Construction of a semi-automatic contour of areal objects on hyperspectral satellite images.

Bin Lei
CETC Les Information System Co., Ltd
40220720@qq.com

Wei Wan
CETC Les Information System Co., Ltd
1271130252@qq.com

Artiom Nedzved
FAMCS of Belarusian State University
Minsk, Belarus
artiom.nedzved@gmail.com

Alexander Nedzved
FAMCS of Belarusian State University
Minsk, Belarus
NedzvedA@tut.by

Alexei Belotserkovsky
Dept. ofIntelligent Information Systems United Institute
ofInformatics Problems of NASof
Minsk, Belarus
alex.belot@gmail.com

***Abstract*—In this article, we formalize the problem of constructing a semi-automatic contour of areal objects on satellite hyperspectral images, and present a solution algorithm using PCA and Dijkstra's algorithm. The contour is considered as the boundary of the object, which can be used for its segmentation and classification. The semi-automatic contour assumes reference points set by the operator. The formalization of the algorithm has been completed.**

Keywords—hyperspectrum, multispectrum, contour, object, feature space

I. INTRODUCTION

Multispectral and hyperspectral photography occupies an important place in solving applied problems using image analysis. Hyperspectral images are multidimensional data containing information about the spectral characteristics of each pixel of the image. Multispectral remote sensing data obtained using satellite images allow us to study the characteristics of objects on the Earth's surface that do not appear in panchromatic mode or in images of individual channels. One of the main applied tasks in the field of images is object detection. The solution of this problem on multispectral images is limited to channels and variations of their arithmetic combinations. Nevertheless, the algorithms for segmentation and selection of objects in the image are guided by the contour. A contour is the boundary of an object that can be used for its segmentation and classification. We will consider the task of selecting objects with a contour. To simplify the task, in this article we will consider the selection in semi-automatic mode, based on the methods of contour selection on a bitmap image.

II. DEFINITION OF INFORMATIVE ZONES OF SPECTRAL SPACE

From a number of articles [18-20] it has been established that the most suitable segmentation methods are energy methods of active circuits. Most often, the contour is built on the basis of the brightness gradient of a halftone image, and to build a contour on a hyperspectral image there are many problems that are caused by the dimensionality of space. The solution to this problem is based on lowering the dimension of the hyperspectral image feature space. Lowering the dimension of the space reduces to the following subproblem: find a subspace of a given dimension, in an orthogonal projection on which the root-mean-square distance between each pair of points is maximal. It can be solved using PCA (Principal Component Method). First of all, it is necessary to determine the reflection coefficients of individual images of bands for different objects and get some idea about the selection of objects in images of different spectral bands. To do this, the hyperspectral image is represented as spectral channels as in Figure 1.

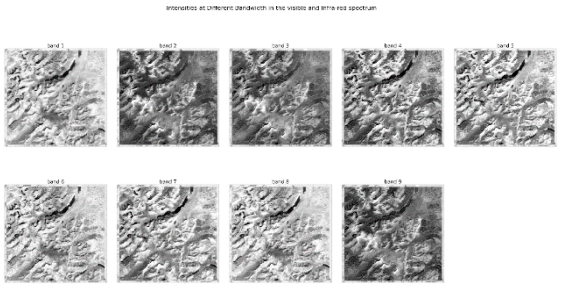


Fig. 1. Example of nine spectral channels of a satellite image, Figure 6 - Pairwise relationship in the initial spectral ranges.

In this case, the principal component method consists in the transformation of the axes, which occurs during its operation. The dot diagram shown in Figure 5 shows the correlation between the green and red stripe data. The original axes (X1, Y1) are now transformed into the axes of the main components (X2, Y2) defined by the eigenvectors of the covariance matrix of variables, and the data projected onto these new axes are the main components. It is important to note that the correlation that existed in the original data is eliminated after conversion to the space (X2, Y2). It should be noted that the variance of the main components along the X2 axis is higher than the variance of the original data along the X1 and Y1 axes, which means a better representation of the surface features, while the variability of the main components along the Y2 axis is less significant, therefore contains relatively less information about the surface features.

Before applying the principal component method, it is necessary to bring the data to a single format through standardization. The purpose of this is to make sure that variables are internally consistent with each other regardless of their type. Standardization is carried out by centering the variable by subtracting the average value, and then bringing them to a common scale by dividing the standard deviation.

Even if the images of spectral bands that are processed have the same range, standardization is not really necessary, but still its application guarantees a stable result. Variables that are two-dimensional arrays of images need to be transformed into a one-dimensional vector by smoothing to facilitate the calculation of the matrix. To do this, we create a variable matrix with a size of 935000 X 9 (the number of pixels in the spectral band, where X is the number of bands) and store these three-dimensional vectors in it. Now, by the method of principal components, the eigenvectors and the corresponding eigenvalues of the covariance matrix are calculated. At this stage, data is compressed and their dimension is reduced. If we look at the eigenvalues, we can see that the values are completely different. These values give us the order of significance of the eigenvectors or directions, i.e. the axis along the eigenvector with the largest eigenvalue is the most significant axis of the principal component, and so on. The next step is to arrange the eigenvectors by their eigenvalue, from largest to smallest, in order to rearrange the components in order of importance, as in Figure 2.

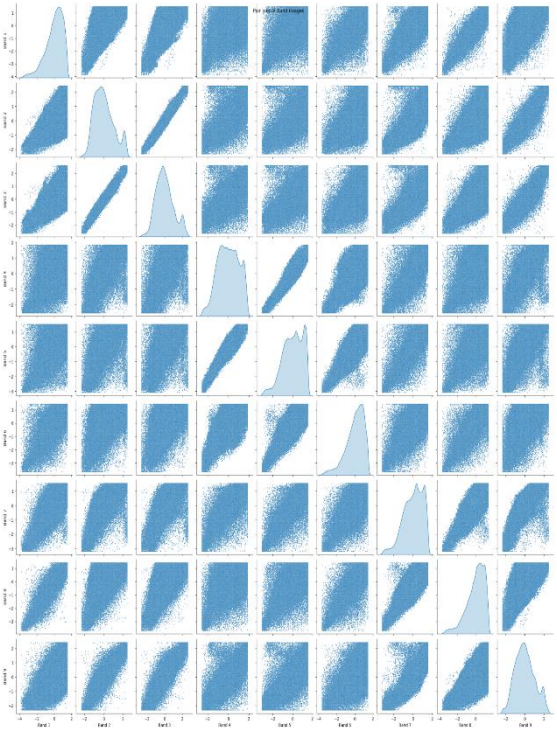


Fig. 2. Pairwise relationship in the initial spectral ranges.

It is necessary to project data in the directions of ordered eigenvectors, which, in turn, lead to the main components. Next, you need to check the components to check for redundancy reduction, as well as the extent to which data compression has been achieved. The dot graphs in Figure 7 show the pairwise relationship in the original ranges, and compare the same with the pairwise relationship of the PC to check for dependencies.

These paired graphs demonstrate the dependencies between variables that exist in the source data. They disappeared into the main components. thus, it is obvious that the principal component method was able to significantly reduce the dependence. The distribution graphs along the diagonal tell us that the principal component method also managed to extract deviations from the original data of high dimension, which may be associated with the possibility of

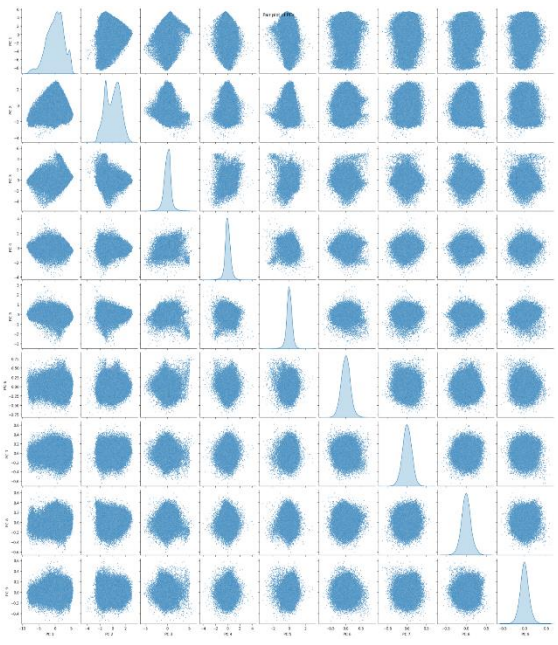


Fig. 3. Pairwise relationship in rebuilt components.

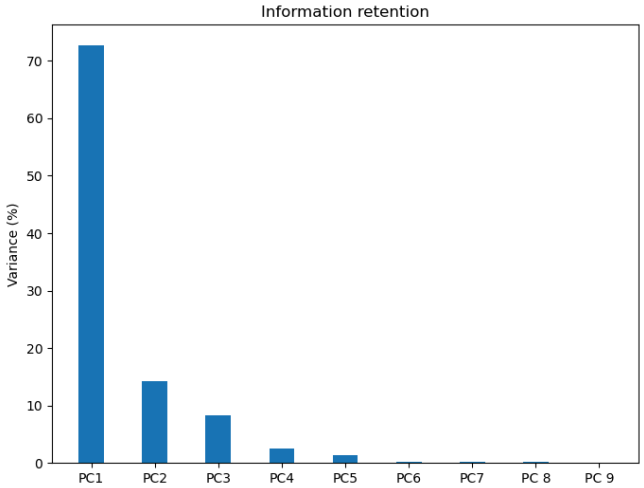


Fig. 4. Informativeness in images by new spectral components (plans) after the principal component method.

compression. The informativeness of the new components is shown in Figure 4.

These paired graphs demonstrate the dependencies between variables that exist in the source data. They disappeared into the main components. thus, it is obvious that the principal component method was able to significantly reduce the dependence. The distribution graphs along the diagonal tell us that the principal component method also managed to extract deviations from the original high-dimensional data, which may be associated with the possibility of compression. The informativeness of the new components is shown in Figure 8.Next, it is necessary to return the original image shape to the one-dimensional components and normalize the main component in the range from 0 to 255, which coincides with the range of the original image, in order to make it possible to visualize the image, as in Figure 5.

III. Building a contour on the gradient of the image

In [23, 25], the authors use the following characteristics to calculate the local communication price: the modulus and

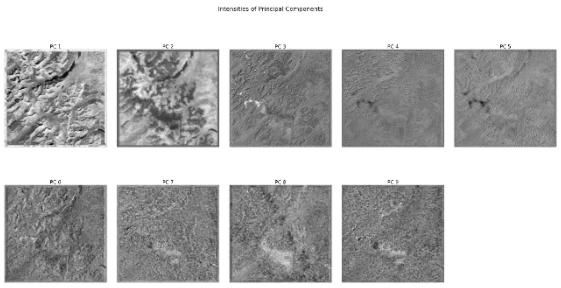
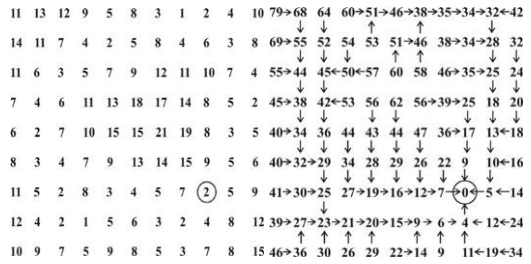


Fig. 5. Images of new spectral components (plans) after the principal component method.

direction of the gradient, pixel intensity, the value of the second derivative, and other characteristics, including those calculated during training and training of the algorithm.



a) b)

Fig. 6. Matrix of local weights and matrix of accumulated weights and direction map.

Figure 6a shows a local matrix of weights of a weighted graph that characterizes the local properties of image pixels. After the operator specifies the seed point (highlighted in a circle in Figure 6a), a direction map is calculated (Figure 6b), with which the optimal path from any point in the image to the seed point is constructed. Each number in Figure 6b corresponds to the cost of the path from this point to the seed point. To calculate the direction map specifying the optimal path, a search in a four-connected area was used. Dijkstra's algorithm is used to determine the optimal path. The optimal path is calculated and displayed in real time on the display from the current cursor position to the seed point. If the image is very noisy or contains objects of complex shape, then several boundary segments may be needed to define a segmenting contour. If the resulting segment adequately describes a part of the boundary of the object, then a new seed point is indicated for the next selected boundary segment.

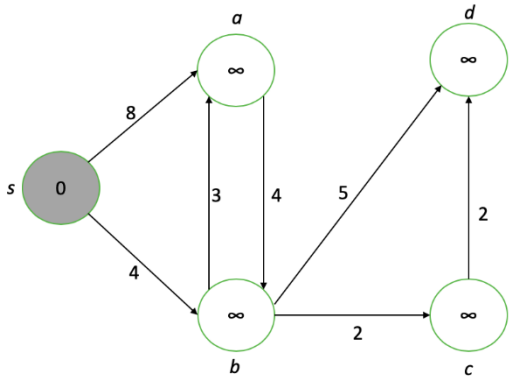


Fig. 7. The first iteration of Dijkstra's algorithm, to demonstrate the interaction of graph nodes

The initial node is the raw node with the lowest value (shown in gray), i.e. s. First, each adjacent vertex is weakened to the node of interest, updating their values to the minimum of their current value or the value of the node of interest plus the length of the connecting edge.

Node s is now completed, and its neighbors a and b take on new values. The new node of interest is b, so we repeat the process of "weakening" the neighboring nodes of b and finalizing the value of the shortest path for b. After going through each node, we will eventually get a graph showing the shortest path length from the source to each node, as in Figure 12.

In general , the algorithm can be represented as:

Input: w(p,q) is a function that sets the price of the transition from the node of the graph p to q; ps is the initial node; N(p) is a function that returns the neighbors of the node p.

Output: MD directions map.

Intermediate structures: AL − active list of nodes; E − array of processed nodes; S − array of total cost; Stmp − intermediate value of cost.

As a result of the algorithm's operation, a contour can be constructed between any two given points in the image, which will take into account typological features as in Figure 8.

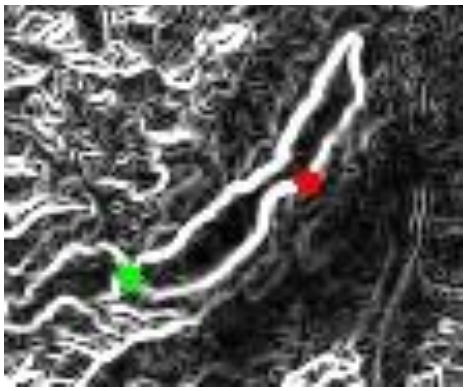


Fig. 8. Determination of points on the gradient image and construction of a contour fragment using Dijkstra's algorithm.

IV. Using the Template General Scheme of the Algorithm

Thus, the algorithm starts working based on the analysis of the hyperspectral image as in Figure 9. The principal component method makes it possible to obtain halftone images with the most pronounced properties in hyperspectral space. This image is used by Dijkstra's algorithm to determine a fragment of a contour enclosed between two points for a hyperspectral image.

V. Conclusion

In this article, we examined the algorithm for searching for hyperspectral boundaries in a multispectral image. The problem of image transformation in hyperspectral space has been studied, and a principal component method has been proposed to determine the informative zones of spectral space. This method allows you to identify the most significant components that can be used for further analysis.

To find the optimal contour, we used Dijkstra's algorithm, which allows us to efficiently find the minimum path in the

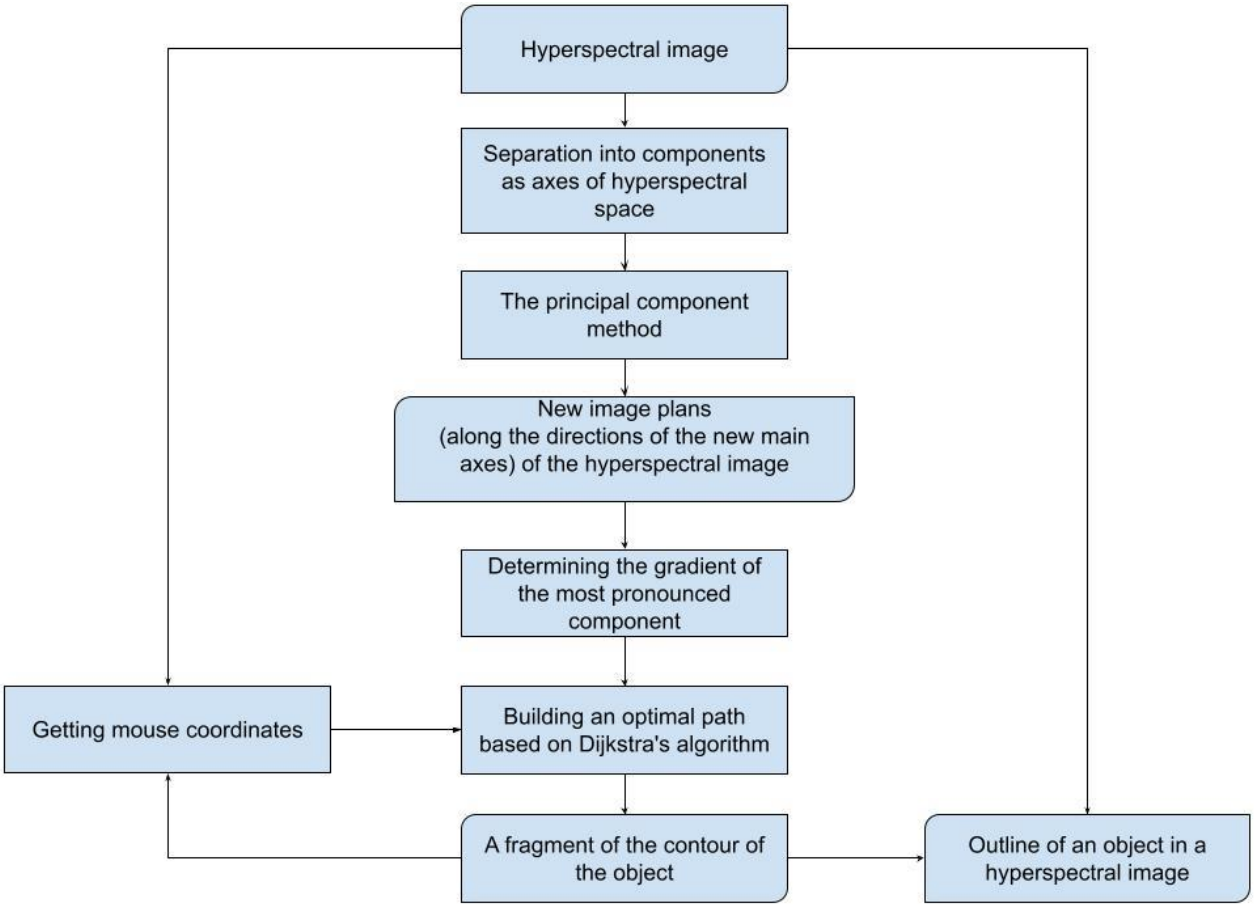


Fig. 9. Diagram of the algorithm for constructing a contour on a hyperspectral image

graph. This algorithm allows us to determine the boundaries of objects in a hyperspectral image, which is an important step in its analysis.

The general scheme of the algorithm was presented, combining all the above methods into a single system. This allows us to efficiently and accurately determine the boundaries of objects in multispectral images and conduct further data analysis.

REFERENCES

[1] I. V. Ognev, "Image processing by methods of mathematical morphology in an associative oscillatory medium," I.V. Ognev and N. A. Sidorova, Technical sciences. Computer science and computer engineering, 2007.

[2] T. V. Fisenko, "Computer processing and image recognition: textbook," V.T. Fisenko and T.Y. Fisenko. St. Petersburg. ITMO State University, 2008, p. 192.

[3] G. Potpourri, "Detection of contours oriented along edges and lines: the current state. Computing images and visual perception," G. Papary and N. Petkov, 2011, pp. 79-103.

[4] H. Cheng, H. Jiang and Yu. Song, "Segmentation of color images: achievements and prospects," H. Cheng, H. Jiang and Yu. Song et al., Pattern recognition, 2001, No. 34, pp. 2259-2281.

[5] H. V. Kang and Y. S. Sun, "Expanded bandwidth: interactive segmentation of images by constructing a map of incremental sections," Graphic Models, 2003, No. 64, pp. 282-303