Depth Evaluation from Pattern Projection Optimized for Automated Electronics Assembling Robots

Jong-Rul Park[1] and Jun Dong Cho[2]

[1] College of Information and Communication Engineering, Sungkyunkwan University / Seoul, South Korea jrpark@vada.skku.ac.kr
[2] College of Information and Communication Engineering, Sungkyunkwan University / Seoul, South Korea jdcho@skku.edu

***** Corresponding Author: Jun Dong Cho

Received December 20, 2013; Revised February 27, 2013; Accepted May 12, 2014; Published August 31, 2014

** Regular Paper: This paper is invited by Joonki Paik and Seon Wook Kim, the editors-in-chief.*

** Review Paper: This paper reviews the recent progress possibly including previous works in a particular research topic, and has been accepted by the editorial board through the regular reviewing process.*

** Extended from a Conference: Preliminary results of this paper were presented at the IEEE ISCE 2014. This present paper has been accepted by the editorial board through the regular reviewing process that confirms the original contribution.*

***Abstract*:** This paper presents the depth evaluation for object detection by automated assembling robots. Pattern distortion analysis from a structured light system identifies an object with the greatest depth from its background. An automated assembling robot should prior select and pick an object with the greatest depth to reduce the physical harm during the picking action of the robot arm. Object detection is then combined with a depth evaluation to provide contour, showing the edges of an object with the greatest depth. The contour provides shape information to an automated assembling robot, which equips the laser based proxy sensor, for picking up and placing an object in the intended place. The depth evaluation process using structured light for an automated electronics assembling robot is accelerated for an image frame to be used for computation using the simplest experimental set, which consists of a single camera and projector. The experiments for the depth evaluation process required 31 *ms* to 32 *ms*, which were optimized for the robot vision system that equips a 30-frames-per-second camera.

***Keywords*:** Connected component, Image labeling, Robot vision, Image binarization, Pixel searching, Structured light system, Relative depth evaluation, Automated robot

1. Introduction

The image categorization algorithm in robotics is based on the visual information acquired from the system that operates robots [1]. The robot vision system suites for target tracking in automated surveillance systems and traffic-jam detection for traffic controls [2]. The robot should identify objects detected from its vision system. Connected component labeling is one of fundamental issues in image processing, including computer vision, and machine intelligence [3, 4].

The robot vision system provides visionary information for the control part of the vision system via object recognition process. The object recognition process acts similar to the process of the vision system of human beings because the recognition process distinguishes objects from the background of the objects. A robot vision system should recognize objects within several milliseconds for the proper operation of robotics.

Automated electronic parts assembling robot requires a faster object detection process. Faster object detection contributes more to the manufacturing capability of an automated robot. For a real-time object recognition process, the image pixel data to be processed within the system should be simplified to reduce the computational load.

First, to achieve faster object detection, an image binarization process is used as a pre-processing stage. The image binarization achieves a lower computational load of the vision system of an automated assembling robot.

During the image binarization process, the shadowed areas near an object tend to be detected as a part of the object. Although studies on shadow detection have been performed, there has been insufficient research on shadow removal to achieve significant development [5]. Removing the shadows carries a risk of deleting some parts of the original object that generated the shadow. Shading occurs as a result of the texture of the material of an object interfering with a light source [5]. The shadow removal method proposed by Finlayson [6] performs operations for an illumination invariant image. Because the edges in a shadowed area are not a part of the illumination invariant image, the shadows are canceled out. The illumination invariant image has the following assumptions. The camera input sensor responses are proportional to the luminance of the light source with the narrow setting of camera sensor response, which is similar to the Dirac delta function, and the source of light is based on Planckian lighting [5]. The grayscale intensity distribution depends on the light reflection ratio between an object and its background. The following formula [5] represents a light illumination model considering light source illumination, object reflectance, and vision sensor characteristics. $E(\lambda)$ represents the spectral power distribution of the illumination, which is approximated by Planck's law, $S(\lambda)$ represents the surface reflectance distribution, $Q_k(\lambda)$ represents the sensor sensitivity distribution, and the variable k represents the red, green, blue light sources [5].

$$\rho_k = \int E(\lambda)\, S(\lambda)\; Q_k(\lambda) d\lambda \qquad (1)$$

In the next stage, pattern analysis of structured light is applied to select an object with the greatest depth for multiple objects on the same background image. Structured light is projected from a light source with uniform patterns to both the target object and its background part.

Among the different depth evaluation values within the entire image frame, an object with the highest depth from the background is selected first for automated manufacturing robots. The automated manufacturing robot equips its arm to pick up an object and place the object in the target place in the right direction to assemble electronics. The object is normally a kind of small chip, so the chips are stacked together.

While an arm of the assembling robot picks up a stacked object, the robot arm damages the other stacked objects. On the other hand, when the robot arm picks up an object with the highest depth, physical contact between the robot arm and neighbored objects is reduced. The reduced physical contact minimizes the chances of physical damage during the picking action of an automated assembling robot.

The connected component labeling [7] stage then detects objects in an image frame. The image labeling process selects informative objects by discarding the background image within the same frame of image. The connected component labeling utilizes the concept of grouped pixel coordinates, which represents a unique object within an image frame. Connected component labeling checks both the grouped pixels and their surrounding pixels to determine if they are composed of connected component pixels for an object. The connected component algorithm scans each pixel, and assigns the temporary labels. The value of each temporary label is checked for the connectivity between pixels. When the value of the temporary labels among the connected pixels is different from each other, the second pixel scan replaces each temporary label [8].

Finally, the highest depth location from pattern analysis is matched with the location of the detected object. An object with the highest depth is then selected. The proposed labeling stage determines the connected component pixels of the contour of objects. The contour map of an object with the greatest depth is then given to an automated assembling robot. The remainder of this paper is organizes as follows. Section II introduces the proposed algorithm with an analysis of the experimental results. The concluding remarks are reported in the final section.

2. The Algorithms

Image binarization stage discriminates the objects from their background with a 2-bit level of grayscale intensity, which helps minimize the computational load for automated electronics robots. The image binarization includes a grayscale intensity distribution comparison between an object and shadow area to regard the shadow as a background, not a part of an object. The image labeling stage then represents each detected object, which is from a binary image, in several grayscale intensities to label each object. The relative depth information among the labeled objects results from different pattern reflection angles of the objects.

This proposed approach for a relative depth evaluation for an automated assembling robot evaluates the range of pixel coordinates with the greatest depth among stacked objects from distorted pattern due to the different pattern reflection angles.

The pixel coordinates range with the most depth information is then matched with the pixel coordinates of the labeled objects. The grayscale intensity in the highest depth region is then extracted. Only those pixels showing the same grayscale intensity in the greatest depth region remain still and the other pixels are regarded as a background part. As a result, the remaining pixels represent the contour of the labeled object.

Finally, an automated electronics assembling robot recognizes the shape and location of the object contour, and attempts to pick up the object. An arm of the robot has a laser based proxy sensor to detect how close the target objects are. When most of the object is close enough to be picked up, the robot begins its proper actions. Fig. 1 shows the algorithms used for the depth evaluation for a robot vision application. Each block diagram represents a modification of a conventional algorithm for an automated electronics assembly robot.

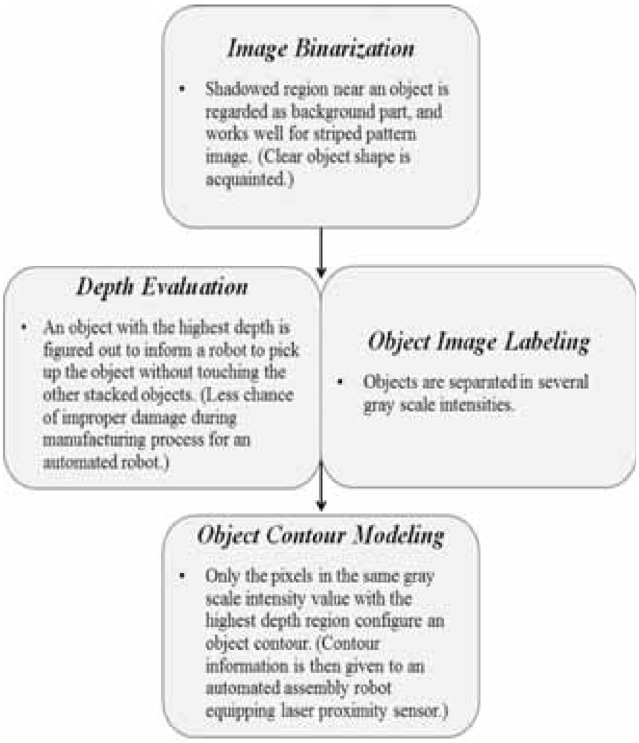


Fig. 1. Flowchart of the proposed algorithm.

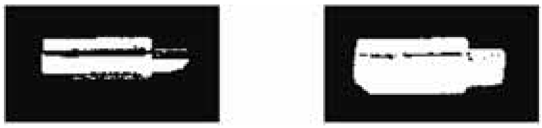
2.1 Image Binarization with Shadow Removing

Image binarization, which converts an image composed of pixels in 0 and 255 grayscale intensities, reduces the computational load for faster further image processing. Because a background image and objects show a different distribution of grayscale intensities, an algorithm is used to determine the objects from their common background image.

The binary image lowers the amount of imprecise visionary information among every object in one frame of an image, because a binary image first distinguishes the detectable objects from their background image. The binary image helps shorten the time needed for to detect objects when using a connected component algorithm. The classic connected components labeling algorithms in image processing and computer vision area rely mostly on the scanning of two connected components in a binary image [8].

In this experiment, an object in an image frame generates a shadow due to the limited source of light. The shadowed area is not filtered by the threshold value of the grayscale intensity variance between the object and its background. On the other hand, a comparison of the variance of grayscale intensity between the object part and its shadowed part cancels out the shadowed area in the proposed image binarization scheme.

The image binarization process by comparing different grayscale intensity distributions between the object part and its shadow part cancels out the shadow, as shown in Fig. 2(a). On the other hand, the image binarization process considering only the grayscale intensity distribution between an object part and its background part detects the object including its shadow, as shown in Fig. 2(b), which causes an object location error for automated assembling robots.



(a) (b)

Fig. 2. Shadowed area of an object (a) Binary image with shadow removal, (b) The shadowed region is included as an object during image binarization.

Two separated groups of grayscale distributions within the same image frame, which shows an object with its background, represent an object region and its background region. The proposed image binarization stage decides the probability of the existence [9] of an object from the grayscale histogram, which shows the distinctive distribution between an object and its background.

During the image binarization process, the grayscale intensity threshold value, which separates an object from its background and shadow, is set. The region-of-interest is then limited to the pixels of an object in the final binary image frame. In this paper, the object parts in the binary image are the only pixels to be considered for the next stage, which is the image labeling process.

2.2 Image Labeling with Noise Pixel Cancelling

To distinguish the objects from their backgrounds, this paper proposes an algorithm that determines the connectivity among objects with multiple pixel searching, memory space management and gray-scale intensity correction by detecting the connected components among pixels during the discontinuous iterative pixel searching stage.

The pixels in each image are represented by m × n format, where m and n are integers. To obtain information on the image used in the experiment, there should be a pixel searching stage. The pixel search aims to discriminate the meaningful image information from meaningless image information, i.e., background image. The background image for a robot vision system includes images of an empty conveyor-belt system, floor surfaces with no obstacles, or empty walls. For outdoor use of the robot vision system, a uniform sky is the background image for the proposed algorithm, and clouds provide the meaningful image information to recognize as an object.

2.2.1 Pixel Searching Stage

For a 640 × 480 binary image, which is used in this experiment, the pixel search is started from the left most and the uppermost pixel. An algorithm for pixel searching varies the pixel searching indices, i, and j. Whole pixels with 255-level of grayscale intensity, which are within the 640 × 480 binary image, are the pixels to be detected for the image labeling process. Only the objects in an image frame, which are shown in white color, are considered during the pixel searching stage.

The proposed algorithm uses the pixel searching indices, i and j. The variation of index j first searches each row, and index i searches each column. In the proposed algorithm, index j has its maximum value as much as the width information of the loaded binary image, and index i has its maximum value as much as the height information of the loaded binary image. The width of the loaded image is related to a horizontal element in the searching pixels within the loaded binary image.

The memory space for 640 columns in the image frame used in this experiment is initialized as 641 arrays before the pixel searching part of the image labeling algorithm. The last array, which is the 641st array, is for preparing to be initialized as the first array in the next row. The searching index, i, is increased as the searching range is shifted to the next row.

The memory space for 480 rows is also guaranteed by the form of an array in the proposed algorithm. The array for the entire rows in the loaded binary image constitutes a two-dimensional array with the array for columns of the loaded binary image. The two-dimensional array in a 640 × 480 binary image, which has two memory spaces for 640 columns and 480 rows, generates memory overflow phenomenon in the proposed algorithm. As a solution of overflow in the memory spaces, the number of the memory arrays only should be limited to several rows, and the memory arrays are reset when all allocated memory spaces are full. The overflow prevention part of the proposed algorithm utilizes an integer variable with its range of zero to three. The 480 memory spaces in the array formats for 480 rows become equivalent to the memory space with 4 memory spaces.

The integer variable used for overflow prevention is initially set to zero. After completing the search of the first row, the integer value is increased by one. The integer value becomes 1 for the first row of the loaded binary image, 2 for the second row, and 3 for the third row. After the third iteration for searching the three rows, the pixel searching process for the fourth row should change the integer value to 1 because the remainder of 4 over 3 is 1. The integer value for the fourth iteration for searching the pixels in the fourth row then becomes 1.

Because a 640 × 480 binary image is used, this paper proposes searching the pixels of the first row with 640 iterations. The pixels in the first row with 640 columns are searched to count the pixels with a 255-level of grayscale intensity, which represent object part pixels in the binary image. The search index, j, is increased each time when the current pixel is changed to the next pixel within the same row.

2.2.2 Image Labeling Stage

The detection of two or more neighboring 255-level of grayscale intensity pixels within the same 3 × 3 pixel searching window means the existence of an object. The pixel searching stage checks the connectivity within the window and finds the connected component pixels to detect the objects. After the pixel searching stage for object detection, the proposed algorithm is prepared for the image labeling stage. The labeling scheme distinguishes each detected object with different grayscale intensities. Different labeling index numbers are used to distinguish the separated objects in different grayscale intensities, as shown in Fig. 3(b).

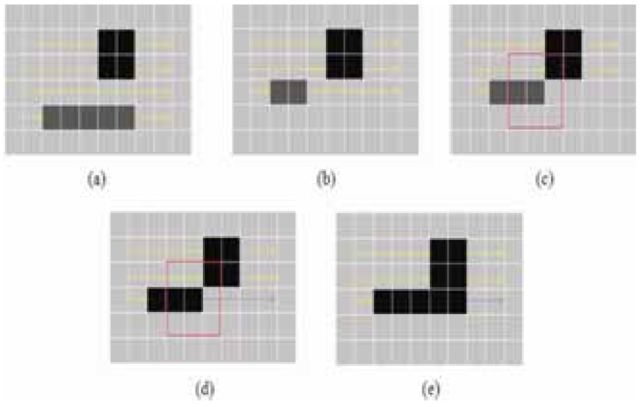


Fig. 3. Grayscale intensity correction of the connected components (a) One object is labeled in black, and the other object is labeled in gray, (b) Another image labeling case for two separate objects, (c) Two objects are labeled separately until the yellow arrow reaches the red square, (d) Because the connected components are detected in the red square, all the objects receive the same label. The blue arrow shows that the connected component finding algorithm with re-labeling continues for the next pixels, (e) Another example of an object shows similar procedure as an object in (d).

The iteration of the sliding process of pixel searching window ensures that a pixel has its neighboring 255-level of grayscale intensity pixel, and the neighboring pixel has at least another neighboring pixel with the same grayscale intensity. Finding the neighboring 255-level of grayscale intensity pixels determines the existence of an object. When the objects are detected from a binary image, the proposed algorithm is prepared for a grayscale intensity correction to represent the connected pixel components.

Fig. 3(a) shows the pixel search by indexing each pixel. Because two objects were detected, they are in different pixel colors. In Fig. 3(b), two objects are closer to each other, and the closest case is shown in Fig. 3(c). Until the second row of an image frame, label indexing remains same with the former cases, showing no grayscale intensity correction.

The center pixel of the labeling window shown in the red box in Fig. 3(c) detects a pre-labeled object. When a 3 × 3 pixel searching window is at the pixel connecting separate objects, grayscale intensity representing each object should be corrected to represent that both objects are parts of the same object. Therefore, the current component is re-labeled as same as the pre-labeled object, as shown in Fig. 3(d). In Fig. 3(e), the labeling indices became the same because they are connected components. The same labeling indices represent the same level of grayscale intensity.

A 3 × 3 pixel searching window, which is a red box

Table 1. Comparison of The Proposed Labeling Schemes.

Labeling Scheme	Operation Time [*ms*]
Lumina [10]	174.2
Shirai [11]	38.8
Hybrid [12]	15.1
The proposed labeling scheme	7
Contour tracing labeling [13]	2.2

Fig. 4. Labeled image of Lena from our connected-component labeling scheme

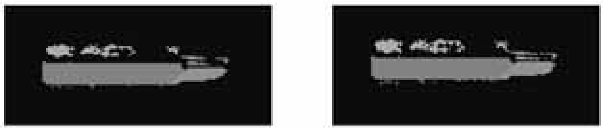
shown in Figs. 3(c) and (d), is applied to all pixels that are represented as 255-level of grayscale intensity pixels from the initial binary image. Sliding the 3 × 3 window among the pixels in the loaded binary image frame is equivalent to a 255-level of grayscale intensity pixel detection process in the image frame.

Table 1 lists the labeling process time comparison for the Lena image. The proposed algorithm required 7*ms*, and the result is shown in Fig. 4. The image does not fit the exact model for robot vision applications due to the various shapes of thin or thick objects in the Lena image.

Lumina's two-scan algorithm [10] employs a table for storing labeling data for each row of an image. Shirai [11] proposed a two-scan scheme depending on the labeled area. The hybrid algorithm [12] uses a one-dimensional table with a four times raster-scan for each pixel. The algorithm in [13] applies the contour tracing scheme of the contour-point pixels. Because contour tracing labeling accesses each pixel in an irregular manner, the algorithm does not suit pipelining and parallel processing [7] for hardware verification. The proposed algorithm achieves fast hardware implementation through a sequential search of the neighboring pixels in a two-scan scheme.

2.2.3 Noise Cancellation Stage

The reflection or shadow within each object in an image frame is likely to generate visible artifacts in the binary images [14]. In Fig. 5(b), the noise elements mostly on the lower side of an experimental object are labeled as objects. On the other hand, Fig. 5(a) shows that noise pixel elements were removed. The object used in Fig. 5 is a metal object so that uniform color detection from a camera is almost impossible due to the range of light reflections against the source of light. For automated electronics assembling robots, most working parts to be assembled are chips, which are non-reflective except for small



(a) (b)

Fig. 5. Grayscale labeling with noise pixel cancelling (a) Noise is cancelled during image labeling, (b): A labeled object with noise pixels.

conducting area that is surrounded by a non-reflective part.

The labeling index values, i and j, used for searching during the connected component labeling process, counts both the connected pixels in the objects parts and a single pixel noise. Whenever a binary image shows a single in a 255-level grayscale intensity pixel noise, the proposed algorithm detects it as noise to reconsider it as a background from sliding a 3x3 pixel searching window, which is a red box shown in Figs. 3(c) and (d). For a noise compartment, the noise pixel in 255-levels of grayscale intensity is surrounded by 0-levels of grayscale intensity pixel within the same 3x3 pixel search window. For this case, the algorithm changes the grayscale intensity of an artifact as 0-intensity to regard as a background.

2.3 Relative Depth Evaluation from Pattern Projection

Structured light system projects a user-defined pattern with a paired image sensor. Projecting a pattern is achieved using an optical beam projector, and pattern sensing is achieved by capturing an image frame from a camera. The camera captures objects and background that are reflecting the projected pattern. The pattern reflected from an object is easily distorted due to the different reflecting angle from the background. The distorted pattern only occurs on the surfaces and edges of an object. The large surface area of an object, which has the same light reflecting angle as the background surface, reflects the same or similar pattern as the pattern reflected in the background. On the other hand, the surrounding edges of the surface distort the projected pattern, and help detect the change in depth of an image frame. The depth evaluation process has cues, such as pattern distortion from the object surfaces and edges.

Each pixel in the projector is manipulated by controlling the optical reflective mirror, which is arranged for each pixel. This light reflecting scheme is called Digital Light Processing (DLP). Turning on the signal simply reflects light to the light pixel, and turning off signal cuts off light for the dark pixel. Each mirror is turned on and off ten thousand times in one second [15]. The ratio between the turning on and off a mirror for a pixel determines the grayscale intensity of the pixel. To project the color patterns, light passes the color filter.

The proposed structured light pattern is based on a binary bit; a black color pattern for signal '0' and a white color pattern for signal '1'. The signals for a binary bit pattern are well suited for hardware implementation due to the assigning of just two more variables. Moreover, both the software algorithm and hardware implementation of the proposed structured light pattern with connected component labeling process guarantees high-speed operation considering the automated electronics assembling robots. The automated robots are deployed for quicker operation and a more intensive laboring time than humans can do.

Considering the vision system for automated electronics assembling robot applications, the system should offer information on which objects to be picked up first for the robot. Assigning an object picking order is to reduce the interference among stacked composites of electronics. The electronics composite, positioned on top of the other composites, has the greatest depth causing the highest pattern distortion. This paper aims to determine which objects have the greatest depth to order an automated electronics assembling robot to determine which electronics parts to pick up first by its arm.

2.3.1 Experimental Set Preparation

To determine which electronics composites are located on top of the stacked objects, this paper proposes the utilization of relative depth information among the objects from the projecting pattern beam to those stacked objects. Fig. 6 shows the pattern image used to project the pattern beam. The pattern is iterating black and white stripes.

The projector with DLP technology reflects the source light pixel-by-pixel from its micro-mirror device. In the image of projected beam, shown in Fig. 7, the white stripe patterns shows edge shapes of the micro-mirror device.

Each square is equivalent to one pixel, and the white stripe patterns are equivalent to 255-level of grayscale intensity. The white stripes in Fig. 7 are the result of source light reflection with a color frequency shift by the color filter. Because the light source is close to the white color-temperature, the white stripes show less color frequency shift than the black stripes. Therefore, the white stripes are brighter than the black stripes.

A camera captures an image frame containing an object and its background surface with a projected pattern image. Because the locations of a projector, object, and background screen are fixed, the location of a camera influences the depth evaluation process. The depth information is processed by the amount of distortion of the reflected pattern from each surface of an object. Varying the location of a camera reduces the efficiency of the structured light system. A camera captures a reference image frame, which is Fig. 7, for the current experiment.

Electronic chips are set together as shown in Fig. 8(a), which considers the electronics composites condition of automated electronics assembling robots. The background image of Fig. 8(a) is just a screen from the experimental set, and this screen will reflect a projected pattern from a beam projector. As the location of a beam projector and a camera should differ to maximize the disparity between them, the camera was located on the left side of the projector so that the objects in Fig. 8(a) would appear like they are located on the right side of the image frame.

2.3.2 Image Labeling

The objects are labeled in grayscale intensities, as shown in Fig. 8(b). Because the surrounding light during an image frame capturing process effects the grayscale intensity distribution, the dark part of the largest object in Fig. 8(a) is shown in several parts as Fig. 8(b). On the other hand, the grayscale labeling is used for the pre-

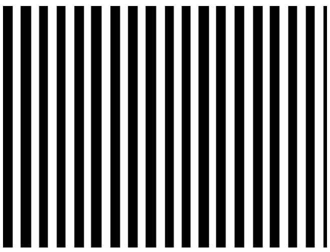


Fig. 6. Original pattern image for pattern projection.

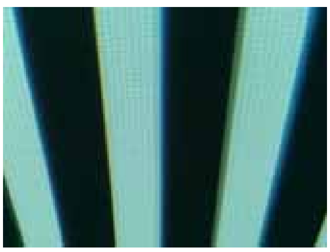


Fig. 7. Pattern projection image frame captured by a camera.

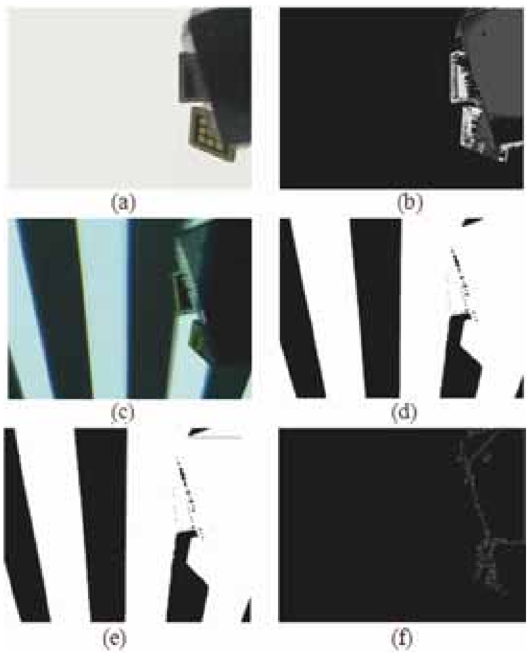


Fig. 8. Depth evaluation procedure using structured light for automated assembling robots (a) Experimental objects are stacked together, (b): Image labeling result of (a), (c) Pattern projection, (d) Binarized image of (c), (e) The greatest depth region is shown in the gray rectangular box, (f): Contour modeling for the object with the greatest depth.

detection of a structured light system to determine the location with the highest disparity between the beam projector and a camera in the structured light system.

2.3.3 Pattern Projection

Pattern projection to a background screen is captured as another image frame from a camera used for the structured light system, which is shown in Fig. 8(c). Possible pattern distortion occurs when black stripes are overlapped at the position of the white stripes. The objects cause pattern distortion because of the light reflection angle between each surface of an object and uniform standstill background screen.

The gray scale distribution among objects, black stripe pattern, and white stripe pattern differ from each other. Such distinctive grayscale distribution is used for the image binarization stage. Image binarization is once again used after the pattern projection to reduce the computational load. Fig. 8(d) presents the binary image of Fig. 8(c).

2.4 Relative Depth Evaluation Stage

Fig. 8(c) shows that the reflective pattern shown in the background part differs from the reflective pattern shown in the objects part. A possible pattern distortion occurs when black stripes are overlapped at the position of the white stripes. The objects cause pattern distortion because of the light reflection angle between each surface of an object and uniform standstill background screen. The maximum pattern distortion occurs when the angle between the pattern projection and pattern reflection from an object surface becomes 45°. An image frame, which shows an object that reflects a distorted pattern, as shown in Fig. 9(b), is captured by a camera to determine the location with high disparity.

Fig. 9 shows a binarized image set of the reference pattern and distorted pattern. Because image binarization re-assigns a 2-bit grayscale intensity for both the reference and pattern distortion image, the computational load for image processing becomes lower and achieves faster operation. The proposed relative depth evaluation stage utilizes the input images, as shown in Figs. 9(a) and (b), to determine where the highest depth region is.

The black stripe pattern regions in the input image set, as shown in Fig. 9, are the main regions of interest in the proposed depth evaluation stage. The other regions with white stripes in Fig. 9 do not need to be considered for the depth evaluation because the white stripe pattern acts as a background image.

Considering only the regions of interest during the depth evaluation lowers the computational load. For the reference and pattern distortion image frames, the location of the black and white stripes remains still with the exception of the pattern distortion and objects region. The pixels showing pattern distortion are used for the depth evaluation stage. The resulting image is presented in Fig. 8(e), which shows a rectangular box that represents the highest depth region. The required computational time was 31ms, as shown at the 5th row in Fig. 10.

Each row of the image set of Fig. 9 was compared line-

(a) (b)

Fig. 9. Binarized image set of reference pattern reflection and pattern distortion (a) A binary image result for reference pattern projection, (b) A binary image for pattern projection onto our experimental objects.

by-line. The 0-level of the grayscale intensity regions in Fig. 9(a) were compared directly with those of Fig. 9(b). Because the image set is a binary image, the grayscale intensity of the distorted black stripe patterns in Fig. 8(c) becomes a 255-level grayscale intensity when binarized. For each row of images in Fig. 9, the number of pixels showing opposite grayscale intensity is counted and saved in an array in this algorithm. The row showing the highest number of pixels showing the opposite grayscale intensity is noted as a row for an object with the greatest depth. The row showing the greatest depth is then saved as an integer-type variable.

Every column of Fig. 9(a) is compared with the same location of the column of Fig. 9(b). Only the black stripe pattern in Figs. 9(a) and (b) is similar to the row-range depth evaluation. The black pattern of Fig. 9(a) was compared directly with the same pixel location in Fig. 9(b). The number of pixels showing the opposite grayscale intensity between Figs. 9(a) and (b) are counted column-by-column. The maximum number of pixels showing the opposite grayscale intensity for each column is saved in the form of a variable.

An evaluation of the range of pixels representing an object with the greatest depth involves searching each row and column one more time after counting the pixels showing pattern distortion. Because the number of pixels showing pattern distortion is evaluated for each row and column, the second iteration process for searching each row and column aims to define the precise location of pixels with the highest pattern distortion. The process compares each row and column to determine the maximum value of the pixels representing pattern distortion in the binarized image of the pattern reflection by objects.

An object causing the maximum pattern distortion becomes the target range of the highest depth region. While searching for the maximum row and column showing the greatest depth, those rows and columns become flag indices, showing that they would be included in the range of the highest depth information. The flagged rows were compared with the neighboring rows, and the flagged columns were compared with neighboring columns. The discovered range of high depth information was then marked in rectangular form, as shown in Fig. 8(e).

Fig. 10 shows the simulation result. The greatest depth region is determined in the range of pixel coordinates for

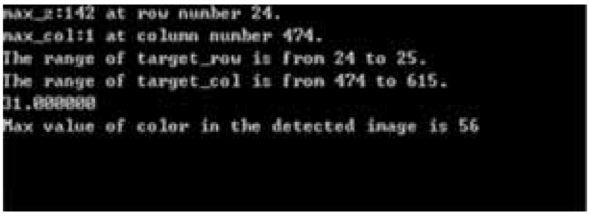


Fig. 10. Simulation result of the relative depth evaluation stage.

both the row and column. The flag index number for the row was 42, which means that the maximum pattern distortion region that showed 42 pixels in the 24th row changed to the opposite grayscale intensity due to pattern distortion. The rectangular area in Fig. 8(e) was initially a black stripe region, but was substituted with a white stripe region, where pattern distortion occurred.

2.5 Contour Generation of an Object with The Greatest Depth

The labeled image frame for the objects, which is shown in Fig. 8(b), becomes an input for a grayscale intensity evaluation of the range of pixels with the maximum pattern distortion. The maximum pattern distortion shows the area of the greatest depth among stacked objects.

The grayscale intensity values within the maximum pattern distortion area are then saved in array format in the proposed algorithm, and compared with each other to determine their brightest grayscale color. The brightest grayscale intensity within the maximum pattern distortion area was 56, as shown at the last row in Fig. 8. The pixels in the labeled image frame, which is shown in Fig. 8(b), are searched to select the pixels in same grayscale intensity with the brightest grayscale intensity, which is 56 in these experimental results, and are located in the maximum pattern distortion area.

The pixels in the same grayscale intensity with the labeled object causing the maximum pattern distortion are represented in a new output image frame, as shown in Fig. 8(f). The contour shape in Fig. 8(f) was defined as (125, 125, 125) considering the black-color background.

Because every object is captured from a camera, a grayscale intensity difference naturally occurs among pixels within the same object. Pixels showing the same grayscale intensity that which is within the object contour boundary generates a contour line. For example, direct light reflection of an object appears like a contour line representing light reflection.

As an automated electronics assembling robot equips a laser-proximity sensor near the arm of the robot, the robot receives locational information in the format of a contour line. The object locational information is defined as the pixel coordinates. The robot then receives object contour shape information to grab the target object.

An object depth evaluation requires pattern projection, and an analysis distorted pattern reflection. The pattern used for pattern projection is equivalent to structured light. Various structured light applications were compared in

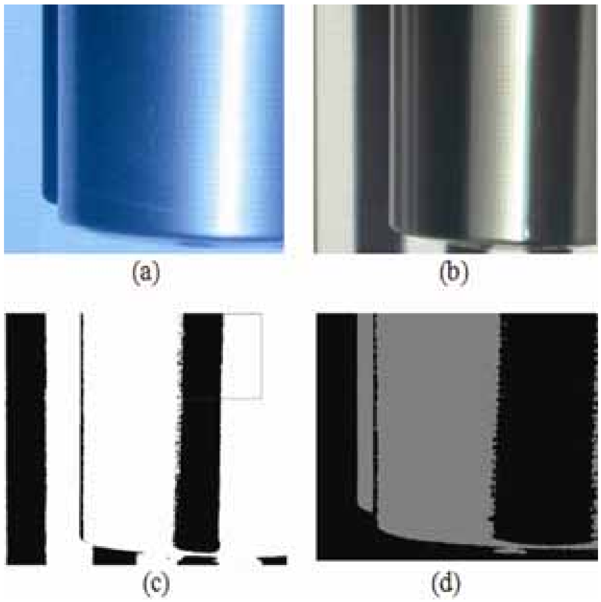


Fig. 11. Edge contour reconstruction of a cylinder-shaped object (a) Experimental object with a reflective cylindrical shape, (b) Pattern projection, (c) Depth evaluation result in the binary image frame of (b), (d) Contour modeling for shape reconstruction.

Table 2. Structured Light Applications.

Application	The Number of Image Frame Sets	The Number of Cameras	The Number of Projectors
Robot vision (proposed)	1	1	1
Shape reconstruction [17]	Multiple	1	1
Stereo vision [16]	1 or more	2	1
Gesture recognition [16]	1 or more	1	2

[16]. An additional camera is needed to match the correspondence in stereo vision, and an additional projector is needed for gesture recognition. Stereo vision has high computation, and gesture recognition requires more patterns [16].

Shape reconstruction [17] requires several image frames. The three image frames were used for cylindrical shape reconstruction in reference [17]. The algorithm requires just one image frame, as shown in Fig. 11(b), for shape reconstruction. The reconstructed shape is shown in Fig. 11(d), which shows a seamless contour object shape. The depth evaluation result is the rectangular form in Fig. 11(c) from pattern distortion analysis, which requires 32*ms*.

Table 2 lists the proposed structured light system for robot vision application aiming fast processing of depth evaluation due to the minimum number of equipment requirement. The experimental equipment for robot vision application was composed of one set of an image frame, camera and a projector.

3. Conclusion

Pattern projection from a projector is fixed perpendicular to the background screen, and different pattern reflections from various surfaces of objects generate pattern distortion. An analysis of a pattern distortion image frame, which is captured by the camera, results in depth evaluation. An automated assembling robot should be informed of an object with the greatest depth to pick up the object first so that the arm of the robot has less chance of harming objects, which are stacked at a lower depth.

The depth evaluation process with the experimental objects required 31 ms for Fig. 8(e), and 32 ms for Fig. 11(c). Because a 30-frame-per-second camera captures an image frame in every 33.33 ms, the proposed depth evaluation algorithm was optimized for an automated assembling robot using a camera. Until the camera captures the next image frame, the depth evaluation stage is performed for the current image frame and becomes ready for the next image frame.

References

[1] J. Wu, H. Christensen, and J. Rehg, "Visual Place categorization: problem, dataset, and algorithm," IEEE/RSJ Internationl conference on Intelligent Robots and Systems, St. Louis, USA, October, 2009. Article (IEEE Xplore Link)

[2] L. He, Y. Chao, and K. Suzuki, "A new two-scan algorithm for labeling connected components in binary images," World Congress on Engineering, London, U.K., 2012. Article (International Association of Engineers Link)

[3] T. Asano, and H. Tanaka, "In-place algorithm for connected components labeling," School of Information science, JAIST, Japan, Journal of Pattern Recognition Research 1(2010) 10-22, 2010. Article (Journal of Pattern Recognition Research Link)

[4] R. Gonzalez, and R. Woods, "Digital image processing," Addison Wesley, 1992. Article (ACM Digital Library Link)

[5] S. Audet, "Removing shadows from images," Electrical and Computer Engineering Department, Quebec, Canada, Image Processing and Communication ECSE-529A, Project Report, 2005. Article (Tokyo Institute of Technology Link)

[6] G. Finlayson, S. Hordley, M. Drew, and E. Tj, "Removing shadows from images," European Conference on Computer Vision – Part IV, London, U.K., 2002. Article (Springer Link)

[7] L. He, Y. Chao, K. Suzuki, and K. Wu, "Fast connected-component labeling," Pattern Recognition, vol. 42, 2009. Article (Science Direct Link)

[8] A. Rakhmadi, N. Othman, A. Babe, M. Rahim, and I. Amin, "Connected component labeling using components neighbors-scan labeling approach," University Technology Malaysia, Johor Bahru, Malysia, Journal of Computer Science 6 (10): 1099-1107, 2010. Article (CrossRef Link)

[9] N. Otsu, "A threshold selection method from gray-level histograms," IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-9, No. 1, 1979. Article (University of Aizu Link)

[10] R. Lumina, "A new three-dimensional connected components algorithm," Computer Vision Graphics Image Process 23(2), 207-217, 1983. Article (CrossRef Link)

[11] T. Gotoh, Y. Ohta, M. Yoshida, and Y. Shirai, "Component labeling algorithm for video rate processing processing," Proceeding of the SPIE, Advances in Image Processing, vol. 804, pp.217–224, 1987. Article (CrossRef Link)

[12] K. Suzuki, I. Horiba, and N. Sugie, "Linear-time connected-component labeling based on sequential local operations," Computer Vision Image Understanding, vol. 89, 2003. Article (ScienceDirect Link)

[13] F. Chang, C. Chen, and C. Lu, "A linear-time component-labeling algorithm using contour tracing technique," Computer Vision Image Understanding, vol. 93, 2004. Article (ScienceDirect Link)

[14] M. Wu, E. Tang, and B. Lin, "Data hiding in digital binary image," Electrical Engineering Dept., Princeton Univ., NJ, USA et al., 2000. Article (IEEE Xplore Link)

[15] Texas Instruments, "How DLP technology works," web, accessed on 1st, April, 2014. Article (Texas Instruments Link)

[16] C. Kim, "Structured light based depth edge detection for object shape recovery," Department of Computer Engineering, Sungkyunkwan University, Seoul, South Korea, 2004. Article (Sungkyunkwan University Library Link)

[17] H. Kawasaki, R. Furukawa, R. Sagawa, and Y. Yagi "Dynamic scene shape reconstruction using a single structured light pattern," IEEE conference on Computer Vision and Pattern Recognition, 2008. Article (IEEE Xplore Link)

Jong-Rul Park received his B.S. degree in Electronics and Electrical Engineering from Sungkyunkwan University, Seoul, Korea, in 2013. In 2009, he was an exchange student at the University of California, Berkeley. Currently, he is a graduate student of College of Information and Communication Engineering at Sungkyunkwan University, Seoul, Korea. He was involved in projects including robot vision and 3D vision system as an assistant researcher. His research interests include robot systems, mobile ubiquitous healthcare computing, and human-computer interface.

Jun Dong Cho received a BS degree in electronic engineering, Sungkyunkwan University in Seoul, Korea, 1980, an MS degree from Polytechnic University, Brooklyn, NY, 1989, and a PhD degree from Northwestern University, Evanston, IL, 1993, both in computer science. He was a senior CAD engineer at Samsung Electronics, Co., Ltd. He is now professor of Department of Electronic Engineering, Sungkyunkwan University, Korea. He is also a founder and chairman of a new graduate program called "Human ICT Convergence" since 2013. He was a visiting scientist of IBM T.J. Watson Research Center, from 2000 to 2001. He has been an IEEE senior member since April 1996. His research interests are in the area of VLSI/SoC of multimedia and communication and smart-care wearable devices.