IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 16, 2021, accepted August 3, 2021, date of publication August 6, 2021, date of current version August 17, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3103049

Stereo Laryngoscopic Impact Site Prediction for Droplet-Based Stimulation of the Laryngeal Adductor Reflex

JACOB F. FAST[1,2], (Graduate Student Member, IEEE), HARDIK R. DAVA[2], ADRIAN K. RÜPPEL[2], DENNIS KUNDRAT[3], (Member, IEEE), MAURICE KRAUTH[2], MAX-HEINRICH LAVES[4], SVENJA SPINDELDREIER[2], LÜDER A. KAHRS[5,6], AND MARTIN PTOK[1]

[1]Department of Phoniatrics and Pediatric Audiology, Hannover Medical School, 30625 Hannover, Germany
[2]Leibniz Universität Hannover, Institute of Mechatronic Systems, 30823 Garbsen, Germany
[3]Department of Research Funding, Knowledge and Technology Transfer, Hannover Medical School, 30625 Hannover, Germany
[4]Institute of Medical Technology and Intelligent Systems, Hamburg University of Technology, 21073 Hamburg, Germany
[5]Centre for Image Guided Innovation and Therapeutic Intervention (CIGITI), The Hospital for Sick Children, Toronto, ON M5G 1X8, Canada
[6]Department of Mathematical and Computational Sciences, University of Toronto Mississauga, Mississauga, ON L5L 1C6, Canada

Corresponding author: Jacob F. Fast (fast.jacob@mh-hannover.de)

This work was supported by German Research Foundation under Grant KA 2975/6-1 and Grant PT 2/5-1.

ABSTRACT The laryngeal adductor reflex (LAR) is a vital reflex of the human larynx. LAR malfunctions may cause life-threatening aspiration events. An objective, noninvasive, and reproducible method for LAR assessment is still lacking. Stimulation of the larynx by droplet impact, termed *Microdroplet Impulse Testing of the LAR* (MIT-LAR), may remedy this situation. However, droplet instability and imprecise stimulus application thus far prevented MIT-LAR from gaining clinical relevance. We present a system comprising two alternative, custom-built stereo laryngoscopes, each offering a distinct set of properties, a droplet applicator module, and image/point cloud processing algorithms to enable a targeted, droplet-based LAR stimulation. Droplet impact site prediction (ISP) is achieved by droplet trajectory identification and spatial target reconstruction. The reconstruction and ISP accuracies were experimentally evaluated. Global spatial reconstruction errors at the glottal area of (0.3 ± 0.3) mm and (0.4 ± 0.3) mm and global ISP errors of (0.9 ± 0.6) mm and (1.3 ± 0.8) mm were found for a rod lens-based and an alternative, fiberoptic laryngoscope, respectively. In the case of the rod lens-based system, 96% of all observed ISP error values are inferior to 2 mm; a value of 80% was found with the fiberoptic assembly. This contribution represents an important step towards introducing a reproducible and objective LAR screening method into the clinical routine.

INDEX TERMS Aspiration, biomedical engineering, computer aided diagnosis, endoscopes, image guidance, laryngeal closure reflex, reflex latency, stereo endoscopy.

I. INTRODUCTION

According to the World Health Organization, lower respiratory infections were on the fourth position of the global list of cause-specific mortality in 2019 [1]. The laryngeal adductor reflex (LAR) is one of the vital protective reflexes of the human airways. This reflex leads to a fast closure of the larynx after impact of a foreign particle, thus preventing the external matter from entering the lower respiratory system. A dysfunctional LAR was found to correlate with a 6.8-fold increase of the risk of developing aspiration pneumonia [2]. Alterations in the sensorimotor pathways of the LAR may be related to conditions such as chronic cough, dysphonia, and laryngospasm [3]. Clinical research on the LAR has allowed to identify two, presumably independent, components of the LAR [4]. These components, commonly referred to as (L)R1 and (L)R2 [5], [6], are distinct with regard to their location and onset latency. In humans, a mean latency of 13 ms to 18 ms was observed for the early (L)R1 component; the (L)R2 component was reported to occur 51 ms to 70 ms

The associate editor coordinating the review of this manuscript and approving it for publication was Kumaradevan Punithakumar.

after electrical stimulation [5], [6]. (L)R1 is characterized as an adduction of the ipsilateral vocal fold with respect to the stimulation site, whereas (L)R2 is described as a bilateral response [5], [6]. In humans, the superoposterior supraglottis (glottis: area between the vocal folds) and the epiglottis seem to be especially sensitive for LAR stimulation [7]. A targeted and reproducible LAR stimulation may help to further elucidate the underlying neuronal pathways and the sensory receptors of the LAR, which are still subject of research [7]–[9]. Establishing normal LAR latency values may also enable the detection of neurological disorders [6]. Thus, a noninvasive, repeatable, and objective LAR screening procedure is highly desirable and the motivation of the present work.

A. SCREENING METHODS OF THE LARYNGEAL ADDUCTOR REFLEX

The previous paragraph demonstrates a strong clinical motivation for a noninvasive, quantitative, and reproducible evaluation of the LAR. Unfortunately, a method for laryngeal esthesiometry, yielding a standardized and quantitative LAR performance metric, has yet to be established. Still, several useful methods for LAR elicitation and subsequent reflex evaluation exist: In the context of reflex characterization and nerve monitoring, the stimulation of the target area with contact electrodes is a common approach [10]–[12]. Direct neural electrostimulation has also been used in animal and human studies [13]–[15]. A less invasive, tactile method was introduced by Aviv *et al.* [16]. This method, termed *Flexible Endoscopic Evaluation of Swallowing with Sensory Testing* (FEES-ST), relies on the application of air bursts [17], [18]. The invisibility of these bursts impedes the targeted and reproducible stimulation of a desired area. If a laryngoscope tip is used for mucosal stimulation, occlusions can hinder the assessment of the obtained frame sequences and a poor interrater reliability has recently been reported, even after specialized training of the rating personnel [19]. Manual, tactile stimulation of the laryngeal mucosa has also been reported to show a poor reproducibility [20]. Very recently, efforts have been made in an interesting study by Strohl *et al.* to deliver calibrated tactile stimuli with a buckling monofilament [21]. However, the precise LAR stimulation time, required for highly accurate reflex latency quantification, remains difficult to identify with such a manual approach, especially at low videolaryngoscopy frame rates [22].

B. MICRODROPLET IMPULSE TESTING OF THE LAR

To overcome the limitations of the available procedures described above, a new method termed *Microdroplet Impulse Testing of the LAR* (MIT-LAR) has recently been proposed by our research group [23]. After stimulation of the laryngeal mucosa by droplet impact at time t_1, the onset of the glottal reaction at time t_2 is recorded by a high-speed laryngoscope [23], [24]. The subsequent analysis of the acquired frame sequence allows to calculate the LAR onset latency

$$\Delta t_{\mathrm{LAR}} = t_2 - t_1, \tag{1}$$

a metric which has found widespread application for the quantitative characterization of the LAR [25]. In case of a very high LAR onset latency ($\Delta t_{\mathrm{LAR}} \gg 100\,\mathrm{ms}$), or if no reflexive motion is detected at all ($\Delta t_{\mathrm{LAR}} \to \infty$), a reflex pathology should be suspected and further actions, such as aspiration surveillance, can be undertaken to reduce the risk of developing pneumonia.

Thus far, MIT-LAR was limited by instabilities of the ejected fluid volume, which may lead to a stimulation of several areas in the larynx by multiple droplets. A suppression of laryngeal reflexes was reported in the case of repeated stimulation [24], [26], potentially leading to biased latency measurements. Hence, to enable unambiguous LAR latency quantification and intersubject reflex threshold comparisons, a single, stable droplet must be formed at a known and adjustable muzzle energy E_0, which depends on droplet mass m_{dr} and droplet muzzle velocity v_0 as follows:

$$E_0 = \frac{1}{2} m_{\mathrm{dr}} v_0^2. \tag{2}$$

In our recent work, we optimized the droplet applicator module of the original MIT-LAR system [23] to achieve stable droplet formation and muzzle energy variation [27].

C. RELATED WORK IN LARYNGOSCOPY

A prototype of a stereoscopic laryngoscope for laryngeal measurements, based on two analog cameras, was presented as early as 1983 [28]. Stevens *et al.* used a setup comprising a synthetic vocal fold model and a stereoscopic reference system to evaluate the spatial reconstruction performance of a stereo endoscope [29]. High-speed stereo endoscopic image data were used by Sommer *et al.* to estimate the spatial kinematics of the vocal folds [30]. Tokuda *et al.* employed stereo laryngoscopic reconstruction to obtain the topography of the glottal area [31]. Furthermore, stereo endoscopy was used by Nakano *et al.* to assess velopharyngeal movements [32]. George *et al.* and Semmler *et al.* were able to explore the oscillation cycle of the human vocal folds by combining a laser projector and a monoscopic laryngoscope [33], [34]. Bohr *et al.* discussed using several other spatial reconstruction methods, such as *structured light* or *structure from motion*, in laryngoscopy [35]. Very recently, Kim *et al.* proposed a smartphone-based system for monoscopic high-speed laryngoscopy [36].

D. PURPOSE OF THIS CONTRIBUTION

To avoid LAR assessment bias, which may be caused by a habituation of the stimulated region to the stimulus [24], [26], [37], the desired target zone must reliably be impacted by the first droplet ejected by the MIT-LAR system. Hence, the lack of a visual guidance system constitutes a fundamental limitation of MIT-LAR. A guidance system would also reduce the risk of adverse events such as an iatrogenically provoked aspiration, caused by incidentally shooting a droplet into the trachea.

The current state of research lacks an endoscopic system which is adapted to the specific requirements of the droplet-based, targeted stimulation of the LAR. Hence, in the present work, we substantially extend our previous contribution [38] by proposing an integrated system comprising a fiberoptic, distally bendable, stereoscopic laryngoscope and a physically motivated approach for droplet trajectory approximation to enable a highly accurate droplet impact site prediction (ISP). We compare the stereo reconstruction accuracy of the novel, fiberoptic system with a rigid, rod lens-based stereo laryngoscope. The ISP accuracy at different droplet muzzle energies and endoscope shaft roll angles is systematically quantified, based on a large number of laboratory experiments. The proposed system is not limited to LAR stimulation but could also be used in an adapted form in areas such as drug delivery and chemical cauterization [39].

II. MATERIALS AND METHODS

Part II-A describes the design, manufacturing, and assembly of two different stereo laryngoscopes according to the specific requirements of a targeted MIT-LAR procedure. These optical systems were combined with an endoscopic droplet applicator module, characterized previously [27], incorporating a solenoid valve and an external fluid system pressure control circuit for stable and reproducible droplet formation. The components of the setup are shown in Fig. 1.

A physically motivated model of the droplet flight is derived in Sec. II-B. The stereoscopic image data acquired by the stereo laryngoscopes are algorithmically processed, enabling droplet trajectory identification and stimulation site prediction for a safe and targeted MIT-LAR procedure as detailed in Sec. II-C and II-D.

A. DESIGN, MANUFACTURING, AND ASSEMBLY OF TWO STEREOSCOPIC LARYNGOSCOPES

In the present work, the stereoscopic reconstruction of the laryngeal space is used to enable stimulation site prediction and visualization to assist the clinician during a MIT-LAR procedure. As no stereoscopic high-speed laryngoscope with an appropriate direction of view for MIT-LAR (70° to 90°) is commercially available, we designed and assembled two alternative stereoscopic devices, providing distinct sets of optical and geometrical properties. The first device is a rigid, rod lens-based system, which was presented previously in an early design stage without internal illumination and roll angle sensing capabilities [38]. A rod lens-based design offers efficient light transmission, which is especially desirable for high-speed imaging situations with low exposure times, and a theoretically diffraction-limited spatial resolution. However, the rigid configuration of such a system does not allow the construction of a laryngoscope shaft which is adapted to the curved anatomy of the human larynx.

Hence, we designed and assembled a second system, incorporating two flexible fiber bundles for light transmission. This fiberoptic stereo laryngoscope enables the manual bending of its distal shaft elements with the goal of increasing the anatomical compliance. The final laryngoscope designs were obtained by iterative rapid prototyping using a high-precision stereolithography system (*Objet350 Connex3*, Stratasys, Ltd., Eden Prairie, USA). Both laryngoscopes were sequentially evaluated in association with the high-speed camera *Os*^{v3} *7-S1* (IDT Vision, Pasadena, USA), which enables high-speed sequence recording at a maximum frame rate of 1350 Hz and a resolution of 1920 × 1280 pixels.

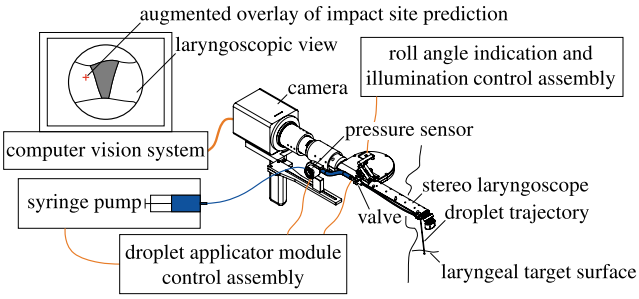


FIGURE 1. Schematic overview of system components for a safe and targeted MIT-LAR procedure. Control/power/data connections shown in orange, fluid ducts shown in blue.

1) ROD LENS SYSTEM

The first of the two custom-built stereoscopic laryngoscopes is based on two identical rod lens assemblies (Blazejewski MEDI TECH GmbH, Sexau, Germany) with an outer diameter of 5.5 mm. The rod lens systems contain distal prisms for deflection of the incoming light, yielding a direction of view of 90° and a field of view of approximately 70°, which is adequate for transoral laryngoscopy. At their proximal end, an intermediate image is formed by each of the two individual rod lens assemblies, which are horizontally and axially aligned. Both images are magnified and projected onto the CMOS chip of the high-speed camera by a pair of aspherized achromats (Edmund Optics, Barrington, USA) with focal lengths of 30 mm and 40 mm, respectively. The assembled, rod lens-based stereo laryngoscope is depicted in Fig. 2. The optical and geometrical properties of this system are listed in Table 1. In the table, the stereo base corresponds to the distance between the two monoscopic optical systems. The raw circular images are the two individual monoscopic images formed side-by-side on the CMOS sensor. The extracted subimages are the inscribed squares of these circular images, which were used for further processing.

2) FIBEROPTIC SYSTEM

The second stereo laryngoscope enables a more anatomically adapted, curved shaft design by employing fiberoptic light transmission and a mechanism for manual distal bending. Two identical fiber bundles (*B424943*, SCHOTT AG, Mainz, Germany) with an outer bundle diameter of 1.4 mm were integrated, each containing approximately 18 000 single fibers. The bundles are equipped with objectives (XION GmbH, Berlin, Germany) providing a field of view of approximately 80°. The intermediate images formed by the fiber

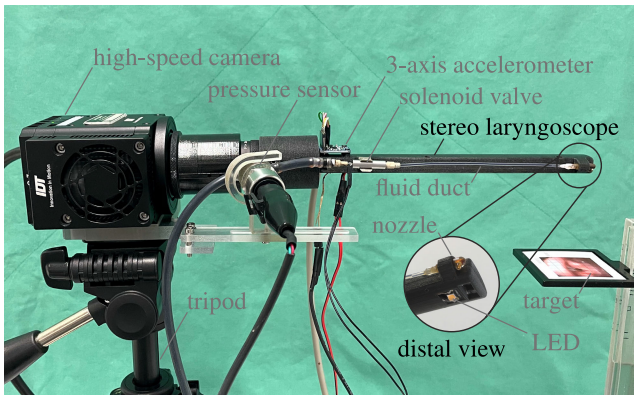


FIGURE 2. Components of the rod lens-based stereo laryngoscope for targeted MIT-LAR droplet application. Components shared between systems indicated in gray.

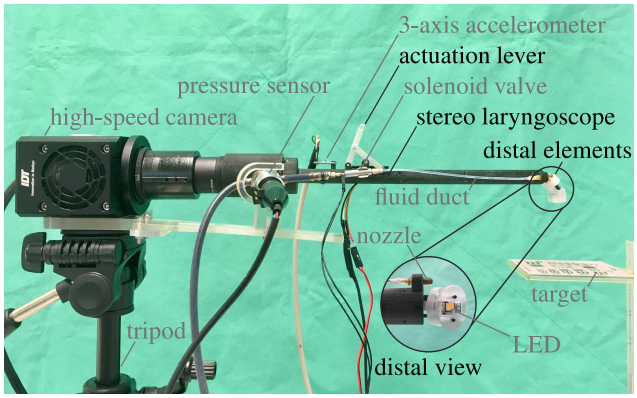


FIGURE 3. Components of the bendable, fiberoptic stereo laryngoscope for targeted MIT-LAR droplet application. Components shared between systems indicated in gray.

bundles are magnified and projected onto the CMOS chip by the same pair of aspherized achromats used with the rod lens-based system.

The bending function was designed to enable the visualization of the periglottal target region while avoiding contact with the posterior pharyngeal wall, as this may lead to an unintended stimulation of the pharyngeal reflex. We chose a kinematic approach inspired by the recent work of Gu *et al.* [40] for the bending mechanism. Two parallel polyamide filaments (∅ 0.7 mm) enable a manual, stepless actuation of the two distal elements by a hand lever located in the proximal laryngoscope shaft. The elements are connected and stabilized by rotatory joints and contain guide channels for the fiber bundles. An easily removable clamping pin at the hand lever allows locking the system in the fully bent state. The components of the fiberoptic stereo laryngoscope are shown in Fig. 3. The low diameter of the fiber bundles allowed us to reduce the laryngoscope shaft width, including the droplet applicator module, by 1.5 mm (final width: 27.5 mm) compared to the rod lens system, while increasing the stereo base by 49 %. The optical and geometrical parameters of the fiberoptic system are provided in Table 1.

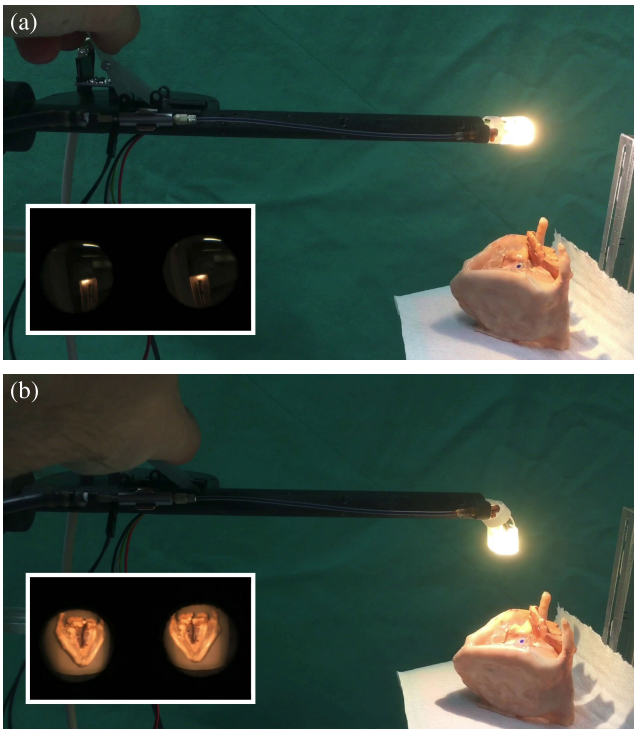


FIGURE 4. Still frames extracted from the demonstration video of the manual actuation of the fiberoptic stereo laryngoscope (provided in [41]). The stereoscopic view of a larynx phantom is shown in the bottom left for each frame. (a) minimum bending angle (0°), (b) maximum bending angle (90°).

TABLE 1. Optical, geometrical, and gravimetric parameters of the stereo laryngoscopes. Stereo base: distance between individual optical channels; raw images: circular images formed on camera sensor; extracted subimages: inscribed squares of raw images, which were used for further processing. Shaft dimensions include droplet applicator module.

parameter	rod lens system	fiberoptic system
stereo base	6.7 mm	10.0 mm
monoscopic field of view	≈ 70°	≈ 80°
∅ of raw images	770 px	220 px
size of extracted subimages	544 × 544 px	156 × 156 px
direction of view	90° (fixed)	0° to 90° (variable)
total shaft height × width	12 × 29 mm	12 × 27.5 mm
total system weight	≈ 1.2 kg	

The manually controlled, continuous bending of the distal tip of the fiber-based MIT-LAR laryngoscope is demonstrated in a supplementary video sequence represented by the still frames shown in Fig. 4. The laryngoscopic view of a phantom of the human larynx is overlaid in the bottom left for each frame. The video sequence is provided at *IEEE DataPort* [41].

3) DROPLET APPLICATOR MODULE

A droplet applicator module, enabling the reproducible formation of a single stimulation droplet at a known muzzle energy, which was developed and characterized in a previous contribution [27], was attached to the side of the stereo laryngoscopes. The stability and visibility of the formed droplets

was enhanced by addition of sodium chloride and the non-toxic food colorant E 150c [27].

4) ROLL ANGLE MONITORING AND INTERNAL ILLUMINATION

The orientation of the droplet applicator module with respect to the direction of gravity influences the droplet's flight trajectory. Hence, the angle α_0 between the direction of gravity and the droplet's muzzle velocity vector $\mathbf{v_0} \in \mathbb{R}^3$ must be known during the droplet trajectory identification and ISP phases. A 3-axis accelerometer (*LIS3DH*, STMicroelectronics N.V., Amsterdam, The Netherlands) was thus mounted on the shaft of the stereo laryngoscopes to obtain measurements of α_0 during the quasi-static handling of the system in a MIT-LAR procedure. As the longitudinal nozzle axis is perpendicular to the (horizontal) laryngoscope shaft in both systems, the droplet shooting angle α_0 is approximately equal to the laryngoscope shaft roll angle of both systems.

A high-power light-emitting diode (LED, *XHP35*, high density version, Cree, Inc., Research Triangle Park, USA) was added to the stereo laryngoscopes to provide internal illumination. A platinum resistance thermometer was attached to the back of the LED to monitor its thermal strain. The hardware components required for roll angle measurement and illumination control were grouped together and controlled by a single-board microcontroller (*Arduino Uno*, Arduino S.r.l., Strambino, Italy). This modular approach allows using the peripheral components interchangeably with both optical systems.

B. DROPLET TRAJECTORY MODELING APPROACH

The expected droplet trajectory must be known to enable droplet ISP and visualization during a MIT-LAR procedure. In this section, the physical background of a droplet's trajectory after ejection with muzzle velocity $\mathbf{v_0}$ is studied and a closed-form trajectory model is derived.

1) BALANCE OF EXTERNAL FORCES

Neglecting all external forces acting on the droplet after ejection would lead to a linear trajectory model T_1, following the droplet muzzle velocity vector $\mathbf{v_0}$ as indicated in Fig. 5. However, the assumption of a linear trajectory may cause important modeling errors [42], except for the singular case of a muzzle velocity vector $\mathbf{v_0}$ which is exactly collinear to the gravitation vector $\mathbf{g}$. Hence, a total of four external forces, acting on the droplet after ejection from the droplet applicator module, are considered here to derive a more accurate model of the droplet trajectory: the force $\mathbf{F_G} \in \mathbb{R}^3$ due to gravity, the force $\mathbf{F_B} \in \mathbb{R}^3$ of buoyancy, the force $\mathbf{F_D} \in \mathbb{R}^3$ of air drag, and the force $\mathbf{F_R} \in \mathbb{R}^3$ due to patient respiration. These forces acting on the droplet are shown in the free-body diagram of Fig. 6.

Whereas the forces of gravity $\mathbf{F_G} = m_{\mathrm{dr}}\mathbf{g}$ and buoyancy $\mathbf{F_B} = -\rho_{\mathrm{air}}/\rho_{\mathrm{dr}}\mathbf{F_G}$, with respective air and droplet densities ρ_{air} and ρ_{dr}, can be considered constant over the course of a droplet flight (no fluid evaporation due to short flight time and low vapor pressure of droplet fluid), the drag forces $\mathbf{F_D}(t)$ and $\mathbf{F_R}(t)$ are functions of the instantaneous droplet velocity $\mathbf{v_{dr}}(t)$ and the patient's respirational cycle, respectively.

Due to a low density ratio $\rho_{\mathrm{air}}/\rho_{\mathrm{dr}} \approx 1 \times 10^{-3}$, the force of buoyancy $\mathbf{F_B}$ is neglected in the scope of this work. Furthermore, we assume that the patient's respiratory cycle is monitored, e.g., by using an elongation-sensing abdominal belt [39], and droplet shooting is only allowed at instants of

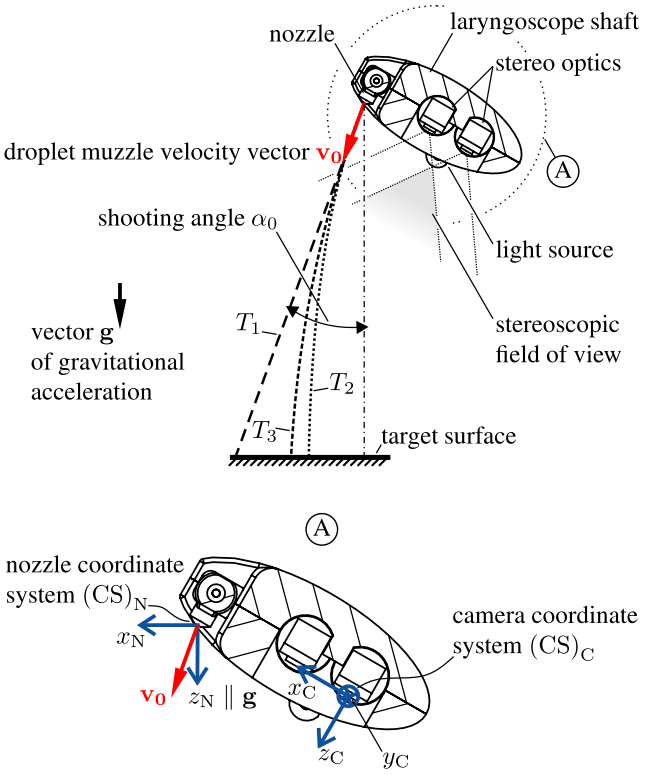


FIGURE 5. Parameters influencing the droplet's trajectory with gravitation vector g, muzzle velocity vector $\mathbf{v_0}$, and droplet shooting angle α_0. T_1: linear extension of $\mathbf{v_0}$, T_2: parabolic model without air drag, T_3: ballistic model with air drag. Coordinate systems shown in magnified view Ⓐ.

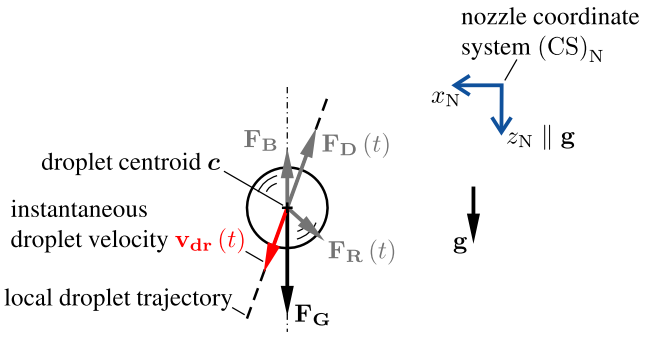


FIGURE 6. Free-body diagram of a droplet in midair with constant forces $\mathbf{F_G}$ and $\mathbf{F_B}$ of gravity and buoyancy, respectively, and drag forces $\mathbf{F_R}(t)$ and $\mathbf{F_D}(t)$ caused by patient respiration and droplet motion at velocity $\mathbf{v_{dr}}(t)$, respectively. Negligible forces shown in gray. Vectors not true-to-scale.

minimum breathing activity. Under this assumption, $\mathbf{F_R}(t)$ is also considered negligible. The magnitude F_D of the force of air drag caused by the droplet motion is in the range of 10 % of the magnitude F_G of the force of gravity [27]. Therefore, only the influence of the droplet's gravity is considered in the following.

2) DERIVATION OF TRAJECTORY MODEL

By taking only the constant force of gravity $\mathbf{F_G}$ into account, the parabolic trajectory model T_2, sketched in Fig. 5, is obtained. The well-known equation of motion for this example of a free fall with initial velocity can be given in closed form in the Cartesian nozzle coordinate system $(\mathrm{CS})_\mathrm{N}$ with axes x_N and $z_\mathrm{N} \parallel \mathbf{g}$ as indicated in detail Ⓐ of Fig. 5. The droplet coordinates $x_{\mathrm{N,dr}}$ and $z_{\mathrm{N,dr}}$ in $(\mathrm{CS})_\mathrm{N}$ are coupled as follows with droplet muzzle velocity magnitude v_0 and shooting angle α_0:

$$z_{\mathrm{N,dr}}\left(x_{\mathrm{N,dr}}\right) = \frac{g}{2\left(v_0 \sin\left(\alpha_0\right)\right)^2} x_{\mathrm{N,dr}}^2 + \frac{x_{\mathrm{N,dr}}}{\tan\left(\alpha_0\right)}. \tag{3}$$

This parabolic approach has the advantage of being available as a closed-form solution, whereas a ballistic approach T_3 as shown in Fig. 5, incorporating the (minor) influence of air drag, cannot readily be provided as an analytical equation for ISP. Thus, the parabolic approach is retained for droplet trajectory approximation.

C. DROPLET TRAJECTORY IDENTIFICATION

This section describes the algorithm for the approximation of the spatial trajectory of a droplet, which is assumed to be highly reproducible after droplet applicator module optimization [27], to enable the visualization of the predicted droplet impact site for the clinician performing MIT-LAR. Fig. 7 gives a high-level overview of the proposed method for LAR stimulation site prediction, including a system calibration phase (shown in blue) and a MIT-LAR application phase (shown in orange).

The calibration phase comprises traditional stereo calibration of the laryngoscopic system, yielding the intrinsic and extrinsic parameters of the stereo setup. An asymmetrical circle grid with an intercircle distance of 2.0 mm and a circle diameter of 1.5 mm is used as calibration pattern. The calibration is performed with frame sequences showing the calibration pattern in different poses to obtain a high accuracy. The reprojection error is calculated to confirm successful calibration. After acquisition of 10 high-speed sequences showing the droplet flight at identical system conditions, the spatial positions of the droplet are stereoscopically triangulated in the Cartesian coordinate system $(\mathrm{CS})_\mathrm{C}$ of the left virtual camera of the stereo laryngoscope (see Fig. 5). Based on these spatial sampling points, a parabolic trajectory approximation is calculated and stored. The application phase begins with the stereoscopic reconstruction of the laryngeal target structures. Based on the resulting point cloud and the previously identified droplet trajectory, an impact site can be predicted and visualized in the laryngoscopic image to guide the MIT-LAR operator.

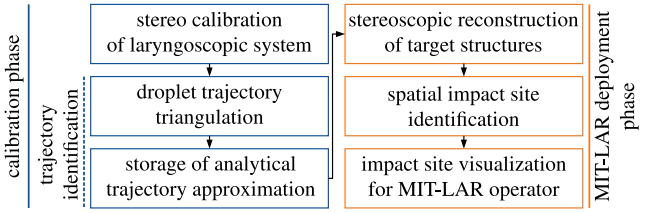


FIGURE 7. High-level overview of workflow for LAR stimulation site prediction.

1) STEREOSCOPIC DROPLET DETECTION

The stereoscopic laryngoscopes enable a straightforward, spatial triangulation of the droplet centroid $\mathbf{c} \in \mathbb{R}^3$ along its trajectory at a high temporal resolution. The following method is applied to locate the droplet in the individual frames of a high-speed sequence showing its trajectory. Spatial sampling points are subsequently derived. All operations are executed individually on the left and right subimages.

First, a simple background subtraction is performed, using a frame without droplet as background frame. Second, the droplet is segmented in the resulting image by fixed-level binary thresholding. Third, a blob detector is used to locate the centroid of the segmented area. Based on the known intrinsic and extrinsic stereo camera parameters and the droplet centroid location in the rectified left and right subimages, $\mathbf{c}$ is then spatially triangulated in $(\mathrm{CS})_\mathrm{C}$. The method is repeated for all available frames, yielding a set S_samp of spatial sampling points. Image pairs in which more than one blob per image is detected, or in which the epipolar constraint of the identified droplet centroid positions is violated, are discarded.

2) TRAJECTORY PLANE APPROXIMATION

The proposed method for droplet trajectory approximation is represented schematically in Fig. 8 to guide the reader through the following sections. Based on the set S_samp of spatial sampling points, acquired as described above, an analytical approximation of the droplet trajectory can be identified. As all forces acting on the droplet in midair are coplanar, the trajectory corresponds to a planar curve in the Cartesian space spanned by $(\mathrm{CS})_\mathrm{C}$ [43]. Thus, in a first step, the droplet trajectory plane is approximated by a fit plane Π, shown in Fig. 8, which minimizes the sum of the orthogonal distances to the available sampling points. As a high-speed camera is used to record the droplet along its trajectory, more than three spatial, noncollinear sampling points describing the droplet trajectory are generally available, leading to an overdetermined plane identification problem. Furthermore, the sampling points are obtained by optical triangulation and are thus affected by noise. Hence, singular-value decomposition (SVD) is used to identify the plane Π: First, the geometrical center $\mathbf{c_{geo}}$ of the set S_samp of spatial sampling points is determined. The center coordinates are subsequently

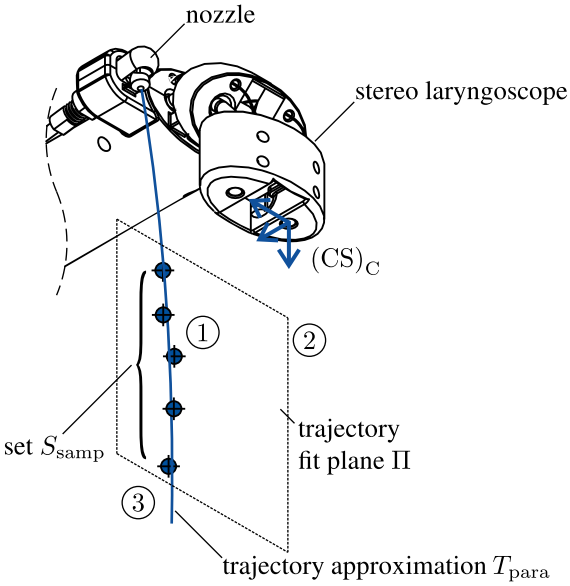


FIGURE 8. Schematic representation of (offline) droplet trajectory identification in left virtual camera coordinate system $(CS)_C$ with ① triangulation of sampling point set S_{samp}, ② calculation of trajectory fit plane Π, ③ identification of trajectory fit function T_{para}.

subtracted from each individual sampling point to obtain the set S^*_{samp} of n zero-mean sampling points. The SVD of the $3 \times n$ matrix containing the shifted sampling points yields a normal vector $\mathbf{n_\Pi}$ of the approximated trajectory plane Π. For straightforward spatial distance calculations during MIT-LAR, the Hesse normal form of Π is preferable. In this form, the normal vector $\mathbf{n_\Pi}$ must be normalized and pointing away from the origin of $(CS)_C$. Hence, a new unit vector $\hat{\mathbf{n}}_\Pi$, satisfying this condition, is calculated. The Hesse normal form of Π can now be formulated with position vector $\mathbf{r_\Pi}$ of an arbitrary point on Π and scalar $q = \mathbf{c_{geo}} \cdot \hat{\mathbf{n}}_\Pi$:

$$\Pi = \left\{ \mathbf{r_\Pi} \in \mathbb{R}^3 \mid \mathbf{r_\Pi} \cdot \hat{\mathbf{n}}_\Pi - q = 0 \right\}. \tag{4}$$

In a final step, q and $\hat{\mathbf{n}}_\Pi$ are stored for efficient ISP during the MIT-LAR application phase, as the droplet impact site must be close to Π (see Sec. II-D).

3) DROPLET TRAJECTORY APPROXIMATION

In general, orthogonal distance fitting of a geometrical object which is more complex than a line/plane or a circle/sphere to a set of spatial sampling points is an analytically and computationally challenging task [44]. We propose a parabolic trajectory approximation T_{para} in parametric form, allowing for a simple, yet accurate droplet trajectory identification method, which can directly be executed in $(CS)_C$. We further propose to take advantage of the time stamps t_i, associated with each triangulated droplet position along its trajectory, to enable a geometrical fitting of T_{para} to the spatial sampling points while avoiding nonlinear optimization. The time stamps t_i are provided with very high precision, as the camera frame rate is certified to vary by less than $\pm$ 0.15 Hz at the frame rate of the

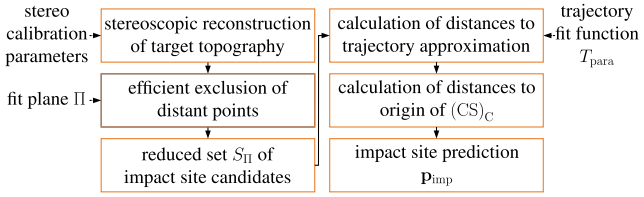


FIGURE 9. Workflow for laryngoscopic droplet impact site prediction.

droplet flight sequence recordings (set to 1000 Hz). At least three sampling points along the droplet trajectory must be available to identify three vectors $\mathbf{f}_i \in \mathbb{R}^3$ with $i \in \{0, 1, 2\}$, defining the position vector $\mathbf{r_{para}}(t)$ on T_{para} at a given time t:

$$T_{para} = \left\{ \mathbf{r_{para}}(t) \in \mathbb{R}^3 \mid \mathbf{r_{para}}(t) = \mathbf{f_0} + t\mathbf{f_1} + t^2\mathbf{f_2} \right\}. \tag{5}$$

This approach corresponds to the three-dimensional generalization of (3) and allows the analytical modeling of any parabolic trajectory in $(CS)_C$. As more than three sampling points are generally available, a set of least-squares solution vectors $\hat{\mathbf{f}}_i$ of the resulting overdetermined system of linear equations, built with the set S_{samp} of sampling points and the associated time stamps t_i (starting at $t = 0$), is subsequently identified using an SVD approach. These vectors are then stored to enable online ISP during the MIT-LAR application phase.

D. DROPLET IMPACT SITE PREDICTION

After the plane and the geometry of the droplet's trajectory have been estimated in an offline calibration phase, the droplet's impact site can now be predicted online. The proposed ISP workflow for a targeted MIT-LAR procedure is depicted in Fig. 9. This universal method can be applied with any stereoscopic system containing any kind of applicator module enabling a reproducible trajectory of the stimulation projectile, provided that the relative poses of this applicator module and the optical parts of the system remain constant. The individual steps of the method are presented in Fig. 10 and described in the following.

1) STEREOSCOPIC RECONSTRUCTION OF TARGET TOPOGRAPHY

The left and right subimages are extracted from the raw image data. The intrinsic and extrinsic parameters of the two different stereo laryngoscopes, previously identified by stereo camera calibration as described in Sec. II-C, enable a true-to-scale, spatial reconstruction of the laryngeal target structures. The epipolar geometry constraint is exploited by image rectification to accelerate the stereoscopic correspondence search process. As brightness differences are likely to be present due to the manual assembly and adjustment of the two stereo laryngoscopes used in the present work, a variant of the well-known semi-global block matching (SGBM) algorithm proposed by Hirschmüller [45], operating on census transformed image data, is used for the laryngoscopic

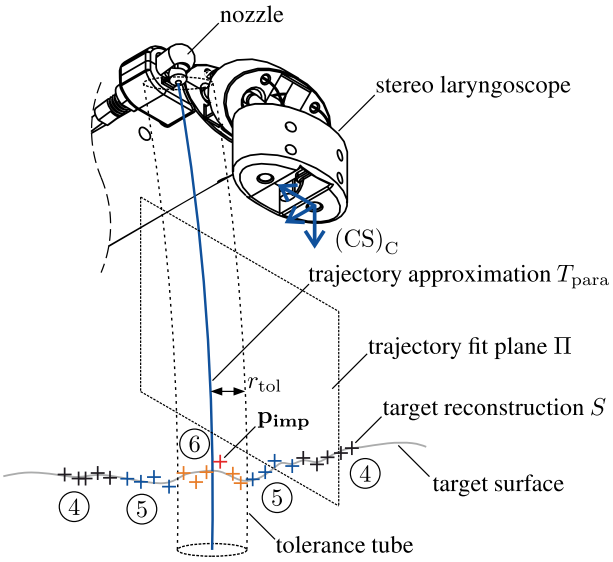


FIGURE 10. Schematic representation of (online) droplet impact site prediction procedure with ④ efficient elimination of impossible impact sites (indicated in black) from target point cloud S using fit plane Π, ⑤ elimination of residual candidate points (indicated in blue) outside tolerance tube with radius r_{tol}, ⑥ identification of final impact site prediction $\mathbf{p}_{imp}$ (indicated in red) among set of points located inside tolerance tube (indicated in orange).

reconstruction of the target. A center-symmetric census transform kernel is applied. The resulting disparity map is post-processed in a weighted least squares (WLS) filtering step [46] before conversion into the final, spatial point set S containing j individual points.

2) EFFICIENT EXCLUSION OF DISTANT POINTS

To achieve a low ISP computation time in the time-critical *in vivo* application phase, the planar nature of the droplet trajectory is exploited. The predicted impact site must be close to the trajectory plane Π. As the Hesse normal form of Π, given in (4), is available, the distance $d(\Pi, \mathbf{p}_i)$ between a given point $\mathbf{p}_i \in S$ and Π can readily be calculated:

$$d(\Pi, \mathbf{p}_i) = |\mathbf{p}_i \cdot \hat{\mathbf{n}}_{\mathbf{\Pi}} - q|. \tag{6}$$

Only the $k \leq j$ points $\mathbf{p}_i \in S$ closer than an arbitrary threshold $\tau_\Pi \in \mathbb{R}$ to Π are retained in the set S_Π of potential impact site candidates:

$$S_\Pi = \{\mathbf{p}_i \in S \mid d(\Pi, \mathbf{p}_i) \leq \tau_\Pi\}. \tag{7}$$

3) PREDICTION OF DROPLET IMPACT SITE

The orthogonal distance $d_\perp(T_{para}, \mathbf{p}_i)$ between a point $\mathbf{p}_i \in S_\Pi$ and the parabolic trajectory approximation T_{para} can be calculated as follows: In a first step, the parameter value $t^* > 0$, yielding the point $\mathbf{r}_{\mathbf{para}}(t^*)$ on T_{para} which minimizes the distance to $\mathbf{p}_i$, is identified. In a second step, the orthogonal distance $d_\perp(T_{para}, \mathbf{p}_i)$ is readily obtained:

$$d_\perp(T_{para}, \mathbf{p}_i) = |\mathbf{r}_{\mathbf{para}}(t^*) - \mathbf{p}_i|. \tag{8}$$

The droplet cannot penetrate the target surface represented by the reduced impact site candidate set S_Π. To make sure that a physically possible ISP is obtained in the presence of measurement noise in S_Π, the point $\mathbf{p}_{imp}$ belonging to the subset $S_{para} \subseteq S_\Pi$ of $\ell \leq k$ points located inside of a tolerance tube with arbitrary radius $r_{tol} \in \mathbb{R}$, centered around T_{para}, with the lowest distance $z_{C,i}$ to the origin of $(CS)_C$ in direction of the z_C-axis among S_{para}, is retained as most probable impact site as depicted in Fig. 10:

$$\mathbf{p}_{imp} = \underset{\mathbf{p}_i \in S_{para}}{\arg\min}\, z_{C,i}. \tag{9}$$

The predicted impact site $\mathbf{p}_{imp}$ is then reprojected into the left laryngoscopic image and visualized as augmented reality crosshairs to guide the operator of the MIT-LAR system.

The algorithms developed in this work can be executed on a personal computer running a recent version of the *Ubuntu* operating system without requiring a dedicated GPU. Image processing was performed on a standard desktop PC (*Optiplex 7010*, Dell Technologies Inc., Round Rock, USA) with an 8-core CPU (*i7-3770* @ 3.4 GHz, Intel Corp., Santa Clara, USA) and 8 GB of RAM. Similar hardware is readily available in many otorhinolaryngology departments. The open-source image processing library *OpenCV* [47] (version 3.3.1) and the *Point Cloud Library* [48] (version 1.11.1-dev) were used with the Linux distribution *Ubuntu* (version 18.04.3 LTS) to implement the methods described in this contribution. The source code developed and used in the present work can be found at https://github.com/JFast/MIT-LAR_IEEE_Access.

III. EXPERIMENTAL EVALUATION

This chapter details the methods for an experimental performance assessment of the proposed system for the targeted, droplet-based stimulation of the LAR.

A. STEREO RECONSTRUCTION ACCURACY

To evaluate the stereo reconstruction accuracy that can be achieved with both stereo laryngoscopes, a three-dimensional phantom of the human larynx, based on publicly available MRI data [49], was stereolithographically fabricated and used as a surrogate test object. A skin-colored silicone layer was applied to obtain a realistic surface texture. Four circular marker points ($\varnothing \approx 2$ mm) were applied in the glottal area of the phantom. A textured, high-density ground truth (GT) point cloud reconstruction of this larynx phantom was acquired at a resolution of approximately 60 points per mm^2 with a laser scanning device (*Desktop 3D Scanner 2020i*, NextEngine, Inc., Santa Monica, USA) and the associated reconstruction software (*ScanStudio PRO 2.0.2*, NextEngine, Inc., Santa Monica, USA). Spatial reconstructions of the larynx phantom were subsequently generated with both stereo laryngoscopes with the same reconstruction algorithm which was applied for ISP (see Sec. II-D1).

The coordinates of the four marker points were manually identified in both point clouds using *MeshLab 64bit_fp*

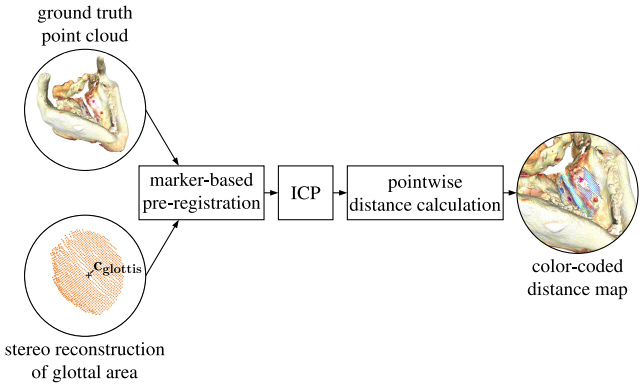


FIGURE 11. Workflow for stereo reconstruction accuracy quantification after marker-based pre-alignment and registration refinement by an iterative closest point (ICP) approach.

v2020.12 (Visual Computing Lab, ISTI-CNR, Pisa, Italy). The coordinates of the glottis center $\mathbf{c}_{\mathbf{glottis}}$ were manually determined in the stereoscopically reconstructed (SR) point clouds. Six subsets S_1 to S_6 of the SR point clouds, delimited by spheres centered around $\mathbf{c}_{\mathbf{glottis}}$ and with diameters of $\{100, 120, 140, 160, 180, 200\}$ % of the approximate vocal fold length $\ell_{\mathrm{vf}} = 16\,\mathrm{mm}$ of the phantom, were retained for the quantification of the reconstruction accuracy with respect to the GT point cloud for increasing sizes of the exploited periglottal zone. In the next step, a marker-based, rigid pre-registration of S_1 to S_6 and the GT point cloud was performed. Iterative closest point (ICP) registration was subsequently applied to obtain the final alignment.

The distances of all points of the six SR point clouds S_1 to S_6 to the nearest neighbors of the high-resolution GT point cloud were sequentially calculated and stored for statistical evaluation. The SR point clouds were color-coded from dark blue to bright red, according to the local distance to the GT point cloud. Finally, the intracloud distances to the nearest neighbors were calculated for the largest subset S_6 for both stereo laryngoscopes to quantitatively evaluate the density of the SR point cloud. The workflow for the quantitative evaluation of stereo reconstruction performance is shown in Fig. 11.

B. TRAJECTORY APPROXIMATION AND IMPACT SITE PREDICTION ACCURACY

To obtain the droplet trajectory fit plane Π and the trajectory approximation T_{para} required for ISP, a total of 10 droplet flight sequences were recorded at constant system conditions at a frame rate of 1000 Hz. This procedure was followed for three different droplet muzzle velocities $v_0 \in V_0$ with $V_0 \approx \{0.5, 0.6, 0.7\}\,\mathrm{m\,s^{-1}}$ (termed low/medium/high in the following) at two different endoscope shaft roll angles α_0 of 0° and 45°, respectively. Hence, a set of 60 trajectory frame sequences was obtained for each stereo laryngoscope, each sequence yielding 41 sampling points on average. The bending angle of the fiberoptic system was fixed at 90° as required in the planned intralaryngeal application. No target was used in this calibration step to enable the observation of a large trajectory segment. The plane and the geometry of the average droplet trajectory were approximated for each of the six system configurations by calculating the global fit plane Π and the global fit function T_{para}, based on all available sampling points obtained by stereoscopic triangulation (see Sec. II-C1) in the 10 recordings.

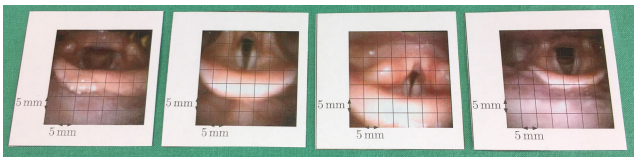


FIGURE 12. Planar surrogate targets with scale indication grid for impact site prediction accuracy measurements.

1) GOODNESS OF FIT OF TRAJECTORY MODELS

To quantify the accuracy of the identified fit planes Π and parabolic approximations T_{para}, the orthogonal distances of each individual sampling point to Π and T_{para} were calculated for each system configuration according to (6) and (8).

2) IMPACT SITE PREDICTION ACCURACY

A total of 120 experiments were performed with each of the two stereo laryngoscopes to assess the ISP accuracy for MIT-LAR in a realistic scenario for the six combinations of α_0 and v_0 listed above. Planar surrogate targets, consisting of four laryngoscopic views, were used. The use of planar targets allowed the straightforward quantification of the ISP accuracy. The targets are shown in Fig. 12. The laryngoscopic views were recorded in four different subjects with the original, monoscopic MIT-LAR system in previous clinical studies [24], [50]. Five copies of each target were prepared using a high-resolution color laser printer (MP C5504exSP, Ricoh Co., Ltd., Tokyo, Japan), yielding a total of 20 targets for each system configuration. The targets were consecutively positioned perpendicularly to the laryngoscopic direction of view at a distance of approximately 66 mm and 51 mm for the rod lens-based system and the fiberoptic system, respectively. These values correspond to the expected average distances between the laryngoscope tip and the glottal area [51]. Based on the trajectory approximation T_{para} and the stereoscopic reconstruction of the target, the predicted droplet impact site was calculated and superimposed onto the undistorted and rectified left laryngoscopic image. The actual impact site after droplet ejection was located in the same image by blob detection and manual verification. The targets included a rectangular grid with known dimensions, thus allowing image scale calculation and direct quantification of the prediction accuracy, corresponding to the distance between the predicted and actual droplet impact sites.

IV. RESULTS

This section presents the results of the stereo reconstruction, trajectory approximation, and ISP accuracy evaluation. In all box plots, whiskers follow the Tukey convention.

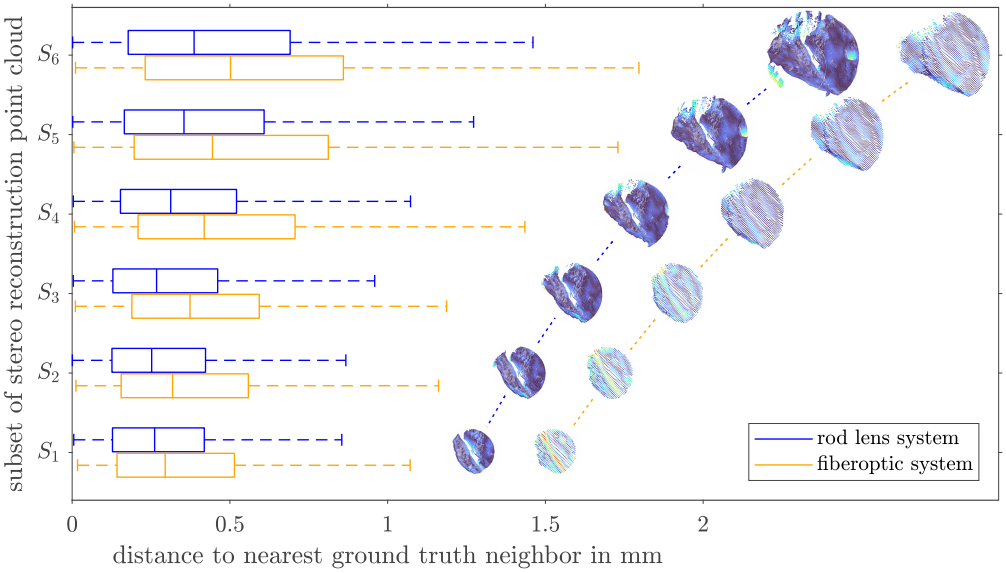


FIGURE 13. Accuracy of stereoscopically obtained point clouds with respect to the ground truth (GT) point cloud representing a larynx phantom (lower is better). The size of the evaluated periglottal zone increases from S_1 to S_6. The color-coded point clouds show the local distance to the nearest GT point (distance increasing from dark blue to bright red). Outliers (values beyond whiskers, ≤ 6.4% of sample size) not shown for clarity.

A. STEREO RECONSTRUCTION ACCURACY

The accuracy of the SR point cloud representations of the larynx phantom with respect to the GT point cloud is shown in Fig. 13 for six spherical zones S_1 to S_6 with increasing size, centered around the glottis as described in Sec. III-A. As the size of the evaluation zone increases, the median of the distances to the nearest neighbor in the GT point cloud also slightly increases from 0.26 mm to 0.39 mm for the rod lens system and from 0.29 mm to 0.50 mm for the fiberoptic system. Notable differences regarding the point density of the SR clouds can be observed between the two systems. Global intracloud distances of (0.12 ± 0.05) mm and (0.48 ± 0.13) mm were obtained with the rod lens-based and the fiberoptic stereo laryngoscope, respectively. The GT and SR point clouds are provided as supporting material in an *IEEE DataPort* dataset [41]. The mean/minimum/maximum values and the sample standard deviations of the distances from the SR point cloud subsets to the nearest neighbors of the GT point cloud and of the intracloud distances (calculated for SR point cloud density evaluation) are provided for both systems in Table 2.

TABLE 2. Mean/minimum/maximum values and sample standard deviations s of the distances from the SR point cloud subsets S_1 to S_6 to the nearest neighbors of the GT point cloud for both stereo laryngoscopes. n_{total} represents the number of available points in each SR point cloud subset. ICD: intracloud distance of S_6.

rod lens-based system				
subset	mean dist. $\pm$ s [a]	min. dist.	max. dist.	n_{total}
S_1	(0.31 ± 0.25) mm	0.01 mm	4.82 mm	1.3×10^4
S_2	(0.32 ± 0.32) mm	0.00 mm	6.33 mm	1.8×10^4
S_3	(0.37 ± 0.44) mm	0.00 mm	7.37 mm	2.4×10^4
S_4	(0.43 ± 0.55) mm	0.00 mm	8.51 mm	3.2×10^4
S_5	(0.50 ± 0.68) mm	0.00 mm	9.56 mm	4.1×10^4
S_6	(0.60 ± 0.85) mm	0.00 mm	11.26 mm	5.3×10^4
ICD	(0.12 ± 0.05) mm	0.08 mm	1.66 mm	5.3×10^4
fiberoptic system				
subset	mean dist. $\pm$ s	min. dist.	max. dist.	n_{total}
S_1	(0.38 ± 0.32) mm	0.02 mm	1.40 mm	7.2×10^2
S_2	(0.41 ± 0.33) mm	0.02 mm	2.04 mm	9.9×10^2
S_3	(0.44 ± 0.35) mm	0.01 mm	2.96 mm	1.3×10^3
S_4	(0.51 ± 0.43) mm	0.01 mm	6.75 mm	1.6×10^3
S_5	(0.57 ± 0.50) mm	0.01 mm	6.84 mm	2.1×10^3
S_6	(0.62 ± 0.53) mm	0.01 mm	6.75 mm	2.6×10^3
ICD	(0.48 ± 0.13) mm	0.40 mm	4.13 mm	2.6×10^3

[a] Physical distances are limited to $\mathbb{R}_{\geq 0}$.

B. TRAJECTORY APPROXIMATION AND IMPACT SITE PREDICTION ACCURACY

In this section, the results of the goodness of fit of the analytical trajectory approximations and of the ISP accuracy measurements are described.

1) GOODNESS OF FIT OF TRAJECTORY MODEL

The goodness of fit of the parabolic trajectory model, i.e., the orthogonal distances between the model function T_{para} and the triangulated droplet positions S_{samp}, is shown in Fig. 14 for both systems. Global values of (0.5 ± 0.4) mm and (0.8 ± 0.7) mm were found for the rod lens-based and the fiberoptic system, respectively. The results are additionally provided in greater detail in Table 3.

2) IMPACT SITE PREDICTION ACCURACY

An ISP error of below 2 mm is desirable to successfully stimulate the left or right vocal fold with approximate dimensions of 13 × 4 mm [52]. The results of the experimental evaluation of the ISP accuracy for all system configurations

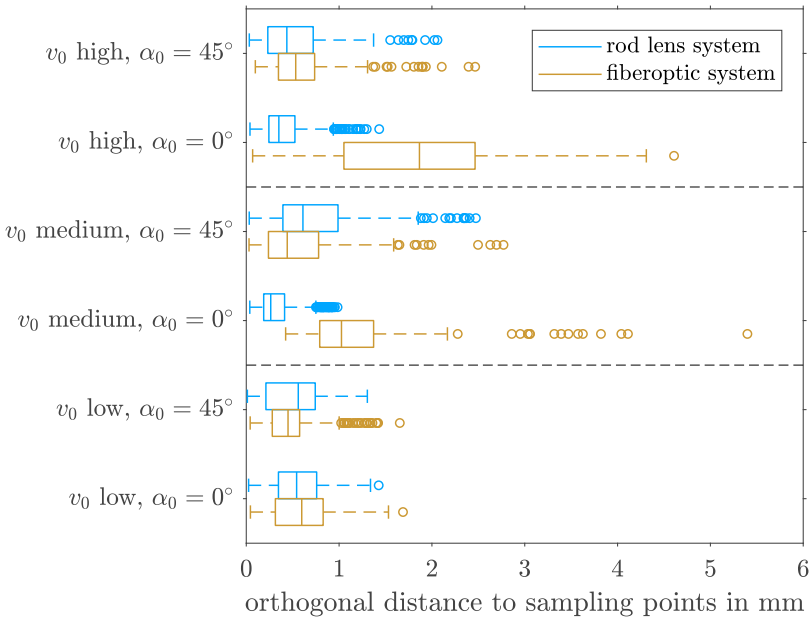


FIGURE 14. Global goodness of fit (orthogonal distances to set S_{samp} of droplet positions) of identified parabolic trajectory approximations T_{para} at different system configurations (lower is better). Outliers displayed as circles. One outlier beyond figure limits not shown for clarity.

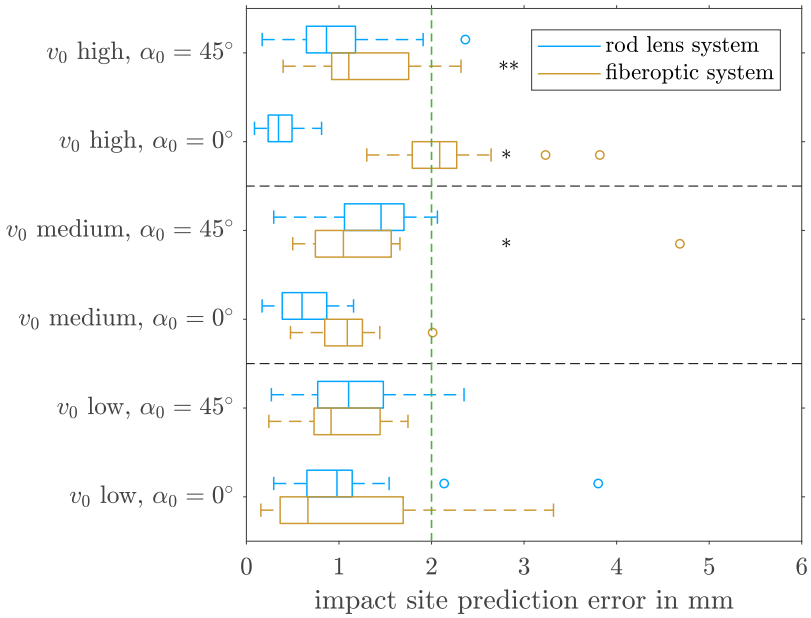


FIGURE 15. Results of experimental quantification of impact site prediction accuracy at different system configurations (lower is better). Dashed vertical line represents 2 mm threshold. $\tau_\Pi = 0.4$ mm, $r_{tol} = 0.2$ mm used with the rod lens system; these threshold values were doubled for the fiberoptic system to account for lower system resolution. Results based on 20 droplet shooting experiments per system/setting. Outliers displayed as circles. $*/**$ indicates cases where no impact site prediction could be identified in one/two experiments, respectively.

are shown in Fig. 15. A dashed vertical line indicates the 2 mm error threshold. In the case of the rod lens-based laryngoscope, 96 % of all observed ISP error values fall below this threshold. For the fiberoptic system, a value of 80 % was found. Whereas all frame sequences recorded with the rod lens-based system yielded a successful ISP, no valid impact site could be predicted in 4 of the 120 sequences recorded with the fiberoptic system as indicated by asterisks in Fig. 15.

TABLE 3. Mean values and sample standard deviations of orthogonal distances of all acquired sampling points to global fit plane Π and global fit function T_{para} for different combinations of muzzle velocity v_0 and droplet shooting angle α_0 for both systems. n_{total} represents the number of available sampling points (acquired in 10 droplet shooting experiments at constant conditions).

rod lens-based system				
v_0	α_0	dist. to global Π	dist. to global T_{para}	n_{total}
low	0°	(0.30 ± 0.23) mm	(0.56 ± 0.28) mm	503
med.	0°	(0.16 ± 0.12) mm	(0.34 ± 0.23) mm	419
high	0°	(0.24 ± 0.18) mm	(0.42 ± 0.27) mm	502
low	45°	(0.31 ± 0.26) mm	(0.51 ± 0.30) mm	538
med.	45°	(0.31 ± 0.17) mm	(0.76 ± 0.47) mm	457
high	45°	(0.17 ± 0.14) mm	(0.51 ± 0.37) mm	513
fiberoptic system				
v_0	α_0	dist. to global Π	dist. to global T_{para}	n_{total}
low	0°	(0.28 ± 0.25) mm	(0.60 ± 0.33) mm	317
med.	0°	(0.41 ± 0.29) mm	(1.29 ± 0.88) mm	149
high	0°	(0.69 ± 0.64) mm	(1.84 ± 0.98) mm	279
low	45°	(0.20 ± 0.15) mm	(0.48 ± 0.28) mm	415
med.	45°	(0.22 ± 0.19) mm	(0.56 ± 0.44) mm	386
high	45°	(0.29 ± 0.21) mm	(0.60 ± 0.37) mm	374

TABLE 4. Mean/minimum/maximum values and sample standard deviations s of the impact site prediction errors for different combinations of muzzle velocity v_0 and droplet shooting angle α_0 for both systems. n_{seq} represents the number of sequences with successful impact site identification.

rod lens-based system					
v_0	α_0	mean error $\pm$ s	min. error	max. error	n_{seq}
low	0°	(1.1 ± 0.8) mm	0.3 mm	3.8 mm	20/20
med.	0°	(0.6 ± 0.3) mm	0.2 mm	1.2 mm	20/20
high	0°	(0.4 ± 0.2) mm	0.1 mm	0.8 mm	20/20
low	45°	(1.1 ± 0.5) mm	0.3 mm	2.4 mm	20/20
med.	45°	(1.4 ± 0.5) mm	0.3 mm	2.1 mm	20/20
high	45°	(1.0 ± 0.5) mm	0.2 mm	2.4 mm	20/20
fiberoptic system					
v_0	α_0	mean error $\pm$ s	min. error	max. error	n_{seq}
low	0°	(1.1 ± 1.1) mm	0.2 mm	3.3 mm	20/20
med.	0°	(1.1 ± 0.4) mm	0.5 mm	2.0 mm	20/20
high	0°	(2.1 ± 0.6) mm	1.3 mm	3.8 mm	19/20
low	45°	(1.0 ± 0.5) mm	0.2 mm	1.7 mm	20/20
med.	45°	(1.3 ± 0.9) mm	0.5 mm	4.7 mm	19/20
high	45°	(1.3 ± 0.6) mm	0.4 mm	2.3 mm	18/20

Global ISP errors of (0.9 ± 0.6) mm and (1.3 ± 0.8) mm were found with the rod lens-based and the fiberoptic system, respectively. The mean/minimum/maximum values and sample standard deviations of the ISP errors are provided individually for each system setting in Table 4.

The computation times for the WLS-filtered disparity map, the reduced candidate set S_Π, and the final ISP step are given in Table 5 for both systems. In the case of the rod lens-based system, the mean impact site computation time could be reduced from 43 ms to 1 ms by efficiently reducing the set of impact site candidate points according to (7). A demonstration video of online ISP is made available at *IEEE DataPort* [41].

V. LIMITATIONS

The number of repetitions for MIT-LAR ISP accuracy quantification was limited by the workload associated with each individual experiment and the subsequent manual analysis. Still, we performed and evaluated a total of 120 experiments for each of the two systems presented in this work. The proposed, straightforward approach for the identification of a parabolic approximation of a droplet trajectory, based on a set of trajectory recordings, relies on the presumption that the first sampling point is triangulated at a similar spatial location in each of the recordings to obtain a uniform range of the parameter t, starting at $t = 0$ for the first sampling point. This assumption holds in the present work, as the optical system and the droplet trajectory were unchanged/highly reproducible between the individual recordings. The prototypical stereo laryngoscopes presented in this contribution have to be sealed against external substances and a safety assessment has to be completed before a clinical trial, involving animal or human test subjects, can be performed. Such a trial will allow to quantify the *in vivo* ISP accuracy, which lies well beyond the scope of the present work. It will further allow to obtain statistically significant results with respect to the clinical benefit of the proposed system as well as the usability perceived by this group of end users.

TABLE 5. Computation times for WLS-filtered disparity map, reduced impact site candidate set S_Π, and final impact site prediction (ISP) step for both laryngoscopic systems. Times given as (mean $\pm$ sample standard deviation), based on the 20 sequences recorded at low v_0 and $\alpha_0 = 0°$.

system	disp. map	S_Π	ISP on S_Π
rod lens-based system	(265 ± 8) ms	2 ms [a]	1 ms [a]
fiberoptic system	(17 ± 2) ms	< 1 ms	< 1 ms

[a] Sample standard deviation below 1 ms.

VI. CONCLUSION

In this final chapter, the methods and results described above are discussed and indications for further research are provided.

A. DISCUSSION

To the best of the authors' knowledge, this manuscript represents the first work on the targeted, droplet-based stimulation of the LAR using a physically motivated model of the droplet trajectory. In the laboratory setting, the systems and methods presented in this contribution yielded global mean ISP accuracies that are below the critical threshold of 2 mm enabling a targeted stimulation of the vocal folds. This fundamentally advances the current state of research.

In the case of the fiberoptic assembly, the lateral boundary of the vocal fold should be targeted to account for the higher ISP errors observed with this system. The proposed fiberoptic stereo laryngoscope is the first MIT-LAR system offering improved anatomical compliance thanks to an increased structural flexibility.

The comparison of the goodness of fit and the ISP results obtained with the fiberoptic stereo laryngoscope for a vertical droplet trajectory (see Fig. 14 and 15) shows that a high goodness of fit during the droplet trajectory

approximation step correlates with a high ISP accuracy. This expected relationship shows the importance of a reproducible droplet trajectory, as this factor immediately influences the prediction accuracy during the clinical use of the system.

The two laryngoscopic systems are characterized by discrete sets of properties. The rigid, rod lens-based system offers a high image resolution and thus, a high density of the spatial reconstruction (see Fig. 13). The lower image resolution of the fiber-based system can partly be compensated by a larger stereo base at reduced overall shaft dimensions, resulting in a comparatively high stereo reconstruction accuracy with a GT point cloud distance of (0.6 ± 0.5) mm for the largest point cloud subset S_6, compared to a value of (0.6 ± 0.9) mm obtained with the rod lens-based system. The median and the spread of the reconstruction error increase when the exploited periglottal zone is expanded. This can be explained by a nonuniform image quality, which may be due to the manual adjustment of the optical devices. The point cloud obtained with the fiberoptic system has a notably lower density with an intracloud point distance of (0.5 ± 0.1) mm, compared to (0.1 ± 0.1) mm for the rod lens-based system. This was accounted for by choosing higher values of the thresholds τ_Π and r_{tol} for the fiberoptic device (see Fig. 15), yielding a slightly elevated global ISP error of (1.3 ± 0.8) mm, compared to a value of (0.9 ± 0.6) mm for the rod lens-based system. Still, we consider these results quite remarkable for manually assembled prototype devices.

B. SUGGESTIONS FOR FUTURE WORK

Initial feedback provided by two expert laryngologists without prior experience in handling the system indicates a potential benefit of an inward orientation of the nozzle. This would enable targeting an area which is more centrally located in the endoscopic image. Furthermore, GPU implementation and code parallelization may help diminish the lag observed in the endoscopic live view. To further increase the clinical utility of MIT-LAR, the three-dimensional motions of the laryngeal structures could readily be determined for diagnostic purposes using the computed spatial reconstruction. These structures could also be tracked, paving the way for a robotic MIT-LAR system which would autonomously stimulate pre-defined targets of interest in the larynx. By addition of an adequate algorithm for the analysis of the high-speed frame sequence recorded after LAR stimulation, the LAR latency of a given individual could ultimately be determined with very limited human involvement [53], thus drastically increasing the efficiency and objectiveness of routine aspiration risk screening.

ACKNOWLEDGMENT

The authors would like to thank Sven Beuchler and Tim Haubold (Institute of Applied Mathematics, Leibniz Universität Hannover) for their assessment of the parabolic trajectory modeling approach, Jorge Badilla-Solórzano, Leon Budde, Sontje Ihler, and Michael Jungheim for helpful discussions, Robin Szymanski for his work on the camera-computer interface, and Paul Loose for compiling additional literature.

DECLARATION OF COMPETING INTERESTS

J. F. Fast, Dennis Kundrat, Lüder A. Kahrs, and Martin Ptok are listed as inventors in the MIT-LAR-related European patent EP 3 536 226 [39].

REFERENCES

[1] World Health Organization. (2020). *The Top 10 Causes of Death*. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death

[2] A. Kaneoka, J. M. Pisegna, H. Inokuchi, R. Ueha, T. Goto, T. Nito, C. E. Stepp, M. P. LaValley, N. Haga, and S. E. Langmore, "Relationship between laryngeal sensory deficits, aspiration, and pneumonia in patients with dysphagia," *Dysphagia*, vol. 33, no. 2, pp. 192–199, Apr. 2018, doi: 10.1007/s00455-017-9845-8.

[3] A. S. Domer, M. A. Kuhn, and P. C. Belafsky, "Neurophysiology and clinical implications of the laryngeal adductor reflex," *Current Otorhinolaryngology Rep.*, vol. 1, no. 3, pp. 178–182, Sep. 2013, doi: 10.1007/s40136-013-0018-5.

[4] T. Yamashita, E. A. Nash, Y. Tanaka, and C. L. Ludlow, "Effects of stimulus intensity on laryngeal long latency responses in awake humans," *Otolaryngol. Head Neck Surg.*, vol. 117, no. 5, pp. 521–529, 1997. [Online]. Available: https://journals.sagepub.com/doi/full/10.1016/S0194-59989770025-1

[5] C. L. Ludlow, F. Van Pelt, and J. Koda, "Characteristics of late responses to superior laryngeal nerve stimulation in humans," *Ann. Otology, Rhinology Laryngology*, vol. 101, no. 2, pp. 127–134, Feb. 1992, doi: 10.1177/000348949210100204.

[6] B. Carey, L. Sulica, A. Wu, and R. Branski, "A novel electrodiagnostic assessment of the laryngeal closure reflex," *Muscle Nerve*, vol. 47, no. 3, pp. 432–436, Mar. 2013, doi: 10.1002/mus.23544.

[7] C. F. Sinclair, M. J. Téllez, and S. Ulkatan, "Human laryngeal sensory receptor mapping illuminates the mechanisms of laryngeal adductor reflex control," *Laryngoscope*, vol. 128, no. 11, pp. E365–E370, Nov. 2018, doi: 10.1002/lary.27248.

[8] R. D. Andreatta, E. A. Mann, C. J. Poletto, and C. L. Ludlow, "Mucosal afferents mediate laryngeal adductor responses in the cat," *J. Appl. Physiol.*, vol. 93, no. 5, pp. 1622–1629, Nov. 2002, doi: 10.1152/japplphysiol.00417.2002.

[9] L. F. Santoso, S. Jafari, D. Y. Kim, and D. Paydarfar, "The internal superior laryngeal nerve in humans: Evidence for pure sensory function," *Laryngoscope*, vol. 131, no. 1, pp. E207–E211, Jan. 2021, doi: 10.1002/lary.28642.

[10] C. F. Sinclair, M. J. Téllez, O. R. Tapia, and S. Ulkatan, "Contralateral R1 and R2 components of the laryngeal adductor reflex in humans under general anesthesia," *Laryngoscope*, vol. 127, no. 12, pp. E443–E448, Dec. 2017, doi: 10.1002/lary.26744.

[11] C. F. Sinclair, M. J. Téllez, O. R. Tapia, S. Ulkatan, and V. Deletis, "A novel methodology for assessing laryngeal and vagus nerve integrity in patients under general anesthesia," *Clin. Neurophysiol.*, vol. 128, no. 7, pp. 1399–1405, Jul. 2017, doi: 10.1016/j.clinph.2017.03.002.

[12] C. F. Sinclair, M. J. Téllez, and S. Ulkatan, "Noninvasive, tube-based, continuous vagal nerve monitoring using the laryngeal adductor reflex: Feasibility study of 134 nerves at risk," *Head Neck*, vol. 40, no. 11, p. 2498—2506, 2018, doi: 10.1002/hed.25377.

[13] M. Suzuki and C. T. Sasaki, "Effect of various sensory stimuli on reflex laryngeal adduction," *Ann. Otology, Rhinology Laryngology*, vol. 86, no. 1, pp. 30–36, Jan. 1977, doi: 10.1177/000348947708600106.

[14] Y.-H. Kim, J. W. Kang, and K.-M. Kim, "Characteristics of glottic closure reflex in a canine model," *Yonsei Med. J.*, vol. 50, no. 3, pp. 380–384, 2009, doi: 10.3349/ymj.2009.50.3.380.

[15] M. J. Téllez, S. Ulkatan, A. Blitzer, and C. F. Sinclair, "Unearthing a consistent bilateral R1 component of the laryngeal adductor reflex in awake humans," *Laryngoscope*, vol. 128, no. 11, p. 2581—2587, 2018, doi: 10.1002/lary.27249.

[16] J. E. Aviv, J. H. Martin, M. Debell, M. S. Keen, and A. Blitzer, "Air pulse quantification of supraglottic and pharyngeal sensation: A new technique," *Ann. Otology, Rhinology Laryngology*, vol. 102, no. 10, pp. 777–780, Oct. 1993, doi: 10.1177/000348949310201007.

[17] T. Kim, K. Goodhart, J. E. Aviv, R. L. Sacco, B. Diamond, S. Kaplan, and L. G. Close, "FEESST: A new bedside endoscopic test of the motor and sensory components of swallowing," *Ann. Otology, Rhinology Laryngology*, vol. 107, no. 5, pp. 378–387, May 1998, doi: 10.1177/000348949810700503.

[18] J. E. Aviv and J. H. Martin, "Measurements of respiratory sensory discrimination thresholds," U.S. Patent 5 970 978, Oct. 26, 1996.

[19] J. C. Borders, M. B. O'Dea, E. McNally, E. Norberg, M. Kitila, M. Walsh, R. Liu, and J. M. Pisegna, "Inter- and intra-rater reliability of laryngeal sensation testing with the touch method during flexible endoscopic evaluations of swallowing," *Ann. Otology, Rhinology Laryngology*, vol. 129, no. 6, pp. 565–571, Jun. 2020, doi: 10.1177/0003489419901145.

[20] A. Kaneoka, J. M. Pisegna, G. P. Krisciunas, T. Nito, M. P. LaValley, C. E. Stepp, and S. E. Langmore, "Variability of the pressure measurements exerted by the tip of laryngoscope during laryngeal sensory testing: A clinical demonstration," *Amer. J. Speech-Lang. Pathol.*, vol. 26, no. 3, pp. 729–736, 2017, doi: 10.1044/2017_AJSLP-16-0006.

[21] M. P. Strohl, V. N. Young, C. D. Dwyer, A. Bhutada, E. Crawford, J. L. Chang, C. A. Rosen, and S. W. Cheung, "Novel adaptation of a validated tactile aesthesiometer to evaluate laryngopharyngeal sensation," *Laryngoscope*, vol. 131, no. 6, pp. 1324–1331, Jun. 2021, doi: 10.1002/lary.28947.

[22] M. P. Strohl, J. L. Chang, C. D. Dwyer, V. N. Young, C. A. Rosen, and S. W. Cheung, "Laryngeal adductor reflex movement latency following tactile stimulation," *Otolaryngol. Head Neck Surg.*, Jul. 2021, Art. no. 1945998211025517, doi: 10.1177/01945998211025517.

[23] M. Ptok and S. Schroeter, "[Deliberate release of the laryngeal adductor reflex via microdroplet impulses. Development of a device] Gezielte Auslösung des laryngealen Adduktionsreflexes per Mikrotröpfchenabgabe: Entwicklung einer Vorrichtung," *HNO*, vol. 64, no. 3, pp. 149–155, 2016, doi: 10.1007/s00106-016-0130-1.

[24] M. Ptok and S. Schroeter, "[Characterization of the laryngeal adductor reflex by stimulation with microdroplets impulses (microdroplet impulse testing)] Charakterisierung des laryngealen Adduktionsreflexes durch Stimulation mit Mikrotröpfchen-Impulsen (microdroplet impulse testing)," *Laryngo-Rhino-Otologie*, vol. 95, no. 7, pp. 482–489, Feb. 2016, doi: 10.1055/s-0041-111516.

[25] M. Ptok, S. Bonenberger, S. Miller, D. Kühn, and M. Jungheim, "[Laryngeal adduction reflex] Der laryngeale Adduktionsreflex," *Laryngo-Rhino-Otologie*, vol. 93, no. 7, pp. 446–449, Jul. 2014, doi: 10.1055/s-0034-1370928.

[26] P. R. Kearney, E. A. Mann, C. J. Poletto, and C. L. Ludlow, "Suppression of thyroarytenoid muscle responses during repeated air pressure stimulation of the laryngeal mucosa in awake humans," *Ann. Otology, Rhinology Laryngology*, vol. 114, no. 4, pp. 264–270, Apr. 2005, doi: 10.1177/000348940511400403.

[27] J. F. Fast, K. A. Westermann, M.-H. Laves, M. Jungheim, M. Ptok, T. Ortmaier, and L. A. Kahrs, "Droplet applicator module for reproducible and controlled endoscopic laryngeal adductor reflex stimulation," *Biomicrofluidics*, vol. 14, no. 4, Jul. 2020, Art. no. 044112, doi: 10.1063/5.0004351.

[28] M. Sawashima, H. Hirose, K. Honda, H. Yoshioka, S. R. Hibi, N. Kawase, and M. Yamada, "Stereoendoscopic measurement of the laryngeal structure," in *Vocal Fold Physiology: Contemporary Research and Clinical Issues*, D. M. Bless and J. H. Abbs, Eds. San Diego, CA, USA: College-Hill Press, 1983, pp. 264–276.

[29] K. A. Stevens, R. Shimamura, H. Imagawa, K.-I. Sakakibara, and I. T. Tokuda, "Validating stereo-endoscopy with a synthetic vocal fold model," *Acta Acustica United Acustica*, vol. 102, no. 4, pp. 745–751, Jul. 2016, doi: 10.3813/AAA.918990.

[30] D. E. Sommer, I. T. Tokuda, S. D. Peterson, K.-I. Sakakibara, H. Imagawa, A. Yamauchi, T. Nito, T. Yamasoba, and N. Tayama, "Estimation of inferior-superior vocal fold kinematics from high-speed stereo endoscopic data *in vivo*," *J. Acoust. Soc. Amer.*, vol. 136, no. 6, pp. 3290–3300, Dec. 2014, doi: 10.1121/1.4900572.

[31] I. T. Tokuda, M. Iwawaki, K.-I. Sakakibara, H. Imagawa, T. Nito, T. Yamasoba, and N. Tayama, "Reconstructing three-dimensional vocal fold movement via stereo matching," *Acoust. Sci. Technol.*, vol. 34, no. 5, pp. 374–377, 2013, doi: 10.1250/ast.34.374.

[32] A. Nakano, K. Mishima, R. Shiraishi, and Y. Ueyama, "Quantitative analysis of velopharyngeal movement using a stereoendoscope: Accuracy and reliability of range images," *Comput. Aided Surg.*, vol. 20, no. 1, pp. 29–33, Jan. 2015, doi: 10.3109/10929088.2015.1076041.

[33] N. A. George, F. F. M. de Mul, Q. Qiu, G. Rakhorst, and H. K. Schutte, "Depth-kymography: High-speed calibrated 3D imaging of human vocal fold vibration dynamics," *Phys. Med. Biol.*, vol. 53, no. 10, pp. 2667–2675, Apr. 2008, doi: 10.1088/0031-9155/53/10/015.

[34] M. Semmler, S. Kniesburges, V. Birk, A. Ziethe, R. Patel, and M. Dollinger, "3D reconstruction of human laryngeal dynamics based on endoscopic high-speed recordings," *IEEE Trans. Med. Imag.*, vol. 35, no. 7, pp. 1615–1624, Jul. 2016, doi: 10.1109/TMI.2016.2521419.

[35] C. Bohr, M. Döllinger, S. Kniesburges, and M. Traxdorf, "3D-Visualisierung und Analyse von Stimmlippenschwingungen," *HNO*, vol. 64, no. 4, pp. 254–261, Apr. 2016, doi: 10.1007/s00106-016-0122-1.

[36] Y. Kim, J. Oh, S.-H. Choi, A. R. Jung, Y. S. Lee, J.-G. Lee, and J. K. Kim, "High speed vocal fold imaging using smartphone-based laryngoscope: A preliminary study," in *Proc. Conf. Lasers Electro-Opt., Appl. Technol.*, Washington, DC, USA, May 2020, pp. 1–2, doi: 10.1364/CLEO_AT.2020.JTu2F.2.

[37] C. L. Ludlow, T. Yamashita, G. M. Schulz, and F. W.-B. Deleyiannis, "Abnormalities in long latency responses to superior laryngeal nerve stimulation in adductor spasmodic dysphonia," *Ann. Otology, Rhinology Laryngology*, vol. 104, no. 12, pp. 928–935, Dec. 1995, doi: 10.1177/000348949510401203.

[38] J. F. Fast, A. K. Rüppel, C. Bärhold, M. Jungheim, T. Ortmaier, M. Ptok, and L. A. Kahrs, "Endoscopic guidance system for stimulation of the laryngeal adductor reflex by droplet impact," *Proc. SPIE*, vol. 10951, Mar. 2019, Art. no. 109510M, doi: 10.1117/12.2512852.

[39] L. A. Kahrs, D. Kundrat, M. Ptok, and J. F. Fast, "Medical device, method of predicting an impact point and computer program," European Patent 3 536 226, Feb. 17, 2021.

[40] X. Gu, C. Li, X. Xiao, C. M. Lim, and H. Ren, "A compliant transoral surgical robotic system based on a parallel flexible mechanism," *Ann. Biomed. Eng.*, vol. 47, no. 6, pp. 1329–1344, Jun. 2019, doi: 10.1007/s10439-019-02241-0.

[41] J. F. Fast, H. R. Dava, A. K. Rüppel, D. Kundrat, M. Krauth, M. Laves, S. Spindeldreier, L. A. Kahrs, and M. Ptok, "Dataset associated with IEEE Access contribution 'Stereo laryngoscopic impact site prediction for droplet-based stimulation of the laryngeal adductor reflex," *IEEE DataPort Dataset*, doi: 10.21227/nbcn-dd59.

[42] H. Li, X. Zhang, X. Zhang, and L. Lu, "Model optimization and calculation method of multi-target curved motion parameters measurement using multi-screen intersection test mechanism," *IEEE Access*, vol. 8, pp. 147872–147879, 2020, doi: 10.1109/ACCESS.2020.3016090.

[43] R. Murray, "Computational and laboratory investigations of a model of blood droplet flight for forensic applications," M.S. thesis, Fac. Grad. Studies Mod. Comput. Sci., Univ. Ontario Inst. Techn., Oshawa, ON, Canada, 2012. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.357.4888&rep=rep1&type=pdf

[44] S. J. Ahn, *Least Squares Orthogonal Distance Fitting of Curves and Surfaces in Space* (Lecture Notes in Computer Science), vol. 3151. Berlin, Germany: Springer, 2004, p. 2, doi: 10.1007/b104017.

[45] H. Hirschmuller, "Stereo processing by semiglobal matching and mutual information," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 30, no. 2, pp. 328–341, Feb. 2008, doi: 10.1109/TPAMI.2007.1166.

[46] D. Min, S. Choi, J. Lu, B. Ham, K. Sohn, and M. Do, "Fast global image smoothing based on weighted least squares," *IEEE Trans. Image Process.*, vol. 23, no. 12, pp. 5638–5653, Dec. 2014, doi: 10.1109/TIP.2014.2366600.

[47] G. Bradski, "The OpenCV library," *Dr Dobb's J. Softw. Tools*. [Online]. Available: https://github.com/opencv

[48] R. B. Rusu and S. Cousins, "3D is here: Point cloud library (PCL)," in *Proc. IEEE Int. Conf. Robot. Autom.*, May 2011, pp. 1–4, doi: 10.1109/ICRA.2011.5980567.

[49] W. S. Selbie, S. L. Gewalt, and C. L. Ludlow, "Developing an anatomical model of the human laryngeal cartilages from magnetic resonance imaging," *J. Acoust. Soc. Amer.*, vol. 112, no. 3, pp. 1077–1090, Sep. 2002, doi: 10.1121/1.1501586.

[50] F. J. Meisoll, M. Jungheim, J. F. Fast, S. Miller, and M. Ptok, "Upper esophageal sphincter response to laryngeal adductor reflex elicitation in humans," *Laryngoscope*, vol. 131, no. 6, Jun. 2021, doi: 10.1002/lary.29166.

[51] D. Diers, J. F. Fast, F. Götz, L. A. Kahrs, S. Miller, M. Jungheim, and M. Ptok, "Euclidean distances of laryngopharyngeal structures obtained from CT data for preclinical development of laryngoscopic devices," *Surgical Radiolog. Anatomy*, vol. 42, no. 6, pp. 695–700, Jun. 2020, doi: 10.1007/s00276-019-02397-3.

[52] H. Eckel, C. Sittel, P. Zorowka, and A. Jerke, "Dimensions of the laryngeal framework in adults," *Surgical Radiologic Anatomy*, vol. 16, no. 1, pp. 31–36, Mar. 1994, doi: 10.1007/BF01627918.

[53] J. F. Fast, M. Ptok, M. Jungheim, R. Szymanski, T. Ortmaier, and L. A. Kahrs, "Towards fully automated determination of laryngeal adductor reflex latencies through high-speed laryngoscopy image processing," in *Bildverarbeitung für die Medizin*, A. Maier, T. Deserno, H. Handels, K. Maier-Hein, C. Palm, and T. Tolxdorff, Eds. Berlin, Germany: Springer, 2018, pp. 121–126, doi: 10.1007/978-3-662-56537-7_41.

JACOB F. FAST (Graduate Student Member, IEEE) was born in Herdecke, Germany, in 1991. He graduated with the baccalauréat général scientifique in Pauillac, France, in 2008. He received the B.S. and M.S. degrees in mechanical engineering from Leibniz Universität Hannover, Germany, in 2012 and 2016, respectively. He is currently pursuing the Ph.D. degree in mechanical engineering.

He gained first insights into the biomedical engineering industry at Fresenius Medical Care AG & Co. KGaA, Bad Homburg vor der Höhe, Germany, from 2012 to 2013, and the aerospace industry at Airbus S.A.S., Toulouse, France, in 2015. Since 2016, he has been a Research Assistant with the Institute of Mechatronic Systems, Leibniz Universität Hannover. Since 2019, he has additionally been a Research Assistant with the Department of Phoniatrics and Pediatric Audiology, Hannover Medical School. His research interests include stereoscopic endoscopy, 3D reconstruction, and computer aided diagnosis.

Mr. Fast is a member of the Association of German Engineers (VDI) and the Association for Electrical, Electronic and Information Technologies (VDE).

HARDIK R. DAVA received the B.E. degree in mechanical engineering from Gujarat Technological University, Ahmedabad, India, in 2017, and the M.S. degree in optical technology from Leibniz Universität Hannover, Germany, in 2020.

In his master's thesis, he worked on the quantitative assessment of the stereo reconstruction accuracy of different laryngoscopic prototypes. He is currently working at Corvitac GmbH as a Machine Learning Engineer. His research interests include techniques of 3D measurement, augmented reality, and computer vision.

ADRIAN K. RÜPPEL was born in Wertheim, Germany, in 1992. He received the B.E. degree from Baden-Wuerttemberg Cooperative State University (DHBW) and the M.S. degree from Leibniz Universität Hannover, Germany, both in mechanical engineering.

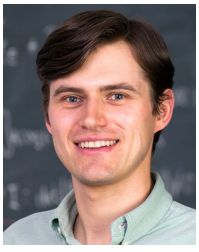
From 2017 to 2018, he explored stereoscopic reconstruction and impact site prediction methods in a student research project and coauthored an SPIE conference contribution on this topic. Since 2019, he has been a Research Assistant with the Laboratory for Machine Tools and Production Engineering (WZL), RWTH Aachen University, Germany. His current research interests include model predictive control for force control in milling and data driven tool wear modeling.

DENNIS KUNDRAT (Member, IEEE) was born in Hannover, Germany. He received the Diploma degree in mechanical engineering with a focus on mechatronics from Leibniz Universität Hannover, Germany, in 2012. He is currently pursuing the Ph.D. degree in mechanical engineering.

After his Diploma degree, he was a Research Assistant with the Institute of Mechatronic Systems, Leibniz Universität Hannover. From 2018 to 2020, he was a Research Associate with the Hamlyn Centre for Robotic Surgery, Imperial College London, U.K., to pave the way for MR-compatible robotic devices in endovascular applications. Recently, he joined Hannover Medical School, Germany, as a Technology Transfer Manager. His main research interests include medical robotics, navigation, and assistive devices in the field of preclinical and translational evaluation.

MAURICE KRAUTH was born in Hannover, Germany, in 1995. He received the B.S. and M.S. degrees in engineering and business administration from Leibniz Universität Hannover, Germany, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree with the Chair of Rail and Public Urban Transport, Technische Universität Dresden, Germany.

In his master's thesis, he developed a non-actuated, fiberoptic stereo laryngoscope, and experimentally compared it to a rod lens-based system. He presented his results at the annual meeting of the German Society of Phoniatrics and Pediatric Audiology (DGPP), in 2019. Since 2020, he has been working as a Research Assistant with the Chair of Rail and Public Urban Transport, Technische Universität Dresden. He is currently researching the control of railcar flows in networks.



MAX-HEINRICH LAVES was born in Hannover, Germany, in 1988. He received the B.S. and M.S. degrees in mechanical engineering from Leibniz Universität Hannover, Hannover, in 2012 and 2015, respectively. He is currently pursuing the Ph.D. degree with a focus on uncertainty estimation in deep learning methods for medical image analysis.

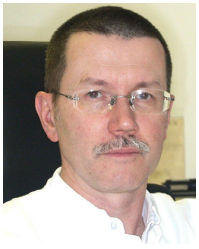
Since his bachelor's, he has been interested in medical technology and was a Research Assistant at the Institute of Mechatronic Systems, Leibniz Universität Hannover, from 2015 to 2020. In 2021, he joined the Institute of Medical Technology and Intelligent Systems, Hamburg University of Technology, to work further on machine learning methods in medicine.

SVENJA SPINDELDREIER (NÉE TAPPE) was born in Bielefeld, Germany, in 1988. She received the B.S. degree in mechatronics from Ostwestfalen-Lippe University of Applied Sciences, Lemgo, Germany, in 2011, and the M.S. degree in mechatronics and the doctoral degree (Dr.-Ing.) in mechanical engineering from Leibniz Universität Hannover, Germany, in 2013 and 2020, respectively.

She joined the Institute of Mechatronic Systems, Leibniz Universität Hannover, in 2013. Starting as a Research Assistant, she became the Head of the Research Group Robotics and Autonomous Systems in 2017, and moved to the Medical Technology and Image Processing Group as a Research Team Leader in 2021. Her current research interests include intelligent mechatronic systems for medical applications, including design, modeling, path planning, and control.

LÜDER A. KAHRS received the Diplom-Physiker degree from Universität Bremen, Germany, in 2002, and the doctoral degree (Dr.-Ing.) in computer science from Universität Karlsruhe (TH), Germany, in 2009.

From 2009 to 2010, he was Postdoctoral Fellow with University Hospital Düsseldorf, Germany, and with the Vanderbilt University Medical Center, Nashville, TN, USA, from 2010 to 2012. From 2011 to 2013, he held a fellowship from the German Research Foundation. From 2013 to 2019, he was the Head of the Research Group Medical Technology and Image Processing, Institute of Mechatronic Systems, Leibniz Universität Hannover, Germany. He is currently an Assistant Professor with the Department of Mathematical and Computational Sciences, University of Toronto Mississauga, Mississauga, Canada. His research interests include medical technology, computer and robot assisted diagnosis and therapy as well as endoscopic and microscopic vision.



MARTIN PTOK was born in 1955. He studied law and medicine at the University of Göttingen, Germany. He completed his residency in otorhinolaryngology at the University of Würzburg, Germany, and his residency in phoniatrics and pediatric audiology at the University of Tübingen, Germany.

He received funding from the German Research Foundation for his stay as a Visiting Scientist at the Kresge Hearing Research Institute, University of Michigan, Ann Arbor, MI, USA. Since 1994, he has been the Director of the Department of Phoniatrics and Pediatric Audiology, Hannover Medical School.