# Analysis of Xvid Video Codec for Clinical Quality Assessment in Tele-Echocardiography

R Leza[1], A Alesanco[1], P Serrano[2], L Ramos[3], A Portolés[4], C Aured[4], M García[4], C Hernández[1], TC Marzo[1], J García[1]

[1] Communications Technologies Group, Aragon Institute for Engineering Research, Zaragoza, Spain
[2]Lozano Blesa Clinic Hospital, Zaragoza, Spain
[3]Santa Marina Hospital, Bilbao, Spain
[4]Miguel Servet Hospital, Zaragoza, Spain

## Abstract

*The use of a tele-echocardiography system requires the use of lossy compression in order to reduce the transmission rates. That's the reason because guaranteeing clinical quality is a highly desired goal. In this paper we introduce two types of tests designed to obtain cardiologists' opinion and translate them to a clinical index. After that, we use this index to establish the recommendable transmission rate when using Xvid video codec in a tele-echocardiography system. Results show that the recommended rate depends on the echocardiography mode: 384 kbps for 2D and M mode, 256 kbps for color Doppler mode and 128 kbps for pulse and continuous Doppler mode.*

## 1. Introduction

Cardiovascular diseases are the main cause of dead in industrialized countries. Diagnostic techniques like echocardiography are very extended due to its low cost and non-invasive characteristics of capturing dynamic images of the heart in real time. Reducing storage requirements and transmitting data in real time are two of the main motivations for applying compression to echocardiographic videos. In order to obtain data rates useful for transmission in a tele-echocardiography project, lossy compression has to be used. However, lossy compression is only acceptable when echocardiographies preserve the relevant diagnostic information. Usually, mathematical distortion indices have been used for evaluating the reconstructed echocardiographic video quality, but these indices do not reflect real quality in clinical interpretation. There are several studies where echocardiography quality has been assessed taking into account cardiologists' opinion [1-2]. Houston et al. [1] used blind test. They divided structures into echocardiography modes and that structures were scored as to whether or not provided sufficient information to be clearly defined. Finley et al. [2] also used blind tests and they rated structures according to their clarity but not to an accurate measurement or diagnostic content. In this paper, we introduce more complete clinical tests mixing blind and semi-blind tests to obtain a clinical quality index. With these tests we are able to obtain a Mean Opinion Score (MOS) index that reflects the cardiologists' opinion about the reconstructed echocardiographic videos. Besides, we analyze Xvid video codec working at different transmission rates using the new proposed clinical index so as to obtain suitable transmission rates for a tele-echocardiography project.

## 2. Methods

In a standard echocardiography examination, four modes of operation can be distinguished. On one hand, 2D mode is the two-dimensional image that represents the heart and its movement and M mode represents a cross section that allows accurate measurements of the heart chambers. Summarizing 2D mode and M mode permit assess the size, thickness and movement of heart structures. On the other hand, color Doppler studies allow to evaluate the blood flow-velocity through the heart, and pulse and continuous mode (we group pulse and continuous like a single mode owing to the similar image characteristics) permits make velocity measures in a specific portion.

MOS tests are used whenever user opinions are needed to evaluate a system. In this paper we are interested in evaluating the effects of using compression methods in clinical diagnoses. We have designed two kind of tests, one blind and the other semi-blind, in order to assess reconstructed echocardiographic videos. The aim of blind test is to obtain the cardiologists' evaluations without knowing the transmission rate used. It is presented in Fig. 1. The aim of semi-blind test is to obtain the cardiologists' evaluations comparing directly the

reconstructed video with the original. Fig. 2 shows the semi-blind test. Thanks to the designed MOS tests we are able to obtain a quality index for each echocardiography mode.

A weighted $MOS_{error}$ was calculated for each mode taking into account the results obtained from the blind and semi-blind tests. It is defined as:

$$MOS_{error} = \frac{\sum_{k=1}^{N} MOS_k^B + \sum_{k=1}^{N} MOS_k^{SB}}{2N}$$

where $N$ is the number of cardiologists who evaluated the echocardiographies and $MOS_k^B$ and $MOS_k^{SB}$ are the results obtained from the *kth* cardiologist in the blind and semi-blind tests respectively. These values are given by

$$MOS_k^B = factor \times \max\left\{\frac{Q_o - Q_r}{Q_o}, 0\right\} \times 100 + (1 - factor)\frac{\sum_{i=1}^{M} \left|\mathrm{sgn}(I_{oi} - I_{ri})\right|}{M} \times 100$$

$$MOS_k^{SB} = factor \times \frac{5 - C}{5} \times 100 + (1 - factor) \times (1 - D) \times 100$$

where

- $Q_o$ is the general quality score of the original video for the mode being evaluated;
- $Q_r$ is the general quality score of the reconstructed video for the mode being evaluated;
- $I_{oi}$ is the interpretation of the *i*th parameter of the original video. This interpretation is translated to a numeric factor in order to be used in the equation. 0 values are assigned if the cardiologist can't answer the question due to the bad video quality. The rest of questions are assigned a value greater than 0.
- $I_{ri}$ is the interpretation of the *i*th parameter of the reconstructed video;
- $M$ is the number of groups that form each echocardiography mode. $M$ is 9 for the 2D mode, 6 for the M mode, 5 for the color Doppler mode and 6 for the pulse and continuous Doppler mode; Table 1 shows the parameters included in the blind tests. Parameters 1, 2, 5, 8, 9, 10, 11, 18 y 19 belong to 2D mode; parameters 1, 2, 3, 4, 6 y 7 belong to M mode; parameters 13, 16, 17, 18 y 19 belong to color Doppler mode and parameters 12, 13, 14, 15, 16 y 17 belong to pulse and continuous Doppler mode;
- $C$ is the measure of similarity between the original video and the reconstructed one (1-5) for the mode being evaluated;
- $D$ is the answer to the Boolean question about the diagnoses (0-YES, 1-NO) for the mode being evaluated;
- *factor* is the weighting coefficient between the measure of quality and the interpretation in the case of blind test and between the measure of similarity and the Boolean question in the case of semi-blind test (*factor* value used is 0.5).

Table 1 Parameters included in blind tests.

| Number | Parameter | Number | Parameter |
|---|---|---|---|
| **1** | **Aortic root (transversal diameter)** | 11 | Pericardial effusion |
| **2** | **Left atrium (transversal diameter)** | 12 | Pulmonary flow (Max. velocity) |
| **3** | **LVTD** | 13 | Mitral flow (Regurgitation) |
| **4** | **LVSD** | 14 | Mitral flow (E wave A wave pattern) |
| **5** | **Global contractility** | 15 | Aortic flow (Max. velocity) |
| **6** | **IVS** | 16 | Aortic flow (Regurgitation) |
| **7** | **PW** | 17 | Tricuspid flow (Regurgitation) |
| **8** | **Mitral valve morphology** | 18 | Septal defects (ASD) |
| **9** | **Aortic valve morphology** | 19 | Septal defects (VSD) |
| **10** | **Tricuspid valve morphology** | | |

The lower the value of $MOS_{error}$, the better the quality evaluation of the reconstructed echocardiographic videos. Table 2 shows the video quality that is defined by dividing the $MOS_{error}$ into four quality groups. For this election we counted with an experienced cardiologists' opinion.

Table 2 Quality groups defined by $MOS_{error}$.

| $MOS_{error}$ | **0** | **12** | **24** | **36** |
|---|---|---|---|---|
| ***General Quality*** | **very good** | **good** | **bad** | **Very bad** |

For evaluation purposes 7 motion echocardiographic videos were selected, which were representative of typical normal and abnormal findings in cardiovascular field. An experienced cardiologist in echocardiography recorded and stored original videos with a digital ultrasound system (Sonosite SonoHeart Elite). Each echocardiography was about three minutes duration and contained the four representative modes in echocardiography and each mode was, approximately, one minute duration 2D and color Doppler modes and 30 seconds duration M and pulse/continuous Doppler mode.

Xvid codec was used to compress original videos to standard ISDN linkages at four different rates (corresponding to 128, 256, 384 and 768 kbps) maintaining original image format (24 bit color, 720x576 pixel) and frame rate (25 fps). Audio was excluded.

Three expert cardiologists in echocardiography viewed in an individually way the uncompressed and compressed

**Analysis of echocardiogram video xxx.avi**

Details of tester
Name:
Date:

**1 General quality score for 2D mode (highlight one number)**

**1** – very bad , **2** – almost tolerable, **3** – tolerable, **4** – good, **5** – excellent

**2 a) General quality score for M mode (highlight one number)**

**1** – very bad , **2** – almost tolerable, **3** – tolerable, **4** – good, **5** – excellent

**2 b) Give an interpretation (circle one interpretation for each parameter)**

| | | | | | |
|---|---|---|---|---|---|
| **M and 2D study** | **Aortic root (transversal diameter)** | | | | Normal |
| | | | | | Dilated |
| | | | | | Dissected |
| | **Left atrium (transversal diameter)** | | | | Normal |
| | | | | | Dilated |
| | | | | | Intra-atrial mass-clot |
| | **Left ventricle** | **LVDD** | | | Normal |
| | | | | | Dilated |
| | | **LVSD** | | | Normal |
| | | | | | Dilated |
| | **Global contractility** | | | | Hyperkinetic |
| | | | | | Normal |
| | | | Depressed | | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | | Asynergy |
| | **Left ventricle wall thickness** | **IVS** | | | Thin |
| | | | | | Normal |
| | | | Left ventricular hypertrophy | Concentric | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | Asymmetric | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | **PW** | | | Thin |
| | | | | | Normal |
| | | | Left ventricular hypertrophy | Concentric | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | Asymmetric | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | **Valve morphology** | **Mitral** | | | Normal |
| | | | | | Abnormal |
| | | **Aortic** | | | Normal |
| | | | | | Abnormal |
| | | **Tricuspid** | | | Normal |
| | | | | | Abnormal |
| | **Pericardial effusion** | | | | Ye |
| | | | | | No |

**3 General quality score for color Doppler mode (highlight one number)**

**1** – very bad , **2** – almost tolerable, **3** – tolerable, **4** – good, **5** – excellent

**4 a) General quality score for pulse and continuous Doppler mode (highlight one number)**

**1** – very bad , **2** – almost tolerable, **3** – tolerable, **4** – good, **5** – excellent

**4 b) Give an interpretation (circle one interpretation for each parameter)**

| | | | | | |
|---|---|---|---|---|---|
| **Doppler study** | **Pulmonary flow** | **Max. velocity** | | Normal | |
| | | | | Stenosis | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | Insufficiency | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | **Mitral flow** | **Systolic** | **Regurgitation** | Yes | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | No | |
| | | **Diastolic** | **E wave A wave pattern** | Normal | |
| | | | | Pseudonormal | |
| | | | | Relaxation Abnormality | |
| | | | | Restrictive | |
| | **Aortic flow** | **Systolic** | **Max. Velocity** | Normal | |
| | | | | Stenosis | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | **Diastolic** | **Regurgitation** | Yes | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | No | |
| | **Tricuspid flow** | **Regurgitation** | | Yes | Light |
| | | | | | Moderate |
| | | | | | Severe |
| | | | | No | |
| | **Septal defects** | **ASD** | | Yes | |
| | | | | No | |
| | | **VSD** | | Yes | |
| | | | | No | |

**5 General quality score for the echocardiography**

**1** – very bad , **2** – almost tolerable, **3** – tolerable, **4** – good, **5** – excellent

**6 Comments**

Fig. 1 Blind MOS test.

**Comparison of echocardiogram video xxx.avi with its original video**

Details of tester
Name:
Date:

**1 a) The measure of similarity between original 2D mode video and the reconstructed one (highlight one number)**

**1** completely different **2** **3** **4** **5** identical

**1 b) Would you give a different diagnosis with the tested video if you had not seen the original 2D mode video? (highlight YES or NO)**

YES NO

**2 a) The measure of similarity between original M mode video and the reconstructed one (highlight one number)**

**1** completely different **2** **3** **4** **5** identical

**2 b) Would you give a different diagnosis with the tested video if you had not seen the original M mode video? (highlight YES or NO)**

YES NO

**3 a) The measure of similarity between original color Doppler mode video and the reconstructed one (highlight one number)**

**1** completely different **2** **3** **4** **5** identical

**3 b) Would you give a different diagnosis with the tested video if you had not seen the original color Doppler mode video? (highlight YES or NO)**

YES NO

Fig. 2 Semi-blind MOS test.

**4 a) The measure of similarity between original pulse and continuous Doppler mode video and the reconstructed one (highlight one number)**

**1** completely different **2** **3** **4** **5** identical

**4 b) Would you give a different diagnosis with the tested video if you had not seen the original pulse and continuous Doppler mode video? (highlight YES or NO)**

YES NO

**5 a) The measure of similarity between original video and the reconstructed one (highlight one number)**

**1** completely different **2** **3** **4** **5** identical

**5 b) Would you give a different diagnosis with the tested video if you had not seen the original video? (highlight YES or NO)**

YES NO

**6 What do you think about the diagnosis quality of the reconstructed video in order to be used in a tele-echocardiography system?**

**a) 2D mode video:** **1** (useless) **2** **3** **4** **5** (excellent)

**b) M mode video:** **1** (useless) **2** **3** **4** **5** (excellent)

**c) Color Doppler mode video:** **1** (useless) **2** **3** **4** **5** (excellent)

**d) Pulse and continuous Doppler mode video:** **1** (useless) **2** **3** **4** **5** (excellent)

**e) All video:** **1** (useless) **2** **3** **4** **5** (excellent)

**7 Comments**

videos to complete the clinical tests. A total of 35 videos were evaluated in a blind condition; 28 pair of videos were evaluated in a semi-blind condition. The evaluation conditions (size, resolution, brightness and contrast of the monitor as well as the illumination in the working environment) were constants along all the process and similar to the clinical routine. Playback was performed using two applications specially developed for echocardiography evaluation. The video viewing order was selected to minimize sources of bias in observers such us fatigue and learning effects.

## 3. Results

$MOS_{error}$ index was estimated to measure the compressed cardiac videos quality using the designed tests. Tables 3-6 show $MOS_{error}$ values obtained for the seven echocardiographic videos divided into its respectively modes at the different transmission rates used for Xvid codec. Mean value was also obtained for each rate.

Table 3 $MOS_{error}$ values for 2D mode.

| Bit rate | V1 | V2 | V3 | V4 | V5 | V6 | V7 | Vmean |
|---|---|---|---|---|---|---|---|---|
| 128 kbps | 50,32 | 40,14 | 28,10 | 43,89 | 56,16 | 39,96 | 48,71 | 43,90 |
| 256 | 40,79 | 29,4 | 13,38 | 37,56 | 26,25 | 34,31 | 34,08 | 30,82 |
| 384 | 28,47 | 11,71 | 12,64 | 29,45 | 18,98 | 22,92 | 25,70 | 21,41 |
| 768 | 16,16 | 12,36 | 11,39 | 10,37 | 12,93 | 12,45 | 15,00 | 12,95 |

Table 4 $MOS_{error}$ values for M mode.

| Bit rate | V1 | V2 | V3 | V4 | V5 | V6 | V7 | Vmean |
|---|---|---|---|---|---|---|---|---|
| 128 kbps | 53,89 | 35,97 | 43,61 | 50,83 | 69,17 | 31,81 | 44,30 | 47,08 |
| 256 | 27,36 | 14,72 | 15,42 | 47,09 | 26,39 | 24,30 | 16,11 | 24,48 |
| 384 | 20,83 | 13,05 | 13,75 | 29,72 | 15,69 | 14,72 | 16,91 | 17,81 |
| 768 | 11,25 | 9,58 | 13,47 | 6,39 | 10,30 | 8,89 | 12,36 | 10,32 |

Table 5 $MOS_{error}$ values for color Doppler mode.

| Bit rate | V1 | V2 | V3 | V4 | V5 | V6 | V7 | Vmean |
|---|---|---|---|---|---|---|---|---|
| 128 kbps | 27,5 | 32,92 | 28,89 | 23,89 | 26,95 | 14,72 | 13,89 | 24,11 |
| 256 | 14,31 | 16,53 | 15,83 | 20,70 | 16,94 | 9,86 | 20 | 16,31 |
| 384 | 12,64 | 11,53 | 9,86 | 7,78 | 15,28 | 15,97 | 11,39 | 12,06 |
| 768 | 8,19 | 5 | 7,5 | 6,67 | 8,33 | 3,33 | 7,77 | 6,68 |

Table 6 $MOS_{error}$ values for pulse and continuous Doppler mode.

| Bit rate | V1 | V2 | V3 | V4 | V5 | V6 | V7 | Vmean |
|---|---|---|---|---|---|---|---|---|
| 128 kbps | 27,08 | 25,14 | 27,22 | 22,22 | 24,44 | 17,92 | 16,11 | 22,88 |
| 256 | 14,44 | 14,44 | 16,53 | 16,22 | 11,67 | 10,70 | 17,64 | 14,52 |
| 384 | 12,5 | 7,92 | 9,72 | 9,58 | 10,97 | 12,36 | 8,37 | 10,20 |
| 768 | 6,25 | 9,31 | 5,56 | 7,78 | 5,83 | 7,64 | 7,92 | 7,18 |

$MOS_{error}$ were then compared across the 4 compression levels with ANOVA statistics. The $MOS_{error}$ obtained for the different rates showed significant differences ($P<0.05$) along all studied modes: 2D mode, M mode, color Doppler mode and pulse and continuous Doppler mode. Interrater agreement was tested with the Pearson correlation. Strong correlation was obtained between the three observers.

## 4. Discussion and conclusions

It is interesting to note (see tables 3-6) that modes 2D and M (normally used to measure heart structures) require a higher transmission rate than both Doppler modes (normally used to measure blood flows), color and pulse/continuous, to obtain a good clinical quality result. This fact could be explained if we take into account that more resolution is needed to emit a diagnosis in modes 2D and M compared to Doppler modes. Looking closer to the modes that measure heart structures, 2D mode obtain worse results than M mode for the same transmission rate. This effect could be understood if we realize that M mode is more static than 2D mode, thus compression algorithm performs better. Finally, in modes used to measure blood flows the differences are not significant indicating than both modes perform equal in the compression process.

For a clinical use, at least a good quality has to be guaranteed in a tele-echocardiography system. In this way, table 7 shows the recommended transmission rates to achieve this target.

Table 7 Recommended transmission rates to obtain good quality.

| 2D | M | Color Doppler | Pulse and Continuous Doppler |
|---|---|---|---|
| **384 kbps** | 384 kbps | 256 kbps | 128 kbps |


## Acknowledgements

This work was supported by project TSI2004-04940-C02-01 from Comisión Interministerial de Ciencia y Tecnología (CICYT) and Fondos Europeos de Desarrollo Regional (FEDER), project PI051416 from Fondo de Investigación Sanitaria (FIS) and European IST Project PULSERS Phase II (IST - 027142).


## References


[1] Houston A, McLeod K, Richens T, Doig W, Lilley S, Murtagh E, Wilson N. Assessment of the quality of neonatal echocardiographic images transmitted by ISDN telephone lines. Heart 1999; 82: 222-5.

[2] Finley JP, Justo R, Loane M, Wootton R. The effect of bandwidth on the quality of transmitted pediatric echocardiograms. Journal of the American Society of Echocardiography 2004; 17:227-230.



Address for correspondence

Raquel Leza
Communications Technologies Group (GTC).
Aragon Institute for Engineering Research (I3A).
University of Zaragoza.
C/ María de Luna, 1. 50018 - ZARAGOZA. SPAIN
rleza@unizar.es