# Chapter V
# Modality Conversion:
## Toward the Semantic Frontier of UMA

**Truong Cong Thang**
*Information and Communications University, Korea*

**Yong Man Ro**
*Information and Communications University, Korea*

## ABSTRACT

Content adaptation currently appears to be the key solution to support the quality of service (QoS) for multimedia services over heterogeneous environments. In this chapter, we study modality conversion as an important adaptation method. We point out two main challenging issues of the problem of modality conversion: (1) the quantification of the content value (quality) when contents are drastically scaled and/or converted to other modalities and (2) the method to accurately decide the modality and content value for each object given that quantification. Accordingly, we will present in detail the solutions as well as the open questions to these two issues. Moreover, we discuss the benefits and obstacles as well as future trends of modality conversion in realizing the goal of universal multimedia access.

## INTRODUCTION

Multimedia contents are increasingly being created in different formats, standards, modalities, and complexities. Meanwhile, the explosive growth of the Internet and wireless networks brings about the diversity and heterogeneity of terminals and network connections. This fact results in big challenges for both providers and users in accessing and disseminating the huge amount of multimedia contents today. The concept of Universal Multimedia Access (UMA) has been introduced in the literature to address the need for solutions to these challenges (Vetro, 2004).

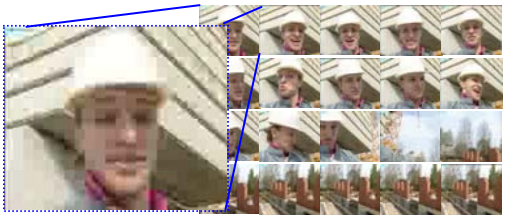
Content adaptation currently appears to be the key solution to support the quality of service (QoS) for multimedia services over heterogeneous environments (Chang & Vetro, 2005). In this trend, the rich multimedia contents are adapted to meet various constraints of terminals, network connections, and user preferences, so as to provide the best possible presentation to the user at anytime and anywhere. It should be noted that in the literature, the term *transcoding* sometimes implies format conversion (e.g., MPEG-2 to MPEG-4). In this chapter, *transcoding* will be used interchangeably with *content adaptation*.

Obviously, when a terminal cannot support the original modality of a content, that modality could be converted to another modality that can be displayed by the terminal. Besides, modality conversion (also called transmoding) can be used to meet some resource constraint while maximizing the quality for users. Given a bitrate constraint, the provider normally (down)scales a content to meet the constraint. However, in some cases, the quality of the scaled content is unacceptable, and a possible solution for this problem is to convert the contents into other modalities. For example, when the connection bitrate is too low, sending a sequence of "important" images would be more appropriate than streaming a scaled video of low quality. This is a typical case of video-to-image conversion.

An illustration for the video-to-image conversion is shown in Figure 1. Here, a Foreman video stream consisting of 20 frames (MPEG-4 format) is compared with a sequence of eight "important" images (JPEG format); both versions are encoded at the bitrate of 13.3kbps. We can see that both versions have the same semantics, but the visual clarity of the image sequence is much better than that of the video stream.

In this chapter, we try to provide a comprehensive picture of modality conversion. As modality conversion lies within the context of content adaptation/transcoding, we first present the key conceptual components of a content adaptation framework. Among these components, we will delve into the decision module, which can be considered as the QoS manager of the whole system. We identify the factors affecting modality conversion and show how they are technically related in a general framework. Especially, we point out two important issues of modality conversion for QoS management; namely,

*Figure. 1. Comparison of a video stream and an image sequence at 13.3kbps (Thang et al., 2004)*



(a) Video stream

(b) Image sequence