

DEVELOPMENT OF THE EFFECTIVE SET OF FEATURES CONSTRUCTION TECHNOLOGY FOR TEXTURE IMAGE CLASSES DISCRIMINATION

E. Biryukova¹, R. Paringer^{1,2}, A.V. Kupriyanov^{1,2}

¹Samara National Research University, Samara, Russia

²Image Processing Systems Institute - Branch of the Federal Scientific Research Centre "Crystallography and Photonics" of Russian Academy of Sciences, Samara, Russia

Abstract. The paper proposes a method for linear transformation of the features space, based on the use of discriminant analysis algorithm. For the test set of images the use of the obtained features space instead of the initial one has reduced the clustering error from 30% to 10%.

Keywords: texture images, discriminant analysis, statistical features.

Citation: Biryukova EV, Paringer RA, Kupriyanov AV. Development of the effective set of features construction technology for texture image classes discrimination. CEUR Workshop Proceedings, 2016; 1638: 263-269. DOI: 10.18287/1613-0073-2016-1638-263-269

Introduction

Texture is one of the important characteristics in the field of image analysis. However, there is still no clear definition of the term, and for each specific task of texture analysis a different method of definition is applied.

Currently, the problem of texture image discerning is quite common, but it is often addressed not effectively enough. This problem cannot be solved due to the use of the inappropriate features.

In this paper, we propose a technology which constructs a set of effective features, relevant for the classification of a large set of texture images, and thus reduces the error in texture recognition.

To achieve the goal, the following tasks were formulated:

- To develop a method for evaluating the effectiveness of a feature for discerning classes of texture images;
- To apply discriminant analysis algorithm to improve the efficiency of a features set;
- To develop a method for constructing of a set of informative features;

- To develop a technology of building a effective features set;
- To analyse the technology on a set of texture images.

Description of the technology

The proposed technology is based on the use of discriminant analysis algorithm for the formation of an effective set of features. It applies the first separability criterion to estimate the informativeness of the features [1].

Discriminant analysis techniques proved their efficiency in many areas, for example in the medical diagnostics [2-6].

The discriminant analysis is used to eliminate the correlation between the features and, as a consequence, reduce the dimension of the attribute space. The implementation of this algorithm allows us to maintain the informativity value of the set of features for the classification, as well as to reduce the number of features, so that simpler methods of classification can be applied, and the value of the classification error is lower.

Suppose there is a sample consisting of n elements, divided into g classes, and containing p signs. In the discriminant analysis, measure of the effectiveness of the sample is criteria of separability, which are calculated according to the formulas:

$$J = \text{tr}((\mathbf{T})^{-1}\mathbf{B})$$

where $\mathbf{T} = \mathbf{B} + \mathbf{W}$,

\mathbf{B} is the intergroup dispersion matrix. The elements of this matrix are calculated according to the formula:

$$b_{ij} = \sum_{k=1}^g n_k (\bar{x}_{ik} - \bar{x}_i)(\bar{x}_{jk} - \bar{x}_j), \quad i, j = \overline{1, p}$$

\mathbf{W} is the intragroup dispersion matrix. The elements of the matrix are calculated according to the formula:

$$w_{ij} = \sum_{k=1}^g \sum_{m=1}^{n_k} (x_{ikm} - \bar{x}_{ik})(x_{jkm} - \bar{x}_{jk}), \quad i, j = \overline{1, p}$$

x_{ikm} is the value of the i -th feature for the m -th element of k class,

$\bar{x}_{ik} = 1/n_k \sum_{m=1}^{n_k} \bar{x}_{ikm}$ is the mean value of the i -th feature of k class,

$\bar{x}_i = 1/n \sum_{k=1}^g n_k \bar{x}_{ik}$ is the mean value of the i -th feature in all the classes, and n_k is the number of elements in k class.

The higher the value of the criterion is, the more the separability of the classes grows.

Suppose $\mathbf{x} = [x_1 x_2 \dots x_p]^T$ is the initial feature vector. Consider the algorithm for formation of new features $\mathbf{y} = [y_1 y_2 \dots y_m]^T$.

1. Calculate the values of eigenvectors \mathbf{v}_i , $i = \overline{1, p}$ for matrix $\mathbf{T}^{-1}\mathbf{B}$.
2. Calculate the standardized coefficient vectors $\boldsymbol{\beta}_i = [\beta_0, \beta_1 \dots \beta_p]$, $i = \overline{1, m}$, such that the elements β_i , $i = \overline{0, p}$ are calculated as following:

$$\beta_0 = - \sum_{i=1}^p \beta_i \bar{x}_i, \quad \beta_i = v_i \sqrt{n-g}, \quad i = \overline{1, p}$$

3. Calculate the elements of the new attribute vector by the formula:

$$y_i = \beta_0 + \beta_1 x_1 + \dots + \beta_p x_p, \quad i = \overline{1, m}.$$

The number of the formed features cannot exceed the number of the initial ones.

We can describe the main stages of the technology for the construction of the effective features set in the form of the following algorithm:

1. Select the initial data: a features set and texture images divided into classes.
2. Calculate the features for each image. Form the original sample.
3. Compose a set of informative initial features (Figure 1):
 - (a) select two classes from the original sample;
 - (b) calculate individual separability criteria for/of each feature for those classes;
 - (c) store the best feature depending on the value of the separability criterion;
 - (d) repeat steps (a) – (c) for all possible pairs of classes from the original sample;
 - (e) eliminate recurring features. The result is the desired set.
4. Eliminate from the initial set all spurious features that do not belong to the set obtained in step 3
5. Apply the discriminant analysis algorithm to generate the set of new features.
6. Calculate the individual separability criteria for each new feature of the generated set.
7. Eliminate the features with low separability criterion value from the set.

Thus, we get effective features set for discerning the classes of texture images.

Experimental results

The choice of the initial set of statistical features is not very important because, with the use of discriminant analysis, we obtain a linear combination of features that provides maximum separability of the classes. Therefore, in the current research we used statistical features to discern the classes of texture images. Textural features have shown good results in recognizing biomedical images and their further diagnosis [7-11]. Those features include the first moment, contrast, correlation, inertia [12] along with the features calculated by the formula

$$\lambda(\Delta x, \Delta y, n) = \frac{\sum |f(x, y) - f(x + \Delta x, y + \Delta y)|^n}{N}$$

where f is the image intensity function, N is the number of image pixels.

In the following research we used the features λ , calculated at $\Delta x, \Delta y = 0, \pm 1, \pm 2$, $n = 1, 2, 3$. The general set contained 40 features.

Texture images are from “Kylberg Texture Dataset v. 1.0” [13]. This dataset consists of 28 classes of various textures. Each class contains 160 unique images. All images are grayscale and have the resolution of 576 by 576 pixels. Examples of texture images of each class are shown in Figure 2.

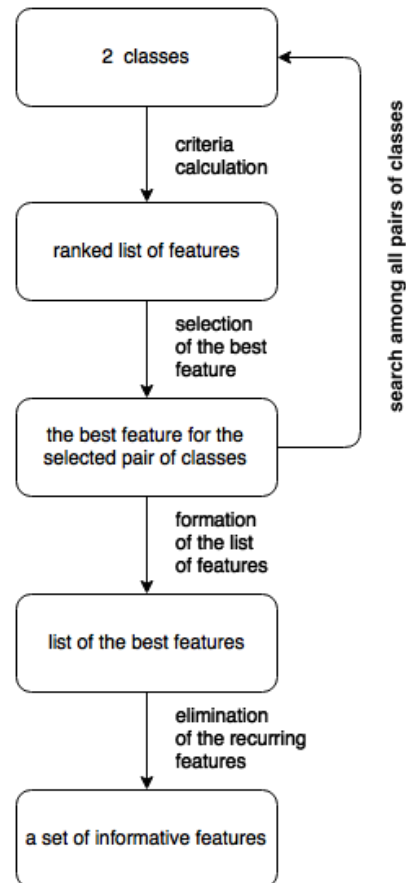


Fig. 1. Set of informative initial features

In order to calculate the features, the program generating the original sample was developed.

Individual separability criteria for each pair of classes from the original sample were calculated in accordance with Step 3 of the algorithm. Thus 378 features were selected. Then, after the elimination of the recurring features, we get the set of 28 features. After we had applied the discriminant analysis to the sample obtained at Step 4, a new features space was formed.

To complete the formation of the effective features set, we calculate the individual separability criteria from the obtained features space. After that we eliminate those features, the criterion value of which is less than 0.5.

Thus, we obtained a set of 12 effective features for texture classes discrimination for the images from “Kylberg Texture Dataset texture image v. 1.0”.

The chosen method for evaluating the effectiveness of the features proved to be effective in [14], where it was implemented on the same set of texture images.

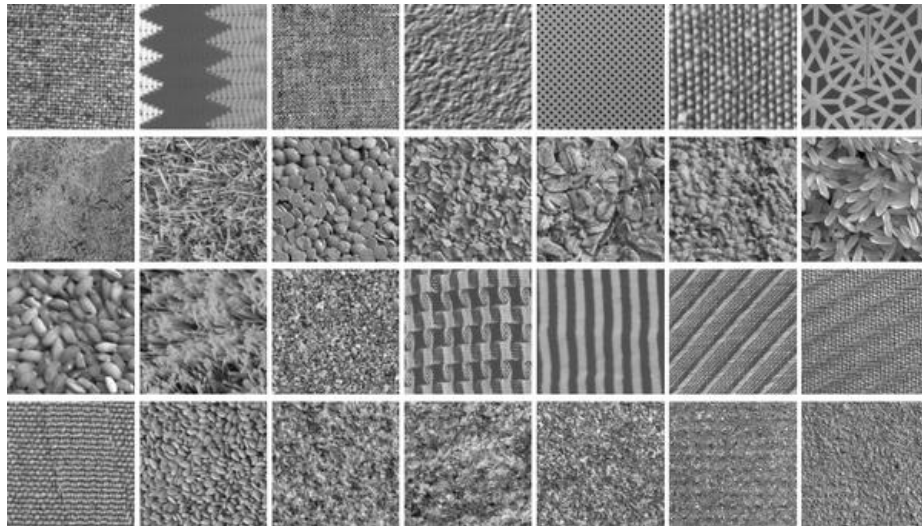


Fig. 2. Examples of texture images for each class of “Kylberg Texture Dataset”

To study the effectiveness of the proposed technology we carried out an evaluation test of the features sets, obtained at different steps of the algorithm. The evaluation was conducted by calculating the clustering error with the use of the k-means algorithm. The centres of the initial classes were used as the entry conditions [15]. The clustering error was calculated after Steps 4 and 7 and amounted to 30% and 10% respectively. This figures confirm the effectiveness of the developed technology.

Conclusion

Thus, the technology of constructing an effective features set has been developed and implemented. This technology consists of several steps, at which we develop a method for evaluating the features effectiveness, increase the effectiveness of the set using the discriminant analysis, generate the informative features set, and form new features from the obtained set by the linear combination of the initial features.

The obtained features set can be used to describe texture differences of a large number of textures.

The technology has been evaluated with the use of a dataset of texture images. The implementation of the technology reduces the error from 30 to 10%.

The proposed technology can be applied for texture image description and for general studies of textures.

Acknowledgements

This work was partially supported by the Ministry of education and science of the Russian Federation in the framework of the implementation of the Program of increasing the competitiveness of SSAU among the world's leading scientific and educational centers for 2013-2020 years; by the Russian Foundation for Basic Research grants (# 14-07-97040, # 15-29-03823, # 15-29-07077, # 16-57-48006); by the ONIT RAS program # 6 "Bioinformatics, modern information technologies and mathematical methods in medicine" 2016.

References

1. Fukunaga K. Introduction to statistical pattern recognition. San Diego: Academic Press, 1990; 592 p.
2. Ilyasova NYu, Kupriyanov AV, Paringer RA. The Discriminative analysis Application to Refine the Diagnostic Features of Blood Vessels Images. *Optical Memory & Neural Networks (Information Optics)*, 2015; 24(4): 309-313.
3. Ilyasova NYu, Kupriyanov AV, Paringer RA. Formation of features for improving the quality of medical diagnosis based on discriminant analysis method. *Computer Optics*, 2014; 38(4): 851-856. [in Russian]
4. Ilyasova NYu, Kupriyanov AV. The Big Data mining to improve medical diagnostics quality. *CEUR Workshop Proceedings*, 2015; 1490: 346-354. DOI: 10.18287/1613-0073-2015-1490-346-354.
5. Ilyasova NYu, Paringer RA. Research effectiveness of features for the vascular pathologies diagnosis. *Scientific Journal of "Proceedings of the Samara Scientific Center of the Russian Academy of Sciences"*, Samara Scientific Center of the Russian Academy of Sciences, 2015; 17; No 2(5): 1015-1020. [In Russian]
6. Gaidel AV, Zelter PM, Kapishnikov AV, Khramov AG. Computed tomography texture analysis possibilities of the chronic obstructive pulmonary disease diagnosis. *Computer Optics*, 2014; 38(4): 843-850. [In Russian]
7. HeiShun Yu, Tischler B, Qureshi MM, Soto JA, Anderson S, Daginawala N, Li B, Buch K. Using texture analyses of contrast enhanced CT to assess hepatic fibrosis. *European Journal of Radiology*, 2016; 85(3): 511-517.
8. Acharya UR, Ng EY, Tan JH, Sree SV, Ng KH. An integrated index for the identification of diabetic retinopathy stages using texture parameters. *Journal of Medical Systems*, 2012; 36(3): 2011-2020.
9. Kutikova VV, Gaidel AV. Study of informative feature selection approaches for the texture image recognition problem using Laws' masks. *Computer Optics*, 2015; 39(5): 744-750 [In Russian]. DOI: 10.18287/0134-2452-2015-39-5-744-750.
10. Gaidel AV. A method for adjusting directed texture features in biomedical image analysis problems masks. *Computer Optics*, 2015; 39(2): 287-293 [In Russian]. DOI: 10.18287/0134-2452-2015-39-2-287-293.
11. Gaidel AV, Pervushkin SS. Research of the textural features for the bony tissue diseases diagnostics using the roentgenograms. *Computer Optics*, 2013; 37(1): 113-119. [In Russian]
12. Haralick RM, Shanmugam K, Dinstein I. Textural features for image classification. *IEEE Trans. on Systems, Man and Cybernetics*, 1973; 3: 610-621.

13. Kylberg texture dataset. URL: <http://www.cb.uu.se/~gustaf/texture/>.
14. Biryukova EV, Paringer RA, Kupriyanov AV. Research efficiency for features istinguishing texture images. *Advanced Information Technologies and Scientific Computing (PIT 2016)*, Proceedings of the International Scientific Conference, Samara: Samara Scientific Center of RAS, 2016; 1052: 45-49. [in Russian]
15. Mandel ID. *Cluster analysis*. Moscow: Finance and Statistics, 1988; 176 p. [in Russian]