World Academy of Science, Engineering and Technology
International Journal of Electronics and Communication Engineering
Vol:5, No:11, 2011

Fast 2.5D model reconstruction of assembled parts with high occlusion for completeness inspection

Matteo Munaro, Stefano Michieletto, Edmond So, Daniele Alberton, and Emanuele Menegatti

Abstract—In this work a dual laser triangulation system is presented for fast building of 2.5D textured models of objects within a production line. This scanner is designed to produce data suitable for 3D completeness inspection algorithms. For this purpose two laser projectors have been used in order to considerably reduce the problem of occlusions in the camera movement direction. Results of reconstruction of electronic boards are presented, together with a comparison with a commercial system.

Keywords—3D quality inspection, 2.5D reconstruction, laser triangulation, occlusions.

I. Introduction and related work

ONE emerging application of machine vision is quality control for checking the completeness (presence/absence of parts, correct type, position, orientation) of assemblies. Existing systems usually exploit 2D cameras that provide a monochrome or color image. These images lack depth information, and consequently have problems when dealing with non-rigid objects (hoses, cables) or low contrast between the background and the inspected parts. They also often do not provide an optimal view on each single part of the assembly. For these reasons the industrial market is increasingly turning to inspection techniques which exploit a 3D representation of the object to be scanned.

Coordinate Measuring Machines (CMMs) [1] are contact based techniques, in which the 3D model of the part is obtained with a physical probe scanning the part by touching it. They are widely used for fulfilling inspection tasks, but it takes a long time to do the inspection and these machines are usually located out of the production line. Moreover, contact inspection may damage the part's surface.

In this work, we propose a non-contact acquisition system that can be located inside a production line and meets typical industrial needs such as high-speed, real-time data analysis, high precision, reliability and ability to measure complex surfaces.

Among non-contact techniques, passive vision systems are not well suited for being used in an industrial framework. *Shape-from-shading* [2] and *shape-from-motion* [3] have limited ability to cope with non-uniform surface textures, while *passive stereo vision* [4] relies on feature extraction and matching which are complex and computationally intensive operations [5]. Our system is based on the active triangulation paradigm [6], that is relatively simple and allows high data acquisition rates and uniform sampling of object points. The main weakness of this technology is the loss of data arising from occlusions [7], either with respect to the light projector or with respect to the camera. This problem is even more significant when dealing with electronic assemblies where some taller components occlude part of the board. Some researchers ([8], [9]) try to reduce occlusions with respect to the camera by adding a second camera, but this method generates twice as many images to be processed. Instead, we use one camera with two laser line projectors located on both sides of the camera. This is a less expensive solution and requires only one stream of images to be processed.

In [10] two parallel laser projectors are used for error compensation in seam tracking for laser welding, while in [11] two lasers are placed at the opposite sides of a camera for acquiring richer range models of objects that do not deal with high self-occlusions. Moreover, Park *et al.* [12] addressed the problem of occlusion for arbitrary 3D modeling. Our acquisition system exploits a similar configuration for providing both range and texture of assembled objects that can be used by 3D inspection algorithms. All this information can be acquired in one scan with the scanner directly inserted in the production line.

The remainder of the paper is organized as follows: in Section II, our dual laser triangulation prototype is described, together with the algorithms for efficient composition of 2.5D textured models. Section III reports some composition results on real assemblies, a comparison between our single and dual triangulation systems and a comparison with a commercial product for this kind of tasks. Some results of inspection algorithms that can be performed on this type of data are also described. Conclusions are reported in Section IV.

II. Dual laser triangulation system

A dual laser triangulation system (DLTS) has been developed for acquiring both range and texture information of an object. It has been designed to be flexible and easy to install within a production line for scanning objects while they move along a conveyor belt. For building more accurate models and reducing the problem of laser occlusion in the movement direction we decided to use two laser projectors, each of them projecting a laser line on the scene.

A. Configuration

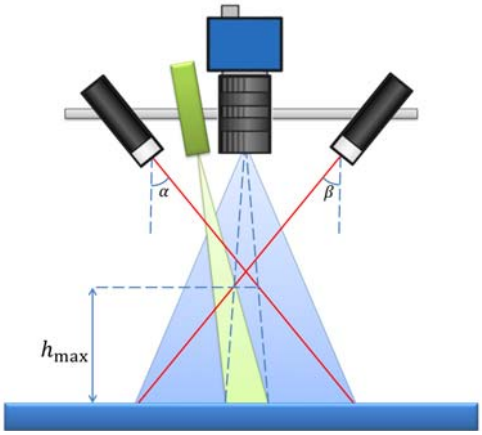
The chosen configuration for our scanner features a downward-looking camera and the two lasers at its sides,

The research leading to these results has received funding from IT+Robotics (www.it-robotics.it) and the European Union's 7th Framework program managed by REA-Research Executive Agency (http://ec.europa.eu/research/rea - FP7/2007-2013) under grant agreement No. 262009.

M. Munaro, S. Michieletto, E. So, D. Alberton, and E. Menegatti are with the Department of Information Engineering (DEI), University of Padova, Italy, 35131 - Padova (IT), e-mail: {munaro, michieletto, so, alberton, emg}@dei.unipd.it.

as illustrated in Fig. 1(a). This camera position allows the acquired texture images to be also suitable for typical 2D inspection tasks.

Range information is easier to extract from images where the laser projectors are the only source of light, while a diffused illumination of the scene is required for texture acquisition. In order to be able to extract both range and texture in a single scan, we placed a collimated white light next to the camera. This light allows us to perform acquisitions in dark lighting conditions and illuminate only a small part of the object where texture can be correctly acquired. An example of an image featuring this configuration is reported in Fig. 1(b), where the two laser lines can be seen above and under the illuminated region that reveals the texture. The laser projectors must be positioned in such a way that the two laser lines intersect above the object to be scanned, in order to increase the laser-camera angle, thus improving the triangulation accuracy and resolution. The laser points should also not overrun the zone reserved for texture. This positioning can be easily done by checking where the laser lines lie in correspondence to the tallest part of the object to be scanned.



(a) Hardware setup.

(b) Image example.

Fig. 1. a) Our prototype components and configuration. The camera is drawn in blue, the laser projectors in black and the collimated light in green. b) An example of an optimally acquired image with the proposed setup.

B. Calibration

The calibration of the acquisition system is composed of three steps:

- camera calibration;
- laser-camera calibration for both lasers;
- scanner-world calibration.

The objective of camera calibration is to determine a set of parameters that describe the mapping between 3D world coordinates and 2D image coordinates. In many cases, the overall performance of a machine vision system strongly depends on the accuracy of the camera calibration. To perform the calibration, the camera is modeled as a pin-hole camera. It is based on the principle of collinearity, where each point in the object space is projected by a straight line through the projection centre into the image plane and it can be represented by (1), onto:

- (u, v) are the coordinates of the point on the image plane,
- K is a intrinsic parameters matrix,
- R is a rotation matrix,
- t is a translation vector,
- (x, y, z) are the coordinates of a 3D point in the real world.

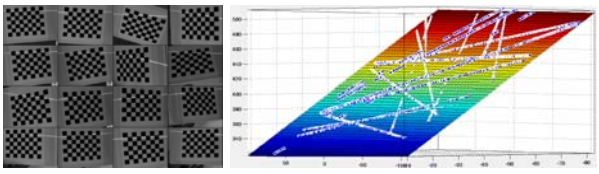
$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = K \, [R|t] \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix} \tag{1}$$

The solution to this equation system was discussed by several authors. In this work, the approaches from Zhang [13] and Heikkilä and Silven [14] are used to obtain good calibration parameters.

The laser points one can see in the image belongs both to the laser plane and to the object to be digitalized. In order to calculate the 3D world coordinates of the laser points in the image, it is necessary to identify the laser plane. It is modeled by the general equation of a plane expressed in the world coordinate system (2).

$$cAx_w + cBy_w + cCz_w + cD = 0 \tag{2}$$

We took different pictures of the laser line projected on a checkerboard (Fig. 2(a)) to estimate the laser plane equation. Our algorithm can identify the rototranslation of the checkerboard respect to the camera coordinate system by decomposing the homography between image and checkerboard plane. Then, the 3D coordinates of the laser points are the intersection between the checkerboard plane and the camera rays passing through the laser point in the image. A fitting algorithm [15] is used to find the best plane through the laser points (Fig. 2(b)).



(a) Calibration images. (b) Laser plane estimation.

Fig. 2. a) An example set of images used for laser-camera calibration. b) Laser plane estimation by plane fitting on the estimated laser points.

The last step is a scanner-world calibration process. It aims to estimate the direction of movement between the acquisition system (or of the object to inspect) and the world. For this

purpose we acquire two images of the same planar object, e.g. a checkerboard, from two different scanner positions with respect to the acquisition plane. From the photos some local features are extracted and matched to estimate the planar transformation between the two images. Actually we deal with a pure translational motion, so rotation is equal to zero. For this reason, we estimate an homology to find the corresponding 3D camera translation. This translation vector gives us the direction of movement. We decided to use an homology in order to reduce eventual estimation errors. If the scanning condition is different, it is possible to find both the rotation and translation by decomposing an estimated homography matrix.

C. Reconstruction steps

Here, the algorithms that allow us to produce range and texture data from the acquired images are described. First, the laser points are detected in the image. Then, they are triangulated in order to estimate their 3D position in the camera reference system. The camera pose is estimated for every frame, so all the 3D profiles can be referred to the same reference system to compose the whole set of points. From the resulting point cloud a mesh is created and images are projected onto it for texturing.

Alternatively, a range image can be created by mapping the triangulated points onto a 2D image, and a texture image can be built by stitching together all the images acquired according to the estimated motion of the camera.

1) Dual laser detection: The objective of laser detection is to provide for every image a list of (x, y) image points where the laser has been detected. The geometry of the system was chosen such as the laser lines are projected along the horizontal direction of the image, and the points coming from the two lasers lie in different parts of the image.

Laser detection is a relatively easy task in controlled environments and on objects of uniform materials. However, in industrial environments, it becomes more difficult because of variations in lighting conditions, noise, and objects being composed of materials with different reflectivity. The algorithm we implemented, as the one in [16], has been shown to be robust to these conditions, while also being accurate.

Differences in the object distance from the camera and the light absorption properties of the underlying components results in laser lines with different widths in the image. We chose to consider every image column as a separate signal and to extract only the peak pixel of each laser for every column.

We apply a threshold on the red channel to neglect low intensity points, which arise mainly due to noise or reflections. Then, the points with the highest intensities are selected between the remaining ones and their mean position is computed.

The detected laser line can present some local discontinuities due to missing or imperfect detections, caused by adverse conditions such as noise or laser diffusion. When dealing with electronic assemblies we can assume that most of the components have planar surfaces, so that the resulting laser line in the image should be many fragments of locally horizontal segments. Based on this observation, a line filter is applied to the detected laser points with the aim to fill some holes and preserve the continuity of the laser line.

An outlier rejection algorithm is also applied in order to remove isolated detections, mainly due to noise.

Fig. 3 gives an example of the result of the initial detection, the outlier rejection and the line filter. It can be seen how our algorithm can successfully cope with imperfect peak detections, which can be a significant problem in an industrial context.

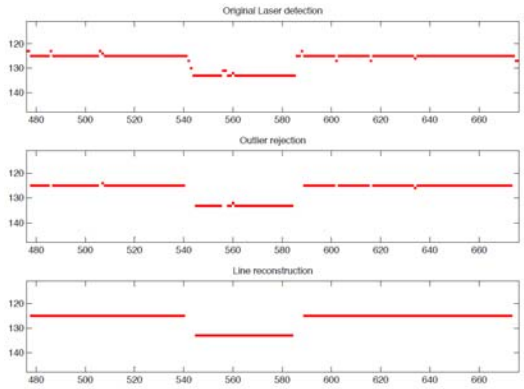


Fig. 3. Example of laser detection result after the initial detection, the outlier rejection and the line filter.

2) Triangulation and model building: Once the laser points are detected in the image, the laser-triangulation principle is applied to estimate their 3D positions in the camera frame. This is done by intersection of the camera rays passing through the laser points in the image plane and the corresponding laser plane estimated in the calibration phase. In Fig. 4, triangulation is illustrated for a couple of image points.

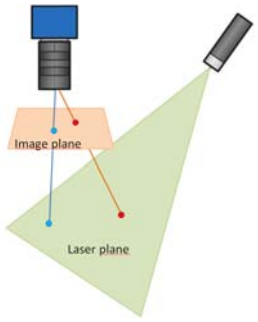


Fig. 4. Illustration of the triangulation principle for two image points.

Once the 3D laser points have been found for every image they are referred to the same reference system by a translation along the direction estimated with the scanner-world calibration. After this step a point cloud for each laser is obtained.

3) Point clouds registration: If the calibration parameters are perfectly estimated, the point clouds obtained from the two lasers will be correctly aligned and they can be merged into a single cloud. However, if any slight error is made during calibration (e.g.: bad calibration images, bad RANSAC estimation), the two point clouds will be misaligned and need a correction before being merged.

The *Iterative Closest Point* (ICP) [17] method can be applied

to estimate the relative rotation and translation between the two point clouds. This is not usually a real-time algorithm, but the rototranslation needed for the registration is fixed, so it can be found off-line and then be applied to every new model that is composed.

As an additional check, we apply a control on the Z coordinate distribution of the two point clouds. The points belonging to the higher parts of a board should be present in both point clouds because these are points that should not be occluded. Thus, if we check the histograms of the Z component, they should coincide for high values of Z. In Fig. 5, an example of the histograms when the point clouds are not well aligned in the Z direction is shown. As we can see, the histogram peaks appears to be shifted by a constant offset that can then be removed from a point cloud in order to perform the alignment. This easy check can be used as a further refinement of the point clouds registration and it can also provide information when the registration with the ICP method is no longer valid, e.g. because something changed in the hardware configuration due to external factors.

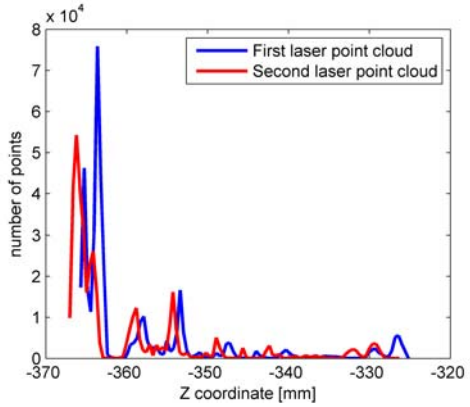
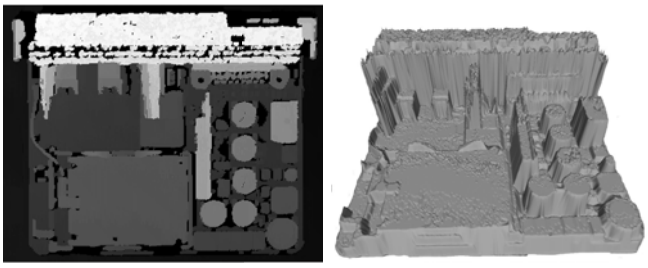


Fig. 5. Example of Z coordinate distribution for not well-aligned point clouds.

From the whole point cloud, a gray-coded height map can be immediately derived, thus obtaining a *range image* that is commonly used for 3D inspection. For better 3D visualization of the composed model or for creation of an orthoimage, a mesh can be also composed through Delaunay triangulation. The mesh, unlike the range image, contains interpolated data in the presence of occlusions. An example of both of these model representations for an automotive electronic board is shown in Fig. 6.

4) Texturing: As we have proposed two different representations of the acquired range data, we developed two corresponding methods for texture creation. For the range image, an associated texture image is created by stitching together all the acquired images. This method gives very good results when the camera/object displacement between consecutive images is small, so the perspective effect can be reduced by using the central part of every image for creating this kind of image.

Conversely the mesh can be textured by means of image projection on the 2.5D mesh and simple blending techniques. An orthoimage, that is an image without perspective effects



(a) Range image. (b) Mesh.

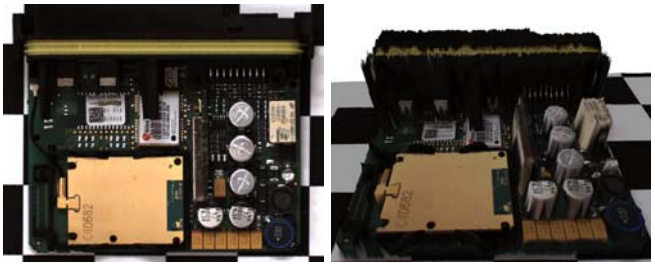
Fig. 6. a) Range image and b) mesh for an automotive electronic board.

TABLE I
ACCURACY AND RESOLUTION IN A VISUAL FIELD OF 150 mm ×340 mm.

Accuracy (mm)	0.5
Horizontal resolution (mm/px)	0.095
Vertical resolution (mm/px)	0.105

that can be used for 2D inspection, is then created by selecting a top view of the textured model. An example of a texture image and a textured mesh for the automotive board is shown in Fig. 7.

In a production line, if the time constraint for the model



(a) Texture image. (b) Textured mesh.

Fig. 7. a) Texture image and b) textured mesh for the automotive test case.

composition is very tight (e.g. of the order of a few seconds), the range and texture image can be easily created in real-time, while the mesh and the orthoimage are intended to be used if time permits and a better visualization of the model is needed.

III. EXPERIMENTS

In this section some further details about our system are provided and some of the experiments we performed are described.

Basic characteristics to know about a visual instrument are its accuracy and resolution. In order to evaluate the accuracy of the system we compared some estimated values from our DLTS to the real measured values. From these comparison we can also estimate horizontal and vertical resolution in a visual field of 150 mm ×340 mm. The results are listed in Table I.

In Table II, the computational times of the algorithms that we implemented for laser detection and triangulation on a single image are reported. Table III contains an example of time needed for computing range and texture data from 350 images acquired while moving the scanner on top of the automotive

TABLE II
MATLAB COMPUTATIONAL TIME FOR THE PROCESSING ALGORITHMS APPLIED TO THE SINGLE FRAME (IN MILLISECONDS).

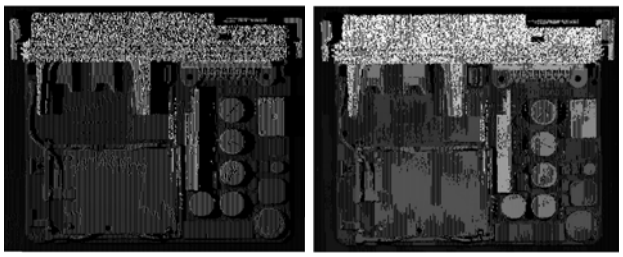
Data for a single image of 800x1600 pixels	***time (ms)***
Laser detection for the first laser	100.0
Laser detection for the second laser	100.0
Triangulation for the first laser	48.8
Triangulation for the second laser	48.8
Total	297.6

TABLE III
MATLAB COMPUTATIONAL TIME FOR THE WHOLE CREATION OF RANGE AND TEXTURE DATA FROM 350 IMAGES (IN SECONDS).

Data for a model of 350 images	***time (s)***
Detection and triangulation from every image	104.16
Range image creation	0.27
Texture image creation	0.71
Total	105.14

test case represented in Fig. 6 and 7. The whole process takes 105 seconds in our non-optimized MATLAB implementation, but we are already working on a C++ implementation within an high-efficiency framework that should reduce this time to a few seconds.

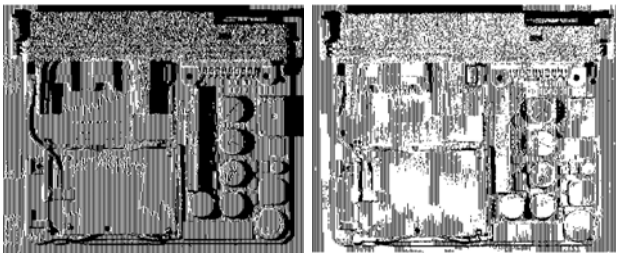
In Fig. 8 the range images produced for our automotive board are compared between using the points from one single laser and from both lasers. A visible improvement can be seen with two lasers in terms of a denser reconstruction of the board and of filling of parts that are occluded with only one laser. The amount of new points introduced with the second laser (here around 35%) can be better visualized by looking at the valid maps shown in Fig. 9. These valid maps are binary masks that are set to 1 (white) for points in the range image for which we have a range value. The parts of the board where the problem of occlusion is eliminated can be immediately detected. In Fig. 10 the effects of occlusion can be seen also for the mesh creation, where missing points are created by interpolating existing ones. Two of the artefacts that arise because of this interpolation are highlighted with red ellipses.



(a) Single laser (b) Both lasers

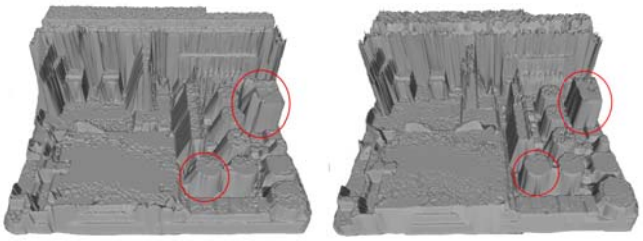
Fig. 8. a) Range image obtained with one laser and b) range image obtained with two lasers for the automotive test case.

Thus, our dual laser triangulation system allows us then to obtain richer range images than with a single laser. This information is very useful for many 3D inspection tasks, as we can see for example in Fig. 11 for edge detection. In this figure the results of the Canny edge detector is reported when



(a) Single laser (b) Both lasers

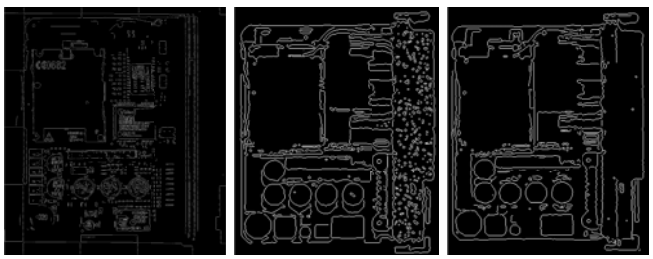
Fig. 9. Valid map for the automotive test case for a) the range image obtained with one laser, b) the range image obtained with two lasers.



(a) Single laser (b) Both lasers

Fig. 10. a) Mesh obtained with one laser and b) mesh obtained with two lasers for the automotive test case. The red ellipses highlight two artefacts that arise in the mesh composed with only one laser due to the problem of occlusions.

applied to a texture image (a), a range image obtained with one laser (b) and a range image obtained with two lasers (c). As we can seen, complementary information can be extracted from range and texture data. In particular, from the texture image it is difficult to segment components that have low contrast with the board or with the background. For such cases the range image can be used, but, if some occlusions are present, some components cannot be seen or some fake edges due to missing points are created when performing edge detection.

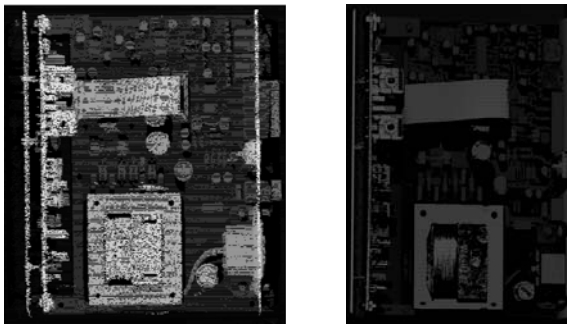


(a) (b) (c)

Fig. 11. Results of the Canny edge detector for the automotive test case on a) the texture image, b) the range image obtained with one laser, c) the range image obtained with two lasers.

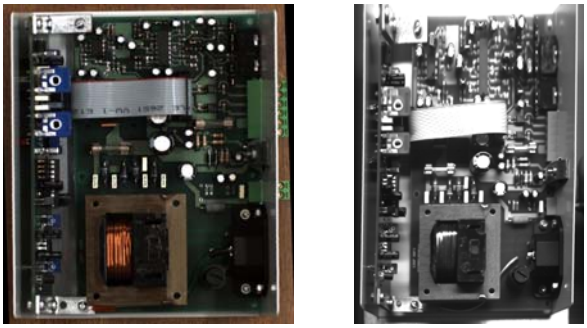
In Fig. 12 and 13, we compare our results for another board from the electronic sector with the output given by a *SICK Ranger E55*, which is a state-of-the-art product used for 2.5D reconstruction within a production line. It also exploits the laser triangulation principle, but it is provided with proprietary routines for laser detection by exploiting the embedded electronics that allows for only one laser. In fact, some occluded (black) zones can be noticed in its range image. From this

comparison, we see that our system gives results comparable in terms of resolution and accuracy with this commercial device. Our system is much slower in acquisition and reconstruction, but it is much cheaper than the Ranger. Moreover, having two lasers our system can build complete models even in presence of strong occlusions along the camera/object movement direction.



(a) Our output (b) Ranger's output

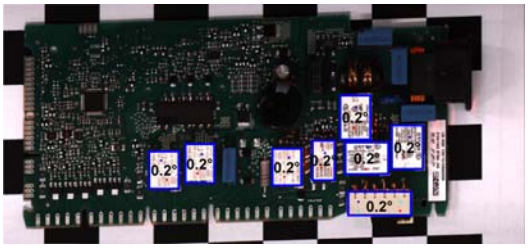
Fig. 12. Range image obtained for an electronic board with our system (a) and with the *SICK Ranger E55* (b).



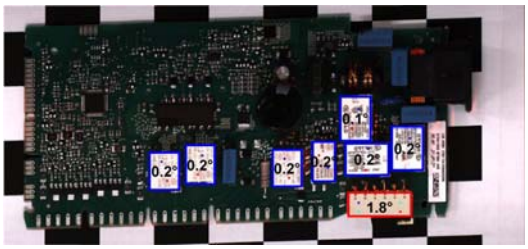
(a) Our output (b) Ranger's output

Fig. 13. Texture image obtained for an electronic board with our system (a) and with the *SICK Ranger E55* (b).

As a further proof that the range and texture data provided by our system are suitable for 3D inspection, we performed a scan of an electronic board for the white goods sector and checked the orientation of its components. For doing that, we used the texture image for locating the components and the range image for performing an orientation check to the top surface of every component. In Fig. 14, the results of this control are reported for a defected board (a) and a good one (b). For the good case all the components have a consistent orientation, while in the defected board, there is a badly assembled component that is correctly detected as faulty and surrounded by a red bounding box.

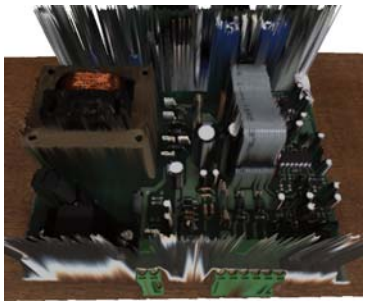


(a) Inspection results for a good board

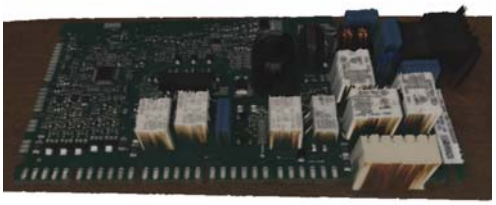


(b) Inspection results for a defected board

Fig. 14. Example of orientation inspection on a board for the white goods sector. The good components are surrounded by a blue bounding box, while the defected ones are highlighted in red.



(a)



(b)

Fig. 15. Textured meshes for two boards of the electronics (a) and white goods (b) sector.

IV. Conclusion

In this paper, we presented our Dual Laser Triangulation System which exploits a camera, two lasers and a collimated white light for acquiring range and texture data from assembled parts with high occlusions. Hardware and software have been developed for being used within a production line and providing suitable data for 3D completeness inspection. The reported results prove the effectiveness of our system in composing good models even in presence of strong occlusions in the camera/object movement direction and that our reconstructions are comparable to those obtained with state-of-the-art but more expensive solutions.

3D inspection algorithms and a C++ implementation of the described techniques within a high-efficiency framework are foreseen as future works.

REFERENCES

[1] W. E. Singhose, W. P. Seering, and N. C. Singer, *The effect of input shaping on coordinate measuring machine repeatability*. In Proceedings of the 1995 IFToMM World Congress on the Theory of Machines and Mechanisms. Pages 2930-2934.

[2] Q. Zheng, and R. Chellappa, *Estimation of illuminant direction, albedo, and shape from shading*. In Proceedings of IEEE Conference on Computer Vision and Pattern Recognition 1991 (CVPR'91), Maui, HI, USA. Pages 540-545.

[3] L. Zhang, B. Curless, A. Hertzmann, and S. M. Seitz, *Shape and motion under varying illumination: unifying structure from motion, photometric stereo, and multi-view stereo*. In Proceedings of the 9th IEEE International Conference on Computer Vision (ICCV 2003), Nice, France. Pages 618-625.

[4] D. Scharstein, R. Szeliski, and R. Zabih, *A taxonomy and evaluation of dense two-frame stereo correspondence algorithms*. In Proceedings of Stereo and Multi-Baseline Vision 2001 (SMBV 2001), Kauai, HI, USA. Pages 131-140.

[5] Z. M. Bi, and L. Wang, *Advances in 3D data acquisition and processing for industrial applications*. Journal on Robotics and Computer-Integrated Manufacturing (2010). Volume: 26, issue: 5, publisher: Elsevier. Pages: 403-413.

[6] F. Blais, *Review of 20 years of range sensor development*. In Journal of Electronic Imaging, 13(1): 231-240. January 2004.

[7] G. Sansoni, M. Trebeschi, and F. Docchio, *State-of-the-art and applications of 3D imaging sensors in industry, cultural heritage, medicine, and criminal investigation*. Sensors, Volume: 9, Issue: 1, Pages: 568-601, 2009.

[8] F. J. Brosed, J. J. Aguilar, D. Guillomia, and J. Santolaria, *3D geometrical inspection of complex geometry parts using a novel laser triangulation sensor and a robot*. Sensors, Vol. 11, No. 1. (23 December 2010), pp. 90-110.

[9] J. Vilaca, F. Jaime, and P. Antonio, *Non-contact 3D acquisition system based on stereo vision and laser triangulation*. Journal on Machine Vision and Applications (2010). Volume: 21, issue: 3, publisher: Springer, pages 341-350.

[10] S. Gao, M. Zhao, L. Zhang, and Y. Zou, *Dual-beam structured light vision system for 3D coordinates measurement*. In Proceedings of the 7th World Congress on Intelligent Control and Automation (2008), Chongqing, China.

[11] A. Peiravi, and B. Taabbodi, *A reliable 3D laser triangulation-based scanner with a new simple but accurate procedure for finding scanner parameters*. In Journal of American Science 2010;6(5), 2010.

[12] J. Park, G. N. Desouza, and A. C. Kak, *Dual-beam structured-light scanning for 3-D object modeling*. In Proceedings of the 3rd International Conference on 3-D Digital Imaging and Modeling (2001).

[13] Z. Zhang, *Flexible camera calibration by viewing a plane from unknown orientations*. In Proceedings of IEEE International Conference on Computer Vision, 1999 (ICCV'99), Kerkyra, Greece. Pages 666-673.

[14] J. Heikkilä, and O. Silven, *A four-step camera calibration procedure with implicit image correction*. In Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition 1997 (CVPR'97), San Juan, Puerto Rico. Page 1106.

[15] M. A. Fischler, and R. C. Bolles, *Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography* . In Communication of ACM, Volume: 24, Issue: 6, Pages: 381-395, 1981.

[16] J. Molleda, R. Usamentiaga, D. F. Garcia, and F. G. Bulnes, *Real-time flatness inspection of rolled products based on optical laser triangulation and three-dimensional surface reconstruction*. Journal of Electronic Imaging 19(3), 031206 (Jul-Sep 2010).

[17] Z. Zhang, *Iterative point matching for registration of free-form curves*. In International Journal of Computer Vision, Volume 13 Issue 2, Oct. 1994.