

Routine tests for both planning and evaluating image quality in tele-echocardiography

Sandra Morelli^(a), Andrea Giordano^(b) and Daniele Giansanti^(a)

^(a)Dipartimento di Tecnologie e Salute, Istituto Superiore di Sanità, Rome, Italy

^(b)Dipartimento di Bioingegneria, Fondazione Salvatore Maugeri, IRCCS, Veruno (NO), Italy

Summary. Both in real-time and “store & forward” tele-echocardiography (T-E), a coding process has to be applied to the echocardiographic videoclips in order to limit the bandwidth needed and adapt it to the bandwidths furnished by network providers. The compression process degrades the videoclips, affecting thus the quality of the videoclips and potentially compromising the diagnostic accuracy of the T-E. In this work the authors investigated on the use of automatic tools for the video quality assessment by means of objective methods with particular care to the role of the system administrator. As the use of tests on video quality assessment (based on subjective methods) is hampered by the high number of needed resources (persons, laboratories and time). The use of valid objective methods is thus desirable. The study reviewed different tools with this specific aim. One of the more suitable tool was found to be represented by a software package designed by the Institute of Telecommunication Sciences and the National Telecommunication and Information Administration, the NTIA/ITS VQM tool. This tool gives back objective-quantitative data as outcomes, however embeds models emulating the subjective perception. This study reviewed and analyzed in depth the functionalities of the tool to improve the image quality in TE over the network. The tool was also found suitable for a more general process of T-E assessment, from a health technology assessment (HTA) perspective.

Key words: tele-echocardiography, video quality assessment, video quality metric, health technology assessment.

Riassunto (*Test di routine per la pianificazione e la valutazione della qualità delle immagini in tele-ecocardiografia*). In tele-ecocardiografia (T-E), sia quella in modalità di trasmissione “real-time” che “store & forward”, deve essere applicato un processo di codifica ai videoclip ecocardiografici provenienti dall’elettrocardiografo, per limitare la banda necessaria alla trasmissione e per adattarsi così alle bande fornite dai fornitori di servizi di rete. Il processo di compressione degrada i videoclip, influenzando quindi la qualità di essi e compromettendo potenzialmente l’accuratezza diagnostica della T-E. Lo studio ha investigato l’uso di strumenti informatici automatici per la valutazione dei videoclip ottenuta tramite metodi oggettivi con particolare attenzione al ruolo dell’amministratore di sistema. Dato che l’uso di test basati sulla valutazione umana è limitato dal numero elevato di risorse disponibili (persone, laboratori, tempo) l’uso di metodi oggettivi è pertanto desiderabile. Lo studio ha rivisitato diversi strumenti con questo obiettivo. Uno degli strumenti più adatti è stato giudicato il pacchetto software disegnato dall’ Institute of Telecommunication Sciences and the National Telecommunication and Information Administration (USA), ossia il “NTIA/ITS VQM tool”. Lo studio ha analizzato con successo le funzionalità dello strumento con particolare riferimento al miglioramento della qualità delle immagini in T-E. Lo strumento è stato ritenuto anche adatto ad essere utilizzato in un più generale processo di valutazione del sistema di T-E secondo una prospettiva tipica dell’*health technology assessment*.

Parole chiave: tele-ecocardiografia, valutazione della qualità video, metrica della qualità video, *health technology assessment*.

INTRODUCTION

The tele-echocardiography (T-E) consists in the transmission of the echocardiographic video sequences, from the echocardiograph to a remote computer. Today, the most common applications of T-E are performed with one of the two different transmission services: “store & forward” and “real-time” [1-3], depending on the clinical setting. The T-E application

is based also on videoconferencing equipment used to transmit videoclips from the echocardiographic equipment to remote sites for diagnosis. The employment of T-E, especially the real-time T-E, implies the compression of the original videoclips obtained from the echocardiograph with consequent possible degradation of the image quality, that could affect diagnostic accuracy.

Problem definition

The evaluation of the image quality is basic for the assessment of the diagnostic accuracy and preliminary to every qualification procedure to introduce a T-E application in the National Health Care System (NHCS). In [4] it has been shown the need of a multi-face protocol to assess the diagnostic accuracy, based not only on purely quantitative evaluation but also on subjective evaluations. In fact the diagnostic accuracy was also found as a function of the neural internal models of the experts in T-E [4]. It has thus been introduced a protocol [4] based on: 1) simulations of the T-E systems; 2) a phantom based analysis; 3) mostly used quantitative parameters; 4) receiver operator characteristics (ROC). The protocol showed to run well and covered the exigencies of the image quality assessment in T-E. However its use in routine applications showed limits caused by high time and costs in the tests. In any case, in order to assure a good diagnostic accuracy of a tele-echocardiographic video sequence, one should evaluate the quality of the transmitted video, generally corrupted by the compression process in relation to the bandwidth of the transmission system and/or the transmission system conditions, such as noise and traffic.

The quality of a video is generally assessed by means of objective methods (quantitative measurements of selected quality parameters) and/or subjective methods (qualitative evaluation of the quality of the video). As it is well known, the subjective evaluation represents the gold standard for the video quality assessment, but the use of this approach results very complex and requires large amount of human resources and time to carry out the tests. This implies different laboratories where to carry out the tests for the evaluation of the video quality; video instrumentation of high quality, such as professional monitors, mixer, video players; a considerable number of subjects for the quality evaluation. Beside subjective assessment, different objective assessment methods were proposed for measuring the digital video quality, based on the human perception [5-9]. These methods were based on a metric that provides an objective measure of the quality of videos, by using also parameters that emulate the human process of estimation. More specifically, the ITS (www.its.bldrdoc.gov) and the National Telecommunications and Information Administration (NTIA; www.ntia.doc.gov) have developed a metric which has been implemented in a software tool named Video Quality Metric (NTIA/ITS VQM) [10]. The video quality measurement technique developed by the ITS has been adopted as an ANSI standard T1.801.03-2003 [11] and was included, in the 2004, in two International Telecommunication Union (ITU, www.itu.int/net/home/index.aspx) Recommendations [12, 13]. The NTIA/ITS VQM software has been extensively tested on a wide range of video systems and bandwidths and the validation has been performed through subjective assessments

of multimedia applications using non-expert viewers judging scenes from typical television programs, e.g. sport, action movies etc. [14]. Among the different objective metrics emulating the human perception, the VQM furnishes outcomes very near and comparable to the subjective ones. The VQM tool uses different objective measurements and furnishes results of the video quality in a "scale of impairment" of the videos, in a proprietary scale or in scales used in some subjective methods.

For clarity, we reported a brief description of the most used methods, both objective and subjective, of video quality assessment.

Objective methods

Among the objective methods the most accepted is based on the computation of the peak signal to noise ratio (PSNR): the procedure for this computation is described in [11, 15]. The PSNR is one of the most accepted quality measurement for a picture and appropriate for evaluating the degradation. Referring to a transmitted videoclip with corruption, the PSNR (expressed in dB) is calculated to assess the quality of the received videoclip, according to the formulae (1) and (2):

$$PSNR = 10 \cdot \log\left(\frac{I_{Max}^2}{MSE}\right) \quad (1)$$

where I_{Max} is the maximum achievable pixel grey-level-value of the videoclip ($I_{Max} = 255$ for 8-bit representation) and the MSE is the mean square error;

$$MSE = \frac{1}{K} \sum_{k=1}^K \left\{ \frac{1}{M \cdot N} \sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [f_{rec}(i, j, k) - f_{tran}(i, j, k)]^2 \right\} \quad (2)$$

where K is the number of frames constituting the videoclips, M and N represent the dimensions of the frames; f_{rec} and f_{tran} indicate respectively the pixel values of the frames of the received clip (the corrupted clip) and the transmitted clip (the original clip). The MSE measures the distortion caused by the digital video system (transmission and/or compression), by means of the square distances pixel by pixel; the PSNR measures the normalized distortion, that is the MSE value referenced to the maximal value of a pixel. Higher PSNR value provides a higher image quality: typical values for the PSNR in lossy image and video compression are between 30 and 50 dB; acceptable values for wireless transmission quality loss are considered to be about 20 dB to 25 dB [16].

Subjective methods

For the subjective assessment, different test methods were proposed by the ITU and standardized in the ITU-R BT.500-11 Recommendation [17] to assess the video quality, each one should be used to address particular assessment problems. Each test method is more properly used with a specific assessment problem and each one uses an "impairment scale", that grades the impairment perceived by the subject in different range, from the highest grade

(imperceptible impairment, high quality) to the lowest grade (very severe impairment, poor quality). Table 1 reported the subjective methods described in the ITU-R BT.500-11 Recommendation.

Aim of the work

In this work it has been investigated the use of the NTIA/ITS VQM tool for evaluating the quality of echocardiographic MPEG4 compressed video clips. The MPEG4 compression process was chosen because this codec today is the most employed one in many applications, as personal communications (videoconferencing, video-telephony) and Internet streaming services (Web TV, video on demand).

This study also faced the potentialities of the VQM objective evaluation method in T-E especially in perspective of its employment in an overall automatic process of video quality assessment (VQA-TE). The authors have already exposed the need of objective methods for the assessment of the image quality in T-E [4] and have tested the effectiveness of the NTIA/ITS VQM metric in a preliminary study [18]. In the specific the authors compared the results of the video quality assessment obtained from the VQM tool with those one from the subjective evaluation and the objective PSNR method, and the authors found the suitability of the VQM as appropriate. A previous work had investigated other objective methods proposed by the ITU for the quality assessment in medical video sequences [19]. In details the authors of this work investigated on the quality assessment

of angiographic MPEG4 compressed video data, by means of the PSNR measurement, revealing that the PSNR was a reliable objective measure of quality of medical data. An in-depth study of validation of the NTIA/ITS VQM metrics in T-E was described in a recent work. In this work [20], the authors analyzed the MPEG2 and MPEG4 compressed echocardiographic videoclips: the VQM results of video quality were comparable with the results coming from subjective evaluations. The VQM tool resulted thus appropriate for the video quality assessment.

The NTIA/ITS VQM tool offers many functionalities never explored in depth for investigating specific features of the video sequences, such as the chromatic and/or statistical ones, which directly could affect the diagnostic content of the medical videos. These are basic issues for the design of testing procedures to place in a health technology assessment (HTA) in T-E. The specific aim of this paper was thus to widen the investigation of the use of the NTIA/ITS VQM tool in the process of the VQA-TE, considering also these functionalities of the tool, with also particular care to the figure/role of the designer and/or administrator of the telemedicine T-E system who has a central role in the telemedicine service and as possible tool for HTA of a telemedicine application.

MATERIALS AND METHODS

We evaluated the quality of different digital echocardiographic videoclips by means of the NTIA/ITS

Table 1 | Recommended subjective test methods for the video quality assessment proposed in the ITU-R BT.500-11 Recommendation

Assessment problem	Method used	Description
Measure the quality of systems relative to a reference	Double-stimulus continuous quality-scale (DSCQS) method(1)	Rec. ITU-R BT.500, par. 5
Measure the robustness of systems (i.e. failure characteristics)	Double-stimulus impairment scale (DSIS) method(1)	Rec. ITU-R BT.500, par. 4
Quantify the quality of systems (when no reference is available)	Ratio-scaling method(2) or categorical scaling (under study)	Report ITU-R BT.1082
Compare the quality of alternative systems (when no reference is available)	Method of direct comparison, ratio-scaling method(2) or categorical scaling (under study)	Report ITU-R BT.1082
Identify factors on which systems are perceived to differ and measure their perceptual influence	Method under study	Report ITU-R BT.1082
Establish the point at which an impairment becomes visible	Threshold estimation by forced-choice method or method of adjustment (under study)	Report ITU-R BT.1082
Determine whether systems are perceived to differ	Forced-choice method (under study)	Report ITU-R BT.1082
Measure the quality of stereoscopic image coding	Double stimulus continuous quality-scale (DSCQS) method(3)	Rec. ITU-R BT.500, par. 5
Measure the fidelity between two impaired video sequences	Simultaneous double stimulus for continuous evaluation (SDSCE) method	Rec. ITU-R BT.500, par. 6.4
Compare different error resilience tools	Simultaneous double stimulus for continuous evaluation (SDSCE) method	Rec. ITU-R BT.500, par. 6.4

⁽¹⁾ Some studies on contextual effects were carried out for the DSCQS and the DSIS methods. It was found that the results of the DSIS method are biased to a certain degree by contextual effects. More details are given in Appendix 3 to Annex 1

⁽²⁾ Some studies suggest that this method is more stable when a full range of quality is available

⁽³⁾ Due to the possibility of high fatigue when evaluating stereoscopic images, the overall duration of a test session should be shortened to be less than 30 min

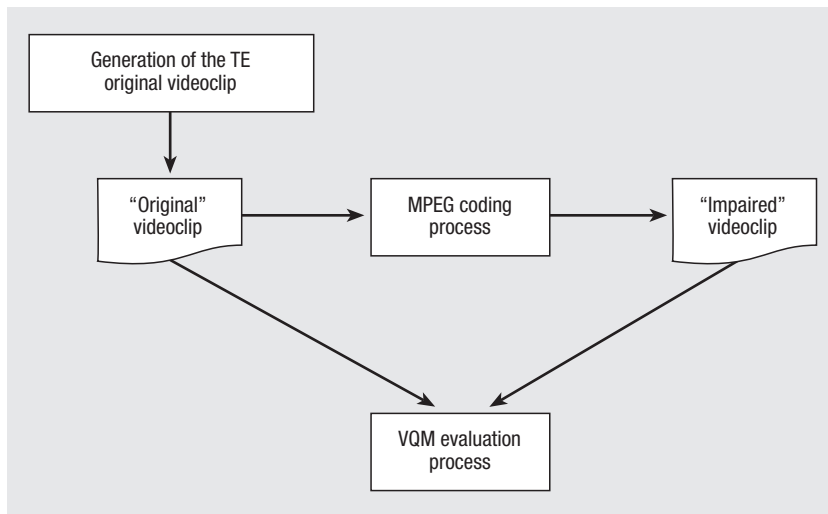


Fig. 1 | The flow of the tele-echocardiographic video quality assessment process.

VQM software, in order to furnish a quality video assessment procedure for evaluating the feasibility of the use of an objective method for the VQA-TE. In fact, as highlighted in detail in the previous section, the VQM metric has been designed for video systems widely ranging in quality and bitrate and it showed a high correlation with the subjective metrics [14].

The NTIA/ITS VQM software is able to compare two different videos, one named “original” and one named “impaired”. The “original” video is an acquired video from the scene under exam; this video can be analogue (coming from a videotape) or digital (coming from an analogue-digital acquisition card). The “impaired” video is intended as the original video which has been submitted to a digital video transmission system or a codec, thus to a degradation process represented by a lossy compression procedure. The NTIA/ITS VQM evaluates, in order to assess the reliability of the digital video system, the quality of the impaired video, by comparing the values of different parameters of the original and impaired videos. Some of the adopted parameters are based on the statistical features of the different frames composing the videos (e.g. PSNR); other parameters are based on some video features perceived by the human eye and they are used to mimic the human evaluation.

We analyzed different MPEG4 compression schemes evaluating the quality of the compressed echocardiographic videoclips. The quality of the videoclip was assessed via the VQM software provided by the NTIA and the ITS available at the URL www.its.blrdoc.gov/n3/video/VQM_software.php for 6-months free license. Different echocardiographic videoclips have been used as “original” and the compressed videoclips have been used as “impaired”. The original videoclip was generated from the sequence of echocardiographic images coming from the echocardiograph, during an exam. These “original” videoclips were compressed with different MPEG4 algorithms and used in the VQM evaluation as “impaired” videoclips. The abovementioned operations have been

done using the license free software “VirtualDub” (www.virtualdub.org).

The entire tele-echocardiographic video quality assessment process we performed, consisted in the following step:

1. generation of the “original” videoclip;
2. encoding process (generation of the “impaired” videoclip);
3. VQM evaluation.

The TE-VQA process is fully described in the following and is depicted in *Figure 1*.

Generation of “original” videoclip

The most common T-E scenarios are the transmission over LAN or/and over geographic networks and the transmission with video-conferencing.

Nine echocardiographic exams were analyzed and the image sequence for each exam was used to produce a digital videoclip lasting 5-second, to generate the “original” videoclip for the VQM software. The “original” videoclip was generated with the VirtualDub software in the uncompressed AVI format, with RGB24 color format, that is one of the color format accepted by the VQM tool. The original videoclips were generated in two frame sizes: the same of the images coming from the echocardiograph (720×512 pixels), here denoted full resolution (FR) video format, and 352×288, named common intermediate format (CIF) video format, that is the format used in the videoconferencing systems.

Encoding process

The original videoclips were compressed with different MPEG4 algorithms and at different bitrate. Two MPEG4 codecs (DivX 6.8.3 DivXNetworks Inc, Windows Media Video 9 Microsoft) and six bitrate were considered (384, 640, 1280, 1500, 2000, 3000 kb/s) in order to estimate the compression effect on the quality of the compressed videos. The adopted bitrates were chosen according to the more easy available transmission bandwidths furnished

by the telecommunication service providers, with the aim to evaluate the quality of the T-E service also over heterogeneous networks including public IP geographic networks.

We evaluated the compressed videoclips that must be streamed over a network with limited bandwidth (like the Internet). The bitrate of the compressed videoclips should correspond to an expected transfer rate, according to the available bandwidth. Specifying a constant bitrate (CBR) for a codec, it causes the image quality to fluctuate somewhat in order to ensure the imposed bitrate, but the content encoded in this manner is well suited to streaming over networks. Conversely, the variable bitrate (VBR) encoding furnishes a better quality of the compressed videos, but do not assure the proper bitrate in tune with the available bandwidth. For this reason, we selected, for each codec, the encoding mode that assured a constant average bitrate of the compressed videos according to an imposed prefixed bitrate, even if at the expense of some quality of the compressed videoclips and some *encoding parameters* needed to assure the compatibility of the format between the original video and the compressed video.

DivX Codec (DIVX)

The DivX codec (DivX 6.8.3 DivXNetworks Inc, www.divx.com) is used in all cases in which it is wanted to be stored videos of good quality in a little space, for archiving on the most diffuse CD or exchanging them through Internet. We selected the *encoding mode* in order to impose a prefixed bitrate and we imposed the format of the compressed video, by selecting the parameter "*certification profile*", in order to assure the compatibility of the compressed video with the player device (the monitor of the remote PC). Among the available encoding modes from the DivX codec (1-Pass, 1-Pass Quality-based, Multi-pass), the *1-pass encoding mode* was preferred. The 1-pass encoding is the mode particularly used for the live sources, when the control on the bitrate is needed and the encoding speed is required, like in the real-time T-E. This encoding mode produces an average bitrate, that is the closest to the imposed bitrate, generally lower, at the expense of some quality of the compressed video. In addition we used, as "*certification profile*", the "*home theater profile*", that assures compatibility for a maximum resolution of 720×480 pixels at 30 fps or 720×576 pixels at 25 fps and for a maximum average bitrate of 4000 kb/s.

Windows Media Video 9 codec (WMV9)

The Microsoft WMV9 codes (© 2003 Microsoft Corporation, www.microsoft.com/windows/windows-media/it/9series/codecs.aspx) was originally designed for Internet streaming applications, as a competitor to RealVideo, but it is today normally used for compressing any audio-video content, also for the high-definition video content. Among the available encoding modes from the WMV9 codec (One-pass CBR, Two-

pass CBR, One-pass VBR, Two-pass VBR, Peak-constrained two-pass VBR), the *One-pass CBR* encoding mode was preferred. The quality of One-pass CBR is constrained by the bitrate and buffer size settings and results quicker of the Two-pass CBR and most suitable in real-time applications. The buffer size (in ms) was adjusted for obtain an average bitrate of the compressed videos lower to the imposed bitrate. For assuring the compatibility of the format of the compressed video clips with the player devices, the WMV9 employs the combination of the two parameters "*decoder complexity*" and "*decoder level*". We used the default setting ("Main" and "Auto"), assuring the compatibility of the compressed videos: maximum resolution of 720×480 pixels at 30 fps or 720×576 pixels at 25 fps, maximum bitrate of 10 Mb/s.

VQM evaluation

The VQM furnishes a measure of the video quality. The VQM offers two methods for measuring the video quality. In the following the method of the VQM evaluation is described, extracted from the user's manual of the NTIA/ITS tool [21]. The first method is applied when both the original and processed (that is the "impaired" one) videoclips are available: the "original and processed videoclips" VQM method; the second method is applied when the only original videoclips is available: the "original videoclips only" VQM method. The VQM software simulate a "video system" applied to the videoclips, in order to test the reliability of such video system. The video system could be intended such as a digital video transmission system (as in the "original videoclips only" VQM method) or a codec (as in the "original and processed videoclips" VQM method). This video system is referred to in the video quality community, as well as in the VQM tool, as a hypothetical reference circuit (HRC). The definition of HRC has been proposed by the ITU [22] and it is present in many ITU-R-Recommendations.

In the video quality assessment carried out in this work, the "original and processed videoclips" VQM method was applied and the VQM, by comparing the two videoclips, the "original" and the "impaired", performed the following steps:

1. spatial and temporal calibration between the videoclips;
2. features extraction of the perceived quality of the "impaired" videoclips;
3. "quality parameters" computation;
4. VQM score computation in the "impaired" videoclips.

The steps 1 and 2 are operations accomplished on each frame, constituting the two videoclips to be compared (frame-based). The steps 3 and 4 provide a global computations on all frames of the videoclips.

1. Spatial and temporal calibration

The *spatial calibration* is used in order to determine, for each frame, space movements, both in vertical and horizontal directions, of the processed

video regarding to the original and to discard portions of the processed videoclips on which video quality measurements should not be made. Many video systems spatially shift the image, hence a corrective action is accomplished to detect and remove the spatial shift. In addition many digital video systems fail to transmit all the image pixels around the edge of the image and then insert black image pixels to form a black border around the image. The spatial calibration determines a “processed valid region” (PVR), that is the region containing the information of the frame, in order to limit the extraction of the characteristics to the pixels of this region.

The *temporal calibration* is used to correct the “gain” of the luminance (said contrast in the television) and the “offset” of the luminance (said brightness in the television) of the processed video, in order to be able to compare the original and processed videoclips. In addition it is used to estimate and correct temporal shift (that is video delay) of the processed video regarding the original. The sequence of the spatial calibration and the temporal calibration defines the “temporal valid region” (TVR), that is a sub-sequence of frames within the processed videoclip that contain valid video.

2. Features extraction of the perceived quality

The features extraction of the perceived quality are parameters of quality that are indicative of the perceived changes of video quality, that is the perceived changes in the space, in the time and the colours (*i.e.* features based on the chrominance, the contrast) of the videoclips. Each feature is extracted from spatial-temporal (S-T) regions containing an integer number of frames. Since the most used frame rates are 10 fps, 15 fps, 25 fps and 30 fps, a time interval having an integer number of frames is equal to a 1/5 of second. For example, at 25 fps, a S-T region could be of 8x8 pixels x5 frames. In such a manner the correlation between the final VQM assessment and the subjective assessment is the most high, whereas for S-T regions of greater dimension, the correlation decreases [23, 24].

3. Quality parameters computation

From the extracted features, for measuring distortions in video quality, a set of “quality parameters” are computed. Firstly, for each S-T region, some “comparing functions” of features are calculated: “gain” (if the difference between the “original” and “processed” feature value is positive), “loss” (if the difference between the “original” and “processed” feature value is negative) and “Euclidean distance” (the length of the difference vector between the original feature vector and the corresponding processed feature vector). Then, a single quality parameter of the video, referring to a feature, is derived from the corresponding “comparing functions”, by computing the mean value, the standard deviation for each frame (over space) and thus among frames (over time), or extracting values below or above prede-

fining threshold levels (perceptibility threshold, simulating the human perception) and computing thus the mean or the standard deviation of the extracted values (in space or time). The quality parameters were computed for the entire videoclip and for each frame.

4. VQM score computation

The VQM score is the measure of the video quality. The score produced by the VQM is accomplished by means of a weighted linear combination of video quality parameters and is a numerical value that ranges from a minimum value (no perceived impairment) to a maximum value (maximum perceived impairment). For computing the score of the video quality, the VQM defines five different models of video quality: a VQM model is intended as a specific algorithm for computing the VQM score.

VQM model

Each VQM model is optimized and standardized for a particular application. The five VQM models, fully described in Section 6 of NTIA Report 02-392 [23], are [21]:

1. general: optimized using a wide range of video quality and bitrates;
2. developer: optimized using a wide range of video quality and bitrates with the added constraint of fast computation;
3. video conferencing: optimized for video conferencing (*e.g.*, H.263, MPEG4);
4. television: optimized for television (*e.g.*, MPEG2);
5. PSNR: optimized using a wide range of video quality and bitrates.

More specifically, the VQM models 1 to 4 furnish the VQM scores as a weighted linear combination of different set of video quality parameters, while the PSNR model uses the peak-signal-to-noise-ratio (PSNR) measurement, as is described in [15, 23]. The models 1 to 4 also allow the computation of the root cause analysis (RCA). The RCA is a procedure generally used in subjective tests for determining the probable causes of degradation of the quality, by means of evaluation of the presence or absence of some perceived video artifacts in the processed video, like as blurring, tiling, jerky motion or dropped frames. RCA produces, for each possible artifact, the percentage of the artifacts presence: 100% means that the artifact was perceived as being the primary artifact by all the viewers, 50% means that the artifact was perceived as being a secondary artifact, and 0% means the artifact was not perceived. The VQM software furnishes, for each of the VQM models 1 to 4, the list of a specific set of artifacts, each one with the relative percentage of presence [23], as defined in [25]. The PSNR model computes the PSNR present in the processed videoclip.

The two video quality models “General” and “PSNR” (denoted in the following VQM-G and VQM-P) were used, in order to compare the VQM results more simi-

lar to the human assessment (from the General model) with those more similar to the objective assessment (from the PSNR model). In the following, we report a brief description of the two model adopted in our study.

General model

This model computes the VQM score by means of a linear combination of seven video-quality parameters chosen to take into account the spatial distortion and changes in the distribution of color. The RCA of this model furnishes the list of four artifacts: 1) “blurring”: the contrast loss originated by the reduced sharpness of edges and spatial details; 2) “jerky or unnatural motion”: *i.e.* the dropping of frames causing a continuous motion perceived as a series of distinct still frames; 3) “global noise”: the contrast gain originated by distortions in proximity of edges; 4) “block distortion”: type of distortion observed when transmission errors are present.

PSNR model

The PSNR (peak signal to noise ratio) model is based on the PSNR measurement and the VQM score is computed via the formulas:

$$VQM - P = \frac{1}{1 + e^{0.1701(PSNR - 25.6675)}} \quad 10 \leq PSNR \leq 55$$

The bounds on PSNR in the above equation are those ones recommended in [23].

Procedure for VQM score computing

In the following, the steps to obtain a VQM score of video quality are listed [21]:

1. creating a library containing the “original” and “impaired” videoclip;
2. creating an HRC file, containing information about the simulated circuit;
3. running a test;
4. saving the results in a report.

1. Library. The created library is a repository of original and processed videoclips. We created one library for each original videoclip (corresponding to an echocardiographic exam) and each MPEG compressed videoclip: to each original videoclip (9 echocardiographic exams), different MPEG codecs (2 codecs) and, for each codecs, different bitrates (6 bitrates) were applied, providing a total of 108 libraries to be applied to carry out the VQM score. The following parameters should be fixed: frame rate, video color format, video size, and video type (progressive or interlaced). These values were used: frame rate of 25 f/s, RGB24 video color format, FR (720x512) and CIF (352x288 pixels) video format, progressive video type.

2. HRC. After the library has been created, it needs to create the Hypothetical Reference Circuit (HRC) file containing information about the simulated circuit. The HRC can assume one of the three

types of expected video quality: Television Quality, Video Conferencing Quality, Unknown Quality. We used the Unknown Quality type, that resulted more appropriate for general application fields, when a specific video system was not defined.

3. Test. In running a test, first the spatial and temporal calibration was accomplished and thus, on the basis of the selected quality model, the VQM score was furnished. If, during calibration, some values are outside of the expected range, an alert arises; in this case the user after analyzing the calibration errors may decide if carry on the test.

4. Report. The results produced by the VQM tests were recorded in report files, containing information about the analyzed videoclips (*e.g.* number of frames, frame rate, frame size) and different types of results (*e.g.* VQM score, calibration results, RCA and PSNR results) in the relative sections, described in the following.

Report - Section of VQM score. The VQM software produces the quality score in four different impairment scales: the native scale (0-1; 0: no impairment, 1: maximum impairment) and three different scales agreed with some of the subjective test methods, reported in *Table 1*. The three subjective scales are: 1) Double Stimulus Impairment Scale (1 to 5; 5: no impairment); 2) Double Stimulus Continuous Quality (0 to 100; 0: no impairment); 3) Double Stimulus Comparison Scale (-3 to +3; -3: the lowest quality value for the processed videoclip). The Double Stimulus Impairment Scale (DSIS) has been used to assess the quality of the tele-echocardiographic videos; the same scale had already been adopted in a previous study of some authors facing the subjective assessment in T-E [18]. The DSIS method can be used to measure the impairment of a video caused by a transmission path and it is capable of evaluating the robustness of the transmission systems. This method uses a five-grade impairment scale ranging from 1 (no impairment) to 5 (maximum impairment) according to the following scale:

- 5 - imperceptible
- 4 - perceptible, but not annoying
- 3 - slightly annoying
- 2 - annoying
- 1 - very annoying

According to this method, the VQM score, both for the VQM-G and VQM-P models, has been furnished as a real value ranging from 1 to 5. As commonly adopted for this scale, the threshold value for the defining videos of good quality was set at value of 3: score values greater than 3 denoted a good quality level for the compressed videoclips.

Report - Section of calibration results: the results coming from the calibration process are the values of the parameters (relative to the spatial and temporal and color information) used to determine the valid region to be processed. In general the entire videoclip is considered as valid region, whereas the limits of the valid region were indicated, if calibration errors were detected.

Table 2 | The VQM scores from for the full resolution video format. The VQM scores ranged from 1 to 5 and were referred to the General model (s-VQM-G) and to the PSNR model of the VQM

Full Resolution video format					
Bitrate (kb/s)	CODEC DIVX		Bitrate (kb/s)	CODEC WMV9	
	s-VQM-G (1÷5)	s-VQM-P (1÷5)		s-VQM-G (1÷5)	s-VQM-P (1÷5)
384	3.50	4.01	384	3.43	3.36
640	3.93	4.23	640	3.60	3.78
1280	4.36	4.46	1280	4.11	4.23
1500	4.44	4.50	1500	4.18	4.31
2000	4.59	4.58	2000	4.34	4.41
3000	4.73	4.66	3000	4.59	4.61

Report - Section of RCA and PSNR results: the RCA results were furnished only for the VQM-G model while, for the VQM-P model, PSNR values were furnished.

RESULTS

In the following the different types of results coming from the VQM analysis were described: the outcomes from VQM score, the outcomes from the calibration process and the outcomes from the RCA and PSNR analysis.

Outcomes from the VQM score

Table 2 and Table 3 report the global values of the different results obtained, respectively for the Full Resolution and CIF video format. For each codec and each bitrate the following results are indicated: the mean values over echocardiographic videoclips of the VQM score from the General Model (s-VQM-G) and the mean values over echocardiographic videoclips of the VQM score from the PSNR Model (s-VQM-P).

The two charts of the Figure 2 show the plot of the VQM score values versus the bitrates for the

Table 3 | The VQM score for the CIF video format. The VQM scores ranged from 1 to 5 and were referred to the General model (s-VQM-G) and to the PSNR Model of the VQM

CIF video format					
Bitrate (kb/s)	CODEC DIVX		Bitrate (kb/s)	CODEC WMV9	
	s-VQM-G (1÷5)	s-VQM-P (1÷5)		s-VQM-G (1÷5)	s-VQM-P (1÷5)
384	4.35	4.35	384	4.09	4.09
640	4.61	4.52	640	4.35	4.34
1280	4.80	4.67	1280	4.69	4.63
1500	4.81	4.69	1500	4.74	4.67
2000	4.84	4.75	2000	4.79	4.73
3000	4.88	4.81	3000	4.81	4.78

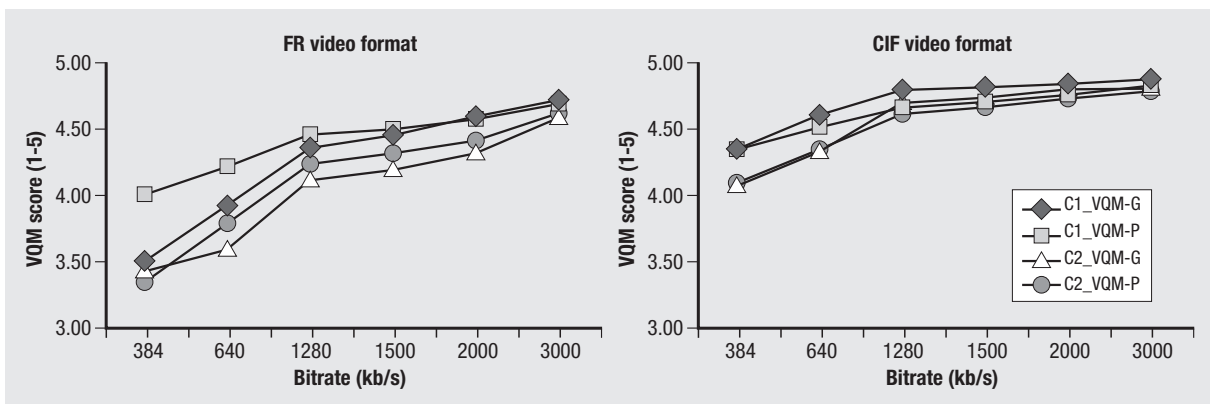


Fig. 2 | Plot of the VQM score values versus bitrate for the two used codecs, in the FR and the CIF video format. VQM scores for the General and PSNR models for the DIVX codec (C1_VQM-G, C1_VQM-P) and WMV9 codec (C2_VQM-G, C2_VQM-P).

Table 4 | The Pearson correlation coefficients for each video format (FR, CIF). C1_G-P: Pearson correlation coefficient between the model VMQ-G and VQM-P for the codec DIVX; C2_G-P: Pearson correlation coefficient between the model VMQ-G and VQM-P for the codec WMV9

Pearson's coefficients	C1_G-P	C2_G-P
FR	0.9998	0.9831
CIF	0.9856	0.9980

two different used MPEG4 codecs (C1=DIVX e C2=WMV9 in figure) and for the two video formats (FR, CIF): for all bitrates the VQM score resulted greater than the threshold (equal at 3), indicating an adequate quality of the compressed videoclips; for each codec the VQM score increased with the bitrate, as it could be foreseen. Figure 2 elucidates that the VQM score seems to be independent to the used codec. These results indicate that the T-E service is practicable over networks (LAN and/or geographic), once set the offered bandwidths by the network service providers.

Table 4 shows the Pearson correlation coefficients between the VQM scores from the two VMQ models, for each video format: the high values of the coefficient, close to 1, indicating a high correlation and thus that the two models run in a comparable manner.

Outcomes from the calibration process

No significant problems of spatial and temporal calibration have been produced and, when needed,

the VQM software automatically computed the PVR and the TVR regions to be processed. In all cases but one the PVR and the TVR coincided for the overall videoclip. In only one case, the FR video format videoclip compressed at 384 ks/s both with the DIVX and WMV9 codec, some border pixels of each frame were excluded from the processing and the final PVR and TVR automatically were determined. However in any cases no time delay between the original and the processed videoclip was encountered.

Outcomes from the RCA and PSNR analysis

Table 5 shows the results coming from the RCA analysis and the PSNR values for the FR video format. In the table, the mean values and, in brackets, the minimum and the maximum values are reported. The same results for the CIF format are presented in Table 6.

Except for the only codec WMV9 at 384 kb/s, the RCA parameters were always lower than 50%, indicating that no one of these parameters was perceived as primary (100%) or secondary ($\geq 50\%$) artifact of degradation, but only as an artifact slightly perceived. This fact agreed with the VQM score that judged all the considered bitrates producing videoclips of good quality. As expected, the more the bitrate increased, the more the value of the RCA parameters decreased. From the RCA results, it was showed that in all cases the "blurring" was the main artifact of degradation, whereas the "Global noise" resulted nearly absent. The WMV9 codec resulted worse than the DIVX codec, both for the FR and CIF video format, as especially illustrated

Table 5 | The RCA and PSNR results for the FR video format, for the two applied codecs (DIVX, WMV9). The indicated values are: the mean values and, in brackets, the minimum and maximum values

Full resolution video format					
CODEC DIVX	RCA from VQM-G				PSNR from VQM-P
BITRATE (kb/s)	Blurring (%)	Jerky or unnatural motion (%)	Global noise (%)	Block distortion (%)	PSNR (dB)
384	29.56 (24-38)	16.89 (15-19)	0.78 (0-1)	26.44 (20-35)	32.19 (31.51-32.89)
640	20.67 (16-27)	12.22 (10-15)	0.22 (0-1)	19.67 (15-26)	34.07 (32.77-34.69)
1280	11.78 (9-16)	7.56 (5-9)	0 (0-0)	9.33 (5-13)	36.55 (35.43-37.18)
1500	10.56 (8-15)	6.56 (5-8)	0 (0-0)	6.78 (3-10)	37.15 (35.99-37.87)
2000	8.11 (6-11)	5.22 (4-7)	0 (0-0)	2.33 (0-5)	38.33 (37.23-39.04)
3000	6.00 (5-8)	4.22 (3-5)	0 (0-0)	0.44 (0-2)	39.72 (38.73-40.33)
CODEC WMV9	RCA from VQM-G				PSNR from VQM-P
Bitrate (kb/s)	Blurring (%)	Jerky or unnatural motion (%)	Global noise (%)	Block distortion (%)	PSNR (dB)
384	34.44 (29-41)	33.56 (22-58)	0 (0-0)	38.78 (29-51)	27.81 (26.97-28.47)
640	29.78 (25-35)	22.78 (18-36)	0.22 (0-1)	29.67 (21-40)	30.51 (29.22-32.08)
1280	19.44 (16-22)	14.78 (14-16)	0 (0-0)	20.44 (11-30)	34.09 (33.14-35.13)
1500	18.11 (16-20)	13.44 (13-15)	0 (0-0)	18.44 (10-29)	34.85 (34.34-35.74)
2000	15.11 (13-18)	11.00 (10-12)	0 (0-0)	14.33 (8-24)	35.94 (35.33-36.94)
3000	10.00 (7-14)	7.89 (7-9)	0 (0-0)	3.00 (0-10)	38.74 (37.95-39.63)

Table 6 | The RCA and PSNR results for the CIF video format, for the two applied codecs (DIVX, WMV9). The indicated values are: the mean values and, in the quotes, the minimum and maximum values

CIF video format					
CODEC DIVX	RCA from VQM-G				PSNR from VQM-P
Bitrate (kb/s)	Blurring (%)	Jerky or unnatural motion (%)	Global noise (%)	Block distortion (%)	PSNR (dB)
384	11.78 (8-16)	10.11 (8-12)	0.33 (0-1)	7.00 (4-11)	35.30 (34.11-36.10)
640	8.44 (6-12)	7.22 (5-11)	0 (0-0)	1.89 (0-4)	37.35 (35.85-38.25)
1280	5.11 (4-6)	4.44 (3-6)	0 (0-0)	0.67 (0-1)	39.77 (39.06-40.48)
1500	4.56 (4-6)	3.89 (3-5)	0 (0-0)	1.33 (0-3)	40.31 (39.76-41.06)
2000	3.78 (3-4)	2.89 (1-4)	0 (0-0)	1.67 (0-4)	41.49 (40.83-42.47)
3000	2.67 (2-3)	2.89 (2-3)	0 (0-0)	0 (0-0)	43.33 (42.73-44.14)
CODEC WMV9	RCA from VQM-G				PSNR from VQM-P
Bitrate (kb/s)	Blurring (%)	Jerky or unnatural motion (%)	Global noise (%)	Block distortion (%)	PSNR (dB)
384	20.33 (16-25)	19.33 (18-23)	0 (0-0)	11.78 (7-20)	32.86 (32.23-33.77)
640	16.33 (13-21)	13.89 (12-16)	0 (0-0)	7.22 (2-14)	35.18 (34.32-36.28)
1280	10.56 (8-13)	6.67 (6-8)	0 (0-0)	0.67 (0-2)	39.06 (38.12-39.92)
1500	9.67 (8-12)	5.67 (5-7)	0 (0-0)	0.33 (0-1)	39.79 (38.88-40.63)
2000	8.44 (7-11)	4.33 (3-5)	0 (0-0)	0 (0-0)	41.07 (40.36-42.02)
3000	7.56 (6-9)	2.33 (2-3)	0 (0-0)	0.11 (0-1)	42.50 (42.19-42.85)

by the “blurring”, the “jerky or unnatural motion” and the “block distortion” parameters: the values of these parameters were greater for the WMV9 codec. Moreover, the artifact of block distortion seemed more evident in the FR video format: the difference between the values of the “block distortion” parameter was greater for the FR video format respect to the CIF video format ones; a further deepening of analysis in the filed could face the low-resolution format may mask some distortion artifacts.

The PSNR values ranged from 26.97 to 44.14 dB, showing typical PSNR values (30-50 dB) of compressed videos. The worst values of the PSNR were

obtained for the FR video format videoclips compressed at 384 kb/s with the WMV9 codec (27.81 dB; 26.97-28.47 dB), while the best values were for the CIF video format videoclips compressed at 3000 kb/s with the DIVX codec (43.33 dB; 42.73-44.14 dB), also in this case a further deepening of analysis in the filed could face the low-resolution format may mask artifacts. *Figure 3A* shows the PSNR frame by frame for the FR video format videoclips n. 5, having the worst global PSNR value (26.97 dB): it could be noted the low values of PSNR with a wide variability among the frames. For this videoclips, the artifacts from the RCA were largely perceived: the RCA parameters presented the maximum values (blurring =

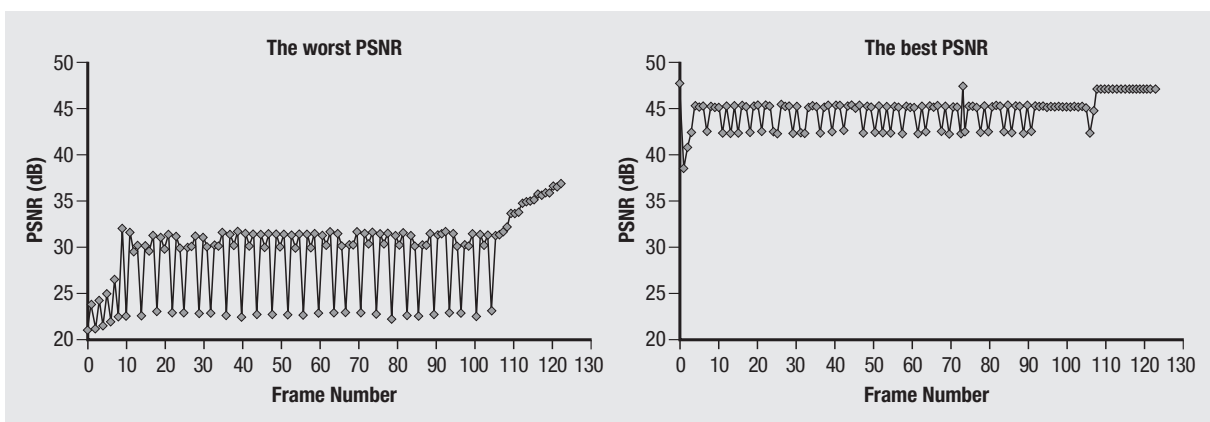


Fig. 3 | Plot of the PSNR values versus the frames in two videoclips. A): the FR video format videoclips n. 6 compressed at 384 kbls with the WMV9 codec; B) the CIF video format videoclips n. 7 compressed at 3000 kbls with the DIVX codec.

31%; Jerky or unnatural motion = 58%; global noise = 0%; block distortion = 49%). *Figure 3B* shows the PSNR frame by frame for the CIF video format videoclip n. 6, having the best global PSNR (44.14 dB): the PSNR values were greater than those ones in the *figure 3A* and presented a lower variability. In agreement with this behavior, the RCA parameters resulted very slightly perceived: blurring = 2%; jerky or unnatural motion = 2%; global noise = 0%; block distortion = 0%.

In conclusion the values of the RCA parameters were lowest for the CIF video format, whereas the PSNR values were comparable for the two tested video formats. The RCA parameters are representative of specific artifacts, this intuitively was foreseeable. In fact, the artifacts are more appreciable in the FR format respect to the CIF format due to the lowest resolution. On the contrary, the PSNR is a general index of degradation and it is not representative of a specific artifact; this is the reason for the absence of significant differences between the two FR and CIF formats.

DISCUSSION

This study was relevant to the use of an automatic tool to assess the image quality, with objective methods based on models of the human subjective perception. The quality of the MPEG4 compressed videoclips in T-E was investigated, in order to explore the feasibility of the quality assurance of the T-E service, both over LAN and geographic network. The complexity of using the human assessment (a lot of human resources and time), demand the need of automating the overall video quality assessment process in T-E, for a better modeling the T-E system and monitoring the T-E service, from a HTA perspective. The VQM use resulted suitable for this aim and should be taken into account also as a test component in HTA processes in T-E.

Comparing with the subjective assessment

The subjective assessment in general shows many limitations for applying to a in-depth video quality assessment in T-E. Up to now, in the studies asserting that the accuracy of the T-E was satisfactory, the validation of the T-E was retrospective [26] and often referring to validation of a predefined T-E system, such as low-bandwidth transmission [27] or videoconferencing [28, 29]. In these cases no suggestions can be assumed for modeling the T-E service, for instance in the bandwidth choice. One study reported a subjective assessment of pediatric echocardiograms transmitted at different bitrates (256, 384, 512, 768 kb/s), investigating on the influence of the bandwidth on the quality [30]: the observers discriminated on the quality of the transmitted echocardiographic exams, while no detailed quantification of the quality has been performed. However besides the high cost and time needed

to effectuate the subjective evaluation, the inter-observer and the intra-observer variability should be also considered: these factors should cause incorrect or discordant evaluations of the degraded videoclips. Papers reported study of validation of the T-E, based on subjective assessment, presenting some discrepancies in the evaluation. In a study [20], some "original" videoclip has been judged as slightly "impaired" by some observers. Another study asserted that the subjective scoring was a function of the original image resolution [31]. The study also enlightened an improving of the subjective performance with the increasing of the number of trials, seemed to be caused by a training effect [31]. The above-mentioned erroneous evaluations obtained by the human assessment could be originated also from images of suboptimal quality. However the use of the VQM tool, via the RCA analysis and the calibration results as enlightened in this study, could be of aid also particularly in these sub-optimal situations with light differences between the original and compressed videoclip, showing new potentialities respect to the previous study [18-20].

Improvement of the activity of the designer of the application through modeling the requirement setting of a T-E system

The RCA analysis offers the operator the confidence about sources of affection occurred in the compression process and gives more information than the VQM overall score (like so an any overall subjective quality score), as well as the overall PSNR value (like so from the objective measurements). This analysis thus could furnish to the T-E system designer the indication how the compression process could hardly affect the spatial and temporal content, compromising the diagnostic accuracy. The VQM tool could allow to investigate on how the encoding techniques and bitrates correlate with the informative content of the different echocardiographic video sequences (*e.g.* high/low motion, color/black-white images, resolution). These characteristics of the video sequences are closely related to the echocardiographic techniques and could affect the types of diagnosis to be carried out, only detecting diseases or also measuring physiological and functional anatomical features. Respect to these aspects, the VQM tool could be useful in the designing dedicate systems for T-E applications (*i.e.* in the bandwidth and codec choice), tailored to the requirement of the clinical setting, that is for instance neonatal, pediatric or adult T-E, types of heart diseases to be examined, types of echocardiographic exams. The encoding process could affect also the quantitative echocardiographic measurements. The accuracy of such measurements was previously investigated in MPEG1 encoded echocardiographic studies [32], resulting that these measurements were comparable with measurements in digitized sVHS videotape. The VQM

tool, thanks the abovementioned embedded tools, useful to modeling a transmission process (intrinsically degrading) could also be of aid, to explore how the different coding process could affect images features (like as edge structures, contrast and chromatic aspects), prejudicing the accuracy of the semi-automatic echocardiographic measurements.

Improvement of the activity of the network administrator during the monitoring of the T-E service

In the telemedicine applications the network parameters should be considered in order to design a network running the application with suitable performance [33]. Generally in planning a telemedicine service, some trade-off have to be forecasted [34], including the quality of service (QoS) of the available network. In telecommunication the term "quality of service" is intended as the network ability to guarantee a certain level of quality of the data flow, in particular audio and video. For this aim some network parameters, such as bit error rate, delay, jitter, packet dropping probability can be monitored for evaluating the network performance. Some authors carried out studies on the general QoS in healthcare services [35] and other focused on the QoS of T-E in real-time ultrasound video transmission over wireless network [36]. Another study investigated on the influence of the bitrate and the network parameters on the diagnostic quality of real-time transmitted tele-echocardiographic images [37]. The VQM could be used thus to deeply investigate how the coding process (used codec and bitrate) and network parameters (*i.e.* high traffic, noise, packet loss) could affect the diagnostic content of the echocardiographic videoclips, determining, for the adopted coding process, network parameters thresholds assuring an adequate diagnostic content of the videos (compressed and transiting in the network). In a perspective of continuing quality assurance of a T-E service, the continuous monitoring of the QoS could be desirable.

Improvement of the activity of the assessors by integrating the VQM tool in a HTA process

The VQM tool could be used by the T-E system designers in order to contemplate the opportunity of integrating this tool in a process of HTA in T-E. The VQM acceptance could be focused on the usability of the tool, in term of both cost and time savings and on its potentiality to guide in the choice of the best compression process preserving the diagnostic accuracy. The VQM tool thus could be employed in some activities of telemedicine technology assessment [38-40], such as in testing the diagnostic effectiveness (*i.e.* via ROC analysis) for the diagnostic accuracy assurance or in maintaining (technical service) the provided telemedicine service. For these purposes, the VQM tool could be properly configured and inserted in a control

procedure of the T-E network performance, especially in real-time T-E, in order to guarantee to the experts, videoclips of adequate quality for a right diagnosis and to avoid non-diagnostic videoclips.

CONCLUSION

In this work it has been investigated the use of the NTIA/ITS VQM tool for evaluating the quality of echocardiographic MPEG4 compressed video clips. The VQM is an automatic tool of video quality assessment with objective methods, based on subjective models. The MPEG4 compression process was preferred because this codec today is the most employed one in many multimedia applications. This study faced the potentialities of the VQM objective evaluation method in T-E especially in perspective of its use in an overall automatic process of video quality assessment (VQA-TE). As the NTIA/ITS VQM tool offers many functionalities never explored in deep for investigating specific features of the video sequences (such as the chromatic and/or statistical ones, which directly could affect the diagnostic content of the medical videos) the study also investigated in details these functionalities. These are basic issues for the design of testing procedures to place in a more general validation process for T-E, which could be used by the designer and/or administrator of the T-E telemedicine service as a whole. All the investigated functionalities of the tool resulted reliable for the video quality assessment in T-E and in particular the tool showed many instruments useful for deeply investigating on the impairments of the tele-echocardiographic compressed videoclips showing an improvement of the previous studies focused on VQM starting to our previous congressional presentation up to the validation of Moore *et al.* described in [18-20]. In particular the study enlightened also the usefulness of the VQM software as a tool useful for the figures involved in the process of designing, maintaining and assessing the application, such as the designer of the application network administrator the activity of the assessors. Anyway, beside the feasibility of the use of the VQM tool, further investigation with this tool could be planned for echocardiographic exams with specific diagnostic content, like as routine pathologic exams as "critical" cases, in order to assess effectiveness of the T-E also for a semi-automatic assisted diagnosis approach. In general, in telemedicine, beside the use of tests of image quality, it is important to plan a more general investigation of the telemedicine application with the aim to integrate it in a routine clinical service. With this consideration the VQM tool could be useful to be integrated in a HTA [41]; dedicated to T-E which should consider a wide range of aspects, from technical to socio-economical aspects.

References

- Sable C. Digital echocardiography and telemedicine applications in pediatric cardiology. *Pediatr Cardiol* 2002;23:358-69.
- Garrett PD, Boyd SY, Bauch TD, Rubal BJ, Bulgrin JR, Kinkler ES Jr. Feasibility of real-time echocardiographic evaluation during patient transport. *J Am Soc Echocardiogr* 2003;16(3):197-201.
- Woodson KE, Sable CA, Cross RR, Pearson GD, Martin GR. Forward and store telemedicine using motion pictures expert group: a novel approach to pediatric tele-echocardiography. *J Am Soc Echocardiogr* 2004;17(11):1197-200.
- Giansanti D, Morelli S, Macellari V. A protocol for the assessment of diagnostic accuracy in tele-echocardiography imaging. *Telemed J E Health* 2007;13(4):399-405.
- Tektronix. A guide to maintaining video quality of service for digital television programs. Application note Tektronix 2000. Available from: www.tek.com/Measurement/App_Notes/Technical_Briefs/digital_QoS/25W_14000_1.pdf. Last visited: 11/12/08.
- Watson AB, Hu J, McGowan JF III. Digital video quality metric based on human vision. *J Electronic Imag* 2001;10(1):20-9. (Journal site: <http://spiedl.aip.org>).
- Van den Branden Lambrecht CJ, Verscheure O. Perceptual quality measure using a spatio-temporal model of human visual system. In: *Proceeding of SPIE*. San Jose (USA), February 1996. SPIE: 1996. vol. 2668. p. 450-61.
- Hekstra AP, Beerends JG, Ledermann D, De Caluwe FE, Kohler S, Koenen RH, Rihs S, Ehrsam M, Schlauss D. PVQM: a perceptual video quality measure. *Signal Process: Image Comm* 2002;17(10):781-98.
- Wang Z, Lu L, Bovik AC. Video quality assessment using structural distortion measurement. Special issue on "Objective video quality metrics". *Signal Process: Image Comm* 2004;19(2):121-32.
- NTIA/ITS VQM Software: www.its.bldrdoc.gov/n3/video/VQM_software.php. Last visited: 11/12/08.
- ANSI T1.801.03 - 2003. *Digital Transport of One-Way Video Signals - Parameters for Objective Performance Assessment (Revision of T1.801.03-1996)*. Washington, DC: American National Standard Institute/Alliance for Telecommunications Industry Solutions ATIS; 2003.
- ITU-T Recommendation J.144 (03/04). *Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference*. Geneva: Recommendations of the International Telecommunication Union/Radiocommunication Sector; March 2004.
- ITU-R Recommendation BT.1683 (06/04). *Objective perceptual video quality measurement techniques for standard definition digital broadcast television in the presence of a full reference*. Geneva: Recommendations of the International Telecommunication Union/Radiocommunication Sector; June 2004.
- Pinson MH, Wolf S. A new standardized method for objectively measuring video quality. *IEEE Transactions on Broadcasting* 2004;50(3):312-22.
- ATIS Technical Report T1.TR.74 - 2001. *Objective video quality measurement using a peak signal-to-noise ratio (PSNR) full reference technique*. Washington, DC: Alliance for Telecommunications Industry Solutions ATIS; 2001.
- Wang Z, Sheikh HR, Bovik AC. Objective video quality assessment. In: Taylor & Francis (Ed.). *Handbook of video databases: design and applications*. Boca Raton (FL): CRC Press; 2003. p. 1041-78.
- ITU-R Recommendation BT.500-11 (06/02). *Methodology for the subjective assessment of the quality of television pictures*. Geneva: Recommendations of the International Telecommunication Union/Radiocommunication Sector; 2002.
- Martelli F, Giordano A, Giansanti D, Morelli S, Macellari V. Evaluation of a new metric for telemedicine video quality assessment. In: *Proceeding of 2nd Annual Meeting on Health Technology Assessment International*. Rome, June 2005. p. 20-2.
- Frankewitsch T, Söhnlein S, Müller M, Prokosch HU. Computed Quality Assessment of MPEG4-compressed DICOM Video Data. *Stud Health Technol Inform* 2005;116:447-52.
- Moore PT, O'Hare N, Walsh KP, Ward N, Conlon N. Objective video quality measure for application to tele-echocardiography. *Med Biol Eng Comput* 2008;46(8):807-13.
- ITU-R Recommendation F.390-4 (07/82). *Definitions of terms and references concerning hypothetical reference circuits and hypothetical reference digital paths for radio-relay systems*. Geneva: Recommendations of the International Telecommunication Union/Radiocommunication Sector; July 1982. Suppressed on 10/10/07 (CACE/435).
- Wolf S, Pinson M. *NTIA Report 02-392 - Video quality measurement techniques*. Boulder, Colorado: US Department of Commerce - National Telecommunications and Information Administration; June 2002. Available from: www.its.bldrdoc.gov/n3/video/documents.htm. Last visited: 11/12/08.
- Wolf S, Pinson MH. *The relationship between performance and spatial-temporal region size for reduced-reference, in-service video quality monitoring systems*. Institute for Telecommunication Sciences, National Telecommunications and Information Administration. 325 Broadway, Boulder, CO 80305, USA.
- Pinson M, Wolf S, Austin PG, Allhands A. *Video quality measurement PC user's manual*. Santa Clara, CA: Intel Corporation; 2002.
- ANSI T1.801.02 - 1996. *American National Standard for Telecommunications - Digital Transport of Video Conferencing/Video Telephony Signals - Performance Terms, Definitions, and Examples*. Washington, DC: American National Standard Institute/Alliance for Telecommunications Industry Solutions ATIS; 1996.
- Lewin M, Xu C, Jordan M, Borchers H, Ayton C, Wilbert D, Melzer S. Accuracy of paediatric echocardiographic transmission via telemedicine. *J Telemed Telecare* 2006;12(8):416-21.
- Kosutic J, Rigby ML, Mijin D, Weatherburn G, Jowett V, Vukomanovic V, Rakic S, Markovic G. Low-bandwidth teleconsultations for patients with complex congenital heart diseases. *J Telemed Telecare* 2007;13(3):113-8.
- Geoffroy O, Acar P, Caillet D, Edmar A, Crepin D, Salvodelli M, Dulac Y, Paranon S. Videoconference pediatric and congenital cardiology consultations: a new application in telemedicine. *Arch Cardiovasc Dis* 2008;101(2):89-93.
- McCrossan BA, Grant B, Morgan GJ, Sands AJ, Craig B, Casey FA. Diagnosis of congenital heart disease in neonates by videoconferencing: an eight-year experience. *J Telemed Telecare* 2008;14(3):137-40.
- Finley JP, Justo R, Loane M, Wootton R. The effect of bandwidth on the quality of transmitted pediatric echocardiograms. *J Am Soc Echocardiogr* 2004;17(3):227-30.
- Barbier P, Alimento M, Berna G, Celeste F, Gentile F, Mantero A, Montericchio V, Muratori M. High-grade video compression of echocardiographic studies: a multicenter validation study of selected motion pictures expert groups (MPEG)-4 algorithms. *J Am Soc Echocardiogr* 2007;20(5):527-36.
- Garcia MJ, Thomas JD, Greenberg N, Sandelski J, Herrera C, Mudd C, Wicks J, Spencer K, Neumann A, Sankpal B, Soble J. Comparison of MPEG-1 digital videotape with

- digitized sVHS videotape for quantitative echocardiographic measurements. *J Am Soc Echocardiogr* 2001;14(2):114-21.
33. Gemmill J. Network basics for telemedicine. *J Telemed Telecare* 2005;11(2):71-6.
 34. Harnett B. Telemedicine systems and telecommunications. *J Telemed Telecare* 2006;12(1):4-15.
 35. Martinez I, Garcia J, Viruete E, Fernandez J. Application parameters optimization to guarantee QoS in e-Health services. *Conf Proc IEEE Eng Med Biol Soc* 2006;1:5222-5.
 36. Hernández C, Alesanco A, Abadia V, García J. The effects of wireless channel errors on the quality of real time ultrasound video transmission. *Conf Proc IEEE Eng Med Biol Soc* 2006;1:6457-60.
 37. Main ML, Foltz D, Firstenberg MS, Bobinsky E, Bailey D, Frantz B, Pleva D, Baldizzi M, Meyers DP, Jones K, Spence MC, Freeman K, Morehead A, Thomas JD. Real-time transmission of full-motion echocardiography over a high-speed data network: impact of data rate and network quality of service. *J Am Soc Echocardiogr* 2000;13(8):764-70.
 38. Giansanti D, Morelli S, Macellari V. Telemedicine technology assessment Part II: tools for a quality control system. *Telemed J E Health* 2007;13(2):130-40.
 39. Giansanti D, Morelli S, Macellari V. Telemedicine technology assessment Part I: setup and validation of a quality control system. *Telemed J E Health* 2007;13(2):118-29.
 40. Giansanti D, Morelli S, Bedini R, Macellari V. *Un'esperienza italiana di controllo di qualità in telemedicina: il progetto eR-METE*. Roma: Istituto Superiore di Sanità; 2004. (Rapporti ISTISAN, 08/23).
 41. Giansanti D, Morelli S, Maccioni G, Guerriero L, Bedini R, Pepe G, Colombo C, Borghi G, Macellari V. A web based health technology assessment in tele-echocardiography: the experience within an Italian project. In: Giansanti D, Morelli S (Ed.). *Digital tele-echocardiography today: successes and failures*. *Ann Ist Super Sanità* 2009;45(4):392-7.