AUTOMATIC TARGET RECOGNITION ALGORITHM FOR LOW–COUNT TERAHERTZ IMAGES

V.E. Antsiperov
Kotel’nikov Institute of Radioengineering and Electronics, Moscow, Russia, Russian Academy of Sciences

Abstract

The paper presents the results of developing an algorithm for automatic target recognition in broadband (0.1–10) terahertz images. Due to the physical properties of terahertz radiation and associated hardware, such images have low contrast, low signal–to–noise ratio and low resolution – i.e. all the characteristics of a low–count images. Therefore, standard recognition algorithms designed for conventional images work poorly or are not suitable at all for the problem considered. We have developed a fundamentally different approach based on clustering 2D point clouds in accordance with a set of predefined patterns. As a result, we reduce the problem of target recognition to the problem of maximizing the image data likelihood with respect to the classes of model objects up to the size and position. The resulting recognition algorithm has a structure close to that of the well-known EM algorithm; its formal scheme is at the end of the paper.

Keywords: automatic target recognition, concealed objects detection, low–count images, THz imaging, EM–algorithm, classification, image recognition.


Acknowledgments: The author wants to express his gratitude for the financial support of the RFBR, grant 16-29-09626 ofi_m.

Introduction

Terahertz (THz) radiation consists of electromagnetic waves with frequencies in the range between infrared and microwave radiation – from 0.1 to 10 THz [1]. Since this frequency range comprises wavelengths from 3 mm to 30 μm (see Fig. 1), the terahertz range is often referred to as a submillimetre range.

<table>
<thead>
<tr>
<th>Wavelength (μm)</th>
<th>Radio</th>
<th>Microwave</th>
<th>THz</th>
<th>Infrared</th>
<th>Visible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (GHz)</td>
<td>1670</td>
<td>100–1000</td>
<td>1–10</td>
<td>100–1000</td>
<td>1–10</td>
</tr>
</tbody>
</table>

Fig. 1. The terahertz (THz) spectral range

A number of useful properties of THz radiation are the same as properties of neighboring ranges on the spectrum. As well as infrared and microwave radiation, THz radiation is propagated in the line of sight (LOS) and is nonionizing (as opposed to X rays). Like microwave radiation, THz radiation can penetrate through nonconductive material: clothes, paper, wood, plastic, etc. (however, it should be borne in mind that, as a rule, THz penetration depth is slightly smaller, and, moreover, it cannot penetrate through liquids [2]). Since THz radiation exhibit good penetrating power, it can be used to obtain images of hidden objects. For this reason it is a good basis for automatic target recognition (ATR) systems [3] developed for early detection and warning of threats. Because detection of such threats is one of key issues of public places security, it is obvious the great interest in such systems. A lot of them either have been developed or are at the evaluation stage. Accordingly, in recent years there is an extraordinary growth of publications on this subject, and we observe it both as in foreign editions [1, 4–7] as in Russian ones [8–10].

However, a number of existing problems partially diminish the initial optimism about the THz–based ATR. First of all, THz radiation is usually weak (including the case of active illumination). Therefore, THz images are characterized by low signal–to–noise ratio. This leads to a low contrast, fuzzy shapes of the objects on the background scene. Secondly, because of limited sensitivity of THz detectors, the THz image should be formed by time–consuming scan procedures (sometimes forming individual pixels). As a result, the obtained images contain a small number of resolution elements – they belong to the class of low–count images. Therefore, in contrast to the images in the visible range, which can be approximated with good accuracy by a continuous distribution of intensity within the image plane, THz images are rather 2D clouds of discrete points with a small number of graduations of intensity, often binary (0/1) images, see Fig 2 (A – visible image, B – THz image, C – infrared image).

Fig. 2. Image quality for different spectral ranges: visible, THz and infrared images [3]

From the above arguments it follows that the standard recognition algorithms designed for conventional images suitable for ATR will work poorly in the THz range. Therefore, the task of developing new THz image processing algorithms and techniques in the context of ATR is crucial today.

1. Automatic target recognition for low–count images

Automatic target recognition (ATR) technologies [11], as a rule, comprise the use of computer hardware for systems of detection and recognition of controlled, hidden objects of interest by processing data images from cameras, antennas, radars and other sensors, for example THz sensors.

Automatic target recognition algorithm for low–count terahertz images

V.E. Antsiperov
Kotel’nikov Institute of Radioengineering and Electronics, Moscow, Russia, Russian Academy of Sciences

Abstract

The paper presents the results of developing an algorithm for automatic target recognition in broadband (0.1–10) terahertz images. Due to the physical properties of terahertz radiation and associated hardware, such images have low contrast, low signal–to–noise ratio and low resolution – i.e. all the characteristics of a low–count images. Therefore, standard recognition algorithms designed for conventional images work poorly or are not suitable at all for the problem considered. We have developed a fundamentally different approach based on clustering 2D point clouds in accordance with a set of predefined patterns. As a result, we reduce the problem of target recognition to the problem of maximizing the image data likelihood with respect to the classes of model objects up to the size and position. The resulting recognition algorithm has a structure close to that of the well-known EM algorithm; its formal scheme is at the end of the paper.

Keywords: automatic target recognition, concealed objects detection, low–count images, THz imaging, EM–algorithm, classification, image recognition.


Acknowledgments: The author wants to express his gratitude for the financial support of the RFBR, grant 16-29-09626 ofi_m.

Introduction

Terahertz (THz) radiation consists of electromagnetic waves with frequencies in the range between infrared and microwave radiation – from 0.1 to 10 THz [1]. Since this frequency range comprises wavelengths from 3 mm to 30 μm (see Fig. 1), the terahertz range is often referred to as a submillimetre range.

<table>
<thead>
<tr>
<th>Wavelength (μm)</th>
<th>Radio</th>
<th>Microwave</th>
<th>THz</th>
<th>Infrared</th>
<th>Visible</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency (GHz)</td>
<td>1670</td>
<td>100–1000</td>
<td>1–10</td>
<td>100–1000</td>
<td>1–10</td>
</tr>
</tbody>
</table>

Fig. 1. The terahertz (THz) spectral range

A number of useful properties of THz radiation are the same as properties of neighboring ranges on the spectrum. As well as infrared and microwave radiation, THz radiation is propagated in the line of sight (LOS) and is nonionizing (as opposed to X rays). Like microwave radiation, THz radiation can penetrate through nonconductive material: clothes, paper, wood, plastic, etc. (however, it should be borne in mind that, as a rule, THz penetration depth is slightly smaller, and, moreover, it cannot penetrate through liquids [2]). Since THz radiation exhibit good penetrating power, it can be used to obtain images of hidden objects. For this reason it is a good basis for automatic target recognition (ATR) systems [3] developed for early detection and warning of threats. Because detection of such threats is one of key issues of public places security, it is obvious the great interest in such systems. A lot of them either have been developed or are at the evaluation stage. Accordingly, in recent years there is an extraordinary growth of publications on this subject, and we observe it both as in foreign editions [1, 4–7] as in Russian ones [8–10].

However, a number of existing problems partially diminish the initial optimism about the THz–based ATR. First of all, THz radiation is usually weak (including the case of active illumination). Therefore, THz images are characterized by low signal–to–noise ratio. This leads to a low contrast, fuzzy shapes of the objects on the background scene. Secondly, because of limited sensitivity of THz detectors, the THz image should be formed by time–consuming scan procedures (sometimes forming individual pixels). As a result, the obtained images contain a small number of resolution elements – they belong to the class of low–count images. Therefore, in contrast to the images in the visible range, which can be approximated with good accuracy by a continuous distribution of intensity within the image plane, THz images are rather 2D clouds of discrete points with a small number of graduations of intensity, often binary (0/1) images, see Fig 2 (A – visible image, B – THz image, C – infrared image).

Fig. 2. Image quality for different spectral ranges: visible, THz and infrared images [3]

From the above arguments it follows that the standard recognition algorithms designed for conventional images suitable for ATR will work poorly in the THz range. Therefore, the task of developing new THz image processing algorithms and techniques in the context of ATR is crucial today.

1. Automatic target recognition for low–count images

Automatic target recognition (ATR) technologies [11], as a rule, comprise the use of computer hardware for systems of detection and recognition of controlled, hidden objects of interest by processing data images from cameras, antennas, radars and other sensors, for example THz sensors.
The fundamental problem of ATR is detection and identification of objects (targets) of interest in the context of complex scenes with other objects registered, often in a very noisy environment. The precise definition of the concepts of a target, scene and noise depends on a particular application. In case of ATR the term classification is often used instead of the term identification, and although there are subtle differences (identification is a more precise category), when solving practical problems, these differences are usually ignored, and the term recognition is used instead of above two terms.

Practical ATR systems, as a rule, include a pipeline of operations as shown in Fig. 3 [11]. Ideally, the targets of the original image are subsequently detected, recognized by pipeline operations and included in the output list of targets. While original data are moving through the pipeline, data processing becomes more specific and focused on certain target attributes. As a result, the amount of data associated with non-target objects must gradually decrease. Since there is usually a lot of non-relevant objects in the scene and very few target objects (sometimes none), very sophisticated, non-trivial algorithms for original data processing, image segmentation and object recognition are required.

Fig. 3. Conceptual diagram of data processing pipeline in ATR systems

This paper presents the results of developing an algorithm for one of the procedures executed in the pipeline – the algorithm of automatic target objects (briefly targets) recognition. A special feature of the algorithm is that it is initially focused on the specifics of terahertz images containing low contrast, low-count objects with a low signal-to-noise ratio [12], such as presented in Fig. 4 (A – a ceramic knife and a handgun hidden under the clothes; B – a handgun and a rectangular piece of radio anechoic material under the clothes; C – a ceramic knife also hidden under the clothes). It is assumed that the algorithm receives at its input a fragment of the scene already containing an object detected at the preceding steps of the processing pipeline. The purpose of the algorithm is to recognize the object according to the specified classification database (DB), including as a rule the classes of target objects.

Fig. 4. THz images, the radiometric temperature is shown on the right in grayscale [12]

2. Classification DB

Classes of objects in Classification DB are families of low–contrast, low–count images of similar subjects, which can be arbitrarily positioned or scaled.

Representatives of such classes (a, b, c), for example, for the ATR system that recognizes the images in Fig. 4, are shown in Fig. 5 (a – ‘knife’, b – ‘rectangular’, c – ‘handgun’, respectively). Let us first specify what is meant by similarity of subjects and leave the questions of their permissible locations and sizes for the next section.

Fig. 5. Formalized description of three classes of objects in the form of Gaussian mixture parameters (centers and elliptical contours of four Gaussian components are shown over images)

A main recognition problem is the fact that original data are usually presented in a form that is hardly suitable for their immediate recognition (in our case in a form of binary low–count images). Recognition algorithms usually require a high-level representation of objects. Such representation is carried out by some formalized description. Note that, depending on the degree of generality of descriptions, one (formalized) description can correspond to the whole family of (similar) subjects.

In our previous works [13] related to the processing of discrete space–distributed data that can be represented as a point cloud, we developed a new method of their economic description in a convenient for computations form using a Gaussian mixtures [14]. The basic idea of the method is that a set of cloud points is considered as a set of independently sampled, random coordinate vectors \( \{x_i\} \). These vectors are considered as identically distributed according to the weighted sum (mixture) of Gaussian densities:

\[
p(x_i | \theta) = \sum_{j=1}^{N} \pi_j \cdot \frac{\det B_j}{(2\pi)^{p/2}} \cdot \exp \left( -\frac{1}{2} Q_j (x_i) \right);
\]

where \( N \) is the number of mixture components, \( \theta \) is a set of parameters \( \{\pi_j, \mu_j, B_j\} \) of all \( N \) components, \( P \) is the dimension of coordinate space (in [13] \( P = 3 \)). As a result, it is possible to use \( N \cdot P \) parameters for the relatively rough description of the cloud instead of a list of coordinates \( \{x_i\} \) of all its points. In case of small \( N \) it is more practical. If the same idea is applied to the binary low–count images (\( \bar{P} = 2 \)), then each object will be associated with some Gaussian mixture whose parameters can be considered as a formalized object description (see Fig. 5).

Besides the fact that this description will be much more practical and convenient in terms of computation, it provides an opportunity to introduce a quantitative crite-
tion of similarity. That is, two objects are considered similar if Gaussian mixture (1) corresponding to image of one of them, describe well the image of other and vice versa. As a numerical parameter of the description quality, it is possible to take, for example, the value of likelihood ratio of both images (better its logarithm) with given Gaussian mixture of one of them. The level of this value can be compared to the likelihood ratio of that object image and, for example, an empty image of the same size – an image with a uniform probability distribution – considered as natural point of reference.

However, the most important argument in favor of the Gaussian mixture is a remarkable fact that there are effective algorithms to determine Gaussian mixtures that describe given randomly sampled data in the best way. These algorithms are known as the family of EM (Expectation-Maximization) algorithms [15]. A common feature of this family of algorithms is iterative (recurrent) nature of computations. At consequent iterations, the EM algorithm improves the estimation of parameters $\theta$ for distribution describing the selected data in the best way. The value of the likelihood function is chosen as a criterion of quality in EM (EM implements maximum likelihood method). Mixtures schematically shown in Fig. 5 and describing the consequent fragments in Fig. 4 were obtained exactly using the classical EM algorithm with $N=4$.

3. Recognition algorithm

Successfully solving the problem of describing classes of a certain domain DB (solving the problem of system learning), the proposed method, in the foregoing form, is unfortunately not very suitable for recognition. It would seem that if we find an appropriate mixture (1) for the subject of recognition and test it using the above method for similarities with the available classes in a database generated in the course of learning, the class that is the most similar to the object will solve the problem. However, such "rough" recognition falls down even in case of small displacements of the object on the image plane or changes in size. It turns out that similarity by size and location (on the image) are a much more important factor than the similarity of shape and other more subtle details.

For this reason, we expanded the concept of object classes by adding the ability of scaling the size and random displacement in a plane to each object. Conventionally speaking, if an object of a certain class is represented by a cloud of points with coordinates $\{\bar{x}_i\}$, the object with the coordinates $\{\bar{y}_j\}$, $\bar{y}_j = k\bar{x}_i + \bar{m}$, obtained from the initial object after extension of the image plane with coefficient $k$ and subsequent shift by vector $\bar{m}$ is also considered belonging to this class.

It is easy to deduce that if the mixture with the parameters $\{\pi_j, \mu_j, B_j\}$ corresponds to the initial object, then after described $k-m$ conversion mixture will have parameters $\{\pi_j, k\tilde{\mu}_j + m, B_j/k^2\}$. Accordingly, the test for the class formally described by a set $\{\pi_j, \mu_j, B_j\}$ will include predetermined of $k$ and $\bar{m}$ maximizing the likelihood function (better logarithm) of the mixture:

$$
\rho(\bar{x}_i | 0, k, \bar{m}) = \sum_{j=1}^{N} \pi_j \frac{\det B_j}{(2\pi)^k} \exp \left(-\frac{1}{2} Q_j(\bar{x}_i) \right),
$$

and subsequent analysis of the obtained values of the likelihood function with a purpose to determine the class of the tested object (maximum likelihood class).

3.1. Algorithm scheme

As mentioned above, our algorithm extends EM algorithm [15], so it has analogous scheme. It is also an iterative method for finding maximum likelihood (ML) estimates of parameters in a set of statistical models (from classification DB) like (2). For each model our algorithm iteration alternates between performing an expectation (E) step, which creates (implicitly) a function for the expectation of the log-likelihood evaluated using the current estimate for the weights $\{\pi_j\}$ and $k$, $\bar{m}$ parameters and maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E step. These parameter-estimates are used to determine the distribution of the latent variables in the next E step.

More specifically, if the object given by a set of coordinate vectors $\{\bar{x}_i\}$ is tested with respect to the definite class having formalized (model) description $\{\pi_j, \mu_j, B_j\}$, then the successive steps of the algorithm are as follows:

![Fig. 6. The results of the recognition of a new object of a 'handgun' type using the database shown in Fig. 5](image)
**Initialization:** set \( \{ \pi_j \} \), \( k \), \( \hat{m} \) to some initial estimates, for example exactly equal to class parameters:
\[
\pi_j^{(0)} = \pi_j, \quad \hat{m}_j^{(0)} = \hat{m}, \quad k^{(0)} = 1, \quad j = 1, \ldots, N. \tag{3}
\]
Further calculations proceed iteratively with the iterations counter \( n \).

**Step E:** with the values of parameters have been found at the previous iteration \( n \), calculate discrete conditional distribution of the latent variables (component indicators) for each \( x_i \):
\[
\Lambda_{ij} = \exp \left( - \frac{\left( \bar{x}_i - \bar{M}_j^{(n)} \right)^T B_j (\bar{x}_i - \bar{M}_j^{(n)})}{2k^{(n)}} \right).
\tag{4}
\]
\[
\pi_{ij} = \frac{\pi_j^{(n)} \sqrt{\det B_j \Lambda_{ij}}}{\sum_{i=1}^{N} \pi_j^{(n)} \sqrt{\det B_j \Lambda_{ij}}}
\]
\[
\bar{m}_j = \frac{\sum_{i=1}^{M} \pi_{ij} x_i}{\sum_{i=1}^{M} \pi_{ij}}, \quad j = 1, \ldots, N.
\tag{5}
\]

**Step M:** based on found distribution \( \{ \pi_j \} \) (4) recalculate weights \( \{ \pi_j \} \) and find \( \{ \hat{m}_j, A_j \} \) – analogues of EM estimates:
\[
\pi_j^{(n+1)} = \frac{1}{M} \sum_{i=1}^{M} \pi_{ij},
\]
\[
\hat{m}_j = \frac{1}{M \pi_j^{(n+1)}} \sum_{i=1}^{M} \pi_{ij} x_i
\]
\[
A_j = \frac{1}{M \pi_j^{(n+1)}} \sum_{i=1}^{M} \pi_{ij} (\bar{x}_i - \hat{m}_j) (\bar{x}_i - \hat{m}_j)^T.
\]

where \( M \) denotes a number of coordinate vectors \( \{ x_i \} \). Using (5) calculate auxiliary vectors \( \bar{v} \) and \( \bar{w} \):
\[
\bar{v} = \left[ \sum_{j=1}^{N} \pi_j^{(n+1)} B_j \right]^{-1} \sum_{j=1}^{N} \pi_j^{(n+1)} B_j \hat{m}_j,
\]
\[
\bar{w} = \left[ \sum_{j=1}^{N} \pi_j^{(n+1)} B_j \right]^{-1} \sum_{j=1}^{N} \pi_j^{(n+1)} B_j \bar{m}
\]

using these vectors \( \bar{v} \) and \( \bar{w} \) and estimates (5) find coefficients \( \alpha \) and \( \beta \) of the quadratic equation for \( k \):
\[
\gamma = \frac{1}{2} \sum_{j=1}^{N} \pi_j^{(n+1)} \text{Sp} \left\{ A_j B_j \right\},
\]
\[
\alpha = \frac{1}{2} \sum_{j=1}^{N} \pi_j^{(n+1)} (\hat{m}_j - \bar{v})^T B_j (\hat{m}_j - \bar{v}) + \gamma,
\]
\[
\beta = \frac{1}{2} \sum_{j=1}^{N} \pi_j^{(n+1)} (\hat{m}_j - \bar{w})^T B_j (\hat{m}_j - \bar{w}),
\]
\[
(k^{(n+1)})^2 + \beta k^{(n+1)} - \alpha = 0,
\]

that finally allows us to find all the values of key parameters and \( k, \hat{m} \):

After (8) \( n \) incremented, next iteration begins.

### 3.2 Computational experiments

To estimate characteristics of proposed algorithm a number of computational experiments was performed. Simulation of THz low–count images in accordance with their described in the introduction properties was carried out in two stages (see Fig. 7). At first, to reduce the contrast of the image and to perform its fuzzyfication (blurring), the Gaussian smoothing was performed. After that, a random sample of counts for a given sample size \( M \) was generated according to the intensity of smoothed image. Two stages \( a–b \) and \( b–c \) of this simulation procedure are shown in Fig. 7 (a – a visible source image of 'handgun'; \( b \) – low contrast, fuzzy shaped image – Gaussian smoothing of source image \( a \); \( c \) – 300 count size Poisson sample of smoothing image \( b \)).

![Simulation of THz low–count images in accordance with their described in the introduction properties](image)

**Fig. 7. Modeling low–count images such as THz scans**

To generate a random sample of counts in stage \( b–c \), Poisson sampler was used. The main reason for this is the simplicity of Poisson sampler implementation: all counts are generated independently and each count is formed in accordance with the probability distribution equal to the normalized intensity over the image. Tree typical results generated for different values of the sample size \( M \) are shown in Fig 8 (\( a \) – 300 counts; \( b \) – 1000 counts; \( c \) – 3000 counts).

![Different sample size low–count images of 'handgun' from Fig. 7](image)

**Fig. 8. Different sample size low–count images of 'handgun' from Fig. 7**

Algorithm under discussion was tested on image models like presented in Fig. 8. Preformed classification DB contained four–component Gaussian mixture parameters \( \{ \pi_j, \hat{m}, A_j \} \) \( j = 1, \ldots, 4 \) for three classes of objects (images \( \sim 1000 \) counts) represented in Fig. 5: \( a \) – "knife", \( b \) – "rectangular", \( c \) – "handgun". These parameter sets were found as a classical EM algorithm (100 iterations) output.

In these computational experiments for each of three different sample size low–count image of 'handgun'
Nevertheless, we have implemented this algorithm in certain was somewhere designed and investigated. Because of the evidence of this idea, such algorithm, except may be some no principal details, seems signed. Because of the evidence of this idea, such algorithm, except may be some no principal details, seems signed.

The above discussion is a particular manifestation of the object ‘knife’ is 60% of the likelihood for class ‘gun’ gives much bigger likelihood values and its similarity coefficients are closer to reality than for the alternative algorithm.

A careful analysis of these facts shows that in the case of wrong classes both algorithms behave equally badly - trying to distribute all the image counts in the only component (cluster) assigning zero weights to other components. Therefore, in both wrong cases, both algorithms give the same results. In the case of a class corresponding to the object tested, the algorithm based on Mahalanobis metric clustering, is still trying to assign all the counts to one cluster, while the proposed algorithm intelligently distributes them over suitable components as shown in Fig. 9 (a – algorithm based on Mahalanobis metric clustering (‘hard’ clustering) and b – proposed algorithm (‘soft’ clustering)). So our algorithm gives the best performance. The above discussion is a particular manifestation for a particular problem of a deeper general ideas [16] concerning the advantages of the ‘soft’ clustering methods over the ‘hard’ ones (randomized clustering procedures over deterministic clustering).

Table 1. Recognition characteristics of algorithm

<table>
<thead>
<tr>
<th>Image sample size</th>
<th>DB class</th>
<th>similarity</th>
<th>likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>M = 300</td>
<td>a ('knife')</td>
<td>10.88</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>6.18</td>
<td>141</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>4.26</td>
<td>324</td>
</tr>
<tr>
<td>M = 1000</td>
<td>a ('knife')</td>
<td>10.70</td>
<td>340</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>5.62</td>
<td>575</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>3.82</td>
<td>1239</td>
</tr>
<tr>
<td>M = 3000</td>
<td>a ('knife')</td>
<td>10.44</td>
<td>901</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>5.74</td>
<td>1357</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>3.94</td>
<td>3265</td>
</tr>
</tbody>
</table>

In Table 1 the characteristic named ‘similarity’ is the value of coefficient $k^{(E)}(8)$. The characteristic ‘likelihood’ actually means the logarithm of likelihood function ratio where nominator is probability of recognized object given the Gaussian mixture density $p(x_i | \theta, \Lambda)$ (2) and denominator is a uniform probability distribution $p_0(x_i) = \text{const}$ (empty image class).

Table 1 shows that for each of three different sample size low–count images the likelihood for class c ‘gun’ is maximal. It is more than twice more than likelihood of class b ‘rect’, and about four times bigger the likelihood of class a ‘knife’. This means that the algorithm is quite reliably detects the presented object like a gun. Note that with increasing sample size, the likelihood grows almost proportionally. This allows formulating the almost obvious conclusion: to improve the quality of algorithm in the noisy environment necessary to increase the number of image counts. In this regard, it is worth to note that the similarity parameter – estimation of the coefficient $k$ with a change in the number of counts varies relatively slightly.

For a deeper understanding of the algorithm proposed it was compared with alternative algorithm which use the image counts clustering based on Mahalanobis metric, similar to k-means method. As it is well known, and it is evident from the Algorithm scheme above, the conditional distribution $p_0$ calculated on step E assumes some randomized procedure for image counts distributing across a Gaussian components and evaluation on this base of desired parameters on step M. If instead of the randomized procedure for the distribution of counts would be used a deterministic one, for example based on the criterion of minimum Mahalanobis distance of counts from the center of component, the alternative algorithm would be designed. Because of the evidence of this idea, such algorithm, except may be some no principal details, seems certainly was somewhere designed and investigated. Nevertheless, we have implemented this algorithm in software and a series of numerical experiments similar to that described above was carried out for it also. The results of computational experiments (500 iterations) are shown in Table 2:

Table 2. Recognition characteristics of alternative

<table>
<thead>
<tr>
<th>Image sample size</th>
<th>DB class</th>
<th>similarity</th>
<th>likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>M = 300</td>
<td>a ('knife')</td>
<td>10.88</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>6.20</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>6.92</td>
<td>268</td>
</tr>
<tr>
<td>M = 1000</td>
<td>a ('knife')</td>
<td>10.70</td>
<td>340</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>5.61</td>
<td>574</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>6.29</td>
<td>962</td>
</tr>
<tr>
<td>M = 3000</td>
<td>a ('knife')</td>
<td>10.44</td>
<td>901</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>5.74</td>
<td>1357</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>6.37</td>
<td>2504</td>
</tr>
</tbody>
</table>

Comparing Tables 1 and 2 it is interesting to note that in both series of numerical experiments the results of testing object ‘handgun’ by wrong classes ‘knife’ and ‘rectangle’ are almost the same. On the contrary, the proposed algorithm for the correct class ‘gun’ gives much bigger likelihood values and its similarity coefficients are closer to reality than for the alternative algorithm.

In Table 1 the characteristic named ‘similarity’ is the value of coefficient $k^{(E)}(8)$. The characteristic ‘likelihood’ actually means the logarithm of likelihood function ratio where nominator is probability of recognized object given the Gaussian mixture density $p(x_i | \theta, \Lambda)$ (2) and denominator is a uniform probability distribution $p_0(x_i) = \text{const}$ (empty image class).

Table 1 shows that for each of three different sample size low–count images the likelihood for class c ‘gun’ is maximal. It is more than twice more than likelihood of class b ‘rect’, and about four times bigger the likelihood of class a ‘knife’. This means that the algorithm is quite reliably detects the presented object like a gun. Note that with increasing sample size, the likelihood grows almost proportionally. This allows formulating the almost obvious conclusion: to improve the quality of algorithm in the noisy environment necessary to increase the number of image counts. In this regard, it is worth to note that the similarity parameter – estimation of the coefficient $k$ with a change in the number of counts varies relatively slightly.

For a deeper understanding of the algorithm proposed it was compared with alternative algorithm which use the image counts clustering based on Mahalanobis metric, similar to k-means method. As it is well known, and it is evident from the Algorithm scheme above, the conditional distribution $p_0$ calculated on step E assumes some randomized procedure for image counts distributing across a Gaussian components and evaluation on this base of desired parameters on step M. If instead of the randomized procedure for the distribution of counts would be used a deterministic one, for example based on the criterion of minimum Mahalanobis distance of counts from the center of component, the alternative algorithm would be designed. Because of the evidence of this idea, such algorithm, except may be some no principal details, seems certainly was somewhere designed and investigated. Nevertheless, we have implemented this algorithm in software and a series of numerical experiments similar to that described above was carried out for it also. The results of computational experiments (500 iterations) are shown in Table 2:

Table 2. Recognition characteristics of alternative

<table>
<thead>
<tr>
<th>Image sample size</th>
<th>DB class</th>
<th>similarity</th>
<th>likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>M = 300</td>
<td>a ('knife')</td>
<td>10.88</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>6.20</td>
<td>142</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>6.92</td>
<td>268</td>
</tr>
<tr>
<td>M = 1000</td>
<td>a ('knife')</td>
<td>10.70</td>
<td>340</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>5.61</td>
<td>574</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>6.29</td>
<td>962</td>
</tr>
<tr>
<td>M = 3000</td>
<td>a ('knife')</td>
<td>10.44</td>
<td>901</td>
</tr>
<tr>
<td></td>
<td>b ('rect')</td>
<td>5.74</td>
<td>1357</td>
</tr>
<tr>
<td></td>
<td>c ('gun')</td>
<td>6.37</td>
<td>2504</td>
</tr>
</tbody>
</table>
Conclusions

The computational results obtained by algorithm proposed demonstrate that our approach is a good basis for the ATR system development, including the systems for security screening to detect the presence of a variety of threats, such as weapons or explosives, or illicit items, ranging from drugs to illegal immigrants [17]. The proposed approach for algorithm synthesis is clear in theoretical concepts and computationally efficient. For this reasons, we hope that the proposed methodology will be further developed and will be used for the relevant applications.

References


Author’s information

Viacheslav Evgenievich Antsiperov (b. 1959) graduated from Moscow Physical-Technical Institute in 1982, majoring in «Automation and Electronics». In 1987 defended PhD thesis on the specialty «Radio physics, including quantum radio physics». Currently he works as the leading researcher at Kotelnikov Institute of Radio-engineering and Electronics of Russian Academy of Sciences. Research interests are in digital signal and image processing, biomedical data processing, information systems design, computer aided systems, mobile systems and sensor nets. E-mail: antsiperov@cplire.ru .


Received May 14, 2016. The final version – October 16, 2016.