동영상 복사본 검출을 위한 시공간 정보를 이용한 동영상 서명

동심원 구획 기반 서술자를 이용한 동영상 복사본 검출 기술

Video Signature using Spatio-Temporal Information for Video Copy Detection

조익환,Ik-Hwan Cho*, 오원근, Weon-Geun Oh**, 정동석, Dong-Seok Jeong***

요약 본 논문에서는 동영상 복사본 검출을 위해 시공간 정보를 이용한 새로운 동영상 서명 방법을 제안한다. 제안된 동영상 복사본 검출 방법은 각 키프레임에 대하여 동심원 구획 방법에 기반을 두고 있다. 우선 입력 동영상으로부터 일정한 간격으로 시간적 이중 선형 보간법을 이용하여 키프레임을 추출하고 각 프레임은 동심원 형태로 구획된다. 구획된 세부 영역에 대하여 상대적인 관계를 이용하여 평균 픽셀값, 평균 픽셀값의 차분값, 대칭적 차분값, 원형 차분값 분포의 4 가지 특징 분포를 추출한다. 최종적으로 이 특징 분포들은 간단한 해시 함수를 이용하여 이진 수열 형태로 변형되고 순서대로 병합된다. 제안된 동영상 서명에 대한 유사도 거리는 간단한 해밍 거리를 이용하여 계산되고 따라서 매우 빠른 정합 속도를 보인다. 실험 결과로부터 제안된 방법이 다양한 변형에 대하여 평균 97.4%의 높은 검출 성공률을 보이는 것을 알 수 있다. 그러므로 제안된 방법이 동영상 복사본 검출분야에서 폭넓게 사용될 수 있을 것으로 기대된다.

Abstract This paper proposes new video signature using spatio-temporal information for copy detection. The proposed video copy detection method is based on concentric circle partitioning method for each key frame. Firstly, key frames are extracted from whole video using temporal bilinear interpolation periodically and each frame is partitioned as a shape of concentric circle. For the partitioned sub-regions, 4 feature distributions of average intensity, its difference, symmetric difference and circular difference distributions are obtained by using the relation between the sub-regions. Finally these feature distributions are converted into binary signature by using simple hash function and merged together. For the proposed video signature, the similarity distance is calculated by simple Hamming distance so that its matching speed is very fast. From experiment results, the proposed method shows high detection success ratio of average 97.4% for various modifications. Therefore it is expected that the proposed method can be utilized for video copy detection widely.

핵심어: video signature, concentric circle partitioning, copy detection, duplicate detection

*주저자 : 인하대학교 전자공학과 박사과정 e-mail: teddydino@inhaian.net

**공동저자 : 한국전자통신연구원 연구원 e-mail: owg@etri.re.kr

***교신저자 : 인하대학교 전자공학과 교수; e-mail: dsjeong@inha.ac.kr

1. Introduction

As the distribution of image and video through the internet increase gradually, multimedia data management technology has been requested. Since digital data can easily be copied, illegal copies of multimedia data have been made recently. Therefore, new technology for automatic detection of copies multimedia video has been in necessity and the various related research has been conducted up to date [1-3]. In this manuscript, we propose a new video copy detection algorithm, which is derived from an image copy detection algorithm to be discussed by extending it to video. Image copy detection algorithm partitions image into concentric circle shape and describe unique pattern of image by using spatial relative relation among partitioned sub-regions. In order to extend image-based method to video, the proposed method adds periodic key-frame extraction method.

In section 2, we describe image-based signature in detail and explain the additional methods for extension to video in section 3. And experiment results are depicted in section 4 and we concluded in section 5.

2. Image signature for copy detection

First step of the proposed image-based signature extraction is to partition image into concentric circle using polar coordinate conversion. And from the partitioned sub-regions, 4 feature distributions to describe spatial relation are generated and the extracted distributions are converted into binary string using Hash function. Finally the converted binary strings are concatenated sequentially. The detail explanations are depicted as following sub-sections.

2.1 Concentric circle partition

Input images are resized into fixed resolution using bilinear interpolation and images are partitioned into concentric circle shape. For the partitioning of image, image pixels on Cartesian coordinate are converted into polar coordinate by using Eq. (1). And for converted image, it is partitioned into sub-regions in according to various radius and angles. Figure 1 shows example of image partitioned as concentric circle shape.

$$x = r\cos\theta$$
$$y = r\sin\theta$$
$$r = \sqrt{x^2 + y^2}$$
$$\theta = \begin{cases} \arctan\left(\frac{y}{x}\right) & \text{if x>0 and } y \geq 0 \\ \arctan\left(\frac{y}{x}\right) + 2\pi & \text{if x>0 and } y < 0 \\ \arctan\left(\frac{y}{x}\right) + \pi & \text{if x<0} \\ \frac{\pi}{2} & \text{if x=0 and } y > 0 \\ \frac{3\pi}{2} & \text{if x=0 and } y < 0 \end{cases} \quad (1)$$

Figure 1. Example of concentric circle partition

2.2 Extraction of feature distribution

For partitioned sub-regions, 4 feature distributions are generated by using spatial relative relation among them.

2.2.1 Average intensity distribution

First feature of the proposed method is average intensity distribution through various radius level (Eq.1).

$$C_i = \frac{1}{angleLevel} \sum_{j=0}^{angleLevel} P_{i,j} \quad (1)$$

$P_{i,j}$ is all pixel value in ith radius level and jth angle level.

2.2.2 Intensity difference distribution

Second feature is intensity difference distribution through various radius levels and the difference value between neighborhood ring regions is calculated with first average intensity distribution (Eq. 2).

$$V_i = abs(C_{i+1} - C_i) \quad (2)$$

C_i is all average intensity level in i*th* radius level.

2.2.3 Average symmetric difference distribution

For each ring region, average symmetrical difference is calculated by summing all difference values between sub-regions located in symmetrical position each other (Eq.3). Figure 2(a) shows the example of difference between symmetrical regions.

$$S_i = \frac{1}{N_a/2} \sum_{j=0}^{N_a/2-1} abs\left(C_{i,j} - C_{i,j+N_a/2}\right) \quad (3)$$

$C_{i,j}$ is all average intensity level in ith radius level and jth angle level and N_a is angle levels.

2.2.4 Average circular difference distribution

Circular difference value on specific ring region is obtained by calculating average difference between neighborhood sub-regions into circular direction (Eq.4). The example of difference between neighborhood regions on circular direction is depicted in Figure 2(b).

$$R_i = \frac{1}{N_a} \sum_{j=0}^{N_a-1} abs\left(C_{i,j} - C_{i,(j+1) \bmod N_a}\right) \quad (4)$$

$C_{i,j}$ is all average intensity level in ith radius level and jth angle level and N_a is angle levels.

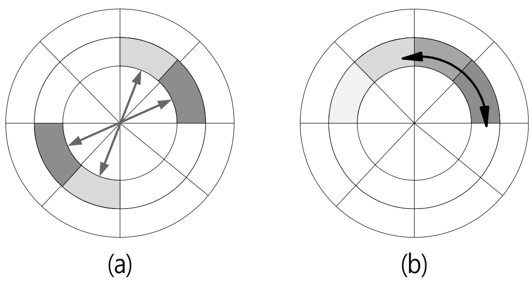

Figure 2. (a) Difference between symmetrical regions in one circle, (b) difference between neighborhood regions on circular direction.

2.2.5 Binary conversion

4 kinds of scalar distributions are converted into binary string by applying hash function of Eq. 5.

$$B_i = \begin{cases} 1, & M_{i+1} > M_i \\ 0, & M_{i+1} \le M_i \end{cases} \quad (5)$$

M_i is distribution value in i*th* index.

3. The proposed video signature

Video signature can be derived by extending image signature to video. For extension of image signature, periodic key frames are selected and for each key frame, binary shaped image signature is extracted. And final video signature is obtained by concatenating individual signatures. Figure 3 shows overall process for extraction of the proposed video signature.

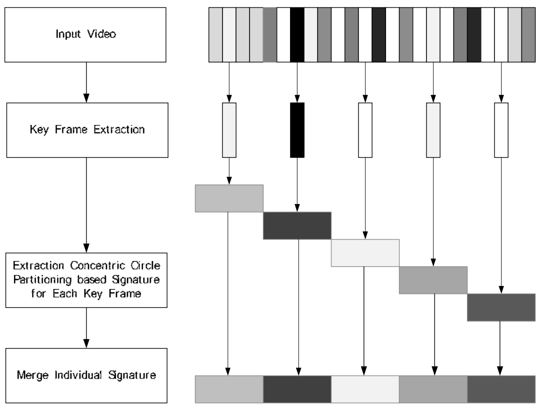

Figure 3. The proposed video signature extraction process.

3.1 Key frame selection

In video signal processing, key frame selection is one of most important research topics. Roughly key frame selection methods are separated into scene change-based method and periodic selection method. The former is very popular method for video sequence analysis such as video retrieval and indexing. In general, video sequence has several scenes and frames in each scene are very similar each other so that it is possible to represent whole scene with only one frame in each scene as the representative image. However it is impossible to get precise temporal information from key frame in specific scene. So in this paper, we select key frames periodically from input video sequence in order to get accurate time information. And from the selected key frames, the proposed image signature is extracted individually. At last the final video descriptor is obtained by merging each image signature together.

3.2 Matching of video signature

Matching of the proposed video signature is carried out frame-by-frame Hamming distance (Eq. 6) [4]. And firstly individual distance is calculated by shifting overlapped range and then final distance is obtained by selecting minimum value of them.

$$D = \frac{1}{N}\sum_{j=0}^{N-1} R_j \otimes Q_j \qquad (6)$$

N is the number of bits of descriptor and R_j and Q_j is bit of reference and query in jth index.

4. Experiment results

For the performance test of the proposed method, we used totally 130 music video clips with large motion and frequent scene changes. And for these video sequences, 13 kinds of modifications are applied and detection success ratio for modified video query is used as final performance measure. For the performance evaluation, detection success ratio is used and it is defined as Eq. (7).

$$\text{Detection success ratio} = \frac{K}{M} \qquad (2)$$

M is the numbe of original sequences and K is the number of sequences retrieved correctly.

For correct retrieval, the distance between query and its original should be smaller than any other pair and it should be possible to find out accurate position of query sequence in original one (Figure 4 **오류! 참조 원본을 찾을 수 없습니다.**).

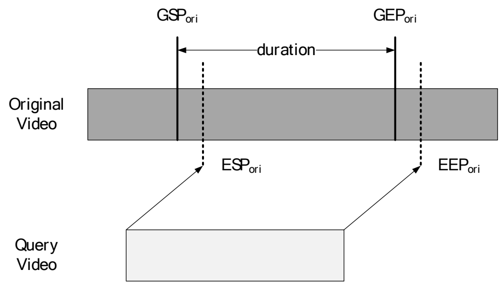

Figure 4. Detection of the matched section between query and original clips.

Table 1 shows several parameters used in this paper.

Table 1. Parameters for the proposed video signature

Parameters	Values
Key frame interval	100 msec
Max radius of circle	256
Radius levels	32, 16
Angle levels	36

Table 2 represents detection success ratio for various modifications of the proposed method. The proposed method shows high accuracy for various modifications and the average detection accuracy is about 97.4%.

Table 2. Detection success ratio for various modifications

Modifications	Detection Success Ratio
Bitrate conversion (150kbps)	0.992
Bitrate conversion (300kbps)	0.984
Frame conversion (10 fps)	0.977
Frame conversion (15 fps)	0.984
Brightness change (+30%)	0.984
Brightness change (−30%)	0.969
Contrast change (+30%)	0.984
Contrast change (−30%)	0.977
Temporal cutting (60sec)	1.000
Temporal cutting (30sec)	0.992
Resize (CIF)	0.907
Resize (QCIF)	0.953
Image logo embedding (transparent 80%)	0.953
Average	0.974

5. Conclusion

In this paper, we proposed a new video signature for video copy detection. The proposed method is originally based on image signature. And basic image signature uses concentric circle partitioning method and is made by describing spatial relation among partitioned regions. Final signature has binary string types so that it can make high speed matching. Periodic key frame selection method makes it possible to extend image copy detection method to video. From the experiment results, the proposed method shows average 97.4% robustness for various modifications. Therefore it is expected that the proposed video signature is very useful for video copy detection.

Acknowledgment The presented research is supported by Research Grant of Electronics and Telecommunication Research Institute (ETRI).

References

[1] L. Zhenyan and T. Yap-Peng, "Content-based video copy detection with video signature," presented at Circuits and Systems, 2006. ISCAS 2006. Proceedings. 2006 IEEE International Symposium on, 2006.

[2] J. Law-To, L. Chen, A. Joly, I. Laptev, O. Buisson, V. Gouet-Brunet, N. Boujemaa, and F. Stentiford, "Video copy detection: a comparative study," Proceedings of the 6th ACM international conference on Image and video retrieval, pp. 371-378, 2007.

[3] K. Iwamoto, E. Kasutani, and A. Yamada, "Image Signature Robust to Caption Superimposition for Video Sequence Identification," presented at Image Processing, 2006 IEEE International Conference on, 2006.

[4] R. W. Hamming, "Error detecting and error correcting codes," Bell Syst. Tech. J, vol. 29, pp. 147-160, 1950.