Med Biol Eng Comput (2008) 46:1227–1237
DOI 10.1007/s11517-008-0407-y

ORIGINAL ARTICLE

Fetal ultrasound image segmentation system and its use in fetal weight estimation

Jinhua Yu · Yuanyuan Wang · Ping Chen

Received: 18 March 2008 / Accepted: 17 September 2008 / Published online: 11 October 2008
© International Federation for Medical and Biological Engineering 2008

Abstract A semi-automated fetal ultrasound image segmentation system is developed to improve the estimation of fetal weight (EFW). Four standardized fetal parameters are measured by the proposed segmentation system: biparietal diameter, head circumference, abdominal circumference and femur length. Computerized measurements of 215 fetuses are compared with manual measurements in term of fitness analysis and difference analysis. Among 215 cases, computerized measurements of 103 fetuses within 3 days of delivery are utilized in the fetal weight estimation. The EFW based on computerized measurements and manual measurements are compared by using regression analysis, artificial neural network and support vector regression. By using different estimation methods, the computerized measurements decrease the EFW errors about 40–70 g. The lowest mean absolute percentage error of EFW decrease from 6.71% for manual measurements to 4.66% for computerized measurements. The proposed fetal ultrasound image segmentation system can provide more accurate EFW in antepartum examination.

Keywords Antepartum ultrasonogram · Fetal weight estimation · Ultrasound image segmentation

J. Yu · Y. Wang (✉)
Department of Electronic Engineering,
Fudan University, 200433 Shanghai, China
e-mail: yywang@fudan.edu.cn

P. Chen
Ultrasound Department, The
First Maternity and Infant Health Hospital,
200433 Shanghai, China

1 Introduction

Fetal weight is an important parameter to predict birth weight and neonatal outcome. Accurate estimation of fetal weight (EFW) is very beneficial in labor management and decision making. In clinical application, EFW is obtained by regression formulas with several ultrasonic measurements of fetal parts serving as independent variables. Widely used fetal parameters in ultrasound examination include biparietal diameter (BPD), head circumference (HC), abdominal circumference (AC) and femur length (FL) [9]. In modern ultrasound scanners, fetal measurements are made by manually extracting diameters or contours from ultrasound images. Because of the characteristic of ultrasonic imaging process, ultrasound images suffer from a very low signal-to-noise ratio (SNR) that makes its observation and interpretation difficult. Manual measurements of fetal parts are usually inaccurate and inconsistent. Several studies [6, 9, 19] have suggested that the large random errors in manual measurements are a major cause of inaccurate EFW, making it necessary to minimize the measurement error.

Several recent literatures have shed some light on the fetal weight estimation based on automated or semi-automated measurements of fetal parameters by utilizing different image segmentation techniques. For instance, an iterative randomized Hough transform was proposed to measure the BPD and HC [14]. A method based on maximum likelihood estimation of a parametric deformable model was developed to measure the FL and HC [11]. A method that integrates four types of image segmentation techniques was utilized to measure the AC [25]. All these literatures suggested that automated or semi-automated ultrasound image segmentation can produce more consistent and accurate fetal parameter measurements than

manual measurements. However, no systematic attempt has been made to evaluate the efficacy of these image segmentation techniques in fetal weight estimation. The purposes of this study are: (1) to develop a semi-automated image segmentation system to measure BPD, HC, AC and FL. (2) To systematically compare the computerized measurement with the manual measurement. (3) To assess the value of ultrasound image segmentation system in fetal weight estimation.

2 Methods

2.1 Background of methods

The semi-automated image segmentation system involves measurements of three fetal body parts (head, abdomen and leg) with each having its own physiological shape and acoustic characteristic. It is well known that ultrasound images are corrupted by heavy artifacts such as speckle, attenuation, shadow, and signal missing. The difference of segmentation problems and the low SNR of ultrasound images make the automated measurement of fetal parameters difficult. To improve results of fetal ultrasound image measurement, several image processing techniques have been utilized in our image segmentation system. This section gives a brief description about the background of these techniques.

2.1.1 Speckle reducing anisotropic diffusion

Speckle noise is known as multiplicative noise with a complex statistical distribution. As the most pronounced artifact, speckle noise not only obscures diagnostically important details in images but also complicate subsequent image segmentation. Anisotropic diffusion (AD) [15, 16] is commonly used to remove speckles in ultrasound images. An improved anisotropic diffusion filter referred to as speckle reducing anisotropic diffusion (SRAD) [26] is used in this study to improve the accuracy of image segmentation and measurement. SRAD is derived by integrating spatially adaptive filters into nonlinear diffusion. The diffusion process is determined by a diffusivity function:

$$c(q) = \frac{1}{1 + \left[q^2(x,y;t) - q_0^2(t)\right] / \left[q_0^2(t)(1+q_0^2(t))\right]} \tag{1}$$

where $q(x, y; t)$ is the instantaneous coefficient of variation (ICOV) used for edge detection and $q_0(t)$ is the speckle scale function used as the diffusion threshold. The diffusivity function $c(q)$ encourages homogeneous-region smoothing and inhibits smoothing across the edges. The prominent speckle noise reduction ability of SRAD mainly contributes to the effectiveness of its edge detection operator ICOV or $q(x, y; t)$ in Eq. 1:

$$q(x,y;t) = \sqrt{\frac{\left|\left(\frac{1}{2}\right)\|\nabla I\|^2 - \left(\frac{1}{16}\right)\left(\nabla^2 I\right)^2\right|}{\left(I + \left(\frac{1}{4}\right)\nabla^2 I\right)^2}} \tag{2}$$

where ∇ ,∇^2 , $\|\|$, and $|$| are the gradient, Laplacian, gradient magnitude, and absolute value, respectively.

Recently, the state-of-the-art methods such as Rayleigh-maximum-likelihood filtering [1] have been proposed to reduce speckle for ultrasound images. Compared with other methods, the most desirable feature of SRAD for the proposed system is that an effective edge detection strategy (ICOV) has been integrated into the diffusion process, which means we do not need extra operation to extract edges from denoising images for the following segmentation procedures.

2.1.2 Fuzzy C-means algorithm

The fuzzy C-means (FCM) [2] and one of its improved versions have been used in this study to simplify detected edges and extract parts of interests from background. Let $X = \{x_1, x_2, \ldots, x_n\}$ be a set of n data points, and c be the total number of clusters or classes. The standard FCM [2] is formulated by minimizing an objective function J_{FCM} with respect to the membership values μ_{ij} and cluster prototypes m_j, $j = 1, 2, \ldots, c$

$$J_{\mathrm{FCM}} = \sum_{j=1}^{c}\sum_{i=1}^{n} \mu_{ij}^{b}\left\|x_i - m_j\right\|^2 \tag{3}$$

subject to

$$U = \left\{ \mu_{ij} \in [0,1] \,\middle|\, \sum_{j=1}^{c} \mu_{ij} = 1 \;\; \forall i \;\text{ and }\; 0 < \sum_{i=1}^{n} \mu_{ij} < N \;\; \forall j \right\}. \tag{4}$$

The objective function (Eq. 3) is minimized when high membership values are assigned to pixels whose intensities are close to the centroid of its class and low membership values are assigned when the pixel value is distant from the centroid. FCM works well on images with low levels of noise, but it does not incorporate spatial information, making it sensitive to noise and other imaging artifacts. To improve the robustness of FCM, many algorithms have been introduced, with majority of them trying to impose different spatial constraints into the objective function of FCM [7].

2.1.3 Randomized Hough transform

Hough transform (HT) [13, 27] is a method to extract global curved segments from an image. The main

drawbacks of HT are its high computational complexity, memory requirements, and limited accuracy. A new class of HTs called Randomized Hough Transform (RHT) [24] can overcome several disadvantages of the original HT algorithm. RHT depends on random sampling and many-to-one mapping from the image space to the parameter space in order to achieve effective object detection. In this paper, RHT is utilized to detect ellipse (HC and AC) and line segment (FL) from fetal ultrasound images. The steps in RHT for the general case of detecting a p-dimensional curve are outlined below.

1. Scan the binary image and record coordinates of each point $\boldsymbol{d}_i = (x_i, y_i)$ into matrix D.
2. Set the valid trail counter $k = 0$. Empty the accumulator space $H(z_1,\ldots, z_p)$.
3. Select p points $\boldsymbol{d}_i = (x_i, y_i)$ $(i = 1,\ldots, p)$ randomly from the matrix D.
4. If points do not satisfy the predefined sampling condition, go to (3); otherwise continue.
5. Solve a set of equations $f(\boldsymbol{a}, \boldsymbol{d}_i) = 0$ to obtain parameters $\alpha = [\alpha_1,\ldots, \alpha_p]^{\mathrm{T}}$.
6. Convert α into its natural parameters $[z_1,\ldots, z_p]$ and check its validity. If parameters are not reasonable, go to (3); otherwise the corresponding count in H is increased by one and $k = k + 1$.
7. If the number of the valid trail k is less than a predefined constant K, go to (3); otherwise continue.
8. The $[z_1,\ldots, z_p]$ corresponding to the count maximum in $H(z_1,\ldots, z_p)$ is chosen as parameters of the target curve. The implementation details for each particular fetal parameter will be discussed in the next Section.

2.1.4 *Gradient vector flow Snakes*

The snake algorithm is employed in this study to find the best fit between extracted contour and the actual shape of the fetal part. In the snake algorithm, a contour is treated as an elastic band whose deformation is guided by external constraint forces computed from the image data and influenced by image forces originating from the contour itself. The gradient vector flow (GVF) [23] is an external force field that can deal with problems associated with the initialization and poor convergence to boundary concavities of the original snake algorithm [12]. GVF is defined as the vector field $\mathbf{v}(x, y)$ that minimizes the energy functional

$$E = \int\int \mu|\nabla\mathbf{v}|^2+|\nabla f|^2|\mathbf{v} - \nabla f|^2 \mathrm{d}x\mathrm{d}y \tag{5}$$

where μ is a parameter governing the tradeoff between the first term and the second term in the integrand, ∇ is the gradient operator applied to each component of $\mathbf{v}$ separately, and f represents the edge map. The snake algorithm using the GVF field as its external force is referred to as a GVF snake, which is defined as the parametric curve $\mathbf{x}(s) = [x(s), y(s)]$ satisfying the dynamic equation

$$\mathbf{x}_t(s, t) = \alpha\mathbf{x}'' - \beta\mathbf{x}'''' + \mathbf{v}(\mathbf{x}). \tag{6}$$

Here, $\mathbf{x}$ is treated as a function of time t as well as s. α and β are weighting parameters that control the snake's tension and rigidity, respectively, and $\mathbf{x}''$ denotes the second derivative of $\mathbf{x}$.

2.2 Segmentation and measurement of BPD, HC and AC

In antepartum ultrasound examination, BPD is measured as the distance from the outer edge of the fetal skull bone nearest to the transducer to the outer edge of the cranium farthest from the transducer. HC and AC are measured by fitting an ellipse to the outer contour of head and abdomen, respectively.

The segmentation and measurement of HC and AC require fitting an ellipse to the outer contour of head and abdomen. However, there are several differences in the segmentation algorithms for HC and AC. First, the abdominal contour formed by skin has a weaker contrast in the image than the head contour formed by skull. Second, compared with fetal head, the abdominal contour lacks regularity due to deformation under the pressure from other maternal organs and fetus itself. This suggests that, to achieve more accurate abdominal contour extraction, an additional procedure should be employed to adapt the extracted contour to real edges of the abdomen. Figure 1 summarizes the implementation procedure of segmentation algorithms for head and abdomen measurements.

Because of the complexity of the segmentation algorithm, performing the algorithm over the whole image is not feasible. A well-defined region-of-interests (ROI) not only reduces the computational time but also helps reject interferences from irrelevant organs or tissues in the same scanning plane. In the first step, a rectangle, enclosing the contour of object, is given by the operator (as shown in Figs. 2a, 3a), then a ROI is defined in the form of an

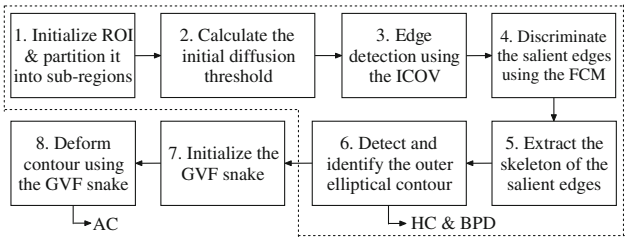

Fig. 1 Flowchart of segmentation algorithms for head and abdomen measurements

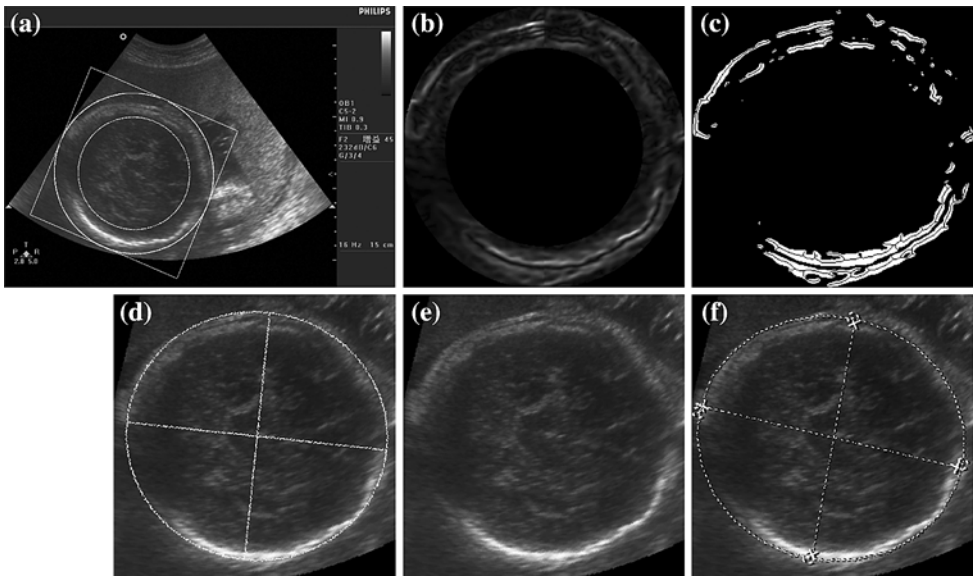

Fig. 2 Segmentation and measurement of fetal head. **a** ROI definition, **b** Edge map detected by the ICOV, **c** Salient edges map with its skeleton, **d** Head contour extracted by program, **e** Original image, **f** Manually extracted head contour

elliptical ring in accordance with the manually defined rectangle. The outer ellipse of the ring is formed such that it is inscribed within the rectangle. The inner ellipse is obtained by shrinking the outer ellipse by 30% along its major and minor axes. The shrinking percentage is determined by 'trial-and-error'. A 30% is a choice which satisfies the requirements of both algorithm simplification and generalization ability.

To be adaptive to edge strength variation, the ROI is equally partitioned into eight sub-regions according to the central angle. Then, in the second step, the initial diffusion threshold q_0 for each sub-region is formulated as

$$q_0 = \frac{\sqrt{\text{var}[\text{SR}]}}{\overline{\text{SR}}} \tag{7}$$

where var[SR] and $\overline{\text{SR}}$ are the intensity variance and mean of each sub-region SR. $q_0(t)$ is approximated by $q_0(t) = q_0 \exp[-t/6]$, here t represents the discrete diffusion time step. In the third step, speckle reduction and edge detection are applied using SRAD filter. The edge maps are then obtained by scaling the magnitude of ICOV from 0 to 255 (as shown in Figs. 2b, 3b).

In the fourth step, FCM clustering algorithm is utilized to discriminate salient edges from weak edges and background. This is realized by classifying the edge map into three categories (salient edges, weak edges and background). Then, the category with the highest value is defined as salient edges. The number of cluster ($c = 3$) is determined by the distribution decomposition of image histogram. We assume that the histogram of images is a mixture of c Gaussian distribution. Then, expectation maximum (EM) algorithm is used to estimate parameters for the mixture distribution. When c equals to 3, the mixture distribution fits the histogram best. Since setting c to 3 can produce consistently satisfactory results and the distribution decomposition based on EM algorithm will cost memory and processing time, we set c to 3 as a prior knowledge. In the next step, salient edges are thinned to serve as the input of the next processing. As shown in Figs. 2c and 3c, white pixels represent the cluster of salient edges, and black curves are the skeleton of the clustering result.

In the sixth step, RHT is used to detect and locate the outer contour of HC and AC. The parametric function of the ellipse used in this study is formulated as

$$a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + 1 = 0. \tag{8}$$

The RHT is iteratively applied to update the ROI and identify the outer ellipse of the edge map. At the end of each RHT iterations, the skeleton image is updated by discarding pixels within an ellipse which is obtained by shrinking the detected ellipse by 5% along its major and minor axes. The implementation details can be found in our previous work [25]. The BPD is measured as twice of minor axes, and HC is measured as the perimeter of the detected ellipse by iterative RHT. The detected ellipses for HC and AC are shown in Figs. 2d and 3d, respectively. Hereto, the segmentation and measurement of BPD and HC are obtained.

To obtain more accurate contour extraction for AC, in the seventh and eighth steps, a GVF snake is employed to adapt the detected ellipse to real edges of abdominal contour. The final segmentation result by the GVF snake is shown in Fig. 3e. For comparison convenience, the original image in ROI and the manual contour outlining are shown in Figs. 2e, f and 3f, g, respectively.

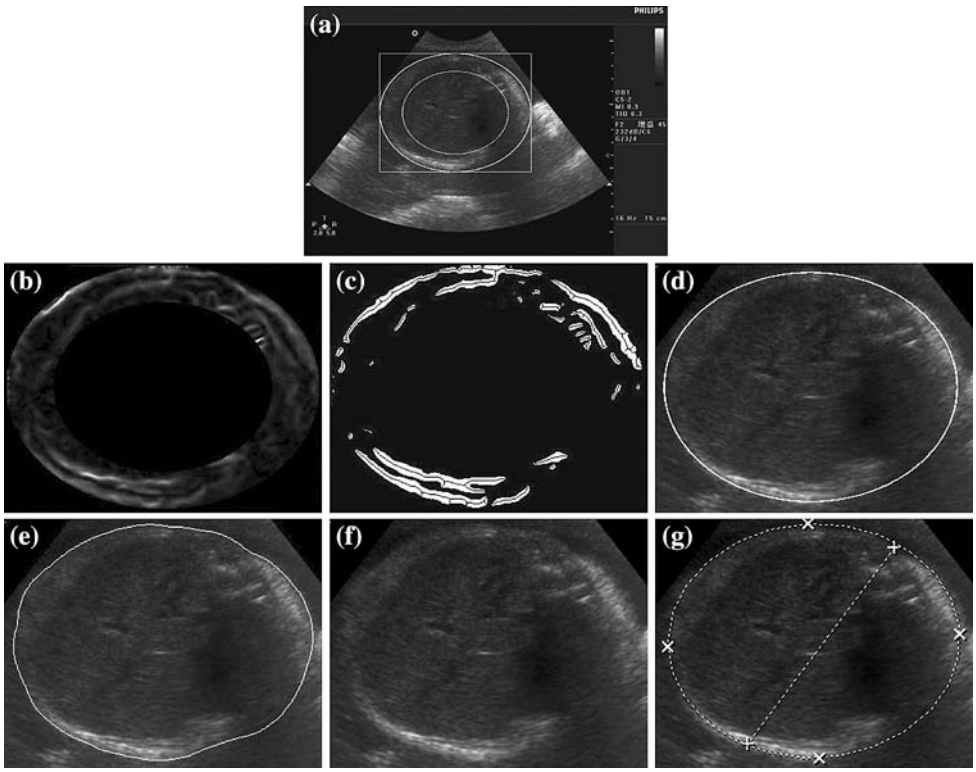

Fig. 3 Segmentation and measurement of fetal abdomen. **a** ROI definition, **b** Edge map detected by ICOV, **c** Salient edges map with its skeleton, **d** Initial contour obtained by RHT, **e** Abdominal contour extracted by program, **f** Original image, **g** Manually extracted abdominal contour

2.3 Segmentation and measurement of FL

Femur length is measured as the distance between two endpoints of the femoral diaphysis. It can be observed that a fetal femur formed as a straight line segment for younger fetuses and a slightly curved arc for older fetuses [17]. The femur curvature is ignored in manual measurements. With the femur curvature increases, the accuracy of manual measurements could decrease. We propose a segmentation algorithm that takes the curvature of femoral diaphysis into consideration. Figure 4 summarizes the implementation procedure of the proposed algorithm.

In the first step, a rectangle covering the femoral diaphysis is defined by the operator (as shown in Fig. 5a). In the second step, an improved FCM is used to extract the bone from the background. As mentioned in the previous section, the original FCM algorithm is sensitive to noise and outliers. The spatial constraint, which takes into account the spatial relationships among pixels in classification, can improve the performance of the FCM under noisy conditions. An improved version of FCM denoted as FCM_S [7] was proposed by modifying the objective function in 3 as follows:

$$J_{\mathrm{FCM\_S}} = \sum_{j=1}^{c}\sum_{i=1}^{n} \mu_{ij}^{b}\left\|x_i - m_j\right\|^2 + \alpha \sum_{j=1}^{c}\sum_{i=1}^{n} \mu_{ij}^{b}\left\|\bar{x}_i - m_j\right\|^2 \tag{9}$$

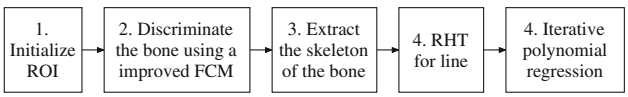

Fig. 4 Flowchart of the segmentation algorithm for femur measurement

where $\bar{x}_i$ is the median of neighboring pixels lying within a window around x_i, α controls the effect of the spatial constraints. FCM_S is used to classify each pixel into one of three categories: bright object, gray object and background, then bright objects are extracted. According to the distribution decomposition based on EM algorithm, c is also set to 3 as a prior knowledge for all images that we tested. The clustering result is thinned in the third step to serve as the input of the next processing (as shown in Fig. 5b).

In the fourth step, the RHT for line is used to detect the best line candidate in the image. The parametric function of the line is formulated as

$$y - a_1 x - a_2 = 0 \tag{10}$$

where a_1 and a_2 are the slope and intercept of the detected line, respectively. Then the detected line is used to define a new ROI which only includes the skeleton pixels falling near the line. Figure 5c illustrates the detected line by the RHT and the new ROI definition. In this image, the thicker white line is the detected line; the region enclosed by two

thin lines is the new ROI. The two thin lines are parallel to the detected line and have equal distance to the detected line. The distance between the two thin lines equals to one-fourth length of the detected line. The parameter 1/4 is chosen based on the observation that the bending degree of femur is proportional to its length and this degree will not exceed 1/4 of femur length except for some severe abnormality of fetal thigh. Through the new ROI definition, most pixels corresponding to other tissues are excluded.

In the fifth step, a polynomial regression is used to fit a curve to skeleton pixels in the new ROI. A conic function is used:

$$y - b_1x^2 - b_2x - b_3 = 0 \tag{11}$$

where b_1, b_2 and b_3 decides the curvature, slope and intercept of the curve, respectively. The polynomial regression is iteratively applied to adaptively update the ROI. In this way, a curve that gradually fit to the femur shape can be obtained after few times of iteration. Figure 5d illustrates the detected curve obtained by the first iteration of the polynomial regression and the updated ROI. The distance between the two thin curves in Fig. 5d equals to one-fifth length of the detected line by the RHT. The final detected curve is shown in Fig. 5e. The end points of the curve are obtained by comparing the detected curve with the clustering pixels in the final ROI. For comparisons convenience, the original image in the ROI and the manual contour outlining are shown in Fig. 5f, g.

2.4 Fetal weight estimation based on ultrasound measurement

As a widely accepted technique, the regression analysis has been utilized routinely for the ultrasonically based fetal weight estimation. Varieties of regression formulas, incorporating different ultrasonic measurements, have been studied extensively. It was first pointed out that AC is best correlated with the fetal growth and is the most effective parameter for the birth weight estimation [4]. They proposed a formula based on AC measurement:

$$\ln(\mathrm{EFW}/1{,}000) = -4.564 + (0.282 \times \mathrm{AC}) - \left(0.00331 \times \mathrm{AC}^2\right). \tag{12}$$

It was also indicated that the growth rate of femur length illustrates a decreasing growth rate as the gestational age increases [22]. A regression formula based on FL measurement has the form as:

$$\ln(\mathrm{EFW}) = 4.6914 + \left(0.151 \times \mathrm{FL}^2\right) - \left(0.0119 \times \mathrm{FL}^3\right). \tag{13}$$

Although regression formulas only based on fetal head measurements are rarely reported, many multi-parameters regression formulas have been reported by incorporating BPD and HC. Among these formulas, one of the most accurate is the one integrated with all standardized parameters [10]:

$$\begin{aligned}\mathrm{Log}_{10}\mathrm{EFW} = {} & 1.3596 - 0.00386(\mathrm{AC})(\mathrm{FL}) + 0.0064(\mathrm{HC}) \\ & + 0.00061(\mathrm{BPD})(\mathrm{AC}) + 0.0424(\mathrm{AC}) \\ & + 0.174(\mathrm{FL}).\end{aligned} \tag{14}$$

In recent years, several alternative methods have also been proposed to estimate the fetal weight. For instance, a general artificial neural network (ANN) model was developed [8]. The ANN is a computerized analog of the biological neural structures of the central nervous system. For given input data, a flow of activation is forwarded from the input layer to the output layer via hidden layers. Then weights and biases of the network are adjusted according to errors in the output. When errors between network outputs and expected outputs are small enough which means the relationship between inputs and outputs has been memorized, the learning process is ceased. Thus,

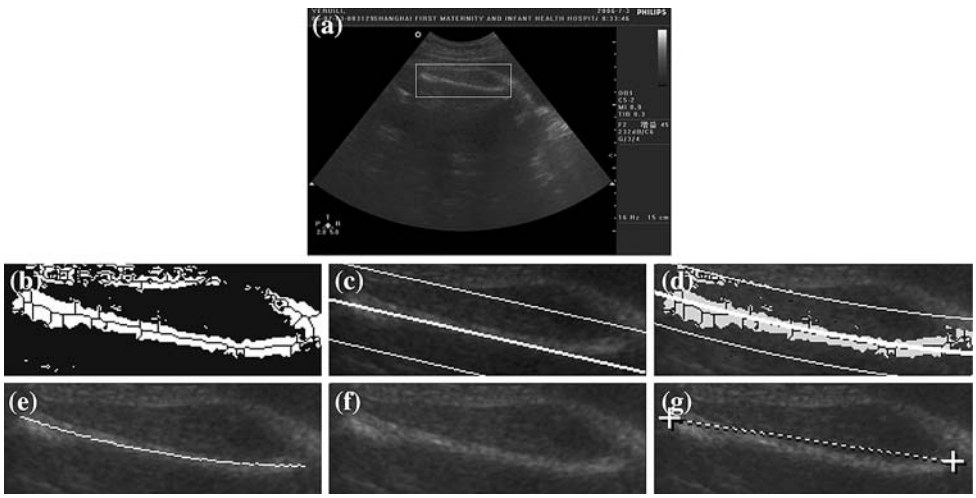

Fig. 5 Segmentation and measurement of femur length. **a** ROI definition, **b** Bone extraction by FCM_S, **c** Line detection by RHT and ROI update, **d** First iteration of polynomial regression and ROI update, **e** FL measured by program, **f** Original image, **g** Manually measured FL

Table 1 The patient characteristics of two datasets

Characteristics	Data set I	Data set II
Patients' age	29.39 ± 4.08 (21 – 45)	29.62 ± 4.04 (21–45)
Gestational age, GA (weeks)	35.43 ± 5.87 (16 – 41)	39.01 ± 1.22 (33–41)
Actual birth weight, ABW (g)	–	3,299.17 ± 405.21 (1,850–4,870)
Cases percentage of ABW (%)		
<2,500 g (%)	–	1.94
2,500–4,000 g (%)	–	95.15
>4,000 g (%)	–	2.91
Manual measurements of fetal parameters (cm)		
BPD (cm)	8.76 ± 1.37 (3.57–10.50)	9.49 ± 0.40 (7.92–10.30)
HC (cm)	30.82 ± 4.60 (12.20–36.80)	33.23 ± 1.20 (29.00–36.80)
AC (cm)	30.62 ± 5.73 (10.40–38.70)	33.86 ± 2.03 (25.80–38.50)
FL (cm)	6.45 ± 1.21 (1.78–7.89)	7.11 ± 0.29 (6.04–7.89)

meaningful answers can be supplied when other data have been put into the network.

The support vector machine (SVM) [20] also has been used in the fetal weight estimation [18]. The SVM is based on the statistic learning theory and Structural Risk Minimization (SRM) principle which pursues minimizing the upper bound on the expected risk. The SRM principle equips the SVM with a good ability of the generalization and classification precision. When the SVM is utilized into addressing the regression problem, it is called as the support vector regression (SVR).

3 Materials and results

3.1 Materials

The subjects consisted of 215 women undergoing ultrasound examination in the Ultrasound Department of Shanghai First Maternity and Infant Health Hospital, China. Among the whole subjects, 103 mothers delivered within 3 days of ultrasound examination. The actual birth weight (ABW) of these 103 fetuses was used as ground truth for evaluating the validity of semi-automated image segmentation system in fetal weight estimation. All mothers were given informed consent to participate in this study, and the study was approved by the Ethics Committee of our institution. The following criteria were used to select the subjects: (1) normal singleton pregnancy without maternal complications; and (2) fetuses without major congenital or chromosomal anomalies. Ultrasound measurements were carried out using the commercial 2-D ultrasound scanner EnVisor 2540A (PHILIPS, ShenYang, China) with a 3.5 MHz trans-abdominal probe by experienced technicians or physicians. All images were stored in Microsoft Bitmap (BMP) format of a size 800 × 564 with 24 bit per pixel.

All the images of total 215 patients were processed and compared with the manual measurements, and computerized measurements of 103 fetuses were used to estimate fetal weight. We refer to the data set including all 215 patients as data set I and that including the 103 patients as data set II. The patient characteristics of the two data sets are shown in Table 1.

3.2 Comparisons between computerized measurements and manual measurements

In this part, the fitness between computerized measurements and manual measurements was carried out by using the linear regression and correlation analysis. The regression analysis is carried out with the SPSS statistical software package (version 13.0 for Windows). Assuming the computerized measurement and manual measurement of parameter X can be represented by X_C and X_M, respectively (X can be BPD, HC, AC, and FL), then the linear regression equations between X_M and X_C can be written as:

$$\begin{aligned} \mathrm{BPD\_C} &= 1.006(\mathrm{BPD\_M}) + 0.171, \\ \mathrm{BPD\_M} &= 0.975(\mathrm{BPD\_C}) - 0.006 \end{aligned} \tag{15}$$

$$\begin{aligned} \mathrm{HC\_C} &= 0.971(\mathrm{HC\_M}) + 0.527, \\ \mathrm{HC\_M} &= 1.011(\mathrm{HC\_C}) + 0.053 \end{aligned} \tag{16}$$

$$\begin{aligned} \mathrm{AC\_C} &= 1.122(\mathrm{AC\_M}) - 1.352, \\ \mathrm{AC\_M} &= 0.863(\mathrm{AC\_C}) + 2.139 \end{aligned} \tag{17}$$

$$\begin{aligned} \mathrm{FL\_C} &= 1.031(\mathrm{FL\_M}) - 0.042, \\ \mathrm{FL\_M} &= 0.937(\mathrm{FL\_C}) + 0.253 \end{aligned} \tag{18}$$

Correlation between computerized measurements and manual measurements is determined by Pearson's correlation coefficient (r). Correlations and differences are considered significant when the significance probability P was less than 0.05. The correlations between manual

Table 2 Difference analysis between two measurement systems

	Mean	Deviation	Minimum	Maximum
BPD (cm)	0.25 (2.89%)	0.16 (1.84%)	0.00 (0.03%)	0.76 (10.03%)
HC (cm)	0.57 (1.85%)	0.47 (1.44%)	0.01 (0.03%)	2.37 (7.02%)
AC (cm)	2.38 (7.53%)	1.35 (3.97%)	0.00 (0.00%)	6.84 (23.65%)
FL (cm)	0.22 (3.34%)	0.18 (2.67%)	0.00 (0.00%)	1.13 (15.88%)

measurements and computerized measurements are ($r = 0.991$, $n = 215$, $P < 0.0001$) for BPD, ($r = 0.990$, $n = 215$, $P < 0.0001$) for HC, ($r = 0.984$, $n = 215$, $P < 0.0001$) for AC and ($r = 0.983$, $n = 215$, $P < 0.0001$) for FL.

Then the difference analysis was carried out by using the mean, standard deviation, minimum and maximum of absolute error and absolute percentage error. Absolute percentage error is calculated as abs (X_C − X_M)/X_M × 100%. Table 2 summarizes the difference between computerized measurements and manual measurements for BPD, HC, AC and FL. The comparison suggests that, despite of high correlations between computerized measurements and manual measurements, there exist considerable differences between two measurement systems.

3.3 Fetal weight estimation based on two measurement systems

Data set II is used to evaluate the validity of computerized measurements in the fetal weight estimation. At first, the correlation between computerized measurements and the ABW is compared with the correlation between manual measurements and the ABW. Figure 6a summarizes the comparison results. It can be seen that correlations between computerized measurements and the ABW are always higher than ones between manual measurements and the ABW. This can be a valid indirect proof to infer that the proposed segmentation system provides more consistent and accurate fetal parameters than manual measurements.

Next, we compare prediction results generated from the regression analysis by using two measurement systems. Despite better correlations between computerized measurements and the ABW have been produced, computerized measurements cannot guarantee to yield more accurate EFW by using existing regression formulas such as Eqs. 12–14. The reason is that computerized measurements have considerable systematical differences with manual measurements, and existing regression formulas are all based on manual measurements. To fairly compare two measurement systems, we use computerized measurements

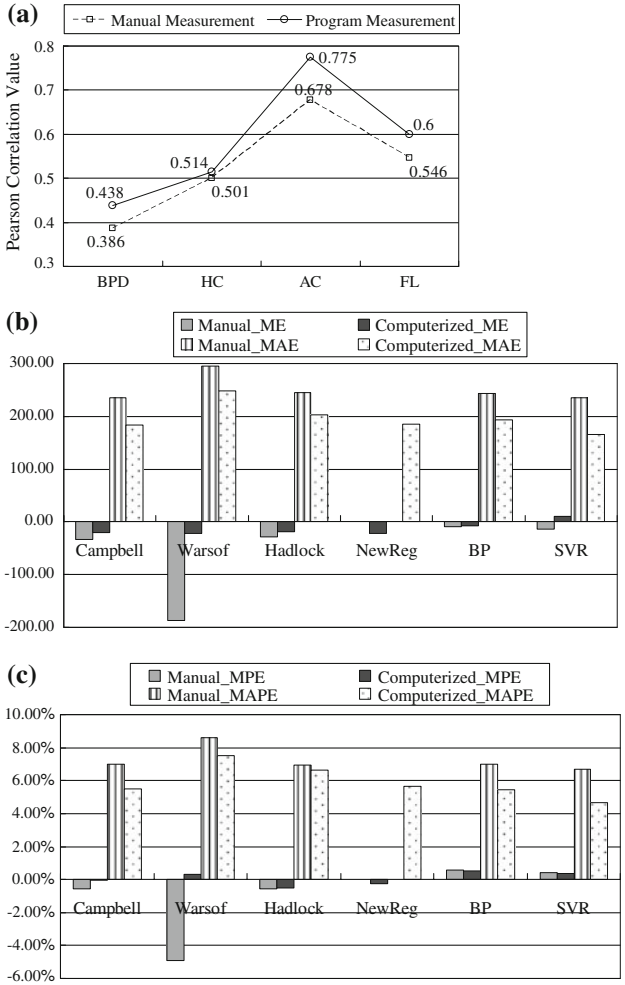

Fig. 6 Comparisons of two measurement systems in fetal weight estimation. **a** The correlation between the ABW and computerized measurements versus manual measurements, **b** ME and MAE of EFW based on two measurement systems in testing group, **c** MPE and MAPE of EFW based on two measurement systems in testing group

to adjust coefficient values of regression formulas Eqs. 12–14. A total of 103 samples are separated into three groups, and two groups are randomly chosen as formula founding set, one as testing set. The best-fit EFW formulas derived from the regression analysis based on Campbell model (Eq. 12), Warsof model (Eq. 13) and Hadlock model (Eq. 14) have forms as:

$$\ln(\mathrm{EFW}/1{,}000) = -3.479 + (0.211 \times \mathrm{AC}) - \left(0.00228 \times \mathrm{AC}^2\right) \quad (19)$$

$$\ln(\mathrm{EFW}) = 4.850 + \left(0.155 \times \mathrm{FL}^2\right) - \left(0.0129 \times \mathrm{FL}^3\right) \quad (20)$$

$$\mathrm{Log}_{10}\mathrm{EFW} = -0.2016 - 0.01066(\mathrm{AC})(\mathrm{FL}) + 0.0015(\mathrm{HC}) + 0.00044(\mathrm{BPD})(\mathrm{AC}) + 0.0881(\mathrm{AC}) + 0.43(\mathrm{FL}). \quad (21)$$

Moreover, because of the considerable differences between two measurement systems, existing regression

models (the choice of parameters and the combination between parameters) may not suitable for the computerized measurements. We also develop a new regression formula for computerized measurements by using stepwise regression analysis, which has the form:

$$\begin{aligned}\mathrm{Log}_{10}\mathrm{EFW} = {} & 0.00169(\mathrm{HC}) + 0.08464(\mathrm{AC}) \\ & + 0.40103(\mathrm{FL}) + 0.000022(\mathrm{BPD})^2(\mathrm{AC}) \\ & - 0.00988(\mathrm{AC})(\mathrm{FL}).\end{aligned} \tag{22}$$

The ANN and SVR are also used to compare the validity of computerized measurements and manual measurements in the fetal weight estimation. Following the notion of the reference [8], the architecture of the BPNN used in our work is composed of three layers: one input layer with four inputs (BPD, HC, AC and FL), one hidden layer with eight neurons, and one output layer with the prediction result. The transfer function is the hyperbolic tangent sigmoid function. For the SVR model, the regularization parameter is set to 10^5, the insensitivity zone is 0.001, the kernel function is chosen as a radial basis function with the parameter of 10. A total of 103 samples are also used in a cross-validation with two-third samples serving as the training group, one-third as the testing group. Figure 6b illustrates comparisons of the mean error (ME) and the mean absolute error (MAE) based on two measurement systems in the testing group. Figure 6c illustrates comparisons of the mean percent error (MPE) and the mean absolute percent error (MAPE) based on two measurement systems in the testing group.

As shown in Fig. 6, each method using computerized measurements produces a lower estimation error than using manual measurements. The SVR model based on computerized measurements has the lowest mean absolute error (165.68 g, $P < 0.0001$) and mean absolute percentage error (4.66%, $P < 0.0001$). When the SVR model is based on data obtained by manual measurements, the mean absolute error and mean absolute percentage error are (235.32 g, $P < 0.0001$) and (6.71%, $P < 0.0001$), respectively. This means that the accuracy of the EFW has been improved by computerized measurements about 70 g (2.1%) for the SVR model. If estimation errors produced by five methods have been averaged, the improvement of the EFW accuracy based on computerized measurements reaches to 55 g (1.52%).

4 Discussions

In this study, a semi-automated fetal ultrasound image segmentation system has been developed, and the potential of this system in the fetal weight estimation has been evaluated. Findings in this study can be summarized as follows: (1) Computerized measurements are highly correlated with manual measurements. However, there exist considerable system differences and random differences between two measurement systems. (2) Computerized measurements can produce more accurate EFW than manual measurements. Computerized measurements produce a higher correlation with the ABW and decrease the EFW errors around 40–70 g. (3) Methods such as the ANN and SVR can provide a more accurate EFW compared with the regression analysis. One current limitation of the proposed image segmentation system is that the measurement procedure is not totally automated and the operator should define a ROI that encloses the target area.

Many efforts have been made to improve the EFW accuracy. These attempts can be divided into two main groups: pursuing powerful estimation models which can characterize the nonlinear relationship between fetal parameters and EFW more effectively [8, 18]; attempting to obtain more accurate fetal measurements either from the improvement of imaging instruments [3, 5] or from image processing techniques such as image denoising, image enhancement, or image segmentation [11, 14, 21, 25].

Although the regression analysis has been taken as a daily routine for ultrasonically based fetal weight estimation, there is a prominent disadvantage of regression formulas in fetal weight estimation. Most of the fetal weight estimation formulas were derived from a certain population. When a formula was applied to a population different from the one which the formula was derived from, the estimation accuracy might be degraded because of the differences on nutritional condition or socioeconomic circumstances. This might be the main reason of why a fix formula cannot produce consistent estimation accuracy in different studies. Compared with the regression analysis, some alternative estimation models such as ANN and SVR have better capabilities in capturing nonlinear relationship between input variables and output outcomes. In study of [8], an ANN model was developed to estimate the fetal weight. By using this model, mean absolute percentage error (MAPE) was reduced around 1.6% compared with four regression formulas. Reference [18] reported 2.7% reduction in MAPE compared with radial basis function (RBF) networks by using SVR to estimate fetal weight. However, no compassions with regression formulas have been made in this literature.

Another important factor that disturbs the EFW accuracy is inaccurate ultrasonic measurements of fetal parameters. There are a lot of reasons to generate these inaccurate measurement data, such as the incorrect probe position and inappropriate scanning plane due to the deficient experience of sonographer, incorrect head measurement when fetal heads were deeply engaged in the

pelvis, inaccurate abdominal circumference measurement when fetal bodies were compressed by placenta or limbs of fetus itself. So, some studies try to improve the EFW by pursuing more accurate ultrasonic fetal measurement. In reference [5], an ANN model was also utilized to characterize the relationship between the fetal measurements and fetal weight, and then a multi-normal probability model which is based on the trained ANN parameters is designed to correct human errors in ultrasonic measurements. By using this software, measurement errors could be detected and corrected on real-time, thus the MAPE could be reduced around 2.1% based on eight regression formulas. For obtaining more accurate fetal measurements, three-dimensional (3D) ultrasound drew adequate attentions. Early in 1984, reference [3] indicated that 3D ultrasound measurements had the potential to improve the EFW remarkably. However, obtaining 3D measurement of fetal parts is kind of challenging itself due to the noisy nature of ultrasonic data. In reference [21], a method that integrated diffusion filter, light absorbing function, and texture-based rendering has been developed to enhance the visualization of 3D fetal ultrasound images. This method might be an interesting technique to improve the accuracy of 3D measurement, and its potential in fetal weight estimation should be studied.

Compared with conventional manual measurements, the MAPE of EFW obtained from computerized measurements decreases from 0.34% of Hadlock formula to 2.1% of SVR model. If the new formula (Eq. 22) is used, the MAPE decreases 1.33% compared with Hadlock formula based on manual measurements. Among four fetal parameters, computerized measurements of AC and FL significantly increase the accuracy of formulas. Although the accuracy improvements of EFW obtained from regression formulas equipped with computerized measurements do not always excel previous literatures [5, 8], the estimation results generated from BP and SVR are quite competitive. Compared with the three regression formulas, the reduction of MAPE reaches to 2.87% if the SVR equipped with computerized measurements is used. In conclusion, the fetal ultrasound image segmentation system can provide more reliable and accurate fetal measurements than manual measurements. Computerized measurements effectively decrease the random error in ultrasound examinations. With the decreasing of the measurement error, the accuracy of the EFW can be improved considerably.

Acknowledgments This work was supported by the National Basic Research Program of China (No. 2006CB705707), Natural Science Foundation of China (No. 30570488), Shanghai Leading Academic Discipline Project (No. B112) and Postgraduate Innovation Fund of Fudan University (No. EYH1220001).

References

1. Aysal TC, Barner KE (2007) Rayleigh-maximum-likelihood filtering for speckle reduction of ultrasound images. IEEE Trans Med Imaging 26(5):712–727. doi:10.1109/TMI.2007.895484
2. Bezdek JC (1980) A convergence theorem for the fuzzy ISODATA clustering algorithm. IEEE Trans Pattern Anal Mach Intell 1(2):1–8
3. Brinkley JF, McCallum WD, Muramatsu SK, Liu DY (1984) Fetal weight estimation from lengths and volumes found by three-dimensional ultrasound measurements. J Ultrasound Med 3(4):163–168
4. Campbell S, Wilkin D (1975) Ultrasonic measurement of the fetal abdomen circumference in the estimation of fetal weight. Br J Obstet Gynaecol 82:689–697
5. Cevenini G, Severi FM, Bocchi C, Petraglia F, Barbini P (2008) An informative probability model enhancing real time echobiometry to improve fetal weight estimation accuracy. Med Biol Eng Comput 46:109–120. doi:10.1007/s11517-007-0299-2
6. Chang TC, Robson SC, Spencer JA, Gallivan S (1993) Ultrasonic fetal weight estimation: analysis of inter- and intra-observer variability. J Clin Ultrasound 21:515–519. doi:10.1002/jcu.1870210808
7. Chen SC, Zhang DQ (2004) Robust image segmentation using FCM with spatial constraints based on new kernel-induced distance measure. IEEE Trans Syst Man Cybern 34(4):1907–1916. doi:10.1109/TSMCB.2004.831165
8. Chuang L, Hwang JY, Chang CH, Yu CH, Chang FM (2002) Ultrasound estimation of fetal weight with the use of computerized artificial neural network model. Ultrasound Med Biol 28(8):991–996. doi:10.1016/S0301-5629(02)00554-9
9. Dudley NJ (2004) A systematic review of the ultrasound estimation of fetal weight. Ultrasound Obstet Gynecol 25(1):80–89. doi:10.1002/uog.1751
10. Hadlock FP, Harrist RB, Sharman RS, Deter RL, Park SK (1985) Estimation of fetal weight with the use of head, body, and femur measurements—a prospective study. Am J Obstet Gynecol 151:333–337
11. Jardim SMGVB, Figueiredo MAT (2005) Segmentation of fetal ultrasound images. Ultrasound Med Biol 31(2):243–250. doi:10.1016/j.ultrasmedbio.2004.11.003
12. Kass M, Witkin A, Terzopoulos D (1988) Snakes: active contour models. Int J Comput Vis 1(4):321–331. doi:10.1007/BF00133570
13. Kiryati N, Eldar Y, Bruckstein AM (1991) A probabilistic Hough transform. Pattern Recognit 24(4):303–316. doi:10.1016/0031-3203(91)90073-E
14. Lu W, Tan J, Floyd R (2005) Automated fetal head detection and measurement in ultrasound images by iterative randomized Hough transform. Ultrasound Med Biol 31(7):929–936. doi:10.1016/j.ultrasmedbio.2005.04.002
15. Michailovich OV, Tannenbaum A (2006) Despeckling of medical ultrasound images. IEEE Trans Ultrason Ferrolelectr Freq Control 53(1):64–78
16. Perona P, Malik J (1990) Scale space and edge detection using anisotropic diffusion. IEEE Trans Pattern Mach Intell 12(7):629–639. doi:10.1109/34.56205
17. Sanders R, James A (1985) The principles and practice of ultrasonography in obstetrics and gynecology. Appleton Century Crofts, Connecticut
18. Song XF, Han P, Zou L, Chen DZ, Hu SX (2004) A new method for estimation of fetal weight using support vector machine. Chin J Biomed Eng 23(6):516–522

19. Stetzer BP, Thomas A, Amini SB, Catalano PM (2002) Neonatal anthropometric measurements to predict birth weight by ultrasound. J Perinatol 22(5):397–402. doi:10.1038/sj.jp.7210754
20. Vapnik VN (1998) Statistical learning theory. Wiley, New York
21. Wang SR, Sun YN, Chang FM (2008) Artifact removal and texture-based rendering for visualization of 3D fetal ultrasound images. Med Biol Eng Comput 46:575–588. doi:10.1007/s11517-007-0286-7
22. Warsof SL, Wolf P, Coulehan J, Queenan JT (1986) Comparison of fetal weight estimation formulae with and without head measurements. Obstet Gynecol 67(4):569–573
23. Xu CY, Prince JL (1998) Snake, shapes, and gradient vector flow. IEEE Trans Image Process 7(3):359–369. doi:10.1109/83.661186
24. Xu L, Oja E, Kultanen P (1990) A new curve detection method: Randomized Hough transform (RHT). Pattern Recognit Lett 11(5):331–338. doi:10.1016/0167-8655(90)90042-Z
25. Yu JH, Wang YY, Chen P, Shen YZ (2008) Fetal abdominal contour extraction and measurement in ultrasound images. Ultrasound Med Biol 34(2):169–182. doi:10.1016/j.ultrasmedbio.2007.06.026
26. Yu YJ, Acton ST (2002) Speckle reducing anisotropic diffusion. IEEE Trans Image Process 11(11):1260–1270. doi:10.1109/TIP.2002.804276
27. Yuen HK, Illingworth J, Kitter J (1989) Detecting partially occluded ellipse using the Hough transform. Image Vis Comput 7(1):31–37. doi:10.1016/0262-8856(89)90017-6