3D Visualization and Evaluation of Remote Sensing Data

Stefan Growe   Peter Schulze   Ralf Tönjes

Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung
University of Hannover, Appelstrasse 9A, 30167 Hannover, Germany
E-mail: growe|toenjes@tnt.uni-hannover.de, peter@yin.atlas.de

Abstract

The parallel evaluation of remote sensing data from multiple sensors is complicated by different image scales and orientations. Thus a simultaneous visualization of the multi-sensor data is envisaged. For a spatial impression of the observed area the images are co-registered, projected onto a 3D terrain model and displayed stereoscopically. The presented 3D evaluation system allows the real-time navigation in the virtual landscape. The user is able to manipulate the 3D model by blending and mosaicking the images from different sensors. An online access to a geoinformation system (GIS) offers the possibility to display GIS data.

1. Introduction

The analysis of remotely sensed images for environmental and agricultural monitoring and map update represents a major topic of remote sensing. In addition to the conventional visual or infrared cameras the synthetic aperture radar (SAR) sensor becomes more and more popular because it works at night and penetrates clouds. These days digital images of the same observation area are evaluated in parallel by displaying them on different monitors together with a map. Due to different image scales and orientations it is difficult for the human interpreter to explain the image content or to recognize corresponding structures.

To improve the evaluation facilities the combination of the multi-sensor data on a single screen is envisaged. The simultaneous visualization eases the explanation of unknown structures and the detection of temporal changes because the image data is compared directly with other images and the map. To give a spatial impression of the observed terrain a three-dimensional real-time visualization should be provided. This enables the human interpreter to explore and evaluate the multi-sensor image data in a virtual landscape. To ease the image interpretation vector data of a Geographic Information System (GIS) is made accessible. The presented work exploits the German digital landscape model DLM 25 of ATKIS (Authoritative Topographic and Cartographic Information System) which mirrors the content of the 1:25000 map.

Various 3D visualization systems for geographical data have been developed in the past. They differ in several aspects like their application, the used datasets, visualization techniques etc.. The system ViRXIS [3], for example, serves as base for a VR based Tourist Information System (ViRTIS) or a VR based Geographic Information System (VirGIS) respectively. It uses a Digital Elevation Model (DEM) and raster data like satellite imagery or scanned maps. To guarantee a high visualization performance the elevation model is represented in different levels of detail (LOD).

Nüesch et al. [2] developed a system for the real-time landscape visualization based on Landsat TM images and a DEM. Additionally vector data like roads and buildings are inserted into the landscape model by modifying the triangulated irregular network (TIN) representing the surface. Beside the LOD management the use of geotypical texture taken from a database instead of the photo texture from the image data increases the display performance. A similar approach is described by McKeown et al. [1]. In addition to satellite imagery they use aerial images. The given vector data set containing information about buildings and roads is extended by automatically extracted features which are integrated in the TIN. Visualization is performed with artificial textures.

In contrast to the pure visualization of a landscape model the Virtual GIS developed by Koller et al. [4] provides an interactive component. The navigation in the virtual world is enhanced by an overview map inset or an overlaid coordinate grid. Additionally the system offers a popup menu interface to the GIS database to access geographical information about selected objects.

Our goal is the simultaneous 3D visualization of remote sensing images from multiple sensors. To optimize the interactive evaluation facilities the system should provide features to manipulate, blend and mosaic the images and an

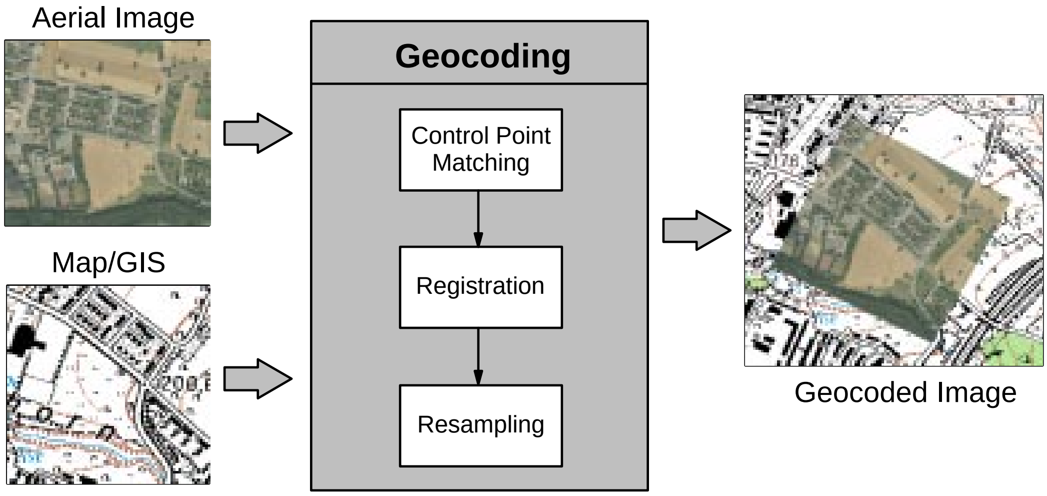

Figure 1. Geocoding of remote sensing data

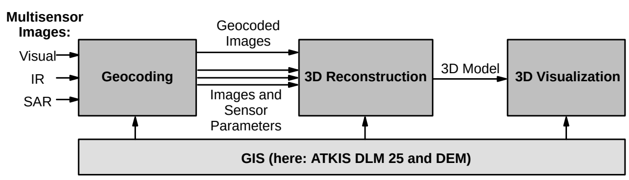

Figure 2. Processing pipeline for the 3D evaluation of remote sensing data from multiple sensors

online access to the GIS database. Therefore all images are registered in a common geocoordinate system (*geocoding*) (Fig. 1). In a second step a 3D terrain model which reconstructs the height relief of the observed area is generated out of the image data and the DEM. Finally the 3D landscape model is visualized by a stereoscopical projection system based on a *SGI Onyx Infinity Reality* platform. A special user interface allows the operator to evaluate the data interactively. In all processing steps the GIS data is used to improve the results or to support the user. The complete pipeline shown in Figure 2 is described in the following chapters.

2. Geocoding

In general the multi-sensor image data to be evaluated are taken at different moments from different positions and heights. SAR images are processed from complex raw data recorded during the flight. Thus the remote sensing images differ in scale and orientation. For the simultaneous visualization and evaluation the images have to be transformed from their image coordinates into a common coordinate space. Because information from a map and a GIS is supposed to be included, a geographical coordinate system like UTM or the German Gauss-Krüger is used. The process to generate a georeferenced image is called geocoding.

Prerequisite for the image-to-map registration is the detection of corresponding points in the image and the map. These control points must be distributed equally and defined accurately. Nowadays this is mostly done manually. But the quantity of data and the short update periods ask for methods that automate the control point matching. In the literature various approaches for image registration have been suggested [5] [6]. We developed a knowledge based approach to search control points for geocoding [7]. The image interpretation system AIDA [8] is used to select crossroads from the GIS and to find the corresponding structures in the image to be registered.

AIDA provides methods for the explicit representation of knowledge about the objects expected in the image. This prior knowledge is defined by so called semantic nets which consist of nodes representing the scene objects and links inbetween that describe their decomposition in parts or specialized objects. A semantic net for the detection of control points (here: crossroads) was implemented that defines crossroads as the junction of three or more roads. Each road is represented in the GIS database with its 3D geometry on the one hand and by a linear feature in the remote sensing image on the other hand. Problem independent rules exploit this knowledge to generate hypotheses in a model driven manner that are consecutively verified or falsified in the image data. In this application roads forming a crossroads are extracted from the GIS and projected into the image using

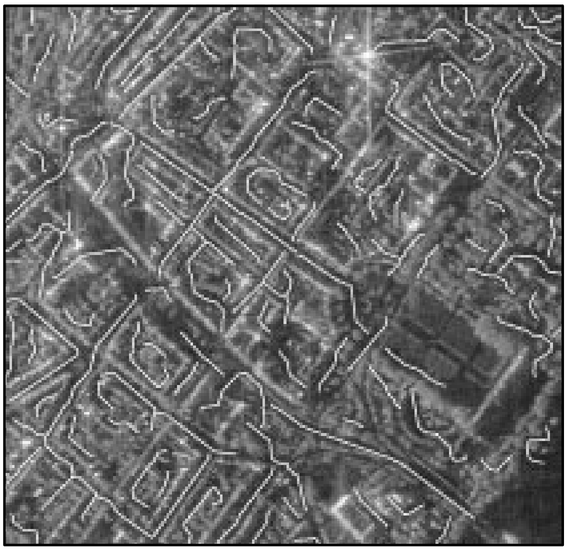

Figure 3. Road candidates segmented in a SAR image

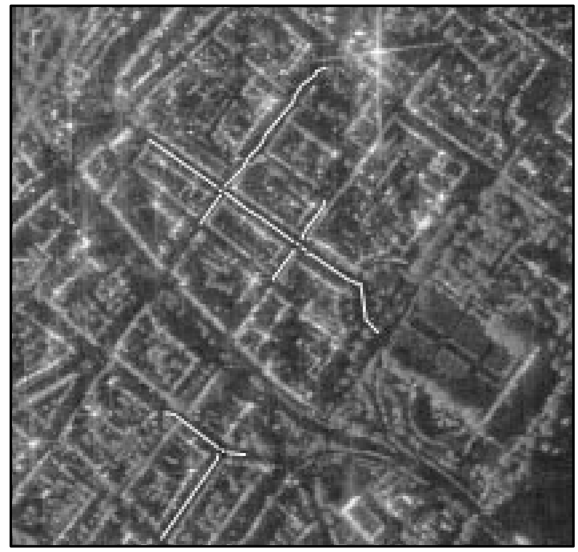

Figure 4. Selected crossroads used as control points

a coarse registration derived from general flight parameters like course, attitude and image resolution. These expectations are used to select suitable road candidates segmented in the image (Fig. 3). Competing alternatives are judged and an A*-algorithm selects the most promising interpretation for further investigation. Finally the most certain crossroads are used as control points where the 3D coordinate is taken from the GIS data and the 2D coordinate is given by the estimated intersection of the roads in the image. Results of the control point detection are shown in Fig. 4 for a SAR image.

Based on the detected control points the parameters of a sensor specific mapping model are estimated. For an aerial image, for example, the parameters of the central projection like position, orientation, and focal length of the camera are calculated. For SAR imagery we use a polynomial approximation of the Range-Doppler equations [9] where the orbit of the sensor platform during data acquisition and the Doppler polynomial coefficients are estimated. For a set of model parameters the control points are projected into the image and the distances to the corresponding image points are measured. These residuals are minimized by varying the model parameters and using least square optimization techniques. The optimal set of model parameters is taken to transform each geocoordinate of the region covered by the future geocoded image into the original image space. Using bilinear interpolation the pixel value of the geocoded image is calculated from the original image. Assuming a precise geocoding different images of the same area can be combined to 2D image mosaics or projected onto a map (Fig. 1). For a three-dimensional visualization however a realistic digital terrain model is necessary which approximates the height relief of the landscape accurately.

3. 3D Reconstruction

Digital terrain models are available from the German ordnance survey offices with a raster of 50 m. But these models contain only the height at ground level and disregard elevations caused by vegetation and buildings. They are insufficient for a realistic 3D visualization in the close range because of the missing detail. Hence the images to be visualized are used itself to improve the 3D reconstruction of the observed scene.

The common approach to recover height information from aerial images is to use stereo triangulation techniques for overlapping images [10] [11]. The focal point of the left and right camera and the observed point form a plane that intersects the left and right image in the epipolar line. Hence the search for corresponding (homologous) points is reduced to the search along the epipolar line. To ease the search for homologous points the two overlapping aerial images are rectified in a way that the epipolar line coincides with the image scanline. The required camera orientation is known from the registration process.

The correspondence analysis determines the height dependent parallax of homologous points. Due to the limited height difference of the observed terrain the search space is reduced by a parallax limit. Using normalized cross correlation as cost function for matching homologous points the optimum for each scanline is found by dynamic programming yielding a parallax map for the image [12]. Finally the parallax map is transformed to a depth map using binocular camera geometry. The depth τ_l describes the distance between the spatial point $\boldsymbol{P}$ and the focal point $\boldsymbol{C}_l$ of the left camera along the line of sight $\boldsymbol{S}_l$ (Eq. (1)). The line of sight is defined by the normal vector pointing from the focal point of the camera to the projection of $\boldsymbol{P}$ in the image plane. Inserting τ_l in equation (Eq. (2)) returns the geocoordinate.

$$\tau_l = \frac{(\mathbf{C}_l - \mathbf{C}_r)^T \cdot (s_{lr} \cdot \mathbf{S}_r - \mathbf{S}_l)}{1 - s_{lr}^2} \quad \text{with: } s_{lr} = \boldsymbol{S}_l^T \cdot \boldsymbol{S}_r \tag{1}$$

$$\boldsymbol{P} = \boldsymbol{C}_l + \tau_l \cdot \boldsymbol{S}_l \tag{2}$$

To obtain 3D models for efficient visualization the height map is subsampled and approximated by a triangle mesh in space. The detailed photometric properties are modelled by projecting the aerial images onto the triangle surfaces.

However the height map is erroneous and incomplete. Furthermore reconstruction suffers from inaccuracy of the image data and from occlusions of 3D objects in their 2D projections. Thus the 3D reconstruction employs surface models to correct erroneous measurements and to complete the sparse height map. Assuming piecewise smooth surfaces regions of continuous height are interpolated [13]. This requires the detection of discontinuities. The presented systems uses prior knowledge about the objects in the scene to constrain the geometry. For example, the knowledge that forest edges coincide with a height step constitutes a strong constraint.

Furthermore the mesh approximation has to consider the discontinuities. Otherwise the models do not meet the expectations of the human observer who knows, for example, that roads run continuously and that houses have a particular shape. Faulty breaklines are caused if the mesh approximation does not correspond with the object boundaries.

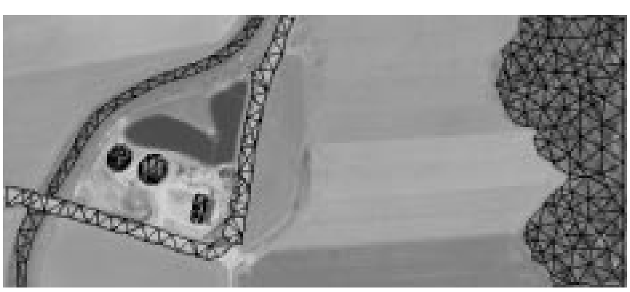

Figure 5. Separate triangulation of roads and forests

The presented approach employs again the interpretation system AIDA [8] to assign a semantic to image regions. The system uses GIS data as expectations and assigns semantic meanings to the image primitives derived from image processing algorithms. Interpretation yields the segmentation of aerial images into various regions, such as forests, grassland, roads and buildings. The location of these regions is stored in image masks. Scene reconstruction uses these image masks to apply object specific constraints to the height map obtained by stereoscopic correspondence analysis. The prior knowledge forces a height step between forests and grassland or roads. Furthermore the object semantic controls mesh generation. Roads and rivers are approximated by a separate mesh to ensure a continuous course. At the edges of forests a vertical mesh for the height step is inserted. Figure 5 shows the net for selected objects.

The semantics attached to the model parts allow an object specific post processing: Objects can be artificially refined by adding details which are invisible to the sensor from computer graphic libraries. For close-up views of forest edges, for example, synthetic trees with fine transparent leaf structure are placed in front of the edges (Fig. 6).

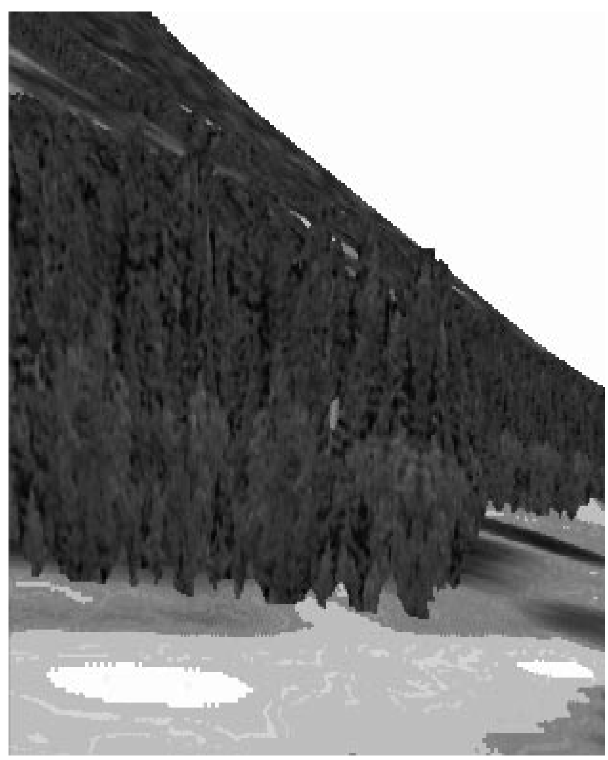

Figure 6. Refinement of forest edges by synthetic trees

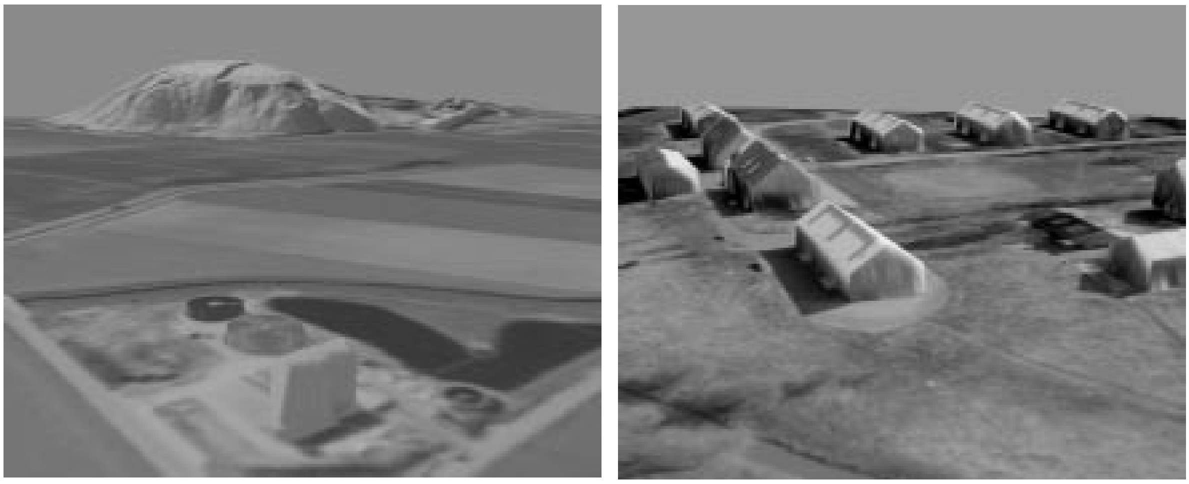

Figure 7. Synthesized view of 3D terrain models with adapted house models and continuous roads

Buildings can have quite complex shapes. Their reconstruction employs a generic building model composed of polyhedrons. The parameters are adapted to the image data combining model and data driven strategies (Fig. 7 and Fig. 8). The object specific 3D reconstruction is described in detail in [14] and [15].

Figure 8. Adapted parametric house models

4. 3D Visualization

The 3D landscape model existing of the geocoded multisensor images and the digital terrain model is visualized stereoscopically and in real-time to the human interpreter. The 3D evaluation system should provide the following functionalities:

- navigation in the virtual landscape
- separate activation and deactivation of the sensor data
- blending of the different images
- interactive creation of image mosaics
- realization of simple measurement tasks, e.g. distances
- online access to GIS data

The features mentioned above demand for a particular user interface on the one hand and a special model representation form on the other hand. Furthermore the system has to provide an interface to the GIS database.

Visualization is performed on a SGI graphic engine with a stereoscopic projection device. Two images, for the left and right eye, are rendered separately, polarized perpendicularly and projected onto a screen. The human observer wears polarization glasses that separates the two images. Thus he has a spatial impression of the visualized scene.

4.1. Model Representation

To achieve real-time performance the visualization toolkit *IRIS Performer* [16] is used which is optimized

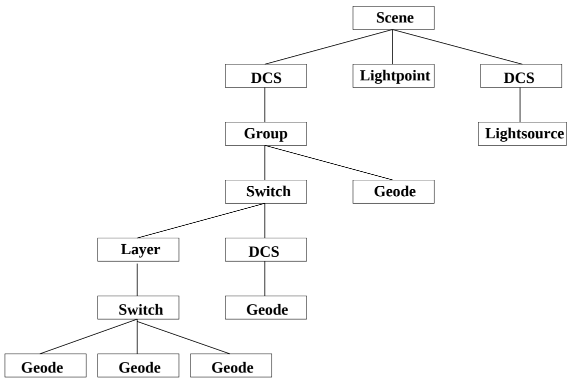

Figure 9. Scene Graph in *IRIS Performer*

for SGI multiprocessor platforms. In *Performer* the virtual world to be visualized is represented by the so called *scene graph*. This graph describes the hierarchy of the scene elements as a tree (Fig. 9). The geometry and the texture of a 3D model is represented by a geometry node (*Geode*). Other node types contain the information about light sources (*Lightsource*) or coordinate transformations (*DCS*). The scene graph is processed from top to bottom and from left to right. Hence the arrangement of the nodes in the scene graph determines the order of visualization and thus the visibility of the scene elements.

Because Performer uses the graphic libraries *IRISGL* or *OpenGL* the 3D geometry is described by triangle meshes consisting of vertices. The triangles possess attributes like colour, texture, transparency etc.. The real-time simulation is performed by the repeated visualization of the scene graph with a frame update rate as constant as possible. The supported motion models allow to fly and drive through the virtual landscape and to rotate and zoom the model.

The human interpreter should have the possibility to activate and deactivate the remote sensing images separately to improve the comparison between different sensors. Because only one texture can be defined for a 3D model, the scene graph contains separate geometry nodes for each image to be visualized. The geometry nodes reference the same triangle mesh but different textures resulting in overlapping multi-sensor layers.

For the representation of the 3D landscape model the node types *Layer* and *Switch* are used in the scene graph (Fig. 10). For each texture (visual, SAR, IR, or map) one switch exists which is again connected with the different geometry nodes. Hence the models can be activated or deactivated separately by configuring the switches. Because of the processing order of the scene graph the models are visualized from left to right. Consequently an active node occludes all nodes positioned left of it in the scene graph. In the given example model 2 appears above model 0 and model 1.

The separate geometry nodes with their triangle meshes and textures allocate a large amount of memory. The number of geometry nodes raises quadratically with the number of textures to be visualized. Hence the nodes are not duplicated, but they are merely referenced by the switch nodes (Fig. 11). This improves the performance during simulation and reduces the file size by 70% for a model consisting of three layers.

By now the human interpreter is able to activate and deactivate the textures. For a simultaneous visualization of the multi-sensor data the textures have to be blended. Therefore a transparency value is defined for each texture which can be modified continuously by a slider in the graphical user interface. By reducing the transparency value of a texture it is mixed with the textures beneath.

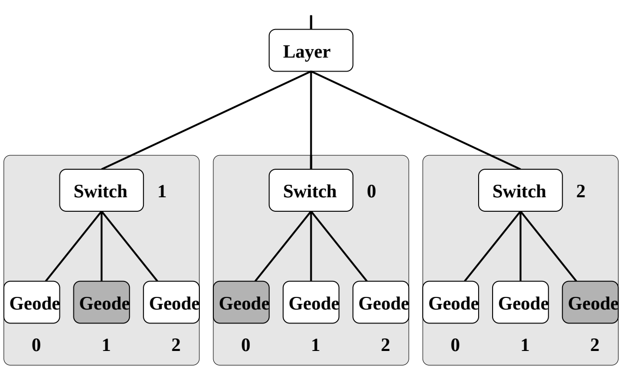

Figure 10. Scene Graph of the layered model with separate geometry nodes

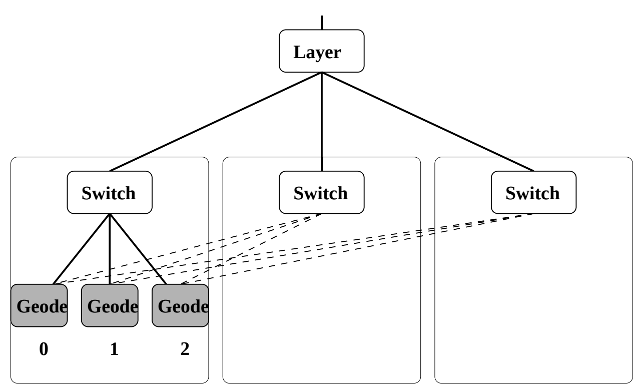

Figure 11. Scene Graph with referenced geometry nodes

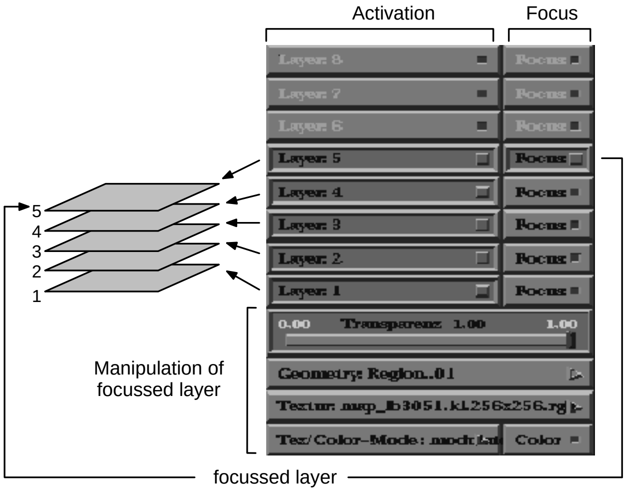

Figure 12. Activation of model layers for the manipulation of texture, transparency, etc.

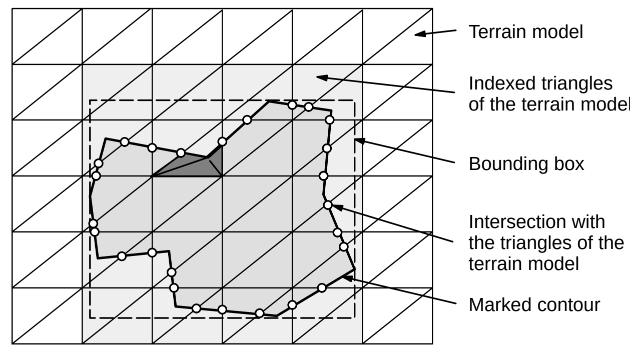

Figure 14. Creation of a new contour by insertion of the intersections between marked contour and the triangle mesh of the terrain model

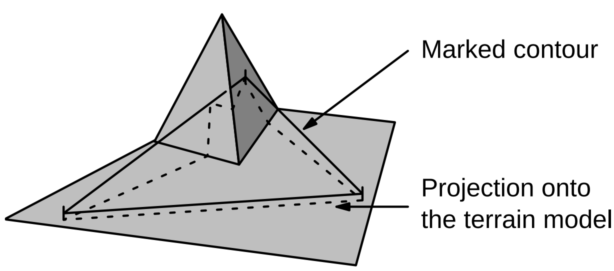

Figure 13. Projection of marked contour onto the terrain model

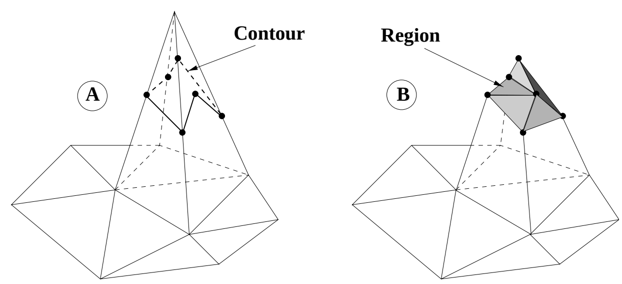

Figure 15. Erroneous triangulation of a region out of a given contour

Figure 12 shows a part of the graphical user interface. Each texture is represented by a layer which can be turned on and off by toggling the corresponding button. The order of the layers represent the sequence of visualization. By focussing a layer the widgets to define the geometry, texture, colour, and transparency get valid. Thus the human operator can explore the 3D model interactively by navigating through the virtual landscape and manipulating the textures.

4.2. Interactive Creation of Regions

To create arbitrary image mosaics or to request informations from the GIS database the user has to mark a region within the virtual landscape. Similar to 2D image manipulation programs a region is defined by its polygonal contour. In the three-dimensional case the marked region is not planar but describes a new 3D model to be created which inherits the height information from the enclosed landscape model.

The corners of the contour marked interactively are visualized by adding 3D models of rods to the scene. In a first step the contour marked by the user is projected onto the surface of the landscape model. As depicted in Figure 13 a new polygon is built (dotted line).

To adapt it to the height relief the points of intersection between the original contour and the triangles of the terrain model are inserted as new vertices (Fig. 14). For an increased performance only those triangles of the terrain model are searched for intersections which are covered by the bounding box of the marked contour (light grey in Fig. 14).

After the adaptation of the contour the defined region has to be triangulated to generate a 3D model. Because of the enclosed height relief it is not sufficient to connect all vertices of the contour with each other. This would yield erroneous results (Fig. 15). Thus all triangles of the enclosed part of the terrain model have to be used for the triangulation. If a triangle is contained entirely in the region it can be utilized without changes. If the contour intersects a triangle of the terrain model the generated polygon is triangulated using the vertices of the contour, of the terrain model, and their points of intersection. This is illustrated for one triangle in Fig. 14 (dark grey region).

The 3D model of the marked region is combined with the currently visible texture and added to the scene graph with a further switch node. Thus each interactively created region represents a new layer of the 3D landscape model to which an arbitrary texture, colour, and transparency can be assigned.

The basic facility to define points and regions interactively makes it possible to realize various features. As the model is registered in geocoordinates the user can measure the distance between two points or the area of a region in world coordinate units like meters. He is able to request GIS data for the selected area. Finally it is possible to create 3D image mosaics interactively. For example the human interpreter is able to superpose the aerial image on shadow regions of the corresponding SAR image.

4.3. Interface to Geoinformation System

For the evaluation of the remote sensing data it is useful to integrate data from a geoinformation system (GIS). For example the user might visualize the road net of a region of interest to explain the linear structures in a SAR image. The 3D visualization system provides an online interface to the used GIS *SICAD/open* giving access to the GIS data during evaluation.

To formulate a GIS request the human interpreter marks a region as described in chapter 4.2. After selecting an object class to be searched for like roads, forests, bridges etc. the request is started. The arbitrarily formed region is transformed in a rectangle defined by the bounding box. The given object class is translated in numerical codes of the German ATKIS. The data returned from the GIS database is gathered and the geometrical information is extracted. The visualization system can only process regions approximated by a triangle mesh, but streets for example are modeled in the GIS by their central axes. Thus all linear objects are transformed in stripes using the information about their width. The GIS data is transferred to the visualization system which triangulates each region and includes it in the scene graph. The GIS information represents a separate layer of the model which can be manipulated and deactivated in the same way as all other layers.

5. Results

The described system for the 3D evaluation of remote sensing data was realized on a *SGI Onyx Infinite Reality* multiprocessor platform using *IRIS Performer* and *OpenGL* [17]. To guarantee a constant frame update rate the rendering pipeline is processed on a separate CPU. The calculation of intersections and the handling of GIS requests are distributed on a second and third processor respectively.

The features of the visualization system are demonstrated in the following snapshots of the test-site *Koblenz*. Figure 16 shows a synthesized view of the 3D landscape model with mixed SAR and map texture. The triangulation result of an interactively marked region is illustrated for the Rhine river in Fig. 18. The Mosel river is displayed with map texture. Another example for a 3D image mosaic is shown in Fig. 17 where a part of the SAR image is superposed on the aerial image. The current geocoordinate of the mouse is visualized in a head-up-display (Fig. 17). Thus the interpreter is able to measure distances in the virtual landscape. Finally the result of an online GIS request is illustrated in Fig. 19. For a region of interest the roads are displayed in the 3D landscape model textured with the SAR image.

6. Conclusion

A system for the interactive 3D evaluation of remote sensing data from multiple sensors was presented. The images to be visualized are geocoded and projected onto a 3D terrain model. For the automatic search of control points and the object specific 3D modelling of the scene a knowledge based image interpretation system called AIDA is used.

The textured 3D landscape model is visualized stereoscopically and in real-time to the human interpreter. An optimized representation form of the layered 3D model allows the manipulation of the multiple textures. Via the graphical user interface the textures can be exchanged and blended. The human interpreter has the possibility to mark arbitrary regions interactively. Hence he can create image mosaics to combine the multi-sensor data. An online interface to a geoinformation system allows the visualization of GIS data for a region of interest.

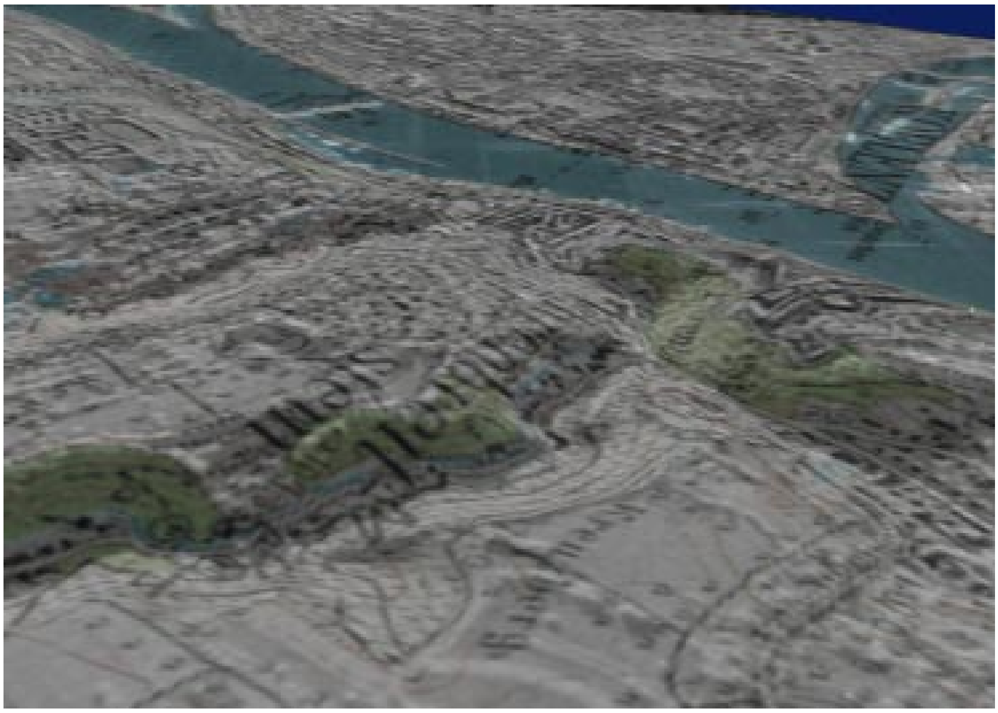

Figure 16. Interactive blending of a SAR image with a map

Figure 17. Interactively created 3D image mosaic of an aerial and SAR image

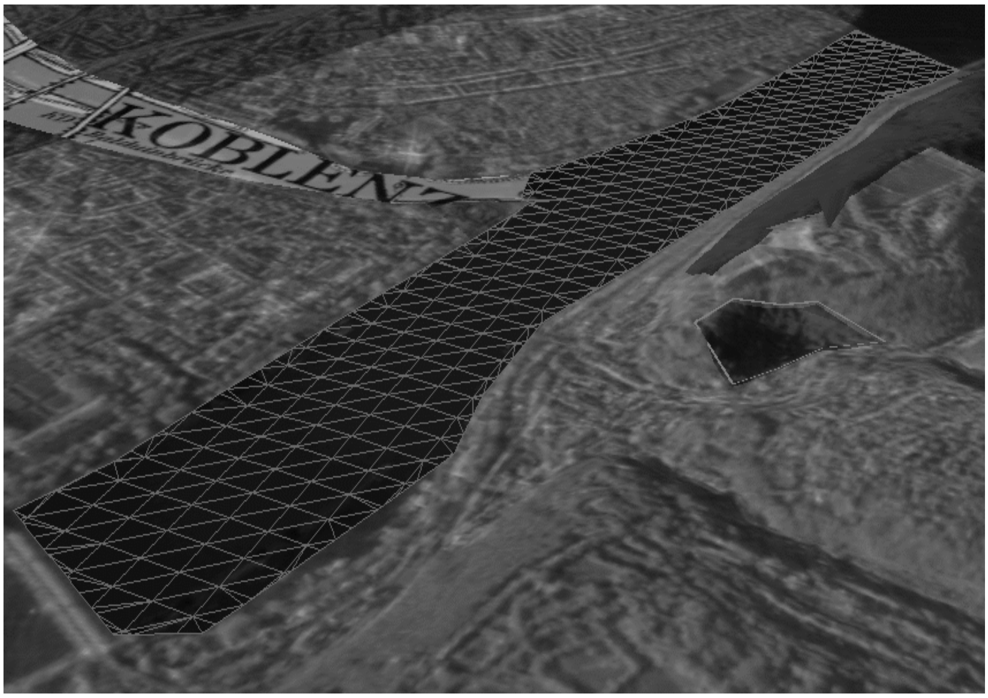

Figure 18. Triangulation of an interactively marked region and map overlay

Figure 19. online visualization of GIS data (here road net) for a selected region of interest

References

[1] Polis, M., Gifford, S., McKeown, D., "Automating the Construction of Large Scale Virtual Worlds", *Proc. of ARPA Image Understanding Workshop 94*, Monterey, CA, pp. 931-946, 1994

[2] Suter, M., Nüesch, D., "Automated Generation of Visual Simulation Databases Using Remote Sensing and GIS", *Proc. of IEEE Visualization '95*, Atlanta, GA, pp. 86-93, 1995

[3] Szabó, K., Stucki, P., Aschwanden, P., Ohler, T., Pajarola, R., Widmayer, P., "A Virtual Reality based System Environment for Intuitive Walk-Throughs and Exploration of Large-Scale Tourist Information", *Proc. of the Int. Conference on Information and Communication Technologies in Tourism (Enter'95)*, Innsbruck, Austria, January 1995

[4] Koller, D., Lindstrom, P., Ribarsky, W., Hodges, L., Faust, N., Turner, G., "Virtual GIS: A Real-Time 3D Geographic Information System", *Proc. of ARPA Image Understanding Workshop 94*, Monterey, CA, pp. 94-100, 1994

[5] Li, S.Z., Kittler, J., Petrou, M., "Automatic Registration of aerial photographs and digitized maps", *Optical Engineering*, Vol. 32, No. 6, pp. 1213-1221, June 1993

[6] Perlant, F., McKeown, D., "Scene Registration in Aerial Image Analysis", *Photogrammetric Engineering and Remote Sensing*, Vol. 56, No. 4, pp. 481-493, April 1990

[7] Growe, S., Tönjes, R., "A Knowledge Based Approach to Automatic Image Registration", *Proc. of International Conference on Image Processing (ICIP'97)*, Vol. III, October 1997, Santa Barbara, USA

[8] Liedtke, C.-E., Bückner, J., Grau O., Growe S., Tönjes, R., "AIDA: A System for the Knowledge Based Interpretation of Remote Sensing Data", *3rd Int. Airborne Remote Sensing Conference & Exhibition*, Copenhagen, Denmark, Vol II, pp. 313-320, July 1997

[9] Raggam, J., Strobl, D., Hummelbrunner, W., "Product Quality Enhancement and Quality Evaluation", in *SAR Geocoding: Data and Systems*, Schreier, G. (Eds), pp. 187-207, Wichmann, 1993

[10] Ackermann, F., Krzystek, P. , "MATCH-T: Automatic Mensuration of Digital Elevation Models", *Proceedings of Technical Seminar of the Sociedad Espanola de Cartografia Fotogrammetria y Teledeteccion*, pp.67-73, Barcelona, April 1991

[11] Lemmens, M., "A survey on Stereo Matching Techniques", *International Archives of Photogrammetry and Remote Sensing*, Vol. XXVII, Part 5, pp. 11-23, 1988

[12] Falkenhagen, L., "3D Object-Based Depth Estimation from Stereoscopic Image Sequences", *International Workshop on stereoscopic and three-dimensional imaging*, Santorini, Greece, 1995

[13] Terzopoulos, D., "The Computation of Visible-Surface Representations", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, Vol. 10 No.4, pp. 417-438, 1988

[14] Tönjes, R., "3D Reconstruction of Objects from Aerial Images using a GIS", *ISPRS Joint Workshop on "Theoretical and Practical Aspects of Surface Reconstruction and 3-D Object Extraction"*, Haifa, Israel, September 9-11, 1997

[15] Tönjes, R., "Knowledge Based Modelling of Landscapes", *International Archives of Photogrammetry and Remote Sensing*, Vol. XXXI, Part B3, pp. 868-873, Vienna, 1996

[16] Silicon Graphics Inc., *IRIS Performer TM Programmer's Guide*, 1995

[17] Schulze, P., "Entwicklung eines Systems zur interaktiven Auswertung von multisensoriellen Fernerkundungsdaten mittels 3D-Visualisierung", Master Thesis, University of Hannover, Institut für Theoretische Nachrichtentechnik und Informationsverarbeitung, 1997