

Characteristics Extraction from MPEG-Compressed Bitstream and Its Application: An Intelligent Transcoder

Kan-Li Huang, Yi-Shin Tung, Ja-Ling Wu, Senior Member, IEEE, Po-Kang Hsiao, and Hsien-Shuo Chen

*Communication and Multimedia Laboratory,
Department of Computer Science and Information Engineering,
National Taiwan University, Taipei, Taiwan*

ABSTRACT

In this paper, we present a frame-based characteristics extraction tool, which can retrieve meaningful information from MPEG-compressed bitstreams. Taking advantage of the extracted information for content analyses, such as scene change detection and bitrate allocation, we implemented a practical and useful application: an intelligent video transcoder. This representative application demonstrates the potential usage of such frame-based characteristics, and it is believed that these characteristics can also benefit MPEG-7 related applications.

INTRODUCTION

Video compression techniques are getting mature after years of efforts on development of successful standard and de-facto codecs. The well-known moving picture expert group (MPEG) has published three versions of video codecs (MPEG-1; MPEG-2, and MPEG-4). They are broadly used for digital video storage, presentation and transmission. By exploring these codecs in depth, we find that they not only reduce the data rate explicitly but also hide some meaningful information implicitly. The information hidden in the MPEG syntax is rich and worth to investigate further as the compression performance getting better. And they may benefit many applications, such as scene change detection [1], video indexing [2], meta-data extracting [3], transcoding [4], and etc.

THE PROPOSED INTERMEDIATE FORMAT AND INTELLIGENT VIDEO TRANSCODER

First, we define the low-level characteristics extractable from MPEG bitstreams and categorize them into three classes: Temporal Similarity, Spatial Property, and Region Perceptibility. These three characteristics, defined as "Low-level Descriptors", are described by XML (Extensible Markup Language) for interchangeability and stored in the atom of standardized MPEG-4 file format (MP4) for streaming applications. Fig. 1 depicts the relationship diagram of these descriptors.

Then, an intelligent video transcoding system is implemented by using these characteristics. This transcoder not only can convert syntax (MPEG-1/2 to MPEG-4) and bitrate (high bitrate to low-bitrate), but also can augment some meta-descriptions into the generated streams simultaneously (c.f. Fig. 2). Moreover, it can also encode individual semantic video segment separately by intelligent frame type conversion. The overall transcoding process and its timing map are shown in Fig. 3. We adopt a working window of several frames to overlap the following four stages. The first stage, called video decoding process, is to decode input video and to extract low-level characteristics, and then working window slides forward. The second stage, called scene change detection process, is to detect scene boundary after neighboring frames are decoded. The scene change detection can be done either by using extracted description [1] or detection in decompressed video frames [2]. At the third stage, named coding type determination process, the new coding type is decided not only to facilitate the cut-and-paste functionality of an arbitrary scene in the output bitstream but also to encode video efficiently. If scene change doesn't occur, we will reuse coding patterns in the original video, but if scene change is detected on some frames, the coding types of those frames are changed to "I" type and those of their immediately previous frames to "I" type or "P" type depending on the original frame coding pattern. Under the assumption that object motion in several frames of the same scene can be well approximated by linear motion model, special coding mode (direct mode) of MPEG-4 can operate well by this way (c.f. Fig. 4). The final stage is named re-encoding process. Two works are done in this stage. One is reusing frame coding pattern and macroblock coding mode to re-encode input video. The other is performing rate control during re-encoding process. Based on extracted characteristics such as Region Perceptibility, effective rate control scheme can be integrated into well-known rate control algorithms, such as TM5, to get better performance.

CONCLUSIONS

Our work describes the MPEG coding properties in a

semantic way, and uses this information in the applications of transcoding and video content analyzer. Furthermore, the proposed system can transcode and annotate the MPEG coded stream intelligently. And this system can be treated as a gateway service at the VOD server side to fit underlying channel bandwidth, or integrated into DSP as a new functionality of current settop-box to enable storing user-preferred programs with their desired bitrate or quality (c.f. Fig. 5).

REFERENCES

- [1] Soo-Chang Pei and Yu-Zuon Chou, "Efficient MPEG compressed video analysis using macroblock type information," in IEEE Trans. Multimedia, vol. 1, pp. 321-333, Dec. 1999.
- [2] R. Brunelli, O. Mich, and C. M. Modena, "A Survey on the Automatic Indexing of Video Data," in Journal of Visual Communication and Image Representation 10, pp. 78-112, 1999.
- [3] ISO/IEC JTC1/SC29/WG11, Coding of Moving Pictures and Audio, CD 15938 MPEG-7 Multimedia Description Interface - Part 2 Description Definition Language, Part 3 Visual, Part 5 Multimedia Description Schemes.
- [4] T. Shanableh and M. Ghanbari, "Heterogeneous video transcoding to lower spatial-temporal resolutions and different encoding formats," IEEE Transaction on Multimedia, vol. 2, no. 2, pp. 101-110, Jun. 2000.

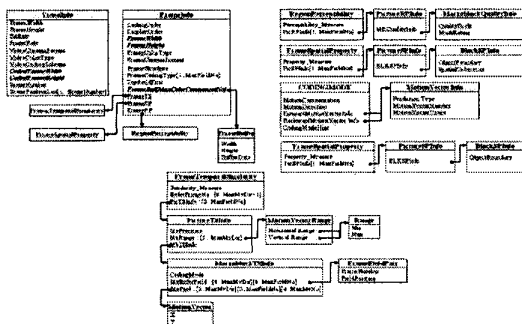


Figure 1: The Relationship Diagram of Low-Level Descriptors.

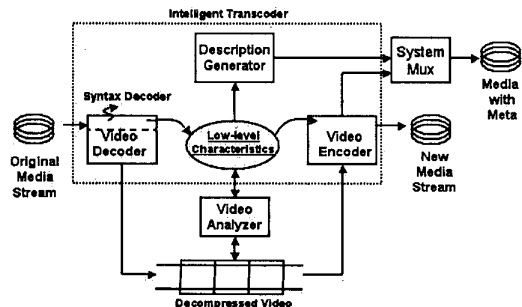


Figure 2: The Block Diagram of the Proposed Intelligent Transcoder:

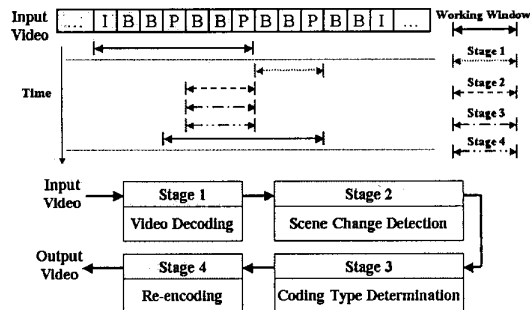


Figure 3: The Flowcharts of Video Transcoding Process and Timing Map.

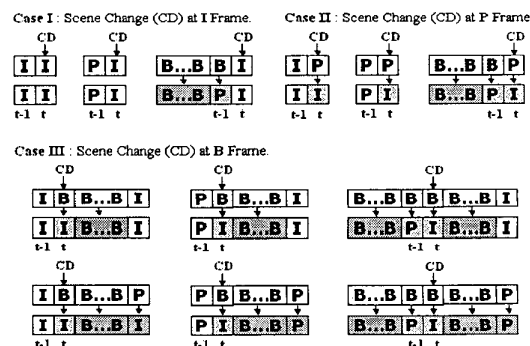


Figure 4: Coding Pattern Selection Rules. Based on coding type of scene-change frame, all situations can be classified into three cases.

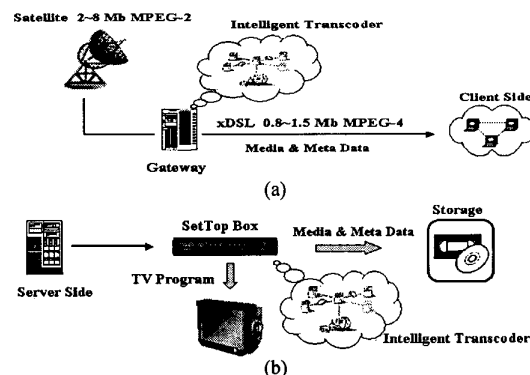


Figure 5: The Applicative Scenarios of the Intelligent Transcoder. (a) The gateway of the server application. (b) The digital settop-box application.