MULTITEMPORAL INTERPRETATION OF REMOTE SENSING DATA

Sönke Müller[a,*], Guilherme Lucio Abelha Mota[b] and Claus-Eberhard Liedtke[a]

[a]Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover, Germany - {mueller, liedtke}@tnt.uni-hannover.de, [*]corresponding author
[b]Department of Electrical Engineering, Pontifical Catholic University of Rio de Janeiro, Brazil - guimota@ele.puc-rio.br

KEY WORDS: Multitemporal, Land Use, Classification, Urban, Aerial, Knowledge Base.

ABSTRACT

The automated interpretation of aerial image data is a task with increasing significance for several applications, e.g. quality control and automatic updating of GIS data, automatic land use change detection, measurement of sealed areas for public authority uses, monitoring of land erosion etc. The use of additional sensors could improve the performance of the automated classification; however, because of additional costs or simple unavailability of data, this approach should be avoided. One possibility to stabilize an automatic image analysis is using remote sensing data of the same region of different dates that is often existing. This paper presents a method how a monotemporal knowledge representation can be expanded by a temporal component to take advantage of previous classifications of the same scene and knowledge about the time dependency of the object classes. The present approach proposes the combination of a semantic network, representing the generic description of the scene, and a state transition diagram, modeling the possible state transitions for each one of the classes of interest. The probabilities of the state transition diagram are introduced as a priori knowledge in a statistical classification procedure. Experimental results from a series of three aerial images from 1983 up to 2001 of a suburban region near Hannover are shown in order to illustrate the potential of the proposed multitemporal approach.

1 INTRODUCTION

The quality of geodata refers to the geometric and semantic correctness as well as to its up-to-dateness. Among these features, most part of the effort is devoted to keep the consistence between the geodata and the respective area. In order to achieve this aim, it is desirable to develop an image processing system that is able to generate automatically up-to-date geodata.

The quality of the outcome of an automatic image analysis depends on the used input data and on the knowledge about the investigated scene. In the present approach, actual aerial or satellite images are the standard input for the scene interpretation. In many cases, the increment of the cost or unavailability hinder the utilization of additional sensor data. Moreover, the fact that aerial and satellite images are produced in standardized intervals and quality permits the system to work on data acquired in different time instances.

In this paper we restrict on aerial images with a resolution of $0,3125\frac{m}{pixel}$ and differentiate the following object classes :

- Inhabited area
- Forest
- Agriculture

Figure 1 presents example input data. These images were acquired in 1983, 1988 and 2001. While the images of 1983 and 1988 are gray scale, the image of 2001 is originally colored. In order to standardize these data, the image of 2001 was converted to gray scale.

In this paper, the knowledge based system GEOAIDA is employed. The system was developed to interpret a scene considering aerial photographs or other raster data. The original conception of GEOAIDA (Bückner et al., 2002) aims at interpreting remote sensing data by exploiting an exclusively hierarchical description of the problem given by a semantic network. The system is being extended actually in order to incorporate features to manipulate temporal data.

Inside GEOAIDA, the interpretation of remote sensing data from a given scene aims at finding out its structural and pictorial descriptions. The structural description has the same structure of the semantic network and is bound to the pictorial description. This approach allows simultaneous access to information about the object type, the geocoordinates and all other attributes calculated during the analysis.

The temporal approach proposed is based on (Pakzad, 2002) which employs a transition graph to describe the temporal dependencies between the classes of interest. Such strategy enables the user to formulate temporal a priori knowledge and to use it during the automatic analysis in connection to an older classification of the scene. Thus, in the present paper, besides the structural knowledge, knowledge about temporal dependencies is exploited to refine decisions in the interpretation process. The temporal knowledge is used in a statistical classification process and directly influence the interpretation result. In a previous work (Müller et al., 2003) the usage of temporal knowledge was limited to generate hypothesis for possible new states of region, without the usage of real probabilities.

The remaining part of this paper is organized as follows. Section 2 describes briefly the approach. Section 3 presents the classification strategy. Section 4 presents the experimental results and section 5 the conclusions.

2 PROPOSED APPROACH

The proposed approach is based on the knowledge based image interpretation system GEOAIDA (Bückner et al., 2002) developed at the Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, University of Hannover. In GEOAIDA,

Figure 1: Registered Input Images (1983, 1988, 2001)

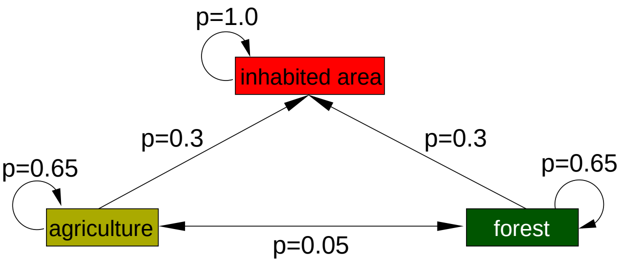

Figure 2: State Transition Diagram $t \rightarrow t+\Delta t$, $p \,\widehat{=}$ Probability for State Transition

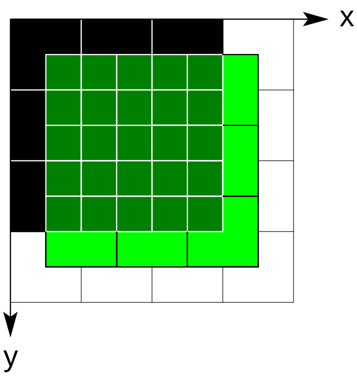

Figure 4: Segments for Classification

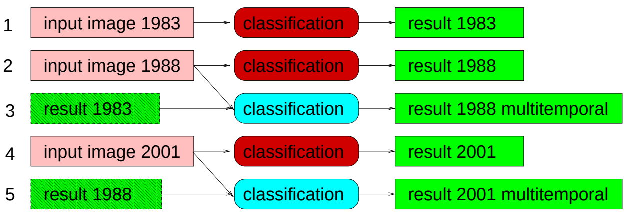

Figure 3: Data Flow Diagram

the structure of the scene to be interpreted is modeled in a semantic network, allowing an effective hypothesis handling. The used state transition diagram is shown in figure 2. The arrows indicate assumed possible state transitions for regions between two images of different dates and its respective probabilities are placed near the arrows. Here inhabited area always stays inhabited area, the demolition of buildings and change to agriculture is very unusual. Changes from vegetation class to inhabited area are possible with probability of $p = 30\%$. Additionally, changes from forest to agriculture and from agriculture to forest are possible with probability of $p = 5\%$.

In figure 3 the data flow is diagramed. First an aerial image of 1983 is processed in the normal way, without any multitemporal knowledge. For the next step the interpretation result of 1983 is used as a priori temporal knowledge for the interpretation of the aerial image of 1988. The same is done with the input image of 2001. To assess the contribution of the multitemporal knowledge, the images of 1988 and 2001 are also interpreted without temporal knowledge.

3 CLASSIFICATION

Two different approaches are followed to classify an input image. A region-based classifier works on segments of the image and classifies these segments into the considered 3 classes inhabited area, forest and agriculture. For an assumed application of a GIS verification system, the result is combined with a structural operator (see 3.2). The structural classifier searches directly for buildings that indicate the class inhabited area in the image. In the following section a description of both approaches is given.

3.1 Region-based Classifier

The region-based classification operator starts with a segmentation of the entire image into segments of predetermined size. The size of a segment is chosen equivalent to that of an average house including a small garden. For each segment features are calculated that are basis for the following linear regression classifier. The linear regression classifier uses a priori probabilities from the state transition diagram (see figure 2) with use of a previous classification. To get a more reliable interpretation the classification is repeated on a shifted segmentation of the image (see figure 4). For the black and green illustrated segments hypothesis are generated that overlap half of a segment size. The emerging conflicts for one segment are solved by GEOAIDA, the most probable segment is taken as instance. The classification procedure is described in the next two sections.

3.1.1 Feature Extraction Aerial images of different time are subjected to different illumination conditions and camera parameters. The input images (see figure 1) are preprocessed by a contrast stretch algorithm that unifies the color distribution in the histogram for the following feature extraction operators.

The used classifier uses four features that are calculated for each segment:

- A measurement for the shadiness,
- a measurement for the uniformity,
- a measurement for the contour angularity and
- a measurement for straight contour lines.

The shadiness is calculated by use of the grey value histogram see in figure 5. The threshold t in equation 1 is defined as the first local minimum in the grey value histogram $H(i)$, where i is the grey value. The measurement is the ratio of shade pixels to all pixels in the image. Buildings cause a characteristic high shadiness in inhabited areas regions.

$$shadiness = \frac{\sum_{i=0}^{t} H(i)}{\sum_{i=0}^{255} H(i)} \quad (1)$$

The uniformity is calculated by a mean m_x of the local variances σ_x^2 of the image grey values. In figure 5 a contour image is shown that is calculated from the local variance matrix with a threshold decision.

$$uniformity = m_x(\sigma_x^2) \quad (2)$$

The calculation of a measurement of angularity is based on a linear Hough transformation for straight lines (Duda and Hart, 1972) of the contour image (cp. Hough space in figure 5). A straight line forms a local maximum in the Hough space, orthogonal lines form maxima in the Hough space. The values of the Hough space are totalized over the angle and illustrated as a histogram in figure 5. The discrete Fourier transformation (DFT) is calculated in the next step and the $DFT(k = 2)$ is taken as a measurement for the angularity, because vertical lines have an angle distance of 90° in the Hough space.

$$angularity = \frac{DFT(k = 2)}{imagesize^2} \quad (3)$$

The calculation of a measurement of straight contour lines $MSCL$ is done by searching local maxima in the Hough space $HS(x, y)$ because straight lines shape a local maximum in the Hough space. The results of the region based classifier are shown in figure 7 and discussed in chapter 4. The results for the years 1988 and 2001 are achieved in two ways with and without use of a previous classification.

$$MSCL = \frac{\sum_{localmax} HS(x, y)}{(size(HS(x, y)))^2} \quad (4)$$

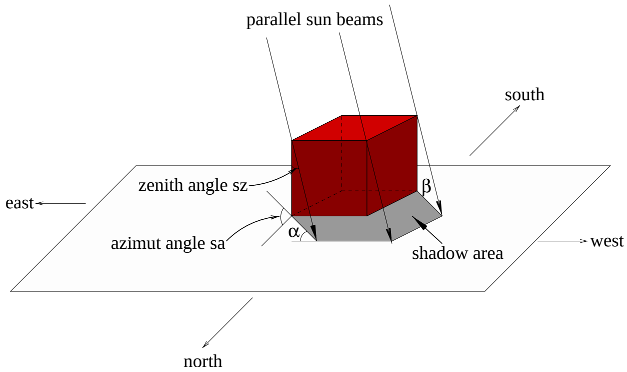

Figure 6: Illumination Model for Buildings

3.1.2 Linear Regression Classification The features of each region are combined in a feature vector that is basis for the linear regression classification algorithm method described in (Meyer-Brötz and Schürmann, 1970). The method was implemented by using the scilab (Gomez, 1998) environment. The a priori probabilities are taken from the state transition diagram considering the previous classification. They are empirical set, because the random sample was limited. Result of the classifier is a vector with probability for each differentiated class. GEOAIDA decides for the most probable class for each segment. Here both segmentations are taken into account.

3.2 Structural Classifier

The structural building extraction operator models buildings as complex structures consisting of different parts (cp. (Müller et al., 2003)). It assumes an illumination model shown in figure 6. The angles α and β are calculated from the exact date and time of the image capture and the sun angle. Hypotheses for shades and roofs are generated using two different image segmentation operators. To get the buildings, the roofs are grouped with one or more shades. The neighborhood relations regard the illumination model presented in figure 6.

Shades of buildings are derived with a threshold decision in the image. The threshold can be calculated automatically from the histogram, so that images with different illumination can be processed. Since shades are generally not visible in a green color channel, the green color has been masked during shade detection. Pixel taken as shade have to fulfill the condition $grey\ value < threshold$ and $hue < 90^\circ \vee hue > 150^\circ$.

Roofs are generated in a more complex procedure. Here the so-called color structure code (Priese et al., 1994), (Rehrmann and Priese, 1998) is used to segment the entire image. Additionally greenish areas are masked and roofs are accepted only in the other parts. An additional size criterion restricts acceptable roof hypotheses considering the size.

Shades generated by buildings have a limited area, so for example shade near a forest can be excluded. The compactness and orthogonality of roof labels is additionally measured to validate buildings. The grouping of shades and roof labels leads to validated buildings. The neighboring position of a shade to a roof has to fulfill the illumination model. Sometimes it is not possible to differentiate between the roof of a building and for example an adjacent parking area. In that case the expected size for a building is exceeded, does not fulfill the model, and the grouping is rejected during the analysis.

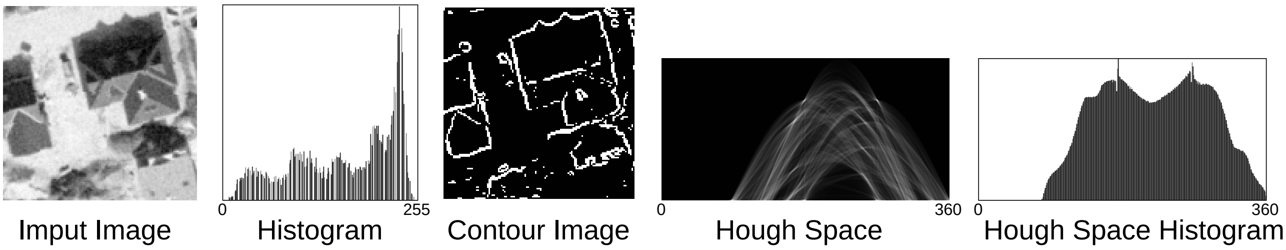

Figure 5: Intermediate Results for a Cut of an Input Image of 1988

Table 1: Results of Building Detection

Number	Evaluated Result	Detection Percentage	Branch Factor
1	1983	75.3%	42.9%
2	1988	86.2%	36.2%
3	1988 multitemporal	82.7%	36.0%
4	2001	96.8%	37.9%
5	2001 multitemporal	97.2%	41.4%

4 RESULTS

To evaluate the multitemporal approach a building detection tool for a GIS verification application was assumed. For this task the region-based and structural classifier were combined. Only detected building hypothesis that overlap with the class inhabited in the region-based classification result are taken as a building (cp. figure 7). The others are unconsidered for the evaluation.

The evaluation is based on manually segmented buildings in the input images of 1983 to 2001 (see manually detected buildings in the input image of 2001 in figure 7). Two measurements for a detection evaluation described in (Lin and Nevatia, 1998) were made:

$$detection\ percentage = \frac{100 \cdot TP}{TP + TN} \tag{5}$$

$$branch\ factor = \frac{100 \cdot FP}{TP + FP} \tag{6}$$

The two measurements are calculated by making a comparison of the manually detected buildings and the automatic results, where TP (true positive) is a building detected by both a person and GEOAIDA , FP (false positive) is a building detected by GEOAIDA but not a person, and TN (true negative) is a building detected by a person but not by GEOAIDA . A building is considered detected if the main part (min 50%) of the building is detected; an alternative could be to require that a certain fraction of the building is detected.

5 CONCLUSIONS

The developed approach shows how temporal knowledge can be used in an automatic image interpretation system. Temporal knowledge is modeled in a state transition diagram, the probabilities for state transitions are used as a priori knowledge for a linear regression classifier.

The approach was tested on a dataset containing aerial images acquired in 1983, 1988 and 2001. Three object classes are differentiated: Inhabited area, forest and agriculture. A region-based linear regression classifier uses features like shadiness, uniformity, contour angularity and straight contour lines to interpret the images.

For the images of 1988 and 2001 temporal knowledge in terms of a previous classification and a state transition diagram was used. Both images were also processed without temporal knowledge to compare the results.

To evaluate the multitemporal approach a building detection tool for a GIS verification application was assumed. The results in table 1 show that the proposed multitemporal approach is applicable. The multitemporal result of 2001 shows less confusion between the classes agriculture and forest than the monotemporal result. Additional tests are necessary to measure the advantage of a multitemporal approach in comparison with a monotemporal.

The multitemporal approach was tested in the focus of GIS verification, other possible applications are the detection of alteration, environmental studies, the development of urban areas and the examination of natural disasters.

REFERENCES

Bückner, J., Pahl, M., Stahlhut, O. and Liedtke, C.-E., 2002. A knowledge-based system for context dependent evaluation of remote sensing data. In: L. J. V. Gool (ed.), DAGM-Symposium, Lecture Notes in Computer Science, Vol. 2449, Springer, Zurich, Switzerland.

Duda, R. O. and Hart, P. E., 1972. Use of the hough transformation to detect lines and curves in pictures. Commun. ACM 15(1), pp. 11–15.

Gomez, C., 1998. Engineering and Scientific Computing with Scilab. Birkhauser Boston.

Lin, C. and Nevatia, R., 1998. Building detection and description from a single intensity image. Computer Vision and Image Understanding: CVIU 72(2), pp. 101–121.

Meyer-Brötz, G. and Schürmann, J., 1970. Methoden der automatischen Zeichenerkennung. K. Becker-Berke and R. Herschel, R. Oldenbourg Verlag, München-Wien.

Müller, S., Weis, M., Liedtke, C.-E. and Pahl, M., 2003. Automatic quality surveillance of GIS data with GEOAIDA. In: Proc. PIA 2003 - ISPRS Conference on Photogrammetric Image Analysis, Munich, Germany.

Pakzad, K., 2002. Knowledge based multitemporal interpretation. In: Proc. ISPRS2002, Graz, Austria.

Priese, L., Klieber, J., Lakmann, R., Rehrmann, V. and Schian, R., 1994. New results on traffic sign recognition. In: Proceedings of the Intelligent Vehicles Symposium, pages 249–254. IEEE, 1994.

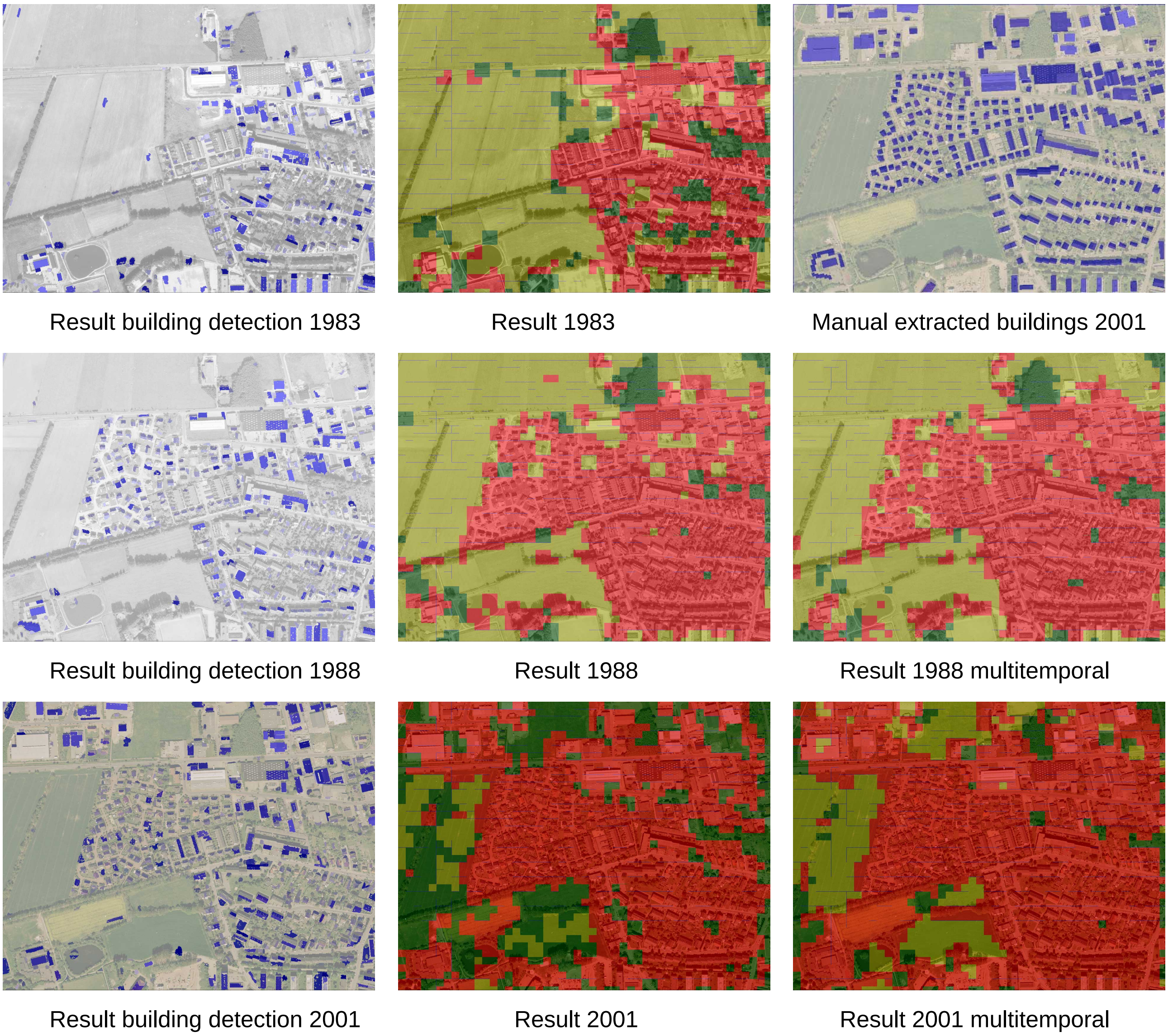

Figure 7: *Left column:* Result of Building Detection Algorithm (1983, 1988, 2001); *Middle Column:* Region Based Classification Monotemporal (1983, 1988, 2001); *Right Column:* Manual Extracted Buildings of 2001 and Region Based Classification Multitemporal (1988, 2001) - Overlays with Input Images

Rehrmann, V. and Priese, L., 1998. Fast and robust segmentation of natural color scenes. In: ACCV (1), pp. 598–606.