§34. Improvement of Similarity Retrieval in Fusion Experiment Multimedia Data Archive

Hochin, T., Nomiya, H. (Kyoto Inst. of Tech.),
Nakanishi, H., Nagayama, Y., Emoto, M.,
Ohsuna, M., Ito, T.

Experiments of the fusion phenomena produce a lot of sequences of time-varying values which form waveforms. If the waveforms similar to a desired one can be obtained by using computer system, the burden of researchers in searching similar waveforms will extremely be decreased. We have addressed to the issue on this kind of retrieval. The method using three dimensional (3D) Fourier transformation has been developed for the similarity retrieval of movies of plasma discharges, which are called plasma movies [1]. The method using 3D Fourier transformation, however, has the problem that the processing cost is very high. This paper examines the method not using it.

This paper uses the frame hashing method [2]. Frames are divided into $n \times m$ blocks. Luminosity of each block is averaged in a block. The average value is binarized by the mean luminosity of the frame. In this paper, instead of a frame, we use the mean luminosity of a video. After that, $n \cdot m$ bit binary digits can be obtained for each frame. This is the hash value of a frame. A series of frames having similar hash values are gathered into a shot.

First, the shots whose hash values are equal to zero, and the shots which are shorter than the minimum shot length, which is five here, are removed. Next, the shots are sorted in ascending order of the shot length. The weight of a shot s in a video v is calculated by the equation: $W_{\mathrm{shot}}(s) = rank(s)/n(v)$, where $rank(s)$ is the rank of the shot s in the shots of the video v, and $n(v)$ is the number of shots in the video v. The similarity of shots $s1$ and $s2$ based on hash values is calculated by the equation: $Sim_{\mathrm{H}}(s1, s2) = 1 - (\mathrm{XOR}(H(s1), H(s2))/n\_blocks)$, where $H(s)$ is the hash value of a shot s, XOR is the exclusive OR operation, and n_blocks is the number of blocks in a frame. The weight of the difference of the lengths of shots is calculated by the equation: $W_{\mathrm{Diff_{Len}}}(s1, s2) = min(l(s1), l(s2))/max(l(s1), l(s2))$, where $l(s)$ is the length of a shot s. When this value is less than 0.3, zero is set to it. The similarity of luminosity of shots $s1$ and $s2$ is calculated by the equation: $Sim_{\mathrm{Lum}}(s1, s2) = min(lum(s1), lum(s2))/max(lum(s1), lum(s2))$, where $lum(s)$ is the luminosity of a shot s. When this value is less than 0.4, it is made zero. The similarity of videos $v1$ and $v2$ based on hash values is calculated by the equation: $Sim_{\mathrm{HVideo}}(v1, v2) = \Sigma_{n=1}^{N}\Sigma_{m=1}^{M}(Sim_{\mathrm{H}}(m, n) \times W_{\mathrm{Diff_{Len}}}(m, n) \times (W_{\mathrm{shot}}(m) + W_{\mathrm{shot}}(n)))/M \cdot N$, where M and N are the numbers of shots in videos $v1$ and $v2$, respectively. The similarity of videos $v1$ and $v2$ based on luminosity is calculated by the equation: $Sim_{\mathrm{LumVideo}}(v1, v2) = \Sigma_{n=1}^{N}\Sigma_{m=1}^{M}(Sim_{\mathrm{Lum}}(m, n) \times W_{\mathrm{Diff_{Len}}}(m, n) \times (W_{\mathrm{shot}}(m) + W_{\mathrm{shot}}(n)))/M \cdot N$. The similarity of videos $v1$ and $v2$ based on the number of shots is calculated by the equation: $Sim_{\mathrm{NVideo}}(v1, v2) = min(n(v1), n(v2))/max(n(v1), n(v2))$. The total similarity of two videos including a minimum-length shot is calculated by the equation: $Sim_{\mathrm{com_{min}}} = Sim_{\mathrm{HVideo}}(1, 1) \times Sim_{\mathrm{LumVideo}}(1, 1) \times Sim_{\mathrm{NVideo}}(2, 2)$, and that of two videos without any minimum-length shot is calculated by the equation: $Sim_{\mathrm{com}} = Sim_{\mathrm{com_{min}}} \times Sim_{\mathrm{HVideo}}(4, 4)$, where two numbers of the arguments are the number of divisions of a frame.

Plasma video evaluation criteria (PVEC) include seven criteria [1]. We prepared five plasma videos (KEY1-KEY5) as key videos, and fifteen videos as database videos, which include the key videos. Five male subjects saw key videos and database videos, and answered the similarity between key and database videos along the seven evaluation criteria of PVEC in three steps: 0: dissimilar, 1: a little similar, and 2: similar [1]. The average values of similarity were obtained. These values are called the similarities obtained through the evaluation experiments.

Correlation coefficients between the total similarities obtained based on the luminosity of every two pixels and the similarities obtained through the evaluation experiments are shown in Table I. As for KEY2, $Sim_{\mathrm{com_{min}}}$ is used as the total similarity. As the average values of the correlation coefficients are around 0.7, the total similarities calculated are correlated to the similarities obtained through the evaluation experiments. It is considered that the total similarity proposed could give good similarities to two videos.

1) Shiroshita, T. *et al.*: Similarity Retrieval of Plasma Videos and Its Evaluation, Proc. of 3rd Int'l Conf. on Applied Comp. and Inf. Tech. (ACIT 2015), pp.229–235 (2015).

2) Chida, T. *et al.*: Fast Detection of Query Videos from Video Streams, Proc. of DEIM Forum 2013,D1-5 (2013).

Table I: Correlation coefficients between the total similarities and the similarities obtained through the evaluation experiments

key video	Cr1	Cr2	Cr3	Cr4	Cr5
KEY1	0.50	0.56	0.63	0.56	0.67
KEY2	0.60	0.56	0.58	0.62	0.53
KEY3	0.66	0.80	0.79	0.75	0.86
KEY4	0.81	0.76	0.77	0.80	0.86
KEY5	0.69	0.55	0.75	0.81	0.80
Average	0.65	0.65	0.70	0.71	0.75