

SEGMENTATION OF INTERFEROMETRIC SAR DATA FOR BUILDING DETECTION

Uwe SOERGEL, Ulrich THOENNESSEN, Hermann GROSS, Uwe STILLA
 FGAN-FOM Research Institute for Optronics and Pattern Recognition
 D 76275 Ettlingen, Germany
{soe,thoe,gross,usti}@fom.fgan.de

Working Group II/4

KEY WORDS: Visualisation, Interferometry, Data fusion, Segmentation, Buildings.

ABSTRACT

The improved quality of InSAR data suggests to utilise such data for building detection. But, the phase information from which the height data is calculated, is often severely disturbed, depending on the signal to noise ratio. In this paper we refer to investigations to stabilise and improve the InSAR height data.

After speckle filtering, a segmentation of the intensity data is carried out. With these segments height data are masked and an average height is calculated using the related intensity values as weights. In a post-processing step, possible existing over- and under-segmentation are corrected. Adjacent objects with similar heights are merged and objects including shadow areas are split.

Different tasks are distinguished for utilisation of the derived 3D information. A test site including the airport of Frankfurt (Main) was chosen. For the visualisation purpose, the segmentation result is shown. The results are compared to a vector map and differences are depicted and discussed.

1 INTRODUCTION

Photo interpretation of remotely sensed data can be supported by fusing data from different sensors and knowledge sources together with an appropriate visualisation of this data. Such a visualisation can be e.g. mapping of imagery on a 3D-model of urban areas. Multi spectral aerial images (Figure 1) can be combined with laser altimeter data. Additionally, vector maps and laser elevation data may be used to generate prismatic models as a coarse 3D-description of buildings [Stilla & Jurkiewicz, 1999]. In this paper we focus on improved visualisation of InSAR data and combining InSAR data and vector maps.

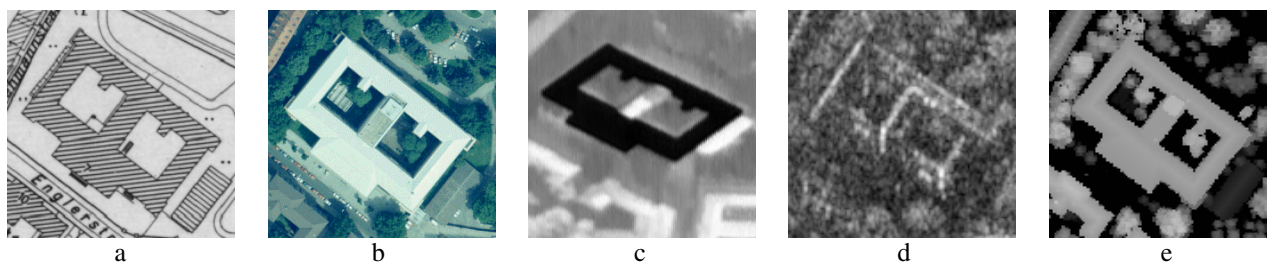


Figure 1: Knowledge Sources: a) Map, b) Optical Image, c) IR Image, d) SAR Image, e) Laser Altimeter Data

The improved quality and ground resolution of InSAR suggests to use this data for supporting the interpretation of man-made objects. But, in the case of SAR interferometric data many additional problems have to be taken into account, e.g. radiometric correction, layover effects, multi-backscattering, speckle, phase-unwrapping errors, relations and dependencies of intensity and noise in the height information [Schreier, 1993] [Bamler et al., 1998].

Even small errors in the phase lead to big errors in the height values. Thus, the phase information has to be prepared for further processing. Smoothing of the height data, like phase multilook processing [Goldstein et al., 1988], blurs edges of man-made objects (e.g. building walls) and as a consequence, details may be lost. On the other hand, object borders can often be extracted in the intensity data because of the different backscattering behaviour of different materials. Additionally, in many cases areas with comparatively homogeneous intensity distributions correspond to regions, respectively objects, with same height information. Such regions are extractable by edge-preserving region-oriented image processing algorithms. In our approach we segment regions with similar intensity values in the speckle-filtered intensity data using an edge-sensitive region growing algorithm [Levine & Shaheen, 1981]. We are interested in man-

made objects with constant height present in the scene like flat-roofed buildings. Hence, the height of one segmented object is set to the mean height over the segment, weighted with the intensity values [Soergel et al., 2000].

The derived depth map can for example be used to support an interpreter in map updating tasks. A map is compared with the depth map, which has to be geocoded for this purpose. A possible systematic bias of the height data, which might be present because of erroneous navigation data, can be corrected by incorporating tie points in the map. The depth map is split in two complementing depth maps by masking it with the given ground plans of the buildings in the map. The part containing the buildings is used to verify the buildings in the map and detect missing buildings. Large compact areas in the complementing depth map with significant height over ground are hints to recently built up buildings not represented in the map.

First the interferometric principle is briefly recapitulated to point out the dependency of InSAR data and the signal to noise ratio (SNR). Then, the segmentation process in the intensity data and the smoothing of the height data are described. Afterwards, the incorporation of the evaluated smoothed depth data for the map updating task is suggested. Finally, results of the approach are presented and discussed.

2 INTERFEROMETRIC PRINCIPLE

In this paper we focus on interferograms derived from airborne single pass measurements. Figure 2 illustrates the basic principle of SAR interferometry. An aeroplane carries two SAR antennas which are displaced by a base-line B . One of the antennas illuminates the scene and both antennas receive the backscattered complex signals s_1 and s_2 . The interferogram S is calculated by a pixel by pixel complex multiplication of the two received signals:

$$S = |s_1| \cdot |s_2| \cdot \exp(j\Delta\varphi) \tag{1}$$

The object height h can be expressed as a function of the phase difference $\Delta\varphi$:

$$h = \frac{\lambda \cdot r \cdot \cos(\theta)}{2\pi \cdot B} \cdot \Delta\varphi \tag{2}$$

with parameters distance r , wavelength λ , effective baseline B and depression angle θ .

SAR interferometry makes only sense in case of significant correlation respectively coherence between the two complex SAR images. The coherence γ depends on the expectation values of the signals. It can be estimated from the data by window-based computation of the magnitude of the complex cross-correlation coefficient of the SAR images [Hellwich, 1999]

$$\hat{\gamma} = \frac{\left| \sum_{n=1}^N s_1^{(n)} \cdot s_2^{(n)*} \right|}{\sqrt{\sum_{n=1}^N |s_1^{(n)}|^2 \cdot \sum_{n=1}^N |s_2^{(n)}|^2}} \tag{3}$$

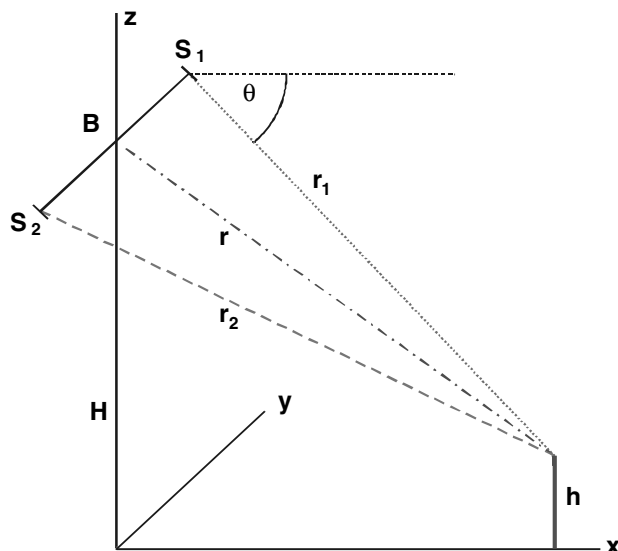


Figure 2: Geometry of InSAR Measurement

In case of a single pass airborne system temporal and geometrical contributions to the coherence coefficient can be neglected compared to the influence of noise. Assuming additive thermal noise, the complex signals s_i can be modelled as consisting of a correlation part c and noise part n_i

$$s_i = c + n_i \tag{4}$$

The absolute value of γ can be expressed as a function of the SNR [Rodriguez & Martin, 1992]:

$$|\gamma| = \frac{1}{1 + \frac{1}{SNR}}, \quad \text{with} \quad SNR = \frac{|c|^2}{|n|^2} \tag{5}$$

If no phase-unwrapping errors occur, the standard deviation σ_h of the height measurement is:

$$\sigma_h = \frac{\lambda \cdot r}{2\pi \cdot B \cdot \sqrt{SNR}} \tag{6}$$

Thus, the height accuracy declines with decreasing SNR.

3 SEGMENTATION OF INTERFEROMETRIC SAR DATA

Smoothing directly in the height data without regarding the intensities reduces σ_h , but also blurs sharp height jumps in the scene like those between building roofs and the ground.

Due to the interferogram calculation (Equation 1) the height and intensity data are co-registered. Therefore, an exploitation of height and intensity data in combination is immediately possible. Concerning a model-based segmentation of man-made objects, we propose as a first approach that regions with similar intensity values correspond to regions with equal height information. The regions are segmented in the intensity data with an edge-sensitive region growing algorithm. Because of the additive noise model in equation 4 we consider segments with low intensity values as noisy and brighter areas as more reliable regions with higher *SNR*. Thus, segments with too small mean intensities are masked. Inside each remaining segment the weighted mean height is calculated, using the intensity values as weights. Before the segmentation step, the data has to be prepared by speckle filtering to avoid an over-segmented result. Furthermore, adjacent segments with similar heights are merged.

3.1 Pre-processing

To limit the number of segments in the subsequent region growing step, the intensity data has to be filtered to reduce the speckle effect. We tested several filters described in the literature. Best results for our purposes were achieved using a median and the ESTEC filter [Desnos & Matteini, 1993]. After despeckling, the intensities are linear scaled to 8 bit to achieve identical start configurations in the region growing step for different images.

In the filtered image the gradient is calculated. Pixel with a gradient higher than a threshold are considered as possible region borders. The threshold th_b is calculated from the average intensity (*mean*) and standard deviation *sd* of the entire intensity image:

$$th_b = mean + 2 \cdot sd \quad (7)$$

The border candidate pixels are skeletonised by morphological operations [Glazer, 1993][Schulz & Thoennesen, 2000]. The result is a binary image with thin borders.

3.2 Segmentation

In the speckle filtered scaled intensity image an adaptive region growing step is carried out. For a detailed description of the algorithm the reader is referred to Levine & Shaheen [1981]. The only parameter of the algorithm is the grey level distance threshold th_r for a pixel to the mean of a adjacent region. If the distance is smaller than th_r the pixel is included in the region. The choice of th_r is a trade-off between possible over- and under-segmentation: On the one hand, every object of interest in the scene should be extracted. This indicates a high th_r , which might lead to a too large number of small segments. On the other hand, a low th_r might result in a few large segments only and details are lost. Furthermore, the large segments tend to "bleed" over the entire image.

In this first approach we are interested in a course 3D description only. Hence, a low th_r was chosen. To prevent under-segmentation, the binary border image is incorporated in the algorithm. As additional constraint, the region growing is not allowed to cross the borders. Our goal is to derive a coarse prismatic 3D-model of the man-made objects present in the scene. Therefore, small details like chimneys are ignored and masked. This is also true for areas with low coherence, which are identified by mean intensity values below a threshold th_i :

$$th_i = mean - sd \quad (8)$$

3.3 Smoothing of height data

Inside the remaining segments the height h_s of the segment with a number of N pixels is calculated by the mean of the heights h weighted with the intensity i :

$$h_s = \frac{\sum_{n=1}^N i^{(n)} \cdot h^{(n)}}{\sum_{n=1}^N i^{(n)}} \quad (9)$$

Holes inside the weighted heights which result from the masking step are set to the height value of the surrounding region. The strategy towards the remaining holes depends on the task. Because of their unreliable information they should not be considered for further processing. However, for visualisation purposes it is often better to set them to a default height, because holes in the evaluated depth map might irritate an interpreter.

3.4 Post-processing of the depth map

In a region segmentation based approach four types of false segmentation are possible: missing objects, noise objects, over- and under-segmentation [Geibel 2000]. Over-segmentation occurs for example if objects consist of different materials with different backscatter behaviour like rooftops consisting of concrete and tin parts. Thus, two separated segments are extracted in the region growing step. Therefore, in a post-processing step adjacent segments with similar heights are merged.

Object parts, like several height levels of flat roofed buildings, with the same roof material on every level, are often under-segmented because they can not be distinguished in the intensity data. Such objects can be identified if they include shadow areas. Shadow of buildings appears dark in the intensity data. Such areas can be used for a model-based object reconstruction [Hoepfner 1999]. Shadow leads to either long or L-shaped segments, depending on the aspect. Their width is a function of depression angle and the height difference between the object and the background. In this approach, segments which contain a possible shadow area (long or L-shaped dark segment) are further investigated. Two different manners of discriminating different height levels were tested. Firstly, the histogram of the median filtered height values of such a segment is analysed. If it shows a bimodal curve and the segment can be separated in two closed segments in the height data, then the segment is split into two closed segments. If the data suffers from noise or if the height difference is small the histogram might not be discriminated into two separated closed parts. Hence, in a second approach a region growing step in the median filtered heights is performed, utilising the border towards the shadow region as seeding points. After the correction of over- and under-segmentation the depth map is recalculated in the manner described above.

4 UTILISATION OF THE SEGMENTATION RESULTS

Three different states of a priori information of the scene are considered:

1. No map information available.
2. 2D vector maps available (e.g. building ground plans)
3. 2D vector maps and DEM available

In the first case significant height jumps between regions in the depth map are hints to buildings present in the scene. Hints which coincide with compact regions exceeding a minimum extension are further investigated. For a visual interpretation by an operator such regions can be highlighted in the intensity data for example.

Due to erroneous navigation data a systematic bias might be present in the height data. In the cases 2 and 3 mentioned above the bias can be corrected by incorporating tie points in the map respectively DEM. For the change detection and map update task, the depth map is split in two complementing depth maps by masking it with the given ground plans of the buildings extracted from the map [Stilla & Jurkiewicz, 1999]. The part containing the buildings is used to verify the buildings in the map and detect missing buildings. Again, large compact areas in the complementing depth map with significant height over ground are hints to recently constructed buildings. In the third case the given DEM might be updated with the derived depth map.

5 RESULTS

Rhein-Main airport in Frankfurt was chosen as test site. A map of the test area is shown in Figure 2a. The map was scanned from a printed map of scale 1:10000. The scene consists of parts of the airfield traffic net and several, mostly flat-roofed, extended buildings. Furthermore, a highway, forests and agricultural areas are present in the scene.

The test data set consists of intensity and height images (Fig 2 b,c) of the area, taken in November 1998, plus collateral data containing data acquisition and InSAR processing information. The sensor was carried by an aeroplane (Transall) in approximately 3000m height above ground. It operated in X-band with wavelength $\lambda = 3\text{cm}$, effective baseline $B = 1\text{m}$ and depression angle $\theta = 43^\circ$. The single pass ground range data has a resolution of $1\text{m} \times 1\text{m}$. Slant to ground range transformation was carried out by projecting the slant range data to a flat plane, not incorporating the height information of the InSAR data. The data was not geocoded.

According to equation 2 this sensor configuration results in an unambiguous (phase difference $\Delta\varphi < 2\pi$) height range of approximately 100m in mid swath. Therefore, no phase ambiguities had to be taken into account to calculate the height from the phase information in the comparatively flat scene with about 30m height difference. A maximum SNR of 100 (20 dB) is assumed in areas with high intensities in the interferogram. The lowest value for σ_i achievable is then about 2.1m according to equation 6.

The depth map in figure 2d illustrates the result of the segmentation and weighting process before the refinement step. The grey values of the segmented regions correspond to the averaged height. Some segments, like the airfield traffic net, are considered as unreliable because of low intensity values. But they were not masked for this visualisation to avoid holes in the depth map. Their height was set to the ground level taken from the map. To compensate over-segmentation, adjacent segments with height differences lower than 1m were merged. Compared to the original height data in Figure 2c the smoothed and averaged heights are more appropriate for visual interpretation.

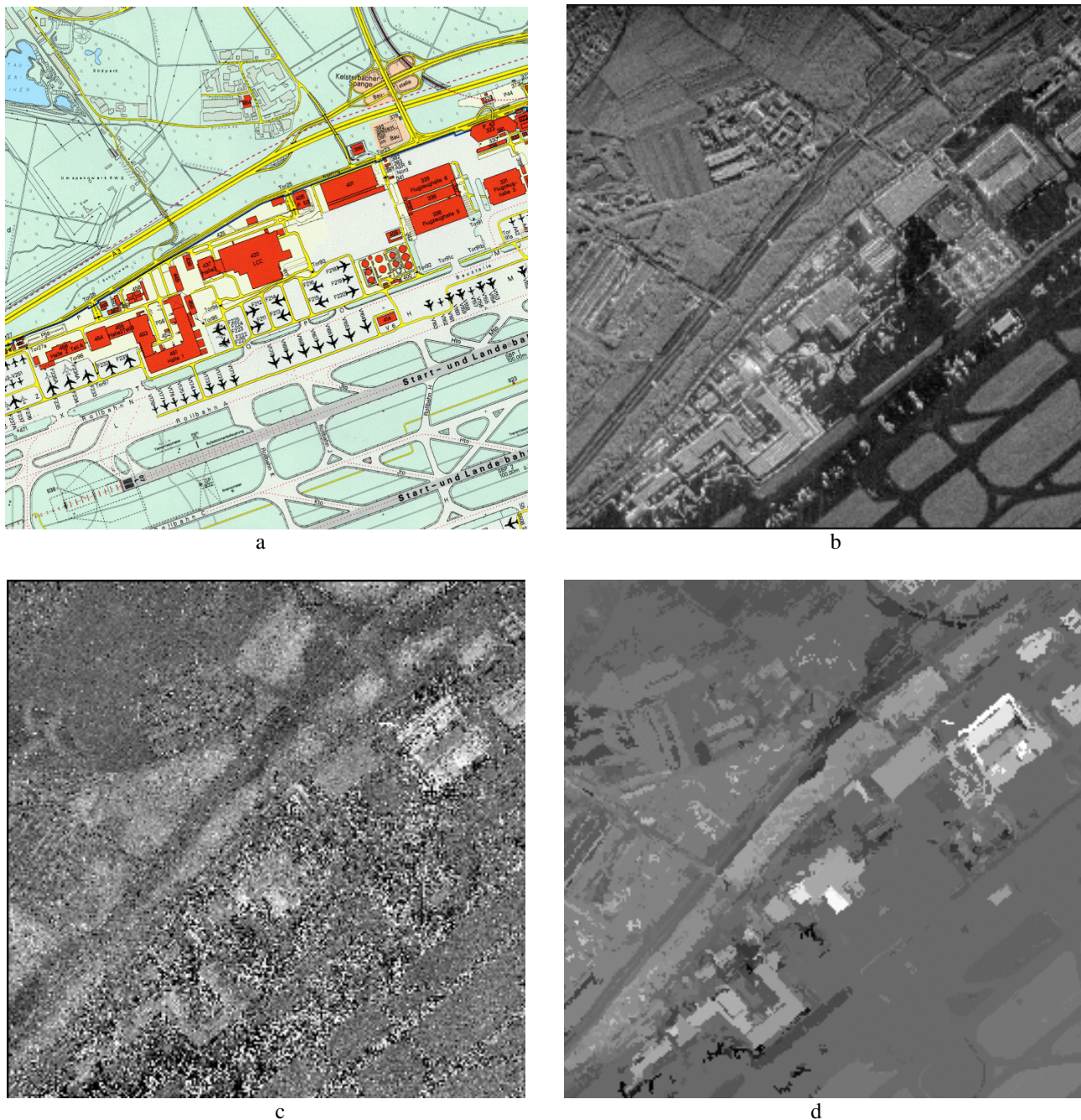


Figure 3: Scene Frankfurt Airport: a) Map, b) Intensity Data, c) Height Data, d) Depth Map

Several tie-points with height information can be extracted from the map. Only tie-points on the ground were chosen. Correspondences of these tie-points were found in the intensity data. At the related positions in the height data, the height values were accumulated in a histogram representation. The histograms showed gaussian distribution with a significant offset of the mean value of about plus 8 m in contrast to the map information. Hence, all calculated heights in the depth map were corrected by subtracting 8m. After the offset correction, the depth values of the large gassy-areas in between the airfield traffic net met about $\pm 2m$ the true heights.

The printed map was digitised. A vector map was manually extracted from the scanned map. It contains a layer of the generalised representation of the building footprints, restricted to the biggest buildings. The InSAR data is slightly rotated compared to the map. For the correction of the rotation and to avoid artefacts caused by foreshortening and layover, the vector map is aligned manually on the intensity image. From the aligned building layer of the vector map a binary mask is generated.

Figure 4 illustrates the process for a subset of the data in the lower left of the scene around the question-mark-shaped large building. To ensure that the buildings are covered entirely by the building mask, the mask is dilated by two pixel in each direction. The binary mask is used to separate the depth map in two complementing parts, a “*building expectation area*” and a “*non-building expectation area*”.

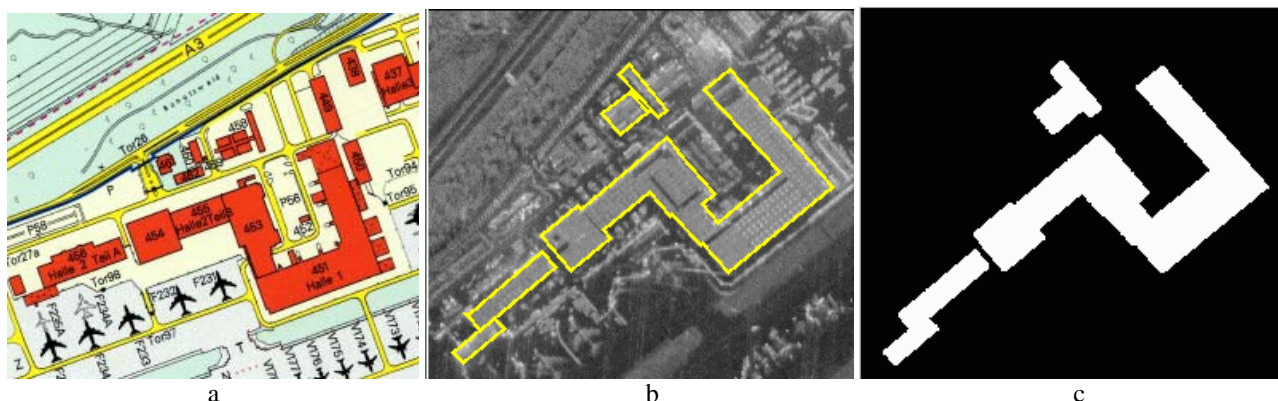


Figure 4: Generating a binary mask: a) Printed Map, b) Aligned Building Layer Superimposed on Intensity, c) Mask

In Figure 5a the depth map of the interesting scene detail is shown. Objects with a height of 5m above the ground are considered as possible buildings (5b). For a better discrimination of the segments their heights are pseudo-colour coded. A comparison with the map indicates that the elevated objects correspond indeed with the buildings and with groups of tall trees in the forests at both sides of the highway. The two complementing sets of building candidates are illustrated in Figure 5c,d.

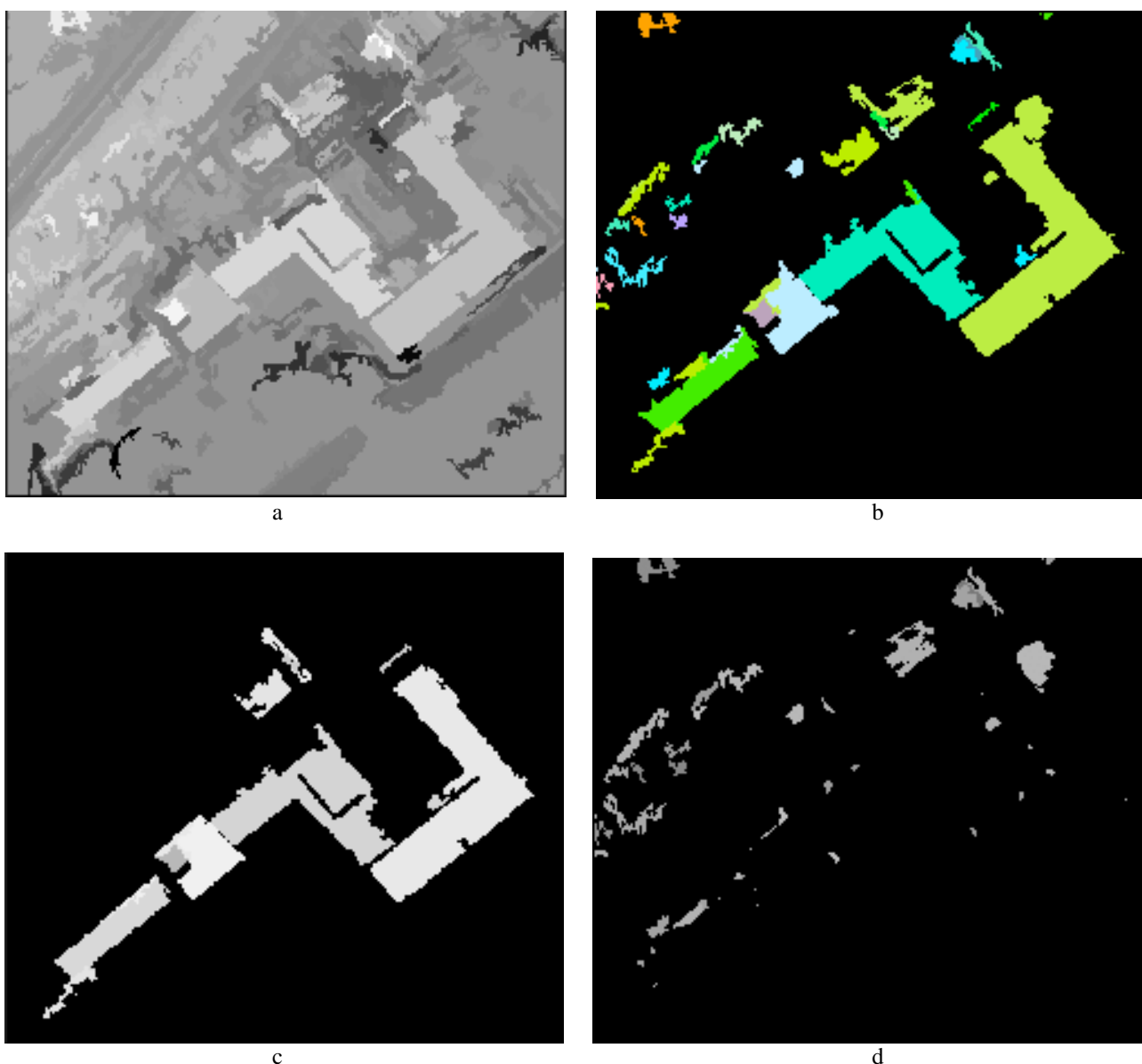


Figure 5: Scene Detail: a) Detailed Depth Map, b) Objects with Height, c) Masked Objects, d) Complement of c)

The presence of buildings in the scene can be verified in almost the entire “*building expectation area*”. Only a small part at the top of the largest building is missing, due to shadowing effects. However, the weighted means of the extracted object parts differed significantly from the truth heights taken from detailed building ground plans. The calculated heights tended to be about 5m below the expected values, even in segments which fitted almost perfectly the referring truth data. This seems to be surprisingly considering the better results mentioned above achieved for the extended grassy-areas inside the airfield traffic net. A possible reason for this effect is that the signal response from buildings suffers more from artefacts like multi-backscattering or layover. This behaviour is object of further studies. On the other hand, in the “*non-building expectation area*” a hint to a recently constructed building is found. It fits like a symmetric twin to an existing building in the middle on top of the scene detail. It came out that indeed the building was built up in the time period between the production of the paper map and the InSAR measurement campaign. The large region in the middle of the question-mark-shaped building contains a shadow region and was obviously under-segmented in the first iteration. In the refinement step this region is further analysed (Figure 6). The histogram shows no bimodal distribution. However, the distribution of the heights indicate the presence of two different height levels. In the region growing step, starting at the shadow area, two large regions can be segmented, besides some small regions nearby the shadow. In Figure 6e the final result is shown. The small regions are filtered out and the remaining two segments became filled.

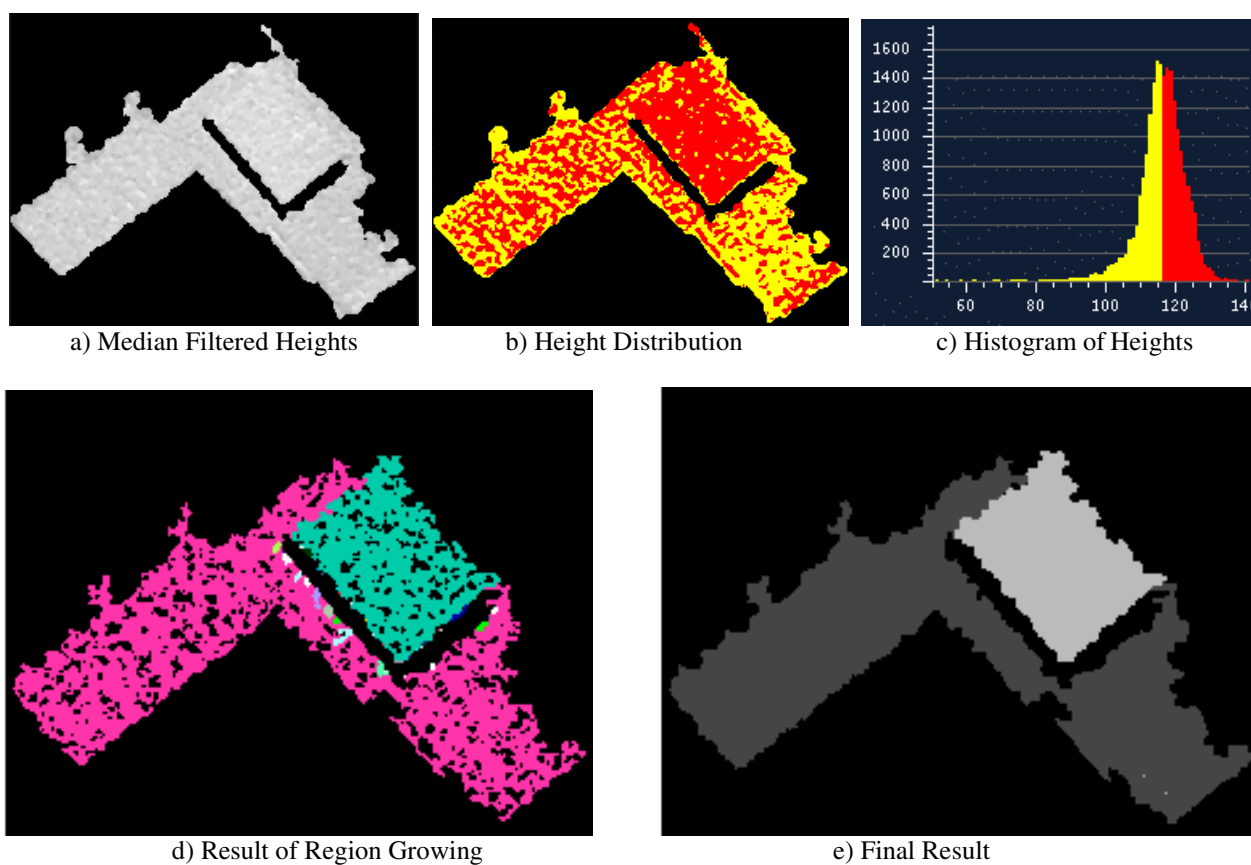


Figure 6: Histogram of Heights and Region Growing Result of a Scene Detail Containing Shadow

6 CONCLUSIONS AND FUTURE WORK

The approach of improving the interpretation of InSAR height information, using a model-based segmentation process in the intensity data, achieved promising first results. The visual impression is improved compared to the noisy original data.

After geocoding, the smoothed InSAR height data can be used for change detection purposes, at least for extended buildings. But the results indicate that the approach is limited to coarse scene descriptions only. For image based detail analysis, like roof reconstruction, the data is still too inaccurate. If a further improvement of the accuracy is achievable in the future, e.g. by better motion compensation incorporating more sophisticated INS, more detailed object models become appropriate. Gradient-based segmentation methods can be used for example to detect sloped objects like gable roofs. In the present state, the approach is restricted to flat terrain. In the case of undulated terrain, surrounding areas have to be examined to determine the elevation of buildings over ground.

In future investigations we will extend our approach to slant range data. The extracted objects in the InSAR data could be matched with referring objects in a given 3D description of the scene. Afterwards, a geometric and radiometric [Ulander, 1993] correction of the SAR intensity is performed.

The segmented regions might be incorporated in the phase unwrapping step. Shadow and layover areas, which correspond to high intensity and low coherence, could be masked. Only residues in the remaining data are considered.

ACKNOWLEDGMENTS

We want to thank Dr. Ender (FGAN-FHR Research Institute for High Frequency Physics and Radar Techniques) for providing us the InSAR image data. The data were recorded by the AER II experimental system of FGAN [Ender, 1998]. We also thank Mr Konetzki of the survey department FIP-P1/V of the Flughafen Frankfurt Main AG for kindly supplying us the paper map.

REFERENCES

- Bamler R, Hart P (1998) Synthetic aperture radar interferometry. *Inverse problems* 14(4): 1-54.
- Desnos YL, Matteini V (1993) Review on structure detection and speckle filtering on ERS-1 images", *EARSel Advances in Remote Sensing*, 2(2): 52-65.
- Ender JHG (1998) Experimental results achieved with the airborne multi-channel SAR system AER-II. *Proc. EUSAR'98*, 315-318
- Geibel R, Stilla U (2000) Segmentation of laser altimeter data for building reconstruction: different procedures and comparison, *This issue*
- Glazer F (1993) Morphological Image Processing in the KBVision System. Amerinex Artificial Intelligence, Inc., Technical Report
- Goldstein R, Zebker H, Werner C (1988) Satellite radar interferometry: Two dimensional Phase unwrapping. *Radio Science*, 23(4): 713-720.
- Hellwich O (1999) Basic principles and current issues of SAR interferometry. *ISPRS Joint Workshop: Sensors and Mapping from Space 1999*, No. 18, Publications of the institute for photogrammetry and engineering surveying of the university of Hannover, Hannover
- Hoepfner KB (1999) Recovery of building structure from IFSAR-derived elevation maps. *CMPSCI Technical report 99-16*, Amherst: University of Mass., Computer Science Department
- Levine MD, Shaheen SI (1981) A modular computer vision system for picture segmentation and interpretation. *IEEE Trans. PAMI-3* (5): 540-554.
- Rodriguez E, Martin JM (1992) Theory and design of interferometric synthetic aperture radar. *IEEE Proceedings-F*, 139(2): 147-159.
- Schreier G (1993) Geometrical properties of SAR images. In: Schreier G (ed.) *SAR Geocoding: Data and Systems*, Wichmann, Karlsruhe, pp. 103-134
- Schulz K, Thoennesen U (1999) Segmentation of detail structures of extended objects in SAR images. *EUSAR'00* (in press)
- Soergel U, Schulz K, Thoennesen U, Stilla U (2000) 3D-Visualization of interferometric SAR data. *Proc. EUSAR'00* (in press)
- Stilla U, Jurkiewicz K (1999) Reconstruction of building models from maps and laser altimeter data. In: P. Agouris P, Stefanidis A (eds) *Integrated spatial databases: Digital Images and GIS*. Berlin: Springer, 39-46.
- Ulander LMH, Hagberg JO (1993) Use of InSAR for radiometric calibration over sloping terrain. *Proc. of CEOS SAR Workshop*, 147-159.