บทความวิจัย

การศึกษาวิธีการเปรียบเทียบวัตถุสามมิติบนภาพถ่ายด้วยแผนที่ UV ด้วยการใช้คุณลักษณะและเครือข่ายนิวรัล

นราวิชญ์ ไมตรีจิตร และ จันทนา จันทราพรชัย*
ภาควิชาวิศวกรรมคอมพิวเตอร์ คณะวิศวกรรมศาสตร์ มหาวิทยาลัยเกษตรศาสตร์

* ผู้นิพนธ์ประสานงาน โทรศัพท์ 08 3010 8294 อีเมล: fengcnc@ku.ac.th DOI: 10.14416/j.kmutnb.2024.09.001
รับเมื่อ 6 เมษายน 2566 แก้ไขเมื่อ 3 กรกฎาคม 2566 ตอบรับเมื่อ 10 สิงหาคม 2566 เผยแพร่ออนไลน์ 2 กันยายน 2567

บทคัดย่อ

งานวิจัยนี้เป็นการออกแบบวิธีการเปรียบเทียบวัตถุสามมิติบนภาพถ่ายด้วยแผนที่ UV โดยการนำแผนที่ UV ในเฟรมที่ต้องการเปรียบเทียบมาเทียบกับแผนที่ UV ที่เป็นตัวอ้างอิงหรือแผนที่ UV ที่ได้จากเฟรมซึ่งมีการตรวจสอบว่าถูกต้องแล้วด้วยมนุษย์ ซึ่งในการเปรียบเทียบวิธีการดั้งเดิมได้ใช้การเปรียบเทียบแบบพิกเซลต่อพิกเซลและการเปรียบเทียบโดยใช้วิธีการจับคู่ลักษณะเด่น แล้วนำผลลัพธ์ที่ได้จากการเปรียบเทียบไปคำนวณหาความคลาดเคลื่อนของตำแหน่งวัตถุสามมิติ เป้าหมายคือ การนำอัลกอริทึมสำหรับการเปรียบเทียบแผนที่ UV ไปต่อยอดกับการติดตามวัตถุสามมิติบนภาพถ่ายเพื่อลดเวลาที่ใช้และเพิ่มความแม่นยำในขั้นตอนการจับการเคลื่อนไหวของตำแหน่งภาพของการทำเทคนิคพิเศษทางภาพ ในงานวิจัยนี้ศึกษาวิธีการเปรียบเทียบแผนที่ UV ซึ่งแบ่งเป็น 2 ประเภทใหญ่ ได้แก่ การใช้คุณลักษณะทางภาพและการใช้เครือข่ายนิวรัล ได้ออกแบบการทดลองทั้งหมดเป็นจำนวน 5 แบบ ได้แก่ 1) การเปรียบเทียบแบบพิกเซลต่อพิกเซลด้วยการลบแบบปกติ (Subtract) 2) การเปรียบเทียบแบบพิกเซลต่อพิกเซลด้วยค่าสัมบูรณ์ของผลต่าง (Absdiff) 3) การเปรียบเทียบโดยใช้วิธีการจับคู่ลักษณะเด่นด้วย SIFT 4) การเปรียบเทียบโดยใช้วิธีการจับคู่ลักษณะเด่นด้วย SIFT กับ Ratio test และ 5) การเปรียบเทียบโดยใช้วิธีการจับคู่ลักษณะเด่นด้วย SuperPoint กับ SuperGlue ซึ่งแต่ละวิธีให้ความแม่นยำในการเปรียบเทียบสำหรับการเปรียบเทียบทั้งแผนที่ UV เท่ากับ 50%, 100%, 33.33%, 16.67% และ 91.67% ตามลำดับ และการเปรียบเทียบเฉพาะด้านที่มีเหมือนกันเท่ากับ 67.5%, 100%, 25%, 8.33% และ 100% ตามลำดับ

คำสำคัญ: เทคนิคพิเศษทางภาพ การประมวลผลภาพ แผนที่ UV การจับคู่ลักษณะเด่น

การอ้างอิงบทความ: นราวิชญ์ ไมตรีจิตร และ จันทนา จันทราพรชัย, “การศึกษาวิธีการเปรียบเทียบวัตถุสามมิติบนภาพถ่ายด้วยแผนที่ UV ด้วยการใช้คุณลักษณะและเครือข่ายนิวรัล,” *วารสารวิชาการพระจอมเกล้าพระนครเหนือ*, ปีที่ 35, ฉบับที่ 2, หน้า 1–12, เลขที่บทความ 252-036860, เม.ย.–มิ.ย. 2568.

Research Article

Comparison Study of 3D Object Matching Techniques for UV Projection Mapping: Image Feature-based and Neural Network-based Approaches

Narawich Maitreejitr and Chantana Chantrapornchai*
Department of Computer Engineering, Faculty of Engineering, Kasetsart University, Bangkok, Thailand

* Corresponding Author, Tel. 08 3010 8294, fengcnc@ku.ac.th DOI: 10.14416/j.kmutnb.2024.09.001
Received 6 April 2023; Revised 3 July 2023; Accepted 10 August 2023; Published online: 2 September 2024

Abstract

This work describes a method for comparing 3D objects on an image with a UV map by comparing the UV map in the desired frame with the reference UV map or the frame-derived UV map verified to be valid by humans. For the comparison method, pixel-by-pixel comparison and feature-matching comparison were performed. Then the results obtained from the comparison were used to calculate the discrepancy of the position of 3D objects. The primary objective is to apply the UV map comparison algorithm to the 3D object matching on the image to reduce the time spent and increase the accuracy of the tracking step in Visual Effect (VFX). In this work, we studied two algorithm styles: Image feature-based and neural network-based approaches. Totally, five UV map comparison methods were observed: 1) pixel-by-pixel comparison with Normal Subtract, 2) pixel-by-pixel comparison with Absolute Value of Difference (Absdiff), 3) feature matching method with SIFT, 4) feature matching method with SIFT and 5) Ratio test, and comparison with feature matching method with SuperPoint and SuperGlue. Each of these methods yielded a comparison accuracy for a comparison of the entire UV map of 50%, 100%, 33.33%, 16.67%, and 91.67%, respectively. The specific side comparison yielded 67.5%, 100%, 25%, 8.33%, and 100%, respectively.

Keywords: Visual Effect, Image Processing, UV Map, Feature Matching

Please cite this article as: N. Maitreejitr and C. Chantrapornchai, "Comparison study of 3D object matching techniques for UV projection mapping: Image feature-based and neural network-based approaches," *The Journal of KMUTNB*, vol. 35, no. 2, pp. 1–12, ID. 252-036860, Apr.–Jun. 2025 (in Thai).

1. Introduction

Visual Effects (VFX) have been with the movie world since the dawn of filmmaking. Visual special effects create something that does not exist on the movie screen, in advertising, or in other moving media, such as water, fire, smoke, dust, and explosions. In the past, images were written directly onto film or superimposed with mirrors and many more. Nowadays, with the advancement of technology, things that are not possible in the conceptual dimension can be done with a computer process.

Visual Effects, which enjoy great popularity nowadays, are special visual effects created by superimposing CG images (Computer Generated or special effects using computer programs) and recorded video images (Live Action). For the images created with visual effects to look harmonious and realistic and not deceiving. It is necessary to calculate the coordinates of the object in the image based on 3D coordinates and camera movement, it is necessary to apply the method of motion detection of the image position (tracking).

The problem with tracking is that it takes a lot of time, especially when tracking three-dimensional objects in the photo (3D object tracking), because it is a process that requires high precision to seamlessly overlay CG images on top of actual video footage (Live Action). Currently, this step is done by humans, including validation, which leads us to find different methods to reduce the time required for this process. The appealing technique is the use of UV maps to support 3D object tracking.

The UV map can tell the surface of a three-dimensional object if the model looks the same in terms of shape, texture color, and each side of the 3D object is unwrapped (UV unwrap) in the same way. The resulting UV map has the same appearance, which is the main idea of a 3D object matching technique using UV projection mapping.

Several works have been done related to 3D object matching. Some uses approximate method some uses machine learning or deep learning to minimize deformation or find the matching.

Least Squares Conformal Maps (LSCM) [1], a technique used in computer graphics to automatically generate texture maps for 3D models. The LSCM algorithm finds a conformal map between the 3D model and a planar domain by minimizing the squared distances between them with a new quasi-conformal parameterization method based on a least-squares approximation of the Cauchy-Riemann equations. This allows the creation of texture maps that can be applied to the surface of the 3D model to create a detailed appearance.

SIFT (Scale-Invariant Feature Transform) [2], a computer vision algorithm used for feature detection and extraction in images. The algorithm identifies and extracting distinctive features from an image that are invariant to scaling, rotation, and illumination changes, making it possible to recognize and identify objects even when they are partially occluded or viewed from different angles. The algorithm consists of several steps, including scale-space extrema detection, keypoint localization, orientation assignment, and feature descriptor generation.

SuperPoint [3], a self-supervised deep learning algorithm for interest point detection and description in images. It learns to detect and describe interest points without the need for explicit human supervision but uses Homographic Adaptation, a multi-scale, multi homography approach for boosting interest point detection repeatability and performing cross-domain adaptation (e.g., synthetic-to-real). The algorithm consists of a detector and a descriptor, and it can handle images with significant changes in scale and orientation. SuperPoint has been shown to achieve state-of-the-art performance on several benchmark datasets for interest point detection and description, with better performance compared to SIFT [2], LIFT [4], and ORB [5].

SuperGlue [6], a state-of-the-art feature matching method in computer vision that uses a Graph Neural Network (GNN) to learn correspondences between pairs of images. Unlike traditional feature matching methods that rely on hand-crafted descriptors, SuperGlue learns a deep neural network that maps local features to a high-dimensional embedding space. Then, it uses the GNN to establish correspondences between features by considering their spatial relationships and similarities in the embedding space. SuperGlue has achieved superior performance on several feature matching benchmark datasets for feature matching, demonstrating the effectiveness of the GNN-based approach.

2. Materials and Methods

2.1 System Design

In this work, we focus on the technique comparison in the aspect of the accuracy. The data for the experiments were setup. First, we generate a UV map from 3D objects and photos (Generate

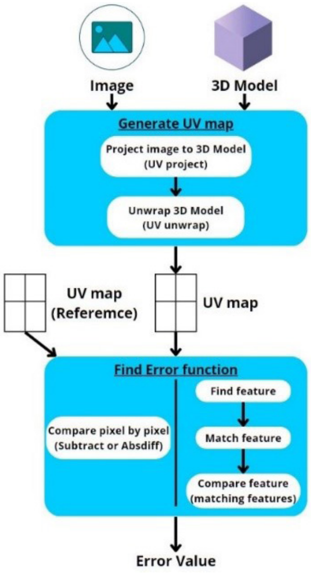

Figure 1 System overview.

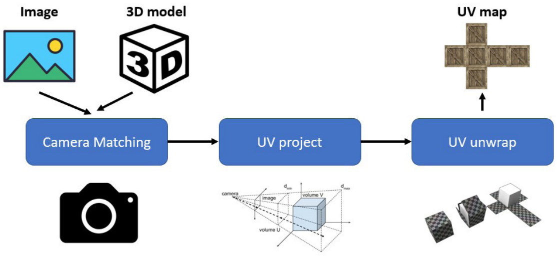

Figure 2 Procedure for generating UV map.

UV map). Then, comparing the UV map to find errors (Find Error function) is done, as shown in Figure 1.

2.1.1 Generate UV map

We prepare the data before importing to the comparison process by converting the 3D objects and photos to a UV map. The procedures are camera matching, UV project and UV unwrap, as shown in Figure 2.

2.1.2 Find Error Function

A comparison is made between the reference UV map (the UV map was verified by humans) and the frame-derived UV map to be compared. The

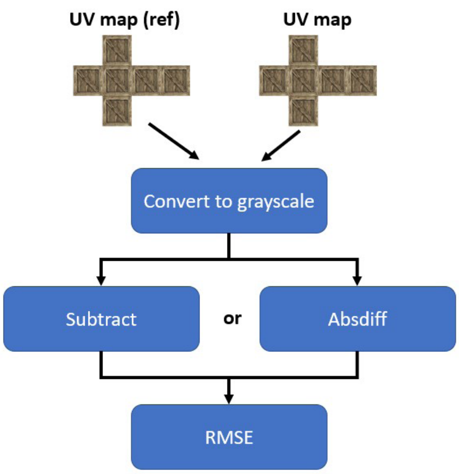

Figure 3 Procedure for comparing UV maps with pixel-by-pixel comparison.

result of the comparison is the $RMSE$ value for n points as Equation (1).

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Error)^2} \quad (1)$$

There are two comparison methods for UV maps: pixel by pixel comparison and feature matching.

The pixel-by-pixel comparison of the UV map comparison is the difference between the color values at the same pixel position of the UV map being compared. Then we take the difference from each pixel to calculate the RMSE value. There are 2 methods to find the difference, namely the normal subtraction (Subtract) and the absolute value of the difference (Absdiff), with the procedure as shown in Figure 3.

The comparison of UV maps using the feature-matching method consists of examining the two input UV maps to find their features, and then finding a pair of features by feature matching

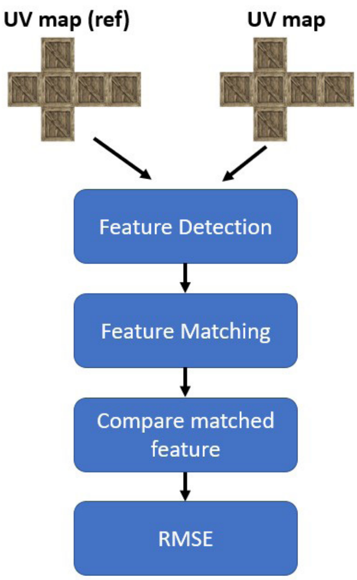

Figure 4 Procedure for comparing UV maps by the feature matching.

between the two maps. Then we use the obtained dominant pairs to find the difference between the positions and determine the RMSE value using the procedure shown in Figure 4. Two algorithms were selected for the feature detection and feature matching phases: SIFT [2] with brute-force matching or brute-force matching with ratio test and SuperPoint [3] with SuperGlue [6] (the procedure is shown in Figure 5 and Figure 6)

Figure 5 shows that the feature matching procedure is different: Brute-force matching results is dominant pairs, while Brute-force matching and ratio testing use KNNMatch, which provides k best matches (using k = 2) for each descriptor and performs a ratio test to remove insufficiently unique features (see example in Figure 7).

there are two parameters to be defined in the “Compare matched feature” step shown in Figure 4, namely N (number of feature pairs) and filter_range

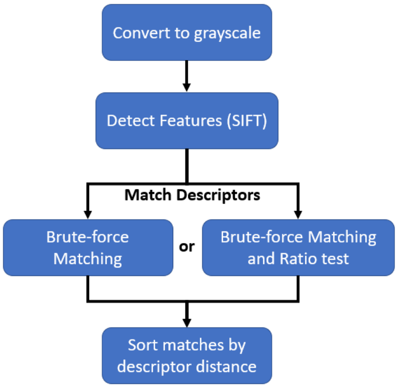

Figure 5 Feature detection and feature matching steps using SIFT.

Figure 6 Feature detection and feature matching steps with SuperPoint and SuperGlue.

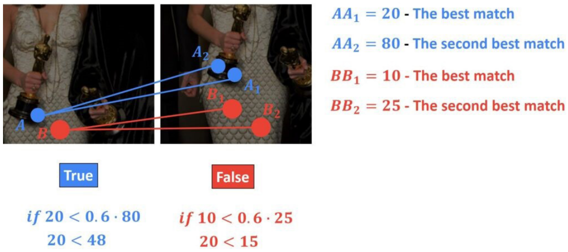

Figure 7 Example of brute-force matching and ratio test when k = 2 and ratio = 0.6.

(Maximum possible value of absolute difference). It is used to filter out feature pairs that are unlikely to match, since the feature pairs obtained from the UV map should have similar positions in both maps.

For comparison, we use the absolute value of the X position difference and the absolute value of the Y position difference in all feature pairs and then take all absolute values to determine the RMSE value.

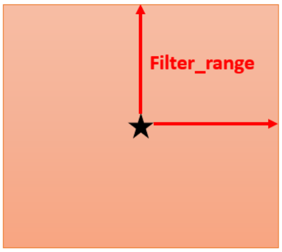

Figure 8 filter_range area.

2.2 Measurement Approaches

To test different ways of comparing UV maps with the same input data (Figure 9–16), a comparison of the entire UV map and a specific side comparison were performed. The UV map comparison model to be tested consisted of

- Pixel by Pixel Comparison with Normal Subtract (Subtract)
- Pixel by Pixel Comparison with Absolute Value of Difference (Absdiff)
- Comparison using the Feature Matching method with SIFT
- Comparison using the Feature Matching method with SIFT (Ratio test: ratio = 0.5, k = 2)
- Comparison using the Feature Matching method with SuperPoint and SuperGlue

When comparing with the feature matching method, different N values were used 20, 30, 50, and tested in the given case and without filter_range.

The UV map comparison test was divided into 4 times, each time using the correct UV map from the camera angle as the reference map. It is compared with the UV map obtained from a different camera angle, comparing 3 maps at a time: the correct UV map, the low error UV map, and the high error UV map. The test is performed 4 times as follows.

Figure 9 Photos used to generate UV maps for testing.

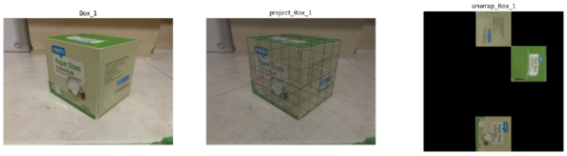

Figure 10 Generate correct UV map from the 1st camera angle box image.

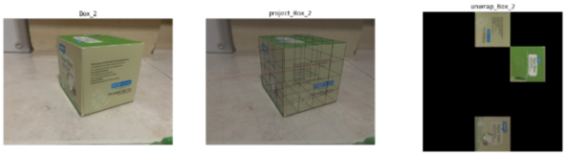

Figure 11 Generate correct UV map from the 2nd camera angle box image.

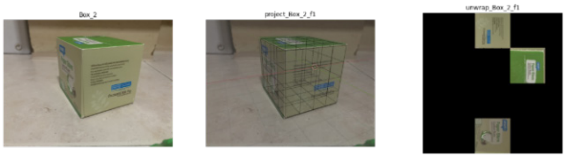

Figure 12 Generate a low-error UV map from the 2nd camera angle box image.

- Correct UV map from Camera 1 compared to all UV maps from Camera 2
- Correct UV map from Camera 3 compared to all UV maps from Camera 2
- Correct UV map from Camera 1 compared to all UV maps from Camera 3
- Correct UV map from Camera 2 compared to all UV maps from Camera 3

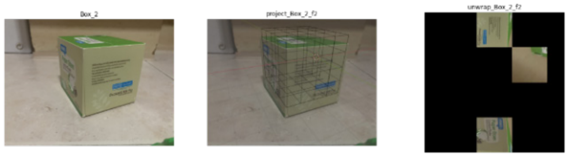

Figure 13 Generate a High-error UV map from the 2nd camera angle box image.

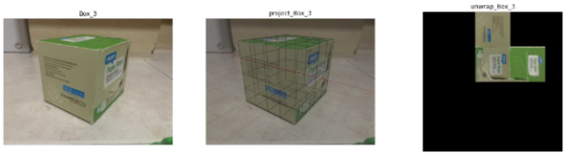

Figure 14 Generate correct UV map from the 3rd camera angle box image.

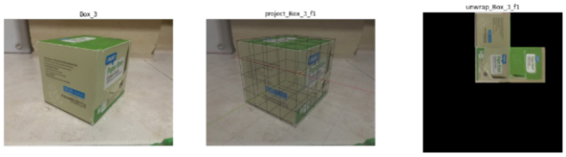

Figure 15 Generate a low-error UV map from the 3rd camera angle box image.

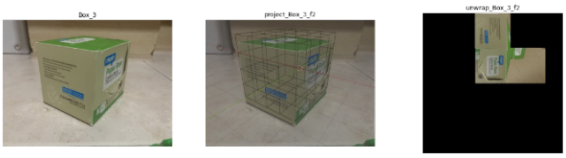

Figure 16 Generate a high-error UV map from the 3rd camera angle box image.

3. Results

From 2.2, there are a total of 5 UV map comparison methods, 4 tests per method. The result of the test is the RMSE value resulting from comparing the correctly generated UV map, the RMSE value resulting from comparing the correctly generated UV map with a low-error UV map, and the RMSE values resulting from comparing the correctly

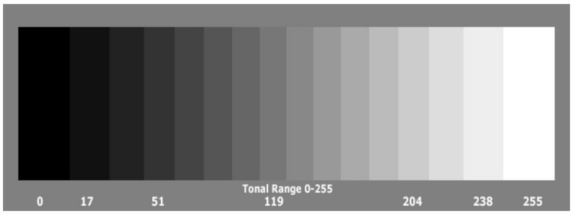

Figure 17 Grayscale color range.

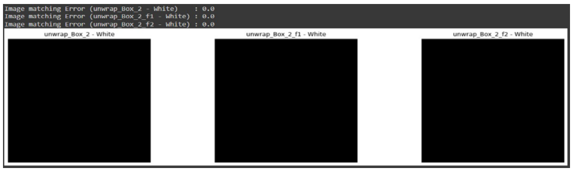

Figure 18 Result when subtract UV map from all 2nd camera angles with the white image.

generated UV map with a high-error UV map. We divide the experiments into the following aspects.

3.1 Pixel-by-pixel Comparison of UV Maps with Normal Subtracting (Subtract)

Table 1, there are some tests where the UV map comparison between the correctly created UV maps does not have the smallest RMSE value. This is due to the Subtract function used. The result obtained by the Subtract function is only in the range of 0–255 (Figure 17), which makes the result time negative. The function rounds to 0 instead, making the value used to calculate the RMSE inaccurate.

Table 1 RMSE values from pixel-by-pixel UV map comparison with normal subtracting (Subtract).

UV ref	UV	UV ref – UV		UV – UV ref	
		Normal	Compare Area	Normal	Compare Area
1	2	2.60558	2.60558	2.05103	2.05103
1	2_f1	2.59655	2.59655	2.50781	2.50781
1	2_f2	2.80160	2.80160	2.67399	2.67399
3	2	2.81479	1.73451	3.34316	2.32703
3	2_f1	2.77493	1.66924	3.46654	2.56385
3	2_f2	2.97194	1.96571	3.46160	2.60327
1	3	3.08443	2.14733	2.79858	1.70775
1	3_f1	3.07234	2.13074	3.06218	2.05384
1	3_f2	3.13409	2.21696	3.38507	2.52241
2	3	3.34316	2.32703	2.81479	1.73451
2	3_f1	3.36374	2.35655	3.01138	1.97706
2	3_f2	3.39616	2.40113	3.37069	2.50252

This problem occurs when the set value is smaller than the subtractor, as in Figure 18, where the UV map is subtracted from the white image (the color value is 255 for the entire image), resulting in a black or the value 0 in the color range of the gray scale.

3.2 Pixel-by-pixel UV Map Comparison with Absolute Value of Difference (Absdiff)

Table 2 shows that the RMSE values of all tests were as expected, i.e., the RMSE from the comparison between the correctly constructed UV maps was the smallest and the RMSE from the comparison between the correctly constructed UV map and the severely mis generated UV map had the largest value. However, there may be a problem if the normal comparison is chosen and the UV map has different sides, leading to discrepancies in the calculation of the RMSE as shown in Figure 19.

3.3 Results from UV Map Comparisons Using the SIFT Feature Matching Method

Table 3 and Table 4 show that the RMSE was mostly correct when comparing with using the SIFT feature matching method in the case of Compare Area only and the filter_range was defined, but mostly wrong in the other cases.

The reason for the error in the RMSE values is due to the accuracy of the feature matching,

Figure 19 An example of the result from the comparison with the Absdiff by not Compare Area.

Table 2 RMSE values from pixel-by-pixel UV map comparison with absolute value of difference (Absdiff).

UV ref	UV	| UV ref – UV |	
		Normal	Compare Area
1	2	3.31599	3.31599
1	2_f1	3.60987	3.60987
1	2_f2	3.87288	3.87288
3	2	4.37033	2.90234
3	2_f1	4.44040	3.05936
3	2_f2	4.56236	3.26206
1	3	4.16482	2.74362
1	3_f1	4.33777	2.95944
1	3_f2	4.61316	3.35820
2	3	4.37033	2.90234
2	3_f1	4.51478	3.07605
2	3_f2	4.78492	3.46815

where some of the feature pairs were not matched correctly as shown in Figure 20. If these are not filtered out beforehand, the RMSE value will be quite high. Defining filter_range and Compare Area only is a must. This will reduce the probability of unmatched feature pairs in the calculations. Because of this problem, N also affects the RMSE value depending on whether the selected feature pair is a valid match or not.

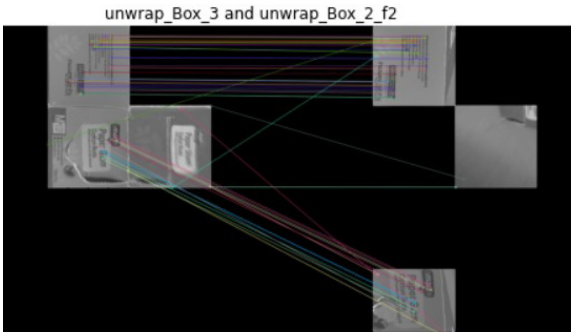

Figure 20 SIFT mismatch example.

Table 3 RMSE values from UV map comparisons using the SIFT feature matching.

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
1	2	106.34	87.53	68.63	1.71	1.83	7.38
1	2_f1	83.73	112.69	117.56	14.42	13.53	15.73
1	2_f2	109.37	131.57	114.30	31.22	31.03	30.26
3	2	146.56	150.45	185.93	4.31	4.16	10.92
3	2_f1	235.63	233.60	226.15	7.11	11.55	15.03
3	2_f2	223.52	228.89	226.04	30.30	29.82	30.70
1	3	246.51	244.57	278.40	23.44	20.49	21.31
1	3_f1	295.16	288.65	293.14	28.92	27.74	29.07
1	3_f2	311.66	286.63	275.34	23.29	21.11	19.86
2	3	146.56	150.45	185.93	4.31	4.16	10.92
2	3_f1	197.22	196.59	210.36	5.52	14.59	18.15
2	3_f2	200.98	209.57	198.34	14.25	13.93	15.92

Table 4 RMSE values from UV map comparisons using SIFT feature matching and Compare Area only.

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
1	2	106.34	87.53	68.63	1.71	1.83	7.38
1	2_f1	83.73	112.69	117.56	14.42	13.53	15.73
1	2_f2	109.37	131.57	114.30	31.22	31.03	30.26
3	2	59.01	18.64	63.52	3.49	3.79	10.81
3	2_f1	122.41	105.77	82.84	7.33	7.61	10.31
3	2_f2	81.99	99.60	85.75	26.15	27.01	27.51
1	3	74.69	62.98	50.53	15.85	19.36	17.38
1	3_f1	68.44	58.07	54.62	20.40	24.08	24.23
1	3_f2	21.78	22.55	29.56	21.78	22.55	20.96
2	3	59.01	48.64	63.52	3.49	3.79	10.81
2	3_f1	65.46	67.54	65.61	5.66	5.36	12.79
2	3_f2	37.34	31.68	26.28	12.96	13.11	13.57

3.4 UV Map Comparisons Using the SIFT Feature Matching Method with Ratio Test

Table 5 and Table 6, there is no difference from comparing UV maps using the SIFT method for feature matching without using Ratio test when only Compare Area and filter_range were assigned. Most of the RMSE values are correct, but in the remaining cases, most of the RMSE values obtained are incorrect. The Ratio test filters out outstanding features that are not clear enough. This makes the resulting feature pair better, but still not accurate enough.

Table 5 RMSE values from UV map comparisons using the SIFT feature matching with ratio test.

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
1	2	101.56	83.60	65.19	12.95	10.70	8.43
1	2_f1	24.31	21.40	22.12	15.46	14.33	18.32
1	2_f2	117.20	98.04	80.22	31.42	31.45	34.60
3	2	192.89	212.51	238.36	4.14	4.08	4.06

Table 5 RMSE values from UV map comparisons using the SIFT feature matching with ratio test (Continued).

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
3	2_f1	245.49	244.99	246.21	11.97	12.09	12.09
3	2_f2	235.14	264.34	264.73	33.27	33.27	33.27
1	3	296.67	309.61	311.51	5.17	5.17	5.17
1	3_f1	308.56	320.86	321.14	17.96	17.96	17.96
1	3_f2	295.39	307.40	306.62	13.94	13.94	13.94
2	3	204.59	228.65	237.51	4.16	3.97	3.91
2	3_f1	191.30	232.26	250.74	8.01	8.22	8.22
2	3_f2	192.00	203.71	221.85	13.87	13.94	13.89

Table 6 RMSE values from UV map comparisons using SIFT feature matching with Ratio test and Compare Area only.

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
1	2	101.56	83.60	65.19	12.95	10.70	8.43
1	2_f1	24.31	21.40	22.12	15.46	14.33	18.32
1	2_f2	117.20	98.04	80.22	31.42	31.45	34.60
3	2	39.08	32.26	26.40	3.43	3.91	4.50
3	2_f1	7.36	10.15	10.28	7.36	10.15	10.28
3	2_f2	88.99	82.42	82.42	30.69	30.27	30.27
1	3	39.46	33.73	33.73	6.31	6.24	6.24
1	3_f1	60.11	60.11	60.11	14.92	14.92	14.92
1	3_f2	13.05	13.05	13.05	13.05	13.05	13.05
2	3	3.43	3.81	25.00	3.43	3.81	4.38
2	3_f1	5.51	48.55	47.08	5.51	7.59	7.51
2	3_f2	12.97	13.30	13.19	12.97	13.30	13.19

3.5 UV Map Comparison Using SuperPoint and SuperGlue Feature Matching

Tables 7 and 8 show that comparing UV maps with the feature matching methods SuperPoint and SuperGlue resulted in RMSE values that were incorrect only when Compare Area only and filter_range were not used. Feature matching with SuperPoint and SuperGlue is more accurate and quantifies feature pairs than SIFT, which leads to better results. However, it requires more time. N values have the same impact on RMSE values as comparisons with SIFT.

Table 7 RMSE values from UV map comparison using SuperPoint and SuperGlue feature matching.

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
1	2	2.32	2.45	2.83	2.32	2.45	2.83
1	2_f1	10.44	13.96	16.81	10.44	13.96	16.81
1	2_f2	38.64	38.18	36.79	38.64	38.18	36.79
3	2	235.94	203.11	222.94	5.39	5.25	5.34
3	2_f1	247.85	237.61	247.44	9.48	9.13	9.91
3	2_f2	324.54	322.32	315.93	31.54	32.16	32.16
1	3	285.38	281.97	274.49	7.24	7.43	7.33
1	3_f1	289.79	270.44	283.01	7.29	8.30	8.95
1	3_f2	337.48	325.05	331.45	15.35	15.29	15.20
2	3	235.94	203.11	222.94	5.39	5.25	5.34
2	3_f1	275.95	269.54	287.36	10.64	10.31	11.11
2	3_f2	301.31	302.51	310.12	14.39	15.16	15.49

Table 8 RMSE values from UV map comparison using SuperPoint and SuperGlue feature matching and Compare Area only.

UV ref	UV	Normal			filter_range = 100		
		N = 20	N = 30	N = 50	N = 20	N = 30	N = 50
1	2	2.32	2.45	2.83	2.32	2.45	2.83
1	2_f1	10.44	13.96	16.81	10.44	13.96	16.81
1	2_f2	38.64	38.18	36.79	38.64	38.18	36.79
3	2	4.42	4.72	5.19	4.42	4.72	5.19
3	2_f1	5.92	8.36	10.40	5.92	8.36	10.40
3	2_f2	29.80	30.50	31.71	29.80	30.50	31.71
1	3	5.90	6.23	6.46	5.90	6.23	6.46
1	3_f1	7.28	7.39	7.75	7.28	7.39	7.75
1	3_f2	13.98	14.06	14.29	13.98	14.06	14.29
2	3	3.11	3.96	4.73	3.11	3.96	4.73
2	3_f1	6.28	7.84	8.50	6.28	7.84	8.50
2	3_f2	13.46	13.68	14.25	13.46	13.68	14.25

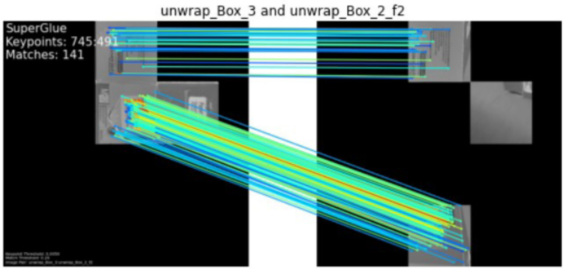

Figure 22 An example of SuperGlue's feature mismatch.

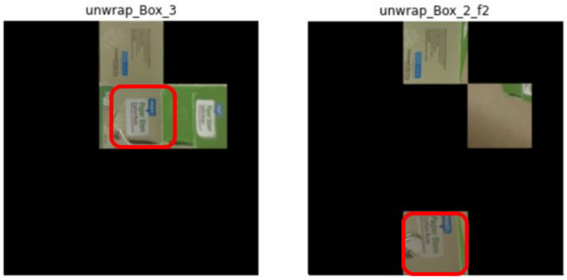

Figure 21 Examples of different sides but similar characteristics.

In the test, the error is caused by non-matching on the same side, so if filter_range is omitted, non-matching feature pairs are computed. Moreover, it may be on the test image because ithas similar sides, but not the same side.

In Figure 21, the red squares of the two UV maps are similar, resulting in the mismatch of the features as in Figure 22.

3.5 Discussion

The accuracy of each comparison method was determined by taking the number of valid tests (the RMSE value resulting from the comparison between the correctly generated UV map was the smallest and the RMSE value resulting from the comparison between the correctly generated UV map and high-error UV map is the largest) relative to the total number of tests. In the comparison of the UV maps with the feature matching method with *N* parameters, the accuracy obtained is the mean of all N (20, 30, 50) accuracy.

Table 9 shows that two methods have the highest comparison accuracy: comparing the UV map pixel by pixel with the absolute value of the difference and comparing using the feature matching methods SuperPoint and SuperGlue with filter_range defined, both of which have 100% accuracy for both normal UV map comparisons and Compare Area comparisons.

Table 9 UV Mapping accuracy of various methods. Comparison

Methods	Algorithm	Normal	Compare Area
Pixel by Pixel	Subtract	50.00%	67.50%
Pixel by Pixel	Absdiff	100.00%	100.00%
Feature Matching	SIFT	33.33%	25.00%
Feature Matching	SIFT with filter_range=100	58.33%	75.00%
Feature Matching	SIFT and Ratio Test	16.67%	8.33%
Feature Matching	SIFT and Ratio Test with filter_range=100	75.00%	75.00%
Feature Matching	SuperPoint and SuperGlue	91.67%	100%
Feature Matching	SuperPoint and SuperGlue with filter_range=100	100%	100%

4. Conclusion

We have presented a method for comparing 3D objects on an image with a UV map by comparing the UV map in the desired frame with the reference UV map or the frame-derived UV map verified as valid by humans. Our experiments show that 1) it is possible to use UV maps to support 3D Object Tracking, 2) there are several comparison methods that can be used to compare UV maps, and 3) the results obtained from the comparison can indicate a discrepancy.

The future work can study more aspects in comparison such as execution time and practicability Also, the preprocessing using RANSAC can be used to improve the algorithm efficiency.

References

[1] B. Lévy, S. Petitjean, N. Ray, and J. Maillot, “Least squares conformal maps for automatic texture atlas generation,” *ACM Transactions on Graphics*, vol. 21, no. 3, pp. 362–371, 2002.

[2] D. Lowe, “Distinctive Image Features from Scale-Invariant Keypoints,” *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[3] D. DeTone, T. Malisiewicz and A. Rabinovich. (2018). SuperPoint: Self-supervised interest point detection and description. *arXiv*. [Online]. Available: https://arxiv.org/pdf/1712.07629.pdf

[4] K. M. Yi, E. Trulls, V. Lepetit, and P. Fua. (2018). LIFT: Learned invariant feature transform. *arXiv*. [Online]. Available: https://arxiv.org/pdf/1911.11763.pdf

[5] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, “ORB: An efficient alternative to SIFT or SURF,” in *Proceedings of the 2011 International Conference on Computer Vision*, 2011, pp. 2564–2571.

[6] P. Sarlin, D. DeTone, T. Malisiewicz, and A. Rabinovich. (2020). SuperGlue: Learning feature matching with graph neural network. *arXiv*. [Online]. Available: https://arxiv.org/pdf/1911.11763.pdf