Introduction to MPEG-7: Multimedia Content Description Interface

Article · January 2002

CITATIONS
884

READS
893

3 authors:

B. S. Manjunath
University of California, Santa Barbara
440 PUBLICATIONS **22,129** CITATIONS
SEE PROFILE

Philippe Salembier
Universitat Politècnica de Catalunya
196 PUBLICATIONS **6,536** CITATIONS
SEE PROFILE

Thomas Sikora
Technische Universität Berlin
185 PUBLICATIONS **5,463** CITATIONS
SEE PROFILE

1

Introduction to MPEG-7: Multimedia Content Description Interface

Leonardo Chiariglione

Telecom Italia Lab, Italy

Established in 1988, the Moving Picture Experts Group (MPEG) has developed digital audiovisual compression standards that have changed the way audiovisual content is produced by manifold industries, delivered through all sorts of distribution channels and consumed by a variety of devices.

MPEG is intimately connected to digital audio and video. However, when MPEG made its first breath, bits were already abundant, but they were kind of 'heavy' bits. They were part of Pulse Code Modulation (PCM) samples of music stored on compact discs. No one thought of storing or moving a song around in digital form when this meant to store or move 50 MB, unless this was done in a special environment like a studio. The only known way of moving audio and video was in the shape of analog waveforms.

MPEG-1 and MPEG-2 changed this environment radically. Audio files became manageable, the more so if the user was willing to get the music with some artifacts in exchange of a reduced file size or reduced transmission time. The number of television programs started multiplying by orders of magnitude. First, because more television programs in digital form could be packed in the bandwidth that used to carry one television program and second, because of the ability to make new offerings, thanks to the new economies of scale made possible by audio and video in digital form. Compact discs could be used to store movies, and new types of compact discs were even invented to store movies in new forms. Then MPEG-4 extended the possibility to deliver audio and video to new environments such as the Internet and mobile channels.

With the full range of bit-rates covered by the three standards, it could have been expected that MPEG follow the advice of an interested party: to take a sabbatical or, better, to simply disband. The latter option leaves the field open to anybody claiming: 'My algorithm performs 1.5 times better than MPEG' (typical claim of a conference paper) or 'My algorithm performs 3 times better than MPEG' (typical claim of a naive salesman) or 'My algorithm performs 10 times better than MPEG' (typical claim of a

confessed liar). As a corollary, the MPEG Convenor could have followed another piece of advice from the same interested party and look after his vineyard.

Not so. Already in November 1993, when MPEG was still busy finalizing MPEG-2 and working to define the scope of MPEG-4, a proposal had been made to Subcommittee 29 of JTC 1 (the parent body of MPEG). This proposal requested that MPEG investigate the need to develop a standard that would allow users to identify the content that would be present in '500 channels' (The slogan of those times that represented the effect of digitization of television on the business of Community Antenna Television, CATV). That was the beginning of the idea, but MPEG-7 or 'Multimedia Content Description Interface' only started officially in 1997.

Having completed MPEG-2, at least the most crucial parts of its audio-video-systems parts of it, MPEG could mobilize its energies to plan its work beyond MPEG-4. That work turned out to be the ideal continuation of the work proposed in 1993, that is, the definition of an audiovisual information representation standard that would describe or express the semantic meaning of the information and therefore enable people to discover what is in a set of audiovisual objects without having to actually access the information itself.

New MPEG standards have always presented cultural challenges. The development of MPEG-1 provided the first opportunity for people from different industry backgrounds, mostly consumer electronics and telecommunications, to come together and develop a coding standard. Probably the most innovative concept, both from the technical and business viewpoint, was the clear separation of what would constitute the standard – the decoder – and what was left unspecified and open to competition – the encoder. MPEG-2 was the opportunity for a massive intake of people from the television world, both broadcasters and consumer electronics, to join together and make the resulting team work successfully. At the time of MPEG-4, more people from 'new' industry segments came to MPEG to develop the multimedia and mobile standard.

In the new 'MPEG-7 community' that gradually built up, it was not clear what the difference was between what was needed in an algorithm or an implementation and what was required in a standard. It was also not clear which were the interfaces that the standard made reference to. Even less clear was what characterized an 'MPEG-7 encoder' and which were the exact functions that were left for 'encoder optimization' and which were the subject of standardization because they made reference to the 'MPEG-7 decoder'. The role of the textual representation versus binary representation was also not clear because MPEG-7 applications were not just of the communication type in which binary encoding is paramount for practical applications but were also for local database applications for which binary coding is not necessarily a requirement. The demarcation between Audio or Visual Descriptors and Description Schemes seemed impossible to find.

The reason was that MPEG-7 is, conceptually, an audiovisual information representation standard, but the representation satisfies very specific requirements.

The difficulties highlighted notwithstanding, MPEG-7 has turned out to be a very solid and effective standard. The Audio and Video parts provide standardized 'audio only' and 'visual only' descriptors, the Multimedia Description Schemes (MDS) part provides standardized description schemes involving both audio and visual descriptors,

the Description Definition Language (DDL) provides a standardized language to express description schemes, and the Systems part provides the necessary glue that enables the use of the standard in practical environments. Lastly, the Reference Software contains the huge contributions made by the community to develop the MPEG-7 'Open Source' code.

MPEG-7 will fulfill a key function in the forthcoming evolutionary steps of multimedia. As much as MPEG-1, MPEG-2 and MPEG-4 provided the tools through which the current abundance of audiovisual content could happen, MPEG-7 will provide the means to navigate through this wealth of content. MPEG-7 is not the only shop in town in this area. However, the other 'metadata' initiatives have all been developed to serve the specific needs of one business environment. But this is the age in which content companies are by no means constrained by their own traditional delivery mechanisms. In the same way, content consumers are no longer tied to a single source of content. This is the real value of MPEG-7, a value that is in line with past MPEG standard of providing a generic solution that is technically agnostic of the environment.

This book has been written by some of the most authoritative individuals playing key roles in the development of the MPEG-7 standard. Even though this book is not a substitute to the standard itself, it provides a general overview, as well as insights, into the critical elements of the standard and information that is pertinent to understanding MPEG-7 and its practical use.