REVIEW

Computer vision and augmented reality in gastrointestinal endoscopy

Nadim Mahmud¹, Jonah Cohen², Kleovoulos Tsourides³ and Tyler M. Berzin²,*

¹Department of Internal Medicine, Brigham and Women’s Hospital, Harvard Medical School, Boston MA, USA, ²The Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center, Harvard Medical School, Boston MA, USA and ³Brain and Cognitive Sciences, Massachusetts Institute of Technology, Boston MA, USA

*Corresponding author. The Center for Advanced Endoscopy, Beth Israel Deaconess Medical Center, 330 Brookline Avenue, Boston, MA 02215, USA.
Tel: +1-617-632-8623; Fax: +1-617-632-9199; Email: tberzin@bidmc.harvard.edu

Abstract

Augmented reality (AR) is an environment-enhancing technology, widely applied in the computer sciences, which has only recently begun to permeate the medical field. Gastrointestinal endoscopy—which relies on the integration of high-definition video data with pathologic correlates—requires endoscopists to assimilate and process a tremendous amount of data in real time. We believe that AR is well positioned to provide computer-guided assistance with a wide variety of endoscopic applications, beginning with polyp detection. In this article, we review the principles of AR, describe its potential integration into an endoscopy set-up, and envisage a series of novel uses. With close collaboration between physicians and computer scientists, AR promises to contribute significant improvements to the field of endoscopy.

Key words: augmented reality; gastrointestinal endoscopy; colonoscopy; computer vision

Introduction

Augmented reality (AR) is the process of superimposing computer-generated objects and data over existing, real structures [1]. This differs from virtual reality, where the basic elements of the environment are entirely computer-generated in an effort to simulate their existence [2]. Augmentation typically operates within the semantic context of environmental elements. A simple example of this is displaying ‘live’ scores on a televised sports match. With the help of advanced AR technology, such as computer vision and object recognition, information about the environment can become interactive and easier to manipulate from a computational perspective [3–6]. AR has empirically sought to seamlessly integrate reality with analytical information, to improve a user’s ability to perform a task in real time. Consider, for example, the earliest application of AR when Mark VIII fighters in World War II had live radar information displayed in the pilot’s line of sight. This improved the pilot’s ability to locate other airplanes in the sky and identify enemy aircraft [7].

Despite its long history, AR has only recently made its debut in medical practice, being applied primarily to navigational surgery [8]. This involves taking data from pre-operative imaging and using anatomical anchors in the operating field to link—or register—the two representations in real time. The neurosurgical and otolaryngological fields have used AR to map 3D images of a patient’s paranasal sinus or neuroanatomy on monitors, and mobile devices to assist with various surgical procedures, including prototypes to display the ventricular system for drain placement and brain tumors for resection planning [9, 10]. Cabrilo et al. utilized AR in 33 patients, using representations of the cerebral arteries to assist with arteriovenous malformation...
resection and aneurysm clipping [11, 12]. Finally, several studies describe landmark-based AR systems for endoscopic sinus or skull-base surgeries [13–15].

Although few studies in general surgery have taken AR to the operating room, environments have been developed for both laparoscopic and open procedures [16]. López-Mir et al. used AR to improve the accuracy of trocar placement for cholecystectomy [17], while Kang et al. integrated laparoscopic ultrasound imaging in order to interrogate the liver, gallbladder, biliary tree, and kidneys below the visible surface [18]. Volonte et al. reported a successful cholecystectomy using AR-based stereoscopic images with the da Vinci robot [19], and Simpfendörfer et al. generated AR images from transrectal ultrasound to perform a radical prostatectomy [20]. The demand for AR in open surgeries has been more limited, but Peterhans et al. performed partial hepatic resections in nine patients, with AR used to display vascular and biliary anatomy [21], and Marzano et al. completed a pancreaticoduodenectomy using AR to highlight important vascular structures [22].

The role of AR in gastrointestinal (GI) endoscopy to date has been limited to localization of GI tumors and novel transluminal surgical approaches. Kranzfelder et al. described an AR system that successfully registered CT data with upper endoscopy imaging in 24 patients with GI tumors [23]. Azagury et al. developed an image registration (IR) system to assist with natural orifice transluminal endoscopic surgery (NOTES) [24]. IR-NOTES was used to target various intra-abdominal organs with the endoscope in 35 cadavers. This group and Vosburgh et al. found that easy transluminal access to the kidneys, gallbladder, liver, and pancreas could be achieved with close AR guidance, raising the possibility of novel endoscopic surgical procedures [24, 25].

While the potential of navigational surgical excitement is, we believe that GI endoscopy is fertile ground for a vast array of new applications, including detection and identification of polyps, performance tracking, pathology scoring and more. In this article we review the components of AR systems, elaborate on our proposed applications for AR, highlight technical challenges, and outline a path towards innovation.

**Technical considerations for AR in the endoscopy unit**

Important to the fundamentals of AR are the technologies of image processing and computer vision. Image processing refers to the deconstruction of image data into a series of parameters or properties relating to the image. ‘Computer vision’ refers to high-level image processing in which a computer deciphers the contents of an image or sequence of images and uses this information to make intelligent decisions [26–29]. In essence, these technologies allow computers to ‘see and understand’ environments, and to make complex judgments for further output [30]. In the case of video- and image enhancement through graphical overlay, the output is termed ‘augmented reality’.

There is a wide variety of AR set-ups and applications but they have in common a data input source, a processor, and a display [31]. The input source is commonly a camera, which provides information for the computer to compare with image databanks, effectively allowing it to ‘see’ what the user sees. The display is the medium for combining reality with virtual information. In different clinical settings, the display may be a standard video monitor or an optical head-mounted display, Google Glass being a recent example of the latter.

The basic hardware set-up for an endoscopy procedure is well suited to adaptation to an AR environment because high-definition video capture is already a core capability of most modern endoscopy units. A high-definition camera, at the distal tip of the flexible endoscope, supplies image data to a camera control unit (CCU), where the data are processed and formatted for output to high-definition monitors. The CCU is further connected to a quick-swap memory unit and a computer or keyboard to accept user input. An added central processing unit can be inserted in series between the CCU and monitor, to enable image processing and computer vision capabilities. This will house the signal processing, image analysis, and decision-making capabilities of the system prior to modified high-definition output with graphical overlay.

**AR to improve polyp detection**

Colonoscopy is the ‘gold standard’ for early detection of colorectal adenomas [32], and more than 14 million colonoscopies are performed in the United States annually [33]. Despite this volume, there is a significant adenoma miss rate (AMR), according to the literature ranging from 6–27% and depending on a variety of polyp and operator characteristics [34]; for example, smaller polyps [26, 35], flat polyps [37], and left colonic location [38] may be associated with an increased miss rate. Sessile serrated adenomas are a topic of particular concern, given their higher pre-dilection towards neoplastic change [39, 40] and, unlike pedunculated polyps, are more frequently missed in the right side of the colon [41]. Operator experience and fatigue are also significant considerations; endoscopists are more likely to miss polyps in the afternoon, as compared to morning cases [42]. Importantly, there is convincing evidence that having ‘more eyes’ on the video monitor increases the adenoma detection rate. Observation of the video monitor by nurses has been shown to increase polyp detection by ~30% [43, 44] and, participation by a gastroenterology trainee has been shown to do the same [45]. These findings suggest that individual endoscopists routinely miss polyps that are visible in the monitor.

Increasing adenoma detection rate improves the preventive efficacy of colonoscopy and polypectomy [46, 47]. During a colonoscopy, there are two general reasons why a polyp might be missed: (i) it was never in the visual field or (ii) it was in the visual field but not recognized.

Several hardware innovations have sought to address the first problem through expanded visualization of the colonic lumen. This includes the Third Eye retroscope camera, designed to identify polyps hidden behind folds in the bowel wall [48], and full-spectrum endoscopy (FUSE), which provides a wider, 330° left-to-right endoscopic view [49]. The second problem—unrecognized polyps within the visual field—has been more difficult to address. In addition to the available data on the benefit, in terms of polyp detection, of additional observers in the endoscopy suite, multiple studies have attempted to use chromoendoscopic dyes [50–52] or narrow-band imaging [53] to make flat or isochromatic polyps more apparent. There exists a great deal of controversy surrounding the comparative effectiveness of these approaches [54].

Augmented reality represents an important opportunity to improve adenoma detection rate. By aggregating a large volume of polyp images from colonoscopies, it is possible to implement machine-learning and computer vision algorithms to assist with polyp detection. This would be the first line of approach to innovation within the context of the high-resolution video data that is routinely acquired during colonoscopy. Using AR, the endoscopist could enjoy real-time visual assistance with polyp detection in the form of overlaid images on the primary HD
monitor or on an adjacent one. In more advanced iterations, suspected polyp type may also be displayed, with color-coding or other visual information used to represent the level of confidence in the analysis (Figure 1).

Several important features of this technology can be highlighted. First, modern processing capabilities would enable an AR system to function in real time during a colonoscopy. Second, flat or isochromatic polyps, which might be visually occult to the less experienced or fatigued endoscopist, would be parsed by the AR system in identical fashion to any other polyp; that is to say that, while humans have a tendency to “see what they want to see” [53], an AR system will “see what it is trained to see” with fidelity and without bias. Third, as more image data is acquired for analysis, the AR system will become increasingly efficient and accurate in detecting polyps of all varieties.

Finally, AR is complementary to contemporary hardware innovations designed to improve adenoma detection rate through expanded visualization. Indeed, the Third Eye and FUSE systems involve additional cameras and force the operator to scan multiple screens displaying live video. As mentioned above, AR could draw the endoscopist’s eye, in real time, to lesions that give rise to concern, effectively creating ‘extra sets of eyes’ on all aspects of the video data.

Other potential AR applications in GI endoscopy

Computer vision for polyp classification

The American Society for Gastrointestinal Endoscopy recommends that all neoplastic polyps in the colon be resected [56]; however, distinguishing between neoplastic and non-neoplastic polyps can be extremely challenging at the time of endoscopy. Up to 35% of colonic polyps are non-neoplastic, including hyperplastic, inflammatory, and mucosal polyps [57]. Polypectomy followed by histopathological analysis of all acquired specimens is currently the ‘gold standard’ approach to polyp classification. To minimize unnecessary polyp resection—which prolongs procedures and increases the risk of morbidity—several novel technologies have emerged to distinguish neoplastic from non-neoplastic polyps. These include narrow-band imaging and endoscopy, both of which require specialized training and additional hardware [58–60]. Definitive polyp classification through AR could provide similar benefit with less expense. As detailed above, AR can work with existing endoscopy set-ups and would require minimal training.

Multimodality image enhancement

There are several imaging techniques used during gastrointestinal endoscopy, which could benefit from integration with augmented reality. Endoscopic ultrasound, for example, can be used to characterize organs or masses adjacent to the gastrointestinal tract. Structural images obtained using EUS could be registered in order to display a spatial projection of the object in the live colonoscopy video feed. Registration of EUS images for computer-guided navigation has previously been described [61], and could simplify targeting for needle biopsies or cyst drainage. Similarly, in endoscopic retrograde cholangiopancreatography (ERCP), both fluoroscopic and endoscopic images are used simultaneously during evaluation of the pancreaticobiliary system. Applying AR solutions to fluoroscopic images has been explored in the arena of cardiac interventions [62], and image analysis tools to measure the diameter and length of fluoroscopically identified biliary strictures might facilitate accurate stent selection for optimal treatment of strictures.

Performance improvement and tracking

As governmental and insurance regulations surrounding accountable care take form, there will be ever-increasing pressure for endoscopists to track their polyp detection rates and other performance metrics; thus another critical aspect of endoscopic AR could be to track, test, and validate polyp detection performance. This could assist in quality reporting for the purpose of tracking and improving outcomes in gastrointestinal endoscopy [63]. Tracking could also generate automated data regarding colonoscopy withdrawal time, which has been significantly linked to the quality of adenoma detection [64]. This information could be archived in order to inform endoscopists about performance measures in real time, as well as during aggregate performance reviews.

Digital ruler for improved accuracy of measurement

Accurate estimation of the size of various anatomical structures during endoscopy has the potential to greatly enhance current practice. In the arena of polyp detection, determining the size of a given polyp plays a critical role in dictating follow-up intervals for a patient’s next surveillance colonoscopy. Recent data suggest that, with visual estimation, substantial variations in recorded polyp size occur, leading to incorrect surveillance intervals in 10% of cases [65, 66]. To date there is no digital tool to assist with polyp sizing and traditional methods, such as comparison with biopsy forceps and snares, add little to accuracy [67]. While a ‘digital ruler’ using AR might require a second camera, several endoscopes already incorporate multiple lenses, and innovative reference points may ultimately enable accurate measurement with standard single-camera endoscopes.

Standardized scoring/grading systems

Standardized endoscopic scoring systems exist for many disease states of the gastrointestinal tract. Unfortunately, their utility is often limited by subjective interpretation and impracticality in the clinical setting. Inflammatory bowel disease (IBD)
scoring systems, for example, are used to assess disease response in clinical trials but have been difficult to apply in clinical practice [68, 69]. AR-assisted IBD scoring could determine patient response to novel therapeutics through objective assessment of mucosal healing. A second application that would benefit from objective scoring is the grading of bowel preparation quality. Despite its critical importance in informing surveillance intervals, current grading is operator-dependent and highly variable [70, 71]. AR-based scales, based on stool burden, would minimize subjectivity and globally improve adherence to surveillance guidelines.

**Dynamic braking for capsule endoscopy**

Wireless capsule endoscopy is indicated to interrogate portions of the small bowel that are not easily accessed through standard endoscopy—usually to search for obscure gastrointestinal bleeding or signs of Crohn’s disease. Recent advances have granted some degree of control over maneuvering the capsule whilst it is in the gastrointestinal tract. This has been accomplished through techniques such as electro-stimulation and by using magnetic fields [72, 73]. Although video data from wireless capsules are currently processed after completion of the endoscopy, it is likely that future generations of capsules will support real-time video analysis. When this is available, computer vision and augmented reality could be used to automatically detect pathology and trigger slowing of the capsule using the methods noted above. This would maximize the video data acquired to better delineate lesions of interest.

**The road ahead**

While no software platforms to integrate augmented reality into endoscopy are currently available, this is an active area for research and innovation. We have only just entered into an era in which computer vision technology may surpass human ability in terms of facial and object recognition [74]. The combination of computer vision and endoscopy thus offers vast potential to enhance the abilities of the average endoscopist. For this technology to integrate into regular practice, however, numerous challenges must be overcome. First, in order to optimize the signal-to-noise ratio and minimize false positive (during polyp recognition, for example), enormous image and video repositories will be needed for reference and machine learning. Second, careful quality control must be paramount in all aspects of novel AR applications. Finally, creating a useful AR assistance tool for endoscopists will necessarily require close collaboration between clinicians, software engineers and computer vision experts. Careful attention to these potential pitfalls may bring AR to the forefront of the next era of gastrointestinal endoscopy.

**Conflict of interest statement:** NM, JC, KT, and TB are advisors to EndoLayers.

**References**

2. Wright WG. Using virtual reality to augment perception, enhance sensorimotor adaptation, and change our minds. *Front Syst Neurosci* 2014;8:56.


