31st Annual International Conference of the IEEE EMBS
Minneapolis, Minnesota, USA, September 2-6, 2009

Objective and Subjective Evaluations of Quality for Speckle Reduced Echocardiography

Seán Finn, *Student Member, IEEE,* Edward Jones, and Martin Glavin, *Members, IEEE*

Abstract— **This paper presents the results of a study into the subjective effect of speckle reduction filtering in echocardiography, as assessed by clinical experts. Echocardiographic videos were filtered by a number of speckle reduction methods to produce a set of test videos with varying levels of speckle content. Six practicing cardiac technicians were asked to rate each video in three scoring categories, with the aim of quantifying their subjective evaluation of the diagnostic usefulness and the level of speckle in each video. The change in expert scores due to filtering is analyzed.**

Inter-expert difference in the evaluation is investigated, and intra-expert analysis of the association between each score category is also performed. In addition, a number of objective quality metrics are applied to the filtered videos, and the correlation between these metrics and the expert scores is determined. Results indicate that, while there are inter-expert differences, strong intra-expert relationships exist between the score categories. Furthermore, each of the three subjective score categories is strongly associated with one of the objective quality metrics.

I. INTRODUCTION

ULTRASOUND imaging is a widely used medical diagnostic tool, due to its non-invasive and non-ionizing nature, relatively low cost and real-time capabilities. A common problem in the manual and automated analysis of ultrasound images is the presence of speckle noise. Speckle is the interference pattern of sub-resolution sized scatterers, and is observed as a fine granular pattern. Manual analysis can be impaired by speckle by the obscuring of fine structures and boundaries as a result of small grey-level difference masking [1]. Other research found an eight-fold reduction in lesion detectability [2].

A considerable amount of research has resulted in a number of methods to reduce speckle, including [3]-[7]. The motivation for reducing speckle content for purposes of automated analysis usually is the improvement of border detection and segmentation, and it has been widely applied in this field [8]-[11]. However, there has been relatively little work done on the impact on expert clinical opinion of the use of speckle reduction methods. Related work in carotid artery ultrasound speckle reduction [12] uses two experts to rate still ultrasound, filtered by a number of different methods. It is found that only one filter improves subjective quality of the ultrasound images. If the use of automated speckle reduction improves the diagnostic accuracy by experts such as cardiologists and cardiac technicians, speckle reduction methods could be applied in a more widespread fashion in echocardiography.

This paper first examines the effect of speckle reduction methods on the subjective quality of echocardiography, as perceived by experts in the analysis and interpretation of this modality. The objective is to determine if speckle reduction filters are actually perceived to be of benefit in the context of subjective clinical evaluation. Six experts were shown a set of forty-eight echocardiographic videos which had been filtered to reduce their speckle content by four different methods. Videos were evaluated to quantify each expert's opinion on the change in speckle level, change in the clarity of diagnostically important details, and change in overall quality due to the filtering.

Overall trends in these areas are examined, along with an investigation of inter-expert difference in video evaluation. The relationship between each evaluation category is also determined for each expert. In addition, the relationship between expert scores and a number of objective quality metrics is also investigated, in order to determine if these objective metrics are of benefit in assessing the subjective effects of speckle reduction filtering.

The structure of this paper is as follows: Section II outlines the speckle reduction methods used in this study. Section III details the evaluation procedure used to obtain subjective expert opinion, and also the set of objective quality metrics applied to the processed videos. Section IV specifies the results of the evaluation test, including statistical analysis of results. Finally, section V concludes the paper.

II. SPECKLE REDUCTION FILTERS

Four speckle reduction filters were chosen for their suitability for medical ultrasound. They are based on multiscale and diffusion approaches, two of the most widely used methodologies in this area [3]-[7].

A. Speckle Reducing Anisotropic Diffusion (SRAD)

Anisotropic diffusion, introduced by Perona and Malik [13] aims to smooth the noise in an image while preserving edges by evolving a partial differential equation conceptually similar to heat diffusion. Yu and Acton's SRAD method [3] extends this filtering approach to the

Manuscript received April 7, 2009. This work was supported by IRCSET, the Irish Research Council for Science, Engineering and Technology, under the National Development Plan.

The authors are with the Bioelectronics Research Cluster, National Centre for Biomedical Engineering Science, and the Department of Electrical & Electronic Engineering, National University of Ireland, Galway, Republic of Ireland.

Seán Finn can be contacted at: phone: +353 91 49 2728; fax: +353 91 49 4511; e-mail: sean.finn@nuigalway.ie. E. Jones and M. Glavin can be contacted by e-mail at: edward.jones@nuigalway.ie; martin.glavin@nuigalway.ie.

requirements of speckle corrupted images. The diffusion process is controlled by the instantaneous coefficient of variation, which is based on a discriminator used in radar imagery. Further details can be found in [3].

B. Method of Abd-Elmoniem et al.

Another diffusion based method is that of Abd-Elmoniem *et al.* [4]. This adapts to the local speckle content based on speckle categorization. Similar to the coherence enhancing method of Weickert [14], the diffusion tensor is used to describe image structure. The function used to control the diffusion process is also tensor valued, and is calculated by modifying the values of the structure tensor eigenvalues. This allows control of diffusion strength in both the normal and tangential directions relative to the local gradient direction.

C. Generalized Likelihood Method

The generalized likelihood method of Pižurica *et al.* [5] removes noise in the wavelet domain. A general model of the noise is used, and denoising is based on estimating the probability distribution functions of signal and noise pixels using the generalized likelihood method.

D. Nonlinear Multiscale Wavelet Diffusion (NMWD)

This method combines the signal/noise separation capabilities of wavelet analysis with the iterative noise smoothing of anisotropic diffusion. Wavelet decomposition is performed using the discrete wavelet transform of Mallat and Zhong [15]. The wavelet coefficients are modified using anisotropic diffusion to remove noise, before reconstruction of the denoised image.

III. Methods

A. Expert Evaluation

Two transthoracic echocardiographic videos, a short-axis view of the mitral valve and a two-chamber long-axis view, were used as input to the speckle reduction filters. For each input video a number of filtered videos with varying levels of speckle reduction are produced by tuning filter parameters. A total of forty-eight output videos are produced.

Six cardiac technicians, skilled in the analysis and interpretation of echocardiographic videos, assessed the set of videos. The three subjective criteria for evaluation were based on the opinion of one clinical expert as to what constitute important clinical factors:

1. *Speckle Level.* This quantifies the expert's assessment of the speckle level in each video.
2. *Detail Clarity.* Quantifies the subjective resolvability of diagnostically important details.
3. *Overall Quality.* This quantifies the Overall Quality of the video, and includes any other clinical considerations.

Each expert was shown the filtered and corresponding unfiltered videos in random order, and asked to assign a score in each category on a scale of one to ten.

For each of the three expert scores, the non-parametric Kruskal-Wallis test is applied to determine if the differences between experts are significant. Any possible intra-expert association between the three expert scores is examined by using Spearman's method. The level of significance in all cases is chosen as 1% (p=0.01).

B. Objective Quality Metrics

A set of commonly-used quality metrics are calculated to determine if there is any relationship between objective measures and the subjective opinions of the participants. Metrics are calculated on a frame-by-frame basis, and averaged over all the frames in each video.

1) Mean Squared Error (MSE)

The MSE is a commonly used metric signal processing. Here it quantifies the modification of a frame due to filtering, averaged over all pixels:

$$MSE(I_1, I_2) = \sum_{i=1}^{X} \sum_{j=1}^{Y} \{I_1(i,j) - I_2(i,j)\}^2 \qquad (1)$$

I_1 and I_2 are the filtered and unfiltered frames (both of size $X \times Y$ pixels), and (i,j) are the spatial co-ordinates.

2) Edge Region MSE

This is the MSE as above, but it is calculated only over the edge region pixels in a frame. Each frame of unfiltered video is split into edge and non-edge regions by thresholding a homogeneity measure, defined as $h = \sigma^2/\mu$. Here σ^2 and μ are the local variance and mean of the unfiltered image, calculated using an 11×11 pixel window. A binary edge map I_{edge} is determined using an experimentally chosen threshold τ, as:

$$I_{edge}(i,j) = \begin{cases} 1, & \text{if } h(i,j) > \tau \\ 0, & \text{otherwise} \end{cases} \qquad (2)$$

3) Pratt's Figure of Merit (FOM)

A measure of edge preservation between two images, the FOM is calculated by determining an edge map for both images [16]. This metric is defined as:

$$FOM(I_1, I_2) = \frac{1}{\max(N_1, N_2)} \sum_{i=1}^{N_1} \frac{1}{1 + d_i^2 \alpha} \qquad (3)$$

where N_1 and N_2 are the number of edge pixels in the edge maps of the filtered and unfiltered frames. α is a constant (usually set to 1/9) and d_i is the Euclidean distance between the i^{th} edge pixel in the filtered and unfiltered edge maps. The FOM ranges from 0 to 1, with higher values better. Edge maps were found using the Canny edge detector [17].

IV. RESULTS

A. General Trends

Table I shows the breakdown of each expert score into three categories, corresponding to an increase, decrease, or no change relative to the same expert's score for the unfiltered version of that video. Here scores are aggregated over all experts to determine general trends.

Almost two thirds of Speckle Level scores indicate a reduction in perceived Speckle Level. In 36% of cases, the Overall Quality before and after filtering was deemed unchanged, and in a very small number of cases Overall Quality was judged to be improved by speckle reduction. A similar distribution is observed for the perceived Detail Clarity, with less than 10% exhibiting an increase.

TABLE I
AGGREGATED SCORE DISTRIBUTION, N = 288

	Increase	No Change	Decrease
Overall Quality	16 (5.6%)	104 (36.1%)	168 (58.3%)
Detail Clarity	22 (7.6%)	113 (39.3%)	153 (53.1%)
Speckle Level	15 (5.2%)	86 (29.9%)	187 (64.9%)

B. Inter-Expert Differences

Fig.1 shows the summary of each expert's scores, for the three scoring categories. A degree of inter-expert variability in scores is observed. Application of the Kruskal-Wallis H-test yields the values shown in Table II. The inter-expert differences in the Detail Clarity and Speckle Level scores are significant. Some difference is seen in expert scores on Overall Quality, but these are not significant at the 1% level.

C. Intra-Expert Subjective Score Relationships

Investigation of intra-expert association between the three scores using Spearman's method results in the ρ and significance values of Table III. For all of the experts, a statistically significant and strong positive correspondence between Overall Quality and Detail Clarity is observed. A statistically significant positive relationship is observed for four of the six experts between Overall Quality and Speckle Level. Five expert's scores had a positive association between Detail Clarity and Overall Quality.

TABLE II
INTER-EXPERT KRUSKAL-WALLIS VALUES WITH SIGNIFICANCE

	Overall Quality	Detail Clarity	Speckle Level
Kruskal-Wallis H values	H = 17.76, NS	H = 37.19, p = 547.8E-9	H = 51.53, p = 673.9E-12

NS = Not Significant at 1%

D. Relationship Between Objective and Subjective Scores

Results of analysis of the correlation between expert scores and the objective metrics using Spearman's method are shown in Table IV. For the scores of all experts, a strong positive relation is observed between the FOM metric and Overall Quality. A strong negative association is observed between MSE and subjective Speckle Level, and also between the edge region MSE and perceived Detail Clarity.

V. CONCLUSIONS

The effect of speckle reduction filtering on both the Overall Quality and Detail Clarity scores is negative in over half of cases. Based on these aggregate scores, the expert participants do not judge a reduction in speckle to have a positive effect on the Overall Quality of echocardiographic video.

The lack of a significant inter-expert difference on Overall Quality indicates a similarity in assessment of general clinical quality. There are however significant inter-expert differences in Detail Clarity and Speckle Level. This suggests a variability in how experts subjectively asses the visibility of diagnostically relevant details, and also the strength of the speckle component in echocardiographic video.

While expert opinion on Overall Quality is not associated with a decrease in perceived Speckle Level in this study, there is a strong association between perception of Overall Quality and Detail Clarity in all expert participants. This indicates that subjective Detail Clarity is a strong indicator of diagnostic quality in echocardiography, and perhaps the primary consideration in expert evaluation of echocardiographic usefulness.

For most of the experts there is a statistically-significant positive association between Speckle Level and both Overall

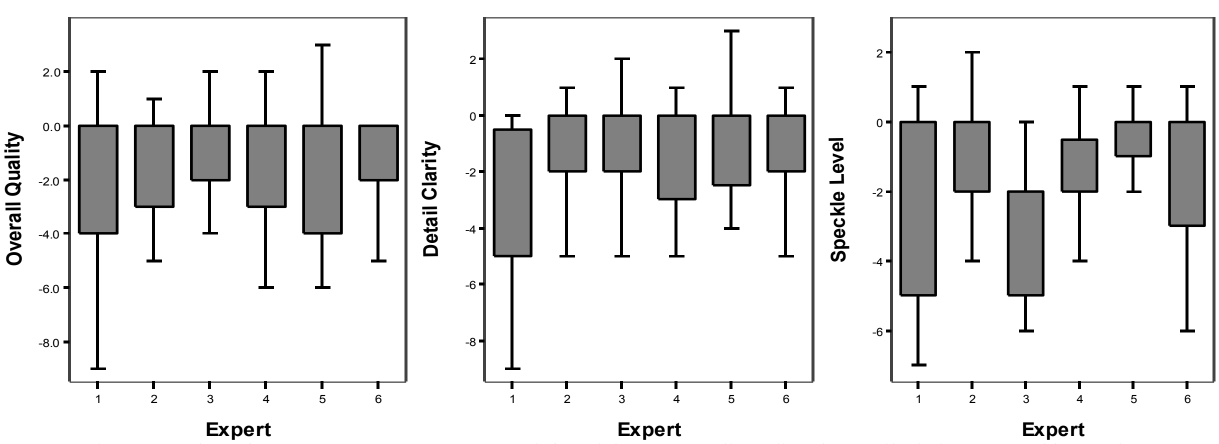

Fig. 1. Boxplots of Expert Scores, per expert. From left to right: (a) Overall Quality, (b) Detail Clarity, (c) Speckle Level.

TABLE III
INTRA-EXPERT ASSOCIATIONS BETWEEN SCORES USING SPEARMAN'S METHOD

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6
Overall Quality and Detail Clarity	ρ = 0.947, *p* = 3.97E-15	ρ = 0.888, *p* = 365.2E-19	ρ = 0.715, *p* = 113.2E-10	ρ = 0.936, *p* = 195.9E-24	ρ = 0.783, *p* = 474.0E-13	ρ = 0.825, *p* = 487.6E-20
Overall Quality and Speckle Level	ρ = 0.625, *p* = 156.1E-8	ρ = 0.355, NS	ρ = 0.609, *p* = 411.8E-8	ρ = 0.217, NS	ρ = 0.571, *p* = 225.5E-7	ρ = 0.759, *p* = 627.4E-16
Detail Clarity and Speckle Level	ρ = 0.644, *p* = 389.6E-9	ρ = 0.454, *p* = 117.7E-5	ρ = 0.534, *p* = 935.7E-7	ρ = 0.233, NS	ρ = 0.493, *p* = 374.9E-6	ρ = 0.777, *p* = 514.8E-16

NS = Not Significant at 1%

TABLE IV
INTRA-EXPERT ASSOCIATIONS BETWEEN SCORES AND METRICS USING SPEARMAN'S METHOD

	Expert 1	Expert 2	Expert 3	Expert 4	Expert 5	Expert 6
Overall Quality and FOM	ρ = 0.74, *p* = 176.5E-11	ρ = 0.59, *p* = 123.0E-7	ρ = 0.53, *p* = 124.1E-6	ρ = 0.72, *p* = 607.0E-11	ρ = 0.55, *p* = 452.7E-7	ρ = 0.82, *p* = 992.7E-15
Detail Clarity and Edge Region MSE	ρ = -0.67, *p* = 164.0E-9	ρ = -0.62, *p* = 240.0E-8	ρ = -0.49, *p* = 402.0E-6	ρ = -0.76, *p* = 346.8E-12	ρ = -0.49, *p* = 332.9E-6	ρ = -0.83, *p* = 205.6E-15
Speckle Level and MSE	ρ = -0.69, *p* = 531.1E-10	ρ = -0.47, *p* = 726.0E-6	ρ = -0.85, *p* = 306.8E-16	ρ = -0.47, *p* = 710.1E-6	ρ = -0.64, *p* = 992.1E-9	ρ = -0.87, *p* = 214.3E-17

Quality and Detail Clarity. This suggests perceptions of Overall Quality and Detail Clarity are higher for videos with less speckle reduction. These associations are not significant for all experts however, and are weaker than the association between Overall Quality and Detail Clarity.

The correlation between Pratt's FOM metric and the Overall Quality scores of all experts demonstrates this metrics applicability to echocardiography. The negative correlation between edge region MSE and Detail Clarity (i.e. higher Edge MSE implies lower Detail Clarity) for all experts suggests that this metric is a good indicator of expert opinion on Detail Clarity. The association between the overall MSE and perceived Speckle Level indicates that perception of Speckle Level decreases with increasing amounts of filtering as measured by this metric.

This study demonstrates some of the factors considered in expert evaluation of speckle reduction in echocardiographic filtering. Speckle filtering is shown to reduce perceived Speckle Level; however this can lead to a reduction in overall subjective quality. Furthermore, the results also indicate a reasonably strong correlation between a number of objective metrics and subjective clinical opinion, suggesting that these objective metrics may have clinical utility in echocardiography.

REFERENCES

[1] C. B. Burkhardt, "Speckle in ultrasound B-mode scans," *IEEE Trans. Sonics Ultrason.,* vol. 25, no. 1, pp. 1–6, Jan. 1978.

[2] J. C. Bamber and C. Daft, "Adaptive filtering for reduction of speckle in ultrasonic pulse-echo images." *Ultrasonics*, pp. 41–44, Jan 1986.

[3] Y. Yu and S. Acton, "Speckle reducing anisotropic diffusion," *IEEE Trans. Image Process.,* vol. 11, no. 11, pp. 1260–1270, Nov. 2002.

[4] K. Z. Abd-Elmoniem, A.-B. M. Youssef, and Y. M. Kadah, "Real-time speckle reduction and coherence enhancement in ultrasound imaging via nonlinear anisotropic diffusion," *IEEE Trans. Biomed. Eng.,* Sep 2002.

[5] A. Piẑ urica, W. Philips, I. Lemahieu, and M. Acheroy, "A versatile wavelet domain noise filtration technique for medical imaging," *IEEE Trans. Med. Imag.,* vol. 22, no. 3, pp. 323-331, March 2003.

[6] X. Zong, A. F. Laine, and E. A. Geiser, "Speckle reduction and contrast enhancement of echocardiograms via multiscale nonlinear processing," *IEEE Trans. Med. Imag.,* vol. 17, no. 4, pp. 532–540, Aug 1998.

[7] Y. Yue, M.M. Croitoru, A. Bidani, J.B. Zwischenberger, and J.W. Clark, Jr., "Nonlinear multiscale wavelet diffusion for speckle suppression and edge enhancement in ultrasound images," *IEEE Trans. Med. Imag.,* vol. 25, no. 3, pp. 297–311, March 2006.

[8] N. Archip, R. Rohling, P. Cooperberg, and H. Tahmasebpour, "Ultrasound image segmentation using spectral clustering," *Ultrasound in Medicine & Biology*, Vol. 31, Iss. 11, pp. 1485-1497, Nov. 2005.

[9] M. Alemán-Flores, L. Álvarez, and V. Caselles, "Texture-Oriented Anisotropic Filtering and Geodesic Active Contours in Breast Tumor Ultrasound Segmentation", *Jour. Math. Imaging Vis.,* vol. 28, Issue 1, pp. 81-97, May 2007.

[10] S. Balocco, O. Basset, C. Cachard, and P. Delachartre, "Spatial anisotropic diffusion and local time correlation applied to segmentation of vessels in ultrasound image sequences," *Ultrasonics, 2003 IEEE Symposium on,* vol.2, no. 5-8, pp. 1549-1552 Oct. 2003.

[11] A. Kissi, S. Cormier, L. Pourcelot, A. Bleuzen, and E. Tranquart, "Contrast enhanced ultrasound image segmentation based on fuzzy competitive clustering and anisotropic diffusion," *26th Ann. Conf. IEEE EMBS,* vol.1, no. 1-5, pp.1613-1615, Sept. 2004.

[12] C.P. Loizou *et al.*, "Comparative evaluation of despeckle filtering in ultrasound imaging of the carotid artery," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control,* vol. 52 no. 10 pp.1653–1669, Oct. 2005.

[13] P. Perona and J. Malik, "Scale-space and edge detection using anisotropic diffusion," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 12, no. 7, pp. 629–639, Jul 1990.

[14] J. Weickert. *Anisotropic Diffusion in Image Processing.* Teubner-Verlag, Stuttgart, Germany, 1998.

[15] S.Mallat and S. Zhong, "Characterization of signals from multiscale edges," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol.14, no.7, pp.710-732, Jul 1992.

[16] I.E. Abdou and W.K. Pratt. "Quantitative design and evaluation of enhancement/thresholding edge detectors, " *Proceedings of the IEEE*, vol. 67 no. 5 pp. 753–763, May 1979.

[17] F. J.Canny. "A Computational Approach to Edge Detection," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 8, no. 6, pp. 679–698, Nov. 1986.