Automatic Grading of Placental Maturity Based on LIOP and Fisher Vector

Baiying Lei, Xinyao Li, Yuan Yao*, Shengli Li, Siping Chen, Yongjin Zhou, Dong Ni, and Tianfu Wang*

Abstract— **Currently, the evaluation of placental maturity has mainly focused on subjective measure, which highly depends on the observation and experiences of the clinicians and not reliable. This paper proposes a new method for grading placenta maturity in B-mod ultrasound (US) images automatically based on local intensity order pattern (LIOP) and fisher vector (FV). After extracting invariant LIOP feature from the affine covariant region, the feature is encoded by FV to improve the classification accuracy and reduce the processing time. Experimental results show the effectiveness of the proposed method with an accuracy of 0.9375, a sensitivity of 0.9804 and a specificity of 0.9375 for the placental maturity grading. Moreover, experimental results demonstrate that the LIOP feature outperforms the traditional SIFT feature for grading.**

I. INTRODUCTION

Ultrasound (US) imaging has been widely used in prenatal diagnosis due to its radiation-free, direct-use and low cost. Placental maturity grading based on B-mode ultrasound image is the most frequently used functional evaluation of placental abnormalities such as fetal death, still birth, small gestational age and various pregnancy complications. Placental function is an important index to directly assess the fetal growth and development, which reflects intrauterine growth conditions to ensure fetus health. However, this method relies too much on visual observation of placental ultrasound images to determine their degree of calcification, and may cause errors by doctor's misjudgment and discrepancies. The subjective evaluation requires high level and experienced doctors, but it is challenging in the relatively backward areas. To resolve the adverse effects, the computer-assisted classification method is developed to make the judgment not only from doctor's subjective diagnosis experience, but also the extracted feature, which can be more precise.

In the literature, automatic classification algorithm for ultrasound placental maturity classification has been developed to reduce the incidence of judgment error, standardize medical test and reduce the doctor's workload. For example, the first placental maturity classification was proposed by Grammum [1] in 1979 to divide the chorionic plate, substance and basal plate of the placenta into four levels. However, this method relied on visual observation of placental ultrasound images to determine the calcification degree, which was highly dependent on the subjective judgment of the operator. Liu et al. [2] proposed to automatically classify placenta maturity using SVM classifier. A classification rate of 90% had obtained based on the three quantitative parameters: gray variance, distortions, and kurtosis. However, this classification result is not accurate enough. Overall, the current methods have not been applied for clinical practice. Developing a practical computer-based automatic placental maturity grading method is very desirable.

To develop a new grading algorithm for placental maturity, interest points should be first detected, and then invariant feature descriptors are calculated based on these points. There are lots of methods available for affine covariant region detection such as Harris [3] and difference of Gaussian [4]. As to the feature descriptor, the widely used SIFT [4] obtains better performance than derivative or moment based descriptor. Placental ultrasound images are subject to complex illumination modification, exposure time change and specular reflection by the imaging process, which is more challenging, and hence SIFT descriptor may be not suitable for this task. Meanwhile, local intensity order pattern(LIOP) has proved to be a very effective method successfully applied in object classification [5]. LIOP descriptor is robust to many variations and distortions. Feature point detection algorithm based on Harris-Laplace is more desirable than Harris due to its invariance to change. Therefore, Harris-Laplace feature point detection algorithm and LIOP local image descriptors are integrated for placental grading. In order to further boost the grading performance, the extracted features are encoded before representing them by a histogram of occurrence. The most popular encoding methods for grading are bag of visual words (BoVW) [6], aggregated codes of BoVW extensions such as vector of locally aggregated descriptor (VLAD) [7] and fisher vector (FV) [4].

Placental grading based on B-mode gray-scale ultrasound images are related with calcification, image quality constraints and other external conditions. To the best of our knowledge, there is no uniform standard and successful application for automatic grading of placental maturity in the clinical practice. According to gestational stages of placental chorionic plate, the 4-grades placental maturity [1, 8] is shown in Table 1 based on placental variations, chorionic plate, placental substance and basal layer. The aim of this paper is to develop an automatic technique to grade the placenta maturity based on the grading standards specified in Table 1.

[1]B. Lei, X. Li, S. Chen, Y. Zhou, D. Ni, and T. Wang* are with Department of Biomedical Engineering, School of Medicine, Shenzhen University, National-Regional Key Technology Engineering Laboratory for Medical Ultrasound, Guangdong Key Laboratory for Biomedical Measurements and Ultrasound Imaging, Shenzhen, China.

[2]Y. Yao* and S. Li are with Department of Ultrasound, Affiliated Shenzhen Maternal and Child Healthcare, Hospital of Nanfang Medical University, Shenzhen, China.

This work was supported partly by National Natural Science Foundation of China (Nos. 61101026, 61372006, 60871060 61031003 and 81270707), Shenzhen Key Basic Research Project (Nos. 201101013, 201202076, JCYJ20130329105033277, and JSE201109150013A), China Postdoctoral Science Foundation Funded Project (No. 2013M540663) and National Natural Science Foundation of Guangdong Province (No. S2013040014448).
(Email: *yaoyuan90@126.com, *tfwang@szu.edu.cn).

TABLE1 CHARACTERISTICS OF 4 GRADES OF PLACENTAL MATURITY.

Grade	Chorionic plate	Substance	Basal layer
0	Straight, smooth and chiseled	uniform	No echo
1	Slight undulating	unvenly distributed, scattered, point-like	No echo
2	In a serrated form, may extend into the substance of the placenta, but not the basal layer	Linearly echogenic, comma-like densities	Linear aligned, point-like echo
3	Jaggered, stretched into basal layer	Circular densities, halo with cast acoustic shadow	Large, confluent with basal layer, Acoustic shadow

II. METHODOLOGY

A. Grading Framework

As shown in the system framework of Figure 1, the input images are first pre-processed such as noise reduction. After Harris interest detection on placental images, the LIOP features are extracted and then encoded by FV. The procedure of feature vector formation is illustrated in Figure 2. The input placental image is first partitioned into patches, and each patch is represented by the patch descriptor using LIOP. Gaussian mixture model (GMM) is applied to generate k Gaussians based on the assumption of diagonal covariance matrix. A set of LIOP feature is encoded by FV into a single feature vector. Moreover, histogram of occurrence is created by k-means to cluster FV representatives. Feature normalization is applied to improve the classification accuracy. After encoding, the classification task is completed by the widely used support vector machine (SVM) method as it can locate the global optimal values and solve the over-fitting problem.

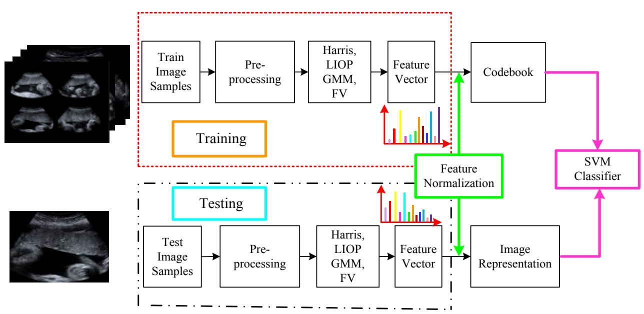

Figure 1. System framework.

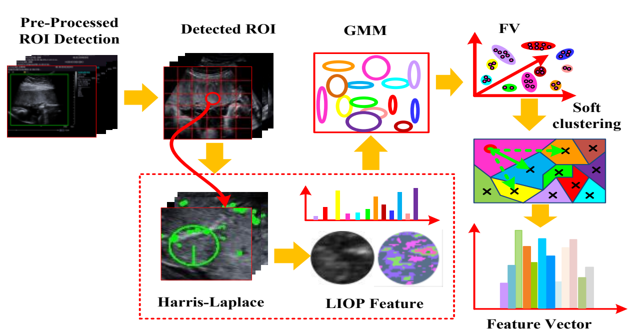

Figure 2. Feature vector formation.

B. Harris-Laplace Feature Point Detection

In the feature point detection stage, the Harris-Laplace algorithm [1] is adopted to detect key points of placental image since it is invariant to illumination changes, image noise addition and scale change. Harris-Laplace operator based on Harris corner detection algorithm employs scale space theory to find the maximum response point at multi-scale space, adaptive scale correlation matrix is defined as follows:

$$C(x,y,\sigma_I,\sigma_D)=\sigma_D^2 g(x,y,\sigma_I)*\begin{bmatrix} L_x^2(x,y,\sigma_D) & L_xL_y(x,y,\sigma_D) \\ L_xL_y(x,y,\sigma_D) & L_y^2(x,y,\sigma_D) \end{bmatrix} \quad (1)$$

where σ_I is integral scale, σ_D is differential scale, $\sigma_D{=}s\sigma_I$. The corresponding Gaussian function is $g(x,y,\sigma_I)$. The approximate image gradient in the x-direction and y-direction is $L_x(x,y,\sigma_D)$ and $L_y(x,y,\sigma_D)$. Laplace operator is defined:

$$J(x,y,\sigma_D)=\left|\sigma_D{}^2 L_{xx}(x,y,\sigma_D)+L_{yy}(x,y,\sigma_D)\right| \quad (2)$$

Laplace operator will get the maximum point above the threshold as the characteristic intensity scale α.

C. LIOP

LIOP algorithm is a method characterizing the local image luminance of order information [5]. The overall brightness of ROI is divided into a plurality of sequence information in each sub-region. As a local image descriptor, LIOP is obtained by sorting the selected image samples in increasing intensity based on the concept of local order pattern. This feature is invariant to light, monotonic intensity change of image, perspective changes, lossy compression and image blur. The order patterns are rotation invariant [5] by grouping the neighborhood sample around a pixel x, which is illustrated in Figure 3. The points are anticlockwise sampled on a circle at a radius of r.

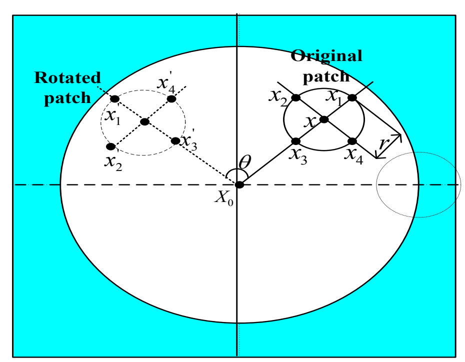

Figure 3. Layout of LIOP descriptor: shaded area is input patch (square area), white area is circular measurement region, and blue area is local neighborhood of a point.

Let N represent N sample points around the current pixel x, $Ind(\pi)$ is index value in the index table, $V_{N!}^{i}$ represents a feature vector. LIOP characteristics are defined as:

$$LIOP(x)=\phi(\gamma(P(x)))=V_{N!}^{Ind(\gamma P(x))}=(0,\cdots,0,\underset{Ind(\gamma P(x))}{1},0,\cdots,0) \quad (3)$$

where $P(x)=\{I(x_1),I(x_2),\cdots I(x_N)\}\in P^N$, P^N is N integers, $I(x_i)$ represents a pixel value in the sampling point. Permuted pattern based on local order to sort the neighbors: $I(x_{\sigma(1)})\le I(x_{\sigma(2)})\le ...I(x_{\sigma(n)})$. All elements in the feature vector are 0, except the i-th element is 1. LIOP groups local order pattern and describes the distinctive and invariant image rotation regions, which is very suitable for the placental maturity classification. The image of a local LIOP feature vector is defined as：

$$LIOPdescriptor=(des_1,des_2,\cdots,des_B) \tag{4}$$

$$des_i=\sum\nolimits_{x\in bin_i} LIOP(x) \tag{5}$$

where B is a patch number, dimension of LIOP feature vector is $N!\times B$ .Once local order patterns are computed for all pixels x in the image, they can be pooled into a histogram to form an image descriptor. Moreover, pooling discards spatial information resulting in a warp-invariant statistics.

D. Fisher Vector

Inspired by remarkable results in [4], GMM model is implemented first to enhance the classification performance. Assuming a codebook learned by k-means: $\{\mu_k, k=1,...,K\}$, a set of local descriptors: $\{x_m, m=1,...,N\}$, the feature vector is extracted as follows:

1) Assign neighboring:

$$NN(x_m)=\underset{\mu_k}{\operatorname{argmin}}\|x_m-\mu_k\|. \tag{6}$$

2) Compute v_k :

$$v_k=\sum_{x_m:NN(x_m)=\mu_k} x_m-\mu_k. \tag{7}$$

3) Concatenate v_k and normalize all feature vectors.

To better fit data by a GMM model, higher order statistics (i.e. derivative) are concatenated together. Gaussian means and variances of the first and second order derivatives [4] between features and GMM center are computed by:

$$\Phi_k^{(1)}=\frac{1}{N\sqrt{w_k}}\sum_{m=1}^{N}\gamma_m(k)\left(\frac{x_m-\mu_k}{\sigma_k}\right), \tag{8}$$

$$\Phi_k^{(2)}=\frac{1}{N\sqrt{2w_k}}\sum_{m=1}^{N}\gamma_m(k)\left(\frac{(x_m-\mu_k)^2}{\sigma_k}-1\right), \tag{9}$$

where $\{w_k,\mu_k,\sigma_k\}$ are the GMM mixture weights, means, and diagonal covariance. $\gamma_m(k)$ is the soft assignment weight of the m-th feature x_m of the k-th Gaussian. The derivatives of the log-likelihood of the GMM model are encoded by FV. The main purpose of the encoding is to discriminate the distribution difference between a specific test image and all fitted training image. FV ϕ is obtained by concatenating the difference vectors together: $\phi=[...,\Phi_1^{(1)},\Phi_1^{(2)},...,\Phi_k^{(1)},\Phi_k^{(2)},...]$. Essentially, FV is soft assigned VLAD with high-order statistics and an extension of BoVW. For D dimensional feature vector, the main difference between the BoVW and FV can be represented as:

$$\phi_{BoVW}(x_m)=[0,...,0,1,0,...,0], \tag{10}$$

$$\phi_{FV}(x_m)=\left[0,...,0,\underbrace{\Phi_k^{(1)},\Phi_k^{(2)}}_{2D\text{ non-zero dim}},0,...,0\right]. \tag{11}$$

PCA is performed to reduce the dimension of feature vector as well as processing time. Since the uncorrelated features and GMM covariance matrices of diagonal assumption are consistent, PCA whitening is also applied to ensure that diagonal covariance matrix assumption is satisfied.

III. Experimental Results

A. Experiment Setup

In our experiment, there are a total of 443 images of placenta, 187 images of grade 0, 135 images of grade 1, 85 images of grade 2, and 36 images of grade 3. All images were acquired by an ultrasound scanner from a commercial US scanner (Acuson Sequoia 512, Siemens Medical Solutions, USA) from Shenzhen Maternal and Child Health Hospital. Fetal gestational age ranges from 18 to 40 weeks. Conventional US sweep was performed to obtain the images on pregnant women in the supine position by a radiologist with more than five years of experience in US obstetrics. The placental image samples of 4 classes are shown in Figure 4. Our system was implemented by the mixed programming technology of Matlab and C++. The interest detection and feature extraction time for an image (size: 1024×768) is 6 seconds (32GBs RAM, double quad-core multithreaded server with a single CPU). The whole processing time for the testing step requires less than 1 second on a single CPU core.

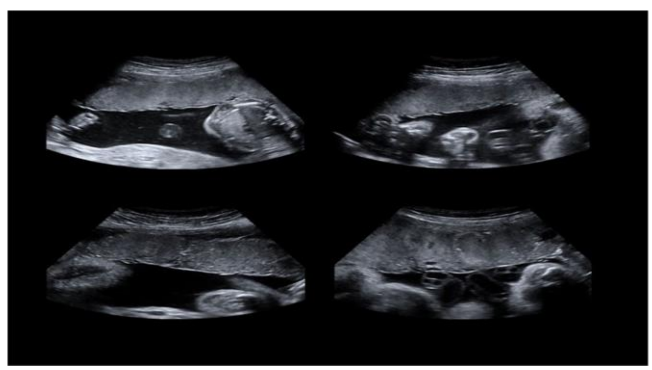

Figure 4. Image samples of 4 grades, the left upper is grade 0 image, right upper is grade 1 image, the left bottom is grade 2 image, and the right bottom is grade 3 image.

The placental maturity grading problem is quantitatively evaluated by the classification accuracy (the ratio between the number of correctly classified samples and the actual number of samples in each class). Moreover, quantitatively expressed classification metrics such as mean average precision (mAP), average accuracy, sensitivity and specificity are adopted for performance evaluation. The experiments are repeated at least 10 times and the average results are reported.

The classification results with precision and recall curve are shown in Figure 5, which demonstrates that the highest grading result is achieved by grade 0. We can see that grades 0-2 are more easily to be discriminated than grade 3. Grade 3 is the most confused in placental maturity classification, which requires a lot discriminative power to separate it.

The confusion matrix of the placental grading is presented in Figure 6, where rows denote actual grading levels, the columns mean the predicted grading levels. The diagonal elements denote the mean grading accuracy for each grade. As shown in the confusion matrix, the mean accuracy of 4 grades is 0.9375. Moreover, the overall classification accuracy for each grade is very high. The obtained remarkable grading scores indicate the effectiveness of the LIOP feature and FV algorithm.

Table 2 shows the classification results for the four grades in terms of accuracy, mAP, sensitivity and specificity. It is observed that mAP is often higher than accuracy in the planental maturity grading. It is noteworthy that the LIOP feature outperforms traditional SIFT feature in grading the plaental maturity, namely, grading performance can be boosted by taking advantage of LIOP feature. Given the comparison result, highest classification results have been achieved using FV encoding method. In general, aggregating vectors methods (VLAD and FV) demonstrate better classification results than the traditional BoVW method. The obtained high grading result based on the proposed method demonstrates the practical application in the clinical practice. It also indicates that FV is very suitable and effective for the placental grading.

IV. CONCLUSIONS

In this paper, an automatic grading system for placenta maturity is presented based on LIOP and FV. The experimental results demonstrate that the proposed method accurately grade the placenta with promising performance. Computer-assisting placenta grading in US fetal image not only provides effective diagnosis, but also saves time, tedious work and labor associated with the diagnosis. Furthermore, this method is generalized and can be extended to other classification task.

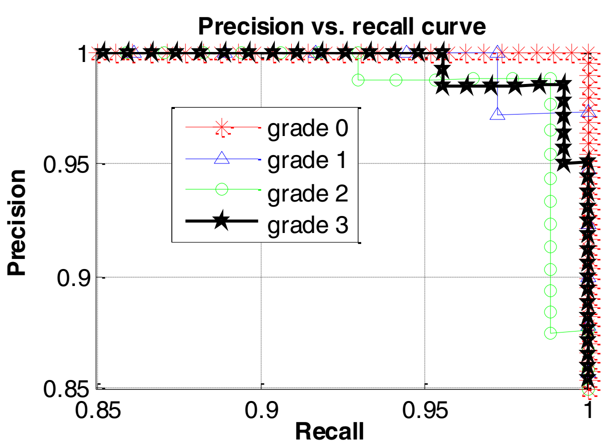

Figure 5. Precision vs. recall curve of the grading algorithm.

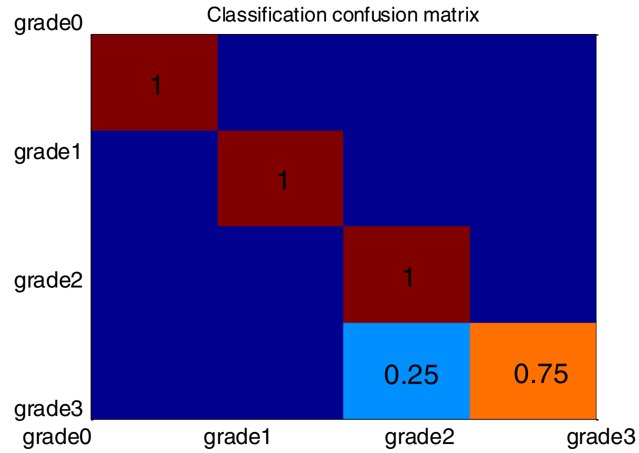

Figure 6. Confusion matrix of placental maturity grading.

TABLE 2. PLACENTAL MATURITY GRADING RESULTS.

Method	Feature	mAP	Accuracy	Sensitivity	Specificity
BoVW	SIFT	0.8519	0.7235	0.9116	0.724
VLAD	SIFT	0.8851	0.7307	0.917	0.731
FV	SIFT	0.9531	0.7932	0.9261	0.788
BoVW	LIOP	0.95	0.8365	0.9725	0.9143
VLAD	LIOP	0.9706	0.9182	0.975	0.907
FV	LIOP	1	0.9375	0.9808	0.9375

REFERENCES

[1] P. A. Grannum, R. Berkowitz, and J. C. Hobbins, "The ultrasonic changes in the maturing placenta and their relation to fetal pulmonic maturity," *Am. J. Obstet. Gynecol.*, vol. 133, pp. 915-22, 1979.

[2] Z. Liu, H. Zheng, and S. Lin, "Application of Multi-Classification Support Vector Machine in the B-Placenta Image Classification," in *Proc. of Int. Conf. on Computational Intell. and Software Eng.*, 2009, pp. 1-4.

[3] K. Mikolajczyk and C. Schmid, "Scale & affine invariant interest point detectors," *Int. J. of Comput. Vis*., vol. 60, pp. 63-86, 2004.

[4] J. Sánchez, F. Perronnin, T. Mensink, and J. Verbeek, "Image classification with the fisher vector: theory and practice," *Int. J. of Comput. Vis*., pp. 1-24, 2013.

[5] Z. Wang, B. Fan, and F. Wu, "Local intensity order pattern for feature description," in *Proc. of Int. Conf. on Comput. Vis.*, 2011, pp. 603-610.

[6] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: spatial pyramid matching for recognizing natural scene categories," in *Proc. of IEEE Conf. on Comput. Vis. and Pattern Recognit*., 2006, pp. 2169-2178.

[7] H. Jégou, F. Perronnin, M. Douze, and C. Schmid, "Aggregating local image descriptors into compact codes," *IEEE Trans. on Pattern Anal. and Mach. Intell.*, vol. 34, pp. 1704-1716, 2012.

[8] P. A. Linares, P. J. McCullagh, N. D. Black, and J. Dornan, "Feature selection for the characterization of ultrasonic images of the placenta using texture classification," in *Proc. of Int. Symp. on Biomed. Imag.*, 2004, pp. 1147-1150.