Empirical Evaluation of Real-Time Video Foveation

Ayub Bokani
Computer Science and Engineering School
The University of New South Wales, Sydney, NSW, Australia
National ICT Australia (NICTA)
abokani@cse.unsw.edu.au

ABSTRACT

The *foveated* video compression techniques were developed to deliver high-quality video at lower bitrates, matching the nonuniform sampling in human's visual system (HVS). In these techniques the *fovea* region i.e., the area of about 2-5° around gaze centre is compressed at the highest bitrate. Allocating lower bitrates to other regions based on their distances from fovea decreases the overall bitrate. This technology could be used in streaming high quality videos over the Internet with minimal bandwidth consumption. In this study we evaluate our implementation of real-time foveation algorithm as the first step for foveated video streaming.

Keywords

Foveation; Video Streaming

1. INTRODUCTION

The distribution of photoreceptors on our retina is highly nonuniform. In our visual system, a small region of 2-5 degree around the centre of the retina, called fovea, is responsible for sensing visual information in highest detail. With increasing distance from the centre, retinal sensitivity, i.e., its ability to process visual details, falls off logarithmically. This motivates the foveated video, an increasingly researched video coding technique that allocates more bits to a small region in the frame where the viewer is likely to point his eyes (centre of gaze), but increasingly less bits to other parts further away from the centre of gaze.

There are two important benefits of foveation: First, for the same network bandwidth, foveation significantly increases the perceived quality of experience (QoE); Second, similar QoE can be maintained using significantly less network bandwidth.

2. REAL-TIME FOVEATION

Foveated video streaming requires a foveation toolbox which compresses the video images online. In such toolbox, a mul-
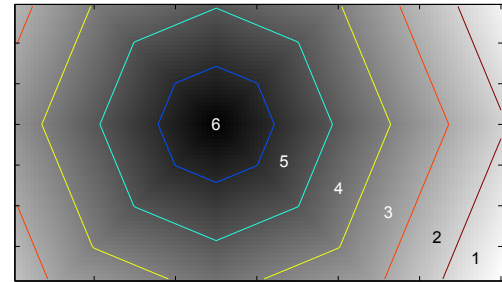
Figure 1: Using *chamfer* distance function to allocate 6 different bitrates based on their distances from the *fovea* centre (*frame 1946, the Big Buck Bunny movie*)

tiple bit allocation scheme allocates different bitrates to every pixels of image based on their distances to the fovea centre [1]. Figure 1 demonstrates how 6 different quality levels are allocated to different regions of an image with higher values indicating higher bitrates.

For this purpose, we use the freely available MATLAB foveation toolbox [2]. This software uses the mouse cursor to define the fovea centre and generates the foveated images in real-time basis. With some modifications, we connected it to our eye-tracker device to dynamically change the fovea region based on subject's real time gaze data.

3. DEMO SETUP AND RESULTS

To evaluate the real-time video foveation, we set up a foveation lab and collect different subjects' quality assessments based on their observations as following:

3.1 Subjective Quality Assessment

A Tobii-X60 eye-tracker is mounted on a 22" screen and adjusted in 60-70 cm distance to the subject's eyes (Figure 3). Two different Full-HD videos, the *Big Buck Bunny* and *Tower* were shown to subjects in both *foveated* and *regular* (i.e., uncompressed) modes in two different qualities 5 and 10. These numbers indicate the lowest quality level in a frame where the uncompressed frames have a uniform quality 15. In a foveated video with quality 10, the fovea region has the same quality as uncompressed frames. Ten subjects including five males and five females aged 24 - 38 were participated in this experiment. In each session, 200 video frames were played continuously in a streaming loop. The frames
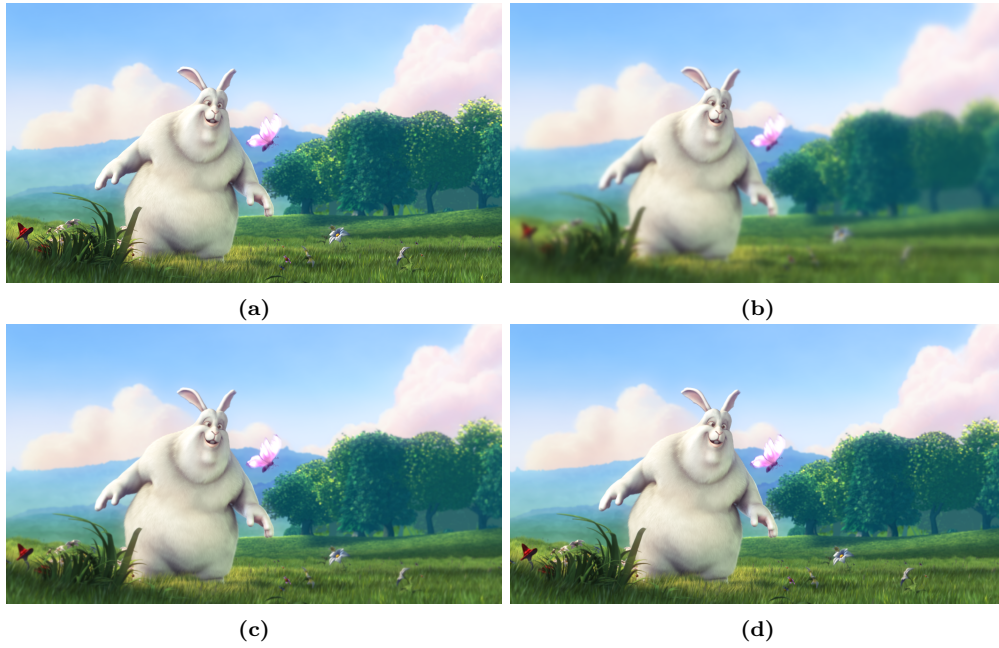
Figure 2: Foveation - frame 1946 of Big Buck Bunny movie; (a) Uncompressed; (b), (c) and (d) are foveated frames with qualities 1, 5 and 10 respectively.
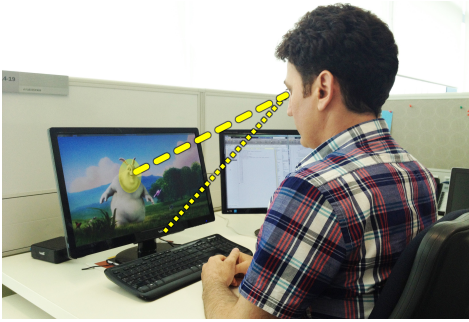

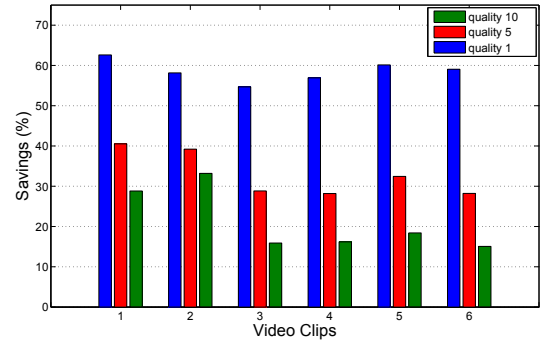**Figure 3:** Real-time foveation with Tobii-X60 eye tracker

Figure 4: Saving rates - Y column: saving % ; X column: Video clips, 1. *Big Buck Bunny* (1080×1920), 2. *Tower* (1088×1920), 3. *Avner* (304×720), 4. *Caesar* (400×640), 5. *Hana* (340×640), 6. *Kathrine* (340×640)

were played back in the *regular* mode first and introduced as reference video to the subjects. Then, after watching the same video in *foveated* mode, they were asked to rate the similarity of tow videos from 1-10 with 10 indicating 100% similar perceptual experience. In average, the videos with quality 10 were considered as 92% similar to the reference videos and video quality 5 achieved 74 % similarity.

3.2 Savings

We used 6 different video streams in 3 different qualities and resolutions to measure the saving rates and validate the foveated streaming idea. As it's shown in Figure 4, in high resolution foveated videos with quality 10 we can save up to 33.21% bandwidth. The three quality levels as well as uncompressed version of frame 1946 of Big Buck Bunny video are shown in Figure 2 for comparison.

4. CONCLUSIONS

In this study we have empirically evaluated the performance of real-time video foveation. With subjective quality assessment, we found that foveated videos with very close perceptual qualities to the original videos can be delivered to the end user with significantly lower bandwidth requirements. This technique could be used in high quality video streaming which is a challenging task as the client's real-time gaze data can not be accessed in DASH. This problem will be addressed in our future work.

5. REFERENCES

[1] Z. Li, S. Qin, and L. Itti. Visual Attention Guided Bit Allocation in Video Compression. *Image and Vision Computing*, 29(1):1–14, 2011.

[2] C. f. P. S. The University of Texas at Austin. Space variant imaging system: Foveation toolbox. [Online accessed 04-September-2014], URL: http://svi.cps.utexas.edu/software.shtml.