Proceedings of the Fifth IIEEJ International Workshop
on Image Electronics and Visual Computing 2017
Da Nang, Vietnam, February 28 - March 3, 2017

NON-PLANAR MARKER MATCHING WITH PARTIAL AFFINE FOR AR

Tao Iwasaki† Hiroki Takahashi†
†Graduate School of Informatics and Engineering,
The University of Electro-Communications, Japan

ABSTRACT

This paper proposes a method to deal with locally different affine parameters on a non-planar marker. The method assumes SIFT points are uniformly distorted by affine parameters. Marker detection is the most important process for vision-based AR. It is necessary for the non-planar marker detection to compare obtained feature points between an original planar image and a target non-planar image. Since the target image has a variety of partial distortions and occlusions, it becomes a difficult process. ASIFT is an affine invariant SIFT. It, however, deals with a unique parameter for a whole image. This paper, therefore, proposes a method that enables a marker with a variety of distortions to be detected precisely. The planar marker is split into small rectangular meshes. Each affine transformed mesh is compared with the non-planar marker. The proposed method has the same computational complexity to ASIFT in feature points matching. The almost same numbers of matched feature points for a planar, multifaceted and curved marker are obtained by the proposed method.

1. INTRODUCTION

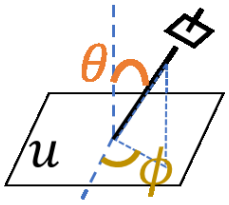
AR (Augmented Reality) system attracts attention as an effective method for information presentation in various fields such as entertainment, advertising and industry. Many kinds of AR applications are proposed [1, 2]. AR applications with arbitrary distorted markers are compatible with intuitive experience. Especially, various vision-based AR system application has been developed since it requires only a marker and a camera. In typical AR system using a marker, the marker is detected from a camera image, and estimated for its position and posture. Then a synthesized image is overlaid. Since the AR system is easy to handle, recent researches are widely using vision-based AR system. Some recent researches have proposed marker detection methods for designed markers, partially occluded markers [3] and curved artificial markers[4]. A few of them, however, target a deformable marker since that is a challenging problem. Many locally different projective transformations are performed on a deformed marker when the marker is projected onto a camera image plane. It is, therefore, difficult to apply traditional marker detection algorithms to the deformed maker directly. A. Lee et al. [3] proposed realtime tracking plane with feature points. They concluded combining SIFT (Scale Invariant Feature Transform) [5] with their feature tracking method for more robust feature tracking. Although SIFT is invariant feature for scale, rotation and intensity change, it is not invariant for more complex transformation such as affine and projective transformations. ASIFT (Affine-SIFT) [6] has been proposed in order to deal with affine transformation for SIFT. ASIFT samples affine space by camera optical axis longitude and latitude based on geometrical affine camera model. Fig.1(a) illustrates geometrical affine camera model which is used for the affine parameters sampling. The u denotes marker image plane. Angle θ and ϕ represent camera optical axis for longitude and latitude, respectively. It generates affine transformed images from sampled affine parameters and calculates SIFT from all transformed images. The matching examined made between each of the affine distorted markers images and an input image by using SIFT feature points. Furthermore, the maximum number of matching points of a sampling affine transformed image is selected as ASIFT correspondence. In summary, ASIFT practically can obtain affine invariant SIFT features. Moreover, ASIFT allows to match different view image of non-planar shape marker. It is, however, hard for ASIFT to match a planar template maker with a non-planar marker. ASIFT cannot obtain sufficient correspondence points for non-planar marker detection, even though it compares the non-planar marker image with a planar template marker. Thus, this paper proposes a method that matches split marker images with locally different affine parameters to a target marker.

2. KEYPOINTS MATCHING ALGORITHM

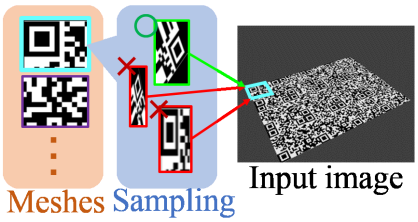
This paper uses affine space sampling on Affine camera model and matching methods based on ASIFT[6]. The cameramodel is given by Eq.(1) and affine transromation depends on the (θ, ϕ).

$$A = \begin{bmatrix} \cos(\psi) & -\sin(\psi) \\ \sin(\psi) & \cos(\psi) \end{bmatrix} \begin{bmatrix} \cos(\theta) & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\phi) & -\sin(\phi) \\ \sin(\phi) & \cos(\phi) \end{bmatrix} \quad (1)$$

The ψ denotes camera rotation with respect to the optical axis of the camera passing through the center of the image plane. In this paper, the ψ is not considered because



(a): Rotational transfer model



(b): Matching method overview

Figure 1: Model and mesh matching method

our method applies the SIFT. The θ is maked by camera optical axis with normal to the image plane u. The angle is represented $\theta = \arccos(1/t)$ by tilt t. The tilt has a one-to-one relation to the θ. The tilt and θ entail a strong image distortion. In addition, the tilt t is changed according to $t_{k+1} = \Delta t t_k$, where $t = [1, 4\sqrt{2}]$. The sampling step Δt is experimentally assigned $\sqrt{2}$. Furthermore, ϕ depends on $\Delta\phi$. The sampling step $\Delta\phi$ is also experimentally decided as $\Delta\phi = 72°/t$. ϕ is changed according to $\phi_{k+1} = \Delta\phi + \phi_k$, where $\phi = [0°, 180°)$.

A non-planar marker image is regarded as a combination of partially distorted images with different affine parameters. This subsection proposes a method to deal with locally different affine parameter on each part of non-planar marker image. The proposed method divides a simple plane marker image into $N = P \times Q$ grid mesh. Then, each of the mesh is distorted with sampled affine parameter. The matching is examined conducted between each mesh region of sampled affine distorted marker and an input image by using SIFT feature points. Each maximum correspondence points is obtained from each mesh and total correspondence points are given by the maximum correspondences for all meshes. Fig.1(b) shows overview of our proposed method. The left images represent N grid meshes. The images in the middle illustrate sampled distorted meshes with different affine parameters and the right one shows an input image. A green line indicates an ASIFT correspondence and red lines show miss-matching pairs. Assuming that the numbers of affine samplings are N_s, SIFT key points are detected for all N_s affine distorted images. n_j $(j = 1, \ldots, N_s)$ key points are extracted from an affine distorted image A_j. Moreover n key points are extracted from the input image. Since nn_j times comparisons are necessary for one distorted marker, it becomes $n\sum_{j=1}^{N_s} n_j$ comparisons in total. In contrast, dividing into grid mesh, SIFT key points are detected for all $N_s N$ affine distorted images. Besides, $n_{i,j}$ $(i = 1, \ldots, N)$ is a number of SIFT points from a mesh $M_{i,j}$ of A_j. Notice that the relation between the number of SIFT points n_j in an affine image A_j and all of N meshes is given by Eq.(2).

$$n_j = \sum_{i=1}^{N} n_{i,j} \tag{2}$$

Thus, a number of comparisons for maximum correspondences by our method becomes almost $n\sum_{j=1}^{N_s} n_j$ and the number is nealy equal to ASIFT comparison method,

All distorted meshes with sampled affine parameters are matched with an input image. This paper examines a full search algorithm based on Euclidean distance. A keypont measure employs a ratio of distance $r = closest/nextclosest$ [5]. The best matched distorted i-th mesh in all $M_{i,j}$ means that the number of matching pairs is maximum and the ratio of distance is minimum in the distorted meshes $M_{i,j}$.

3. EXPERIMENTAL RESULTS

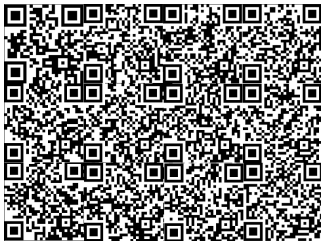
This section evaluates the performance of the proposed detection algorithm. A performance is compared between the proposed method and ASIFT in simulation environment.

3.1. EXPERIMENTAL DATA

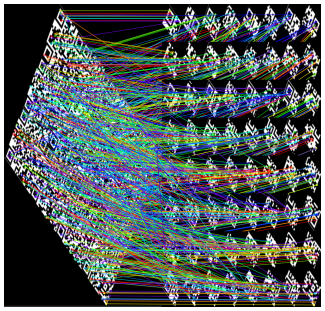
In this paper, each experiment uses a few a kinds of markers as shown in Fig.2(a) and Fig.3. To evaluate the performance clearly, the marker image shown in Fig.2(a) is employed. The QR code marker in Fig.2(a) includes 12 QR codes which are different with each other without margin. The QR marker is used to get SIFT feature on each mesh evenly. Three deformed markers are used as shown in Fig.3 with in CG simulation environment. Additionally, in the simulation environment, each image is captured by a camera that is rotational transformed depending on affine sampling algorithm based on Fig.1(a). This paper examines the correspondence of each method with sampling number $N_s = 44$. The image size is 800×600 pixel and the grid mesh N is $8 \times 8 = 64$. Moreover, a threshhold ratio of distance of SIFT keypoint matching is set to 3.5.

3.2. MATCHING PERFORMANCE EVALUATION

This section discribes Exp.2 matching result for Fig.2(a). The proposed method is evaluated for three kinds of surface shapes, i.e. planar, multifaceted and curved surfaces as shown in Fig.3. Furthermore, the proposed method is also compared with ASIFT. First, a matching result of planar

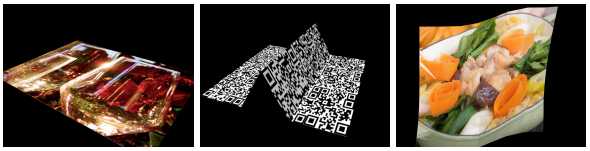


(a): QR code marker



(b): A matching result

Figure 2: Experimental marker and matching result

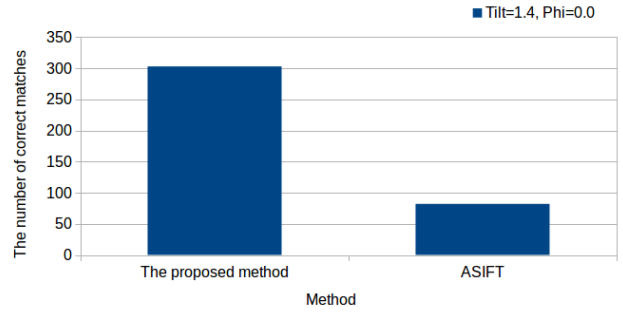


(a): Planar **(b):** Multifaceted **(c):** Cuved

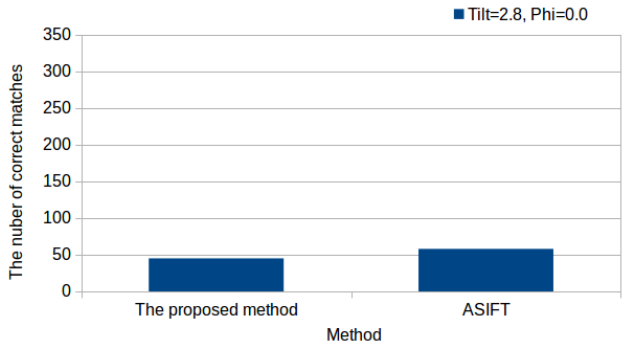
Figure 3: Experimental simulation images

marker is shown in Fig.4. Each horizontal axis in the graph shows total matched and correct matches of the proposed method and ASIFT. Each vertical axis in the graph shows the number of matching keypoints and correct matches. The blue bar shows the proposed method and the orange bar shows ASIFT. The correct matches is obtained from total matches that is below the matching distance threshold. In ths paper, the threshold is 100. The proposed method obtains larger number of total matched keypoints and correct one than ASIFT whrere $t = 1.4$ and $\phi = 0.0$ as shown Fig.4(a). The proposed method, however, decrease the number of them whrere $t = 2.8$ and $\phi = 0.0$ as shown Fig.4(b). The number of correct matched keypoints, additionally, less than ASIFT. The proposed method, hence, is hard to match keypoints with planar marker, where the camera view larger tilt t.

Fig.5 shows each matching result of method where camera view parameters are $t = 1.4$ and $\phi = 0.0°$. The proposed method match with larger region on perspective trans-



(a): Num of matches, where camera view $t = 1.4, \phi = 0.0°$



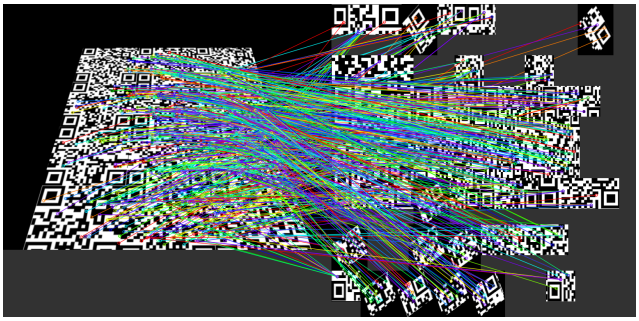
(b): Num of matches, where camera view $t = 2.8, \phi = 0.0°$

Figure 4: Comparison results for planar surface

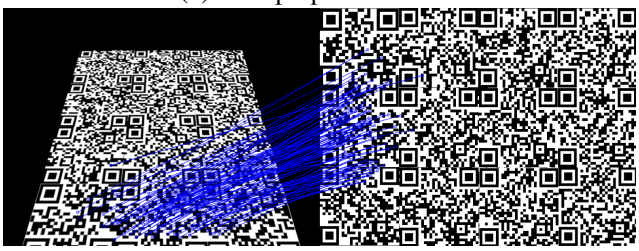
formed marker than ASIFT in this case as shown Fig.5(a). The proposed method, however, is hard to deal with strong perspective trasformation. Fig.6 shows each matching result of method where camera view parameters are $t = 2.8$ and $\phi = 0.0$. In this case, the proposed method and ASIFT obtain fewer number of correct matches than Fig.**??** Fig.**??**

Fig.7 shows matching result of ASIFT at sampling parameters $t = 5.657$ and $\phi = 12.728°$. The result obtained 15 correct matches. The proposed method, however, obtained no matches. Due to mesh size is small, each mesh is similar each others. Moreover, our method is hard to deal with large number of tilt. It is hard to match because meshes are divided into small rectangles. SIFT keypoints around long edge cannot be detected because the edge is divided into several meshes. Therefore, our method is hard to detect planar marker with large number of tilt.

Second, a matching result for multifaceted marker is shown in Fig.8 and Fig.9. Fig.8 shows the number of correct matches of the proposed method and ASIFT. The proposed method similar equal to the number of correct matches of ASIFT. Fig.9(a) shows matching lines of the proposed method. The proposed method matched with a partial region on constraction plane of multifaceted marker in the fromt and the other region one. ASIFT, on the other hand, only one patial region on constraction plane of multifaceted

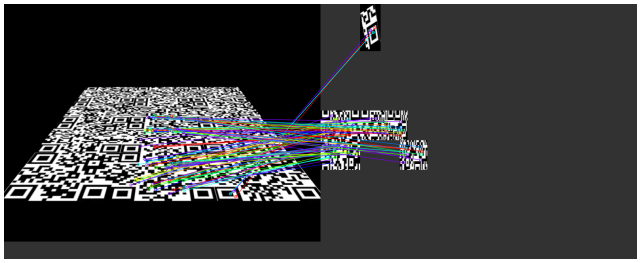


(a): The proposed method

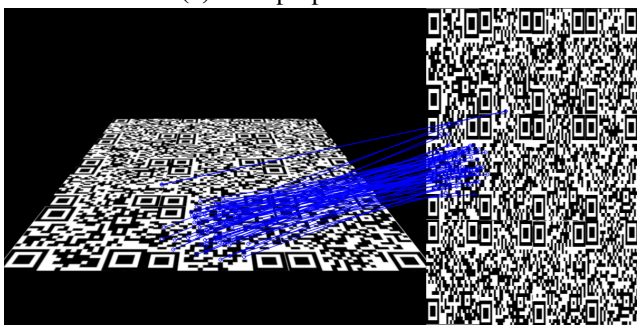


(b): ASIFT

Figure 5: Matching results at $t = 1.40$ and $\phi = 0.0°$



(a): The proposed method



(b): ASIFT

Figure 6: Matching results at $t = 2.80$ and $\phi = 0.0°$

marker in the front as shown Fig.9(b).

Finally, a matching result of curve distorted marker, where camera view parameters are tilt $t = 1.4$ and $\phi = 152°$, is shown in Fig.10 and Fig.11. The proposed method obtains greater the number of correct matches than ASIFT. ASIFT matched only ceartain parts that seem to be undergone almost similar transformations because the curved marker has individually different transformation parameter meshes. Fig.11 shows results of matched keypoints pairs for the curved marker from camera view parameters are tilt $t = 1.4$ and $\phi = 152°$. The proposed method, on the other hand, matches with various meshes of the curved marker as shown Fig.11(a).

However, our method obtains many matched keypoints including errors in almost experiments as shown Fig.12. A green rectangle illustrates correct correspondence of mesh and input marker image. A red rectangle illustrates incorrect correspondence. A yellow arrow illustrates error matching. As a possible cause, the proposed method obtains SIFT key points locally features in mesh because the propsed method divides a marker that is resolution 800×600 into 8×8 meshes that is resolution 100×75. In other words, the mesh resolution is too small for SIFT keypoints extraction.

Therefore, the SIFT cannot calculate the necessary scale image pyramid to extract keypoints.

4. CONCLUSION & FUTURE WORKS

This paper proposed non-planar marker matching method based on ASIFT. The computational complexity of the proposed method is equal to that of the ASIFT. The result matching performance of the proposed method suggested that it is valid method for detection non-planar marker like multifaceted shape and curved shape. However, our method is difficult to match preciously with planar marker. Further studies are needed to match key points more preciously.

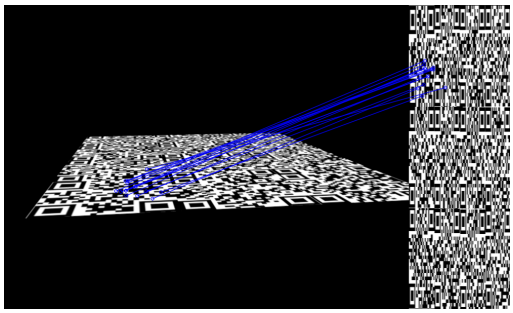


Figure 7: ASIFT matching result

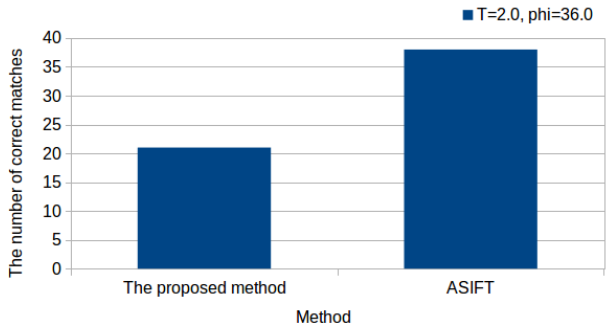


Figure 8: Comparison the number of matches for multifaceted surface

5. REFERENCES

[1] J. Steimle, A. Jordt, and P. Maes, "Flexpad : Highly Flexible Bending Interactions for Projected Handheld Displays," Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp.237–246, 2013.

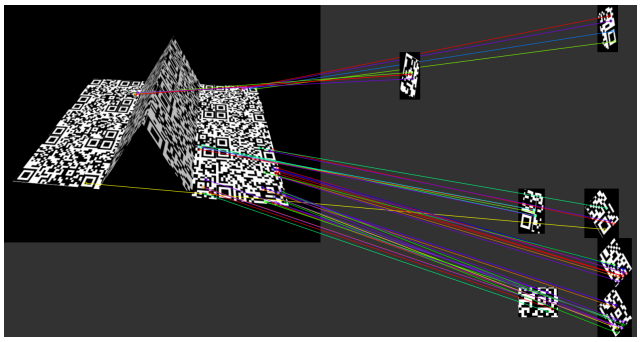
[2] Y. Fujimoto, J. Miyazaki, T. Taketomi, H. Kato, B.H. Thomas, G. Yamamoto, and R.T. Smith, "Geometrically-correct projection-based texture mapping onto a cloth," IEEE Virtual Reality, VR, vol.20, pp.169–170, 2014.

[3] A. Lee, J.-Y. Lee, S.-H. Lee, and J.-S. Choi, "Markerless augmented reality system based on planar object tracking," 17th Korea-Japan Joint Workshop on Frontiers of Computer Vision, FCV, pp.1–4, 2011.

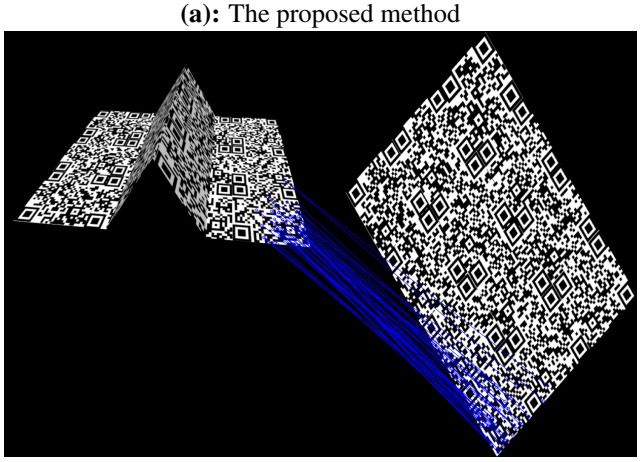
[4] H. Uchiyama and E. Marchand, "Deformable random Dot markers," 10th IEEE International Symposium on Mixed and Augmented Reality, ISMAR, pp.237–238, 2011.

[5] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol.60, no.2, pp.91–110, 2004.

[6] J.-M. Morel and G. Yu, "ASIFT: A New Framework for Fully Affine Invariant Image Comparison," SIAM Journal on Imaging Sciences, vol.2, no.2, pp.438–469, 2009.



(a): The proposed method



(b): ASFT

Figure 9: Matching results at $t = 2.00$ and $\phi = 32.0°$

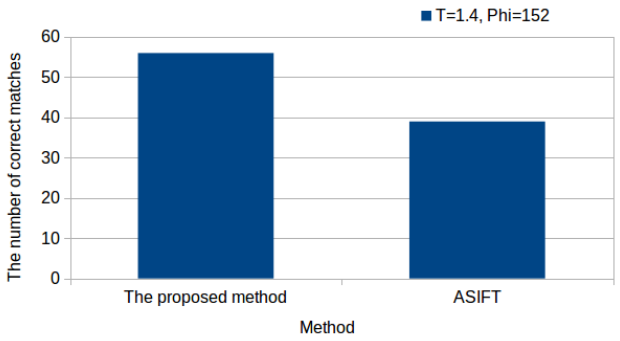
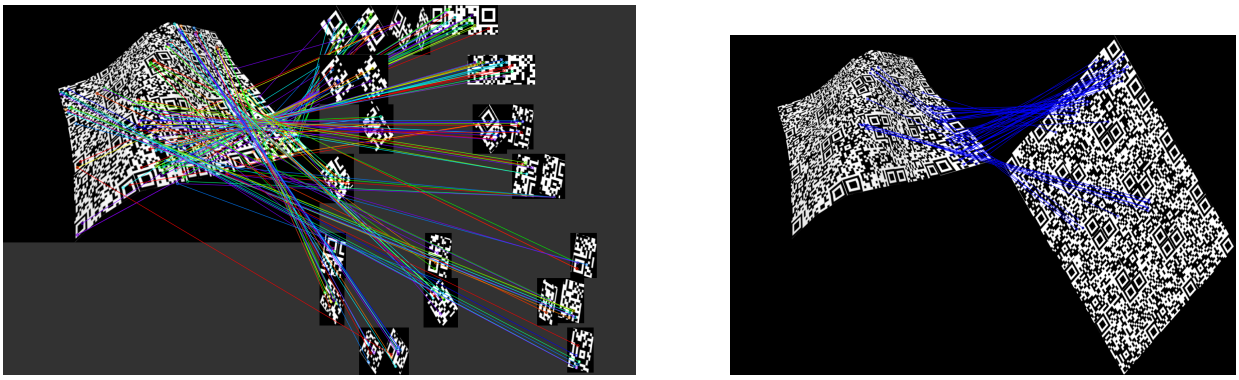


Figure 10: Comparison results for curving surface



(a): The proposed method **(b):** ASIFT

Figure 11: Matching results at $t = 1.41$ and $\phi = 152.0°$

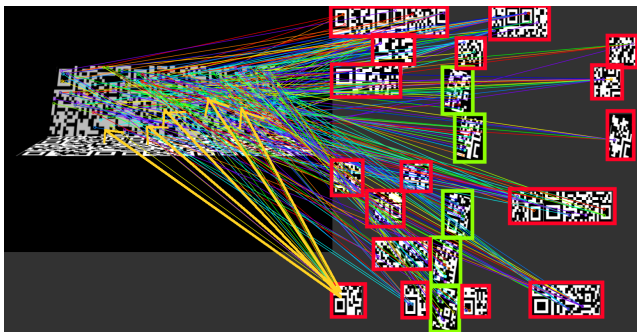


Figure 12: Example of mesh matching