# Micro-Doppler Classification for Ground Surveillance Radar Using Speech Recognition Tools

Dalila Yessad[1], Abderrahmane Amrouche[1], Mohamed Debyeche[1], and Mustapha Djeddou[2]

[1] Speech Communication and Signal Processing Laboratory, Faculty of Electronics and Computer Sciences, USTHB, P.O. Box 32, El Alia, Bab Ezzouar, 16111, Algiers, Algeria
[2] Communication Systems Laboratory, Ecole Militaire Polytechnique, BEB, Algiers, Algeria
namrouche@usthb.dz,
{yessad.dalila,mdebyeche,djeddou.mustapha}@gmail.com

**Abstract.** Among the applications of a radar system, target classification for ground surveillance is one of the most widely used. This paper deals with micro-Doppler Signature ($\mu$-DS) based radar Automatic Target Recognition (ATR). The main goal for performing $\mu$-DS classification using speech processing tools was to investigate whether automatic speech recognition (ASR) techniques are suitable methods for radar ATR. In this work, extracted features from micro-Doppler echoes signal, using MFCC, LPC and LPCC, are used to estimate models for target classification. In classification stage, two parametric models based on Gaussian Mixture Model (GMM) and Greedy GMM were successively investigated for echo target modeling. Maximum a posteriori (MAP) and Majority-voting post-processing (MV) decision schemes are applied. Thus, ASR techniques based on GMM and GMM Greedy classifiers have been successfully used to distinguish different classes of targets echoes (humans, truck, vehicle and clutter) recorded by a low-resolution ground surveillance Doppler radar. Experimental results show that MV post processing improves target recognition and the performances reach to 99, 08% correct classification on the testing set.

**Keywords:** Automatic Target Recognition (ATR), micro-Doppler Signatures ($\mu$-DS), Automatic Speech Recognition (ASR), Gaussian Mixture Model (GMM), Greedy GMM, Maximum a Posteriori (MAP), Majority Vote (MV).

## 1 Introduction

Target classification using radar signatures has potential applications in air/ marine traffic and ground surveillance radar. The goal for any target recognition system is to give the most accurate interpretation of what a target is at any given point in time. ATR is a crucial task for both military and civil applications.

C. San Martin and S.-W. Kim (Eds.): CIARP 2011, LNCS 7042, pp. 280–287, 2011.

In the acquisition stage, each target is illuminated with a frequency stepped signal and the returned echoes are then received. The radar operator identifies targets from the audio representation of the echoes signal. Mechanical vibration or rotation of a target may induce additional frequency modulations on the returned radar signal. This phenomenon, known as the micro-Doppler effect, generates sidebands at the target Doppler frequency.

Techniques based on micro-Doppler signatures [1], [2] are used to divide targets into several macro groups such as aircrafts, vehicles, creatures, etc. An effective tool to extract information from this signature is the time-frequency transform [3]. The time-varying trajectories of the different micro-Doppler components are quite revealing, especially when viewed in the joint time-frequency space [4]. Anderson [5] used micro-Doppler features to distinguish among humans, animals and vehicles. In [6], analysis of radar micro-Doppler signature with time-frequency transform was discussed. The time-frequency signature of the micro-Doppler provides additional time information and shows micro-Doppler frequency variations. Thus, additional information about vibration rate or rotation rate is available for target recognition. Gaussian mixture model (GMM)-based classification methods are widely applied to automatic speech and speaker recognition [7]. Mixture models form a common technique for probability density estimation. In [8], it was proved that any density can be estimated using finite Gaussian mixture. A Greedy learning of GMM based target classification for ground surveillance Doppler radar, recently proposed in [9], overcomes the drawbacks of the Expectation Minimization (EM) algorithm. The greedy learning algorithm does not require prior knowledge of the number of components in the mixture, because it inherently estimates the model order.

In this paper, we investigate the micro-Doppler radar signatures in order to obtain best classification performances. The classification algorithms are implemented using three kinds of features; Mel-Frequency Cepstral Coefficients (MFCC), Linear Prediction Coding (LPC) and Cepstrum Coefficient feature sets (LPCC), extracted from echoes signals recorded by Doppler radar. These features are fed respectively to GMM and greedy GMM parametric and statistical classifier approaches for multi-hypotheses problem. The classification tasks include the determination of the statistical modeling of extracted features distribution and the application of Maximum a posteriori (MAP) rule. As a post-processing enhancement method, a majority vote technique is proposed.

This paper is organized as follows: in section 2, features extractions and classification schemes are presented. In Section 3, we describe the experimental framework including the data collection. Experimental results are drawn in section 4.

## 2 Classification Scheme

In this paper, a supervised classification process was performed and two decision methods were implemented.

### 2.1 Feature Extraction

In practical case, a human operator listen to the audio Doppler output from the surveillance radar for detecting and may be identifying targets. In fact, human operators classify the targets using an audio representation of the micro-Doppler effect, caused by the target motion. As in speech processing a set of operations are taken during pre-processing step to take in count the human ear characteristics. Features are numerical measurements used in computation to discriminate between classes. In this work, we investigated three classes of features namely, LPC, LPCC, and MFCC.

**Linear Prediction Coding (LPC).** Linear prediction is the process of predicting future sample values of a digital signal from a linear system. It is therefore about predicting the signal $x(n)$ at instant $n$ from $p$ previous samples as in (1).

$$x(n) = \sum_{k=1}^{p} a_k x(n-k) + e(n) \tag{1}$$

So the coding by linear prediction consists in determining coefficients $a_k$ that minimize the error $e(n)$. LPC are expected to give very accurate formant information of acoustic signals. We considered the LPC up to the $16^{th}$ order (excluding the zero coefficient) and applied it directly to the radar signal.

**Cepstral Linear Prediction Coding (LPCC).** The cepstrum coefficients $\{ceps_q\}_{q=0}^{Q}$ can be estimated from the LPC coefficients $\{a_q\}_{q=1}^{p}$ using a recursion procedure:

$$ceps_q = \begin{cases} ln(G), & q=0 \\ a_q + \sum_{k=1}^{q-1} \frac{k-q}{q} a_k ceps_{q-k}, & 1 \leqslant q \leqslant p \\ \sum_{k=1}^{p} \frac{k-q}{q} a_k ceps_{q-k}, & p < q \leqslant Q \end{cases} \tag{2}$$

Where $G$ is the gain term in the LPC model, $p$ the LPC model order, and $Q+1$ the number of cepstrum coefficients.

**Mel Frequency Cepstral Coefficients (MFCC).** The most commonly used feature vector in speech recognition is composed of Mel-Frequency Cepstral Coefficients (MFCC). Fig.1 is a block diagram of the MFCC generation process from micro-Doppler signal. The MFCC extraction is done in three steps:

1. Step 1-a: Cut up the signal in several overlapping windows;
2. Step 1-b: To decrease the spectral distortion a Hamming windowing is applied to signal frames;

$$W(n) = 0.54 - 0.46 * cos(\frac{2\pi n}{N-1}) \tag{3}$$

   Where $N$ is the window size.

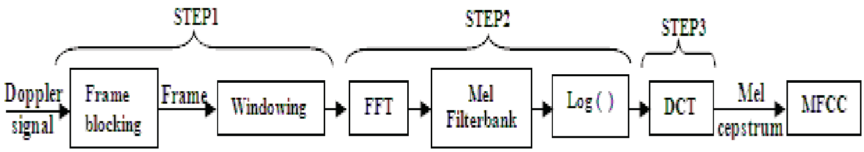


**Fig. 1.** MFCC generation

3. Step 2-a: Apply the FFT ;
4. Step 2-b: The Mel-frequency scale is applied to obtain an appropriate signal representation In fact psychophysical studies have shown that human perception of the frequency content of sounds does not follow a linear scale. The recognition model thus reflects the behaviour of the brain in this respect and is equally applicable to both speech and radar Doppler. We use the following transformation formula;

$$mel(f) = 2595 * log10(1 + \frac{f}{700}) \tag{4}$$

5. Step 2-c: Apply the logarithm after the Mel scale;
6. Step 3: Finally, obtain the discrete cosine transform (DCT) of the output signal.

### 2.2 Modelisation

In the present work, each target class is represented by two parametric models; GMM and Greedy GMM.

**Gaussian mixture model (GMM).** Gaussian mixture model (GMM) is a mixture of several Gaussian distributions. The probability density function is defined as a weighted sum of Gaussians:

$$p\,(x;\theta) = \sum_{c=1}^{C} \alpha_c N(x;\mu_c,\Sigma_c) \tag{5}$$

Where $\alpha_c$ is the weight of the component $c$, $0 < \alpha_c < 1$ for all components, and $\sum_{c+1}^{C} \alpha_c = 1$. $\mu_c$ is the mean of components and $\Sigma_c$ is the covariance matrix. We define the parameter vector $\theta$:

$$\theta = \{\alpha_1, \mu_1, \Sigma_1, ..., \alpha_c, \mu_c, \Sigma_c\} \tag{6}$$

Estimating the Gaussian mixture parameters for one class can be considered as an unsupervised learning in the case where samples are generated by individual components of the mixture distribution. The expectation maximization (EM) algorithm is an iterative method for calculating maximum likelihood distribution parameter. This algorithm starts from an initial guess $\theta^0$ for the distribution parameters and the log-likelihood is guaranteed to increase at each iteration until it converges. The initialization is one of the crucial problems of the EM algorithm. The selection of $\theta^0$ determines where the algorithm converges or hits the boundary of the parameter space producing singular, meaningless results. An elegant solution for the initialization problem is provided by the greedy learning of GMM [10].

**Greedy Gaussian mixture model (Greedy GMM).** The greedy algorithm starts with a single component and then adds components into the mixture one by one. The optimal starting component for a Gaussian mixture is trivially computed, optimal meaning the highest training data likelihood. The algorithm repeats two steps: insert a component into the mixture, and run EM until convergence. Inserting a component that increases the likelihood the most is thought to be an easier problem than initializing a whole near-optimal distribution. Component insertion involves searching for the parameters for only one component at a time. Recall that EM finds a local optimum for the distribution parameters, not necessarily the global optimum which makes it initialization dependent method [10].

## 2.3 Classifiers

A classifier is a function that defines the decision boundary between different patterns (classes). Each classifier must be trained with a training dataset before being used to recognize new patterns, such that it generalizes training dataset into classification rules. Two decision methods were examined. The first one suggests the maximum a posteriori probability (MAP) and the second uses the majority vote (MV) post-processing after classifier decision.

**Decision.** If we have a group of targets represented by the GMM models: $\lambda_1, \lambda_2, ..., \lambda_\xi$, The classification decision is done using the posteriori probability (MAP):

$$\hat{S} = \arg \underline{\xi} \max p(\lambda_s|X) \tag{7}$$

According to Bayesian rule:

$$\hat{S} = \arg\max \frac{p(X|\lambda_s)p(\lambda_s)}{p(X)} \tag{8}$$

$X$: is the observed sequence.

Assuming that each class has the same a priori probability ($p(\lambda_s) = 1/\xi$) and the probability of apparition of the sequence is the same for all targets the classification rule of Bayes becomes:

$$\hat{S} = \arg\max p(X|\lambda_s) \tag{9}$$

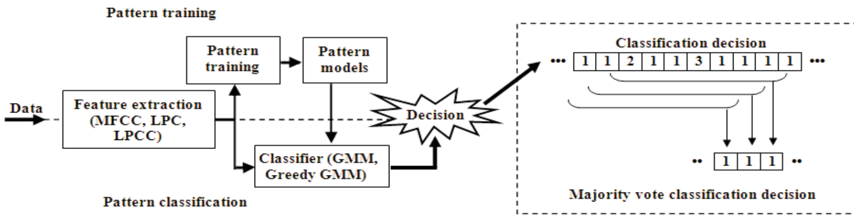


**Fig. 2.** Majority vote post-processing after classifier decision

**Majority Vote.** The majority vote (MV) post-processing can be employed after classifier decision. It uses the current classification result, along with the previous classification results and makes a classification decision based on the class that appears most often. A plot of the classification by MV (post-processing) after classifier decision is shown in Fig.2.

## 3 Measurements and Data Collection

Data were obtained using records of a low-resolution ground surveillance radar. The target was detected and tracked automatically by the radar, allowing continuous target echo records from the following targets: 1, 2, and 3 persons, vehicle, truck and clutter. We first collected the Doppler signatures from the echoes of six different targets in movements namely, one, two, and three persons, vehicle, truck and vegetation clutter. The target was detected and tracked automatically by a low-power Doppler radar operating at 9.72 GHz, sweep in azimuth 30 at 270 and emission power is 100mW. When the radar transmits an electromagnetic signal in the surveillance area, this signal interacts with the target and then returns to the radar. After demodulation and analog to digital conversion,

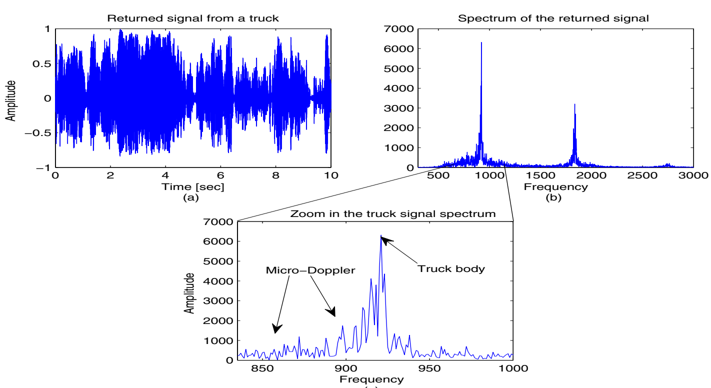


**Fig. 3.** (a)Returned signal from a truck (b) Spectrum of the returned signal (c) Zoom in the truck signal spectrum

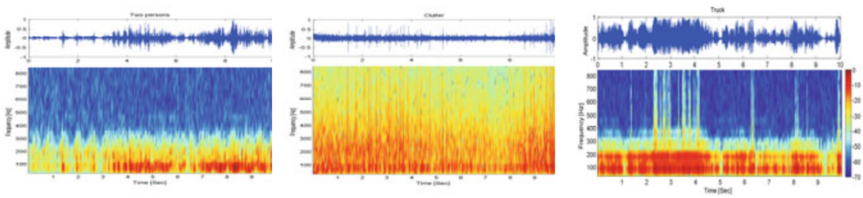

**Fig. 4.** Radar echo samples and the typical spectrograms of three moving targets; a) Two persons, b) Clutter, c) Truck

the received echoes are recorded in wav audio format; each record has a duration of 10 seconds. By taking the Fourier transform of the recorded signal, the micro-Doppler frequency shift may be observed in the frequency domain. An illustration of a measurement and its spectrum is shown in Fig.3. The change of the properties of the returned signal reflects the characteristics of the target. When the target is moving, the carrier frequency of the returned signal will be shifted due to Doppler effect. The Doppler frequency shift can be used to determine the radial velocity of the moving target. If the target or any structure on the target is vibrating or rotating in addition to target translation, it will induce frequency modulation on the returned signal that generates sidebands about the target's Doppler frequency. This modulation is called the micro-Doppler ($\mu$-DS) phenomenon. The $\mu$-DS phenomenon can be regarded as a characteristic of the interaction between the vibrating or rotating structures and the target body. Fig.4 (a)-(c) show the temporal representation and the typical spectrograms of three targets for two persons, clutter and truck. Each target class has unique time-frequency characteristic which can be used for classification. These particular plots are obtained by taking a succession of FFTs and using a sampling rate of 8 KHz, FFT size of 256 points, overlap of 128, and a hamming window.

## 4 Results

In this work, target class pdfs were modeled by GMMs using both greedy and EM estimation algorithms. MFCC, LPCC and LPC coefficients were used as classification features. The MAP and the majority voting decision concepts were

**Table 1.** Confusion matrix of Greedy GMM-based classifier with MFCC coefficients and MV post-processing after MAP decision rule for six-class problem

| Class / Decision | 1Person | 2Persons | 3Persons | Vehicle | Truck | Clutter |
|---|---|---|---|---|---|---|
| 1Person | 96.30 | 1.85 | 0 | 1.85 | 0 | 0 |
| 2Persons | 0 | 100 | 0 | 0 | 0 | 0 |
| 3Persons | 0 | 0 | 100 | 0 | 0 | 0 |
| Vehicle | 1 .85 | 0 | 0 | 98.15 | 0 | 0 |
| Truck | 0 | 0 | 0 | 0 | 100 | 0 |
| Clutter | 0 | 0 | 0 | 0 | 0 | 100 |

examined. Table 1 presents the confusion matrix of Greedy GMM based classifier with MFCC coefficients and MV post-processing after MAP decision for six class problem. Greedy GMM outperform GMM classifier. To improve classification accuracy, majority vote post-processing can be employed. The resulting effect is a smooth operation that removes spurious misclassification. Indeed, the classification rate improves to 99.08% for Greedy GMM after MAP decision following majority vote post-processing, 97.93% for GMM after MV decision.

## 5 Conclusion

Acoustics features like LPC, LPCC and MFCC are used to exploit the micro-Doppler signatures issued from moving target in order to provide separation among the target classes like humans, vehicles, trucks and clutter. Speech recognition techniques, using GMM and Greedy GMM including the MAP decision rules, have been successfully applied for ground surveillance radar. Experimental results show that the Greedy GMM using MFCC features gives the best classification performances. However, it fails to avoid all classification errors, which we are bound to eradicate through MV-post processing which guarantees a 99.08% classification rate for six-class problem presented in this work.

## References


1. Thayaparan, T., Abrol, S., Riseborough, E., Stankovic, L., Lamothe, D., Duff, D.: Analysis of radar micro-Doppler signatures from experimental helicopter and human data. IEE Proc. Radar Sonar Navigation 1(4), 288–299 (2007)
2. Natecz, M., Rytel-Andrianik, R., Wojtkiewicz, A.: Micro-Doppler analysis of signal received by FMCW radar. In: International Radar Symposium, Germany (2003)
3. Boashash, B.: Time frequency signal analysis and processing, 1st edn. Elsevier Ltd. (2003)
4. Chen, V.C., Ling, H.: Time frequency transforms for radar imaging and signal analysis. Artech House, Boston (2002)
5. Anderson, M., Rogers, R.: Micro-Doppler analysis of multiple frequency continuous wave radar signatures. In: SPIE Proc. Radar Sensor Technology, vol. 654 (2007)
6. Chen, V.C.: Analysis of radar micro-Doppler signature with time-frequency transform. In: Proc. Tenth IEEE Workshop on Statistical Signal and Array Processing, pp. 463–466 (2000)
7. Reynolds, D.A., Quatieri, T.F., Dunn, R.B.: Speaker verification using adapted Gaussian mixture models. Digit. Signal Process. 10, 19–41 (2000)
8. Li, J.Q., Barron, A.R.: Mixture density estimation. In: Advances in Neural Information Processing Systems, vol. 12. MIT Press, Cambridge (2000)
9. Bilik, I., Tabrikian, J., Cohen, A.: GMM-based target classification for ground surveillance Doppler radar. IEEE Trans. on Aerospace and Electronic Systems 42(1), 267–278 (2006)
10. Verbeek, J.J., Vlassis, N., Krose, B.: Efficient greedy learning of Gaussian mixture models. Neural Computation 5(2), 469–485 (2003)