Complex Printed Uyghur Document Image Retrieval Based on Modified SURF Features

Aliya Batur[1], Patigul Mamat[2], Wenjie Zhou[1], Yali Zhu[1],
and Kurban Ubul[1](✉)

[1] School of Information Science and Engineering,
Xinjiang University, Urumqi 830046, China
kurbanu@xju.edu.cn
[2] School of Mathematics and Information,
Hotan Normal College, Hotan 848000, China

Abstract. As an important part of information retrieval, it is important to improve the accuracy of document image retrieval system. This paper proposes a document image retrieval method based on modified SURF features. Firstly, FAST+SURF features are extracted from the image, and then the similarity degree is retrieved by using different kinds of distances and matching points respectively. With the change of size, angle and illumination, the FLANN bidirectional matching and KD-Tree +BBF matching are implemented for its feature points; finally, based on these two kinds of retrieval methods, various Uyghur document image databases that have been collected and retrieved are searched. The experimental results indicated that both search methods can achieve accurate search requirements, but in computational complexity based on the matching number of retrieval is more convenient. At the same time, the comparison experiment proves that the proposed method is superior to the original feature in the retrieval time.

Keywords: SURF feature · FALNN bidirectional match
KD-Tree and BBF match · Complex document image retrieval

1 Introduction

With the rapid development of multimedia information technology, document images have become the main information resource, which also causes the explosive growth of document image. How to obtain document image content efficiently has become a hot research topic in domestic and overseas research. Xiaoxiao et al. [1] compared 64-dimensional vectors to describe the feature points that were more suitable for image data processing. Two modified SVM algorithms were used to extract information from matched images and compare with traditional SVM algorithm. Zhao et al. [2] first extracted the 64-dimensional SURF feature points, and based on the FLANN algorithm for bidirectional matching, matching pairs for PROSAC analysis, excluding mismatched pairs to improve the image matching accuracy, and effectively reduce the matching time. Cheon et al. [3] proposed an enhanced Fast Robustness Feature (e-SURF) algorithm to save memory and increase speed. Zhang et al. [4] proposed an

© Springer Nature Switzerland AG 2018
J.-H. Lai et al. (Eds.): PRCV 2018, LNCS 11258, pp. 99–111, 2018.
https://doi.org/10.1007/978-3-030-03338-5_9

modified matching algorithm based on SURF (Speeded Up Robust Features) feature point matching, which combined SURF and RANSAC (Random Sample Consensus) algorithm. Chen et al. [5] proposed to improve the detection of SURF key points, extract the feature points of the image detail region, and achieve accurate matching based on KD-Tree bidirectional matching. Luo et al. [6] modified the SURF descriptor using the DAISY descriptor, and matched the target image with nearest neighbor distance ratio (NNDR), with a maximum matching rate of 95.78%. Wang et al. [7] proposed a robust feature (SURF) based on improved accelerated fast image matching algorithm, The RELIEF-F algorithm is used to reduce and simplify the improved SURF descriptors to achieve image registration, and finally the improved algorithm is verified by the experiments of real-time and robustness.

This paper analyzes the Uyghur complex document image without layout analysis, proposes to the modified SURF features to achieve the key points extraction, and to achieve effective retrieval from the large-scale image database. The algorithmic flow of this paper is shown as in Fig. 1.

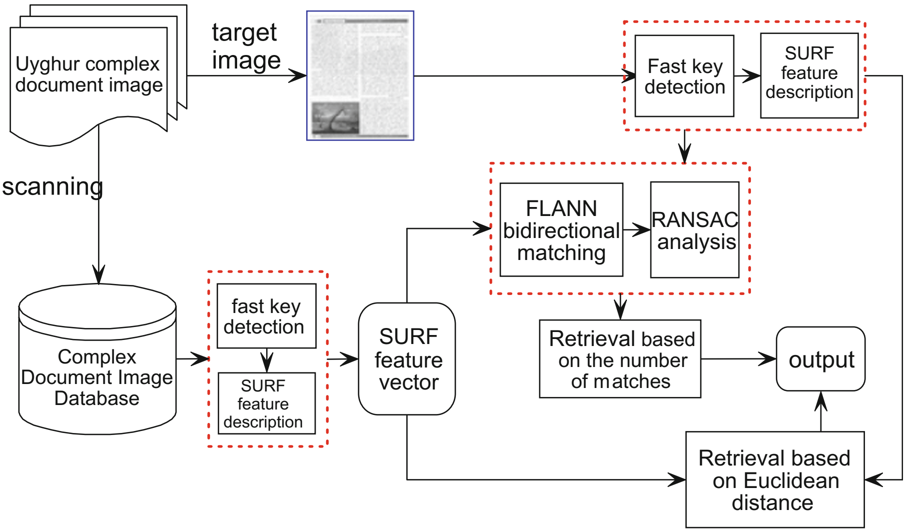

Uyghur complex document image
target image
scanning
Fast key detection
SURF feature description
FLANN bidirectional matching
RANSAC analysis
Complex Document Image Database
fast key detection
SURF feature description
SURF feature vector
Retrieval based on the number of matches
output
Retrieval based on Euclidean distance

Fig. 1. The block diagram of Uyghur printed complex document image retrieval based on modified SURF feature.

2 Fast and SURF Feature Extraction

The process of fast robust feature extraction (SURF) is similar to SIFT, and consists of two parts: key point detection and feature description. However, it maintains the same image size and changes the size and scale of the box filter in multiples relationship. Based on the integral image, the proportional space is filtered so that the feature detection takes much less time than SIFT. And the key points detected in the scale space have the size translational robustness. In the feature description, the Haar wavelet response value in the fan-shaped area is calculated, the main direction of the key points is determined, and the computational complexity is reduced.

However, shortening time parameter is not ideal for the complicated document images of text and video. Therefore, in order to quickly detect the key points of the image in the complex layout, the author makes full use of pixel gray level information, detects the corners based on the FAST algorithm, and describes the sub-description with the SURF descriptor to form the 64 dimension FAST and SURF feature, effectively shortening the features Extraction time [8]. The Flow chart of modified SURF feature key point detection is shown in the following Fig. 2.

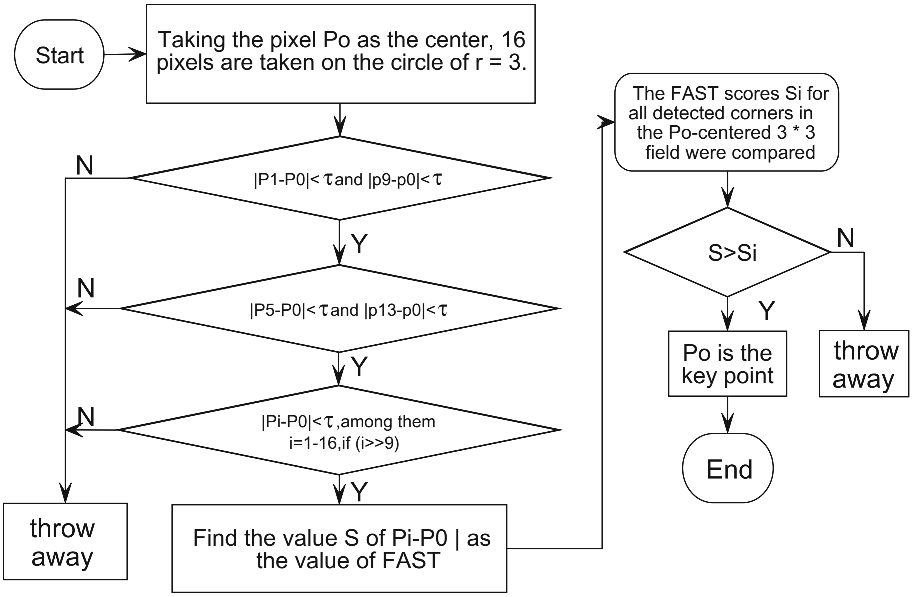

Start
Taking the pixel Po as the center, 16 pixels are taken on the circle of r = 3.
N
|P1-P0|< τand |p9-p0|< τ
Y
N
|P5-P0|< τand |p13-p0|< τ
Y
N
|Pi-P0|< τ,among them i=1-16,if (i>>9)
Y
throw away
Find the value S of Pi-P0 | as the value of FAST
The FAST scores Si for all detected corners in the Po-centered 3 * 3 field were compared
S>Si
N
Y
Po is the key point
throw away
End

Fig. 2. Flow chart of modified SURF feature key point detection

3 Fast and SURF Feature Matching Analysis

To improve the matching speed of Uyghur complex document images, the author implements two-way fast approximate nearest neighbor (FLANN) matching for different layout images, and compares the results with KD-Tree and BBF matching results, from the performance of matching system to establish a retrieval system, and realize the effective retrieval of Uyghur complex document images.

3.1 Bidirectional FLANN Match

Due to the SURF feature vector is a high-dimensional vector, the matching process is equivalent to the nearest neighbor search problem in high-dimensional space, and the operation is complex. Therefore, this paper starts from the rapidness of FLANN matching, and matches in two directions successively to get the location information of matching pair. By comparing the location of the matching point to determine whether it is correct. In order to effectively remove the mismatched point pairs, the author uses the RANSAC algorithm to calculate the distance between the matched points and the projection matrix, and compares it with the threshold value, effectively eliminating the outer points and improves the matching accuracy. The original image FALNN bidirectional matching results are shown in Fig. 3.

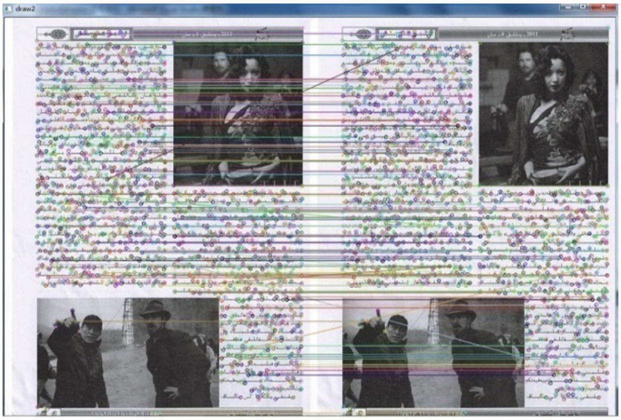


Fig. 3. Schematic diagram of modified SURF features bidirectional FLANN matching

3.2 KD-Tree and BBF Match

KD-Tree is a tree structure for realizes K-nearest neighbor search and matching in large-scale high-dimensional eigenvector space. Its research mainly consists of two parts, namely, the establishment of tree structure and the nearest neighbor search. With the increase of image feature vector dimension, the KD-Tree search ability is greatly reduced. Therefore, starting from the modified KD-Tree, this paper finds the nearest neighbor distance within the limit of maximum backtracking times, and compares the distance ratio with a predetermined threshold to determine whether it is a matching key point [9]. In this paper, the process of improving KD-Tree matching by improving 64-dimensional SURF features is shown in Fig. 4.

The matching efficiency of the matching system under different transformation conditions is evaluated by the matching rate, the correct matching rate and the false matching rate, and its mathematical expression is as follows:

$$\text{Match rate} = \frac{\text{The total number of matched pairs}}{\text{The total number of feature points detected}} \tag{1}$$

$$\text{Correct match rate} = \frac{\text{The total number of correct matched}}{\text{The total number of matched pairs}} \tag{2}$$

$$\text{Mismatch rate} = \frac{\text{The total number of error matched}}{\text{The total number of matched pairs}} \tag{3}$$

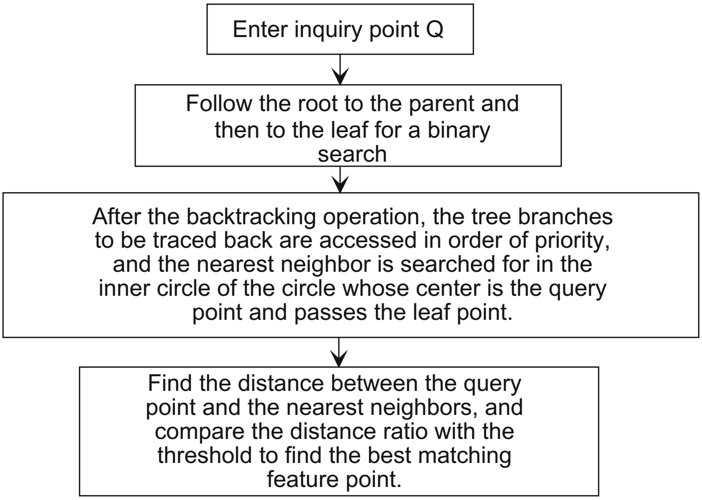

Enter inquiry point Q
Follow the root to the parent and then to the leaf for a binary search
After the backtracking operation, the tree branches to be traced back are accessed in order of priority, and the nearest neighbor is searched for in the inner circle of the circle whose center is the query point and passes the leaf point.
Find the distance between the query point and the nearest neighbors, and compare the distance ratio with the threshold to find the best matching feature point.

Fig. 4. The description of modified KD-Tree match

4 Uyghur Complex Document Image Retrieval Method

In this paper, the distance-based similarity measure and the matching number-based similarity measure are used. Four eigenvector distance similarity measures algorithms are selected and they are as follows:

$$\text{Euclidean distance} = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \tag{4}$$

$$\text{Manhattan distance} = |x_1 - x_2| + |y_1 - y_2| \tag{5}$$

$$\text{Chebyshev distance} = \max(|x_1 - x_2|, |y_1 - y_2|) \tag{6}$$

$$\text{Cosine distance} = \frac{x_1 x_2 + y_1 y_2}{\sqrt{x_1^2 + y_1^2} \times \sqrt{x_2^2 + y_2^2}} \tag{7}$$

In the retrieval system that based on the matching number, the correct number of matches between the document image and each image in the image database is calculated from the correct number of matches, and then the number of correct matches between each image in the image database is sorted and sorted in descending order to effectively retrieve the document image. The more similar complex document images, the greater the number of matches. The calculation of retrieval system is:

$$\text{Retrieval rate} = \frac{\text{N} - \text{S}}{\text{N} - 1} * 100\% \tag{8}$$

Where N is the size of the complex document image database, and S is the position number of the target document image output in the retrieval system window.

5 Experimental Results and Analysis

5.1 Experimental Data

Collection of Uyghur complex layout of books, magazines, documents, scanned with a resolution of 100 dpi to form a depth of 8. bmp format of weaving complex document images, construction of 1000 complex document image database. The system is in 4 GB memory, Windows7_64 bit operating system environment, and Visual Studio 2010 programming.

5.2 Matching Analysis Under Various Transformation Conditions

The original SURF feature detection relies on the choice of Octaves, Intervals, and thresholds. Under the different thresholds (Octaves, Intervals, Init-sample, THRES), the number of feature points to be acquired varies greatly. To test and verify the feasibility of FAST and SURF features, the original SURF features were extracted at (4, 4, 2, 0.0004f) thresholds for complex text documents, in order to obtain the same layout with different sizes, and compared with FAST and SURF Features for performance analysis. The experimental results are shown in Table 1.

Table 1. Number of FAST+SURF key points and time statistics of different sizes image under different threshold [10].

Image size	Feature					
	Performance	SIFT	SURF	FAST (50) +surf	FAST (100) +surf	FAST (150) +surf
803 × 1145	Key points	3276	3537	9767	7195	4960
	Occupation time (S)	30.405	15.866	0.021	0.01	0.009
1606 × 2290	Key points	10320	11516	22028	9414	9299
	Occupation time (S)	105.450	51.141	0.031	0.019	0.017
3212 × 4581	Key points	27820	38115	52967	17764	7839
	Occupation time (S)	250.492	162.491	0.082	0.04	0.035

In order to detect the robustness of the extracted features to the rotation, scale and illumination transformation, the modified SURF eigenvectors of the Uyghur complex document image with the size of 1606 × 2290 were extracted under different transformations. Based on FLANN bidirectional matching, KD-Tree and BBF matches the number of exact match pairs. When the threshold $\gamma = 0.1$, the results of FLANN bidirectional matching and KD-Tree and BBF under the dimensional transformation are shown in Table 2.

Table 2. Uyghur document image different matching results of FAST+SURF features under scale transform condition [10].

	FLANN bidirectional matching			KD-Tree and BBF match		
	Whole	*1:1/2*	*1:1/4*	*Whole*	*1:1/2*	*1:1/4*
The total number of key points	9414	4369	2082	9414	4369	2082
The total number of matching pairs	1145	454	200	9335	4264	2017
Correctly matched pairs	759	363	172	7582	4151	573
Correct match rate (%)	66.29	79.96	86	81.22	97.35	28.40

As can be seen from Table 2 that the image area decreases, the number of feature points detected decreases, and the total number of matches also cut back. Therefore, for thousands of key points, the stability of FLANN bidirectional matching is stronger than KD-Tree and BBF matching. To test and verify the rotation invariance of the selected features, the complex document images are rotated anticlockwise or clockwise in different angular ranges, and matching based on different matching algorithms. The experimental results are shown in Table 3.

Table 3. Two kinds of FAST (100)+SURF feature points matching results under Uyghur document image rotation transform

FLANN bidirectional matching					
	0°	*+5°*	*+10°*	*−5°*	*−10°*
The total number of key points	9414	10373	10614	9339	10423
The total number of matching pairs	1145	1213	1164	1070	1254
Correctly matched pairs	759	791	757	752	785
Correct match rate (%)	66.29	65.21	65.03	70.28	62.60
KD-Tree and BBF match					
	0°	*+5°*	*+10°*	*−5°*	*−10°*
The total number of key points	9414	10373	10614	9339	10423
The total number of matching pairs	9335	6044	6691	6005	6794
Correctly matched pairs	7582	1802	2112	2015	1845
Correct match rate (%)	81.22	29.81	31.56	33.56	27.16

Rotating the image of a complex text document in the anti-clockwise or clockwise direction can enlarges the image area. Therefore, the number of the detected key points is appropriately increased and the position of the key point is changed. From Table 3, it can be seen that FLANN bidirectional matching performance is better than KD-Tree and BBF matching under the condition of rotation transformation. In order to verify the robustness of the feature under light conversion conditions, the brightness of the original document image is adjusted. The experimental results are shown in Table 4.

Table 4. Two types of FAST (100)+SURF feature points matching results with Uyghur document image illumination transform

FLANN bidirectional matching					
	0	*20*	*40*	*−20*	*−40*
The total number of key points	9414	9882	9411	8963	9647
The total number of matching pairs	1145	1073	1139	999	952
Correctly matched pairs	759	749	758	803	784
Correct match rate (%)	66.29	69.80	66.55	80.38	82.35
KD-Tree and BBF match					
	0	*20*	*40*	*−20*	*−40*
The total number of key points	9414	9882	9411	8963	9647
The total number of matching pairs	9335	4892	4931	5591	5789
Correctly matched pairs	7582	3556	1869	3499	3281
Correct match rate (%)	81.22	72.69	37.90	62.58	56.68

The change of illumination is the lightness and darkness of the image. From Table 4, it can be seen that the KD-Tree and BBF matching performance is better than FLANN matching under the key point matching under light conversion conditions, and the matching number is large and the matching rate is high.

5.3 Analysis of Search Results

Due to the large size of the original image collected, the number of feature points obtained by feature extraction is too large, which has a great influence on the number of final matching points. Therefore, in order to assess the performance of the retrieval system, two modifications were made to the overall Uyghur complex document image database by compressing each image and cutting each image into 256 * 256 size, as shown in Fig. 5, and constructed two kinds of Uyghur complex document image database.

In Fig. 5, Fig. 5(b) is sheared image from Fig. 5(a). For the above two improved Uyghur complicated document image databases, based on the number of matches, Euclidean distance and cosine distance similarity measures, the user-specific target document images are retrieved. The retrieval test results are shown in Tables 5 and 6.

(a) Compressed image sample (b) Sheared image sample

Fig. 5. The sample instance of modified database

Table 5. The statistical results of the sheared Uyghur document image retrieval experiment

Retrieve performance indicators	Match the number of search	Euclidean distance search	Cosine distance search
Retrieval rate	100%	100%	100%
Total search time (s)	1000	854	861
Average index time (s)	1	0.854	0.861

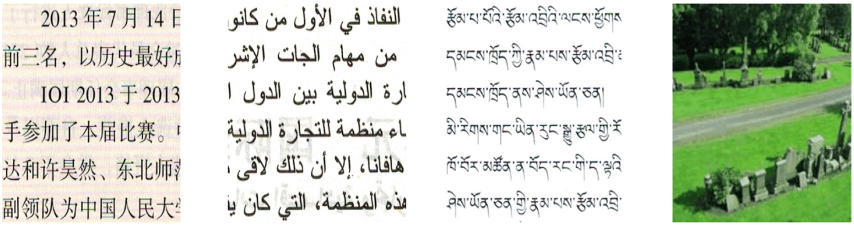
Table 6. The statistical results of the compressed Uyghur document image retrieval experiment

Retrieve performance indicators	Match the number of search	Euclidean distance search	Cosine distance search
Retrieval rate	100%	100%	100%
Total search time (s)	1636	599	607
Average index time (s)	1.636	0.599	0.607

It can be seen from Table 5 above that all three retrieval methods in the cut-structured Uyghur complex document database achieve a retrieval rate of 100%, but the search occupancy time is different. The matching needs to find the nearest neighbor and the next nearest neighbor matching point of each key point, and it need to compare the distance ratio with the first threshold to determine whether they match. Therefore, the system consumes more time than the distance similarity metric retrieval algorithm. The experimental results of compressed Uyghur document image are indicated in Table 6.

It can be seen from Table 6 that the retrieval system based on the number of matches consumes more time than the distance similarity metric retrieval system. For the two databases, matching number based retrieval system, the more the number of image feature points is, the greater the number of matching and the greater the system matching index time. In terms of similarity measure of distance between feature vectors, although the number of vector images in compressed image is larger, the output target document image can be searched within a shorter time than the cut image.

In this paper, in order to further validate the effectiveness of the FAST and SURF algorithm proposed in this paper, a cut-file image database of 256 * 256 size Arabic, Chinese, Tibetan and natural images is collected, each of which has a size of 1000 frames. The sample example is shown in Fig. 6 below:

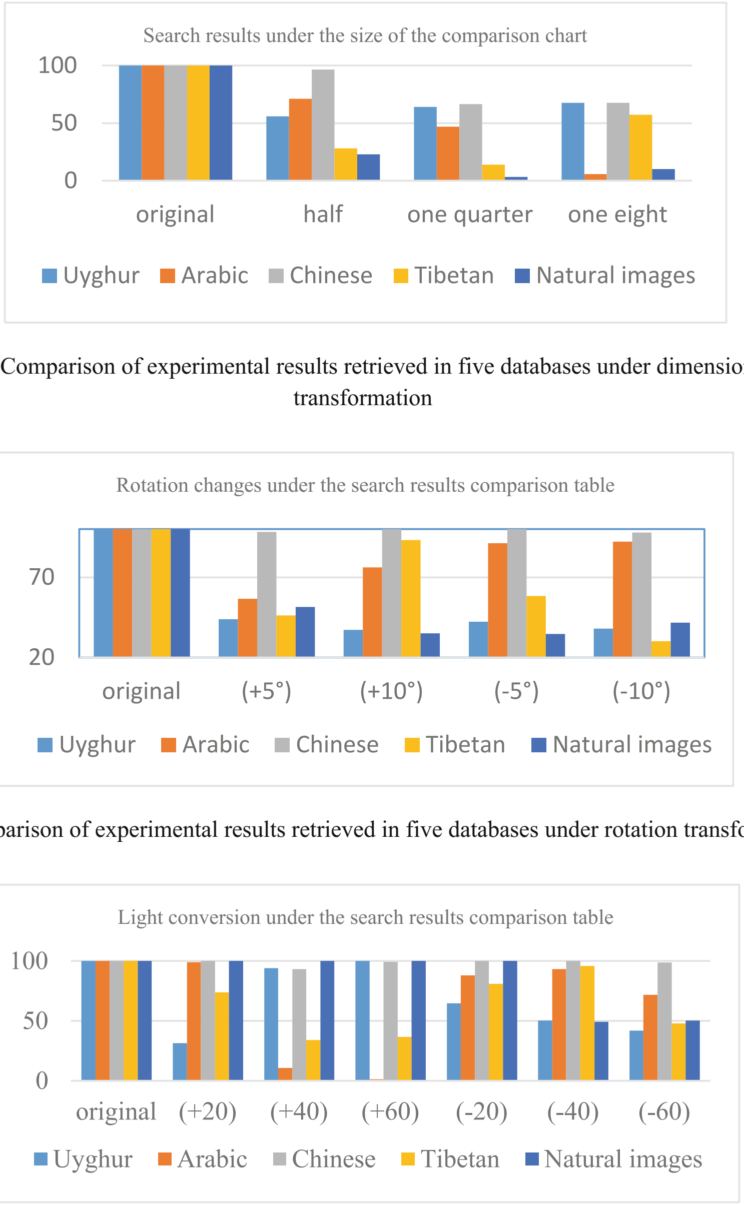

2013 年 7 月 14 E
前三名，以历史最好
IOI 2013 于 2013
手参加了本届比赛。
达和许昊然、东北师
副领队为中国人民大
النفاذ في الأول من كانو
من مهام الجات الإشر
ارة الدولية بين الدول ا
اء منظمة للتجارة الدولية
هافانا، إلا أن ذلك لاقى ،
هذه المنظمة، التي كان يا

(a) Chinese (b) Arabic (c) Tibetan (d) Natural images

Fig. 6. Comparative experimental database sample diagram

A number of examples of the experimental sample were transformed, such as size (2, 4, 8), illumination (20, 40, 60, −20, −40, −60), and rotation angle (5°, 10°, −5°, −10°) transformation, the retrieval results under different transformations are compared with the retrieval experiments of the Uyghur-cut complex document images, Validate the validity of the retrieval algorithm. The comparison result of the experimental results of retrieving the output target image is shown in Fig. 7(a) to (c).

As can be seen from Fig. 7, the letters of Uyghur, Arabic and Tibetan are more irregular than those of Chinese characters, and the differences in the gray-level values of the neighborhood pixels vary greatly. The Chinese language has horizontal and vertical Coherence; the difference in gray value is small. Therefore, the retrieval rate of Chinese query images after many transformations is larger than that of other databases. There are many transformations on the query, and the average indexing time for finding the target image based on the modified retrieval system is 0.013 (0.018), 0.041 (middle), 0.043 (hide) and 0.003 (natural) respectively. Compared with the average retrieval time of the retrieval system of the original features, it is 35.38 (original), 27.81 (a), 15.61 (middle), 16.05 (hide), 123.33 (natural) times. It can be seen that the retrieval system of FAST+SURF features makes it easy to find the target image quickly and accurately, which shows that this article proposes the effective and reliable method of improving ideas.



(a) Comparison of experimental results retrieved in five databases under dimensional transformation

(b) Comparison of experimental results retrieved in five databases under rotation transformation

(c) Comparison of experimental results retrieved in five databases under light conversion

Fig. 7. Comparison of modified FAST and SURF retrieval platform under various transformations experimental results comparison chart

6 Conclusion

In order to make up for the gap in Uyghur complex document image research, this paper proposes a document image retrieval method which is to match retrieval of printed Uyghur composite document images using SURF and the modified SURF features. It is combined the FAST corner detection and SURF description, and two kinds of matching of the selected 64-dimensional feature vectors are performed, and the matching ratio is compared under the condition of size, rotation and light conversion to analyze the performance of the two matching systems. In the end, two retrieval systems were proposed, that is, retrieval scheme based on multiple distance metrics and matching number. The original 100 document images, 1000 compressed images and 1000 document images are retrieved respectively. The matched number of searches takes more time than the distance-based search, but it has a good retrieval rate. Therefore, the focus of the further work is to reduce the retrieval time while ensuring the high retrieval rate of the system.

Acknowledgments. This work was supported by the National Natural Science Foundation of China (No. 61563052, 61363064, 61163028), and Hotan Normal University Scientific Research Plan Project (No. 1076515160).

References

1. Xiaoxiao, M.A., Gang, Y.U., Changchun, L.I.: A data processing algorithm for unmanned aerial vehicle images based on SURF and SVM. J. Henan Polytech. Univ. (2017)
2. Zhao, L.L., Geng, G.H., Kang, L.I., A-Jing, H.E.: Images matching algorithm based on SURF and fast approximate nearest neighbor search. Appl. Res. Comput. **30**(3), 921–923 (2013)
3. Cheon, S.H., Eom, I.K., Ha, S.W., Yong, H.M.: An enhanced SURF algorithm based on new interest point detection procedure and fast computation technique. J. R.-Time Image Process. 1–11 (2016)
4. Zhang, H.M., Yang, L., Li, M.L.: Improved SURF algorithm and its application in seabed relief image matching. **12**, 05017 (2017)
5. Chen, J., Han, X.: Image matching algorithm combining FAST-SURF and improved k-d tree nearest neighbor search. J. Xian Univ. Technol. (2016)
6. Luo, N., Sun, Q.S., Chen, Q., Ze-Xuan, J.I., Xia, D.S.: Image matching algorithm combining SURF feature point and DAISY descriptor. Comput. Sci. **41**, 286–290 (2014)
7. Wang, D., Yan, S., Ming, M.: A fast image matching algorithm based on improved SURF. In: Tenth International Conference on Computational Intelligence and Security, pp. 3643–3647 IEEE (2015)
8. Weisheng, A.N, Rangming, Y.U., Yuling, W.U.: Image registration algorithm based on FAST and SURF. Comput. Eng. (2015)
9. Dong, H., Han, D.Y.: Research of image matching algorithm based on SURF features. In: International Conference on Computer Science and Information Processing, pp. 581–584 IEEE (2012)
10. Batur, A.: Research on Uyghur printed complex document image retrieval based on local feature. Xinjiang University (2017)

11. Ren, K., Hu, M.: Color image registration algorithm based on improved SURF. J. Electron. Meas. Instrum. (2016)
12. Ma, Y.L.S.: Research on image based on improved SURF feature matching. In: Seventh International Symposium on Computational Intelligence and Design, pp. 581–584. IEEE (2015)
13. El-Gayar, M.M., Soliman, H., Meky, N.: A comparative study of image low level feature extraction algorithms. Egypt. Inform. J. **14**(2), 175–181 (2013)
14. Huang, L., Chen, C., Shen, H., He, B.: Adaptive registration algorithm of color images based on SURF. Measurement **66**, 118–124 (2015)
15. Zheng, C., Jin, W., Fang, F., Tang, C., Ling, Y.: Robust visual tracking algorithm based on structural multi-scale features adaptive fusion in co-training. In: International Conference on Information Science and Control Engineering, pp. 588–592. IEEE (2016)
16. Pandey, R.C., Singhm, S.K., Shukla, K.K., Agrawal, R.: Fast and robust passive copy-move forgery detection using SURF and SIFT image features. In: International Conference on Industrial and Information Systems, pp. 1–6. IEEE (2015)
17. Darve, N.R., Theng, D.P.: Image processing on eye image using SURF feature extraction. **3297**, 2738–2741 (2015)
18. Horak, K.: Classification of SURF image features by selected machine learning algorithms. In: International Conference on Telecommunications and Signal Processing, pp. 636–641 (2017)
19. Shanmugam, B., Rathinavel, R., Perumal, T., Subbaiyan, S.: An efficient perceptual of CBIR system using MIL-SVM classification and SURF feature extraction. Int. Arab. J. Inf. Technol. **14**(4), 428–435 (2017)