Virtually Transparent Epidermal Imagery (VTEI): On New Approaches to *In Vivo* Wireless High-Definition Video and Image Processing

Adam L. Anderson, *Member, IEEE*, Bingxiong Lin, *Student Member, IEEE*, and Yu Sun, *Senior Member, IEEE*

Abstract—This work first overviews a novel design, and prototype implementation, of a virtually transparent epidermal imagery (VTEI) system for laparo-endoscopic single-site (LESS) surgery. The system uses a network of multiple, micro-cameras and multi-view mosaicking to obtain a panoramic view of the surgery area. The prototype VTEI system also projects the generated panoramic view on the abdomen area to create a transparent display effect that mimics equivalent, but higher risk, open-cavity surgeries. The specific research focus of this paper is on two important aspects of a VTEI system: 1) *in vivo* wireless high-definition (HD) video transmission and 2) multi-image processing—both of which play key roles in next-generation systems. For transmission and reception, this paper proposes a theoretical wireless communication scheme for high-definition video in situations that require extremely small-footprint image sensors and in zero-latency applications. In such situations the typical optimized metrics in communication schemes, such as power and data rate, are far less important than latency and hardware footprint that absolutely preclude their use if not satisfied. This work proposes the use of a novel Frequency-Modulated Voltage-Division Multiplexing (FM-VDM) scheme where sensor data is kept analog and transmitted via "voltage-multiplexed" signals that are also frequency-modulated. Once images are received, a novel Homographic Image Mosaicking and Morphing (HIMM) algorithm is proposed to stitch images from respective cameras, that also compensates for irregular surfaces in real-time, into a single cohesive view of the surgical area. In VTEI, this view is then visible to the surgeon directly on the patient to give an "open cavity" feel to laparoscopic procedures.

Index Terms—Analog systems, biomedical devices, high-definition video, image processing, wireless communications.

I. INTRODUCTION

MINIMALLY INVASIVE SURGERY (MIS)—utilizing small incisions in the body for placement and manipulation of surgical equipment—has been widely adapted and performed as an alternative to open-cavity surgery because it provides a tremendous public benefit for: minimizing trauma, shorter hospitalizations, and faster recoveries; however, these operations often take longer to complete than equivalent open operations, with associated patient risks to contamination. The MIS procedure also poses challenges to surgeons in many aspects: limited view and limited number of view points fixed by the insertion sites, an overhead monitor that displays the video from the videoscope but does not have a consistent and clear indication of orientation of the video. Spawning from MIS is Natural Orifice Transluminal Endoscopic Surgery (NOTES) or Laparo-Endoscopic Single Site (LESS) surgery—recently developed MIS techniques—whereby "scarless" abdominal operations can be performed with multiple endoscopic tools passing through a natural body orifice, such as the umbilicus, as the insertion point. Though there are advances in MIS-related equipment [1]–[6], including flexible tip endoscopes or robotic surgical platforms, surgeons often have to rely heavily on their experience to sense the locations of tools relative to the internal surgical area.

We are currently working on a virtually transparent epidermal imagery (VTEI) system [7], [8] as an enhancement to NOTES or LESS surgery, that will be composed of several micro wireless cameras to mitigate the bottleneck issue. Wireless image sensors are inserted at the beginning of the procedure and tethered to the inside of the abdomen. Tethering needles are also used to power the devices though newer versions of the sensors will be completely non-invasive (i.e., no additional puncture points). These video images from the wireless laparoscopes are transmitted wirelessly to some external receiver and then reprojected onto the abdomen of the patient. This gives surgeons a disorientation-free view of the surgical area and lessens the dexterity issue of these advanced surgical procedures.

There are many open problems to resolve before surgeons can benefit from a full-HD system on real patients. The geographical locations of the wireless endoscopes need to be carefully planned before surgery to provide an optimal view of the surgical regions, instrument pathway, and other related regions. How to use both imaging system and efficient display of videos need to be explored in future work. Improvements to the current image-based rendering approach to generate more reliable mapping with less distortion—the current approach is very sensitive to the accuracy of feature detection. An exhaustive list of possible research directions is not appropriate here but the richness of the research applications should be apparent. This current paper will focus on two key aspects of the VTEI system that need to be addressed: 1) *in vivo* wireless

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

2 IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS

high-definition (HD) video transmission and 2) multi-image processing. In other words, *how do we get the images out of the body and what do we do with them once they are out*?

Traditionally the focus of digital wireless communications schemes, including real-time wireless HD video, has been to maximize the data rate over some bandwidth with the minimum amount of power necessary [9]. These metrics of power, rate, and bandwidth are at the heart of wireless communications since these are also primarily the scarce resources in the system. While this is true is most cases the current work studies a scenario where this traditional approach to optimized wireless links may not be as applicable as other system metrics such as in *in vivo* wireless biomedical devices. *In vivo* biomedical devices are almost intrinsically wireless in nature due to the medium in which they reside—our own bodies. The approaches taken to establish a wireless link for data transfer with an *in vivo* device or sensor can take the form of different wireless methods. One such method uses various forms of induction in order to create a link between the device inside the body and the outside world [10]–[12] which can also be applicable to power coupling or harvesting. Due to the highly absorptive nature of human tissue, the short transmission ranges, and low required data rates in some sensors, acoustic transmissions have also been proposed as a method to procure data [13], [14]. Finally, classical radio-frequency (RF) wireless links have been proposed for use in *in vivo* biomedical devices [15]–[17] that include their own strengths, often due to the maturity of the field, but also some weaknesses specifically from this unique wireless medium.

The motivation for working on this aspect of the VTEI system is feasibility of a zero-latency high-definition (HD) wireless video sensor that has the additional requirement of a small hardware footprint. With these requirements of HD video in a tiny footprint, approaches such as with acoustic and magnetic induction are most probably precluded and classic RF approaches remain the primary possibility. Several RF approaches have been taken to HD wireless in non-biomedical related fields [18]. To avoid latency, some have proposed uncompressed digital transmission, which are high bandwidth, but are transmitted at a high carrier as well to compensate [19]. Commercial companies like SiBeam are pushing for this type of data transmission in the WirelessHD or WiHD standardization. In addition to some latency, the large chipsets do not satisfy the footprint requirement and the high 60 GHz carrier frequency may have issues propagating through human tissue. Another school of thought uses proprietary encoding or other techniques like UWB and transmits in lower frequency bands [20]–[22]. The commercial company Amimon is pushing for Wireless High-Definition Interface or WHDI standards at 5 GHz for uncompressed HD video. While this will help with absorption issues at 60 GHz, the chipsets are still too large and the reported latency may be unacceptable for the surgical scenario in question.

We propose an alternative approach to HD wireless specifically for *in vivo* devices but practically applicable to other high data rate links. The first concept of this approach is to keep the video signals in their analog formats prior to transmission. This allows significant savings in hardware by reducing the re-quired analog-to-digital (ADC) conversions and latency caused by data compression and error correction. Voltage-division multiplexing (VDM) is proposed to transmit multiple analog signals simultaneously in their "voltage space". Instantaneous orthogonality of VDM encoded signals is enforced by conversion of the VDM signals via frequency modulation; thus, the proposed scheme is referred to as Frequency-Modulated Voltage-Division Multiplexing (FM-VDM). By exploiting unused portions of the spectrum in an instantaneous manner multiplexing of analog signals is allowed in an otherwise occupied bandwidth. This overlapping of spectrum will potentially allow several analog signals representing the HD video to be transmitted in a bandwidth that is otherwise impossible and also in a small hardware footprint at the transmit sensor due to the minimalistic nature of analog FM radio circuits.

After reception of the HD video signal from various sensor sources, image processing must be carefully performed to give surgeons the requisite view of the surgical field. Various approaches [23]–[26] for visualization in image-guided interventions have been proposed to achieve "seeing through" effect. Applying the concept of augmented reality has obvious benefits, including: enabling the surgeons to focus on the surgical site without dividing his or her attention between the patient and a separate monitor, providing hand-eye coordination as the surgeon observes the operation area, and giving a virtual feel of open-cavity procedures. For example, a computed tomography (CT) image of a patient overlay with the patient, and appearing at the location of the actual anatomy, is a possible use of advanced imaging [27], [28]. The location of the surgeical tool can be tracked and graphically drawn as a virtual tool overlayed on the CT image to guide surgeons during operations [29], [30]. For obvious reasons, if the mapping does not align correctly with the patient and the surgical tool, the visualization could be dangerous. It is very challenging to achieve satisfactory, accurate alignment between the tracking data and the image since it requires precise models of the patient and models of instruments.

This work presents a novel application of a Homographic Image Mosaicking and Morphing (HIMM) algorithm in order to achieve the imaging goals of the VTEI system. We realize three special properties of our system: the wireless cameras we use are small, the depth change of the scene is not large and the cameras are anchored on the inflated abdomen wall, which is "far" from the scene. Based on these three properties, we will be able to apply the traditional image mosaicking technique to generate a panoramic view [31] and do so in a real-time fashion. Additionally, the proposed HIMM algorithm will have a tight correspondence between images thus mitigating the surgical tool overlay problem seen in other approaches.

Section II overviews the current state of the VTEI system and provides details on the laboratory prototype implementation. Section III presents a mathematical description of the theoretical FM-VDM as well as a possible detector circuit and system derived from traditional FM PLL detection with parallel feedback. Section IV describes the image processing algorithm used to carefully mosaic the received signals into a cohesive view using lower resolution COTS wireless cameras. The paper is concluded in Section V.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ANDERSON *et al.*: VTEI: ON NEW APPROACHES TO *In Vivo* WIRELESS HIGH-DEFINITION VIDEO AND IMAGE PROCESSING 3
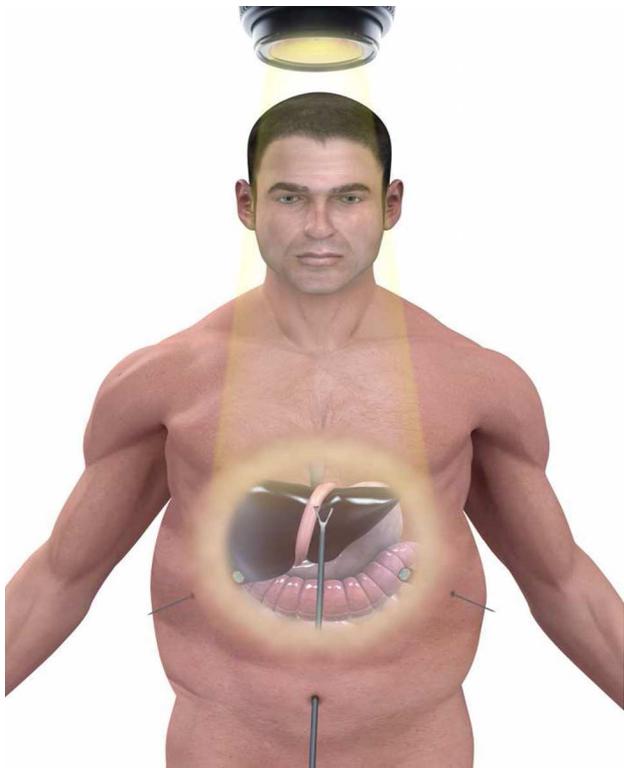

Fig. 1. A visual representation of the virtually transparent epidermal imagery (VTEI) system. Implanted wireless camera images are reprojected onto the patient to give the impression of open cavity viewing. All medical devices involved in the procedure are sized such that they can be inserted and removed through an incision in the umbilicus.
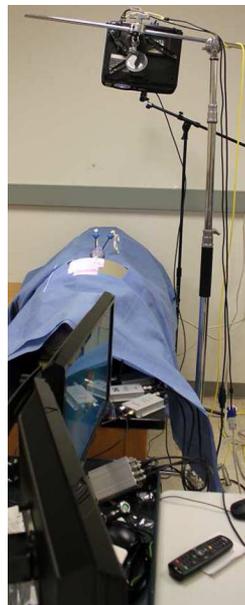
Fig. 2. An expanded view of the VTEI lab prototype. Care was taken to recreate the feel of an actual surgical room.

II. VIRTUALLY TRANSPARENT EPIDERMAL IMAGERY (VTEI)

We present the VTEI approach that composites several *in vivo* micro-cameras that capture the surgical anatomy and the surgical instruments at the same time and in the same frame (see Fig. 1 for a graphical representation). This approach does not encounter the difficult instrument mapping and alignment problem seen by other advanced augmented reality approaches and more closely approaches the "feel" of open cavity surgeries. The goal of the current lab prototype is two-fold: 1) develop an image-based rendering approach to mosaic the videos from wireless scopes to form a panoramic view of the surgical area and then project this panoramic video onto the outside abdominal wall and 2) quantize the benefits of this approach in a surgical training environment led by one of the thought leaders, and current team mate, in LESS surgeries—Dr. Alex Rosemurgy, M.D. However, the goals of the current prototype do not cover the research problems addressed in Sections III and IV.

A. Lab Prototype

The goal of the system is to provide a visual feedback to the surgeons about where the surgical instruments are relative to the *in vivo* organs. A network of wireless cameras are placed inside the abdominal wall model via serial insertion through the trocar. Since the cameras are anchored on the abdomen with a thin needle (<1 mm), they will leave no scar on live models. The panoramic video merged from the wireless cameras are then processed to have the correct orientation so that the abdomen

projection provides a correct hand-eye correlation for surgeons. Fig. 2 shows a photo demonstrating the VTEI lab prototype; care was taken on creating a prototype area that mimics a surgical room.

Fig. 2 is a test platform setup with a surgical simulator, which has a simulated inflated abdomen, plastic internal organs, a LESS multi-port trocar with a couple of laparoscopic surgical instrument. Three wireless cameras are anchored on the abdomen with a LED light source. The wireless video signals are collected with three receivers and a 4-channel USB Sensoray 2255S frame grabber. The images are processed with a Core i7-950 Quad-Core 3.06 GHz PC with three Nvidia GeForce GTX 470 video cards and then projected on the abdomen with a BenQ MX761 XGA 3D DLP Projector with a Point Grey Flea video camera as distortion feedback.

Close-up views of the prototype are shown in Fig. 3. Fig. 3(a) shows an abdominal model with insufflation and access to the internal organ models. Fig. 3(b) shows the projection result on the inflated abdomen of our surgical simulation setup with distortion compensation for the merged video. The image was taken with a natural indoor lighting condition from surgeon's point of view.

The images are accomplished when videos from three wireless cameras looking at the region of interest from different viewing points are stitched together with partial overlapping areas to create a seamless panoramic video with high resolution [32]. This inital prototype uses scale invariant feature transform (SIFT) and random sample consensus (RANSAC) based matching techniques to automatically compute optimal global alignment for the mosaicking of videos from different cameras with Levenberg-Marquardt nonlinear minimization algorithm. As shown in Fig. 4, the video from the wireless camera focusing on the area of surgery and the videos from the cameras monitoring the surgery instrument pathway and surround organs are

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

4   IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS
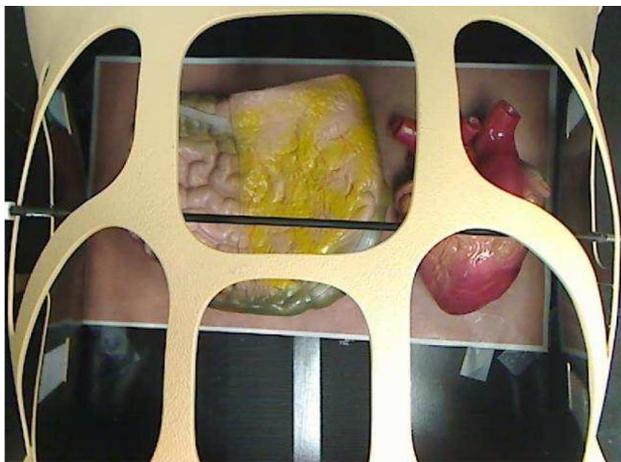
(A)
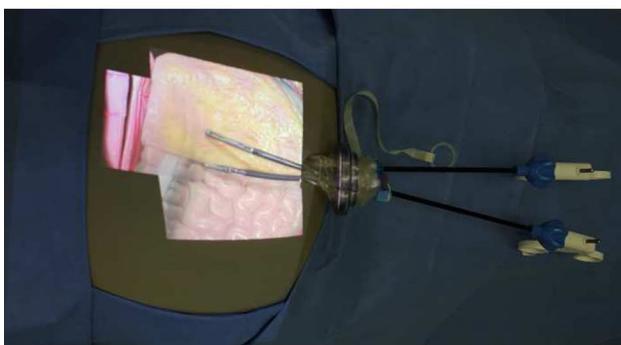
(B)

Fig. 3. (a) A transparent abdominal model with easy access to model organs. (b) An opaque abdominal model where the surgical area video is projected on the abdomen via cameras right above the surgical region but inside the abdominal model.

(A)   (B)   (C)

(D)

Fig. 4. (a) Image from the camera focusing on the area of interest. (b), (c) Images from the cameras monitoring the pathway of the surgery tools and their surrounding organs. (d) The merged view of the whole surgical related regions from all three wireless cameras.

merged together to generate a panoramic view of the surgical related regions. The computed mapping relationship for the mosaicking remains the same during a surgery if the abdomen is stationary. Otherwise, the mapping has to be recomputed when there is any motion for the abdomen (e.g., induced breathing). However, the computation can use the computed mapping result as initial point for more efficient optimization and surgical motion is relatively slow.

Before the panoramic video is fed into the projector, it is processed to prevent color or geometrical distortion for the convex abdomen surface with the feedback from the Point Grey Flea camera. The camera provides a visual feedback for projection distortion compensation and orientation alignment. For distortion calibration, the computer sends a checkerboard image to the project and the camera captures the projected checkerboard image as part of surgery preparation. The locations of the checker corners in both images are automatically detected, and then a mapping between the source image and the projected image is built for future use of distortion compensation [33].

Our work has demonstrated a new method of perceiving the surgical area, which opens an easier and safer scar-less minimally invasive surgery. The presented framework will ultimately allow MIS surgeons to benefit from the "feel" of open surgeries while maintaining the safety of LESS. The system is currently being evaluated by two MIS surgeons on our research team for surgical training.

B. VTEI Open Problems

As demonstrated by the lab prototype VTEI system, there are many rich aspects to be explored before VTEI can become a reality in clinical trials. The current prototype utilizes standard-definition (SD) video sensors to extract image data from the peritoneal cavity. Many surgeons have grown accustomed to HD-quality wired endoscopes and feel it is imperative that next-generated biomedical systems in their fields continue this trend. As mentioned previously, wireless HD video in its current commercial form is not appropriate to this system due its large physical footprint and associated delay. Section III details a new approach to wireless video that may prove capable of achieving the HD-quality imaging required by surgeons while also maintaining the necessary footprint to fit through the tiny surgical incisions.

The problem of using multiple cameras for image mosaicking is exacerbated by the odd shape of the insufflated abdomen; thus, new hybrid stitching/morphing algorithms are needed to accurately reproduce the correct display. This research focus is imperative so that surgeons are not gaining increased viewing area but then losing perception via distortion. Section IV presents a homographic image mosaicking technique that also morphs the combined image to "settle" correctly on the concave surface of the abdomen. This algorithm is optimized to allow real-time updates which are necessary for both periodic movements such as by the induced breathing apparatus and also aperiodic movements by the surgeon.

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ANDERSON *et al.*: VTEI: ON NEW APPROACHES TO *In Vivo* WIRELESS HIGH-DEFINITION VIDEO AND IMAGE PROCESSING 5

III. FREQUENCY-MODULATED VOLTAGE-DIVISION MULTIPLEXING (FM-VDM)

The nature of the wireless scenario described for using VTEI in MIS procedures presents an optimization problem that is non-traditional for wireless communications. Obviously the image sensors must be small enough to fit through the insertion point but other metrics such as power and bandwidth are not necessarily optimized as expected. The following metrics need to be addressed for the wireless HD image sensor:

- **Size**—The sensor spatial footprint is of paramount importance. Only those devices that can fit through the tiny umbilicus incision can be used in LESS operations.
- **Power**—Ironically, a low-power scheme is less important for these first-generation devices. A tiny micro-needle is used to tether the sensor inside the abdomen and has double-use to power the device. Next-generation devices will be completely wireless but still need less power consideration than other battery-powered devices since sensors can be swapped out mid-surgery.
- **Bandwidth**—Ultimately, multiple camera sensors will work in tandem to give surgeons better imaging capabilities; thus, a single sensor cannot occupy the entire available frequency band.
- **Heterogeneous**—Where transmit sensors need to be small and replaceable the receiver architecture is not constrained by the same metrics listed here.
- **Latency**—It is our experience from working with our own surgeons, that laparoscopic surgeons in general have zero tolerance of video latency. Any lag in the video must be outside the realm of human perception.
- **Resolution**—Typical wired endoscopes provide a bright, clean high-definition image to the surgeons. This same resolution is expected in next-generation wireless endoscopes; thus, wireless HD image sensors are being pursued.

Though each of the above metrics is important, size and latency are critical in that they can quickly preclude different hardware approaches to a VTEI system. Current digital wireless HD chipsets are massive when considering the small incisions surgical tools must pass through. Also, typical in some wireless digital systems, bit encoding/compression, bit-error-correction and packet retransmission may lead to a latency unacceptable to our surgeons. To address these critical metrics sufficiently, analog-based FM-VDM is proposed in the following section.

One of the main emphases behind FM-VDM is to keep parallel data streams as analog signals rather than their digital counterparts in order to bypass hardware overhead and potential signal latency described above. Such an approach is especially well-suited for video links where image sensors can naturally output analog data. For example, imagine a wireless link where the luminance (Y) and chrominance (UV) signals, or other analog representations of HD video, are all compounded into a single bandwidth and transmitted simultaneously. For practical systems, multiple cameras could then be multiplexed at different carriers allowing for stereoscopic or panoramic surgical scenes. If possible, such an approach would go far in achieving the metrics listed above since wireless analog systems are often smaller and simpler than their digital counterparts. Indeed, a
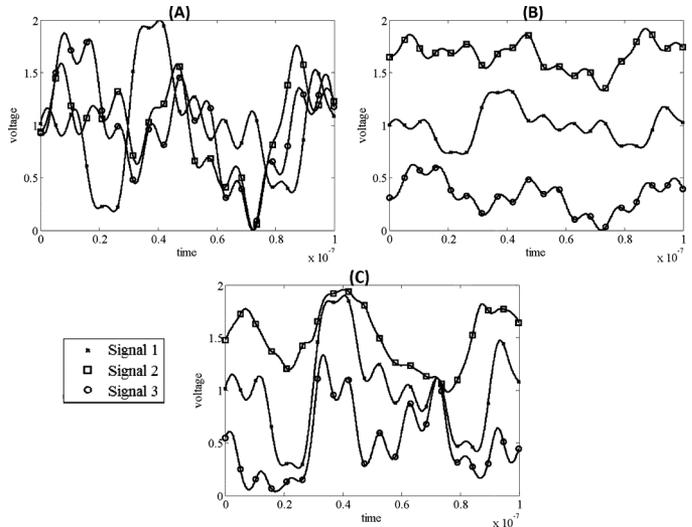

Fig. 5. Shown are various possible voltage multiplexing schemes. (a) No voltage guardbands where all signals occupy all voltage space. (b) Maximum guardbands where no signals occupy other voltage space. (c) Encoding to allow some overlap of voltage space but no overlap of voltage signals.

wireless analog transmitter can be as simple/small as a few surface-mount passive and active components.

To facilitate a discussion of FM-VDM we will consider a scenario with only three analog signal streams; however, it should be noted that addition of more signals is certainly possible and an easy extension of this discussion. To accomplish "voltage multiplexing" in the time domain consider a possible technique where the encoding scheme is

$$
\begin{aligned}
\tilde{m}_1(t) &= \tilde{m}_2(t) + v_1 m_1(t) \frac{v_{\max} - \tilde{m}_2(t)}{v_{\max}} \\
\tilde{m}_2(t) &= v_2 m_2(t) \\
\tilde{m}_3(t) &= \tilde{m}_2(t) - v_3 m_3(t) \frac{\tilde{m}_2(t)}{v_{\max}}
\end{aligned}
\tag{1}
$$

where $m_i(t)$ are the desired message signals, v_i are scalars analogous to modulation indices in AM/FM radios, $v_{\max}$ is the maximum voltage of signal 2, and $\tilde{m}_i(t)$ are the signals to be modulated. Note that with this approach it is possible to select values of v_i such that message voltages are still kept distinct-no signal overlap in voltage—but signals will take advantage of open voltage space when it becomes available; thus, a greater utilization of the "voltage spectrum" is achieved.

Fig. 5 shows VDM signals using the above approach for various voltage guardbands; since VDM signals are always kept on parallel lines (not added together) Fig. 5 shows the signals in the same voltage space in the time domain. Fig. 5(a) contains no voltage guardbands and all signals occupy the entire voltage space. Fig. 5(b) has maximal guardbands meaning no signal is allowed to occupy another signal's voltage space at any time. Fig. 5(c) is selected from (1) such that no signal overlap occurs, but when available voltage space presents itself, the signal is allowed to occupy that space until another signal comes along to push it out. With this approach to FM-VDM a more efficient use of voltage, and, as will be shown, the frequency spectrum,

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

6   IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS

becomes possible. Indeed, we will see that the voltage guard-bands can be chosen to produce the same total bandwidth but individual signal bandwidths will be vastly different.

A problem with VDM is that, though the signals in Fig. 5 appear to be orthogonal, such orthogonality does not exist in the voltage space. For example, unlike frequency—or time-division multiplexing, if VDM signals are added together they lose all orthogonality. To orthogonalize the VDM signals, each signal is individually modulated using a simple FM transmitter. In other words, after VDM encoding, each signal is upconverted to the carrier frequency via FM

$$x_i(t) = A_i \cos\left(2\pi f_c t + 2\pi k_f \int_0^t \tilde{m}_i(\tau)d\tau\right) \quad (2)$$

where A_i is the carrier amplitude, f_c is the carrier frequency and k_f is the conversion "gain" from instantaneous voltage to instantaneous frequency (IF) (not to be confused with interme-diate frequency).

In order to separate each FM-VDM signal we first look at the composite signal arriving at the receiver. Keeping in mind that FM-VDM is not limited to three signals the received signal will look like

$$y(t) = \sum_{i=1}^N A_i \cos\left(2\pi f_c t + 2\pi k_f \int_0^t \tilde{m}_i(\tau)d\tau\right) + n(t) \quad (3)$$

where $n(t)$ is additive-white Gaussian noise (AWGN) and for this discussion on FM-VDM we choose $N = 3$. To detect FM-VDM signals a receiver architecture is needed that can simultaneously track N multiple signals as different instanta-neous frequencies (or phases) traverse the frequency band in the presence of noise.

A new type of circuit is required to successfully demod-ulate the type of modulated signals proposed here—a fre-quency-modulated voltage-division multiplexing phase-locked loop (FM-VDM-PLL). Fig. 6 shows the proposed circuit for just two input signals; it is straightforward to extend to more signals. Three types of states are of interest at the output of the FM-VDM-PLL: lock, ambiguity, or catastrophic events. The locked state means the FM-VDM-PLL is successfully tracking the phase errors, and thereby the message signals, of all encoded signals. An ambiguity event occurs when the FM-VDM-PLL is in a locked state but the tracked signals may be ambiguous with each other; in other words, branch 1 of the FM-VDM-PLL may not be tracking signal 1. The final catastrophic event occurs when the FM-VDM signal falls out of lock and no tracking is possible of any of the signals. Each of these events is illustrated in Fig. 7. Fig. 7(a) shows the original VDM signals prior to transmission while Fig. 7(b) shows the tracked signals on the output of the FM-VDM-PLL in a successful locked state. Fig. 7(c) shows an ambiguity event when the voltage guardband was too small, or noise was too large, and lock is maintained but the signal outputs have been flipped. This event is not necessarily catastrophic if the receiver can remove the ambiguity using characteristics of the signal. In Fig. 7(d) the
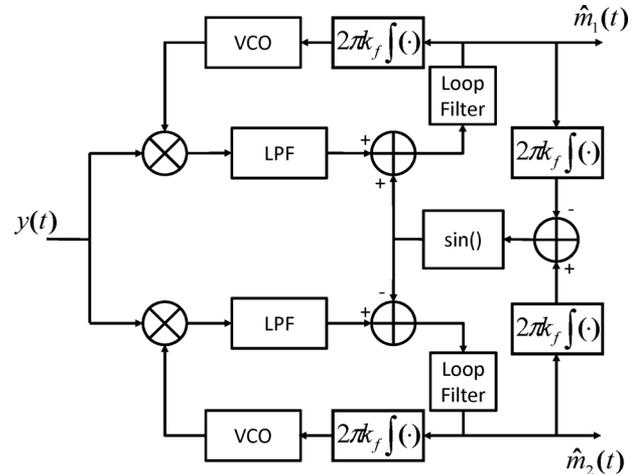
Fig. 6.   A system block diagram of the FM-VDM-PLL to track all signal phases. Individual tracking PLL blocks are interconnected to remove interfer-ence caused by adjacent instantaneous frequencies.
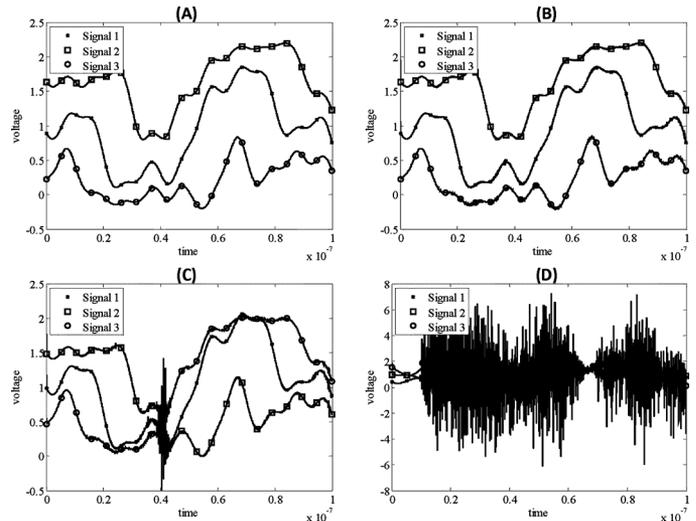
Fig. 7.   Events of interest for the FM-VDM-PLL. (a) Original voltage multi-plexed signals prior to modulation and transmission. (b) Tracked and locked signals at the output of the FM-VDM-PLL. (c) An ambiguous event where lock is reestablished but uncertainty in signal numeration. (d) A catastrophic event where the signals are not tracked correctly.

catastrophic event is shown where the FM-VDM-PLL loses lock completely. Such events can be mitigated by inserting training frames into the signals. For example, consider an HD video stream transmitting at 30 frames-per-second (30 fps). By allowing one (or more) frame(s) to be known at the receiver locked states can be more greatly increased by "training" the FM-VDM-PLL.

To get an intuitive feel for FM-VDM consider the FM-VDM-PLL output shown in Fig. 8. For this simula-tion three randomly generated baseband signals with equal bandwidth were generated using the procedure shown in Fig. 5(c). These signals were taken over a large time window to show how FM-VDM attempts to more fully utilize the voltage space. Power is equally divided between the three carriers and transmitted via FM. The received signal is passed through the FM-VDM-PLL and each message estimate is shown "stacked"

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ANDERSON *et al.*: VTEI: ON NEW APPROACHES TO *In Vivo* WIRELESS HIGH-DEFINITION VIDEO AND IMAGE PROCESSING 7
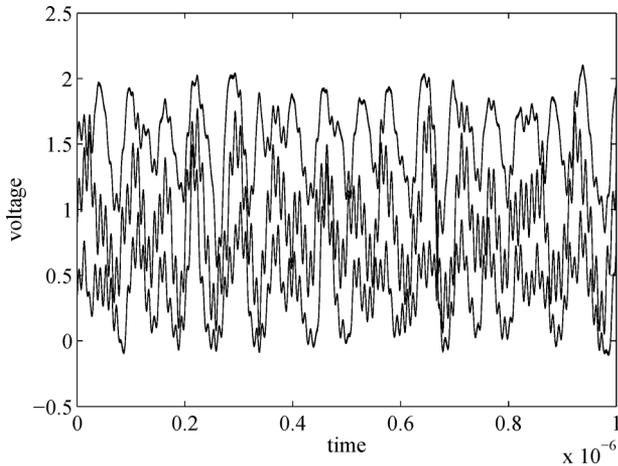

Fig. 8. FM-VDM-PLL outputs for a 10 dB carrier SNR and FM-VDM encoding such that voltages do not overlap but voltage spaces do. The plot was taken over a large time window to demonstrate the impression of voltage space reuse with FM-VDM. No catastrophic or ambiguous events occurred; in other words, the top, middle, and bottom signals correspond to unique input messages.

on the same timeline in Fig. 8. It should also be noted that the input FM-VDM signal occupies the same bandwidth as the largest bandwidth signal when frequency modulated; additionally, the latency of the FM-VDM-PLL output is negligible since, though there is some small transient behavior, the PLL operates on analog signals in real-time.

IV. HOMOGRAPHIC IMAGE MOSAICKING AND MORPHING (HIMM)[1]

The previous section proposed a possible method to extract images from *in vivo* sensors, in this section, the processing done on these images to make them coherent for surgeons is discussed. The method we use to approximate the traditional image mosaicking is introduced and the technical detail of the mosaicking is also described. The wirelessly transmitted videos from three cameras looking at the region of interest from different but close viewing points are stitched together with partial overlapping areas to create a seamless panoramic video with high resolution.

Image mosaicking is a well-studied research topic in computer vision [31] and has been widely applied in medical image analysis [34]–[37]. For example, a real time mosaicking method is used in a tubular environment [34] and applied to general endoscopic procedures. However, to achieve the best mosaicking results, either the different cameras must share a common optical center, or the scene must be planar [38] or tubular as in [34]. If none of the requirements is satisfied, the mosaicking result is not guaranteed and may produce unsatisfactory distortion. Our setup attempts to approximate these two scenarios at the same time to allow for satisfactory mosaicking. First, we place the wireless cameras close to each other, which makes their optical centers close and will be shown below as a good assumption.

[1]Due to the nature of the supporting grant, results for only the surgical simulator will be shown and focus is on the signal and image processing. Live animal model results are used in other work.

Second, the cameras are anchored on the simulator's abdominal wall to make the depth change of the scene as small as possible compared to the distance of the camera to the scene.

Besides the approximation mentioned above, we designed a new image mosaicking method specifically for our VTEI setup. Compared with traditional image mosaicking technique, our method has two major modifications. Firstly, to overcome the difficulties of feature point matching between low-texture images, recently proposed vector field concensus (VFC) [39] is used to reject mismatches from scale invariant feature transform (SIFT) [40] features. Traditionially, mismatch rejection is solved using random sample consensus (RANSAC) [41]. The usage of VFC makes the mosaicking results more robust against image texture differences. Secondly, the depth change of the scene is relatively small compared with the distance of the camera on the abdominal wall to the scene. We exploit this property and propose an adaptive scheme to iteratively update the homography parameter when the scene changes or the cameras are moving. Our experimental results show that these modifications allow image mosaicking to work well while allowing new environments to be explored.

To begin the HIMM treatment, assume at any given time, there is a set of images transmitted to the computer from the cameras. The images will be mosaicked together piecewisely. The procedure of HIMM is summarized as follow, more details about the traditional mosaicking method can be found in [31]. First, a set of SIFT feature points are detected in two adjacent images using the method presented in [40]. The points in the first image are denoted as x_i, y_i, and corresponding points in the second image as x'_i, y'_i. Because we have assumed the cameras share the same optical center, the points between the two images satisfy the homography relationship

$$\begin{bmatrix} x' \\ y' \\ w' \end{bmatrix} \sim \begin{bmatrix} h_0 & h_1 & h_2 \\ h_3 & h_4 & h_5 \\ h_6 & h_7 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \qquad (4)$$

where $\sim$ indicates equality up to some scale factor. The above equation can be rewritten in the following form to better visualize the relationship between correspondence points:

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} \frac{h_0 x + h_1 y + h_2}{h_6 x + h_7 y + 1} \\ \frac{h_3 x + h_4 y + h_5}{h_6 x + h_7 y + 1} \end{bmatrix}. \qquad (5)$$

In order to estimate the homography matrix we must minimize the error, E, defined as the sum of the squared residuals. This can be written as

$$\min E = \min \sum_{i=1}^{n} (x''_i - x'_i)^2 + (y''_i - y'_i)^2 \qquad (6)$$

where x''_i, y''_i represents the predicted corresponding point by substituting (x, y) into (4). Since the relationship between (x''_i, y''_i) and (x_i, y_i) is nonlinear, the Gauss-Newton optimization method is used to iteratively estimate the homography parameters. The initial estimate of the parameters for the

nonlinear optimization is obtained by solving the following linearized equation:

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -x''x & -x''y \\ 0 & 0 & 0 & x & y & 1 & -y''x & -y''y \end{bmatrix} \begin{bmatrix} h_0 \\ h_1 \\ \ldots \\ h_7 \end{bmatrix}.$$

(7)

Since there are usually more than eight corresponding points, the problem is a typical over constrained system, whose solution can be found by least squares. One of the major difficulties of the application of image mosaicking on low-texture images from our setup is the large number of outliers of feature correspondences. Traditionally, RANSAC [41] is used to solve this problem which we have replaced with the more effective VFC method.

The major strength of our mosaicking technique is that the homography mapping is adaptively calculated when a new environment is explored. We assume that the homography matrices of successive frames are similar and the homography matrix from the previous frame is used as the initial value for the current frame. For each frame, rather than directly matching SIFT features, we observe that the initial homography can be used to guide and improve the feature matching. For each point in the left image, the homography matrix maps to its corresponding point in the right image. Due to the approximation nature, the mapped point does not overlay exactly over its true corerponding point. However the mapped point and the true corresponding point should be very close and hence we only search feature points around the mapped point within a predefined radius in the right image. This greatly speeds up feature matching process and improves the feature matching accuracy.

After the homography matrices are computed, we will be able to warp one image to another in a piecewise fashion. First, a reference frame is selected and all the others are warped back into the reference image's coordinate. There are two ways of warping: forward and backward. Since forward mapping might create black seams for those pixels that are not mapped, the backward mapping is used in this paper, which is incorporated with bilinear interpolation. The overlap area of the final composite image might look blurry due to exposure differences and inaccurate registration. To solve this problem, we have found that simple feathering in the overlapped area is fast and effective enough. During feathering process, each pixel's distance to the edge of its original image is calculated and denoted as d, d', respectively. The weight is proportional to the distance as

$$w = \frac{d}{d + d'}$$
$$w' = \frac{d'}{d + d'}.$$

(8)

Finally, once mosaicking has been accomplished distortion is still introduced by the very nature of VTEI. Planar 2D images are projected directly onto the insufflated abdomen to give surgeons an open-cavity view of the surgical area. This projection

(a)   (b)
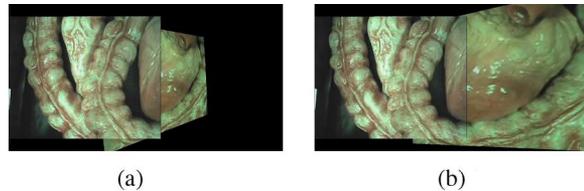
Fig. 9. Incorrect feature point matches cause discontinuities in the mosaicked image. Image blending is intentionally absent on these images to clearly isolate the effect of image warping. In other words, without blending, the black vertical line in the middle of the mosaicked image clearly separates the first and the second image. (a) Without VFC. (b) With VFC.

(a)   (b)
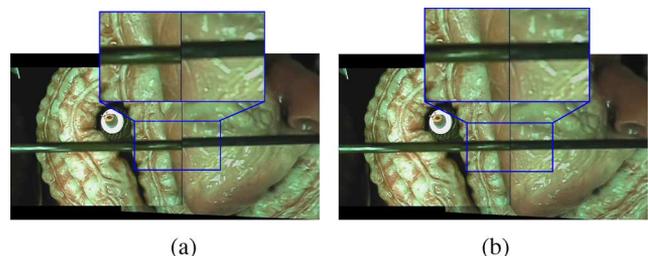
Fig. 10. Without an adaptive update of the homography mapping, misalignment of the surgical tool is possible. An expanded view of the instrument is also shown. (a) Static mosaicking. (b) Adaptive mosaicking.

over 2D images onto a quasi-sphere introduces a depth/distortion effect that is disorienting for the surgeon. To mitigate this effect, the mosaicking is combined with a morphing algorithm to compensate for the curvature of the projection surface. This combination of mosaicking and morphing into a single algorithm that can be run in real-time is referred to as HIMM.

With these improvements, our HIMM method shows more significant performance than the traditional mosaicking method. The usage of VFC effectively removes mismatches and ensures the accuracy of the calculated mapping. Without accurate filtering from VFC, the mismatches would be fed to the least square optimization. Depending on the distance of error point and the correct point, the distortion in the final image can be very severe as shown in Fig. 9(a). The image mosaicking results with VFC method removes the influence from mismatches as shown in Fig. 9(b).

For traditional static mosaicking technique, when the scene is changed or the camera moves, the computed mapping might not be usable for the new scene. Compared with static mosaicking technique, our adaptive mosaicking method shows more accurate results in a dynamic scene and the errors are within three pixels and unnoticeable. For example, Fig. 10(a) shows result from static mosaicking algorithm. With a pre-calculated mapping, when the scene changes, such as when a surgical tool passes through the field-of-view, the distortion becomes obvious. The adaptive method is able to update the homography mapping and reduce the distortion, such as seen on the tool in Fig. 10(a). The adaptive scheme also allows the motion of the cameras in our setup—motion is induced by slowly moving the simulator "abdomen" to simulate breathing.

Finally, the composite HIMM algorithm is shown via screenshots taken of the region of interest as the imaged area moves relative to the position of the cameras as seen in Fig. 11(a)–(d).

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

ANDERSON *et al.*: VTEI: ON NEW APPROACHES TO *In Vivo* WIRELESS HIGH-DEFINITION VIDEO AND IMAGE PROCESSING
9
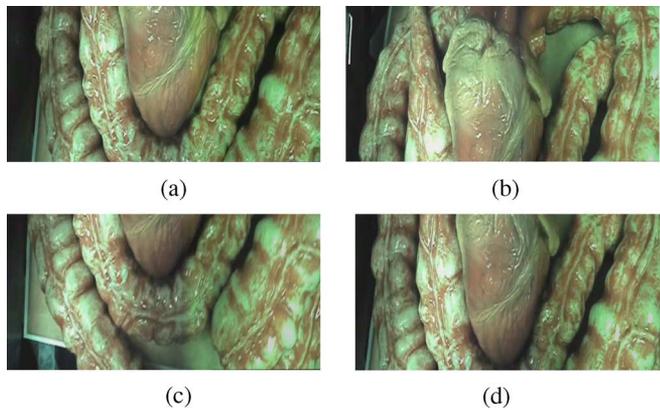
Fig. 11. Screen captures of real-time video as produced by the HIMM algorithm from three different cameras. Still frames are captured in (a)–(d) from the mosaicked image that show different quadrants of the internal surgical area. (a) Upper-left quadrant. (b) Upper-right quadrant. (c) Lower-left quadrant. (d) Lower-right quadrant.
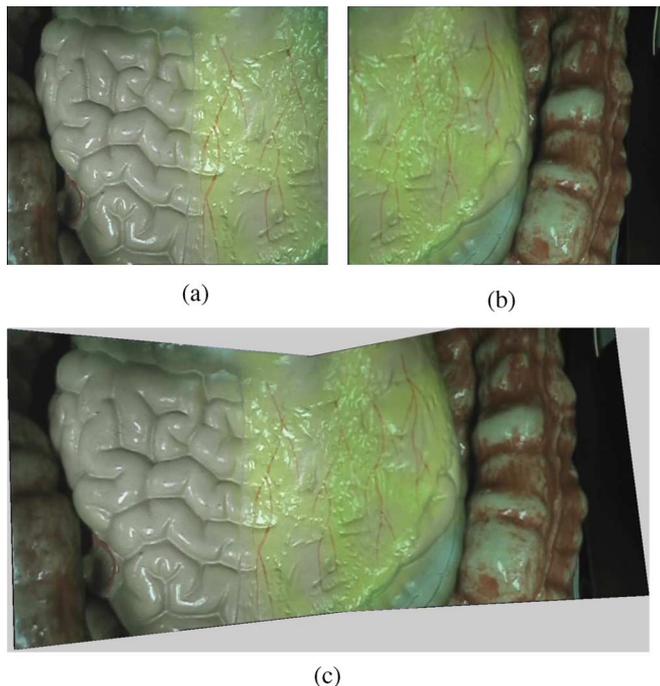
Fig. 12. Resulting surgical area after images from (a) left and (b) right cameras are processed with the HIMM algorithm. Pixel correspondence is nearly perfect with imperceptible differences in (c). (a) Image captured by the left camera. (b) Image captured by the right camera. (c) Lower-left.

These images show that not only can the HIMM algorithm maintain a correspondence between the various inputs but the panaromic view can also be virtually panned as dictated by the needs of the surgeon. To demonstrate the proposed HIMM algorithm and the improved performance over our initial prototyping (Fig. 4) images are taken from two different cameras as shown in Fig. 12(a) and (b). The final blending result is shown in Fig. 12(c) where the original left and right images are seamlessly mosaicked and morphed into a unified display of the operating region.

V. CONCLUSION

Next-generation biomedical devices and systems for advanced minimally invasive surgeries will require wireless links to aggregate data from *in vivo* image sensors. This work has proposed a transmission scheme for *in vivo* HD wireless that allows the sensor footprint to remain small by utilizing traditional analog communication techniques, occupy a nominal bandwidth by a novel multiplexing approach, and also provide low-latency for delay-sensitive operations. This frequency-modulated voltage-division multiplexing (FM-VDM) keeps video data in its analog format and multiplexes multiple streams in voltage before RF transmission via FM. A new circuit architecture, FM-VDM-PLL receiver, though suboptimal, is able to track each signal and provide a greater bandwidth to each than is possible via FDM. However, reception of the video images is only part of the problem. Image processing techniques must occur in real-time and provide a surgical view that lowers distortion and increases perception to the surgeon. This work has adapted homographic image mosaicking techniques and morphing algorithms to display the *in vivo* surgical area directly on the patient model. This overall VTEI system will give practitioners a more open-cavity experience without actually exacting the associated trauma on the patient.

REFERENCES

[1] W. Artibani, S. Fracalanza, S. Cavalleri, M. Iafrate, M. Aragona, G. Novara, M. Gardiman, and V. Ficarra, "Learning curve and preliminary experience with da Vinci-assisted laparoscopic radical prostatectomy," *Urol. Int.*, vol. 80, no. 3, pp. 237–244, 2008.

[2] S. Kaul, R. Laungani, R. Sarle, H. Stricker, J. Peabody, R. Littleton, and M. Menon, "Da Vinci-assisted robotic partial nephrectomy: Technique and results at a mean of 15 months of follow-up," *Eur. Urol.*, vol. 51, no. 1, pp. 186–192, 2007.

[3] T. Hu, P. Allen, N. Hogle, and D. Fowler, "Insertable surgical imaging device with pan, tilt, zoom, and lighting," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2008.

[4] M. Rentschler, J. Dumpert, S. Platt, S. Farritor, and D. Oleynikov, "Natural orifice surgery with an endoluminal mobile robot," *J. Surg. Endosc.*, vol. 21, no. 7, pp. 1212–1215, 2007.

[5] X. Chen, X. Zhang, L. Zhang, X. Li, N. Qi, H. Jiang, and Z. Wang, "A wireless capsule endoscope system with low-power controlling and processing ASIC," *IEEE Trans. Biomed. Circuits Syst.*, vol. 3, no. 1, pp. 11–22, Feb. 2009.

[6] L.-R. Dung and Y.-Y. Wu, "A wireless narrowband imaging chip for capsule endoscope," *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 6, pp. 462–468, Dec. 2010.

[7] Y. Sun, A. Anderson, C. Castro, B. Lin, and R. Gitlin, "Virtually transparent epidermal imagery for laparo-endoscopic single-site surgery," in *Proc. Int. Conf. IEEE Engineering in Medicine and Biology Soc.*, Aug. 2011.

[8] C. Castro, S. Smith, A. Alqassis, T. Ketterl, Y. Sun, S. Ross, A. Rosemurgy, P. Savage, and R. Gitlin, "MARVEL: A wireless miniature anchored robotic videoscope for expedited laparoscopy," in *Proc. IEEE Int. Conf. Robotics and Automation*, 2012.

[9] J. G. Proakis, *Digital Communications*. New York, NY, USA: McGraw-Hill, 2001.

[10] M. Ghovanloo and K. Najafi, "A wideband frequency-shift keying wireless link for inductively powered biomedical implants," *IEEE Trans. Circuits Syst.. Reg. Papers*, vol. 51, no. 12, 2004.

[11] C.-C. Wu, C.-C. Li, and F.-J. Kao, "The high frequency magnetic induction driven miniature power module," in *Proc. OptoeElectronics and Communications Conf.*, Jul. 2011.

[12] S. Mandal and R. Sarpeshkar, "Power-efficient impedance-modulation wireless data links for biomedical implants," *IEEE Trans. Biomed. Circuits Syst.*, vol. 2, no. 4, pp. 301–315, Dec. 2008.

[13] P.-J. Shih and W.-P. Shih, "Design, fabrication, and application of bioimplantable acoustic power transmission," *J. Microelectromech. Syst.*, vol. 19, no. 3, pp. 494–502, 2010.

[14] L. Galluccio, T. Melodia, S. Palazzo, and G. Santagati, "Challenges and implications of using ultrasonic communications in intra-body area networks," in *Proc. 9th Annu. Conf. Wireless On-Demand Network Systems and Services*, Jan. 2012.

[15] P. Si, A. Hu, S. Malpas, and D. Budgett, "A frequency control method for regulating wireless power to implantable devices," *IEEE Trans. Biomed. Circuits Syst.*, vol. 2, no. 1, pp. 22–29, 2008.

[16] R. Bashirullah, "Wireless implants," *IEEE Microw. Mag.*, vol. 11, no. 7, pp. S14–S23, 2010.

[17] Q. Zhang, P. Feng, Z. Geng, X. Yan, and N. Wu, "A 2.4-GHz energy-efficient transmitter for wireless medical applications," *IEEE Trans. Biomed. Circuits Syst.*, vol. 5, no. 1, pp. 39–47, Feb. 2011.

[18] G. Lawton, "Wireless HD video heats up," *Computer*, vol. 41, no. 12, pp. 18–20, Dec. 2008.

[19] J. Gilbert, C. Doan, S. Emami, and C. Shung, "A 4-Gbps uncompressed wireless HD A/V transceiver chipset," *IEEE Micro*, vol. 28, no. 2, pp. 56–64, Mar.–Apr. 2008.

[20] N. Geri, "Wireless HDTV-compressed or uncompressed? That is the question," *AMIMON Ltd.*, Nov. 2006.

[21] S. Shimu, K. Wahid, and A. Dinh, "Wireless transmission of HD video using H.264 compression over UWB channel," *J. Real-Time Image Process.*, 2011, DOI: 10.1007/s11554-011-0191-z.

[22] N. Flaherty, "Hi-def goes wireless—[Comms wireless HD]," *Eng. Technol.*, vol. 3, no. 15, pp. 74–75, 6, 2008.

[23] H. Fuchs, M. A. Livingston, R. Raskar, D. Colucci, K. Keller, A. State, J. R. Crawford, P. Rademacher, S. H. Drake, and A. A. Meyer, "Augmented reality visualization for laparoscopic surgery," in *Proc. Medical Image Computing and Computer Assisted Intervention Conf.*, 1998, pp. 11–13.

[24] H. Hoppe, G. Eggers, T. Heurich, J. Raczkowsky, R. Marmuller, H. Worn, S. Hassfeld, and L. Moctezuma, "Projector-based visualization for intraoperative navigation: First clinical results," in *Proc. 17th Int. Congress and Exhibition on Computer Assisted Radiology and Surgery*, 2003, p. 771.

[25] S. Nicolar, X. Pennec, L. Soler, and N. Ayache, "A complete augmented reality guidance system for liver punctures: First clinical evaluation," in *Proc. Medical Image Computing and Computer Assisted Intervention Conf.*, 2005, pp. 539–547.

[26] M. Blackwell, C. Nikou, A. Digioia, and T. Kanade, "An image overlay system for medical data visualization," *Med. Image Anal.*, vol. 4, no. 1, pp. 67–72, 2000.

[27] F. Sauer, F. Wenzel, S. Vogt, Y. Tao, Y. Genc, and A. Bani-Hashemi, "Augmented workspace: Designing an ar testbed," in *Proc. IEEE and ACM Int. Symp. Augmented Reality*, 2000, pp. 47–53.

[28] O. Dandekar and R. Shekhar, "Fpga-accelerated deformable image registration for improved target-delineation during ct-guided interventions," *IEEE Trans. Biomed. Circuits Syst.*, vol. 1, no. 2, pp. 116–127, Jun. 2007.

[29] G. Fischer, A. Deguet, C. Csoma, R. Taylor, C. Fayad, J. Carrino, S. Zinreich, and G. Fichtinger, "MRI image overlay: Application to arthrography needle insertion," *Comput. Aided. Surg.*, vol. 12, no. 1, pp. 2–14, 2007.

[30] J. Marmurek, C. Wedlake, U. Pardasani, R. Eagleson, and T. Peters, "Image-guided laser projection for port placement in minimally invasive surgery," *Stud. Health Tech. Inform*, vol. 119, pp. 367–372, Aug. 2006.

[31] R. Szeliski, "Image alignment and stitching: A tutorial," *Found. Trends. Comput. Graph. Vis.*, vol. 2, no. 1, pp. 1–104, Jan. 2006.

[32] R. Szeliski, "Video mosaics for virtual environments," *IEEE Comput. Graph. Appl.*, vol. 16, no. 2, pp. 22–30, 1996.

[33] R. Raskar, J. Baar, P. Beardsley, T. Willwacher, S. Rao, and C. Forlines, "ilamps: Geometrically aware and self-configuring projectors," *ACM Trans. Graph*, vol. 22, no. 3, pp. 809–818, 2003.

[34] S. Seshamani, W. W. Lau, and G. D. Hager, "Real-time endoscopic mosaicking," in *Proc. Medical Image Computing and Computer Assisted Intervention Conf.*, 2006, pp. 355–363.

[35] W. Konen, M. Naderi, and M. Scholz, "Endoscopic image mosaics for real-time color video sequences," in *Proc. Computer Assisted Radiology and Surgery*, 2007.

[36] P. C. Amd, C. H. Bay, L. V. Gool, and G. Szkely, "Retina mosaicing using local features," in *Proc. Medical Image Computing and Computer Assisted Intervention Conf.*, 2006.

[37] A. Kolar, "Prototype of video endoscopic capsule with 3-d imaging capabilities," *IEEE Trans. Biomed. Circuits Syst.*, vol. 4, no. 4, pp. 239–249, 2010.

[38] R. Szeliski and H.-Y. Shum, "Creating full view panoramic image mosaics and environment maps," in *Proc. 24th Annu. Conf. Computer Graphics and Interactive Techniques*, 1997, ser. SIGGRAPH'97.

[39] J. Zhao, J. Ma, J. Tian, J. Ma, and D. Zhang, "A robust method for vector field learning with application to mismatch removing," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition*, Jun. 2011, pp. 2977–2984.

[40] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vision*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[41] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, Jun. 1981.

Adam L. Anderson (S'00–M'10) received the B.S. and M.S. degrees in electrical engineering from Brigham Young University, Provo, UT, USA, in 2002 and 2004, respectively, and the Ph.D. degree in electrical engineering from the University of California, San Diego, San Diego, CA, USA.

He has worked as a Research Assistant Professor at the University of South Florida, Tampa, FL, USA, and is currently an Assistant Professor at Tennessee Technological University, Cookeville, TN, USA. His current research interests focus on MIMO systems in mobile ad hoc networks, sensors and sensing systems, and biomedical devices.

Bingxiong Lin (S'11) received the B.S. degree in applied mathematics from Jilin University, Changchun, China, and the M.S. degree in computational mathematics from the Dalian University of Technology, Dalian, China, in 2008 and 2010, respectively.

Currently, he is working toward the Ph.D. degree in computer science and engineering at the University of South Florida, Tampa, FL, USA. His research includes medical image processing and surgical robot vision.
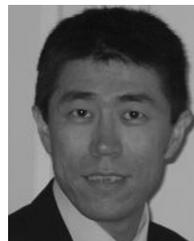
Yu Sun (S'03–M'07–SM'12) received the B.S. and M.S. degrees in electrical engineering from the Dalian University of Technology, Dalian, China, in 1997 and 2000, respectively, and the Ph.D. degree in computer science from the University of Utah, Salt Lake City, UT, USA, in 2007.

Currently, he is an Assistant Professor in the Department of Computer Science and Engineering at the University of South Florida, Tampa, FL, USA. He was a Postdoctoral Associate at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, from December 2007 to May 2008 and a Postdoctoral Associate in the School of Computing at the University of Utah from May 2008 to May 2009. His research interests include medical imaging, robotics, haptics, computer vision, virtual reality, human computer interaction, and medical applications.

Dr. Sun serves as an associate editor of the IEEE ROBOTICS AND AUTOMATION MAGAZINE.