

Integrating Perceptual, Device and Location Characteristics for Wireless Multimedia Transmission

Tacha Serif, Gheorghita Ghinea

School of Information Systems, Computing and Mathematics
Brunel University.
Uxbridge, UK
email {tacha.serif, george.ghinea}@brunel.ac.uk

Anders Fongen

Norwegian School of Information Technology.
Oslo, Norway
email andfon@nith.no

Abstract- In this paper, we describe an investigation exploring user experiences of accessing streamed multimedia content, when that content is tailored according to perceptual, device and location characteristics. To this end, we have created pre-defined transmission profiles and stream perceptually tailored multimedia content to three different locations, each characterised by different infotainment requirements. In the light of our results, we propose that multimedia transmission to mobile and wireless devices should be made based on pre-defined profiles, which contains a combination of static (perceptual, device type, CPU speed, and display specifications) and dynamic information (streamed content type location of the device/user, context of the device/user). The evaluation of such a system showed that the users and service providers can gain from such an approach considerably, as user perceptions of quality were not detrimentally affected by QoS degradations. Consequently, service providers can utilise this information to effectively manage local network traffic and bandwidth.

I. INTRODUCTION

As computing environments become mobile and information access on the move commonplace, computing research has extended to cover contextual information such as a user's location, situation and circumstances. Although multimedia has been previously tailored according to user location [1];[2] and device [3];[4];[5], such tailoring until now has ignored perceptual quality considerations, even though, as we have shown in previous work [6] bandwidth – a scarce resource in a ubiquitous computing environment – could be more efficiently utilised this way. Nonetheless, to the best of our knowledge, no research has explored user multimedia perception and experiences when supported by a GPS-based, location aware, mobile guide – the focus of our study, described in this paper.

II. PERCEPTUAL-BASED MULTIMEDIA PROFILING

The main aim of our study was to evaluate the user experience of accessing streamed multimedia content tailored according to perceptual requirements as well as location and device type. It seems natural to include perceptual-based considerations in the wireless transmission of multimedia, for user take-up will be reduced, if the perceived quality is deemed to be low.

Recognising that all multimedia content is situated somewhere along an *infotainment* continuum, we have devised the Quality of Perception (QoP) metric [8], which assesses subjective multimedia quality and has two components, both

expressed as percentage measures: QoP-IA (information assimilation), representing informational analysis and synthesis, and QoP-S (satisfaction), representing subjective satisfaction with multimedia content. The impact of a multimedia clip on the user experience is then expressed as:

$$\text{Overall } QoP_{\text{clip}} = \text{Information}_{\text{clip}} \times QoP\text{-IA} + \text{Entertainment}_{\text{clip}} \times QoP\text{-LoE} \quad (1)$$

where

$$\text{Information}_{\text{clip}} = \overline{D} \times W_{I,D} + \overline{A} \times W_{I,A} + \overline{V} \times W_{I,V} + \overline{T} \times W_{I,T} \quad (2)$$

$$\text{Entertainment}_{\text{clip}} = \overline{D} \times W_{E,D} + \overline{A} \times W_{E,A} + \overline{V} \times W_{E,V} + \overline{T} \times W_{E,T} \quad (3)$$

Here, \overline{D} , \overline{A} , \overline{V} and \overline{T} stand for the average dynamism, audio, video and textual content respectively of the clip in question, as determined through subjective tests [7], while $W_{I,X}$ and $W_{E,X}$ reflect the relative importance of multimedia component X for information and entertainment, respectively.

We have taken the view that clips which are highly dynamic tend to imply that they contain predominantly entertainment material, as extracting informational content out of highly dynamic clips has proved to be problematic, whilst fast-paced, dynamic content usually has a high entertainment factor [8]. At the other end of the spectrum, it is relatively unlikely that the appearance of text in content watched predominantly for entertainment purposes adds to the entertainment factor of the clip. Of course, textual information is usually important in content watched mainly for informational purposes. Thus, in our work we have chosen the weights $W_{E,D}=3$, $W_{E,T} = 1$ for entertainment purposes, with $W_{I,D}=1$, $W_{I,T} = 3$ for informational purposes. Previous research [11];[12] has shown the importance of audio cues over video ones, in informational tasks, hence $W_{I,A} = 2$ and $W_{I,V} = 1$. On the other hand, true to the dictum 'that a picture is worth a thousand words', we have decided that the video stream has primacy over the audio, when content is viewed for entertainment purposes. Thus, we have $W_{E,V} = 2$, $W_{E,A} = 1$.

In previous work of ours[9] [10], we have determined frame rates which lead to optimal QoP values, for a wide variety of

multimedia infotainment content, and it is these values which make up the perceptual transmission profiles used in our study.

III. CAMPUS GUIDE APPLICATION

In our study, we developed a Campus Guide application which the user ran on a WiFi and GPS-enabled device. Initially the application connects to the content server via the university's wireless Internet access points and downloads a brief description about the buildings within the campus and

their locations, using the HTTP. Following this, it activates a GPS connection with a satellite and reads the current geographic coordinates of the user. Based on the location data sent by the satellite, the system provides the user with textual information about the surrounding buildings or areas (Fig. 1). Whenever the user is around a building recognised by the system, the application updates its interface and enables the audio-visual streaming for the specific location. Upon

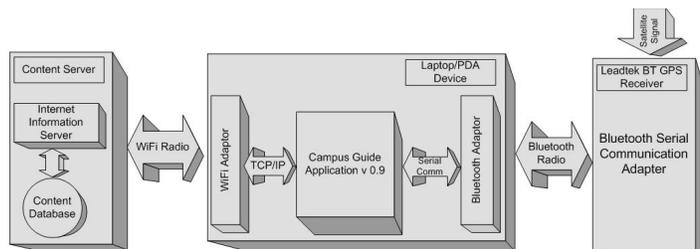


Figure 1: Application structure diagram

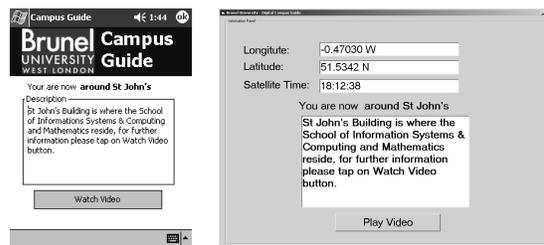


Figure 2: Campus Guide application interface (PDA and HMD)



Information Intensive (59,41) Clip



Infotainment (50,50)Clip



Entertainment Intensive (39,61)Clip

Figure 3: Video clips used in our experiments – in brackets, the (*Information, Entertainment*) breakdown

receiving the streaming request, the content server parses the parameters and identifies the type of device and the location of the user. Accordingly, the server selects suitable content using a knowledge base of pre-defined perceptual transmission profiles, which resides in its backend database, and streams the multimedia content to the user.

IV. EXPERIMENTAL STUDY

Two experimental variables were manipulated in our study – these were type of device and user location. Accordingly, three types of display devices were considered in our experiments (representing varying levels of user mobility).

A. Experimental Devices

A laptop, a Personal Digital Assistant (PDA) and a head-mounted display were used in our study. The laptop PC used in our experiments was selected from the inventory of the School of Information Systems, Computing and Mathematics. The selection was made among a Toshiba Satellite P35-S611, a considerably heavy HP Compaq NX9020s and an IBM T20s with a low processing and memory specifications. Accordingly, we picked a Toshiba Satellite P35-S611 Notebook PC, which ran Microsoft Windows XP Professional, and had a 3.33GHz Pentium IV processor, 512MB memory,

100GB hard drive, 17-inch XGA TFT Active Matrix Display and 64MB graphics adapter memory.

The HP iPAQ 5450 PDA has 16-bit touch sensitive transfective thin film translator (TFT) liquid crystal display (LCD) that supports 65,536 colours. The display pixel pitch of the device is 0.24 mm and its viewable image size is 2.26 inch wide and 3.02 inch tall. It runs the Microsoft Windows for Pocket PC 2003 operating system on an Intel 400Mhz XSCALE processor and allows the user complete mobility. Furthermore, it comes with a full wireless connectivity suite, which is made of WiFi (802.11b/g, Bluetooth and Infrared). By default it contains 64MB standard memory and 48MB internal flash ROM. Additionally in the course of some of our experiments a 128 MB secure digital memory card was used for multimedia data storage purposes

The Olympus Eye-Trek FMD 200 head-mounted display uses two 0.55-inch LCD displays and simulates a display view that is equivalent to a 52-inch screen at a distance of 2 meters. Each one of the displays contains 180,000 pixels and the viewing angle is 30.0° horizontal, 27.0° vertical. It supports Phase Alternating Line (PAL) format and has a display weight of 85g.

B. Experimental Locations

Three different locations for multimedia access were selected: an information intensive one, an entertainment intensive one, and one which users accessed an equal mix of information and entertainment. In our study, these were given by the Departmental Building (St John's), the University Library (Bannerman Centre), and the university lawn, respectively.

C. Participants

36 users, of average computing ability, took part in the study. These were split into three equal-sized groups, corresponding to the particular device used, with each user employing that device to access streamed multimedia content in each of the three locations considered in our study.

Participants were provided with a one-page Campus Guide experiment how-to manual, a QoP questionnaire and the guide device. In the manual participants were also provided with an approximate route to follow to reach their targeted destinations, thus ensuring that they went through all the key locations where streamed content was provided. At each such location, after requesting and watching the streamed video using their specific devices (Fig. 2), participants then answered the QoP questions (in terms of information assimilation and enjoyment) for each of the location-specific clips (Fig. 3). The whole process took on average 49 minutes per participant.

V. RESULTS SUMMARY

For brevity, we present the main results obtained in our study. These highlighted that device type had no significant impact on either user QoP-IA or QoP-S (Figures 4 and 5).

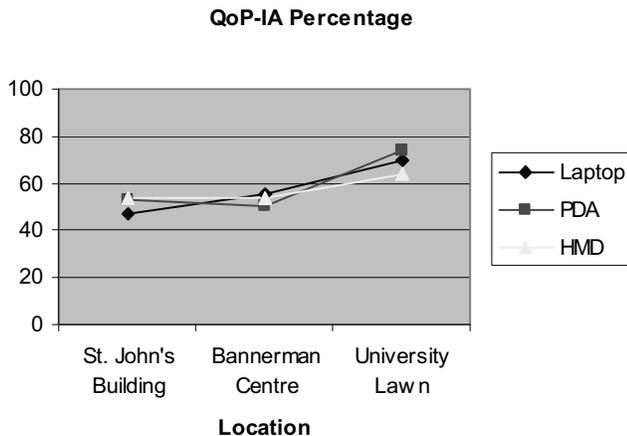


Figure 4. Overall quality of perception information assimilation (QoP-IA) based on device type

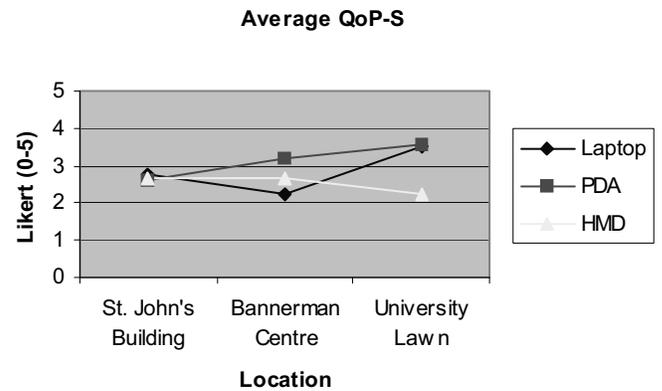


Figure 5. Average perceived satisfaction based on device type

This therefore shows that the user experience of using small screened devices is not necessarily detrimentally affected, if the content is appropriately tailored according to perceptual-location and technical requirements. On the other hand, when comparing the results obtained from the study described in this paper, with that of a control group, in which no tailoring took place [10], analysis revealed statistically significant influences of experiment type and device type on user overall QoP (QoP was higher in the tailored scenario), thus highlighting the benefits of streaming perceptually tailored multimedia content.

VI. CONCLUSION

In this paper we have presented the results of a study which tailored multimedia content to both location and device characteristics, when these are matched through the prism of perceptual quality requirements. Results thus far have highlighted the benefits of this approach and have shown that better network resource usage could result when performing such matching in practice.

REFERENCES

- [1] Abowd, D. A., Atkeson, C. G., Hong, J., Long, S. and Pinkerton, M., "Cyberguide: A Mobile Context-Aware Tour Guide", *Wireless Networks*, Vol. 3, No 5, pp. 421-433 (1996).
- [2] Burigat, S. and Chittaro, L., "Location-aware visualization of VRML models in GPS-based mobile guides", in *Proceedings of the tenth international conference on 3D Web technology*, pp. 57-64 (2005).
- [3] Buyukkoken, O., Garcia-Molina, H., Paepcke, A., Winograd, T., "Power Browser: Efficient Web Browsing for PDAs", *Digital Libraries Lab (InfoLab), Stanford University*, in *Proceedings of ACM CHI 2000 The Hague, Amsterdam*, pp. 430-437 (2000).
- [4] Fox, A., Goldberg, I., Gribble, S.D., Lee, D.C., Polito, A. and Brewer, E.A., "Experience With Top Gun Wingman: A proxy-based Graphical Web Browser for the 3Com PalmPilot", in *Proceedings of Middleware '98, Lake District, England* (1998).
- [5] Freire, J., Kumar, B. and Liewen, D., "WebViews: Accessing Personalized Web Content and Services", in *Proceedings of the 10th International Conference on World Wide Web*, pp. 576-586, (2001).
- [6] Ghinea, G. and Thomas, J., "Improving Perceptual Multimedia Quality with an Adaptable Communication Protocol", *Journal of Computing and Information Technology* 13 (2), pp. 149-161, (2005).
- [7] Ghinea, G., "Quality of Perception – An Essential Facet of Multimedia Communications", PhD Thesis, Department of Computer Science, The University of Reading, UK (2000).

- [8] Ghinea, G., and Thomas, J.P., "QoS Impact on User Perception and Understanding of multimedia Video Clips", in Proceedings of ACM Multimedia '98, Bristol UK, pp. 49- 54, (1998).
- [9] Ghinea, G., and Thomas, J.P., "Quality of Perception: User Quality of Service in Multimedia Presentations", IEEE Transactions on Multimedia, 7(4), 786-789, 2005.
- [10] Gulliver, S.R., Serif, T. and Ghinea, G., "Pervasive and Standalone Computing: The Perceptual Effects of Variable Multimedia Quality", International Journal of Human Computer Studies 60, pp. 640-665 (2004).
- [11] Kawalek, J. A., "User perspective for QoS Management", in Proceedings of the QoS Workshop aligned with the 3rd International Conference on Intelligence in Broadband Services and Network (IS&N 95), Crete, Greece (1995).
- [12] Watson, A., and Sasse, M.A., "Multimedia conferencing via multicasting: Determining the quality of service required by the end user", in Proceedings of AVSPN '97, Aberdeen, Scotland, pp. 189-194, (1997).