HEVC Encoding for Reproducible Medical Ultrasound Video Diagnosis

Conference Paper · November 2013

CITATIONS
3

READS
41

3 authors:

C. S. Pattichis
University of Cyprus
431 PUBLICATIONS **5,663** CITATIONS
SEE PROFILE

Andreas S. Panayides
University of Cyprus
65 PUBLICATIONS **525** CITATIONS
SEE PROFILE

Marios S Pattichis
University of New Mexico
334 PUBLICATIONS **3,291** CITATIONS
SEE PROFILE

HEVC Encoding for Reproducible Medical Ultrasound Video Diagnosis

A. Panayides[1], M.S. Pattichis[2], and C. S. Pattichis[3]

[1]Department of Electrical and Electronic Engineering, Imperial College, London, UK
[2]Department of Electrical and Computer Engineering, University of New Mexico, Albuquerque, USA
[3]Department of Computer Science, University of Cyprus, Nicosia, Cyprus

***Abstract*— In this study, we investigate how the recently introduced high efficiency video coding (HEVC) standard can be used in the design of reproducible medical ultrasound video diagnosis systems. We introduce a set of clinical criteria that are used to validate the clinical capacity of HEVC encoded atherosclerotic plaque ultrasound videos. The clinical criteria include the assessment of atherosclerotic plaque geometry, morphology, and motion. The efficiency of HEVC compared to prior video coding standards is then demonstrated based on the use of objective ratings. Ultimately, the goal is to support diagnostically lossless compression at the clinically acquired video resolution and frame rates for storage, streaming, and follow-up visualization.**

Keywords: HEVC, video quality assessment, m-health, medical video communications, ultrasound video coding.

I. INTRODUCTION

The recently introduced High Efficiency Video Coding (HEVC) [1] standard is expected to play a significant role in the success of future mobile-health (m-health) medical ultrasound video systems and services [2]. Thus far, standard practice in m-health medical video communication systems dictated the spatial and temporal down-sampling of the original ultrasound video, hence compromising clinical quality [2], [3]. This was attributed to the inability of wireless systems to accommodate ultrasound video communications at the acquired video resolution and frame rate (bounded by the upload data rates of 3G and early 3.5G-HSDPA wireless channels) and the inefficiency of prior standards to provide for real-time, high-resolution encoding. The former has been timely addressed via the introduction of 4G (and beyond) wireless technologies [4], while the latter by the release of the new HEVC standard. New wireless networks now provide upload data transfer rates that can support low-delay and higher quality video communications, thus far feasible only using wired infrastructure.

HEVC was specifically designed for (beyond) high-definition (HD) video coding and therefore is suitable for encoding ultrasound video at the ultrasound machine's resolution and frame rate. Increased coding efficiency compared to its predecessor, H.264/AVC standard, allows 35.4% bitrate gains for equivalent objective, Peak Signal-to-Noise (PSNR) ratings, while it extends to 49.3% based on subjective, perceptual evaluation [5]. For real-time communication, HEVC is the first standard to support parallel processing tools that can provide real-time encoding for HD video transmission.

As recently highlighted in [2], using the clinically-acquired spatial resolution and frame rate will support wider adoption of videos in clinical practice. Such advancement was not feasible in the previous decade due to wireless infrastructure and video coding standards limitations. Indicatively, in recent studies reported by our group [6], [7], limited frame rate prohibited efficient assessment of clinical motion for atherosclerotic plaque ultrasound video, which was excluded from the clinical assessment criteria. The reduced frame rate -of fifteen frames per second- did not compromise the assessment of the degree of the artery stenosis (carotid disease) and the classification of the plaque's type. Instead, plaque morphology characteristics assessment can be compromised by lower video resolution (such as QCIF (176x144) and CIF (352x288)) and/or the selected compression level [7]. As a result, prior to HEVC, clinical capacity of the communicated video, was in most cases a-priori bounded by the source encoding parameters, resulting in analogous modifications of the clinical assessment process, deviating from the protocol followed during in-hospital examinations. Consequently, there is an imperative need to investigate medical video transmission at the acquired resolution and frame rate and provide for m-health video systems that will rival the quality of in-hospital examinations.

Materializing the aforementioned m-health systems and services requires the development of clinical video quality assessment (c-VQA) metrics that will reliably evaluate the diagnostic capacity of the compressed ultrasound video. Efficient c-VQA metrics need to consistently correlate to the medical expert's ratings. In other words, to be able to reproduce the diagnosis performed by the medical experts. In order to achieve this, a clinically established protocol for each assessed medical video modality should be in place. Ideally, this protocol should reflect procedures and criteria followed during typical ultrasound examinations.

We highlight the contributions of this study in the following areas:

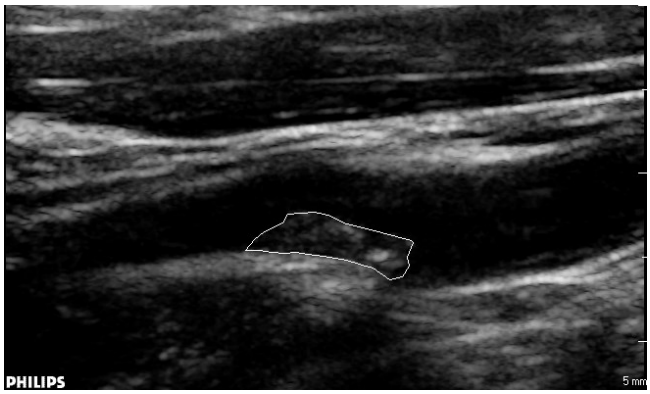


Fig. 1. Atherosclerotic plaque ultrasound video frame. The atherosclerotic plaque d-ROI is outlined using the automated segmentation algorithm described in [8].

1) *Investigating the efficiency of the HEVC standard for medical video communications*: An ultrasound video data set of five videos is used to estimate the bitrate gains achieved by HEVC when compared to prior video coding standards based on objective evaluation (PSNR and BD-rate metric).
2) *Reproducible clinical video assessment including motion quality assessment*: We perform a comprehensive evaluation of the compressed video that aims to support reproducibility to the level of in-hospital examinations. This is the first time that clinical motion is assessed using high frame rate encoding that does not compromise the clinical ultrasound's video motion patterns. The employed clinical assessment protocol reflects current practices for in-hospital examination.

The rest of the paper is organized as follows: Section II introduces the criteria for reproducible, clinical video quality assessment for atherosclerotic plaque ultrasound video. Section III describes the methodology while Section IV summarizes the results. Finally, Section V provides concluding remarks and future work.

II. Clinical Criteria for Reproducible Clinical Video Quality Assessment

During an ultrasound examination of the common carotid artery, the medical expert and/or vascular technician is mainly interested in visualizing key areas and that will allow him/her to evaluate the severity of the disease and/or in the case of follow-up examination, the progression of the carotid disease. These clinically sensitive regions can be described as diagnostic region(s)-of-interest (d-ROI) (see Fig. 1). For atherosclerotic plaque ultrasound video we summarize these d-ROI:

1) *Atherosclerotic plaque region:* The primary d-ROI is used to determine the plaque's type by assessing the plaque's morphology and texture (see Fig. 1). Moreover, it provides for assessing the plaque's stability, which corresponds to the likelihood of a plaque's rupture leading to a stroke incident. Plaque's stability is determined by evaluating motion patterns both on the plaque's boundaries but also within plaque motions, where different plaque components may move in opposite directions. It is important to note here that unstable plaques are often subject to different within plaque motions (concordant/discordant). On the other hand, stiff plaques tend to be safer.
2) *Near and far wall regions for visualizing wall motions and stenosis:* Visualizing the artery walls supports the assessment of the degree of stenosis. Significant stenosis can be associated with stroke events. In addition to stenosis, the motion differences between the arteries and the plaques can be used to recognize plaque instability.
3) *ECG region for visualizing ECG waveform:* The ECG region (subject to ultrasound machine's display options), supports clinical assessment during the cardiac cycle. In other words, the diagnosis involves recording how stenosis and motion patterns of different plaque components change during the cardiac cycle.

Based on the afore-described clinically sensitive regions, the following clinical video quality assessment criteria are used for establishing the reproducibility of the diagnosis:

1) The *degree of the artery stenosis*: the percentage of the artery that is blocked by the plaque's presence, obscuring blood flow.

2) The *plaque's morphology*: the medical expert examines the plaque morphology and geometry as described in [9]. From the examination, the doctor can determine the plaque's type and infer the possible composition of the plaque.

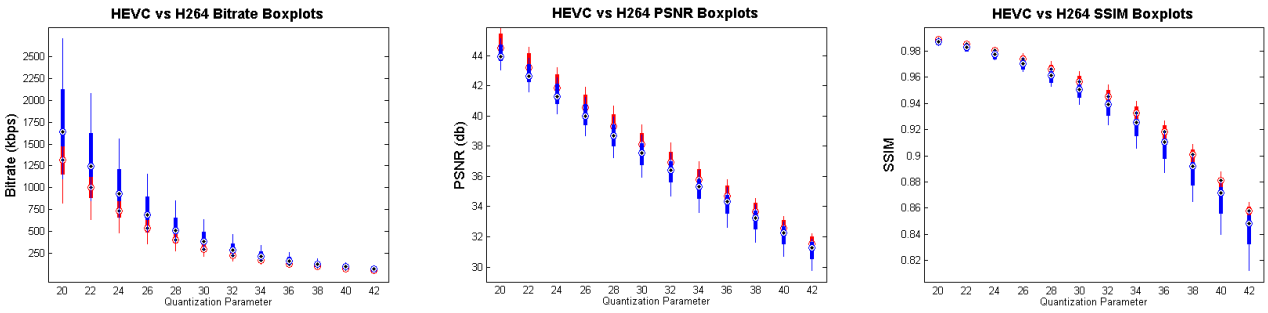
3) The *plaque's motion characteristics*: the medical expert evaluates motions patterns of different plaque components. Based on the motions patterns, the medical expert can determine whether the plaque motion is *concordant* or *discordant* as described in [10].

III. Methodology

This section describes the undertaken methodology employed for obtaining the objective video quality assessment (VQA) results used for video coding standard's comparison and the clinical assessment procedure used to validate the added clinical value of HEVC compared to prior standards.

A. HEVC Encoding of Ultrasound Video Data Set

HEVC encoding setup is summarized in Table I. The maximum coding unit (CU) width and height, partition depth, and search range is set to 64. New coding tools introduced in HEVC including sample adaptive offset (SAO), asymmetric motion partitioning (AMP), transform skip (TS), and temporal motion vector prediction (TMVP) are enabled. In other words, all key coding tools found in HEVC main profile which are responsible for HEVC increased coding efficiency are used. On the other hand, parallel processing tools such as tiles or wavefront parallel processing (WPP) are disabled, as coding speed and



(a) (b) (c)

Fig. 2. HEVC vs H264/AVC boxplots for the entire data set and 12 different QPs. HEVC measurements appear in red, whereas H.264/AVC in blue. (a) Bitrate requirements, (b) PSNR ratings, and (c) SSIM scores. HEVC lowers bitrate requirements while increasing PSNR and SSIM scores.

TABLE I. HEVC ENCODING PARAMETERS AND CODING TOOLS SETUP

Parameter	Value	Coding Tools	ON/OFF
Search Range	64	Sample Adaptive Offset	ON
Max CU Width/Height	64	Asymmetric Motion Partitioning	ON
Max Partition Depth	64	Transform Skip	ON
GOP	8	Temporal Motion Vector Prediction	ON
Intra Period	48	Parallel Processing	OFF

[a] CU: Coding Unit, GOP: Group of Pictures.
[b] Encoding setup for the all investigated video coding standards appears in [5].

TABLE II. AVERAGE BITRATE GAINS OF HEVC STANDARD COMPARED TO PRIOR VIDEO CODING STANDARDS.

	H.264/AVC	H.263	MPEG-4	MPEG-2
HEVC	33.1%	54.5%	58.6%	71%

[a]Results obtained using the BD-Rate algorithm for a data set of five atherosclerotic plaque ultrasound videos and 12 rate points for each video.

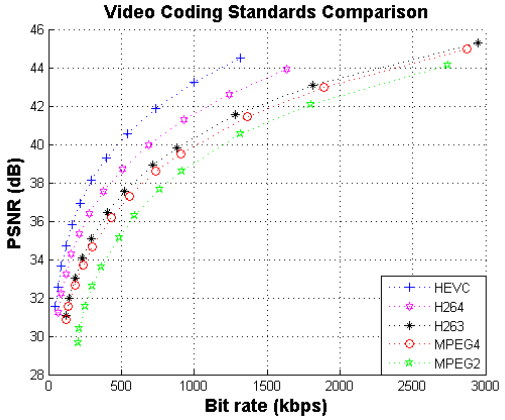


Fig. 3. Rate-distortion curves for the investigated video coding standards for a single video at a video resolution 560x416@50 fps. The new HEVC standard significantly lowers bitrate requirements witout comprimising objective quality (based on PSNR). The trend was the same for all examined videos.

complexity are outside the scope of the current study. For more information regarding the performance of each new coding tool and impact on the overall coding efficiency, the reader is referred to [1], [5]. In [5], the encoding setup of H.264/AVC, H.263, MPEG-4 and MPEG-2 is also addressed. All video coding standards have been tuned for optimum performance, so that a fair evaluation is provided. Here, it is important to highlight that HEVC is the first video coding standard that goes beyond the macroblock (MB), 16x16 block, which was the core coding unit in all previous video coding standards. It is replaced by the coding tree block (CTB) in HEVC, which has typically higher values, up to 64x64.

For each video, we generate 12 different rates by varying the Quantization Parameter (QP) from 20 to 42. To compare the achieved rate-distortion performance with other clinical video coding standards, we use the same QP range for H.264/AVC, while a QP range of 2 to 31 is used for H.263, MPEG-4, and MPEG-2.

The methods are compared for five atherosclerotic plaque ultrasound videos, with a spatial resolution of 560x416 at 50 frames per second (fps) spanning over 4 cardiac cycles. The spatial resolution represents the default output resolution, which the particular ultrasound machine used in this study used for storing uncompressed video sequences. The frame rate has been chosen to be sufficiently high to support motion evaluation. Given that the video is expected to be periodic over the cardiac cycle, the use of four cardiac cycles will eliminate potential, transient artifacts that may appear over a single cardiac cycle.

The Peak Signal-to-Noise Ration and the Structure Similarity Index (SSIM) VQA metrics are used to measure the objective quality of the encoded videos [11]. The bjontegaard-rate (BD-rate) algorithm [12]is used to quantify the average bitrate gains achieved by the recently introduced HEVC video coding standard.

B. Clinical Criteria for Reproducible HEVC Encoding

The medical expert was asked to blindly evaluate the compressed videos without knowledge of the encoding method. To support reproducibility, we use the three clinical criteria described in Section II, with individual scores ranging from one (lowest score) to five (highest score). A rating of five is reserved for clinical quality that is equivalent to that of the uncompressed ultrasound video that is acquired at the original spatiotemporal resolution of the ultrasound device. A rating of four corresponded to acceptable loss of minor details that did not prevent the

TABLE III. CLINICAL EVALUATION FOR THE SINGLE VIDEO DEPICTED IN FIG. 1. A CLINICAL SCORE FROM 1 TO 5 IS ASSIGNED FOR THE THREE CLINICAL CRITERIA DESCRIBED IN SECTION II. HERE, FOR THREE QPS: 42, 32, AND 28.

	Stenosis			**Morphology**			**Motion**		
Quantization Parameter	**42**	**32**	**28**	**42**	**32**	**28**	**42**	**32**	**28**
HEVC	5	5	5	4	5	5	3	5	5
H.264/AVC	5	5	5	3	4	5	4	5	5
MPEG-4	3	5	5	2	5	5	2	4	5
H.263	3	4	5	3	5	5	3	5	4
MPEG-2	3	4	4	2	4	4	2	3	3

TABLE IV. AVERAGE CLINICAL RATINGS FOR THREE HEVC ENCODED VIDEOS USING DIFFERENT QUANTIZATION PARAMETERS. FOR QP OF 27, CLINICAL QUALITY IS IDENTICAL TO THE ORIGINAL, UNCOMPRESSED VIDEO.

	Quantization Parameter		
Clinical Criteria	**37**	**32**	**27**
Stenosis (structure)	5	4.67	5
Morphology (type/ulceration)	5	5	5
Motion (surface motion, concordant/discordant)	4.67	5	5

medical expert to reach a confident diagnosis. At the opposite scale, a rating of one marked the evaluated video unworthy of any clinical usage. The rating scale is described in more detail in [6]. Here, it has been extended to include motion quality assessment.

Formally, reproducible video encoding can be achieved for:

$$C_i = 5 \text{ for all } i,$$

where C_i denotes the i-th clinical criterion and a rating of 5 signifies that the clinical evaluation is equivalent to that of the uncompressed, original video without any spatiotemporal conversion. Furthermore, the clinical criteria need to address ***geometric measurements (e.g., stenosis), image appearance (e.g., plaque morphology), and motion.***

All evaluations were performed using laptop equipment, at the original video resolution with screen brightness set to maximum, in a mildly dark environment, where windows were shut. The medical doctor's eyes had adjusted to the current lighting conditions. The laptop's spatial resolution was 1920x1080. The viewing distance was approximately one meter. Overall, the viewing conditions were compatible with a routine clinical exam.

IV. RESULTS & DISCUSSION

A. Video Coding Standards Comparison for Medical Ultrasound Video Communications

1) Objective Assessment

Table II depicts the bitrate demands reductions achieved by the new, HEVC standard, when compared to previous video coding standards. Compared to its predecessor, H.264/AVC standard, HEVC reduces bitrate requirements approximately 33.1%. Bitrate gains are higher for H.263 and MPEG-4 standards, extending to bitrate gains of 54.5% and 58.6%, respectively. When compared to earlier MPEG-2 standard, HEVC lowers bitrate requirements by as much as 71%.

Fig. 2 verifies the afore-mentioned observations by recording the bitrate requirements, PSNR, and SSIM ratings of the examined data set, for the whole range of the employed quantization parameters. Fig. 2(a) shows the bitrate demands of HEVC and H.264/AVC encodings. As evident in the depicted boxplots, HEVC -appearing in red- involves lower bitrate in all investigated cases. Similarly, Fig. 2(b) displays the slightly higher PSNR ratings achieved by HEVC when compared to the H.264/AVC standard. In Fig. 3(c), the corresponding SSIM measurements verify the coding efficiency of the new video coding standard compared to its predecessor.

Fig. 3 shows typical rate-distortion curves of the investigated video coding standards for the single video depicted in Fig. 1. As evident in Fig. 3, HEVC achieves higher PSNR scores (objective quality) for significantly lower bitrate. The trend is the same for all investigated videos as discussed for Fig. 2.

2) Clinical Evaluation

Table III summarizes the clinical ratings assigned by a neurovascular expert to a representative sample of the ultrasound video instances of the video depicted in Fig. 1. The video instances were derived by encoding the video using different video coding standards and compression levels. In contrast to objective ratings, the clinical evaluation is currently the only reliable assessment of the clinical capacity of the compressed and/or communicated medical videos. As documented widely in the literature for conventional videos, PSNR fails to achieve the desirable level of correlation to the perceived video quality [13]. This was also the case reported in [6] for atherosclerotic plaque ultrasound video. Therefore, in the absence of efficient c-VQA methods, reproducible diagnosis for processed videos is assessed by means of clinical evaluation.

In addition to the questionable use of PSNR for medical video evaluation, HEVC coding efficiency is significantly enhanced based on perceptual quality. More specifically, the target goal set during the development phase of HEVC for halving bitrate requirements compared to H.264/AVC for equivalent perceptual quality, is achieved when bitrate gains are estimated based on visual quality [5]. Here, we present preliminary clinical evaluation results that verify that the clinical capacity of HEVC encoded videos is higher than that

of rival video coding standards for comparable objective, PSNR ratings.

As depicted in Table III, HEVC achieves the best clinical ratings. It is the only standard that achieves the highest rating of 5 for all clinical criteria, for a QP of 32. H.264/AVC and MPEG-4 attain diagnostically lossless ratings of 5 in all assessed clinical criteria for a QP of 28. Lower compression levels (lower QP) is required for H.263 and MPEG-2 to reach clinical performance identical to the original, uncompressed video. Using a frame rate of 50 fps is overwhelming for the assessment of clinical motion as the medical expert noted.

While this is a single measurement that remains to be verified by additional experimentation, the preliminary results show that it is highly possible for HEVC to achieve higher bitrate gains based on perceptual, clinical video quality.

B. Clinical Evaluation for Diagnostically Lossless Compression

Table IV presents preliminary results investigating HEVC compression levels that support reproducible diagnosis. The depicted results correspond to mean opinion scores averaged over three videos. For this limited sample, diagnostically lossless compression is achieved for a QP of 27. For higher compression, clinical ratings fall below the optimum rating of 5, for different clinical criteria assessment in a subset of the investigated cases. However, a higher number of ultrasound videos and a more dense compression step size are more appropriate to determine the diagnostically lossless compression thresholds. This is a matter of ongoing work, also including multiple medical experts.

Here, it is important to note that we seek only ratings of 5, which correspond to clinical quality identical to that of the original, uncompressed video. For medical video communications, only the highest level of clinical quality during compression is desirable, so that if minor errors appear during transmission (e.g. packet drops), reproducible clinical quality is still maintained. As a result, a clinical rating of 4 is not acceptable for determining the diagnostically lossless threshold.

V. Concluding Remarks

The use of HEVC supports clinical video encoding for reproducible diagnosis. Based on three independent clinical criteria that reflect the standard in-hospital evaluation procedure, and evaluate plaque geometry, morphology, and motion, the highest ratings can be achieved using the HEVC main profile. Increased coding efficiency provides for significant bandwidth demands reductions. Indicatively, for this particular study, HEVC achieves bitrate gains of 33.1% compared to the H.264/AVC standard. Realizable compression levels for diagnostically lossless encoding are well within today's 3.5G upload data rates and therefore well suited for wireless medical video communications.

Future work includes validation with multiple experts over a larger database of ultrasound videos. Moreover, there is a need to investigate the efficiency of HEVC for m-health medical video communications over LTE and LTE-Advanced cellular networks, both in terms of clinical quality and network's resources utilization.

Acknowledgment

This work was supported by the Marie Curie Actions – Intra European Fellowships (IEF), FP7-PEOPLE-2011-IEF call, 301476, under the "Diagnostically Robust Ultrasound Video Transmission over Emerging Wireless Networks"-DRIVEN project.

References

[1] G. J. Sullivan, J.-R. Ohm, W.-J. Han, and T. Wiegand, "Overview of the High Efficiency Video Coding (HEVC) Standard," *IEEE Trans. Circuits and Systems for Video Tech*., vol.22, no.12, pp.1649-1668, Dec. 2012.

[2] A. S. Panayides, M. S. Pattichis, and C. S. Pattichis, "Mobile-Health Systems Use Diagnostically Driven Medical Video Technologies [Life Sciences]," *Signal Processing Magazine, IEEE*, vol.30, no.6, pp.163-172, Nov. 2013. DOI: 10.1109/MSP.2013.2276512.

[3] A. Panayides, M.S. Pattichis, C.S. Pattichis, and A. Pitsillides, "A Tutorial for Emerging Wireless Medical Video Transmission Systems [Wireless Corner]," *IEEE Antennas & Propagation Magazine*, vol. 53, no. 2, April 2011, pp. 202-213.

[4] Rysavy Research, LLC, "Mobile Broadband Explosion: The 3GPP Wireless Evolution," Aug. 2013. Available: http://www.4gamericas.org/.

[5] J.-R. Ohm, G. J. Sullivan, H. Schwarz, T. K. Tan, and T. Wiegand, "Comparison of the Coding Efficiency of Video Coding Standards – Including High Efficiency Video Coding (HEVC)," *IEEE Trans. Circuits and Systems for Video Tech.*, vol.22, no.12, pp.1669- 1684, Dec. 2012.

[6] A. Panayides, M.S. Pattichis, C.S. Pattichis, C.P. Loizou, M. Pantziaris, and A. Pitsillides, "Atherosclerotic Plaque Ultrasound Video Encoding, Wireless Transmission, and Quality Assessment Using H.264", *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, no. 3, pp.387-397, May 2011.

[7] A. Panayides, Z. Antoniou, Y. Mylonas, M. S. Pattichis, A. Pitsillides, and C. S. Pattichis, "High-resolution, low-delay, and error-resilient medical ultrasound video communication using H.264/AVC over mobile WiMAX networks," *IEEE J. Biomed. Health Informat.*,vol. 17, no. 3, pp. 619–628, May 2013.

[8] C.P. Loizou, C.S. Pattichis, M. Pantziaris, and A. Nicolaides, "An integrated system for the segmentation of atherosclerotic carotid plaque," *IEEE Trans. on Inform. Techn. in Biomedicine*, vol. 11, no. 5, pp. 661-667, Nov. 2007.

[9] E. Kyriacou, M. S. Pattichis, C. S. Pattichis, A. Mavrommatis, C. I. Christodoulou, S. Kakkos, and A. Nicolaides, "Classification of atherosclerotic carotid plaques using morphological analysis on ultrasound images," *Applied Intelligence*, vol. 30, no. 1, pp. 3-23, Feb. 2009.

[10] H. Nasrabadi, M.S. Pattichis, P. Fisher, A.N. Nicolaides, M. Griffin, G.C. Makris, E. Kyriacou, and C.S. Pattichis, "Measurement of motion of carotid bifurcation plaques," *Bioinformatics & Bioengineering (BIBE), 2012 IEEE 12th International Conference on* , vol., no., pp.506,511, 11-13 Nov. 2012 doi: 10.1109/BIBE.2012.6399729.

[11] Metrix_mux objective video quality assessment software, Available: http://foulard.ece.cornell.edu/gaubatz/metrix_mux/.

[12] G. Bjøntegaard, "Improvements of the BD-PSNR model," ITU-T SG16 Q.6 Document, VCEG-AI11, Berlin, Germany, July 2008.

[13] K. Seshadrinathan, R. Soundararajan, A. C. Bovik, and L. K. Cormack, "Study of subjective and objective quality assessment of video," *IEEE Trans. Image Process.*, vol. 19, no. 6, pp. 1427–1441, Jun. 2010.