Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

J. Vis. Commun. Image R. 16 (2005) 180–211

JOURNAL OF VISUAL Communication & IMAGE Representation

www.elsevier.com/locate/jvci

High quality, low delay foveated visual communications over mobile channels☆

Sanghoon Lee[a,*], Alan C. Bovik[b], Young Yong Kim[a]

[a] *School of Electrical and Electronic Engineering, Yonsei University, Seoul 120-749, Korea*
[b] *The University of Texas at Austin, Austin, TX 78712, USA*

Received 28 March 2003; accepted 22 June 2004
Available online 21 August 2004

Abstract

Foveated video streams are created by selectively varying the spatio-temporal resolution of video data according to the assumed or measured fixation patterns of the intended viewers. The significant potential of foveated video lies in the considerable entropy reduction relative to the original video data while minimizing the loss of apparent visual information. By exploiting new human–machine interface techniques, such as visual eyetrackers, combined with foveated video compression and communication protocols, more efficient visual multimedia services operating over low bandwidths should become available in the near future. In this paper, we introduce a prototype foveated visual communication system suitable for implementation as a core element of an interactive multimedia wireless communication environment. We demonstrate the benefit of using foveation in noisy wireless low bandwidth applications, using measured fading statistics from the downtown area of Austin, Texas as an example. Based on a maximum source throughput criterion, the source and channel video transmission rates and the target video coding bit rate are adaptively decided according to the channel dynamics. In the simulations, we use the channel throughput, the spatial/temporal resolution, and the transmission delay as criteria to compare the performance of the foveated approach relative to normal (non-foveated) video. The results clearly underline the significant potential of foveated video communication protocols for wireless multimedia applications.
© 2004 Elsevier Inc. All rights reserved.

☆ This work was supported (in part) by the Ministry of Information and Communication, Korea, under the Information Technology Research Center (ITRC) Support Program.
* Corresponding author. Fax: +82-2-313-2879.
E-mail addresses: slee@yonsei.ac.kr (S. Lee), bovik@ece.utexas.edu (A.C. Bovik).

1047-3203/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jvcir.2004.06.001

Keywords: Foveated visual communications; Foveated video; Wireless video; Foveation; Video compression; Image compression; Digital video; Rate control

1. Introduction

Research into real-time visual communications over wireless or wireline channels has made significant recent progress because of significant advances in understanding human vision, video compression, network design, communication protocols, and the interplay between these areas.

In applications involving point-to-point visual communications, each subscriber is allocated a given bandwidth that is determined according to the required quality of the reconstructed video data. Because of the increased availability of very low-bandwidth digital communications channels coupled with multimedia applications that require high-quality visual service, the development of more powerful video compression and transmission techniques remains of significant topical interest. One promising means for increasing the efficacy of video transmission and compression takes into account the interface between the human user and the displayed video, as well as an important property of the human visual system.

The retina of the human eye samples visual information non-uniformly in space: the resolution is highest where the optical axis of the eye strikes the retina (known as the *fovea*) and it decreases as a function of angular distance (*eccentricity*) from this point. The repurcussion of this is that a human observer sees only a small part of any scene with high resolution at a given moment, with the rest of the scene serving as lower-resolution context or memory from fixations at prior moments. Likewise, a human observer of a video display of sufficient size (relative to the viewing distance) sees only a part of the display at high resolution. It therefore makes engineering sense that, if the fixation point of the observer(s) could be known, then the video could be transmitted with a variable resolution decreasing from fixation, with a consequent reduction in the requisite bandwidth. Indeed, this idea is not new, and several authors have proposed methods for increasing compression efficiency by exploiting the non-uniform sensing apparatus of the human eye via the creation of *foveated images* and *foveated video* (Ebrahimi and Kunt, 1998; Hartung et al., 1998; Kunt et al., 1985; Silsbee et al., 1993).

The goal of foveated video compression algorithms is generically to improve the apparent (subjective) visual quality of the data by allocating more resources to perceptually important regions, which usually means those regions which are being fixated, but which could also encompass other measures of "interest" such as image motion, spatial images features, or even context.

Recently, foveated video has been actively studied as a promising means for managing high-quality video transmission in lowbit rate video applications involving human interaction (Geisler and Perry, 1998; Lee et al., 2002, 2001; Reeves and Robinson, 1996). For point-to-point visual communication applications, a fixation point can be chosen by an end user using an interactive approach such as a visual eye tracker, a mouse, or a touch screen. Among these, only the eyetracker may be

Download English Version:

https://daneshyari.com/en/article/10360215

Download Persian Version:

https://daneshyari.com/article/10360215

Daneshyari.com