Abstract—Limited contrast, along with speckle and acoustic noise, can reduce the diagnostic value of echocardiographic images. This study introduces dynamic histogram-based intensity mapping (DHBIM), a novel approach employing temporal variations in the cumulative histograms of cardiac ultrasound images to contrast enhance the imaged structures. DHBIM is then combined with spatial compounding to compensate for noise and speckle. The proposed techniques are quantitatively assessed (32 clinical data sets) employing (i) standard image quality measures and (ii) the repeatability of routine clinical measurements, such as chamber diameter and wall thickness. DHBIM introduces a mean increase of 120.9% in tissue/chamber detectability, improving the overall repeatability of clinical measurements by 17%. The integrated approach of DHBIM followed by spatial compounding provides the best overall enhancement of image quality and diagnostic value, consistently outperforming the individual approaches and achieving a 401.4% average increase in tissue/chamber detectability with an associated 24.3% improvement in the overall repeatability of clinical measurements. (E-mail: A.Perperidis@hw.ac.uk) © 2017 The Authors. Published by Elsevier Inc. on behalf of World Federation for Ultrasound in Medicine & Biology. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Key Words: Ultrasound, Echocardiography, Histogram processing, Image enhancement, Contrast enhancement, Noise suppression.

INTRODUCTION

Transthoracic echocardiography, although a valuable tool for the assessment of cardiac morphology and function, suffers from a range of artifacts because of the interaction of the transmitted ultrasound with structures such as bone, lung and fat. These interactions may cause acoustical noise and speckle, reduce contrast and limit the delineation of fine anatomic detail. Such artifacts may therefore limit the (i) diagnostic value and (ii) effectiveness of post-processing tools on cardiac ultrasound images. Although advances in data acquisition technologies have substantially improved cardiac ultrasound data, a systematic study performed at the Echocardiography Department of the Western General Hospital (Edinburgh, Scotland) indicated that a considerable (<30%) portion of routine cardiac scans generate low-quality images of limited diagnostic value (corrupted structures, limiting or even prohibiting clinical measurements). The results have been used for educational purposes within the department and have not been published yet. Consequently, there is research interest in the development of post-processing methods that address these limitations, enhancing the image quality and diagnostic value of cardiac ultrasound.

Numerous approaches to enhancement of cardiac ultrasound images have been suggested (Perperidis 2016). Spatial compounding is a popular method that suppresses noise by combining partially decorrelated images produced by imaging the target region of interest from different viewing angles. Tissue structures present in all the partially decorrelated views are enhanced, whereas artifacts not present in all views are suppressed. There are studies that have employed spatial compounding through transducer repositioning for the enhancement of 3-D cardiac ultrasound data (Gooding et al. 2010;...
Mulder et al. 2014; Rajpoot et al. 2011; Szmigielski et al. 2010; Yao et al. 2010). Some studies have attempted to enhance 2-D cardiac ultrasound images by averaging temporally consecutive frames (Achmad et al. 2009; dos Reis et al. 2009; Li et al. 1994; Petrovic et al. 1986). Other studies have utilised the repeated rhythmic contractions of the heart to acquire and compound multiple partially decorrelated 2-D images of the same cardiac phase over consecutive cardiac cycles through a single acoustic window. The process has been referred to as temporal compounding (Abiko et al. 1997; Amorim et al. 2009; Klingler et al. 1989; Olstad 2002; Perperidis et al. 2009; Rigney and Wei 1988; Unser et al. 1989; van Ocken et al. 1981; Vitale et al. 1993).

Although spatial compounding is primarily a tool for noise and speckle suppression, it has also been found to compensate for other artefacts in cardiac ultrasound, including partial shadowing and reverberations. However, spatial compounding has no noteworthy effect in enhancing the cardiac tissue/chamber contrast. Limited contrast between cardiac tissue and chambers constitutes a major limitation in cardiac ultrasound images, potentially making (when combined with noise) the delineation of cardiac structures very challenging. Non-contrast tissue harmonic imaging (THI) is a well-established approach (sometimes employed as the standard acquisition mode) for enhancing the contrast and delineation between cardiac tissue and chambers (Averkiou et al. 1997; Becher et al. 1998; Caidahl et al. 1998; Franke et al. 2000; Ward et al. 1997). However, challenging cases still arise where effective post-processing techniques for the contrast enhancement between cardiac tissue and chambers could be beneficial. A number of filters (Finn et al. 2011; Tay et al. 2010; Yue et al. 2005; Zong et al. 1998) and spatial compounding approaches (Rajpoot et al. 2009; Szmigielski et al. 2010; Yao and Penney 2008) have been claimed to enhance the structure boundaries along with the contrast between cardiac tissue and chambers. However, many of these noise suppression methods actually perform contrast and boundary preservation rather than enhancement. Consequently, dedicated methods that suppress noise in cardiac chambers while enhancing the contrast and detectability between cardiac tissue and chamber structures are desirable.

To distinguish and enhance the contrast between tissue and chamber pixels, a method to derive an effective tissue/chamber threshold is required. Sezgin and Sankur (2004) categorised thresholding methods into six groups: (i) histogram shape, (ii) clustering, (iii) entropy, (iv) image attributes, (v) spatial information and (vi) local characteristics. Histogram-based methods are very popular because of their simplicity and efficiency. Therefore, the multiframe nature of cardiac ultrasound data (exceeding 100 frames/s on state-of-the-art scanners) makes histogram based thresholding very appealing.

Abdullah-Al-Wadud et al. (2007) and Ibrahim and Kong (2007) introduced contrast enhancement techniques that analyse the shape of the histogram and partition it based on local minima and maxima, respectively. They both then assigned specific gray-level ranges to each partition before performing histogram equalisation on them individually. Both techniques performed well in low-contrast images. However, the high levels of noise along with speckle and the subsequently more complex histograms in cardiac ultrasound data impose a severe limitation on the potential of similar histogram shape-based techniques (Hammoudeh 1998). Zwirn and Akselrod (2004, 2005) introduced the adaptive brightness transfer function (ABTF) based on the assumption that the histogram of cardiac ultrasound images can be approximated by the sum of three overlapping Gaussian distributions. The intersections between these three Gaussians would act as thresholds, with each section of the histogram being individually processed using a number of techniques, including histogram equalisation, specification and scaling. The results presented on suppressing chamber noise and enhancing contrast between cardiac tissue and chambers were promising. However, like most histogram shape-based approaches, ABTF makes a strong assumption about the shape of the histograms of the processed images. Although some images may conform to this paradigm, such an assumption may result in considerable pixel misclassification, especially in cases of large overlap between the fitted Gaussians. Moreover, ABTF assumed constant illumination throughout the field of view, which in many scans is not the case because of suboptimal manual setting of the time gain compensation (TGC).

The aim of this study was to introduce a cardiac ultrasound image enhancement approach that suppresses noise and speckle while increasing contrast and detectability between cardiac tissue and chambers. Initially, dynamic histogram based intensity mapping (DHBIM) is introduced as a novel, simple and efficient method for suppressing cardiac chamber noise, enhancing tissue speckle and increasing tissue/chamber detectability. Unlike previous approaches, DHBIM employs variations over time in the cumulative histograms of cardiac ultrasound image sequences to derive a tissue/chamber threshold, avoiding strong assumptions on histogram shape. DHBIM is then combined with a temporal compounding (Perperidis et al. 2015) approach that also makes use of image variations over time to compensate for noise and speckle, as well as other artefacts that may momentarily appear, such as shadowing and reverberations. The results of the individual (DHBIM and temporal compounding) approaches as well as the integrated approach are finally compared.
METHODS

Data acquisition and classification

Cardiac data from 32 patients were acquired according to the standards adopted by the British Society of Echocardiography (Feigenbaum et al. 2005; Henry et al. 1980; Wharton et al. 2012) at the Western General Hospital, Edinburgh. All data sets used in this study were selected from fully anonymised cine loops recorded during routine clinical examinations in the course of normal care with no intention to use them in research at the time of collection. Consequently, no National Health Service (NHS) ethics approval was required under the terms of the Governance Arrangements for Research Ethics Committees: A Harmonised Edition (Department of Health Research and Development Directorate of England et al. 2011). The data sets were representative of patients examined in the department and so included data over the full diagnostic quality range of high (12), average (12) or low (8) (as specified by an experienced echocardiographer). B-Mode data of at least 25 cardiac cycles of the parasternal long-axis (PLAX) view were selected with no other quality-related selection criteria being used. The PLAX view was utilised because it is common in clinical practice and enables (i) the visualisation of multiple cardiac structures and (ii) the acquisition of a range of clinical measurements that are essential during the diagnostic process. All data were acquired using a GE Vivid 7 Dimension ultrasound scanner (GE Healthcare, Little Chalfont, UK) along with a 3-MHz phased array probe employing tissue harmonic imaging as standard acquisition mode. Images were captured at 25 frames/s (fps). Other acquisition parameters such as acquisition depth, focus depth, sector width, gain and TGC were manually optimised for each subject. Finally, B-mode image sequences of $434 \times 636$ pixels were exported in DICOM format with no compression applied.

Data analysis: Contrast enhancement

Dynamic histogram-based intensity mapping was implemented using a three-step process (Fig. 1a): (i) identification and compression of any underutilised intensity levels within the frame sequence, (ii) derivation of a cardiac tissue/chamber intensity threshold and (iii) enhancement of the contrast between cardiac tissue and chamber structures.

A B-mode cardiac cycle can be described by an ordered sequence of $n$ 2-D images $S(x,y)$ with a fixed field of view $Q_S$ and an acquisition time $t$ in the temporal direction. The resulting image sequence can be viewed as 2-D + time structure $S(x,y,t)$ defined on the spatiotemporal domain $Q_S \times [t_1,t_n]$. The goal was therefore to derive an intensity transformation function $T$ of the form:

$$T : S(x,y,t) \rightarrow S'(x,y,t)$$

where $S(x,y,t)$ represented the gray level of the original image at point $(x,y)$ and time $t \in [1,M]$, and $S'(x,y,t)$ represented the corresponding gray level on the processed data of image. Transformation $T$ mapped the intensity level $S(x,y,t)$ at any point in an image sequence into a corresponding intensity level $S'(x,y,t)$. Any parameters required to define the intensity transfer function $T$ (intensity mapping) were automatically derived by analysing the dynamic variations in the image histograms through a sequence of consecutive cardiac ultrasound frames. The normalized histogram of a digital image with gray levels in the range $[0,N]$ ($N = 255$) was a discrete function (Gonzalez and Woods 2001)

$$p(r_k) = \frac{n_k}{n}$$

where $r_k$ is the $k$th gray level, $n_k$ is the number of pixels in the image having gray level $r_k$ for $k \in [0,N]$ and $n$ denotes the total number of pixels in the image (i.e., $p(r_k)$ denotes the probability of occurrence of gray level $r_k$).

![Fig. 1. Block diagram of (a) the proposed dynamic, contrast enhancement approach followed by (b) temporal compounding to a holistic image enhancement methodology.](image-url)
Identification and compression of underutilised intensities within the frame sequence

Underutilised intensity levels can be defined as intensity levels that either are not utilised because of limited imaging dynamic range or correspond to regions that are not affected by the constant motion of the cardiac walls (such as central regions of chambers). Underutilised intensity levels within the B-mode frame sequences were identified by analysing the dynamic temporal variations in the image histograms corresponding to a sequence of consecutive cardiac ultrasound frames. A normalised histogram was generated for each frame in the B-mode sequence:

\[ p(r_k,t) = n_k(t)/n \]  

The standard deviation (SD) of the variations in each intensity level over consecutive frames was then derived:

\[ \sigma_p(r_k) = \frac{1}{M} \sum_{i=1}^{M} (p(r_k,t) - \mu)^2 \]  

where \( M \) is the total number of frames, and \( \mu \) is the mean \( p(r_k) \) value across the B-mode frame sequence. A heuristic threshold set as <1% of the maximum value observed over the derived \( \sigma_p(r_k) \) profile was utilised to identify such underutilised intensity levels:

\[
\begin{align*}
  th_l &= \left\{ \max(i), i \in [1, N] : [\sigma_p(r_k), \sigma_p(r_j)] < 0.01 \times \max_k \sigma_p(r_k) \right\} \\
  th_u &= \left\{ \min(j), j \in [1, N] : [\sigma_p(r_i), \sigma_p(r_j)] < 0.01 \times \max_k \sigma_p(r_k) \right\}
\end{align*}
\]

where \( th_l \) and \( th_u \) are the estimated lower and upper thresholds, respectively. Subsequently, a simple piecewise-linear transformation \( T_1 \) function was used as an intensity mapping between the original and the processed frame sequences

\[ T_1(r_k) = \begin{cases} 0, & \text{if } k \leq th_l \\ N, & \text{if } k \geq th_u \\ (a \times r_k) + b, & \text{if } th_l < k < th_u \end{cases} \]

where \( r_k \) is the \( k \)th intensity level, \( a=N(th_u-th_l) \) and \( b=(N \times th_l)/(th_u-th_l) \) represents the slope and y-intercept of the linear contrast stretching, respectively. \( T_1 \) compressed all intensity levels below or equal to \( th_l \) to 0 (black-level expansion) and all intensity levels above or equal to \( th_u \) to \( N=255 \) (white-level expansion), increasing the dynamic range of the imaged cardiac structure.

Identification of tissue/chamber intensity threshold

The novel approach proposed in this study was based on some basic observations in cardiovascular ultrasound images. More precisely, the cardiac tissue/chamber intensity threshold \( th_{c,c} \) was identified by analysing the dynamic variations in the image histograms over a sequence of consecutive cardiac ultrasound frames. Let \( n_c(t_d) \) and \( n_c(t_s) \) represent the true number of cardiac chamber pixels within an end-diastole (ED) \( (t_d) \) and an end-systole (ES) \( (t_s) \) frame, respectively (corresponding to a correct tissue/chamber gray-level threshold, there are three possible cases.

In the first case, \( th_c < th_{c,c} \), therefore,

\[
DS_c = n_c(t_d) - n_c(t_s) = (n_c(t_d) - n_p(t_d)) - (n_c(t_s) - n_p(t_s))
\]

where \( n_c(t_d) \) and \( n_c(t_s) \) are the total number of pixels identified as chamber pixels \( (t_e,c) \), and \( n_p(t_d) \) and \( n_p(t_s) \) are the number of chamber pixels falsely identified as tissue pixels in the ED and ES frames, respectively. Because the number of chamber pixels reaches its maximum in ED and its maximum in ES, it is expected that \( (n_p(t_d) - n_p(t_s)) > 0 \) and, therefore, \( DS_c < DS_c \).

In the second case, \( th_c > th_{c,c} \) therefore,

\[
DS_c = n_c(t_d) - n_c(t_s) = (n_c(t_d) + n_p(t_d)) - (n_c(t_s) + n_p(t_s))
\]

where \( n_c(t_d) \) and \( n_c(t_s) \) are the number of tissue pixels falsely identified as chamber pixels in the ED and ES frames, respectively. Given that the number of tissue pixels reaches its minimum in ED and its maximum in ES, it is expected that \( (n_p(t_d) - n_p(t_s)) < 0 \), and therefore,
$D_S < D_{Sr}$. Consequently, in the third and last case, where $th_c = th_c$, $D_S$ is expected to attain its maximum value, which is equal to $D_{Sr}$.

In this study, the number of pixels identified as “chamber” $n_i$ for a given threshold $th_c$ (pixel intensities less than threshold $th_c$) was derived over the full range of possible threshold values ($i.e., th_c \epsilon [0, N]$). $N = 255$ utilising the cumulative image histogram

$$n_k = cH(r_k) = \sum_{i=0}^{k} n_i \quad [9]$$

where $r_k$ is the $k$th gray level, and $n_i$ is the number of pixels in the image having gray level $i$. Considering the high noise level along with other artifacts (such as shadowing) present in cardiac ultrasound data, a more robust measure of determining the variation in the number of chamber pixels between ED and ES (compared with the measure of determining the variation in the number of high noise level along with other artifacts (such as shading) pixels in the image having gray level $r_k$) was utilised: for each intensity level within the cumulative histogram (SD) of the population variations over consecutive frames $D_{Sr}$, the SD was expected to have a maximum value for $th_c$ which is equal to $D_{Sr}$.

Consequently, in the third and last case, where $th_c = th_c$, $D_S$ is expected to attain its maximum value, which is equal to $D_{Sr}$.

In this study, the number of pixels identified as “chamber” $n_i$ for a given threshold $th_c$ (pixel intensities less than threshold $th_c$) was derived over the full range of possible threshold values ($i.e., th_c \epsilon [0, N]$), $N = 255$ utilising the cumulative image histogram

$$n_k = cH(r_k) = \sum_{i=0}^{k} n_i \quad [9]$$

where $r_k$ is the $k$th gray level, and $n_i$ is the number of pixels in the image having gray level $i$. Considering the high noise level along with other artifacts (such as shadowing) present in cardiac ultrasound data, a more robust measure of determining the variation in the number of chamber pixels between ED and ES (compared with $D_S$) was required. As a result, the standard deviation (SD) of the population variations over consecutive frames for each intensity level within the cumulative histogram was utilised:

$$\sigma_n(r_k) = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (n_{ni}(r) - \mu)^2}. \quad [10]$$

Here $M$ is the total number of frames, and $\mu$ is the mean $n_{ni}$ value over the B-mode frame sequence. Similar to $DS_r$, the SD was expected to have a maximum value for $th_c$ equal to the actual tissue/chamber threshold ($th_c = th_c$):

$$th_c = k : \max_{k \in [0, N]} (\sigma_n(r_k)). \quad [11]$$

To further enhance the robustness of the approach, making it less susceptible to outliers, data from multiple ED–ES sequences were utilised during the derivation and allocation of the maximum SD value.

**Enhance contrast between cardiac tissue and chamber structures**

Having identified threshold $th_c$, defining the intensity levels corresponding to cardiac chamber and tissue, a piecewise-linear transformation function (intensity mapping) was employed to (i) suppress noise in cardiac chambers; (ii) increase the displayed dynamic range in cardiac tissue, enhancing the available tissue speckle information; and (iii) enhance the contrast between cardiac tissue and chamber. The shape of the transformation $T_2$ was controlled by threshold $th_c$ and was defined as

$$T_2(r_k) = \begin{cases} r_k/2, & \text{if } k \leq th_c, \\ \alpha \times r_k + \beta, & \text{if } k > th_c \end{cases} \quad [12]$$

where $r_k$ is the $k$th intensity level, $\alpha$ represents the slope of the linear contrast stretching $\left( \alpha = 1 + \frac{th_c}{2(N-th_c)} \right)$ and $\beta$ represents its $y$-intercept $\left( \beta = -N \times \frac{N \times th_c}{2(N-th_c)} \right)$.

**Subsectioning**

From experience, the level of chamber noise as well as the tissue/chamber contrast can vary considerably across the imaged field of view. Such variations can be attributed to the complex interaction of each structure with the transmitted ultrasound signals, as well as non-optimal TGC settings during data acquisition. To compensate for such potential variations, each step was performed on a number of non-overlapping subsectors along the $y$-axis of the cardiac scan (Fig. 2). Each non-overlapping sector would act as a binary mask, with no information outside the sector being used throughout the tissue/chamber threshold estimation process. A prerequisite for each subsector was that it contain both cardiac tissue and chamber structures so that temporal histogram variations could generate a representative tissue/chamber threshold. Identifying a suitable number of subsectors was essential for the effectiveness and efficiency of the DHBIM process. The sectors were therefore empirically selected to include different cardiac sections imaged in a parasternal long-axis cardiac ultrasound scan. A linear interpolation between the individual extracted thresholds was employed across the boundaries of each sector to avoid any intensity level discontinuities.

Fig. 2. Partition of scan in four non-overlapping subsectors along the $y$-axis.
Data analysis: Noise suppression

To further enhance the image quality and diagnostic value of the imaged cardiac structures, temporal compounding was employed as a noise/speckle suppression technique. Temporal compounding was implemented using a three-step process (Fig. 1b), including (i) temporal alignment of the multicycle data to a reference cardiac cycle, (ii) spatial alignment of the temporally aligned frames and (iii) spatial compounding of the spatiotemporally aligned data. Detailed information can be found in Perperidis et al. (2015).

Step 1: Temporal alignment. The temporal alignment process was further divided into four steps.

Identification of ED and ES frames. All ED and ES frames within a multicycle data set were identified utilising intensity information (a robust adaptation of normalised cross-correlation [NXC]) from the B-mode image sequence. The method was based on the periodic ventricular deformation during the cardiac cycle and required the manual identification of only one ED (ED1) and one ES (ES1) frame. The remaining end-diastolic frames within the B-mode sequence were automatically identified demonstrating maximum similarity with ED1 and minimum similarity to ES1. Similarly, the remaining end-systolic frames were automatically identified demonstrating maximum similarity with ES1 and minimum similarity to ED1.

Selection of a representative reference cardiac cycle. For each cardiac cycle, a weighting factor was defined comprising the difference (in number of frames) of the current diastole and systole lengths from the mean diastole and systole lengths over the whole

Fig. 3. (a) Example measurements of the interventricular septum thickness (IVSd), left ventricular internal dimension (LVIDd) and left ventricular posterior wall (LVPWd) during end diastole (ED). (b) Example measurements of left ventricular internal dimension (LVIDs) and left atrial dimension (LADs) during end systole (ES). All measurements were made across the parasternal long-axis view of the heart.

Fig. 4. Mean absolute difference (MAD) in threshold estimation for increasing number of cardiac cycles. Separate curves for each of the four non-overlapping regions are provided.
Identification of additional control points. To provide a reliable representation of the seven independent stages of a cardiac cycle (Berne et al. 2004; Bray et al. 1999; Guyton 1991; Guyton and Hall 1997), two additional frames were introduced at regular temporal intervals in each of the contraction and relaxation phases of the reference cycle (i.e., 7 points in total). Image similarity (NXC) was then employed to identify equivalent frames within the remaining cardiac cycles. These frames would act as additional control points during the interpolation process, generating an alignment independent of the nature of the temporal variations throughout the multicycle data.

Interpolation process. The final stage of the temporal alignment between two frame sequences established a correspondence between frames in the reference and the aligned cardiac cycles. Both global and local temporal variations were addressed by applying a 1-D interpolating cubic B-spline curve (Barsky 1982; Caglar et al. 2006) between the reference and all the remaining cardiac cycles within a multicycle frame sequence.

Step 2: Spatial alignment. A rigid body transformation was found to correct for most of the spatial misalignments observed throughout the multicycle data sets (mostly caused by respiratory motion). More precisely, Nelder and Mead (1965) simplex approach was employed to derive the optimal transformation that maximised the similarity (NXC) between the registered images.

Step 3: Spatial compounding. Each frame within the reference cardiac cycle was replaced by a compound frame generated by averaging the 12 most similar spatiotemporally aligned frames, one from each cardiac cycle. Intensity averaging was employed as a well-established and effective compounding method for noise/speckle suppression in ultrasound data sets. Dissimilar frames (potentially caused by spatiotemporal misalignments) were automatically identified, employing a threshold on their image similarity with the reference frame, and discarded before spatial compounding to minimise the tissue/chamber boundary blurring being introduced on the compound images.

Clinical assessment
An experienced echocardiographer assessed the effect of proposed image enhancement approaches on the diagnostic value of cardiac ultrasound images. Quantitative assessment was achieved by performing routine clinical measurements on ED and ES frames from both the original unprocessed and the processed data. More precisely, a sequence of ED frames was presented, and (i) interventricular septal thickness (IVSd), (ii) left ventricular internal dimension (LVIDd) and (iii) left ventricular posterior wall (LVPWd) measurements were performed on each frame (Fig. 3). Similarly, a sequence of ES frames was presented, and (i) left ventricular internal dimension (LVIDs) and (ii) left atrium dimension (LADs) measurements were performed on each frame (Fig. 3). The selected measurements are widely used during routine clinical cardiac ultrasound examinations and provide valuable information on the state and function of the examined heart. More information is available in Feigenbaum et al. (2005). Each frame sequence contained one original and three enhanced frames (compounded, contrast enhanced and combination) for each of the data sets (128 frames in total). The order of the frames was randomised to ensure no bias in the results. All clinical measurements were performed twice, according to the standards adopted by the British Society of Echocardiography (Feigenbaum et al. 2005; Fuster et al. 2008; Henry et al. 1980; Wharton et al. 2012), to enable the examination of measurement repeatability and agreement among the techniques (Bland and Altman 1986).

Table 1. Numbers and percentages of 32 cardiac ultrasound data sets exhibiting underutilised gray levels within the image intensity range

<table>
<thead>
<tr>
<th>Sector</th>
<th>No. of data sets with ( t_h &gt; 0 )</th>
<th>No. of data sets with ( t_h &lt; 255 )</th>
<th>No. of data sets with underutilised intensities</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>14 (43.8%)</td>
<td>7 (21.9%)</td>
<td>18 (56.3%)</td>
</tr>
<tr>
<td>2</td>
<td>22 (68.8%)</td>
<td>6 (18.8%)</td>
<td>26 (81.3%)</td>
</tr>
<tr>
<td>3</td>
<td>9 (28.1%)</td>
<td>19 (59.4%)</td>
<td>25 (78.1%)</td>
</tr>
<tr>
<td>4</td>
<td>5 (15.6%)</td>
<td>11 (34.4%)</td>
<td>13 (40.6%)</td>
</tr>
<tr>
<td>Total</td>
<td>50 (39.1%)</td>
<td>43 (33.6%)</td>
<td>82 (64.1%)</td>
</tr>
</tbody>
</table>

Table 2. Range of total number of underutilised intensity levels within the 32 cardiac ultrasound data sets

<table>
<thead>
<tr>
<th>Underutilised level</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Mean</th>
<th>Standard deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0</td>
<td>146</td>
<td>28</td>
<td>34</td>
</tr>
</tbody>
</table>
To ensure that an accurate and robust tissue/chamber threshold was derived, images from multiple cardiac cycles were utilised. More precisely, the mean absolute difference (MAD) of the derived threshold $t_{th,c}$ for increasing number cardiac cycles was estimated as

$$MAD(t) = \frac{1}{K} \sum_{k=1}^{K} \sqrt{(t_{th,c}(t) - t_{th,c}(t-1))^2} \quad [13]$$

with $K = 32$ being the number of examined data sets and $t$ being the number of cardiac cycles employed in the threshold $t_{th,c}$ estimation of the $k$th data set. Figure 4 illustrates the MAD for increasing number of cardiac cycles for each of the four non-overlapping subsectors. In all subsectors, the threshold estimation converges (MAD < 0.5) within eight cardiac cycles. Sector 1 at the top of the field of view exhibits (because of its small size) the most variate behavior in the threshold estimation over series of eight or more cardiac cycles. Twelve cardiac cycles have previously been reported to provide a good trade-off between image enhancement and acquisition and processing requirements in temporal compounding (Perperidis et al. 2015). Although fewer cardiac cycles (e.g., 8 cycles) would suffice when 12 cardiac cycles cannot be obtained, images from 12 cardiac cycles were used in the estimation of the results presented in the remainder of this study.

Each step of the algorithm was performed on a number of non-overlapping subsectors along the y-axis of each cardiac scan. A prerequisite for each subsector was that it contain both cardiac tissue and chamber structures so that temporal histogram variations could generate a representative tissue/chamber threshold. After thorough examination, it was empirically estimated that four subsectors (Fig. 2) were the most computationally efficient option (smallest number of subsectors) that provided a fair representation of the (i) near-field noise, (ii) right ventricle (RV) chamber, (iii) left ventricle (LV) chamber and (iv) left atrium (LA) chamber.

Table 3. Comparison of the effects of individual and integrated methods on image quality metrics

<table>
<thead>
<tr>
<th></th>
<th>Temporally compounded</th>
<th>Contrast enhanced</th>
<th>Contrast enhanced and temporally compounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tissue SNR</td>
<td>87.1% (63.0%)*</td>
<td>−37.0% (15.9%)</td>
<td>14.1% (38.3%)</td>
</tr>
<tr>
<td>Chamber SNR</td>
<td>143.1% (75.0%)</td>
<td>−26.8% (17.4%)</td>
<td>78.2% (59.7%)</td>
</tr>
<tr>
<td>Contrast</td>
<td>−3.1% (16.4%)</td>
<td>73.8% (48.1%)</td>
<td>72.1% (56.4%)</td>
</tr>
<tr>
<td>SDNR</td>
<td>128.6% (85.1%)</td>
<td>120.9% (52.3%)</td>
<td>401.4% (182.3%)</td>
</tr>
</tbody>
</table>

SDNR = signal difference-to-noise ratio; SNR = signal-to-noise ratio.
* Mean (standard deviation).

RESULTS

To ensure that an accurate and robust tissue/chamber threshold was derived, images from multiple cardiac cycles were utilised. More precisely, the mean absolute difference (MAD) of the derived threshold $t_{th,c}$ for increasing number cardiac cycles was estimated as

$$MAD(t) = \frac{1}{K} \sum_{k=1}^{K} \sqrt{(t_{th,c}(t) - t_{th,c}(t-1))^2} \quad [13]$$

with $K = 32$ being the number of examined data sets and $t$ being the number of cardiac cycles employed in the threshold $t_{th,c}$ estimation of the $k$th data set. Figure 4 illustrates the MAD for increasing number of cardiac cycles for each of the four non-overlapping subsectors. In all subsectors, the threshold estimation converges (MAD < 0.5) within eight cardiac cycles. Sector 1 at the top of the field of view exhibits (because of its small size) the most variate behavior in the threshold estimation over series of eight or more cardiac cycles. Twelve cardiac cycles have previously been reported to provide a good trade-off between image enhancement and acquisition and processing requirements in temporal compounding (Perperidis et al. 2015). Although fewer cardiac cycles (e.g., 8 cycles) would suffice when 12 cardiac cycles cannot be obtained, images from 12 cardiac cycles were used in the estimation of the results presented in the remainder of this study.

Each step of the algorithm was performed on a number of non-overlapping subsectors along the y-axis of each cardiac scan. A prerequisite for each subsector was that it contain both cardiac tissue and chamber structures so that temporal histogram variations could generate a representative tissue/chamber threshold. After thorough examination, it was empirically estimated that four subsectors (Fig. 2) were the most computationally efficient option (smallest number of subsectors) that provided a fair representation of the (i) near-field noise, (ii) right ventricle (RV) chamber, (iii) left ventricle (LV) chamber and (iv) left atrium (LA) chamber.

Fig. 5. Binary masks overlaid on top of B-mode end-diastole frames of (a,b) low, (c,d) average and (e) high diagnostic value.
Underutilised intensities

Table 1 enumerates the data sets identified as having underutilised gray levels within the image intensity range. The threshold identification algorithm was independently executed in four non-overlapping subsectors of each cardiac scan. Table 2 provides the range of the total number of underutilised intensity levels within the 32 cardiac ultrasound data sets.

Effect on SNR, contrast and SDNR

Two 11 × 11-pixel regions of interest (ROIs) corresponding to the IVS and the RV chamber were manually defined on each of the 32 patient data sets. The pixel intensity values within each ROI were used to estimate the tissue and chamber signal-to-noise ratio (SNR = μ/σ), where μ and σ refer to the mean and standard deviation of the corresponding ROI intensity values. In a similar manner, the tissue/chamber contrast (C = |μT − μC|/(μT + μC)/2) and signal difference-to-noise ratio (SDNR = |μT − μC|/σC), also referred to as the detectability index, were derived, where μT and μC corresponded to the mean tissue and chamber intensity levels, respectively, and σC corresponds to the chamber standard deviation. Table 3 summarises and provides a direct comparison of the mean effects of the individual and integrated image enhancement methods on all four image quality metrics.

Visual effect

Figure 5 illustrates characteristic examples of the binary masks distinguishing between cardiac tissue and chambers over a range of image qualities. Figures 6 to 10 display five example ED frames before and after three image enhancement methods (compounding, contrast enhancement, combination) are applied to their corresponding multicycle data sets. Data sets covering a range of image and diagnostic qualities (2 low, 2 average and 1 high) were selected to best illustrate the effect of each enhancement approach on cardiac ultrasound data.

Effect on diagnostic value

Bland–Altman analysis (Bland and Altman 1986) was employed for the quantitative assessment of the

Fig. 6. Images characterised as of very low diagnostic value. (a) Original, (b) compounded, (c) contrast-enhanced and (d) combined end-diastole frames.
effect of each enhancement approach on routine clinical measurements. The Bland–Altman analysis derives the coefficient of repeatability (CR) denoting (i) the level of repeatability of clinical measurements performed using a single approach, and (ii) the level of agreement between corresponding measurements performed using different approaches. In all cases, the lower the CR, the higher are the measurement repeatability and agreement. Moreover, the mean difference indicates the presence of any bias in the corresponding measurements. Table 4 summarises the bias and repeatability level coefficients derived from the individual plots for each clinical measurement. Similarly, Table 5 summarises the bias and agreement of each clinical measurement when performed on enhanced images, compared with measurements on the original, unprocessed data.

**DISCUSSION**

In this study, tissue/chamber contrast enhancement was applied to the original B-mode data before spatial compounding. Inverting the process is expected to generate a similar effect on cardiac ultrasound data. However, spatial compounding can sometimes introduce minor tissue/chamber boundary blurring, which in turn can have a limiting effect on the derivation of an effective threshold. Furthermore, the effectiveness of compounding is heavily based on the accurate alignment of the multicycle data. Suppressing chamber noise and enhancing tissue/chamber contrast before spatial compounding can therefore increase the accuracy of the spatiotemporal data registration, resulting in more effective image enhancement.

*Identify underutilised intensities*

Table 1 indicates that 64.1% of the data sets included underutilised intensity levels spread fairly evenly in the lower (33.6%) and higher (39.1%) ends of the [0–255] image intensity range. On average, 11% (28 levels) and up to 57% (149 levels) of the available intensity range were estimated to be underutilised. A substantial 72% of the cases demonstrating $th_t > 0$ occurred on the top two subsectors of the field of view, whereas 70% of the cases demonstrating $th_t < 255$ occurred on the bottom two subsectors. This observation suggests that signals at low-depth regions tend to be overamplified, whereas signals at larger depths are underamplified. This behaviour

*Fig. 7. Images characterised as of low diagnostic value. (a) Original, (b) compounded, (c) contrast-enhanced and (d) combined end-diastole frames.*
can be partially attributed to non-optimal TGC manually set during data acquisition. Identifying and compensating for such underutilised intensity levels can provide a partial, indirect solution to the problem. Moreover, it provides a simple process increasing the dynamic range in the imaged structure for a more robust subsequent tissue/chamber identification.

Effect on SNR, contrast and SDNR

As outlined in Table 3, temporal compounding suppressed speckle and noise, increasing the average tissue (87.1%) and chamber (143.1%) SNR, as well as the associated tissue/chamber detectability (128.6%). On the other hand, temporal compounding had a negligible effect on tissue/chamber contrast. The proposed DHBIM enabled a large mean increase in contrast (73.8%) and detectability (120.9%) between cardiac tissue and chambers. The drop observed in the estimated SNRs should not necessarily be associated with increase in the noise within the signal. The tissue SNR increase is mostly generated by enhancing the underlying speckle patterns, providing valuable (e.g., in speckle tracking applications) information on the imaged structure and relevant motion patterns. On the contrary, the chamber SNR is mostly noise, and its suppression is always desirable. However, spurious high-intensity noise in cardiac chambers can sometimes be misclassified (Fig. 5a,b,e) and enhanced as tissue, but, as illustrated in Figures 6 to 10, without any noticeable detrimental effect in the processed images. Simple morphologic filtering in the binary mask before the intensity mapping can also partially compensate for such spurious pixels. Finally, the integrated approach, combining contrast enhancement and temporal compounding, provides the only method improving all four image quality metrics, yielding a large mean increase in chamber SNR (78.2%) and tissue/chamber contrast (72.1%), as well as an impressive mean increase of 401.4% in tissue/chamber detectability (SDNR).

Visual effect

Figures 6 to 10 provide characteristic examples of the effects of each enhancement approach on clinical data (12 consecutive cardiac cycles) over a range of image quality and diagnostic value. Visual examination of the 32 data sets suggests that temporal compounding, as suggested by the relevant SNR increases (Table 3), can significantly reduce tissue speckle as well as noise in cardiac tissue and chambers. Furthermore, it can also enhance structures whose boundaries are hard to delineate because of high levels of noise or shadowing, such as...
as the RV and IVS (Figs. 8 and 10), the LVPW (Fig. 7) and the aortic valve in (Fig. 8). In Figure 6 is an example of very limited diagnostic value; although the quality of the data set remained low, structures such as the RV chamber and the LVPW were marginally enhanced, making them easier to delineate. On the other hand, temporal compounding has no noteworthy effect on tissue/chamber contrast and may introduce modest tissue/chamber boundary blurring, mostly around rapidly moving structures, such as valves (mitral valve in Figs. 8–10).

Visual examination of 32 data sets indicated that DHBIM can (i) generate a representative binary cardiac tissue mask (Fig. 5), (ii) suppress the cardiac chamber noise levels and (iii) increase the delineation of cardiac tissue. In general, the lower the image corruption levels are, the more representative the binary mask is. However, even in highly corrupted cases, the underlying tissue structure is fairly delineated (Figs. 5a and 6). Noticeable noise suppression occurs in the top half of the RV chamber, typically corrupted by near-field static noise (Figs. 6–10). Applying the algorithm in four non-overlapping sectors partially compensated for the variations in mean intensity and noise level among cardiac structures. A limitation of DHBIM occurs when high levels of noise, along with artifacts such as shadowing, mask out certain cardiac structures (IVS in Fig. 8). In such cases, the algorithm can misclassify and suppress the relevant pixel intensities of the corrupted tissue. Nevertheless, DHBIM has not been observed to suppress vital clinical information that is present in the original data set.

As illustrated in Figures 6 to 10, the integrated image enhancement method benefits from the advantages, whereas it compensates for the limitations of each individual method. In the first instance, the chamber noise is suppressed and the tissue/chamber contrast and detectability are enhanced, making cardiac structures clearly distinguishable. Likewise, tissue speckle as well as acoustic noise is further suppressed through spatial averaging. Averaging of spatiotemporally aligned frames also compensates for tissue structures that are heavily corrupted by noise or shadowing, which are otherwise misclassified and suppressed as chamber noise (aortic valve in Fig. 7, IVS in Fig. 8) by DHBIM. Furthermore, suppressing chamber noise and enhancing tissue/chamber contrast before image alignment result in a moderate reduction to the boundary blurring introduced during spatial compounding (mitral valve in Figs. 8 and 10). Nevertheless, low blurring levels are still observed, mostly around fast moving structures, such as valves, with some blurring identified along

![Fig. 9. Images characterised as of average diagnostic value. (a) Original, (b) compounded, (c) contrast-enhanced and (d) combined end-diastole frames.](image-url)
chamber walls. The blurring effect is partially due to quantification errors as a result of the limited acquisition frame rate. Cardiac ultrasound scanners can acquire B-mode sequences at frame rates $\geq 100$ Hz. Such an increase in temporal resolution (currently 25 Hz) is expected to further decrease the level of the tissue/chamber boundary blurring introduced.

**Effect on diagnostic value**

The repeatability and agreement plots derived for each of the five clinical measurements revealed (i) no major outliers and (ii) no substantial or systematic bias within each method or among the original and processed data. Furthermore, as suggested by Tables 4 and 5, the measurement agreement of each enhancement approach with the original measurement is on the same scale as the relevant measurement repeatability levels. Consequently, the results indicate a strong potential for the original and processed data to be interchangeable when performing cardiac measurements, and act as an additional aid in the diagnostic process, especially in challenging cases with corrupted imaged structures.

![Fig. 10. Images characterised as of high diagnostic value. (a) Original, (b) compounded, (c) contrast-enhanced and (d) combined end-diastole frames.](image-url)
Table 5. Agreement of clinical measurements performed on processed images compared with measurements on original data

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Compounded Bias</th>
<th>Compounded Agreement</th>
<th>Contrast enhanced Bias</th>
<th>Contrast enhanced Agreement</th>
<th>Integrated Bias</th>
<th>Integrated Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>IVSd</td>
<td>−0.36</td>
<td>3.63</td>
<td>−1.18</td>
<td>2.98</td>
<td>−1.00</td>
<td>3.62</td>
</tr>
<tr>
<td>LVIDd</td>
<td>0.19</td>
<td>7.17</td>
<td>0.60</td>
<td>5.73</td>
<td>−0.02</td>
<td>7.24</td>
</tr>
<tr>
<td>LVPWd</td>
<td>−0.24</td>
<td>3.37</td>
<td>−0.72</td>
<td>3.24</td>
<td>−0.45</td>
<td>3.51</td>
</tr>
<tr>
<td>LADs</td>
<td>0.48</td>
<td>3.94</td>
<td>−1.52</td>
<td>3.12</td>
<td>−1.18</td>
<td>2.72</td>
</tr>
<tr>
<td>LVIDs</td>
<td>−0.57</td>
<td>5.13</td>
<td>−1.28</td>
<td>5.03</td>
<td>−1.81</td>
<td>6.16</td>
</tr>
<tr>
<td>Combined</td>
<td>−0.08</td>
<td>4.92</td>
<td>−0.81</td>
<td>4.43</td>
<td>−0.87</td>
<td>5.09</td>
</tr>
</tbody>
</table>

IVSd = end-diastole interventricular septal thickness; LVIDd = end-diastole left ventricular internal dimension; LVPWd = end-diastole left ventricular posterior wall; LADs = end-systole left atrium dimension; LVIDs = end-systole left ventricular internal dimension.

particular, all three approaches improved the repeatability of the clinical measurements, with the integrated approach achieving the highest overall improvement of 24% (and highest improvement in three of the five measurements). On the other hand, DHBIM contrast enhancement achieved the lowest overall improvement with a reduction of the repeatability in two of the five measurements (LVPW and LAD). This reduction is potentially a result of the inability of DHBIM to compensate for heavily corrupted tissue structures that are therefore misclassified and suppressed as chamber noise. However, as suggested by their visual effect on cardiac images and indicated by the associated increase in LVPW and LAD measurement repeatability (13% and 16% respectively), the integrated approach can compensate for this limitation. Clinical measurement repeatability is expected to improve as the familiarity of the echocardiographer with the processed images increases. Given no observed degradation of high-quality images, the proposed integrated method has the potential to replace or act as an adjunct to existing image processing and display methods in ultrasonic scanners.

**Limitations and future work**

There are a number of limitations that, although beyond the scope of this work, will need to be addressed in future studies. In particular, the proposed algorithms can be further improved by making temporal compounding a fully automated approach. Currently, a single pair of ED and ES frames need to be manually identified by the clinician. The remaining ED and ES frames are then automatically identified using image similarity measures. Additional information, such as ECG data, LV shape/size parameters and tissue movement direction, can be incorporated to automatically classify all ED and ES frames. Furthermore, the current implementations of the proposed algorithms were aimed at the off-line processing and enhancement of cardiac ultrasound images. However, the real-time aspect of cardiac ultrasound constitutes a major advantage over other imaging modalities. Implementing the algorithms for the real-time or near real-time enhancement of cardiac ultrasound images can be of great benefit to the diagnostic process. Finally, the proposed algorithms have been thoroughly evaluated, on both image quality metrics as well as routine clinical measurements. However, a more extensive clinical assessment of the algorithms, including (i) direct comparison (on the same data) with other state-of-the-art enhancement approaches, (ii) on a wider range of imaging views (in addition to PLAX), and (iii) with a larger number of assessing expert echocardiographers will be of great value. Any subsequent study needs to ensure that images are acquired using a state-of-the-art device with optimal acquisition parameters such as axial, lateral and temporal resolution.

**CONCLUSIONS**

This study introduced DHBIM, a novel approach that employs variations of the cumulative histograms over time to estimate an intensity threshold between cardiac tissue and chambers and subsequently enhance their respective contrast. DHBIM reduces noise in cardiac chambers and improves the tissue/chamber contrast and detectability levels considerably. However, DHBIM does not compensate for tissue speckle and other common cardiac ultrasound artefacts such as shadowing and reverberations. Furthermore, there are cases in which heavily corrupted cardiac tissue structures are misclassified and suppressed as noise. Combining DHBIM with a temporal compounding to an integrated methodology can compensate for many of their individual limitations generating low-noise and high-contrast images with well-delineated cardiac structures. Assessment of the effect of each of the approaches on the image quality and diagnostic value of cardiac ultrasound data verified that although all three approaches can enhance the processed images, the integrated approach provided the best overall improvement.

**Acknowledgments**—This work was supported by the Engineering and Physical Sciences Research Council (EPSRC) through a Biologically Inspired Acoustic Systems grant (EP/C523776/1).
Dynamic enhancement of B-mode cardiologic US images • A. Perperidis et al. 15

REFERENCES


Caglar H, Caglar N, Elfaituri K. B-Spline interpolation compared with finite difference, finite element and finite volume methods which applied to two-point boundary value problems. Appl Math Comput 2006;175:72–79.


