Original Research Paper

Direct Sequence and Frequency Hopping Signals Classification Based on Co-Occurrence Matrix and Clustering Techniques

[1]Haidy S. Fouad, [2]Hend A. Elsayed and [1]Shawkat K. Girgis

[1]Department of Information Technology,
Institute of Graduate Studies and Research (IGSR), Alexandria University, Egypt
[2]Department of Communication and Computer Engineering,
Delta University for Science and Technology, Mansoura, Egypt

Abstract: Spread Spectrum techniques (SS) were first developed for military applications, but currently, they have commercial applications. SS provides secure communication and allows multiple accesses for same radio spectrum. So, most Wireless Local Area Network (WLAN) systems use it, as do Cognitive Radios (CR), space systems and Global Positioning Systems (GPS). Direct Sequence Spread Spectrum (DSSS) and frequency hopped spread spectrum (FHSS) are the two most-used techniques today. Nowadays radio spectrum has become very crowded and so now there is a need for spectrum efficiency. Automatic SS classification presents a rather difficult problem, especially if the parameters, such as the signal power, carrier frequency, etc., are unknown. This research takes a new direction; it deploys the Gray Level Co-occurrence Matrix (GLCM) to capture statistical features of SS signals. Using GLCM, 22 features are extracted for each vector of signal. Analyzing the signals is done in the time domain which measures the variation of amplitude of signals with time. Therefore, the main contribution is to apply and show how GLCM improves the identification accuracy of the two signals in presence of noise. The proposed model achieves considerably accurate results even with a low SNR. GLCM features help classifiers to achieve average accuracy 84% and reach 100% signal identification at a zero SNR. To prove the superiority of these features, a variety of clustering methods are applied, such as centroid, connectivity, model-based and message-passing models. Clustering performance results based on GLCM features are compared With Principal Components Analysis (PCA), Kernel-based Principal Components Analysis (KPCA) and fast Independent Components Analysis (Fast-ICA). Clustering results are evaluated with external and internal validity indices. The accuracy was tested over 26 levels of Signal-to-Noise Ratios (SNR).

Keywords: Clustering, Cluster Validity, Direct Sequence Spread Spectrum (DSSS), Frequency Hopped Spread Spectrum (FHSS), Gray Level Co-Occurrence Matrix (GLCM)

Introduction

Spread Spectrum signals (SS) have an extra modulation called Code Division Multiple Access (CDMA), which offers one of the main methods for managing multiple accesses in broadcasts. This modulation has advantages, such as resisting interference and making communications secure and difficult to detect since the communications resemble noise (Gibson, 2013). Consequently, SS is a favorable choice for many mobile communication networks and intelligent receivers, such as Software-Defined Radio (SDR) and Cognitive Radio (CR), which are devices capable of learning and adapting to their environment (Liljana *et al.*, 2013; Pandit and Singh, 2017). Fette (2009) discussed the concept of CR as input in different

fields, such as smart antennas and hardware architectures. The main task of both CR and SDR is identification of the signal received. The most practical and leading SS systems are direct sequence (DSSS) and Frequency-Hopping Systems (FHSS) (Torrieri, 2018). In both, noise-like sequences are deployed through data spreading and modulation. Pseudo-Random Binary Sequences (PRBSs) or Pseudo-Noise (PN) is the most common methods (Hasan *et al.*, 2016). DSSS is used as an example in IEEE 802.11b and IEEE 802.15.4. FHSS is used in Bluetooth communications.

Gray Level Co-occurrence Matrix

Texture features extraction can be divided into the following four categories of computation: Statistical, geometrical, model-based and signal-processing methods. GLCM features consider the most widely used statistical methods (Chaki *et al.*, 2015) and are used in many arenas. It provides a robust method for face recognition (Eleyan and Irel, 2011). It has been applied to 3D printing (Fastowicz and Okarma, 2016), Digital Panoramic Radiograph Images (Enny and Sutarman, 2017), mammogram analysis (Pratiwi *et al.*, 2015) and audio processing (Lacerda and Mello, 2017). This research takes a new direction, deploying GLCM to capture statistical features of SS signals.

GLCM is a way of extracting second-order statistical texture features. It transforms the data into a Gray Level co-occurrence matrix that acts as a basis for features calculation. GLCM counts the occurrence frequency of two pixels: One is the reference and the other is the neighbor pixel. The former has intensity "*I*" and the latter has intensity "*j*", they are separated by a pixel distance of (Δx, Δy), so that the matrix element $P(i, j|\Delta x, \Delta y)$ is the relative frequency. Then, the matrix is reconstructed to be symmetrical around the diagonal. Lastly, normalization is done by dividing every element over the sum of values, to turn it into probabilities and form the final GLCM elements.

Based on this matrix, Haralick earlier presented fourteen features (Haralick *et al.*, 1973). Only four features were measured with MATLAB, namely, contrast, correlation, energy and homogeneity. Later, five scalar texture features were introduced by Conners *et al.* (1984). Then, many features have been added, such as trace, which is introduced in (Sebastian *et al.*, 2012) and used in (Korchiyne *et al.*, 2014). A selective group of these features have been deployed in many fields. Energy, correlation, sum entropy and sum variance are used in mammograms, based on a t-test (Pratiwi *et al.*, 2015). Angular second moment, entropy and sum entropy are selected based on Receiver Operating Characteristic (ROC) analysis (Mashrur *et al.*, 2017). Likewise, energy, contrast, correlation and entropy are used in satellite images (Zhang *et al.*, 2017).

Pattern recognition includes two basic learning techniques, namely, supervised and unsupervised. The difference between the two lies in the training process. The choice of learning technique is directed by the availability of information. This research uses clustering, which is an unsupervised method, as it has no need for complex training methods or ground-truth labels during the learning process. It is based on investigating characteristics and the intrinsic structure of the data. Many of the top unsupervised methods use feature-based models (Berg-Kirkpatrick *et al.*, 2010).

The problem considered in this research is the identification of spread spectrum communication signals with unknown parameter values in the presence of the noise. The proposed solution is organized in five sections. Section 1 presents the introduction. The Clustering Strategy is presented in Section 2, followed by the related work in Section 3. The Methodology is presented in Section 4. Lastly, Section 5 presents the study's conclusions.

Clustering Strategy

This research applied a variety of clustering techniques to prove the effectiveness of GLCM□s features.

a) Partitioning-Based Methods

K-means, k-medoids, PCM and FPCM all yield unsustainable results when the initial cluster centers have virtual and random characteristics. Therefore, some researchers tried data points with minimum entropy values, while others used k-means++ or evolution algorithms to find the initial center. In this research, two methods are applied to obtain the initial centers – choosing data points randomly or by using subtractive clustering:

- Fuzzy c-means (FCM): This is a soft fuzzy method which allows a data point to belong to a cluster in varying degrees, specified by a membership grade. There are several improved versions, such as Possibilistic C-Means clustering (PCM), which deals with the problem of noise in the FCM algorithm. However, the solution still has many quality-related drawbacks, such as coincident clusters. Fuzzy Possibilistic C-Means (FPCM) was introduced to solve problems in both FCM and PCM algorithms (Shouraki, 2003), (Zhang and Chen, 2014). The adaptive possibilistic approach based on average similarity is also used in to allow discovering non-convex shaped clusters (Mahfouz *et al.*, 2018)
- K-means: The k-means algorithm clusters data by minimizing the mean distance. Each cluster is relocated by its center point
- K-medoids: It picks data points as centers that are most centrally located and minimizes a sum of pairwise dissimilarities

b) Subtractive Clustering

It is an enhancement of mountain clustering in that it reduces the computation process (Dhanachandra *et al*., 2015). In this study, subtractive clustering is used as an isolated clustering technique and as a method to generate the initial center for use in all centroid-based clustering.

c) Hierarchical Clustering

It only requires a measure of similarity between groups of data points as a key and then builds a binary tree of them. Linkage methods encode the tree hierarchy in a dendrogram, which represents the cluster groups. Its strength is that the clusters are not assumed to be globular. The number of clusters must be known, because this is the cut-off point of the desired hierarchy level. Basically, this method can be a divisive clustering method, or an agglomerative clustering method, which is considered as one of the most common methods of hierarchical clustering. The agglomerative method is the used method in this research. Hierarchical methods integrate with other techniques, such as CURE and CHAMELEON. Recently, a scalable BIRCH-like adaptive possibilistic clustering algorithm was introduced as ExBIRCH (Mahfouz, 2018).

d) Density-Based Spatial Clustering

Unlike the partitioning methods, Density-Based Spatial Clustering of Applications with Noise (DBSCAN) does not have the risk of randomness in cluster centroids initialization and it does not assume globular clusters. In addition, it does not require every point to be assigned to a cluster and has scalable performance. It controls clustering by the radius of the cluster and the minimum number of data objects required within the cluster. The values for those parameters are viewed as having an optimization problem, because they are highly affected by the final result. Differential Evolution was used in (Karami and Johansson 2014).

e) Model-Based Clustering

A Gaussian Mixture Model (GMM) is used in this method. It makes soft assignments. As such, it provides more flexibility. The cluster is often referred to as a component distribution. Each component has the parameters of a positive integer dimensional mean, a covariance matrix and a mixing proportion. The algorithm estimates them by applying maximum likelihood. It tries to infer these distributions from the data and by optimizing the fit between the data and the model (Tian *et al*., 2017).

f) Affinity Propagation (AP)

AP is a non-parametric clustering algorithm proposed by (Frey and Dueck, 2007; Dueck, 2009). AP uses graph distances between points. Clustering is achieved by passing two types of real-valued messages between them responsibility and availability. Unlike k-means or k-medoids, AP does not require specifying the initial cluster centers in advance. Instead, it regards all data points equally as potential exemplars (cluster centers) and groups the clusters by the degree of similarity among data point. Through its objective function it aims to maximize the net similarity that is indicated by affinities. As such, the choice of distance has a big influence on clustering performance. Its weakness is in big data handling (Sun, 2017) due to its message-exchange technique.

After applying these various clustering techniques, we evaluated clustering accuracy with external and internal indices (Rezaei, 2016), to demonstrate the quality of the proposed features.

Strategy of the Comparative Study

Clustering results based on GLCM features are compared with the results that are obtained by the most common feature extraction techniques. These techniques are; Principal Components Analysis (PCA), Kernel-based Principal Components Analysis (KPCA) and fast Independent Components Analysis (Fast-ICA) (Yang and Cheng, 2015; Alam and Fukumizu, 2014). PCA is implemented in MATLAB, KPCA is based on (Wang, 2014) and Fast-ICA on (AU, 2005).

Related Work

Spectrum identification by its nature is a very challenging problem (Watson, 2013). Clustering or blind signal separation is potential pattern recognition methods for this purpose (Jajooa *et al*., 2017). The selected solution is directed by the available parameters, as shown in Table 1. Carrier or center frequency, signal bandwidth and cyclostationary signals are some features are used in (Weng *et al*., 2014). A cyclostationary signal is one with statistical properties that change with a certain periodic pattern, such as symbol rates, coding schemes and training symbol sequences. Others identify signals based on Automatic Modulation Classification (AMC) approaches (Zhu and Nandi, 2015). In DSSS-detection algorithms, especially at low Signal-to-Noise Ratios (SNR), PCA and sequence synchronization as a semi-blind method is proposed in (Vlok and Olivier, 2012; Vlok, 2014). Additionally, correlation is used in (Javed and Khalid, 2018), Eigen-decomposition and Chirp Z Transform (CZT) in case the period and chip rate of PN sequence are present (Shen and Wang, 2015).

Many methods become useless when there is not much prior information, in which case blind identification becomes especially significant. In these

cases, another direction based on image processing is raised, such as extracting the parameters of Frequency-Hopping (FH) signals based on the spectrogram obtained by Short-Time Fourier Transform (STFT), in time–frequency domain (Fu *et al.*, 2017). After the feature extraction stage, researchers deploy various methods for SS signals clustering, such as K-means, x-means algorithms and DPMM-based, Bayesian non-parametric classification. However, in all of them, 100% signal clustering accuracy was reached only in SNR = 6 Decibel (dB) (Jayaweera, 2014).

The basic steps of the proposed framework are summarized below.

Methodology

To demonstrate the superiority of the proposed solution, clustering performance is measured with a 26 SNR level. Signals are simulated using MATLAB R2017a. Figure 1 illustrates the basic workflow of the pattern recognition process using the proposed GLCM feature extraction techniques.
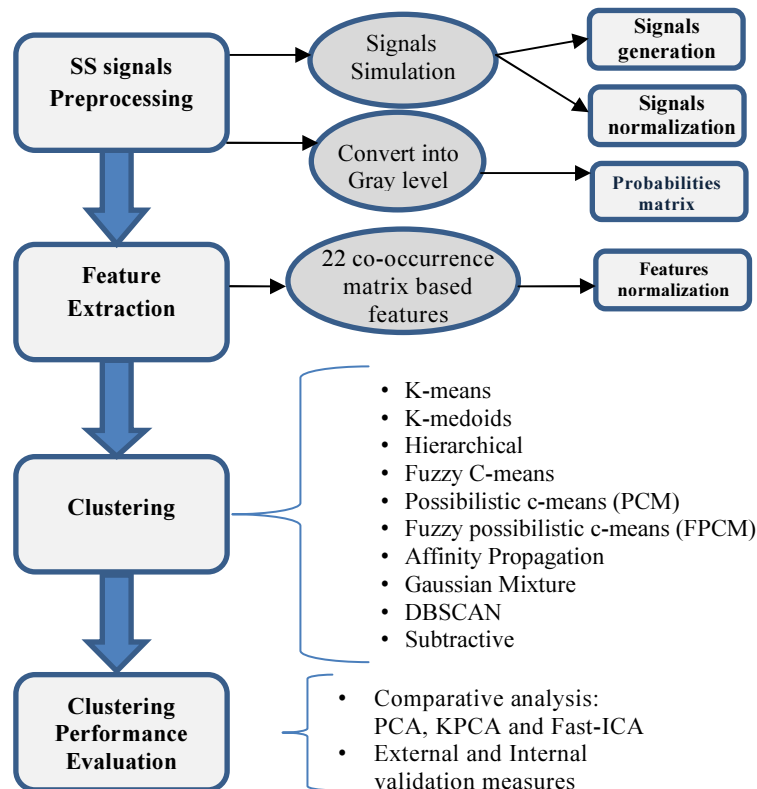
Fig. 1: Diagram of the proposed pattern recognition framework

Table 1: Spread spectrum signal radio identification in cognitive

Sensing method	Dimension	Sensing parameter	Result
Energy Detection	Time/frequency	Signal energy	Detection only
Matched Filter	Time	Time domain signal structure and characteristics, such as pulse shape and guard time	Detection and Identification
Cyclostationary Feature	Frequency/code	Data and chip rate, coding schemes, modulation type and carrier number	Detection and Identiaction
Statistical Tests	Time	Signal distribution	Detection only
Entropy Based	Frequency	Signal entropy	Detection only
Autocorrelation	Time/angle	Signal eigenvalues and direction of arrival	Detection only
Template Matching	Time	Cyclic prefix and PN sequence	Detection and Identification
Template Matching	Frequency	Frequency domain	Detection and Identification
Multilayer Based	Frequency	Signal energy	Detection only
Wavelet	Frequency	Signal energy	Detection only
Multiband Joint Detection	Frequency	Signal energy	Detection only

**Step 1: SS** Signals preprocessing

Signals simulation

Direct Sequence Spread Spectrum (DSSS)

First, the original signal bit sequence d(t) is generated from ten bits values of -1 and 1, each having a duration t which equals 100 chip pattern, multiplied by a pseudorandom sequence, which spreads the bandwidth. The spreading is applied by using PN-Sequence $c(t)$ which defined as a coded sequence of 1s and 0s. Then the output signal $s(t)$ is modulated by Binary Phase-Shift Keying (BPSK) using sinusoidal carrier wave w_c. The modulation is done by using two sinusoidal carrier waves. The simplest signals modulation is BPSK (Swamy *et al.*, 2013). The equation is in (1):

$$s(t) = d(t)c(t)\cos(w_c t) \tag{1}$$

The output is the modulated signal $x(t)$, which transmits over the AWGN channel, with multiple SNR from -15 to 10. Finally, the transmitted signal $y(t)$ is obtained, as shown in Fig. 2. The receiver does not have these parameters and the signal appears as noise during the transmission, as shown in Fig. 3.
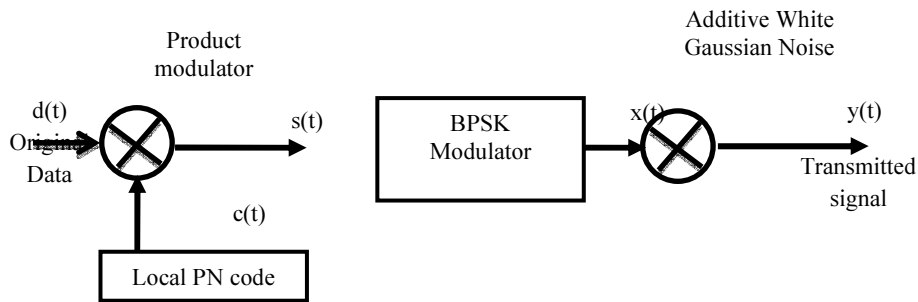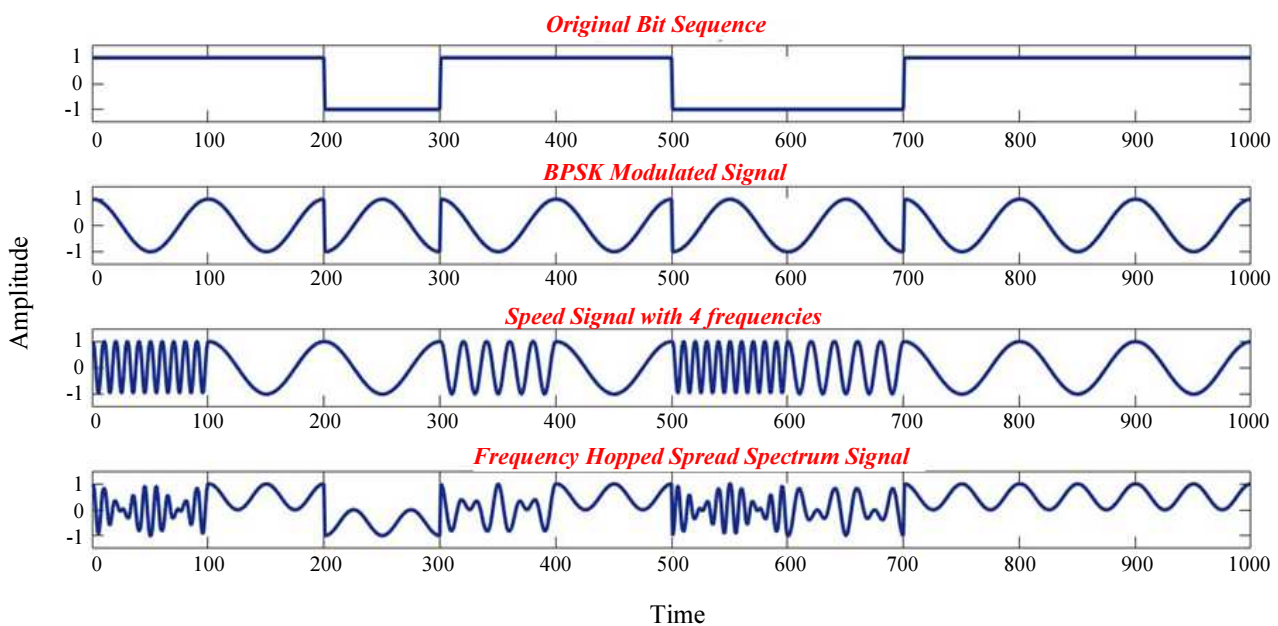

Fig. 2: Diagram of the transmitted DSSS
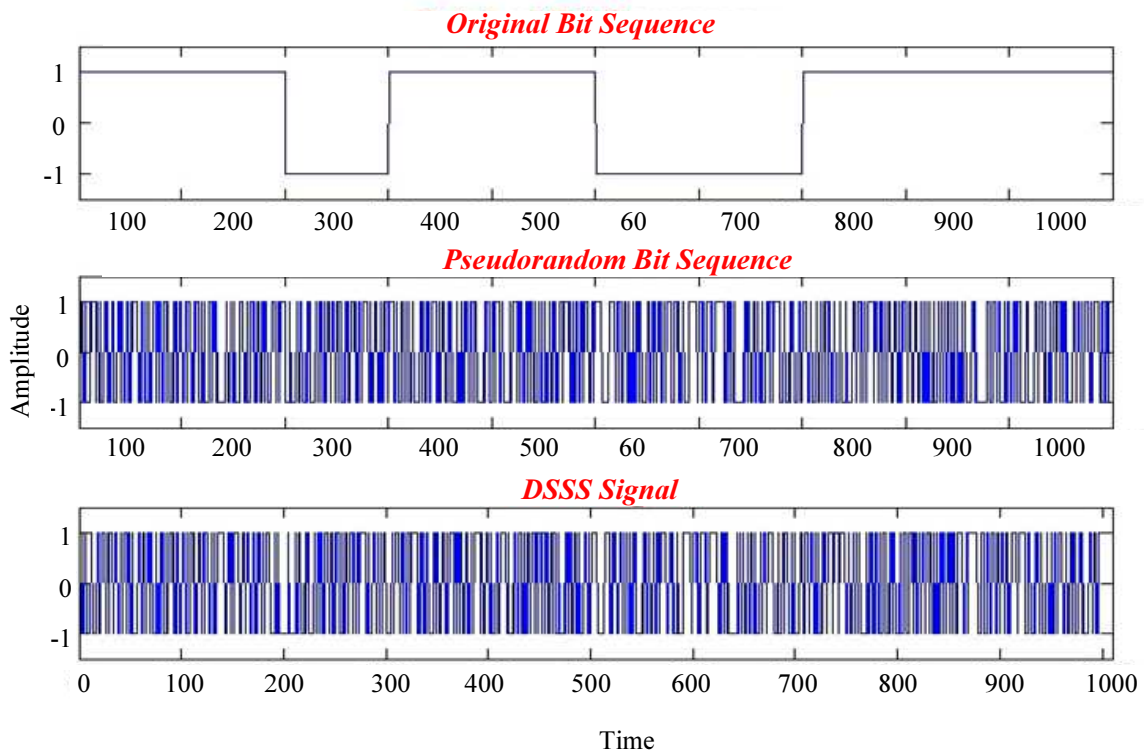
Fig. 3: Simulation of the DSSS transmission technique
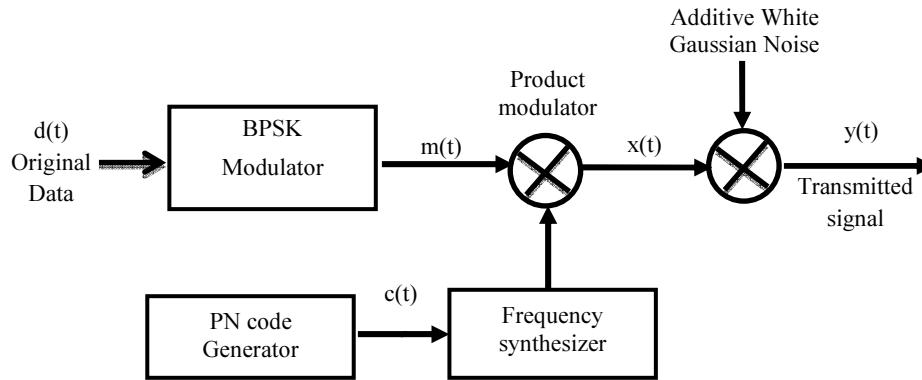
Fig. 4: Diagram of the transmitted FHSS

Fig. 5: Simulation of the FHSS transmission technique

The signal is simulated by a binary data stream and obtained from a random number generator, from ten bits takes values of -1 or 1, then the signal is directly modulated onto a four sinusoidal carrier by BPSK to get the modulated signal $m(t)$. The frequency of the carrier is rapidly switched between many frequency channels in a pseudorandom manner and at a specified time interval so that the pattern is not fixed; the list of different frequencies is maintained in a table called „hopping". As in the DSSS analysis, this method is performed with additive white Gaussian noise (AWGN), with SNR added from -15 to 10 dB. Figure 4 presents a diagram of the transmitted FHSS. The total number of samples taken was 200 signals at each noise level. The number of DSSS signals was 100, the same as the FHSS signals. For purposes of comparison, a default random number generator in MATLAB is used. Figure 5 shows the FHSS spreading with 4 frequencies.

Signals Normalization

In both SS signals, basic signals are generated on time domain with time-series lengths N = 1000 samples. The modulated signals are normalized with a scaling factor that is based on average power, using "modnorm" function in Matlab. Finally, the signal constellation is performed by multiplying the modulated signal by the scaling factor. Normalizing the signals gives them the comparative ability and does not make the signal power change with the modulation scheme (Stukavec, 2010). For each signal, 100 signals vectors are generated under every specified noise level between -15 and 10 dB.

Convert the Signals into Gray Level

To capture the signal pattern, each simulated signal is transformed into an 8-bit gray level to build the Gray Level Co-occurrence Matrix (GLCM). Creating the GLCM from the signal vector is done by using the "graycomatrix" function in MATLAB. The size of the GLCM is equal to the number of gray-levels. The increasing of the levels provides more accurate textural information but with a computational cost.

Step 2: Features Extraction

Extract Signal Features using GLCM

The experiment was conducted on a computer with the following specifications: AMD A8-7600; 3.10 GHz processor; 12 GB of RAM.

Features are extracted from the simulated FHSS and DSSS signals for 26 different SNR ratios ranging from -15 to 10 S/R. In this research the features vector is formed from the 22 features. The features are calculated using the well-known GLCM functions. Figure 6 shows the complete process of features' calculation and illustrates each equation's reference. Some features have the same names but have different definitions and equations. The equations of the 22 features are listed in Table 2.
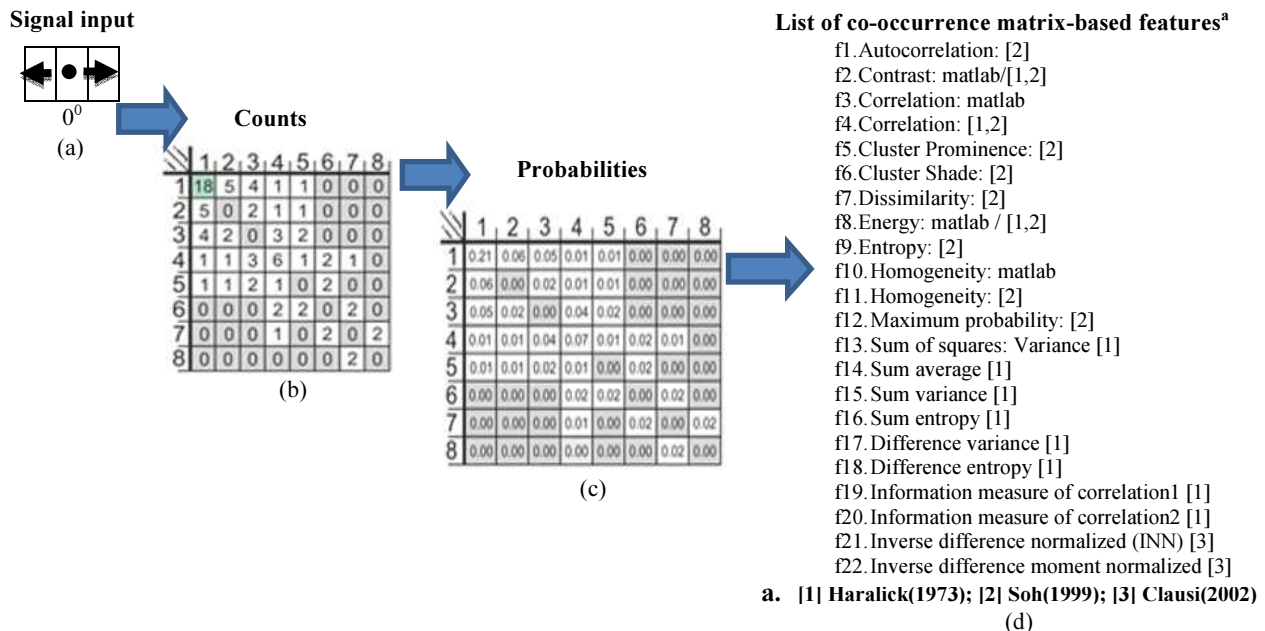
Where:

- i and j are the gray level values in the input data window
- P_{ij} is the properties of the element i, j in the normalized symmetrical GLCM
- N is the number of gray levels
- μ is the GLCM mean
- C is the Correlation feature
- $\sum_i is \sum_{i=1}^{N} and \sum_j is \sum_{j=1}^{N}$
- μ_i, μ_j and σ_i, σ_j are the mean and standard deviations of probability matrix GLCM along row wise x and column wise y.
- $P_{X+y}(i)$ is the probability of matrix coordinates summing to $x+y$
- $p_{x+y}(k) = \sum_i \sum_j p(i,j)$ where $i+j = k = \{0,1, 2, ..., 2(N-1)\}$
- $p_{x-y}(k) = \sum_i \sum_j p(i,j) = k = |i-j| = \{0,1, 2,....,(N-1)\}$

Features Normalization

The signals dataset has 200 samples by 22 features for each level of SNR. Feature scaling is an important pre-processing step before clustering is performed. Standardization is considered to avoid erroneous influencing and to simplify the range in the distance-based classification. It imparts the same importance to all variables and leads to more efficient and accurate cluster results. Z-score scales the features to have a standard normal distribution with $\mu = 0$ and $\sigma = 1$, representing the mean and standard deviation from the mean, respectively. All features are scaled before clustering. The z-score is calculated as in (2):

$$Z = (X - \mu) / \sigma \qquad (2)$$

List of co-occurrence matrix-based features[a]
f1. Autocorrelation: [2]
f2. Contrast: matlab/[1,2]
f3. Correlation: matlab
f4. Correlation: [1,2]
f5. Cluster Prominence: [2]
f6. Cluster Shade: [2]
f7. Dissimilarity: [2]
f8. Energy: matlab / [1,2]
f9. Entropy: [2]
f10. Homogeneity: matlab
f11. Homogeneity: [2]
f12. Maximum probability: [2]
f13. Sum of squares: Variance [1]
f14. Sum average [1]
f15. Sum variance [1]
f16. Sum entropy [1]
f17. Difference variance [1]
f18. Difference entropy [1]
f19. Information measure of correlation1 [1]
f20. Information measure of correlation2 [1]
f21. Inverse difference normalized (INN) [3]
f22. Inverse difference moment normalized [3]

a. [1] Haralick(1973); [2] Soh(1999); [3] Clausi(2002)
(d)

Fig. 6: The Flow of the process of GLCM features calculation (a) Input vector of normalized signals (b) Initial GLCM using eight grey levels in (d = 1, θ = 0°) (c) Normalized GLCM (d) List of co-occurrence matrix-based features

Table 2: GLCM features equations

Formula	Formula
f1 Autocorrelation = $\sum_i \sum_j (i,j) p(i,j)$	f12 Maximum probability = max(p)
f2 Contrast = $\sum_i \sum_j (i-j)^2 p_{i,j}$	f13 Sum of Squares variance = $\sum_i \sum_j (i-\mu)^2 p_{ij}$
f3 Correlation = $\sum_{ij} \frac{(i-\mu_i)(j-u_j) p_{ij}}{\sigma_i \sigma_j}$	f14 Sum Average = $\sum_{i=2}^{2N} i P_{x+y}(i)$
f4 Correlation = $\frac{\sum_i \sum_j (ij) p(i,j) - \mu_x \mu_y}{\sigma_i \sigma_j}$	f15 Sum Variance = $\sum_{i=2}^{2N} (i-\mu_{x+y})$
f5 Cluster Prominence = $\sum_{i,j} \{i+j-\mu_x \mu_y\}^4 * p(i,j)$	f16 Sum Entropy = $-\sum_{k=2}^{2N} (p_{x+y}(i) \log\{p_{x+y}(i)\})$
f6 Cluster Shade = $\sum_{i,j} \{i+j-\mu_x-\mu_y\}^3 * p(i.j)$	f17 Difference variance = $\frac{1}{2} \sum_{I,J} (iP(i,j)+jP(i,j))$
f7 Dissimilarity = $\sum_{i,j=0}^{N-1} P_{i,j} |i-j|$	f18 Difference entropy = $-\sum_{i=0}^{N-1} p_{x-y} \log\{p_{x-y}(i)\}$
f8 Energy = $\sum_i \sum_j p_{ij}^2$	f19 Information measures of correlation 1 = $\frac{HXY - HXY1}{MAX\{HX,HY\}}$ $HXY = -\sum_i \sum_j p(i,j) \log(p(i,j))$ $HXY1 = -\sum_i \sum_j P(i,j) \log\{P_x(i) P_y(j)\}$
f9 Entropy = $-\sum_i \sum_j p_{ij} \log_2 p_{ij}$	f20 Information measures of correlation 2 = $(1 - \exp[-2(HXY2 - HXY)])^{\frac{1}{2}}$ $HXY2 = -\sum_i \sum_j P_x(i) p_y(j) \log\{p_x(i) P_y(i)\}$
f10 Homogeneity = $\sum_{i,j} \frac{p_{i,j}}{1+|i-j|}$	f21 Inverse difference normalized (INN) = $\sum_{i=1}^N \sum_{j=1}^N \frac{p_{i,j}}{1+|i-j|}$
f11 Homogeneity = $\sum_i \sum_j \frac{1}{1+(i-j)^2} p_{ij}$	f22 Inverse difference moment normalized (IDN) = $\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$

Step 3: Clustering implementation

There are special notes related with some techniques that were used

Hierarchical Clustering

This technique has its own measure to assess the solution quality, which is the cophenetic correlation coefficient function. It was good; it was 0.68 in -15 SNR and in SNR 10 reached 0.996. This function reflects how the tree represents the dissimilarity between observations. A high-quality solution has a value close to one.

Gaussian Mixture Model (GMM)

The k-means++ algorithm is used in the initial parameter values for a fitted Gaussian mixture model, which are the mean, covariance and mixing proportion.

DBSCAN Clustering

This method has sensitivity to noise and outliers. It has parameters to control the final result, such as the neighbourhood radius and the minimum number of a neighbourhood of an object. Therefore, we used the function that was implemented in (Hutchins, 2013).

Assess feature extraction techniques with a bundle of classifiers. besise apply different techniques as PCA, KPCA and Fast-ICA for comparative analysis.

Step 4: Clustering performance evaluation

		Clustering result	
		Negative	Positive
Actual class	Negative	TN	FP
	Positive	FN	TP

Fig. 7: The confusion matrix

First, for clustering result assessment, the overall accuracy rate is used. It is an external measure that evaluates the clustering based on how well the cluster labels produced by the algorithm match the ground-truth labels. It can be obtained by the "classperf" function in MATLAB. It can also be obtained using Equation 3, which measures the success rate *Cs* by dividing the sum of samples correctly classified *TP* and *TN* by the total number of samples, which is divided into true positive or negative and false positive or negative, as shown in the confusion matrix in Fig. 7. The complement of this measure is called the misclassification rate:

$$Cs = (TP+TN)/(TP+FP+TN+FN) \qquad (3)$$

Parameter of Clustering Techniques

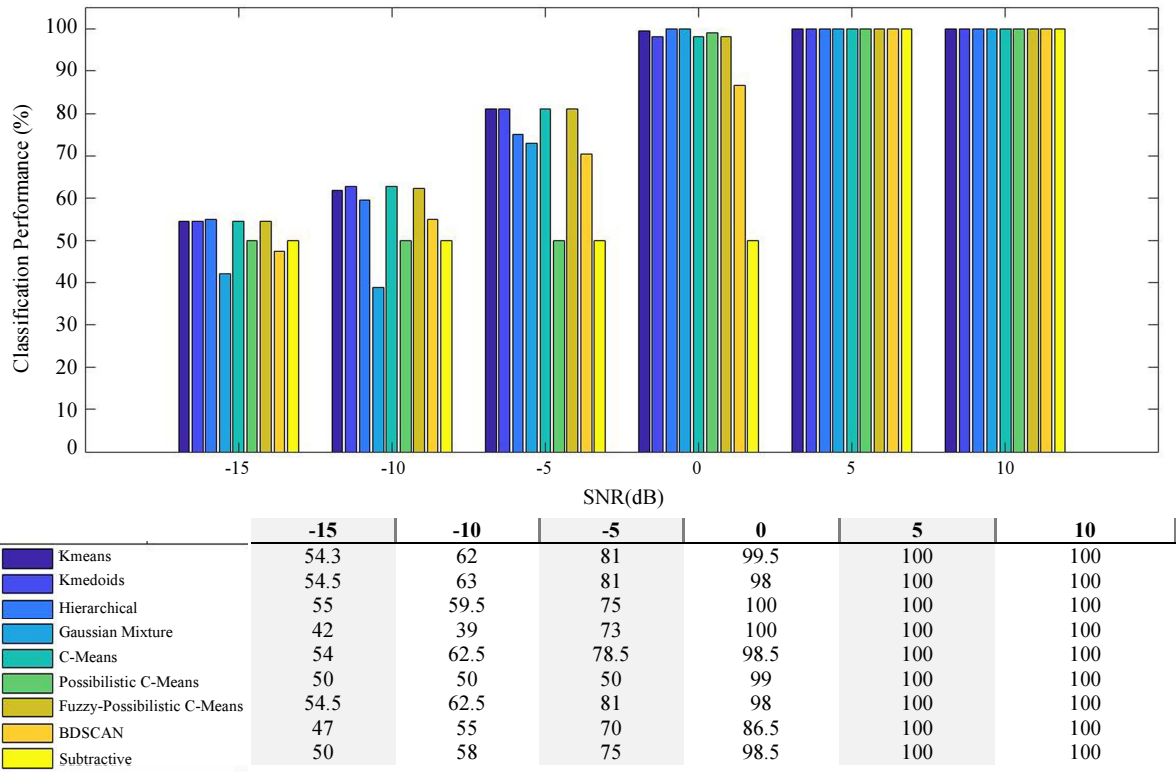Parameter of the clustering techniques can be summarized as shown in Table 3.

Comparative Analysis

Compared with the results of Jayaweera (2014) that used two features (carrier frequency and estimated bandwidth), the GLCM features enhance clustering performance. As shown in Table 4, at a -4 SNR, the accuracy of k-means with GLCM features reached 84.5%, compared to 57.3% for the two-feature k-means. Additionally, GLCM features achieved 100% at 2 SNR, while the other only achieved that level at 6 SNR. By

using GLCM features, the GMM and hierarchical methods achieve accurate results at 0 SNR.

For more analysis, Table 5 presents the averages of overall accuracy grouped in three categories; from -15 to -1, at 0 and above 0 SNR. K-means achieved the same result as k-means-sub above -4 SNR. The hierarchical and Gaussian mixture techniques reach 100% accuracy more quickly, with good accuracy at low levels of SNR (68.8% and 61.8%, respectively). Figure 8 gives wide view to the accuracy in transition points -15, 10,5,0,5 and 10.

	-15	**-10**	**-5**	**0**	**5**	**10**
Kmeans	54.3	62	81	99.5	100	100
Kmedoids	54.5	63	81	98	100	100
Hierarchical	55	59.5	75	100	100	100
Gaussian Mixture	42	39	73	100	100	100
C-Means	54	62.5	78.5	98.5	100	100
Possibilistic C-Means	50	50	50	99	100	100
Fuzzy-Possibilistic C-Means	54.5	62.5	81	98	100	100
BDSCAN	47	55	70	86.5	100	100
Subtractive	50	58	75	98.5	100	100

Fig. 8: GLCM-based clustering accuracy at different Signal-to-Noise Ratio (SNR) levels

Table 3: Parameters of the used clustering techniques

Method	Parameters	Metric properties
K-Means	clusters number = 2; distance = cosine	Distances between points
Affinity propagation	damping factor = 0.9, important to avoid oscillations	Graph distance
	data point similarities = Pearson correlation coefficients, 100 iterations	
	data point similarities = Pearson correlation coefficients, 100 iterations	
c-mean	Exponent for the fuzzy partition matrix = 3	Distances between
Agglomerative	threshold for cutting = 2, linkage type = ward	points
clustering	distance = Euclidean	Distances between
BSCAN	Neighborhood radius = 7.	nearest points
Gaussian	Component covariance structure= Diagonal; number k = 2; Regularization Value = 0.01	Mahalanobis
Mixtures (GMM)	The maximum number of iterations is 100	distances to centers
Subtractive	Range of influence = 0.9	Pairwise distance
	Squash factor of 8.0; Accept ratio 0.9; Reject ratio 0.2; Distance = City block	
PCM &FPCM	Number of clusters = 2, Weighting exponent= 2. The clustering process stops when	Distances between
	the maximum number of iterations equals 100 or the termination	points
	threshold equals 0.00001.	
	For FPCM: typicality weight = 4	

Table 4: Comparative performance analysis between GLCM features and other techniques under different SNR

SNR=	−4	−2	0	2	4	6	8
K-means[a]	57.3	76.3	84.2	96.9	97	100	100
GLCM-Based							
K-means	84.5	94.5	99.5	100	100	100	100
K-medoids	85.5	93.5	98	100	100	100	100
Hierarchical	78.5	96	**100**	100	100	100	100
FPCM	84.5	92	99	100	100	100	100
Affinity	84.5	90.5	98	100	100	100	100
GMM	79	96.5	**100**	100	100	100	100
Subtractive	74.5	91	98.5	100	100	100	100
DBSCAN	70.5	70.5	86.5	100	100	100	100
C-means	84	92.5	99	100	100	100	100
PCM	50	50	90	100	100	100	100

[a](Jayaweera, 2014)

Table 5: Average accuracy percentages using GLCM features under different SNR

Techniques	SNR = -15: -1	SNR = 0	Above 0 SNR
K-means	71.5	99.5	**100**
K-means-sub	68.88	99.5	**100**
K-medoids	71.4	98	99.95
FPCMC- Sub	71.5	99.5	**100**
Hierarchical	68.8	**100**	**100**
FPCMC-subtractive	71.3	98	99.95
FPCMC	62.5	99	99.95
Affinity	**69.2**	99.5	**100**
GMM	61.8	**100**	**100**
DBSCAN	60.5	86.5	98
C-means	60.07	98.5	99.95
PCM	53.1	99	100
Subtractive	61.16	98.5	100

Comparing the clustering results based on 22 features extracted by PCA, KPCA and Fast-ICA, the best average performance results were all under 68%. PCA gets its best performance by using Euclidean distance and only achieved an average accuracy of 38.78%. Subtractive clustering based on PCA and KPCA have unstable accuracy starting with 36.5% in the lowest SNR level. The Fast-ICA results were ignored as it had poor results with different classifiers; its best average accuracy was 65.02% using hierarchical clustering.

To clarify the results, Table 6 presents average accuracy calculated over 26 SNR levels. Classifiers achieved higher identification ratios based on GLCM features than they did based on other features. K-means and FPCMC achieved the highest average performance. K-means had the best average time of 0.004 seconds. Affinity consumes time 2. 13 so it neglected as time is consider an important factor. Subtractive as initial center technique does not add a value, so it will not be mentioned. Figure 8 presents the clustering accuracy of basic nine techniques based on GLCM in -15, 10,5,0,5, 10 SNR.

However, overall accuracy was not enough to demonstrate the quality of clustering. Therefore, other external measures were used, such as Normalized Mutual Information (NMI), precision, recall, F-measure, the Adjusted Rand Index (ARI), Jaccard and the Fowlkes Mallows Index (FMI). The highest values achieved are shown in bold in Table 6. All range between 0 to 1 and the best are equal to 1. Higher values indicate greater similarity between the clusters and the benchmark classifications.

Table 6 presents the results for the average performance of clustering methods using seven external indices. The results represent average scores for 26 noise levels. PCA and KPCA validity measures implemented with k-medoids only achieved the best overall accuracy. All GLCM feature-based clustering techniques have excellent results with these indices. K-means achieved many high scores, including the maximum scores for NMI (information theoretic), F-measure (set-matching) and ARI (pair-counting measures).

Internal Clustering Evaluation

For deeper clustering evaluation, internal cluster validity measures have been applied, as they rely only on the information that is in the data itself and clustering structures. The selected indices are Davies-Bouldin (DB), Dunn, Calinski-Harabasz (CH), Han component of Hartigan (Han) and a component of the Krzanowski-Lai (KL) index. Table 7 presents the indices' values.

DB evaluates both compactness and separation in summary form by averaging all the cluster similarities. uses the average between and within the cluster sum of squares. It measures separation using the maximum distance between cluster centres and measures compactness using the sum of distances between objects and their cluster centre. Dunn is the ratio between the minimum inter-cluster distance to the maximum intra-cluster distance and measures inter-cluster separation by minimum pairwise distance between points in different clusters and intra-cluster compactness by maximum diameter among all clusters. Han is based on the Sum-of-Squares Within a cluster (SSW) and the Sum-of-Squares Between clusters (SSB), respectively. KL shows the sum-of-squares within a cluster. K-medoids, Hierarchical and FPCMC achieved the best results. There are many techniques that achieved similar results and good clustering.

Table 6: Average accuracy percentages using GLCM features under different SNR

Features extraction method	Techniques	Average accuracy	Average time (seconds)
GLCM features	K-means	**84**	**0.004**
	FPCM	83.59	0.012
	Affinity	**82.2**	2.13
	Hierarchical	82	0.016
	K-means- sub	81	**0.005**
	GMM	78	0.064
	FPCMC- sub	83.05	0.008
	K-means	77.1	**0.004**
	Subtractive	77.5	0.011
	FCM	76.9	0.006
	K-medoids	76.7	0.011
	DBSCAN	76.38	0.048
	PCM	72.9	0.017
PCA	K-medoids	**67.64**	0.009
	Affinity	65.02	3.349
	FPCM	61.9	0.004
	PCM	58.30	0.008
	K-means	55.44	0.155
	FCM	54.05	0.003
	Hierarchical	50	**0.002**
	GMM	50	0.040
	DBSCAN	38.79	0.011
KPCA	K-medoids	**67.73**	0.009
	Affinity	65.02	4.942
	FPCM	55.44	0.009
	FCM	54.82	**0.002**
	K-means	50.08	0.061
	Hierarchical	50	0.003
	Gaussian Mixture	50	0.043
	PCM	45.87	0.009
	DBSCAN	38.79	0.015

Table 7: Internal clustering evaluation

Clustering technique	DB	CH	Dunn	KL	Han
K-means	0.85	1767.39	4.26	1946.24	1827.38
K-means-sub	0.85	1767.33	4.26	1947.05	1828.14
K-medoids	0.84	1766.14	**4.27**	1960.00	1840.30
Hierarchical	**0.84**	1766.14	**4.27**	1960.00	1840.30
GMM	1.38	1734.53	3.96	2368.38	2223.74
FPCMC	0.85	**1767.77**	**4.27**	**1943.07**	**1824.40**
FPCMC-Sub	0.85	1765.87	4.26	1963.37	1843.47
Affinity	0.85	1765.87	4.26	1963.37	1843.47
DBSCAN	NaN	1635.22	NaN	3142.24	3126.04
PCA_ K-medoids	20.41	6.53	0.26	4545.60	4268.00
KPCA_K-medoids	46.28	0.88	0.11	4665.58	4380.66
Fast-ICA_K-medoids	28.19	0.35	0.08	4677.91	4392.24
Optimal value	Min	Max	Max	Min	Min

Table 8: External clustering evaluation

Clustering technique	NMI	F-measure	Recall	Precision	ARI	Jaccard	FMI
K-means	**0.55**	**0.84**	**0.84**	**0.84**	**0.57**	0.79	**0.70**
K-means-sub	0.54	**0.84**	**0.84**	**0.84**	0.57	0.78	**0.70**
K-medoids	0.54	0.83	0.84	**0.84**	0.57	0.79	**0.70**
Hierarchical	0.54	0.83	0.84	**0.84**	0.35	**0.80**	0.67
DBSCAN	0.31	0.54	0.65	0.50	0.35	0.80	0.65
Affinity	0.54	0.84	**0.84**	0.83	0.35	0.79	0.69
GMM	0.54	0.81	0.81	0.81	0.35	0.78	**0.70**
FPCMC	0.54	0.83	0.83	**0.84**	0.35	0.78	0.69
FPCMC-Sub	0.54	**0.84**	**0.84**	**0.84**	0.57	0.78	0.69
Pca_K-medoids	0.09	0.68	0.68	0.68	0.05	0.56	0.39
KPCA_K-medoids	0.09	0.68	0.68	0.68	0.05	0.56	0.39
Fast-ICA_K-medoids	0.00	0.50	0.50	0.50	0.00	0.49	0.33

Conclusion

Without prior knowledge of many parameters, the GLCM features, as a statistical measure of texture, can improve identification of both FHSS and DSSS signals, even with a low SNR. These features achieve an accurate SS signal clustering result with a variety of classifiers, such as k-means, FCM, PCM, FPCM, agglomerative hierarchical clustering, affinity propagation, Gaussian mixture, subtractive clustering, k-medoids and DBSCAN. Both external and internal indices demonstrate the high quality of the clustering performance. GLCM features help classifiers to achieve 100% signal identification at a zero SNR. A comparative analysis proves the superiority of the GLCM features results compared to other features extracted by PCA, KPCA and Fast-ICA.

The improved Direct Sequence Spread Spectrum (DSSS) and Frequency Hopped Spread Spectrum (FHSS) identification enhances the bandwidth and data flow to devices that are likely to implement them, such as mobile communication networks and intelligent receivers.

Author's Contributions

Haidy S. Fouad: Prepared the extracted features and applied different clustering techniques with suitable tuning and refining. Writing the manuscript.

Hend A. Elsayed: Designed the research plane.

Shawkat K. Girgis: Supervision and critical review

Ethics

This article is original and contains unpublished material. There are no ethical issues involved.

References

Alam, M. and K. Fukumizu, 2014. Hyperparameter selection in kernel principal component analysis. J. Comput. Sci., 10: 1139-1150. DOI: 10.3844/jcssp.2014.1139.1150

AU, 2005. Department of Information and Computer Science, Aalto University.

Berg-Kirkpatrick, T., A. Bouchard-Côté, J. DeNero and D. Klein, 2010. Painless unsupervised learning with features. Proceedings of the Annual Conference of the North American Chapter of the Association for Computational Linguistics, Jun. 02-04, Los Angeles, California, pp: 582-590.

Chaki, J., R. Parekh and S. Bhattacharya, 2015. Plant leaf recognition using texture and shape features with neural classifiers. Patt. Recognit. Lett., 58: 61-68. DOI: 10.1016/j.patrec.2015.02.010

Clausi, A., 2002. An analysis of co-occurrence texture statistics as a function of grey level quantization. Canadian J. Remote Sens., 28: 45-62. DOI: 10.5589/m02-004

Conners, R., M. Trivedi and C. Harlow, 1984. Segmentation of a high-resolution urban scene using texture operators. Comput. Vis. Graph. Image Process., 25: 273-310. DOI: 10.1016/0734-189X(84)90197-X

Dhanachandra, N., K. Manglem and Y. Chanu, 2015. Image segmentation using k-means clustering algorithm and subtractive clustering algorithm. Proc. Comput. Sci., 54: 764-771. DOI: 10.1016/j.procs.2015.06.090. 54: 764-771

Dueck, D., 2009. Affinity propagation: Clustering data by passing messages. PhD Thesis, University of Toronto, Canada.

Eleyan, A. and H. Irel, 2011. Co-occurrence matrix and its statistical features as a new approach for face recognition GLCM Features. Turk J. Electr. Eng. Comput. Sci., 19: 97-105. DOI: 10.3906/elk-0906-27

Enny, S. and M. Sutarman, 2017. Extracting the potential features of digital panoramic radiograph images by combining radio morphometry index, texture analysis and morphological features. J. Comput. Sci., 14: 144-152. DOI: 10.3844/jcssp.2018.144.152

Fastowicz, J. and K. Okarma, 2016. Texture based quality assessment of 3d prints for different lighting conditions. Proceedings of the International Conference on Computer Vision and Graphics, Sept. 19-21, Springer, Warsaw, Poland, pp: 17-28. DOI: 10.1007/978-3-319-46418-3_2

Fette, B., 2009. Cognitive Radio Technology. 2nd Edn., Elsevier, USA, ISBN-13: 9780080923161, pp: 848.

Frey, B. and D. Dueck, 2007. Clustering by passing messages between data points. Science, 315: 972-976. DOI: 10.1126/science.1136800

Fu, W., X. Li, N. Liu, Y. Hei and J. Wei, 2017. Parameter blind estimation of frequency-hopping signal based on time–frequency diagram modification. Wireless Personal Communications, 97: 3979-3992. DOI: 10.1007/s11277-017-4710-5

Gibson, J., 2013. Basic Principles. In: Mobile Communications Handbook, CRC Press, Boca Raton, FL, USA, ISBN-13: 9781439817247.

Haralick, R., K. Shanmugam and I. Dinstein, 1973. Textural features for image classification. IEEE Trans. Syst. Man Cybernet., 3: 610-621. DOI: 10.1109/TSMC.1973.4309314

Hasan, M., J. Thakur and P. Podder, 2016. Design and implementation of FHSS and DSSS for secure data transmission. Int. J. Signal Process. Syst., 4: 144-149. DOI: 10.12720/ijsps.4.2.144-149

Hutchins, M., 2013. https://github.com/mlhutchins/functions/blob/master/dbscan.m

Jajooa, G., Y. Kumar, S. Kumar, B. Adhikari and A. Kumar, 2017. Blind signal modulation recognition through clustering analysis of constellation signature. Expert Syst. Applic., 90: 13-22.

Javed, H. and M. Khalid, 2018. Blind scramble code detection and period estimation in non-cooperative interception of a scrambled DSSS signal. Proceedings of the 15th International Bhurban Conference on Applied Sciences and Technology, Jan. 9-13, IEEE Xplore Press, Islamabad, Pakistan, pp: 846-849 DOI: 10.1109/IBCAST.2018.8312323

Jayaweera, S., 2014. Signal Classification in Wideband Cognitive Radios. In: Signal Processing for Cognitive Radios, Jayaweera, S.K. (Ed.), John Wiley and Sons, Inc., New Jersey, ISBN-13: 978-1-118-82493-1, pp: 429- 471.

Karami, A. and R. Johansson, 2014. Choosing DBSCAN parameters automatically using differential evolution. Int. J. Comput. Applic., 91: 1-11. DOI: 10.5120/15890-5059

Korchiyne, R., S. Farssi, A. Sbihi, R. Touahni and M. Alaoui, 2014. A combined method of fractal and GLCM features for MRI and CT scan images classification. Signal Image Process. Int. J., 5: 85-97. DOI: 10.5121/sipij.2014.5409

Lacerda, E. and C. Mello, 2017. Automatic classification of laryngeal mechanisms in singing based on the audio signal. Proc. Comput. Sci., 112: 2204-2212. DOI: 10.1016/j.procs.2017.08.115

Liljana, G., V. Atanasovski, I. Macaluso and L. DaSilva, 2013. Learning and reasoning in cognitive radio networks. IEEE Commun. Surveys Tutorials, 15: 1761-1777. DOI: 10.1109/SURV.2013.030713.00113

Mahfouz, A., 2018. ExBIRCH: Scalable non-centroid BIRCH-like algorithm for clustering gene expression data based on average correlation. Int. J. Pure Applied Biosc., 6: 37-46. DOI: 10.18782/2320-7051.6308

Mahfouz, M., A. Mohamed, Y. El-Sonbaty and M.A. Ismail, 2018. Symbolic approach to reduced bio-basis. Int. J. Data Min. Bioinformat., 20: 47-66. DOI: 10.1504/IJDMB.2018.092158

Mashrur, G., E. Elahi, S. Kalra and Y. Yang, 2017. Texture classification of MR images of the brain in ALS using CoHOG. Computer Vision and Pattern Recognition.

Pandit, G. and G. Singh, 2017. Spectrum Sensing in Cognitive Radio Networks: Potential Challenges and Future Perspective. In: Spectrum Sharing in Cognitive Radio Networks, Pandit, S. and G. Singh (Eds.,) Springer, Switzerland, ISBN-13: 978-3-319-53147-2, pp: 35-37.

Peterson D., P. Ghosh and R. Maitra, 2017. Merging K - means with hierarchical clustering for identifying general-shaped groups, The ISI's J. for the Rapid Dissemination of Statistics Research. DOI: 10.100X/sta.0000.

Pratiwi, M., J.H. Alexander and S. Nanda, 2015. Mammograms classification using gray-level co-occurrence matrix and radial basis function neural network. Proc. Comput. Sci., 59: 83-91. DOI: 10.1016/j.procs.2015.07.340

Rezaei, M., 2016. Clustering validation. PhD Thesis, University of Eastern Finland.

Sebastian, B., A. Unnikrishnan and K. Balakrishnan, 2012. Gray level co-occurrence matrices: Generalisation and some new features. Int. J. Comput. Sci. Eng. Inform. Technol., 2: 151-175. DOI: 10.5121/ijcseit.2012.2213

Shen, B. and J. Wang, 2015. Estimation of PN sequence in DSSS signals with unknown carrier frequency under narrow band interferences. J. Electr. Inform. Technol., 37: 1556-1561. DOI: 10.11999/JEIT141322

Shouraki, S., 2003. ce.sharif.edu/~m_amiri/project/yfcmc/index.htm

Soh, L. and C. Tsatsoulis, 1999. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Trans. Geosci. Remote Sens., 37: 780-795. DOI: 10.1109/36.752194

Stukavec, R. and T. Kratochvil, 2010. Simulation and measurement of the transmission distortions of the digital television DVB- part 2: Hierarchical modulation performance. Radio Eng., 19: 429-436.

Sun, L., C. Guo, C. Liu and H. Xiong, 2017. Fast affinity propagation clustering based on incomplete similarity matrix. Int. J. Knowledge Information Syst., 51: 941-963. DOI:10.1007/s10115-016-0996-y

Swamy, M., M. Deepthi, V. Mounika and R. Saranya, 2013. Performance analysis of DSSS and FHSS Techniques over AWGN Channel. Int. J. Electr. Commun. Instrument. Eng. Res. Dev., 4: 56-65.

Tian, K., S. Zhou and J. Guan, 2017. DeepCluster: A general clustering framework based on deep learning. Proceedings of the Joint European Conference on Machine Learning and Knowledge Discovery in Databases, Sept. 18-22, Skopje, Macedonia, pp: 809-825. DOI: 10.1007/978-3-319-71246-8_49

Torrieri, D., 2018. Principles of Spread-Spectrum Communication Systems. 4th Edn., Springer, ISBN-10: 9783319705682, pp: 727.

Vlok, D. and J. Olivier, 2012. Non-cooperative detection of weak spread-spectrum signals in additive white Gaussian noise. IET Commun., 6: 2513-2524. DOI: 10.1049/iet-com.2011.0614.

Vlok, J., 2014. Detection of direct sequence spread spectrum signals. PhD Thesis, University of Tasmania, Hobart.

Wang, Q., 2014. https://www.mathworks.com/matlabcentral/fileexchange/39715-kernel-pca-and-pre-image-reconstruction

Watson, C., 2013. Signal detection and digital modulation classification-based spectrum sensing for cognitive radio. PhD Thesis, Northeastern University Boston, Massachusetts, USA.

Weng, Z., P. Orlik and K. Kim, 2014. Classification of wireless interference on 2.4GHz spectrum. Proceedings of the IEEE Wireless Communications and Networking Conference. Apr. 6-9, IEEE Xplore Press, Istanbul, Turkey, pp: 786-791. DOI: 10.1109/WCNC.2014.6952168

Yang, J. and Q. Cheng, 2015. A comparative study of independent component analysis with principal component analysis in geological objects identification, Part I: Simulations. J. Geochem. Explorat., 149: 127-135. DOI: 10.1016/j.gexplo.2014.11.013

Zhang, Q. and Z. Chen, 2014. A distributed weighted possibilistic c-means algorithm for clustering incomplete big sensor data. Int. J. Distributed Sensor Netw.

Zhang, X., J. Cui, W. Wang and C. Lin, 2017. A study for texture feature extraction of high-resolution satellite images based on a direction measure and gray level co-occurrence matrix fusion algorithm. Sensors, 17: 1-15. DOI: 10.3390/s17071474

Zhu, Z. and K. Nandi, 2015. Blind Modulation Classification. In: Automatic Modulation Classification: Principles, Algorithms and Applications, Wiley, UK, ISBN-13: 978-1-118-90649-1, pp: 97-105.