Perceptual Grouping of Regular Structures for Automatic Detection of Man-Made Objects

Examples from IR and SAR

U. Stilla, E. Michaelsen, U. Soergel, K. Schulz
FGAN-FOM Research Institute for Optronics and Pattern Recognition
76275 Ettlingen, Germany
e-mail: stilla@fom.fgan.de

***Abstract*— Human observers perceive man-made objects in images from the visual spectrum domain as well as in IR or SAR imagery. Mechanisms like perceptual grouping are crucial to this capability. In this paper two examples for grouping in different image sources are discussed. The first example is activity estimation in urban areas from thermal IR images. The grouping of vehicles into rows is performed along the margins of the roads. The other example is related to the detection of industrial buildings from InSAR data. Such buildings often show salient regular patterns of strong scatterers on their roofs. A previous segmentation which uses the intensity, height and coherence information extracts building cues. Strong scatterers are filtered by a spot detector and localized by a cluster formation. These scatterers are grouped in rows by a process that uses the contours of the building cues as context.**

Keywords-perceptual grouping; image understanding; infrared; SAR; urban areas

I. INTRODUCTION

Human observers often show remarkable capability to discriminate man-made objects from natural objects or random clutter in images. This holds for images from the visual spectrum domain as well as for images from infrared (IR) or SAR - even if the observer is not an expert. Perceptual grouping considering simultaneously radiometric as well as geometric aspects may be one key feature for explaining this performance.

Almost hundred years ago perceptual psychology researchers began to investigate the principles of human vision, e.g. figure-background discrimination and the composition of objects. Wertheimer [6] showed some fundamental features of object relations which determine the way in which objects tend to be grouped. Marr [1] introduced basic ideas resulting from perception research into computer vision in the 80's: "Vision is the construction of efficient symbolic descriptions from images of the world".

Exploiting the principles of perceptual grouping in automatic image analysis systems allows to improve the discrimination of man-made objects with regular structures from other objects even if their radiometric properties are similar.

II. PERCEPTUAL GROUPING

Given a set of n extracted objects from an image the problem of finding regular structures arises. In principle this is a search for subsets i.e. members of the power set 2^n. Human vision tends to prefer just a very small number of certain subsets with special properties common to all subjects. Wertheimer found proximity, similarity, good continuation and symmetry as key relations controlling the grouping process.

Fig. 1a shows a set of 16 dots that are perceptually decomposed into 4 columns of 4 members. This occurs due to the proximity. Fig. 1b shows the same geometrical setting. The similarity of the objects in the horizontal rows leads to an alternative decomposition. The property of 'good continuation' is evident in Fig. 1c. In spite of the competing relations proximity and similarity here the diagonal group is preferred.

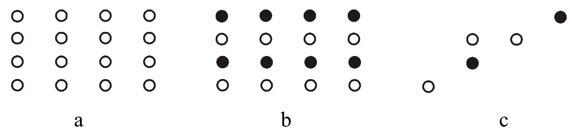

Figure 1. Examples of a grouping

These relations can be used to prune the search in the power set. Nevertheless, in real data from images due to noise, missing objects, spurious objects, and competing perceptual relations there remain several possibilities for grouping. Contextual knowledge for preferred locations or directions of grouping may further reduce an stabilize the search.

III. KNOWLEDGE REPRESENTATION

Structural relations of objects can be captured by rules or so-called productions. The production defines how a given configuration of objects is transformed into a single more complex object. In the condition part of a production, geometrical or topological relations are examined. If the condition part of the production holds, an object specific generation function is executed to generate a new object. An object concept represents a type of an object and defines a frame for the attribute values of an individual object (instance). The hierarchical organization of object concepts and productions can be visualized by a production net [3]. A basic

0-7803-7930-6/$17.00 (C) 2003 IEEE

production net for a grouping process is depicted in Fig. 2. In production p1 the relations proximity and similarity of objects SINGLE are utilized to select preferred pairs and to generate objects PAIR only from these. Production p2 additionally uses the relation 'good continuation' to select only those objects SINGLE that fit into the spacing and the direction of the object PAIR. Such an object GROUP can be recursively prolonged by additional objects SINGLE executing production p3 to arbitrary length. More complex structures can be composed by extending the net and grouping the objects GROUP in objects META_GROUP.

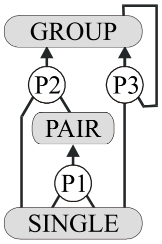

Figure 2. Production net

IV. APPLICATIONS

In this paper two examples for grouping from different applications and image sources are presented. The first example is given from activity estimation in urban areas from thermal infrared images and the other example is related to the detection of buildings from InSAR data. A precondition for object grouping in images is the extraction of objects itself. Depending on the recognition task different types of primitive objects like points, spots, lines, corners, or more complex objects can be considered. For the imagery of both examples a simple spot detector has been used.

A. Thermal Infrared Images

Monitoring traffic in dense urban areas is a difficult task. The detection of moving cars is being studied for a long time, but the activity of cars cannot only be described by their movement. Stationary cars may still be active, i.e. waiting at a traffic light or in a traffic jam. An important feature for the activity of a car is its temperature. Temperature can be captured by IR-sensors. An approach to assess the activity of non-moving vehicles based on airborne image sequences from an infrared camera is described in [4].

At the resolution of approximately one meter vehicles appear as elongated spots. In urban areas many additional other objects have the same property. Large scale vector maps of roads provide a proper means to exclude most of the non-vehicle spots. Interest regions were derived from the map and used to discriminate vehicles candidates from clutter in the IR-image.

Cars tend to be placed in equidistantly spaced rows along the margin of roads. This criterion allows to discriminate them from other remaining spot-shaped objects. Grouping of such spots into rows of arbitrary length is a generic operation. Fig. 3a shows a section of an IR image containing a row of cars. All detected warm and cold spots in this section are displayed in Fig. 3b. Spots caused by cars constitute only a subset of these. Fig. 3c shows those spots, that are sufficiently close to a road margin. The grouping starts only from spots, which exceed a minimal mass. The grouping direction is constrained by the road margin. Only those spots fitting into the straight and equidistantly spaced row model are grouped.

Still there may be several alternatives of grouping, e.g. if two spots are close to another in a location consistent with the model (see Fig. 3c, most right member of the row). Among the alternatives one group is selected based on an assessment, which is calculated from the number of spots, total mass, regularity in spacing and straightness and consistency in orientation of the spots. Fig. 3d shows the best assessed group containing seven spots.

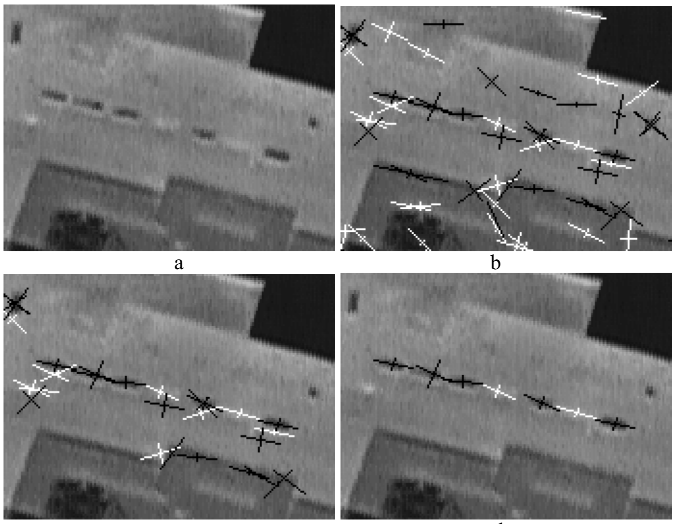

Figure 3. The benefit of grouping: a) section of an IR-image, b) all spots constructed in that region, c) spots in the interest region given by fusion with the map, d) car-spots remaining consistent with the row model after grouping.

B. SAR Images

Detecting buildings in InSAR data is a very active field of research. But the reconstruction of small buildings or buildings in dense urban areas from such data is limited by certain geometric properties of the SAR acquisition [5]. Large buildings in rather open areas like industrial sites or airports often show flat roofs or roofs with flat superstructures. One other important feature is that most of such buildings have rectilinear outlines with long planar vertical faces leading to a significant step in height. These features can be used to discriminate building candidates from many other large objects An approach exploiting the characteristic appearance of certain buildings to support the recognition process is described in [2].

Particularly industrial sites often exhibit roofs with regular structure of strong scatterers. This feature is perceived very dominantly by human observers in the intensity data (Fig. 4b). The scatterers correspond to ventilation, air-condition or natural lighting facilities necessary for the purpose of the building. They tend to be placed in equidistantly spaced rows parallel to the outline of the building. This feature provides very strong evidence for large well-organized man-made assemblies making it suitable to infer an industrial building with high confidence. Large buildings close to such an industrial building are assumed to be industrial buildings, too. The recognition of such a complex of industrial buildings can therefore use the industrial buildings with regular patterns of scatterers as a seed for a larger aggregate forming an industrial site.

First, the intensity image is segmented using a region growing approach. Then the elevation data are averaged within the segments obtained from the intensity data and weighted by the coherence data. Segments of significant height above ground are building cues. From the outlines of the building cues an orientation histogram is assembled giving two major orientations rectangular to each other. The margins of the segments are classified into these two. Fig. 4e shows the resulting rectilinear polygons in black.

Strong scatterers are filtered by a spot detector (Fig. 4c) and localized by a cluster formation. Fig. 4d shows examples of such spot clusters displayed as white circles. The corresponding part of the intensity image is used as background.

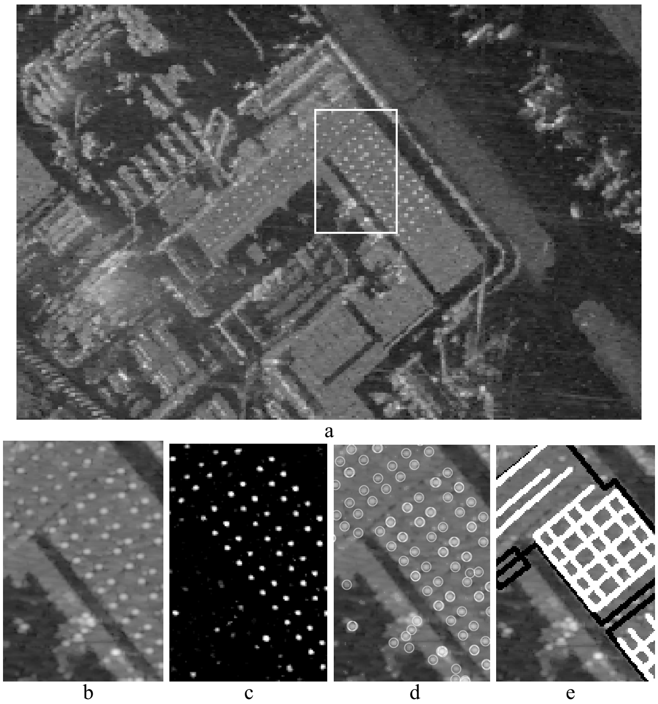

Figure 4. Detection and grouping of salient scatterers from an industrial building: a) section of SAR image (intensity) from Frankfurt Airport, b) section of a, c) result of spot filtering, d) cluster formation, e) grouping results (white)

Grouping such rows of arbitrary length is performed by successively adding spots to the row. This is initialized by rows with only a single member, where the direction information is provided by the major orientations of the building. With growing numbers of members the search areas become narrower and better regularity assessments can be calculated. A building will be accepted as industrial building if a row is assembled on its roof with sufficient regularity. Fig. 4e displays these rows as white lines. Some of them contain up to 20 member spots which are very precisely and regularly aligned.

Regular structures above each other on the facade of buildings (e.g. Fig 5b) may cause groups of scatterers aligned with the range direction of SAR (Fig. 5c). Detected spots are shown in Fig. 5d. Such groups may be grouped along the building contour again.

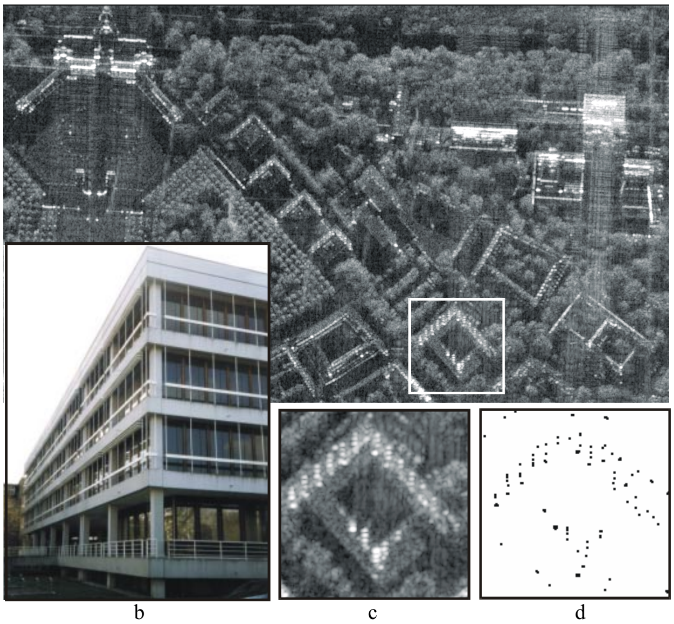

Figure 5. Groups of scatterers caused by buildings

V. CONCLUSION

The mentioned examples have shown that the detection of regular structures by perceptual grouping can support the recognition process of man-made objects. Performing a recursive grouping without context may lead to a high computational effort. This can be significantly reduced if context is provided that limits possible locations and orientations of grouping. In the first example road areas of a corresponding map were used to determine expectation locations of vehicles and road margins were used to direct the grouping orientation. In the second example elevated areas (building candidates) limit the search space and the contour provides constrains for the orientation of the grouping process.

REFERENCES

[1] Marr D. "Vision:". San Francisco: Freeman, 1980

[2] E. Michaelsen, U. Soergel, U. Stilla, "Grouping salient scatterers in InSAR data for recognition of industrial buildings" In: R. Kasturi, D. Laurendeau, C. Sun (eds) 16th Int. Conf. on Pattern Recogn., ICPR 2002, vol II, 2002, pp. 613-616.

[3] U. Stilla, E. Michaelsen "Semantic modelling of man-made objects by production nets" In: A. Gruen, E.P. Baltsavias, O. Henricsson (eds) Automatic extraction of man-made objects from aerial and space images (II) Basel: Birkhaeuser, pp. 43-52, 1997

[4] U. Stilla, Michaelsen E, "Estimating vehicle activity using thermal image sequences and maps", Symposium on geospatial theory, processing and applications. International Archives of Photogrammetry and Remote Sensing, 2002

[5] U. Stilla, U. Soergel, U. Thoennessen U, Potential and limits of InSAR data for building reconstruction in built up-areas. ISPRS Journal of Photogrammetry and Remote Sensing, 58 (1-2), pp. 113-123, 2003

[6] M. Wertheimer "Laws of Organisation in Perceptual Forms" First published as "Untersuchungen zur Lehre von der Gestalt II". Psychologische Forschung 4, pp 301-350, 1923. Translation published in W. Ellis (1938) A source book of Gestalt psychology, pp. 71-88, London: Routledge & Kegan