

Tutamen: An Integrated Personal Mobile and Adaptable Video Platform for Health and Protection

David Palma¹, Joao Goncalves¹, Luis Cordeiro¹, Paulo Simoes²,
Edmundo Monteiro², Panagis Magdalinos³, and Ioannis Chochliouros⁴

¹ OneSource, Consultoria Informatica, Lda.
{palma, john, cordeiro}@onesource.pt

² Centre for Informatics and Systems of the University of Coimbra
{psimoes, edmundo}@dei.uc.pt

³ National and Kapodestrian University of Athens
panagis@di.uoa.gr

⁴ Hellenic Telecommunications Organization (OTE) S.A.
ichochliouros@oteresearch.gr

Abstract. A framework for mobile and portable High-Definition Video streaming is proposed, developed and assessed. Suitable for emergency scenarios, involving for instance ambulances and fire-fighters, the presented framework resorts to a state-of-art platform which considers off-the-shelf hardware and available video codecs for High-Definition Video. The obtained results show that the proposed architecture is able to efficiently support rescuing teams in the demanding scenarios where they operate, guaranteeing video quality and ease of use. This solution is particularly useful for situations where experts in the fields can accurately provide their insights and contributions remotely and in a timely fashion.

Keywords: HD video, Streaming, Smart Cities, Mobile First Responders.

1 Introduction

Everyday advances in technology and its availability within an increasing number of communities has motivated new approaches towards technological innovation in many sectors. A valuable contribution in this aspect is the definition of Smart Cities, where technologies can be used to provide dynamic and adaptive services and contexts. Among several topics, the LiveCity project, in which this work has been developed, is concerned with sustainability as well as with a reliable, smart and secure cities. By exploring the advances in wireless communications, in conjunction with the portability of increasingly more powerful computers, this work aims at extending the benefits of integrated personal mobility for demanding scenarios involving aspects such as healthcare and civil protection.

Previous works have already proposed many end-to-end mechanisms connecting hospital experts with ambulances and similar agents in emergency scenarios.

For instance, not only simulation based but also real-time non-diagnostic systems have been developed, capable of providing trauma and echo-cardiogram videos [1]. Nonetheless, these solutions are strongly dependent on the entire infrastructure provided by ambulances, being limited in terms of their operational area, requiring patients to be inside the vehicle. Moreover, the addition of features such as the monitoring of other vital signals, a clear image of the patients' face or any non-digital medical equipment is non-trivial in these solutions.

Nowadays a big challenge is the design of mobile and ubiquitous telemedicine, versus the known store and forward telemedicine which typically demands tightly-coupled services and rigid architectures [2]. The Tutamen proposal, takes its name from the same latin word which stands for protection, defence and caring. It tackles flaws such as lack of real-time communication between rescuing teams, the portability and flexibility that will allow its usage in a seamless and effective manner, defining a new perspective towards the exploitation of technology for effectively improving the quality of living both in urban and remote areas. Moreover, this proposal is expected not only to allow a physician to be available anywhere at anytime, but also to follow a well defined methodology with field trials and real feedback from the involved professionals, which are typical disregarded by other solutions.

Other recent solutions, such as the eBag [3], require teams operating in the field to carry a suitcase equipped with a full-size laptop and other equipment. This not only creates difficulties in transportation but also in setup. When responding to an emergency scenario, timing is critical and other similar solutions such as the use of robotic agents [4], makes the operation more cumbersome, possibility creating trust issues with patients regarding their efficacy.

The technological challenges raised by the need of providing an ubiquitous and real-time high definition (HD) stream have been addressed in three different approaches, guaranteeing that a modular solution would be achieved. By separating the proposed Tutamen architecture into three blocks - video acquisition, streaming and interfacing - the personal video-to-video (v2v) system was designed taking into account the possibility of future hardware and software updates.

This work presents the Tutamen framework, concerning both wearable and flexible hardware and the software mechanisms used for adaptable real-time video streaming, taking into account the available network conditions.

A general description of the LiveCity project in which the mobile video platform has been developed is presented in Section 2, followed by a detailed presentation of the proposed solution in Section 3, considering a preliminary use case for ambulances, requirements and methodology in Section 4. The assessment of the performance and main characteristics of the Tutamen personal unit are included in Section 5. Finally, Section 6 highlights the main contributions provided by this work.

2 The LiveCity Project

Empowering inhabitants of a city by providing them means to interact with each other in a farther prolific, effectual and socially useful way, by the use of high quality v2v over the Internet, could be presented as the main goal of LiveCity Project. The proposed augmentation could be used to save patients lives, improve city administration, reduce fuel costs, reduce carbon footprint, enhance education and improve city experiences for tourists and cultural consumers.

LiveCity promotes the context of a modern world where live HD interactive v2v can be easily available on a variety of display devices and where a video call with HD quality can be as globally reachable to a city environment as a plain old telephony call; or even a world where any video screen coupled with a video camera in a city could connect an HD full screen video call (where HD video represents moving pictures with at least a resolution similar to 1280 pixels wide and 720 pixels height) at an attractive cost with seamless, utterly simple usability. This is the vision of the essential LiveCity approach. To realize its strategic aims, LiveCity considers standard video encoding already available in off-the-shelf devices.

The LiveCity effort aims to fill the above matters, per separate option, by providing appropriate responses and by suitably promoting experiences and results gained from pilot actions.

3 Camera Data Acquisition Prototype

Initially developed for emergency first responders - namely paramedics units deployed in emergency ambulances - so as to provide a balanced compromise between computational resources (necessary to acquire, encode and transmit live HD video), mobility (wearable device, battery autonomy and hands-free operation) and good performance under extreme conditions, mostly due to the inclusion of rugged devices, video camera able to self-adjust to variable lighting environments and to the possibility of integration with emergency responders equipment, as shown in Figure 1.



Fig. 1. Equipped Member

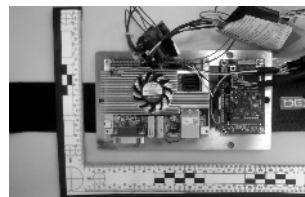


Fig. 2. Micro Computer

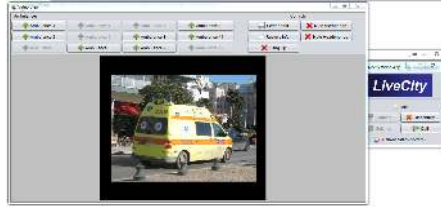


Fig. 3. Application layout deployed in the hospital side

The system, with its core component depicted in Figure 2, supports any type of Global System for Mobile Communications (GSM) or Long Term Evolution (LTE) based networks by the means of a Universal Serial Bus (USB) dongle. Composing the system is an x86 compliant dual core board, coupled with an High-Definition Multimedia Interface (HDMI) Mini Card capture device which interfaces with a Mini Peripheral Component Interconnect (PCI) Express expansion card needed by the main board to support this kind of device. External to the main component, we have buttons connected to the machine, allowing the user to control and change states of the current call. Travelling through the back of the user, is a sleeve with cabling to connect the external camera and headset, providing the user with audio feedback and the hospital with a realtime video and audio feed. This setup has been developed within the LiveCity project and its performance results for v2v communication are presented in this paper.

4 Use Case – Ambulances

During the so called "golden hour", patients suffering from an heart attack, on the verge of having a stroke, or suffering from some type of environment induced trauma, could significantly benefit from speed about decision making.

Early recognition of stroke features and opportune referral to a stroke unit by the means of v2v enabled technology should improve overall outcome, saving lives and improving the patient chance of independently functioning, as well as an expedited advice concerning optimal pain management and requirement of evacuation on trauma victims.

4.1 Requirements and Methodology

The most important step in the design and implementation of the prototype was the extraction of a detailed list of requirements from the involved user groups. Due to the focus of the application, this effort was of paramount importance and had to be designed and implemented in every possible detail. Most of these requirements were assessed by working closely with emergency personnel from several different areas, but mostly, first responders.

From a methodological point view we applied an iterative and incremental process (closely resembling the concept of Rational Unified Process [5]) implemented

through consecutive cycles of interviews and subsequent software releases. The latter facilitated timely feedback from the involved user groups while in parallel assured that the development process will not result in an application which although functional is useless to the end user.

In the context of this paragraph we will try to briefly present this process focusing primarily in end user requirements, how they were captured and what they implied for the subsequent design and implementation process. The end product is a v2v software application with an end user interface which offers the doctor an opportunity to intervene early in emergency situations and offers the paramedic the opportunity to get medical assistance in a difficult situation.

We look at the three use case scenarios of heart attack, stroke and trauma together for the purposes of the hardware and software development. The storyboard for the emergency use case is the following:

- The paramedic initiates a video-call to the hospital using a portable device.
- In case the doctor is directly available, an alert is sent to his smart phone.
- The doctor proceeds to the emergency room and answers the call.
- The paramedic speaks into the microphone held on his headband and explains the issue.
- The doctor looks at the patient via the monitor and instructs the paramedic to focus on the part of the patient of interest.
- The doctor asks questions to the patient in order to assess his condition.
- The doctor instructs the paramedics on how to proceed.

The latter provided the first basis for design of the application. However, more user specific requirements came after the first release of the application. The involved user groups have different requirements; paramedics need a simple application with a maximum number of two or three buttons that will enable them to communicate with the hospital with minimal overhead; on the other hand the doctors required an application that supports full management of the incoming video and audio streams. Thus, following the well known Model-View-Controller software design concept, two different flavours for the end user application were designed.

From the hospital perspective, the application caters for stream management (video and audio stream management), notifications to the end user (SMS texting in case of absence, highlight of emergency call), capturing of still images from the video source, manual and/or dynamic adaptation of the video/audio quality and encrypted storage of patient information. Finally, the doctor can initiate a call to an ambulance; however the latter is permitted if and only if a call has been received by that ambulance. Figure 3 provides the layout.

From the paramedics' point of view, the application should be simple and easy to use with a minimum number of buttons facilitating communication with the hospital. All notifications and interruptions should take place through a simple beep-code thus enabling the ambulance crew to minimize the amount of time being distracted because of the application.

Not only the application has gone through the scrutiny of the paramedics', but also the hardware configuration of the whole platform. Targeted at personnel

with different requirements, one of the most requested features was modularity. Being able to switch the position of the computer box and how the cabling flows through the belt and hoses, became mandatory.

5 Main Results

Aiming to provide a proper evaluation of the platform, measurements regarding the overall performance of the platform during the process of video acquisition, processing and transmission will be performed. Two different types of measurements will be made; firstly the resulting Peak Signal-to-Noise Ratio (PSNR) of the codec while encoding the acquired video into the different bit rates and resolutions shall be measured; secondly a measurement of the transmitted video's PSNR, using different bit rates and resolutions, will be obtained as well. For the sake of simplicity, only the performance of one video codec will be reviewed and the choice will fall upon the open implementation of the H.264/MPEG-4 AVC video codec, x264, since it is considered to be very robust for video streaming in wireless networks [6].

In order to measure the encoding performance of the codec, its internal measuring tool will be used to retrieve the global PSNR, but with the caveat of not enabling its optimizations for PSNR monitoring. The reason behind this decision has to do with the fact that it is intentional to measure the codec performance with the optimizations that were selected for the Tutamen live streaming, and not the optimizations that achieve a most advantageous PSNR during the encoding process.

An additional quality metric that is typically considered is the Mean Opinion Score (MOS) which indicates the quality of a video and is quantified from 1 (bad quality) up to 5 (excellent quality). However, this metric is a sensorial metric, which means it depends how the quality perceived by humans, therefore not necessarily providing accurate results when calculated [7].

The video quality assessment will be performed taking into account the degradation introduced both by the transcoding and streaming processes. For comparison purposes, the raw FullHD video will be considered as a reference value and compared against the transcoded videos at different bit-rates prior to and after streaming. In fact, variable bit-rates will be used during the streaming of the video, adapting to the available network conditions. However, the reduction of the videos' bit-rate will lower the initial PSNR, therefore the bit-rate reduction must be sensible. In order to measure the PSNR resulting from the transcoding and streaming process, the Evalvid tool [8] and its known methods will be used.

5.1 Usability / Wearable Characteristics

Bearing in mind the use cases proposed for the Tutamen platform, its usability and overall look and feel experience is of paramount importance. In fact, the definition of a comfortable and intuitive device for either physicians, fireman or any other unit working under emergency situations, has influenced not only

the design of the wearable communication unit but also of its mechanisms and interaction procedures with the remote units in control centres. Therefore, the results obtained from the process of creating the hardware-based interface and the communication module, provide important insights for future related works.

The developed work was closely followed by doctors and physicians which provided their contribute and knowledge regarding the main requirements and challenges in emergency situations. A key point was the ease in using the entire platform, not only should the device be easily carried, being light in both weight and size, but it should also be intuitively operated, requiring as few buttons as possible. These requirements resulted in a solution slightly exceeding 1kg that can be used by fitting an adjustable belt, capable of being worn with any uniform in any atmospheric conditions. Moreover, taking into account the ease of operation, a simple three button system was developed.

With the available buttons, the units carrying the Tutamen device are able to start and stop the entire system by using the power button as well as to start and stop a direct call to the control centre or hospital. Moreover, since some disaster scenarios lack the desirable lighting, the third button can be used to enable a powerful lighting system, capable of providing good quality video. Finally, due to the harsh conditions in which emergency units usually operate, the buttons module is protected with a cover and each operation (i.e. button press), requires a button press of at least two seconds, thus avoiding accidental operations.

An important requirement identified during the development of this work, in conjunction with emergency teams, was that the equipment should be as hands-free as possible. Therefore, the video acquisition unit, as well as the audio and lighting systems, have considered taking into account this aspect. Regarding the video and lighting systems, they have been designed so that both can be attached to a working helmet or to a headband, allowing the individual wearing the device to transmit and illuminate the desired regions of interest without any adjustments or configurations. Additionally, in order not to distract the person operating Tutamen, each action in the system is transmitted through a headset that is also used to transmit audio to the control centre.

5.2 Technical Performance

In order to correctly address the difficulties found by the operating teams in the field, the performance of the used hardware has been taken into consideration, influencing the choices for each component included in the equipment. The used camera is able to provide FullHD Video at several frame rates, guaranteeing a 170° field of view through a wide angle lens, for up to 2.5 hours of continuous video. The camera has a dedicated power source which can easily be replaced or recharged.

Coupled with the camera acquisition unit, a supportive light system has been developed, providing a 52 Lumen output with a wide viewing angle as well. The lighting system not only has a small power consumption (1.1 Watt), which increases the total working time of the Tutamen solution, but also it is able to operate for 50,000 hours without having to be replaced.

Regarding the overall system's autonomy, the used batteries are able to support the platform up to 3 hours in full load, offering the possibility of being replaced by other charged batteries while the empty batteries can be recharged in a separate module. The running time of the system can be vastly extended if the calls to the coordination centre are only periodic. Moreover, the system can be easily shut-down and booted into working mode in a matter of seconds.

Considering the audio system, which will provide most of the feedback to the physician, a single-sided ear coupling Sennheiser headset is used for system related events (i.e. warnings and notifications such as system ready, new call arrived, among others), while at the same time it is also used for the communication between the physician and the control centre. The headset includes also an adjustable microphone, a button to mute the microphone and a volume regulator for a better experience.

5.3 Calling Performance

The establishment of call between the field teams and the control centres involves the transmission of both audio and FullHD video. Since audio over the internet has already been widely studied, the call performance analysis will focus on the challenges posed by the transmission of high-definition video. Taking into account the variable and limited bandwidth of 3G Networks, different tests were performed with different encoding bit-rates of the video codec, reflecting the available bandwidth at a given time. This feature of the Tutamen architecture aims at guaranteeing the best possible quality at any given time.

Generally, when connecting to a 3G Wireless Network, the device knows the available bandwidth. However, the announced bandwidth does not always correspond to the real available bandwidth, for instance, while the High-Speed Downlink Packet Access (HSDPA) standard may announce 7.2Mb/s bandwidth, this value only represents the maximum bandwidth than can be achieved, providing no actual guarantees. The presented results show the quality of different bit-rate videos, transmitted through the same network conditions of 7.2Mb/s of bandwidth. Since many interferences may occur, the desired transmission bit-rate should be below the maximum available bandwidth.

The transcoding process at lower bit-rates may reduce the overall quality of the video. However, the results we present in this paper show that less harmful to previously adapt the video quality to the available network conditions, rather than maintaining a higher quality and incur packet losses that may carry key frames. In fact, the resulting degradation of a higher bit-rate video being transmitted over a lower bandwidth network, may be higher than the the degradation of a video previously transcoded with a lower bit-rate, thus with a forced reduction of the video quality before being streamed, which during the streaming process has few or no losses that compromise the video integrity.

Table 1 presents the percentage of degradation obtained for different bit-rate videos, considering the same network conditions. This illustrates the need to correctly adapt videos' bit-rates during the transcoding process, guaranteeing that the number of losses is reduced and that overall perceived quality is the best

Table 1. Degradation Comparison over a network of 7.2 Mbps

Bit Rates Mb/s	% of Degradation		
	Transcoding	Streaming	Overall
12.0	2.17	33.65	35.09
7.2	4.64	22.47	26.06
5.5	6.02	18.45	23.36
2.0	11.08	16.61	25.85
.384	18.37	0.00	18.37

possible. In particular, by analysing the results in this figure, it becomes clear that bit-rates higher or closer to the network's maximum bandwidth, despite having a lower degradation during the transcoding process, register a higher degradation after being streamed. On the other hand, moderate adjustments of bit-rate are able to provide a better quality after being streamed, presenting a lower degradation.

Nonetheless, as previously mentioned, the sent videos' bit-rate must be enough such that it does not affect the original quality significantly. For example, for a 2Mb/s bit rate, it is clear in Table 1 that even though the lower bit-rate video is within the limits of the network, with few packet losses, it has an initially higher degradation than for 5.5Mb/s, not being worth the using a lower bit rate.

An interesting result occurred for the 384kb/s bit rate as despite having a higher transcoding degradation, there were no registered packet losses. The overall quality is a result of the x264 codec which is able to maintain an acceptable quality regarding the measured PSNR, suggesting that other measurement mechanisms should be used with this codec.

6 Conclusions

The constant development of lighter and more powerful hardware has motivated the creation of new approaches for improving citizens life quality. The work presented in this paper introduces a portable platform for high quality video transmission in emergency scenarios. By resorting to state-of-the-art hardware and video codecs, 3G networks are used to guarantee the expertise of highly qualified professionals in remote locations in timely matter.

Resulting from the feedback obtained from different emergency teams, the designed equipment and architecture stand-out by being extremely portable and simple to use, while allowing High-definition Video to be seamlessly transmitted between teams in the field and experts in the control centre.

In order to assess the quality provided by the Tutamen platform, comprehensive video tests were performed, evaluating the quality of the transmitted videos with different bit-rates in different network conditions. The perceived video quality is of paramount importance as it will be the main input source at the control centre. The obtained results revealed that the platform is capable

of successfully encoding and transmitting HD Video with a good MOS (higher than 4), adapting to different network conditions.

An important contribution from the developed work concerns the integration of off-the-shelf hardware with the defined software platform, guaranteeing real-time adaptable video transmission that guarantees the best possible video quality.

Future work involves the extension of this platform to other services such as security and education. Moreover, in order to further improve the adaptable characteristics of the presented solution, an additional feature would be the reduction of the video's resolution when FullHD is not mandatory or whenever network conditions are very poor.

Acknowledgements. This work was performed in project LiveCity which has received research funding from the Community's Seventh Framework programme. This paper reflects only the authors' views and the Community is not liable for any use that may be made of the information contained therein. The contributions of colleagues from LiveCity consortium are hereby acknowledged.

References

1. Panayides, A., Pattichis, M.S., Pattichis, C.S., Schizas, C.N., Spanias, A., Kyriacou, E.: An overview of recent end-to-end wireless medical video telemedicine systems using 3g. In: 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), October 31 - September 4, pp. 1045–1048 (2010)
2. El Khaddar, M., Harroud, H., Boulmalf, M., Elkoutbi, M., Habbani, A.: Emerging wireless technologies in e-health trends, challenges, and framework design issues. In: 2012 International Conference on Multimedia Computing and Systems (ICMCS), pp. 440–445 (May 2012)
3. Sugita, N., Yoshizawa, M., Kawata, H., Yambe, T., Konno, S., Saijo, O., Abe, M., Homma, N., Nitta, S.: Telemedicine system necessary in disaster areas. In: 2012 Proceedings of SICE Annual Conference (SICE), pp. 1661–1664 (August 2012)
4. Kumar, S., Krupinski, E.A.: Teleradiology. Springer (2008)
5. Kruchten, P.: The Rational Unified Process: An Introduction, 3rd edn. Addison-Wesley Longman Publishing Co., Inc., Boston (2003)
6. Wang, Y.K., Even, R., Kristensen, T., Jesup, R.: RTP Payload Format for H.264 Video. RFC 6184 (Proposed Standard) (May 2011)
7. Bernardo, V., Sousa, B., Curado, M.: Voip over wimax: Quality of experience evaluation. In: IEEE Symposium on Computers and Communications, ISCC 2009, pp. 42–47 (July 2009)
8. Klaue, J., Rathke, B., Wolisz, A.: EvalVid – A framework for video transmission and quality evaluation. In: Kemper, P., Sanders, W.H. (eds.) TOOLS 2003. LNCS, vol. 2794, pp. 255–272. Springer, Heidelberg (2003)