Imaging the Body: Embodied Vision in Minimally Invasive Surgery

Helena M. Mentis & Alex S. Taylor
Socio-Digital Systems
Microsoft Research Cambridge, United Kingdom
[hementis, ast]@microsoft.com

ABSTRACT
Recent years have seen the possibilities of new imaging and interaction technologies for minimally invasive surgery such as touchless interaction and high definition renderings of three-dimensional anatomy. With this paper we take a step back to review the historical introduction and assimilation of imaging technologies in the surgical theatre in parallel with the productive and cross-referential nature of surgical practice and image use. We present findings from a field study of image use during neurosurgery where we see that the work to see medical images is highly constructed and embodied with the action of manipulating the body. This perspective lends itself to a discussion of the directions for new imaging interaction technologies.

Author Keywords
Embodiment; Vision; Movement; Health; Surgery; Imaging.

ACM Classification Keywords
H5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
Although images have a long history of appropriation in medicine, surgery is an area that has been delayed in its use of images as an integral part of its procedures. As the practice of surgery has historically been a process of exposing the internal body in order to fix or remove anatomy, the need for images has been confined to the pre- or post-operative stages. However, advanced imaging technologies, enabling the capture of images throughout a procedure (such as fluoroscopic X-rays), as well as the introduction of new mechanisms for visualizing and interacting with preoperative images (such as CT scans) are providing the basis for intraoperative image use.

This may seem like a relatively straightforward technical challenge; one that is reliant on the generation of more sophisticated imaging techniques and implementation of automated processes for image registration in order to represent increasingly accurate anatomical data to surgeons. However, these techniques and systems invite questions about what is actually being represented, and to what extent the representations have the capacity for unmediated and complete views of the body. Similarly, questions are raised about the situations in which these techniques are deployed. The data are obviously subject to interpretation, contingent on numerous other situational factors, as well as a negotiation between other parties and sources of data. With these questions in mind, the representations of data and the mechanisms designed to support interacting with them take on a level of investigative importance.

These are not entirely new points to make in the social studies of science. There is a large and growing body of sociological studies examining "scientific images" showing that they are products of active interventions and not merely passive representations of some stable "reality out there" (see for example, [18], [19], [28], [17]). The arguments provided here are intended as a contribution to this corpus of sociological work. The contribution, specifically, is to do with the introduction of new technologies for 'seeing' the body, and the 'professional vision' [6] these technologies necessitate. More to the point, this paper aims to situate the uptake and use of new intraoperative imaging technologies alongside the historical and social use of medical images in surgery, and, in doing so, convey something of the embodied skills and ways of knowing that are both demanded of and enacted by surgeons.

This is an important contribution for HCI as we begin to introduce the next generation of intraoperative imaging technologies in minimally invasive surgery. As the point of surgery is visually separated from the surgeon, the reliance on the images is increased. Moreover, as the sophistication of these images increases – for instance, interactive three-dimensional renderings – the questions are raised about the appropriate location, time, and mechanism of interaction. Drawing on insights from the literature, we present findings from a field study of image use during neurosurgery. Our analysis focuses on the active use of imaging systems. We use this to discuss further implications for imaging interaction technologies that relate in particular to embodied use.

MOVING IMAGES TO THE CENTRE STAGE

On the face of it, medical imaging offers an unadulterated means of seeing inside the patient's body, exposing the anatomy beyond what is visible on the surface through non-invasive methods. With the sophisticated, highly technical achievements that allow the internal structures and workings of bodies to be visualized, it is hardly surprising that they are now considered a common part of, if not the *de facto* standard, in modern medical diagnosis and practice. However, it may be surprising to know that images were not always held in such high regard.

Historical accounts of the assimilation of medical images, such as X-rays, emphasize the gradual and negotiated process by which meaning and acceptance of medical images arose in the medical community at the turn of the 20th century [22] [28] [5]. This rise was accompanied (or perhaps assisted) by a growing rhetoric where the status of seeing was raised above the other senses in diagnosis and assessment in addition to the push for a more scientific profession based on a visual verification of anatomical results as opposed to a physician's subjective testimony. This 'visual as gold standard' is in contrast to modern Western science's history of placing visual knowledge lower in the hierarchy below that of text and numbers [24] [28]. However, as we will elaborate later, modern medical images are anything but simply visuals.

In order to use this new form of knowledge, surgeons had to learn to 'see' the body and its ailments from a visual, radiological perspective. Moreover, such forms of 'seeing' enabled them to find further useful applications for the rays themselves [22]. By the 1920s, X-ray images had established themselves in regular medical practice and came to be a fixture in surgical procedures – particularly in planning and assessment. X-rays provided the surgeon with a clear and recordable/verifiable understanding of the body's condition before an operation as well as a clear view of the resulting condition post-intervention, all without any invasive method. This in turn led to the rapid development of new diagnosis and viewing criteria [28]. For instance, the use of X-rays by orthopaedic surgeons in Germany were used to produce "new definitions of what constituted a satisfactory treatment" through the introduction of sequences of images as far more persuasive visual evidence than a patient in the flesh standing before them ([28], p.3). Thus, the images came to be seen as true and accurate representations of the body—literally standing in for the patient—and, in so doing, established a basis for knowledge and an organising practice [22].

In modern day surgical interventions, medical imaging has come to play an increasingly important role, particularly with it enabling minimally invasive procedures. These are procedures that enter the body through small incisions or existing body cavities in order to minimize anatomical damage, reduce recovery time, and minimize the risk for infection.

Interpretation & Negotiation of What is 'Seen'

Despite the trajectory of these technological innovations, the meaning of medical images is still a matter of negotiation and interpretation [23] [14] [1]. There is a substantial body of literature that addresses vision as situated and interactional (see [6] [7] [8] [9] [10] [21] [25] [26]). It is beyond both the scope and purpose of this paper to offer complete coverage of this work. Our intent here is to highlight some of the key theoretical works that particularly influence, motivate, and inform the contributions in this paper.

Studies of medical image use discuss the negotiations around and interpretation of images in clinical and surgical work and the ways these unfolding interventions manifest themselves as integral to medical decision-making [15] [11]. The practice of surgery for instance is reliant on the ability of the surgeons to resolve the body before them with the abstract representations of anatomy from books and diagrams. Hirschauer [11] describes in his sociological account of the practices of surgery that dissection is a process through which the internal body is made visible and identifiable in relation to these abstract representations. This is called 'exposition' or 'making anatomy' [11]. This method of viewing the body is a part of the training and visual learning of human anatomy that a surgeon must undergo – to see the body as parts that are separable from the whole body [23]. Evidence of this in surgery is also seen in how the body part that is of interest is delineated during surgery by removing all other parts from view through sterile coverings and drapes [11].

But, the body itself can also present facts in which the images must be resolved. Thus, the process of uncovering, identifying, and resolving between the material body and the representation of the body is reflexive and iterative. As Hirschauer [11] later describes it:

> *"In the process of operating, both the impressions gained of the concrete body and the single images, which make up the abstract body, are only of transitory relevance. Perception in surgery is a continuous superimposition of one over the other: a permanent cross-fading of experience and representation" (p.310).*

Thus, surgery is a complex process of resolving the physical body as it lies before the surgeon with what one knows – both learned from experience and expected from preliminary work – and this does not occur only once or twice, but continually throughout a medical intervention. Although by 'representations' and 'images' Hirschauer is referring to those that lie in books that are learned and memorized, the relationship between the physical body and the intraoperative digital representation of that body has much in common, as we shall soon show. This is the complex situation facing surgical practice with the exposed, physical body. However, in the case of minimally invasive surgery, the contributing factors to this complexity appear to be exacerbated.

To reiterate, in minimally invasive surgeries the body is not available for exposition. Thus the focus of manipulating and annotating is on the supporting images, and the consequence is the production of what might be thought of as a *docile object* [18]. Lynch refers to scientific images as docile objects when they have become observable, measurable, and quantifiable. In other words, the body is visualized and configured in such a way that it conforms to the needs of the surgeon. As we argue in this paper, simply producing and presenting an image does not directly translate to an effective medical gaze. Disciplining an image is first necessary before they can be of use in guiding procedures [23].

Apparent, then, is how new digital visualization technologies in minimally invasive surgery may be having an impact on the disciplining of the image, and on how the docile object is seen and constituted. Technologies such as computer tomography (CT), magnetic resonance (MR) and digital X-ray technology seem, on the face of it, to fundamentally alter what it means to see the body [3]. A new visual regime is emerging that changes both the way the image is captured and represented as well as how one interacts with these new digital images in relation to the patient body.

THE SLIPPERY IMAGE PATH

Traditionally, minimally invasive procedures are supported by intra-procedure imaging (X-ray fluoro, ultrasound), which provide real-time feedback, but the images have limited information due to their often lower-fidelity representations. In contrast, new image-guidance systems provide location information through a location aware probe superimposed on pre-operative MR or CT images, which provide the necessary clarity and detail.

Whereas X-ray and ultrasound images both entailed a camera-like technique to produce images of the body, new technologies deal with cross-sections of the body [4]. For instance, MR images are computer-generated visual representations of signals such as relaxation times, proton density, or diffusion of blood or other fluids. Even though the resulting image is displayed in two-dimensions, the result is a three-dimensional representation, as each image is made up of contiguous slices through a body part and can be viewed from more than one plane.

Thus, the ‘images’ are in actuality visualizations of data – models, which are, in a manner of speaking, “once or even twice removed from reality” [16] (p. 829). This new visual regime is further instituted with MR’s ability to image soft tissue, allowing the medical gaze to bypass the need for surgical or autopsical interventions. The fundamental shift from reflection or absorption of light or other electromagnetic waves to the calculation of various parameters per voxel is being pushed beyond that which is achievable by the human eye to produce further reconfigurations of the body (for example, producing clearer images of anatomical parts by suppressing the signals from body fat [27]).

As the newer imaging technologies have moved towards data representation as opposed to optical/aural reflection due to their numerical basis, these ‘images’ have secured a place as objective and superior knowledge [24]. However, it is important to note that these images are not equivalent to the physical body or provide a view of the body that can be known. They do not offer unmediated access to a body that exists outside of language and human actions and surgeons do not look upon them as providing certainty and definitiveness. Although they engage in popular rhetorical practices regarding their vision through the images, physicians and other professionals close to the production and use of images understand that imaging technologies do not render the body transparent [14] and that context plays an important role in the way an image is discussed and viewed [2]. They understand that medical imaging technologies do not *reveal* the inner body as much as *produce* the body; and that in order to produce that body they need to discipline it.

As Prasad [23] showed in the seeing practices of the radiologist, cross-referencing of different images and diagnostic data constitute one’s ability to detect and fix pathology. However, even in the radiology suite, cross-referencing through differential analysis is not contained within, but extends beyond – through, for instance, epidemiological data on the incidence of disease with respect to age, sex and other demographics, or through the diagrams and notations that are used to interpret MR images [23]. Alač [1], in a study of neuroscientists’ use of fMRI brain images, observed the gesturing and manipulation of digital displays during discussion in order to collaboratively understand and make meaningful what they were seeing. Their gestural engagements are phenomenal action to meld the digital, abstract images and the physical, concrete body through bodily orientation, gestures, and discussion. This can be seen in the way surgeons step through slices of a head or abdominal MR images as demonstrated by Johnson et al.’s [13] discussion of interventional radiologists viewing the temporal sequencing of images when stepping through fluoroscopic X-ray images.

These perspectives on medical images are what underpin and motivate the work that follows. As minimally invasive surgery becomes more prevalent, the need for sophisticated intraoperative images increases; however our understanding of the practices of disciplining and interacting with those images is lacking. How, exactly, do surgeons ‘see’ images in the context of surgical practice and what does this imply for the interactive mechanisms and techniques applied to these images, in particular in relation to the interaction with the patient body? In the following work, we explore these questions in the realm of neurosurgery. Although neurosurgery is only one sub-specialty of surgical practice, the need for imaging to guide the surgeon's tools around delicate anatomical parts is representative of the relationship between the physical and digital towards a common goal. From these insights we discuss the interactive needs around imaging and possible directions for new and integrative interaction methods.

NEUROSURGERY FIELDWORK

The examples presented are from fieldwork we have previously reported on in the neurosurgery department of a large hospital in the UK [20]. In our previous work we focused on the spatial organization of work in surgery, whereas this work is focusing on the construction of images through body and instrument manipulation. The vignettes used in the two papers are different and thus we are presenting new findings as well as using a different lens to analyze those findings.

During the fieldwork we employed observations and open-ended interviews – both of which we video recorded for further analysis in addition to our field notes. During the observations, we were in the surgical theatre with freedom of movement to observe the operations and the use of the images. We also had the opportunity to ask questions of the surgeons at appropriate moments during the surgery regarding image use. A total of 25 hours of observations over five different surgeries were conducted. Each surgery had at least two neurosurgeons – throughout our findings we refer to the consultant surgeon as S1, the resident surgeon as S2, and if there was a second consultant surgeon we refer to him as S3. As a further step in data collection and validation, the neurosurgeons were presented with findings and given opportunities for comment during the interview.

As neurosurgery is concerned with any portion of the nervous system including the brain, spinal column, spinal cord, peripheral nerves, and extra-cranial cerebrovascular system, the extent to which images can be used in the context of surgical practice are quite varied. For our needs we focused on those procedures that were reliant on images and, across the array of procedures observed, the use of a variety of imaging technologies. Again, the focus of the analysis presented here is on how the neurosurgery team viewed and constructed the body through the images.

FINDINGS

For the following, we will specifically discuss three vignettes that occurred in three different surgical procedures: a keyhole spinal fixation surgery, an endo-nasal spinal tumor excision, and a brain tumor biopsy. All of these surgeries are reliant on various imaging techniques in order to 'see' the interior of the body and perform the necessary procedures. Each of the vignettes also entails a relationship with movement of tools and tactile perceptions of the patient body. They differ however, in the closeness of this relationship.

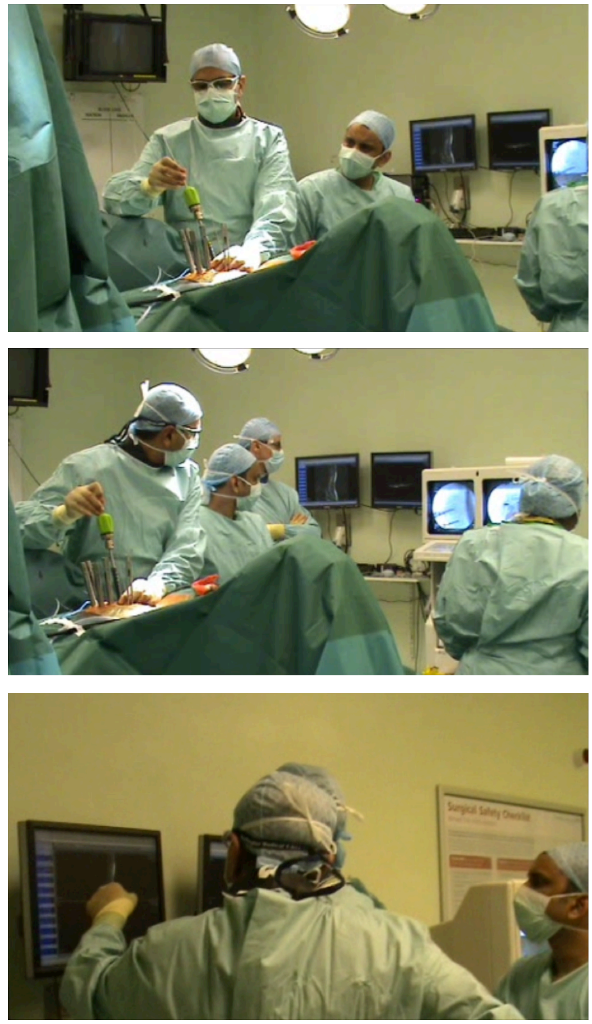

Figure 1. Seeing as a Material Achievement

Seeing as a Material Achievement

Oftentimes the use of images in the context of surgical intervention entails the visualization of instrument placement with regards to specific anatomy. This visualization is reliant on a close relationship between the manipulation of the instruments in the body and the image production. This allows for the images to have a direct relationship to the action of the surgery and a relational aspect to the anatomy in question. However, there are other times when preoperative images are brought to bear. These images typically have no relationship to the point of surgery and therefore have to be constructed into the fold of the bodily manipulation and other image seeing practices. The job of the surgeon then is to do the work of constructing a story that takes in all of the evidence in order to make a decision. In the following vignette, we see the use of two imaging systems – each viewing the body in very different ways, but, through iterative viewing of the images, discussion between the surgical team, and physical manipulation and touch of the body, the surgeon constructs a vision of the unseen focus of work.

During the keyhole spinal fixation surgery we observed, X-rays were produced from a mobile unit called a C-arm. This provided maximum flexibility in positioning the imaging system around the patient (e.g., lateral or coronal). A workstation unit used to store and view the images was connected to the C-arm through a long lead, which allowed the screens to be positioned at the end of the surgery table for the surgeons to view the output. A technician was on hand to take a new X-ray when the surgeon requested one. A keyhole spinal fixation surgery consists of several vertebrae being anchored together with a device in order to reduce vertebral mobility. Small half-inch incisions are made in the back where smaller surgical instruments are passed through and thus the need for the continual X-ray images for identifying the location of the instruments in the spine.

The surgery was proceeding on pace with almost all of the fixation screws having been placed. Towards the end of the surgery, though, the primary surgeon (S1) realized one of the wires used for guiding the screws is loose. After a bit of inspection, it became clear that one of the inserted wires had slipped out of a hole drilled into the patient's spinal pedicle and the surgical team was forced to decide whether or not they should continue with inserting that screw or continue on without it.

As the other two surgeons look on, S1 pulls the wire out. He inspects the tip for a moment and then hands the wire to the nurse. He puts the pedicle tap back into the hole and holds it vertical. He then turns to the X-ray displays.

> *S1: Show me that? [He angles the pedicle tap slightly to the left.] Show me that? Give me a wire?*

S1 inserts the wire through the hollow pedicle tap back into the pedicle. He slowly prods the opening with the end of the wire while looking down at the body. He removes it slightly and then puts it back in with small prods. He then holds it in place and turns to the display of x-rays.

> *S1: Show me that?*
>
> *S2: There should be a hole in the pedicle.*
>
> *S1: Yeah, the thing is, I think it is uh, it slipped out.*

He repeats this vigorous manipulation of the wire and viewing the scans a number of times pausing on several occasions, presumably in contemplation. Then, inserting and prodding the wire, again, and pushing down on it:

> *S1: Show me that? [pushes more] Show me that? [pushes more] Show me that? [prods a bit more and then takes the wire out and hands it to the nurse and begins to unscrew the pedicle tap] Can I have [another] wire?*
>
> *S2: The angle was right.*
>
> *S1: Yeah, well the angle was right, but it is out.*
>
> *S1: [Probing with the new wire] I don't feel the thread where it slipped out. ... Ah. [holds in place] show me that? [all turn to look at the x-rays] Yeah, I think this is uh, somewhere along the way it slipped out. [takes out the wire] Take that please, Grace [to nurse]. [Turns to S2] Did you feel that the bone was soft on your side?*
>
> *S2: Soft. Yeah, it was soft.*

S1 starts to prod with the wire again. He leans into the body and continues to prods with the tool.

What emerges over the course of this excerpt is that the wire may have been displaced because of a problem with the patient's bone. Rather than just slipping out by surgical mistake, the wire seems to have come out because of a worrying softness in the bone (a sign of cancer degeneration). With the goal of clarifying this, the two consultant surgeons walk over to a PACS system, mounted on a wall, to view the patient's pre-op MR images. The system has two screens where the one on the left shows a side view and the one on the right shows the corresponding cross-section.

Removing his glove, S2 uses a mouse to navigate the images on the screens, talking with S1 to locate the area of interest. Once they have found the area of the spine in question, they discuss the quality of bone and whether they can still do what they set out to do. In the midst of this, S1 returns to the table and probes with the wire, this time purposefully using it to determine the softness of the area. After a little while, he returns to the MR images with his colleagues to assess if the bone is diseased.

> *S2: It doesn't look that bad [on the MRI].*
>
> *S1: I'm not so sure…*

In this sequence of actions and talk, we see the reliance on images, but also quickly glean a sense of the need for other forms of intervention. The X-rays are meant to confirm the anatomy and position of the bony structures of the spine as well as support the planned surgery. The surgeons are, in effect, manipulating the shapes on the screen through the use of their hands on the body.

However, the X-rays do not provide enough information when faced with the problem of whether a bone is diseased or not or whether it is solid enough to hold a screw. Indeed, both the X-rays and pre-op MRIs failed to identify any problem beforehand. It takes some unmistakably physical probing with wires and the pedicle tap, as well as a good deal of talk to build up a representation of the body, to 'see' the body, so to speak. The questions of whether the bone was diseased; whether the wire had simply slipped out (a problematic event, but solvable); and whether the bone could hold the screws are not found within an image. Nor are they answered through a number of images. The necessary vision of the body emerges through the unfolding relations between the surgeon's tactile prodding of the body, an iteratively taken set of X-rays in relation to the body for location guidance, and the multiple viewings of MR images, sequenced to produce a view of bone density. Moreover, a sequence of orientating statements, questions and answers between the surgical staff weave between these activities, in all, producing an image – although it has to be said fuzzy image – of the body.

What is more, as the MR images are physically separated from the body, exploratory activities are undertaken by the surgeon moving back and forth between stepping through the pre-operative MRI slices and the body – inserting, feeling, moving, and reflecting. This interleaved exploration is how images come to constitute surgical practice and the professional surgical vision. Not simply a view of the body, but an object (or set of objects) brought into relation with others and set in a sequence of actions that then enacts a view onto and into the body.

Seeing-in-Motion

In the second of our examples, the imaging system being used is not producing new images throughout the surgery, but instead is using images as a basis for localization and directional planning. Again we see that the use of the imaging system is not simply an act of viewing, but an act of construction. This construction is maintained through a tight coupling between the manipulation of instruments in and on the body and corresponding changes to the displayed images.

For the brain tumor biopsy, the surgical staff are performing a craniotomy (in which openings are created in the skull) and introducing a thin syringe into the brain matter to retrieve a sample. In order to guide the syringe, the surgical team uses a Medtronic StealthStation® Treon™ Navigation System (from here referred to as the StealthCam). During surgery, this system displays the real-time location of a probe on pre-existing, pre-op CT scans. Thus, it allows the operator to determine where and to what depth something such as a syringe should be inserted *vis-à-vis* the patient's external body and internal organs.

To accomplish this in practice, an infrared camera is positioned and calibrated with respect to the patient's head (the latter fixed in position using a clamp). The probe is then inserted into the opening in the skull, and static images of the probe's tip superimposed on the CT scans are produced by depressing a foot-pedal. The location is represented on three planes on a connected display: axial, coronal, and sagittal.

In our observations, the assisting surgeon (S2) begins by touching the tip of the probe to the outside of the brain matter. He makes a series of touches within close proximity to one another as he steps on the footpedal and looks at the corresponding location on the StealthCam screens. Once he identifies an optimal location for insertion, he begins to go deeper into the brain matter for further probes to determine an optimal trajectory: inserting the probe, stepping on the pedal, looking at the screen, removing the probe, changing the angle ever so slightly and inserting it again. When satisfied with an angle of approach, the surgeon inserts the probe down to the point of the tumor – continuing to monitor his progress on the StealthCam as he descends. When he stops, S1 and S2 look at the display.

Figure 2. Seeing in Motion

S1: Is that good? Yeah, OK? This is the entry here. [points to screen]

S2: So it is in line to [he squints and leans towards the display] 48mm.

S1: That seems like an awfully long way. What is that for?

S2: Because you want to take from the depth of it.

S1: From the middle of it?

S2: Yeah.

S1: [Considers this for a moment and then nods] Let's see if it comes out normal.

After they put a marker on the biopsy needle at the appropriate depth – which also can act as the probe in conjunction with the StealthCam – S2 slowly inserts the biopsy needle to extract the tissue for testing. S1 looks back and forth between the screen and the procedure site.

S1: So you must be there. Is that the depth?

S2: [Stops and looks up at the StealthCam display] Yeah. [removes tool]

Both look at what was extracted and seem happy that it is not brain tissue and appears abnormal (as expected). The sample is given to the nurse to send to the lab for testing.

The point of interest here is that the body, and in this case the patient's brain, is being produced as a space to navigate.

Unlike the previous example where images are being configured to locate a source and examine the features of that source, in this case the representation is being enlisted to coordinate the movement of a tool through the body. That is, the StealthCam is operated to compose not only a spatial representation of the body, but movement through it.

Thus, the surgeon devises an understanding of the spatial trajectory of work. A vision of the space to be navigated is pieced together through movement. The movement of the probe helps to construct a representation of the area, spread over the three planes on the StealthCam display. But, there is still negotiation as to where to enter and where to stop the descent; the images do not indicate 'start here' or 'end here'. As we can see from both the decision making with the probe and the subsequent discussion between the two surgeons, where to start and where to end is constructed in relation to the images and agreed between the surgical team. In fact, this process started much earlier, at the point of decision as to where to perform the craniotomy. Creating a hole in the skull of the patient in an appropriate location and size required the same process of planning with multiple measures with the StealthCam probe.

In practice, then, the instruments (and the other assembled agents and actors) are brought into relation to one another to enact particular visions of and into the body. The surgical staff find themselves manipulating the instruments they have to hand to produce the desired effects on the images, to 'see' the body in the way they need to.

Seeing as a Relational Achievement

Above, we have observed the need for multiple imaging systems during an operation. We have also seen the surgeon using physical manipulation and haptic senses to make sense of the images as well as coordinate his movements. In the following example, we see a similar use of several imaging systems and a physical engagement with the body being operated on. However, what we want to draw out here are the ways multiple representations and material interventions are used in concert with one another. More particularly, we want to show how a 'professional vision' of the patient's body is achieved through an unfolding set of relations between hands, instruments and images (as well as, as seen above, the interactions between surgical staff).

Performing a spinal tumor excision, the neurosurgery team is using an endonasal approach – i.e. using the nostrils as the entry point – to create a corridor through the nasal passages into the back of the sinus to the point of the tumor. To aid this work, both the StealthCam and an endoscope are used. The latter consists of a tiny camera at the end of a long flexible tube. It has its own light source and displays the captured image on a large monitor at the surgical table.

During the procedure, the surgeons stand by the operating table with their eyes affixed on the screens before them. Their movements with the tools in their hands are in relation to the regularly updated images. They continuously watch the endoscopic display – observing the progress the surgical instruments are making in creating a corridor to the tumor. The StealthCam probe lies within the cavity. After a bit of drilling, they retreat slightly and move the StealthCam's probe towards the back of the corridor. Both surgeons shift their eyes to the StealthCam display. S1 moves the probe, steps on the pedal and repeats this act a few more times.

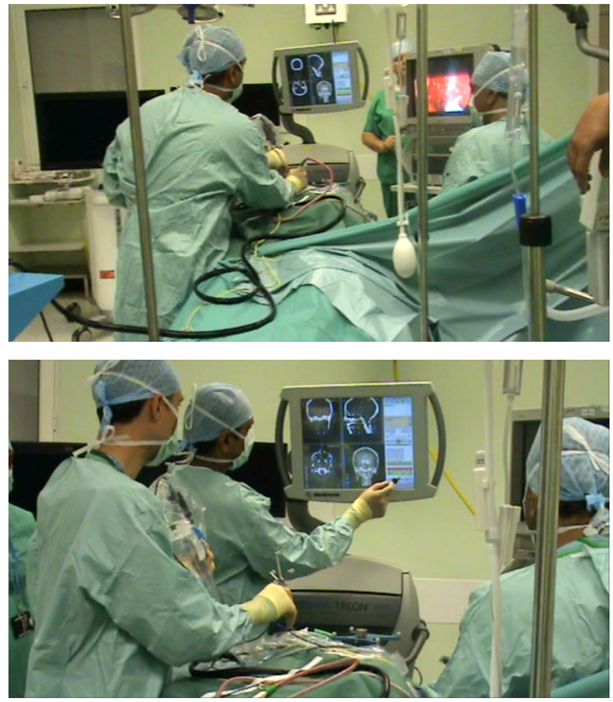

Figure 3. Seeing as a Relational Achievement. Watching the endoscope video (top); using the StealthCam (bottom).

S1: I think ... move up a bit and to the right.

S2: Yeah, OK. Much further than we are.

S1: Yeah.

They shift their attention back up to the endoscopic display, pull back the probe, and move the drill back into place to continue drilling. They continue this process throughout the first stage of the surgery – shifting their attention between the StealthCam display and the endoscopic real-time view, and, at the same time, making micro adjustments to their interactions with the patient's body.

Crucially, we see that the endoscopic video and the StealthCam are used in coordination with the surgeon's actions. The visual representations only make sense together, and in combination with the surgeon's physical exploration of the body and temporal-spatial movement through it. Critical to the surgery then is the co-registering not only of the different images, but also the material sensations felt through instruments and the surgeon's fingers and hands.

As the drilling work progresses through to the back of the nasal passage, the surgeon changes the way in which he is operating the instruments. Whereas earlier his drilling was in extended bursts and a good deal of force was used to move the drill through the bone, he now begins to apply shorter and lighter drilling movements while viewing the endoscopic display. These movements are interspersed with moving the StealthCam probe into position to check the depth of the drill. Thus, we see a repeated process of drilling, stopping, placing the probe, scrutinizing the StealthCam images, tapping the bone with the probe and then starting to drill again. Eventually the point is reached where soft tissue is felt instead of bone, and the surgeon stops drilling and presses the instrument forward as he stares at the endoscope display. He then turns to the StealthCam display and moves the probe forward to the place where the drill has just been. Making these real-time comparisons between touch and spatially dislocated visual feedback, he becomes satisfied that the spinal cavity has been reached and that a larger corridor can be made to remove the tumor.

In this way, we catch sight of how 'seeing' the body is emergent, achieved through a sequence of actions and interactions. The surgeon juxtaposes the various real-time and pre-operative images against one another, and sets these against his own material sensations felt from manipulating the instruments. Exploratory actions with both the body and the images come together to form an understanding of what to do next. The StealthCam conveys the location of the activity, while the endoscope provides a view onto the action and an indication of the type of tissue being acted upon. Moreover, these representations are made meaningful through the concurrent interventions of drilling, tapping, applying gentle (and sometimes not so gentle) pressure, etc. Without such coordinated actions and the entangling of actions, tools and visual representations, the images remain passive and, arguably, of limited use in the surgical work.

In fact, the distinctions between representations and interventions appear blurred in the surgical work, with the visual (and other sensual information) shifting between both passive object and interventional tool (see [12]). At times the displayed images are passive representations of the area being worked on, but at others they serve as tools for navigation. Similarly, the body is both the 'object' being acted upon and an assortment of different textures and resistances, thus taking on the role of guiding instrument. Seeing the body then is an active accomplishment and one in which an assemblage of artifacts and sensations are in a continuous process of being the object(s) under examination and the tools with which the surgery is conducted. The body is only 'seen' insofar as this coordinated and unfolding work occurs. Seeing, if you will, is achieved over time, in materially bound action, and through the relations between multiple actors and artifacts.

DISCUSSION

Our work has aimed to demonstrate the nuanced role played by medical imaging in surgical procedures. It aimed to show that the evermore sophisticated ways of scanning and displaying medical images are only part of the act of 'seeing' the body, of constructing a professional vision of the body as a site for medical work. Bound up in such acts of seeing are an interleaving and emergent set of relations between hands and instruments on the patient body and the separated images that give shape to the body and how it must be operated on.

The broad implication, here, is that medical imaging systems, such as those used throughout surgical interventions, should be seen not as standalone technical achievements, 'solving' *the* problem of seeing into the body; rather, imaging the body must be treated as *one* of an assemblage of activities used to constitute a professional vision in surgical practice. Moreover, 'seeing' the body appears to be more than just capturing a visual representation of some body organ or part. Seeing has different features to it. In our second example, for instance, we saw how seeing can be about navigation, about how to work one's way through different organs, and plan for an intervention. This seeing-in-motion can be contrasted with efforts to determine the condition of an organ, to extract a tumor, etc. The point here is imaging systems should be designed to accommodate and support these different forms of seeing as opposed to treating the professional vision as a singular achievement.

Our last example draws these points out still further. It conveys how 'seeing' is performed or enacted, that it is an unfolding achievement accomplished by continually setting certain sensations, actions and tool-use on the body and the separated images in relation to one another. The body is not simply captured and represented. Seeing a patient's body is something that progressively unfolds through feeling the changing texture of tissue, the resistance of movement, the pressure applied to a tool, assessing the updated images, etc. Imaging systems then might be best designed not to indicate a somehow final and complete view of the body, but rather to suggest the view into the body as something constructed through coordinated and largely embodied or material interactions. Seeing the body might be represented in an interactive system as unfolding and assembled, where the material interventions are part of 'seeing', not just performed in support of seeing.

This perspective on medical imaging systems stands in stark contrast to the common notion that surgeons simply need to view the data and that design's role is to optimize the visualization of such data. Our work indicates that images are not full simulacrums of the necessary data and that the problem is not simply about accessing this information; rather, images are useful only as a part of an unfolding practice with relation to the patient body. Thus, whereas the emphasis in the imaging sciences is on improving the clarity and detail of image representations, our findings show the necessity of supporting 'seeing' as an embodied and unfolding practice in surgery. By privileging image fidelity and detail over the active use of images, one falls susceptible to forgetting the importance of exploration, interpretation, and negotiation in the practice of surgery [14] [20] [13]. Moreover, one overlooks the work involved in constructing a professional vision of a patient's body in surgery, of actively 'seeing'.

Interactive Vision Systems in Medical Seeing

In closing we want to discuss this standpoint *vis-à-vis* some of our own ongoing interests in using Microsoft's Kinect as part of a surgical imaging system. Particularly we want to elucidate the design of interaction mechanisms for medical images in tandem with interactions with the patient body. While space does not permit us to go into the technical details, we want to draw out just a few of the issues the work above raises for the implementation of such real-world systems.

It is evident from a small but growing body of research that touchless interaction, facilitated by vision systems such as Kinect, are providing innovative solutions for manipulating (and annotating) images by the surgical team at the operating table [13] [20]. Sophisticated techniques are being developed that support gesture-based manipulation of two-dimensional images or three-dimensional renderings, modifying the addition or presentation of various overlays, and moving images between screens with a wave of the hand.

These solutions, however, remain centered on viewing and manipulating images of the body. Aiming to extend the application of Kinect beyond this, our work points to the value of vision systems in supporting the active forms of 'seeing' we have so far discussed. For instance, it encourages design possibilities that may in some way account for interactions with the body, so that manipulations of the body also alter the representations – supporting a tighter integration of hand-work, tool-use and images.

For example, the trajectory of instruments and tools could overlay the three-dimensional views of the brain produced using pre-operative CT scans as is currently supported through the StealthCam. An innovation here would be for the image representation to change perhaps in orientation or detail in relation to the bodily orientation of the surgeon with regards to the patient body or in relation to the instrument-based work that unfolds; in other words, that manipulations of the body and instruments on the body in kind manipulate the representations. For instance, during a spinal tumor extraction, the orientation of the images presented on an overhead display matches that of the orientation of the surgeon. When the surgeon is on the left side of the body, the images are presented in front of the surgeon with the head to the left. When the surgeon switches sides in order to have a better angle of approach, the images are presented to the surgeon with the head to the right.

Another example would be in the repair of an aneurysm. Based on the current trajectory of instruments, a three-dimensional rendering of the brain rotates to show the view from the opposing side of the approach. In a case such as this, the surgeon is not concerned with what they can physically see before them due to the craniotomy, but are concerned with what they cannot see on the other side of the aneurysm. Thus, with the representation being manipulated by the hand-work of the surgeon, he can see what lies on the other side of the aneurysm – seeing what they cannot see for themselves in relation to their current actions and trajectories.

In this sense, the image of the body is treated not as something singular, but as multiple and evolving. Moreover, any effect on the representation would aim to be dependent on the imaging technique used, so that the gestural interaction would complement the particular properties and uses of the different imaging systems. For example, using pre-operative images, the work with instruments may be more instrumental as an aid for pointing and annotating anatomy – allowing for exposition of the images during surgery. Whereas intra-procedure imaging that provide real-time feedback would be more passive in their presentation of data based on the surgeon's bodily orientation to the patient's body and anatomy. Not only is the intention in this to give serious consideration to the coupling of gestures and images for each imaging system. It also encourages a sensitivity to the multiple and concurrent use of systems, and how each might represent the interventional work in different but complimentary ways.

While these ideas are relatively abstract and general, our hope is they provide an indication of what might be envisioned when the practice of embodied seeing of medical images is carefully considered. As further advances are made, the possibilities will certainly become more sophisticated, but the lesson here is not to confine the development of imaging systems to the representation of data, but rather, the interaction that can be afforded with that data in relation to the body during the surgical procedure.

CONCLUSION

Intraoperative imaging systems for minimally invasive surgery are changing the professional vision of the surgeon. It is changing what it means to perform surgery on the body and to 'see' the body through new imaging techniques. In our studies of this new practice we have found that the 'seeing' of images is an embodied process achieved through a coordination of visual information about the body and instruments and explorative actions with instruments on and in the body.

From our observations of imaging systems in neurosurgery we elucidated the relationship between how surgeons 'see' images in the context of surgical practice and embodied exploration as a part of professional vision. We saw how images do not hold within them the information that simply needs to be accessed; that images, no matter their form or function, are useful only as a part of practice. Thus, we showed that (1) there is a close relationship between the manipulation and interaction with the body and the use of the images and (2) images are not seen, but constructed through this relationship. The importance of the tightly related actions with medical images contrasts with the popular notion that surgeons simply need to view the data; thus our intention has been to lay a foundation for the innovation of new interactive mechanisms for intraoperative image use.

REFERENCES

1. Alač, M. (2008). Working with brain scans: Digital images and gestural interaction in fMRI laboratory. *Social Studies of Science, 38*(4), 483-508.
2. Beaulieu, A. (2002). Images are not the (only) truth: Brain mapping, visual knowledge, and iconoclasm. *Science, Technology & Human Value, 27*(1), 53-86.
3. Burri, R.V. (2008) Boundary work and symbolic capital in radiology. *Social Studies of Science, 38*(1), 35-62.
4. Crary, J. (1990). *Techniques of the observer: On vision and modernity in the nineteenth century*. Cambridge, MA: MIT Press.
5. Golan, T. (1998). The authority of shadows: The legal embrace of the X-ray. *Historical Reflections, 24*(3), 437-458.
6. Goodwin, C. (1994). Professional vision. *American Anthropologist, 96*(3), 606-633.
7. Goodwin, C. (2000). Practice of color classification. *Mind, Culture, and Activity, 7*(1-2), 19-36.
8. Goodwin, C. (2003). The semiotic body in its environment. In J. Coupland & R. Gwyn (Eds.), *Discourses of the body.* New York: Macmillan.
9. Heath, C. & Hindmarsh, J. (2000). Configuring action in objects: From mutual space to media space. *Mind, Culture and Activity, 7*(1-2), 81-104.
10. Hindmarsh, J. & Heath C. (2000). Embodied reference: A study of deixis in workplace interaction. *Ournal of pragmatics, 32*, 1855-1878.
11. Hirschauer, S. (1991). The manufacture of bodies in surgery. *Social Studies of Science, 21*(2), 279-319.
12. Hacking, I. Representing and intervening: Introductory topics in the philosophy of natural science. Cambridge: Cambridge UP, 1983
13. Johnson, R., O'Hara, K., Sellen, A., Cousins, C., & Criminisi, A. (2011). Exploring the potential for touchless interaction in image-guided interventional radiology. *Proceedings of the Conference on Human Factors in Computing*, Vancouver, Canada (pp. 3323-3332).
14. Joyce, K. (2005). Appealing images: Magnetic resonance imaging and the production of authoritative knowledge. *Social Studies of Science, 35*(3), 437-462.
15. Kaplan, B. (1995). Objectification and negotiation in interpreting clinical images: Implications for computer-based patient records. *Artificial Intelligence in Medicine, 7*, 439-454.
16. Kassirer, J. P. (1992). Images of clinical medicine. *New England Journal of Medicine, 32*(6), 829-830.
17. Latour, B. (1990). Drawing things together. In M. Lynch & S. Woolgar (Eds.), *Representation in Scientific practice* (pp. 19-68). Cambridge, MA: MIT Press.
18. Lynch, M. (1985). Discipline and the material form of images: An analysis of scientific visibility. *Social Studies of Science, 15*, 37-66.
19. Lynch, M. (1990). Drawing things together. In M. Lynch and S. Woolgar (Eds.), *Representation in a scientific practice* (pp. 19-68). Cambridge, MA: MIT Press.
20. Mentis, H.M., O'Hara, K., Sellen, A., & Trivedi, R. (2012). Interaction proxemics and image use in neurosurgery. *Proceedings of the Conference on Human Factors in Computing Systems* (CHI), Austin, Texas, New York: ACM.
21. Nishizaka A. (2000). Seeing what one sees: Perception, emotion and activity. *Mind, Culture, and Activity, 7*(1-2), 105-123.
22. Pasveer, B. (1989). Knowledge of shadows: The introduction of x-ray images in medicine. *Sociology of Health and Illness, 11*(4), 360-381.
23. Prasad, A. (2005). Making images/making bodies: Visibilizing and disciplining through magnetic resonance imaging. *Science, Technology & Human Values, 30*, 2, p. 291-316.
24. Stafford, B. (1991). *Body criticism: Imaging the unseen in Enlightenment art and medicine.*Cambridge, MA: MITPress.
25. Suchman, L. (2000). Embodied practices of engineering work. *Mind, Culture and Activity, 7*(1-2), 4-18.
26. Ueno, N. (2000). Ecologies of inscriptions: Technologies of making the social organization of work and the mass production of machine parts visible in collaborative activity. *Mind, Culture, and Activity, 7*(1-2), 59-80.
27. Vlaardingerbroek, M. & Boer, J. (1999). *Magnetic resonance imaging: Theory and practice*. New York: Springer.
28. Warwick, A. (2005). X-rays as Evidence in German Orthopedic Surgery, 1895-1900. *ISIS, 96*, 1-24.