Authoring of scalable multimedia documents

Benoit Pellan, Cyril Concolato

Institut TELECOM; TELECOM ParisTech; CNRS LTCI

{Benoit.Pellan, Cyril.Concolato}@telecom-paristech.fr

Abstract: With the diversity of usage conditions affecting the consumption of multimedia content, the adaptation of dynamic and interactive multimedia presentations is essential. The challenge consists in allowing a multimedia presentation to adapt in all its dimensions: spatially, temporally and interactively, without lessening its attractiveness and still giving the author the control over adapted versions. Additionally, the authoring of adaptable content should not increase the complexity of authoring. To address this challenge, we propose to transpose the concept of scalability to the world of multimedia documents by introducing the so-called *Scalable MSTI* model. In this paper, we show the properties of this model and how, from an authoring point of view, scalable multimedia documents can be created to address a wide range of usage conditions.

KeyWords Multimedia Scalability, Scalable Documents, Content Adaptation

1 Introduction

Usability [26][27] is crucial when designing multimedia services and constitutes one of the greatest concerns of content producers. Indeed, although the value of a multimedia content greatly relies on the quality of the media it conveys, only an appropriate composition of all these media elements into a well-suited presentation will lead users to actually enjoy them. As a consequence, ergonomic aspects, aesthetics and legibility properties constitute a significant part of the authoring efforts when developing multimedia contents. In the scope of this paper, the spatial and interactive composition of media elements over time and the styling properties applied to them form a multimedia presentation, also called a multimedia scene.

Multimedia scene usability cannot be simply guarantied by a smartly designed and appealing presentation but it also requires a technical interoperability between the multimedia scene and usage environments. The common approach which is largely deployed in nowadays authoring chains consists in designing multimedia presentations that only feature very well supported functionalities in order to guaranty the expected document playback to the user. In fact, this “one-content-fits-all” strategy is very frustrating for service providers who must author highly constrained multimedia documents to fit the most restricted usage environments. It is also unsatisfactory for end users who cannot have access to an advanced multimedia experience although they are appropriately equipped.

As an extension to our work introduced in [20], this paper presents our scalable multimedia document model that allows addressing the most constrained usage environments with a suitable presentation but also provides multimedia enhancements for advanced usage conditions. The underlying multimedia model of our scalable documents is called the *Scalable MSTI* model and can be used to generate Rich-Media presentations suitable for various usage environments such as heterogeneous terminal capabilities (screen size, processing power, battery, memory, player capabilities…) and various transport mechanisms (file casting, progressive downloading, real-time streaming, broadcasting…). The *Scalable MSTI* model (*Media Spatial Temporal* and *Interactive*) is based on state-of-the-art multimedia document models [4][2][29] that clearly separate media elements and logical structures from the multimedia presentation. Hence, the *Spatial*, *Temporal* and *Interactive (STI)* descriptions of our model reference media elements in the *Media* description to simply apply their presentation properties. Multimedia scalability is additionally provided by dividing these *STI* descriptions intro progressive layers that can be described in a generic XML-based layered structure.

In this paper, we propose some authoring techniques that address the new challenges raised by the production of scalable presentations based on the *Scalable MSTI* model. Indeed, the *Spatial*, *Temporal* and *Interactive* scalability axes introduced in [20] define a three-dimension space that need to be appropriately filled in. In fact, a mapping between adaptation parameters, which represent the complete set of targeted usage constraints, and the *STI* layers must be performed during authoring and will determine the adaptation possibilities of the generated scalable multimedia document. This authoring can be formally divided in four major steps. First, a *Base* document that must address the most constrained usage environments is created. The presentation properties of this *Base* document are separated in *Spatial*, *Temporal* and *Interactive* descriptions. Second, three main adaptation parameters (e.g. display resolution) are assigned to *STI* axes (e.g *Spatial* axis), thus defining adaptation axes. At that stage, one (*Complete*) *STI* presentation enhancing the *Base* document can already be authored to generate a coarse-grained scalable document. Third, a total order relation is associated with each selected adaptation parameter to define the granularity of the adaptation axes and therefore the scalable *STI* layers in the adaptation space. The fourth and last authoring phase consists in filling each defined *STI* layer with presentation properties that meet its adaptation parameter value according to its category (*Spatial*, *Temporal* or *Interactive*). This final authoring task can be driven by a *Complete STI* presentation previously authored and outputs a fine-grained scalable multimedia document that is ready to be visualized in various usage environments with an optimum perceived quality (as chosen by the content creator).

The remainder of this paper is organized as follows. Section 2 introduces related work in the field of adaptation of multimedia scenes. Section 3 summarizes the properties of the *Scalable MSTI*. Section 4 presents the scalable authoring techniques we are proposing by formally decomposing the generation of scalable multimedia documents into four authoring steps. Section 5 concludes this paper and introduces perspectives for future work.

2 Related Work

The authoring of adaptable multimedia content can be tackled using three different principles depending on service scenarios as described in [8]:

- The first approach, also called “device-specific authoring”, consists in preparing presentations for each targeted device configuration. In that case, alternative versions of the same presentation can be automatically selected during content negotiation by identifying requesting devices. Such authoring approach is straightforward and efficient in controlled usage environments but does not scale well as the number of possible device configurations increases.
- The second authoring technique, also called “automatic re-authoring”, lets software system automatically operate a transformation of the presentation according to each specific device configuration [8][14]. This approach lightens the burden of content creators by taking care of usage environment constraints but can only be applied to simplistic (e.g. audio/video only) or structure-driven presentations (e.g. XHTML-based layout) where the analysis of the multimedia content is straightforward. Additionally, no control over adapted presentation of the document is provided to the content creator.
- The third and last approach, also called “multi-device authoring”, does not consider a set of targeted devices but consists in generating presentations with a set of configurable properties that can be selected to address various usage environments [23]. In practice, this approach does not claim to provide the best possible presentation for all device configurations. Indeed, it usually covers a range of adaptation possibilities where prevalent device configurations are implicitly targeted and more specific devices are still able to retrieve a satisfactory presentation.

Our authoring approach can be classified as a “multi-device authoring” approach because it proposes a flexible trade-off between the author’s control over designed presentations and the capacity to address the unpredictable variety of usage environments. In the following, we provide a state-of-the-art of multi-device authoring techniques that tackle the multimedia adaptation problem either at the media level (Section 2.1) or at the scene level (Section 2.2).

2.1 Authoring of adaptable multi-media documents

The main challenge when authoring adaptable multimedia documents is to preserve semantics while altering the presentation to address different usage environments. A lot of approaches propose to address this issue by focusing on the adaptation of media elements according to usage environments. For instance, the adaptability modeling defined in Tiempo [31] relies on media *selection groups* and *Quality of Service* ranges that define priorities on media elements and help in computing an optimal decision at a given time. In a similar way, the synchronized template for adaptable multimedia presentation (STAMP) model [2] defines *data*, *composition* and *decorative* dimensions in addition to traditional spatial, temporal and interactive properties in order to query and select related media elements based on a nucleus-satellite metaphor. A last example is the multimedia adaptation framework based on MPEG-21 standard that can perform generic and dynamic media transformations based on Perceived Quality of Service metrics [22][25]. All these approaches do not directly tackle the adaptation of multimedia presentations but focus on the combined adaptation of several media elements because they assume that an underlying presentation model will smartly take into account these changes operated at the media level. In practice, these approaches mostly rely on automated layout rules (usually based on highly structured languages such as SMIL [6], Madeus [12], HTML or highly-templated documents). For instance, the definition of a *multimedia document specification* given in [14] restricts content production to structured documents to enable automated high-level transformations: *"A multimedia document specification s={O,C} is made of a set of multimedia objects O and a set of constraints C between these objects"*. In this paper, we do not necessarily rely on such structured documents as defined in [7]. Instead, we propose adaptable authoring techniques that also take into account advanced multimedia scenes (such as SVG [1], BIFS [11] or Flash[1]) where the spatial, temporal and interactive properties of a presentation are essential and allow content creators to carefully design their presentations.

2.2 Authoring of adaptable Rich-Media documents

The adaptation of a multimedia document is a transformation that is applied to the description of its presentation. This transformation can result from a dynamic decision based on a straightforward analysis of the presentation but is often driven by authoring hints that are provided along with the presentation so as to preserve authors intentions and content semantics. A significant amount of research efforts have been spent on defining abstract models and metadata that can be used to describe the semantics of multimedia presentations and enhance the adaptation process: media priorities, possible media modalities, alternative layouts… For instance, the multimedia content model (MMCM) [17] proposes a hierarchical model to represent multimedia presentations in order to drive the adaptation process with *semantical* information (rather than *physical* information). In a similar way, multimedia scenes can be annotated with semantic information to guide the adaptation process and preserve the consistency and meaningfulness of the presentation [13]. These authoring approaches rely on two important assumptions. First, content creators have to provide this optional semantic information although they can't clearly control their impact in terms of content usability. Second, adaptation decision-taking engines are supposed to generate an optimal multimedia document based on these adaptation preferences. In practice, the scope of such application scenarios is limited to specific domain [9] where content production, based on specialized authoring tools allowing appropriate content annotations, is closely coupled with adaptation tools leveraging these specific semantic descriptions. Since we are mainly targeting broadcast scenarios where adaptation use cases requires straightforward and generic adaptation algorithm, our approach does not rely on such loose adaptation hints but defines generic adaptation rules (implicitly conveying semantic information) that ensure a deterministic control of adapted presentations.

In the field of multi-device authoring, we have identified two state-of-the-art approaches that focus on advanced multimedia scenes (also called Rich-Media documents) and propose the authoring of generic adaptation rules: constraint-based rules [5] and interpolation-based rules [17].

2.2.1 Constraint-based adaptation rules

The authoring of constraint-based adaptable multimedia documents consists in creating several versions of media elements and scene configurations (style, spatial layout…) along with a set of one-way or multi-way constraints that must be maintained when transforming the document for adaptation purposes. In this approach, the adaptation process takes into account usage environment variables by submitting them to a constraint-solving algorithm generated during the authoring that computes, at playback time, adapted presentation properties. The main difficulty raised by the authoring of constraint-based adaptable documents is that the specifications of document constraints require a high level of expertise and should be automated to be usable. For instance, the *explicit* and *global constraints* that guide the adaptation process in [2] can be provided by an application developer but not a content creator (e.g. *cut(element[media="text"]) = time_sequence*). Based on this conclusion, a customized constraint-based authoring tool has been developed for diagrams [16] and allows defining default adaptation behaviors for multimedia presentations based on several diagram models. However, *"a key question is [still] how the author*

[1] http://www.adobe.com/products/flash/

can specify diagram specific adaptive layout behavior" without having *"to write textual attribute expressions when, say, specifying one-way constraints in CSVG"* [15].

We did not select this approach to author adaptable multimedia documents for two main reasons. First, as explained in [21], constraint-based adaptation requires processing capabilities and bandwidth requirements that are not optimized for restricted devices in Broadcast scenarios. Indeed, even if the complexity of one-way constraint-solving techniques is not high, it might not be acceptable for very limited terminals. Furthermore, a description of adaptation constraints, provided as a JavaScript code in [16], requires a significant bandwidth when frequently repeated in a broadcast carousel. Second, constraint-based adaptive documents are not designed to be updated over time. Indeed, whenever a new object, which was not considered during authoring, is to be dynamically inserted to the multimedia document, the whole constraint-based algorithm description has to be updated and the complete adaptation has to be reprocessed.

2.2.2 Interpolation-based adaptation rules

The authoring of interpolation-based adaptable multimedia document consists in creating several key scene configurations corresponding to typical usage environments. For adaptation purpose, key scenes can be selected or even interpolated to address usage environments that have not been initially covered during authoring. An application of such interpolation-based techniques is the adaptation of Graphical User Interfaces to receivers resolution or screen size (various scene sizes) [10]. The interpolation-based approach extends the scalability features of vector graphics languages such as SVG by providing intermediate versions that are specifically designed for targeted scene sizes. It also leverages authoring efforts of content creators who are used to provide their multimedia documents in several versions (large and small logo of a company, top or left banner for a Flash commercial on a web site…) and advocate incremental authoring approaches described in [19][30].

We did not select this approach to create adaptable multimedia documents for two main reasons. First, although the interpolation-based approach can be seen as a simplification of the constraint-based approach described in Section 2.2.1 (only input parameters of the one-way interpolation process need to be provided), our previous experiments [21] showed that the update mechanism used in our *Scalable MSTI* model is more efficient on very limited receivers and therefore better addresses Broadcast requirements. Second, the adaptation parameters of interpolation-based adaptable multimedia documents are restricted to the spatial layout and *orthogonal interpolations* as defined in [10] have some awkward limitations in terms of flexibility (e.g. rotations, cross-axis constraints, non-linear constraints). However, the interpolation-based spatial adaptation technique and the spatial scalability properties of our document model can be combined to achieve non-linear and infinite spatial scalability.

3 The Scalable MSTI model

The *Scalable MSTI* model proposed in [20] describes multimedia documents composed of several media elements (audio sequences, video clips, images, graphics, text…) and enables author-driven adaptation by introducing the concept of scalability. The *Scalable MSTI* model is based on state-of-the-art multimedia document models [29][4][3] that clearly separate media elements and logical structures from the multimedia presentation. This model is also simply called the *MSTI* model when we want to insist on its core components: *Media, Spatial, Temporal* and *Interactive* descriptions. A scalable *MSTI* document can be defined as a logical structure referencing several media elements (*Media*) on which a set of transformations (*Spatial, Temporal* and *Interactive*) can be applied to generate a multimedia presentation that addresses various usage environments. The *Scalable MSTI* model extends the core components of the *MSTI* model by dividing its *Spatial, Temporal* and *Interactive* descriptions into progressive or incremental layers in order to provide multimedia document scalability. In the following, an overview of the three components of the *Scalable MSTI* model and of three scalable adaptation axes described in [20] is provided in Section 3.1 and in Section 3.2, respectively. Additionally, the concept of adaptation graph and the principles of the *STI* composition are illustrated in Section 3.3 and in Section 3.4, respectively.

3.1 STI components

The *Scalable MSTI* model separates the presentation properties of a multimedia document into *Spatial, Temporal* and *Interactive (STI)* descriptions by grouping media elements, structures and semantics of a multimedia document in a *Media* description. The specification of the *STI* descriptions of our model is driven by perceptual properties and identified multimedia functionalities of existing languages such as SVG, BIFS, SMIL or Flash. Hence, the *Spatial* description defines the layout for all media elements that are part of the document; the *Temporal* description provides the timing of the multimedia presentation as a whole but also the timing of each individual media elements; and the *Interactive* description adds interactive aspects to the multimedia document and triggers the behavior associated with interactions. These three *Spatial*, *Temporal* and *Interactive* components are all related to the same *Media* description and possibly share logic elements as illustrated in Figure 1.

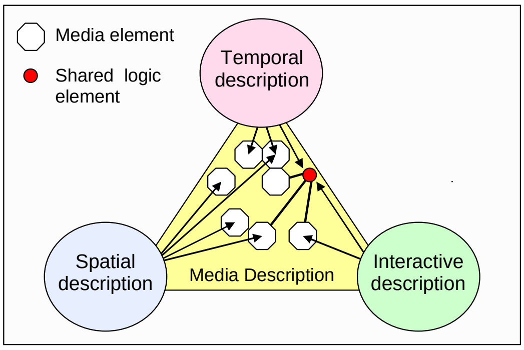


Figure 1. A multimedia document in the MSTI model

3.2 Adaptation axes and scalability axes

In the context of scalability, the *Spatial*, *Temporal* and *Interactive* components of the *MSTI* model are called *STI* axes. The term “scalable” and “scalability” should be understood as in the SVC standard [24] and *“refer to the removal of [video] bit-stream in order to adapt it to the various needs or preferences of end users as well as to varying terminal capabilities or network conditions”*. As a consequence, the *Scalable MSTI* model is based on order relations that organize *STI* axes in a progressive manner according to adaptation parameters so as to enable a straightforward adaptation process. The scalability along an axis can be coarse-grained or fine-grained according to the length of adaptation steps. Each grain (or step) is called a layer. Each layer of a scalable multimedia document is associated with an adaptation parameter related to its applicable usage environments. Since three scalability axes are defined in the *Scalable MSTI* model, at most three adaptation axes have to be defined during document authoring and mapped to *STI* scalability axes. In [20], some typical adaptation axes have been described: resolution (Section 3.2.1), processing power (Section 3.2.2) and memory consumption (Section 3.2.3).

3.2.1 Spatial scalability axis

A typical adaptation parameter which can be mapped onto the *Spatial* axis of the *Scalable MSTI* model is the targeted screen resolution. Since this parameter is bi-dimensional, we restrict targeted spatial dimensions to incremental values as for the generalized spatial scalability of SVC [24]. As a consequence, *“neither the horizontal nor the vertical resolution can decrease from one layer to the next”*. Doing so, a layered description of the spatial properties of multimedia documents can address several screen resolutions. For example, from one layer to another (S_0,S_1,S_2), the *Spatial* description can move objects to better fit the updated screen size, or increase the size of some objects, possibly in a non-linear way as illustrated in Figure 2.

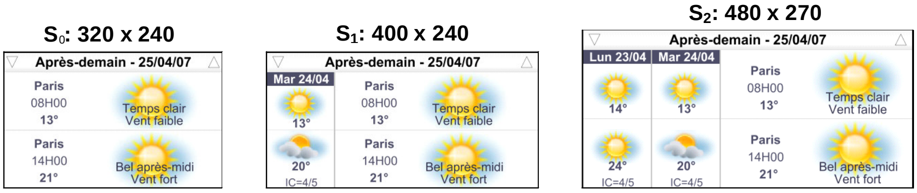


Figure 2: A spatially scalable document driven by the screen resolution

3.2.2 Temporal scalability axis

The transformation of the *Temporal* component of a *MSTI* document into a scalability axis can be performed by providing progressive timed properties. Examples of adaptation parameters that can be mapped onto the *Temporal* axis are the available processing power or battery state. In that case, *Temporal* layers of the scalable multimedia document are ordered in terms of processing requirements. For instance, the timing of an animation can be composed of different *Temporal* layers that define incremental levels of smoothness: from a two-state approach (T_0) to the continuous rendering of complex animations (T_2) as illustrated in Figure 3.

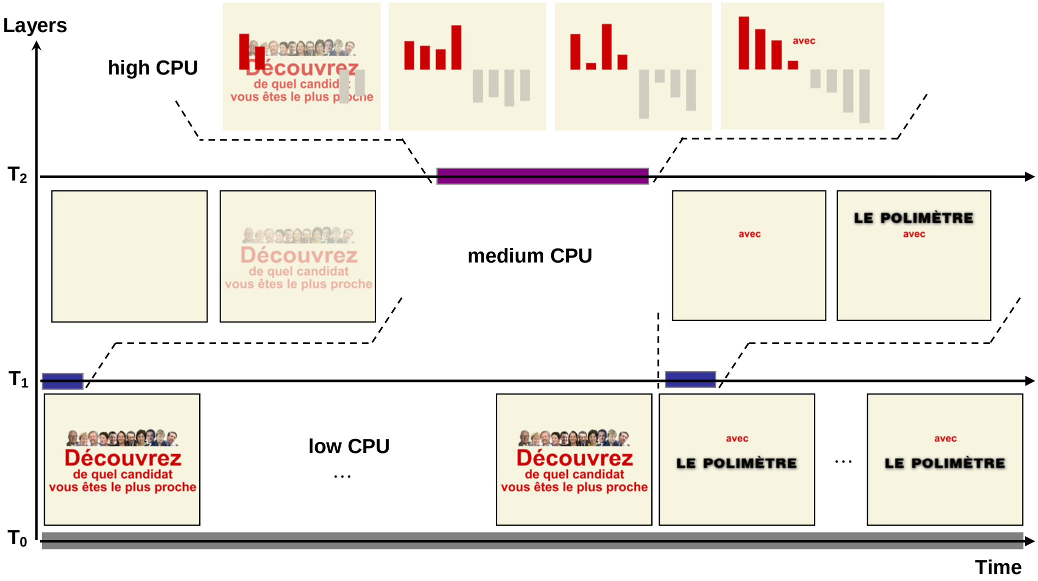
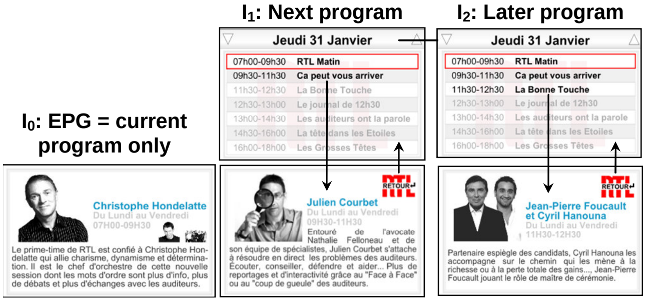


Figure 3: A temporally scalable document driven by processing capabilities

3.2.3 Interactive scalability axis

The *Interactive* component of the *MSTI* model can be divided into layers by progressively providing interactivity functions to the user. An adaptation parameter that can be mapped onto the *Interactive* axis is the number of media elements required to be loaded when accessing to auxiliary media elements through interactions. For instance, accessible documents through user interactions can be prioritize according to author's wishes and defined in incremental *Interactive* layers *(I_0,I_1,I_2)* so that memory requirements are progressively evaluated during playback as illustrated in Figure 4.



07h00-09h30	RTL Matin
09h30-11h30	Ca peut vous arriver
11h30-12h30	La Bonne Touche
12h30-13h00	Le journal de 12h30
13h00-14h30	Les auditeurs ont la parole
14h30-16h00	La tête dans les Etoiles
16h00-18h00	Les Grosses Têtes

Christophe Hondelatte
Du Lundi au Vendredi
07H00-09H30

Le prime-time de RTL est confié à Christophe Hondelatte qui allie charisme, dynamisme et détermination. Il est le chef d'orchestre de cette nouvelle session dont les mots d'ordre sont plus d'info, plus de débats et plus d'échanges avec les auditeurs.

Julien Courbet
Du Lundi au Vendredi
09H30-11H30

Entouré de l'avocate Nathalie Felloneau et de son équipe de spécialistes, Julien Courbet s'attache à résoudre en direct les problèmes des auditeurs. Écouter, conseiller, défendre et aider... Plus de reportages et d'interactivité grâce au "Face à Face" ou au "coup de gueule" des auditeurs.

Jean-Pierre Foucault et Cyril Hanouna
Du Lundi au Vendredi
11H30-12H30

Partenaire espiègle des candidats, Cyril Hanouna les accompagne sur le chemin qui les mène à la richesse ou à la perte totale des gains..., Jean-Pierre Foucault jouant le rôle de maître de cérémonie.

Figure 4: An interactively scalable document driven by memory capacities

3.3 Adaptation graph

Whatever adaptation parameters is associated with the *Spatial*, *Temporal* and *Interactive* scalability axes, a scalable multimedia document always has a generic structure composed of an initial *Base* document and *STI* enhancement layers. In the *Scalable MSTI* model, the *Base* presentation of this document is composed of the lowest *STI* layers (S_0,T_0,I_0) and is designed to address the most constrained usage environments. Starting from the *Base* presentation, *STI* enhancement layers can be successively applied on each scalability axis. The relationships between *STI* enhancement layers can be represented by an adaptation graph as illustrated in Figure

5. A path in the graph defines the adapted presentation of the document that results from the selection of incremental layers according to usage conditions. When an adaptation path reaches an end, the corresponding presentation is called a *Complete* document. The adaptation graph example provided in Figure 5 shows a scalable multimedia document composed of four *Spatial* layers, one *Temporal* layer and two *Interactive* layers.

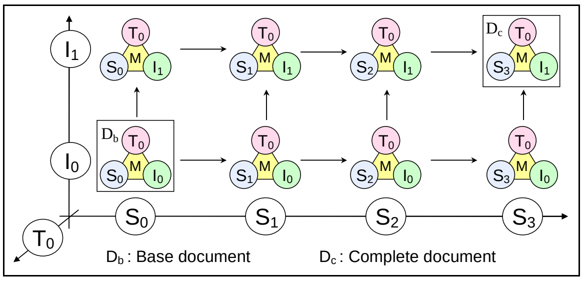


Figure 5. An adaptation graph example and its adaptation paths

3.4 Scalable STI composition

The transformation that consists in applying the *Spatial*, *Temporal* and *Interactive* descriptions to a *Media* description to create presentable multimedia documents is called the *STI* composition. The division of *STI* components into scalable layers requires an incremental or cascaded transformation of the *Media* description to progressively compose the adapted multimedia document. In fact, the cascading of *STI* compositions is possible because an *MSTI* document can be considered as a *Media* description as far as *STI* composition is concerned. Indeed, as illustrated in [20], they are both described in the same format and contain the same identifiers referenced by *STI* layers. The cascading of *STI* compositions is illustrated in Figure 6. It can be noticed on this figure, that the adaptation of scalable multimedia document is driven by presentation aspects and consists in a straightforward adaptive filtering of *STI* layers that is operated by matching their adaptation parameter values (APV) with usage conditions.

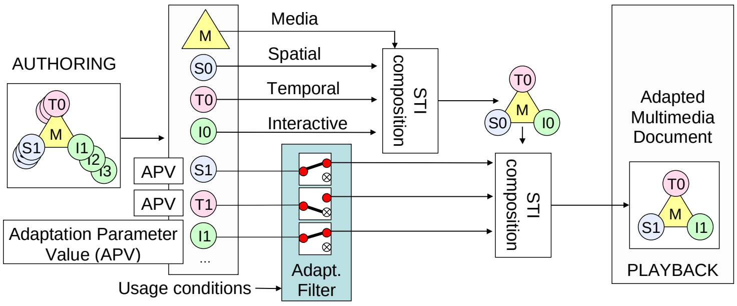


Figure 6. Cascaded STI composition of a Scalable MSTI document

4 Authoring of scalable documents

Section 3 gives an overview of the scalability properties and related adaptation features that the *Scalable MSTI* model, as defined in [20], can offer. In this section, we present authoring techniques that address the new challenges raised by the production of adaptable presentations based on this model. In the following, four major steps in the authoring of scalable multimedia documents are identified and described (see Figure 7). The first step (step 1) follows traditional authoring paradigms: it consists in selecting media elements and designing an appropriate presentation that is simple enough to fit the most constrained usage environments. This initial authoring phase aims at generating the *Base* presentation of a scalable multimedia document where the presentation is split into *Spatial*, *Temporal* and *Interactive* descriptions. To create such *Base* presentation, we present in Section 4.1 a classification of the presentation properties of the XML-based SVG, BIFS and SMIL multimedia languages. The second authoring step (step 2) adds one enhancement layer for each *STI* axis to create a *Complete* presentation that targets the most advanced usage conditions. General authoring guidelines to produce such coarse-grained scalable multimedia documents are described and illustrated on a Digital Radio multimedia service in Section 4.2. The two last authoring steps aim at giving adaptation granularity to the generated multimedia documents in order to address a large set of usage environments. They consist in defining targeted adaptation scenarios (step 3) and designing incremental presentation versions that will ensure an appropriate presentation in most usage environments (step 4). These authoring guidelines and their adaptation properties related to *Scalable MSTI* model are described in Section 4.3 and Section 4.4 respectively.

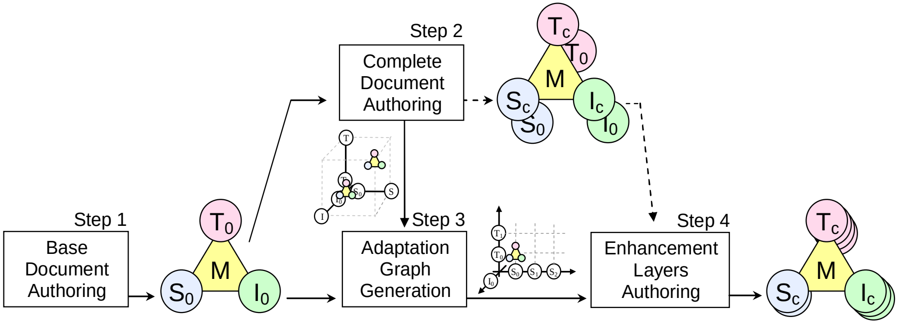


Figure 7: Four-step authoring of a scalable multimedia document

4.1 Step1: Base document authoring

A multimedia presentation encompasses composition properties, which organize in space and time media elements into an interactive presentation, but also styling properties which make the content correspond to a graphical charter, when applied to the media elements. The generation of a non-adaptable document requires its presentation properties to be compatible with all targeted usage environments and in particular the most constrained usage conditions. As a consequence, such one-content-fits-all strategy does not allow the design of an advanced presentation but guaranty minimal content usability in various usage environments. When designing a scalable multimedia document, such content is named a *Base* document and is composed of the *Base Spatial*, *Temporal* and *Interactive* layers. The *Base STI* layers of a scalable multimedia document are always required because they guaranty a fallback presentation in very limited environments. The authoring of this *Base* document does not rely on any adaptation principles but require the separation of spatial, temporal and interactive presentation properties according to the *MSTI* model. In the following, several authoring guidelines are provided to properly classify XML elements and attributes into a media description (Section 4.1.1), styling properties (Section 4.1.2) and composition properties (Section 4.1.3).

4.1.1 Media description

The initial phase in the authoring of a multimedia document consists in selecting key media elements that will be composed into a presentation. The *MSTI* model separates media elements from presentation properties. Indeed, this presentation-agnostic view of a document is a common authoring practice: e.g. defining the layout of web pages using W3C Cascading Style Sheets; or using SMIL Time Sheet to apply temporal properties to an existing document or using an external JavaScript to add behavior. In the *MSTI* model, selected media elements and their metadata are described (or referenced) in the *Media* description that constitute a structured XML document that can be queried and repurposed independently from its presentation. With such a structure, media elements can be associated with metadata such as defined in W3C Ressource Description Framework[2] (RDF) in order to ease semantic search and can also be used in several application scenarios where their presentation will be different (possibly using the <*defs*> element in SVG). An example of media description is given in Figure 8.

[2] http://www.w3.org/RDF/

```
<svg
xmlns="http://www.w3.org/2000/svg"
version="1.2" baseProfile="tiny"/>

  <metadata id="meta-rdf">
   <rdf:RDF
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:dc = http://purl.org/dc/elements/1.1/>
    <rdf:Description
     dc:date="2008-10-04"
     dc:publisher="RTL"
     dc:language="fr"
     dc:Title="DMB Digital Radio at Paris Motor Show 2008">
    </rdf:Description>
   </rdf:RDF>
  </metadata>

  <g id="g_0">
   <rect id="rect_0"/>
   <video id="video_0" xlink:href="salonAutoRTL.h264"/>
   <img id="img_0" xlink:href="koleos-RTLNumerique.jpg"/>
   <text id="text_0" xlink:href="pressRelease.svg#text"/>
  </g>
</svg>
```

Figure 8: SVG Media description including RDF metadata

4.1.2 Styling properties

Styling properties[3] are decorative features that are applied to media elements (e.g. images, text or polygons) and that determine the overall aspect of the presentation which is usually based on a graphical charter. From an authoring point of view, the *MSTI* model is very lenient about styling properties because they do not have any consequences on spatial, temporal or interactive properties of the *MSTI* model. Two main reasons have driven our choice to not mandate a specific style description. First, some styling properties can be closely linked to the semantics of the presentation and they cannot be easily replaced without reconsidering the entire presentation or applied to another document without making sure that both documents have similar semantics. For instance, two key media elements (e.g. an image and its caption) might be semantically linked in a presentation by encompassing them in rectangle with a border style that materializes their close relationship. Such a border style would not be applied to the rectangles of another *Media* description unless similar media relationships were identified. A second reason is that styling properties are often defined according to a particular layout or designed to create a visual effect that is dynamically triggered through interactive or timed events. For instance, the width of a rectangle's border differs depending on its size to match human visual system and the color of a button will be different if it is pressed or not [10]. As for now, we suggest to define all styling properties of media elements in the *Spatial*[4] description of *MSTI* document except for those that are dynamically or interactively assigned and that can be defined respectively in the *Temporal* or *Interactive* descriptions. An example of static and dynamic styling properties assignment is given in Figure 9 where the *Spatial* and *Temporal* description can be applied to the *Media* description provided in Figure 8

```
<Spatial>
   <Replace ref_id="rect_0" attribute="fill">blue</Replace>
   <Replace ref_id="rect_0" attribute="stroke">green</Replace>
   <Replace ref_id="rect_0" attribute="stroke-width">4</Replace>
</Spatial>
```

```
<Temporal>
  <Insert ref_id="text_0">
     <animate id="animate_text_0_color"
        attributeName="fill" from="blue "to="green"
        repeatCount="indefinite" begin="0s" dur="2s" fill="freeze"/>
   </Insert>
</Temporal>
```

Figure 9: Styling properties in Spatial and Temporal descriptions

[3] This definition of styling properties does not cover layout properties that can be specified using Cascading Style Sheet (CSS).

[4] The term *Spatial* could be renamed *Perceptual* but for consistency reasons, the term *Spatial* is preferred.

A non-exhaustive list of styling properties for SVG [1], BIFS [11] and SMIL [6] is given in Table 1 and gives an overview of the XML elements/attributes that can be considered as styling properties when authoring *Scalable MSTI* documents.

Table 1: Styling properties of SVG, BIFS and SMIL languages

SVG
The <solidColor>, <linearGradient>, <radialGradient> elements The <font> and <font-face> elements fill, fill-rule, fill-opacity, stroke, stroke-width, stroke-linecap, stroke-linejoin, stroke-miterlimit, stroke-dasharray, stroke-dashoffset, stroke-opacity, viewport-fill and viewport-fill-opacity attributes
BIFS
The <Material2D> element*(emissiveColor, filled, lineProps, transparency*) <Color>, <LineProperties>, <LinearGradient>, <RadialGradient> elements The <FontStyle> elements <Background2D> attribute: *backColor*.
SMIL
<region> attributes: *backgroundColor* and *showBackground* <root-layout> and <topLayout> attributes: *backgroundColor*

4.1.3 Composition properties

The composition properties of a presentation define the “choreography” in space and time of media elements and according to user interaction. Once media elements have been selected, an important part of the authoring efforts is concentrated in assigning composition properties to these media elements to achieve a coherent, usable and attractive multimedia presentation. Every content creator has its own technique to achieve these high-level objectives but all composition properties must be divided into *Spatial*, *Temporal* and *Interactive (STI)* descriptions when authoring *MSTI* documents. The authoring of the *STI* components of *Base* documents does not have to be performed by the content creator himself because an extraction of the spatial, temporal and interactive properties of authored documents can be automatically performed for XML-based multimedia languages. This section provides the results of our experiments that have been conducted on SVG [1], BIFS [11] and SMIL [6] and shows how the elements and attributes of these XML-based languages can be successfully separated into *STI* descriptions. The classification, we are proposing, is partially[5] given in Table 2, Table 3 and in Table 4.

This informative *STI* classification of the SVG, BIFS and SMIL multimedia languages tries; as much as possible; to create independent *STI* descriptions by gathering related XML elements and attributes into the same *STI* component. Most of the elements and attributes of the XML-based multimedia standards we have studied can be easily classified in *STI* descriptions. However, some difficulties and exceptions have been identified and are discussed in the following:

- Multimedia formats often offer generic syntax in order to improve language learning time or to enable coding efficiency. Such generic elements can be used in all three *STI* components but we chose to place them in the description they related to: *Spatial*, *Temporal* or *Interactive*. Typical examples are the multipurpose *<script>* elements of SVG which modify *STI* descriptions, the *<Route>* element of BIFS which can link *STI* events and the *begin* attribute of the SMIL animation elements which can contains time or interactive values. Any element that does not fit into one single *STI* component due to its usage context is then placed in the *Media* description.

- It is sometimes necessary to carefully extract the attributes of a single element and to assign them different *STI* descriptions. The animation of the position of an object in SVG or SMIL is a perfect illustration of this point. An animation requires targeting a parameter of the presentation (e.g. the position of an object). It also defines the initial position, the path of the animation and the final position, which are obviously part of the *Spatial* description because they depend on the spatial layout of the document. The timing of the animation is part of the *Temporal* description (start time, duration, end time). Workarounds to such STI component dependencies have been proposed in [20] and basically consists in sharing presentation states in the *Media* description so that *STI*

[5] The full results of our experiments can be found at:
http://www.tsi.enst.fr/mm/MSTI/MMPropertyClassification.html

components never link directly to each other. However, animations do not always relate to spatial properties as in the above example and can also be used to enable/disable the focus on a media element over time. As a consequence, the *attributeName* attribute of a *<set>* SVG element cannot be defined as part of the *Spatial* or the *Interactive* description and will be classified according to the value it has.

- Some multimedia formats define optimizations properties that are essential to improve usability on specific usage environments: the *image-rendering* attribute in SVG allows indicating author preference between quality and speed of rendering; or the <*QuantizationParameter*> element in BIFS enables authoring the trade-off between quality and bandwidth. These properties are not directly related to *STI* components (and their support is often optional on receiver's side) but have an impact on applicable usage conditions for *STI* descriptions. Therefore, these parameters are classified according to the adaptation parameters applied to *STI* descriptions: e.g. screen resolution, processing power and interactive means. With such requirements, quantification parameters can be defined in the *Spatial* description because they have a direct impact on the highest acceptable screen resolutions (based on a minimum perceived quality criteria) and rendering optimization would be part of the *Temporal* description.

Table 2: Spatial properties of SVG, BIFS and SMIL languages

SVG
Positioning attributes: transform, x, y, preserveAspectRatio, viewbox, rotate, text-anchor, translate **Sizing attributes:** width, height, scale, skewX, skewY, font-size **Visibility attributes:** display, visibility, initialVisibility <animateTransform> attributes: type, from, to <animation> attributes: x, y, width, height, preserveRatio,
BIFS
Positioning attributes: *translation* **Sizing attributes:** *rotationAngle, scale, scaleOrientation, radius* **Visibility attributes:** transparency, whichChoice The <PathLayout> element and The <IndexedLineSet2D> attributes:*coord, colorIndex, colorPerVertex, coordIndex* <ScalarInterpolator>, <CoordinateInterpolator2D>, <PositionInterpolator2D> attributes: *key* and *keyValue* <ScalarAnimator>, <PositionAnimator2D>: *fromTo, key, keyType, keySpline, keyValue, keyOrientation, weight, keyValueType, offset*
SMIL
Positioning attributes: top, bottom, left, right, width, height, fit, z-index, regPoint, regAlign, mediaAlign, soundAlign <area> attributes: shape, coords <animate> attributes: from, to, by, accumulate, additive, calcMode, values, path, keySplines <transition> attributes: type, subtype, fadeColor, horzRepeat, vertRepeat, borderWidth, borderColor and <param> element (name, value)

Table 3: Temporal properties of SVG, BIFS and SMIL languages

SVG
The <animate>, <set>, <animateMotion>, <animateColor>, <animateTransform> elements The <discard> elements **Attributes to control the timing of the animation**: begin, dur, end, min, max, restart, repeatCount, repeatDur, fill **Time-related events**: beginEvent, endEvent, repeatEvent, SVGTimer
BIFS
The <TimeSensor> (cycleInterval, loop, startTime, stopTime) element The <TemporalTransform> element (startTime, optimalDuration, active, speed, scalability, stretchMode, shrinkMode, maxDelay) The <TemporalGroup> elements (costart, coend, meet) <MovieTexture> and <AnimationStream> attributes: loop, speed, startTime,

stopTime, isActive <Layout> attributes: scrollRate, smoothScroll, loop, scrollVertical scrollMode
SMIL
The <animate>, <set>, <animateMotion>, <animateColor> elements The <seq>, <excl>, <par> elements The <prefetch> element **Attributes to control the timing of the animation**: begin, end, dur, repeatCount, repeatDur, keyTimes, accelerate, decelerate, speed, **media attributes**: mediaRepeat, clipBegin, clipEnd, min, max, restart, restartDefault **Attributes that define animation values over time:** dur, startProgress, endProgress, direction, transIn, transOut **Time-related events**: beginEvent, endEvent, repeatEvent, repeat

Table 4: Interactive properties of SVG, BIFS and SMIL languages

SVG
The <a> element **Graphics and text content elements attributes:** focusable, nav-next, nav-prev, nav-up, nav-up-right, nav-right, nav-down-right, nav-down, nav-down-left, nav-left, nav-up-left, focusHighlight **Interactive events:** DOMFocusIn, DOMFocusOut, DOMActivate, click, mousedown, mouseup, mouseover, mousemove, mouseout, mousewheel, textInput, keydown, keyup, scroll The <handler> and <listener> elements <text> and <textarea> attribute: editable
BIFS
<Anchor>, <TouchSensor> and <PlaneSensor2D> elements <InputSensor> elements (KeySensor, StringSensor, Mouse type)
SMIL
<a> and <area> elements (show, target, accesskey, tabindex, fragment, alt) **Media attribute:** sensitivity **User-related events**: focusInEvent, focusOutEvent, activateEvent, inBoundsEvent, outOfBoundsEvent

4.2 Step2: Complete document authoring

The authoring guidelines described in Section 4.1 (step1) aim at helping the generation of a valid Scalable *MSTI* documents that is designed to address a large heterogeneity of usage conditions. In fact, such MSTI document does not feature any adaptation functionalities but is suitable for the most constrained environments because of important limitations in presentation functionalities. Based on this initial document, the second authoring phase described in this section gives authoring guidelines to generate adaptable *MSTI* documents using the *Scalable MSTI* model. The authoring of these adaptable multimedia documents consists in the generation of a *Complete* document that enhance the *Base* document as illustrated in Figure 10. This authoring relies on the choice of three adaptation parameters that are associated to the *Spatial*, *Temporal* and *Interactive* axes respectively. In fact, *MSTI* document produced during this authoring phase are scalable multimedia documents where each *STI* scalability axis is (possibly) composed of two layers: a *Base* layer (S_0,T_0,I_0) and a *Complete* layer (S_c,T_c,I_c). Even though these scalable multimedia documents are simple, an appropriate authoring of *STI* description is required to later-on leverage the adaptation functionalities offered by the *Scalable MSTI* model. These general authoring principles are described in Section 4.2.1 and illustrated in Section 4.2.2 on a multimedia Digital Radio example.

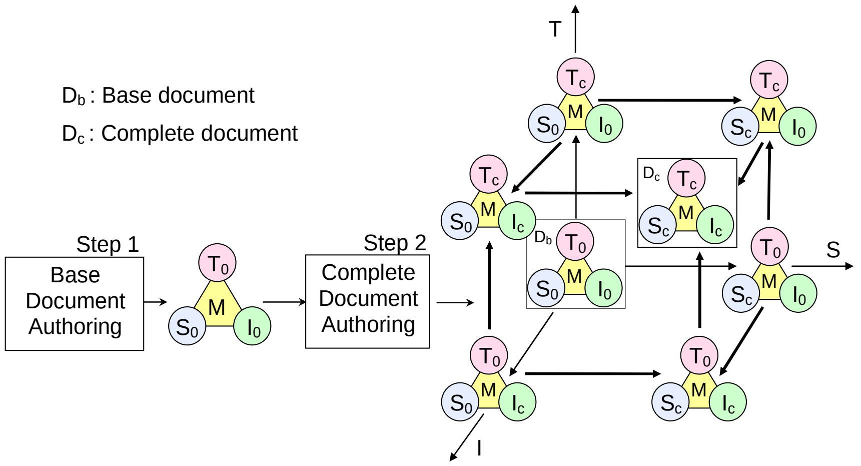


Figure 10: From a Base Document to a 2-layer scalable multimedia document

4.2.1 Simple document adaptability

In our approach, the generation of adaptable multimedia document is driven by authoring principles that ensure a simple and controlled adaptation of the content. This approach is generic because it does not make any assumption on the adaptation features of targeted formats: system attributes (e.g. *systemBitrate*) in SMIL, the *requiredFeatures* attribute in SVG or the *<TermCap>* element in BIFS Instead, the *Scalable MSTI* approach only relies on a straightforward filtering adaptation process that can be progressively performed during the playback of the document either based on specific format capabilities (such as listed above) or through an external adaptation engine (MPEG-21 DIA [22], W3C XSL Transformations (XSLT), JavaScript code…). Additionally, the adaptation process of a scalable multimedia document can be dynamic. In the case of increasing capabilities, it just consists in continuing the execution of *STI* layers. In the case of decreasing capabilities, Random Access Layers (RAL) can be used to revert to a previous state without having to restart the decoding of the adaptable document from the beginning. The authoring of such layers is described in Section 4.3.2. In practice, three essential authoring principles shall be respected to leverage the intrinsic adaptation properties of scalable multimedia documents:

- A scalable document should not assume any temporal or interactive properties to be visualized. Indeed, although a *Complete* multimedia presentation may heavily rely on animations or interactive items, the *Spatial* description of a scalable document should be self-working. This simple view of a multimedia document is very important when authoring adaptable document because it might be the only possible version that will fit highly constrained environments. If this guideline cannot be fully respected, it means that the *Spatial* description either depends on the *Temporal*, *Interactive* or both. In that case, the constraints layer attributes of the *Scalable MSTI* model (*dependsOn{S|T|I}LayerValue* and *requiredFor{S|T|I}LayerValue*) that are described in [20] should be used to define dependent layers that will reduce possible paths in the adaptation graph.

- A scalable document should provide all possible means to access media elements. In most cases, all media elements available in the presentation cannot be directly visualized. Some media elements may require the user to wait for some time before being visible and some others are only visible through user interaction. Since there is no guaranty that the temporal properties (visual activations over time) or the interactive properties (visual activation by user interaction) can be activated in the playback usage environment, it is important that content access to media elements is possible either way. For instance, a workaround to the lack of an appropriate timing module on a terminal would be a scalable document where the *Spatial* and *Interactive Complete* layers are activated whereas the playback of the same presentation will still be possible on receivers without a user interface by activating only the *Spatial* and the *Temporal Complete* layers of the scalable document. If this guideline is not taken into account, the generated scalable multimedia document still remain valid but will not offer an optimal user experience in all adaptation scenarios.

- A scalable document should take advantage of the cooperation between temporal and interactive properties. When authoring a scalable multimedia document, one option consists in defining two independent access paradigms: automated mode (based on a timeline) or user-centric mode (based on interactions). However, the *Scalable MSTI* model does not mandate such an opposition. On the contrary, a scalable document featuring its full capabilities on the *Temporal* and on the *Interactive* axis can be authored to provide content enhancements that are only available when it is composed of all its *Complete STI* components. For instance, automated transitions from one media element to another can be performed through animations but the same timed animation can also be used to achieve transition on user requests. When this guideline is ignored, generated scalable documents are perfectly correct but they will not make use of all its possible multimedia functionalities.

4.2.2 Digital Radio multimedia scalable example

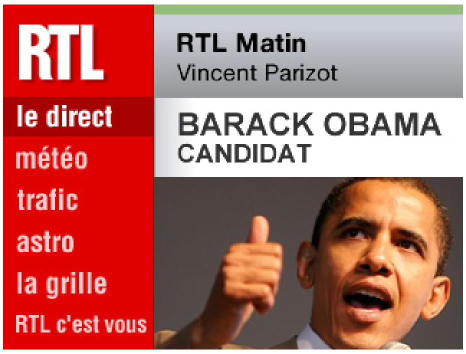


Figure 11: A multimedia Digital Radio program

In order to illustrate the three authoring principles described in Section 4.2.1 on an example, we consider the scalable menu of a DMB Digital Radio program (see a snapshot on Figure 11). The menu items define several topics that are part of the presentation: 'live' (le direct), 'weather forecast' (météo), 'traffic information' (trafic), 'astrology messages' (astro) and a 'program guide' (la grille). This *scalable MPEG-4 BIFS* document has been authored following the principles laid down in Section 4.1 (Step1) and Section 4.2.1 (Step2) so as to ensure content usability in a large number of application scenarios:

- The style and the layout of the presentation have been designed to maximize user experience on typical DMB handheld receivers (e.g. QVGA display). This configuration constitutes the *Base Spatial* layer of the document that could be enhanced by a *Complete Spatial* layer targeting a VGA display for instance. Since, key media elements are part of the 'live' page, the 'live' menu item is activated when connecting to the service. This activation can be noticed thanks to the style of the 'live' menu item (darker color of the background and brighter color of the text). Additionally, the *Base* layers of the *Temporal* and *Interactive* axes are empty since there may be some usage environments where users may want to disable timed sequences (e.g. battery saving) and may not be allowed to interact with the presentation (e.g. in-car use cases).

- The *Temporal* and the *Interactive Complete* descriptions of the scalable document define two alternative mechanisms to access to weather forecast, traffic information, astrology messages and program scheduling. Either the presentation automatically switches from one menu item to another based on a pre-defined timed sequence or the user directly selects topics by clicking on menu items (or using a keypad). The authoring of both *Temporal* and *Interactive* descriptions is important because some receivers may not be able to render animations that notify the user with topic updates (because of processing power shortage) or may not have any interactive means at all.

- Two different types of cooperation between temporal and interactive properties are defined in the menu of the scalable document illustrated in Figure 11. First, the scalable menu item has been authored so as to perform a dynamic switching between the two navigation modes. In particular, a *Complete* presentation of the scalable document featuring its *Complete Spatial*, *Temporal* and *Interactive* descriptions will start an automated slideshow displaying the weather forecast, traffic information, astrology messages and program scheduling only if no interaction has been detected within the last 5 minutes. Second, animations described in the *Temporal* component to activate fading effects on menu items can also be triggered on user interaction.

4.3 Step3: Adaptation graph generation

The two authoring steps described in Section 4.1 (step1) and in Section 4.2 (step2) aim at generating coarse-grained scalable documents which rely on a *Base* presentation that can be enriched by one enhancement layer on each *Spatial*, *Temporal* and *Interactive* (*STI*) scalability layer. This approach is straightforward and provides simple but essential adaptation functionalities. An advanced authoring of the scalability properties proposed by the *Scalable MSTI* model can significantly extend these adaptation possibilities. In fact, these last authoring phases can be seen as a refinement process that inserts enhancement layers between *Base* and *Complete STI* layers. As illustrated in Figure 12, the last authoring phases described in this Section 4.3 (step3) and in Section 4.4 (step4) leverage the adaptation properties of the scalable multimedia document generated as output of (Step 2) by refining its application scenarios.

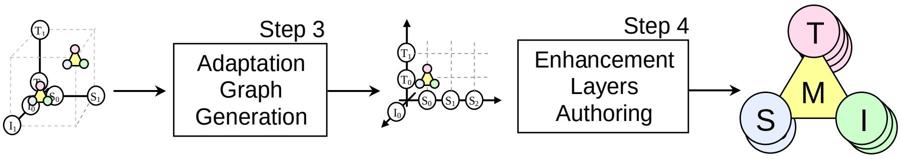


Figure 12: Fine-grained scalable multimedia document authoring

When generating a coarse-grained scalable multimedia document, the content creator only needs to select three essential adaptation parameters that will determine what usage conditions are mandated to have access to a document. When trying to address usage environments with fine-grained scalability, the content creator needs to better foresee possible adaptation scenarios. Unfortunately, content creators may have little knowledge (if any) about the actual usage environments their documents will have to cope with. They are usually asked to address some typical use cases such as some type of terminals (with typical resolutions, processing power capabilities…) or typical delivery mechanisms (with various bandwidths…). In this section, we propose authoring techniques to map these high-level adaptation parameters onto a scalable *STI* graph before starting the actual authoring of the scalable multimedia document. The generation of this adaptation graph requires mapping the three main adaptation parameters onto scalability axes (Section 4.3.1) and to identify key adaptation states (Section 4.3.2). The *Scalable MSTI* also provides tools to cope with exceptions in the adaptation graph. Two exceptions have been identified: when only few values of an adaptation parameter do not fit in the given general ordering; or when no order relationship can be defined for some adaptation parameters. For these exceptions, authors can use *auxiliary* adaptation parameters as introduced in Section 4.3.1 or define additional *digressing* or *forking* adaptation parameters as described in Section 4.3.3 and in Section 4.3.4, respectively.

4.3.1 Scalability and adaptation axes

Scalable multimedia authoring consists in defining and filling the successive layers of *Spatial*, *Temporal* and *Interactive* axes of a *Scalable MSTI* documents. These incremental layers represent improvements of the multimedia presentation but also assume increasing capabilities in users usage environments. Therefore, a specific type of adaptation parameter has to be defined for each *STI* axis. Typical device-oriented examples of adaptation parameters for *STI* axes are: resolution or memory consumption (*Spatial*), processing power (*Temporal*), interactive means or memory consumption (*Interactive*). However, these mappings are not mandated and may be very different from one use case to another. For instance, user-oriented adaptation parameters could be: accessibility to partially-sighted people *(Spatial),* presentation duration *(Temporal)* and level of expertise *(Interactive).* Once an adaptation parameter has been assigned to each *STI* axis, a total order relation needs to be chosen. Given an order relation, chosen adaptation parameter values can be ordered into a progressive manner and mapped onto individual scalability layers. This process determines the expected number of scalability layers for each *STI* axis. It should be noted that parameter values do not have to be accurate (e.g. processing power values). In fact, only the ordering of values is essential since precise parameter values may only be known at playback time (e.g. memory usage depends on the implementation).

When considering adaptation parameters that are mapped onto the three *STI* axes of a scalable multimedia document, it might not be satisfactory to discard some parameters just because there are too many of them. For instance, the three selected adaptation axes might be: resolution (*Spatial*), processing power (*Temporal*) and bandwidth (*Interactive*) but memory requirements and the availability (or not) of a sensitive screen might also be critical for some devices. The *Scalable MSTI* model enables adaptation scenarios to be driven by more than three adaptation parameters. Indeed, several types of adaptation parameters can be assigned to one *STI* layer. In that case, the additional parameters are called *auxiliary* adaptation parameters. As illustrated in Figure 13, such auxiliary adaptation parameter can either be present on a single layer (e.g. sensitive screen applied to the layer I_1 of the *Interactive* axis) or distributed on several layers (e.g. internal memory requirements that depends on *Spatial* and *Interactive* scalability axes, e.g. I_1 and S_2). However, it should be noted that these *auxiliary* adaptation parameters are only adaptation hints and that the *Scalable MSTI* model may not provide the optimal presentation when considering them. In fact, the *Scalable MSTI* model guaranties that the best presentation is selected only when the adaptation process is driven by the three scalable adaptation parameters. For instance, the auxiliary memory value (RAM) provided in Figure 13 indicates a measure corresponding to (S_2,T_0,I_1). As a consequence, (S_3,T_0,I_0) might be the optimal presentation for a device with a 480x320 resolution, an 8 kbps bandwidth and 8 Mbytes of available RAM but an adaptation engine, which cares about memory, could stop at (S_2,T_0,I_0) since S_2 is labelled with a 8 Mbytes RAM auxiliary adaptation parameter. Therefore, *auxiliary* adaptation parameters can be used to guide the adaptation of scalable multimedia documents with much more than the three parameters defined as the adaptation axes of the scalable multimedia document but they also introduce limitations in some usage environments.

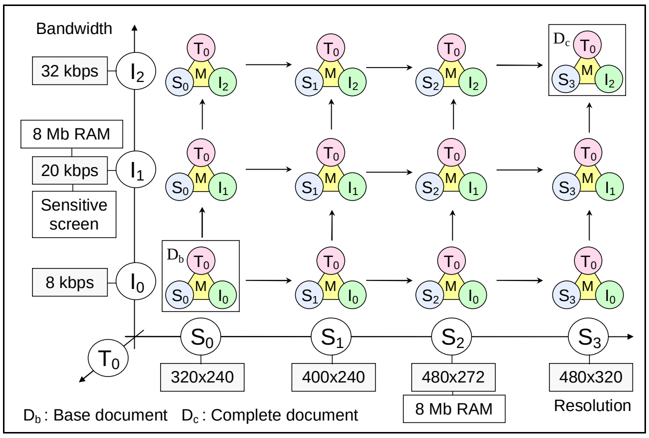


Figure 13: Scalable adaptation axes example

4.3.2 Key adaptation point

The adaptation graph of a scalable multimedia document defines layers for each *Spatial*, *Temporal* and *Interactive* (*STI)* axis. In fact, it defines a set of incremental presentation versions that can be selected according to usage environments. This set of presentation versions does not aim to be optimal for all usage configurations but the ordering of the three main adaptation parameters and the mapping between progressive parameters values and document *STI* layers guaranty that a suitable presentation can be found in any case, given that the *Base* scalable document addresses the lowest possible capabilities. However, content creators are usually asked to focus on some specific usage configurations that must be optimally addressed (e.g. the iPhone device). Such configuration must then be addressed by a point in the adaptation graph and our approach guaranty that the customized document will be reached. However, because the author might want to ease or favour such targeted players when they perform their navigation in the adaptation path, the *Scalable MSTI* model lets content creators indicate key points by using Random Access Layers as described in [20].

From a playback point of view, the Random Access Layer indication helps in two situations. First, if the path to this key point is long in the adaptation graph, a fast and direct access to the presentation is enabled. Second, if the receiver follows a blind strategy to navigate in the adaptation graph, by processing the incremental *STI* layers of the scalable document one after another without matching its capabilities with their adaptation parameters, key points enable fallbacks that avoid the reloading of the whole document when a maximum layer has been identified. The added value of the RAL tool is illustrated in Figure 14. However, it should be noted that Random Access Layers reduce the coding performances as explained in [20] because of information redundancy.

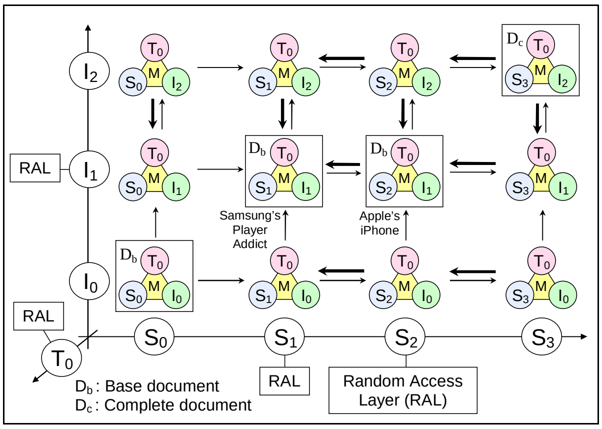


Figure 14: Identifying key adaptation states

4.3.3 Digressing adaptation parameter values

When preparing the adaptation graph of a scalable multimedia document, three main adaptation parameters are selected with their order relations, a set of progressive parameters values are defined and mapped onto the scalability layers of the *Scalable MSTI* model. However, it may happen that one specific parameter value does no fit with the overall ordering of its adaptation axis. For instance, a scalable multimedia document may target several resolutions ordered first by their horizontal dimension: (S_0)240x240, (S_1)320x240, (S_2)400x240, (S_3)480x272 and (S_4)480x320. However, the content creator may be asked to provide a vertical layout for one of these targeted resolutions (e.g. 400x240) because some receivers have the ability to dynamically switch from an horizontal layout to a vertical layout (e.g. Samsung's Player Addict). Such vertical resolution (e.g. 240x400) does not fit with the ordering of the resolution-based adaptation axis but the presentation associated with this specific configuration may roughly look like some other presentations available in the adaptation graph (S_2 in our example). In that case, this adaptation parameter value can be defined as a digression in the adaptation paths. In the *Scalable MSTI* model such *digressing* adaptation parameter values are signalled by *skippable* layers as introduced in [20] and illustrated in Figure 15. In fact, these layers related to a digressing adaptation parameter value depend on their preceding layers, may be applied (or not) and are still connected to the main path. From an authoring point of view, digressing adaptation parameter values provide additional flexibility as far as adaptation is concerned since they increase the number of possible paths in the adaptation graph. However, *skippable* layers cause some redundancy (in S_3 in the example) and should therefore be used carefully.

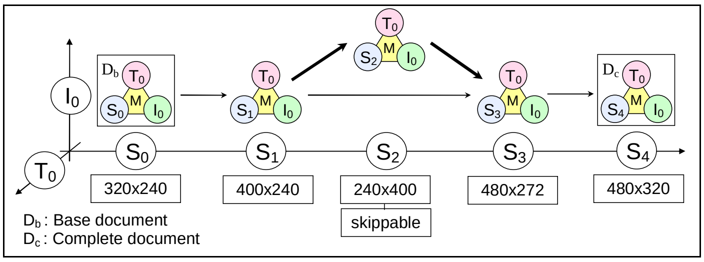


Figure 15: A digression in a one-way adaptation graph

4.3.4 Forking adaptation parameter values

The *Scalable MSTI* model enables three main incremental adaptation axes. All adaptation parameters that cannot be mapped onto an existing presentation or onto a presentation similar to an existing one in the adaptation graph cannot be simply defined as *auxiliary* adaptation parameters or *digressing* values as introduced in Section 4.3.1 and Section 4.3.3 respectively. Instead, a fork needs to be defined in order to extend the adaptation graph with presentation exceptions that are entirely incompatible with the selected scalable adaptation parameters. In fact, *forking* adaptation parameter values are mapped onto scalable layers that depend on their preceding layers, that may be applied (or not), as *digressing* adaptation parameter values do, but that cannot be updated to return to the main adaptation path. Such isolated *Spatial, Temporal* or *Interactive (STI)* layers can be used when further enhancements are not envisioned or if embedded presentation properties would significantly impact succeeding layers (and therefore introduce undesirable redundancy).

Technically, these isolated layers can be defined by the *Scalable MSTI* model through the combined use of *skippable* and *blocking* layers or originate from diverging adaptation paths associated with a combination of *skippable* and *dependent* layers that were individually introduced in [20]. Both approaches are illustrated in the following from an authoring point of view.

First, a *forking* adaptation parameter can be related to a very specific usage configuration that requires a customized presentation which clearly differs from the *Complete* presentation of scalable multimedia document. For instance, a typical adaptation axis for the *Interactive* layers can provide progressive and complementary interactive means: 2-state keypad (left-right), touch screen (click), 5-state keypad (left-right-up-down-ok), trackball (isOver) as illustrated in Figure 16. Integrating the iPhone's multi-touch navigation paradigm as a *digressing* adaptation parameter in such an adaptation graph is possible but would not be fitted because the whole interactivity mechanism will be very different. Indeed, the iPhone's interface releases left-right-up-down button requirements and very specific interactive behaviours can be triggered by the built-in accelerometer. As illustrated Figure 16, such specific adaptation parameter values can be mapped onto layers identified as *skippable* and *self-blocking*.

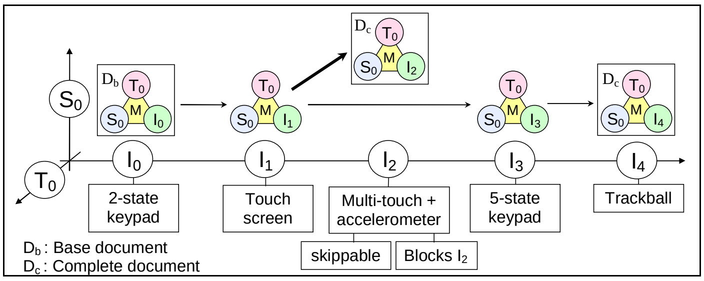


Figure 16: A short fork in a one-way adaptation graph

Second, forking adaptation parameters can also be defined to initiate a scalable alternative to the main adaptation axis. In fact, one (or more) forks could be defined in parallel to their main adaptation axis in order to define alternative scalable versions of the same document. One typical example of such forking adaptation parameter is document internationalization. Indeed, multi-lingual support cannot be considered as an adaptation axis since an order relation cannot be defined between several languages. However, language selection has significant consequences on document layout that would impact the *Spatial* component of a scalable multimedia document. To cope with such a use case, a language-based fork can be defined in the adaptation graph as illustrated in Figure 17. In that case, this *forking* adaptation parameter would be mapped onto layers identified as *dependent* and *skippable*. It should be noted the loose definition of layers dependency introduced in [20] is here restricted to immediate dependencies. This restriction reduces possible adaptation paths but simplifies the playback of *Scalable MSTI* document since only the last applied layer needs to be remembered.

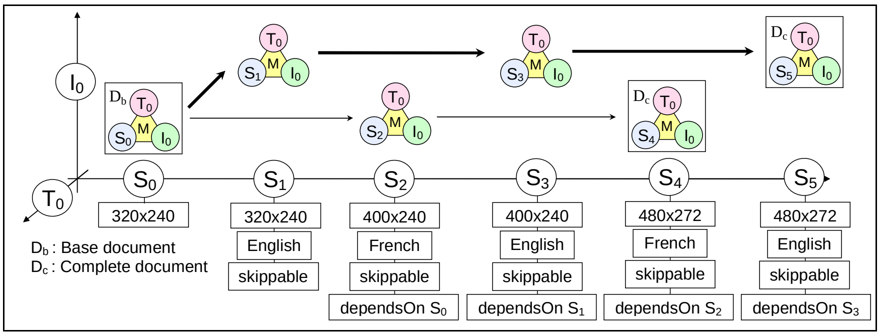


Figure 17: A long fork in a one-way adaptation graph

4.4 Step4: authoring STI enhancement layers

The last authoring phase of a scalable multimedia document is all about creativity. Indeed, it technically consists in filling *STI* layers with presentation properties that are compatible with the adaptation parameter values they are assigned by the adaptation graph. This authoring phase does not require any specific knowledge about the *Scalable MSTI* model and should be guided by the general authoring principles provided in Section 4.1 and in Section 4.2. Additionally, the *Scalable*

MSTI model entirely relies on existing multimedia languages such as SVG, BIFS or SMIL. *Scalable MSTI* documents can therefore be authored by advanced content creators who sometimes prefer to use an XML editor to control the generated code for those standardized languages or through an authoring tool that features a traditional user interface used to manipulate such formats.

This content production phase can be driven by the *Complete* document generated during the coarse-grained scalable multimedia document generation (Step 2). However, the preliminary authoring of such coarse-grained scalable multimedia document is absolutely not required to achieve fine-grained scalability. Indeed, an incremental authoring of the presentation properties of progressive layers of the adaptation graph is an authoring approach that also makes sense for designers [10][19].

5 Conclusion and discussions

In this paper, we have briefly recapped the *Scalable MSTI* model and presented authoring techniques that address the new challenges raised by the production of scalable presentations based on this model. First, the authoring of a scalable multimedia document is initiated by the creation of a *Base* document that clearly separate *Spatial*, *Temporal* and *Interactive* descriptions by referring to media elements grouped in the *Media* description. This *Base* document is essential because it aims at providing a multimedia presentation simple enough to address the most constrained usage environments. Starting from this *Base* document, the authoring of a scalable multimedia document consists in defining enhancement layers that will enrich this *Base* presentation according to usage conditions. As a consequence, a set of adaptation scenarios must be envisioned during the authoring phase. The core of our proposal is to design a scalable adaptation graph that is formed of three adaptation axes composed of ordered adaptation parameters. In order to efficiently fulfill the largest number of adaptation scenarios, some additions to the main scalability features of the adaptation graph can even be authored by defining *auxiliary*, *digressing* and *forking* adaptation parameters values. This scalable adaptation graph is important because it determines the adaptation properties of the scalable multimedia document. Starting from this adaptation graph, the author creativity is now required to design *Spatial*, *Temporal* and *Interactive* (*STI*) layers that will determine the document presentation in various environments.

As previously said in the state-of-the-art section, our authoring approach is clearly driven by the strong opinion that multimedia adaptation cannot be driven by the media level unless only austere presentations are to be proposed to the user. Therefore, we propose authoring techniques that defines the multimedia scene as the mainstay of the document. In fact, a Rich-Media presentation can include key media elements (e.g. video, audio, images, text) but also graphical elements (e.g. rectangles, circles, polygons…) with multiple styling properties and flexible layouts that can vary over time and depend on user interactions. Such advanced multimedia scene therefore requires an important authoring effort compared to straightforward presentations but significantly improves the perceived quality of the whole content. For instance, these high-quality requirements for multimedia services are reality on internet through Flash web sites, on 3G mobile phones through LASeR on demand multimedia services and on DMB Radio receivers through BIFS interactive broadcast program. The authoring techniques proposed in this paper clearly target the production of such advanced multimedia documents and take as essential requirement the guarantee of a complete control of the quality of the adapted presentations during its lifetime. We believe that this additional authoring cost for adaptation is mandated because any adaptation technique that do not offer a transparent control over generated presentations will simply fail to achieve the high-quality expectations of a designer. Hence, our approach combines the definition of the adaptive behavior of a presentation and its authoring process in order to minimize authoring complexity and leverage the incremental design of a multimedia document. Obviously, our adaptation authoring techniques can also be used to produce simple adaptable presentations (e.g. Audio/Video only presentations) or highly structured presentations in languages like XHTML. For sure, our approach will not be the simpler option from an authoring point of view in that case and that is the reason why multimedia document scalability may be combined with other adaptation techniques such as constraint-based and interpolation-based approaches depending on application scenarios.

This paper presented general authoring principles that are essential when producing *Scalable MSTI* documents and authoring guidelines that result from the underlying properties of the *Scalable MSTI* model. It did not aim at describing an authoring tool for scalable multimedia document generation. Nevertheless, we have started the development of a WYSIWYG scalable document authoring tool and some important issues can already be raised: how to visually represent the three dimensions of a scalable multimedia model? How to efficiently help authors in managing deltas between layers? How to extend the notion of scalability to template-based authoring? How to compose scalable sub-presentations? This will be the objectives of our future work, in particular considering incremental transformations [30].

6 References

[1] Andersson O, Berjon R, Dahlström E et al. (2006) Scalable Vector Graphics (SVG) Tiny 1.2 Specification. W3C, http://www.w3.org/TR/SVGMobile12

[2] Bes F, Jourdan M, Khantache F (2001) A generic architecture for automated construction of multimedia presentation. In: Proc of the 8th conference on Multimedia Modeling (MMM), Amsterdam, Holland, pp. 229-246

[3] Bilasco IM, Gensel J, Villanova-Oliver M (2005) STAMP: A model for generating adaptable multimedia presentations. In: Multimedia Tools and Applications, vol. 25, no. 3. Springer Netherlands, pp. 361-375

[4] Boll S, Klas W (2001) ZYX-A multimedia document model for reuse and adaptation of multimedia content. In: IEEE Trans. On Knowledge and Data Engineering, vol. 13, no. 3, pp. 361-382

[5] Borning A (2000) Constraint-based document layout for the Web. In: Multimedia Systems special issue on Multimedia authoring and presentation techniques, vol. 8 , no. 3, pp. 177-189

[6] Bulterman DCA, Grassel G, Jansen J et al (2005), Synchronized Multimedia Integration Language (SMIL 2.1). W3C, http://www.w3.org/TR/SMIL

[7] Bulterman DCA, Hardman L (2005) Structured Multimedia Authoring, ACM Trans. on multimedia computing, communications, and applications, vol. 1, no. 1, pp. 89-109

[8] Decneut S, Hendrickx F, Nachtergaele, Van Assche S (2004) Targeting heterogeneous multimedia environments with web services. In: Proc. of IEEE International conference on Web services (ICWS), San Diego, California, pp. 682-689

[9] Deltour R, Guerraz A, Roisin C (2006) Multimedia Authoring for CoPs. In: Proc. of the 1st international workshop on building technology enhanced learning solutions for communities of practice, Crete, Greece

[10] Dragicevic P, Chatty S, Thevenin D, Vinot JL (2005) Artistic Resizing: A Technique for Rich Scale-Sensitive Vector Graphics. In: Proc. of the 18th annual ACM Symposium on User Interface Software and Technology, Seattle, USA, pp. 201-210

[11] ISO/IEC 14496-11:2005, Information technology - Coding of audio-visual objects - Part 11: Scene description and application engine

[12] Jourdan M, Layaïda, Roisin C, Sabry-Ismaïl L, Tardif L (1998) Madeus, an authoring environment for interactive multimedia documents. In: Proc. of the 6th ACM international conference on Multimedia, Bristol, United Kingdom, pp. 267-272

[13] Kimiaei-Asadi M, Dufourd JC (2005) Context-Aware Semantic Adaptation of Multimedia Presentations. In: Proc. of the IEEE International Conference on Multimedia and Expo (ICME), Amsterdam, Holland, pp. 362-365

[14] Laborie S, Euzenat J (2008) An incremental framework for adapting the Hypermedia structure of multimedia. In: Advances in Semantic Media Adaptation and Personalization, Studies in Computational Intelligence, vol. 93. Springer, Berlin Heidelberg, pp. 157-176

[15] Marriott K, Meyer B, Stuckey PJ (2005) Towards Flexible Graphical Communication Using Adaptive Diagrams. In: Advances in Computer Science - ASIAN 2004, vol. 2220 of LNCS. Springer, Berlin Heidelberg, pp. 380-394

[16] McCormack C, Marriott K, Meyer B (2008) Authoring Adaptive Diagrams. In: Proc. of the ACM symposium on Document engineering (DocEng), Sao Paulo, Brazil, pp. 154-163

[17] Meijering E (2002) A chronology of interpolation: From ancient astronomy to modern signal and image processing. In: Proc. of IEEE, vol. 90, no. 3, pp. 319-342

[18] Metso M, Koivisto A, Sauvola J (2001) A Content Model for the Mobile Adaptation of Multimedia Information. In: Journal of VLSI Signal Processing Systems, vol. 29, no. 1-2. Kluwer Academic Publishers, pp. 115 - 128

[19] Nanard M, Nanard J, King PR, Gaillard L (2007) Genre Driven Multimedia Document Production by Means of Incremental Transformation, In: Proc. of ACM Symposium on Document Engineering (DocEng), Winnipeg, Canada, pp. 111-120

[20] Pellan B and Concolato C (2008) Adaptation of Scalable Multimedia Documents. In: Proc. of the ACM symposium on Document engineering (DocEng), Sao Paulo, Brazil, pp. 32-41

[21] Pellan B and Concolato C (2008) Spatial Scene Adaptation in Broadcast Environment, In: Proc. of the International Conference on Multimedia & Expo (ICME), Hannover, Germany, pp. 389-392

[22] Ransburg M, Timmerer C, Hellwagner H (2008) Dynamic and Distributed Multimedia Content Adaptation based on the MPEG-21 Multimedia Framework. In: Multimedia Semantics - The Role of Metadata, vol.101 of LNCS. Springer, Berlin Heidelberg, pp. 3-23

[23] Scherp A, Boll S (2005) MM4U: A Framework for Creating Personalized Multimedia Content. In: Managing Multimedia Semantics. IRM Press, pp. 246-287

[24] Schwarz H, Marpe D, Wiegand T (2007) Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. In: IEEE Trans. on Circuit Systems for Video Technology, vol. 17, no. 9, pp. 1103-1120

[25] Sidibé M, Koumaras H, Kofler I Mehaoua A, Kourtis A, Timmerer C (2008) A novel monitoring architecture for media services adaptation based on network QoS to perceived QoS mapping. In: Journal on Signal, Image and Video Processing (SIVP), vol. 2, no. 4. Springer London, pp. 307-320

[26] Sullivan T, Matson R (2000) Barriers to use: usability and content accessibility on the Web's most popular sites. In: Proc. of the 2000 ACM conference on Universal Usability, Arlington, Virginia, pp. 139-144

[27] Sutcliffe A (2001) Heuristic evaluation of website attractiveness and usability. In: Interactive systems: design, specification, and verification, vol. 2220 of LNCS. Springer, Berlin Heidelberg, pp. 183-198

[28] Vetro A, Christopoulos C, Ebrahimi T (2003), eds., Special Issue on Universal Multimedia Access, IEEE Signal Processing Magazine, vol. 20, no. 2, March 2003

[29] Villard L, Roisin C, Layaïda N (2000) A XML-based multimedia document processing model for content adaptation. In: Digital Documents and Electronic Publishing, vol. 2023 of LNCS. Springer, Berlin Heidelberg, pp.104-119

[30] Villard L, Layaïda L (2002) An incremental XSLT transformation processor for XML document manipulation. In: Proc.of the 11th international conference on World Wide Web, Honolulu, Hawaii, pp. 474 - 485

[31] Wirag S (1997) Modeling of adaptable multimedia documents. In: Interactive Distributed Multimedia Systems and Telecommunication Services, vol. 1309. Springer, Berlin Heidelberg, pp. 420-429