2026 IEEE International Conference on Robotics and Automation (ICRA 2026)
June 1-5, 2026. Vienna, Austria

LATIOS: Latency-Aware Telemonitoring for Injection in Ophthalmic Surgery - An Adaptive Motion Scaling Approach

Korab Hoxha[1], Angelo Henriques[1], Junjie Yang[1], Ahnaf Naheen[2], Mahdi Tavakoli[2] and M. Ali Nasseri[1,2]

***Abstract*— Communication latency in long-distance telerobotic surgery poses a critical safety risk, particularly in high-precision procedures like retinal surgery where tool overshoots can cause irreversible patient injury. This paper introduces the Latency-Aware Telemonitoring for Injection in Ophthalmic Surgery (LATIOS) framework, which enhances safety by adaptively scaling the surgical robot's velocity along the critical axis of tool insertion. Our core contribution is a control algorithm that dynamically modulates the velocity scaling factor based on two real-time, coupled variables: the tool-tip-to-retina distance, estimated via a non-contact, shadow-based method, and the measured communication delay. We validated this system in a transatlantic user study where six participants in North America teleoperated a surgical robot in Europe to perform a series of simulated retinal punctuation tasks. The results demonstrate that LATIOS provides a statistically significant reduction in applied punctuation forces compared to constant control (p = 0.005). This objective safety improvement is achieved through a deliberate safety-efficiency trade-off, with the system enforcing a more cautious pace under high-latency conditions. Our work presents a robust, context-aware safety framework that addresses a key barrier to the clinical adoption of long-distance telerobotic surgery.**

I. INTRODUCTION

In recent years, extensive surgical and medical research has focused on enhancing operational safety by using robotic systems. These platforms enable surgeons to perform complex operations with higher accuracy and safety. An extension of this paradigm, robotic telesurgery, presents an opportunity to bridge geographical gaps between surgeons and patients in remote locations. This technology enables remote delivery of specialized surgical care. The feasibility of long-distance surgery was demonstrated over two decades ago in a transatlantic cholecystectomy [1].

Despite this success, the widespread adoption of long-distance surgery is hindered by a fundamental challenge: communication latency. The delay between the surgeon's command and the robot's action can degrade performance and increase error rates. In delicate procedures, this leads to dangerous tool overshoots, increases the operator's cognitive load, and creates a critical risk of forceful, unintended contact with tissue [2], [3]. These core challenges of long-distance teleoperation are illustrated in Fig. 1.

This work is supported by the state of Bavaria. Corresponding author: K. Hoxha (korab.hoxha@tum.de)

[1]K. Hoxha, A. Henriques, J. Yang and M. Ali Nasseri are with Augenklinik und Poliklinik, Klinikum rechts der Isar of the Technische Universität München, München, 81675, Germany.

[2] A. Naheen, M. Tavakoli and M. Ali Nasseri are with the Department of Biomedical Engineering, University of Alberta, Canada.

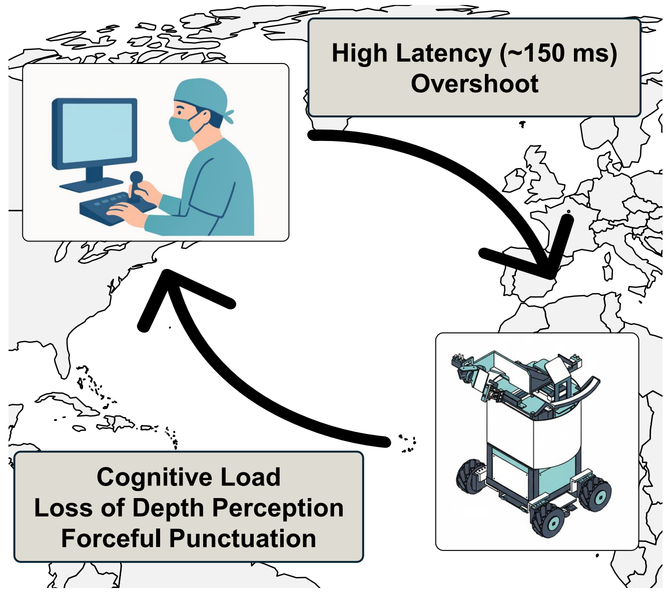


Fig. 1. **The Core Challenges of Long-Distance Telerobotic Surgery.** The diagram illustrates the vicious cycle of high-latency teleoperation. The surgeon's commands are delayed, leading to a high risk of robotic tool overshoot. The resulting sensory feedback is also delayed and lacks critical depth cues, which in turn increases the surgeon's cognitive load and the likelihood of forceful, unintended contact.

The problem of latency is critically amplified in high-precision domains like vitreoretinal surgery, where maneuvers are measured in micrometers and there is zero tolerance for error [4]–[12]. This is particularly relevant for novel therapies targeting debilitating conditions such as age-related macular degeneration (AMD), the leading cause of blindness in developed nations, which is projected to affect 288 million people by 2040 [13]. While robotic systems offer the necessary precision for such procedures, the introduction of significant latency during teleoperation creates a hazardous environment where the risk of forceful punctuation of the delicate retina is high.

Existing approaches to managing latency often involve methods like uniform motion scaling, which constantly dampens the robot's movements. While this can improve accuracy, a static scaling factor does not adapt to the dynamic and context-dependent risks of the surgical task, where danger is highest when the instrument is moving towards the retina at close proximity.

To address this safety challenge, we present the Latency-aware Telemonitoring for Injection in Ophthalmic Surgery (`LATIOS`) framework, which introduces an adaptive and context-aware approach to motion scaling. Specifically, we modulate the velocity of the surgical instrument along the critical z-axis based on two real-time factors: 1) the tool-tip-to-retina distance, estimated via a non-contact, shadow-based method, and 2) the instantaneous communication delay. This allows for fluid motion when the tool is far from the retina while dynamically enforcing safety as the risk of contact grows. We hypothesize that by fusing these real-time spatial (distance) and temporal (latency) data streams, our adaptive motion scaling algorithm will significantly reduce dangerous punctuation forces in high-latency conditions, albeit with a quantifiable trade-off in task completion time, thereby creating a safer and more robust system for long-distance telesurgery.

We validated our system through a transatlantic user study where six participants in Canada used a haptic input device to control a surgical robot in Germany. The study was designed to quantify the performance of our method across three key metrics: punctuation force, task completion time, and subjective workload. The contributions of this paper are:

1) The `LATIOS` framework, which fuses real-time, shadow-based depth estimation and communication latency to adaptively scale robot velocity and enhance safety in long-distance telesurgery.
2) The design and execution of a transatlantic user study providing a real-world evaluation of the system's performance under significant and variable latency.
3) Quantitative results from the study that validate a statistically significant increase in safety (reduced force) is achieved via a deliberate and quantifiable safety-efficiency trade-off (increased time).

II. RELATED WORK

Robotic systems have become instrumental in advancing the state-of-the-art in vitreoretinal surgery, with platforms like the Preceyes Surgical System demonstrating remarkable success in increasing surgical precision by scaling motion and filtering physiological tremor [14]. While such platforms have been successfully teleoperated, their design has been focused on co-located operation where the surgeon is close to the robotic system, thereby not addressing the challenges introduced by long-distance communication latency.

An important aspect of ensuring safety in retinal surgery, whether local or remote, is the accurate perception of the distance from the tool tip to the retinal surface. To overcome this challenge, researchers have explored using visual cues present in the scene. A promising approach is to leverage the shadow cast by the instrument on the retinal surface. Yang et al. have proposed systems that use the instrument's shadow as a visual guide for depth estimation and for predicting the needle's punctuation point [15], [16]. These methods show the feasibility of extracting critical 3D information from a 2D video feed [17], a principle we adopt for our telemonitoring framework.

Beyond depth perception, ensuring safe interaction requires managing the minuscule forces at the tool tip during punctuation. Recent work by Pannek et al. has highlighted that these forces are often imperceptible to a human operator, motivating the development of robotic systems with integrated force sensing to provide crucial feedback [18]. The challenge is further complicated by the dynamic nature of the tissue itself. State-of-the-art research in autonomous local surgery has demonstrated the need for real-time adaptation to physical phenomena such as tissue deformation during insertion [19] and physiological motion from breathing or heartbeat [20]. These advanced methods typically rely on high-frequency intraoperative Optical Coherence Tomography (iOCT) to track the retinal layers in real-time. While powerful, these approaches present challenges for long-distance teleoperation, where high-bandwidth, low-latency iOCT streaming is often not feasible. Our work approaches this safety problem from the perspective of the teleoperator, focusing on mitigating the primary source of error in a high-latency environment: the operator's own control inputs.

While depth perception addresses the spatial challenge, the temporal challenge of communication delay requires different solutions. High latency in teleoperation is known to increase task completion times, error rates, and the cognitive load on the surgeon [2], [3]. An effective method to mitigate these issues is motion scaling, where the motion from the input device is reduced before being executed by the robotic system. Numerous studies have confirmed that reducing the scaling factor improves precision and accuracy, albeit often at the cost of task completion time [21], [22]. More recent work has focused on optimizing this scaling factor. For instance, Lim et al. investigated adapting the scaling factor based on communication latency and even personalized it to surgeon tendencies [23].

Our research builds upon and integrates these distinct fields. While prior work has established the utility of shadow-based depth cues and the effectiveness of adaptive motion scaling for latency, these concepts have been explored largely in isolation. To our knowledge, no prior work has proposed a framework that directly couples real-time, vision-based depth perception with network latency to create an adaptive safety layer. We address this gap by introducing a telemonitoring framework where the motion scaling factor is not static, nor is it only a function of network delay. Instead, we dynamically modulate the tool's velocity along the critical z-axis based on both the measured latency and the real-time, shadow-based estimation of the tool-tip-to-retina distance. This creates a context-aware safety system that gracefully restricts motion only when danger is imminent, addressing both spatial and temporal challenges simultaneously to enhance safety in long-distance telerobotic surgery.

III. METHODS

To validate our telemonitoring framework, we developed a transatlantic telerobotic system and designed a user study to quantify its impact on safety and performance. This section details the physical hardware of our experimental setup, the core algorithms, and the setup of the user study.

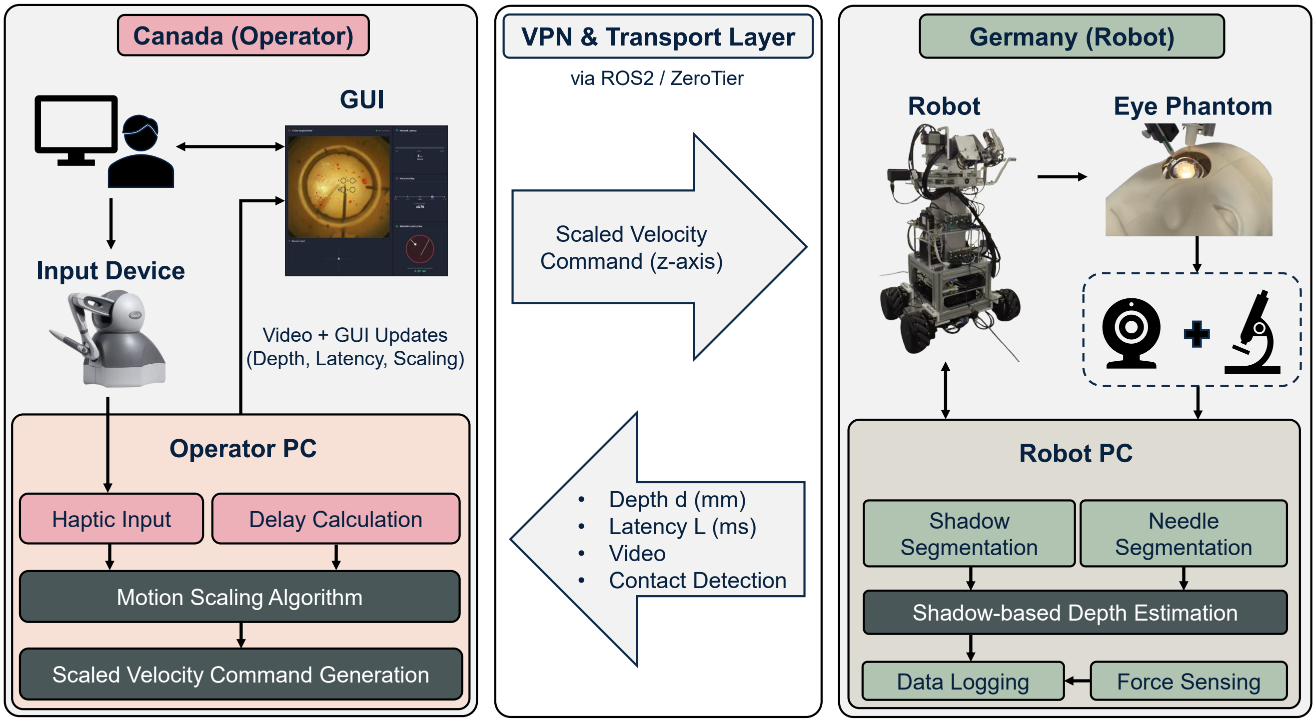


Fig. 2. **The Transatlantic Telerobotic System Architecture.** The operator site (left) in Canada consists of a user, a haptic input device, and an operator PC that runs the motion scaling and delay calculation. The robotic site (right) in Germany includes the surgical robot, an eye phantom, and a robot PC that performs shadow-based depth estimation and force sensing. The two sites are connected via a transport layer (center) that transmits scaled velocity commands and returns the video stream, depth values, retina contacts, and latency measurements.

A. System Overview

The experimental platform, illustrated in Fig. 2, comprised two interconnected sites: an operator station in Canada, and a robotic station in Germany. The operator site consisted of a PC connected to a 3D Systems Touch haptic device. The device provides intuitive stylus control and supports a remote center of motion (RCM) constraint. The operator PC captured motion commands, executed the scaling algorithm, and displayed the graphical user interface (GUI).

The robotic site featured a robotic system, which was specifically developed for high-precision ophthalmic procedures such as retinal surgery [4]. The robot offers high-resolution motion control, capable of executing movements in the micrometer range, a critical requirement for delicate intraocular maneuvers. For this study, the robot's end-effector was equipped with a state-of-the-art 6-axis HEX-8 force/torque sensor (Resense, Germany), which held a standard surgical instrument. The sensor's primary role was for data acquisition, providing precise measurements of the forces exerted during punctuation. The robot was positioned over a physical eye model designed to mimic the constraints of a real surgery. The model replicates the globe's suspension by extraocular muscles using a spring-based mechanism and simulates tissue properties with a silicone eye, all housed within a face mask to ensure anatomically correct instrument access [24]. Visual feedback of the operating area was provided by a Basler camera (A1920-150uc) mounted on a surgical microscope. The live video feed was processed by a dedicated robot-side PC, which was tasked with running the shadow-based depth estimation algorithm and executing the final commands sent from the operator.

The two sites were connected over a public internet connection using a virtual private network established with ZeroTier. All data, including motion commands, video streams, and sensor values, were exchanged as topics within the Robot Operating System 2 (ROS2) framework. The natural transatlantic connection introduced a baseline round-trip latency that typically fluctuated between 150 ms and 200 ms during the experiments. Note that L in our control loop represents this measured network round-trip time.

B. Shadow-Based Depth Estimation

To provide the system with real-time awareness of the tool's proximity to the retina, we implemented a depth estimation algorithm that leverages the shadow cast by the surgical instrument according to [15]. This approach uses the existing microscope video feed instead of additional hardware such as iOCT.

The algorithm operates on the geometric principle that, under a fixed light source, the 2D pixel distance between the instrument tip and its shadow is directly proportional to the instrument's 3D height above the retinal surface. The algorithm identifies the centroids of the tool tip and its shadow in the microscope video (30 fps). This pixel distance is then converted into a real-world distance in millimeters using a pre-calibrated camera model. This calculated tool-to-retina distance, denoted as d, serves as a critical input to our motion scaling controller. It is important to note that this method's accuracy diminishes as the tool gets extremely close to the surface ($d < 1$ mm) due to self-occlusion of the shadow, a limitation we address in our control algorithm.

C. Adaptive Velocity Scaling

The core of our contribution is an adaptive motion scaling algorithm designed to enhance safety without being overly restrictive. Unlike uniform motion scaling, which constantly dampens the operator's input, our method modulates the scaling factor, s, based on two real-time variables: the tool-to-retina distance d and the measured round-trip communication latency L. The scaling is applied exclusively to the velocity along the z-axis (the direction of approach towards the retina). While tangential motion (x-y axis) at low clearance can also induce tissue tearing, the z-axis dictates the primary insertion force during the highest-risk phase of the procedure. Risk depends on both tool proximity and operator reaction time, which is affected by latency. The system allows for fluid, unscaled motion ($s = 1.0$) when the tool is far from the retina. As the tool crosses a predefined safety threshold, the scaling factor is smoothly and linearly reduced, proactively dampening the robot's velocity to prevent dangerous overshoots.

This behavior is governed by the following relationship:

$$s(d, L) = \max\left(s_{\min}, \frac{k \cdot d}{T_{\text{reaction}} + T_{\text{stop}} + L}\right) \qquad (1)$$

where d is the tool-to-retina distance and L is the measured round-trip latency. The total delay in the denominator accounts for both network and human-in-the-loop factors. The human reaction time, T_{reaction}, is based on established literature for complex human-in-the-loop control tasks and is set to a conservative value of 0.5 s [25], [26]. This value is consistent with telerobotics literature that identifies delays of 500 ms and greater as significant for operator performance, and is justified by cognitive models that account for the multiple processing stages in complex choice-reaction tasks. The robot's motor stop time, T_{stop}, was experimentally characterized by measuring the time required for the robot to decelerate to a complete stop from maximum velocity. Based on these measurements, we selected a conservative value of 0.1 s to ensure the model robustly accounts for the system's physical inertia. The term k is a constant gain factor used to map the function to the desired operational range, and $s_{\min}$ is a minimum scaling factor of 0.2, ensuring a slow and deliberate final approach as described previously.

As shown in Fig. 3, this formulation creates a scaling behavior that is directly responsive to risk. A higher measured latency L increases the denominator, causing the scaling to become more aggressive—that is, the slowdown begins earlier (at a greater distance d) and the velocity is reduced more steeply. This preemptively accounts for the increased total delay in the control loop. To handle the known limitations of our vision-based sensor at very close range and to ensure maximum safety during contact, the scaling factor is clamped at a predefined minimum value, $s_{\min} = 0.2$, whenever the distance d falls below 1 mm. This mandatory "safe mode" is indicated by the dashed vertical line in Fig. 3. This approach balances operational efficiency and patient safety in high-latency teleoperation.

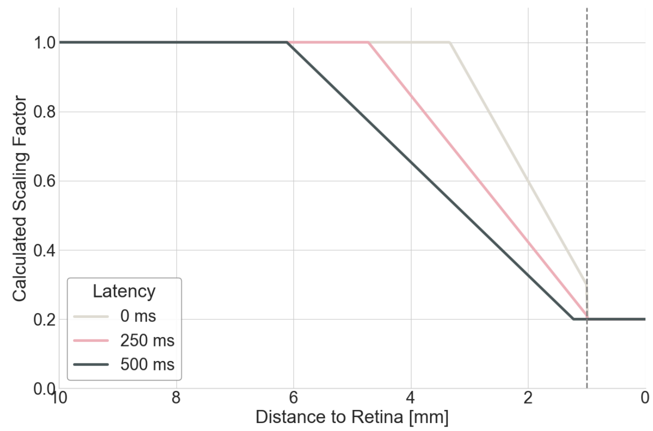


Fig. 3. **Adaptive Scaling Factor as a Function of Distance and Latency for `LATIOS`.** The graph illustrates the relationship between the calculated scaling factor s and the tool-to-retina distance d for different latencies. Our proposed `LATIOS` mode becomes more conservative the higher the latency and initiates the slowdown earlier. The dashed line indicates the 1 mm threshold where the system enters a mandatory "safe mode" with a minimum scaling factor.

D. User study and Evaluation

We conducted a user study with six participants to evaluate the effectiveness of our system. The study protocol was approved by the Research Ethics Board at the University of Alberta, and all participants provided informed written consent prior to the experiment. The cohort included two expert vitreoretinal surgeons, two medical students, and two engineering students to capture a range of user experience and to assess the system's intuitiveness across varying levels of clinical and technical expertise. Prior to the experiment, each participant received a ten-minute training session to familiarize themselves with the haptic device, the concept of the Remote Center of Motion (RCM), and the functionalities of the user interface. Each participant, located in Canada, was tasked with performing a series of punctuation maneuvers on the eye phantom in Germany. The primary goal was to touch four designated circular targets on the "retina" as quickly and as gently as possible. To provide the operator with full situational awareness, we developed a custom Graphical User Interface (GUI), shown in Fig. 4. The GUI is composed of several key components:

- **Live Surgical Feed:** The central panel displays the live video stream from the Basler camera, showing the surgical instrument inside the eye phantom. Overlaid on this feed are the four punctuation targets, a green "RETINA CONTACT DETECTED" banner that appears upon successful contact, and an orange "PROXIMITY WARNING" that activates when the tool is less

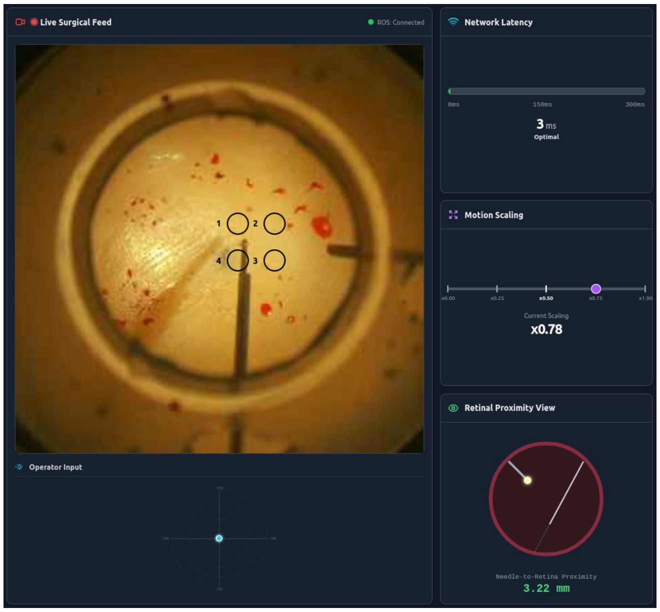


Fig. 4. **The Operator's Graphical User Interface (GUI).** This custom interface provides the user with comprehensive situational awareness by integrating the live surgical feed with real-time visualizations of network latency, the active motion scaling factor, and an intuitive retinal proximity view.

than 1 mm from the retinal surface.

- **Network Latency:** A dedicated panel on the right displays the measured round-trip communication delay in real-time.
- **Motion Scaling:** This panel visualizes the current z-axis velocity scaling factor, allowing the user to see how the system is modulating their commands.
- **Retinal Proximity View:** A side-view schematic of the eye that provides an intuitive visualization of the tool's distance to the retina, complementing the numerical depth value.
- **Operator Input:** A crosshair display at the bottom shows the user's current planar (x-y) input from the haptic device.

The study was conducted as a within-subjects experiment where each participant performed the task under three distinct assistance methods, each defined by a different scaling function $s(d, L)$:

- **Constant Scaling:** A baseline condition with a fixed scaling factor, representing unassisted control:

$$s = 1.0 \tag{2}$$

- **Linear Scaling:** A condition where the velocity was scaled only as a linear function of depth d. The scaling was active in the operational range from $d_{\text{start}} = 5\,\text{mm}$ down to the retinal surface ($d = 0\,\text{mm}$):

$$s(d) = \min\left(1.0, \frac{d}{d_{\text{start}}}\right) \tag{3}$$

- **LATIOS:** Our full proposed method, which scales based on both depth d and latency L and includes a mandatory "safe mode" clamping the factor at $s_{\min}$ for $d < 1\,\text{mm}$, as previously defined in Eq. 1.

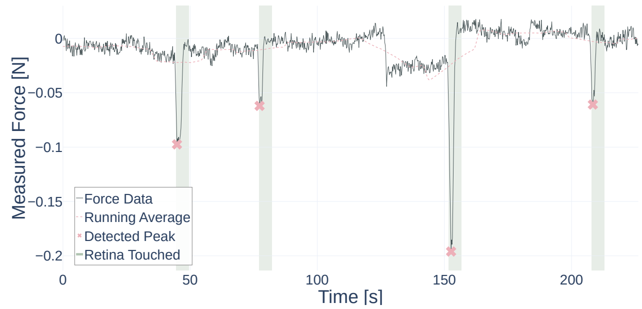


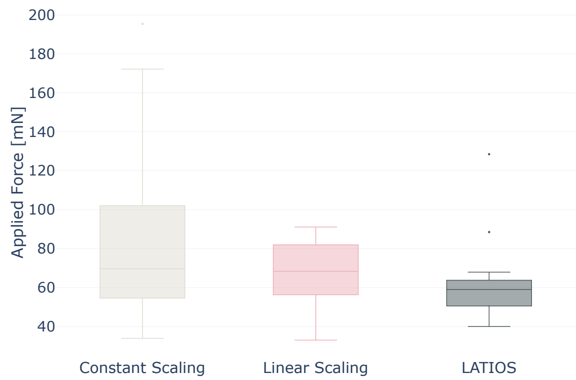
Fig. 5. **Automated Punctuation Force Detection.** The force spike detection method used for objective data analysis. A punctuation event is detected when the force data (black line) sharply deviates from its stable baseline (dashed red line). The algorithm then identifies the peak force of the spike (red 'x'), which is used as the quantitative safety metric.

To mitigate learning effects, the order in which these three methods were presented was randomized for each participant. To specifically test the system's reaction to changing network conditions, each participant performed two punctuation targets under the natural transatlantic latency, and two targets under an additional artificial delay of 500 ms (introduced via the Linux network emulator `netem`) for each of the three modes. This resulted in 12 analyzed puncture events per participant, totaling 72 events across the study.

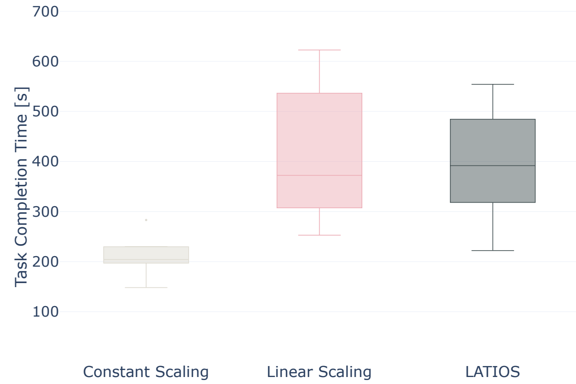
For our evaluation, we collected several objective and subjective metrics. The primary safety metric, punctuation force, was measured using the integrated 6-axis force/torque sensor. To robustly detect each contact event and quantify the force, we implemented a spike detection algorithm, illustrated in Fig. 5. The algorithm calculates a stable baseline of the force signal using a moving average window of 1500 ms and identifies a contact event when the live signal deviates from this baseline by a predefined empirical threshold of 30 mN, chosen to robustly filter sensor noise while capturing initial tissue contact. The subsequent peak force of this spike is then recorded for analysis. This event also momentarily displays the "RETINA CONTACT DETECTED" banner in the user's GUI. In addition to force, we recorded task completion time and collected subjective feedback on cognitive workload using the NASA-TLX questionnaire after each condition.

IV. Results

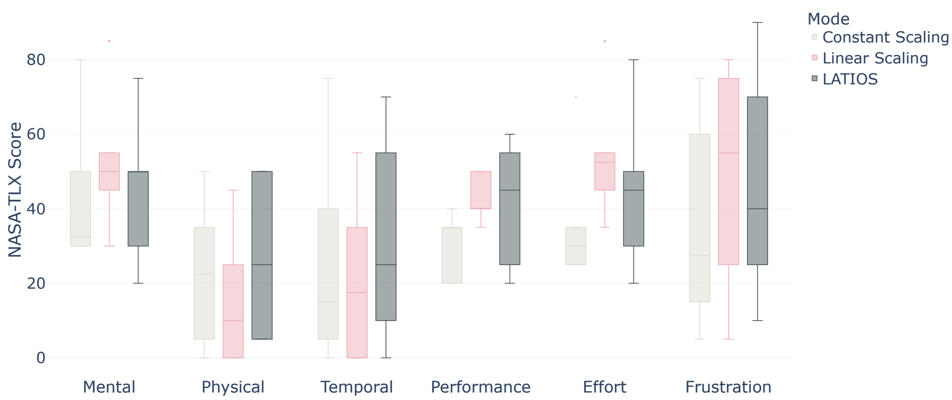
The user study provided a comprehensive dataset across our three primary metrics: punctuation force for objective safety, task completion time for efficiency, and the NASA-TLX score for subjective workload. The analysis shows that the adaptive assistance method improves safety by modulating intelligently tool speed. The most critical finding relates to patient safety, as measured by the applied punctuation force. As illustrated in the boxplot in Fig. 6a, the choice of assistance method had a clear and significant impact on the forces exerted on the eye phantom. The `Constant Scaling` mode exhibited not only the highest



(a) Punctuation Force



(b) Total Completion Time



(c) NASA-TLX Subscales

Fig. 6. **Experimental Results from the User Study.** (a) Distribution of applied forces, showing our `LATIOS` method is objectively the safest. (b) Distribution of total task completion times, illustrating the safety-efficiency trade-off. (c) NASA-TLX subscale scores (lower is better), indicating that while unassisted control felt less demanding, users also assessed their own performance as poorest in that mode.

median force but also the greatest variability, including the highest recorded force peak of nearly 200 mN. Introducing the `Linear Scaling` method provided a noticeable improvement, lowering the median force and reducing the spread of the data. Our proposed `LATIOS` method achieved the lowest median force and the tightest interquartile range, indicating the most consistent and safe performance.

TABLE I
MEAN PUNCTUATION FORCE (± STD. DEV.) UNDER LOW AND HIGH LATENCY CONDITIONS.

Setup	**Low Latency [mN]**	**High Latency [mN]**	**Increase [%]**
Constant Scaling	71.6 ± 17.7	93.7 ± 50.8	30.8
Linear Scaling	59.7 ± 15.7	76.5 ± 11.8	28.2
LATIOS	59.9 ± 10.5	61.9 ± 21.6	3.4

Because the experiment followed a within-subjects design, the data were analyzed using a linear mixed-effects model

TABLE II
MEAN PER-TASK TIME (± STD. DEV.) UNDER LOW AND HIGH LATENCY CONDITIONS.

Setup	**Low Latency [s]**	**High Latency [s]**	**Increase [%]**
Constant Scaling	48.4 ± 19.8	57.1 ± 16.7	18.0
Linear Scaling	81.0 ± 32.1	124.3 ± 78.1	53.4
LATIOS	82.0 ± 32.8	114.9 ± 46.5	40.0

with assistance method as a fixed effect and participant as a random intercept. The analysis revealed a significant influence of assistance method on applied force (p = 0.005). Post-hoc comparisons indicated that the `LATIOS` condition produced significantly lower forces than the `Constant Scaling` condition (mean difference = 21.7 mN, 95% CI [6.6 mN, 36.9 mN], $p = 0.005$). To understand the system's performance under challenging network conditions, we analyzed the data separately for low and high latency tasks, with the results summarized in Table I. While forces tended to increase with high latency across all modes, the magnitude of this increase varied dramatically. For the `Constant Scaling` and `Linear Scaling` modes, forces rose by a substantial 30.8% and 28.2%, respectively. In stark contrast, our `LATIOS` method demonstrated remarkable robustness, with forces increasing by a mere 3.4%. This highlights the effectiveness of the latency-aware component of our algorithm in mitigating the primary safety risks of long-distance telesurgery.

This enhanced safety is achieved through a deliberate safety-efficiency trade-off, which is evident in the task completion time data shown in Fig. 6b. The `Constant Scaling` mode was, on average, the fastest. Both assisted methods resulted in longer completion times. Because the study used a within-subjects design, task completion time was analyzed using a linear mixed-effects model with assistance method as a fixed effect and participant as a random intercept. The analysis revealed a significant effect of assistance method on completion time ($p < 0.001$), with the unassisted condition being significantly faster than the `LATIOS` condition. The reason for this increased time becomes clear when analyzing the impact of latency, as detailed in Table II. High latency induced a substantial increase in task time for the assisted modes (`Linear Scaling`: +53.4%, `LATIOS`: +40.0%), while the effect on the `Constant Scaling` mode was less pronounced (+18.0%). This demonstrates that the system functions as intended, enforcing a more cautious and deliberate pace when the combination of high latency and proximity to the retina presents the greatest risk.

Finally, to assess the human experience, we analyzed the NASA-TLX workload scores. The overall workload scores showed no statistically significant difference between the methods ($p = 0.091$). The analysis of the subscales, presented in Fig. 6c, reveals that the `Constant Scaling` mode was rated lowest across most dimensions of workload. This suggests users perceived the direct, unfiltered control as the least demanding. This indicates a potential mismatch between perceived ease and objective safety, highlighting the value of an assistance system that ensures a high safety standard.

V. Discussion

Our transatlantic user study demonstrates that `LATIOS` improves the safety of long-distance telerobotic surgery. The core contribution of this work is the fusion of real-time depth estimation and network latency to create an adaptive safety system.

Our primary finding, a statistically significant reduction in applied punctuation forces ($p = 0.005$), directly addresses the most critical risk in delayed teleoperation: unintended, forceful contact with delicate tissue. The data from Table I is particularly telling: our full `LATIOS` algorithm maintained force levels at a near-constant, safe baseline even when high latency was introduced (3.4% increase), a feat the other modes could not achieve. This confirms the need for safety control that adapts to both robot state and network latency. This objective safety improvement is achieved through a deliberate safety-efficiency trade-off, which was clearly quantified in our time analysis ($p < 0.001$). While `LATIOS` was not the fastest method, this is by design. The system functions as intended by enforcing a slower, more cautious approach under high-risk conditions. Our system is not uniformly slow; it is intelligently and selectively slow only when necessary, a critical feature for maintaining user acceptance and efficiency in less critical phases of the task. The analysis of the subjective metrics provides a crucial window into the human experience. The NASA-TLX data revealed a fascinating insight into the potential for operator overconfidence. The `Constant Scaling` mode was perceived as the least mentally demanding, yet it was objectively the most dangerous. This indicates a potential misalignment between perceived ease and objective safety. Our `LATIOS` system helps to resolve this conflict by enforcing a high standard of objective safety.

It is important to acknowledge the limitations of this study. The user cohort, while diverse, was small, and the experimental task serves as a foundational primitive for evaluating axial depth control and force mitigation. A larger study would be necessary to draw more generalizable conclusions about the effects of surgical expertise. Furthermore, our experiments were conducted on a high-fidelity phantom without trocars to simplify the measurement of punctuation forces. In a real clinical scenario, friction between the tool and the trocar cannula could introduce forces that might interfere with our current contact detection threshold. However, it is important to note that the absence of these confounding forces likely made the unassisted `Constant Scaling` mode feel deceptively controllable. In a clinical setting, the added complexity of trocar friction would likely increase the risk of overshoots, further underscoring the necessity of an intelligent assistance system like `LATIOS` to ensure patient safety. Future work should address this by developing more sophisticated force-sensing algorithms capable of decoupling tool-trocar friction from tool-tissue contact. Finally, the study was limited to a velocity-based mapping between the haptic device and the robot.

VI. Conclusion and Future Work

In this paper, we presented and validated `LATIOS`, a novel adaptive motion scaling system for long-distance telerobotic surgery. By dynamically modulating the robot's velocity based on both shadow-based depth estimation and measured communication delay, our system significantly improves safety, as demonstrated by a statistically significant reduction in punctuation forces. Our user study confirms that this safety is achieved through an intelligent and quantifiable safety-efficiency trade-off. This framework represents an important step toward safe long-distance telerobotic surgery. Looking ahead, we see several promising avenues for research. One key goal is to improve the balance of the safety-efficiency trade-off. We plan to investigate alternative scaling functions, potentially incorporating non-linear curves or predictive models of user intent, to further decrease task completion time while maintaining the high safety standard. Another significant area for improvement lies in the human-robot interface. A future study could explore the use of a specialized haptic device capable of rendering RCM-constrained forces. Such a device would allow for a comparison between the current velocity-based control and a more complex position-based mapping, which could offer a more intuitive control paradigm. Finally, we plan to expand our user study to a larger cohort of clinical professionals to investigate how expertise influences the interaction with our system and to further refine `LATIOS` for eventual clinical translation.

VII. ACKNOWLEDGEMENT

The authors acknowledge the use of Google's Gemini model for manuscript structuring and generating illustrative elements for Fig. 1. All outputs were reviewed and finalized by the authors.

REFERENCES

[1] Jacques Marescaux, Joel Leroy, Michel Gagner, Francesco Rubino, Didier Mutter, Michel Vix, Steven E Butner, and Michelle K Smith. Transatlantic robot-assisted telesurgery. *Nature*, pages 379–380, 2001.

[2] Mehran Anvari. Remote telepresence surgery: the canadian experience. *Surgical Endoscopy*, pages 537–541, 2007.

[3] Song Xu, Manuela Perez, Kun Yang, Cyril Perrenot, Jacques Felblinger, and Jacques Hubert. Determination of the latency effects on surgical performance and the acceptable latency levels in telesurgery using the dv-trainer® simulator. *Surgical endoscopy*, pages 2569–2576, 2014.

[4] M Ali Nasseri, Martin Eder, Suraj Nair, EC Dean, Mathias Maier, Daniel Zapp, Chris P Lohmann, and Aaron Knoll. The introduction of a new robot for assistance in ophthalmic surgery. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5682–5685, 2013.

[5] Xingchi He, Daniel Roppenecker, Dominikus Gierlach, Marcin Balicki, Kevin Olds, Peter Gehlbach, James Handa, Russell Taylor, and Iulian Iordachita. Toward clinically applicable steady-hand eye robot for vitreoretinal surgery. In *ASME International Mechanical Engineering Congress and Exposition*, volume 45189, pages 145–153, 2012.

[6] Amir Molaei, Ebrahim Abedloo, Marc D de Smet, Sare Safi, Milad Khorshidifar, Hamid Ahmadieh, Mohammad Azam Khosravi, and Narsis Daftarian. Toward the art of robotic-assisted vitreoretinal surgery. *Journal of ophthalmic & vision research*, page 212, 2017.

[7] Franziska Ullrich, Christos Bergeles, Juho Pokki, Olgac Ergeneman, Sandro Erni, George Chatzipirpiridis, Salvador Pané, Carsten Framme, and Bradley J Nelson. Mobility experiments with microrobots for minimally invasive intraocular surgery. *Investigative ophthalmology & visual science*, pages 2853–2863, 2013.

[8] Taiga Nakano, Naohiko Sugita, Takashi Ueta, Yasuhiro Tamaki, and Mamoru Mitsuishi. A parallel robot to assist vitreoretinal surgery. *International journal of computer assisted radiology and surgery*, 4:517–526, 2009.

[9] Wei Tech Ang, Cameron N Riviere, and Pradeep K Khosla. Design and implementation of active error canceling in hand-held microsurgical instrument. In *Proceedings 2001 IEEE/RSJ International Conference on Intelligent Robots and Systems. Expanding the Societal Role of Robotics in the the Next Millennium (Cat. No. 01CH37180)*, volume 2, pages 1106–1111, 2001.

[10] Wei Wei, Roger Goldman, Nabil Simaan, Howard Fine, and Stanley Chang. Design and theoretical evaluation of micro-surgical manipulators for orbital manipulation and intraocular dexterity. In *Proceedings 2007 IEEE international conference on robotics and automation*, pages 3389–3395, 2007.

[11] Russell Taylor, Pat Jensen, Louis Whitcomb, Aaron Barnes, Rajesh Kumar, Dan Stoianovici, Puneet Gupta, ZhengXian Wang, Eugene Dejuan, and Louis Kavoussi. A steady-hand robotic system for microsurgical augmentation. *The International Journal of Robotics Research*, pages 1201–1210, 1999.

[12] A Gijbels, Emmanuel B Vander Poorten, Benjamin Gorissen, Alain Devreker, Peter Stalmans, and Dominiek Reynaerts. Experimental validation of a robotic comanipulation and telemanipulation system for retinal surgery. In *5th IEEE RAS/EMBS International Conference on Biomedical Robotics and Biomechatronics*, pages 144–150, 2014.

[13] Wan Ling Wong, Xinyi Su, Xiang Li, Chui Ming G Cheung, Ronald Klein, Ching-Yu Cheng, and Tien Yin Wong. Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: a systematic review and meta-analysis. *The Lancet Global Health*, pages e106–e116, 2014.

[14] TL Edwards, K Xue, HCM Meenink, MJ Beelen, GJL Naus, MP Simunovic, M Latasiewicz, AD Farmery, MD De Smet, and RE MacLaren. First-in-human study of the safety and viability of intraocular robotic surgery. *Nature biomedical engineering*, pages 649–656, 2018.

[15] Junjie Yang, Zhihao Zhao, Siyuan Shen, Daniel Zapp, Mathias Maier, Kai Huang, Nassir Navab, and M Ali Nasseri. Eyels: Shadow-guided instrument landing system for target approaching in robotic eye surgery. *IEEE Robotics and Automation Letters*, pages 3664–3671, 2024.

[16] Junjie Yang, Satoshi Inagaki, Zhihao Zhao, Daniel Zapp, Mathias Maier, Kai Huang, Nassir Navab, and M Ali Nasseri. Shadow maintenance for automatic light-probe control in ophthalmic surgeries using only 2d information. In *2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 7220–7226, 2024.

[17] Takashi Tayama, Yusuke Kurose, Murilo M Marinho, Yuki Koyama, Kanako Harada, Seiji Omata, Fumihito Arai, Koichiro Sugimoto, Fumiyuki Araki, Kiyohito Totsuka, et al. Autonomous positioning of eye surgical robot using the tool shadow and kalman filtering. In *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 1723–1726, 2018.

[18] Simon Pannek, Shervin Dehghani, Michael Sommersperger, Peiyao Zhang, Peter Gehlbach, M Ali Nasseri, Iulian Iordachita, and Nassir Navab. Exploring the needle tip interaction force with retinal tissue deformation in vitreoretinal surgery. In *2024 IEEE International Conference on Robotics and Automation (ICRA)*, pages 16999–17005, 2024.

[19] Demir Arikan, Peiyao Zhang, Michael Sommersperger, Shervin Dehghani, Mojtaba Esfandiari, Russell H Taylor, M Ali Nasseri, Peter Gehlbach, Nassir Navab, and Iulian Iordachita. Real-time deformation-aware control for autonomous robotic subretinal injection under ioct guidance. In *2025 IEEE International Conference on Robotics and Automation (ICRA)*, pages 10531–10537, 2025.

[20] Demir Arikan, Mojtaba Esfandiari, Peiyao Zhang, Michael Sommersperger, Shervin Dehghani, Russell H Taylor, M Ali Nasseri, Peter Gehlbach, Nassir Navab, and Iulian Iordachita. Towards motion compensation in autonomous robotic subretinal injections. In *2025 International Symposium on Medical Robotics (ISMR)*, pages 66–72, 2025.

[21] Florian Richter, Ryan K Orosco, and Michael C Yip. Motion scaling solutions for improved performance in high delay surgical teleoperation. In *2019 International Conference on Robotics and Automation (ICRA)*, pages 1590–1595, 2019.

[22] Ryan K Orosco, Benjamin Lurie, Tokio Matsuzaki, Emily K Funk, Vasu Divi, F Christopher Holsinger, Steven Hong, Florian Richter, Nikhil Das, and Michael Yip. Compensatory motion scaling for time-delayed robotic surgery. *Surgical endoscopy*, pages 2613–2618, 2021.

[23] Jason Lim, Florian Richter, Zih-Yun Chiu, Jaeyon Lee, Ethan Quist, Nathan Fisher, Jonathan Chambers, Steven Hong, and Michael C Yip. Optimal motion scaling for delayed telesurgery. *arXiv preprint arXiv:2506.21689*, 2025.

[24] Korab Hoxha, Alireza Alikhani, Satoshi Inagaki, Manuel Ferle, Mathias Maier, and M Ali Nasseri. Modelling and development of a mechanical eye for the evaluation of robotic systems for surgery. In *2023 45th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, pages 1–4, 2023.

[25] Andries F Sanders and Andries Sanders. *Elements of human performance: Reaction processes and attention in human skill*. Psychology Press, 2013.

[26] Thomas B Sheridan. Space teleoperation through time delay: Review and prognosis. *IEEE Transactions on robotics and Automation*, pages 592–606, 2002.