A NEW STEP TOWARDS MULTIMEDIA DOCUMENTS GENERATION

Muriel Jourdan and Frederic Bes

Unité de Recherche INRIA Rhône-Alpes
ZIRST, 655 avenue de l'Europe, Montbonnot, 38334 St Ismier CEDEX, France

Tel: 33 476 61 53 58; Fax: 33 476 61 52 07
E-mail : {Muriel.Jourdan,Frederic.Bes}@inrialpes.fr

ABSTRACT

We are interested in applications that automatically generate a dynamic multimedia presentation adapted to the needs of the user from a database of XML fragments of basic elements (video, images, paragraphs, ...). We present in this paper the design of an architecture which goal is to support developers of such applications. This architecture takes benefits from both the use of style sheets, to associate different renderings that depend on discrete parameters, and the use of a constraint solver, to handle continuous parameters and to optimize the content selection taking into account global criteria.

Keywords: Presentation Generation, Optimization, Multimedia Document.

INTRODUCTION

TV-like multimedia presentations will be more and more used to communicate information : training courses, medical reports, news, on-line shopping catalogues, ... However, designing high-quality multimedia presentations is known to be a complex, time-consuming and error prone task [3][5][8] whatever the used authoring tool is. Moreover, the dynamic nature of some information (they are regularly updated) like news, medical information, ... and the information overload problem associated with the Internet, raise a crucial need of information services that are customized for the needs of individual users [4][12]. Typical examples of such information services are : educational courses adapted to the student's level, automatic design of medical reports, personalized TV-news, ...

Whereas a lot of works have been done in those areas on the run-time adaptation of the navigational structure of a hypermedia document [7], few works focus on the automatic generation of dynamic (i.e. temporal) multimedia presentations. Our aim in this paper is to propose an architecture, for helping the development of personalized information services. This proposition is based on a detailed analysis of works already done in this area. It will give the opportunity to the developers of such kind of applications to find a real software environment that ease their work.

The first section presents the main components of an application that automatically generates multimedia presentation from a set of basic objects. The second one is devoted to a structured presentation of related works. The third section presents the architecture we proposed.

1 A FUNCTIONAL DESCRIPTION OF A PERSONALIZED MULTIMEDIA PRESENTATION SYSTEM

We consider in this paper that a personalized multimedia presentation system is an application that takes some user parameters (content preferences, capabilities, ...) and technical parameters (screen size, bandwidth ...) and generates a dynamic multimedia presentation from a set of XML fragments. Dynamic multimedia presentation means that objects are organized both in time, space and hyperspace (hyperlinks structure).

The process for automatic multimedia presentation generation can be divided into 4 main functions :

content selection : selects which objects must participate to the final multimedia presentation;

temporal ordering : defines a temporal order between the selected objects;

layout definition : defines the spatial organization of the selected objects;

hyperlink definition : defines the navigational structure between the information fragments.

One difficulty encountered in such applications is to know in which order content selection and content organization (through the three dimensions) must be done. For instance, deciding if an object with a duration d has to be present in a final presentation which duration equals to D (this is a user's request) strongly depends on the presentation temporal organization. On the other hand the temporal order between two objects may depend whether another object is present or not in the final presentation.

The two following examples will help the reader to understand the kind of applications we are interested in :

- *personalized sport report* : deliver a TV-like report after a championship day to a user who can choose the total report duration, the teams for which she/he wants to have detailed information (her/his favorite teams), the kind of the report : an exhaustive one (at least some minimal information about each match is given) or not (the maximum information about each favorite team is given).

- *personalized medical survey* : deliver a multimedia presentation that describes the main patient surgeries by using some synchronizations between a graphical representation of a patient body, some speech and textual information [5].

2 RELATED WORKS

2.1 Programming based approach

It is obvious that classical programming languages like JAVA can be used to develop a system that generates personalized multimedia presentation. On the Web, this approach is mainly used by using CGI scripts. Well known disadvantages of this approach are : their lack of independence between the programming code and the piece of information; the difficulty to reuse some parts of an application for another one; …

2.2 Using templates and selection instructions

Thanks to the use of more descriptive languages like XSLT [19], Northolk [15], … or more imperative ones like PHP [14], JSP [9], ASP [11] it is possible to write virtual documents (they are dynamically generated on demand) that merge static data, i.e. information that does not change, with dynamic ones, i.e. information dynamically extracted from external data sources. This is done by integrating in some presentation language (HTML for instance or SMIL for multimedia presentation) a way to make queries on some data sources and to filter the results of such queries. This approach can be thought as the design of a template in a presentation language that contains the static parts of the document and some selection instructions that express how to collect the dynamic data. This approach suits well to application in which content selection can be split into several local database requests. But it is not easy to handle global content selection criteria like: choose the set of videos that maximize the user interest and such that the sum of video duration is lower than D. Moreover even if one implements it through successive requests, the time performances of the application may become very slow.

2.3 Using a set of style sheets

Style sheets are a well known way to associate different presentation rendering (different layout for instance) with the same set of objects. These presentation rendering are usually deduced from a logical structure (for instance a DTD) associated with the objects. Personalization can be achieved by associating one style sheet with each parameter value that controls the application. An example of such kind of architecture can be found in [15] in which the user can choose between three different styles to present the same set of objects. The obvious limitation is that it is only possible to handle discrete parameter. In [20] more flexible style sheets are presented. They are based on the use of constraints. For instance the layout of the document can be express in the style sheet by using high-level constraints (centering, alignment, …) that will be solved by using a constraint solver. Doing so, the application has some possibilities to adapt the document to the size of the user's screen. However some drawback remains, since it does not provide the designer with a powerful way to select information, only filtering on some local criteria can be easily expressed.

2.4 Optimization algorithms

A lot of works in the area of automatic generation application put the emphasis on its content selection step and formalize it as a optimization process [6][10][12] For example, in [10] the system tries to create a TV news program that fits with the user interests (a lot of international political news followed with few sport news for instance) and a maximal duration constraint. The database contains a set of individual news classified by categories. Moreover a numerical value is associated with each news to express its degree of importance. The content selection problem is then equivalent to find a news subset that maximize the interest of the user while respecting the duration constraint. Then the authors propose an heuristic that makes the resolution of this NP problem more efficient. Once the content selection has been done, the TV news program simply consists in concatenating the selected elements (i.e. sequential presentation). A similar approach is used in [12] to generate music program (a concatenation of music …) adapted to the user preferences (at least 30% female-type voice, at most 20% "Jazz style", …).

These works seem to be very efficient to select content on global criteria but they mainly address some specific applications (the proposed heuristic strongly depends on the addressed problems) and do not propose some generic architecture to ease the design of various kind of automatic generation applications. Moreover, they only consider a simple and unique way to temporally organize the selected content (spatial and hyperlink organization are not at all addressed). Only [6] takes the temporal organization issue into account but in such a way that the content selection is done without any knowledge on the future organization of the object (content selection is performed before temporal organization decision). As a consequence the selected content is often not adapted to the user request. For instance, the user wants a multimedia document of 30 min but get something significantly shorter than the wished 30 min.

2.5 Knowledge based approach

People from knowledge based area makes some propositions to use their experiment for solving "automatic generation like" problem. [1][4][5][16]. The idea is to use different knowledge bases (on the content properties, on some principles for temporal organization, on graphical rules, …) in order to be able to answer to a large variety of requests (even those not planned when the system was designed) by decomposing the initial request in some subgoals to reach. Thus, the content selection decisions are distributed through the application and we do not see how easily and efficiently handling global criteria. Moreover this approach is certainly efficient to deal with semantic criteria but it is not easy to see how dealing with other kind of criteria.

2.6 Synthesis

Style sheets based approaches like templates based ones have the same kind of advantages and limitations : they suit well to associate different final rendering (depending on discrete parameters) to a set of (may-be locally filtered) existing objects. However, they fail to easily express global criteria. At the opposite, formalizing the automatic generation application as an optimization problem provides the designer with a powerful way to base content selection upon global criteria. However, only specific applications have been yet handled by such an approach and always with very simple and unique temporal organization of selected content.

In this paper, we propose an architecture that take benefits from both the use on style sheets (to associate different final rendering to the same set of objects) and the expressive power of optimization tools.

3 A GENERIC ARCHITECTURE FOR AUTOMATED CONSTRUCTION OF MULTIMEDIA PRESENTATION

The main principle of our architecture is that both a transformation sheet (a more general kind of style sheets) and a constraints set controls the content selection step and the organization decisions through the three dimensions. The transformation sheet aims to fix the temporal, spatial and hyperlink organizations of a set of objects. The constraints express a set of rules that has to be as most as possible satisfied while selecting the basic objects on which the transformation sheet will be applied (duration of the presentation must be less than 10 min, no selection of objects with sound properties). In our architecture the transformation sheet is selected before the content selection step in order to be able to select objects that fits with the final presentation.

Parameters	Constraints	Transformation sheets
No sound	Dur(Sound Object) = 0	-
Presentation duration = D	Dur(Doc) = D	If D < Dmin then TS1 or TS3 else TS2 or TS4
Team exhaustiveness	∀t∈ Team, Dur(t) > 0	-
Screen Size = S	Size(Doc) = S	If S < Smin then TS3 or TS4 else TS1 or TS2

Figure 1 : Example of parameters association

The designer has to define :

- The input parameters that control the application;
- A set of different transformation sheets that are related to some values of input parameters (TS1, TS2, … in Figure 1 and Figure 2);
- Several set of constraints that depend on input parameters values (CS1, CS2 and CS3 in Figure 2);
- A set of constraints that does not depend on the input parameters value (General Constraints in Figure 2);
- The XML structure of the basic objects;

Figure 1 illustrates some parts of this work on the personalized sport report presented in the introduction. We suppose that the designer has defined 4 transformation sheets that depends on whether the requested duration is short or long and the size screen is large or not. When the user choose a presentation duration, the following constraint (`Dur(Doc) = D`) is added to the set of general constraints that is used as input by the content selection component. Moreover, some transformation sheets are selected and globally analyzed by the negotiation component to select one transformation sheet at the end.

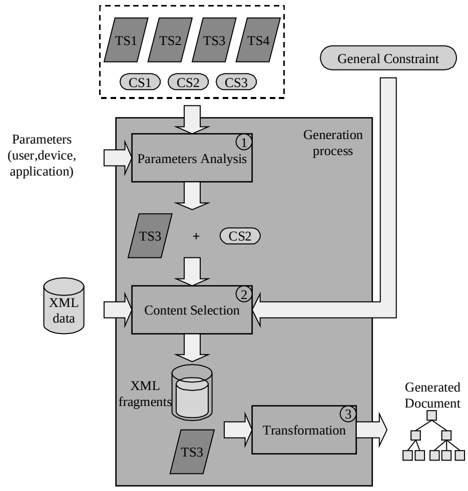

Figure 2: Architecture

The proposed architecture (Figure 2) consist of three main components.

The *Parameters analysis* component analyses the input parameters in order to select one transformation sheet and to build a constraints set that merged general constraints with constraints that are associated with the given parameters value;

The *Content selection* component aims to select a subset of XML fragments from the general XML structure and to compute some attributes value associated with this selected XML data (objects duration, objects size, …). It is decomposed in three layers:

- The translation layer which goal is to translate all the high level constraints into constraints manageable by the optimization layer. This translation strongly depends on the temporal and spatial organization defined by the selected transformation sheet. For instance, the constraint Dur(doc) = D must be translated into the sum of objects duration if they are in sequence in the transformation sheet.
- Optimization layer which goal is to solve the constraints issued from the translation layer in an optimal way. We currently experiment the use of

Cassowary [2] a constraint solver that is able to efficiently solve linear constraints.

- Filter layer which goal is to use the result of the optimization layer to extract the XML fragments from the XML structure. Moreover it gives a value to each attribute that corresponds to a variable of the constraints system solve by the optimization layer.

The *Transformation* component corresponds to the transformation processor. The document in the target format (SMIL [17] for instance) is build from the transformation sheet chosen by the Analysis parameters component and the selected data given by the *Content Selection* component by using tools like Xalan XSLT processor [21].

The main feature of this architecture is the use of a constraint solver to globally optimize the content selection and also to perform some parts of the formatting phase in order to handle continuous parameter like global durations, screen size, … Moreover we choose to select the transformation sheet before selecting the content in order to constrain this selection with some information deduced from the transformation sheet.

4 CONCLUSION AND FUTURE WORKS

After a detailed analysis of existing works on automatic generation of multimedia documents we have proposed a generic architecture to ease the design of such kind of applications. This architecture take benefits from both the use of style sheets and the expressive power of constraint solvers to easily and efficiently handle global criteria of optimization. Thanks to this architecture, the designer is able to generate dynamic multimedia document which content and organization (even the temporal one) suits well to the user's requests. The next short-term step of this work is to evaluate its performance on large scale applications. Another direction of our future works will be to consider fragments of transformation sheets in order to reduce the designer's tasks and to get a final document still more adapted to the user's request. Our final goal is to provide the designer with a more complete support from easy authoring of transformation sheets [19] to run-time adaptation of multimedia documents.

REFERENCES

[1] E. André and T. Rist, "Generating Coherent Presentations Employing Textual and Visual Material", AI Review 9, 1995, pp. 147-165.

[2] G. J. Badros and A. Borning, "The Cassowary Linear Arithmetic Constraint Solving Algorithm: Interface and Implementation", Technical Report UW-CSE-98-06-04, June 1998.

[3] S. Boll and W. Klas, "ZYX - A Multimedia Document Model for Reuse and Adaptation", In Transactions on Knowledge and Data Engineering, DS-8 Special Issue, IEEE, 2000, to appear.

[4] M. Bordegoni et al., "A Standard Reference Model for Intelligent Multimedia Presentation System", Computer Standards and Interfaces, Special Issue on Intelligent Multimedia Presentation Systems, 18(6-7), December 1997, pp. 477-496.

[5] M. Dalal et al., "Negotiation for Automated Generation of Temporal Multimedia Presentations. ", Proceedings of ACM Multimedia'96, ACM Press, 1996, pp. 55-64.

[6] V. Hakkoymaz, J. Kraft and G. Ozsoyoglu, "Constraint-Based Automation of Multimedia Presentation Assembly", Proceedings of ACM Multimedia'99, ACM Press, Orlando, USA, November 1999.

[7] N. Henze and W. Nejdl, "Adaptivity in the KBS Hyperbook System", 2nd Workshop on Adaptive Systems and User Modeling on the WWW, Toronto, May 1999.

[8] M. Jourdan et al., "Authoring SMIL documents by direct manipulations during presentation", World Wide Web, Balzer Science Publishers, 2 (4), December 1999.

[9] JSP, http://java.sun.com/products/jsp, 2001.

[10] B. Merialdo et al., "Automatic Construction of Personalized TV News Programs", Proceedings of ACM Multimedia'99, ACM Press, Orlando, USA, November 1999.

[11] Microsoft ASP, http://www.microsoft.com, 2001.

[12] F. Pachet, P. Roy and D. Cazaly, "A Combinatorial approach to content-based music selection", IEEE Multimedia, March 2000.

[13] F. Paradis and A.-M. Vercoustre, "A Language for Publishing Virtual Documents on the Web", International Workshop on the Web and Databases, Valencia, Spain, March 1998.

[14] PHP, http://www.php.net, 2001.

[15] L. Rutledge et al., "Adaptable Hypermedia with Web Standards and Tools", The Active Web -- A British HCI Group Day Conference, London, January 1999.

[16] L. Rutledge et al., "Generating Presentation Constraints from Rhetorical Structure", Proceedings of the 11th ACM conference on Hypertext and Hypermedia, San Antonio, USA, June 2000, pp. 19-28.

[17] SMIL, http://www.w3.org/TR/smil20/, 2001.

[18] A.-M. Vercoustre and F. Paradis, "Metadata for Photographs: from Digital Library to Multimedia Application", European Conference on Digital Libraries (EuroDL99), Paris, September 1999.

[19] L. Villard, C. Roisin and N. Layaïda, "A XML-based multimedia document processing model for content adaptation", Digital Documents and Electronic Publishing (DDEP00), September 2000.

[20] L. Weitzman and K. Wittenburg, "Automatic Presentation of Multimedia Documents Using Relational Grammars", Proceedings of 2nd ACM Conference on Multimedia, ACM Press, San Francisco, California, October 1994, pp. 443-451.

[21] Xalan, http://xml.apache.org, 2001.

[22] XSLT, http://www.w3.org/TR/xslt.html, 2001.