MITSUBISHI ELECTRIC RESEARCH LABORATORIES
http://www.merl.com

Transcoding, Scalable Coding & Standardized Metadata

Anthony Vetro

TR2003-113 September 2003

Abstract

The need for content adaptation will exist for quite a while. In this position, views on the relation between transcoding and scalable coding, standardized metadata to assist content adaptation, and thoughts on the future of content adaptation are given.

Transcoding, Scalable Coding & Standardized Metadata

Anthony Vetro

MERL – Mitsubishi Electric Research Labs
201 Broadway, Cambridge, MA 02138, USA

avetro@merl.com

Abstract. The need for content adaptation will exist for quite a while. In this position, views on the relation between transcoding and scalable coding, standardized metadata to assist content adaptation, and thoughts on the future of content adaptation are given.

1 Introduction

As the number of networks, types of devices and content representation formats increase, interoperability between different systems and different networks is becoming more important. Thus, devices such as gateways, multipoint control units and servers, must provide a seamless interaction between content creation and consumption. Transcoding of video content is one key technology to make this possible – scalable coding is another. In the following, the relation between these two seemingly competing techniques is discussed. It is maintained that both solutions can coexist. Furthermore, regardless of which format and technique is used for delivery, standardized metadata to support the adaptation process will be needed.

2 Transcoding vs. Scalable Coding

Scalable coding specifies the data format at the encoding stage independently of the transmission requirements, while transcoding converts the existing data format to meet the current transmission requirements. The holy grail of scalable video coding is to encode the video once, then by simply truncating certain layers or bits from the original stream, lower qualities, spatial resolutions, and/or temporal resolutions could be obtained. Ideally, this scalable representation of the video should be achieved without any impact on the coding efficiency. While current scalable coding schemes fall short of this goal, preliminary results based on exploration activity within MPEG indicate that the possibility for an efficient universally scalable coding scheme is within reach. Besides the issue of coding efficiency, which seems likely to be solved soon, scalable coding will need to define the application space that it could occupy. For instance, content providers for high-quality mainstream applications, such as DTV and DVD,

have already adopted single-layer MPEG-2 Video coding as the default format, hence a large number of MPEG-2 coded video content already exists. To access these existing MPEG-2 video contents from various devices with varying terminal and network capabilities, transcoding is needed. For this reason, research on video transcoding of single-layer streams has flourished and is not likely to go away anytime soon. However, in the near-term, scalable coding may satisfy a wide range of video applications outside this space, and in the long-term, we should not dismiss the fact that scalable coding format could replace existing coding formats. Of course, this could depend more on economic and political factors rather than technical ones. The main point to all of this is that the scalable coding and transcoding should not be viewed as opposing or competing technologies. Instead, they are technologies that meet different needs in a given application space and it is likely that they will coexist.

3 Standardized Metadata

The use of standardized metadata to assist the content adaptation process is quite central to the distribution of content to diverse and heterogeneous environments regardless of the coding format, i.e., single-layer or scalable coded video. On one hand, MPEG-7 offers tools for the description of multimedia formats, tools that allow summaries to be expressed, tools that provide transcoding hints, and tools that indicate the available variations of multimedia content. On the other hand, MPEG-21 Digital Item Adaptation is standardizing tools to describe the usage environment, which includes terminal capabilities, network characteristics, user characteristics, as well as characteristics of the natural environment. It is expected that these tools will be used in conjunction with each other to allow for negotiation and understanding between both source and destination, which in turn will steer the adaptation process to output content suitable for delivery and consumption within the given usage environment.

4 Thoughts on the Future

Looking to the future of video transcoding, there are still quite a number of topics that require further study. One problem is finding an optimal transcoding strategy. Given several transcoding operations that would satisfy given constraints, a means for deciding the best one in a dynamic way has yet to be determined. One important technique needed to achieve this is a means to measure and compare quality across spatio-temporal scales, possibly taking into account subjective factors, and account for a wide range of potential constraints (e.g., terminal, network and user characteristics). Another topic is the transcoding of encrypted bitstreams. The problems associated with the transcoding of encrypted bitstreams include breaches in security by decrypting and re-encrypting within the network, as well as computational issues. Also, considering the rules and conditions that govern the use/reuse of content, new requirements on the adaptation process may arise.