Identification of Ultrasonic Echolucent Carotid Plaques Using Discrete Fréchet Distance Between Bimodal Gamma Distributions

Xiaowei Huang, Yanling Zhang, Long Meng, Ming Qian, Member, IEEE, Kelvin K. L. Wong, Derek Abbott, Fellow, IEEE, Rongqin Zheng, Hairong Zheng* Senior Member, IEEE and Lili Niu*

Abstract—Objective: Echolucent carotid plaques are associated with acute cardiovascular and cerebrovascular events (ACCEs) in atherosclerotic patients. The aim of this study was to develop a computer-aided method for identifying echolucent plaques.

Methods: A total of 315 ultrasound images of carotid plaques (105 echo-rich, 105 intermediate and 105 echolucent) collected from 153 patients were included in this study. A bimodal gamma distribution was proposed to model the pixel statistics in the grayscale images of plaques. The discrete Fréchet distance features (DFDFs) of each plaque were extracted based on the statistical model. The most discriminative features (MDFs) were obtained from DFDFs by linear discriminant analysis, and a k-nearest-neighbor classifier was implemented for classification of different types of plaques.

Results: The classification accuracy of the three types of plaques using MDFs can reach 77.46%. When a receiver operating characteristics (ROC) curve was produced to identify echolucent plaques, the area under the curve was 0.831.

Conclusion: Our results indicate potential feasibility of the method for identifying echolucent plaques based on DFDFs. Significance: Our method may potentially improve the ability of noninvasive ultrasonic examination in risk prediction of ACCEs for patients with plaques.

Index Terms—Bimodal gamma distribution, Carotid plaque, Discrete Fréchet distance, Ultrasound imaging.

I. INTRODUCTION

ACUTE cardiovascular and cerebrovascular events (ACCEs) are major causes of disability and premature death worldwide for subjects without obvious symptoms [1, 2]. Plaque echogenicity assessed by B-mode ultrasound has been reported to be associated with cardiovascular and cerebrovascular events in previous studies [3-5]. Echolucent plaques have higher lipid and macrophage, signifying a higher risk for ACCEs [6-9]. Therefore, we proposed a computer-aided method to identify echolucent plaques, which may potentially improve risk prediction of ACCEs for patients with plaques.

In recent years, the gamma distribution has been widely adopted for modeling gray level distribution in synthetic aperture radar images [10, 11], mammogram images [12], and ultrasound images [13-17]. Li et al. have reported that the gamma distribution is useful for the statistical modeling of amplitude synthetic aperture radar images [10]. Qin et al. have indicated that the gamma distribution is a flexible empirical model for synthetic aperture radar images, and the categories of synthetic aperture radar images can be discriminated using the Kullback–Leibler distance between gamma distributions [11]. Gumaei et al. have shown that the gamma distribution is suitable for describing symmetric and non-symmetric mammogram images, and can improve the accuracy of breast cancer detection. Tao et al. model cardiac ultrasound images based on four families of distributions (gamma, Weibull, normal, and log-normal), and find that the gamma distribution demonstrates improved performance in fitting the data and a lower misclassification rate in classifying blood and tissue [15]. The study of Vegas-Sánchez-Ferrero et al. have shown desirable performance of the generalized gamma distribution in characterizing the speckle of blood and myocardial tissue in ultrasonic images [16]. They also propose a plaque characterization method for intravascular ultrasound images based on the gamma distribution [17]. The unimodal gamma distribution is characterized by two parameters and is useful for describing most of the homogeneous images, but it may be inadequate for modelling images with complex structures [18, 19]. The distribution of gray levels in B-mode ultrasound images of carotid plaques is complicated as the plaques have heterogeneous histologic components (e.g. calcifications, lipids, hemorrhages, fibrous tissue etc.) [20]. Thus, a bimodal gamma distribution with five parameters may be sufficient to describe the ultrasound images of plaques. Shankar et al. report 33 ultrasound images of plaques indicating that a bimodal gamma distribution model is effective in modeling hard and soft plaques [13]. However, this is an incomplete scheme for feature extraction and classification of plaques with different echogenicity. According to the criteria of the European carotid...
plaque study group, the plaques can be classified into echo-rich, intermediate and echolucent plaques [21]. In our study, the bimodal gamma distribution is proposed to model the pixel statistics of the gray scale images of the three different types of plaques. There are three curves representing the echo-rich, intermediate and echolucent plaques in the model. Furthermore, previous studies have indicated that discrete Fréchet distance is feasible for quantitative assessment of the similarity of two curves [22, 23]. The discrete Fréchet distance features (DFDFs) were extracted from the relationship between the cumulative distribution curve of the pixel gray value distribution of each plaque and the three curves in the statistical model, respectively.

Our study aims to develop a computer-aided method based on DFDFs for identifying echolucent plaques, which may potentially improve the ability of ultrasonic examination in risk prediction of ACCes for patients with carotid plaques. The flowchart of the method is shown in Fig. 1.

![Flowchart of the method](image)

**Fig. 1. Flowchart for the classification of echolucent carotid plaques based on the discrete Fréchet distance features.**

**II. MATERIALS AND METHODS**

**A. Materials**

A total of 315 ultrasound images of carotid plaques (105 echo-rich, 105 intermediate and 105 echolucent) collected from 153 subjects were analyzed in our study. The plaques were randomly divided into a training group (70 echo-rich, 70 intermediate and 70 echolucent images) and a test group (35 echo-rich, 70 intermediate and 35 echolucent images). In order to improve reliability of the results, a 3-fold cross validation test was performed, and the mean accuracy was taken as the final result.

**B. Images Acquisition and Preprocessing**

From February 2014 to October 2015, ultrasound carotid plaque images were collected using an Aplio XG (SSA-790A, Toshiba Medical Systems, Japan) equipped with a 5-12 MHz (center frequency is 8 MHz) linear-array transducer (PLT-805AT) and a MyLab90 (Esaote Medical Systems, Italy) equipped with a 4-13 MHz (center frequency is 8 MHz) linear-array transducer (LA523) by a sonographer with 5 years of experience in vascular imaging. The carotid artery was examined with the head tilted slightly upward in the mid-line position. The transducer was manipulated so that the near and far walls were parallel to the transducer footprint, and the lumen diameter was maximized in the longitudinal plane. All participants provided written informed consent. The study protocol was approved by the Institutional Review Board of the third affiliated hospital of Sun Yat-sen University (Guangzhou, China).

Plaque echogenicity was visually classified into type 1 (echo-rich), type 2 (intermediate) and type 3 (echolucent) according to the criteria of the European carotid plaque study group [21]. The assessment of plaque echogenicity was performed by two sonographers with at least 5 years experience in vascular imaging, and the controversial plaques were removed. A Cohen’s kappa coefficient (κ) was calculated to evaluate the inter-observer agreement. To improve the comparability of the images and the reliability of our results, the images were normalized according to the scheme proposed by Sabetai et al. [24]. After normalization, the gray-scale median (GSM) of the blood range from 0 to 5, and the GSM of adventitia range from 185 to 195.

**C. Statistics of Gray Value Distribution of Plaque Pixels**

The region of interest within each plaque image was manually defined in gray scale, and the statistics of the gray value distribution of each plaque pixel was obtained. Let \( x \) denote pixel value, where \( f(x) \) is the probability density function (PDF) of \( x \), and the empirical cumulative distribution function (CDF), \( F_{\text{plaque}}(x) \), can be expressed by:

\[
F_{\text{plaque}}(x) = \sum_{i=0}^{255} f(x_i), \quad x_i = 0, 1, \ldots, 254, 255. \tag{1}
\]

**D. Bimodal Gamma Distribution**

For a parameter \( \alpha > 0 \), gamma function \( \Gamma(\alpha) \) is defined by:

\[
\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx. \tag{2}
\]

A variable \( x \) that is gamma-distributed with shape \( k \) and scale \( \theta \) is denoted by \( X \sim \Gamma(k, \theta) \). Its PDF is defined as

\[
f(x; k, \theta) = \frac{x^{k-1} e^{-x/\theta}}{\theta^k \Gamma(k)}, \quad x > 0, k > 0, \theta > 0. \tag{3}
\]

The corresponding CDF is calculated by Equation (4). For a bimodal gamma distribution, its PDF and CDF can be expressed by Equations (5) and (6), respectively.
where $M$ is the plaque label in the training set, $M$ respectively. Further, we proposed a statistical model that given by:

$$F_{\text{model-gamma}}(x;eta, k_1, \theta_1, k_2, \theta_2) = \int_0^x \left[ \beta \frac{x^{k_1-1} e^{-x/\theta_1}}{\theta_1^k \Gamma(k)} + (1-\beta) \frac{x^{k_2-1} e^{-x/\theta_2}}{\theta_2^k \Gamma(k)} \right] dx.$$  

(6)

E. Statistical Model of Plaques

We investigated the ability of a bimodal gamma distribution in modeling the gray scale distribution of the plaque. There are five parameters ($\beta$, $k_1$, $\theta_1$, $k_2$ and $\theta_2$) that can be adjusted for curve-fitting implemented between the CDF of the bimodal gamma distribution and the empirical CDF of the plaque. In order to guarantee the reliability of curve-fitting, the parameters were obtained by minimizing the error (all $\varepsilon < 0.05$), given by:

$$\varepsilon = \sum_{x=0}^{255} \left[ F_{\text{plaque}}(x) - F_{\text{gamma}}(x) \right]^2$$  

(7)

where $F_{\text{plaque}}(x)$ and $F_{\text{gamma}}(x)$ were obtained from (1) and (6), respectively. Further, we proposed a statistical model that can be used for classifying the three types of plaques.

F. Feature Extraction

1) Discrete Fréchet Distance

Assume a path $Z = (z_1, \ldots, z_r)$ of $r$ vertices, a $t$-walk along $Z$ means $Z$ is partitioned along the path into disjoint nonempty subpaths $\{Z_j\}_{j=1,\ldots,t}$ such that $Z_j = (z_{r_j+1}, \ldots, z_{r_j})$ and $0 = r_0 < r_1 < \cdots < r_r = r$. For two paths $A = \{a_1, \ldots, a_n\}$ and $B = \{b_1, \ldots, b_m\}$, a paired walk $w = \{A_j, B_j\}$ along $A$ and $B$ is a $t$-walk $\{A_j\}_{j=1,\ldots,t}$ along $A$ and a $t$-walk $\{B_j\}_{j=1,\ldots,t}$ along $B$ for some $t$, such that, for $1 \leq j \leq t$, $|A_j| = 1$ or $|B_j| = 1$ (i.e., either $A$ or $B$ contains exactly one vertex). The discrete Fréchet distance between two paths $A$ and $B$ is:

$$\text{dis}_F(A, B) = \min_{w} \max_{j \in (a, b) \in A \times B} \text{dis}(a, b)$$  

(9)

where $\text{dis}(a, b)$ represents the Euclidean distance between two points $a$ and $b$. In this study, the discrete Fréchet distance was proposed to measure the similarity between CDF of each plaque and the three model curves. The discrete Fréchet distance between two paths $F_{\text{model-gamma}}(x)$ or $F_{\text{plaque}}(x)$ is:

$$\text{DFDF} = \text{dis}_F(F_{\text{model-gamma}}(x_i), F_{\text{plaque}}).$$  

(10)

2) Discrete Fréchet Distance Feature Extracted from Piecewise Curves

In order to acquire more effective features, we extracted the DFDFs from the piecewise curves. The CDF of each plaque and the model curves were divided into step lengths of 255, 50 or 30 gray levels, as shown in Table I.

3) Gray-scale Median

The GSM of a carotid plaque is a common metric for evaluating plaque echogenicity. We measured the GSM of each plaque in this study. The effectiveness of GSM in identifying the echoluent plaques was investigated and compared with that of DFDFs.

G. Feature Reduction and Classification

1) Linear Discriminant Analysis

Before classification, feature reduction was necessary to consider the relative large amount of DFDFs. Linear discriminant analysis (LDA) is a useful tool for pattern recognition [25]. It is widely used in process of feature reduction by providing a linear transformation of the feature space. Here, LDA was implemented for DFDFs and the most discriminating features (MDFs) were obtained.

Let $\text{DFDF}_g^c$ denote the DFDF of the $g$-th plaque in class $c$, and $c = 1, 2, 3$. Define the within-class scatter matrix $S_w$ as:

$$S_w = \frac{1}{N_{\text{all}} \sum_{g=1}^{N_c} \sum_{c=1}^{N_c} (\text{DFDF}_g^c - \mu_c)(\text{DFDF}_g^c - \mu_c)^T}$$  

(11)

where $\mu_c$ is the mean vector of the $c$-th class, $N_{\text{all}}$ is the number of all the sample, $N_c$ is the number of the sample in the $c$-th class. Define the between class scatter matrix $S_B$ as:

$$S_B = \frac{1}{N_{\text{all}} \sum_{c=1}^{N_c} N_c (\mu_c - \mu)(\mu_c - \mu)^T}$$  

(12)

where $\mu$ is the mean vector of the pooled data. The aim of LDA is to find a linear transform matrix $\Theta$ such that the objective function as maximizing the following:

$$\Theta = \max_{\Theta} \frac{\text{tr}(\Theta^T S_B \Theta)}{\text{tr}(\Theta^T S_w \Theta)}$$
\[ J(\omega) = \arg\max_\omega \frac{\omega^\top S_B \omega}{\omega^\top S_W \omega}. \] (13)

It can be proved that such a transform \( \Phi \) is composed of eigenvectors corresponding to largest eigenvalues of \( S_B S_W \). The MDFs can be produced by the transformation:

\[ \text{MDFs} = \Phi^\top \text{DFDFs}. \] (14)

**III. RESULTS**

**A. Visual Classification**

The visual classification of the carotid plaques (\( n = 330 \)) into three different types showed a good inter-agreement (Table II). The inter-observer reproducibility was 97.57% (\( \kappa = 0.964 \)). A total of 8 controversial plaques were excluded, and the remaining 105 echo-rich, 105 intermediate and 105 echolucent consensual plaques were randomly selected for the following analysis.

**B. Model Curves of Plaque Based on Bimodal Gamma Distribution**

Fig. 2 shows that the CDF of bimodal gamma distribution is effective in fitting the empirical CDF of the echo-rich, intermediate and echolucent plaques. The statistical model of the gray value distribution of plaque was shown in Fig. 3. The CDF of bimodal gamma distribution with parameters \( (\beta = 0.5, k_1 = 4.74, \theta_1 = 18.19, k_2 = 13.14, \theta_2 = 6.34), (\beta = 0.5, k_1 = 3.76, \theta_1 = 29.19, k_2 = 11.41, \theta_2 = 8.03) \) and \( (\beta = 0.5, k_1 = 6.73, \theta_1 = 15.27, k_2 = 9.51, \theta_2 = 16.48) \) represents the echolucent, intermediate and echo-rich plaques. The distance between the three model curves showed a trend of first increase and then decrease as gray value increasing (Fig. 3). And the curve of echolucent plaque can be distinguished from the other two model curves.

**TABLE II VISUAL CLASSIFICATION OF 330 CAROTID PLAQUES BY TWO OBSERVERS**

<table>
<thead>
<tr>
<th>Second observer</th>
<th>First observer</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Echo-rich</td>
</tr>
<tr>
<td>Echo-rich</td>
<td>106</td>
</tr>
<tr>
<td>Intermediate</td>
<td>1</td>
</tr>
<tr>
<td>Echolucent</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>107</td>
</tr>
</tbody>
</table>

\( \kappa = 0.964 \)
Fi 2. Cumulative distribution function (CDF) of bimodal gamma distribution fitting empirical CDF of the gray value distributions of the plaques. The empirical CDFs in (a) was drew according to the echolucent, intermediate and echo-rich plaques defined in (b), (c) and (d), respectively. The parameters of fitting curve were $\beta = 0.31$, $k_1 = 4.44$, $\theta_1 = 13.79$, $k_2 = 20.49$, $\theta_2 = 4.44$, $\epsilon = 0.0020$ (echolucent); $\beta = 0.19$, $k_1 = 21.93$, $\theta_1 = 6.69$, $k_2 = 19.54$, $\theta_2 = 4.32$, $\epsilon = 0.0028$ (intermediate); $\beta = 0.42$, $k_1 = 13.84$, $\theta_1 = 11.76$, $k_2 = 20.23$, $\theta_2 = 4.94$, $\epsilon = 0.0019$ (echo-rich).

Fig. 3. Model curves of the three types of plaques based on bimodal gamma distribution. The parameters of CDF are $\beta = 0.5$, $k_1 = 4.74$, $\theta_1 = 18.19$, $k_2 = 13.14$, $\theta_2 = 6.34$ (echolucent); $\beta = 0.5$, $k_1 = 3.76$, $\theta_1 = 29.19$, $k_2 = 11.41$, $\theta_2 = 8.03$ (intermediate); $\beta = 0.5$, $k_1 = 6.73$, $\theta_1 = 15.27$, $k_2 = 9.51$, $\theta_2 = 16.48$ (echo-rich).

C. Feature Extraction and Classification

For each plaque, a total of 42 DFDFs (feature set 1 with 3 features, feature set 2 with 15 features, and feature set 3 with 24 features) were obtained. LDA was performed for each feature set, and two MDFs (MDF1 and MDF2) were obtained. All MDF1 and MDF2 of feature set 1 to 3 were significantly different between the three types of plaques (all, $p < 0.001$) (Table III). Fig. 4 illustrates the two-dimensional scatter-plots of 315 plaques using MDF1 and MDF2. The scattered point distributions based on MDFs obtained from feature set 2 and set 3 constituted three more distinct congregate areas compared with feature set 1. Accordingly, the classification accuracy of three types of plaques were 68.25%, 75.87% and 77.46% when using MDFs of feature set 1 to 3 for training.

D. Identification of Echolucent Plaques

Previous studies proved that echolucent plaques were potentially unstable and were regarded as high-risk plaques, whereas echo-rich and intermediate plaques were considered as low-risk [6, 8]. When echolucent plaques were identified using feature set 2, the accuracy, sensitivity and specificity were 85.09%, 77.14% and 89.04%, which were higher than these using feature set 1 (79.05%, 67.62% and 84.76%) and feature set 3 (83.17%, 75.24% and 88.09%) (TABLE IV). Also the identification of echolucent plaques based on GSM showed a sensitivity of 72.38%, a specificity of 80.95% and a specificity of 63.33%. The receiver operating characteristic (ROC) is a standard method for assessing the sensitivity and specificity of diagnostic procedures, which provides a curve to describe the inherent tradeoff between the sensitivity and specificity of a diagnostic system. The ROC analysis was implemented to examine the ability of our method and GSM in identifying echolucent plaques. Fig. 5 shows the ROC curves for the KNN classifier when feature set 1 to 3 and GSM were used to train the classifier. The areas under the curve (AUC) for the KNN classifier were 0.762 and 0.831 and 0.812, which was higher than that for GSM (AUC = 0.712).
IV. DISCUSSION

In this study, a statistical model based on a bimodal gamma distribution was proposed to identify echo-rich, intermediate and echolucent plaques (Fig. 3). Once the model was built, the problem of classification of plaques with different echogenicity was transformed into a mathematical problem for distinguishing similarity between the CDF of each plaque and the model curves, and the novel DFDFs were extracted for evaluating the similarity. The classification accuracy can reach 77.46%, when classifier was trained with the DFDFs. Furthermore, previous studies suggest that patients with echolucent plaques have a high risk of carotid bifurcation lesions [26] and ischemic cerebrovascular events [7]. Moreover, plaque echolucency is useful for predicting coronary events [3] and future strokes [9]. Fig. 5 shows that our method has a high potential for identifying echolucent plaques (AUC = 0.831). Therefore, a computer-aided method based on DFDFs may be promising in risk prediction of ACCEs for patients with plaques.

Visual identification is a common method for classification of plaques with different echogenicity [21, 27-29]. Mayor et al. visually classify 95 plaques into 5 types: type 1 (uniformly echolucent); type 2 (predominantly echolucent);
type 3 (predominantly echogenic); type 4 (uniformly echogenic); and type 5 (unclassified plaques owing to calcification and producing acoustic shadows) [28] and they find that the plaques of type 1-5 present intermediate mean GSMs, respectively of 33, 58, 100, 127 and 163, respectively. The mean GSM is linear related with the types (Spearman $r = 1, p < 0.05$), and GSMs among different types of plaques show a statistical significance ($p < 0.02$). Plaques are classified visually into: type 1 (echo-rich), type 2 (intermediate) and type 3 (echolucent) in this study [21]. We compared the ability of DFDFs and GSM for identifying echolucent plaques. When the classifier was trained using MDFs of feature set2, the accuracy (85.09%) and the specificity (89.04%) were the highest, and a moderate sensitivity (77.14%) was achieved (TABLE IV). The identification of echolucent plaques based on GSM showed a highest specificity of 80.95%, but it had a lower sensitivity (72.38%) and specificity (63.33%). TABLE IV shows the 3-fold cross validation analysis in detailed. Overall, the classification based on DFDFs feature set 2 had a higher stability than based on feature set 1, feature set 3 and GSM in identifying echolucent plaques.

In order to improve the comparability of the images acquired from different ultrasonic condition, images standardization is necessary before data processing. According to the method proposed by Elatrozy et al. [30], images were standardized manually by linearly adjusting the image so that the median gray level value of the blood was 0–5, and the median gray level value of the adventitia (artery wall) was 185-195. Moreover, Sabetai et al. have indicated that such an image normalization can decrease the variability between storage media and between probes [24]. In our study, all ultrasound images were normalized as the same as in [31-33].

In 1906, the Fréchet distance was defined as a measure of similarity between two parametric curves [34]. In 1994, Heikki et al. presented the discrete Fréchet distance, which is used for approximately computing Fréchet distance between two arbitrary curves using the discrete nodes along the curves for the measurements [35]. Many recent studies have proven the effectiveness of discrete Fréchet distance [22, 23]. In this study, the most discriminating features extracted from the DFDFs showed statistical difference between plaques with different echogenicity (Table III). Furthermore, Irie et al. [36] in a study with 287 patients investigated the relationship between the echogenicity of carotid plaque and the occurrence of CVD events in detail. They divided the GSM values into quartiles (Q1: $\geq 59$, Q2: 48-58, Q3: 38-47, and Q4: $\leq 37$) and found that the lowest GSM quartile (Q4: GSM $\leq 37$) has much higher risk for CVD as compare to the other GSM quartiles. Ruiz-Ares et al. [37] analyzed 42 patients indicate that the unstable plaques have lower echogenicity than the stable plaques (GSM = 23 vs. 37, $p < 0.001$). These studies suggest that plaque with different GSM had different level of risk, which may consistent with our results since more effective DFDFs can be obtained from the piecewise CDF of each plaque and the model curves in different gray level ranges.

The main limitation of present study is that the model curves were built with relatively few samples, and a more precise model is the next step for investigation. Moreover, we only took 255, 50 and 30 gray levels as step lengths to segment the curves, and DFDFs were extracted from piecewise curves. In order to obtain an optimal step length, the curve segmentation scheme needs detailed investigation. Our results may be verified with additional ultrasound images of carotid plaques.

V. CONCLUSION

Our results demonstrate the potential feasibility of the method for identifying echoluent plaques based on DFDFs, which may potentially improve the ability of noninvasive ultrasonic examination in risk prediction of ACCs for patients with plaques.

REFERENCES


