3-D Robotic Tele-Surgery and Training over Next Generation Wireless Networks

Maria G. Martini, Chaminda T. E. R. Hewage, Moustafa M. Nasralla,

Ralph Smith, Iain Jourdan, and Timothy Rockall

***Abstract*— The latest advances on robotic surgery enable the performance of many surgical procedures by utilizing minimally invasive techniques. In particular, recent 3-D endoscopes have improved the performance of minimally invasive surgical procedures. Based on these advances, performing or visualizing in real-time surgical procedures at a distance can be envisaged. In this paper, we present a performance evaluation of 3-D robotic tele-surgery and training over next generation wireless networks, namely wireless networks based on the long term evolution (LTE) 3GPP standard. Different scheduling strategies are compared and results are analyzed in term of the resulting quality of experience (QoE) for the surgeon.**

I. Introduction

The enhanced depth perception produced by recent 3-D endoscopes has been demonstrated to improve the performance of minimally invasive surgical procedures. Three-dimensional imaging also facilitates the training of minimally invasive surgery and may lessen the learning curve of these technically demanding procedures. The increased depth of field facilitates intricate minimally invasive surgical procedures [1], it allows better recognition of tissue layers and may facilitate complex maneuvers such as laparoscopic suturing or knot tying [2]. Skill tests performed assessing laparoscopic suturing and knot tying demonstrated a 25% increase in speed and accuracy of these laparoscopic tasks when utilizing a three-dimensional video system as compared to a standard two-dimensional endoscopic video system [3]. Three-dimensional video systems facilitate surgical tasks in general and benefits are particularly evident for inexperienced laparoscopic surgeons.

After the very first tele-surgery test with a human patient [4], tele-surgery tests are reported in [5], where video was encoded according to the MPEG-2 standard, through DVTS uncompressed quality video, or through commercial video codecs. For 3-D video, left and right views were transmitted separately and the trade-off delay-quality was recognized as a major issue. It is then important that a remote surgery system is designed with the goal of meeting a required QoE determined by the users (medical specialists and trainees).

Lossless compression techniques are often considered in the medical imaging area, although these have the disadvantage of a reduced compression ratio. Therefore, when transmission is over band-limited and error-prone channels, a compromise must be made between compression fidelity and protection and resilience to channel errors and packet loss. It has been estimated that lossy compression ratios of 1:5 to 1:29 do not result in a lowering of diagnostic accuracy for medical images [6].

In this paper, lossy video compression is considered for medical video transmission over wireless channels for telemedicine and medical education purposes. Furthermore, simulcast/parallel encoding of left and right stereoscopic video is considered in this work. This enables independent control of left and right views. In order to provide better diagnostic quality, a low Quantization Parameter (QP) is employed. The requirement of good image quality for diagnosis and surgical procedures can hence be fulfilled by a simulcast configuration at low compression levels. We consider the case where 3-D medical video sequences acquired through a robotic surgery system are transmitted real-time over a wireless system and we investigate the acceptability of the results in terms of quality of experience for the case of medical education and for wireless robotic tele-surgery. We refer to the latest 3GPP standard for wireless cellular communications, i.e., the Long Term Evolution (LTE) standard and its advanced version LTE-Advanced (LTE-A). Specifically, we test the performance of real-time 3-D surgery video transmission with different scheduling strategies and we compare the relevant results. The requirement of near real-time transmission derives from the need of interaction with the real scene in the case of surgery and it is a desirable feature also when video is transmitted for education purposes (training of surgeons).

The remainder of this paper is organized as follows. Section II introduces the proposed transmission strategy for 3-D medical/surgical video over an LTE-A wireless network. The experimental setup and results are discussed in Section III. Section IV concludes the paper.

II. Proposed Transmission Strategy

The proposed 3-D medical/surgical video transmission system over LTE is illustrated in Figure 1. Left and right video sequences captured from a 3-D endoscope are fed into the 3-D video encoder module for compression. The 3-D video encoder consists of two video encoders running in parallel. The output of the 3-D video encoder module is then fed into the multiplexer module. This module multiplexes the encoded left and right bit-streams together, and performs packetization. Finally the multiplexed packets are transmitted over the LTE network to the intended destination. Figure 2 depicts the network architecture used

* M. G. Martini and C. Hewage acknowledge the support of the EU FP7 Programme (CONCERTO project).

Maria G. Martini, Chaminda T. Hewage, and Moustafa M. Nasralla are with Kingston University London, UK (email: m.martini@kingston.ac.uk).

Ralph Smith, Iain Jourdan, and Timothy Rockall are with Minimal Access Therapy Training Unit (MATTU), UK.

when 3-D videos are transmitted over LTE to multiple users in the cell.

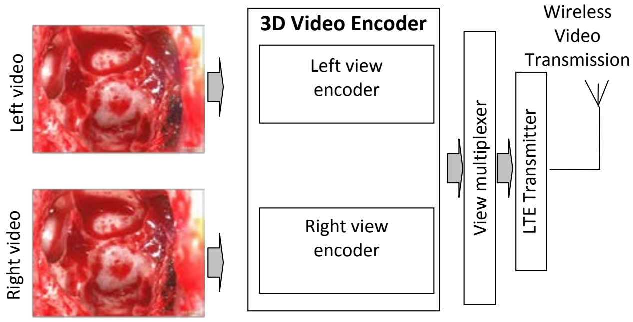

Figure 1. Block diagram of the proposed 3-D medical/surgical video transmission system.

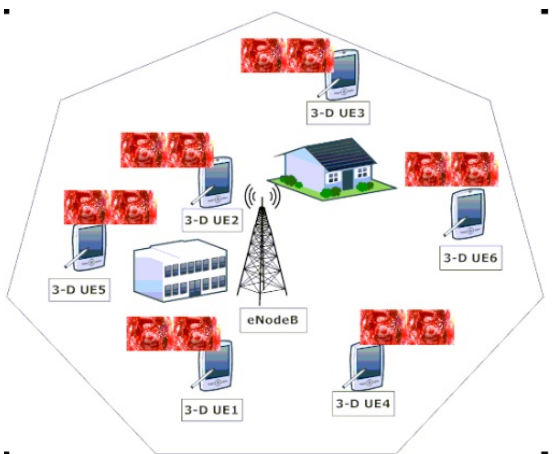

Figure 2. 3-D video users in the LTE radio network.

A. *LTE Scheduling algorithms*

In conjunction with the increasing growth of multimedia, internet and real time services, the LTE wireless standard has emerged to cope with these services efficiently. LTE has been introduced by the Third Generation Partnership Project (3GPP) as the next technology after the 3.5G (HSPA+) cellular networks. The system architecture of the 3GPP LTE system contains several base stations called "eNodeB" (evolved node B) where the packet scheduling process is performed along with other radio resource management (RRM) tasks. Resource allocation is performed by every transmission time interval (TTI) whose duration is one ms.

Various packet scheduling algorithms have been developed to support real-time (RT) and non real time (NRT) flows, such as Proportional Fair (PF), Modified-Larges Weighted Delay First (M-LWDF), and exponential PF (EXP/PF) [7][8][9]. In the aforementioned schedulers, each radio bearer is assigned a priority value by considering specific metrics. The bearer with the best metric is scheduled first at the next TTI.

Scheduling algorithms differ in the way they calculate the metric considered for the assignment of the resources to the different users. In most of the cases, scheduling algorithms require the knowledge of the average transmission data rate (R_i) of each flow associated to user i, and the instantaneous available data rate for each sub-channel. As a result, when information about the performance guaranteed in the past for each flow is available, the system can perform fairness balancing.

The PF scheduler assigns radio resource blocks by taking into account both the experienced channel quality and the past user throughput [7]. The goal of this scheduler is to optimize the total network throughput and to guarantee fairness among the flows. The M-LWDF scheduler was developed to support multiple data users with different QoS requirements [8]. When RT flows are considered, a packet delay threshold t_i is considered in the scheduler's metric. Hence, this scheduler considers the best channel condition and the highest delay for the RT flows` Head of Line (HoL) packet in its allocation scheme. The EXP/PF scheduler is developed to maximize the priority of RT flows with respect to NRT ones. Also, this scheduler is designed to deal reliably with both RT and NRT users [9].

III. Experimental Setup, Results and Discussions

A. *3-D Medical Video*

Left and right based stereoscopic video content is considered for this investigation. The 3-D medical video contents are provided by the Visionsense Corp., USA. Visionsense produces a miniature stereoscopic (3-D) sensor that optically maps the surgical field. Two sample left and right frames produced using the Visionsense 3-D endoscope are shown in Figure 3. These two selected sequences (acquired during different surgical procedures) are used in the experiments.

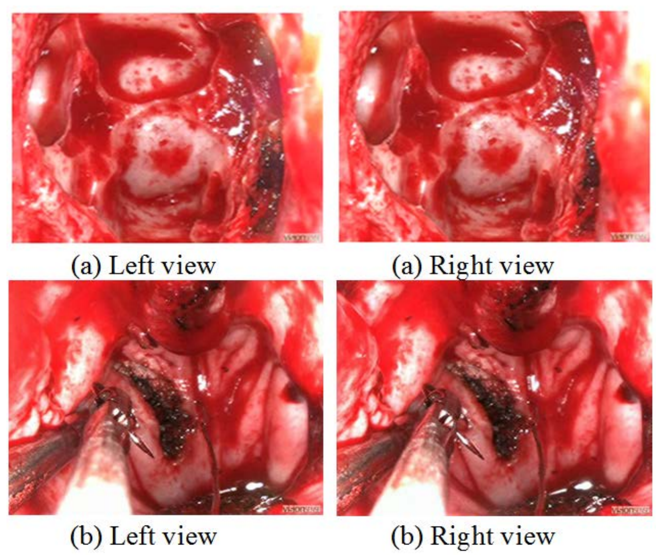

Figure 3. Two sample 3-D endoscope videos (a) Pituitary (b) Clivus_biopsy_suture.

The 3-D test sequences are encoded using the H.264/AVC video coding standard (JM reference software Version 16.0). Two codecs are employed to encode left and right views separately. Twenty-second long sequences (i.e., 1000 frames at 50 fps) are encoded with *IPPP…IPPP*... sequence format, using quantization parameter (*QP*) value 30 for both intra-coded (*I*) and predicted (*P*) frames. An *I* frame is inserted by every 1s of the video sequence to minimize the effect from temporal error propagation. 500 Bytes Slices are also introduced in order to make the decoding process more robust to errors. Our previous study reported in [10] suggests that 500 Bytes packet size is suitable to provide a good 3-D video quality over LTE-like networks.

In order to obtain stable results, simulations are run for several times and the image quality is averaged over 1000 frames. The corrupted bit-streams are later decoded using the JM reference software decoder. Slice copy is enabled during decoding to conceal the missing slices/packets of the corrupted bit-stream. At each error condition the left and right view quality is measured using the peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) quality metrics.

In order to obtain true opinions of medical practitioners, subjective tests are conducted by the surgeons at The Minimal Access Therapy Training Unit (MATTU). MATTU is an internationally recognized centre of excellence created for the teaching of innovative laparoscopic surgical techniques. The Mean Opinion Scores (MOS) are then considered alongside objective measures to evaluate the performance of 3-D medical video transmission over LTE. Subjective tests were performed using the Double Stimulus Impairment Scale (DSIS) method. The sequences are displayed on a LGCF 3-D projector and a 3x2m silver screen with passive circular polarizing eyewear. The content is evaluated in a room with low lighting levels at a distance of 5 meters.

B. LTE system

This paper investigates the performance of PF, M-LWDF, and EXP/PF when 3-D traffics are transmitted over LTE systems. The scenario used in this process is a single cell with interference. We have set up a scenario where there are 6 3-D users. Users are randomly distributed with uniform distribution in the cell with distance from the *eNodeB* ranging from 10 m to 900 m. We assume users are moving at a pedestrian speed of 3 km/h in random directions (*i.e.*, the speed direction is chosen randomly for each user and it remains constant during the time and moves towards the boundary area. Once the user reaches the boundary area, the user chooses a new speed direction). The LTE propagation loss model is composed by four different models (path loss, penetration loss, shadowing, and multipath) [11].

The serving LTE base station (*eNodeB*) is located at the center of the cell. In the *eNodeB* a scheduler controls all the available physical resource blocks (PRBs) by assigning them to the active flows which are competing for resources. The LTE-Sim simulator is used to simulate this scenario [12]. LTE-Sim gives the possibility to make use of resource allocation in a time-frequency domain. In the time domain, resource allocation is performed by every TTI (*i.e.*, every 1 ms), which is exactly by every two consecutive resource blocks. Two time slots of 0.5 ms compose one TTI which consists of 14 orthogonal frequency division multiplexing (OFDM) symbols with short cyclic prefix. Ten consecutives TTIs compose the LTE frame. Simulation parameters are shown in TABLE I.

C. Results

The performance of the transmission of 3-D tele-robotic surgery videos using three different scheduling algorithms for LTE is evaluated using both objective and subjective quality evaluation methods for the received 3-D video. Quality evaluation studies are carried out using the sequence captured during the “Pituitary” surgical procedure. The average PSNR and SSIM of left and right video are considered as a measure of objective quality of 3-D video. The average PSNR and SSIM image quality results for the scenarios of “4 Users” and “5 Users” are shown in TABLE II and TABLE III respectively. According to TABLE II and TABLE III, most of the users achieve reasonable PSNR and SSIM image quality ratings. The average received image quality is better in the scenario of “4 Users” compared to “5 Users”. User 2, 4 and 5 in TABLE III achieve a lower quality left and right image (less than 30dB) and this may not be suitable for clinical applications. The reduced quality in the “5 Users” scenario is mainly due to the scarce resources in the LTE transmission system for the given bandwidth (10MHz). It can be observed that the “M-LWDF” scheduling algorithm achieves better image quality compared to “EXP/PF” and “PF” methods. Therefore, it can be concluded that the “M-LWDF” scheduling method provides improved 3-D image quality over next generation wireless networks such as LTE. This suggests that the improved quality achieved with “M-LWDF” method will facilitate 3-D tele-robotic surgery with respect to other conventional scheduling methods for next generation wireless systems.

TABLE I
LTE DOWNLINK SIMULATION PARAMETERS

Parameters	**Values**
Simulation duration	40 s
Flows duration	19.980 s
Frame structure	FDD
Cell radius	1 km
E-UTRAN frequency band	1 (2.1 GHz)
Bandwidth	10 MHz
Number of PRBs	50
Slot duration	0.5 ms
Scheduling time	1 ms
Max delay	0.1 s
Average video bit-rate	2201.19 kbps
Propagation loss/ Channel model	Typical Urban /Pedestrian A

TABLE II
AVERAGE PSNR AND SSIM QUALITY RESULTS FOR “4 USERS”

"4 Users"	**Average Left and Right PSNR (dB)**			
Method	**User 1**	**User 2**	**User 3**	**User 4**
MLWDF	31.905	34.075	31.905	36.520
EXPPF	31.010	32.820	31.175	36.520
PF	29.640	32.950	30.830	36.520
"4 Users"	**Average Left and Right SSIM**			
Method	**User 1**	**User 2**	**User 3**	**User 4**
MLWDF	0.8985	0.9194	0.9026	0.9383
EXPPF	0.8824	0.9149	0.8869	0.9383
PF	0.8715	0.9159	0.8834	0.9383

Surgeons at MATTU took part in the subjective evaluation of 3-D surgical video transmitted using different scheduling algorithms. The opinion scores (OSs), using the DSIS method, were marked on a 5-point categorical scale, namely: 5. "Imperceptible" (same as the original); 4. "Perceptible, but not annoying"; 3. "Slightly annoying"; 2. "Annoying"; 1. "Very annoying". All individual Opinion Scores were later averaged to obtain the MOS value for a given test.

TABLE III
AVERAGE PSNR AND SSIM QUALITY RESULTS FOR "5 USERS"

"5 Users"	Average left and right PSNR (dB)				
Method	**User 1**	**User 2**	**User 3**	**User 4**	**User 5**
MLWDF	33.82	28.07	31.77	27.57	27.18
EXPPF	33.98	26.28	30.74	27.88	27.64
PF	33.38	24.83	31.88	27.31	27.66
"5 Users"	**Average left and right SSIM**				
Method	**User 1**	**User 2**	**User 3**	**User 4**	**User 5**
MLWDF	0.92	0.84	0.91	0.85	0.85
EXPPF	0.92	0.83	0.90	0.86	0.85
PF	0.91	0.81	0.90	0.85	0.85

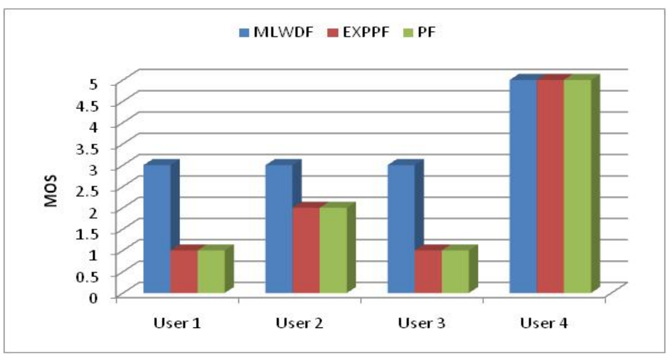

Figure 4. Subjective 3-D video quality for "4 Users" case.

Figure 4 shows the MOS for three scheduling algorithms over LTE. The MOS values are low in general due to packet losses occur during wireless transmission. Since errors in tele-robotic surgical video can lead to wrong diagnosis/observations and inaccurate interventions, the surgeons have provided quite low quality scores for all the test cases. However, the "M-LWDF" scheduling algorithm provides the best result compared to "EXP/PF" and "PF" methods. This suggests that the transmission of tele-robotic surgical video over resource limited wireless channels may affect the performance of the surgical procedure. The received quality, however, was considered acceptable in the case of four users for training purposes. Optimized exploitation of network resources and suitable scheduling algorithm as suggested by this paper (*i.e.*, "M-LWDF" method) could be effectively used to enable 3-D tele-surgical training over next generation wireless channels. Furthermore, it is expected that increasing the fairness of the scheduling algorithm further would result in suitable medical quality for all the users in the cell. With the considered bandwidth, the maximum number of 3-D medical users supported with reasonable quality is four (the surgeons were not satisfied of the results in the scenario with 5 users in the cell – results of MOS tests are not reported here for brevity).

IV. Conclusion

The paper presented a performance evaluation of 3-D robotic surgery video transmission over a wireless network. Three different scheduling strategies have been compared. The performance was evaluated in terms of network parameters, objective video quality metrics, and subjective video quality assessment tests. The paper also enables a mapping among the different quality metrics used in the paper, from network parameters, to 2-D image quality metrics, and 3-D perceptual quality.

It has been shown that the M-LWDF scheduling strategy results in the best performance with all the evaluation criteria and, for the case of an LTE system with 10MHz bandwidth, a maximum of four users can be supported with acceptable quality from the clinical point of view (MOS quality evaluation from surgeons > 3 for all users). The results also suggest that it is expected that increasing the fairness of the scheduling algorithm further would result in a globally improved medical quality for the users in the cell.

References

[1] 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.

[2] G. Janetschek. A. Reissial, R. Peschel, H. Drax and G. Bartsch, "Chip on a stick technology: First clinical experience with this new video-laparoscope", *J. Endourol.*, vol. 7, 1993.

[3] R.K. Bahayan, A.W. Chiu, J. Este-McDonald and D.H. Birkett, "The comparison between 2-dimensional and 3-dimensional laparoscopic video systems in a pelvic trainer", *J. Endourol.*, vol. 7, 1993.

[4] J. Marescaux, J. Leroy, M. Gagner et al., "Transatlantic robot-assisted telesurgery", *Nature*, vol. 413, no. 6854, pp. 379-380, September 2001.

[5] J.R. Sterbis, E.J. Hanly, B.C. Herman, M.R. Marohn, T.J. Broderick, S.P. Shih, B. Harnett, C. Doarn, and N.S. Schenkman, "Transcontinental Telesurgical Nephrectomy Using the da Vinci Robot in a Porcine Model", *Urology*, `vol. 71, no. 5 pp. 971-973, 2008.

[6] P.C. Cosman, H.C. Davidson, C.J. Bergin, et al., "Thoracic CT images: effect of lossy image compression on diagnostic accuracy", *Radiology*, vol. 190, no. 2, pp. 517-524, Feb. 1994.

[7] J.G. Choi and S. Bahk, "Cell-throughput analysis of the proportional fair scheduler in the single-cell environment". *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 766–778, Mar. 2007.

[8] P. Ameigeiras, J. Wigard, and P. Mogensen. "Performance of the M-LWDF scheduling algorithm for streaming services in HSDPA". *IEEE 60th Vehicular Technology Conference (VTC2004-Fall)*, vol. 2, Los Angeles, USA, September 2004, pp. 999-1003.

[9] R. Basukala, H. Mohd Ramli and K. Sandrasegaran. "Performance Analysis of EXP/PF and M-LWDF in Downlink 3GPP LTE System". *IEEE F. Asian Himalayas Conf.*, Kathmandu, Nepal, Nov. 2009.

[10] C.T.E.R. Hewage, M.G. Martini, and N. Khan, "3-D medical video transmission over 4G networks", *In: 4th International Symposium on Applied Sciences in Biomedical and Communication Technologies*, Barcelona, Spain, 26 - 29 Oct 2011.

[11] Tech. Specif. Group Radio Access Network 3GPP. "*Physical layer aspect for evolved universal terrestrial radio access (utra) (release 7)*", Technical report, 3GPP TS 25.814.

[12] G. Piro, L. A. Grieco, G. Boggia, F. Capozzi, and P. Camarda, "Simulating LTE cellular systems: an open source framework", *IEEE Transactions on Vehicular Technology*, vol. 60, no. 2, pp. 498 -513, October 2010.