UTILIZING HIGH-LEVEL KNOWLEDGE IN MIDDLE-LEVEL IMAGE ANALYSIS

Yaonan Zhang

Faculty of Geodesy, Delft University of Technology
Thijsseweg 11, 2629 JA Delft, The Netherlands

Commission III

ABSTRACT:

This article proposes a unified architecture for image analysis, which enable us to: 1) integrate high-level knowledge in image segmentation; 2) use structural information for stereo matching in image/object dual spaces; 3) integrate image segmentation with stereo matching; 4) combine the edge-based and region-based segmentation. In order to design an integrated image analysis system, we must solve the theoretic problem on how to combine the knowledge from different sources (e.g. image intensity, object shape, structural information), as well as the implementation problem (e.g. how many layers or modules should be used, how to negotiate the objectives with each module and how to control the system, etc.). Some of the questions have been answered in this paper and some proposals have been made to solve the remaining problems.

KEYWORDS : Image Analysis, Image Processing, Machine Vision.

1. INTRODUCTION

The tasks of computational vision often rely on solving the following problem: from one to more images of a scene, derive an accurate geometric description of the scene and quantitatively recover the properties of the scene that are relevant to the given task. This problem (referred as middle-level image analysis) is hard because of several reasons [Aloimonos]:

1). During the image formation process the three-dimensional world is mapped into two dimensions, and one dimension is lost. This create many problems when we try to solve the inverse (ill-posed) problem of recovering the world from the image.
2). Even well posed (or regularized) visual computations are often numerically unstable, if noise is present in both the scene and the image. As a result, many problems which theoretically have unique solutions become very unstable in the presence of input noise.
3). Visual objects are hard to define. Object modelling techniques have been developed in the artificial intelligence and computer graphics to represent the 3-D objects, but it is still very difficult to use these techniques to describe a variety of natural objects.

Model-based vision allows high-level knowledge of the shape and appearance of specific objects to be used during the process of visual interpretation. Reliable identification can be made by identifying consistent partial matches between the models and features extracted from the image, thereby allowing the system to make inferences about the scene that go beyond what is explicitly available from the image. By providing this link between perception and high-level knowledge of the components of the scene, model-based recognition is an essential component of most potential application of vision.

This paper proposes an integrated architecture for image analysis and addresses the problems on how to integrate the information from different sources to improve the performance of objectives associated with middle-level analysis. The scheme presented in this paper is the combination and extension of the work described in author's other papers [Zhang,91a,91b,92a,92b], which mainly focus on the image segmentation and stereo matching. The paradigm in this paper would allow us to integrate a variety sources of knowledge and different kind of techniques. Under such scheme, we want to carry out the following integrations:

1) integrate high-level knowledge into image segmentation.
2) use structural information for stereo matching.
3) integrate image segmentation and stereo matching.
4) combine the edge-based and region-based segmentation.

We first in section 2 present an architecture on the integrated image analysis. In section 3 we examine several principles or criterions from probability and information theory on the possibility as a unified measurement to combine the information from different sources. Finally we in section 4 point out several applications based on the system we proposed.

2. AN INTEGRATED ARCHITECTURE FOR IMAGE ANALYSIS

In Fig.1 (next page), an integrated architecture for image analysis is proposed. In the following, the system is explanted in detail.

Description of each layer

- original image. Raster images are the most common input for the image analysis, which can be in format of binary (2-valued), grey, or in multispectral forms. In our research, we only deal with 2-D images, not with 3-D images, such as range images and medical images. Each pixel on 2-D image is indexed from left to right and from top to bottom.

- segmented image. In order to interpret a 2-D image, the image is first partitioned into regions, and each region is uniform and homogeneous with respect to some criterions. For the purpose of incoperating the high level knowledge into the segmentation, an adequate data structure should be designed to represent the segmented image, which should fulfil the requirements: 1), it should be able to be used as the linkage between the original raster image and vector representation of objects implicitly contained on the image, which means that the data structure should represent each region directly and it should be easy to calculate the every kind of properties associated with regions such as the region boundary list, area of region, intensity mean of region, etc.; 2), the data structure should be in a hierarchic fashion. Such requirement is based on the observation that segmentation is an evolving procedure which usually starts from original raster image and gradually groups small regions into more meaningful regions. During such evolution, some grouping or decision making may go wrong due to a variety of reasons. Therefore it should be possible to return to more primitive status and make a new decision. Bearing these requirements in mind, a "N-node tree" has been developed (Fig.2).

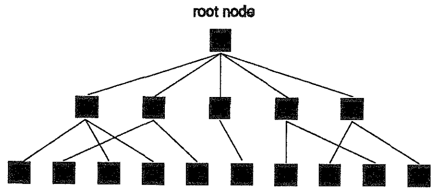

Fig.2

One can regard the "N-node tree" as the extension of quad tree, and under such extension the number of children under a node is changeable. The whole tree consists of a number of levels, with each level representing segmentation results at different stages. Each node on one level describes a complete region which has no overlapping with other regions on the same level. For a node on one level, one can find out its associated original image pixels by tracing down the tree through its children until the lowest level is reached where each node represents the image pixel indexed from left to right and from top to bottom on the image. In practice, in order to facilitate the segmentation, the pixels corresponding to a region is stored as one of properties of a node. In addition, a label image, which has the same size as the original image, is created and valued by its corresponding label. By such strategy, it is easy to refer the label by image pixel or refer image pixels by a region node.

- vector representation. Vector representation is the critical step for shape analysis. Shape is a function of the position and direction of a simply connected curve defined within a two dimensional field. A simply connected curve is one in which any point on the curve has at most two neighbours which lie on the curve. The coding of shape may involve a description of a closed boundary or the pixels which lie with it. In order to incoperate the result from edge detection and point detection, we extent the concept of "regions" to include the lines and points by representing the region boundary using the edges between the region boundary pixels instead of pixels themselves (Fig.3).

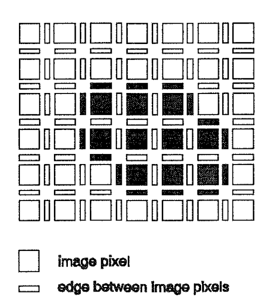

Fig.3

- 2-D structural description. The individual description on each object region is often not sufficient for the final goals of many applications, because such description may be

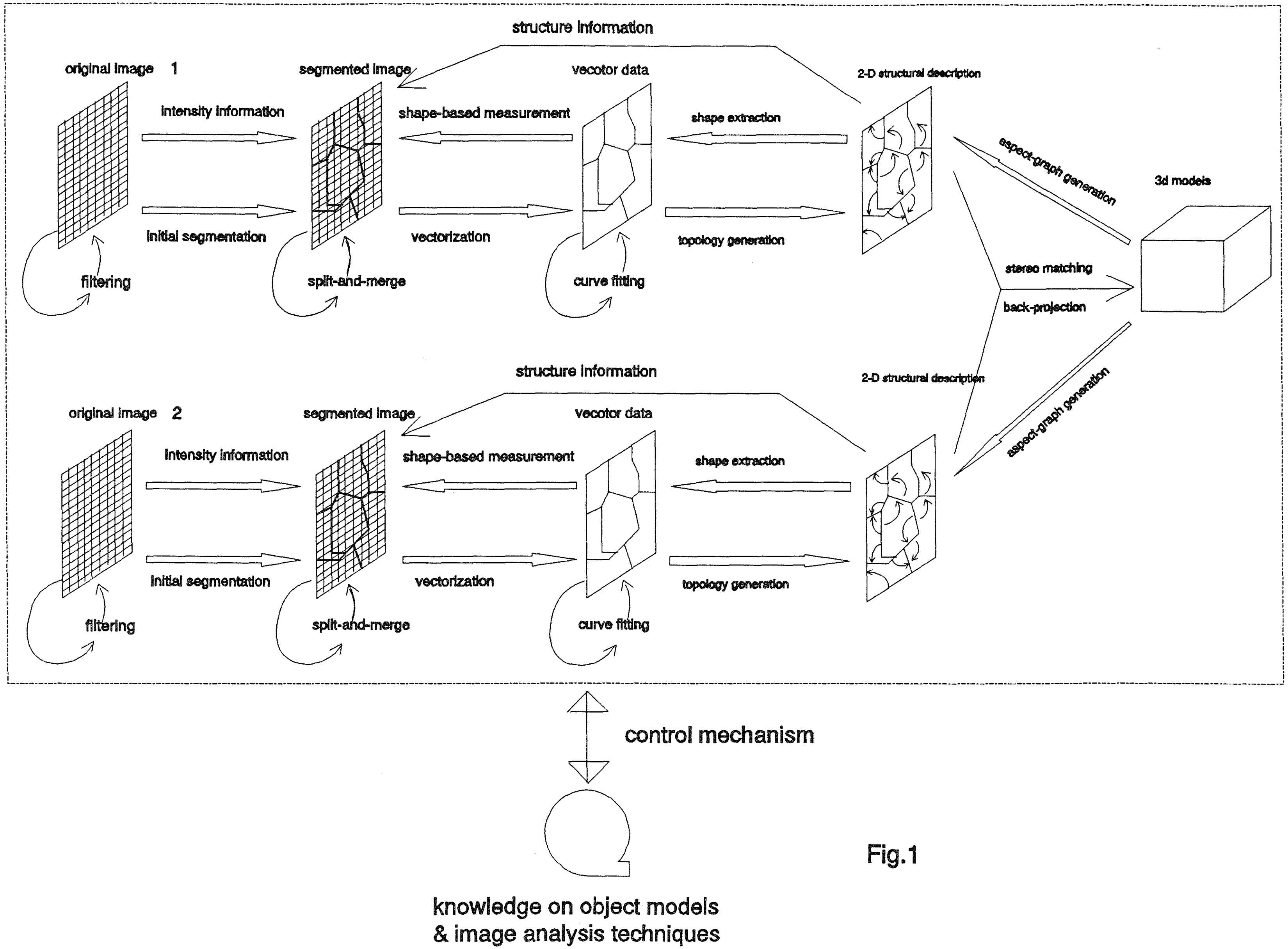
structure information
original image 1
segmented image
vecotor data
2-D structural description
intensity information
shape-based measurement
shape extraction
aspect-graph generation
3d models
initial segmentation
vectorization
topology generation
filtering
split-and-merge
curve fitting
stereo matching
back-projection
structure information
2-D structural description
original image 2
segmented image
vecotor data
intensity information
shape-based measurement
shape extraction
aspect-graph generation
initial segmentation
vectorization
topology generation
filtering
split-and-merge
curve fitting
control mechanism
knowledge on object models
& image analysis techniques

Fig.1

to vague for the purpose of high-level analysis such as matching and object recognition. Structure is used to represent the interrelationship of the boundaries, which provides the basis for many kinds of tasks in computer vision, specially for matching, object recognition purpose.

- 3-D structural description of models. The representation scheme for 3d world is the typical topic in symbolic artificial intelligence and computer graphics. A structural description of an object consists of the descriptions of its parts and their interrelationships. The parts of an object can be primitive (nondecomposeable) or they may be further broken down into subparts. When the parts of an object are not primitives, the structural description of the object consists of one level of descriptions for each level of subparts. Such a multilevel description is called a hierarchic description and is useful for complex objects with many repetitions of parts and subparts.

Operations on one layer

By "operations on one level", we mean that input and output of the operation are in the same format, e.g., from raster image to raster image, from vector data to vector data, etc. During the processing, the content of representation may change.

- operations on the original raster images. These operations may include 1) calculating the characteristics of image (e.g. histogram transformation, etc); 2) image quality improvement (e.g. enhancement, noise suppression by filtering, etc).

- operations on the segmented images. Split-and-merge is the main mechanism in the segmentation procedure, which merges small regions into more meaningful big regions, or split the big regions into small regions if necessary.

- operations on the vector data. This refers to the line fitting or curve fitting algorithm, which reduces the data needed while keeping the result as closed as possible to the original data. To be useful for high-level analysis, these vector data must be approximated so as to overcome local noise, and be represented in a more manageable form. The more comprehensive that representation is, the better the performance of the analysis would be.

Operations between the layers

The operations under this category changes or transfers the data from one representation to another representation, which are the essential parts of image analysis.

- operations between the original image and segmented image. These operations are generally called segmentation which usually is in two kinds of forms: a), edge detection and line following. This category of techniques study various of operators applied to raw images, which yield primitive edge elements, followed by a concatenating procedure to make a coherent one dimensional feature from many local edge elements; b), Region-based methods, which depend on pixel statistics over localized areas of the image. Regions of an image segmentation should be uniform and homogeneous with respect to some characteristic such as grey tone or texture. Region interiors should be simple and without many small holes. Adjacent regions of a segmentation should have significantly different values with respect to the characteristic on which they are uniform. Boundary of each segment should be simple, not ragged, and must be spatially accurate [Haralick].

- operations between the segmented image and vector data. This so-called vectorization procedure traces along the each region boundary to get the boundary position and the position is represented in chain code, which is used later by shape analysis. On the other hand, in order to integrate shape constraint into segmentation, there is another information flow which transfers the result of curve fitting into the region growing. The principle of encoding shape is described in section 3.

- operations between the vector data and 2d structural description. These operations build the structural description by performing a geometric analysis on the vector data. Vector-based perceptual grouping can be also included in this category, which organizes the fragmented low-level descriptions into meaningful higher level descriptions by mimicking the human visual system in detecting geometric relationship such as collinearity, parallelism, connectivity, and repetitive patterns in an otherwise randomly distributed set of image elements, some relevant work can be referred to [Mohan].

- operations between the 2d and 3d structural descriptions. Matching two or more than two images of the same scene from different viewing positions in order to recover the

three-dimensional geometry of the scene, is the correspondence (stereo matching) problem. According to the space where the matching takes place, the existing techniques for solving the correspondence problem roughly fall into two categories: image space based and object space based. In the image space based matching, the primitives of one image are compared with ones on the another image. Many solutions to the matching have been proposed in the image space. The methods vary with different choice of primitives: area-based (intensity-based), feature-based and structure-based (relational matching). Recently, several articles are devoted to the object space based matching. This method emerged originally from the task of reconstructing digital terrain model from a pair of digital images, independently developed by Wrobel [Wrobel] and Helava [Helava], etc. Helava used the concept of "groundel" as a unit in object space similar to the "pixel" in the image space. The image intensities corresponding to each groundel can be analytically computed, if all pertinent geometric and radiometric parameters (including groundel reflectance, etc.) are known. A least square method is adopted to determine a set of unknown quantities or improvements to their approximate values used in the analytical prediction process.
The mapping of 3-D structure into 2-D description is formulated by the concept "aspect graph" which was introduced by Koenderink and van Doorn [Koenderink] [Gigus], based on the idea of using the aspect graph of topologically distinct views of an object to represent its shape. Informally, at each vertex of the aspect graph there is a view - an aspect - that is representative of the projections of the object from a connected set of viewpoints from which the object appears qualitatively similar. Two aspects are adjacent in the graph if the corresponding sets of viewpoints are adjacent. A visual event is said to occur when the view changes as the observer moves between adjacent sets.

Control mechanism

For a complex system, a control mechanism is always required to control the reactions on the components as well as the ones between the components inside the system. In Fig.1, such control mechanism is monitored by a database which has the description of two kinds of knowledge, that is, knowledge about objects and about analysis tools (i.e. image analysis techniques). The control structure concerns the model to integrate diverse levels and control the proper information flow and scheme. The characteristics of image and specification and requirement from the application can also be integrated. This is sometimes regarded as "knowledge-based image processing system" [Matsuyama,87,89] [Nicolin].

3. A UNIFIED MEASUREMENT BASED ON MDL PRINCIPLE

In the last section, we have proposed a architecture for an integrated image analysis. We want to emphasis on the interactive reactions between different levels or layers. The operations from low layer to high layers have been addressed in a lot of literature. In our research, we are interested to integrate the information from high layers into low layers' operations. In order to do so, we must answer the following theoretic questions:

1) What kind of knowledge can be integrated in low and middle-level processing.
2) What is the proper language which can describe the information from different layers on the common ground.

We here examine the several tools or criterion provided by probability theory and information theory.

MAP, BE, ML, LS

The Maximum A Posteriori (MAP) criterion selects the best solution on the model X that maximizes the conditional probability of the model given the data Y: $P(X/Y)$. The MAP criterion leads to three important estimation methods, namely Bayes Estimation (BE), Maximum Likehood (ML), and Least Square (LS).

applying Bayes' theorem gives

$$P(X/Y) = P(Y/X)P(X)/P(Y)$$

where $P(Y/X)$ is the conditional probability of getting data Y given the model X and $P(X)$ the priori probability of the model X. If we assume $P(Y)$ is constant, maximizing $P(X/Y)$ is equivalent to maximize the

$$P(Y/X)P(X),$$

which is the principle of Bayesian Estimation.

Further, under the specification that the priori probability $P(X)$ are all the same (constant), the MAP criterion leads to the simpler maximum likehood principle of maximizing

$$P(Y/X)$$

If the random variables to which the data Y refer are normally distributed, the maximum likehood method will give the same results as the least squares estimation which has widely been used.

MDL

Minimum Description Length (MDL) principle studies estimation based upon the principle of minimizing the total number of binary digits required to rewrite the observed data, when each observation is given with some precision. Instead of attempting at an absolutely shortest description, it looks for the optimum relative to a class of parametrically given distributions.

The MDL principle can be generally expressed as

$$L(x,\Theta) = L(x/\Theta) + L(\Theta) \quad (1)$$

where

L(.) is a measure of the uncertainty of an event and its unit is "bit".

$L(x,\Theta)$ is the total number of bits to describe the observed data when we introduce the model. "x" expresses the observed data, and "Θ" represents the model parameters.

$L(x/\Theta)$ is the number of bits to describe the data if assuming the model is known.

$L(\Theta)$ is the number of bits to describe the model.

$L(x,\Theta)$ is the least information content required to remove the uncertainty in the observation and describe the model. Thus, the number of bits in a description required for the interpretation of the observation becomes a measure of simplicity.

The main power of the MDL principle is that it permits estimates of the entire model, its parameters, their number, and even the way the parameters appear in the model; i.e., the model structure [Ressanen,78,83,84].

Bayesian Estimation has been successfully used in scene reconstruction [Zheng], while MDL principle has been applied in the image segmentation and feature extraction [Fua,89a,89b,91] [Leclerc,89,90] [Keeler]. The reason on the different applications is that it is reasonable to apply probability theory to analysis the raster image when we assume the image is a stochastic process, but is has not been proven that BE can be applied to analysis structural phenomena, which is sometime considered in the integration of statistic and structural pattern recognition. The flexibility of MDL is that we can treat (1) simply as a new criterion, while forgetting its background on statistics and information theory. If we can not derive a precise descriptive language or encoding scheme to describe an event by a number of bits based on Shannon's first theorem, an approximated scheme can still be used. The resulted estimation is the relative optimal result constrained by the approximated encoding scheme.

Examples of encoding using MDL

- encoding the image intensity

  The interior intensities of an image region can be modelled by a smooth intensity with a Gaussian distribution of deviations from the surface. The formula for this problem has been solved by Hua and Hanson [Fua,91]:

$$L_I = n_1(\log\sigma + c) + 8n_2 + [n_1\log(\frac{n_0}{n_1}) + n_2\log(\frac{n_0}{n_2})] + N_p \quad (2)$$

Where

$$c = \frac{1}{2}\log 2\pi e$$

L_I	:	number of bits to describe intensity information in a region
n_0	:	total number of pixels in one region
n_1	:	number of pixels in the Gaussian peak
n_2	:	number of outliers
N_p	:	number of bits to describe surface model
σ	:	standard deviation of intensity noise

In (2), first item is the cost of Huffmann-encoding the pixels in a Gaussian peak, second item is the cost of encoding the pixel outliers, third one is the entropy for encoding the pixel on whether it is or is not a anomalous and fourth term specifies the coding of the model.

- encoding shape of region

  In this paper, we give an example modelling on shape on which the ideal region boundary composes a number of straight line segments. For a straight line, its digitalization fulfils the chord property which states [Hung,85] that "a digital arc A is said to have the chord property if for every two digital points, the chess-board distance of any point of arc A to the straight line nowhere exceeds 1". We consider the points on a curve which do not meet the chord property as outliers, and such outliers are constrained by their neighbours.

  The cost required to describe the image region shape is formulated as:

$$L_S = 0 + 2m_2 + [m_1\log(\frac{m_0}{m_1}) + m_2\log(\frac{m_0}{m_2})] + \log(D_xD_y)m_S \quad (3)$$

Where

- L_S : number of bits describing the shape of a region
- m_0 : total number of points on the boundary
- m_S : number of straight line segments on the boundary
- m_1 : number of points fulfilling the chord property
- m_2 : number of outliers
- D_x,D_y: number of pixels along x and y direction of the image

In accordance with (2), we also use 4 items encoding the shape of regions. For the points on the curve which meet the chord condition, no additional coding is needed as far as the nodes specifying the straight line segments are known, so first item in this case is zero. The second item in (3) is the number of bits describing the outliers. If the boundary is encoded in Freeman chain code, 3 bits is required to store each pixel (for 8 directions). But if we store the edges between the pixels instead of pixels themselves, only 2 bits are necessary (for 4 directions). The third term is in the same meaning as equation (2). The final component is used for the coding of nodes connecting straight line segments.

4. APPLICATIONS

Segmentation

The segmentation quality can be significantly improved by utilizing the image intensity as well as high level knowledge about the objects contained on the image. We have successfully integrated the shape constraint into segmentation using three layers in Fig.1, i.e. original image, segmented image and vector data, on which
the segmentation is the result from the operations carried on segmented image and inter-reactions from the original image and vector data. Split-and-merge is the main mechanism in the segmentation procedure, which merges small regions into more meaningful big region, or split the big region into small regions when required. An initial segmentation is performed to get basic regions from the original image. After each level of split-and-merge, vectorization procedure transfers the region boundary into vector description, followed by a curve fitting algorithm which derives a compacted vector data based on the generic model. Based on the result of curve fitting, a measurement is calculated using MDL principle to describe the uniformity of region by shape constraints. Such measurement is integrated with the information derived from the original image intensity to improve the decision making of split-and-merge of regions. For the detail of this part of work, reader is referred to [Zhang,92b].

Stereo matching

The stereo matching (or correspondence) remains one of permanent problems in Computer Vision. In [Zhang,91a,92a], author presents a new approach to solve the problem, which incorporates the image space based matching techniques with the high level knowledge about the objects. The low-level processing (edge detection, feature extraction) and candidate matching are carried out in image space, while the final matching is determined in object space as solving a consistent labelling problem which results from the integration of candidate matching, high level constraints of objects and other constraints of image matching. One of the innovative features in our approach lies in back-projecting (back tracing) the line pairs from candidate matching into the object (scene) space, and combining all the constraints in a unified process. We substitute the concept of "figure continuity" usually used in the image matching with the high level knowledge from the object space.

Integration of segmentation and stereo matching

Our experiment has shown that segmentation is one of major difficulties in matching, among other reasons. One of possible solution is to integrate the segmentation with matching interactively as proposed in the following: after an initial segmentation which forms the lowest level in "N-node tree", a candidate stereo matching is carried out, which assigns the corresponding regions from one image to another images by using simple criterions such as shape, intensity difference, etc. During the next step of segmentation, stereo information is included, that is, in considering the merging of one region with its neighbouring region, the corresponding regions in the candidate pools are extracted and a unified measurement is calculated which integrate the intensity and shape information from both images as well as the some invariant properties constrained by the central projection geometry [Forsyth] [Boyer]. After each stage of segmentation, stereo matching is performed which introduces the other matching constraints such as uniqueness, together with the constraints described by object models. Such segmentation and matching procedure continue interactively until the result does not change.

Integration of edge-based and region-based segmentation

It is observed by a lot of researchers that no single method can provide a complete interpretation of image. However, each method may provide a subset of the information necessary to produce a more meaningful interpretation of the scene [Chu,90,91] [Hsieh] [Shufelt]. It is reasonable to expect that there will be complication in fusing the result from edge detection and region-based segmentation. Most research on this subject falls into four categories: (1) pixel-wise logical operations; (2) algorithm for specific imaging modalities or processing techniques; (3) theoretical approaches using a priori information and probabilistic models; and (4) techniques using high-level knowledge. Our future scheme will be oriented on using high-level constraints to enhance the quality of segmentation.

5. CONCLUSIONS

In this article, we have proposed an architecture for integrated image analysis which enable us to carry out image analysis from low-level to high-level and from high-level to low-level interactively. We are especially interested in the problems of: 1) integrate high-level knowledge into image segmentation; 2) use structural information for stereo matching; 3) integrate image segmentation and stereo matching; 4) combine the edge-based and region-based segmentation. One of major theoretic obstacles is on how to combine the knowledge from different sources, which has partially answered in this paper.

ACKNOWLEDGEMENT

Without the consistent support from Prof.dr.ir. M.J.M. Bogaerts, the work presented here would be not possible. This research is jointly supported by CCGM (Centre for Computer Graphics and Mapping), FRANK project and Photogrammetry Laboratory, Faculty of Geodesy, TU Delft. FRANK is a registered trademark of FRANK system supported by Geeris Holding Netherlands BV.

REFERENCES

Aloimonos, J.Y., Davis, L.S., and Rosenfeld, A., "Maryland progress in image understanding", Image Understanding workshop 1990.

Boyer, K.L, and Kak, A.C., "Structural stereopsis for 3-D vision", IEEE Transactions on PAMI, Vol.10, no.2, March 1988.

Chu, C.-C. and Aggarwal, J.K., "The integration of region and edge-based segmentation", International Conference on Computer Vision 1990.

Chu, C.-C. and Aggarwal, J.K., "The interpretation of laser radar images by a knowledge-based system", Machine Vision and Applications (1991) 4:145-163.

Forsyth, D., Mundy, J.L., Zisserman, A., Coelho, C., Heller, A. and Rothwell, C., "Invariant descriptions for 3-D object recognition and pose", IEEE Transactions on PAMI, Vol.13, No.10, Oct. 1991.

Fua, P., and Hanson, A.J., "Objective Functions for Feature Discrimination: Theory", Image Understanding Workshop 1989a.

Fua, P., and Hanson, A.J., "Objective Functions for Feature Discrimination: Applications to semiautomated and Automated Feature Extraction", Image Understanding Workshop 1989b.

Fua, P., and Hanson, A.J., "An Optimization Framework for Feature Extraction", Machine Vision and Applications (1991) 4:59-87.

Gigus, Z., and Malik, J., "Computing the aspect graph for line drawings of polyhedral objects", IEEE Transactions on PAMI, Vol.12, no.2, Feb. 1990.

Helava, U.V., "Object-space least-square correlation", Photogrammetric Engineering and Remote Sensing, Vol.54, No.6, Part1, June 1988, pp.711-714.

Haralick R.M. and Shapiro, L.G., "Image segmentation techniques", Computer Vision, Graphics, and Image Processing, vol. 29, 1985, 100-132.

Hsieh, Y.C., Perlant, F., and McKeown, D.M., "Recovering 3D information from complex aerial imagery", Image Understanding Workshop 1990.

Hung, S.H.Y., "On the straightness of digital arcs", IEEE Transactions on PAMI, Vol. PAMI-7, no.2, pp.203-215, 1985.

Keeler, K., "Map representations and coding-based prioris for segmentation", Inter. Conference on Computer Vision and Pattern Recognition, Hawaii, 1991.

Koenderink, J.J., and Doorn, A.J. van, "The internal representation of solid shape with respect to vision", Biol. Cybern., vol.32, pp.211-216, 1979.

Leclerc, Y.G., Image and boundary segmentation via minimal-length encoding on the Connection Machine, Proceedings, Image Understanding Workshop (Palo Alto, CA, May 23-26, 1989), Morgan Kaufmann, San Mateo, CA, 1056-1069.

Leclerc, Y.G., "Region grouping using the minimum-description-length principle", Image Understanding Workshop 1990.

Matsuyama, T., "Knowledge-based aerial image understanding systems and expert system for image processing", IEEE Transactions on Geoscience and remote sensing, Vol.GE-25, No.3, may 1987

Matsuyama, T., "Expert systems for image processing: knowledge-based composition of image analysis processes", Computer Vision, Graphics and

Image
Processing 48, 22-49 (1989).

Nicolin, B. and Gabler, R., "A knowledge-based system for the analysis of aerial images", IEEE Transactions on Geoscience and remote sensing, Vol.GE-25, No.3, may 1987.

Rissanen, J., "Modelling by shortest data description", Automatica, Vol.14, pp.465-471, 1978.

Rissanen, J., "A universal priori for integers and estimation by minimum description length", The annals of statistics, 1983, Vol.11, no.2, 416-431.

Rissanen, J., "Universal coding, information, prediction, and estimation", IEEE Transactions on Information Theory, Vol.IT-30, No.4, July, 1984.

Shufelt, J., and McKeown, D.M., "Fusion of monocular cues to detect man-made structures in aerial imagery", Image Understanding Workshop 1990.

Wrobel, B., "Facet Stereo Vision (FAST vision) - a new approach to computer stereo vision and to digital photogrammetry", In proceeding of intercommission conference on Fast Processing of Photogrammetric Data, Interlaken, Switzerland, pp.231-258, 1987.

Zhang, Y., "Matching in image-object dual spaces", SPIE Vol.1526 Industrial Vision Metrology, 1991a.

Zhang, Y., "A pool-based searching approach for self calibration of multi-cameras and motion analysis under possible incorrect correspondences", Technical report, Faculty of Geodesy, Delft University of Technology, 1991b.

Zhang, Y., "A new stereo matching approach in image/object dual spaces", XVII ISPRS congress, Commission III, Washington, D.C., U.S.A., 1992a.

Zhang, Y., "Deformable model for image segmentation", XVII ISPRS congress, Commission IV, Washington, D.C., U.S.A., 1992b.

Zheng, Y.-J and Hahn, M., "Surface reconstruction from digital images in the presence of discontinuities, occlusions and deformations", In proceeding of ISPRS Commission II conference, Wuhan, China, 1990.