Interactive Control of Dance Groups Using Kinect

Asako Soga and Itsuo Yoshida
Faculty of Science and Technology, Ryukoku University, 1-5 Yokotani, Oe-cho, Seta, Otsu, Japan

Keywords: Kinect, Motion Data, Dance Groups, Gesture Recognition.

Abstract: This paper describes a system that can control the CG animation of dance groups in a virtual world by recognizing gestures using Kinect in real time. The system can change the formation of CG dancers by right-arm gestures and their motions by left-arm gestures. Moreover, a virtual camera is controlled by the gestures of both arms. The rotational data of the joints are used to recognize each arm gesture to control the dance groups, and the translational data of the joints are used to control the virtual camera. Since the system recognizes both arm gestures at the same time, users can control two or more objects by a combination of gestures.

1 INTRODUCTION

In recent years, many cultural ceremonies and traditional dances have been archived with digital technologies. Motion capture technologies are used for archiving traditional dances, and digital contents with realistic motion have been widely used. Many researches support about education and the creation of traditional dances using archived motion data.

The purpose of this research is to develop systems that can intuitively control CG animations by human actions using motion data and interactive techniques. We have been developing simulation systems for dance groups to simulate multiple CG characters (Soga, 2010) (Soga, 2012). In this paper, we describe a control system for dance groups using Kinect in real time.

Simulation systems for groups and crowds have already been developed (Kwon, 2008). Many studies have used such general human motion data as walking and running as well as interactive authoring of human groups (Ulicny, 2004). However deriving complex human motions in composition remains a challenge.

Since it is easy to capture human body motion with Kinect devices, applications using them are being used for entertainment, rehabilitation, and dance training (Marquardt, 2012). Our system simulates dance groups and can change human motions as well as the formation of dance groups. By recognizing human gestures with Kinect, a user can interactively change the formations and the motions of CG dancers arranged in 3D space. This paper describes a method to recognize gestures by Kinect and to control dance groups.

2 SYSTEM OVERVIEW

This system simulates the motion of multiple dancers. The motion or formation of dance groups in the virtual world is changed by gestures that are recognized by capturing human motions and tracking them by a Kinect device in real time. The camera view in the 3D world is also changed by gestures. Since the system simultaneously recognizes both arm gestures, users can control two or more objects by a combination of gestures. They can also change the formation of the groups by right-arm gestures and the motions of virtual dancers by left-arm gestures. Therefore, users can manipulate the virtual dancers like a stage director and control the dance motions and formations by the postures that are often used in ballet.

Our system runs on a personal computer with a Windows OS and is operated with a Kinect. We developed it using Visual C++, DirectX, and Kinect for Windows SDK. The DX Library handles DirectX. Figure 1 shows the system configuration.

Dance animations were obtained from a professional dancer using an optical motion capture system. Motion clips of classical ballet steps for the dance groups were prepared for the system. Users can arbitrarily select the dance motions of virtual dancers.

Soga A. and Yoshida I..
Interactive Control of Dance Groups Using Kinect.
DOI: 10.5220/0005357303620365
In *Proceedings of the 10th International Conference on Computer Graphics Theory and Applications* (GRAPP-2015), pages 362-365
ISBN: 978-989-758-087-1

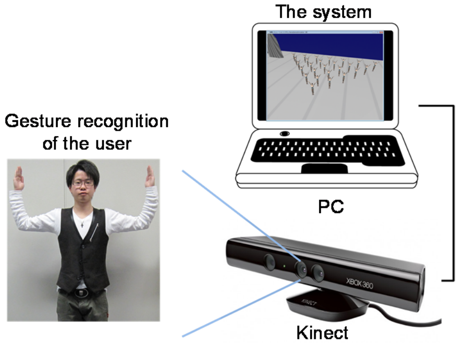


Figure 1: Configuration of system.

3 GESTURE RECOGNITION

3.1 Process

Figure 2 shows the process of gesture recognition. This system has two modes: group control and camera control. Although users usually operate it in the group control mode, they can switch to the camera control mode by recognizing a specific pose only if they want to change to the camera view. In the gesture recognition process, first, the system tracks the skeletons of users and gets the positions of each joint. If the system recognizes the gesture for changing the mode, the mode is changed; otherwise the system runs the process of the selected mode.

In the group control mode, each joint angle is calculated by the captured joint positions. For example, if the system recognizes a gesture defined by the angles of the right arm, it changes the formation of the dance groups. In the same way, if it recognizes a gesture defined by the angles of the left arm, it changes the motion of the dance groups. These processes are executed for each frame.

In the camera control mode, the system recognizes the position of the captured joint to adapt the moving distance to the camera angle. First, the system records the positions of both hand joints when the mode is changed to this mode and calculates the distance between the recorded and current positions. The angle of the camera view is changed based on the moving distance.

3.2 Gesture Recognition by Angles

In the control mode of the dance groups, joint angles, which are used to recognize gestures, are calculated by the inner product of vectors between the joints of the base pose and the captured joints. The rotational angles of the shoulder and elbow joints are calculated

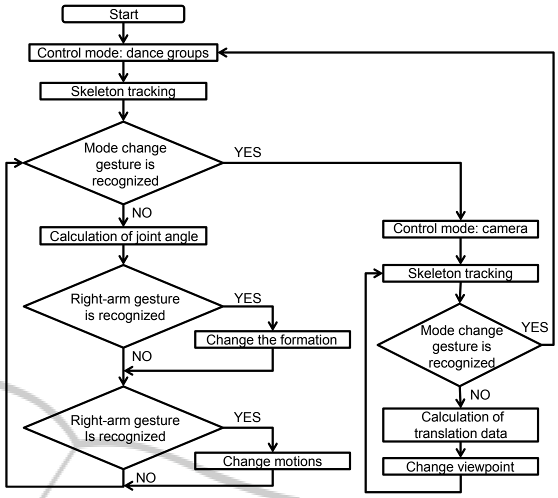


Figure 2: Process of gesture recognition.

Table 1: Arm gestures and their rotational angles.

gesture	shoulder	elbow
horizontal	0	0
upper circle	-45	-45
lower circle	45	45
right angle	0	-90
upper oblique	-45	0
lower oblique	45	0

with the positions of the shoulder, elbow, and hand and used to recognize the arm gestures. These gestures are recognized by comparing these angles to the angles of each pose that are predefined in the template file. Table 1 shows the predefined six gestures and their rotational angles for the Z-axis. All angles for the X and Y-axes are defined as zero degree. Figure 3 shows examples of arm gestures.

The rotational angles are calculated by the average angle of ten frames to avoid misrecognition. The recognition succeeds if the difference between the predefined angle and the calculated average angle is less than 15 degrees since users have difficulty inputting the complete pose.

3.3 Gesture Recognition by Positions

In the camera control mode or the mode change, the relational positions of joints are used to recognize gestures. When an arm is stretched outward, the mode is changed to the translation mode. When both hands are placed in front of the chest, the mode is changed to the rotation mode.

To switch to the rotation mode, the pose must satisfy three conditions; the distance of both hands is narrower than that of the shoulders, both hands are

(a) horizontal (b) lower circle (c) right angle

Figure 3: Examples of arm gestures.

placed in front of the torso, and the height of each hand is between the chest joint and the top of the head. To switch to the translation mode, the pose must satisfy two conditions; a hand is placed in front of the torso, and the depth distance of the hand from the torso is longer than the half of the arm.

To release these modes, the following conditions are required. To release the rotation mode, the distance between both hands has to be wider than that of the shoulders. For the translation mode, the distance between both hands should be less than ten centimeters. After switching to the camera control mode, the virtual camera can be moved and rotated by the hand in each direction, such as up-down and right-left.

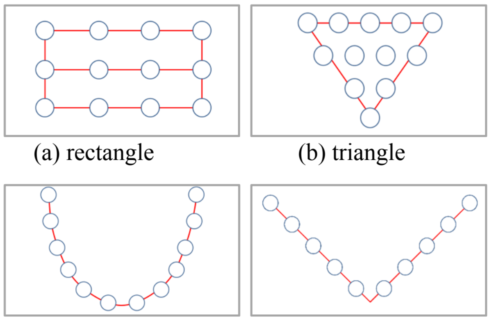
When switching the mode, the system records the hand position, and the distance between the recorded and current positions is applied to the translation or the rotation of the camera. In the rotation mode, the moving distance for each axis is adapted to the rotation angle, and in the translation mode, the moving distance for the XZ plane is adapted to the translation of each axis. To release the camera mode, both hands are required to overlap.

4 INTERACTIVE CONTROL BY GESTURES

The user changes the formation of the CG dancers by right-arm gestures and their motions by left-arm gestures. When a right-arm gesture is recognized, each CG dancer moves to the defined position based on the formation and the number of CG dancers. When a left-arm gesture is recognized, the motion of each CG dancer is changed based on the gesture.

We prepared six typical formation patterns: straight line, circle, rectangle, triangle, curve, and V-letter line. Figure 4 depicts examples of the formation patterns. The system automatically calculates the dancer positions based on the number of dancers.

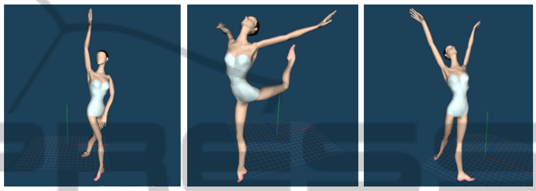
Easy and intuitive gestures are assigned to the six formations. For example, horizontally stretching the right arm arranges the CG dancers in a line. Ballet



(a) rectangle (b) triangle

(c) curve (d) V-letter line

Figure 4: Formation patterns for dance groups.



(a) pose (b) attitude (c) port de bras

Figure 5: Examples of dance motions.

motions likely to be performed with the formation are assigned to the six left-arm gestures, since it is easy to perform the same gestures for both arms. Figure 5 shows examples of the dance motions. If the user makes the same gesture for both arms, the CG dancers perform ballet steps that are often seen in a formation corresponding to the right-arm gesture.

Table 2 shows the gestures and the corresponding formations and ballet motions. Figure 6 shows a combination of gestures. Figure 6(a) is an example of the same gestures for both arms, and Figure 6(b) is different gestures for both arms. The corresponding results are shown in Figure 7.

5 EXPERIMENT

We conducted an experiment with nine students who seldom play videogames with Kinect. After explaining the experiment process, the required task, and the system usage, they tried the system and performed the task. Then the unused functions of the system were explained, and they answered questionnaires. They mainly tried the formation change, the dance motion change, and manipulations of both arms. As a task, the students arranged the CG dancers in the indicated formation and with the required dance motion. The following three factors were evaluated on five levels (1: bad, 5: good): (a) suitability of manipulation and gestures, (b)

Table 2: Gestures and corresponding formations and ballet motions.

Gestures	Formation	Dance
horizontal	straight line	pose
upper circle	circle	attitude
lower circle	curve	arabesque
right angle	rectangle	port de bras
upper oblique	V-letter line	pas de chat
lower oblique	triangle	pas de buorrée

(a) identical gestures (b) different gestures

Figure 6: Examples of gestures of both arms.

simplicity and memorability of manipulation, and (c) velocity and suitability of reactions.

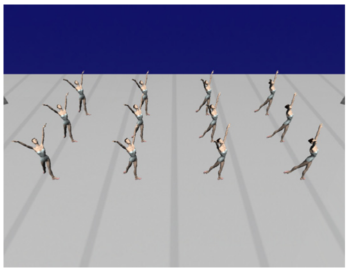
The averages of the nine students were (a) 3.6, (b) 3.8, and (c) 2.7. In terms of (a) the suitability of manipulation and gestures and (b) the simplicity and memorability of the manipulation, the results were mostly good. However, (c) the velocity and the suitability of the reactions was worse than the others.

User feedback suggested that we include finger gestures, but different gestures for both arms were acceptable. Negative comments complained that the manipulations were complex, and that the gestures were difficult to memorize. However, most students seemed to use our system easily after they got used to its manipulations.

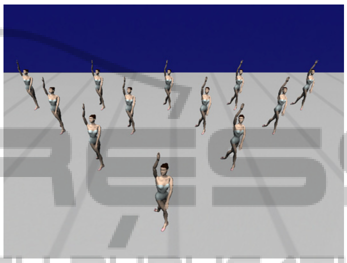
6 CONCLUSIONS

This paper described a system that can control the CG animation of dance groups in a virtual world by recognizing gestures using Kinect in real time. By recognizing human gestures, a user can interactively change the formations and motions of CG dancers arranged in 3D space. Moreover, a virtual camera is controlled by the gestures of both arms. We experimentally verified that (a) the suitability of the manipulation and the gestures and (b) the simplicity and the memorability of the manipulation of the system are better than (c) the velocity and the suitability of the reactions.

However, this system only recognizes human poses and cannot recognize human motions. In future work, the recognition of the motions of arms and legs will be supported to improve our system. In addition, the motions of the person who is controlling the dance groups might become a type of motion or enhance the prepared motion.



(a) identical gestures



(b) different gestures

Figure 7: Results for controlling dance groups.

REFERENCES

Soga, A., Boulic, R. and Thalmann, D., 2010. Motion Planning and Animation Variety using Dance Motion Clips, In *CW '10, Proceedings of the 2010 International Conference on Cyberworlds*, pp. 421-424.

Soga, A., Yoshida I., 2012. A Simulation System for Dance Groups using a Gamepad, In *GRAPP2012, Proceedings of 7th International Conference on Computer Graphics Theory and Applications*, pp.365-368.

Kwon, T., Lee, K. H., Lee, J. and Takahashi, S., 2008. Group Motion Editing. In *ACM Transactions on Graphics*, Vol. 27, No. 3, Article 80.

Ulicny, B., Ciechomski, P. H. and Thalmann, D., 2004. Crowdbrush: Interactive Authoring of Real-time Crowd Scenes. In *SCA '04, Proceedings of the 2004 ACM SIGGRAPH/Eurographics Symposium on Computer Animation*, pp. 243-252.

Marquardt, Z., Beira, J., Em, N., Paiva, I. and Kox, S., 2012. Super Mirror: A Kinect Interface for Ballet Dancers. In *CHI EA '12, Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems Extended Abstracts*, pp. 1619-1624.