Observer-driven texture analysis in CT imaging

Kupinski, Matthew, Garrett, Zachary, Fan, Jiahua

Matthew A. Kupinski, Zachary Garrett, Jiahua Fan, "Observer-driven texture analysis in CT imaging," Proc. SPIE 11316, Medical Imaging 2020: Image Perception, Observer Performance, and Technology Assessment, 1131610 (16 March 2020); doi: 10.1117/12.2549042

SPIE.

Event: SPIE Medical Imaging, 2020, Houston, Texas, United States

Observer-driven Texture Analysis in CT Imaging

Matthew A. Kupinski[a], Zachary Garrett[a] and Jiahua Fan[b]

[a]College of Optical Sciences, University of Arizona, Tucson, AZ, USA;
[b]GE Healthcare, Waukesha, WI, USA

ABSTRACT

We have implemented a technique for analyzing and characterizing the textures in medical images. This technique generates a list of characteristic textures and sorts them from most important to least important for the task of detecting a specific signal in the image. The effects of the human-visual system can be incorporated into this method through the use of an eye filter. The final set of sorted textures can be quickly utilized to analyze new sets of images and make comparison regarding performance on the same task. This analysis is based upon whether the new set of images contains textures that are similar or dissimilar to that of the original set of images. We present the method for analyzing and sorting textures based on how well signals can be distinguished. We also discuss the importance of the most "obscuring" textures that make signal-detection difficult. Results and comparisons of task performance are presented.

Keywords: Model observers, texture analysis, image quality, CT imaging

1. INTRODUCTION

It is well established that observer performance on signal-detection tasks partially depends on the texture of the background and noise statistics.[1–4] Different types of noise textures have varying effects on the observer's ability to perform the task. Texture is typically characterized with noise-power spectra (NPS) which plots the power (magnitude squared) of the average Fourier transform of the random backgrounds.[5] Unfortunately, NPS is constructed from a linear, shift-invariant theory of image formation and it does not directly translate to systems that are pixelated, shift variant, or non-linear. Modern CT reconstructions are voxelized, contain shift-variant noise features, and are non-linear in the image data even if a linear image-reconstruction algorithm is utilized due to an initial log transform of the projection data.[6] Thus, these images do not satisfy the constraints needed to apply NPS analysis.

An alternative to utilizing NPS to characterize texture, is to employ the image covariance matrix. Unfortunately, difficulties arise when using this approach as the covariance matrix is large and often needs to be inverted. We have implemented a practical technique for analyzing and characterizing the textures in medical images. This technique generates a list of characteristic textures and sorts them from most important to least important for the task of detecting a specific signal in the image. This set of sorted textures can be quickly utilized to analyze new sets of images and make comparison regarding performance on the same task. This analysis is based upon whether the new set of images contains textures that are similar or dissimilar to that of the original set of images.

2. METHODS

We define an image as the vector $\boldsymbol{g}_i$ where the index i indicates which image in a database of images. Practically, these are not entire CT volumes, but smaller ROIs within the CT reconstruction. This allows multiple $\boldsymbol{g}_i$s to be extracted from a single reconstruction although care must be taken since noise textures in reconstructed images are spatially variant and can have long-range correlations. The size of the ROI must be large enough to encompass the textures we wish to evaluate. For many of these studies, the ROI size was set at 80×80 (~ 2.5cm on a side) but larger ROI sizes can be accommodated. A random image is characterized very simply by its first- and second-order statistics $\overline{\boldsymbol{g}}$ and K_g, respectively. Textures are primarily a second-order statistics phenomena, and we assume that the mean is fairly constant over the size of the ROI. It is clear the higher-order statistics play a role in textures and observer performance; others have studied these effects.[7,8] However, we will focus on

For further author information, please contact M.A.K.: e-mail: mkupinski@optics.arizona.edu

Medical Imaging 2020: Image Perception, Observer Performance, and Technology Assessment,
edited by Frank W. Samuelson, Sian Taylor-Phillips, Proc. of SPIE Vol. 11316, 1131610
 · doi: 10.1117/12.2549042

a second-order statistical analysis of textures as voxel correlations. The covariance matrix can be decomposed using eigenanalysis into

$$K_g = \Phi D(\boldsymbol{\lambda})\Phi^\dagger, \tag{1}$$

where Φ is an orthonormal eigen-basis and the λ values represent the eigenvalues. The columns of the matrix Φ can be thought of as individual texture components, each one orthogonal from the others.

Typically, when analyzing textures with covariance decompositions, the textures are sorted based on eigenvalues which makes some intuitive sense. The texture (column of Φ) with the highest eigenvalue contributes the most to the overall K_g and is thus the most "important" texture. However, this cursory analysis doesn't account for the task. To account for task performance, we note that the Hotelling observer (i.e., the ideal linear observer) has a performance that is given by,

$$d^2 = \boldsymbol{s}^\dagger K_g^{-1} \boldsymbol{s}, \tag{2}$$

where $\boldsymbol{s}$ is the signal to detect and we utilize the inverse of the covariance matrix. Using our eigen-decomposition above, we can rewrite the above equation as,

$$d^2 = \boldsymbol{s}^\dagger \Phi D(\boldsymbol{\lambda})^{-1} \Phi^\dagger \boldsymbol{s}, \tag{3}$$

where the $D(\boldsymbol{\lambda})^{-1}$ is a trivial inverse because it is a diagonal matrix. An informative way of rewriting the Eqn. 3 is as,

$$d^2 = \sum_{j=1}^{J} (s_j^\Phi)^2/\lambda_j \tag{4}$$

where s_j^Φ is the jth component of the signal represented in the texture basis of Φ. Instead of sorting by eigenvalues, we choose to sort based on the terms in the above sum. Thus, the most important texture is the one that contributes the most to the Hotelling-observer task performance. Equation 4 is a generalization of the ideal-observer performance relationship shown in ICRU Report 54.[9]

2.1 Eye filter

The above analysis relates to the Hotelling observer which is the ideal linear observer. It is well known that human observers are not always ideal linear observers. Thus, to ensure that our sorting of textures is relevant to human-observer perception of textures, we include an eye filter which filters out frequencies in the signal that we do not use effectively in perception tasks. Thus, we rewrite the above summation as,

$$d^2 = \sum_{j=1}^{J} (\tilde{s}_j^\Phi)^2/\lambda_j \tag{5}$$

where $\tilde{s}_j$ is the signal after being filtered by the eye filter.

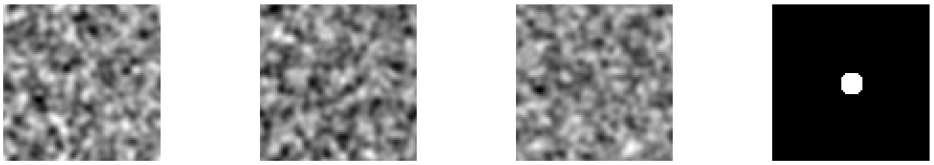
2.2 Using the texture basis

We sort the terms in the above summation and index them with l instead of j to denote that we have sorted the terms from largest to smallest. Then, observer performance can be approximated by using only the first L terms in the sorted summation, i.e.,

$$d^2 = \sum_{l=1}^{L} (\tilde{s}_l^\Phi)^2/\lambda_l. \tag{6}$$

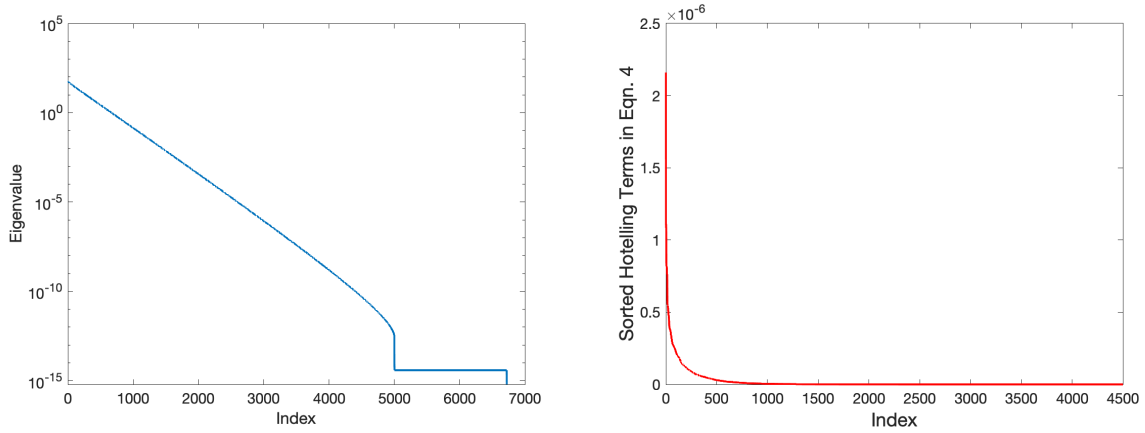
The eigenvectors associated with the top L terms represent a texture basis that contains the textures that were most "useful" for a human observer to detect the signal. We call this basis $\tilde{\Phi}$ where the dimension of this set of channels is $M \times L$ where M is the total number of pixels in the ROI.

For this method to be useful, we need to be able to use $\tilde{\Phi}$ in a meaningful way to characterize new sets of images for the same task. It should be noted that for the original set of images $\boldsymbol{g}_i$, that the covariance of



(a) (b) (c) (d)

Figure 1. (a)-(c) Example images used in the simulation to validate the method. (d) The signal to be detected.



(a) (b)

Figure 2. (a) The calculated eigenspectrum on a log scale. (b) The sorted terms that make up the Hotelling observer performance. Each term is associated with a separate and orthogonal texture. Also note that the Hotelling performance (sum of the terms in (b)) can be truncated without loss of accuracy.

$\tilde{\Phi}^\dagger \boldsymbol{g}$ is given by a diagnonal matrix $D(\tilde{\boldsymbol{\lambda}})$ where $\tilde{\boldsymbol{\lambda}}$ denotes the L eigenvalues sorted as described in the previous section. If we define $\boldsymbol{v} = \tilde{\Phi}^\dagger \boldsymbol{g}$, then $\boldsymbol{v}$ is an L vector representing the amount of the top L textures present in an individual image. The overall variance attributed to these texture components is simply

$$\sigma^2 = \mathrm{tr}\left[K_v\right] = \sum_{l=1}^{L} \lambda_l. \tag{7}$$

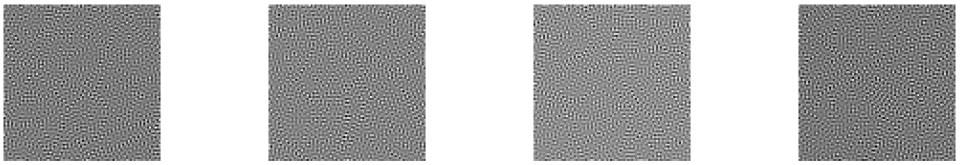
If we have a new set of images $\boldsymbol{g}'_j$ that have different texture characteristics, we can analyze this set of images by defining $\boldsymbol{v}'_j = \tilde{\Phi}^\dagger \boldsymbol{g}'_j$. The overall variance can be approximated by,

$$\sigma'^2 \approx \mathrm{tr}\left[K_{v'}\right]. \tag{8}$$

The above expression is not exact because with this new set of images we are not guaranteed to have uncorrelated channel outputs since the $\tilde{\Phi}$ basis was derived from a set of images with different statistical characteristics.

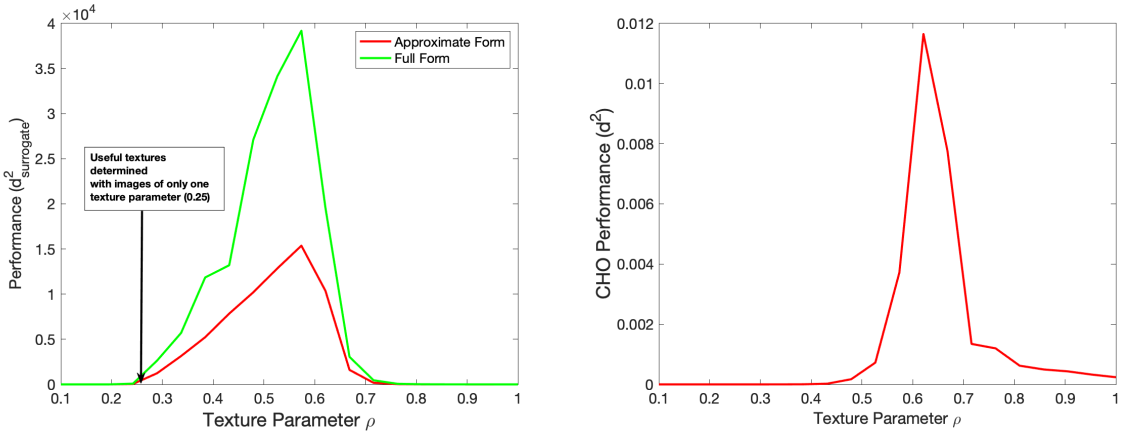
It is straightforward to use $K_{v'}$ to compute the surrogate figure of merit similar to the expression shown in Eqn. 6. However, we can no longer assume a summation formulation because the channel outputs are no longer uncorrelated. Furthermore, this surrogate figure of merit will not have an absolute value that is correct for a quantitative measure of image quality since we are truncating the expression and using a finite number of channels. However, the trends will follow that of human observers and it can thus be used for optimization and comparison on the same signal. This figure of merit can be computed using the covariance matrix $K_{v'}$ described above as well as the signal, processed by the eye filter and projected into the $\tilde{\Phi}$ basis – this is the signal that we already computed in Equation (6) and thus does not require any additional computational effort. The overall performance can be written as,

$$d^2_{\text{surrogate}} = [\tilde{\boldsymbol{s}}^{\tilde{\Phi}}]^\dagger K_{v'}^{-1} [\tilde{\boldsymbol{s}}^{\tilde{\Phi}}] \approx \sum_{l=1}^{L} \frac{[\tilde{\boldsymbol{s}}_l^{\tilde{\Phi}}]^2}{[K_{v'}]_{ll}}, \tag{9}$$



(a) (b) (c) (d)

Figure 3. The top four textures that contribute to the observer performance. These represent the textures that are easily distinguished from the signal. Because the signal is fairly large, these textures are simply the high-frequency white noise that was simulated.



(a) (b)

Figure 4. (a) A plot of both methods of computing the surrogate figure of merit shown in Eqn. 9 for image sets of varying background textures. (b) A plot of the Hotelling observer performance over the same background textures. The peak of all methods occurring at around the same texture-parameter value indicates that the method is a reasonable figure of merit and thus is a reasonable method for sorting the various texture components in the example images.

where the final approximation comes from assuming the correlation terms in $K_{v'}$ are negligible. We will evaluate both expressions shown in Eqn. 9 as a surrogate figure of merit.

3. SIMULATIONS

We validate the method using simulations of Gaussian textures. To accomplish this, we generated 1000 signal-absent ROIs (80×80 pixels, 2.5cm×2.5cm) with a constant mean and known covariance matrix. The covariance matrix is parameterized by a single parameter which we will call the correlation length which was set to $\rho = 0.25$pixels for the generation of the training images. The signal to be detected is generated separately since the technique does not require signal-present images. We utilized a pillbox signal with a diameter of 3.6mm. Example images and the signal used are shown in Fig. 1. The results of the SVD resulted in an eigenspectrum and corresponding texture (eigenvectors). The calculated eigenspectrum is shown in Fig. 2a and we see a characteristic rapid falloff in the spectrum values. The corresponding sorted spectrum (see Eqn. 6) is shown in Fig. 2b. Finally, the top four textures associated with observer performance are shown in Fig. 3.

We tested the application of these optimal textures to new sets of images using Eqn. 9. To accomplish this we generated many sets of signal-absent images with texture parameters ρ ranging from 0.1 to 1.0 (0.25 was used to determine the set of useful textures that contribute to signal detectability). We then applied Eqn. 9 to these new sets of images and compared the results to the Channelized Hotelling observer (CHO) with dense difference-of-Gaussian channels. We chose this comparison because the CHO with this channel set has been shown to mimic human-observer performance but is computed independently of the calculations we present here. Results are shown in Fig. 4 which shows that the peak performance of both the Hotelling observer and our surrogate figure of merit occur at approximately the same location.

We tested the repeatability of this result by rerunning these simulations over 100 times with different sets of images. The results remained consistent with the peaks of the CHO and the surrogate figure of merit within a ρ of 0.05. Note that the magnitudes of the surrogate figure of merit did change substantially from run to run but we do not expect this method to have an absolute value near the actual performance.

4. DISCUSSION

We have presented a method for sorting textures in images based on how they contribute to overall signal detectability. Textures that are easily distinguished from the signal and those that obscure the signal are useful for analyzing new sets of images for the task of detecting the same signal. This analysis represents a method that can be quickly applied to small sets of images to obtain a surrogate figure of merit and can potentially be used to help optimize image-reconstruction parameters in different CT algorithms. Initial results show the potential utility of the method. Future analysis on real image data will be presented in a separate paper.

Acknowledgements

The authors acknowledge Eric Clarkson for his helpful comments and careful analysis of the methods.

REFERENCES

[1] Wagner, R. F., Insana, M. F., Brown, D. G., Garra, B. S., and Jennings, R. J., "Texture discrimination: radiologist, machine and man," *Vision: Coding and efficiency* , 310–318 (1990).

[2] Wagner, R. F., Insana, M. F., and Brown, D. G., "Progress in signal and texture discrimination in medical imaging," in [*Application of Optical Instrumentation in Medicine XIII*], **535**, 57–64, International Society for Optics and Photonics (1985).

[3] Burgess, A. E., "Visual signal detection with two-component noise: low-pass spectrum effects," *JOSA A* **16**(3), 694–704 (1999).

[4] Solomon, J., Bochud, F., and Samei, E., "Design of anthropomorphic textured phantoms for ct performance evaluation," in [*Medical Imaging 2014: Physics of Medical Imaging*], **9033**, 90331U, International Society for Optics and Photonics (2014).

[5] Riederer, S. J., Pelc, N. J., and Chesler, D. A., "The noise power spectrum in computed x-ray tomography," *Physics in Medicine & Biology* **23**(3), 446 (1978).

[6] Razifar, P., Sandström, M., Schnieder, H., Långström, B., Maripuu, E., Bengtsson, E., and Bergström, M., "Noise correlation in PET, CT, SPECT and PET/CT data evaluated using autocorrelation function: a phantom study on data, reconstructed using FBP and OSEM," *BMC medical imaging* **5**(1), 5 (2005).

[7] Abbey, C. K., Bakic, P. R., Pokrajac, D. D., Maidment, A. D., Eckstein, M. P., and Boone, J. M., "Evaluation of non-gaussian statistical properties in virtual breast phantoms," *Journal of Medical Imaging* **6**(2), 025502 (2019).

[8] Xu, Y., Raj, A., and Victor, J. D., "Systematic differences between perceptually-relevant image statistics of brain mri and natural images," *Frontiers in neuroinformatics* **13**, 46 (2019).

[9] Sharp, P., Barber, D. C., Brown, D. G., Burgess, A. E., Metz, C. E., Myers, K. J., Taylor, C. J., Wagner, R. F., Brooks, R., Hill, C. R., Kuhl, D. E., Smith, M. A., Wells, P., and Worthington, B., "Report 54," *Journal of the International Commission on Radiation Units and Measurements* **os28**, NP–NP (04 2016).