Predicting Generalized Contrast-to-Noise Ratios in Frame-Averaged Photoacoustic Images

Mardava R. Gubbi* and Muyinatu A. Lediju Bell*†‡
*Department of Electrical and Computer Engineering, Johns Hopkins University, Baltimore, MD
†Department of Computer Science, Johns Hopkins University, Baltimore, MD
‡Department of Biomedical Engineering, Johns Hopkins University, Baltimore, MD

2022 IEEE International Ultrasonics Symposium (IUS) | 978-1-6654-6657-8/22/$31.00 ©2022 IEEE | DOI: 10.1109/IUS54386.2022.9958666

***Abstract*—Our related journal article presents a theoretical framework to relate the performance of photoacoustic-based computer vision systems to the parameters of the underlying photoacoustic imaging systems. However, software-based adjustments were not previously included. In this paper, we extend the previous framework to include software-based noise reduction algorithms (e.g., frame averaging). We derive expressions for generalized contrast-to-noise ratio (gCNR) predictions in frame-averaged photoacoustic images. These gCNR predictions are validated against histogram-based gCNR measurements, with mean absolute errors of 4.4×10^{-3} and 7.7×10^{-3} in the phantom and *ex vivo* environments, respectively. This extension reduced the mean absolute prediction error by 7.74% and 27.15% when compared to predictions with the previous photoacoustic framework applied to frame-averaged photoacoustic images from phantom and *ex vivo* data, respectively. Results demonstrate that our extended framework more accurately predicts target detectability in frame-averaged photoacoustic images, which is beneficial for photoacoustic-based computer vision systems.**

***Index Terms*—photoacoustic imaging, image processing**

I. Introduction

THE integration of computer vision, robotics, and photoacoustic imaging for medical applications has garnered increasing interest in recent years. One problem of interest involves predicting the performance of computer vision-based tasks with photoacoustic images as inputs [1]–[5]. Our recently published journal paper presented a theoretical framework which leveraged the generalized contrast-to-noise ratio (gCNR) [6], [7] to relate computer vision-based system performance to the parameter values of the underlying photoacoustic imaging system [8].

This framework was initially limited to analyzing the impact of hardware-based system parameters (e.g., laser energy, channel SNR) on image quality and system performance [8]. However, software algorithms (e.g., beamforming, frame averaging) are additionally implemented to improve image quality and interpretability. These algorithms are known to affect the target and background statistics in the resulting images, which in turn impacts the resulting gCNR and achievable system performance. Thus, understanding the impact of software components will further expand our ability to design and optimize photoacoustic imaging systems.

In this paper, we extend our previously presented theoretical framework to analyze the effect of software algorithms on photoacoustic image quality. In particular, we explore the application of frame averaging as a software-based noise reduction technique. We derive the effect of the frame averaging technique on the statistics of the target and background regions of the image, and provide mathematical expressions for gCNR predictions in frame-averaged photoacoustic images. Finally, we demonstrate the improved ability of this expanded framework over the previous ultrasound-based [6], and photoacoustic-based [8] frameworks to predict gCNR in frame-averaged photoacoustic images.

II. Theory

A. General Framework

Fig. 1 shows our framework, expanded to include software-based techniques as a parameter of the imaging system. The framework consists of four layers (i.e., system parameters, target and background power distributions, image quality metrics, and computer vision-based task performance). The solid black lines denote relationships investigated in this paper to derive theoretical expressions for gCNR in frame-averaged photoacoustic images, while the dashed gray lines represent relationships outside of the scope of this paper.

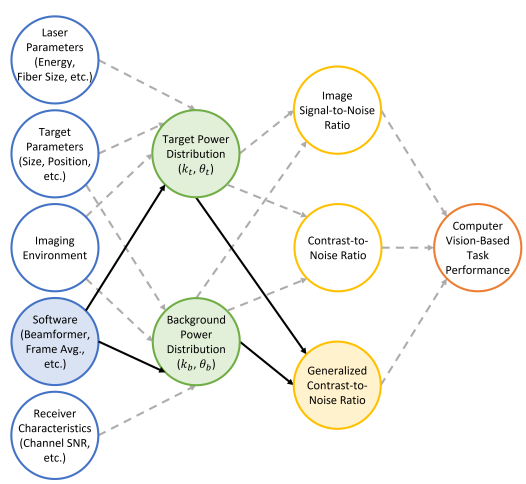

Fig. 1. Directed graph illustrating relationships among the system parameters, signal power distributions, image quality metrics, and performance of computer vision tasks. The solid black lines denote relationships reported in this paper. The dashed gray lines denote relationships which are known to exist, but lie outside the scope of this paper. k_t = shape of target power distribution; θ_t = scale of target power distribution; k_b = shape of background power distribution; θ_b = scale of background power distribution.

B. Modeling gCNR in Frame Averaged Photoacoustic Images

1) Approach Overview: The gCNR metric was introduced as a normalized measurement of the highest achievable performance of an optimal two-class classifier on a given photoacoustic image [6], [7], represented as,

$$\text{gCNR} = 1 - \int_{x=0}^{\infty} \min\{p_t(x), p_b(x)\}dx, \quad (1)$$

where x is the pixel power, and p_t and p_b are the probability density functions (PDF) of the target and background, respectively. The performance of this classifier depends on the target and background power statistics in the photoacoustic image. Therefore, to predict the gCNR of a frame-averaged photoacoustic image, we first model the target and background power statistics (Section II-B2), then determine decision boundaries and predicted gCNR from the modeled statistics (Section II-B3), similar our previously implemented approach [8].

2) Target and Background Power Statistics: We assume that N frames of photoacoustic channel data are acquired from a stationary target in the set S_{ch}:

$$S_{ch} = \{C_1, C_2, ..., C_i, ..., C_N\}, \quad (2)$$

where C_i is the i^{th} frame of channel data acquired in the set S_{ch}. N channel data frames in S_{ch} are averaged prior to delay-and-sum (DAS) beamforming (represented as operator f_D) to form the frame-averaged photoacoustic image, Y_{avg}:

$$Y_{avg} = \sum_{i=1}^{N} \frac{Y_i}{N}, \quad (3)$$

where $Y_i = f_D(C_i)$ is the image reconstructed using DAS beamforming from the individual channel data frame C_i.

We utilize the gamma distribution to model the target power statistics in Y_{avg} with probability density function (PDF) p_t and parameters k_t and θ_t similar to our previously presented framework [8]. In addition, we model the background power statistics using the gamma distribution with PDF p_b and parameters k_b and θ_b.

3) Computing Decision Boundaries: With the target and background power statistics presented in Section II-B2, we equate the target and background PDFs to obtain the following expression for the decision boundaries ϵ_m of the optimal classifier,

$$e^{\left(\frac{1}{\theta_t} - \frac{1}{\theta_b}\right)\epsilon_m} = \left(\frac{\theta_b^{k_b}\Gamma(k_b)}{\theta_t^{k_t}\Gamma(k_t)}\right)\epsilon_m^{k_t - k_b}. \quad (4)$$

Eq. (4) can be satisfied with four possible cases derived from the values of k_t, θ_t, k_b, and θ_b.

Case 1: $k_t \neq k_b$ and $\theta_t \neq \theta_b$. In this first case, there are up to two points of intersection between the target and background power distributions, given by:

$$\epsilon_m = -\frac{c}{a} \times \text{LambertW}\left(\left(-\frac{a}{c}\right)\left(\frac{1}{b}\right)^{\frac{1}{c}}\right). \quad (5)$$

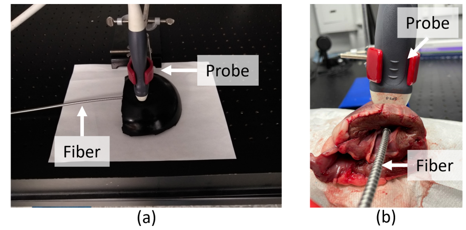

Fig. 2. Experimental setups showing a 2 mm-diameter optical fiber bundle inserted into (a) a plastisol phantom and (b) an *ex vivo* caprine heart. Both experimental setups were imaged with an Alpinion SP1-5 probe.

where $a = \frac{1}{\theta_t} - \frac{1}{\theta_b}$, $b = \left(\theta_b^{k_b}\Gamma(k_b)\right) / \left(\theta_t^{k_t}\Gamma(k_t)\right)$, and $c = k_t - k_b$. This case is similar to Case 1 in [8] and leverages the same method to compute the numerical values of the decision boundaries. However, the gamma-distributed model of the background power is responsible for differences in the values of a, b, and c when compared to Case 1 in [8], which employs an exponential background power distribution.

Case 2: $k_t \neq k_b$ and $\theta_t = \theta_b$. In this second case, with $k_t \geq 0$, $\theta_t \geq 0$, and $k_b \geq 0$, Eq. (4) is satisfied by the expression,

$$\epsilon_0 = \theta_t \times \left(\frac{\Gamma(k_t)}{\Gamma(k_b)}\right)^{\frac{1}{k_t - k_b}}. \quad (6)$$

The single non-negative value of ϵ_m which satisfies Eq. (6) forms the decision boundary in Case 2.

Case 3: $k_t = k_b$ and $\theta_t \neq \theta_b$. In this third case, Eq. (4) is satisfied by the expression,

$$\epsilon_0 = k_t\theta_t\theta_b\left(\frac{\ln\theta_b - \ln\theta_t}{\theta_b - \theta_t}\right), \quad (7)$$

which yields the single decision boundary for Case 3.

Case 4: $k_t = k_b$ and $\theta_t = \theta_b$. The fourth case occurs in photoacoustic images with sufficiently low channel SNR such that the target and background regions are indistinguishable, resulting in identical PDFs for the target and background power. This case is similar to Case 4 in [8], with no optimal decision boundaries for classification.

Once the decision boundaries of the optimal classifier were computed, the strategy presented in Section II-B5 of [8] was implemented to compute gCNR predictions for frame-averaged photoacoustic images.

III. Methods

To investigate the relationship between gCNR and frame averaging, we interfaced one end of a 2 mm-diameter fiber bundle to an LS-series pulsed laser diode (PLD) (Laser Components, Bedford, NH, USA). The other end was inserted into a plastisol phantom and an *ex vivo* caprine heart as shown in Fig. 2. An Alpinion SP1-5 probe (Alpinion, Seoul, South Korea) was placed in contact with each imaging environment. The probe was interfaced to an Alpinion E-CUBE 12R ultrasound scanner. The PLD was pulsed at a wavelength of 905 nm, a fixed repetition rate of 20 Hz, and 99 and

88 unique laser energy levels in the phantom and *ex vivo* environments, respectively. At each laser energy level, 100 frames of photoacoustic data were acquired by the ultrasound system. The corresponding laser energy values were then measured using an EnergyMax laser energy meter (Coherent, Inc. Santa Clara, CA, USA).

The acquired channel data frames were averaged and photoacoustic images were reconstructed from the frame-averaged channel data using DAS beamforming. Circular regions of interest (ROI) were manually selected in the target and background of each resulting image. The diameters of these ROIs were 1.6 mm and 0.8 mm for the phantom and *ex vivo* images, respectively. These diameters were chosen to match the sizes of the targets in each dataset. The target and background ROIs were located at the same depth, with a lateral offset of 5 mm between them in each image. This offset was selected to minimize the overlap between signal and noise while ensuring that the ROIs were as close to each other as possible. Measurements of the gCNR metric were obtained in each image using the histogram-based approach described in [8]. In addition, gCNR was predicted using the mathematical expressions derived in Section II. The mean absolute error (MAE) Δg_{FA} was computed between the gCNR predictions and measurements. The gCNR metric was also predicted using the photoacoustic and ultrasound frameworks presented in [8] and [6], respectively. The corresponding MAE values Δg_{PA} and Δg_{US} were computed for each gCNR measurement and prediction using the previous photoacoustic [8] and ultrasound [6] frameworks , respectively. Finally, the relative improvement Δg of the extended framework over the previous photoacoustic framework was computed in the phantom and *ex vivo* environments.

IV. Results

Fig. 3 shows the photoacoustic images of the optical fiber in a plastisol phantom and *ex vivo* caprine heart, acquired with laser energies of 0.36±0.03 μJ and 0.37±0.02 μJ, respectively, with 1, 10, and 100 channel data frames averaged prior to DAS beamforming. Images of the phantom demonstrated an improvement in gCNR predictions from 0.68 to 0.94 to 100 with no, 10, and 100 frames averaged, respectively. Images of the *ex vivo* images were noiser with no frame averaging, producing a gCNR of 0.41. The noise variation in the selected background region increased then decreased when 10 and 100 channel data frames were averaged, resulting in gCNRs of 0.33 and 0.69, respectively. This gCNR improvement of 0.36 achieved with 100 frames averaged is comparable to the gCNR improvement of 0.32 achieved with 10 frames averaged in the phantom environment. These results indicate that the effectiveness of frame averaging varies among imaging environments, even for similar laser energies.

Fig. 4 shows the measured and predicted gCNRs obtained from frame-averaged phantom and *ex vivo* images as functions of the number of frames averaged, separated by laser energy levels. The following three distinct trends are observed. First, in Fig. 4(a), the laser energy levels were too low for frame

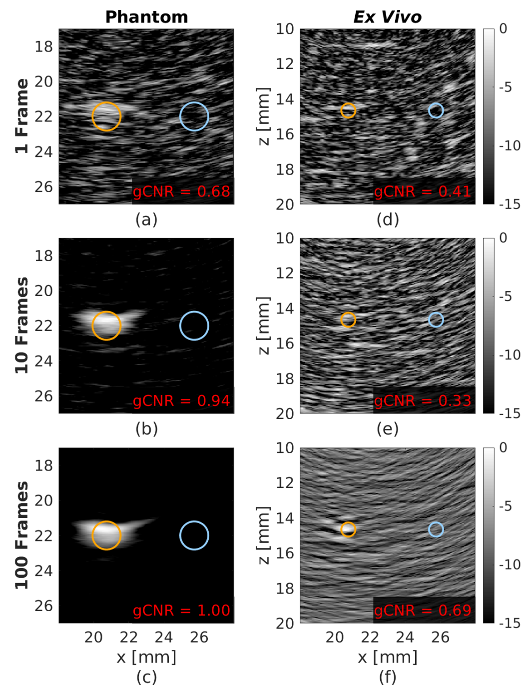

Fig. 3. Photoacoustic delay and sum (DAS) beamformed images of a 2 mm-diameter optical fiber bundle in a plastisol phantom and an *ex vivo* caprine heart, acquired with laser energies of 0.36±0.03 μJ and 0.37±0.02 μJ, respectively, with (a,d) 1, (b,e) 10, and (c,f) 100 channel data frames averaged prior to DAS beamforming. The target and background regions of interest are denoted by the orange and blue circles, respectively.

averaging to significantly improve gCNR in the phantom, with the mean predicted gCNR remaining between 0.11±0.05 and 0.25±0.09. Second, in Figs. 4(b) and 4(d), image quality improved as the number of frames averaged increased from 1 to 100, with the mean predicted gCNR monotonically increasing from 0.12±0.01 to 0.45±0.07 with phantom data and from 0.21±0.01 to 0.66±0.07 with the *ex vivo* data. Third, in Fig. 4(e), the improvement in gCNR achieved with 20 frames averaged (0.30±0.10 to 0.84±0.07) was higher then the improvement achieved by increasing the number of frames from 20 to 100 (0.84±0.07 to 0.94±0.04). Finally, in Figs. 4(c) and 4(f), the laser energy was sufficiently high to yield high gCNR values without frame averaging. Thus, the gCNR improvement as the number of frames increased from 1 to 100 was minimal in comparison to the previous trends (i.e., 0.98±0.05 to 0.99±0.01 and 0.86±0.18 to 1.00±0.01 in the phantom and *ex vivo* tissue, respectively). These results demonstrate that the dependence of frame averaging on both laser energy and imaging environment.

Table I reports the MAE between histogram-based gCNR measurements and gCNR predictions using the framework introduced herein, our previous photoacoustic framework [8], and the original ultrasound-based framework [6]. Our framework outperformed both previously introduced frameworks in both the phantom and *ex vivo* imaging environments. Thus, modeling the background power statistics with the gamma distribution rather than the exponential distribution utilized

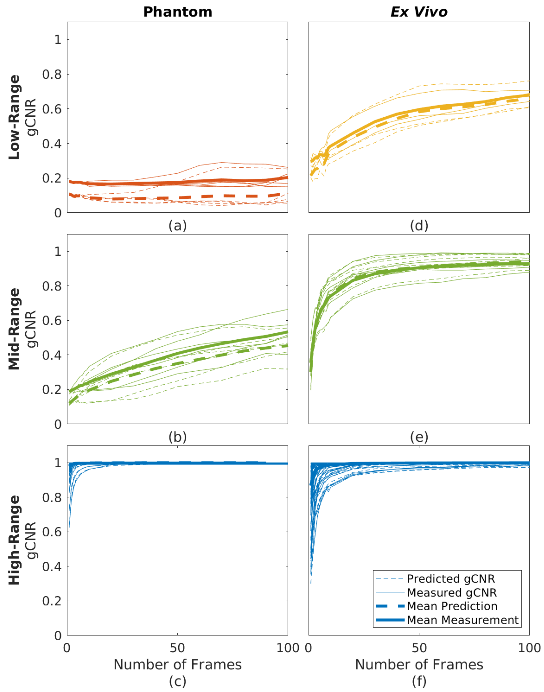

Fig. 4. gCNR as a function of the number of frames averaged prior to DAS beamforming of a 2 mm-diameter optical fiber bundle with one end interfaced to a Pulsed Laser Diode (PLD) and the other end inserted into (a,b,c) a plastisol phantom and (d,e,f) an *ex vivo* caprine heart. The plots are separated by input PLD settings with measured laser energies in the ranges (a) 0.07±0.03 μJ to 0.10±0.04 μJ, (b) 0.16±0.04 μJ to 0.27±0.03 μJ, (c) 0.36±0.03 μJ to 10.95±0.04 μJ, (d) 0.37±0.02 μJ to 0.62±0.02 μJ, (e) 0.79±0.02 μJ to 1.25±0.02 μJ, and (f) 1.4±0.02 μJ to 10.7±0.02 μJ.

TABLE I
MEAN ABSOLUTE ERRORS BETWEEN PREDICTED AND MEASURED GCNR VALUES IN PHANTOM AND EX VIVO DATASETS

	Δg_{FA}	Δg_{PA} [8]	Δg_{US} [6]	Δg (%)
Phantom	4.4×10^{-3}	4.7×10^{-3}	2.9×10^{-2}	7.74
Ex Vivo	7.7×10^{-3}	1.1×10^{-2}	9.6×10^{-2}	27.15

in the previous frameworks improves the accuracy of gCNR predictions for frame-averaged photoacoustic images.

V. DISCUSSION & CONCLUSIONS

This paper presents an expansion of our recently introduced theoretical gCNR framework to characterize the effect of software algorithms (i.e., frame averaging) on photoacoustic image quality. We presented theoretical derivations of mathematical expressions for gCNR predictions in frame-averaged photoacoustic images, and compared the performance of the extended framework with the previous ultrasound-based and hardware-based photoacoustic frameworks in phantom and *ex vivo* imaging environments. This framework has the potential to be extended to other software-based image quality improvement algorithms, allowing for the prediction of the performance of computer vision-based systems in an increasingly wide variety of clinical applications.

The results in Figs. 3 and 4 demonstrate the overall effectiveness of frame averaging to improve photoacoustic image quality. However, this improvement is tempered by lower frame rates, which may be concerning for the clinical applicability of real-time computer vision-based systems. In addition, the effectiveness of frame averaging was demonstrated to depend on factors such as the imaging environment (Fig. 3) and laser energy (Fig. 4). These results indicate that photoacoustic-based computer vision systems would benefit from allowing users to configure the number of frames averaged prior to image reconstruction, which would allow users to select a frame-averaging setting corresponding to the desired image quality and achievable frame rate for a given surgical or interventional application.

The extension of our previous photoacoustic framework reduced the errors between gCNR predictions and measurements in frame-averaged images, as shown in Table I. These results demonstrate the impact of statistical models on the accuracy of our theoretical gCNR predictions. This paper presents a working strategy to further expand our theoretical framework to accommodate additional software algorithms, beginning with a statistical analysis of the target and background statistics, followed by derivations of mathematical expressions of gCNR predictions in the resulting images, which is expected to increase the accuracy of predictions surrounding the performance of photoacoustic imaging systems.

ACKNOWLEDGMENTS

This work is supported by the NSF CAREER Award No. ECCS-1751522 and NSF SCH Award No. NSF IIS-2014088.

REFERENCES

[1] M. Graham, F. Assis, D. Allman, A. Wiacek, E. Gonzalez, M. Gubbi, J. Dong, H. Hou, S. Beck, J. Chrispin, and M. A. L. Bell, "In vivo demonstration of photoacoustic image guidance and robotic visual servoing for cardiac catheter-based interventions," *IEEE Transactions on Medical Imaging*, vol. 39, no. 4, pp. 1015–1029, 2019.
[2] M. R. Gubbi and M. A. L. Bell, "Deep learning-based photoacoustic visual servoing: Using outputs from raw sensor data as inputs to a robot controller," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 2021.
[3] M. A. L. Bell, "Photoacoustic imaging for surgical guidance: principles, applications, and outlook," *Journal of Applied Physics*, vol. 128, no. 6, p. 060904, 2020.
[4] A. Wiacek and M. A. L. Bell, "Photoacoustic-guided surgery from head to toe," *Biomedical Optics Express*, vol. 12, no. 4, pp. 2079–2117, 2021.
[5] S. E. Salcudean, H. Moradi, D. G. Black, and N. Navab, "Robot-assisted medical imaging: A review," *Proceedings of the IEEE*, vol. 110, no. 7, pp. 951–967, 2022.
[6] A. Rodriguez-Molares, O. M. H. Rindal, J. D'hooge, S.-E. Måsøy, A. Austeng, M. A. L. Bell, and H. Torp, "The generalized contrast-to-noise ratio: a formal definition for lesion detectability," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 67, no. 4, pp. 745–759, 2019.
[7] K. M. Kempski, M. T. Graham, M. R. Gubbi, T. Palmer, and M. A. L. Bell, "Application of the generalized contrast-to-noise ratio to assess photoacoustic image quality," *Biomedical Optics Express*, vol. 11, no. 7, pp. 3684–3698, 2020.
[8] M. R. Gubbi, E. A. Gonzalez, and M. A. L. Bell, "Theoretical framework to predict generalized contrast-to-noise ratios of photoacoustic images with applications to computer vision," *IEEE Transactions on Ultrasonics, Ferroelectrics, and Frequency Control*, vol. 69, no. 6, pp. 2098–2114, 2022.