

Advanced Virtual Endoscopic Pituitary Surgery

André Neubauer, *Student Member, IEEE*, Stefan Wolfsberger,
Marie-Thérèse Forster, Lukas Mroz, Rainer Wegenkittl, and Katja Bühler

Abstract—Endoscopy has recently been introduced to endonasal transsphenoidal pituitary surgery as a minimally invasive procedure for the removal of various kinds of pituitary tumors. To reduce morbidity and mortality with this new technique, the surgeon must be well-trained and well-prepared. Virtual endoscopy can be beneficial as a tool for training, preoperative planning, and intraoperative support. This paper introduces STEPS, a virtual endoscopy system designed to aid surgeons in getting acquainted with the endoscopic view of the anatomy, the handling of instruments, the transsphenoidal approach, and challenges associated with the procedure. STEPS also assists experienced surgeons in planning a real endoscopic intervention by getting familiar with the individual patient anatomy, identifying landmarks, planning the approach, and deciding upon the ideal target position of the actual surgical activity. The application provides interactive visualization, navigation, and perception aids and the possibility of simulating the procedure, including haptic feedback and simulation of surgical instruments.

Index Terms—Applications, visualization techniques and methodologies.

1 INTRODUCTION

LOCATED centrally at the base of the brain, the pituitary is the central gland for the hormonal metabolism of the whole body. Pituitary adenomas are benign slow-growing tumors that arise within the pituitary gland. Comprising around 12.5 percent of all intracranial tumors, they are a common and important neurosurgical problem: If left untreated, these tumors can cause a variety of symptoms, from mild headaches and progressive visual loss due to the mass effect of the tumor to death due to electrolyte imbalance. The most common symptoms are those brought about by pituitary dysfunction, causing either a lack of hormones, yielding, e.g., fatigue and sexual dysfunction, or an excess of certain hormones with Cushing's disease, gigantism/acromegaly, and amenorrhoea-galactorrhoea as clinical syndromes.

Minimally invasive endoscopic procedures are constantly gaining importance in modern medicine [1]. Endonasal transsphenoidal pituitary surgery is a relatively new endoscopic procedure used for the removal of pituitary tumors (see Figs. 1 and 2): A rigid endoscope and surgical instruments are inserted into the patient's nose and advanced through the natural nasal airway along the *nasal septum* (the wall dividing the nasal cavity into halves) into the *sphenoid sinus*. This cavity of the skull must be entered through a small opening, the *sphenoid ostium*. In order to gain sufficient access, the natural ostium usually must be

widened using a bone punch (*rongeur*). With the transition through the ostium, the so-called *nasal phase* of the procedure is completed and the *sphenoid phase* starts. The sphenoid sinus often consists of chambers separated by bony *septa* (dividing walls). The septa usually must be removed, again using the punch, in order to get sufficient access to the *sellar floor*, a bony structure separating the sphenoid sinus from the pituitary gland and the tumor. The sellar floor is then opened using a drill or a small chisel. After enlarging this opening using the rongeur, the tumor can be divided from the surrounding tissue and removed through the patient's nose.

The procedure is associated with some risk: The *internal carotid arteries* (ICAs), which supply large parts of the brain and the optic nerves, pass just laterally to the pituitary gland behind the sellar floor and are therefore invisible to the surgeon. Damage to these structures and the pituitary gland itself must be prevented.

In order to keep risk at a low level, it is obviously of utmost importance for the surgeon to be skilled and experienced as well as to be familiar with the patient's anatomy prior to each surgical intervention. Acquisition and analysis of both MRA (Magnetic Resonance Angiography, a type of Magnetic Resonance Imaging (MRI), which depicts blood at a higher intensity level, with no contrast agent needed) and CT (Computed Tomography) scans of the patient's head are usually part of the routine preoperative diagnostic workup. Besides standard 2D viewing techniques, the surgeon can, therefore, without any additional radiological examinations, use virtual endoscopy, which can be of benefit in three different ways:

1. *Preoperative Planning*: While careful slice-by-slice-analysis of the raw images obtained from the modalities can give a sound overall picture of the situation to a skilled surgeon, this picture can still be enhanced by additionally using simulated 3D views as provided by virtual endoscopy. This helps to get

• A. Neubauer and K. Bühler are with the VRVis Research Center, Donaueystraße 1, 1220 Wien, Austria.
E-mail: {neubauer, buehler}@vrvis.at.

• S. Wolfsberger and M.T. Forster are with the Department of Neurosurgery, Medical University Vienna, A-1090 Wien, Währinger Gürtel 18-20, Austria. E-mail: stefan.wolfsberger@univie.ac.at, marie.johanna@chello.at.

• L. Mroz and R. Wegenkittl are with Tiani Medgraph AG, Diefenbachgasse 35/3/5, A-1150 Wien, Austria.
E-mail: {lukas, rainer}@tiani.com.

Manuscript received 1 Nov. 2004; revised 12 Feb. 2005; accepted 18 Feb. 2005; published online 11 July 2005.

For information on obtaining reprints of this article, please send e-mail to: tcvg@computer.org, and reference IEEECS Log Number TVCG-0134-1104.

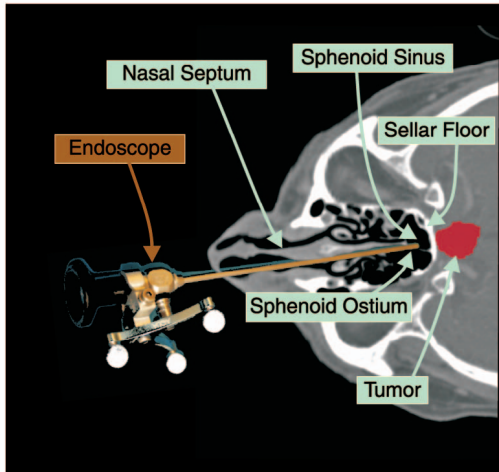


Fig. 1. The endoscope and the transsphenoidal approach.

familiar with the anatomy of the patient, decide on the path to take, and plan exactly where to open the sellar floor.

Also, semitransparency can be used: Rendering the tumor, the pituitary gland, the ICAs, and the optic nerves behind the semitransparent sellar floor (see Fig. 3, left image) can give the surgeon a profound impression of how to minimize the risks associated with the procedure.

2. *Intraoperative Support:* Virtual endoscopy can be used during the real surgical procedure both as a navigation aid and to decide on the exact opening site of the sellar floor. If the position of the real endoscope is tracked, the viewing parameters can be retrieved and transferred to the virtual endoscopy system.
3. *Training:* Safe execution of the procedure requires training. It is, for instance, a hard task for beginners to find the correct path to the sphenoid ostium. Simulation of the real surgical operation by not only providing the visual impression but also modeling haptic parameters and constraints of movement and simulating surgical instruments can serve as a tool to exercise the complete approach. Furthermore, virtual endoscopy can help the trainee to get used to visual

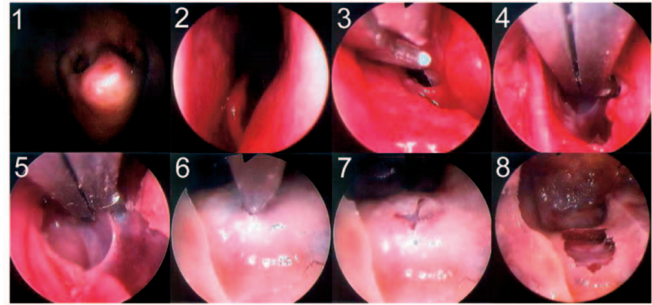


Fig. 2. Images from a real surgical intervention: 1) The endoscope is entered into the nose. 2) The endoscope is advanced along the nasal septum. 3) The sphenoid ostium is too small to pass through. 4) The ostium is enlarged using a punch. 5) Removal of sphenoid septa. 6) The sellar floor is opened using a small chisel. 7) Result after application of the chisel. 8) The hole has been widened using a rongeur.

parameters offered by endoscopy: These include image distortion, nonstereoscopic 3D images, and a limited range of light.

Virtual endoscopy has been reported to be usable for visualization of the nasal cavity and the paranasal sinuses [2]. Talala et al. assessed the benefits of a combination of virtual endoscopy with nonperspective semitransparent imaging of surface landmarks for transsphenoidal pituitary surgery [3].

This paper is an extended and revised version of the previously published article "STEPS" [4]. STEPS (Simulation of Transsphenoidal Endonasal Pituitary Surgery) is a virtual endoscopy system designed to aid neurosurgeons as a training and planning device for endonasal transsphenoidal pituitary surgery. Section 2 describes data preprocessing. Section 3 introduces the virtual endoscopy system. Section 4 discusses benefits and problems experienced by neurosurgeons working with the application. Section 5 deals with planned extensions.

2 DATA PREPROCESSING

Radiological images of the patient's head are generated using both CT and MRA. CT data give information about skull anatomy. The pituitary gland, the tumor, the optic nerves, and the ICAs are usually well captured by MRA. The size of a typical CT data set amounts to

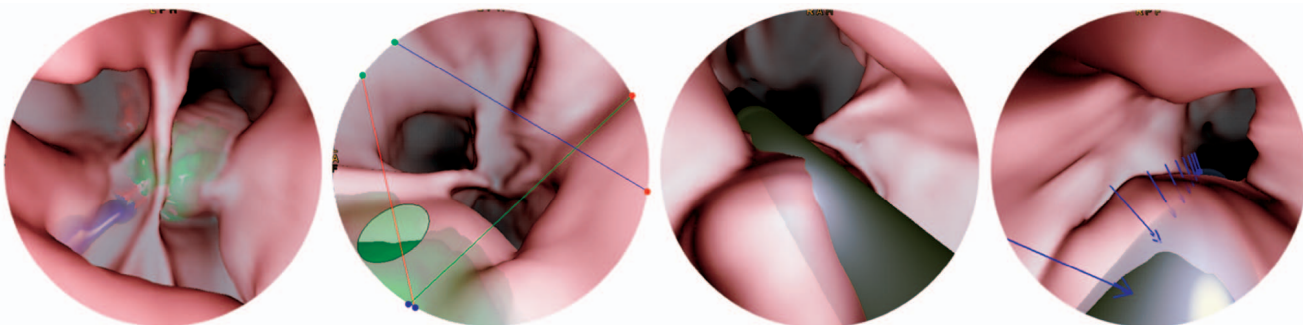


Fig. 3. Left: Inside the sphenoid sinus, the tumor (green), the internal carotid arteries (red), and the optic nerves (blue) are in the background. Center left: Preview of the virtual bone punch—the part of the preview painted dark green shows the part of the tumor that is within reach of the punch. Center right: The simulation is suspended, the endoscope (gray) has been added to the scene. Right: The endoscope has deformed soft tissue, the arrows point in the direction of deformation.

approximately $512 \times 512 \times 170$ voxels, with a voxel size of $0.44\text{mm} \times 0.44\text{mm} \times 1\text{mm}$. The MRA data sets used comprise approximately $512 \times 512 \times 100$ voxels, with a voxel size of approximately $0.4\text{mm} \times 0.4\text{mm} \times 1\text{mm}$.

Objects of interest are extracted (segmented) from the MRA image. Since the visualization of the boundaries of the investigated cavities during the virtual endoscopy procedure is based on the CT-data, the two data sets must be geometrically aligned (registered) to ensure exact correlation of object locations. After registration, objects of interest can be segmented in the MRA image and displayed at correct positions during the CT-based virtual fly-through.

STEPS was implemented as a collection of three modules, embedded into the Java-based medical workstation J-Vision (Tiani Medgraph, Vienna, Austria). Registration, segmentation, and virtual endoscopy were each encapsulated into one module. The following paragraphs describe the registration and segmentation modules. The virtual endoscopy module is covered in detail in Section 3.

2.1 Registration

The registration module aligns two three-dimensional data sets by using rigid body transformations with six degrees of freedom (three translations and three rotations). The alignment is achieved by optimizing a similarity criterion (based on mutual information) using adaptive simulated annealing.

Capek and Wegenkittl [5] developed this fast and robust module for voxel-based registration. Intuitive user interaction has been added. This allows performing a manual repositioning of the data set to be registered. To ensure finding the global optimum, usually only very rough manual preprocessing is required. A skilled user performs this task in less than two minutes. No manual definition of landmark points or similar is required. The automatic registration process takes approximately 10 seconds on a standard PC.

2.2 Segmentation

Segmentation is one of the most important problems of medical image processing and visualization. There is a constant on-going search for methods to extract certain parts of the anatomy with as little user interaction needed as possible. Automatic approaches usually rely on models describing the usual appearances of certain anatomic structures. These models are extracted from training sets and mirror the statistical variability of object shape and texture in these training sets [6].

However, in regions where tumors grow and deform the neighboring anatomy in a chaotic way, hardly describable by statistical measures, appearance models are of very limited use. Since this is often the case in the presented project, STEPS relies on semi-automatic segmentation.

The so-called *watershed-from-markers* technique is used: The watershed algorithm is a classical algorithm in mathematical morphology originally described for 2D data sets. The original algorithm suffers from oversegmentation. The watershed-from-markers technique improves this by adding simple user interaction: Markers are placed into the volume, each marker related to either an object of interest or the background. This additional input adds information

about object connectivity to the algorithm and reduces the probability of detecting false borders. Felkel et al. [7] developed an extremely memory efficient and fast implementation of the watershed-from-markers algorithm, extended to 3D, with acceptable performance even for large data sets. The increased speed (the calculation takes about 25 seconds on average on a P4-3000 MHz) allows multiple passes of editing the markers and running the algorithm until the result is satisfactory. Minor adaptations can be (and usually are) carried out manually in a paint/erase fashion. Manual editing is important to retain complete control over the segmentation process and to achieve a medically sound segmentation result. Manual adaptation takes about 50 percent of the time spent on segmentation, i.e., approximately 15 minutes per patient.

User interaction in the segmentation module takes place via an MPR (multiplanar reconstruction, see also Section 3.3.1). A 3D preview of the segmented object can be viewed at any time during segmentation.

3 THE VIRTUAL ENDOSCOPY SYSTEM

3.1 Visualization

To be applicable, a virtual endoscopy system must provide instant reaction to user input. This results in the need for high frame rates and, thus, fast visualization. Many existing systems therefore render isosurfaces [1]. Roughly two groups of isosurfacing algorithms can be identified: those which use hardware-accelerated polygonal surface rendering [8], [9] and those using first-hit ray casting [10].

In first-hit ray casting, for each pixel, a viewing ray is tracked through the volume and intersected with the isosurface. As soon as the first intersection has been found, ray tracking is stopped and the pixel is shaded. Plain first-hit ray casting without special hardware-acceleration is usually too CPU-intensive to be performed at interactive frame rates. Therefore, optimization strategies which either reduce ray traversal distances or increase ray tracking efficiency have to be applied [11], [12]. There are a number of approaches which pay for increased performance by reducing image quality [13], [14], [15] or flexibility [16].

For STEPS, two optimized first-hit ray casting techniques, both without special hardware acceleration, are applied for reasons of hardware independence. One of the ray casting techniques is used to depict the foreground (the inner walls of the investigated cavities, an isosurface in the original CT data) and one to visualize the presegmented background objects.

Since the foreground always covers the complete image and, thus, each pixel has to be processed, an optimized image-order (pixel by pixel) ray casting technique is applied [4]. The technique achieves interactive frame rates by exploiting interpixel coherency and accelerating memory access and empty space traversal. Rays are tracked efficiently from cell-boundary to cell-boundary [17] in order to avoid missing thin structures.

Background visualization is performed on a synthetic background volume which is generated by reconstructing objects of interest from binary segmentation masks. For objects which can be represented by an isosurface in the

original MRA or CT data (e.g., the ICAs in the MRA data set), the mask is comprised of the objects and its surroundings—the original values of voxels inside the mask are simply transferred to the background volume and the object boundaries can be extracted during the virtual endoscopy by interactively adjusting the isovalue for the object. Other objects are reconstructed by generating synthetic isosurfaces. The boundaries of these objects are defined by the binary segmentation masks.

Background objects cover a varying number of pixels, depending on their shapes and their positions relative to the viewing frustum. Therefore, the background is generated by applying the optimized object-order algorithm *cell-based first-hit ray casting* [18], [19]. For each object, an isosurface is displayed; the threshold can be adapted individually for each object.

In each frame, the two resulting images (foreground and background) are merged using a z-buffer. The difference of the z-values between foreground and background determines foreground opacity. For more detail, please refer to the previous publication [4].

3.2 Surgery Simulation

The user can decide to either move freely inside the patient's head or to simulate the endonasal transsphenoidal procedure. With STEPS, the process of advancing the endoscope through the narrow nasal cavities and finding the correct way to the sphenoid sinus can be simulated, as well as the process of widening the sphenoid ostium, the removal of sphenoid septa, and one of the most crucial parts of the actual surgical operation, the process of opening the sellar floor. Simulation is used mainly for training purposes, but can also provide information about accessibility and the risks and obstacles to expect, which also makes it useful for preoperative planning. For the simulation, it is important to model the degrees of freedom and constraints experienced when working with a rigid endoscope as accurately as possible. The user should be confronted with similar challenges as in the real procedure and guided by haptic feedback. In STEPS, the virtual endoscope is maneuvered using a force-feedback joystick (WingMan Force 3D, Logitech, <http://www.logitech.com>).

3.2.1 Control

The simulation starts with the user selecting a *start point* p_{start} and a *target point* p_{target} in the CT volume. The start point defines the position where the endoscope is entered into the patient and is usually located inside one of the patient's nostrils. The target is on the sellar floor. After selection of the two points, the eye point p_{eye} is automatically moved to p_{start} and the viewing vector points toward p_{target} . The user can move forward (into the viewing direction) and backward using joystick buttons and can rotate the virtual endoscope by moving the joystick handle.

The endoscope is modeled as the line segment connecting p_{start} and p_{eye} . It can only be rotated around p_{start} (see Fig. 4).

3.2.2 Collision Detection and Deformation

On its way to the sellar floor, the endoscope passes through a number of cavities in the skull. These cavities are separated from each other by bony or cartilaginous walls,

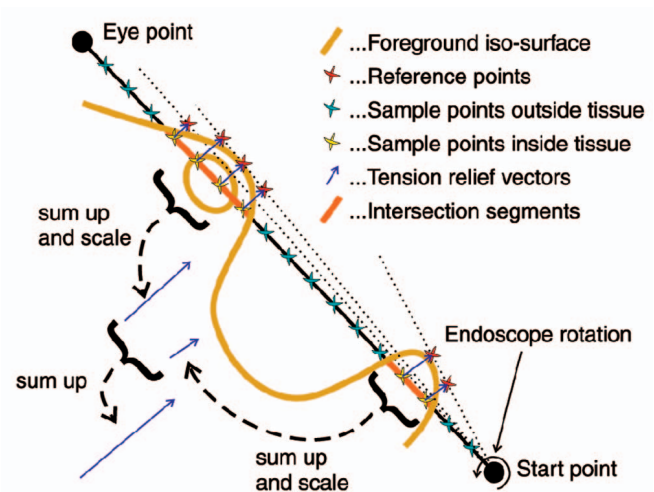


Fig. 4. Calculation of the tension relief vector after endoscope rotation.

so-called *septa*. Bones form barriers for the movement of the endoscope. Soft (i.e., deformable) tissue (e.g., mucosa and the cartilaginous part of the nasal septum) also poses resistance to movement, but can be deformed to some extent. In fact, soft tissue must be deformed to gain access to the sphenoid ostium.

STEPS currently simulates soft tissue deformation in an implicit way. For performance reasons, the deformation is not modeled geometrically, but manifests itself only in relieved constraints of movement and in force-feedback. Volume adjustments resulting in optically or even physically correct deformation of the isosurface are computationally expensive and not straightforward. Implicit deformation as performed by STEPS is less costly and yields equivalent results as long as it does not affect the eye point.

STEPS uses two thresholds, a soft threshold, t_{soft} , and a firm threshold, t_{firm} , to define the data range of soft tissue, with $-1024HU < t_{soft} < t_{firm} < 3071HU$, HU being short for *Hounsfield Units*. Threshold t_{soft} is equal to the foreground isovalue. Threshold t_{firm} represents the boundary between soft tissue and bony structures. Apart from the eye-point, the virtual endoscope may be pushed into soft tissue, i.e., implicitly deform it. No part of the virtual endoscope is allowed to be moved to intensity greater than t_{firm} . Because p_{eye} may not cross the foreground isosurface, tissue deformation can only be performed via endoscope rotations and not by advancing the endoscope. In the vicinity of p_{eye} , geometric deformation would therefore be of significant advantage to prevent artificial limitations of endoscope movement and to increase the overall realism of the simulation. A suitable model is currently being developed (see Section 5).

Because the bones encasing the sphenoid sinus are very thin, they are captured by the CT at a lower intensity level than more massive bones. The intensity level of these nondeformable barriers is, in fact, similar to the intensity level of larger deformable structures, as experienced near p_{start} , like the cartilaginous nasal septum. Therefore, the deformability of virtual structures cannot be made dependent on CT-intensity values alone, but also spatial

information must be taken into account. In other words, the value t_{firm} must change along the path of the endoscope. The same also holds, to a lesser extent, for the threshold t_{soft} . To achieve a behavior similar to reality, the value t_{firm} is defined as a function of space, the threshold $t_{firm}(V)$ used at voxel V depending on the distance d_{sv} between V and p_{start} . Currently, the following step-like function is used, values are in HU:

$$t_{firm}(V) = \max(f(d_{sv}), t_{soft} + 1) \quad (1)$$

with

$$f(d_{sv}) = \begin{cases} 450 & d_{sv} \leq \frac{8}{10} d_{st} \\ -450 & d_{sv} \geq \frac{9}{10} d_{st} \\ -\frac{1}{d_{st}} 9,000 d_{sv} + 7,650 & \text{otherwise,} \end{cases} \quad (2)$$

where d_{st} denotes the distance between p_{start} and p_{target} . The formula is motivated as follows: Soft tissue that must be deformed is usually encountered within the first 80 percent of the path of the endoscope. The transition through the ostium takes place after about 90 percent of the traveling distance. So, when p_{eye} is inside the sphenoid sinus, the tip of the endoscope experiences a small threshold, t_{firm} , compared to the part of the endoscope that is inside the nose. This means that, during the sphenoid phase, parts of the endoscope close to p_{start} can still deform tissue, while movement of the tip of the endoscope is bounded by almost any visible structures, which is realistic because the bones encasing the sphenoid sinus are usually only covered by a very thin layer of mucosa. As a result, freedom of movement during the sphenoid phase is confined by the size of the (enlarged) ostium, as is also the case in reality.

The threshold used for foreground visualization, $t_f = t_{soft}$, is usually selected such that soft tissue, as seen during the nasal phase, and thin bones, as seen especially during the sphenoid phase, are depicted. The intensity levels are similar, but there may not be an isovalue that is equally suitable for both. For this reason, the user can select to make t_f and, therefore, t_{soft} , a linear function of time, depending on the distance between p_{start} and p_{eye} . Thus, t_{soft} might slightly change in time, with p_{eye} moving.

In the simulation, actions which would cause the eye point to move across the visualized foreground isosurface are disallowed. To prevent the eye-point from jumping past thin walls, the path from the previous to the current eye point is sampled after each movement. If an intersection with the foreground isosurface defined by t_{soft} is found, the endoscope is returned to its previous position. Sampling is performed by casting a viewing ray such that any visible structures are encountered with a very high probability. To prevent further disallowed collisions, this process is repeated for a number of points evenly distributed along the endoscope, with the only difference being that, for points other than the eye-point, the threshold t_{firm} is used instead of t_{soft} .

Collision detection is only performed for the center axis of the virtual endoscope and not for its complete diameter. This is compensation for the fact that the surgeon in reality has an increased amount of freedom by performing slight intuitive movements.

3.2.3 Force-Feedback

If a part of the virtual endoscope is inside soft tissue, force-feedback is exerted with the goal to press the endoscope back into regions of lower image intensity. The direction and intensity of force-feedback are calculated using the following algorithm, illustrated in Fig. 4: The sorted set of intersection points of the virtual endoscope with the foreground isosurface is a sequence of, alternately, *entry points* (of the endoscope into tissue denser than t_{soft}) and *exit points* (out of tissue denser than t_{soft}). An *intersection segment* is a segment of the virtual endoscope that is associated with tension because tissue has been deformed. An intersection segment is always confined by an entry point and an exit point, but may span more than two isosurface intersections for the following reason: Inside soft tissue, there may be, due to noise, small connected sets of voxels whose values are smaller than t_{soft} . Intersections with these regions do not break an intersection segment apart. For example, the endoscope depicted in Fig. 4 intersects the isosurface six times (three entry points and three exit points). Still, there are only two intersection segments.

The direction into which the endoscope should be rotated in order to reduce the tension caused by the i th intersection segment S_i most effectively is defined by the *tension relief vector* $v_{tension}^i$.

Tension relief vectors are calculated as follows: The CT data is sampled at evenly spaced points along the virtual endoscope, from p_{start} to p_{eye} . The intersample distance d remains constant throughout the simulation. The λ th sample point p_{sample}^λ is therefore always located at distance $\lambda \cdot d$ from p_{start} . For each λ , a reference point p_{ref}^λ is stored, indicating the last position of a λ th sample point that was outside soft tissue. The i th tension relief vector is calculated as follows:

$$v_{tension}^i = c^i \frac{v}{|v|} \quad (3)$$

with

$$v = \sum_{\{\lambda | p_{sample}^\lambda \in S^i\}} (p_{ref}^\lambda - p_{sample}^\lambda) \quad (4)$$

and the coefficient c^i being a function directly proportional to the largest intensity value in the (slightly low-pass filtered, to reduce the impact of noise) sequence of intensity samples inside S_i and directly proportional to the distance between the center point of S_i and p_{start} (simulating torque). Finally, the sum of all tension relief vectors is transformed to screen space and directly passed on to the force-feedback-joystick. The intensity of force-feedback is determined by the Euclidean norm of the resulting vector.

3.2.4 Suspended Simulation

During simulation, the user's range of view is very limited. It can, as in real endoscopy, be enhanced by switching to an angled endoscope (see Section 3.3.3). This, however, is usually not sufficient to acquire a complete picture of the current state of the simulation, which, however, might be useful, especially for training purposes.

STEPS provides the possibility of suspending the simulation. This means that the current state of the simulation is frozen and the application switches to a temporary free-flight mode. During suspension of the simulation, the endoscope is added to the rendered scene as an additional background object (see Fig. 3, third image). The user can view the current situation and assess the state of the simulation from arbitrary view points. If the way of the endoscope is blocked or endoscope rotations are prevented, suspending the simulation helps to find out the reasons.

Since, currently, there is no geometrical soft tissue deformation, the virtual endoscope, pushed far enough into soft tissue, might make the impression of having sunk into the tissue (see Fig. 3, right). This can be quite confusing and it is sometimes hard to see from which direction soft tissue has been deformed. To give a more intuitive picture, the direction of deformation is depicted using arrows. For this, the information determining the application of force-feedback is reused: For each intersection segment, a tension relief vector was calculated. This vector points, approximately, in the direction from which the endoscope came when being pushed into soft tissue. Each arrow points accordingly in the direction of the inverted tension relief vector to the surface of the endoscope. Arrow length corresponds to the length of the tension relief vector and, therefore, to the amount of force-feedback caused by the corresponding intersection segment.

As soon as the simulation is continued, the endoscope is removed by restoring the original voxel values from the hard disk.

3.3 Additional Features

The following features are available during both free flight and simulation:

3.3.1 Linked Views

The screen is separated into four parts (see Fig. 5): the 3D-view displaying the simulated endoscopic image and a multiplanar reconstruction (MPR) consisting of three *section views* displaying raw CT data and selected background objects on three mutually orthogonal cross-sections of the volume. The current eye point is the intersection point of the three cross-sections. The cutting planes defining the cross-sections can be arbitrarily rotated around the eye point by the user. The section views are linked to each other—the intersection between any pair of cross-sections is displayed in both section views. The user can also link any of the section views to the 3D view. If a section view is linked, the intersection of its cutting plane with the viewing frustum is displayed in both the 3D view and the section view. This improves navigation tremendously since structures seen in the 3D view are mapped to their locations in the CT data in a very intuitive way. This technique also serves as an instrument to find a suitable threshold for the foreground isosurface.

3.3.2 Simulation of Surgical Instruments

Surgical instruments are needed on several occasions during the procedure: The sphenoid ostium must be widened using a bone punch in order to gain sufficient

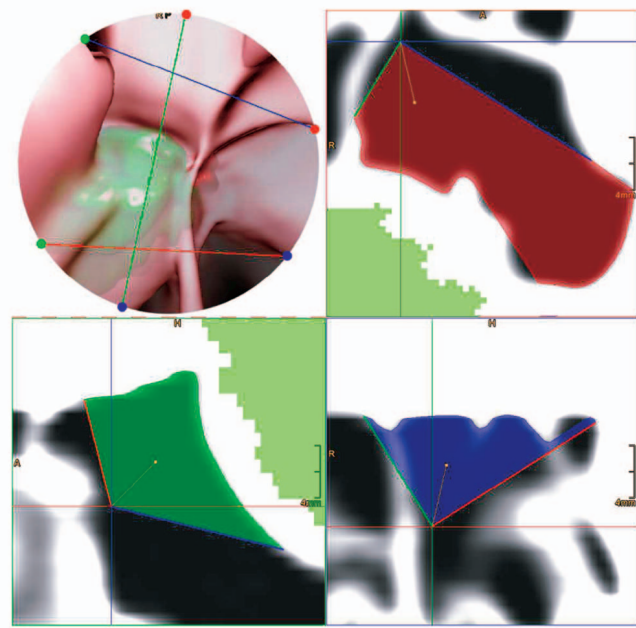


Fig. 5. The 3D view (top left) and the three section views are linked to each other. Note the septum on the left-hand side of the 3D view, rendered incompletely due to a too high threshold, as confirmed by the blue section view (bottom right).

access to the sphenoid sinus. Inside the sphenoid sinus, sphenoid septa might have to be removed to uncover the complete sphenoid anatomy, again using the bone punch. Finally, the sellar floor must be opened. Since the bone punch cannot be used to create, but only to enlarge an opening, a small hole is created using a drill or chisel, then the hole is widened using the rongeur.

In the simulation, there is no need to use separate instruments, i.e., both drill and bone punch. Instead, STEPS simulates a bone punch with a built-in drill, which can also be used to create an opening.

The user can select the radius r and the depth d of the rongeur to use. The virtual rongeur is maneuvered using the mouse. In previous versions of the system, the punch was activated using a joystick button [4]. Control of the rongeur was, however, completely delegated to the mouse since, also in reality, the endoscope and the rongeur are two disjoint instruments. When the mouse cursor is moved over the endoscopic image, a preview is shown underneath the cursor, indicating the expected impact of the punch (see Fig. 3, second image). A mouse-click leads to activation of the rongeur at the *activation point* $p_{activation}$, i.e., the position currently rendered below the mouse cursor.

The bone punch is, logically, a rigid instrument, like the endoscope. Therefore, it can be maneuvered with the same constraints as the endoscope. During the simulation, the punch preview is only shown and the punch can only be activated if the mouse cursor points to a position on the surface which can be reached from p_{start} . For testing the reachability of the rongeur, the same rules are applied as for testing the reachability for the simulation of endoscope movement (see Section 3.2).

The preview of punch impact is generated using the following algorithm: The punch vector v_p is identified from

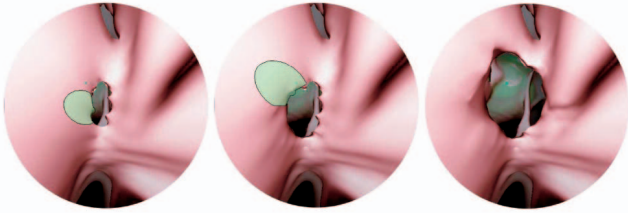


Fig. 6. Using the virtual bone punch, the sphenoid ostium is enlarged to gain sufficient access to the sphenoid sinus.

the coordinates of the mouse cursor. It points from p_{start} (during simulation) or from p_{eye} (during free flight) to $p_{activation}$. If the simulation mode is active, the reachability is checked by testing whether this vector intersects the isosurface defined by threshold t_{firm} .

If the result of the test is negative (i.e., there is no intersection), n ray source points are established. These points are evenly distributed on a base circle which is located on a plane perpendicular to v_p and initially centered on $p_{activation}$. The radius of the base circle is the user-specified radius r . The base circle is then projected orthogonally, in direction of $\pm v_p$, to the isosurface. The projection is displayed as a polygon in black color (see Fig. 6). This polygon marks the *optical region of impact (OROI)*. Next, the plane containing the base circle is shifted in the direction of its normal ($-v_p$) until the base circle does not intersect the isosurface any more. At the end of this procedure, it marks the beginning of the *physical region of impact (PROI)* of the punch. The PROI is comprised of all parts of the surface that are rendered inside the OROI and whose distance d_s from p_{start} satisfies the condition $d_{sb} \leq d_s \leq d_{sb} + d$, with d_{sb} being the distance between p_{start} and the shifted base plane.

In the punch preview, pixels depicting a part of the PROI are highlighted (light green color). If a part of a background object is inside the PROI, the preview shows the color of the object (see second image, Fig. 3).

The user can activate the punch with a mouse click. The punch cuts a cylindrical piece of the specified radius and depth out of foreground tissue. This is done by changing voxel values in the CT volume accordingly—the impact can therefore also be observed in the MPR. Information about each activation of the punch is stored to disk. Thus, any volume manipulations can be undone, allowing, for instance, a restart of the simulation.

Fig. 6 shows the process of virtually widening the sphenoid ostium. Fig. 7 depicts the simulation of the process of opening the sellar floor.

3.3.3 Angled Endoscopes

In reality, the optical system of the endoscope can be adapted to satisfy special needs of the surgeon. Special objectives attached to the endoscope cause the principal viewing vector to deviate from the axis of the endoscope by a certain angle. The angle can be selected to be 12, 30, or 45 degrees for the forward oblique view, 70 or 90 degrees for the lateral view, and 120 degrees for the retrospective view. In case of an angled objective, rotation of the endoscope around its axis does not result in simple rotation of the endoscopic image around its center with constant

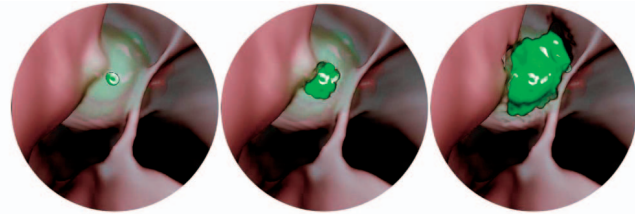


Fig. 7. The virtual bone punch is used to open the sellar floor and uncover the tumor.

view direction, but the view direction is also rotated around the endoscope axis. This tremendously enlarges the surgeon's field of view, but, at the same time, makes handling of the endoscope far more complicated. The larger the angular deviation, the less intuitive the visual response to movements of the endoscope is. Therefore, usually, the following strategy is pursued: In the beginning, a non-angled endoscope is inserted to identify the path to the point of application of surgical activity. As soon as instruments have been added and the path is laid, the endoscope can be retrieved, the objective replaced, and the endoscope inserted again, next to the surgical instruments.

STEPS supports simulation of angled endoscopes and, therefore, helps the user get used to the difficulties arising with maneuvering an endoscope with an angled viewing axis. The angle can be selected at any time. The strategy used in the virtual case is as follows: Use an angle of 0 to get to the position of interest, then change the angle and start looking around by rotating the endoscope around its axis (twisting the joystick handle).

3.3.4 Surface Rigidity

A smoothly shaded isosurface gives no information apart from shape. All structures, regardless of the types of tissue they represent, are displayed in the same way and in the same color. The information given can be enhanced by adding meaningful texture to the surface. Information about tissue rigidity, for example, helps the surgeon in a variety of ways. First of all, the intuitive mapping between the 3D view and reality is enhanced. In real endoscopy, it is easy to differentiate between soft and rigid tissue—if there are not sufficient visual hints, a structure can be tested for rigidity quite easily by checking the deformability. This cannot yet be done in STEPS. However, showing surface rigidity as a texture reduces that disadvantage. If the simulation is suspended, the texture can help identify why endoscope rotation is prevented in a certain direction. The texture also gives information about isosurface reliability: The more rigid the displayed structure is, the steeper the gradient in the region of the isosurface is, implying a smaller change of the surface as a response to isovalue adaptation. On the other hand, parts of the surface which are far from rigid structures, can be expected to change significantly if the isovalue is slightly changed.

STEPS can generate texture showing tissue rigidity in various ways, namely, volume rendering, double-isosurfacing, and simple color-mapping. Each of these has its advantages and disadvantages.

Volume rendering is performed by extending each foreground viewing ray a few steps beyond the isosurface,

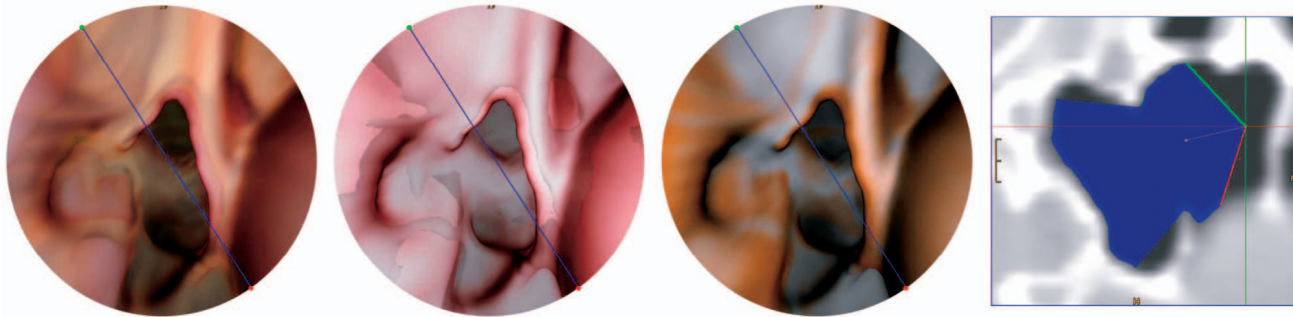


Fig. 8. Displaying surface rigidity: Left: Volume rendering. Center left: Double isosurfacing. Center right: Color-coding. Right: Section through original CT data.

using a transfer function to accumulate a pixel color. This accumulated color is then blended with the color assigned by foreground isosurfacing. The extended ray is stopped either after reaching the user-defined maximum number of traversal steps, after reaching full opacity, after encountering a data value lower than the foreground isovalue, or after entering a background object. Volume rendering is the most flexible (the transfer function can be adapted by the user) but slowest alternative, increasing rendering cost by 20 to 60 percent, depending on the maximum number of accumulation steps.

Double isosurfacing is performed by defining a second foreground isovalue t_{ff} , representing the transition between soft tissue and bone. Similarly to volume rendering, rays are extended by a limited number of steps and stopped when a background object is entered. An isosurface intersection is performed as soon as a data value larger than t_{ff} is encountered. Double isosurfacing gives a good impression of the proximity of bone and works faster than volume rendering (speed decreases by about 13 percent), but generates less intuitive images.

The third method, simple color-coding, is the fastest one, with a speed decrease of about 10 percent: Tissue rigidity is mapped to surface color. The color scale ranges from red (soft tissue) to white (massive bone). Again, each foreground ray is elongated by a limited number of steps. After each step, the tissue density at the current ray position is sampled using trilinear interpolation. The calculated value is decreased by $c \cdot d$, with c being a positive constant and d the distance to the foreground isosurface. This emphasizes rigid tissue that is very close to the isosurface. The largest value encountered for each ray is used to calculate the texture color. In order to add to the understanding, the number of sampling steps, and the constant c , the sensitivity to distance can be adjusted by the user.

Fig. 8 shows example images for all three methods and a slice of original CT data showing part of the scene depicted in the 3D renderings. The number of ray steps beyond the isosurface was 20 for the volume rendering and 10 for the other two methods. Currently, we consider simple color mapping and volume rendering the most effective methods. Color mapping provides an intuitive visual result with little computational overhead. Volume rendering is more expensive, but gives a better impression of distances between bones and the surface. Both techniques, especially volume

rendering, in fact, were accepted very well by the neurosurgeons.

3.3.5 Depth Shading

For the endonasal approach, an endoscope with a relatively small diameter (4mm or less) must be used to maximize maneuverability. Thin endoscopes, however, have the disadvantage of reduced light transport capabilities. The range of light emitted by the endoscope is therefore very limited, as can be seen in Fig. 2.

STEPS uses depth shading to simulate this phenomenon and to enhance depth perception. The left image in Fig. 9 was rendered without using depth shading. Due to the shaded surface, depth is anticipated, but only very vaguely. By darkening the scene with increasing distance from the eye point (Fig. 9, center and right images), depth is communicated much more effectively. The user can interactively adjust the intensity of depth shading.

4 DISCUSSION

STEPS has been developed in close cooperation with neurosurgeons and tested within the scope of a medical research project [20]. This section deals with problems and questions brought up during development and evaluation and assesses the value of the application for training and preoperative planning.

4.1 Time Consumption

Preprocessing of objects is time-consuming. Registration usually takes no more than 3 minutes, even if manual preprocessing is required. Segmentation, however, is still a tedious task because, usually, much manual work must be done. Segmentation of the ICAs is generally performed quickly since the boundaries do not have to be specified

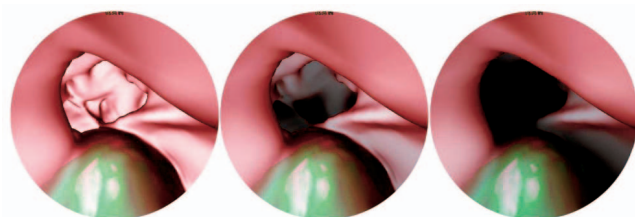


Fig. 9. Left: No depth shading. Center: Low intensity depth shading. Right: High intensity depth shading.

exactly. Segmentation of the tumor, the gland, and the optic nerves, however, takes more effort. All in all, the time needed for segmentation amounts to approximately 30 minutes per patient. The virtual endoscopy is then usually set up within 3 minutes, including reconstruction of all objects. The time needed for preoperative investigation depends heavily on the surgeon's experience and patient anatomy, but can be estimated as roughly 30 minutes.

STEPS attempts to keep above a minimum frame rate during interaction by rendering images of adaptive size and scaling them to fit the screen. The minimum frame rate is adjusted by the user. Typically, a circular image with a diameter of 500 pixels can be rendered at 10 fps on a P4-3000 MHz.

4.2 Visualization

Feedback from the medical doctors concerning the visualization was predominantly positive. The surgeons observed a good correlation of the intraoperatively observed and the virtually reconstructed anatomy. Rendering quality and speed is considered sufficient.

4.2.1 Background Visualization

There is an on-going discussion among people associated with this project about whether background objects reconstructed from binary segmentation masks should be smoothed or not.

Smoothing alters object appearance as communicated by the segmentation mask created by the physician. Smoothing is always associated with a low-pass filtering process and, therefore, reduces the amount of information given. This can, in a medical application, lead to severe problems. On the other hand, smoothing makes objects look more realistic. This increases the acceptance among physicians and their patients. Therefore, if the virtual endoscopy system is used for training purposes, in our opinion, background objects can and should be smoothed to some extent. In the case of preoperative planning, there must be the possibility of adding unsmoothed objects. If the physician chooses to smooth objects, the additional error introduced by smoothing must be constrained and the maximum error must be communicated to the physician. Also, physicians must (and usually do) keep in mind that the image provided by virtual endoscopy does not depict reality, but a snapshot reconstructed from discretized scalar data. In the worst case, the errors made during scalarization, discretization, recording, segmentation, (possibly) smoothing, and reconstruction add up. The result is an approximation which must not be mistaken for a completely accurate representation. Therefore, an assumption like the one that cutting the sellar floor at a certain position is safe because, in the virtual endoscopy, a vessel is missed by a few pixels is not at all justifiable, regardless of whether objects were smoothed.

In STEPS, the user can select, for each reconstructed background object, whether it should be smoothed and adjust the amount of smoothing applied. The filtering process is explained in a previous publication [18].

The way some objects (e.g., the ICAs, see Section 3.1) are reconstructed in STEPS, by passing the original MRA data almost through the complete rendering pipeline and

handing the crucial decision, the definition of the boundaries, i.e., the adjustment of the isovalue, to the physician, helps the physician correctly interpret the visual result, its relation to the data, and its expressiveness. Also, these objects do not have to be reconstructed from a binary mask; the appearance of the object boundaries as recorded by the MRA is maintained. The question of whether smoothing should be applied does not arise for these objects. The segmentation result can always be cross-checked in the MPR in which the voxels belonging to a visible background object are highlighted (see Fig. 5). Therefore, this way of segmenting is no more error-prone than performing the segmentation solely upon original MRA data.

4.2.2 Foreground Visualization

A well-known problem of isosurfacing based on intensity values as obtained from a CT is that isosurfaces generally do not entirely correspond to boundaries as observed in real endoscopy. The reason is that optimal threshold values differ between structures, even between different structures seen in the same image. Due to the fact that CT data appear slightly low-pass filtered, a small variation of the threshold can have quite a drastic effect on image appearance. So, a slightly wrong threshold can result in a certain structure appearing quite different from its real shape as seen during the real endoscopic procedure. Adding an additional dimension to user interaction by providing interactive threshold adjustment significantly reduces the impact of this problem. Also, again, the increased amount of control over the visual outcome handed to the user helps them get a feel for the relation between raw data and resulting images and makes them aware of the fact that what they see is not equal to reality, but an approximation. On the other hand, it is worth noting that it takes some time and training for medical users to get used to the provided amount of control. However, according to our experience, after this learning phase, they are convinced of the advantages, as pointed out above, and they also have a better appreciation and understanding of the strategies and processes underlying the visual and haptic feedback.

4.3 Application as a Training Device

STEPS was recognized to be beneficial as a training tool. The traditional way to train transsphenoidal pituitary surgery is cadaver training. Advantages of virtual endoscopy over cadaver training include lower cost and effort, unbounded reproducibility, an improved learning curve by utilizing navigation and perception aids which can, to some extent, replace an instructor, and the possibility of suspending the training simulation and viewing the situation from perspectives unreachable for a real endoscope. These advantages make virtual endoscopy a valuable addition to cadaver training.

Still, training as currently offered by STEPS suffers from some drawbacks: A force-feedback-joystick is a cheap, but not ideal, solution to the problem of modeling haptic feedback as experienced in real life. Near-exact simulation would require using some highly sophisticated haptic feedback device which comes at a very high cost. However, although force-feedback as applied by STEPS is not completely realistic, it manages to give the user a very

intuitive notion of soft tissue deformation, which substantially increases the value of the simulation. Currently, only a part of the simulation is enhanced using force-feedback. There is no force-feedback triggered by the virtual bone punch and by tissue encounters during advancement of the endoscope.

The lack of geometric tissue deformation reduces the maneuverability of the virtual endoscope. It, for instance, sometimes prevents the virtual endoscope from being retreated. The reason is that if the endoscope intersects the visible isosurface (see Fig. 4), pulling it in the direction of the inverted viewing vector, toward p_{start} , would at some point cause p_{eye} to cross the isosurface, which is forbidden. In real endoscopy, this restriction does not exist. A possible fix to this problem would be to use the reference points used for the calculation of force-feedback to define a path back to the nose which leads around the deformed structure. Retreating along this path would, however, cause rotations of the endoscope not triggered by the user. This has not yet been implemented.

4.4 Application for Preoperative Planning

The value of STEPS in preoperative planning has been assessed in a series of patients. The system was found useful in a variety of ways. For details, refer to our earlier publications [18], [20].

5 FUTURE WORK

A model for geometric deformation of soft tissue is being developed. This is expected to greatly enhance the realism of the simulation and further increase the applicability of the system as a training device. Geometric tissue deformation will, during the simulation, be applied only in the vicinity of the eye point, i.e., if the effects are visible to the user. An exception to this will be the suspended simulation. When the simulation is suspended, geometrical deformation will be exerted everywhere the endoscope touches the surface, offering the user a better notion of the position of the endoscope and the dynamics of the simulation.

Another planned extension is to connect the application to a tracking system to optimize intraoperative support.

ACKNOWLEDGMENTS

Parts of this work were carried out within the research project "Effective Medical Visualization" at the VRVis research center (<http://www.vrvis.at>) in Vienna, Austria, which is funded by the Austrian governmental research project Kplus.

REFERENCES

- [1] D. Bartz, "Virtual Endoscopy in Research and Clinical Practice," *Eurographics State-of-the-Art-Reports 2003*, (S4), 2003.
- [2] M.D. Nicola, L. Salvolini, and U. Salvolini, "Virtual Endoscopy of Nasal Cavity and Paranasal Sinuses," *European J. Radiology*, vol. 24, pp. 175-180, 1997.
- [3] T. Talala, T. Pirila, V. Karhula, E. Ilkko, and I. Suramo, "Preoperative Virtual Endoscopy and Three-Dimensional Imaging of the Surface Landmarks of the Internal Carotid Arteries in Trans-Sphenoidal Pituitary Surgery," *Acta Otolaryngol.*, vol. 120, pp. 783-787, 2000.

- [4] A. Neubauer, S. Wolfsberger, M. Forster, L. Mroz, R. Wegenkittl, and K. Bühler, "STEPS—An Application for Simulation of Transsphenoidal Endonasal Pituitary Surgery," *Proc. IEEE Visualization 2004*, pp. 513-520, 2004.
- [5] M. Capek and R. Wegenkittl, "Robust and Fast Medical Registration of 3D-Multi-Modality Data Sets," *Proc. Medicon '01*, pp. 515-518, 2001.
- [6] T.F. Cootes, G.J. Edwards, and C.J. Taylor, "Active Appearance Models," *Proc. European Conf. Computer Vision 1998*, vol. 2, pp. 484-498, 1998.
- [7] P. Felkel, R. Wegenkittl, and M. Bruckschwaiger, "Implementation and Complexity of the Watershed-from-Markers Algorithm Computed as a Minimal Cost Forest," *Proc. Eurographics '01*, pp. 26-35, 2001.
- [8] R. Hietala and J. Oikarinen, "A Visibility Determination Algorithm for Interactive Virtual Endoscopy," *Proc. IEEE Visualization '00*, pp. 29-34, 2000.
- [9] L. Hong, S. Muraki, A. Kaufman, D. Bartz, and T. He, "Virtual Voyage: Interactive Navigation in the Human Colon," *Proc. Int'l Conf. Computer Graphics and Interactive Techniques (SIGGRAPH '97)*, pp. 27-34, 1997.
- [10] M. Levoy, "Display of Surfaces from Volume Data," *IEEE Computer Graphics and Applications*, vol. 8, no. 5, pp. 29-37, 1988.
- [11] B.T. Stander and J.C. Hart, "A Lipschitz Method for Accelerated Volume Rendering," *Proc. IEEE Visualization '95*, pp. 107-114, 1995.
- [12] K.J. Zuiderveld, A.H.J. Koning, and M.A. Viergever, "Acceleration of Ray-Casting Using 3D Distance Transforms," *Visualization in Biomedical Computing II, Proc. SPIE 1808*, pp. 324-335, 1992.
- [13] K. Kreeger, I. Bitter, F. Dachille, B. Chen, and A. Kaufman, "Adaptive Perspective Ray Casting," *Proc. Symp. Volume Visualization*, pp. 55-62, 1998.
- [14] H. Qu, M. Wan, J. Qin, and A. Kaufman, "Image Based Rendering with Stable Frame Rates," *Proc. IEEE Visualization 2000*, pp. 251-258, 2000.
- [15] A. Vilanova Bartroli, R. Wegenkittl, A. König, and M.E. Gröller, "Perspective Projection through Parallely Projected Slabs for Virtual Endoscopy," *Proc. SCCG '01—Spring Conf. Computer Graphics*, pp. 287-295, Apr. 2001.
- [16] R. Wegenkittl, A. Vilanova, B. Hegedüs, D. Wagner, M.C. Freund, and M.E. Gröller, "Mastering Interactive Virtual Bronchoscopy on a Low-End PC," *Proc. IEEE Visualization 2000*, pp. 461-464, 2000.
- [17] J. Amanatides and A. Woo, "A Fast Voxel Traversal Algorithm for Ray Tracing," *Proc. Eurographics '87*, pp. 3-10, 1987.
- [18] A. Neubauer, M. Forster, R. Wegenkittl, L. Mroz, and K. Bühler, "Efficient Display Of Background Objects for Virtual Endoscopy Using Flexible First-Hit Ray Casting," *Proc. Eurographics/IEEE TCVG Symp. Visualization (VisSym '04)*, pp. 301-310, 2004.
- [19] A. Neubauer, L. Mroz, H. Hauser, and R. Wegenkittl, "Cell-Based First-Hit Ray Casting," *Proc. Eurographics/IEEE TCVG Symp. Visualization (VisSym '02)*, pp. 77-86, 2002.
- [20] S. Wolfsberger, M.-T. Forster, M. Donat, A. Neubauer, K. Bühler, J. Hainfellner, and E. Knosp, "Virtual Endoscopy Is a Useful Device for Training and Preoperative Planning of Transsphenoidal Endoscopic Pituitary Surgery," *Minimally Invasive Surgery*, no. 47, pp. 214-220, 2004.



André Neubauer received the master's degree in computer science from the Vienna University of Technology in 2001. He is currently working as a junior researcher at the VRVis Research Center and is about to complete his PhD thesis about virtual endoscopy. His research interests include medical visualization and image segmentation. He is a student member of the IEEE.



and cadaver dissections.

Stefan Wolfsberger graduated from Vienna Medical School in 1997. He started his neurosurgical training in the Department of Neurosurgery at the Vienna Medical University, where he was board certified in 2003 and appointed an associate professor in 2005. His clinical interests are pituitary tumors, endoscopic surgery, and neuronavigation. He is currently running an international yearly course on endoscopic pituitary surgery with the aid of virtual endoscopy



Marie-Thérèse Forster received the MD degree from the Medical University Vienna in 2003 following her thesis about virtual pituitary endoscopy. Two years earlier, she spent about one year in Paris, where she had her first experiences in neuroscience, including a short stay at the famous Necker's Hospital working with Professor A. Kahn. Currently, she is preparing to move to the University Hospital of Geneva.



volume visualization at Tiani Medgraph. His main interests include volume visualization of medical data, portable volume visualization for Web applications, and usability issues of volume rendering in medicine.

Lukas Mroz graduated in 1998 from the Vienna University of Technology (TU Wien), Austria, as Dipl. Ing (MSc) in the field of computer graphics. In 2001, he received the degree of Dr. techn (PhD) from the same university after research projects on real-time volume visualization for medical and mathematical data. After research at the Institute of Computer Graphics and Algorithms, TU Wien, and the VRVis Research Center, he continues work in the area of medical



Visualization Department at VRVis. His main research interest is medical visualization.

Rainer Wegenkittl received the MSc degree in technical mathematics in 1994 and the PhD degree in computer science in 1997. He worked from 1996 to 1997 as a research assistant in the field of visualization of dynamical systems at the Institute of Computer Graphics and Algorithms at the Vienna University of Technology. In 1998, he joined TIANI Medgraph, a PACS vendor, where he now leads the R&D department. Between 2000 and 2003, he also led the Medical



University of Technology. Her current research topics are motivated by real-world applications in the medical field and focus mainly on techniques for computer-aided diagnosis and surgery simulation, including specialized solutions for segmentation and visualization.

Katja Bühler received the master's degree in mathematics from the University of Karlsruhe, Germany. In 2001, she received the PhD degree in computer science from the Institute of Computer Graphics and Algorithms, Vienna University of Technology, for her work on reliable geometry processing. Since 2003, she has been head of the Medical Visualization Department at the VRVis Research Center for Virtual Reality and Visualization and an external lecturer for medical visualization at the Vienna

► **For more information on this or any other computing topic, please visit our Digital Library at www.computer.org/publications/dlib.**