

Interactive Reparation of a Distorted Street-View Panorama for Efficient 3D Façade Modeling

Qing Zhu, Mier Zhang, Han Hu and Feng Wang

Abstract—Façade features are important in large-scale LoD-3 reconstruction in urban environments, and street-view panoramas are arguably the best option for detailed 3D façade modeling. However, despite the plethora of street-view panoramas available, few studies have explored the metric capabilities of panoramas. This is due in part to the complexities of system integration and in part to problems associated with projection (e.g., distortion at the tops of buildings) and deformation (e.g., the bending of straight structures). In an effort to solve these problems, this study introduces a flexible and practical solution using only a single panorama. The key is to efficiently rectify panoramas using image-space line constrained deformation inspired by the as-rigid-as-possible deformation of surface meshes. The image is then re-projected using gnomonic projection on a properly selected tangent plane. The proposed approach requires a reasonable amount of user interaction to select and position the vertical line segments. The tangent point is also chosen empirically for each panorama. The rectified images can then be imported into off-the-shelf 3D modeling solutions as reference images for interactive sketching. Experimental evaluations reveal the effectiveness of the image-space rectification: after proper scaling, the semantic-aware 3D façade models achieve decimeter-level accuracy with respect to the reference surface mesh.

Index Terms—3D Modeling, Panorama, Façade, Distortion Rectification

I. INTRODUCTION

The 3D modeling of building façades in urban environments is part of LoD-3 (Level-of-Detail) reconstruction [1] and an important prerequisite for applications ranging from urban design and disaster management [2] to forms of entertainment such as virtual reality and movies [3]. Unlike the traditional nadir images designed for building roofs [4], aerial oblique images, mobile laser scanning and street-view panoramas are common datasets that have the capability to reveal building façades. However, aerial oblique images suffer from severe occlusion, limited spatial resolution and perspective deformation [5], and mobile laser scanning is confined by the inhomogeneity of the point samples and the incompleteness of

Manuscript received: September 03, 2019; revised: October 29, 2019. This work was supported in part by the National Natural Science Foundation of China (Projects No.: 41631174, 6102392, 41871291) and in part by the National Key Research and Development Program of China (Project No.: 2018YFB0505404). (Corresponding author: Han Hu)

Qing Zhu, Mier Zhang, and Feng Wang are with the Faculty of Geoscience and Environmental Engineering, Southwest Jiaotong University, Chengdu, China. (e-mail: zhuq66@263.net; mierzhang@my.swjtu.edu.cn; wangfeng19920903@my.swjtu.edu.cn)

Han Hu is with the State Key Laboratory of Rail Transit Engineering Informationization, China Railway First Survey and Design Institute Co. Ltd., Xi'an, China and the Faculty of Geoscience and Environmental Engineering, Southwest Jiaotong University, Chengdu, China. (email: han.hu@swjtu.edu.cn)



Fig. 1: Deformation and projection problems of panorama images. In an equirectangular projection, the top parts shrink and the vertical lines bend due to deformation.

the data acquisition, and is thus more suitable for road features [6] than façades. Street-view panoramas, because of their 360° field of view and wide availability in the community [7], are arguably the best option for detailed façade modeling.

Previous work has produced 3D reconstructions of façades by generating textured models represented as irregular triangles [8]. Most previous approaches, such as original perspective images [9], laser scanner [10], LoD-2 models [11] and radar [12], require a lot of auxiliary information, which decreases their flexibility.

Although panoramas are readily available and most 3D modeling software supports image referencing for efficient modeling [13], there are few 3D modeling solutions that use a single panorama. This is due to inherent problems of panoramas that have not yet been solved, as shown in Figure 1: (1) *Image deformation*. Street-view panoramas are generally mosaicked from multiple cameras. In theory, if the extrinsic and intrinsic parameters of the cameras were calibrated perfectly, then the vertical lines on building façades should be preserved. However, in practice, it is quite common for the vertical lines to be bent. (2) *Projection of spherical cameras*. The equirectangular projection of spherical cameras severely distorts the upper part of panoramas, and cannot preserve metric properties.

To solve these problems with panoramas, this study proposes an image reparation approach using as-rigid-as-possible (ARAP) deformation and gnomonic rectification. The repaired images can then be used in existing solutions such as SketchUp [13] to improve the efficiency of interactive modeling. Inspired by the prominent ARAP deformation of surface meshes [14], we cover the panorama with a gridded mesh and cast the problem of straight line reparation as an ARAP-like deformation of the planar mesh [15], [16]. As the purpose is to straighten bent façade structures, interactively selected line constraints are also explicitly positioned at the desired locations in the proposed method. To solve the anisometric problem caused



Fig. 2: Overview of the proposed method. The ARAP deformation and gnomonic projection are responsible for repairing curved structures and rectifying anisometric parts of the images, respectively.

by the equirectangular projection, this study proposes a piecewisely gnomonic rectification with just a mild assumption that façades are generally aligned vertically. The proposed method empirically chooses suitable tangent planes for the gnomonic projection and solves the metric problem without additional position and attitude information being required from the sensor.

In summary, this study makes two main contributions: (1) an improved ARAP-like deformation of images with explicit line constraints and (2) a flexible tangent plane determination approach for the gnomonic projection of panoramas. The remainder of this paper is organized as follows. Section II presents the details of the proposed method. Section III demonstrates the experimental evaluations, and conclusions are given in Section IV.

II. METHODOLOGY

To improve the flexibility of the method for use with existing 3D modeling solutions [13], we confine the method to the minimum amount of auxiliary information that needs to be collected or distributed. Therefore, this study does not address the problems of the systematic processing, calibration and integration of mobile mapping systems, and only considers façade modeling from community street-view panoramas such as Google Street View [7]. An overview of the proposed methods, consisting of the ARAP deformation and gnomonic projection, is shown in Figure 2.

A. As-rigid-as-possible deformation for the straightening of bent structures

Inspired by the ARAP deformation [14] of surface mesh, we aim to unbend façade structures while keeping the adjacent pixels as rigid as possible. Therefore, we cover the original image with a gridded mesh [15] $\mathcal{M} = \{\mathcal{V}, \mathcal{F}\}$ where $\mathcal{V} = \{v_i(x_i, y_i)\}$ are the vertices and $\mathcal{F} = \{f_i\}$ the quadrangle faces. In general, 40×20 grids are placed on each panorama. To repair bent structures, line segments $\mathcal{L} = \{l_i(o_i, e_i)\}$ are detected using the popular line segment detector (LSD) [17], where o and e are the starting and



Fig. 3: Illustration of a line constrained ARAP deformation.

end point of a segment, respectively. We then look for new positions of all of the vertices, including v' , o' and e' , under the energies defined below. The deformation of the image is cast as a series of similarity transformations of the mesh vertices. The final image is then achieved by conducting corresponding perspective transformations for each quadrangle using homographic transformation [18]. An illustration of a line constrained ARAP-deformation is shown in Figure 3.

a) *Mesh rigidity energy $E_r(v')$* : The first term in the optimization is the mesh rigidity, taken from [15], which imposes a constraint on the quadrangle to preserve the original shape under a certain similarity transformation SE(2), e.g., *scale*, *translation* and *rotation* $se(v) = s(\begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}) v + (\begin{pmatrix} t_x \\ t_y \end{pmatrix})$. Other deformations that lead to a non-square shape are penalized. By substituting the four parameters with the least-squares solver [15], E_r can be written concisely as,

$$E_r(v') = \frac{1}{|\mathcal{F}|} \sum_f \| (A_p(A_p^T A_p)^{-1} A_p^T - I) V'_f \|^2, \quad (1)$$

where $A_p \in \mathbb{R}^{8 \times 4}$ and $V'_f \in \mathbb{R}^8$ comprise the original and deformed coordinates of the four vertices enclosed by face f , i.e., v and v' , as in the following equation [15],

$$A_p = \begin{bmatrix} x_1 & -y_1 & 1 & 0 \\ y_1 & x_1 & 0 & 1 \\ \vdots & \vdots & \vdots & \vdots \\ x_4 & -y_4 & 1 & 0 \\ y_4 & x_4 & 0 & 1 \end{bmatrix}, V'_f = \begin{bmatrix} x'_1 \\ y'_1 \\ \vdots \\ x'_4 \\ y'_4 \end{bmatrix}. \quad (2)$$

b) *Vertical and parallel line rigidity energy $E_v(o', e')$ and $E_p(o', e')$* : For the vertical rigidity constraints $E_v(o', e')$, a subset of the detected or traced line segments $\mathcal{L}_v = \{l(o, e)\}$ is selected interactively and placed into new positions. To keep the joint parts changing smoothly, we also optionally allow several endpoints $\mathcal{V}_e = \{e_l\}$ that are kept fixed as boundary constraints. Specifically, for each vertical line segment l , we apply a transformation with scale and rotation to the line and constrain it to be rigid with regard to the original line, because the rotation angle θ_l is directly computed from the misalignment to the vertical direction. The only additional unknown parameter is the scale s . Similar to Equation 1, by substituting the scaled line segments with $\tau = e - o$ and using $R_l = \begin{pmatrix} \cos \theta_l & -\sin \theta_l \\ \sin \theta_l & \cos \theta_l \end{pmatrix}$ to represent the known rotation matrix, we have a simplified rigidity constraint as in Equation 3,

$$E_v(o', e') = \frac{1}{|\mathcal{L}_v|} \sum_l \| (R_l \tau (\tau^T \tau)^{-1} \tau^T R_l^T - I) \tau' \|^2. \quad (3)$$

For the parallel lines induced by façade attachments, the target orientation is set to the average of all of the parallel lines rather than the vertical direction. We do not use a fixed number

of bins [16], [19], because a horizontal parallel segment may have different orientations in a panorama. The parallel rigidity energy E_p is defined in the same way as E_v , except that the matrix R is computed from the target orientation.

c) *Line coherent energy*, $E_c(\mathbf{v}', \mathbf{o}', \mathbf{e}')$: The line segments $l(\mathbf{o}', \mathbf{v}')$ and the mesh grids \mathbf{v}' are still separated. Inspired by the rigid constraints of mesh grids that aim to preserve the shape of a quadrangle, we also keep the relative position of an endpoint fixed with respect to both the original and deformed grids. Specifically, the original endpoints \mathbf{o} and \mathbf{e} can be expressed by a bi-linear interpolation of the vertices of the grid, $\mathbf{o} = a_{11}\mathbf{v}_1 + a_{12}\mathbf{v}_2 + a_{21}\mathbf{v}_3 + a_{22}\mathbf{v}_4 = A\mathbf{V}$. The coefficient vector A is used for the deformed segments in $E_c(\mathbf{v}', \mathbf{o}', \mathbf{e}')$ by

$$E_c(\mathbf{v}', \mathbf{o}', \mathbf{e}') = \frac{1}{2|\mathcal{L}|} \sum_{o,e} \|(\mathbf{o}' - A_o \mathbf{V}_o') + (\mathbf{e}' - A_e \mathbf{V}_e')\|^2. \quad (4)$$

d) *Optimization*: The total energy is a weighted summation of all of the energies in Equations 1, 3 and 4, including the mesh rigidity E_r , vertical rigidity E_v , parallel rigidity E_p and coherent E_c . Each term is normalized to the unit variance by applying *a priori* standard deviations adding it to the total energy as,

$$\min_{\mathbf{v}', \mathbf{o}', \mathbf{e}'} \lambda_r E_r + \lambda_v E_v + \lambda_p E_p + \lambda_c E_c, \quad (5)$$

where λ s are the weights. The above optimization is the standard least-squares problem and is solved using the Ceres Solver from Google Inc. [20]. The fixed endpoints in \mathcal{V}_e are explicitly constrained by setting the parameter blocks to be constant, which is supported by the Ceres Solver. For the parameters λ , we follow previous work [16] and award larger weights to the line structures. $\lambda_r = 0.1, \lambda_v = 0.4, \lambda_p = 0.4, \lambda_c = 0.1$ are used in all of the experiments that follow.

B. Gnomonic projection for panoramic images

Generally, a panorama is obtained by mapping multiple images acquired at the same time to a unit sphere and then unwrapping them to a 2D image plane by equirectangular projection. Although the original panoramic image maintains straight structures after the ARAP deformation constrained by line segments, the images are still distorted due to the projection problem, especially for objects that are distant from the central horizontal line. This study proposes a piece-wise gnomonic projection method to correct the distortion.

For the equirectangular projection, a point in the object space is first projected onto a sphere with a radius $R = W/2\pi$, and then the corresponding azimuth θ and zenith φ angles are directly used to unwrap the sphere onto a raster image with a width of $W = 2\pi R$ and height of $H = \pi R$. Therefore, the pixel coordinates (x, y) of the panoramas have explicitly one-on-one correspondences with the spherical coordinates (θ, φ) .

Rather than directly using the unscaled coordinates (θ, φ) for the raster unwrapping, gnomonic projection maps the sphere onto a tangential plane. The rationale behind adopting this approach for the rectification of the façade area is that, as long as the selected tangential plane is approximately parallel



Fig. 4: Selection of the tangent points. (a) Tangent points are located directly on the inflection points of the road marks and sidewalks and (b) when occluded, they may also be found from the inflection on the curves of vehicles and façades.

to the façade, the distortion can be reduced. Therefore, the key to a satisfactory projection is to find the best tangent points $T(\theta_0, \varphi_0)$. The projection is then determined as,

$$\begin{aligned} y &= R \cot \varphi \\ x &= R \tan(\theta - \theta_0). \end{aligned} \quad (6)$$

Although no auxiliary navigation or mapping information is available, the determination of the best tangent points is quite obvious in the original panoramas due to the presence of inflection points of the boundaries of road features (Figure 4a), such as marks and sidewalks. Even if these features are occluded, the tangent point can be approximated from building façades (Figure 4b). We also assume that the sensor is leveled and the tangent points lie in the center of a panorama, i.e., $\varphi_0 = 90^\circ$; therefore, only the horizontal coordinates of the tangent point are selected. If the moving direction of the vehicle that collects the panorama is parallel to the road, the two tangent points should have an interval of 180° in the spherical coordinates (or half of the width of the image). However, we do not adopt this assumption and always select two or more tangent points in special scenarios, such as crossroads.

Another practical issue for the gnomonic projection is the choice of receptive area on the tangent plane, or the horizontal and vertical field of view. Because the areas exceeding $|\varphi - \varphi_0| = 90^\circ$ or $|\theta - \theta_0| = 90^\circ$ project to a point at an infinite distance, the ranges should be chosen properly. To prevent unnecessary up-sampling and avoid ground areas, the range of φ and $\theta - \theta_0$ must be limited. For instance, if $\varphi \in [15^\circ, 105^\circ]$ and $(\theta - \theta_0) \in [-60^\circ, 60^\circ]$ are selected and the distance from the sensor to the façade is 10 m, the horizontal and vertical receptive areas cover about 35 m and 37 m, respectively. A practical way of extending the horizontal areas is by mosaicking adjacent images.

III. EXPERIMENTAL EVALUATIONS

A. Results for image deformation and rectification

Four scenarios from both metropolis and old-style cities (Hong Kong, Toronto, London and Paris) are considered. The purpose of the image deformation is to unbend the façade structures. Figure 5 shows the results of the image reparation. The two enlarged areas in the right column of each subfigure compare the panoramas before and after the ARAP-like deformation. The bottom row of each subfigure denotes the results after rectification, and can be used directly as a

reference image for 3D modeling. The differences between the original and deformed panoramas are quite subtle. This is to be expected because two out of a total of four energies in the ARAP-like deformation (Equation 5) are trying to keep their original shapes. However, in the enlarged yellow regions, the oblique and curved façade structures have been straightened satisfactorily, with axis-aligned lines serving as references for better interpretation.



Fig. 5: Reparation of the four datasets. Areas enclosed by the yellow and cyan rectangles are enlarged on the right column and bottom row, respectively.

However, it is still difficult to use the deformed images in Figure 5 for 3D modeling, and rectification using the gnomonic projection with proper tangent points and fields of view are necessary. The enlarged cyan regions in Figure 5 demonstrate the effects of the rectification of distortion and correction of the orientation of the façades. After rectification, the sliced images from different panoramas can be mosaicked using the same scale settings. This can be achieved by most off-the-shelf 3D modeling solutions.

B. Results of 3D modeling

The repaired images are directly imported into [13] as reference images for 3D façade modeling. The textured models are shown in Figure 6. Even when using a single panorama, it is possible to create photorealistic and semantic-aware façade models. The LoD-3 attachments, such as windows and doors, are reconstructed satisfactorily [1]. The environmental light reflections on the windows as shown in the right column of Figure 6 are essential for many applications requiring realistic rendering, and can only be achieved with semantic-aware LoD-3 models. More 3D modeling results are given in the supplementary materials.

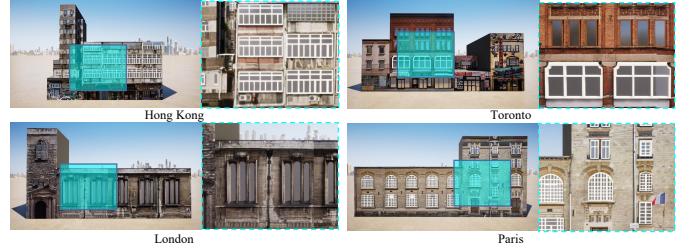


Fig. 6: Comparison with and without the panoramas as the background image for sketching.

To further evaluate the metric capability of the gnomonic projection, we also use the photogrammetric mesh models from aerial oblique images [5], which are available on Google Earth, as a 3D reference. The images are scaled by the height of the buildings as measured from Google Earth. The photogrammetric meshes and façade models are registered in the local coordinate system. The center positions and diagonal lengths of the windows of a five-story building are compared. The results are shown in Table I. The root mean square error (RMSE) for the diagonal lengths and center offsets is 0.15 m and 0.04 m, respectively, which indicates a satisfactory metric accuracy. This good alignment is expected, because the rectification has no distortion as long as all of the objects are coplanar, which is generally tenable for windows on the same façade.

TABLE I: Accuracies of the diagonal length (shaded cells) and center offset of the windows of a five story building, with differences measured in meters.

	Column #1	Column #2	Column #3		
Floor #1	0.02	0.02	0.01	0.02	0.03
Floor #2	0.09	-0.01	0.13	0.08	0.09
Floor #3	0.18	-0.04	0.21	0.06	0.19
Floor #4	0.21	-0.04	0.19	0.05	0.2
Floor #5	0.18	0.02	0.14	0.07	0.15
					-0.02

3D modeling in this paper is an interactive procedure that involves sketching directly onto a reference image. To further evaluate the effectiveness of the proposed strategy, we also compare an alternative strategy that uses the original panorama shown in the web browser as the reference. The Hong Kong and London datasets are used in this comparison. Figure 7 shows a comparison of the runtime and number of operations for each dataset. We also count the time taken for the image reparation. The proposed method shows an improved efficiency of interactions of about 30% to 50%, and also a reduced number of operations, at about 10% to 55%. The proposed method is also metric aware, on the contrary, the operator can only empirically determine the size of the objects from the web browser.

C. Discussion and Limitations

Based on the experimental evaluations, we discuss some properties and limitations of the proposed methods.

Flexibility: The proposed method only requires a single panorama, obtained from community distributors; auxiliary

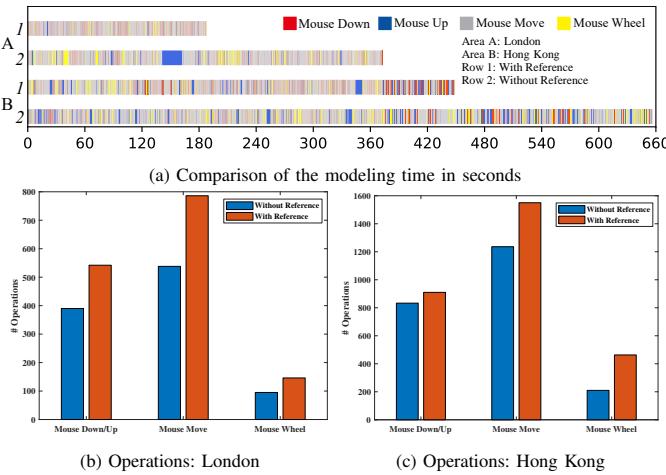


Fig. 7: Comparison with and without the panorama as the background image for sketching.

information may be helpful, but come at the cost of the complexities in the system design, which lead to problems in the consumption by off-the-shelf 3D modeling solutions. In order to fulfil this requirement, the panorama has to be repaired with only image-space processing techniques, e.g. the line constrained ARAP deformation and the gnomonic projection with suitable tangent points.

Limitations: Because the coordinates are defined in the tangent space, as seen in Equation 6, the height of the reprojected image quickly approaches infinity when the zenith angle is around zero. This issue limits the use of the proposed method for high-rise buildings in metropolitan cities. The non-compensated leveling problem and geometric warping in the image mosaicking also lead to systematic distortions. In the Paris dataset, the building is located on a gentle slope, and therefore the roof is also slanted. Further projective or affine transformations could be used to correct such distortions.

IV. CONCLUSIONS

To fully exploit street-view panoramas for LoD-3 reconstruction, this study proposes an image-space deformation and projection approach to repair distorted panoramas. To make the approach more flexible and easier to use with existing 3D modeling solutions, no auxiliary information is required. The underlying solutions essentially consist of (1) a line constrained ARAP deformation that corrects bent structures caused by intrinsic sensor defects and extrinsic orientation inaccuracies, and (2) the rectification of distortion using gnomonic projection by proper tangent point estimation. Experimental evaluations show that the proposed image-space correction methods balance the two competing desires of wanting to keep the original images and fix the curved structures. The metric accuracy is also acceptable, even when the scale is only measured from a known distance. However, this method still has some limitations for the high-rise buildings. Future research directions include the automatic detection of façade objects using deep learning and the alignment (or snapping) of structured objects in a consistent manner [11]. Sensor

calibration is also a viable approach [21], which may be investigated in the future.

REFERENCES

- [1] T. H. Kolbe, G. Gröger, and L. Plümer, “Citygml: Interoperable access to 3d city models,” in *Geo-information for disaster management*. Springer, 2005, pp. 883–899.
- [2] D. Duarte, F. Nex, N. Kerle, and G. Vosselman, “Damage detection on building façades using multi-temporal aerial oblique imagery,” *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2019.
- [3] T. Kelly, P. Guerrero, A. Steed, P. Wonka, and N. J. Mitra, “Frankengan: Guided detail synthesis for building mass-models using style-synchronized gans,” *arXiv preprint arXiv:1806.07179*, 2018.
- [4] P. Gurram, E. Saber, and H. Rhody, “A segment-based mesh design for building parallel-perspective stereo mosaics,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 3, pp. 1256–1269, 2009.
- [5] H. Hu, Q. Zhu, Z. Du, Y. Zhang, and Y. Ding, “Reliable spatial relationship constrained feature point matching of oblique aerial images,” *Photogrammetric Engineering & Remote Sensing*, vol. 81, no. 1, pp. 49–58, 2015.
- [6] F. Li, M. Lehtomäki, S. O. Elberink, G. Vosselman, A. Kukko, E. Puttonen, Y. Chen, and J. Hyppä, “Semantic segmentation of road furniture in mobile laser scanning data,” *ISPRS journal of photogrammetry and remote sensing*, vol. 154, pp. 98–113, 2019.
- [7] D. Anguelov, C. Dulong, D. Filip, C. Frueh, S. Lafon, R. Lyon, A. Ogale, L. Vincent, and J. Weaver, “Google street view: Capturing the world at street level,” *Computer*, vol. 43, no. 6, pp. 32–38, 2010.
- [8] S. Verstockt, M. Gerke, and N. Kerle, “Geolocalization of crowdsourced images for 3-d modeling of city points of interest,” *IEEE geoscience and remote sensing letters*, vol. 12, no. 8, pp. 1670–1674, 2015.
- [9] T. K. Dang, M. Worring, and T. D. Bui, “A semi-interactive panorama based 3d reconstruction framework for indoor scenes,” *Computer vision and image understanding*, vol. 115, no. 11, pp. 1516–1524, 2011.
- [10] Y. Li, Q. Hu, M. Wu, J. Liu, and X. Wu, “Extraction and simplification of building façade pieces from mobile laser scanner point clouds for 3d street view services,” *ISPRS International Journal of Geo-Information*, vol. 5, no. 12, p. 231, 2016.
- [11] S. Hensel, S. Goebels, and M. Kada, “Facade reconstruction for textured lod2 citygml models based on deep learning and mixed integer linear programming,” *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, pp. 37–44, 2019.
- [12] Y. Jia, G. Cui, L. Kong, and X. Yang, “Multichannel and multiview imaging approach to building layout determination of through-wall radar,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 5, pp. 970–974, 2013.
- [13] SketchUp, “Sketchup: 3d design software — 3d modeling on the web,” <https://www.sketchup.com/>, 2019, accessed: August 6th.
- [14] O. Sorkine and M. Alexa, “As-rigid-as-possible surface modeling,” in *Symposium on Geometry processing*, vol. 4, 2007, pp. 109–116.
- [15] G.-X. Zhang, M.-M. Cheng, S.-M. Hu, and R. R. Martin, “A shape-preserving approach to image resizing,” in *Computer Graphics Forum*, vol. 28, no. 7. Wiley Online Library, 2009, pp. 1897–1906.
- [16] K. He, H. Chang, and J. Sun, “Rectangling panoramic images via warping,” *ACM Transactions on Graphics (TOG)*, vol. 32, no. 4, p. 79, 2013.
- [17] R. G. Von Gioi, J. Jakubowicz, J.-M. Morel, and G. Randall, “Lsd: a line segment detector,” *Image Processing On Line*, vol. 2, pp. 35–55, 2012.
- [18] R. Hartley and A. Zisserman, *Multiple view geometry in computer vision*. Cambridge university press, 2003.
- [19] C.-H. Chang, Y. Sato, and Y.-Y. Chuang, “Shape-preserving half-projective warps for image stitching,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 3254–3261.
- [20] S. Agarwal, K. Mierle *et al.*, “Ceres solver,” ceres-solver.org, 2012, accessed: August 6th.
- [21] R. Ishikawa, T. Oishi, and K. Ikeuchi, “Lidar and camera calibration using motions estimated by sensor fusion odometry,” in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 7342–7349.