Automatic detection of complete and measurable cardiac cycles in antenatal pulsed-wave Doppler signals

Eleonora Sulas\textsuperscript{a,*}, Monica Urru\textsuperscript{b}, Roberto Tumbarello\textsuperscript{b}, Luigi Raffo\textsuperscript{a}, Danilo Pani\textsuperscript{a}

\textsuperscript{a} Department of Electrical and Electronic Engineering, University of Cagliari, Italy
\textsuperscript{b} Division of Pediatric Cardiology, San Michele Hospital, Cagliari, Italy

**ARTICLE INFO**

Article history:
Received 1 October 2019
Revised 21 December 2019
Accepted 9 January 2020

Keywords:
Ultrasound
Pattern recognition
Doppler
Fetal echocardiography

**ABSTRACT**

Background and objective: Pulsed-wave Doppler (PWD) echocardiography is the primary tool for antenatal cardiac diagnosis. Based on it, different measurements and validated reference parameters can be extracted. The automatic detection of complete and measurable cardiac cycles would represent a useful tool for the quality assessment of the PWD trace and the automated analysis of long traces.

Methods: This work proposes and compares three different algorithms for this purpose, based on the preliminary extraction of the PWD velocity spectrum envelopes: template matching, supervised classification over a reduced set of relevant waveshape features, and supervised classification over the whole waveshape potentially representing a cardiac cycle. A custom dataset comprising 43 fetal cardiac PWD traces (174,319 signal segments) acquired on an apical five-chamber window was developed and used for the assessment of the different algorithms.

Results: The adoption of a supervised classifier trained with the samples representing the upper and lower envelopes of the PWD, with additional features extracted from the image, achieved significantly better results ($p < 0.0001$) than the other algorithms, with an average accuracy of $98 \pm 1\%$ when using an SVM classifier and a leave-one-subject-out cross-validation. Further, the robustness of the results with respect to the classifier model was proved.

Conclusions: The results reveal excellent detection performance, suggesting that the proposed approach can be adopted for the automatic analysis of long PWD traces or embedded in ultrasound machines as a first step for the extraction of measurements and reference clinical parameters.

© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)

1. Introduction

The assessment of fetal cardiac function in early pregnancy represents an important clinical challenge. Congenital heart diseases (CHDs) are the most common type of birth defect [1], with a prevalence of approximately nine per 1000 live births [2]. Excluding infections, CHD accounts for the largest number of deaths during the first year of life [3]. Despite its intrinsic operator-dependency [4], fetal echocardiography is the elective tool for the early diagnosis of CHD during the second trimester in high-risk pregnancies, but several studies advocate its use also in low-risk pregnancies [5–7]. In this case, fetal echocardiography can be used to detect up to 40% of the CHDs [1].

In fetal echocardiography, pulsed-wave Doppler (PWD) can aid the diagnosis of CHD, functional problems and the assessment of cardiac rhythm [8,9]. Compared with other modalities, blood flow through the heart and great vessels, along with the movement of heart tissues, can be more objectively studied using Doppler ultrasound. To this aim, several parameters, computed from quantitative measurements of the PWD spectrum over complete cardiac cycles have been proposed and validated [10]. In order to enable their automatic computation, complete and clinically meaningful cardiac cycles must be first identified. This process is usually performed by the operator via visual inspection. However, the automatic identification could be introduced in the echocardiographs as a first step of a measurement chain. Remarkably, this process is less important in adult ultrasound examinations, since the measurement is easier, because of the position and size of the heart, and more stable, thus requiring less manual skills to the cardiologist.

\* Corresponding author.
E-mail address: sulaseleonora1992@gmail.com (E. Sulas).

https://doi.org/10.1016/j.cmpb.2020.105336
0169-2607/© 2020 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license. (http://creativecommons.org/licenses/by-nc-nd/4.0/)
Moreover, several works [11,12] studied the PWD signal by multimodal recordings including the fetal electrocardiogram (ECC). In this case, PWD can represent a reference signal for non-invasive fetal ECG analysis in early pregnancy, by providing clues on the cardiac rhythm and heart phase. In this context, characterized by long traces to be analyzed and compared, the automatic detection of the complete and measurable cardiac cycles would represent an invaluable tool for the researchers.

A preliminary investigation on the automatic identification of uncorrupted, complete, and measurable cardiac cycles in antenatal PWD signals was presented in [13,14]. This work significantly elaborates on the previous findings by proposing and comparing three different algorithms: the previous approach [13,14] and two other methods, characterized by a reduced number of features and overall computational cost. The three methods are based on a preliminary extraction of the PWD velocity spectrum envelopes, followed by different detection approaches: (i) template matching with relevant waveshapes associated with atrial systole, (ii) supervised classification over a reduced set of relevant waveshape features (fiducial points), and (iii) supervised classification on the whole waveshape potentially representing a cardiac cycle, with additional features extracted from the PWD spectrum image. The robustness of methods (ii) and (iii) was verified using two different classifiers: an artificial neural network (ANN) and a support vector machine (SVM).

Even though several methods have been presented thus far for the extraction of the PWD envelope, to the best of our knowledge, none of them aim at the automatic recognition of complete and measurable cardiac cycles. Conversely, they require the manual selection of the cardiac cycles for additional information. For instance, most require an ECG because they are ECC-gated. However, for antenatal studies, no fetal ECG utility has been established yet [9], revealing an informational gap.

The performance of the proposed methods was assessed on a custom dataset comprising real signals acquired from 25 pregnant women between the 21st and 27th weeks of gestation at the Division of Pediatric Cardiology of the San Michele Hospital in Cagliari, Italy.

2. Background

For the routine Doppler examination of patients with suspected valvular heart disease, it is usually best to begin by using an apical view (see Fig. 1a) [15]. In particular, an apical five-chamber view allows for recognizing the four cardiac chambers (in Fig. 1a VDX, ADX, VSN, ASN) and first part of the aorta (in Fig. 1a AO). By choosing this view, it is possible to analyze the diastolic and systolic functions, producing morphologically well-defined envelopes in the PWD signal, as reported in Fig. 1b and Fig. 1c. While the systolic phase produces a single peak in the flow velocity waveform (V), caused by blood flow through the aortic valve, the diastolic phase results in a typical biphasic waveform [16], with an early peak (E) determined by the passive filling of the left ventricle through the mitral valve due to the differential pressure between the two chambers, and a second peak (A) during the active atrial contraction. Overall, the E/A wave represents the mitral inflow, whereas the V wave represents the aortic outflow, and the two waves present opposite polarity.

Typical quantitative cardiac performance indexes, extracted from the PWD signal, are the heart rate [17], the E/A ratio [18], the mechanical AV conduction time interval [19] and the myocardial performance index [20]. Beyond helping in diagnosing specific diseases, time indexes are also useful indicators of fetal wellbeing [21]. In healthy fetuses, each atrial event is followed by a ventricular event, which occurs within a well-defined time-interval in normal 1:1 atrioventricular (AV) conduction [22].

2.1. State of the art on automatic Doppler velocity envelope extraction

To the best of our knowledge, there are no published works addressing the automatic detection of complete and measurable cardiac cycles from PWD, neither for adults nor for fetuses. Conversely, several works deal with the automatic tracing of the Doppler velocity envelope, which is a preliminary step required by our algorithms. As such, in this section a short review of the approaches representing the state of the art in the field is presented.

The first studies addressing the automatic tracing of Doppler envelopes were based on the edge detection techniques [23–31]. In [23], the method comprised the following steps: (i) image filtering using a Gaussian-shaped low-pass filter with $\sigma = 1.5$ to remove the high-frequency noises, (ii) detection of the horizontal axis, (iii) edge detection using a non-linear Laplace edge detector [32], (iv) suppression of spurious edges, and (v) extraction of the overall envelope function by using the biggest-gap algorithm [24].

In [24], the main steps used in [23] were adopted, but with more focus on the contrast enhancement. A k-means clustering algorithm was used to cluster the image intensities into three groups: the weak and strong parts of the signal and the pixels forming the background. From the centroids of the three groups, the authors derived two thresholds beyond the contrast enhancement to stretch the contrast and reduce the noise in the background. Moreover, [24] used a combination of the Sobel operator [33] and non-linear Laplace edge detector [32]. The envelope was obtained using a linking processing step, aimed at producing a continuous velocity envelope. Then, it was parameterized using model fitting with a sum of cosines: four for the systolic tricuspid valve regurgitation flow and five for the diastolic mitral valve inflow. Both [23] and [24] compared the obtained envelope with a manually traced envelope by a Bland-Altman analysis.

Another work [29] exploited 2D Doppler echocardiography for the noninvasive automated assessment of the aortic regurgitation severity. It was based on the following steps: (i) image filtering using a low-pass filter with a Gaussian kernel $5 \times 5$ and $\sigma = 1.5$, (ii) filtered image converted to grayscale, (iii) morphological operation of the image with disk approximations, (iv) subtraction of the images obtained during the first step from those obtained during the second and intensity adjustment, and (v) edge detection on this final image performed by Canny algorithm ($\sigma = 0.5$) to identify the peak velocity. As in the previous works, the results were compared with manual tracing of the envelopes.

The work [30] analyzed 30-s-long adult Doppler traces acquired using an apical five-chamber window positioned over the left ventricular outflow tract by exploiting the same workflow of the aforementioned works. In the region of interest of the PWD video frame, converted to grayscale, the connected areas presenting fewer than 500 pixels were removed. Then, the maximum velocity profile was extracted from the resulting image by the biggest-gap method [24]. High frequency noise was removed from the initial velocity profile using a first-order low-pass Butterworth filter, the cut-off frequency of which was estimated from the cardiac cycle length. Any frequency ten times greater than the fundamental frequency of the heart motion was filtered out. Finally, a third-order Gaussian model was adopted to fit to the velocity profile.

The study in [31] focused on the flow velocity estimation in the carotid artery and compared four different methods to binarize the image following the main steps described for the other studies. The results were evaluated through visual inspection.

Further, learning-based and probabilistic-framework algorithms for the automatic detection and segmentation of the Doppler traces were reported. In [34], an algorithm for automatically tracing the envelope of mitral valve inflow Doppler spectra was presented. The
algorithm was built on the techniques used in the probabilistic hierarchical and discriminant framework [35]. In this case, a single triangular object represented an isolated E- or A-wave, whereas a double triangle represented a pair of overlapping E and A waves. A single triangle model comprised three points: a left root, a right root and a peak. A double triangle model had five points: a left root, a right root, a left peak, a right peak, and an intersection point. Three detectors were trained: left root detector, right root detector, and peak detector. Additionally, two global box detectors were trained: a single triangle detector and a double triangle detector. Each candidate was associated with a posterior detection probability. For each shape candidate, either specified or inferred, the shape profile model was invoked to score it with a posterior probability. Based on these two probabilities, the algorithm selected the best candidates from the single and double triangle candidate pools. To quantify the performance of the algorithm, four measurements were computed: E-wave peak velocity (EPV), E-wave deceleration time, A-wave peak velocity, and A-wave duration. These values were compared with the measurements individually computed using annotations made by two experts.

Some related works focused on model-based image segmentation algorithms. In fact, knowing the expected shape can improve velocity envelope tracing. In [36], a novel model-based feedback and adaptive-weighting tracing algorithm using the Kalman filter was proposed. The algorithm incorporated a non-parametric statistical comparison of image intensities to estimate edges in noisy PWD signals, as well as a statistical shape model learnt in an offline process using manually-traced envelopes. The results were compared with the manual tracing.

Overall, the main limitation of the algorithms presented to review the state of the art is that they were not conceived for the automatic identification of clinically valuable heartbeats with adequate quality for the measurement of parameters associated with the fetal heart function. Furthermore, most of these algorithms are ECG-gated, which represents a critical problem for the fetal examination in early pregnancy, when the possibility to obtain a trustable fetal ECG trace is still a goal of the research in the field. Remarkably, this was not an issue for the discussed works, since most of them were conceived to work on adult’s signals rather than on fetal ones. Nevertheless, the importance of these algorithms lies in the capability to extract a good Doppler velocity envelope that, according to the methods studied in this work, represents the starting point for the automatic identification of the complete and measurable cardiac cycles. In the revised version, according to the Reviewers suggestion, these aspects are discussed at the end of the related section.

3. Materials and methods

In this section, the three techniques developed for the automatic identification of the complete fetal cardiac cycle in PWD signals are presented. The three methods share the same pro-
cEDURE for the extraction of the envelope from the PWD video (see Fig. 2). First, the common aspects will be described; then, the different methods will be presented in detail. All methods and video/data processing were performed using MATLAB v2017b (MathWorks Inc., Natick, MA, USA).

3.1. Data acquisition protocol

A custom dataset comprising 43 signals was adopted in order to develop and validate the three methods. The study was approved by the Independent Ethical Committee of the Cagliari University Hospital (AOU Cagliari) and performed following the principles outlined in the Helsinki Declaration of 1975, as revised in 2000. All the volunteers provided their signed informed consent to the recording protocol. According to the approved protocol, PWD signals were collected from 25 low-risk voluntary pregnant women between the 21st and 27th weeks of gestation, only with cardiology logically healthy fetuses. Two factors influenced the choice of the time frame for acquiring these signals. First, below 20 weeks of gestation, the heart of the fetus is still too small to guarantee any cardiac signal acquisition. Second, the used dataset consists of both PWD and non-invasive fetal ECG recordings. It is well known that, for the latter measurement, the signal quality is affected by the layers of different biological media existing between the fetus and the electrodes, which attenuate the ECG signal in its propagation towards the maternal abdomen. The lowest conductivity layer of the volume conductor is the vernix caseosa, a protective film that completely covers the fetus between the 28th and 32nd weeks of gestation, making it difficult to record the ECG non-invasively. Finally, the weeks of gestation between the 21st and 27th are those typically adopted for the first and most important antenatal cardiological screening by echocardiography, then the weeks in which the proposed approach could be more effectively used, also because of the more intense fetal motility.

The recording was performed on the abdomen with the subject in a comfortable semi-sitting position. The duration of the recordings was variable, from 6.4 s to 119.8 s, depending on the occurrence of fetal movements. The PWD, based on an apical five-chamber view, was performed using an iE33 ultrasound machine (Philips, Amsterdam, The Netherlands). The sweep speed and frame rate were set to 75 mm/s and 60 Hz, respectively. All of the other machine settings (e.g., gain, axis scaling, baseline, etc.) were subject-specific and maintained throughout the recording. The iE33 video had a native resolution of 1680 × 1050 pixels, and the frame rate on the video output ports was 30 Hz. All the frames were captured from the DVI output using an USB3HDACP Video Capture Device (StarTech, ON, Canada) capable of recording 1080p HD videos at a frame rate up to 60 frames per second and applying H.264 encoding.

3.2. Video-to-image conversion

The acquired video was initially converted into a single wide image. The PWD signal on the screen was updated from left to right. Once the updating reached the rightmost end of the dedicated area, the signals wrap back from the leftmost end, overwriting the previous trace. Initially, the video was decomposed into single frames, and each frame was cropped to isolate the region of interest that contained only the PWD signal. Such a region was the same for all the subjects with the adopted instrumental setting. Despite the reduction of the image size, image cropping does not reduce the resolution (in terms of pixels per inch) and then does not negatively affect the subsequent processing steps.

Using a threshold-based approach made it possible to derive an index representing the position of the updating front of the PWD velocity spectrum in every single frame. This index increases until its value suddenly decreases when the updating front wraps to the left of the image. When this condition occurred, the previous frame was saved. This frame was appended to the previously saved image. These steps were repeated until the end of the signal. At the end of the video processing, the result was an image containing all the PWD without interruption. Since the position of the fetus in the womb and consequently the position of the probe on the maternal abdomen were different for each subject, the mitral blood inflow could be directed towards the probe or moving away from it. Therefore, the EA - V can exhibit a positive or negative balance when there is a positive EA - wave or the opposite polarity occurs, respectively. All the signals with a negative balance were inverted to have a homogeneous dataset with positive balance only.

3.3. Envelope extraction

On the single wide image representing the whole PWD signal of interest, the velocity envelopes were extracted in the form of two 1D signals: one for the upper (positive) part of the signal and the other for the lower (negative) part. The envelope extraction involved the following three steps to finally identify the boundary that separates the PWD spectrum from the background.

1. Image binarization

Initially, the single wide PWD image was converted to greyscale. The definition of the binarization threshold on this image was limited by the different device settings (i.e., different gain setting, leading to different intensities among the datasets) and by the presence of background noise. Based on 2D Otsu’s method, we adopted a global threshold from a gray-level-median histogram [37]. After a comparison with the original Otsu’s method and a fixed manual threshold, this threshold was selected as the most efficient and robust regarding noise and information preservation [13].

2. Area opening

The binarized image was processed using a morphological operator (area opening) to reduce the presence of small spurious areas spread over the image, caused, for instance, by a high gain setting. Such noise significantly limits envelope extraction. Rather than using a higher threshold for binarization, this solution is more conservative regarding the information content of the image. We empirically chose to remove all four connected components with a small area (70 pixels). In addition to the noise, this step removed the dotted line characterizing the vertical grid in the video.

3. Edge detection

At this point, a binary and nearly noise-free image of the PWD spectrum was available. The image was converted into the two 1D envelopes representing its upper and lower pro-
Table 1
Characteristics of the methods assessed for the identification of the complete and measurable fetal heartbeats from 5-chamber apical view PWD traces.

<table>
<thead>
<tr>
<th>Technique</th>
<th>Method 1</th>
<th>Method 2</th>
<th>Method 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Classifier</td>
<td>Template matching</td>
<td>Supervised classification</td>
<td>Supervised classification</td>
</tr>
<tr>
<td>Parameters</td>
<td>None</td>
<td>SVM, ANN</td>
<td>SVM, ANN</td>
</tr>
<tr>
<td>Features</td>
<td>5 templates</td>
<td>15 features</td>
<td>264 features</td>
</tr>
</tbody>
</table>

file, respectively:
\[ G_u(x) = \arg \max_y \{ f(x, y) = 1 \} - y_b \]  
\[ G_l(x) = \arg \min_y \{ f(x, y) = 1 \} - y_b \]

where \( y_b \) is the baseline position, that is the horizontal axis line that separates the image into two parts containing the negative and the positive waves, whose polarities depend on the direction of the blood flow, respectively if it is moving away or going towards the transducer. Fig. 1d shows the PWD signal with the associated \( G_u \) and \( G_l \) curves. Finally, the two envelopes were normalized, between 0 and 1 and between -1 and 0, respectively, for the automatic detection of the complete fetal cardiac cycle only.

Based on the time information and image resolution, the envelopes presented 284 samples per second, which can be assumed to be the sampling rate for these signals.

3.4. Automatic detection methods

Based on the original single wide PWD image, the envelopes were affected by different artifacts. For instance, with a moderate increase in gain, the velocity tracing thickened and the spectrum broadened [38]. Additionally, a mirror image artifact could appear. Furthermore, with the chosen apical five-chamber view, mitral and aortic valve closure were typically present as spikes in the PWD envelopes. Regardless of the method, the presence of these residual artifacts hampers automatic detection. The occurrence of such artifacts must be considered as a disturbance in the adopted dataset, but no specific technique was implemented in order to reduce their effect on the signals. Moreover, due to the approved recording protocol, only the best settings were used in the recordings, since the signals were acquired during real clinical examinations. For this reason, a point assessment of the robustness of the methods with respect to the occurrence and entity of these artifacts could not be performed in the scope of this work.

In the following sections, the details of the three methods developed and tested in this work are presented. Table 1 summarizes the main differences between them.

3.4.1. Method 1: template matching

The morphology of the PWD presents a typical EA-V pattern of the chosen apical five-chamber view of healthy fetuses. In particular, the atrial activity is characterized by two peak velocities, \( E \) and \( A \), which form an “M” shape. The shape of the \( V \) wave, that is present after the \( EA \)-wave, is reasonably the same across subjects, changing only for its amplitude. However, \( EA \)-wave can slightly change, as can be seen from Fig. 3a. By identifying different templates for the atrial activity, well represented in the available signals, it is possible, in principle, to identify the complete fetal heart beats by a template matching approach. Fig. 3b shows the selected set of atrial activity templates (upper envelope) used for this method, that respectively corresponds to the real PWD atrial parts in Fig. 3a. Such templates were chosen by looking at the PWD signals available in the adopted dataset, according to the pediatric cardiologist’s indications, in order to represent the largest part of EA-patterns.

The \( E \) and \( A \) waves of the transmural spectral Doppler envelope could be completely detached, with an \( E \)-wave that is occasionally barely identifiable as a separate peak. By considering the envelopes sampling rate and the mean cardiac cycle duration, which is approximately 128 samples, and considering that the available dataset includes only healthy fetuses with a 1:1 AV conduction, on average, the length of the templates for the fetal atrial activity is about 64 samples.

A sample-by-sample sliding window to compute the normalized cross-correlation between the five templates and the whole upper PWD envelope was adopted. A complete and measurable fetal cardiac cycle was detected using this method when at least one cross-correlation signal in that point was above a threshold (equal to 0.6) chosen to maximize the performance on the available dataset. This simple method did not require any classification stage downstream as opposed to the other methods.

3.4.2. Method 2: classification based on fiducial points

A more sophisticated delineation-based approach was developed to characterize the envelopes over a time frame lasting as long as a full cardiac cycle. To this end, a supervised classification approach was adopted by extracting meaningful features from each signal frame. Eight fiducial points characterizing the two envelopes (five for the upper envelope and three for the lower envelope, see Fig. 3c) were identified:

1. Onset of the \( E \) wave (\( E_o \))
2. \( E \) peak (\( E \))
3. The intersection point between the \( E \) and the \( A \) waves (\( EA \))
4. \( A \) peak (\( A \))
5. End point of the \( A \) wave (\( Ae \))
6. Onset of the \( V \) wave (\( Vo \))
7. \( V \) peak (\( V \))
8. End point of the \( V \) wave (\( Ve \)).

In order to smooth the velocity profile, a three-tap median filter was applied to the initial velocity profile. Then, the average fetal cardiac cycle duration was estimated for that signal by detecting the \( V \) wave in the lower envelope and computing the \( V-V \) intervals. Unfortunately, valve closure clicks can affect the estimation of the \( V-V \) interval, which is used to set some thresholds needed for the fiducial points detection. Fiducial points were detected in the order reported in the aforementioned list.

Because the onset of the \( E \) wave lays on the \( x \) axis, it was detected by selecting the first non-zero sample after a sequence of zeros. Since, in the presence of noise, the onset is no longer on the
x axis, if such method fails, the E wave onset was detected by imposing the constraint that the distance between two consecutive onsets should be not larger than the duration of two average cardiac cycles, and detecting the local minimum within a window of reasonable length, according to physiological constraints. Furthermore, because of the presence of the valve clicks, the detection of some E-wave onset points could be erroneously marked at the onset of the valve click rather than at its end. To correct these cases, the valve click was detected by using the first derivative of the upper envelope and the E-wave onset identified at the end of the valve click.

The A peak was detected as the local maximum in the interval between the E onset and half of the average cardiac cycle duration. Then, the E peak was detected as the local maximum in the interval between the E wave onset and the A peak.

The conjunction between the E and A waves was detected as a minimum (see the first, second, fourth and fifth templates in Fig. 3b) or a flat portion (see the third template in Fig. 3b) of the curve. The last point of the upper envelope was the end of the A wave, recognized by the sharp decrease in the envelope, taking the minimum reached in the declining slope.

Regarding the lower envelope, the first point to be searched was the onset of the V wave owing to the known position of the end of the previous A wave. Then, the end of the V wave was found in the interval lasting as long the average cardiac cycle duration and starting from the onset of the V wave. Again, the first derivative of the velocity curve was used to correct the position of that point in case of the presence of valve click artifacts.

Between the onset and end of the V wave, the local minimum was recognized as the V peak. These eight fiducial points can be described using 16 coordinates (time and amplitude). All the time coordinates can be referred to the E wave onset of that cardiac cycle, such that their values are not dependent on their occurrence over time in that cycle. This implies that the E onset time is no longer useful for the remainder of the algorithm; as such, seven time features can be extracted from the eight fiducial points, i.e., the latency from the E wave onset. The other eight amplitude features can be extracted, as the amplitudes of the eight fiducial points, for a total of 15 features. Then, a supervised classifier was trained to identify the complete and measurable cardiac cycles in the PWD signal from these features. Since this last step is common to the third method as well, it will be described only once in Section 3.5.

3.4.3. Method 3: classification based on the envelope, area and pixels

Residual ultrasound artifacts, including mirror images, may hide the cardiac cycle information, limiting the fiducial point detection and reducing the effectiveness of template matching. The third method was created to overcome these limitations exploiting as classifier features the complete envelopes of the PWD signal. This approach, first investigated in [13] and [14], provides the largest set of available information to the classifier without any preliminary identification of the waveforms of interest or any delineation of
the PWD trace, the errors of which could impact the quality of the final solution. Simultaneously, the dimensionality of the problem was significantly reduced using the envelopes rather than the image. The upper and lower envelopes were considered over a time frame representative of the duration of a fetal cardiac cycle in the studied gestational weeks, as previously performed for the template matching approach. Then, an analysis window size of 128 samples was adopted, leading to 256 features for the classifier (128 for each of the upper and lower envelopes).

Moreover, four additional features associated with the area under the curve were introduced. Based on an AV conduction of 1:1 in healthy fetuses so that the ventricular activity was always expected after the atrial activity, the areas between the x axis and upper and lower profile envelopes were computed on the first and second half of the analysis window.

Furthermore, four additional features based on the pixel analysis were added to improve the classifier performance (shown in the work [13]). The brightness information is very useful for the recognition of the meaningful part of the image. In fact, ultrasound artifacts could be less bright than the portion of the signal of interest; moreover, the cardiac cycle parts are characterized by the same brightness across different heart beats in the same trace. Therefore, the mean brightness of the pixel enclosed in the four areas were computed. Overall, 264 features were chosen: 128 samples of $G_u$, 128 samples of $G_l$, four area features and four brightness features.

3.5. Pattern recognition

Two different classifiers were adopted to assess the robustness of the second and third methods described in the previous sections for the automatic identification of complete and measurable fetal cardiac cycles in the PWD signal: ANN [39] and SVM [40]. Therefore, the classification problem has a binary output: 0 for an incomplete, malformed or unrecognized cardiac cycle and 1 for a complete and measurable one. The use of an ANN was previously investigated in [13] for the third method, whereas the SVM was introduced in this work because of its capability to solve non-linearly separable classification problems by the adoption of non-linear kernels.

Depending on the method (exploiting the fiducial points or whole envelope), the ANN was characterized by a different number of input neurons. However, in both cases, ten hidden neurons and two output nodes were always used, and the scaled conjugate gradient backpropagation algorithm was exploited for training. The number of neurons in the hidden layer was empirically chosen to obtain an acceptable performance on the available dataset without incurring an overfitting problem.

A second classifier was also tested, i.e., an SVM with a Gaussian radial basis function kernel, $\nu = 0.5$, and box constraint set to 1. Classification was performed using the MATLAB Neural Network Toolbox and MATLAB Statistics and Machine Learning Toolbox.

3.6. Dataset labelling

For a quantitative assessment of the algorithm performance, 174,319 signal windows were labelled by using a custom graphical user interface developed using MATLAB. The interface allowed the cardiologist for scrolling over the trace to label multiple cardiac cycles. In order to represent the window under examination to the clinician, the interface showed the PWD image with two rectangular canvases, one enclosing the atrial activity (red box, 64 sample long) and one the ventricular activity (blue box, 64 sample long). The resulting window size was 128 samples, as it was explained in Section 3.4.3. If the two boxes overall contained a well-recorded cardiac cycle in terms of both atrial and ventricular activities, thus useful for clinical inspection and parameters computing, the cardiologist could label the selected window as “complete and measurable”, otherwise she could keep scrolling over the trace.

Fig. 4 shows some examples of cardiac cycles labelled as “complete and measurable”, Fig. 4a, and a segment of PWD signal where no one was labelled in that way, Fig. 4b. As can be seen from Fig. 4a, complete cycles are neither noise-free nor perfect from a morphological perspective, conversely being affected by artifacts but still meaningful for diagnostic purposes.

Moreover, for each window labelled as complete and measurable, the tool also applied the same label to those preceding and following it, up to 15 samples. Similarly, incomplete or unreadable beats were taken on the remainder of the traces. Overall, 87,736 signal windows represented complete and measurable cycles, and 86,583 windows represented incomplete or malformed fetal cardiac cycles. Such an approach led to a balanced dataset.

3.7. Comparative analysis method

To assess the performance of the different methods, the accuracy (Acc) was computed as follows:

$$\text{Acc} = \frac{TP + TN}{P + N} \times 100 \quad (3)$$

where $TP$ is the number of true positive detections, $TN$ is the number of true negative detections, $P$ is the number of positive samples, and $N$ is the number of negative samples. For the template matching technique, the results were compared directly with the annotated cardiac cycles leading to a single accuracy value.

Instead, the validation of the classifiers was performed using a leave-one-subject-out scheme, which in this case is equivalent to a 25-fold scheme without random sampling. This allowed the achievement of a fairer test as no PWD signal from a given subject was used for training the classifier when that subject was used for testing; however, the same would not have occurred in the case

(a) Cardiac cycles labelled as “complete and measurable”.

(b) Segment of a PWD trace where no cardiac cycle was labelled as “complete and measurable”.

Fig. 4. Labelling procedure.
of random sampling for training and testing. Remarkably, since the dataset included 43 traces taken from 25 subjects, there could be more than one trace per subject. Based on this validation approach, 25 different error predictions for the 25 folds were obtained, and the results are presented using box-and-whiskers plots reporting the median, the 25th and 75th percentiles (box), the smaller and higher values not representing outliers (whiskers), and the outliers (crosses).

Furthermore, statistical analysis was performed to investigate the significant differences between the following: (i) the two classifiers (ANN and SVM) among the second and third techniques and (ii) the second and third methods. First, the normality of the distributions was inspected by using the Lilliefors test. Since the outcome of the test allowed to safely reject the normality hypothesis, we adopted the Wilcoxon signed rank test, a non-parametric statistical test, to study the significance of the results, considering a significance level $p < 0.05$.

4. Results

The accuracy for the first method (template matching) is limited to 48.5%. The confusion matrix, which represents this result (Fig. 5a), shows a really high number of the FP with the chosen threshold. By increasing the threshold, the number of FP decreases, but the number of FN grows. The selected threshold was chosen to be a good compromise.

From a comparison of the second and third methods, the superiority of the latter compared with the former is evident ($p < 0.0001$ regardless the classifier model). Fig. 5b allows appreciating the robustness of the proposed solutions with respect to the chosen classifier model, excluding a large bias in the results due to this choice. The SVM and ANN performed similarly. However, from the statistical analysis it is possible to exclude the significant superiority of a classifier compared with the other for the second method, where the achieved $p$-value was 0.97, whereas the $p$-value was 0.001 for the third method, which revealed a significantly better performance of the SVM with respect to the ANN. Feature reduction by principal component analysis (PCA) applied before the classifier did not lead to any significant difference compared to the adoption of the whole feature set for both the approaches.

5. Discussion

Regarding the first method, the poor performance could be attributed to the selection of templates based on the fetal heart rate. To exclude the possible occurrence of this effect, different heart rates within the dataset were checked to prevent the presence of outliers that could affect the performance of the algorithm. Since no outliers were detected, the results of the first method were not influenced by the chosen length of the templates. Another explanation for this unsatisfactory result could be the set of templates identified to represent the atrial activity patterns. However, the templates were extracted as those resembling the largest part of the atrial activities in the beats annotated as complete and measurable by the cardiologist. Further analyses, possibly including the ventricular activities, did not improve the algorithm performance (data not shown). Another parameter that could have influenced the results is the template matching threshold, but it was selected to be a good trade-off in the classification performance. In the end, the noisy nature of the signals, imperfect envelopes, and presence of the valve clicks could have contributed to this poor performance.

The reduced feature set for the second method, possibly along with an imprecise identification of the fiducial points in the presence of noise, was probably responsible for its relatively poor performance. This problem could be overcome by a more sophisticated algorithm for the detection of the fiducial points characterizing $EA$ and $V$ waves. However, either the optimization of the delineation algorithm or its validation through manual annotation (requiring a huge annotation work by the cardiologists) was beyond the scope of this research. Conversely, the third method, not requiring complex feature extraction stages, was able to accurately distinguish complete and measurable fetal cardiac cycles from incomplete or not meaningful cycles, reaching a mean accuracy of 98% with a standard deviation of 1% using the SVM classifier (97.63% ± 2.09% for ANN). Remarkably, the number of features is significantly lower than the number of pixels in the images corresponding to the extracted windows, so the complexity of the classifier model is reduced. However, the number of features is in any case largely higher for the third method than for the second one, so the better results could be ascribed to the wider knowledge available by the third technique, even after possible reduction in the number of adopted features by PCA. Nevertheless, the performed comparison was needed in order to evaluate if a classifier exploiting a reduced feature set providing the most relevant information about the PWD envelope could be effectively used, since in this case the computational complexity would be substantially reduced. Even though, in the light of a possible integration of the proposed methods in an ultrasound device, this aspect is of paramount importance, the results suggest to avoid considering a reduced set of features in order to preserve the maximum possible accuracy.
By looking at the presented results, some important remarks should be discussed. The achieved results could have been biased by the labelling performed by a single cardiologist. Having a second expert would avoid any bias but, considering the amount of time for the labelling procedure and the required expertise, it was impossible to obtain further labelling from other clinicians. Nevertheless, considering that no measurements were performed by the cardiologists, but only visual inspection of the heart cycles, the bias introduced is negligible from a clinical perspective.

It is also important to underline that the work focused the attention on antenatal PWD, but the same method could also work on the adult cardiac PWD, where the pattern is almost the same, apart from the inversion of the E/A ratio. From this perspective, it is worth to note that the adoption of the proposed method on adults would be definitely easier. In fact, on fetuses, the considered sample volume for a five-chamber apical window is tiny and the fetus is often moving in the womb, whereas the sample volume in adults is bigger and its position is more stable. Remarkably, no compensation for fetal movements can be introduced since their occurrence simply leads to loosing the correct sample volume.

Finally, an important aspect to be considered is that, by dealing with the mitral flow, which is biphasic and characterized by high inter-patient variability in the shape of the PWD, this problem is more challenging than the detection of the monophasic waves [30], so the accuracy results are significantly high and important.

The main limitation of this study is that the available dataset included only healthy fetuses, which is motivated by the specific data acquisition protocol approved by the Ethical Committee of the Cagliari University Hospital. A comprehensive study including a large population with CHD and functional problems, showing major changes in the apical five-chamber window PWD pattern, is beyond the scope of the present work. However, it is conceivable that a hierarchical classification scheme can identify the different patterns from incomplete or noisy portions of the recordings. The most significant complexity of this study extension would be the incidence of such diseases in antenatal screening, limiting the acquisition of a significant number of cases for classifier training. Nevertheless, we planned to investigate such an extension in the future.

6. Conclusion

This work presented and compared different approaches for the automatic detection of complete and measurable cardiac cycles in PWD traces, acquired based on an apical five-chamber window. The results reveal how the adoption of a supervised classifier trained with the samples representing the upper and lower envelopes of the PWD, with eight additional features based on the area under the curve of the envelopes and the brightness of the pixels enclosed in such areas, helps in achieving excellent detection results. The results are significantly better ($p < 0.0001$) than those of similar approaches based on reduced feature sets that exploit only some fiducial points on the signal, and dramatically better than the simplest template matching technique. Due to the importance of the PWD envelope shape in the definition of the features used by the classifier-based approaches, a future step of the research will be aimed at the evaluation of the impact of this processing on the final outcome of the automatic detection.

Once a well-formed fetal heartbeat is identified by the proposed algorithm, its measurement can be performed manually or automatically. As such, the algorithm can be useful for the analysis of long PWD traces or as an aid for the ultrasound operator with limited skills on fetal echocardiography for the identification of fetal heartbeats that can be analyzed for the extraction of clinical parameters useful for the quantification of the cardiac function.

We assume that the proposed algorithm could be integrated as a software plugin in ultrasound machines. Because of the characteristics of the processing, the algorithm can work on-line. However, substantial changes would affect the extraction of the PWD envelope in this case since the raw signal is available in another form inside the machine. The automatic recognition algorithm could be implemented on the same hardware architecture in charge to perform the usual ultrasound processing or could also run on a dedicated small hardware module.

Future steps of this research are also pursuing the automatic extraction of the clinical measurements from the well-formed heartbeats in the PWD trace. To this aim, we are currently developing a dataset of fetal cardiac PWD traces manually annotated by pediatric cardiologists with experience in the antenatal assessment of the cardiac function.

Conflict of Interest

The authors declare no competing interests.

Acknowledgments

The authors wish to thank the team headed by Dr. Roberto Tumbarello, Division of Paediatric Cardiology, AOI Hospital (Cagliari, Italy), for their support. Eleonora Sulas is grateful to Sardegna Regional Government for supporting her PhD scholarship (P.O.R. F.S.E., European Social Fund 2014–2020).

Supplementary material


References


