THPM 17.2

A TV DRAMA DIGESTING METHOD BASED ON THE MENTAL SIDE OF THE CONTENT

Tsuyoshi MORIYAMA and Masao SAKAUCHI

Institute of Industrial Science, University of Tokyo,
7–22–1 Roppongi Minato-ku, Tokyo, 106–8558 Japan
moriyama@sak.iis.u-tokyo.ac.jp

ABSTRACT

In this study, the method to digest TV drama based on the mental side of the picture content is proposed and implemented. The video structure model which consists of the tracks of actor's utterances, BGM, offset noise, effect sounds and video shots is defined and the relation between the chosen frames in the digest and their structural features in a live drama video and the broadcasted digest made by the professional beforehand is clarified. The system, which implements the relation, determines the candidate frames for the digest of the video based on the relation information and makes the digest of the specified length.

1. INTRODUCTION

Considering the recent and near future arrangement of the broadcasting services such as coming digital broadcasting, it is supposed that the technology to sum the content of a video program up including its essential points increasingly grows its importance in order to make viewing all over programs easy and fast.

Also in the past studies about digesting video or the video skimming technology [1][2], such technologies have been investigated. But there have so far never been the idea to make the digest of a TV program from the psychological side of the video content, because almost of those have treated the kind of programs in which the meaning information is the most important component such as news programs, then the researchers have mattered the meanings like the names, places and the accidents they tell. On the other hand about a TV drama, it is supposed that the important keys of the content exists not in the physical aspect such as scene changes but rather in the psychological unfolding, so it is effective to truncate such the frames which include psychologically important components.

In this paper, the digesting method based on the mental side of the video content is proposed and implemented, which treats TV drama programs.

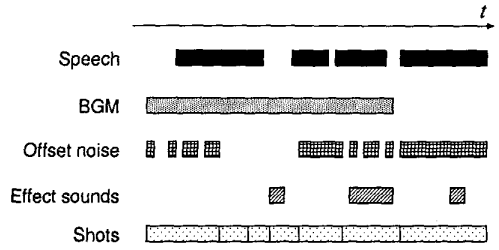

Figure 1: Structure of the drama video

2. STRUCTURE OF TV DRAMA

The structure model of a drama video is shown in Fig.1. A drama video mainly consists of the auditory component (speech by the actors, offset noise, effect sounds and BGMs) and the visual component (shots [3]). And it is assumed in this study that those components are the physical factors which are concerned with making the psychological unfolding of the drama.

In this study, it is also assumed that the physical structure made by those structural components, which are called tracks, is more essential to express the psychological description than the meanings of the speech or the content of the lyrics of BGMs. Namely, under the tense situation, for example, nonexistence of any BGMs and speech and frequent scene changes (rapid camera works) is supposed to be characteristic. Accordingly, to extract the relation between the mental side and the structural features of the TV drama video is the purpose of this study.

3. METHOD TO DIGEST DRAMA

Fig.2 indicates the method of making a digest of a TV drama.

3.1. Acquisition of the relation information

In this study, the structure of the video data is related to the psychological impression (sad, tense and so on), which

0-7803-6301-9/00 $10.00 © 2000 IEEE

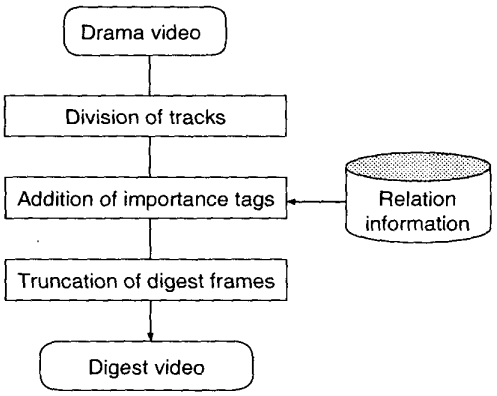

Figure 2: Method of making digest of a TV drama

is necessary to observe in order to make digests, and the relation information is applied to put the weights on each scenes in the drama. And the candidate frames are determined based on the weights.

Two methods of obtaining the relation information are putting the standard on

1. the broadcasted digest made by professionals and
2. the result of the psychological experiments.

1. is adopted at first here. Namely the structural features of the broadcasted digest frames in the substance of the TV drama are investigated and the relation information is extracted.

3.2. Division of tracks

This process is dividing video to the tracks in Fig.1. The shots in video are judged and the timings of each speech, effect sounds, offset noise and BGMs are observed manually here. Moreover, picture effects such as 'disolve' are observed simultaneously.

3.3. Addition of importance tags

Based on the relation information mentioned above, the partial video candidates of the digest are marked.

3.4. Truncation of digest frames

According to the specified length of the digest, the boundary frames on the candidate partial video are changed.

4. CONCLUSION

In this study, the method of making a digest of a TV drama was proposed and implemented. The structural model of a TV drama was shown in the method, and the relation information between the psychological description and the physical structure of the video data was obtained. The truncating procedure of the important scenes for digesting the drama was shown.

In the future, the automatic method of dividing the tracks is investigated and the digesting system proposed here is implemented to the online automatic system such as the EPG in the shape of plug-ins.

5. REFERENCES

[1] Y. Nakamura, T. Kanade "Semantic Analysis for Video Contents Extraction – Spotting by Association in News Video." Proc. ACM Multimedia, pp. 393-401, 1997.

[2] Zhang,H., Low,C., Smoliar,S. and Wu,J. " Video Parsing, Retrieval and Browsing: An Integrated and Content-Based Solution." Proc. ACM Multimedia, pp. 15-24, 1995.

[3] Y.Gong and M.Sakauchi, "A method for color moving image classification using the color and motion features of moving Images," ICARCV'92, 1992.

[4] T.Satou and M.Sakauchi, "GOLS: A hyper media platform which can process television pictures," Inst.of Television Engineers of Japan(ITEJ), Tech. Report, vol.16.no.80, 1992.

[5] T.Satou and M.Sakauchi, "Video Acquisition on Live Hypermedia, " Proc.of ICMCS95, 1995.

[6] H.Davis, W.Hall, I.Heath, and G.Hill, "Towards an integrated information environment with open hypermedia systems," Proc. of ACM Conference on Hypertext(ECHT'92), pp.181-190, 1992.

[7] D.Lowe and A.Ginige, "HyperImages: using object recognition to navigate through images in multimedia,"IS&T/SPIE Symposium on Electronic Imaging, pp.5-10, 1995.

[8] H.Zhang, Y Gong, S.W.Smoliar and S.Y.Tan, "Automatic parsing of news video, " Proc. of IEEE Multimedia '94, pp.45-54, 1994.

[9] D.Swanberg, C.F.Shu, and R.Jain, "Knowledge guided parsing in video databases, "Proc.SPIE IS&T Symp. on Storage and Retrieval for Image and Video Databases, 1993.

[10] Y.Yaginuma and M.Sakauchi, "Moving TV image analysis based on multimedia fusion focusing on extracted common concepts, " IECON'93, pp.1803-1807, 1993.