PREFERENCE-BASED MULTIMEDIA CONTENT ADAPTATION FOR SCREEN SIZE INDEPENDENT ACCESS

Eiji Kasutani
Media and Information Research Labs.
NEC Corporation
e-kasutani@cq.jp.nec.com

Touradj Ebrahimi
Signal Processing Institute
Swiss Federal Institute of Technology Lausanne
touradj.ebrahimi@epfl.ch

ABSTRACT

This paper proposes a scheme for adaptation of audiovisual contents and their metadata to any screen size so as to optimally match users' preferences. The main objective of the scheme is to balance the presentation of a requested video content and its associated metadata, such that either, the total amount of presented information, or the modality, or the presentation layout is optimized. The optimal balance is determined in a way that the total value of all the adapted contents is maximized. The total value is defined as a function of the presentation balance, where the total value for each presentation balance can be calculated from the value of adapted audiovisual content and its metadata under the given balance, and the subjective importance of each modality to the user. This scheme can be used for device independent access services to audiovisual contents with their associated metadata.

1. INTRODUCTION

The recent innovations of information technologies have enabled increasing opportunities to access and to consume audiovisual contents. The number of new appliances to consume these audiovisual contents, such as digital TV, PC, PDA, mobile phone, etc. continues to grow. At the same time audiovisual contents are being distributed with associated textual metadata. Metadata assists audiovisual content browsing by presenting additional information about the content. As there is a wide variation of the capabilities among these new devices, new technologies are needed to better bridge the gap between appropriate usage of rich contents and the capabilities offered by a specific device.

Since screen size is one of the most relevant limitations on consumer devices, it is important to maximize the user experience in presenting an audiovisual content and their associated metadata in screens of any size. A typical approach to solve this problem is to prepare different versions of the content for each type of devices [1]. The system identifies the user's terminal capabilities and automatically adapts to the most appropriate variation among already available variations. The problem with such an approach is that a large amount of effort is necessary to prepare and to manage different variations for all possible devices. Another approach is to transform the content on the fly [2-5] in a way that meets limitations of display capabilities. Such approaches aim at transforming the content so that it includes as much information as possible within the limited capabilities offered by the device. A problem with this approach is that the adapted content is always adapted in a same way to the same device without taking into account the individual tastes and browsing preferences.

This paper proposes a scheme for adaptation of audiovisual content and its metadata to any screen size in a way that maximizes user's browsing experience. The proposed scheme balances the presentation of the requested video content and its textual metadata adapting any screen size considering usage preferences.

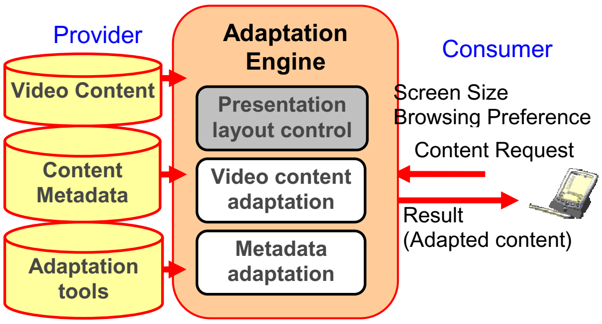

Figure 1. Overview of screen size adaptation system.

2. SCREEN SIZE ADAPTATION FOR BROWSING A VIDEO CONTENT WITH ITS METADATA

Figure 1 presents the overview of a screen size adaptation system. This system enables the consumer to access the desired video content with its metadata in an optimal way by adapting them to the device screen size as well as to his/her preferences.

Three types of data are necessary to be prepared beforehand. The video contents, their associated metadata and adaptation tools. The content metadata describe the structure of the video content. The text annotations for both the entire content and each component used for presentation are also described. The adaptation tools include hint information for guiding the adaptation process to efficiently adapt a desired content to the consumer. On the consumer side, the screen size of the device (browser) and browsing preferences are described and used for adaptation.

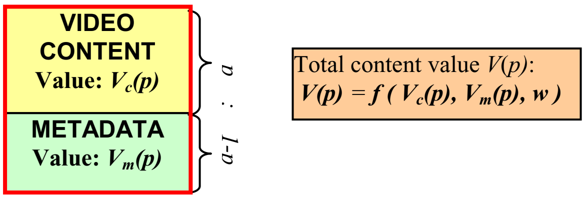

Figure 2. Definition of total content value as a function of the balance of the pane size.

Two types of browsing preferences are considered. The first one is the preferred layout structure, that is, the preferred balance between the video presentation area (video pane) and the metadata presentation area (metadata pane) within a given screen size. The second one is the balance of the importance of the video and its associated metadata for the consumer.

The adaptation procedure is as follows. First, the consumer requests the desired content. At that time, the browsing preferences and the screen size are sent to the provider to let the provider know about device capabilities and user preferences. The adaptation engine dynamically adapts these different modalities and balance their presentation to maximize browsing experience within a given screen size. The adapted video content and the adapted metadata are wrapped and delivered to the consumer in a suitable description format (e.g. SMIL, HTML) considering their presentation balance.

3. MODELING TOTAL CONTENT VALUE

To optimize the content adaptation process, a quality metric to evaluate how well the content is adapted, and 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 desired. This metric should define a value for the total adapted content evaluating the amount of fidelity of the overall adapted content when compared with the original content. Thus, the optimization process becomes that of determining the way of adaptation that maximizes the total content value meeting the given constraints.

To balance the presentation of the video content and its associated metadata, it is necessary to dynamically determine how to divide the screen. Two aspects should be taken into account in dividing the screen: the size and the layout. To express the size of each pane, a variable p is defined as a normalized size of the video pane (see Fig. 2). This variable is normalized to be between 0 (the whole screen is devoted to metadata display) and 1 (the whole screen is devoted to video display). Thus the balance between the size of the video pane and metadata pane can be expressed as $p:1\text{-}p$. The screen can be divided both horizontally and vertically.

Once the pane size to present video and 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.

Three parameters affect the total content value: 1) the adapted video value, 2) the adapted metadata value, and 3) the preferred balance of their importance. The total value function $V(p)$ can be modeled as follows;

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

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) .

The total content function $V(p)$ evaluates the balance between the adapted video and metadata, and takes into account the subjective importance of each modality for the consumer.

Considering the video content and associated metadata values, 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.

Any definition of the adapted video content value function $V_c(p)$ and metadata value function $V_m(p)$ is applicable. In this paper, we define $V_c(p)$ as a scaling ratio that maximizes the video content resolution inside the given video pane size. $V_m(p)$ is defined as a balancing function between the amount of information and number of pages necessary to present all the information. Value w is a normalized value of the video importance which is provided by the user, through a mechanism such as browsing preferences as explained in section 2. The balance between the importance on video content and on metadata can be expressed as $w : 1\text{-}w$ $(0<w<1)$.

4. OPTIMIZING CONTENT ADAPTATION CONSIDERING BROWSING PREFERENCES

4.1. Overview

Considering the user preferences on browsing, various approaches for optimization of the content adaptation process can be categorized into three types. These approaches depends on which of the following are most important for the consumer: the total amount of the presented information, the preferred modality (video or text), and the layout of the whole content. Hereafter they will be referred to as: a) *information-centric approach*, b) *modality-centric approach*, c) *layout-centric approach*, respectively. The approach applied for the adaptation process is automatically determined by analyzing the browsing preferences, namely, the values p and w. Figure 3 shows its determination procedure.

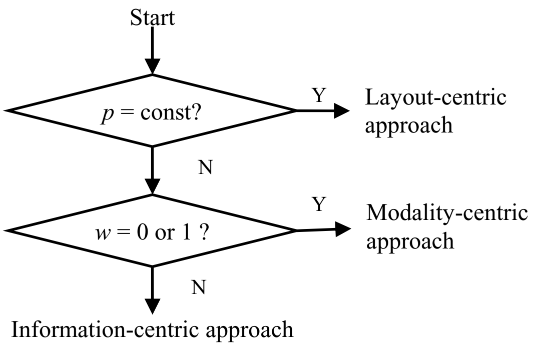

Figure 3. Decision of adaptation process approach based on browsing preferences

4.2. Information-centric approach

Information-centric approach is applied when 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 $0<w<1$.

In information-centric approach, the total content value is optimized so as to provide as much information as possible to the consumer. In this approach, we try to find the value p that maximizes function (1) as described in section 3. The total content value for every possible p is calculated and the value p which maximizes the total content value is selected as the optimal balance between video and metadata pane.

4.3. Modality-centric approach

Modality-centric approach is applied when the user wants to browse mainly a single modality and the other modality is just optional. A typical example of this approach is when someone watches a video content and he/she doesn't care much about the metadata. This approach is applied when $w = 0$ or $w = 1$, and p is flexible.

The layout is determined in a way that maximizes value of the modality the user is interested in. When $w=0$, the user is mainly interested in metadata, and the video is just ignored, devoting the entire screen to presentation of metadata. When $w=1$, the user is mainly interested in video, the video content resolution is maximized to fit the whole screen. If there is enough room left in the screen after presenting the adapted video content, the associated metadata is presented in that area. If the area left is too smaller, the metadata is just ignored. It should be noted that the aspect ratio of the adapted video content is always kept the same. We try to maximize the video content resolution while keeping the video content size within the screen. There are two possibilities that could maximize the video content size inside the screen. The first case is that the width of the adapted video content is identical to the screen width X. The second case is that the height of the adapted video content is identical to the screen height Y. The layout and video content scaling ratio b are determined as follows;

b = min (b1, b2) *where* $b1 = X / x$ *and* $b2 = Y / y$
if $(b > b_{min})$
  if (b == b1) vertical layout.
  if (b == b2) horizontal layout.
if $(b > b_{max})$ $b = b_{max}$;
if $(b < b_{min})$ $b = 0$

Parameters:
- b: Video content scaling ratio.
- (X, Y): Width and height of the screen.
- (x, y): Width and height of the original video.
- b_{max}, b_{min}: Maximum, minimum allowed scaling ratio.

4.4. Layout-centric approach

Layout-centric approach is applied when the user wants to have video and metadata presented with a pre-defined balance and/or in a pre-defined position. A typical example would be when he/she always likes to have the screen in two third covered with the video content and the rest with metadata. This approach is applied with a constant value p reflecting the pre-defined layout preference. Therefore, the video content value and metadata values are separately maximized within the given pane size for each modality.

Two methods can be used to determine the layout pattern and the scaling size of the video content for the fixed pane balance. The methods differ depending on the value of w. In the first method, $0<w<1$ and in the second $w=1$. The methods are described in the following.

0<w<1 case:

1. Maximize both video content and metadata value independently in each pane.
2. Calculate the total content value for both horizontal layout and vertical layout.

Horizontal layout: $V_H(a) = f(V_c(a), V_m(a), w)$
$\quad = f(v_c(aX, Y), v_m((1-a)X, Y), w)$

Vertical layout: $V_V(a) = f(v_c(X, aY), v_m(X, (1-a)Y), w)$

3. Adopt the one with higher content value.
   $V_H(a) > V_V(a)$: Horizontal layout.
   $V_V(a) > V_H(a)$: Vertical layout.

w=1 case:

For $w=1$ case, first, one has to decide between a horizontal or a vertical layout considering the maximum video content size that can be displayed in each layout. For fast calculation, the layout and the video content scaling ratio b are determined in the following way;

$b5 = min\ (b1, b2)$ where $b1 = X / x$ and $b2 = aY / y$
$b6 = min\ (b3, b4)$ where $b3 = aX / x$ and $b4 = Y / y$
$b = max\ (b5, b6)$
if $(b > b_{max})\ b = b_{max}$
if $(b > b_{min})$
  if **$(b == b5)$** ***vertical layout.***
  if **$(b == b6)$** ***horizontal layout.***
if $(b < b_{min})\ b = 0$

Parameters:

a: Balance of video content and metadata presentation area given by the user (video : metadata = a : 1-a. a = constant value).
b: Video content scaling ratio.
(X, Y): Width and height of the screen.
(x, y): Width and height of the original video content.
b_{max}, b_{min}: Maximum, minimum allowed scaling ratio.

4. EXPERIMENTAL RESULTS

Three news contents from MPEG-7 dataset were used in our experiments. As a content metadata for presentation, a title and a detailed text were generated for each news content, as well as for every news items. The video and its associated metadata were adapted to several screen sizes. The resolution of the video content, the number of titles and that of detailed text to be presented were then selected considering the screen size and browsing preferences.

Figures 4 to 6 present examples of video and metadata adaptation to the same screen size based on different user browsing preferences. Figure 4 shows an adaptation example using information-centric approach. Figure 5 illustrates an example of video-centric approach where video resolution has to be always maximized. In metadata presentation, only the titles of each news item are displayed due to the small area left after the display of the adapted video content. An example of adaptation by layout-centric approach (video : metadata = 1:1) is shown in Fig. 6.

The computational cost of the adaptation process on our test bed were evaluated using a Linux machine with Pentium 4 3GHz CPU. The average processing time for adaptation was 3.39 msec on a screen size of 240x320, 6.96 msec for 640x480, and 17.0 msec for 1280x1024.

5. CONCLUSION

We have proposed a scheme for efficient adaptation of audiovisual contents and their associated metadata to devices with any screen size considering user preferences. The proposed scheme dynamically balances the presentation of the requested video content and its metadata within a given screen size in a way that maximize the content value to the user. The content value for the user is maximized by applying different adaptation methods taking into account the most important parameters to the user between information, modality and layout. Future work includes user studies to validate our approach and objective means to assess the quality of experience provided to the user.

Figure 4. An example of adaptation by "information-centric approach": This approach tries to provide as much total information as possible.

Figure 5. An example of adaptation by "modality-centric approach": The video size is always maximized.

Figure 6. An example of adaptation by "layout-centric approach": This user prefers an equally balanced video content and associated metadata in display size (1 : 1 ratio).

REFERENCES

[1] R.Mohan, J.Smith, C.S.Li, "Adapting Multimedia Internet Content for Universal Access", IEEE Transactions on Multimedia, pp. 104-114, March 1999.
[2] K.Lee, H.Chang; S.Chun, H.Choi, S.Sull, "Perception-based Image Transcoding for Universal Multimedia Access", IEEE Int. Conf. on Image Processing 2001, vol.2, pp. 475 -478, 2001.
[3] X.Fan, X.Xie, W.Ma, H.Zhang, H.Zhou, "Visual Attention Based Image Browsing on Mobile Devices", IEEE Int. Conf. on Multimedia and Expo 2003, vol.1, pp. 53-56, 2003.
[4] A.Cavallaro, O.Steiger, T.Ebrahimi,"Semantic Segmentation and Description for Video Transcoding", IEEE Int. Conf. on Multimedia and Expo 2003, vol. 3, pp. 597-600, 2003.
[5] H.Liu, X.Xie, W.Ma, H.Zhang: "Automatic Browsing of Large Pictures on Mobile Devices". ACM Multimedia 2003: pp.148-155, 2003.