Augmented Reality in Surgical Training

Jonathan Reuvers
reuve026@morris.umn.edu
Division of Science and Mathematics University of Minnesota, Morris
Morris, Minnesota, USA

Abstract

Augmented reality is a developing technology that allows for the user to gain extra information about their environment. This aspect can be and is being put to use in the medical field to assist with the training of surgical procedures. Training can be improved through the ability to experience a more realistic feeling of created scenarios as opposed to a full virtual reality method. There is also the ability to instantly gain access to information in real time for feedback or for surgical needs. This paper will explore the usage of augmented reality in the training and performance of surgery.

Keywords: augmented reality, surgery

1 Introduction

Augmented reality (AR) allows for the user to interact with an altered world without being fully immersed into a virtual space. This contrasts it from virtual reality and allows for potentially greater fidelity in feedback of performing actions to interact with the augmented reality. These differences are what makes AR an appealing addition to develop and apply to medical facilities, specifically within the area of surgery. AR allows for external information that the surgeon would need to retrieve from elsewhere be available and even accessible without having to stop performing whatever action they are in the midst of. This can be a valuable time saving and life saving aspect that makes this technology worthwhile. It allows for students of surgery to experience a more realistic feedback due to the interactions taking place in a real space. The realism comes from the fact that it is actual reality they are interacting with.

AR has not been a very consumer accessible technology [2]. It is limited by a lot of factors with it being explored more fully in recent years through different means. One of the larger reasons for the greater accessibility stems from the development of smart phones. Smart phones have allowed for more access to augmented reality due to the powerful processing they have in addition to the cameras they house. In addition to that, they are more generally affordable and therefore accessible by the average consumer. The combination has made it possible to develop different kinds of consumer friendly augmented reality. The most accessible example of Augmented Reality from an average point of view would be the mobile game sensation of Pokemon GO. As AR becomes more accessible to the general consumer it is emerging as a feasible option for use in professional fields. Its usage within the medical field is continuing to be developed and explored for its ability to benefit in the training for surgery.

Minimally invasive surgery and neurosurgery are the most targeted area of use. Neurosurgery utilizes AR for its ability to combine the data from scans in real time to allow for more precise and safe "neuronavigation" [5]. Minimally invasive surgery is where this paper will focus the most on its applications. Minimally invasive surgery involves limiting access to a targeted area using few small incisions. From there a laparoscope, a tube-like instrument with a lens and light, can be inserted to provide guidance. The lack of space and viewing through a camera is where the use of AR can benefit the technique. By augmenting the view provided by a laparoscope, it can afford the surgeon a better ability to navigate the space.

2 Description of Augmented Reality

Augmented reality, as defined by Azuma [1] in 1997, contains three primary features:

1. Real world and virtual elements are combined.
2. Can be registered in three dimensions.
3. It can be interacted with in real time.

This definition allows for separation from virtual reality (VR) as well as not considering something like computer generated effects in movies from being augmented reality. The AR allows for additional information to be placed into the user's reality, hence the term augmented. It is acting to supplement what is real in front of them with digitally displayed images, information or other inserts.

By Ghazwani and Smith's definition a surgeon would most often be classified as a passive user [2]. This implies that the user is not directly manipulating the virtual content that is being displayed to them, but rather just receiving it. It's important to make this distinction because it allows for the specification of what these medical AR devices would be accomplishing. They are not attempting to perform the surgery, but rather assist in the operation by providing information about what is being looked at or about the patient in general.

2.1 Differences from Virtual Reality

Virtual reality does not function in the same manner as augmented reality. This is due, as stated previously, to the difference in the objects and environments that are being interacted with. VR provides a great ability to allow for simulations of artificial scenarios that is immersive for a learning student [5]. VR does allow for an increase in ability when used for training, but somewhat lacks in its ability to provide accurate tactile feedback to the user. It is also something that is not exactly feasible to provide benefit to a surgeon during an operation. The need to construct the space virtually and be placed fully within that space does not lend itself for use like that. AR on the other hand can provide information in addition to what the user is already seeing in the real world allowing for it to be used in both training and actual performances. These differences create a use for both of the technologies to have a purpose depending on what is needed, but there is a notable distinction between AR and VR.

3 Use in Surgical Training

In the context of surgical training there are a few different avenues that AR can be used for. The ability to teach individuals how and what to perform in certain situations is highly beneficial. There are traditional options, such as cadavers that allow people to practice medical procedures, but it is hard to be precise or have access to teachers who are at the top of their field for a specific procedure. By introducing AR, students can gain access to the ability to practice space navigation increasing their efficiency and accuracy. There is also the opportunity to gain guidance from experts at distant location through direct interactions in real time.

3.1 Visual Interactive Presence and Augmented Reality

Specifically an interesting developmental project called, visual interactive presence and augmented reality (VIPAR) [4], is something that allows for surgeons to be able to digitally overlay their hands to assist with students in other locations. This allows for a much larger educational influence of top experts around the world. It is accomplished by taking visual inputs from two locations and mapping the images over one another. Then both of the users are required to view the area through a viewfinder. When looking through as shown in figure 1, both parties will see all of the included hands as shown in the capturing area. The display of the hands allows for a remote mentorship of any desired situation as the mentor can directly show and point out objects in real time due to the AR qualities of the device.

The software is required to perform standardized calibrations to match the two locations within the same three dimensional space. Specifically it requires removing the effects of video compression and transmission in an attempt to restore the images that are being transmitted back to the original state. From there the software has a continuous buffer of the most recent state so that it can consistently use the most recently available image or frame. These images are then merged by overlaying a semitransparent version of the remote location's image onto the local one. The result

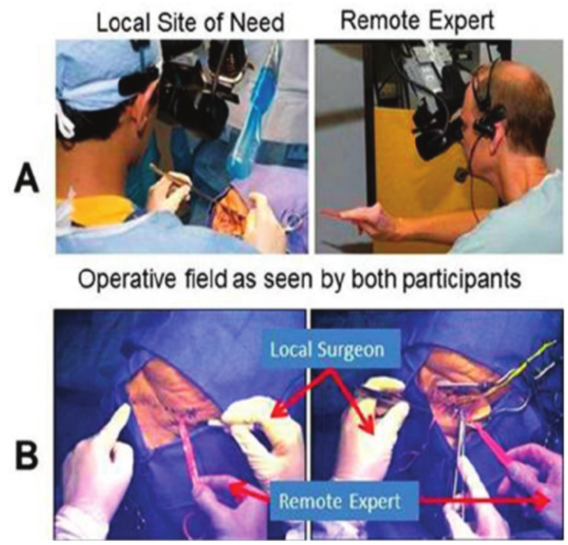

Figure 1. The local student (left) and remote mentor (right) (A) can see each others hands actions merged into one display for a collaborative experience (B) [4]

is contained in another buffer that allows for a consistent visual output for the user.

Some limitations with this specific technology are due to the frame rate and latency that can occur. Both are reliant on solid hardware and internet connections from each end of the connected users. Each side may see the other's movements as being delayed or choppy if either hardware or internet cannot handle the information being processed. The delay in movements can be eliminated by having slow movements to some extent, but the latency in actions should be accounted for. Estimates of latency have been performed from Birmingham, Alabama to Indianapolis, Indiana and found to be within a single frame [4]. The other thing that was tested and noted is the the visual accuracy in locations between the remote and local users. This was found to be sufficiently accurate with inaccuracies usually being due to different cameras at the two sites [4]. The cameras may use different compression algorithms (codecs). The change in compression can cause the images to not properly align or appear distorted in comparison to each other.

3.2 Training for Navigating Three Dimensional Space

Using AR can be somewhat difficult currently for the user if they are not accustomed to the lack of depth perception. Stemming from the fact that most interfaces are currently a console or viewing apparatus displaying a monocular view, it restricts the ability to determine the depth of motion in some situations. The inability to determine different depths of objects can be hazardous when working on the internals of a patient. To remedy the unfamiliarity, there is training with AR that allows for the users to get more experience working within the three dimensional space while it is being viewed from the two dimensional screen. The visual guidance provided by creating reference guides could also be considered for use during a real operation as well [7].

Humans utilize different aspects of our vision to be able to determine depth perception. A lot of the time human's depth perception is equated to binocular vision, cues from both eyes. This method of overlaying information over the display feed leans into the monocular cues that humans can benefit from. Binocular vision is not easily achieved during minimally invasive surgery due to the use of laparoscopes [7]. They can only provide a singular camera feed at a given time, this requires a method of guiding that can utilize humans monocular vision cues. These would include the ability to compare relative sizes, occlusion, and similar cues that can be achieved by one eye. By taking advantage of these, the user may be able to navigate the three dimensional space through one camera.

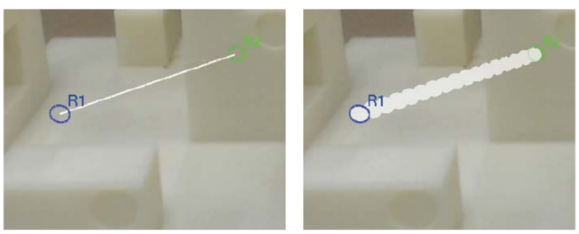

Figure 2. Paths with no references [7]

In figure 2, the depiction of two points are displayed at separate location in both the x and z planes. It is fairly difficult to determine from the two dimensional imaging provided by a picture. Wagner and Rosenblit decided upon using these as the starting point for developing a system that allows for the current location of a tool to be identified. These two examples have no reference points added to them and are just the path that the user is intended to navigate along. By comparison, figure 3 depicts the four methods of adding references to guide the user. Each of the reference displays are constructed in a different way.

3.2.1 2D Squares Overlay. The two dimensional square overlay, the top left of figure 3, depicts the usage of fixed size two dimensional squares that show the direct position of the instrument and the next intended position. It does not change size based upon depth of the points nor does it change its overlap between the two squares meaning that the red square indicating the user's instrument is always on top. This may lead to confusion by the user due to expectation of one position always being in front of the other [7].

3.2.2 3D Squares Overlay. The three dimensional square overlay, the top right of figure 3, depicts the usage of squares that are constrained to the world XZ plane [7]. In doing this, the squares have a fixed side length of 1.5 cm, but are rendered into the image to appear larger or smaller on screen based upon being closer or further to the screen respectively. This is done by computing the fixed side length into the world frame before they are placed into the image plane. This allows for the feeling for depth for the user.

3.2.3 Cubes Overlay. The cube overlay, the bottom right of figure 3, depicts the usage of cubes that are rendered into three dimensions. This uses a similar method as the three dimensional squares, but now includes translucent faces and a sense of depth to the cube itself. The cubes now gains the benefit of providing the user information through occlusion by allowing for the cubes to overlap and intersect as shown in the example.

3.2.4 Post Overlay. The post overlay, the bottom left of figure 3, depicts the usage of posts that are drawn vertically as lines along the designated world Z axis from a specified "ground" plane. This essentially means that a point in the air that is associated to the instrument or future intended path is being marked by a line that drops from its three dimensional XYZ position. The line drops down until it touches the ground where it is casting a "shadow".

3.2.5 Results. These four methods, along with the no reference methods from figure 2, were used to assist in subjects ability to navigate a path in three dimensional space. The goal of the study was to have each subject navigate the instrument through a course route that was built from an algorithm. The results show a significant decrease in total path length of the instrument by using the AR reference guides as well as overall less deviation from the intended path [7]. With a total decrease in path length, it is showing that the user is being more accurate on the designated point to point path they were instructed to follow. The two dimensional square method performed the worst of the reference models by an average of 5 cm. The performance is still noticeably better than the paths with no reference imaging which on average had a total path length of approximately 170 cm. In comparison, the two dimensional squares had an average total path length of approximately 140 cm. To compare the maximum distance traveled, the no reference methods reached a total distance of 320 cm while the other methods achieved no more than 180 cm total distance.

3.3 Robotic Surgery Training

In a similar fashion to the previously described VIPAR project, there is development of remote devices that allows for a controller to operate through robotic limbs. This is done to provide a higher level of accuracy and precision compared to traditional surgeries [3]. The operator accesses a console that allows for both viewing and controlling the robotic limbs. The visual display is similar in that it is taking a video feed

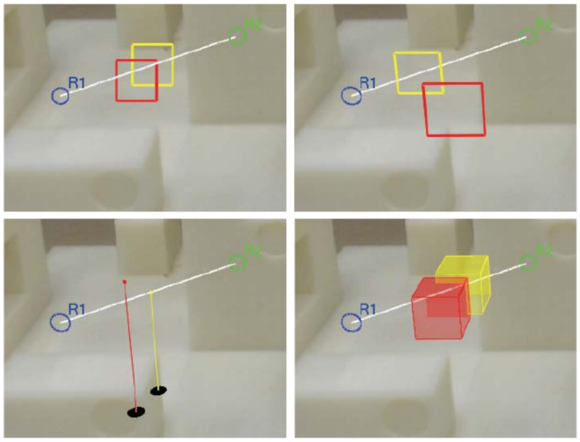

Figure 3. Spatial reference guide. The red indicates the instruments tip and the yellow indicates the reference position. [7]

and displaying it with the necessary imaging virtually being placed into it. The base model of machine that is being utilized to perform the operation is the da Vinci Surgical System developed by Intuitive Surgical Inc. [3]. It is a three piece cart. One part of the device allows for the operator to view what they are working with through a stereoscopic view. The operating table with attached limbs and an external console that displays what the surgeon is viewing make up the other two parts. When a student is being taught to use these however, the trainer only has access to the two dimensional output from a standalone console, this is a similar issue as brought up when talking about navigating the 3D space. The intention of allowing a trainer the ability to interact in the same fashion as the main operator to allow for a better method of communication and instruction.

This is achieved by first developing a virtual environment for the instructor to interact with similar to what is in the original da Vinci Surgical System. This was accomplished by utilizing a similar control scheme and then directly interfacing with the system's stereoscopic vision. The AR functionality is then realized by taking the feeds and mapping them to specific viewports as shown in figure 4. This allows for the virtually constructed environment to be mapped to the live feed from the cameras on the da Vinci Surgical System [3]. In doing so, the instructor and student are able to interact in three dimensional space at the same time rather than the instructor being restricted to drawing on a two dimensional plane to try and guide a student. The feed displays in 60 frames per second, but struggled with latency, approximately 164 ms of delay on average [3]. The testers found the system to be somewhat cumbersome to use and had some trouble distinguishing the virtual robotic limbs for the real ones.

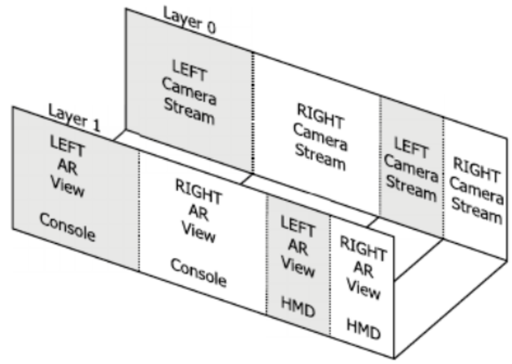

Figure 4. A depiction of how the views are being overlaid [3]

3.4 Outlining

Projection of desired information such as labeling or outlining is another possible application of AR that can benefit the surgical training. Outlining is done by software helping to emphasize locations internally and help to outline them for the user. The software requires the user to manually position the overlay to help ensure that is is correctly identifying what it is seeing. In figure 5, the landmark being used is the ribs [6]. The program then is able to place a dotted line around the outline of the indicated internals to help better navigate the space. This can be challenging as shifts in the body, such as respiratory movements, can cause key guiding points to become misaligned. In a study on utilizing outlining in regards to digestive surgery, an outlined tumour's movement could be reduced to within 1mm [6]. This display of information is depicted by a three dimensional visualization being superimposed onto a video view of a patient.

In order to register the location for the AR system to display an outline, the user must, as stated, manually align it to a landmark. In this case, it was the ribs. In other situations it might not be optimal to have to manually reposition to get proper readouts and could lead to error, so the development of automatic tracking was looked into. This was accomplished by using 25 markers that can be seen in both CT scans and video images are viewed from two cameras. Multiple algorithms are then applied and allow for the markers to be matched and registered. This allows for the system to be able to "see" the patient. In addition to this, the instruments need to be tracked to appear on the AR system as well.

With this system in place, participants were made to complete needle targeting modeled tumours. This was done by using a synthetic liver inside of a abdominal model. Ten consecutive attempts were made by the participants and an acceptable average error of 1.8mm was found [6]. On average it took below 30 seconds to reach the target [6].

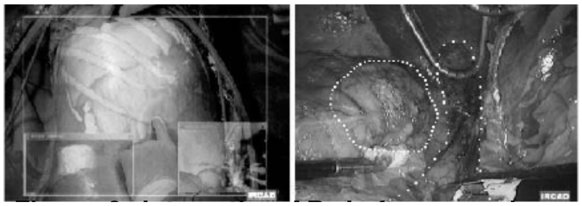

Figure 5. An example of manual registration (left) and structures emphasized with an outline (right) [6]

4 Limitations

Although the developments of AR can provide benefits to the surgical training field, there are currently limitations on what it can accomplish in a live setting. These stem from the usability or fluidity of the product. In addition, the pieces of equipment can be quite expensive. The technologies that have been covered are not free from of these flaws and as such should be addressed.

4.1 Calibration

Many of the systems in AR require some form of anchoring to the space that is being observed. This is commonly done through the usage of camera feeds as described in section 3.1. In order for these to be useful, there has to be a maintained location or some ability to adapt to the movement of the target. If there is no consistency involved, or a method as described in section 3.4, to continue to track the movement of the subject, then the usage of AR becomes detrimental as the readings and information it provides loses accuracy.

4.2 Power Consumption

Another important limiting factor to note about AR devices is that of power consumption. These are not mobile devices and are generally limited in location and transportation. This can limit the usability as well as these are generally not battery operated device leading to cumbersome cable management and other restrictions that could impede the ability of the surgeon to efficiently perform an operation [5].

4.3 Privacy

One of the last limitations or rather problems that can arise from AR assisted surgery and training is the increased risk to patient information security and privacy. In training it is less impactful, but if these surgical methods move to live operations there is the risk of losing data privacy. Due to the nature of AR being mostly digital and video feeds there opens more areas for potential breaches of patient privacy [5]. There is a need for proper encryption and data security that upholds the requirements of the medical field.

5 Conclusion

The pursuit of AR allows for the education of new students in the surgical field to both gain experience and confidence in the tasks that they will need to perform. More developments should be pursued to allow for a greater ability for augmented reality technology to be accessed by more individuals. By in improving on this area of training, surgeons, both experienced and inexperienced, will gain more precision and efficiency within the procedures they must perform overall resulting in a safer and less invasive surgeries. This in turn reduces the risk to patients as well as the potential recovery period needed.

The visual interactive presence and augmented reality project allows for students to be mentored through situations miles away from the instructor. By allowing for explicit demonstration as a form of distant education it reduces the need for travel from the experts of the field. It greatly simplifies access to those who would be best provide guidance for a specific procedure. The technology is mostly reliant on having the equipment available at two sites which in turn does not require a great deal of prior setup if the technology becomes a widespread form of instruction.

Navigating three dimensional space can be eased through training, but is made even more efficient through the specific ability to layer images on top of the field of view to help assist the navigation. Introducing the referencing images provides the user with greater accuracy to their movements which allows for more minimally invasive operations. In developing this technology further, there is the potential capability to expand and improve upon the displayed location. In doing so, it may eventually enable an almost realistic one to one reference in live time. In order to achieve that level of accuracy, some aspects of the technology, both hardware and software, would need to be improved, but it is a potential future feat.

Acknowledgments

Thank you to Elena Machkasova, Hussam Ghunaim, and Scott Steffes for their feedback and advice.

References

[1] Ronald T. Azuma. 1997. A Survey of Augmented Reality. *Presence: Teleoperators and Virtual Environments* 6, 4 (1997), 355–385. https://doi.org/10.1162/pres.1997.6.4.355

[2] Yahya Ghazwani and Shamus Smith. 2020. Interaction in Augmented Reality: Challenges to Enhance User Experience. In *Proceedings of the 2020 4th International Conference on Virtual and Augmented Reality Simulations* (Sydney, NSW, Australia) *(ICVARS 2020)*. Association for Computing Machinery, New York, NY, USA, 39–44. https://doi.org/10.1145/3385378.3385384

[3] Florin Octavian Matu, Mikkel Thøgersen, Bo Galsgaard, Martin Møller Jensen, and Martin Kraus. 2014. Stereoscopic Augmented Reality System for Supervised Training on Minimal Invasive Surgery Robots. In *Proceedings of the 2014 Virtual Reality International Conference* (Laval, France) *(VRIC '14)*. Association for Computing Machinery, New York, NY, USA, Article 33, 4 pages. https://doi.org/10.1145/2617841.2620722

[4] Mahesh Shenai, R. Shane Tubbs, Barton Guthrie, and Aaron Cohen-Gadol. 2014. Virtual interactive presence for real-time, long-distance

surgical collaboration during complex microsurgical procedures Technical note. *Journal of neurosurgery* 121 (06 2014), 1–8. https://doi.org/10.3171/2014.4.JNS131805

[5] Khor Wee Sim, Benjamin Baker, Kavit Amin, Adrian Chan, Ketan Patel, and Jason Wong. 2016. Augmented and virtual reality in surgery—the digital surgical environment: applications, limitations and legal pitfalls. *Annals of Translational Medicine* 4, 23 (2016). https://atm.amegroups.com/article/view/12851

[6] L. Soler, S. Nicolau, J. Schmid, C. Koehl, J. Marescaux, X. Pennec, and N. Ayache. 2004. Virtual Reality and Augmented Reality in Digestive Surgery. In *Proceedings of the 3rd IEEE/ACM International Symposium on Mixed and Augmented Reality (ISMAR '04)*. IEEE Computer Society, USA, 278–279. https://doi.org/10.1109/ISMAR.2004.64

[7] Adam Wagner and Jerzy W. Rozenblit. 2017. Augmented Reality Visual Guidance for Spatial Perception in the Computer Assisted Surgical Trainer. In *Proceedings of the Symposium on Modeling and Simulation in Medicine* (Virginia Beach, Virginia) *(MSM '17)*. Society for Computer Simulation International, San Diego, CA, USA, Article 5, 12 pages.