

UDC 629.7.014.066-519:004.93(045)

Philip Prystavka¹
Anastasia Rogatyuk²**THE MATHEMATICAL FOUNDATIONS OF FOREIGN OBJECT RECOGNITION IN THE VIDEO FROM UNMANNED AIRCRAFT**^{1,2}National Aviation University

1, Kosmonavta Komarova ave., Kyiv, 03680, Ukraine

E-mails: ¹chindakor@mail.ru, ²arogatyuk@gmail.com

Abstract. The paper considers an analysis of object and phenomena recognition methods. We focus at the main problems of objects recognition in real-time in the video. As part of identification procedure paper considers separating mixtures of normal distributions iterative method which based on histogram estimation method.

Keywords: objects identification; recognition methods; separating mixtures; video processing

1. Introduction

Object recognition is of great significance for automation process of human activity related to environmental object identification. For automation recognition process needs to develop specific systems, for example, there is a topical problem of object recognition in the video from unmanned aircraft. In general, process of creating a recognition system could be divided into the following main steps [1]:

1. Create a complete list of features that characterize objects or phenomena, for which a recognition system was made. Feature of objects can be deterministic, probabilistic, logical and structural. Structural features are some elements (characters) of object structure.

2. Make the initial classification of objects or phenomena which need to recognize. Realization of initial objects separation process into classes based on priori data, in other words compose priori alphabet of classes.

3. Define priori dictionary of feature (which describes each class).

4. Divide priori feature space into regions which correspond to the classes of priori alphabet.

5. Choose such recognition algorithm that will include object or phenomenon to a particular class or a certain sequence.

6. Determinant the working classes alphabet and features dictionary of recognition system.

7. Develop the algorithms for control the system.

8. Choose the system performance evaluation and estimation of their value.

If to look more specifically the problem of object recognition for photo and video frames, it is clear that the feature will have the probability character, since the same moving object at different times have little variation characteristics as size and color com-

ponent. It is therefore advisable to consider it further statistical recognition methods.

2. Materials and methods

In [2] are the main stages of statistical recognition: the formation of feature space, obtaining standard classes (classes priori alphabet) and building regulations decision on the observed object class.

Consider the classical methods of statistical recognition. Let priori dictionary of feature is an ordered set of parameters of objects x_1, \dots, x_N . Then when features are probability, the descriptions of classes is conditional probability distribution density values of variables x_1, \dots, x_N for each class $\Omega_1, \dots, \Omega_M$, that is mean function $f(x_1, \dots, x_N)$. In case after some observation and analysis of sample sets of values can set the density function, then the aim is to obtain estimates of distribution parameters, in this case using parametric methods of recognition, otherwise when the form is unknown density function are used nonparametric methods of recognition. Classical parametric methods of recognition are discussed in [2] the method maximum likelihood estimation and Bayesian criterion.

The simplest nonparametric is histogram estimation method that based on an assessment of medium density distribution in the region. The advantage of this method is the ease of constructing histograms. The downside is that with a large number of considered attributes of the objects, assessment density function as a histogram can be considered satisfactory only if a sufficiently large sample [3].

Parzen method relates to methods of local estimation and is builds function for each point of the class, which reaches a maximum at this point and decreases rapidly with distance from it in a certain

neighborhood. The disadvantage of this method is the high sensitivity of the assessment to the choice of the initial range and that the amount of its surroundings, satisfactory for one point can be completely unsatisfactory for another point.

When using the nearest neighbor rule with a specific set of data resulting classification error rate will depend on the characteristics of random sampling. However, in an infinite sample size standard errors for the nearest neighbor rule will never be worse Bayesian more than doubled.

Speaking about the specific task of object recognition in video in [4] considered the methods of image segmentation into homogeneous areas. In general, for video processing such approaches are very costly in terms of computational complexity. Appears issue of the task that takes into account the performance calculation.

3. Purpose of the paper

The purpose of this paper is to review such method of object recognition on the video that would gives effective results with low computational complexity and considering objects and phenomena recognition classic approaches.

4. Discussion

Let there is a digital video that is presented as sequence of frames $V = \{V_k; k = \overline{1, K}\}$, where V_k is video frame; K is the number of frames. Each V_k frame is a two-dimensional array of pixels $V_k = \{v_{k,n,m}; n = \overline{1, N}, m = \overline{1, M}\}$, $k = \overline{1, K}$, where N and M are the size of frame and each pixel $v_{k,i,j}$ is represented as values of three raster components in the RGB color model, where:

$$v_{k,n,m} = R_{k,n,m} + 256 \cdot G_{k,n,m} + 65536 \cdot B_{k,n,m}$$

Any object or area (texture) on video, can be represented as a one-dimensional array of any color components. Let Δ is the texture on the frame with N_p on M_p pixels size and p is any of the elements of color raster (R, G or B). So we have an array of intensities of color component

$$P = \{p_{n,m}; n = \overline{1, N_p}, m = \overline{1, M_p}\}, p_{n,m} \in [0; 255]$$

or the same as a one-dimensional array

$$P = \{p_l; l = \overline{1, L}\}, p_l \in [0; 255],$$

where $L = N_p \cdot M_p$ is the number of all pixels of texture Δ . Thereafter texture will be understood as a specific visible region on the frame and as digital representation in P form.

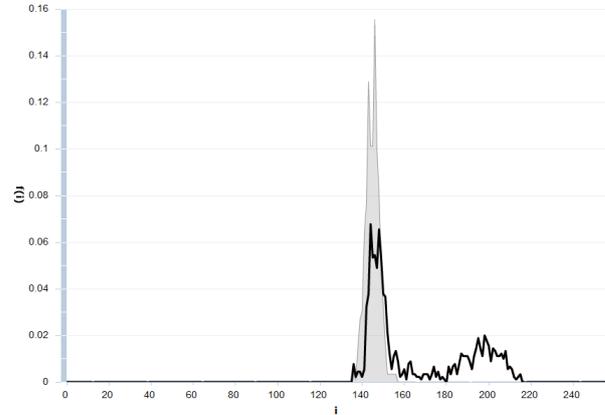


Fig. 1. Texture histogram:

a) texture «grass» (thin line); b) texture «grass and sand» (thick line).

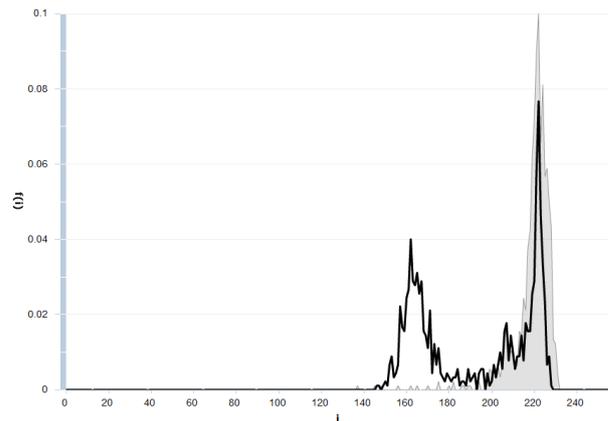


Fig. 2. Texture histogram:

a) texture «sand» (thin line); b) texture «sand ra grass» (thick line).

To make a conclusion about some texture that presented as P array need calculate histogram estimation. Individual dataset is homogeneous or not can define by analyzed histogram view. In particular, if the histogram has only one mode then texture P is homogeneous, else if has more than one mode (multimode) then P texture is heterogeneous. Figure 1 shows a unimodal and multimodal histograms: first histogram describes fragment frame which shows grass (texture "grass"), the second histogram describes mixed textures "grass and sand." Similarly, can see (Figure 2) histograms of "sand" and "sand and grass" textures.

Homogeneous areas which oft-times show up on V_k video are typical textures and the set of them is an array of standard textures

$$E = \{E_q; q = \overline{1, Q}\},$$

if standard texture determined uniquely, or

$$E = \{e_{q,w}; q = \overline{1, Q}, w = \overline{1, W_q}\},$$

if the standard texture can include any of the W_q similar textures among Q types.

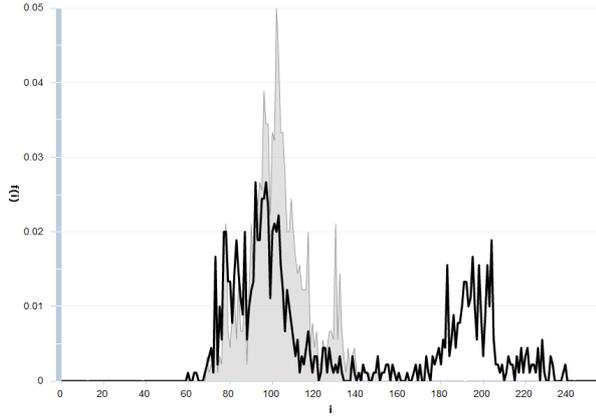


Fig. 3. Texture histogram:

a) texture «forest» (thin line); b) texture «the house in the woods» (thick line).

Any homogeneous area on video that statistically not equivalent any E_q texture from array of standard texture can either be added to the E set or considered as texture with foreign objects and belong to some set of objects P^* that are searching on V .

Non-uniform textures on video frames can be characterized by the presence of homogeneous components that can be clearly separated on the image, like on texture "Grass and Sand" (Figure 1, 2), with all corresponding histograms fragments that equated simple components and spaced on x-axis of intensity. These textures can be identified based on analysis by type of homogeneous components, so will assume that such area belong to some set \tilde{P} . Thus, the set of texture \tilde{P} components are elements which individually owned either E , P^* , E or P^* (Figure 3).

Let raises the problem of identification the presence of a foreign object in a certain texture P . Then its solution consists of three stages:

1. calculate histogram estimation for P ;

2. test if belong P to some of the sets E , P^* or \tilde{P} ;
3. if $P \subset \tilde{P}$, get to know if P contains an element of the set P^* or composed entirely of textures belonging to E .

Consider the widely formulated problem. The initial stage of processing is the histogram estimation. A histogram is a graphical representation of the frequency of values depending on the brightness of pixels, the analyzed interval of grouping represented 256 brightness values of the frame. Practical applications have two types of histograms, frequency and relative frequencies.

Frequency histogram calculation in this way:

$$n(i), \quad i = \overline{0, 255},$$

where $n(i)$ the number of pixels, with quantum amplitude intensity level i within the aperture of the analyzed image.

Histogram relative frequencies allows us to estimate the empirical probability of a pixel that has the i -th level of intensity, that is mean:

$$f(i) = \frac{n(i)}{L}, \quad i = \overline{0, 255}. \quad (1)$$

Histogram relative frequencies provides a valid comparison textures of different sizes, so subsequent presentation will bear in mind this type of histogram.

Histogram gives an idea of the shape density function for the intensity of the color components $f(i)$. In general, the intensity distribution for color components homogeneous texture can put a normal distribution with density function

$$f(p) = \frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{(p-\bar{p})^2}{2\sigma^2}},$$

where

$$\bar{p} = \sum_{i=0}^{255} i \cdot f(i);$$

$$\sigma = \sqrt{\frac{L}{L-1} \sum_{i=0}^{255} (i-\bar{p})^2 \cdot f(i)},$$

is evaluation of the expectation and variance P .

Introduction normal law to describe the intensity distribution color components "simple" texture allows naturally parameterize the dataset E , that is

mean characterize each standard texture deuce ratings

$$E_q : (\bar{p}_q, \sigma_q), \quad q = \overline{1, Q}.$$

The intensity of color components of heterogeneous texture mixture model can be described by normal distributions

$$f(p) = \sum_{j=1}^r \alpha_j f_j(p), \quad j = \overline{1, r}, \quad (2)$$

where

$$\sum_{j=1}^r \alpha_j = 1;$$

$$f_j(p) = \frac{1}{\sqrt{2\pi}\sigma_j} e^{-\frac{(p-\bar{p}_j)^2}{2\sigma_j^2}}.$$

Note that in the trivial case of heterogeneous intensity texture (Figure 1-3) model is a mixture of two normal distributions.

In the second stage the problem of determining the presence of a foreign object in the texture P comes to testing statistical hypotheses of belonging to a particular set of: E , P^* or \tilde{P} .

Main hypothesis about homogeneity of P is:

$$H_0 : P \subset E \cup P^*,$$

with alternative

$$H_1 : P \not\subset E \cup P^*.$$

For this hypothesis error of the first kind is that we take the texture for the "simple" when in reality it is not so. The error of the second kind believe that texture is homogeneous, when in reality it is heterogeneous. The value of the probability of error of the first kind depends on the amount L an P array, but not more than the value of probability 0,05.

For statistical characteristics hypothesis can choose value σ . Then for each color component raster texture P do checking

$$\sigma \leq C_1,$$

where C_1 level of variability that characterizes homogeneous texture and the main hypothesis is correct when inequality holds for a certain value, for example $C_1 = 7$.

The question of belonging of a simple texture P specifically to E or P^* can be solved empirically, for example, based on frequency of occurrence of two (\bar{p}, σ) when analyzing the texture of video frames V .

Proximity options (\bar{p}, σ) to a certain standard (\bar{p}_q, σ_q) , $q = \overline{1, Q}$ is easy to verify by f -test hypothesis

$$H_0 : \sigma_q = \sigma, \text{ with alternative } H_1 : \sigma_q \neq \sigma$$

and t -test hypothesis

$$H_0 : \bar{p}_q = \bar{p}, \text{ with alternative } H_1 : \bar{p}_q \neq \bar{p}.$$

F -test carry after calculating statistics

$$f = \begin{cases} \frac{\sigma}{\sigma_q}, & \sigma > \sigma_q, \\ \frac{\sigma_q}{\sigma}, & \sigma_q > \sigma \end{cases}$$

and checking the condition

$$f \leq f_{\alpha, \nu_1, \nu_2},$$

where α is probability of error of the first kind in making the hypothesis of statistical variability same two textures compare; $\nu_1 = \nu_2 = L - 1$ the number of degrees of freedom.

The basis of the t-test laid checking the condition

$$|t| \leq t_{\alpha, \nu},$$

where

$$t = \frac{\bar{p} - \bar{p}_q}{\sqrt{\sigma^2 + \sigma_q^2}} \sqrt{L},$$

$$\nu = L - 2.$$

If inequality is satisfied, then finally conclude on statistical equality of control and reference textures. Conversely if the t-test is not passed for $\forall q = \overline{1, Q}$, is statistically significant at the probability α texture P does not get to the set standard and can be considered as being belongs to P^* .

If does not start from the normal distribution model of homogeneous texture intensities, the question of comparing two histograms can be solved based on the distance metric in Euclidean space. If

$f(i)$ and $f_q(i)$, $i = \overline{0, 255}$, are histogram textures P and E_q , then distance d between them for each color component is defined as:

$$d = \sum_{i=0}^{255} \sqrt{(f(i) - f_q(i))^2}. \quad (3)$$

Then value d compared with some threshold D and if the inequality

$$d \leq D,$$

then accepted assumptions about the proximity of the two histograms, thus and of belonging P to sequence E . Note that could set $D = 0,6$.

The questions of statistical equality histograms $f(i)$ and $f_q(i)$, $i = \overline{0, 255}$ solved by testing static hypotheses

$$H_0 : f(i) = f_q(i), i = \overline{0, 255}, \text{ with alternative}$$

$$H_1 : f(i) \neq f_q(i), i = \overline{0, 255},$$

which can be made based on any criterion of agreement, such as Pearson. To implement Pearson criterion calculate statistical characteristics

$$\chi^2 = \sum_{i=0}^{255} \frac{(n(i) - n_q(i))^2}{n_q(i)}.$$

The main hypothesis is accepted if the inequality

$$\chi^2 \leq \chi^2_{\alpha, v},$$

where $\chi^2_{\alpha, v}$ is quantile distribution Pearson, α is probability of error of the first kind for the hypothesis H_0 , $v = 255$ the number of degrees of freedom Pearson statistics.

Implementation Pearson criterion can have drawbacks associated with the number of degrees of freedom. With the increase of $v \approx 10$ number of degrees of freedom power criterion is reduced, so is recommended [5] define statistics χ^2 , for witch

$v \leq \frac{L}{5}$. Therefore, for effective application of Pearson criterion, there is a need to conduct histogram frequencies and relative frequencies for a smaller number of intervals grouping than 256.

To test accessories textures P to the set of heterogeneous textures, like \tilde{P} formulate main hypotheses

$$H_0 : P \subset \tilde{P}, \text{ with alternative}$$

$$H_1 : P \not\subset \tilde{P}.$$

For this hypothesis error of the first kind is this: when we consider that the texture P is not homogeneous components that can be isolated and identified when in reality these components are. The error of the second kind believe that a homogeneous texture components, which are separated by the horizontal axis, when in reality this is not observed.

Hypothesis testing H_0 carried out by comparing variability σ with some threshold C_2

$$\sigma \geq C_2,$$

and for certainty can be put $C_2 = 28$.

A further issue to finding a foreign object on texture, when $P \subset \tilde{P}$, is to implement procedures play a mixture of normal distributions, that is mean obtaining vector ratings

$$\hat{\Theta} = \{(\bar{p}_j, \sigma_j), \alpha_j; j = \overline{1, r}\}$$

with follow definition of homogeneous components of certain components to sequence E or P^* .

To implement in software it is important to choose the method that has high performance in use.

For example, consider the iterative method of separating a mixture of normal distributions, based on the evaluation of the resulting histogram of relative frequencies, and does not depend on the number r of mixed distributions in the mixture (2). The purpose of this method is to determine the location of modes in the histogram to be considered tops normal distributions, and the subsequent assignment of each element of the sample to the appropriate division forming mixture.

In the first stage of the method is carried histogram estimation sample by the formula (1) of subdivision into 256 classes. Based on the histogram formed the original array \tilde{I} location of local modes:

$$\tilde{I} = \{\tilde{i}_j; j \leq 128\},$$

where

$$\tilde{i}_j = i,$$

if

$$f(i) > f(i-1), f(i) > f(i+1).$$

The maximum number of r distributions in the mixture at this stage may be 128. The next step method removes vertices that are within a homogeneous sample. If for some $j \in [2, \max\{j\} - 1]$ will

$$f(\tilde{i}_j) > f(\tilde{i}_{j-1}), f(\tilde{i}_j) > f(\tilde{i}_{j+1})$$

or

$$f(\tilde{i}_j) > f(\tilde{i}_{j-1}), j = \max\{j\}, \\ f(\tilde{i}_j) > f(\tilde{i}_{j+1}), j = 1,$$

Then from array \tilde{I} remove element with such index \tilde{i}_{j+1} , for which will

$$\begin{cases} |\tilde{i}_j - \tilde{i}_{j-1}| < 2,5 \cdot C_1, \\ |\tilde{i}_j - \tilde{i}_{j+1}| < 2,5 \cdot C_1, \end{cases} \quad (5)$$

For an updated array $\tilde{I} = \{\tilde{i}_j\}$ this procedure is repeated for as long as the condition (5) will deviate for all local modes remaining. With the amount of distributions that form the mixture (2) choose value

$$r = \max\{j\}.$$

The last step of the method is to classify each element texture P the relevant components of the mixture. The condition

$$\min_j |p_l - \tilde{i}_j|, \quad i \in [0, 255], \quad l = \overline{1, L},$$

calculate index j distribution, which include p_l . The result is r texture components, which then assume homogeneous. Further, each of these components must be tested to its belonging to the sets E or P^* .

5. Conclusions

Considered and analyzed the classical methods of object recognition. Considered the problem formulation of object recognition in video that has effective results with low computational complexity. Invented new method of foreign object recognition in video in real time which using etalon textures. As part of identification procedure was considered and offered simple in implementation separating mixtures of normal distributions iterative method which based on histogram estimation method.

References

- [1] *Gorelik A.* Recognition methods: Textbooks for schools / Gorelik A., Skripnik V. – 3rd ed., Revised and appended. – Moscow: Higher School, 1989. – 232 p. (in Russian).
- [2] *Vasilyeva I.* Methods of Pattern Recognition: Tutorial Laboratory workshop. – Kharkiv.: National Aerospace University "Kharkiv Aviation Institute", 2008. – 57 p. (in Russian).
- [3] Statistical methods for pattern recognition [Online resource]. – URL:
- [4] <http://yury.name/modern/07modernnote.pdf> (in Russian).
- [5] *Gruzman I.* Digital image processing in information systems. – Novosibirsk.: Tutorial – 2002. – 352 p. (in Russian).
- [6] *Kobzar A.* Applied Mathematical Statistics. For engineers and scientists. – M.: FIZMATLIT, 2006. – 816 p. (in Russian).
- [7] *Prystavka A.* Mixtures and spline on heterogeneous distribution of normal spaces – D.: Dnepropetrovsk State. University, 1987. – 233 p. Dep. VINITI 11.01.88, №33–V88. (in Russian).
- [8] *Glumov H.* Image processing, pattern recognition. // Computer Optics – 2008 – part 32, № 3 – p. 277–282. (in Russian).

Received 17 December 2014.

Amended after peer review 15 June 2015.

П.А. Приставка¹, А.А. Рогатюк². Математичне забезпечення розпізнавання чужорідного об'єкта на відео з безпілотного повітряного судна

^{1,2}Національний авіаційний університет, проспект Космонавта Комарова 1, Київ, Україна, 03680

E-mails: ¹chindakor@mail.ru; ²arogatuk@gmail.com

Проаналізовано класичні методи розпізнавання об'єктів і явищ. Виділено основні завдання при створенні системи розпізнавання. Розглянуто параметричні і непараметричні методи, такі як метод максимальної правдоподібності, баєсівський критерій, гістограмний, метод Парзена, правило найближчого сусіда. На основі аналізу класичних методів виявлена необхідність і сформульована постановка задачі розпізнавання в режимі близького до реального часу об'єктів на відео. В рамках етапів ідентифікації розглянуто ітераційний метод розділення суміші нормального розподілу, оснований на гістограмній оцінці.

Ключові слова: ідентифікація об'єктів; метод розпізнавання; обробка відео; розділення суміші

Ф.А. Приставка¹, А.А. Рогатюк². Математическое обеспечение распознавания чужеродного объекта на видео с беспилотного летательного аппарата

Национальный авиационный университет, проспект Космонавта Комарова 1, Киев, Украина, 03680

E-mails: ¹chindakor@mail.ru; ²arogatuk@gmail.com

Проанализированы классические методы распознавания объектов и явлений. Выделены основные задачи при создании системы распознавания. Рассмотрены параметрические и непараметрические методы, такие как метод максимального правдоподобия, байесовский критерий, гистограммный, метод Парзена, правила ближайшего соседа. На основе анализа классических методов выделена необходимость и сформулирована постановка задачи распознавания в режиме близкого к реальному времени объектов на видео. В рамках этапов идентификации рассмотрен итерационный метод разделение смеси нормальных распределений, основанный на гистограммной оценке.

Ключевые слова: идентификация объектов; методы распознавания; обработка видео; разделение смеси

Prystavka Philip (1974). Doctor of Technical Sciences. Professor.

National Aviation University, Kyiv, Ukraine.

Education: Dnipropetrovsk National University, Dnipropetrovsk, Ukraine (1996).

Research area: applied mathematics and information technologies.

Publications: 130.

E-mail: chindakor@mail.ru

Rogatuk Anastasia (1989). Postgraduate Student.

National Aviation University, Kyiv, Ukraine.

Education: National Aviation University, Kyiv, Ukraine (2011).

Research area: information technology.

Publications: 3.

E-mail: arogatuk@gmail.com