

## Macroblock Importance Based Method for Stereo Video Congestion Control

*Xiaodong Wang, Ting Yao, Shaobo Li, Lianjun Zhang, Chaolin Wang*

*Faculty of Information Science and Engineering, Ningbo University, Ningbo, 315211 China  
Emails: wangxiaodong@nbu.edu.cn yaoting99@sina.cn li287184@sina.com  
zhanglijun@nbu.edu.cn wcl900107@163.com*

**Abstract:** *Transporting a stereo video stream over an unreliable channel can degrade severely the video quality. Network congestion is one of the important factors affecting the video quality. To improve the stereo video quality, a stereo video congestion control method, based on the macroblock importance, is proposed. Firstly, the importance of every macroblock is calculated according to its influence on the current frame and future frames. Then, combined with the importance information, all macroblocks of a frame are divided into three levels. Finally, the packets are transported with the help of RTP/RTCP protocol. We discard some not so important packets actively according to the importance levels when a congestion occurs. The results from NS2 simulation platform show that the stereo video quality is improved with the method proposed.*

**Keywords:** *Stereo video, macroblock importance, congestion control, NS2.*

### 1. Introduction

With the progress of network communication technology and multimedia technology, stereo video applications and researches have gained rapid development [1, 2]. Stereo video data are huger than traditional business, so it is easy to cause a congestion in an unreliable network, thus affecting the quality of the decoder video. An efficient congestion control method is needed to solve the problem of transporting stereo video with a limited bandwidth resource and improve the terminal perception quality of the stereo video.

Most of the traditional congestion control strategies are improved by modifying the protocols or models. TCP congestion control method of the AIMD (Additive Increase Multiplicative Decrease) [3] algorithm is well-known. But it does not keep TCP-friendliness. To solve these problems, many researchers have

proposed a TCP-friendly congestion control protocol, such as TFRC (TCP-friendly rate control protocol) [4], HCC TCP (Hybrid Congestion Control TCP) [5]. The queue management mechanisms control the congestion degree at each router through different kinds of AQM (Active Queue Management) mechanisms, e.g., Drop Tail [6], RED (Random Early Detection) [7], AVQ (Adaptive Virtual Queue) [8], etc. As obvious from the analysis of the above strategies, the traditional congestion control strategies do not take into account the characteristics of the video, they treat all the packets in the same way, and this is very bad for video reconstruction [9]. In fact, the importance of the video data is different. For example, when some of the data is lost, this will cause the entire video become unidentifiable; while if other data is lost, the video could be recovered by some simple error resilience methods. Therefore, designing congestion control algorithms according to the characteristics of the video data has become an important research direction. According to the video encoding dependencies, [10] and [11] thought that I frame is the most important, next follows P frame and B frame is the least important. So they proposed the strategies of discarding the low-important frames actively to reduce the possibility of congestion. In [12] the authors proposed a definition of the weighted expired-time packet loss rate by combining the structure importance of the video packets, it dropped the expired-time packets when a congestion occurred. The above methods just distinguish the frame importance based on the type and location of a frame; however they do not take into account the subjective feeling of the human eyes. In [13] a new scalable video congestion control mechanism is proposed based on the visual motion characteristics. It calculates the frame importance according to the sensitive degree of the human eyes to different video contents, then adjusts the packet loss policy according to the network conditions and the importance of the video packets. It could improve the network throughput and the video broadcast quality efficiently.

The above congestion control strategies either do not consider the video features, or just consider the importance of the whole frame. However, the importance of a macroblock in a frame is different, it is rough to consider only the frame importance. In this paper we propose a stereo video congestion control method based on the macroblock importance. Firstly it calculates the macroblock importance according to the distortion when a macroblock is lost and the relationship between the frames, then it marks the corresponding priority. When a network congestion occurs, the low-important packets are discarded actively, so that the transmission of high priority packets is ensured. Based on this, we build NS2 simulation platform for transporting a stereo video. To improve the video quality of the decoder, we judge the network congestion situation of the sender in real time and drop the low-important packets actively according to the macroblock importance. We use the RTP/RTCP protocol to encapsulate the data and for feedback quality information.

The remainder of the paper is organized as follows. In Section 2 a congestion control method based on the macroblock importance is introduced. In Section 3 we conduct an experiment to verify the method proposed. Finally, in the last section, we briefly present our conclusions.

## 2. Stereo video congestion control method based on the macroblock importance

### 2.1. Decision method of the stereo video macroblock importance

The current video is encoded and transmitted by taking the macroblock as a unit; each macroblock has different importance. In this paper the importance of each macroblock is calculated according to its influence on the current frame and future frames. We use IPPP coding structure, and if a macroblock in a frame is lost, then a simple error concealment method is used to recover the video quality: when the lost macroblock is in a frame of the left viewpoint, the error concealment strategy of an inter-frame macroblock direct copy is used; otherwise, the error concealment strategy of an inter-view macroblock direct copy is used.

#### 2.1.1. Decision method of a left viewpoint macroblock importance

Firstly, we calculate its influence on the current frame when a macroblock in the left viewpoint is lost. When the macroblock  $m$  in frame  $t$  of the left viewpoint is lost, the error concealment strategy of an inter-frame macroblock direct copy is used, so that the real signal at the decoder is  $\tilde{F}(s, t-1, m, i)$ , and its influence on the current frame is:

$$(1) \quad I_c^L = E \left\{ \left[ \tilde{F}(s, t, m, i) - \tilde{F}(s, t-1, m, i) \right]^2 \right\} = \\ = E \left\{ \left[ \tilde{F}(s, t, m, i) - \tilde{F}(s, t-1, m, i) \right]^2 \right\} = \text{RFD}(t, t-1, m),$$

where  $\text{RFD}(t, t-1, m)$  is the reconstructed frame difference between the encoder-reconstructed frames  $t$  and  $t-1$  in the macroblock  $m$ ;  $\tilde{F}(s, t-1, m, i)$ , and  $\tilde{F}(s, t, m, i)$  are approximately equal in size since only the current macroblock is lost.

Then we calculate its influence on the future frames, when a macroblock of the left viewpoint is lost. It encodes based on the block in H.264. The more times the reference pixels in a macroblock are referred to by the subsequent frames for motion compensation, the greater the effect, generated by this macroblock on the image quality to subsequent frames is. Based on the above discussion, we define the impact on the following frames, when a macroblock is lost as: the reference region area in each macroblock which is used for motion compensation for subsequent frames, namely the total number of pixels in the region. For a macroblock in the left viewpoint, it will not only be referenced by the following first frame of the left viewpoint, but it may also be referenced by the same time frame of the right viewpoint. So the importance of a macroblock in the left viewpoint consists of the region area in this macroblock which is used for motion compensation for a macroblock in the subsequent left viewpoint and right viewpoint, namely  $S_{ll}$  and  $S_{lr}$ , where  $S_{ll}$  can be calculated as

$$(2) \quad S_{ll} = d_{x_l} \times d_{y_l} = ((x_l/16+1) \times 16 - x_l) \times ((y_l/16+1) \times 16 - y_l),$$

where  $(x_i, y_i)$  are the axes at the top left corner of the matching block  $m'$  of a macroblock  $m$ . It can be obtained by the axes of  $(x_c, y_c)$  and the motion vector, as shown in Fig. 1.

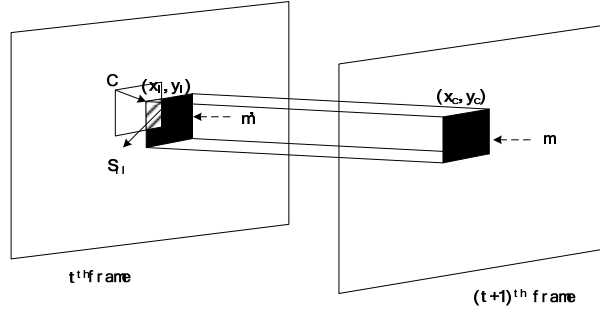


Fig. 1. The importance calculation of MB in the left viewpoint

In the same way  $S_{lr}$  can be calculated. So we can obtain the importance of the current macroblock  $C$  by cumulating the total area which the current macroblock  $C$  has used as motion reference by the macroblocks in the left viewpoint and right viewpoint:

$$(3) \quad I_p^1 = \sum_{i=1}^{M1} S_{ll}^i + \sum_{j=1}^{M2} S_{lr}^j,$$

where  $M1$  denotes that there are  $M1$  macroblocks of the left viewpoint referring to the current macroblock  $C$ ,  $M2$  denotes that there are  $M2$  macroblocks of the right viewpoint referring to the current macroblock  $C$ ,  $S_{ll}^i$  denotes the reference region area of macroblock  $i$  in the left viewpoint, and  $S_{lr}^j$  denotes the reference region area of macroblock  $j$  in the right viewpoint.

According to Equations (1) and (3) we define a parameter  $I_L$  to evaluate the importance of a macroblock, called an impact factor of the macroblock in the left viewpoint,

$$(4) \quad I_L = \alpha * I_C^1 + (1 - \alpha) * I_p^1,$$

where  $\alpha$  is a weight with a value between 0 and 1. If  $\alpha$  equals 0, the impact factor entirely refers to the impact on the current frame, else it entirely refers to the impact on the later frame.

### 2.1.2. A decision method of the right viewpoint macroblock importance

Firstly, we calculate its influence on the current frame when a macroblock in the right viewpoint is lost. When the frame  $t$  of the macroblock  $m$  in the right viewpoint is lost, the error concealment strategy of the inter-view macroblock direct copy is used, so that the real signal at the decoder is  $\tilde{F}(s-1, t, m, i)$  and its influence on the current frame is

$$\begin{aligned}
(5) \quad I_C^R &= E \left\{ \left[ \tilde{F}(s, t, m, i) - \tilde{F}(s-1, t, m, i) \right]^2 \right\} = \\
&= E \left\{ \left[ \tilde{F}(s, t, m, i) - \tilde{F}(s-1, t, m, i) \right]^2 \right\} = \text{RFD}(s, s-1, m),
\end{aligned}$$

where  $\text{RFD}(s, s-1, m)$  is the reconstructed frame difference between the viewpoints  $s$  and  $s-1$  in the macroblock  $m$ ;  $\tilde{F}(s-1, t, m, i)$  and  $\tilde{F}(s-1, t, m, i)$  are approximately equal in size since only the current macroblock is lost.

Then we calculate its influence on the future frames when a macroblock of the right viewpoint is lost. For the macroblock in the right viewpoint, it will only be referenced by the following first frame of the right viewpoint. So the importance of this macroblock depends on the region area used for motion compensation of the subsequent right viewpoint frame, called  $S_{rr}$ , where the calculation methods of  $S_{rr}$  and  $S_{ll}$  are similar. So we can obtain the value of importance of the current macroblock  $C$  by cumulating the total area which is used for motion compensation

$$(6) \quad I_p^r = \sum_{h=1}^{M3} S_{rr}^h,$$

where  $M3$  denotes that there are  $M3$  macroblocks of the right viewpoint referring to the current macroblock  $C$  for inter-frame motion compensation;  $S_{rr}^h$  denotes the reference region area of macroblock  $i$  in the right viewpoint.

According to Equations (5) and (6), we define a parameter  $I_R$  to evaluate the importance of a macroblock, called an impact factor of the macroblock in the right viewpoint as follows:

$$(7) \quad I_R = \beta * I_C^r + (1 - \beta) * I_p^r.$$

The meanings of  $\alpha$  and  $\beta$  are similar,  $\beta$  is also a weight with a value between 0 and 1.

## 2.2. Stereo video congestion control method based on the macroblock importance

The destination of the congestion control is to let the applications make a rational use of the network bandwidth and get the best transmission performance. There are two transport layer protocols over IP: TCP and UDP. TCP is a connection-oriented protocol; it provides reliable and ordered data transmission services. But its retransmission mechanism will introduce a large time delay, so it is not suitable for video transmission. To solve it, the researchers have proposed some new transportation protocols to meet the video transmission. We often use RTP/RTCP and RTSP for transporting streaming media at present. RTP is a kind of a streaming media protocol, architect of the UDP protocol. During the RTP session, each participant periodically sends RTCP packets. The main function of RTCP is to provide a feedback about the service quality. The network terminal can adapt to different network conditions according to the feedback information. In this paper we adopt a RTP/RTCP transport protocol.

The sender and receiver detect the network status in real time and discard the low-important packets actively when a network congestion occurs. So it can reduce

the transmission bandwidth and ensure a minimum impact on the quality of the decoder. The steps of the stereo video congestion control method based on the macroblock importance are as follows:

(1) We classify the macroblocks into different levels of importance according to the macroblock importance value calculated in the above section. We need to choose a threshold TH in the paper, TH is the largest importance value of a video sequence. The packaging rules are as follows:

1) if the macroblock importance value is  $I > \frac{2}{3}TH$ , then we argue that this macroblock is the most important one, marking the important level as 1;

2) if the macroblock importance value is  $\frac{1}{3}TH < I < \frac{2}{3}TH$ , then we argue that this macroblock is the second one of importance, marking the important level as 2;

3) if the macroblock importance value  $I < \frac{1}{3}TH$ , then we argue that this macroblock is the least important, marking the important level as 3.

We can obtain a record file after packaging as shown in Fig. 2. Each column in the file represents the frame number, frame type, packets bytes, packets number, sending time and importance level respectively. The congestion control strategy of this paper decides to discard the packets according to their importance level.

|   |   |      |   |          |   |
|---|---|------|---|----------|---|
| 1 | I | 8877 | 9 | 0.001000 | 1 |
| 1 | I | 135  | 1 | 0.031000 | 2 |
| 1 | I | 163  | 1 | 0.031000 | 3 |
| 2 | P | 12   | 1 | 0.031000 | 1 |
| 2 | P | 48   | 1 | 0.031000 | 2 |
| 2 | P | 3558 | 4 | 0.061000 | 3 |
| 3 | P | 733  | 1 | 0.091000 | 1 |
| 3 | P | 695  | 1 | 0.091000 | 2 |
| 3 | P | 198  | 1 | 0.091000 | 3 |
| 4 | P | 289  | 1 | 0.091000 | 1 |
| 4 | P | 945  | 1 | 0.091000 | 2 |
| 4 | P | 356  | 1 | 0.121000 | 3 |

Fig. 2. The mark of importance levels

(2) Transmitting the packaged data by using RTP protocol, the sender provides statistics of RTCP feedback information in real time. Then we calculate the number of lost packets within a certain period of time Plost, and judge the congestion status according to Plost.

(3) If Plost equals 0, then the network is not congested, we send packets normally; If Plost is less than the threshold Pmin, then the network is light congested, we drop the packets with importance level of 3; if Plost is between Pmax and Pmin, then the network is more seriously congested at this time, and we drop the packets in proportion with an importance level of 2 and 3. If Plost is greater than or equal to the threshold Pmax, then the network is seriously congested, so we drop the packets in proportion of all three levels. The drop ratio is proportional to the importance level. The specific algorithm is as follows:

```

If(Pmin<Plost<Pmax)
{
    disrate=L:M:N;
    If(pri==3&&n<N) { Discard(); n++;}
    If(pri==2&&m<M) { Discard(); m++;}
    Else
        send();
}
If(0< Plost < Pmin)
{
    disrate=L: M: N;
    If(pri==3&&n<N) {Discard(); n++;}
    Else
        send();
}
If(Plost>Pmax)
{
    disrate=L: M: N;
    If(pri==3&&n<N) {Discard(); n++;}
    If(pri==2&&m<M) {Discard(); m++;}
    If(pri==1&&l<L) {Discard(); l++;}
    Else
        send();
}
Else
    send();

```

### 3. Experimental results and analysis

To verify the efficiency of the stereo video congestion control method based on the macroblock importance, we have built a stereo video transmission simulation system on NS2 network simulation software. The network topology is shown in Fig. 3.

The sending node 0 transmits stereo video streaming, going through the intermediate nodes 2 and 3 to the receiving node 1. We adopt RTP/RTCP as a streaming media transmission protocol. We control it with the proposed stereo video congestion method when we detect a network congestion. Taking JM code and the coding structure IPPP, the test sequences include Book, Akko, Crowd and so on, GOP's size is 15, and the frame rate is 28 fps. To ensure the validity of the experiment, the receiver uses the same error concealment method, when the lost macroblock is in a frame of the left viewpoint, the error concealment strategy of an inter-frame macroblock direct copy; otherwise, the error concealment strategy of an inter-view macroblock direct copy is used. Adopting a random packet loss method and the proposed macroblock importance active packet loss method under the same network environment with the movement video sequence of Book and the dramatic

video sequence Crowd, the packet loss rate is 1%, 3%, 5%, 10%. Fig. 4 shows the results.

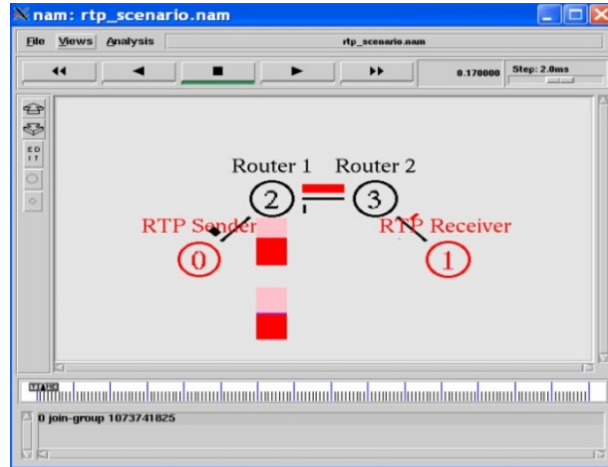


Fig. 3. Stereo video network simulation topology

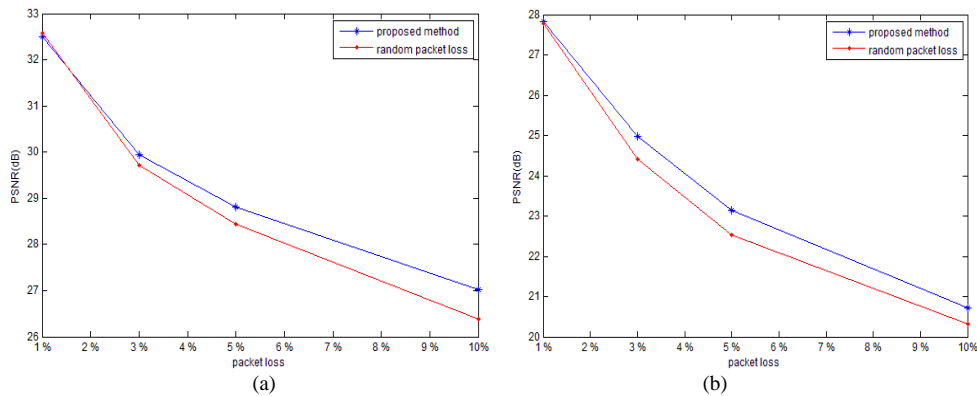


Fig. 4. PSNR comparison in the decoder: Book sequence (a); crowd sequence (b)

We can see from Fig. 4 that the video quality keeps falling down with the increase of the packet loss rate, so that PSNR keeps decreasing. When the packet loss rate is 1%, the network congestion is hardly detected, so the results of the proposed method and random packet loss method are fair. With the increase of the packet loss rate, PSNR is significantly higher when using the proposed method than using the random packet loss method. It illustrates that this algorithm provides a better video transmission service and improves the video quality in the decoder. In order to compare the efficiency and applicability of the proposed congestion control method, we perform experiments on more sequences, the results being shown in Tables 1-3.



Table 1. Comparison of the experimental results with Akko sequence

| Packet loss rate | Method             | PSNR  | +PSNR |
|------------------|--------------------|-------|-------|
| 3%               | Proposed method    | 28.57 | +0.37 |
|                  | Random packet loss | 28.20 |       |
| 5%               | Proposed method    | 27.17 | +0.67 |
|                  | Random packet loss | 26.50 |       |
| 10%              | Proposed method    | 24.13 | +0.35 |
|                  | Random packet loss | 23.78 |       |

Table 2. Comparison of the experimental results with Exit sequence

| Packet loss rate | Method             | PSNR  | +PSNR |
|------------------|--------------------|-------|-------|
| 3%               | Proposed method    | 30.52 | +0.74 |
|                  | Random packet loss | 29.78 |       |
| 5%               | Proposed method    | 29.22 | +0.63 |
|                  | Random packet loss | 28.59 |       |
| 10%              | Proposed method    | 26.66 | +0.46 |
|                  | Random packet loss | 26.20 |       |

Table 3. Comparison of the experimental results with Xmas sequence

| Packet loss rate | Method             | PSNR  | +PSNR |
|------------------|--------------------|-------|-------|
| 3%               | Proposed method    | 29.07 | +0.30 |
|                  | Random packet loss | 28.77 |       |
| 5%               | Proposed method    | 26.61 | +0.16 |
|                  | Random packet loss | 26.45 |       |
| 10%              | Proposed method    | 24.39 | +0.42 |
|                  | Random packet loss | 23.97 |       |

The above three sequences, Akko and Exit are video sequences contained in a fast foreground motion, and Xmas is a video sequence in which the whole scene moves fast. Tables 1-3 also show the downward trends in the video quality of the decoder with the increase of the packet loss rate. Under the same packet loss rate, the method proposed can get better video quality than the random packet loss, PSNR increases about 0.5 dB in average. It can be further verified that the proposed method is applicable to a video sequence, whether or not it contains fast or slow motion.

#### 4. Conclusion

A stereo video congestion control method based on the macroblock importance is proposed in the present paper. Compared with the traditional congestion control algorithms, it distinguishes the importance by considering the distortion of the macroblock and the error diffusion characteristics. When a network congestion occurs, the packets are discarded according to the network status and the packets importance. NS2 simulation results show that the method proposed has obtained better PSNR than the random active loss. It illustrates that the proposed method can protect an important macroblock in the video sequence and improve the video quality of the decoder.

**Acknowledgments:** The work was supported by the National Science-technology Support Plan (No 2012BAH67F01), National Nature Science Foundation of China (No 60832003, No 61071120), Zhejiang Science bureau/Innovation team design project (No 2012R10009-08), Zhejiang Nature Science Foundation (No Y1110161), Ningbo Science-technology Innovation team research project (No 2011B81002), Zhejiang Education Bureau Scientific research project (No Y201327703).

## References

1. Yilmaz, G. N., H. K. Arachchi, S. Dogan et al. 3D Video Bit Rate Adaptation Decision Taking Using Ambient Illumination Context. – Engineering Science and Technology, Vol. **17**, 2014, No 3, pp. 105-115.
2. Vosoughi, A., V. Testoni, P. C. Cosman et al. Multiview Coding and Error Correction Coding for 3D Video over Noisy Channels. – Signal Processing: Image Communication, Vol. **30**, 2015, pp. 107-120.
3. Edmonds, J. On the Competitiveness of AIMD-TCP within a General Network. – Theoretical Computer Science, Vol. **462**, 2012, No 30, pp. 12-22.
4. Sathiaselan, A., G. Fairhurst. TCP-Friendly Rate Control (TFRC) for Bursty Media Flows. – Computer Communications, Vol. **34**, 2011, No 15, pp. 1836-1847.
5. Xu, W. J., Z. D. Zhou, D. T. Pham et al. Hybrid Congestion Control for High-Speed Networks. – Journal of Network and Computer Application, Vol. **34**, 2011, No 4, pp. 1416-1438.
6. Jacobson, V. Congestion Avoidance and Control. – AMC SIGCOMM Computer Communication Review, Vol. **18**, 1988, No 4, pp. 314-329.
7. Floyd, S., V. Jacobson. Random Early Detection Gateways for Congestion Avoidance. – IEEE Transactions on Networking, Vol. **1**, 1993, No 4, pp. 397-413.
8. Kunniyur, S., R. Srikant. Analysis and Design of Adaptive Virtual Queue Algorithm for Active Queue Management. – ACM SIGCOMM Computer Communication Review, Vol. **31**, 2001, No 4, pp. 123-134.
9. Jacobs, S., A. Eleftheriadis. Streaming Video Using Dynamic Rate Shaping and TCP Congestion Control. – Visual Communication and Image Representation, Vol. **9**, 1998, No 3, pp. 211-222.
10. Zare, S., A. G. Rahbar. Congestion Control in IPTV over PON Using Digital Fountain Forward Error Correction. – Journal of Network and Computer Applications, Vol. **37**, 2014, pp. 240-252.
11. Lin, T. L., J. Shin, P. Cosman et al. Packet Dropping for Widely Varying Bit Reduction Rates Using a Network-Based Packet Loss Visibility Model. – In: Proc. of Data Compression Conference (DCC), 2010, pp. 445-454.
12. Pei, C. J., H. C. Lu, P. I. Hong. Deadline-Constrained Active Queue Management Mechanism in Wireless Networks. – Journal of Electronics & Information Technology, Vol. **35**, 2013, No 5, pp. 1069-1075.
13. Zhou, Y., J. H. Zhang, Y. S. Liang et al. Visual Motion Characteristics Based Scalable Video Congestion Control Mechanism. – Journal of Shenzhen University Science and Engineering, Vol. **31**, 2014, No 4, pp. 352-360.