VISUAL CONTENT ADAPTATION FOR LOW VISION USERS IN MPEG-21 FRAMEWORK

Truong Cong Thang and Yong Man Ro

Multimedia Group, Information and Communication University (ICU), South Korea
Email: yro@icu.ac.kr

ABSTRACT

This paper deals with the visual content adaptation, in the context of MPEG-21 standard, to help low vision users have better accessibility to the contents. The proposed adaptation targets at two low vision symptoms, loss of fine detail and lack of contrast. Specifically, we present an adaptation framework describing the problem space, and then a systematic contrast-enhancement method to improve the content visibility for low vision users. The experiment results show that the proposed framework and method are effective for the low vision users.

1. INTRODUCTION

Accessibility is an important requirement for any multimedia system. To this end, we have developed the description tools to describe the visual impairments of users in the MPEG-21 framework [1]. Visual impairments consist of two major categories: *color vision deficiency* and *low vision*. The overview of visual impairments and content adaptation was presented in [2], and the adaptation methods for color vision deficiency were proposed in [3]. This paper continues the track by considering the adaptation for low vision. Basically, low vision means that even with regular glasses, medicine, or surgery, the patients still have difficulties in everyday tasks, of course including accessing the existing multimedia contents.

It should be noted that in the context of MPEG-21 Digital Item Adaptation (DIA), one has to consider 1) how to get the metadata/description and 2) how to use the metadata to enable the automatic adaptation process. In this paper, we will mainly deal with the second issue.

Instead of listing various unfamiliar names of low vision conditions, the low vision description consists of seven popular symptoms: loss of fine detail, lack of contrast, light sensitivity, need of light, central vision loss, peripheral vision loss, and hemianopia. The rationale for this symptom-based description can be found in [4][2].

Among these symptoms, *loss of fine detail* and *lack of contrast* are the most popular ones and they are the focus of this paper. Loss of fine detail causes the perceived image blurred; its level can be described by the clinical visual acuity normalized into the range [0,1]. Lack of contrast causes the dynamic range of image intensity to be reduced; its level can be described by the clinical contrast sensitivity normalized into the range [0,1]. In the standardized description, the higher the level is, the more severe the symptom becomes. The terms *low vision user* and *patient* are used synonymously in this paper to mean the user having these two symptoms.

Though not explicitly mentioned, the researches on image (and video) enhancement for low vision have so far dealt mainly with the problems related to the two above symptoms [5][6][7]. However, the current methods are still application-specific and difficult for the automatic adaptation. Besides, most of other methods of image enhancement only target at normal users (e.g. [8]), even disregard the human visual characteristics. Moreover, a framework that clarifies the problem space for this kind of adaptation has not been provided before.

In this paper we first propose a content adaptation framework specific for lack of contrast and loss of fine detail, with the focus on the adaptation possibilities. Then we propose a systematic method that automatically adapts the visual contents for coping with these two symptoms.

The paper is organized as follows. In section 2 the adaptation framework for low vision is discussed. Section 3 presents a method of contrast enhancement, which is efficient for low vision user. Some experiments are shown in section 4, and finally conclusions and future works are given in section 5.

2. ADAPTATION FRAMEWORK FOR LOW VISION

The general adaptation framework is depicted in Fig. 1. First, the user specifies the symptoms and the associated severity levels using MPEG-21 description tool. These parameters will be used in the content adaptation process. Normally, user's model of human visual system (HVS) is constructed based on the specified parameters. Then the contents are adapted according to the obtained HVS model. As a result, the low vision user perceives the adapted contents with better visibility. The HVS model can also be used to simulate low vision user's perception.

0-7803-8554-3/04/$20.00 ©2004 IEEE.

Thus the verification of adaptation process can be done through both simulations and tests with real patients. This framework is in fact common for any visual impairment.

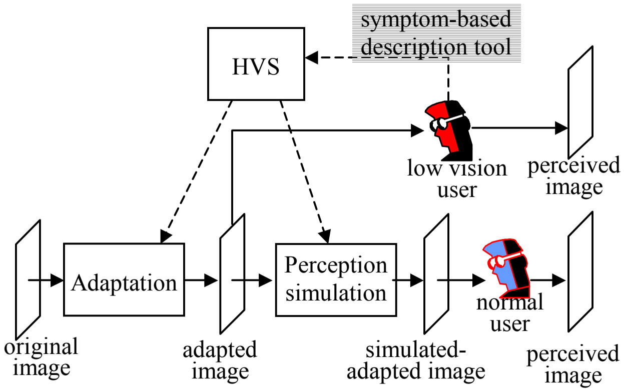

Fig. 1: General adaptation framework for low vision.

To handle the symptoms loss of fine detail and lack of contrast, we employ the contrast sensitivity function (CSF), a popular representative model of HVS [6]. This function shows the user's contrast sensitivity across different spatial frequencies. We can note that the clinical contrast sensitivity reflects the contrast sensitivity near the peak of the function and clinical visual acuity reflects the sensitivity at the right-hand limit of the function [9]. Denote the normalized levels of lack of contrast and loss of fine detail by L_c and L_d, and the CSF of a normal user by $CSF(u)$, where u is the spatial frequency measured in cycles/degree. We construct the contrast sensitivity function $CSF_{lv}(u)$ of a low vision user as follows:

$$CFS_{lv}(u) = (1-L_c)\times CSF(\frac{u}{1-L_d}) \qquad (1)$$

We see that when L_c increases, the function is scaled down; and when L_d increases, the function is shrunk. However the mappings from clinical measures of contrast sensitivity and visual acuity to values of L_c and L_d are not simply linear. We currently carry out the clinical experiments to find the accurate mapping tables between the clinical measures and the normalized levels of the symptoms. Anyway, it should be noted that the research on low vision CSF construction and the research on adaptation can be relatively separated. In the following, the focus is on the systematic adaptation based on a CSF of low vision user, regardless of how the CSF is obtained.

The perception simulation (or simulation for short) of low vision closely relates to the HVS. Several simulation models have been discussed in [5][6][7]. The multiscale model based on the local band-limited contrast seems very promising. However, this is a threshold model so it cannot represent correctly the apparent contrast (or perceived contrast) [10]. In this work, we use the CSF as the linear filter and the simulation is based on the ratio of a patient's CSF to a normal user's CSF [6].

For the adaptation, we consider a more detailed version of the general framework as in Fig. 2. The image can be represented by a characteristic spectrum (either the conventional frequency spectrum or the contrast spectrum [5]). The HVS of patient obviously tends to reduce 1) the amplitude, especially at the high frequencies, and 2) the width of the image spectrum. So the possible adaptation operations on the spectrum to enhance the content visibility are:

- Increasing the amplitude of the image spectrum by increasing the contrast and sharpness of image.
- Reducing the width of image spectrum by enlarging the spatial size of image.

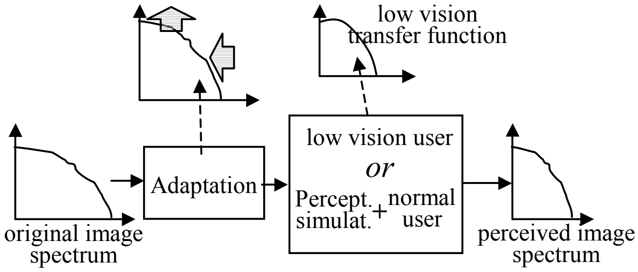

Fig. 2: Adaptation for loss of fine detail and lack of contrast.

In content adaptation, enlargement is not necessarily zooming the image based on interpolation, instead the provider can select a high quality version of larger size and send it to user. In fact, the user can change the viewing distance instead of using enlargement. Anyway, controlling of image size is still necessary because in many cases the user cannot change viewing distance, e.g. user is using a wearable computer headset.

The visual content adaptation for low vision is not only limited to the above low-level image processing. It may further include the semantics-based methods. For example in an image, there may be only one region of interest (ROI) that should be enhanced or enlarged, the remaining area can be removed or left intact. Even, if the user's symptoms are too severe, the visual contents can be converted to different modalities (e.g. audio) [11]. In essence, these adaptation possibilities are valid for different kinds of simulation models. In the next section, we will present a systematic contrast-enhancement method to improve the visibility of natural image.

3. AUTOMATIC CONTRAST ENHANCEMENT OF NATURAL IMAGE FOR LOW VISION

Before presenting an enhancement method for low vision users, we look at some previous works. The sharpness of the image can be improved for low vision user by a deblurring filter which pre-compensates for the degradation in HVS. However, this method is very limited [7]. This is because the dynamic range of display device is limited while the natural image is usually sharp already, and the HVS of low vision user often discards lots of pre-compensated high frequencies.

In [6], the adaptive filtering (unsharp masking) is used to enhance the contrast of face images for low vision users. However the adaptability of the method is actually not utilized as stated by the authors. The method essentially increases, by a fixed gain, the contrasts at frequencies visible by low vision user and crucial for the face recognition. That means this method is application-specific. Further, a pre-emphasis filter is employed in [7]. In fact this method is nearly the same as the previous adaptive filtering. The new features in the pre-emphasis filter are that it filters out the invisible frequencies and has the different gains for different critical frequencies. These methods show useful insights into low vision perception, however they are not general and not appropriate for the automatic adaptation.

Meanwhile, contrast enhancement methods devised by signal-processing experts just target at normal user cases where perception simulation is unnecessary. Among those, JGACE, an adaptive unsharp masking method proposed in [8], is shown to systematically adapt to both the image's characteristics and user's HVS.

In this section, we propose a method that combines the advantages of these previous works. Similar to the previous ones, our method is basically unsharp masking.

As noted in [7], the high frequencies are invisible to low vision user so they should be removed to give more dynamic range for other frequencies. Then the original image I_0 is first low-pass filtered with the cutoff frequency f_{up} similar to that of the CSF-based filter in simulation. The resulting low-pass image I_1 is the one that is actually enhanced by unsharp masking and then presented to user.

The unsharp mask I_2 is obtained by low-pass filtering I_1 with the cutoff frequency f_{lo}, which is 16 cycles per picture (c/p) in [8]. This frequency f_{lo} would limit the enhancement for the low vision user because with the user having a high value of L_d, the contrasts at low frequencies would be the only means to perceive the contents. Thus, we modify f_{lo} accordingly to the loss of fine detail as follows:

$$f_{lo} = (1-L_d)\times 16 \qquad (2)$$

Now f_{lo} becomes lower when L_d increases.

The image I_3 obtained as the difference of I_1 and I_2 is actually the "band-pass" (not high-pass) image with upper and lower cutoff frequencies being f_{up} and f_{lo}. Then the local contrast and local frequency of every region of I_3 are computed in the same way as part 3 of JGACE [8]. Note that low vision user often has certain loss of fine detail, that is, the small noise in usual images is likely unnoticeable by the user and it would be removed by the cutoff frequency f_{up}.

Here, we use the normal user CSF (i.e. reciprocal of threshold contrast function) in [8]:

$$CSF(u) = (1/c_o)\times e^{-0.166u} \qquad (3)$$

where c_o is now normalized to be 1. From (1) and (3), $CSF_{lv}(u)$ of low vision user is computed as follows:

$$CSF_{lv}(u) = \frac{(1-L_c)}{c_o}\times e^{\frac{-0.166u}{1-L_d}} \qquad (4)$$

This function is also used in perception simulation. Because we do not consider the noise, the local contrast gain G for every region is only calculated by the equation (12) of [8], in which $CSF_{lv}(u)$ is used instead of $CSF(u)$.

Finally the amplified band-pass image I_3 is added back to I_2 to get the final enhanced image. Also we find that the contrast gain for low vision often very strong, so increasing the global image contrast as in [8] is not necessary. Moreover, it is no longer necessary to distinguish the smooth and detail regions.

As mentioned, this proposed method has the advantages from the previous ones: compared to method of [8], our method is less complex and more appropriate for the low vision; while compared to those in [6][7], this method is more systematic and adaptive.

4. EXPERIMENTS

Experiments with various values of L_d and L_c are carried out to verify the automation and effectiveness of the proposed framework and image enhancement method. Currently the adaptation is checked through perception simulation only. In our experiments, the test images have size of 256x256, and the viewing distance is 50cm.

Fig. 3 illustrates an example experiment with Lena image when the patient has $L_d = 0.8$ and $L_c = 0.2$ (e.g. a cataract patient). Image 3(a) is the original image seen by normal user, 3(b) is the simulation of original image perceived by the patient, 3(c) is the enhanced image provided by the proposed method, and 3(d) is the image perceived by the patient from 3(c). We can see that the enhanced image 3(c) has high contrasts in the very frequencies visible by the patient, and so the patient can see more information as shown by image 3(d).

Another example is shown in Fig. 4 with a lake image, when $L_d = 0.1$ and $L_c = 0.9$ (e.g. an patient with multiple sclerosis). In this case the contrast of the perceived image 4(b) is very low. Accordingly, the enhanced image 4(c) has very high contrast at the details of suitable frequencies, not just at the edges of objects in the image. Especially, this enhanced image highlights a small cloud in the sky, which is even difficult to be noticed by normal user in original image 4(a). This shows the adaptability of the method to the local contrast, which is not possible with the methods in [6][7]. As a result, many more details are noticeable in the perceived image 4(d) of the enhanced image. Note that the results displayed on monitor screen are much clearer than the printed ones.

Fig. 3: Experiment result with $L_d = 0.8$ and $L_c = 0.2$.

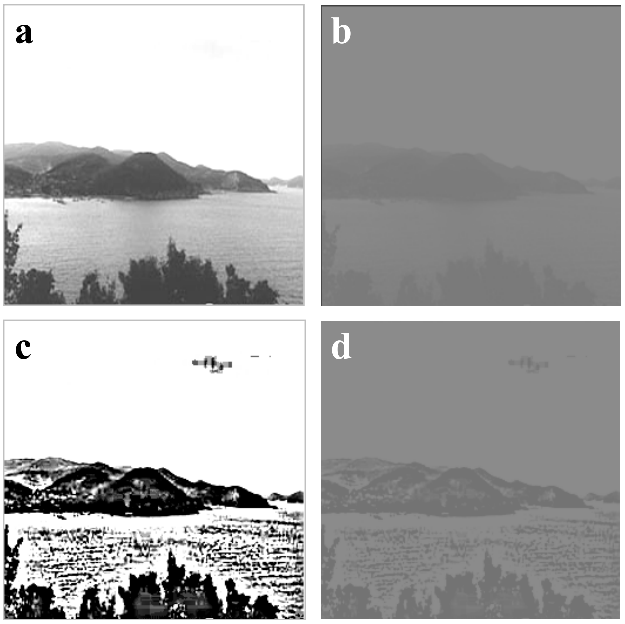

Fig. 4: Experiment result with $L_d = 0.1$ and $L_c = 0.9$.

5. CONCLUSIONS AND FUTURE WORKS

We have presented a content adaptation framework to provide the accessibility to low vision users having the symptoms loss of fine detail and lack of contrast. We point out the possible adaptation solutions in the framework, and then propose a systematic contrast-enhancement method to enhance the visibility of images for the users. The experiment results demonstrate the effectiveness of the proposed method and framework.

In future, we will study on the content-based enhancement methods that are more effective for low vision users. Size enlargement is also well effective with the natural images, thus we are exploring the tradeoff between contrast enhancement and size enlargement. The optimally combined adaptation would depend on some “contrast measure” of image, on the user’s HVS, and on the user’s preference. Further, computational efficiency of the adaptation should be improved for real-time requirement. To handle the other specified symptoms of low vision, more complex HVS models will be considered. Finally, the current work, which processes only gray image/video, will be extended to support the colored contents.

6. REFERENCES

[1] ISO/IEC 21000-7 FDIS, “Information Technology – Multimedia Framework – Part 7: DIA,” Dec. 2003.

[2] S. J. Yang, T. C. Thang, and Y. M. Ro, “Content Adaptation for Visual Impairment in MPEG-21”, Proc. of SPIE Electronic Imaging, vol. 5292, Jan. 2004.

[3] S. J. Yang, and Y. M. Ro, “Visual Contents Adaptation for Color Vision Deficiency”, Proc. of ICIP’03, Sept. 2003.

[4] J. Nam, T. C. Thang, J. Song, Y. M. Ro, J. W. Hong, “Report on CE for Visual Impairment,” ISO/IEC JTC1/SC29/WG11 m9190, Awaji, Japan, December, 2002.

[5] E. Peli, “Contrast in Complex Images”, Journal of Opt. Soc. America A, vol. 7, pp.2032-2040, 1990.

[6] E. Peli et al., “Image Enhancement for The Visually Impaired: Simulations and Experimental Results,” Invest. Ophthalmol. Vis. Sci., vol. 32, pp.2337–2350, 1991.

[7] E. Peli. "Limitations of Image Enhancement for The Visually Impaired.", Optometry & Vision Science, vol. 69, pp.5-24, 1992.

[8] T.-L. Ji, M.K. Sundareshan, H. Roehrig, “Adaptive Image Contrast Enhancement Based on Human Visual Properties”, IEEE Trans. Medical Imaging, vol. 13, pp.573-586, Dec. 1994.

[9] P. Lennie and S. B. Van Hemel, “Visual Impairments: Determining Eligibility for Social Security Benefits”, National Academy Press, Washington DC, 2002.

[10] S. N. Pattanaik et al., "A Multiscale Model of Adaptation and Spatial Vision for Realistic Image Display", Proc. of SIGGRAPH'98, pp. 287-298, Orlando, July 1998.

[11] T. C. Thang, Y. J. Jung, J. W. Lee, Y. M. Ro, “Modality Conversion for Universal Multimedia Services”, Proc. Int. Workshop on Image Analysis for Multimedia Interactive Services, Lisboa, April, 2004.