Color Pattern Diffusion：Verification for Wide Baseline Matching

Liu Xinhai
Engineering Research Center of Metallurgical Automation and Measurement Technology，Ministry of Education；
College of Information Science and Technology
Wuhan University of Science and Technology
Wuhan, China
liu.xinhai8@gmail.com

Fang Kangling
Engineering Research Center of Metallurgical Automation and Measurement Technology，Ministry of Education；
College of Information Science and Technology
Wuhan University of Science and Technology
Wuhan, China
klfangcn@yahoo.com.cn

Xu Wangming
Engineering Research Center of Metallurgical Automation and Measurement Technology，Ministry of Education；
College of Information Science and Technology
Wuhan University of Science and Technology
Wuhan, China
hyuan19@163.com

Wu Jin
Engineering Research Center of Metallurgical Automation and Measurement Technology，Ministry of Education；
College of Information Science and Technology
Wuhan University of Science and Technology
Wuhan, China
wujin1988@163.com

***Abstract*—After image retrieval based on wide baseline match, verification is required since there are still many false match pairs. The popular verification method, geometry, is limited for example in a weak match situation where the number of match pairs is less than enough. In order to improve retrieval results in such cases, this paper puts forward color pattern diffusion verification. Firstly, HSI Color moments in the original region where two corresponding SIFT feature points are detected are extracted from both the query image and the database image and then matched with each other. Secondly, color moments in both diffusion regions are extracted and matched as well. Finally, if the matching similarity of above color moments keeps invariant with color pattern diffusion in image space, the two interest points is taken as a real corresponding pair. Experiments demonstrated this approach works better with near real-time performance in many practical situations, especially in the weak match situation context.**

***Keywords*—image retrieval; wide baseline matching; verification; color pattern diffusion**

I. INTRODUCTION

Wide baseline matching for Content-Based Image Retrieval has been extensively studied in recent years[1][2].Wide baseline matching[1] refers to the type of algorithms that seek the images in the database that are more similar to the ones taken by the mobile camera. In order to match under such conditions, the number of reference images in the scene model should be kept to a minimum. Most wide baseline matching approaches are based on so-called invariant regions, especially, the intensity invariant regions. These regions are constructed around interest points, such as corners, so as to adapt their shapes to the viewpoint and keep the part of the scene they enclose fixed. These regions are then described by a descriptor vector, the elements of which are invariant under combinations of geometric and photometric changes. They can be matched efficiently between views taken from different viewpoints and under different illumination. The crux of the matter is that these affine invariant regions are determined solely on the basis of a single image, i.e. no information about the other view(s) is necessary during extraction.

Many recent studies on invariant feature descriptors have focused on so called scale-invariant feature transform, briefly, SIFT[2].

In order to match efficiently, some image retrieval approaches are usually employed. Vector quantization and clustering[4] over SIFT features of the database images, are involved in the building process of retrieval structure. While the near neighborhood search method is concerned in the query stage. There are many effective image retrieval method, such as Kd-tree, ball tree and vocabulary tree. Usually after retrieval, a list of candidate matched images in the database is offered.

It is obvious that there are many false match results after image retrieval due to feature vector quantization error. Thus, some verification methods are adopted to re-check the match. A central issue is geometry constraints. In Goedeme's paper[1], a verification method by RANSAC based on Fundamental Matrix, is used. Lower[2] adopted object pose verification by Hough Transform. According to Josef[6], he verified the image match with spatial consistency of corresponding features distribution. Apart from geometry verification, Bay[8] and Ferrari[7] put forward topological constraint for verification. Although considerable research has been devoted to such conditions as above where the number of real match feature pairs is more than enough, rather less attention has been paid to some practical situations, especially , the weak match situation.

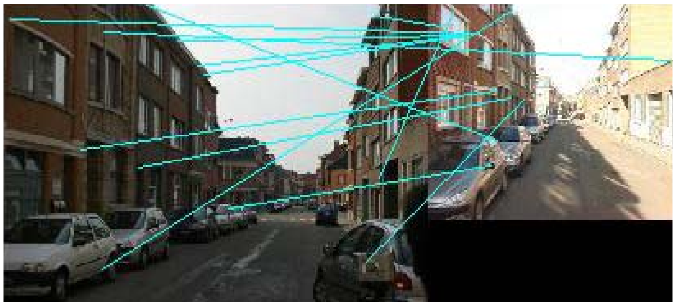

Figure 1 the number of real match points pairs is less than 7

As shown in Fig. 1, the number of real match point pair is less than 7 while actually it is only 4. Under such condition, both geometrical verification and topological verification seem invalid.

Color is an important component for distinction between objects. If the color information in an object is neglected, a very important source of distinction may be lost. In some case, pure gray based geometry description can cause confusion between two different features while their color patterns seem very distinctive, as is shown in Fig. 2.

Figure 2 Distinction in color pattern

At the same time, the color constraint for feature match has no requirement about the number of real match feature pairs. (Although color information for match has been employed in paper[9] where color information is directly employed for originally matching, color information is more reliable for verification than for first matching when the color constraint is not very strict.)

Therefore, this paper explores a novel verification algorithm based on color pattern diffusion. The remainder of the paper is organized as follows. Section 2 discusses the feature descriptor and image retrieval method employed here. Section 3 explains the verification algorithm base on color pattern diffusion. Experiments follow in Section 4. We conclude with section 5.

II. IMAGE RETRIEVAL TECHNOLOGY

A. *SIFT features*

SIFT features are based on the appearance of the object at particular interest points, and is invariant to image scale and rotation. They are also robust to changes in illumination, noise, occlusion and minor changes in viewpoint. In addition to these properties, the local features given by SIFT are be highly distinctive, relatively easy to extract, allow for correct object identification with low probability of mismatch and are easy to match against a (large) database of local features. They are also robust to occlusion; as few as 3 features from an object are enough to compute its location and pose. In addition to object recognition, local features can be used for matching, which is useful for tracking and 3D scene reconstruction. Recognition can be performed in close to real time, at least for small databases. There has been an extensive study done on the performance evaluation of different local descriptors, including SIFT, using a range of detectors[3]. The evaluation carried out suggests strongly that SIFT-based descriptors, which are region-based, are the most robust and distinctive, and are therefore best suited for feature matching.

B. *Kd-tree –the retrieval structure*

A Kd-tree of the reference image data is built of SIFT features. We used the on-line available package ANN by Mount et al.[4] for this. Searching in this kind of database is very fast. Also, the database can contain data of more than one reference image, so we can match one image in parallel with several images.

III. VERIFICATION ALGORITHM

Since there are still many mismatch point pairs existing, verification methods are required to reject these mismatch point pairs, thereby rejecting the mismatch images pair. SIFT features are based on gray intensity pattern. If there are many repeatable patterns existing, such as outdoor buildings image of a city, SIFT features seems not robust enough to match. While the color information is a good supplementary to SIFT because many color pattern around SIFT features which are repeatable gray intensity patterns are quite a bit different. Therefore a verification method based on color information is put forward in this paper.

A. *Generalized color moments*

"Generalized color moments" were introduced in paper[5], which implicitly characterize the shape, the intensity and the color distribution of the pattern in a uniform manner; and a broad set of moment invariants can be extracted that never call upon high powers of either intensities or spatial coordinates. Under certain condition, experiments demonstrated that these moment invariants can cope with relatively high degrees of photometric distortions(strong partial shadowing, high intensity lighting, etc) and with fair amounts of image blur. In paper[1], it presents a new kind of color descriptor: C_{RG}, C_{RG} and C_{GB} with

$$C_{PQ} = \frac{\int_{\Omega} PQd\ \Omega \int_{\Omega} d\Omega}{\int_{\Omega} Pd\ \Omega \int_{\Omega} Qd\ \Omega} \quad (1)$$

Where $P, Q \in \{R, G, B\}$, i.e. the red, green, and bluer color bands, centralized around their means. The selected three combinations of color moments have 0th order and 2nd degree, and are the lowest-order invariants yielding invariance to photometric changes according to the definition in paper[5].

B. Generalized Color moments in HSI color space

The original color space of common image is RGB space. Compared with RGB space, HSI space attempts to describe perceptual color relationships more accurately than *RGB*, while remaining computationally simple. Furthermore, HIS space has more physical meaning.

In this paper, we extend the generalize color moment from RGB color space to HSI color space due to the good performance of HSI. Thus a class of new color descriptor is introduce as following:

$$C_U = \frac{\int_{\Omega} Ud\Omega}{\int_{\Omega} d\Omega} \quad (2)$$

$$C_{WV} = \frac{\int_{\Omega} WVd\Omega \int_{\Omega} d\Omega}{\int_{\Omega} Wd\Omega \int_{\Omega} Vd\Omega} \quad (3)$$

Where $U, W, V \in \{H, S, I\}$, i.e. the hue, saturation and intensity bands, centralized around their means. There are totally six elements in this color descriptor over certain region in an image: $C_{\Omega} = \{C_H, C_S, C_I, C_{HS}, C_{SI}, C_{HI}\}$.

C. Color pattern diffusion

Concerning the same object and same scene in different views, color pattern should keep invariant. As before, some color or gray intensity descriptor is extracted from the image region where an object or a scene lies. These descriptors should be very similar in terms of vector space. But there are also false similarities due to repeatable pattern. If the size of above image region is enlarged to some degree, some descriptors are re-extracted. With the regions' growing, the same color pattern will diffused while the different color patterns will become more distinctive. The color pattern diffusion property could be measured by the stability of the similarity of these descriptors two times.

During matching process, the similarity of corresponding color vectors is measured by Euclidean distance defined as following:

$$S = \left| C_{\Omega 1} - C_{\Omega 2} \right| \quad (4)$$

Where $C_{\Omega 1}$ is a color vector over the region $\Omega 1$ in one image and $C_{\Omega 2}$ is a color vector over the corresponding region $\Omega 2$ in another image.

As the region $\Omega 1$ diffuses to region $\Omega 1'$ in the 1st image, the region $\Omega 2$ diffuses to region $\Omega 2'$ in the 2nd image. The similarity after diffusion is defined as

$$S_d = \left| C_{\Omega 1'} - C_{\Omega 2'} \right| \quad (5)$$

Based on color pattern diffusion invariance within certain degree, for example, if $S/S_d > T$, the match relationship of point pair is rejected , otherwise , it is kept, where T is a threshold.

Figure 3 match under color pattern diffusion

Fig. 3 displays two matching correspondences by SIFT. The first one which is real match pair is plotted in solid line while the second one which is false match is plotted in dashed line. Around each interest point or SIFT feature point, there are two circles: the inner circle is the original region and the outer circle is the diffusion region. It is obvious that the color pattern similarity from inner circle to outer circle in the first corresponding points varies a little and that in the second corresponding points is very different.

IV. IMPLEMENTATION AND EXPERIMENT

A. implementation of above algorithm

Based on the HSI space, some kinds of color moments are generated from the interest point detected by SIFT. At the same time, in order to keep the scale invariant, the bound for color moments calculation is proportional to the SIFT scale of the same interest point. It is easy to implement this task. During adding a SIFT feature to a Kd-tree, its scale and location in the image are added to the Kd tree as well. As a result, this color pattern verification has the same scale invariance property as SIFT.

Since Saturation element of HSI color model is very sensitive to illumination variation, then the color invariants containing saturation element are not employed in actual application.

B. Experimental result

In this section, we evaluate the performance of above algorithm by querying 120 local images from a retrieval system which contains more than 2000 images, comparing it with other algorithms. In addition, both querying image and retrieval system are captured from outdoor context in order to show how the algorithm is robust to the complicate environment.

TABLE I. COMPARISION OF RETRIEVAL RESULTS

Measurement	**Match method**		
	SIFT	***GEOV***	***CPDV***
Precision	0.4195	0.7524	0. 7386
False Alarm Rate	0.5805	0.3476	0.3614

Table 1 shows comparison of experimental result with different methods: SIFT, Geometry Verification (GEOV) and Color Pattern Diffusion Verification (CPDV). Where Precision is defined as the ratio of verified positive feature pairs to the total number verified and False Alarm Rate means the ratio of both verified false negative and verified false positive feature pairs to the total number verified. As can be seen, the results of GEOV and CPDV are better than that of only SIFT. It is obvious that from both sides, the result of Geometry verification seems slightly best. However GEOV would fail under the weak match situation while CPDV is robust, typically, in the case of the number of match correspondences less than 7. Thus color pattern diffusion is the suitable choice for verification there. The relative lower Precision among above methods might be due to the fact that the match situation under outdoor context is more complicate to some degree.

Furthermore, the running time of CPDV is similar to that of GEOV: on a 1GHz machine on 640×480 color image, the operation time of verifying one query image is about 0.446 second on average.

V. CONCLUSION

A new verification approach based color pattern diffusion is presented in this paper. This feature descriptor contains more information than its counterpart in intensity pattern. The method owns the good property of scale invariant, and reduces some repeatable features to some degree during its verification. Compared with geometrical verification method, it specially works well in the cases where the feature match situation is very weak, which is very common in the research of wide baseline matching. Furthermore, it has almost the same efficiency as geometrical verification.

On the other hand, this approach is not robust enough to outdoor context, such as very strong illumination, occlusion, deformation and so on. Therefore further work is still ongoing.

REFERENCES

[1] Toon Goedeme, Tinne Tuytelaars, and Luc Van Gool, “Fast Wide Baseline Matching for Visual Navigation,” CVPR’S 04(IEEE Computer Vision and Pattern Recognition 2004)

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

[3] Mikolajczyk, K., and Schmid, C, "A performance evaluation of local descriptors," IEEE Transactions on Pattern Analysis and Machine Intelligence, 10, 27, pp. 1615--1630, 2005.

[4] S. Arya, D. Mount, N. Netanyahu, R. Silverman, and A.Wu, “An optimal algorithm for approximate nearest neighborsearching,” J. of the ACM, vol. 45, pp. 891-923, 1998, http://www.cs.umd.edu/ ~ mount/ANN/.

[5] F. Mindru, T. Moons, and L. Van Gool, “Recognizing color patters irrespective of viewpoint and illumination,” CVPR’S 99 vol. 1, pp. 368-373.

[6] Josef Sivic and Andrew Zisserman. “Video Google: Efficient Visual Search of Videos,” Toward Category-Level Object Recognition 2006: pp.127-144.

[7] Vittorio Ferrari, Tinne Tuytelaars, and Luc Van Gool, “Wide-baseline muliple-view Correspondences,” CVPR’S 03

[8] Herbert Bay, Vittorio Ferrari, Luc Van Gool, “Wide-Baseline Stereo Matching with Line Segments,” CVPR’S 05.

[9] A. E. Abdel-Hakim, and A. A. Farag,” CSIFT: A SIFT Descriptor with Color Invariant Characteristics ,” CVPR’S 06, Vol. 2 (2006), pp. 1978-1983.