Feature Representation for Quick and Robust Media Searching and Processing

Takayuki Kurozumi
School of Information Science,
Japan Advanced Institute of Science and Technology

September, 2007

Abstract

This thesis describes feature representations of signals used to search media data quickly and accurately from among vastly numerous multimedia data such as audio and video recordings.

The query data used for seeking information are assumed to be audio or video signals. This kind of retrieval approach can be performed without using metadata such as keywords. Moreover, such an approach is applicable to various applications using query signals, such as surveillance of pirated multimedia contents on the internet.

The stored signal in the database is assumed to comprise a huge amount of audio or video data that extend into tens of thousands of hours. The query signal is assumed to be an audio or video segment, such as an approximately ten-second signal captured from audio or video of TV broadcasts, or captured with a speaker and a microphone, or a display and a camera in the real world. The search process using such query signals is obtainable as metadata of the stored signal in the database by detecting and locating a similar signal in the database.

Major problems in this kind of retrieval approach are that the features of the query signal might include noise or distortion caused by environmental conditions and that the calculation time is too long because of the huge parameter space that must be searched.

This paper proposes a solution to those problems using feature degeneration, feature generation, and sparse feature selection. Feature degeneration absorbs noise and distortion; feature generation creates noise and distortion. Moreover, sparse feature selection chooses robust features based on statistical values. This thesis describes feature degeneration, feature generation, and sparse feature selection, along with their effectiveness as shown through experimentation. Subsequently, a combination of feature degeneration, feature generation, and sparse feature selection is proposed along with discussion of present and future work.

Key Words: **signal search, degeneration, generation, feature selection, normalization, occlusion, signal lack**