*XXIV Simpozijum o novim tehnologijama u poštanskom i telekomunikacionom saobraćaju – **PosTel 2006**, Beograd, 12. i 13. decembar 2006.*

PERSONALISED MULTIMEDIA COMMUNICATION SERVICES

Urban Burnik, Matevž Pogačnik
University of Ljubljana, Faculty of Electrical Engineering

Abstract: *An overview of techniques for automated adaptive operation of multimedia telecommunication services respecting individual user requirements is presented in the paper. Future user and media descriptors will be gathered automatically, introducing understanding of service objects and subjects rather than plain recognition of isolated descriptors. Thus, the personal multimedia communication services will improve and become easier to use.*

Keywords: *universal media access, adaptive services, personalised services, MPEG–7, MPEG–21*

1. Introduction

Modern communication services rely on application of personal communication devices that enable communication with other service users, content distribution and information browsing. Up until recent time, telecommunication services could be categorised into 3 main categories, (1) broadcast, (2) phone services and more recently (3) Web browsing. For all mentioned categories, one may consider a telecommunication service as distribution of information over a large distance, where the nature of service differs in terms of "one–to–many", "one–to–one" and "interactive" approach. The "triple–play" idea of sharing the Internet as a common distribution media can be recognised as a first move towards truly integrated media services. The concepts of telecommunication services are to be unified; as an example only, Googling for someone's next tourist destination may directly continue by a conversation to a former visitor of the resort and later be finished by interchange of personal information, photographs etc.

One of the most challenging problems of future is to effectively manage the vast amount of information nowadays available to the service users. Together with the amount of available content grows the number of service users and in modern decentralised services also the number of nodes offering access to the content and information. Current search

mechanisms result in a large number of potential hits among which majority does not satisfy personal requirements of a service user. Besides, many of them are not suitable to be presented on a specific communication device currently in use. Modern services are to be designed universally, regardless on the available hardware or network connection. Universal telecommunication services call for built–in intelligence in sense of automated content and user understanding, with respect to available resources, and thus to become more friendly, natural to the service user.

2. Content markup, classification and retrieval

A general trend in communication services is to provide information anytime and anywhere. The information source is not exclusively owned by media production professionals anymore. Each service user equally acts as a content consumer as well as a content provider – the fact which raises both legal and technical challenges [11, 12, 15]. Enhanced and systematic approach to content authoring, processing and adaptation is required to provide users with adequate information in heterogeneous environments. To avoid information over flooding procedures are needed which control data priorities with respect to specific user requirements, service conditions and available terminal equipment. The framework for these universal services is to exploit mechanisms for content recognition combined with user modelling and advanced user interfacing.

Successful progress depends on

- Standardisation of metadata describing content, equipment and user preferences.
- Automated procedures to provide and to exploit the above mentioned descriptors.

There exist a number of systems for text document markup. Automated search and description procedures are often exploit one or more labels [1, 2], and are based on statistical models, information theory, probability and methods of multiple labelling. Manual labelling is also in use, which is more time consuming however still provides more accurate descriptions and higher level of abstraction.

The remaining multimedia document types are fundamentally different from text documents. Their low–level features like brightness, colour, contrast or texture are not in a string correlation to high–level features of an information search space. In this sense, regardless of the present MPEG–7 [3] standard, metadata structures for image and video description are not widely exploited. Most successful, efficient and reliable systems for multimedia markup and retrieval are still based on additional, textual metadata predominately provided by a human operator (TV Anytime [4] for TV program guides).

Despite the first promising results, the reliable procedures on automated document markup and retrieval are yet to be developed. The research is to be oriented towards automated understanding of data, identification of relations among data primitives and structural notation.

The efficiency of user–initiated queries is directly related to concise use of query expression. Manually generated expressions require a skilled service user. User expectations often include, besides basic keywords, individual user properties, his personal interests, previous knowledge on the topics, purpose of search, current service conditions etc. The additional information may be provided by extensive use of user modelling procedures.

2.1 User modelling

A good model of user behaviour is a key towards friendlier and more manageable telecommunication services. The data required to form a model of a particular user can be obtained in several ways. It is best not to obstruct the user with extensive manual insertion of data as the required information can be obtained interactively during the use of the service. Based on simultaneous analysis of user behaviour the model of user expectation and requirements can be built. The mechanisms of behaviour observation may be part of a standard user interface, and the return channel may be provided:

- Explicitly by using direct evaluation of query results (form–based selection).
- Implicitly by evaluation of query results based on further user actions (delete, select, skip…).

A combination of the above mentioned methods is also an option.

The complexity of user models is varied. Simplest approaches are restricted to a list of selected/unselected items. More structured solutions take advantage of decision trees, hierarchical structures, keyword vectors and other similar approaches. The first standardised approach to user modelling is a part of MPEG–7 user profile [3] and is to be presented throughout this paper.

Procedures for content selection strongly rely on the use of user models. These can be autonomous numerical procedures (like calculating a degree of similarity) or can be a part of user mode structure (like a decision tree). In general, two fundamental approaches for content evaluation and selection are known. The first is autonomous and is referred to as "**Content Based Filtering**" (CBF). In this case, evaluation of content suitability for a particular user is ensured by direct comparison of content medatada and user metadata Alternative approach evaluates similarity of individual user of a system. The content recommendation is then based on the results of evaluation formerly performed by similar users. The approach is known as "**Collaborative Filtering**" (CF).

2.2 Content personalisation

The advances in telecommunication infrastructure like Internet and predominately by the raise of interactive access to the Web initiated strong research on adaptive (personal) hypermedia systems and services. Besides a personalised approach within individual (hypermedia) documents personalised services are currently dominating the area of electronic news (newsgroups, online newspapers) and also e–mail spam fighting [1].

Currently these user model based services are most developed in the field of textual information, however first audiovisual (A/V) systems are presented. A special challenge is the field of “personal TV” [5, 6] where methods like »content–based selection« by using textual metadata descriptors, »collaborative selection« by using grades provided within specific user groups and combined approaches are known.

3 Personalisation and standards

It is clear that user–oriented systems have a future and that they will appear in most of the content retrieval domains. Regardless of the content types the personalisation approaches used today have similar usage scenarios. Users browse content listings, implicitly or explicitly rate content items and get content suggestions. The drawbacks of today's systems are context unaware keyword approaches, relatively primitive means of interaction with systems resulting in poor implicit feedback, unawareness of user's mood etc. In addition to that the systems are ignorant of the device types that are being used to access the content. The introduction of mobile handheld devices requires additional mechanisms that will adapt the content to the type of the device and enable content access from anywhere in the world using the same user 'identities'.

3.1 Universal Multimedia Access

The UMA (Universal Multimedia Access) concept is one of the key concepts that drive the development on the filed of multimedia technology. It represents the notion that any content should be available anytime, anywhere. This may require adaptation of the content, in order to meet the limitations of the user’s terminal or network [10].

However, the multimedia delivery is evolving from simple user content access to the delivery of a “best experience” to a user in a given context. This evolution is encompassed by the UME concept, which represents the notion that a user should have equivalent, informative experience anytime, anywhere. The critical difference between UMA and UME is that the latter clearly acknowledge that the end point of a universal multimedia consumption is the user and not the terminal [10].

There is a good match between the goals of UMA and UME concepts on one side and the newly emerging standard MPEG–21 on the other side. The vision of MPEG–21 is to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities, in a secure environment [11].
MPEG–21 is based on two essential concepts:

- The concept of Digital Item (DI), defined as a fundamental unit of distribution and transaction
- The concept of users interacting with DIs.

The DIs can be considered the “what” and the users can be considered the “who” of the multimedia framework [12].

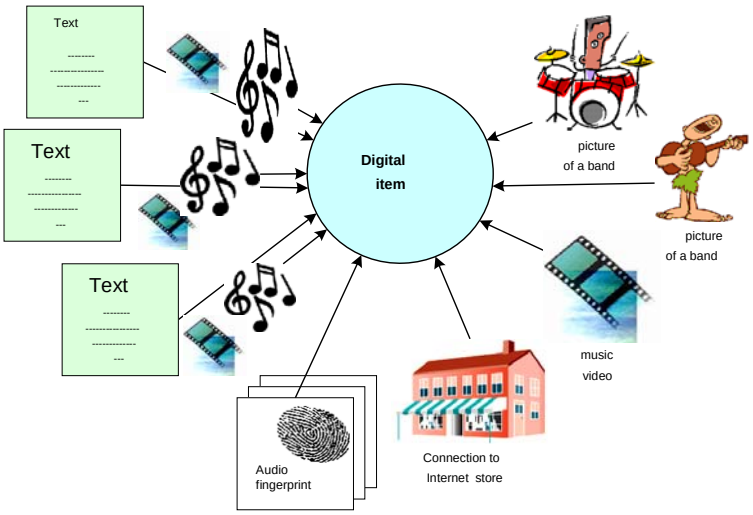

Figure 1: Example of a Digital Item

A DI is a structure digital object with a standard representation, identification, and associated metadata within the MPEG–21 framework. Part 2 (Digital Item Declaration) of the MPEG–21 defines a powerful and flexible model for DIs, that can accommodate the myriad forms content can take. An MPEG–21 User is any entity that interacts within the MPEG–21 environment and makes use of DIs. MPEG–21 makes no distinction between a "content provider" and a "consumer" – both are Users. However, Users are identified specifically by their relationship to another User for a given interaction. The MPEG–21 User concept focuses on two basic fields of interaction with the DIs: the IPMP (Intellectual Property Management & Protection) and the DIA (Digital Item Adaptation) [12].

The DIA part of the MPEG–21 (Part 7 – Digital Item Adaptation) contains the technology for DI adaptation, enabling UMA scenarios for adaptation of media resources at the server, network, and terminal. The DIA addresses requirements that are specific to the usage environment descriptions. MPEG–21 shall support the description of the following: terminal capabilities (acquisition properties, device type, output properties, hardware properties…), network capabilities (delay characteristics, error characteristics, bandwidth characteristics…), delivery capabilities (transport protocols, connections supported…), User characteristics (User preferences, demographic information…), natural environment characteristics (location, illumination and acoustic properties affecting User or terminal…), service capabilities (User roles, types of services…), interactions and relations among Users [11]. Figure 2 presents the basic concepts related to DIA.

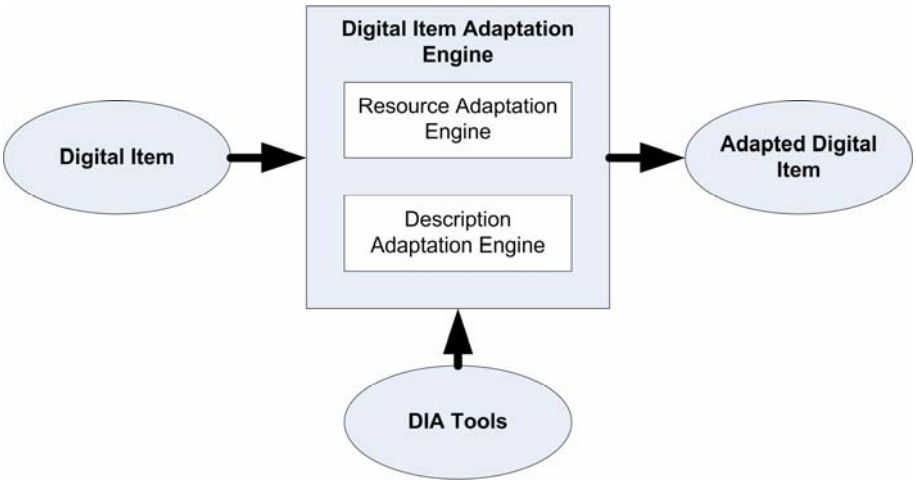

Figure 2: Digital Item Adaptation

Such comprehensive usage environment description allows developers to create a MPEG–21 based multimedia system that can fulfil the visions of the UMA and UME concepts.

3.2 Resource Description Format

The problem of contextual understanding of keywords is mostly present in case of textual documents, especially web pages. Current situation in the web is such that data is generally "hard coded" in HTML files. The concept terms used are semantically ambiguous, so there is no way of telling, which of the possible meanings is the right one. This ambiguity is transferred into the user profile when information about "preferred" keywords is extracted from web pages and stored. In order to resolve this issue, an initiative called the Semantic web has been started [9]. To put it simply, the idea is to describe specific term meanings, their relationships and context information with the help of schemas and ontologies. The descriptions are made using RDF (Resource Description Format). With the development and use of inference logic this approach may become a very powerful tool for information processing and retrieval. Similar situation is with MM content types. A solution to this problem is being offered by the MPEG–7 standard [3, 7]. Namely, the MPEG–7 standard aims to describe all types of MM content (audio and speech, moving video, still pictures, graphics and 3D models) including information how objects are combined in scenes, their semantic meanings and relationships (temporal, spatial, etc.). The standard has powerful mechanisms for content description, but is on the other hand very complex, so widespread use is still questionable.

3.3 User profile

Despite the development in the field of personalized content selection there are surprisingly few standards regarding the descriptions of user preferences. Apart from the widespread use of word vectors there is actually only one standard in this field: the MPEG–7 user profile. The idea was to standardize the description of user preferences and the means of usage of MM content [3].

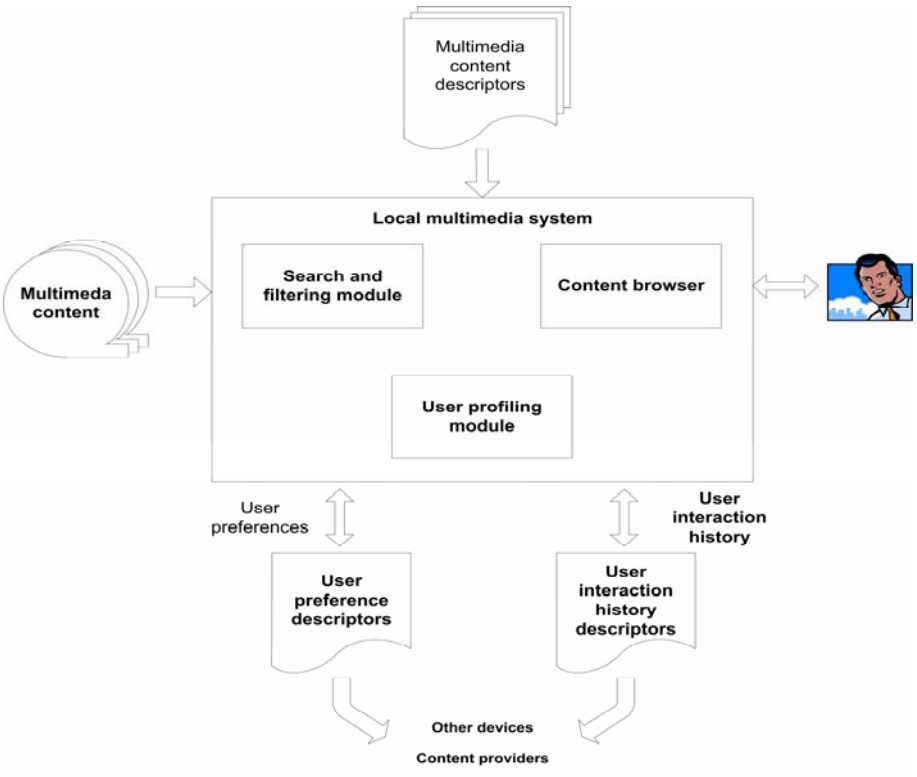

Figure 3: MM system and user interaction [3]

In order to enable the exchange of information about user preferences with 3rd party services they decided to standardize also the information about content usage history. This part of the standard is specifically oriented towards the user interaction with the personal digital video recorder or similar devices. For this purpose two description schemes (DS) were designed: The UsageHistoryDS and the UserPreferencesDS. The UsageHistoryDS enables annotation of user actions (play, record, delete etc.) with respect to particular content item, while the UserPreferencesDS enables annotations of user preferences regarding content creation (favourite titles, actors, directors, locations of content creation etc.), content classification (favourite genres, subjects, languages etc.), source preferences (favourite media formats, dissemination mediums etc.) and some others. The standard does not mandate the algorithms used for mapping of usage history to user preferences. Figure 3 presents typical data and content flow of a personalized retrieval system using user preference description and usage history description.

4. Understanding the users' needs

The standards mentioned above by no means prescribe the actual procedures of fitting the content to a specific user, or, in other words, make no algorithms for user or content recognition recognition. These are only meant to provide the necessary data required to perform these procedures. The intelligent telecommunication services of the future should, by using advanced software tools, make a step from mechanical fulfilment of requests towards understanding of specific user expectations about a specific telecommunication service. A strong framework for next–generation smart multimedia communication systems is to be built. The fulfilment of this main objective tackles with important problems at different levels:

Bridging the semantic gap. There exists a gap between the low–level item descriptors and the real world situations they belong to. The challenge is to use tools based on ontology, data mining and computational intelligence in order to automatically generate high–level semantic descriptors based on low and medium–level descriptors and on the available semantic annotated content. Automated audiovisual data annotation should be using concepts and levels of abstraction similar to human thinking, maybe using some clues from multimedia ontologies.

Using adaptive user modelling. Tools for adaptive user modelling, based on ontologies, using and gathering context aware information about users' preferences or dislikes are foreseen.

The unification of content representation and identification. Content representation and identification are currently being considered independently. However, for most future multimedia services, both problems will require a unified framework, and new tools should be used to define a common structure supporting both the content representation (coding) and its identification (description). This involves the tasks of developing presentation tools using metadata, content–based encoding and transmission strategies, and of a code supporting visualization and search/retrieval functionalities.

Building retrieval mechanisms based on multiple audio–visual features. A combination of audio–visual low–level feature descriptors should be employed in the classification process for the purpose of semantic object and scene identification and automatic summarization of video sequences. The extracted features and the inferred concepts should represent a new challenge for retrieving audio–visual (A/V) content. The additional media descriptors obtained automatically strongly overcome currently known annotations of audiovisual material. Having more data on the material brings even more focus on proper search mechanisms; a balanced combination of features and concepts is to be used to provide an optimal answer to a user's query.

Building context aware retrieval mechanisms. Most search engines provide queries by keywords without being able to express their contextual meaning and position; the use of semantic structures such as content metadata and ontologies can improve the retrieval

process. New tools and approaches for content retrieval are required based on structured or unstructured high–level descriptions using topic–oriented ontologies.

Integrating the activities of the communication chain. Content understanding tasks should be incorporated in the integral communication chain. We therefore require supporting techniques for content, presentation and distribution management, supported by fundamental signal processing algorithms running preferably in real–time.

5. Conclusion

Personalised content retrieval is becoming an important mechanism of satisfying customer requirements using multimedia communication services. In the near future one may expect significant improvements in the domain of user–to–system interaction, based on contextual understanding of content and concepts, leading towards a unique personal experience on any communication device. For wide implementation of such services standard frameworks are to be exploited to the widest extent, supported by automated content indexing, which will include high level object recognition and semantic understanding of multimedia content. The experience of Digital Signal, Image and Video Processing Laboratory in the presented domain is rich and has been acquired on many European research projects.

References

[1] Gudivada N. et al. "Information Retrieval on the World wide web", IEEE Internet Computing, Vol. 1, No. 5, 1997

[2] Hand D.J. et al. "Principles of data mining", The MIT Press, Cambridge, Massachusetts, USA, 2001

[3] Manjunath B.S. et al. "Introduction to MPEG–7, Multimedia description interface", John Wiley & sons Ltd., England, 2002

[4] "TVAnytime: Specification series: S–3 metadata, SP0003 v1.3, part A", 2002, ftp:// ftp.bbc.co.uk/pub/Specifications/SP003v13.zip

[5] O'Sullivan D. et al.: "Improving the quality of the personalized electronic program guide", User Modelling and user adapted interaction, Vol.14, 2004

[6] Cotter P., Smyth B. "PTV: Intelligent personalised TV guides", Proceedings of the 12th Innovative Applications of Artificial Intelligence (IAAI–2000) Conference. AAAI Press, 2000

[7] Martínez J.M. (ed.) "MPEG–7 overview", ISO/IEC JTC1/SC29/WG11, CODING OF MOVING PICTURES AND AUDIO, Pattaya, March 2003, http://www.chiariglione.org/mpeg/standards/mpeg-7/mpeg-7.htm

[8] TiVo home page, http://www.tivo.com, 2003

[9] Palmer S., "The semantic web: An Introduction", 2001, http://infomesh.net/2001/swintro

[10] Pereira, F., Burnett, I., Universal multimedia experiences for tomorrow, IEEE Signal Processing Magazine, volume: 20 issue: 2, March 2003, pp. 63 – 73

[11] Bormans J., Gelissen J., Perkis A., "MPEG–21: The 21st Century Multimedia Framework", IEEE Signal Processing Magazine, volume: 20, issue: 2, pp. 53 – 62, March 2003
[12] Bormans J., Hill K., "MPEG–21 Overview", http://www.chiariglione.org/mpeg/standards/mpeg-21/mpeg-21.htm
[13] Burnik, U., Optimalno kodiranje za prenos slik po ozkopasovnih prenosnih medijih, PhD Dissertation. University of Ljubljana, Ljubljana, 2002
[14] Vetro, A., Christopoulus, C., Ebrahimi, T., Universal Multimedia Access, IEEE Signal Processing Magazine, volume: 20, issue: 2, March 2003, pp. 16
[15] Dobravec, Š., Ogrodje novih multimedijskih tehnologij – MPEG–21, MSc Thesis, University of Ljubljana, Ljubljana, July 2003
[16] Vetro, A., Christopoulus, C., Sun, H., Video Transcoding Architectures and Techniques: An Overview, IEEE Signal Processing Magazine, volume 20 issue: 2, March 2003, pp. 18 – 29
[17] The MPEG home page, http://mpeg.telecomitalialab.com/
[18] Koenen, R., From MPEG–1 to MPEG–21: creating an interoperable multimedia infrastructure, http://mpeg.telecomitalialab.com/documents/from_mpeg-1_to_mpeg-21.htm
[19] Koenen, R., editor, MPEG–4 Overview, http://mpeg.telecomitalialab.com/standards/mpeg-4/mpeg-4.htm
[20] van Beek, P., Smith, J.R., Ebrahimi, T., Suzuki, T., Askelof, J., Meta-data-driven multimedia access, IEEE Signal Processing Magazine, volume: 20 issue: 2, March 2003, pp. 40 – 52
[21] Ferman, A.M., Errico, J.H., van Beek, P., Sezan, M.I., Content-based filtering and personalization using structured metadata, Proceedings, 2nd ACM/IEEE-CS Joint Conf. Digital Libraries, Portland, July 2002, pp. 393
[22] Good, N., Schafer, J.B., Konstan, J.A., Borchers, A., Sarwar, B., Herlocker, J., Riedl, J., Combining collaborative filtering with personal agents for better recommendations, Proceedings, Conf. Am. Assoc. Artificial Intelligence, Orlando, Florida, July 1999, pp. 439 – 446
[23] Wang, Y., Liu, Z., Huang, J.C., Multimedia content analysis, IEEE Signal Processing Magazine, volume: 17, November 2000, pp. 12 – 36
[24] Dimitrova, N., Zhang, H., Shahraray, B., Sezan, I., Huang, T., Zakhor, A., Applications of video content analysis and retrieval, IEEE Multimedia, volume: 9, July–September 2002, pp. 42 – 55

Sadržaj: *Ovaj rad predstavlja pregled metoda za automatsko prilagođenje ponašanja multimedijalnih telekomunikacionih usluga, imajući u vidu individualne potrebe pojedinih korisnika. Planira se da će se u budućnosti multimedijalni sadržaji i modeli korisnika generisati automatski, razumevanjem objekata i subjekata usluge, umesto da se koriste izolovane ključne reči. Osim toga, očekuju se bolje personalne multimedijalne telekomunikacione usluge kojima će moći korisnici lakše da se služe.*

Ključne reči: *univerzalan pristup do medija, prilagodljive usluge, personalizacija usluga, MPEG–7, MPEG–21*

PERSONALNE MULTIMEDIJALNE TELEKOMUNIKACIONE USLUGE

Urban Burnik