Interventional Video Tomography

Michael Truppe, Ferenc Pongracz, Artma Medizintechnik GmbH, Vienna, Austria
Oliver Ploder, Arne Wagner, Rolf Ewers, University of Vienna, Dept. Maxillofacial Surgery

ABSTRACT

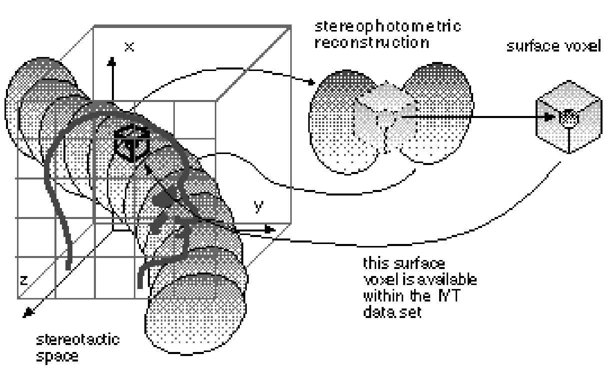
Interventional Video Tomography (IVT) is a new imaging modality for IMAGE DIRECTED SURGERY to visualize in real-time intraoperatively the spatial position of surgical instruments relative to the patient's anatomy.

The video imaging detector is based on a special camera equipped with an optical viewing and lighting system and electronic 3D sensors. When combined with an endoscope it is used for examining the inside of cavities or hollow organs of the body from many different angles. The surface topography of objects is reconstructed from a sequence of monocular video or endoscopic images. To increase accuracy and speed of the reconstruction the relative movement between objects and endoscope is continuously tracked by electronic sensors. The IVT image sequence represents a four-dimensional data set in stereotactic space and contains image, surface topography and motion data.

In ENT surgery an IVT image sequence of the planned and so far accessible surgical path is acquired prior to surgery. To simulate the surgical procedure the cross sectional imaging data is superimposed with the digitally stored IVT image sequence.

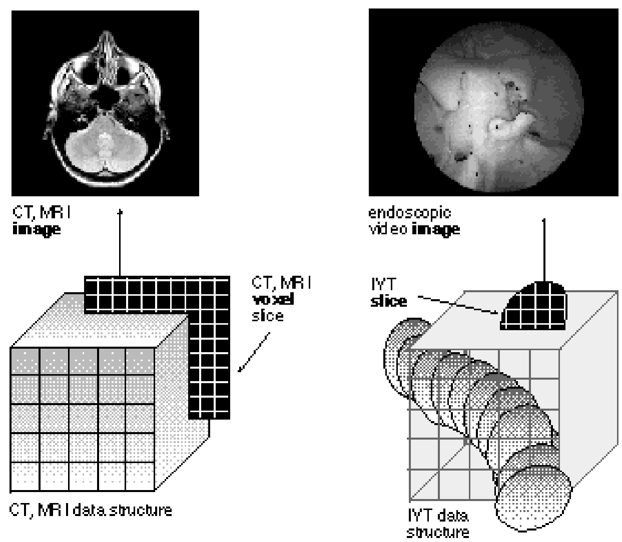
During surgery the video sequence component of the IVT simulation is substituted by the live video source. The IVT technology makes obsolete the use of 3D digitizing probes for the patient image coordinate transformation.

The image fusion of medical imaging data with live video sources is the first practical use of augmented reality in medicine. During surgery a head-up display is used to overlay real-time reformatted cross sectional imaging data with the live video image.



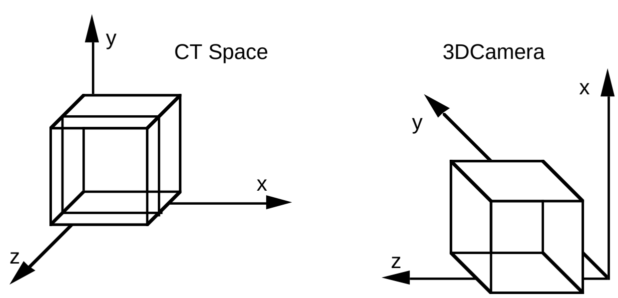
IVT TECHNOLOGY

The native IVT sequence contains all information for off-line reconstruction, imaging and electronic sensor data. Any two corresponding images are correlated and a three dimensional stereophotometric reconstruction of the surface topography is completed.



COORDINATE TRANSFORMATION

The calibration in the Artma Virtual Patient™ system is defined as the computation of the transformation parameters between the cross sectional imaging data space (CT space, considered as "world space") and the



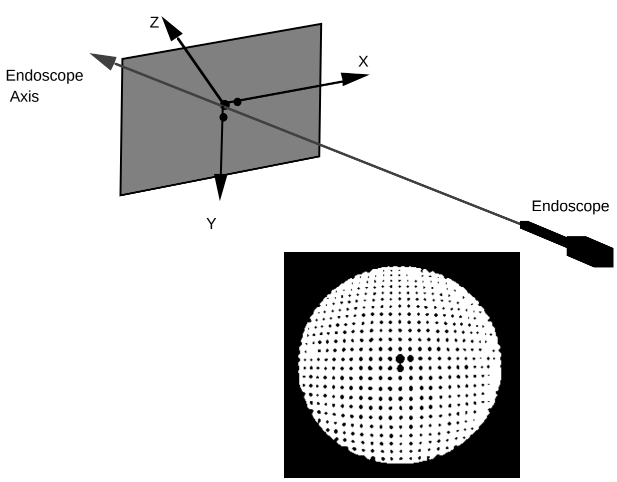
camera space. Both spaces are referred to as 3D spaces despite the fact that the camera space can be manipulated by the user in only 2 view coordinates.

The "connector points" are the minimum number of points, represented both in the CT space and camera space, necessary to calculate the transformation parameters with a certain accuracy. The user can associate the object position in the CT space with its position in the camera space once the transformation parameters are determined. Our method allows registering a CT space position by dragging a point in a single camera view, which appears a little bit different from the conventional stereo photogrammetric methods. The CT space and the camera space have always a common reference point to tracking a cursor position back projected from the camera view into the CT space. The solution for the transformation matrix is based on the method of Adbel-Aziz and Karara (1971). The reliability of the Z coordinate in the CT space is determined by the thickness of CT slices which can be in the mm range. A simple redundancy rule was included to enhance the accuracy of the Z coordinates.

Endoscope Calibration

3 planes were distinguished during the calculation of distortion parameters (Bowman and Forrest, 1987):

A. Detected image plane
B. Grid image plane
C. Distorted image plane

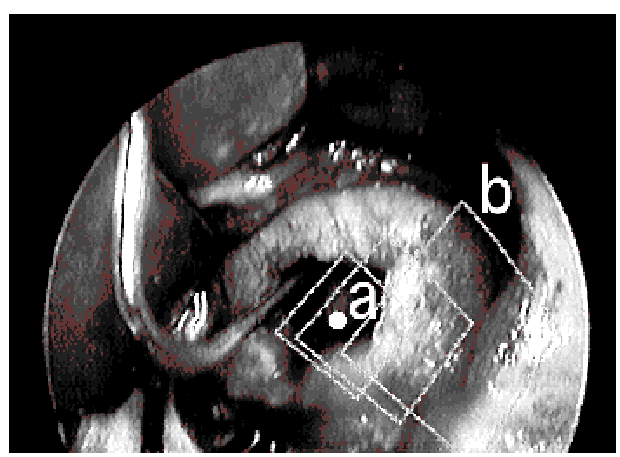


The detected image plane represents the observed image plane actually visible through lenses with radial distortion. This detected plane is identified by computer vision algorithms (first the center dot and the directions of the X-Y axes are determined, then by a sorting algorithm each calibration dot is addressed by an index). The grid plane is defined by the original flat reference grid by setting up the right orientation with a (linear) matrix calibration (to find the right orientation parameters we include a small, undistorted central part of the picture into the analysis; the center of the grid should be matched by the optical center of the endoscope). The distorted image plane is an abstraction representing an optimal approximation of the detected image plane. The goal of distortion analysis is to find a transformation procedure to relate a point in the grid plane to the distorted plane. The “input plane” is the grid plane, the “output plane” is the distorted plane and the “ideal plane” is the detected image plane.

To get a statistically dependable solution, large number of grid points (200-600) is used in our calibration routine. The algorithm solves the inverse problem too: the point defined in the detected image plane can be transformed (in a least-square sense) to the appropriate location in an undistorted image plane which is an optimal approximation of the grid plane. This step makes the quantitative analysis possible, that is, the real length of lines and the real values of an area can be determined from drawing in the distorted endoscopic image. In this case the roles of planes compared to the first case are exchanged: the “input plane” is the detected plane, the “output plane” is the undistorted plane and the “ideal plane” is the grid plane oriented and scaled by linear matrix calibration.

Case Study

The patient had a dislocated radix in the right paranasal sinus. We decided to localize the structure with stereotactic navigation of the endoscope. In this case the radix *(a)* was identified in two planar radiographs. To reconstruct the three dimensional coordinates of the location (equivalent to the CT space) a stereophotometric reconstruction was performed based on the method of Longuet-Higgins (1989). The planned



trajectory path *(b)* of the endoscope was simulated in the computer system. The position of endoscope and surgical instruments relative to the anatomical structures were tracked by 3D sensors in real-time. Because 3D sensors were also attached to the patient the movement of the head was not constrained during surgery.

The live video image fusion concept eliminates the need for additional computer monitors for stereotactic navigation information. The integration of the Head-Up display was demonstrated as work in progress.

References

Adbel-Aziz,Y.I. and Karara,H.M. Direct linear transformation into object space coordinates in close-range photogrammetry. In Proceedings of the Symposium on Close-Range Photogrammetry, University of Illinois, pages 1-18, January 1971.

Bowman, M.E. and A.K. Forrest: Transformation calibration of a camera mounted on a robot. (1987) Image and Vision Computing 5(4): 261-266.

H. C. Longuet-Higgins: A computer algorithm for reconstructing a scene from two projections. Nature. 1981, 293: 133-135.