

TWIN SNAKES FOR DETERMINING SEAM LINES IN ORTHOIMAGE MOSAICKING

Martin KERSCHNER

Vienna University of Technology, Austria
Institute of Photogrammetry and Remote Sensing
mk@ipf.tuwien.ac.at

KEY WORDS: orthophoto, radiometric quality.

ABSTRACT

As a last step in an orthophoto production process, called mosaicking, neighboring and partly overlapping images of a scene are merged to one image. This should be done in a way that the transition from one to the next image cannot be seen. The production line of orthophotos consists of several steps, each of which can introduce differences in overlapping orthophotos. For mosaicking these orthophotos, a seam line has to be delineated through the overlap area. In this paper, criteria for such an optimal seam line in color orthophotos are elaborated. The main requirements are maximum color similarity of the images (mainly in hue and intensity), and maximum texture similarity. Especially those areas are avoided, where many edges in radial direction can be found in one image, because these edges have another orientation in the neighboring image. The specified criteria are formulated in the energy function of snakes. Two snakes that attract one another (twin snakes) are initialized in a hierarchical strategy for detecting the seam line. The potential of the method is shown on an example.

1 INTRODUCTION

Orthophotos have become a popular planning instrument and a visualization document. They combine the rich information content of photographs with the geometric properties (ground projection) of maps. Thus, they can easily be combined with additional information from plans and maps to an orthophoto map.

An orthophoto production line consists of several processing steps. First, a digital terrain model (DTM) has to be produced. Often the DTM is derived from the same photographs that are used in the second step for producing orthoimages. Using the DTM, the photos are rectified geometrically with respect to a global ground coordinate system by resampling them to new images. The pixel size of the new images has to be specified according to the resolution of the photographs and the final scale of the orthophoto map. The accuracy of rectification strongly depends on the quality of the DTM. Those details of the surface, that are smoothed or even not contained in the generalized DTM, cannot be rectified.

Large areas can usually not be covered by a single orthophoto in the requested scale and resolution. Instead, many orthophotos are produced from a series of aerial photographs. Finally, neighboring and partly overlapping images of a scene are merged to one orthophoto mosaic. If the orthophotos have been rectified with respect to the same ground coordinate system, no geometric errors occur when mosaicking. Difficulties arise in balancing the colors of neighboring orthophotos with the goal that the transition from one to the next image cannot be seen any more.

This mosaicking step has not received much attention in the literature. In computer vision literature about mosaicking, the emphasis lies on finding the geometric transformations between the images to be arranged (e.g. Szeliski, 1996). This task is solved for orthophoto mosaicking, as the transformation of the photos to a common coordinate system was determined in advance by photogrammetric image rectification. According to Baltsavias & Käser (1998), problems remaining in photogrammetric systems concern mainly the radiometric equalization of orthoimages during mosaicking and a convenient, automatic determination of the seam lines.

In this paper, a method is presented for solving the second problem, detecting suitable seam lines in aerial color orthophotos for arranging them to a larger orthophoto mosaic. In the next section, we try to specify and formulate criteria for optimal seam lines. In section 3, these criteria are integrated in an energy function for making snakes detect optimal seam lines. A practical, hierarchical solution of delineating the seam line by twin snakes is described. Section 4 presents results.

2 CRITERIA FOR OPTIMAL SEAM LINES

Some parts of the scene, for which an orthophoto mosaic is created, are covered by more than one photograph. For these parts, separate orthoimages are produced independently. For arranging these orthoimages to an entire mosaic, a seam line has to be determined. Unfortunately, these orthoimages differ from each other more or less in color and texture. An optimal seam line is placed in regions of maximum similarity between the images. In the following, we want to find out explanations for the differences appearing when orthoimages are composed from different photographs. Further, we try to draw the conclusions for detecting an optimal seam line in color images.

2.1 Reasons for Differences in Orthoimages

The influences described below are ordered chronologically concerning the production line of orthophotos from the time when the photo is taken to the resampling step.

Changing object and illumination over time: Multiple aerial photographs can hardly be taken synchronously. The aircraft flies in form of meanders over the region. Photos are taken one after the other. The time interval between two neighboring photos in the same flying strip is quite short, but the time interval between two neighboring strips can be some minutes. If photos from only one flight are taken, the changes of the natural surface itself in such a short time can be ignored. Illumination usually doesn't change, either, except in case of partly sunny and partly cloudy weather, because shadows of clouds on the surface can move significantly further. The differences can be found mainly in the intensity channel of the image.

Different remission rates of the surface for different viewing directions (Kraus, Schneider, 1988): The amount of light reflected by the natural surface depends on the orientation of the surface and the position of the light source, as well as on the type of the surface (Horn, 1977). The natural surface is not a Lambertian diffuser which would have a constant remission in all directions. Instead, the surface looks differently bright from various viewing directions. The surfaces can even have some specular or mirror-like reflection producing hot spots in the image. The more diffuse the illumination when taking the photo, the less this effect appears. Therefore good conditions for flying photos for producing orthoimages are under a closed cloud cover. However, the contrast in such photos is low. That's why the flight is usually done under sunny weather conditions.

In the northern hemisphere, the light source is somewhere in the south. Thus, the northern part of the photo (where viewing direction and illumination direction are similar) appears brighter than the southern part. Two reasons can be found for that. In the southern part of the photo, the viewer looks on the shaded side of the surface and furthermore more shadows themselves can be seen. Before resampling the photos, a global trend of color intensity fall-off from north to south should be corrected as much as possible. However, differences will always remain. The greatest discrepancies of two neighboring orthophotos occur when the southern part of the first is to be merged with the northern part of the second one.

Lens effects: The light falling on the camera has to pass through the lens system before exposing the film. On the way through the lenses, a part of the light is absorbed. The absorption rate is low in the center of the photo, when the imaging ray is parallel to the optical axis of the lens system. It increases with a function of the cosine of the angle between imaging ray and optical axis (Kraus, Schneider, 1988). Thus, the color intensity decreases towards the image borders. For wide-angle lenses this effect is even larger than for normal-angle lenses. Modern lens systems reduce the effect. The radial intensity fall-off can be modeled and approximately corrected in the photographs.

Separate photo development and scanning: Although aerial photographs of the same flight are usually developed together and scanned with the same parameters, slight color differences may occur. Analogously, further image processing should always be applied with the same parameters. Fortunately, modern photographic processes and scanners are quite stable.

Deficiencies in the DTM used for geometric rectification: For the resampling process from the aerial photograph to an orthoimage (geometric rectification), the three-dimensional coordinates of each pixel of the orthoimage have to be known. The orthoimage is defined on a regular grid with respect to a global ground coordinate system. A square unit of this grid corresponds to one pixel of the orthoimage. Thus, the planar coordinates of the center of each pixel are given. The third coordinate (the height) is derived from a DTM. Given the x and y coordinate, the z coordinate is interpolated. The orientation parameters of the aerial photograph allow transforming this point into the local image coordinate system of the photograph. In this way, the regular grid of pixels of the orthoimage is transformed to an irregular grid in the photograph. Finally, for each point of the transformed grid, the color value for the orthoimage is selected from the photograph (Kraus, 1993).

The demands on the DTM depend on the scale of the orthophoto. Details of the surface that are smoothed in the DTM produce location errors in the orthophoto. Thus, a DTM with break lines should be used. If objects on the surface (e.g. buildings, trees) are not included in the terrain model, they are disregarded in the image rectification so that they appear turned down in the orthoimage (cf. figure 1). The amount of displacement ΔR depends mainly on the height above the modeled terrain ΔZ , the focal length c of the camera, the ground slope α in radial direction, and increases with the distance ρ' from the principal point. In a photo with vertical viewing direction, there is:

$$\Delta R = \frac{\Delta Z}{\frac{c}{\rho'} + \tan \alpha} \quad (1)$$

All points above the DTM are displaced in radial direction from the principal point of the photo towards the border. Consequently the directions of displacement are different across the overlapping images, which can be seen in figure 2. If two images are put together along a seam line that is chosen in urban or wooded areas, these geometric errors cause striking effects, because in the vicinity of the seam line, the objects appear turned down in contrary directions. Some of the objects turned down may even be cut. Thus, the optimal seam line goes through areas, where no objects are above the modeled surface. If not possible, the optimal seam line goes at least through areas, where the displacement is low, i.e. near the principal point.

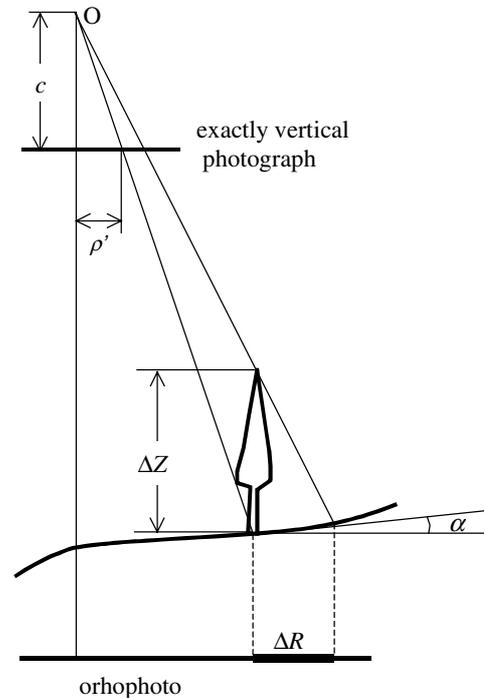


Figure 1: Radial displacement of details (Kraus, 1993).

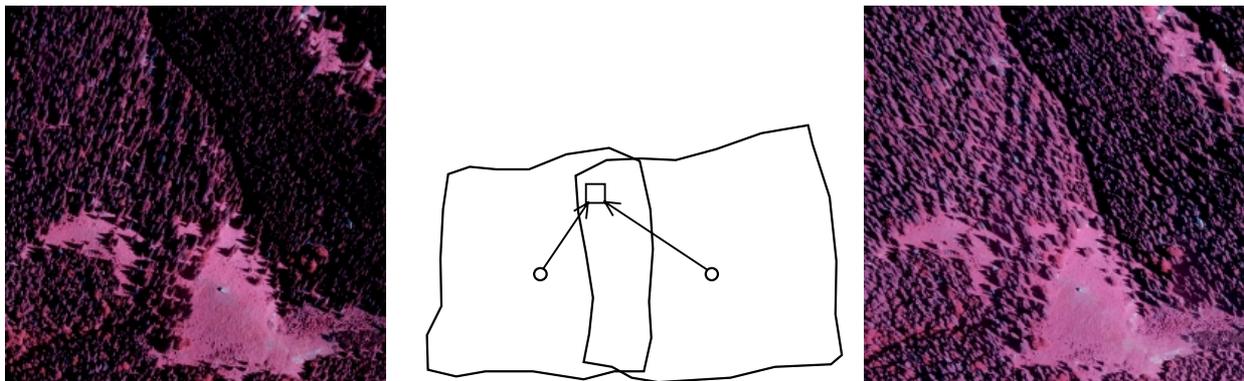


Figure 2: Parts of two neighboring orthophotos, where trees are turned down in contrary directions.

Flat view on hillsides: An orthophoto is produced by resampling the original photo. As explained in the previous paragraph, the regular grid of the orthophoto is transformed to an irregular grid in the aerial photo. These grid points do not coincide with the pixel centers of the photograph. Usually, the color value for the pixel of the orthophoto is determined by the color of the nearest pixel in the photograph (nearest neighbor). Another resampling method would be bilinear interpolation. Some constellations of local surface and aerial photograph are not well suited for resampling. If many of the transformed grid points fall into the same pixel of the original photo, this pixel is blown up in the orthoimage and the region appears blurred. An example can be seen in figure 3. Such constellations again occur near the image borders for hillsides whose surface normals meet the viewing direction of the photograph at an angle near 90 degrees. In other words: constellations, where the camera looks at a hillside in a flat view. Fortunately, such regions appear sharp in orthoimages originating from neighboring photographs. The optimal seam line is placed in a way, that the sharp orthoimage is taken for the mosaic. Again, it avoids regions near the image border.

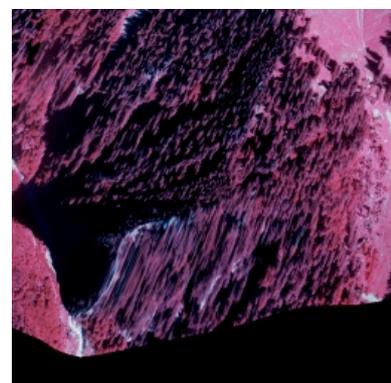


Figure 3: Blurred orthoimage region.

2.2 Criteria for Regions of Maximum Similarity

Drawing the conclusions for each of the factors causing image differences, we see that the similarity often depends on the distance of a pixel from the principal point of the photograph (north-south intensity fall-off because of non-constant remission rate, radial fall-off because of lens effects, displacement because of deficiencies in the DTM, and blurred regions because of flat view). The quality of each of the orthophotos is best in the vicinity of its principal point. On the other hand, this region is far from the principal point in neighboring orthophotos (or even outside). A compromise has to be found to be as close as possible to both principal points. From this point of view, the perpendicular bisector between the two principal points in the ground coordinate system builds the optimal seam line (figure 4).

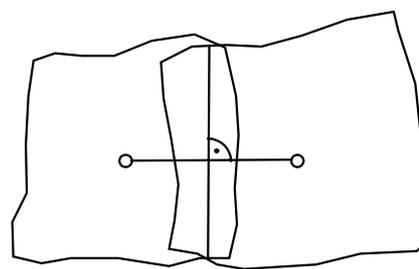


Figure 4: Perpendicular bisector of principal points.

Color differences can be divided into a global trend that can be modeled by an analytic function and into a local irregular part. Although the differences caused by the trend can be balanced smoothly, the optimal seam line should try to find regions with maximum color similarity. In digital images, colors are coded by three values, representing in general the channels red, green and blue (RGB). False color infrared photos (IR-photos) map the remission in the near infrared, red and green band to the channels red, green and blue. The RGB space is not well suited for calculating differences. Colors are better compared in the IHS space (intensity, hue and saturation of the color). As we found in the first paragraph of the previous chapter, moving cloud shadows cause significant disparities mainly in the intensity channel. An optimal seam line avoids regions, where the absolute value of the intensity difference is great. With $\Delta I = I_2 - I_1$ (the difference of the intensity values of the corresponding pixels in the overlapping orthoimages), one criterion that should be respected is:

$$|\Delta I| \rightarrow \min \tag{2}$$

Another channel, where differences may occur, is the color hue. The color hue is measured as angle on the color wheel. For formulating a criterion that minimizes the difference in hue, we have to take into account that the difference of orientations on the color wheel is calculated:

$$|\Delta H| = \begin{cases} |H_2 - H_1| & \forall H_1, H_2 : |H_2 - H_1| < 180^\circ \\ 360^\circ - |H_2 - H_1| & \forall H_1, H_2 : |H_2 - H_1| > 180^\circ \end{cases} \tag{3}$$

Balancing an image in hue changes its impression tremendously, especially in medium intensity range. On the other hand, the hue is not significant for black (intensity I_b) and white (intensity I_w). The hue difference can be scaled by the following factor s reflecting the full influence for medium intensity I_m and less influence in darker and brighter areas:

$$s = 1 - \frac{|I - I_m|}{I_m} \quad \text{with} \quad I_m = \frac{I_b + I_w}{2} \tag{4}$$

Seam lines should be attracted to regions, where:

$$|\Delta H| \cdot s \rightarrow \min \tag{5}$$

Dissimilarities in the images can be found not only in the color domain, but also in the texture domain. Haralick & Shapiro (1992) summarize approaches to texture analysis of the past. A formal approach or precise definition of texture does not exist. In orthophotos of a scene, the natural texture is more or less the same in the different images, except for the texture caused by deficiencies in the DTM. Especially in wooded or in settled areas, the objects are turned down and generate much texture in radial direction. As the radial direction is different in the neighboring image, texture discrepancies arise.

In order to separate this radial texture from the common natural texture, edge maps are calculated and subtracted. For each of the orthophotos, a gradient image is calculated from the overlap area. Both components of the gradient, its magnitude and its orientation are stored. With F_x and F_y the first derivatives in x- and y-direction of the image function F , the magnitude m and the orientation o are given by:

$$m = \sqrt{F_x^2 + F_y^2}, \quad o = \arctan\left(\frac{F_y}{F_x}\right) \tag{6}$$

Edges in the image produce high gradient magnitudes. Thus, the magnitude value can be called "edginess" of a pixel. The orientation of the gradient is perpendicular to the edge direction. In figure 2, a forest in two neighboring IR-orthophotos can be seen. The gradient orientations were divided into eight classes with a width of 25 gon each and for each of these classes the gradient magnitudes were accumulated. In figure 5, the result is drawn in form of a histogram in the respective edge direction (perpendicular to the gradient orientation) as dashed line.

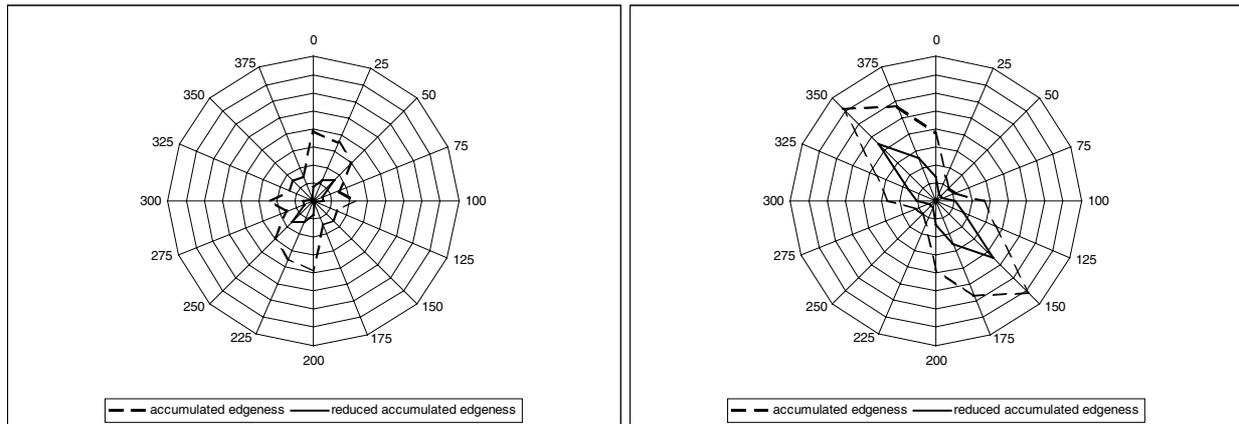


Figure 5: Gradient magnitude histogram divided in 8 orientation classes (plus the eight symmetric classes) for the images of figure 2. The solid line is the histogram reduced by the gradients of the other orthophoto.

The solid line shows the histogram of edginess remaining after subtracting the respective edginess of the neighboring orthophoto. For each pixel, the gradient orientation is compared to the orientation in the other image. If the orientations are the same, the magnitude of the other image is subtracted. For reasons of robustness against numerical problems, the orientation of the neighboring image was allowed to be in a range of ± 25 gon to be accepted as equal. The reduced magnitudes are defined as:

$$m_{1r} = \begin{cases} m_1 & \forall m_1, o_1, m_2, o_2 : o_2 \notin [o_1 - 25^g, o_1 + 25^g] \\ m_1 - m_2 & \forall m_1, o_1, m_2, o_2 : o_2 \in [o_1 - 25^g, o_1 + 25^g] \wedge m_2 < m_1 \\ 0 & \forall m_1, o_1, m_2, o_2 : o_2 \in [o_1 - 25^g, o_1 + 25^g] \wedge m_2 > m_1 \end{cases} \tag{7}$$

$m_{2r} = \dots$

Furthermore, not only the pixel in the same location of the neighboring photo is taken for subtraction, but also neighboring pixels in a 3x3 neighborhood in order to compensate geometric discrepancies. In the resulting solid histogram, a sharp peak dominates in the respective radial direction. A second peak which occurred in east-west direction of both images could be reduced. This peak was mainly due to the horizontal shadows of single trees, which are the same in both images. Based on the results of this example, the third criterion can be formulated:

$$m_{1r} + m_{2r} \rightarrow \min \tag{8}$$

Unfortunately, the measures used in the criteria (2), (5) and (8) are not independent. Instead, they formulate redundant information. Especially a high texture difference will produce high intensity difference and hue difference, too. Nevertheless, a seam line that minimizes all three criteria simultaneously and which is as close as possible to the perpendicular bisector of the two principal points, can be seen as optimal solution.

3 MINIMISATION BY SNAKES

The results of the considerations of the previous chapter are now integrated in the energy function of snakes. We found four main criteria for optimally delineating a line through the overlap area of two orthophotos. In this chapter an

algorithm is presented for automatically detecting this seam line based on these criteria. The algorithm makes use of two snakes attracting each other (twin snakes). After a short introduction into the basic concept of snakes, and its extension to twin snakes, the above criteria are formulated in a special energy function for seam line detection.

3.1 Twin Snakes Concept

Snakes (Kass et al., 1988) are active contours for linear feature extraction which were originally designed as a semi-automatic measurement tool. Starting from an initial estimate of a curve's shape and location the snake wriggles through an image in an iterative process. This evolution of the snake (the change of its shape and position) is driven by minimization of an energy function that is composed from several terms of different nature: internal terms E_{int} , photometric terms E_{pho} and external forces E_{ext} . The energy is calculated for each vertex $v(s)$ and integrated over the whole length of the snake:

$$E_{snake}^* = \int_0^1 E_{snake}(v(s)) ds = \int_0^1 (E_{int}(v(s)) + E_{pho}(v(s)) + E_{ext}(v(s))) ds \quad (9)$$

The internal energy tries to preserve a smooth shape of the curve. Photometric terms usually evaluate edge strength or similar measures in the examined image and try to pull the snake to salient image features. External forces can be introduced by user interaction and are responsible for global controlling and guiding the snakes evolution.

Twin Snakes (Kerschner, 1998) are an extension to the traditional concept, in which attraction forces are introduced as external forces. Two snakes approach to the desired image feature from opposite sides. Their initial states can lie far beside the resulting curve as long as the desired curve lies in between. During their evolution, the two snakes attract one another and move towards one another while always embracing the line of minimum energy. If one of the two snakes has reached a significant local minimum regarding internal and photometric energy, it is captured there, because the attraction forces are too low to pull it away. Only the second snake moves further until it reaches a local minimum, too. If the desired distance of the two snakes is already reached, the process terminates. Otherwise, single curve segments, that have not yet reached the desired distance are pulled out of their local minima, and the iterative process is continued until all segments fulfill the termination requirements.

3.2 Twin Snakes for Detecting a Seam Line

The twin snakes are now modified for detecting an optimal seam line in low and equally textured regions with similar colors in the participating images. According to the first optimality criterion of chapter 2, the snakes should be pulled towards the perpendicular bisector of the principal points. This can implicitly be integrated in the twin snakes concept by initializing the two snakes symmetrically to the bisector at (or slightly outside) the borders of the overlap area. The desired distance of the two snakes is set to zero, i.e. the minimum external energy is reached when the two twins coincide. The attraction force of the two snakes to one another let them move closer in the energy minimization process. If there were no photometric forces, the iterative process would end in the bisector.

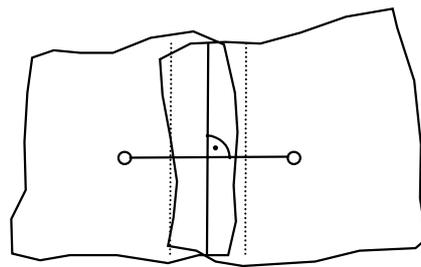


Figure 6: Initial state of twins (dotted).

The remaining criteria are integrated in form of separate energy terms. Three terms are used when calculating the current energy, one for the difference of intensity E_I , one for the normalized difference of hue E_H and the third for the texture difference E_T . For reasons of efficiency, for each pixel of the overlap area the combined energy is calculated in advance:

$$E_{pho} = a E_I + b E_H + c E_T \quad (10)$$

Finding the appropriate weighting coefficients a , b and c from the three photometric energy terms is a more or less empirical process. In the examples presented later, all three image terms were weighted equally and the combined term E_{pho} was weighted relatively high in comparison to the internal terms and to the attraction force.

3.3 Hierarchical solution

The energy minimization process is performed hierarchically. First, an image pyramid is calculated from the image representing the photometric energy terms. A reduction factor of 5 is used with an average filter of extent 5x5 pixels, i.e. 25 pixels are averaged and build one pixel of the next higher level. In the highest level, the twins are initialized

according to figure 6. Each pixel corresponds to one vertex of the snake. The vertices are then allowed to move only horizontally or only vertically, depending whether the snake has a vertical or horizontal orientation, respectively.

For performing the minimization a dynamic programming algorithm is used (Kerschner, 1998). After a few iterations, both snakes have reached a local minimum. At this point, the attraction forces are switched off, so that only photometric and internal terms participate in the energy calculation. Further iterations are performed and the total energy value of the resulting snakes are stored for later comparisons.

If segments of the snake do not yet coincide, these segments are first pulled out of the minimum and then further iterations are calculated. In this way other local minima may be detected. If the total energy of the snake is lower than the stored one, a "better" minimum has been found. The best result is taken for deriving the initial state of the snakes in the next higher level of the image pyramid. The snake is moved backward and forward by a few pixels in the lower resolution corresponding to five times more pixels in the higher resolution. This procedure is repeated fully automatically for all levels of the pyramid until the final solution is found on the lowest level.

In that way, a path through the overlap area is found, where the color and texture of the two images are similar. The still remaining jumps in hue and the differences in intensity and saturation have to be leveled by smooth interpolation in the vicinity of the seam line. After having processed all overlapping areas, the results can be merged and the orthoimages can be arranged to the image mosaic.

4 RESULTS

As example, a representative pair of IR-orthophotos of a larger project from the Schneealpe is shown. It is a mountainous region in the north-east of the Austrian Alps. The overlap area is north-south oriented. In the northern part, forests are dominating. The central part is covered by alpine meadows and crossed by a road. In the south, a valley starts and goes down. In figure 7, the whole overlap area is shown in both photos, and the "energy image" representing the photometric energy terms is displayed in between. The seam line that was found by the algorithm is overlaid. Figure 8 shows one northern, critical part of the image in detail.



Figure 7: Orthophotos 7472 (left), 7468 (right) and the image representing the photometric energy terms in the overlap area (middle): the darker the image, the higher the energy. The displayed line is the seam line as proposed by the algorithm.

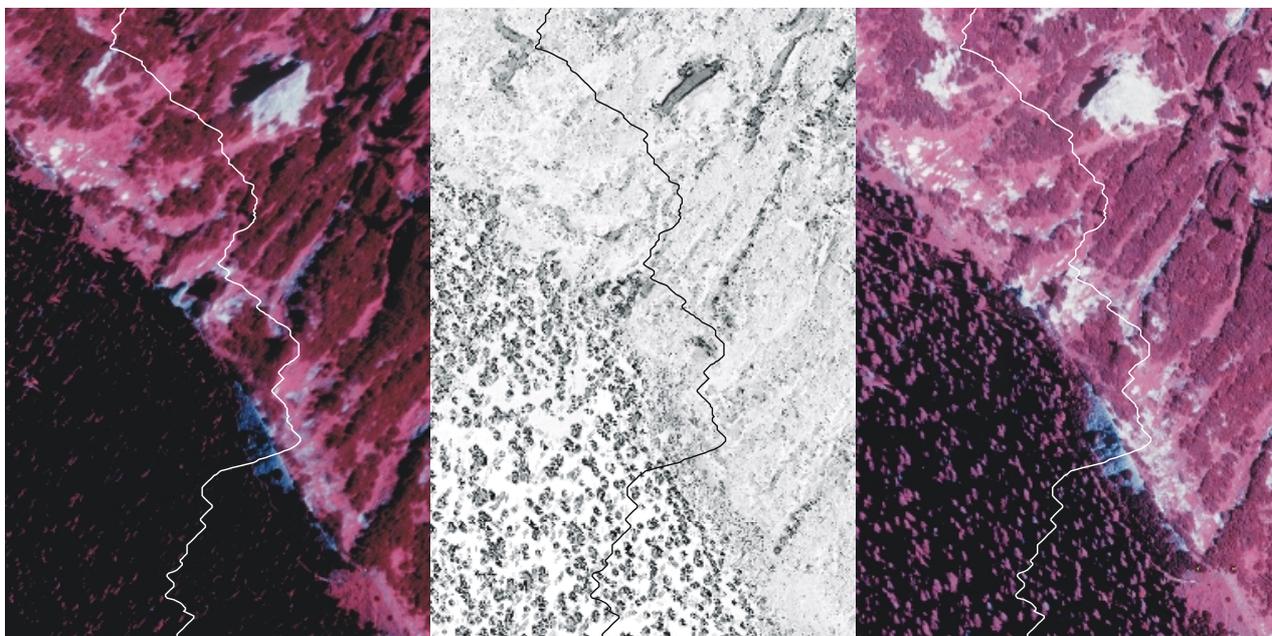


Figure 8: A difficult area for delineating the seam line in 7472 (left), 7468 (right) and in the "energy image" (middle).

5 CONCLUSIONS

The presented approach is an appropriate method for solving one problem in the production of orthophoto mosaics: detecting seam lines in a sophisticated manner. Snakes are well suited for detecting the line. The success depends on an energy function, in which all criteria for an optimal line are correctly formulated. The presented algorithm is designed for color orthophotos. For gray-level photos, a similar strategy can be used. The method works without any user interaction and can be applied in automatic processing chains.

ACKNOWLEDGMENTS

The research was funded by the Austrian Science Fund (FWF) in the project P13725-MAT.

REFERENCES

- Baltsavias E., Käser Ch., 1998. DTM and orthoimage generation – a thorough analysis and comparison of four digital photogrammetric systems, IAPRS, Vol. 32, Part 4, Stuttgart, FRG, pp. 42-51.
- Haralick R.M., Shapiro L.G., 1992. Computer and robot vision, Addison-Wesley, Reading, MA, USA.
- Horn B.K.P., 1977. Understanding image intensities, *Artificial Intelligence* 8, pp. 201-231.
- Kass M., Witkin A., Terzopoulos D., 1988. Snakes: active contour models, *International Journal of Computer Vision*, 1(4), pp. 321-331
- Kerschner M., 1998. Homologous twin snakes integrated in a bundle block adjustment, IAPRS, Vol. 32, Part 3/1, Columbus, OH, USA, pp. 244-249
- Kraus K., Schneider W., 1988. Fernerkundung, *Physikalische Grundlagen und Aufnahmetechniken*, Band 1, Dümmler, Bonn, pp. 18-30, 36-51, 86-93.
- Kraus K., 1993. Photogrammetry, fundamentals and standard processes, Vol. 1, Dümmler, Bonn, chapter 6.
- Szeliski R., 1996. Video mosaics for virtual environments. *IEEE Computer Graphics and Applications* 16(2), pp. 22-30.