A Study on the Sport Demonstration System Based on Three-dimensional Reconstruction

Lei Geng* & Junjie Li

Department of Public Fundation, Langfang Health Vocational College, Langfang, Hebei, China

ABSTRACT: Sport demonstration system can help improve the effect of teaching and training in physical sports. In practical athletic training, the two-dimensional sport demonstration system has been widely applied. However, there's certain visual deficiency existing in the accurate motion positioning and the kinematic analysis of the two-dimensional sport demonstration system. In order to help coaches and students clearly understand the subtle points of physical sports in an all-around way, this paper studied the sport demonstration system based on three-dimensional reconstruction. It analyzed the camera imaging principle and the three-dimensional reconstruction principle in theory, and has reached the conclusion that the key in the three-dimensional reconstruction of two-dimensional digital video lies in the segmentation of two-dimensional videos. The selective analyses of this paper are about the three-dimensional reconstruction of segmented images. Based on this reason, this paper proposed the calibration technique of application camera and the means of stereo matching, so as to achieve the precise matching of accurate space ray equation and corresponding pixels. In the final part, this paper took diving as the example and designed the sport demonstration system composed of analysis module, three-dimensional simulation, interface module and data management module on basis of three-dimensional reconstruction.

Keywords: camera calibration, stereo matching, three-dimensional reconstruction, visual tracking, virtual athlete

1 INTRODUCTION

Haidong Zhang (2013) has pointed out that the application of scientific thinking and microcosmic perspective in analysis of sport techniques can help improve athlete's performance. In the meantime, it can also provide theoretical proof for coach's teaching. [1] In order to provide a sport demonstration system for coaches and athletes in an all-around way, this paper studied the establishment of the three-dimensional sport demonstration system based on two-dimensional sport videos.

Certain studies have been made according to the three-dimensional reconstruction problem of three-dimensional videos. Among the studies, Zhongren Wang et al (2015) have pointed out that the internal and external parameters of the measurement system can be confirmed by the circular dot markers; the images shall be precisely selected from the interest area; and the surface shall be reconstructed through precise positioning, so as to reduce wild point error [2]. Lei Yang et al (2015) have proposed a matching method used among multiple sequences facing scene reconstruction which can solve the problem of scene information [3]. Xufei Zhang et al (2013) have pointed out that the three-dimensional reconstruction of billiard table in binocular vision can be realized by the application of the reconstruction method provided by the computer vision database OpenCVD [4]. Jun Wang et al (2013) have proposed a three-dimensional reconstruction method of matrix decomposition based on

*Corresponding author: 35372989@qq.com

target SAR image sequence. If more image frames can be reconstructed in the application of this method in the three-dimensional reconstruction process of sequence image, better reconstruction property can be achieved [5]. Based on previous research, this paper took two-dimensional digital video of diving as the example to reconstruct the sport demonstration system through three-dimensional technique.

2 ANALYSIS OF THE THREE-DIMENSIONAL RECONSTRUCTION PRINCIPLE BASED ON TWO-DIMENSIONAL IMAGES

The content of computer vision [6] is to restore the three-dimensional geometrical information of the filmed objects by using several two-dimensional digital images and the internal camera parameters of the digital images, and thus to apply computer vision in specialized analysis. The main content of the research in this paper was the three-dimensional reconstruction principle based on two-dimensional images. It abstracted the algorithm and established the sport demonstration system through the methods which have been realized by computer science.

2.1 *Camera Imaging Principle*

Camera imaging process can be regarded as the projection process of space [7]. In general, this process can meet the physical principle of pin-hole imaging. The

projection equation is shown in Equation(1) [8] in which the $\mathbf{R}$ refers to the rotation matrix from the world coordinate to the camera coordinate; $\mathbf{t}$ refers to translation vector; $\mathbf{A}$ refers to the intrinsic parameter matrix of the camera; k refers to proportionality coefficient; $\widetilde{m}$ refers to the homogeneous coordinate of image point; and P refers to the homogeneous coordinate of p. The intrinsic matrix A of the camera is shown in Equation (2):

$$\widetilde{m} = k\mathbf{A}[\mathbf{R},\mathbf{t}]\widetilde{P} \tag{1}$$

$$\mathbf{A} = \begin{pmatrix} \alpha_u & S & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{pmatrix} \tag{2}$$

There're five parameters contained in the intrinsic matrix component elements of the camera in Equation (2). These five parameters are irrelevant with camera movement, and thus are called the intrinsic parameters of the camera. Coordinate (u_0, v_0) refers to the coordinates of the camera center in the image while coordinate (α_u, α_v) refers to the scale factors of the horizontal axis and the vertical axis. S refers to inclination factor.

According to the principle mentioned above, if the coordinate of a point naming P in the space is (x_w, y_w, z_w), the image point $p(u, v)$ of this point in a two-dimensional image is the angular point between the camera screen and the line connected by P and the camera optical center O. If it is known that the image points of a point P in a binocular camera are $p_1(u_1, v_1)$ and $p_2(u_2, v_2)$, P stands for intersection point of the extension cord of space ray $o_1 p_1$ and space ray $o_2 p_2$. The three-dimensional coordinate of P can be calculated by calculating the intersection point of these two space camera shooting images [9] as shown in Figure 1:
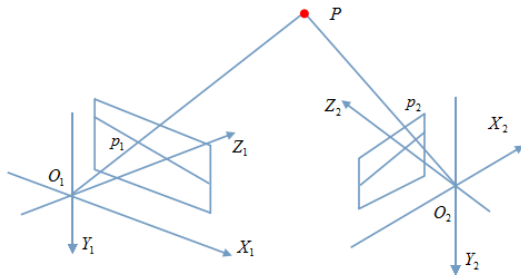

Figure 1. Schematic diagram of the camera geometrical relationship in binocular vision [10]

It can be seen that the vision of these two cameras can be regarded as the projection process of the image coordinates in the three-dimensional world.

The three-dimensional reconstruction is the inverse process of projection [11]. The intrinsic parameter matrix of the camera can be calculated according to Equation (1). When the camera coordinates move within the world coordinate system, the space point of the three-dimensional space can be restricted on a two-dimensional line. When there're two camera coordinate systems, the location of any space point can be obtained by calculating the intersection point of the two lines where the point stands. In order to realize the three-dimensional reconstruction mentioned above, camera calibration problem and stereo matching problem need to be solved [12]. The goal of the former is to obtain the equations of the two space rays while the goal of the latter is to match two pixels to one point.

2.2 Camera Calibration and Stereo Matching

Camera calibration process is the process of confirming camera projection matrix [14]. As there's randomicity contained in the angle selection and location selection in the shooting process of camera[16], the relationship between the camera coordinate system and the world coordinate system cannot be assured. The key is how to describe the relationship between these two systems in a more convenient way.

This paper took a rotation matrix $\mathbf{R}$ as an example to describe the angle selection of camera and used translation vector $\mathbf{t}$ to describe the location selection of camera. By this time, the relationship between the coordinates of P —— $P(x_w, y_w, z_w)$ $P(x_c, y_c, z_c)$ ——is shown in Equation (3). The $\mathbf{R}$ in Equation (3) refers to a 3×3 orthogonal rotation matrix; $\mathbf{t}$ refers to a 3×1 translation vector; $\mathbf{0} = (0,0,0)$; and $\mathbf{M}$ refers to a 4×4 matrix.

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} \mathbf{R} & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} = \mathbf{M} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix} \tag{3}$$

Based on the equation above, the perspective projection relationship between the camera coordinate of the object and the image coordinate [17] can be presented as Equation (4):

$$z_c \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} \tag{4}$$

Camera calibration steps are as follows:

STEP1. Select a standard test object and confirm a proper world coordinate system. Mark several test points on the standard test object.

STEP2. Confirm the three-dimensional coordinate of each test point. Confirm the pixel coordinate of each test point in the image by means of man-computer interaction.

STEP3. Obtain N equations of the 12 components in the $\mathbf{M}$ matrix. Use the least square method to obtain

each component of **M** matrix and complete the camera calibration.

The key part of stereo matching is to assign a pixel $p_1(u_1,v_1)$ and its pairing pixel point $p_2(u_2,v_2)$ respectively in the left figure and the right figure of Figure 1, so as to form a stereo point pair. If the stereo point composed of pixel pint $p_1(u_1,v_1)$ and pixel point $p_2(u_2,v_2)$ can match the same space point $P(x_w,y_w,z_w)$, it can be mapped to the original space point.

Based on the stereo matching realization of epipolar constraint, the imaging matrixes of the left camera and the right camera M_1 and M_2 need to be set as $M_1=[M_{11},m_1]$ and $M_2=[M_{21},m_2]$, among which both M_{11} and M_{21} are 3×3 matrixes while both m_1 and m_2 are 3×1 matrixes. The projection equations of the two cameras are shown in Equation (5):

$$\begin{cases} z_{c1}U_1 = M_1P_w = (M_{11} \quad m_1)P_w \\ z_{c2}U_2 = M_2P_w = (M_{21} \quad m_2)P_w \end{cases} \quad (5)$$

$P_w = [x_w \quad y_w \quad z_w \quad 1]^T$ in Equation (5) refers to the homogeneous coordinates of the space point P and $U_1 = (u_1 \quad v_1 \quad 1)^T$ refers to the image coordinates of one point in the left image while its dual point coordinates in the right image are $U_2 = (u_2 \quad v_2 \quad 1)^T$. If there's $T = (x_w,y_w,z_w)^T$, the extreme line equation shown in Equation (6) can be obtained from the relation described in Equation (5):

$$U_2^T\left[(m_2 - M_{21}M_{11}^{-1}m_1)\times(M_{21}M_{11}^{-1}U_1)\right] = 0 \quad (6)$$

The equation set of cancellation z_{c1} and z_{c2} shown in Equation (7) can be obtained by method of elimination. The three-dimensional coordinates of P which are (x_w,y_w,z_w) can be calculated by the equation set.

$$\begin{bmatrix} u_1m_{31}^1-m_{11}^1 & v_1m_{31}^1-m_{21}^1 & u_2m_{31}^1-m_{11}^1 & v_2m_{21}^1-m_{21}^1 \\ u_1m_{33}^1-m_{12}^1 & v_1m_{32}^1-m_{22}^1 & u_2m_{32}^1-m_{12}^1 & v_2m_{32}^1-m_{22}^1 \\ u_1m_{33}^1-m_{13}^1 & v_1m_{33}^1-m_{23}^1 & u_2m_{33}^1-m_{13}^1 & v_2m_{33}^1-m_{23}^1 \end{bmatrix}^T$$

$$\times\begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} m_{14}^1-u_1m_{34}^1 \\ m_{24}^1-v_1m_{34}^1 \\ m_{14}^2-u_2m_{34}^2 \\ m_{24}^2-v_2m_{34}^2 \end{bmatrix} \quad (7)$$

3 THREE-DIMENSIONAL RECONSTRUCTION OF TWO-DIMENSIONAL SPORT DEMONSTRATION VIDEO IMAGES

As the three-dimensional reconstruction of two-dimensional sport demonstration videos aims at conducting three-dimensional decomposition analy-

sis of sport motions, so as to provide comprehensive sport guidance, this paper took diving as the example to conduct reconstructing analysis of diving. It studied the sport demonstration system based on three-dimensional reconstruction in order to provide theoretical guidance for the three-dimensional reconstruction of two-dimensional video images.

3.1 Process Design of Three-Dimensional Reconstructed Sport Demonstration System

First, the system is composed of analysis module, 3D sport simulation, interface module and database management module. The application target of analysis module is to effectively segment videos and extract the pixel data information from segmented video images. The application target of 3D sport simulation is to display the processed two-dimensional images in a three-dimensional way. The application target of interface module is to realize image editing. The application target of database management module is to realize the effective receiving, issuing and storage of image data images. Based on the above, Figure 2 shows the diving sport demonstration system process designed in this paper on basis of three-dimensional reconstruction:

Description of the system design process is as follows:

1) Import the existing diving demonstration video to the main interface and segment the video.

2) Conduct sport tracking based on video segmentation. The target of the segmentation is to obtain the kinematic parameters of each arthrosis point of the moving object.

3) Build mathematical models of kinematic parameters; establish the kinematic parameters; and build the calculating models of posture parameters.

4) Drive the 3D virtual human with a set of kinematic parameters obtained by tracking. Use the virtual human to simulate the whole process of athlete's demonstration motions.

3.2 Video Segmentation, Model Tracking and Kinematic Simulation

Taking the features of diving demonstration videos into consideration, this paper applied a three-parameter camera kinematic model to rapidly obtain the overall kinematic parameters. It used a binocular camera and established a skin color model for processing. This paper extracted the athlete's two-dimensional images in the current frame and in the previous frame, and then made closing operation on the two images to obtain the contiguous images. In the next step, it did Macroblock matching to the images of which most foreground images had been removed. See Figure 3 for the segmentation results of the diving demonstration kinematic test sequence.

The purpose of video segmentation is to lay a foundation for tracking. By video tracking, the kinematic parameters of the moving object's each arthrosis point can be effectively obtained from the original
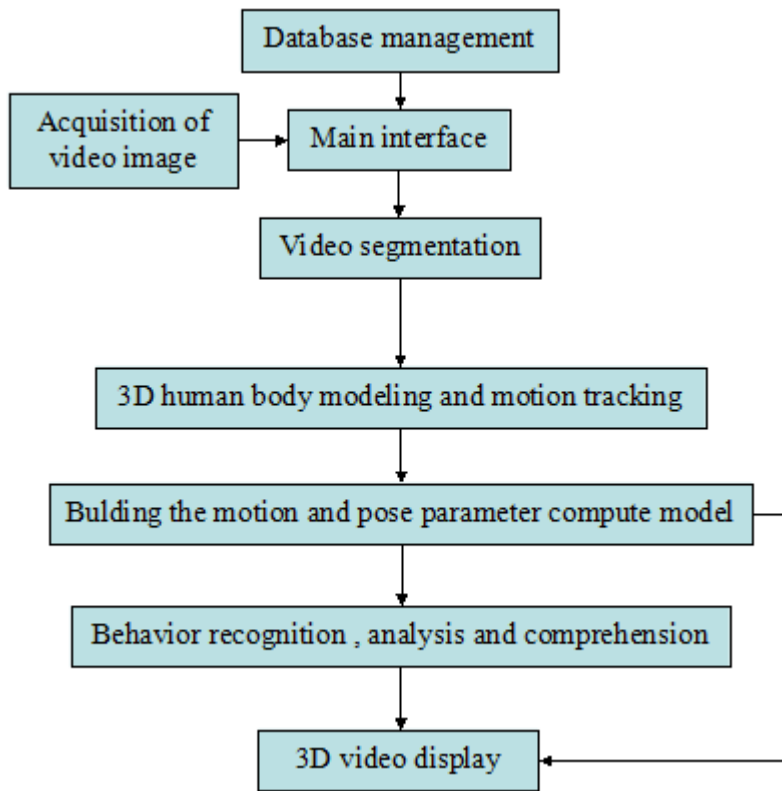
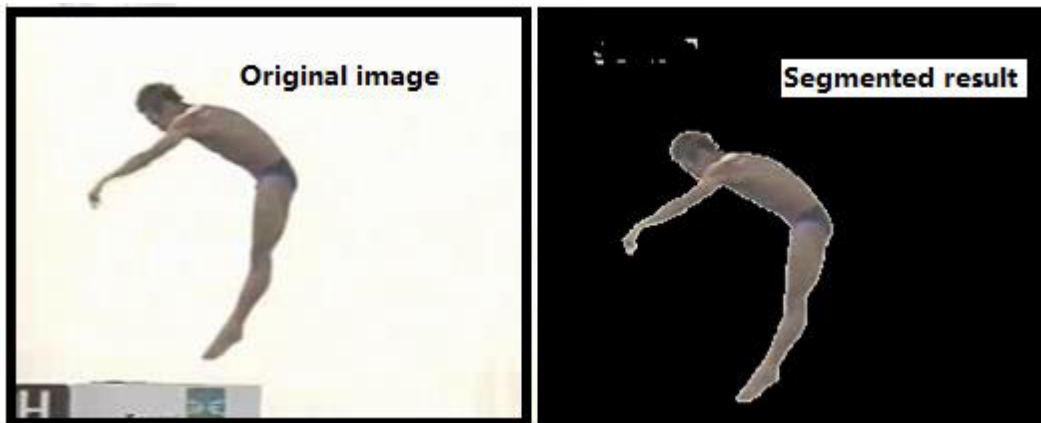Figure 2. Design process of sport demonstration system

Figure 3. Comparison diagram of the diving demonstration kinematic test sequence before and after segmentation

video images. See Figure 4 for the human skeleton pattern established in this paper.

In video tracking, the kinematic model can be used to predict the posture of the next frame based on the current frame. By applying the simple second-order autoregressive model presented in Equation (8), this paper assumed that each status parameter would vary by constant accelerated velocity:

$$x_t - x_{t-1} = x_{t-1} - x_{t-2} + \eta w_t \qquad (8)$$

The ηw_t in the equation refers to noise component and w_t complies with the Gaussian distribution of zero mean value [20].

The key of kinematic simulation is to obtain the kinematic parameters of each arthrosis point during the human kinematic process. This paper applied hu-
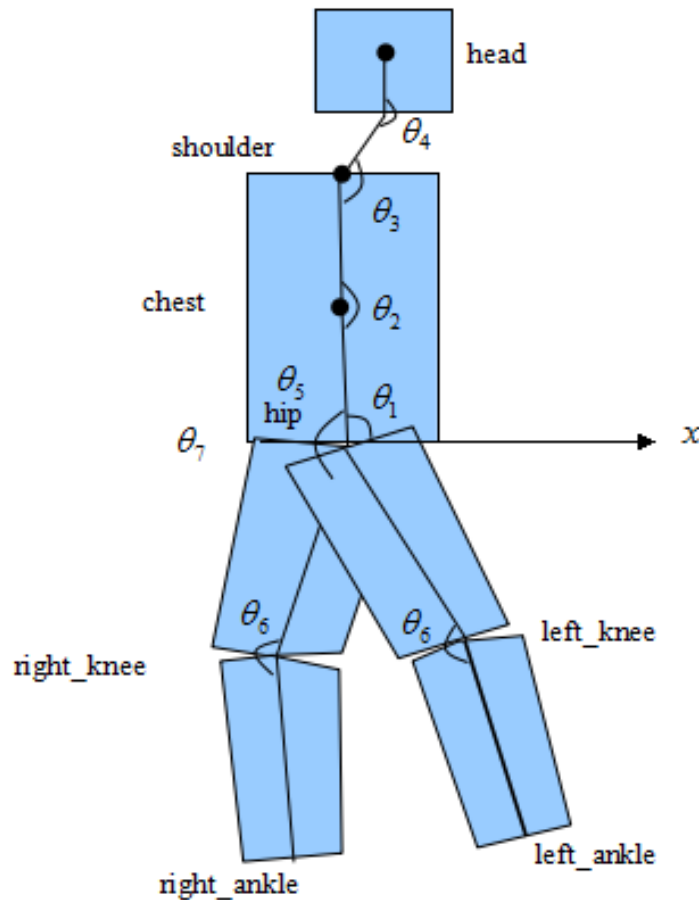
Figure 4. Human skeleton model [19]

man skeleton model to conduct three-dimensional reconstruction of image sequence. According to human characteristics, it is assumed that the segment length between the athlete's two arthrosis joints during diving will remain the same as shown in Equation (9). In the equation, $\mathbf{r_{ij}}$ refers to the vector quantity between arthrosis point p_i to arthrosis point p_j; l_{ij} refers to the segment length between arthrosis i and arthrosis j; and $p_i(x_i,y_i,z_i), p_j(x_j,y_j,z_j)$ respectively refer to the three-dimensional coordinates of the two arthrosis points.

$$r_{ij}^T r_{ij} - l_{ij}^2 = 0$$

$$\chi \qquad\qquad (9)$$

$$(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2 - l_{ij}^2 = 0$$

As human structure is very complicated, the following steps shall be obeyed in order to obtain more precise simulating demonstration motions of diving. Firstly, establish a database of elite athletes' motions by motion capturing. Secondly, complete the editing of the motions by kinematic editing techniques, such as displacement mapping and motion tiling; and then provide kinematics verification and analysis of the revised motions. At last, conduct comparison analysis

of the simulation motions and the athletes' training videos on the same screen.

In order to reach the goal of realizing visualized interactive motion design in a more convenient way, it shall be assumed that for the human motion $motion(t)$, the posture $posture(t_i)$ needs to be modified to obtain a new posture $posture'(t_i)$. Mouse can be used to select and drag the corresponding the rigid body of skeleton on the $X \times Y$ user window. Assume that the variation quantities of the mouse along the horizontal direction and along the vertical direction are $\Delta x, \Delta y$ respectively during dragging. According to euler therein, the eulerian angles $\langle \alpha, \beta, \gamma \rangle$ of the three coordinate directions can be used to present this rotation, so as to confirm the new posture. The calculation of the three angles mentioned above is shown in Equation (10) in which a, b, c refer to the influence degree that $\Delta x, \Delta y$ has on the eulerian angles $\langle \alpha, \beta, \gamma \rangle$ along the three directions.

Use displacement mapping technique to obtain new motions and estimate human's inertial parameters according to the athletes' physical parameters. Verify the reasonability of the new motions by back tracking. If the first step is started from the example motion
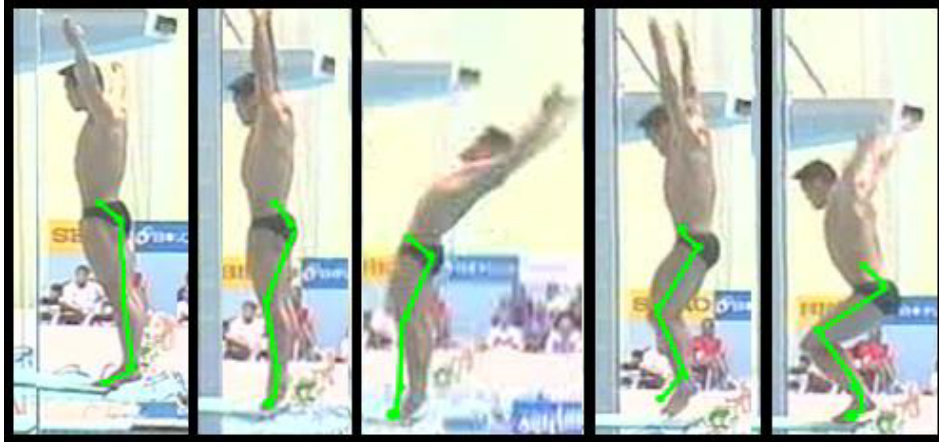
Figure 5. Binocular Camera Video Tracking Results of the Athlete during Diving
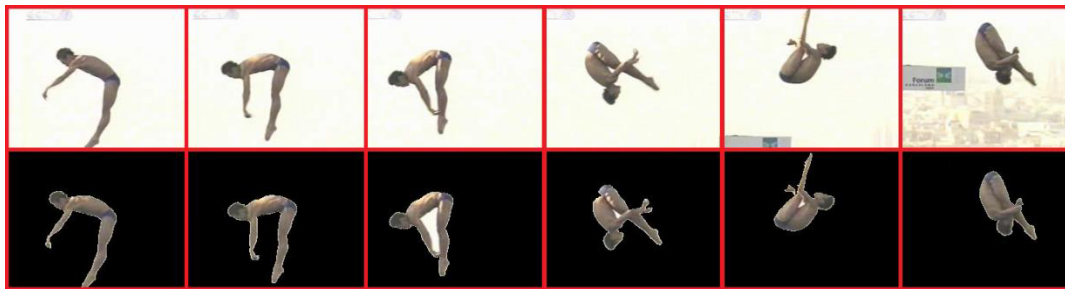

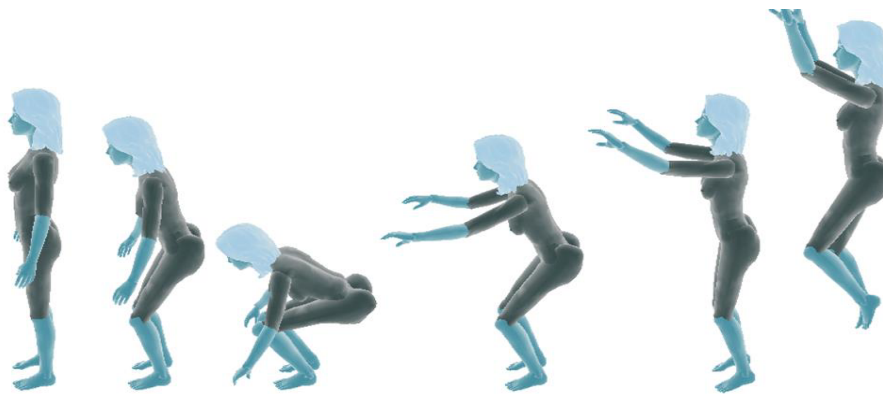Figure 6. Effect contrast of diving motion video segmentation

Figure 7. Video tracking results of the virtual human jumping process made in three-dimensional econstruction

database, the "demonstration" motions which are related to the athletes' single skilled motion can be easily established. Video files are the main storage media of athletes' motion data [21]. They video files can present the geometric attributes of sport motions. However, the analogue simulation results of three-dimensional reconstruction must reflect the physical attributes of the actual motions. Therefore, three-dimensional simulating motions need to be es-tablished on the same screen for the diving kinematic videos.

$$\begin{cases} \alpha = \arcsin \dfrac{a\Delta x + (1-a)\Delta y}{X+Y} \\ \beta = \arcsin \dfrac{b\Delta x + (1-b)\Delta y}{X+Y} \\ \gamma = \arcsin \dfrac{c\Delta x + (1-c)\Delta y}{X+Y} \end{cases} \quad (10)$$

3.3 *Experimental Results and Analysis*

See the tracking results of diving kinematic videos in Figure 5. It can be seen that the video tracking effect is good. Furthermore, the comparison diagram of video segmentation results as shown in Figure 6 can show that the method of using binocular cameras and establishing skin color models to extract motion foreground is very applicable in video segmentation.

As shown in Figure 7, the results are tracked by the virtual human's jumping video completed by the virtual human file which can meet H-anim standard. In the actual operation interface, visual angles can be adjusted by the operator's mouse operation, so as to provide the virtual human's three-dimensional dynamic in an all-around way.

4 CONCLUSION

The motion segmentation of two-dimensional sport videos and the three-dimensional reconstruction of the demonstration motions can help improve the effect of image data in sport guidance and thus to help coaches and students build deeper understanding of sports in an all-around way. This paper analyzed the three-dimensional reconstruction of two-dimensional images in theory and analyzed the camera calibration technique and the stereo matching technique, providing technical support and theoretical foundation for the three-dimensional reconstruction of two-dimensional motion videos. This paper took diving as the example and has designed the sport demonstration system composed of analysis module, three-dimensional simulation, interface module and data management module on basis of three-dimensional reconstruction. See the conclusions reached in this paper as follows:

1) Two-dimensional digital video files are actually ordered arrangement of two-dimensional images. The key of establishing three-dimensional sport demonstration videos of two-dimensional video files lies in the segmentation of two-dimensional videos and the three-dimensional reconstruction of two-dimensional images.

2) The key technologies contained in the three-dimensional reconstruction of two-dimensional digital images are the calibration of internal and external camera parameters and the stereo matching technique. The content processed in the former technology is to obtain accurate space ray equation while that of the latter technology is to realize the precise matching of corresponding pixels.

3) In this paper, the applied segmentation technology of two-dimensional video files was visual tracking method. Aiming at completing the whole process of the 3D virtual athlete's demonstration motions, the visual tracking method can be applied for obtaining the three-dimensional virtual human's kinematic parameters.

4) During the process of the three-dimensional reconstruction for the three-dimensional motion demonstration videos of diving, the editing and calibration of the videos are necessary steps. This paper applied displacement mapping technique to obtain the calibrated videos and complete the three-dimensional simulation demonstration motions displayed on the same screen.

REFERENCES

[1] Zhang D.H. 2013. Computerized Automatic Recognition of the Three-Dimensional Reconstruction in Sport Technique Analysis. *Bulletin of Science and Technology*. 29(6):146-148.

[2] Wang Z.R., Zhou Y.B. & Li B. 2015. The Three-Dimensional Reconstruction Methods Based on Linear Structure-Light and Sport Platform. *Tool Engineering*. 49(1):81-83.

[3] Yang L., Li G.J. & Wang L.R. 2015. Multi-Sequencing Rectification Facing Scene Reconstruction. *Optics and Precision Engineering*. 23(2):557-565.

[4] Zhang X.F., Wang Z.L. & Yuan W. 2013. The Three-Dimensional Reconstruction of the Billiard Table in Binocular Vision. Computer Systems & Applications. 22(3):173-176.

[5] Wang J., Wen Y.Y., Yu J. & Cai D.D. 2013. A Three-Dimensional Reconstruction Method Based on the ISAR Image Sequence of the Moving Object. *Journal of System Simulation*. 25(4):809-816.

[6] Chen Q. 2015. The Three-Dimensional Reconstruction Based on Binocular Vision. *Modern Computer (Professional Edition)*. 05(02):32-37.

[7] Yan L. 2011. The Simulation Technique of Camera Imaging Process. Journal of Shandong University (Engineering Edition). 12(03):51-54.

[8] Tanng L.M. & Zhang T.H. 2010. Target Imaging Simulation Based on OpenGL. *Scientific & Technology Information*. 32(15):122-125.

[9] Ke C.Z. & Xu P. 2014. Realization of a Binocular Camera Based on Three-Dimensional Vision. *Innovation and Application of Science and Technology*. 12(30):12-15.

[10] Lu Y., Wang H.Q., Tong W. & Li J.J. 2014. The Camera Calibration Algorithm Based on Harris Zhang Zhengyou Plane Calibration. Journal of Xi'an University of Architecture and Technology (JCR Science Edition). 35(06):72-76.

[11] Zhu Q.S., Luo D.J., Ge L. & Liu J.F. 2010. The Three-Dimensional Reconstruction Based on Multi-Images. *Computer Engineering and Design*. 17(10):131-135.

[12] Lin J., Jiang K.Y., Lin J.Y. & Li L.J. 2013. The Fast Stereo Matching of Phase Position Based on Three-Dimensional Rectification. Journal of Guizhou Normal University (JCR Science Edition). 14(02):28-31.

[13] Xiahou K.S., Chen S.X. & Wu Y.L. 2015. The Modeling and Simulation of the Target Tracking System Based on Binocular Platform Camera. *Journal of System Simulation*. 01(02):44-47.

[14] Li M., Zhou B., Meng Z.D. & Dai X.Z. 2015. *Study of the Calibration of Ternary Stereo Camera*. 03(02):17-21.

[15] Zhi J.H., Dong X.M., Kong X.W. & Wang X.F. 2014. Analysis of the Outside Influential Factors in Camera Calibration. *Applied Optics*. 2014.22(02):86-89.

[16] Zheng F.J., Yu T., Yuan G.T. & Lan J.J. 2010. Survey of Camera Geometric Calibration Methods. *Taiyuan Science and Technology*. 33(02):131-134.

[17] Hu G.Y., He P.A., Wang B.L. & Xun J.P. 2004. The Camera Calibration Problems in Optical Measurement. *Optics & Optoelectronic Technolog*y.26 (03):42-45.

[18] Xiong Y., Gong H.J. & Shen Y.Q. 2007.Study of the Digital Camera Parameter Calibration Methods in Three-Dimensional Reconstruction. *Computer Simulation*. 35(09):72-78.

[19] Wang Z.Q., Zhang Y.D. & Xia S.H. 2005. System of Three-Dimensional Human Motion Simulation during Physical Training and Video Analysis. *Journal of Computer Research and Development*. 42(02):344-352.

[20] Yang F., Ding L., Yang C.X. & Yuan X.G.2005. The Application of Image Sequence and Human Body Model Reconstruction in Three-Dimensional Human Motion. *Journal of Biomedical Engineering*. 22(2):307-311.

[21] Liu G.Y., Chen R., Deng Y. & Li H. 2006. The Reconstruction and Simulation of the Three-Dimensional Human Body Model in Weight Lifting. *Journal of System Simulation*. 18(12):301-305.