

# Team Spirit Model Using MPEG Standards for Video Delivery

Neetu Singh, Piyush Chauhan, Nitin Rakesh, Nitin Nitin

*Department of CSE & IT, Jaypee University of Information Technology, Waknaghat, India*

*Email: [nitin.rakesh@gmail.com](mailto:nitin.rakesh@gmail.com), [delnيتين@ieee.org](mailto:delnيتين@ieee.org)*

*Received December 10, 2009; revised January 18, 2010; accepted February 21, 2010*

## Abstract

Today the multimedia content delivery chain faces many challenges. There are increasing terminal diversity, network heterogeneity and pressure to satisfy the user preferences. Therefore, the need of customized contents comes in light to provide users the best possible experience. In this paper, we address the problem of multimedia customization. For the customized content, we suggest the team spirit model (TSM) that uses the web edition content description, MPEG-4 standards and the MPEG-21 multimedia framework. It efficiently implements the video customization

**Keywords:** TSM, MPEG-4, MPEG-21

## 1. Introduction

Users are already swamped by the quantity of material available on conventional mass media. In today's world, the larger range of multimedia material is on hand. This can only serve up to offer customers with so much choice that they are overwhelmed by it. In addition, obviously they expect to be provided with services that grant access to information. The service might be due to the increased accessibility of technologies on a broad level. It arises the need to customize those services with regarding to the user's receiving device, which could be from a handheld to a widescreen terminal. If not so then the provided information could be inadequate. If at users widescreen device the least standard is applied, it would suffer under-represented information while a maximum standard of handheld devices might not be able to properly view or use the service at all [1]. That is why the obvious way of optimizing and adapting are needed. The provided service matches to the device to maintain or even enhance user-friendliness. It is the only way to cope with the technical convergence of devices, which would provide access to the similar services. In this paper, our aim probe into the use of dynamic components involved, which should take into user characteristics and assure an efficient way to "access any information from any terminal", eventually after some kind of adaptation.

In this paper, we are presenting the team spirit model, which take input as web edition content that stores all

content information, MPEG-4 Advanced video coding (AVC), and File Format for audio and video and lastly the MPEG-21 usage environment for other necessary description to efficiently implement adaption and finally achieve the customization in the well-mannered way. In TSM scenarios, it is essential to customize more easily and efficiently for the desired content. The contents have available descriptions of the parts that fit to be matched/bridged: the content and the usage environment. The major objective of this paper probes into the use and development of a TSM system that has capable of managing and customizing content so that any information can be delivered to different terminals and networks.

## 2. Related Work

Video customization is demand of today's world. The customization can be achieved efficiently by using various models like hybrid multimedia access (HMA) Model and Universal Multimedia Access (UMA) Model using standards MPEG-7 and MPEG-21 [1-4]. In this paper, we are focusing on Team Spirit as an adaptation engine to implement the video customization using some advance technology based on MPEG standards. The major objective of this paper is to discuss the role of the various MPEG standards in the context of multimedia customization adaptation and to contribute for a better organization and understanding of the multimedia customization for TSM.

### 3. MPEG Standards Used

MPEG-4 enables different software and hardware developers to create multimedia objects possessing better abilities of adaptability and flexibility to improve the quality of such services and technologies as digital television, graphics, the World Wide Web and their extensions. This standard enables developers to better control their content and to fight more effectively against copyright violation. Data network providers can use MPEG-4 for data transparency. With the help of standard procedures, MPEG-4 data can be interpreted and transformed into other signal types compatible with any available network.

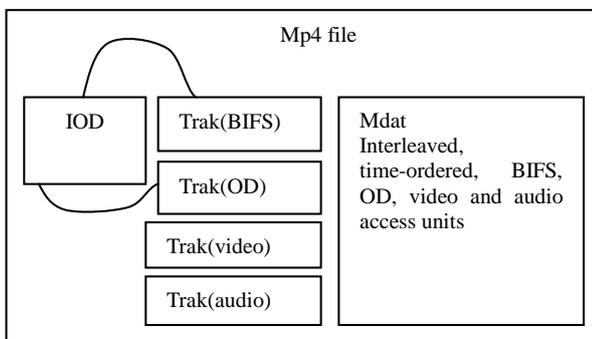
#### 3.1. MPEG-4 (Part-14) File Format

The MP4 file format is intended to contain the information regarding the media of an MPEG-4 presentation in a flexible, extensible format that in turn facilitates interchange, management, editing, and presentation of the media. The presentation may be view as ‘local’ to the system containing the presentation, or may be via a network or other stream delivery mechanism. The file format is designed to be independent of any particular delivery protocol while enabling efficient support for delivery in general. The **Figure 1**, gives an example of a simple interchange file, containing three streams.

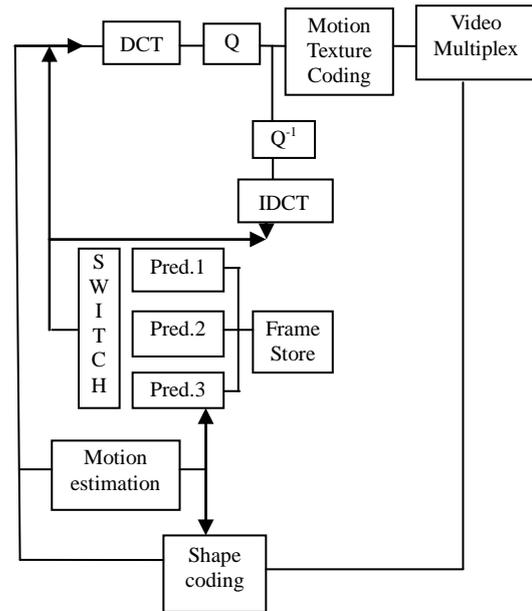
The composition of the Mp4 file format is of object-oriented structures called ‘atoms’. A unique tag and a length identify each atom. Most atoms describe a hierarchy of metadata giving information such as index points, durations, and pointers to the media data. Now the ‘movie atom’ can be defined as the collection of these atoms. The media data itself is located elsewhere; it can be in the MP4 file, contained in one or more ‘mdat’ or media data atoms, or located outside the MP4 file and referenced via URL’s [5].

#### 3.2. MPEG-4 Video Image Coding Scheme

**Figure 2** below outlines the description of the MPEG-4



**Figure 1.** Example of a simple interchange file.



**Figure 2.** Basic block diagram of MPEG-4 video coder.

video algorithms to encode rectangular as well as arbitrarily shaped input image sequences.

The basic coding structure involves shape coding (for arbitrarily shaped VOs) and motion compensation as well as DCT-based texture coding (using standard  $8 \times 8$  DCT or shape adaptive DCT). The basic coding structure involves shape coding (for arbitrarily shaped VOs) and motion adaptive DCT). The basic coding structure involves shape coding (for arbitrarily shaped VOs) and motion compensation as well as DCT-based texture coding (using standard  $8 \times 8$  DCT or shape adaptive DCT).

An important advantage of the content-based coding approach MPEG-4 is that the compression efficiency can be significantly improved for some video sequences by using appropriate and dedicated object-based motion prediction “tools” for each object in a scene.

A number of motion prediction techniques can be used to allow efficient coding and flexible presentation of the objects [5]:

- Standard  $8 \times 8$  or  $16 \times 16$  pixel block-based motion estimation and compensation, with up to  $\frac{1}{4}$  pel accuracy.

Global Motion Compensation (GMC) for video objects: Encoding of the global motion for a object using a small number of parameters. GMC is based on global motion estimation, image warping, motion trajectory coding and texture coding for prediction errors.

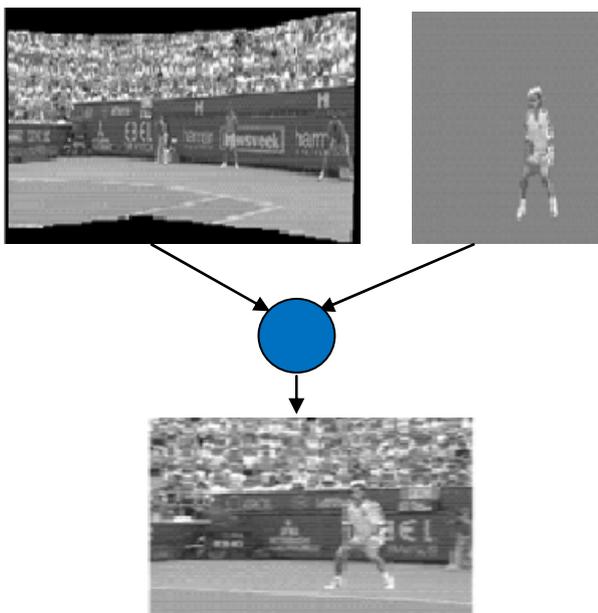
- Global motion compensation based for static “sprites”. A static sprite is a possibly large still image, describing panoramic background. For each consecutive image in a sequence, only 8 global motion parameters describing camera motion are coded to reconstruct the object. These parameters represent the appropriate affine transform of the sprite transmitted in the first frame.

- Quarter Pel Motion Compensation enhances the precision of the motion compensation scheme, at the cost of only small syntactical and computational overhead. A accurate motion description leads to a smaller prediction error and, hence, to better visual quality.

- Shape-adaptive DCT: In the area of texture coding, the shape-adaptive DCT (SA-DCT) improves the coding efficiency of arbitrary shaped objects. The SA-DCT algorithm is based on predefined orthonormal sets of one-dimensional DCT basis functions.

**Figure 3** depicts the basic concept for coding an MPEG-4 video sequence using a sprite panorama image. It is assumed that the foreground object (tennis player, image top right) can be segmented from the background and that the sprite panorama image can be extracted from the sequence prior to coding. (A sprite panorama is a still image that describes as a static image the content of the background over all frames in the sequence). The large panorama sprite image is transmitted to the receiver only once as first frame of the sequence to describe the background - the sprite remains is stored in a sprite buffer. In each consecutive frame, only the camera parameters relevant for the background are transmitted to the receiver. This allows the receiver to reconstruct the background image for each frame in the sequence based on the sprite [5]. The moving foreground object is transmitted separately as an arbitrary-shape video object. The receiver composes both the foreground and background images to reconstruct each frame (bottom picture in figure below). For low delay applications, it is possible to transmit the sprite in multiple smaller pieces over consecutive frames or to build up the sprite at the decoder progressively.

Subjective evaluation tests within MPEG have shown



**Figure 3.** Sprite coding of video sequence.

that the combination of these techniques can result in a bit stream saving of up to 50% compared with the version 1, depending on content type and data rate.

### 3.3. MPEG-4 (Part-10) Avc/H.264

The intent of the H.264/AVC project was to create a standard capable of providing good video quality at substantially lower bit rates than previous standards without increasing the complexity of design so much that it would be impractical or excessively expensive to implement. An additional goal was to provide enough flexibility to allow the standard to be applied to a wide variety of applications on a wide variety of networks and systems, including low and high bit rates, low and high-resolution video, packet networks. The H.264 standard is a “family of standards”: the members of which are the profiles described below.

A specific decoder decodes at least one, but not necessarily all, profiles. The decoder specification describes which of the profiles can be decoded. Scalable video coding as specified. H.264/AVC allows the construction of bitstreams that contain sub-bitstreams that conform to H.264/AVC. For temporal bitstream scalability, *i.e.*, the presence of a sub-bitstream with a smaller temporal sampling rate than the bitstream, complete access units are removed from the bitstream when deriving the sub-bitstream. In this case, high-level syntax and inter prediction reference pictures in the bitstream are constructed accordingly. For spatial and quality bitstream scalability, *i.e.* the presence of a sub-bitstream with lower spatial resolution or quality than the bitstream, NAL (Network Abstraction Layer) removed from the bitstream when deriving the sub-bitstream. In this case, inter-layer prediction, *i.e.*, the prediction of the higher spatial resolution or quality signal by data of the lower spatial resolution or quality signal, is typically used for efficient coding.

### 3.4. MPEG-21 Multimedia Framework

MPEG-21 framework is based on two essential concepts: the definition of a fundamental unit of distribution and transaction (the Digital Item) and the concept of Users interacting with Digital Items. The Digital Items can be considered the “what” of the Multimedia Framework (e.g., a video collection, a music album) and the Users can be considered the “who” of the Multimedia Framework. In practice, a Digital Item is a combination of resources, metadata, and structure. The resources are the individual assets or (distributed) resources. The metadata comprises informational data about or pertaining to the Digital Item as a whole or to the individual resources included in the Digital Item.

Finally, the structure relates to the relationships among the parts of the Digital Item, both resources and metadata. Within MPEG-21, a resource is defined as an individu-

ally identifiable asset such as a video or audio clip, an image, or a textual asset. Hence, terms like multimedia content, media data, image, graphics, video, movie, visual content, audio data, speech content, etc. become unnecessary and should be avoided in the context of MPEG-21 when referring to resources with no specific context. Within any system (such as MPEG-21) that proposes to facilitate a wide range of actions involving “Digital Items”, there is a need for a very precise description for defining exactly what constitutes such an “item”.

Clearly, there are many kinds of content, and probably just as many possible ways of describing it to reflect its context of use. This presents a strong challenge to lay out a powerful and flexible model for Digital Items, which can accommodate the myriad forms that content can take (and the new forms it will assume in the future). Such a model is only truly useful if it yields a format that can be used to represent any Digital Items defined within the model unambiguously and communicate them, and information about them, successfully. The Digital Item Declaration specification (part 2 of ISO/IEC 21000) provides such flexibility for representing Digital Items. Part 7 of the MPEG-21 standard Digital Item Adaptation (DIA) have been created mainly to address the usage environment description problem [6].

### 3.5. Web Edition Content Description

Web edition is a database-grounded web content management system where in a central database; it stores all “content information” including text, images, graphics, flash movies, style sheets, etc. The plus point of this system is that the page itself does not have to be edited if the content has to be changed. As an alternative, only the database entry has to be changed and the corresponding page re-saved. In addition, the content stored in the database can be quickly searched or used in other ways. Therefore, it can be added that a central database contains all the content on the web site user wish to manage with web edition. However, user hardly notices the database, because it runs initially in the background. Users only see their web sites and the input boxes on the actual web pages. The web edition program only has to be installed on the provider's (ISP) server once [7].

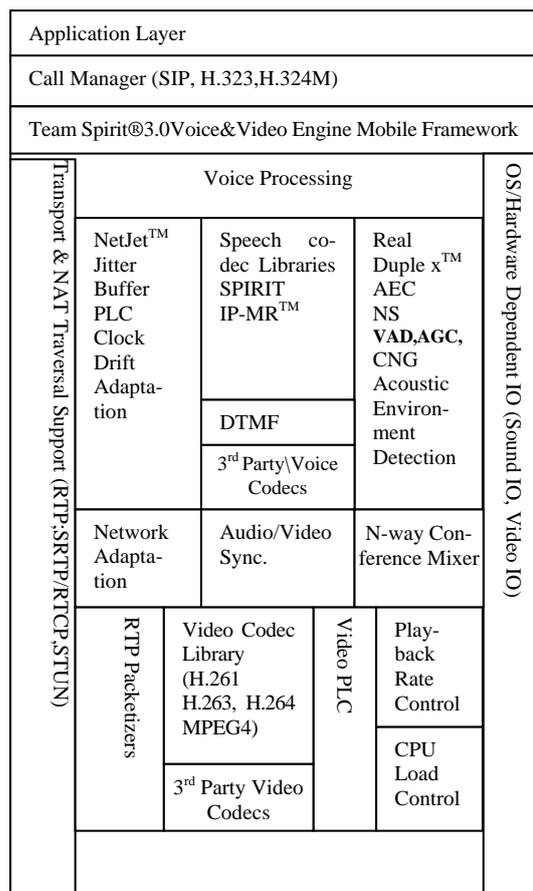
### 4. Team Spirit Engine

Team Spirit® **Figure 4** Mobile is a comprehensive voice and video engine SDK, including a library of highly optimized low-MIPS consuming voice and video codecs to enable mobile video over IP communication on a broad range of handsets, speech enhancement and network adaptation algorithms [8]. To speed up deployment and secure flawless voice quality on every handset, the engine can use tuning Wizard for automatic adjustment of AEC

parameters. MHz-efficiency and Wi-Fi adaptation are extremely important for resource constrained mobile devices, intended for usage in public Wi-Fi networks, more susceptible to delays, jitter, packet loss, aggregated echo, noise and congestion. The SPIRIT's Engine handles Wi-Fi-specific network impairments and makes video transmission over Wi-Fi/WiMAX more stable and reliable. This Engine can be run even on 200 MHz processors. Simultaneously, the Engine supports both software and hardware accelerators inside popular application processors to deliver CIF quality video on mobile devices.

### 5. Team Spirit Model for Video Customization

Starting with the more traditional coding tools such as MPEG-1, and MPEG-2, the recent scalable video coding tools is MPEG-4 Fine-Grain-Scalability (FGS), and passing through web edition content description.MPEG standardization culminates with the MPEG-21 multimedia framework which offers a wrapper to allow all the pieces in a multimedia customization chain to integrate and interact with each other.



**Figure 4. Layered architecture of team spirit.**

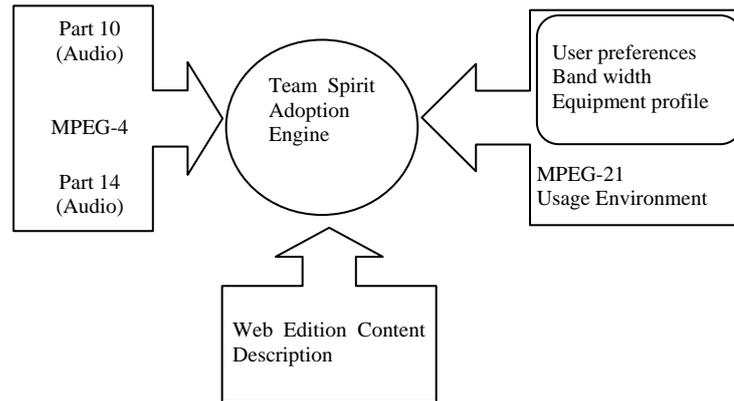


Figure 5. Team spirit model.

Figure 5 depicts a possible configuration of a multimedia customization chain using all MPEG standards. Basically adaptation engine require three inputs, *i.e.*, firstly, MPEG-4 AVC and file format for audio and video processing so that whatever the audio and video requirement like formatting, compression, coding and other essentials regarding both. Secondly, content description web Edition is a database-grounded web content management system. Web edition stores all “content information”. Additionally, the content stored in the database can be rapidly searched or used in other ways. Lastly, at the end user side, the MPEG-21 UED describes the environment (terminal, network, natural environment, and user preferences) where the content is to be consumed.

When the user performs a query or defines a choice, the UED, thus enabling a customizing application to explore this information to create the right content variation to provide the best possible user experience, accompanies request. Finally, the TSM adaptation engine at the center of in Figure E is responsible for matching the user query either by selecting the most adequate available variation, or by performing some adaptation. As processing a user query, the customizing application creates an adapted variation of the Item to be sent to the user—the new variation and its corresponding description may be added to the resources available at the server. The user query response may be delivered through a network, eventually using a real-time connection. In this case, the streaming module will stream the scalable or non-scalable content to the user; in the case real-time transcoding is been performed, it may happen that real time adjustments to the transcoding process are implemented using measures which characterize, for example, the network fluctuations. This is how the customization is achieved.

## 6. Conclusions and Future Work

In this paper, we present a team spirit model for efficient video customization using MPEG standards and Web

edition content description. The MPEG-4 (part-10) is used for advance video coding and MPEG-4 (part-14) for file formatting. For the usage environment description, we have MPEG-21 standards. Hence, these standards act as a base for development of a TSM system.

The TSM systems are capable of managing and customizing content efficiently so that any information can be delivered to different terminals and networks like for television, phones and over other communication media where video customization is needed. In future, after adding some more functionality to it, the model is capable to deploy all essential features related to video delivery in smart phones, iPhones etc.

## 7. References

- [1] S. Kim and Y. Yoon, “Video Customization System Using MPEG Standards,” *International Conference on Multimedia and Ubiquitous Engineering*, 2008.
- [2] S. Kim and Y. Yoon, “Universal Multimedia Access Model for Video Delivery,” *The 3rd International Conference on Grid and Pervasive Computing - Workshops*, 2008.
- [3] B. S. Manjunath, P. Salembier and T. Sikora, “Introduction to MPEG 7: Multimedia Content Description Language,” Wiley, New York, 2002.
- [4] “MPEG Requirement Group, MPEG-21 Multimedia Framework,” *Part 1: Vision, Technologies and Strategy, Proposed Draft Technical Report*, 2nd Edition, MPEG Waikaloa Meeting, USA, December 2003.
- [5] R. Koenen, “MPEG Requirement Group, MPEG-4 Overview,” V.21–Jeju Version, March 2002.
- [6] “MPEG MDS Group, Multimedia Framework,” *Part 7: Digital Item Adaptation, Final Draft International Standard*, MPEG Waikaloa Meeting, USA, December 2003.
- [7] <http://www.macupdate.com/info.php/id/7902/webedition>
- [8] “Spirit DSP, Voice and Engine Expert Group,” *Datasheet*, Juniper Research, February 2002.