Project acronym: CONCERTO

Project full title: Content and cOntext aware delivery for iNteraCtive multimEdia healthcaRe applications

Grant Agreement no.: 288502

Deliverable 6.5

Final System Validation

Contractual date of delivery to EC	T0+39
Actual date of delivery to EC	18/03/2015
Version number	1.0

Lead Beneficiary	KU
Participants	KU, UNIPG, BME, VTT, TCS, SIEMENS, CNIT

Estimated person months:	18.4
Dissemination Level:	PU
Nature:	R
Total number of pages	57

Keywords list: Demonstration, Quality evaluation, ultrasound, emergency, tele-consultation.

Executive Summary

This deliverable presents the results of the final CONCERTO system validation obtained via the implementation of the CONCERTO wireless multimedia platform that has been developed during the project in order to demonstrate the real-time implementation of the CONCERTO solution for different scenarios and, more in general, the last activities realized to validate the CONCERTO system.

The results of system validations realized via simulations are provided in Deliverable 6.3, "CONCERTO Simulator: Final Architecture and Numerical Result" [4]. The demonstration platform has been introduced in CONCERTO Deliverable 6.4, "Demonstrator description and validation plan" [5], where a description of the platform was provided, together with an overview of the hardware and software modules integrated and their interfaces. The demonstration platform, as well as the system level simulator, has been built according to the CONCERTO cross-layer system architecture elaborated in the scope of WP2 and according to the specifications provided in D6.2 "Specification of the demonstrator" [3].

An overview of the results obtained via the CONCERTO simulator and the results obtained with the CONCERTO demonstration platform in different scenarios are presented in this deliverable. In particular, three demonstration scenarios, namely Emergency, Ubiquitous Tele-Consultation, and 3D medical data storage and encoding, have been considered. These scenarios, described in Section 3 of deliverable D6.4 [5], are based on the use cases described in Deliverable D2.1 [1], each covering more than one use case. The results of the field tests and of the validation phases carried out in collaboration with the medical doctors of the hospital of Perugia are presented in detail in terms of different performance indicators associated to the quality of service (throughput, packet loss ratio, delay) and to the quality of experience (e.g., medical image and video quality), including subjective tests performed with the medical specialists.

The results show that the CONCERTO solution enables the provision of an adequate quality for the end users of the considered systems (medical doctors in particular), outperforming on several metrics the selected benchmark approaches (i.e., without CONCERTO optimizations).

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1 Introduction

The main purpose of the CONCERTO project was to develop novel methodologies, algorithms, and architectures to foster the use of multimedia healthcare applications. The innovative strategies developed within the project, based on the concepts of content- and context-awareness, aimed at providing a better quality of experience (QoE) to both medical doctors and citizens and a higher efficiency in the processing and transmission of both medical and classic multimedia contents. The developed solutions have been individually validated through analysis and dedicated simulations and the relevant results have been reported in the previous CONCERTO technical deliverables (WP3, WP4 and WP5 deliverables).

The combined performance of different blocks and the full system have then been validated for different use cases in two different modalities: via the CONCERTO system level simulator [4] and via the CONCERTO multimedia demonstration platform, through field trials and real time demonstrations.

The adoption of these two instruments enabled to assess the impact of the proposed techniques considering different levels of control over the communication chain. For example, some simulations have been carried out to evaluate the impact of different multi-user resource allocation strategies within the LTE eNodeB. While it has been demonstrated that the effect of different policies is relevant for the final QoE, during the final over-the-air demo session it was not possible to control this aspect of the communication link between the emergency area and the hospital, since, using a commercial network, we had no control on the LTE eNodeB. Hence, the complementary tools adopted in the project enabled the evaluation of the proposed techniques in a more complete way.

While the specifications of the demonstration platform have been provided in Deliverable D6.2 [3] and the description of the platform is provided in Deliverable D6.4 [5], the objective of this deliverable is to present the results obtained via the developed wireless multimedia demonstration platform in field trials of the selected demonstration scenarios (based on the CONCERTO use cases). An overview of the results obtained with the CONCERTO simulator is also presented in this deliverable. Subjective results obtained with the medical doctors in the Hospital of Perugia are also presented.
This deliverable is organized as follows: Section 2 presents the main results obtained with the CONCERTO simulator, providing a detailed performance assessment of the CONCERTO solution with respect to the considered benchmark solution (detailed results are available in deliverable D6.3). Section 3 presents the results of the demonstrator in the three different scenarios considered, in terms of quality of service metrics and objective image / video quality (e.g., throughput, packet loss ratio, delay, PSNR).Finally, Section 4 presents the feedback obtained from the end users, i.e. the subjective (diagnostic) quality results obtained with the medical specialists in the hospital. The medical doctors have evaluated the CONCERTO results for different video compression and transmission methodologies. Different types of subjective tests have been performed, and the relevant results are presented in detail.

2 Simulator results

This section gives a brief overview of the final simulator architecture, the selected algorithms used, a summary of the use cases simulated and results of the key performance indicators (KPI) obtained from the simulator.

The architecture of the CONCERTO simulator is reported in Figure 1 and described in details in the deliverable D6.3 [4].

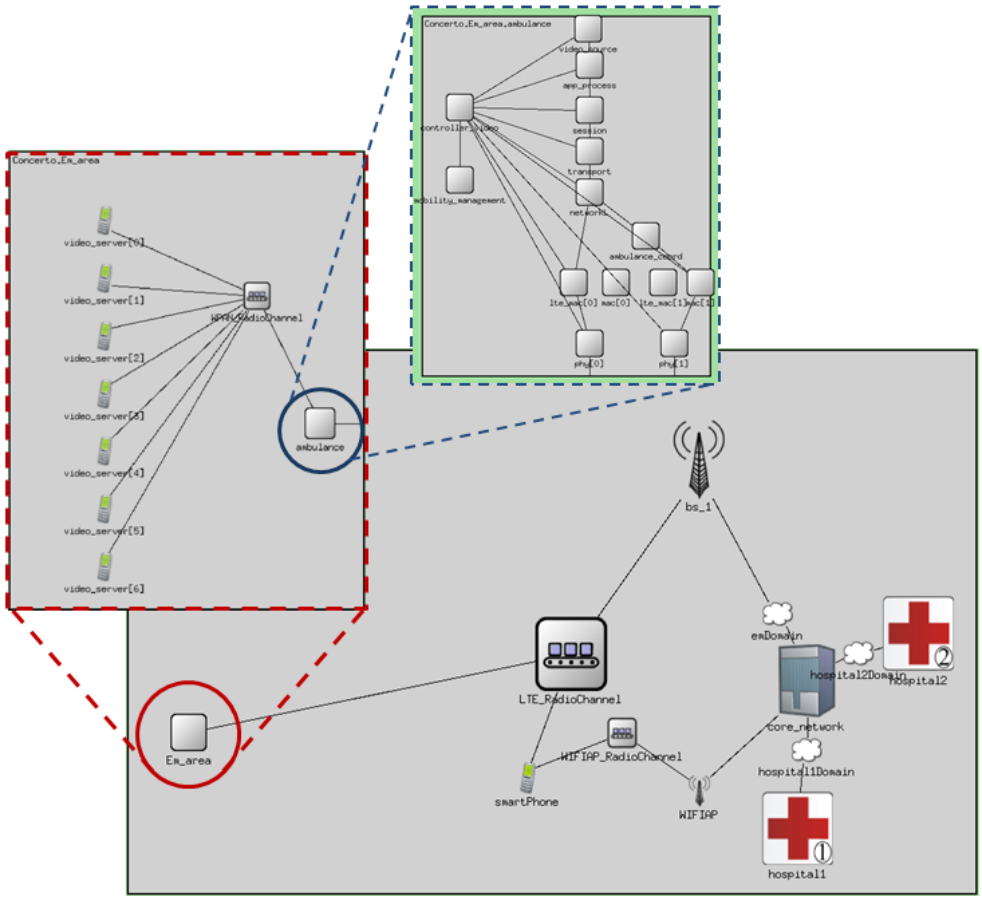

Figure 1 – Architecture of the CONCERTO simulator.

As reported in D6.3, for each module of the CONCERTO simulator different optimization techniques have been implemented, allowing a performance comparison between the solutions proposed within the project and traditional schemes. The main characteristics of the options implemented in the CONCERTO simulator are summarized in the following.

- ***Coordination Centre*** – When source ranking is activated (**SR**), the CC transmits a prioritized list of video sources to the ambulance adaptation unit, in charge of enabling and adapting all the cameras in the emergency area. The ranking techniques developed in CONCERTO are described in D4.2 and D4.3. Control signalling between CC and ambulance is based on the DDE. In addition, the CC includes an HTTP streaming server that permits to compare an end-to-end RTP communication to an MPEG-DASH solution implemented within the hospital. In the latter case the RTP stream received from the ambulance/emergency area is transcoded to MPEG-DASH by a dedicated processing module at the CC.

- ***LTE eNodeB*** – Two scheduling and radio resource allocation techniques have been implemented in the simulator: a traditional proportional fair (**PF**) approach and the **GBR-ALRE** solution described in D4.3. In the latter case, the ambulance is treated by the 4G access network as a privileged user, and a guaranteed bitrate is provided by the system based on a low-complexity processing scheme fully compatible with the LTE uplink standard.

- ***Ambulance*** - The processing unit on-board the ambulance is responsible to collect and aggregate the multimedia streams from all the cameras available in the emergency area and to transmit them through the LTE link, toward the CC/hospital. To perform its processing, the ambulance exploits the ranking information received from the CC and the achievable throughput across the LTE link. Two techniques have been compared through simulation to perform rate adaptation, namely equal rate (**ER**) and quality fair (**QF**). For the medical stream higher quality is selected with respect to the ambient videos, in order to allow an effective tele-diagnosis support.

2.1 KPIs addressed through simulations

The main KPIs addressed by the simulation campaign can be classified into two groups: end-to-end quality metrics and transmission quality metrics. The end-to-end video quality is evaluated through multiple objective metrics, namely PSNR and SSIM. In addition, the end-to-end delivery delay has been considered as a factor contributing to QoE. Low latency enables real-time or quasi real-time services, characterized by the possibility for the remote specialist to provide an interactive support to the on-site staff. Radio link quality has been evaluated by means of traditional QoS metrics, such as the achieved throughput and link delay.

2.2 Simulated scenarios and results

Three scenarios have been selected for validation through numerical results, as they are representative of most use cases identified in the project.

The Ambulance and Emergency Area scenario mainly focus on multimedia communication between an ambulance on the move and a remote hospital, where a doctor can utilise the received multimedia contents from the ambulance for different purposes such as monitoring, coordination and/or consultation. The main video and medical data streams are originated by pieces of equipment on the ambulance and are received by a client at the hospital. Since the ambulance is typically moving, sometimes also at high speed, a 4G link is supposed to be available to communicate with the hospital. The simulation was focused on transmission of two ambient video streams and one ultrasound sequence acquired inside the ambulance. In this way we simulated the medical assistance performed after loading the patient on the vehicle with the support of advanced tele-diagnosis services. Besides a detailed description of the performed simulations in deliverable D6.3, a comparison of full-RTP and HTTP solutions has been provided. Moreover, the benefits offered by the adoption of a PL-FEC technique based on nonbinary LDPC codes have been presented. As an example of results, the final end-to-end video quality obtained in the first scenario is reported in Figure 2. More results can be found in D6.3.

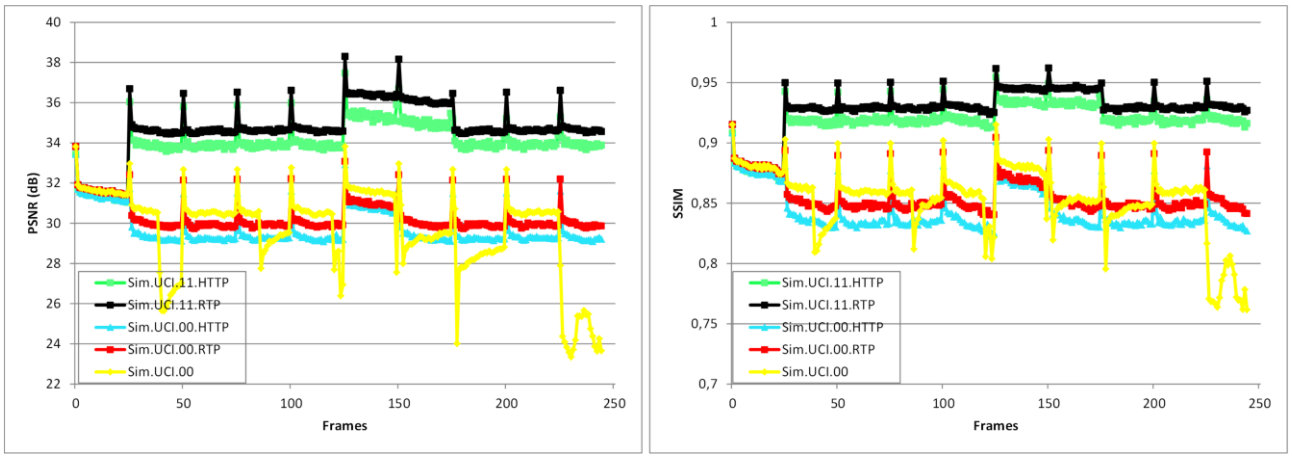

Figure 2 – Ambulance and emergency area scenario: PSNR and SSIM quality comparison obtained with the CONCERTO simulator.

As reported in D6.3, the second scenario addressed by simulation is the case of Emergency Area with Multiple Casualties. The simulations have been carried out considering two operative phases: In the first phase, the first aid staff assisted two injured people on the ground, located some distance apart. Four ambient cameras were deployed in the area and available to acquire and transmit information. Several simulations have been done considering different sets of transmitted videos: first we assumed that the entire set of video streams has to be sent to the CC; then we evaluated the

case in which just two videos sources out of the available four were selected through fixed or dynamic mechanisms and transmitted to the remote user. The second simulation phase consisted in the transmission of two ambient video streams and one ultrasound sequence acquired inside the ambulance. In this way we simulated the medical assistance performed after loading the patient on the vehicle with the support of advanced tele-diagnosis services.

Eight different combinations of optimization techniques were compared and detailed results have been provided in the deliverable D6.3. As an example, in Figure 3 we report the Phase 2 performance obtained through simulations, comparing a traditional communication link (benchmark) with the CONCERTO optimized solution. In particular, we represent the end-to-end video quality received at the hospital, in terms of SSIM. We note how the CONCERTO solutions allow improving the quality significantly of the medical video sequence with respect to the benchmark, guaranteeing satisfying diagnosis accuracy while still providing a reasonable high quality to the ambient videos. Moreover, the resulting end-to-end packet delay is significantly reduced, as the 99th percentile of the packet delay moves from 2.5 s (benchmark case) to 230 ms (CONCERTO solution). A detailed description of the performed simulations as well as the additional numerical results is reported in the deliverable D6.3.

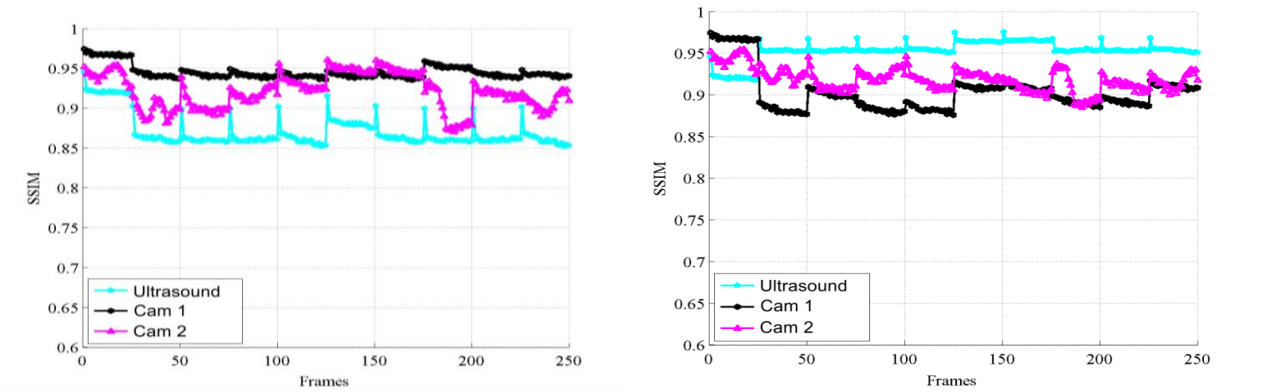

Figure 3 – Received video quality in the benchmark case (left) and with the CONCERTO solution (right).

In [7] we have derived SSIM vs. differential mean opinion score (DMOS) mapping for ultrasound video sequences. DMOS represents the quality difference perceived by the medical doctors between the impaired and the original image. Assuming the mapping in [7], an SSIM value of 0.87 (average value for the benchmark solution) corresponds to a DMOS value of 72 out of 100, i.e. non satisfactory perceived quality, while a value of SSIM 0.96 (average value for the CONCERTO solution) corresponds to a DMOS value of 40 out of 100, i.e. satisfactory perceived quality. Hence, the CONCERTO solution enables to provide a satisfactory quality to the end-users (medical doctors), while the benchmark solution does not.

In the third simulation scenario implemented, the ubiquitous tele-consultation use case (which corresponds to use case 4 in deliverable D2.1) was addressed. In this use case the uplink has been evaluated between a smartphone and the Coordination Centre exploiting the contemporary presence of multiple radio access networks and implementing the CONCERTO seamless handover solution based on the network discovery service. The transmitting smartphone has multiple wireless interfaces contemporary active and, based on DDE and NIS signalling, the network discovery service enables the automatic and seamless switch of the current wireless connection. By exploiting the less congested network among those available, it is possible to reduce the average end-to-end delivery latency, thus improving the communication interactivity and the QoE. In particular, in our simulations, we passed from an average delay of 2.79s, obtained in an overloaded LTE cell, to 0.3s, obtained by switching the communication to an available WiFi network.

3 Demonstrator results

3.1 The CONCERTO demonstration platform

The multimedia platform built to demonstrate CONCERTO concepts is fully described in deliverable D6.4 and represented in Figure 4.

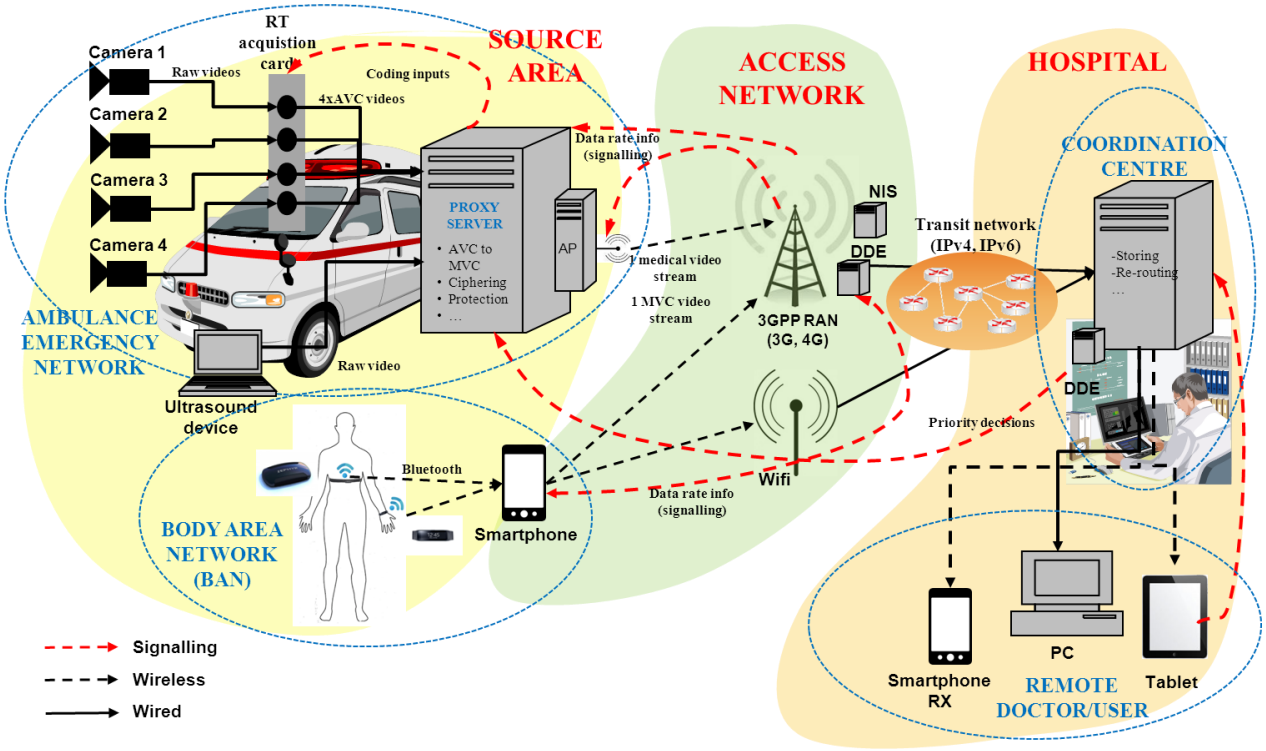

Figure 4: CONCERTO demonstration platform

Three main areas are identified in the demonstration platform:

- The **Source Area**, which includes all the possible sources of environmental and medical data that can be used by remote doctors to perform real-time or offline diagnosis;
- The **Access Network**, which is composed by the different Radio Access Networks (RAN) that can be exploited to route the traffic between the Emergency Area and the Hospital;
- The **Hospital**, which represents the place where the traffic is received, stocked and analysed and is composed by a Coordination Centre (CC) and by the wired and wireless network connecting the Coordination Centre with remote doctors.

Following the principle of the CONCERTO solution, the demonstrator has been conceived in a modular way. The different components can be combined or disabled/removed in order to implement different scenarios and to satisfy different requirements.

3.2 KPIs addressed through demonstration

The CONCERTO test bed has been mainly assessed via end-to-end quality metrics. The end-to-end video quality is evaluated through multiple objective metrics, namely PSNR and SSIM, and subjective quality metrics (mean opinion scores on the whole sequence and over time).
In addition, the end-to-end delivery delay has been considered as a factor contributing to QoE and bit-rate and packet losses have been measured where they could provide added value to end-to-end quality measurements.

3.3 Selected demonstration scenarios

The field trials realized along with the project allowed to test all the different components and configurations of the demonstrator in order to validate the integrated functionalities under various conditions. In particular three demonstration scenarios, covering several use cases and highlighting different key functionalities, had been selected for the final field trials realized by the project.
These scenarios are described in detail in deliverable D6.4 [5] and are summarized in the following, together with the associated use cases. The demonstration scenarios have been selected in order to show key capabilities and functionalities at different levels and to validate the effectiveness in healthcare use cases of the CONCERTO system.

3.3.1 Emergency scenario

This scenario allowed reproducing the full chain from medical acquisition in an emergency area to real-time reception and a navigation of a doctor at the hospital on a mobile device. It has been initially tested both with a customized access network (i.e., proprietary 3G femtocell) and finally realized using a commercial network (i.e., 4G network operated by an Italian operator).

Requirements derived from two use cases are mapped into this demonstration scenario, namely the "ambulance and emergency area" and the "emergency area with multiple casualties" use cases. The demonstration includes two main areas: the emergency area, where an ambulance is deployed and the hospital area, where the Coordination Centre and mobile medical doctors are situated.

Different types of image/video data have to be transmitted from the ambulance to the hospital in order to allow an accurate diagnosis to help the personnel on the field and to speed up the acceptance process once back at the hospital. Hence, multiple data flows need to be transmitted from the ambulance to the hospital; this includes both environment videos (from cameras deployed inside and outside the ambulance) and medical videos (i.e., ultrasound videos).

This scenario is considered for both the final evaluation via simulation and for the final demonstration.
In the final demonstrator configuration, this scenario considers the transmission of image and video data from the emergency area in which an ambulance is deployed to a remote hospital using a 4G commercial network and transmissions inside the hospital from the Coordination Centre to doctors in mobility through Wi-Fi.

3.3.2 Ubiquitous tele-consultation scenario

This scenario showed the interest of the exploitation of heterogeneous wireless access networks and of the network solutions developed in CONCERTO.
This mainly refers to CONCERTO use case No. 4 (Ubiquitous tele-consultation), where preliminary consultation with nursing-home staff, patients and other doctors may be carried out using hand-held devices, which provide access to multiple heterogeneous wireless networks, regardless whether he/she is in an ambulance, office, airport, traffic jam, or other place. However, optimization solutions applied in this demonstration are also applicable in other CONCERTO use cases like 1, 2, and 7.

3.3.3 3D medical data storage and encoding

This scenario gives an example of the functionalities available at the Coordination Centre. The effective medical data encoding system could be applied in all use cases, where large data medical sets are stored or transmitted (Use cases: 1, 2, 3, 4, 6, 7). In Use case 5 ("Surgical assistance") near-real-time communication is essential, so solutions based on high-complexity algorithms cannot be applied. This demonstration has the main focus on showing some of the capabilities developed at the Coordination Centre.

In the following sections the results of the final field trials realized for these three demonstration scenarios are described.

3.4 Emergency demonstration scenario

This demonstration scenario considers the transmission of image and video data from an emergency area in which an ambulance is deployed to a remote hospital. The final measurements have been realised using for the transmission between the emergency area and the hospital a 4G/LTE commercial network. However, previous tests were also realized using a customized 3G femtocell. These previous tests validated the positive impact of the integration of the DDE also at the RAN which allows to have better feedbacks and to adapt the encoding and the transmission at the ambulance side more efficiently. For the final tests it was decided to use a commercial network, to show that CONCERTO solution can be easily implemented using existing infrastructures and then requiring a limited investment.

In the scenario, different types of image/video data have to be transmitted from the ambulance to the hospital in order to allow an accurate diagnosis to help the personnel on the field and to speed up the acceptance process once back at the hospital. Hence, multiple data flows need to be transmitted from the ambulance to the hospital; this includes both environment videos (from cameras deployed inside and outside the ambulance) and medical videos (i.e., ultrasound videos). For the demonstration, the transmission of one ultrasound image stream and two ambient video streams have been realized. Once received at the hospital by the Coordination Centre, the videos are stored and forwarded in real-time to a remote doctor on a mobile device using the hospital Wi-Fi network. Both RTP/RTSP and MPEG-DASH (HTTP) transmissions have been implemented and tested.

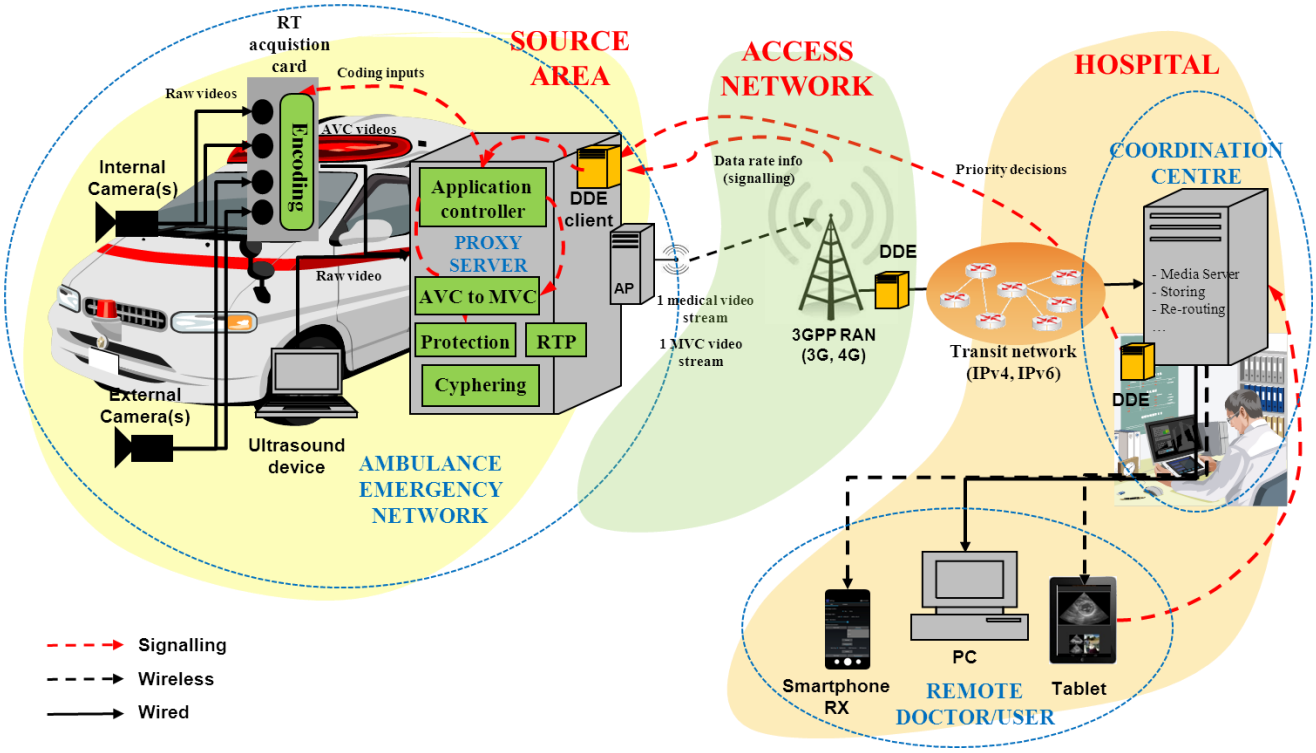

Figure 5: Demo setup of emergency scenario.

3.4.1 Real-time transmission of ultrasound image and ambient video streams from an ambulance to the hospital

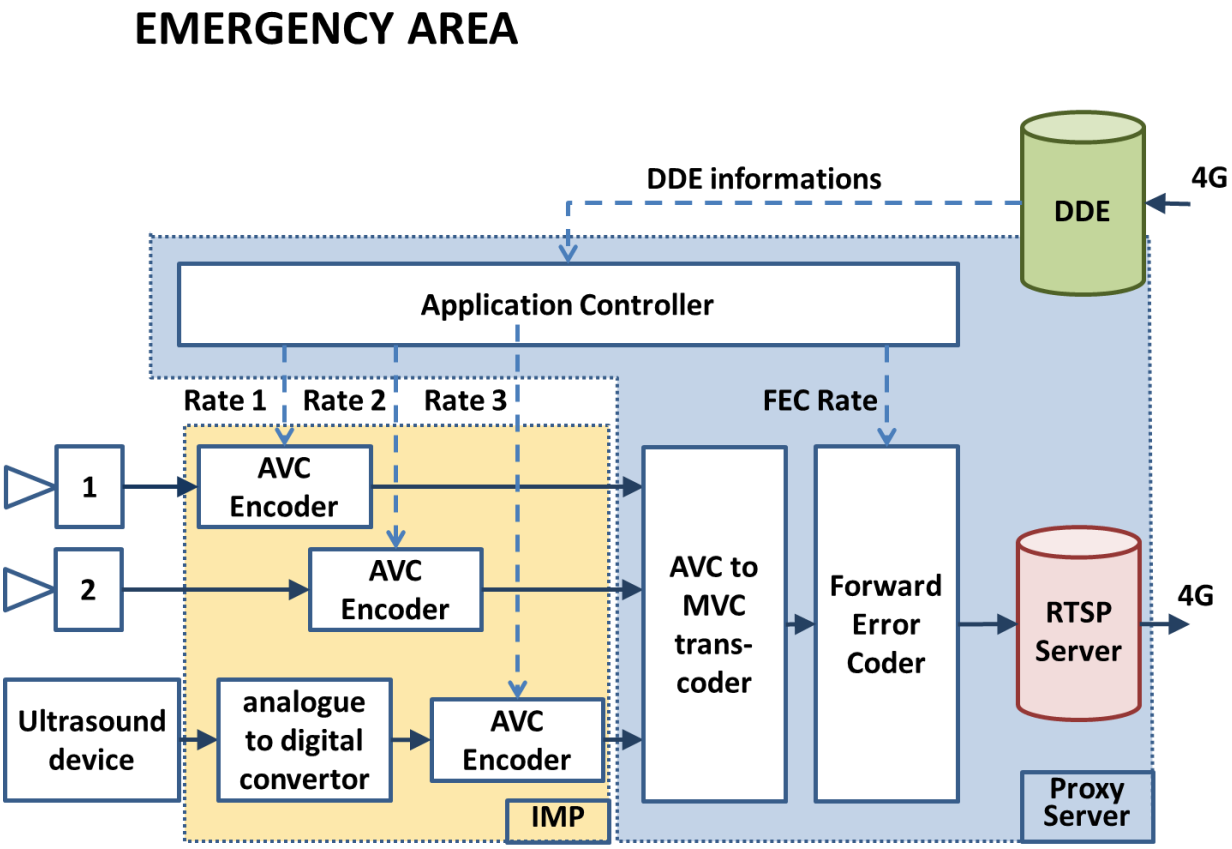

Figure 6 – Configuration of the demonstrator representing the emergency area.

In Figure 6 represents the configuration of the demonstrator used for the final validation trials for the ambulance emergency area. Two video cameras (analogue PAL format) were deployed both inside the ambulance to capture ambient videos of the patient. The cameras were wired connected to a compression server, called IMP box, which can encode in real-time four different video sources. The IMP box transfers one encoded video source per camera in AVC to a laptop PC, namely as the Proxy Server. A portable ultrasound device was also available at the ambulance and used to acquire medical ultrasound videos directly in the emergency area. This device provides a resolution sufficient to doctors to perform preliminary diagnosis and has the advantage to have reduced size (around the dimensions of a standard laptop). The medical videos acquired were converted starting from the analogue real-time video flow to digital (VGA to PAL format), and then converted again by VITEC Optibase MGW PICO to provide the digital videos to the Proxy Server through an Ethernet cable. The specifications of the different devices used are provided in deliverable D6.4.

The Proxy Server, deployed at the ambulance, was in charge to centralize all the different flows acquired in the emergency area, and to transform them into a unique MVC stream. At the Proxy Server the Application Controller is implemented to provide quality control and adaptation and to add adequate protection before transmission by using Forward Error Correction (FEC). The used FEC, called PL-FEC, has been described in D4.2 and D4.3. The Proxy Server, thanks to the application controller, acts taking into account cross-layer signalling information on QoS and priority requirements received through the DDE. Moreover, another module was also present to cypher the different streams for the security aspect. The Proxy Server was then connected to an access point who managed the connection to the 4G commercial RAN.

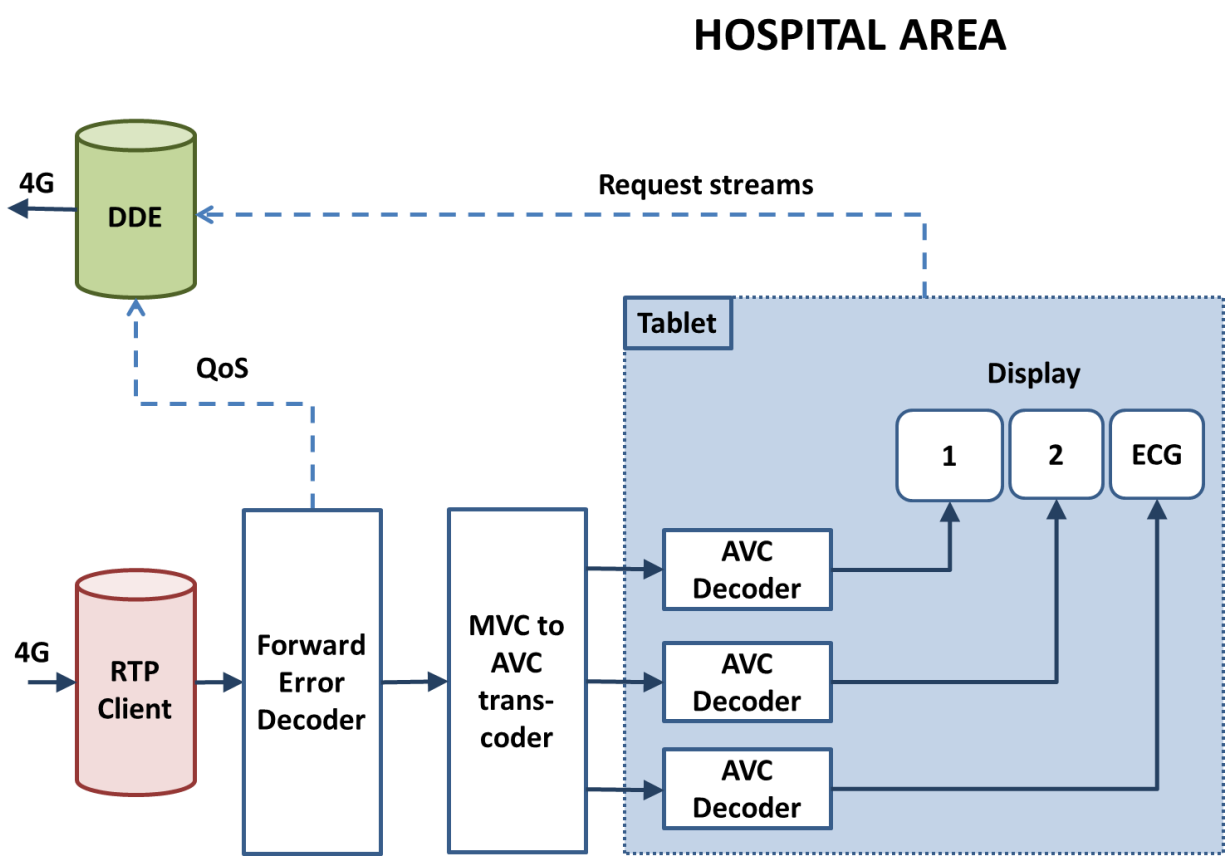

Figure 7 - Configuration of the demonstrator representing the Hospital area.

At the hospital side (Figure 7), input media streams are received from the Ambulance Emergency area as one MVC RTSP/RTP stream. Then the streams are decoded by the PL-FEC decoder to recover erasures and provide QoS information concerning the instantaneous quality of the 4G link. This information is reported through DDE to the Proxy Server at the ambulance where the Application Controller uses it to select the best rate and level of protection to apply. After passing through the PL-FEC decoder, the H.264/MVC stream is converted in the original H.264/AVC streams which can be either ambient video feeds or ultrasound medical video streams. Finally, the streams available are transmitted through the hospital Wi-Fi as an RTP/RTSP stream to a doctor's tablet and decoded. The doctor can select which videos to receive in full quality: its preferences are transmitted through the DDE to the Proxy Server to select the videos to transmit.

3.4.1.1 Validation results

Since we used a commercial 4G network and it was then impossible to provide two identical 4G links during our experiments, the evaluation of our solution was made in two steps. The first step was to implement the benchmark scheme i.e. to test the nominal way of video transmission without CONCERTO's optimizations. Then we have made second tests with all the optimizations provided by CONCERTO (Application controller, FEC, QoS signalling through the DDE…). Moreover, since it was not possible to record raw video data from ambient cameras and ultrasound devices, the evaluation was done on the loss of quality from the original compressed video stream and the received one. So we used delta PSNR (dPSNR) which is a PSNR computed by subtracting the original compressed video stream and the received one. Separate tests to evaluate the performances of the encoding and compression strategies have been realized in the past and reported in previous technical deliverables.

The following results show the different behaviour of our solution and the benchmark scheme (i.e., transmission without CONCERTO optimizations) in terms of dPSNR and SSIM estimated with the compressed video stream and the received stream. Note that if a reconstructed frame at receiver side (hospital area) is exactly the same as the reconstructed frame at the emission side (emergency area), the dPSNR should be equal to infinity (and the SSIM is equal to 1), so we substituted this value with an arbitrary value of 50 to make it clearer.

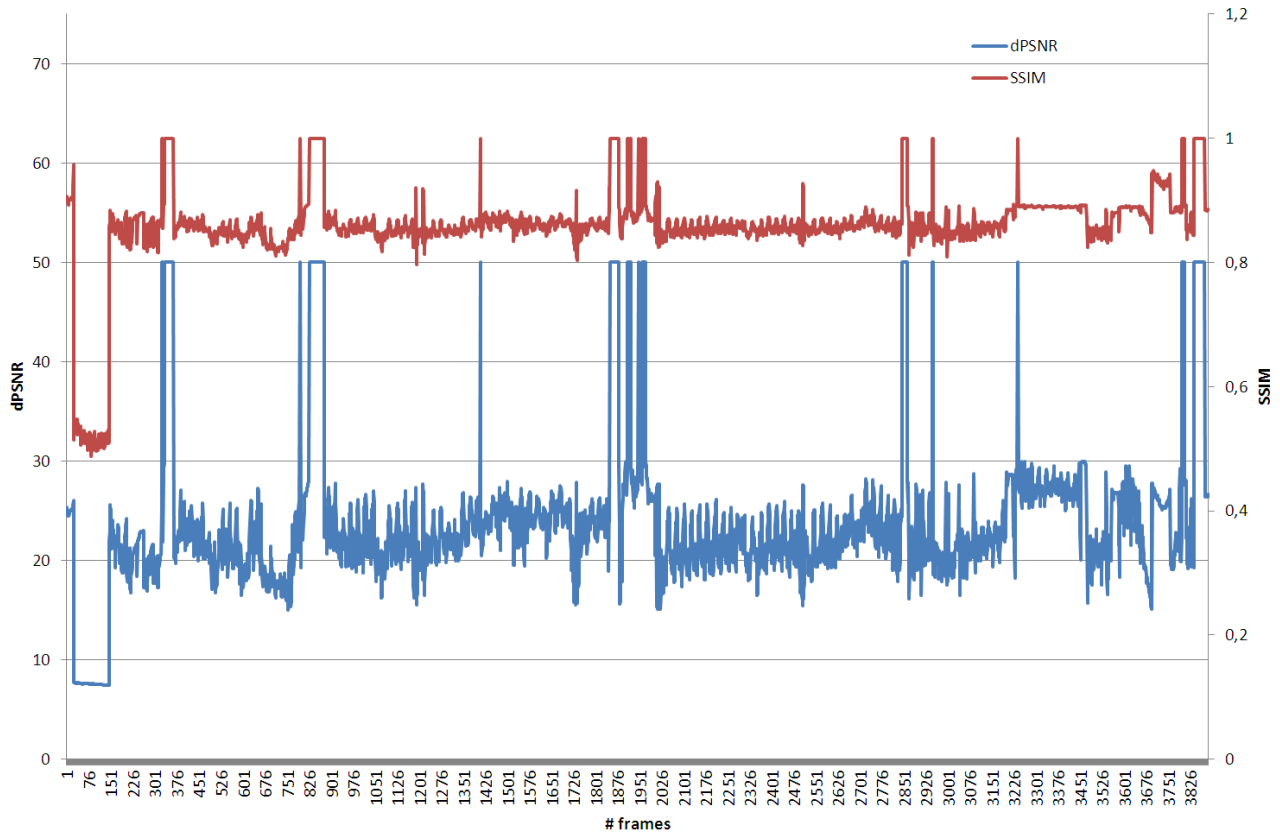

Figure 8 - Evolution of dPSNR and SSIM for the medical video stream (ultrasound) without CONCERTO's optimization.

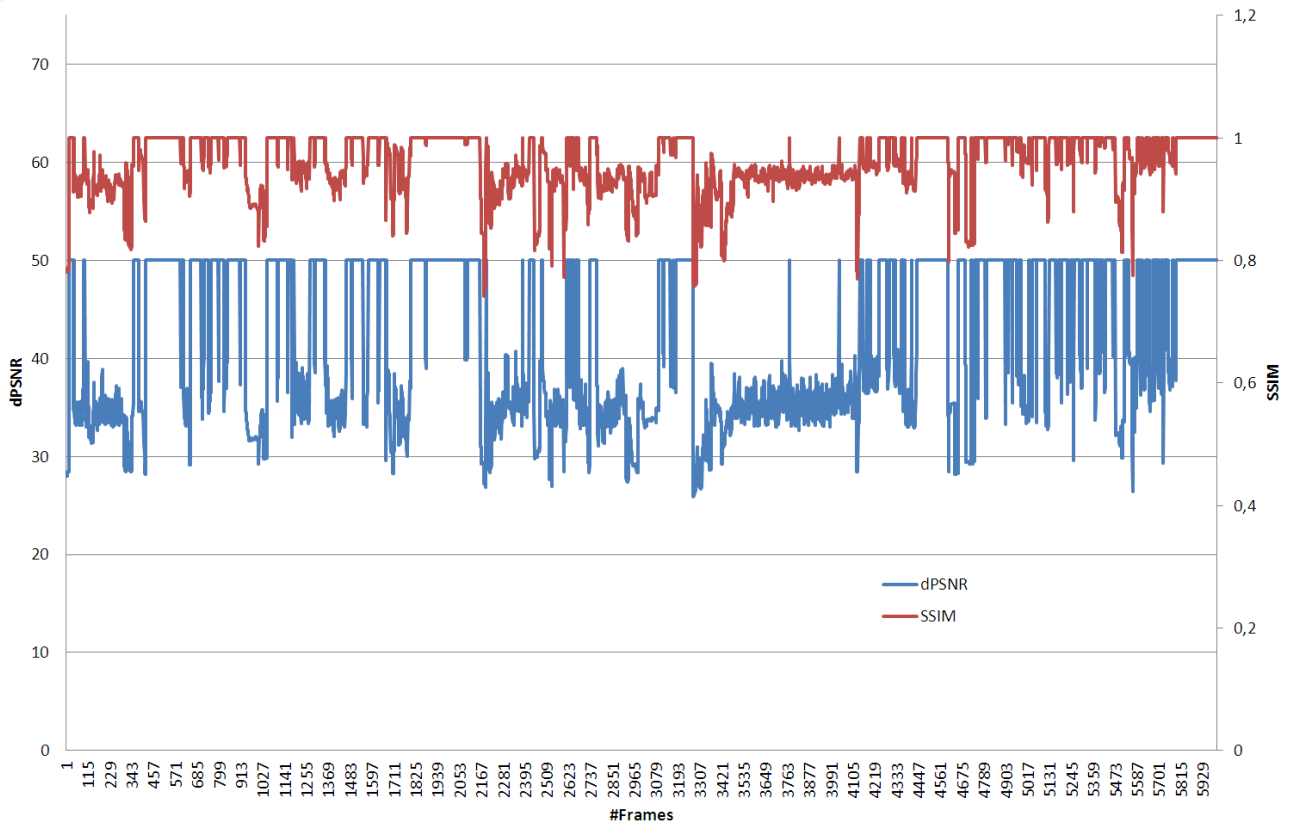

Figure 9 - Evolution of dPSNR and SSIM for the medical video stream (ultrasound) with CONCERTO's optimization.

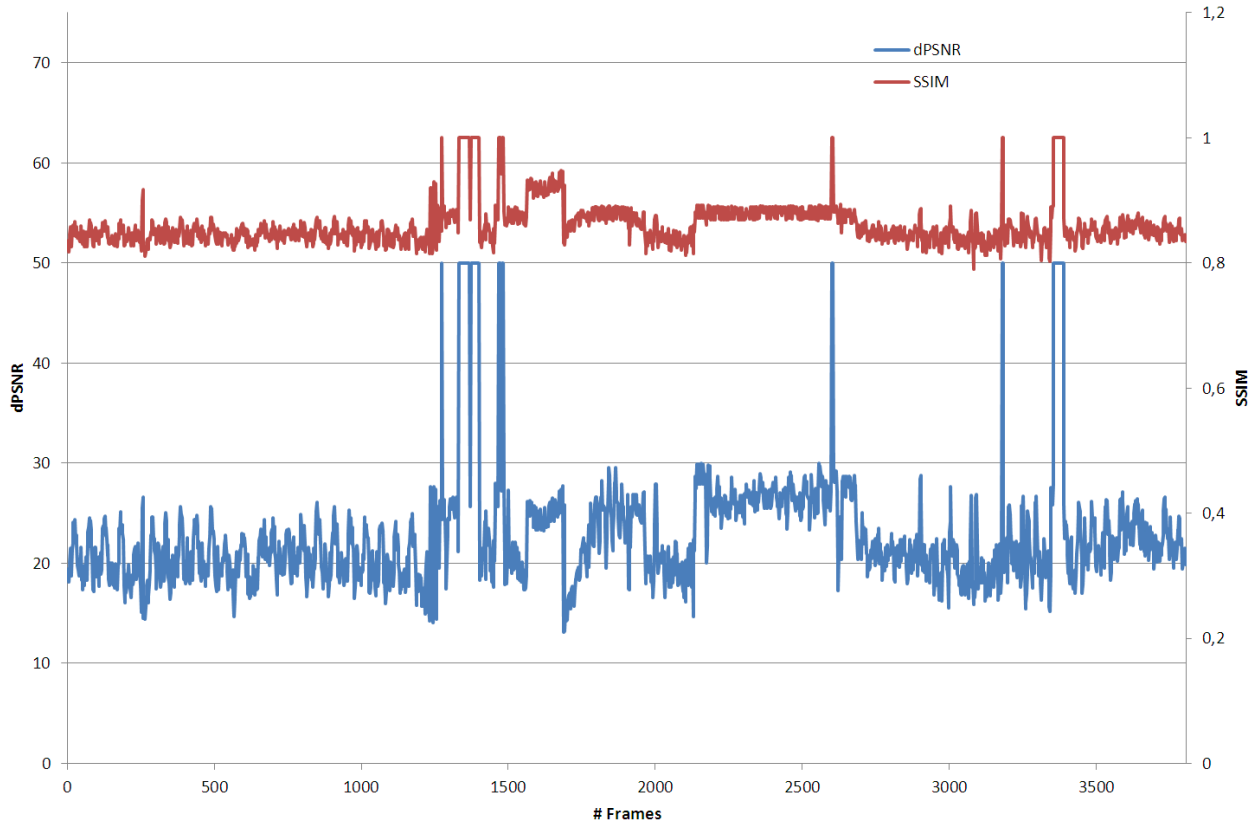

Figure 10 - Evolution of dPSNR and SSIM for the1st ambient video camera without CONCERTO's optimization

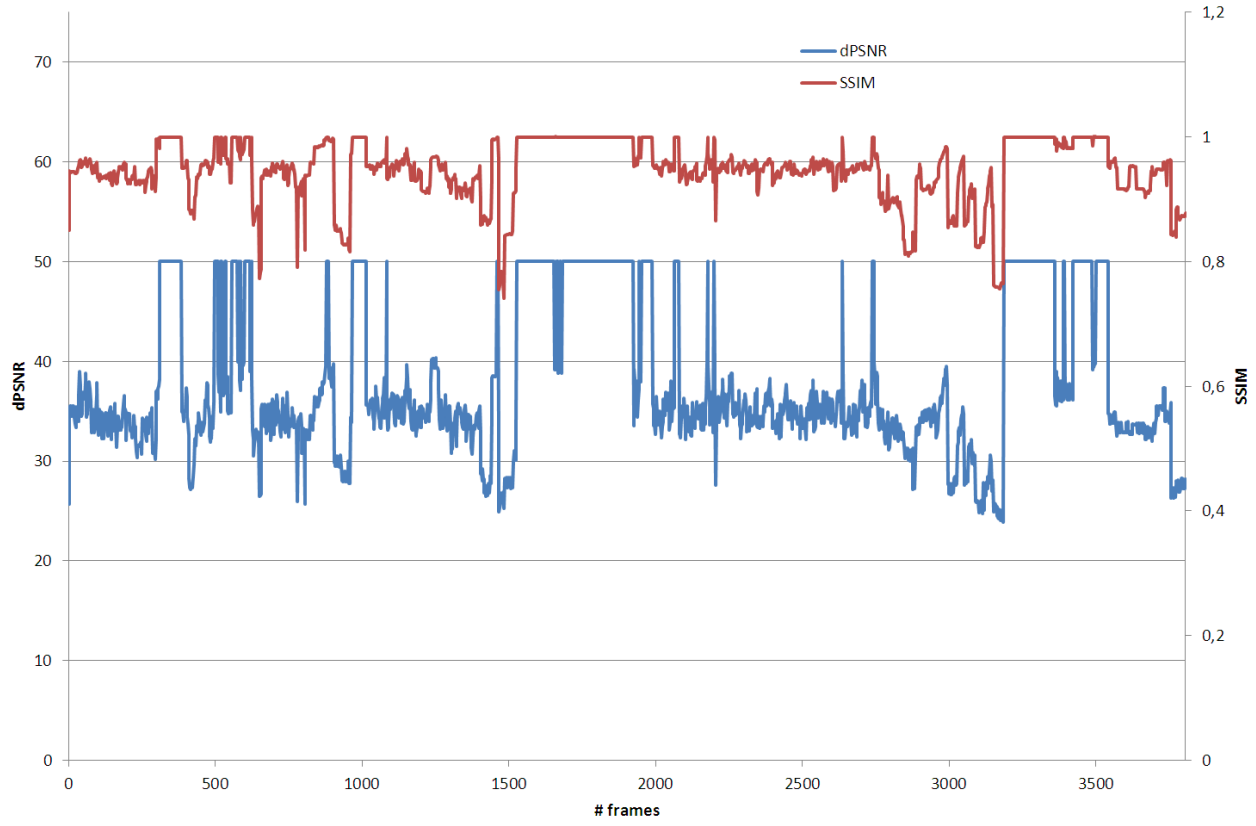

Figure 11 - Evolution of dPSNR and SSIM for the 1st ambient video camera with CONCERTO's optimization

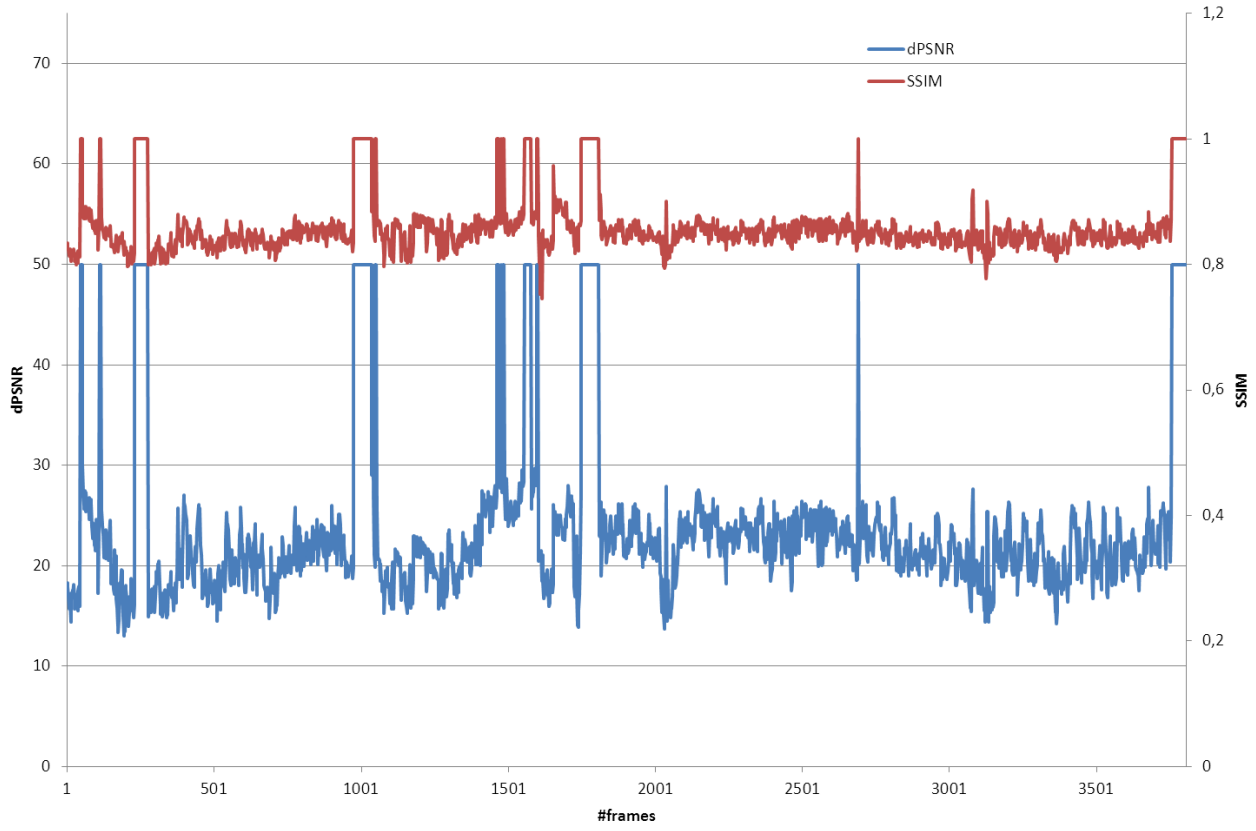

Figure 12 - Evolution of dPSNR and SSIM for the 2nd ambient video camera without CONCERTO's optimization

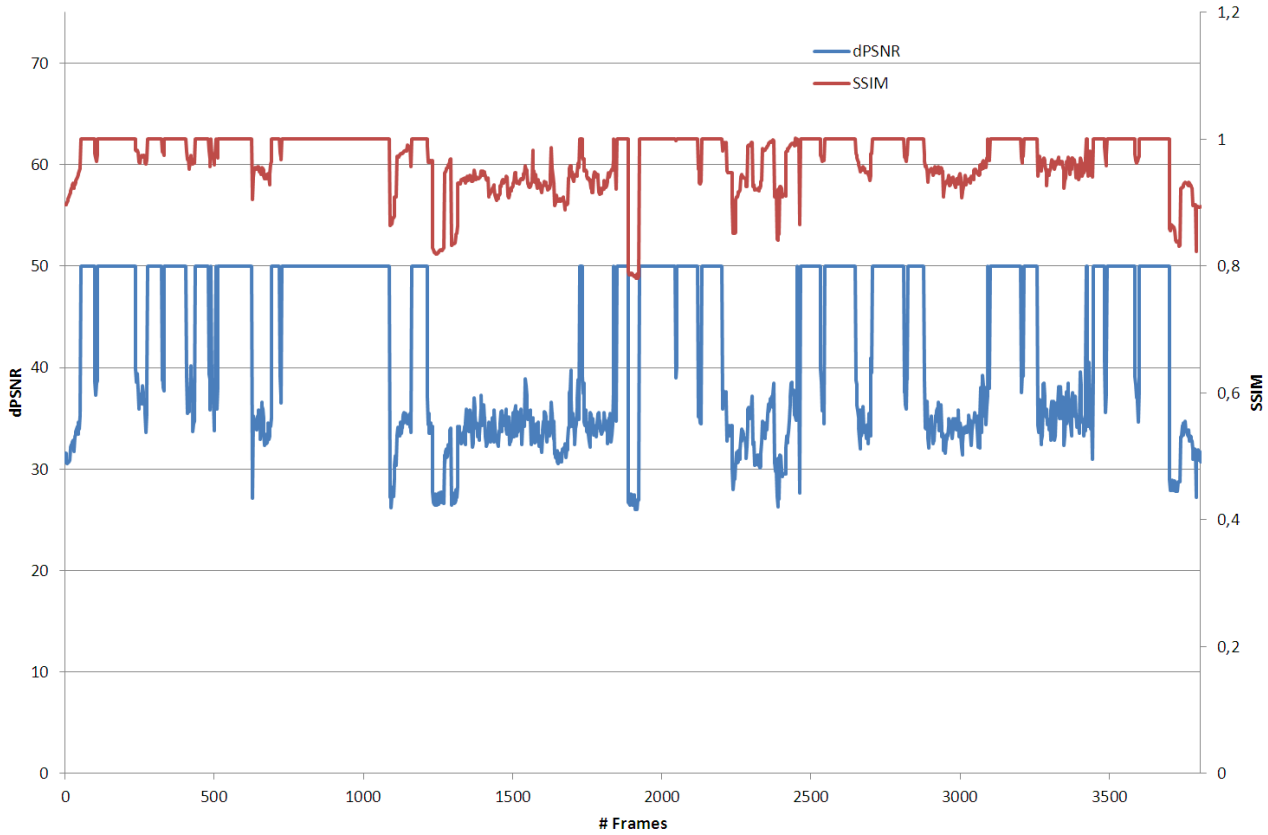

Figure 13 - Evolution of dPSNR and SSIM for the 2nd ambient video camera with CONCERTO's optimization

Figure 8, Figure 10, and Figure 12 show the evaluation of dPSNR and SSIM respectively for the received videos for the medical video source and the first and second ambient video cameras in the benchmark configuration. In this configuration the application controller and the optimizations based on received DDE messages on the link quality are disabled. The total source bitrate is equal to 2Mb/s (respectively 1Mb/s for the medical video stream and 0.5 Mb/s for each ambient video camera) transmitted on 4G link.
Figure 9, Figure 11 and Figure 13, instead, show the evaluation of dPSNR and SSIM with the same overall bitrate but with CONCERTO optimizations and adaptations enabled. As we can notice, CONCERTO's solution provides a significant gain on the mean dPSNR (~40dB) and on the mean SSIM (~0.95) when comparing to the benchmark configuration (respectively ~22dB for dPSNR and ~0.86 for SSIM).

3.4.2 Real-time video delivery from Coordination Centre using RTP/RTSP

As shown in Figure 14 (and described in Deliverable D6.4), the streams sent from the emergency area (i.e., the ambulance) are received by an RTP relay in order to be able to duplicate the streams for multiple usage (i.e., live viewing, mixing and storage). The relay also includes an FEC decoder as explained above.

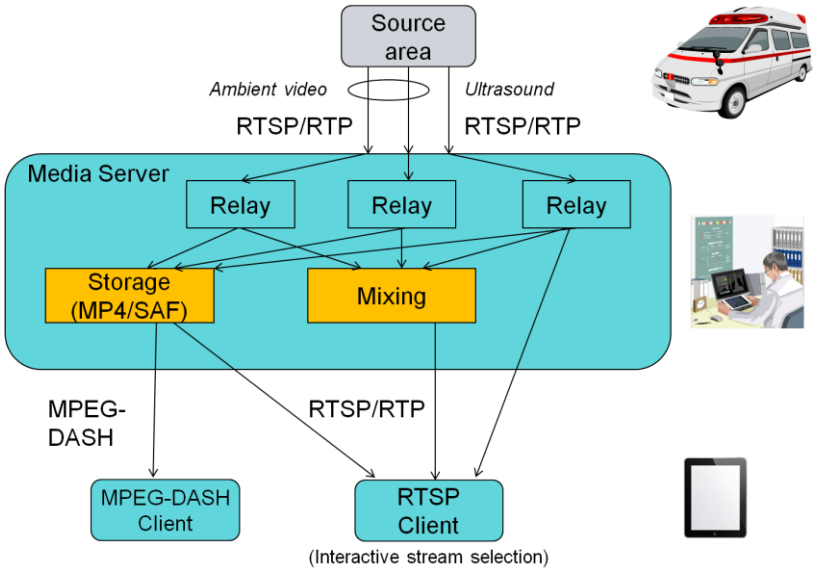

Figure 14 – Single CONCERTO video with one medical ultrasound and two AVC streams.

All live input streams are mixed to achieve a stitched output stream. The stitched output video is presented as an RTSP/RTP stream to the clients in order to gain an overview of the situation. The stitched video also serves as a grid for interactive selection of the input streams, which are then shown in full resolution.

The data rate selected at the ambulance side in order to meet the requirements for transmission over LTE does not seem to be a challenge for the further transmission through the hospital WiFi network. Due to the good transmission quality and sufficient data rate of the WiFi network at hospital side, the quality of the streams received at the RTSP/RTP client (i.e., the tablet) is nearly the same as the streams received at the Coordination Centre. Only very few packet losses on the WiFi link can be observed, which degraded the video quality for short time instance.

The overalls delay for the RTP transmission was assessed to be about 1.5 seconds. The delay at hospital side (from RTP relay to display on the RTSP/RTP client) is estimated to about 200 ms. This low delay is a key advantage of RTP streaming over MPEG-DASH streaming (with a delay of about 4s). This allows for an interaction with the scene (i.e., the interactive selection of the most important video stream).

3.4.3 SAF recording and RTP/RTSP streaming for low-delay navigation

At the Coordination Centre, all input video streams are stored in the MP4/SAF format as also shown in Figure 14. The streams can be retrieved by a dedicated RTSP/RTP client. Again, the quality at the client is virtually the same as the quality of the recorded streams received at the Coordination Centre. In addition, it is possible to interactively select the time instance for replay. The seeking delay is assessed to be between 250 ms and 1 second, which is again significantly lower compared to MPEG-DASH.

3.4.4 Near-real-time video delivery from Coordination Centre using MPEG-DASH

As already mentioned, in this scenario multiple data flows are transmitted from the ambulance to the Coordination Centre using RTP/RTSP streaming via the 4G network. In the realized field trials, the flow includes one medical ultrasound video and two AVC videos from the web cameras located inside the ambulance. For the transmissions of the video flows received at the Coordination Centre to a specialist on the move at the hospital (using Wi-Fi), an alternative to the already mentioned RTP/RTSP approach is the use of HTTP streaming, This approach, which guarantees a near-real-time and reliable transmission, has also been implemented and tested.

On the hospital side (i.e., in the Coordination Centre), all video streams are received, each at a dedicated RTSP/RTP interface, which serves as a relay. On the one hand, the received streams are mixed in order to create a single video stream containing all the three streams (one medical and two AVC). This stream is sent directly to a RTP2DASH server hosting also as a HTTP server. In this server, the stream is transcoded from RTP to MPEG-DASH format suitable for HTTP streaming. Furthermore, this HTTP stream can be viewed with an Android client inside the hospital Wi-Fi capable of MPEG-DASH playback. Figure 15 illustrates the stitched CONCERTO video stream and Figure 16 the overall MPEG-DASH streaming scenario in the CONCERTO demonstration setup.

Figure 15 – Single CONCERTO video with one medical ultrasound and two AVC streams.

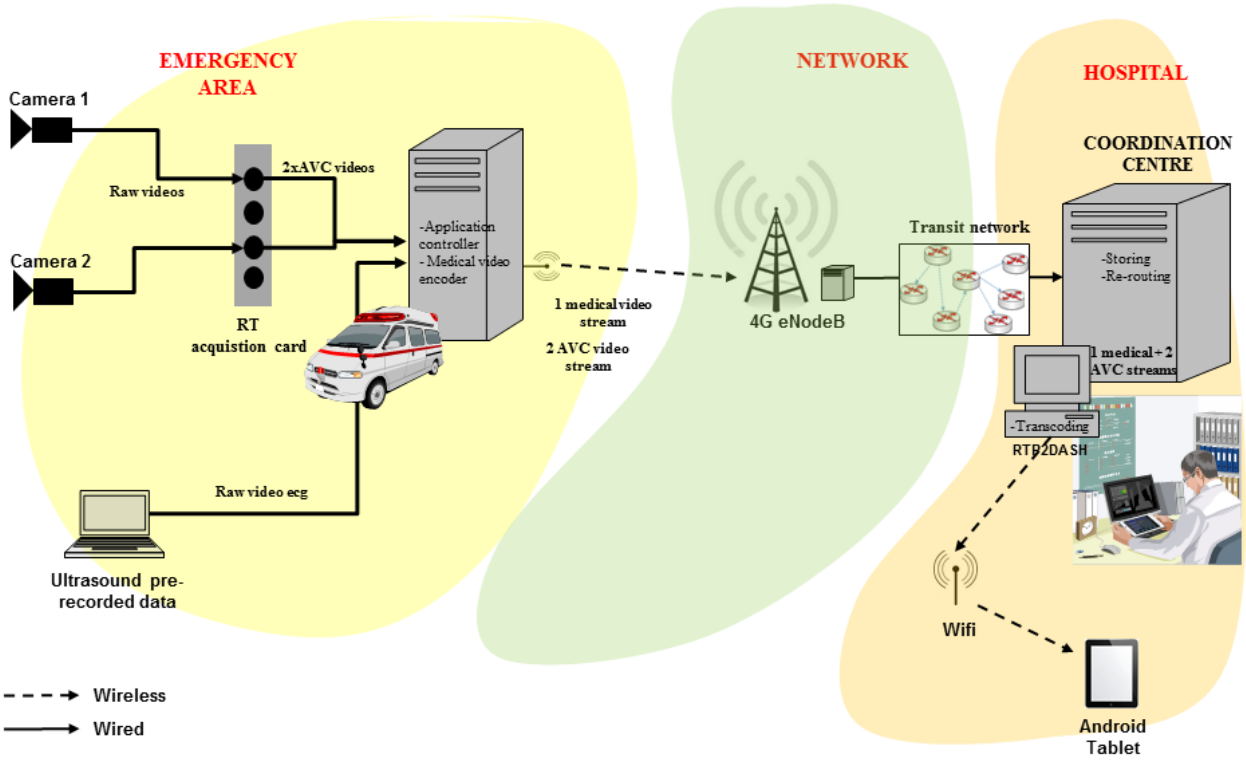

Figure 16 – MPEG-DASH in the CONCERTO demo scenario.

3.4.4.1 Validation results

This section presents the numerical results of our activities with MPEG-DASH in the demonstrator. The RTP/RTSP input stream from the Coordination Centre is transcoded into three MPEG-DASH representations with dynamic (live) playlist generation using the RTP2DASH module. By this, only the video bitrate is altered while maintaining the original video resolution (720x192) and frame rate (25 fps). Figure 17 illustrates the bitrates for the three MPEG-DASH representations. We set the highest representation (V1) with bitrate 1000 Kbit/s, which is close to the original RTP stream. The second representation (V2) is targeted to 500 Kbit/s and the lowest representation (V3) to 200 Kbit/s.

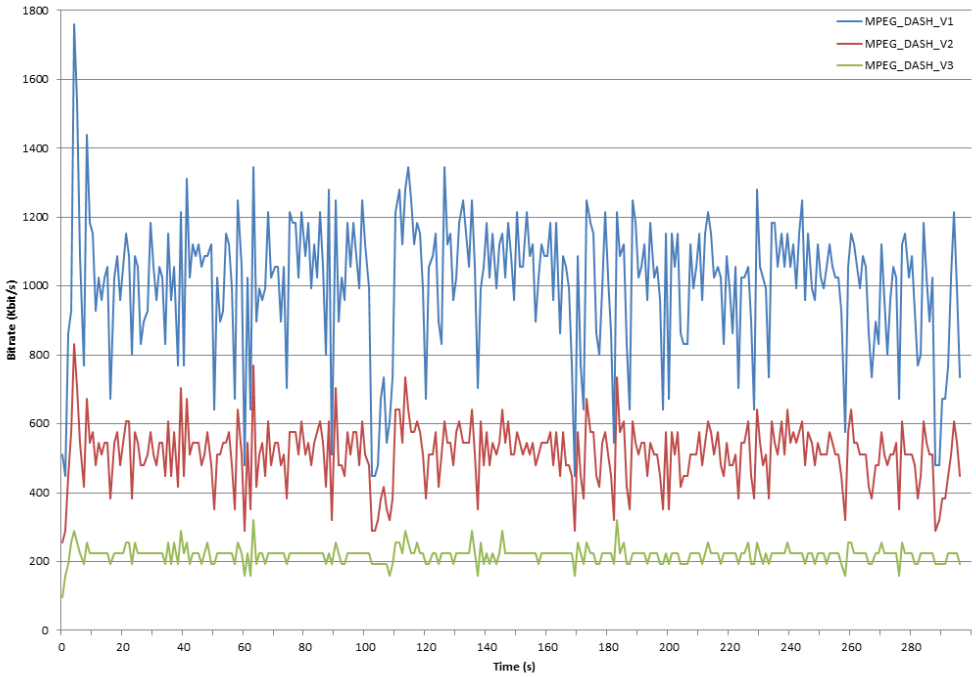

Figure 17 –Bitrates for the three MPEG-DASH video representations.

Figure 18 and Figure 19 present the objective quality evaluation curves (PSNR and SSIM) for the different MPEG-DASH representations. The average PSNRs for V1, V2 and V3 are 49.4, 44.2 and 38.8, respectively. Similar average values for the SSIM are 0.996, 0.990 and 0.976. Video quality has been estimated by comparing the transcoded videos at different rates to the video received at the Coordination Centre. This means that the absolute PSNR (and SSIM) values indicated in the result curves are not comparable with general PSNR qualities. The main purpose for the quality evaluation has been to estimate the quality difference between the videos at different rates, not the absolute values, which was not possible in the demonstration setup.

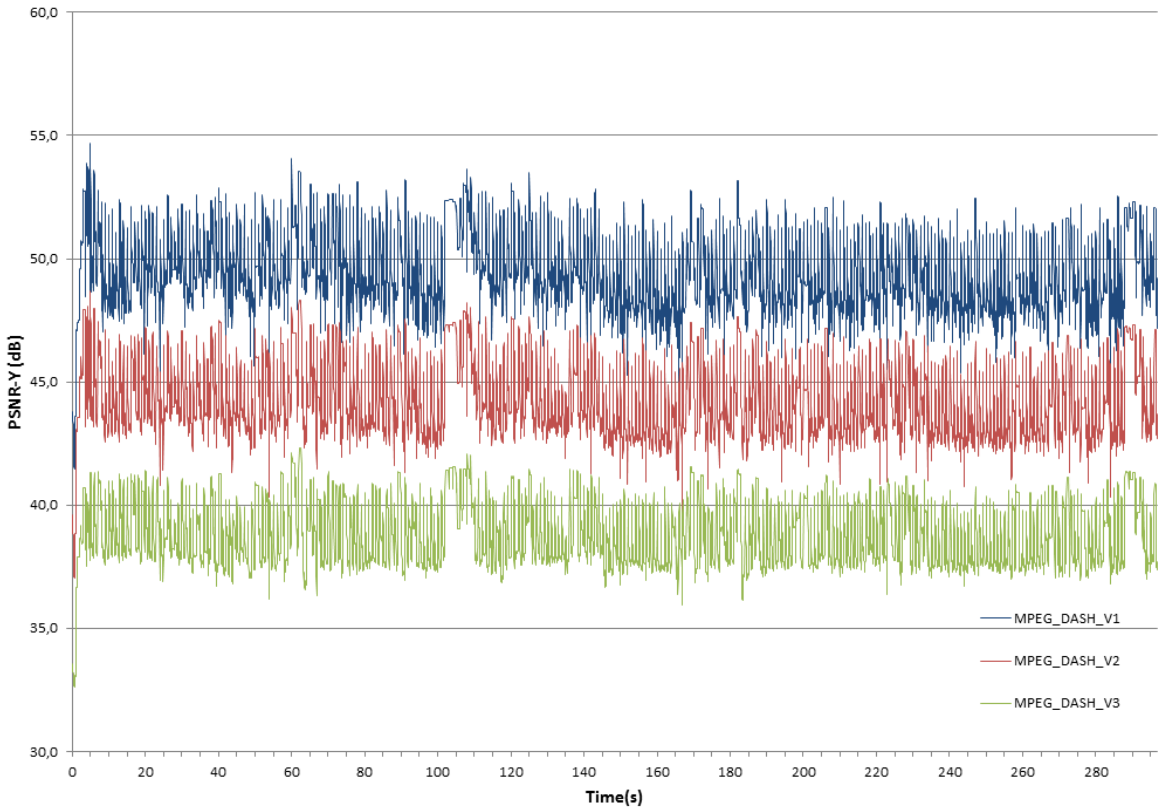

Figure 18 – PSNR quality values for the three MPEG-DASH video representations.

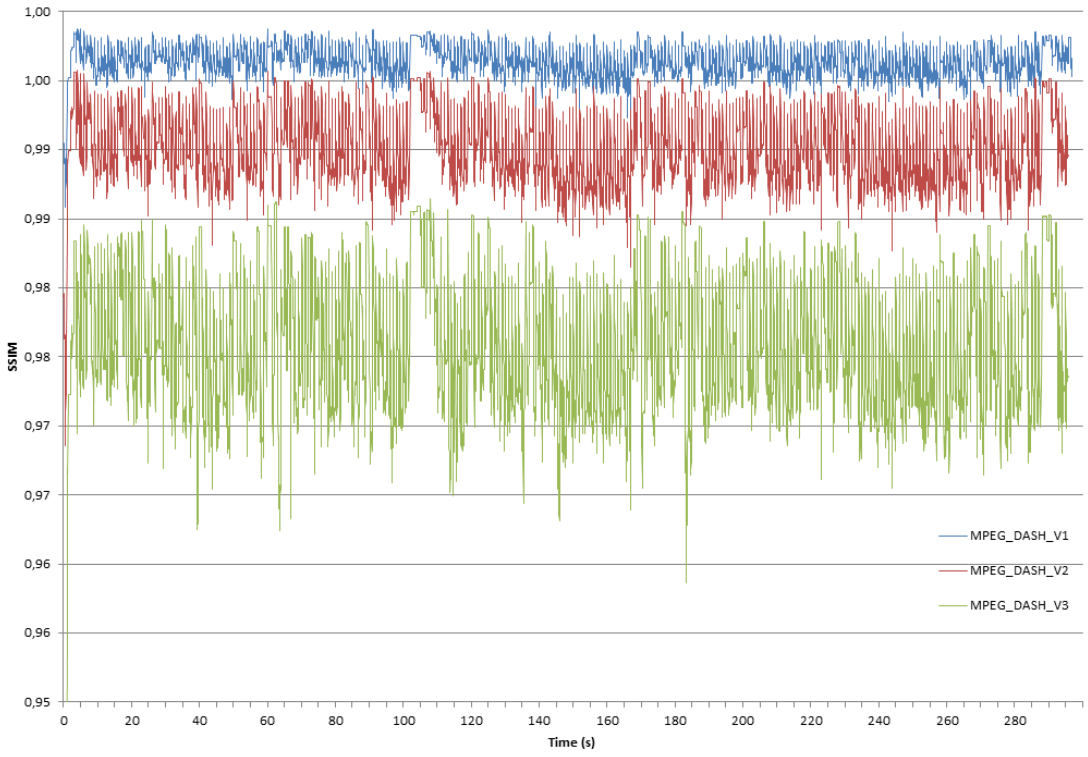

Figure 19 – SSIM quality values for the three MPEG-DASH video representations.

The MPEG-DASH demonstration scenario consisted of a laptop with five cores running Apache web server and GPAC software responsible of creating dynamic (live) transcoding of the RTP/RTSP content. Furthermore, we had an Android

Nexus 7 tablet running with Android 4.4 KitKat and modified OSMO4 MPEG-DASH player installed. We also controlled the network bandwidth between the RTP2DASH server and the tablet device. By this, we were able to adapt the stream between the three representations using simple adaptation algorithm available in the OSMO4 MPEG-DASH player. We utilised one second MPEG-DASH segment and fragment length. At the client side, three segment (3s) buffering was applied. We validated the MPEG-DASH demonstration scenario by monitoring the uplink throughput from the RTP2DASH server with Wireshark packet analyser.

Figure 20 b) presents the model for the available bitrate of Wi-Fi link between the Coordination Centre and wireless client on the move that was used in our experiments. The model is a simple example to demonstrate how fast adaptation algorithm reacts to changes in the bandwidth to estimate the delay. First, we have the full bandwidth at use (2M) which overruns the V1 representation bitrate (1M). At the point of 80s, we limit the throughput to 400K meaning that only V3 can be requested by the wireless client, which monitors the network throughput and is then able to adapt to the situation by requesting for smaller bitrate video representation defined in the playlist. Subsequently at the point of 110s, we limit the throughput to 600K, which triggers the adaptation routine to order V2 segments. Finally, at the point of 140 sec, we raise the throughput back to its original value 2M that triggers also the adaptation routine to order highest bitrate segments (V1).

Figure 20 a) illustrates the Wireshark throughput capture of our demonstration. As can be seen, the MPEG-DASH adaptation routine adapts well to the changes in the available network bandwidth. The peaks at the beginning and in representation switchovers are due to fast downloading to fulfil the playback buffer of the MPEG-DASH.

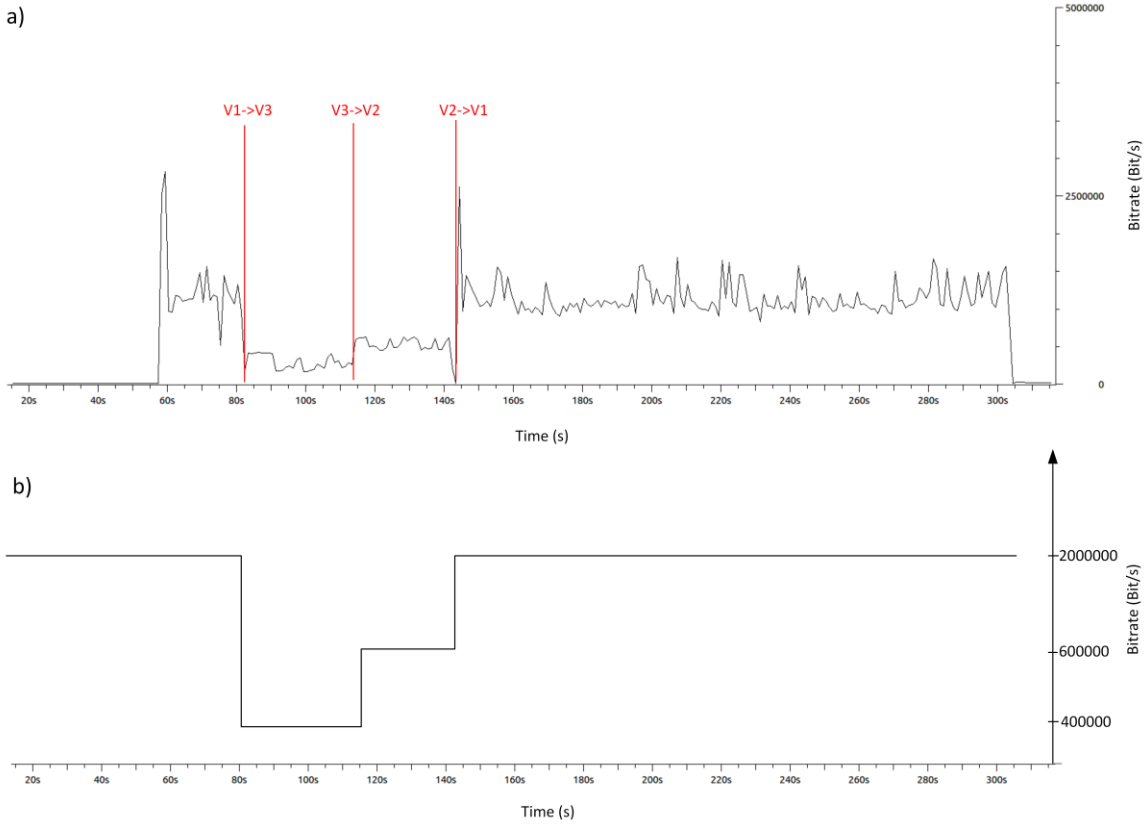

Figure 20 - Validation curves: a) Wireshark throughput capture for the sent MPEG-DASH stream towards Android client with adaptation indicators and b) The available bitrate of the Wi-Fi connection for the validation.

Figure 21 presents the objective quality curve (PSNR) for the demonstrated scenario. The red dotted line illustrates the same model as depicted in Figure 20 b). The most interesting metric of our evaluation is the delay introduced by MPEG-DASH compared to the case when RTP is used. We observed that the playback delay when adaptation is performed is approximately 4 seconds, which basically yields from the one second segment length plus three second player buffering delay. Our experiments confirm that adaptive HTTP streaming (in our case MPEG-DASH) can provide reliable alternative alongside the RTP/RTSP streaming with slight delay to be used in medical scenarios as indicated with simulations in earlier deliverables of WP6.

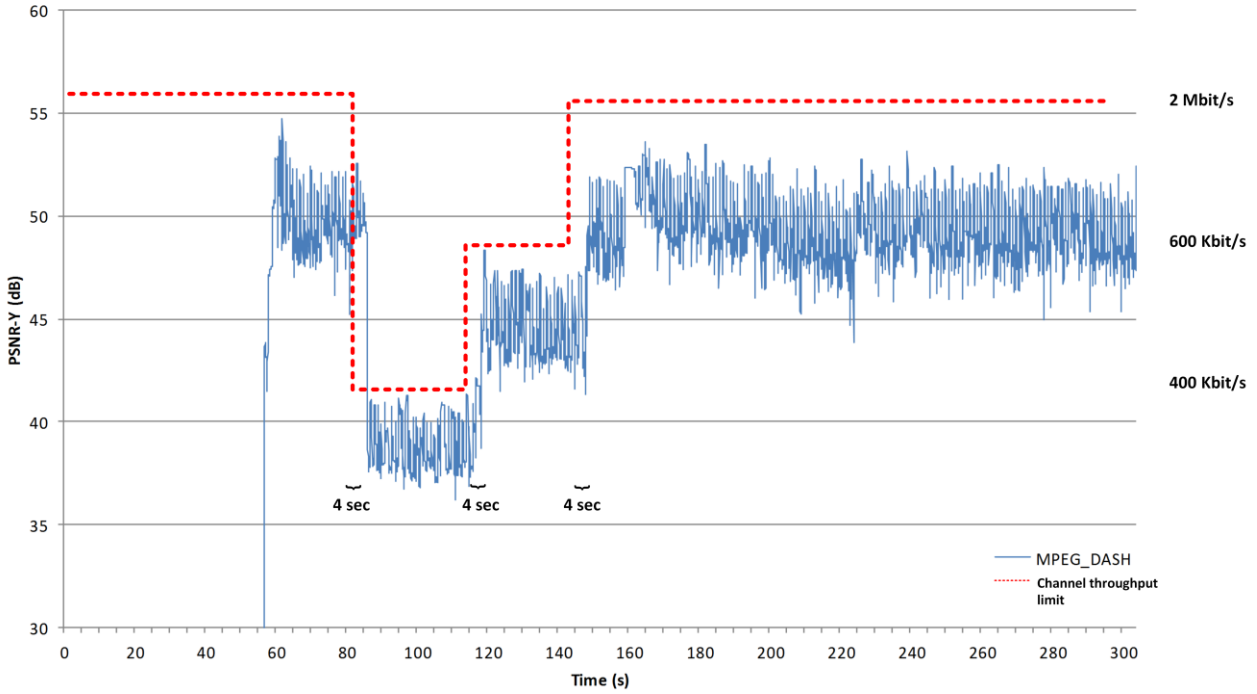

Figure 21 – Objective PSNR quality and observed delay in the Android client.

3.5 Ubiquitous tele-consultation and remote patient monitoring demonstration scenario

Recently, multi-access mobile devices interconnected with different wearable vital sensors via Body Area Network (BAN, also referred to as a Body Sensor Network or BSN) played an increasingly important role in healthcare and provide significant solutions for home healthcare, remote patient monitoring and real-time tele-consultation. The spreading of heterogeneous and overlapping wireless access technologies make possible for mobile healthcare applications to use the available network resources efficiently and call into being the ubiquitous Internet connection for these applications. These facts motivated us to design and implement a context-aware IPv6 flow mobility management scheme for multi-sensor-based mobile patient monitoring and tele-consultation. We designed and implemented a cross-layer optimization platform and a flow-aware, client-based mobility management mechanism for Android-based systems interconnected with different medical body sensors and Smartphone CCD video cameras. This demonstration scenario is focused on context-aware flow mobility mechanisms for multi-sensor mHealth services and highlights the effectiveness of CONCERTO solution in the context of ubiquitous tele-consultation applications with real-time patient monitoring support. A detailed description on our use-cases can be found in [5].

3.5.1 Signalling framework

Figure 22 presents the operation of our network-assisted Wi-Fi access point selection scheme. The first step according to the proposed signalling framework is the registration of the mHealth application. With this step the application notifies our framework to manage the network connections using the smart access point selection scheme. After the registration the algorithm starts the network context discovery session. The Context Discovery Module (CDM) queries information about the available Wi-Fi networks using NIS, which provides static information of wireless accesses such as cell ID, network SSID, geographical location of base station, channel, frequency band, IP and MAC address. Based on these parameters the CDM collects dynamic information of the selected network(s) through the DDE events. The most important DDE parameters are throughput in the uplink/downlink, rate of erroneously received and discarded packets, number of discarded packets and number of users, all measured at the Wi-Fi access points. The AHP Decision Engine makes decision about the optimal Wi-Fi AP based on the collected measurement data and directs the Network Management System to connect the mobile to the selected network. Figure 22 presents that in the first phase our Android device is connected to the AP1. The smartphone then notices through the DDE events that the current access point gets congested. The packet loss values of AP1 exceed the predefined threshold. The Network Management System directs the Android OS to connect to AP2, which provides suitable QoS parameters. The reader is advised to refer to [5] for a more detailed description on our cross-layer optimized signalling framework.

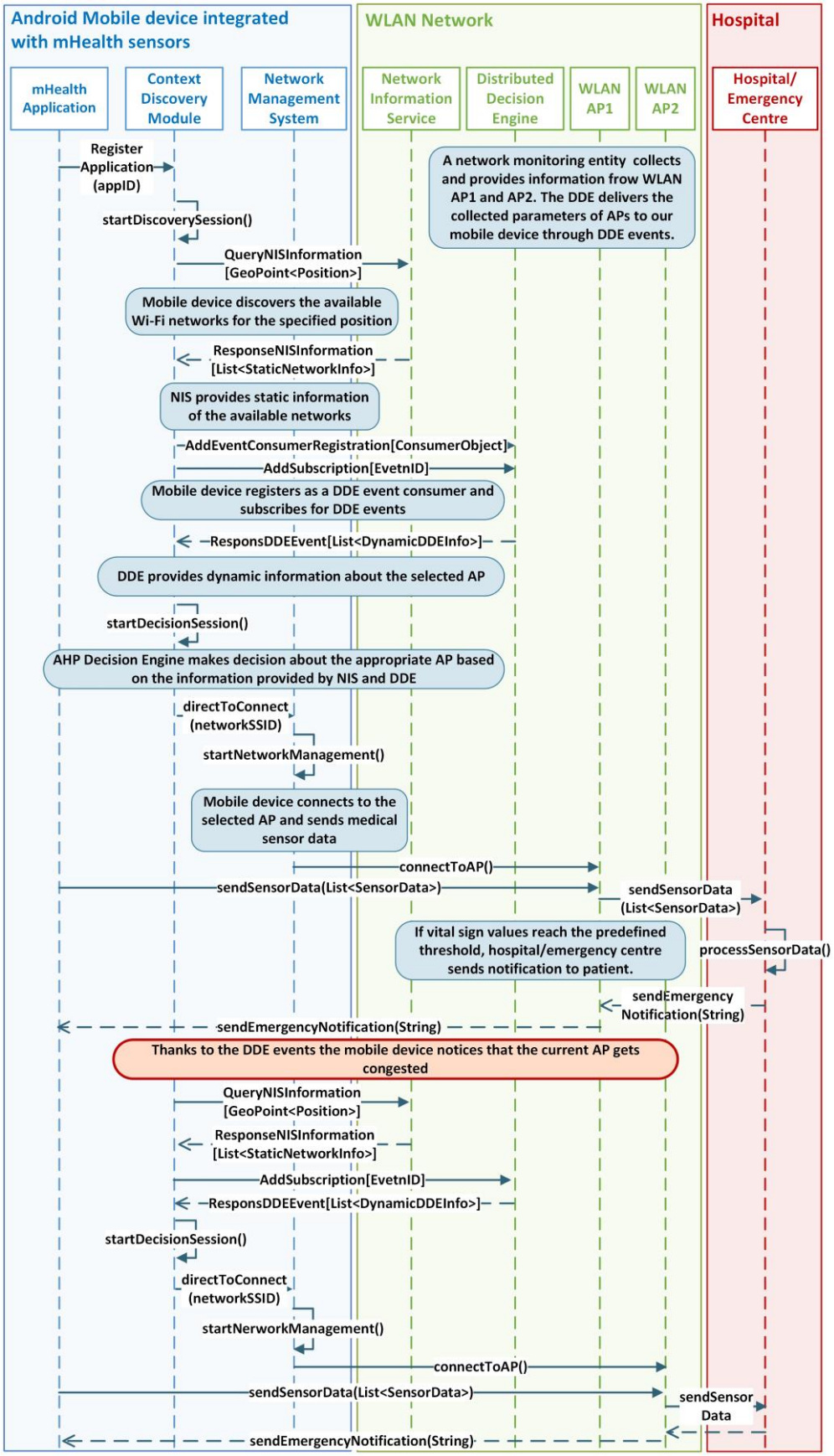

Figure 22 - The proposed and validated network-assisted access discovery and selection mechanism.

3.5.2 Testbed topology

Figure 23 presents the overall testbed topology designed and implemented for evaluating the proposed multi-sensor based mobile patient monitoring and tele-consultation system. The Home Agent is realized by a Dell Inspiron 7720 notebook running a MIP6D-NG daemon configured for Home Agent functionality. A Samsung Galaxy Note 3 Smartphone plays the role of the MN. The core parameters of our Smartphone: 3G and LTE/Wi-Fi connectivity, FullHD display resolution, 3 GB RAM, Qualcomm Snapdragon 800 Quad-core 2.3 GHz processor, 13 MP primary and 2 MP secondary camera, BT4.0LE, GPS. The two sources are heart rate information from the Zehpyr HxM and HD live stream from the Smartphone's camera. A Samsung Gear Fit is used to show notifications sent by the hospital/emergency centre. The Gear Fit communicates via Bluetooth 4.0 LE with host Android Smartphone. It has a 1,84 Curved sAMOLED touchscreen display and possesses different built-in sensors, such as accelerometer, gyroscope pedometer and heart rate monitor. Also Zephyr prefers Bluetooth communication, is able to measure heart rate, speed and distance. Bluetooth low energy or Bluetooth LE (or BT Smart) is intended to provide reduced power consumption. BT LE is used in the healthcare, fitness, security and home entertainment applications. It uses 2.4 GHz frequencies, and the maximum of data rate is 1 Mbit/s. Video information is provided by the Smartphone's CCD sensors of the primary and secondary integrated cameras.

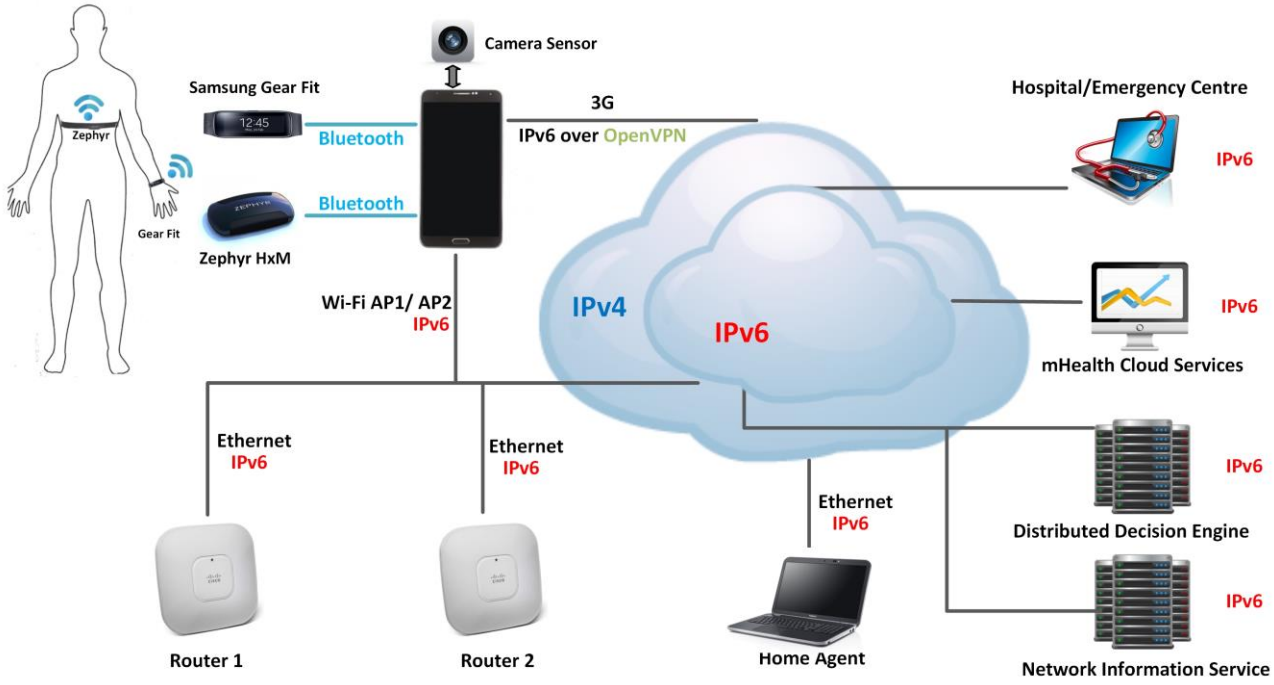

Figure 23 - Evaluation Testbed Topology.

3.5.3 Validation methodology and results

3.5.3.1 DDE integration scenarios

We designed and implemented a network-assisted wireless access network selection framework for mHealth services, which is able to select the appropriate available access network using a multi-criteria decision engine. The proposed engine relies on the Distributed Decision Engine (DDE) and the Network Information Service (NIS) providing both static and dynamic parameters of the available networks. In order to present the efficiency of our mobility framework and to evaluate the proposed multi-criteria decision engine, we designed and implemented two different measurement scenarios. In the first scenario we simultaneously transfer medical sensor information (heart rate values provided by Zephry HxM), high resolution real-time video stream (using the built-in camera of our Android device) and numerous VoIP flows as background traffic. We measured the delay and packet loss of the currently used AP on the correspondent node which runs and controls the background traffic.

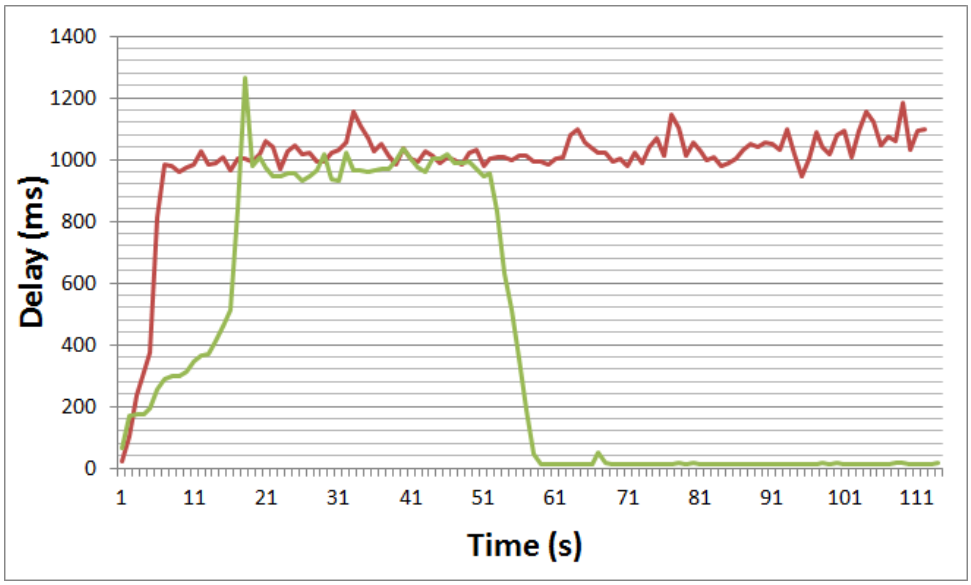

Figure 24 - Delay on AP1 without mobility [red line] and with the proposed framework [green line].

Figure 24 shows that without the proposed Smart Access Point Selection mechanism the delay is high during the overall measurement session. Using our mHealth framework the Android device is able to detect the inconvenient QoS parameters of AP1 and executes a handover from AP1 to AP2. In the first period of the measurement the average delay is high, approximately 1000 ms. However, after the handover from AP1 to AP2 only the background traffic is on AP1, which induces 5-10 ms delay. We limited the capacity of AP1 to 5.5 Mb/s in order to reach the saturation level with less background traffic. The AP2's capacity was 54 Mb/s. We stressed only AP1 with the background traffic. Packet loss values of AP1 in the aforementioned scenario are presented on Figure 25.

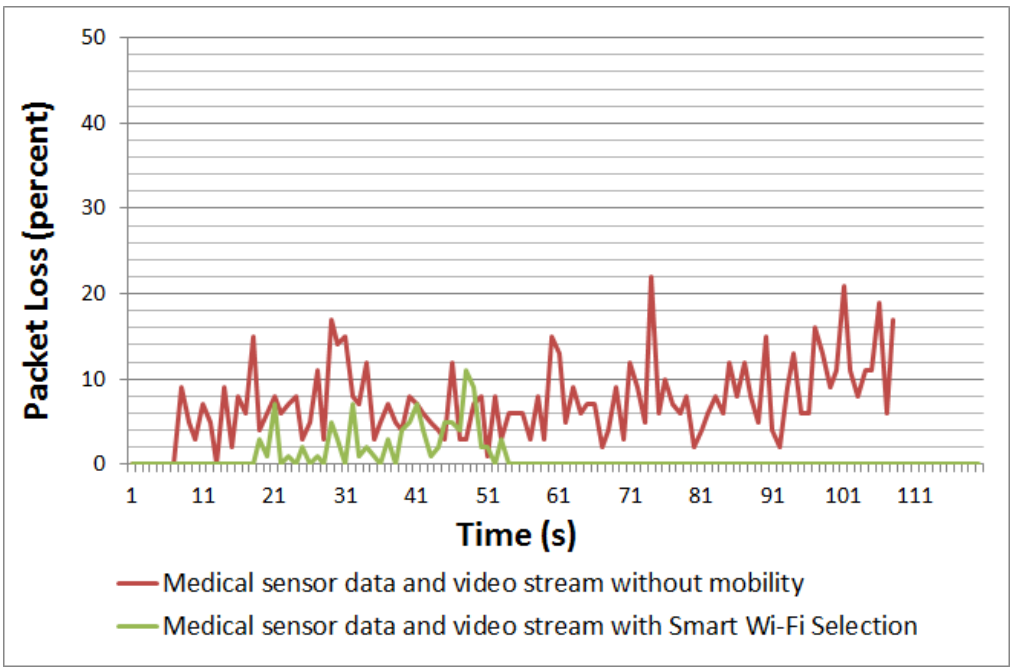

Figure 25 - Packet loss in AP1 without mobility [red line] and with Smart Access Point Selection framework [green line].

In the second scenario we monitored the delay and packet loss values of AP1 and AP2 on our Android device using the cross-layer feedback information from the DDE events. Based on the measured data the proposed Smart Access Point Selection executes handover from AP1 to AP2. Figure 26 shows that there is quite high delay and packet loss in AP1, which are not acceptable for the mHealth services.

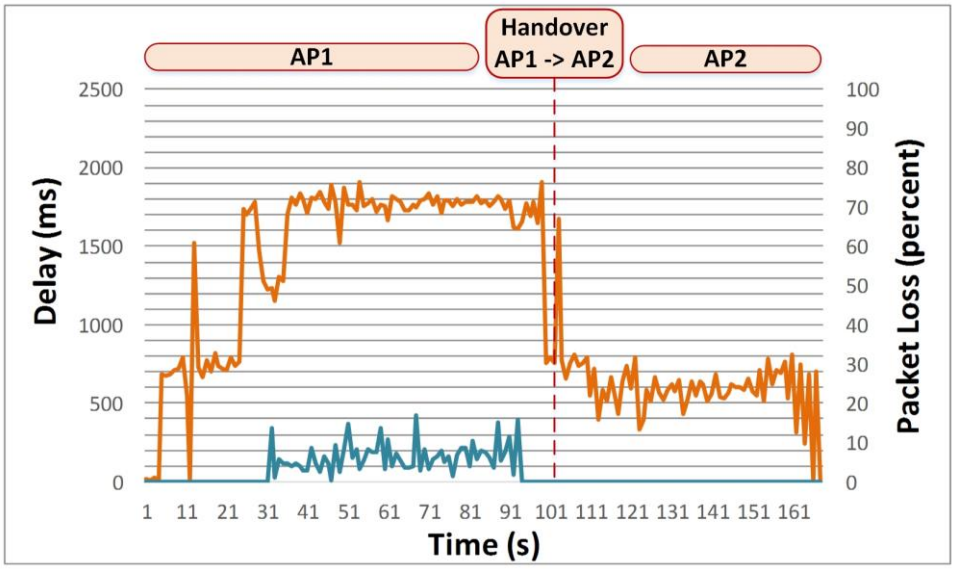

Figure 26 - Delay [orange] and packet loss [blue].

After the packet loss crossed the threshold, our Network Manager System executes a handover from AP1 to AP2. The QoS parameter values of AP2 are suitable for our application, because there are no packet losses and also the delay is normal.

Figure 27 depicts the communication delay on the currently used access point and also highlights the heart rate data of our biovital sensor. In this scenario we transfer a camera stream and medical sensor data towards to the hospital. Due to the background traffic in AP1, the delay starts to grow continuously. Our framework detects the decreased quality of service and initiates a handover execution.

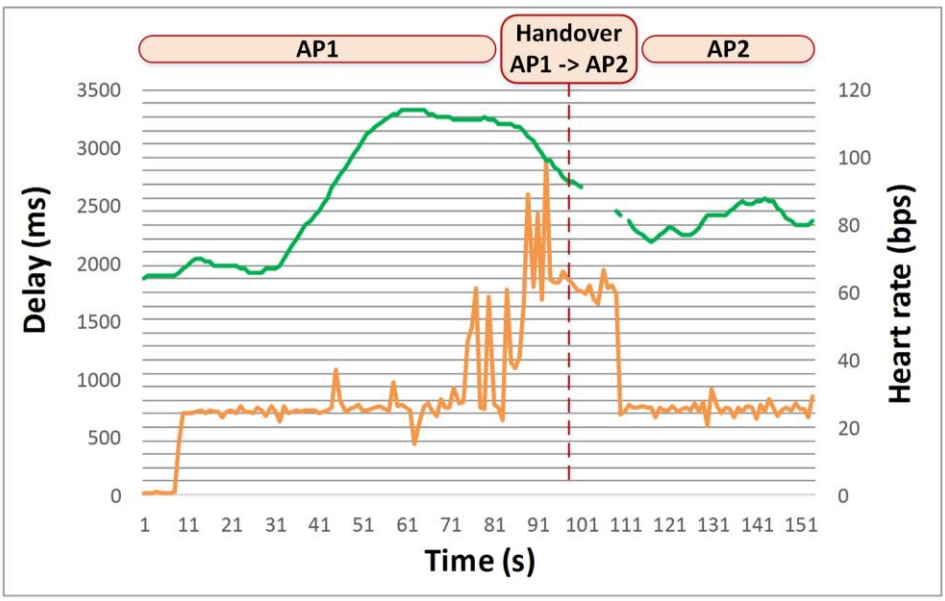

Figure 27 - Delay [orange] and heart rate traffic [green].

3.5.3.2 Performance tests

Figure 28 presents the components of the average handover latency measured during flow mobility events. The two considerable parts are the latency between the Java and the native layer, and the delay between the reception of the flow update command and the sending of FBU message towards the HA by the MIP6D-NG daemon.

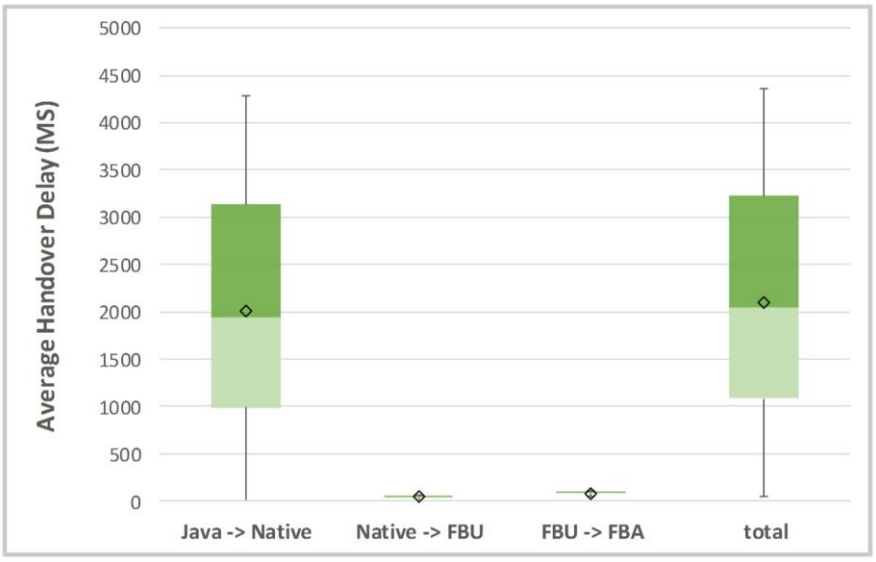

Figure 28 - The total handover latency in its three main components on Samsung Note 3.

The quite high (~2000 ms) latency does not imply serious practical issues, because during these procedures the registered flows will not be affected: data transfer will harmlessly continue till the third phase (FBU->FBA signalling) starts. This last component is the delay between the sent FBU and the received FBA messages. Figure 29 shows only the native components of the handover process measured during the flow mobility event.

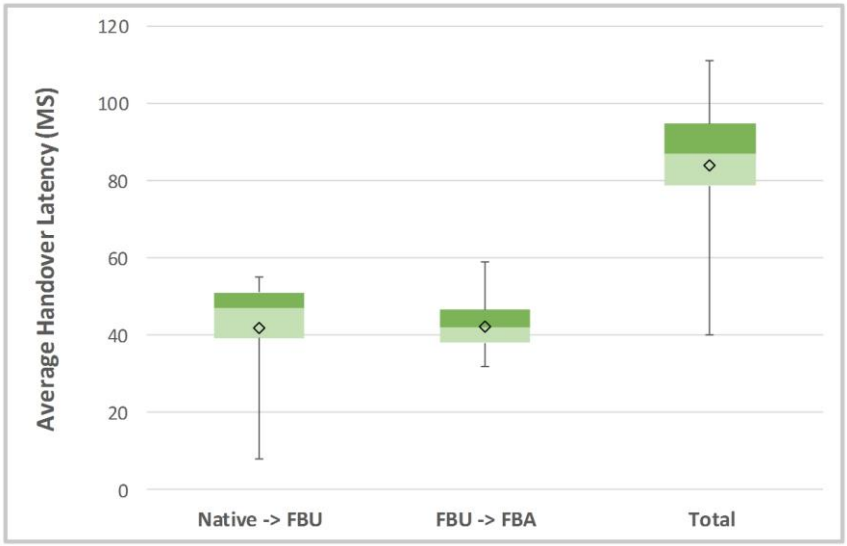

Figure 29 - The native components of the total handover latency on Samsung Note 3.

3.5.3.3 Framework tests

In this scenario we measured the quality of medical sensor and high-resolution video flows transmitted from patient's smartphone toward to the hospital/emergency centre during different mobility events. In the first phase of the proposed mobility framework our smartphone initiates both data flows on the only available 3G/LTE interface. Due to the limited network resources of 3G/LTE both data flows suffer high packet loss, thus neither video flow nor sensor data values can be transmitted with medical quality. Our smartphone detects a new Wi-Fi network and discovers the static and dynamic parameters of this AP using DDE and NIS. The proposed decision engine finds appropriate this AP and moves the video flow from 3G to Wi-Fi (Figure 30).

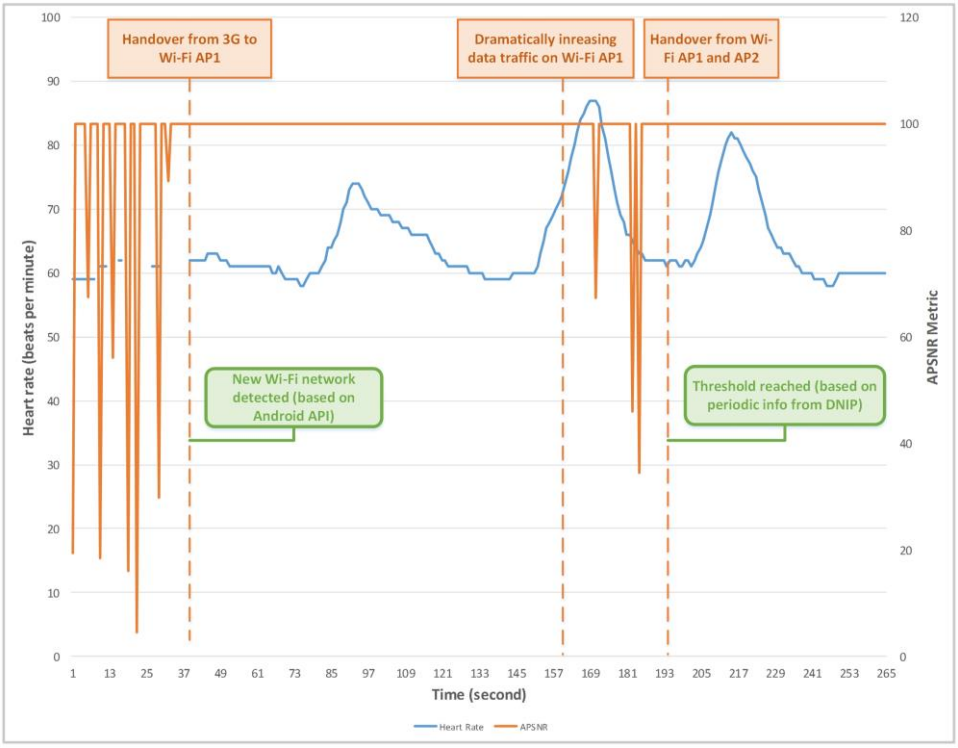

Figure 30 - Evaluation of video and medical sensor data flow in heterogeneous wireless environment.

Our framework detects dramatically increasing data traffic, the predefined threshold is exceeded, which indicates another handover, thus our cross-layer optimized mobility management system moves the video flow from AP1 to AP2.

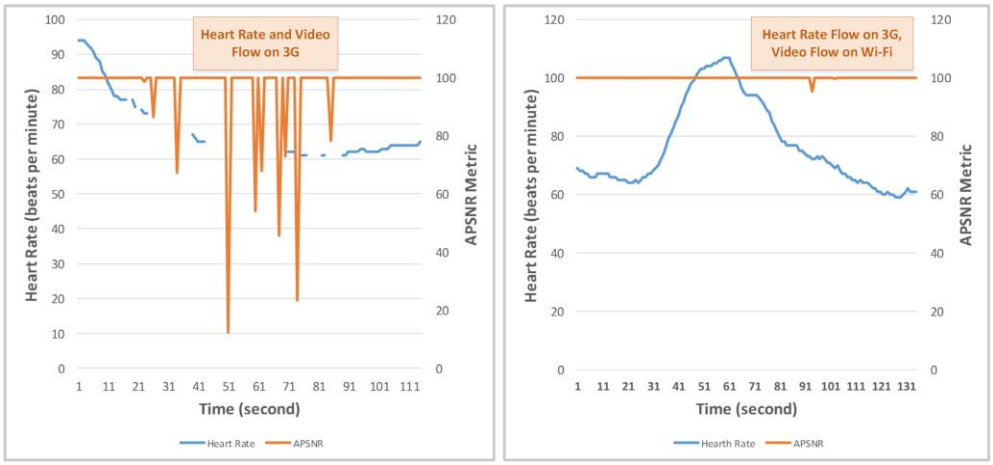

Figure 31 - Video and medical sensor data flow initiated on different wireless interfaces.

Figure 31 shows the heart rate and the APSNR metric [12] of the video in two different scenarios. The graph on the left depicts a scenario, when flows of the video stream and of the heart rate are both assigned to the 3G interface (3G is the only available network at the beginning of this case). The bandwidth of 3G is not enough for both flows, thus the quality of video degrading and also the heart rate information flow suffers severe packet loss. The scenario, when the two flows are assigned to different interfaces (3G and Wi-Fi) is presented by the second graph. Then an appropriate Wi-Fi network appears that will be detected by our framework. The proposed algorithm moves the flow of the video stream from 3G to Wi-Fi, which results in continuously high video quality and unharmed heart rate transmission. The scenario, when the two flows are originally assigned to different interfaces (3G and Wi-Fi) is presented by the second graph. Note, that such scenarios require overlapping radio accesses even from the beginning of the operation.

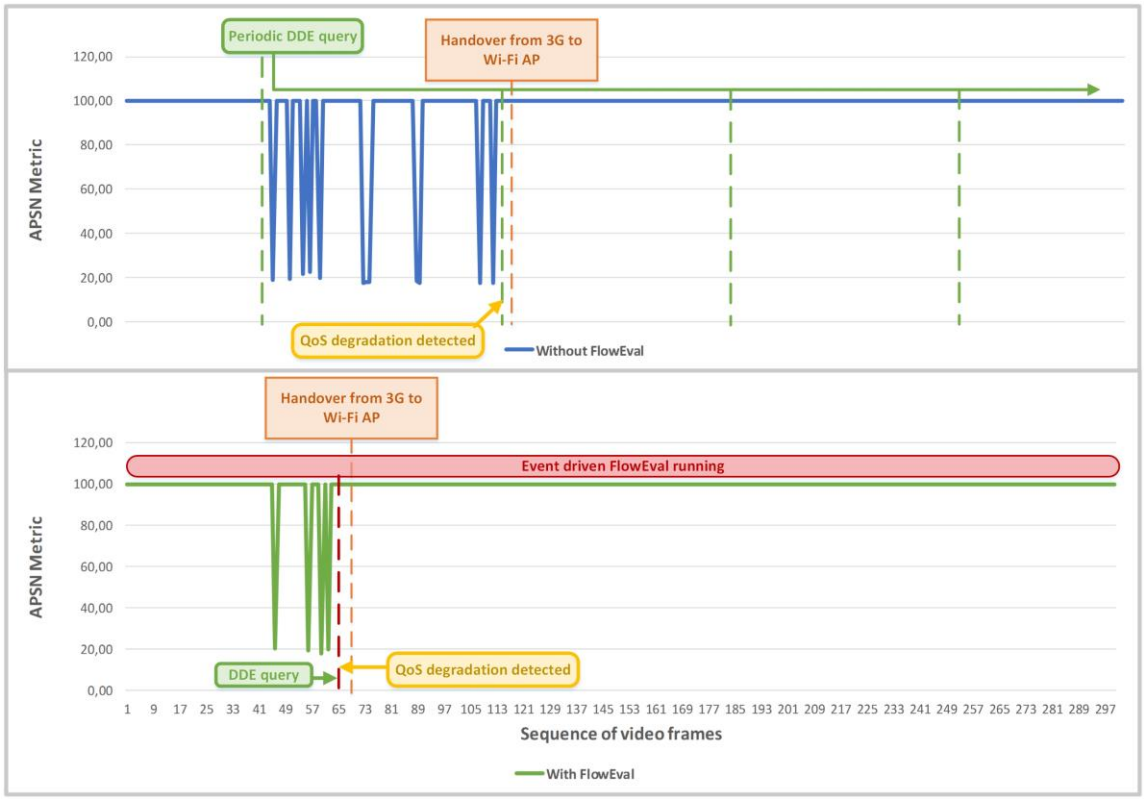

Figure 32 - Periodic and event-driven data collection scheme.

Figure 32 presents the comparison between periodic and event-driven data collection mechanism. The so called FlowEval module of the proposed architecture runs measurements in the background, which are able to detect the degradation of the current used network. With this mechanism our solution can react to the changed environment faster. FlowEval requires more resources, but can be more efficient in emergency situations.

The main purpose of the field trials for this demonstration scenario, was to validate the CONCERTO Smart Wi-Fi Selection mobility infrastructure for real-time mHealth services where multi-sensor based mobile patient monitoring and tele-consultation is the main scope. The decisions taken by the proposed framework rely on the Distributed Decision Engine and the Network Information Service. Results show that our framework and algorithms help to provide medical level QoS/QoE for the quality-sensitive real-time, interactive mHealth services.

3.6 3D medical data storage and encoding test and results

The 3D medical data storage and encoding test system has a web-based user interface, where input (original 3D medical data) files can be selected, 2D preview montage can be generated, parameters for directory-based and JPEG2000 encoding can be set, volume rendered image can be generated and statistics about the test results can be viewed. The interface has a simple menu-system, from where each important service is available (Figure 33). This system is implemented at the Coordination Centre where the medical data are stored.

Figure 33 - Web interface of the 3D medical data and storage demo.

The 3D medical data storage and encoding test system has the following functionalities: viewer for 2D Medical data (1 slice), montage-viewer for set of 2D medical data, free point of view volume renderer for 3D medical data, transfer function changeability for volume renderer, efficient 3D medical data storage with JPEG2000 and H264 video coding and 3D RAW data encoder and decoder.

3.6.1 Measurement story-line

The 3D medical data sets are stored in the database of our test server. One 3D data set consists of slices through an object, where each 2D slice is normally presented as a grey scale image. We can view the 2D images one-by-one, or get a montage of the 2D slices (as a preview), in 9, 16, or 64 pictures in array (Figure 34).

Figure 34 - Preview from 16 and 9 slices.

In order to view the original 3D data, we apply a texture-based volume rendering technique, where we perform the sampling and compositing steps by rendering a set of slices inside the volume.

Figure 35 - Rendering teeth with higher intensity.

Using our volume rendering engine, we can render a free point of view 3D image from the original data set. We use an *optical model* to map data values to *optical properties* (like color and opacity). Optical parameters are calculated by applying *transfer functions* to the data. For example, with changing the transfer function, we can render the tooth with more intensity than the surrounding bones (Fig. 14).

In our tests, at first, user selects the original 3D data to be processed. An optional montage of 2D slices is available, where the user sets the number of slices to be viewed. A free point-of-view volume rendered 3D image is also applicable, with some optional settings for the rendering (like parameters of the alpha blending). After choosing the applied directory-base methods (lzma, gz2, etc.), and setting the parameters for JPEG-2000 and H264-based encoding, the medical data file is encoded in several different ways, and statistics are generated about the different compression methods. The complete flow-chart of the above mentioned procedures are depicted in Figure 36.

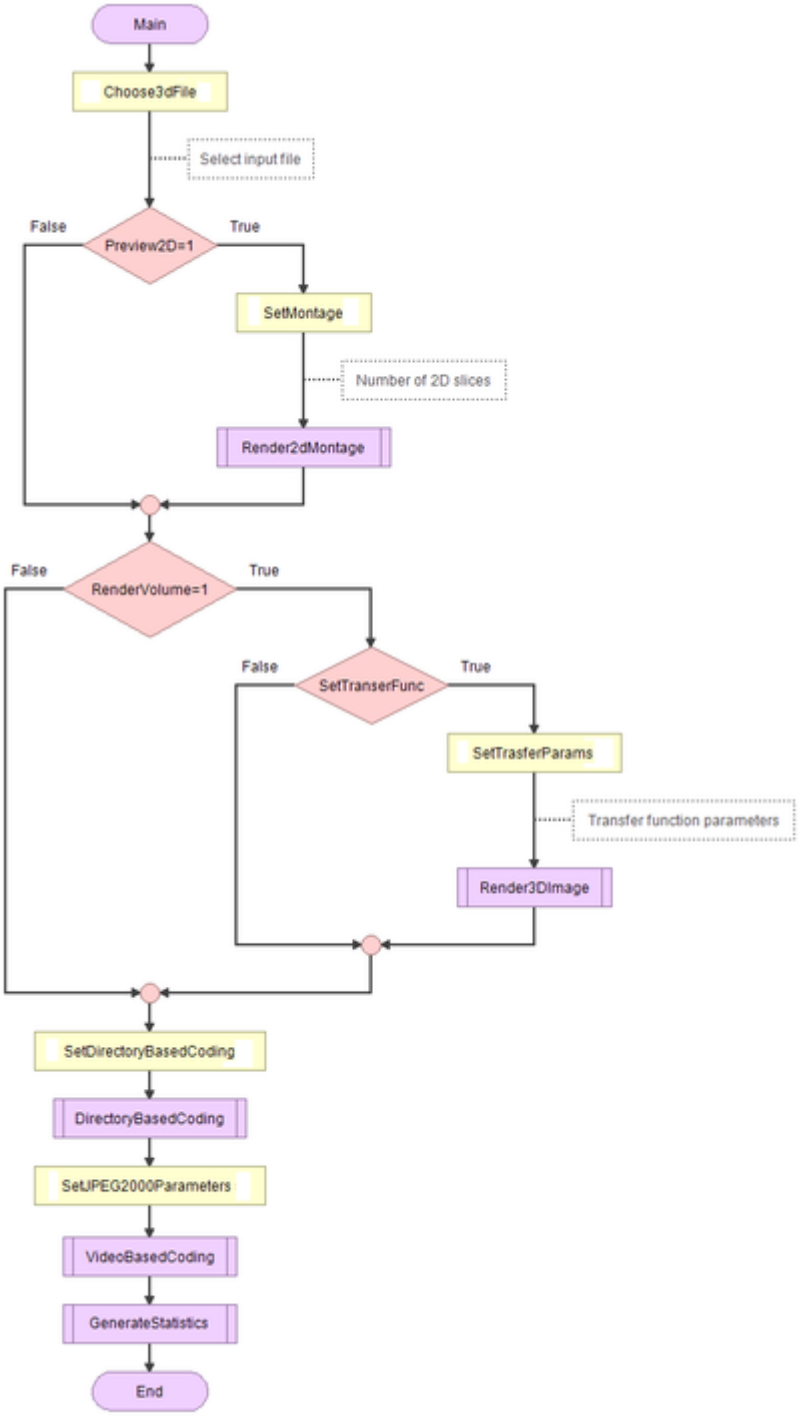

Figure 36 - Flowchart of the test.

3.6.2 Test results

In order to design an efficient storage environment of 3D medical data, we applied and compared several compression methods using our environment. The used compression methods are: lzma, gzip, bz2, lzma with –e parameter, and lossless jpeg2000. Our test system generates statistics from the efficiency of compression. The directory-based compression methods are proved to be quite effective because compared with continuous-tone images there are significant sized black blocks in the original 3d medical data sets (voxels inside the bounding box, but outside the patient's body). According to our test results, the jpeg2000 techniques are quite efficient for encoding medical data sets.

We have calculated the Shannon-entropy of the test files and the average code word length to find out the theoretical minimum of the lossless encoded files. The theoretically minimal file size is more than 99.99% of the j2kencoded file,

so the JPEG-2000 lossless encoding could be considered quasi-optimal and practically applicable (against the Huffman-coding).

In addition to the efficient lossless compression, the JPEG-2000 techniques provide near-lossless and highly scalable lossy compression.

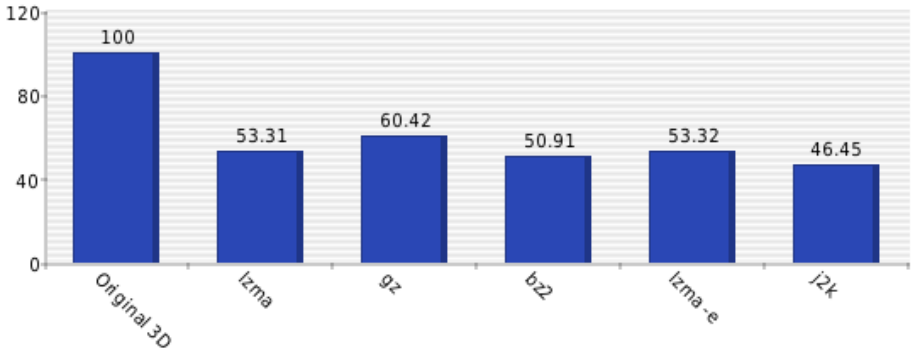

Figure 37 - Applied lossless file compressions - file sizes in percentage.

We investigated the lossy JPEG-2000 compression file sizes at different Peak Signal to Noise Ratio. At 80dB PSNR (and above) the file size is about equal to the lossless compressed file. So if the medical application requires at least 80dB PSNR, the lossless compression is more relevant.

Under 40dB PSNR the degradation of quality makes the compressed file useless. Between 40 and 80 dB, the file size increases sharply (almost linearly). The shape of the function reminds us the (transformed) arctan function.

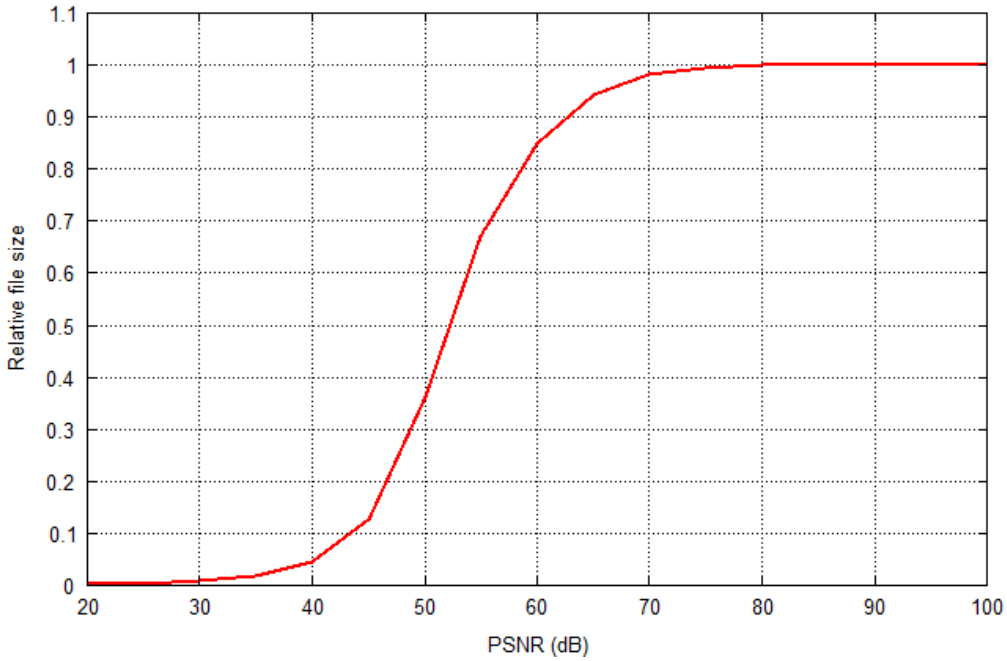

Figure 38 - Relative file size - PSNR function.

We have investigated the required compression time at different PSNRs (Figure 39). According to our measurements, the compression time is nearly independent of the PSNR (constant).

There are local extreme points and increasing and decreasing intervals on the function, but the test was not performed with a dedicated machine, so fluctuations may have been caused by other background processes. It could be said that with lossy compression the compression time is irreducible.

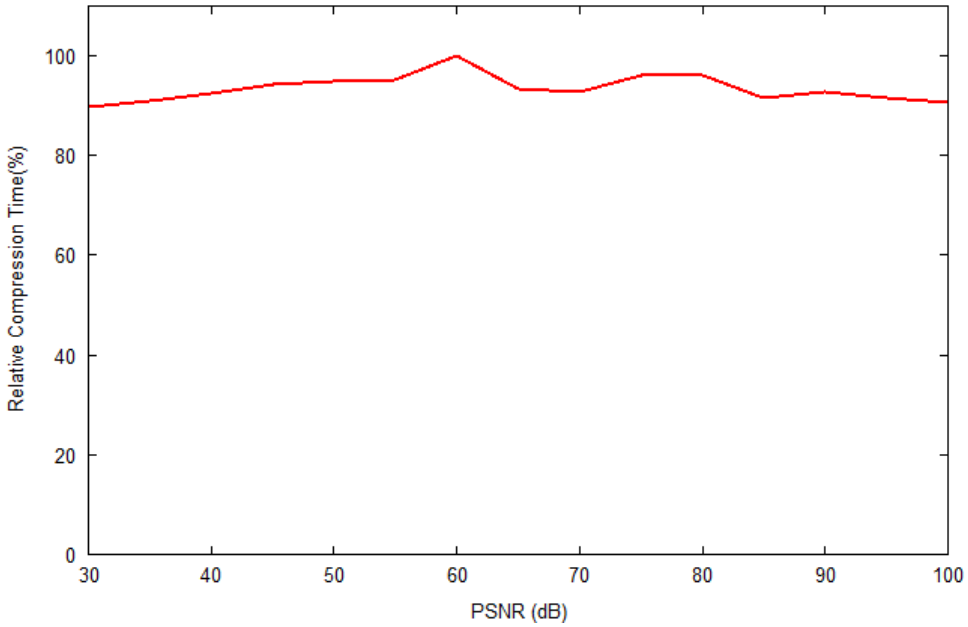

Figure 39 - Compression time - PSNR function.

Investigating the decompression time, we obtained a near-linear, slightly increasing tendency (Figure 40). Reducing the image quality, the decrease of the decompression time is significantly lower than the decrease of file size.

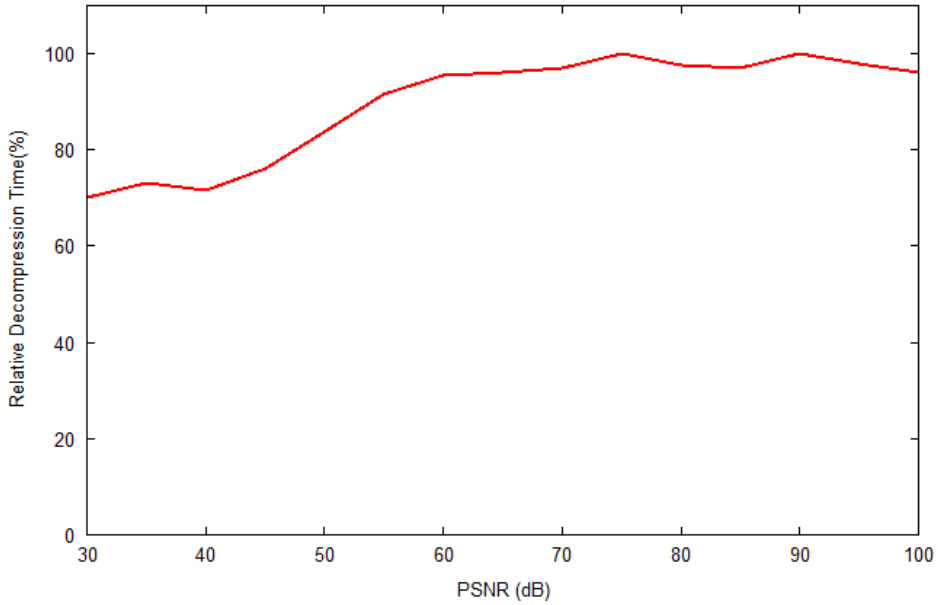

Figure 40 - Decompression time - PSNR function.

Because of the correlation between neighbouring slices, we can interpret the sequence of 2D slices as frames of a video stream. The differences between frames could be interpreted as motion vectors. The high correlation between frames implies effectively calculated and stored motion vectors. In our test system, we generate a video from each 3D data set, using lossless video encoding (H.264 with ultrafast and very slow presets, and ffv1).

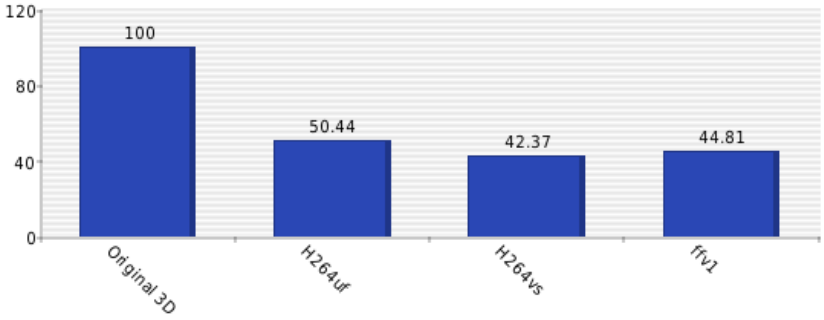

Figure 41 - Applying video coding for 3D data sets file sizes in percentage.

According to our test results, the H.264 with ultrafast preset is much faster and near as efficient as the best directory-based methods. The H.264 with very slow preset is more efficient compared to the directory-based methods.

4 Medical image/video quality assessment tests and results

This section presents quality assessment results obtained from the hospital for different CONCERTO use cases. In particular, the quality of the video sequences obtained following compression via HEVC and H.264 of the cardiac ultrasound sequences provided by the hospital has been assessed by the medical doctors in the hospital via subjective tests and the relevant results are presented in this section. Following discussion with the cardiologists in the hospital, two different methodologies have been used for the tests (DSCQS and simultaneous direct comparison of all the impaired versions).

The video sequences received following transmission via (adaptive) HTTP streaming have also been assessed by the medical doctors via subjective tests. The relevant results are presented and commented in the second part of this section. Finally the medical doctors participated to the field trials realized with CONCERTO demonstrator and stated about the received quality of the medical videos from a diagnostic point of view.

4.1 Use cases considered

Quality evaluation is essential in all scenarios where compression and transmission of medical images and videos is involved. In the use cases 1, 2, 3, 4, 6, and 7, compression and transmission of medical data is done. In these use cases, the medical video quality evaluation is implemented at the receiver side, i.e., when the doctors receive the video sequences via transmission or store the video sequences following compression. In particular, these results are relevant for the scenarios on emergency (ambulance) and tele-consultation.

4.2 Quality Assessment of Medical Ultrasound Videos compressed via HEVC

Video compression is one of the ways of reducing the bandwidth requirements for transmission of videos over communication channels. Basically, video compression is the removal of redundant and also irrelevant data (and sometimes relevant data) from the original video in order to represent it in lower bitrate without significantly affecting the perceptual quality of the video. Lower bitrate video would require less channel resources for transmission and thereby contributes in improving the transmission performance. High Efficiency Video Coding (HEVC) is the latest video compression standard developed by the Joint Collaborative Team on Video Coding (JCT-VC) as a joint project of "ITU-T SG 16 WP 3" and ISO/IEC Moving Picture Experts Group (MPEG). HEVC is a direct successor of the previous video compression standard, H.264/AVC, developed with the aim of achieving bitrate reductions of 50% with respect to its predecessor, in particular to facilitate better compression ratios for High Definition (HD) and beyond-HD video formats. Due to its remarkable bitrate saving capability, HEVC is expected to be widely used both for research and commercial purposes.

The effect of compression and transmission of videos often results in a reduced video quality. Therefore, it is essential that the quality of the medical images and videos received is monitored via Video Quality Assessment (VQA) techniques, so that, along with other objective quality evaluation measures, it can contribute to facilitate the design of future medical multimedia services and applications. Typically, for VQA of medical videos two approaches are widely used, namely, objective and subjective methods. Objective VQA approaches use mathematical models mainly designed to evaluate the perceptual quality of the video. Subjective VQA approaches for medical videos generally involve medical experts (and sometimes non-experts) evaluating the quality of the video mainly in terms of its diagnostic importance.

We study the performance of the HEVC standard on medical ultrasound videos via rate-distortion and rate-quality analysis to evaluate the diagnostic and perceptual quality of the video sequences. The subjective assessment of the medical videos is done by both medical experts and non-experts from whom subjective scores are obtained. Some relevant results have been presented in deliverable D3.3, Section 3.1.1 [2] and in [7]. Differently from the works in [2] [7] where the focus is on comparing the results of subjective tests with the results obtained via different objective metrics, the focus is here on rate-distortion and rate-quality analysis.

4.2.1 Test Methodology

Video Sequences

The performance of HEVC is evaluated on nine original medical ultrasound videos with a frame resolution of 640 × 416, each compressed at eight different quality levels. Each video sequence has 100 frames, encoded at 25 frames per

second (fps). Of the nine ultrasound videos, three videos are related to the heart and liver each, two for kidney, and one video is related to the lung. An example frame from three sequences is shown in Figure 42.

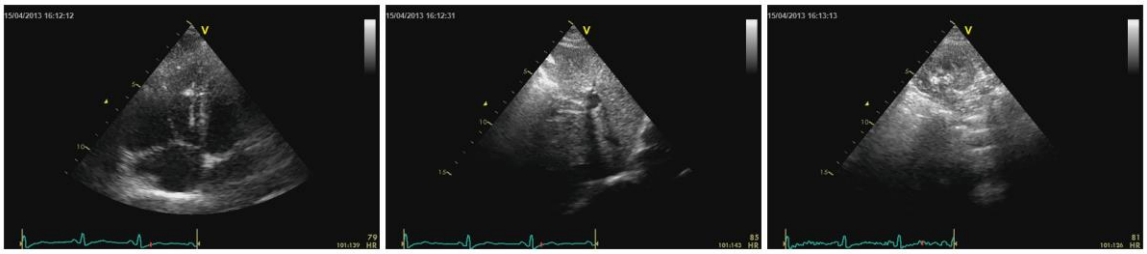

Figure 42 - An example frame of some of the sequences used in the tests. Left to Right: (a) Echocardiography: 4 chambers view. The right ventricle is dilated. (b) Echocardiography: the subcostal view displays the liver and the inferior vena cava. (c) Renal ultrasound: cortical and medullary view.

The compression of the sequences is done at eight different Quantization Parameter (QP) levels using the JM reference software provided by the Joint Collaborative Team on Video Coding (JCT-VC) team. The QP values chosen are 27, 29, 31, 33, 35, 37, 39, and 41. As the QP value increases, the compression ratio increases which in turn gives lower quality videos. In the tests, we used: 9 video sequences, compressed at 8 different QPs, i.e., $9 \times 8 = 72$ impaired medical video sequences.

Subjective Test

The compressed video sequences were subjectively evaluated for the visual and diagnostic quality by both medical experts and non-medical experts who provided their opinion scores on a scale of a specified range. The subjective evaluation was done using the Double Stimulus Continuous Quality Scale (DSCQS) - type II, which is one of the methodologies recommended by the International Telecommunication Union (ITU) in the document ITU-R BT.500-115. The DSCQS method was adopted in our tests because in this method the subjective scores are less sensitive to the context, i.e., the ordering and the level of impaired sequences has less influence on the subjective ratings [8].

The DSCQS methodology uses a Just Noticeable Difference (JND) approach in which the medical expert is presented with two videos side by side, typically the original and a processed video. The subject is asked to assess the quality of both the videos. One of the sequences is the reference video, i.e. unimpaired video, whereas the other sequence is impaired. The subject is asked to rate both the sequences on two separate scales of 1 to 5, where 1 corresponds to the lowest and 5 to the highest quality. The subject is unaware of which one is the reference video (the reference video is displayed randomly either at the left or at the right end side). The Moscow State University (MSU) perceptual quality tool 11 was used to document the score obtained in the subjective study. The ratings obtained were then used to get the mean scores and other desired statistics.

Subjective Scores

For the subjective evaluation, four medical opinions and sixteen non-medical opinions were collected. The experts rated the video sequences mainly for their diagnostic quality, whereas the non-experts more likely rated based on the perceived visual quality. In the DSCQS method, for each video sequence, two ratings were obtained. One of the scores corresponds to the reference video and the other to the impaired video. The difference between the subjective scores of the original and the impaired sequence for each subject is calculated and the mean across all scores taken to obtain the Differential Mean Opinion Scores (DMOS).

4.2.2 Performance Evaluation

The Rate-Distortion performance of HEVC for the considered medical ultrasound videos are shown in Figure 43 for three heart and three liver sequences for QPs 29, 33, 37, and 41.

The rate-distortion performance of HEVC for the heart and liver sequences is assessed by considering the bitrate and the PSNR of the compressed sequences. It can be observed that the rate-distortion performance is influenced by the spatio-temporal complexity of the sequences. For the liver sequences which have relatively lower spatio-temporal complexity, the rate-distortion performance appears to be better.

The rate-quality performance of HEVC for the considered medical ultrasound videos are shown in Figure 44.

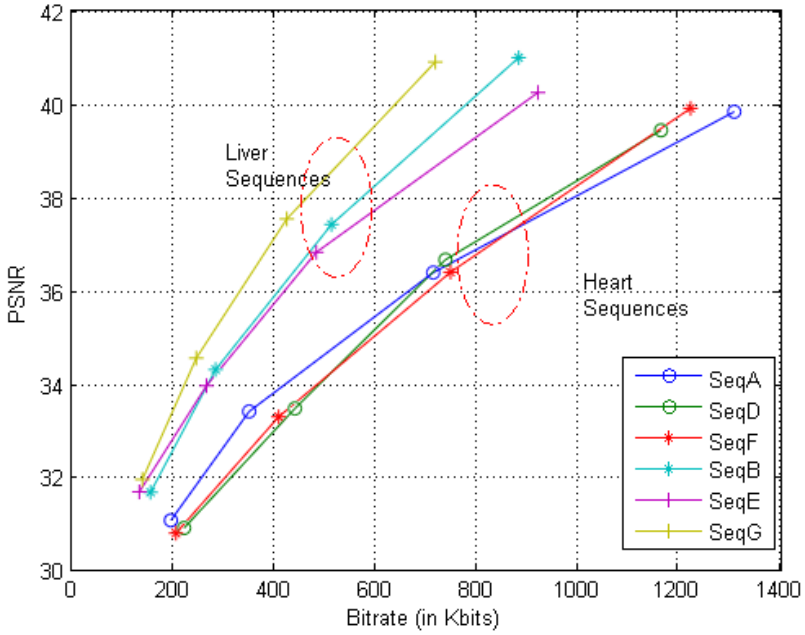

Figure 43 - Rate-Distortion curves for three cardiac and three liver sequences.

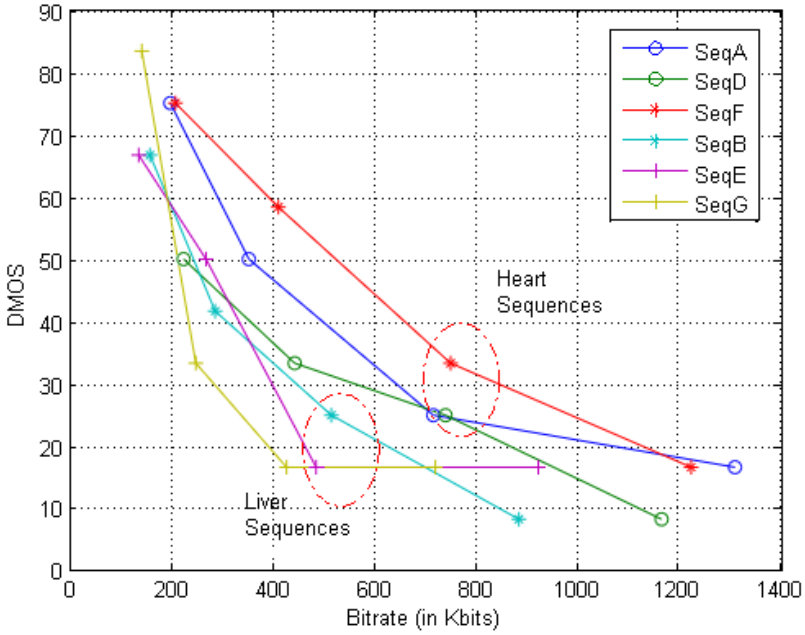

Figure 44 - Rate-Quality curve depicting the variation of DMOS with bitrate for three heart and three liver sequences each.

In Figure 44, considering DMOS = 40 as the threshold for acceptable quality, it can be noted that for the corresponding sequences in Figure 43, the PSNR threshold for liver sequences is approximately 35 − 36dB and for heart sequences approximately around 34 − 37dB. The normally accepted PSNR quality is ≈ 35dB for medical video sequences 12, 13, 14. Based on this criterion, in Figure 43, it can be seen that for the liver sequences HEVC delivers an acceptable PSNR-quality (35 dB) video at bitrates approximately over 300 kbps. For the heart sequences this is approximately over 600 kbps. Therefore, from this result it appears that rate-distortion performance is considerably influenced by the spatio-temporal complexity of the video.

Figure 44 shows the rate-quality curves for three heart and three liver sequences, again for QPs 29, 33, 37, and 41 only. The DMOS values 20, 40, 60, and 80 represent the diagnostic quality levels of the videos at Excellent, Good, Annoying, and Very Annoying range respectively. Using this quality reference level, Figure 44 shows that excellent diagnostic quality video for the considered liver sequences could be obtained at the bitrate range of 400 - 600 kbps and good diagnostic quality videos at ≈ 240 − 360 kbps. Similarly, for the considered heart sequences, excellent diagnostic video quality could be obtained at the bitrate range of 900 - 1200 kbps and good quality heart sequences were obtained with the 380 - 700 kbps range. Further calculation shows that excellent diagnostic quality video sequences were obtained at compression ratios between ≈ 140:1 and 420:1 via HEVC. Again, it should be noted that the spatio-temporal complexities of the sequences can have considerable influence on the compression ratio.

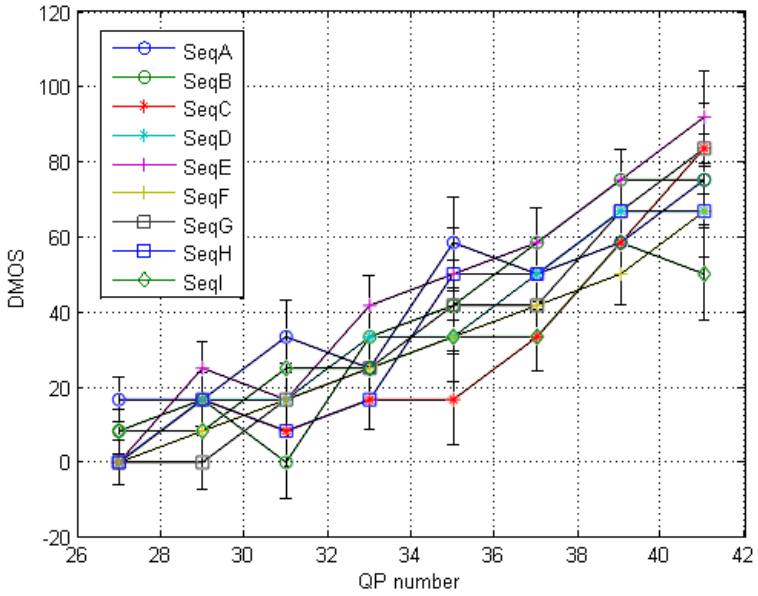

Figure 45 - DMOS of experts vs. QP.

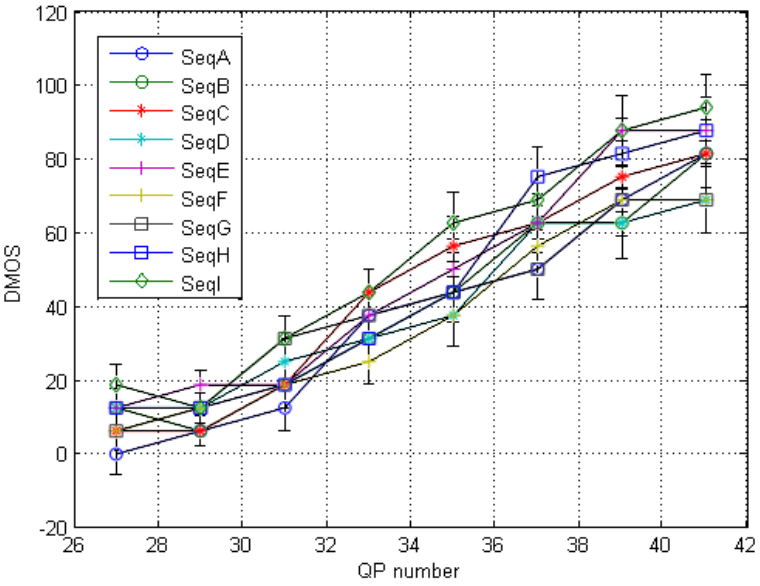

Figure 46 - DMOS of non-experts vs. QP.

Figure 45 and Figure 46 shows the variation of DMOS with respect to the QPs for all the nine sequences considered in the tests with their associated confidence intervals. Since the expert DMOS was obtained from four experts, for a fair comparison, out of the 16 available non-expert scores, only DMOS of four non-expert subjects were randomly selected. It can be seen that the DMOS of non-experts is slightly higher than the expert DMOS across most QPs. This implies

that the non-experts gave lower ratings to the video sequences than the experts. It can also be observed that the DMOS of non-experts is relatively more consistent than expert DMOS across QPs. The relatively higher inconsistency in expert DMOS is possibly because of the different experience levels of individual experts influencing the Differential Opinion Score (DOS).

The variation in DMOS indicates how medical experts and non-experts perceive the quality of compressed medical videos. The relatively lower DMOS range of the experts implies that the experts were able to deduce considerable diagnostic information even from the videos which were perceived to be annoying from non-experts. The expert DMOS also shows that experts with lower experiences are likely to rate the diagnostic quality of the video much lower than the experienced ones.

4.3 Quality Assessment of Medical Ultrasound Videos compressed via H.264

The video sequences mentioned in the previous section were also compressed using H.264/AVC video codec, which is the predecessor of the HEVC video compression standard. The sequences were compressed using the same test conditions described in the previous section, i.e. 8 QP levels, 27, 29, 31, 33, 35, 37, 39, 41.

4.3.1 Subjective Tests & Scores

The subjective evaluation for the H.264 compressed video sequences were done by three medical experts. Following the request of the medical doctors, in this set of tests all the impaired video sequences were displayed simultaneously along with the reference video sequence on multiple displays. The medical experts were then asked to rate the video sequences based on their diagnostic quality on a scale of 1 to 5. Three medical doctors participated in this study. The mean of all the subjective scores for each impaired sequences was computed to obtain the Mean Opinion Score (MOS).

4.3.2 Test Results

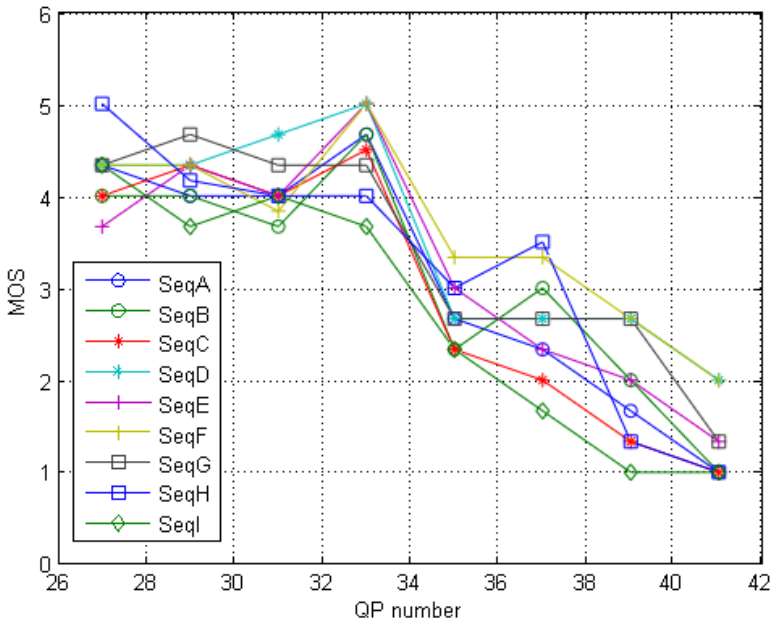

Figure 47 – MOS vs. QP for sequences compressed via H.264.

In Figure 47, the MOS of nine sequences against different QP impairment levels is presented. It can be seen that there is a significant decrease in the MOS score between the QP scores 33 and 35. This implies that the experts found the diagnostic quality of the video to be remarkably influenced approximately for QP values higher than 33 for H.264 compression.

4.4 Quality of Service (QoS) and Quality of Experience (QoE) for HTTP medical adaptive video streaming based on the MPEG-DASH standard

This section presents a study on the Quality of Service (QoS) and Quality of Experience (QoE) for HTTP medical adaptive video streaming based on the MPEG-DASH standard. The QoS is analysed through the Peak-Signal-to-Noise Ratio (PSNR) of the streamed frames, the video throughput, and the rebuffering frequency, whereas the QoE is presented through the Mean Opinion Score (MOS) from subjective tests. The subjective tests have been performed with experts in the hospital of Perugia.

The focus of this study is the QoS and QoE dependency on the encoding and segmentation of the DASH dataset. The encoding is controlled by setting the Group of Picture (GOP) size (G) and the segmentation is performed by using different segment size (T). The presented results show the QoS and QoE for HTTP medical video streaming when using different GOP and segment sizes.

4.4.1 Evaluation setup

The evaluation setup comprises the encoding and segmentation of the video content and the adaptation strategy for the HTTP medical video streaming. Based on the output of the setup, an impaired medical video sequences were generated and used for the subjective tests.

The settings of the encoding and segmentation of the cardiac ultrasound video sequence, in the following "US", are as follows. The GOP sizes are set as: $G = T$, $G = T/2$, and $G = T/4$, where the segment sizes is $T = 1$; $T = 2$; $T = 4$; $T = 8$, expressed in seconds (s). In addition, the following five representation levels were used (200; 400; 800; 1000; 1500), expressed in (kbps). The DASHEncoder [9] was used for the encoding and segmentation. The US sequence has resolution 636x240 pixels, duration 32 seconds, and the encoding frame rate is set at 25 frames per second (fps).

The adaptation at the client side considers the network throughput at the decision instance τ (kbps), the segment size T (s), the buffer state B_{state} expressed as number of buffered frames, which depends on the playout rate F_r (fps), the current representation level ρ^n (kbps) before the decision, and the available representation set $\{\rho^i, i = 1,2,..,M\}$ (kbps) composed of the representation levels, where we adopt $\rho^i < \rho^{i+1}$. The output of the algorithm is the representation level selected at the client ρ^* for the following segment, therefore $\rho^{n+1} = \rho^*$. The algorithm considers two types of triggers used to monitor the changes in the buffer and in the network, i.e., a buffer trigger (BT) and a throughput trigger (TT). The possible values for the triggers are 0 and 1, where 0 is associated to a decrease of the network throughput and of the number of buffered frames, and 1 is associated to an increase of the network throughput and of the number of buffered frames compared to the current condition. If both triggers have value 1, and the current representation level is $\rho^n = \rho^i$, then the algorithm chooses to switch to a higher representation level ($\rho^* = \rho^{i+1}$), i.e., $\rho^{n+1} = \rho^{i+1}$. If both triggers have value 0, then switching to a lower representation level is performed ($\rho^* = \rho^{i-1}$), i.e., $\rho^{n+1} = \rho^{i-1}$. In the end, if one trigger has value 1, and the other has value 0, then the same representation level is kept $\rho^* = \rho^i$), i.e., $\rho^{n+1} = \rho^i$. The triggers are calculated based on (1) and (2) :

$$BT = \begin{cases} 0, B_{state} \le NT \dfrac{\rho^{\mathrm{n}}}{\rho^1} \\ 1, B_{state} > NT \dfrac{\rho^{\mathrm{n}}}{\rho^1} \end{cases} \quad (1)$$

$$TT = \begin{cases} 0, \tau \le \rho \\ 1, \tau > \rho \end{cases} \quad (2)$$

The buffer trigger is used to monitor the depletion rate of the buffer and it is used to mitigate rebuffering, whereas the throughput trigger is used to adapt the switching according to the variable throughput in the network.

4.4.2 Evaluation results

The emulation of the adaptive video streaming session is based on a channel with a variable (modelled) bandwidth capacity, where the bandwidth model is based on [10]. The definition of the observed QoS metrics is as follows:

- PSNR of the streamed frames: The PSNR of the streamed frames at the client.

- Video bitrate: Based on the switching between the representation levels, we consider the bitrate of the adaptive video streaming session.
- Rebuffering frequency: We consider the percentage of rebuffered frames at the client, where a frame that is not downloaded fully by the time it needs to be played out at the client side is a rebuffered frame.

The MOS of the HTTP adaptive video streaming is obtained through subjective tests with 3 experts (cardiologists in the hospital of Perugia).

QoS metrics

The GOP size and the segment size used in the content encoding and segmentation have a significant effect on the PSNR. As shown in [9] smaller segment sizes decrease the PSNR values, and by using GOP sizes smaller than the segment size ($G = T/2$; $G = T/4$), the overall PSNR is further reduced. The usage of smaller segment sizes increases the number of potential switching points between different representations, thus the adaptation can be performed more efficiently. This leads to improvements in the overall PSNR and bitrate, as shown in Figure 48 for the extreme cases (G = T and G = T=4). Although from the encoding point of view smaller segment sizes decrease the average PSNR of the segment, the higher number of switching points provides better adaptation to the fast bandwidth changes. This allows the client to select representations with higher quality more frequently.

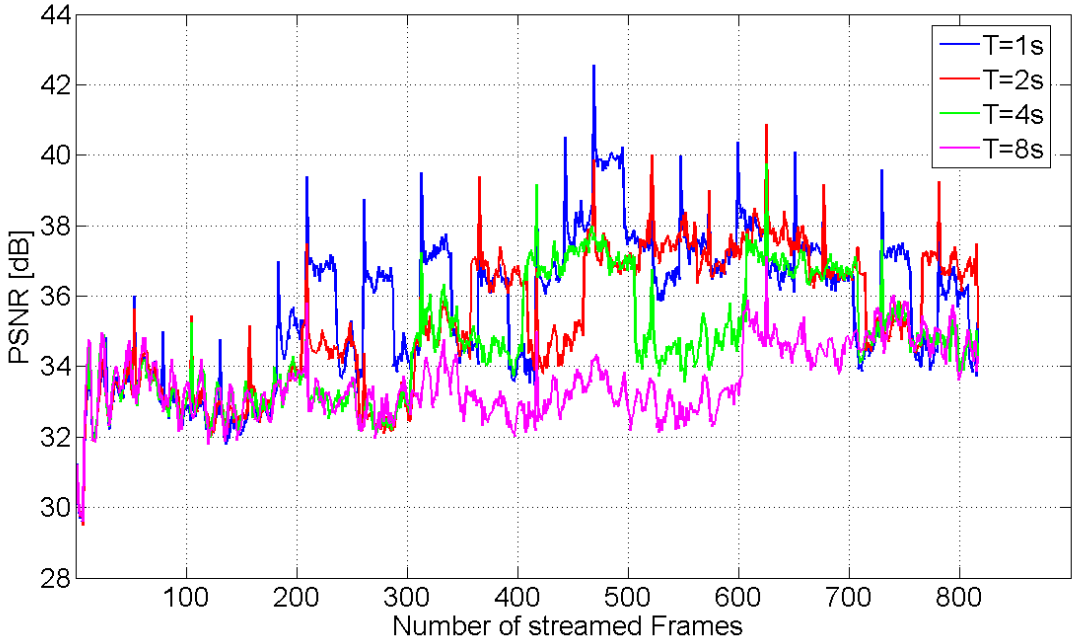

Figure 48 - PSNR of the streamed frames when using different segment sizes and $G=T$.

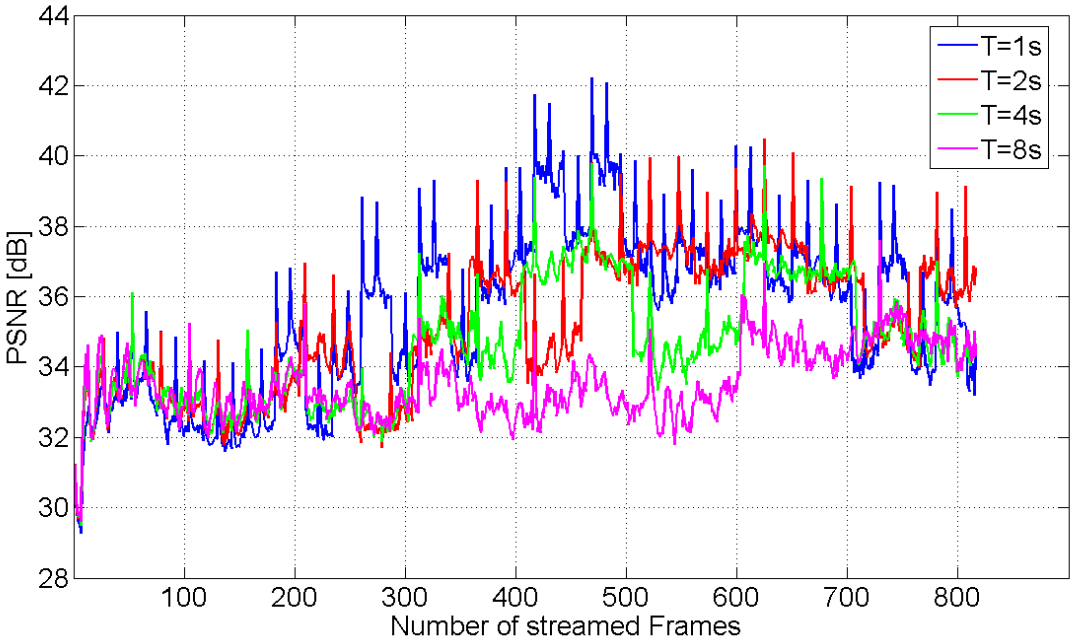

Figure 49 - PSNR of the streamed frames when using different segment sizes and *G*=*T*/4.

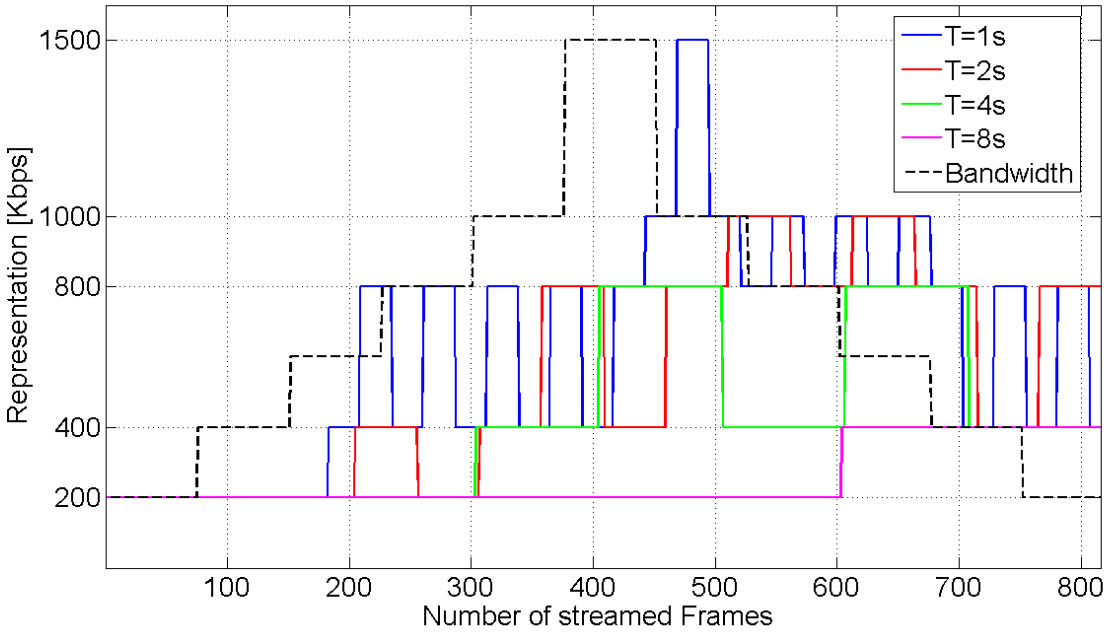

Figure 50 - Bitrate of adaptive video streaming when using different segment sizes and *G*=*T*.

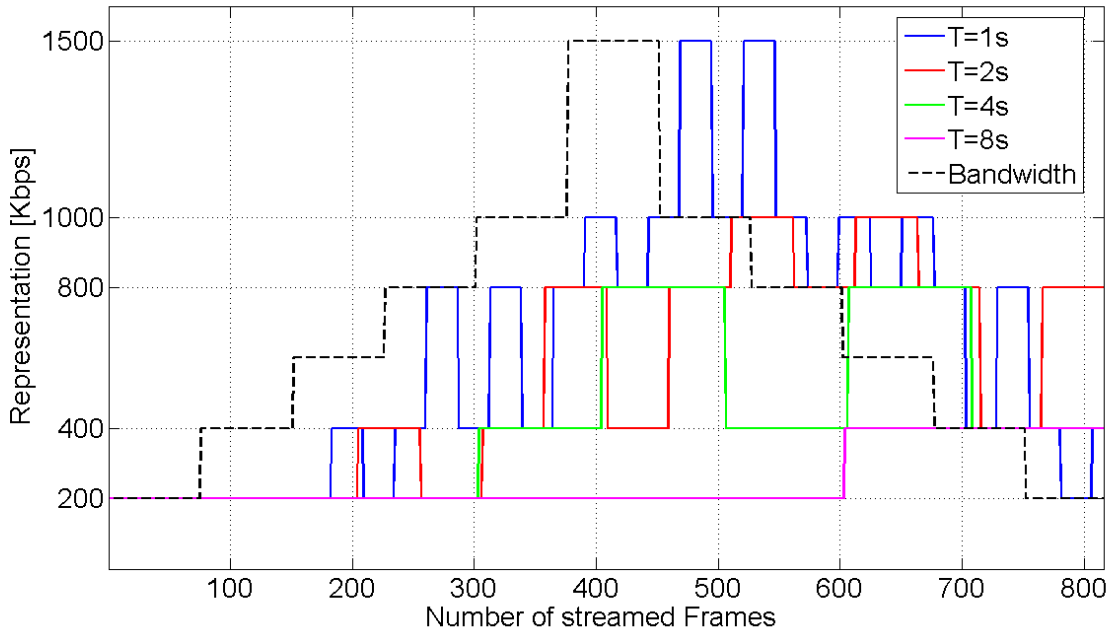

Figure 51 - Bitrate of adaptive video streaming when using different segment sizes and *G*=*T*/4.

The averages of the PSNR and the video bitrate, as well as the rebuffering frequency of the medical video streaming with different segment sizes and GOP sizes are shown in Table 1.

Table 1. QoS metrics for adaptive HTTP medical video streaming

	T=1s			*T*=2s			*T*=4s			*T*=8s		
	G=*T*	*G*=*T*/2	*G*=*T*/4	*G*=*T*	*G*=*T*/2	*G*=*T*/4	*G*=*T*	*G*=*T*/2	*G*=*T*/4	*G*=*T*	*G*=*T*/2	*G*=*T*/4
PSNR [dB]	35.64	35.45	34.84	35.26	35.12	34.84	34.68	34.61	34.84	33.53	33.50	33.42
Bitrate [kbps]	626.2	648.5	629.4	537.5	537.5	537.5	424.7	424.7	424.7	252.2	252.2	252.2
Rebuff. freq. [%]	0.85	0.98	1.22	2.45	1.34	0.85	2.45	1.71	0.98	4.16	4.28	2.57

The more efficient adaptation with the smaller segments provides better performances when compared to the case with larger segments. Furthermore, we observe that the ratio between the GOP size and the segment size has insignificant effect on the video bitrate, but it can impact the PSNR. The results for the rebuffering frequency show that smaller segment sizes reduce the frequency of rebuffering due to more efficient adaptation to the bandwidth changes. Furthermore, we observe that smaller GOP size with respect to the segment size decreases the rebuffering frequency only for the large segment sizes. These results were obtained with no initial buffering at the client. However, in practice, the clients wait a certain time for the initial buffering of the video before starting the playout. The initial buffering is useful approach to smooth out the video playout by reducing the rebuffering. Figure 52 shows these effects, where initial buffering at the client is performed until smooth streaming, i.e., the video streaming completes without rebuffering. The required initial buffering to achieve smooth streaming is decreased for smaller segment sizes. Further, reducing the GOP size for the small segment sizes is not an efficient approach, however in the case of large segments, reducing the GOP size decreases the required initial buffering for smooth streaming.

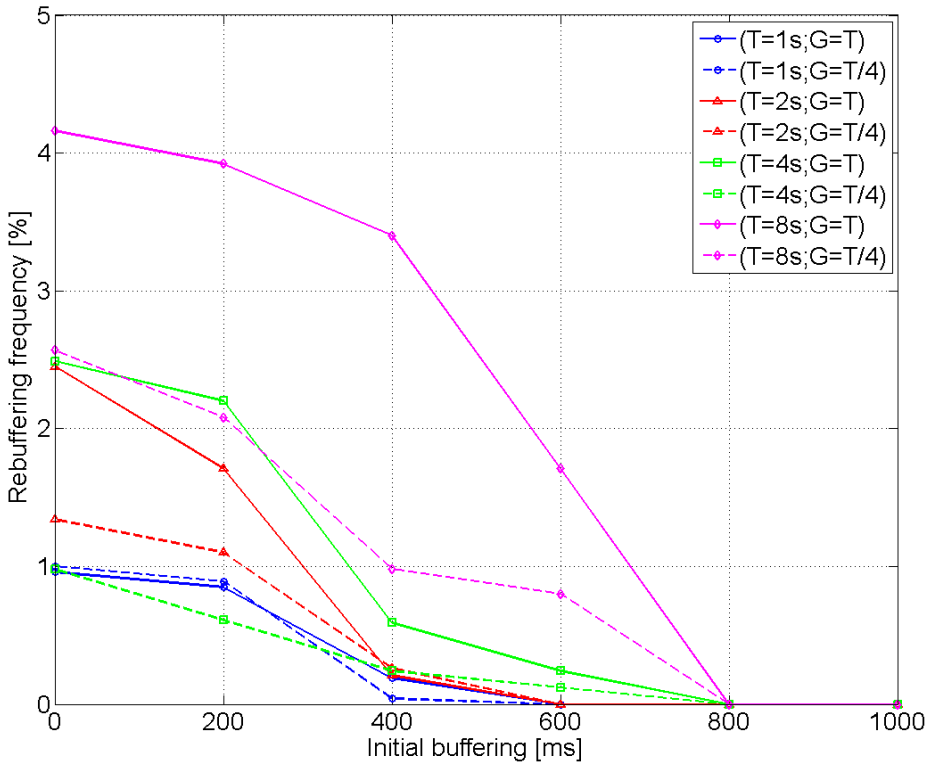

Figure 52 - The impact of the initial buffering on the rebuffering frequency when using different segment sizes and *G*=*T*/4.

QoE for adaptive HTTP medical video streaming

The video sequences were subjectively evaluated in terms of diagnostic quality by medical experts who provided their opinion scores on a scale of a specified range (1-5). The subjective evaluation was done using the Double Stimulus Continuous Quality Scale (DSCQS) type II [8]. The evaluation was performed in a room that the specialists use to visualize video sequences and perform diagnosis accordingly.

The results from the subjective tests are shown in Figure 53, where we observe the impact of different GOP sizes and segment sizes on the MOS score at the end of the streaming. Large segment sizes in the considered setup reduce the MOS due to the inefficient adaptation to the changing bandwidth. Hence, higher MOS is obtained for T = 1s and T = 2s regardless of the video sequence, therefore these segment sizes are preferable for medical video streaming in the considered setup. Further, we observe that reducing the GOP size with respect to the segment size decreases the MOS regardless of the segment. Therefore, in the considered setup it is preferable to set *G*=*T*.

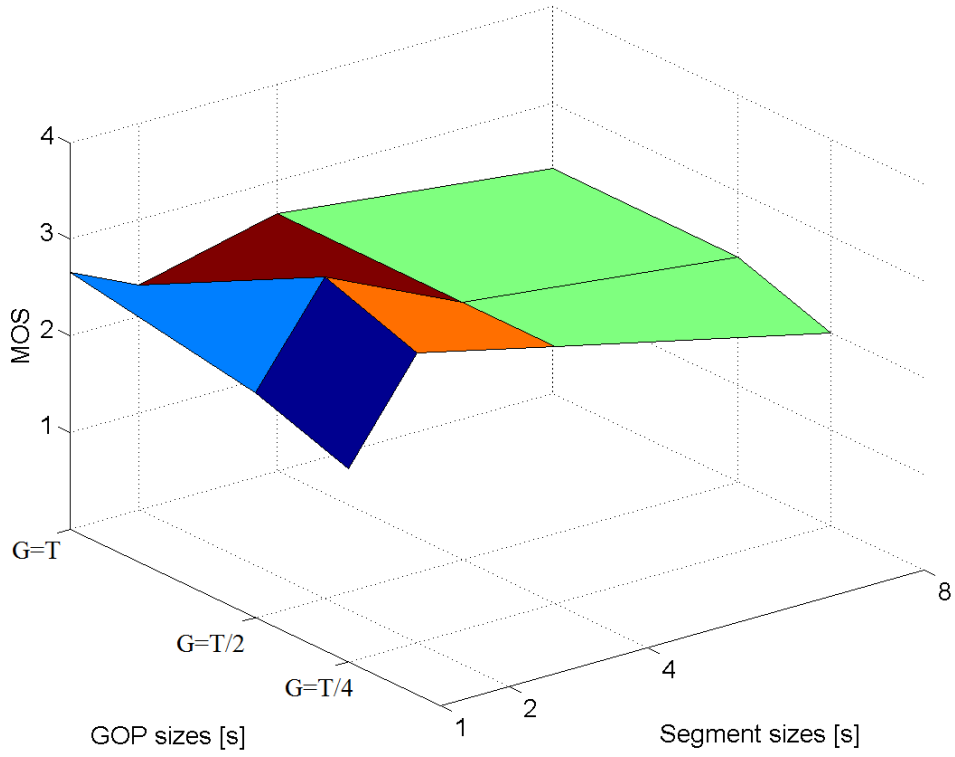

Figure 53 - The MOS scores when using different GOP sizes and segment size.

By adopting the setting T=G=1s, we performed additional subjective tests for two adaptive video streams, denoted as AS1 and AS2. The same rebuffering impairments were inserted in both streams, where the rebuffering frequency ranges from null to 10%. The capacity model in [10] was used for AS1, whereas a trace from mobile HTTP throughput measurements [11] was used for AS2. The AS1 is obtained with a high number of switches between the representations (corresponding to the capacity), whereas AS2 is obtained with a smaller number of switches between the representations, in such way that the average bitrate of both streams is approximately the same. The MOS scores for both streams are shown in Figure 54. We observe that high number of switches (case AS1) reduces the MOS. In this case, even 2% rebuffering decreases the MOS below the acceptable level (MOS=3). Small number of switches (case AS2) results in acceptable MOS even for rebuffering up to 10%.

In the end, we conducted subjective tests to investigate the acceptable MOS for fixed bitrate streaming in the same rebuffering frequency range. The idea is to analyse the thresholds for the required minimal bitrate and maximum rebuffering frequency resulting in acceptable MOS. These inputs are also relevant for tailoring the adaptation strategy during adaptive streaming by controlling the number of switches and targeting an average bitrate and rebuffering frequency. The results are shown in Table 1. Fixed bitrate streaming with minimum 400 kbps bitrate produces acceptable MOS in the entire rebuffering frequency range. In addition there is no need to increase the bitrate above 800 kbps since the MOS is not significantly affected.

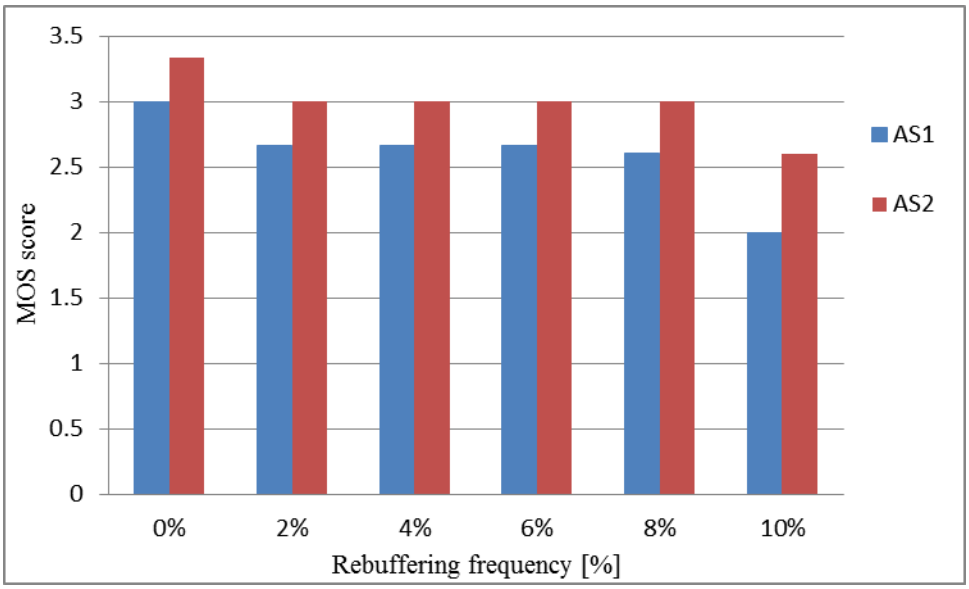

Figure 54 - MOS scores for adaptive video streaming.

Table 2. MOS scores for fixed bitrate video streaming.

	Rebuffering frequency					
Bitrate [kbps]	0%	2%	4%	6%	8%	10%
200	2.3	2	2	2.3	2	2.3
400	3.3	3.6	3.6	3.6	3.6	3.6
800	4.3	4.3	4.3	4	4	4
1000	3.3	4	4	4	4	4
1500	4	4	4	4.3	4.3	4.3

4.5 Medical Quality of Experience results for the sequences acquired during the demo session at Perugia hospital (emergency scenario)

This section reports the results of the subjective tests performed with the medical doctors during the demo session in the hospital of Perugia.

Different types of subjective scores have been acquired: global opinion scores on the medical and ambient video sequences and real time measurements acquired via a tablet with the Android application developed by Kingston University.

4.5.1 Real time subjective quality acquisition via Android-based application

As highlighted in previous WP3 deliverables, there are different methodologies to perform subjective quality assessment tests. In the first part of the project, the focus was on subjective tests where a single score was provided for each sequence at the end of its visualisation. We mainly used three types of tests: Single Stimulus Continuous Quality Scale, Double Stimulus Continuous Quality Scale tests and Multiple Stimulus Continuous Quality Scale tests (as requested by the medical doctors in the hospital). In Single Stimulus Continuous Quality Scale tests only the impaired sequence is visualised and a score is provided at the end of the visualisation of the sequence. In Double Stimulus Continuous Quality Scale tests each sequence is compared with the corresponding original one and a score is provided for both the impaired sequence and the original one. The subject (medical doctor in our case) is not aware of which sequence is the original. The difference between the score assigned to the original sequence and to the impaired one is considered. In the Multiple Stimulus Continuous Quality Score tests requested by the medical doctors all the impaired versions were shown to the medical doctors along with the original sequence in multiple displays and the medical

doctors provided scores for all the sequences, including the original (they were not aware of the position of the original sequence).
All these methodologies provide a general indication of the quality of the sequence, but in the case of large variations in the quality, as in the case where rate control or dynamic adaptive streaming is used and where transmission is over a highly time-varying transmission channel (such as the wireless channel), it is important to assess also the quality variation of the received video sequence over time. For this reason, in the last part of the project also real-time quality evaluations have been done, thanks to the Android application developed in Kingston University (see Deliverable D6.4, Section 2.3.4 [5]).

The quality meter is developed as an Android application with a graphical interface, enabling to evaluate the quality in a categorical scale (1 to 5) using a slider. Figure 55 shows a snapshot of the graphical user interface of the developed Android App.

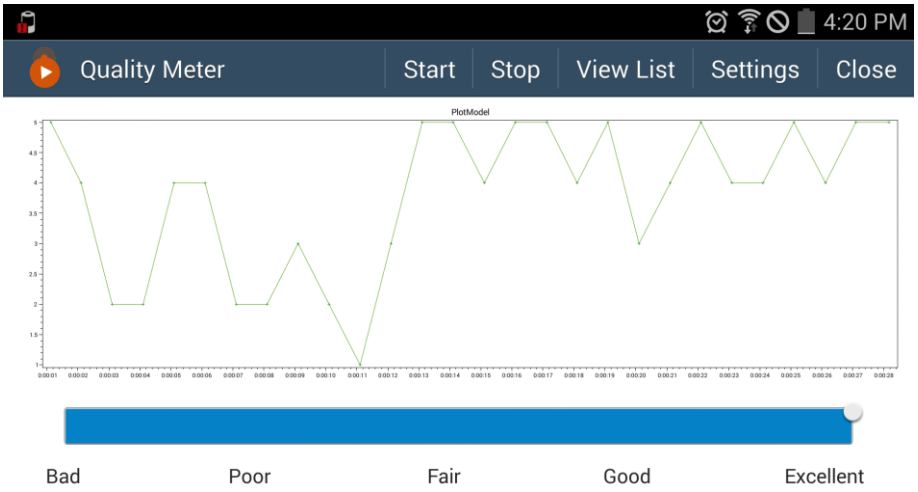

Figure 55 - Graphical user interface of the real-time quality meter.

The Opinion Scores (OS) from the users can be stored locally or transmitted to a remote server for further analysis. Figure 56 shows a medical doctor in the hospital providing her real time opinion score about a cardiac video sequence. Figure 57 shows part of a file stored in the server after the demo tests in the Perugia hospital. The file reports the opinion scores as well as the date and time of the acquisition and the name of the medical evaluator as file name.

Figure 56 – Real time acquisition of subjective scores via Android application from medical doctor at Hospital of Perugia.

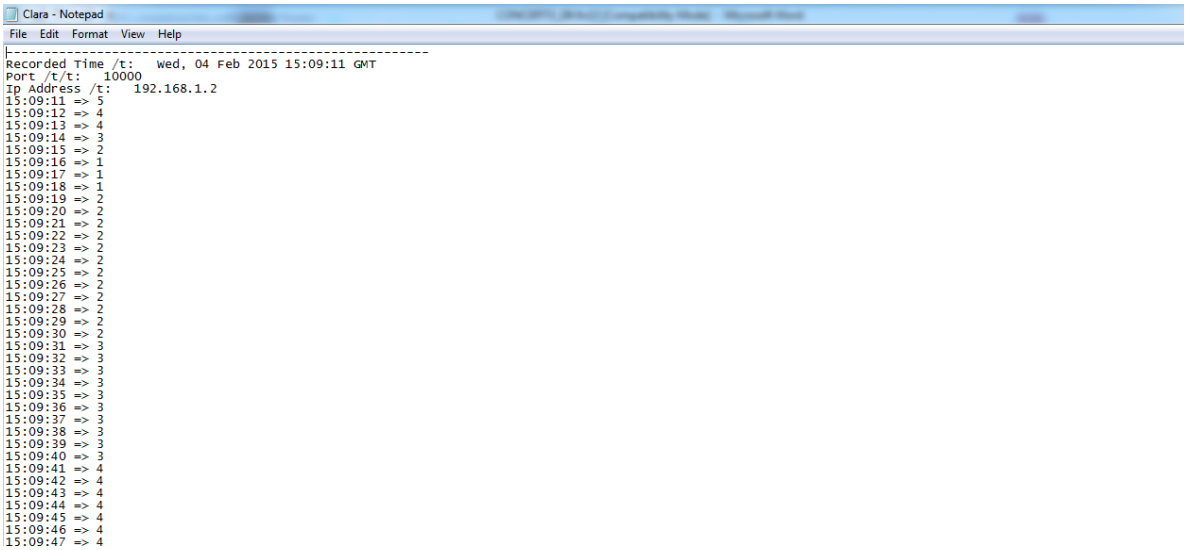

Figure 57 – Example text file stored in the Server.

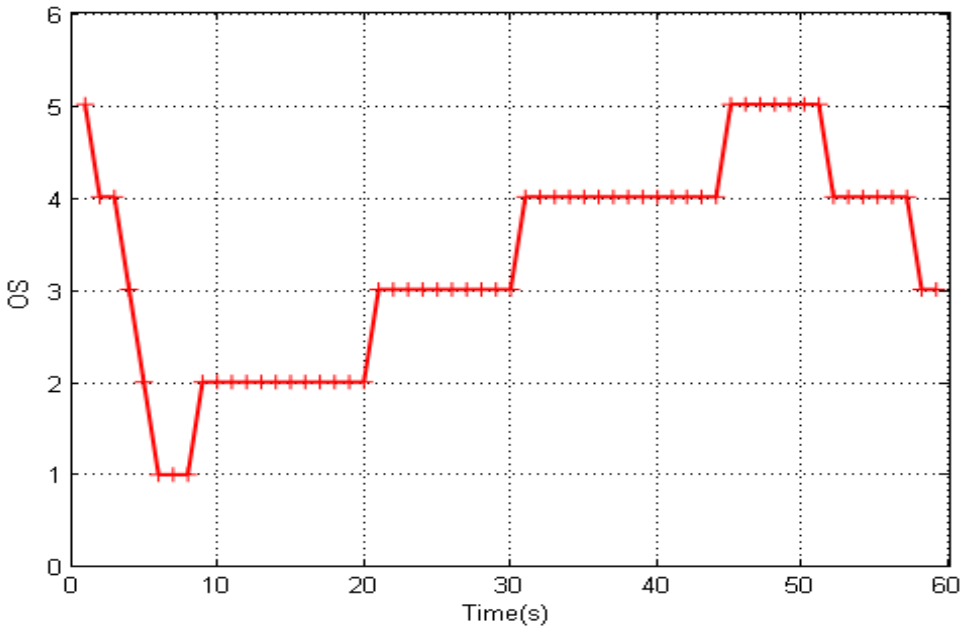

Figure 58 – Real time subjective quality results (OS vs. time) for ultrasound video sequence received from ambulance in emergency scenario.

Figure 58 reports the plot of the scores provided by the medical doctor in the hospital for approximately the first minute (59s) of the streaming session during the demo session in February 2015 (corresponding to the file in Figure 57).
We can observe in Figure 58 that the initial quality was bad (the medical doctor moved the cursor, initialised at excellent quality, to the score corresponding to bad quality, i.e. 1) since the probe was not yet on appropriate position on the patient's body at the beginning of the streaming session. Then the sonographer in the ambulance started searching for a significant section of the heart to display (initially the diagnostic quality was "fair") and then identified a section whose information was relevant for the medical doctor in the hospital. At this point, the quality varied from 3 to 5 (good to excellent) depending on the quality of the LTE network and the complexity of the ultrasound video content transmitted.
The percentage of time where the quality was satisfactory for an emergency diagnosis is 78% after the start of the actual examination. The medical doctors highlighted that the quality was adequate (score 3 in OS scale) for them to perform a diagnosis in emergency. They also commented that instantaneous reductions of quality do not affect the diagnosis as long as this can be compensated by cardiac cycles with good quality.

Figure 59 reports a longer evaluation performed by another cardiologist in the Hospital of Perugia on one of the sequences received during the demo session in Perugia (sequence also considered in Section 3.4). The evaluation was performed on more than five minutes video (5min 20s). The subjective quality ranges from excellent to bad and was globally assessed as adequate for an emergency scenario by the medical specialist. The medical doctor reiterated that, although the quality was bad at some points, the cardiac cycles with good quality enabled a proper emergency diagnosis.

The quality of the received ambient video acquired from the two different cameras in the ambulance is reported in Figure 60 and Figure 61 in terms of subjective scores over time. The quality was evaluated globally as "very good" for the purpose of checking the conditions of the patient in the ambulance.
The subjective results hence confirm that, even in the conditions during the demo session (ambulance located between two buildings and bad weather conditions), the CONCERTO solution provided an adequate diagnostic quality for both ultrasound and ambient video.

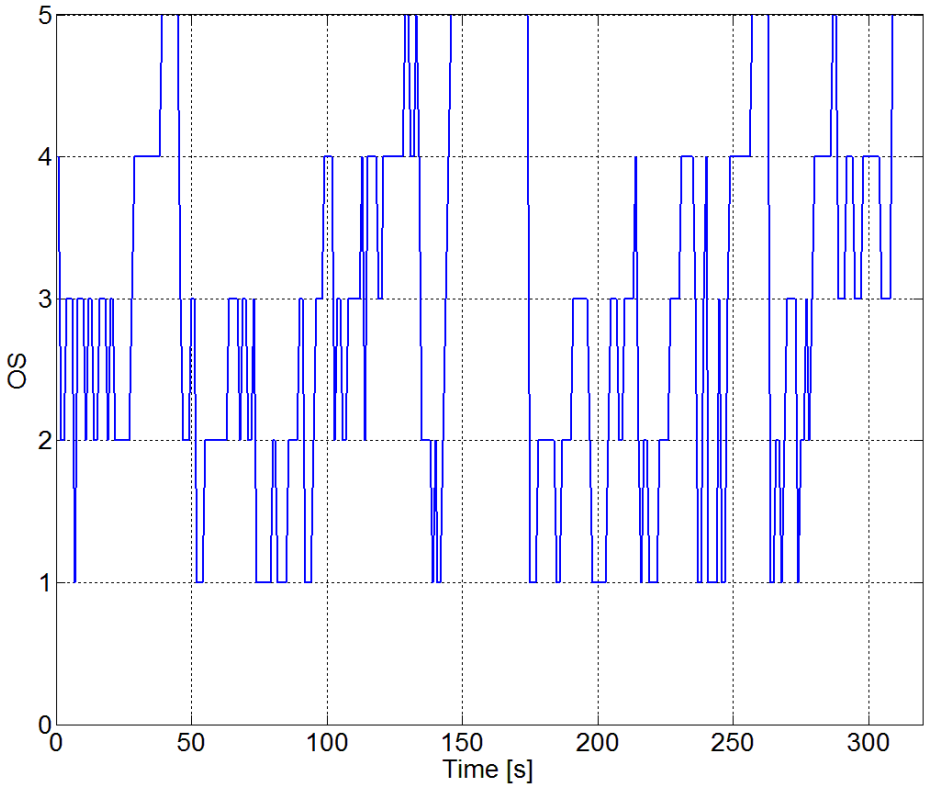

Figure 59 – Opinion scores for received ultrasound video.

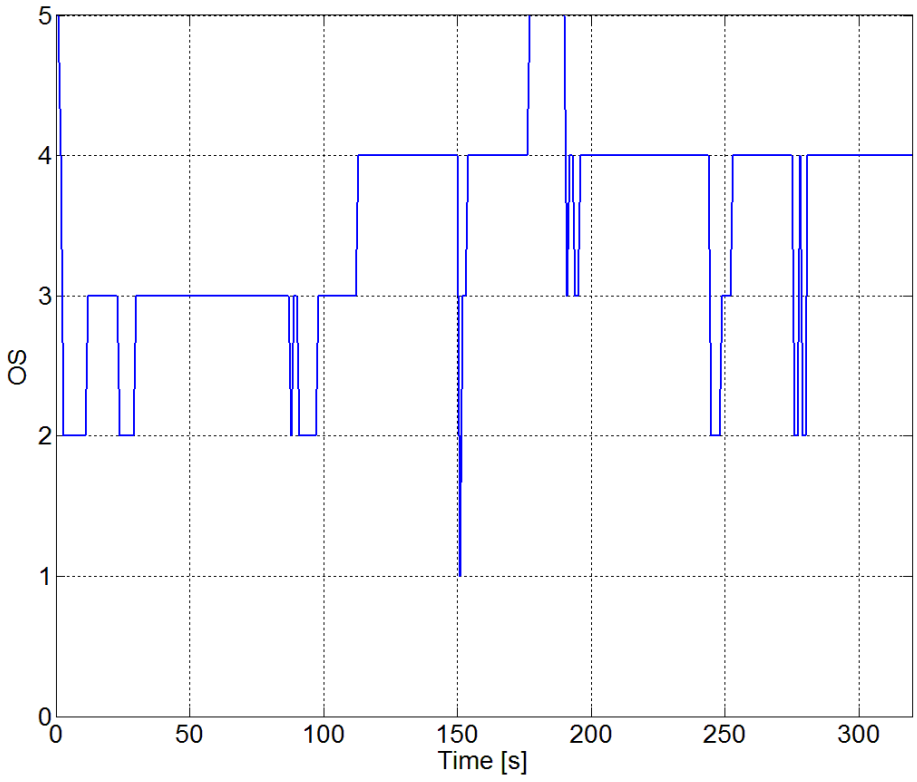

Figure 60 – Opinion scores for ambient video from one of the two cameras (A) in the ambulance.

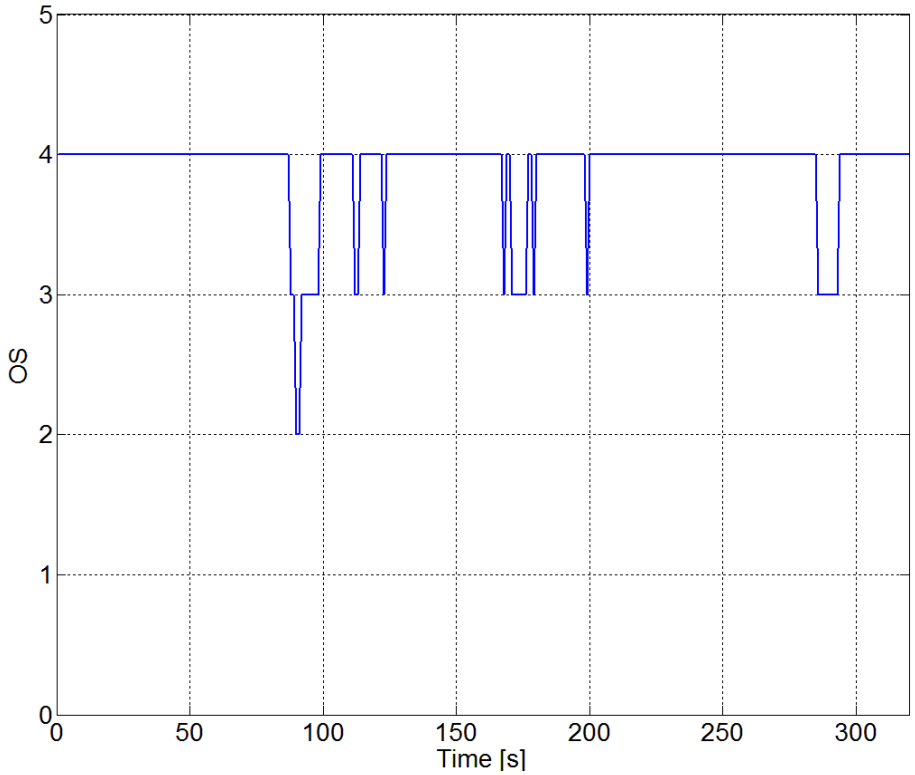

Figure 61 - Opinion scores for ambient video from one of the two cameras (B) in the ambulance.

5 Conclusions

This deliverable presented the CONCERTO final system validation results, which have been obtained through an extensive tests campaign, both using the system simulator developed during the project (full results obtained via the simulator were presented in [4]) and via measurements in a real environment performed via the final implementation of the CONCERTO wireless multimedia platform, developed within the project in order to demonstrate the CONCERTO solution for different scenarios on a real-time implementation. Besides performance indicators like throughput, packet loss ratio, delay, importance has been given to video quality, reporting several curves referring to objective video quality metrics, allowing to directly evaluate the end-to-end improvements due to the techniques developed within the CONCERTO project and to compare them with more traditional transmission schemes.

The quality of the received video sequences has been measured in particular in terms of PSNR and SSIM. Finally, subjective video quality tests have been performed on a subset of videos in order to complete this evaluation. The latter tests required the involvement of a number of medical doctors in the Hospital of Perugia.

The advantages of the CONCERTO solutions are evident. The results obtained with the simulator show that the improvement obtained enables achieving a video quality that is satisfactory for the medical doctors to perform a diagnosis, remarkably improving on the benchmark solution.

The results obtained with the demonstrator show that, even considering unfavourable radio channel conditions (as those during the demo for the ambulance scenario in the Hospital of Perugia, given the specific location of the ambulance and weather conditions), the introduction of the CONCERTO solution enabled a quality that was assessed as “good” by the medical doctors. At the opposite, the benchmark scheme (i.e., without CONCERTO’s optimizations) provided a really poor received quality: the received videos were so corrupted that subjective quality evaluations are meaningless on them. The demonstrator results also reveal the significance of HTTP (MPEG-DASH) streaming in order to provide reliable video transmission with slightly longer delay compared to RTP/RTSP streaming.

6 References

[1] CONCERTO deliverable D2.1 - Use cases, usage models and related requirements, November 2012.

[2] CONCERTO deliverable D3.3 - Final description and evaluation of solutions for QoE-aware image/video coding, July 2014.

[3] CONCERTO deliverable D6.2 - Specification of the demonstrator, August 2013.

[4] CONCERTO deliverable D6.3 - CONCERTO Simulator: Final Architecture and Numerical Result, December 2014.

[5] CONCERTO deliverable D6.4 - Demonstrator description and validation plan, December 2014.

[6] E. A. Ayele and S. Dhok, "Review of proposed high efficiency video coding (HEVC) standard", *International Journal of Computer Applications*, vol. 59, no. 15, pp. 1-9, 2012.

[7] M. Razaak, M.G. Martini and K. Savino, "A Study on Quality Assessment for Medical Ultrasound Video Compressed via HEVC,"*IEEE Journal of Biomedical and Health Informatics (J-BHI)*, vol. 18, no. 5, pp. 1552-1559, Sep 2014.

[8] M. H. Pinson, and S. Wolf, "Comparing subjective video quality testing methodologies", *SPIE Visual Communications and Image Processing*, pp. 573–582, 2003.

[9] S. Lederer, C. Müller and C. Timmerer, "Dynamic adaptive streaming over HTTP dataset," *Proceedings of the 3rd Multimedia Systems Conference*, pp. 89-94, February 2012.

[10] F. De Simone and F. Dufaux, "Comparison of DASH adaptation strategies based on bitrate and quality signalling," *15th International Workshop on Multimedia Signal Processing*, pp. 87-92, September 2013.

[11] R. Haakon, P. Vigmostad, C. Griwodz and P. Halvorsen, "Commute path bandwidth traces from 3G networks: analysis and applications," *Proceedings of the 4th ACM Multimedia Systems Conference*, pp. 114-118, February 2013.

[12] Y. A. Syahbana, A. Herman, A. A. Rahman, and K. A. Bakar. "Aligned-PSNR (APSNR) for Objective Video Quality Measurement (VQM) in video stream over wireless and mobile network." In *World Congress on, Information and Communication Technologies (WICT)*, pp. 330-335, 2011.

7 Glossary

AHP	Analytic hierarchy process	Decision making system
AVC	Advanced Video Coding	H.264 video compression standard
BAN	Body Area Network	
BSN	Body Sensor Network	
CCD	Charge coupled device	
CDM	Context Discovery Module	
DASH	Dynamic Adaptive Streaming over HTTP	
DDE	Distributed Decision Engine	Event dissemination framework
DMOS	Differential Mean Opinion Score	
DSCQS	Double Stimulus Continuous Quality Scale	Methodology for subjective quality assessment
GOP	Group of Pictures	
GPAC	GPAC Project on Advanced Content	gpac.wp.mines-telecom.fr
GPS	Global positioning system	
HD	High Definition	
HEVC	High Efficiency Video Coding	
HTTP	Hypertext Transfer Protocol	
IPv6	Internet Protocol version 6	
ITU	International telecommunications Union	
JCT-VC	the Joint Collaborative Team on Video Coding	
JND	Just Noticeable Difference	
JPEG	Joint photography expert group	Image compression standard
JSCC	Joint Source and Channel Coding	
LTE	Long term evolution	3GPP standard
MAC	Media Access Control	
MAP	Maximum A Posteriori	
MOS	Mean Opinion Score	
MPEG	Moving Picture Experts Group	
MSU	Moscow State University	
NIS	Network Information Service	
OS	Opinion Score	
PSNR	Peak Signal-to-Noise Ratio	
QoE	Quality of Experience	
QoS	Quality of Service	
RTP	Real-time Transport Protocol	
RTSP	Real Time Streaming Protocol	
SSCQS	Single Stimulus Continuous Quality Scale	Methodology for subjective quality assessment
SSID	Service Set Identifier	
SSIM	Structural Similarity	
VQA	Video Quality Assessment	
WLAN	Wireless local area network	Wi-Fi is a commercial name for a WLAN realization.