Chapter 8

Management of Multimedia Semantics Using MPEG-7

Uma Srinivasan, CSIRO ICT Centre, Australia

Ajay Divakaran, Mitsubishi Electric Research Laboratories, USA

ABSTRACT

This chapter presents the ISO/IEC MPEG-7 Multimedia Content description Interface Standard from the point of view of managing semantics in the context of multimedia applications. We describe the organisation and structure of the MPEG-7 Multimedia Description schemes which are metadata structures for describing and annotating multimedia content at several levels of granularity and abstraction. As we look at MPEG-7 semantic descriptions, we realise they provide a rich framework for static descriptions of content semantics. As content semantics evolves with interaction, the human user will have to compensate for the absence of detailed semantics that cannot be specified in advance. We explore the practical aspects of using these descriptions in the context of different applications and present some pros and cons from the point of view of managing multimedia semantics.

INTRODUCTION

MPEG-7 is an ISO/IEC Standard that aims at providing a standard way to describe multimedia content, to enable fast and efficient searching and filtering of audiovisual content. MPEG-7 has a broad scope to facilitate functions such as indexing, management, filtering, authoring, editing, browsing, navigation, and searching content descrip-

tions. The purpose of the standard is to describe the content in a machine-readable format for further processing determined by the application requirements.

Multimedia content can be described in many different ways depending on the context, the user, the purpose of use and the application domain. In order to address the description requirement of a wide range of applications, MPEG-7 aims to describe content at several levels of granularity and abstraction to include description of features, structure, semantics, models, collections and metadata about the content.

Initial research focused on feature extraction techniques influenced the description of content at the perceptual feature level. Examples of visual features that can be extracted using image-processing techniques are colour, shape and texture. Accordingly, there are several MPEG-7 Descriptors (Ds) to describe visual features. Similarly there are a number of low-level Descriptors to describe audio content at the level of spectral, parametric and temporal features of an audio signal. While these Descriptors describe objective measures of audio and visual features, they are inadequate for describing content at a higher level of semantics to describe relationships among audio and visual descriptors within an image or over a video segment. This need is addressed through the construct called Multimedia Descriptions Scheme (MDS), also referred to simply as Description Scheme (DS). Description schemes are designed to describe higher-level content features such as regions, segments, objects and events, as well as metadata about the content, its usage, and so forth. Accordingly, there are several groups or categories of MDS tools.

An important factor that needs to be considered while describing audiovisual content is the recognition that humans start to interpret and describe the meaning of the content that goes far beyond visual features and cinematic constructs introduced in films. While such meanings and interpretations cannot be extracted automatically, because they are contextual, they can be described using free text descriptions. MPEG-7 handles this aspect through several description schemes that are based on structured free text descriptions.

As our focus is on management of multimedia semantics, we look at MPEG-7 MDS constructs from two perspectives: (a) the level of granularity offered while describing content, and (b) the level of abstraction available to describe multimedia semantics. The second section provides an overview of the MPEG-7 constructs and how they hang together. The third section looks at MDS tools to manage multimedia semantics at multiple levels of granularity and abstraction. The fourth section takes a look at the whole framework from the perspective of different applications. The last section presents some discussions and conclusions.

MPEG-7 CONTENT DESCRIPTION AND ORGANISATION

The main elements of MPEG-7 as described in the MPEG-7 Overview document (Martínez, 2003) are a set of tools to describe the content, a language to define the syntax of the descriptions, and system tools to support efficient storage and transmission, execution and synchronization of binary encoded descriptions.

Related Content

Rank-Pooling-Based Features on Localized Regions for Automatic Micro-Expression Recognition

Trang Thanh Quynh Le, Thuong-Khanh Tranand Manjeet Rege (2020). *International Journal of Multimedia Data Engineering and Management (pp. 25-37).*

www.irma-international.org/article/rank-pooling-based-features-on-localized-regions-for-automatic-micro-expression-recognition/267765

Counterfactual Autoencoder for Unsupervised Semantic Learning

Saad Sadiq, Mei-Ling Shyuand Daniel J. Feaster (2018). *International Journal of Multimedia Data Engineering and Management (pp. 1-20).*

www.irma-international.org/article/counterfactual-autoencoder-for-unsupervised-semantic-learning/226226

Board Game Supporting Learning Prim's Algorithm and Dijkstra's Algorithm

Wen-Chih Chang, Te-Hua Wangand Yan-Da Chiu (2012). *Methods and Innovations for Multimedia Database Content Management (pp. 256-270).*

www.irma-international.org/chapter/board-game-supporting-learning-prim/66698

Interactive Multimedia and Listening

Yen (Ingrid) Vo (2018). *Digital Multimedia: Concepts, Methodologies, Tools, and Applications (pp. 752-773).*

www.irma-international.org/chapter/interactive-multimedia-and-listening/189503

Processor for Mobile Applications

Ben Abdallah Abderazek, Arquimedes Canedoand Kenichi Kuroda (2009). *Handbook of Research on Mobile Multimedia, Second Edition (pp. 510-522).*

www.irma-international.org/chapter/processor-mobile-applications/21025