

# **A Multimedia Service with MPEG-7 Metadata and Context Semantics**

Yiwei Cao, Ralf Klamma, and Maziar Khodaei

Lehrstuhl für Informatik 5  
RWTH Aachen University  
Ahornstr. 55  
52056 Aachen Germany  
{cao|klamma|khodaei}@dbis.rwth-aachen.de

**Abstract:** With the emergent and rapid development of mobile technologies, more and more multimedia applications run on handheld devices in mobile networks. This raises new challenges to mobile multimedia information systems. In this paper three technical aspects are considered as tightly intertwining and crucial for developing mobile multimedia applications: context awareness, multimedia adaptation, and mobility. An approach to integrating multimedia metadata standards such as MPEG-7 and MPEG-21 into context-aware ontologies effectively and seamlessly is proposed and related algorithms are extended. A context-aware mobile multimedia application (CA3M) is implemented and evaluated as a proof of concept. The evaluation results of the prototypes show the soundness of the concept. Open issues and development potentials for further research in the area of context-aware mobile multimedia information system development are addressed.

## **1 Introduction**

Mobile applications have been developed with advancement and rapidly in recent years. Mobile devices like iPhones make a revolution of mobile application development. The conventional barriers of mobile application development like limited capacities have made great progress. It is quite normal to have an 8 GB, 16 GB hard disk or even larger. The display sizes have also been maximized within the device size available. All that innovative progress makes mobile multimedia applications more and more attractive and important for users.

Nowadays mobile devices and applications are featured with new focuses such as integrity, multimodality, “one device fitting all” or “only carrying one gadget at a time” etc. Multimodality is realized both in information input and output channels. Google supports speech-based keyword search, while iPhone has three built-in sensors, the accelerometer, the proximity sensor, and the ambient light sensor, to detect movements even movement intentions of the mobile device, besides GPS and camera etc. Making a call is becoming the side function of cell phones, because cell phones are also MP3-players, video players, cameras, recorders, gaming-consoles, digital books, personal digital assistants to organize contacts and appointments, and have much other functionality.

A recent report shows that the variety of functionality available on cell phones raises usage complexity as well. And many users still only stick to the phone call function of their cell phones<sup>1</sup>. Thus, the intuitiveness of the mobile user interfaces is still limited. Mobile multimedia information systems with enhanced context awareness are the potential approaches to reduce the complicated user device interaction.

In addition, fault tolerance and speed are identified as the most critical aspects for general multimedia applications, because both audiovisual media and metadata including control information are processed simultaneously [24]. Hence, we consider the context uncertainty problems as an important aspect at mobile multimedia application development. Furthermore, three technical aspects are considered as tightly intertwining and crucial for developing mobile multimedia applications: context awareness, multimedia adaptation, and mobility. Uncertainty reasoning is based on context reasoning which is enabled both by metadata based multimedia adaptation and context-aware multimedia adaptation (cf. Figure 1).

The rest of the paper is organized as follows. The state-of-the-art technologies are discussed in Section 2. An approach to integrating multimedia metadata standards such as MPEG-7 and MPEG-21 into context-aware ontologies effectively and seamlessly is proposed in Section 3. Section 4 provides an insight into implementation and evaluation of a context-aware mobile multimedia application (CA3M). CA3M delivers a service of context-aware multimedia search and multimedia adaptation which can be accessed from mobile devices. Section 5 concludes this paper with open issues and development potentials for further research in the area of context-aware mobile multimedia information system development.

## 2 Related work

Much research work has been done in the area of context modeling, context-aware information systems, mobile computing, and multimedia systems, multimedia metadata standards, and ontologies. Research progresses in these areas contribute to the ubiquitous computing as well as pervasive computing paradigms. The context information might be uncertain which needs to be reduced via an appropriate context model.

---

<sup>1</sup> Rhizomik Initiative, <http://rhizomik.net>.

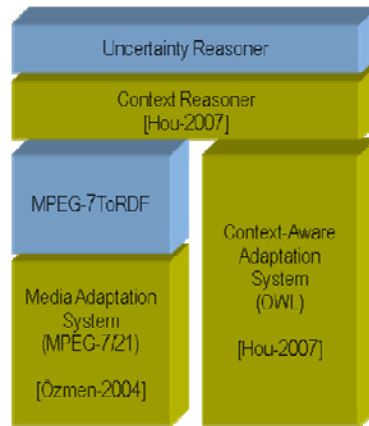


Figure 1: Conceptual approach

## 2.1 Context awareness, mobility and multimedia adaptation

First of all, these three main concepts are reviewed from related literature. Any piece of information for interaction between user communities and applications can be observed as a context. Dey defines context as piece of information characterizing certain entity including person, location, or a physically assessable object [7]. Correspondingly, context-aware systems refer to any information system that deals with and makes use of context information. Representation of context information is an important part of research in pervasive computing and considered as a sub field of knowledge engineering, within knowledge management and artificial intelligence. Context-aware mobile systems can be applied in such fields as presenting context information to users, services adaptation to mobile users, as well as handing context information at real time [25].

Mobility assures mobile device users of reaching information according to certain context such as user communities, location, temporal and device capacities anywhere at any time [19]. Kakiyara and Sorensen associated mobility with three dimensions which act as a whole: spatial, temporal, and context-dependent [11]. Among them, multimedia content is of high interest. Multimedia refers to content and data from more than one digital resource such as text, photographs, graphics, animation, audio, and video etc. [24]. Multimedia annotation, adaptation and different interactions are main aspects for multimedia applications. In order to deliver right information to right person at right time and right location, multimedia search results are adapted to users' preferences and up-to-data context information in certain environments.

## 2.2 Context uncertainty

It is difficult to avoid the problems with data uncertainty especially in context-aware systems. Inconsistency occurs between models and the real world, as well as among different local environment models. Four main factors are stated to lead to that inconsistency in [9]:

- *unknown* when no information about the property is available;
- *ambiguous* when several different reports about the property are available (for example, when two distinct location readings for a given person are supplied by separate positioning devices);
- *imprecise* when the reported state is a correct yet inexact approximation of the true state (for example, when a person's location is known to be within a limited region, but the position within this region cannot be pinpointed to the required (application-determined) degree of precision); or
- *erroneous* when there is a mismatch between the actual and reported states of the property.

Correspondingly, Quality of Context (QoC) or Quality of Information is often discussed in these cases and quality of context information depends greatly on the change of context sources, by which context is provided [14].

## 2.3 Context modeling

In order to deal with the context information as well as context uncertainty mentioned before, specific frameworks and data structures are needed to capture, manage, process and retrieve context information. In the field of knowledge engineering and artificial intelligence, context information must be at first collected and presented to the application in order to enable efficient context-aware adaptation. Therefore, a common representation format for the context information is required [17]. A well-defined context model is needed to define and store context data in machine readable forms in order to enhance interoperability. In [23] Strang and Linnhoff-Popien made an in-depth survey across several context-aware systems and compared the most relevant context modeling approaches including Key-Value, Markup Scheme, Graphical, OO, Logic-based and Ontology-based models. These approaches are based on different data structures and represent context information for machine processing and reasoning.

## 2.4 MPEG-7, MPEG-21 and RDF

With regard to various multimedia metadata, MPEG-7 provides a large set of pre-defined elements to describe multimedia contents. In particular, these elements are composed of two different types: Description Schemes (DS) and Descriptors (D). Depending on the fields of applications, a specific description scheme can be defined by freely combinable descriptors (i.e. tags). Each descriptor itself refers to a specific feature or attribute of multimedia content [12, 15]. In addition, the Description Definition Language (DDL) of MPEG-7 makes this standard more powerful than other metadata standards, since it allows the creation of new descriptors and description schemes within the standard. Hence, it makes the vocabulary of MPEG-7 to be extensible by employing the XMLSchema.

Together with MPEG-7, MPEG-21 provides a comprehensive framework for multimedia adaptation. In general, MPEG-21 consists of twelve parts [2, 12]. These parts are independent of each other. So the excerpts of the standard might also be applied. Alone for the purpose of managing multimedia contents in mobile end devices we confine the system to the MPEG-21 Digital Item Adaptation (DIA) and MPEG-21 Digital Item Declaration Language (DIDL). The MPEG-21 DIDL contains six top level descriptors (tags), of which the first four are particularly important for mobile data management.

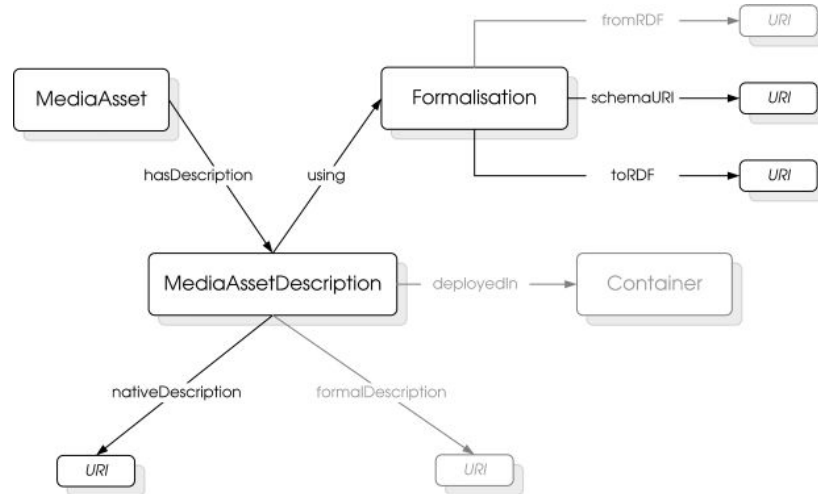


Figure 2: The ramm.x vocabulary [16]

However, neither MPEG-7 nor MPEG-21 multimedia metadata standard defines ontology which can be used for context modeling. Several ontology-based MPEG binding specifications are surveyed in the related research fields.

*RDFa-deployed Multimedia Metadata* (ramm.x) offers mapping from multimedia standards such as MPEG-7 to the Semantic Web whose documents are often described in RDF (Resource Description Framework). The ramm.x can be used for reuse of existing multimedia metadata, converting, validation and exchange of metadata via an easy-to-use vocabulary set (cf. Figure 2). Main requirements of ramm.x include enabling the resulting description to be consumable by a Semantic Web agent through being encoded in RDF; embedding reference to description in existing multimedia metadata format in (X)HTML; providing reference to services capable of mapping between a specific multimedia metadata format and RDF; and multiple descriptions available for one media asset (e.g. in different formats, covering different aspects), etc. Based on the ramm.x, an MPEG-7 ontology demonstrates the first successful practical realization [18] with MPEG-7 schema version of 2001. The MPEG-7 ontology specified by Rhizomik [18] shows its advantage in comparison to those approaches like MPEG-7 Ontology [10] Core Ontology for Multimedia (COMM) [4, 5] and aceMedia Visual Descriptor Ontology (aceMedia VDO) [1] (cf. Table 1).

Framework	Supporting format	Mapping	Comments
MPEG-7 Ontology [10]	OWL-full	-	-
COMM [4]	OWL-DL	-	-
aceMedia [1]	RDFS-DS	-	-
Rhizomik MPEG-7 Ontology [18]	OWL-full	XSD2OWL, XML2RDF	Confidence value can be set for uncertainty reasoning

Table 1 A comparison of MPEG-7 ontologies

### 3 System design of a mobile context-aware multimedia service

A service is designed and implemented to meet the following requirements. How can preferences of user communities, spatial and temporal context information contribute to the context-aware multimedia search? How can Semantic Web and multimedia metadata standards enhance multimedia search together?

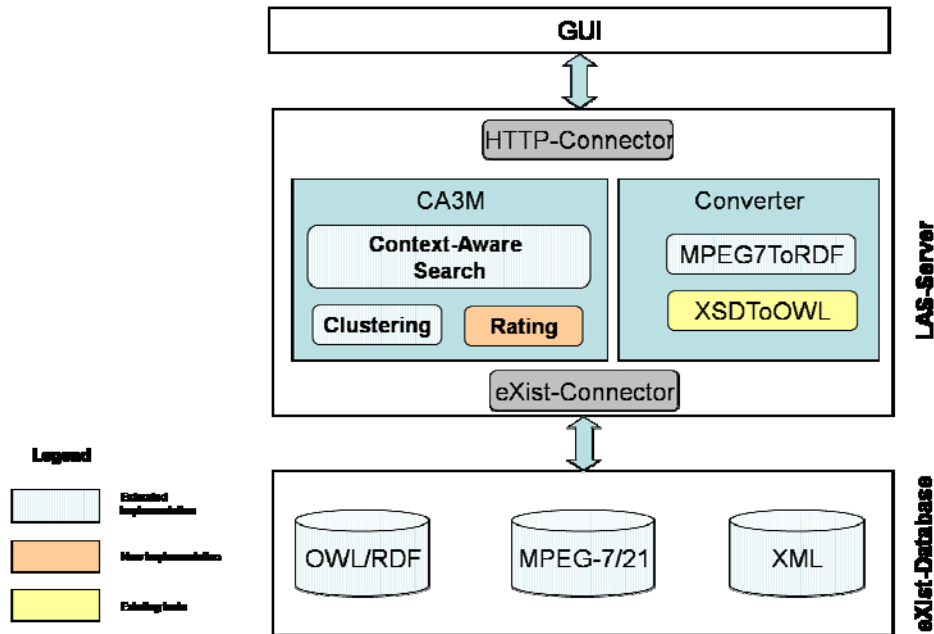


Figure 3: CA3M system architecture

### 3.1 Conceptual approach of CA3M

The CA3M (Context-Aware Mobile Multimedia) service enables multimedia search based on keywords specified by users and users' context. CA3M uses the Web Service technologies in order to provide easy access to mobile applications on the client side. On the server side, a set of services provide functionality such as context acquirement, context reasoning, and context querying. On the client side, mobile user interfaces can process users' requests and retrieve demanding multimedia results. For rapid prototyping we employ our previous research result, the Lightweight Application Server (LAS) [20] as the basic framework. LAS provides HTTP and SOAP connectors, which makes the implementation of client-server communication easier. In addition, the LAS components provide functionality such as the connector to a database. Ontologies using OWL, RDF and a media repository with metadata are maintained by a native XML database, e.g. eXist [8]. Moreover, the main benefit of LAS is that the services on it can be flexibly extended for specific applications by using the LAS Java APIs (cf. Figure 3).

Thus, CA3M is designed and implemented as LAS services within the LAS framework. Figure 4 illustrates the information flow in CA3M. Besides Context reasoning and multimedia search functionality, rating and clustering are provided. As the first step, user communities are grouped into different clusters according to their own interests or preferences. Users' rating to multimedia search results is carried out based on different user clusters and individual users. The rating mechanism by user communities can reduce context uncertainty via users' feedback to certain multimedia search results.

### 3.2 The MPEG-7 to RDF converter

The binding of MPEG-7 and Semantic Web encoded in RDF is one of the basic components beside the CA3M service. The contribution of this research work is mainly in two aspects. One is the adapted converting service which enables MPEG-7 and MPEG-21 multimedia standards to be understandable in order to realize the Semantic Web. This is an extension of the existing research by the Rhizomik Initiative<sup>1</sup>. The converter consists of transformation from RDF to HTML, XML Schema to OWL, and XML to RDF. An ontology schema based on a mapping of MPEG-7 schema to RDF is additionally specified.

The other contribution is that CA3M supports both XQuery (XML Query Language) and SPARQL (Simple Protocol and RDF Query Language). SPARQL is a tree-structured query language for RDF documents. The semantics of SPARQL is similar to SQL, while the query processing mechanisms of both are different [21, 22].

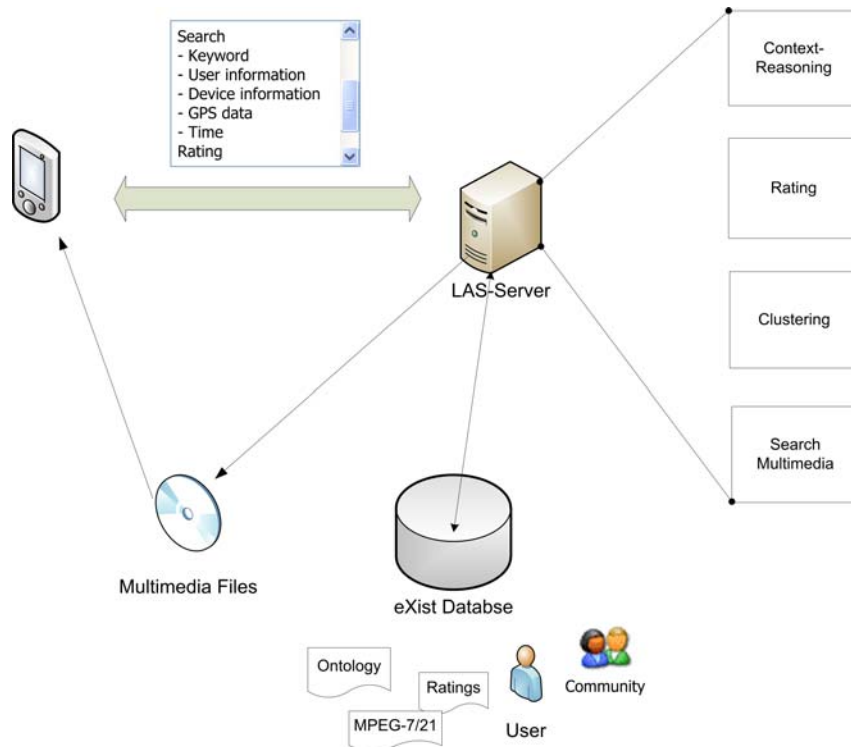


Figure 4: CA3M information flow



## 4 Implementation and evaluation

MPEG-7 Schema is mapped to ontology before MPEG-7 documents are transformed into RDF. Our converter is adapted and realized for the latest MPEG-7 Schema of 2004. For the context-aware query processing, two steps are carried out sequentially. Firstly, all queries related to context information are expressed and evaluated in XQuery. The to-be-searched multimedia metadata at this phase is in MPEG-21 DIDL format. Second, SPARQL queries are prepared for executing on RDF documents, after the converter has processed the metadata documents.

The MPEG-7 to RDF converter works within two packages, *XSD2OWL* and *XML2RDF*. *XSD2OWL* facilitates the transformation of MPEG-7 Schema into MPEG-7 Ontology, while *XML2RDF* transforms XML based MPEG-7 documents into RDF formats. A LAS service called MPEG7ToRDF Service has been implemented.

A set of evaluation work is done on the CA3M service prototype. In our project Virtual Campfire [6], there are a great amount of multimedia files with metadata stored on streaming servers, FTP servers, HTTP servers, and several multimedia or XML databases in a distributed computing environment. Over 300 main multimedia metadata files are converted from MPEG-7 into RDF. Then SPARQL queries are executed to obtain context-aware multimedia search results across the 300 MPEG-7 documents and their related MPEG-7 documents which e.g. save information about location using MPEG-7 *SemanticPlaceType*. The first evaluation result shows that more efforts could be made to optimize the service execution, because around 30 seconds are needed to handle the SPARQL queries at the current development stage.

## 5 Conclusions and outlook

Context awareness, mobility and multimedia adaptation are discussed as three main factors to enable context-aware multimedia adaptation and search on mobile platforms. This work is based on two accomplished prototypes of multimedia information systems with metadata standards based multimedia adaptation and context modeling and reasoning in the previous work [13, 3]. To make good use of and to advance the existing tools, a conceptual approach is proposed to enhance multimedia adaptation and search combining comprehensive multimedia standards such as MPEG-7/21 and context modeling and context awareness. Both technologies intertwine and are able to support context reasoning in a better way. Furthermore, they make it possible and promising to shape and perform measurements in order to reduce context uncertainty.

In our ongoing research, we will focus on mobility issues and context uncertain problems related to multimedia and context awareness. Within the UMIC (Ultra High-speed Mobile Information and Communication) research cluster of the German Excellence Initiative, a lot of research work for mobile context-aware multimedia services is challenging. P2P data management with regard to MPEG-7/21 metadata standards will be addressed. Data uncertainty and context uncertainty problems need to be in-depth identified, analyzed and handled. Mobile (Web) services will be developed to bridge the mobile social software on the higher application layer and the mobile wireless network technologies on the lower network layer. The performance of these services should also be improved.

## Acknowledgment

This work has been supported by the UMIC Research Centre, RWTH Aachen University. We would like to thank our colleagues for the fruitful discussions.

## Reference

- [1] Ace Media Project, <http://www.acemedia.org/aceMedia>, {December 2008}.
- [2] Burnett, I.; Van de Walle, R.; Hill, K.; Bormans, J.; Pereira, F.: MPEG-21: goals and achievements. *IEEE Multimedia*, 10(4): 60–70, 2003.
- [3] Cao, Y.; Klamma, R.; Hou, M.; Jarke, M.: Follow Me, Follow You - Spatiotemporal Community Context Modeling and Adaptation for Mobile Information Systems, In: *Proc. of the 9th International Conference on Mobile Data Management*, April 27-30, 2008, Beijing, China, pp. 108-115.
- [4] COMM: Core Ontology for Multimedia, <http://comm.semanticweb.org/>, {December, 2008}.
- [5] Bloehdorn, S.; Petridis, K.; Saathoff, C.; Simou, N.; Tzouvaras, V.; Avrithis, Y.; Handschuh, S.; Kompatsiaris, Y.; Staab, S.; Strintzis, M. G.: Semantic Annotation of Images and Videos for Multimedia Analysis. In: *Proceedings of the Second European Semantic Web Conference (ESWC 2005)*, Springer, 2005, pp. 592-607.
- [6] Cao, Y.; Spaniol, M.; Klamma, R.; Renzel, D.: Virtual Campfire - A Mobile Social Software for Cross-Media Communities, K. Tochtermann, H. Maurer, F. Kappe, A. Scharl (Eds.): *Proceedings of I-Media'07, International Conference on New Media Technology and Semantic Systems*, Graz, Austria, September 5 - 7, 2007, J.UCS (Journal of Universal Computer Science) *Proceedings*, 2007, pp. 192-195.
- [7] Dey, A. K.; Abowd, G. D.: Towards a Better Understanding of Context and Context-Awareness. In: *HUC '99: Proceedings of the 1st international symposium on Handheld and Ubiquitous Computing Bd. 1707*. London, UK: Springer, 1999; pp. 304-307.
- [8] eXist, Open Source Native XML Database, <http://exist.sourceforge.net/>, {December 2008}.
- [9] Henriksen, K.; Indulska, J.: Modelling and Using Imperfect Context Information, In: *PERCOMW 04: Proceedings of the Second IEEE Annual Conference on Pervasive Computing and Communications Workshops*. Washington, DC, USA: IEEE Computer Society, 2004, pp. 33-37.
- [10] Hunter, J.: Adding Multimedia to the Semantic Web - Building an MPEG-7 Ontology. In: *Proceedings of the First Semantic Web Working Symposium (SWWS)*, Stanford, USA, 2001, pp. 261-281.

- [11] Kakihara, M.; Sorensen, C.: Mobility: an extended perspective. In: Proceedings of the Hawaii International Conference on System Sciences, New York, NY, USA: IEEE Computer Society, January 2002; pp. 1756-1766.
- [12] Kosch, H.: Distributed Multimedia Database Technologies Supported by MPEG-7 and MPEG-21, CRC Press, 2003.
- [13] Klamma, R.; Spaniol, M.; Cao, Y.: Community Aware Content Adaptation for Mobile Technology Enhanced Learning, In: W. Nejdl, K. Tochtermann (Eds.): Innovative Approaches to Learning and Knowledge Sharing, Proceedings of the 1st European Conference on Technology Enhanced Learning (EC-TEL 2006), Hersionissou, Greece, October 1-3, LNCS 4227, Springer-Verlag, 2006, pp. 227-241.
- [14] Lei, H.; Sow, D. M.; Davis, J. S.; Banavar, G.; Ebling, M. R.: The design and applications of a context service. In: SIGMOBILE Mobile Computing Communication, Rev. 6 (2002), No. 4, pp. 45-55.
- [15] Martinez, J. M.; Gonzalez, C.; Fernandez, O.; Garcia, C.; de Ramon, J.: Towards universal access to content using MPEG-7, In: Proceedings of the 10th ACM International Conference on Multimedia, ACM Press, 2002, pp. 199-202.
- [16] ramm.x: RDFa-deployed Multimedia Metadata. <http://sw.joanneum.at/rammx/>, {December 2008}.
- [17] Rothmel, K.; Bauer, M.; Becker, C.: Sonderforschungsbereich 627: Nexus - Umgebungsmodelle für mobile kontextbezogene Systeme. it - Information Technology 45(5): 293-; 2003.
- [18] Roberto Garcia Gonzalez @ Rhizomik. <http://rhizomik.net/~roberto/>, {December 2008}.
- [19] Roy, N. L. S.; Scheepers, H.; Kendall, E.; Saliba, A.: A comprehensive model incorporating mobile context to design for mobile use. In: Proceedings of the 5th Conference on Human Computer Interaction in Southern Africa, January, 2006; pp. 22-30.
- [20] Spaniol, M.; Klamma, R.; Janßen, H.; Renzel, D.: LAS: A Lightweight Application Server for MPEG-7 Services in Community Engines, In: Proceedings of I-KNOW '06, 6th International Conference on Knowledge Management, Graz, Austria, September 6 - 8, ser. J.UCS (Journal of Universal Computer Science) Proceedings, K. Tochtermann and H. Maurer, Eds. Springer-Verlag, 2006, pp. 592-599.
- [21] SPARQL Query Language for RDF. <http://www.w3.org/TR/rdf-sparql-query/>, {December 2008}.
- [22] SPARQL Tutorial. <http://jena.sourceforge.net/ARQ/Tutorial/>, {December 2008}.
- [23] Strang T.; Linnhoff-Popien, C.: A Context Modeling Survey, In: First International Workshop on Advanced Context Modelling, Reasoning and Management at UbiComp, Nottingham, UK, September 2004.
- [24] Steinmetz, R.; Nahrstedt, K.: Multimedia Systems, Springer-Verlag, 2004.
- [25] Zhang, D.; Gu, T.; Pung, H.: A Middleware for Building Context-Aware Mobile Services. In: Proceedings of IEEE Vehicular Technology Conference, 2004.