

CLASSIFICATION IN THE GABOR TIME–FREQUENCY DOMAIN OF NON–STATIONARY SIGNALS EMBEDDED IN HEAVY NOISE WITH UNKNOWN STATISTICAL DISTRIBUTION

EWA ŚWIERCZ

Faculty of Electrical Engineering
Białystok Technical University, ul. Wiejska 45D, 15–351 Białystok, Poland
e-mail: ewasw@pb.edu.pl

A new supervised classification algorithm of a heavily distorted pattern (shape) obtained from noisy observations of non-stationary signals is proposed in the paper. Based on the Gabor transform of 1-D non-stationary signals, 2-D shapes of signals are formulated and the classification formula is developed using the pattern matching idea, which is the simplest case of a pattern recognition task. In the pattern matching problem, where a set of known patterns creates predefined classes, classification relies on assigning the examined pattern to one of the classes. Classical formulation of a Bayes decision rule requires *a priori* knowledge about statistical features characterising each class, which are rarely known in practice. In the proposed algorithm, the necessity of the statistical approach is avoided, especially since the probability distribution of noise is unknown. In the algorithm, the concept of discriminant functions, represented by Frobenius inner products, is used. The classification rule relies on the choice of the class corresponding to the *max* discriminant function. Computer simulation results are given to demonstrate the effectiveness of the new classification algorithm. It is shown that the proposed approach is able to correctly classify signals which are embedded in noise with a very low SNR ratio. One of the goals here is to develop a pattern recognition algorithm as the best possible way to automatically make decisions. All simulations have been performed in Matlab. The proposed algorithm can be applied to non-stationary frequency modulated signal classification and non-stationary signal recognition.

Keywords: non-stationary signals, signal classification, pattern recognition, time-frequency transforms.

1. Introduction

The term “classification” has two distinct meanings. We may be given a set of observations with the aim of establishing the existence of classes or clusters in the data. On the other hand, we may know that there are many classes, and the aim is to establish a rule whereby we can classify a new observation on the basis of observed *attributes* or *features* into one of the existing classes. The former type is known as *unsupervised learning* (or *clustering*), the latter as *supervised learning*. The construction of a classification procedure from a set of data for which the true classes are known has also been variously termed as *pattern recognition* or *discrimination*. Generally, pattern recognition is a science for the description or classification/recognition of measurements interrelated with statistical, syntactical or structural and neural approaches (Bishop, 2006; Duda *et al.*, 2001; Fukunaga, 1990; Jain *et al.*, 2000; McLachlan, 1992).

Despite different kinds of classifiers (statistical or

non-statistical), a uniform formula of classification can be described by a set of discriminant functions $g_i(x)$, $i = 1, \dots, K$, where x is an element of a feature space extracted from observations. For each point x in a feature space, an i -th class is chosen with the corresponding largest (smallest) discriminant measure $g_i(x)$ calculated from some optimisation algorithm.

In the statistical approach with a feature vector x and two kinds of conditional probability functions, discriminant functions $g_i(x)$ can be formulated as

- the Bayes rule $g_i(x) = -\mathfrak{R}(\alpha_i|x)$, where \mathfrak{R} is a Bayesian risk,
- a maximum *a posteriori* (MAP) rule $g_i(x) = P(i|x)$,
- a maximum likelihood (ML) rule $g_i(x) = P(x|i)$,

with the following classification rule: Choose an i -th class for $\max g_i(x)$.

In the statistical approach, it is assumed that the form of a relevant probability distribution is known. Now, assume that the form of the discriminant function is known and that it is linear either in components or functions of x . In such cases, linear discriminants LDF are relatively easy to compute and analytically attractive. A linear discriminant function in the related literature (Duda *et al.*, 2001) is written in the form

$$g_i(x) = w_{i,0} + w_i^T x, \quad (1)$$

where x is a vector of features and w_i is a pattern vector.

The intercept $w_{i,0}$ can be absorbed by augmenting the vector x with an additional constant dimension, and $g_i(x)$ can be represented by the inner product of vectors $\langle \cdot, \cdot \rangle$,

$$\begin{aligned} g_i(x) &= w_0 + w_i^T x = [w_0 \ w_i^T] \begin{bmatrix} 1 \\ x \end{bmatrix} \\ &= a_i^T y = \langle a_i^T, y \rangle \end{aligned} \quad (2)$$

The selection of maximum $g_i(x)$ for $i = 1, 2, \dots, K$ can be proposed as a simple classification rule. This means that the feature space is divided into K classes and current data belong to the i -th class with $g_i(x)$ being the largest.

The generalisation of the inner product to matrices—the Frobenius inner product—is proposed as a discrimination function for matrices. The Frobenius inner product (denoted as $A : B$) is defined as

$$\begin{aligned} A : B &= \sum_i \sum_j A_{ij} B_{ij} \\ &= \text{trace}(A^T B) = \text{trace}(AB^T). \end{aligned} \quad (3)$$

The discrimination function $g_i(x)$ based on a matrix X of features and a matrix W_i of the pattern for an i -th class can be reformulated as

$$g_i(X) = W_i : X. \quad (4)$$

The selection of maximum $g_i(X)$ assigns data to the i -th class.

Shape matching is a method of classification especially useful in computer vision (Belongie *et al.*, 2002; Demirci *et al.*, 2007; Hagedoorn and Veltkamp, 1999; Jain *et al.*, 2000). Matching is a generic operation in pattern recognition which is used to determine the similarity between two entities (points, curves, or shapes) of the same type (Basri *et al.*, 1998; Gdalyahu and Weinshall, 1999; Latecki and Lakamper, 2000; Liu and Srinath, 1990; Umeyama, 1993; Younes, 1999). Features which are used for shape description can be very different, for example, algebraic moments, area, circularity, eccentricity, compactness, major axis orientation, Euler number, concavity tree, shape numbers (Jain *et al.*, 2000; Zhang

and Lu, 2003). The main difficulties are that many descriptions are sensitive to noise.

Similarity between two shapes can rely on comparing each point on the first shape with the “best” matching point on the second shape. Many similarity measures of shapes are based on the L_p distance between two points (x, y) (Fry, 1993; Hagedoorn and Veltkamp, 1999; Santini and Jain, 1999; Veltkamp, 2001). The L_p distance can be treated as equivalent to a discriminant function. For two points (x, y) in \mathbb{R}^k , the L_p distance is defined as

$$L_p(x, y) = \left(\sum_{i=0}^k |x_i - y_i|^p \right)^{1/p}. \quad (5)$$

This is also often called the Minkowski distance. For $p = 2$, this yields the Euclidean distance L_2 , while for $p = 1$, we get the Manhattan, city block, or taxicab distance L_1 .

Typically, two sets of points A (points on the first shape) and B (points on the second shape) are of a different size, so that no one-to-one correspondence exists between all points. In that case, a dissimilarity measure that is often used is the Hausdorff distance (Santini and Jain, 1999; Veltkamp, 2001; Zhang and Lu, 2004). The Hausdorff distance is defined not only for finite point sets, but it is also defined on non-empty closed bounded subsets of any metric space. The directed Hausdorff distance $\vec{h}(A, B)$ is defined as the lowest upper bound (supremum) over all points in A of the distances to B : with $d(a, b)$ the underlying distance, e.g., the Euclidean distance L_2 . The Hausdorff distance $H(A; B)$ is the maximum of $\vec{h}(A, B)$ and $\vec{h}(B, A)$:

$$H(A, B) = \max \left\{ \vec{h}(A, B), \vec{h}(B, A) \right\}, \quad (6)$$

where

$$\vec{h}(A, B) = \sup_{a \in A} \inf_{b \in B} d(a, b).$$

The Hausdorff distance is not translation, scale and rotation invariant (Zhang and Lu, 2004). Unfortunately, it is also very sensitive to noise and some modifications of this distance are introduced to get less sensitivity (Zhang and Lu, 2004), so that this distance is impractical in the problems considered in this paper.

One way to describe a shape is to locate a finite number of points, so-called landmarks, on the shape curve. The shape correspondence can be defined as matching from the set of landmarks on one shape to that on the next (Gdalyahu and Weinshall, 1999; Petrakis *et al.*, 2002). In (Xie *et al.*, 2008), the authors developed a mechanism to generate coarse segment matching between different instances of an object, based on representative skeletal features. This approach is especially useful in anatomical modelling and shape retrieval of living beings.

Most shape matching techniques concentrate much more on distortions coming from scaling, rotation and

shifting, but not on heavy noise. There are few papers describing the influence of random noise on deformation of a shape (Manay *et al.*, 2006; Sebe and Lew, 2002).

In this paper, shapes are the results of a two-argument transformation of 1-D signals, leading to 2-D representation. It is assumed that the signals considered are deterministic with a frequency modulation law. The single component of a non-stationary signal $s(t)$ is described by the relation

$$s(t) = \exp(j 2 \pi \int_0^t f(\tau) d\tau) = \exp(j 2 \pi \phi(t)), \quad (7)$$

where $f(t)$ is an instantaneous frequency (IF) (Auger *et al.*, 1996).

The received signal (an observation) $r(t)$ is modelled by

$$r(t) = s(t) + n(t) = \exp(j 2 \pi \int_0^t f(\tau) d\tau) + n(t), \quad (8)$$

where $n(t)$ is additive noise with an unknown statistical description.

Among other methods, 1-D non-stationary signals can be processed by time-frequency algorithms, which create a 2-D time-frequency structure on the plane (t, f) , understood as an individual image of a signal. It allows us to display changeable features of a signal (coming from its non-stationarity), which is not possible to be exposed in the original time domain or the Fourier domain. For mono-component, non-stationary signals coming from frequency modulations (FM), time-frequency transforms create a trajectory of IF on the (t, f) plane. Unfortunately, these transformations, because of their properties and numerical implementation, smear an ideal trajectory, creating a specific shape rather than an ideal curve. Numerical implementations of these transformations are represented by matrices corresponding to a displayed image, where each element of a matrix corresponds to a pixel of an image. This convention allows us to work with images in a way similar to working with any other type of matrix data. Therefore, an image and a matrix will be used interchangeably in this paper.

Here, mono-component signals with non-linear frequency modulation laws are addressed and the discrete Gabor transform is used as an example of 2-D transformation useful in developing a new classification algorithm, based on the pattern matching idea.

2. Selected topics of classification based on time-frequency representations

In many pattern recognition applications, features are traditionally extracted from standard quadratic time-frequency representations (TFRs) of a signal $r(t)$ from a

Cohen's class $C_r^\phi(t, f)$,

$$\begin{aligned} C_r^\phi(t, f) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \exp(j 2 \pi \nu(u-t)) \phi(\nu, \tau) \\ &\quad \cdot r(u + \frac{\tau}{2}) r^*(u - \frac{\tau}{2}) \exp(-j 2 \pi f \tau) d\nu du d\tau, \end{aligned} \quad (9)$$

where $\phi(\nu, \tau)$ is a two-dimensional function called a kernel of the representation defined in another time-frequency plane (ν, τ) and $r(t)$ is a transformed signal.

There exist several methods based on kernel optimisation of TFRs, which lead to a minimum classification error or a minimum probability of the error for non-stationary signal classification. In (Breakenridge and Mesbah, 2003), data-driven, time-frequency representations are proposed based on kernel optimisation (an exponential kernel and a Gaussian-shaped kernel), which lead to the minimum classification error for non-stationary signal classification. The optimisation of the parameterised kernel, which best separates the classes based on the available training signals, is governed by an objective function. In (Breakenridge and Mesbah, 2003), an estimated probability of the error and Fisher's discriminant ratio (FDR) are considered as objective functions. The classifier uses a properly defined distance d (Basseville, 1989; Doncarli *et al.*, 2001). The proposed time-frequency (TF) classifier uses the following decision rule.

The observation $r(t)$ is assigned to the i -th class with the rule

$$i = \arg \min_{i=1, \dots, K} d(C_r^\phi, \bar{C}_i^\phi), \quad (10)$$

where \bar{C}_i^ϕ is a representative TFR characterising the i -th class and C_r^ϕ is a real-valued TFR from Cohen's class. Several TF classifiers use the general decision rule based on Eqn. (10) (Richard and Lengell, 1999; Davy and Doncarli, 1998), or a similar rule (Flandrin, 1988). In (Heitz, 1995; Richard and Lengell, 1999), the classifier relies on TF correlations, the kernel parameters being optimised using a contrast criterion. The kernel types are smoothed pseudo-Wigner-Ville distribution and spectrogram, respectively.

In (Vincent *et al.*, 1994), TF distance measures are implemented, the TFR being the Wigner-Ville distribution. In (Gillespie and Atlas, 2001), a special class-dependent kernel is computed directly in the plane (τ, ν) by selecting discriminant (τ, ν) locations. The decision rule involves the Mahalanobis distance. In (Colas and Gelle, 2004), the L-Wigner distribution (LWD4) and the fourth-moment Wigner distribution (MWD4) as statistical expectation of LWD4 are defined. The classification scheme is a bank which contains a number of L MWD4—"energy compensated" detectors (to ensure normalisation) for which the outputs can be viewed as a special case of

the “minimum distance”,

$$d_i^2 = \sum_n \int_f (\text{MWD}_{4_r}(n, f) - \text{MWD}_{4_{\text{ref}_i}}(n, f))^2 df, \quad (11)$$

In the literature there are studies of time-frequency tools such as a Gabor filter or a wavelet transform which are used for extracting features from the texture of an image (Grigorescu *et al.*, 2002; Huang *et al.*, 2003; Kyrki *et al.*, 2004; Li and Shawe-Taylor, 2005; Tai, 2007). The classifier is trained based on some labelled texture features as the training set, used to classify unlabelled texture features of images into some pre-defined classes.

Here another time-frequency representation—the discrete Gabor transform (DGT)—is selected for the classification. The finite, discrete Gabor transform of a discrete signal $r \in \mathbb{C}^L$ (\mathbb{C}^L is the space of L -dimensional complex vectors) is given by a set of coefficients (Gröchenig, 2001; Sondergaard, 2006; Qian and Chen, 1993),

$$c(m, n) = \sum_{l=0}^{L-1} r(l) \text{win}^*(l - an) \exp(-2\pi j m l / M). \quad (12)$$

Here *win* is an analysis window that localises the signal in both time and frequency. We focus on the case when *win* and r are equally long (Sondergaard, 2006). Using the Gabor frame approach, we find an easy way to construct the inverse discrete Gabor transform (IDGT). The idea is to have analysis mapping and synthesis mapping, where the frames $(\text{win}_{m,n})$ and $(\gamma_{m,n})$ are dual (Gröchenig, 2001; Sondergaard, 2006; Werther *et al.*, 2005).

A finite, discrete Gabor frame is a family of functions $\text{win}_{m,n}$ which come from a reference function *win* by its translation in time and modulation in frequency,

$$\text{win}_{m,n} = \text{win}(l - na) \exp(2\pi j m l / M), \quad (13)$$

for $m = 1, \dots, M$ and $n = 1, \dots, N$, where $L = aN = bM$ for some $M \in \mathbb{N}, N \in \mathbb{N}$. The sampling (shift) parameters are also denoted in this paper as $b = \Delta M = L/M$ and $a = \Delta N = L/N$. If $MN > L$, the frame is redundant (oversampled). A large redundancy $R = L/(ab)$ is assumed, which gives better accuracy but requires greater computational efforts. The inverse Gabor transform with the synthesis window $\gamma_{m,n}$ is calculated according to the formula

$$r = \sum_{n=0}^{N-1} \sum_{m=0}^{M-1} c(m, n) \exp(2\pi j m l / M) \gamma(l - an). \quad (14)$$

The accuracy of the Gabor transform and the inverse Gabor transform depends on compact localisation of synthesis and analysis windows in time and in frequency. In experiments, a Gaussian window is chosen as the analysis window *win*.

3. Concept of a new classification scheme of deterministic, non-stationary signals based on the Gabor transform

The algorithm presented here is yet another proposal to resolve a noisy matching shape problem with a low SNR ratio for the classification task. 2-D noisy shapes come from time-frequency transformation of 1-D non-stationary data heavily distorted by random noise. Rigid template matching cannot be performed because of large unpredictable deformation and the loss of many points in noisy shapes compared with to non-noisy ones. That is why a more heuristic approach should be taken into account.

The complete, discrete multi-classification algorithm in terms of discriminant functions can be executed in the following way:

1. Initial step

- Establish the length L of an observation window (assumed time of observations) longer than the length of K examined signals $r_i, i = 1, \dots, K$. (This assumption allows the signal to appear at a random time moment in the observation window.)
- Establish a redundancy $R = L/(ab)$ of the time-frequency Gabor transformation, where $M = L/b$ is the number of frequency samples, $N = L/a$ is the number of time samples in the time-frequency plane.
- Establish K matrices $KM_{M \times N, i}$ for $i = 1, \dots, K$ by the Gabor transforming of K ideal non-noisy useful signals. (This step is equivalent to the training step for creating patterns in different classification algorithms.)
- Create a bank of K masks called also ideal patterns (ideal shapes) represented by the matrix $PM_{m \times n, i}$ based on selected sub-matrices of $KM_{M \times N, i}$ of lower dimensions $m_i \times n_i, i = 1, \dots, K$, containing only non-zero Gabor coefficients. Transform the matrix $PM_{m \times n, i}$ to binary form, replacing non-zero elements by ones.

2. Preparation of auxiliary matrices for classification

- Compute the Gabor transform of the received signal r to create the image on the time-frequency plane represented by the matrix $RM_{M \times N}$ of the received signal.
- Establish an auxiliary threshold th_{aux} lower than the maximum value of Gabor coefficients.
- Determine a secondary plane (the thresholding plane) located at th_{aux} , represented by the secondary matrix $SCM_{M \times N}$, leaving time-frequency coefficients crossing the secondary plane and setting the rest of coefficients to zero.

3. Preparation of the set of discriminant functions

- Find the best localisation of the useful signal (localisation is random because of a random moment of the signal arrival in the observation window) in the secondary plane by computing a few Frobenius inner products by slightly shifting the mask matrix $PM_{m \times n, 1}$ over the secondary matrix $SCM_{M \times N}$ in the neighbourhood of the biggest values (the most probable localisation of the useful signal). Only translations along the time axis and the frequency axis are allowed from the physical point of view. In the background, the matrix of the noisy mask $NM_{m \times n, 1}$ (the ideal mask distorted by heavy noise) is computed. Select the max Frobenius inner product denoted as MFP_1 . Repeat the same process for $i = 2, \dots, K$.
- Take the collection of $MFP_i, i = 1, \dots, K$ as the set of discriminant functions $MFP_i \Rightarrow g_i(r)$ treated as a two-dimensional extension of linear discrimination functions with weights represented by ideal pattern images and features represented by images of received signals. For simplicity, discriminant functions are further denoted as $g(r)$ instead of $g(X)$, where X is a matrix of Gabor coefficients of a received signal r .

4. Final step—classification

- Assign the received observation to the i -th class with the max $g_i(r)$.

Motivating applications of signals with different frequency modulations range from communications to a sonar and a radar. For example, a low probability of intercept radar (LPI) is designed to be difficult to detect. An LPI radar works with such frequency modulated signals which are acknowledged to be hard to detect. Unfortunately, such signals are secret. Signals with different frequency modulation laws are also met in the animal world. These signals are recorded by researchers and are rarely accessible. That is the main reason why that multi-classification has been limited to three accessible signals embedded in heavy noise generated in simulations in this paper.

In the 3-class classification task, three discrete signals with different lengths and approximately polynomial frequency modulation and hyperbolic frequency modulation laws were used. The non-noisy real radar signal RAD (the signal lasts 460 samples) was received by a radar receiver. The original non-noisy bat signal BAT (the signal has 1024 samples) with an approximately hyperbolic frequency modulation law was received by a specialised device and a synthetic signal with the exactly hyperbolic frequency modulation law called shortly HFM (the signal has 256 samples) was formed by a Matlab function,

according to the formula describing its spectral content, varying with time (Auger *et al.*, 1996),

$$HFM(n) = \exp\left(j 2\pi\left(f_0 n + \frac{c}{\log(|(n)|)}\right)\right). \quad (15)$$

These three accessible signals are assumed to be ideal in this paper and are used for forming ideal patterns.

The corresponding time representations of the signals considered are shown below. Additionally, as functions of time, the signals are randomly put in the observation window of the length $L = 1024$, simulating an unknown moment of the signal arrival during observation. All signals are normalised to the range $[-1, 1]$.

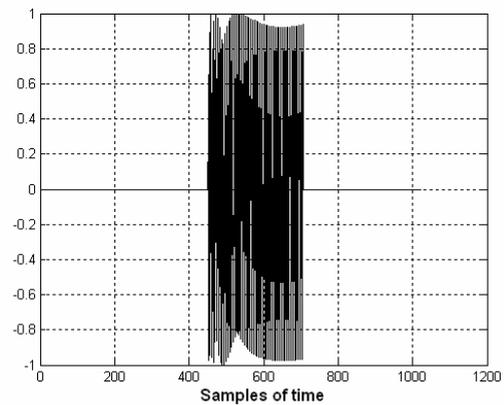


Fig. 1. Real part of the HFM 256-sampled signal.

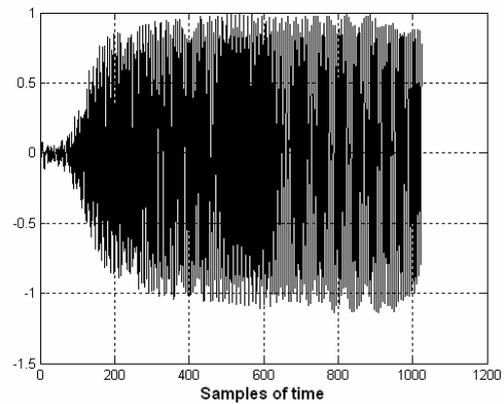


Fig. 2. Real 1024-sampled BAT signal.

Having three signals, three ideal shapes (ideal masks) are created from regions occupied by Gabor coefficients. These shapes are strongly dependent on the assumed values of redundancy R ($R = 128$ in the experiments) of the Gabor transform. In the numeric computational environment, the Gabor transform is represented by the matrix $KM_{M \times N}$ with dimension $M \times N = 512 \times 256$.

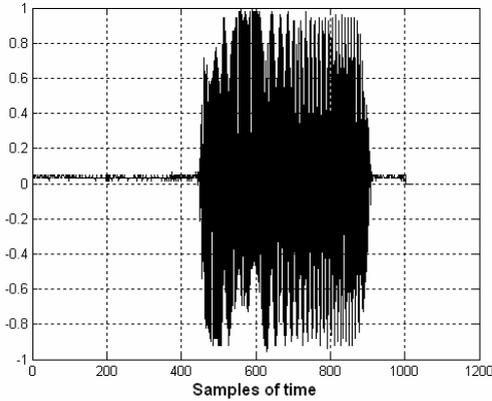


Fig. 3. Real 460-sampled RAD signal.

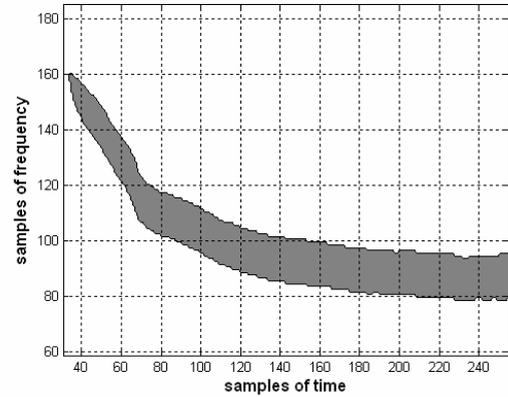


Fig. 6. Form of a non-noisy BAT signal for $R = 128$, $th_{ini} = 3.5$.

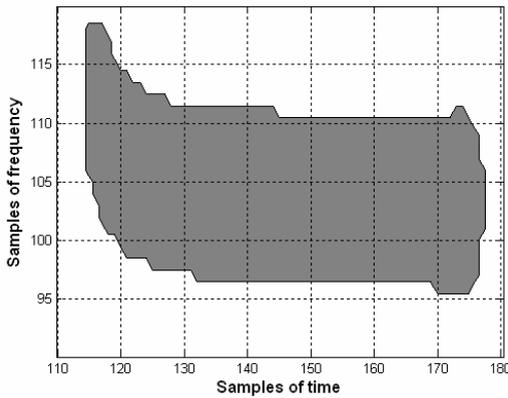


Fig. 4. Ideal shape of the HFM signal for $R = 128$, $th_{ini} = 3.5$.

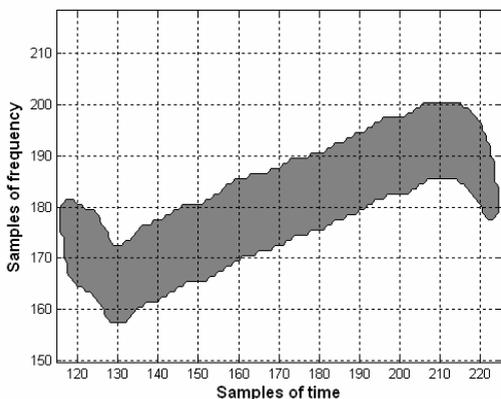


Fig. 5. Ideal shape of the non-noisy RAD signal for $R = 128$, $th_{ini} = 3.5$.

Images of patterns (sub-matrices) are limited to essential pixels for particular signals. The initial threshold $th_{ini} = 3.5$ is taken in order to reject negligible Gabor coefficients (caused, e.g., by computational inaccuracies).

Matrices $PM_{m \times n, i}$, are transformed to binary form by setting non-zero values to ones. In this way, a bank of three masks is created, represented numerically by matrices $PM_{m \times n, i}$, $i = 1, 2, 3$.

Next, signals are gradually emerged in noise with a decreasing SNR ratio, which causes the deterioration of the Gabor transformation. Noise with a Weibull distribution (a typical distribution for a radar clutter) is chosen as a model of disturbances. Noise modelled by Rayleigh, Rice or normal distributions behaves similarly. Random noise tends to spread its energy over the entire time-frequency domain, while signals concentrate their energy within limited time intervals and frequency bands (Auger *et al.*, 1996). Coefficients of the Gabor transform of a signal with additive noise are a sum of coefficients of the Gabor transform of a signal (signal Gabor coefficients) and coefficients of the Gabor transform of noise (noise Gabor coefficients), because of the linearity of this transform.

In the region occupied by signal coefficients in the ideal mask, we can only have signal coefficients distorted by noise coefficients. Outside this region, we have only noise Gabor coefficients. An example of the relation between the time observation for $SNR \approx -20$ dB and the ideal RAD signal in the background is visible in Fig. 7. For this value of SNR (noise + signal), Gabor coefficients and noise Gabor coefficients are utterly undistinguishable, which can be seen in Fig. 8. For comparison, the Gabor transform of the non-noisy RAD signal is shown in Fig. 9.

The BAT and RAD signals are real and their complex analytic form should be computed before computing the Gabor transform, otherwise the effect of repeating a negative spectral component appears (Fig. 9).

According to rules in the complete classification algorithm, a secondary plane (the thresholding plane) should be determined, located at the threshold level th_{aux} , slightly lower than $\max(\text{abs}(c(m, n)))$. Gabor (noise + signal) coefficients should be theoretically

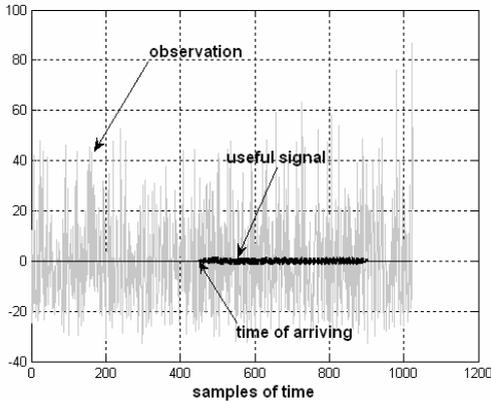


Fig. 7. Relation between time observation and the ideal RAD signal in the background for SNR ≈ -20 dB.

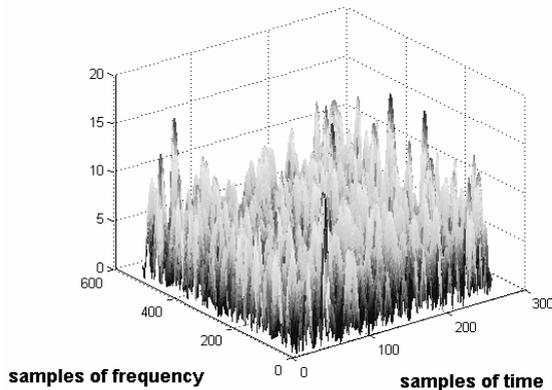


Fig. 8. Gabor transform of observation with the embedded RAD signal for SNR ≈ -20 dB.

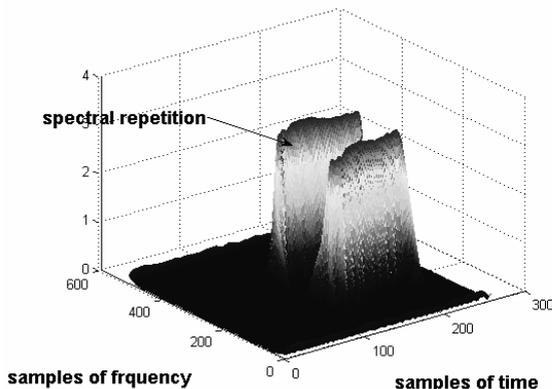


Fig. 9. Gabor transform of the real, non-noisy RAD signal (only signal coefficients) with spectral repeat effects, $R = 128$.

slightly bigger compared with only Gabor noise coefficients. In this secondary plane, represented by the secondary matrix $SCM_{M \times N}$, time-frequency coefficients crossing the secondary plane are left and the rest of coefficients are set to zero. The thresholding causes that many (but not all) of Gabor noise coefficients are rejected. In the secondary plane, the area occupied by Gabor (noise + signal) coefficients creates a heavily distorted shape called a noisy mask, represented by the matrix $NM_{m \times n, i}$.

According to the earlier assumptions that a signal can appear at a random time instant in an observation window, the localisation of a noisy mask is also random. This localisation should be found automatically in the algorithm. Thus the local Frobenius inner product is computed a few times for a slightly shifted pattern matrix (mask) $PM_{m \times n, i}$, $i = 1$ over a secondary matrix $SC_{M \times N}$ in the neighbourhood of the biggest values (only translations along the time axis and the frequency axis are allowed from the physical point of view). The biggest local Frobenius inner product MFP_1 indicates the most probable localisation of Gabor (noise + signal) coefficients (best matching localisation). The same operations have to be repeated for the two remaining ideal masks for computing MFP_2 and MFP_3 . The collection of MFP_i , $i = 1, 2, 3$ is also the collection of discriminant functions $g_i(r) = MFP_i$. The received signal is assigned to the i -th class, $i = 1, 2, 3$, with $\max g_i(r)$. This means that the noisy mask $NM_{m \times n}$ for the signal really embedded in noisy observation has just been found.

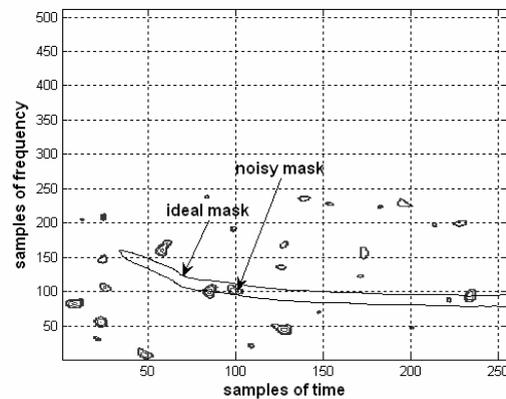


Fig. 10. Noisy mask of an analytic form of the BAT signal with the best matching with the ideal mask corresponding to SNR ≈ -16 dB.

In Figs. 10–12 it is clearly visible how an ideal cohesive pattern shape has been changed into a noisy mask because of the disturbing noise. It is obvious that the higher SNR, the less distorted the noisy mask and the easier the matching process.

In the experiments, two kinds of noisy masks were considered: $NM_{m \times n}$ with absolute values of Gabor coef-

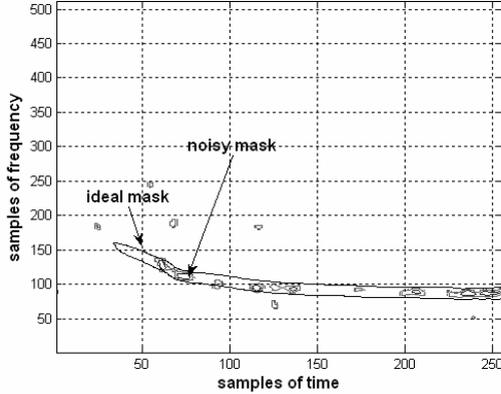


Fig. 11. Noisy mask of the analytic form of the BAT signal with the best matching with the ideal mask for SNR ≈ -5.5 dB.

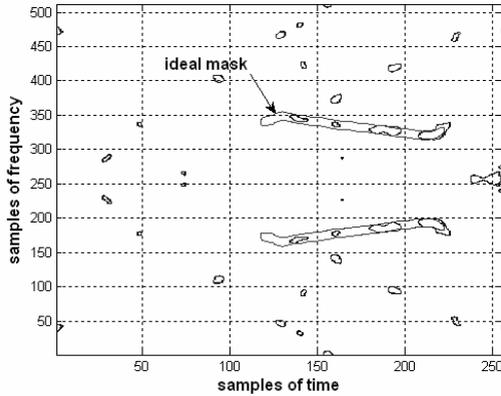


Fig. 12. Noisy mask of the real RAD signal with the best matching with the ideal mask for SNR ≈ -8 dB.

ficients and $NMB_{m \times n}$ with ones instead of absolute values of Gabor coefficients. This assumption results in two variants of classification rules:

$$\alpha = \max_{i=1,2,3} (g_i(r)) = \max_{i=1,2,3} MFP_i \quad (16)$$

provided that $NMB_{m \times n, i}$ is used,

$$\beta = \max_{i=1,2,3} (g_i(r)) = \max_{i=1,2,3} MFP_i \quad (17)$$

provided that $NM_{m \times n, i}$ is used.

Generally, in a multi-class classification problem the same classifications rules could be taken for $i = 1, \dots, K$.

4. Numerical evaluation of the proposed classification method

The number of proper classifications versus the SNR ratio is assumed as the criterion of effectiveness of the proposed

classification algorithm. The dependence of the number of proper classifications for RAD, BAT, HFM signals in 200 trials for each value of SNR versus increasing values of SNR ratios shown in Figs. 13–15 presents experimental evaluation of the proposed classification algorithm.

The classification of the HFM signal presented in Fig. 15 requires some explanations. It is worth noticing that the mask of the HFM signal for chosen parameters of computation might be completely contained in the mask of the BAT signal. In such a situation, a classifier could make a wrong decision taking the BAT signal instead of the HFM signal. Classification was performed assuming that HFM is always located outside the BAT mask, like in Fig. 16 in numerical experiments. The threshold th_{aux} is selected as $th_{aux} = \xi \max(\text{abs}(c(m, n)))$. Classification was performed with two variants of classification rules α and β from Eqns. (16)–(17) and a varying value of ξ . The results of classification for $\xi = 0.5, 0.7, 0.85$ respectively using the example of the RAD signal are shown in Figs. 14 and 17–18. The value of ξ has to be carefully established. If ξ is too small, more undesired noisy Gabor coefficients will appear in the secondary plane. If ξ is too high, less desired signal coefficients will appear in the secondary plane.

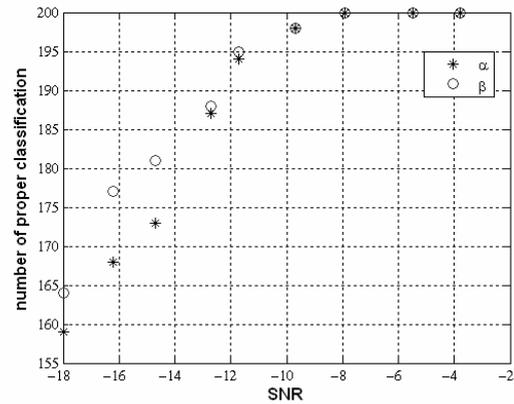


Fig. 13. Classification according to the rules α and β in 200 trials when the BAT signal is present in observations, $th_{aux} = 0.7 \max(\text{abs}(c(m, n)))$.

The comparison of Figs. 13–15 shows that the proposed algorithm is more effective in the range -18 dB to -12 dB for longer signals (the BAT signal is the longest signal whereas the HMF is the shortest one). For shorter signals, the possibility of incorrect matching of non-noisy and noisy masks substantially increases.

Experiments reveal that this strategy allows us to obtain almost as good as 100% classification accuracy for SNR higher than -8 dB, independently of the value th_{aux} for the three examined values of ξ . Comparing the classification results, $\xi = 0.7$ turns out to be the reasonable choice.

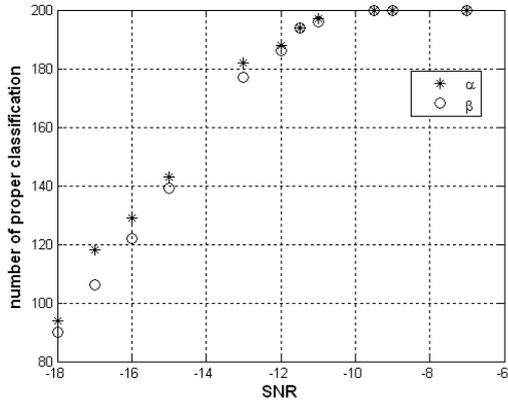


Fig. 14. Classification according to the rule α and β in 200 trials when the RAD signal is present in observations, $th_{aux} = 0.7 \max(\text{abs}(c(m, n)))$.

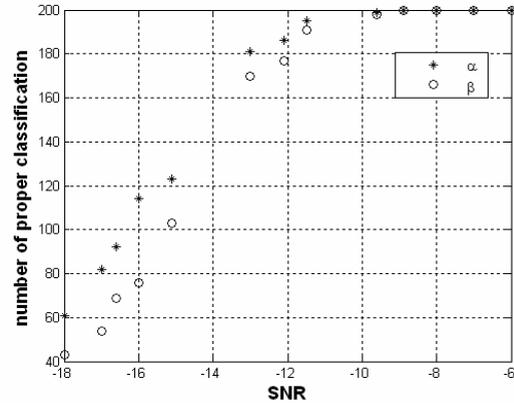


Fig. 17. Classification according to the rule α and β in 200 trials when the RAD signal is present in observations, $th_{aux} = 0.5 \max(\text{abs}(c(m, n)))$.

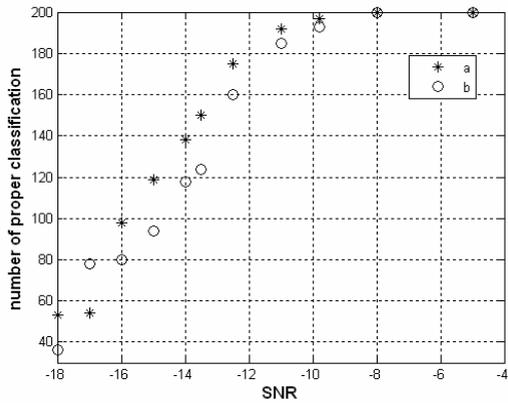


Fig. 15. Classification according to the rule α and β in 200 trials when the HFM signal is present in observations, $th_{aux} = 0.7 \max(\text{abs}(c(m, n)))$.

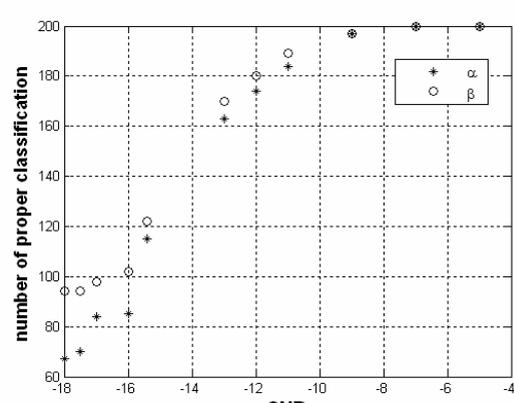


Fig. 18. Classification according to the rule α and β in 200 trials when the RAD signal is present in observations, $th_{aux} = 0.85 \max(\text{abs}(c(m, n)))$.

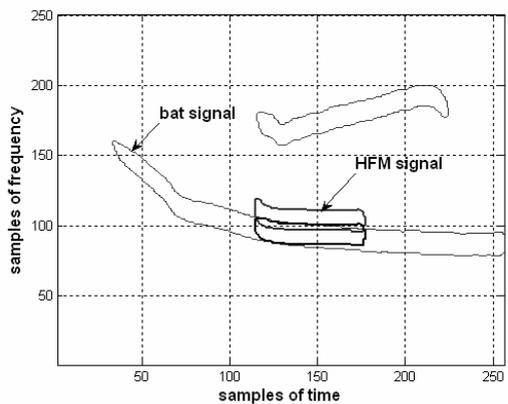


Fig. 16. Forbidden and allowed positions of the HFM signal in relation to the BAT signal.

The proposed algorithm was compared with the classification algorithm based on discrete wavelet decomposition and the LVQ neural network with the same experimental conditions. The structure of this algorithm was built using typical approaches to signal processing, which can be seen in Fig. 19, originated from (Sejdic et al., 2009).

Looking at the scheme in Fig. 19, the algorithm presented in the paper can be roughly matched to the following path: *Time-Frequency Domain* \Rightarrow *Amplitude Levels in TF bands* \Rightarrow *Distance Measures* represented in the paper by the Frobenius inner product in the *TF plane* \Rightarrow *Decision*. The classification algorithms based on the wavelet decomposition for creating a feature vector and artificial intelligence as a classifier were chosen as representative methods for comparison. This algorithm can be also seen in the following path in Fig. 19: *Time-Frequency Domain*

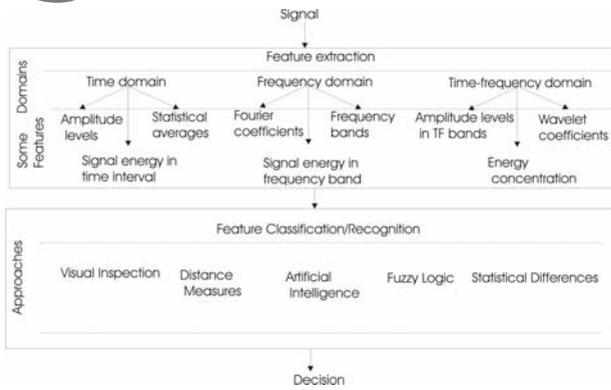


Fig. 19. Signal processing for pattern classification in signal applications, from (Sejdic *et al.*, 2009).

⇒ *Wavelet Coefficients* ⇒ *Artificial Intelligence* ⇒ *Decision*. It is worth pointing out that classification in the scheme *wavelet decomposition jointly with the LVQ network classifier* has some restrictions—the algorithm requires an established point of occurrence of all signals in time, e.g., at the beginning of the observation window. The algorithm presented in the paper does not impose such constraints; the signal can start at a random time point, which corresponds to a random time point in an observation window and a random localisation in the time-frequency plane.

The wavelet decomposition algorithm was used with the same three signals as described in previous experiments, i.e., the real life bat signal (BAT), the received radar signal (RAD) and the synthetic signal with hyperbolic modulation of frequency (HFM). It was found by experience that Daubechies wavelets are very useful in the proposed classification experiments. Using the Daubechies No. 4 mother wavelet up to the fourth level of decomposition, the feature vector was created from variances of detail coefficients for four decomposition levels.

A learning vector quantisation (LVQ) neural network was chosen as an intelligent classifier. The LVQ network has a first competitive layer and a second linear layer. The LVQ structure with the four-element input (number of features) and the three-element output (number of classes) was trained with 11 training signals in each class. Among the examined LVQ architectures, the one with three neurons (one neuron per class) in the competitive layer and one neuron per class in the linear output layer turned out to be the simplest and sufficient configuration for the classification. The training set was created by adding to the original signals small low-pass noise, which can be roughly treated as measurement noise.

The detection ability of the LVQ network was evaluated using the set of signals RAD, BAT, HFM with approximately the same range of SNR ratios as in the experiments in the proposed algorithm. The results of this

classification were quantified in terms of trials classified correctly.

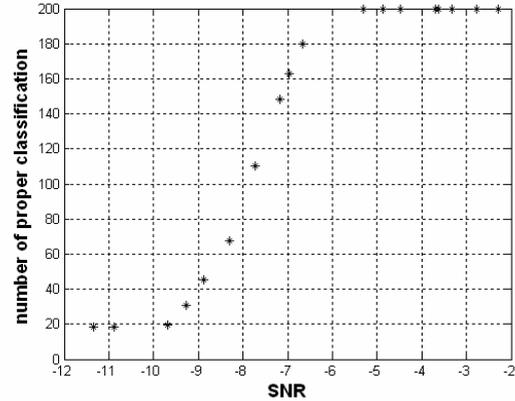


Fig. 20. Classification in 200 trials for each value of the SNR ratio when the BAT signal is present using wavelet decomposition and the LVQ algorithm as the intelligent classifier.

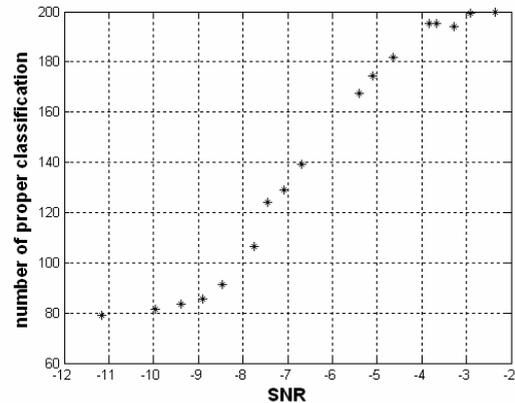


Fig. 21. Mean classification accuracy—jointly for three signals in 200 trials for each value of the SNR ratio using wavelet decomposition and the LVQ algorithm as the intelligent classifier.

Figures 20–22 display the results obtained with two methods: the one proposed in the paper (Fig. 13–18) and the one based on wavelet decomposition (Fig. 20–21). Comparing the results of the classification using the example of the BAT signal in two methods, cf., Fig. 13 and 20, it is clearly seen that the proposed method is more efficient especially for a low SNR ratio. For SNR ≈ -10 dB, the LVQ network utterly loses classification abilities. The confusion matrix for SNR < -9 dB has permanently the same form as presented in Table 1.

The overall classification rate is about 9.3% for SNR < -9 dB, which is the result of losing the classification ability.

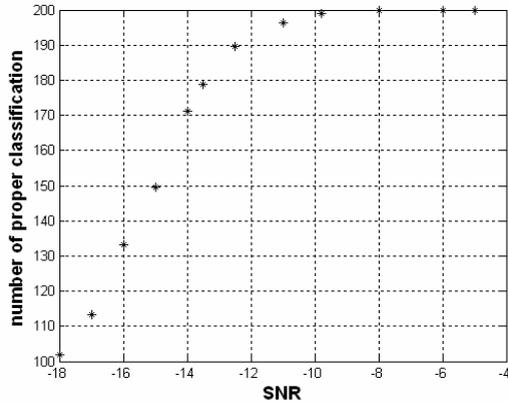


Fig. 22. Mean classification accuracy—jointly for three signals in 200 trials for each value of the SNR ratio using the proposed algorithm.

Table 1. Confusion matrix for the test set; SNR < -9 dB.

Class	Assignment of RAD to a class [%]	Assignment of BAT to a class [%]	Assignment of HFM to a class [%]
RAD	9.09	0.00	0.00
BAT	90.91	100.00	90.91
HFM	0.00	0.00	9.09

There is no point in using the proposed classification method with some intuitive approach in small noise environments. There are numerous efficient classification algorithms, which are based on precise mathematical relations, but they can fail in heavy noise conditions.

5. Discussion and conclusions

In this paper, a study of a new classification algorithm based on the matching shape idea of non-stationary signals available from observations was presented. The developed algorithm is suitable for non-stationary signals, well characterised by a frequency modulation law (FM). For this reason, not every non-stationary signal could be classified by the proposed algorithm. For example, speech cannot be classified by it.

The proposed algorithm requires the following steps:

- Create an ideal pattern (ideal mask) for an individual class by creating a characteristic image for the class on the Gabor time-frequency plane.
- Create an image of features by the Gabor transformation of received observations.
- Define a set of discriminant functions (computed as Frobenius inner products).

- Find a maximal discriminant function and make a decision about signal recognition embedded in observations.

The key idea is based on the comparison of shapes (ideal and noisy) obtained from the Gabor transform, which enhances frequency properties of signals belonging to the class. In the examined problem, there is no relation between each sample point on one shape and the corresponding sample point on the noisy shape, because of deformation caused by heavy noise. This kind of unpredictable deformation and the loss of a lot of points in noisy shapes cannot be explained in an easy way in the Gabor plain. Then, rather intuitive comparison of the pattern shape to the noisy shape is adopted to derive the overall classification algorithm.

The algorithm was demonstrated for three class classification of signals embedded in heavy noise and characterised by approximately polynomial and hyperbolic frequency modulation laws. Despite a low number of cases in the data set, the algorithm turned out to be effective for a low SNR ratio, as expected. The developed procedure gives reasonable results for the three non-stationary signals considered, and performance analysis showed that classification accuracy in numerical experiments is about 100% for SNR above -8 dB. In the same experiment conditions, classification based on wavelet decomposition and LVQ classification was performed. This LVQ classifier lost the classification ability for the SNR value, for which the proposed algorithm performs the classification task with a good classification rate.

It seems that the proposed method is general and can be successfully used in the recognition/classification of different frequency modulated, non-stationary signals.

References

Auger F., Flandrin P., Goncalves P. and Lemoine O. (1996). *Time-Frequency Toolbox for Matlab*, CNRS, Rice University, Houston, TX, <http://iut-saint-nazaire.univ-nantes.fr/~auger/tftb.html>.

Basri R., Costa L., Geiger D. and Jacobs D. (1998). Determining the similarity of deformable shapes, *Vision Research* **38**(15-16): 2365-2385.

Basseville M. (1989). Distance measures for signal processing and pattern recognition, *Signal Processing* **35**(3): 349-369.

Belongie S., Malik J. and Puzicha J. (2002). Shape matching and object recognition using shape contexts, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(4): 509-522.

Bishop C. M. (2006). *Pattern Recognition and Machine Learning (Information Science and Statistics)*, Springer Science + Business Media LLC, New York, NY.

Breakenridge C. and Mesbah M. (2003). Minimum classification error using time-frequency analysis, *Proceedings of*

- the 3rd IEEE International Symposium on Signal Processing and Information Technology (ISSPIT 2003), Darmstadt, Germany, pp. 717–720.
- Colas M. and Gelle G. (2004). A multitime-frequency approach for detection and classification of neighboring instantaneous frequency laws in a noisy environment, *Signal Processing Letters* **11**(2): 71–74.
- Davy M. and Doncarli C. (1998). Optimal kernels of time-frequency representations for signal classification, *Proceedings of the International Symposium Time-Frequency and Time-Scale*, Pittsburgh, PA, USA, pp. 581–584.
- Demirci M. F., van Leuken R. H. and Veltkamp R. C. (2007). Shape indexing through laplacian spectra, *Proceedings of the International Conference on Image Analysis and Processing Workshops (ICIAPW 2007)*, Modena, Italy, pp. 21–26.
- Doncarli C., Davy M. and Boudreaux-Bartels F. (2001). Improved optimization of time-frequency-based signal classifiers, *IEEE Signal Processing Letters* **8**(2): 52–57.
- Duda R. O., Hart P. E. and Stork D. G. (2001). *Pattern Classification, 2nd Edition*, John Wiley & Sons, Inc., New York, NY.
- Flandrin P. (1988). A time-frequency formulation of optimal detection, *IEEE Transactions on Acoustics, Speech and Signal Processing* **36**(9): 1337–1384.
- Fry D. (1993). *Shape Recognition Using Metrics on the Space of Shapes*, Ph.D. thesis, Harvard University, Cambridge, MA.
- Fukunaga K. (1990). *Introduction to Statistical Pattern Recognition, 2nd Edition*, Academic Press, London.
- Gdalyahu Y. and Weinshall D. (1999). Flexible syntactic matching of curves and its application to automatic hierarchical classification of silhouettes, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(12): 1312–1328.
- Gillespie B. and Atlas L. (2001). Optimizing time-frequency kernels for classification, *IEEE Transactions on Signal Processing* **49**(3): 485–496.
- Grigorescu S. E., Petkov N. and Kruizinga P. (2002). Comparison of texture features based on Gabor filters, *IEEE Transactions on Image Processing* **11**(10): 1160–1167.
- Gröchenig K. (2001). *Foundations of Time-Frequency Analysis*, Birkhäuser, Boston, MA, pp. 83–142.
- Hagedoorn M. and Veltkamp R. C. (1999). Reliable and efficient pattern matching using an affine invariant metric, *Journal of Computer Vision* **31**(2/3): 203–225.
- Heitz C. (1995). Optimum time-frequency representations for the classification and detection of signals, *Applied Signal Processing* **2**(3): 124–143.
- Huang Y., Chan K. L. and Zhang Z. (2003). Texture classification by multi-model feature integration using Bayesian networks, *Pattern Recognition Letters* **24**(1–3): 393–401.
- Jain A. K., Duin R. P. W. and Mao J. (2000). Statistical pattern recognition: A review, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(1): 4–7.
- Kyrki V., Kamarainen J.-K. and Klviinen H. (2004). Simple Gabor feature space for invariant object recognition, *Pattern Recognition Letters* **25**(3): 311–318.
- Latecki L. J. and Lakamper R. (2000). Shape similarity measure based on correspondence of visual parts, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22**(10): 1185–1190.
- Li S. and Shawe-Taylor S. (2005). Comparison and fusion of multiresolution features for texture classification, *Pattern Recognition Letters* **26**(5): 633–638.
- Liu H. and Srinath M. (1990). Partial shape classification using contour matching in distance transforms, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **12**(2): 1072–1079.
- Manay S., Cremers D., Hong B.-W., Yezzi A. J. Jr. and Soatto S. (2006). Integral invariants for shape matching, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **28**(10): 1602–1618.
- McLachlan G. J. (1992). *Discriminant Analysis and Statistical Pattern Recognition*, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., New York, NY.
- Petrakis E. G. M., Diplaros A. and Milios E. (2002). Matching and retrieval of distorted and occluded shapes using dynamic programming, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **24**(11): 1501–1516.
- Qian S. and Chen D. (1993). Discrete Gabor Transform, *IEEE Transactions on Signal Processing* **41**(7): 2429–2438.
- Richard C. and Lenggell R. (1999). Data driven design and complexity control of time frequency detectors, *Signal Processing* **77**(1): 37–48.
- Santini S. and Jain R. (1999). Similarity measures, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **21**(9): 871–883.
- Sebe N. and Lew M. S. (2002). Maximum likelihood shape matching, *Proceedings of the 5th Asian Conference on Computer Vision (ACCV2002)*, Melbourne, Australia, Vol. 1, pp. 713–718.
- Sejdic E., Djurovic I. and Jiang J. (2009). Time-frequency feature representation using energy concentration: An overview of recent advances, *Digital Signal Processing* **19**(1): 153–183.
- Sondergaard P. (2006). *Time-Frequency Toolbox for Matlab*, Technical University of Denmark, Lyngby, <http://www2.mat.dtu.dk/people/P.Sondergaard/toolbox/>.
- Tai C.-F. (2007). Image mining by spectral features: A case study of scenery image classification, *Expert Systems with Applications* **32**(1): 135–142.
- Umeyama S. (1993). Parameterized point pattern matching and its application to recognition of object families, *IEEE Transactions on Pattern Analysis and Machine Intelligence* **15**(2): 136–144.
- Veltkamp R. C. (2001). Shape matching: Similarity measures and algorithms, *Technical Report UU-CS-2001-03*, Utrecht University, Utrecht.

- Vincent I., Doncarli C. and Carpentier E. L. (1994). Non-stationary signals classification using time-frequency distributions, *Proceedings of the International Symposium on Time-Frequency and Time Scale, Paris, France*, pp. 233–236.
- Werther T., Eldar Y. C. and Subanna N. K. (2005). Dual Gabor frames: Theory and computational aspects, *IEEE Transactions on Signal Processing* **53**(11): 4147–4158.
- Xie J., Hengb P.-A. and Shah M. (2008). Shape matching and modeling using skeletal context, *Pattern Recognition* **41**(5): 1773–1784.
- Younes L. (1999). Optimal matching between shapes via elastic deformations, *Image and Vision Computing* **17**(5–7): 381–389.
- Zhang D. and Lu G. (2003). A comparative study of curvature scale space and Fourier descriptors for shape-based image retrieval, *Journal of Visual Communication and Image Representation* **14**(1): 41–60.
- Zhang D. and Lu G. (2004). Review of shape representation and description techniques, *Pattern Recognition* **37**(1): 1–19.



Ewa Świercz received her M.Sc. degree in electronics and the Ph.D. degree in telecommunications from the Faculty of Electronics (nowadays the Faculty of Electronics and Information Technology) at the Warsaw University of Technology, Poland, in 1977 and 1985, respectively. She currently works as a senior lecturer at the Department of Telecommunications and Electronic Equipment, Electrical Faculty, Białystok University of Technology. Her research interests include application of non-stationary signal processing methods and algorithms (e.g., time-frequency and time-scale decomposition) for detection of structure and parameter changes in dynamic systems and discovery of weak signals embedded in noise. She is an author and co-author of more than 60 research papers in journals and conference proceedings.

Received: 29 July 2008

Revised: 6 July 2009