Science Translational Medicine

Title: Hexachromatic bio-inspired camera for image-guided cancer surgery

Authors

Steven Blair[1], Missael Garcia[1], Tyler Davis[1], Zhongmin Zhu[1], Zuodong Liang[1], Christian Konopka[2,3], Kevin Kauffman[4], Risto Colanceski[5], Imran Ferati[5], Borislav Kondov[5], Sinisa Stojanovski[6], Magdalena Bogdanovska Todorovska[7], Natasha Toleska Dimitrovska[5], Nexhat Jakupi[5], Daniela Miladinova[6], Gordana Petrusevska[7], Goran Kondov[5], Wawrzyniec Lawrence Dobrucki[2,3,10], Shuming Nie[1,2,8,9], and Viktor Gruev[1,3,10†]

Affiliations

[1] Department of Electrical and Computer Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[2] Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[3] Beckman Institute for Advanced Science and Technology, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[4] Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA

[5] University Clinic Hospital, Department of Thoracic and Vascular Surgery, Ss. Cyril and Methodius University of Skopje, 1000 Skopje, Republic of North Macedonia.

[6] University Clinic Hospital, Institute of Pathophysiology and Nuclear Medicine, Ss. Cyril and Methodius University of Skopje, 1000 Skopje, Republic of North Macedonia.

[7] University Clinic Hospital, Department of Pathology, Ss. Cyril and Methodius University of Skopje, 1000 Skopje, Republic of North Macedonia.

[8] Department of Chemistry, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[9] Department of Materials Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

[10] Carle Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL 61820, USA

[†]Corresponding Author Email: vgruev@illinois.edu

Related resources:
10.1126/scitranslmed.aaf3936;
10.1126/scitranslmed.aav2169;
10.1126/scitranslmed.aaf6169;
10.1126/scitranslmed.3010633

Overline: Cancer Imaging

One Sentence Summary

A color/near-infrared camera inspired by the mantis shrimp visual system enables tumor detection and lymph node mapping during cancer surgery.

Editor's summary:

Through a shrimp eye brightly

A camera for intraoperative imaging of tumors could improve surgical outcomes, but some imaging technologies have been difficult to translate to clinical practice. Blair et al. designed an imaging system based on the eye of the mantis shrimp. This system detected multiple near-infrared fluorescent signals simultaneously and was tested in a mouse model of human prostate cancer. In support of clinical feasibility, the authors showed that fluorescently labeled sentinel lymph nodes could be detected by the sensor in patients with breast cancer undergoing surgical resection. This bio-inspired imaging sensor could offer a flexible tool for image-guided surgical removal of tumors.

Abstract

Cancer affects 1 in 3 people worldwide. Surgery remains the primary curative option for localized cancers, but good prognoses require complete removal of primary tumors and timely recognition of metastases. To expand surgical capabilities and enhance patient outcomes, we developed a six-channel color/near-infrared image sensor inspired by the mantis shrimp visual system that enabled near-infrared fluorescence image guidance during surgery. The mantis shrimp's unique anatomy, which maximizes the number of photons contributing to and the amount of information contained in each glimpse of its surroundings, is recapitulated in our single-chip imaging system that integrates arrays of vertically stacked silicon photodetectors and pixelated spectral filters. To provide information about tumor location unavailable from a single instrument, we tuned three color channels to permit an

intuitive perspective of the surgical procedure and three near-infrared channels to permit multifunctional imaging of optical probes highlighting cancerous tissue and other features characterizing surrounding tissue. In nude athymic mice carrying human prostate tumors, our image sensor enabled simultaneous detection of two tumor-targeted fluorophores, distinguishing diseased from healthy tissue in an estimated 92% of all cases, and permitted extraction of near-infrared structured illumination, mapping the tumors' and the surgical site's three-dimensional topography to within 1.2 mm error. In the operating room, during surgical resection in eighteen patients with breast cancer, our image sensor further enabled sentinel lymph node mapping using clinically approved near-infrared fluorophores. The flexibility and performance afforded by this simple and compact architecture highlights the benefits of biologically inspired sensors in image-guided surgery.

Introduction

As humans have pursued rapid advancements in digital cameras, we have also found new biomedical applications for those cameras with surgical guidance provided by intraoperative imagers (*1-4*) and pathological samples identified by lab-on-a-chip devices (*5*). For diseases like cancer, where removal of primary tumors and confirmation of negative margins are critical to treatment, these aids to surgery and pathology offer great promise; however, there is a notable gap in the translation of many imaging technologies to clinical practice.

As an example, near-infrared fluorescence image-guided surgery has shown enormous potential for cancer surgery due to the low autofluorescence and scattering of tissues at near-infrared wavelengths, enabling large signal-to-background ratios and imaging depths while eschewing damaging ionizing radiation (*6-10*). With spectral sensitivity spanning the visible and near-infrared, digital cameras would seem well-suited to color imaging of surgical sites and near-infrared imaging of fluorescently labeled tumors; however, state-of-the-art near-infrared fluorescence imaging systems suffer from complex architectures and narrow feature sets that have hindered adoption of next generation tumor-targeted fluorophores as well as currently available nonspecific fluorophores. The first fluorescence-guided surgery was performed over 70 years ago (*11*), yet these issues have prevented near-infrared fluorescence image-guided surgery from being widely accepted as standard of care. Most physicians thus continue to rely on their unaided senses of sight and touch (*3, 12, 13*) which are, unfortunately, not optimized for the task: incomplete tumor resection occurs in 25% of patients with breast cancer, 35% of patients with colon cancer, and 40% of patients with head and neck cancer *(14)*. It would seem, then, that the solution to near-infrared

fluorescence image-guided surgery lies somewhere between the digital cameras that engineering has manufactured and the eyes that biology has conceived.

The optimization of both manmade digital cameras and naturally evolved eyes has been driven by selection toward maximum fitness. Digital cameras have developed over several decades to satisfy consumers documenting daily life in real time and at high resolution (*15*), while eyes have evolved over millions of years to facilitate the complex decisions that permit survival of the species (*16*). Such divergent evolutionary pathways have created vastly different visual systems. Engineers have achieved frame rates and spatial resolutions that surpass nature (*17*) while offering quantum efficiencies, dynamic ranges, and signal-to-noise ratios that can be tuned over tremendous ranges. Biology has in turn seized the upper hand with visual systems that detect spectral information with extremely efficient and effective architectures (*18, 19*); for example, the mantis shrimp visual system, which fits 16 types of cone cells onto compact eyestalks, is unmatched by any manmade spectral camera in power consumption and information density. The marriage of the optimized optoelectronics built into digital cameras and the diverse functionality observed in eyes has borne imaging systems with sensory capabilities exceeding those previously provided by engineers and biology alike (*10, 20, 21*), suggesting a future for near-infrared fluorescence image-guided surgery.

Results

A bio-inspired image sensor for near-infrared fluorescence image-guided surgery

Two key trends jeopardize potential solutions for near-infrared fluorescence image-guided surgery. The first trend involves architectures: existing instruments rely on two architectures that are poorly matched to the clinical context (table S1). In division-of-time architectures, a filter wheel isolates different spectral channels at different times, permitting all spectral channels to be captured by the same image sensor at distinct time points. This collection of spectral information over time requires reduced exposure times with increasing channel counts, maintaining real-time frame rates but reducing fluorescence signal-to-noise ratios, and introduces co-registration error between spectral channels due to interframe motion. In division-of-optical-path architectures, dichroic beamsplitters or dichroic/trichroic prisms direct different spectral channels along different optical paths, permitting all spectral channels to be captured by distinct image sensors at the same time point. This collection of spectral information across space guarantees increased device sizes with increasing channel

counts, demanding additional relay optics that impact image quality. The second trend involves feature sets: existing instruments typically image a single fluorescent probe (tables S1 and S2), providing insufficient sensitivity and specificity for real physiology. Due to variations in inter- and intratumor biomarkers, a single tumor-targeted probe may fail to detect metastatic tumors or may fail to detect the boundaries of primary tumors. This was confirmed by a recent meta-analysis indicating that up to 40% of metastatic tumors and involved lymph nodes can have biomarkers different from those of the primary tumor (*22*).

To address these issues, we turned to the mantis shrimp visual system as a blueprint for an artificial image sensor. The mantis shrimp benefits from a simple physical phenomenon that imbues spectral sensitivity: shorter-wavelength photons (i.e., blue light) penetrate shorter distances into their compound eyes than longer-wavelength photons (i.e., red light) due to a wavelength-dependent absorption coefficient (*23*). As a result, the photosensitive cells at the top of the mantis shrimp's compound eyes preferentially register shorter-wavelength photons, while the photosensitive cells at the bottom preferentially register longer-wavelength photons (Fig. 1A,B). Crystalline cones above the compound eye provide fine-grained spectral filtering, enabling different parts of the eye to see different bands of light (Fig. 1A,B). Ultimately, this unique combination of vertically stacked photodetectors with spectral filters enables 16 spatiotemporally co-registered spectral bands to be constantly probed by microscopic cells, offering increased optical throughput as multiple spectral observations are made without rejecting photons, increased data throughput as multiple spectral observations are made at every point, and increased spatial resolution as spectral observations are distributed vertically as well as laterally.

To mimic the mantis shrimp's multispectral capabilities, we designed and fabricated an image sensor by monolithically integrating an array of vertically stacked silicon photodetectors with an array of pixelated spectral filters (Fig. 1C,D). The absorption length of silicon strongly varies with wavelength, with 99% of blue photons (~400 nm) absorbed within 0.5 μm of penetration and 99% of red photons (~650 nm) absorbed within 16 μm of penetration. Recognizing the high density of photocharge generation by blue photons in a thin sheet near the surface and by red photons in a thick slab farther below the surface, the location and extent of photodiodes in the silicon can be selected to measure the amounts of blue light and red light in a scene: a blue-sensitive photodiode is shallow and of limited extent, whereas a red-sensitive photodiode is deeper and of greater extent. Green photons experience an intermediate absorption coefficient and can be detected with an intermediate

168 photodiode. The spectrum of incident photons can be further shaped with spectral filters,
169 enabling a wide range of distinct responses.

170 An array of vertically stacked photodiodes was fabricated by epitaxial growth of three
171 positively doped (p+) silicon layers with individual thicknesses of 0.8 μm, 2.8 μm, and
172 4.3 μm, followed by ion implantation to generate negatively doped (n+) regions (Fig. 1C,D).
173 The photodiodes were most responsive at ~430 nm, ~550 nm, and ~700 nm but maintained
174 responsiveness at longer wavelengths. An array of pixelated spectral filters was produced
175 by stacking submicron layers of the dielectric materials silicon dioxide, silicon nitride, and
176 hafnium oxide in different quantities and thicknesses on a glass substrate (Fig. 1C-F). The
177 filters were organized in an alternating grid of short-pass filters, with passbands below 700
178 nm, and long-pass filters, with passbands above 700 nm. The monolithically integrated
179 image sensor was finally produced by flip-chip bonding the arrays of vertically stacked
180 photodiodes and pixelated spectral filters (Fig. 1C,D).

181 **Optoelectronic characterization**

182 The spectral response of the complete image sensor was benefitted by the high in-band
183 transmission (~95%) and high out-of-band rejection (optical density ~4) of the pixelated
184 spectral filters (Fig. 2A). The three photodiode layers under the short-pass filters maintained
185 distinct spectral responses that peaked under blue light, green light, and red light; the three
186 photodiode layers under the long-pass filters exhibited similar spectral responses that
187 peaked at ~700 nm (Fig. 2B), but the differences in the spectral responses across the
188 photodiode layers were substantial enough to enable three different observations of the near-
189 infrared spectrum. Together, the three visible channels under the short-pass filters and the
190 three near-infrared channels under the long-pass filter accounted for a total of six spectral
191 channels—providing hexachromatic vision. As needed, additional notch filters (optical
192 density >6) were placed over the whole sensor to suppress photons from the excitation light
193 source.

194 The stacked photodiodes' spectral responses compared favorably against scientific cameras
195 optimized for color reproduction, with the centroid wavelengths of our imaging system's
196 color channels (i.e., 507 nm, 563 nm, and 607 nm) (Fig. 2B) falling near those of Teledyne
197 QImaging's MicroPublisher 6's color channels (i.e., 464 nm, 540 nm, and 608 nm) (*24*).
198 This correspondence was confirmed in color images (like that in fig. S1) where complex
199 natural scenes with diverse color content appeared natural. The optimization toward broadly

shaped quantum efficiencies in the visible spectrum ensured sufficient shaping in the spectral responses at the shorter end of the near-infrared spectrum, allowing the United States Food and Drug Administration-approved fluorescent dyes methylene blue ($\lambda_{em} \approx 690$ nm) and indocyanine green ($\lambda_{em} \approx 810$ nm) to induce different signals. Nevertheless, the location and extent of the bottom photodiode offered a benefit farther into the near-infrared spectrum where the spectral response did not degrade until beyond 1000 nm (Fig. 2B), providing a meaningful increase for weak fluorescent signals at the expense of a computationally correctable red hue in color images. Further tuning of the photodiode parameters would enable an application-specific trade-off between color accuracy and near-infrared discrimination when the selected parameters prove insufficient, while maintaining the additional benefits to optical throughput, data throughput, and spatial resolution afforded by the stacked photodiodes.

The image sensor exhibited a read noise of 62 electrons and a linear full well capacity of 101,473 electrons at the middle photodiode, indicating a dynamic range of 64 dB or 10.7 bits at this worst-case photodiode; furthermore, the image sensor exhibited a maximum fixed pattern noise of 0.6% at half linear full-scale range (Fig. 2C). Both the read noise and the fixed pattern noise were affected by difference double sampling within the readout circuit: at the price of an effective doubling of the read noise at each pixel, there was a substantial drop in threshold voltage variation across the image sensor. An additional charge-transfer transistor, a pinning layer, and a floating diffusion for each photodiode would permit correlated double sampling, which would simultaneously reduce read noise and correct offsets, as demonstrated in other complementary metal-oxide-semiconductor (CMOS) fabrication processes using embedded photodiodes (*25, 26*). The dark current of the image sensor was 4,107 electrons per second in the worst-case at the top photodiode; however, the requirements for image-guided surgery often included a real-time frame rate of ~25 frames per second (~40 milliseconds per frame), limiting the quantity of dark charge that could be accumulated and the power of dark noise that could be observed after any single exposure. Therefore, common corrections like image sensor cooling would yield a minimal reduction in an already small dark current, producing little improvement in the accumulated dark charge, the observed dark noise, and the resulting signal-to-noise ratio. The quantum efficiencies coupled with this noise performance permitted detection of 201 pM of indocyanine green at 200-ms integration time under 20 mW/cm^2 of 780 nm laser excitation (Fig. 2D).

Intraoperative imaging of cancers requires simultaneous recording of the plentiful visible photons that outline a patient's anatomical features and the sparse fluorescent photons that highlight the locations of tumors. Visible photons from the surgical lighting undergo a single reflection at the surface of the surgical site before entering the imaging system, whereas near-infrared photons from the fluorescence excitation undergo a complex series of scattering, absorption, excitation, and emission events before collection, ensuring that the fluorescent photons experience far more losses. However, the International Electrotechnical Commission has specified a lower bound on the illuminance of surgical lighting and an upper bound on the irradiance of fluorescence excitation that prevents changes to the light sources sufficient to overcome such losses (*27, 28*). These asymmetries in optical paths and lighting specifications yield a large dynamic range between the strong visible lighting and the weak fluorescent signal that exceeds ~90 dB in the best case, beyond the 80-dB capability of state-of-the-art sensors. Even if sufficiently large dynamic ranges could be achieved in conventional sensors, for example, by reducing the readout noise and increasing the full well depth, the requirement that the color image remain unsaturated would practically guarantee that the near-infrared image would occupy the lower half of the dynamic range. Due to the Poisson distribution of the shot noise, the signal-to-noise ratio of the near-infrared image would be minimal, increasing the risk that weakly labeled fluorescent targets may not be detected. As a result, our image sensor included programmable readout circuitry that independently controlled the exposure times for the color pixels and near-infrared pixels. This enabled our image sensor to simultaneously acquire high signal-to-noise-ratio images in both the visible and near-infrared spectra.

Simultaneous imaging of multiple near-infrared fluorescent signals

The most important information about the near-infrared spectrum was encoded in the comparative variation across and the overall intensity of the near-infrared channels, not the absolute value of each channel; thus, a transformation was required to map the measured photodiode response to a more intuitive quantity. In practice, the transformation from red-green-blue (RGB) to hue-saturation-value (HSV) used in color image processing proved adept at separating different signals encoded in the near-infrared channels. Under this transformation, hue and saturation represented ratios of differences between the three channels, providing information about the incident spectrum, while value represented the response of the strongest channel, providing information about the incident intensity. Taken together, hue and saturation enabled differentiation of multiple targets (e.g., two fluorescent

markers) based on the targets' spectral characteristics (e.g., different emission wavelengths), while value permitted a comparison between the targets' intensities. At the same time, value measured the targets' intensities with maximal signal-to-noise ratio and was preferred when the overall intensity of multiple dyes, not the differences between those dyes, was of interest during decision making. Even without the physical meaning of the total energy across spectral bands, these abstract quantities of hue, saturation, and value permitted application-specific descriptions for image-guided surgery while enabling full use of the near-infrared measurements at each pixel. Images could thus be formed from either the raw quantities or informative transformations of them and presented to the surgical staff on a display.

To demonstrate that the image sensor was sensitive to small perturbations of the incident spectrum, we prepared the near-infrared fluorophore indocyanine green in two different solvents, deionized water and fetal bovine serum. When excited with the same light source, the emission peaks from these samples were spectrally separated by 8 nm, but the emission spectra from these samples were clearly distinguished by our image sensor, as evidenced by the separation between the 95% tolerance ellipses for the hues and saturations (Fig. 3A). To illustrate further that the hue-saturation response varied continuously with a smooth change in the incident spectrum, we prepared the near-infrared fluorophores IRDye 680RD and IRDye 800CW in seven volumetric ratios from 10-to-0 to 0-to-10. When excited with two light sources, the single degree of freedom in volumetric ratios dictated that the emitted spectrum from the fluorophore mixture was the volume-weighted average of the emission spectra from the individual fluorophores; indeed, our image sensor returned hues and saturations confined about a one-dimensional manifold with limited overlap between the 95% tolerance ellipses associated with different ratios (Fig. 3B), permitting quantification of the relative concentrations of the two dyes. Both of these results were achieved due to the spectral sensitivity, temporal noise performance, and spatial noise performance of the imaging system.

The image sensor's ability to distinguish emission spectra has notable applications, but its ability to differentiate emission spectra from excitation spectra is also important. Given that the hue and saturation provide information about the incident spectrum, and that the fluorescent excitation spectrum and the fluorescent emission spectrum are always different, the reflection from an excitation light source and the emission from a fluorophore map to different hue-saturation points, permitting discrimination in the hue-saturation space. This has applications in surgery, where the small number of excitation photons that reflect from

the surgical site and pass through the emission filter compete with the potentially small number of emission photons that reach the image sensor. A divergence in response was exhibited in the benchtop experiment shown in Fig. 3C to G, where pixels exposed only to fluorescence excitation or fluorescence emission laid in separate clusters in the hue-saturation space and pixels exposed to a mixture of excitation and emission laid between. This permitted labeling of pixels that were useful (observed mostly emission), useless (observed mostly excitation), and of potential use (observed a mixture of emission and excitation).

Tumor detection in a small-animal model of human prostate cancer

To evaluate the sensitivity to and utility of multiple fluorophores in a surgical context, our imaging system was used to identify fluorescently labeled tumors in a small-animal model of human prostate cancer. Human prostate cancer cells (LNCaP and PC3) were subcutaneously implanted in nude athymic mice, and ~2 weeks later two near-infrared fluorescent optical probes (IRDye 680RD labeled with human recombinant epidermal growth factor [EGF] and IRDye 800CW labeled with 2-deoxy-D-glucose [2-DG]) were intravenously injected into the same mice. After 24 hours, the mice underwent whole body fluorescent imaging with our imaging system before the tumor tissue and healthy tissue were surgically removed and imaged separately. Imaging occurred under 665 nm excitation (targeting IRDye 680RD EGF), 780 nm excitation (targeting IRDye 800CW 2-DG), and mixed 665 nm-780 nm excitation (targeting both IRDye 680RD EGF and IRDye 800CW 2-DG).

Using the near-infrared information from our imaging system, each tissue sample was scored and classified as either tumor tissue or healthy tissue, and these predictions were compared against ground truth. Receiver operating characteristic (ROC) curves (Fig. 4A) indicated an area under the curve (AUC) of 0.92 under mixed 665 nm-780 nm excitation (IRDye 680RD EGF and IRDye 800CW 2-DG together) compared to 0.77 for 665 nm excitation (IRDye 680RD EGF alone) and 0.75 for 780 nm excitation (IRDye 800CW 2-DG alone). Accounting for the paired design of the imaging study, a dual-tracer strategy using EGF and 2-DG together significantly outperformed single-tracer use of EGF ($p = 0.022$) and single-tracer use of 2-DG ($p = 0.016$) when comparing AUCs. Considering the difference between the AUC for the EGF plus 2-DG scenario and the AUCs for the single-tracer scenarios, the 95% confidence interval was (0.03–0.29) for the EGF scenario and

(0.04–0.36) for the 2-DG scenario, indicating a statistically significant improvement in tumor scoring.

Images composited from color and near-infrared observations of a mouse along with histograms computed from near-infrared observations within its tumors are provided in Fig. 4B. Although IRDye 680RD EGF and IRDye 800CW 2-DG were taken up by two tumors, the tracers exhibited substantial heterogeneity between and within those tumors. IRDye 680RD EGF produced a consistently stronger fluorescent signal in the tumor on the left flank than the tumor on the right flank, whereas IRDye 800CW 2-DG behaved in the opposite way, producing a stronger signal in the right tumor than the left tumor. Regardless of the overall brightness, the dyes produced a stronger signal in some regions and a weaker signal in others with the intensity generally dropping from the medial region of the tumors to the lateral region. The application of a dual-tracer strategy would counteract the heterogeneity of these single-tracer strategies. In some cases, the two tracers would complement each other, with one tracer accentuating tissues missed by the other tracer; in other cases, the two tracers would supplement each other, with both tracers emphasizing similar cues that should improve confidence in surgical decision making.

All three visible channels were required for color imaging of the surgical site, whereas two near-infrared channels were required for near-infrared imaging of the fluorescent dyes (IRDye 680RD EGF with $\lambda_{em} \approx 693$ nm and IRDye 800CW 2-DG with $\lambda_{em} \approx 780$ nm), accounting for five of the image sensor's spectral channels. The third near-infrared channel was dedicated to three-dimensional reconstruction using structured illumination from a near-infrared projector (at $\lambda \approx 900$ nm), ensuring full use of all six channels. The projector illuminated the surgical site with a sinusoidal pattern advancing along the wavefront normal, while an algorithm compared the observed phase change to the expected phase change and extracted the phase change induced by the three-dimensional topography; an inverse model then facilitated conversion of these phase changes into a height map. These three-dimensional reconstructions could be generated at a spatial resolution of 1280 by 720 pixels and a frame rate of 30 frames per second while exhibiting an average root-mean-square error of 1.179 mm and an average signal-to-noise ratio of 36.

Figure 4C shows representative images of a tumor-bearing mouse before surgery indicating both the fluorescently labeled tumors and the three-dimensional reconstruction; an animation of these images is provided in video S1. Three visible channels worth of information were used to visualize the mouse, while two near-infrared channels worth of

information were used to detect the tumors (Fig. 4D); the remaining near-infrared channel worth of information was used to extract the shape of the mouse and the tumors within the mouse (Fig. 4E). In Figure 4, the tumors were highlighted in either green or blue to emphasize preferential accumulation of either IRDye 680RD EGF or IRDye 800CW 2-DG as measured by normalized fluorescence under 665 nm excitation and 780 nm excitation, emphasizing that different tumors may be more easily detected with different dyes. In the operating room, however, the tumors would normally be identified together under mixed 665 nm-780 nm illumination, permitting the greater diagnostic performance demonstrated by the ROC analysis. This diagnostic information, coupled with inherently co-registered shape information, may improve tumor detection despite intertumoral and intratumoral variations and optimize surgical strategies according to the tumors' size and extent.

Clinical feasibility of sentinel lymph node detection in patients with breast cancer

To explore translation of a multifluorophore workflow to a clinical environment, our imaging system was used to visualize fluorescently labeled sentinel lymph nodes in patients with early-stage breast cancer. In conventional surgical practice, sentinel lymph node mapping involves the peritumoral administration of radioactive tracers like technetium-99m-labeled human serum albumin colloid (^{99m}Tc-HSA colloid) that can be identified with a gamma probe and dark-colored dyes like indocyanine green and methylene blue that can be observed with the naked eye. Given that indocyanine green and methylene blue are both near-infrared fluorophores, though, it is possible to detect these dyes by their near-infrared fluorescence, enabling identification of sentinel lymph nodes without the radioactivity associated with radiotracers and the limited depth of visualization associated with visible stains. According to the standard of care, ^{99m}Tc-HSA colloid, methylene blue and indocyanine green were administered prior to breast cancer surgery. Based on radioactive signals and visual cues, all suspected lymph nodes were removed from the patient and imaged with our imaging system under 780 nm excitation (for indocyanine green detection), 665 nm excitation (for methylene blue detection), and mixed 780 nm-665 nm excitation (for dual indocyanine green and methylene blue detection). Immediately thereafter, the suspected lymph nodes underwent histopathologic analysis to determine whether the resected tissues were lymphatic structures and to evaluate whether they exhibited metastases.

A photograph of our imaging system in the operating room is provided in Fig. 5A. Representative images of sentinel lymph nodes in vivo during and ex vivo after surgery are

provided in Fig. 5B to F, and a video of a sentinel lymph node resection is provided in video S2. In vivo, indocyanine green and methylene blue collected in different regions of the surgical site at different moments during the operation. From the injection of the dyes through the massage of the tissues to the beginning of resection, the dyes migrated from the injection site to the lymph node as scalpels, retractors, and other tools were introduced throughout the procedure. In such a dynamic environment, where fluorophores were covered and uncovered, modulating the fluorescent emission, and where tools were constantly moving, absorbing and reflecting the fluorescent excitation, it was critical that relevant fluorescent emission be detected and that irrelevant fluorescent excitation be rejected, all while under surgical illumination. Our imaging system, which could detect weak fluorescence emission and differentiate between fluorescence emission and excitation, identified lymph nodes under such demands. Ex vivo, indocyanine green and methylene blue concentrated at different points in different lymph nodes. As the dyes traveled through the lymphatic basin into a lymph node, they exhibited differences in accumulation based on disparities in molecular weight, hydrodynamic radius, and charge; such variation could cause one dye to exhibit a homogeneous response and another dye a highly heterogeneous response even though both dyes were injected at a similar time and place. Again, our imaging system, which could spectrally differentiate the dyes and use that spectral information for application, identified lymph nodes in such a context.

Ultimately, our imaging system succeeded at detecting sentinel lymph nodes, providing an AUC of 88% for ex vivo samples with indocyanine green, as well as a true positive rate of 84% and a false positive rate of 0% at maximal Youden's index (Table 1).

Discussion

When successfully translated to the medical field, new sensing technologies can decrease patients' medical expenses and increase their quality of life. In most hospitals, the primary sensing modalities during surgical procedures are the surgeon's eyesight and touch, leading to subjective discrimination between cancerous and healthy tissues. As a result, there is continued interest in the development of imaging systems that can eliminate the positive tumor margins associated with tumor regrowth and the overly negative tumor margins associated with iatrogenic effects. Unfortunately, the adoption of imaging technologies into the operating room has faced little success. The imaging architectures that have come to

dominate within the last 50 years of semiconductor and optical research have produced biomedical imaging systems that are incompatible with the clinical environment and thus lack diagnostic performance under operative conditions (*8-10*).

Here, we have presented a bio-inspired image sensor with hexachromatic vision for image-guided surgical applications. The sensor, which is constructed from arrays of vertically stacked photodetectors and pixelated spectral filters, enables the differentiation of multiple fluorophores across a surgical site without slowing down or otherwise impeding a surgeon's workflow. Near-infrared images are inherently co-registered with visible images in time and space thanks to this architecture, eliminating the ambiguity between anatomical features and fluorescently labeled structures. Tumor detection in a preclinical mouse model of human prostate cancer indicated that this technology could exploit tumor-targeted markers to enhance resection of multifocal, multicentric, and infiltrative cancers. Sentinel lymph node detection for the clinical management of breast cancer in the operating room demonstrated that this technology could detect such markers against a complex background and elucidate their heterogeneous response. With its compact footprint and low weight, which permit seamless integration into the operating room, our bio-inspired imager offers opportunities for quality-based health care as contrast agents continue to evolve.

Recent progress in scientific CMOS image sensors has enhanced optoelectronic performance, with some sensors offering sub-electron read noise. Continuous refinement to the CMOS fabrication and post-CMOS integration of these image sensors and their color filter arrays has also facilitated pixels approaching submicron pitches. A previously underexplored technology, stacked photodiode CMOS image sensors have not seen the full benefit of the scientific CMOS image sensor's improvements. Nevertheless, circuit-, system-, and process-level similarities between the two sensors indicate that the largest hurdle to equalizing performance is economic, not scientific (*29*). Comparing a conventional spectrometer to our imager, there is additional concern that the similarities in the quantum efficiencies for the three near-infrared channels, as well as the limited magnitudes of those quantum efficiencies at longer wavelengths, may negatively impact sensitivity to and discrimination between near-infrared fluorophores. These characteristics have not inhibited our ability to discriminate between common fluorophores, but additional work will be required to identify our sensor's spectral limits as new fluorophores emerge. In its current state, our stacked photodiode imaging system poses solutions to important clinical problems, and with additional improvements, the ability to detect fluorophores at

lower concentrations, distinguish fluorophores with closer emission spectra, and capture information in two stacked photodiode pixels that would normally require six conventional pixels should advance patient outcomes. Throughout this transition, our imaging system should further enable lines of inquiry that have previously faced technical obstacles.

One such area of exploration is in multifluorophore detection of sentinel lymph nodes, which was investigated as a part of our clinical study. Early instances of sentinel lymph node mapping used either a blue dye or a radiotracer, whereas modern guidance suggests that both agents should be used to improve identification rates and reduce false-negative rates (*30, 31*). Likewise, early studies into sentinel lymph node staging used a single targeted agent to highlight metastatic lymph nodes, but recent research indicates that paired agents should be used to improve sensitivity and specificity (*32-34*). However, there have been no clinical studies exploring the complementary use of multiple fluorophores for either sentinel lymph node mapping or sentinel lymph node staging. For example, indocyanine green and methylene blue migrate to lymph nodes via the same mechanisms, namely, drift and diffusion from the injection site through the lymphatic capillaries. However, indocyanine green and methylene blue exhibit notable physiochemical differences, with indocyanine green (775 Da) and human serum albumin (66.5 kDa) forming large complexes and methylene blue (320 Da) experiencing no such interaction with serum proteins. Even though the literature largely treats the two fluorophores as equivalent for sentinel lymph node mapping, these similarities and differences cause them to accumulate in a correlated, but not identical, manner in lymph nodes—a distinction which may carry subtle but meaningful diagnostic information. If fluorescence image-guided sentinel lymph node biopsy is to enter practice, clinicians must understand whether such complementary use can offer sustained improvements and how it can be properly translated, using both nonspecific and tumor-targeted probes. Our study did not fully explore these possibilities, looking only at those nonspecific probes available in the clinic as they are used in the standard of care since tumor-targeted probes have not yet been approved by regulatory bodies. Nevertheless, imaging systems like ours that are equipped for multifluorophore detection can enable both the fundamental studies into probe pharmacokinetics and the practical studies into patient outcomes needed to understand and evaluate these techniques as the field advances.

Another area of exploration is in three-dimensional reconstruction of surgical sites, which could not be investigated as a part of our clinical study. In remote surgeries or minimally invasive surgeries where the physician cannot directly view the surgical site, simultaneous

three-dimensional profilometry and fluorescence imaging obviates the need for the physician to mentally map the surgical volume from multiple video streams on a flat-screen display, while in ordinary open surgeries, co-registration of the tissue's profile with its fluorescence frees the surgical staff from the need to constantly turn from the patient to a display. It is unclear how dramatically three-dimensional imaging and fluorescence imaging, when combined in this way, would affect surgical outcomes. Surgeons may experience reduced learning curves for some procedures, observe improved standardization in other procedures, and receive new views of surgical sites that enable previously unidentified treatment approaches, but relevant technologies have been rare and underutilized and relevant studies have faced limitations and yielded inconsistent results (*35*). Our study did not examine whether surgical performance differed based on the presence or absence of a three-dimensional profile, but imaging systems like ours can facilitate this line of inquiry. In open surgeries, our imaging system can be integrated with a near-infrared projector in the surgical lighting system to enable three-dimensional reconstruction of the full surgical field, while in minimally invasive surgeries, our imaging system can be integrated with a near-infrared light channel and a microelectromechanically-actuated grating in an endoscopic lighting system to enable similar reconstruction of the surgical cavity.

Materials and Methods

Study design

The goal of this study was to develop a camera based on the compound eye of the mantis shrimp and to apply this camera for multi-functional imaging during cancer surgery. The camera's development was verified by optoelectronic characterization, while its application involved two series of experiments. The first series of applied experiments established that the camera could record multiple fluorophores and extract the complementary information encoded in their combined fluorescence. This included benchtop experiments that evaluated whether the emission spectra of various fluorophores could be distinguished and animal experiments that tested whether tumor detection with two tumor-targeted fluorescent probes was superior to that with only one tumor-targeted fluorescent probe. The second series of applied experiments established that the camera could function in the operating room, providing features that were conducive to clinical translation and performance that was equivalent to the current standard of care. This included a benchtop experiment that evaluated whether the excitation spectrum and emission spectrum of a fluorophore could be

distinguished and a human study that tested whether sentinel lymph node mapping with near-infrared fluorescent dyes was non-inferior to that with two dark colored dyes and a radiotracer. Additional experiments also explored three-dimensional profilometry with structured illumination and intraoperative imaging during surgical workflows.

Surgeons were blinded to our technology during resection. Analysts were not blinded to the tissue's histopathological status during data processing; instead, statistical methods like cross-validation were used to mitigate overfitting and related effects. Randomization was not necessary given that all animals and patients underwent the same procedures.

All animal experiments were performed under protocols approved by the University of Illinois Institutional Animal Care and Use Committee. The studies complied with all relevant ethical regulations. All human studies were performed under protocols approved by the Institutional Review Board at the University of Illinois at Urbana-Champaign and the Agency for Drugs and Medical Instruments in Skopje, Republic of North Macedonia; the clinical studies were registered on clinicaltrials.gov (trial ID no. NCT03619967). The studies complied with all relevant ethical regulations and adhered to approved guidelines. All patients gave informed consent before joining this HIPAA-compliant study. Inclusion criteria included: early or progressive stage of breast cancer and ability to understand and willingness to sign written informed consent documents. Exclusion criteria included: presence of inflammatory cancerous tissue, history of allergic reactions to iodide or seafood, failure to detect sentinel lymph nodes with radiocolloid and static gamma camera, history of breast surgery, pregnancy, and unwillingness to enter study.

Animal experiments

Male nude athymic J:NU mice (2-3.5 months old, >30 grams each; Jackson Laboratory) were used for in vivo tumor imaging and ex vivo receiver operating characteristic (ROC) analysis. PC3 and LNCaP cells were incubated, grown to 90% confluence, detached, suspended in Matrigel (Corning), and delivered via subcutaneous injection. The suspension consisted of 2×10^7 cells per 100 μL of 10% Matrigel, and the injection comprised of 50 μL of cell-Matrigel solution. For some mice, two subcutaneous injections of LNCaP cells on the left flank and two subcutaneous injections of PC3 cells on the right flank were delivered. For the remaining mice, two subcutaneous injections of PC3 cells, one on each flank, were delivered. The mice were monitored daily for tumors and were imaged when tumors grew to ~1 cm diameter.

Twenty-four hours before imaging, the animals were anesthetized with 1-2% isoflurane, and their skin was prepared using an alcoholic iodine solution and opened during a jugular vein cut-down procedure. For each mouse, two fluorophores were injected in sequence: 200 μL of 0.1 nmol/1 μL IRDye 800CW 2-DG (labeled with 2-deoxy-D-glucose) followed by 100 μL of 2 nmol/100 μL IRDye 680RD EGF (labeled with human recombinant epidermal growth factor). Using a surgical microscope for direct visualization, each fluorophore was administered using a 30-gauge insulin syringe needle inserted into the jugular vein. After injection, gentle pressure was applied to the injection site with a cotton-tip swab for approximately 30 seconds before the skin was closed with a sterile 6-0 Prolene suture. All procedures were done under aseptic conditions.

At imaging time, the mice were anesthetized under 1-2% isoflurane on a heated bed; in vivo images were captured; the surgeon euthanized the mice and removed tissue samples, marking each as either cancer-positive tumor or cancer-negative muscle; and ex vivo images were captured. To mimic the case where IRDye 800CW 2-DG alone was administered, ~20 mW/cm^2 intensity of 780 nm excitation was provided; to mimic the case where IRDye 680RD EGF alone was administered, ~20 mW/cm^2 intensity of 665 nm excitation was provided; and to observe both IRDye 800CW 2-DG and IRDye 680RD EGF, both excitations were provided. The imaging system and excitation sources were set up at a ~0.5-m working distance. The image sensor was equipped with a Canon EF 50mm f/1.2 lens that was focused on the sample and adjusted to an aperture between ~f/2.0 and ~f/4.0. The imaging system was configured to provide a ~200-ms to ~400-ms exposure for the near-infrared pixels to ensure high signal-to-noise ratio in the fluorescence signal. A total of 7 mice were involved, and a total of 48 tissue samples were collected; these tissue samples included 24 tumor tissues (14 PC3 and 10 LNCaP) and 24 healthy tissues.

A multistep procedure was required to locate the tissue sample, extract the near-infrared features, and distinguish the tumor tissues from healthy tissues. Raw images were demosaiced and temporally averaged, and regions of interest corresponding to the tissue samples were selected by hand; these regions were spatially averaged. Near-infrared channels were then mapped from red-green-blue representations to hue-saturation-value representations; ultimately, the value (intensity information) was dropped during the classification procedure because the hue and saturation (spectral information) provided sufficient information about the tissue samples. Generation and analysis of ROC curves then followed the procedures described in the Statistical Analysis subsection below.

Human study

Women with breast cancer undergoing concurrent curative surgery and sentinel lymph node dissection were recruited for in vivo imaging and ex vivo imaging. Before the surgical procedure, technetium-99m-labeled human serum albumin colloid (^{99m}Tc-HSA colloid, via Senti-Scint; 834 μCi), indocyanine green (ICG; 2 mL at 0.5 mg ICG per mL saline), and methylene blue (MB; 1 mL at 10 mg MB per mL water) were injected into the patient's tumor area, followed by 5 min of site massage; at 10–15 min postinjection, the surgeon proceeded with the surgery per standard of care. When suspected sentinel lymph nodes were identified using the radioactivity measured with a gamma probe (Europrobe, EuroMedical Instruments) and the visible contrast provided by MB and ICG, the tissues were removed. All resected samples were analyzed using histopathology to confirm or disaffirm sentinel lymph node status.

For one set of patients (7 patients; mean ± standard deviation age: 58 ± 12 years), images of tissues before and during resection were captured. Under operating conditions, bright light from broadband surgical lamps, ~10 mW/cm^2 from a 780 nm laser, and ~10 mW/cm^2 from a 665 nm laser were provided. The image sensor was configured to provide a 25-ms exposure for color pixels and a ~100- to ~200-ms exposure for near-infrared pixels and was equipped with a Canon EF 50 mm f/1.2 lens adjusted to an aperture of ~f/2.2. The imaging system and excitation sources were set up at a ~2-m working distance. These in vivo samples were not incorporated into the ROC analysis. For the remaining patients (11 patients; mean ± standard deviation age: 61 ± 15 years), images of tissues after resection were captured. To mimic the case where ICG alone was administered, ~5 mW/cm^2 intensity of 780 nm excitation was provided; to mimic the case where MB alone was administered, ~5 mW/cm^2 intensity of 665 nm excitation was provided; and to observe both the ICG and MB, both excitations were provided. The image sensor was configured to provide a ~200-ms exposure for near-infrared pixels and was equipped with a Rokinon 50mm f/1.4 lens adjusted to an aperture of ~f/2.0. The imaging system and excitation sources were set up at a ~0.5-m working distance. These ex vivo samples were incorporated into the ROC analysis: 55 tissue samples were collected of which 49 were lymph nodes and 6 were nonlymphatic structures.

A multistep procedure was required to locate the tissue sample, extract the near-infrared features, and distinguish the sentinel lymph nodes from nonlymphatic structures. Raw images were demosaiced and temporally averaged, and regions of interests corresponding

to the tissue samples were selected by isolating a rectangular window at the image's center and extracting the largest contiguous region with average near-infrared response in the 95th percentile and above; these regions were spatially averaged. Near-infrared channels were then mapped from red-green-blue representations to hue-saturation-value representations. Generation and analysis of ROC curves then followed the procedures described in the Statistical Analysis subsection below.

Statistical analysis

For generation of 95% tolerance ellipses during the fluorescence sensitivity studies, data points were defined as tuples of hues and saturations that laid in a polar coordinate system (i.e., where the hue was the angular coordinate and the saturation was the radial coordinate); however, data points were confined to a sufficiently small region such that the tuples of hues and saturations could be taken to lie in a rectangular coordinate system (i.e., where the hue was the x-coordinate and the saturation was the y-coordinate). All data points were assumed drawn from a binormal distribution. A 95% tolerance ellipse was taken as an ellipse generated by a method that had an approximately 95% probability of covering at least 95% of the population when the empirical mean and covariance were taken as sample estimators of the population statistics. As a result, the 95% tolerance ellipse could be computed using equation (5.9) reported in (*36*).

Generation of ROC curves in both the animal study and human study followed similar procedures. To map the near-infrared signal in the hue-saturation-like/hue-saturation-value-like color space to a classification score appropriate for an ROC analysis without making assumptions on class covariance, quadratic discriminant analysis was used to model probabilities of class membership; to mitigate the risk of overfitting, stratified cross-validation was used in which the data set was broken into folds with equal class distributions and in which the discriminator was trained on out-of-fold samples and tested on in-fold samples. Conventional algorithms for ROC analysis could then be applied to the cross-validated probabilities to evaluate the area under the curve (AUC) for the scorer and to determine the optimal cut-point for a classifier.

Statistical analysis of ROC curves was facilitated using a resampling scheme based on the bootstrap method. A family of ROC curves was produced by generating stratified bootstrap samples of the tissue samples and generating ROC curves for each bootstrap sample. Confidence intervals for the AUCs, the true positive rates, and the false positive rates were

computed using the bias-corrected and accelerated bootstrap interval: a bias-correction parameter related to statistical bias and an acceleration parameter related to statistical skewness were estimated from the data set and used to correct confidence intervals computed from bootstrapped samples. Confidence intervals were generated at 95% confidence.

Comparison of AUCs was facilitated using a bootstrap-based paired difference test. For each scenario of dye administration, stratified bootstrap samples were generated with the same tissue samples, and ROC curves and AUCs were computed for each bootstrap sample. To compare any two dye administration scenarios, the difference between the AUCs observed over all tissue samples was computed, and the probability under the null hypothesis of the observed difference or a more extreme difference was found by finding the corresponding percentile in the distribution of differences between the paired AUCs observed for the bootstrap samples. These tests were run as one-tailed tests of superiority where, for example, the two-dye administration scenario could be presumed better than the one-dye administration scenario. All tests were run at a significance level of $\alpha = 0.05$.

Supplementary Materials

Supplementary Materials and Methods

Fig. S1. Demonstration of the bio-inspired imaging system's color reproduction for natural scenes.

Video S1. Tumor detection in small-animal model of human prostate cancer showing overlaid information from color images of mouse, near-infrared images of fluorescently labeled tumors, and three-dimensional shape of mouse and tumors.

Video S2. Sentinel lymph node mapping in patients with breast cancer using near-infrared fluorescence from indocyanine green.

Table S1. Optical performance of state-of-the-art near-infrared fluorescence (NIRF) imaging systems compared with our bio-inspired NIRF imager.

Table S2. Optical performance of pixelated near-infrared fluorescence (NIRF) imaging systems compared with our bio-inspired NIRF imager.

687 Data S1. Classification scores from the receiver operating characteristic (ROC) analysis of
688 tumor detection in a small-animal model of human prostate cancer.

689 Data S2. Classification scores from the receiver operating characteristic (ROC) analysis of
690 sentinel lymph node detection in patients with breast cancer.

691 **References and Notes**

692 1. L. V. Wang, J. Yao, A practical guide to photoacoustic tomography in the life sciences.
693 *Nature Methods* **13**, 627-638 (2016).
694 2. H. Tu *et al.*, Stain-free histopathology by programmable supercontinuum pulses. *Nature*
695 *Photonics* **10**, 534-540 (2016).
696 3. A. L. Vahrmeijer, M. Hutteman, J. R. van der Vorst, C. J. H. van de Velde, J. V. Frangioni,
697 Image-guided cancer surgery using near-infrared fluorescence. *Nature Reviews Clinical*
698 *Oncology* **10**, 507-518 (2013).
699 4. S. H. Yun, S. J. J. Kwok, Light in diagnosis, therapy and surgery. *Nature Biomedical*
700 *Engineering* **1**, (2017).
701 5. R. Daw, J. Finkelstein, Lab on a chip. *Nature* **442**, 367 (2006).
702 6. G. Hong, A. L. Antaris, H. Dai, Near-infrared fluorophores for biomedical imaging. *Nature*
703 *Biomedical Engineering* **1**, (2017).
704 7. G. M. van Dam *et al.*, Intraoperative tumor-specific fluorescence imaging in ovarian cancer
705 by folate receptor-α targeting: first in-human results. *Nature Medicine* **17**, 1315-1319
706 (2011).
707 8. A. V. DSouza, H. Lin, E. R. Henderson, K. S. Samkoe, B. W. Pogue, Review of
708 fluorescence guided surgery systems: identification of key performance capabilities beyond
709 indocyanine green imaging. *Journal of Biomedical Optics* **21**, (2016).
710 9. R. R. Zhang *et al.*, Beyond the margins: real-time detection of cancer using targeted
711 fluorophores. *Nature Reviews Clinical Oncology* **14**, 347-364 (2017).
712 10. M. Garcia *et al.*, Bio-inspired imager improves sensitivity in near-infrared fluorescence
713 image-guided surgery. *Optica* **5**, 413-422 (2018).
714 11. G. E. Moore, W. T. Peyton, L. A. French, W. W. Walker, The Clinical Use of Fluorescein
715 in Neurosurgery. *Journal of Neurosurgery* **5**, 392-398 (1948).
716 12. M. Koller *et al.*, Implementation and benchmarking of a novel analytical framework to
717 clinically evaluate tumor-specific fluorescent tracers. *Nature Communications* **9**, (2018).

13. P. S. Low, S. Singhal, M. Srinivasarao, Fluorescence-guided surgery of cancer: applications, tools and perspectives. *Current Opinion in Chemical Biology* **45**, 64-72 (2018).
14. R. L. Siegel, K. D. Miller, A. Jemal, Cancer statistics, 2017. *CA: A Cancer Journal for Clinicians* **67**, 7-30 (2017).
15. E. R. Fossum, CMOS image sensors: electronic camera-on-a-chip. *IEEE Transactions on Electron Devices* **44**, 1689-1698 (1997).
16. C. Darwin, *On the origin of species by means of natural selection, or, The preservation of favoured races in the struggle for life*. (John Murray, London, England, 1859).
17. Y. M. Song *et al.*, Digital cameras with designs inspired by the arthropod eye. *Nature* **497**, 95-99 (2013).
18. H. H. Thoen, M. J. How, T.-H. Chiou, J. Marshall, A Different Form of Color Vision in Mantis Shrimp. *Science* **343**, 411-413 (2014).
19. T. W. Cronin, S. Johnsen, N. J. Marshall, E. J. Warrant, *Visual Ecology*. (Princeton University Press, Princeton, NJ, 2014).
20. J. A. Lenero-Bardallo, D. H. Bryn, P. Hafliger, Bio-Inspired Asynchronous Pixel Event Tricolor Vision Sensor. *IEEE Transactions on Biomedical Circuits and Systems* **8**, 345-357 (2014).
21. T. York *et al.*, Bioinspired Polarization Imaging Sensors: From Circuits and Optics to Signal Processing Algorithms and Biomedical Applications. *Proceedings of the IEEE* **102**, 1450-1469 (2014).
22. N. Niikura *et al.*, Loss of Human Epidermal Growth Factor Receptor 2 (HER2) Expression in Metastatic Sites of HER2-Overexpressing Primary Breast Tumors. *Journal of Clinical Oncology* **30**, 593-599 (2012).
23. N. J. Marshall, M. F. Land, C. A. King, T. W. Cronin, The compound eyes of mantis shrimps (Crustacea, Hoplocarida, Stomatopoda). I. Compound eye structure: the detection of polarized light. *Philosophical Transactions of the Royal Society B* **334**, 33-56 (1991).
24. Teledyne QImaging, "MicroPublisher 6™ CCD Camera Datasheet," (2019).
25. H. E. Rhodes, H. Nozaki, S. Manabe, "US2009/0200580A1 - Image sensor and pixel including a deep photodetector," (Omnivision Technologies, Inc., United States, 2009).
26. T. Dai, H.-C. Tai, S. Manabe, H. Nozaki, H. E. Rhodes, "US8228411B2 - Circuit and photo sensor overlap for backside illumination image sensor," (Omnivision Technologies, Inc, United States, 2012).

27. International Electrotechnical Commission, "Medical electrical equipment - Part 2-41: Particular requirements for the safety of surgical luminaires and luminaires for diagnosis (IEC 60601-2-41:2000)," (2000).
28. International Electrotechnical Commission, "Safety of laser products - Part 1: Equipment classification, requirements and user's guide (IEC 60825-1:1993+AMD1:1997+AMD2:2001 CSV)," (2001).
29. R. M. Turner, R. J. Guttosch, in *ICIS '06, International Congress of Imaging Science*. (Society for Imaging Science and Technology, Rochester, NY, 2006), pp. 175-181.
30. G. H. Lyman *et al.*, American Society of Clinical Oncology Guideline Recommendations for Sentinel Lymph Node Biopsy in Early-Stage Breast Cancer. *Journal of Clinical Oncology* **23**, 7703-7720 (2005).
31. H. S. Cody *et al.*, Complementarity of Blue Dye and Isotope in Sentinel Node Localization for Breast Cancer: Univariate and Multivariate Analysis of 966 Procedures. *Annals of Surgical Oncology* **8**, 13-19 (2001).
32. J. Baeten, J. Haller, H. Shih, V. Ntziachristos, In Vivo Investigation of Breast Cancer Progression by Use of an Internal Control. *Neoplasia* **11**, 220-227 (2009).
33. E. N. Savariar *et al.*, Real-time In Vivo Molecular Detection of Primary Tumors and Metastases with Ratiometric Activatable Cell-Penetrating Peptides. *Cancer Research* **73**, 855-864 (2013).
34. K. M. Tichauer *et al.*, Microscopic lymph node tumor burden quantified by macroscopic dual-tracer molecular imaging. *Nature Medicine* **20**, 1348-1353 (2014).
35. A. Arezzo *et al.*, The use of 3D laparoscopic imaging systems in surgery: EAES consensus development conference 2018. *Surgical Endoscopy* **33**, 3251-3274 (2019).
36. V. Chew, Confidence, Prediction, and Tolerance Regions for the Multivariate Normal Distribution. *Journal of the American Statistical Association* **61**, 605-617 (1966).
37. S. Zhang, S.-T. Yau, Generic nonsinusoidal phase error correction for three-dimensional shape measurement using a digital video projector. *Applied Optics* **46**, 36-43 (2007).
38. R. M. Goldstein, H. A. Zebker, C. L. Werner, Satellite radar interferometry: Two-dimensional phase unwrapping. *Radio Science* **23**, 713-720 (1988).
39. H. Du, Z. Wang, Three-dimensional shape measurement with an arbitrarily arranged fringe projection profilometry system. *Optics Letters* **32**, 2438-2440 (2007).
40. Z. Chen, X. Wang, R. Liang, RGB-NIR multispectral camera. *Optics Express* **22**, 4985-4994 (2014).

41. Z. Chen, N. Zhu, S. Pacheco, X. Wang, R. Liang, Single camera imaging system for color and near-infrared fluorescence image guided surgery. *Biomedical Optics Express* **5**, 2791-2797 (2014).

Acknowledgments: The authors thank James Hutchinson and Patricia J. Watson for manuscript editing. **Funding:** This work was funded by grants from the U.S. Air Force Office of Scientific Research (FA9550-18-1-0278), Congressionally Directed Medical Research Programs (W81XWH-19-1-0299), National Science Foundation (2030421), and Office of Naval Research (N00014-19-1-2400). **Author contributions:** V.G. conceived the idea behind the image sensor and oversaw the entire project. M.G. and V.G. designed the image sensor. Optical evaluation of the image sensor, including optoelectronic characterization and fluorescence sensitivity, was performed by S.B., M.G., T.D., Z.Z., Z.L., and V.G. Cell cultures were maintained by C.K. and L.W.D. Animal studies were designed, performed, and evaluated by S.B., M.G., T.D., C.K., K.K., L.W.D., S.N., and V.G. The three-dimensional reconstruction algorithm was developed and evaluated by S.B., M.G., and T.D. Human studies were designed, performed, and evaluated by S.B., M.G., T.D., R.C., I.F., B.K., S.S., M.B.T., N.T.D., N.J., D.M., G.P., G.K., and V.G. Statistical analysis was provided by S.B. and T.D. The manuscript was written by S.B., M.G., and V.G. All authors proofread the manuscript. **Competing interests:** M.G. and V.G. are listed as inventors on Patent US20200120293A1, "Multispectral imaging sensors and systems." The authors are not aware of any other competing interests. **Data and materials availability:** All data associated with this study are present in the paper or the Supplementary Materials. No materials were used under a material transfer agreement or other license.

Figures and Figure Captions

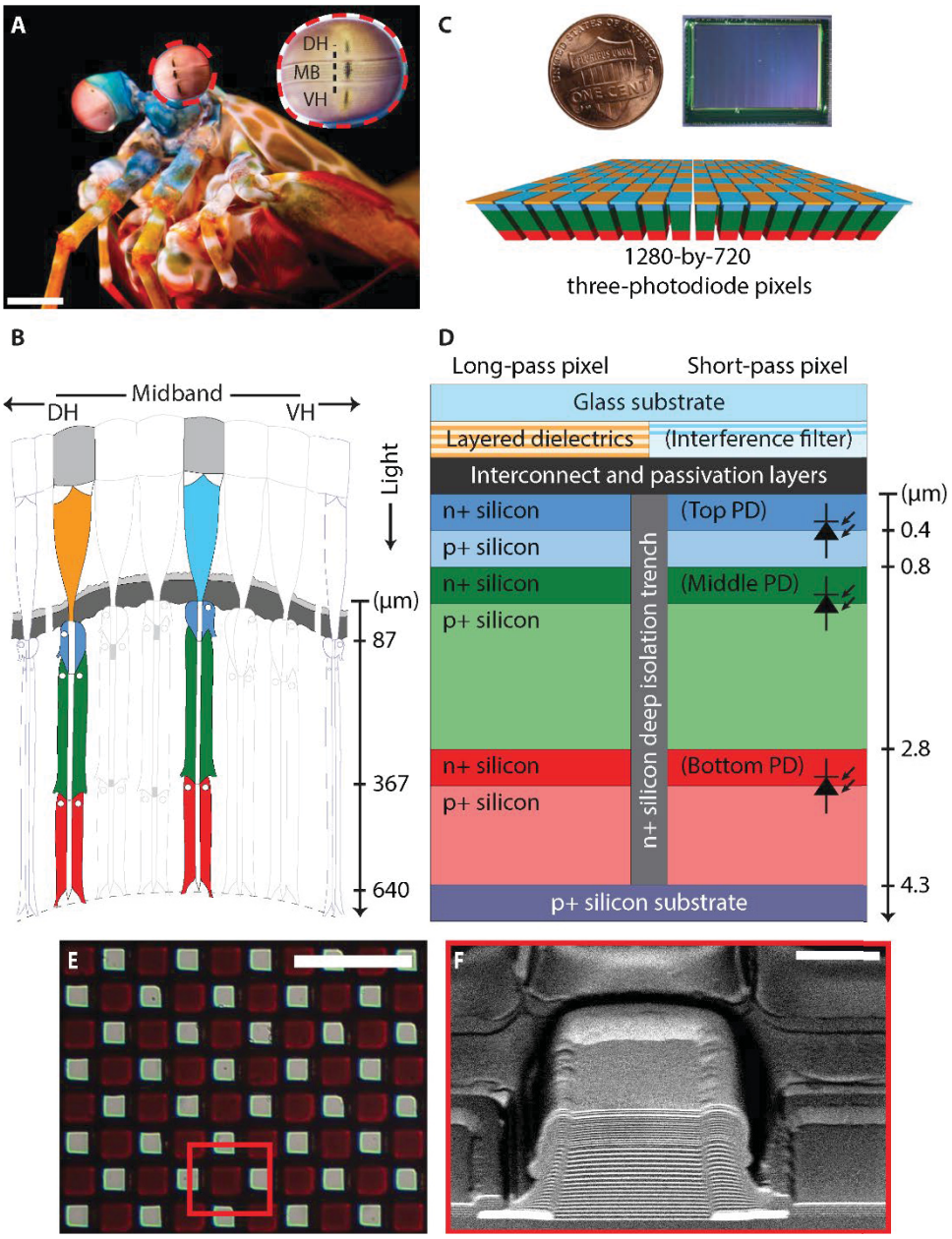

Fig. 1. Comparison of the mantis shrimp compound eye and the bio-inspired image sensor. (A) A photograph of the mantis shrimp *Odontodactylus scyllarus*, and a magnified photograph of its compound eye (inset) are shown; scale bar indicates 5 mm. (B) A cross-section of the compound eye of the mantis shrimp *Gonodactylus chiragra* showing the photosensitive cells and spectral filters; structures without dimensions are not to scale. The photosensitive cells at the bottom, middle, and top have been colored red, green, and blue to emphasize that longer wavelengths penetrate farther than shorter wavelengths. (C) A photograph of the bio-inspired imager (top), and a magnified illustration of the pixel array (bottom) are shown; scale is indicated by the penny. (D)

A cross-section of the bio-inspired imager showing the vertically stacked photodiodes, composed from layers of n-doped silicon and p-doped silicon, and the interference filters, composed from layers of different dielectrics; structures without dimensions are not to scale. The photodiodes at the bottom, middle, and top have been colored red, green, and blue to emphasize that longer wavelengths are absorbed deeper than shorter wavelengths. (E) An optical microscope image of the pixelated optical filters showing the long-pass filters as dark colored squares and the short-pass filters as bright colored squares; scale bar indicates 50 μm. (F) A scanning electron microscope image depicting a long-pass filter in the center and short-pass filters on both sides; scale bar indicates 5 μm. Panel A credit, Michael Bok, University of Lund. In panel B, the cross-section is adapted from (*23*). In panels C and D, the magnified illustration and cross-section present a conceptual view of the device, omitting some details for clarity. DH, dorsal hemisphere; MB, midband; PD, photodiode; VH, ventral hemisphere.

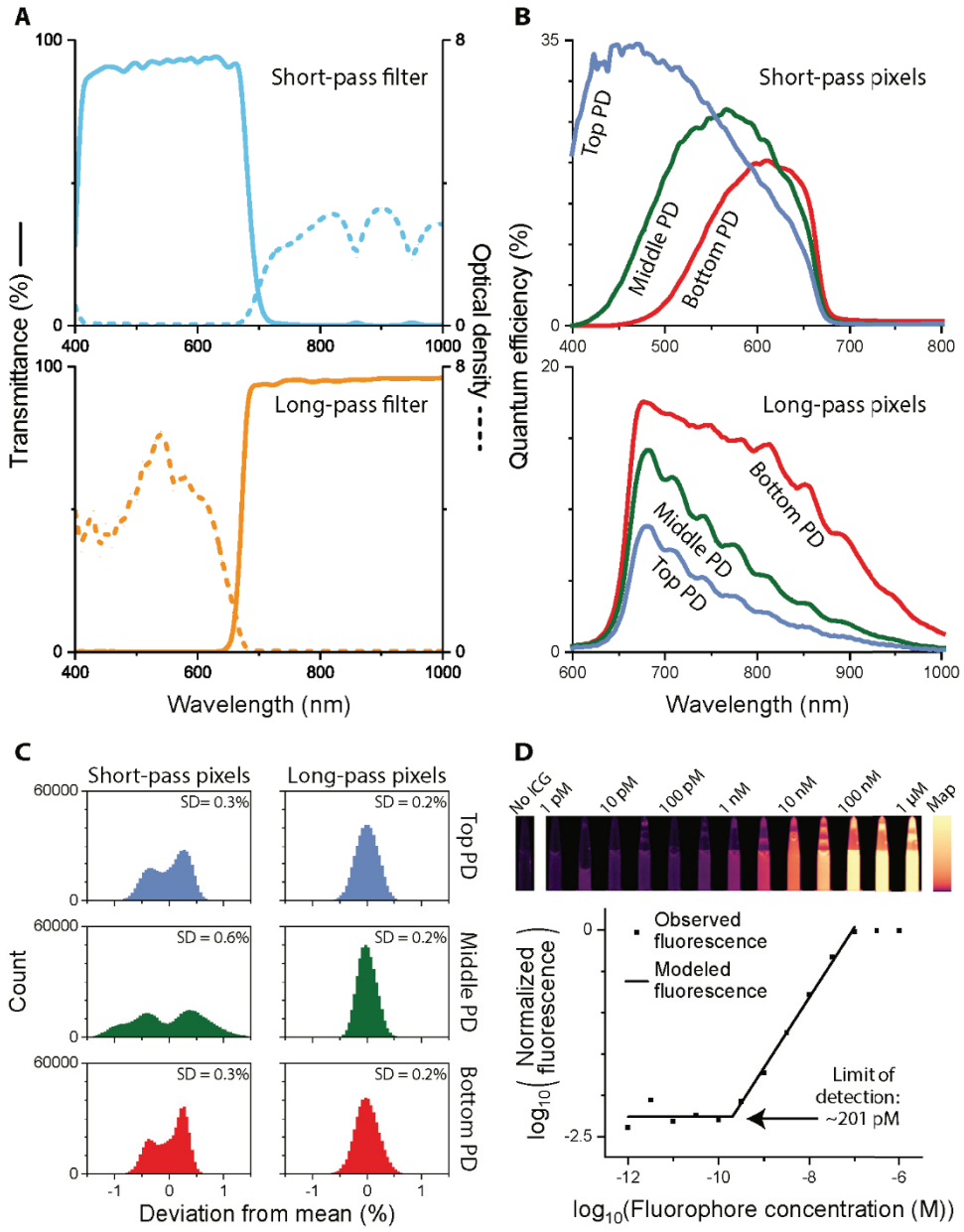

Fig. 2. Optoelectronic characteristics of the bio-inspired imager. (A) Transmission and optical density curves for the short-pass filters (top) and long-pass filters (bottom) are shown. (B) Quantum efficiency curves for the three photodiode (PD) layers under the short-pass filters (top) and long-pass filters (bottom) are presented. (C) Fixed pattern noise histograms at half linear full-scale range for the three photodiode layers under the short-pass filters (left) and long-pass filters (right) are shown. The fixed pattern noise is the standard deviations of the histograms. (D) Provided is the normalized fluorescence observed and modeled across different concentrations of indocyanine green (ICG); temporally averaged intensity images are displayed against concentration (top), and spatiotemporally averaged intensities are plotted against concentration (bottom). The detection

843 limit (201 pM) is the transition point between the constant region below the noise floor and the
844 linear region above the noise floor. SD, standard deviation.

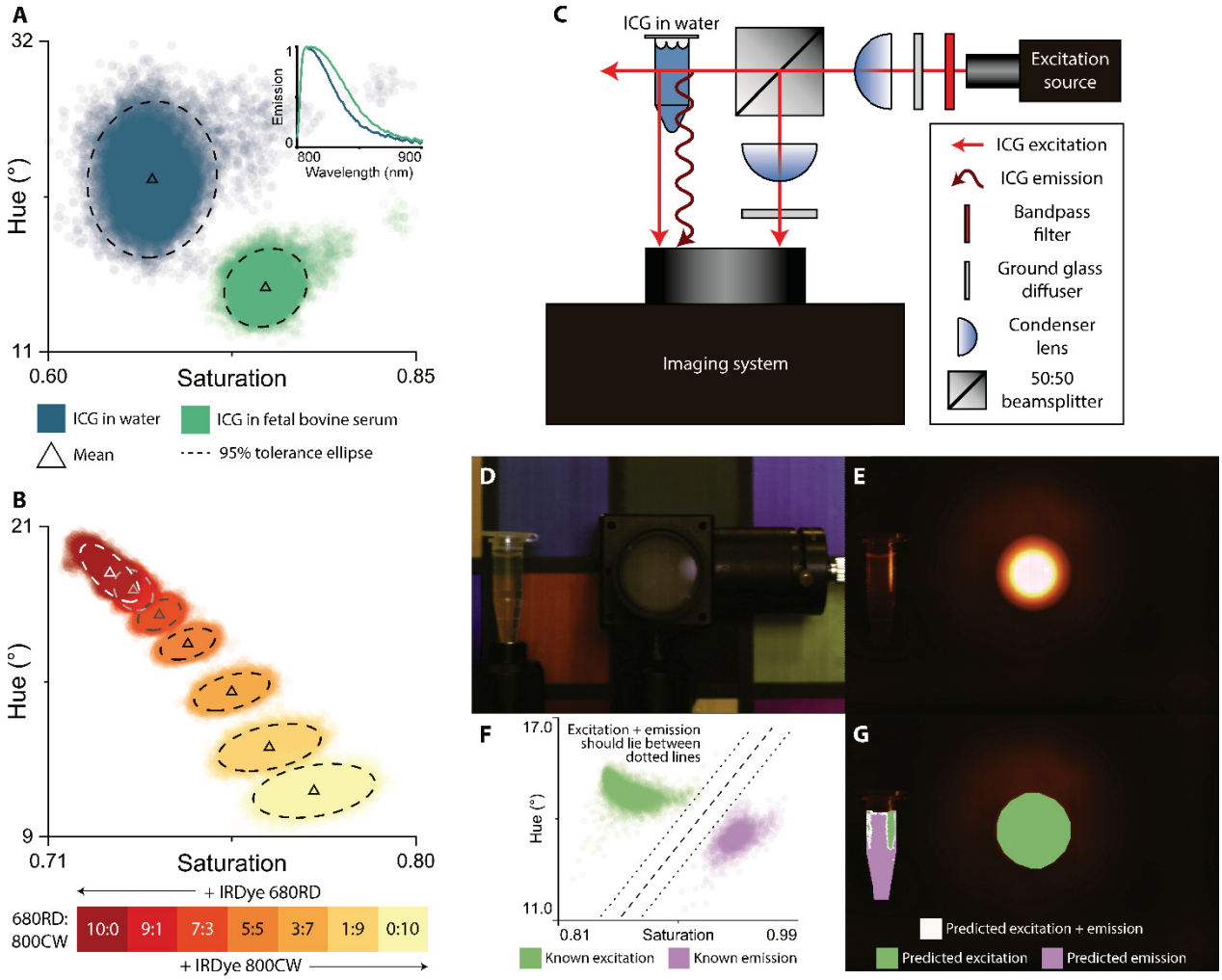

Fig. 3. Imaging capabilities of the bio-inspired imager. (A) Discrimination of indocyanine green (ICG) diluted in two different media (deionized water and fetal bovine serum) showing the uncertainty created by temporal and spatial noise. Observations of different solutions appear in different clusters and are circled by 95% tolerance ellipses ($n_{observations}$ = 27,000 per solution). Emission spectra for the two solutions are displayed in the top-right inset. (B) Discrimination of IRDye 680RD and IRDye 800CW mixed in different volumetric ratios showing the uncertainty created by temporal and spatial noise. Observations of different ratios appear in different clusters and are circled by 95% tolerance ellipses ($n_{observations}$ = 27,000 per solution). Emission spectra for IRDye 680RD and IRDye 800CW are strongest at ~694 nm and ~794 nm, respectively. (C-G) Discrimination of indocyanine green excitation and emission in the presence of spatial noise. (C, D) Shown is a schematic (C) and a color photograph (D) of the experimental setup. Excitation is projected from the port; emission is generated in and excitation is reflected from the vial. (E) Near-infrared image of the experimental setup. Pure excitation is observed inside the port; pure excitation, pure emission, and mixed excitation-emission are observed inside the vial. (F)

Temporally-averaged observations of pure excitation and pure emission cluster differently, permitting classification via a simple margin classifier such as a linear support vector machine ($n_{observations}$ = 8,246 for excitation, 2,665 for emission). Observations of mixed excitation-emission should lie between and are classified accordingly. (G) Labels of pure excitation, pure emission, and mixed excitation-emission can be applied to the near-infrared image using the classifier.

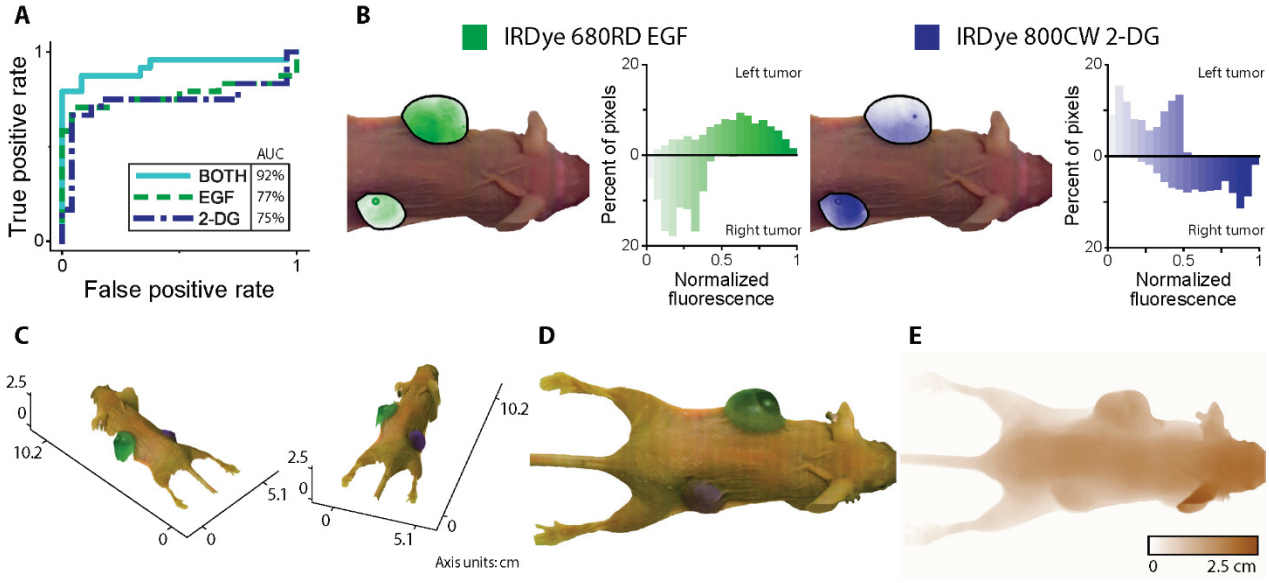

Fig. 4. Human tumor detection in nude mice using the bio-inspired imaging sensor. (A) Receiver operating characteristic (ROC) curves for tumor detection in mice using the two dyes IRDye 680RD epidermal growth factor (EGF) and IRDye 800CW 2-deoxy-D-glucose (2-DG) together or each dye alone. Area under the curve (AUC) improves for two dyes compared to one dye ($p < 0.05$; $n_{tumor} = 24$, $n_{healthy} = 24$). (B) Shown is the normalized fluorescence from human PC3 prostate tumors in a nude athymic mouse overlaid on a color image and plotted in a histogram. IRDye 680RD EGF (left) and IRDye 800CW 2-DG (right) are taken up by the two tumors but the dyes exhibit substantial heterogeneity between and within those tumors. IRDye 680RD EGF dominates in the left tumor, whereas IRDye 800CW 2-DG dominates in the right tumor. (C) Two color images of the same tumor-bearing mouse showing the estimated three-dimensional profile with the tumors highlighted and the dominant targeted probes indicated (see video S1). (D) Color image of the tumor-bearing mouse showing the tumors highlighted along with the dominant targeted probes as indicated by normalized fluorescence. (E) Three-dimensional profile of the tumor-bearing mouse indicating the out-of-plane height as extracted from structured illumination.

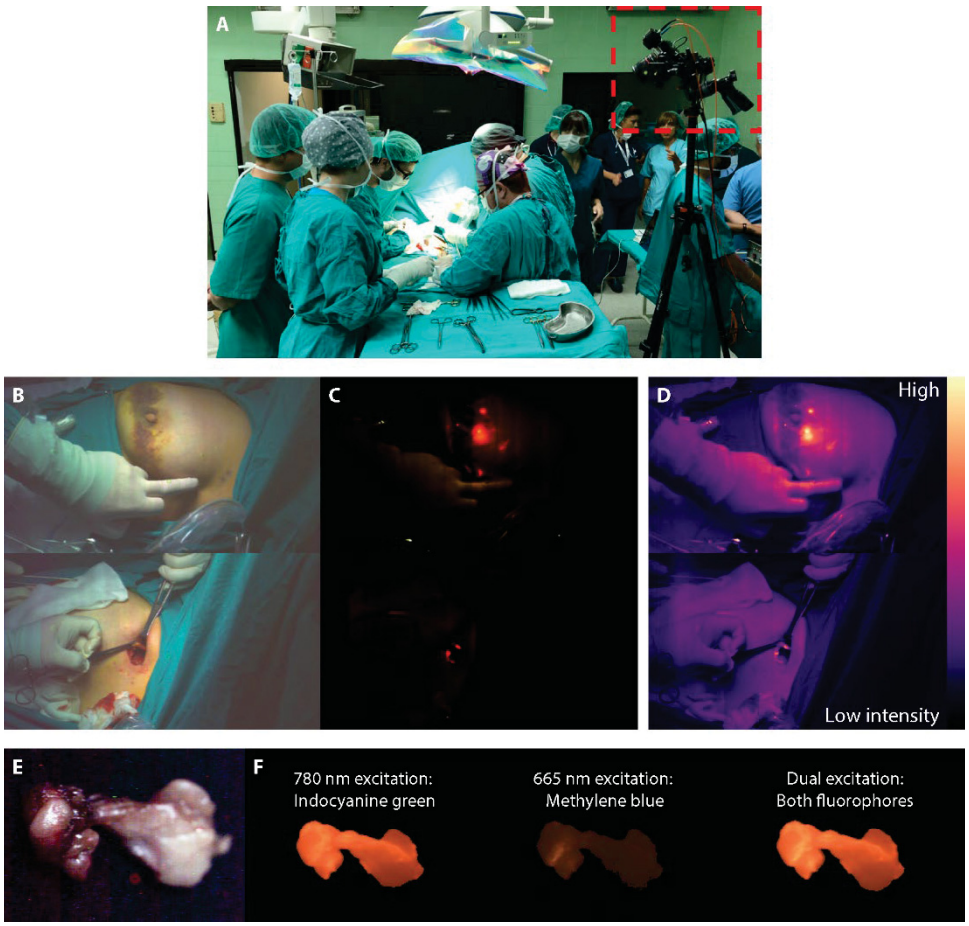

Fig. 5. Clinical feasibility study using the bio-inspired imaging sensor to map sentinel lymph nodes. (A) Color photograph of the imaging system (red box, top-right) integrated into the operating room. (B-D) In vivo images before and after the surgical incision to remove lymph nodes from a patient with breast cancer, with the fluorophores indocyanine green and methylene blue administered by injection at the surgical site (see video S2). (B, C) Shown is a color photograph (B) and near-infrared image in trichromatic format (i.e., with spectral discrimination) (C) of the surgical site; the trichromatic formatting presents the individual responses from the three near-infrared channels like a color image, preserving the spectral information. Dye injection points and sentinel lymph nodes exhibit high signal-to-background ratio with no false positives due to specular reflections. (D) Shown is a near-infrared image in monochromatic format (i.e., without spectral discrimination) of the surgical site; the monochromatic formatting presents the aggregate response across the three-near infrared channels like an intensity image, eliminating the spectral information. Surgical instruments appear as false positives due to the strong signal from specular reflection. (E, F) An ex vivo resected lymph node presented as a color photograph (E) and as near-infrared images (F) representing indocyanine green alone (left), methylene blue alone (center), and both dyes together (right). The two dyes accumulate differently in different tissue sites. In vivo images are

representative of 7 patients; ex vivo images are representative of 49 resected lymph nodes collected from eleven patients.

Table 1. Clinical feasibility study with the bio-inspired imaging sensor to map sentinel lymph nodes. Receiver operating characteristic analysis for sentinel lymph node mapping using indocyanine green in eleven patients with breast cancer undergoing surgical resection ($n_{lymph} = 49$, $n_{non\text{-}lymphatic} = 6$). The optimum true positive rate and optimum false positive rate were evaluated at the threshold with maximal Youden's index.

Estimate	*Mean*	*95% confidence interval*
Area under the curve	88%	77.9–96.6%
Optimum true positive rate	84%	73.5–93.9%
Optimum false positive rate	0%	0–0%

SUPPLEMENTARY MATERIALS

Supplementary Materials and Methods

Assembly of imaging system

The array of three vertically stacked photodiodes was fabricated using a specialized 180 nm complementary metal-oxide-semiconductor (CMOS) process and packaged in a ceramic housing by Sigma Corporation, and the array of pixelated long-pass filters and short-pass filters was produced using alternating steps of deposition and lift-off on a glass substrate by Salvo Technologies. The imager and filters were aligned using a submicron-resolution, six-degree-of-freedom positioning stage and were flip-chip bonded using ultraviolet-curing optical adhesives at the University of Illinois at Urbana-Champaign. The stacked photodiodes were fabricated with a pitch of 7.8 μm, while the spectral filters were produced with a pitch of 15.6 μm. During the initial alignment, the two arrays were adjusted so that each spectral filter was centered over a stacked photodiode; subsequently, images were formed from those pixels where a spectral filter completely covered a stacked photodiode. Diagrams of and specifications for the imaging system do not include those pixels where multiple spectral filters covered a stacked photodiode.

Two microscope images were collected of pixelated filters that were not integrated into the complete imaging system. An optical microscope image was taken without modifications to the pixelated filters, while a scanning electron microscope image was taken after a small area of roughly 50 μm by 20 μm was milled via focused ion beam.

Operation of imaging system

As appropriate, several transformations were applied to images captured with the system according to accepted practice in the imaging community. To account for dark currents and other offsets, dark images were removed from images before additional processing. To mitigate the effects of spatial nonuniformity in the image sensor, an affine transformation (i.e., multiplication by a gain and addition of an offset) was applied across all photodiodes and all channels so that every photodiode in each channel would return an identical value when illuminated by a uniform source. Since observations from the visible spectrum can only be made under the short-pass filters and observations from the near-infrared spectrum can only be made under the long-pass filters, a demosaicing routine was used to extract each set of observations from the corresponding set of pixels and to bilinearly interpolate each set of observations across the entire pixel array. To bring the spectral response provided by the image sensor in line with the spectral response provided by the human eye, an affine transformation (i.e., multiplication by a color correction matrix and addition of an offset vector) was applied to convert the three measurements made at every pixel in the color channel to the color measurements that would be made at every pixel by a CIE 1931 standard observer.

The transformation from a red-green-blue (RGB) color space to a hue-saturation-value (HSV) color space was used to map the three near-infrared observations under each long-pass filter to two components that described the near-infrared spectrum (hue and saturation) and one component that described the near-infrared intensity (value). The observation from the bottom photodiode was taken as R, the observation from the middle photodiode was taken as G, and the observation from the top photodiode was taken as B; then, the hue H, the saturation S, and the value V were defined as:

$$H = \begin{cases} 0 & \text{if } \max(R,G,B) = \min(R,G,B) \\ 60° \times \left(0 + \dfrac{G-B}{\max(R,G,B) - \min(R,G,B)}\right) & \text{if } \max(R,G,B) \neq \min(R,G,B) \text{ and } \max(R,G,B) = R \\ 60° \times \left(2 + \dfrac{B-R}{\max(R,G,B) - \min(R,G,B)}\right) & \text{if } \max(R,G,B) \neq \min(R,G,B) \text{ and } \max(R,G,B) = G \\ 60° \times \left(4 + \dfrac{R-G}{\max(R,G,B) - \min(R,G,B)}\right) & \text{if } \max(R,G,B) \neq \min(R,G,B) \text{ and } \max(R,G,B) = B \end{cases}$$

$$S = \begin{cases} 0 & \text{if } \max(R,G,B) = 0 \\ \dfrac{\max(R,G,B) - \min(R,G,B)}{\max(R,G,B)} & \text{if } \max(R,G,B) \neq 0 \end{cases}$$

$$V = \max(R,G,B)$$

Optoelectronic characterization

The transmissions and optical densities of the spectral filters were reported by the manufacturer and are reproduced here.

To evaluate the image sensor's quantum efficiencies, an integrating sphere (IS200-4, Thorlabs) was used to couple light from a computer-controlled monochromator (Acton SP2150 w/ TS-428 bulb, Princeton Instruments) to either the image sensor under test or a power meter (PM100D w/ S130VC sensor, Thorlabs). The monochromator was programmed to sweep from 300 nm to 1100 nm in 3-nm steps, and the image sensor and power meter were used to collect a video and measure the irradiance at each step. First, a responsivity curve was produced for each channel. The videos were linearized and spatiotemporally averaged (over 7 frames and 64,516 pixels near the image centroid), providing a per-pixel digital number for each channel, while the irradiances were multiplied by the pixel's fill factor, photosensitive area, and integration time and divided by the appropriate photon energy, providing a per-pixel photon count; the responsivity curve for each channel was then produced by dividing the per-pixel digital numbers for that channel by the per-pixel photon counts. Second, a quantum efficiency curve was produced for each channel. A conversion gain, defined as the digital number induced by a single photoelectron, was computed for each channel using a photon transfer method; the quantum efficiency curve was then computed for each channel by dividing the responsivity curve for that channel by the conversion gain for that channel.

To evaluate the image sensor's linear full well capacity and linear full-scale range, an integrating sphere (819D-SF-4 or 819D-IS-5.3, Newport) was used to couple light from color light emitting diodes (LEDs) to the image sensor under test. The full well capacity and full-scale range for the photodiodes were measured under the short-pass filters and assumed equal under the long-pass filters since identical photodiodes were used under the short-pass filters and the long-pass filters. When evaluating the full well capacity alongside the read noise, a blue LED and a red LED (M470L4 and M625L4, Thorlabs) were simultaneously mixed to produce equal responses across all three photodiode layers, and when evaluating the full-scale range alongside the fixed pattern noise, a blue LED, a green LED, and a red LED (LZ4-00B208, LZ4-00G108, and LZ4-00R108, LED Engin) were independently enabled to probe the top photodiode layer, the middle photodiode layer, and the bottom photodiode layer, respectively; note that these values should have been identical across both experiments but were independently confirmed each time. Current was driven into or switched between the LEDs as the integration time was adjusted on the image sensor, providing videos for each LED configuration at multiple integration times. The videos were spatiotemporally averaged at each integration time (over all pixels and 95–100 frames), and a plot was produced of spatiotemporal average versus integration time; a piecewise function

composed of a linear function (modeling the linearity of the image sensor) and an inverse exponential function (modeling the saturation of the image sensor) was fitted to the data. The bottom of the full-scale range was found by extrapolating the piecewise function's linear portion to shorter integration times and finding the digital value at zero integration time, while the top of the full-scale range was found by extrapolating the piecewise function's linear portion to larger integration times and finding the digital value where the percent residual between the extrapolated line and actual data exceeded ±1% for the last time. The full-well capacity was defined as the difference between these two digital numbers divided by a conversion gain representing the digital number induced by a single electron, while the full-scale range was simply defined as the difference between these two digital numbers.

To evaluate the image sensor's read noise and dark current, a lens cap was installed over the image sensor under test. The image sensor was used to collect videos at multiple integration times. To compute the read noise, the temporal standard deviation was computed over each video (over 95 frames), the spatial average (via the median) was computed over the temporal standard deviations (over all pixels), and a plot was produced of spatially averaged temporal standard deviation versus integration time; a linear function was fitted to the data, and the read noise was taken as the spatially averaged temporal standard deviation extrapolated to minimum integration time. To compute the dark current, the spatiotemporal average was computed over each video (over 95 frames and all pixels), and a plot was produced of spatiotemporal averages versus integration time; a linear function was fitted to the data, and the dark current was taken as the slope. The read noise and the dark current were converted from quantities in digital numbers to quantities in electrons by dividing through by a conversion gain representing the digital number induced by a single electron.

To evaluate the image sensor's fixed pattern noise, an integrating sphere (819D-SF-4 or 819D-IS-5.3, Newport) was used to couple light from an incandescent lamp (OSL1 w/ OSL1B bulb, Thorlabs) to the image sensor under test. The incandescent lamp generally induced a larger response under the long-pass filter than under the short-pass filter and in the deeper photodiodes than in the shallower photodiodes, so smaller currents/shorter integration times were generally used for those channels that were more responsive than for those channels that were less responsive. Current was driven into the incandescent lamp as the integration time was adjusted on the image sensor, providing videos for each current at multiple integration times. The video was identified where the spatiotemporal average fell closest to half linear full-scale range; this video was temporally averaged (over 100 frames), spatially cropped (to 95% of the frame), and both subtracted by the spatial mean (providing an array of deviations from the mean) and divided by the spatial mean (providing an array of percent deviations from the mean). A histogram was produced over the array (producing the fixed pattern noise histogram), and the standard deviation was computed over the histogram (producing the fixed pattern noise).

Fluorescence sensitivity

For the experiments in this section, the image sensor was equipped with a photographic lens (50mm f/1.4 AS IF UMC, Rokinon, or EF 50 mm f/1.2L USM, Canon), and samples were kept in 1.5-mL microcentrifuge tubes (022363204, Eppendorf). Unless otherwise noted, a diode laser (BWF2-780, B&W Tek), a laser-shaping filter (FF01-769/41 Brightline bandpass filter, Semrock), and a diffusing lens (ACL2520U-DG15, Thorlabs) were used to produce 780 nm illumination, while another diode laser (BWF2-665, B&W Tek), a laser-shaping filter (665.85-1 OD6 Ultra Narrow Bandpass Filter, Alluxa), and a diffusing lens (ACL2520U-DG15, Thorlabs) were used to produce 665 nm illumination.

To determine the detection limit for indocyanine green (ICG) in fetal bovine serum (FBS), solutions of ICG in 90% FBS were produced via half-log dilution from 1 μM to 1 pM, and samples of ICG in FBS were collected at each concentration (13 total samples). The lens (Rokinon) was adjusted to f/2.0 aperture, the sensor was configured with 200-ms exposure time, the 780 nm illumination was tuned to 20 mW/cm^2, and the samples were imaged one-by-one. In this work, the limit of detection was taken as the fluorophore concentration where the average fluorescent signal, as a function of fluorophore concentration, changes from a constant function (below the noise floor) to an increasing function (above the noise floor). Videos of each sample were converted from a red-green-blue representation to a hue-saturation-value representation; the intensity information encoded in the value was retained, while the spectral information encoded in the hue and saturation were discarded. The average signal was established at each concentration by taking the temporal mean at each pixel (over 25 frames) and the spatial mean across pixels (over 1769 pixels near the vial centers), and the average signal, normalized to unit maximum, was plotted against concentration on log-log axes. The concentration where the average signal became constant or began decreasing was identified as the quenching concentration, and a piecewise linear function defined by a zero-slope segment below a threshold, a positively sloped segment above a threshold, and continuity at the threshold was fit using non-linear least squares below the quenching concentration; the limit of detection was taken as the piecewise linear function's threshold. (Note that some fluid was lost during transfer of the 3.16 pM sample from the stock flask to the sample vial; however, the region of interest within the vials was selected so that this reduced volume did not affect the results.)

To demonstrate the capability to differentiate ICG dissolved in two different media, solutions of ICG in deionized water and in FBS were produced at a concentration of 3 μM. One sample of each solution was collected. The lens (Canon) was adjusted to f/2.2 aperture, the sensor was configured with 200-ms exposure time, and the 780 nm illumination was tuned to 25 mW/cm^2. Videos were captured of each sample and converted from a red-green-blue representation to a hue-saturation-value representation; the spectral information encoded in the hue and saturation was retained, while the intensity information encoded in the value was discarded. For each solvent, a collection of hue-saturation points was produced by this process; a region of interest of 30-by-30 pixels across 30 frames (27,000 total pixels) was selected, and a 95% tolerance ellipse was obtained using the procedures described in the Statistical Analysis subsection below. The hue-saturation points and the tolerance ellipses were plotted to reveal differentiability between the two solutions. To demonstrate the capability to differentiate between different ratios of IRDye 680RD and IRDye 800CW, solutions of IRDye 680RD and IRDye 800CW in phosphate-buffered saline were produced at a concentration of 1 μM, and mixtures of the stock solutions were made at volumetric ratios of 10:0, 9:1, 7:3, 5:5, 3:7, 1:9, and 0:10. One sample of each mixture was collected. The lens (Canon) was adjusted to f/2.2 aperture, the sensor was configured with 200-ms exposure time, and the 780 nm and 665 nm illumination were each tuned to 20 mW/cm^2. Videos were captured of each sample and analyzed as previously described.

To demonstrate that excitation and emission spectra could be differentiated, a scene containing a mixture between excitation and emission spectra was assembled. The 780 nm illumination from a diode laser (BWF2-780, B&W Tek) was passed through a laser-shaping filter (FF01-769/41 Brightline bandpass filter, Semrock), a ground glass diffuser (DG05-120, Thorlabs), and a condenser lens (ACL25416U, Thorlabs) before incidence on a 50:50 beamsplitter (CM1-BS1, Thorlabs). One half of the total power was directed toward a condenser lens (ACL25416U, Thorlabs) and a ground glass diffuser (DG05-120, Thorlabs) such that this light could be projected outward; the other half was directed toward ICG in

deionized water stored in a microcentrifuge tube such that the light could generate fluorescence emission within or reflect off the microcentrifuge tube. The lens (Rokinon) and sensor were set up such that they viewed the projected excitation from the first optical path as well as the fluorescence emission and reflected excitation from the second optical path. A 50-frame color/near-infrared video of the scene was captured, and the temporally averaged near-infrared frame was transformed from a red-green-blue representation to a hue-saturation-value representation. Pixels were selected from two regions of interest that corresponded to pure excitation signal (collected over the projection from the first optical path) and pure emission signal (collected from the center of the vial in the second optical path); pixels were discarded if the value was too close to the noise floor or too close to saturation. The value and its intensity information were trimmed away, while the hue and saturation and their spectral information were retained; a binary linear support vector machine capable of predicting pure excitation signal or pure emission signal was then trained on the hues and saturations. Recognizing that the mixed excitation-emission signal should lie between pure excitation signal and pure emission signal and that the binary linear support vector machine's margin should lie between these two classes, a three-class classifier was derived that predicted mixed excitation-emission within the margin and either pure excitation or pure emission outside the margin. Pixels were then selected from another region of interest covering the entirety of the projection from the first optical path and the entirety of the vial in the second optical path. The value and its intensity information were trimmed away, while the hue and saturation and their spectral information were retained; the three-class classifier was then tested on the hues and saturations, labeling each pixel as mixed excitation-emission, pure excitation, or pure emission. Finally, a scatterplot was generated over the hue-saturation space using the pixels used to train the classifier, and an image was produced marking the pixel-wise classification from the pixels used to test the classifier.

Cell culture

For the animal procedures, prostate cancer cell lines PC3 and LNCaP were obtained from the American Type Culture Collection (ATCC) and used for experiments within 6 months of their receipt. Cell lines were commercially tested for pathogens, including mycoplasma, through the University of Illinois Department of Animal Resources and Charles River Research Animal Diagnostic Services. All cells were cultured under the recommended conditions in a 37 ˚C humidified incubator with 5% CO_2 atmosphere.

Animal experiments

A separate J:NU mouse with 2 PC3 tumors was used to explore the fluorescent signal's spatial heterogeneity; that mouse was handled similarly to those mice used for in vivo tumor imaging and ex vivo receiver operating characteristic analysis. The fluorescent signal within the tumors was computed by taking the value from the hue-saturation-value representation, while the normalized fluorescent signal within the tumors was computed by shifting/scaling the fluorescent signal such that the 1^{st} and 99^{th} percentile, computed across both tumors, equaled 0 and 1, respectively. For image display (done for both tumors at the same time), the normalized fluorescent signal was mapped to false color then composited over a color image; for histogram display (done for one tumor at a time), the pixels were counted for which the normalized fluorescent signal laid within a given bin, and the pixel count for each bin was divided by the total pixel count across all bins.

Three-dimensional reconstruction

To permit three-dimensional surface reconstruction, structured illumination was used. A near-infrared projector with a 900 nm lamp (EKB Technologies Ltd., Israel) and our imaging system with 900 nm sensitivity were synchronized such that a sinusoidal pattern could be projected onto an object by the near-infrared projector and captured at a high rate by our imaging system. After each frame, the projected pattern underwent a known phase shift, causing the captured pattern to experience an unknown phase shift associated with the object's topology; after 6 frames, a least-squares algorithm provided wrapped values for the unknown phase shift (*37*), and Goldstein's branch cut method provided unwrapped values for the unknown phase shift (*38*). Plugging in the phase shift at every pixel coordinate, a 12-term rational function provided the height at every pixel coordinate (*39*).

To facilitate calibration of the reconstruction model (i.e., the unknown coefficients in the 12-term rational function) and evaluation of the reconstruction error (i.e., between the predicted height and the actual height) a gauge object with known topology was utilized. The gauge object was designed with six ramps leading to six ledges and was fabricated in plastic via stereolithography (with tolerances of ±0.140 mm). Three ledges (at heights of 9.525 mm, 28.575 mm, and 41.275 mm) were designated for calibration of the reconstruction model, and three ledges (at heights of 15.875 mm, 22.225 mm, and 34.925 mm) were designated for evaluation of the reconstruction error. Sinusoidal fringes were projected onto the gauge object (with a period of ~14.50 mm along the ground plane) and were shifted across the gauge object (in discrete steps of 2.42 mm along the ground plane) as near-infrared images were captured (with ~5.78 pixels spanning 1 mm along the ground plane). The observed deformation of the sinusoidal fringes along the ramps and the known heights of the ledges were used to populate the reconstruction model via the Levenberg-Marquardt algorithm and to calculate the reconstruction error using the fitted model. The root-mean-square error and signal-to-noise ratio over multiple measurements were evaluated at each point on each evaluation step and were spatially averaged across each step and across all evaluation steps.

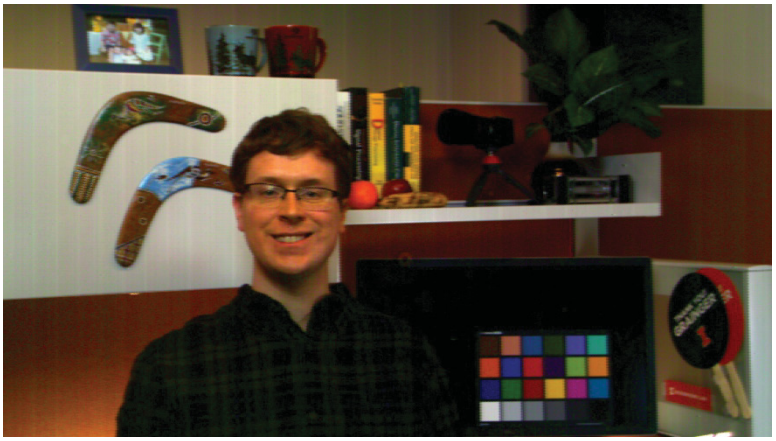

Fig. S1. Demonstration of the bio-inspired imaging system's color reproduction for natural scenes.

[Placeholder; See attached files.]

Video S1. Tumor detection in small-animal model of human prostate cancer showing overlaid information from color images of mouse, near-infrared images of fluorescently labeled tumors, and three-dimensional shape of mouse and tumors.

[Placeholder; See attached files.]

Video S2. Sentinel lymph node mapping in patients with breast cancer using near-infrared fluorescence from indocyanine green.

Table S1. Optical performance of state-of-the-art near-infrared fluorescence (NIRF) imaging systems compared with our bio-inspired NIRF imager.

Device	*Fluobeam 800*	*PDE Neo*	*SPY Elite*	*Iridium*	*Spectrum*	*Solaris*	*Bio-inspired*	*Bio-inspired*
Company or university	Fluoptics	Hamamatsu	Novadaq	VisionSense	Quest	PerkinElmer	Illinois & WUSTL	Illinois
Reference	(*8*)	(*8*)	(*8*)	(*8*)	(*8*)	(*8*)	(*10*)	This article
FDA/EMA approved	Yes	Yes	Yes	Yes	Yes	No	No	No
Instrument architecture	Single sensor	Single sensor w/ swappable filters	Single sensor	One–two sensors w/ dual optical paths	Three sensors w/ trichroic prism	Two sensors w/ beamsplitter	Single sensor	Single sensor
Sensor type	CCD	CCD	CCD	CCD	CCD	sCMOS	CMOS	CMOS
Permitted surgical lighting	Dim	Dim	Dim	Dim	On	On	On	On
Supports RGB imaging	No	Yes	No	Yes	Yes	Yes	Yes	Yes
Supports real-time RGB/NIR overlay	No	No	No	Yes	Yes	Yes	Yes	Yes
Number of NIR channels	1	1	1	1	2	2	1	3
Number of simultaneous NIR channels	1	1	1	1	2	1	1	3
Excitation source	750-nm laser	760-nm LEDs	805-nm laser	805-nm laser	LEDs	Filtered LEDs	780-nm laser	665- and 780-nm lasers
Fluorescence detection limit	~5 nM[a]	~15 nM[a]	~5 nM[a]	~50 pM[a]	~10 nM[a]	~1 nM[a]	100 pM[b]	201 pM[c]
Resolution (px)	720 x 576	640 x 480	1024 x 768	960 x 720	NS	1024 x 1024	1280 x 720	1280 x 720
Bit depth	8	8	8	12	12	16	14	14
Maximum frame rate (FPS)	25	NS	NS	NS	20	100	40	30

[a] IRDye 800CW fluorescence marker was used for the experiment. Limit of detection was computed by plotting the average signal versus concentration and finding the intersection between lines fit to low concentrations and high concentrations.

[b] Indocyanine green fluorescence marker was used for the experiment. Limit of detection was computed by modeling the average pixel's response to each concentration and finding the concentration where the average pixel fails to detect the concentration.

[c] Indocyanine green fluorescence marker was used for the experiment. Limit of detection was computed by plotting the average signal versus concentration and finding the intersection between lines fit to low concentrations and high concentrations.

Abbreviations: CCD, charge-coupled device; CMOS, complementary metal-oxide semiconductor; EMA, European Medicines Administration; FDA, U.S. Food and Drug Administration; FPS, frames per second; Illinois, University of Illinois at Urbana-Champaign; NIR, near infrared; NS, not stated; RGB, red, green, blue; sCMOS, scientific complementary metal-oxide semiconductor; WUSTL, Washington University in St. Louis.

Table S2. Optical performance of pixelated near-infrared fluorescence (NIRF) imaging systems compared with our bio-inspired NIRF imager.

Device	*Piranha4*	*OV4682*	*RGB-NIR*	*Bio-inspired*	*Bio-inspired*
Company or university	Teledyne DALSA	OmniVision	Univ. of Arizona	Illinois & WUSTL	Illinois
Reference	Company website	Company website	(*40*), (*41*)	(*10*)	This article
Sensor type	Line scan CMOS	Full frame CMOS	Full frame CMOS	Full frame CMOS	Full frame CMOS
Filter material	NS	NS	Polymer	Dielectric interference filters	Dielectric interference filters
Supports multiple integration times	Yes	No	No	Yes	Yes
Supports RGB imaging	Yes	Yes	Yes	Yes	Yes
Number of NIR channels	1	1	1	1	3
Maximum OD of NIR filters in visible spectrum	NS	NS	NS	~2.5	~6.2
Maximum OD of RGB filters in NIR spectrum	NS	NS	NS	~2.1	~3.3
ICG detection limit	NS	NS	130 nM[a]	100 pM[a]	201 pM[b]
Resolution (px)	2048 x 4	2688 x 1520	2304 x 1296	1280 x 720	1280 x 720
Bit depth	12	10	12	14	14
Maximum frame rate (FPS)	NA	90	60	40	30

[a] Limit of detection was computed by modeling the average pixel's response to each concentration and finding the concentration where the average pixel fails to detect the concentration.
[b] Limit of detection was computed by plotting the average signal versus concentration and finding the intersection between lines fit to low concentrations and high concentrations.

Abbreviations: CMOS, complementary metal-oxide semiconductor; FPS, frames per second; ICG, indocyanine green; Illinois, University of Illinois at Urbana-Champaign; NIR, near infrared; NA, not applicable; NS, not stated; OD, optical density; RGB, red, green, blue; WUSTL, Washington University in St. Louis.

Data S1. Classification scores from the receiver operating characteristic (ROC) analysis of tumor detection in a small-animal model of human prostate cancer. Tissue-level data is provided for the ROC analysis presented in Fig. 4 and described in the text. This data breaks down the number and type of tissues collected from each mouse as well as the classification scores assigned via quadratic discriminant analysis for the three dye administration strategies (IRDye 680RD EGF and IRDye 800CW 2-DG together, IRDye 680RD EGF alone, and IRDye 800CW 2-DG alone). Quadratic discriminant analysis modeled the conditional probability distribution $P(X|y = k)$ where X was the data (a 2-tuple of the hue and saturation) and y was the class (either healthy tissue or tumor tissue), and Bayes' rule provided the posterior probability distribution $P(y = k|X)$. The classification scores were then taken as the decision function $\log P(y = \text{Tumor}|X) - \log P(y = \text{Healthy}|X)$ such that more positive classification scores corresponded to tumor tissues and more negative classification scores to healthy tissues.

[Placeholder; See attached files.]

Data S2. Classification scores from the receiver operating characteristic (ROC) analysis of sentinel lymph node detection in patients with breast cancer. Tissue-level data is provided for the ROC analysis presented in Table 1 and described in the text. This data breaks down the number and type of tissues collected from each patient as well as the classification scores assigned via quadratic discriminant analysis for the administration of indocyanine green alone. Quadratic discriminant analysis modeled the conditional probability distribution $P(X|y = k)$ where X was the data (a 3-tuple of the hue, saturation, and value) and y was the class (either lymph node or nonlymphatic structure), and Bayes' rule provided the posterior probability distribution $P(y = k|X)$. The classification scores were then taken as the decision function $\log P(y = \text{Lymph}|X) - \log P(y = \text{Nonlymphatic}|X)$ such that more positive classification scores corresponded to lymph nodes and more negative classification scores to nonlymphatic structures.

[Placeholder; See attached files.]