



Illuminating precise stencils on surgical sites using projection-based augmented reality

Muhammad Twaha Ibrahim ^{a,*}, Aditi Majumder ^a, M. Gopi ^a, Lohrasb R. Sayadi ^b, Raj M. Vyas ^b

^a School of Information and Computer Science, University of California, Irvine, United States of America

^b School of Medicine, University of California, Irvine, United States of America



ARTICLE INFO

Keywords:

Projection Augmented Reality
Spatially Augmented Reality
Cleft lip
Real-time
Surgical stencils
Remote surgery
Telemedicine

ABSTRACT

In this paper we propose a system that connects surgeons to remote or local experts who provide real-time surgical guidance by illuminating salient markings or stencils (e.g. points, lines and curves) on the physical surgical site using a projector. The projection can be modified in real time by the expert using a GUI and can be seen by all in the operating room (OR) without the use of any wearables. This system overcomes the limitations of AR/VR headsets which can overlay information through a headset, but are obtrusive, not very accurate with movements, and visible only to the surgeon excluding others in the room. Overlaying information, at high precision, directly on the physical surgical site that can be seen by everyone in the OR can become an useful tool for skill transfer, expert consultation and training, especially in telemedicine.

In addition to the projector, the system comprises of a RGB-D camera (e.g. Kinect) for feedback, together designated as the PDC (Projector Depth Camera) unit. The PDC is driven by a PC. The RGB-D camera provides depth information in addition to an image at video frame rates. A high resolution mesh of the surgical site is captured using the PDC unit initially. During the surgical planning, training or execution session, this digital model can be marked by appropriate incision markings on a tablet or monitor using touch based or mouse based interface, on the same local machine or after being transmitted to a remote machine. These markings are then communicated back to the PDC unit and illuminated at high precision via the projector on the surgical site in real time. If the surgical site moves during the process, the movement is tracked and updated quickly on the surgical site. Our method specifically overcomes the obtrusive, exclusive, and indirect attributes of headsets and displays while maintaining high accuracy of registration with movements.

1. Introduction

Augmented Reality (AR) headsets are being used to assist surgeons in overlaying patient scans (like CT, MRI etc.) on the human body (Andrews, Khan, Mallela, Maroon, Biehl, Fernandes-Cabral, Gersey, Abou-Al-Shaar, Gardner, & Zenonos, 2018; Figueira, Ibrahim, Majumder, & Gopi, 2022; Khan, Andrews, et al., 2022; Khan et al., 2023). In addition to visualization, this overlay can also aid in pre-operative surgical planning. Instead of marking the incisions and blood vessels by referring to CT scans separately,

* Corresponding author.

E-mail addresses: muhammti@uci.edu (M.T. Ibrahim), majumder@ics.uci.edu (A. Majumder), gopi@ics.uci.edu (M. Gopi), rsayadi@hs.uci.edu (L.R. Sayadi), rajv1@hs.uci.edu (R.M. Vyas).

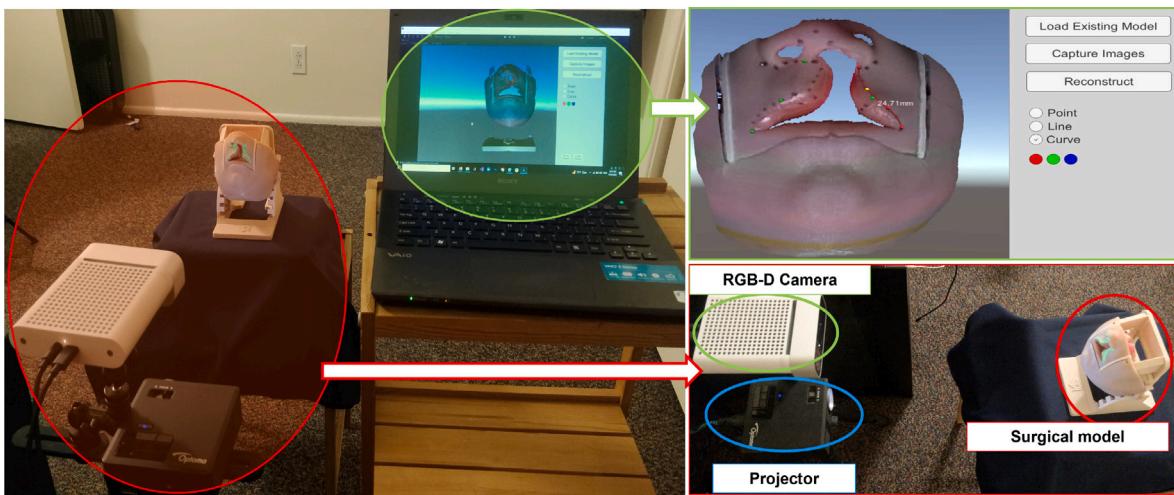


Fig. 1. The setup comprising a RGB-D camera and a projector illuminating a model of the surgical area.

AR/VR headsets allow marking using cues from the overlaid digital information coming from the headset increasing the accuracy of the markings (Khan, Biehl, Andrews, & Babichenko, 2022). Though immersive, wearables are cumbersome and become a barrier to accessing the surgical site directly. Further, the overlaid information is only visible to the surgeon and not to other people (e.g. surgical staff, residents etc.) making communication or collaboration difficult amongst the team. In contrast, our digital overlay via the projector appears directly on the patient's body, making the information visible to everyone. Our projection conforms to the shape of the surgical site putting the digital data in physical alignment with the 3D shape of the surgical site providing the best possible context to improve surgical understanding. It can also be used in preoperative planning for iterative revision of the incision marks that is currently achieved by drawing directly on the surgical site using a marker and erasing and redrawing for changes. This process is not only time-consuming and cumbersome, but often leads to a 'dirty' surgical site where legibility of markings are impacted.

The proposed system falls in the broad umbrella of Projection-based Augmented Reality (PAR) where projectors are used to augment real objects to create visualizations that can be seen together by multiple people (Ibrahim, Gopi, & Majumder, 2020, 2023b, 2024; Ibrahim, Gopi, Vyas, Sayadi, & Majumder, 2023; Ibrahim, Majumder, & Gopi, 2022; Tehrani, Ibrahim, Majumder, & Gopi, 2023). Use of display-based AR (DAR) where a user has to look at a tablet, monitor or headsets to surgical assistance, navigation and visualization have been explored in recent years. However, prior PAR systems have been explored in a very limited manner. They are usually heavily engineered using custom devices, need preoperative specialized scans of the surgical site which are usually registered manually using fiducials or markers in the operating room (OR). In some cases, 3D shapes are not even taken into account resulting in a loosely registered projection. No practical method exists to bind the projection accurately with movement of the surgical site.

We first use the PDC for a structured light scan to create a high-resolution digital model of the surgical area in the pre-processing stage in around 3 min. This model is then loaded in an GUI on another machine (desktop, tablet or laptop) being used remotely or locally by an expert. The expert surgeon then marks the 3D model in the GUI using touch-based or pen-based or mouse based interactions in a tablet or computer (Fig. 1). The high resolution model is registered to the lower-resolution depth camera frame capture throughout the entire session of the interaction. Any markings made on the digital model shows up at a high precision at the surgical site in real time. Movements of the surgical site is common during the surgery. Therefore, we track the moving surgical area and update the projection quickly so that the digital overlay binds to the surgical site accurately even after the movement.

We demonstrate the PDC based surgical guidance capability on a cleft lip clinical model focusing on craniofacial surgery. The choice of craniofacial surgery for initial demonstration and testing is driven by the computational challenges posed by enabling precise 3D projection on a very small and richly three-dimensional surgical site of infants. Cleft surgery, in particular, is one of the most challenging and impactful surgeries that require precision preoperative planning in order to ensure a correct and balanced repaired lip (Vyas & Warren, 2014). Errors in preoperative planning can result in imbalanced and asymmetric repair which requires follow-up corrective surgeries. Due to the sensitive nature of the preoperative planning for cleft repair, surgeons often have to revisit and revise the surgical markings, a process that can take 20%–30% of OR time (e.g. 1.5–2 h in a 5–7 h surgery) due to repeated erasing and redrawing the markings. Using optics based marking from the proposed system also eliminates the special care and attention that has to be devoted to the sterility of the markers and the erasers. Fig. 2 shows the anthropometric landmarks for a unilateral nasolabial repair. Note that a real surgical site on a 1–3 month old infant will be smaller than the size of this image on a A8 letter-sized paper.

Every step of this work has been done in collaboration two renowned reconstructive surgeons, Dr. Raj M. Vyas and Dr. Lohrash R. Sayadi, who were the motivation for the application of the PAR technology in this domain. Dr. Vyas is the Chief of Plastic Surgery

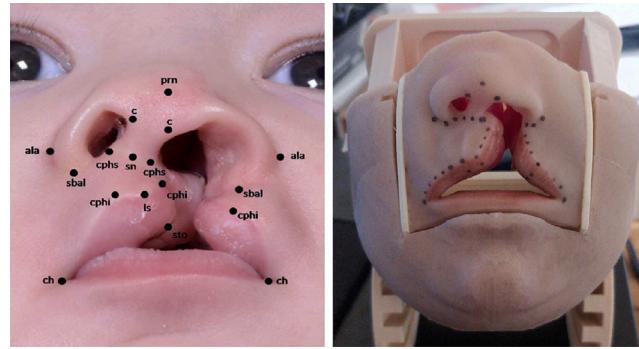


Fig. 2. (Left) Anthropometric landmarks on a unilateral cleft patient. (Right) Unilateral cleft model.

Table 1

This table shows the comparison of our method with existing PAR systems.

Method	Projection based	Automatic Calibration	Handles Movement	3D-shape Reconstruction	Markerless
Tardif, Roy, and Meunier (2003)	✓	N/A			✓
Gavaghan, Peterhans, Oliveira-Santos, and Weber (2011)	✓	✓		✓	
Hummelink et al. (2015), Hummelink, Kool, and Ulrich (2016), Sotsuka, Matsuda, Fujita, Fujiwara, and Kakibuchi (2014), Tabrizi and Mahvash (2015), Volonté, Pugin, Bucher, Sugimoto, Ratib, and Morel (2011)	✓				
Wen, Chui, Ong, Lim, and Chang (2013)	✓	✓		✓	✓
Edgcumbe, Pratt, Yang, Nguan, and Rohling (2015)	✓			✓	✓
Chae, Ganhewa, Hunter-Smith, and Rozen (2018)	✓				✓
Fukuhara et al. (2019)	✓	✓			✓
Ours	✓	✓	✓	✓	✓

at CHOC Children's Hospital and Professor of Plastic Surgery and Neurosurgery at the School of Medicine, University of California, Irvine. He specializes in adult and pediatric facial reconstruction of congenital, traumatic and oncologic conditions. Dr. Sayadi is a resident physician in the Plastic and Reconstructive Surgery Department at the University of California, Irvine. Both Dr. Vyas and Dr. Sayadi have been working with us by providing us the models for development of our system and in marking the models with accurate real surgical stencils which are used for evaluation. The system we present has undergone multiple rounds of testing and evaluation by the doctors themselves and their comments are documented in the results section.

Main Contributions: The main contributions of the proposed work are as follows.

1. To the best of our knowledge, we have built the first end to end system that allows digital surgical guidance markings on a 3D model to be transferred accurately on a physical object via a PDC unit at high accuracy in real time;
2. We have developed a method that uses a combination of computer vision and optimization techniques to compute the camera pixel to projector pixel correspondence at the precision required in surgical applications;
3. We have developed a method that uses the PDC unit to project precision surgical markings that is updated quickly with movement;
4. We have designed a user-friendly GUI that is readily usable by surgeons for skill transfer;
5. We have tested and validated our system using feedback from real surgeons who played with the developed prototype. We have evaluated the system efficiency and functionality using two reputed reconstructive surgeons who anticipate using this system in the OR. Their initial feedback on the accuracy and effectiveness is documented in the results section and is extremely encouraging in terms of scaling the system towards other types of surgeries.

2. Related work

2.1. Display-based AR (DAR)

DAR systems overlay information on a digital representation of the surgical site (e.g. image, video, 3D mesh) and show it on a dedicated display device (e.g. a tablet or monitor or 3D display). Wang et al. (2013) uses a regular tablet or monitor display to

visualize a needle as it is inserted into a patient's teeth during an oral and maxillofacial (OMS) surgery. They use a calibrated stereo camera pair to track the contours of the patient's teeth and reconstruct the 3D model in real time. By tracking the needle using a set of dot markers and registering the reconstructed teeth with a preoperative model, they can visualize the position of the needle tip as it is inserted into the patient on the display. In Wang et al. (2014) and Wang, Suenaga, Liao, et al. (2015), Wang et al. extend their work by using 3D display device (i.e. a lens array monitor) through which the patient can be seen. The 3D display is used to overlay and visualize the teeth model on the patient seen through the display. The teeth contours are tracked every frame by a stereo camera setup so that the teeth model realigns with patient movement. However, the visualizations are only visible when viewed through the 3D display. In Wang, Suenaga, Yang, et al. (2015), Wang et al. improve the tracking system by using a single monochrome camera (instead of stereo cameras) and a novel image registration technique (Wang, Shen, & Yang, 2019) to achieve the digital overlay. In these systems, since the digital representation and the physical object are not overlaid, error in geometry recovery or registration between them, especially with patient movement, cannot be evaluated robustly. Further, while DAR systems are useful to visualize 3D models, on a surgical area, they require users to look into a display. This constant context switch from surgical site to the display impedes the hand eye coordination increasing the time for surgery (Vyas, Sayadi, Bendit, & Hamdan, 2020).

2.2. Projection-based augmented reality (PAR)

The limitations of DAR systems can be overcome by projection-based augment reality (PAR) systems. When describing PAR systems we do not consider systems that use projectors only for structured light scanning purposes and not for augmenting physical objects. For example, Edgcumbe et al. (2015) develop a small device using a pico projector, called the PicoLantern, for use in laproscopic surgeries. Using the camera in an endoscope and the pic projector, they can perform structured light scanning of organs inside the body. However, the projector is not used to add information on to the 3D surface.

When considering information augmentation by projectors on surgical sites, Tardif et al. (2003) propose a rudimentary system to project content on a patient. Using a camera atop the surgical site, they first obtain projector-camera pixel correspondences using structured light patterns. These correspondences are used to warp the content to look correct from the camera's viewpoint. Since the 3D shape of the surgical site is not considered, it can only be useful for projecting on roughly flat areas of the body (e.g. abdomen) and will look correct only from the single view of the camera. Also, since they do not calibrate the projector and camera pair, they cannot bind the projection with patient movement when they repeat the process to find new projector-camera correspondences. Chae et al. (2018) proposes another basic system to project CT scan data on the patient using a pico projector. They assume that the projector is looking directly down on the patient and adjust the height manually until the projection aligns correctly. Fukuhara et al. (2019) use a pre-calibrated projector-camera pair to project blood vessels on the abdomen. They compute the camera and projector pose by matching the outline of the abdomen model in the camera image to the CT scan of the abdomen.

Gavaghan et al. (2011) design a handheld projector device called an Image Overlay Device (IOD) which they use to project onto a static surgical area. The 3D shape of the surgical site is scanned apriori using a custom high resolution depth camera. The scan is also registered with a camera atop the surgical site apriori. The IOD has a 3D tracker attached to it which is tracked by the aforementioned camera. The same camera is also used to detect markers on the surgical site to align to the digital 3D shape with the physical surgical site. Using the tracked projector pose and orientation and the marker-based alignment of the digital shape with the physical, correct projections can be achieved on the surgical site. Since the camera atop the surgical site, the 3D model and the markers are all tightly calibrated apriori, the system must be recalibrated if the camera, surgical site or the markers move.

Wen et al. (2013) propose a projection-based visual guidance system to assist with radiofrequency (RF) ablation needle insertion. The surgical site, i.e. the abdomen, is surrounded by a rig of pre-calibrated stereo cameras and a projector. Structured light is used to find the 3D shape of the abdomen from the calibrated rig. Then, they project the critical structures of the inside of the abdomen on the abdomen itself, including the tumor and the needle trajectory in a view-dependent manner by tracking the viewer's position. In case of movement of the abdomen, the system must be re-calibrated. Other similar systems include Tabrizi and Mahvash (2015) for neuronavigation and image-guided neurosurgery. They compute the warping for the projector image by manually registering the projection using five fiducial markers placed on the patient head.

Other works (Sotsuka et al., 2014; Volonté et al., 2011) project slices of a CT scan on the body to visualize the internal body structures. They manually align the projection to match physical features on the body e.g. abdomen or breast, and mark important structures inside the body on the skin using a pen/marker. Hummelink et al. (2015) also project CT scans on the human body by manually aligning the projector while Hummelink et al. (2016) make projector registration easier by marking four points on the CT scan and placing markers at the corresponding physical locations on the body.

In contrast, our system does not require any markers or pre-calibration and can bind to the surgical site with movement. We achieve this by designing methods that can leverage the low-resolution depth information provided by consumer RGB-D camera (e.g. Kinect) in addition to high resolution color information at video rates. Table 1 shows all the PAR systems compared with our proposed system.

3. System overview

Our setup (see Fig. 1) consists of the PDC unit driven by a PC arranged in a manner such that the surgical area is covered by the field of view of both devices. We assume that the depth camera provides registered color and depth images, as is available in most consumer devices (e.g. Kinect). The user (local or remote) is anticipated to use the system via a graphical user interface (GUI) on a tablet or desktop i.e. a client connected to the PC server. The GUI shows a high resolution 3D digital twin of the surgical site. The

user draws surgical guidance marks or stencils (e.g. points, lines, curves, areas) on this digital twin using a preferred interaction modality (e.g. mouse, pen or touch) which are then illuminated precisely on the physical surgical site using the PDC unit in real-time. Also, if the surgical site moves, the stencils bind accurately to it even after movement. Formally, the main objective of our method is to determine the set of projector pixels Ω at each time step which when turned on, illuminate the intended stencils on the surgical site. *We assume uncalibrated projector and camera with no apriori knowledge of the 3D shape of the surgical site.*

Fig. 3 summarizes our system in a flowchart. It consists of five different modules. First, in the *System Parameter Reconstruction* module, we reconstruct the shape of the surgical site (i.e. digital twin) and recover the parameters of the uncalibrated projector and depth camera in the PDC. The digital twin is then displayed by the *Graphical User Interface* that presents it to the user and allows them to inspect it (via different transformations like rotation and zoom) and mark the surgical stencils on it. The marked stencils and recovered device parameters are fed to the *Illumination Module* that finds the desired Ω in real time and projects it appropriately to illuminate the surgical site accurately by with the appropriate projector image generation module. The surgical area is continuously monitored by the PDC unit. If it moves, the *Tracking Module*, running in parallel with the illumination module, recovers the new location of the surgical site and projects an updated imagery for the illumination module to bind the stencils to the moved surgical site quickly (i.e. within seconds). This is achieved by computing a transformation that takes the structured light reconstruction of the surgical site to its moved location and orientation. The following sections explain each of the above modules in detail.

4. System parameter reconstruction

The system parameter reconstruction module uses well known structured light reconstruction (SLR) techniques (Geng, 2011; Rubinsztein-Dunlop, Forbes, Berry, Dennis, Andrews, Mansuripur, Denz, Alpmann, Banzer, Bauer, et al., 2016; Zuo et al., 2018) for shape reconstruction. Known patterns are projected from a projector on the surgical site and captured by a camera. The captured patterns are decoded to compute the projector to camera pixel correspondence at every 3D location. These pixel correspondences are triangulated to generate a 3D model (see **Fig. 3**). Thus, structured light scanning techniques differ mainly on the pattern design and the decoding algorithm.

In our work, we use Micro phase-shift (MPS) technique (Gupta & Nayar, 2012) that uses multiple high-frequency sinusoidal patterns in a narrow frequency range. MPS is robust to global illumination effects, inter-reflections and projector defocus, all of which are challenges associated with scanning the human body, especially in a surgical setting. However, it is important to note that our method is independent of the SLR technique used and can be replaced with other methods like Embedded Phase Shift (Moreno, Son, & Taubin, 2015).

4.1. PDC calibration

The reconstruction of the model can only be as accurate as the SLR pixel correspondence decoding and the projector-camera calibration parameters. Since we start from an uncalibrated PDC unit, we need to calibrate it first. In order to compute the PDC calibration parameters, we again use the pixel correspondence decoding from SLR and approximate the 3D model/depth given by the RGB-D camera available in the PDC unit. We capture the depth map I_D of the surgical area using the RGB-D to get the corresponding 3D point $d_i = I_D(c_i) \in \mathbb{R}^3$ at the camera pixel c_i . Most RGB-D cameras provide this information via their API. Thus, we establish correspondences between the camera, projector and 3D points as a 3-tuple $\{c_i, p_i = L(c_i), d_i = I_D(c_i)\}$ which we use to refine the calibration parameters of the projector-camera pair. The camera calibration parameters, denoted by $\mathbf{M}^c = \{f^c, o^c, k^c\}$ include the focal length $f^c = (f_x^c, f_y^c)$, principal point $o^c = (o_x^c, o_y^c)$ and radial distortion coefficients $k^c = (k_1^c, k_2^c)$. For the projector, the calibration parameters include the intrinsics $\mathbf{M}^p = \{f^p, o^p, k^p\}$ along with the pose (R^p, T^p) relative to the camera.

We assume that the RGB camera is at the origin and looking down the positive Z -axis. Thus, we only need to determine the camera intrinsics. While most RGB-D camera intrinsics are provided by the manufacturer, we noticed that they may not always be accurate and therefore, we optimize for the camera intrinsics using $d_i = I_D(c_i)$ and a standard reprojection solver.

We start with an estimate of the projector intrinsics obtained from (Ibrahim, Gopi, & Majumder, 2023a). Using the pixel correspondences $\{p_i, d_i\}$ established during the SLR, we get an initial estimate of the extrinsics using the Perspective n-Point algorithm (Marchand, Uchiyama, & Spindler, 2015). Note that the extrinsics have the correct scale since we use depth information from the depth camera. Then, we use a standard reprojection solver to optimize the projector parameters. Finally, we perform bundle adjustment using the pixel correspondences and refine the projector extrinsics further by fixing the camera extrinsics. This gives us a good set of calibration parameters which we use for triangulating the SLR pixel correspondences and obtain an accurate 3D digital twin of the surgical area.

Discussion: Despite the presence of the depth camera in the PDC unit, we perform a SLR for the digital twin motivated by the following. (a) While depth cameras are becoming increasingly accurate, they do not capture complex surfaces (e.g. the human face) accurately enough, primarily due to their lower resolution (e.g. the depth camera in Kinect has a resolution of (640×576) that is much lower than the RGB camera at (4096×3072)). **Fig. 4** compares the 3D shape from the structured light scan reconstruction and the depth camera output. Note how the nose is extra pointy in the depth camera, while some parts of the cheek have high error. In order to project precisely onto the surgical area, we need to have an accurate digital twin at the inception to prevent errors due to the inaccurate shape from creeping in. (b) Aside from accuracy, depth cameras also show high quantization errors. The depth at each point is quantized to a whole number instead of a floating point. This loss of precision shows up as inaccuracies in the final projection. SLR solves both these problems. By determining the pixel correspondences and triangulating them, we get an accurate, high resolution 3D shape reconstruction of the surgical site at a high precision.

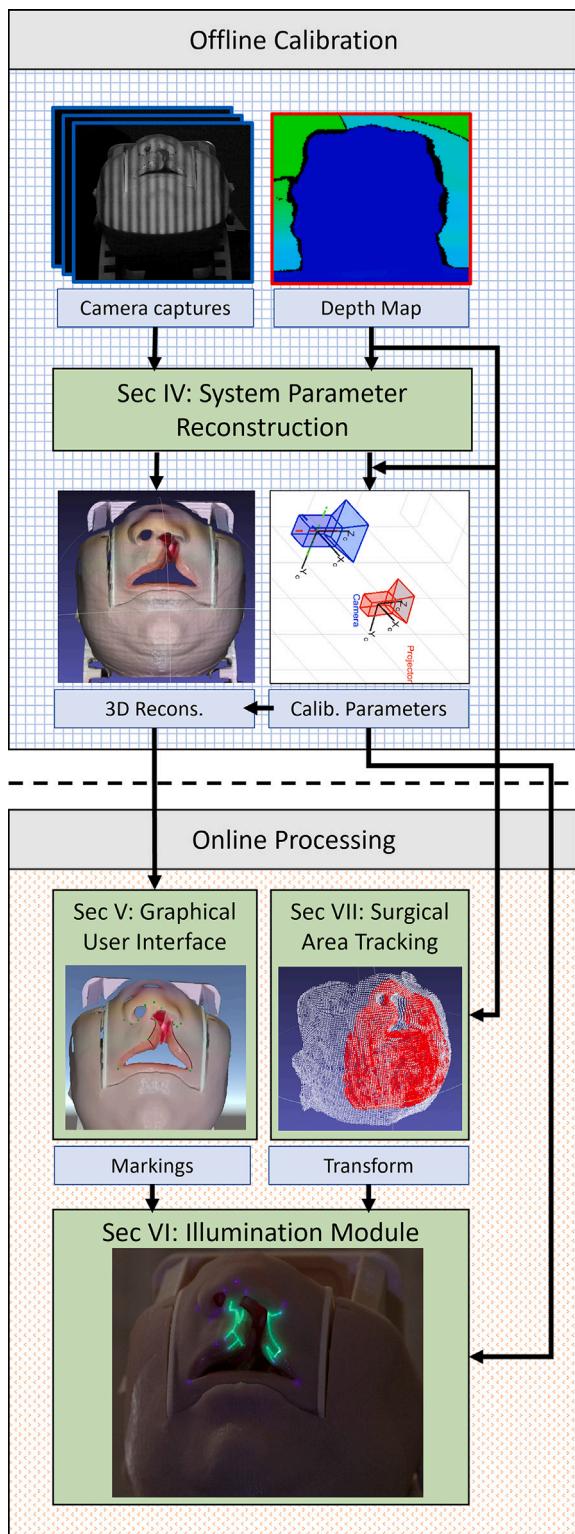


Fig. 3. The flowchart of our system.

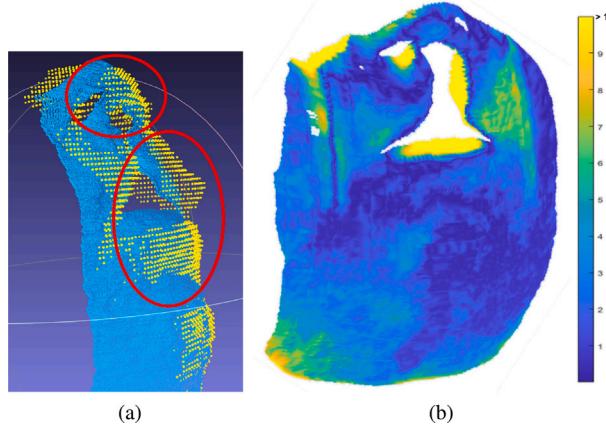


Fig. 4. Comparison of the depth camera output with the reconstruction using a calibrated PDC unit and structured light. (a) The structured light point cloud (cyan) is shown superimposed with the depth camera point cloud (yellow). The red circles highlight some of the mismatched regions. (b) The distance (in mm) between the structured light reconstruction and the depth map of the cleft face. Notice how the depth map does not reflect the face geometry accurately, especially in non-planar regions like the lips.

5. Graphical user interface (GUI)

The high resolution precise digital twin of the surgical area is presented to the user (remote or local) through a GUI, where they can inspect the model by zooming in and rotating it. Additionally, the user can mark points or curves directly on the digital twin. The GUI adds the 3D coordinates corresponding to the marked points and curves in two sets $P \in \mathbb{R}^{N \times 3}$ and $Q \in \mathbb{R}^{L \times Q \times 3}$ respectively, where N is the number of points, L is the number of curves and Q is the number of points per curve.

Marking Points: To mark points, the user hovers the mouse pointer over the 3D surface. We cast a ray into the 3D scene from the rendering camera center through the 2D mouse position. The point where the ray intersects the 3D model is highlighted for the user by rendering a small sphere. When the user clicks, the sphere is anchored at that point on the 3D model and the 3D coordinates are added to the set P . Fig. 5 shows an example of a user having marked several points (green spheres) on the model. In order to *modify* an existing point, a user clicks on an existing sphere and can place it anywhere else on the model. The corresponding 3D point is updated in P . In order to *remove* a point, a user right-clicks on an existing sphere and the corresponding point in P is removed.

Drawing Curves: When drawing curves, our goal is to draw geodesic curves so that it also lies on the surface of the reconstructed digital twin. We use a piecewise linear representation of the geodesic curve and decipher it in two steps. First, the user clicks and drags the mouse across the model surface. The camera shoots rays into the 3D scene at equal intervals between the start and end points of the curve inclusive. Let there be R such rays, and let r_i denotes the 3D point where the i th ray intersects the 3D model. We render a sphere at each r_i (red and yellow spheres in Fig. 5) and fit a 3D Bezier curve of degree $(r - 1)$ through all r_i 's. If we render this Bezier curve as is, there is no guarantee that it will conform to the surface. Parts of the curve may be below or above the 3D model. The former ones will not even be visible to the user. Therefore, we evaluate the Bezier curve for $Q \gg R$ points densely spaced along the curve and project each 3D point onto the 3D model. Let q_i denote each projected 3D point. Then, we render a line between each consecutive pair of the projected points $\{q_i, q_{i+1}\}$. This ensures that the final rendered curve is geodesic to the surface geometry. The set of projected points q_i comprising the curve are added to the set Q . Fig. 5 shows an example of a user drawing a curve. In order to *modify* an existing curve, the user can select and move any of the points r_i that the Bezier curve passes through. Upon changing any r_i , all q_i 's are updated each frame while the rendered curve is displayed to user. Once satisfied, the user clicks again to finalize the changes. The corresponding curve in Q is also updated. If a user right-click's on an existing curve, the curve is *deleted* from the model as well as from Q . **Please watch the anonymous video at Demo Video (2023) for a demonstration of the proposed system.**

6. Illumination module

The 3D points marked by the user $\{P, Q\}$ and the projector parameters $\{M^P, R^P, T^P\}$ are used to determine the final set of pixels Ω that must be illuminated to project on the intended regions on the surgical area using

$$\bar{W}_i = \begin{bmatrix} R_P & T_P \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} W_i \\ 1 \end{bmatrix} \quad (1)$$

$$\omega_i = \text{project}(\bar{W}_i; M^P)$$

where W_i is a point in $\{P, Q\}$, \bar{W}_i is the same 3D point after applying the projector extrinsics, $\omega_i \in \mathbb{R}^{2 \times 1}$ is the i th point in Ω and $\mathbf{0} \in \mathbb{R}^{1 \times 3}$. All pixels Ω are illuminated in the final image which is then projected onto the surgical area.

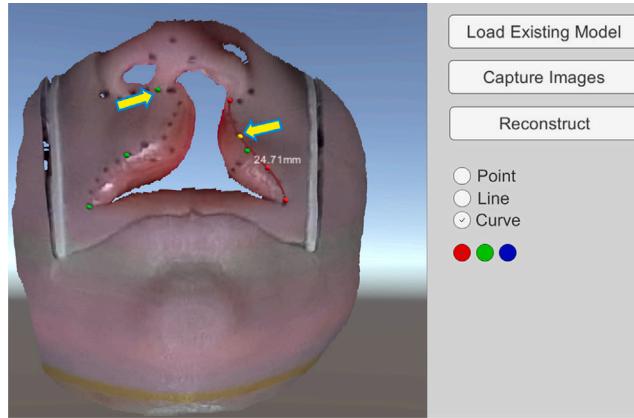


Fig. 5. A screenshot of our UI. The user can mark points (green spheres) and/or draw curves and lines (red curve).

Table 2

Projection accuracy evaluation in mm. θ (in degrees) and δ (in mm) are the rotation and translation of the model from its original position. Nomenclature: c' : cleft side, nc' : non-cleft side, m' : medial, l' : lateral.

Keypoint	$\theta = 0$ $\delta = 0$	$\theta = 0$ $\delta = 30$	$\theta = 15$ $\delta = 50$	$\theta = 15$ $\delta = 100$	$\theta = 30$ $\delta = 200$
prn	0	0.16	0.42	0.38	0.19
sn	0	0.26	0.21	0.28	0.25
nc'sbal	0	0.10	0.22	0.12	0.15
c'sbal	0	0.15	0.35	0.13	0.18
m'cphi	0	0.23	0.11	0.25	0.15
l'cphi	0	0.16	0.18	0.40	0.10
c'cphi	0	0.39	0.15	0.40	0.25
nc'ch	0	0.40	0.50	0.58	0.68
c'ch	0	0.47	0.48	0.46	0.48

7. Surgical area tracking

During surgery or planning thereof, it is likely that the surgical area will move with respect to the PDC unit. Therefore, we track the surgical area in 3D space using the depth camera. We determine the rigid transformation that minimizes the error between the current depth map from the camera and the depth map captured at the time of SLR. In order to achieve this we use the Iterative Closest Point (ICP) algorithm represented by:

$$\epsilon_{ICP}^k = \sum_i \sum_j v_{ij}^k \|d_i^k - S^k \cdot d_j^{k-1}\| \quad (2)$$

where d_i^k is a 3D point in the depth map at frame k , v_{ij}^k is a binary variable that is 1 only if ICP matches points d_i^k to d_j^{k-1} , and $S^k \in \mathbb{R}^{3 \times 4}$ is the rigid transformation computed at frame k . In order to adapt to the movement we modify the projection mapping equation (Eq. (1)) to include S^k as:

$$\begin{aligned} \bar{W}_i &= \begin{bmatrix} R_P & T_P \\ \mathbf{0} & 1 \end{bmatrix} S^k \begin{bmatrix} W_i \\ 1 \end{bmatrix} \\ \omega_i &= project(\bar{W}_i; M^P) \end{aligned} \quad (3)$$

where W_i is a point in $\{P, Q\}$, \bar{W}_i is the same 3D point after applying the rigid transform and the projector extrinsics, $\omega_i \in \mathbb{R}^{2 \times 1}$ is the i th point in Ω and $\mathbf{0} \in \mathbb{R}^{1 \times 3}$. All pixels Ω are illuminated in the final image which is then projected onto the surgical area.

To achieve this efficiently, we use the lower resolution depth map captured during the SLR rather than the high resolution model recovered by the SLR, even though the latter is a more accurate representation of the surgical area. This density of the point cloud from the SLR is much more than the point cloud captured by the depth camera. This results in ICP computing incorrect matches resulting in imprecise S^k as well as becoming inefficient. Further, when trying to match the depth camera point cloud with the high resolution point cloud from SLR, the noise between consecutive depth camera frame manifests itself in noisy computation of S^k . This results in jittery projection. Using the depth map captured during SLR avoids all these aforementioned issues.

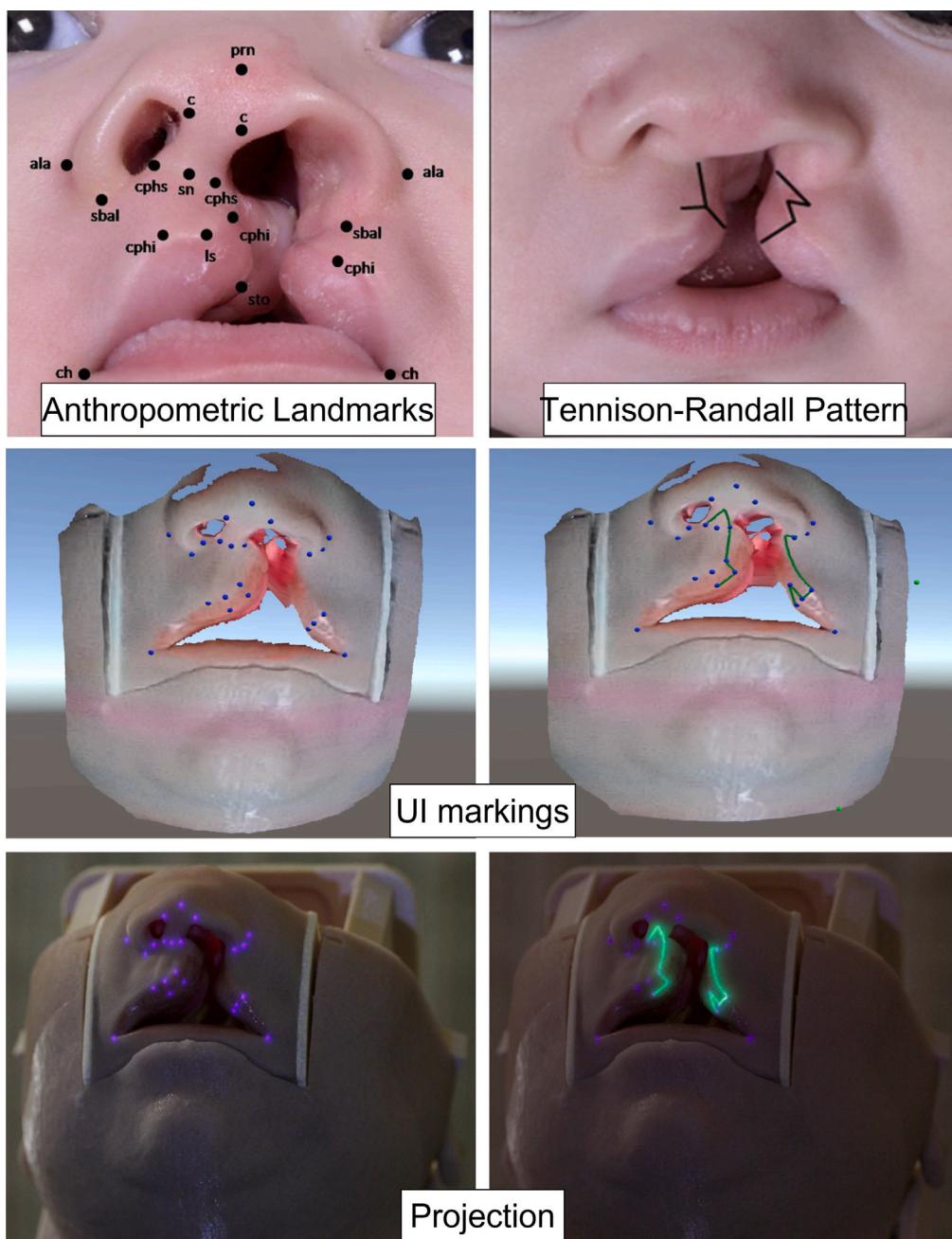


Fig. 6. Anthropometric landmarks and the Tennison-Randall pattern marked on the SLS through the UI and projected onto the model using our system.

8. Implementation and results

In our setup, we used the Microsoft Azure Kinect RGB-D camera along with an Optoma Technology ML750 DLP projector. The Azure Kinect consists of a depth camera and a RGB camera. It provides different operating modes depending on the capture resolution ([Microsoft, 2022](#)). Note that the highest RGB camera resolution (4096×3072) is much higher than the corresponding depth camera resolution (640×576). We used the highest resolution (4096×3072) during SLR for a high resolution model. However, for ICP, we use the lowest resolution of the Kinect RGB i.e. (1280×720). The projector resolution is (1920×1080). The surgical models were typically placed around 1.5–2 ft away from the PDC unit to accommodate for the projector focus and Kinect depth sensor limitations, which can measure the depth for objects more than 1.5 ft away.

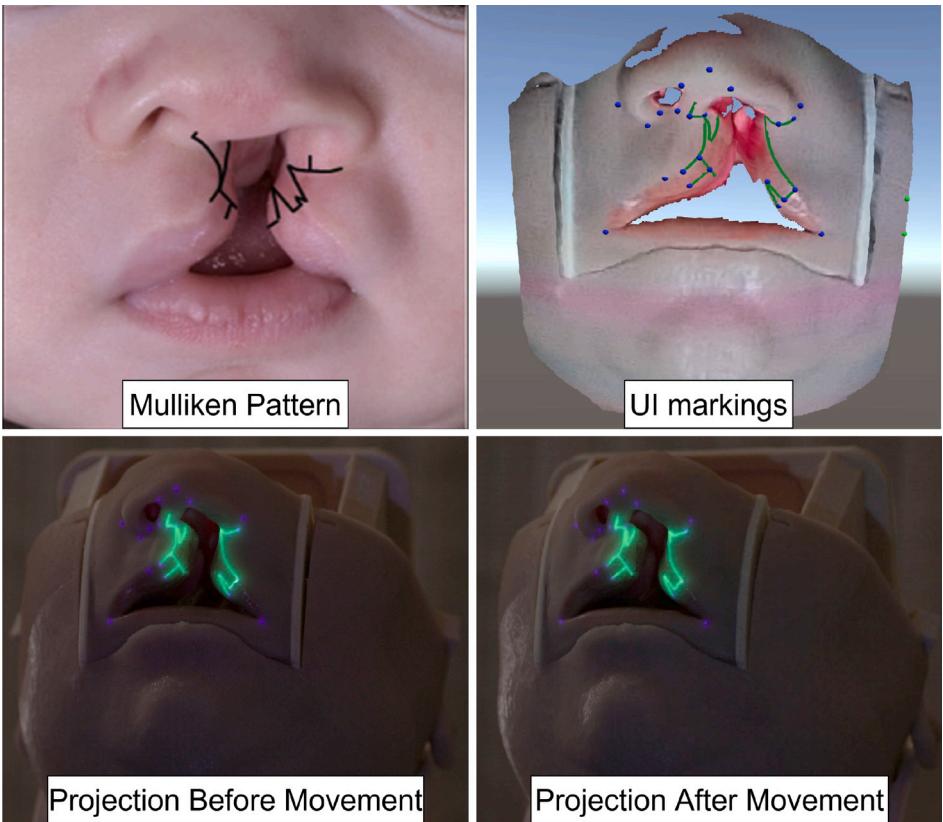


Fig. 7. Our system can handle movement of the surgical area as well. The projection of the Mulliken pattern marked through the UI and projected onto the model realigns after the model is moved.



Fig. 8. This figure shows some frames of the Mulliken projection realigning after it is moved (from left to right).

The intrinsics and extrinsics of the RGB and depth cameras can be accessed through the Kinect API. We use the RGB camera intrinsics to initialize the camera parameter optimization. The Kinect has a depth resolution of up to a millimeter and provides a low resolution raw depth data as well as a higher resolution interpolated depth map rendered from the RGB camera's viewpoint. During calibration, we use the raw depth map since the interpolated depth map has interpolation artifacts like depth quantization. The PDC calibration optimizations were implemented in MATLAB. For computing an initial estimate of the projector intrinsics, we used the projector-camera calibration technique proposed by (Ibrahim et al., 2023a, 2023b) that projects ArUCO markers on a deformable surface to compute calibration parameters. The GUI and 3D model interaction was developed using Unity rendering engine. For tracking the surgical area, we used OpenCV and the ICP implementation provided by *libpointmatcher* (Pomerleau, Colas, Siegwart, & Magnenat, 2013; Pomerleau, Magnenat, Colas, Liu, & Siegwart, 2011).

Fig. 6 shows the projection of the 21 anthropomorphic landmarks and the Tennison-Randall pattern on the cleft model. The user marks the points and curves on the SLS through the UI. Notice how the projection illuminates the models at the marked points precisely. Fig. 7 shows the projection of the Mulliken pattern marked through the UI by the user. Even after the model is rotated and translated, the projection realigns to illuminate the correct areas. Fig. 8 shows the frame-by-frame movement of the projection after the model is moved.

The proposed system can also be used for other surgeries. Fig. 9 shows the proposed system being used to projecting patterns for face reconstruction and breast reduction surgeries. The system reconstructs the surgical area from the SLR, the user marks the points and curves on it and the system projects onto the surgical areas.

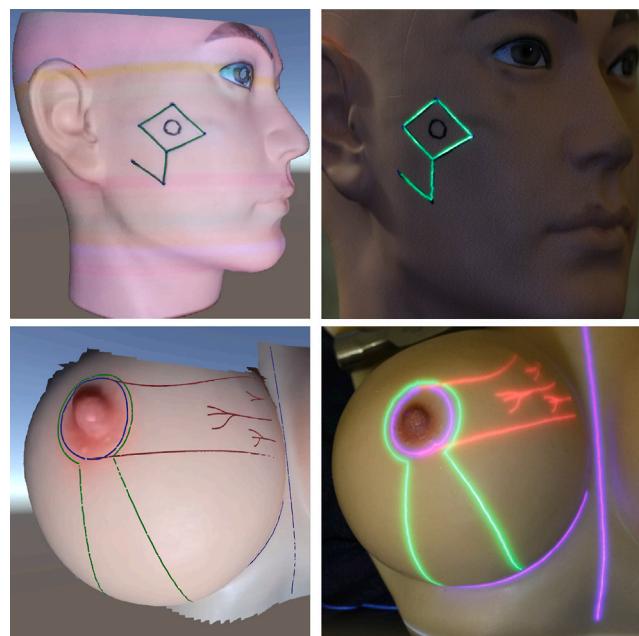


Fig. 9. The proposed system can be used for other surgical areas as well e.g. the face projected with the rhomboid pattern and a model of a human breast for breast reduction surgery.

8.1. Performance and evaluation

Timing: On average, the image acquisition for SLR took 1 min while decoding took 3 min. The entire calibration and reconstruction process following decoding took 1.5 min. Thus, the one-time SLR and calibration took 5–6 min. The average time taken to compute a rigid transformation with movement is 0.75 s.

Projector Calibration Accuracy: In order to verify the projector-camera calibration accuracy, we performed SLR of a planar board and compared the 3D reconstruction with the depth camera output. The average error was 2.11 mm. This was in part due to errors in the depth camera. The mean reprojection error in the projector after calibration was 0.1821 pixels. This sub-pixel error shows the accuracy of our calibration and 3D reconstruction methods, enabling precise projection on the actual surgical site.

Projection Accuracy: To verify the accuracy of the projection, we marked the 21 anthropomorphic landmarks on the cleft model with a black pen and ran the calibration pipeline. This resulted in the black markers showing up in the digital twin recovered by SLR. Then, we marked those 21 points on the digital twin and measured the distance between the actual marked surface points and their corresponding projections using a digital caliper for different positions of the model. Table 2 shows the average error (in mm) for 9 of the 21 landmarks. In general, we noticed that the error for all 21 points was less than 0.5 mm, even after rotation and translation. This satisfies the requirements of surgeons, who want sub-millimeter accuracy of the projections and is similar to the accuracy reported by DAR systems as well (Wang, Suenaga, Liao, et al., 2015). However, we did notice a higher error of 0.5–0.6 mm for *c'ch* and *nc'ch* keypoints. This is because the edges of the lip have more error in the SLR decoding and the depth camera. Note that the error is less than the maximum resolution of 1 mm that can be detected by the depth camera. This shows that our method yields extremely accurate results limited by the resolution of the depth camera.

Tracking Accuracy: To verify the accuracy of the model tracking, we attached a board printed with several ArUCo markers to the cleft model. We tracked the ArUCo markers as the model was moved and compared the transforms provided by ICP and the transform computed using the ArUCo markers. The rotation error was 0.1° and the translation error was 0.47 mm.

8.2. Domain validation

We have been working closely with two reconstructive surgeons, Dr. Raj M. Vyas and Dr. Lohrasb R. Sayadi, who have many years of experience with craniofacial surgery. They have been providing feedback at every level of the development process. All the markings shown in the aforementioned images were performed by the doctors on the computer monitor using our UI and projected on the model by our system. Thus, we received feedback on all salient aspects of the system – ease of use, accuracy of projection and so on. Their comments are as follows:

- (1) *In the field of plastic and reconstructive surgery, cleft lip repair marking require the greatest degree of precision and accuracy. These markings are defined by both specific anatomical landmarks of the lip as well as geometrical considerations. To make things even more rigorous these points must be placed in the anatomically complex three dimensional space of the child's lip in an area of less*

than 10 cm². The aforementioned technology presented in this manuscript can accurately place these markings with consistency and with minimal error fit by our standards for the operating room. The ability to project on the complex contour of the lip opens up this technology to the whole human body.

- (2) In its current state, this technology is useful for a variety of plastic surgery procedures spanning from the head to the toes. Once we find the ability to project onto surfaces other than the skin (i.e. fat, fluid) in addition to projecting within crevices, then we can expand the technology to intra-thoracic and intra-abdominal surgeries. What sets this technology apart is the lack of headsets which are cumbersome, ergonomically disadvantageous and intrusive to the surgeon. The fact that everyone in the operating room can see the projection allows for better communication between surgical team members. We are looking forward to test this technology in the operating room.
- (3) The accuracy the PAR system achieves is phenomenal. Cleft lip surgery has perhaps the most rigorous surgical markings given the number of markings in such small anatomical confines. The demonstrated ability to achieve low error rates in cleft lip markings bodes well for the clinical translation of this technology to a wide variety of surgical procedures on the surface of the human body. We have not tested the device in the operating room. However, our colleagues who have seen the mannequin projections are amazed by the technology and see a clear room for its application in plastic surgery for a wide variety of surgeries including craniofacial, hand and microsurgery.

9. Limitations and future work

The proposed work is just the beginning of an effort to build a device that can be integrated in the OR (for example with the OR lights) to project surgical guidance directly on surgical sites. However, there are several important technical challenges that have to be overcome which are not addressed by the proposed system.

In the OR, we will have to deal with occlusion of projection due to obstruction by surgeons. Detecting occlusion and compensating for it from other non-obstructing projectors is a possible solution. This builds the motivation to design a system that scales to multiple PDC units. Another limitation of the proposed system is the depth camera, whose limited resolution and noisy reconstruction can cause minor errors in tracking the surgical site. Using noise removal methods and super-resolution techniques to combine depth from multiple PDCs can be a way to alleviate this issue. Finally, the proposed system assumes that the surgical site is semi-rigid and retains its shape. Therefore, it cannot handle soft tissue whose shape can change.

CRediT authorship contribution statement

Muhammad Twaha Ibrahim: Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Aditi Majumder:** Methodology, Software, Supervision, Writing – original draft, Writing – review & editing. **M. Gopi:** Methodology, Software, Supervision. **Lohrasb R. Sayadi:** Conceptualization, Supervision, Validation, Visualization. **Raj M. Vyas:** Conceptualization, Methodology, Supervision, Validation, Visualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

No data was used for the research described in the article.

References

- Andrews, E. G., Khan, T., Mallela, A., Maroon, J. C., Biehl, J., Fernandes-Cabral, D., et al. (2018). Streaming 2D-Endoscopic Video into an Augmented Reality Headset Display: A Feasibility Study. *Journal of Neurological Surgery Part B: Skull Base*, 83 (S 01) A107.
- Chae, M. P., Ganhewa, D., Hunter-Smith, D. J., & Rozen, W. M. (2018). Direct augmented reality computed tomographic angiography technique (ARC): an innovation in preoperative imaging. *European Journal of Plastic Surgery*, 41(4), 415–420.
- (2023). Demo Video can be viewed at <https://youtu.be/g0bl2dciZ5A>. (Accessed 30 November 2023).
- Edgcumbe, P., Pratt, P., Yang, G.-Z., Nguan, C., & Rohling, R. (2015). Pico Lantern: Surface reconstruction and augmented reality in laparoscopic surgery using a pick-up laser projector. *Medical Image Analysis*, 25(1), 95–102.
- Figueira, I., Ibrahim, M. T., Majumder, A., & Gopi, M. (2022). Augmented reality patient-specific registration for medical visualization. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology*.
- Fukuhara, R., Kaneda, K., Tamaki, T., Raytchev, B., Higaki, T., Nishimoto, S., et al. (2019). A projection mapping system onto a human body for medical applications. In *Eurographics (posters)* (pp. 29–30).
- Gavaghan, K. A., Peterhans, M., Oliveira-Santos, T., & Weber, S. (2011). A portable image overlay projection device for computer-aided open liver surgery. *IEEE Transactions on Biomedical Engineering*, 58(6), 1855–1864.
- Geng, J. (2011). Structured-light 3D surface imaging: a tutorial. *Advances in Optics and Photonics*, 3(2), 128–160.
- Gupta, M., & Nayar, S. K. (2012). Micro phase shifting. In *2012 IEEE conference on computer vision and pattern recognition* (pp. 813–820). IEEE.
- Hummelink, S., Hameeteman, M., Hoogeveen, Y., Slump, C., Ulrich, D., & Kool, L. S. (2015). Preliminary results using a newly developed projection method to visualize vascular anatomy prior to DIEP flap breast reconstruction. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 68(3), 390–394.
- Hummelink, S., Kool, L. S., & Ulrich, D. (2016). Displaying inguinal lymph nodes before transplantation in a deep inferior epigastric perforator flap breast reconstruction using an innovative projection method. *Journal of Plastic, Reconstructive & Aesthetic Surgery*, 69(3), 376–380.

- Ibrahim, M. T., Gopi, M., & Majumder, A. (2020). Dynamic projection mapping of deformable stretchable materials. In *Proceedings of the 26th ACM Symposium on Virtual Reality Software and Technology*.
- Ibrahim, M. T., Gopi, M., & Majumder, A. (2023a). Projector-camera calibration on dynamic, deformable surfaces. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*.
- Ibrahim, M. T., Gopi, M., & Majumder, A. (2023b). Self-calibrating dynamic projection mapping system for dynamic, deformable surfaces with jitter correction and occlusion handling. In *22nd IEEE International Symposium on Mixed and Augmented Reality 2023*.
- Ibrahim, M. T., Gopi, M., & Majumder, A. (2024). Real-time seamless multi-projector displays on deformable surfaces. *IEEE Transactions on Visualization and Computer Graphics*.
- Ibrahim, M. T., Gopi, M., Vyas, R., Sayadi, L. R., & Majumder, A. (2023). Projector illuminated precise stencils on surgical sites. In *2023 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops*.
- Ibrahim, M. T., Majumder, A., & Gopi, M. (2022). Dynamic projection mapping on deformable stretchable materials using boundary tracking. *Computers and Graphics*.
- Khan, T., Andrews, E. G., Gardner, P. A., Mallela, A. N., Head, J. R., Maroon, J. C., et al. (2022). AR in the OR: exploring use of augmented reality to support endoscopic surgery. In *Proceedings of the 2022 ACM international conference on interactive media experiences* (pp. 267–270).
- Khan, T., Biehl, J. T., Andrews, E. G., & Babichenko, D. (2022). A systematic comparison of the accuracy of monocular RGB tracking and LiDAR for neuronavigation. *Healthcare Technology Letters*, 9(6), 91–101.
- Khan, T., Zhu, T. S., Downes, T., Cheng, L., Kass, N. M., Andrews, E. G., et al. (2023). Understanding effects of visual feedback delay in ar on fine motor surgical tasks. *IEEE Transactions on Visualization and Computer Graphics*.
- Marchand, E., Uchiyama, H., & Spindler, F. (2015). Pose estimation for augmented reality: a hands-on survey. *IEEE Transactions on Visualization and Computer Graphics*, 22(12), 2633–2651.
- Microsoft (2022). Microsoft azure kinect. <https://docs.microsoft.com/en-us/azure/Kinect-dk/hardware-specification>. (Accessed 11 September 2022).
- Moreno, D., Son, K., & Taubin, G. (2015). Embedded phase shifting: Robust phase shifting with embedded signals. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 2301–2309).
- Pomerleau, F., Colas, F., Siegwart, R., & Magnenat, S. (2013). Comparing ICP variants on real-world data sets. *Autonomous Robots*, 34(3), 133–148.
- Pomerleau, F., Magnenat, S., Colas, F., Liu, M., & Siegwart, R. (2011). Tracking a depth camera: Parameter exploration for fast ICP. In *2011 IEEE/RSJ international conference on intelligent robots and systems* (pp. 3824–3829). IEEE.
- Rubinstein-Dunlop, H., Forbes, A., Berry, M. V., Dennis, M. R., Andrews, D. L., Mansuripur, M., et al. (2016). Roadmap on structured light. *Journal of Optics*, 19(1), Article 013001.
- Sotsuka, Y., Matsuda, K., Fujita, K., Fujiwara, T., & Kakibuchi, M. (2014). Image overlay of deep inferior epigastric artery in breast reconstruction. *Plastic and Reconstructive Surgery Global Open*, 2(10).
- Tabrizi, L. B., & Mahvash, M. (2015). Augmented reality-guided neurosurgery: accuracy and intraoperative application of an image projection technique. *Journal of Neurosurgery*, 123(1), 206–211.
- Tardif, J.-P., Roy, S., & Meunier, J. (2003). Projector-based augmented reality in surgery without calibration. Vol. 1, In *Proceedings of the 25th annual international conference of the IEEE engineering in medicine and biology society* (IEEE cat. no.03CH37439), (pp. 548–551).
- Tehrani, M. A., Ibrahim, M. T., Majumder, A., & Gopi, M. (2023). 3D gamut morphing for non-rectangular multi-projector displays. *IEEE Transactions on Visualization and Computer Graphics*.
- Volonté, F., Pugin, F., Bucher, P., Sugimoto, M., Ratib, O., & Morel, P. (2011). Augmented reality and image overlay navigation with OsiriX in laparoscopic and robotic surgery: not only a matter of fashion. *Journal of Hepato-Biliary-Pancreatic Sciences*, 18(4), 506–509.
- Vyas, R. M., Sayadi, L. R., Bendit, D., & Hamdan, U. S. (2020). Using virtual augmented reality to remotely proctor overseas surgical outreach: building long-term international capacity and sustainability. *Plastic and Reconstructive Surgery*, 146(5), 622e–629e.
- Vyas, R. M., & Warren, S. M. (2014). Unilateral cleft lip repair. *Clinics in Plastic Surgery*, 41(2), 165–177.
- Wang, J., Shen, Y., & Yang, S. (2019). A practical marker-less image registration method for augmented reality oral and maxillofacial surgery. *International Journal of Computer Assisted Radiology and Surgery*, 14(5), 763–773.
- Wang, J., Suenaga, H., Hoshi, K., Yang, L., Kobayashi, E., Sakuma, I., et al. (2014). Augmented reality navigation with automatic marker-free image registration using 3-D image overlay for dental surgery. *IEEE Transactions on Biomedical Engineering*, 61(4), 1295–1304.
- Wang, J., Suenaga, H., Liao, H., Hoshi, K., Yang, L., Kobayashi, E., et al. (2015). Real-time computer-generated integral imaging and 3D image calibration for augmented reality surgical navigation. *Computerized Medical Imaging and Graphics*, 40, 147–159.
- Wang, J., Suenaga, H., Yang, L., Liao, H., Ando, T., Kobayashi, E., et al. (2015). 3D surgical overlay with markerless image registration using a single camera. In *Workshop on augmented environments for computer-assisted interventions* (pp. 124–133). Springer.
- Wang, J., Suenaga, H., Yang, L., Liao, H., Kobayashi, E., Takato, T., et al. (2013). Real-time marker-free patient registration and image-based navigation using stereovision for dental surgery. In *Augmented reality environments for medical imaging and computer-assisted interventions* (pp. 9–18). Springer.
- Wen, R., Chui, C.-K., Ong, S.-H., Lim, K.-B., & Chang, S. K.-Y. (2013). Projection-based visual guidance for robot-aided RF needle insertion. *International Journal of Computer Assisted Radiology and Surgery*, 8(6), 1015–1025.
- Zuo, C., Feng, S., Huang, L., Tao, T., Yin, W., & Chen, Q. (2018). Phase shifting algorithms for fringe projection profilometry: A review. *Optics and Lasers in Engineering*, 109, 23–59.