CONTEXTUAL INFORMATION IN VIRTUAL COLLABORATION SYSTEMS BEYOND CURRENT STANDARDS

Anna Carreras[1], Maria Teresa Andrade[2], Tim Masterton[3], Hemantha Kodikara Arachchi[4],
Vitor Barbosa[2], Safak Dogan[4], Jaime Delgado[1], Ahmet M. Kondoz[4]

[1] Universitat Politècnica de Catalunya, c/ Jordi Girona 1-3 Mòdul D6, 08034 Barcelona, Spain
{annac, jaime.delgado}@ac.upc.edu
[2] INESC Porto, Rua Dr. Roberto Frias 378, 4200-465 Porto, Portugal
{maria.andrade, vhsb}@inescporto.pt
[3] Thales Research and Technology (UK) Ltd, Reading RG2 0SB, Berkshire, UK
{Tim.Masterton}@thalesgroup.com
[4] I-Lab, CCSR, University of Surrey, Guildford GU2 7XH, Surrey, UK
{H.Kodikaraarachchi, S.Dogan, A.Kondoz}@surrey.ac.uk

ABSTRACT

Context-aware applications are fast becoming popular as a means of enriching users' experiences in various multimedia content access and delivery scenarios. Nevertheless, the definition, identification, and representation of contextual information are still open issues that need to be addressed. In this paper, we briefly present our work developed within the VISNET II Network of Excellence (NoE) project on context-based content adaptation in Virtual Collaboration Systems (VCSs). Based on the conducted research, we conclude that MPEG-21 Digital Item Adaptation (DIA) is the most complete standardization initiative to represent context for content adaptation. However, tools defined in MPEG-21 DIA Usage Environment Descriptors (UEDs) are not adequate for Virtual Collaboration application scenarios, and thus, we propose potential extensions to the available UEDs.

1. INTRODUCTION

In today's diverse multimedia application scenarios, users wish to access contents at anytime and anywhere. This trend has led to an increasing heterogeneity of coexisting terminals and networks, which has also rendered the use of content adaptation techniques essential in such scenarios.

Contextual information is used in many different context-aware applications. Its use is especially significant for supporting content adaptation, as it can help to take the optimum adaptation decision based on the usage scenario, which in turn has the potential for maximizing user's satisfaction.

In this paper, we first analyze the existing standardization initiatives contributing to the identification and representation of context for multimedia content adaptation. Subsequently, we present a context-aware content adaptation platform developed within the VISNET II NoE project [1]. During the following discussions, we first describe the Virtual Collaboration application scenario, in which our work is applied. Then, we propose potential extensions to the available MPEG-21 DIA UEDs, as our previous research showed that the existing UEDs do not present adequate descriptions for Virtual Collaboration application scenarios [2].

2. STATE-OF-THE-ART

Contextual information can be any kind of information that characterizes the complete delivery and consumption environment. Based on previous research [3], this information can be grouped into four main context classes: Computing, User, Physical, and Time. Contextual information is generated by sensors. Any hardware appliance, software component or entity capable of generating information describing some aspect of the context of usage can be designated as a sensor [4].

One of the main challenges faced in context-aware computing has been, until recently, the lack of standardized models to represent and process, context. Although recently, standardization bodies, such as the World Wide Web Consortium (W3C) and Motion Picture Experts Group (MPEG), have started to work on specifications to represent and exchange context, they provide mechanisms with limited capabilities. More specifically, there is almost no support for establishing relationships among contextual information, nor with constraints. Systems tend to simply use the sensed context in an uncoordinated way. This effectively limits the understanding that these systems have of the complete context of usage. The same contextual information sensed from the network could lead to different conclusions if combined with contextual information regarding the natural environment and/or the terminal. For the same sensed network context, different adaptation decisions could thus be taken. Accordingly, current standards provide efficient frameworks for developing moderately simple context-aware systems. However, they

do not answer the needs of more complex and demanding context-aware applications.

MPEG-21 [5], [6] is the ISO/IEC standard currently under its final phase of development within MPEG. It provides a set of specifications, descriptions and tools to facilitate the transactions of multimedia content in heterogeneous network environments. Part 7 of the standard (Digital Item Adaptation, DIA [7]) provides a set of tools to describe characteristics and capabilities of networks, terminals, and environments, as well as preferences of users. The tools: Usage Environment Descriptor (UED), Adaptation Quality of Service (AQoS), and Universal Constraints Descriptor (UCD), allow to represent both contextual information in the form of descriptions and constraints, as well as operations that can be performed upon the content and the results that can be expected. The W3C is presently focusing on the development of technologies to implement the Semantic Web. The accomplishment of this goal requires the use of common semantics to describe content, services or any resource on the Web. As such, the work of W3C is very much connected with context awareness. The Resource Description Framework (RDF) [8], the RDF Schema (RDFS) [9], and the Web Ontology Language (OWL) [10], are specifications of the W3C that constitute general description frameworks suitable to express metadata, and hence context. The W3C has also developed the Composite Capability/Preferences Profiles (CC/PP) specification [11], which defines an RDF-based framework for describing terminal capabilities and user preferences.

The meaning of context-awareness is to make use of sensed context to build understanding at higher levels of conceptual abstractions, the way humans do [12]. This kind of reasoning usually requires the use of models with rules and relations, to allow combining and relating sensed context. Ontologies can provide the support for reasoning and concepts' formalization, thus enabling to build complex context models based on knowledge. The work presently conducted by some research groups, thus addresses the use of ontologie to obtain high-level semantic descriptions of the context of usage [13]-[15].

3. VISNET II NOE CONTEXT-AWARE CONTENT ADAPTATION PLATFORM

3.1. Architecture

A distributed architecture is proposed for the VISNET II NoE context-aware content adaptation platform. As shown in Fig. 1, it consists of four different functional modules: Adaptation Decision Engine (ADE), Adaptation Authoriser (AA), Context Providers (CxPs), and Adaptation Engine Stacks (AESs) comprising Adaptation Engines (AEs) within. These modules have been developed as independent units that interact with each other through Web Services-based interfaces. The most notable advantage of the proposed distributed modular architecture is that it ensures

scalability. Hierarchical distribution of each functional module at strategic locations within the network is possible with this architecture. For example, an ADE could act as a Service Provider located somewhere in the network, dealing with various AESs situated near each of the terminals, and with only one AA (as shown in Fig. 4); but the same platform could support many other different topographies, such as, for example, having several ADEs located on the user/terminal side with only one AES situated on the network and several AAs. Well-defined interfaces based on open standards guarantee interoperability and flexibility of freely adding, removing, and migrating modules.

Based on contextual information received from various Context Providers (CxPs), the Adaptation Decision Engine (ADE) determines the optimum adaptation options that can maximize the user satisfaction across the network. An ontology based approach is used to implement the adaptation decision taking algorithm. The use of ontologies in this ADE provides the platform with context-aware analysis capabilities closer to real-world situations. Here, the two-layer ontology model shown in Fig. 2 has been utilized.

This model is developed using the OWL specification. The core ontology layer provides descriptions of generic concepts and rules that can be used in any multimedia content consumption scenario. This layer is based on MPEG-21 DIA, in particular the UED tool, with additional information related to the content being consumed (as of MPEG-7 Multimedia Description Scheme (MDS) media format descriptors). The external layer, *i.e.,* domain-specific layer, provides rules dedicated to a given application. Multiple domain-specific ontologies can thus co-exist in this layer.
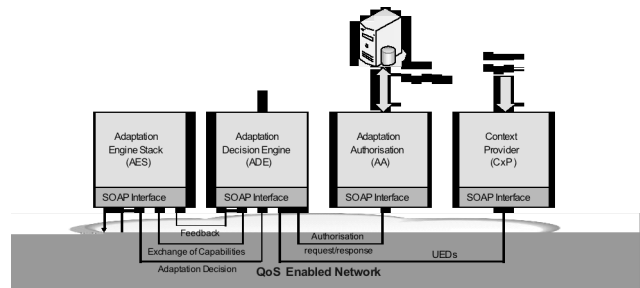

Fig. 1. Functional architecture of the proposed platform for context-aware and DRM-enabled multimedia content adaptation
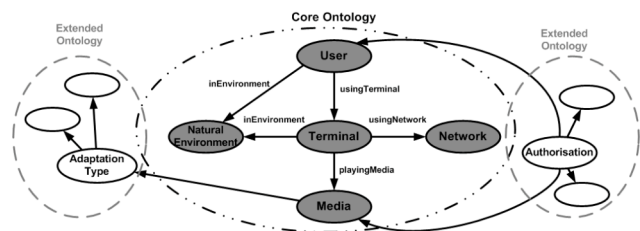

Fig. 2. Context ontology overview

One of the innovative features of the proposed context-aware content adaptation platform is the Adaptation Authoriser (AA), which ensures the governed use of protected content through Digital Rights Management (DRM). The only standardisation initiative to integrate DRM and adaptation also comes from the MPEG-21 DIA standard. It offers the possibility of extending MPEG-21 REL [16] licenses with more fine-grained descriptions of the permitted conversions and their associated limit constraints. A detailed description of our work on the Adaptation Authorisation can be found in [17].

The Adaptation Engine Stack (AES) enables the execution of a variety of adaptations that can be dynamically configured and requested on the fly.

CxPs are software or hardware based entities that are able to generate and provide explicit contextual information. The low-level contextual information generated by these CxPs, once acquired and represented according to a standard format, will be used to infer higher-level concepts, and thus assist the adaptation decision operation. The next subsection elaborates the types of context profiles made available to the ADE.

3.2. MPEG-21 DIA based context profiles

The use of standards is instrumental to enable interoperability among systems and applications, and across services. The standardized format used in this work is the MPEG-21 DIA specification. It specifies appropriate XML schemas to represent the low-level contextual information. In particular, the MPEG-21 DIA UED tool provides four main types of descriptors: *User*, *Terminal*, *Network*, and *Natural Environment*. Based on this division, four context profiles have been created, as illustrated in Fig. 3. With these profiles, each CxP needs only to know and implement its own sphere of action resulting in a level of interoperability enhancement.
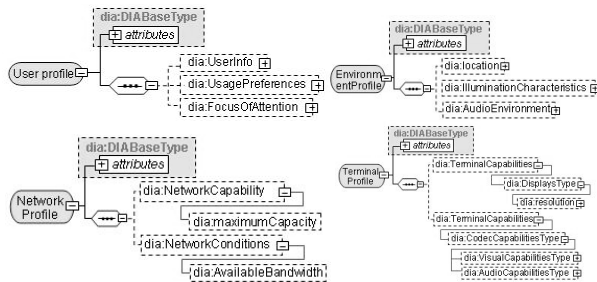

Fig. 3. Virtual Collaboration context profiles

4. VIRTUAL COLLABORATION APPLICATION SCENARIO

4.1. Virtual Office for Collaboration

Virtual Collaboration allows remotely located partners to meet in a virtual environment using the state-of-the-art audiovisual communication technologies coupled with shared electronic documents and applications. Fig. 4 shows a diagram of how the VISNET II NoE context-aware content adaptation platform, as described in the previous section, could be applied in a Virtual Collaboration scenario. In this paper, we focus particularly on its application in a business/office context, in which a number of office workers and/or clients of an organization act as the different agents involved in the collaboration.

Virtual Collaboration Systems (VCSs) typically have a number of users with heterogeneous terminal types communicating and exchanging media with each other through heterogeneous communications networks. For simplicity here, we identify three main types of user interface terminals: 1) Large terminals, which have a fixed location and a high data rate wired connection to the VCS. These terminals support multiple users, and an example would be the VCS terminals at a Headquarters (HQ) Meeting Room (Fig. 4). 2) Small terminals, which can be portable yet have a fixed location during the collaboration session with wireless or wired connection to the VCS, as represented by the Home-based Worker in the figure. 3) Mobile terminals, which can be moved around while in use and are designed for a single user, as for example the one used by the Mobile Field Worker in the same figure.

Furthermore, in a Virtual Office for Collaboration, we can identify the following types of users: Team Manager, Team Leader and Deputy Team Leader, Team Member, Customer, Interviewer, and Interviewee.

In brief, we have described a highly heterogeneous scenario here, in which context-aware content adaptation is required. In the next subsection, we will focus on identifying the contextual information that can be gathered for use in this kind of a scenario.

Fig. 4. Virtual Office for Collaboration

4.2. The use of context in a Virtual Collaboration application scenario

In this section, we identify the contextual descriptors used in the described application scenario. Aligned with the proposed MPEG-21 DIA context profiles, our work

distinguishes context related to the terminal, network, user, and natural environment. Tools defined in MPEG-21 DIA UEDs are not adequate for Virtual Collaboration application scenarios, and thus we propose potential extensions to the available UEDs as shown in Fig. 6.

The ADE receives concrete contextual information, such as the display size (in order to adjust the resolution), the maximum frame rate, the battery time remaining, and the codecs supported, from the terminals. On the network side, context is slightly more abstract. The current MPEG-21 network context provides information on available bandwidth, delay, packet loss rate and bit error rate, but a more sophisticated measure of the link conditions including the pattern of packet losses is required to make an intelligent assessment of the forward protection required on the transmitted communications stream. For example forward error correction is beneficial in a random packet loss scenario but detrimental in a congested (burst) packet loss scenario because the extra packets add to the congestion.. Therefore, we propose to add a new element inside the MPEG-21 DIA UED *NetworkConditions* (Fig. 3) named *Congestion* that will specify the maximum burst lengths with errors so as to know whether it is caused by noise or by congestion. An example of the contextual information that the ADE receives from the network is shown in Fig. 5.

The user and the natural environment related contexts are equally important. As detailed in the previous section, VCS allows multiple users to log into a given interface terminal simultaneously, and these users have multiple assigned roles (e.g., manager, team member, etc). However, they also have multiple assumed roles (e.g., speaker, inattentive participant, etc). Thus, we consider it is necessary to include some user context descriptors in order to assist with identifying and making use of these assumed roles for content adaptation.

```
| <DIA xmlns="urn:mpeg:mpeg21:2003:01-DIA-NS"
 xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
 xmlns:mpeg7="urn:mpeg:mpeg7:schema:2001">
   <Description xsi:type="UsageEnvironmentType">
     <UsageEnvironmentProperty xsi:type="NetworksType">
       <Network>
         <NetworkCharacteristic xsi:type="NetworkCapabilityType" maxCapacity="512000"
 minGuaranteed="32000" errorCorrection="true" errorDelivery="false"/>
         <NetworkCharacteristic xsi:type="NetworkConditionType"
 startTime="2008-12-10T13:13+00:00" duration="PT330N1000F">
           <AvailableBandwidth maximum="256000" average="128000"/>
           <Delay packetTwoWay="150" delayVariation="32"/>
           <Error packetLossRate="0.15"/>
           <Congestion maxBurstLength="12"/>
           <!-- New descriptor maxBurstLength in number of packets -->
         </NetworkCharacteristic>
       </Network>
     </UsageEnvironmentProperty>
   </Description>
 </DIA>
```

Fig. 5. Example of network's contextual information

With the integration of audiovisual processing techniques developed within the VISNET II NoE project such as gesture recognition, face detection, or speech recognition, we propose the extraction of abstract contextual concepts such as the "user attentiveness" or "has floor" descriptors.

These can be mapped to user identities and roles, allowing network communication and terminal processing resources to be focused on the information of most relevance to the users by adaptation of the streams that each one receives. For example, this allows the audio and video input capturing to be focused on a particular individual without a detailed manual camera control activity.
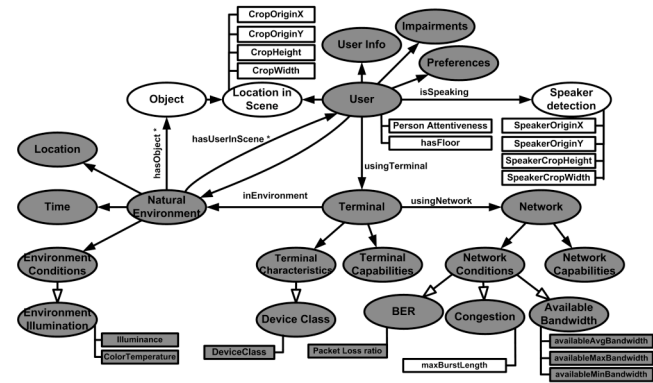

Fig. 6. Context Aware Ontology Model

Beyond the complete description of users, objects in a scene are also important elements of a VC scenario; therefore, we include the presence of objects as new context descriptors. This is due to the fact that, a user does not always focus on all of the VCS participants during a meeting session, but also sometimes on the white-board, a screen or an object while discussing various matters depending on how the meeting discussions progress.

The extended context descriptors proposed for the Virtual Office application scenario that have been justified so far are shown in white in Fig. 6. The elements in dark-grey were already identified in Section 3, and are expressed in MPEG-21 DIA UED. Furthermore, Table 1 sums up the motivation for the proposed extensions, their nature, and their intended use in terms of adaptation.

System condition	Nature	Profile	Extended Descriptor	Adaptation	
				Without the extended descriptor	**With the extended descriptor**
Congestion in the communication network	Static or dynamic conditions of the physical environment	Network	Dia:Congestion	With the existing descriptors (error rate, delay and available b/w) it is not possible to determine whether the packet errors	By means of error pattern analysis (correlation of error patterns over time or burst length detection) the cause of the losses can be inferred.

				occur in bursts or at random	
User X is not paying attention	Preferences of the user	User	Dia:UserAttentiveness	With existing descriptors only the speaker or floor holder can be singled out for special treatment	With an attentiveness measure the speaker can be alerted when some group inattentiveness threshold is reached or content could be prioritised for those who are paying attention
User Y starts speaking	Preferences of the user	User	Dia:hasFloor	Equal priority level for all of the participants for error resilience, focus of attention, scaling, etc adaptations	Higher priority level for those who have the floor over the others

5. CONCLUSIONS

This paper justifies the need of an extension of the MPEG-21 DIA UED descriptors when context-aware content adaptation is applied in VCSs. Thus, we have firstly presented our context-aware content adaptation platform developed within the VISNET II NoE project, and then made an accurate description of the Virtual Collaboration application scenario. Our conclusion is to propose an amendment of new context descriptors related to the presence of objects or multiple users in a scene, and to the recognition of the speaker in the MPEG-21 DIA UED, as they are a core part of Virtual Collaboration sessions. We also justify why MPEG-21 DIA is the most complete standard for context-aware content adaptation.

6. ACKNOWLEDGEMENT

7. REFERENCES

[1] "VISNET II NoE," 2007. [Online]. Available: http://www.visnet-noe.org/.

[2] M.T. Andrade et al., "Using context to assist the adaptation of protected multimedia content in Virtual Collaboration applications," in *Proc. the 3rd IEEE International Conference on Collaborative Computing: Networking, Applications and Worksharing*, New York, NY, USA, 12-15 Nov. 2007.

[3] G. Chen and D. Kotz, "Survey of Context-Aware Mobile Computing Research," Dartmouth College, Dept. of Computer Science, Technical Report TR2000-381, 2000.

[4] S. Jang, "Implementation of Context-Aware Application Model in Smart Environments," Gwangju Institute of Science and Technology, PhD thesis 2005.

[5] MPEG-21 part-1: ISO/IEC TR 21000-1:2004, Information Technology, Multimedia Framework: Vision, Technologies and Strategy, Nov. 2004.

[6] I. Burnett, F. Pereira, R. Walle, and R. Koenen, "The MPEG-21 Book," John Wiley and Sons Ltd, 2006.

[7] MPEG-21 part-7: ISO/IEC 21000-7:2007 Information Technology, Multimedia Framework, Digital Item Adaptation, Dec. 2007.

[8] Recommendation W3C: "Resource Description Framework (RDF): Concepts and abstract syntax," Feb. 2004.

[9] D. Brickley and R.V. Guha, W3C Recommendation, "RDF Vocabulary Description Language 1.0: RDF Schema," 2004.

[10] L. McGuinness and F. van Harmelen, W3C Recommendation, "Web Ontology Language Overview," Feb. 2004.

[11] Recommendation W3C: "Composite Capability / Preference Profiles (CC/PP): Structure and vocabularies 1.0," Jan. 2004.

[12] L. Harvel1, L. Liu, G. D. Abowd, Yu-Xi Lim, C. Scheibe and Chris Chatham, "Context Cube: Flexible and Effective Manipulation of Sensed Context Data", Lecture Notes in Computer Science, Springer Berlin, Vol. 3001, pp. 51-68, March 2004.

[13] T. Chaari, F. Laforest, and A. Celentano, "Adaptation in Context-Aware Pervasive Information Systems: The SECAS Project," Int. J. of Pervasive Computing and Commun., vol. 2, no. 2, June 2006.

[14] K.-C. Lee, J.-H. Kim, J.-H. Lee and K.-M. Lee, "Implementation of Ontology Based Context-Awareness Framework for Ubiquitous Environment", IEEE International Conference on Multimedia and Ubiquitous Engineering (MUE'07), Seoul, Korea, April 2007.

[15] J. Serrano, J. Serrat, J. Strassner, "Ontology-Based Reasoning for Supporting Context-Aware Services on Autonomic Networks", IEEE International Conference on Communications 2007 (ICC-2007), Glasgow, Scotland, 24-28 June 2007.

[16] A. Carreras and J. Delgado, "A new type of contextual information based on the adaptation authorisation," in *Proc. the 9th Int. Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS 2008)*, Klagenfurt, Austria, 7-9 May 2008.

[17] MPEG-21 part-5: ISO/IEC 21000-5, Information Technology, Multimedia Framework, Rights Expression Language, Mar. 2004.