# SMIL Content Adaptation for Embedded Devices

Tayeb Lemlouma
OPERA Project
Zirst 655 Avenue de l'Europe – 38330
Montbonnot, Saint Martin, France
Tel: +33 4 7661 5281
Tayeb.Lemlouma@inrialpes.fr

Nabil Layaïda
OPERA Project
Zirst 655 Avenue de l'Europe – 38330
Montbonnot, Saint Martin, France
Tel: +33 4 7661 5384
Nabil.Layaida@inrialpes.fr

## Abstract

In this paper, we propose a SMIL content adaptation framework for embedded devices such as PDA, cellular phones, etc. This framework is based on a three-pier organization of the SMIL content access which allows decoupling two levels of performance: a wired network with abundant resources such as bandwidth and processing power and, a wireless network with variable bandwidth and limited resources. Clients are given access to content from a variety of terminals with different capabilities. In order to adapt the SMIL content to these devices, the content access traverses an intermediate entity called adaptation proxy where the adaptation occurs. An experimental implementation of this framework shows how to achieve SMIL content adaptation using client profiles descriptions, negotiation strategies, simple SMIL adaptation and media transcoding.

## 1. Introduction

One of the most active fields these days is ubiquitous and embedded computing. The proliferation of a variety of devices such as PDAs and cellular phones together with wireless technology such bluetooth, 802.11, GPRS and the upcoming UMTS at the enabling factors of this research. These various technologies has created a need to provide traditionally desktop accessed content on these devices. First applications ported to the embedded world were e-mail, agenda synchronization and more recently the attention is increasingly given to a more complex environment such as the Web. In particular, as these devices are becoming more multimedia capable, one of the interesting challenges is the multimedia content delivery on these embedded devices.

On the other hand, SMIL 2.0 has become a World Wide Web Consortium recommendation in 2001. It is the dominant representation in Web technology for describing timing and synchronization of multimedia presentations. A careful attention has been given, in the design of SMIL, to modularity and extensibility of the recommendation and three language profiles have been proposed [3]. Most notably, SMIL Basic profile is a collection of modules together with a scalable framework, which allows customizing the document profile to the capabilities of the device. Providing an adaptive content is still under investigation as some general mechanisms such as content negotiation, universal profile (document, user, network, terminal) descriptions and processing are not yet well established.

Several simple content selection mechanisms have been introduced in SMIL to provide a better flexibility. But in most cases, SMIL adaptation is achieved at the client side. This supposes that the client is adaptation capable and that the profiles and the client capabilities are set somehow. In addition, adaptations do not necessarily belong to the same layer of a document presentation. One can start by designing a device independent document layer and generate, once the profiles are identified, the SMIL content representation. It is also possible to operate adaptation within a SMIL document instance beyond what is provided by the format. It is also possible to modify the content itself to fit bandwidth and display limitations. This is particularly useful when the content is legacy and authored without the content selectors provided by the language. The global movement today in content adaptation is to make explicit declaratively the capabilities in terms of user, document, client and server profiles. Another visible movement is the separation of the different aspects of the document and their modularization. For an efficient content access from various devices, there is a need to reconcile the different profiles by providing negotiation and adaptation methods at every layer of this information system.

In this paper, we present a framework capable of such adaptations at the format and media content level. Adaptation is handled at a proxy where profiles are processed and adaptations achieved. The application developed according to this framework includes an embedded SMIL Basic player, an adaptation proxy together with content adaptation methods.

## 2. The Framework Overview

The application architecture that we consider in our framework is composed by content servers, embedded clients given access to the content through a wireless network and intermediate entities that perform particular tasks. A content server is the entity where documents and resources are stored. Multimedia documents existing in the server side, may contain many media types: text, pictures, video, audio, animation, scripts, etc. The stored content can be client specific. It can also be defined in a more abstract as semantic relationships between resources instantiated during content negotiation and delivery.

In the considered environment, clients can be very different regarding their hardware and software capabilities and preferences. As the content that we consider follows the SMIL model, the client used in our architecture is an embedded player called PocketSMIL. PocketSMIL is a SMIL 2.0 player which has been developed within the Opera project. It is a conforming SMIL 2.0 Basic player developed for experimental purposes to meet resource constrained clients such as personal device assistants, pagers, mobile cellular phones, etc.

In a browsing session, clients send their requests to the content server via an intermediate entity called a proxy. In wireless environment, little assumptions are made on the connection speed, reliability and bandwidth. This means that clients and servers may interact in a low bandwidth and unreliable environment.

A third entity called proxy exists between clients and servers (Figure 1). From the communication point of view, a proxy plays the role of a client in a proxy-server session, and the role of a server in a proxy-client configuration.

The proxy (Figure 1) is a good place to deal with the heterogeneity of clients and servers. Another advantage is that clients and servers need not to be modified and all the environment characteristics that already exist still the same. The presence of proxies adds new interesting functionalities and facilities to make the client/server interaction more efficient. In the context of SMIL content adaptation the proxy-based architecture represents a very suitable solution. The proxy is the entity responsible of retrieving client requests and contexts and performing possible adaptation on the content received from the server. The adapted content is then sent to the client with respect to its characteristics. The proxy can transform on the fly (with live streams) existing multimedia content and thus existing content does not have to be re-created for each context. All the proxy tasks are generally transparent for clients and content servers.

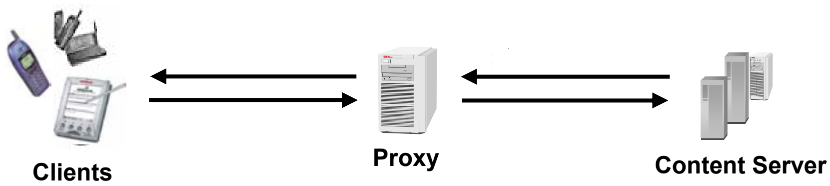



**Figure 1. Overall architecture**

In our approach, the target context constraints such as client capabilities and preferences are described using the universal profiling schema UPS [6]. UPS is based on RDF [9] and CC/PP [8][4] and allows also to describe other elements involved in the adaptation such as the document profile, the network profile, etc. [6]. During the content negotiation and profiles processing applied by the proxy, the client profile and the profile of the requested content are parsed and the set of included constraints are resolved.

The proxy checks then if the resource (media or document) is compatible with the client capabilities or not. In the positive case the resource is sent directly to the client without any modifications. In the negative case the proxy checks, using UPS structure [5] and SMIL facilities, if it is possible to have a version of the resource that can meet the client requirements. If the proxy succeeds to find a variant that fits the client requirements, the original resource is substituted by the variant. Otherwise the proxy tries to adapt the original resource using external adaptation methods. As we will see in the rest of this paper, SMIL flexibility [7] and UPS offer a fairly efficient approach for the delivery of adapted content for constrained environments and embedded devices.

## 3. Application Architecture

As we have seen in the previous section, our architecture is build on top of three main elements: the client, the content server and the intermediary proxy. The proxy entity includes an adaptation and negotiation module called ANM. This later can cooperate with an optional module, called UCM (User Context Module), which can reside on the client. It allows the retrieval of the client profile (the client capabilities and preferences). The client identification and thus profiles determination can be done merely by ANM using the network information (such as the device IP address, the session identifier, the device location, etc.) which can be obtained from the client session. Profiles can be stored everywhere in the mobile environment: in the server side, in a common profile repository (Figure 2), etc. In our architecture the proxy retrieves the profiles thanks to the device IP address which identifies a unique profile. For efficiency reasons, we apply also a simple caching strategy to avoid network overhead due to frequent profiles transfer.

The communication proxy ensures the interaction task with clients and original servers. The proxy receives client requests and adapts the SMIL content of the server using ANM. To perform the adaptation task properly, ANM needs to process the information related to the characteristics of the user but also the capabilities and the functionalities of the content (the document profile). In our approach, we propose to add semantic metadata related to the content that exists at the server side (Figure 2). This additional information is directly related to the adaptation mechanisms and includes: the SMIL internal relationships defined between different objects of the content and additional metadata that can relate SMIL objects with other external objects available at the server side.

The approach that we propose in this paper transfers the SMIL content adaptation to the intermediate proxy. The proxy receives the user request sent by the PocketSMIL client and the corresponding content from the original server. The received content can not be sent directly to the client since it belongs to the SMIL 2.0 language profile. Consequently, the proxy must perform necessary adaptation techniques to make the content understandable by the target device, SMIL Basic in this case.

The approach aims to make the proxy responsible of the whole content adaptation task even those to be achieved by the client. For instance, even the Basic Content Control module [1] which is a part of the SMIL 2.0 Basic profile [10] should be processed by the proxy. This processing takes into consideration the limitation of embedded devices that can not achieve advanced adaptation mechanisms such as video manipulations and transcoding [2]. Furthermore, the exchange of adaptation oriented messages (content profile transfer, device capabilities declaration, etc.) between servers and client should be minimized since wireless networks suffers from many resource limitations. Therefore, it is preferable to restrict this type of interactions between the proxy and the server located in the wired part of the network.

In another side, the adaptations using proxies is more efficient since it considers better the characteristics of the environment. The proxy can have a global picture of the environment resources: the network latency and bandwidth, the content size to be transported before the final delivery, the user preferences and user changes that may occur, etc. The use of the proxy ensures also the profiles processing and content caching.

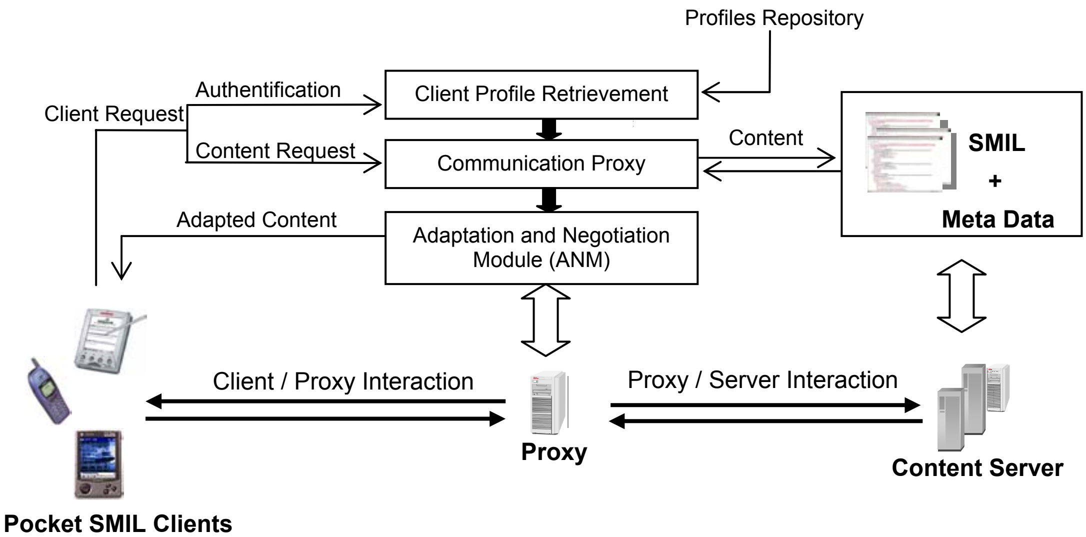
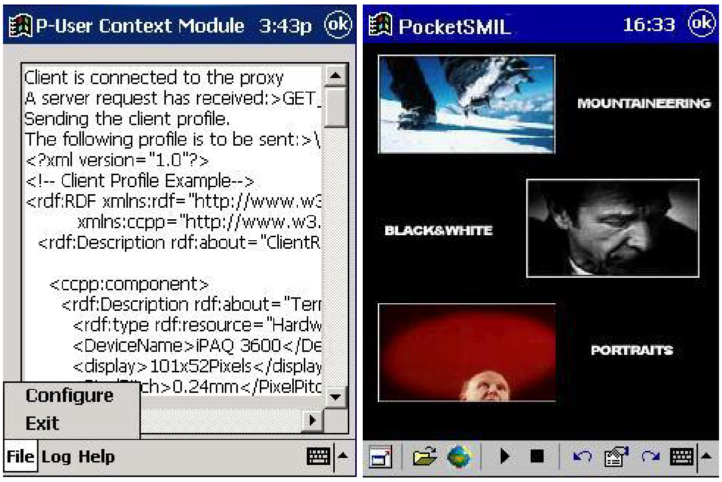



**Figure 2. Proxy based adaptation process**

In our architecture, ANM runs permanently on an intermediate proxy and allows content delivery after receiving client requests. ANM processes the client request and retrieves the requested SMIL content from the original server. Using a predefined content negotiation and adaptation strategy, the original content is matched to the client preferences and capabilities and eventually to other parameters extracted from UPS profiles [5]. The adapted content is finally delivered to the client. In the case where the client supports additional modules and in order to achieve advanced and optimized negotiation exchange between the client and the proxy, another module (the user context module: UCM) is used (Figure 3.a). In this situation and before that the client player, the PocketSMIL in our case (Figure 3.b), starts to request SMIL documents; UCM should run using the ANM host address and the context listening port of the ANM. As we said earlier, UCM is optional. It was developed for embedded devices to facilitate content negotiation such as supporting the user context change and the user profile manipulations (Figure 3.a).




**(a) User Context Module: UCM** **(b) PocketSMIL player**

**Figure 3. The embedded client**

Once the last configuration settings are achieved, the ANM adaptation process can adapt dynamically SMIL content received from the server and stored temporarily in the ANM host using and/or combining appropriate transformation methods: style sheets or programs. In order to facilitate the UPS profiles processing, the utile information (i.e. different constraints) are extracted and stored in memory. The ANM adaptation strategy uses directly the memory form of UPS profiles.

## 4 Proxy Adaptations

After collecting the different profile and the original SMIL content, the proxy applies its adaptation process and matching strategy. The adaptation process includes: version selections, document transformation and media adaptation.

### 4.1 Versions selection

This approach consists to choose the best variant of the multimedia content or object on behalf of the user agent. Generally the selection is applied on the available variant set and based on variants description and the user requirements. Selection criteria may include the language, the media type, the char-set, etc. Unlike other models like HTML, SMIL offers many facilities to achieve this task. SMIL 2.0 Content Control Module [1] allows using the *switch* element to specify inside the document a collection of alternative elements. As we delegate the adaptation task to the proxy, ANM processes the switch element and chooses, when possible, the first acceptable element. An element (a timing structure or media object) with no test attribute is always acceptable. The content selection can be expressed using the SMIL system test attributes (the seven attributes of SMIL 1.0 and the new five attributes added by SMIL 2.0). ANM evaluates the test using the information extracted from the different UPS profiles. For example, if we take the following SMIL document:

```
. . .
<par>
        <audio src="welcome_to_inria.wav" . . . />
        <switch>
                <img src="inria_1024_1280.gif" systemScreenSize="1024X1280"/>
                <img src="inria_480_640.gif" systemScreenSize="480X640"/>
                <img src="inria_240_320.gif" systemScreenSize="240X320"/>
                <img src="inria_default.gif"/>
        </switch>
</par>
. . .
```

If the extracted screen size from the UPS client profile equals to 240X320, the ANM proxy will result in the delivery of the following SMIL document:

```
. . .
<par>
        <audio src="welcome_to_inria.wav" . . . />
        <img src="inria_240_320.gif" />
</par>
. . .
```

Note, the variants selection is applied and the SMIL document is restructured by eliminating the switch elements and the non selected alternatives. Another way to apply elements selection is the use of SMIL in-line test attributes. In this situation, ANM acts in the same manner as with the switch element. However, the processing here is easier because the element to be selected exists independently from the switch. Consequently, the element is either left or removed according to the test evaluation. If the evaluation returns true, the element is left and the test attributes are removed from the original document. Example:

```
. . .
<par>
        <textstream src="presentation_speech_translation.rt" systemLanguage="fr" />
        <audio src="presentation_speech.mp3" . . . />
        <video src="presentation.mpg" . . . />
</par>
. . .
```

Here, the proxy removes the textstream element if according to the UPS client profile, the language of the user is different than French. In the opposite case, the proxy delivers the following document:

```
. . .
<par>
        <textstream src="presentation_speech_translation.rt" />
        <audio src="presentation_speech.mp3" . . . />
        <video src="presentation.mpg" . . . />
</par>
. . .
```

As we can note, performing alternative selections by the proxy rather than by the target client has many benefits such as the efficient evaluation of the test attributes, like the systemBitrate attribute, and thus allows the optimization of the resources consumption of the environment. Unfortunately, the versioning approach puts burden on content authors. They are required to produce multimedia objects separately for each individual device and context.

## 4.2 Documents transformation

This approach concerns transformations that are applied on the global document organization or tree and can modify its structure. The SMIL structural transformation applied by the proxy can either keep the same media resource used by the original SMIL document, filter it or require an external media resource transformation to adapt the media for the end user context. Document manipulations are generally applied using appropriate languages designed for structure and tree modification such as XSLT.

To show how can the proxy behave, we give here a simple example of a client that uses the PocketSMIL player through a PDA device. The client retrieves the content of the server in a wireless network. After the necessary configuration settings (pointing the connection to ANM, etc.) the client requests a content in the form of a SMIL document located at the address: http://opera.inrialpes.fr/Smil/T/t1.smi. The document is presented as follows:

```
. . .
<body>
<seq>
    <par>
        <audio id="audio" src="tayeb.mp3" begin="0s" end="12s"/>
        <img id="img1" region="r1" src="image01.jpg" begin="0s" dur="12s"/>
        <img id="img2" region="r2" src="image02.jpg" begin="00:00:6" dur="6s"/>
    </par>
    <video region="r4" src="iceage.mpeg" begin="3s" end="30s"/>
    <par>
        <img id="img3" region="r3" src="image03.jpg" begin="0s" dur="6s"/>
        <img id="img4" region="r5" src="image04.gif" begin="3s" dur="3s"/>
    </par>
</seq>
</body>
. . .
```

As we can note, the above document includes different types of media: image, audio and video presented in different ways. The document was tested first, with a UPS client profile which does not contain any constraint concerning SMIL documents. The multimedia presentation was played correctly with respect to the source document semantic. The client profile (represented by the *ClientProfile.xml* file) was modified to declare that video and audio media resources are not supported anymore. As it is shown in the following:

```
. . .
<ccpp:component>
  <rdf:Description rdf:about="NonSupportedResources">
  <rdf:type rdf:resource="Resources" />
  <neg:NonSupportedResources>
       <rdf:Bag>
          <rdf:li rdf:parseType="Resource">
          <neg:type>video</neg:type>
          <neg:format>mpeg</neg:format>
          <neg:profile>device-profiles/mpeg-profile.xml</neg:profile>
          </rdf:li>
          <rdf:li rdf:parseType="Resource">
          <neg:type>audio</neg:type>
          <neg:format>wav</neg:format>
          <neg:profile>device-profiles/wav-profile.xml</neg:profile>
          </rdf:li>
          . . .
       </rdf:Bag>
  </neg:NonSupportedResources>
  </rdf:Description>
</ccpp:component>
. . .
```

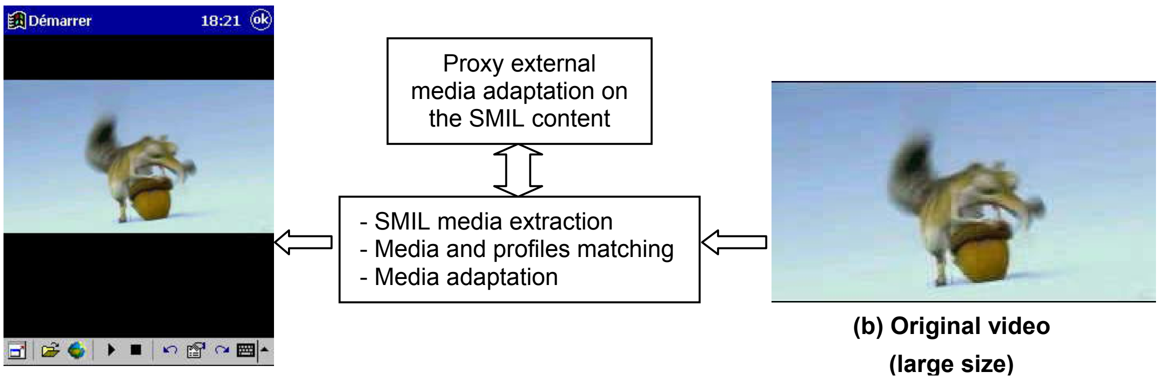
The PDA requests then the same SMIL document with the new profile. The multimedia presentation received from ANM is played without neither audio nor video with the same temporal semantic of the original presentation: for instance the semantic of the parallel and the sequential media were not violated. Another document transformation which has been tested also, is to transcode SMIL2.0 documents to SMIL 2.0 Basic by removing some non supported elements. For example, the *t1.smi* document uses the regPoint element added to the SMIL Basic profile by the HierarchicallLayout module [11]. The client profile was modified to indicate that the client does not support that element. The adapted document received from ANM proxy is a SMIL Basic document.

## 4.3 Media Resources Adaptation

Usually SMIL multimedia presentations include media objects (images, audio files, video streams, etc.) by reference. Media resources of different SMIL documents should not be sent directly to the end users if they do not meet their needs and do not satisfy the target context constraints. In order to make these objects available to the client and adapted to the environment characteristics such as the mobility [2] and low available bandwidth, media resources can be substituted, removed or transformed to an acceptable format using available adaptation techniques and methods. Semantic relations can also be used in order to complete the role of direct adaptation programs. The minimal semantic vocabulary that expresses these media relationships includes the "equivalent" and "adapted to" relations. An "equivalent" relation means simply that the two related objects have the same role in the multimedia presentation. An "adapted to" relation means that the target object represents a constrained version of the content. A semantic parameter indicating the level of the adaptation can be associated with the "adapted to" relation. This can be exploited in the content negotiation strategy if a media resource can be adapted to more than one object.

Using the proposed proxy based architecture, the adaptation of media objects referenced by the SMIL content includes video and image adaptation. Video adaptation considers the characteristics of the wireless network [2] where the bandwidth must be optimized unlike fast wired networks. Furthermore, adapting the video used by SMIL documents considers the client capabilities in terms of screen dimensions and color resolution.

The following figure shows how the media adaptation can serve to deliver adapted SMIL content that imports video objects. Here, the original content is represented by a simple SMIL document that includes a full resolution video as it shown in figure 4.b. In this case and in order to fit client requirements, the proxy resizes the video first and sends the output to the client as it shown in figure 4.a. One of the important advantages of this process is that the adaptation and the content delivery can be achieved on the fly using real time streaming protocols such as RTP and RTSP which takes into account the current characteristics of the network. Using such mechanisms allows to enable small devices to access some desktop SMIL presentations.




**(a) Received content by PocketSMIL**

**(b) Original video (large size)**

**Figure 4. Video adaptation inside SMIL content**

In addition to the video adaptation, the proxy allows also images transformation including additional image compression and resizing. The following SMIL presentation, located in http://opera.inrialpes.fr/Smil/T/SMILDocSample.smi, uses a large image with high resolution:

```
. . .
<body>
        <par dur ="120s">
                <audio src="Frozen.mp3"/>
                <img src="Ray_of_Light_CD.jpg"/>
        </par>
</body>
. . .
```

As we can see below (figure 5.b), the original SMIL document is not adapted to the client characteristics (the screen dimensions here) as it is on a desktop (figure 5.a). Indeed, using the PocketSMIL on a personal device assistant having 240x320 as screen dimensions, the image object used by the received SMIL document, takes a large area and can not be displayed entirely. Furthermore, the transport of the original media object consumes a lot of resources of the wireless network: bandwidth and transport time.


**(a) Original SMIL document on a PC** **(b) Original SMIL document on a PDA** **(c) Adapted SMIL document on a PDA**

**Figure 5. Image adaptation inside SMIL content**

The proxy adaptation of the SMIL document gives as a result the presentation shown in the figure 5.c. The adaptation is performed using a resizing and a JPEG compression of the image used in the original document. Thought that the size of the original image equals to 39284 bytes, the image resizing and compression take an average time of 321 milliseconds and 140 milliseconds respectively which gives a size gain of 33272 bytes.

As we have seen, applying media resources adaptation in SMIL has many benefits that help considering client capabilities. However, the application of these methods requires additional delays that are related to the processing of the original media and the output of the adapted one.

In order to evaluate these adaptations, we measured their cost. Several adaptations were experimented on different image resources. Each content adaptation method is applied -on the fly- at the proxy side and results in a new adapted media delivered to the client. The experimented adaptation is represented by a JPEG compression of the images content.

The following table (Figure 6) gives the measurements concerning the adaptation and the delivery times of four images. Image *a* represents the original media referenced by the requested SMIL document. The other images are generated by the proxy using different level of compressions applied to *a*. The given values of the delivery time do not include the time taken for the transformation.

| Media Resource | Size (bytes) | Applied Compression (%) | Transformation Time (millisecond) | Delivery time (millisecond) |
|---|---|---|---|---|
| Image a | 24635 | 0 | 0 | 376 |
| Image b | 23366 | 50 | 362,20 | 356 |
| Image c | 13998 | 80 | 360,20 | 214 |
| Image d | 9776 | 90 | 297,16 | 149 |

**Figure 6. Media adaptation and delivery time**

The delivery time gain obtained by delivering an adapted media rather than delivering the original one depends to three parameters: the adaptation time of the original media resource, the current network bandwidth and the difference of size between the original and the adapted media. If $a$ is the value of the size difference and $b$ is the media adaptation time, the delivery time gain will be positive while the network bandwidth value is less than a/b. Otherwise the delivery of the adapted resources used inside the document will take more time than the delivery of the original content. Consequently, when the bandwidth becomes modest, it is efficient to apply adaptation methods in order to deal with the network situation. In future work, this formula will be used to determine the level of the adaptation to apply in the situation where the user agent requires a maximal delay time regarding the reception of the content.

## 5. Conclusions

In this paper we presented a SMIL document adaptation framework for embedded devices. This architecture is based on an intermediate entity, the proxy, which intercepts and adapts content transfers between clients and servers.

As we have seen, SMIL has several advantages. It allows adaptability and provides flexibility thanks to modularization of the language profiles. In our approach, we have used these facilities with other mechanisms, such as the capabilities and preferences description using UPS, to deliver adapted content for embedded devices. Delivered content is in the form of SMIL Basic vocabulary which is ideal to this kind of limited devices.

Our proxy adaptation of the SMIL content has considered two levels of the multimedia content: the structure of SMIL documents and the media objects used inside the document. This approach has permitted to obtain fairly good results given the client and network limitations. The adaptations applied to the media objects offer new interesting applications for embedded devices. For example, live video broadcasts can be aggregated and adapted on the fly and according to the available network resources. As a future work in adaptation, we plan the design and use of a complete vocabulary that includes the necessary set of semantic metadata.

## References


[1] Dick B. and Jeffrey A. The SMIL 2.0 Content Control Modules. http://www.w3.org/TR/smil20/smil-content.html

[2] Hagimont D. and Layaïda N. Adaptation de serveur multimédia par un code mobile. *Technique et Science Informatiques (TSI)*, numéro spécial Agents et code mobile, vol. 21, num. 6, 2002.

[3] Layaïda N. and Van Ossenbruggen J. SMIL 2.0 Language Profile. http://www.w3.org/TR/smil20/smil20-profile.html.

[4] Lemlouma T. and Layaïda N. Basic Ideas for User Constraints Specification in CC/PP. Research report, June 2002. http://opera.inrialpes.fr/people/Tayeb.Lemlouma/Papers/Constraints%20Expressing.pdf

[5] Lemlouma T. and Layaïda N. Universal Profiling for Content Negotiation and Adaptation in Heterogeneous Environments. W3C Workshop on Delivery Context, W3C/INRIA Sophia-Antipolis, France, 4-5 March 2002.

[6] Lemlouma T. and Layaïda N. Universal Profiling Schema. http://opera.inrialpes.fr/people/Tayeb.Lemlouma/NegotiationSchema/index.htm, January 2002.

[7] Rutledge L., Hardman L. and Ossenbruggen J. V. The use of SMIL: Multimedia Research Currently Applied on a Global Scale.CWI (Centrum voor Wiskunde en Iformatica), NL-1090 GB Amesterdam, The Netherlands.

[8] W3C. Composite Capability/Preference Profile. http://www.w3.org/TR/CCPP-struct-vocab/, W3C Working Draft 15 March 2001.

[9] W3C. Resource Description Framework (RDF) Model and Syntax Specification. http://www.w3.org/TR/1999/REC-rdf-syntax, W3C Recommendation 22 February 1999.

[10] W3C. SMIL 2.0 Basic Profile and Scalability Framework. http://www.w3.org/TR/smil20/smil-basic.html.

[11] W3C. Synchronized Multimedia Integration Language (SMIL 2.0), W3C Recommendation 07 August 2001.