COLOUR AERIAL IMAGE SEGMENTATION USING A BAYESIAN HOMOGENEITY PREDICATE AND MAP KNOWLEDGE

F. Quint S. Landes
Institute for Photogrammetry and Remote Sensing
University of Karlsruhe
76128 Karlsruhe, Germany
quint@ipf.bau-verm.uni-karlsruhe.de

Commision III, Working Group 2

KEY WORDS: Colour, Aerial, Image, Model, Vision, Colour Aerial Image, Aerial Image Segmentation, Bayesian Model

ABSTRACT

In this article we present a homogeneity predicate for segmentation purposes. It is based on the probability of a pixels to fulfill the model assumptions for a region. For some practical relevant models, a closed formula for this probability is given. The homogeneity predicate is used in a region growing procedure to segment colour aerial images. In this application, estimates for the position of initial seed regions and the model type to be used are extracted from topographical maps.

KURZFASSUNG

In diesem Artikel wird ein Homogenitätsprädikat für die Segmentierung von Bildern vorgestellt. Es beruht auf der Wahrscheinlichkeit für einen Bildpunkt, daß er den getroffenen Modellannahmen entspricht. Für einige praxisrelevante Modelle kann eine geschlossene Formel zur Berechnung dieser Wahrscheinlichkeit angegeben werden. Das Homogenitätsprädikat wird in einem Flächenwachstumsverfahren zur Segmentierung von Farbluftbildern verwendet. Schätzwerte für anfängliche Kristallisationspunkte des Flächenwachstumsverfahrens und die zu verwendenden Modelle werden aus Karten gewonnen.

1 INTRODUCTION

Segmentation of images into physically meaningful regions is one of the most often addressed problems in computer vision literature. The periodically appearing review articles give a good overview of the domain, see e.g. (Haralick and Shapiro, 1985),(Pal and Pal, 1993).

Haralick and Shapiro (Haralick and Shapiro, 1985) categorize the different segmentation procedures according to the control algorithm they use, in:

- measurement space guided spatial clustering,
- region growing,
- spatial clustering and
- split and merge schemes.

In (Pal and Pal, 1993) the different image segmentation techniques are reviewed according to the used homogeneity predicate. It is made distinction between:

- gray level thresholding,
- iterative pixel classification,
- surface based segmentation,
- segmentation of colour images,
- edge detection based approaches and
- methods based on fuzzy sets.

The method presented in the current article is a region growing scheme. As homogeneity predicate we use the a-posteriori probability for the features of an image pixel to fulfill an a-priori model of a region. Similar approaches for the homogeneity predicate, embedded in different segmentation schemes have also been made in (Silverman and Cooper, 1988),(LaValle and Hutchinson, 1995).

In section 2 we describe our model assumptions. A closed formula for calculating the probability of homogeneity is derived in section 3. After a brief look at computational issues in section 4, we give in section 5 an example for a simple, planar model. Finally we show how the developed procedure can be used for segmenting colour aerial images. Initial seed regions for the region growing scheme and information on the model type to be used are extracted from map knowledge.

1.1 Segmentation procedure

Our definition of segmentation follows (Pavlidis, 1977): it is the partition of the image in pairwise disjunct regions $\mathcal{R}_r$, which, in their union cover the whole image. In order to assign a pixel to a region, it must fulfill two conditions:

- it must be neighbour with at least one other pixel of the region (connectedness condition)
- a homogeneity predicate between the pixel and the region must evaluate to true (homogeneity condition)

We implement our segmentation procedure as a region growing scheme: For each pixel of the image which is not already marked as belonging to a region and which is neighbour to at least one region a homogeneity predicate is tested. The pixel is marked as belonging to the region for which the tested predicate evaluates to true. The procedure stops when all pixels are assigned to a region.

The homogeneity predicate is calculated using the a-posteriori probability of a pixel for belonging to the current region. This probability is calculated according to a model described in the following section. We calculate the a-posteriori probability for all regions, which the pixel is neighbouring. The homogeneity predicate evaluates to true for the region with the highest probability.

2 THE MODEL

2.1 Image formation

For simplicity of expression, we will call the quantities forming the image light intensities. The presented scheme however is

not limited only to the segmentation of optical images, it can also be applied to the segmentation of range images or other types of images.

In an ideal image formation process the light intensities of points in the scene form the intensity $I(x_k, y_k)$ at pixel (x_k, y_k) in the image. Because of degradation, the „true" intensities $I(x_k, y_k)$ are not observable, accessible are only the gray values $g(x_k, y_k)$ of the image. For simplicity we will denote a location (x_k, y_k) only with its index k, e.g. instead of $g(x_k, y_k)$ we write g_k.

We assume that the degradation is due to additive white noise with a Gaussian probability density function (pdf) and zero mean value. The noise is statistically independent from the light intensities $I(x_k, y_k)$. Nonlinearities due to saturation, aliasing and quantization effects are neglected. Accordingly, we have for the gray values in the image:

$$g_k = I_k + n,$$

where n is a realization of the Gaussian white noise. This leads for the a-posteriori pdf of the gray values in the image to:

$$f_g(g_k \mid I_k) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(g_k - I_k)^2}{2\sigma^2}\right),$$

where σ^2 is the variance of the Gaussian noise.

We also need a prior model for the light intensity I_0 of the pixel (x_0, y_0), for which the homogeneity condition is tested. The prior model reflects our expectations in the value of the intensity I_0 *before* the pixel was assigned to a particular region. Since a-priori we have no reason to believe that some intensities are preferred, we assume a uniform density on the bounded definition space D_I. With $\Delta I = I_{max} - I_{min}$, we have:

$$f_I(I_0) = \begin{cases} \frac{1}{\Delta I} & : \ I_0 \in D_I \\ 0 & : \ \text{otherwise.} \end{cases}$$

2.2 Region model

Our model for a region $\mathcal{R}$ is a parametric model. The „true" light intensities of the pixels belonging to the same region satisfy the equation:

$$I(x_k, y_k) = \sum_{j=1}^{J} a_j \, \phi_j(x_k, y_k) \qquad (1)$$

with $\{k \mid (x_k, y_k) \in \mathcal{R}\}, a_j \in \mathbb{R}$.

The functions $\phi_j(x, y)$ are arbitrary, real-valued functions, which are supposed to be known for a given region. However, it is not necessary that these functions are the same for all regions in the image. In our task of map based segmentation of aerial images, we choose the model of a region (i.e. the functions $\phi_j(x, y)$) according to knowledge gained from maps.

The parameters $a_j, j = 1, ..., J$ in equation (1) are unknown and have to be estimated. However, as we will show later, if we are interested only in the segmentation of the image and not in the parametric description of the regions, the explicit calculation of their values is not necessary. We assume that these parameters are random variables over the set of regions in the image and have an a-priori Gaussian pdf with mean m_j and standard deviation σ_j:

$$f_a(a_j) = \frac{1}{\sqrt{2\pi}\sigma_j} \exp\left(-\frac{(a_j - m_j)^2}{2\sigma_j^2}\right).$$

2.3 Homogeneity predicate

We now define the predicate used for testing the homogeneity condition. Let $(x_k, y_k), k = 1, ..., K$ be the pixels already marked as belonging to region $\mathcal{R}_r$. Their (unmeasurable) light intensities $I(x_k, y_k)$ (or short I_k) fulfill equation:

$$I_k = \sum_{j=1}^{J} a_j^{(r)} \phi_j^{(r)}(x_k, y_k). \qquad (2)$$

We denote with (x_0, y_0) the pixel for which the homogeneity predicate is tested in the current step. Its gray value is g_0 and its light intensity is I_0. The homogeneity predicate $\mathcal{H}_r$ for pixel (x_0, y_0) and region $\mathcal{R}_r$ evaluates to true ($\mathcal{H}_r = 1$), if

$$I_0 = \sum_{j=1}^{J} a_j^{(r)} \phi_j^{(r)}(x_0, y_0). \qquad (3)$$

Otherwise, $\mathcal{H}_r$ evaluates to false ($\mathcal{H}_r = 0$). According to this definition, the conditional probability of the predicate $\mathcal{H}_r$ is:

$$P_{\mathcal{H}}(\mathcal{H}_r = 1 \mid a_j^{(r)}, I_0) = \begin{cases} 1 \ : \ I_0 = \sum_{j=1}^{J} a_j^{(r)} \phi_j^{(r)}(x_0, y_0) \\ 0 \ : \ \text{otherwise.} \end{cases}$$

$P_{\mathcal{H}}(\mathcal{H}_r = 0 \mid a_j^{(r)}, I_0)$ and $P_{\mathcal{H}}(\mathcal{H}_r = 1 \mid a_j^{(r)}, I_0)$ are complementary.

The random variables needed for testing the homogeneity predicate according to equation (3) are unmeasurable. Accessible are only the gray values g_k of the image. Hence, we redefine our homogeneity predicate and consider the a-posteriori probability $P_{\mathcal{H}}(\mathcal{H}_r = 1 \mid g_k)$, $k = 0, ..., K$. We call this expression *probability of homogeneity*. If the calculated value for the probability of homogeneity exceeds a given threshold we take the decision, that pixel (x_0, y_0) belongs to the region $\mathcal{R}_r$

3 PROBABILITY OF HOMOGENEITY

To illustrate the dependencies between the different random variables which appear in the calculation of the probability of homogeneity, we represent them in a Bayesian network (see e.g. (Pearl, 1986)). The nodes of the network contain the random variables. If there exists a direct causal influence of one random variable on the behavior of a second one, an arc of the graph leads from the node of the first variable to the node of the second one. The strengths of the dependencies are quantified by conditional probabilities.

Consider the situation, where the homogeneity predicate for pixel (x_0, y_0) and region $\mathcal{R}_r$ is tested. The region $\mathcal{R}_r = \{(x_k, y_k) \mid k = 1 \ldots K\}$ already contains K pixels. The corresponding Bayesian network is given in Figure 1. The probability for the homogeneity predicate to evaluate to true given the gray values of the image (i.e. the probability of homogeneity) is calculated considering the dependencies given in the network. After successful predicate testing the Bayesian network is updated since the number of pixels in the region has increased. Each decision situation has its particular Bayesian network.

The probability of homogeneity can be written as:

$$P(\mathcal{H}_r = 1 \mid \{g_k\}, g_0) = \frac{P(\mathcal{H}_r = 1, \{g_k\}, g_0)}{\sum_{\mathcal{H}} P(\mathcal{H}_r, \{g_k\}, g_0)} = \frac{P_Z}{P_N}. \qquad (4)$$

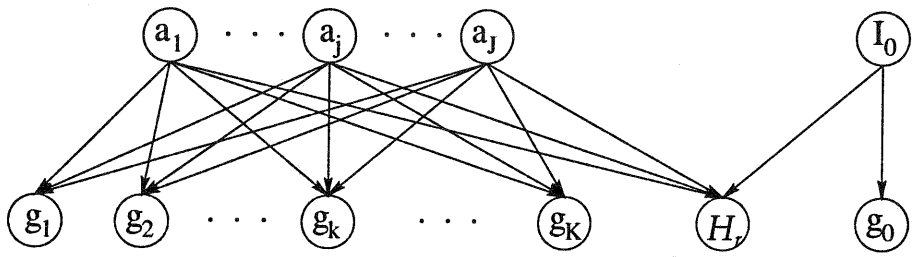

Figure 1: Bayesian network for a particular decision situation

The nominator P_Z of equation (4) is calculated by marginalizing the joint probability distribution:

$$P_Z = \sum_{a_j^{(r)}} \sum_{I_0} P(\mathcal{H}_r = 1, \{a_j^{(r)}\}, I_0, \{g_k\}, g_0)$$

with $j = 1, ..., J$, $k = 1, ..., K$. Considering the dependencies between the random variables in the Bayesian network in Figure 1, the joint probability distribution $P\big(\mathcal{H}_r, \{a_j^{(r)}\}, I_0, \{g_k\}, g_0\big)$ of the random variables in the network results to:

$$P\big(\mathcal{H}_r, \{a_j^{(r)}\}, I_0, \{g_k\}, g_0\big) = P_{\mathcal{H}}(\mathcal{H}_r \mid \{a_j^{(r)}\}, I_0) \times P(g_0 \mid I_0)\, P(I_0) \Big(\prod_{k=1}^{K} P\big(g_k \mid \{a_j^{(r)}\}\big)\Big) \Big(\prod_{j=1}^{J} P(a_j^{(r)})\Big).$$

Using the probability density functions as given in section 2 and observing that for the pixels (x_k, y_k), $k = 1, ..., K$ already assigned to region $\mathcal{R}_r$ equation (2) is fulfilled, the expression for P_Z becomes to:

$$P_Z = \int_{a_1^{(r)}} \cdots \int_{a_J^{(r)}} \int_{I_0} P_{\mathcal{H}}(\mathcal{H}_r = 1 \mid \{a_j^{(r)}\}, I_0)\, \frac{1}{\Delta I} \times \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(g_0 - I_0)^2}{2\sigma^2}\right) \frac{1}{\left(\sqrt{2\pi}\sigma\right)^N} \times \left(\prod_{k=1}^{N} \exp\left(-\frac{\left(g_k - \left(\sum_{j=1}^{J} a_j^{(r)} \phi_j^{(r)}(x_k, y_k)\right)\right)^2}{2\sigma^2}\right)\right) \times \frac{1}{\left(\sqrt{2\pi}\sigma_j\right)^J} \left(\prod_{j=1}^{J} \exp\left(-\frac{(a_j^{(r)} - m_j)^2}{2\sigma_j^2}\right) da_j\right) dI_0$$

Because $P_{\mathcal{H}}(\mathcal{H}_r = 1 \mid \{a_j^{(r)}\}, I_0) = 1$ only for $I_0 = \sum_{j=1}^{J} a_j^{(r)} \phi_j^{(r)}(x_0, y_0)$, after calculating the integral with respect to I_0, the expression for P_Z results to:

$$P_Z = \int_{a_1^{(r)}} \cdots \int_{a_J^{(r)}} \frac{1}{\Delta I} \frac{1}{\left(\sqrt{2\pi}\sigma\right)^{N+1}} \frac{1}{\left(\sqrt{2\pi}\sigma_j\right)^N} \times \left(\prod_{k=0}^{N} \exp\left(-\frac{\left(g_k - \left(\sum_{j=1}^{J} a_j^{(r)} \phi_j^{(r)}(x_k, y_k)\right)\right)^2}{2\sigma^2}\right)\right) \times \left(\prod_{j=1}^{J} \exp\left(-\frac{(a_j^{(r)} - m_j)^2}{2\sigma_j^2}\right) da_j\right) \quad (5)$$

The results of the integrals in equation (5) can be expressed in a closed form. The detailed calculation is given in (Quint, 1994). The integrals which appear in the denominator of equation (4) are calculated in a similar way. One finally obtains for the probability of homogeneity:

$$P(\mathcal{H}_r = 1 \mid \{g_k\}, g_0) = \frac{1}{f_{g_0}\sqrt{2\pi}\sigma} \frac{N^{\frac{3}{2}}\sqrt{\det \mathbf{C}}}{(N+1)^{\frac{3}{2}}\sqrt{\det \mathbf{C}^*}} \times \exp\left(\frac{N \det \mathbf{C_{ext}}}{2\sigma^2 \det \mathbf{C}} - \frac{(N+1) \det \mathbf{C^*_{ext}}}{2\sigma^2 \det \mathbf{C}^*}\right) \quad (6)$$

Being a probability, the values taken by expression (6) are in the domain: $P(\mathcal{H}_r = 1 \mid \{g_k\}, g_0) \in [0, 1]$.

In equation (6), the factor f_{g_0} is defined as:

$$f_{g_0} = \frac{1}{2} \operatorname{erf}\left(\frac{g_0 - I_{min}}{\sqrt{2}\sigma}\right) + \frac{1}{2} \operatorname{erf}\left(\frac{I_{max} - g_0}{\sqrt{2}\sigma}\right)$$

where $\operatorname{erf}(x)$ is the Gaussian error function:

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2)\, dt.$$

Using images with eight bits per pixel, the minimal and maximal intensity values are: $I_{min} = 0$ and $I_{max} = 255$.

The elements of the matrices appearing in equation (6) are given in Table 1. The matrix $\mathbf{C} = (c_{ij})$ is $J \times J$ and composed of the elements c_{ij}, $i, j = 1, ..., J$ given in Table 1. The matrix $\mathbf{C_{ext}} = (c_{ij})$ is $(J+1) \times (J+1)$. The upper left $J \times J$ submatrix of $\mathbf{C_{ext}}$ is identical with the matrix $\mathbf{C}$. Column and row $J+1$ respectively are composed of the elements $c_{j,J+1}$ given in Table 1. The matrices $\mathbf{C}^*$and $\mathbf{C^*_{ext}}$ are constructed in a similar way, but now the elements c_{ij}^* from Table 1 are used. All matrices are symmetrical. For computing the matrix elements c_{ij}, the summation has to be done over the product of the functions $\phi_i^{(r)}$ and $\phi_j^{(r)}$ at all pixel locations $(x_k, y_k), k = 1, ..., N$ already marked as belonging to region $\mathcal{R}_r$. In addition to this, for calculating c_{ij}^* the summation is extended over the pixel (x_0, y_0) for which predicate testing is under way.

4 COMPUTATIONAL ASPECTS

For calculating the probability of homogeneity for a region $\mathcal{R}_r$ and a pixel (x_0, y_0), partial knowledge of the regions model is necessary: the functions $\phi_j^{(r)}$ have to be known for the region. However, this does not assume, that these functions are the same for all possible regions of the image. The complete model of a region is given if one also knows the coefficients $a_j^{(r)}$ in the linear combination (2). These coefficients could be estimated from the gray values of the image. Since we are only interested in the segmentation of the image and not in the actual values of these coefficients, in our segmentation

$c_{jj} = \frac{1}{N}\sum_{k=1}^{K}\left(\phi_j^{(r)}(x_k,y_k)\right)^2 + \frac{\sigma^2}{N\sigma_j^2}$	$c^*_{jj} = \frac{1}{N+1}\sum_{k=0}^{K}\left(\phi_j^{(r)}(x_k,y_k)\right)^2 + \frac{\sigma^2}{(N+1)\sigma_j^2}$	$j = 1,...,J$
$c_{ij} = \frac{1}{N}\sum_{k=1}^{N}\phi_i^{(r)}(x_k,y_k)\,\phi_j^{(r)}(x_k,y_k)$	$c^*_{ij} = \frac{1}{N+1}\sum_{k=0}^{N}\phi_i^{(r)}(x_k,y_k)\,\phi_j^{(r)}(x_k,y_k)$	$i,j = 1,...,J$ $i \neq j$
$c_{j,J+1} = \frac{1}{N}\sum_{k=1}^{N}\phi_j^{(r)}(x_k,y_k)\,g_k + \frac{\sigma^2\,m_j}{N\sigma_j^2}$	$c^*_{j,J+1} = \frac{1}{N+1}\sum_{k=0}^{N}\phi_j^{(r)}(x_k,y_k)\,g_k + \frac{\sigma^2\,m_j}{(N+1)\sigma_j^2}$	$j = 1,...,J$
$c_{J+1,J+1} = \frac{1}{N}\sum_{k=1}^{N} g_k^2$	$c^*_{J+1,J+1} = \frac{1}{N+1}\sum_{k=0}^{N} g_k^2$	

Table 1: Definition of the matrix elements used in the calculation of the probability of homogeneity

procedure the estimation is not done explicitly. The calculation of the probability of homogeneity can be performed without knowing the actual values of the coefficients $a_j^{(r)}$ for the current region.

It is possible to calculate the matrix elements given in Table 1 iteratively. One observes that after a successful assignment of a pixel to a region, the matrix elements c^*_{ij} of the current step will become the matrix elements c_{ij} of the next step. Thus, for testing the homogeneity predicate, only the elements c^*_{ij} have to be calculated. These elements can be calculated iteratively using the elements c_{ij} from the previous step.

The computational complexity depends from the size of the model used, i.e. from the number J of functions used in the linear combination (2). The main effort spent in the calculation of the probability of homogeneity is for the calculation of the determinant of a $(J+1) \times (J+1)$ matrix. Due to iterative calculation the effort is independent from the size (number of pixels N) of a region.

5 A SIMPLE EXAMPLE

To illustrate the usage of our approach with a simple example, we assume that the image is composed part by part of planar surfaces. Although this is a limitation, it can be used with good approximation for range images of scenes with mostly planar surfaces or even for light intensity images of objects without textured surfaces. It is a reduction to the simplest case of a polygonal approximation of surfaces, which is often used in the segmentation of range images, see e.g. (Besl, 1988), (Silverman and Cooper, 1988).

In this case, the model functions for the regions to be segmented are:

$$\phi_1(x_k,y_k) = x_k$$
$$\phi_2(x_k,y_k) = y_k$$
$$\phi_3(x_k,y_k) = 1$$

and $J = 3$. According to the definitions given in Table 1 the matrices $\mathbf{C}$ and $\mathbf{C_{ext}}$ take the form:

$$\mathbf{C} = \begin{pmatrix} s_x & l_{xy} & m_x \\ l_{xy} & s_y & m_y \\ m_x & m_y & 1+\frac{\sigma^2}{N\sigma_3^2} \end{pmatrix}$$

$$\mathbf{C_{ext}} = \begin{pmatrix} s_x & l_{xy} & m_x & l_{xg} \\ l_{xy} & s_y & m_y & l_{yg} \\ m_x & m_y & 1+\frac{\sigma^2}{N\sigma_3^2} & m_g+\frac{m_3\sigma^2}{N\sigma_3^2} \\ l_{xg} & l_{yg} & m_g+\frac{m_3\sigma^2}{N\sigma_3^2} & s_g \end{pmatrix}$$

and the matrix elements are defined as follows:

$$s_x = \frac{1}{N}\sum_{k=1}^{K} x_k^2 + \frac{\sigma^2}{N\sigma_1^2} \qquad s_y = \frac{1}{N}\sum_{k=1}^{K} y_k^2 + \frac{\sigma^2}{N\sigma_2^2}$$

$$s_g = \frac{1}{N}\sum_{k=1}^{K} g_k^2 \qquad l_{xy} = \frac{1}{N}\sum_{k=1}^{N} x_k y_k$$

$$l_{xg} = \frac{1}{N}\sum_{k=1}^{N} x_k g_k + \frac{m_1\sigma^2}{N\sigma_1^2} \qquad l_{yg} = \frac{1}{N}\sum_{k=1}^{N} y_k g_k + \frac{m_2\sigma^2}{N\sigma_2^2}$$

$$m_x = \frac{1}{N}\sum_{k=1}^{N} x_k \qquad m_y = \frac{1}{N}\sum_{k=1}^{N} y_k$$

$$m_g = \frac{1}{N}\sum_{k=1}^{N} g_k.$$

The elements of the matrices $\mathbf{C}^*$ and $\mathbf{C^*_{ext}}$ are computed in a similar way. For these matrices, the summations in the definition of their elements start with index $k = 0$.

To test the sensitivity of our approach with respect to degradation with noise and with respect to violations of the model assumptions, we have used a synthetic image. In the upper left area of the image, the gray values rise from background level with a constant slope of two gray values per column until the middle column of the image. In continuation of the first region, the gray values fall in region 2 with the same slope until they reach background level. The gray values in region 3, which is situated in the middle of the image, violate the model assumptions: they depend upon a parabolic rule from their position. In region 4, situated in the bottom of the image, the gray values have a slope of 1.5 in both row and column direction. Within each region absolute gray value differences up to 200 occur.

The segmentation results for the synthetic image degraded with Gaussian white noise of different variances and for different parameter settings are given in (Landes, 1995). In Figure 2, the segmentation result for the synthetic image degraded with Gaussian white noise with the variance $\sigma_n^2 = 30$

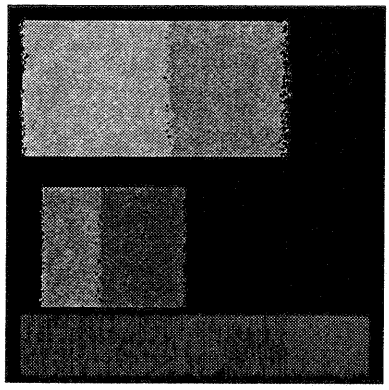

Figure 2: Segmentation result of the synthetic image degraded with Gaussian white noise with $\sigma_n^2 = 30$

is given. As model parameters we have used in this case: $m_1 = m_2 = 0$, $m_3 = 128$, $\sigma_1 = \sigma_2 = \sigma_3 = 3$ and $\sigma^2 = 30$. We have used the value $\delta = 0.8$ as the decision threshold in the homogeneity predicate testing.

Experiments have shown that the segmentation results for regions for which the correct model was chosen are good up to values for the standard deviation of the added noise which are three times higher than the gradient of the gray values within the region. At the left border of region 1 and the right border of region 2 there appear inaccuracies which are expected since the gray values of the two regions reach at these borders background level. Model violations, as shown with region 3 of the synthetic image, are partly tolerated.

It is mainly the parameter σ^2 which controls the amount of noise or model violation tolerated by the segmentation algorithm. For optimality, this parameter should be chosen equal to the actual noise variance in the image. Choosing this parameter smaller than the variance of the actual noise results in a segmented image containing many single points rejected by the algorithm. However, these points could be eliminated in a following stage by morphological operations. Choosing this parameter bigger than the variance of the actual noise is more critical since in this case different regions could be merged in the segmented image.

6 AERIAL IMAGE SEGMENTATION

We are using the homogeneity predicate described in this article for the segmentation of colour aerial images. The regions gained this way are used together with line segments as primitives in our model based aerial image understanding system MOSES (Quint and Sties, 1995).

As a control algorithm for the segmentation process a region growing scheme is used. The process starts with a set of initial seed regions. For all regions, pixels are sought which are neighbour to at least one region and which are not yet marked as belonging to a region. The probability of homogeneity is calculated for these pixels and each of their neighbouring regions. If this probability exceeds the decision threshold δ, the pixel is marked as belonging to the corresponding region. If the initial regions cannot be extended any longer new seed regions are chosen in areas with small gray value differences.

The digital images used in our project are acquired by scanning aerial colour photographies and have a raster size of $30\,cm \times 30\,cm$ on the ground. The image in Figure 3 shows a part of the campus of University of Karlsruhe. The German Topographic Base Map 1:5000 which is available in digital form is used to gain estimates for the positions of initial seed regions. In order to obtain stable values in the calculation of the probability of homogeneity, seed regions should have a minimum size. Experiments have shown, that an initial region size of 5×5 pixels is suitable.

Map knowledge is also used to choose the model for a given region according to the known class of the objects. For the segmentation of the image in Figure 3 we have used two types of models: the planar model presented in section 5 for regions corresponding to buildings, parking areas and streets, and a Markov Random Field (MRF) model for wood and grass regions. MRF approaches already have been used in previous work (see e.g. (Cohen and Fan, 1992), (Herlin et al., 1994)) for the segmentation of textured surfaces.

In our approach we use a second order MRF model:

$$\sum_{l=-1}^{1} \sum_{m=-1}^{1} a_{lm} \left(I(x_k - l, y_k - m) - \mu_k \right) = 0.$$

Since the light intensities I_k are unmeasurable they are replaced with the gray values at the corresponding pixel location. Hence, the model functions $\phi_j(x_k, y_k)$ in equation (1) are:

$$\phi_j(x_k, y_k) = g(x_k - l, y_k - m) - \mu_k$$

with $l, m \in \{-1, 0, 1\}$ excepting the pair $(l, m) = (0, 0)$. There are eight model functions and thus for the probability of homogeneity determinants of 8×8 and 9×9 matrices have to be calculated. For the parameter μ_k we use the local mean of the gray values in the neighbourhood. The variance σ^2 of the noise in the three channels of the images is estimated using the method described in (Brügelmann and Förstner, 1992).

For each channel we calculate the corresponding probability of homogeneity. The value used for testing the homogeneity predicate is obtained in analogy to the law of total probability as a linear combination of the three probabilities of homogeneity. The factors in this linear combination are chosen inverse proportional to the variance of the noise in the corresponding channel.

Figure 4 gives the segmentation result of the aerial image of Figure 3. Pixels belonging to the same region are marked in Figure 4 with the same gray value. As a decision threshold the value $\delta = 0.8$ was used. A number of 14 initial seed regions were extracted from the map. After our segmentation the image was divided in 86 regions. As one can observe, man made objects like buildings, streets and walking ways, for which the planar model was used, are segmented with good accuracy. The MRF model provided good results in the area with regular planted trees in the lower left corner of the image, but difficulties arise in the wood area in the upper part of the image. The gray values in this area are very inhomogeneous and cannot be represented by the used model. As a result, the wood area was splitten into several regions.

7 SUMMARY AND CONCLUSION

Our approach for a homogeneity predicate is based on the a-posteriori probability for a pixel to fulfill the model assumptions for a region. For some practical relevant models (polygonal surfaces, MRF models) we have derived a closed formula to calculate the probability of homogeneity. Using this

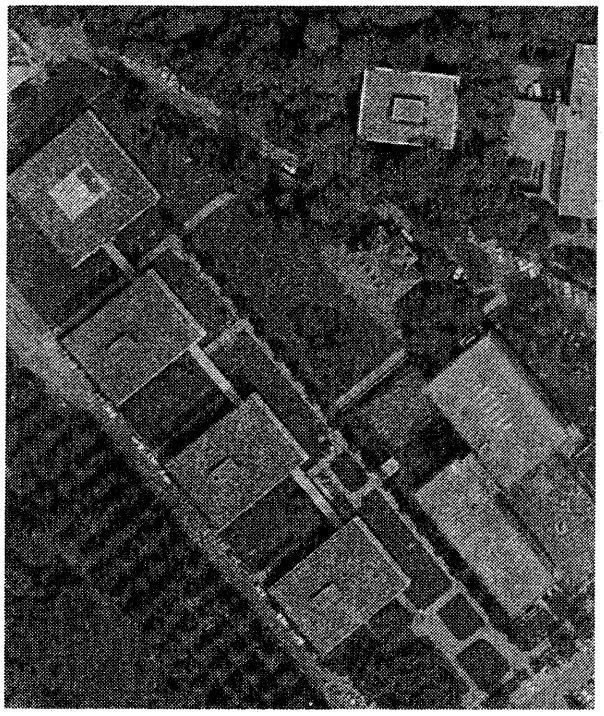

Figure 3: Aerial image

Figure 4: Segmentation result

formula, the computational effort depends only from the size of the model and is independent from the size of the segmented region.

The homogeneity predicate is used in a region growing scheme, but it can also be used in other control algorithms for image segmentation or clustering. Experiments with synthetical images have shown, that the most important parameter of our approach is the variance of the noise in the image. For segmenting aerial images, this variance is estimated using an algorithm from the literature. Initial seed regions and the model type to use is extracted from map data. The segmentation results are good for non-textured areas and for areas with regular texture. For irregular textured surfaces experiments with higher order MRF-models will be performed.

ACKNOWLEDGMENT

This work is funded by the Deutsche Forschungsgemeinschaft (DFG).

References

Besl, P. (1988). *Surfaces in range image understanding.* Springer, New York.

Brügelmann, R. and Förstner, W. (1992). Noise estimation for color edge extraction. In Förstner, W. and Ruwiedel, S., editors, *Robust computer vision*, pages 90–106. Wichmann, Karlsruhe.

Cohen, F. and Fan, Z. (1992). Maximum likelihood unsupervised textured image segmentation. *Computer Vision, Graphics and Image Processing*, 54:239–251.

Haralick, R. M. and Shapiro, L. G. (1985). Survey, image segmentation techniques. *Computer Vision, Graphics and Image Processing*, 29:100–132.

Herlin, I., Bereziat, D., Giraudon, G., Nguyen, C., and Graffigne, C. (1994). Segmentation of echocardiographic images with Markov random fields. In Eklundh, J.-O., editor, *Computer Vision – ECCV '94*, pages 201–206, Berlin. Springer.

Landes, S. (1995). Entwicklung eines Flächenwachstumsverfahrens zur Segmentierung von Luftbildern. Master's thesis, IPF, Universität Karlsruhe.

LaValle, S. and Hutchinson, S. (1995). A Bayesian segmentation methodology for parametric image models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 17:211–218.

Pal, N. R. and Pal, S. K. (1993). A Review on Image Segmentation Techniques. *Pattern Recognition*, 26(9):1277–1294.

Pavlidis, T. (1977). *Structural Pattern Recognition.* Springer, Berlin.

Pearl, J. (1986). Fusion, propagation and structuring in belief networks. *Artificial Intelligence*, 29:241–288.

Quint, F. (1994). Bildsegmentierung mit einem Bayesschen Ansatz. Technischer Bericht IPF-FQ-9/94, IPF, Universität Karlsruhe.

Quint, F. and Sties, M. (1995). Map–based semantic modeling for the extraction of objects from aerial images. In Grün, A., Kübler, O., and Agouris, P., editors, *Automatic Extraction of Man-Made Objects from Aerial and Space Images*, pages 307–316. Birkhäuser, Basel.

Silverman, J. and Cooper, D. (1988). Bayesian clustering for unsupervised estimation of surface and texture models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 10:482–495.