Automatic Multimedia Creation Enriched with Dynamic Conceptual Data

Angel Martín, Haritz Iribas, Ion Alberdi and Naiara Aginako

Digital Tv & Multimedia Services, Vicomtech-IK4 GraphicsMedia.net, Donostia-San Sebastián, Spain

Abstract — **There is a growing gap between the multimedia production and the context centric multimedia services. The main problem is the under-exploitation of the content creation design. The idea is to support dynamic content generation adapted to the user or display profile. Our work is an implementation of a web platform for automatic generation of multimedia presentations based on SMIL (Synchronized Multimedia Integration Language) standard. The system is able to produce rich media with dynamic multimedia content retrieved automatically from different content databases matching the semantic context. For this purpose, we extend the standard interpretation of SMIL tags in order to accomplish a semantic translation of multimedia objects in database queries. This permits services to take benefit of production process to create customized content enhanced with real time information fed from databases. The described system has been successfully deployed to create advanced context centric weather forecasts.**

Keywords — **Context-aware Multimedia, Dynamic Multimedia Production, Content Driven Database, SMIL**

I. INTRODUCTION

In the recent years, motivated by the rapid rise of the Internet, new challenges have arisen from the increasing amount of audiovisual data that are becoming available. Furthermore, the traditional multimedia production has experienced a deep change due to the arrival of new communication paradigms and the evolution and convergence of new technologies. In this context, distributed multimedia presentations are getting more and more popular for web users. SMIL (Synchronized Multimedia Integration Language) [1] is a W3C standard, designed for describing multimedia presentations which combine audio, video, images, text or any other media.

When the creation is concerned, in order to reduce time and cost of multimedia production, both the possibility of working on them collaboratively and reusing existing content are mandatory. For reutilization purpose the use of meta-information has become a necessity [2]. The increase of multimedia content, and the growing difficulty to search, filter and manage such content, require an effective and efficient multimedia search and retrieval system [3]. In addition, the Web [4] may also be employed as a medium to connect a group of authors, which may be geographically distributed, empowering users to be more creative and fostering social interactions enabling cooperative production [5]. One of the main advantages of team working is to gain efficiency in time to solve complex problems, in order to reach a better final result. Currently, the majority of authoring tools goes online enabling real time creative collaboration processes through cooperative frameworks.

Audiovisual producers gather different multimedia contents and graphic objects composing more attractive and enriched contents in order to engage the audience while consumers are willing to embrace new technologies inherited from other application areas to get richer experiences [6]. However, content-aware research is usually placed at the end of the production process, instead of integrating the solution within the production chain where the different multimedia objects involved are still isolated. This fact facilitates multimedia analysis workout through exploiting content knowledge inside the description of the creation process, in terms of timeline and spatial schedule of multiple objects in the scene [7]. A drawback related to the majority of solutions is the interoperability, they operate over private formats that hide the creation process and makes not feasible to exploit this information for other purposes. Another major trend of context-aware multimedia research that could take advantage of multimedia creation descriptors focus on dynamic generation of multimedia presentations.

This paper proposes an automatic dynamic content generator platform based on SMIL standard. Regarding automatic generation of video content in order to address the gap between visualization layer and dynamic data retrieval for databases, we have designed standard-based definitions. The main benefit is the optimization of the costs of the producing process, as well as saving time. In our context, the creation of multimedia presentations, the simplest way for that is the reusability of previously created multimedia videos and templates into new ones. Besides, in this

automatic generation of content, we are able to insert new 2D-3D objects as the spatial relation among different objects of the scene is defined in the production process. Our work has been tested in a weather forecast platform.

The rest of the paper is organized as follows. In Section II, we introduce the related work involved in the development, while in Section III we explain the implementation which includes architecture and the main features of our platform. In section IV we state the validation of the platform in the meteorological domain. Finally, we give some conclusion remarks in Section V.

II. RELATED WORK

We can classify the related work section into several broad categories: collaborative SMIL authoring tools; SMIL players; and the metadata exploitation for search and retrieval and Content Management System (CMS).

There are several advantages of using the web standard SMIL instead of tailoring a custom XML. First, the standard compliance eases the interoperability with other solutions or platforms and future developments. Furthermore, it provides access to a huge audience through web browsers. There are several authoring tools for developing multimedia presentations in SMIL. Some of the research solutions like H-SMIL-Net [8], LimSee2 [9] and SMILAuthor2 [10] provide SMIL authoring interface for multimedia objects. They are designed to interpret the temporal behavior of a multimedia presentation, associate hyperlinks with media objects, enable interactivity, and describe the layout of the presentation on a screen. Other authoring tools [11], [12], add to these works the opportunity to work collaboratively to make the SMIL file enriching the final multimedia presentation with different points of view, while reducing the cost and production time. Our approach lives on the previously works aggregating support to additional features described below.

The players can be divided into two categories. On the one hand, the specific SMIL players like GRiNS [13], Ambulant [14] and SmilingWeb [15]. SmilingWeb is cross-platform player for SMIL 3.0 presentations contained in web pages. On the other hand, it should be mentioned that some of the most important media players support SMIL, such as Real Player, Totem, QuickTime and Windows Media Player. However, none of them supports 3D animations and virtual environments rendering that enhance the experience generated by our platform.

The amount of multimedia content that has to be managed has already become unaffordable without a good CMS. This requires an extensive use of multimedia metadata [16]. The SMIL standard provides extra information of the different multimedia objects involved in a content using metadata [17]. Some studies have analyzed how SMIL metadata can apply to the indexing and abstracting of multimedia documentation [18]. Besides, it can be seen how SMIL metadata is used to store information of different templates in databases in some papers [19]. It facilitates the search and retrieval of these templates, making the system more efficient to reuse content. Here, the proposed platform gathers the human knowledge on semantic tags through the multimedia objects contextualization performed in the authoring tool enabling the connection with databases that collect real time information.

The platform described in this paper integrates a SMIL collaborative authoring tool, a multimedia presentations player and a metadata tag aggregator to facilitate the reuse of contents, saving the information of templates in a database.

In addition, we have created a system that using ontologies allows retrieving data from different databases dynamically, so multimedia presentations can be automatically generated using the generated SMIL templates. It should be mentioned that the platform produces videos from these multimedia presentations. Other features provided by our approach include the possibility to add an avatar or virtual character, that makes the created videos more friendly, but also an integrated text to speech converter.

III. IMPLEMENTATION

The architecture of our platform is based on SMIL standard and Gstreamer [20] open source multimedia processing framework. Our approach has extended the semantic meaning of some tags of SMIL in order to perform dynamic content creation while keeping standard compliance. It is achieved by bridging semantic context connections of the different contents involved, such as multimedia elements defined in the outline and scheduled in the layout, with data stored in a database. Once the required data are retrieved from the database according to semantic matching algorithms, Gstreamer is responsible for the creation of the video designed in the SMIL template.

Meanwhile, as we described in the previous section, the SMIL document itself helps the platform to facilitate the work among different people. SMIL describes not only the play out outline but also the production process so it is perfect to trace the activity of each team member in the cooperative creation of multimedia contents. Regarding concurrency and consistency, the SMIL structure centered on parallel or sequence relations between the different elements involved eases the required control. This means, each member can place a multimedia content inside a sequence without timing constraints or in parallel with others establishing just the z-order, from background to foreground, and the position in the layout. The changes are

updated in the timeline and the different lines where concurrent contents are represented in order to provide the team current situation awareness. The authoring process driven by a visual representation of the timeline and the different lines available for each content helps to build a consistent SMIL document that depicts the visual designed outline and layout.

Taking into account that the present research in collaborative environments focuses on authoring tools and the required communication and concurrency control on top of XML modules based on Web Services (WS), we have developed a Service Oriented Architecture (SOA) which integrates WS communication and control solutions. However, we will focus on the real novelty around dynamic content generation with data retrieved from databases.

A. Extending SMIL specification

For our approach, SMIL document capabilities potential in terms of content management is twofold depending on the metadata granularity. On the one hand, SMIL provides suitable features in terms of CMS where the metadata, that describe the document content, enable indexing of a large library of generated documents while metadata also provide recommendation solution basis in a similar way that MPEG-7 [2] does but easing web publishing and processing. On the other hand, SMIL document templates also support creation of dynamic multimedia reports and contents through the semantic definition of the different multimedia contents involved. This means, aggregated metadata can be exploited in order to achieve dynamic generation of multimedia contents enabling temporal and spatial visual templates that can be fed with real time retrieved data from databases in order to perform updated and customized videos according to user preferences [21]. This way, our developed platform bridges the gap between data visualization and data storage leveraging data available from databases.

However, no tags for this kind of circumstances are conceived at all in SMIL. So, we use existing tags within a new semantic context in order to keep backwards compatibility with current players. These tags would ease the mapping translation from concepts within the SMIL document, defined like a template, to database (DB) queries.

We use the standard <meta> and <metadata> tags as they are suitable from the point of view of aggregating semantic meaning for the different visual objects scheduled in the document while keep fully standard compliance.

Three different semantic layers have been defined:

1) *First layer:* Metadata related to the general document template suitable for ownership and semantic keyword indexing based CMS.

2) *Second layer:* Semantic keywords defining Primary Key (PK) field name and related fields from DB tables.

3) *Third layer:* Multimedia objects and DB fields mapping.

The different roles of the added attributes are scheduled in the Fig. 1.

The meta-information included in the first layer is stored in a database for more exhaustive searches, and to facilitate reuse of content. In the depicted example the SMIL document represents a "weather report" template to show the "weather figures forecast".

Not just semantic keywords are defined for the general SMIL document, but also for different multimedia objects involved in the second layer. This way advanced CMS indexing is supported easing advanced search of objects inside produced contents by the authoring tool.

The subtemplate structures, defined in the second layer, provide a gathered semantic scheme of required fields from the DB ("attribute") around a PK ("subtemplate_key") where the keywords ("content") provide semantic concepts related to DB table fields. In the example depicted the concept employed as PK is "city" and the required fields to fulfill the necessary data are "city name", "forecast icon" and "wind speed". So, these definitions establish the items in order to infer the SQL database queries.
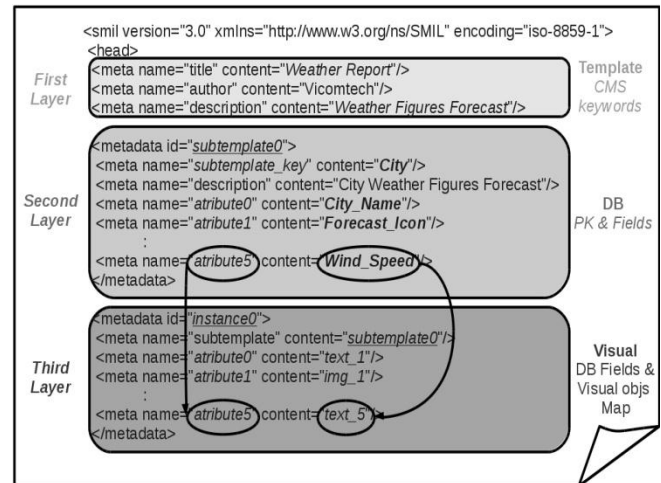

Fig. 1.   Semantic data for dynamic performance.

Finally, in the third layer, multimedia object IDs involved in the SMIL document ("content") are mapped to the DB fields ("attribute"). So contents from layer 2 and 3 are linked through the attributes correspondence. Again, according to the Fig. 1, the values of the layer 2 attributes like "wind speed" retrieved from the fields of different DB tables for a specific PK, in this case a "city", would be linked to a layer 3 text ("text_5") respectively. This way the aggregated metadata enable the DB middleware to translate layer 2 structure to DB queries such as SQL

$$select\ f(X)\ from\ T(Y)\ where\ f(Y) = target\_P$$

where:
- f(X) is the result of semantic matching of keyword X with field names from the DB.
- T(Y) is the result of semantic matching of keyword Y with table names from the DB.
- target_PK is the target primary key.

In our example X would be "wind speed", Y "city" and target_PK the city-ID whose forecast parameters will be introduced in the resulting multimedia forecast report.

The definition of subtemplates in the second layer permits the reutilization of subtemplates when creating new multimedia contents in order to reuse a previous design within the same template document. The subtemplate structure has been conceived like a "Class" of object oriented programming languages while instance structure has been designed as an instance of object oriented programming languages. Therefore, the same dataset definition could have different visual instances within the same template depending on the PK value provided. For example, according to the elements defined in the SMIL document depicted in the Fig. 1 if the designer is interested in create a multimedia forecast for several cities, he should just copy the existing instance and add as many as he wants changing the "instance id" and keeping the "subtemplate id" reference.

B. Dynamic multimedia platform architecture

Once we have defined the extension to SMIL standard in order to both CMS features to reuse the content and make a platform that can create videos with dynamic data automatically; in this section, we give an overview of the developed platform based on Gstreamer to render the result.

This platform provides a solution for:
- User edition through the authoring tool to create both multimedia templates and contents.
- Applications and services that request the automatic generation of a certain content according to the context of the user or periodic time scheduled events.

Fig. 2 shows the general architecture of the approach where individual building blocks are described below:

1) *Authoring tool:* Here, users can work collaboratively in the implementation of the multimedia content. Moreover, to facilitate user work, and remove learning curve of the SMIL language, we have developed a multimedia edition tool based on SMIL standard that supports aggregation of metadata tags according to an ontology in order to accomplish DB incrustation.

2) *Context Scheduler:* Third party application or user context aware service that wants to provide user-centric multimedia content.

3) *DB Translator:* Semantic translator from layer 2 metadata and specified PK values, from the authoring tool or a context scheduler, to DB queries. The translator engine is deployed on top of OWL-API and Pellet solutions [23]. The response is the result of perform the query.

4) *Template DB*: SMIL document templates in a database organized around layer 1 metadata.

5) Template Manager: CMS that manages the library of SMIL document templates.

6) *Content DB:* Multimedia DB with contents for specific applications such as weather forecast reports.

7) *Gstreamer Render:* Multimedia processing platform able to mix and orchestrate contents both online and offline contents (Content DB).

8) *SMIL Parser:* Web requests and responses processor to create a content in SMIL format.

9) *Server:* Web service manages two kind of requests:

a) *Available multimedia templates:* Request includes target topics. So, level 1 keyword based multimedia template search is performed by the Template Manager. The response includes all the recommended templates and their layer 2 and 3 data to be translated to DB queries.

The designed workflow starts when the Authoring Tool for human active creation or Context Scheduler for automatic multimedia generation, requests through a WS some available multimedia templates to provide a multimedia content to a specific user. Here context is performed by means of some semantic keywords that define the multimedia target type.

The Template Manager retrieves SMIL documents exploiting information from layer 1 matching with the request keywords. The response includes exclusively the layer 2 and 3 data of the filtered SMIL templates.

b) *Multimedia content creation:* Request includes level 3 with real DB values for each multimedia item. The generation of the video according to the involved SMIL templates and the data provided is performed. The response is the URL with the generated video.

In case there is a request for multimedia content creation from dynamic data, DB Translator takes charge of processing layer 2 data to transform them into DB queries. Afterwards layer 3 elements are filled by Context Scheduler or Authoring Tool with the figures and values resulting from the DB queries. Here it is important to highlight that values from layer 3 linked to a text object will produce text rendering. While if it is related to image or video objects it refers an URL to an existing content. Last but not least, a text value referred to an audio will unleash a Text To Speech (TTS) audio generation.
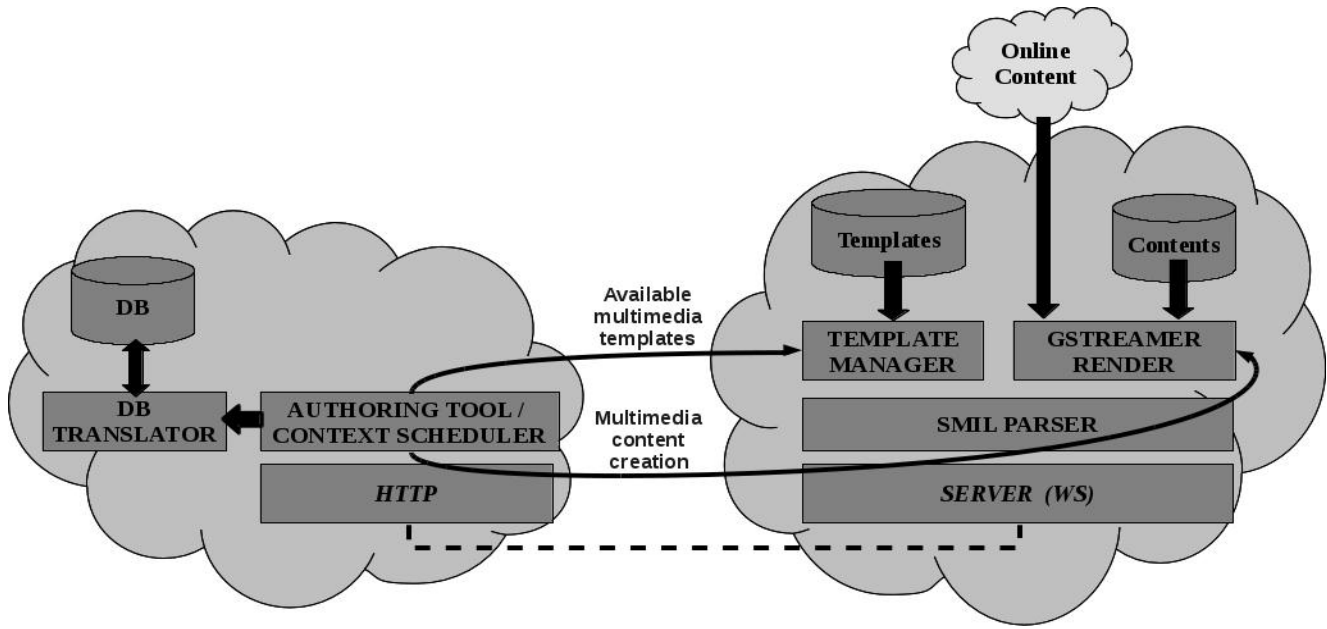
Fig. 2. Architecture for dynamic multimedia creation.

c) Design of ontology: We have designed an ontology in order to define the concept relations between the dynamic content saved in the databases. Therefore, it is possible to manage the information of different databases in order to extract more exact information. The template editor has a limited access to the definition of DB queries so is necessary to implement anarchitecture in order to develop a more powerful tool.

In the meteorological context, we have defined an ontology drawing all the relations between the parameters of weather stations. Consider a real weather station, it contains several meteorological or environmental sensors that measure different forecast parameters. But only measures are relevant, we have to consider more information as sensor model, sensor picture, geolocalization... The definition of the attributes of the concepts is highly relevant.

Ontology permits inserting PK in a total transparent manner from the user point of view. When creating a new template, user inserts a meteorological station ID and the system, using the defined ontology and its relations, shows the possible parameter definitions.

Attributes give the capacity of adequate the presentation of the parameters to the user. Therefore, depending on the nature of each concept, different rules are applied to new objects.

Once all the required data of the desired templates form the final fulfilled SMIL document, it is employed to request the multimedia SMIL driven creation through the WS.

For the process of creating the video, we have developed a SMIL Parser that reads the SMIL document and creates a DOM [22] representation. It extracts the data into a format that we can easily use to create a Gstreamer pipeline. After the platform generates the pipeline with the marks of time and space indicators, GStreamer module executes the pipeline. The GStreamer Render generates the video results and responds

with the URL of the video. The Gstreamer core module, gnonlin is in charge of managing the temporal and spatial positions of the different multimedia elements that make up the output video.

To sum up, we have designed a context scheduler that according to time, user preferences and context defines the required PKs in the automatic generation case; or the user himself, in case of on demand authoring, to later perform the renderization request by inserting the retrieved data from the DB inside the corresponding instance of the SMIL document. In order to accomplish our objectives, we have developed all the modules defined in our platform which permits the total control of the automatic content creation.

C. Content-aware production inline

Multimedia applications' next move focuses on 3D-2D fusion and user centric and context-aware that needs of advanced multimedia analysis algorithms to achieve realistic experience and automatic creation of dynamic contents customized to the user context and preferences.

The availability of isolated visual objects within the production chain eases the introduction of 2D-3D objects through the <z-index> SMIL attribute associated to each object. We have exploited the intrinsic information from the production process to remove issues associated to research solutions usually placed at the end of the production chain, such as segmentation, clustering, depth map and more in particular for 3D-2D fusion T-unions, optical flow, partial occlusions, etc. As we keep information of the production process in the SMIL document, we manage to make the later fusion of 3D-2D objects and effects realistically.

During the design process of our content generator platform, we have defined a module to control the z-index of the objects completing a scenario. Therefore, during the creation process of new dynamic content we are able to insert new 2D-3D data

to enrich the already defined scenario making the content production affordable.

D. Authoring Tool

The authoring tool to create new objects has been designed as a video editor adapted to user needs. This video editor is divided in two parts. On the one hand, we have the usual editor that permits the mix of audios, videos, texts, images and avatars in the easiest way possible for the multimedia items creators. On the other hand, the collaborative editor permits the reuse of old video templates and videos templates that other collaborators made, and the blend of them for creation of new material.

The whole editor has three main areas structure. On the right there is a preview and options area, on the left there is an elements list and bellow those areas, there is an edition area with the buttons that are necessary for it.

- Elements List

On the top-left side of the interface there is a list of elements with the content library that will be used for the creation of the templates. With the below buttons we can add and remove elements, with the buttons plus and minus respectively, to the list so that the user could use them in the edition.

- Edition Lists

In the video editor there are three elements lists, in the first list are audio elements, in the second list are background elements and the speeches of avatars and in the third list lessons main elements.

The user can manage intuitively all the elements; add, erase and change elements position and duration with the buttons placed at the bottom of the interface. It is so easy to use that e.g. if the user wants to add an element to any list, the system identifies the type of the element and adds to the corresponding list, if it is an audio element it adds in the first list while if it is a video or an image the user just has to define if he wants to add it as a background element or as a main element of the learning object.

- Edition in Collaborative Web Editor

SMIL has different tags for the definition of timing process and the most important for our collaborative system is <par> and <sec> tags.

The elements inside the tag <par> are reproduced all in parallel being grouped to support complementary pedagogical information. The tag <sec> permits the reproduction of elements in sequence enabling sequential outline construction. It is really important because permits to reproduce an element after other without an explicit definition of the start time. The system knows that as soon as one element ends next one must start, maintaining the established order. The user just has to define the order of the elements and when saves it, the system takes the information of the different files and saves in a resultant file.

To sum up, according to the strategy of keeping the developed platform simple, experts can create new objects through defining new multimedia templates or modifying existing ones through a desktop editor application described before. Afterwards the created objects are pushed automatically to the server.

IV. VALIDATION

To test the approach tailored for weather forecast production, validation work has been mainly driven by semantic translation of multimedia objects in database queries by checking correct connection of meteorological concepts, defined by metadata, to real time acquisition databases. SMIL rendering compliance was also performed, where we have created different templates with all kinds of items, including pictures, videos, texts and audios.

Using the authoring tool, we have designed a wide range of meteorological report templates in which different weather maps have been generated as depicted in the Fig. 3. On the one hand, we have produced weather maps where the multimedia elements are not interrelated with semantic information layer. On the other, we defined maps with the items related to weather stations from which data are obtained. In this case tests performed achieve multimedia results automatically adapted to geographical user context and preferences. We also proved the correct implementation of the defined ontologies concepts by retrieving the requested parameters using semantic high-level concepts as weather stations.

As mentioned above, templates are very useful because allow us to create new presentations each day with the same structure, but with new data. These data are acquired from the database, using the weather stations as a PK identifier. In other words, when creating a new video, system translates the metadata attached to a multimedia template item, for example that visualizes the atmospheric pressure, and it takes the value automatically from the database of a specific station according to the user context system.

After testing, we have seen that when introducing elements, the definition of <z-index> in SMIL document has allowed us to use 3D elements, such as avatars, in a direct manner. We resolve the problem of occlusions as the depth of the objects in the scene is already defined.

About the metadata aggregation, the platform makes even easier the use of templates and databases. Access to a database via the web platform makes it more intuitive for inexperienced users. User can provide human understandable meteorological concepts that would be translated to real time acquired atmospheric measures. This way the user just has to provide a geographic context to establish the database PK. After selecting this information, the system automatically creates the queries for the database so that the user does not have to worry about anything. In addition to manually creation of videos, it has proven the periodic creation of videos in an automatic manner.

Fig. 3.   Platform produced weather report.

This testbed based in the weather information platform does not have real time performance restrictions or high scalability requirements, but it faces dynamic contexts describing a solution extensible from time and localization features to context user-centered multimedia applications. The automation of this process has enormously reduced the cost of producing new contents for weather forecasts and the time needed for the creation of new videos which implies that new content can be updated with higher frequency.

V. CONCLUSION

Creating new multimedia contents and keeping them as a template in order to ease reutilization, decreases dramatically the amount of digital content replicated in edition platforms. Moreover, the capability of creating this content collaboratively permits new envisaged creation possibilities and sharing of content. The SMIL documents considered as templates also ease the reutilization of production work according to reuse and recycle philosophy of green computing minimizing the set of needed resources.

The redefinition of the metadata tag of SMIL standard permits the declaration of objects in SMIL presentations. Giving a semantic meaning to these metadata, reveals the possibility of creating new updated multimedia content, including dynamic data, that automatically minimizes the production costs of lots of contents. This property becomes crucial in repetitive processes which must be updated with new data, for example in weather reports.

The proposed format to describe the content outline makes feasible to exploit this information for other purposes related to content-aware research, such as auto-summaries, 2D-3D fusion or visual near-duplicated research. This one may be relevant for rights holders to provide control of their content enabling copyright policies and content management through automatic processes to identify, claim, and apply policies to partial or entire content.

Future work will include more complex relations in terms of database structures that will also need of enhanced semantic translators from context based keywords to existing databases.

We also plan to take benefit of the multimedia object segmentation provided by the SMIL templates in order to boost visual analytics applications improving the search algorithms.

REFERENCES

[1]   W3C, SMIL Specification, http://w3.org/AudioVideo
[2]   Cornejo, C. M., Ruiz-Rube, I., Dodero, J. M. "Semantic Management of Digital Contents for the Cultural Domain", Recent Trends in Information Reuse and Integration, Springer, pp 211-226, 2012
[3]   HongJiang Z., "Multimedia search: Past and current approaches", IEEE 10th Workshop on Multimedia Signal Processing, 2008
[4]   Olson, G.M, Zimmerman, A., and Bos, N., Eds., "Scientific Collaboration on the Internet" MIT press, Cambridge, MA, 2008
[5]   Sung ,M.Y., Lee D.H., "A Collaborative Multimedia Authoring System", Lecture Notes in Computer Scien-ce, Volume 3033/2004, pp.311-318, 2004
[6]   Jansen, J., Cesar, P., Bulterman, D., Stevens, T., Kegel, I., and Issing, J., "Enabling Composition-Based Video-Conferenfing for the Home", IEEE transactions on Multimedia, IEEE, Volume 13, pp. 869-881, 2011
[7]   R. Kernchen, S. Meissner, K. Moessner, P. Cesar, I. Vaishnavi, M.u Boussard and C. Hesselman, "Intelli-gent Multimedia Presentation in Ubiquitous Multide-vice", IEEE Computer Society, pp. 52-62, April-June 2010
[8]   S. Bouyakoub , A. Belkhir, "H-SMIL-Net: A Hierarchical Petri Net Model for SMIL Documents", Proceedings of the Tenth International Conference on Computer Modeling and Simulation, p.106-111, 2008
[9]   Deltour, R., Layaida, N., Weck, D., "LimSee2: A Cross-Platform SMIL2.0 Authoring Tool", The European Research Consortium for Informatics and Mathematics – ERCIM News. n. 62 , 2005
[10] Chun-Chuan, Y., Chen-Kuei, C, Yung-Chi, W., "Extension of Timeline-based Editing for Non-deterministic Temporal Behavior in SMIL2.0 Authoring," Journal of Information Science and Engineering, vol. 24, no. 5, pp. 1377-1395, 2008
[11] Sung, M. Y., Lee, D. H. "A collaborative authoring system for multimedia presentation", Proceedings of the IEEE International Conference on Communications, 1396-1400, 2004
[12] El Saddik, A., Rahman, Md.A., Hossain, M.A., "Authoring Multimedia Objects in Collaborative Ambient Intelligent Virtual Environment", IEEE International Workshop on Haptic Audio Visual Environments and their Applications, 2005
[13] Bulterman, D.C.A., Hardman, L., Jansen, J., Mullender, K.S., Rutledge, L., "GRiNS: A GRaphical INterface for creating and playing SMIL documents" WWW7 Conference, Computer Networks and ISDN Systems, vol. 30, pp. 519–529, 1998
[14] Bulterman, D.C.A., Jansen, J., Kleanthous, K., Blom, K., Benden, D., "Ambulant: a fast, multi-platform open source SMIL player", Proceedings of the 12th annual ACM international conference on Multimedia, 2004
[15] O. Gaggi, L. Danese. "A SMIL player for any web browser", Proceedings of International Conference on Distributed Multimedia Systems, p 114-119, 2011
[16] Garcia, R., Celma, O. "Semantic Integration and Retrieval of Multimedia Metadata", 5th Int.Workshop on Knowledge Markup and Semantic Annotation, 2005
[17] Bulterman, D., Rutledge,L., SMIL 3.0 Flexible Multi-media for Web, Mobile Devices and Daisy Talking Books
[18] Gil, B., Pérez, J. R., Sánchez, R., "Gestión de documentación multimedia con Synchronized Multimedia Integration Language (SMIL)." // Scire: representación y organización del conocimiento. 10:2, p 87-97, 2004
[19] Jiang, D., Liu, Z., Zhong, L., "A smil-based multimedia system for mobile education", 2nd IEEE International Conference on Broadband Network & Multimedia Technology, IC-BNMT '09, p 724-728, 2009
[20] Gstreamer web page, http://gstreamer.freedesktop.org/
[21] Steele, R., Lubonski, M., Ventsov, Y., and Lawrence, E., "Accessing SMIL-based Dynamically Adaptable Multimedia Presentations from Mobile Devices", Proc. Of Int. Conf. on Inf. Technology: Coding and Compu-ting, ITCC04, 2004
[22] W3C, DOM Specification, http://www.w3.org/DOM/

[23] OWL api engine website, http://owlapi.sourceforge.net/

A. Martin. Angel Martin is with the Department of Digital Tv & Multimedia Services, Vicomtech-IK4 GraphicsMedia.net, Spain. He received his engineering degree in 2003 from University Carlos III of Madrid, Spain. He has published many papers on multimedia technology in many conferences. His research interests are in the areas of multimedia services, interactive multimedia and rich media contents creation and transmission.

In 2003 he starts his Phd in the Communications and Signal Processing Department of the University Carlos III of Madrid in the video coding research area. At the same time he collaborates with the company Prodys developing a standard MPEG-4 AVC/H.264 codec for DSP architectures, and web applications for coding management of IP surveillance cameras. In 2005 he starts to work on Telefonica I+D in projects related to multimodal interactive service projects for Home Network environments and web services. Later, also in Telefonica I+D, goes deeper into image processing area in terms of 3D video and multiview coding. From 2008 to 2010 in Innovalia Association as R&D Project consultant developing projects and applications for smart environments, ubiquitous and pervasive computing in the artificial vision area. Currently he is on Vicomtech managing and developing R&D projects around multimedia content services.

I. Alberdi Ion Alberdi is a research assistant at Vicomtech-Ik4 in the area of Digital TV & Multimedia Services. He graduated in Computer Engineering in 2011 from UPV-EHU where he undertook his undergraduate dissertation based on analysis of cartographic imagery at the Italian Università degli Studi di Padova. His research interests are in the areas of multimedia services and computer vision.

H. Iribas. Haritz Iribas is is with the Department of Digital Tv & Multimedia Services, Vicomtech-IK4 GraphicsMedia.net. He graduated in Technical Telecommunications Engineering in 2008 at the Public University of Navarre (UPNA-NUP). Between the 2006 and 2010 he teached sciences at the private schools "Laguntza" and "International college". Since July of 2010 he has been working in Vicomtech-IK4. His research interests are in the areas of multimedia services, interactive multimedia and computer vision.

N. Aginako. Naiara Aginako is with the Department of Digital Tv & Multimedia Services, Vicomtech-IK4 GraphicsMedia.net, Spain. Between 2000 and 2005 she studied Ingenieria Superior de Telecomunicación (Computer Engineering applied to Telecommunication) at the University of the Basque Country (UPV-EHU). She was cooperating with the Electronics and Communications Department of the above mentioned Faculty between 2003 and 2005 where she undertook her undergraduate dissertation. Since October of 2005 she has been developing and leading R&D projects in Vicomtech-IK4. She undertook her postgraduate course in Informatics Engineering at the University of the Basque Country (UPV-EHU) during 2008-2010. Since 2006, she also teaches junior level courses at the University of the Basque Country at the Electronics and Telecommunications Department.