J. Lundén, L. Terho and V. Koivunen, Classifying pulse compression radar waveforms using time-frequency distributions, in Proceedings of the 39th Annual Conference on Information Sciences and Systems (CISS), Baltimore, USA, March 16-18, 2005.

# Classifying Pulse Compression Radar Waveforms Using Time-Frequency Distributions

Jarmo Lundén, Liisa Terho[1] and
Visa Koivunen
Signal Processing Laboratory
SMARAD CoE
Helsinki University of Technology
P.O. Box 3000
FIN-02015 HUT, Finland
e-mail: jrlunden@wooster.hut.fi

*Abstract* — **In this paper the classification of pulse compression radar waveforms using features extracted from the Choi-Williams time-frequency distribution is studied. In addition, a feature based on the symmetry properties of polyphase waveforms is introduced. The pulse compression waveforms examined are the Frank, P1, P2, P3, and P4 codes. The discrimination capability of the features is evaluated using an ensemble averaging early-stop committee of 10 multilayer perceptrons. The classifier achieves over 96 % overall correct classification rate in signal-to-noise ratio of 3 dB on data similar to the training data.**

## I. Introduction

Automatic waveform recognition is an important task in spectrum management and surveillance, communication and radar emitter identification, electronic support applications as well as in software defined radio (SDR) and intercept receivers. A key task in waveform recognition is to find a set of features that distinguish various waveforms from each other. In addition, computationally efficient algorithms for computing the features have to be derived.

Recognition of radar waveforms using time-frequency analysis has received attention during the past few years, e.g. [1–3]. In this paper a method for recognizing the polyphase pulse compression waveforms using features derived from the Choi-Williams time-frequency distribution (CWD) and symmetry properties of radar signals is proposed. The polyphase pulse compression waveforms considered are the Frank, P1, P2, P3, and P4 codes. The actual supervised classifier is an ensemble averaging early-stop committee of 10 multilayer perceptrons (MLPs). The CWD is treated as an image and features that possess desirable invariance properties are computed from those images.

The paper is organized as follows. The signal model is given in Section II. The polyphase waveforms are introduced in Section III. The CWD representation is described in Section IV. Normalization procedure needed as a pre-processing step is outlined in Section V. The actual features are introduced in Section VI and simulation results are presented in Section VII.

[1]Liisa Terho is with the Finnish Defence Forces Technical Research Centre.

## II. Signal model

The channel is assumed to be additive white Gaussian noise (AWGN) channel. In addition, it is assumed that the signal has been transferred to the baseband, i.e. the complex envelope $y(t)$ of the received signal is given by:

$$y(t) = x(t) + w(t), \tag{1}$$

where $x(t)$ is the complex envelope of the transmitted signal containing only one code period and $w(t)$ is a complex circular additive white Gaussian noise process.

## III. Polyphase waveforms

The transmitted complex phase coded signal can be expressed in the following form:

$$s(t) = Ae^{j(2\pi f_c t + \phi_i)}, \tag{2}$$

where $A$ is the amplitude, $f_c$ is the carrier frequency, and $\phi_i$ is the discrete phase sequence. Each phase has the same time duration. In the following, five different polyphase pulse compression codes are presented. The codes presented are the Frank, P1, P2, P3, and P4 codes.

The Frank code is a step approximation to a linear frequency modulation (LFM) waveform using $N$ frequency steps and $N$ samples per frequency. Thus, the total number of samples in a Frank code is $N^2$. The phase of the $i$th sample of the $j$th frequency of a Frank code is given by [4]:

$$\phi_{i,j} = \frac{2\pi}{N}(i-1)(j-1), \tag{3}$$

where $i = 1, 2, \ldots, N$, and $j = 1, 2, \ldots, N$. The pulse compression ratio of the Frank code is $N^2$.

The P1 code is also derived from a step approximation to an LFM waveform using $N$ frequency steps and $N$ samples per frequency. The phase of the $i$th sample of the $j$th frequency of a P1 code is given by [4]:

$$\phi_{i,j} = -\frac{\pi}{N}[N-(2j-1)][(j-1)N+(i-1)], \tag{4}$$

where $i = 1, 2, \ldots, N$, and $j = 1, 2, \ldots, N$. The P1 code has a pulse compression ratio of $N^2$.

The phase of the $i$th sample of the $j$th frequency of a P2 code is given by [4]:

$$\phi_{i,j} = \left[\frac{\pi}{2}\cdot\frac{N-1}{N} - \frac{\pi}{N}(i-1)\right](N+1-2j), \tag{5}$$

where $i = 1, 2, \ldots, N$, and $j = 1, 2, \ldots, N$, and where $N$ is even. For odd values of $N$ the autocorrelation sidelobes are high [4]. The pulse compression ratio of the P2 code is $N^2$. The P2 code has the property of being palindromic.

The P3 code is derived by sampling an LFM waveform. The phase of the $i$th sample of a P3 code is given by [4]:

$$\phi_i = \frac{\pi}{\rho}(i-1)^2, \tag{6}$$

where $i = 1, 2, \ldots, \rho$ where $\rho$ is the pulse compression ratio.

The P4 code is derived from the same waveform as the P3 code. The phase of the $i$th sample of a P4 code is given by [4]:

$$\phi_i = \frac{\pi}{\rho}(i-1)^2 - \pi(i-1), \tag{7}$$

where $i = 1, 2, \ldots, \rho$ where $\rho$ is the pulse compression ratio.

## IV. Choi-Williams distribution

The Choi-Williams distribution of a continuous time signal $x(t)$ is defined [5]:

$$W(t,\omega) = \iint \frac{1}{\sqrt{4\pi\tau^2/\sigma}} \exp\left(-\frac{(\mu-t)^2}{4\tau^2/\sigma}\right) \cdot x\left(\mu+\frac{\tau}{2}\right) x^*\left(\mu-\frac{\tau}{2}\right) \exp\left(-j\omega\tau\right) d\mu d\tau, \tag{8}$$

where $\sigma$ ($\sigma > 0$) is a scaling factor.

Fig. 1 shows the Choi-Williams distributions of different polyphase codes. The differences among the classes are clear from the figure. For instance, the Frank, P1, and P2 codes have a distinctive blocky structure due to the fact these codes are derived from a step approximation to a linear frequency modulation waveform. The P3 and P4 codes, on the other hand, are derived by sampling a linear frequency modulation waveform, and thus have a smoother signal component in the CWD. The shape of the CWD remains relatively similar within the code classes. The largest difference being the number of frequency blocks in the CWDs of the Frank, P1, and P2 codes. The number depends on the length of the code. In addition, the sampling frequency and the signal bandwidth have an effect on the shape of the CWD.

The scaling factor $\sigma$ had a value 0.05 when calculating the CWDs in Fig. 1. This relatively small value was experimentally found to produce a good separation among the different classes. It provides effective cross-term suppression but smearing of the auto-terms is considerable. However, since image processing techniques are applied to the CWD image, it is preferred that the signal objects are a little bit larger. Hence, a small value for $\sigma$, such as 0.05, is reasonable.

The CWD can be treated as a 2-D image. However, in order to effectively extract features from the CWD image, it has to be normalized first. The normalization procedure is explained next.

## V. Normalization of the CWD image

The goal of the normalization is to minimize the influence of the signal bandwidth and the sampling frequency on the final CWD image. The normalization procedure consists of three steps:

1) Thresholding of the CWD image
2) Time gating and frequency filtering of the thresholded image, i.e. removal of areas not containing signal components from the edges of the image

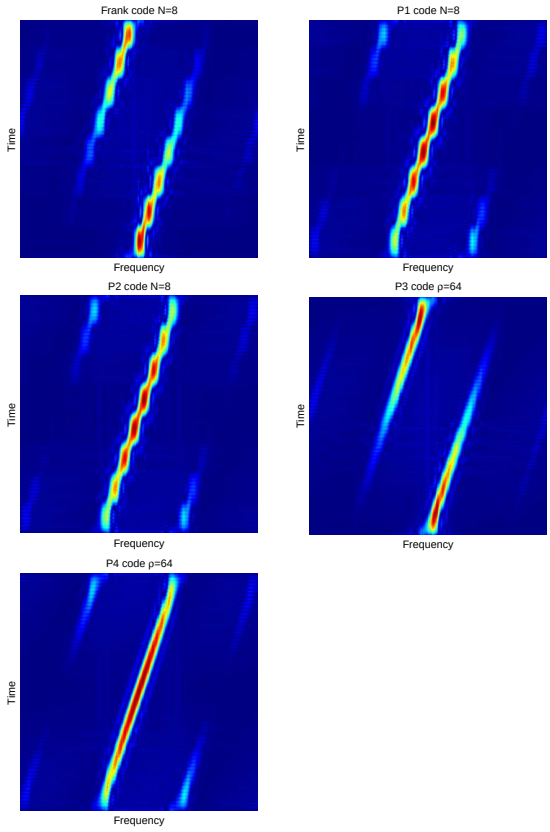


Fig. 1: Choi-Williams distributions of different polyphase coded signals. Pulse compression ratio of all signals is 64. The scaling factor $\sigma$ had a value 0.05.

Fig. 2: Three normalization steps: 1) thresholding, 2) time gating and frequency filtering, and 3) aspect ratio normalization.

3) Aspect ratio normalization of the final binary image to 1.

Fig. 2 illustrates the three normalization steps for a P3 coded waveform.

The thresholding step is crucial to the overall success of the whole normalization algorithm. The thresholded image should contain only the signal objects without any isolated noise points since the outcome of the second step can be very sensitive to such noise. The threshold value plays a key role in the outcome of the first step. The following iterative algorithm was used to obtain a global threshold $T$ [6]:

1. Select an initial estimate for $T$.
2. Segment the image using $T$ to two groups $G_1$ and $G_2$ where $G_1$ consists of all pixels with gray level values

$> T$ and $G_2$ consists of all pixels with gray level values $\leq T$.

3. Compute the average gray level values $\mu_1$ and $\mu_2$ for the pixels in regions $G_1$ and $G_2$.
4. The new threshold value is given by $T = \frac{1}{2}(\mu_1 + \mu_2)$.
5. Repeat steps 2 through 4 until convergence.

The initial estimate for $T$ can be assigned to the average of the maximum and minimum gray level values of the CWD image.

Global thresholding alone, however, cannot guarantee complete removal of isolated noise. Thus, before time gating and frequency filtering the thresholded image should be treated to remove harmful noise points. The desired result can be achieved, for example, with a combination of morphological *opening* and labeling of the binary objects followed by a removal of objects that are not of sufficient size (e.g. at least 10 % of the size of the largest object). The morphological opening operator (i.e. *erosion* followed by *dilation*) smoothens the image objects and removes the possibly occurring horizontal and vertical lines (especially in low signal-to-noise ratios) caused by the form of the Choi-Williams kernel, see equation (8). The mask used in the morphological operations was a $3 \times 3$ square mask. If the threshold used in removing small objects is chosen high enough (e.g. 10 % of the size of the largest object), the removal operation should also eliminate the possible side-terms that can be observed in the CWDs of the P1, P2, and P4 codes (see Fig. 1). Existence of these sideterms in the final binary CWDs can be detrimental to the discrimination capability of the features.

In the second step of the normalization procedure, areas not containing signal objects are removed from the edges of the image. And finally, in the third step the aspect ratio of the binary image is normalized to 1 by resizing the image. In this paper the *nearest neighbor interpolation* method was employed. The size of the resized binary image is $M \times M$ where $M$ is the minimum dimension of the image resulting from the second step of the normalization procedure.

## VI. Features

The features that have been used in this study consist of a number of pseudo-Zernike moments, three other features targeting specific properties observed in the CWDs, and a feature based on the differences in the symmetry properties of the polyphase codes.

The pseudo-Zernike moments have been used for object recognition in images (e.g. [7]). The pseudo-Zernike moments that have been used in this study are invariant to translation, scaling, rotation, and mirroring. Consequently, they are very suitable to the problem in question.

The pseudo-Zernike moments are invariant to translation and scaling since they are calculated using scaled geometric moments. The geometric moments of order $p+q$ of a digital image $f(x,y)$ are defined as:

$$m_{pq} = \sum_x \sum_y f(x,y) x^p y^q. \tag{9}$$

The translation and scale invariant central geometric moments are defined as [7]:

$$G_{pq} = \frac{1}{m_{00}^{(p+q+2)/2}} \sum_x \sum_y f(x,y)(x-\bar{x})^p (y-\bar{y})^q, \tag{10}$$

where

$$\bar{x} = \frac{m_{10}}{m_{00}}, \quad \bar{y} = \frac{m_{01}}{m_{00}}. \tag{11}$$

The translation and scale invariant radial-geometric moments are defined as [7]:

$$R_{pq} = \frac{1}{m_{00}^{(p+q+3)/2}} \sum_x \sum_y f(x,y)(\tilde{x}^2 + \tilde{y}^2)^{1/2} \tilde{x}^p \tilde{y}^q, \tag{12}$$

where $\tilde{x} = x - \bar{x}$ and $\tilde{y} = y - \bar{y}$.

The pseudo-Zernike moments of order $n$ with repetition $m$ ($|m| \leq n$) can be computed using the translation and scale invariant central geometric and radial-geometric moments with the following equation [7, 8]:

$$\begin{aligned} Z_{nm} =& \frac{n+1}{\pi} \sum_{\substack{s=0 \\ n-s-m=even}}^{n-|m|} D_{nms} \sum_{a=0}^{k} \sum_{b=0}^{m} \\ & (-j)^b \binom{k}{a} \binom{m}{b} B_{nms} G_{2k-2a+m-b,2a+b} \\ & + \frac{n+1}{\pi} \sum_{\substack{s=0 \\ n-s-m=odd}}^{n-|m|} D_{nms} \sum_{a=0}^{d} \sum_{b=0}^{m} \\ & (-j)^b \binom{d}{a} \binom{m}{b} B_{nms} R_{2d-2a+m-b,2a+b}, \end{aligned} \tag{13}$$

where

$$k = \frac{n-s-m}{2} \tag{14}$$

$$d = \frac{n-s-m-1}{2} \tag{15}$$

and

$$B_{nms} = (-1)^s \frac{(n-s)!}{s! \left(\frac{n+|m|}{2} - s\right)! \left(\frac{n-|m|}{2} - s\right)!} \tag{16}$$

$$D_{nms} = (-1)^s \frac{(2n+1-s)!}{s!\,(n-|m|-s)!\,(n+|m|+1-s)!}. \tag{17}$$

Other methods for obtaining scale invariant pseudo-Zernike moments have been reviewed in [9].

Rotation invariance is achieved by taking the absolute value of the pseudo-Zernike moments $Z_{nm}$ [8]. The dynamic range can be reduced by taking the logarithm which gives the final features:

$$\hat{Z}_{nm} = \log_e |Z_{nm}|. \tag{18}$$

The following pseudo-Zernike moments of the binary CWD image were selected as features: $\hat{Z}_{20}$, $\hat{Z}_{22}$, $\hat{Z}_{30}$, $\hat{Z}_{31}$, $\hat{Z}_{32}$, $\hat{Z}_{33}$, and $\hat{Z}_{43}$.

Number of image objects in the binary image is a distinctive feature. Provided the normalization has succeeded, the Frank and P3 codes have two signal objects in the binary CWD image while the other polyphase codes have only one. To increase the robustness of the feature, objects smaller than 20 % of the size of the largest object are discarded.

Another distinctive feature is the time location of the peak power in the CWD. The P1, P2, and P4 codes have the peak power relatively close to the center of the code while the Frank and P3 codes have the highest powers in the ends of the code. This feature cannot be calculated from the binary image, and

thus does not require the full normalization of the CWD image. Merely time gating is required. The feature is calculated from the CWD image $W_{CW}(x, y)$:

$$t_{\max} = \frac{1}{N-1} \arg\max_x \{W_{CW}(x, y)\}, \tag{19}$$

where $x$ is the time axis, $y$ the frequency axis, and $N$ is the time axis length of $W_{CW}(x, y)$. The division by $N - 1$ normalizes the value of the feature between 0 and 1.

The final CWD feature is targeting the blocky structure observed in the CWDs of the Frank, P1, and P2 codes. It measures the standard deviation of the width of the objects in the binary image. This feature distinguishes the P3 and P4 codes from the rest of the polyphase codes. The feature is calculated as follows.

First the objects of the image are labeled. Then the procedure explained next is done for each one of the objects in the binary image separately, i.e. by first masking the rest of the objects away such that the image contains only the desired object. The procedure is done separately for each image object because the principal components that will be calculated would not be the ones desired for some code classes (i.e. the Frank and P3 codes) if they would be calculated from the image containing all the objects.

The principal components of the binary CWD image are calculated. The principal components of a binary image $B(x, y)$ are the eigenvectors of the covariance matrix which can be calculated as:

$$\boldsymbol{C} = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} (\boldsymbol{z} - \boldsymbol{c})(\boldsymbol{z} - \boldsymbol{c})^T B(x, y), \tag{20}$$

where the size of the image is $N \times N$, and

$$\boldsymbol{z} = (x, y)^T, \quad \boldsymbol{c} = (\bar{x}, \bar{y})^T, \tag{21}$$

where $\bar{x}$ and $\bar{y}$ define the center of the image, and can be calculated using (11).

Then the image is rotated so that the principal axes are parallel to the horizontal and vertical axes of the image. Because of the discrete pixel locations, rotation requires interpolation. The nearest neighbor interpolation method was employed again.

After rotation, the standard deviation of the width of the image object can be calculated. This is done by calculating the row or the column sum of the image depending on which direction the image is rotated. In the following it is assumed that the image is rotated such a way that the first principal axis corresponding to the highest principal component is parallel to the vertical axis. That is, the row sum is calculated as:

$$r(x) = \sum_{y=0}^{N-1} \hat{B}(x, y), \quad x = 0, 1, \ldots, N-1, \tag{22}$$

where $\hat{B}(x, y)$ is the rotated binary image.

Then $r(x)$ is normalized between 0 and 1:

$$\hat{r}(x) = \frac{r(x)}{\max r(x)}. \tag{23}$$

Finally, the standard deviation of the width of the image object in the binary CWD image is defined as:

$$\sigma_{\text{obj}} = \sqrt{\frac{1}{M} \sum_x \hat{r}^2(x) - \left( \frac{1}{M} \sum_x \hat{r}(x) \right)^2}, \tag{24}$$

Tab. 1: Theoretical values of the time lag of the maximum cross-correlation between pulse and time-reversed pulse. The values have been tested to hold at least for codes $N$ = 3–100 (Frank, P1, P2), and $\rho$ = 3–200 (P3, P4).

| **Frank** | **P1** | **P2** | **P3** | **P4** |
|---|---|---|---|---|
| N+1 | 1 | 0 | 1 | 1 |

where the sums are taken over non-weak samples, i.e. $\hat{r}(x) \geq T_{\text{obj}}$, and $M$ is the number of non-weak samples. The weak samples are discarded since they influence the estimate heavily. Especially the rows (or columns depending on rotation) that do not contain any signal have large impact on the estimate. The value used for the threshold $T_{\text{obj}}$ was 0.3, i.e. 30 % of the maximum of $\hat{r}(x)$. The final feature value is the average of $\sigma_{\text{obj}}$ over the image objects.

Early classification results using the features derived from the Choi-Williams distribution indicated that the P1 and P2 codes could not be reliably distinguished using only these features. Hence, a new feature based on the different symmetry properties of the P1 and P2 codes was developed. The feature calculates the cross-correlation between a symbol rate sampled pulse and its time-reversed counterpart. The time lag of the maximum of the above cross-correlation is a distinctive feature among the polyphase codes. The feature is calculated as follows.

Denoting a single symbol rate sampled discrete time complex envelope code period with $y(n)$, $n = 0, 1, \ldots, N-1$, cross-correlation between pulse and time-reversed pulse is given by:

$$\hat{r}_y(\tau) = \begin{cases} \sum_{n=0}^{N-\tau-1} y(n+\tau) y^*(N-1-n), & \tau \geq 0, \\ \sum_{n=-\tau}^{N-1} y(n+\tau) y^*(N-1-n), & \tau < 0, \end{cases} \tag{25}$$

for all $|\tau| \leq N - 1$. The final feature is the time lag of the maximum cross-correlation:

$$\tau_{\max} = \arg\max_\tau |\hat{r}_y(\tau)|. \tag{26}$$

The feature is invariant to constant rotation (i.e. each sample $y(n)$, $n = 0, 1, \ldots, N-1$ is rotated the same amount). In order to get the same feature value for time-reversed pulses, the absolute value of $\tau_{\max}$ can be used as a feature.

Table 1 lists the theoretical values for different polyphase classes. The feature requires symbol rate sampling (i.e. sampling at the subpulse frequency). The symbol rate can be estimated with a cyclic correlation based symbol rate estimator, e.g. [10].

## VII. Simulation results

In order to test the discrimination capability of the features, a supervised classifier was employed. The purpose of the classifier is to classify the received radar signal based on the pulse compression waveform of the signal to five different classes: the Frank, P1, P2, P3, and P4 codes. The classifier was an ensemble averaging early-stop committee of 10 multilayer perceptrons (i.e. the output of the committee is the average of the outputs of the committee members) with different partitions of the data to the training and validation sets for each MLP. The validation sets consisted of 10 % of the original training

Tab. 2: Simulation parameters.

| Parameter | Parameter value |
|---|---|
| **General parameters** | |
| Sampling frequency | 12000 Hz |
| **Polyphase waveforms** | |
| Carrier frequency $f_c$ | Uniform(3000,4000) Hz |
| Subpulse frequency | $\frac{1}{4}f_c$ when $N \in [3,7]$ |
| | $\frac{1}{2}f_c$ when $N \in [8,10]$ |
| Frank, P1 codes $N$ | 3–10 |
| P2 codes $N$ | 4, 6, 8, 10 |
| P3, P4 codes $\rho$ | $N^2$ where $N \in [3,10]$ |

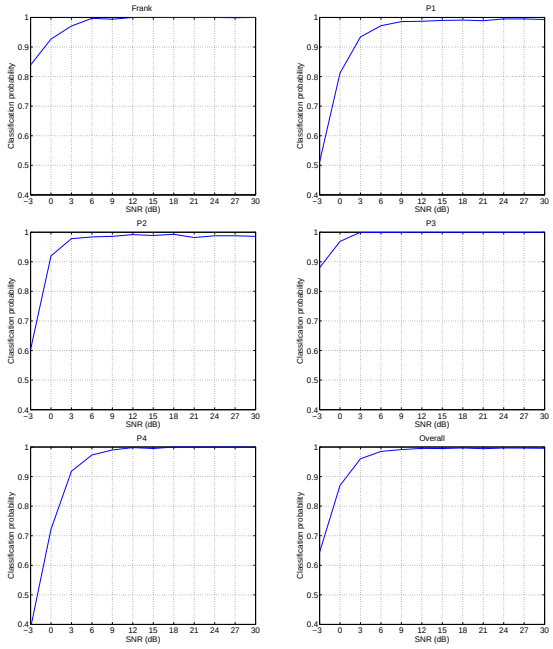


Fig. 3: Classification performance as a function of the SNR on data similar to the training data. The overall correct classification rate exceeds 96 % already in SNR of 3 dB.

data. Each MLP had one hidden layer with 30 hidden neurons. The number of inputs was the number of features and the number of outputs was the number of classes. The hidden and output layer activation functions were the hyperbolic tangent and the softmax activation functions, respectively. The cross-entropy error function was used. The training algorithm employed was the scaled conjugate gradient algorithm [11].

A large set of different codes from each class were generated for both training and testing purposes. In total 1000 signals from each class for signal-to-noise ratios (SNRs) of -3, 0, ..., 30 dB were generated for testing purposes. Table 2 lists the values of the parameters used for generating the training and testing signals.

The classifier performance was measured as a function of the SNR in AWGN channel given in (1). Fig. 3 depicts the performance of the classifier on data similar to the training data. The training data consisted of 600 samples from each class with different SNRs between 0 and 20 dB.

Tab. 3: Confusion matrix for SNR of 3 dB. The overall correct classification rate was 96.0 %.

| | **Frank** | **P1** | **P2** | **P3** | **P4** |
|---|---|---|---|---|---|
| **Frank** | 97.1 | 0.1 | 0.5 | 1.8 | 0.5 |
| **P1** | 0 | 93.4 | 0.2 | 0.5 | 5.9 |
| **P2** | 0.1 | 2.0 | 97.8 | 0 | 0.1 |
| **P3** | 0 | 0 | 0 | 100 | 0 |
| **P4** | 0.5 | 7.4 | 0.1 | 0.2 | 91.8 |

Fig. 3 indicates that the classifier performs reliably: the overall correct classification rate is over 96 % and the correct classification rates of the invidual classes are well over 90 % in SNR of 3 dB. However, some minor confusion still exists even at high SNRs. Approximately 1–2 % of the P2 codes are incorrectly classified to the P1 codes at high SNR regime. The confusion is caused by the errors made in symbol rate estimation. In addition, roughly 1 % of the P1 codes are incorrectly classified to the P4 codes at high SNRs.

Some of the above confusion might be removed by increasing the number of training signals. Inspection of the feature values indicated that there are clear differences in the values of the time-frequency distribution features even within the classes. Hence, the training data has to be as diverse as possible to achieve the best possible classification performance.

Table 3 reports the classification percentages in SNR of 3 dB. The table indicates that discrimination between the P1 and P4 codes is most difficult.

## VIII. Conclusion

In this paper the classification of polyphase pulse compression radar waveforms using the Choi-Williams time-frequency distribution was studied. The statistics of the CWD were found to be very suitable for the task. However, the P1 and P2 codes could not be distinguished reliably using only the Choi-Williams time-frequency distribution (see the CWDs from Fig. 1). Hence, an additional feature dichotomizing between the P1 and P2 codes was introduced. The feature is based on the different symmetry properties of the polyphase codes.

The discrimination capability of the features was evaluated using an ensemble averaging early-stop committee of 10 MLPs. The classifier performed reliably: the overall correct classification rate exceeded 96 % in SNR of 3 dB on data similar to the training data.

## References


[1] P. E. Pace, *Detecting and classifying low probability of intercept radar*, Artech House, 2004.

[2] G. López-Risueño, J. Grajal, and O. Yeste-Ojeda, "Atomic decomposition-based radar complex signal interception," *IEE Proc. – Radar Sonar Navig.*, vol. 150, no. 4, pp. 323–331, Aug. 2003.

[3] G. López-Risueño, J. Grajal, O. A. Yeste-Ojeda, Á. Sanz-Osorio, and J. A. Moreno, "Two digital receivers based on time-frequency analysis for signal interception," in *Proc. Int. Radar Conf. 2003*, Sep. 2003, pp. 394–399.

[4] B. L. Lewis, F. F. Kretschmer, Jr., and W. W. Shelton, *Aspects of radar signal processing*, Artech House, Inc., 1986.

[5] H.-I. Choi and W. J. Williams, "Improved time-frequency representation of multicomponent signals using exponential kernels," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 37, no. 6, pp. 862–871, Jun. 1989.

[6] R. C. Gonzalez and R. E. Woods, *Digital image processing*, Prentice-Hall, Inc., 2002.

[7] M. Dehghan and K. Faez, "Farsi handwritten character recognition with moment invariants," in *Proc. 13th Int. Conf. Digital Signal Processing*, Jul. 1997, vol. 2, pp. 507–510.

[8] R. B. Bailey and M. Srinath, "Orthogonal moment features for use with parametric and non-parametric classifiers," *IEEE Trans. Pattern Anal. Machine Intell.*, vol. 18, no. 4, pp. 389–399, Apr. 1996.

[9] C.-W. Chong, P. Raveendran, and R. Mukundan, "The scale invariants of pseudo-Zernike moments," *Pattern Anal. & Applicat.*, vol. 6, no. 3, pp. 176–184, 2003.

[10] P. Ciblat, P. Loubaton, E. Serpedin, and G. B. Giannakis, "Asymptotic analysis of blind cyclic correlation-based symbol-rate estimators," *IEEE Trans. Inform. Theory*, vol. 48, no. 7, pp. 1922–1934, Jul. 2002.

[11] M. F. Møller, "A scaled conjugate gradient algorithm for fast supervised learning," *Neural Networks*, vol. 6, no. 4, pp. 525–533, 1993.