ABSTRACT

Most statistical models for applications rely on the Gaussian assumption. Yet, in many realistic situations, the underlying variation or uncertainty is essentially non-Gaussian. In detection problems, for instance, the Gaussian assumption leads to false alarms in cases where the tail is a fatter one, such as in the case of the Laplace density function. In classification problems, the Gaussian model for variability may be too restrictive, and other models, such as the Generalized Gaussian density function, are more appropriate. We will present examples of such models as applied to applications with multiple images, and show performance in two applications: functional magnetic resonance imaging, and stem cell classification.

1. INTRODUCTION

Gaussian models are ubiquitous in statistical signal processing and hypothesis testing applications. The major justification given is the Central Limit Theorem (CLT) [8]. This theorem enables ready statistical characterization for finite sums of random variables. It states that the density function of the sum of random variables, normalized by the square root of the sample size, approaches the Gaussian density function as the sample size increases. The most common assumptions for the theorem are that the random variables be independent, identically distributed, of zero mean, and of finite variance. Yet, in reality, mathematical tractability is the real attractive feature of this model; probability density functions (pdf’s) of sufficient statistics that are derived from detection tests based on the Gaussian assumption, such as the $\chi^2$, $t$, and $F$ statistics, are readily available. These pdf’s can be used to compute performance, including false alarm and missed detection probability tradeoffs.

For these reasons, the errors that accompany the Gaussian assumptions are, even when acknowledged, conveniently ignored. In reality, given a sample size, the Gaussian pdf provides a better approximation of the actual probability density function (pdf) of the sum in the central limit theorem around the center or the mode, but is considerably less accurate near the tails [9]. Outcomes of hypothesis testing based detection problems, however, depend on tail behavior. This is particularly true if the underlying pdf has a tail fatter than the Gaussian, as is the case with the Laplacian density. In such a case, the probability that the random variable takes larger values is higher than for Gaussian variables. The Gaussian approximation thus leads to higher false alarm or false positive rates, or equivalently, a higher probability of type I statistical error.

In this paper, we review results obtained for detection [3, 6] and retrieval/classification problems [7, 11, 12] using non-Gaussian models and demonstrate how they were used for enhanced performance in two motivating important medical imaging applications: functional magnetic resonance imaging (fMRI) [2, 4, 5], and non-invasive stem cell image based classification [11, 12, 14]. The underlying models for both cases are the Generalized Gaussian (GG) probability density functions; these comprise a unimodal and symmetric family of pdf’s that embrace a wide range of tail behavior. The Gaussian is but one member of this family. For functional MRI, the problem is a signal detection problem, and the GG pdf is used to model the noise. For the stem cell application, it is used to model the coefficients of the wavelet decomposition of an image of stem cell colony or nucleus texture. We provide performance results for both applications.

One traditional reason for the reluctance to use non-Gaussian detectors is computational tractability of non-Gaussian methods. This obstacle is not as severe now that computers are gaining speed by the day. But we will show that the GG based matched filter detectors we present here are made tractable, thanks to a lemma in [3]. Specifically, we review derived statistics based on underlying GG random variables that are Generalizations of the Gaussian based $\chi^2$ and $t$ or $F$ statistics.

The paper is organized as follows. In Section 2, we describe and discuss the family of Generalized Gaussian densities. In Section 4, we present the functional MRI application. In section 5, we discuss the classification of stem cell colonies and nuclei using a GG model. Section 6 presents concluding remarks.
The Generalized Gaussian(GG) density function for width factor $w = 1$ and various values of the shape parameter.

2. THE GENERALIZED GAUSSIAN DENSITIES

For a $K$ dimensional random vector, the Generalized Gaussian density function [10] is given by

$$f_p(x|m, \omega) = \left(\frac{p}{2\omega \Gamma(1/p)}\right)^K \exp\left(-\frac{\|x - m\|_p^p}{\omega}\right),$$

for $p \in (0, \infty)$, and where $\Gamma$ is the Gamma function given by

$$\Gamma(k) = \int_0^\infty t^{k-1} \exp(-t) \, dt.$$  

and for an arbitrary vector $y$, $\|y\|_p$ is defined as

$$\|y\|_p = \left(\sum_i |y_i|^p\right)^{1/p}$$

(2)

Here $m$, $\omega$, and $p$ are respectively the location, width factor and shape or decay parameters of the density function in (1).

In particular, the Laplacian and Gaussian density functions with standard deviation $\sigma$ are obtained when $(p, \omega) = (1, \sigma/\sqrt{2})$ and $(p, \omega) = (2, \sqrt{2}\sigma)$, respectively. The uniform density function is approximated by large values of $p$. More generally, varying $p$ allows us to take into account various pdf variability considerations as desired. Note that while the width factor $\omega$ and the location parameter $m$ may be estimated, we assume that the shape parameter $p$ is known. Methods for estimating the shape parameter are found, for instance, in [16, 17].

In order to gain more insight into the role of the parameter $p$, consider a set of measurements for a scalar variable plotted on the real line as shown in Figure 2. We are to estimate a location parameter based on maximum GG likelihood of the measurements. If $p$ is unknown, what effect will the choice of $p$ have on the estimate? As the figure shows, the location parameter estimate is very dependent on our choice of $p$. If $p < 1$, then the estimate will be located within the largest cluster of closely bunched up measurements. If $p = 1$, then the estimate will be exactly the median assuming the number of measurements is odd, or between the two middle ranked measurement values if it is even. If $p = 2$, then the estimate is the mean. Further, the larger the value of $p$, the more weight will outliers have. At the limit, as $p$ approaches infinity, then the estimate consists of the midpoint of the minimum and maximum of the measurements. Thus, while lower values of $p$ tend to produce an estimate that is robust to outliers, a higher value choice implicitly assumes that outliers are not an issue. This suggests that the shape parameter $p$ can be used as a design option. For instance, if outliers are a concern, then smaller values of $p$ may be assumed.

3. THE GENERALIZED GAUSSIAN MODELS IN ROBUST DETECTION

Consider the following signal model

$$x = s\theta + U\psi + v$$

(3)

where $x$ is a $N$ dimensional measurement vector, $s$ is a known one-dimensional subspace $R^N$, $\theta$ is an unknown scalar signal strength, $U$ is an unknown interference subspace, $\psi$ is the unknown interference vector, and $v$ is the noise vector. We will adopt the Generalized Gaussian model for the noise. We consider the following hypothesis test,

$$H_0 : \theta = 0; \psi \neq 0$$

(4)

$$H_1 : \theta \neq 0; \psi = 0$$

(5)

The above test is more general than conventional matched filter detector tests [15] in two ways: first, the noise is not Gaussian, and second it is derived from game theoretic principles [3]. It is robust to interference residing in unknown subspaces. In practice, only the components of the interference
residing in the subspace orthogonal to can be distinguished from the signal of interest. So, an implicit assumption is that \( U \) is simply the null space of \( s \). The case where the interference space is known is also treated in [3, 6].

When the noise width parameter \( w \) is unknown, the likelihood ratio test yields a statistic that is a generalization of the \( t \) or \( F \) statistic normally associated with Gaussian noise based **Constant False Alarm Rate (CFAR)** matched filter detectors. Thus, skipping details [3], we have

\[
t_p = \frac{\min_{\psi} \| x - U \psi \|_p}{\min_{\theta} \| x - s \theta \|_p} \quad (6)
\]

\[
t_p = \frac{\min_{\theta} \| x - s \theta \|_p}{\min_{\theta} \| x - s \theta \|_p} \quad (7)
\]

where \( q \) is such that \( 1/p + 1/q = 1 \) if \( p > 1 \), or \( q = \infty \) otherwise. Equation (6) is simply the generalized likelihood ratio for (5). As is, the ratio requires a numerical search for the one dimensional scalar \( \theta \), and another search in \( N - 1 \) dimensions for \( \psi \). As stated in a lemma in [3], the latter search is not needed, since the ratio can be reduced to (7). Thus, we have a CFAR matched subspace detector for an the entire class of Generalized Gaussian distribution that is computationally implementable in real time, since only a search for a scalar is required on line.

We also note that when \( p = 2 \), we have \( t_2 \) or the well known \( t \) statistic with degrees of freedom equaling the dimension of \( s \); this is the CFAR detector shown in [15]. Finally, robust detection in the context of dynamic plants has been addressed in [1, 13].

### 4. DETECTOR PERFORMANCE AND APPLICATION TO FUNCTIONAL MRI

**Performance.** The Generalized Gaussian detectors’ performance is shown in Figures 3 and 4. Figure 3 considers a situation where the true noise is Laplacian, and compares the performance of a Gaussian and a Laplacian detector. There is no interference present in this scenario. The figure plots the probability of detection vs. the probability of false alarm for a signal of small magnitude. Thus, for a probability of detection of 0.95, the Gaussian detector gives a false alarm probability of 0.15, while the Laplacian detector’s probability of false alarm is 3 times smaller, or 0.05; the results are a clear indication of the deleterious effect on performance of the widely held practice of assuming Gaussian noise and using Gaussian based detectors.

Figure 4 shows the effect of taking interference into account. The noise is Laplacian. Two scenarios are considered: interference absent, and interference present. Two detectors are considered, one that is robust to interference, the other is conventional (or optimal), meaning it assumes the ideal situation, and ignores the presence of interference.

In the idealized situation, where no interference is present, the conventional detector performs best (solid blue curve). However, the situation is unlikely to obtain. In the more realistic situation, where interference is present, the conventional detector’s performance suffers considerably (dashed blue curve). By contrast, the robust detector’s performance is similar under both scenarios, interference absent and present. The robust detector’s performance is therefore more predictable and consistent over a wider range of scenarios.

**Application to functional MRI.** In [2, 4, 5] we developed new detectors for a functional MRI application. This is a decision making challenge of a low signal to noise ratio and non-Gaussian noise environment. The objective is to determine which voxel in the brain responds to a particular stimulus signal. As this is a low signal to noise ratio (SNR) environment where fatter than Gaussian tail noise dominates, the conventional Gaussian detector generated false alarms and failed to identify responsive voxels (Top, Figure 5). By contrast, the Laplacian detector was sensitive to voxels responsive to the stimulus and robust to false alarms in the presence of interference.

Figure 5 shows that the conventional detector classifies an area outside the skull as a voxel that responds to the stimulus, though it is not inside the brain, while missing responsive voxels inside the brain. By contrast, the robust detector did not make such mistakes and had superior performance in terms of enhanced detection, meaning less missed detections, and minimal false alarms.

### 5. GENERALIZED GAUSSIAN MODELS FOR STEM CELL CLASSIFICATION

**Stem cell classification.** Texture classification is another phenomenon that displays non-Gaussian behavior. In [7], tex-
Nominal vs. robust performance of CFAR detectors

Fig. 4. Probability of detection vs. Probability of false alarm (ROC) performance comparison of Laplacian (solid, blue) and Gaussian (dashed, red) detectors.

tures were decomposed using wavelets, and it was shown that at each subband except the lowest one, the coefficients of the wavelet decomposition are randomly distributed with mean zero according to the Generalized Gaussian density function (1). Thus, for a given texture, two parameters for each wavelet band, the scale factor and the shape parameter, describe a texture.

Texture analysis is an approach we pursue for analyzing the level of pluripotency in stem cells with images of whole colonies and nuclei images [11, 12, 14]. A stem cell is pluripotent if it has not yet specialized, meaning it still has the ability to give rise to types of cells that develop into three germ layers: mesoderm, endoderm, and ectoderm. From these cell types, all the cells of the body arise. A cell is differentiated if it has specialized into one of the many types of the body’s cells. We are classifying here human embryonic stem cells (hESCs) in tissue culture.

Currently, there are two methods for classifying stem cells: the microscopist’s eye, and chemical tests. The first requires a trained specialist, and is therefore expensive. In addition to cost, speed and consistency are two limitations of manual classification. The second method, chemical tests, is accurate, but destructive, thus preventing the use of the cells for further research, therapy, etc.

Quantitative textural differences between pluripotent and differentiated stem cell colonies enable autonomous, non-invasive non-destructive accurate classification at high speeds. Figure 6 illustrates the textural difference of three classes of colony: pluripotent, differentiated, and an intermediate (middle) colony. The pluripotent colony is tight and of finer grain, and has circular, crisp boundaries. The differentiated colony is more diffused, with marshland borders [11, 14].

GG densities and Stem cell texture multi-resolution wavelet decomposition. The wavelet coefficients of different types of colonies have GG density functions with different parameters. The same is true for single cell nuclei textures [12]. Consider the problem of assigning a texture to one of two classes characterized by different Generalized Gaussian density parameters for their wavelet decomposition. Assume we have $B$ subbands and $S$ samples for this texture. Denote the associated random variables $x_{sb}, s = 1, \ldots, S$, and $b = 1, \ldots, B$. As in [7], we assume for simplicity that the subband decompositions are independently distributed. Let $X_s$ be the vector of random variables from each of these samples $s = 1, \ldots, S$. We formulate the following binary hypothesis test

$$H_0 : X_1, \ldots, X_b, \ldots, X_S \sim f_0 = \prod_{b=1}^{B} \prod_{s=1}^{S} f_{0b}(x_{sb})$$

$$H_1 : X_1, \ldots, X_b, \ldots, X_S \sim f_1 = \prod_{b=1}^{B} \prod_{s=1}^{S} f_{1b}(x_{sb})$$

where each of the densities $f_{0b}$ and $f_{1b}, b = 1, \ldots, B$ is a GG density function given in (1), with different width and
shape parameters. The log likelihood ratio for this test is given by [12]
\[ \Lambda(X_1, \ldots, X_S) = \chi^2_0 - \chi^2_1 \]  
(10)
where \( \chi^2_0, \chi^2_1 \) are GG generalization of the \( \chi^2 \) random variables.

Specifically, the random variable \( \chi^2 \) is the sum
\[ \chi^2 = \sum_{i=1}^{N} \chi^2_i = \sum_{i=1}^{N} \left| \frac{x_i}{w_i} \right|^{p_i} \]  
(11)
where the \( x_i \)'s have each a GG pdf with parameters \( w_i, p_i \). If all shape parameters take the value \( p = 2 \), we have \( \chi^2 \) random variables with the appropriate degree of freedom. The generalized \( \chi^2 \) pdf is given by
\[ f_{\chi^2}(x) = \frac{1}{\Gamma(\sum_{i=1}^{N} 1/p_i)} e^{-x\left| x \right|^{\sum_{i=1}^{N} 1/p_i}}, \quad x \geq 0 \]  
(13)
In addition to the likelihood statistic, it is also possible to use the Kullback-Leibler distance [7]. In [11, 12, 14], we classify stem cells using the above ideas. Figure 7 shows how differences between colonies, once quantitated using statistics from the density functions (8), (9), or (13), can be visualized. The multiresolution statistical texture analysis described above was conducted on three colonies such as those of Figure 6. The color scale is an indication of the relative statistical distance between the colonies. Thus, dark blue in the diagonal block indicates that each colony is similar to itself. The off-diagonal light blue blocks indicate the distance between the pluripotent colony and the colony in transition. The dark brown indicates the largest distance is between the pluripotent and the differentiated colony.

6. CONCLUDING REMARKS

We have demonstrated that non-Gaussian detection and classification methods are applicable to two classes of problems in medical and biological imaging: functional MRI and stem cell classification. Both applications relied on the Generalized Gaussian density model. We have reviewed the new detectors or statistics that can be derived from the Generalized Gaussian densities, and given sample results. In both applications, we have demonstrated that the methods are computationally tractable, and yield good performance.

7. REFERENCES


