Video Retrieval

Yoshinobu Hotta
Image Computing Laboratory
Fujitsu Laboratories Ltd.

Currently, digital video contents are fast becoming the main contents transferred on the broadband Internet. Users can watch a variety of digital media: digital videos, DVD, digital television and so on. Moreover, the quantity of digital video contents increases exponentially.

Since users want to retrieve a video or a scene with ease, an efficient retrieval technique is necessary for large digital video archives.

As for TV programs, a method of using caption character information is effective, since closed captioning is not widely used in Japanese TV program. When a caption appears in video, it is recognized as text and saved with video as meta-data. Then we can retrieve the video by keyword. Along with the increase of video, a lot of captions in various forms have appeared. So it is necessary to robustly recognize various forms of captions.

On the other hand, caption characters are not always attached to all video scenes. In this case, visual features such as color, shape and texture features can be clues to retrieve video scenes.

In this talk, first, caption recognition method for multi-color characters on complex background is introduced. Our method extracts decomposed binary images from input color caption image by color clustering. Then character candidates that are composed of combination of connect components are extracted by using recognition certainty. Finally, characters are selected by beyond-color Dynamic Programming method in which weight on recognition certainty, character alignment, and continuity of color are used.

Second, interactive video retrieval system using visual features is introduced where similar scene images are arranged close to each other in a virtual 3D space by the self-organizing maps (SOM) so that users can retrieve scenes intuitively and efficiently.

Third, a digest video generation method is introduced for efficiently retrieving desired scenes and watching short clips from lengthy video sequences. This method is applicable to video application services in a wide variety of businesses such as TV programs, movies, homemade videos and surveillance footages. And highlight scenes extraction for sports programs is shown as an example.