Structural Analysis of Aerial Photographs
(HB47 Computer Vision: Assignment)

Xiaodong Lu, Jin Yu, Yajie Li
Master in Artificial Intelligence
May 2004

Table of Contents

1 Introduction

The task of this assignment is concerned with the analysis of multispectral aerial photographs by an image processing system. Based on the approach presented by Nagao and Matsuyama in [1], we develop an aerial photograph analysis program in Matlab, which is designed to automatically classify a multispectral aerial photograph into different categories, such as roads, houses, cars, grassland, shadow regions, etc. The analysis process implemented in our program can be divided into 5 steps:

(1). Edge-preserving smoothing using anisotropic diffusion
(2). Marker-controlled watershed segmentation
(3). Basic property calculation
(4). Extraction of characteristic regions
(5). Fuzzy rule-based object recogniton

Fig.1: Aerial photograph

In this paper, image processing methods implemented in our program will be discussed. For testing purpose those methods are applied to an aerial image in 4 color channels (see Fig.1).

2 Edge-Preserving Smoothing

Image smoothing is the set of local pre-processing methods whose predominant use is to remove image noise. A traditional way to smooth an image is to convolve the original image with a Gaussian kernel. The major drawback of this approach is that it blurs sharp edges. In order to avoid this problem we implement a smoothing method which is edge preserving.

In our program, an anisotropic diffusion approach is used, which is introduced by Perona and Malik in [2]. The approach is to use a conductivity function whose diffusion coefficient is chosen to vary spatially in such a way that the noise in the image is removed as the details of region boundaries are still preserved. In order to privilege high contrast edges, we implement the exponential conductivity function ($g(\Delta I) = \exp(-(\|\Delta I\| / k)^2)$). Compared to the original image (see Fig.2), as can be seen in Fig.3 the noise appears in the forest region is removed as the rectangle grassland boundaries remain sharp.

Fig.2: Original Image

Fig.3: Smoothed Image

3 Marker-Controlled Watershed Segmentation

After the noise is removed, segmentation processing is conducted on the smoothed image. One of the methods for obtaining a segmented image is to use watershed transform which works on gradient magnitude images. However, the basic watershed transform has the problem of oversegmentation. One solution to this problem is to “mark” regions of interest, which constrains the “water” to emerge only from those marked regions. But, the oversegmentation problem may still remain in some marker-controlled watershed segmentation (e.g. watershed segmentation using morphological techniques). This is mainly due to the existence of irrelevant markers inside objects.

Here we present a modified watershed algorithm, which merges irrelevant minima inside an object by using connected pixels which are inside objects as foreground markers. That is, these foreground markers correspond to interior areas of objects. In order to locate these interior (homogeneous) areas, we compute a homogeneous threshold by analyzing the histogram of the gradient magnitude image. As indicated in Fig.5, large amounts of elements (about 30% of total pixels) are contained in the first 4 bins. The reason for the high concentration of elements in the low-gradient region is that the number of interior pixels is much larger than the boundary pixels, and these interior pixels have relatively low gradient magnitude because of their similar gray levels, that means most of these low-gradient elements correspond to pixels inside objects and can be used as foreground markers. So we define the homogeneous threshold as the gradient magnitude which divides the histogram into two regions with 30% of the total elements in the low-gradient region. Adjacent pixels with gradient value below the threshold are merged and then labeled as markers (see Fig.4). Compared to the segmented image produced by normal marker-controlled watershed transform (see Fig.6), our segmentation method gives more sensible result (see Fig.7).

Fig.4: Highlighted foreground markers

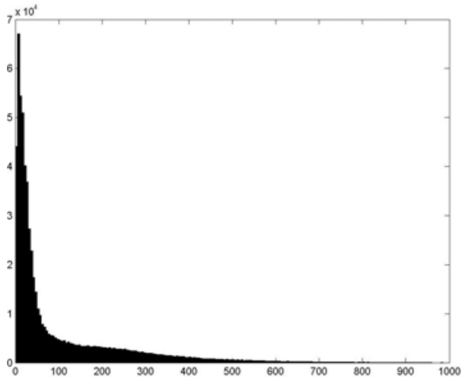

Fig.5: Histogram of gradient magnitude

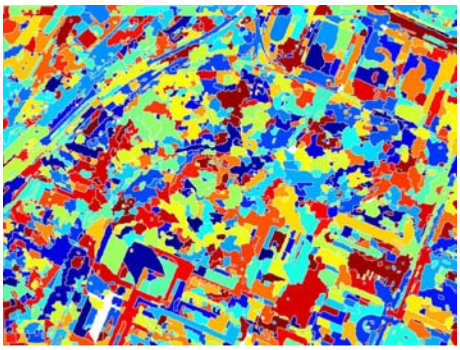

Fig.6: Segmentation using morphological techniques

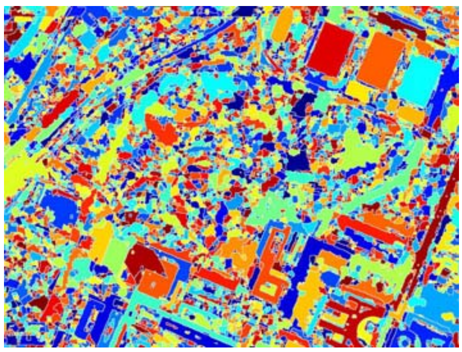

Fig.7: Segmentation by homogeneous thresholding

4 Basic Property Calculation

After segmentation, we get a segmented image which is referred to as the “label picture” [1]. The basic property calculation is conducted for each labeled region, and the calculation result is recorded in the “property table” [1]. These generic property data will be used for the further analysis of the image. In our program the following properties are computed:

- Location: corner coordinates of the smallest rectangle containing the region
- Shape feature: FIT, DIREC (referring to [1] for the naming) and ELONG (We define ELONG as the ratio between the length of major axis and the minor axis of the boundary ellipse (see Fig.8))
- Gray level: average gray level in each color channel (red, green, blue and infrared) and the average brightness among four channels

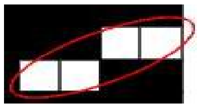

Fig.8: Boundary ellipse

In the program, we model the property table by a structure array which holds the property data in several fields. The Matlab function regionprops is used to facilitate the calculation.

5 Extraction of Characteristic Regions

Based on the generic property data of elementary regions, several kinds of “characteristic regions” [1] are extracted by analyzing the segmented picture. Because each characteristic region is defined in the nature language rather than by precise criteria, we integrate the fuzzy logic into the analyzing process so that the vague and imprecise concepts (e.g. large, elongated) which exist in the nature language can be properly handled. So, the result of the analysis is added to the property table in the form of membership values.

5.1 Large Homogeneous Regions

Large homogeneous regions are extracted by area size thresholding. The threshold value is determined by the “valley-detection” algorithm [1]. A membership function is defined which makes the region with larger area size get higher membership value. The region whose area size is larger than the threshold gets the highest membership value (1.0). Fig.9 shows large homogeneous regions with membership value higher than 0.9.

Fig.9: Large homogeneous regions

5.2 Elongated Regions

For extracting elongated regions, a membership function for elongatedness is defined. The elongatedness is measured based on the value of ELONG (referring to section 4 for the definition). Fig.10 shows elongated regions having membership value higher than 0.9.

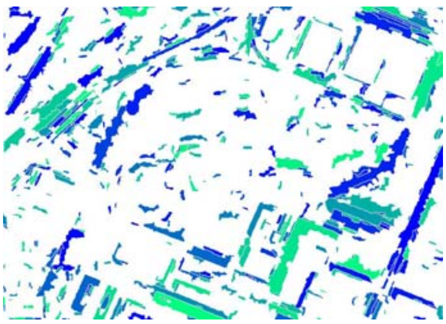

Fig.10: Elongated regions

5.3 Shadow Regions

Shadow regions are located by average brightness thresholding. The brightness threshold is determined using Otsu's method [3], which chooses the threshold to maximize the interclass variance. The membership function defined for the shadowiness favors darker regions and gives a region the highest membership value when its brightness is lower than the threshold. Fig.11 shows shadow regions with membership value higher than 0.9.

5.4 Shadow-Making Regions

Shadow-making regions are defined as the regions which have a long common boundary with an adjacent shadow region in the direction away from the sun. The selection of shadow-making regions is based on the analysis result of shadow regions whose membership value of shadowiness is higher than 0.9. Instead of using one criteria (the number of common boundary pixels larger than 5) to specify a long common boundary, we apply different criteria, which are empirically chosen. The effect of using different criteria is that the larger the region is the more common boundary pixels it is required to share with the adjacent shadow region. Fig.12 shows the most possible shadow-making regions.

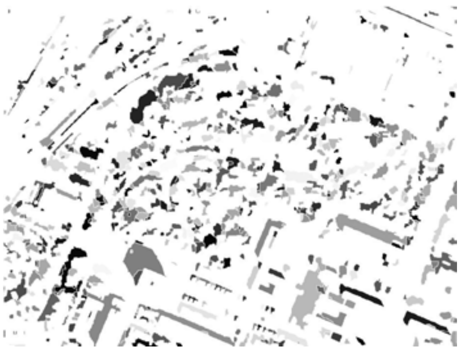

Fig. 11: Shadow regions

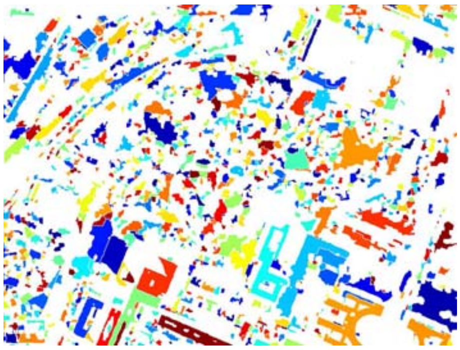

Fig. 12: Shadow-making regions

5.5 Vegetation Regions

As vegetation regions show prominent characteristics in the multispectral properties, such as the reflectance rate in the red band is much smaller than that in the infrared band, we define a membership function which privileges regions whose gray level in the infrared band is higher than that in the red band and gray level in the blue band is not very high. Fig.13 shows vegetation regions with membership value higher than 0.6. After extracting individual vegetation regions, adjacent vegetation regions with membership value higher than 0.6 are merged into one region. By applying large area thresholding, large vegetation areas are extracted. Fig.14 shows large vegetation areas with membership value larger than 0.9.

Fig.13: Vegetation regions

Fig.14: Large vegetation regions

5.6 High-Contrast Region

Areas containing diverse textures are referred to as high-contrast regions. So the density of elementary regions within a predefined window can be used to extract those areas. We implement the extraction procedure according to the five-step method presented in [1] except that in step 2 we use the number of different regions within one window instead of the number of boundaries points as the measure of texture diversity. As can be seen in Fig.15, high-contrast (non-homogeneous) regions (membership value higher than 0.9) are extracted. After extracting individual high-contrast regions, adjacent high-contrast regions with membership value higher than 0.9 are merged into one region. And the large area thresholding is applied to select large high-contrast regions. Fig.16 shows large high-contrast regions with membership value of being a large high-contrast region higher than 0.9.

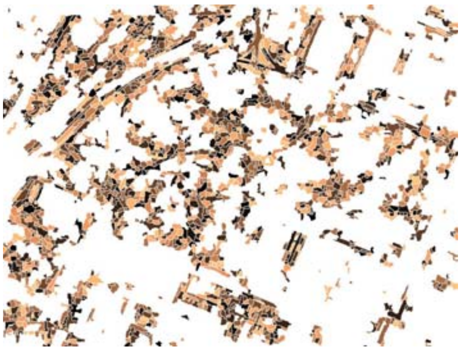

Fig.15: High-contrast regions

Fig.16: Large high-contrast regions

5.7 Water Regions

Although the photo we use in our program does not contain water regions, following the procedure presented in [1] we implement the water region extraction based on multispectral properties of elementary regions.

6 Fuzzy Rule-Based Object Recognition

After the extraction of various characteristic regions, fuzzy rule-based object recognition is conducted. As the basis of fuzzy inference, a set of fuzzy rules are used to specify candidate regions. We apply those fuzzy rules to every elementary region, and for each region, the outputs of fuzzy inferences are defuzzified by membership value thresholding. For instance, if the membership value of being a crop field is higher than 0.6, the region will be selected as a candidate region for crop field. After extracting candidate regions, further recognition processing (calculation of new features, recognition judgment and correction of segmentation errors, etc.) is triggered. Due to the time limit, we only implement the candidate region extraction in our program.

6.1 Crop Field

The candidate regions for crop field (see Fig.17) are specified by the following logical expression:

- (LargeHomogeniousRegion) & (VegetationRegion) & (~ShadowRegion) & (~ShadowMakingRegion)

6.2 Bare-Soil Field

The candidate regions for bare-soil field (see Fig.18) are specified by the following logical expression:

- (~VegetationRegion) & (~ShadowMakingRegion) & (~ShadowRegion)

6.3 Building

Generally, buildings in an aerial photo display a relatively high gray level. So, first of all we filter out regions whose average brightness is lower than 80, then we apply the fuzzy rule to the remained regions. The candidate regions for building (see Fig.19) are specified by the following logical expression:

- (ShadowMakingRegion) & (~LargeHighContrast) & (~VegetationRegion)

6.4 House

The candidate regions for house (see Fig.20) are specified by the following logical expression:

- (ShadowMakingRegion) & (~ShadowRegion) & (~VegetationRegion) &(~ LargeHomogeniousRegion)

6.5 Forest

The candidate regions for forest (see Fig.21) are specified by the following logical expression:

- (LargeVegetationRegion) & (HighContrastRegion) & (~LargeHomogeniousRegion)

6.6 Grassland

The candidate regions for grassland (see Fig.22) are specified by the following logical expression:

- (VegetationRegion) & (~ShadowMakingRegion)

6.7 Road

The candidate regions for road (see Fig.23) are specified by the following logical expression:

- (ElongatedRegion) & (~ShadowMakingRegion) & (~VegetationRegion)

6.8 Car

In addition to applying the following fuzzy rule, the contextual information that cars are usually on roads is also used for extracting candidate regions for car (see Fig.24).

- (~LargeHomogeniousRegion) & (~VegetationRegion) & (~ShadowRegion) & (~ShadowMakingRegion) & (~ HighContrastRegion) & (~ElongatedRegion)


Fig.17 Crop field

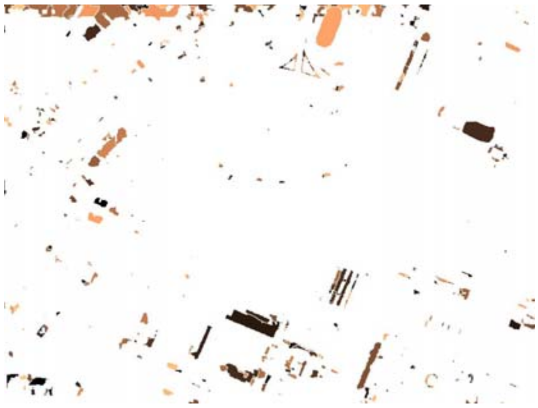
Fig.18 Bare-soil field

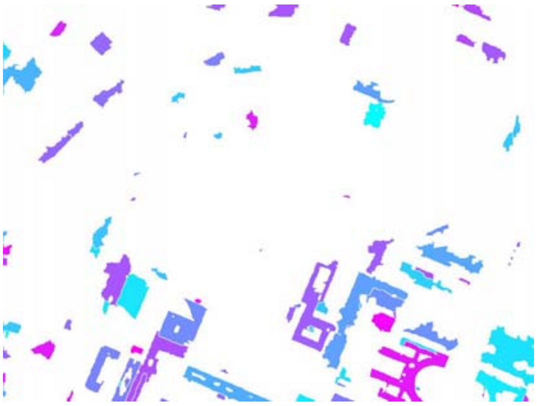
Fig.19 Building


Fig.20 House

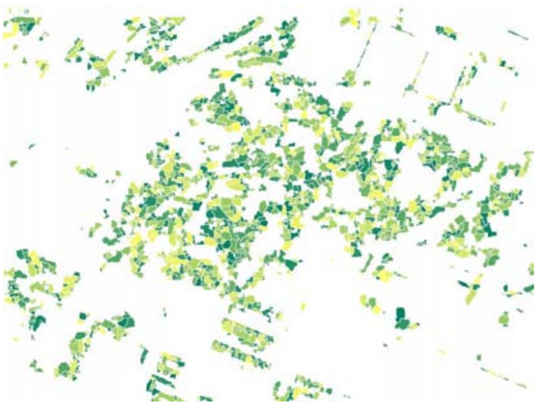
Fig.21 Forest

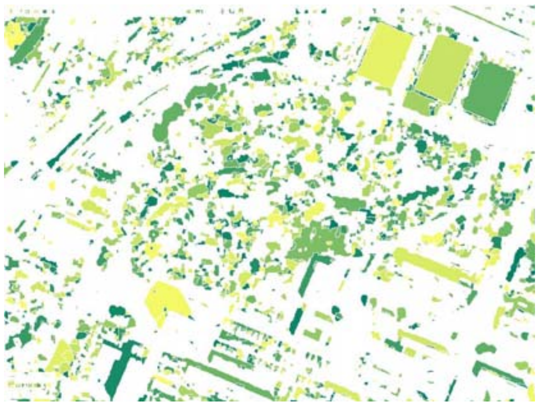
Fig.22 Grassland

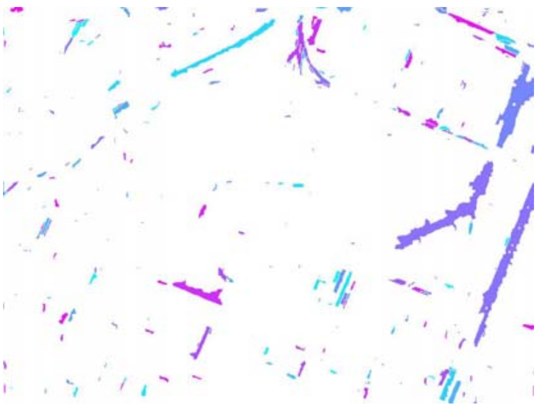
Fig.23 Road


Fig.24 Car

7 Conclusion

The main purpose of this assignment is to implement an image processing system for the structural analysis of complex aerial photographs. After the thorough study of various image processing algorithms and the approach introduced by Nagao and Matsuyama in [1], we develop a Matlab program, which can be used to analyze aerial photographs of urban areas. In order to enhance the performance of the program, we present some new processing methods, such as using a homogeneous threshold to create markers for watershed segmentation, using the boundary ellipse instead of the minimum bounding rectangle to measure the elongatedness of a region, etc. These methods prove to be effective according to the processing result. In addition, we extend the knowledge-based object recognition approach introduced in [1] by integrating fuzzy inference into the analysis procedure, which enables the program to handle vague and imprecise concepts. Although due to the time limit the assignment has not been completed, the candidate regions chosen by the current program can already give a sensible description of the urban area under analysis.

Reference

[1]. Makoto Nagao, Takashi Matsuyama.
A Structural Analysis of Complex Aerial Photographs.

[2]. P. Perona and J. Malik(1990).
Scale-Space and Edge Detection Using Anisotropic Diffusion.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 12(7):629–639.

[3]. N. Otsu(1979).
A Threshold Selection Method from Gray-Level Histograms.
IEEE Transactions on Systems, Man, and Cybernetics, 9(1): 62-66.