CrossMark

Comparison of six display modes for a multi-resolution foveated laparoscope

Sangyoon Lee[1] · Hong Hua[1] · Mike Nguyen[2] · Allan J. Hamilton[3]

Abstract

Background To overcome the field of view and ergonomic limitations of standard laparoscopes, we are developing a multi-resolution foveated laparoscope that can simultaneously obtain both wide- and zoomed-in-view images through a single scope. To facilitate the effective access to the dual views of images with different resolution and field coverage acquired by our laparoscope, six different display modes have been developed. Each of the six display modes has inherent advantages and disadvantages. This study compares the six display modes through a human-subject experiment, which was conducted with an emulated laparoscope using a 4K camera.

Methods Twenty-four subjects without medicine background performed three evaluation trials of a touching task with each of the display modes. Various objective measurements including task completion time, the number of collisions, gaze position, and tooltip position, and subjective preference for the display modes were recorded.

Results For all the measurements except for task completion time and moving speed of tooltip, there were statistically significant differences among the display modes. Although the focus plus warped context view mode was selected as one of the least preferred modes, it showed the best task performance.

Conclusions The unblocked wide context view was useful to provide a situational awareness even when it was severely distorted in some of the display modes, and information continuity played an important role in improving task performance. Moreover, the position change of viewing window coupled to the location of region of interest helped improve task performance, by providing an additional cue for spatial awareness.

Keywords Laparoscope · Display interface · Dual views · Multi-resolution visualization · Focus + context · Overview + detail

Laparoscopy has been widely utilized as the most successful means for minimally invasive surgical procedures. Compared to open surgery, it provides many advantages and benefits to patients, including reduced pain, shorter recovery and hospital stay time, and reduced cost [1, 2]. From the surgeon's perspective, however, it has several disadvantages or limitations [2, 3]. One of its major limitations is a trade-off between limited field of view (FOV) for high spatial resolution versus wide FOV for situational awareness but with diminished resolution [2–4]. Standard laparoscopes lack the ability to acquire both wide-angle and high-resolution images simultaneously through a single scope.

In the current clinical practice, the FOV limitation is addressed by manually moving the entire laparoscope in and out of the camera port to obtain either close-up views for details or wide-angle overviews for orientation. This practice requires a trained assistant for holding and maneuvering the camera. The practice of frequently maneuvering the camera using a trained assistant can introduce ergonomic conflicts with hand cross-over between the surgeon and the assistant holding the camera [5].

✉ Sangyoon Lee
  sylee@optics.arizona.edu

✉ Hong Hua
  hhua@optics.arizona.edu

1   College of Optical Sciences, University of Arizona, 1630 E University Blvd, Tucson, AZ 85721, USA

2   USC Institute of Urology, Keck School of Medicine, University of Southern California, Los Angeles, USA

3   Division of Neurosurgery, Department of Surgery, Arizona Simulation Technology and Education Center, Arizona Health Sciences Center, University of Arizona, Tucson, USA

To overcome the FOV and ergonomic limitations of standard laparoscopy, robotically assisted techniques, such as voice, foot pedal, or head motion-activated cameras, have been developed. However, these techniques suffer from delays in task performance due to errors in voice recognition or robotic control of camera speed, and requires significant practice to become efficient with setup and use [6]. There have also been prior attempt to develop cameras that have a low external profile with high definition (HD) picture and automatic focusing [7]. These scopes, however, still require advancing and withdrawing of the lens to obtain magnification, which can lead to problems with inadvertent and restricted movement of the laparoscope due to collisions with other instruments internally and issues with hand collisions externally [8]. It has also been suggested that varying the magnification of the laparoscope and making it a low profile could reduce the effect of crowding [9], but this approach alone could compound the problem of loss of situational awareness from zoomed-in surgery. A better approach to laparoscopic imaging which balances high spatial resolution with preservation of situational awareness is needed.

As an alternative and better approach to existing ones, we are developing a multi-resolution foveated laparoscope (MRFL). Figure 1A shows a schematic layout of an MRFL in clinical use, and Fig. 1B shows a schematic optical layout of our MRFL design consisting of two imaging probes, one for wide-angle and one for high-resolution, which are fully integrated through shared objective and relay lens groups. The wide-angle probe captures a wide-FOV surgical field, while the high-resolution probe obtains images of a subregion of the wide-angle field at a high resolution. A two-axis optical scanner steers the high-resolution field toward a selected region of interest (ROI) within the wide-angle field [10]. Overall, our MRFL technology can (1) simultaneously obtain both wide-angle and high-magnification (or optically zoomed) images of a surgical area in real-time in a single, fully integrated scope, (2) yield ultra-high spatial resolution which is over three times better than a standard laparoscope at a close-up working distance, (3) automatically scan and engage the high magnification probe to any sub-region of the surgical field through region-of-interest tracking capabilities, (4) vary the optical magnification of the high-resolution probe without the need of physically advancing or withdrawing the scope, and (5) maintain a low-length profile, optimized for operating at a long working distance, providing both superb spatial awareness and high-magnification views simultaneously. Figure 1C shows two prototypes of a MRFL design, one with a normal full-length profile and one with low-length profile where the wide-angle probe captures an 80° FOV and the high-resolution probe covers a 26° FOV of interest. The MRFL optics was optimized at a 120 mm working distance with working range of 80–180 mm. At a 120 mm working distance, the wide-angle probe captures a surgical field of $\sim 150 \times 112$ mm^2, which is more than nine times as large as that of a standard laparoscope under
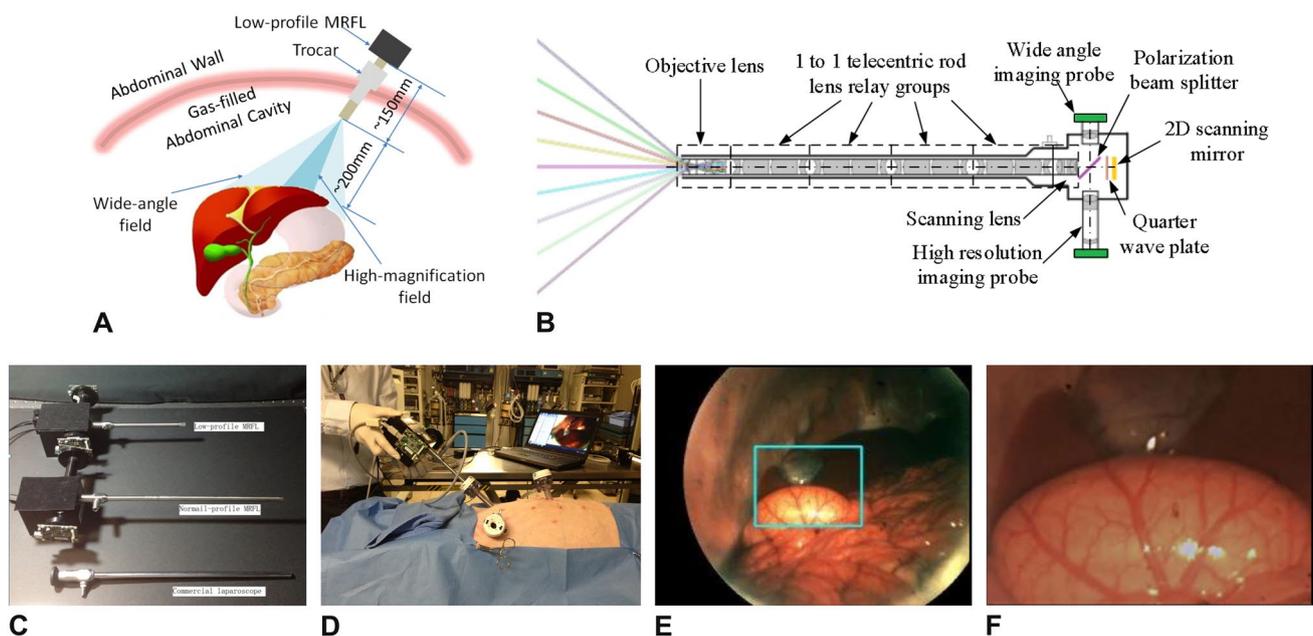

Fig. 1 Demonstration of a multi-resolution foveated laparoscope (MRFL) for minimally invasive surgery. **A** Schematic layout of MRFL in clinical use. **B** Schematic optical layout of a MRFL design. **C** MRFL prototypes in comparison to a standard laparoscope. **D** Clinical setup for testing the MRFL prototypes with a porcine model. **E, F** A wide-angle view and a zoomed-in view captured by the wide-angle and high-resolution imaging probes, respectively

comparison at a typical 50 mm working distance, while the foveated probe captures a surgical field of about 45×35 mm^2, equivalent to the coverage of a standard scope and yields a 7 lps/mm resolution, about three times as good as a standard laparoscope under comparison. Figure 1D shows the setup of clinical test of the prototype with a live porcine model at the live animal lab in the Keck School of Medicine at the University of Southern California, while Fig. 1E and F shows the wide-angle and high-resolution views captured by the scope. The Cyan-color box in Fig. 1E marked the corresponding ROI captured by the foveated probe for Fig. 1F.

Since our MRFL can simultaneously obtain both wide-angle and narrow-angle zoom images, it is imperative to investigate and understand how to effectively display the dual-view multi-resolution images using limited display resources in a clinical environment where a single monitor is preferred due to space constraints. Inspired by the previous work regarding the display interface mechanisms for information visualization [11] and our previous research on multi-scale visualization interfaces for 3D displays [12], we have implemented six different display modes for display interfaces of our MRFL, namely overview plus detail (o+d), picture-in-picture overview plus detail (o+d (pip)), focus plus occluded context (f+oc), fixed focus plus occluded context (ff+oc), focus plus warped context (f+wc), and fixed focus plus warped context (ff+wc) modes. These

display modes are demonstrated in Fig. 2A–F, respectively. They mainly differ in terms of window management and positioning the zoom-view window relative to the viewer. Note that in all the display modes, the zoomed-in area in the surgical field is automatically determined by the position of instrument which is automatically tracked in the wide-view image using computer vision techniques [13] and is considered as a center of the region of interest.

In the o+d mode, the wide and zoom views are presented on two separate windows side-by-side (Fig. 2A). The zoom-view window is substantially larger than the wide-view window to provide surgeons a centralized focus on a surgical field. In the current setup, the ratio between the two window sizes is set to 2:1 in both horizontal and vertical directions. In the o+d (pip) mode, the wide-view window is overlaid on the upper-left corner of the zoom-view window with the size ratio of 3:1 (Fig. 2D). This and the remaining modes were adapted from the well-known picture-in-picture technique, embedding a view window inside another view window. In both modes, a cyan-colored box is drawn inside the wide-view window to mark the ROI determined by the tracking algorithm for the corresponding zoom view, while a purple-colored box is drawn to highlight the boundaries of the wide-view window in the o+d (pip) mode. The o+d and o+d (pip) modes are intuitive, but the o+d mode requires a user to fully switch his/her attention between the two views since the two view windows
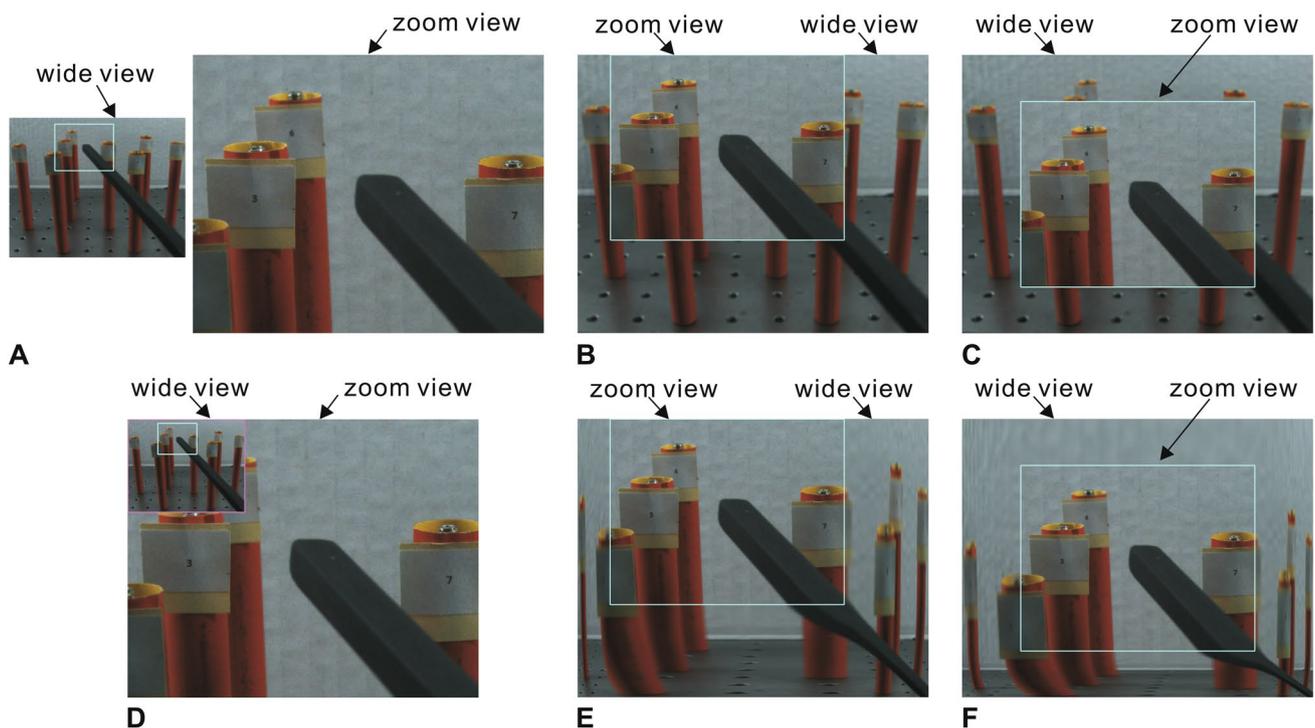

Fig. 2 Demonstration of display modes developed for our MRFL. **A** Overview plus detail (o+d) mode. **B** Focus plus occluded context (f+oc) mode. **C** Fixed focus plus occluded context (ff+oc) mode. **D** Picture-in-picture overview plus detail (o+d (pip)) mode. **E** Focus plus warped context (f+wc) mode. **F** Fixed focus plus warped context (ff+wc) mode

are physically apart, and the o+d (pip) mode requires to block a specific area of the zoom view.

In the f+oc and ff+oc modes, the zoom-view window is overlaid on the wide-view window (Fig. 2B, C, respectively). In these modes, the wide-view window is larger than the zoom-view window (size ratio of zoom-view window to wide-view window = 2:3 in both directions). In the f+oc mode, the zoom-view window (i.e., focus) is embedded inside the wide-angle view window (context) and its position follows the position of instrument (Fig. 2B). In contrast, in the ff+oc mode, the zoom-view window is fixed at the center of the wide-view window (Fig. 2C). The f+oc mode presumably imposes minimal cognitive tolls on correlating the spatial relationship of the two views. However, the position change of the zoom-view window may impose an adverse effect because the viewer has to visually track the location of the zoom-view window. The ff+oc mode may adequately address the adverse effect of the f+oc mode. To display the foveated zoom-in view at its higher magnification than the wide-angle view, a disadvantage of the f+oc and ff+oc modes is that a major portion of the wide view is blocked by the embedded zoom-in view, which causes information discontinuity and may diminish the advantage of context awareness.

The f+wc and ff+wc modes are similar to the f+oc and ff+oc modes, respectively, in terms of window arrangement. In these modes, however, the wide-view image was severely warped such that the information hidden behind the zoom-in view window will become visible in a highly distorted region surrounding the zoom-in window and information continuity is maintained (Fig. 2E, F). To generate the distorted wide-view image, a distortion-based visualization technique [14] was adopted. These modes may address the view-blocking issue in the f+oc and ff+oc modes, but their clinical effectiveness remains unclear due to severe distortion.

Each of the six display modes has inherent advantages and disadvantages. There is no clear and definite understanding how the different display modes may affect surgical task performance. In this paper, we compare the six display modes through a human-subject experiment, which was conducted with an emulated MRFL using a 4K camera. The emulated MRFL was carefully configured with a 4K imaging sensor and camera lens such that its imaging capabilities such as magnification ratio of zoom view to wide view and spatial resolution were equivalent to the actual MRFL being developed in our lab.

Materials and methods

Participants

This study was approved by the Institutional Review Board at University of Arizona. Twenty-four participants (16 males and 8 females) were recruited from the population of staffs and students at University of Arizona. Their ages ranged from 19 to 34 years ($M = 23.5$ and $SD = 4.3$). Each participant voluntarily signed the informed consent form. All the participants were right-handed (except for one) and had no surgical experience. The majority of participants had backgrounds in either biomedical engineering (11 participants) or optical sciences (9 participants). The rest of the participants had backgrounds in either information management system, material science, electrical engineering, or mechanical engineering. Among the participants who had biomedical engineering backgrounds, three had a little experience of using laparoscopes in simulated environments only.

Experimental setup and user task

The MRFL system demonstrated in Fig. 1 is expected to offer two significant practical advantages to a standard laparoscope. First of all, it should permit a surgeon to maintain a centralized focus of enhanced visual acuity and magnification while maintaining overall visual awareness of the surgical field in his or her peripheral vision. This first benefit should reduce the potential for inadvertent collisions with objects and structures in the periphery of the surgical field. Second, the foveated view with high levels of magnification should permit the execution of surgical tasks without the need to excessively manipulate the laparoscope nearer the target while having to retract the scope to re-establish visualization of the peripheral surgical field. To fully utilize the capabilities offered by an MRFL, we need to design a novel experimental setup and user task that requires magnification while placing a premium on situational awareness in a study platform. To focus on evaluating how the different display modes may affect a user's task performance and behavior, we chose to use participants with no medical training and use simple generic tasks rather than surgical tasks. Additionally, we need to collect data not only on the time required for task completion but also on number of inadvertent collisions with objects in the peripheral field, gaze trajectory, and tool trajectory.

To meet these requirements, the experimental design was modified and simplified from the standardized pegboard transfer task (PPT) from the Fundamentals of Laparoscope Surgery technical skill training module. Figure 3 shows our experimental setup. Similar to the original pegboard design, we arranged 8 red posts, each of which had a diameter of 1 cm and a height of 10.5 cm, on a board in a test area of 19 cm × 16 cm. The distances between adjacent posts were ranged in 4.7–6.6 cm. The posts provide the occluders and objects of occlusion. By fixing the test camera at 26 cm from the test area, it is able to capture the overall scene of 8 posts without the need of manipulating the camera position, which simulates the wide-view of the MRFL scope and provides
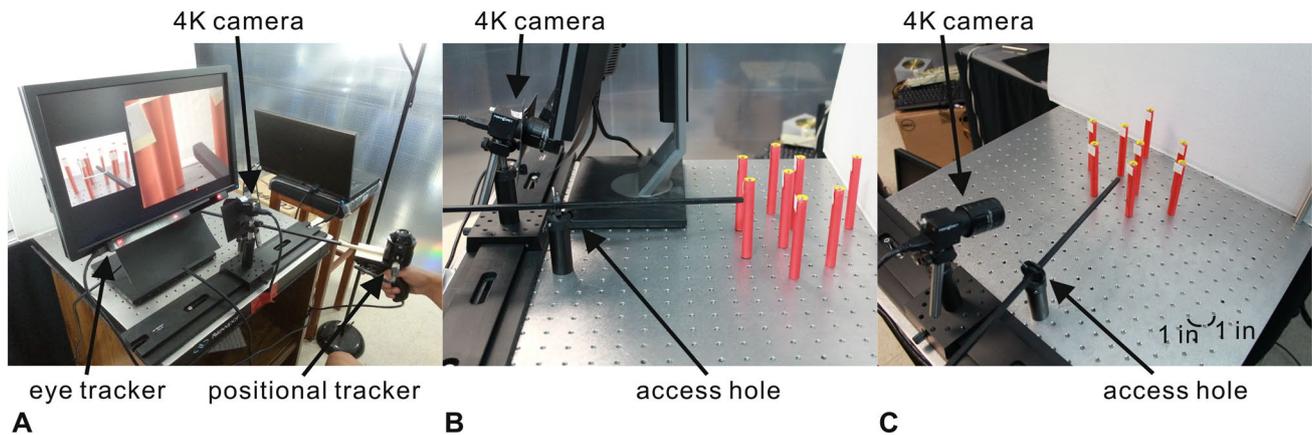
Fig. 3 Experimental setup: **A** a user-side view, **B** a right side view, **C** a view where the computer monitor was removed to clearly show the positional relationship among the camera, access hole (for wooden stick), and test area (posts). In the experiment, the computer monitor was used for blocking direct view of the working area, as well as providing views captured by the camera

the intact awareness of the test field. In the original PPT design, a subject is tasked to lift six triangular objects with a laparoscopic forceps instrument and to transfer them from six pegs on one side of the board to the corresponding pegs on the other side. To simplify the task, we attach a label to the top of each post and a unique number (between 1 and 8) was printed on each label. The numbers on labels can only be clearly seen in the zoom view only with adequate magnification. During the experiment, the participants were not allowed to directly look at the 8 physical posts. The positions of the camera and red posts were fixed, whereas the numbers on labels were changed across task trials.

A wooden stick painted black with a length of 60 cm and a thickness of 0.6 cm was used as an alternative to an actual surgical tool. A six degrees of freedom HiBall 3000 tracker (3rdTech Inc., Chapel Hill, NC) was attached to the wooden stick to track the tip position. We used the wooden stick rather than an actual surgical tool, since we considered only a touching task in this experiment, and it was convenient to attach the tracker to the wooden stick. The participants could reach the posts only through an access hole (diameter = 1.2 cm) located at the right side of the camera (8 cm away from the camera), while they were seated about 40 cm away from the monitor.

In the experiment, a 4K Flea3 FL3-U3-88S2C camera (pixel size of imaging sensor = 1.55 μm) with 6-mm lens (PointGrey Research Inc., Richmond, Canada) was installed rather than an actual multi-resolution foveated laparoscope (since the actual laparoscope was still being under development at the time of this experiment). From an image taken by this camera, two separate images for wide and zoom views were obtained. While a zoom-view image was clipped from an original image that had a resolution of 2700×2160, a wide-view image was generated by down-sampling the same original image. The pixel resolution of both views was 900×720. The spatial resolution of the zoom view for the test area was 7 lps/mm, which is consistent with the spatial resolution of the actual MRFL system. Though the pixel resolution of both views was the same, the wide-view image covers an area nine times as large as that by the zoom-view. As a result, the lateral magnification ratio of the zoom view to the wide view was 3:1 in both horizontal and vertical directions, which is consistent with the magnification ratio of the MRFL system. The wide- and zoom-view images were displayed on a 19-inch wide LCD monitor that had a pixel resolution of 1440×900, where the zoom view is always displayed in a window of the same pixel resolution as that of the generated zoom-view image and the wide-view is displayed at in a window of different pixels for different display modes. An EyeX eye tracker (Tobii Inc., Stockholm, Sweden) was attached to the monitor to track participants' gaze positions on the screen. Besides the reason of readiness, am emulated MRFL was chosen over the actual MRFL prototype for its optical rigidity without introducing irrelevant experiment variables such as the scanning speed, resolution or accuracy of the optical scanner used for steering the foveated imaging probe.

The user task was to touch all the red posts in numerical order (from 1 to 8) as fast as possible, but without colliding with posts other than a designated post. To maintain the same optimal travel distance of tooltip across task trials, the tooltip was initially located at a lower-right corner in the wide view, and it was required to return back to the initial position after touching each designated post.

Independent and dependent variables

Display mode was considered as a primary independent variable, and repetition was considered as a secondary independent variable. In the experiment, each participant

carried out three evaluation trials for each display mode. During each evaluation trial, a computer recorded tooltip position and participants' gaze position on the display, and the experimenter recorded task completion time and the number of collisions. From these data, the following measurements were produced as dependent variables:

- Task completion time
- The number of collisions
- Ratio between the times spent in gazing at the wide-view and zoom-view windows (= time in gazing at the wide view/time in gazing at the zoom view)
- Frequency of view changes between the wide- and zoom-view windows
- Length of gaze point trajectory on the display screen
- Moving speed of gaze point on the display screen
- Length of tooltip trajectory
- Moving speed of tooltip

In addition to the listed objective measurements, we also obtained subjective preference for the display modes using a post-test questionnaire. The questionnaire asked to select the least preferred and the most preferred display modes.

Experimental design and procedure

A 6 × 3 (display mode × repetition) within-subjects repeated measures design was used. Each participant started his/her session by reading and signing the consent form, followed by a demographic questionnaire. Afterward, the experimenter explained to the participant the six display modes and the user task. The eye tracker was then calibrated. In the calibration, the participant was asked to gaze at seven different positions (one by one) on the monitor screen, without moving his/her head. After the calibration of eye tracker, the participant carried out one training trial followed by three evaluation trials for the first display mode. S/he then again carried out one training trial followed by three evaluation trials for the second display mode. This task trial step was repeated until the task trials for the sixth display mode were finished (6 training and 18 evaluation trials in total). During the task trials, slight head movements were allowed. The participant was allowed to take a short break between trials. The presentation order of the six display modes was arranged with a digram-balanced Latin square. After carrying out all the task trials, the participant filled out the post-test questionnaire asking about his/her preference for the display mode. Upon completing the session, the participant was paid $20. A session lasted 1.5–2 h.

Statistical analysis

Statistical analysis was performed using SPSS version 21 (IBM Corp., Armonk, NY). On the objective measurements, two-way (display mode × repetition) repeated-measures ANOVAs were used for investigating differences among the display modes, and interactions between the display mode and the repetition. Mauchly's test was used for measuring sphericity on the data. Where significant violation of sphericity was found, the degrees of freedoms were corrected using Greenhouse–Geisser procedure. Following the ANOVAs, post hoc multiple comparison tests with Šidák correction were used for pair-wise comparisons among the display modes. For all the statistical tests, $\alpha = 0.05$ was used as a significance level.

Results

According to the statistical analysis, there were no significant interactions between the display mode and repetition for all the objective measurements [$F_{(4.592, 105.609)} = 1.07$, $p = 0.380$ for task completion time, $F_{(6.155, 141.572)} = 0.91$, $p = 0.493$ for the number of collisions, $F_{(3.518, 80.909)} = 0.447$, $p = 0.750$ for gaze ratio of wide view to zoom view, $F_{(5.935, 136.505)} = 0.73$, $p = 0.625$ for frequency of view changes, $F_{(4.471, 102.827)} = 0.97$, $p = 0.431$ for length of gaze point trajectory, $F_{(5.166, 118.810)} = 1.13$, $p = 0.351$ for moving speed of gaze point, $F_{(10, 230)} = 0.72$, $p = 0.709$ for length of tooltip trajectory, and $F_{(10, 230)} = 0.28$, $p = 0.985$ for moving speed of tooltip]. Since in this experiment, we were interested in main effects of display mode and interactions between the display mode and repetition, main effects of repetition are not presented and considered in this paper.

Task completion time

The participants carried out the task most quickly in the f + wc mode and most slowly in the ff + oc and o + d (pip) modes (Fig. 4A). In the ff + wc mode, they carried out the task more quickly than in the o + d and f + oc modes. However, the ANOVA showed that there was no significant difference among the display modes [$F_{(2.930, 67.394)} = 2.05$, $p = 0.117$].

The number of collisions

The f + wc mode showed the smallest number of collisions, followed by the ff + wc mode, then the o + d mode, then the o + d (pip) mode and the f + oc mode. The ff + oc mode showed the largest number of collisions (Fig. 4B). The ANOVA revealed that there were
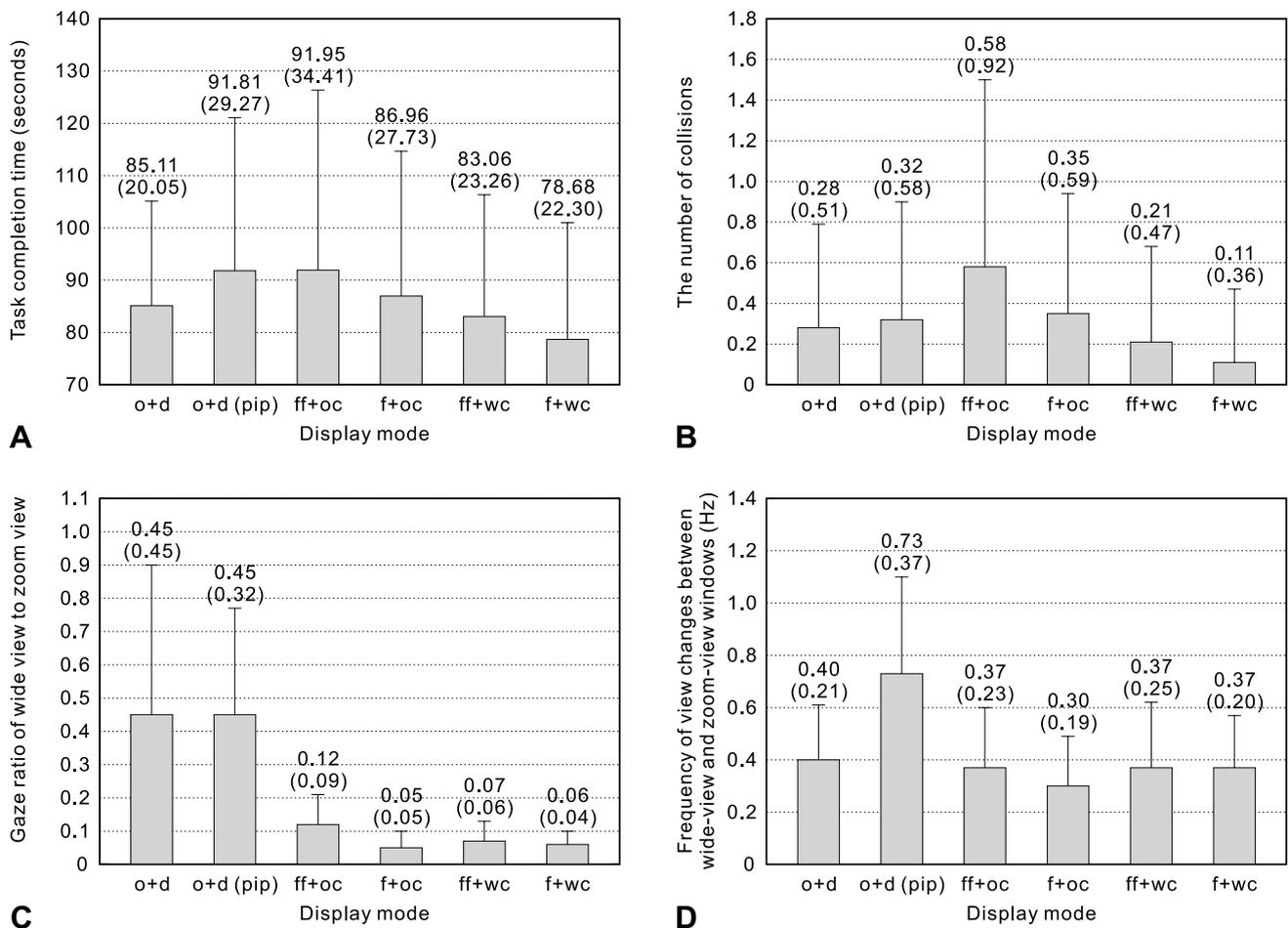
Fig. 4 Means and standard deviations of task completion time (**A**), the number of collisions (**B**), ratio between the times spent in gazing at the wide-view and zoom-view windows (**C**), and frequency of view changes between the wide-view and zoom-view windows (**D**)

significant differences among the display modes [F(2.451, 56.364) = 6.35, $p = 0.002$]. According to the multiple comparison tests, the f + wc mode showed significantly smaller number of collisions than the f + oc, ff + oc, and o + d (pip) modes only.

Ratio between the times spent in gazing at the wide-view and zoom-view windows

In all the display modes, the participants spent more time to gaze on the zoom view than the wide view (Fig. 4C). In other words, the zoom view was more used than the wide view in all the display modes. The ANOVA revealed that there were significant differences among the display modes in the ratio. According to the multiple comparison tests, the wide view was significantly more used in the o + d and o + d (pip) modes than in the other modes, and in the ff + oc mode than in the f + oc mode.

Frequency of view changes between the wide-view and zoom-view windows

The participants made the most frequent view changes in the o + d (pip) mode, then in the o + d mode, then in the ff + oc, ff + wc, and f + wc modes. They made the least frequent view changes in the f + oc mode (Fig. 4D). The ANOVA revealed that there were significant differences among the display modes [F(2.699, 62.079) = 17.44, $p < 0.0001$]. The multiple comparison tests showed that the view window was significantly more frequently changed in the o + d (pip) mode than in the other modes.

Length of gaze point trajectory on the display screen

The participants' gaze points on the display screen were most moved in the o + d mode, then in the o + d (pip) mode, then in the ff + oc mode, then in the f + oc mode, then in the

f + w mode, then in the ff + w mode (Fig. 5A). The ANOVA showed that there were significant differences among the display modes [F(2.893, 66.547) = 35.32, p < 0.0001]. According to the multiple comparison tests, the gaze points were significantly more moved in the o + d and o + d (pip) modes than in the other modes, and more moved in the o + d mode than in the o + d (pip) mode. In addition, they were significantly more moved in the ff + oc mode than in the ff + wc mode.

Moving speed of gaze point on the display screen

The participants' gaze points were fastest moved in the o + d mode, then in the o + d (pip) mode, then in the f + oc and ff + oc modes, then in the f + wc mode, then in the ff + wc mode (Fig. 5B). The ANOVA revealed that there were significant differences among the display modes [F(2.424, 55.754) = 144.96, p < 0.0001]. The multiple comparison tests showed that the gaze points were significantly faster moved in the o + d and o + d (pip) modes than in the other

modes, and were faster in the o + d mode than in the o + d (pip) mode. In addition, they were significantly slower moved in the ff + wc mode than in the other modes.

Length of tooltip trajectory

The tooltip was most moved in the ff + oc mode, then in the f + oc mode, then in the o + d (pip) mode, then in o + d mode, then in the ff + wc mode, then in the f + wc mode (Fig. 5C). The ANOVA showed that there were significant differences among the display modes [F(2.974, 68.413) = 4.10, p = 0.010]. According to the multiple comparison tests, the tooltip was significantly more moved in the ff + oc and o + d (pip) modes than in the f + wc mode.

Moving speed of tooltip

The tooltip was fastest moved in the ff + oc mode and most slowly in the o + d (pip) mode (Fig. 5D). However, the
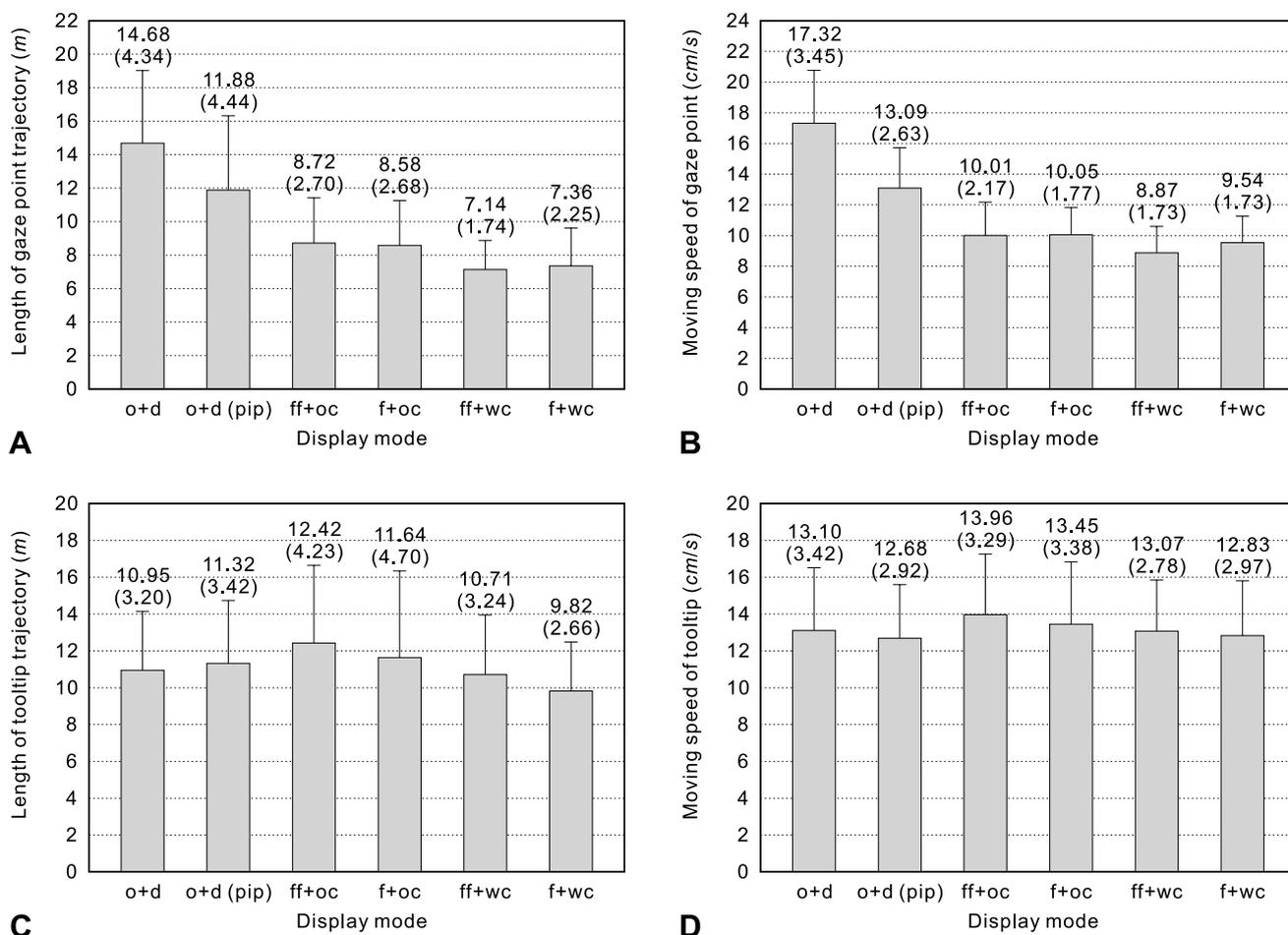

Fig. 5 Means and standard deviations of length of gaze point trajectory on the display screen (**A**), moving speed of gaze point on the display screen (**B**), length of tooltip trajectory (**C**), and moving speed of tooltip (**D**)

ANOVA revealed that there was no significant difference among the display mode [$F(5, 115) = 2.07$, $p = 0.075$].

Preference for display mode

The majority of subjects selected the o + d mode as the most preferred one, and the ff + oc and ff + wc modes as the least preferred one (Fig. 6).

Discussion

For laparoscopic surgical tasks, a conventional display mode using only a single monitor (or view) was compared with a dual-display mode using two separate monitors (or views) by Cao et al. [15] and Shah et al. [16]. Their experiments showed that the dual-display mode had a better performance than the single-display mode. In our experiment, all the display modes had dual views although they used a single monitor only. Our experimental results showed that the f + wc mode had a better performance than the other display modes in terms of both task completion time and the number of collisions. Considering the o + d mode as a display setup analogous to the dual-monitor setup in [15, 16], it is implied that the f + wc mode would also be superior to the dual-monitor setup.

In the f + wc and ff + wc modes, the wide-view image was distorted to show the area occluded by the zoom-view window. The participants reported that they did not prefer these two modes because they were not familiar with the distorted image provided in the modes. However, they showed the best performance in the distorted-image modes. This indicates that the overall context view played an important role to provide a contextual awareness, even if it was distorted. We think that the preference problem of the f + wc and ff + wc modes can be resolved through intensive training.
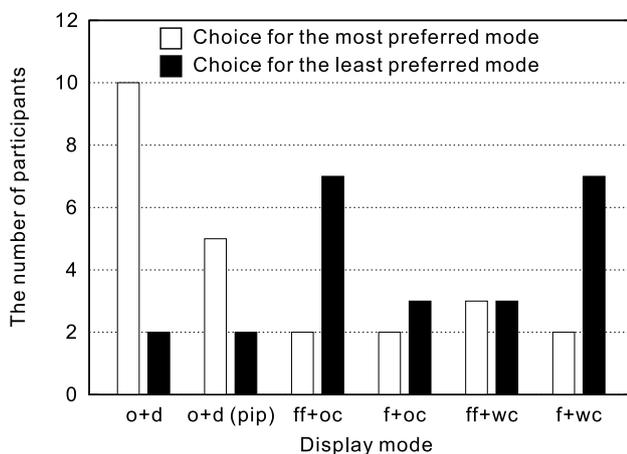

Fig. 6 Subjective preference for display mode

The participants showed a better performance in the f + wc mode than in the ff + wc mode. The difference between the two modes was whether the position of the zoom-view window followed the tooltip or not. We think that the position change of the zoom-view window depending on the tooltip position gave to the participants information about the tooltip position in the overall context. Due to this information, the participants could guess the tooltip position in the environment, and thus they performed the task better in the f + wc mode where the position of the zoom-view window was changed. Note that the tooltip position was extracted from the undistorted wide-view image and was not distorted. The effect of position change of zoom-view window can also be found in the f + oc and ff + oc modes. The difference between these two modes was the same as the difference between the f + wc and ff + wc modes. In the f + oc mode where the position of zoom-view window was changed, the participants showed a better performance than in the ff + oc mode where the position was fixed at the center of the wide-view window.

In the o + d mode, the participants showed a lower performance than in the f + wc and ff + wc modes (especially for the task completion time). This result can be explained with the distance between the wide-view and zoom-view windows. In the o + d mode, the distance between the two windows was farther than in the f + wc and ff + wc modes, and thus the participants' eyes and/or heads (or gaze point on the display screen) had to more largely move to switch between the wide-view and zoom-view windows. The increase in the time of eye movement led to increase the task completion time. Our explanation is supported by the results of the length of gaze point trajectory and frequency of view changes between the wide-view and zoom-view windows. The results showed that the gaze point moved significantly longer in the o + d mode than in the f + wc and ff + wc modes, while the frequency of view changes was similar in those modes. According to the result of moving speed of gaze point, the longer the gaze point moved, the faster the moving speed of the gaze point. However, the decrease in the time of eye movement due to the fast gaze speed did not compensate for the increase in the time of eye movement due to the long gaze point trajectory.

Although the distance between the wide-view and zoom-view windows was shorter in the o + d (pip) mode than in the o + d mode, the participants showed a lower performance in the o + d (pip) mode than in the o + d mode. The performance degradation in the o + d (pip) mode can be explained by the occlusion of the zoom-view window. In the task, it was necessarily required to observe the zoom-view window since the numbers on the labels were clearly shown in the zoom view only. Unlike in the o + d mode, in the o + d (pip) mode, the upper-right part of the zoom-view window was occluded by the wide-view window. Where the target

number or post was located in the occluded area, the participants had to change the zoomed area by carefully moving the tooltip to locate the target number in the visible area. This additional movement of tooltip led to increase the task completion time. According to the results of length of tooltip trajectory and its moving speed, the participants moved the tooltip more at a slower speed in the o + d (pip) mode than in the o + d mode. The slower moving speed of the tooltip in the o + d (pip) mode was due to the shape of visible area in the zoom-view window. Unlike in the other modes, in the o + d (pip) mode, the visible area was not a rectangle. To locate the target number or post in the non-rectangle area, it was required for the participants to move the tooltip more carefully.

In the ff + oc mode, the participants showed the worst performance. We think that this was due to the lack of the context information. In this mode, a large area of the wide-view window was occluded by the zoom-view window, and thus the context information was limited. Considering that a conventional laparoscope provides only a single zoom-view window without any context information, we expect that its performance would be even lower than the ff + oc mode.

In this experiment, we did not measure the degrees of eye and arm fatigue, which might be another performance metrics. From the results regarding gaze point and tooltip movements, however, the degrees of eye and arm fatigue depending on the display modes can be indirectly compared. The fatigue typically results from physical movement. This means that the more eyes and arm move, the more the eye and arm fatigue, respectively, is felt. According to the results, the participants' eyes moved the most at the fastest speed in the o + d mode and the least at the slowest speed in the ff + wc mode. This implies that their eyes felt fatigued the most in the o + d mode and the least in the ff + wc mode. In the f + wc mode, the participants' eyes moved more than in the ff + wc mode. This was because additional eye movement was required in the f + wc mode to follow the zoom-view window of which position was changed. Nevertheless, the amount and speed of eye movement in the f + wc mode was comparable with those in the ff + wc mode. Similarly, the degrees of arm fatigue for the display modes can also be compared using the results regarding the tooltip movements. The results imply that the participants' arms felt fatigued the most in the ff + oc mode and the least in the f + wc mode.

In the o + d and o + d (pip) modes, the participants spent significantly more time in gazing at the wide-view window than in the other modes. We think that this was because the wide-view image was undistorted and fully shown in those modes. Due to the familiarity with the undistorted and unoccluded imaging, the participants more actively used the wide-view window in those modes. Although the ratios between the times spent in gazing at the wide-view and zoom-view windows were similar for the o + d and

o + d (pip) modes, the participants made significantly more frequent view changes in the o + d (pip) mode than in the o + d mode. In the o + d (pip) mode, there was a possibility that the participants unintendedly gazed at the wide-view window since it was overlaid on the zoom-view window. For instance, when the participants intended to observe the upper-right part of the zoomed area, they might have automatically gazed at that part, which was occupied by the wide-view window. The moment they recognized that they were looking at the wide-view window rather than the zoom-view window, they would switch their view and try to change the position of tooltip. We think that this unintended gazing at the wide-view window might cause the significantly more frequent view changes in the o + d (pip) mode.

According to the statistical analysis, there were no significant interactions between the display mode and repetition for all the objective measurements. This indicates that the repetition did not significantly change the main effects of the display mode, and thus our experimental results were reliable.

Several empirical studies supporting our results can be found in the literature, although the tasks considered in those studies were different from ours and not directly related to laparoscopic operations. For example, Hornbæk and Frøkjær [17] compared the f + wc and o + d modes for document-reading tasks, and found that with the f + wc mode subjects spent less time in reading the details and thus read the documents more quickly than with the o + d mode. Another example is the study conducted by Gutwin and Skopik [18], in which they compared the ff + wc and o + d (pip) modes for path-steering tasks. They found that in the ff + wc mode subjects completed the tasks significantly faster than in the o + d (pip) mode, without losing accuracy. More recently, Lie et al. [19] and Butscher and Reiterer [20] conducted experiments to compare the f + wc and o + d (pip) modes, and the ff + wc and o + d modes, respectively. Lie et al. [19] showed that for a classification task, subjects carried out the task faster in the f + wc mode than in the o + d (pip) mode. Butscher and Reiterer [20] found that in a reconstruction task where the positions of objects were recalled after navigating a multi-scale space, the ff + wc mode showed a lower reconstruction error than the o + d mode. In both experiments, however, there were no statistically significant differences between the f + wc and o + d (pip) modes, and between the ff + wc and o + d modes.

In conclusion, we understood how the different modes to display the dual-view multi-resolution images affected the task performance in the touching task. It was found that the unblocked wide context view was useful even though it was severely distorted, and the information continuity played an important role to more effectively provide a situational awareness. Moreover, the position change of viewing window coupled to the location of tooltip or region of interest

helped improve the task performance, by providing an additional cue for spatial awareness. Since our study focused on the display interface for a multi-resolution foveated laparoscope rather than the efficiency in specific surgical tasks, a generalizable task was used in our experiment. Considering that the touching (or tooltip movement) task must be performed prior to any specific laparoscopic tasks, we expect that our findings are transferable to surgery-related tasks and with surgeons. Our future work will be to further validate the results with laparoscopic surgeons performing simulated surgical procedures.

Compliance with ethical standards

Disclosures Dr. Hong Hua has a patent Multi-Resolution Foveated Endoscope/Laparoscope (pending) and a patent Optical Article and Illumination System for Endoscope (pending), and has no other related conflicts of interest or financial ties to disclose. Drs. Sangyoon Lee, Mike Nguyen, and Allan J. Hamilton have no conflicts of interest or financial ties to disclose.

References

1. Cuschieri A (1991) Minimal access surgery and the future of interventional laparoscopy. Am J Surg 161:404–407
2. Heemskerk J, Zandbergen R, Maessen JG, Greve JWM, Bouvy ND (2006) Advantages of advanced laparoscopic systems. Surg Endosc 20:730–733
3. Parker WH (2010) Understanding errors during laparoscopic surgery. Obstet Gynecol Clin N Am 37(3):437–449
4. Pierre SA, Ferrandino MN, Simmons WN, Fernandez C, Zhong P, Albala DM, Preminger GM (2009) High definition laparoscopy: objective assessment of performance characteristics and comparison with standard laparoscopy. J Endourol Soc 23:523–528
5. Kim SP, Shah ND, Weight CJ, Thompson RH, Moriarty JP, Shippee ND, Costello BA, Boorjian SA, Leibovich BC (2011) Contemporary trends in nephrectomy for renal cell carcinoma in the United States: results from a population based cohort. J Urol 186(5):1779–1785
6. Bernie JE, Venkatesh R, Brown J, Gardner TA, Sundaram CP (2005) Comparison of laparoscopic pyeloplasty with and without robotic assistance. JSLS-J Soc Laparoendosc Surg 9:258–261
7. Tsai FS, Johnson D, Francis CS, Cho SH, Qiao W, Arianpour A, Mintz Y, Horgan S, Talamini M, Lo Y-H (2010) Fluidic lens laparoscopic zoom camera for minimally invasive surgery. J Biomed Opt 15(3):030504
8. Goel R, Lomanto D (2012) Controversies in single-port laparoscopic surgery. Surg Laparosc Endosc Percutan Tech 22(5):380–382
9. Canes D, Desai MM, Aron M, Haber G-P, Goel R, Stein J, Kaouk JH, Gill IS (2008) Transumbilical single-port surgery: evolution and current status. Eur Urol 54:1020–1029
10. Qin Y, Hua H, Nguyen M (2013) Development of a laparoscope with multi-resolution foveation capability for minimally invasive surgery. Proceedings of SPIE 8573, design and quality for biomedical technologies VI, 857309, February 2–3 2013, San Francisco, CA
11. Cockburn A, Karlson A, Bederson BB (2009) A review of overview + detail, zooming, and focus + context interfaces. ACM Comput Surv 41(1):2:1–2:31
12. Oh J-Y, Hua H (2008) Usability of multi-scale interfaces for 3D workbench displays. Presence 17(5):415–440
13. Jiang X, Zheng B, Atkins MS (2015) Video processing to locate the tooltip position in surgical eye-hand coordination tasks. Surg Innov 22(3):285–293
14. Liu F, Gleicher M (2005) Automatic image retargeting with fisheye-view warping. Proc 18th Annu ACM symp user interface softw technol (UIST 05), October 23–27 2005, Seattle, WA, pp. 153–162
15. Cao A, Ellis RD, Klein ED, Auner GW, Klein MD, Pandya AK (2008) Comparison of a supplemental wide field of view versus a single field of view with zoom on performance in minimally invasive surgery. Surg Endosc 22:1445–1451
16. Shah RD, Cao A, Golenberg L, Ellis RD, Auner GW, Pandya AK, Klein MD (2009) Performance of basic manipulation and intracorporeal suturing tasks in a robotic surgical system: single- versus dual-monitor views. Surg Endosc 23:727–733
17. Hornbæk K, Frøkjær E (2003) Reading patterns and usability in visualizations of electronic documents. ACM Trans Comput-Hum Interact 10(2):119–149
18. Gutwin C, Skopik A (2003) Fisheye views are good for large steering tasks. Proc ACM SIGCHI Conf human factors in computing systems (CHI'03), April 05–10 2003, Ft. Lauderdale, FL, pp. 201–208
19. Liu C, Chapuis O, Beaudouin-Lafon M, Lecolinet E, Mackay W (2014) Effects of display size and navigation type on a classification task. Proc ACM SIGCHI Conf human factors in computing systems (CHI'14), April 26-May 01 2014, Toronto, ON, pp. 4147–4156
20. Butscher S, Reiterer H (2016) Applying guidelines for the design of distortions on focus + context interfaces. Proc Int Working Conf advanced visual interfaces (AVI'16), June 07–10 2016, Bari, pp. 244–247