Stochastic Resonance and Suboptimal Radar Target Classification

Ismail Jouny
ECE Dept., Lafayette College, Easton, PA, 18042

ABSTRACT

Stochastic resonance has received significant attention recently in the signal processing community with emphasis on signal detection. The basic notion is that the performance of some suboptimal detectors can be improved by adding independent noise to the measured (and already noise contaminated) observation. The notion of adding noise makes sense if the observation is the result of nonlinear processing, and there exist proven scenarios where the signal-to-noise ratio improves by adding independent noise. This paper reviews a set of parametric and nonparametric sub-optimal radar target classification systems and explores (via computer simulation) the impact of adding independent noise to the observation on the performance of such sub-optimal systems. Although noise is not added in an optimal fashion, it does have an impact on the probability of classification error. Real radar scattering data of commercial aircraft models is used in this study. The focus is on exploring scenarios where added noise may improve radar target classification.

Keywords: Stochastic resonance, radar target classification.

1. INTRODUCTION

Well established signal detection and classification systems that rely on optimal (risk minimizing) decision theoretic schemes are available for achieving optimal signal detection in various noise scenarios. Such optimal detectors require prior knowledge of the noise statistics and knowledge of the a priori probability of occurrence of each hypothesis. Unfortunately, often neither of these important features is known exactly. Although some of these parameters can be estimated, the estimation error does have an impact on detection performance. Furthermore, there are numerous scenarios where the notion of a decision theoretic optimal detector or classifier is simply unfeasible. The problem is worse when dealing M-ary hypotheses testing where decision spaces are far more complicated and performance statistics are more sensitive to model uncertainty. In addition to incomplete statistical knowledge, hardware and software complexities often prohibit relying on statistically optimal decision schemes. An attractive alternative, particularly when dealing with complex classification problems is to rely on suboptimal recognition system that may be parametric or non-parametric. Even if the classification is done in a decision-theoretic parametric manner, the assumptions made may render the classifier suboptimal.

Suboptimal detectors and classifiers are commonly used in several practical pattern recognition schemes including radar. One of the possible ways to improve classifier performance is to use stochastic resonance assuming that the classifier underlying features permit improvability. The idea that was explored by physicists [1] and later imported into signal processing and engineering [2-9] makes more sense when the observation is the result of nonlinear operation, where it has been shown that the signal-to-noise ratio may be improved by adding independent noise to the measured data. This concept may make more sense in image recognition where discerning images may improve by sampling adding certain types of noise to the already noisy image.

2. STOCHASTIC RESONANCE

Features used for detection and classification that are the output of nonlinear physical systems may become more discernable by adding independent noise [2-6]. The improvement is not always in terms of the signal-to-noise ratio (SNR), and the definition of SNR differs across disciplines and applications, but the improvement may be better observed in the probability of detection or classification, or in more than one aspect of pattern recognition performance. It is certainly the case that, in some image recognition problems, some patterns are more discernable if some form of
noise is added to the image. This improvement is not limited to images of objects and scripts but includes images of
landmines, mammograms, etc. Some researchers have shown improvement using added Gaussian noise, non-Gaussian
noise, and constant noise [2-9]. Others have described the optimal added noise for specific sub-optimal detectors and
measurement statistics.

The basic notion that the improvement is likely when the observation is the result of a nonlinear physical phenomenon
makes sense in radar because ultimately the radar interrogation process begins with a narrowband (single frequency) or
wideband (impulse type) waveform that may be corrupted by noise before it impinges on a target and energy is scattered
every direction, some of which is scattered back to a radar receiver. The scattered energy depends on the geometrical
shape of each target subcomponent and on the communications channel noise. This process is a nonlinear process that
produces the observation vector $X$. One of the models for scattering from a target is

$$X = \sum_{p=1}^{P} (j \omega)^{n_p} e^{j(\omega_p + \theta_p)} + w$$

Where $P$ is the number of scattering centers and $n$ determines the geometrical nature of each scattering center. Some are
dispersive, others are impulsive, and some has a filtering effect. The noise is represented by the vector $w$ which may or
may not be Gaussian. This observation vector (which is nonlinearly dependent on the number of scattering centers and
their geometrical attribute) is representative of a particular target at a specific aspect (azimuth) angle. Changes in
azimuth may yield very different backscatter response with possibly a different number of scattering centers. Yet, it is
the same target class that yields considerably different signatures. The nonlinearity of the scattering model and the
change with azimuth makes statistically optimal classifiers unattainable from a practical point of view. The alternative
is a set of suboptimal detectors that have been explored for radar target recognition purposes. So, stochastic resonance
which was originally developed in [1] may have an impact on classifying radar targets when associated with sub-optimal
classifiers. Exploring stochastic resonance in detection theory has been substantial in recent years but only focusing on
binary hypothesis testing problems [2,3]. The study in [4] is one of very few studies that explore stochastic resonance for
M-ary Hypothesis testing in a Minimax framework. One of the most important questions to answer before taking
advantage (if any) of stochastic resonance is what type of noise to add. Unfortunately, the answer depends on the
existing noise statistics, signal type, and the suboptimal detector or classifier being used. The classifiers explored below
are M-ary parametric and non-parametric suboptimal classifiers.

### 3. SUB-OPTIMAL RADAR CLASSIFICATION

#### 3.1 Maximum likelihood classifier

Assuming that the added noise is Gaussian, it is possible to model the probability density function of the radar
backscatter as multivariate Gaussian with $N$ independent observations ($N$ being the length of $X$), so the likelihood that
the vector $X$ belongs to target class $k$ at the $m$-th catalogued azimuth position

$$P(X / \omega_k, \theta_m) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma}} \exp[(x(n) - s_{m,k}(n))^2 / (2\sigma^2)]$$

Where the mixed Gaussian distribution is
\[ P(X / \omega_k) = \frac{1}{N_s} \sum_{m=1}^{N_s} P(X / \omega_k, \theta_m) P(\theta_m) \]

Is the likelihood function that the vector \( X \) belongs to target class \( k \). The classifier maximizes

\[ P(\omega_k / X) = \frac{P(X / \omega_k) P(\omega_k)}{p(X)} \]

over all possible targets \( k=1,\ldots, M \). Assuming equi-probable target occurrence, the classifier becomes a maximum likelihood classifier. This classifier is suboptimal because of the mixed Gaussian distribution over azimuth. The target azimuth may not be known, and only a finite number of catalogued returns at specific azimuth positions is usually available, so the mixed Gaussian distribution assuming knowledge of the a priori probability of each azimuth position renders the classifier suboptimal.

In assessing the performance of this classifier under additive independent noise \((Y=X+Q)\) where \( Q \) is independent iid noise vector, three possibilities were considered; first the additive noise is Gaussian, and second, constant noise with impulse like probability density function, and third assuming additive Rician distributed noise. The choices in the second case are based on the work reported in [4] and [5] although the systems considered in [4] use the Minimax approach and the work in [5] is for binary hypotheses detection type problem.

### 3.2 Nearest neighbor classifier

The nearest neighbor classifier is based on minimizing the Euclidean distance between the unknown target backscatter observation vector and that of each of the catalogued target backscatter responses at various recorded azimuth positions, i.e choose target \( k \) if

\[ d(X, \omega_{k,m}) = \min_{k} \{ \min_{m} (d(X, S_{k,m})) \} \quad \text{with} \]

\[ d(X, S_{k,m}) = (X - S_{k,m})^T (X - S_{k,m}) \]

where \( S \) represents the catalogued response for target class \( k \) at azimuth \( m \).

This classifier is optimum if the additive noise is Gaussian and the azimuth position is exactly known, otherwise it is considered a non-parametric decision-theoretic suboptimal classifier. The work reported in [3] generates some attributes of the optimal additive noise that may be used in assessing any possible improvements in the performance of the above classification scheme. Two possible noise scenarios were examined here, additive Gaussian noise and constant additive noise with impulse type pdf. Scenarios where the exact target zero-time reference position is known or unknown are being considered. The target reference position represents the zero time point when the target impulse response is displayed. The target impulse response can be obtained by taking the inverse Fourier transform of the recorded backscatter data, and it represents the returned energy as a function of the down range profile of the target which is proportional to the time of propagation along the target. Exact knowledge of the target zero-time reference implies knowledge of the phase of the recorded backscatter at each frequency. So, if the phase of the unknown target backscatter and that of the catalogue target are not exactly aligned then the probability density function is Rician

\[ P(X / \omega_k, \theta_m) = \prod_{n=1}^{N} \frac{x(n)}{\sigma^2} \exp\left(-\frac{(x(n) + s_{k,m}(n))}{2\sigma^2}\right) I_0\left(\frac{x(n) s_{k,m}(n)}{\sigma^2}\right) \]

and the nearest-neighbor classifier is unusable.
3.3 Correlation based classifier

Scenarios where the target exact zero-time reference is not exactly known lead to another sub-optimal classifier that is non-parametric in the sense that no information about the additive noise is needed. This classifier relies on maximizing the normalized or the un-normalized cross-correlation between the observation vector $X$ and each of the possible catalogue vectors $S$. So,

$$\text{choose } \omega_k \text{ if } \max_k \{ \max_m \text{correlation}(X, S_{k,m}) \}$$

There are scenarios where both vectors should be normalized with respect to their power to prevent the signature of one dominant target from being selected every time. In this case, this would be the signature of the Boeing 747 target model that has high energy and resembles most of the other target models. Unfortunately, any normalization does reduce the difference between target signatures rendering the classifier less effective. This type of classifier has not been explored in the stochastic resonance framework, so noise of various characteristics were added in a trial and error type approach.

3.4 Multi-layer perceptron classifier

Neural networks have had successes in radar target recognition and can be utilized in classifying unknown target backscatter into one of $M$ possibilities. The multi-layer perceptron trained with the back-propagation algorithm has been used for this purpose. The data being complex in nature is presented into the neural net as a long vector of real and imaginary components. The number of neurons in the hidden layer is selected as ten based on previous experimentation with target recognition. The three layer neural net has four output neurons representing each of the four possible classes. Clearly, the architecture of this neural net is not optimal and so is the choice of learning rate and presentation mode of the data. The key here to test the same neural net architecture with or without additive independent noise, and assess the performance change if any.

![Effect of adding independent Gaussian noise with varying intensity--nearest neighbor](image)

Fig. 1. Performance of nearest neighbor classifier when independent Gaussian noise with varying power (parameter $c$) was added to the observed target backscatter.
Fig. 2. Performance of cross-correlation classifier when independent Gaussian noise with varying power (parameter c) was added to the observed target backscatter.

Fig. 3. Performance of maximum likelihood classifier (assuming mixed Gaussian distribution over azimuth) when independent Gaussian noise with varying power (parameter c) was added to the observed target backscatter.
Effect of adding independent Gaussian noise with varying intensity—neural net

Fig. 4. Performance of neural net classifier when independent Gaussian noise with varying power (parameter c) was added to the observed target backscatter. The multi-layer perceptron neural net has 400 input nodes, 10 hidden nodes, and 4 output nodes.

Effect of adding independent discrete type noise with varying intensity—nearest neighbor

Fig. 5. Performance of nearest neighbor classifier when independent two-level dc noise with varying power (parameter c) was added to the observed target backscatter.
Fig. 6. Performance of cross-correlation classifier when independent two-level dc noise with varying power (parameter $c$) was added to the observed target backscatter.

Fig. 7. Performance of maximum likelihood classifier (assuming mixed Gaussian distribution vs azimuth) when independent two-level dc noise with varying power (parameter $c$) was added to the observed target backscatter.
Fig. 8. Performance of nearest neighbor classifier when independent Rician noise with varying power (parameter $c$) was added to the observed target backscatter.

Fig. 9. Performance of cross-correlation classifier when independent Rician noise with varying power (parameter $c$) was added to the observed target backscatter.
Fig. 10. Performance of maximum likelihood classifier (assuming mixed Gaussian distribution vs azimuth) when independent Rician noise with varying power (parameter c) was added to the observed target backscatter.

Fig. 11. Performance of neural net classifier when independent Rician noise with varying power (parameter c) was added to the observed target backscatter.
Effect of adding independent discrete type noise with varying intensity—nearest neighbor

Fig. 12. Performance of nearest neighbor classifier when independent two-positive-level noise with varying power (parameter c) was added to the observed target backscatter.

Effect of adding independent Rician noise with varying intensity—correlation based

Fig. 13. Performance of cross correlation classifier when independent positive two-level dc noise with varying power (parameter c) was added to the observed target backscatter.
Effect of adding independent discrete type noise with varying intensity—mixed Gaussian

Fig. 14. Performance of maximum likelihood classifier (assuming mixed Gaussian distribution vs azimuth) when independent positive two-level noise with varying power (parameter c) was added to the observed target backscatter.

Effect of adding independent discrete noise with varying intensity—neural net

Fig. 15. Performance of neural net classifier when independent positive two-level noise with varying power (parameter c) was added to the observed target backscatter.
4. DISCUSSION AND CONCLUSIONS

The data used in this study represents radar backscatter recorded in an anechoic chamber at high signal-to-noise ratio using miniature target models of four commercial aircraft namely Boeing 707, Boeing 727, Boeing 747, and DC10. The data is in the 2-12 GHz range with 50 MHz increments where both the amplitude and phase of the return are available. This frequency range corresponds to the UHF radar range when considering full size targets which is known as the resonance scattering region where some of the target scattering features are discernable but others may not be. This range of frequencies supports the nonlinearity assumption made earlier and motivates the interest in determining whether additive noise can enhance the classification performance. The data can be used in the frequency domain format, or converted into the time domain using inverse Fourier of windowed data segment. No further feature selection has been implemented beyond what is experimentally recorded. The range resolution is about 1.5 cm and the model targets lengths are between 10 and 15 cm.

Four types of noise were added to the observed backscatter in addition to the simulated additive white noise. The added noise signals are additive Gaussian with varying standard deviation C, additive Rician noise, additive two-level dc noise, and additive two positive levels dc noise. All four classifiers were simulated, nearest neighbor, cross-correlation, maximum likelihood assuming mixed Gaussian distribution, and neural net. The three-layer neural net has 10 hidden neurons, and 4 output neurons trained at 0dB SNR with additive Gaussian noise. The learning rate is held constant at 0.5. To ensure accurate estimate of error rate, 1000 experiments were conducted per signal-to-noise ratio level and per additive noise scenario, which corresponds to about 3% confidence interval for the error estimate.

Figures 1-15 show the effects of adding independent noise to the Gaussian noise corrupted observations. The probability distribution of the added noise did take some of the results in [2-9] into account but ultimately the search for the optimal noise distribution was ad-hoc for the simple reason that this problem is a M-ary hypotheses classification problem and not a binary hypotheses detection problem, and the suboptimal classifiers simulated have not been examined in the context of adding independent noise. Figures 1-15 show that the performance of parametric or non-parametric radar target classifiers is affected by the additive independent noise power and probability distribution function. Furthermore, the improvement in error rate if any is dependent on the signal-to-noise ratio of the observed backscatter. These results indicate that additive independent noise does have a role to play in radar target recognition studies. What is not clear is whether the improvement if any is significant enough to justify the added computational cost and added complexity of the classification system. This work does however motivate a theoretical investigation of the problem of M-ary parametric or nonparametric recognition using observations corrupted with additive Gaussian noise.

REFERENCES