# Quality evaluation of compressed 3D surgical video

Chaminda T.E.R. Hewage, Harsha D. Appuhami
and Maria G. Martini
Wireless Multimedia & Networking (WMN) Research Group
Kingston University, Kingston upon Thames
KT1 2EE, London, UK
Email: c.hewage@kingston.ac.uk

Ralph Smith, Iain Jourdan
and Timothy Rockall
Minimal Access Therapy Training Unit
Post Graduate Medical School, University of Surrey
Manor Park, Guildford
Surrey, GU2 7WG, UK

***Abstract*—3D medical video was forecasted to be one of the groundbreaking 3D video applications. These range from tele-consultation to 3D robotics surgery. Enabling 3D video in e-health applications results in the provision of more natural viewing conditions, improved diagnosis and accurate interventions in surgical procedures. The deployment of 3D video services in healthcare is made possible to some extent by the advanced capturing devices (e.g., 3D endoscopes), recent advances in wireless communication technologies (e.g., LTE-Advanced(LTE-A)) and 3D video display technologies. Remote robotic assisted surgery and surgery training (education for surgeons) can benefit in particular from 3D video technologies due to the added dimension of depth. This paper analyzes the quality of compressed 3D surgical video. Moreover, asymmetric encoding of 3D medical video without compromising the medical quality of experience (M-QoE) is investigated in this paper. The quality of the compressed 3D medical video with the proposed method is evaluated using a comprehensive subjective quality evaluation test involving 12 medical surgeons. The results show a slightly better perception with the proposed asymmetric coding method compared to reference symmetric compression method, however the difference is statistically insignificant.**

## I. Introduction

3D medical video was forecasted to be one of the groundbreaking 3D video applications. These range from tele-consultation to 3D robotics surgery. 3D video in medical applications allows improved visual perception, better diagnosis, effective interventions and training of medical professionals (e.g., training of surgeons). The research published in [11] shows that 3D video increases performance of surgical procedures (e.g., knot tying). However, there are many challenges to overcome in order to enable 3D medical applications over band-limited and unreliable communication channels. Some of the identified challenges are discussed below;

- Need for miniaterized and accurate 3D surgical video capturing devices (miniaturized 3D video lences/cameras for capturing and recording surgeries, e.g., 3D endoscopes for Minimally Invasive Surgery (MIS))
- 3D video requires higher bandwidth and storage spaces compared to 2D medical video: for example, theoretically, stereoscopic video application requires twice the bandwidth of a 2D monoscopic video application due to the availability of two left and right video streams. Moreover, medical video is often captured with high resolution and pixel depth for better perception. The use of compression algorithms should not visibly degrade the quality of the original medical 3D video. This is the main aspect addressed in this paper. This paper investigates the employability of asymmetric coding for stereoscopic surgical video applications to ease the storage and bandwidth requirements of such applications.
- The bit-errors and packet loss errors occur over unreliable communication channels may affect the quality of reconstructed 3D video in a different way than that of conventional 2D video applications [9][10] The effects of channel induced errors could be much severe for surgery application which could lead for inaccurate diagnosis and interventions. The delay and jitter introduced by these channels would also affect the applications such as tele-surgery (e.g., delay is still too high for haptics/control signals).
- 3D video quality can be described as a collection several perceptual attributes such as image quality, depth perception, naturalness, comfortness, presence, etc. Therefore, measurement of 3D video quality of experience (QoE) is a challenge. Automated diagnostic quality measurements in medical applications are still a major challenge to be addressed.
- Due to the conflict of accommodation and convergence, viewing 3D video over long time on flat screent displays may cause discomfort and eye strain. This creates a major challange to surgical procedures which could take a few hours and visual comfort in these situations is yet to be realized. The emerging display technologies with and without viewing aids may be able to provide comfortable viewing experience in these situations.

Many of these obstacles mentioned above are being gradually overcomed with emerging novel technologies and standards. For instance, at present 3D endoscopes can be used in Minimally Invasive Surgery (MIS). In these procedures, surgeons perform procedures by making only a few small incisions or using natural body openings through which they insert endoscopes and surgical devices [1][2]. The fesibility of these procedures when using 3D video is discussed in [11]. For example, it allows better recognition of tissue layers and may facilitate complex maneuvers such as laparoscopic suturing or knot tying as discussed in [11].

The latest communication technologies/standards (e.g., LTE/LTE-A or WiMAX) provide high bandwidth, which could support 3D video medical services to a certain limit. In addition to this, these technologies provide a range of error resilient features (e.g., Hybrid ARQ, Admission control policies), which enable reliable transmission of 3D medical video. Performance of 3D surgical video transmission over next generation communication channels is evaluated in [13]. This study shows that with specific channel setups or configurations, we could achieve improved 3D medical QoE over unreliable and band-limited wireless channels [13].

For the requirement of reduced bandwidth, stereoscopic video could be compressed by exploiting inter-view redundancies in addition to removing spatial and temporal redundancies. For these we could use simulcast encoding, or frame-packed encoding (e.g., side-by-side format) or layered encoding (e.g. Multi-View Coding (MVC)) mechanisms [12]. These mechanisms have their own advantages and disadvantages such as the possiblity to remove inter-view redundancies and backward compatibility. Furthermore, the left and right views of stereoscopic 3D video could be compressed taking into account the inherent characteristics of 3D video; for instance, the binocular suppression concept can be considered without affecting the overall 3D perception [3]. Binocular suppression theory suggests that the overall 3D perception is driven by the best quality view up to a certain limit [3]. Asymmetric coding results in mixed-resolution video sequences, which meet bitrate limitations imposed by the user terminals, applications and the underlying network. In this paper, asymmetric/multi-resolution coding of stereoscopic surgery video is analyzed as a method for reducing the storage and memory requirements and the bandwidth required for stereoscopic medical video streaming/transmission over band-limited networks. Mixed-resolution video coding for left and right videos is based on the response of the human visual system for depth perception [7]. According to [7] the high quality view is dominating the stereoscopic video quality without any degradation to depth perception. Stelmach et al have shown in [8] that the spatial filtering of one channel of a stereo video-sequence may be an effective means of reducing transmission bandwidth. The H.264/AVC based stereoscopic video coder described in [4] used an asymmetric combination of temporal and spatial formats to increase the compression efficiency. The results in [4] show that stereoscopic video sequences can be encoded at a rate about 1.2 times the monoscopic video using spatial and temporal filtering of one image sequence. Moreover, [5] examines the bounds of asymmetric stereo video compression for a 3-D video system based on H.264/AVC view coding. The mixed-resolution coding concept has been applied in [6] to colour plus depth video in different perspectives, i.e., to encode the colour and depth video based on their influence towards better perceptual quality. However, the possibility to employ asymmetric coding for 3D medical video and its consequences have not been thoroughly investigated to date. Unlike other 3D video applications, medical applications need a significant attention since inaccurate perception can lead to wrong diagnosis and ineffective interventions in surgical procedures. In this paper, the effect of asymmetric coding on 3D medical video is analyzed using subjective quality measurements. The opinions of medical professionals are considered as ground truth data in this work.

A prior study reported in [14] analyzes asymmetric compression of 3D surgery video. However, this study focuses on high bitrates, where symmetric coding performs better in high to moderate bitrates (at high bitrates, both left and right video can be compressed at high quality without the need for asymmetric compression). The study presented in this paper considers moderate bitrates and performance of asymmetric and symmetric coding is compared at the same overall bitrates. The number of surgeons involved in this study is also higher than the number of medical professionals involved in the investigation reported in [14]. Hence, this would provide reliable quality scores. Furthermore, eye dominance could affect user perception when viewing asymmetrically encoded 3D video. For instance, the highly compressed right view may not be ideal for a person with right eye dominance. Therefore, in order to null the effect of eye dominance both left and right eye dominance asymmetric coding configurations are considered in this study.

This paper is organized as follows. Section II presents the investigated asymmetric coding approach for 3D medical video. The experimental work and results are discussed in Section III. Section IV concludes the paper.

## II. Asymmetric compression of 3D Medical/Surgical Video

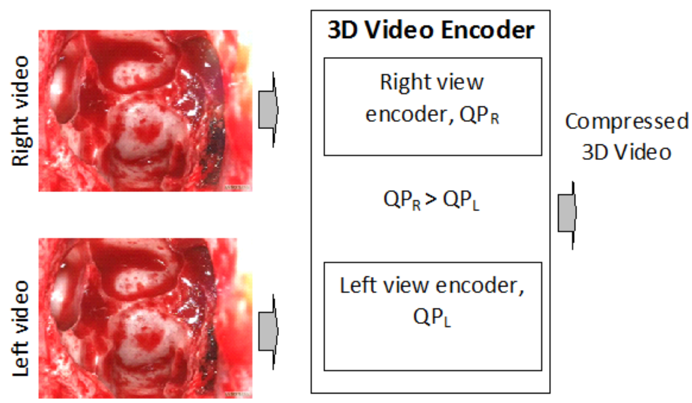



Fig. 1: Asymmetric coding with coarsely quantized right video - Left eye dominance case ($QP_R$: QP of the Right view encoder & $QP_L$: QP of the Left view encoder)

Asymmetric coding could be performed on temporal, spatial or quality domains. For instance, the most important view can be encoded at the highest spatial resolution whereas the other view can be compressed at a lower spatial resolution. This study focuses on quality asymmetricity where one view is encoded at the highest possible quality whereas the other view is encoded at slightly lower quality. This can be done by varying the quantization parameter (QP) of a block based codec. The proposed asymmetric coding approach for 3D medical video is illustrated in Figure 1. Two Quantization Parameters (QPs) are employed in the left and right view encoders to achieve unbalanced compression for left and right videos. The left video is quantized using lower QP values during encoding compared to the QP values used for the right video encoding. This results in a higher bitrate (and better quality) for the left image sequence and in a relatively lower bitrate (and reduced quality) for the right image sequence. The eye dominance of subjects could have an effect on depth perception due to this asymmetric coding. Therefore, in order to null the effect from eye dominance, another set of asymmetrically encoded sequences are created with inverse effect. In these sequences, the right video is kept at higher quality whereas the left video is encoded at a higher compression rate to produce a reduced quality video.

The asymmetric coding approach described above reduces the overall bitrate required for 3D medical video. The effect of asymmetric coding for better diagnosis and intervention can be investigated using subjective quality evaluation methods. As a reference method, 3D medical video coding with symmetric coding is considered in the study. In this case, the left and right image sequences are encoded using the same QP value to achieve similar compression levels (as well as comparable quality). In order to evaluate the performance of both encoding approaches, the overall bitrates of each approach is kept at the same value.

The left and right videos are encoded using two H.264/AVC encoders running in parallel (i.e., simulcast encoding method). During the encoding of the left/right video sequences, four QP modes (i.e., QP-10, QP-20, QP-30, and QP-40) are used. In the proposed asymmetric coding approach, QP-10 case is referred as the best compressed quality in this study. This is chosen based on the results of [14] which shows even QP-10 provides comparable results with QP-0 case (i.e., lossless compression). This is performed to keep the quality of the best image sequence close to its original quality, which is important for 3D medical video applications. The highest QP is set to 40 since QPs above 40 produced low Mean Opinion Scores (MOS) in our previous study [14]. The left and right compression pairs for left eye dominance case are listed below;

- Left QP-10/Right QP-10
- Left QP-10/Right QP-20
- Left QP-10/Right QP-30
- Left QP-10/Right QP-40

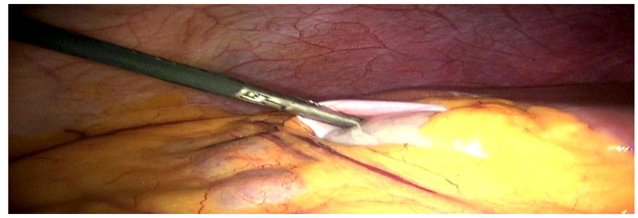
And for the right eye dominance, the compression pairs are as follow;

- Left QP-10/Right QP-10
- Left QP-20/Right QP-10
- Left QP-30/Right QP-10
- Left QP-40/Right QP-10

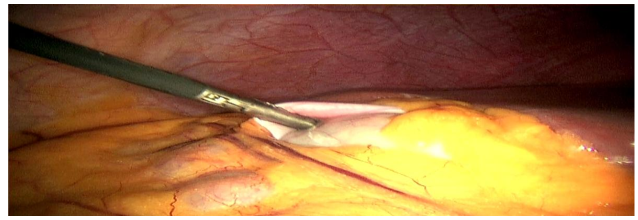
For the symmetric encoding case following compression pairs are employed in order to match the overall bitrate with the asymmetric case.

- Left QP-10/Right QP-10
- Left QP-14/Right QP-14
- Left QP-16/Right QP-16
- Left QP-17/Right QP-17

The quality is evaluated at these matching bitrates to see the performance improvement of the proposed asymmetric compression method compared to symmetric encoding approach.



(a) Left view: Sequence A



(b) Right view: Sequence A

(c) Left view: Sequence B

(d) Right view: Sequence B

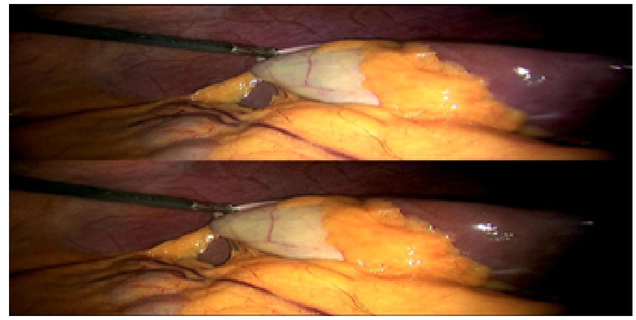
Fig. 2: Two samples from the 3D medical endoscope videos (a) Sequence A: Left view, (b) Sequence A: Right view, (c) Sequence B: Left view, and (d) Sequence B: Right view

## III. Experimental setup, Results and Discussions

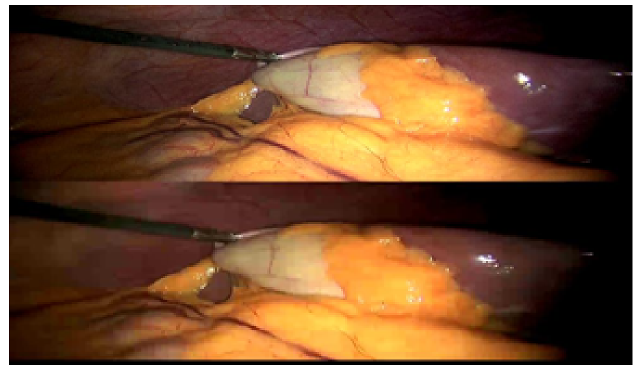
In order to evaluate the performance of the proposed asymmetric encoding approach for 3D surgery video, experiments are carried out using two samples (Sequence A and Sequence B) from the same 3D surgery sequence provided by the Minimal Access Therapy Training Unit (MATTU), Surrey, UK. The sample frames of these sequences are shown in figure 2. The resolution of each image frame is 1440x540 pixels. This medical video sequence covers a surgical procedure, named Laparoscopic Cholecystectomy. These 3D video sequences are encoded using the H.264/AVC video coding standard (using JM reference software Version 16.0). Two video codecs are employed to encode left and right views separately. Ten-seconds long video sequences (i.e., 250 frames at 25 fps) are encoded with IPPPIPPP sequence format, using the QP combinations mentioned in Section II for both I and P frames. Slices (i.e. 3 MB lines = 1 slice) are also introduced in the encoding process to make the decoding more robust to errors (for future work on 3D medical video transmission over unreliable communication channels).

In order to obtain opinion scores (OSs) from medical practitioners, subjective tests are conducted using the sequences obtained with different asymmetric and symmetric compression combinations. Twelve senior medical surgeons participated in subjective quality evaluation tests conducted at MATTU LAB, Surrey, UK. The subjective tests are performed using the Double Stimulus Impairment Scale (DSIS) quality evaluation method as described in the recommendation ITU−R BT.500−12. The five points categorical scale (i.e., Imperceptible (same as the original)-5, Perceptible but not Annoying-4, Slightly Annoying-3, Annoying-2 and Very Annoying-1) is utilized to obtain the opinions of the medical practitioners. The sequences are displayed in 3D using the LGCF 3D projector and the 3m x 2m silver screen with passive circular polarising eyewear for 3D viewing. All test sequences are viewed with low room lighting levels at a distance of 5 meters. Subjective quality (i.e., the overall 3D video quality), is recorded for the sequence under different compression combinations (both symmetric and asymmetric cases). The OSs of individual subjects were averaged to obtain the Mean Opinion Score (MOS) for all test cases during the result analysis.

The overall bitrates of Sequence A and Sequence B for asymmetric and symmetric encoding approaches are listed in Table I and Table II respectively. The overall bitrates of both symmetric and asymmetric encoding approaches are kept approximately the same as depicted in these tables. As expected the overall bitrate decreases with the increase of QP values for all the cases considered. The quality degradation caused by these compression levels should be taken into account when evaluating the feasibility of the proposed approach as presented below.



(a) Original left (top) and right (bottom) images



(b) Original left image (top) and right image compressed with QP=40 (bottom)

Fig. 3: $100^{th}$ frame of Sequence A (a) Original left (top) and right (bottom) images, and (b) Original left image (top) and right image compressed with QP = 40 (bottom)

Figure 3 shows the $100^{th}$ frame of Sequence A; 3(a) when using both original left (top) and right (bottom) images and 3(b) original left image (top) and right image compressed with QP=40 (bottom). It is visible that the bottom image in subfigure 3b is not as sharp as the bottom image in subfigure 3a (e.g., blockiness is clearly visible). However, when viewing in 3D, the higher quality left image (the top image in subfigure 3b) could drive the 3D perception and mask the effects of the lower quality right view/bottom image (i.e., binocular suppression theorem). On the other hand unbalanced images with asymmetric coding can lead to increased binocular rivalry and affect 3D viewings (could lead to visual discomfort). The effect of eye dominance of the viewers is minimized by doing the asymmetric coding for both left and right eye dominance subjects. In the case of symmetric coding, both images are subjected to similar compression artifacts. However, the amount of compression that can be achieved should be derived based on perceptual bounds without affecting the true 3D perception. Especially for a 3D medical video application, further precautions have to be taken, since these videos are used for surgical procedures. It is difficult to assess the true 3D perception by analyzing the quality of individual left and right images, therefore, subjective quality evaluation tests are conducted to obtain the practitioner's point of view. Both sequence A and B are selected for these tests.

The MOS for left and right eye dominance asymmetric coding are shown in Figure 4. A high MOS score means that the quality is close to the original 3D image sequence (i.e., the difference in quality with respect to the original 3D image sequence is low), whereas lower MOS means that the image quality is not as good as the original 3D image sequence (i.e., the difference between the original and processed sequences is high). In this figure, "Bitrate 1" refers to the high overall bitrate whereas "Bitrate 4" refers to the lowest overall bitrate. According to this figure, it is evident that there is a significant difference between left and right eye dominance encoding approaches. This may be due to the eye dominance of participated subjects. According to Student T-Test, the results of both approaches are statistically insignificant. However, the probability of this result, assuming the null hypothesis, is 0.63. The Pearson correlation coefficient and Spearman rank orders between these results are 0.3924 and 0.4248 respectively. This suggests that eye dominance should be taken into account when selecting asymmetric encoding approaches.

In order to compare asymmetric encoding with symmetric compression, the obtained MOSs are evaluated and reported in Figure 5. In order to analyze the results both the left and right eye dominance cases were also averaged to obtain a single scores for the asymmetric case (listed as global asymmetric case in this figure). According to 3D subjective quality results in this figure, it can be observed that, asymmetric coding performs slightly better compared to symmetric encoding. However, the difference is insignificant according to the Student T-test. In case of the left eye dominance case, correlation with symmetric encoding;

TABLE I: Overall bitrates of asymmetric and symmetric encoding approaches: Sequence A

| Sequence A | Bitrate (Kbps) Asymmetric coding (Left eye dominance) | Sequence A | Bitrate (Kbps) Asymmetric coding (Right eye dominance) | Sequence A | Bitrate (Kbps) Symmetric coding |
|---|---|---|---|---|---|
| Original | 622080 | Original | 622080 | Original | 622080 |
| Left QP-10/Right QP-10 | 79911 | Left QP-10/Right QP-10 | 79911 | Left QP-10/Right QP-10 | 79911 |
| Left QP-10/Right QP-20 | 53260 | Left QP-20/Right QP-10 | 52771 | Left QP-14/Right QP-14 | 53647 |
| Left QP-10/Right QP-30 | 41185 | Left QP-30/Right QP-10 | 41557 | Left QP-16/Right QP-16 | 43824 |
| Left QP-10/Right QP-40 | 40118 | Left QP-40/Right QP-10 | 40498 | Left QP-17/Right QP-17 | 38212 |

TABLE II: Overall bitrates of asymmetric and symmetric encoding approaches: Sequence B

| Sequence B | Bitrate (Kbps) Asymmetric coding (Left eye dominance) | Sequence B | Bitrate (Kbps) Asymmetric coding (Right eye dominance) | Sequence B | Bitrate (Kbps) Symmetric coding |
|---|---|---|---|---|---|
| Original | 622080 | Original | 622080 | Original | 622080 |
| Left QP-10/Right QP-10 | 78340 | Left QP-10/Right QP-10 | 78340 | Left QP-10/Right QP-10 | 78340 |
| Left QP-10/Right QP-20 | 51904 | Left QP-20/Right QP-10 | 52418 | Left QP-14/Right QP-14 | 52780 |
| Left QP-10/Right QP-30 | 40013 | Left QP-30/Right QP-10 | 41842 | Left QP-16/Right QP-16 | 43196 |
| Left QP-10/Right QP-40 | 38735 | Left QP-40/Right QP-10 | 40564 | Left QP-17/Right QP-17 | 37733 |

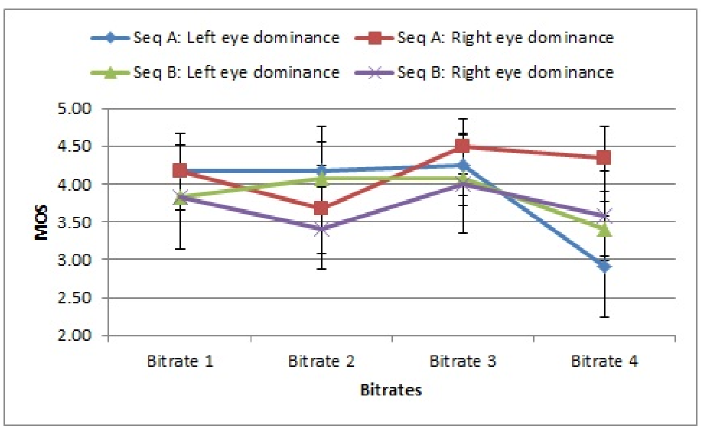



Fig. 4: Left-eye dominance vs. Right-eye dominance encoding approaches

- T-Test statistically insignificant (The probability of this result, assuming the null hypothesis, is 0.89)
- Spearman rank 0.4448
- Pearson correlation 0.3684

In case of the right eye dominance case, correlation with symmetric encoding;

- T-Test statistically insignificant (The probability of this result, assuming the null hypothesis, is 0.72)
- Spearman rank 0.5654
- Pearson correlation 0.5263

However, low ccorrelation indices indicate that they are loosely correlated with each other. The high MOSs of asymmetric encoding suggests that there is still an opportunity to further compress the 3D surgery video. This will further gain in quality compared to symmetric encoding. These results suggest that asymmetric video coding approaches as described in this paper could effectively be deployed for reducing the bandwidth requirements of 3D surgery video without affecting the diagnostic or perceptual quality at moderate bitrate levels in general. In addition, it is necessary to analyze the complete range of bitrates to find the true perceptual bounds of asymmetric encoding which will be addressed in our future work. The reliability of the results could be further improved by inviting more medical professionals to evaluate these processed sequences (i.e., surgeons). The confidential interval is high in general for the obtained dataset. Therefore, the reliability of these results could be further improved by inviting more surgeons to evaluate these sequences.

## IV. Conclusion

This paper elaborates on potential compression techniques and methodologies which can be used with 3D medical video. The challenges faced by emerging 3D medical video applications are discussed in brief. An asymmetric coding approaches for 3D medical video is evaluated in this paper for reducing the space/bandwidth needed for 3D medical video storage and transmission applications. The simulations and quality evaluations are performed to investigate the performance of the proposed asymmetric coding approach. The symmetric coding method is considered as the reference method in this study. The eye dominance of subjects is also considered in this study to see its effect. Subjective results provided by surgeons show

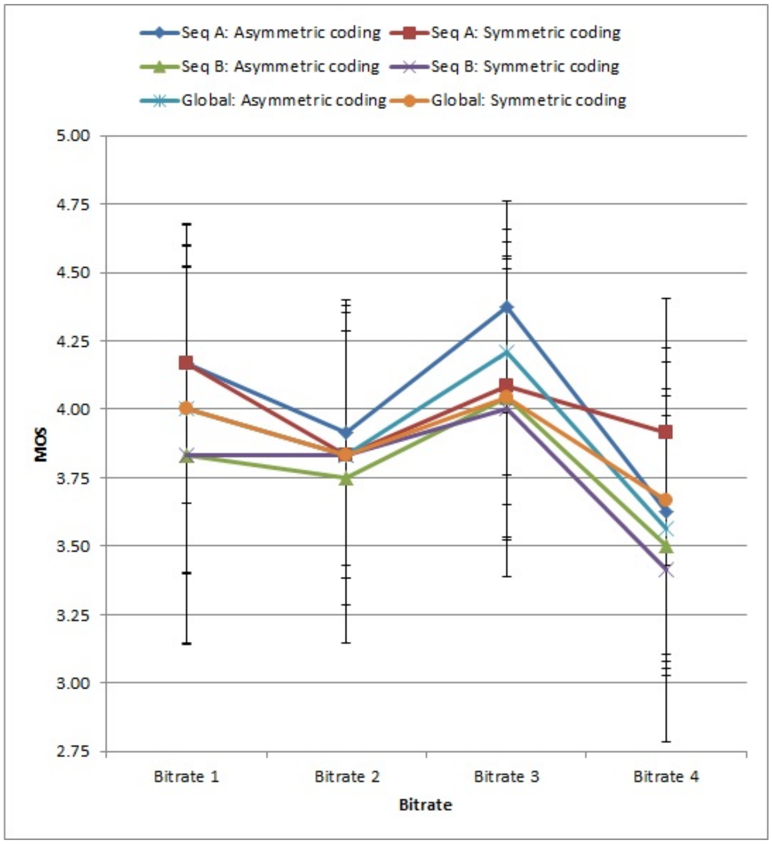



Fig. 5: Asymmetric vs. symmetric encoding approaches

that asymmetric encoding slightly outperforms the symmetric encoding approach at all the considered bitrates. However, the results are statistically insignificant. The results for left and right eye dominance show a significant differences between each other. This suggests that eye dominance plays a crucial role in selecting asymmetric encoding approaches. Further tests are necessary to find the optimum compression range for asymmetric coding to outperform its counterpart symmetric coding approach in the quality range of interest.

## Acknowledgment

This work was supported in part by the EU FP7 Programme, CONCERTO (Grant no: 288502).

## References

[1] Visionsense Corp., Internet: http://www.visionsense.com/profile.html, 2011.

[2] Wasol Ltd., Internet: http://www.wasol.co.kr/english/index_eng.php, 2010.

[3] B. Randolph, A neural theory of binocular rivalry, *Psychological review*, vol. 96, no. 1, pp. 145-167, 1989.

[4] A. Aksay, C. Bilen, E. Kurutepe, T. Ozcelebi, G. B. Akar, M. R. Civanlar, and A. M. Tekalp, Temporal and Spatial Scaling for Stereoscopic Video Compression, *IEEE 14th Eur. Signal Process. Conf. EUSIPCO 2006*, Florence, Italy, Sept. 2006.

[5] L. Christodoulou, L. Mayron, H. Kalva, O. Marques, and B. Furht, Design and Evaluation of 3D Video System Based on H.264 View Coding, *International Workshop on Network and Operating Systems Support for Digital Audio and Video (NOSSDAV 2006)*, Newport, Rhode Island, May. 2006.

[6] C.T.E.R. Hewage, S. Worrall, S. Dogan, H. Kodikara Arachchi and A.M. Kondoz, "Stereoscopic TV over IP", *Proceedings of the 4th IET European Conference on Visual Media Production (CVMP'2007)*, London, UK, Nov. 2007.

[7] X. Changman, Z. Zhaoyang. Data compression of stereoscopic image pairs, *Proceedings of SPIE*, vol. 4551, pp. 90 94, 2001.

[8] L. Stelmach, W.J. Tam, D. Meegan, A. Vincent. Stereo Image Quality: Effects of Mixed Spatio-Temporal Resolution, *IEEE Transactions on circuits and systems for video technology*, vol. 10, no. 2, pp. 188 - 193, 2000.

[9] C.T.E.R. Hewage and M.G. Martini. "Quality of experience for 3D video streaming." *Communications Magazine*, vol. 51, no. 5, pp. 101 - 107, 2013.

[10] C.T.E.R. Hewage S. Worrall, S. Dogan, H. Kodikaraarachchi, and A.M. Kondoz, "Stereoscopic TV over IP." *4th IET European Conference on Visual Media Production, IETCVMP*, 2007.

[11] R. Smith, A. Day, T. Rockall et al, Advanced stereoscopic projection technology significantly improves performance of minimally invasive surgical skills, *Surg. Endoscopy, Springer*, vol. 26, no. 6, pp. 1522-1527, June 2012.

[12] C.T.E.R. Hewage, H.A. Karim, S. Worrall, S. Dogan, and A.M. Kondoz, "Comparison of stereo video coding support in MPEG-4MAC, H.264/AVC and H.264/SVC", *In: 4th IET Visual Information Engineering conference*, 25-27 July 2007, London, UK.

[13] M. Martini, C.T.E.R. Hewage, N. Moustafa, R. Smith, I. Jourdan and T. Rockall, "3-D robotic tele-surgery and training over next generation wireless networks"', *In: 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 3 - 7 July 2013, Osaka, Japan.

[14] C.T.E.R. Hewage, H.D. Appuhami Ralalage, M.G. Martini, R. Smith, I. Jourdan, and T. Rockall, "Quality evaluation of asymmetric compression for 3D Surgery Video"', *In: 2013 IEEE 15th International Conference on e-Health Networking, Applications and Services (Healthcom)*, 9-12 Oct 2013, Lisbon, Portugal.