, such as dancing and gymnastics, 2) investigation of the features of basic motion, and 3) the motion word selection which reduces the rate of incorrect detection.

Acknowledgments This research was partially supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Young Scientists (B), 15700076, 2003.

References

[1] O. Arikan and D. Forsyth: "Interactive motion generation from examples," In *Proceedings of the 29th annual conference on Computer Graphics and Interactive Techniques*, pp.483-490, 2002.

[2] Motion Capture Database RIKIYA: http://www.viewworks.co.jp/rikiya/

[3] Motion Capture Database MOCAPPERS: http://www.anystyle.jp/mocappers/

[4] T. Yukawa, T. Kaiga, K. Nagase, and H. Tamamoto: "Human Motion Description System by Using BUYO-FU," *IPSJ Journal*, vol.41, no.10, pp.2873-2880, 2000.

[5] T. Yukawa, N. Obara, and H. Tamamoto: "Automatic Detection of Motion Primitive Boundaries from Human Motion Capture Data," *IPSJ Journal*, vol.45, no.4, pp. 1198-1201, 2004.