Distributed Cluster Multidimensional Transcoding for Stereoscopic Video Stream

Fengfeng Duan[1,2,a]

[1]College of Computer Science, Communication University of China, Beijing 100024, China

[2]Cultural Resources Research and Development Center of Hunan, Hunan Normal University, Changsha 410081, China

[a]dffeng2010@126.com

Keywords: Multidimensional Transcoding; Stereoscopic Video; Distributed Cluster

Abstract: In order to improve the quality and efficiency of the distributed cluster collaborative production for video stream, a multidimensional transcoding framework of stereoscopic video is proposed. In the framework, the transcoding technologies, including decoding and encoding, are applied based on multiple levels. At the same time, a new scheduling algorithm is introduced for the transcoding orderly and efficiently. For the stereoscopic video transcoding effects evaluation, several methods are introduced and the experimental results show that the proposed method has high accuracy and efficiency.

Introduction

With the development of digital technology, the production and application of high-quality digital video becomes more widespread. However, complexity of production and conversion as well as the problem of large tasks are concerned. The development of distributed cluster technology and transcoding technology may be able to solve these problems. However, these technologies are not perfect. The innovation of the study is that the introduction of multidimensional transcoding and scheduling in distributed cluster production for stereoscopic video. At the same time, SSIM method is used to evaluate the quality of video.

The rest of this paper is organized as follows. In section 2, the description of distributed cluster collaborative production and multidimensional transcoding technology is given. In section 3, we present the architecture of multidimensional transcoding scheduling in distributed cluster system. Afterwards, the scheduling algorithm and process is described. The implementation and evaluation is presented in section 4. Finally we give a conclusion of this paper in section 5.

Distributed Cluster System and Multidimensional Transcoding

Distributed Cluster Collaborative Production for Stereoscopic Video Stream. Video production has experienced several stages, such as manual editing, electronic editing, computer non-linear editing, online and network editing. Nowadays, stand-alone or online collaborative editing technologies are used prevalently in video production. Recently, distributed cluster collaborative production technology develops rapidly. In distributed cluster system, the clients and workstations in WAN can take part in the collaborative work of designing, editing, rendering and publishing in real-time or non-real-time. At the same time, high-speed network communications and transmission, content integration and distribution can be realized through the network servers, as well as the task of production can be completed.

There are many disadvantages for distributed cluster collaborative production. These production methods are not perfect in time-consuming, efficiency, quality and so on. They won't be able to meet the needs of development in stereoscopic video production with the increasing of quality requirements, the expansion of the production range and the raising of complexity. Stereoscopic video format conversion is one of the most prominent problems. It is inevitable for stereoscopic video format conversion in distributed cluster collaborative production and we need encoding and decoding for different formats of stereoscopic video. Stereoscopic video transcoding technology is an effective measure in order to improve the efficiency of video format conversion and video quality.

Multidimensional Transcoding Framework and Technology. The standards of video encoding and decoding are varieties which result in many different formats. With the development and widespread of video transmission, application, manner of sharing in distributed cluster collaborative production, it is inevitable for stereoscopic video conversion among different formats. To achieve the conversion, the technologies of transcoding must be implied. The video transcoding can be divided into homogeneous transcoding and heterogeneous transcoding. The homogeneous video transcoding happens among the same format or the standards of same series, while the heterogeneous video transcoding occurs in different standards or formats. In [1], the application of heterogeneous video transcoding is discussed.

Video transcoding methods mainly including re-quantization to reduce the bit rate, the down-sampling to reduce the spatial resolution and drop B-frames to reduce the frame rate. The research on video transcoding gradually increased at home and abroad since 2000. In China the research is still in its initial stage with the study of technologies and the fields continue to expand. There are many articles on video transcoding and the research mainly about the architecture, key technologies, problems and related solutions, useful algorithms, and specific applications. However, the three review articles can give some ideas about these studies. In references [2, 3, 4], the points of transcoding techniques are described and analyzed, including bit-rate reduction, spatial resolution and temporal resolution. The standards transcoding and transcoding quality optimization are also expounded in some of them. Additionally, error-resilience video transcoding technologies gradually become a new research direction. All in all, these studies are mainly focus on one-dimensional transcoding.

There are many shortcomings for one-dimensional transcoding technology and we propose a multidimensional transcoding technology. Multidimensional transcoding is a perfect cascade for more than one-dimensional transcoding through specific methods and techniques. It can be constructed as an efficient codec. It can be modified and converted for video parameters in multiple dimensions, which are better able to achieve homogeneous or heterogeneous video transcoding.

According to [5], we propose an optimized framework for multidimensional transcoding. In this framework, we make a concatenation of bit-rate reduction, spatial resolution reduction, temporal resolution reduction and detail resolution reduction to constitute a transcoder. Bit-rate reduction can effectively reduce the bit rate (bandwidth savings), so that the output bit rate can match requirements in the network, wireless transmission or digital devices and storage applications. The drop rate transcoding can effectively reduce complexity and drift reduction. The spatial resolution is to reduce the resolution of each frame and can be achieved by down sampling algorithms. Temporal resolution can be realized through the frame dropping strategy, frame type or field frame conversion, the motion vector synthesis and optimization algorithm. The detail resolution is mainly to increase the quantitative level. We just give a practical case and transcoding dimension can be more than four or may also be less than four-dimensional. The following is implementation architecture of multidimensional transcoding.

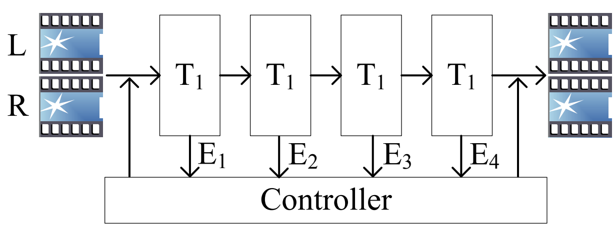

Fig. 1 Implementation architecture of multidimensional transcoding

As we can see from the transcoding architecture, T_i is passive transcoder modules and E_i is event. The controller is the active component. At the beginning, the original stereoscopic video is input into the transcoder and then decoded and encoded in T_1, T_2, T_3 and T_4 respectively. E_1, E_2, E_3, E_4 are the events which they produce and these events are transmitted to the controller. The controller can keep the operating state of the whole transcoder and give feedback or handling for the issues timely. Eventually it will generate the desired files.

We present the architecture of multidimensional transcoding above. The transcoding based on the bit rate, the spatial resolution, temporal resolution and detail resolution is very complex. The following diagram is the frame process of multidimensional transcoding and each stage of the frame status. In fact, it is the transcoding chains of frame transition.

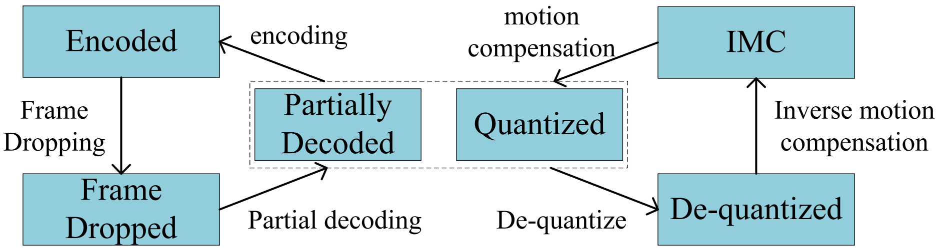

Fig. 2 Frame status of multidimensional transcoding

At the beginning, the DCT is implemented on reference frame. The decoding may become inaccurate, because the error blocks in reference frame encoded still depends on the calculation of the original DCT value and it can be avoided by drift free ring. I-frames can not be discarded as they are the key frames. B-frames are not used as reference frames, so the redundant B-frames can be discarded and the type needs to be decoded only. But if P-frames are discarded, the motion information will lose and the following motion vector P-frames and B-frames are invalid. In this case the motion vector can be re-defined. The full decoding and encoding should be avoided in the multidimensional transcoder to improve the efficiency. De-quantize and inverse motion compensation is help for eliminate errors to improve video quality [6].

Scheduling of Multidimensional Transcoding in Distributed Cluster System

The Architecture of Multidimensional Transcoding Scheduling. The function of transcoding scheduling can be realized in the cluster environment. However, the complexity of multidimensional transcoding in distributed cluster system and the issues of cascade compatibility may lead to a decline in efficiency. Therefore, we proposed a more effective and accurate multidimensional transcoding scheduling algorithm. The transcoding scheduling architecture is presented in Fig. 3.

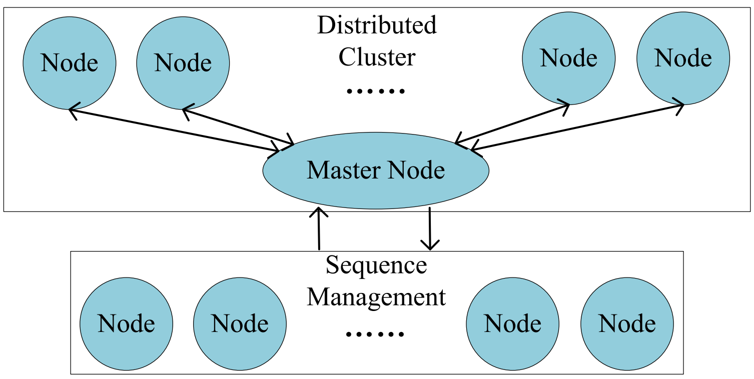

Fig. 3 The transcoding scheduling architecture

The Scheduling Algorithm and Process. The source sequence is divided into many sub-tasks according to the different particle size (GoP or Slice). A typical GoP frame structure is used [7]. The material must be analyzed to acquire the information of material format, including package format, video and audio encoding format and video parameters. The separation of video or the extraction of the key frames usually needed to get efficient scheduling.

Multidimensional transcoding in distributed cluster system is the cascade of several one-dimensional components with specific parameters. The implementation of transcoding tasks in the various components need to coordinate and follow the queue order. It is important to performing the resource scheduling according to the processing ability of the CPU and RAM. For example the fragment sequences can be described as $\{S_i\}=\{S_1,S_2,...,S_3\}$, and they are entered into multidimensional transcoder and executed. $\Delta U_{\mathrm{CPU}\_j}$ and $\Delta U_{\mathrm{RAM}\_j}$ are occupancy of CPU and RAM respectively for the sequence S_j when it is processed. γ_{CPU} and γ_{RAM} are weighted coefficients of RAM and CPU, respectively, which indicates the occupancy rate of them. ε_i is the environment level according to the processing node *i*. With the measurement and application of the three parameters dynamically and programming, the scheduling of transcoding is optimized [8]. The values of γ_{CPU} , γ_{RAM} and ξ_i can be defined respectively by [9]:

$$\gamma_{\mathrm{CPU}} = \frac{\left(\sum_{j=1}^{k} \Delta U_{\mathrm{CPU}\_j}\right) \Big/ k}{\left(\sum_{j=1}^{k} \Delta U_{\mathrm{CPU}\_j}\right) \Big/ k + \left(\sum_{i=1}^{k} \Delta U_{\mathrm{CPU}\_i}\right) \Big/ k}, \tag{1}$$

$$\gamma_{\mathrm{RAM}} = \frac{\left(\sum_{j=1}^{k} \Delta U_{\mathrm{RAM}\_j}\right) \Big/ k}{\left(\sum_{j=1}^{k} \Delta U_{\mathrm{RAM}\_j}\right) \Big/ k + \left(\sum_{i=1}^{k} \Delta U_{\mathrm{RAM}\_i}\right) \Big/ k}, \tag{2}$$

$$\xi_i = \frac{\left(\sum_{j=1}^{k} \Delta U_{\mathrm{CPU}\_j\_i}\right) \Big/ k + \left(\sum_{j=1}^{k} \Delta U_{\mathrm{RAM}\_j\_i}\right) \Big/ k}{\left(\sum_{j=1}^{k} \Delta U_{\mathrm{CPU}\_j\_s}\right) \Big/ k + \left(\sum_{i=1}^{k} \Delta U_{\mathrm{RAM}\_j\_s}\right) \Big/ k}. \tag{3}$$

With the frame dropping, the changing of spatial resolution and temporal resolution as well as improving of quantization level, the task is converted into video stream of another format. The results of transcode will be transmitted to the server after the transcoding tasks are finished by each node. The scattered video clips must be reassembled into a complete video file in server. At the same time it also should be packaged into a new multimedia material together with audio [10].

Simulation and Quality Assessment

Construction of Experimental Environment and Simulation. In the experiment, binocular stereoscopic videos 'Facial Masks', 'Decorated Archway' and 'Cheongsam' which are shot and produced by China Art Science and Technology Institute are applied as test sequences for implementation. The length, resolution, number of frames of each test sequence is shown in Table 1.The distributed cluster environment is built with Hadoop and it includes four nodes: a Master and 3 Salves. The nodes are connected by LAN. For each node, we use Windows 7 of 32-bit dual-core processor and 3.3GHz frequency. IP address of each node is shown in Table 2. The Xvid is used as transcoding software and the encoding and decoding can be realized. The scheduling algorithm proposed in part three is also implemented.

Table 1 Specific information of test sequence

Test sequence	Length (s)	Resolution	Frames
Facial Masks	375	720×576	9386
Decorated Archway	288	1920×1080	7215
Cheongsam	501	1920×1080	12544

Table 2 PC name and IP address in distributed cluster environment

PC name	IP address
Master.Hadoop	192.168.1.2
Salve1.Hadoop	192.168.1.3
Salve2.Hadoop	192.168.1.4
Salve3.Hadoop	192.168.1.5

Experimental Results. There are certain disadvantages and deficiencies in the assessment methods of Signal to Noise Ratio (SNR) and Peak Signal to Noise Ratio (PSNR). The picture quality is very low although the value is very effective in certain cases. So, Structural Similarity Index Measurement system (SSIM) is used as the evaluation method of video transcoding in the paper. The larger its value, the better the image quality, and the maximum value is 1. According to SSIM, structure information is defined as the realization of the structural theory. It reflects the structural attributes of the objects in the scene independent of the brightness and contrast. The distortion modeling is constructed with the combination of brightness, contrast and structure. x_i and y_i denote the pixel values before and after transcoding. The mean of luminance, standard deviation of contrast and covariance of structural similarity are defined respectively by:

$$\mu_x = \frac{1}{N}\sum_{i=1}^{N} x_i \, , \tag{4}$$

$$\sigma_x = \left(\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \mu_x)^2 \right)^{1/2} , \tag{5}$$

$$\sigma_{xy} = \left(\frac{1}{N-1}\sum_{i=1}^{N} (x_i - \mu_x)(y_i - \mu_y) \right) . \tag{6}$$

L is one of the dynamic range of the pixel values (255 for 8-bit grayscale images) and $K_1 \ll 1$. Then C_1, SSIM of x and y can be defined respectively by:

$$C_1 = \left(K_1 L\right)^2 , \tag{7}$$

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)} . \tag{8}$$

The value of SSIM relatively closes to 1 after video transcoding according to calculation of a large number of test sequences. It is more stability compared to the SNR and PSNR. The value of SSIM closer to 1, the better the image effect and the stability is the advantage of such evaluation mode. At the same time, the quality of stereoscopic video frames after transcoding is good enough and the efficiency is also high. The average percentage of pixels matching for transcoded test frames with original frames and time consumption are shown in Table 3.

Table 3 The average percentage and runtime of pixels matching

Test sequence	Matching percentage (%)	Runtime (s)
Facial Masks	88.2	0.26
Decorated Archway	90.6	0.11
Cheongsam	87.5	0.37

The stereoscopic video sequence frames of mpeg are converted into the sequence frames of H.264 through the multidimensional transcoding in distributed cluster system. The images before and after transcoding are shown in Fig. 4. We found that the image quality does not distort significantly by subjective observation and the image quality close to the original.

a. Effects before transcoding (left video frames of Decorated Archway)

b. Effects after transcoding (left video frames of Decorated Archway)

Fig. 4 Effects Comparison before and after Transcoding

Conclusion & Future Work

The method that we proposed is effective according to analysis as well as experimental results. It is necessary to build and use multidimensional transcoding technology and properly construct CPU resource scheduling algorithm for multidimensional transcoding scheduling in distributed cluster system. If so, distributed transcoding can be efficiently implemented. The higher quality video and more efficiency of task completion can also be acquired. Future work of the paper will focus on the improvement of scheduling algorithm.

Acknowledgements

This work was financially supported by the National Key Technology Research and Development Program of the Ministry of Science and Technology of China (2012BAH37F02) and the Base Project of Hunan Province Social Science (14JD38).

References

[1] T. Shanableh, M. Ghanbari: Heterogeneous video transcoding to lower spatial-temporal resolutions and different encoding formats. IEEE Transaction on Multimedia, vol. 2, no. 2(2000), p. 101-110.

[2] R. Choupani, S. Wong, M. Tolun: Video coding and transcoding: a review. Computer Engineering Laboratory on http://ce.et.tudelft.nl.

[3] J. Xin, C. W. Lin, M. T. Sun: Digital video transcoding. Proceedings of the IEEE, vol. 93, no.1(2005), p. 84 -97.

[4] A. Vetro, C. Christopoulos, H. Sun: Video transcoding architectures and techniques: an overview. IEEE Signal Processing Magazine, vol. 20, no. 2(2003), p.18-29.

[5] J. Brandt, L. Wolf, P. Halvorsen: Multidimensional transcoding for adaptive video streaming. ACM SIGMultimedia Records, vol.1, no.4(2009), p.16-17.

[6] Yong Ju Jung, Truong Cong Thang, Yong Man Ro: Distortion measures in MPEG-compressed domain for multidimensional transcoding. 2005 IEEE 7th Workshop on Multimedia Signal Processing, (2005), p.1-4,.

[7] Yong Ju Jung, Yong Man Ro: Joint control for hybrid transcoding using multidimensional rate distortion modeling. International Conference on Image Processing, vol.4(2004), p.2789-2792.

[8] YongQing Liang, Yap-Peng Tan: A new content-based hybrid video transcoding method. International Conference on Image Processing, vol.1(2001), p.429 - 432.

[9] Zhiyi Tan，Jianxin Song: Research on scheduling algorithms for distributed video transcoding service. Video Engineer, vol.36, no.9(2012), p.112-116.

[10] Y. Wang, S.F. Chang, A. C. Loui: Subjective preference of spatio-temporal rate in video adaptation using multi-dimensional scalable coding. IEEE International Conference on Multimedia and Expo, vol.3(2004), p.1719-1722.