Signal Processing 126 (2016) 12–17

Contents lists available at ScienceDirect

Signal Processing

journal homepage: www.elsevier.com/locate/sigpro

Anomaly detection based on probability density function with Kullback–Leibler divergence

Wei Wang*, Baoju Zhang, Dan Wang, Yu Jiang, Shan Qin, Lei Xue

College of Electronic and Communication Engineering, Tianjin Normal University, Tianjin, PR China

ARTICLE INFO

Article history:
Received 28 August 2015
Received in revised form
3 January 2016
Accepted 14 January 2016
Available online 22 January 2016

Keywords:
Anomaly detection
Kullback–Leibler divergence
Principal component analysis
Through wall human detection
UWB radar

ABSTRACT

Anomaly detection is a popular problem in many fields. We investigate an anomaly detection method based on probability density function (PDF) of different status. The constructed PDF only require few training data based on Kullback–Leibler Divergence method and small signal assumption. The measurement matrix was deduced according to principal component analysis (PCA). And the statistical detection indicator was set up under iid Gaussian Noise background. The performance of the proposed anomaly detection method was tested with through wall human detection experiments. The results showed that the proposed method could detection human being for brick wall and gypsum wall, but had unremarkable results for concrete wall.

1. Introduction

Anomaly detection which aims to look for the different pattern with the desired pattern in data, also known as novelty detection、anomaly mining、noising mining has been applied in many fields, such as credit card fraud detection, medical diagnostics information anomaly detection, industrial equipment fault detection and structural defect detection, network intrusion detection and novel theme of text mining. Currently, anomaly detection methods include Density-based techniques [1], Subspace and correlation-based outlier detection [2], one class support vector machines [3], Cluster analysis-based outlier detection [4],and Ensemble techniques [5] and so on.

Based on the number of samples with different patterns for data mining, there are supervised learning method, semi-supervised method and unsupervised method. But in anomaly detection, the abnormal patterns are usually few and are present for only a little time. So, it is very difficult to acquire the anomaly through supervised learning method, even with semi-supervised method.

As a typical application of anomaly detection, through wall detection of human being is of great interest for many applications which used the Ultra-wideband (UWB) radar with strong anti-interference ability, high resolution performance and good target recognition capabilities. In Military field, it could be used for detecting and locating the terrorists and hostages in the building to improve the success rate for the hostage rescue greatly. Through wall human detection also can be used for disaster search and rescue operations such as people trapped under building during earthquake, explosion or fire.

Many experts have taken extensive researches on the through wall target detection. Most researchers are focusing on the two dimension imaging or three dimensions imaging for through wall human detection [6,7]. And some researches are used the time domain signal to detect human directly. Such as: M.Y.W.Chia detected the heart beats and breathing rate distortions through 8 cm thick wall obstacles with a UWB radar prototype within FCC's mask. [8] A.G. Yarovoy measured the spatial variations of UWB radar signals to detect and locate a person within 0.6 cm due to breathing or a novel motion [9]. In [10], an

* Corresponding author.
E-mail address: weiwangvip@163.com (W. Wang).

http://dx.doi.org/10.1016/j.sigpro.2016.01.008
0165-1684/

algorithm for detection of trapped persons under rubble based on enhancing specific signal features with an adequate filtering through an M-sequence UWB radar. In [11,12], some UWB radar signal processing algorithm for through wall human detection such as raw radar data processing, background subtraction, target detection, target trace estimation and localization has been provided to detect a walking person along a given trajectory behind wall. And further the multiple target detection based on the UWB radar has been introduced in [13]. Wei Wang has used the UWB radar for through wall human detection and proposed an detection criterion with wavelet packet transform based on statistical process control theory [14]. In [15], Ultra-Wideband Pseudo-Noise radar were used for through wall human detection and tracking of moving person. There are two insufficient in these achievements. In one hand, the noise is unavoidable added in radar signal. And it will affect the detection accuracy. In the other hand, some methods needed a few of signals with different patterns which was unpractical in the reality.

Some experts have used the transform domain of UWB radar signal for through wall human detection. Such as: Qilian Liang provided Doppler based method, short time Fourier transform method and singular value decomposition method were used for through wall human detection with UWB radar for different types of walls [16]. Wei Wang collected the UWB radar signal with compressive sensing algorithm and acquired the singular values of the compressed radar data to detect human behind Gypsum wall [17]. In [18], human being detection behind the wall with UWB radar was proposed based on the spectrum of breathing human. But these method needed more anomaly patterns too, and it is also impossible in the reality.

As a typical case of anomaly detection, the above through wall human detection algorithms showed that the methods were usually based on some training data with different patterns and had poor noise resistance. However, in practice, this information may not be available. In this paper, we propose a novel method for anomaly detection which construct the Probability Density Function (PDF) of different pattern using the Kullback–Leibler (KL) divergence theory. This method is good at detection while less information is needed. The remainder of the paper is organized as follows. In Section 2 we will construct the PDF of different pattern and apply it to anomaly detection. In Section 3 we set up a through wall human detection system with UWB radar under three kind of walls to test the effective of our detection algorithm. The discussion about the experimental results is presented in Section 4. Conclusion and future work are discussed in Section 5.

2. Methods

2.1. PDF estimation of different status

Considering that we observe the output data as follows:

$$x = s + n \tag{1}$$

where $x \in R^N$ is a snapshot of the sensor network with N sensors at one time or a length N time series signal of a sensor. In this paper, we used one couple of UWB radars as emitter and receiver for through wall human detection. So we received a signal x with length N. Signal x consists of two parts: s characterizes the radar signal and n represents the noise. Previous studies on through wall human detection with UWB radar have found that n is iid Gaussian noise with variance σ^2 [14,16,17].

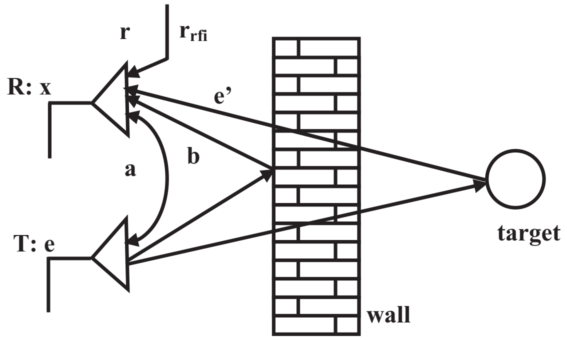

Fig. 1. Schematic diagram of through wall human detection with UWB radar.

For anomaly detection, we distinguish between two hypotheses H_0 and H_1, which represent normal status and anomaly status respectively, based on the measurement of $T(x)$. And we have:

$$T(x) = Gx \tag{2}$$

where G is the transform matrix and it can be used for acquire the principal information of the output data signal.

Assume that we have sufficient training data $T(x)$ under H_0. It is usually acceptable because we can observe data under normal status easily. And thus, we can have a good estimate of PDF of $T(x)$ under H_0. But we usually can not have enough training data to estimate the PDF of $T(x)$ under H_1 due to the anomaly status is present for only a small portion of the time. Thus our goal it to find the PDF of $T(x)$ with few available information under H_1.

The PDF of $T(x)$ under H_0 and H_1 are defined as $p(t;H_0)$, $p(t;H_1)$ respectively. To uniquely specified the $p(t;H_1)$ based on $p(t;H_0)$, we assume that $p(t;H_1)$ is close to $p(t;H_0)$ which is small signal assumption [19]. In practice, the small signal assumption is reasonable. Such as: the through wall human detection is based on the fact that the human body is always in a state of motion even if it sleeps or is trapped because of breathing. These tiny human motions would cause the scattering and reflection changes of electromagnetic wave which is emitted by UWB radar and pass through walls to reach the human body target [20]. Obviously, the difference of received signals between without person and with person is small.

Now, we use the Kullback–Leibler (KL) Divergence theory to find the $p(t;H_1)$ based on $p(t;H_0)$ [20].

$$D(\hat{p}(t;H_1)\|p(t;H_0)) = \int \hat{p}(t;H_1)\ln\frac{\hat{p}(t;H_1)}{p(t;H_0)}dt \tag{3}$$

When the $D(\hat{p}(t;H_1)\|p(t;H_0))$ is minimized with the constraint that

$$E_{\hat{p}(t;H_1)}(T(x)) = E_{p(t;H_1)}(T(x)) \tag{4}$$