Non-parametric Vertical Box Algorithm for Detecting Amplitude Information in the Received Digital Signal

Maria Beljaeva *maria.beljaeva29@gmail.com*
Scientific Centre of Applied Electrodynamics
Saint Petersburg, 190103, Russian Federation

Abstract

Vertical boxes algorithm (*VBA*) for non-cooperative automatic discrimination between digital signals with amplitude information (*AI*) from those without AI is presented. The problem is not new and several solutions have been proposed. Unlike them VBA needs no information about propagation conditions and the received signal parameters. No SNR and noise distribution assumptions are made, no carrier frequency and no thresholds are required. The only assumption made is the symbol rate interval which may be as wide as desired. The signal is considered only in the time domain. VBA also distinguishes between some other classes of signals.

Keywords: Modulation Recognition, Modulation Classification, Vertical Box Control Chart, Non-parametric.

1. INTRODUCTION

The less we know the more difficult it is to make a well-founded decision. Therefore while studying a high-uncertainty problem one hardly can resist the temptation to introduce several, sometimes unfounded, assumptions that simplify the problem statement. Modulation recognition (MR) problem of determining the received signal modulation type is not an exception. Far less methods are proposed for blind modulation recognition than for MR problems with some *a priori* information about the received signal.

There are two sources of uncertainty in MR problem. The first one is the channel nature that influences noise character and its intensity. MR methods of both *Maximum Likelihood* (ML) and *Feature Based* (FB) classes [1] are based on strict assumptions about noise distribution. The former because they are based on *Parametric Hypothesis Testing* [2] and the latter because they use thresholds to which the sample characteristics (*features*) of the signal are compared. Threshold values are calculated by means of simulation [3], and simulation requires information, in particular information about corresponding random variable distributions.

The Gaussian distribution of the background noise is the most common assumption. But *'To begin, distributions are never normal'* [4]. In fact the shot effect showed by Rice [5] to be the source of Gaussian noise is not the only reason of signal corruption. The received radio signal has passed through the medium which properties are varying in time and space and cannot be controlled completely. Nobody can foreknow the proper noise distribution kind and the only reason for using the Gaussian one is that '*Everyone believes in the normal law, the experimenters because they imagine that it is a mathematical theorem, and the mathematicians because they think it is an experimental fact*' [6]. Some empirical data show non-gaussianity of real HF signals in [7].

Other distributions are assumed sometimes [8,9] but none of them is suitable for general case.

Ergo, free distribution (non-parametric) methods are to be used for modulation recognition problems.

The second source of uncertainty is the transmitter emitting the signal. It produces different kinds of uncertainty: the set of possible modulating schemes and the values of signal parameters such as career, baud rate or SNR. The receiver knows some of this data in the *cooperative case.* The lack of knowledge *of any kind* complicates MR problem and requires special methods for its solution. For example many ML and FB methods [1] assume the carrier and phase offset are known and cannot work if they aren't; cluster methods [10] work well if number of clusters is known *a priory* etc. In the *non-cooperative* case the receiver knows nothing except the sequence of corrupted samples. The uncertainty in this situation is of greatest possible level. Many authors (see [1]) recommend prior processing (*preprocessing*) of the received signal in order to estimate some of its aforementioned parameters and then use the obtained values. Thus MR becomes a two-stage process. It seems to be too long and complicated because it requires additional studies of the robustness of MR methods to unavoidable estimation errors. We prefer another way in this paper that is to analyse the sample sequence itself without intermediate steps.

So it would be useful to form some idea of received signal modulation type when nothing is known except the distorted signal itself, not even the kind of distortion.

MR algorithms often have binary tree structure [3,11-15,19,21]. At each node decision is made whether the received signal belongs to certain class or not. Does the signal contain amplitude information (*AI*) or does it not is usually among these alternatives. Hereafter *AI* modulation set will mean all kinds of AM and QAM while *nonAI* one will mean frequency and phase modulations or just a non-modulated carrier.

Deciding if AI is present in the received signal or not is an important stage of MR, something like a stage of differential diagnosis in medicine when exclusion of some diagnosis reduces the set of alternatives.

To distinguish between AI and nonAI modulation of the received signal in non-cooperative conditions several methods were proposed. Aisbett [16] and Chan & Gadboys [17] assumed noise to be Gaussian and exploited its characteristic properties.

In order to decide is there AI in the received signal or not Ketterer [18] compares the variance of envelope with threshold. This method requires knowledge of SNR which is not available in real non-cooperative situation. Azzouz & Nandi [3] calculate statistical characteristic (*extracted feature*) γ_{max} of the received signal and compare it with the threshold $t(\gamma_{max})$ which value is determined by means of simulation, and again Gaussian assumption is made. So $t(\gamma_{max})$ may become useless if the noise has some other distribution. Another matter is that the signal envelope spectrum is used to calculate γ_{max}. This approach is viable only if symbol rate r_s is big enough so that few symbols would be located in the time interval $T_{FFT} = \frac{N_0}{f_s}$ (here N_0 is Fourier transform number of points and f_s is sampling rate). Indeed, the example in [3] has $f_s = 1200\text{kHz}$, N_0 =2048, $r_s = 12.5\text{kHz}$ so that about 21 symbols come during N_0 samples. But if $r_s = 500\text{Hz}$ then AI modulation would not be recognized, since symbol duration would exceed T_{FFT}. In more recent papers [19, 20] several time domain extracted features intended to solve AI/nonAI problem were proposed, however without any decision rules with them. In [21] peak number of the amplitude component derived from the *DoE* (Difference of Estimator filter) is used as extracted feature but, as before, the threshold is calculated under the noise normality assumption.

In this paper we propose a new *Vertical boxes algorithm (VBA)* for distinguishing between AI and nonAI modulation in the received digital signal in case when available information uncertainty is maximal i.e. the only data used is the sequence of corrupted real samples $x(t_i)$. The only assumption made is about the symbol frequency interval which can be chosen as wide as desired. No SNR or noise distribution assumptions are required, no carrier frequency estimation is done.

VBA seems to be a promising algorithm. Developed only to solve AI vs nonAI problem it distinguishes also between any modulated signal and career or noise and detects PM2 in the received signal. VBA may be efficiently applied to every recognition problem reducible to detecting of low and high periods existence in time series.

The rest of the paper is organized as follows. Section 2 presents VBA, in Section 3 we discuss it's holes and shortcomings, and describe necessary improvements. Section 4 presents simulation results, the uses of VBA to recognize some other signals are considered in Section 5. Section 6 concludes the paper.

2. VBA

2.1. Algorithm Description

To detect, without going to frequency domain, whether the received signal is modulated in any way one needs to find out whether signal oscillogram fluctuations cannot be explained only by noise. AI oscillogram consists of periods of different height. For example there are *low* and *high* periods in AM2 signal depending on the symbol transmitted (figure1).

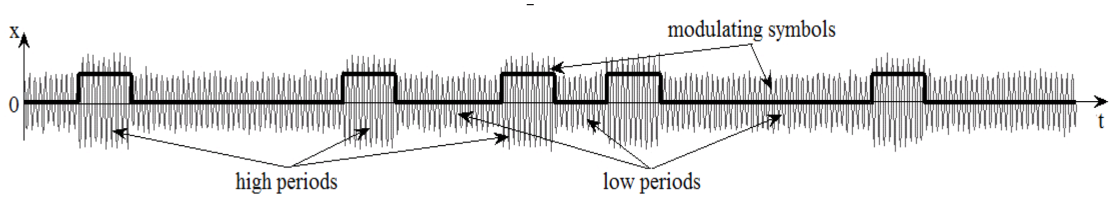

FIGURE 1: High and Low Periods of AM2 Signal, SNR=20dB.

The less is SNR the more difficult it is to detect the presence of two kinds of periods especially because we do not know the moments when symbols change. Thus the AM2 recognition problem reminds the *change-point* one of identifying stochastic process changes at unknown times [22]. In our recognition algorithm we use the idea of *Vertical Box Control Chart* [23], that is, controlling the number of random process observations which fall into the box moving along the time axis.

In-control state remains until this number *jumps* i.e. changes for more than a certain threshold, then the *out-of-control* state is diagnosed. In our approach we replace one moving box by many stationary ones and use non-parametric statistical test instead of comparing with threshold. Another statistical test is used to process results.

*Vertical box (V-bo*x) is a rectangle with one side on horizontal axis. Let $[f_{smb\,min}, f_{smb\,max}]$ be the interval of possible symbol frequency, f_s the sample frequency, and

$$t_{box} = 1/f_{smb\,max}\ . \qquad (1)$$

Consider the received signal waveform $x(t)$ over interval $(0,T)$, where $T = Kt_{box}$, K is some integer, the choice of which will be discussed below. Now sample size is $N = [Tf_s]$, and time sample moments are $t_i = \frac{i}{f_s}$, $i = 1,...,N$. VBA consists of six steps:

Step a). ***Center the signal and flip it positive*** (figure 2a).

Let

$$s(t_i) = \left| x(t_i) - \frac{1}{N}\sum_{j=1}^{N} x(t_j) \right|$$

Now average values in periods of different heights are different. Centering is necessary because noise distribution may be asymmetrical.

Step b)*.* ***Build a sequence of V-boxes over s(t)*** (figure 2b)***.*** Let $s_{max} = \max_{i=1,...,N} s(t_i)$. Draw *K* vertical boxes with height s_{max} based on consecutive subintervals of length t_{box}. Divide each V-box in two halves by a horizontal line.

Step c). ***Split all V-boxes into low and high classes according to how many samples fall under the middle line*** (figure 2c)***.*** Let v_i be the number of samples in the ith V-box below the middle line:

$$v_i = \sum_{j:(i-1)t_{box} \le j \le i\,t_{box}} \chi_j\ , \qquad (2)$$

where

$$\chi_j = \begin{cases} 1 & \text{if } s(j) < \frac{s_{max}}{2} \\ 0 & \text{else} \end{cases}.$$

Let $m_v = \frac{1}{K}\sum_{i=1}^{K} v_i$. We call the ith V-box *low* if $v(i) > m_v$ and *high* otherwise. Now we have a sequence of low and high boxes. Neighboring boxes may belong either to different or to the same classes.

Step d). ***Transform the sequence of boxes into the sequence of coffers of alternating classes*** (figure 2d). Join every subsequence of the same class boxes into one *coffer* assigning the same class to it. That gives us a sequence of coffers of varying length with alternating classes. By M denote number of coffers.

Step e). ***Make coffer borders more precise i.e. more close to symbol borders*** (figure 2e). Join each coffer $C_m = (a_m, a_{m+1}]$ with its right neighbor's left half $C' = \left(a_{m+1}, \frac{a_{m+1}+a_{m+2}}{2}\right]$ and compare the union's mean with m th coffer's mean: if

$$\frac{1}{L(C_m)} \sum_{j:t_j \in C_m} s(t_j) < \frac{1}{L(C')} \sum_{j:t_j \in C_m \cup C'} s(t_j)$$

where $L(C)$ is number of samples into C, then the m th coffer's right border is left in place, otherwise it is shifted to $\frac{a_{m+1}+a_{m+2}}{2}$. Do the same thing with the left borders. Now coffers are close to symbols if there is AM (figure 2e) and entirely random otherwise. In the first case sample distributions in neighboring coffers differ, in the second they do not. Note that some symbol may be missed (as the 9th one in figure 2e) if SNR is low.

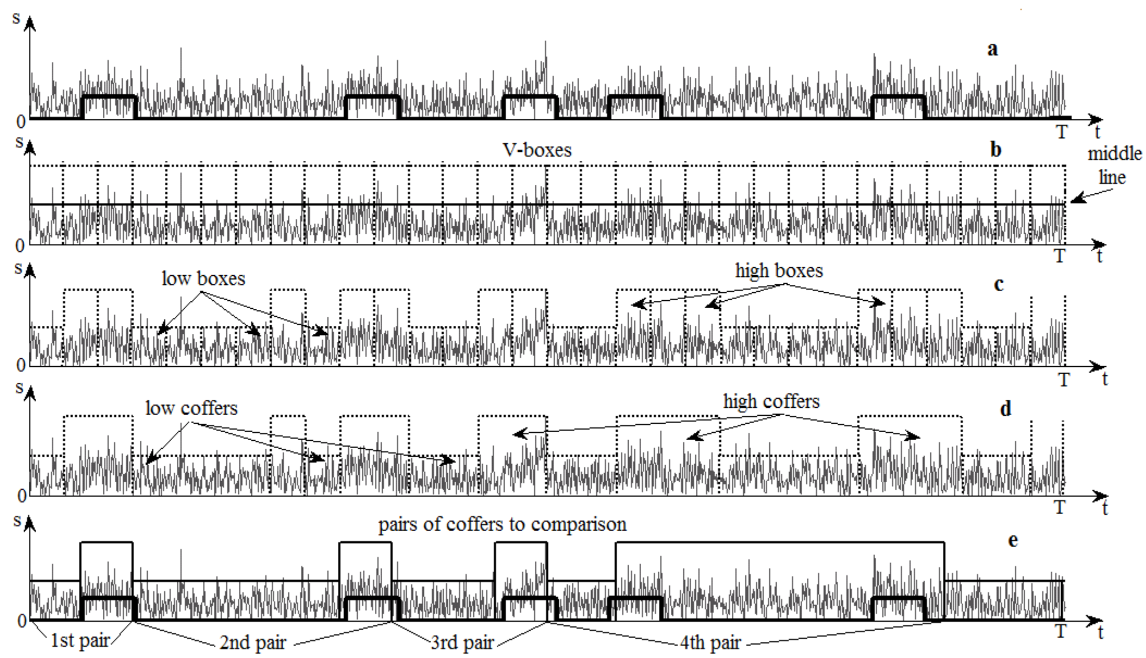

FIGURE 2: (a)-(e) Steps of VBA. SNR=3dB.

Step f). ***Compare sample distributions inside neighboring coffers content by statistical test***. We split all M coffers into $\mu = \left[\frac{M}{2}\right]$ successive non-overlapping pairs, each of them consisting of two coffers of different classes.

Then we use the well-known Mann–Whitney-–Wilcoxon (MWW) non-parametrical test [24] with significance level U, implemented by the MATLAB function *ranksum,* to test sample distribution sameness within each pair. μ test results form vector $R = \{R_1, R_2, \ldots R_\mu\}$ of random independent components where $R_k = 1$ ("failure") if distributions in k th pair are different or $R_k = 0$ ("success") otherwise. Let μ_1 be number of units in R. Obviously we decide "AI" if $\mu_1 = \mu$ and "nonAI" if $\mu_1 = 0$. But simulations show that such certain results usually occur when SNR is high enough (more than 10-15dB). Otherwise R usually becomes a mixed array of zeroes and ones. For example in figure 2 $R = \{1110\}$, which means that distributions in the 1st & 2nd coffers are not equal ($R_1 = 1$), and the same goes for 3rd & 4th and 5th & 6th distributions ($R_2 = R_3 = 1$). Distributions in 7th & 8th coffers passed MWW test successfully and could therefore be equal. The 9th coffer has no pair and is not taken into consideration.

Step g). *Use another statistical test to analyze the results of step "f"*. The conditional probability $P\{R_i = 1 | \text{nonAI}\}$ is less than U by the construction of R. We check the hypothesis that the success probability p_s is also less than U, using a *test involving a single binomial probability* [2] which is implemented by the MATLAB function *finv*. If the hypothesis is confirmed on significance level U' our algorithm says "nonAI", otherwise it says "AI". Simulations show that in the absence of noise VBA recognizes AM2 signals with confidence when symbol length is no less than t_{box}.

2.1. The choice of K

Simulations show that for $SNR > 5dB$ VBA gives good results if M is near 10. Obviously M is strongly related to symbol variation number (SVN). When noise is absent these numbers are equal or differ by 1. The more is noise the more they differ. The distribution of SVN is binomial, and according to [24] (table D) K_0=30 symbols are needed to obtain 10 symbol variations with probability 0.95 assuming that symbol probabilities are equal. So $K = K_0$ is recommended in a special case when symbol length is equal to t_{box} (1).

Now suppose symbol length equals Lt_{box},. where $L > 1$. Again K_0 symbols are needed therefore K should be equal to $[K_0 L]$. Now if symbol length varies from $1/f_{smb\,max}$ to $1/f_{smb\,min}$ we would need

$$L = \frac{f_{smb\,max}}{f_{smb\,min}} \tag{3}$$

to recognize AM2. Thus the total sample number is

$$\hat{N}_{VBA} = [K_0 L] \times [t_{box} f_s]. \tag{4}$$

Note that the 1st factor is the number of V-boxes, while the 2nd one is the number of samples per V-box. Using (1,2), we get

$$\hat{N}_{VBA} = \left[K_0 \frac{f_s}{f_{smb\,min}} \right].$$

3. VBA DISCUSSION AND IMPROVEMENTS

3.1. VBA Holes and Shortcomings

First hole is the use of MWW test, which requires sample groups to be independent. That doesn't seem to be the case for coffers contents. At least the numbers of samples in different class coffers are correlated. For example, if the ith box is high and the jth one is low, then ν_i, ν_j (2) are order statistics (ν_i has the smaller index than ν_j). Since the latter are correlated, so are the sample arrays within V-boxes, and, therefore, coffers' contents. But correlation is weak and medians of sample distributions within low and high coffers differ much more in AI case than in nonAI, which is why MWW test distinguishes between them (see Section 4).

Second hole is the interpretation of MWW test results. Strictly speaking sample distributions in different class boxes are not the same even in nonAI case. The means of these distributions are different, again due to the way classes are chosen. But the difference is rather small and the test usually doesn't "feel" it when sample number within one V-box is not too large.

First VBA shortcoming is that it's sensitive to signal aliasing. Signal without modulation could be mistakenly determined to be amplitude modulated (figure.3).

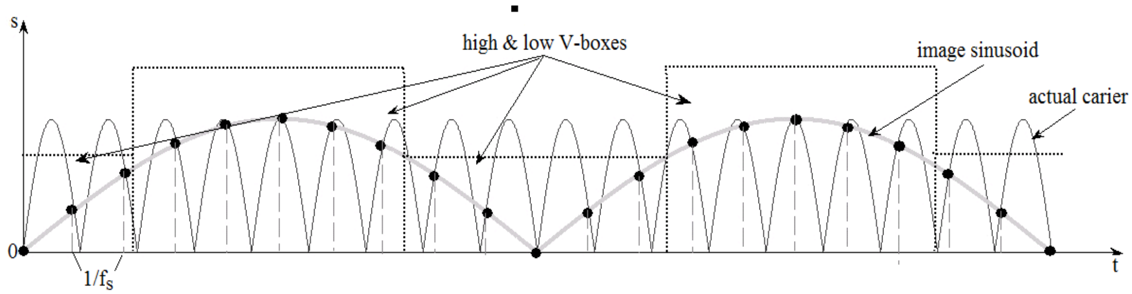

FIGURE 3: Non-modulated Carrier Looks like AI Signal Because of Aliasing.

Second shortcoming is that VBA is computationally slow because of large number of proceeded samples. For example let $f_{smb\,min} = 100\text{Hz}$, $f_{smb\,max} = 3\text{KHz}$, $f_s = 20\text{MHz}$. According to (4) $\hat{N}_{VBA} = 6 \cdot 10^6$; there are about 900 V-boxes with more than 6500 samples per box. Both shortcomings can be removed by the same trick namely the randomization of sample times (see below). The number of samples is lowered even more by using V-boxes of different lengths.

3.2. Improving VBA

Decreasing the Number of Samples per V-box

To determine the box's class it's enough to take much less samples than $\left[f_s / f_{smb\,min} \right]$ if they are taken in uniformly distributed moments. Let us take D pseudorandom numbers $\xi_j, j = 1,...,D$, uniformly distributed in *(*0,1*),* and sample the signal at the moments $t_{D(i-1)+j} = (i - 1 + \xi_j) t_{box}$. Thus there would be D samples within each V-box. Note that the same array $\{\xi_i\}$ is used for all V-boxes. Therefore the total sample size decreases significally from $\hat{N}_{VBA}$ to $N_{VBA} = KD$. What's more, this approach removes the first shortcoming too due to nonregularity of time sampling.

Randomization also helps to diminish the second hole i.e. to reduce the tendency of the MWW test to give the false positive result i.e. to detect AI when there is none. Create another sequence of D' samples exactly the same way as the first one (using another array of pseudorandom numbers). This second sequence would then be used as input for the MWW test. If there is no AI in the received signal then the only difference between low and high coffers would come from the noise, hence it will be neutralized while comparing samples taken in another array moments within coffers. Otherwise, if there is AI in the signal, changing the sample moments array will not decrease the difference greatly. Now there are $D + D'$ samples per V-box. Simulations show that $D = D' = 200$ is a good choice.

Decreasing V-box number

Let us divide the symbol length interval $\left[t_{smb\,min}, t_{smb\,max} \right] = \left[1/f_{smb\,max}, 1/f_{smb\,min} \right]$ to m subintervals and denote points of division by $\tau_1,...,\tau_{m-1}$, $\tau_0 = t_{smb\,min}$, $\tau_m = t_{smb\,max}$. Then we use VBA with $t_{box} = \tau_i$ in order to recognize AM in the received signal with the symbol length belonging to $[\tau_i, \tau_{i+1}]$. Now the total number of different length V-boxes is $K_0 I(m, \tau_0,...,\tau_m)$, where

$$I(m, \tau_0, ..., \tau_m) = \sum_{i=0}^{m-1} \left[\frac{\tau_{i+1}}{\tau_i} \right]. \quad (5)$$

Now the problem is to find the best way of division.

Theorem. Let

$$\alpha = [\ln L],$$

$$m^* = \begin{cases} \alpha & \text{if } \frac{\alpha}{\alpha+1} L^{\frac{1}{\alpha(\alpha+1)}} < 1 \\ \alpha + 1 & \text{otherwise} \end{cases} \quad (6)$$

$$\lambda = \sqrt[m^*]{L}.$$

Then

$$\min_{m, \tau_1, ..., \tau_{m-1}} I(m, \tau_0, ..., \tau_m) = I\left(m^*, \{\tau_0 \lambda^i\}_{i=0}^{m^*} \right) = \lambda m^* \quad (7)$$

Proof

Note that we can only vary $\tau_1, ..., \tau_{m-1}$ since τ_0 and τ_m are fixed. Assume that $m = 2$. According to (5) $I(2, \tau_0, \tau_1, \tau_2) = \frac{\tau_1}{\tau_0} + \frac{\tau_2}{\tau_1}$.

It has a minimum when $\tau_1 = \sqrt{\tau_0 \tau_2}$. It follows that for any $m > 2$, $I(m, \tau_0, ..., \tau_m)$ would be minimal when the points of division satisfy simultaneous equations $\tau_i = \sqrt{\tau_{i-1}\tau_{i+1}}, i = 1, ...m-1$. Therefore

$$\tau_i = \tau_0 L^{\frac{i}{m}}, i = 1, ...m-1. \quad (8)$$

Substituting (8) to (5), we obtain

$$I(m, \tau_0, ..., \tau_m) = m L^{\frac{1}{m}}. \quad (9)$$

Function $f(x) = x \cdot \sqrt[x]{L}$ has a unique minimum at $x = \ln L$. Since m takes only integer values then, in order to minimize I, it must be equal to α if $f(\alpha) < f(\alpha+1)$ and to $\alpha + 1$ otherwise. □

Now the total number of V-boxes is $\lambda K_0 m^*$ and the total sample number is equal to

$$N_{VBA} = \lambda K_0 m^* (D_1 + D_2) \quad (10)$$

instead of (4). Note that (10) does not include sampling rate because of sample time randomization. Under the conditions of section 3 example $m^* = 3, \lambda \approx 3.1$. If $K_0 = 30$, $D_1 = D_2 = 200$ then about 280 V-boxes and 112000 samples are needed. So in this example the improved VBA reduces the required samples number in more than 50 times.

It's possible to use parallel computations to speed up signal processing. There are m^* subsequences of λK_0 V-boxes in each one. The length of V-boxes in ith subsequence is $\tau_{i-1} = \tau_0 \lambda^{i-1}, i = 1, ...m^*$. If all subsequences start to be proceeded simultaneously then the length of the required signal piece T is equal to the length of the longest subsequence i.e $K_0 \tau_0 \lambda^{m^*-1}$.

Under the conditions of section 3A example $T = 30\tau_2 \approx 0.1$ s. Keep in mind that such a long time is the charge for symbol rate uncertainty. If, hypotheticaly, the symbol rate is known (as in [3]) then the signal duration about $\frac{K_0}{r_s} = 0.0024$ s would be enough. It is about the same as required by the algorithm described in [3]: the latter needs $\frac{2048}{f_s} \approx 0.0017$ s but cannot recognize AI when the symbol rate interval is as wide as in our example.

While proceeding m^* subsequences we obtain m^* results: $r_1, r_2, ..., r_{m^*}$, where $r_i = 1$ if AI is recognized for symbol length belonging to (τ_{i-1}, τ_i), $r_i = 0$ otherwise. The final decision is made as following: if $m^* \le 2 \,\&\, \varepsilon \ge 1$ or $m^* > 2 \,\&\, \varepsilon \ge 2$ where $\varepsilon = \sum_{i=1}^{m^*} r_i$ then signal contains AI, otherwise it does not.

Since VBA is built around the recognition of different height oscillogram periods it can detect the presence of AI not only in AM2 signals but also in other AI modulated signals. Naturally the detection probability depends on the difference between average heights of high and low periods.

4. SIMULATION

Simulation was performed under the conditions of section 3 example. Received signal was simulated as follows. Sequence of equiprobable and independent zeros and ones modulated the carrier, the result was corrupted by additive noise. 100 realizations were generated for every variant.

Modulation Types and Noise.

VBA was tested on 10 types of modulation: AM2, 4-level PAM, PSK2, PSK4, FSK2, FSK4, three kinds of QAM8 shown at figure 4a-c, QAM16 (figure 4d), and non-modulated carrier.

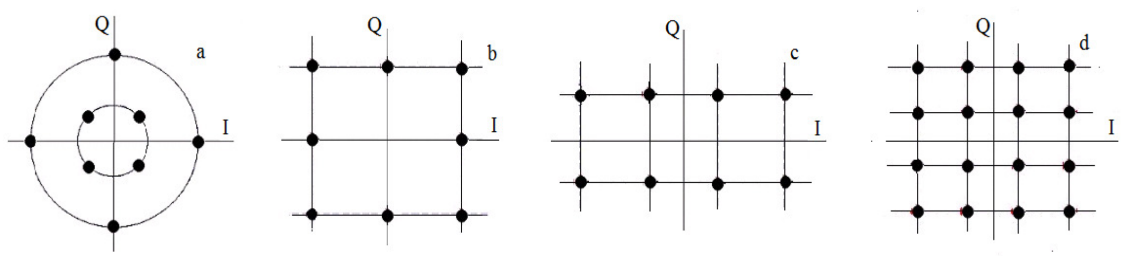

FIGURE 4: Simulated Kinds of QAM.

Two kinds of noise distributions (Gaussian and a mixture of uniform and Rayleigh distributions) and three SNR levels (10, 5, and 0 dB) were simulated.

Algorithm Parameters.

$D = D' = 200$, $K_0 = 30$, $U = U' = 0.05$.

A=1; carrier=7.5 MHz, modulation index 0.6 for AM2; modulation indices 0.2, 0.4, 0.6 for PAM. Symbol
frequency do not influence the results thus their values were chosen arbitrarily.

Simulation Results are presented in Table 1. The numerator in each cell of Table 1 is the probability for Gaussian noise while the denominator is the same for mixture noise. We see that VBA confidently detects AI in all considered signals when $SNR \ge 5dB$. As SNR decreases the correct decision probability decreases as well. Its rate of decrease is greater if the average difference between low and high periods in the signal is smaller. That's why AI in PAM4, QAM8(b), QAM16 is recognized with much less confidence than in the other modulation types when SNR=0 (see the last line in Table 1).

	carrier	AM2	PAM4	PM2	PM4	FM2	FM4	QAM8(1, 2, 3)	QAM16
10dB	$\underline{0.97}$ 0.97	$\underline{1.0}$ 1.0	$\underline{0.98}$ 1.0	$\underline{0.97}$ 0.94	$\underline{0.96}$ 0.98	$\underline{0.97}$ 0.99	$\underline{0.99}$ 0.98	$\underline{1.0}$, $\underline{0.99}$, $\underline{1.0}$ 1.0 1.0 1.0	$\underline{0.99}$ 0.98
5dB	$\underline{0.98}$ 0.96	$\underline{0.99}$ 1.0	$\underline{0.99}$ 0.97	$\underline{0.95}$ 0.99	$\underline{0.98}$ 0.97	$\underline{0.97}$ 0.98	$\underline{0.98}$ 0.95	$\underline{0.99}$, $\underline{1.0}$, $\underline{1.0}$ 1.0 1.0 1.0	$\underline{0.99}$ 1.0
0dB	$\underline{0.98}$ 0.97	$\underline{0.94}$ 0.92	$\underline{0.34}$ 0.23	$\underline{0.98}$ 0.97	$\underline{0.96}$ 0.98	$\underline{0.97}$ 0.97	$\underline{0.99}$ 0.98	$\underline{1.0}$, $\underline{0.75}$, $\underline{1.0}$ 0.98 0.67 0.97	$\underline{0.76}$ 0.68

TABLE 1: Probability of Correct Decision «AI vs. nonAI». $K = \lambda K_0 = 93$

It is difficult to compare these results with those obtained previously because of difference in basic assumptions. Methods [3,16,17,18-21] proposed for the *AI vs non AI* problem are based on Gaussian hypothesis. Besides, some signal characteristics are assumed to be known, e.g. carrier frequency in [3,20,21], and symbol rate in [20], and SNR in [16,18], although not all referenced articles mention it. Moreover the probability of correct decision is not estimated in some articles at all [16,19], while in others the problem is considered as a part of binary tree algorithm [3,19,21] so only integral estimates are presented. Thus the only characteristic we can use to compare these methods is signal-to-noise ratio, specifically its minimum value SNR_{min} for which good recognition results are claimed to be obtained. It equals -1dB in [16], 0dB in [19,20], 5dB in [3], 7dB in [17], and 10dB in [18,21]. Note that the best results in [16,20] require *a priori* knowledge of the received signal characteristics. In the case SNR=0, [19] shows good results, but only for those modulation kinds VBA recognizes with high probability too.

Now we can conclude that VBA detects the presence or absence of AI in the signal well when neither it characteristics nor the noise level and distribution are known.

5. THE USE OF VBA FOR SOME OTHER PROBLEMS

Not only signals carrying true amplitude information are classified by VBA as AI modulated signals but also all other signals with high and low periods. We shall call such signals *a la AI* ones.

Let $s^{(+\Delta t)}$ be the sum of signal and its shifted to Δt copy and let $s^{(*\Delta t)}$ be their product:

$$s^{(+\Delta t)}(t) = s(t) + s(t - \Delta t) \tag{11}$$

$$s^{(*\Delta t)}(t) = s(t)s(t - \Delta t) \tag{12}$$

Note that if $s(t)$ is pure carrier or noise then $s^{(+\Delta t)}(t)$ is not *a la AI* signal (fig.5).

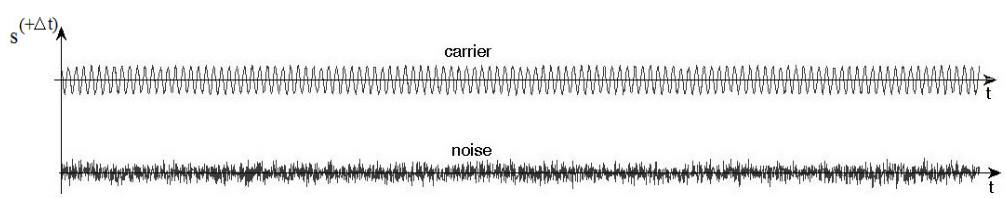

FIGURE 5: Sums of Carrier and Gaussian Noise With Their Shifted Copies; $\Delta t = 1.2 t_{smb\,min}$; SNR=20dB.

On the contrary, $s^{(+\Delta t)}$ is *a la AI* signal when $s(t)$ carries frequency or phase modulation (fig.6). Being flipped such a signal remains *a la AI* one, and VBA detects AI even when SNR is quite low (e.g. 3dB in fig.7; R={1110}). Thus the presence of phase or frequency modulation in the received signal is recognized by means of VBA.

To distinguish between phase and frequency modulations in the received signal let's consider $s^{(*\Delta t)}$ (fig.8).

Note that $s^{(*\Delta t)}$ is *a la AI* for all modulated signals and $\left|s^{(*\Delta t)}\right|$ is not *a la AI* only for PM2 signal. So to distinguish PM2 from the other signals one has to apply VBA not only to the flipped signal $\left|s^{(*\Delta t)}\right|$ as it is described in Section 2 but also to the signal shifted up:

$$s_{up}(t,\Delta t) = s^{(*\Delta t)}(t) + \left|\min_{t\in(0,T-\Delta t)} s^{(*\Delta t)}\right|. \qquad (13)$$

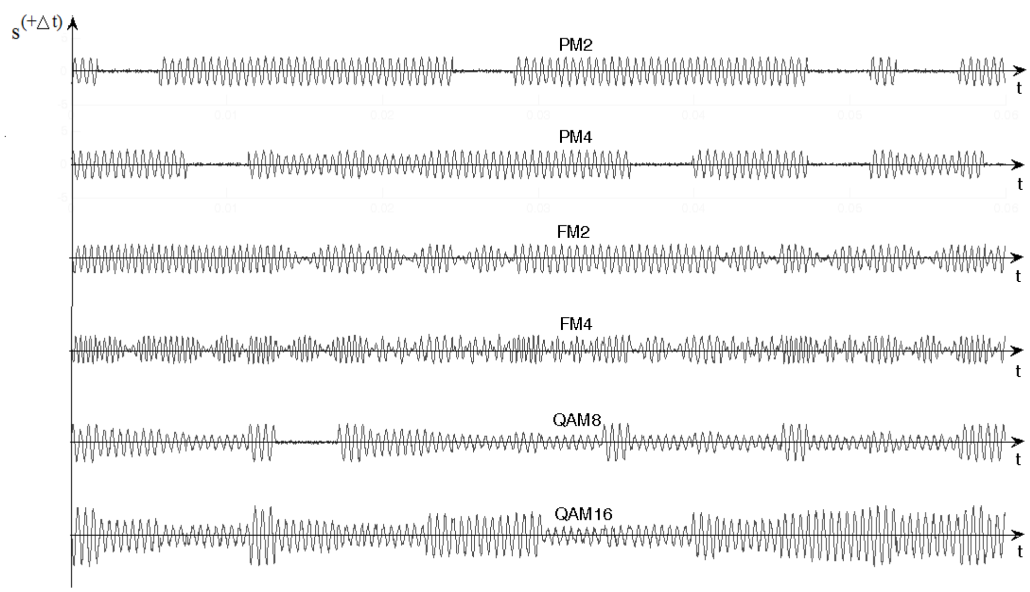

FIGURE 6: Sums of Modulated Signals and Their Shifted Copies; $\Delta t = 1.2t_{smb\,min}$, SNR=20dB.

If the first answer is nonAI and the second is AI then the signal is PM2.

To distinguish between PM and FM signals the more accurate VBA modification is needed. It may be achieved by breaking up boxes in more than two parts (step ***e***, Section II), and it would take more time. If we build coffers for $s_{up}(t,\Delta t)$ accurately enough, the scatter of samples within them would differ when the received signal is FM and not differ when it is PM. So to distinguish one from another one would have to compare scatters by some non-parametrical test for variances [24].

These and other distinguishing characteristics analyzable by means of VBA would be the subject to further investigation. One more idea to be examined in the future is applying VBA to analogous modulation recognition.

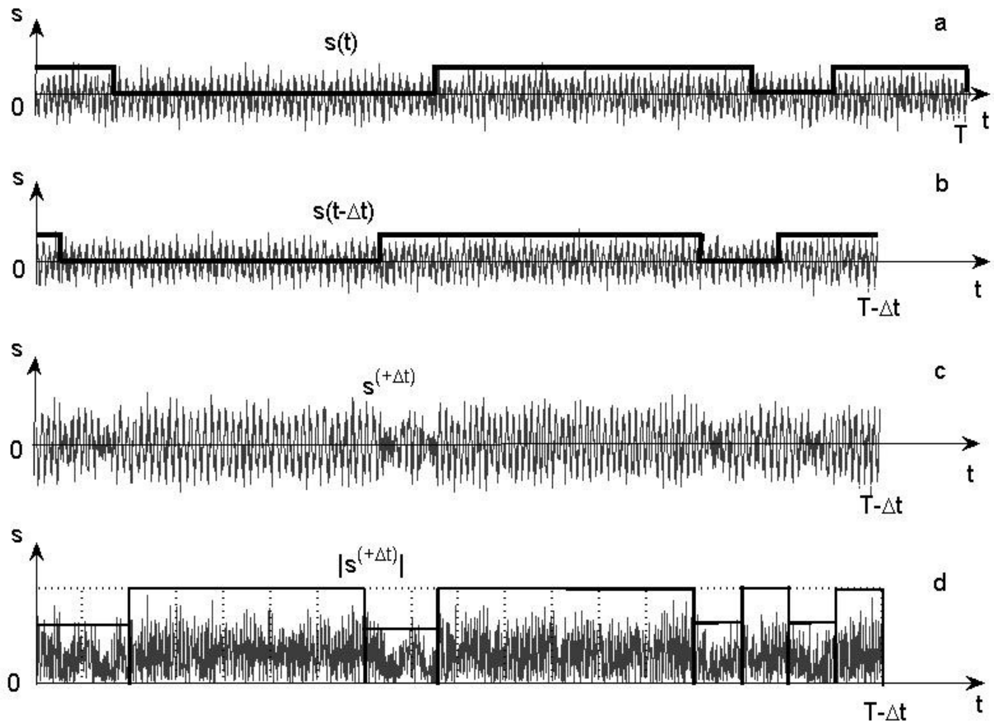

FIGURE 7: a) PM2-signal b) Its Shifted Copy c) Their Sum d) The Result of VBA: Four Pairs of Coffers.

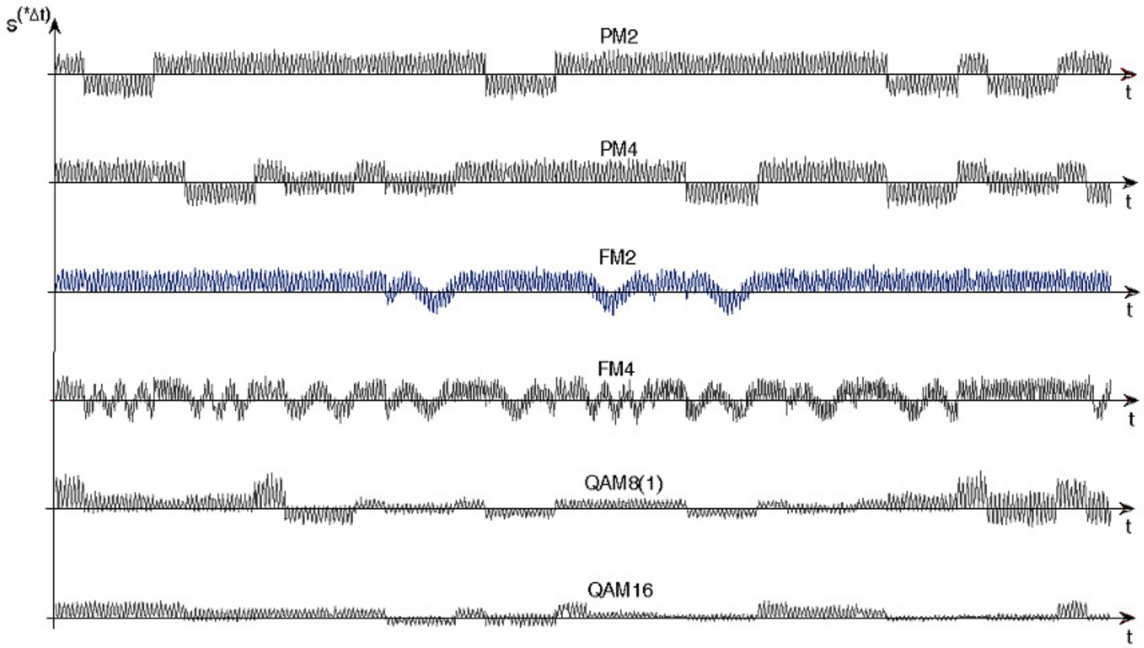

FIGURE 8: Products of Modulated Signals and Their Shifted Copies; $\Delta t = 1.2 t_{smb\,min}$, SNR=20dB.

6. CONCLUSION

The idea of VBA is clear enough: the algorithm "sees" AI as a man does. In fact method [3] does the same in frequency domain; the difference is that man looks at spectrum, not at oscillogram. However in time domain it is possible to use non-parametrical statistical tests while at the moment there are no such tests for frequency domain. That's why there is no method of recognizing modulation type in frequency domain automatically without choosing some thresholds. This approach inevitably limits the assortment of considered modulation kinds and types of noise while VBA realizes more general approach.

In practice, VBA's advantage is that it doesn't require any information not available in reality, whilst other algorithms do. At the same time VBA works well under the same SNR conditions as other algorithms.

VBA is simple and rough enough, and that is quite normal: problems with high degree of uncertainty need rough solutions.

7. REFERENCES

[1] O. A. Dobre, A. Abdi,. Y. Bar-Ness, and W. Su. "Survey of automatic modulation classification techniques: classical approaches and new trends." *Communications, IET,* vol. **1**, iss. 2, 2007, pp. 137-156.

[2] J. H. Pollard. *Handbook of Numerical and Statistical Techniques***.** Cambridge: Cambridge University Press, 1977, pp. 144-145.

[3] E. E. Azzouz and A. K. Nandi. *Automatic Modulation Recognition of Communication Signals*. Boston, Mass, USA: Kluwer Academic Publishers, 1996.

[4] R. Wilcox., *Introduction to Robust Estimation and Hypothesis Testing*. 2nd ed, San Diego, CA: Elsevier Academic Press, 2005, p. 2.

[5] S. O. Rice. "Mathematical analysis of random noise." *Bell Systems Technical Journal*, vol. 23, 1944, pp. 282–332.

[6] G. Lippmann. "Conversation with Henri Poincaré." In Henri Poincaré, *Calcul des Probabilités*, 1896, p. 171.

[7] J. Giesbrecht, R. Clarke, and D. Abbot. "An empirical study of the probability density function of HF noise." *Fluctuation and Noise Letters*, vol. 6, no. 2, 2006, pp.117-125.

[8] V G. Chavali and C.R.C.M. da Silva. "Maximum-likelihood classification of digital amplitude-phase modulated signals in flat fading non-Gaussian channels." *IEEE Transactions on Communications, vol. 59, iss. 8,* 2011, pp. 2051-2056.

[9] A. Hazza, M. Shoaib, A. Saleh, and A. Fahd. "Classification of Digitally Modulated Signals in Presence of Non-Gaussian HF Noise." in *Proc. of ISWCS, 2010, pp.815-819.*

[10] B. G. Mobasseri. "Constellation Shape as a Robust Signature for Digital Modulation Recognition." *In Proc. of IEEE Veh. Technol. Conf.*, vol.1, 1999, pp.442-446.

[11] D. Boudreau, C. Dubuc, F. Patenaude, M. Dufour, J. Lodge, and R. Inkol, "A fast automatic modulation recognition algorithm and its implementation in a spectrum monitoring application." in *Proc. of IEEE MILCOM'2000*, 2000, pp. 732-736.

[12] L. De Vito, S. Rapuano, and M. Villanacci. "An improved method for the automatic digital modulation classification." in *Proc. of IEEE IMTC'2008*, 2008,pp. 1441-1446.

[13] N. Björsell, P. Daponte, L. De Vito, and S. Rapuano. "Automatic signal recognition for a flexible spectrum management." in *Proc.of IEEE IMEKO*, 2009, pp. 568-573.

[14] L. Anton and I. Vizitiu. "Some aspects of signal recognition." in *Proc. of Distance Learning, Simulation and Communication International Conference,* 2011, Brno, Czech Republic, pp .9-15.

[15] Y. Xu L. Ge, and B. Wang. "Digital modulation recognition method based on tree-structured neural networks." in *Proc of ICCSN'09*, 2009, pp. 708 – 712.

[16] J. Aisbett. "Automatic modulation recognition using time-domain parameters." *Signal Processing*, vol. 13, No. 3, 1987, pp. 323-329.
[17] Y. T. Chan and L. G. Gadbois. "Identification of the modulation type of a signal." *Signal Processing,* vol. 16, No. 2, 1989, pp. 149-154.

[18] H. Ketterer, F. Jondrall, and A. H. Costa, "Classification of modulation modes using time-frequency methods," in *Proc. of IEEE ICASSP'1999*, vol.5, 1999, pp. 2471-2474.

[19] Y. Jin, S. Li, Z. Yang, and Y. Wang. "Study of a novel key feature in non-cooperative modulation automatic recognition." in *Proc. of WiCom'200*7, 2007, pp. 1240-1243.

[20] J. Bagga and N. Tripathi. "Analysis of digitally modulated signals using instantaneous and stochastic features, for classification." *International Journal of Soft Computing and Engineering (IJSCE)*, vol.1, no. 2, May, 2011, pp. 57-61.

[21] M. Kang, C. Lee, and J. Joo. "Automatic recognition of analog and digital modulation signals using DoE filter." in *Proc. of ISCIT'2009*, 2009, Incheon, Korea, pp. 609-614.

[22] U. Jensen and C. LÄutkebohmert. "Change-Point Models," in *Encyclopedia of Statistics in Quality and Reliability*, vol. 1, New York: Wiley, 2007, pp. 306-312.

[23] E. Rafajlowicz, M. Pawlak, and A. Steland. “Nonparametric sequential change-point detection by a vertically trimmed box method.” IEEE T Inf Th, vol 56, no 7, 2010, pp 3621-3634.

[24] S. Siegel and J. Castellan. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill Book Company, 1988, pp. 128-137.