SIFT Feature Extraction Algorithm for Image in DCT Domain

Wu Zhen

National Computer Network and Information Security Administration Center,

Beijing 100029, China

Xu Zhe,Zhang Rui-nian,Li Shao-Mei

National Digital switching System

Engineering & Technological R&D Center,

Zhengzhou 450002, China

Abstract—**Image feature extraction is an important technology in image matching and retrieval. For the problem of high computational complexity of spatial domain image feature extraction using the SIFT algorithm, and by studying the relationship between DCT coefficient matrix and image, the paper designed the DCT coefficients reduced matrix of image and proposed the algorithm of SIFT feature extraction in DCT domain reduced image. Experiments showed that with the low loss of accuracy in image matching and retrieval, the method proposed can significantly improve the computational efficiency of feature extraction.**

Keywords-Image feature extraction; DCT coefficients reduced matrix; DCT domain reduced image; SIFT algorithm

I. INTRODUCTION

Image features can be divided into global and local features, global features mainly reflect the statistic information of the whole image while local features describe the details of the structure and texture of local areas of the image [1], and local features have a higher robustness and a broader application [2]. Existing studies have shown that in the local features SIFT has the highest accuracy [3] [4] [5], but it also has a high computational complexity and a serious consumption of the computing resources [6] [7].

Discrete Cosine Transform (DCT) has a good de-correlation and energy focused performance [8], and all international standards for image and video encoding now have used the DCT-based transform coding [9]. In the process of image compression, the input image is firstly decomposed into blocks of 8×8 or 16×16, then each image block is carried out with the two-dimensional DCT transformation, DCT coefficients quantization and encoding. The image restoration is an inverse process as mentioned above. This paper designed the DCT coefficients reduced matrix of image and proposed an algorithm of SIFT feature extraction in DCT domain, which reduced the number of SIFT key points and improved the efficiency.

II. DCT DOMAIN REDUCED IMAGE FEATURE EXTRACTION AND MATCHING BASED ON SIFT

A. DCT domain reduced image construction and SIFT feature extraction

DCT transformation is an important mathematical basis of data compression standards such as JPEG, MPEG. In the compression algorithm, after color space conversion, input images firstly is decomposed into 8x8 image blocks, then with two-dimensional DCT transformation each image block is converted into 64 DCT coefficient values and these values are quantized, lastly the compression task finishes with lossless encoding. Decompression is the inverse process as mentioned above.
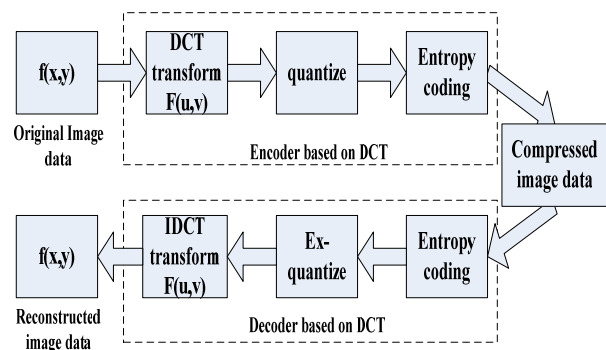

Figure 1. DCT-based compression and de-compression model

In DCT coefficients matrix, the main components of the spectral coefficients of image data which represent the low frequency part, locate in the upper left corner of the matrix. And the coefficients reflect high frequency part with small amplitude locate in the lower right corner. Since the main energy of the image is concentrated in the low frequency, we can discard some of the AC coefficient values which are close to or equal to 0. After reducing the DCT coefficients matrix in according with (1), we can get the DCT coefficients reduced matrix which is proposed in this paper.

$$D(u,v,s)=F(u,v) \qquad (u,v=1,...,s;1<s\leq 8) \qquad (1)$$

Where $F(u,v)$ is the DCT coefficient matrix of the image block, and $D(u,v,s)$ is the DCT coefficients reduced matrix, s is the reduction scale. As shown in Fig.2.

Because the DC coefficients between the adjacent image blocks of the DCT domain coefficients reduced image is close, and the AC coefficients are relatively small, the DCT domain coefficients reduced image only has the basic outline of the spatial domain image. As shown in Fig.3.

(a)DCT coefficients reduced process

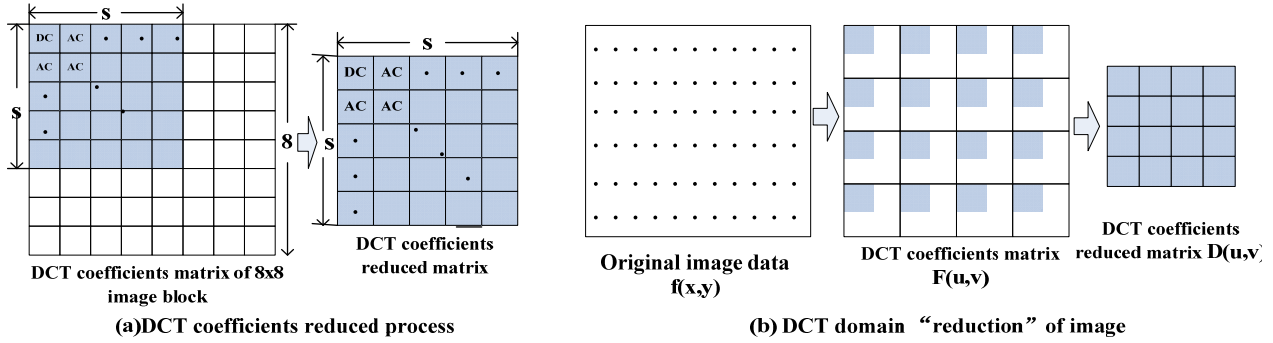**(b) DCT domain "reduction" of image**

Figure 2.   Reduced Image DCT coefficients matrix

As the correlation between the adjacent 8x8 image blocks of the image transformed by DCT is strong, it is necessary to smooth the image block using the DC coefficient values, which aims to spread the low-frequency energy evenly in the matrix. The smoothed DCT domain coefficients reduced image is called the DCT domain reduced image in this paper.

$$D(u,v,s)=\begin{cases} F(u,v)/8 & (u,v{=}1) \\ F(u,v){+}F(1,1)/8 & (u,v{=}2,...,s) \end{cases} \quad (1{<}s{\le}8) \quad (2)$$

Construct the DCT domain reduced image according to (2). In Fig.3(c) we can find that the DCT domain reduced image and the spatial domain image is almost same, with a few small differences in the image quality, such as the screen size and the image details. As shown in Fig.3 (e), this paper focuses on designing SIFT feature extraction algorithm for DCT domain reduced image.
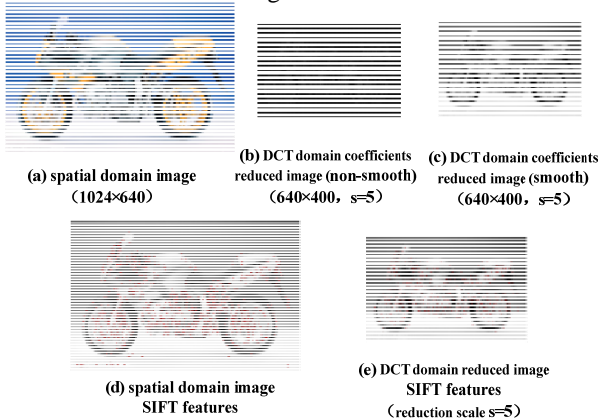

(a) spatial domain image
(1024×640)

(b) DCT domain coefficients reduced image (non-smooth) (640×400，s=5)

(c) DCT domain coefficients reduced image (smooth) (640×400，s=5)

(d) spatial domain image SIFT features

(e) DCT domain reduced image SIFT features （reduction scale s=5）

Figure 3.   DCT domain image construction and SIFT feature extraction

B. Performance parameters

In order to compare the performance of the SIFT feature between the DCT domain image and the spatial one. This paper defines the following three performance parameters.

1) *Matching degree.* Assuming that the two images are A and B, their SIFT vector number are respectively $KeyNum_A$ and $KeyNum_B$, the matching feature number is $KeyMatch_{AB}$, then *Matching degree* θ of the two images is

$$\theta = \frac{KeyMatch_{AB}}{\min(KeyNum_A, KeyNum_B)} \times 100\% \quad (3)$$

2) *DCT-space fitting degree.* Assuming that the *matching degree* of DCT domain and spatial domain are respectively θ_{dct} and θ_{space}, then *DCT-space fitting degree* λ is

$$\lambda = \frac{\theta_{dct}}{\theta_{spac}} \times 100\% \quad (4)$$

If $\lambda > 1$, the effectiveness of the former is better than the latter, otherwise in the case of $\lambda < 1$, the effectiveness of the latter is better than the former.

3)   *Efficiency upgrade degree.* If the processing times to extract the SIFT feature in DCT domain and spatial domain are respectively T_{dct} and T_{space}, then the *efficiency upgrade degree* of DCT domain η is:

$$\eta = \frac{T_{space} - T_{dct}}{T_{dct}} \times 100\% \quad (5)$$

If $\eta > 0$, the method based on DCT domain has a shorter processing time and a higher efficiency than the spatial one, moreover, the larger the value of η is, the higher the *efficiency upgrade degree* is. If $-1 < \eta < 0$, it is opposite.

III. EXPERIMENTS AND ANALYSIS

Two types of experiments, image matching and image retrieval, are proposed in this paper to test the performance of DCT domain SIFT feature. In the image matching experiment, experimental data come from the Mikolajczyk[3] data set, while in the image retrieval experiment, we used UCID data set. Experiments performed on machine with Intel 8-core CPU with 2.33GHz and 2.93GB memory,

TABLE I.   DCT DOMAIN AND SPATIAL DOMAIN IMAGE SIFT FEATURE EXTRACTION AND MATCHING METHOD

COMPARISON OF MIKOLAJCZYK DATA SETS MATCHING RESULTS

	Average matching degree (%) (DCT domain/spatial domain)	Average fitting degree (%) (DCT domain/spatial domain)	Average processing time(s) (DCT domain/spatial domain)	Average efficiency upgrades degree (%) (DCT domain/spatial domain)
Graffiti	25.38/35.42	72.33	0.2888/2.5014	610.28
Wall	25.59/34.74	74.69	0.2737/2.1655	683.40
Boats	24.33/29.32	82.31	0.2385/1.7612	635.99
Barks	19.32/27.25	73.97	0.2014/2.1827	894.77
Bikes	55.14/61.17	93.05	0.3175/1.8074	466.28
Trees	26.46/19.41	144.78	0.4422/4.7998	920.39
Leuven	44.96/62.58	72.43	0.2102/1.0663	407.44
UBC	45.95/31.74	169.40	0.2591/2.0205	688.34

and the programs coded with Matlab and C mixed language.

A. Reduction scale factor determination

In the DCT coefficients reduced matrix and DCT domain reduced image which are proposed in this paper, reduction scale s is an important parameter. In order to determine the optimal s value, we conduct the following experiments.

Select eight images from each of the two data sets above, test the *DCT-space fitting degree* on the total 16 images, and calculate their *efficiency upgrade degree*. The average *fitting degree* and *efficiency upgrade degree* are shown in Fig.4.
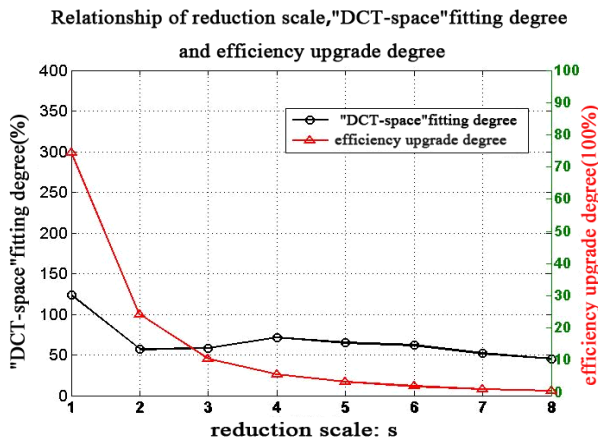

Figure 1.   Determination of the reduction scale S.

In Fig.4, *efficiency upgrade degree* constantly decreases as s increases, but *fitting degree* is relatively stable. When $s =1$, we can get a high *fitting degree*, with the *efficiency upgrade degree.* So $s =1$ is set in the following experiments in this paper.

B. Image matching experiments

Mikolajczyk data set comprises eight image subsets of different scene types with different geometric transformations and brightness conversions. Each subset contains six images, and the corresponding six transformations are: View, Scale, Rotation, Blur, Illumination and JEPG Compression.

In the first experiment, firstly, extract the DCT domain and spatial domain SIFT feature for each two image samples of each image data subset, then match them. It means this process was repeated 15 times, and respectively compute the matching degree and the processing time after each time, then get the average values of these 15 times. The results are shown in Table 1.

In Table 1, the matching fitting degree of the two methods are both higher than 72% in the image matching process, of which image subset of Trees and UBC are both more than 100%, but the average *efficiency upgrade degree* using DCT domain method is 4 to 9 times higher than the spatial domain one.

C. Image retrieval experiments

The UCID dataset contains a total of 1338 different content types of images. Firstly, 20 sets of images were sampled from the dataset, each consists of three images taken from different angles on a same content type, so the total image number is 60. Then extract one image from each set to generate a total of 20 images as query image and, the remaining 40 images as the database image. The experimental results are shown in Table 2.

TABLE II.   COMPARISON OF RETRIEVAL RESULTS BETWEEN DCT DOMAIN AND SPATIAL DOMAIN SIFT FEATURE EXTRACTION MATCHING

	DCT domain	Spatial domain
Correctly retrieved number(first 3)	13/11/4	13/13/2
Correct retrieval rate(%)	65	67
Time consumed(s)	5.74	74.23

In Table 2, correct retrieval rate of the DCT domain method was 2% which is lower than that of the space domain method, but the processing time is 12.93 times shorter than the spatial one. It is because the number of SIFT key points for DCT domain method is much less than the spatial domain, resulting in a much smaller amount of computation in feature extraction and even in feature retrieval and

matching. Thus, with a slight reduction of retrieval accuracy, method using DCT domain SIFT feature can significantly improve the retrieval speed and efficiency.

IV. CONCLUSIONS

In this paper, we present a DCT domain SIFT feature extraction algorithm. The algorithm firstly constructs a DCT coefficients reduced matrix, and then smooths the DC coefficients to generate the DCT domain coefficients reduced image, based on which SIFT feature is extracted. The experimental results show that with low loss of accuracy, the DCT domain SIFT feature extraction algorithm can significantly improve the computational efficiency of feature extraction, and will have great practical value in the condition of high real-time requirements.

ACKNOWLEDGEMENT

REFERENCES

[1]. Li Shanshan. Research of feature design and similarity measurement in computer vision[D]. PhD thesis, University of Science and Technology of China, 2010.

[2]. Cheng Lei. Target Recognition Method Based on Structure of Local Feature [D], University of Science and Technology of China, 2009.

[3]. Mikolajczyk K,Schmid C.A Performance Evaluation of Local Descriptors[J].IEEE Trans.on Pat.Analysis and Machine Intelligence,2005,27(10):1615-1630.

[4]. Mikolajczyk K, Tuytelaars T, Schmid C, et al. A comparsion of affine region detectors [J]. International Journal of Computer Vision, 2005, 65(1):43-72.

[5]. Douze M, Jegou H, Schmid C. An image-based approach to video copy detection with spatio-temporal post-filtering [J].IEEE Transactions on Multimedia, 2008, 12(4):257-266.

[6]. Wang Jinde, Li Xiaoyan, Shou Lidan, Chen Gang. A SIFT Pruning Algorithm for Efficient Near-Duplicate Image Matching [J]. Journal of computer-aided design & computer graphics, 2010,22 (6), 1042-1049.

[7]. Zheng Yongbin, Huang Xinsheng, Feng Songjiang. An Image Matching Algorithm Based on Combination of SIFT and the Rotation Invariant LBP [J]. Journal of computer-aided design & computer graphics, 2010, 22 (2): 286-291.

[8]. Wan Yuan,Wu Chuansheng.Approach of MPEG- 4 video based on DCTQ module. Computer Engineering and Applications, 2007, 43(12): 42- 44.

[9]. Zeng Hui, Mu Zhichun, Wang Xiu-qing. A Robust Method for Local Image Feature Region Description [J]. Acta automatica Sinica, 2011, 37(6): 658-644.