SHOT CUT DETECTION WITH SLIDING CORRELATION IN DCT DOMAIN

Jianfeng Wang, Xiaorong Zhao

ChongQing Aerospace Polytechnic, China,400021

ARTICLE DETAILS

ABSTRACT

In this paper, we propose a shot cut detection algorithm in JPEG compressed domain. We show that, with the same content-based shot cut detection algorithm, adding sliding correlation to the existing operational process can improve the performance of content-based shot cut detection significantly. In comparison with the latest research efforts in shot cut detection, the experimental results show the superiority of the proposed method in terms of precision and processing speed.

KEYWORDS

1. INTRODUCTION

Numerous techniques have been developed in image retrveval over the past decades, providing successful solutions for a range of applications in image processing and computer vision. Via extensive literature survey, it is revealed that existing research on image segmentation can be mainly divided into two categories: edge-based approaches and region-based approaches, and all the retrieval techniques are developed in pixel domain [1-6]. A researcher proposed a texture-based feature extraction technique via multi-resolution nature of wavelets [7]. A researcher used a statistical structuring approach to organize images inside a database for content-based shot cut detection. Based on a research, a region-based shot cut detection technique, and developed DCT-based shot cut detection algorithms by exploiting the relationship between pixel domain and DCT domain [8-9]. With the growing demand for high resolution images and automatic archiving for digitized images and videos, effective and efficient image content similarity measurements are becoming increasingly important for researchers across a range of subject areas, including digital library, image processing, computer vision, and information systems [10,11].

2. THE PROPOSED ALGORITHM DESIGN

2.1 Related Work

We use our previous work as the basic research platform and a benchmark to start our investigation. This algorithm extracts statistics features directly in compressed domain by calculating the mean and variance form each block of DCT coefficients to construct a feature vector with 42 elements. Such a feature vector is then used as an indexing key to address all images inside a database When a query image comes in, the algorithm constructs the same content feature vector and search for the best match in terms of distance measurement inside the database. Like many other algorithms based on low-level content features, there exists significant discrepancy between the content feature vector match and the actual visual content match. Figure-l-(a) illustrates one typical example, from which it can be seen that two retrieved images have obvious different visual content in comparison with that of the query image (located at the top left corner).

2.2 defined of DCT transform

A 2D DCT defined in a block of NxN pixels can be expressed as

$$F(u,v) = (\frac{2}{N})\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}A(i)A(j)\cos[\frac{\pi \cdot u}{2N}(2i+1)]\cos[\frac{\pi \cdot v}{2N}(2j+1)]f(i,j) \quad (1)$$

Its inverse transform is:

$$f(i,j) = (\frac{2}{N})\sum_{i=0}^{N-1}\sum_{j=0}^{M-1}A(i)A(j)\cos[\frac{\pi \cdot u}{2N}(2i+1)]\cos[\frac{\pi \cdot v}{2N}(2j+1)]F(u,v) \quad (2)$$

2.3 Sliding Correlation

To improve its content match performance and judge the different between current frame and next frame, we introduce a rotational correlation test into the existing distance-based similarity measurement. Give two content feature vectors constructed by the below histogram,

$Q = \{q_1, q_2. . . . . .q_{35}\}$ and

$I_k = \{i_1^k, i_2^k . . . . . .i_{35}^k\}$ k∈[1,N],we propose a rotational correlation value as follows to test the level of correlation between these two feature vectors:

$$R_{Q,I}(\tau) = \frac{\sum_{j=1}^{35} q_j i^k \bmod_{35}(j-\tau)}{\sqrt{\sum_{j=1}^{35}(q_j)^2 \sum_{j=1}^{35}(i_j)^2}} \quad (3)$$

Where $Q = \{q_1, q_2. . . . . .q\}$ is the feature vector for the N frame, $I_k = \{i_1^k, i_2^k . . . . . .i_{35}^k\}$ is the feature vector for the N-1 frame, and t∈[0,35] indicates the left-rotation displacement of all elements inside the feature vector Ik . A complete left-rotation process for feature vector Ik is illustrated below:

$$I_k = \{i_1^k, i_2^k \qquad .i_{35}^k\}$$
$$I_{k-1} = \{i_2^k, i_3^k \qquad .i_{35}^k, i_1^k\} \quad (4)$$

… … …

$$I_{k-34} = \{i_{35}^k, i_1^k, i_2^k \qquad .i_{34}^k\}$$

The modular operation inside equation (4) can be further described as:

$$\bmod_{35}(j-\tau) = \begin{cases} j-\tau \\ j-\tau+35 \end{cases} \quad (5)$$

Before applying distance measurement, we propose the following correlation test between the N frame and N-1 frame, Next we compare the

r_k with the threshold T

$$r_k = \max(R_{Q,I}(\tau)) \quad (6)$$

Following that, all images are arranged in a descending order according to their correlation test value rk . As a result, by introducing the proposed correlation test, we have two similarity measurements to achieve visual content match for shot cut detection. One is based on distance measurement and the other on correlation test. To this end, the shot cut detection algorithm can be designed with two options. The first option is correlation test first followed by distance match to retrieve the closest images. In this option, we can select the first M images out of the image list produced by the correlation test to do the distance-based search and retrieval. In this way, distance search can be carried out among these M images only, which are highly correlated with the query image proved by the proposed correlation test. In this option, we essentially pre-screen all images by their distance match, and use correlation test to select the final images to be retrieved. Specific operation is similar to the first one in the sense that images are firstly arranged into an ascending order in terms of their distance match with the query image. We then select M images out of the list to do the correlation test and further arrange them into a descending order before final images are retrieved from the top of this list.

3. ALGORITHM EVALUATIONS VIA EXPERIMENTS

Along with the above design, the proposed algorithm was implemented in MATLAB as an off-line system. To evaluate the performance of the proposed idea, experiments were carried out using a database of 50000 images with various sizes. The database includes natural scenes, animals, flowers, people, cars etc. The variety of images in the database prevents any bias toward a particular type of images.

In general, there are two main factors to evaluate the performance of a CBIR (Content-Based Shot cut detection) system. One is retrieval efficiency, which is measured by retrieving speed, and the other is the effectiveness, which emphasizes the accuracy of the retrieval. To assess the performance of the proposed algorithm, we design a query image set with ground truth inside the database as follows.

Out of the 50000 images inside the database, we manually select 5000 images and divide them into 5 categories, containing 1000 image with human face indexed as category-1;1000images with trees indexed as category-3; 1000 images with outdoor scenes indexed as category-4; and 1000images with cars indexed as category-5. For the convenience of description, we refer these 1000 images as a query data set.

Table 1: Experimental Results for Retrieving Effectiveness

	Car(EA1)	People(EA2)	Flower(EA3)	Scene(EA4)	Animal(EA5)	ASP
Benchmark	0.66	0.65	0.59	0.75	0.67	0.66
Option-1 (M=72)	0.75	0.77	0.91	0.84	0.85	0.82
Option-2 (M=48)	0.81	0.79	0.88	0.89	0.82	0.84

Main code of algorithm

```
for(k=1:10000)
    s=num2str(k);
    s=strcat(s,'.jpg');
    A=imread(s);

 for(i=1:10000)
   result1(1,i)=jieguo(1,i);
end
for (i=1:10000)
   for(j=1:10000-i)
   if (jieguo(1,j)>jieguo(1,j+1))
     temp=jieguo(1,j);
     jieguo(1,j)=jieguo(1,j+1);
     jieguo(1,j+1)=temp;
   end,
    end
end

for (i=1:10000)
    if (jieguo(1,1)==result1(1,i))
    s=num2str(i);
    s=strcat(s,'.jpg');
    subplot(2,6,1);
    imshow(s);
    end,
end
```

4. CONCLUSIONS

We propose a new similarity measurement by introducing a rotational correlation test into the conventional distance matching mechanism for content-based shot cut detection and illustrate that such proposed scheme achieves significant performance improvement benchmarked by its distance match only counterpart.

ACKNOWLEDGEMENT

REFERENCES

[1] Fan, J.P., Gaoa, Y.L., Luoa, H.Z., Xub, G.Y. 2015. Statistical modeling and conceptualization of natural images, Patten Recognition 38, 865-885

[2] Chul Ko, B.C., Byun, H.R. 2014. FRIP: A Region-Based Shot cut detection Tool Using Automatic Image Segmentation and Stepwise Boolean and Matching, IEEE Transactions on Multimedia, 7, (1).

[3] Jiang, J.M., Feng, G.C. 2014. The Spatial Relationship of DCT Coefficients Between a Balock and Its Sub-block, IEEE Transactions on Signal Processing, 50, (5).

[4] Erol, B., Kossentini, F. 2013. Shape-Based Retrieval of Video Objects, IEEE Transactions on Multimedia, 7, (1).

[5] Faloutsos, C., Barber, R., Flickner, M. 2010. Efficient and effective querying by image content, Journal of Intelligent Information System 3, (3) 231-262.

[6] Jing, F., Li, M., Zhang, H.J., Zhang, B. 2005. A Unified Framework for Shot cut detection Using Keyword and Visual Features, IEEE Transactions on Image Processing, 14, (7) 979-989;

[7] Vasconcelos, N., Lippman, A. 2013. A Multiresolution Manifold Distance for Invariant Image Similarity, IEEE Transactions on Multimedia, 7, (1), 127-142.

[8] Kourosh Jafari-Khouzani, K., Soltanian-Zadeh, H. 2005. Rotation-Invariant Multiresolution Texture Analysis Using Radon and Wavelet Transforms, IEEE Transaction on Image Processing. 14, (6).

[9] Soledad Torres-Guijarro, S., A. Beracoechea-A'lava, J., Luis I. Ortiz-Berenguer, L.I., Javier Casaju's-Quiro's, F. 2005. Inter-Channel de-correlation for perceptual audio coding, Applied Acoustics 66, 889-901.

[10] Yu, Z.C. 2017. An Improved Infrared and Visible Image Fusion Algorithm Based on Curvelet Transform. Acta Informatica Malaysia, 1 (1), 36-38.

[11] Tu, C.X. 2017. Research on the Multilevel Security Authorization Method Based on Image Content. Acta Informatica Malaysia, 1 (2), 17-19.