

Journal of Biomedical Optics

SPIEDigitalLibrary.org/jbo

Toward real-time virtual biopsy of oral lesions using confocal laser endomicroscopy interfaced with embedded computing

Patricia S. P. Thong
Stephanus S. Tandjung
Muhammad Mobeen Movania
Wei-Ming Chiew
Malini Olivo
Ramaswamy Bhuvaneshwari
Hock-Soon Seah
Feng Lin
Kemao Qian
Khee-Chee Soo

Toward real-time virtual biopsy of oral lesions using confocal laser endomicroscopy interfaced with embedded computing

Patricia S. P. Thong,^a Stephanus S. Tandjung,^b Muhammad Mobeen Movania,^b Wei-Ming Chiew,^b Malini Olivo,^{a,c} Ramaswamy Bhuvaneshwari,^a Hock-Soon Seah,^b Feng Lin,^b Kemao Qian,^b and Khee-Chee Soo^a

^aNational Cancer Centre, Division of Medical Sciences, Singapore

^bNanyang Technological University, School of Computer Engineering, Singapore

^cBiomedical Sciences Institutes, Singapore Bioimaging Consortium, Singapore

Abstract. Oral lesions are conventionally diagnosed using white light endoscopy and histopathology. This can pose a challenge because the lesions may be difficult to visualise under white light illumination. Confocal laser endomicroscopy can be used for confocal fluorescence imaging of surface and subsurface cellular and tissue structures. To move toward real-time “virtual” biopsy of oral lesions, we interfaced an embedded computing system to a confocal laser endomicroscope to achieve a prototype three-dimensional (3-D) fluorescence imaging system. A field-programmable gated array computing platform was programmed to enable synchronization of cross-sectional image grabbing and Z-depth scanning, automate the acquisition of confocal image stacks and perform volume rendering. Fluorescence imaging of the human and murine oral cavities was carried out using the fluorescent dyes fluorescein sodium and hypericin. Volume rendering of cellular and tissue structures from the oral cavity demonstrate the potential of the system for 3-D fluorescence visualization of the oral cavity in real-time. We aim toward achieving a real-time virtual biopsy technique that can complement current diagnostic techniques and aid in targeted biopsy for better clinical outcomes. © 2012 Society of Photo-Optical Instrumentation Engineers (SPIE). [DOI: 10.1117/1.JBO.17.5.056009]

Keywords: confocal laser endomicroscopy; 3-D fluorescence imaging; oral cancer; volume rendering; embedded computing; virtual biopsy.

Paper 11663 received Nov. 10, 2011; revised manuscript received Mar. 12, 2012; accepted for publication Mar. 27, 2012; published online May 4, 2012.

1 Introduction

1.1 Oral Cancer Diagnosis

Oral cancer is becoming one of the most common types of cancer worldwide, particularly in developing countries where this form of malignancy ranks as the 7th and 9th most common cancers in males and females, respectively.^{1,2} Risk factors include smoking, drinking alcohol, using smokeless tobacco products and infection with the human papillomavirus.² Recent advances in cancer treatment techniques have not significantly improved the survival rate for oral cancer, which remains at about 50%, due to it often not being diagnosed until a relatively advanced stage. Early diagnosis is therefore crucial for a good treatment outcome. Currently, lesions of the oral cavity are diagnosed using white light endoscopy followed by histopathological examination of biopsy samples. Endoscopic examinations may also be extended to the larynx and esophagus to check for other possible lesions.³

There are some challenges in the existing conventional techniques for oral cancer diagnosis. Firstly, early oral lesions are often flat, making it difficult to distinguish between benign and malignant lesions under white light illumination. Secondly, histopathology is time-consuming and requires specialized skills and experienced, trained personnel. Thirdly, while biopsies are

generally safe, they carry a small risk of patient complications. Finally, it might be difficult to determine the margins of oral lesions, and multiple biopsies are often required to ensure a clear margin during surgical procedures. Therefore there is a need to develop minimally invasive “virtual” biopsy techniques that can provide accurate and real-time diagnosis of oral lesions in the clinic, which will help to target biopsy to abnormal regions, and thus reduce the number of biopsies needed to make a diagnosis. One emerging optical technique that has shown potential as a tool for virtual biopsy and guided biopsy procedures is confocal laser endomicroscopy.

1.2 Confocal Laser Endomicroscopy

Confocal laser endomicroscopy (CLE) is an endoscopic technique that complements conventional endoscopy by enabling *in vivo* imaging of tissue and cellular structures at microscopic resolution (about 1 μm laterally).⁴⁻⁶ Through cellular, structural and molecular imaging, information can be extracted not only from the surface but also from deeper subsurface layers, thus offering a tool for optical or virtual biopsy in the clinic.⁷⁻⁹ Currently there are two commercially available confocal laser endomicroscope systems. In the endoscope-based system (developed by Optiscan Pty, Ltd., Victoria, Australia), the focusing and scanning mechanisms are miniaturized into the distal tip of an endoscope.¹⁰ In the probe-based system (developed by

Address all correspondence to: Patricia S. P. Thong, National Cancer Centre, Division of Medical Sciences, Singapore. Tel.: 65 63266192; E-mail: nmstsp@nccs.com.sg

Mauna Kea Technologies, Paris, France), the scanning is achieved at the proximal end of a fiber optic probe.¹¹

Recent studies demonstrate the potential of CLE as a clinical tool for surveillance and diagnosis of several cancerous and pre-cancerous conditions. These include clinical diagnostic applications in the airways,^{12–14} upper and lower gastrointestinal tracts,^{15–23} bladder neoplasia,^{24,25} cervical intraepithelial neoplasia^{26,27} and the oral and oropharynx.^{28–31} Results have been promising. For example, a recent study by Xie et al. reported that CLE could detect adenomas in colonic polyps at a sensitivity and specificity of 93.9% and 95.9%, respectively, when compared to histopathology results.²² In another recently completed multicenter randomized controlled trial for detection of high-grade dysplasia and early carcinoma in Barrett's esophagus, the combined use of probe-based CLE (pCLE) and high-definition white-light endoscopy (HD-WLE) yielded a favourable sensitivity and specificity of 68.3% and 87.8% compared to 34.2% and 92.7%, respectively, for the use of HD-WLE alone.²³

Confocal laser endomicroscopy has also been used as an aid for targeted biopsy procedures to improve effectiveness and reduce the number of biopsies needed. For example, Gunther et al. reported that targeted biopsies guided by either chromoendoscopy or CLE resulted in higher detection rates of intraepithelial neoplasia in the surveillance of inflammatory bowel disease.³² Confocal endomicroscopy also holds promise for image-guided surgery by aiding the assessment of lesion margins during or following surgical procedures.³³

With the use of suitable fluorescent dyes, confocal fluorescence imaging can be carried out. Fluorescein sodium is commonly used, as it is safe for human use³⁴ and spectrally matched to the 488-nm excitation wavelengths of many confocal endomicroscope systems. However, fluorescein sodium is non-specifically absorbed by all cells and thus may result in false positives during diagnostic imaging. Other possible fluorescent dyes include hypericin, a photosensitizer extracted from the plant commonly known as St John's wort, and the fluorescent pre-cursor, 5-aminolevulinic acid (5-ALA) that metabolises into the fluorescent compound protoporphyrin IX (PpIX). Hypericin and 5-ALA may be more selectively taken up by abnormal cells, and may thus enable fluorescence diagnostic imaging with higher specificity.^{35,36}

1.3 Toward Real-Time Virtual Biopsy of Oral Lesions

We have previously described the use of endoscope-based CLE for confocal fluorescence diagnostic imaging of the human and murine oral cavities.^{28–30} A prototype confocal endomicroscope with a rigid, hand-held probe was used with 5-ALA, fluorescein sodium and hypericin as contrast agents. Hypericin was used in mouse models while 5-ALA and fluorescein sodium were used in rat models.^{28,29} Fluorescence images of the normal rat tongue were compared to those from carcinogen-induced models of oral squamous cell carcinoma (SCC). Images of the normal rat tongue showed regularly arranged filiform papillae, while the architecture of the SCC rat tongue showed up as more irregular. In pilot clinical studies, 5-ALA was topically applied to the oral cavities of healthy volunteers and an oral SCC patient to compare ALA-induced PpIX fluorescence images from the normal and SCC human tongue.^{28,30} The results demonstrated the capability of the confocal endomicroscope to differentiate between the normal and SCC tongue by morphology and tissue architecture, leading to its potential to be used as a minimally invasive

technique for oral cancer diagnosis. This was in agreement with Haxel et al. who reported that CLE could be used for the diagnosis of malignancy in the human oral and oropharyngeal cavity by means of the altered tissue architecture and irregularity in blood vessels.³¹

Conventional bench-top confocal microscopes are equipped with hardware and software to acquire and render 3-D confocal image stacks. However, these systems can only be used for *ex vivo* imaging of tissue sections. On the other hand, confocal endomicroscopes enable *in vivo* imaging but capture and display images from one single focal plane at a time. Real-time 3-D image registration, voxel-based processing and rendering software are unavailable, making it difficult to recognize 3-D structures. In order to bridge this gap and move toward a real-time "virtual" biopsy technique, we developed a 3-D fluorescence imaging system by interfacing a confocal laser endomicroscope to an embedded computing system.^{37–40} We used a high performance multimedia field-programmable gated array (FPGA) board as a reconfigurable platform.³⁹ The FPGA board has the required interfaces, such as dual video support for Digital Video Interface (DVI), Thin-Film Transistor (TFT) flat panel display, Personal System/2 (PS/2) keyboard and mouse ports, a four-line-by-16-character Liquid Crystal Display (LCD) display, eight white user-programmable Light-emitting Diodes (LEDs), and general input/output (I/O) pins. Other peripherals, such as a keyboard, can also be used for user interface. In this study, we describe the development of the endomicroscope-embedded computing system for 3-D fluorescence imaging of the oral cavity.

2 Materials and Methods

2.1 Endomicroscope-Embedded Computing System

An endoscope-based confocal laser endomicroscope system (FIVE1, Optiscan, Australia) was fitted with a short, hand-held rigid probe (model RBK6315A) that is suitable for imaging the oral cavity. The excitation source is a 488-nm laser coupled into a single optical fiber acting as both a point source and as a point detection pinhole for confocal imaging.⁴¹ The lateral resolution is about 0.7 μm . The rigid probe houses the miniaturized components of the x-y scanning mechanism, allowing images to be captured with a field of view of 475 $\mu\text{m} \times 475 \mu\text{m}$. The laser power can be adjusted between 0 and a maximum of 1000 μW at the distal tip of the probe in contact with the tissue sample. Fluorescence signals are collected via a 505- to 750-nm emission filter. Under the normal mode of operation, Z-depth sectioning can be achieved via a footswitch that controls the imaging depth from the surface down to deeper planes with a nominal step size of about 4 μm between consecutive slices. In biological samples, the maximum imaging depth is about 250 μm below the surface, depending on the tissue optical properties.

Figure 1 shows the schematic diagram of the confocal endomicroscope interfaced with an embedded computing system based on an RC340 FPGA (Mentor Graphics Corporation, USA). The embedded FPGA platform functions as the main board and controls the endomicroscope through a Z-depth control circuit called the daughter board.³⁹ The fluorescence signal from the hand-held probe is converted by the endomicroscope into digital data that is displayed on the monitor and sent to the RC340 main board. The main board captures the images for real-time image processing and displays the processed data via the DVI interface. Under this mode of operation, depth control

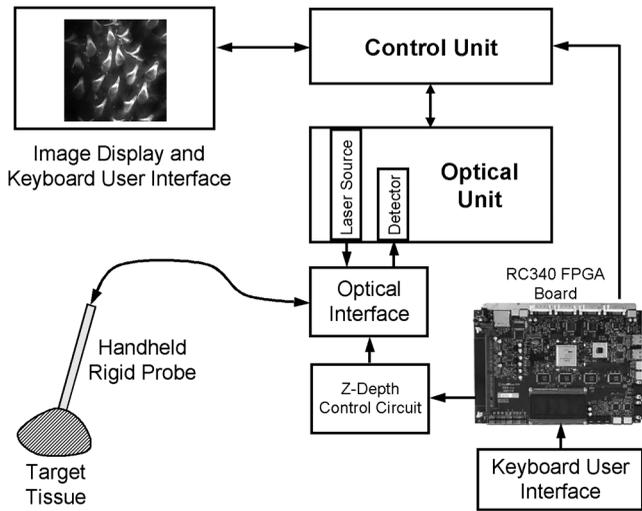


Fig. 1 Schematic diagram showing a confocal laser endomicroscope interfaced with an embedded computing system based on an RC340 field-programmable gated array (FPGA) board and a Z-depth control circuit designed for automated acquisition of confocal image stacks upon initiation by the operator via a keyboard user interface.

via the footswitch is disabled. The RC340 board automatically controls the endomicroscope system to capture confocal image stacks (termed datasets) from the surface to deeper focal planes in the target tissue upon initiation by the operator via a keyboard-user interface.

2.2 Volume-Rendering using GPU

For prototyping, the 3-D visualization of confocal image stacks was first developed on a PC equipped with a graphics processing unit (GPU). We utilized 3-D texture slicing⁴²⁻⁴⁶ to generate high quality volume renderings of the 3-D datasets. Our volume-rendering process starts with filtering the dataset to remove noise and other high frequency artifacts, after which a 3-D texture is generated to store the processed image stack. This stack is then classified using a transfer function.⁴⁷⁻⁴⁹ Finally, the volume is sampled using view-aligned proxy geometry to generate the 3-D texture slicing. These slices are then alpha-blended in back-to-front order using hardware-based alpha blending.^{50,51} The overall volume-rendering process is illustrated in Fig. 2.

The classification of datasets is critical to make sense of the volume-rendering. The process typically assigns a mapping ($R \rightarrow R^3$) from a scalar value to a color value, implemented as a look-up table. We created a customized widget (Fig. 3) for assignment of the transfer function.^{52,53} This widget generates piecewise linear transfer functions and enables the user to create custom mappings for each individual colour channel, alpha, red, green and blue (ARGB), by plotting control points. Each

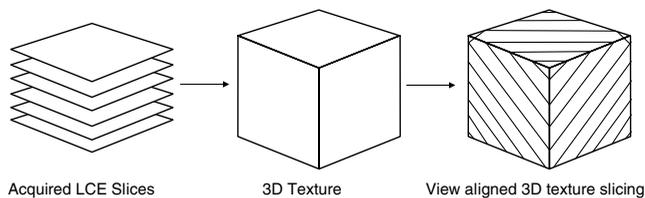


Fig. 2 Illustration of the overall 3-D volume rendering process, beginning with a stack of confocal images that forms the dataset and resulting in a view-aligned texture.

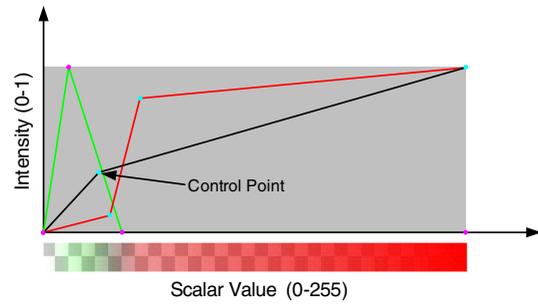


Fig. 3 The customized widget used for assignment of the transfer function during volume rendering using a graphics processing unit (GPU).

channel is rendered in its own colour to help distinguish it. The alpha channel curve is coloured black. The X-axis of the widget refers to the scalar value of the intensity, which ranges from 0 to 255 in our datasets. The Y-value stores the intensity (0-1) as an indication of how much of the value is being assigned. To assist the user in visualizing the output generated from the combined function, a colour preview strip is rendered under the curve while the background is rendered to portray the color's transparency. The higher the background is, the more transparent the colour. In addition, a specialized interpolation scheme is also developed to maintain smooth interpolation across acquired slices.⁵³

2.3 Volume Rendering using FPGA

In our embedded computing solution, we also programmed the RC340 FPGA board to support real-time 3-D visualization while the image stack is being acquired. The rendering of incrementally acquired slices will be produced and displayed “on-the-fly” through the video output on the FPGA board. The volume ray-casting technique⁵⁴ is employed in our system to generate two-dimensional (2-D) projection output images viewed from arbitrary angles in 3-D. In this algorithm, imaginary rays are projected towards the dataset and sampling points along each ray are accumulated to produce the output, which is illustrated in Fig. 4. Each pixel on the image plane is initialized as a ray origin. Rays cast from each origin traverses through the confocal dataset volume. Consistent points on each ray are sampled within the

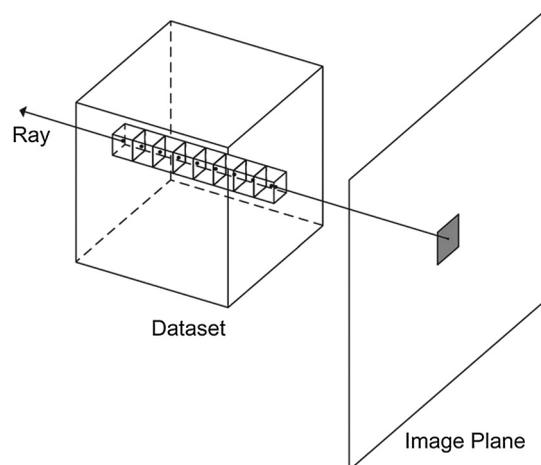


Fig. 4 Illustration of the ray-casting algorithm in which an imaginary ray is cast from each pixel onto the 2-D image plane toward the dataset.

dataset, and the sampled points are integrated using a pre-selected ray function to generate the final output pixel value. The embedded computing module used in the visualization process utilizes hardware parallelization features to reduce the computational time required for calculations, thus providing fast, high-quality, real-time volume-rendering of datasets with resolutions of up to 1024×1024 pixels per frame. While imaging is being performed using the endomicroscope-embedded computing platform, the dataset can be visualized in real-time.

2.4 Fluorescence 3-D Imaging of the Murine Oral Cavity

Fluorescein sodium and hypericin were used as fluorescent agents in 6- to 8-week-old Balb/c nude mouse models. Fluorescein sodium (Novartis Pharma AG, Switzerland) was freshly prepared as a 1% solution while hypericin (Molecular Probes, USA) was prepared as a 0.004% solution. Topical application to the murine oral cavity was carried out by the insertion of cotton buds soaked in the freshly prepared fluorescein or hypericin solutions for 5 to 10 min. Following an incubation period of 30 min, the mice were sacrificed and the tongues excised for imaging. Excised tissue of the mouse tongue was sectioned and processed for haematoxylin and eosin (H&E) staining.

2.5 Fluorescence 3-D Imaging of the Human Oral Cavity

A pilot clinical study to test the prototype system for fluorescence 3-D imaging was approved by the Centralised Institutional Review Board of the Singapore Health Services Pte, Ltd. Four healthy volunteers with no history of oral malignancies and two patients who were undergoing surgical procedures for lesions in the head and neck were recruited for the study following informed consent. Topically applied hypericin was used as the fluorescent agent. Additionally, fluorescein was also used in the volunteer group only to compare the results from topically applied hypericin and fluorescein.

Hypericin (Molecular Probes, USA) was freshly prepared in 1% serum albumin in phosphate buffered saline (PBS) and diluted in saline to give an $8 \mu\text{M}$ instillation solution. The solution was filtered and topically administered to both the volunteer and patient groups by oral rinsing using 100 ml of the solution over 30 min. After a further incubation period of at least 45 min, fluorescence 3-D imaging was carried out.

Fluorescein (Novartis Pharma AG, Switzerland) was freshly diluted in PBS to obtain a 0.1% solution. The solution was filtered and topically administered to the volunteer group only by oral rinsing using 100 ml of solution over 30 min.

After a further incubation period of at least 45 min, fluorescence 3-D imaging was carried out.

3 Results

3.1 Automated Acquisition of Confocal Image Datasets

We have developed a prototype 3-D fluorescence imaging system comprising a confocal endomicroscope interfaced to an FPGA-based embedded computing system. In the normal mode of operation, an operator manually controls the imaging depth via a footswitch to collect individual confocal images at the desired depths. In our endomicroscope-embedded computing system, the image acquisition control is replaced by the daughter board and circuitry programmed on the FPGA board. In place of the footswitch, we have designed an interface with main components that act as relays to isolate the endomicroscope system from the FPGA board electronically. The controller circuitry replaces operator tasks from basic to higher level ones. With this controller, the user only needs to push one button for the FPGA board to capture a confocal image stack (dataset) automatically from the initial imaging depth and sequentially deeper focal planes until the desired imaging depth has been reached and the user initiates a stop signal. The current system is capable of acquiring one image per 1.4 sec, the fastest attainable rate limited by endomicroscope hardware. The automated acquisition of image stacks minimizes the data acquisition time and thus effectively minimizes the chances of movement in between consecutive images.

3.2 Rendering of Murine Datasets using GPU

Pilot testing of the prototype endomicroscope-embedded computing system was carried out using mouse models. Rendering of image stacks acquired from the murine tongue was carried out using a GPU program. Figure 5 shows a series of rendering results with each series having twice as many texture slices. This figure illustrates how adding more slices improves the rendering quality considerably.

Figure 6(a)–6(f) show six consecutive images from a confocal image stack acquired from the dorsal surface of the mouse tongue following topical application of fluorescein sodium. The arrows in (a) indicate filiform papillae on the surface of the mouse tongue. The images show that while bright fluorescence is observed at the surface (a), there is a gradual reduction in fluorescence intensity as the imaging depth increases (b–f). Figure 6(g) shows the 3-D volume rendering result obtained using a GPU-based program and displayed in false colours. The conical shapes of the filiform papillae are well rendered.

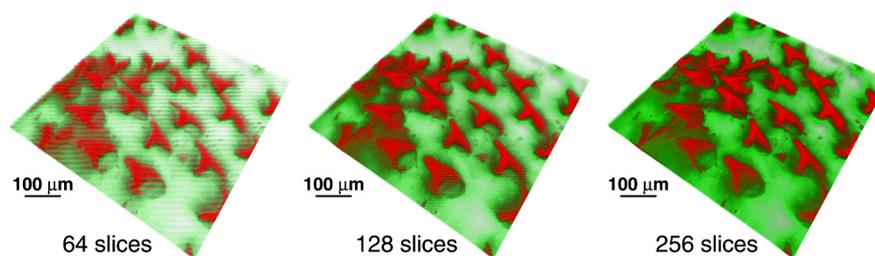


Fig. 5 A murine tongue CLE dataset rendered using 3-D texture slicing. The number of slices for each image shows the total number of texture slices used for the rendering, with more slices improving the rendering quality.

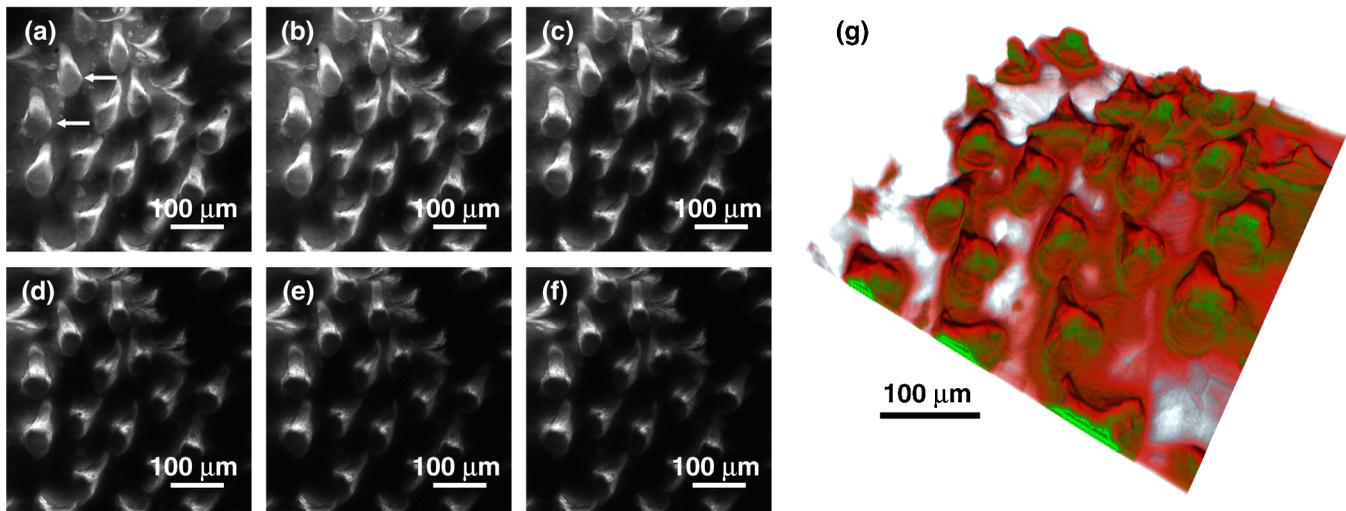


Fig. 6 (a–f) Six consecutive images from a dataset acquired from the dorsal surface of a mouse tongue after topical administration of fluorescein sodium. The arrows in (a) indicate filiform papillae on the tongue. Image (g) shows the 3-D rendering result (displayed in false colours) obtained using a GPU-based program. The conical shapes of the filiform papillae are well rendered and the 3-D image provides topographical information that is not easily visualized from the 2-D images.

Figure 7 shows the H&E stained image of a cross-section of a mouse tongue, showing filiform papillae on the surface (arrows) similar to those seen in the confocal images. Compared to the individual confocal images captured from single focal planes and the H&E stained cross-sectional image, the 3-D image provides additional topographical information that is not easily visualized from the 2-D images.

Figure 8 shows the rendering results for a confocal image stack acquired from the dorsal surface of a mouse tongue following topical administration of hypericin. In addition to the normal composite rendering with classification (a), we also used an omni-directional “light” source to illuminate surface details that are otherwise less visible (b).

3.3 Rendering of Human Datasets using GPU

Confocal fluorescence image stacks of various sites in the human oral cavity were acquired *in vivo* from healthy volunteers following topical administration of fluorescein sodium or hypericin. The sites investigated include the dorsal surface of

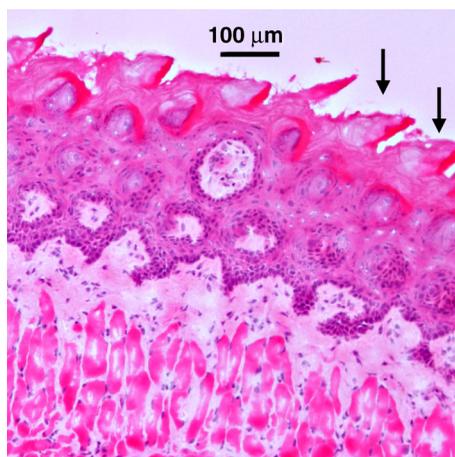


Fig. 7 H&E-stained image of a cross-section of a mouse tongue captured using a 20x objective lens under a light microscope. The arrows indicate filiform papillae on the surface of the mouse tongue.

the tongue, base of tongue, floor of mouth, the buccal mucosa and the lip. Real-time volume-rendering of the acquired image stacks was achieved using a GPU-based program. In Fig. 9(a) and (b) show confocal images from the same dataset that was acquired from the dorsal surface of the human tongue following topical administration of fluorescein sodium. The image in Fig. 9(a) was captured at the surface and Fig. 9(b) image from a focal plane approximately 30 μm below the surface. Filiform papillae [solid arrows in (a)] and cellular structures [dotted arrows in (b)] can be seen in the individual images. The GPU volume-rendering result of the entire dataset, displayed in false colours in (c), show the depth relation of the filiform papillae with respect to the cellular structures below the surface. Such information on depth relation is not easily visualized from the individual 2-D confocal images.

In Fig. 10(a) and (b) show confocal images from the same dataset that was obtained from the buccal mucosa of a healthy volunteer following topical application of fluorescein sodium. Image (a) was captured from the surface and image (b) from a deeper layer approximately 30 μm below the surface. Cellular structures at the surface [solid arrow in (a)] and below the surface [dotted arrow (in (b))] are observable from the individual images. The GPU volume-rendering result of the entire dataset, displayed in false colours in (c), shows the depth relation between the cellular structures from different focal planes.

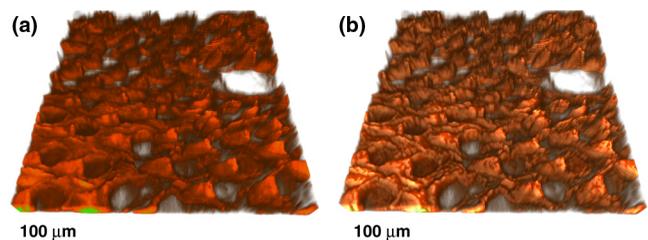


Fig. 8 GPU volume rendering results for a confocal image stack acquired from the dorsal surface of a mouse tongue following topical administration of hypericin. The composite rendering result without “lighting” (a) and the rendering result with “lighting” (b) demonstrate that adding “light” illuminates more surface details.

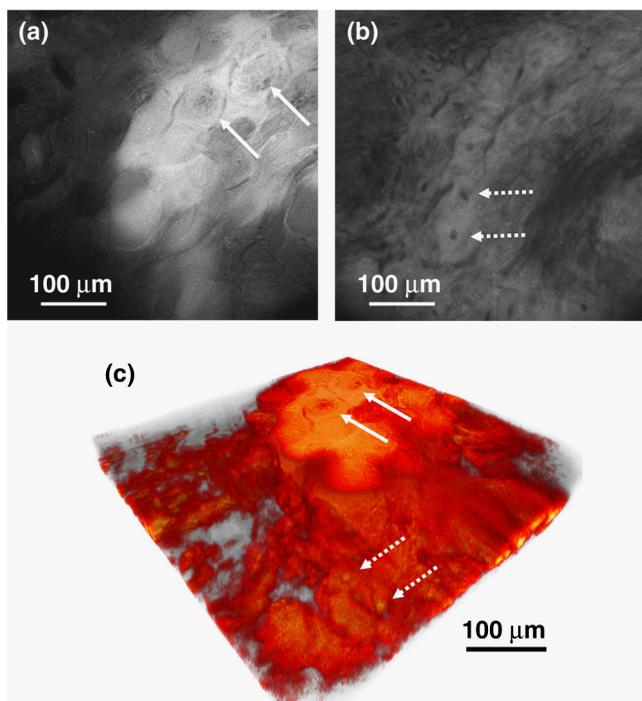


Fig. 9 Confocal images acquired from the dorsal surface of the human tongue following topical administration of fluorescein sodium. Image (a) was captured from the surface and image (b) approximately $30\ \mu\text{m}$ below the surface in the same dataset. Filiform papillae (solid arrows) and cellular structures (dotted arrows) can be observed from the individual images, while the GPU volume rendering result of the entire dataset, displayed in false colors (c), shows the depth relation of the filiform papillae to the cellular structures.

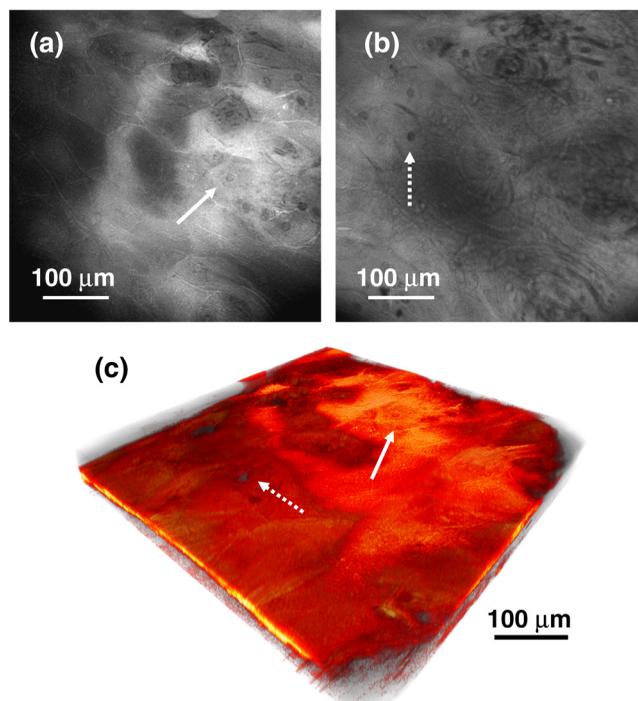


Fig. 10 Confocal images from the same dataset that was acquired from the human buccal mucosa following topical administration of fluorescein sodium. Image (a) was captured at the surface and image (b) approximately $30\ \mu\text{m}$ below the surface. Cellular structures at the surface (solid arrow) and in a deeper focal plane (dotted arrow) can be observed in the individual images. The GPU-volume-rendering result of the entire dataset, displayed in false colors in (c), shows the depth relation between the cellular structures from different focal planes.

Figure 11 shows images acquired from the human buccal mucosa following topical administration of hypericin. Image (a) was captured at the surface and image (b) from a focal plane approximately $15\ \mu\text{m}$ below the surface. These images show cellular structures at the surface [solid arrow in (a)] and below the surface [dotted arrow in (b)], while the GPU volume-rendering result of the entire dataset is displayed in false colours in (c). It is noted that the volume rendering result from the hypericin dataset is shallower compared to the results from fluorescein datasets. This is due to weaker fluorescence signals from hypericin compared to that from fluorescein and the limited maximum imaging depth reached in this dataset.

Pilot testing of the endomicroscope-embedded computing system in a clinical setting was carried out on two patients who were undergoing surgical procedures for lesions in the head and neck. Hypericin was topically applied prior to both *in vivo* and *ex vivo* imaging using the prototype imaging system. As the initial testing yielded weak fluorescence signals and noisy images, the results are not shown here. Further improvements are being made to the system to achieve better performance in a clinical setting.

3.4 Rendering of Image Stacks using Embedded FPGA

The FPGA board in our endomicroscope-embedded computing system was also programmed to support real-time 3-D visualization while the image stack is being acquired. The rendering of incrementally acquired slices will be produced and

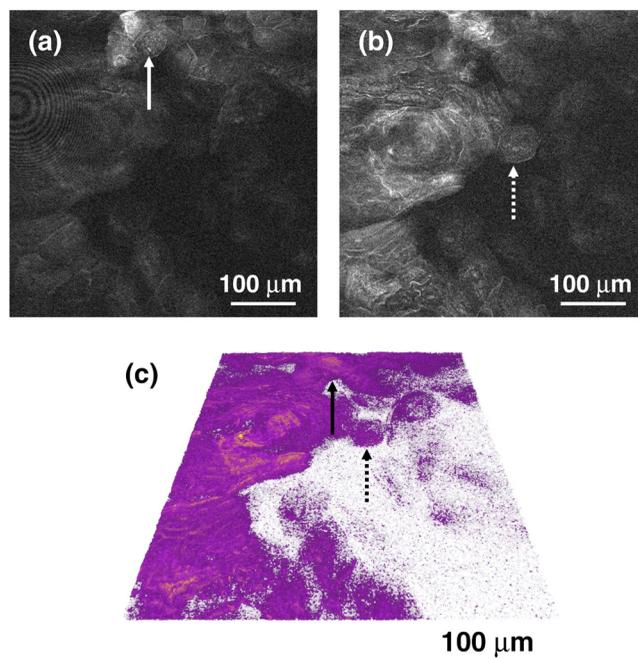


Fig. 11 Confocal images acquired from the human buccal mucosa following topical administration of hypericin. Image (a) was captured at the surface and image (b) approximately $15\ \mu\text{m}$ below the surface in the same dataset. Cellular structures at the surface (solid arrow) and in a deeper layer (dotted arrow) are observed in the individual images. The GPU volume rendering result of the entire dataset, displayed in false colors in (c), is shallow due to the maximum imaging depth reached in this dataset.

displayed “on-the-fly” through the video output on the FPGA monitor. The rendering pipeline developed on the embedded FPGA system is capable of rendering datasets while they are being acquired. The current rendering module deals with gray scale pixel values from the confocal endomicroscope. Shading effects are omitted in the current stage to simplify the rendering pipeline. Upon retrieval of an image slice, it is stored in the

on-board synchronous Dynamic Random-Access Memory (SDRAM) module. The maximum capacity of the SDRAM is 256 MB, which is sufficient for storing up to 256 slices with a resolution of 1024×1024 pixels using direct address indexing. This capacity is sufficient for our typical CLE datasets. Subsequently, the output pixel values from rendering pipeline are stored in a frame buffer driven by the display module.

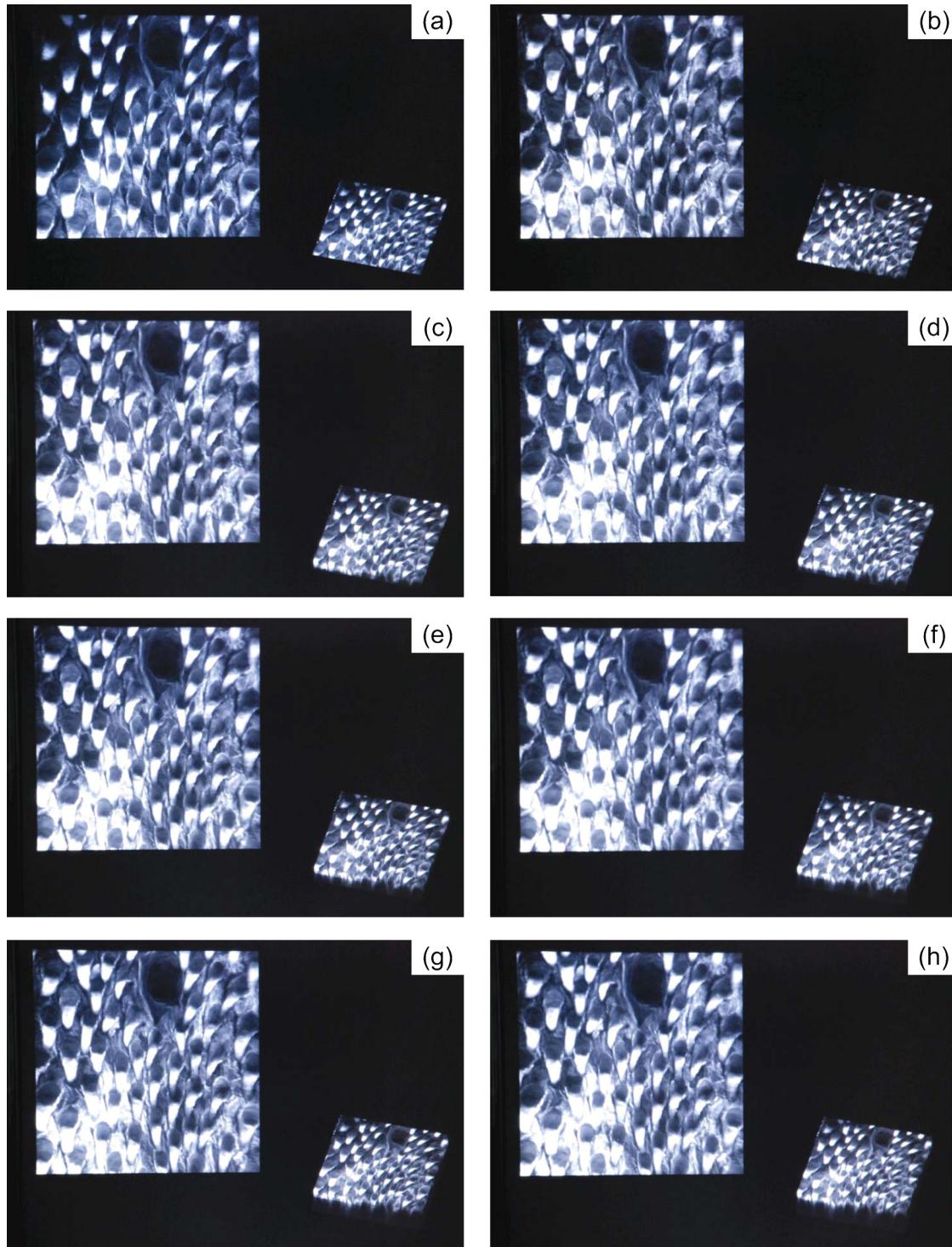


Fig. 12 Real-time volume rendering of a mouse tongue dataset acquired following topical application of hypericin. The results were obtained under different thickness: (a) 2 slices, (b) 10 slices, (c) 15 slices, (d) 20 slices, (e) 25 slices, (f) 30 slices, (g) 35 slices and (h) 38 slices.

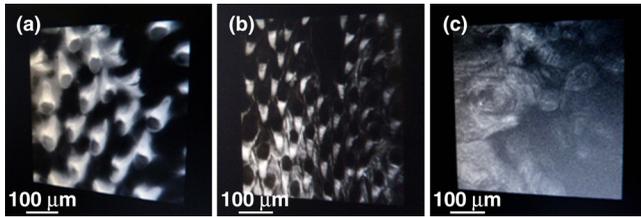


Fig. 13 FPGA- volume-rendering results of mouse and human oral cavity datasets. The conical shapes of filiform papillae are easily distinguishable on the mouse tongue images acquired following topical application of fluorescein sodium (a) and hypericin (b). The rendering result of a dataset acquired from the human buccal mucosa following topical application of hypericin is shown in (c).

Figure 12 shows consecutive rendering of the hypericin mouse tongue dataset while new slices are being captured. Rendering is refreshed at every new slice, and a growth in thickness can be observed as more slices are acquired. The renderings are from an orthogonal perspective with a perpendicular viewing angle towards the slice. The datasets do not have a fixed size, and as more slices are obtained, more storage space is required. To provide illustrative results, the rendering is performed with the maximum intensity projection scheme, where the sampling point with the highest intensity across each ray is used as the output pixel.

Figure 13 shows screen captures of different datasets rendered using the endomicroscope-embedded computing system. Figure 13(a) shows the result from a fluorescein sodium mouse tongue dataset, which corresponds to the GPU-based rendering result in Fig. 6. Figure 13(b) is the result from a hypericin mouse tongue dataset. The conical shapes of the filiform papillae are easily distinguishable in the outputs in (a) and (b). Figure 13(c) shows the result from a hypericin human buccal mucosa dataset and corresponds to the GPU-based rendering result in Fig. 11.

4 Discussion and Conclusion

Oral cancer is becoming one of the most common forms of cancer worldwide and early diagnosis is the key to a good prognosis. Current conventional techniques used for diagnosing oral cancer have their limitations and there is a need to develop newer and better techniques. We have previously shown the potential for confocal laser endomicroscopy (CLE), a minimally invasive endoscopic technique, to be used for fluorescence diagnostic imaging of oral cavity lesions.^{28–30} In this study, we present the further development of both hardware and software to optimize CLE for fluorescence 3-D imaging and to move toward a real-time virtual biopsy of oral lesions in a clinical setting.

We developed a prototype 3-D fluorescence imaging system comprising a field-programmable gated array (FPGA)-based embedded computing system interfaced to a confocal laser endomicroscope. The system is designed for automated acquisition of confocal image stacks and real-time volume rendering and display of 3-D tissue structures. In the normal operation of the endomicroscope, the manual acquisition of image stacks is accomplished by the operator using a footswitch. This process is slow and subject to movement by both the operator and the subject being imaged. With the prototype system, the image stack acquisition control is automated and needs only a start and stop input from the operator. Automation effectively

minimizes the chances of movement in between consecutive images, thus providing for more effective volume rendering in real-time. Even with this automation, image acquisition in the Z plane can be influenced by operator movement during image capture, for example if the pressure on the tissue changes and there is compression of the tissue. Selecting a different resolution changes the speed of image capture and thus may influence the effect of operator movement during the acquisition time. Currently, the best attainable rate limited by endomicroscope hardware is about 1.4 sec per frame. While this rate is slow by real-time standards, our results demonstrate the potential for 3-D endomicroscopic imaging of the oral cavity. Further hardware acceleration to increase the image capture rate will help to improve the performance of the system.

We tested the prototype endomicroscope-embedded system on murine models, healthy human volunteers and patients with head and neck lesions. Fluorescence 3-D imaging was carried out using fluorescein sodium or hypericin, both of which are safe for diagnostic use in humans.^{34,35} Confocal image stacks acquired from the murine and human oral cavities were rendered in real-time using programs developed using a graphics processing unit (GPU)^{52,53} and the embedded FPGA system that has been customized for fluorescence 3-D imaging. Compared to 2-D images from a conventional endomicroscope, the volume-rendered 3-D images highlight topographical information of the tissue being imaged. The 3-D images also provide depth-relation information between tissue structures at different focal planes. Such information is not easily visualized from the conventional 2-D confocal images acquired from a single focal plane. Depth information may be important for the assessment of the depth of neoplastic changes, carcinoma invasion etc. With experience, endoscopists may be able to interpret 3-D images for detection of abnormalities based on altered architecture and morphology.³¹

Fluorescein sodium yielded strong fluorescence signals and bright images since the spectral characteristics of fluorescein sodium match the excitation wavelength of the 488-nm laser in our system. However, fluorescein sodium is non-specifically absorbed by all cells and thus may result in false positives during diagnostic imaging. Moreover, although intravenously applied fluorescein, which is used in conventional CLE, is an option available to patients, the patients selected for this study so far have been reluctant to receive intravenous fluorescein. We have therefore used hypericin for our patient trials. Although the excitation wavelength of hypericin does not match 488 nm, resulting in weaker fluorescence signals, hypericin may be more selectively taken up by abnormal cells.³⁵ Therefore the improved selectivity of hypericin in lesional tissue might allow diagnostic imaging with higher specificity and provide sufficient contrast for diagnosis.

Further development plans include implementing real-time incremental volume-rendering while images are being acquired, feature registration to compensate for sample movement during imaging and hardware acceleration for faster image acquisition. The system setup using the current generation of endoscope-based confocal endomicroscopes does not yet have capability for guided biopsy. Our ultimate aim remains to develop a minimally invasive virtual biopsy method that can complement current diagnostic techniques and be used for fluorescence image-guided surgery and guided biopsy procedures in the oral cavity.

Acknowledgments

The authors would like to thank Cheong Lee Sing of the School of Computer Engineering, Nanyang Technological University, Singapore; Sasidharan Swarnalatha Lucky and Tee Chuan Sia of National Cancer Centre, Singapore; visiting student Agostino Guida of Seconda Università degli Studi, Naples, Italy; and Peter Delaney of Optiscan Pty, Ltd., Australia, for their assistance. This research project was supported by a grant from the Singapore Bioimaging Consortium (SBIC RP C010/2006).

References

- M. Garcia et al., *Global Cancer Facts and Figures 2007*, American Cancer Society, Atlanta, GA, 1–46 (2007).
- A. Jemal et al., “Global cancer statistics,” *CA Cancer J. Clin.* **61**(2), 69–90 (2011).
- B. W. Neville and T. A. Day, “Oral cancer and precancerous lesions,” *CA Cancer J. Clin.* **52**(4), 195–215 (2002).
- B. A. Flusberg et al., “Fiber-optic fluorescence imaging,” *Nat. Methods* **2**(12), 941–950 (2005).
- R. Kiesslich and M. I. Canto, “Confocal laser endomicroscopy,” *Gastrointest. Endosc. Clin. North Am.* **19**(2), 261–272 (2009).
- H. Neumann et al., “Confocal laser endomicroscopy: technical advances and clinical applications,” *Gastroenterology* **139**(2), 388–392 (2010).
- R. Kiesslich, M. Goetz, and M. F. Neurath, “Virtual histology,” *Best Pract. Res. Clin. Gastroenterol.* **22**(5), 883–897 (2008).
- R. Singh et al., “Real-time histology with the endocytoscope,” *World J. Gastroenterol.* **16**(40), 5016–5019 (2010).
- A. L. Carlson et al., “Confocal microscopy and molecular-specific optical contrast agents for the detection of oral neoplasia,” *Technol. Cancer Res. Treat.* **6**(5), 361–374 (2007).
- P. Delaney and M. Harris, “Fiber-optics in scanning optical microscopy,” in *Handbook of Biological Confocal Microscopy*, 3rd ed., J. B. Pawley, Ed., pp. 501–515, Springer Science+Business Media, New York, NY (2006).
- G. Le Goualher et al., “Towards optical biopsies with an integrated fibered confocal fluorescence microscope,” in *Proc. of the MICCAI'04*, Springer-Verlag, Berlin/Heidelberg, 761–768 (2004).
- L. Thiberville et al., “Confocal fluorescence endomicroscopy of the human airways,” in *Proc. Am. Thorac. Soc.* **6**(5), pp. 444–449 (2009).
- L. Thiberville and M. Salaün, “Bronchoscopic advances: on the way to the cells,” *Respiration* **79**(6), 441–449 (2010).
- F. S. Fuchs et al., “Fluorescein-aided confocal laser endomicroscopy of the lung,” *Respiration* **81**(1), 32–38 (2011).
- A. L. Polglase, W. J. McLaren, and P. M. Delaney, “Pentax confocal endomicroscope: a novel imaging device for in vivo histology of the upper and lower gastrointestinal tract,” *Expert Rev. Med. Devices* **3**(5), 549–556 (2006).
- M. Goetz and R. Kiesslich, “Advances of endomicroscopy for gastrointestinal physiology and diseases,” *Am. J. Physiol. Gastrointest. Liver Physiol.* **298**(6), G797–G806 (2010).
- M. I. Canto, “Endomicroscopy of Barrett's Esophagus,” *Gastroenterol. Clin. North Am.* **39**(4), 759–769 (2010).
- R. Kiesslich, “Screening: Endomicroscopy for a reliable diagnosis of colorectal neoplasia,” *Nat. Rev. Gastroenterol. Hepatol.* **7**(8), 422–423 (2010).
- K. Venkatesh et al., “Role of confocal endomicroscopy in the diagnosis of celiac disease,” *J. Pediatr. Gastroenterol. Nutr.* **51**(3), 274–279 (2010).
- Z. Li et al., “Confocal laser endomicroscopy for in vivo diagnosis of gastric intraepithelial neoplasia: a feasibility study,” *Gastrointest. Endosc.* **72**(6), 1146–1153 (2010).
- M. Goetz, A. Watson, and R. Kiesslich, “Confocal laser endomicroscopy in gastrointestinal diseases,” *J. Biophotonics* **4**(7–8), 498–508 (2011).
- X. J. Xie et al., “Differentiation of colonic polyps by confocal laser endomicroscopy,” *Endoscopy* **43**(2), 87–93 (2011).
- P. Sharma et al., “Real-time increased detection of neoplastic tissue in Barrett's esophagus with probe-based confocal laser endomicroscopy: final results of an international multicenter, prospective, randomized, controlled trial,” *Gastrointest. Endosc.* **74**(3), 465–472 (2011).
- G. A. Sonn et al., “Optical biopsy of human bladder neoplasia with in vivo confocal laser endomicroscopy,” *J. Urol.* **182**(4), 1299–1305 (2009).
- C. Wiesner et al., “Confocal laser endomicroscopy for the diagnosis of urothelial bladder neoplasia: a technology of the future?,” *BJU Int.* **107**(3), 399–403 (2011).
- J. Tan, P. Delaney, and W. J. McLaren, “Confocal endomicroscopy: a novel imaging technique for in vivo histology of cervical intraepithelial neoplasia,” *Expert Rev. Med. Devices* **4**(6), 863–871 (2007).
- J. Tan et al., “Detection of cervical intraepithelial neoplasia in vivo using confocal endomicroscopy,” *BJOG* **116**(12), 1663–1670 (2009).
- W. Zheng et al., “Confocal endomicroscopic imaging of normal and neoplastic human tongue tissue using ALA-induced-PPIX fluorescence: a preliminary study,” *Oncol. Rep.* **12**(2), 397–401 (2004).
- P. S. Thong et al., “Development of a laser confocal endomicroscope for in vivo fluorescence imaging,” *J. Mech. Med. Biol.* **7**(1), 1–8 (2007).
- P. S. Thong et al., “Laser confocal endomicroscopy as a novel technique for fluorescence diagnostic imaging of the oral cavity,” *J. Biomed. Opt.* **12**(1), 014007 (2007).
- B. R. Haxel et al., “Confocal endomicroscopy: a novel application for imaging of oral and oropharyngeal mucosa in human,” *Eur. Arch. Otorhinolaryngol.* **267**(3), 443–448 (2010).
- U. Günther et al., “Surveillance colonoscopy in patients with inflammatory bowel disease: comparison of random biopsy vs. targeted biopsy protocols,” *Int. J. Colorectal Dis.* **26**(5), 667–672 (2011).
- N. Q. Nguyen, “Real time intraoperative confocal laser microscopy-guided surgery,” *Ann. Surg.* **249**(5), 735–737 (2009).
- M. B. Wallace et al., “The safety of intravenous fluorescein for confocal laser endomicroscopy in the gastrointestinal tract,” *Aliment. Pharmacol. Ther.* **31**(5), 548–552 (2010).
- M. A. D'Hallewin, L. Bezdetnaya, and F. Guillemin, “Fluorescence detection of bladder cancer: a review,” *Eur. Urol.* **42**(5), 417–425 (2002).
- D. L. Campbell et al., “Detection of early stages of carcinogenesis in adenomas of murine lung by 5-aminolevulinic acid-induced protoporphyrin IX fluorescence,” *Photochem. Photobiol.* **64**(4), 676–682 (1996).
- L. S. Cheong et al., “Embedded computing for fluorescence confocal endomicroscopy imaging,” *J. Signal Process. Syst.* **55**(1–3), 217–228 (2009).
- P. S. Thong et al., “Detection and diagnosis of human oral cancer using hypericin fluorescence endoscopic imaging interfaced with embedded computing,” in *Proc. SPIE* **7380**, 73806U (2009).
- S. S. Tandjung et al., “Synchronized volumetric cell image acquisition with FPGA-controlled endomicroscope,” in *Proceeding of the International Conference on Embedded Systems and Applications*, pp. 61–67, ESA'09, CSREA Press, USA (2009).
- P. S. Thong et al., “Hypericin fluorescence imaging of oral cancer: From endoscopy to real-time 3-dimensional endomicroscopy,” *J. Med. Imag. Health Inf.* **1**(2), 139–143 (2011).
- M. Goetz et al., “In-vivo confocal real-time mini-microscopy in animal models of human inflammatory and neoplastic diseases,” *Endoscopy* **39**(4), 350–356 (2007).
- T. J. Cullip and U. Neumann, *Accelerating Volume Reconstruction with 3-D Texture Hardware*, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA (1994).
- B. Cabral, N. Cam, and J. Foran, “Accelerated volume rendering and tomographic reconstruction using texture mapping hardware,” in *Proc. of the 1994 Symposium on Volume Visualization Tysons Corner*, ACM, Virginia, USA (1994).
- A. Van Gelder and K. Kim, “Direct volume rendering with shading via three-dimensional textures,” in *Proc. of the 1996 Symposium on Volume Visualization*, IEEE Press, San Francisco, CA (1996).
- R. Westermann and T. Ertl, “Efficiently using graphics hardware in volume rendering applications,” in *Proc. of the 25th Annual Conference on Computer Graphics and Interactive Techniques*, ACM, New York, NY, USA (1998).

46. F. Dacheille et al., "High-quality volume rendering using texture mapping hardware," in *Proc. of the ACM SIGGRAPH/EUROGRAPHICS Workshop on Graphics Hardware*, ACM, Lisbon, Portugal(1998).
47. B. Lichtenbelt, R. Crane, and S. Naqvi, *Introduction to Volume Rendering*, Hewlett-Packard Professional Books, Prentice-Hall, Inc., New Jersey, pp. 1–120 (1998).
48. K. Engel et al., *Real-Time Volume Graphics*, A. K. Peters, Ltd., USA, 47–273, (2006).
49. B. Preim and D. Bartz, "Algorithms for direct volume visualization," Chapter 9, in *Visualization in Medicine: Theory, Algorithm and Applications*B. Preim and D. Bartz, Eds. pp. 197–288, Elsevier, Inc., Burlington, MA (2007).
50. M. Meissner, U. Hoffmann, and W. Strasser, "Enabling classification and shading for 3-D texture mapping based volume rendering using OpenGL and extensions," in *Proc. of the Conference on Visualization '99: Celebrating Ten Years*, IEEE Computer Society Press, San Francisco, CA (1999).
51. C. Rezk-Salama et al., "Interactive volume on standard PC graphics hardware using multi-textures and multi-stage rasterization," in *Proc. of the ACM SIGGRAPH/EUROGRAPHICS Workshop on Graphics Hardware*, ACM, Interlaken, Switzerland (2000).
52. M. M. Movania et al., "Automated local adaptive thresholding for real-time feature detection and rendering of 3-D endomicroscopy images on GPU," *Proc. of the 2009 Int. Conf. on Computer Graphics and Virtual Reality, CGVR 2009*, CSREA Press, USA (2009).
53. M. M. Movania et al., "GPU-based surface oriented interslice directional interpolation for volume visualization," *Proc. of the 2nd Int. Symposium on Applied Sciences in Biomedical and Communication Technologies*, Isabel pp. 1–5 (2009).
54. M. Levoy, "Display of surfaces from volume data," *Comput. Graph. Appl. IEEE* **8**(3), 29–37 (1988).