Title: Colour Based Man-Made Object Detection in an Aerial Image
Author: D. Tien, X. Yi and M. Zhao
Author Address: dtien@csu.edu.au
Conference Name: Fourth International Conference on Information Technology & Applications
Year of Conference: 2007
Conference Location: Harbin, China
Pages: pp306-309
URL: http://icita.hlju.edu.cn/
Abstract: paper presents a colour based approach for urban object recognition. The main framework is to separate candidature objects by the nearest neighbour classification and then the desired objects are identified from the candidates with the help of geometry features. For the classification, the points of a targeting classifier and a non targeting classifier are established from training samples. The experiment shows the method is superior to other methods in extracting urban objects with mixed colours or single colour in various illuminations.
Colour based man-made object detection in an aerial image

D. Tien, Y. Xiao and Mansuo Zhao

Abstract—This paper presents a colour based approach for urban object recognition. The main framework is to separate candidate objects by the nearest neighbour classification and then the desired objects are identified from the candidates with the help of geometry features. For the classification, the points of a targeting classifier and a non targeting classifier are established from training samples. The experiment shows the method is superior to other methods in extracting urban objects with mixed colours or single colour in various illuminations.

Index Terms—Image processing, colour segmentation, nearest neighbour classification

I. INTRODUCTION

The extraction of man-made objects such as buildings, swimming pools and roads, from the aerial data is one of the major steps in automatic updating or retrieving spatial information for the successful maintenance of large GIS databases.

Building extraction is the issue frequently discussed in the literature for man made object detection. Many methods have been proposed to solve the problem of building detection and description. Some of which rely on a variety of data sources, such as laser scanner data, multispectral stereoscopic images, multiple view images and GIS data [1,2]. These methods focus on building descriptions and the detection is often carried out locally. Many trials have also been undertaken to work from one grey scale image. Geometric regularity and the features of edges, shadows and corners are the basis of the approaches for target searching. Consequently, they are typically restricted to the case of rectangular buildings. Another problem is the inadequacy in tracing the boundaries of objects as the results of darker coloured roofs or shade or poor contrast between roofs and ground.

Colour images convey more information than grey-level images. They allow one to obtain more meaningful and robust identification results. With the broad spread and easy access of the high-resolution colour aerial images, it is expected the colour aerial information to be used as a principal spectral source for revision, updating and quality assessment of many land use types within GIS databases.

II. RELATED WORK

Various colour models for colour image segmentation have been studied. M. Jolly and A. Gupta [3] assumed colours in a class have multivariate Gaussian distribution, which is characterized with its colour mean and deviation. A pixel is labeled as the targeting object if its colour information satisfies the equation as below.

\[
f(p) = \begin{cases} 
1 & \text{if } \sum_{i=1}^{3} (y_{p,i} - \mu_{i})^{2} \leq \alpha \cdot \sum_{i=1}^{3} (\sigma_{i})^{2} \\
0 & \text{otherwise.}
\end{cases}
\]

where, \(y_{p,i}\) is the colour value in \(i\) th colour channel at location \(p\), \(\mu_{i}\) and \(\sigma_{i}\) are mean and standard deviation of input data, and \(\alpha\) is a given error tolerance parameter.

J. Chen and T. N. Pappas [4] proposed an adaptive clustering algorithm that separates the pixels in the image into clusters based on both their colours and their relative location. A posteriori probability density function is defined for the distribution of regions given the observed image. The probability density function has the form

\[
p(x \mid y) \propto \exp \left\{-\sum_{i=1}^{C} \frac{1}{2\sigma_{i}^{2}} \left[ y_{i} - \mu_{i} \right]^{2} - \sum_{i=1}^{C} V_{C}(x) \right\}
\]

where

\[
V_{C}(x) = \begin{cases} 
-\beta, & \text{if } x_{s} = x_{q} \text{ and } s, q \in C \\
+\beta, & \text{if } x_{s} \neq x_{q} \text{ and } s, q \in C.
\end{cases}
\]

\(\mu_{i}\) is the mean intensity of region \(i\) at location \(s\) and \(\sigma_{i}^{2}\) is the variance of white Gaussian process. \(x_{i} = i\) means that the pixel at \(s\) belongs to region \(i\). \(V_{C}\) is the clique potentials of clique \(C\). A clique is a set of points that are neighbors of each other.

The objects in urban aerial image contain homogeneous colour regions but one object may contain colour under different illumination or even mixed colours, which causes the irregularity and overlap of the data distribution of an object in colour space. For example, for three training sets of red roofs, gray roofs and grass in an aerial image, colours in all the three objects stretch to a wide range and are irregular-shaped. Some distances between two elements within roof class are greater than the distances between red roofs and gray roofs. Also there are some overlaps between grass samples and mixed colour roofs samples as the mixed colours exist in both training samples. Thus, the appearance of an object region is best described by the distribution of feature samples, trained from images containing the desired object regions with a wide variety. While the existing methods such as clustering and statistical models are not sufficient to handle these problems.

III. OUR METHOD

In this paper, we proposed a nearest neighbour rule based method for man made object identification. The proposed method consists of two steps. At the first step, the pixels in the
colour image are classified as desired objects and non desired object (background) according to the nearest neighbour rule in colour space. For classification, the points of a targeting classifier are collected from the typical areas of the desired object in the image and the points in the non targeting classifier are established from the areas of other objects that have similar colours to the targeting object in the image. At the second step, the desired objects are identified from the segmented image by geometric feature measuring.

IV. NEAREST NEIGHBOUR RULE CLASSIFICATION

Nearest neighbour decision rule classification assigns a query point to the class of its nearest neighbour in the measurement space using, most commonly, Euclidean metrics.

\[ \text{d}_{\text{min}}(c(x, y), A) = \min |c(x, y) - c(x, y)| \]

where \( c(x, y) \) is the input of the colour image \( I \), when pixel value is 1, the pixel is classified as a pixel in A. Otherwise the pixel is not in A.

Figure 1 shows a two class (A and B) problem. Query point \( q \) is classified as A as its nearest point is in A. Class B overlaps class A, a point in B has the closest neighbour that belongs to A instead of B.

Nearest Neighbour Decision rule [5, 6] has the advantage of modeling arbitrary spatial pattern, which leads to an accurate solution for data classification despite the diverse nature of the data structure. It has been employed in many applications in pattern recognition and computer vision, such as handwriting recognition, forest statistics and map segmentation. Furthermore, the effective nearest neighbour searching via proximity graph make it possible for more practical applications.

V. MAN MADE OBJECT IDENTIFICATION

In this paper, a colour image is represented by its three colour component value \( r(x, y) , g(x, y) , b(x, y) \) at each pixel \( p(x, y) \). The colour feature vector at the pixel \( p(x, y) \) is denoted as \( c(x, y) = [r(x, y) , g(x, y) , b(x, y)] \).

Training sets are collected by hand. The sets consist of colour feature vectors of known objects from original image. The elements in the training set have higher appearance values.

When segmentation is treated as a classification process, normally one training set is selected. To overcome the misclassification caused by colour overlapping, we add one more training set. Therefore we have three classes: the object interested, the object having similar colours with the object interested and background. In the algorithm, an additional condition is applied to judge the belonging of a point.

Let \( d_{\text{min}}(c(x, y), A) \) be the nearest distance between the distances from \( c(x, y) \) to the points in training set A in colour space. Then,

\[ d_{\text{min}}(c(x, y), A) = \min |c(x, y) - c(x, y)| \]

where \( c(x, y) \) is the input of the colour image \( I \), when pixel value is 1, the pixel is classified as a pixel in A. Otherwise the pixel is not in A.

The examples of the extracted results are illustrated in Figure 2.

VI. EVALUATION

The proposed algorithm has been evaluated with two sets of images taken above two suburbs in NSW, Australia respectively, with the resolutions of 15cm/pixel and 20 meter/pixel. Objects of interest include buildings, swimming pools and roads. Each of the tested images had its own training sets. For training, we selected an average of 10 areas for each object. The areas covered the colour variation of the targeting objects.

As mentioned in Section 5, the threshold \( d_{th} \) needs to be specified. \( d_{th} \) is for separating targeting object and non targeting object regions of the image. A too small or large value of \( d_{th} \) will lead the under-segmentation or over-segmentation of the object regions. In the experiment, we found \( d_{th} = 7 \text{(pixel)} \) gave best results.

The examples of the extracted results are illustrated in Figure 2.
Table 1 listed the object detection rates obtained by using the proposed method.

Pool detection had a high detection rate in both images as the colour elements of pools were well separated from other objects' in colour space. All pools except the very small ones and the pools largely covered by trees were extracted.

For red roof detection, the high resolution image set yielded more accurate results than the low resolution image set as the corresponding training set of high resolution image set covered more colour variation.

In gray roof detection, the high resolution image set achieved much higher detection rate than the low resolution image set. Their undetected mixed colour roofs were the small roofs. The low resolution image set had low detection rate, because the colour elements of gray roofs were very close to or even overlapped the colour elements in grass and paved paths.

Table 1 Detection rate

<table>
<thead>
<tr>
<th></th>
<th>Red roof</th>
<th>Gray roof</th>
<th>Pool</th>
<th>Road</th>
</tr>
</thead>
<tbody>
<tr>
<td>15cm/pixel</td>
<td>100%</td>
<td>92%</td>
<td>98%</td>
<td>89%</td>
</tr>
<tr>
<td>20m/pixel</td>
<td>94%</td>
<td>82%</td>
<td>92%</td>
<td>78%</td>
</tr>
</tbody>
</table>

There were some false alarms caused by colour sharing between different objects, such as gray roofs and road; red roofs and soil. Future studies will consider more information such as depth and texture.

Comparisons of building roof extraction were made among Gaussian model (GM) [3], adaptive clustering algorithm (ACA) [4] and our proposed method. To evaluate the results numerically, the completeness (cp) and correctness (ct) used in [7] were applied. cp measures the ability to find the object regions and ct measures the accuracy of object detection.

Table 4 cp and ct

<table>
<thead>
<tr>
<th></th>
<th>Red roof</th>
<th>Gray roof</th>
</tr>
</thead>
<tbody>
<tr>
<td>Our method</td>
<td>100%</td>
<td>86%</td>
</tr>
<tr>
<td>ACA</td>
<td>98%</td>
<td>76%</td>
</tr>
<tr>
<td>GM</td>
<td>100%</td>
<td>89%</td>
</tr>
</tbody>
</table>

For single colour objects (red roofs, swimming pools), both GM and our method showed good extraction results. ACA gave reasonable result for single colour object with small range of colour variation. For mixed colour objects (gray roofs, roads) our method obtained robust extraction results. However, GM missed some colour elements belong to the desired objects while ACA clustered other object regions that have the similar colour with the desired objects. Our method had a notable improvement of the extraction accuracy for mixed colour object identify, in which the GM and ACA methods failed.

It can be seen from the figures that as the detection is based on colour, it can extract the man-made objects with arbitrary shapes and orientation.

To measure the detection accuracy, we define the detection rate as $R = \frac{\text{number of objects extracted}}{\text{total number of objects targeted}}$. 

![Figure 2](image-url)
VII. CONCLUSION

The objects in urban aerial image contain homogeneous colour regions but one object may contain a colour with different illumination or even mixed colours. The mixed colours on object surfaces make the urban object segmentation more difficult.

The proposed technique represents an object by its actual colour distribution, no matter the variation of the colour. It is salient not only to identify single colour varying in illumination but also to mixed colours.

According to the study results, the proposed method achieved reliable recognition of man-made objects. As the detection is based on colour, it can identify not only the rectangular man-made objects but also the objects with arbitrary shapes and orientations. And as the extraction accuracy for mixed colour object is high, the method is particularly suitable for urban applications maintenance of large GIS databases.

REFERENCES


