Real Time Motion Authoring of a 3D Avatar

Harinadha Reddy Chintalapalli and Young-Ho Chai

Graduate School of Advanced Imaging Science, Multimedia & Film,
Chung-Ang University
{harinath, yhchai}@cau.ac.kr

Abstract. A diverse number of avatars that express human motions are produced in the fields of film, computer games, and virtual reality. In order to simulate human motions realistically, data are collected through a sensor-based method such as motion capture and are organized into a database; motions are then generated on the basis of this database. Human body motions can be reproduced naturally through a vast amount of data, but complicated calculations are necessary to generate these motions, and thus, it is not easy to generate an interactive 3 dimensional (3D) avatar in real time. In order to generate the motions of a 3D mobile avatar in real time, in this study, we implement a dataset with an intuitive classification through pattern recognition and propose the mapping of various avatar motions using a flexible and adaptive hierarchical algorithm. For real time motion authoring, we use a support vector machine (SVM) to classify the types of motion pattern data input from magnetic, angular rate, gravity (MARG) sensor.

Keywords: Mobile avatar, Motion authoring, Pose classification, MARG sensor, Support vector machine

1 Introduction

Creating a 3D avatar with natural motions is an issue frequently dealt with in the fields of film, computer games, and virtual reality. To address this issue, software (such as Poser and Motion Builder) or various sensors are used for extracting and processing data to generate various forms of motions. To generate motions by using this method, the following requirements must be fulfilled: first, a large amount of data is required for generating the natural motions of an avatar; second, the processing and the interpolation of data are necessary to connect the motions smoothly; and finally, the user must be able to generate motions intuitively [8].

In this study, we use a gyro sensor to help users learn the motion patterns [4] of the gyro sensor of their choice and to generate new motions of the avatar in real time from a pre-existing motion dataset through the learned patterns. For the motion data of the avatar, we can use files of various formats or a database management system (DBMS) with motion data. For example, BVH file format includes the skeleton information and motions of skeletons in each frame, and the Wave Front format includes the vertex or texture information. The FBX format is software that includes many types of information such as actual motion capture data and skeleton structure, and the C3D format is a form of data extracted from the motion capture system. This

form of data can be used directly, managed by a DBMS, or searched to generate motions using a query. Various other formats exist, and many studies are being conducted to generate efficient data formats.

In this paper, we propose a hierarchical structure in which real-time motion generation is also possible in mobile de-vices [5-7], by using formats randomly selected from various motion data formats. The basic approach is as follows: the user uses the gyro to make a series of rotary motions and the motion data of one-to-one correspondence is then accessed to form the motion of the avatar. The rotation data of the gyro moved by the user uses SVM [2, 3] to learn, predict, and estimate the rotary motions. Considering that it is a mobile device, we need to ensure that appropriately sampled data are used in the learning of the SVM in order to reduce the load.

To generate the avatar's motions using the gyro, three processes need to be followed: start process, order process, and end process. The start process selects joint groups to facilitate significant motion authoring; the order process sets motions in the selected joint groups; and the end process completes the motion authoring or sets additional options for the motions. The series of processes minimizes the number of patterns created using the 3-axis data, because the numbers or letters that can be immediately remembered by humans are generally 7±2 [7].

In this study, we aim to show the utility of the proposed method by presenting real-time authoring examples for avatar motions in mobile devices. For the experiment, various motion data are used, and the user can use the touch sensor and the gyro sensor to directly and immediately control the avatar. The method proposed in this paper can be implemented relatively simply and used in a wide range of applications.

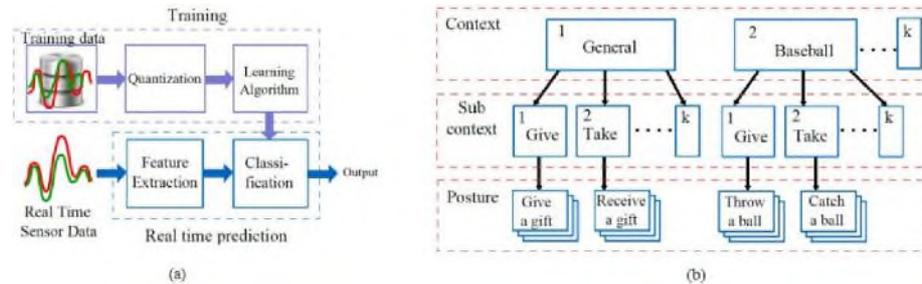## 2 Real time motion pattern recognition

Fig. 1. (a) Motion pattern recognition. (b) Context-aware action classification.

Equations 1 and 2 describe how to compute the attitude quaternion by numerical integration [1], where Δt is the sampling period,

-1 is the previous normalized attitude quaternion, and q0, q1, q2 and q3 represent the elements of quaternion .

$$= \underset{2}{1} \; \sim$$

$$- 1 \otimes \omega \tag{1}$$

$$= [q0 \ q1 \ q2 \ q3 \quad I = !-1 +$$

$$! \qquad\qquad\qquad t\,t$$
$$!2 \ t \qquad\qquad\qquad \textbf{(2)}$$

The implemented motion pattern recognition block diagram is shown in Figure 1(a). Training is the one time procedure that should be done first in order to recognize motion patterns in real time. Training data is the database of the recorded sensor data for different motion patterns; note that this database contains several samples of each motion pattern for training the algorithm. This training data will then be quantized to reduce the amount of data to be trained. This is because the sensor system provides data at higher rates (> 100Hz), using all this data for training will thus not be necessary. After training is finished, the resultant model from the training phase is used for the real time prediction of motion pattern made using the sensor data. The feature extraction module extracts features such as absolute magnitude of the recognized motion pattern and will be used later. We used radial basis function (RBF) kernel for the SVM classifier because inertial sensors data contain zero mean Gaussian noise. The classifier outputs predicted motion pattern label along with an accuracy measure. Depending upon this accuracy and features computed, the predicted motion pattern will be considered as success.

3 Motion dataset generation and results

The avatar's actions are classified according to the context, which is connected to the movements of the avatar. For example, the action of giving and receiving an object can occur in various contexts as shown in Figure 1(b). This study uses the motion data of around 2,000 processed frames by using the Poser software based on motion capture data. For the motion data format, the BVH format is used.
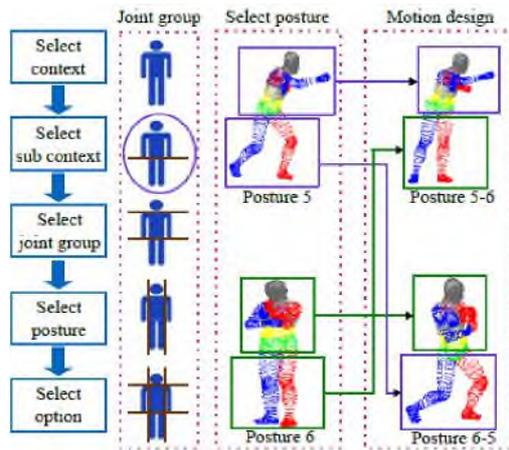

Fig. 2. Posture generation process.

In this study, we use BVH because human motions can be generated through the central axis (hip) position and each skeleton's rotation data, which BVH satisfies.

Figure 2 shows the motion generation process using BVH extracts and structuralizes the skeleton information from BVH and structuralizes it, and then maps the motion information suitable for the generated skeleton in frame units to generate motions. For example, the position of joint x is predicted by adding all offsets of the parent joint and adding the ROOT position in the present frame using Equation 3, where k is the parent joint count.

$$P_{Joint} = P_{Root} + \sum_{=1} \qquad (1)$$

All joints have a position property and a rotation property. The rotation of joint x can be predicted by multiplying the rotation quaternion of ROOT by the rotation quaternion of all parent joints as shown in Equation 4, where Q and k represent quaternion and parent joint count respectively.

$$Q_{Joint} = Q_{Root} \otimes \prod_{=1} , \otimes \qquad (2)$$

4 Conclusion

In this study, we have successfully generated various forms of motions from limited motion data using a class of SVM classifier with our MARG sensor motion data pattern recognition through a context-aware adaptive hierarchical structure. The context-aware structure proposed in this paper would be the optimal method for the real-time interactive avatar's motion authoring through efficient bifurcation with limited resources, and will enable relatively easy motion authoring by classifying the joint groups.

References

1. A. Sabatini: Quaternion-based extended kalman filter for determining orientation by inertial and magnetic sensing. IEEE Transactions on Biomedical Engineering, vol. 53, no. 7, pp. 1346-1356, (2006).
2. Chang C.-C. and Lin C.-J.: Libsvm: A library for support vector machines. ACM Trans. Intell. Syst. Technol., vol. 2, no. 3, (2011), pp. 27:1-27.
3. Cristianini N. and Shawe-Taylor J.: An introduction to support Vector Machines: and other kernel-based learning methods. Cambridge University Press, New York, NY, USA, (2000).
4. Liu J., Zhong L., Wickramasuriya J. and Vasudevan V.: Uwave, Accelerometer-based personalized gesture recognition and its applications. Pervasive Mob. Comput., vol. 5, no. 6, pp. 657-675, (2009).
5. Madgwick S., Harrison A. J. L. and Vaidyanathan R.: Estimation of imu and marg orientation using a gradient descent algorithm. In Rehabilitation Robotics (ICORR), 2011 IEEE International Conference, pp. 1-7, (2011).
6. Purkayastha S. N., Eckenstein N., Byrne M. D., O'Malley M.: Analysis and comparison of

low cost gaming controllers for motion analysis. IEEE/ASME Advanced Intelligent Mechatronics, AIM, (2010).

7. Miller G. A.: The magical number seven, plus or minus two: Some limits on our capacity for processing information. The Psychological Review, vol. 63, no. 2, pp. 81-97, March (1956).
8. Serizawa T., Yanagida Y.: Poster: Authoring tool for intuitive editing of avatar pose using a virtual puppet. Proceedings of the 2008 IEEE Symposium on 3D User Interfaces, pp. 153-154. IEEE Computer Society, (2008).