1. INTRODUCTION

In Chapters 11, 12 and 13 we discussed the goals of the first two phases of the MPEG Audio standard, MPEG-1 and MPEG-2, and we reviewed the main features of the specifications. MPEG-4 is another ISO/IEC standard that was proposed as a work item in 1992 [ISO/IEC MPEG N271]. In addition to audiovisual coding at very low bit rates, the MPEG-4 standard addresses different functionalities, such as, for example, scalability, 3-D, synthetic/natural hybrid coding, etc. MPEG-4 became an ISO/IEC final draft international standard, FDIS, in October 1998 (ISO/IEC 14496 version 1), see for example [ISO/IEC MPEG N2501, N2506, N2502 and N2503]. The second version of ISO/IEC 14496 was finalized in December 1999 [ISO/IEC 14996]. In order to address the needs of emerging applications, the scope of the standard was expanded in later amendments and, even currently, a number of new features are under development. These features will be incorporated in new extensions to the standard, where the newer versions of the standard are compatible with the older ones.

The MPEG-4 standard targets a wide number of applications including wired, wireless, streaming, digital broadcasting, interactive multimedia and high quality audio/video. Rather than standardize a full algorithm and a bitstream as was done in MPEG-1 and 2, MPEG-4 specifies a set of tools, where a tool is defined as a coding module that can be used as a component in different coding algorithms. Different profiles, that represent a collection of tools and refer to a particular application, are defined in the standard.

MPEG-4 Audio includes, in addition to technology for coding general audio as in MPEG-1 and 2, speech, synthetic audio and text to speech
interface technology. Features like scalability, special effects, sound manipulations, and 3-D composition are also included in the standard. While MPEG-1 and 2 Audio typically specify the data rate at the time of the encoding process, the scalability feature in MPEG-4 allows for a system data rate, which is, with some boundaries, dynamically adaptable to the channel capacity. This feature provides significant benefits when dealing with transmission channels with variable capacity, such as internet and mobile channels.

In this chapter, a high level description of MPEG-4, its goals and functionalities are discussed. The development of MPEG-4 Audio is then presented followed by a description of the basic tools and profiles of MPEG-4 Audio. Finally an evaluation of the audio coding tools performance is discussed and intellectual property management issues are introduced.

2. MPEG-4: WHAT IS IT?

The MPEG-4 standard specifies the coding parameters of elements of audio, visual, or audiovisual information, referred to as “media objects”. These objects can be multidimensional, natural or synthetic, i.e. they can be recorded from natural scenes with a microphone and a video recorder or they can be computer-generated [Chiariglione 98].

For example (see Figure 1), a talking person can be represented as the ensemble of basic media objects such as the background image (still image object), the talking person without the background (video object) and that person’s voice plus background noise (audio object).

In addition, the MPEG-4 standard describes the composition of these objects to create groups of media objects that describe an audiovisual scene. For example, the audio object representing the person’s voice can be combined with video object representing the talking person to form a new media object containing both the audio and visual components of the talking person and then further combined into more complex audiovisual scenes.

MPEG-4 defines also the multiplexing and synchronization of the data associated with media objects, so that they can be transported over media channels, and it provides means for interaction with the audiovisual scene generated at the receiver’s end. It incorporates identification of intellectual property and supports controlled access to intellectual property through the requirements specified in the “Management and Protection of Intellectual Property”, IPMP, part of the standard [ISO/IEC 14496-1, ISO/IEC MPEG N2614].