Block Matching Mosaicing for Surface Inspection Using an Autonomous Mobile Robot

Sara Roos-Hoefgeest Toribio
Systems and Automation Engineering
Universidad de Oviedo
Gijón, Spain
roossara@uniovi.es

Ignacio Alvarez Garcia
Systems and Automation Engineering
Universidad de Oviedo
Gijón, Spain
ialvarez@uniovi.es

Rafael C. Gonzalez de los Reyes
Systems and Automation Engineering
Universidad de Oviedo
Gijón, Spain
rcgonzalez@uniovi.es

***Abstract*—Visual inspection of manufactured products is a field in constant expansion. In this work we present a method to create a high resolution panorama (1mm/pixel) of a large rectangular plate using a mobile robot with two RGB-D cameras. The panorama is intended to analyze the surface in search of possible defects and identify areas of interest that have been encircled using a high contrast mark. Identifying which points belong to the surface plane and estimating the amount of distortion caused by the perspective correction we are able to form a panorama of a 4400 by 2500 mm plate with errors lower than 2% and a resolution of 1 mm/pixel.**

***Index Terms*—visual inspection, image mosaicing, autonomous robot.**

I. Introduction

Surface inspection of steel products using computer vision techniques is a well established practice nowadays. There have been a huge effort to analyze different categories of steel products as they are manufactured [1]–[4]. Some defects occurring in slabs or heavy plates may be repaired. Defects are removed manually using an angle grinder. The worker that conducts the operation usually needs to adopt awkward postures. As a consequence, his exposition occupational hazards increases.

We are designing a mobile robot to automate heavy steel plates reparation. To be successful, the robot must be able to explore the plate surface to identify which parts have been marked as defective, or even to look for defects. The thickness of heavy steel plates ranges from 5 to 150 mm. Their width varies from 1400 to 3300 mm and their length may change from 4 to 18 m. To enforce robot and human safety, the robot must complete the analysis moving exclusively within the plate. To complete the taks, the robot must move on top of the surface and take partial surface images. Before the inspection starts, these images have to be stitched together to form a single view called a panorama. This approach has been used in many different applications [5]–[7].

Image registration and stitching is usually solved through the pairing of image features. Those methods usually provide very good results and are considered to be the state of the art. Works by Szeliski [8] or Zitová and Flusser [9] are a good starting point to have an overview of the full process. A common problem when these methods are applied at a large scale is that concatenation of geometric transforms may cause an increasing amount of image distortion [10]. Block matching methods, although not so common, have been successfully applied to register a sequence of images in industrial environments [6] or to determine the motion of glaciers or landslides [11]. In this paper we describe a method to reconstruct the surface of a large rectangular plate using an omnidirectional mobile robot and RGB-D cameras. The method has been developed to be used on heavy steel plates, although real tests have been done using Oriented Strand Board (OSB) wooden panels. This is because they can be easily handled inside our lab and present a repetitive texture where key points are difficult to detect. The areas to be repaired are expected to be enclosed within a high contrast mark.

Our proposal is based on a classical block matching approach, although several new ideas are introduced to account for image distortion and dynamic selection of the reference block to be matched. The cameras are pointing 30° downward with respect to a horizontal plane. We apply an homography to remove the perspective effect. This introduces a distortion in the re-projected image. We use the transformation Jacobian to get an estimation of the distortion. This value is used to reduce the search area within the block matching algorithm and to control how images are merged to build the final panorama. In addition, we classify pixels as surface/not surface points according to their distance to the plane defined by the inspected plate. This clustering is used to further restrict the candidate positions of the template block. To build the full panorama we use a hierarchical approach similar to the one described by Xie et al. [13]. First, we register images from a single camera to form a stripe. Neighboring stripes are stitched together to build a unique panorama for each camera. We finish by joining both partial panoramas into a single one.

II. Methods

A. Algorithm description

Our algorithm, outlined in Fig. 1, follows a hierarchical structure. First we compute the translation between two consecutive frames. The result is used to register the new image within the current stripe panorama. Those stripe panoramas are

Research supported by: Daorje S.L.U inside project code CN-18-011 and a scholarship under the "Severo Ochoa" program for predoctoral research and teaching with Ref: PA-20-PF-BP19-067, financed by Asturias Regional Government .

the result of stitching the images captured by the robot while it moves forward or backwards in the longitudinal direction of the plane. Each time a new stripe is completed, the result is combined with the previous one in a similar way as the frame to frame registration. The result is a partial panorama associated with each camera. To complete the surface image, the partial panoramas obtained for each camera are stitched together in a similar way.

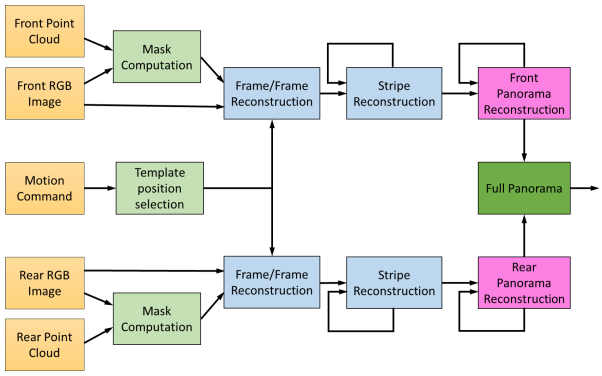

Fig. 1. Outline of the algorithm.

We make the assumption that the surface to be reconstructed lies on a plane and the robot moves on top of that plane and is not allowed to move outside of the surface. An RGB-D camera is placed at the front and another one at the back of the robot. Both cameras have their optical axis parallel to the X-Z axis of the robot, and tilted downwards to the surface. This allows to capture the whole surface with a single boustrophedon type path, as represented in Fig. 2. The overall size of the surface is initially unknown, although it is assumed to be a rectangle. The robot is able to detect the border to control its orientation and the limits of its movements. Lateral displacements are limited to a maximum that guaranties image superposition while reducing the number of longitudinal segments in the trajectory. The last one may be slightly smaller so that the robot keeps within the surface.

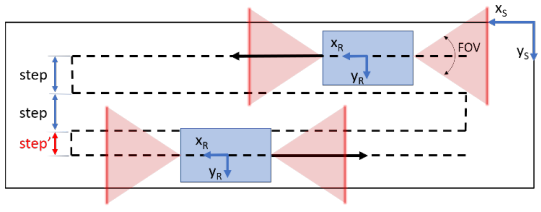

Fig. 2. Trajectory followed by the robot for image acquisition. Using an omnidirectional robot and two cameras allow to simplify the trajectory required to cover the whole surface.

To correct the perspective effect we estimate the value of an homography matrix using a standard checkered pattern and applying RANSAC and SVD decomposition. The board has 8 rows and 11 columns to provide a reasonable number of points. Each pattern square is 60 mm wide. The calibration pattern covers almost all the transformed part of the image. The relationship between an image point, $\mathbf{p}_i = [u,\ v]^T$, and its position in the real surface plane, $\mathbf{p}_s = [x,\ y]^T$, is defined by:

$$\begin{aligned} x &= \frac{h_{11}\ u+h_{12}\ v+h_{13}}{h_{31}\ u+h_{32}\ v+h_{33}} \\ y &= \frac{h_{21}\ u+h_{22}\ v+h_{23}}{h_{31}\ u+h_{32}\ v+h_{33}} \end{aligned} \tag{1}$$

Where h_{ij} are the nine elements that define the homography matrix up to a scale factor. As the cameras are fixed with respect to the robot and it moves on top of the surface to be reconstructed, these geometrical transformations remains constant during robot operation. Therefore, they can be calibrated off-line.

B. Template selection

In our proposal, image registration is based on a template matching approach. To register two consecutive images, we select a reference region in one frame and search for it on the other one. The position and size of the reference region, and the image selected to extract it, is chosen dynamically depending on robot movement direction. When the robot moves in the direction pointed by the camera (forward) image points move towards image bottom, so the reference region is extracted from the last video frame, close the last rows and centered. Its width is chosen to be higher than its height. The same area is expected to lay closer to the image center in the previous video frame. As the robot moves aligned with the surface edges, the displacement of the reference region in the horizontal direction would be negligible. The height of the reference region should be high enough as to enclose sufficient texture and, at the same time, small enough to obtain a sharp peak in the vertical direction. When the robot moves backwards, the reference image is the previous one and the pattern is searched for in the last one.

Similarly, when the robot moves opposite to the positive X camera axis (left), the reference region is extracted from the last image. As image points are expected to move to the right, the reference region is taken from the bottom of the image, close to the rightmost visible part of the image. It should be taller than wide following the same reasoning as above. When the robot moves to the right, the reference region is taken again from the last image, but in this case it is located at the leftmost visible part and at the bottom. Fig. 3 summarizes the rules to select the reference region. Similar rules are followed to joint different strips or to joint each camera partial panorama.

C. Mask Computation

The template obtained is searched for in the search image using normalized cross-correlation. To reduce the number of computations and to improve peak sharpness we carry out the calculation only on points that fulfill two conditions:

1) Should present a small distortion after the perspective effect is corrected.
2) Should belong to the plane defined by the surface we want to reconstruct.

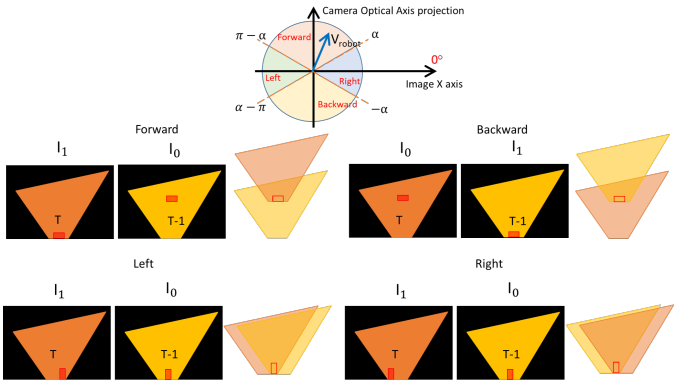

Fig. 3. The reference image (I_1) and the template region is selected according to the angle between the positive horizontal axis and the robot velocity. As an example, when the robot moves forward, the reference image is the last frame, template width is bigger than its height and is located close to be end of the reference image. The template is searched in the search image (I_0) which, in this case is the previous frame.

1) Distortion: Jacobian estimation: The perspective effect introduces a distortion in the image. We use the Jacobian associated to the mapping defined by the homography to estimate the amount of distortion at any pixel in the output image. This value represents the expansion, or shrinkage if the determinant is lower than one, that a pixel in the original image will undergo when it is transformed to its position in the surface. Fig. 4 shows the transformation suffered by the original image and the determinant of the Jacobian for this transform.

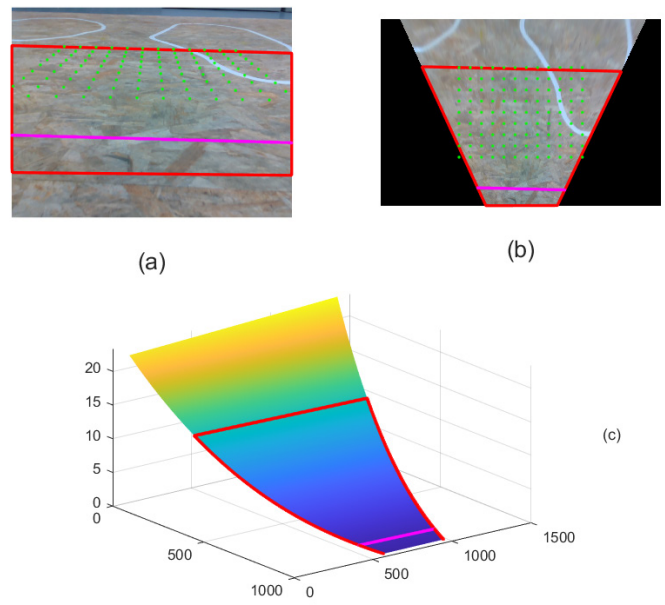

Fig. 4. Perspective correction: (a) shows the original image captured by the camera (640x480). (b) Shows the zenithal view obtained after applying the homography [1500x1000, 1mm/pixel]. (c) Represents the value of the determinant of the Jacobian for every corrected point. Green dots represent the points in the calibration pattern. The red line encircles the area with where the determinant value is below a given threshold (12 in the example). The horizontal magenta line marks the position with no distortion.

The Jacobian at a point $[x,\ y]^T$ of the surface is computed by taking the partial derivatives of the point coordinates with respect to its position in the image $[u,\ v]^T$. Equation 2 shows the value of the Jacobian associated to the homography H at a point $[x,\ y]^T$ on the surface:

$$J_H = \begin{bmatrix} \frac{\partial x}{\partial u} & \frac{\partial x}{\partial v} \\ \frac{\partial y}{\partial u} & \frac{\partial y}{\partial v} \end{bmatrix} = \begin{bmatrix} \frac{h_{11}\sigma_3 - h_{31}\sigma_1}{\sigma_3^2} & \frac{h_{12}\sigma_3 - h_{32}\sigma_1}{\sigma_3^2} \\ \frac{h_{21}\sigma_3 - h_{31}\sigma_2}{\sigma_3^2} & \frac{h_{22}\sigma_3 - h_{32}\sigma_2}{\sigma_3^2} \end{bmatrix} \quad (2)$$

With $\sigma_i = [u,\ v,\ 1]\,[h_{i1},\ h_{i2},\ h_{i3}]^T$. As the determinant of the Jacobian increases and is bigger than one, the original image pixel will expand over a larger portion of the destination image. This will cause a blurring effect in the transformed image which is not desirable. Alternatively, when the value falls below one, a bigger portion of the original image will be mapped to the same output pixel.

2) Plane segmentation: Only points that belong to the plane defined by the surface should be taken into account. During the homography calibration process we compute the vector that represents, in homogeneous coordinates, the plane defined by the corners detected in the calibration pattern. Distance information provided by the RGB-D camera allows to compute the 3D position of these points with respect to the camera reference system. All of them lay in the same plane. This condition can be expressed as:

$$\begin{bmatrix} X_i & Y_i & Z_i & 1 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \\ d \end{bmatrix} = 0 \quad (3)$$

Where $[X_i,\ Y_i,\ Z_i]^T$ is the 3D position of a corner point in the calibration pattern and $[a,\ b,\ c,\ d]^T$ is the vector that describes the plane in homogeneous coordinates. Expressing this condition for every corner point, we get an homogeneous equation system that can be solved applying SVD decomposition. Every time a new frame is acquired, (3) is used to compute the distance of a point to the plane. If the value is lower than a predefined threshold, the point belongs to the plane, otherwise the point is out of the plane. The threshold is proportional to the standard deviation of the distance computed from the points used to adjust the plane. Fig. 5 shows the result of applying this method to segment the surface plane.

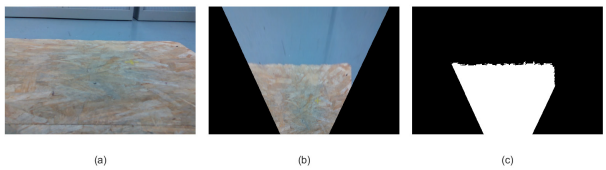

Fig. 5. (a) Captured RGB image. (b) Zenithal view obtained after appling the homographic correction. (c) Points identified as belonging to the surface of interest.

D. Panorama construction

As mentioned before, we follow a hierarchical approach to build the panorama that represents the whole surface. At a first

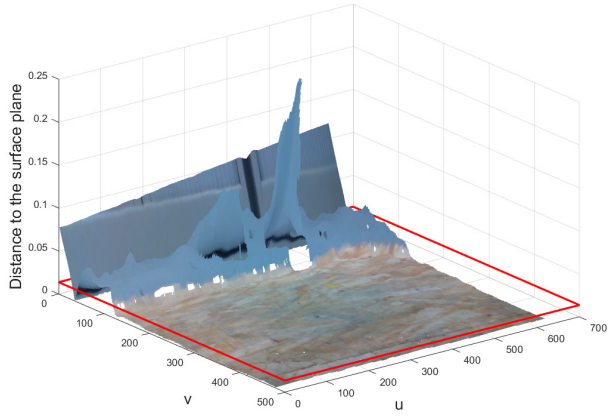

Fig. 6. Computed distance to the plane for the points in Fig. 5.(a). The red line marks the threshold used to classify image points. The proportionality value used is 15.

level, we compute the translation between two consecutive images. This result is used to stitch the new image to the previous one in the same longitudinal path to form a stripe. When a new stripe is completed, we add it to the ones already computed to obtain a panorama for one camera. When the robot completes the scanning trajectory the frontal and rear partial panoramas are stitched in order to get the full panorama of the surface that we want to reconstruct. Throughout the process we also compute a panorama with the values of the determinant of the Jacobian and a mask indicating which pixels belong to the sheet.

1) Stripe panorama: The relationship between the re-projections of the images is always a translation because the robot keeps its heading constant throughout its trajectory. To estimate the translation between two consecutive images, we look for the region in the search image that is more similar to the template extracted from the reference image. We use zero-mean normalized cross-correlation to measure the similarity between two image regions. Only values corresponding to regions completely included in the surface plane are taken into account.

To stitch both images, the template center is moved on top of the highest correlation peak and we select the pixels with a lower Jacobian value to build the new image. Resulting panorama minimizes the distortion induced by the perspective effect. Fig. 7 shows how two consecutive images are stitched together. The example corresponds with a forward motion of the camera. The same transformation is applied to the Jacobian and the plane mask. This is to ensure that it is possible to apply the same process at a higher level.

All the frames that correspond to the same longitudinal section of the trajectory followed by the robot are joined together to build a stripe panorama. The initial image in this panorama is the one that corresponds to a change in the motion direction from left (or right) to a foward (or backguard) one. While new frames are stitched together with a constant movement direction, the translation computed in the frame-to-frame step is used to add the new image to the panorama. Fig. 8 shows an example of a stripe.

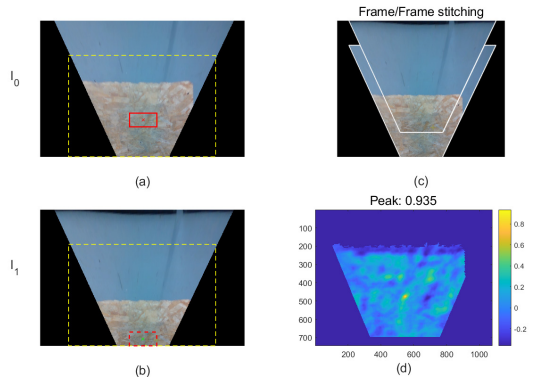

Fig. 7. Two consecutive images taken while the robot moves forward. (a) was taken at time T-1 and is the search image. (b) was taken at time T and is the reference. Dashed yellow rectangle is the bounding box of the mask. Red rectangles show the template (dashed) and its position on the search image (continuous). (c) shows both images stitched. The white outline shows seam line. (d) shows the values of the normalized cross-correlation. Only a single strong peak appears.

2) Camera and Surface panoramas: New stripes from the same camera are stitched together to build the panorama associated to a single camera. The procedure is similar to the one described in the previous section. Fig. 9 exemplifies the process. The main difference is that movements between two neighboring stripes always correspond to a left or right camera movement.

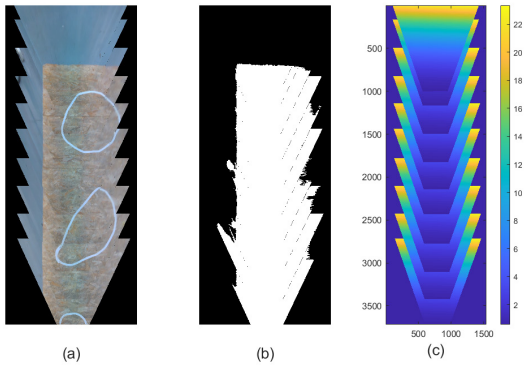

Fig. 8. Example of a stripe panorama. The RGB images are stitched together to form the zenithal view (a). Using the same information, the surface/not surface segmentation masks (b) and the Jacobian values associated to each pixel (c) are stitched together.

Fig. 9.(a) shows the partial camera panorama. Every time a new stripe is completed (Fig. 9.(b)), it is added to the partial camera panorama. In this example, images correspond to the rear camera and the robot was moving to the left. Therefore, the rear camera was moving to the right. As the movement is the horizontal direction, the template height is bigger than its width.

Once the panoramas corresponding to both cameras have been completed, the same procedure is used to stitch them into a single panorama depicting the whole surface to be inspected. Fig. 10 shows one example of the result.

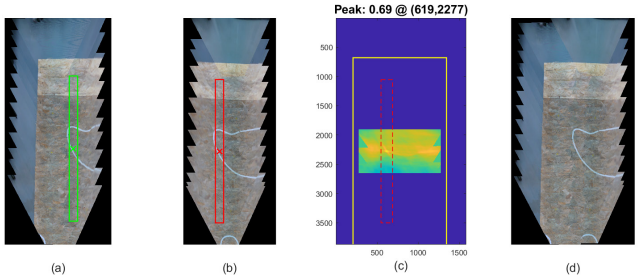

Fig. 9. Neighboring stripes are stitched together to form the panorama of one camera. Initial panorama (a) showing in green the reference region. (b) shows the new stripe and the region paired with the reference one (in red). (c) shows the correlation values. (d) shows the expanded panorama.

III. Experiments and Results

To test the algorithm we have used an omnidirectional Summit XL Steel robot from Robotnik. Two fixed Intel RealSense D435 RGB-D cameras provide the information to reconstruct the surface. Both are oriented 30º to the ground. One is placed at the front of the robot and the other is at the back. The scanned surface is a rectangle of 4400 by 2500 mm formed by seven OSB (Oriented Strand Board) wooden panels. Each panel measures 625 by 2500 mm and is 25 mm thick. Defective areas were marked using white tape.

The algorithms have been developed using MATLAB 2020a and then have been ported to ROS-Kinetic running on a Ubuntu 16.04 LTS operating system. We used the open source OpenCV library to process the images.

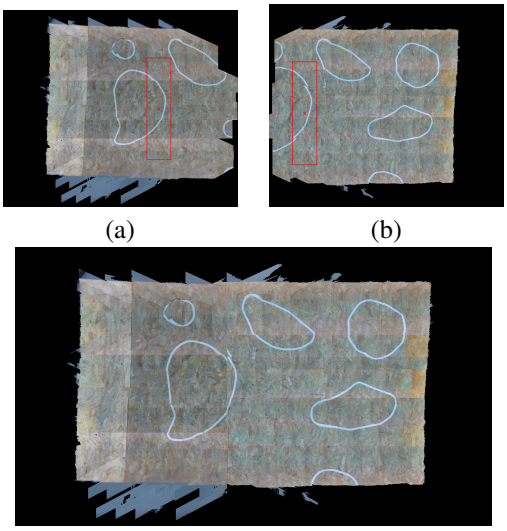

Fig. 10. (a) Rear camera panorama. (b) Front camera panorama. Red rectangles shows the reference region in the front camera and its corresponding one in the rear camera. Both images have been rotated to simplify the representation. In this case, we have considered that the robot is moving backwards. (c) Full surface panorama.

The robot starts at the lower right corner of the surface and align its X axis with detected surface edge using the RGB-D cameras. The length and number of longitudinal runs is not fixed. End points are detected automatically from the information provided by the frontal and rear cameras. The robot moves until its wheels are at a predefined distance from the edge, although the edge may be out of sight. This is to ensure that the camera opposite to the edge maximizes the covered surface while keeping within the surface for security reasons.

To ensure that the robot may cover the whole surface it is compulsory that surface length is big enough. To be sure that the panoramas from each camera overlaps, surface length must be bigger than twice the distance between the fields of view of both cameras. With our robot and camera configuration, the minimum length must be 3000mm. It is possible to reduce this dimension using only one camera, although a more complex scanning trajectory is be needed.

Full panoramas are formed from 100 images and have a resolution of 1mm/pixel. Seven stripes are needed to reconstruct the full width of the surface. The average size of the reconstructed surface is 4339 by 2514 mm. The error in the longitudinal direction is $-1.4 \pm 1.2\%$, while the error in the transversal direction is $0.6 \pm 2.1\%$.

The marks enclosing the areas of interest are correctly reconstructed an positioned on the surface. As these marks present a high contrast, they clearly affect the output of the correlation as shown in Fig. 11. Although in both cases the peak is higher than 0.95, The peak is not so sharp in the direction of the mark. This may cause small localization errors, especially when the mark is aligned with the motion direction.

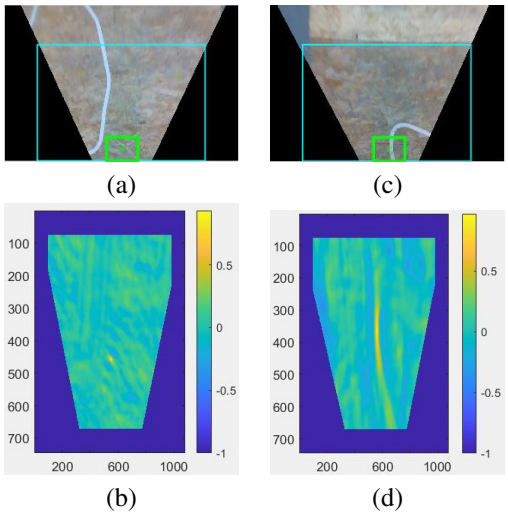

Fig. 11. (a) Rear camera panorama. (b) Front camera panorama. Red rectangles shows the reference region in the front camera and its corresponding one in the rear camera. Both images have been rotated to simplify the representation. In this case, we have considered that the robot is moving backwards. (c) Full surface panorama.

In our experiments, 97.4% of peak values in the frame-to-frame matching are above 0.8, and 70.8% are higher than 0.9. Only 1.3% of peaks values fall below 0.6. In addition, in more than 60% of the frames there is only one peak bigger than 0.5.

Those values fall to a range between 0.4 and 0.7 in the case of stripe stitching. Anyway, a single sharp peak appears in the

correlation values matrix, providing a reliable estimation of camera translation. Correlation peaks between front and rear panoramas drop again and range from 0.25 to 0.4.

IV. Analysis and Conclusions

In this paper we have solved the problem of scanning a big surface for later analysis using a mobile robot equipped with two RGB-D cameras. We assume that the surface has enough texture, although it is not possible to identify reliable image features at the scale provided by each individual image.

Our solution is based on zero-mean normalized cross correlation, a very well known area based registration technique. To improve our results we follow a hierarchical approach based induced by the boustrophedon type path followed by the robot while scanning the surface. At a first stage we compute the translation between to consecutive images as the robot moves forward or backward. Individual images corresponding to the same longitudinal segment of the boustrophedon path are stitched together using the computed translation to form a stripe partial panorama. Neighboring stripes are joined to form a panoramic view of the scanned surface taken by one camera. Front and rear panoramas are finally registered together to form the full surface panorama. Each stripe comprises 11 images on average (ranging from 10 to 12). To cover the full width of our test surface, seven stripes are formed. This yields an average number of 77 images per camera and therefore, 144 images must be stitched together to form a complete panorama.

To reduce ambiguity, we restrict the correlation process to those pixels that lay on the surface. The classification process is based on the depth information provided by the cameras. Using the 3D position of each pixel, we compute the distance to the ground plane. Only those pixels within a small threshold are classified as surface points. The parameters of the ground plane are estimated during the calibration process, using the 3D position of the same points used to solve for the homography that maps the image to the surface plane. This method has proved to be very robust and allow to distinguish the surface of interest from the ground, even though its thickness is only $25mm$. Under some circumstances, some ground points are classified erroneously as surface points as can be seen in Fig. 10.(c). We think that this problem appears due to direct sun illumination which may cause problems in the correct identification of the infrared pattern used by the camera to compute depth.

Our method also introduces the use of the determinant of the Jacobian of the homographic mapping. These value is used as an estimation of the amount of distortion suffered by a pixel when is reprojected from the image to a zenithal view. Values close to one minimize the distortion, while bigger ones produce a blurring effect, as a single pixel is mapped to a big area in the zenithal view. When the value is too small we may loose information because a bigger portion of the input image is mapped to a single pixel after its transformation. In the overlapping areas of the images, the point with a lower Jacobian value survives. The final result is a high resolution surface image with reduced distortion. Estimated mean reconstruction error in the longitudinal direction is -1.4% of the overall length and only 0.6% in the transversal direction.

References

[1] N. Neogi, D. K. Mohanta, and P. K. Dutta, "Review of vision-based steel surface inspection systems," *EURASIP Journal on Image and Video Processing*, vol. 2014, no. 1, p. 50, Nov. 2014. [Online]. Available: https://doi.org/10.1186/1687-5281-2014-50

[2] Q. Luo, X. Fang, L. Liu, C. Yang, and Y. Sun, "Automated Visual Defect Detection for Flat Steel Surface: A Survey," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 3, pp. 626–644, Mar. 2020, conference Name: IEEE Transactions on Instrumentation and Measurement.

[3] Q. Luo, X. Fang, J. Su, J. Zhou, B. Zhou, C. Yang, L. Liu, W. Gui, and L. Tian, "Automated Visual Defect Classification for Flat Steel Surface: A Survey," *IEEE Transactions on Instrumentation and Measurement*, vol. 69, no. 12, pp. 9329–9349, Dec. 2020, conference Name: IEEE Transactions on Instrumentation and Measurement.

[4] X. Sun, J. Gu, S. Tang, and J. Li, "Research Progress of Visual Inspection Technology of Steel Products—A Review," *Applied Sciences*, vol. 8, no. 11, p. 2195, Nov. 2018, number: 11 Publisher: Multidisciplinary Digital Publishing Institute. [Online]. Available: https://www.mdpi.com/2076-3417/8/11/2195

[5] C. Fraga, R. Gonzalez, J. Cancelas, L. Enguita, and L. Loredo, "Camber measurement system in a hot rolling mill," in *Conference Record of the 2004 IEEE Industry Applications Conference, 2004. 39th IAS Annual Meeting.*, vol. 2, Oct. 2004, pp. 897–902 vol.2, iSSN: 0197-2618.

[6] S. Rodríguez-Jiménez, I. Álvarez, R. García, and J. Marina, "Analysis of Block Matching Algorithms for the application of image mosaicing to online surface inspection of steel products," in *2010 IEEE International Conference on Imaging Systems and Techniques*, Jul. 2010, pp. 214–219, iSSN: 1558-2809.

[7] T. Schlagenhauf, T. Brander, and J. Fleischer, "A stitching algorithm for automated surface inspection of rotationally symmetric components," *CIRP Journal of Manufacturing Science and Technology*, vol. 35, pp. 169–177, Nov. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1755581721000833

[8] R. Szeliski, "Image alignment and stitching: A tutorial," *Foundations and Trends® in Computer Graphics and Vision*, vol. 2, no. 1, pp. 1–104, 2006, publisher: Now Publishers Inc. Hanover, MA, USA.

[9] B. Zitová and J. Flusser, "Image registration methods: a survey," *Image and Vision Computing*, vol. 21, no. 11, pp. 977–1000, Oct. 2003. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0262885603001379

[10] M. Xia, J. Yao, R. Xie, L. Li, and W. Zhang, "Globally consistent alignment for planar mosaicking via topology analysis," *Pattern Recognition*, vol. 66, pp. 239–252, Jun. 2017, publisher: Elsevier Ltd.

[11] M. Debella-Gilo and A. Kääb, "Sub-pixel precision image matching for measuring surface displacements on mass movements using normalized cross-correlation," *Remote Sensing of Environment*, vol. 115, no. 1, pp. 130–142, Jan. 2011, publisher: Elsevier. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S003442571000252X

[12] R. G. Llenderrozos, I. l. García, J. M. E. González, and S. R. Jiménez, "Automatic area based registration method and its application to the surface inspection of steel industry products," in *Videometrics, Range Imaging, and Applications XII; and Automated Visual Inspection*, vol. 8791. International Society for Optics and Photonics, May 2013, p. 87911I. [Online]. Available: https://www.spiedigitallibrary.org/conference-proceedings-of-spie/8791/87911I/Automatic-area-based-registration-method-and-its-application-to-the/10.1117/12.2020543.short

[13] R. Xie, J. Yao, K. Liu, X. Lu, Y. Liu, M. Xia, and Q. Zeng, "Automatic multi-image stitching for concrete bridge inspection by combining point and line features," *Automation in Construction*, vol. 90, pp. 265–280, Jun. 2018. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0926580518301237

[14] K. Chaiyasarn, T.-K. Kim, F. Viola, R. Cipolla, and K. Soga, "Distortion-Free Image Mosaicing for Tunnel Inspection Based on Robust Cylindrical Surface Estimation through Structure from Motion," *Journal of Computing in Civil Engineering*, vol. 30, p. 04015045, Aug. 2015.