Active Multimedia Documents for Mobile Services

Omid E. Kia
National Institute of Standards and Technology
Information Technology Laboratory
Gaithersburg, MD 20899
E-mail: omid.kia@nist.gov

Jaakko J. Sauvola
Media Processing Team
Machine Vision and Media Processing Group
Infotech Oulu, University of Oulu
90570 Oulu, Finland
E-mail: jjs@ee.oulu.fi

Abstract - Multimedia content is designed for many purposes and the services that are associated with them rely heavily on technologies that address representation, processing, and transmission of that content. This paper presents a new model that refines the multimedia documents for mobile environments while integrating information technology components to address strict requirements. We present new ways to structure and represent the multimedia content to fit into the traditional document framework so that we can serve the document for an alternative presentation ordering, cross-media translation, and scalable renditions. We propose techniques that utilize the structure of documents across several levels in a manner that applies semantic description to provide scalable multimedia services. A modern multimedia document consumption scenario is introduced where a mobile collaborative working environment is demonstrated utilizing the proposed techniques.

INTRODUCTION

Technological advancements in networking and wireless technology have allowed service providers access to a market that benefits from a large array of informational services. Traditionally, the information service providers could only transmit ASCII character sets to remote terminals, such as beepers or more recently short messages in cellular phones (SMS) for example. The advent of digital cellular telephones audio and data content (low quality audio) has turned the wireless technology into a feasible medium for provision of multi- and cross-media content. Protocol standards exist for specification of wireless transmissions [1] and multimedia communications [2]. Introduction of services for multimedia documents over wireless terminals requires special techniques and (document related) intelligence of the multimedia signal processing in an integrated, layered, and content-based object-on-demand framework. Currently such services over the wireless mobile networks do not exist. A framework that can successfully integrate such demands requires an understanding of information content (in single and cross media content) and presentation characteristics of media components.

In this paper we propose a new active model that uses techniques which extend the hyperdocuments with multimedia content [2] to a new functional domain. The added functionality is in terms of a client-initiated task that would be performed over wireless networks and on mobile terminals. Document knowledge is used to extend the uses of the conventional documents to that of active documents with desired presentation ordering, scalable rendition, and transmission qualities. Proposed techniques contribute to application areas such as remote collaborative working for mobile professionals, command and control capability in disaster management, remote system diagnostics, and document retrieval solutions among a growing field of applications. In this paper we present one of these scenarios supported by the active multimedia document model.

EVOLVING DOCUMENTS

Traditionally, documents are representations of abstract content on a paper-based medium. Books, magazines, newspapers, and related publications form the major part of the paper-based documents. The organization and presentation of such material has evolved to a point that most people can readily determine their reading content by quick scan and layout of the page. Technologies such as character recognition, duplicate detection, and document classification have contributed toward an easier access to document content. This is possible by a number of factors including use of images and icons, variation in font size and style, and content layout. Distribution of multimedia content through the internet has extended this concept to a new level where authors have devised their own style with varying degrees of success. The ease of use has mostly been driven by the functionality of servers and viewers.

We consider documents to be content repositories. In essence, we can formulate a database core that can organize and represent documents efficiently and a client-server paradigm to service user-initiated tasks. The use of multimedia content extends this philosophy to address a wide variety of content provision services. The driving factor for extending document concepts to the multimedia domain from the traditional binary ASCII content is to extrapolate the well understood document layout, organization, and representation. In document domain, the concepts of information content, progressive transmission, and compression is well known [3]. In almost all cases, a good measure of the usable information yields a suitable representation that expedites processing and handling. Similar concepts need to be addressed in the case of multimedia documents.

The basis of a multimedia document is the ability to scale quality and information content in a uniform fashion. While scalable representations exist for almost all representations, a unified representation does not exist. Much like concepts of metadata and object oriented processing, a meta-link addresses the link and terminal capabilities and scales the information appropriately. These can be done either on a per-data-format basis or can be inherent to the gateway. For these tasks we consider scalable renditions, progressive scenarios, and presentation methodologies. In today's browser technologies with custom program and add-on capabilities, this can be well developed and well maintained. Another fundamental functionality is to pro-

vide a two-way communication whereby a user cannot only retrieve information but also submit information for archiving. The reverse scaling of information is fundamental to this and has not been considered yet.

WIRELESS MULTIMEDIA ADAPTATION

Mobile computing is emerging as a mainstream in many forms such as modern offices, travelling professional suites, and remote work capabilities [4]. Such concepts have considered advances in wireless data transmission to benefit them [5]. Utilization of multimedia content in these cases demand a special approach for multimedia adaptation layers and techniques. Requirements for adaptation of end-to-end systems are set by the transmission and terminal properties where a large variety exists e.g. GSM/PCS 9.6Kb/s, and emerging GPRS 170Kb/s, GSM+ 384Kb/s, and third generation wireless data networks (UMTS) 2Mb/s. There exists a wide variety of capabilities when considering cases of wireless IP to IP over ATM. In all these cases, multimedia content is to reach the end user, and quality of rendition and service should be based on the user's link and terminal capabilities. In our system, we approach this problem from a media understanding approach by defining document content and functionality and making it an interactive document.

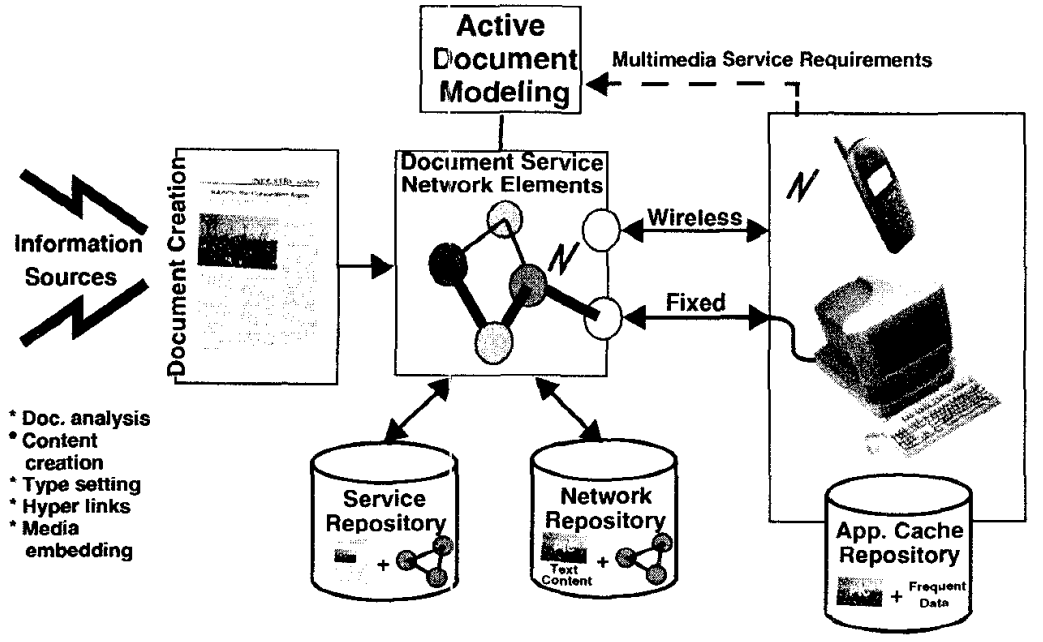


Figure 1. Principle of active document creation and consumption in mobile/fixed domain.

Active documents form the basis for the interactive session keeping in mind that robust content description, semantics, transmission properties, and functionality within document parts will play a large role for mobile environments. Active documents offer an entirely new way to adapt information content into different services where mobile wireless terminals would benefit the most. For documents rich in structural and semantic components, a hyperlinking technique can be utilized along with content-based activity. Presentation over low-bandwidth wireless connection

can be optimized towards the information content given constraints (freedom in this case) on presentation capabilities.

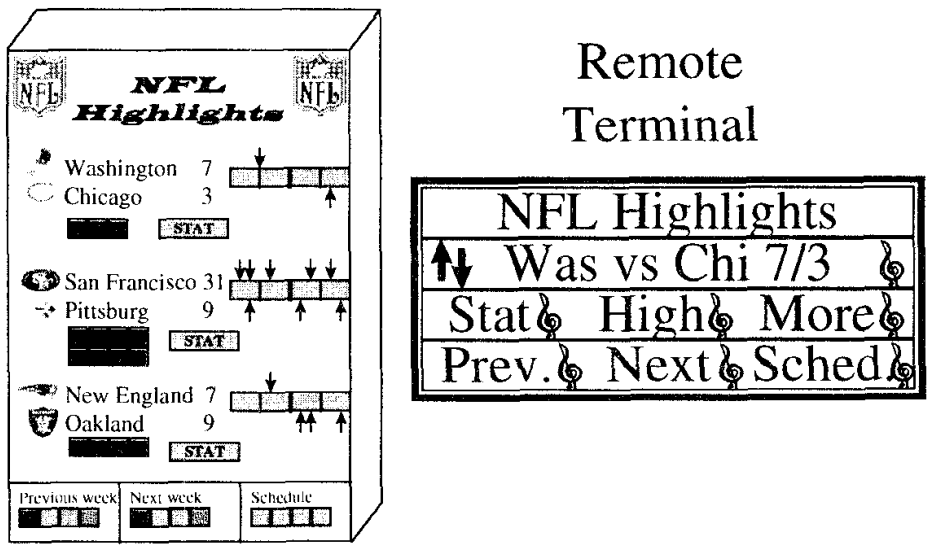
Current development in transmission and service standards, such as packet-based H.323 (media telephony conferencing), can directly be exploited in an active document model. The 'call setup' and 'control' properties are designed for multimedia conferencing over the networks and terminals where the quality of service cannot be quaranteed. Different parts of the active document, meant for the service at hand, can be divided into media and data channels whose implementation in the terminal for presentation is supported by the system contol RAS (call setup), Q.931(channel and transmission setup and control), and H.265 (media conferencing control). While the H.323 is packet-based, the transmission channel reservation and usage can be efficiently performed in service for small wireless terminals. In switching events, when the media stream is directed to other terminal (such as multimedia PC), the control deals with transmission re-routing and scalability by utilizing the 'fast call control' routines. When the channel specifications are accepted, the bandwidth is gained from the gatekeeper, and can be used through the entire presentation. The media understanding-based channel optimization in active documents is one of the most efficient ways to optimize the presentation besides the conventional standards based compression. Since the multimedia document is in a highly structured state as an active document, the parameter reduction, network and terminal caching, content-based compression and transmission, and background transmission operations can be used effectively. At the terminal, presentation objects available as either a local applications or in repositories, augment the received data to increase or refine the information quality that is presented to the user.

EXAMPLE SERVICE SCENARIO

A simple scenario is to access a multimedia database over a slow wireless channel to a terminal capable of playing audio and very few text components. A content provider may organize various data types such as text, images, video, and audio to be available over a wireless network. In turn a user equipped with a handheld phone may access this content, and based on the layout of the content, receive appropriate information. Figure 2 shows an example of such a content used for reporting football scores. In essence the example shows a sample content that has text, image, and video content. For cellular phones, the content is too big and complex to be rendered usefully. An algorithm needs to reorder the content to be presentable on a low complex terminal along with reducing quality to achieve desired rates. Such a translation is shown in Figure 2 where the arrows on the second line denote scroll capability in presentation and the music scale bar denotes availability of audio content as it relates to the textual message.

We intend to develop this concept for use in disaster management scenarios. The organization and effective deployment of emergency personnel during a disaster will determine the effectiveness of relief delivery. In this case a multimedia database is maintained with annotated content that has been gathered by on-sight personnel. These annotations differ in content by terminal and link capabilities and the scaling

algorithm. The underlying document is a local map with layers of rendition quality. Icons identify annotated content and are available for decoding at the terminal based on the rendition capability.



ɜure 2. Active document model utilization example in mobile multimedia

Figure 3 shows an example of such an active document. The underlying document is an area map with labeled major roads. Legends on the right show emergency types, and while we use colors to denote type of emergencies, in real application small icons would be used which denotes the available data types for the emergency situation. The emergency denoted by light blue color on the top right denotes a medical emergency and video, audio, and text is included in the report. The fire emergency in the lower left area of the previous incident carries only image data. The original annotators of this document carried digitization equipment such as keyboard, sound recorder, video recorder, or image capture device. Each device carries the map information which is then used to anchor the emergency event along with the annotation material. For downloading this information, cross-media translation occurs at the data repository and is transmitted down the network path. Note that it is much easier to translate a video signal to a single image than to translate an image to a textual message. However, there exists ways to represent scalable renditions of all data components and it is our intention to create a uniform scalable methodology to conserve as much information as possible.

CONCLUSIONS

Multimedia documents are continually emerging for services provided in wireless environments. This paper discusses a new way to organize document information gained from content understanding in multimedia domain. Using an active document model the contents can be rendered in a new way for services in wireless environment. The proposed model offers a seamless content scaling for different terminal types that is mainly based on value adding information gained from docu-

ment structure and content analysis and refinement information mapped to multimedia services. An example using these techniques is presented that demonstrates the overall efficiency of the proposed model.

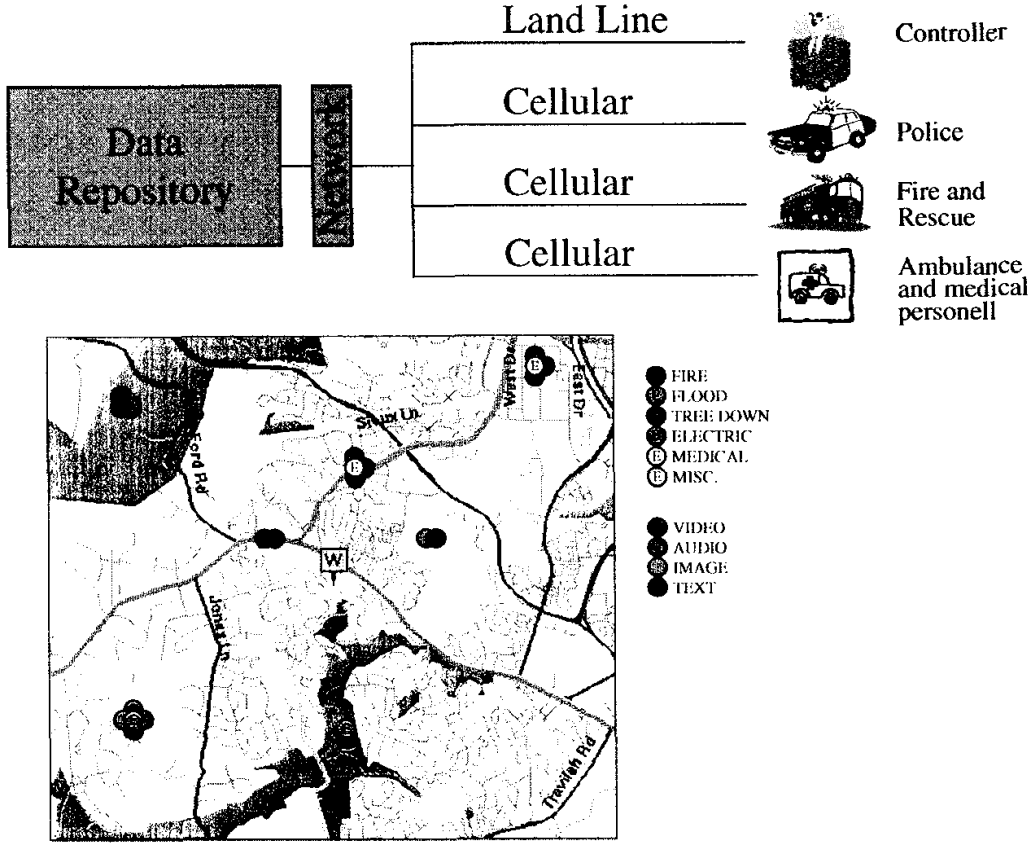


Figure 3. Active document model utilization example in disaster management.

ACKNOWLEDGMENTS

Collaboration with the Language and Media Processing Laboratory of the University of Maryland is appreciated and acknowledged.

References

[1] WAP Forum (1998) WAP standard specification, version 1.0, http:// www.xwap.com/

[2] Wu C. & Irwin J. (1998) Emerging Multimedia Communication Technologies, Prentice Hall, 443 pages.

[3] Kia O. (1997) Document Image Compression and Analysis. PhD Thesis, University of Maryland at College Park, College Park, MD, 1997.

[4] Bates R. & Gregory D. (1998) Voice and Data Communications Handbook, McGraw-Hill, 902 pages.

[5] Blair G. & Stefani J. (1998) Open Distributed Processing and Multimedia, Addison-Wesley, 452 pages.