Object Dimension Estimation for Remote Visual Inspection in Borescope Systems

Hyun-Sik Kim* and Yong-Suk Park
Contents Convergence Research Center, Korea Electronics Technology Institute
Seoul, 03924 - Republic of Korea
[e-mail: {hskim, yspark}@keti.re.kr]
*Corresponding author: Hyun-Sik Kim

Abstract

Borescopes facilitate the inspection of areas inside machines and systems that are not directly accessible for visual inspection. They offer real-time, up-close access to confined and hard-to-access spaces without having to dismantle or destructure the object under inspection. Borescopes are ideal instruments for routine maintenance, quality inspection and monitoring of systems and structures. The main application being fault or defect detection, it is useful to have measuring capability to quantify object dimensions in a target area. High-end borescopes use multi-optic solutions to provide measurement information of viewed objects. Multi-optic solutions can provide accurate measurements at the expense of structural complexity and cost increase. Measuring functionality is often unavailable in low-end, single camera borescopes. In this paper, a single camera measurement solution that enables the size estimation of viewed objects is proposed. The proposed solution computes and overlays a scaled grid of known spacing value over the screen view, enabling the human inspector to estimate the size of the objects in view. The proposed method provides a simple means of measurement that is applicable to low-end borescopes with no built-in measurement capability.

1. Introduction

Endoscopes are slender, tubular optical tools used to view areas that would otherwise not be visible. They are inserted into the item being evaluated through a small opening, using minimally invasive procedures, leaving the item of interest intact. Endoscopic procedures applied in industry have greatly reduced the cost and time for inspection. The term "endoscope" usually refers to instruments used inside the human body for medical purposes. For industrial use, similar instruments are often referred to as "borescopes". By providing real-time images of difficult-to-access areas, borescopes have become indispensable for industrial remote visual inspections. Borescopes are used to non-destructively inspect systems and equipment for condition, quality, security, and safety. Borescopes also enable in-field inspection of safety critical systems or components in their natural usage environment [1]. They are used for intial manufacturing and routine maintenance in various areas of industry, such as aerospace, automobile, power generation, housing and public works, etc., to quantify defects and help determine fitness-for-operation [2].

Defect damage viewing and quantifying has become a key function in endoscopic remote visual inspection for quality control. Recently, deep learning techniques are also being applied to images aquired during borescope inspection for damage detection [3]. The ability to quantify inspection data enables higher levels of system performance. Size and distance information of objects in view can be obtained from stereo imaging and applying triangulation techniques. In order to capture images from different viewing angles, stereo imaging techniques use multi-optic solutions [4, 5]. Many high-end borescopes integrate multi-optic tip adaptors for increased performance and measurement accuracy. However, multi-optic solutions accompany structural complexity and system cost increase, making them impractical to be applied for low-end borescopes. Entry-level, low-end borescopes that use single cameras usually do not provide built-in measurement functionality. Inspectors that use low-end borescopes rely on their own eyes and experience to estimate the size of the object or defects discovered. Therefore, it would be beneficial to have a single-optic solution that enables measurement of objects in view. However, due to the single optical path structure of the borescope, finding the size and working distance of objects viewed is a challenge. In single-optic borescopes systems, images from the observations are generally planar with a large depth of focus, not providing the size and distance information of the viewed objects. The working distance is often unknown and can constantly change during the examination. In addition, it is difficult to have a reference data or a reference point useful to get a length scale.

In this paper, a method is proposed that enables inspectors to estimate the size of objects in view from a borescope with a single camera. The size of the objects are determined using a scaled grid of known spacing that is overlaid over the view image. The dimension of the overlaid grid is adjusted accordingly as the working distance from the tip of the borescope to the object of interest changes. In Section 2, current methods for endoscopic object measurement on are reviewed. Section 3 provides details on the procedures involved for the measurement grid generation. Experimental evaluation results of endoscopic measurements based on the proposed method are presented in Section 4. Concluding remarks and future work are addressed in Section 5.

2. Related Work

Borescope systems used for remote visual inspection consist of a lens and illuminating light source such as fiber optics or LEDs. CCD or CMOS image sensors are usually embedded in borescopes to capture images. Borescopes are connected to computing platforms that can process the images captured for quantification and measurement of objects or defects in view.

Reference comparison method for measurement uses the dimensions of a known object in the inspection image as a reference to infer the dimensions of unknown objects located in the same view, as shown in **Fig. 1**. Dimensions for the known and unknown objects in the image are measured in pixels. From the actual size of the reference known object, the pixels per metric proportion can be obtained. This proportion can then be used to obtain the actual measurements of the unknown objects. The comparison method is only useful when there is a reference object with known dimensions in the same view and plane. The reference object may be placed by the inspector using a probe which introduces additional operational complexities.
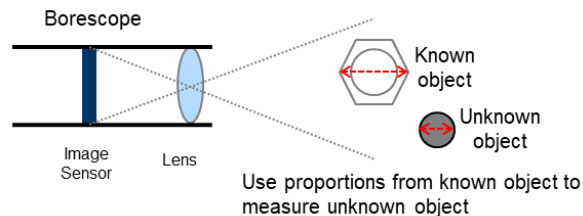

Fig. 1. Reference object comparison method for measurement

In stereo methods, the left and right views of the object under observation are captured and processed for measurement [6, 7]. The borescope embeds a prism or dual lens in its tip which splits the image at a precise angle of separation, as illustrated in **Fig. 2**. The left and right views created are used to obtain relative depth information. Using triangulation geometry on the user indicated point of interest, measurements can be obtained. Stereo imaging typically requires multiple optical channels or other elaborate techniques to capture images of an object from different viewing angles, increasing the endoscope diameter, structural complexity, and system cost.
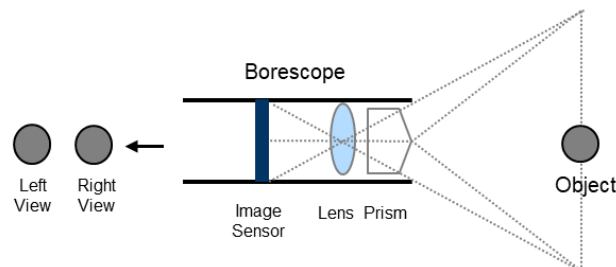

Fig. 2. Stereo imaging method for measurement

In optical phase shifting based measurement, a borescope projects sinusoidal shadow patterns onto the surface which are analyzed to produce a 3D map of the surface. This is based on optical metrology technique known as phase-shifting interferometry [8]. **Fig. 3** shows an illustration of a borescope with optical phase shifting. It involves sequentially projecting a

series of structured light patterns onto a surface, capturing on camera an image of each pattern on the surface. Processing the pattern images or interferograms produces a 3D surface map. Although this method provides high precision and accurate measurements, the phase projection and measurement units needed makes the borescope tip bulky and the device expensive.
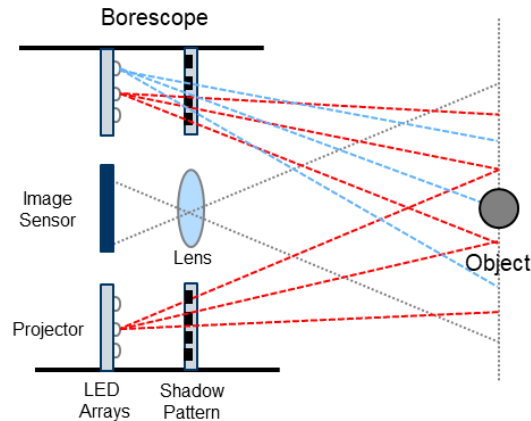

Fig. 3. Optical phase shifting method for measurement

3. Grid Generation for Object Dimension Estimation

3.1 Overview

Current high-end borescopes enable relatively accurate measurement of objects or defects using multi-optic solutions. These solutions accompany increase in borescope tip size, which may limit their use in small, narrow spaces. In addition, the manipulation and operation of borescopes with added measurement capabilities is complex. It is hard to pinpoint the measurement location with precision, and the operator requires training and expertise.

Machine learning could automate the task of measurement in borescopes by recognizing objects in the scene and compute their dimensions. However, in visual inspection scenarios, the objects that need to be identified are defects or foreign debris, which makes machine learning challenging [9]. Most of the times, the object that you are looking for is uncertain, and the objects are hard to classify or categorize.

The design considerations for the proposed borescope measurement solution are as follows. First, the solution should be applicable to single camera borescope systems. Using single camera borescopes reduces operational and structural complexity. Second, the measurement solution should be affordable and inexpensive. Using an image processing based software solution eliminates the need for expensive parts and hardware. In addition, it makes the solution reusable and applicable to other models with minor calibration. The design considerations make a simple measurement functionality available to low-end borescopes, leveraging their capability.

Fig. 4 illustrates the proposed solution that enables measurement in borescopes for visual inspection. Using single vision camera optics, the size of a reference object in the plane of interest can be determined. From the actual size of the object and its corresponding image size in pixels, the pixel to metric or the correspondence of the pixel to actual measurement is

computed. The pixel to metric is then used to construct a grid with known spacing. The grid is overlaid to the view, and the inspector can use the grid as a reference to determine or estimate the size of the object of interest. The grid can be fine-scaled, if higher accuracy is required.
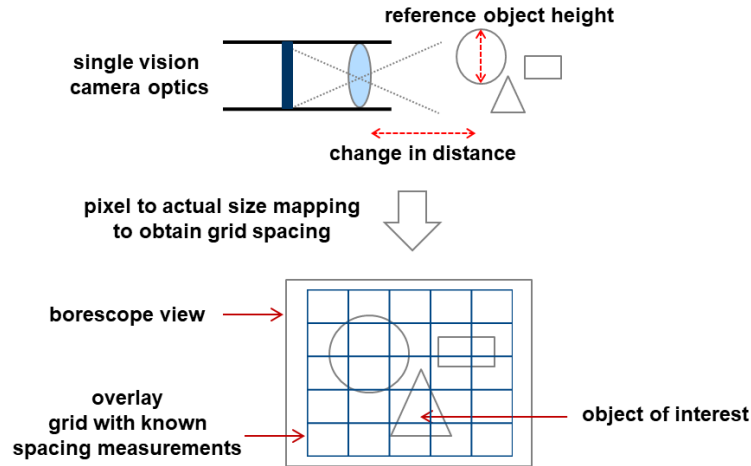

Fig. 4. Proposed object dimension estimation method based on grid overlay

Fig. 5 shows the basic block diagram of the proposed method. The Image Acquisition block consists of the conventional borescope to obtain the endoscopic image. The Distance Measurement block measures the working distance, i.e., the distance from the tip of the endoscope to the object plane. For this particular implementation, single vision camera optics are used for distance measurement. The Grid Scaling block adjusts the dimension of the grid scale depending on the measured distance. The object displayed over the endoscopic image changes in scale based on the working distance. Therefore, the grid needs to be scaled accordingly for accurate measurement. The Image Grid Overlay block overlays the scaled grid over the original endoscope image. The size of the object from the endoscope image can be determined by comparing the object with the size of the grid scale on the endoscopic image.
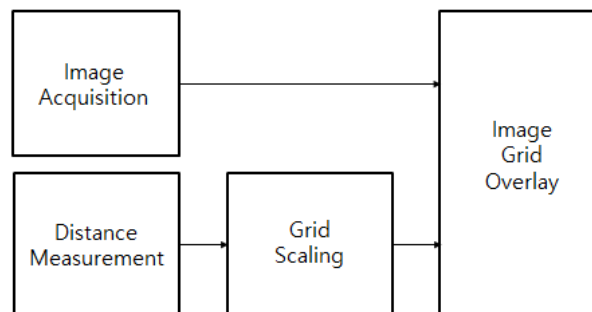

Fig. 5. Grid scale based measuring block diagram

Table 1 provides a comparison of the existing borescope measurement methods and the proposed method. The proposed method provides a way to make estimated measurements unlike other methods that provide actual measurements. However, the proposed method provides simple means of operation and is cost-effective. Therefore, it is particularly suited for low-end endoscope systems with no built-in measurement functionality and that do not require

high accuracy measurements.

<p style="text-align:center">Table 1. Comparison of measurement methods for borescopes</p>

Methods	Dependencies	Type	Tip Size (mm)
Comparison	Distance to object Perpendicularity	2D (height, width)	5
Stereo Image	Prism separation distance Image sharpness & contrast Distance to object	2D or 3D (height, width, depth)	4 to 8
Optical Phase Shifting	Distance to object	3D (height, width, depth)	6
Proposed	Distance to object Perpendicularity	2D (height, width)	5

3.2 Single Vision Camera Optics

Single vision camera optics are used to obtain the parameters necessary for grid computation. **Fig. 6** shows the optical imaging diagram for a single vision camera (i.e. borescope). The left side represents the camera (borescope) where the lens and the image sensor are located. The lens is at a distance f away from the image sensor. This is the focal length, which is required to focus the image on the image sensor. The right side represents the scene of interest where the observation target (object) is located. The object has height h and located at point x. The distance from the borescope lens and the object is d. This creates an image at the image sensor with height a.
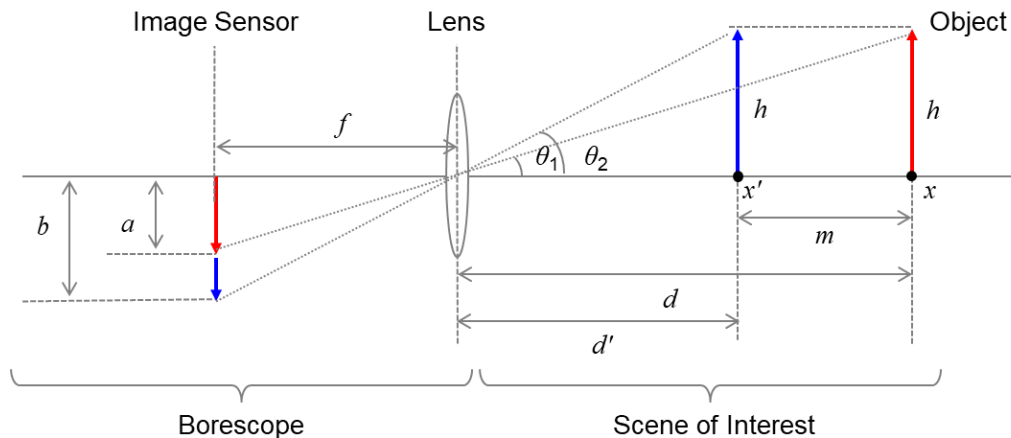

<p style="text-align:center">Fig. 6. Single vision camera (i.e. borescope) imaging model</p>

Moving the borescope towards the object at point x changes the distance between the object and the borescope to d'. This is equivalent to moving the object from point x to point x' by a distance of m, to make the distance between the object and the borescope to d'. When the distance between the object and the borescope changes to d', the height of the image on the image sensor changes to b. The angle created between the height h of the object at point x and the adjacent distance d to the lens is θ_1. The same angle θ_1 is also created between the projected image height a and focal length f. The relationship among h, a, d, f and θ_1 can be established

from trigonometry as shown in (1).

$$\tan\theta_1 = \frac{a}{f} = \frac{h}{d}$$

(1)

The angle created between the height h of the object at point x' and the adjacent distance d' to the lens is θ_2. The angle between the projected image height b and focal length f is also θ_2. Distance d' is equivalent to $d - m$. Equation (2) shows the relationship among h, b, d, f, m, and θ_2 based on trigonometry.

$$\tan\theta_2 = \frac{b}{f} = \frac{h}{d-m}$$

(2)

From equations (1) and (2), the distance d can be derived as shown in (3). Equation (3) shows that the distance from the borescope to the object can be determined by moving the borescope by a distance m and finding the corresponding rate of change a/b in the image sensor.

$$d = \frac{m}{1 - \frac{a}{b}}$$

(3)

The equation for focal length f can also be obtained from equations (1) and (2) as shown in (4). The focal length f of the borescope is a static value. Therefore, if distance d is found from (3) for a reference object with known height h, the focal length can be easily determined from equation (4). Complex calibration procedures using the camera's intrinsic and extrinsic parameters are not necessary.

$$f = \frac{ad}{h} = \frac{b(d-m)}{h}$$

(4)

Equation (5) shows the formula for the height of the target object h, which is obtained from equations (1) and (2). If the focal length f is obtained from (4) using a reference object of known height, f can be used as a constant value in (5).

$$h = \frac{ad}{f} = \frac{b(d-m)}{f}$$

(5)

Therefore, by using single vision camera optics, the distance from the object to the borescope can be obtained, which in turn is used to find the focal length and actual height of the object.

3.3 Contour Detection

In order to figure out the height or size of an object, we need to get its pixel dimensions. To determine the pixel size of the object in the image, contour detection is used. **Fig. 7** shows the steps involved in contour detection.

Fig. 7. Steps involved in contour detection

After loading the image of interest, the image is converted to grayscle and smoothed using the Gaussian filter. The image is smoothed (blurred) to reduce image noise. This is a pre-requisite for edge detection, since edge detection is susceptible to noise. Edge detection identifies discontinuity points in the image at which the image brightness changes sharply. The points with pixels where the intensity changes significantly are turned on, while the others are turned off. Diverse edge detection algorithms may be used, such as Canny edge detection algorithm [10].

For the edge pixels present in an image, there is no particular requirement that the pixels representing an edge are all contiguous. Contours on the other hand can be described as a curve joining all the continuous points along the boundary, having same color or intensity. Therefore, after edge detection, dilation and erosion are performed on the image to close gaps in between object edges. Dilation adds pixels to the boundaries of objects in an image. Erosion removes pixels on object boundaries. From the contiguous edge map created, cotours can be extracted. Contours are abstract collections of points and / or line segments that represent the shapes of the objects found in an image. They are a useful for shape analysis and object detection and recognition.

3.4 Grid Generation

After detecting all contours present in the image, the area of the contours detected are computed. The area can be calculated using the Green's theorem [11]. After calculating the area of all contours, the areas are compared to find the maximum contour. As an example, in **Fig. 8**, the contour marked with a circle is used as the reference since it has the largest contour area.
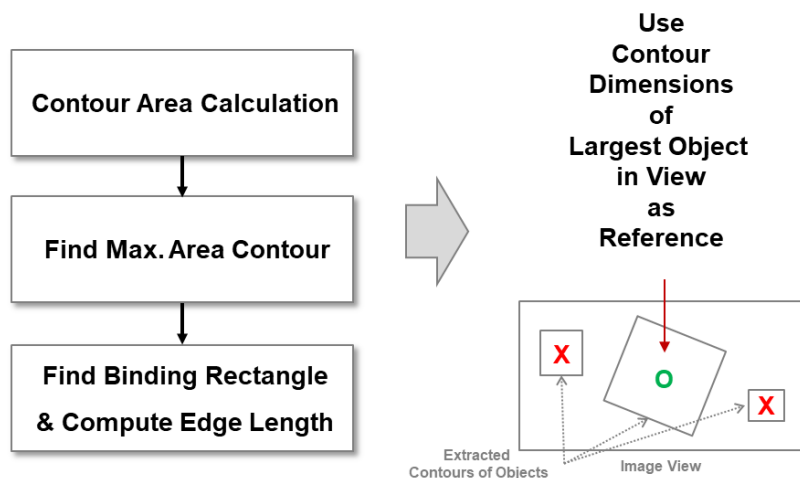

Fig. 8. Finding the maximum area contour to be used as reference object

After finding the largest contour, the rotated binding box or rectangle is computed for the largest contour. The midpoints of the binding rectangle's edges are computed, and by connecting the midpoints the maximum width and height of the contour are obtained.

By taking two images at a distance of m apart, and calculating their maximum contours and respective width and height, heights a and b from equation (3) can be obtained, as shown in **Fig. 9**.
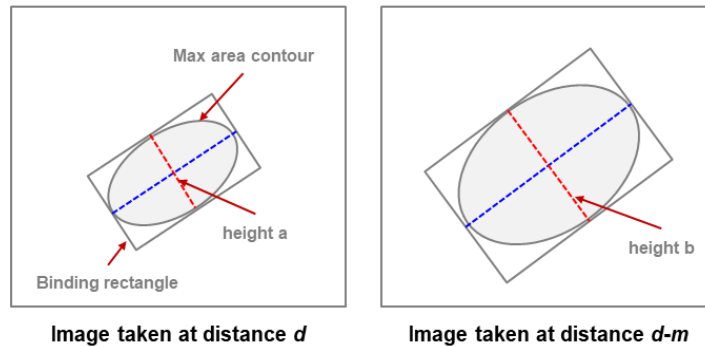

Fig. 9. Determining heights a and b from equation (3) from images taken m distance apart

Using equation (3), the distance d from the borescope to the reference object can be calculated. After obtaining d, equation (5) can be used to calculate the actual height h of the reference object. The actual and pixel heights can be used to calculate the pixel per metric ratio r_{ppm} which is defined as follows:

$$r_{ppm} = \frac{h_{pixel}}{h_{real}}$$

(6)

where h_{pixel} is the reference object contour pixel height and h_{real} is the actual height of the reference object. More specifically, h_{pixel} will be the values of height a or b and h_{real} is the value of heigh h in equation (5). For example, if $h_{pixel} = 90$ px and $h_{real} = 15$ mm, then $r_{ppm} = 6$ px/mm. Grids lines could be generated every 60 pixels to mark spacing of 1 cm. The grid spacing can be fine scaled as necessary. **Fig. 10** illustrates how the generated grid can be used to determine the size of the objects in the image view.
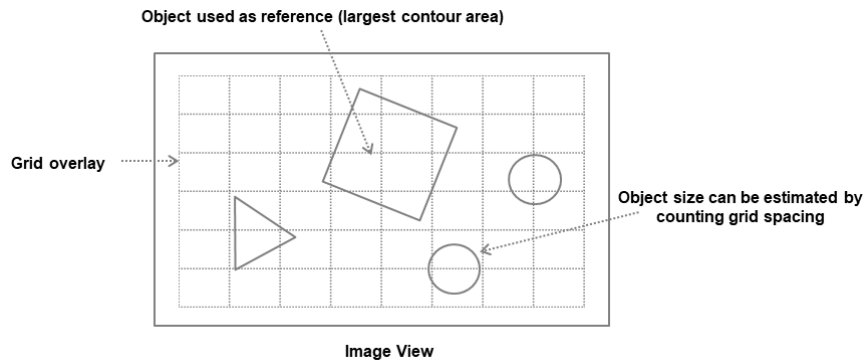

Fig. 10. Overlaying the grid over the image view to estimate the size of objects

4. Experimental Results and Analysis

Experiments were conducted using a low-cost, single vision camera USB borescope with no built-in measurement capability and a laptop as shown in **Fig. 11**.

Fig. 11. Experiment setup showing the borescope, measuring software, and visual inspection target

The borescope contains LEDs for illumination and adjustable LED brightness switch. The resolution of the borescope camera is 1600x1200. The total length of the borescope including the cable is 2 m, and the scope head diameter is 5.5 mm. The laptop is used to process the images sent by the borescope via USB interface. The laptop runs Windows operating system with Intel quad-core processor running at 2.50 GHz and 8 GB of RAM. The measuring software program uses OpenCV 3.4 software library to process computer vision related tasks. A visual inspection target environment was created to simulate remote visual inspection scenarios.
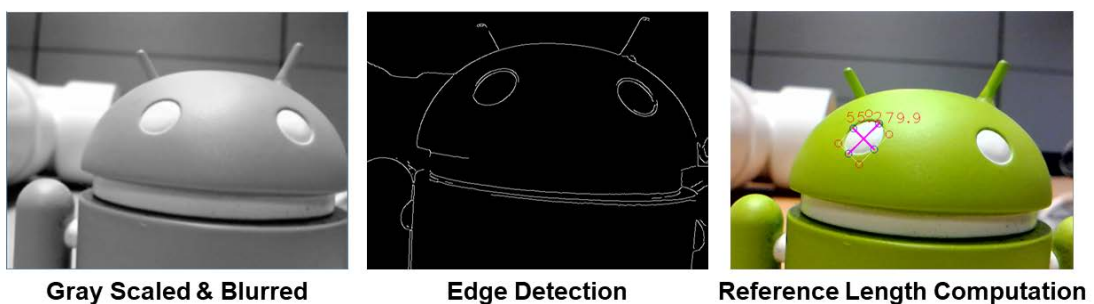

Gray Scaled & Blurred   **Edge Detection**   **Reference Length Computation**

Fig. 12. Reference object (largest contour) pixel width and height determination

Fig. 12 shows the result of reference measurement determination using contour detection. When multiple contours are detected, the contour with the largest area is selected as reference and evaluated to obtain its width and height. The reference object detection worked well in the open environment. However, if the inspection environment contained structures with glossy or

reflective surfaces, contour detection did not work very well. The brightness of the LED light source can be reduced, but at the expense of overall image clarity, so alternative solutions for removing glare need to be applied [12].
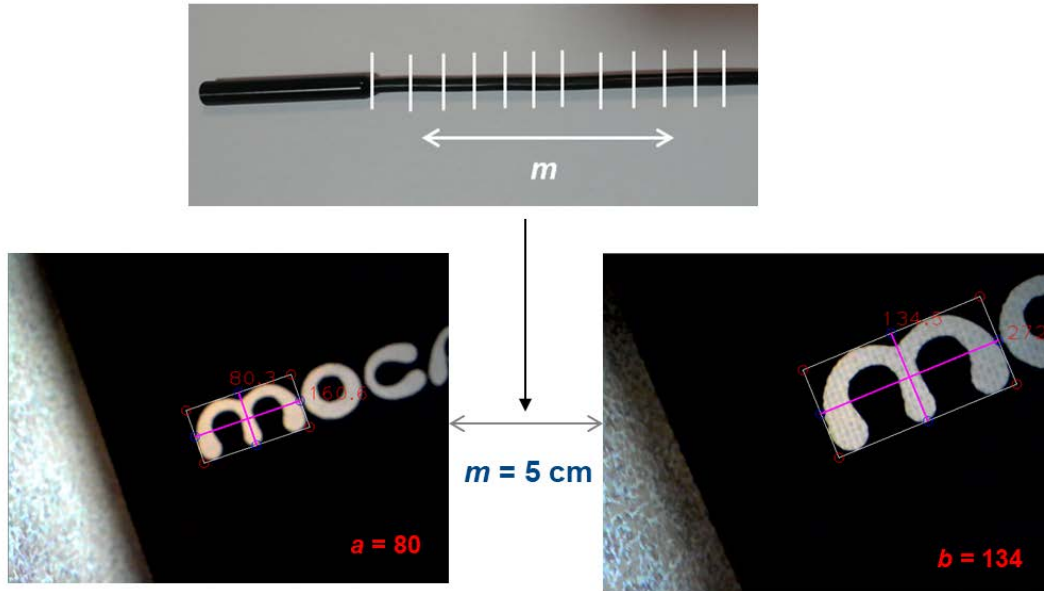

Fig. 13. Testing the accuracy of distance calculation

The single vision camera optics equation (3) used to obtain the distance between the borescope and the reference object was tested as shown in **Fig. 13**. Markers were placed on the borescope every 5 mm to easily determine the distance moved. This gives the value of m in the equation. Two snapshots of the reference object were taken after moving the borescope by m, and the distance was computed. In the example in **Fig. 13**, the distance moved $m = 5$ cm, the height of the reference object at distance d represented by $a = 80$ px, and the height of the reference object at d - m represented by $b = 134$ px. Using equation 3, $d = 12.4$ cm. The actual distance measured was 12.5 cm, which means that the distance calculation has an error rate of 0.8 % for this particular case.

Additional distance calculation results are shown in **Table 2**. The results show that as distance between the reference and the borescope increases, the error rate also increases. When the distance is 94 cm, the error rate is to 4.25%. Since in most cases borescopes operate at close distances, the error rates of large distances are not critical.

Table 2. Error rate of distance calculation

Actual distance (cm)	Computed distance (cm)	Error (%)
12.5	12.4	0.8
30	31	3.33
55	53	3.63
94	94	4.25

The focal length f in equation (4) for the experiment was determined by using a known reference object with dimensions 13 mm x 27 mm. After obtaining the distance value d and

substituting the value of h with the measured height, f can be easily calculated. After calculating f, it can be used to compute the actual height of the object h in equation (5).

Creating the grid overlay becomes a trivial task once the reference object pixel measurements a and b, the distance d, focal length f, and reference object height h are obtained. Knowing a, b, and h, the pixels per metric ratio can be obtained which describes the number of pixels that can "fit" into a given number of metric units, as seen in (6). **Fig. 14** shows the result of grid projection for a view of interest. The grid is scaled according to the changes in distance. The grid lines projected are 5 mm apart. For example, using the grids, it can be inferred that the left most defect in the image is approxmately 10 mm in height and 15 mm in length. As the working distance becomes smaller, the grid is enlarged to reflect the scale change as seen in the right image.
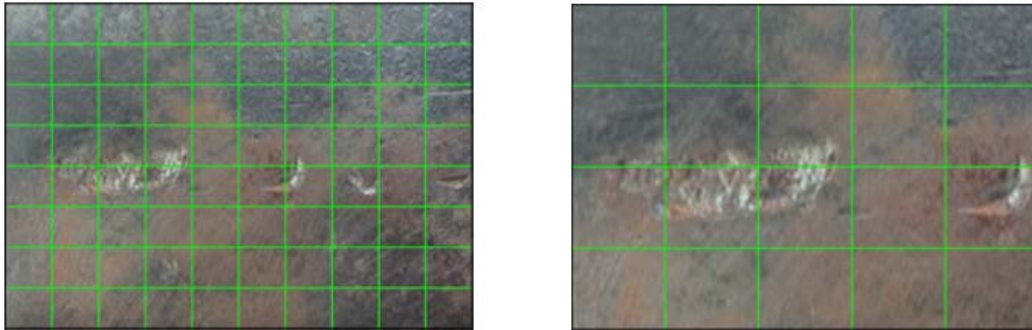

Fig. 14. Grid scaled depending on working distance and projected over image

5. Conclusion

In this paper, a grid scale based measurement method applicable to industrial endoscopy systems for visual inspection is presented. Current high-end endoscope measurement methods provide high-accuracy measurements based on multi-optic solutions. However, these solutions accompany increase in borescope tip size, operational and structural complexity, and system cost. The proposed method provides simple means of on-screen object measurement using grids with scale adjustment based on changing working distance. Using single vision camera optics, the actual size of a reference object in the view of interest can be determined. From the actual size and the pixel image size of the reference object, the pixel to metric ratio is computed. The pixel to metric ratio is used to construct a grid with known spacing. The grid is overlaid to the view, and the inspector can use the grid as a reference to determine or estimate the size of the object of interest. The proposed solution can provide a simple and affordable measuring capability to any single-vision camera borescope system. The software based implementation makes it reusable and easily applicable to other borescope systems with minor calibration. The proposed system is particularly suited for low-end borescopes that do not have built-in measurement capabilities and do not require high-accuracy measurements.

The current solution has been tested for consecutive snapshots of still images. It has not been tested thoroughly for streaming video, so real-time grid adjustment remains as a future task. It has also been observed that if the inspection environment contains structures with glossy or reflective surfaces, contour detection does not work very well. Therefore, solutions for removing glare in images need to be explored as well. In addition, since wide angle lenses

that introduce barrel distortion are widely used in many borescopes for improved visual coverage, lens distortion needs to be taken into account and necessary adjustments need to be made to the grid scale of the proposed system.

References

[1] S. Addepalli, R. Roy, D. Axinte, and J. Mehnen, "'In-situ' inspection technologies: Trends in degradation assessment and associated technologies," *Procedia CIRP*, vol. 59, pp. 35-40, 2017. Article (CrossRef Link).

[2] T. Zhang, Y. Luo, X. Wang, and M. Wang, "Machine vision technology for measurement of miniature parts in narrow space using borescope," in *Proc. of 2010 Int. Conf. on Digital Manufacturing & Automation*, vol. 1, pp. 904-907, 2010. Article (CrossRef Link).

[3] Z. Shen, X. Wan, F. Ye, X. Guan, and S. Liu, "Deep learning based framework for automatic damage detection in aircraft engine borescope inspection," in *Proc. of* 2019 *Int. Conf. on Computing, Networking and Communications* (*ICNC*), pp. 1050-1010, February, 2019. Article (CrossRef Link).

[4] H. Mayer, I. Nagy, A. Knoll, E. U. Braun, R. Bauernschmitt, and R. Lange, "Haptic feedback in a telepresence system for endoscopic heart surgery," *Presence*, vol. 16, no. 5, pp. 459-470, October, 2007. Article (CrossRef Link).

[5] G. A. Triantafyllidis, D. Tzovaras, and M. G. Strintzis, "Occlusion and visible background and foreground areas in stereo: a Bayesian approach," *IEEE Trans. Circuits and Systems for Video Technology*, vol. 10, no. 4, pp. 563-575, June, 2000. Article (CrossRef Link).

[6] H. W. Schreier, D. Garcia, and M. A. Sutton, "Advances in light microscope stereo vision," *Experimental Mechanics*, vol. 44, no. 3, pp. 278-288, June, 2004. Article (CrossRef Link).

[7] W. Choi, V. Rubtsov, and C. J. Kim, "Miniature flipping disk device for size measurement of objects through endoscope," *J. of Microelectromechanical Systems*, vol. 21, no. 4, pp. 926-933, August, 2012. Article (CrossRef Link).

[8] L. Yuhe, C. Yanxiang, T. Xiaolei, and L. Qingxiang, "Phase unwrapping by K-means clustering in three-dimensional measurement," in *Proc. of 2013 Third Int. Conf. on Instrumentation, Measurement, Computer, Communication and Control*, pp. 65-69, September 21-23, 2013. Article (CrossRef Link).

[9] H. Xu, Z. Han, S. Feng, H. Zhou, and Y. Fang, "Foreign object debris material recognition based on convolutional neural networks," *EURASIP J. on Image and Video Processing*, vol. 21, pp. 1-10, April, 2018. Article (CrossRef Link).

[10] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Analysis and Machine Intelligenc*e, vol. PAMI-8, no. 6, pp. 679-698, November, 1986. Article (CrossRef Link).

[11] N. M. Sirakov, "Monotonic vector forces and Green's theorem for automatic area calculation," in *Proc. of 2007 IEEE Int. Conf. Image Processing*, pp. 297-300, September 16-October 19, 2007. Article (CrossRef Link).

[12] E. W. Abel, Y. Zhuo, P. D. Ross, and P. S. White, "Automatic glare removal in endoscopic imaging," *Surgical Endoscopy*, vol. 28, no. 2, pp. 584-591, February, 2014. Article (CrossRef Link).

Hyun-Sik Kim received B.S. and M.S. degrees in information and communication engineering from Inha University, South Korea, in 2002 and 2004, respectively. He received his Ph.D. degree in the electrical & electronic engineering from Yonsei University, South Korea, in 2017. He is working as a managerial researcher at the Contents Convergence Research Center, Korea Electronics Technology Institute (KETI), South Korea from 2004. His research interests are in contents sharing, wearable computing, human-computer interaction, and Machine-to-Machine/Internet of Things.

Yong-Suk Park is a managerial researcher at the Contents Convergence Research Center, Korea Electronics Technology Institute (KETI), Seoul, Korea. Before joining KETI in 2003, he was with I&C Technology and Samsung S1, where he worked in projects relevant to wireless networks and system integration. He received his B.S. and M.S. degrees in electrical and computer engineering from Carnegie Mellon University in 1997 and 1998, respectively. He received the Ph.D. degree in electrical & electronic engineering from Yonsei University, Seoul, Korea, in 2018. His current research interests are in the areas of multimedia processing, media sharing and contents delivery networks.