

# A Mixed Markov Model for Change Detection in Aerial Photos with Large Time Differences

Csaba Benedek<sup>1,2</sup> and Tamás Szirányi<sup>1</sup>

<sup>1</sup>Distributed Events Analysis Research Group, Computer and Automation Research Institute  
H-1111, Budapest, Kende utca 13-17, Hungary, bcsaba@sztaki.hu, sziranyi@sztaki.hu

<sup>2</sup>Currently with Ariana Project-Team INRIA/CNRS/UNSA, B.P. 93, 06902 Sophia Antipolis, France

## Abstract

*In the paper we propose a novel multi-layer Mixed Markov model for detecting relevant changes in registered aerial images taken with significant time differences. The introduced approach combines global intensity statistics with local correlation and contrast features. A global energy optimization process simultaneously ensures optimal local feature selection and smooth, observation-consistent classification. Validation is given on real aerial photos.*

## 1 Introduction

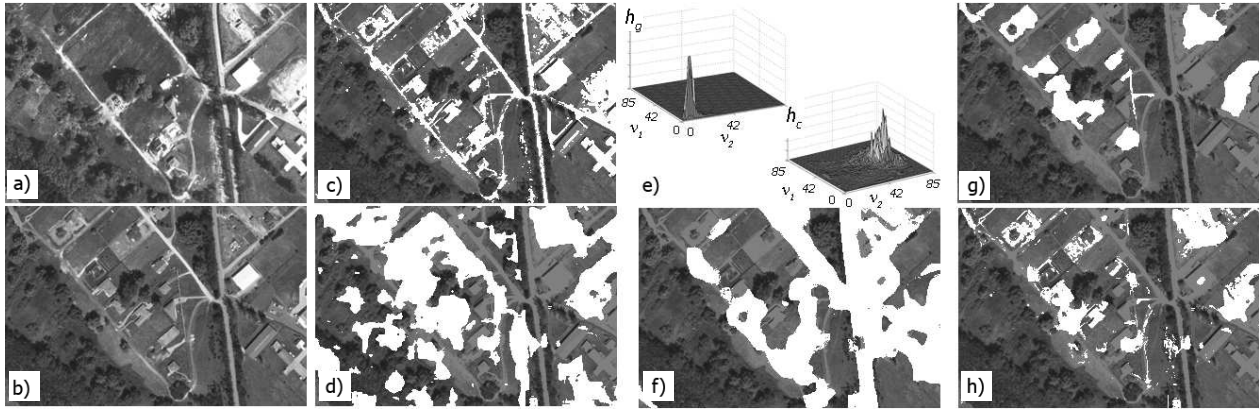
Automatic evaluation of aerial photo repositories is an important field of research, since periodically repeated manual processing is time-consuming and cumbersome in cases of high number of images, and dynamically changing content. The paper deals with change detection in aerial images taken with many years time difference, probably in different seasons and in different lighting conditions. Therefore, 'simple frame differencing' [2, 7] techniques cannot be used, since the observed pixel levels may be significantly different even in the unchanged areas. We focus on standard optical images, unlike [4], which uses multi temporal SAR imagery exploiting its insensitivity to atmospheric conditions. We only assume that the database contains registered orthophotos. On the other hand, most previous models monitor purely natural [6] or urban [8] territories, or they are dedicated to a specific task like detecting new built-in areas [1] or destructions due to earthquakes [5]. As Fig. 1 and 3 show, the current photos include both built-in and unpopulated regions, including forests, fields and agricultural lands as well, presenting various types of differences. Our goal is detecting changes which are 'unusual' in a statistical manner. In a related PCA-based model [8], the authors assumed that the 'unimportant' differences are caused by alteration of illumination and camera settings. Since the

above effects influence the observed sensor values in a multiplicative or additive fashion, they modelled the relationship of the corresponding pixel levels within the unchanged regions by a globally constant linear transform. Similar approaches can be also found in [7]. However, these models disregard that the scene may 'regularly' alter as well, primarily due to the seasonal vegetation changes. Moreover, in agricultural areas which follow crop rotation, the shape and arrangement of the neighboring tracks of a plough-land may be changed significantly. We will show that the regularity of these changes also can be measured in a statistical way, although they may cause significant deviations from the estimated linear approach.

In the paper we propose a robust multi-layer Mixed Markov model [3] to tackle the above change detection problem. We identify the changes through complementary features: global intensity statistics and local correlation. A contrast based selection process is responsible for locally choosing the more reliable feature in the different image regions, while a smooth change map is ensured using local connectivity constraints.

## 2 Image model and feature extraction

Let  $G_1$  and  $G_2$  be the two registered images which we wish to compare.  $G_1$  and  $G_2$  have an identical pixel lattice  $S$ . The gray values are denoted by  $g_1(s)$  and  $g_2(s)$  for a pixel  $s \in S$  of  $G_1$  and  $G_2$ , respectively. Our first task is to extract local features at each  $s \in S$  which give us information for classifying  $s$  as a *changed* (ch) or *background* (bg) i.e. unchanged surface point. Taking a probabilistic approach, we consider the ch/bg classes as random processes generating the features according to different distributions. We start our investigations in the joint intensity domain of the two images. Here, instead of prescribing a global linear transform between  $g_1(s)$  and  $g_2(s)$  for the background areas [8], we give a multi modal description of the observed data. We approximate the 2-D histogram of the



**Figure 1. Feature selection: a)  $G_1$  image, b)  $G_2$  image, c) intensity based change detection, d) correlation based change detection, e)  $h_g$  and  $h_c$  histograms f) local contrast based segmentation, g) ground truth h) change detection with combined features, without considering local connectivity**

$\bar{g}(s) = [g_1(s), g_2(s)]^T$  vectors by a mixture of Gaussians distribution. In this way, we measure which intensity values occur often together in the two images. Thereafter, the probability of the  $\bar{g}(s)$  observation in the background is calculated as:  $P(\bar{g}(s)|bg) = \sum_{i=1}^K \kappa_i \cdot \eta(\bar{g}(s), \bar{\mu}_i, \Sigma_i)$ , where  $\eta(\cdot)$  denotes a two dimensional Gaussian density function with  $\bar{\mu}_i$  mean vector and  $\Sigma_i$  covariance matrix, while the  $\kappa_i$  terms are positive weighting factors. Using a fixed  $K = 5$ , the distribution parameters are estimated automatically by the conventional EM algorithm. On the other hand, any  $\bar{g}(s)$  value may occur in the changed regions, hence the ‘ch’ class is modeled by a uniform density:  $P(\bar{g}(s)|ch) = u$ . Next, we demonstrate the limitations of this feature. We derive the segmentation in Fig. 1(c) as the maximum likelihood (ML) estimate, where the label of  $s$  is  $\phi_g(s) = \arg\max_{\psi \in \{ch, bg\}} P(\bar{g}(s)|\psi)$ . One can observe that this multi-Gaussian intensity based approach (MGI) erroneously marks several unaltered regions as changes compared to the ground truth [Fig. 1(g)]. However, the misclassifications are mainly limited to highly textured regions (e.g. buildings and roads) since the  $\bar{g}(s)$  gray values occurring there are less frequent in the global image statistics.

We derive the next feature,  $c(s)$  as the correlation between the rectangular  $v \times v$  neighborhoods of  $s$  in  $G_1$  and in  $G_2$  (used  $v = 17$ ). Pixels with higher  $c(s)$  values lie more likely in unchanged image regions. Our experiments showed that the  $P(c(s)|bg)$  and  $P(c(s)|ch)$  probabilities can be approximated by different Gaussian distributions. Note that in itself, a simple ML classification based on  $c(\cdot)$  results in a fairly poor segmentation  $\phi_c$  [see Fig. 1(d)]. However, we can see that  $\bar{g}(s)$  and  $c(s)$  are efficient complementary features. In low contrasted image regions, where the noisy  $c(s)$  may be irrelevant, the decision based on  $\bar{g}(s)$  is reliable. In

textured areas one should choose  $c(s)$  instead of  $\bar{g}(s)$ .

We formulate the *contrast based feature selection* in a probabilistic manner. Let  $\nu_i(s)$  be the local contrast at  $s$  in  $G_i$  ( $i = 1, 2$ ), that is measured by the variance of the gray levels in the neighborhood of  $s$ . Let be  $\bar{\nu}(s) = [\nu_1(s), \nu_2(s)]^T$ . Using a few manually segmented training images, one can derive a 2-D histogram  $h_g$  which statistically measures the reliability of the decision based on  $\bar{g}(s)$  as a function of  $\bar{\nu}(s)$ . We calculate  $h_c$  in a similar manner, for the  $c(\cdot)$  feature. As Fig. 1(e) shows, the normalized  $h_g$  and  $h_c$  histograms can be efficiently approximated by 2-D Gaussian density functions:  $P(\bar{\nu}(s)|h_\psi) = \eta(\bar{\nu}(s), \bar{\mu}_\psi, \Sigma_\psi)$ ,  $\psi \in \{g, c\}$ . Thereafter, we create a contrast map as  $\phi_\nu(s) = \arg\max_{\psi \in \{g, c\}} P(\bar{\nu}(s)|h_\psi)$  [see Fig. 1(f): class ‘c’ marked with white], and the combined map  $\phi_*$ , where  $\phi_*(s) = \phi_g(s)$  if  $\phi_\nu(s) = ‘g’$ ;  $\phi_*(s) = \phi_c(s)$  if  $\phi_\nu(s) = ‘c’$  [Fig. 1(h)]. Let observe that  $\phi_*$  is a more improved approximation of the ground truth, however, it is still quite noisy. Therefore, based on the above investigations, we introduce a robust segmentation model in the following.

### 3 A Mixed Markov Random Field image segmentation model

Mixed Markov models [3] extend the modeling capabilities of Markov random fields: they enable using both static and observation-dependent dynamic links between the processing nodes. We can take here the advantage of this property, since the  $\bar{\nu}(s)$  feature plays a particular role: it may locally switch ON and OFF the  $\bar{g}(s)$  respectively  $c(s)$  features into the integration procedure. We consider our task as a composition of four interactive segmentation processes [analogously to Fig. 1(c), (d), (f) and (h)]. Thus we map the



