7th Int. Conf. on Computer Analysis of Images and Patterns CAIP'97,
Kiel, Germany, Sep. 10-12, 1997

Architectural Image Segmentation Using Digital Watersheds

Volker Rodehorst

Technical University of Berlin, Photogrammetry and Cartography
Straße des 17. Juni 135, Sekr. EB 9, D-10623 Berlin, Germany

vr@fpk.tu-berlin.de

Abstract. The use of digital image analysis in architectural photogrammetry allows the automation of monotonous low-level tasks and makes the interpretation of facades much easier. Different segmentation methods for a computer assisted mapping of joint contours are discussed. Exemplary the application of the watershed transform is described and a combined watershed-method will be presented, which is well suited for the extraction of irregular stones.

1. Introduction

The purpose of Photogrammetry is to obtain geometric information of objects by taking measurements from images. One of the earliest applications of photogrammetry is the documentation of cultural monuments. A major objective is to document and to archive historically important objects in such a way, that the actual situation can be reconstructed at any time. Today, the acquisition of geometric information is mainly done by human operators who analyze images using classical photogrammetric equipment. Architectural Photogrammetry deals with spatial structured objects, which are in principle suitable for an automatic evaluation [1].

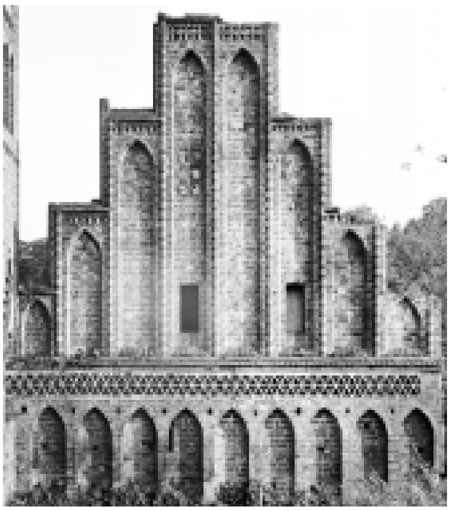

Fig. 1a: Photograph of the northern facade of the brewery house in the Chorin monastery

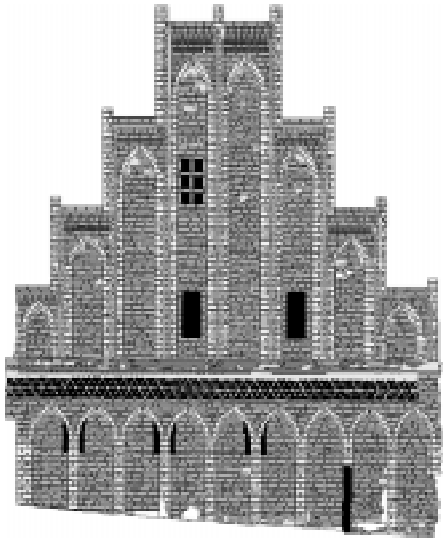

Fig. 1b: CAD-Model of the analytical measured facade

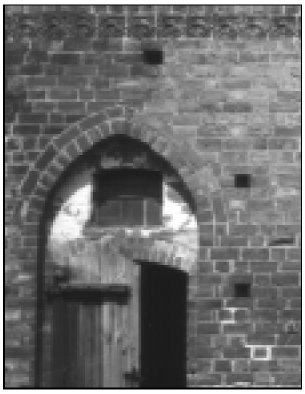

Fig. 2a: Section of the digital image from the southern facade in Chorin

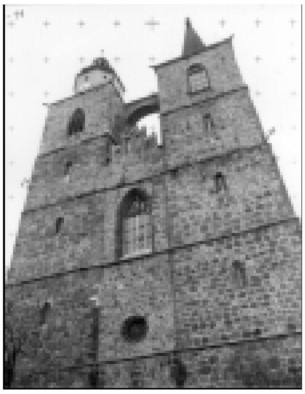

Fig. 2b: Photograph of the Towers of the Church St. Nicolai

Fig. 2c: Selected area of the digital image from the stone facade

The use of special feature extraction algorithms and the development of interpretation models for these tasks are in a preliminary stage until now. The focus of the described work is the computer assisted mapping of joint contours in architectural images (see Fig. 1a). The aim to automate the creation of a facade map (see Fig. 1b) from gray-level photographs can only be achieved using knowledge based image analysis.

For these studies, historical objects of Brandenburg are selected as test material. The Fig. 2b and 2c show the towers from the Church St. Nikolai in Jüterbog. The stones have different sizes, they have approximately circular but irregular shapes and they are arbitrary distributed in the image. This example of an image class is characterized by mainly dark stones, which stand out against the bright background of the extensive joints.

The second example shows a section of the brewery house facade in the Chorin monastery (see Fig. 2a). Here, a definite separation between the objects and a background is not obvious. Differences in the illumination, shadows and weathering make the analysis tasks difficult. The advantages for an automatic processing are given by the simple shape and the regular and symmetric arrangement of the objects.

2. Image Segmentation

The segmentation based on *discontinuities* searches for areas, where the intensity of the image values changes significant locally. The segmentation based on *homogeneity criteria's* joins pixel to areas of maximum size using a homogeneity predicate. It is not sufficient to search for areas with the same photometric characteristic. The recognition of semantic units, e.g. bricks and joints, is more important.

2.1 Point orientated Segmentation

A simple method to separate objects for a segmentation is the usage of *thresholding*. A lot of approaches are discussed in detail by WAHL [11]. Another approach for a point based segmentation is the usage of a *classificator*, which tries to form groups inside a set of feature vectors to fulfill an appropriate target criterion. The dependence on old gray-level images makes the use of a point orientated segmentation difficult.

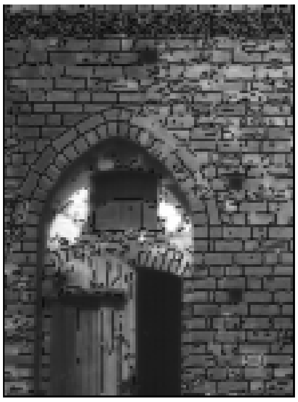

Fig. 3a: Superposition of an image with the result of a line extraction algorithm

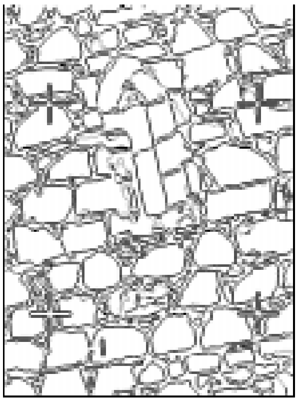

Fig. 3b: Result of an edge detection *(black)* followed by a contour tracing *(gray)*

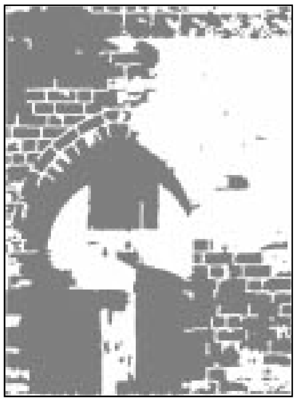

Fig. 3c: Result of an area based segmentation without preprocessing

2.2 Edge-/Line based Segmentation

The edge based segmentation tries to find borders between objects under the assumption that there is an abrupt change in the gray-levels (*optical edge*). Edge filters have different characteristics for the localization and the determination of the amplitude and direction. The resulting contours should be closed, straight and should have a small width. Additional criteria for the quality are the handling of contour crosses and the noise sensitivity of the filter. More information can be found by MARR-HILDRETH[6] and CANNY[3].

The detection of line structures requires neighbored edges with opposite direction. The result of a *line extraction* algorithm by matching of 2D-functions proposed by BUSCH [2] is shown in Fig. 3a. Problems may occur, because the extracted edge or line contours are not closed. The use of *contour tracing* approaches, which are able to close contour gaps using the gradient information, is only successful if the distance is very small (see Fig. 3b).

2.3 Area based Segmentation

The region growing techniques are based on similarities between neighbored pixel to find a connected area. Beginning from a starting point, features are compared with the neighbors to find pixel with the same characteristic (see Fig. 3c). An investigation for knowledge based image analysis of facades using *hierarchical region aggregation, region growing* and *quadtrees* is described by XU [12]. An advantage of this approach is, that edges which are defined by segment borders have closed contours. A problem is the automatic determination of the segment size. Often, some important details merge to one segment while other uniform areas are still splitted.

2.4 Segmentation Using Morphology

An important approach is to find objects using their shape, which requires that the shape is a-priori known. An application for the *matching* of bricks was described by TRAUTHAN [8]. This approach depends on the absolute gray-level intensity and is not very flexible, especially when the bricks are distorted by perspective projection.

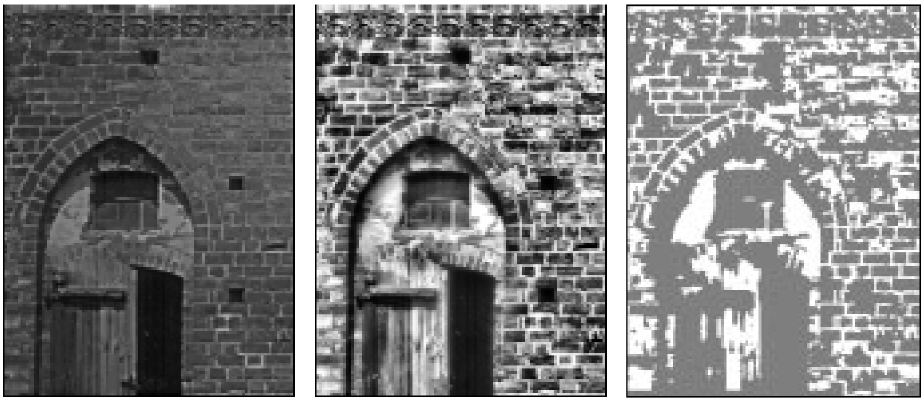

Fig. 4a: High-pass filtered image with the Butterworth-Filter

Fig. 4b: Result of the Dynamic Histogram-Equalization

Fig. 4c: Segmentation improvements after the pre-processing stage

3. Preprocessing

3.1 Homomorphic Filtering

Variations of the lighting or weathering in architectural images are characterized by slow oscillations in the digital image $f(x,y)$. The basic approach for enhancing an image via frequency domain techniques is to compute the *fast fourier transform* (FFT), multiply the result by a filter function $H(u,v)$ and take the inverse transform (IFFT). For sharpening we used the *Butterworth-Filter* of order n, which is defined as

$$H(u,v) = \frac{1}{1 + \left[D_0 / D(u,v)\right]^{2n}} .$$

$D(u,v)$ is the distance from the origin of the frequency plane and D_0 is the value, where $H(u,v) = 0.5$. The basic idea behind *homomorphic filtering* [7] is to operate separately on the frequency components for illumination and reflectance. A logarithm operation provides the measure of separation between the low- and high-frequency components. Finally, the filtering process is defined as follows

$$g(x,y) = \exp\left[\mathrm{IFFT}\left(H(u,v)\,\mathrm{FFT}\left(\ln f(x,y)\right)\right)\right].$$

The enhanced image $g(x,y)$ obtained by homomorphic filtering is shown in Fig. 4a.

3.2 Dynamic Histogram-Equalization

A local contrast manipulation can be achieved by modifying the histogram of intensities. The objective is to produce a transformation function $T(i)$ to take advantage of the full range of intensity levels $i = 0,...,G_{MAX}$. Let $n_L,...,n_H$ denote the given discrete intensities and $G(n_i)$ the frequency of intensity i in the image. The mapping function for a *histogram equalization* [11] is given by

$$T(i) = \frac{G_{MAX} \cdot \left(\sum_{j=n_L}^{n_i} G(j) \right) - \frac{G(n_i) + G(n_L)}{2}}{n} ,$$

Fig. 5a: Over-segmentation using the watershed transform

Fig. 5b: Erroneous merging of segments

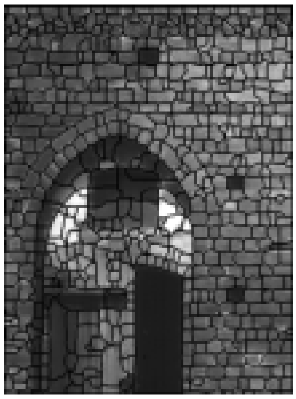

Fig. 5c: Result of the watershed smoothing

where n is the number of pixels in the image. Using the *dynamic histogram equalization* [10] the entire image is divided in several regions and the histogram equalization is computed for every region. In order to get a smooth transition at the boundary between two regions the final intensities are bilinear interpolated. Depending on the region size an optimal local contrast can be achieved (see. Fig. 4b).

4. Watershed Transform

The watershed transform is an approach based on mathematical morphology to divide an image due to discontinuities. In contrast to a classical area based segmentation, the watershed transform is executed on the gradient image. We obtain the gradient information using the algorithm proposed by DERICHE [4]. A digital watershed is defined as a small region that can not assigned unique to an influence zones of a local minima in the gradient image.

For applying the watershed transform, an algorithm proposed by VINCENT [9] was used. The intensity values are sorted and are stored in a table to achieve an efficient execution. The transform can be calculated without the gradient image using the table and neighborhood relations. A result of the watershed transform is shown in Fig. 5a.

4.1 Watershed Transform Using a Smoothed Gradient

It is evident that the segment sizes are very small and that uniform areas are splitted. The reason for the so-called *over-segmentation* is the sensitivity of the gradient to noise. An obvious approach is the smoothing of the gradient image, e.g. using a gauss filter. The noise will be reduced and the segments will grow depending on the *standard deviation* σ. The choice of a very large σ may lead to erroneous merged segments (see Fig. 5b). A segmentation result of the watershed transform using a smoothed gradient with an appropriate σ for the bricks is shown in Fig. 5c.

A problem introduced by the smoothing process are the shifted positions of the contours. For simple object structures a description by graphic primitives using a form factor may be sufficient (see Fig. 6a). The correction of object contours with an arbitrary shape is more difficult.

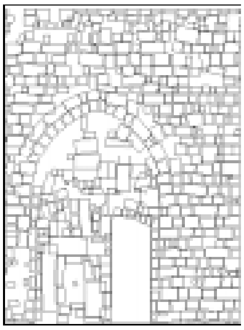

Fig. 6a: Description of the segments using graphic primitives

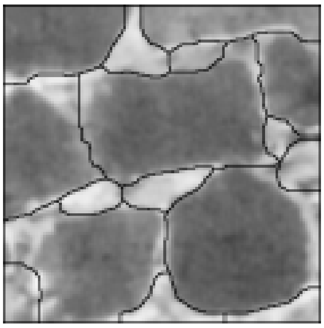

Fig. 6b: Shifted contours of the watershed transform using a smoothed gradient

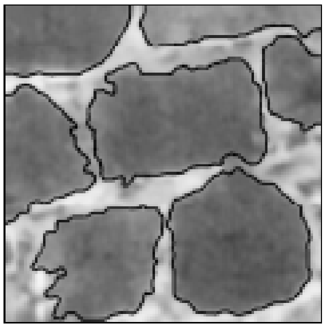

Fig. 6c: Result of the optimization process using Active Contours

4.2 Shape Analysis Using Active Contours

In order to find the real contour of a stone using the active contour model it is important to specify an initial curve. A rough approximation to the desired contour in the image is supplied by the watershed algorithm that also defines an outer boundary (see Fig. 6b). An active contour with N points is represented by $s_i = (x_i, y_i)$, i=1..N. The optimization will minimize the contour's energy based on the *energy function*

$$E_{TOTAL} = \sum_{i=1}^{N} E_{INT}(s_i) + E_{IMG}(s_i) + E_{CON}(s_i)$$

The *internal energy* E_{INT} consists of the term α, which affects the rigidity of the active contour and the maximum change in the tangent vectors of the curve. Additionally, the term β minimizes the distance between contour points and favors the shortest contour length. The *external* forces are defined by the *constraint energy* E_{CON} and the *image energy* E_{IMG} of a contour. E_{CON} attracts points on the active contour to points in the image and E_{IMG} describes how the contour conforms to the gray-level gradient within the image. Details concerning active contours can be found by KASS et al. [5].

The result obtained by the energy minimization using a dynamic programming technique is shown in Fig. 6c. A major advantage of the active contour model is the retention of the viscosity feature of the watersheds. On the other hand, it is not easy to find appropriate parameters and the optimization process is not very robust.

4.3 Merging of Watershed Regions

Another approach to prevent the over-segmentation is the merging of small segments using a fusion criterion. By merging of neighbored segments, erroneous watersheds should be eliminated (see Fig. 7a). After applying the watershed transform this criterion is calculated for every region. The numbers are normalized to the maximum gray level G_{MAX} and the regions are filled with these values.

The resulting *mosaic image* can be used for another watershed transform during an iterative process (see Fig. 7b). The merging criteria is based on region properties (i.e. texture features) and the gradient amplitude between the regions. Using the region mean value, the quality is comparable with the results obtained by a classical region-growing approach (see Fig. 7c).

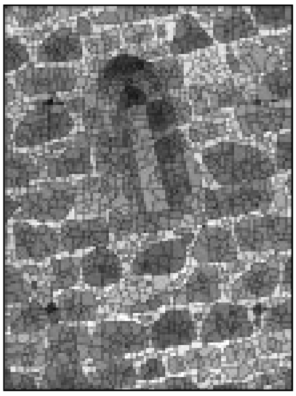

Fig. 7a: Over-segmented result of the watershed transform

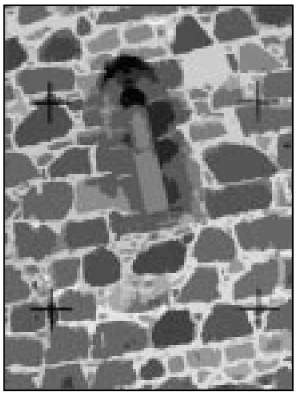

Fig. 7b: Merged watershed regions using the region mean value

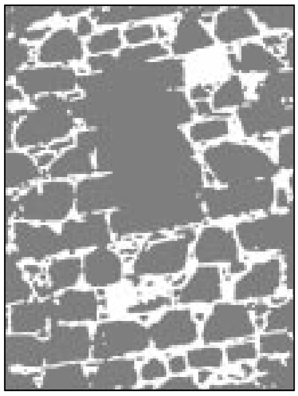

Fig. 7c: Thresholding of the merged watershed regions

4.4 Combination of Watershed Algorithms

Using the watershed smoothing, most of the stones are separated but with shifted segment contours (see Fig. 8a). The mosaic image of the watershed merging process contains the correct contours but several objects are joined erroneously (see Fig. 7c). A combination of these approaches is able to separate objects from a background. First, both algorithms have to be applied to the input image. Using a logical operation the erroneous joined results of the watershed merging can be splitted by the result of the watershed smoothing. A superposition of the joined segments with the shifted watersheds is shown in Fig. 8b. After the separation of the stones a postprocessing stage is necessary. The elimination of remains may be done by a *morphological opening*, which is defined as an erosion followed by a dilation. The final result of the combined watershed algorithm applied to irregular stones is shown in Fig. 8c.

5. Conclusions

Different approaches were presented to support the construction of a facade map by using digital image analysis. We can summarize, that the watershed transformation is well suited for the segmentation of facades. The independence on the absolute gray-levels by using the gradient image and the viscosity while expanding the segments are major advantages. For the known over-segmentation problem different solutions were shown. The smoothing of the gradient image leads to a good separation of relevant structures. For an exact determination of the contour the shifted segment borders have to be revised. Exemplary the active contour method was introduced, which seemed not robust enough for practical use. By merging small watershed regions the original contour position is preserved. The choice of a fusion characteristic may cause a dependence on the absolute gray-levels, which leads to erroneous results. Therefore, appropriate preprocessing techniques for architectural images were shown to improve area based techniques. Finally, the presented combination of the watershed algorithms is able to separate objects from a background and is well suited for the segmentation of irregular stones.

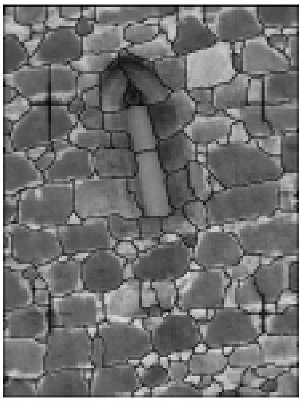

Fig. 8a: Separation of stones using the watershed transform on a smoothed gradient

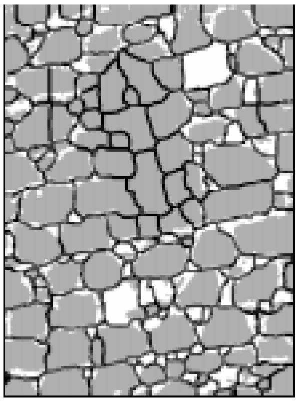

Fig. 8b: Superposition of the merged segments *(gray)* with the watersheds *(black)*

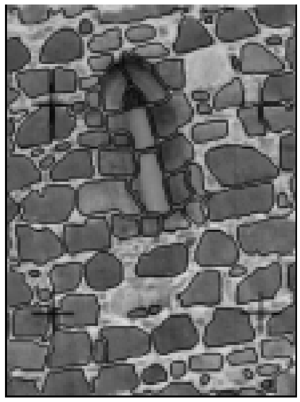

Fig. 8c: Segmentation result for irregular stones using the combined technique

Acknowledgment

This work was funded by the German Research Foundation (*DFG*).

References

1. ALBERTZ, J.: *Invarianten und ihre Bedeutung für Photogrammetrie und Fernerkundung*, Vermessung, Photogrammetrie, Kulturtechnik 83, pp. 295-299, 1985.
2. BUSCH, A.: *Fast Recognition of Lines in Digital Images Without User-Supplied Parameters*, Spatial Information from Digital Photogrammetry and Computer Vision, International Archives of Photogrammetry and Remote Sensing 30, Part 3/1, pp. 91-97, 1994.
3. CANNY, J.: *A Computational Approach to Edge Detection*, Pattern Analysis and Machine Intelligence 8, No. 6, pp. 679-698, 1986.
4. DERICHE, R.: *Using Canny's Criteria to Derive a Recursively Implemented Optimal Edge Detector*, Int. Journal of Computer Vision 1, No. 2, pp. 167-187, 1987.
5. KASS, M., WITKIN, A. AND TERZOPOULUS, D: *Snakes: Active contour models*, Int. Journal of Computer Vision, Band 1, No. 4, pp. 321-331, 1988.
6. MARR, D. AND HILDRETH, E.C.*: Theory of Edge Detection*, Proc. of the Royal Society of London B207, pp. 187-217, 1980.
7. STOCKMAN, T.G.: *Image processing in the context of a visual model*, Proc. IEEE 60, No. 7, pp. 828-842, 1972.
8. TRAUTHAN, F.: *Versuche zur automatischen Erfassung von Steinen einer Hausfassade,* Hrsg.: F.K. List, Publikation zur 13. DGPF-Jahrestagung: Geoinformation durch Fernerkundung, Band 2, Augsburg, pp. 221-228, 1993.
9. VINCENT, L. AND SOILLE, P.: *Watersheds in Digital Spaces: An Efficient Algorithm Based on Immersion Simulation*, IEEE Trans. on Pattern Analysis and Machine Intelligence 13, No. 6, pp. 583-598, 1991.
10. VOSSEPOEL, A.M., STOEL, B.C., MEERSHOEK, A.P.*: Adaptive Histogram Equalization Using Variable Regions*, 9. Int. Conf. on Pattern Recognition, IEEE, pp. 351-353, 1988.
11. WAHL, F.M.: *Digitale Bildsignalverarbeitung*, Nachrichtentechnik 13, Springer, 1984.
12. XU, YONGLONG: *Untersuchung der Bildsegmentation zwecks der nachfolgenden wissensbasierten Bildanalyse*, ISPRS Spatial Information from Digital Photogrammetry and Computer Vision, Munich, pp. 931-938, 1994.