A Virtual Modeling Method of Digital Media Image Synchronization Based on Motion Hybrid Algorithm

Yanyan Yang*, Zhiping Wang, Caixia Yang, Huijun Zhu
Nanyang Institute of Technology, Nanyang 473004, China
*Corresponding author: xuezhi_he@126.com

Abstract. In virtual reality digital media, motion capture system is used to build basic motion library. Then the basic motion is processed by motion editing technology. Motion blending is one of the most practical and complex editing techniques. This paper proposes a real-time synchronization algorithm based on motion blending, so as to better mix motion dynamically, avoid unexpected effects, and create complex virtual digital media animation. This paper adopts the strategy of data and model hybrid drive. The motion generation and control technology of virtual human is studied from the following aspects: Modeling and simulation method of motion system, grasping of virtual human considering the change of whole body posture, fast generation of operation action, automatic interactive generation of virtual human's key frame posture, multi priority editing synthesis and interactive control of virtual human's whole body motion. In this paper, the corresponding control strategy and model are proposed, and the traditional algorithm is improved. Experimental results show that the method improves the efficiency and accuracy of digital media generation. This method provides a reference for the research of digital media image synchronous virtual modeling method based on motion hybrid algorithm.

Keywords: Virtual Reality, Digital Media, Motion Capture, Motion Hybrid Algorithm.

1. Introduction
In distributed video coding (DVC), the quality of side information plays an important role [1-2]. The higher the quality of side information, the less parity bits are needed for decoding, and the smaller the bit rate is. When reconstructing WG frames, the higher the prediction accuracy of side information, the better the reconstruction quality and the lower the rate distortion (RD) [3]. Therefore, how to accurately generate side information at the decoding end is an important problem and a technical problem faced by distributed video coding.

MCTI is the most typical method to generate edge information, which mainly includes three steps: motion estimation, motion vector correction and motion compensation interpolation [4-5]. How to propose an improved algorithm to improve the quality of edge information has been a hot topic for scholars at home and abroad. On the basis of MCTI, many scholars have proposed many improved interpolation and extrapolation edge information generation algorithms [6-8]. These algorithms make use of motion block, sub-pixel, noise relationship between edge information and decoded frame and original frame to improve the quality of edge information significantly. However, these edge
information algorithms are all improved for symmetric motion vectors. They are all linear motion estimation. When there are a large number of macroblock regions in the actual video sequence that do not meet the requirements of linear motion estimation, the edge information generated by linear motion estimation has a large error. To solve the above problems, this paper proposes an edge information generation algorithm, which uses a mixed motion model. When the video sequence motion is regular, linear function is used for motion estimation. When the motion of video sequence is irregular, curve motion function is used instead of linear function to adjust the motion vector, so that the adjusted motion vector is more in line with the real motion situation, and can generate high-quality edge information, so as to improve the rate distortion performance of the system.

2. Motion model
When the video sequence motion is regular, the linear motion model is used to generate the edge information. But when the video sequence motion is irregular, the motion vector will have deviation, and the linear motion model will not be accurate to adjust. This requires other motion models to replace the linear motion model. This paper uses the curve motion model. The linear motion model will not be introduced here, but the curvilinear motion model will be briefly described.

2.1. Curvilinear motion hypothesis
According to the curvilinear motion hypothesis theory proposed in the literature, the curvilinear trajectory of the motion macroblock is shown in Figure 1.

![Figure 1 Curve trace of motion macroblock](image)

How to determine the curvilinear trajectory of a moving macroblock is a key problem. It can be seen from the curve trajectory diagram of the motion macroblock in Fig. 1 that the curve trajectory of the macroblock is modeled as a function of the plane, and the macroblock searches for the motion estimation in the adjacent K frames to obtain the coordinate points \((X_1, Y_1), (X_2, Y_2), (X_3, Y_3)\) of the similar macroblock, so as to use these coordinate points to determine the parameters of the function and obtain the curve trajectory function.

2.2. Curvilinear motion function
There are many motion functions, such as quadratic function, exponential function, logarithmic function, cubic function and so on. This paper takes quadratic function and exponential function as examples to illustrate. To illustrate the curvilinear motion function, the motion vector estimation given in Fig. 2 is used.

In fig. 2, A0 is the wz frame to be tested, a1 is the decoded key frame \(K_{2n+1}\), A2 is the decoded key frame \(K_{2n-1}\), A2 is the reference frame of A1, and a1 is the reference frame of A0. Block f1 moves to block A of WZ frame A0 via motion vector \(MV_0\), and block f2 moves to block E of WZ frame A0 via motion vector \(MV_3\). The middle block between f1 and f2 is f4, and the motion vector of block f4 is \(MV_4 (x_4, y_4)\).
Figure 2 Motion vector estimation

(1) Quadratic function
Suppose that the vertical direction of the image satisfies the quadratic function, and the horizontal direction also satisfies the quadratic function, that is, the macroblock satisfies the quadratic function \( f(y) = ay^2 + by + c \) in the vertical direction. And \( f(0) = y_0, f(3y_0/2) = y_4, f(2y_0 - y_3) = y_3 \), then the motion vector of block \( f_0 \) is \( MV(f(x_0), f(y_0)) \). Where \( f(x_0) \) and \( f(y_0) \) are shown in equations (1) and (2) [9-12]:

\[
\begin{align*}
  f(x_0) &= \frac{y_0}{3} \left( \frac{1}{2} x_0^2 - \frac{1}{2} x_0 x_j + 2x_0 x_j - x_j x_i \right) \\
  f(y_0) &= \frac{y_0}{3} \left( \frac{1}{2} y_0^2 - \frac{1}{2} y_0 y_j + 2y_0 y_j - y_j y_i \right)
\end{align*}
\]

(2) Logarithmic function
It is assumed that the vertical direction in the image satisfies the logarithmic function, and the horizontal direction also satisfies the logarithmic function, that is, the macroblock satisfies the exponential function \( f(y) = a10^{by} \) in the vertical direction. There are \( f(0) = y_0, f(2y_0 - y_3) = y_3 \) in the vertical direction and \( f(0) = x_0, f(2x_0 - x_3) = x_3 \) in the horizontal direction. The motion vector of block \( f_0 \) is \( MV(f(x_0), f(y_0)) \), where \( f(x_0) \) and \( f(y_0) \) are as shown in formula (3) and formula (4):

\[
\begin{align*}
  f(x_0) &= x_0 \log_2 \left( \frac{x_j}{x_0} \right) \\
  f(y_0) &= y_0 \log_2 \left( \frac{y_j}{y_0} \right)
\end{align*}
\]

When \( 2x_0 - x_3 = 0 \) or \( x_0 = 0 \) in equation (3), the horizontal component of MV is replaced by \( x_0 \); when \( 2y_0 - y_3 = 0 \) or \( y_0 = 0 \) in equation (4), the horizontal component of MV is replaced by \( y_0 \), that is, the motion vector of block \( f_0 \) is \( MV(x_0, y_0) \).

3. Edge information generation algorithm based on hybrid motion model
Combined with MCTI algorithm, on the basis of symmetric motion vector (SMV) hypothesis (that is, linear motion) and curve motion model (quadratic function and exponential function are used in this paper), an edge information generation algorithm based on mixed motion model is proposed.
3.1. Algorithm description
Based on MCTI algorithm, this algorithm adds a detection module to detect whether the motion vector meets SMV by setting a decision threshold and comparing the size of SAD with the threshold. If it does, it directly processes the motion vector to obtain side information. If not, search again, and then compare with SAD value to introduce curve motion estimation and get edge information.

The specific steps of the algorithm are as follows:

In step 1, the input video sequence is divided into two groups (grouped according to parity): one is keyframe / K frame (specified as odd frame, represented by X_{2i-1} and X_{2i+1}), and the other is WZ frame (specified as even frame, represented by x_{2i}). The k-frame is processed by the traditional encoding and decoding method (generally H.264 encoding and decoding method), and then the decoded X_{2i-1} key frame and X_{2i+1} key frame are low-pass filtered, and go to step 2;

In step 2, the block matching based symmetric bidirectional motion estimation (SBME) is used to obtain the motion vector of X_{2i+1} frame relative to X_{2i-1} frame, and go to step 3;

In step 3, the sad of x_{2i} + 1 frame relative to x_{2i}-1 frame is calculated, and the threshold THR is set. If sad ≥ THR, go to step 5; otherwise, go to step 4;

Step 4: the motion vector satisfies the assumption of symmetric motion vector (SMV), that is, linear motion. Go to step 7;

In step 5, the best matching block is searched again, and the SAD_{2n+1} of X_{2i} frame relative to X_{2i} + 1 frame and SAD_{2n-1} of X_{2i+1} frame relative to X_{2i+1} frame are calculated. If SAD_{2n-1} ≤ SAD or SAD_{2n+1} ≤ sad, go to step 6; otherwise, go to step 4;

In step 6, the motion function is obtained by curve motion estimation, and the coordinate position of the macroblock (macroblock in WZ frame) is determined. Go to step 9;

Step 7 performs adaptive filtering on the motion vector, and goes to step 8;

In step 8, the block directed by the motion vector is expanded, and the overlapped bidirectional motion compensation is performed to generate the edge information;

Step 9 uses different weights to synthesize the edge information of different curves.

3.2. Algorithm flow
According to the specific steps of the algorithm described above, the flow of the algorithm is shown in Figure 3.

4. Experimental simulation results and analysis
In order to verify the performance of the algorithm proposed in this paper, three standard video sequences, Carphone, Foreman and Football, are used to test, among which Foreman sequence moves slowly. Moreover, both characters and shots shake, Carphone sequence moves generally, the background is mostly static, and football sequence moves violently. The video sequence format is QCIF (176×144), and the frame rate is 30. In the simulation experiment, the number of frames is selected as 100, and the threshold THR is set as 256. The PSNR value of the side information brightness and the code rate transmitted by the generated side information of Carphone, Foreman and Football are tested and simulated. The simulation results of PSNR values of side information brightness of three video sequences are shown in Figure 4.
It can be seen from fig. 4 that the PSNR value of Foreman sequence with gentle motion is increased by 0.44 dB compared with the traditional linear algorithm, and the PSNR value of the mixed algorithm with exponential function is increased by 0.17 dB compared with the traditional linear algorithm. For Carphone video sequences with moderate motion, the PSNR value of side information generated by the hybrid algorithm of quadratic function is 0.716 dB higher than that of the traditional linear algorithm, and the PSNR value generated by the hybrid algorithm of exponential function is 0.344 dB higher than that of the traditional linear algorithm. The PSNR value of the side information generated by the mixed algorithm of quadratic function is 1.14 dB higher than that of the traditional linear algorithm, and the PSNR value generated by the mixed algorithm of exponential function is 0.62 dB higher than that of the traditional linear algorithm for the football video sequence with intense motion.
motion.

It can be seen that the PSNR value of the edge information generated by the proposed algorithm is obviously improved for the video sequence with strong motion, which indicates that the algorithm is especially suitable for the video sequence with strong motion.

The simulation results of the bit rate needed to generate side information for three video sequences are shown in Figure 5.

(a) Comparison of code rates required by Foreman sequence to generate side information by using different algorithms

(b) Compare the code rates needed by Carphone sequence to generate side information by using different algorithms

(c) Comparing the code rates required by the foreman sequence to generate side information by using the identical algorithm

Figure 5 Three video sequences use different algorithms to generate the required bit rate of side information
As can be seen from Figure 5, for foreman sequences with gentle motion and background jitter, the bit rate required for the hybrid algorithm with quadratic function to generate side information is 0.00392 lower than that of the traditional linear algorithm, and the bit rate required for the hybrid algorithm with exponential function to generate side information is 0.00276 lower than that of the traditional linear algorithm. For moderate motion Carphone video sequences, the bit rate of the hybrid algorithm with quadratic function is 0.0063 lower than that of the traditional linear algorithm, and the bit rate of the hybrid algorithm with exponential function is 0.00292 lower than that of the traditional linear algorithm. For football video sequences with intense motion, the bit rate of the hybrid algorithm using quadratic function is 0.00798 lower than that of the traditional linear algorithm, and the bit rate of the hybrid algorithm using exponential function is 0.00404 lower than that of the traditional linear algorithm.

It can be seen from the test results that the mixed model algorithm using quadratic function and exponential function is more suitable for describing the motion trajectory of macroblock, and the edge information obtained is more accurate.

5. Conclusion
In this paper, an edge information generation algorithm based on hybrid motion model is proposed, which improves the quality of edge information by combining linear motion estimation with curve motion estimation. The proposed algorithm fully considers the situation of irregular motion video sequence, and detects by setting threshold. When the video sequence is in regular motion, linear motion model is used for linear motion estimation, while when the video sequence is in irregular motion, curve motion model is used for curve motion estimation. Experimental results show that the proposed algorithm can effectively improve the quality of edge information and significantly reduce the bit rate needed to generate edge information, especially for the video sequences with intense motion. This algorithm provides a theoretical basis for improving the performance of distributed video coding.

Acknowledgement
This work was supported in part by the Industrial Science and Technology Research Project of Henan Province under Grants 162102210358 and 212102310297.

References
[1] Huang Jin, Zhang Xiaobing, Yin Hanchun. 3d Measurement Method Based on Image Processing. Electronic Devices, 2002, 25 (4): 364-368
[2] Chen Bingquan, Liu Hongli, Meng Fanbin. Current Situation and Development Direction of Digital Image Processing Technology. Journal of Jishou University (natural Science Edition), 2009 (01): 63-70
[3] Yang Ming, Song Xuefeng, Wang Hong. Image Processing for Intelligent Transportation System. Computer Engineering and Application, 2001, 37 (9): 4-7
[4] Ji Shouwen, Wang Rongben, Chen Jiajuan. Identification of Weeds in Maize Seedling Field by Using Computer Image Processing Technology. Acta Agriculturae Sinica, 2001, 17 (002): 154-156
[5] Liu Zhonghe, Wang Ruixue, Wang Fengde. Current Situation and Prospect of Digital Image Processing Technology. Computer Era, 2005, 9: 6-8
[6] Zhang Xuming, Xu Binshi, Dong Shiyun. Adaptive Median Filtering for Image Processing. Chinese Journal of Computer Aided Design and Graphics, 2005, 17 (002): 295-299
[7] Cheng Zheng Xing, Lin Yong Ping. Application of Wavelet Analysis in Image Processing. Acta Mathematica Sinica Sinica, 2001, 18 (f12): 57-86
[8] Bai Lifen, Xu Yuxian, Yu Shui. Study on Auto Focusing Method of Microscope Based on Image Processing. Journal of Instrumentation, 1999 (06): 64-66
[9] Long Gang, Xiao Lei, Chen Xuequan. Overview of Curvelet Transform in Image Processing. Computer Research and Development, 2005,42 (8): 1331-1337
[10] Yuan Hongbo, Zhao Nudong, Cheng Man. Research Progress and Prospect of Field Weed Recognition Based on Image Processing. Acta Agriculturalis Sinica, 2020, V. 51 (s2): 330-341
[11] Zhang Guoping, Li Yali, Xu Xiangyi. Research on Application of Image Processing in Partial Discharge Ultraviolet Imaging Detection of Power Equipment. Computer Programming Skills and Maintenance, 2020, No. 425 (11): 146-148
[12] Jiang Haiyang, Wang Huiwen, Guo Chunrong. Research Progress of Tongue Image Processing Technology. China Science and Technology of Traditional Chinese Medicine, 2021, Vol. 28, No. 1, 164-166, ISTIC CA, 2021