Optimal Control for Correlated Wireless Multiview Video Systems

Yi Chen[1(✉)] and Ge Gao[2]

[1] School of Computer Science, Huazhong Normal University, Wuhan, Hubei, China
`chenyi30@mail.ccnu.edu.cn`
[2] School of Computer Science, Wuhan University, Wuhan, Hubei, China
`gaoge@whu.edu.cn`

Abstract. Emerging multimedia Multiview video systems consist of a dense deployment of multiple partial-overlapped wireless cameras, as well as some access points (Aps) and many wireless distributed relay nodes. Correlated views are captured by cameras followed being transmitted to destination by different Aps and networks links. Packet expiration of one camera flow may harm the whole task. To effectively integrate multiple viewpoints into a whole image, the correlated data rate and deadline of flows from multiple cameras are meaningful. There is a trade-off between data redundancy and time deadline among correlated multi-views subjecting to the constraints of limited buffer length. However, most researches in this field have not considered packet expiration suffering from varieties of delays after multipath. In this paper, we conduct this problem to optimally adjust multiple flows of viewpoints by exploiting spatial and temporal correlations among cameras to reduce delay variances. A global optimization algorithm based on joint rate-distortion and delay-distortion model is proposed. Simulation results show that quality of service for Multiview streaming can be improved by allocating suitable transmission rates among correlated cameras as well as appropriate playout deadline. The PSNR quality shows that better performance can be achieved compared with baseline policies.

Keywords: Correlation · Multiview streaming · Packet scheduling · Delay

1 Introduction

In recent years, the progress in multimedia technology has given rise to the demand for Multiview wide-area video applications over wireless networks. We consider a system that generates a wide area scene by acquiring images and depth information from many different viewpoints. However, network is unable to transmit all video packets from each camera. To sure a satisfactory viewing experience for end users, it is important to consider temporal and spatial masking effects to help compressed code work optimally that result in efficient usage of network resource for video service with cost-based quality. Our goal is to improve multi-camera wide area system's performance with respect to network constraints.

S. Wang and A. Zhou (Eds.): CollaborateCom 2016, LNICST 201, pp. 462–467, 2017.
DOI: 10.1007/978-3-319-59288-6_42

For instance, Fig. 1 shows a wide area field image generated by combining sources from more than three cameras through wireless networks.

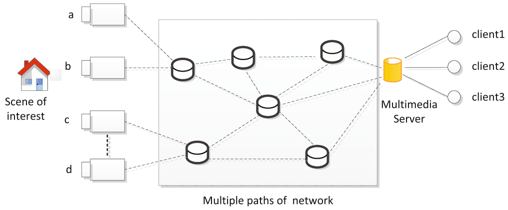


Fig. 1. An example of multimedia wide area view field system with multi-cameras

Resource allocation for video application has been an important research topic [1, 2], and there exists several works addressing the problem of scheduling of correlated video sources [3–5]. Some studies addressed the problem of Multiview transmission and view-switching delay [6–8]. The work in [9] coded multiple views at the minimum level of redundancy in order to speed up the view switching. The work in [10, 11] mainly focused on the transmission of Multiview video coded streams based on multicast whose aim was to maximize the coding optimization to facilitate accessing to different views. However, they have a common shortcoming that the dynamics of networks are not efficiently exploited to improve protocol performance of Multiview video [12].

Although packet loss and transmission delay is considered in [13], it does not solve the problem of correlated multiple flows in Multiview video.

A video transmission control is presented in [14] to jointly control channel rate and relay node. In [15], Yanzhi Dou et al. utilized correlation of application users to monitor control flows and data flows for cognitive radio.

In this paper, a packet scheduling optimization algorithm for correlated multiple flows is proposed, enabling maximizing Multiview video quality performance by exploiting spatial and temporal correlation of multiple flows to reduce the costs of delay variance among flows as well as transmission redundancy among them.

2 Problem Formulation

By using distributed source code, an image $f(t)$ can be recovered from either temporal or spatial correlated frames, under the condition that those reference frames should be in encoder's buffer already. So we have

$$f(t) = \begin{cases} [1 - \rho(f(t))\big|\theta^{S}(t)\,] \cup [1 - \rho(f(t)\big|\theta^{T}(t-\tau))], & if \ \ f(t) \in K \\ f(t), & if \ \ f(t) \notin K \end{cases} \tag{1}$$

where K means pure key frames set. At the decoder side, key frame is decoded independently without the help of other frames, and dependent frame can be recovered from temporal or spatial correlated frames.

We use D_τ to denote this kind of video distortion of wide screen video streaming. Assume the packet length is S, the transmission delay of a camera m is represented by

$$\tau_m = \frac{S \times n(L_m)}{R_m} \tag{2}$$

where $n(L_m)$ is the number of hops from the camera m to decoder.

We use I_m to denote the encoding importance of camera m compared to other cameras, with $\sum_m I_m = 1$. The delay distortion of multi-camera is described as:

$$D_\tau(F(t)) = \sum_m I_m d_\tau(f_m(t)) = \sum_m I_m e^{-\frac{t_{dex}}{t_0 + \tau_m}} \tag{3}$$

We use $\rho(f_m(t))f_k(t_k)$ to describe the recovery degree of an image $f_m(t)$ from its neighbor $f_k(t)$. Then the data amount of $f_m(t)$ is

$$f_m(t) = (1 - \rho_{m,k})f_k(t) \;\; t \geq \tau_m, t \geq \tau_k \, m \neq k. \tag{4}$$

If the image is recovered from time correlated relationship, then the data rate of $f_m(t)$ is

$$f_m(t) = (1 - \rho_{m,m})f_m(t - \tau_{m,m})\, t \geq \tau_m + \tau_{m,m} \tag{5}$$

where $\tau_{m,m}$ represent the gap of time between two successive images from the same camera m.

So the effective frames in the decoder's buffer can be described as:

$$F(t_{dec}) = \sum_{m=1}^{M} \sum_{t=\tau_m}^{t_{dec}} \{\alpha_1(1 - \rho_{m,k})R_k(t) + \alpha_2(1 - \rho_{m,m})R_m(t - \tau_{m,m})\} \tag{6}$$

where

$$\sum_{i=1}^{2} \alpha_2 \leq 1, \quad \alpha_i \geq 0 \tag{7}$$

and

$$R_m, R_k \in [R_L, R_M] \tag{8}$$

In Eq. (7), α_i denotes the weight of decoding dependence.

We express the frame rate distortion as:

$$D_R(F(t)) = D_0 + \frac{\theta_0}{F(t) - R_0} \tag{9}$$

The wide area video is measured in terms of PSNR, which is a monotonically decreasing function of the mean-square error (MSE) [13].

$$PSNR(F) = 10 \log_{10}(\frac{D_{\max}}{D(F)}) \tag{10}$$

The events of packet loss occur only when deadline cannot be met. So the distortion of wide screen video streaming might be represented as

$$D(F(t)) = D_R(F(t)) + D_\tau(F(t)) \tag{11}$$

We can now formulate the problem of maximizing the sum-PSNR of multiple cameras by jointly controlling the video encoding rate, delay distortion of a wide screen as followings.

P: Maximize

$$PSNR_d(F) \tag{12}$$

Subject to: (1)–(11)

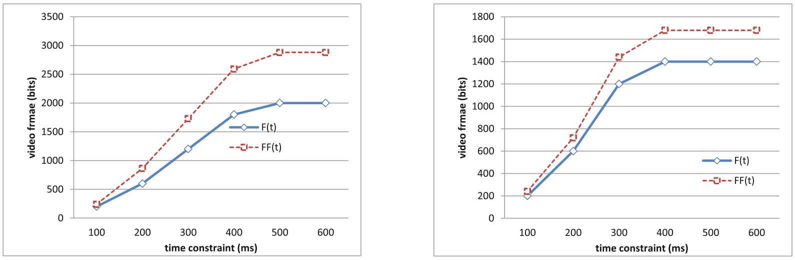
$$\sum_{m \in M} \sum_{t=\tau_m}^{t_{dec}} R_m(t) \le C_{dec} \tag{13}$$

$$\tau_m \le t_{dec} \quad (m \in M) \tag{14}$$

3 Simulation Results

We consider a 3-camera system where video packets experience different delays to reach decoder. Using same packet length and same transmission rate, the delays are proportional to hop number. For simplicity, the arriving intervals of packets from two cameras are calculated in terms of constant time slot ΔT. we use $F(t)$ to denote reconstructed wide area frame rate by conventional camera transmission scheduling [7], which does not consider joint packet scheduling and delay minimization, and use $FF(t)$ to denote our algorithms.

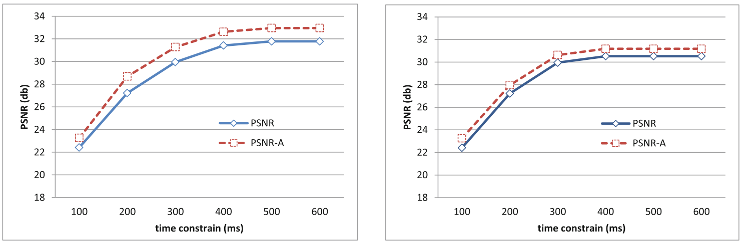
In Fig. 2(a), we observe that when the playout deadline is 400 ms the effective frame rate is maximized. In Fig. 2(b), the playout deadline is 500 ms at the maximal video frame rate. This means that relaxing delay constraint can improve effective frame rate with its maximum determined by decoder capability. We can observe that $FF(t)$ is about 400 bits more than $F(t)$ when time constraints are 400 ms and $C_{dec} = 1800$ bits. A similar phenomenon happens when $C_{dec} = 1400$ bits.



(a) (b)

Fig. 2. Video frame of wide area display for different playout deadline ($\rho_{s,m} = 0.8$, $\rho_{\tau,m} = 0.8$, $R_m = 200\,\text{bit/ms}$). (a) $C_{dec} = 2000$ bits. (b) $C_{dec} = 1400$ bits.

In Fig. 3, we use *PSNR* to denote conventional camera transmission scheduling [7], and $PSNR - A$ to denote our algorithms. With changing values of playout deadline, *PSNR* and $PSNR - A$ become larger. But $PSNR - A$ is still more than *PSNR* about 0.7 db. That is, our algorithm can allow cameras to exploit correlation information to schedule packets for higher PSNR.



(a) (b)

Fig. 3. PSNRs of wide area display for different playout deadline ($\rho_{s,m} = 0.8$, $\rho_{\tau,m} = 0.8$, $R_m = 200\,\text{bit/ms}$). (a) $C_{dec} = 2000$ bits. (b) $C_{dec} = 1400$ bits.

4 Conclusion

We studied correlated time constraints on Multiview multimedia system under certain network limitations. We have proposed a novel rate-distortion and delay-distortion model to take the advantages of the correlation level among cameras themselves and their neighborhood. Based on joint rate-distortion and delay distortion function, a distributed packet scheduling and delay control algorithm is proposed, which can optimally control resource allocation based on network capability and camera correlation. The proposed algorithm adjusts the amounts of video among sources to minimize the transmission delay variances, by which to maximize the effective video frame rate and minimize the packet loss due to exceeding playout deadline. In the simulation, we have analyzed the spatial and time constraints impacts on multimedia system. The results demonstrate the gain of our algorithm compared with classical method. We also have pointed out that the video reconstruction quality can be achieved by manipulating the correlation relationship among cameras.

Acknowledgments. This work was supported by the National Science Foundation of China under Grant (No. 61202470, No. 61471271), Wuhan Science and Technology Project (No. 2013010501010148, No. 2014010202010108), Chin Postdoctoral Science Foundation (No. 2013M531711), Financially supported by self-determined research funds of CCNU from the colleges' basic research and operation of MOE (No.CCNU14A05016, No. CCNU15A02017).

References

1. Chou, P., Miao, Z.: Rate-distortion optimized streaming of packetized media. IEEE Trans. Multimed. **8**(2), 390–404 (2006)
2. Fu, F., van der Schaar, M.: Structural solution for dynamic scheduling in wireless multimedia transmission. IEEE Trans. Circ. Syst. Video Technol. **22**(5), 727–739 (2012)
3. Wang, P., Dai, R., Akyildiz, I.F.: Visual correlation-based image gathering for wireless multimedia sensor networks. In: IEEE Proceedings INFOCOM, pp. 746–749. IEEE Press (2011)
4. Chakareski, J.: Transmission policy selection for multi-view content delivery over bandwidth constrained channels. IEEE Trans. Image Process. **23**(2), 931–942 (2014)
5. Toni, L., Maugey, T., Frossard, P.: Correlation-aware packet scheduling in multi-camera networks. IEEE Trans. Multimed. **16**(2), 496–509 (2014)
6. Kurutepe, E., Civanlar, M.R., Tekalp, A.M.: Client-driven selective streaming of multiview video for interactive 3DTV. IEEE Trans. Circ. Syst. Video Technol. **17**(11), 1558–1565 (2007)
7. De Abreu, A., Frossard, P., Pereira, F.: Optimizing multiview video plus depth prediction structures for interactive multiview video streaming. IEEE J. Sel. Topics Signal Process. **9**(3), 487–500 (2015)
8. Chakareski, J: Scheduling space-time dependent packets in multi-view video streaming. In: IEEE 15th International Workshop on Multimedia Signal Processing (MMSP), pp. 070–075 (2013)
9. Cheung, G., Ortega, V., Cheung, N.-M.: Interactive streaming of stored multiview video using redundant frame structures. IEEE Trans. Image Process. **20**(3), 744–761 (2011)
10. Lou, J.-G., Cai, H., Li, J.: Interactive multiview video delivery based on IP multicast. J. Adv. Multimed. **2007**(1), 13 (2007)
11. Li, Z., Begen, A.C., Gahm, J., Shan, Y., Osler, B., Oran, D.: Streaming video over HTTP with consistent quality. In: Proceedings of ACM Multimedia System Conference, pp. 248–258 (2014)
12. Cheung, G., Ortega, A., Cheung, N.: Interactive streaming of stored multiview video using redundant frame structures. IEEE Trans. Image Process. **3**(3), 744–761 (2011)
13. Stuhlmuler, K., Faber, N., Link, M., Girod, B.: Analysis of video transmission over lossy channels. IEEE J. Sel. Areas Commu. **18**(6), 1012–1032 (2000)
14. Guan, Z., Melodia, T., Yuan, D.: Joint optimal rate control and relay selection for cooperative wireless video streaming. IEEE/ACM Trans. Netw. **21**(4), 1173–1186 (2013)
15. Dou, Y., Zeng, K.C., Yang, Y., Yao, D.D.: MadeCR: correlation-based malware detection for cognitive radio. In: IEEE Conference on Computer Communications INFOCOM (2015)