

Institute of Information Technology @ University Klagenfurt

Project Description Letter

Research Topics of the ADMITS Project

Status: May 2004

1 ViTooKi – The Video Tool Kit

Over the last year, the major research topics concerning the ADMITS project have been unified into one large reference software tool kit, called ViTooKi -- the Video Tool Kit. ViTooKi is an operating system independent, powerful high-level C++ multimedia library developed to simplify the implementation of multimedia applications. For example, ViTooKi supports standard compliant streaming of MPEG-1, MPEG-2, and MPEG-4 videos via RTSP and RTP/UDP, standardized RTP extensions to allow intelligent retransmission of lost frames, real-time adaptation according to the clients' terminal capabilities, MPEG-21 descriptors for describing terminal capabilities and user preferences, and MPEG-7 descriptors for adding semantic information to videos. The tool kit also contains a number of larger multimedia applications which make use of the main library, like an adaptive media server, an adaptive media proxy, and a multi-video player.

Actual Internet drafts on RTP extensions on immediate feedback and retransmissions were implemented and investigated in various network scenarios subjected to packet loss (see [KH04] for technical details). Dynamic stream adaptation was conducted by temporal adaptation, using various different algorithms for different packet prioritizations. Additionally, various kinds of statistics like PSNR calculations, client buffer estimations, and stream bandwidth detection are an integral part of the test software. Keywords like proxy cache replacement strategies, TCP-friendliness, buffer management, stream switching and stream adaptation are all tackled in the test implementation and will be an integral part of the next ViTooKi file release by the end of this year. To make the working prototype for video streaming, proxying and adaptation available to a broad range of interested experts all over the world, the ViTooKi project went open source via the well-established sourceforge.net¹ platform, where CVS access to the source code is provided to everyone.

2 QBIX: A Quality Based Intelligent Proxy for MPEG-4 Videos

QBIX is a quality-based intelligent proxy which introduces a novel concept for a brokering proxy. It implements a "What You Need is What You Get" (WYNIWYG) video service, where users do not get just the available video quality nor the best possible quality, but exactly that quality which they need and which they are ready to pay for. To do so, the proxy takes into account client terminal capabilities, user preferences, and network conditions to serve the appropriate video quality. The proxy thereby may act both as a media gateway and as a proxy cache. If a video is getting popular for users both with high and with low terminal capabilities then QBIX not only transcodes the original video quality to different lower qualities in real-time, but it also caches the generated lower quality variants of the video.

QBIX also incorporates a cost model that controls the decisions of the proxy. To our knowledge this is the first work on a quality aware video cache that combines partial caching in the quality domain with a differentiated model of user preferences, of

¹ For more information please visit the URL: <http://vitooki.sourceforge.net>.

video variations and of the caching costs. A nice feature of this approach is that we get the gateway and a basic filtering functionality dynamically and for free as a "side effect".

The QBIX implementation is an integral part of the ViTooKi software tool kit. More technical details about the system can be found in [SBH+03] and [SBH04].

3 | **ADMS: An Adaptive Distributed Multimedia Streaming Server**

In some cases a defensive quality adaptation of stream content (as e.g. done by QBIX) may not be applicable due to legal constraints, or when the client is not satisfied with a lower stream quality caused by network resource shortages. In this cases an offensive adaptation (i.e. replication or migration) of both streaming services and requested content may help.

ADMS represents a distributed streaming server which supports an offensive adaptation of both streaming services and content in Internet settings. It searches for adequate available server resources at runtime in order to serve the original stream quality even when faced with network resource shortages. To do so, the server is organised in four basic components where each component fulfils one particular task in a streaming scenario. The four basic tasks are data storage and management, data distribution, data collection and streaming, and central system control. A concrete composition of ADMS component instances in a distributed environment defines an ADMS configuration. Taking into consideration a given set of client requests including their QoS requirements, an ADMS configuration can be optimal or sub-optimal. To determine whether a configuration is optimal, server characteristics and network characteristics between server nodes are queried from the underlying middleware **Vagabond2** and compared against the QoS requirements of the client requests. Vagabond2 manages a dynamic set of servers (called "harbours") in the Internet which serve as runtime environment of so-called "adaptive server applications" (ASAs). Each ADMS component is realized as an Vagabond2 ASA in order to be able to execute it on any Vagabond2 server. If an ADMS configuration is not optimal regarding the client requests' requirements, Vagabond2 enables for a replication or migration of the code and/or data of an ASA from one Vagabond2 server to another. Taking into account the mean time needed for an offensive adaptation, the set of non-servable client requests must only include requests which need not be immediately served, but at a certain future time point. As a consequence, changes to an ADMS configuration are of proactive and not reactive nature.

Currently, the existing ADMS system is extended with algorithms for quality aware replication of video variations. To do so, the proactive ADMS server needs to know future behaviour patterns of its users. These patterns can be obtained from outside the system or be predicted from observations about the past. The patterns tell which clients or groups of clients are interested in which video topic at which time instants or periods. Different metrics are examined which allow the server to choose among proactive adaptation steps for future client behaviour.

Technical details about the ADMS and Vagabond2 systems can be consulted at [TSK+03] and [TBG+04].

4 Knowledge-based Media Adaptation

This project operates in the area of *Digital Item Adaptation* within the MPEG-21 framework. Since the scope of MPEG-21 is limited to the definition of description tools, this project deals with the internals of the adaptation process itself. It aims at developing an adaptation decision-taking engine which automatically generates a multi-step adaptation plan for a given media resource to fit the client's needs. Thereby, the problem of finding a suitable set of adaptation steps is viewed as a classical state-space planning problem. The goal of a planner is to find a sequence of actions in order to reach the defined goal state. The start state of the planning process is the original media resource, described by means of MPEG-7 media information and the goal state is the usage environment of the receiving terminal, described by means of MPEG-21 meta data. The actions that have to be performed to arrive from the start state at the goal state are the adaptation steps. To capture the semantics of the transformation, the adaptation steps (i.e. the adaptation algorithms) are also described. The generated adaptation plan is then executed by the adaptation engine, which also transforms the MPEG-7 description in a way that it is valid for the adapted resource. For more technical details about the system architecture see [JLH+04] and [JLH04].

5 Universal Multimedia Access

Universal Multimedia Access (UMA) refers to the ability for any type of terminals to access and consume a rich set of multimedia content. Ideally, this is achieved seamlessly over dynamic and heterogeneous networks and devices, independent of location or time, and taking into account a wide variety of possible user preferences.

This project addresses a number of problems in this context (e.g. diversity of media formats, scalable video formats, and interactive content) by contributing to the development of a generic framework for bitstream adaptation using XML-based descriptions in the scope of MPEG-21's Digital Item Adaptation (DIA). Figure 1 depicts the framework's architecture. A bitstream is defined as a structured sequence of binary symbols. DIA uses XML to describe the high-level structure of a bitstream, i.e., how it is organized in packets or layers of data. The resulting XML document is called Bitstream Syntax Description (BSD) [PHH+03]. In most cases, the BSD does not describe the bitstream on a bit-per-bit basis and is itself scalable. Thus, it may describe the bitstream at different syntactic layers, e.g., finer or coarser levels of detail such as frames or scenes for a video resource, depending on the application.

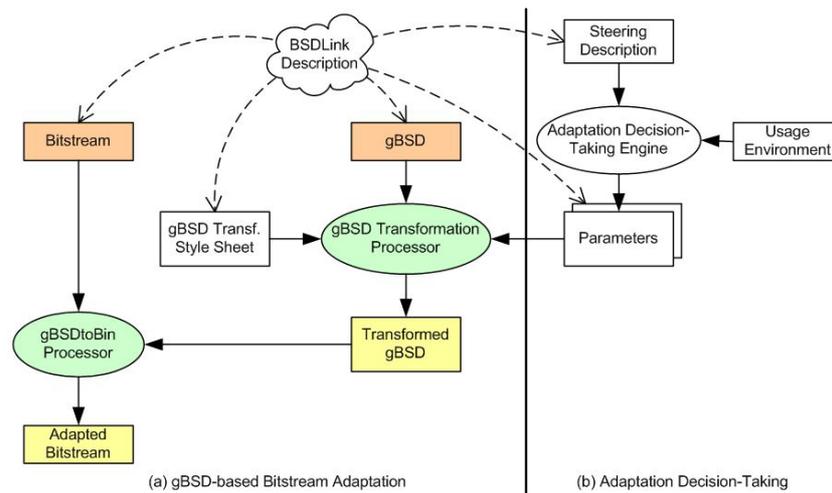


Figure 1: The gBSD-based Bitstream Adaptation Architecture

The major contribution of this project is the development of a generic Bitstream Syntax Schema (gBS Schema), which facilitates generic and coding-format independent BSDs. Consequently, these BSDs are referred to as generic Bitstream Syntax Descriptions (gBSDs) and are applicable to any coding format. The description of a bitstream can also be constructed in a hierarchical fashion that allows grouping of bitstream elements for efficient, hierarchical adaptations. The flexible addressing scheme supports various application requirements and random access into the bitstream. Finally, it enables the development of distributed multi-step adaptations as described in section 4.

6 | MPEG-7 MMDB – A MPEG-7 Multimedia Database

The MPEG-7 Multimedia Database (MPEG-7 MMDB) is an extension of an Oracle database based on their data cartridge technology. It currently consists of four main parts (see Figure 2). The core management system (1) is composed of an MPEG-7 based multimedia database schema, a cost-based query optimizer that approximates the selectivity of range searches and the costs in terms of page accesses of k-nearest neighbour searches, and an indexer that enables for a similarity based retrieval. In addition, a set of internal libraries (2) provides basic functionality such as insertion, deletion and querying of MPEG-7 descriptions.

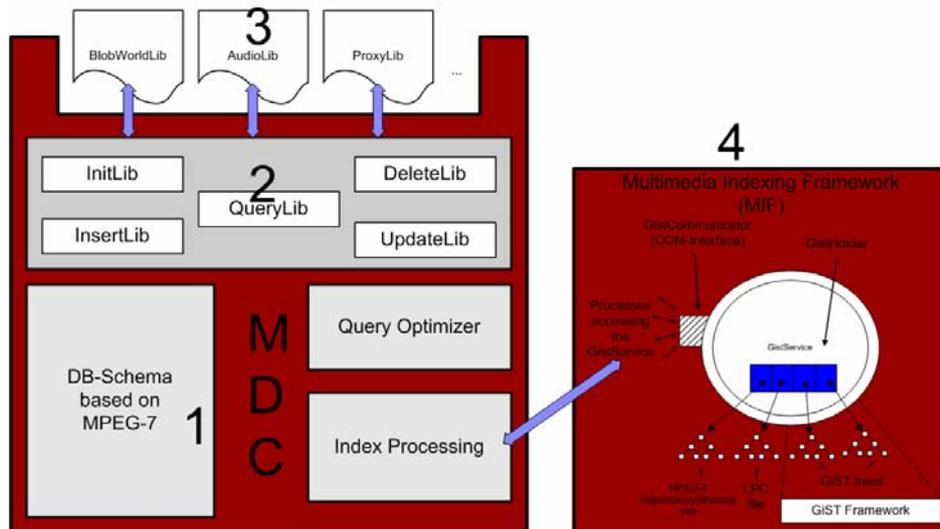


Figure 2: Architecture of the MPEG-7 MMDB

Additional parts of the MPEG-7 MMDB are the Multimedia Indexing Framework (MIF)(4) and several application libraries(3). The indexing system supports fast execution of NN-search, range search and point search. The application libraries serve as interfaces between applications and the MPEG-7 MMDB. At the moment, libraries for the following applications have been implemented: BlobworldLib for a content based image retrieval application, AudioLib for an audio recognition tool and ProxyLib for the QBIX proxy. An application library for video browsing is under development. More information about MPEG-7 MMDB and the mentioned services can be found at the project's home page². Technical information is available in [DK02] and [DK03].

7 | The DAHL Project

The DAHL project aims at demonstrating some of the research achievements at ITEC by extending the Virtual Computer Science Exhibition web application [ITEC03] with content based search mechanisms and an adaptive streaming environment for video data. The search is based on MPEG-7 descriptions of video data and uses the database and interfaces implemented in the CODAC project at ITEC (see MPEG-7 MMDB project description). Video retrieval uses the MPEG-4 conforming adaptive streaming server and proxy implementations developed in the ViTooKi project, which allows to adapt the video stream dynamically to client capabilities, user preferences, and available network bandwidth.

The system to be developed in the DAHL project should provide the following main functionality: segmentation and annotation of video content using MPEG-7 descriptors, searching the MPEG-7 descriptions for meta information about the video content, and retrieving and playing video segments in an adaptive streaming environment. The main components of the desired system and the flow of information

² See URL: <http://medphyt2.uni-klu.ac.at:9090/MDCPortal>

between them is depicted in Figure 3. The video search component is to be integrated into the virtual computer science exhibition.

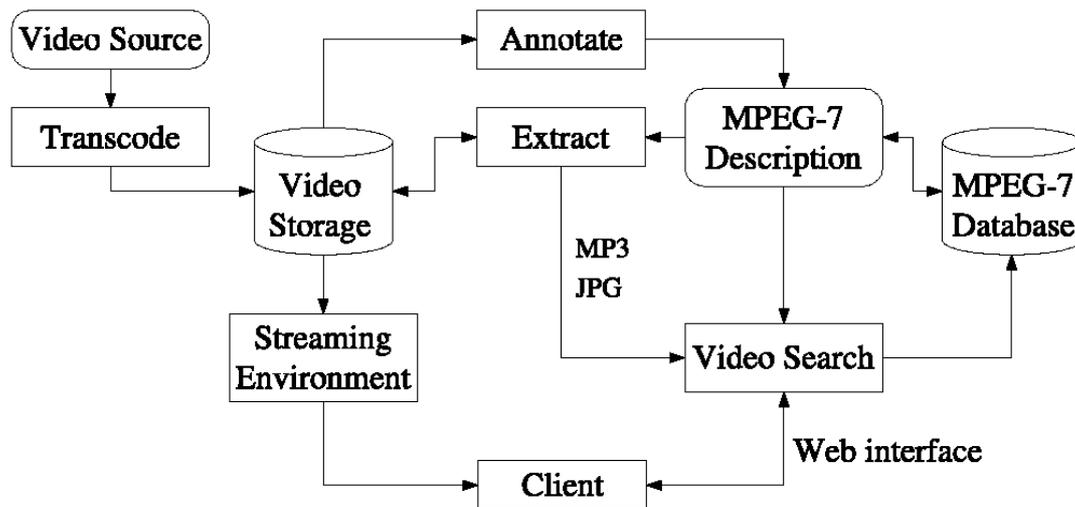


Figure 3: Main Components and Data Flow of the DAHL System

The Extractor component allows for extraction of certain variations of the multimedia content. These include key frames and sound tracks of video segments. They are extracted to separate files and transferred to the web server implementing the video search component.

The user visits the virtual computer science exhibition web site and navigates to a video search form, where he can enter various search criteria for video segment meta information, e.g. title, text annotations describing video scenes, or semantic place or time of video content. Moreover, the user can specify the segmentation level of the desired search results (in analogy to chapters, sections, or subsections of books). The search results are presented as a list of textual meta information about the video segments matching the search criteria. The meta information may include items like segment title, abstract, textual annotations, semantic place and time, and meta information pertaining to the entire video containing the particular segment. And of course, each search result presents a hyperlink to the actual video data.

8 Metadata Supported Content Adaptation in Distributed Multimedia Systems

This project aims to show video variation supported with metadata as a possible approach to adaptation to enable ubiquitous multimedia access. In video variation, the interest is to generate variation videos from a source video with reduced quality and data size by applying reduction or variation methods. Since temporal, spatial, and SNR variations are already well understood, the major contribution of this project is to do investigations in object-based and segment-based video variations.

Object-based variation extracts primary and secondary objects and re-encodes them as one video with two visual objects providing an invariant variation video that can be

used later in the adaptation process to drop objects. Segment-based variation allows the application of reduction methods discriminately on different parts of a video, thereby minimizing quality loss and/or maximizing the gain in data size reduction. The selection of reduction methods to be applied on a segment is based on the physical characteristics of motion, texture, and color of the segment. This requires a proper segmentation of the source video.

A unifying framework of the variation products, called the *Variation Factory*, is developed. It is extended to what is known as the Multi-Step Variation Factory, to allow intermediary videos to be variations in their own right and then serve as sources to further variations. This allows variation videos within the same variation set to have different sources. Hence, the Multi-Step Variation Factory creates a tree of variations. An adaptation architecture using the variation factories, called the Variation Processing Unit (VaPU), is also developed. VaPU is a supplemental server component which is responsible for creating variation videos and extracting meta-information about the variations.

References

- [BDH+02] László Böszörményi, Mario Döller, Hermann Hellwagner, Harald Kosch, Mulugeta Libsie and Peter Schojer: **Comprehensive Treatment of Adaptation in Distributed Multimedia Systems in the ADMITS Project** *In: Proceedings of the 10th ACM Conference on Multimedia, Antibes, France, Nov. to Dec. 2002*
- [BHK+03] László Böszörményi, Hermann Hellwagner, Harald Kosch, Mulugeta Libsie, Stefan Podlipnig: **Metadata Driven Adaptation in the ADMITS Project** *EURASIP Signal Processing: Image Communication, SPECIAL ISSUE ON MULTIMEDIA ADAPTATION, pp 749-766, 2003*
- [DK02] Mario Döller and Harald Kosch. Demonstration of an MPEG-7 Multimedia Data Cartridge. In *10th ACM International Conference on Multimedia*, pages 85–86, Antibes, France, December 2002
- [DK03] Mario Döller and Harald Kosch. An MPEG-7 Multimedia Data Cartridge. In *SPIE Conference on Multimedia Computing and Networking 2003 (MMCN03)*, pages 126 – 137, Santa Clara, CA, January 29-31 2003
- [ITEC03] People Behind Informatics, Virtual Computer Science Exhibition web application, August 2003, URL: <http://cs-exhibitions.uni-klu.ac.at/>
- [KH04] Michael Kropfberger and Hermann Hellwagner, **Evaluation of RTP Immediate Feedback and Retransmission Extensions**, *In: Proceedings of IEEE International Conference on Multimedia and Expo (ICME), June 2004*
- [JLH+04] Dietmar Jannach, Klaus Leopold, Hermann Hellwagner, Christian Timmerer: **A Knowledge Based Approach for Multi-step Media Adaptation** *In: Proceedings of the 5th International Workshop on Image Analysis for Multimedia Interactive Services (WIAMIS), Lisboa, Portugal, April 2004*
- [JLH04] Dietmar Jannach, Klaus Leopold, Hermann Hellwagner: **An Extensible Framework for Knowledge-based Multimedia Adaptation** *To appear in the Proceedings of the 17th International Conference on Industrial & Engineering Applications of Artificial Intelligence & Expert Systems, Ottawa, Canada, May 2004*
- [PHH+03] Gabriel Panis, Andreas Hutter, Jörg Heuer, Hermann Hellwagner, Harald Kosch, Christian Timmerer, Sylvain Devillers and Myriam Amielh: **Bitstream Syntax Description: A Tool for Multimedia Resource Adaptation within MPEG-21** *In: EURASIP Signal Processing: Image Communication, Special Issue on Multimedia Adaptation, pp 721-748, 2003*
- [PB02] Stefan Podlipnig, Laszlo Boeszormentyi: **Replacement Strategies for Quality Based Video Caching** *International Conference on Multimedia and Expo (ICME), Volume 2, 49-53, Lausanne, Switzerland, August 2002*
- [PB03] Stefan Podlipnig, Laszlo Böszörményi: **A Survey of Web Cache Replacement Strategies** *ACM Computing Surveys (CSUR) Volume 35, Issue 4, pp 331 – 373, December 2003*
- [SBH+03] Peter Schojer, Laszlo Böszörményi, Hermann Hellwagner, Bernhard Penz, Stefan Podlipnig: **Architecture of a Quality Based Intelligent Proxy (QBIX) for MPEG-4 Videos** *In ACM WWW2003, Budapest, Hungary, pp 394-402, May 2003*
- [SBH04] Peter Schojer, Laszlo Böszörményi, Herman Hellwagner: **QBIX-G: A Quality Based Intelligent proXY Gateway** *submitted to ACM Multimedia 2004*
- [TSK+03] Roland Tusch, Christian Spielvogel, Markus Kröpfl, László Böszörményi: **An Adaptive Distributed Multimedia Streaming Server in Internet Settings** *In: Proceedings of Information Technologies and Communications (ITCom), Internet Multimedia Management Systems IV, Orlando, USA, pp 312-323, September 2003*
- [TBG+04] Roland Tusch, László Böszörményi, Balázs Goldschmidt, Hermann Hellwagner, Peter Schojer: **Offensive and Defensive Adaptation in Distributed Multimedia Systems** *Computer Science and Information Systems (ComSIS), Vol. 1, No. 1, pp. 49-77, 2004*

Further references can be consulted at our publications home page³.

³ See URL: <http://www.ifi.uni-klu.ac.at/ITEC/Publications>