Human Systems Engineering and Design (IHSED 2026), Vol. 237*,* 2026, 843–853
https://doi.org/10.54941/ahfe1008222

Tactile Feedback for Controlling a Robot With Hand Gestures

Peter Schmid, Celine Bartschinski, Jonathan Kiessling, and Thomas Maier

Institute for Engineering Design and Industrial Design, Dept. Industrial Design Engineering, University of Stuttgart, Stuttgart, Germany

ABSTRACT

This study investigates contactless gesture control combined with tactile ultrasound feedback as an alternative interaction method for controlling a robotic arm. A gesture alphabet was developed for translational and grasping movements. Hand motions were captured using a Leap Motion Controller, while tactile feedback forms were projected onto the user's palm. 5 feedback forms were designed and evaluated regarding perceptibility, mental workload, directional changes, gripper opening width, task fulfillment, and operating time. The study included 32 participants, who first assessed the isolated feedback forms and then performed a grasping and placement task using a robotic arm. The results show that simple geometric forms, particularly points and lines, are perceived and interpreted more reliably than complex shapes such as triangles or helix. A line as Feedback form achieved the most favourable results across several variables, including directional perception, lower mental workload, and shorter average operating times. Rotational coding proved less intuitive and increased mental load, especially when used to represent linear robot movements. Overall, the findings indicate that tactile ultrasound feedback can support contactless robot control, provided that feedback forms are simple, continuous, and spatially intuitive. Future work should examine blind operation, larger samples, and person-specific adaptation based on tactile sensitivity.

Keywords: Human-robot-interaction, Huma-machine-interface, Tactile feedback, Usability, information coding

INTRODUCTION

In 1986, the first robot-assisted surgical procedure was performed (Lanfranco et al., 2004), and with the introduction of the Da Vinci robot in 2000, robot-assisted surgery was revolutionized (Diana and Marescaux, 2015). Nevertheless, the issue of practical, sterile operation remains unresolved so far, meaning that significant development opportunities still exist in the market. To date, three different solutions are available on the market. (Chojecki and Leiner, 2009). On the one hand, sterile personnel can control non-sterile assistance via voice commands. Alternatively, all surfaces can be covered with sterile disposable films. As a third option, the surgeon can remotely control sterile instruments attached to robotic arms from the non-sterile area. However, none of these solutions is optimal. On the one hand, voice commands can lead to misunderstandings during transmission, thereby

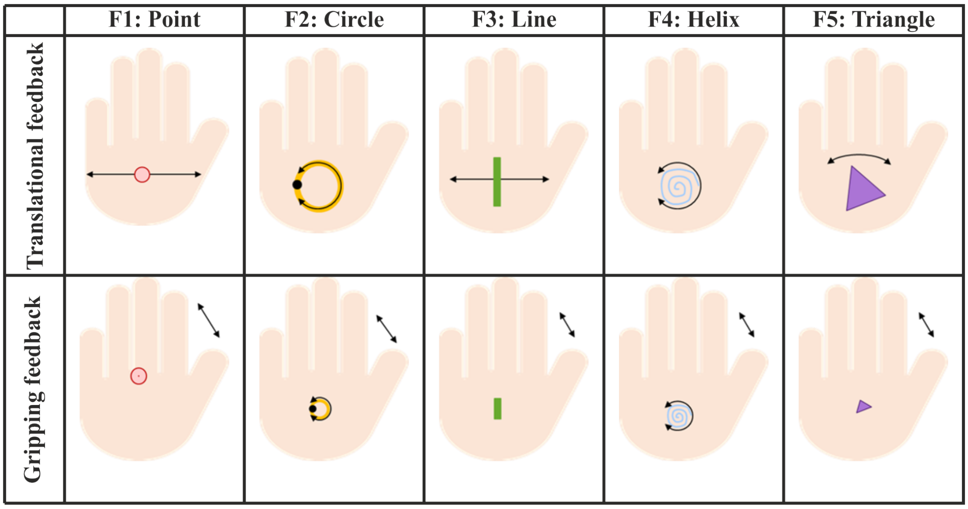
Received May 12, 2026 Revised June 22, 2026 Accepted July 11, 2026 Available online September 25, 2026

hindering a smooth workflow and putting the patient at risk (Christian et al., 2006; Lingard et al., 2004). On the other hand, there is already a high level of background noise in an operating room due to the multitude of technical devices and conversations among staff. De Lima et al. (2021) demonstrated that noise levels in hospitals exceed the threshold classified as safe by the World Health Organization. Furthermore, sterile disposable sheets must be changed after each use, making this approach time-consuming and costly. Remote control from the non-sterile area becomes problematic as soon as unexpected complications, such as uncontrollable bleeding or cardiac arrest, occur. In robot-assisted surgery, 2.8% of cases require a switch to open surgery (Shah et al., 2022). This necessitates the surgeon approaching the operating table and thus entering the sterile field. The challenges outlined above highlight the need for a contactless method of interaction between hospital staff and a robot. In this context, "contactless" refers to the absence of physical contact with an object (Zimmermann, 2011). Head and eye movements are perceived as unintuitive and disruptive, which is why they may hinder smooth and efficient control (Hamza et al., 2023; Aaltonen and Wahlström, 2018). Mariappan et al. (2023) have already achieved positive results in the real-time gesture control of a robotic arm. Furthermore, Korayem et al. (2021) and Vargas and Vivas (2014) investigated this interaction method in a surgical setting. However, previous studies have not integrated clear feedback. The user receives only visual feedback by looking at the robotic arm. Multimodal perception, e.g., through the combination of visual and tactile feedback, can lead to shorter reaction times and lower error rates (Fitch et al., 2011; Prewett et al., 2006). Therefore, the goal of this work is to investigate the combination of gesture control with tactile feedback.

METHOD

To investigate the suitability for transmitting information during user-controlled robot operation, a gesture alphabet for hand movements is conducted. The gesture alphabet for hand movements includes translation, pronation, and supination gestures in all three spatial directions of the Cartesian coordinate system, as well as the grasping motion of the thumb and index finger. The rotational movements of the robotic arm are derived from the orientation of the palm, while the distance between the thumb and index finger additionally controls the opening width of the end-effector. The translational feedback serves to provide the user with information about the robot's direction of movement resulting from the executed hand gesture and, at the same time, to enable an estimation of the robot's current position in space. As soon as the user enters the interaction space with their hand, the palm is detected by the Leap Motion Controller, whereupon the feedback is projected onto the palm as a continuous signal. When the hand moves, the feedback is adjusted accordingly. The grip feedback serves as a source of information for the user regarding the current opening angle of the robot's end-effector. The closer the index finger and thumb are moved together, the further the gripper closes. Since both types of movement can occur simultaneously during interaction, consistent and clearly distinguishable coding of the feedback information is necessary. In this context, shape is the only constant feature, as it cannot be dynamically altered during the ongoing interaction. To ensure that the shapes can be perceived in the palm by all body size classes, the metric data for the feedback shapes are calculated based on the body measurements of a 5^{th} percentile woman. The use of complex shapes and combinations of shapes is avoided. The selection of shapes is primarily based on the findings of existing studies. These studies indicate that dynamic feedback forms, in particular, are easily perceptible. In this context, a moving point along a line and a circular path are specifically recommended as suitable options (Freeman and Wilson, 2021). These are therefore included in the solution space. In addition, an arrow is considered, as it has an established directional function. Shapes such as lines and rectangles are also provided as predefined sensations in Unity and are likewise included in the solution space. To ensure the comparability and traceability of the results, parameters such as frequency and intensity are kept constant for the time being. The information transmission itself remains simple and is limited to the relevant characteristics of shape, position, and their changes as the palm moves over the ultrasound array. A total of 5 feedback shapes were derived. Table 1 shows the variants for both translational and grasping movements.

Table 1: Feedback forms of the study.



DESIGN OF THE STUDY

To investigate the suitability of the gesture alphabet, a comparative evaluation is carried out, taking into account both the participants' subjective assessments and objectively measurable data. Throughout the study, the participant sits at a workstation (see Figure 1). Directly in front of them is the Ultrahaptics Stratos Explore Development Kit. Behind it is the Adeept 5-DoF robotic arm, whose end-effector extends into the workspace. After completing the task, participants provide a subjective evaluation of the set feedback using a questionnaire. The robot used is a serial articulated-arm robot of the Adeept 5-DoF type. It consists of a kinematic chain of rigid links connected by a total of five joints. Each of these joints allows for relative movement with respect to the others and contributes to the overall motion of the robot arm (Adeept, 2025).

Each servomotor has a nominal operating range of 0° to 180°. An exception is the servomotor of the gripper, whose usable range of motion is limited to 0° to 90°, since the gripper is already fully closed at a deflection of 90°. The Leap Motion Controller, which is attached to the top edge of the Ultrahaptics, is used to track the hand.

At the beginning of the study, demographic data from the participants are collected. This is followed by a briefing on the experimental setup and the conduct of the study. The study is divided into two parts. In the first part of the study, the feedback types are projected one after another, in a randomized order, onto the participants' palms. The goal is to give participants the opportunity to familiarize themselves with the feedback shapes and the associated tactile sensations. At the same time, the first evaluation of the individual feedback shapes takes place, focusing on their perceptibility. In the second part of the study, the participants perform a grasping task using the robot. The task consists of grasping the object in Area 1 with the robotic arm, then placing it in Area 2, grasping it again, and returning it to the starting position (see Figure 1). Meanwhile, the feedback forms from Table 1 are projected onto the palm to convey information about the direction of the robotic arm's movement and the gripper's opening width.

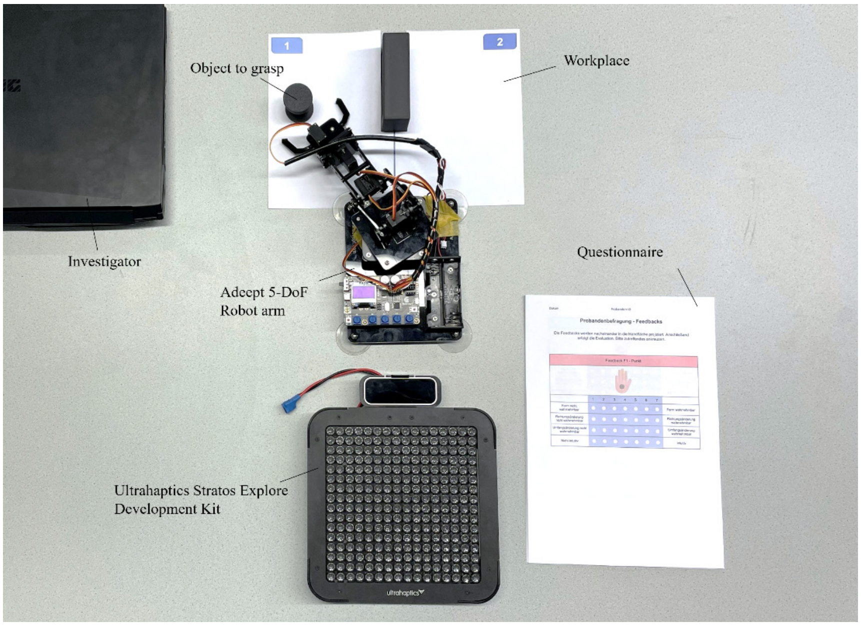


Figure 1: Experimental setup as seen by the subject, featuring a robotic arm and ultrahaptics.

Test Person

The study sample consists of a total of 32 participants. The average age is 28.19 years, with a range of 21 to 57 years. Men account for 53.1% of the sample, and women for 46.9%. The majority are right-handed (87.5%), while 12.5% are left-handed. There are no known impairments of the sense of touch. Regarding smoking status, 96.9% are non-smokers and 3.1% are smokers. 46.9% of the study participants reported having had previous experience with ultrasound haptics.

RESULTS

The collected data are presented descriptively using box plots, supplemented by an inferential statistical analysis to test for significant differences among the five types of feedback examined. Within the corresponding figures, the statistical significance of the feedback types is indicated by asterisks (*$p < 0.05$; **$p < 0.01$, ***$p < 0.001$).

Figure 2a shows the task fulfillment. The object could be transported along the specified path using any of the feedback forms. Only the F4 feedback form produced three errors. This was caused by the fact that the robot's gripper could not be fully closed despite the pinch gesture being executed. As a result, the object was dropped and could not be transported to its final destination.

Figure 2b presents the ratings for pure shape recognition without robot control using box plots. Feedback form F2 has the highest values (Ø = 6.37, SD = 1.21), followed by feedback form F3 (Ø = 6.12, SD = 0.98). Feedback form F1 falls in the middle range (Ø = 5.31, SD = 1.33), while significantly lower ratings are evident for F5 (Ø = 3.72, SD = 1.61) and, in particular, F4 (Ø = 2.19, SD = 1.31). The interquartile ranges indicate comparatively low dispersion for F2 and F3. Additionally, there are outliers in F3 and an extreme outlier in F2. The inferential statistical analysis using the Friedman test reveals significant differences among the five feedback types (χ^2 (4) = 86.27, $p < .001$, W = .67, n = 32). Bonferroni-adjusted post hoc comparisons are performed to precisely identify the significant differences. These reveal highly significant differences in shape recognition for F4-F1 (z = 4.822, $p < .001$, n = 32), F4-F3 (z = 6.997, $p < .001$, n = 32), and F4-F2 (z = 7.550, $p < .001$, n = 32), with form F4 being recognized significantly worse in all cases. Highly significant differences also emerge between the feedback forms F5-F3 (z = 4.625, $p < .001$, n = 32) and F5-F2 (z = 5.178, $p < .001$, n = 32), with F3 and F2 being more easily recognized in each case.

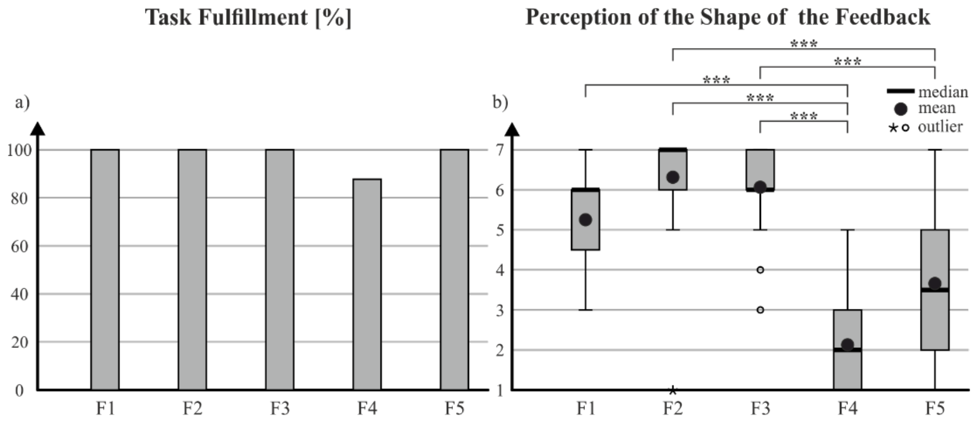


Figure 2: a) Task Fulfillment and b) perception of the shape of the feedback (1 = [barely perceptible]; 7 = [clearly perceptible]).

The operating times for the grasping task using the respective feedback types are presented in Figure 3a in the form of box plots. The longest operating times are observed for the two feedback forms F5 (Ø = 17.20, SD = 5.72) and F4 (Ø = 17.64, SD = 5.35). Feedback forms F1 (Ø = 17.15, SD = 4.11) and F2 (Ø = 17.81, SD = 4.85) are slightly below this, but differ only marginally. Feedback form F3 has the shortest average operating time (Ø = 15.11, SD = 14.08). In particular, feedback forms F1, F2, and F5 show outliers in the upper range of operating times. A repeated-measures General Linear Model was calculated for inferential statistical analysis. The results revealed a significant main effect of the feedback type factor ($P = .303$, $p < .05$, $F = .303$, $n = 32$), which can be classified as a moderate effect. However, the Bonferroni-corrected pairwise comparisons did not reveal any significant differences between the individual feedback forms.

Figure 3b illustrates the mental load by the respective form of feedback in terms of task performance and attention to the feedback. Descriptively, there are small but noticeable differences in perceived mental load between the feedback forms. The highest values are found for F2 (Ø = 4.47, SD = 1.74) and F5 (Ø = 4.59, SD = 1.85). The feedback type F1 falls within the range of moderate strain (Ø = 4.16, SD = 1.97). Feedback form F4 also shows comparable results (Ø = 4.47, SD = 1.89). The least stressful form of information transmission is achieved through feedback F3 (Ø = 3.36, SD = 1.89). Also striking is the wide variation in results across all five feedback forms. Feedback form F1, in particular, shows a large range between the values of 0 and 9. Furthermore, feedback form F3 includes an outlier with a value of 8. The Friedman test also shows that these differences are significant (χ^2 (4) = 15.22, $p < .01$, $W = .119$, $n = 32$). The effect is localized within a small range. In the pairwise comparisons with Bonferroni correction, a significant difference is observed between F3 and F4 ($z = -3.044$, $p < .05$, $n = 32$), with F4 being rated as significantly more demanding.

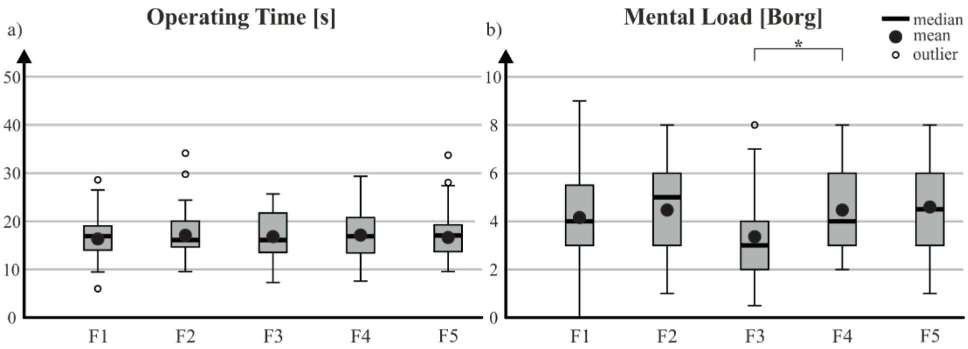


Figure 3: a) Operating time and b) Mental load according to Borg during the rasping task with the robot.

Figure 4a shows the distribution of ratings regarding the perception of changes in direction during robot control using box plots. Feedback form F3 has the highest mean values (Ø = 5.66, SD = 1.18). Feedback forms F1 (Ø = 4.03, SD = 1.84) and F2 (Ø = 3.78, SD = 1.68) fall within the middle range of ratings, while feedback form F4 (Ø = 3.44, SD = 1.66) and F5 (Ø = 2.88, SD = 1.74) have lower values. The Friedman test revealed a significant effect (χ^2 (4) = 40.87, p < .001, W = .319, n = 32), which should accordingly be interpreted as a moderate to large effect. Consistent with the descriptive analysis of feedback form F3, the pairwise comparisons reveal highly significant differences between F5-F3 (z = 5.574, p < .001, n = 32) and F4-F3 (z = 4.862, p < .001, n = 32). Highly significant differences are evident between feedback forms F2 and F3 (z = –3.716, p < .01, n = 32) as well as F1 and F3 (z = 3.439, p < .01, n = 32). The results consistently indicate that the change in direction of feedback form F3 is perceived significantly better during robot control in all cases mentioned.

Figure 4b shows the ratings for the perception of changes in gripper size during opening and closing while the robot is being controlled. With the exception of feedback type F4, the ratings for the remaining feedback types are very close to one another. The medians for feedback types F1, F2, F3, and F5 are at a comparable level in the upper to middle rating range. Similarly, the mean values differ only slightly from one another (F1: Ø = 4.37, SD = 2.14; F2: Ø = 4.75, SD = 1.76; F3: Ø = 4.69, SD = 1.60; F5: Ø = 4.47, SD = 1.65). Feedback F4 stands out, as it falls within the lower rating range in both the median and the mean (Ø = 3.53, SD = 1.69). Overall, the differences are evident less in the central tendencies and more in the dispersion of the ratings. The inferential statistical analysis using the Friedman test reveals a significant effect between the feedback types (χ^2 (4) = 11.75, p < .05, W = .092, n = 32). The Bonferroni-corrected post hoc comparisons indicate that F2 was rated significantly better than F4 in terms of the change in circumference during robot control. This is demonstrated by the significance values of the F4-F2 post hoc comparison (z = 2.965, p < 0.05, n = 32).

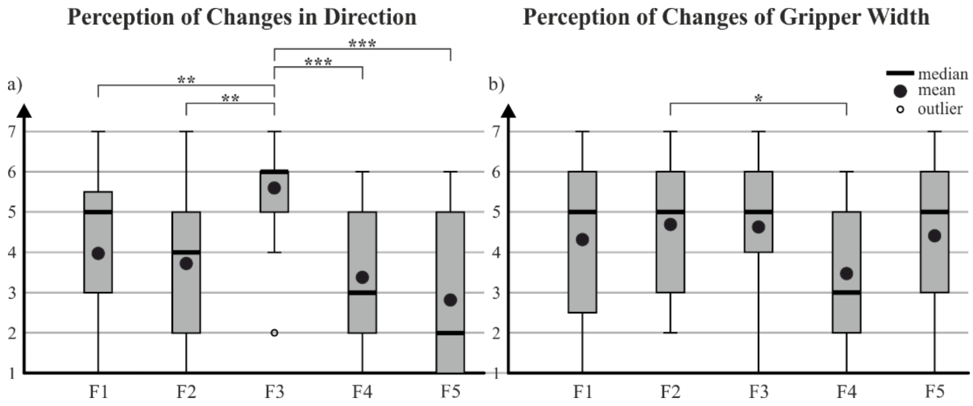


Figure 4: a) Perception of changes in direction and b) perception of changes in gripper width (1 = [barely perceptible]; 7 = [clearly perceptible]).

DISCUSSION

The results of the study show that the perceptibility and usability of tactile feedback as a means of information transmission depend largely on its geometric structure, the spatial resolution on the palm, and the motion logic. The results show that simple forms such as points, circles, and lines (F1–F3) achieve higher median values in shape recognition in the first part of the study than complex shapes such as triangles or helix (cf. Figure 2b). This result could be explained by the limited two-point discrimination ability of the palm. If the individual stimulus points are too close together, they are no longer perceived separately from one another but merge into a flat perception. In the case of complex shapes such as triangles and helix, the distances between the stimulus points are apparently too small to clearly distinguish and perceive distinct edges or changes in direction. Instead of a clearly defined shape, the impression is one of vague, widespread feedback. This explains the low median values for shape recognition in F4 and F5 compared to F2 and F3 (cf. Figure 2b).

Another aspect is the logic of movement. In the dynamic feedback forms F2 and F5, a rotational movement of the feedback is coupled with a translational movement of the robot arm (cf. Figure 4a). From a perceptual point of view, this association is not very intuitive, as there is no direct spatial relationship. In contrast, linear movements, as seen in F1 and F3, represent a clear spatial analogy. Interpreting a direction of rotation requires additional cognitive effort.

In addition, the dynamics of the feedback forms (e.g. F2, F4, F5) with regard to the results should be discussed. The dynamic feedback form F2 presents a mixed impression. On the one hand, their movement increases sensory salience and leads to more intense perception (cf. Figure 2 or 4b). On the other hand, they result in increased mental load, particularly in the case of rotational coding (cf. Figure 3b). The same applies to feedback form F5. While directional perception declines significantly, the perception of changes in gripper size remains comparably good.

However, the processing of directional information varies more significantly across the different feedback forms. In this regard, the continuity of the tactile stimulus is worth noting. In F4 and F5, the size of the feedback form decreases during grasping until, in some cases, it is no longer perceived. As a result, relevant information is lost. In contrast, a continuous stimulus is maintained in F1, F2, and F3, keeping the information content more constant (cf. Figure 4a). Additionally, the hand position during the gripping gesture can interfere with the Leap Motion Controller's detection of the hand's center of mass, which can lead to the interruption of the feedback projection.

Nevertheless, the information content of the directional change is rated lower. This can be explained, on the one hand, by the lack of information regarding the exact position of the end-effector. On the other hand, attention must be directed toward the grasping task in parallel with the feedback. This is more difficult with changes in direction than with changes in the gripping size. To sum up, feedback form F3 has the highest information content due to its simple structure.

According to the Bonferroni-corrected comparisons, the execution time shows no significant differences, but overall exhibits a relatively large variation in results. This is likely due to the randomized study design as well as learning and habituation effects. On average, F3 shows the shortest execution times, while execution under F4 and F5 requires more time. Several participants reported that F3 made the task easier, which can be attributed to the visual similarity of the line to the appearance of the robot arm.

CONCLUSION

In general, the five feedback forms are very good to distinguish. However, the triangle and helix are more frequently confused, while the remaining feedback forms are easily distinguishable. Feedback F3, in particular, can be considered the reference design, as it achieved the most valid results across nearly all variables examined. To sum up, the following design recommendations can be derived:

- Point or line as the basic element
- Avoid rotational coding for linear movements (circle, helix)
- Continuous stimulus presentation
- Combination of feedback with changes in frequency and/or intensity, but to a limited extent, to keep mental load low

Nevertheless the study revealed some technical and ergonomic limitations that may have influenced the results. From a technical perspective, the use of more powerful servo motors should be considered. This would increase the stability and precision of the robot's movements. Another limitation concerns the intensity of the ultrahaptics used, which was at times perceived as too low, depending on the type of feedback. Based on the identified limitations, there are several starting points for further development emerge for future work. The approaches described should be examined in further applied studies, particularly by including a larger sample size, in order to increase the validity of the results. Since this study was conducted with the robot in view, the aspect of blind operation should be taken into account in subsequent studies. During the course of the studies, the measurement of two-point discrimination should be considered to determine correlations with the shapes and the transmission of information. This would allow for a person-specific adjustment of the shapes within the study.

REFERENCES

Aaltonen, I. E. und Wahlström, M. (2018). “Envisioning robotic surgery: Surgeons’ needs and views on interacting with future technologies and interfaces”, The International Journal of Medical Robotics and Computer Assisted Surgery, Bd. 14 Nr. 6.

Adeept. (December 20, 2025). Adeept 5-DOF Robotic Arm Kit. Website: https://www.adeept.com/adeept-5-dof-robotic-arm-kit_p0360.html

Chojecki, P., Leiner, U. (2009). “Berührungslose Gestik-Interaktion im Operationssaal”, Journal of Interactive Media, Bd. 8 Nr. 1.

Christian, C. K., Gustafson, M. L., Roth, E. M., Sheridan, T. B., Gandhi, T. K., Dwyer, K., Zinner, M. J., Dierks, M. M. (2006). A prospective study of patient safety in the operating room. Surgery, Bd. 139 Nr. 2.

De Lima, E. A., da Cunha e Silva, D. C., de Oliveira, R. A., Zannin, P. H. T., Martins, A. C. G. (2021). “Environmental noise in hospitals: A systematic review”, Environmental Science and Pollution Research, Bd. 28 Nr. 16.

Diana, M., Marescaux, J. (2015). “Robotic surgery”, British Journal of Surgery, Bd. 102 Nr. 2.

Fitch, G. M., Hankey, J. M., Kleiner, B. M., Dingus, T. A. (2011). “Driver comprehension of multiple haptic seat alerts intended for use in an integrated collision avoidance system”, Transportation Re-search Part F: Traffic Psychology and Behaviour, Bd. 14 Nr. 4.

Freeman, E., Wilson, G. A. (2021). “Perception of Ultrasound Haptic Focal Point Motion”, in: Proceedings of the 2021 International Conference on Multimodal Interaction (ICMI ‘21). Association for Computing Machinery, New York.

Hamza, H., Baez, V. M., Al-Ansari, A., Becker, A. T., Navkar, N. V. (2023). “User interfaces for actuated scope maneuvering in surgical systems: a scoping review“, Surgical Endoscopy, Bd. 37 Nr. 6.

Korayem, M. H., Madihi, M. A., Vahidifar, V. (2021). “Controlling surgical robot arm using leap motion controller with Kalman filter”, Measurement, Bd. 178.

Lanfranco, A. R., Castellanos, A. E., Desai, J. P., Meyers, W. C. (2004). “Robotic Surgery: A Current Perspective“, Annals of Surgery, Bd. 239 Nr. 1.

Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G. R., Reznick, R., Bohnen, J., Orser, B., Doran, D., Grober, E. (2004). “Communication failures in the operating room: An observational classification of recurrent types and effects”, Quality & Safty in Health Care, Bd. 13 Nr. 5.

Mariappan, R., Gayathri, P., Pushpalatha, P., Sri Sidhvi Rishik, V., Satish, T. (2023). “Real Time Robotic Arm using Leap Motion Controller”, Journal of Physics: Conference Series, Bd. 2466 Nr. 1.

Prewett, M. S., Yang, L., Stilson, F. R. B., Gray, A. A., Coovert, M. D., Burke, J., Redden, E., Elliot, L. R. (2006). “The benefits of multimodal information: a meta-analysis comparing visual and visual-tactile feedback”, Proceedings of the 8th international conference on Multimodal interfaces, Association for Computing Machinery, Banff.

Shah, P. C., de Groot, A., Cerfolio, R., Huang, W. C., Huang, K., Song, C., Li, Y., Kreaden, U., Oh, D. S. (2022). "Impact of type of minimally invasive approach on open conversions across ten common procedures in different specialties", Surgical Endoscopy, Bd. 36 Nr. 8.

Vargas, H. F., Vivas, O. A. (2014). "Gesture recognition system for surgical robot's manipulation", 2014 XIX Symposium on Image, Signal Processing and Artificial Vision.

Zimmermann, J. L. (2011). "A Methodology for the Design of Non-Contact Handling Systems", Herbert Utz Verlag, München.