醫工學會誌：第 14 卷, 第 3 號, 1993
J. of KOSOMBE : Vol. 14, No. 3

研 究 論 文
93-14-3-05

On the Development of a Video Endoscope Having a Swallowable Insertion Tube

Wooshik Kim*, Woochoel Lee*, Byung Goo Min*, & In Sung Song**

=Abstract=

An endoscope is a medical device which observes the inner surface of an organ inside a body. Recently, a new type of endoscope using a CCD image sensor has been developed and turned out to have many strong points over the conventional optical fiber endoscope system. A swallowable insertion tube is an insertion tube having so small a diameter that the head of the insertion tube can reach the end of stomach only by mechanical movement of the esphagus and stomach or the patient's external movement. This paper presents some of the results that the Institute of Biomedical Engineering at Seoul National University has obtained while developing an electronic endoscope having a swallowable insertion tube. After some of the features of the developed system are presented, some of the image processing techniaues are addressed.

1. INTRODUCTION

An endoscope is a device that observes the inner surface of an organ inside a body. This has been developed rapidly since the appearance of optical fibers. Although these conventional endoscopes have an acceptable image quality, they have several shortcomings. First of all, it is not easy to raise the resolution of the endoscopic image. The resolution of the endoscope is largely dependent on the number of the optical fibers in the insertion tube. To have a better resolution, it is necessary to increase the number of the optical fibers. If then, the diameter of the insertion tube will increase accordingly, which may cause some pain to patients. At present, the number of the optical fibers in one insertion tube is approximately 10,000. Thus, the conventional systems cannot have a better resolution without giving more pain to the patient during the test. The second one is that the conventional systems have a short lifetime. As the number of uses of the endoscope increases, the optical fibers in the insertion tube may be cut off or broken so that the number of the pixels decreases. As a result, the insertion tubes should be replaced with new ones on a regular basis. The third one is that, since the images are formed through optical fibers, it is not possible for several people to observe the same image at the same time. Finally, a specially designed camera is required to make a hard copy.

Recently, thanks to the development of the technologies in the semiconductor area, higt-integrated CCD sensors (charge coupled device) were introduced. Using these sensors, a new type of endoscope system, called an electronic endoscope system, was developed. This electronic endoscope has a CCD sensor at the tip of the insertion tube, transforms the

〈접수 : 1993년 3월 18일〉
* Institute of Biomedical Engineering College of Medicine Seoul National University
** College of Medicine Seoul National University
Submitted: March 1993, Revised: June, 1993

optical information into an electrical signal, and sends it to a video display device such as monitor. Usually CCD sensors have a good resolution and therefore, the images coming out of an endoscope are high-quality images, which enable doctors to make more accurate diagnoses and even early diagnoses. Since the information is sent to the display device via several electric lines, the diameter of the insertion tube may be made very small, which means in turn that the pain that a patient may feel may be avoided. In addition to these, since all the parts in the electronic endoscope are reusable, the electronic endoscope system can last longer than the conventional endoscope system. Thus, it is our expectation that the conventional endoscope market will be replaced with the electronic endoscope system in the near future.

An insertion tube is a part of the endoscope that is inserted into a patient's body and transduces optical information into a corresponding electric signal. A swallowable insertion tube is an insertion tube having so small a diameter that the head of the insertion tube can reach the end of stomach only by mechanical movement of the esphagus and stomach and the patient's external movement. To meet this condition, the body of the insertion tube should have a small diameter although its head might have a large diameter. Since the insertion tube has a small diameter, usually the tube does not have a similar controllability as the conventional insertion tube does. To compensate these uncontrollability, there are two things to be considered. First of all, it is desirable that the swallowable endoscope should have an ability to get as much information as possible at one point, which may be solved by implementing a lens system having a wide angle of view. The other is that a powerful software needs to be developed to collect and re-organize all the information.

Exploiting these shortcomings of the conventional endoscope system, possible strong points of the electronic endoscope systems, and the new concept of the swallowable endoscope, the institute of Biomedical Engineering at Seoul National University(abbreviated to the IBMESNU from now on) launched the development of a new endoscope system having a swallowable insertion tube. This paper presents some of the preliminary results after an experimental endoscope was completed. This paper organized as follows. First of all, some of general features of the developed system are introduced. Then, some of the image processing techniques that are developed to process the endoscopic imagesae addressed.

2. DEVELOPED ENDOSCOPE SYSTEM

The endoscope system comprises 4 parts; an insertion tube, the endoscope main body, a computer, and a monitor. The insertion tube is to be inserted into a patient's body, sends lights from the main body to the object, collects the reflected lights, and transduces optical information into a corresponding electric signal. The endoscope main body is concerned with the transformation of an electric signal into a corresponding video signal and the generation of light. A monitor is concerned with the formation of a video signal into an image. Finally, the computer part is responsible for processing of the data including archiving, grabbing, or image processing. Figure 1 is a picture of the developed endoscope system.

2. 1 Insertion tube

The insertion tube is composed of two parts: a tube body and a head. The head is composed of optical fibers, a CCD sensor, a lens system, and a holder. The tube body is composed of optical fibers, several electric lines, and a tube that surrounds and protect the optical fibers and the electric lines.

During a gastro-endoscopic test, the insertion tube reaches the end of the stomach of a patient through one's mouth and esphagus. For normal adults, the lengths of the esphagus is about 40cm and the stomach, about 25.4-28cm, respectively. The diameter of the esphagus and the stomach are about 14-22mm and about 10-12cm[1]. Therefore, the diameter of the tube body should be less than

14mm and the length of the tube should be at least 70cm long. The insertion tube in the endoscope system developed in the IBMESNU was designed to be about 75cm long and 6.1mm wide and the head, 8.6mm wide.

2. 2 Lens System

The lens system is one of the most important parts in the video endoscope system. The main function of the lens system is to collect lights and form images. While an endoscope having a swallowable insertion tube has a strong point that the diameter of the insertion tube can be small, it may not have a good way of controlling the insertion tube except dragging it. As a result, the swallowable endoscope should have a capability of collecting as much information as possible at one point. In other words, the lens system should have a large angle of view. The IBMESNU endoscope is designed to have a 120 degrees of angle of view.

Figure 2 is the diagram of the developed lens system. The first and the second lenses are fisheyes to realize a large angle of view. A biconvex lens is used to collect and focus the dispersed light coming out of the first two lenses. Since the rays passing through these three lenses reveals several aberrations such as comatic aberration, spherical aberration, or chromatic aberration, a doublet(i.e., combination of a bioconvex and a biconcave lens) is inserted between the second lens and the biconvex lens to correct the abberations. Also, a 1mm aperture is inserted between the 4th and the fifth lens to mitigate the effect of the uncorrectible aberrations. The curvatures and the positions of the lenses and the aperture are given in[2].

3 IMAGE PROCESSING

Figure 3 and Figure 4(a) are a picutre of a rectangular grid and its endoscopic output, respectively. Broadly speaking, the images coming out of the endoscope have several features. 1) The endoscopic images have vertical strips and a narrow

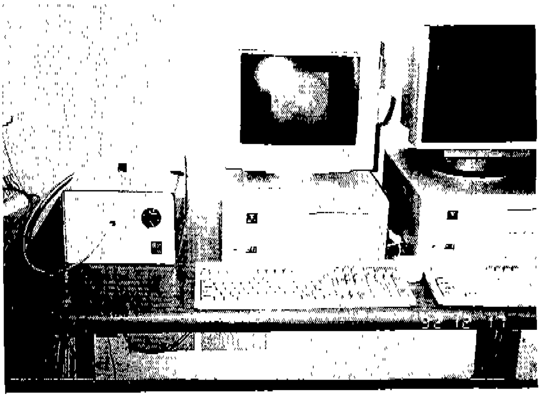


Fig. 1 The endoscope system developed in the IBMESNU

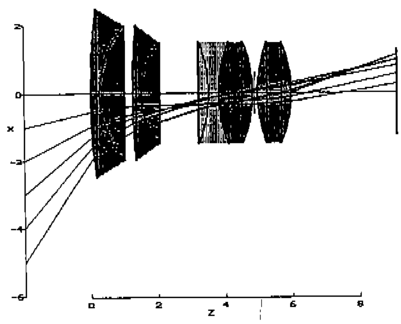


Fig. 2 The designed lens system for the endoscope developed in the IBMESNU

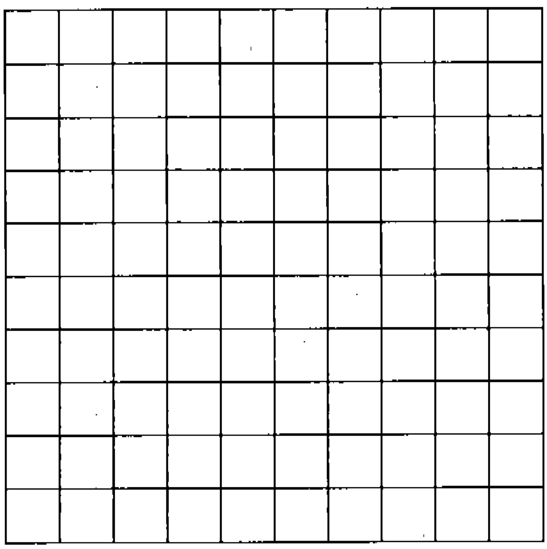


Fig. 3 Grid used for evaluating the developed lens system.

histogram. 2) The endoscopic images are distorted (warped) and blurred. In this section, some of the approaches developed to solve these problems are addressed.

3. 1 Removal of strips and histogram expansion

Strips are put into the endoscopic image while scanning the charges stored in the CCD and making corresponding video signals. To remove these strips, the image model of an endocopic image are assumed to be of the following form

$$y(m,n)=l(m,n)+n(m,n)$$

where $y(m,n)$is the output image of the endoscope, $l(m,n)$ is the output image of the CCD sensor only, and $n(m,n)$ is its noise component. Generally, an output of a system l(m,n) can be written as

$$l(m,n)=\mathrm{T}\{\sum_k\sum_r \chi(k,r)\delta(m-k,n-r)\}$$

where T{ * }is the system response the system. Since the developed system is linear and space-variant[2], the equation above can be simplified as

$$l(m,n)=\sum_k\sum_r \chi(k,r)h(m,n;k,r)$$

where $h(m,m;k,r)$ is the response of the system to the input $\delta(m-k,n-r)$. In a general sense, noise is represented as the signal components other than the desired signal. The nolse component $n(m,n)$ is thought to have two components: the random noise $nr(m,n)$ and non-random noise $nr(m,n)$cluding the strips. The random noise has a stochastic property so that a realization at a momemt has no physical meaning. Instead, only statistical quantities such as mean, standard deviation, or probability density, have its special meaning. On the other hand, non-random noises have deterministic properties such that, if their characteristics are exactly derived, then these noises may be removed. If a non-random noise is additive, then it can be removed by subtraction. One example is in the digital subtraction angiography, where the noise component such as bone or soft tissue is measured and subtracted so that the only image of blood vessels can be extracted[3].

Since the nonrandom noise has the same pattern along the vertical axis, the average along the vertical axis provides the nonrandom component only due to the zero mean property of random noise. *Let* $f_n(m)$ *be*

$$f_n(m)=y(m,n) \ for \ n=0,1,\cdots, N-1$$

where N is the number of rows. Then, this can be written as follows

$$f_n(m)=l_n(m)+n_{rn}(m)+n_{nrn}(m)$$

where $n_{rn}(m)$ and $n_{nrn}(m)$ are the n-th row of the random noise and non-random noise component, respectively. Roughly speaking, $l_n(m)$ may be assumed to have a random distribution with respect to the mean value and $n_{rn}(m)$, with respect to 0 with a uniform distribution. Thus, the non-random noise component can be obtained from the average of $f_n(m)$ along the vertical axis, i.e.,

$$n_{nrn}(m)=\sum_{n=0}^{N-1} f_n(m)/N-average\{l_n(m)\}$$

Figure 4(b) is the picture of the average of Figure 4(a) along the vertical axis. As can be seen in this picture, the strips in the endoscopic images appears as dents in the picture and, if the strips are removed properly, then the dents would disappear, accordingly.

Since there is no light inside a body, it is necessary to enlighten the object. Since the light is transported to the tip via optical fibers, it is not easy to supply light having enough intensity and therefore, it is necessary to focus the light. As a result, the endoscopic images are usually very sensitive to the level of focusing, and have a narrow histogram. The histogram, however, can be expanded from 0 to 255 through the histogram expansion method. The mathematical formula of the histogram expansion is given as

$$\tilde{y}(m,n)=\frac{y(m,n)-f_{min}}{f_{max}-f_{min}} * 255$$

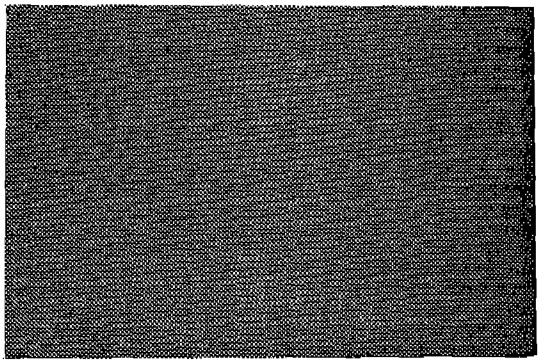
where f_{max} and f_{min} are given as

$$f_{max} = max\{y(m,n)\}$$
$$f_{min} = min\{y(m,n)\}$$

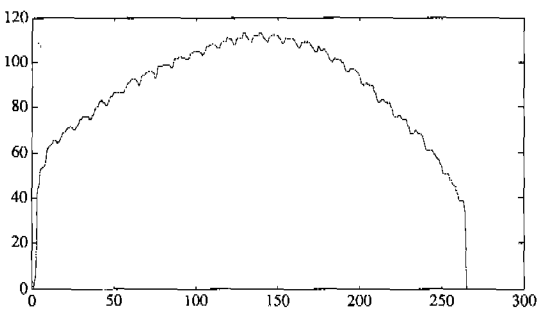
for all $(m,n)\in R[M,N]$. Figure 5(a)is the picture of Figure 4(a) after its strips are removed and histograms are expanded. To get Figure 5(b). $\Sigma_{n=0}^{N-1} f_n(m)/N$ is subtracted from Figure 4(a) and a constant (in this case 128) is added to make all the pixels nonnegative. Figure 5(b) is the average of Figure 5 (a) along the vertical axis. As is expected, the dents that appeared in Figure 4(b) disappeared. The flatness is due to the subtraction of the average term of the original signal $average\{l_n(m)\}$.

3. 2 Dewarping

The images coming out of the endoscope have a special optical characteristics. In particular. the endoscope developed in the IBMESNU uses fisheye lenses to have a large angular field of view, which results in distortion of the output images. This distortion, however, may be corrected by a software approach. Figure 6 is a diagram which shows the relationships among the lens object plane(r plane) and the CCD plane (R plane) in the lens system

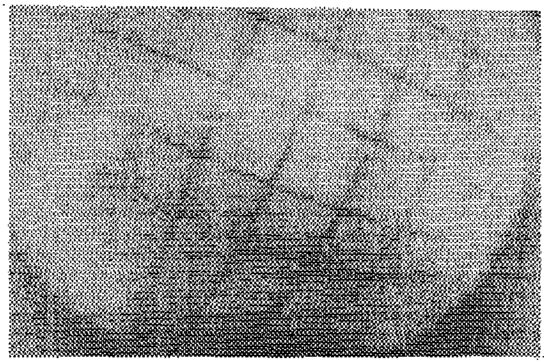


(a)

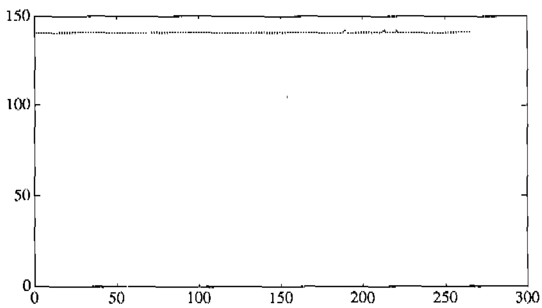


(b)

Fig 4 Image obtained through the endoscope:(a) An endoscopic image of the grid Figure 3 and (b) The avrage of (a) along the the vertical axis



(a)



(b)

Fig. 5 The de-stripped and histogram expanded image of Figure 4 : (a) The de-stripped and histogram expanded image of Figure 4 (a) and (b) the average of (a) along the vertical axis and the real object plane(z-plane) and the image plane(Z-plane) on the endoscope. The measured image is in the image Z-plane. Unfolding the distorted image is the same as mapping the pixels in the distorted image to the corresponding pixels in the real object plane (z-plane). The formula that relates the pixels in the real object plane and the image plane may be derived from the realtionship between coordinates of the object plane and the CCD plane in the lens system, which will be referred to as the parameter estimation procedure. The formula needs to be scaled and translated to comply with the real plane, which is called the coordinates mapping procedure. In the interpolation procedure, the values of the pixels that are omitted during the coordinate mapping procedure are calculated from the surrounding pixels.

(1) Parameter Estimation

Since the lens system is circularly symmetric, let $re^{j\theta}$ and $Re^{j\varphi}$ be the coordinates of the object plane and the CCD plane, respectively, and $R = f(r)$ be the relationship between the two coordinates. To dewarp the distortion, it is necessary to find the inverse relationship $r=f^{-1}(R)$. The inverse function may be approximated as a third-order polynomial,

$$\tilde{r}=f^{-1}(R)=aR+bR^2+cR^3$$

and the error function to be minimized is defined as follows

$$Err=\Sigma[r-\tilde{r}]^2=\Sigma[r-aR-bR^2-cR^3]^2$$

To find the values of a, b, and c that minimize the error, the error function is differentiated with respect to a, b, and c;

$$\frac{\partial Err}{\partial a}=\frac{\partial Err}{\partial b}=\frac{\partial Err}{\partial c}=0.$$

From this and the data from Table 1, a matrix equation can be derived and, by solving this equation, the polynomial is given as

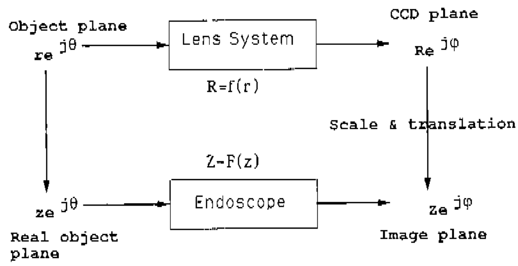


Fig. 6 Relationship among the object plane, the CCD－plane, the real object plane, and the endoscopic image plane

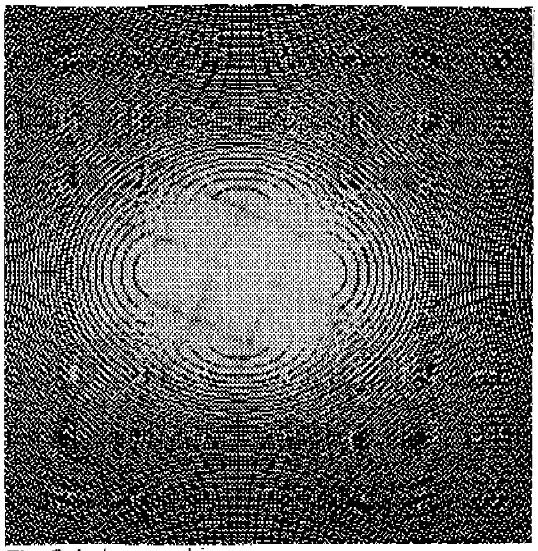


Fig. 7 A dewarped image

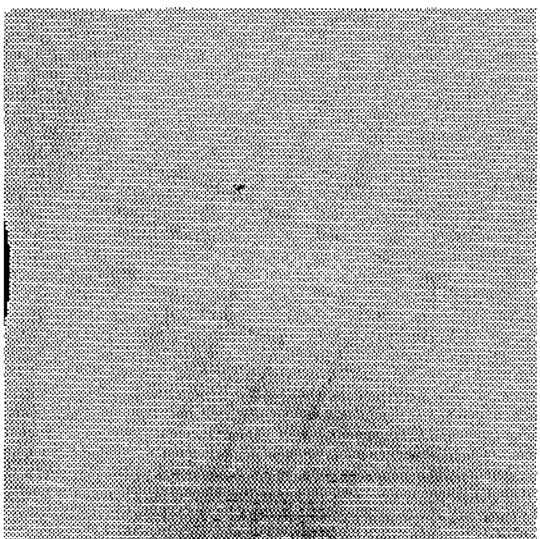


Fig. 8 An interpolated image of the dewaped image

$$\tilde{r} = 3.3863R - 1.0662R^2 + 2.0426R^3$$

and the minimum error value at this time is given as 0.113237.

(2) Coordinate mapping

The next step is to find the relationships among the lens object plane, the real object plane, the CCD plane, and the image plan. The coordinate of 6.0 on the object plane in the lens object system maps to the 1.1689 point on the CCD plane which, in turn, corresponds to about 100 pixels from the origin on the endoscopic image plane and should be about 120 pixels from the center on the real object plane. Fom these relationships,

$$Z = 85.6 \times R$$
$$z = 20 \times r.$$

Using these equations, the formula

$$z = .7912Z - 2.910 \cdot 10^{-3}Z^2 + 6.513 \times 10^{-5}Z^3 \qquad (1)$$

is derived. Figure 7 is a reconstructed image using this formula. As is seen from the image, some of pixels are missing. These pixels do not have their corresponding pixels and their values need to be calculated through interpolation.

(3) Interpolation

The interpolation can be done as follows. First of all, from the real object plane, the coordinate of the missing pixels, say(x,y), is determined. To find the corresponding point in the endoscopic image plane, the distance and the angle of the underlying point from the center of the real object plane is derived as follows

$$z = \sqrt{(x - x_C)^2 + (y - y_C)^2}$$
$$\tan\theta = (y - y_C)/(x - x_C)$$

If (X,Y) is referred as the corresponding coordinate of (x,y) on the image plane, then (X,Y) is given as

$$X = Z\ \mathrm{con}\theta + X_C$$
$$Y = Z \sin\theta + Y_C$$

where (X_C, Y_C) is the center on the endoscopic image plane and Z is the distance between the center and (X,Y) which is derived from the inverse formula of Equation 1.

Note that X and Y may have fractional values. If the integer parts of X and Y is let [X]and [Y], respectively, and their fractionals part ΔX and ΔY, respectively (i,e., $X = [X] + \Delta X$ and $Y = [Y] + \Delta Y$), then the pixel value of the desired point (x,y) is given as

$$p(x,y) = p([X],[Y])\Delta X \Delta Y + p([X]+1,[Y])(1-\Delta X)\Delta Y + p([X],[Y]+1)\Delta X(1-\Delta Y) + p([X]+1,[Y]+1)(1-\Delta X)(1-\Delta Y)$$

where p(m,n) is the pixel value of the point (m,n). Figure 8 is the final dewarped image.

r	tan θ	R
0.0	0.0	0.000
-1.0	0.22	0.2528
-2.0	0.45	0.5434
-3.0	0.64	0.7889
-4.0	0.76	0.9520
-5.0	0.83	1.0982

Table 1 The coordinate of the inputrays in Figure 2 and their corresponding coordinates on the CCD plane.

4. DISCUSSION/CONCLUSION

To test the performance of the developed endoscope system, two animal experiments were performed. In the first animal experiment, the performance of the developed system was proved to be good in gerneral although two problems were pointed out: the one is that the surface of the insertion tube was sometimes contaminated by strange material and the other is that sometimes the head of the insertion tube is stuck in the stomach wall so that the insertion tube would no longer proceed. To solve the first problem, a biopsy channel was attached along the insertion tube and a 25cc air baloon was put on near the head of the insertion tube to solve the second problem. With this, the second animal experiment was made and it was confirmed that the problems were solved.

In conclusion, this paper considered the development of a video endoscope having a swallowable insertion tube using a CCD sensor. Some of the basic researches including the development of the lens system, a swallowable insertion tube, a CCD sensor driver, as well as some of the techniques for processing endoscopic images. In the Institute of Biomedical Engineering at Seoul National University, the upgrade of the developed system using a color CCD is under progress. It is our hope that the know-hows that have been accumulated while developing the video endoscope having a swallowable insertion tube will be stepping stones to the development of other similar devices such as endoscopic sonography or others.

Acknowledgement

This work is supported in part by Korea Science and Engineering Fourdation under contract #8912-0503-001.

REFERENCES

1) D.C. Sabistion. "*Textbook of Surgery*". Saunder's Company, 1982.
2) The Institute of Biomdical Engineering, College of Medicine, Seoul National University. "*Development of a video endoscope and a swallowable insertion tube*". Korea Scince and Engineering Foundation, 1993. Final Report.
3) J.R. Cameron and J.G. Skofronick. "*Medical Physics*" A Wiely-Interscience Publication, John Wiley & Sons, 1978.
4) J.R. Meyer-Arendt "*Introduction of Classical and Modern Optices*". Prentice Hallm, engleweed Cliffe New Jersey 07632, 1989.
5) The Institute of Biomedical Engineering, College of Medicine, Seoul National University. "*Development of video endoscope and a swallowable insertion tube*". Koera Science and Engineering Foundation, 1992. Second year final report.
6) Stellar Software. "*Beam 4: Optical Ray Tracer*", Technical Manual, 1989.
7) W.J. Smith. "*Modern Lens Design: A Resource Manual*". Genesee Optica Software, Inc., 1992.
8) E.L. Hall. "*Computer Image Processing and Recognition*". Academic Press, 1979.
9) H.C. Andrews and B.K. Hunt. "*Digital Image Restoration*". Prentice Hall, 1977.
10) A. Rosenfield and A.C, Kak. "*Digital Picture Processing: Volume 1*". Academic Press, 1982.
11) R.L. Weber. "*Small-area CCD Imager Evaluation using television Monitors*". Application Report, Texas Instruments, May.