Detection and Identification of Human Targets in Radar Data

Sevgi Z. Gürbüz[a], William L. Melvin[b], and Douglas B. Williams[a]

Georgia Institute of Technology, Atlanta, GA, 30332
[a] Center for Signal and Image Processing, School of Electrical and Computer Engineering
[b]Georgia Tech Research Institute

ABSTRACT

Radar offers unique advantages over other sensors, such as visual or seismic sensors, for human target detection. Many situations, especially military applications, prevent the placement of video cameras or implantment seismic sensors in the area being observed, because of security or other threats. However, radar can operate far away from potential targets, and functions during daytime as well as nighttime, in virtually all weather conditions. In this paper, we examine the problem of human target detection and identification using single-channel, airborne, synthetic aperture radar (SAR). Human targets are differentiated from other detected slow-moving targets by analyzing the spectrogram of each potential target. Human spectrograms are unique, and can be used not just to identify targets as human, but also to determine features about the human target being observed, such as size, gender, action, and speed. A 12-point human model, together with kinematic equations of motion for each body part, is used to calculate the expected target return and spectrogram. A MATLAB simulation environment is developed including ground clutter, human and non-human targets for the testing of spectrogram-based detection and identification algorithms. Simulations show that spectrograms have some ability to detect and identify human targets in low noise. An example gender discrimination system correctly detected 83.97% of males and 91.11% of females. The problems and limitations of spectrogram-based methods in high clutter environments are discussed. The SNR loss inherent to spectrogram-based methods is quantified. An alternate detection and identification method that will be used as a basis for future work is proposed.

Keywords: Human detection, personnel detection, target identification

1. INTRODUCTION

Radar offers unique advantages over other sensors for human target detection and identification. Radar can operate far away from potential targets and functions during the daytime as well as nighttime, in virtually all weather conditions. As a result, the use of radar for human detection and identification has become the subject of increased research.

Currently, most radar-based systems for human detection or identification almost always employ some form of spectral processing. For example, Yarovoy[1] used frequency spectrum analysis to detect human respiration. Falconer[2] measured the power spectral density (PSD) of a variety of objects and used the differences in the shaping of the PSD (kurtosis and energy band) to differentiate between targets and infer the activity level (i.e., resting versus moving) of human targets. Sabatini[3] used wavelet transforms to remove the high-frequency components and reconstruct the signal; the error between the original and reconstructed signals was then used to compute a threshold for discriminating between targets. In this latter case, human targets would have a low error, since they are slow-moving targets. Unfortunately, such methods could not reliably differentiate humans from other living creatures or, for that matter, low-frequency oscillations present in the environment, such as a slow-spinning ceiling fan.

By 2003, however, a spectral-based feature had emerged that was unique to human beings: the spectrogram. Research done by Geisheimer,[4] Van Dorp,[5] and others had shown that the human spectrogram could be formed by summing the spectrograms derived from the signals reflected from individual body parts, and that it contained distinct features suitable to distinguish human spectrograms from spectrograms of other animals. For example, Greneker[6] observed that the weight of a load around the waist caused a noticeable change in

Send correspondence to Sevgi Zubeyde Gurbuz: E-mail: gtg973c@prism.gatech.edu

the stride and velocity of a human walking, and correspondingly proposed a spectrogram-based suicide bomber detection system. Otero[7] explored the use of velocity, stride, and height of thigh in distinguishing men from women. However, the applicability and performance of the spectrogram-based features proposed for various identification problems were never actually tested in the context of a practical system; thus, the observations remained as suggestions requiring further research.

In this paper, the use of spectrogram-based features in practical radar environments is evaluated. A kinematic model for human walking is used to compute the received radar signal from a human target. Noise and clutter models are then added to emulate a more realistic environment. The reliability and accuracy of features extracted from the human spectrogram are examined with a gender discrimination system used as an example. In light of the limitations of spectrogram-based methods, a new framework for solving the detection and identification problem is proposed as the basis for future work.

2. SIGNAL MODELLING

The signal received by the radar is generally comprised of three components: the return from the human target, receiver noise, and clutter. The mathematical models employed for each of these component signals is explained next.

2.1. Human Model

A human is a complicated target because of the intricate motion of body parts moving along different trajectories at different speeds. An experimentally verified simplification[4] of the problem is to represent the overall received signal as the sum of the returns from a finite number of point targets, each representing a body part. In this work, a total of twelve basic body parts is utilized: the head, upper arms, lower arms, the torso, thighs, lower legs and feet (Fig. 2). As indicated in the figure, each point target is taken to lie at the center of the corresponding body part.

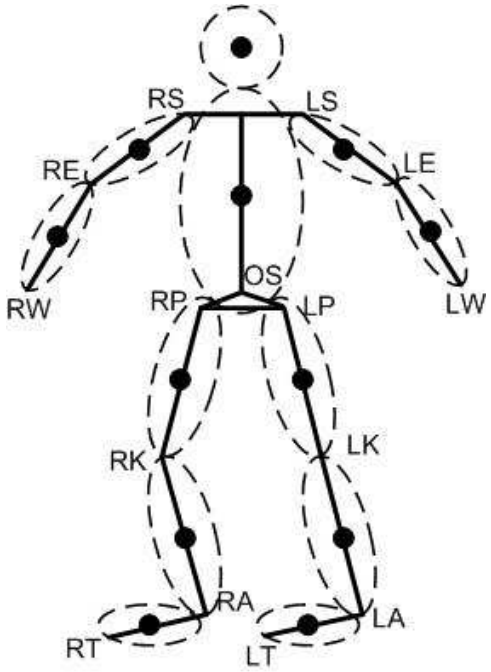


Figure 1. 12-point human model.

The time-varying position of each point target may be computed using the kinematic model of a walking human developed by Thalmann.[8] In the Thalmann model, all positions are referenced to the base of the spine,

denoted by OS. Over the course of one cycle (two steps), the vertical, lateral, and translational position of OS varies sinusoidally. The time-varying angles at the joints between the body parts is also provided by means of charts and equations in Thalmann's paper. Dimensions of the human body parts may then be used in conjunction with these time-varying joint angles to compute the time-varying positions relative to OS. A table of the body part dimensions used in this work may be found in Table 1. Because overall human heights vary according to a Gaussian distribution, the distributions of the body part dimensions are also assumed to be Gaussian. The dimensions for any desired human may be found using interpolation as follows:

Table 1. Human Physical Dimensions[9]

(meters)	95%ile Male	5%ile Male	95%ile Female	5%ile Female
Initial OS	1.180	1.005	1.085	0.930
Height of Thigh	1.000	0.84	0.885	0.740
Torso Height	0.505	0.390	0.460	0.360
Head Height	0.205	0.180	0.190	0.165
Hip Width	0.405	0.310	0.435	0.310
Hip Height	0.295	0.195	0.280	0.185
Thigh Length	0.550	0.440	0.530	0.435
Lower Leg Length	0.490	0.395	0.445	0.355
Foot Length	0.285	0.240	0.255	0.215
Shoulder Length	0.215	0.1825	0.1925	0.1625
Shoulder-Elbow Length	0.395	0.330	0.360	0.300
Elbow-Wrist Length	0.305	0.265	0.270	0.240
Torso Diameter	0.357	0.290	0.340	0.2675
Upper Arm Diameter	0.095	0.080	0.085	0.070
Lower Arm Diameter	0.07125	0.060	0.06375	0.0525
Thigh Diameter	0.185	0.135	0.180	0.125
Leg Diameter	0.1387	0.1013	0.135	0.9375
Foot Diameter	0.110	0.085	0.100	0.080

Define

$$k = \Phi(x) = \frac{1}{2\pi} \int_{-\infty}^{x} e^{-t^2/2} dt. \tag{1}$$

Then the dimension of any body part is

$$d_{bp} = \mu_{bp} + k\sigma_{bp}, \tag{2}$$

where μ_{bp} and σ_{bp}^2 are the mean and variance of the distribution of that body part, as calculated from the values of the 95%ile and 5%ile sizes.

A key factor is that the equations and charts used in Thalmann's model all ultimately depend on just two parameters: the walking speed (V) and height of thigh (HT). As a result, these two features (V and HT) are of special interest when considering the human detection and identification problem.

Given knowledge of the time-varying positions of each point target, the time-delay between the transmission of the radar chirp waveform and the receipt of the reflected signal may be computed as

$$t_d = \frac{2R}{c}, \tag{3}$$

where R is the distance between the antenna and target, and c is the speed of light.

Thus, the received signal from a human target may be mathematically reflected as

$$s_r(n,t) = \sum_{k=1}^{12} a_{t,k} rect(\frac{\hat{t_k} - t_{d,k}}{\tau}) exp(j[-2\pi f_c t_{d,k} + \pi\gamma(\hat{t_k} - t_{d,k})^2]), \tag{4}$$

where n is the slow-time index (pulse number); $\hat{t}$ is the fast-time index relative to the beginning of the pulse; τ is the pulse width of the transmitted chirp; f_c is the center frequency of the transmitted chirp; t_k is time; γ is the frequency slope of the transmitted chirp; and k is an index referencing each body part.

The amplitude of the received signal, $a_{t,k}$, is given by the radar range equation:

$$a_{t,k} = \frac{G\lambda\sqrt{P_t\sigma_k}\sigma_n}{(4\pi)^{1.5}R_k^2\sqrt{L_s}\sqrt{L_a}\sqrt{T_{sys}}} \tag{5}$$

where G is the antenna gain, λ is the wavelength, P_t is the transmitted signal power, σ is the RCS, L_s is the system loss, L_a is the atmospheric loss, T_{sys} is the system temperature, and σ_n is the noise standard deviation.

The antenna gain varies according to the angle that the signal is received, and the atmospheric losses vary with range; however, for simplicity, our simulation assumes that these parameters have constant values. The RCS, on the other hand, is modelled according to the shapes of the body parts. Thus, the head is modelled as a sphere and the other body parts as cylinders. With computation of $a_{t,k}$, the human signal model given by (4) is complete.

2.2. Noise and Clutter Model

Noise is a source of random variations in the received radar signal. There are several sources of noise, such as electronic or thermal noise; in this work, all system noise is modelled as white Gaussian.

Clutter is due to unwanted reflections off of objects in the environment that are not targets, such as trees, asphalt roads, buildings, and multi-path target reflections. Depending on the environment, clutter characteristics may actually vary quite a bit. However, to keep our results as general as possible, an approximate colored Gaussian noise model is used for the clutter. The frequency spectrum (S_c) of the clutter is modelled as a sinc function. Multiple realizations of this clutter spectrum are generated by multiplying with a vector of complex Gaussian random numbers, w_k, where k is a sample index. Thus, the clutter samples may be approximated as follows:

$$c_k = F^{-1}[S_c F(w_k)], \tag{6}$$

where $F(\cdot)$ represents the Fourier Transform of the variable in parenthesis.

3. SPECTROGRAM ANALYSIS

Although in our case the synthetic aperture is not explicitly used for the purposes of increasing resolution, since the data collection geometry is identical to that of synthetic aperture radar (SAR), SAR processing techniques are used to obtain the range-Doppler map for a human target:

1. Pulse Compression
2. Range Migration Correction
3. Quadratic Phase Error Correction (if applicable)
4. Doppler Compression

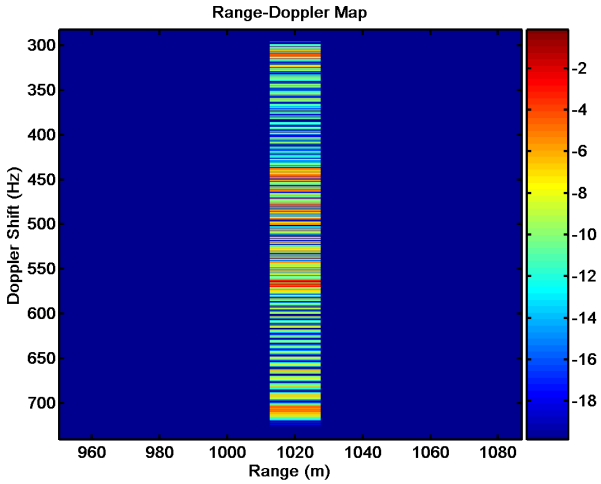


Figure 2. Range-Doppler map for a human target.

An example of the range-Doppler map for a human target is shown in Fig. 2. Notice the significant spread of the Doppler component caused by the varying motion of the appendages.

The spectrogram is computed by taking a slice across slow-time at the range of the target and stacking the fast Fourier transforms (FFT) of short, overlapping time segments. An example of the simulated spectrogram for a human target is shown in Fig. 3. Notice how the strongest return is received from the torso, with its low-frequency, small-amplitude sinusoidal oscillatory motion. On the other hand, the appendages that travel the farthest during a single walking cycle - the feet - appear in the spectrogram with the largest amplitude oscillation. The simulated spectrogram compares nicely with that of a human spectrogram measured by Otero[7] (Fig. 4a) and has distinct features that make it distinguishable from that of other living creatures; for example a dog (Fig. 4b), whose spectrogram has many high-frequency components, almost making it appear random.

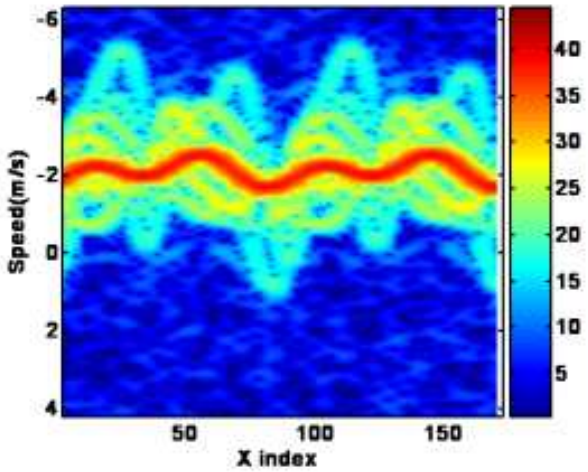


Figure 3. Simulated human spectrogram.

3.1. Extracting Features from the Spectrogram

In fact several authors[6,7,10] have proposed a variety of spectrogram-based features that could be used for detecting or identifying humans, such as velocity, stride, and appendage-to-torso ratio (AppTorR). In particular,

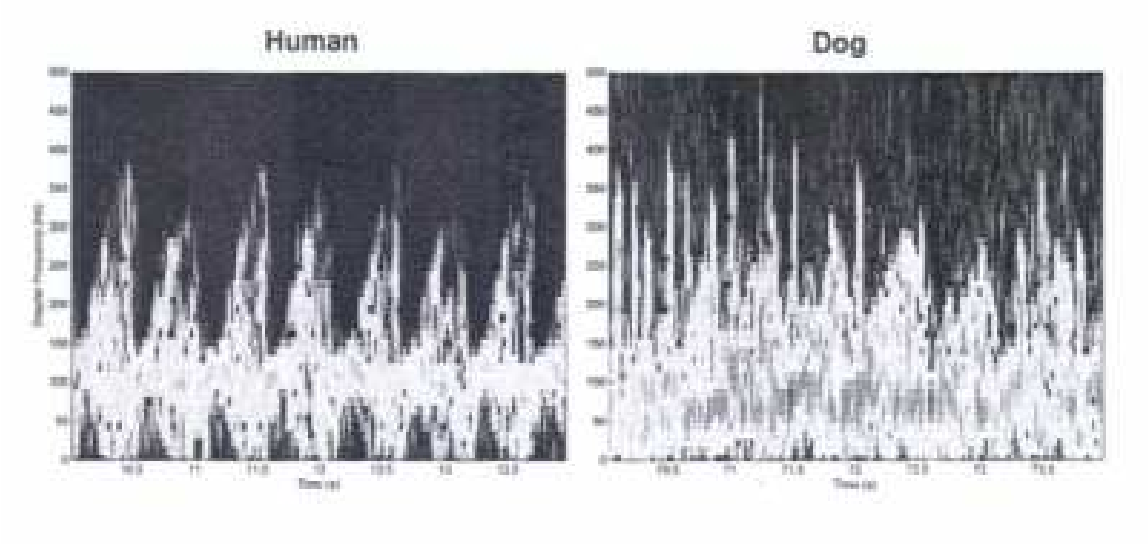


Figure 4. Measured spectrograms for a human (a) and a dog (b), as given in Otero's work.[7]

Otero devised a technique for estimating these features based on the cadence frequency plot derived from the spectrogram by taking its fast Fourier transform across time. A sample cadence frequency plot is shown below in Fig. 5.

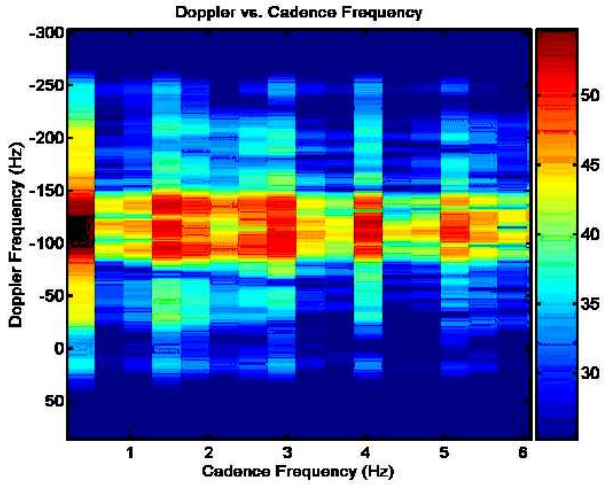


Figure 5. Cadence frequency plot for a human target (2 sec. of data).

The cadence frequency plot highlights the periodic signals present in the spectrogram. The peak at about zero cadence frequency corresponds to the motion of the torso, with the value of the Doppler frequency (f_{Dtorso}) directly related to the walking speed (v):

$$v = \frac{cf_{Dtorso}}{2f_c}. \tag{7}$$

The next peak corresponds to the fundamental frequency, f_m, and is related to the stride length,

$$l_s = \frac{v}{f_m}. \tag{8}$$

Subsequent peaks result from the motion of the other appendages, such as the arms and legs. If we consider the amplitudes of the peaks to be indicative of RCS, then the ATR may be computed as follows:

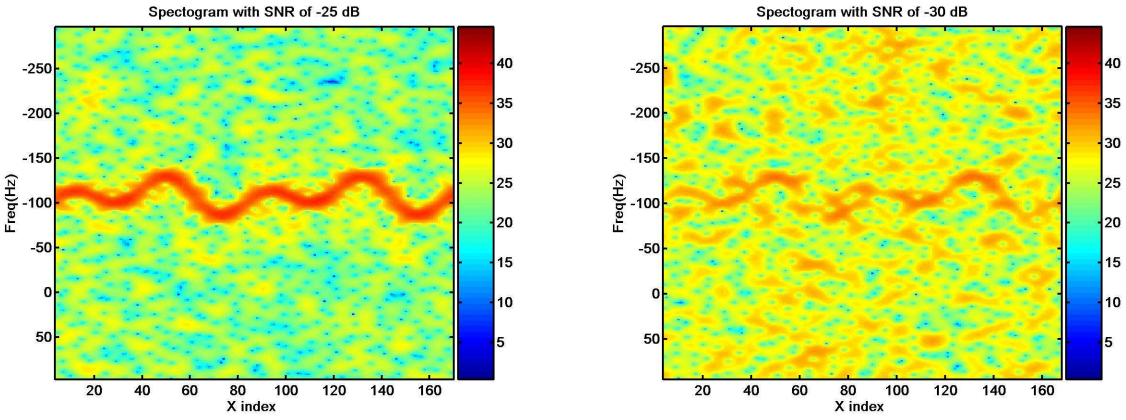
$$AppTorR = \frac{\sum RCS_{appendages}}{RCS_{torso}} \tag{9}$$

Finally, the height of thigh - the second main parameter upon which the Thalmann kinematic model is based - may be computed as

$$HT = (\frac{4}{1.346^2})\frac{v}{f_m^2} \tag{10}$$

At first glance, these features seem like a promising solution to a multitude of detection and identification problems. For example, the target velocity and AppTorR could be used to differentiate humans from other animals. The velocity profiles of dogs, for instance, as well as their AppTorR, is quite different from that of humans. Alternatively, an estimate of HT could be used to try to differentiate human males and females.

Unfortunately, estimates of these parameters derived from the cadence frequency plot tend to be quite unreliable, especially for short or medium dwell scenarios that preclude collecting enough data for a FFT of sufficient length to estimate accurately the fundamental cadence frequency, f_m. Noise and clutter are also extremely detrimental to obtaining an accurate estimate of these parameters. The torso signal is just barely visible at an average pulsewise SINR of -25dB, and at -30dB the spectrogram is completely masked by the noise (Fig. 6). Simulation results also show that the variance of the estimates becomes significantly high as SINR increases.



(a) (b)

Figure 6. Spectrogram in noise. (a) SINR = -25 dB. (b) SINR = -30 dB.

An alternative way of estimating the target velocity and HT is to use a backprojection-style technique in which the human model is used to generate the expected spectrogram for a range of velocity and HT values. These expected spectrograms are then overlayed on the measured spectrogram data, and the corresponding points in the data are summed to yield a statistic. The best match yields the highest statistic (Fig. 7).

The backprojection method yields significantly more robust results in comparison to the estimates directly computed from the cadence frequency plot. The plot of both estimates versus SINR (Fig. 8) clearly shows the performance improvement.

3.2. An Example: Gender Discrimination

To illustrate and benchmark the performance of feature-based decisions, consider the example of gender discrimination based on the HT estimate. Assume that the distribution of HT for men and women both vary according to a Gaussian distribution.

$$P_m(x) = \frac{1}{\sqrt{2\pi\sigma_1^2}} e^{-\frac{1}{2\sigma_1^2}(x-\mu_1)^2}, \tag{11}$$

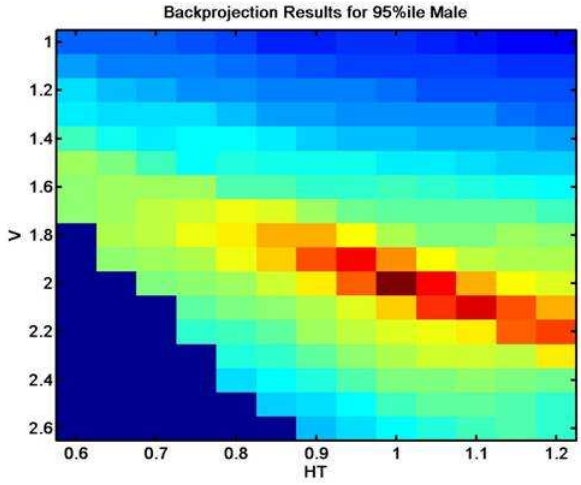


Figure 7. Backprojection statistic for a range of velocity and HT values .

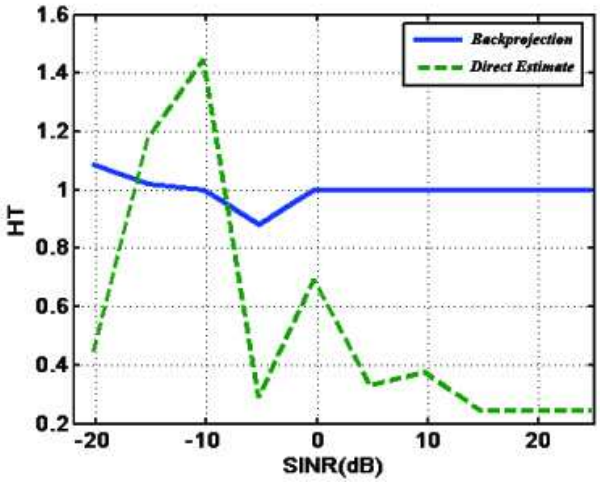


Figure 8. Comparison of backprojection and direct estimates against SINR, for a true HT value of 1 m.

where $P_m(x)$ is the PDF of HT, μ_1 is the average HT, and σ_1 is the variance of HT for a male.

$$P_f(x) = \frac{1}{\sqrt{2\pi\sigma_2^2}} e^{-\frac{1}{2\sigma_2^2}(x-\mu_2)^2}, \tag{12}$$

where $P_f(x)$ is the PDF of HT, μ_2 is the average HT, and σ_2 is the variance of HT for a female.

The Neyman-Pearson detector may be expressed as a choice between two hypothesis: H_1, the person is male, and H_0: the person is female. If the ratio of the probability for H_1 over the probability of H_0 is greater than some threshold, γ, then a decision of "male" will be returned. Mathematically,

$$\frac{P_m(x)}{P_f(x)} > \gamma \tag{13}$$

or

$$T = \frac{(x-\mu_2)^2}{\sigma_2^2} - \frac{(x-\mu_1)^2}{\sigma_1^2} > 2(\gamma + \ln(\frac{\sigma_1}{\sigma_2})) = \gamma'. \tag{14}$$

Based on the human dimensions provided in Table 1, the mean and variance for the male and female PDFs may be computed to be $\mu_1 = 0.92$ m, $\mu_2 = 0.8125$ m, $\sigma_1 = 0.04863$ m, and $\sigma_2 = 0.04407$ m. Notice that the variances for the male and female distributions are approximately the same. In the case when the variances are identical ($\sigma_1 = \sigma_2 = \sigma$), the expression for the test statistic reduces to

$$T = \frac{2}{\sigma^2}(\mu_1 - \mu_2)x + \frac{\mu_2^2 - \mu_1^2}{\sigma^2}, \tag{15}$$

a Gaussian random variable with mean and variance

$$\mu_{T|H_1} = (\frac{\mu_1 - \mu_2}{\sigma})^2 \tag{16}$$

$$\mu_{T|H_0} = -(\frac{\mu_1 - \mu_2}{\sigma})^2 \tag{17}$$

$$\sigma_T^2 = (\frac{2(\mu_1 - \mu_2)}{\sigma})^2. \tag{18}$$

Histogram plots of one million realizations of the test statistic show that the distribution does indeed follow an approximately bell-shaped curve. Thus, $T \approx N(\mu_T, \sigma_T)$ as defined in equations (16), (17) and (18), with $\sigma \equiv (\sigma_1 + \sigma_2)/2$.

The detection threshold may be set with a constant false alarm rate as follows:

$$\gamma' = \sigma_T Q^{-1}(PFA) + \mu_{T|H_0}, \tag{19}$$

where PFA is the probability of false alarm, and the Q-function is defined as

$$Q(\alpha) \equiv \frac{1}{\sqrt{2\pi}} \int_\alpha^\infty e^{-\frac{x^2}{2}} dx \tag{20}$$

Substituting the estimate of HT found by backprojection as x in Eq. 14 to calculate the test statistic, and comparing with the threshold found from Eq. 19, the target may be classified as male or female. In a noiseless environment, for 100,000 people with randomly selected thigh heights, males were correctly detected 83.97% of the time, and females were correctly detected 91.11% of the time.

3.3. Sub-Optimal Nature of Spectrogram-Based Methods

Although the backprojection method works well, even in noise, it is computationally intensive. Thus, this approach is not a very practical solution for real-time detection systems. Parameter estimation directly using the spectrogram is a much quicker approach, but, as we have already shown, yields results that are neither accurate nor robust in noise. In addition to these practical problems, it is important to note that the spectrogram is also sub-optimal.

The spectrogram is an optimal matched filter response only when the signal phase is linear. However, the phase for human targets is not linear, but higher order. The sub-optimal nature of the spectrogram may be quantified by considering the SNR loss relative to an ideal matched filter.

As previously explained, the spectrogram is computed by taking the discrete Fourier transform of small segments of a slice across slow-time of the range compressed data matrix. Thus, for any range bin r_b, this slice may be represented as the sum of the return from each body part, T_k:

$$T_k(r_b) = \alpha \begin{bmatrix} 1 \\ e^{j\theta_1} \\ \vdots \\ e^{j\theta_{N-1}} \end{bmatrix}, \tag{21}$$

where N is the number of pulses transmitted by the radar and α is a complex constant. In general, the phase θ will be nonlinear. The spectrogram, on the other hand, is based on the Fourier Transform, requiring that the phase be linear.

$$s_{t,linear} = \alpha \begin{bmatrix} 1 \\ e^{j\phi_1} \\ \vdots \\ e^{j(N-1)\phi_1} \end{bmatrix}, \tag{22}$$

where α is a constant. Thus, the SNR loss may be calculated as

$$SNR_{loss} = \frac{SNR_{spectrogram}}{SNR_{ideal}} = \frac{|s_{t,linear}^H s_{t,target}|^2}{|s_{t,target}^H s_{t,target}|^2}, \tag{23}$$

where

$$s_{t,target}^H s_{t,target} = \alpha^2 \begin{bmatrix} 1 & e^{j\theta_1} & \dots & e^{j\theta_{N-1}} \end{bmatrix} \begin{bmatrix} 1 \\ e^{j\theta_1} \\ \vdots \\ e^{j\theta_{N-1}} \end{bmatrix} = \alpha^2 N \tag{24}$$

and

$$\begin{aligned} s_{t,linear}^H s_{t,target} &= \alpha^2 \begin{bmatrix} 1 & e^{j\phi_1} & \dots & e^{j(N-1)\phi_1} \end{bmatrix} \begin{bmatrix} 1 \\ e^{j\theta_1} \\ \vdots \\ e^{j\theta_{N-1}} \end{bmatrix} \\ &= \alpha^2 [1 + e^{j(\theta_1 - \phi_1)} + e^{j(\theta_2 - 2\phi_2)} + \dots + e^{j(\theta_{N-1} - (N-1)\phi_1)}] = \alpha^2 mN < \alpha^2 N. \end{aligned} \tag{25}$$

Since the magnitudes of the exponentials in (25) are all equal to one, the sum of the exponentials must be less than N, the total number of terms, and the magnitude of the complex scaling factor m must be less than one.

With these results, the SNR loss can now be expressed as

$$SNR_{Loss} = \frac{SNR_{Spectogram}}{SNR_{Ideal}} = \frac{(|m|N)^2}{N^2} = |m|^2 \tag{26}$$

4. CONCLUSIONS AND FUTURE WORK

Although spectrogram-based methods have shown some success in applications where there is low noise and the target is close to the antenna, thereby ensuring high SNR, in many applications it is common for the SNR to be low. In these situations, accurate and robust estimation of features from the spectrogram or its derivatives is very difficult and leads to poor performance. Fundamentally, the spectrogram is also a sub-optimal method for signals with nonlinear phase, such as human target returns. Therefore, future work will focus on the development of techniques to implement an ideal matched filter by refining parameter estimates with iteration.

REFERENCES

1. A.G. Yarovoy, L.P. Ligthart, J. Matuzas, and B. Levitas, "UWB radar for human being detection," *IEEE Aerospace and Electronic Systems Magazine,* Vol. 21, Issue 3, pp. 10-14, Mar. 2006.
2. D.G. Falconer, R.W. Ficklin, and K.G. Konolige, "Robot-mounted through-wall radar for detecting, locating, and identifying building occupants, *IEEE International Conference on Robotics and Automation,* Vol. 2, pp. 1868-1875, Apr. 2000.
3. A.M. Sabatini, and V. Colla, "A Method for Sonar Based Recognition of Walking People," *Robotics and Autonomous Systems,* Vol. 25, pp. 117-126, 1998.
4. J.L. Geisheimer, E.F. Greneker, and W.S. Marshall, "A High-Resolution Doppler Model of Human Gait," *Proceedings of SPIE, Radar Sensor Technology and Data Visualization,* Vol. 4744, 2002.
5. P. van Dorp and F.C.A. Groen, "Human Walking Estimation with Radar," *IEE Proceedings on Radar, Sonar and Navigation,* Vol. 150, Issue 5, pp. 356-365, Oct. 2003.
6. G. Greneker, "Very low cost stand-off suicide bomber detection system using human gait analysis to screen potential bomb carrying individuals," *Proc. of SPIE,* Vol. 5788, 2005.
7. M. Otero, "Application of a Continuous Wave Radar for Human Gait Recognition," *Proceedings of SPIE, Signal Processing, Sensor Fusion and Target Recognition XIV* Vol. 5809, pp. 538-548, 2005.
8. R. Boulic, M.N. Thalmann, and D. Thalmann, "A Global Human Walking Model with Real-Time Kinematic Personification," *Vis. Comput.* Vol. 6, pp. 344-358, 1990.
9. RoyMech Website: http://www.roymech.cok.uk/Useful_Tables/Human/Human_sizes.html
10. J.L. Geisheimer, W.S. Marshall, E. Greneker, "A Continuous-Wave (CW) Radar for Gait Analysis," *Signals, Systems, and Computers,* Vol. 1, pp. 834-838, 2001.