A prototype design of a wireless capsule endoscope

Y. Chan¹, M. Q.-H. Meng², X. Wang² and J. C.-Y. Wu¹

¹Department of Medicine and Therapeutics, Chinese University of Hong Kong, Shatin, NT, Hong Kong
²Department of Electronic Engineering, Chinese University of Hong Kong, Shatin, NT, Hong Kong

Abstract: Diseases of the gastrointestinal tract, such as intestinal bleeding and ulceration, are very common. To determine the location of the disease, as well as to diagnose the problem, endoscopy is performed. A wireless capsule endoscope is a pill-sized device that is easily swallowed by the patient. It is equipped with a camera and a wireless transmitter so that as it travels through the patient’s gastrointestinal tract, it takes pictures or video images of the tract and transmits the information out of the patient’s body. This article discusses the possible solutions and challenges of a wireless capsule endoscope in terms of the choice of wireless frequency and circuit components. A prototype is built with commercially available components to demonstrate the concept of the device.

Key words: Endoscopy, wireless capsule endoscope, prototype, wireless transmitter, camera.

INTRODUCTION

The human gastrointestinal (GI) tract is a 30-ft long structure that includes the oesophagus, stomach, intestine, colon and rectum. Diseases of the GI tract, such as intestinal bleeding and ulceration, are very common. To determine the location of the disease, as well as to diagnose the problem, endoscopy is usually performed. A traditional endoscope is a long and flexible cable, with a camera fixed at the navigating end and a control unit at the other end. Video images captured by the camera are transmitted to a monitor for display using optic fibres embedded in the cable. While a gastroscope is pushed through the patient’s mouth to examine the first 4 ft of the GI tract, a colonoscope is pushed through the patient’s anus to examine the bottom 6 ft of the tract; the remaining 20 ft of the tract is out of reach. Other imaging and diagnostic techniques for GI tract diseases such as ultrasound and X-rays imaging are neither accurate nor effective. These endoscopic devices also create patient discomfort as they are pushed through the patient’s mouth and anus. Sedation and local anaesthesia are needed for patients undergoing these procedures.

To solve these problems, a biomedical ingested device, wireless capsule endoscope (WCE), has been developed (Iddan et al. 2000; Ginsberg et al. 2002; Given Imaging Ltd. n.d.; Intelligent Microsystem Center n.d.; Olympus Medical System Corporation n.d.). It consists of an image sensor, a lens, light-emitting diodes (LEDs) and a transmitter for wireless transmission of video images captured by the image sensor. Once the capsule is swallowed, it travels the GI tract passively and transmits video images out of the patient’s body through radiotelemetry. Figure 1 shows the already commercialised WCEs, PillCam and EndoCapsule developed by Given Imaging Ltd., Israel and Olympus Medical System Corporation, Japan, respectively. Since the device’s first appearance, a lot of research groups have been working on the system, hoping to further improve the system in terms of imaging and video transmission (Chan et al. 2004; Hu et al. 2004), localisation (Nagaoka and Uchiyama 2003, 2004; Wang et al. 2004), actuation (Sendoh et al. 2003; Chiba et al. 2005; Kim et al. 2005a, 2005b) and prototype design (Gong et al. 2000; Park et al. 2002; Chiu et al. 2004; Meng et al. 2004; Xie et al. 2004; Chan et al. 2005a).

Little discussion, however, has been made in terms of the device’s technical details and the challenges involved in designing the system. This article gives an overview of the system, as well as the possible and potential solutions in developing a WCE prototype. Although the commercial WCEs have everything custom made, this article focuses on the technology that can be purchased off the shelf. The discussion includes the choice of wireless frequency and circuit components. A prototype was built to demonstrate...
Y. Chan, M. Q.-H. Meng, X. Wang and J. C.-Y. Wu

PROTOTYPE OVERVIEW

The prototype shown in Figure 2 has four key components: (1) a video camera, (2) a radio frequency (RF) modulator, (3) an antenna and (4) a lighting system. Once the power of the device is turned on, the video camera starts taking video images of the GI tract. The video signal is then be processed and up converted into an RF signal by the RF modulator. This signal is then transmitted wirelessly out of the human body through an antenna. Finally, the lighting system illuminates the interior of the GI tract so that the video camera is able to capture clear images with suitable brightness.

When designing the prototype, two main constraints have to be considered. The first constraint is the size of the device. The commercial WCEs have a size of 11 mm × 25 mm and the design should not be anything bigger than this. The second constraint is the wireless transmission frequency. Although as a rule of thumb, the frequency should not exceed 1 GHz, as most of the radiation energy is absorbed by the human body tissues above this frequency (Chirwa et al. 2003). A recent article studying the radiation efficiency from an internal source (Chan et al. 2005b) suggested that the 434 MHz, 915 MHz and 2.45 GHz industrial, scientific and medical (ISM) bands are all suitable for wireless transmission from inside the human body. Therefore, this article considers the use of these three ISM bands as the wireless link for video transmission.

Analog and digital video cameras

Traditional video is in the form of an analog video. This includes the broadcasting television (TV) and the video home system. On the other hand, digital video has evolved with the advantage of better quality and accuracy in transmission with products including video compact disc, digital versatile disc and high-definition TV.

As discussed in Chan et al. (2004), with the development of complementary metal-oxide semiconductor (CMOS) and charge-coupled device image sensors, the size of both analog and digital camera modules has decreased drastically. While analog cameras are used for applications such as surveillance cameras, which require less precision, digital cameras are mostly developed to meet the demand of...
digital camera and cellphone camera applications. With the advancement of packaging technology, camera size continues to decrease. CMOS analog image sensors such as OV06910 and OV07648, which have sizes of 11 mm × 11 mm and 5 mm × 5 mm, respectively, were developed by OmniVision Technologies, Inc, US (OmniVision Technologies, Inc n.d.). The later comes with Chip Scale Package, which largely reduces the size of the chip. The picture of a CMOS image sensor is shown in Figure 3. Figure 4 shows an analog camera module that is available commercially, with built-in lens and case. The size of this camera module is about 15 mm × 15 mm. Since analog cameras are mainly developed for surveillance application, manufacturers are less likely to research on smaller modules.

On the contrary, size of digital camera modules continues to decrease, with more and more companies, such as Sony, Hitachi, Fujitsu, Sanyo and Toshiba, Japan joining the competition. For example, Fujitsu announced its newly developed VGA digital camera module to have a size of 5.7 mm × 5.7 mm × 3.5 mm, and Hitachi’s HAM49002 VGA digital camera module has a size of 7 mm × 7.6 mm × 4.9 mm. A Korean company, MCNEX, also developed the world’s smallest camera at the end of year 2005 that has a size of 5.5 mm × 5.5 mm × 2.9 mm. Table 1 summarises the comparison of miniaturised analog and digital cameras.

Digital camera has the greatest advantage in terms of its small size, high resolution and low power consumption, whereas analog camera is extremely easy to interface. The video signal output from an analog camera is a 1-pin output of analog video signal, which can be fed to any video-viewing device such as a traditional TV or a videocassette recorder. In contrast, digital video camera outputs the signal in various digital formats, such as RGB, YUV and YCrCb. Signals are usually outputted in parallel by a 4-pin or 8-pin configuration. Interfacing includes synchronisation of the digital data. Data processing is also needed to reconstruct the video image from digital data, which further complicates the circuit.

**Analog and digital transmitters**

Because of the rich information provided by video images, the bandwidth of an analog video is at least 6 MHz and the data rate of a digital video is at least 3.5 Mbps for a 2 frames per second (fps) system with CIF (352 × 288 pixels) resolution. Therefore, the transmitter for video transmission has to be specially designed for the high volume of data.

A simplest analog video transmitter is a TV modulator that employs an amplitude modulation to up-convert the video signal to RF signal. The RF signal can then be picked up by a normal TV where the video images can be displayed. As a result, no external or special circuit is needed at the receiving end. An example of the simplest
Table 1 Comparison of analog and digital video cameras

<table>
<thead>
<tr>
<th>Video format</th>
<th>Analog (TV and VHS system)</th>
<th>Digital (VCD, DVD, HDTV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Industrial application</td>
<td>Surveillance cameras</td>
<td>PDA, cell phone, digital cameras</td>
</tr>
<tr>
<td>Size</td>
<td>Large (12 mm × 12 mm)</td>
<td>Small (6 mm × 6 mm)</td>
</tr>
<tr>
<td>Resolution</td>
<td>Medium</td>
<td>High</td>
</tr>
<tr>
<td>Power consumption</td>
<td>High</td>
<td>Low</td>
</tr>
<tr>
<td>Circuit interfacing</td>
<td>Simple</td>
<td>Complicated</td>
</tr>
<tr>
<td>Development</td>
<td>Slow</td>
<td>Fast</td>
</tr>
</tbody>
</table>

<VHS = video home system; VCD = video compact disc; DVD = digital versatile disc; HDTV = high-definition TV; and PDA = personal digital assistant.>

Table 2 Comparison of digital transmitters using different technologies

<table>
<thead>
<tr>
<th>Digital wireless transmission technology</th>
<th>ISB transmitter (434 MHz, 915 MHz)</th>
<th>ISB transmitter (2.4 GHz)</th>
<th>Bluetooth (2.4 GHz)</th>
<th>WLAN 802.11b (2.4 GHz)</th>
<th>WLAN 802.11g (2.4 GHz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range of transmission</td>
<td>Short</td>
<td>Short</td>
<td>Short</td>
<td>Medium</td>
<td>Medium</td>
</tr>
<tr>
<td>Highest data rate</td>
<td>500 kbps</td>
<td>1 Mbps</td>
<td>721 kbps</td>
<td>11 Mbps</td>
<td>22 Mbps</td>
</tr>
<tr>
<td>Circuit complexity</td>
<td>Simple</td>
<td>Simple</td>
<td>Medium</td>
<td>Complicated</td>
<td>Complicated</td>
</tr>
<tr>
<td>Video transmission</td>
<td>Compressed video</td>
<td>Compressed video</td>
<td>Compressed video</td>
<td>Uncompressed video</td>
<td>Uncompressed video</td>
</tr>
<tr>
<td>Size</td>
<td>Small</td>
<td>Small</td>
<td>Medium</td>
<td>Large</td>
<td>Large</td>
</tr>
</tbody>
</table>

<ISB = independent sideband; WLAN = Wireless Local Area Network.>

TV modulator that provides pin-to-pin connection is the MC44BS373C developed by Motorola, US with size of 6 mm × 6 mm. The modulator can also be constructed using surface-mounted device type transistor, resistors and inductors, which will give a smaller circuit. The commercially available MC44BS373C transmits RF signal in the frequency range of VHF (30–450 MHz) and UHF (450–880 MHz), whereas a custom-made transmitter can be made to transmit data at almost any frequency at or above the VHF range.

There are a lot of miniature digital transmitters developed specifically for the ISM bands of 434 MHz, 915 MHz and 2.4 GHz, or for other wireless applications. Transmitters designed for the 434 MHz and 915 MHz bands have slower data rate, as they are mainly designed for carrying voice, text and command information. For example, the TRF4903 ISM Band RF Transmitter from Texas Instrument, US has a maximum data rate of 64 kbps, and hence cannot carry digital video data of at least 3.5 Mbps. Chipcon, US CC1100 multi-channel RF transmitter has highest data rate of 500 kbps, which is still too low, although it may be used for compressed video transmission.

Transmitter of 2.4 GHz has a different story. There are 2.4 GHz transmitters designed for ISM and general use, such as the Chipcon’s CC2400 2.4 GHz ISM Band Transceiver and the Nordic Semiconductor, Norway nRF2401, which have programmable data rate of up to 1 Mbps. Other than these generic transmitters, there are other widely used technologies such as Bluetooth and Wireless Local Area Network (WLAN). Bluetooth is designed for short-distance (<10 m) transmission and has a maximum data rate of 721 kbps, whereas WLAN is designed for medium distance (within a building) transmission and has a maximum data rate of 22 MHz. Among the different specifications of WLAN, both the 802.11b and 802.11g standards developed by the Institute of Electrical and Electronics Engineers are most widely used. There are a lot of chipset solutions developed by companies such as Intersil, US and Texas Instruments. The highest data rate of 802.11b standard is 11 MHz and that of 802.11g is 22 MHz. As a result, in terms of data rate, the WLAN technology is most suitable for video transmission. However, WLAN solution is very complex and bulky. Even the smallest chipset solution PRISM 3Ω developed by Intersil has two chips, one of which is an RF IC and the other is a baseband processing chip. Table 2 compares the digital transmitters using different technologies, whereas Table 3 summarises the comparison of analog and digital transmitters.

Since wireless video transmission is the core function of WCE, the design of the endoscope largely depends on
the chosen solution. There are four solutions proposed for wireless video transmission and they are explained as follows.

**Solution 1: analog camera + analog video transmission**
This is the simplest and most straightforward solution for the system. The analog video captured by an analog CMOS image sensor is sent to a TV modulator or an amplitude modulation circuit to generate the RF signal in the range of 30 to 800 MHz. The RF signal is then radiated through an antenna. A block diagram of the system is shown in Figure 5. This system is easy to build and visualise, and requires the least number of components. Although the size of commercially available analog cameras (12 mm × 12 mm) is way larger than the acceptable size of a WCE (11 mm in diameter and 25 mm in length), it is possible to customise cameras to the required size.

**Solution 2: digital camera + analog video transmission**
This solution makes use of a digital camera instead of an analog one. Commercial digital camera has smaller size and can be easily obtained from the market. Because analog video transmission is easy and straightforward, it is also used in this solution. However, a converter is needed for the conversion of the digital video signal outputted from the digital camera to analog video signal so that it can be fed to a TV modulator. As a result, a video digital-to-analog converter (DAC) is needed. There are many video DACs available in the market for different applications such as high-definition TV and digital versatile disc conversions. Only a few of them meet the digital video interface that is specified by individual digital cameras. The smallest one is ADV7179 Video DAC (6 mm × 6 mm) developed by Analog Devices. It accepts 8-bit 27 MHz YCrCb digital data and outputs PAL/NTSC analog video signal. After conversion, the analog signal can be fed to the video transmission module in the same way as the previous solution. Figure 6 shows a block diagram of this solution. This solution is a bit complicated for implementation, and it is about the same size as the previous one, since the Video DAC IC has a small size of 6 mm × 6 mm. In addition, the increase in solution size due to the additional IC can be compensated by the smaller size of digital camera.

**Solution 3: digital camera + digital video transmission using WLAN technology**
This solution makes use of the WLAN technology so that any computer connected to the WLAN network can pick up the video images and display it on a computer monitor. Since WLAN has data rate of up to 22 Mbps, high-quality video images can be transmitted without compression. However, the transmission circuit is very complicated. First, digital camera provides digital data in parallel output. Image processor or data serialiser is needed to serialize the parallel data before transmission. Then the serialised data have to be packaged according to the data communication protocol. As a result, a very fast and powerful central processing unit is required to process the high volume of data. The data can then be sent to the 802.11b/g chipset solution, which processes the data so that it is ready to be transmitted by the antenna at 2.4 GHz. At the

---

![Figure 5 Block diagram of solution 1: analog camera + analog video transmission.](image)

![Figure 6 Block diagram of solution 2: digital camera + analog video transmission.](image)

<table>
<thead>
<tr>
<th>Video transmission</th>
<th>Analog</th>
<th>Digital</th>
</tr>
</thead>
<tbody>
<tr>
<td>Property of video transmission</td>
<td>Large bandwidth (6 MHz)</td>
<td>High data rate (3 Mbps of CIF at the rate of 2 fps)</td>
</tr>
<tr>
<td>RF for uncompressed video transmission</td>
<td>&gt;30 MHz</td>
<td>&gt;2.4 GHz</td>
</tr>
<tr>
<td>Circuit complexity</td>
<td>Very simple</td>
<td>Complicated</td>
</tr>
<tr>
<td>Size</td>
<td>Small</td>
<td>Medium</td>
</tr>
</tbody>
</table>

*CIF = cells in frame; fps = frames per second; and RF = radio frequency.*
receiving end, a computer with the standard WLAN hardware receives the RF data. However, to display the video images in the computer, a program has to be written to process or extract the video information from the received data and reconstruct the images or the video. The system is illustrated in Figure 7.

As shown in Figure 7, this solution is very complicated and requires a lot of work. A program has to be written for the central processing unit to perform the required process, and another program has to be written at the receiving end in order to process and display the received video data. In addition, more circuit components have to be used, such as a serializer and an extra central processing unit. Besides, current 802.11b/g chipset solution is bulky in size, which is impossible to fit into a capsule. Therefore, this solution is not a likely choice in
today’s technology, but has the potential to be used in the future.

**Solution 4: digital camera + digital video transmission with video compression**

Figure 8 shows the block diagram of the solution. With video compression, the video data to be transmitted can be largely reduced. This allows the use of transceivers of lower data rate as well as narrower bandwidth. As a result, by using an image compression chip such as the JPEG (Joint Photographic Expert Group) Encoder OV528 from OmniVision Technologies, Inc, the digital images captured by the camera are compressed into JPEG format. The OV528 Compression Engine converts 15 fps images of
640 × 480 resolutions into serial JPEG data of 115.2 kbps, which can be transmitted by most of the RF transmitters in the ISM bands of 434 MHz, 915 MHz and 2.4 GHz. In recent years, manufacturers have begun to include JPEG encoders in their camera modules to reduce the overall solution size. For example, CoderCam TC5740 and CoderCam TC5747 digital VGA cameras developed by TransChip, Inc., US have built-in JPEG codec so that the output images are already compressed. At the receiving end, similar hardware is needed to receive the signal and extract the JPEG data so that the images can be reconstructed and displayed on the computer.

**Antenna**

An ideal antenna of a WCE is small, has an evenly distributed radiation pattern and has a wide bandwidth. Because of the compact space available in a WCE, the antenna has to be small enough to be encased. When the WCE is swallowed by a patient, it travels the patient’s GI tract passively, and therefore it is randomly oriented. Thus, the direction of radiation transmission is not controlled. Therefore, an omni-directional radiation pattern is needed so that the antenna transmits the signal evenly in all direction. In addition, as mentioned earlier, the amount of video data to be transmitted is huge, and therefore a wide antenna bandwidth is essential. Because of the effect of human body tissues on the antenna propagation characteristics, the transmission frequency may shift during the procedure and a wide bandwidth will be able to cover this shift. Antennas with the above characteristics can be used for WCE application. A simple antenna, for example, a helical antenna (Kwak et al. 2005), may be used.

**Lighting system**

The lighting system is composed of white LEDs and a control circuit. The number of LEDs depends on the brightness of the LEDs and their power consumption. A control circuit can also be implemented to turn the LEDs on and off in accordance to the camera’s frame rate, which is similar to the lighting system of the PillCam and EndoCapsule developed by Given Imaging Ltd. and Olympus, respectively.

**SYSTEM IMPLEMENTATION**

Solution 1 was chosen for system implementation and proof of concept, as it was the simplest and most straightforward solution. The analog video capture by an analog CMOS sensor was sent to a TV modulator or an amplitude modulation circuit to generate the RF signal of 800 MHz. The RF signal was then radiated through a helical antenna.
This system was easy to build and visualise, and required the least number of components.

In this prototype, a 1/4 in CMOS analog camera VC-KTK702C from Vtec Electronic Co, China was used. It had 628(H) × 582(V) of resolution with a size of 17.4 mm × 17.4 mm × 20.4 mm. For the RF modulator, a Motorola TV modulator chip, MC44BC373, was used. It could be programmed to radiate frequencies in the range of 30 to 880 MHz, and its output power was −34.5 dBm, which was enough for demonstration purpose. A helical antenna of 8-mm diameter and 5 turns was tuned to receive the signal at 800 MHz. Finally, for the lighting system, 50 mA of LED forward current was used and the duty cycle of pulses was limited to 10% to give an acceptable brightness. The completed system is shown in Figure 9.

SYSTEM VERIFICATION

The system was put inside a plastic case with a transparent dome as shown in Figure 10 so that it could be used as a stand-alone device. The device was then pushed through a cleaned porcine colon. A traditional TV with a straight antenna was placed beside the colon at a distance of 5 cm to pick up the RF signal transmitted out of the colon. For further distances, a power amplifier has to be added to the circuit or a more sensitive receiving antenna has to be used. A few snapshots of the video captured by the prototype are provided in Figure 11. As seen in the snapshots, images are quite clear and they show details of the colon lining. Although the images have a little noise due to the wireless transmission, it can be filtered at the receiving end by circuit technique or image processing technique.

The set-up of the verification step is simple, as its purpose is to demonstrate the concept of the device. Although a more detailed verification plan can be used for verifying the distance of wireless transmission, adequate illumination from the lighting system, etc., these are not the focus of this article.

CONCLUSIONS

This article gives an overview of the system, as well as the possible and potential solutions in developing a WCE prototype. While the commercial WCEs have everything custom made, this article focuses on the technology that can be purchased off the self. Although the prototype is not comparable to the commercial WCEs in terms of its size and optimisation, it serves as a start-up guide and platform for those who wish to work on this area and those who are working to improve it.

Different technologies for the wireless video transmission have been discussed in this article. Analog modules are generally bulky but easy to implement, whereas digital modules are small in size but complex. Four solutions for wireless video transmission were proposed. Solution 1 uses analog cameras, whereas the remaining use digital cameras. For wireless transmission, both solutions 1 and 2 use analog transmission, while solutions 3 and 4 use digital transmission. Implementation of solutions 1 and 2 is easier because of analog video transmission but their sizes would be too bulky for wireless capsule endoscopy, unless custom-made components are used. Solution 3, which makes use of the WLAN technology, has the advantage of very fast data rate, but its size is large because of the circuit complication. Solution 4 makes use of image compression and the circuit can be made a little bit simpler as a smaller transmitter can be used.

Out of the four proposed solutions, solution 1 was implemented. This solution was the simplest one and could demonstrate the concept of WCE. A 1/4-in CMOS analog camera was used for the analog video capture, and a TV modulator chip, Motorola MC44BC373, was used as the video transmitter. It was programmed to transmit at a frequency of 800 MHz. A helical antenna was tuned to radiate at this frequency and a lighting system consisting of four LEDs were built. The resulting circuit was packaged in a circular plastic case with a transparent dome and pushed inside a porcine colon in vitro. A straight TV antenna was placed at a distance of 5 cm from the colon to pick up the video images transmitted from the device, and the video was displayed on a traditional TV. The function of the prototype was verified by viewing the motion pictures that the prototype captured inside the porcine colon on the TV.

In the future, more work can be done in implementing the remaining three solutions and analyse the challenges involved. Power supply and consumption has not been considered in this article, and therefore more research has to be done to investigate the possibility of minimising power consumption of the device, as well as increasing the power capacity of the batteries. Another suggestion for miniaturisation of WCE is to design an ASCII solution for the system, similar to the one suggested by Xie et al. (2004). The ASCII solution will contain an image-processing module, a control unit and a transmitting module. By developing an ASCII solution, the size of the device can be largely reduced, and an extra room can then be used for other functions such as locomotion and actuation.

ACKNOWLEDGMENT

This project is supported by RGC Competitive Earmarked Research Grant #CUHK4213/04E of the Hong Kong Government awarded to Max Meng.

REFERENCES


Submit your manuscripts at http://www.hindawi.com