SURVEYING AND GEOSPATIAL ENGINEERING JOURNAL

Targetless Coregistration of Terrestrial Laser Scanning Point Clouds Using a Multi Surrounding Scan Image-Based Technique

Bashar Alsadik

Department of Earth Observation Science, ITC Faculty, University of Twente, The Netherlands, b.s.a.alsadik@utwente.nl

Abstract

The coregistration of terrestrial laser point clouds is widely investigated where different techniques are presented to solve this problem. The techniques are divided either as target-based or targetless approaches for coarse and fine coregistration. The targetless approach is more challenging since no physical reference targets are placed in the field during the scanning. Mainly, targetless methods are image-based and they are applied through projecting the point clouds back to the scanning stations. The projected 360° point cloud images are normally in the form of panoramic images utilizing either intensity or RGB values, and an image matching is followed to align the scan stations together. However, the point cloud coregistration is still a challenge since ICP like methods are applicable for fine registration. Furthermore, image-based approaches are restricted when there is: a limited overlap between point clouds, no RGB data accompanied to intensity values, and unstructured scanned objects in the point clouds. Therefore, we present in this paper the concept of a multi surrounding scan MSS image-based approach to overcome the difficulty to register point clouds in challenging cases. The multi surrounding scan approach means to create multi-perspective images per laser scan point cloud. These multi-perspective images will offer different viewpoints per scan station to overcome the viewpoint distortion that causes the failure of the image matching in challenging situations. Two experimental tests are applied using point clouds collected in Enschede city and the published 3D toolkit data set in Bremen city. The experiments showed a successful coregistration approach even in challenging settings with different constellations.

I. INTRODUCTION

Terrestrial laser scanning TLS is widely used nowadays as a standard technique for collecting reliable 3D point clouds of the scanned objects and used for different applications in mapping, mining, as-built surveying, civil engineering, architecture, archaeology, and city modeling, etc.

Since the basic idea of the data acquisition is static, a need for multi-station setups is necessary to ensure the complete coverage of the object of interest. The resulted point clouds at the multi-scan stations are integrated together to compose a one complete point cloud in what is called the coregistration step (Fig.1).

Mathematically, the most commonly adopted solution for 3D registration is the well-known Iterative Closest Point (ICP) procedure. The ICP standard approach performs a fine registration of two coarse aligned overlapping point clouds by iteratively estimating the transformation parameters that relate to the closest Euclidean distances [1, 2].
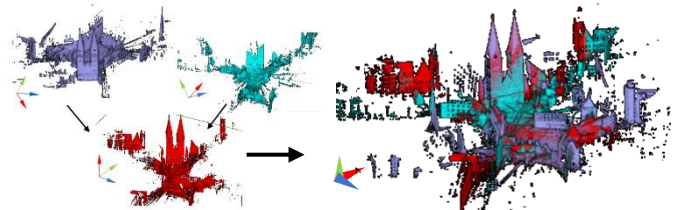

Fig. 1. Coregistration principle of multiple point clouds.

The ICP method showed a high performance in fine coregistration where first a good initial coregistration is applied either manually by selecting a few common points, or by using artificial targets or automatically using targetless techniques. Several improvements and variations of ICP methods have been investigated to automate the coregistration procedure either

ipAcademia
www.ipacademia.org

using target-based or targetless approaches.

Target-based approaches are the most commonly employed and implemented techniques in commercial software packages for static terrestrial laser scanning TLS projects. These targets are normally highly-reflective materials with a certain shape that can be automatically detected and matched between scans [3]. On the other hand, interest is shown to apply the fully automated co-registration scan-to-scan or photo-to scan without using any physical targets. Automated approaches are mainly feature-based methods that establish correspondences by normally comparing the detected feature descriptors either using the 3D features of lines and planes or 2D image features. Generally, these feature-based methods showed interesting results, however, the disadvantage of these geometric primitive based methods is the required processing time.

As mentioned, the majority of 3D coregistration techniques are relying mainly on automatic detection of geometric features like points, lines, or planes out of the TLS point cloud. Among different methods of automatic 3D features detection, there are three conventional Line/Plane detection techniques of: Hough transform, RANSAC, and Region growing (incremental). The authors stated that the best practical approach is to apply a sequence of two or three methods in order to capture as many features as possible. For photo to scan coregistration approach, 3D based methods can rely on the derivation of the image-based 3D model (point cloud or mesh) after applying a dense image matching. Subsequently, 3D planer patches or 3D lines are detected and then a transformation between the TLS point cloud data and the 3D image-based model can be applied.

In the context of 3D targetless coregistration, other researchers tried to avoid point detection since there is no guarantee about the accuracy of detecting these points and what kind of robust descriptors they should use. However, Swart, et al. [4] stated that this is true for static terrestrial laser scanning TLS and not in mobile laser scanning MLS applications where the point cloud density is lower and affects the detection of lines and plane features.

Theiler, et al. [5] presented a valuable work on the automated 3D point detection and description of point cloud registration by what they called Keypoint-based 4-Points Congruent Sets (K-4PCS). As mentioned earlier, the challenge is in the initial registration where the ICP methods can then work efficiently for the fine registration and this is what the authors tried to solve. Some modified ICP versions utilizing lines and planes are also presented and used widely and tried to investigate the use of detected line features in point cloud co-registration [6-8]. [9] discussed the use of what is called ICL (Iterative Closest Line) in the 3D co-registration of TLS pair of point clouds.

Thapa, et al. [10] stated that extracting and matching the geometric features are vulnerable to the variety of point cloud density and symmetries in the scene and can also be computationally expensive. Therefore, they proposed a coregistration method that introduces knowledge to identify the semantic meaning of each geometric feature that will enable a straightforward feature matching. Changchang, et al. [11] presented the novel viewpoint invariant patch (VIP) which provides the required properties to determine the similarity

transformation between two 3D models even in viewpoint significant changes. In their method image textures are rectified with respect to the local geometry of the scene. The produced 3D image-based model is used to create an ortho-texture which is viewpoint independent and the 3D models are then transformed to a set of VIPs, made up of the feature's 3D position, surface normal, patch scale, local gradient orientation in the patch plane, and a SIFT descriptor. The deep information on VIP features makes them suited to 3D similarity transformation estimation.

The second widely used approach is a 2D based coregistration where the point cloud intensity values are projected to create intensity images. These intensity images between point clouds are used to register them together by image feature matching techniques. Further, these intensity images can be used in photo to scan applications by matching them to the RGB images from the same field of view [12]. Böhm and Becker [13] also presented a photo to scan coregistration implementation where a SIFT operator is used in image matching. Houshiar, et al. [14] presented a paper for the scan to scan registration based on 2D feature image matching by using mainly SIFT or SURF detector-descriptor operators. The paper showed an extensive study on the effect of the image projection on the final coregistration results. The results showed that Mercator or Pannini projections are the best compared to equirectangular, stereographic, or azimuthal (Fig.2).
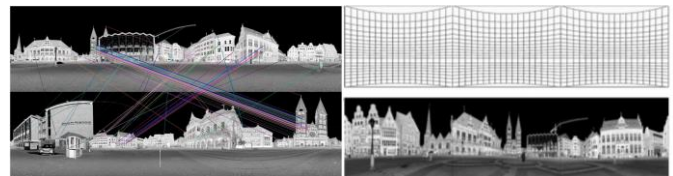

Fig. 2.   Scan to scan registration implementing different map projections [14].

Recent research line [15, 16] based on using the deep learning techniques and convolutional neural networks CNN is rising up but still in its first steps.
Fig.3 summarizes some of the main coregistration techniques of the scan to scan and photo to scan either in 2D or 3D approaches.
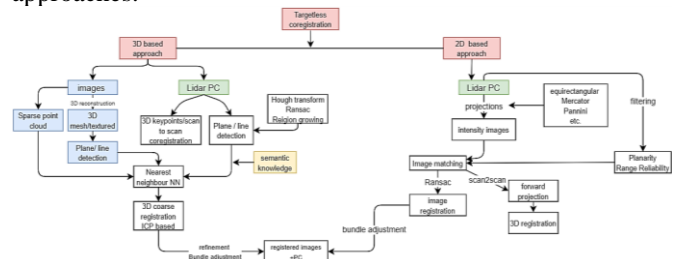

Fig. 3.   A summary of most targetless point cloud coregistration techniques.

As mentioned, when using TLS, it's only feasible to measure the surrounding environment that is visible from the scan device position. Accordingly, occlusions occur when objects around are blocking the line of sight from the scanner. To overcome this problem, multiple setups with 360˚ scanning angles are used to ensure a more complete level of coverage [17]. As presented in the literature, automated point cloud coregistration is still a challenging assignment in many cases

where there is a lack of correspondences between the multiple 360˚ point clouds. As stated, targetless coregistration of the multiple point clouds can be applied by back projecting the object points to create images at the scan stations. The image projection type can be panoramic spherical (equirectangular), cylindrical or perspective (rectilinear). However, these mentioned projections can be insufficient to ascertain a successful image matching due to the perspective distortion in the viewing directions between the overlapped point clouds. Furthermore, the type of projection can add a challenge to the matching operator based on the projection distortion characteristics.

Accordingly, in this paper, a new image-based technique is proposed to solve the coregistration problem of 360˚ point clouds efficiently even in challenging multi-scan configurations. The suggested method is image-based which considers multiple image projections surrounding the original scan station. This idea is expected to increase the chances of successful image matching between point cloud scans since it offers different perspectives and scales that the matching operators and descriptors can benefit as will be shown in the following sections. The suggested method will be presented in detail in section 2 while two experiments are shown in section 3 and followed by discussion and conclusions in section 4.

II. Methodology

In this section, the general concept of the proposed method will be explained in section 2.1 while the mathematical description of projecting a 360˚ point cloud into a planar image will be shown in section 2.2.

A. General Concept

As stated, targetless image-based coregistration can fail due to the lack of feature correspondences between point cloud projected images. The difficulty to match the projected images either cylindrical, spherical or perspective at the TLS 360˚ stations can be avoided by a multi surrounding stations MSS approach as suggested in this paper. The MSS approach concept is to project the point cloud for each scan station into virtual locations around the real scan station where the instrument was set up. The motive is to have more redundant images at different perspectives to increase the chances of the image correspondences between the scan stations. The multiple stations can be created either in 4 or 8 image-configuration (Fig.4) or else based on the user selection and the complexity of the scanned environment. The motivation is to avoid the wide-angle baseline viewing that might affect the success of the matching between the TLS images besides offering different scales and perspectives to the common objects in the corresponding images. Fig. 4 shows the suggested two patterns of the multi images per scan setup. It should be noted that the distance between the original scan station and its surrounding virtual stations is dependent on the available accessibility and the user intuition.

Furthermore, the virtual images are created either as perspective (rectilinear) images that altogether cover the 360˚ space or as panoramic images. Afterward, the image matching technique is applied in the conventional pipeline within the structure from motion SfM approach. The final result is a fully

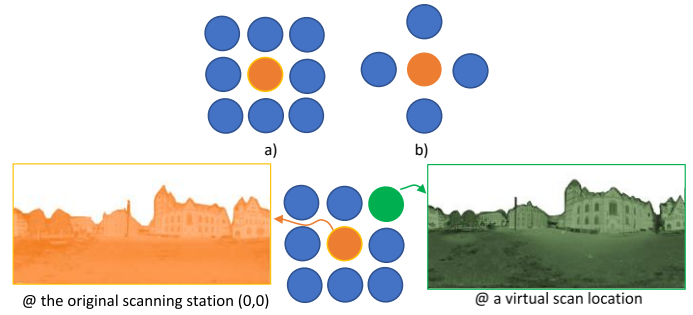aligned network of MSS images that actually represent the coregistered TLS stations.

Fig. 4. The concept of the multi surrounding stations MSS. a) 8 set up configuration. b) 4 set up configuration. (orange=original TLS station, blue=virtual TLS station).

B. Spherical to Planar Projection

As described in the methodology, TLS 360˚ point clouds will be projected into either spherical or cylindrical panoramas or into perspective images. There are different mapping techniques to project a spherical area into a planar surface by assuming cylinders or cones covering the sphere which are then unfolded into that planar surface. However, there is always a magnitude of distortion found in the projections either in areas, angles or scale. One of the conventional projections is called the equidistance cylindrical projection. This projection is neither equal-area nor conformal where the parallels and meridians are equidistant straight lines, intersecting at right angles and the sphere poles are projected into straight lines equal in length to the equator (Fig.5). The distortion in areas and angles increase with distance from the equator of the cylindrical projection while the scale is true along all meridians and equator (equidistant).

Mathematically, the planar projected coordinates x, y are calculated as follows assuming a unit radius [18]:

$$x = (\lambda - \lambda_o) \, cos\varphi_1$$
$$y = (\varphi - \varphi_1) \qquad (1)$$

Where
λ, φ: spherical coordinates.
λ_o : central meridian.
φ_1 : standard parallel.

The simplest form of the mentioned equidistant cylindrical projection is the *Plate Carrée* projection which is also called *equirectangular* or *spherical*. This projection was used often in the 15th century and is quite common today in very simple computer mapping programs. Most of the street panoramic images nowadays are composed and visualized using this projection like Google Street View [19]. Mathematically, the equirectangular projection is the special case where φ_1 in "(1)" is zero [20].
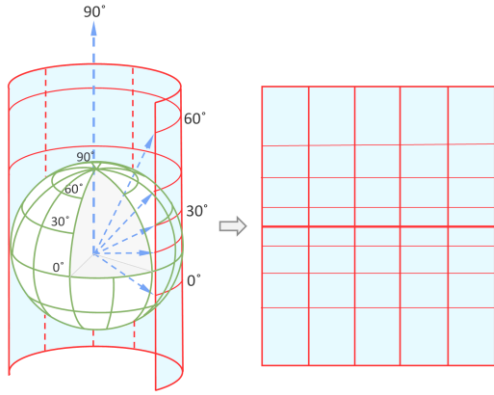
Fig. 5. Equirectangular projection concept.

The pseudo-code to project the 360˚ point cloud into a spherical or cylindrical panorama is shown as follows in Table I.

TABLE I.   ALGORITHM OF THE PROJECTION OF 360˚ POINT CLOUD INTO A PANORAMIC IMAGE.

- Load the scanned point cloud pc.
- Define the size of the projection image (W, H).
- Check the visibility of points.
  For each point i.
- Convert the Cartesian XYZ coordinates into polar coordinates azimuth, elevation, and distance r.
- Transform the polar coordinates into image pixel coordinates using affine.
- Read the associated RGB or intensity values of point I .
- Apply color/intensity image resampling.
  Repeat.

On the other hand, perspective projection can be attained either by projecting the point cloud to the image plane by collinearity equations [21] or by projecting the equirectangular image composed in the former step to a perspective (rectilinear) using what is called Gnomonic map projection [22]. The advantage of the rectilinear projection is to produce an image similar to what a human sees in the world where straight lines stay straight, thus produce undistorted images.

The Gnomonic projection can be achieved by projecting the ray of light from the center of the sphere through its surface at points which touches on a tangent plane (Fig.6). At the tangent point, there is zero distortion, but it increases whenever going away from it and then projects great circles on straight lines.
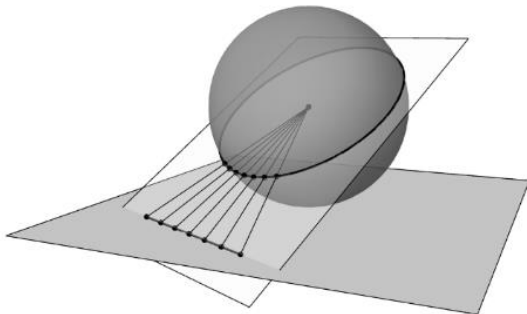

Fig. 6. The projection of a sphere onto a planar surface using Gnomonic Projection [22, 23].

The transformation equation of the plane coordinates x, y using the Gnomonic projection is given as [24]:

$$x = \frac{cos\varphi \, \sin{(\lambda - \lambda_o)}}{\cos c} \tag{2}$$

$$y = \frac{cos\varphi_1 \, sin\varphi - sin\varphi_1 cos\varphi \, \cos{(\lambda - \lambda_o)}}{\cos c} \tag{3}$$

$$cos \, c = sin\varphi_1 \, sin\varphi + cos\varphi_1 \, cos\varphi \, cos \, (\lambda - \lambda_o) \tag{4}$$

Where c is the angular distance of a point (x, y) from the center of the projection.

III. EXPERIMENTAL TESTS

Two experimental tests are applied to illustrate the efficiency of the proposed method. In both experiments, the proposed MSS is applied where the TLS intensity images are processed to finally achieve the coregistration of the point clouds. Moreover, creating the virtual surrounding perspective images is applied using a field of view of 45˚. As a result, eight overlapped images around the horizon are composed. The resolution of 1920×1080 is selected to mimic the conventional HD video frames and to guarantee sufficient image matching.

A. First experiment

This experimental test is applied on a TLS data set collected for the old church in the city center of Enschede in the Netherlands, where nine scan stations are performed using a Velodyne HDL32 Lidar type. The point clouds of some scan stations are shown beside the registered colored point cloud in Fig. 7.
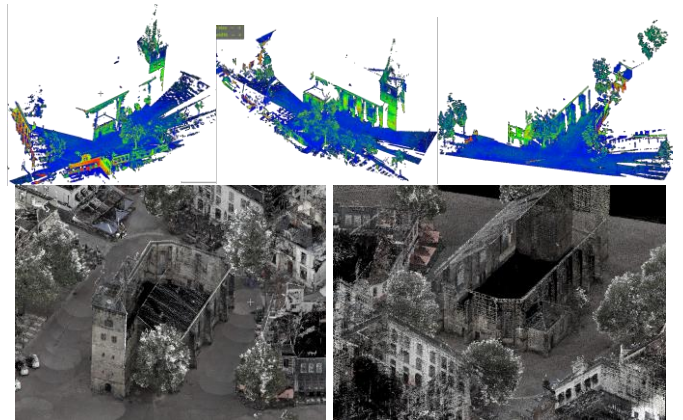

Fig. 7. Some of the scanned point clouds and the coregistered old market church in Enschede.

The 4-MSS constellation concept is applied for the coregistration of the multiple point clouds with the mentioned resolution of 4800×2400 pixels for the spherical images and 1920×1080 for the perspective images. Fig.8 illustrate this configuration for one TLS scan using 5 meters grid distance. Furthermore, the Metashape software [25] is used to apply the alignment task between both types of images after filtering out the low quality images.
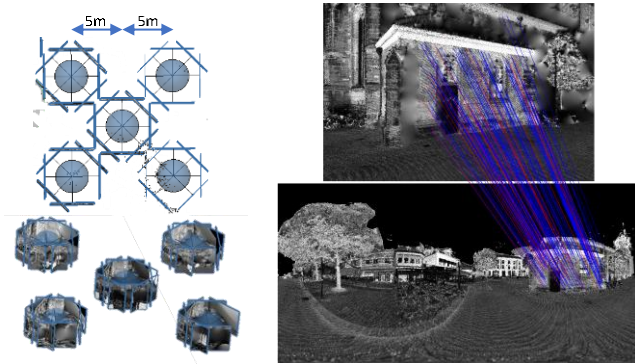
Fig. 8. Left) the 4-MSS constellation around the original station. Right) image matching sample between a virtual perspective image and a spherical TLS scan image.

The final image network of 393 images and the coregistered TLS scan stations are shown in Fig.9 where the perspective and the equirectangular images are aligned together using the SfM technique. It's worth to mention that the average ground sample distance GSD of the images is ≅2 cm.
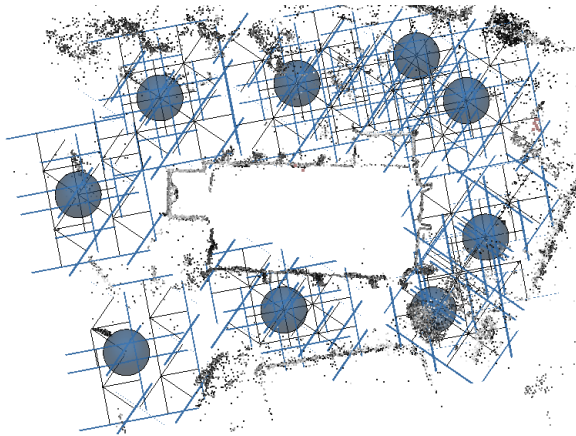

Fig. 9. The final coregistered TLS scans stations using the MSS technique.

B. Second experiment

The second test is applied to a multi 360˚ point cloud data set published by [26] which was recorded using a Riegl VZ400 TLS in the city center of Bremen (Fig.10).
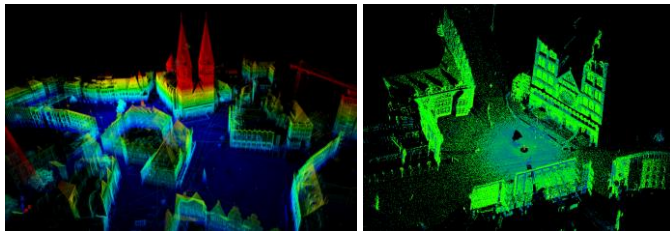

Fig. 10. Left) The full registered point cloud as published in [26]. Right) One TLS point cloud.

This data set has two challenges: TLS point clouds having only intensity values without RGB data similar to the first test and there are some narrow alleys where the corresponding features in the adjacent scanned point clouds are minimal. Therefore, in this test, we used the 8-MMS constellations instead of the 4-MMS shown in the previous test to increase the redundancy and to ensure the successful intensity image matching.

For every 360˚ scan, MSS technique is applied and nine virtual stations are created (including the original scan location) at a separation distance of 10 meters. Then for every virtual scan station, a constellation of eight perspective images covering the horizon is created. Accordingly, we have 12×9×8=864 images in total. Similarly, we can add at each original station, the equirectangular images for strengthening the created image network. Then, the images are aligned using specialized image-based modeling software like the Metashape tool[24, 25]. As mentioned, the perspective images in this test are created at a resolution of 1920×1080 pixels while the equirectangular images are created at 4800×2400 pixels. It's worth to mention that the average ground sample distance GSD of the images is ≅15 cm. Fig.11 illustrates a one scan station equirectangular image and its projection into eight overleaped perspective images.
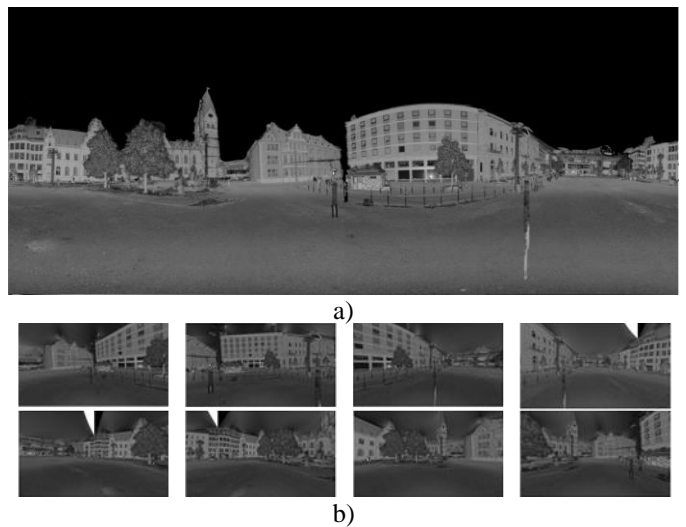

a)

b)

Fig. 11. a) Equirectangular intensity image at a 360˚ TLS point cloud. b) Eight perspective intensity images at the same TLS station.

Fig.12a illustrates an example of the successful matching between two intensity virtual perspective images belongs to two different scan stations at a challenging perspective. The matching was not possible without having multiple virtual images created around the original scan station using the proposed MSS. Fig.12b illustrates another example when we added the equirectangular images for a more redundant matching solution.
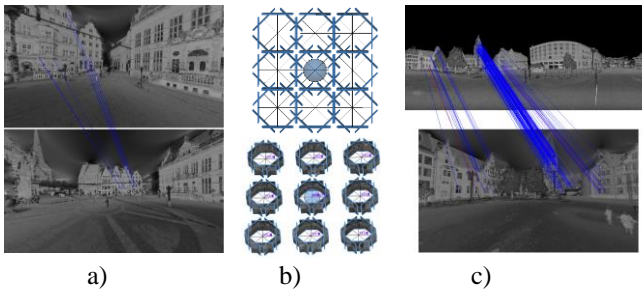
Fig. 12. a) One example of the successful image matching between two virtual perspective images at two difficult to register TLS stations. b) The 8-MSS constellation around the original station. c) Successful image matching between an equirectangular intensity image and a virtual perspective images at two different TLS stations.

After the image alignment step, all the image network was successfully oriented and the TLS scans are coregistered as shown in Fig.13.
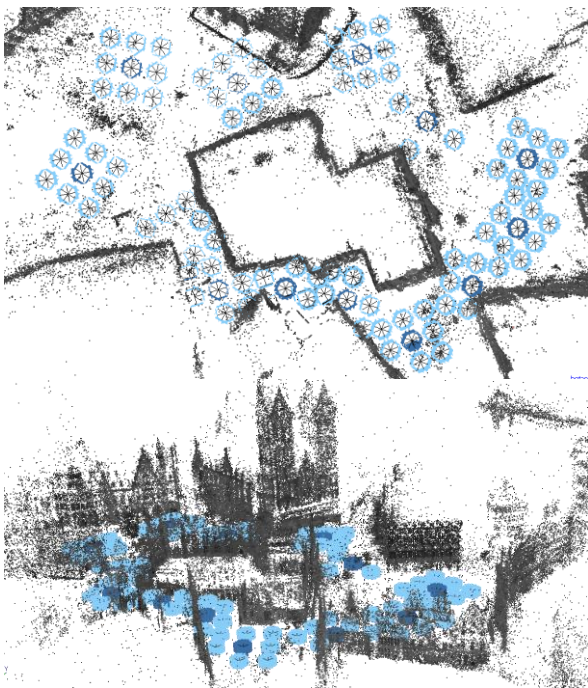

Fig. 13. The complete coregistration of the 12 point cloud dataset using the proposed MSS image-based technique.

IV. DISCUSSION AND CONCLUSIONS

In this paper, a multi surrounding station MSS procedure is proposed for the coregistration of TLS intensity-based point clouds. The suggested MSS procedure was motivated by providing multi-perspective viewpoints at the same scan station to increase the success chances of the image matching operators and then the coregistration. A combination of equirectangular images and perspective images are used in the coregistration to increase the redundancy and to finally get the transformation parameters between the original scans.

The results of both experimental tests shown in Fig. 9 and Fig. 13 illustrated the success of the suggested procedure where the MSS virtual images strengthened the image matching. What is noticed in the experiments is the usefulness of using both types of images: perspective and equirectangular to strengthen the solution and end up with a successful image matching rather than using only one type of image. One disadvantage realized in the second experiment when used the 8-MSS rather than the 4-MSS constellation. This was double the number of images for matching and represented a significant increase in the processing time. However, a guided image matching rather than a full pairwise can be followed to reduce the processing time significantly.

It's worth to mention that for a highly accurate coregistration, it's recommended to follow the MSS results with a fine coregistration using the ICP algorithm.

Future work will continue to investigate the efficiency of using different types of map projections on the success rate of the image matching and possible ways to speed up the processing time for a faster coregistration.

REFERENCES

[1] Yang Chen and G. Medioni, "Object modelling by registration of multiple range images," *Image and vision computing,* vol. 10, no. 3, pp. 145-155, 1992.

[2] P. J. Besl and N. D. McKay, "A Method for Registration of 3-D Shapes," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 14, no. 2, pp. 239-256, 1992, doi: 10.1109/34.121791.

[3] Sayyad Shahzad and M. Wiggenhagen, "Co-registration of terrestrial laser scans and close range digital images using scale invariant features," 2010.

[4] A. Swart, J. Broere, R. Veltkamp, and R. Tan, "Refined non-rigid registration of a panoramic image sequence to a LiDAR point cloud," presented at the Proceedings of the 2011 ISPRS conference on Photogrammetric image analysis, Munich, Germany, 2011.

[5] P. W. Theiler, J. D. Wegner, and K. Schindler, "Markerless point cloud registration with keypoint-based 4-points congruent sets," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.,* vol. II-5/W2, pp. 283-288, 2013, doi: 10.5194/isprsannals-II-5-W2-283-2013.

[6] K.-L. Low, "Linear least-squares optimization for point-to-plane icp surface registration," *Chapel Hill, University of North Carolina,* vol. 4, 2004.

[7] A. F. Habib, M. S. Ghanma, and M. Tait, "Integration of lidar and photogrammetry for close range applications.," presented at the Proceedings of the ISPRS Geo-Imagery Bridging Continents, 2004.

[8] M. Alshawa, "ICL: Iterative Closest Line a Novel Point Cloud Registration Algorithm Based on Linear Features," in *ISPRS 2nd summer school in Ljubljana*, Slovenia, 2007-07 2007, pp. 1-6. Available:https://halshs.archives-ouvertes.fr/halshs-00280659.

[9] M. Alshawa, "ICL: Iterative Closest Line a Novel Point Cloud Registration Algorithm Based on Linear Features," presented at the ISPRS 2nd summer school in, Ljubljana,Slovenia., 2007.

[10] A. Thapa, S. Pu, and M. Gerke, "Semantic Feature Based Registration of Terrestrial Point Clouds," in *Laserscanning '09 Commission III, WG 2*, Paris, France, M. P.-D. F. Bretar, G. Vosselman, Ed., 2009, vol. XXXVIII-3/W8, 2009: ISPRS Archives, pp. 230-235

[11] W. Changchang, B. Clipp, L. Xiaowei, J. M. Frahm, and M. Pollefeys, "3D model matching with Viewpoint-Invariant Patches (VIP)," in *Computer Vision and Pattern Recognition, 2008. CVPR 2008. IEEE Conference on*, 23-28 June 2008 2008, pp. 1-8, doi: 10.1109/CVPR.2008.4587501.

[12] R. Wang, F. P. Ferrie, and J. Macfarlane, "Automatic registration of mobile LiDAR and spherical panoramas," in *2012 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 16-21 June 2012 2012, pp. 33-40, doi: 10.1109/CVPRW.2012.6238912.

[13] J. Böhm and S. Becker, "Automatic marker-free registration of terrestrial laser scans using reflectance," *Proceedings of 8th Conference on Optical 3D Measurement Techniques, Zurich, Switzerland,* pp. 338-344, 2007.

[14] H. Houshiar, J. Elseberg, D. Borrmann, and A. Nüchter, "A study of projections for key point based registration of panoramic terrestrial 3D laser scan," *Geo-spatial Information Science,* vol. 18, no. 1, pp. 11-31, 2015/01/02 2015, doi: 10.1080/10095020.2015.1017913.

[15] W.-C. Chang and V.-T. Pham, "3-D Point Cloud Registration Using Convolutional Neural Networks," *Applied Sciences,* vol. 9, no. 16, p. 3273, 2019.

[16] Y. Wang and J. Solomon, "Deep Closest Point: Learning Representations for Point Cloud Registration," in *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, 27 Oct.-2 Nov. 2019 2019, pp. 3522-3531, doi: 10.1109/ICCV.2019.00362.

[17] C. Thomson. "Improve point cloud registration with targetless scanning." Vercator. https://info.vercator.com/blog/improve-point-cloud-registration-with-targetless-scanning (accessed April 1st, 2020).

[18] J. P. Snyder, "Map projections: A working manual," in "Professional Paper," Washington, D.C., Report 1395, 1987. [Online]. Available: http://pubs.er.usgs.gov/publication/pp1395

[19] "Street View." Google Maps. https://www.google.com/streetview/ (accessed March 9th, 2019).

[20] "Equirectangular projection." Wikipedia. https://en.wikipedia.org/wiki/Equirectangular_projection (accessed March 3rd, 2020).

[21] T. Luhmann, Robson, S., Kyle, S. and Boehm, J., *Close-range photogrammetry and 3d imaging*, 2nd ed. Berlin: De Gruyter (De Gruyter textbook), 2014.

[22] N. S. Mutha. "How to map Equirectangular projection to Rectilinear projection." http://blog.nitishmutha.com/equirectangular/360degree/2017/06/12/How-to-project-Equirectangular-image-to-rectilinear-view.html (accessed March 3rd, 2019).

[23] "Gnomonic projection." Wikiwand. https://www.wikiwand.com/en/Gnomonic_projection (accessed April 11th, 2020).

[24] E. Weisstein. "Gnomonic Projection." Wolfram Mathworld. https://mathworld.wolfram.com/GnomonicProjection.html (accessed March 3rd, 2020).

[25] Agisoft. "AgiSoft Metashape." http://www.agisoft.com/downloads/installer/ (accessed.

[26] D. Borrmann and A. Nüchter. "Robotic 3D Scan Repository." http://kos.informatik.uni-osnabrueck.de/3Dscans/ (accessed December 15th, 2017).