Influence of Packets Losses on Video Quality in Case of Using Multiple Description Coding with Time Division into Two and Three Substreams

Alexey Tarakanov, Olga Gushchina, Ilya Nenakhov
P.G. Demidov Yaroslavl State University
Yaroslavl, Russia
tano@uniyar.ac.ru, olyagushchina@gmail.com, zergoodsound@gmail.com

Abstract—**The paper presents analysis of simultaneous use of multiple description coding (MDC) algorithm and multipath routing for IP-based videoconferencing system. Experimental results are given for time division of video stream in two and three substreams, compressed by H.264 codec with different quality factors (Q) and different packet loss rates in network (from 0% to 50%). Video quality is estimated with the use of two reference quality criterions: Peak Signal-to-Noise Ratio (PSNR) and Video Quality Metric (VQM).**

I. Introduction

Nowadays videoconferencing services are widely used for distance learning, presentations, telemedicine and in many other tasks [1,2]. According to Aberdeen Group research, based on 380 commercial companies, deployment of videoconferencing systems reduces travel costs and time delay for decision making by 17% and 5% respectively (primarily due to the elimination of business trips) [3]. Currently most of such systems are based on IP networks, where real-time data transmission protocols and digital signal processing methods (for audio and video signals) are used [4].

Existing videoconferencing systems could be classified as centralized and decentralized (or distributed) [5-8]. A key element of the centralized videoconferencing is a Multipoint Control Unit (MCU), which serves for connection establishment, signaling, audio and video transmission. In such network organization MCU becomes a “bottleneck” of the videoconferencing system because its resources (network interfaces performance, CPU, RAM, etc.) limit the maximum number of the videoconference participants. It should also be noted, that during the deployment of the videoconference system, to provide future scalability MCU with redundant characteristics is often installed. Such MCU makes the cost of videoconference system deployment high, and that is why centralized systems are currently used only in large industrial and commercial corporations [4, 5].

Above-mentioned limitations of the centralized systems make decentralized or distributed approach more preferable from point of view of scalability and cost.

In a distributed system audio and video streams are transmitted directly between users (rather than going through MCU), i.e. videoconference services are based on peer-to-peer principles. For such systems, the most promising approach combines the use of multipath routing and multiple description coding algorithms [6-8]. Such an approach, as illustrated in Fig. 1, could increase system robustness, eliminating MCU as a single point of failure, and provide acceptable video quality at the receiver side, even in case of packets loss of one or more sub-streams.

II. Multiple Description Coding

With using of multiple description coding (MDC) algorithm initial video sequence can be divided into substreams (or descriptors), each of which independently transmitted through the communication channel (for

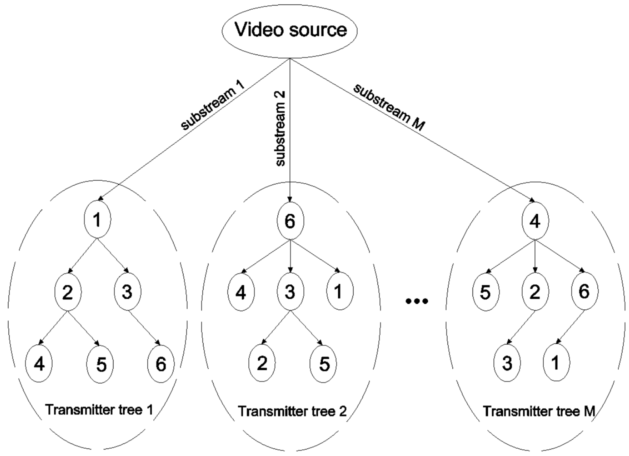


Fig. 1. Video streaming in decentralized videoconference system with multiple description coding

ISSN 2305-7254

example, IP-network). With MDC video stream splitting into substreams can be done in different ways, among the simplest of which are time and spatial separation, as well as separation according to the peak signal-to-noise ratio (PSNR) [9,10].

All MDC algorithms could be classified as balanced or unbalanced. With balanced MDC original video sequence is divided into several descriptors, each of which has an equal importance. In this case, the quality of the reconstructed video only depends on the number of the received descriptors.

With unbalanced MDC video quality at the decoder side depends on the number of received substreams, as well as on the received content (i.e. which descriptors were received).

Among the advantages of the MDC algorithms the following can be listed:

1) Resistance to packets losses and bit errors

MDC video substreams are transmitted through independent parallel channels from source to destination. In this case there is quite small probability, that all of such channels become inaccessible, and the MDC decoder is usually constructed in a way that reconstructed video will have an acceptable quality, even if only one from total number of the descriptors is received.

2) Ability to restore the video with better quality

Reconstructed video quality at the decoder increases with the number of the received substreams.

3) Distributed video storage

The MDC algorithm allows descriptor copies storing on different servers or storage systems. Distribution of stored information could reduce requirements for network attached storage (NAS) devices and at the same time improve reliability of such systems (because a failure of one single storage device does not result in whole system unavailability).

4) Delay decrease

Reduction of transmission delay is achieved by eliminating the need for data retransmission (that is used, for example, in TCP protocol).

At the same time, the use of MDC algorithms may be limited by the following drawbacks:

1) Relatively low coding efficiency

To ensure acceptable quality and possibility of video restoring at the decoder, each substream in MDC scheme must contain some information about original video sequence. Furthermore, coding efficiency is additionally reduces, if the video part contained in one descriptor also transmitted in some other descriptors (i.e. overlap coding is used). Typically, there is a compromise between coding efficiency and information redundancy.

2) Dependence on communication channel parameters

Transmission channels have a number of characteristics, including bandwidth, delay and its variation (jitter), packet loss ratio, etc. More over these parameters may vary in time. Therefore, in practical applications, video coding in MDC scheme is typically carried out with different compression ratios or bit rates and adaptation to network characteristics.

Analysis of papers [9-12] has shown that nowadays even for simplest separation methods for MDC there are no studies devoted to analysis of the video quality in videoconferencing system with MDC and its dependence on the number of the substreams and compression ratio used nor studies that take into account time video correlation. More than that, the time video correlation usually is not taken into account on the restoration stage (at the receiver side), and the quality of the resulted video signal is estimated only with the use of PSNR metric. Therefore, it is relevant to do such analysis for MDC system with time-division multiplexing with two and three substreams, transmitted in the lossy network. It should be noted that the number of substreams are usually limited by computing power of coding/decoding devices and it is especially critical for mobile videoconferencing.

If time division multiplexing is used, the MDC system could be implemented according to block scheme in Fig. 2.

Since data transmitted through IP-based network can be subject to packet loss, at the receiver side additional video processing should be performed to eliminate video quality degradation. At sufficiently high temporal correlation of video frames, restoration of the missing data on MDC decoder can be made by simple replacement of lost frames (which is known as “frame freezing”) or with the use of linear interpolation based on the previous and the following frames [12].

On the receiver side video restoration quality can be analyzed on short time intervals, when the communication channels can be regarded as fully functioning (lossless) of

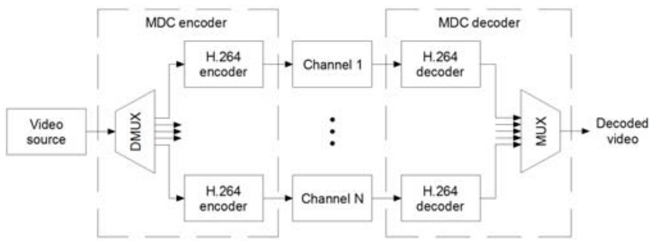


Fig. 2. Scheme of multiple description coding algorithm with time division multiplexing

fully non operable, as well as on long time intervals, when losses are probabilistic in nature [13,14]. The second case is more accurate for practical applications and that is why it is used in the paper.

III. Simulation methodology

During the computer simulation, MDC algorithm with time division multiplexing into two or three substreams without overlapping was explored. Restoration of the video sequence at the receiver side was achieved with the used of linear interpolation for missed frames.

Analysis was done for short time intervals as well as for long intervals. For short time intervals, video restoration was achieved in the assumption that one of the substream is totally loss, while for long intervals losses, arising in the transmission channel, was assumed as independent and uniformly distributed over the substreams, and their value is determined as a percentage of the total number of frames in video sequence.

During the experiment loss rate was varied from 5 to 50%. Values of packet loss greater then 50% usually indicate IP-network inoperability, therefore these cases have not been studied during the experiment [15].

For analysis of the distortion introduced by video compression, the simulation was performed in the absence of the H.264 encoder /decoder (when quality factor Q=0) and with presence of it with compression ratio: low (Q=15), middle (Q=30) and high (Q=50). In order to eliminate the dependence of the video content, the experiment was done several times and results were averaged.

Experiments were performed with using of standard test video sequences “Akiyo”, “Foreman” and “Soccer” [16]. To measure the quality of the video two objective criteria – peak signal-to-noise (PSNR) and video quality metric (VQM) – were used [8,9].

The PSNR can be computed from the mean squared error (MSE) by the following equations:

$$\mathrm{PSNR} = 20\log_{10}\frac{255}{\sqrt{\mathrm{MSE}}}, \tag{1}$$

$$\mathrm{MSE} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i - y_i\right)^2,$$

where N is the number of pixels in the frame, x_i and y_i represents the pixel values of the original (reference) and reconstructed images respectively [9]. It also should be noted, that PSNR value was estimated in Y-component of the image signal (i.e. for one single frame), and the total value of the PSNR for the whole video sequence was calculated by averaging PSNR values of its frames.

PSNR objective quality metric has low computational complexity. At the same time, studies [16-20] show that PSNR does not always accurately measure video quality, especially in the presence of some specific distortions. That is why the second criterion – VQM – was used. Compared to PSNR, VQM metric has sufficiently higher correlation coefficient with subjective quality assessment results and is more versatile to different types of distortions, thereby allows determining video quality more accurately [17].

Calculation of the VQM value is based on transformation of the video sequence into discrete cosine transform domain (DCT) and could be represented by a block-scheme shown in Fig. 3 [21]. It should also be noted, that according to the calculation algorithm, the larger values of the VQM metric correspond to the poor video quality.

Video encoding and decoding during the experiment was performed with the use of FFmpeg software. The results were obtained for different quality factor values Q (i.e. different compression ratios).

IV. Experimental results

A. Short-time analysis

For short-time analysis video reconstruction on the receiver side was carried out in assumption that one channel was totally inoperable. In this case Q value ranges from 5 to 50 (with 5units step).

The results obtained for “Akiyo” test video presented in Fig. 4. The graph shown in Fig. 4 for the scheme without MDC corresponds to so-called single description or standard video transmission over lossless network and takes into account distortions introduced by H.264 encoding algorithm.

The results demonstrated in Fig. 4 show that MDC scheme with 3 descriptors can provide acceptable quality for decoded video even if one of the descriptor was totally loss, and the quality of the reconstructed video in this case is also almost equal to the video quality obtained in lossless environment without MDC (with standard video encoding). In this case PSNR value reduces by 2.7 dB at Q=15 (in comparison with scheme without MDC). At Q=50 PSNR value remains almost constant which can be explained by the fact that in this case main video distortions are caused by H.264 algorithm, rather than information losses.

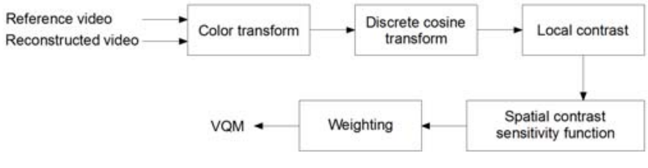


Fig. 3. Scheme of the video quality metric (VQM) calculation

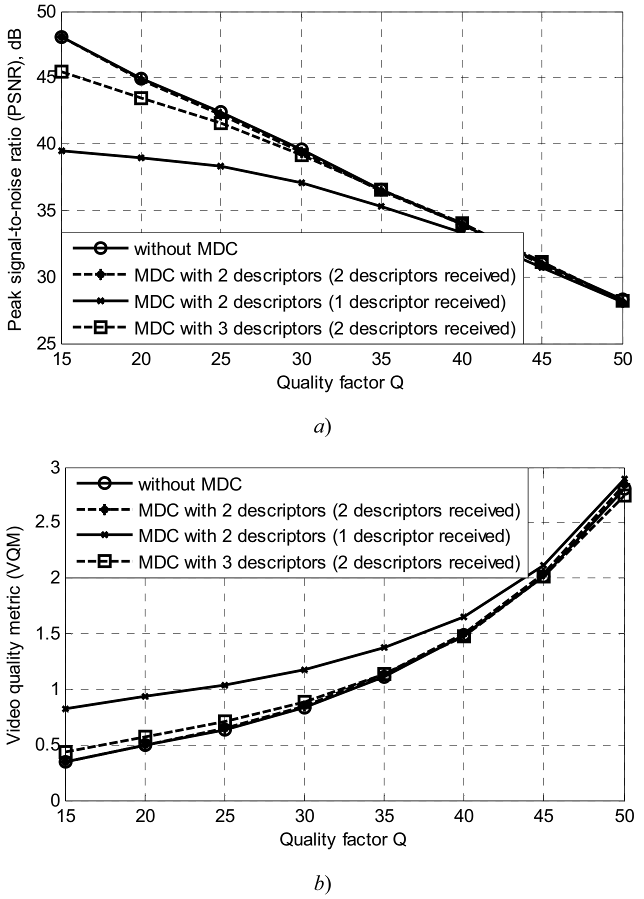


Fig. 4. Restoration performance vs quality factor for “Akiyo” test video in terms of a) peak signal-to-noise ratio (PSNR); b) video quality metric (VQM) values

This is due to the fact that with higher compression ratios (larger values of the quality factor Q) video sequence is characterized by the presence of sufficient artifacts caused mainly by H.264 compression algorithm and not by restoration techniques. In such a case VQM increases not more than 0.09 for all quality factors used.

At the same, efficiency of the MDC scheme with time division of original video sequence is reduced in the presence of rapidly moving object. During the research, such results were observed for “Soccer” test video. Obtained dependencies for “Soccer” video are given in Fig. 5.

As can be seen from the graphs shown in Fig. 5, reconstruction results for MDC scheme with 2 descriptors in case of 2 descriptors received almost coincide with results obtained for scheme without MDC. However, the loss of single descriptor leads to significant video quality degradation.

This is due to the fact that with higher compression ratios (larger values of the quality factor Q) video sequence is

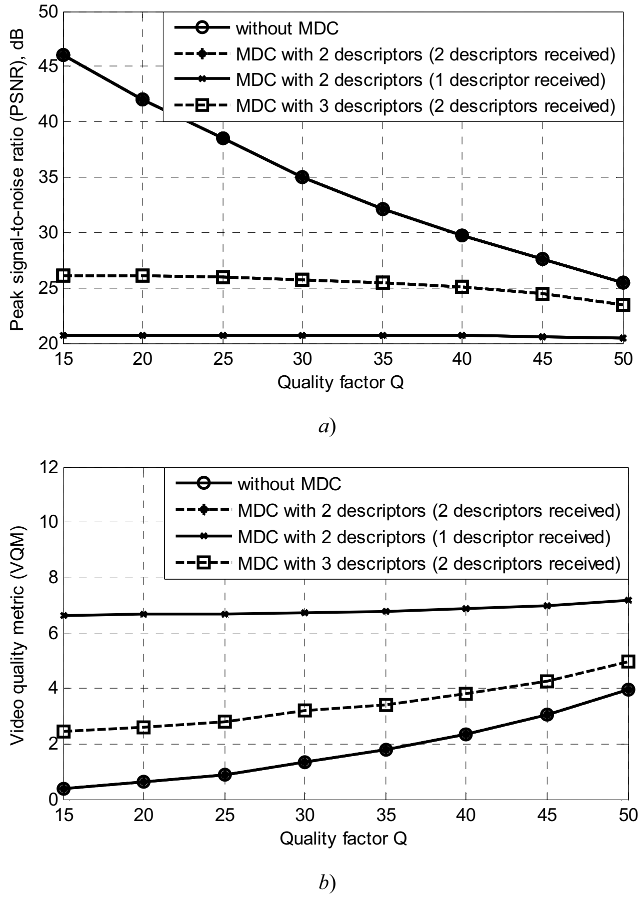


Fig. 5. Restoration performance vs quality factor for “Soccer” test video in terms of a) peak signal-to-noise ratio (PSNR); b) video quality metric (VQM) values

For MDC scheme with 3 descriptors PSNR value decreases by 1.9 dB and 19.9 dB at high (Q=50) and low (Q=15) compression rates respectively. In this case VQM value ranges from 0.99 to 2.05 units.

For MDC scheme with 2 descriptors and 1 descriptor received PSNR value reduction is up to 25.3 dB and remains almost constant with all considered compression rates (PSNR relative change does not exceed 1.5%). At the same time, VQM value increases by 3.27-6.27 units.

B. *Video motion estimation*

Results, presented in previous section, show high efficiency of MDC scheme with time division of original video into substreams (even in the presence of packets losses). At the same time, effectiveness of this method depends on the video content.

For estimation of video motion intensity the time correlation coefficient was used. It is determined for signals X and Y by the following equation:

$$r_{X,Y} = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}(X_{ij}-\overline{X})(Y_{ij}-\overline{Y})}{\sqrt{\left(\sum_{i=1}^{M}\sum_{j=1}^{N}(X_{ij}-\overline{X})^2\right)\left(\sum_{i=1}^{M}\sum_{j=1}^{N}(Y_{ij}-\overline{Y})^2\right)}},$$

where X_{ij} and Y_{ij} represent the current values of the pixel with coordinates i and j; M×N is the frame resolution; and $\overline{X}$ и $\overline{Y}$ are the mean values. Correlation coefficient was estimated in the luminance component between two consecutive video frames.

Estimated time correlation coefficient, its maximum and minimum values, as well as magnitude of its absolute and relative changes for three test videos are given in Table 1.

TABLE I. TIME CORRELATION COEFFICIENT FOR TEST VIDEO SEQUENCES

Test video	"Akiyo"	"Foreman"	"Soccer"
r	0.99	0.95	0.76
r_{max}	0.99	0.99	0.94
r_{min}	0.98	0.72	0.35
Δr	0.01	0.27	0.59
Δr / r	1.0 %	28.4 %	77.6 %

For "Soccer" video sequence time correlation coefficient has been rapidly changing within the entire video sequence. In contrast, for "Akiyo" test video time correlation coefficient remains almost constant. This fact can be explained by static camera position and absence of intensive movements within the video, It also should be noted, that for "Soccer" and "Foreman" sequences r_{min} value was observed at the camera movement frames (i.e. when the main plan was changed).

Obtained dependences between restoration results and video content are shown in Fig. 6. Results demonstrated in Fig. 6 are given for "Akiyo" and "Soccer" video sequences and correspond to analysis on short time intervals without partitioning into substreams (without MDC) and MDC scheme with 3 descriptors in case of one substream loss.

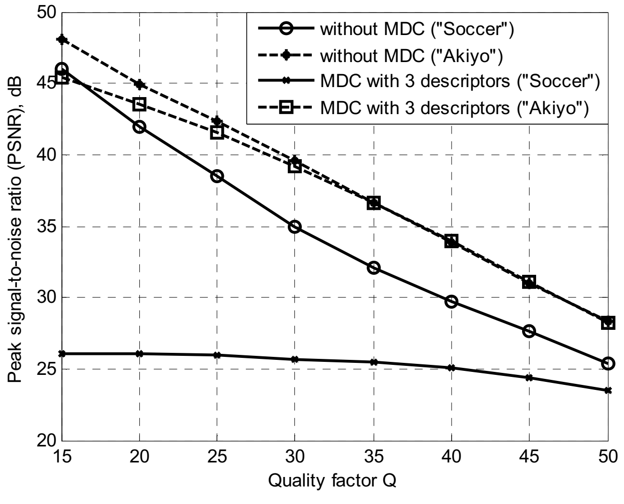
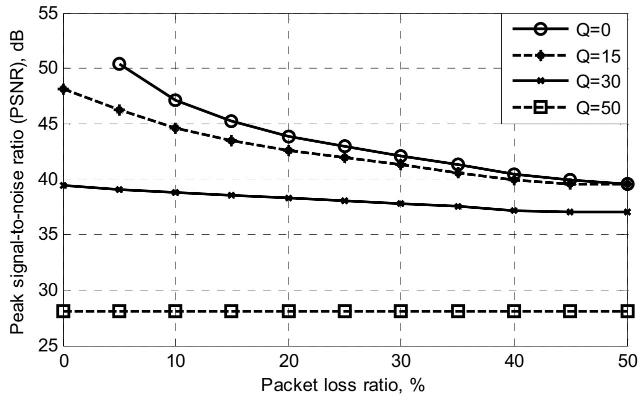


Fig. 6. Restoration performance vs quality factor for "Akiyo" and "Soccer" test videos in terms of PSNR

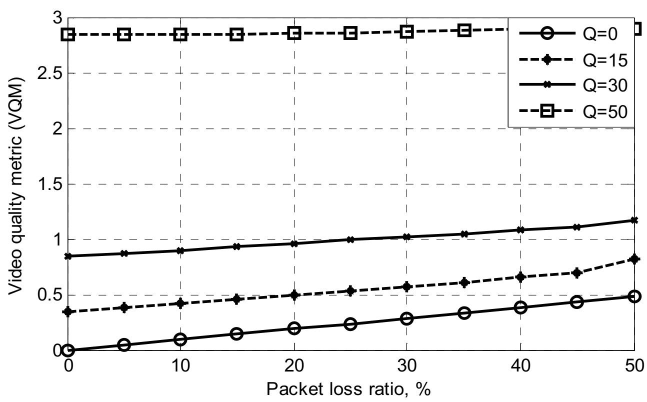
As can be concluded from Table I, for "Akiyo" test video time correlation coefficient remains almost constant, and the efficiency of the MDC algorithm, measured in terms of peak signal-to-noise ratio (PSNR) and video quality metric (VQM) values, is high. But for "Foreman" and "Soccer" videos time correlation coefficient is rapidly changing during the whole sequence, and MDC efficiency decreases. For "Soccer" video sequence r_{min} and r_{max} values correspond to frames No. 110 and 68 respectively. Additional analysis of the reconstructed frames shows that these values also complies with PSNR values, which reaches 15.4 dB for frame No. 110 and 27.6 dB for frame No. 68 (comparing to average PSNR value of 20.7 dB).

C. *Long-time analysis*

The results obtained for "Akiyo" test video, transmitted over lossy network, and 2-description encoding shown in Fig. 7. For 3-descriptor encoding reconstruction results are similar in nature, and that is why not included here. If Q=0



a)



b)

Fig. 7. Restoration performance vs packet loss ratio for "Akiyo" test video in terms of: a) PSNR, b) VQM

and packet loss ratio is 0%, according to formula (1), PSNR reaches the infinite value, because in this case both video compression and data transmission carried out without any losses.

Obtained results indicate that at high compression ratios packet loss ratio completely has no effect on the reconstructed video quality. For example, for Q=50 PSNR value remains constant, and VQM variation is only 0.04 units.

At the same time, without video compression (Q=0) and in case of low (Q=15) and middle (Q=30) compression ratios increasing of the packet loss ratio leads to degradation of the reconstructed video quality. For example, when Q=0, additional 5% of packet loss leads to additional PSNR decrease in 0.5-3.3 dB. For Q=15 quality reduction reaches 0.1-0.6 dB (for PSNR), and for Q=30 it is not more than 0.3 dB. For Q=[0..30] VQM metric value increases by 0-0.05 units for each additional 5% of information losses.

V. CONCLUSIONS

According to recommendations given in [18,19], video quality could be regarded as "excellent" if the PSNR value exceeds 37dB, "good" if it is ranges from 31 to 37 dB, "satisfactory" at 25-31 dB, and "unsatisfactory" if PSNR < 25dB. Thus, with MDC and for Q=15 video can be regarded as "excellent" or "good", even if packet loss ratio reaches 50%. In this case PSNR value does not fall below 32.9 dB even for interpolated frames (a mean value within video sequence PSNR ranges from 39.5 to 46.2 dB).

At middle compression ratio (for Q=30) "excellent" quality of the reconstructed video can be achieved if the packet loss ratio is not more than 30%. Average PSNR value for a given compression ratio does not fall below 37 dB, that is corresponds to "excellent" video quality. At the same time, with 30% loss ratio, PSNR value for interpolated frames reaches only 31 dB, and thus can have a significant impact on subjective quality perception.

At high compression rates (Q=50) average PSNR value is only 28.1 dB, that makes video only acceptable for those applications, where the video quality is not so critical.

Time correlation analysis distinguishes a strong dependence between efficiency of MDC scheme with time division of original video into substreams and motion intensity within a video sequence.

Given analysis demonstrate high efficiency of MDC algorithm with time division multiplexing for videoconferencing. However, it should be noted that in real network this efficiency could be achieved only if MDC algorithm is used with multipath routing schemes.

ACKNOWLEDGMENT

The work is supported by project No. 1060 as part of government task for scientific research in P.G. Demidov Yaroslavl State University.

REFERENCES

[1] V.P. Dvorkovich and A.V. Dvorkovich, *Digital Video Information Systems (theory and practice)*. Moscow: Technosphere, 2012.

[2] V.S. Sinepol and I.A. Tcikin, *Computed Videoconferencing*. Moscow: Mobile Communications, 1999.

[3] Questmark official website, The ROI of Video Collaboration, Web: http://www.questmark.co.uk/cms/wp-content/uploads/Aberdeen-report-ROI-of-TelePresence.pdf.

[4] H. Bai, A. Wang, Y. Zhao, J. Pan and A. Abraham, *Distributed Multiple Description Coding*. London: Springer, 2011.

[5] M. Bernstein, D. Gogia, V. Kamisetty, T. Matsekyukh and N. Udoshi, *Understanding and Implementing High Definition Videoconferencing*. USA: Juniper Inc., 2010.

[6] N.A. Smirnov, "Videoconferencing: from telepresence to communication", *Information Service Director*, vol. 5, 2012, pp. 38-40.

[7] A.N. Tarakanov and P.S. Savasin, "Procedure for calculating metrics in upgraded multiple trees algorithm", *Herald of computer and information technologies*, vol. 5, 2014, pp. 32-37.

[8] A.N. Tarakanov V.V. Tupitsyn and A.L. Priorov, "Construction of network video conferencing system with Dijkstra algorithm usage and the new metric", *Herald of computer and information technologies*, vol. 1, 2013, pp. 34-40.

[9] V.V. Tupitsyn, A.L. Priorov and A.N. Tarakanov, "Analyzes of multipoint videoconferencing scalability based on the multiple trees algorithm", *Design and technology of electronic means*, vol. 4, 2012, pp. 20-24.

[10] E. Magli, S. Campione and L. Peraldo, "Video Multiple Description Coding (MDC)", Web: http://ru.scribd.com/doc/13329246/Video-Multiple-Description-Coding-MDC#scribd.

[11] A. Vitali, "Multiple Description Coding – a new technology for streaming over the Internet", Web: https://tech.ebu.ch/docs/techreview/trev_312-vitali_streaming.pdf.

[12] H.H. Maza'ar and H.N. Elmahdy, "Multiple description coding based network coding", International Journal of Computer Applications, vol. 9, 2010, pp. 35-40.

[13] F.H.P. Fitzek, B. Can, R. Prasad, M. Katz and D.S. Park, "Traffic analysis of multiple description coding of video services over IP networks", in Proc. *Wireless Personal Multimedia Communications (WPMC)* Conf., Italy, 2004, pp. 266-270.

[14] A. Tarakanov and O. Gushchina, "Analysis of multiple description coding algorithm with time division of video stream in task of multipoint videoconferencing", Proc. of *First Int. Scientific-Practical Conf. Problems of Infocommunications Science and Technology*, Ukraine, 2014, pp. 85-87.

[15] A.N. Tarakanov, O.N. Gushchina, I.S. Nenakhov and Y.S. Kornev, "On the problem of multiple description coding algorithms with time division multiplexing for videoconferencing", Proc. of *Radioelectronic means of information obtaining, processing and visualization* Conf., Russia, 2014, pp. 189-193.

[16] H. Rausch "Use of synthetic packet generation tools in measuring IP network performance", Web: https://www.academia.edu/7956528/Use_of_Synthetic_Packet_Generation_Tools_to_Measure_Packet_loss_and_Jitter_in_IP_Networks.

[17] Y. Ou, Y. Zhou and Y. Wang, "Perceptual quality of video with frame rate variation", Proc. of *IEEE Int. Conf. on Acoustic, Speech and Signal Proces.*, 2010, pp. 2446-2449.

[18] Y. Yoon, M. Kim and S. Lee, “Performance analysis of H.264/AVC, H.264/SVC and VP8 over IEEE 802.11 Wireless Network”, Proc. of *IEEE Symp. on Computers and Communications (ISCC)*, 2012, pp. 151-156.
[19] J. Korhonen, N. Burini, J. You and E. Nadernejad, “How to evaluate objective video quality metrics reliably?”, Proc. of *IEEE 4th Int. Works. Quality of Multimedia Experience (QoMEX)*, 2012, pp. 57-62.
[20] E.A. Pavlov, O.N. Gushchina, A.L. Priorov and V.V. Khryashchev, “Non-reference image quality assessment algorithm based on discrete cosine transform”, *Achievements of modern radioelectronics*, vol. 11, 2013, pp. 3-13.
[21] F. Xiao “DCT-based video quality evaluation”, Web: http://compression.ru/video/quality_measure/vqm.pdf.