



Algorithm for Content Adaptation of Multimedia Information

Grigor Mihailov, Teodor Iliev, Elena Ivanova
Department of Telecommunication
University of Ruse
8 Studentska Str.
BULGARIA

Abstract: The Universal Multimedia Access (UMA) addresses the delivery of multimedia resources under different and varying network conditions, diverse terminal equipment capabilities, specific user or author preferences and needs and usage environment conditions. UMA refers to the truly ubiquitous access to and consumption of multimedia content, aiming at guaranteeing unrestricted access to multimedia content from any device, through any network independently of the original content format with guarantees and efficiently and satisfying user preferences and usage environment conditions.

Keywords: multimedia; network; UMA; adaptation; transformation.

I. Introduction

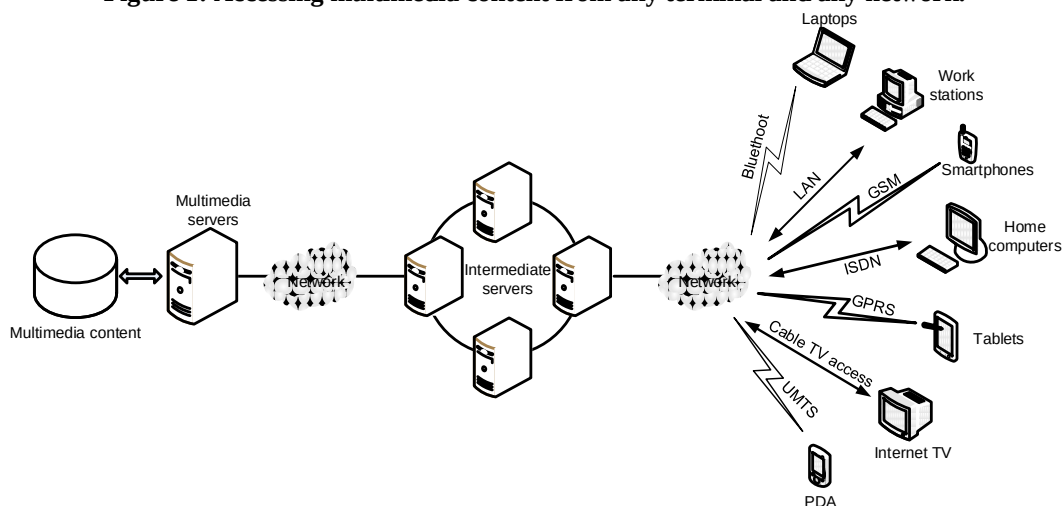
The variety of multimedia information is enormous nowadays. Everyone has a camera, a scanner or another device that almost instantly generates multimedia content. The most of the content author wishes to share its masterpiece. But the variety of terminals and networks may be a problem if he wants everyone to see his work with the best possible quality. This problem will increasingly grow as industry releases more enabling technologies. Late advances in Telecommunications are creating a set of technologies that will enable the development of an all-new set of services with completely different requirements, terminals, capabilities and networks through which terminals will access services in very different conditions.

With the advances in Telecommunications, new protocols and networks are being standardized, providing larger bandwidth and better access (e.g., UMTS, broadband networks, LTE, etc.). The different characteristics of these networks may impose restrictions regarding the content delivery process, e.g. related to bandwidth, error protection, quality of service, etc. Networks can be very different in several aspects. [1]

The recent advances in storage and acquisition technologies are driving the creation and dissemination of large amounts of rich multimedia content. As a result of these advances, there is a growing mismatch between the rich content that is available and the capabilities of the terminals and networks available to the users to access and process the content available. Today, when an author creates some multimedia content, he may have to create several variations of the same content, even with different layouts, different image resolutions and even different media types in order to reach the largest possible number of terminals. [2]

This scenario describes today's situation of accessing multimedia content from any terminal (Fig. 1). When a terminal accesses content to which it was not designed for, the user experience is not so good.

Figure 1: Accessing multimedia content from any terminal and any network.



The user may experience long delays if the content is adapted ignoring the characteristics of the involved parts making its visual presentation less impacting, etc. The scenario described misses a bridging element between all the components involved, which should take into account their characteristics and assure an efficient and consistent inter-working. An efficient way to access any information from any terminal, allowing the delivery of any content (or an adaptation of it) to the user should be provided.

Terminals can be very different in terms of available memory, display colors, etc.; network access can also be very different in terms of bandwidth and so on; each piece of content may be designed with different bandwidth, memory, visualization, rendering, streaming, etc. requirements.

The access to multimedia information by any terminal through any network is a new concept referred by the scientific community as Universal Multimedia Access. The objective of UMA technology is to make available different presentations of the same information, more or less complex e.g. in terms of media types, suiting different terminals, networks and user preferences.

II. Universal Multimedia Access

The purpose of UMA is to enable access to any multimedia content over any type of network with any device, anytime and anywhere. The initial motivation of UMA was to enable terminals with limited communication processing, storage and display capabilities to access rich multimedia contents. Some definitions of UMA are described in the following paragraph: [3]

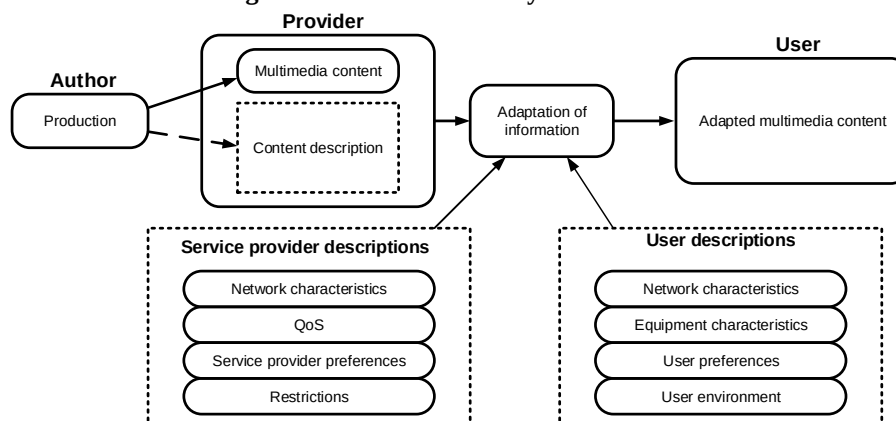
- The concept of UMA is to enable access to any multimedia content over any type of network, such as Internet, Wireless LAN or others, from any type of terminals with varying capabilities such as mobile phones, tablets, PDA, personal computers, and television sets;
- UMA refers to the framework where information is accessed in a suitable form and modality under the current complex and dynamic usage environment such as devices, networks, terminals, preferences, personalization, and other factors of usage environment;
- The primary function of UMA services is to provide the best QoS or QoE by selecting appropriate content formats, or adapting the content format directly, to meet the playback environment, or to adapt the content playback environment to accommodate the content;
- Universal Multimedia Access: Any content should be available anytime, anywhere, even if after adaptation. This may require that content be transcoded form, for example, one bit rate or format to another or transcoded across modalities; e.g., text to speech. UMA concentrates on altering the content to meet the limitations of a user's terminal or network.

The system of UMA allows any customer to reach any multimedia information. The main actors in UMA systems are as follows: [4]

- **Author:** The people or organization who creates the multimedia contents;
- **Provider:** The people or organization who provides multimedia content delivery service;
- **Consumer:** The people or organization who consumes the multimedia contents.

Professional broadcast programs could involve several Authors and Providers, while personal home video could be created and delivered by the same person. The main objective in UMA systems is to maximize the final result satisfaction of all these actors. To achieve UMA systems, multimedia contents must be adapted to the Consumer especially in terms of terminal and network conditions and user characteristics. For easier adaptation of the desired content to the Consumer, it is preferable to have descriptions to fill the gap between media format and terminal, network, user characteristics.

Figure 2: Overview of the system of UMA.

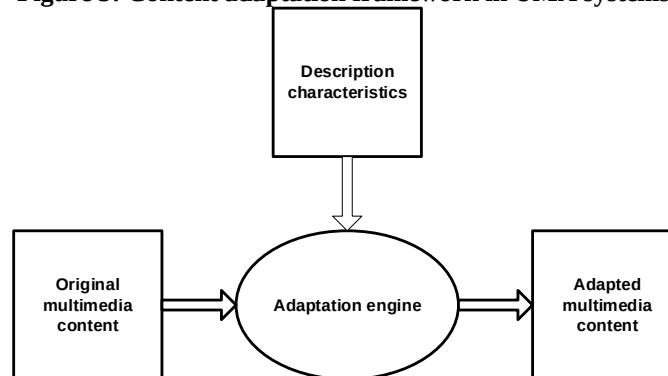


III. Content Adaptation in UMA Systems

Different types of contents in various formats are delivered by many service providers via servers and networks to various terminals of various users. Therefore, a large number of parameters of the contents, service providers and usage environments need to be taken into consideration for adaptation.

Figure 3 illustrates a diagram of content adaptation framework in UMA systems. The multimedia contents and all the descriptions are input into the adaptation engine, and the adaptation engine adapts the contents by selection and transformation based on all descriptions to form a multimedia content adapted to the consumer.

Figure 3: Content adaptation framework in UMA systems.



A. Adaptation engines in UMA systems

There are numerous ways to categorize adaptation type, for example, if: the adaptation should be processed in real-time or not; the target application is push or pull; the adaptation are performed automatically or manually; the adaptation process requires a lot of computation or not and the adaptation is in signal level, perception level or semantic level, and so on.

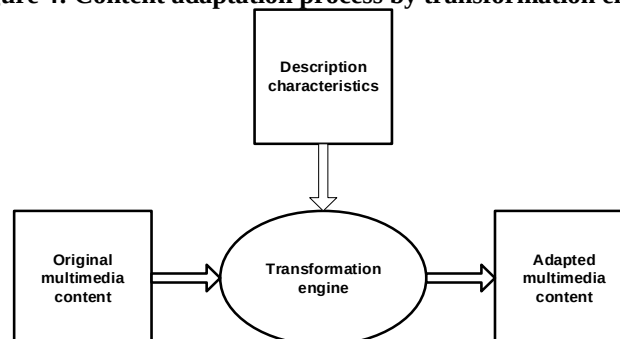
The focus is on what type of multimedia contents the consumer accesses. The first point is if the data to access consists of just one source or many sources. The second point is if there are variations of each content created from a single source or not. Several alternative versions or options derived from a single source may exist (e.g. same program with different frame rate, resolution, number of colors, languages, etc.). [5]

B. Transformation engine

Figure 4 illustrates the diagram of the content adaptation process by transformation engine. The transformation engine transforms a single multimedia content into an adapted multimedia content on-the-fly. The descriptions support the adaptation process to provide the best experience to the user and to reduce calculation cost enough to achieve a real-time adaptation. One relevant application could be broadcast video content delivery to mobile users. [6]

The advantages of content adaptation by transformation are that the storage cost is low because only a single content variation needs to be stored and that an accurate adaptation to every device can be performed considering the device capability in real-time. But the drawbacks are that the engine could have only limited operations that enable real-time or low delay processing, unless it would cause long delay or an engine with quite a lot of computational power would be necessary.

Figure 4: Content adaptation process by transformation engine.



Some content transformation examples are described in the following:

- Adjustment of network capabilities (e.g. bandwidth, delay, error rate) and terminal capabilities (e.g. screen size, terminal power, memory, CPU, decoder) between the service provider and consumer;
- Adjustment of content presentation (visualization) based on user preferences (e.g. preferred mode e.g. small image high frame-rate or large image low frame-rate, all content or summarized content), difficulties in vision or hearing and natural environments (e.g. location, time, weather, color temperature adjustment);
- Content summarization (e.g. 1 hour news in 5 minutes) and highlight (e.g. goal scene in a football match).

Convert the spatial resolution and color depth of the image/video e.g. VGA to QVGA, NTSC to PAL, 24bit to 8bit color depth, color to grayscale, etc. Convert the temporal resolution of the video e.g. 30fps to 24 bps (NTSC

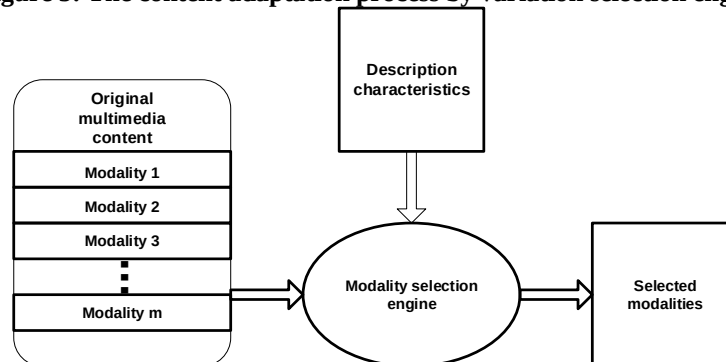
to PAL), 30 fps to 10 fps, etc. Coding parameter conversion frame-rate (e.g. 30 fps to 24 fps). The purpose is the best resolution to the client. [6, 7]

C. Variation selection engine

The variation selection engine selects the best variation from all variations delivered from a single multimedia content. The descriptions support the selection process to provide the best experience to the user. The most typical example of content adaptation by variation selection would be a web site preparing several variations for each type of terminal (PC, PDA, mobile phone) and delivering automatically the most appropriate one by analyzing environmental variables (bandwidth, delay, etc.) Figure 5 illustrates the diagram of the content adaptation process by variation selection engine: [8]

The advantages of content adaptation by selection are that the adaptation process is very quick because the system just has to select the best variation and deliver it to the consumer. At the same time, the system needs just a small amount of computational resources for adaptation.

Figure 5: The content adaptation process by variation selection engine.



On the other hand, the main drawback is that the engine could have only limited variations and sometimes there would be no content that fits the user condition. It is possible to solve this problem by applying a transformation engine to the selected variation in order to increase the consumer's experience. Other problems are that the storage cost and creation cost becomes higher. Furthermore, the management cost is high. If there are some changes on the original content, all of the variations also need to be changed, which requires quite a lot of time and power. [9]

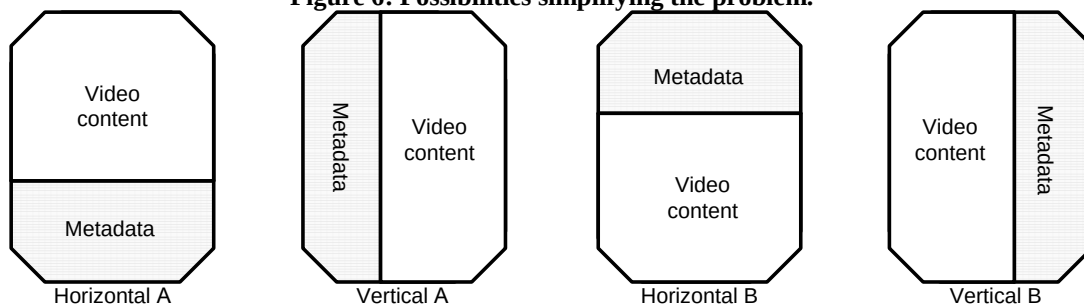
Variation selection examples are basically the same as those of transformation engine. The main difference is that the variations are created beforehand. What to select includes best spatial resolution, temporal resolution, modality, content length, coding format, coding parameters, presentation, content length, important spatial region, temporal region, color selection for the user, and so on.

IV. Layout Control for Balancing Video and Metadata

There are two aspects for determining the layout of how to present the video content and its metadata; the pane layout (horizontal, vertical, etc...) and the size of each pane (video pane, metadata pane). [10]

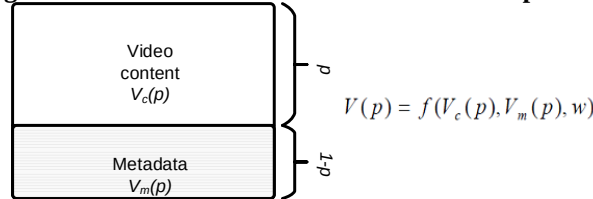
There are many possibilities in how to present the video and the metadata in a fixed screen size. For example, locate the video content and fill all the rest of the screen with text metadata, fill thumbnails in redundant spaces, overlay some text over the video content to maximize the video content size, determine the area for presenting video and metadata not only in rectangular form but also in any form like triangular, circular, and so on. There are four possibilities shown in Figure 6 to simplify the problem; the video, the metadata to be located up, down, left, or right. The screen is simply divided into two parts, horizontally or vertically, and one part is assigned for video and the other for metadata. [11, 12] To optimize the content adaptation process, a quality metric to evaluate how well the content is adapted meeting the constraints on both the provider side (e.g. restrictions on scaling rate) and the consumer side (e.g. device capabilities, user preferences) is emerging.

Figure 6: Possibilities simplifying the problem.



We define a value of the total adapted content to evaluate the fidelity of the overall adapted content compared with the original content. Thus, the optimization process can be described as determining the way of adaptation that maximizes the total content value meeting the given constraints. For expressing the balance of the size of each pane, we define a variable p as a normalized value of the video content pane size (Fig. 7). This value is normalized to be between 0 (whole screen is for metadata) and 1 (whole screen is for video). The balance of the video content pane and metadata pane can be expressed as $p: 1-p$ in both horizontal and vertical layout case. [13]

Figure 7: Normalized value of the video content pane size.



Once the pane size to present the video and that of the metadata are determined, the total content value can be calculated by optimizing the content value of each modality in each given pane size. This means that the value of the total content V can be defined as a function of p . There are three parameters that affect the total content value. The video content value, the metadata value, and the preferred balance of them are the three parameters. The total content value function $V(p)$ is modelled as follows:

$$V(p) = f(V_c(p), V_m(p), w) \quad (0 < p \leq 1), \quad (1)$$

where $V_c(p)$ - value of the adapted video content; $V_m(p)$ - value of the adapted metadata; w - importance of the video content (normalized). [14]

The total content function $V(p)$ evaluates the balance between the adapted video content and the adapted metadata and also takes into account the subjective importance of each modality for the consumer. Considering the video content value and metadata value, as both the video content pane size and the metadata pane size can be calculated from the screen size and the given p , the value of both modalities can also be expressed as a function of p .

The total content value can be determined by calculating the maximum total content value for both horizontal and vertical layout and selecting the one with higher content value. The total content value $V(p)$ becomes:

$$V(p) = \max\{V_H(p), V_V(p)\}, \quad (2)$$

where $V_H(p)$ - total content value in horizontal layout; $V_V(p)$ - total content value in vertical layout.

Fig. 8 illustrates the video content pane size and metadata pane size for both horizontal and vertical layout. By applying the equation (2), the total content value for both layout cases, $V_H(p)$ and $V_V(p)$, can be modeled as: [15, 11]

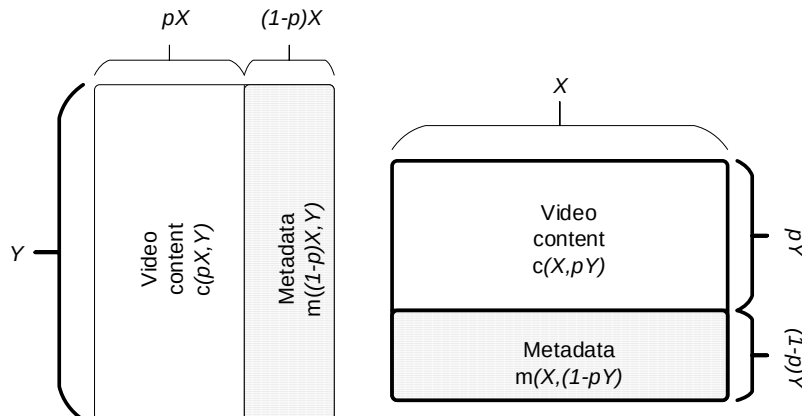
$$\begin{aligned} V_H(p) &= f(c(pX, Y), m((1-p)X, Y), w) \\ V_V(p) &= f(c(X, pY), m(X, (1-p)Y), w) \end{aligned} \quad (3)$$

where (X, Y) - screen size; $c(x, y)$ - maximum video content value in pane size (x, y) and $m(x, y)$ - maximum metadata value in pane size (x, y) .

The value of $V_c(p)$ and $V_m(p)$ can be determined using the following conditions:

$$\begin{aligned} &\text{if } (V(p) = V_H(p)) \\ &\quad \{V_c(p), V_m(p)\} = \{c(pX, Y), m((1-p)X, Y)\} \\ &\text{else} \\ &\quad \{V_c(p), V_m(p)\} = \{c(X, pY), m(X, (1-p)Y)\} \end{aligned} \quad (4)$$

Figure 8: Video content pane size and metadata pane size for both horizontal and vertical layout.



Considering the browsing preferences, we categorize the approaches for optimizing content adaptation into three types. These approaches are based on which of the following are most important for the consumer; the total amount of the presented information, the preferred modality (video or text), or the layout of the whole content. The approach applied for the adaptation process is automatically determined by looking the browsing preferences, to be specific, the value p and w . [13, 11]

A. Information-centric approach. ($p = \text{variable}$, $0 < w < 1$)

Information-centric approach is applied in case the user wants to have as much valuable information as possible regardless of the modality. In this approach, the video content and the metadata are balanced adaptively in a way that maximizes information throughput. This approach is applied when value p is flexible and value w ranges $0 < w < 1$.

B. Modality-centric approach. ($p = \text{variable}$, $w = 0$ or 1)

Modality-centric approach is applied when the user wants to browse mainly a single modality and the other modality is just optional for him/her. The typical case for this approach would be the case what is important for the user is just to watch a video content and he/she doesn't care about metadata at all. This approach is applied when $w = 0$ or $w = 1$, and p is flexible.

C. Layout-centric approach ($p = \text{const}$)

Layout-centric approach is applied when the user wants to have video and metadata presented with a fixed balance and/or in a fixed position. A typical example would be when he/she feels relaxed when metadata is located under the video, or he/she always like to have the screen with two third with the video content and the rest with metadata. This approach is applied when a constant value p is given.

Figure 9: Algorithm for determining the optimal layout and value p .

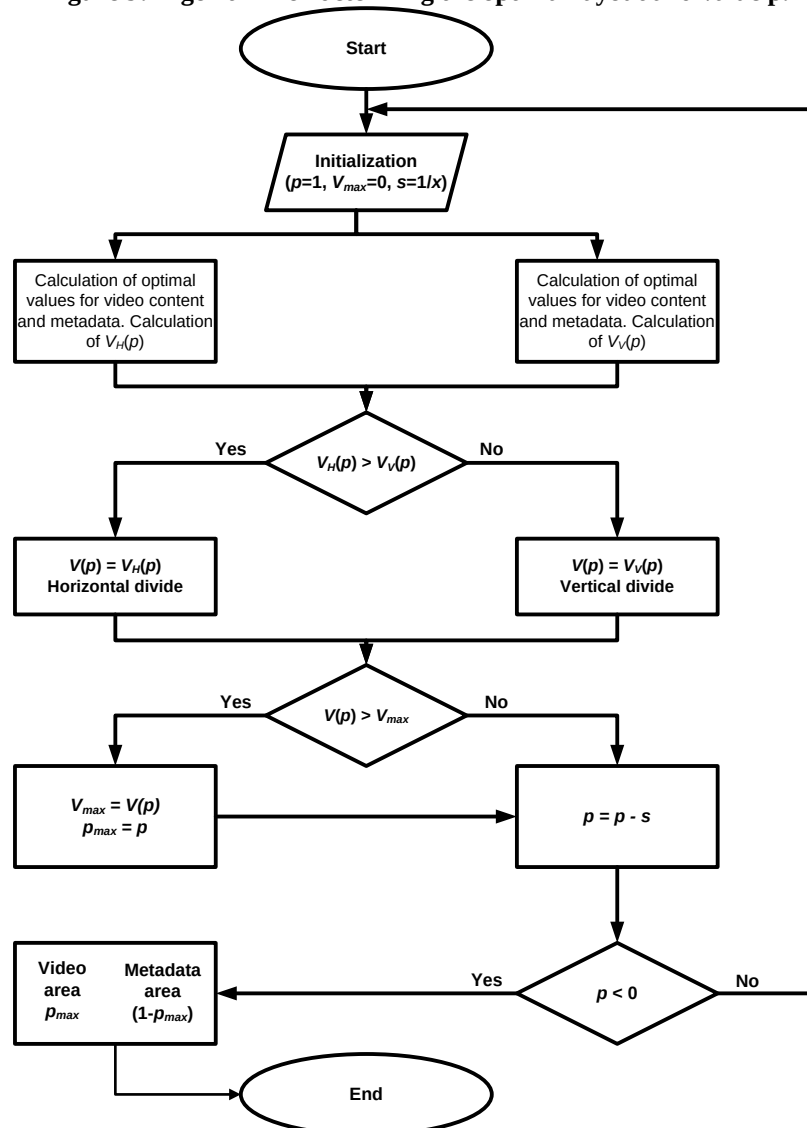


Figure 9 describes the detailed algorithm to find the optimal layout and value p . The total content value for every p is calculated and the value p which maximizes the total content value is selected as the optimal balance between video and metadata pane.

V. Conclusion

The paper presented an analysis of three approaches: Adaptation engines in Universal Multimedia Access systems, transformation engine and variation selection engine. The next part illustrates the necessity of simplify the problem, the two options the video, the metadata to be located up, down, left, or right, division into two parts, horizontally or vertically, and one part is assigned for video and the other for metadata. The mathematical expressions described in (2), (3), (4), determine the pane layout.

The purpose multimedia information to be adapted is reached with the mathematical expressions and evaluation approach, modified algorithm. The algorithm determines the optimal layout of multimedia content and the value of the adaptive variable p , as normalized parameter of video.

References

- [1] Mihaylov Gr., Video transmission over LTE mobile network, Proceedings of the Scientific Student Session 2010, University of Ruse, Ruse, Bulgaria, pp. 59-63, 2010 in Bulgarian.
- [2] Kanev J., Sadinov S., Analysis of current methods and technologies for encoding, distribution and consumption of IPTV services, International Conference ICEST 2012, Sofia, Bulgaria, pp. 74-76, 2012
- [3] Kasutani E., Investigation Report on Universal Multimedia Access, Lausanne EPFL, 2006.
- [4] Kasutani E., New Frontiers in Universal Multimedia Access, Lausanne EPFL, 2006.
- [5] Andrew P., Tor H., Content Adaptation for a Streaming Environment Enabling Universal Multimedia Access, white paper.
- [6] Thang T. C., Jung Y. J., Lee J. W., Ro Y. M., Modality Conversion for Universal Multimedia Services, white paper.
- [7] Li-Qun C., Xing X., Xin F., Wei-Ying M., Hong-Jiang Z., He-Qin Z., A visual attention model for adapting images on small displays, Multimedia Systems, Springer-Verlag, 2003.
- [8] Thang T. C., Jung Y. J., Ro Y. M., Modality Conversion in Content Adaptation for Universal Multimedia Access, white paper.
- [9] Smith J. R., Interoperable Content-based Access of Multimedia in Digital Libraries, white paper.
- [10] Pereira F., Burnett I., Universal multimedia experiences for tomorrow, IEEE Signal Processing Magazine, 2003.
- [11] Raghuveer A., Kang N. O., Du D. H. C., Techniques for Efficient Streaming of Layered Video in Heterogeneous Client Environments, white paper.
- [12] Magalhães J. M., Universal Access to Multimedia Content Based on the MPEG-7 Standard, Master of Science Thesis, Universidade Técnica de Lisboa, 2002.
- [13] Öner Ö. D., Optimal Video Adaptation for Resource Constrained Mobile Devices Based on Utility Theory, Master of Science Thesis, Middle East Technical University, 2003.
- [14] Hellwagner H., Timmerer C., MPEG Standards Enabling Universal Multimedia Access, tutorial description, Klagenfurt University, 2006.
- [15] Dalei W., Song C., Haohong W., Cross-Layer Optimization for Video Summary Transmission over Wireless Networks, IEEE Journal on Selected Areas in Communications, Vol. 25, No. 4, 2007.

Acknowledgments

The present document has been produced with the financial assistance of the European Social Fund under Operational Programme "Human Resources Development". The contents of this document are the sole responsibility of "Angel Kanchev" University of Ruse and can under no circumstances be regarded as reflecting the position of the European Union or the Ministry of Education and Science of Republic of Bulgaria.

Project № BG051PO001-3.3.06-0008 "Supporting Academic Development of Scientific Personnel in Engineering and Information Science and Technologies".