

Ultrasound Video Despeckle Filtering for High Efficiency Video Coding in M-Health systems

A. Panayides^{1,2}, C. P. Loizou³, M.S. Pattichis⁴, E. Kyriacou⁵, C. N. Shizas¹, A. N. Nicolaides⁶, and C. S. Pattichis¹

Abstract— Mobile-Health (m-health) medical video communication systems need to communicate ultrasound video of adequate diagnostic quality for remote diagnosis and care. In this study, we demonstrate how ultrasound video noise reduction, employed as a pre-processing step, used in conjunction with High Efficiency Video Coding (HEVC), can significantly improve diagnostic quality at considerably lower bitrate demands than conventional H.264/AVC source encoding. Three video despeckle filters were investigated, namely: linear filtering (DsFlsmv), hybrid median filtering (DsFhmedian), and speckle reducing anisotropic diffusion (DsFsrad). The methods were tested using ten atherosclerotic plaque ultrasound videos. Speckle filtering was applied at a frame-level basis. Experimental results were obtained using both subjective (by two clinical experts) and objective video quality assessment. The best achieving despeckle filter was DsFlsmv followed closely by DsFhmedian. In both cases, the filters provided for increased clinical performance while associated with significant bitrate demands reductions.

I. INTRODUCTION

Mobile-health (m-health) medical video communication systems have been extensively investigated over the past decade, primarily due to advances in the associated fields of video compression and wireless networks [1]. The H.264/AVC video coding standard introduced network independence design and provided for increased coding efficiency and error resiliency. On the other hand, 3G and beyond wireless networks established wide and reliable coverage and accommodated upload data transfer speeds that could be used to communicate medical video [2].

Despite these advances, there has been limited adoption in clinical practice. The lack of use in clinical practice may be due to the fact that the diagnostic capacity of the communicated ultrasound videos could not rival the standards of in-hospital care, especially due to the absence of effective clinical video quality assessment methods [3].

Ultrasound video at the clinically-acquired resolution and frame rates could not be communicated over existing wireless infrastructure using current compression methods. Moreover, multiplicative noise (speckle) present in ultrasound videos was amplified by the error-prone nature of wireless channels. A decade later, the new HEVC standard [4], used in 3.5G and 4G wireless networks, promises for low-delay systems that can be used to communicate high-resolution and high frame rate video.

In this study, we propose the use of despeckle filtering as a pre-processing step in m-health ultrasound video communication systems. Our goal is to perform noise removal in ultrasound video of the common carotid artery (CCA) to improve clinical capacity as well as investigate the associated reductions in bitrate demands. We also perform an initial comparison of the bitrate demands for H.264/AVC and HEVC, in an attempt to demonstrate the significant bitrate savings due to the use of HEVC. For despeckling, we use three different filters: a) Linear, b) Hyper median, and c) Speckle reducing anisotropic diffusion. The methods are tested on a data set composed of ten ultrasound videos of the CCA acquired at video resolution of 560x416, at 40 frames per second (fps), and duration of 10 seconds.

Despeckle filtering has also been considered for ultrasound video denoising by our group in previous studies with very promising results [5]-[6]. To the best of our knowledge, this is the first attempt to assess the diagnostic implications of despeckle filtering in ultrasound video communications.

The rest of the paper is organized as follows. Section II provides a brief overview of the investigated despeckle filters. Section III summarizes the HEVC standardization initiative. In Section IV, we summarize the methodology, while Section V discusses the achieved results. Finally, we give some concluding remarks and future work in Section VI.

II. VIDEO DESPECKLE FILTERS

In this section we highlight the theoretical background of the three investigated video despeckle filtering methods. A thorough overview and implementation details appears in [5]. For each filter, the filtering window size and the number of iterations were selected based on the clinical evaluation made by two medical experts.

A. Linear Despeckle Filter (DsFlsmv)

This filter utilizes the first order statistics, namely the variance and the mean of a pixel neighborhood and may be described with a multiplicative noise model [7]. Hence the

¹A. Panayides, C.N. Schizas, and C.S. Pattichis are with the Dept. of Computer Science, University of Cyprus, Nicosia, Cyprus (tel:+357-22892697, e-mail: {panayides, schizas, pattichi}@ucy.ac.cy).

²A. Panayides is with the Dept. of Electrical and Electronic Engineering, Imperial College, London, UK (email: a.panagidis@imperial.ac.uk).

³C. P. Loizou is with the Department of Computer Science, Intercollege, 3507 Limassol, Cyprus (e-mail: loizou.c@lim.intercollege.ac.cy).

⁴M.S. Pattichis is with the Dept. of Electrical and Computer Engineering, University of New Mexico, Albuquerque, USA (e-mail: pattichis@ece.unm.edu).

⁵E. Kyriacou is with the Dept. of Computer Science and Engineering, Frederick University Cyprus, Lemesos, Cyprus (email: e.kyriacou@frederick.ac.cy).

⁶A. N. Nicolaides is with the Vascular Screening and Diagnostic Centre, London, UK, and also with the Cyprus Cardiovascular Disease Educational Research Trust, 2368 Nicosia, Cyprus (e-mail: anicolaides1@gmail.com).

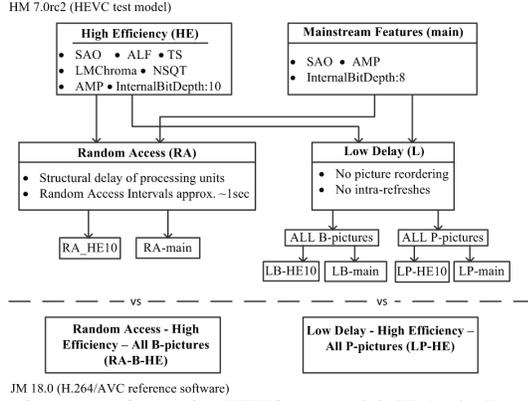


Fig. 1. Summary of emerging HEVC test model (HM 7.0rc2) encoding schemes and corresponding H.264/AVC reference software (JM 18.0) for comparison evaluation.

ALF: Adaptive Loop Filtering, LMChroma: Chroma from Luma intra prediction mode, NSQT: Non-Square Transforms, AMP: Asymmetric motion (see [4], [11] for details).

algorithms in this class may be traced back to the following equation:

$$f_{i,j} = \bar{g} + k_{i,j}(g_{i,j} - \bar{g}) \quad (1)$$

where $f_{i,j}$ is the estimated noise-free pixel value, $g_{i,j}$ is the noisy pixel value in the moving window, \bar{g} is the local mean value of an $N_1 \times N_2$ region surrounding and including pixel $g_{i,j}$, $k_{i,j}$ is a weighting factor, with $k \in [0..1]$, and i,j , are the pixel coordinates. The factor $k_{i,j}$, is a function of the local statistics in a moving window and can be found in the literature [5], [7] as:

$$k_{i,j} = (1 - \bar{g}^2 \sigma^2) / (\sigma^2 (1 + \sigma_n^2)) \quad (2)$$

The values σ^2 and σ_n^2 , represent the variance in the moving window and the variance of noise in the whole image respectively. The noise variance σ_n^2 , may be calculated for the logarithmically compressed image, by computing the average noise variance over a number of windows with dimensions considerable larger than the filtering window [7]. The moving window size is 5x5 and the number of iterations two.

B. Hybrid Median Filtering (DsFhmedian)

The filter DsFhmedian [8], which is an extension of the median filter, computes the average of the outputs generated by median filtering with three different windows (cross shape window, x-shape window and normal window). Here, a 5x5 size moving window was used with the number of iterations applied to each video frame equal to two.

C. Speckle Reducing Anisotropic Diffusion Filtering (DsFsrad)

Speckle reducing anisotropic diffusion was proposed in [9]. It is based on setting the conduction coefficient in the diffusion equation using the local frame gradient and the frame Laplacian. The DsFsrad uses two seemingly different methods, namely the Lee and the Frost diffusion filters [5], [7]. A more general updated function for the output image by extending the partial differential equation versions of the despeckle filter can be found in [5]:

$$f_{i,j} = g_{i,j} + \frac{1}{n_s} \text{div} \left(c_{srad} (|\nabla g|) \nabla g_{i,j} \right) \quad (3)$$

The diffusion coefficient for the speckle anisotropic diffusion, $c_{srad} (|\nabla g|)$, is derived [7] as:

$$c_{srad} (|\nabla g|) = \frac{\frac{1}{2} |\nabla g_{i,j}|^2 - \frac{1}{16} (\nabla^2 g_{i,j})^2}{(g_{i,j} + \frac{1}{4} \nabla^2 g_{i,j})^2} \quad (4)$$

It is required that $c_{srad} (|\nabla g|) \geq 0$. The above instantaneous coefficient of variation combines a normalized gradient magnitude operator and a normalized Laplacian operator to act like an edge detector. High relative gradient magnitude and low relative Laplacian indicates an edge. The DsFsrad filter utilizes speckle reducing anisotropic diffusion after (3) with the diffusion coefficient, $c_{srad} (|\nabla g|)$ in (4).

III. HIGH EFFICIENCY VIDEO CODING STANDARD (HEVC)

The Joint Collaborative Team on Video Coding (JCT-VC), established by the ITU-T VCEG and the ISO/IEC MPEG groups issued a joint call for proposals (CfP) in 2010 to initiate the process of the new video coding standard, termed high efficiency video coding (HEVC). Individual activities had been already launched since 2005. A test model under consideration (TMuC) software was subsequently created in April 2010, composed of the best performing proposals in response to the CfP, before the HEVC test model 1 (HM 1) was generated in October 2010. Coding tools with marginal coding efficiency gains with respect to the incurred complexity were removed. In HM, coding tools development and performance evaluation was primarily performed using two coding modes, namely High Efficiency (HE) and Low Complexity (LC). However, in the new standard, expected to be finalized in January 2013, is likely to define a single profile termed Main [5].

The emerging HEVC standard is based on extended modes of within-frame (intra) and inter-frame (inter) prediction, using enhanced motion vector and motion compensation prediction, followed by residual linear transform, quantization and entropy coding. In other words it maintains the hybrid block-based structure found in all video coding standards since H.261. However, all features and coding tools borrowed from H.264/AVC have undergone significant refinement to reduce complexity, enhance compression efficiency and lower memory requirements. Key to the success of the new standard is the introduction of a new block structure, the coding tree unit (CTU) which allows partitioning of a picture to larger sub-blocks of variable luma size (up to 64x64 luma samples), which replaces the macroblock structure found in previous standards (fixed 16x16 luma samples). A CTU is parted of luma and chroma coding tree blocks (CTBs), and CTBs are further split into smaller coding blocks (CBs) following a tree structure and quadtree signaling. A coding unit (CU) is typically composed of one luma CB and two chroma CBs. A CU also defines the prediction units (PU), the decision whether to adopt intra or inter picture prediction, and transform units (TUs), which describes the block transform coding of the prediction residual. The CBs can then have identical or smaller in size

prediction blocks (PBs) and transform blocks (TBs). A detailed overview and performance evaluation of the emerging HEVC standard can be found in [4], [11].

IV. METHODOLOGY

A. HEVC and H.264/AVC experimental setup

Here, we adopt the common test conditions used during HEVC development to perform coding efficiency comparisons between HEVC and H.264/AVC standards. HM 7.0rc2 used during the experiments of this study defines two primary common test conditions, namely (a) *high efficiency* (HE) and (b) *main*. The former includes more advanced coding tools which are often computationally intensive, while the latter incorporates the mainstream features of HEVC. In addition to these two configurations, performance evaluation for different application further defines *random access* and *low delay* schemes. For the low delay scenarios, in addition to the current default *B-frames* mode, only *P-frames* encoding is also defined. These common test conditions are summarized in Fig. 1.

JM H.264/AVC reference software details two configuration settings to be used for comparison purposes with HEVC. More specifically, random access and low delay schemes, which both employ the high profile, for high efficiency performance. The reasoning is that the maximum coding efficiency achieved by H.264/AVC standard is sought, for a more realistic evaluation of the new standard. For each case, random access and low delay, only B-frames and P-frames are used, respectively.

B. Despeckle Filtering and H.264/AVC and HEVC video coding standards

Despeckle filtering is performed as a pre-processing step to reduce video noise in the ultrasound video. Despeckle filtering is applied at the frame-level, for the entirety of frames parting the ultrasound video. Here, we investigate three different filters, namely a) *Linear Despeckle Filter (DsFlsmv)*, b) *Hybrid Median Filtering (DsFhmedian)* and c) *Speckle Reducing Anisotropic Diffusion Filtering (DsFsrad)*.

Following ultrasound video despeckle, we investigate the associated bitrate demands of the resulting despeckled ultrasound videos. More specifically, the ultrasound video data set is encoded using the afore-described common test conditions. Ultrasound videos are encoded using four rate points for QPs of 22, 27, 32, and 37. Using the bjontegaard-delta peak signal-to-noise ratio (BD-PSNR) algorithm [10], we compute the bitrate gains resulting from the comparison of the pre-processed (despeckled) ultrasound videos for the three investigated filtering approaches, when compared to the original ultrasound video, for all videos parting the data set, for both standards. Moreover, using the procedure depicted in Fig. 1, we compare HEVC random-access and low-delay encodings to analogous H.264/AVC encodings, to showcase the additional bitrate demands reductions which can be achieved using the anticipated HEVC video coding standard.

C. Clinical Video Quality Assessment

Subjective (clinical) video quality assessment was performed by two medical experts. More specifically, a cardiovascular surgeon and a neurovascular specialist were

asked to rate the clinical capacity of the ultrasound video data set prior and after despeckle filtering for the whole data set. The rating scale ranged between one and ten, with one being the lowest rating and ten being the highest. The medical experts provided a clinical rating that reflected the capacity of the ultrasound video under consideration to satisfy the evaluated clinical criteria during standard clinical practice. The ratings were then summed to derive a total clinical score.

V. RESULTS AND DISCUSSION

In this section we discuss the experimental evaluation of the investigated despeckle filtering methods. Firstly, we present the associated bitrate demands reductions resulting from using speckle removal. Next, we showcase the bitrate gains achieved by the emerging HEVC standard compared to H.264/AVC standard, followed by the clinical evaluation.

A. Despeckle Filtering Bitrate Demands in H.264/AVC and HEVC video coding standards

Table I depicts the average bitrate gains of the whole ultrasound video data set resulting from using despeckle filtering as a pre-processing step. The bitrate gains are computed by comparing the despeckled and original videos bitrate requirements. The trend is the same for both H.264/AVC and HEVC standards. At the lowest end, DsFsrad filter achieves bitrate demands reductions ranging from 27% to 30.5%, followed by DsFhmedian filter which extends bitrate savings between 33.2% and 37.6% for the investigated encoding configuration schemes. DsFlsmv filter achieves the best results, with bitrate savings between 42.4% and 47.1%. The additional bitrate can be used to maximize ultrasound video's quality and increase error resiliency.

B. HEVC vs H.264/AVC

Here, we depict the bitrate gains of the emerging HEVC video coding standard. In the leftmost column of Table II we observe that bitrate gains extend up to 32.7% for the low-delay high-efficiency case, when no preprocessing is applied. It is worth noting here that bitrate savings of 35.4% has been recorded in [11] for entertainment applications. The next columns of Table II depict the bitrate gains when despeckle filtering is applied prior to HEVC encoding and the associated bitrate demands are compared to conventional H.264/AVC encoding without any form of pre-processing. DsFsrad which records the lowest savings, requires as much as 43.3% to 52.3% lower bitrate compared to H.264/AVC. DsFhmedian raises bitrate demands savings between 49% and 56.8%, while the best performing despeckle filter DsFlsmv achieves bitrate gains between 57.1% and 63.9%. It is obvious that noise removal is of vital importance for m-health ultrasound video communication systems, not only for improving the visual perception of the ultrasound video but also for lowering the associated bitrate requirements.

C. Clinical Evaluation

The results from the clinical evaluation appear in Table III. As illustrated, DsFlsmv achieves the best mean opinion scores (MOS) followed closely by DsFhmedian. In fact, the 1st expert rated DsFlsmv as the best performing despeckle filtering method while the 2nd expert gave highest scores to DsFhmedian. The latter suggests that more ultrasound cases should be investigated, to limit inter-observer variability.

TABLE I – AVERAGE BITRATE REQUIREMENTS REDUCTIONS (%) AFTER APPLYING DESPECKLE FILTERING COMPARED TO THE ORIGINAL VIDEO FOR H.264/AVC AND HEVC ENCODINGS

		DsFhmedian	DsFlsmv	DsFsrاد
H.264/AVC	<i>JM_LP_HE</i>	33.7%	42.4%	28.0%
	<i>JM_RA_B_HE</i>	34.8%	42.7%	28.7%
	<i>LD_HE10</i>	35.7%	46.7%	29.0%
	<i>LD_main</i>	37.0%	47.1%	30.0%
HEVC	<i>LD_P_HE10</i>	35.8%	46.7%	28.7%
	<i>LD_P_main</i>	37.6%	46.8%	30.5%
	<i>RA_HE10</i>	33.2%	46.1%	26.8%
	<i>RA_main</i>	33.5%	46.0%	27.0%

TABLE II – AVERAGE BITRATE REQUIREMENTS REDUCTIONS (%) AFTER APPLYING DESPECKLE FILTERING AND HEVC ENCODING COMPARED TO THE ORIGINAL H.264/AVC ENCODED ULTRASOUND VIDEO

	Original	DsFhmedian	DsFlsmv	DsFsrاد
LD_HE10 vs JM_LP_HE	32.7%	56.8%	63.9%	52.3%
LD_main vs JM_LP_HE	29.8%	55.5%	62.5%	50.7%
LD_P_HE10 vs JM_LP_HE	30.9%	55.8%	63.0%	51.0%
LD_P_main vs JM_LP_HE	24.3%	52.5%	59.5%	47.3%
RA_HE10 vs JM_RA_B_HE	23.2%	50.5%	58.5%	44.9%
RA_main vs JM_RA_B_HE	21.2%	49.0%	57.1%	43.2%

However, as documented in the objective evaluation, DsFlsmv filter significantly outperforms DsFhmedian in bitrate demands savings. Of paramount importance is the fact that the original ultrasound videos were rated with the lowest scores. This observation strengthens the assumption that noise removal can provide for enhancing the clinical capacity of the examined ultrasound videos. Video image examples of the clinically evaluated videos appear in Fig. 2.

VI. CONCLUDING REMARKS AND FUTURE WORK

This study demonstrates the benefits in terms of bitrate demands reductions resulting from applying despeckle filtering as a pre-processing step, in m-health medical video communication systems. Moreover, this paper highlights the efficiency of the emerging HEVC standard, showing that significant bitrate savings for equivalent perceptual quality can be realized. The best performing despeckle filters are DsFlsmv and DsFhmedian. Both filters attained similar clinical scores. DsFlsmv however achieves higher bitrate reductions, while when used together with HEVC encoding it can provide up to 63.9% bitrate gains compared to H.264/AVC conventional encoding. Future work includes examining additional despeckle filters, also investigating their computational complexity, for a larger ultrasound video data set. In addition, HEVC coded video transmission over 3.5G and 4G wireless networks is currently considered, to showcase the advantages of using HEVC for high-resolution and low-delay communications.

REFERENCES

[1] E. Kyriacou, M.S. Pattichis, C.S. Pattichis, A. Panayides, and A. Pitsillides, "m-Health e-Emergency Systems: Current Status and

TABLE III- CLINICAL EVALUATION OF ORIGINAL AND DESPECKLED ULTRASOUND VIDEOS FOR THE EXAMINED DATA SET

Rating	Original	DsFhmedian	DsFlsmv	DsFsrاد
Expert 1	33	71	75	61
Expert 2	40	75	72	51
MOS	37	73	74	56

1: Lowest Rating, 10: Highest Rating
Results are sums of the ten investigated ultrasound videos

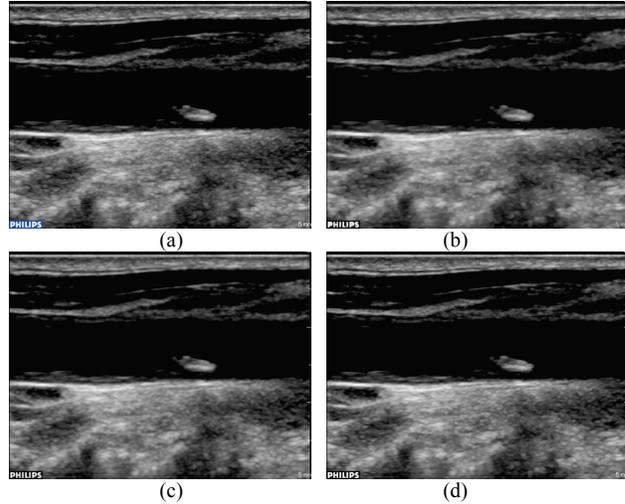


Fig. 2. Ultrasound video image examples of a typical video before and after video despeckle filtering: a) the original ultrasound video, (b) DsFhmedian, (c) DsFlsmv, and (d) DsFsrاد.

Future Directions," *IEEE Antennas and Propagation Magazine*, vol. 49, no.1, pp. 216-231, Feb. 2007.

[2] 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.

[3] 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," *IEEE Transactions on Information Technology in Biomedicine*, vol. 15, no. 3, pp.387-397, May 2011.

[4] 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.*, to be published Dec. 2012.

[5] C.P. Loizou and C.S. Pattichis, "Despeckle filtering algorithms and Software for Ultrasound Imaging", Synthesis Lectures on Algorithms and Software for Engineering, Ed. Morgan & Claypool Publishers, San Rafael, CA, USA, 2008.

[6] C. P. Loizou, T. Kasparis, P. Christodoulides, C. Theofanous, M. Pantziaris, E. Kyriakou, and C. S. Pattichis, "Despeckle Filtering in Ultrasound Video of the Common Carotid Artery," in *Proc of BIBE 2012*, Nov. 11-13, pp. 721-726, Larnaca, Cyprus.

[7] C.P. Loizou, C.S. Pattichis, C.I. Christodoulou, R.S.H. Istepanian, M. Pantziaris, and A. Nicolaidis "Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery," *IEEE Trans. Ultras. Ferroel. Freq. Contr.*, vol. 52, no. 10, pp. 1653-1669, 2005.

[8] A. Nieminen, P. Heinonen, and Y. Neuvo, "A new class of detail-preserving filters for image processing", *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 9, pp. 74-90, 1987.

[9] Y. Yongjian and S.T. Acton, "Speckle reducing anisotropic diffusion," *IEEE Trans. Image Proc.*, vol. 11, no. 11, pp. 1260-1270, Nov. 2002.

[10] G. Bjøntegaard, "Improvements of the BD-PSNR model," ITU-T SG16 Q.6 Document, VCEG-A111, Berlin, Germany, July 2008.

[11] 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.