

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/264986939>

Mobile markerless augmented reality and its application in forensic medicine

Article in International Journal of Computer Assisted Radiology and Surgery · August 2014

DOI: 10.1007/s11548-014-1106-9 · Source: PubMed

CITATIONS
61

READS
1,878

13 authors, including:



Thomas Kilgus
German Cancer Research Center

22 PUBLICATIONS 272 CITATIONS

[SEE PROFILE](#)



Eric Heim
Universität Heidelberg

16 PUBLICATIONS 334 CITATIONS

[SEE PROFILE](#)



Sven Haase
Friedrich-Alexander-University of Erlangen-Nürnberg

24 PUBLICATIONS 366 CITATIONS

[SEE PROFILE](#)



Michael Müller
mbits imaging GmbH

34 PUBLICATIONS 855 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Computer-aided forensic cases analysis and presentation [View project](#)



Blocking and modulation of Spreading Depolarisations in a gyrencephalic brain under high-definition multimodal monitoring including novel multispectral optical and photoacoustic tissue analysis [View project](#)

Mobile Markerless Augmented Reality and its Application in Forensic Medicine

Thomas Kilgus · Eric Heim · Sven Haase · Sabine Prüfer · Michael Müller ·
Alexander Seitel · Markus Fangerau · Tamara Wiebe · Justin Iszatt ·
Heinz-Peter Schlemmer · Joachim Hornegger · Kathrin Yen · Lena Maier-Hein

Received: date / Accepted: date

Abstract Purpose During autopsy, forensic pathologists today mostly rely on visible indication, tactile perception and experience to determine the cause of death. Although computed tomography (CT) data is often available for the bodies under examination, this data is rarely used due to the lack of radiological workstations in the pathological suite. The data may prevent the forensic pathologist from damaging evidence by allowing him to associate, for example, external wounds to internal injuries. To facilitate this, we propose a new multi-modal approach for intuitive visualization of forensic data and evaluate its feasibility.

Methods A range camera is mounted on a tablet computer and positioned in a way such that the camera simultaneously captures depth and color information of the body. A server

Thomas Kilgus · Eric Heim · Justin Iszatt · Michael Müller ·
Tamara Wiebe · Lena Maier-Hein
Junior group: Computer-assisted Interventions,
German Cancer Research Center (DKFZ),
Im Neuenheimer Feld 280, 69120 Heidelberg, Germany
E-mail: t.kilgus@dkfz-heidelberg.de

Alexander Seitel performed this work in
Junior group: Computer-assisted interventions, DKFZ
and is now with the Robotics and Control Laboratory,
University of British Columbia,
2332 Main Mall, Vancouver, BC Canada V6T 1Z4

Markus Fangerau
Division of Medical and Biological Informatics, DKFZ

Heinz-Peter Schlemmer
Division of Radiology, DKFZ

Sven Haase · Joachim Hornegger
Department of Computer Science, Pattern Recognition Lab,
Friedrich-Alexander University Erlangen-Nuremberg,
Schlossplatz 4, 91054 Erlangen, Germany

Sabine Prüfer · Kathrin Yen
Institute for Forensic Medicine and Traffic Medicine,
Heidelberg University Hospital,
Voßstraße 2, 69115 Heidelberg, Germany

estimates the camera pose based on surface registration of CT and depth data to allow for augmented reality visualization of the internal anatomy directly on the tablet. Additionally, projection of color information onto the CT surface is implemented.

Results We validated the system in a post-mortem pilot study using fiducials attached to the skin for quantification of a mean target registration error of 4.4 ± 1.3 mm.

Conclusions The system is mobile, markerless, intuitive and real-time capable with sufficient accuracy. It can support the forensic pathologist during autopsy with augmented reality and textured surfaces. Furthermore, the system enables multi-modal documentation for presentation in court. Despite its preliminary prototype status, it has high potential due to its low price and simplicity.

Keywords Mobile augmented reality · Forensic medicine · Range imaging · Kinect · Iterative closest-point algorithm · Mobile application · Surface documentation

1 Introduction

Forensic medicine aims to distinguish between natural and non-natural causes of death in cases of unknown cause and thus contributes significantly to the detection and solving of crimes. The forensic pathologist (or coroner) performs autopsies to find evidence for homicide, accident, suicide or natural cause of death. Another goal of forensic medicine is to document and present findings in an understandable way for judicial persons and for medical laypersons.

During autopsy, experts mostly rely on (superficial) visible indicators, tactile perception and personal experience. Although computed tomography (CT) scans before autopsy are gradually becoming standard¹, CT images are still rarely

¹ E.g. in the institute for Forensic Medicine and Traffic Medicine in Heidelberg, almost for every case a CT scan is acquired before autopsy.

used in forensic expert reports. This is, to a large extent, due to a lack of appropriate visualization methods during autopsy. It is generally accepted, however, that acquiring a CT scan before the autopsy provides valuable information in addition to that yielded by the customary procedure [44] and that the data can assist the forensic pathologist. As CT provides a non-invasive, non-destructive and objective assessment method, it has the potential of providing the basis for a minimal-invasive autopsy, in analogy to the meanwhile established minimal-invasive surgery. However, today autopsy still represents the gold standard in forensic injury assessment. CT data could be used to guide the forensic pathologist, i.e. depict broken bones, internal injuries, cuts, shot wounds, etc. and prevent them from damaging or destroying evidence [36]. Although post-mortem forensic CT imaging is gaining more interest, only few forensic institutes have their own CT scanner; in most cases, CT data is acquired in clinical units and then transferred to the forensic unit. Workstations are not available at most institutes. Even in those who have easy access to a workstation, the expert has to mentally transfer the 2D CT images onto the human body, which is challenging and requires experience. Examples for computer-assisted systems to support forensic pathologists with CT data during the autopsy include [10, 11], where the Kinect is used to navigate through the CT data with gestures. To our knowledge, there is no augmented reality (AR) system in literature for supporting autopsies by efficiently making use of the CT data.

Another major challenge is to present findings in an adequate way for medical laypersons, as the forensic pathologist also serves as expert witness in courts of law and testifies in civil or criminal law cases. Ampanozi et al. [1] show that cost-effective visualization methods that explain the findings in a simple and intuitive manner will be of increasing interest. For example, a combination of color and CT data could be used as evidence and for intuitive illustration in court for legal professionals and laypeople to support their decision. The fusion of color and 3D CT surfaces is called surface documentation and recent work by Ebert et al. [12] shows that it is applied increasingly, but still only in 5.1% of a total of 475 cases over 11 months. Existing methods of surface documentation include marker-based approaches, such as [43] where radiological markers are used to merge photogrammetry and radiological data, or the so called Virtobot [12, 13]. This fully automatic robot system has shown to be able to perform impressive 3D surface scans, including texture, registration to CT data and even biopsies with high accuracy. Its aim is to fully replace a standard autopsy with a minimally invasive, automatic process. However, the system is expensive, bulky and involves two different types of markers. There are two prototypes of the system in Switzerland, containing complex equipment and several standard and custom safety protocols (e.g. light curtains to stop robot

movement) [13]. Consequently, the Virtobot is rather hard to integrate in small or medium forensic institutes and their daily routine. In a parallel study, Urschler et al. [45] recently presented the first approach to enabling surface documentation at low cost. A digital reference manikin model is applied to combine Kinect data with CT and/or magnetic resonance imaging (MRI) data. The result can be either a patient-specific or an anonymous generic model, however, markers are used for registration.

Despite these initial efforts, computer-assisted forensic medicine is still in its infancy. In particular, there is no prior work on intuitive visualization of multi-modal forensic data and no system combines AR with surface documentation at low cost.

To address these issues, we propose a mobile and markerless concept based on the recently emerging range cameras [3]. For the first time, our concept enables both (1) intuitive AR visualization and (2) surface documentation of forensic data. In contrast to other state-of-the-art systems, the concept is non-invasive, cost-efficient and easy to integrate into the clinical workflow. Our concept involves a range camera, mounted on a tablet PC, to acquire depth and color information in real time, and a server for computation. The idea is to visualize CT data in color images directly over the body part of interest. Because it is portable, the device can be freely placed almost anywhere close to the body and is easy to handle. By fusing color images and CT data, external wounds and internal injuries are shown in an intuitive and non-invasive manner (e.g. shot wound and organ lesions or bruises and fractures), thus avoiding destruction of potential evidence [13]. Our concept facilitates online and offline visualization of multi-modal forensic data. The medical examiner can access this data during the forensic examination (without an extra radiological workstation) until significant deformation, such as a deep cut, occurs and rigid registration becomes impossible. At this point another CT scan is required to update the skin surface. The AR images/videos and the 3D textured CT surfaces can be used for documentation and for intuitive presentation in trial, which seems appropriate according to [1]. To facilitate these features, we present a fully functional prototype and evaluate it in a realistic post-mortem setting.

In the following sections we describe the concept and the prototype hardware implementation (section 2), the evaluation experiments and results (section 3) and conclude with a discussion and summary (section 4).

2 Materials and Methods

The following subsections describe the visualization concept, the hardware setup and the first implementation.

Concept

The idea is to provide physicians with direct on-patient visualization of a given (static) clinical 3D data set acquired by CT or MRI and allow him/her to switch the viewpoint in real time to obtain an overview. Our concept is designed to interfere as little as possible with the current clinical workflow and to be as simple and as intuitive as possible. Its preparation consists of five steps:

1. Perform a CT scan
2. Segment the CT skin and (optionally) further relevant objects (organs)
3. Position the camera and perform initial registration
4. Start intuitive visualization on tablet display
5. Optional: Move the camera freely

During the procedure, a range camera mounted on a tablet PC is positioned to view a region of interest of the body (see Fig. 1 for an illustration of the device). The device continuously acquires color and depth data of the scene and the latter is registered to the static CT skin surface to estimate the camera pose. This allows exact augmentation of the live color data with structures from within the scanned body or the mentioned surface documentation. The portable device can be freely moved along the target by the user at any time as illustrated in Fig. 2. This gives a very vivid and intuitive visualization of multi-modal photographic and tomographic forensic data. By moving the device and watching on the screen the user gains a kind of “x-ray vision” for internal structures. Online resource 1 shows a prototype implementation of the concept proposed for clarification. Several rendering modes are possible and navigation through the 3D data is performed directly at the target. All computations are run remotely on a server (powerful laptop or PC) to cope with real-time requirements.

In previous publications [21,27,28] we introduced the general concept of mobile AR using range imaging but had to complete the performance assessment with *simulated* data due to the lack of appropriate hardware (portable device) and software (real-time registration). In this work, we (1) propose a practical (real-time) implementation of the concept using a general-purpose computing on graphics processing units (GPGPU) [23] approach and (2) apply it for intuitive visualization of multi-modal forensic data. Due to the known camera pose, the color image stream can be continuously projected onto a surface generated from a static imaging modality, such as CT or MRI. This provides orientation for the forensic pathologist and allows for surface documentation and presentation in trial.



Fig. 1 Prototype hardware setup. From left to right: target object, range camera (Kinect) mounted on mobile display (iPad 2), and cable connection to server which provides wireless connection back to display. Note that some manufacturers are already integrating range cameras in tablet computers and the next version of the hardware will be more compact and completely wireless. The 3D printing plan of the custom-designed³ bracket (green) is used to mount a camera (e.g. PMD CamCube, Kinect for Xbox One or Kinect for Xbox 360) on the iPad.

Prototype

A first prototype of this markerless visualization concept was implemented within the open-source Medical Imaging Interaction Toolkit (MITK) [34]², including the two module extensions Medical Imaging Interaction Toolkit (MITK) Time-of-Flight module (MITK-ToF) [41] for range imaging support and MITK-Ocl, which is a generic software module for GPGPU programming with the Open Computing Language (OpenCL) standard in MITK. The prototype covers the basic workflow shown in Fig. 3, including data acquisition, preprocessing and automatic depth data segmentation, global and fine registration, AR rendering and transferring the AR video stream to the mobile display via a wireless connection. The following paragraphs describe the individual components of the prototype in detail. The hardware setup, the segmentation of the skin from CT images, and subsequently all continuously performed steps to estimate the camera pose are explained.

Hardware An overview of the hardware prototype is depicted in Fig. 1. The setup consists of:

- A powerful laptop or PC acting as a server to perform the surface-based registration and rendering in real time.
- A wireless router, which is plugged in the server for data transfer to the display.
- A mobile display (Apple iPad 2) to render the data.
- A range camera (Microsoft Kinect for Xbox 360), which is mounted on the display with a self-constructed bracket³ (Fig. 1) for continuously acquiring RGB and depth data of the target.
- A tripod or a holding device (typically used for endoscopes) can be used to fix the device.

CT skin segmentation First, the CT skin surface has to be generated. Note that the skin is the only segmented surface required by our concept, and this step is performed just once.

² <http://www.mitk.org>

³ We provide design and 3D printing plan publicly at <http://www.thingiverse.com/thing:343425>

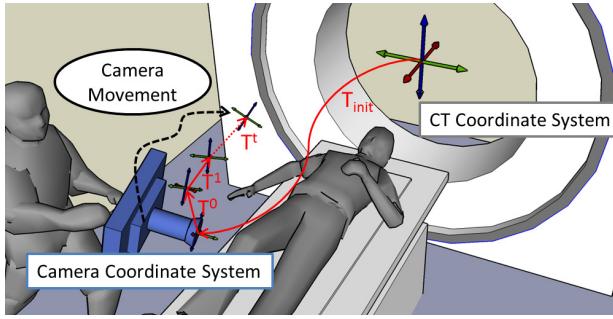


Fig. 2 Concept of mobile augmented reality. A mobile device tracking its pose relative to a target (patient) and rendering internal anatomy from the viewpoint of the device. Continuously updating the virtual camera pose enables a vivid and intuitive visualization of multi-modal (photographic and tomographic) forensic data. T_{init} is the transformation of computed tomography (CT) to camera coordinates and T^0, \dots, T^t are the transformations of the device poses. Better viewed in color.

The CT skin surface is essential for performing registration of depth data (from a range camera) and is generated via a simple 3D upper/lower threshold followed by a closing operation to fill any gaps in the threshold segmentation and surface generation with the marching cubes algorithm [26]. The CT skin surface is remeshed using the algorithm presented by Valette et al. [46], which creates meshes of isotropic and/or anisotropic vertices based on discrete Delaunay criteria. To enable real-time registration at high quality, it is crucial to reduce the number of vertices from $\approx 1 - 5$ millions to $\approx 20k$ while preserving surface features and a homogeneous distribution of vertices.

Data acquisition Before data acquisition, intrinsic calibration is performed once per camera to obtain the intrinsic parameters essential for surface generation and back projection of 3D internal structures to the color camera coordinate system. We use the intrinsic calibration process as described in [30], which is based on the work by Zhang et al. [48]. For stereo calibration (i.e. alignment of color and depth data), default implementations are provided in the Microsoft Kinect for Windows SDK³ and OpenNI⁴. This allows, the depth data to be transformed directly into the color camera coordinate system and the device continuously captures color and depth data in a uniform coordinate system.

Preprocessing and automatic depth data segmentation Let I denote a depth image with dimension $N \times M$ and p_{ij} be a pixel value at index $(i, j) \in I$. In our initial prototype implementation, a simple distance threshold $th_{distance}(p_{ij})$ cuts off pixel values p_{ij} of background $p_{ij} < d_{far}$ and close objects $d_{near} < p_{ij}$ to facilitate the segmentation of the Kinect data. d_{near} and d_{far} are chosen depending on the application.

³ <http://www.microsoft.com/en-us/kinectforwindowsdev/>

⁴ <http://www.openni.org>

Next, the aim is to generate two surfaces - source A and target B as registration input. B is defined as the skin mesh generated via segmentation from CT data and A is produced as follows: The Kinect depth data is converted to 3D world coordinates according to the pinhole camera model and based on [24]. A straightforward triangulation [22] is enhanced with the following simple procedures to allow a fully automatic real-time segmentation of depth data: Let v_{ij} denote the 3D coordinates (relative to the camera coordinates) of the pixel at index $(i, j) \in I$. For each pair of adjacent vertices v_{ij} and v_{kl} within an $n = 1$ neighborhood, the Euclidean distance $d(v_{ij}, v_{kl})$ is calculated and the pair is neglected for triangulation if a certain distance $d_{triangulation}$ is exceeded: $d(v_{ij}, v_{kl}) > d_{triangulation}$. Similar to a region growing, this process is repeated for all v_{kl} accepted for triangulation. As a result, objects are separated at their boundaries. Next, the source A is defined as the partial surface originating from the camera center ($i = N/2, j = M/2$) - i.e. all vertices connected via triangle strips belong to A . To position A in the camera center is a reasonable and valid constraint for the user, as one usually focuses on the region of interest.

Real-time registration Computation of the transformation from source A to target B is the last stage for camera pose estimation. Let $T_{final}^t = \{R, t\}$ denote the rigid transformation (with rotation matrix R and translation vector t) of the virtual camera of the CT scene to the color camera coordinate system of the device with a captured frame at time t .

$$T_{final}^t = (T^t \times \dots \times T^2 \times T^1 \times T^0 \times T_{init})^{-1} \quad (1)$$

(1) Global initialization: First, we need a rough initial alignment T_{init} to position A closely to B (cf. Fig. 2). This part is called *global registration*. It can be performed by placing B at a certain virtual distance from the real device and in the camera center, followed by positioning the actual camera, as proposed by Glocker et al. [14], or based on selecting at least three non-collinear points on both surfaces. Note that T_{final}^t describes the transformation of the virtual camera space to the color camera coordinate system and that both global and fine registration compute the respective transformation from the color camera coordinate system to the virtual camera of the CT coordinates. Hence we use the inverse transformation to obtain T_{final}^t .

(2) Continuous fine registration: Next, A and B have to be fine registered, in order to obtain essential accuracy. Each change of the camera device pose can be described with a rigid transformation $T^t = \{R_t, t_t\}$ and each transformation is computed with the well-established iterative closest-point (ICP) algorithm [4]. For each registration process $t > 0$, A is transformed with T^{t-1} and, together with B , serves as input of the ICP. To achieve a reasonable frame rate, we

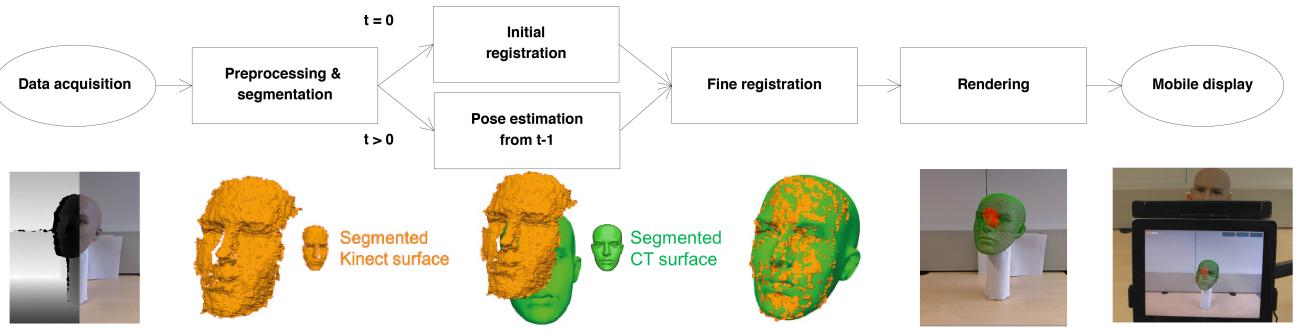


Fig. 3 Workflow of the markerless concept for real-time augmented reality on mobile devices. From left to right: Simultaneous acquisition of depth and color data for time step t . After preprocessing and depth data segmentation, the first time step $t = 0$ requires an initial alignment with the computed tomography (CT) surface. Further time steps $t > 0$ are initialized with the result of the fine registration of the previous time step $t - 1$. The color image is enriched with internal structures of the CT data (here: red tumor, green CT surface), rendered on a server to provide the required real-time processing capacity and transferred to the mobile display via Wi-Fi. Better viewed in color.

apply a GPGPU implementation of the ICP and (optionally) its trimmed variant [7]. The ICP algorithm has become one of the best-known methods for geometric alignment of 3D models. The original algorithm [4] is an iterative procedure that aims to find the rigid transformation τ that minimizes the inter-point distance after registration for two point sets $X = \{\mathbf{x}_1, \dots, \mathbf{x}_{N_x}\}$ and $Y = \{\mathbf{y}_1, \dots, \mathbf{y}_{N_y}\}$. Each iteration consists of two steps: (1) For each $\mathbf{x}_i \in X$ establish a correspondence $\mathbf{z}_i = C(\mathbf{x}_i, Y) \in Y$ based on the Euclidean distance $C(\mathbf{x}_i, Y) = \arg \min_{\mathbf{y}_i \in Y} \|\mathbf{x}_i - \mathbf{y}_i\|$ (2) compute the rigid transformation $\tau_k = [R_k, \mathbf{t}_k]$ that minimizes the Euclidean distance between corresponding points using, for instance, the method of Horn [17] for each iteration k . Steps (1) and (2) are repeated until a convergence criterion, such as a minimal improvement of the fiducial registration error (FRE) for the current iteration, is met. The FRE is defined as root-mean-square distance between all correspondences:

$$FRE^2 = \sum_{i=0}^{N_x} \|(R_k \mathbf{x}_i + \mathbf{t}_k - \mathbf{z}_i)\|_2^2 \quad (2)$$

The trimmed variant of the ICP sorts all correspondences regarding distance and cuts off the worst percentage of matches (i.e. $\#vertices = \#vertices \times partialoverlap$) before computation of the transformation τ_k .

In contrast to previous GPGPU implementations, like the work of Qui et al. [35], which focuses mainly on the parallelization of the correspondence search, we aim for as much parallelization of the steps on the graphics processing unit (GPU) as possible in order to minimize the expensive data transfer between central processing unit (CPU) and GPU [23]. Our approach is designed for NVIDIA hardware to maximize the gain in runtime, but uses the generic architecture of MITK-Ocl and the OpenCL standard to be compatible with other hardware. First, there is a one-time initialization of the algorithm allocating GPU memory once and copying source A and target B to the GPU memory.

For correspondence search of vertices from A and B , we apply a k-dimensional tree implementation on the GPU. The method of Horn [17], requiring center of gravity computation and covariance matrix computation (see [4, 17]), is likewise performed in a massively parallel fashion, applying the concept of parallel reduction [23] wherever possible. The resulting covariance matrix is copied to the CPU to compute the respective transformation τ_k of iteration k of the ICP, because this computation is small and can hardly be executed in parallel. Additionally, the CPU can perform the computation with double accuracy. The transformation matrix τ_k is copied back to the GPU where one thread is started for transforming each vertex of the input A with the matrix. The FRE is then computed with the concept of parallel reduction before being copied back to the CPU to evaluate the convergence criterion. A neat advantage is that, for further time steps $t > 0$ the algorithm is already initialized and only the source A transformed with T^{t-1} has to be copied to GPU memory. The previous registration result is used for initial alignment, which prevents error propagation. Once an ICP convergence criterion is met, the resulting transformation τ is the searched T^t . For the optional trimmed variant of the ICP, we apply a GPGPU implementation of the Bitonic Sort [2] algorithm to sort the correspondences of the ICP algorithm. Bitonic Sort has been chosen because this sort algorithm is convenient for parallelization on the GPU.

For reliable results, we save the FRE e_{valid} of the time step $t = 0$ and check if the FREs $e(t)$ of subsequent time steps $t > 0$ exceed the constraint $e(t) < \alpha e_{valid}$. If they do, the current frame t (T_t) is skipped and T_{final}^t is not updated. α can be chosen depending on the application (we chose $\alpha = 1.5$ and observed stable results in our experiments). If no registration is successful (i.e. $e(t) < \alpha e_{valid}$ is fulfilled), the tracking is assumed to be lost and a new initial registration is required. This allows the user to move the camera almost anywhere and still be able to return to the tracking process once the constraint $e(t) < \alpha e_{valid}$ is met again.

For the final AR visualization, the inverse transformation is computed and T_{final}^t applied to the virtual camera of the CT scene (as shown in equation 1). As mentioned before, the range data is already aligned with the color data of the camera by stereo calibration and thus the color data is now also merged with the CT data.

Visualization To assist the physician, the following visualization techniques are implemented and are performed on the server to cope with real-time requirements:

Mobile augmented reality: The calculated transformation T_{final}^t is used to augment the color video with internal structures from the 3D CT data. To achieve this, the 3D CT data is projected into the color camera coordinate system. The internal CT data can be rendered with several standard visualization techniques, such as semi-transparent surfaces or volume rendering which are provided by MITK.

Surface documentation: For texturing, the color image (which is aligned with the CT data through the registration) is projected onto the CT surface. Let $\mathbf{v}_i = (x_i \ y_i \ z_i)^T$ denote a vertex i from the CT surface in world coordinates, the color camera pose is the origin $\mathbf{o} = (0 \ 0 \ 0)^T$, the up vector $\mathbf{up} = (0 \ 1 \ 0)^T$ and view direction $\mathbf{view} = (0 \ 0 \ 1)^T$. The texture coordinates u and v are assigned to \mathbf{v}_i derived from the generic pinhole camera model [16] as follows:

$$\begin{aligned} u &= \frac{x_i}{z_i} \cdot \frac{f_x}{N} + dx \\ v &= \frac{y_i}{z_i} \cdot \frac{f_y}{M} + dy \end{aligned} \quad (3)$$

where f_x, f_y is the focal length of the color camera achieved from intrinsic calibration, N, M are the dimensions of the color image, and dx, dy is the offset of the principal point $\mathbf{p} = (p_x \ p_y)^T$ (again obtained from intrinsic calibration) caused by small intolerances during camera manufacturing:

$$\begin{aligned} dx &= \frac{N/2 - p_x}{N} \\ dy &= \frac{M/2 - p_y}{M} \end{aligned} \quad (4)$$

Note that the factor $f_x/N, f_y/M$ scales the range of the texture coordinates to $[dx, 1+dx]$ for the horizontal viewing direction of the camera and $[dy, 1+dy]$ for the vertical viewing direction of the camera, which is a common procedure to facilitate a mapping of $[0, 1]$ (in the ideal case $dx = 0, dy = 0$).

If equation 3 is applied to all vertices of the entire CT skin surface, this would result in a completely colored CT surface, and therefore a partly incorrectly textured surface. This becomes relevant once the user is interested in a virtual view of the data and modifies the virtual camera. In other words, the virtual camera pose differs from the color camera pose. To overcome this issue, we add a shader (a program that calculates rendering effects on graphics hardware), setting faces invisible that are front-facing the virtual camera but are not visible to the color camera. Back-facing vertices are rendered black.



Fig. 4 Visualization techniques: Color image of a torso phantom and computed tomography surface are combined to a textured 3D surface and a shader cuts a hole inside the surface (middle). Background is colored black and sites which are not visible for the camera are transparent to view synthetic internal structures (right). Better viewed in color.

As proposed in [28], we provide another shader to cut a hole in the camera center to obtain an “x-ray vision” like effect. The hole size can be adapted. The textured skin provides orientation and improves depth perception. See Fig. 4 for an illustration of the texture projection, the background shader and the hole shader.

The final image is rendered on the server and streamed to the mobile display via Wi-Fi using the VP8 video compression format [39].

3 Experiments and Results

To demonstrate the feasibility of our concept, we conducted several studies with phantoms and a real corpse. We decided not to focus accuracy evaluation on a specific body part, since the concept is meant to work with the complete body. To keep experiments comparable, the following conditions were applied for all studies:

- **Hardware setup:** As server we used a medium-high end PC with an NVIDIA Geforce GTX 560, Intel® Core™ i7-3930K CPU @ 3.20GHz × 12 and 34 GB RAM memory connected to an ASUS RT-N56U router for wireless data transfer. As described in section 2 under hardware, a Microsoft Kinect for Xbox 360 mounted on an Apple iPad 2, which served as mobile display, was the device prototype. To provide a foreground-background contrast in the depth image, the target was placed on flat wooden plates, allowing a distinction between the target and the CT table.
- **CT segmentation:** The complete CT skin segmentation and preparation chain took about 5 minutes. The parameters for the algorithm by Valette et al. [46] were set to sub-sampling = 300 and gradation = 0.0. To render

inner organs and target regions as semi-transparent surfaces, a CT segmentation is required. All CT segmentations for our experiments were generated using MITK [34] and semi-automatic segmentation tools, which took roughly 10 minutes for liver, kidneys, bones and lungs in the post-mortem study.

- Software settings: The distance range for the threshold filter was set to $d_{near} = 0$ m and $d_{far} = 1$ m, while the threshold for triangulation $d_{triangulation}$ was set to 1 cm. The initial registration was achieved by manually selecting four anatomical landmarks (such as nose, chin, fingers, breast or shoulder - depending on their visibility) on target *B* and source *A*, which took roughly 10 seconds. Convergence threshold and maximum number of iterations for the standard ICP and the trimmed ICP were set to 10^{-5} and 200 respectively for all experiments. A partial overlap of 90% was used for the trimmed ICP.
- Evaluation: In order to evaluate FRE and target registration error (TRE) [9], radio dense markers were attached to the targets. The markers are visible in both CT images and Kinect color images (not in range data) and are manually tagged in both modalities. To rigidly register the markers, the method of Horn [17] was applied. Similar to the FRE of the ICP (cf. equation 2), we computed the FRE between the registered markers of the Kinect and the CT. As the method of Horn computes the rigid transformation that minimizes the FRE, this measure can be regarded as ground truth. To estimate the accuracy of the system, the TRE is obtained by applying the computed transformation T_{final}^t (cf. section 2) to the Kinect markers - our artificial targets - and computing the root-mean-square Euclidean distance between corresponding markers based on the computed registration. The markers are not used for computing the transformation T_{final}^t , and hence, the TRE is a suitable and independent measure for accuracy. We evaluate the runtime of the standard and trimmed ICP, since this is the main performance bottleneck and was especially optimized regarding runtime.

Phantom Study We decided to investigate the accuracy and runtime in a well-defined environment with a torso phantom⁵ in order to avoid unpredictable influences, such as movement of body parts or soft tissue deformation. Additionally, this experiment served as preparation for the post-mortem study. Only for evaluation purposes, we attached ten radio dense markers to the phantom. To estimate accuracy, we decided on five sensible and challenging poses (varying view angle, camera distance, reflections and visible parts) with the target in the camera center as only constraint. The five poses are depicted in Fig. 5. Example CT and Kinect surfaces are shown in Fig. 6 (a) and (b). Because, depending on

⁵ Bronchoscopy Model Sick Boy <http://www.gtsimulators.com>

Table 1 The target registration error (TRE) in millimeters and runtime in milliseconds for five targets on torso phantom data achieved by the standard iterative closest-point (ICP) algorithm and the trimmed ICP (trICP). Kinect poses registered to computed tomography, fiducial registration error (FRE) for point-based registration is computed by the method of Horn [17].

Result for pose	FRE (mm)	TRE ICP (mm)	Runtime ICP (ms)	TRE trICP (mm)	Runtime trICP (ms)
1	1.3	5.3	177.5	5.7	543.6
2	2.5	3.9	232.1	4.4	637.3
3	1.7	3.7	173.5	3.7	534.0
4	0.9	4.1	297.9	5.6	810.5
5	1.9	4.2	215.8	4.1	608.5
Mean	1.7	4.2	219.3	4.7	626.8
Std. Dev.	0.6	0.6	50.5	0.9	111.5

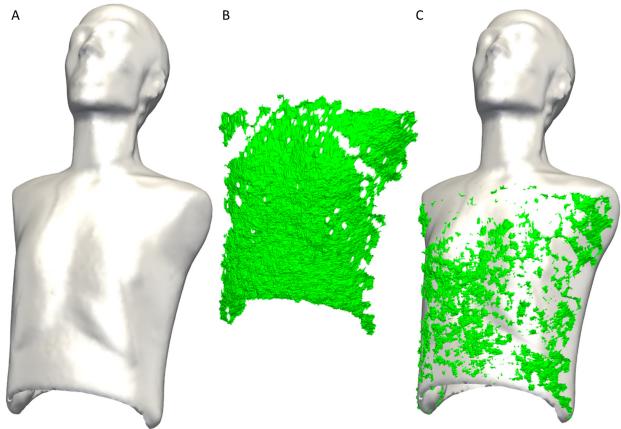


Fig. 6 Computed tomography surface of a torso phantom (a), corresponding Kinect data (b) and result of fine registration (c). Better viewed in color.

the viewpoint, not all markers are visible in the Kinect color image, we used the five markers closest to camera center.

Phantom Results Results for the experiments conducted with the phantom for five camera poses are shown in Table 1. The mean FRE of the torso phantom is 1.7 ± 0.6 mm. The standard ICP achieves a mean TRE of 4.2 ± 0.6 mm for the phantom data with a mean runtime of 219.3 ± 50.5 ms. The trimmed ICP achieves a mean TRE of 4.7 ± 0.9 mm for the phantom data with a mean runtime of 626.8 ± 111.5 ms. A sample data set is shown in Fig. 6. Fig. 7 demonstrates the quality of the texture projection on another head phantom and Fig. 4 shows the shader concept. Online resource 2 shows the phantom data with augmented volume rendering.

Post-mortem Study To show the feasibility of the system in a realistic setup, the Institute for Anatomy and Cell Biology in Heidelberg provided a participant in the willed body program of the University of Heidelberg, which includes post-mortem studies. The body was previously fixed with ethanol



Fig. 5 Phantom data camera poses for evaluation. From left to right: Poses 1 - 5.

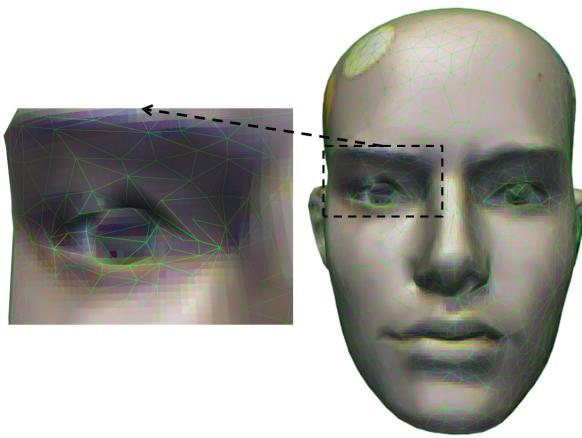


Fig. 7 Microsoft Kinect color image projected on a computed tomography surface of a head phantom with overlaid mesh in green. Note the accuracy around the eye. Better viewed in color.

and, to maintain anonymity, the face was covered. The post-mortem experimental setup is pictured in Fig. 8. Note that the hands of the person are tied on the abdomen to fit the CT scanner, which is a common procedure in forensics. For evaluation purposes only, 22 radiological markers were attached to the body and five camera poses were chosen for accuracy estimation, as for the phantom (varying viewing angle, camera distance, reflections and visible body parts). The five poses are depicted in Fig. 9. In contrast to the phantom data, the post-mortem range data contains parts which are not visible in the CT data. We therefore evaluated only the trimmed ICP, which is made for partially overlapping surfaces. To improve the runtime of the registration, we conducted a series of offline experiments where the range data of the post-mortem study was down-sampled by factors up to 4.

Post-mortem Results The results of the accuracy estimation are presented in Table 2. The mean FRE is 3.0 ± 0.9 mm for the post-mortem data. The system achieves a mean TRE of 4.4 ± 1.3 mm with a mean runtime of 1428.9 ± 334.6 ms for the trimmed GPGPU ICP algorithm. The runtime experiments are presented in Table 3. The accuracy remains stable up to a down-sampling of 2.5 and until runtime is decreased by a factor of 3.9. A down-sampling factor of 4 decreases the runtime by a factor of 7.5 and increases the TRE by about 2.5 mm. An example of the fine registration is depicted in

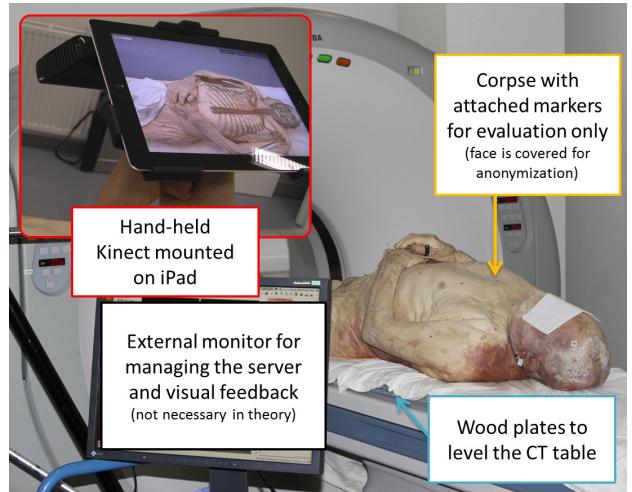


Fig. 8 Post-mortem experimental setup with computed tomography (CT) scanner in background. Red framed photo shows the hand-held device with augmented reality. Better viewed in color.

Table 2 The target registration error (TRE) in millimeters and runtime in milliseconds for five targets on post-mortem data achieved by the trimmed iterative closest-point algorithm. Kinect poses registered to computed tomography. fiducial registration error (FRE) for point-based registration is computed by the method of Horn [17].

Result for pose	FRE (mm)	TRE (mm)	Runtime (ms)
1	1.6	2.4	1808.6
2	2.7	4.2	1566.5
3	3.5	4.2	1379.4
4	3.8	5.5	1489.0
5	3.6	5.6	901.1
Mean \pm Std. Dev.	3.0 ± 0.9	4.4 ± 1.3	1428.9 ± 334.6

Fig. 10. Qualitative results for the augmented reality are depicted in Fig. 11 and for the texture projection in Fig. 12. Online resource 3 shows post-mortem data with augmented semi-transparent surfaces without and with down-sampling enabled.

4 Discussion

We propose the first mobile solution for intuitive visualization of multi-modal forensic data, describe a full workflow of a markerless concept in detail and apply a prototype to a typical post-mortem forensic setting. A range camera in



Fig. 9 Post-mortem data camera poses for evaluation. From left to right: Poses 1 - 5. Face is covered for anonymization.

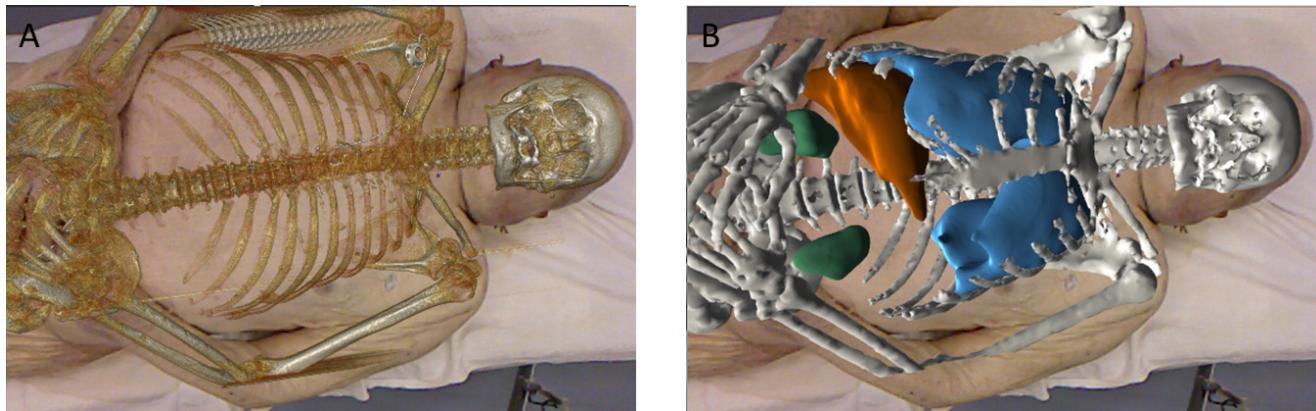


Fig. 11 Examples of augmented reality on a human corpse. (a) Volume rendering and (b) organ surfaces generated from computed tomography. Better viewed in color.

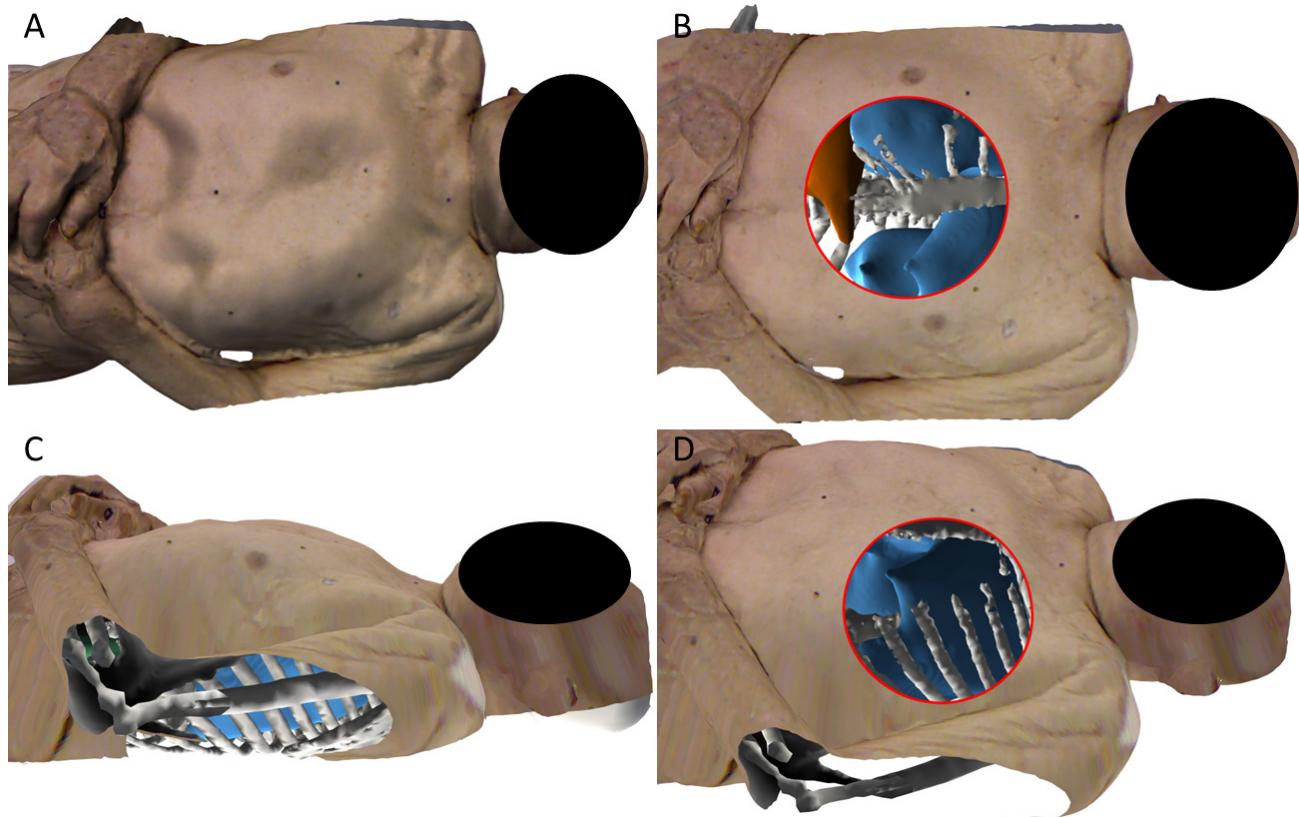


Fig. 12 Kinect color image projected onto a 3D surface generated from computed tomography (CT) of a human corpse. (a) Texture projection only, (b) hole shader allows visualization of internal structures, (c) and (d) side views. Face is covered for anonymization. Better viewed in color.

Table 3 Mean target registration error (TRE), standard deviation of TRE, mean runtime and standard deviation of runtime for different down-sampling factors of the depth images of the corpse data set (five poses). All results are given for the trimmed iterative closest-point algorithm.

Down-sampling	TRE (mm)	Runtime (ms)
1 (none)	4.4 ± 1.3	1428.9 ± 334.6
1.5	4.3 ± 1.2	836.7 ± 247.3
2	4.3 ± 1.2	543.8 ± 201.1
2.5	4.2 ± 1.1	371.1 ± 106.0
3	4.7 ± 1.3	292.5 ± 107.2
3.5	4.9 ± 1.9	241.8 ± 74.7
4	6.9 ± 4.2	194.6 ± 82.4

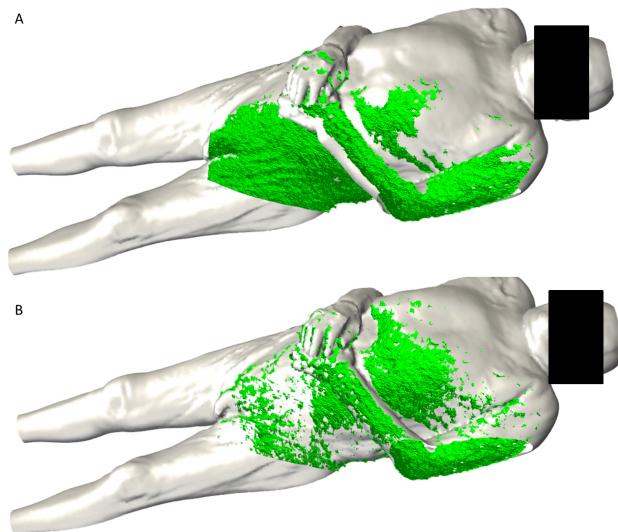


Fig. 10 Example of human body computed tomography (CT) surface and corresponding partial Kinect data (green). Rough initial alignment of the CT surface (a) and fine registration result (b). Face is covered for anonymization. Better viewed in color.

combination with a tablet computer allows real-time visualization of both inner anatomical details and realistic textures on the surface acquired by conventional imaging modalities, such as CT or MRI. We implemented the first prototype and successfully applied it in both phantom and post-mortem experiments. The mean TRE for the phantom and post-mortem experiments is about 4 mm and the registration is performed with a mean runtime of 219.3 ms for the standard GPGPU ICP and 371.1 ms for the trimmed GPGPU ICP (with down-sampled depth images of factor 2.5).

For discussion of the accuracy, let us review the mean FRE, which is 1.7 ± 0.6 mm on phantom data and 3.0 ± 0.9 mm on post-mortem data. For the phantom data, the value seems reasonable considering the accuracy of the Kinect camera [20], calibration error, noise, CT error and user error in tagging markers. We attribute the increase in the FRE on the post-mortem data partly to the fact that the phantom is much smaller than the corpse. Hence, the markers are spread in a far bigger region on the post-mortem data, which in-

creases the FRE. Also, slight surface deformations affecting the markers cannot be fully excluded. In the phantom data, the standard GPGPU ICP provides acceptable runtime and accuracy and is significantly faster than trimmed ICP (see Table 1). On the post-mortem data, the Kinect covers regions which are not present on the CT data and vice versa. For instance, in Fig. 10 the left arm is only captured by the Kinect and the CT surface contains legs and head. At first sight one could argue that the hands are “good” features for the ICP, but in fact they are very challenging. The hands are represented as real 3D objects in the CT data while the Kinect captures only the visible surface (i.e. 2D + t data). Consequently, the Kinect surface (abdomen) may be registered to the (for the Kinect camera) invisible back of the hands, which leads to an inaccurate registration. Consequently, the trimmed ICP is necessary in the post-mortem data to avoid false local minima and achieve acceptable accuracy (see Table 2). As one would expect, the trimmed ICP results in a significantly increased runtime compared to the ICP (cf. runtime in Table 1 and Table 2). It appears to be possible, however, to down-sample the depth data to decrease the runtime by a significant factor (see Table 3).

In the current prototype we optimized the registration runtime as this is one of the main frame-rate bottlenecks of our application, but did not focus on the application runtime. Therefore, the empirically determined frame rate varies depending on the point cloud size between 5 and 12 frames per second (FPS), which seems acceptable for a non-commercial system. The term real time is vague and discussed controversially in literature [8, 15]. An acceptable frame rate depends to a large extend on the application. For a rigid setup (e.g. a fixed device) the required frame rate is lower than for a moving setup (e.g. a hand-held device). In practice, the device is likely to be fixed most of the time. The examiner moves to a region of interest and starts documenting with a fixed device where delay is negligible.

Further offline experiments were made with a generalization of the ICP, accounting for anisotropic uncertainties - the so called anisotropic ICP [29] (A-ICP) - and its respective trimmed version [21]. The A-ICP is intended for surfaces affected by intense noise and copes with different resolutions. As our data is well-tempered in this regard, the A-ICP does not improve accuracy significantly. Due to the much better runtime of the standard ICP, we have thus excluded these results. We emphasize, however, that the ICP is mainly chosen for this first prototype due to real-time requirements, and alternatives for fine registration considering surface normals or texture will be investigated in the future.

Several rendering techniques have been implemented and the visual results are promising. Fig. 11 and 12 show the qualitative results. With recent scene capturing or stitching algorithms, such as Kinect Fusion [18], it will be possible to obtain a fully textured body. It is, however, already possible

to create a video of the whole visible body and repeat the procedure for the back with the current prototype.

An adequate and understandable presentation for laypersons is becoming more important, as shown by multiple sources [1, 12, 45]. Our system enables new ways of creating AR videos and textured 3D CT surfaces and rendering these on a tablet immediately before the autopsy. We plan to enhance the visualization using touch interaction on the tablet. Intuitive zooming and panning in a textured 3D scene would be possible. Even 3D printing for presentation in court is being investigated and could include texture. Further investigation will show which method of presentation is preferable.

Is the tablet necessary for the system? Technically it is not, however, the intuitive and direct feedback of our concept is lost without the mobile display. If the rendering is performed by the device itself, the system works right out of the box, in contrast to a camera with an external monitor. We think that this “feeling” is especially important for a seamless integration in medical routines. Recent studies also reveal that tablet computers, such as the iPad, are well-suited for displaying radiological images and can be an effective alternative to conventional workstations [5, 19]. We further expect integrated depth sensors in tablet computers in the near future. With a very powerful tablet it will, in future, be possible to perform the computation on the device, making the server redundant.

The system is further easy to integrate in clinical routine due to low preparation time, which depends mainly on the CT segmentation of the skin and other potentially variable objects. Note that the CT skin segmentation is the only necessary preparation step that allows for AR volume rendering and textured 3D CT surfaces with gray-scale slices inside similar to [43]. Additional segmentations, such as organs or bones, can be performed depending on the case and after the autopsy for documentation and presentation. The CT skin segmentation is not fully automatic; an automatic segmentation for skin and other organs is desirable, but not topic of this work.

Although, to minimize deformations, the body was not moved off the CT table in the post-mortem experiment, we are confident that this issue can be solved in each case. If rigor mortis sets in or only parts of the body (e.g. head or leg) are found, transport would be possible. In other cases we plan to use a vacuum mattress or develop a special plastic tub to prevent movement of extremities etc. to allow moving the body off the CT table. Furthermore, we are planning to compensate for isometric transformations rather than only rigid transformations in the future.

In contrast to other state-of-the-art AR systems, the system is markerless, does not involve bulky equipment (such as head-/eye-tracking, optical tracking, projectors or robotics, e.g. [33, 42]), is cost-effective, non-invasive, portable and provides direct and intuitive visualization. Although the cur-

rent prototype requires a cable connection and has a moderate frame rate, the results are promising when combined with a holding device. Usage of the system was possible with two different CT scanners without major issues, underlining the system’s mobility. Note that the system can be used with mobile/small CT scanners, such as a DynaCT or C-arm when there is no stationary CT scanner. In contrast to marker-based systems, the system is less sensitive to the line-of-sight problem due to its ability to capture a complete depth scene. The system copes with interfering objects (e.g. instruments or hands) by skipping frames with high ICP FRE values (cf. section 2).

The aim of our concept is to support the forensic pathologist with intuitive visualization and to ease the mental transfer of static 3D data on the body, in contrast to the Virtobot [12, 13], which is designed for a fully automatic autopsy. The prototype cannot compete with the Virtobot in terms of functionalities and accuracy, but is a cost-effective alternative for small and medium-sized forensic institutes.

Future work regarding technical and clinical aspects includes five major issues:

1. Integration of a miniaturized depth sensor, such as the structure sensor⁶, as the current device is relatively heavy and not fully mobile (a cable is still necessary for data transfer and power supply of the Kinect). Replacing the Kinect with the novel structure sensor - a smaller, light-weight version of a structured light camera specially designed for mobile devices - would make the system fully mobile and significantly decrease weight and improve handling. Range imaging is still in its infancy and new devices with interesting features are being released constantly, e.g. the Kinect for Xbox one⁷ with HD color camera.
2. For our prototype system, we chose a manual initial registration, although a fully automatic method should be used in future implementations. Losing the camera pose tracking can be tedious and repetition of manual initial registration is not optimal. Note that this does not apply if the camera and the target are fixed. In literature, there are many promising approaches for automatic initial registration, such as the sample consensus initial alignment algorithm [37], the global optimal ICP [47], and approaches based on surface features, like the one proposed by dos Santos et al. [38]. Automation of the initial registration is definitely preferable, although in practice we observed that such algorithms are often hard to parametrize and suffer from (almost) unacceptable run-times in the order of minutes, which shows the need for further investigation.

⁶ <http://structure.io>

⁷ <http://www.xbox.com/en-US/xbox-one/get-the-facts?xr=shellnav>

3. The preliminary depth data segmentation process was sufficient for all our experiments. However, wood plates had to be used to level the CT table. A regular convex shape would not allow an automatic depth data segmentation which is crucial for the real-time registration. Combining color and depth data seems worthy of further investigation. An air mattress with a very distinctive color could possibly solve this issue, if there are no other materials (e.g. cloths/bandages) involved; in addition, the mattress would ensure fixation of the body.
4. An ethics approval is currently being conducted to evaluate the system on real forensic cases, especially ones with visible injuries; the potential of our approach when used in court trials will be examined. We want to emphasize that the system has potential to be applied to an even broader context of clinical applications, such as: Teaching and learning of anatomy [6,49], intuitive visualization for physicians and patients during ward round, close-to-patient visualization directly before intervention (i.e. actualization of procedure planning and prevention adverse events). In combination with trajectory planning [40] it could be applied for punctures and biopsy [31]. New sensors and mobile devices will help improve our prototype, increase accuracy further and allow for rigid applications like percutaneous nephrolithotomy [31].
5. The main drawback for applications in an even broader field is the rigid registration. In contrast to marker-based approaches, our concept already continuously acquires complete *deforming* surfaces in real time via a depth camera. Further investigations should focus on integration of non-rigid surface registration approaches, such as [25,32], and their real-time application.

To conclude, the proposed markerless and mobile approach has proven to be helpful for intuitive visualization in forensic procedures and enables multi-modal surface documentation which can be used in court. Due to its low costs and generic design, the approach has a good potential for being used in forensics and other clinical applications.

Acknowledgements This study was performed as part of the Research Training Group 1126 *Intelligente Chirurgie* and project PD 15577, both funded by the German Research Foundation (DFG). We further thank all staff - the Division of Radiology of the German Cancer Research Center and the Institute for Anatomy and Cell Biology in Heidelberg - who supported the phantom and post-mortem experiments. Thanks go to our intern, Simon Weiser, for assisting in post-mortem experiments.

Conflict of interest None.

References

1. Ampanozi, G., Zimmermann, D., Hatch, G.M., Ruder, T.D., Ross, S., Flach, P.M., Thali, M.J., Ebert, L.C.: Format preferences of district attorneys for post-mortem medical imaging reports: Understandability, cost effectiveness, and suitability for the courtroom: A questionnaire based study. *Legal Medicine* **14**(3), 116–120 (2012)
2. Batcher, K.E.: Sorting networks and their applications. In: Proceedings of the April 30–May 2, Spring Joint Computer Conference, pp. 307–314 (1968)
3. Bauer, S., Seitel, A., Hofmann, H., Blum, T., Wasza, J., Balda, M., Meinzer, H.P., Navab, N., Hornegger, J., Maier-Hein, L.: Real-time range imaging in health care: A survey. In: Time-of-Flight and Depth Imaging. Sensors, Algorithms, and Applications, pp. 228–254. Springer (2013)
4. Besl, P.J., Neil D., M.: A method for registration of 3-d shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **14**(2), 239–256 (1992)
5. Bhatia, A., Patel, S., Pantol, G., Wu, Y.Y., Plitnikas, M., Hancock, C.: Intra and inter-observer reliability of mobile tablet PACS viewer system vs. standard PACS viewing station-diagnosis of acute central nervous system events. *Open Journal of Radiology* **03**(02), 91–98 (2013)
6. Blum, T., Kleeberger, V., Bichlmeier, C., Navab, N.: mirracle: An augmented reality magic mirror system for anatomy education. In: Virtual Reality Short Papers and Posters (VRW), 2012 IEEE, pp. 115–116. IEEE (2012)
7. Chetverikov, D., Svirko, D., Stepanov, D., Krsek, P.: The trimmed iterative closest point algorithm. In: Pattern Recognition. Proceedings. 16th International Conference on, vol. 3, pp. 545–548. IEEE (2002)
8. Dabrowski, J.R., Munson, E.V.: Is 100 milliseconds too fast? In: CHI EA '01 CHI '01 Extended Abstracts on Human Factors in Computing Systems, pp. 317–318. ACM Press (2001)
9. Danilchenko, A., Fitzpatrick, J.M.: General approach to first-order error prediction in rigid point registration. *IEEE Transactions on Medical Imaging* **30**(3), 679–693 (2011)
10. Ebert, L., Hatch, G., Thali, M., Ross, S.: Invisible touch—control of a DICOM viewer with finger gestures using the kinect depth camera. *Journal of Forensic Radiology and Imaging* **1**(1), 10–14 (2013)
11. Ebert, L.C., Hatch, G., Ampanozi, G., Thali, M.J., Ross, S.: You can't touch this: Touch-free navigation through radiological images. *Surgical Innovation* **19**(3), 301–307 (2012)
12. Ebert, L.C., Ptacek, W., Breitbeck, R., Fürst, M., Kronreif, G., Martinez, R.M., Thali, M., Flach, P.M.: Virtobot 2.0: the future of automated surface documentation and CT-guided needle placement in forensic medicine. *Forensic Science, Medicine, and Pathology* **10**(2), 179–186 (2014)
13. Ebert, L.C., Ptacek, W., Naether, S., Fürst, M., Ross, S., Buck, U., Weber, S., Thali, M.: Virtobot - a multi-functional robotic system for 3D surface scanning and automatic post mortem biopsy. *The International Journal of Medical Robotics and Computer Assisted Surgery* **6**, 18–27 (2009)
14. Glocker, B., Izadi, S., Shotton, J., Criminisi, A.: Real-time RGB-D camera relocalization. In: Mixed and Augmented Reality (ISMAR), IEEE International Symposium on, p. 173–179. IEEE (2013)
15. Handa, A., Newcombe, R.A., Angeli, A., Davison, A.J.: Real-time camera tracking: when is high frame-rate best? In: Computer Vision–ECCV 2012, p. 222–235. Springer (2012)
16. Hartley, R., Zisserman, A.: Multiple view geometry in computer vision. Cambridge University Press (2003)
17. Horn, B.K.P.: Closed-form solution of absolute orientation using unit quaternions. *Journal of the Optical Society of America A* **4**(4), 629–642 (1987)
18. Izadi, S., Kim, D., Hilliges, O., Molyneaux, D., Newcombe, R., Kohli, P., Shotton, J., Hodges, S., Freeman, D., Davison, A., Fitzgibbon, A.: KinectFusion: real-time 3D reconstruction and interaction using a moving depth camera. In: Proceedings of the

- 24th annual ACM Symposium on User Interface Software and Technology, p. 559–568. ACM (2011)
19. John, S., Poh, A.C.C., Lim, T.C.C., Chan, E.H.Y., Chong, L.R.: The iPad tablet computer for mobile on-call radiology diagnosis? auditing discrepancy in CT and MRI reporting. *Journal of Digital Imaging* **25**(5), 628–634 (2012)
 20. Khoshelham, K., Elberink, S.O.: Accuracy and resolution of kinect depth data for indoor mapping applications. *Sensors* **12**(12), 1437–1454 (2012)
 21. Kilgus, T., Franz, A., Seitel, A., März, K., Bartha, L., Fangerau, M., Mersmann, S., Groch, A., Meinzer, H.P., Maier-Hein, L.: Registration of partially overlapping surfaces for range image based augmented reality on mobile devices. In: SPIE Medical Imaging, pp. 83,160T–83,160T. International Society for Optics and Photonics (2012)
 22. Kilgus, T., dos Santos, T.R., Seitel, A., Yung, K., Franz, A.M., Groch, A., Wolf, I., Meinzer, H.P., Maier-Hein, L.: Generation of triangle meshes from time-of-flight data for surface registration. In: Bildverarbeitung für die Medizin 2011, p. 189–193. Springer (2011)
 23. Kirk, D.B., Wen-mei, W.H.: Programming massively parallel processors: a hands-on approach. Newnes (2012)
 24. Kramer, J., Burrus, N., Echtler, F., Daniel, H.C., Parker, M.: Hacking the Kinect. Springer (2012)
 25. Li, H., Sumner, R.W., Pauly, M.: Global correspondence optimization for non-rigid registration of depth scans. In: Eurographics Symposium on Geometry Processing, vol. 27, p. 1421–1430. Wiley Online Library (2008)
 26. Lorensen, W.E., Cline, H.E.: Marching cubes: A high resolution 3d surface construction algorithm. In: ACM Siggraph Computer Graphics, vol. 21, p. 163–169. ACM (1987)
 27. Maier-Hein, L., Fangerau, M., Meinzer, H.P., Seitel, A.: Visualization of anatomical data by augmented reality (2013). US Patent App. 13/891,310
 28. Maier-Hein, L., Franz, A.M., Fangerau, M., Schmidt, M., Seitel, A., Mersmann, S., Kilgus, T., Groch, A., Yung, K., dos Santos, T.R., Meinzer, H.P.: Towards mobile augmented reality for on-patient visualization of medical images. In: Bildverarbeitung für die Medizin 2011, Informatik Aktuell, pp. 389–393. Springer (2011)
 29. Maier-Hein, L., Franz, A.M., Santos, T.R.d., Schmidt, M., Fangerau, M., Meinzer, H.P., Fitzpatrick, J.M.: Convergent iterative closest-point algorithm to accomodate anisotropic and inhomogenous localization error. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **34**(8), 1520–1532 (2012)
 30. Mersmann, S., Seitel, A., Erz, M., Jähne, B., Nickel, F., Mieth, M., Mehrabi, A., Maier-Hein, L.: Calibration of time-of-flight cameras for accurate intraoperative surface reconstruction. *Medical Physics* **40**(8) (2013)
 31. Müller, M., Rassweiler, M.C., Klein, J., Seitel, A., Gondan, M., Baumhauer, M., Teber, D., Rassweiler, J.J., Meinzer, H.P., Maier-Hein, L.: Mobile augmented reality for computer-assisted percutaneous nephrolithotomy. *International Journal of Computer Assisted Radiology and Surgery* **8**(4), 663–675 (2013)
 32. Myronenko, A., Song, X.: Point set registration: Coherent point drift. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **32**(12), 2262–2275 (2010)
 33. Navab, N., Traub, J., Sielhorst, T., Feuerstein, M., Bichlmeier, C.: Action- and workflow-driven augmented reality for computer-aided medical procedures. *Computer Graphics and Applications, IEEE* **27**(5), 10–14 (2007)
 34. Nolden, M., Zelzer, S., Seitel, A., Wald, D., Müller, M., Franz, A.M., Maleike, D., Fangerau, M., Baumhauer, M., Maier-Hein, L., Maier-Hein, K.H., Meinzer, H.P., Wolf, I.: The medical imaging interaction toolkit: challenges and advances. *International Journal of Computer Assisted Radiology and Surgery* **8**(4), 607–620 (2013)
 35. Qiu, D., May, S., Nüchter, A.: GPU-Accelerated nearest neighbor search for 3D registration. In: Proceedings of the 7th International Conference on Computer Vision Systems, p. 194–203. Springer (2009)
 36. Roberts, I.S., Benamore, R.E., Benbow, E.W., Lee, S.H., Harris, J.N., Jackson, A., Mallett, S., Patankar, T., Peebles, C., Roobotom, C., Trail, Z.C.: Post-mortem imaging as an alternative to autopsy in the diagnosis of adult deaths: a validation study. *The Lancet* **379**(9811), 136–142 (2012)
 37. Rusu, R.B., Blodow, N., Beetz, M.: Fast point feature histograms (FPFH) for 3D registration. In: Robotics and Automation, IEEE International Conference on, p. 3212–3217. IEEE (2009)
 38. dos Santos, T.R., Seitel, A., Kilgus, T., Suwelack, S., Wekerle, A.L., Kenngott, H., Speidel, S., Schlemmer, H.P., Meinzer, H.P., Heimann, T., Maier-Hein, L.: Pose-independent surface matching for intra-operative soft-tissue marker-less registration. *Medical Image Analysis* (accepted subject to minor revisions) (2014)
 39. Seeling, P., Fitzek, F.H., Ertl, G., Pulipaka, A., Reisslein, M.: Video network traffic and quality comparison of VP8 and H.264 SVC. In: Proceedings of the 3rd workshop on Mobile Video Delivery, p. 33–37. ACM (2010)
 40. Seitel, A., Engel, M., Sommer, C.M., Radeleff, B.A., Essert-Villard, C., Baegert, C., Fangerau, M., Fritzsche, K.H., Yung, K., Meinzer, H.P.: Computer-assisted trajectory planning for percutaneous needle insertions. *Medical Physics* **38**(6), 3246–3259 (2011)
 41. Seitel, A., Yung, K., Mersmann, S., Kilgus, T., Groch, A., dos Santos, T.R., Franz, A.M., Nolden, M., Meinzer, H.P., Maier-Hein, L.: MITK-ToF - range data within MITK. *International Journal of Computer Assisted Radiology and Surgery* **7**(1), 87–96 (2012)
 42. Sugimoto, M., Yasuda, H., Koda, K., Suzuki, M., Yamazaki, M., Tezuka, T., Kosugi, C., Higuchi, R., Watayo, Y., Yagawa, Y., Uemura, S., Tsuchiya, H., Azuma, T.: Image overlay navigation by markerless surface registration in gastrointestinal, hepatobiliary and pancreatic surgery. *Journal of Hepato-Biliary-Pancreatic Sciences* **17**(5), 629–636 (2010)
 43. Thali, M.J., Braun, M., Wirth, J., Vock, P., Dirnhofer, R.: 3D surface and body documentation in forensic medicine: 3-D/CAD photogrammetry merged with 3D radiological scanning. *Journal of Forensic Sciences* **48**(6), 1356–1365 (2003)
 44. Thali, M.J., Dirnhofer, R., Vock, P.: The virtopsy approach: 3D optical and radiological scanning and reconstruction in forensic medicine. CRC Press/Taylor & Francis (2009)
 45. Urschler, M., Höller, J., Bornik, A., Paul, T., Giretzlehner, M., Bischof, H., Yen, K., Scheurer, E.: Intuitive presentation of clinical forensic data using anonymous and person-specific 3d reference manikins. *Forensic Science International* **241**, 155–166 (2014)
 46. Valette, S., Chassery, J.M., Prost, R.: Generic remeshing of 3D triangular meshes with metric-dependent discrete voronoi diagrams. *IEEE Transactions on Visualization and Computer Graphics* **14**(2), 369–381 (2008)
 47. Yang, J., Li, H., Jia, Y.: Go-ICP: solving 3D registration efficiently and globally optimally. In: International Conference on Computer Vision, pp. 1457–1464. IEEE (2013)
 48. Zhang, Z.: A flexible new technique for camera calibration. *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **22**(11), 1330–1334 (2000)
 49. Zhu, E., Masiello, I., Hadadgar, A., Zary, N.: Augmented reality in healthcare education: an integrative review. *PeerJ PrePrints* **2**, 1–22 (2014)