Modeling Log-Compressed Ultrasound Images for Radio Frequency Signal Recovery

José Seabra and João Sanches

***Abstract*— This paper presents an algorithm for recovering the *radio frequency* (RF) signal provided by the ultrasound probe from the log-compressed ultrasound images displayed in ultrasound equipment.**

Commercial ecographs perform nonlinear image compression to reduce the dynamic range of the Ultrasound (US) signal in order to improve image visualization. Moreover, the clinician may adjust other parameters, such as brightness, gain and contrast, to improve image quality of a given anatomical detail. These operations significantly change the statistical distribution of the original RF raw signal, which is assumed, based on physical considerations on the signal formation process, to be Rayleigh distributed. Therefore, the image pixels are no longer Rayleigh distributed and the RF signal is not usually available in the common ultrasound equipment.

For statistical data processing purposes, more important than having "good looking" images, it is important to have realistic models to describe the data.

In this paper, a nonlinear compression parametric function is used to model the pre-processed image in order to recover the original RF image as well the contrast and brightness parameters.

Tests using synthetic and real data and statistical measures such as the *Kolmogorov-Smirnov* and Kullback-Leibler divergences are used to assess the results. It is shown that the proposed estimation model clearly represents better the observed data than by taking the general assumption of the data being modeled by a Rayleigh distribution.

I. Introduction

Ultrasound (US) imaging is the most prevalent diagnostic tool in almost all hospitals around the world. Unfortunately, the US images present small *Signal to Noise Ratio*(SNR), meaning poor quality, and are corrupted by a type of multiplicative noise called *speckle* [1].

Several statistical models have been proposed [2] to describe the *speckle* noise. A common one, derived from physical consideration on the RF signal generation process, is the Rayleigh distribution [3]. However, the RF signal is not usually available in the common ultrasound equipment and it is usually pre-processed (filtered and compressed) to improve its visualization. Compression is needed to reduce the dynamic range of the RF signal in order to adapted it to the dynamic range of the monitor.

Unfortunately, this pre-processing modifies the distribution of the RF signal. This step depends on a set of parameters such as the brightness, contrast, zoom and dynamic gain. These are tuned by the clinician in order to improve the visualization of an organ or region of interest (see Fig. 1). These processes affect the statistics of the original RF signal which is no longer Rayleigh (Fig. 2).

It is important that these processing operations could be reverted in order to obtain images Rayleigh distributed, independent on the particular parameter selected by the clinician. This is crucial to guarantee objective and reproducible tissue characterizations and realistic models for image reconstruction and de-noising.

Corresponding author: José Seabra (jseabra@isr.ist.utl.pt)
Affiliation: Systems and Robotics Institute / Instituto Superior Técnico, 1049-001 Lisbon, Portugal

Partially supported by FCT, under ISR/IST plurianual funding (POSC program, FEDER).

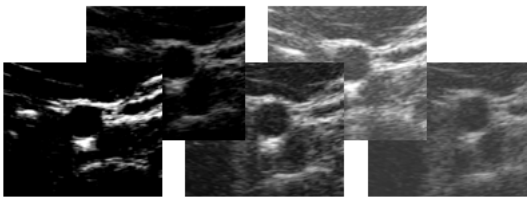

Fig. 1. Ultrasound images showing different contrast and brightness levels.

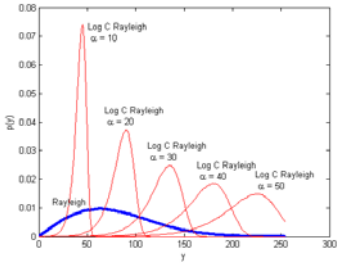

Fig. 2. Probability Density Functions (PDFs) for Rayleigh and log-compressed Rayleigh models.

The statistics of the *speckle* pattern, which depend on the microstructure of the tissue *parenchyma*, can be useful for differentiating between different tissue compositions or types [4]. In [5] the authors describe a method for carotid plaque characterization based on a Rayleigh observation model, which is known to be a reasonable approximation but not entirely correct.

The goal of this paper is to estimate the original RF signal from the observed pre-processed images for which it is known that the Rayleigh distribution is more appropriated to describe it. Here, relevance is not given to the improvement in the visualization obtained by log-compression operation because the main concern is to obtain a realistic model to describe the data.

The paper is organized as follows. Section II describes the US data model proposed here. Parameter modeling and estimation is described in section III and experimental results are presented in section IV. Section V concludes the paper.

II. Ultrasound Signal Modeling

Let Y be the $N \times M$ signal matrix (image) generated by the ultrasound probe of i.i.d. (independent and identically distributed) random variables, which are assumed to be Rayleigh distributed [3],

$$p(y_i|\psi_i) = \frac{y_i}{\psi_i} e^{-\frac{y_i^2}{2\psi_i}}, \qquad (1)$$

where y_i denotes the intensity of the i^{th} pixel of the data before compression and ψ_i is the value of the noiseless image i^{th} pixel , related with the tissue reflectivity at that location [4].

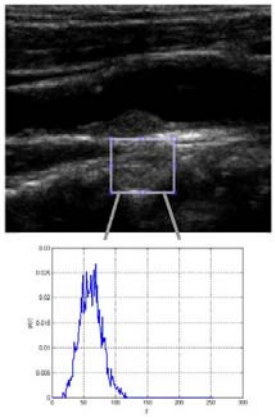

Fig. 3. Pixel histogram of a region selected from an ultrasound image.

Figs. 2 and 3 show clearly that the assumption of a Rayleigh distribution is not realistic. The histogram of an homogeneous region in a real ultrasound image (see Fig. 3), suggests that the data is modeled by a compressed, shifted version of this Rayleigh model.

Let Z be the ultrasound image where statistical independence of its pixels is assumed. This may not be realistic because the *point spread function* (PSF) of the image-acquisition system is larger than the interpixel distance. However, modeling of the pixels dependence highly increases the mathematical complexity of the algorithms without significative improvements in the final results [6].

The pre-processing procedure performed by the ultrasound equipment is modeled as follows,

$$z_{ij} = \alpha \, \log(y_{ij}+1) + \beta, \tag{2}$$

where the log function accounts for the compression, (α, β) are unknown parameters which account for the contrast and brightness, respectively, and z_{ij} is the intensity at the $(i,j)^{th}$ observed pixel. The distribution of z is given by

$$p(z) = \left| \frac{dy}{dz} \right| p(y), \tag{3}$$

where $p(y)$ is the probability density function (PDF) of the original non-compressed image and dy/dz is the value of the derivative of the inverse compression function [7]. By using (1 - 3),

$$p(z) = \frac{\eta\,(\eta+1)}{\alpha\,\psi(x)}\, e^{\frac{-\eta^2}{2\psi}}, \tag{4}$$

where $\eta = e^{\frac{z-\beta}{\alpha}} - 1$ and α and β are the parameters to be estimated.

III. Parameter Estimation

If y is a random variable (RV) which is described by a Rayleigh distribution (1) and the logarithmic compression expression is given by (2), the distribution of z, $p(z)$, is a Fisher-Tippett distribution (double exponential) [8,9]. After rearranging the previous equations, $p(z)$ becomes

$$p(z) = \frac{2}{\alpha}\, e^{-\theta - e^{-\theta}}, \tag{5}$$

with $\theta = \log(2\psi) - 2\frac{y-\beta}{\alpha}$.

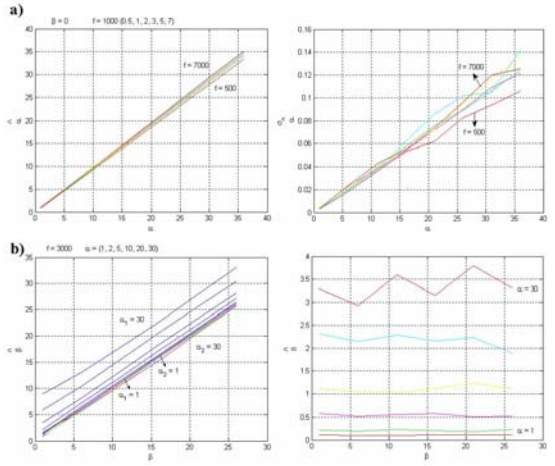

Fig. 4. Mean and SD of the estimators $\hat{\alpha}$ (a) and $\hat{\beta}_0$ (b).

The expressions for the mean and standard deviation for this distribution are

$$\mu_z = \frac{\alpha}{2}\,[\log(2\psi) - \gamma] + \beta. \tag{6}$$

$$\sigma_z = \pi\,\alpha/\sqrt{24} \tag{7}$$

where $\gamma = 0.5772...$ is the Euler-Mascheroni constant. From (7) an estimator for α can be easily derived,

$$\hat{\alpha} = \sqrt{\frac{24}{\pi^2}\,\sigma_z^2}. \tag{8}$$

Assuming that $\eta = e^{\frac{z-\beta}{\alpha}} - 1$ is Rayleigh distributed and therefore that $\eta > 0$, then β must always be smaller than $min(z)$. Therefore the initial guess for β is,

$$\hat{\beta}_0 = min(z). \tag{9}$$

The estimator provides reasonable results, as expected because $\min(z) = \min(\log(y+1)+\beta) = \log(1)+\beta = \beta$.

A set of Monte Carlo tests was performed using uniform synthetic data in order to assess the performance of the estimators for α and β. A set of 150 images with 128 $\times$ 128 pixels, corrupted with Rayleigh noise, compressed and shifted with several pairs of parameters (α, β) was used. For each pair of parameters (α, β), the estimators $(\hat{\alpha}, \hat{\beta})$ were computed for each image and their mean and standard deviation (SD) were displayed. The results are shown in Fig. 4 a). Fig. 4 a) shows the estimations of α for several values of the parameter $\psi = (500, 1000, 2000, 3000, 5000, 7000)$. The estimator of α appears to be unbiased and insensitive to the parameter ψ. The SD of $\hat{\alpha}$ increases with the ψ parameter, but its maximum value is small, which indicates it is an efficient estimator. Fig. 4 b) shows a similar experiment but now to assess the performance of $\hat{\beta}_0$, given by (9). The results prove that there exists a considerable dependence of $\hat{\beta}_0$ with respect to α. However, good results are obtained for low α values. It is also observed that the SD of $\hat{\beta}_0$ grows with α.

As shown in [8] a non biased estimator for β is

$$\hat{\beta} = E(\beta_0) - \alpha A, \tag{10}$$

where $A = \int_0^\infty \frac{e^{-\frac{\tau^2}{2\psi'}}}{\tau+1} d\tau$ depends on $\psi' = \psi/N$. (10) shows that the β_0 estimator is biased and greater than $\hat{\beta}$. The coefficient A (see Fig. 5) depends on the amount of available data N. As the

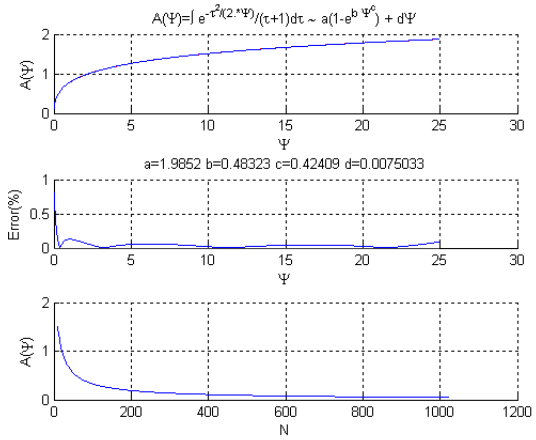

Fig. 5. Numerical computation of the coefficient $A(\psi)$. The plot on the bottom clearly shows that A becomes very small as N grows.

amount of data grows, $\psi' \to 0$, and $A \to 0$. Therefore, the estimator β_0 approximates β when N increases.

The estimator (10), on the contrary, is unbiased but it depends on the unknown ψ. To circumvent this difficulty let us combine (10) and (6),

$$\mu_z - min(z) = \frac{\alpha}{2} log(2\psi) - \frac{\alpha\gamma}{2} - \alpha A(\psi). \quad (11)$$

which is equivalent to the following equation depending on ψ

$$F(\psi) = \frac{\alpha}{2} log(2\psi) - \frac{\alpha\gamma}{2} - \alpha A(\psi) - \mu_z + min(z) = 0. \quad (12)$$

An estimation of ψ can be obtained from (12) using the Newton-Raphson method, $\psi^{t+1} = \psi^t - \frac{F(\psi)}{F'(\psi)}$ to be used in (6) in order to obtain $\hat{\beta}$,

$$\hat{\beta} = \mu_z - \frac{\hat{\alpha}}{2}(log(2\hat{\psi}) - \gamma). \quad (13)$$

IV. Experimental Results

In this section we present three types of experiments,

1) Testing the performance of the estimators of α and β for homogeneous synthetic data;
2) Recovering the RF signal from a set of real medical images;
3) Testing if the proposed compression law is valid by using the Kolmogorov-Smirnov and the Kullback-Leibler conformity tests in synthetic and real data.

Fig. 6 (top) displays a synthetic image, created with random Rayleigh-distributed pixels. The histogram of the pixel intensity is shown below. Fig. 6 (middle) shows the logarithmic compressed image obtained by applying the nonlinear transformation given by (2) to the pixels of the original image, with parameters $\alpha = 10$ and $\beta = 20$. As expected, the histogram is no longer Rayleigh. Fig. 6 (bottom) shows the estimated image using the log-compressed model. As it can be observed, the estimated image is similar to the original one, with both histograms almost identical, proving the ability of the algorithm to recover the original data.

Fig. 7 displays the mean and SD of $\hat{\alpha}$ and $\hat{\beta}$ as a function of the true parameters. This results were obtained by 50 runs of Monte Carlo tests to assess the performance of the estimators for each pair of true α and β parameters. It is experimentally shown (see Fig.7 a)) that the μ_α estimator is unbiased and presents a small variance, independent on β. The SD of $\hat{\alpha}$ increases linearly with the true

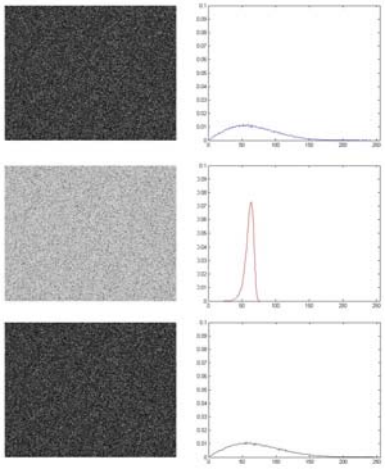

Fig. 6. Images and histograms from restoration with the logarithmic compression law proposed and using estimates of α and β. a) original (Rayleigh distributed), b) observed (log compressed) and c) decompressed.

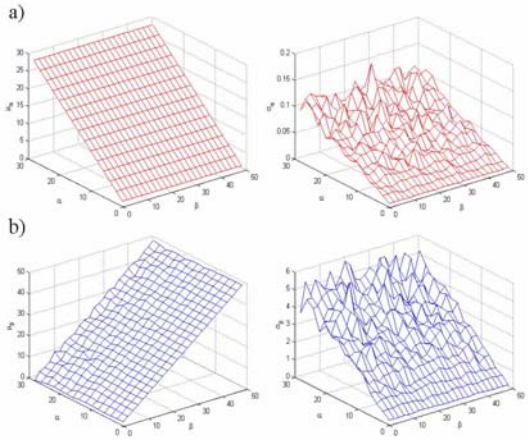

Fig. 7. Mean and SD of the estimators $\hat{\alpha}$ and $\hat{\beta}$ for $\psi = 5000$. Results were obtained with 50 Monte Carlo tests with uniform images with logarithmic compression Rayleigh distribution and different pairs of α and β.

parameter α presenting, however, small values for the whole range tested. The estimator of $\hat{\beta}$ (see Fig.Fig.7 b)), is also unbiased with a SD that increases with α.

Restoration results using medical data are displayed in Fig.8. Fig. 8 a)-c) (top) show observed US images and respective histograms of the thyroid, gall bladder and carotid artery. It is observed that this histograms follow approximately a log compressed Rayleigh distribution and Fig. 8 a)-c) (bottom) shows the estimations of the unobserved RF signal. The histograms of these images clearly show that the decompression function using the estimated parameters α and β makes it possible to recover the original RF signal.

In a final experiment, conformity tests were applied to the US observed images for two different hypotheses: Rayleigh and Fisher-Tippett distributions. To measure the level of fitness between the experimental and theoretical curves, given by the respective PDFs it was used (i) the level of confidence on the null hypothesis, H_0, given by the Kolmogorov-Smirnov statistical test and (ii) the Kullback-Leibler divergence (also called relative entropy).

Conformity test using the original ultrasound data and the estimated Rayleigh and Fisher-Tippet distributions were performed in order to assess which model better represents the observed data.

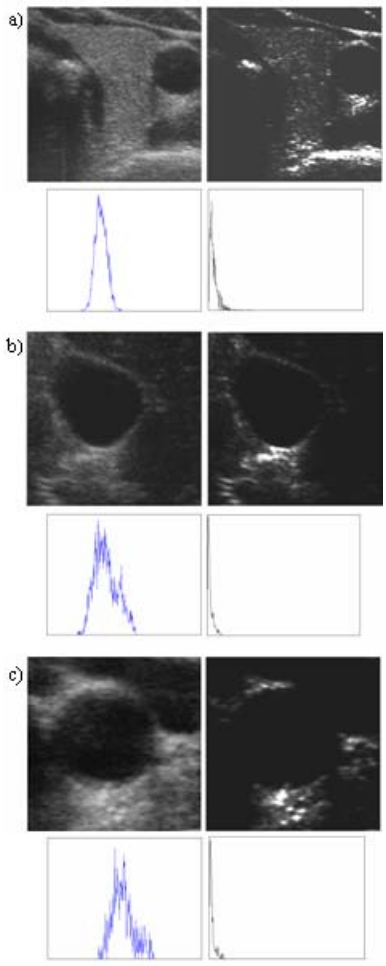

Fig. 8. Restoration results (Original and decompressed images and histograms) using medical data of the tyroid (a), gall bladder (b) and carotid artery (c).

Considering the Kolmogorov-Smirnov conformity statistical test, $P_e = 1 - P_{H_0}$ is the probability of rejection of the null hypothesis, H_0, which is the hypothesis of the data have been generated by the Rayleigh/Fisher-Tippett distributions. Here, $P_{H_0} = Q_{KS}(\lambda)$, $Q_{KS}(\lambda) = 2\sum_{j=1}^{\infty}(-1)^{j-1}e^{-2j^2\lambda^2}$, $\lambda = (\sqrt{(N)} + 0.12 + \frac{0.11}{\sqrt{N}})D$, N is the number of data points and $D = max|c(n) - ch(n)|$, $ch(n)$ are the cumulative probability functions of the Rayleigh/Fisher-Tippett distributions and $c(n)$ is the histogram of the observed image.

The Kullback-Leibler entropy distance is given by, $d = \sum_n p(n) log(\frac{p(n)}{h(n)})$. Here, $p(n)$ are the Rayleigh/Fisher-Tippett PDFs and $h(n)$ is the histogram of the observed image.

Recall that the PDF of the Rayleigh distribution is given by,

$$p(x) = \frac{x}{\psi_R} e^{-\frac{x^2}{2\psi_R}}, \tag{14}$$

where ψ_R is given by the maximum likelihood (ML) estimate, $\psi_R = \frac{1}{2n}\sum y^2$. The PDF of the Fisher-Tippett distribution depends on three parameters, $(\hat{\psi}, \hat{\alpha}, \hat{\beta})$, which were previously modeled and estimated. Table I displays the conformity test results for several synthetic images and also for medical data. From the synthetic images it is visible that there is a good fit between the experimental data and the proposed model (Fisher-Tippett) because the P_e from the Kolmogorov-Smirnov test is very low and the measure of entropy d of the Kullback-Leibler test is also low. For the real data, the fitness is reenforced with the Kullback-Leibler test. As an example, the PDFs of a test using an uniform synthetic image and a real US image are shown (see Fig. 9) to prove the good fit between the experimental data and the estimated model.

TABLE I

KOLMOGOROV-SMIRNOV AND KULLBACK-LEIBLER TESTS FOR SYNTHETIC AND REAL DATA. $(\hat{\alpha}, \hat{\beta})$ ESTIMATES ARE ALSO PROVIDED.

α	β	ψ	$\hat{\alpha}$	$\hat{\beta}$	Pe_R	Pe_{FT}	d_R	d_{FT}
10	10	$1e^3$	9.4848	9.4600	1	$2.9055e^{-04}$	2.8683	0.0030
20	20	$1e^3$	19.0217	20.8884	1	$2.1e^{-03}$	2.7706	0.0038
30	50	$1e^3$	28.3301	51.8945	1	$1.0e^{-03}$	2.6844	0.0038
20	20	$4e^3$	19.2550	21.3793	1	$3.3106e^{-09}$	3.0516	0.0020
10	10	$7e^3$	9.8086	11.6963	1	$2.2766e^{-012}$	3.2600	0.0012
	tyroid		15.0686	48.7970			5.4727	0.2985
	gallbladder		26.6844	28.5519			4.6665	0.6626
	carotid		26.9595	12.6349			4.1438	0.3654

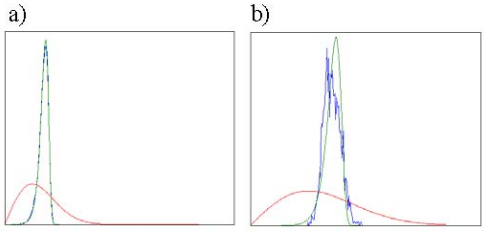

Fig. 9. Comparison of the PDFs of the experimental data (blue), Rayleigh (red) and estimated Fisher-Tippett (green) for uniform synthetic (a) and medical (b) images.

V. CONCLUSIONS

This paper presented an algorithm for modeling ultrasound images in order to recover the RF signal. The experimental results show that the proposed estimation model manages to retrieve the log-compression parameters and to estimate the original signal. The restoration results obtained with our model by considering the compression of the ultrasound data outperform the ones achieved by using the common Rayleigh distribution without compression.

REFERENCES

[1] C. Burckhardt, "Speckle in ultrasound b-mode scans," *IEEE Transations on Sonics and Ultrasonics*, vol. SU-25, no. 1, pp. 1–6, January 1978.

[2] O. V. Michailovich and A. Tannenbaum, "Despeckling of medical ultrasound images," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 53, no. 1, pp. 64–78, 2006. [Online]. Available: http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1588392

[3] J. Abbot and F. Thurstone, "Acoustic speckle: Theory and experimental analysis," *Ultrasound Imaging*, vol. 1, pp. 303–324, 1979.

[4] C. Sehgal, "Quantitative relationship between tissue composition and scattering of ultrasound ," *Acoustical Society of America Journal*, vol. 94, pp. 1944–1952, Oct. 1993.

[5] J. Seabra, J. Sanches, L. Pedro, and J. F. e Fernandes, "Carotid plaque 3d compound imaging and echo-morphology analysis: a bayesian approach," in *Conf Proc IEEE Eng Med Biol Soc*, Lyon, August 2007, pp. 763–6.

[6] R. C. E. Rignot, "Segmentation of polarimetric sunthetic aperture radar data," *IEEE Trans. Image Processing*, vol. 1, no. 1, pp. 281–300, 1992.

[7] T. K. Moon and W. C. Stirling, *Mathematical methods and algorithms for signal processing*. Prentice-Hall, 2000.

[8] J. Sanches and J. Marques, "Compensation of log-compressed images for 3-d ultrasound." *Ultrasound Med Biol*, vol. 29, no. 2, pp. 239–53, 2003.

[9] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th ed. Dover, 1964.