

Pervasive and Standalone Computing: the Perceptual Effects of Variable Multimedia Quality

S. R. Gulliver, T. Serif and G. Ghinea

Department of Information Systems and Computing,
Brunel University,
Uxbridge, Middlesex,
UB8 3PH, UK

E-mail: {Stephen.Gulliver / Tacha.Serif / George.Ghinea}@brunel.ac.uk

Tel: +44 1895 27400 ex 4862 / 2729 / 3635, Fax: +44 1895 251686

ABSTRACT

The introduction of multimedia on pervasive and mobile communication devices raises a number of perceptual quality issues, however, limited work has been done examining the 3-way interaction between use of equipment, quality of perception and quality of service. Our work measures levels of informational transfer (objective) and user satisfaction (subjective) when users are presented with multimedia video clips at three different frame rates, using four different display devices, simulating variation in participant mobility. Our results will show that variation in frame-rate does not impact a user's level of information assimilation, however, does impact a users' perception of multimedia video 'quality'. Additionally, increased visual immersion can be used to increase transfer of video information, but can negatively affect the users' perception of 'quality'. Finally, we illustrate the significant affect of clip-content on the transfer of video, audio and textual information, placing into doubt the use of purely objective quality definitions when considering multimedia presentations.

KEYWORDS

Perception, Head Mounted Display, Personal Digital Assistant, Eye-tracking

1. Introduction

The inclusion of multimedia capabilities in pervasive and mobile communications devices is a feature that, whilst increasing their allure, raises new challenges. Of these, the provision of good quality multimedia, both from a technical and user perspective will be paramount for the take-up and spread of multimedia-enabled pervasive computing. Indeed, we are of the opinion that it is the person and not the machine or the underlying technology which is the ultimate determinant of quality: if an application is perceived to deliver low quality, users will rarely be convinced to pay for the privilege of using it, irrespective of its intrinsic appeal.

Multimedia quality issues are thus important factors to be considered in pervasive computing, none more so if one considers that such environments have traditionally been characterised by their low-bandwidth availability, a feature which goes against the considerable bandwidth requirements that multimedia applications themselves necessitate. Indeed, the proliferation of a variety of pervasive computing devices has brought in a whole new dimension to the quality arena, for not only must multimedia be scaled and appropriately delivered to these devices, but, we contend, perception of multimedia quality may indeed differ across the devices themselves, as a result of user mobility.

Quality, in our perspective, has two main facets in a pervasive multimedia environment: *of service* and *of perception*. The former illustrates the technical side of computer networking and represents the performance properties that the underlying network is able to provide. The latter characterises the perceptual experience of the user when interacting with multimedia applications and forms the focus of this paper, which examines the three-way interaction between the multimedia quality of service provided by the network and its user-centric quality of perception, when mediated across different access devices, with varying level of mobility.

Accordingly, the structure of this paper is as follows: in section 2, we provide a brief introduction to the output device types being considered in this paper, allowing the reader an understanding of the research that has been done, especially relating to the area of multimedia perception. In section 3, the perceptual metric implemented in our empirical study (Quality of

Perception) is introduced allowing the reader to understand why quality of perception is used in our work. Section 4 describes the empirical study undertaken as part of our research, while Section 5 presents the main results obtained. Finally, in Section 6, conclusions are drawn and avenues for future research based on our findings are proposed.

2. Experimental Display Devices

In this section the authors aim to provide the reader with a brief introduction to literature relating to the specific output display devices being considered in this paper, which allow us to consider the perceptual implication of varying mobility (see Table 1). These devices range from a fixed head-position eye-tracker, to a traditional desktop limited mobility monitor, to a head-mounted display (allowing greater autonomy of movement), through to a personal digital assistant, allowing full personal mobility.

Device	Eye Tracker	Generic Monitor	Eye Trek	Personal Digital Assistant
Mobility	Extremely limited mobility	Limited mobility due to the nature of desktop monitor	Provides mobility, yet gives restricted vision and requires supporting equipment.	Causes no mobility restriction. Can be used on the go
Intrusiveness	Great degree of intrusiveness due to chin clamp device	Limited intrusiveness, users have to stand still for the best performance	Slightly restricts the body movements due to the size of the equipment (Battery, control pack, glasses and processor)	No intrusiveness, due to the ergonomic size and weight of the latest personal digital assistants

Table1: Varying mobility of used output display devices.

2.1 Eye-Tracking

The eye naturally fixates on areas that are most likely to be informative (Kaufman, L., Richards, W., 1969). Therefore, monitoring eye movements offers insights into visual perception, as well as the associated attention mechanisms and cognitive processes, and is therefore a logical way of determining factors that affect user perceptual processes.

The process of visual attention itself can be broken into two sequential stages: the pre-attentive stage and the limited-capacity stage (Hoffman, 1978) (Tresman, 1986). The pre-attentive stage of vision subconsciously defines objects from visual primitives, such as lines, curvature, orientation, colour and motion (Salapatel and Kessen, 1966) and allows definition of objects in the visual field. When items pass from the pre-attentive stage to the limited-capacity stage, these items are considered as selected (Glenstrup. and Engell-Nieson, 1995).

Interpretation of eye movement data is based on the empirically validated assumption that when a person is performing a cognitive task, while watching a display, the location of his/her gaze corresponds to the symbol currently being processed in working memory (Just and Carpenter, 1976) and, moreover, that the eye naturally focuses on areas that are most likely to be informative (Mackworth and Bruner, 1970).

There are several approaches to sensing eye movements (Young and Sheena, 1975) including: the use of a camera (Razdan and Kielar, 1998) (Hutchinson, 1993) or imaging systems to visually track some feature of the eye and then a computer or some neural network (Pomerleau, and Baluja, 1993) to do the reverse geometry to determine where the user is looking; the use of contact lenses, either combined with a camera- or magnetic-based feedback system; or alternatively the use of electrodes next to the eyes that sense the electro-oculographic potential (Gips, 1996). As a result, eye-tracking equipment varies considerably depending on the technique being implemented.

Eye tracking is increasingly being used as a tool for obtaining information about human perceptive and cognitive processes (Kowler, 1990) (Pelz, et al., 2000), as it is based on the empirically-validated assumption that the eye naturally centres on areas that are most likely to be informative. Thus, Mackworth and Bruner (Mackworth and Bruner, 1970) studied the eye movement of participants whilst looking at blurred pictures. The visual area was divided into 64 squares, each with an informative weighting. The most informative areas attracted more fixations (Mackworth and Bruner, 1970) (Mackworth, and Morandi, 1967). Mackworth and Morandi identified that informative areas are identified within the first two seconds of observation, a conclusion that has been reported in other studies of eye movement (De groot,

1966) (Yarbus, 1967). Moreover, eye tracking is being employed as an input device in the design of user interfaces, as an efficient interface ensures, for instance, that commonly-used controls are located in areas where the eyes' gaze is most likely to rest (Parkhurst et al., 2000), and that eye movement between these controls is minimal. Additionally, use of eye-based interfaces as a input device also help users (especially disabled) to execute interface input (Majaranta and Rähä, 2002).

Eye-tracking systems can be combined with purely display devices or can provide the user with interactive functionality (Reingold and Loschky, 2002) (Isokoski, 2000). Depending on the equipment, eye-tracking devices can be considered as either intrusive or non-intrusive in nature (Goldberg et al., 2002) and can be developed as either pervasive (Sodhi et al., 2002) or standalone systems. Level of immersion, whilst using eye-tracking equipment, may be high (Hayhoe et al., 2002) or low (Partala et al., 2000), depending on the specific equipment type.

To Allow the perceptual impact of devices of varying mobility to be compared, our study requires a non-mobile / intrusive display device. A camera based eye-tracker, with fixed head position, has been chosen to fulfil this task.

2.2 Head Mounted Display

Head mounted displays are a sub-set of wearable computer technology, which aim to allow hands free access to computer functionality. The users reason for hands free access to computing devices is often varied, and ranges from individuals with a restrictive physical disability (Gips et al., 1996) to individuals working in dangerous or hazardous conditions (Xybernaut, 2003). Integration of wearable mobile devices with network technology, touch pen, speech recognition inputs, interactive glove or face mounted devices (as in the case of Xybernaut's Mobile Assistant) allow extremely adaptable mobile solutions. Devices, such as head mounted displays, have often been considered synonymous with virtual reality development, however, due to falling cost and improved technology, head-mounted displays devices are becoming more commercially available and have recently gained commercial importance for high street companies such as Olympus and Sony.

The head-mounted display is made of two canonical displays, and usually consists of two liquid crystal or cathode-ray tube display screens that are either mounted on a helmet or glasses-frame structure.

There are several attributes that affect the usability of the head-mounted displays. Head-mounted displays can be either binocular, showing the same image to both eyes, or stereoscopic in nature, showing different images to each eye. The choice between binocular or stereoscopic depends on whether three-dimensional interaction or presentation is required. Head-mounted displays use a range of display resolutions. It is important to note, however, that a trade off exists between the resolution used and the field of view, which in turn impacts the perceived level of experienced immersion. A low field of view decreases the experienced level of user immersion, yet a higher field of view involves spreading the available pixels, which can cause distortion on the picture. Finally, ergonomic and usability factors vary considerably between different devices. Issues such as display size, weight and adjustability of physical and visual settings all affect the usability of a particular head-mounted displays for any specific task (Bowman et al., 2002).

Although there is now a wide range of head-mounted displays, there are several drawbacks that prevent their everyday popularity. The current lack of available media that properly facilitates immersive technology, for example: full-motion immersive video imaging emerged during the last ten years (Bohannon, 1997) yet is still not commercially available. The current high cost of the head-mounted displays that display both high resolution and wide field of view is a major factor. The large and encumbering size is an important factor for the users of especially cathode-ray tube based displays (Lantz, 1997). The visual limitation within the real world and reduced interactions with colleagues are also possible reasons that prevent head-mounted displays from regular everyday popularity. Other factors, such as hygiene and weight, also have possible unknown long-term medical implications on the supporting muscles and even on the eyes.

Despite the computational costs and usability drawbacks of the head-mounted displays, they are used widely in active research from virtual environment to wearable Internet applications.

The Smart Spaces (Pablo Research Group, 2002) project promises to implement anywhere / anytime automatic customisable, dynamically adaptable collaboration tools. In order to achieve these goals head-mounted displays, smart spaces, augmented reality and ubiquitous information access devices are being used. The main driving force of this research is information access anytime / anywhere, whilst doing something else.

Hitachi also is involved in the implementation of wearable Internet appliances targeting both industrial users and consumer users. The aim of the WIA (Wearable Internet Appliance) project (Ebina et al., 2002) is to provide mobile Internet and resource access using head mounted displays. Industrial users can facilitate this device to communicate between colleagues as well as access company databases and other centrally stored information (diagrams, equipment explanations) relating to their work. On the other hand, the consumer users' uses these devices to remotely access the Internet, even in the most crowded public location.

Bowman looked at the performance of head-mounted displays, for virtual environment (Bowman et al., 2002). Unlike our work, this research only covers the performance of three dimensional environment displays. In our case, we look at the perception and satisfaction performances of a binocular head-mounted display device. A head-mounted display was chosen for our study as it proves limited mobility, as the user is restricted by limited vision and cumbersome equipment, yet provides a greater autonomy of movement than traditional monitor display devices.

2.3 Personal Digital Assistant

Improvements in technology, especially in the wireless networking, have pushed the barriers of anywhere / anytime information access. Portable information access raises the need for portable information access devices, such as communicator devices and personal digital assistants, which promise to supplant the desktop computer as ubiquitous technology on campuses and in business (Weiser, 1998) - Gartner Research predicts a 260% increase in the unit sales from 9.39 million units in 2000, to 33.7 million units in 2004 (Kim and Albers, 2001).

Personal digital assistants represent a new technology that is still evolving and, as mobile devices, inherit many of their problems related to the distributed systems and mobile computing (Satyanarayanan, 2001). Distributed system problems include:

- *Remote tolerance*, such as protocol layering, the use of timeouts, remote procedure calls (Birrell and Nelson, 1984).
- *Fault tolerance*, such as atomic transactions, distributed nested transactions (Gray and Reuter, 1993).
- *Remote Information access*, such as caching, distributed file systems and databases (Satyanarayanan, 1989).
- *Security*, related encryption-such as mutual authentication and privacy (Needham et al., 1978).

Additional to the problems of distributed systems, personal digital assistants also suffer from issues relating to mobile computing devices. (Satyanarayanan, 2001). These include:

- *Mobile networking*, such as mobile IP (Bhagwat et al., 1996), ad hoc protocols (Royer and Toh, 1999).
- *Mobile information access*, such as disconnected operation (Kistler, and Satyanareyanan, 1992).
- *Support for adaptive applications*, such as transcoding by proxies (Fox et al., 1996).
- *Location sensitivity*, such as location sensing and location-aware system behaviour (Ward et al., 1997).

Personal digital assistants also inherit human-computer interaction and ergonomic related issues, such as small screen size, slow input facilities, low bandwidth, small storage capacity, limited battery lifetime and slow computer processor unit speed, which are all possible obstacles to the success of mobile and pervasive computing objectives (Buyukkokten et al., 2000) (Fulk, 2001) (Fox et al., 1998). For example: Jones et al studied the effect that screen sizes have on web-browsing related tasks (Jones et al., 1999). Their results showed that users with small screens followed hyper-links less frequently than the users with a larger display unit.

The increasing popularity and the above accumulation of problems have made the personal digital assistants a popular area of research. The Power Browser (Buyukkokten, et al., 2000) was created to provide easy navigation in complex web sites using small screen mobile devices, such as a personal digital assistant. This application uses a hypertext transfer protocol proxy that receives the requests from the mobile user and, based on the request fetches of the user, dynamically generates a summary view to be transmitted back to the client. These summary web pages contain both link structure and contents of a set of web pages being accessed.

Top Gun Wingman (Fox et al., 1998) is another transcoder targeted for Palm operating system. Although similar to the Power Browser, this application does not only provide ease of navigation but it converts the pages, images, and files (Zip / PalmDoc) to browser specific suitable format.

The company TV-Anytime (TV-Anytime, 2003) has applications that allow users to access their profiles remotely with a personal digital assistant and wireless Internet access (Kazasis et al., 2003). Logged in users can search the online database for relevant television programmes, documentaries and movies and download to their home appliances. Using the application, the previews of the programmes can be watched online. Also, the user can set the length of the clip according to the network bandwidth and battery lifetime of the personal digital assistant.

Personal digital assistants represent, in our study, a truly mobile device, allowing the user full mobility of movement whilst viewing multimedia information.

2.4 Perceptual impact of multimedia ‘Quality of Service’

Considerable work has been done to investigate the importance of user perception of distributed multimedia quality. Accordingly, Apteker et al. (1995) examined the influence that varying video frame rates have on user satisfaction with multimedia quality and showed that the dependency between human receptivity and the required bandwidth of multimedia video clips is non-linear. Consequently, for certain ranges of human receptivity, a small

perceptual variation leads to a much larger relative variation of the required bandwidth – highlighting the potential for significant bandwidth savings if perceptual considerations are included in the transmission of multimedia.

Steinmetz (1995) presents the perceived effect of synchronisation skews between media, highlighting the relatively large perceptual tolerance here. Kawalek (1995), on the other hand, is more interested in the cut-off rate beyond which the quality of transmitted audio and video becomes unacceptable to human users in desktop conferencing environments: he showed that the perception of media loss is highly dependent on the medium in question. While Bouch et al (2000) have researched the effect of latency on perceived web quality of service, Wijesekera et al. (1999) build on Steinmetz's and Apteker's earlier work and investigate the perceptual tolerance to discontinuity caused by media losses and repetitions, and to that of varying degrees of mis-synchronisation across streams (Apteker et al., 1995).

2.4.1 Personal Digital Assistant

To the best of our knowledge there has been no specific work been done concerning the perceptual issues surrounding personal digital assistants, with the exception of (Elting et al., 2002) which explores the modality-combinations. This looks at the effects that different output modality-combinations have on the devices' effectiveness to transport information and on the user's acceptance of the system being used. It uses three devices personal digital assistant, television and desktop computer to investigate whether the best modality depends on the device. As test data, it uses a web based tourist guide that contains text and images. Their results showed that the most appealing form of information transfer combined picture, text and speech. However due to multi-modal cognitive load, especially when using a personal digital assistant, the most effective form of information transfer was shown to be whilst using just combined picture and speech.

2.4.2 Head Mounted Displays

A number of research studies exist looking at the symptoms related to head-mounted display usage, such as nausea (Regan, 1995), dizziness (Cobb et al., 1995), headaches (Kennedy et al., 1995) and eyestrain (Kolasinki, 1995). However, to the best of our knowledge there has

been no work done concerning perception and satisfaction of head-mounted display usage, with the exception of Geelhoed et al. at Hewlett-Packard Laboratories in Bristol (Geelhoed et al., 2000). This study investigates the comfort level of various tasks, such as text reading and video watching, on two different head-mounted displays using questionnaires on ten volunteers. As a result, it identifies that tasks requiring more long-term attention, such as watching video, causes a greater level of discomfort to the user.

2.4.3 Eye-Tracking

Eye movement data can be used to identify the informative / user selected area within a visual field (Mackworth, and Bruner, 1970). To this end, the relationship between eye movement and user perception of multimedia has been investigated in (Faraday and Sutcliffe, 1999) and (Parkhurst et al., 2000). The former study explores both visual attention (given by eye tracking patterns) and information recall of subjects being presented with a single multimedia educational application, displayed with optimum quality of service parameters. The authors then went on to propose a series of guidelines to be used in web animation based on ‘contact points’ (co-references between text and animation obtained from the initial eye tracking study) (Faraday and Sutcliffe, 1999). Thus, the first study only focuses on the informational assimilation component of ‘Quality of Perception’ (the user-centric perceptual measure used throughout this paper), neglecting the satisfaction side of the multimedia experience. Moreover, users were shown only one particular type of multimedia category with constant, optimum quality of service parameters, which fails to reflect the multitude of multimedia applications and the variety of prevailing network conditions that exist in the 3-way interaction between use of eye-tracking equipment, quality of perception and quality of service. The second study investigated the effect that multi-resolution displays have on user perception. The idea behind this research was to reduce the required resources by not presenting a uniform level of visual detail across the whole display area of a screen, but rather to render a high level of visual detail only around the centre of the user’s gaze. Whilst a spatial service of service parameter was indeed varied, based on user eye gaze location, it was felt that this study neglects the concept of information assimilation and the multimedia

diversity that quality of perception inherently possesses. Although both studies point to the fact that eye tracking results can result in enhancement of the user multimedia experience, to the best of our knowledge no-one has examined the 3-way interaction between use of eye tracking equipment, quality of perception and quality of service.

Our work measures levels of informational transfer (objective) and user satisfaction (subjective) when users are presented multimedia video clips using four different display devices. Multimedia files were presented to users at three different frame rates providing measurements over a variety of quality of service settings. In the following section we present quality of perception, the perceptual metric used in our empirical study.

3. QUALITY OF PERCEPTION

In a distributed setting, quality of digital multimedia has traditionally been measured using quality of service technical parameters, such as jitter, delay, as well as loss and error rates. Although measurable, such objective parameters disregard the user's perception of what defines multimedia quality (Scialfa et al., 2000) (Steinmetz and Nahrstedt, 1995). To date, there has been a common assumption in the computer networking community that many quality issues will be resolved through objective solutions, such as increased bandwidth allocation (Watson and Sasse, 1997). The majority of quality of service research has therefore been systems oriented, focusing on factors such as traffic analysis, scheduling and routing.

Due to the multi-dimensional nature of multimedia, it is therefore impossible to rely on objective factors alone when defining multimedia quality. Multimedia applications are produced for the enjoyment and/or education of human viewers, so their opinion of the presentation quality is important to any quality definition. Therefore, when evaluating multimedia quality, subjective testing by viewers must be considered in combination with objective testing.

In order to explore the human side of the multimedia experience, the authors have used the above notion of 'Quality of Perception'. This more closely reflects multimedia's infotainment (i.e. multimedia applications are located on the informational-entertainment spectrum)

characteristic as well as the fact that objective factors alone are incapable of defining the perceived quality of multimedia video (Ghinea and Thomas 1998) (Bouch et al., 2000) (Watson and Sasse, 1997). Quality of perception uses level of ‘information transfer’ (objective) and user ‘satisfaction’ (subjective) to determine the perceived level of multimedia quality. To this end, quality of perception is a term which encompasses not only a user's satisfaction with the quality of multimedia presentations (‘Satisfaction’), but also his or her ability to analyse, synthesise and assimilate the informational content of multimedia (‘Information Assimilation’).

4. EXPERIMENTAL METHOD

4.1 Participants

Our study involved 48 participants, who were aged between 18 and 56 and were taken from a range of different nationalities and backgrounds – students, clerical and academic staff, white collar workers, as well as a number of retired persons. All participants, however, spoke English as their first language, or to a degree-level qualification, and were computer literate.

4.2. Experimental Variables

Three experimental variables were manipulated in our study – these were type of device, multimedia video frame rate and multimedia content. Accordingly, four types of display devices were considered in our experiments (representing varying levels of user mobility), and three multimedia video frame rates: 5, 15 and 25 frames per second. We chose frame rate as the quality of service parameter of interest within this study, because it is the main factor affecting multimedia bandwidth requirements, of primordial importance, due to its scarcity, in distributed pervasive multimedia environments. As far as multimedia content is concerned, 12 video clips were considered in our experiments and these shall be further described in section 4.4.

4.3 Measuring Quality of Perception

To understand why use of quality of perception is essential in the context of this study, it is important that the reader understands how objective and subjective quality of perception factors were defined and measured.

4.3.1 Measuring Information Assimilation

In our approach, 'Information Assimilation' (level of objective information assimilated) was expressed as a percentage measure, which reflected a user's ability to assimilate specific information from visualised multimedia content. Thus, after watching a particular multimedia clip, the user was asked a standard number of questions (ten, in our case) which examined information being conveyed in the specific clip that had just been presented to the participant. The level of information assimilation was calculated as being the proportion of correct answers that users gave to these specific questions.

To allow objective measurement, all information assimilation questions must, of course, have a definite answer, for example: (from the Rugby video clip used in our experiment) "What was the score of the match at the beginning of the video clip? As this question has an unambiguous answer (England 0 / New Zealand 7), which is clearly presented in the video content of the multimedia clip, it was possible to determine whether the participant had assimilated this information by marking whether they were able to correctly answer the question. As this question is presented in the video content of the clip the participant is awarded a video mark for that specific question if, and only if, the question is answered correctly.

Since, in our experiments, questions can only be answered if certain information is assimilated from specific information sources (for example, the words of a song can only be gained from the audio stream), it is possible to determine the percentage of correctly answered questions that relate to the different information sources within specific multimedia video clips. For each feedback question, the source of the answer was determined as having been assimilated from one of the following information sources:

V : Information relating specifically to the video window, for example, pertaining to the activity of lions in a documentary clip.

A : Information which is presented in the audio stream.

T : Textual information contained in the video window, for example: information contained in a caption (for example: the newscaster's name).

Thus, by calculating the percentage of correctly absorbed information from different information sources, it is possible to determine and compare, over a range of different multimedia content, potential differences that might exist in information assimilation when participants are presented multimedia video clips using different output display devices with a varying level of mobility..

4.3.2 Measuring Subjective Quality of Perception (Satisfaction)

Satisfaction is subjective in nature and, in this study, consists of two component parts: 'Level of enjoyment' (the user's Level Of Enjoyment whilst viewing multimedia content) and 'Level of Quality' (the user's judgement concerning the objective Level Of Quality assigned to the multimedia content being visualised). The authors have used both user level of enjoyment and perceived level of quality to distinguish between a participant's subjective satisfaction concerning the content of the video clip and his/her ability to assess the objective quality of the video clip – both factors of satisfaction.

Measuring Level of Enjoyment

The subjective Level of Enjoyment experienced by a user when watching a multimedia presentation, was polled by asking users to express, on a six-point Likert scale of 0 - 5, how much they enjoyed the presentation (with scores of 0 and 5 representing "complete dislike" and, respectively, "absolute enjoyment" user satisfaction with the multimedia video presentation). A six-point Likert scale was used to ensure a mid-point decision, preventing a completely neutral subjective opinion. This information was used to determine whether a user's ability to assimilate information has any relation to level of enjoyment.

Measuring Level of Quality

In order to measure this, users were asked to indicate, on a six-point Likert scale of 0 - 5, how they judged, independent of the subject matter, the presentation quality of a particular piece of multimedia content they had just seen (with scores of 0 and 5 representing “completely terrible” and, respectively, “absolutely perfect” user satisfaction with the multimedia presentation quality). This information was used to determine whether a user’s ability to assimilate information has any relation to perceived level of video quality.

4.4. Experimental Material

All participants were presented with a series of 12 windowed MPEG-1 video clips, each between 31 and 45 seconds long. The multimedia video clips used in this experiment were chosen to cover the broad possible spectrum of infotainment. Multimedia video clips vary in nature from those that are informational in nature (such as a news / weather broadcast) to ones that are usually viewed purely for entertainment purposes (such as an action sequence, a cartoon, a music clip or a sports event, as detailed in Figure 1). Specific clips were chosen as a mixture of the two viewing goals, such as the cooking clip).

- **Band clip** - this shows a high school band playing a jazz tune against a background of multicoloured and changing lights.
- **Commercial clip** - an advertisement for a bathroom cleaner is being presented. The qualities of the product are praised in four ways - by the narrator, both audio and visually by the couple being shown in the commercial, and textually, through a slogan display.
- **Chorus clip** - this clip presents a chorus comprising 11 members performing mediaeval Latin music. A digital watermark bearing the name of the TV channel is subtly embedded in the image throughout the recording.
- **Cooking clip** - although largely static, there is a wealth of culinary information being passed on to the viewer. This is done both through the dialogue being pursued and visually, through the presentation of ingredients being used in cooking of the meal.

- **Animation clip** - this clip features a disagreement between two main characters. Although dynamically limited, there are several subtle nuances in the clip, for example: the correspondence between the stormy weather and the argument.
- **Weather clip** - this is a clip about forthcoming weather in Europe and the United Kingdom. This information is presented through the three main channels possible: visually (through the use of weather maps), textually (information regarding envisaged temperatures, visibility in foggy areas) and by the oral presentation of the forecaster.
- **Documentary clip** - a feature on lions in India. Both audio and video streams are important, although there is no textual information present.
- **Pop clip** - is characterised by the unusual importance of the textual component, which details facts about the singer's life. From a visual viewpoint it is characterised by the fact that the clip was shot from a single camera position.
- **News clip** - contains two main stories. One of them is presented purely by verbal means, while the other has some supporting video footage. Rudimentary textual information (channel name, newscaster's name) is also displayed at various stages.
- **Rugby clip** - presents a test match between England and New Zealand. Essential textual information (the score) is displayed in the upper left corner of the screen. The main event captured is the score of a try. As is expected, the clip is characterised by great dynamism.
- **Snooker clip** - the lack of dynamism is in stark contrast to the Rugby clip. Textual information (the score and the names of the two players involved) clearly displayed on the screen.
- **Space clip** - this was an action scene from a popular science fiction series. As is common in such sequences it involves rapid scene changes, with accompanying visual effects (explosions).



Figure 1: Shows video frame 500, for the 12 video clips used in our experiment, demonstrating the diversity of multimedia being considered.

4.5 Experimental Quality of Perception Questions

In our experiments, video specific questions can only be answered if certain information is assimilated from specific information sources. To determine how many video, audio and textual questions should be used for each specific multimedia video clip, weightings were defined for each clip (Ghinea and Thomas 1998) and a distribution of questions was allocated (see Table 2).

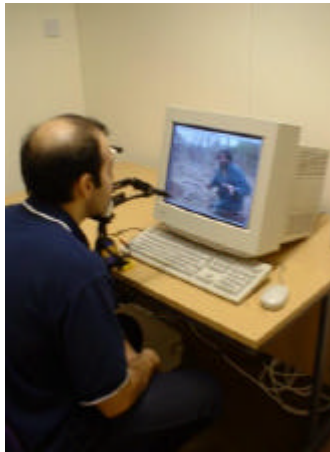
VIDEO CATEGORY	VIDEO (number of questions)	AUDIO (number of questions)	TEXT (number of questions)	DURATION (sec)
Band (Jazz Band)	9	1	0	36
Commercial	7	2	1	31
Chorus – Choir	9	0	1	33
Cooking	7	3	0	39
Animation	6	4	0	32
Weather	2	5	3	45
Documentary	9	1	0	31
Pop	5	2	3	36
News	2	6	2	38
Rugby	7	1	2	34
Snooker	6	2	2	33
Space	9	1	0	36

Table 2: Quality of perception question distribution - Video, Audio and Textual.

4.6 Experimental Set-up

To allow the perceptual comparison of different display equipment on a user's ability to assimilate information from multimedia video, the participants were evenly allocated to four different groups (1, 2, 3 & 4). Within each respective group, users were presented the video clips using a specific display equipment. Group 1 acted as a control group (standard mobility) and was therefore shown the video clips full screen using a normal 15 inch super video graphic array generic computer monitor enabled with a Matrox Rainbow Runner Video Card. Group 2 also viewed the video clips full screen using a computer monitor, however, the participant was simultaneously interacting with a Power Mac G3 (9.2) powered Arrington ViewPoint EyeTracker, used in combination with QuickClamp Hardware (figure 2) providing limited mobility. The ViewPoint EyeTracker allows an accuracy of approximately 0.5° - 1.0° visual arc and has a temporal resolution of 30hz. Group 3 viewed the multimedia video clips using an Olympus Eye-Trek FMD 200 head-mounted display, which uses two liquid crystal displays and allows a greater autonomy of movement than a generic computer monitor. Each one of the displays contains 180.000 pixels and the viewing angle is 30.0° horizontal, 27.0° vertical. It supports PAL (Phase Alternating Line) format and display weight is 85g (figure 3). Group 4 viewed the video clips using a Hewlett-Packard iPAQ 5450 personal digital assistant with 16-bit touch sensitive transfective thin film translator liquid crystal display that supports

65,536 colour. 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 is running Microsoft Windows for Pocket personal computer 2002 operating system on Intel 400Mhz XSCALE processor and allows the user complete mobility. By default it contains 64MB standard memory and 48MB internal flash read only memory. Additionally in the course of this experiment 128 MB secure digital memory card was used for multimedia data storage purposes (figure 4). A pilot test study of 2 participants was used to check and validate the output of all display devices. During this study, both test participants using the personal digital assistant commented that environmental noises interfered with the audio output. As we hoped to provide participants with a consistent audio level, headphones were used to limit interference from the surrounding environment.



a)



b)

Figure 2: a) Power Mac G3 (9.2) ViewPoint EyeTracker, used in combination with QuickClamp Hardware b)



Figure 3: Head Mounted Display Device



Figure 4: Personal Digital Assistant Device.

Participants viewed different video clips at one of three pre-recorded frame rates (5, 15 or 25 frames per second). Thus, each participant viewed four video clips at 5 frames per second, four video clips at 15 frames per second, and four video clips at 25 frames per second, as defined in Table 3. Although this mixed frame rate approach limits sample size it reduces the chance of participant effects in any one specific frame rate group, which allows us to more accurately assess the impact of frame rate on a user's ability to assimilate information from a specific device.

Video	Order 1	Order 2	Order 3
Band (Jazz Band)	25	5	15
Commercial	5	15	25
Chorus – Choir	15	5	25
Cooking	15	25	5
Animation	25	15	5
Weather	5	25	15
Documentary	5	15	25
Pop	15	25	5
News	5	25	15
Rugby	25	5	15
Snooker	15	5	25
Space	25	15	5

Table 3: Frame rate and video order presented to experimental groups.

4.7 Experimental Process

All participants, independent of the display device being used, were asked a number of short questions concerning their sight, which was followed by a basic eye-test to ensure that all participants were able to view menu text on the screen. This was specifically important for the

eye-tracking device as participants were not able to wear corrective spectacles or lenses for the duration of the experiment. Participants were informed that after each video clip they would be required to stop and answer a number of questions that related to the video clip that had just been presented to them. To ensure that the participants did not feel under test conditions, it was made clear that their intelligence was not being tested and that they should not be concerned if they were unable to answer any of the information assimilation questions. After introducing the participant to the experiment, the appropriate system software was loaded and the specific display device was setup for the user's requirements. In the case of the eye-tracker, time was taken to adjust the chin-rest, infrared red capture camera and software settings to ensure that pupil fix was maintained throughout the user's entire visual field.

When display calibration was complete, the participant was asked to get into a comfortable position and in the case of the eye-tracker place his/her chin on the chin-rest. The correct video order was loaded (see Table 3) and the first video was displayed.

After showing each video clip, the video window was closed and the participant was asked a number of quality of perception questions about the video that they had just seen. All questions were chosen in accordance with the quality of perception definition, as described earlier in this paper, to encompass both objective information assimilation and subjective level of enjoyment and level of quality. Once a user answered all questions relating to a specific video clip, and all responses had been noted, experimental setup was repeated and the next video clip was displayed. This process repeated for all 12 videos, independent of the display device.

5. RESULTS

5.1 Effect of Frame-Rate on User Quality of Perception

Level of Information Assimilation

Quality of service was varied in this study, by showing participants video clips at 3 different frame-rates (5, 15 and 25 frames per second). To check the effect of frame-rate on objective quality of perception factors, we used a one-way ANOVA (Analysis of Variance) test, with

information assimilation (Video), information assimilation (Audio) and information assimilation (Text) as dependent variables and frame-rate as the independent variable. This showed that video, audio and textual objective information assimilation was not significantly affected by variation of video frame rate.

This supports the previous finding of Ghinea and Thomas (1998), who found that a significant loss of frames (that is, a reduction in the frame rate) does not proportionally reduce the user's understanding of the multimedia presentation. Current distributed and mobile computing multimedia systems judge quality in terms of objective quality of service settings. The results of our work however suggest that a significant change of objective settings does not necessarily significantly affect users' ability to assimilate information from multimedia video content. This allows us to justify the use of a lower frame-rate, and therefore enabling a reduced bandwidth requirement, for multimedia video presentations, if and only if the level of information assimilation is the primary aim of the multimedia presentation.

Level of Satisfaction

The authors consider the users' satisfaction as essential to any quality definition. Indeed, we are of the opinion that it is the person and not the machine or the underlying technology which is the ultimate determinant of quality: if an application is perceived to deliver low quality, users will rarely be convinced to pay for the privilege of using it, irrespective of its intrinsic appeal. Therefore, we were interesting in looking at whether quality of service variation, in the form of frame-rate, had any significant impact on the end users' perceived level of enjoyment or their opinion concerning the level of objective video quality.

The findings show that although participant enjoyment was not significantly affected by variation in frame rate a users' perception of video quality was significantly affected between frame-rate groups $\{F(1,2) = 4.766, p=0.009\}$.

Our results suggest that, even though a user is aware of degradation in objective 'quality', his or her level of enjoyment is not significantly affected by the frame-rate variation. This allows us to justify the use of a lower frame-rate for multimedia video presentations, if significant

changes in user perceived level of quality are not considered as an important outcome of the multimedia presentation.

5.2 Effect of Device Type on User Quality of Perception

Level of Information Assimilation

Variation of device type was used in this study to identify whether any significant changes occur to user quality of perception, as a result of the device, and therefore the level of mobility, used to display the multimedia presentation (see Table 1). To check the effect of device type on information assimilation, we used a one-way Analysis of Variance test, with information assimilation (Video), information assimilation (Audio) and information assimilation (Text) as dependent variables and device type as the independent variable. As the homogeneity of variance proved that only video information assimilation results to be valid, a K-independence non-parametric test was used to test the significance of information assimilation (Audio) and information assimilation (Text). Results demonstrated that only level of video information assimilation was significantly affected by variation in display device $\{F(1,3) = 3.048, p=0.028\}$ and therefore level of user mobility. Information assimilation (Audio) and information assimilation (Text) was found not to be significantly affected by device type. Post Hoc Tukey tests showed that a significant difference occurred between the mean video information assimilation ratings (mean number of video questions answered correctly per video) of participants using eye-tracker and head-mounted display devices ($p= 0.018$). The head-mounted display and eye-tracker, with respectively mean video information assimilation ratings of 5.78 and 4.75, were identified as respectively the best and worst devices for user video information assimilation.

The authors believe that the reason for the difference is due to the level of immersion available to the user whilst using the two devices. The Olympus Eye-Trek head-mounted display is designed to simulate a 52-inch display monitor, thus proving a high level of user visual immersion. Head-mounted displays also allow full head movement without changing the relative position of the screen and the eye. In comparison, the Arrington ViewPoint

EyeTracker is used in combination with QuickClamp Hardware, which intrusively restricts the movement of the user's head. Although restricted head movement is vital to this specific eye-tracker device to map and interpret eye-gaze location, it is intrusive and far from conducive to user immersion. Additional factors, such as a smaller perceived display screen (15inch generic monitor), as well as the users' conscious awareness of the eye-tracker device are all possible factors that reduce participant visual immersion.

Level of Satisfaction

To check the effect that device type has on satisfaction, we used a one-way Analysis of Variance test, with level of enjoyment and level of quality as dependent variables and device type as the independent variable. A homogeneity of variances test showed that only level of enjoyment results were considered valid, therefore a K-independence non-parametric test was used to check the affect of device type on level of quality.

No significant difference was measured between participants' level of enjoyment, independent of the device type, which suggests that, although device type may facilitate a significant improvement of video information assimilation, variation in level of enjoyment is due to factors other than the display device, and therefore also the level of mobility, used to present the multimedia video presentation.

Non-parametric analysis showed that device type significantly affected a users' level of quality $\{\chi^2(3, N = 576) = 11.578, p = .009\}$, with a significant difference between mean level of quality score for control (3.05) and head-mounted display (2.63) participant groups. Despite facilitating the greatest video information assimilation, users' perceive the head mounted displays as displaying the worst level of quality video – see figure 5.

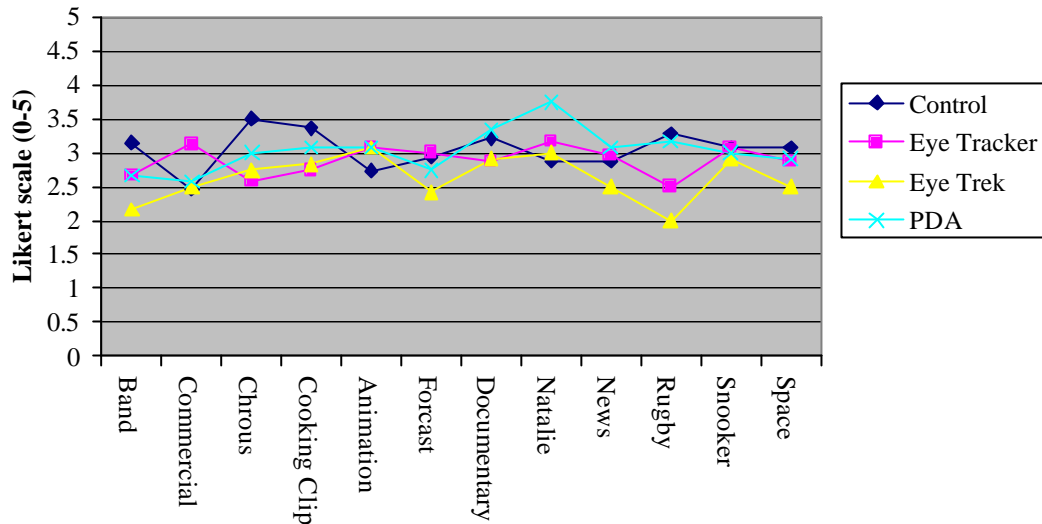


Figure 5: Average perceived quality, for all frame rates, across all video clips.

The authors believe that a significantly lower perceived level of quality may be due to one of two specific issues. The first proposed reason why head mounted displays cause a reduction in the perceived level of quality is as a result of increased level of video immersion. As previously stated, a trade off exists between the resolution used and the field of view. A low field of view decreases the experienced level of user immersion, yet a higher field of view involves spreading the available pixels, which can cause distortion on the picture. The Olympus Eye-Trek head-mounted display is designed to simulate a 52-inch display monitor, providing the user with a high level of video immersion. As the same video clips were used for all devices, the authors suggest that pixel distortion as a result of a higher field of view was perceived as being of bad 'quality'. It is interesting to note that users viewing exactly the same videos on the 2.26 x 3.02 inch personal digital assistant screen, perceived them to be comparatively higher 'quality'. The second proposed reason why head mounted displays cause a reduction in the perceived level of quality is due to physical discomfort. Geelhoed (2000) showed that, whilst using a head-mounted display, tasks requiring more long-term attention, such as watching a video, causes a greater level of discomfort to the user. Irrespective of the reason, the reduction of perceived quality has interesting implications on the future use of head mounted displays.

5.3 Effect of Video Clip Type on User Quality of Perception

Level of Information Assimilation

Clip type obviously affects the quantity of specific questions being used in our experiment (see Table 2), for example: the band clip has no textual content, therefore no textual feedback questions were used. However, when results were displayed as a percentage measure of the number of questions being asked, as defined in the quality of perception definition (see section 4.3), considerable variation was still observed in the level of video, audio and textual information assimilation between the different clips used throughout this experiment.

To analyse this variance non-parametric kruskal-wallis tests were used to check whether clip type significantly impacted video, audio and textual information assimilation. Results showed that the type of clip, independent of frame rate, display type and level of mobility, significantly affects information assimilated from video $\{\chi^2(11, N = 576) = 287.833, p < .005\}$, audio $\{\chi^2(11, N = 576) = 413.210, p < .005\}$ and textual $\{\chi^2(11, N=576) = 427.643, p < .005\}$ content of a multimedia video – see figure 6.

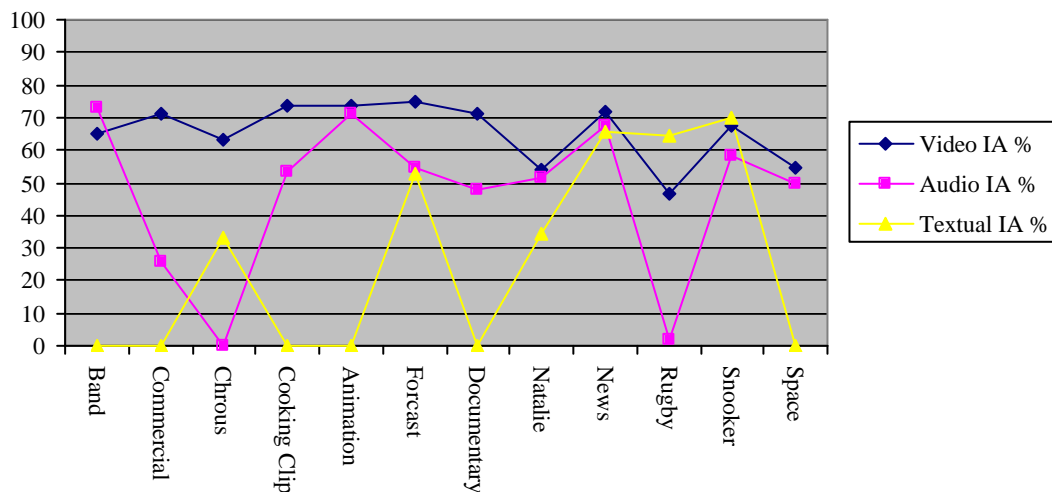


Figure 6: Average video, audio and textual information assimilation, independent of device type and frame-rate, for all video clips.

Our study shows that the video clip type, used as part of a multimedia presentation, has more of a significant affect on a users' level of information transfer than either the frame rate, display device type and therefore the level of mobility being used to present the multimedia presentation. The significant impact of the contents of the clip could be for a number of reasons: user preference leading to greater level of maintained attention, cultural level of pre-knowledge, clearer transfer of relevant information or even capable cognitive load. Although further work is required to determine the relationship between clip contents and level of information assimilation, the authors believe that this result justifies our view that when considering 'multimedia quality', especially in the pervasive multimedia environment, we must consider two main facets: of service and of perception. Quality of digital multimedia has traditionally been measured using quality of service technical parameters, however, our work has implications on using purely objective testing when defining multimedia quality.

Level of Satisfaction

Our study has shown that clip-type has significant implications on user level of information transfer. To identify the impact of clip type on user satisfaction, a one-way ANOVA was used, with perceived level of quality and level of enjoyment as the dependent variables and clip type as the independent variable. This showed that although no significant variation occurred in the perceived level of quality as a result of clip type, a significant variation did occur in the user level of enjoyment $\{F(1,11) = 9.676, p < 0.005\}$ – see figure 7.

As each specific participant has his / her own viewing preference, it seems obvious that the video clip type significantly affects a users' level of enjoyment, however, the fact that the type of video clip does not significantly affect perceived level of quality implies that users' perception of specific video clip 'quality' is not dependant upon the contents of the clip. It is interesting to note that a user's perception of quality is significantly affected by variations in both frame rate, display device and therefore level of mobility, yet is not significantly affected by the video content. It appears from this result that users are able to distinguish between their subjective appreciation of a video clip (enjoyment), and the level of 'quality' with which they

perceive the video clip to possess. This supports the use of quality of perception in context of this study.

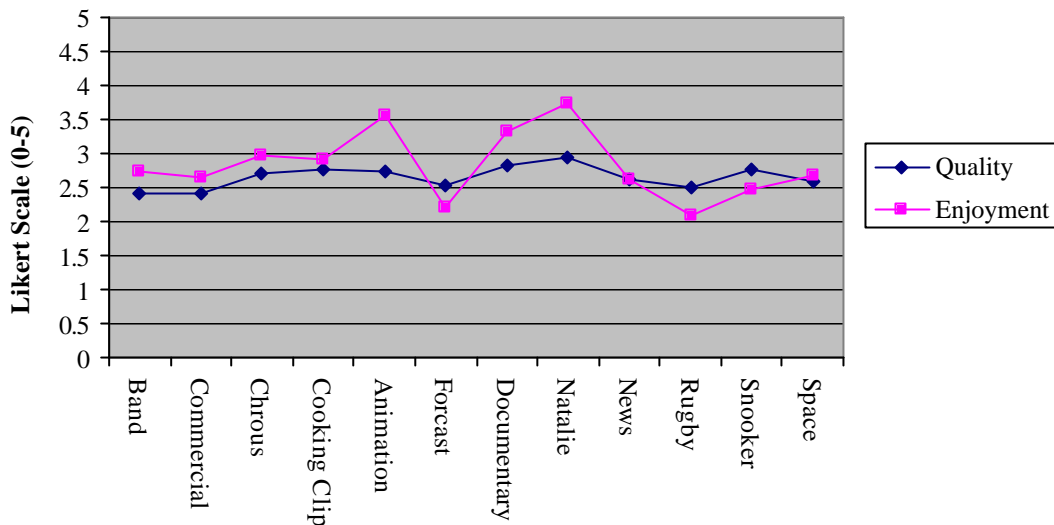


Figure 7: Average perceived level of Quality and Enjoyment, for all frame rates and devices, across all video clips.

6. Conclusion

How does variation in objective multimedia quality impact user perception of multimedia presentations? Can a difference be identified between the perceptual effect of mobile and intrusive computing? Does clip content impact user multimedia perception?

Quality of service was varied in this study, by showing participants video clips at 3 different frame-rates (5, 15 and 25 frames per second), yet despite a considerable loss of frames (that is, a reduction in the frame rate), level of information transfer, essential to a users' understanding of the multimedia presentation, does not significantly fluctuate between groups. This gives justification for a reduction in bandwidth allocation, if and only-if the level of information assimilation is the primary aim of the multimedia presentation. However, if perception of video 'quality' is essential to the outcome of the multimedia video presentation, manipulation of frame-rate should be prevented.

The authors propose that due to granularity the perceptual affect of different device types cannot be generalised by obvious division into defined groups, such as mobile and non-

mobile computing. The impact of device should therefore be considered individually. Results showed that device type has a significant impact on user video information assimilation, which we proposed to be as a result of increased level of user level of immersion. Although variation in device type has no significant impact on user level of enjoyment, head-mounted displays were found to have a significantly lower overall user perceived level of video quality, despite enabling the greatest level video information transfer. If a device is perceived to deliver low quality, despite it's ability to transfer video information, the authors believe that users will rarely be convinced to pay for the privilege of using it. This conclusion has possible implications on the future of fully immersive head-mounted display devices and may result in a slow-down of commercial acceptance. Further research is required to better understand why users perceived head-mounted displays to be of a lower relative quality than small screen devices, such as mobile personal digital assistant.

Our study also considered the impact that clip-contents has on the perception of multimedia video quality. Despite conversion of results, as a percentage measure of the number of questions being asked, results showed that clip type was found to have a more significant affect on a users' level of information transfer than either the frame rate, display device type and therefore level of mobility. Additionally, whilst a user's perception of quality is significantly affected by variations in frame rate and display device being used, it is not significantly affected by the content of the video, implying that users are able to distinguish between their subjective enjoyment of a video clip, and the level of 'quality' with which they perceive the video clip to possess. Further work is required to determine the relationship between clip contents and user quality of perception, however the authors believe that our conclusions justify the use of quality of perception as a means of defining 'multimedia quality' within this and future studies. Quality of digital multimedia has traditionally been measured using quality of service technical parameters, however, this threatens to disregard the user's perception of what defines multimedia quality. Further work is required to identify the impact that both objective and subjective parameters have on user quality of perception, to

help prevent future mobile and pervasive multimedia systems from disregarding the user's own definition of multimedia quality.

REFERENCES

- Apteker R.T., Fisher J.A., Kisimov, V.S., Neishlos, H., 1995. Video acceptability and frame rate. *IEEE multimedia* (2), 32-40.
- Bhagwat, P., Perkins, C. Tripathi, S., June 1996. Network Layer Mobility: An Architecture and Survey. *IEEE Personal Communications* 3(3), 54-64.
- Birrell, A.D., Nelson, B.J., 1984. Implementing Remote Procedure Calls. *ACM Transactions on Computer Systems* 2(1), 39-59.
- Bohannon, W., 1997. Buyers Guide to LCD and DLP Projectors. *Presentations Magazine*, 53-64.
- Bouch, A., Kuchinsky, A., Bhatti, N., 2000. Quality is in the eye of the beholder, in: *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems*, The Hague, The Netherlands, 297-304.
- Bouch, A., Sasse, M.A., DeMeer, H., 2000. Of Packets and People: A User-Centred Approach to Quality of Service, in: *Proceedings of IWQoS 2000*, Pittsburgh, PA, 189-197.
- Bowman, D., Datey, A., Ryu, Y., Farooq, U., Vasnaik, O., 2002. Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments, in: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 2134-2138.
- Buyukkokten, O., Garcia-Molina, H., Paepcke, A., Winograd, T., 2000. Power Brower: Efficient Web Browsing for PDAs. *Digital Libraries Lab (InfoLab)*, Stanford University, in: *ACM CHI 2000 The Hague*, Amsterdam, 430-437.
- Cobb, S., Nichols, S., Wilson, J.R., 1995. Health and safety implications of virtual reality: In search of an experimental methodology, in: *Proceedings of FIVEi95 (Framework for Immersive Virtual Environments)*, University of London, 227-242.
- De groot, A.D., 1966. Perception and memory versus thought: Some old ideas and recent findings. In *Problem solving: Research, method, and theory*, B. Klinmuntz (Eds.). John Wiley, New York.
- Ebina, O., Owada, N., Ohinata, Y., Adachi, K., Fukushima, M., 2002. Wearable Internet Appliances and Their Applications. *Hitachi Review* 51(1), 7-11.
- Elting, C., Zwickel, J., Malaka, R., 2002. Device-Dependent Modality Selection for User-Interfaces – An Empirical Study, in *Proceeding of ACM IUI'02*, 55-62.
- Faraday P., Sutcliffe A., 1996. An Empirical Study of Attending and Comprehending Multimedia Presentations, in: *Proceedings of ACM Multimedia '96*, Boston, Massachusetts, USA, 265-275.
- Faraday P., Sutcliffe A., 1999. Authoring Animated Web Pages using Contact Points, in: *Proceedings of ACM CHI '99*, Pittsburgh, Pennsylvania, USA, 458-465.
- Fox, A., Gribble, S.D, Brewer, E.A., Amir, E., 1996. Adapting to Network and Client Variability via On-Demand Dynamic Distillation, in: *Proceedings of the Seventh International ACM Conference on Architectural Support for Programming Languages and Operating Systems*, Cambridge, MA, 160-170.
- Fox, A., Goldberg, I., Gribble, S.D., Lee, D.C., Polito, A. and Brewer, E.A, September 1998. Experience With Top Gun Wingman: A proxy-based Graphical Web Browser for the 3Com PalmPilot, in: *Proceedings of Middleware '98*, Lake District, England, 407-424.

- Fulk, M., 2001. Improving Web Browsing on Handheld Devices, Gvu Center & College of Computing, Georgia Institute of Technology, in: ACM CHI 2001, 395-396.
- Geelhoed, E., Falahee, M, Latham, K., 2000. Safety and Comfort of Eyeglasses Displays, in : Proceeding of Handheld and Ubiquitous Computing : Second International Symposium, HUC 2000 Bristol, Sept. 2000, 236-247.
- Ghinea, G., Thomas, J.P., 1998. QoS Impact on User Perception and Understanding of multimedia Video Clips, in: Proceedings of ACM Multimedia '98, Bristol, UK, pp. 49-54.
- Gips, J., Olivieri, C. P, and Tecce, J.J. 1993. Direct Control of the computer through electrodes placed around the eyes, in M.J. Smith and G. Salvendy (ed.), Human-Computer Interaction: Applications and Case Studies, Elsevier, Amsterdam, 630-635.
- Gips, J., DiMattia, P., Curran, F.X., Olivieri, C. P., 1996. Using EagleEyes -- an Electrodes Based Device for Controlling the Computer with Your Eyes -- to Help People with Special Needs, in: "Interdisciplinary Aspects on Computer Helping with Special Needs" J. Klausm E Auff, W. Kremser, W. Zagler (eds.) R. Oldenbourg, Vienna, 1996.
- Glenstrup, A. J., Engell-Nieson, T, 1995. Eye Controlled Media: Present and Future State. Retrieved from University of Copenhagen (17 June 2003) <http://www.diku.dk/users/panic/eyegaze/article.html>.
- Goldberg, J.H., Stimson, M. J., Lewenstein, M., Scott, N., Wichansky A. M., 2002. Eye tracking in web search tasks: design implications, in: Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium 2002, New Orleans, Louisiana, 51-58.
- Gray, J. Reuter, A., 1993. Transaction Processing: Concepts and Techniques. Morgan Kaufman.
- Hayhoe, M. M., Ballard, D. H., Triesch, J., Aivar, P., Sullivan, B., 2002. Vision in natural and virtual environments, in: Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium 2002, New Orleans, Louisiana, 7-13.
- Hoffman, J.E., 1978. Search through a sequentially presented visual display. Perception and Psychophysics (23) 1-11.
- Hutchinson, T., 1989. Human-computer interaction using eye-gaze input. IEEE Transactions on Systems, Man and Cybernetics. (19), 1527-1534.
- Isokoski, P, 2000. Text input methods for eye trackers using off-screen targets, Eye Tracking Research & Application, Proceedings of the symposium on Eye tracking research & applications 2000, Palm Beach Gardens, Florida, United States, 15-21.
- Jones, M., Marsden, G., Mohd-Nasir, N., Boone, K. and Buchanan, G., 1999. Improving Web interaction on Small Displays, in: Proceedings of 8th International WWW Conference, 51-59.
- Just, M.A. and Carpenter, P.A., 1976. Eye Fixations and Cognitive Processes, Cognitive Psychology (8), 441-480.
- Kaufman, L., Richards, W., 1969. Spontaneous fixation tendencies for visual forms. Perception and Psychophysics (5), 85-88.
- Kawalek, J., 1995. 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.
- Kazasis, F.G., Moumoutzis, N. Pappas, N., Karanastasi, A, Christodoulakis, S., 2003. Designing Ubiquitous Personalized TV-Anytime Services, in: Eder, J., Mittermeir, R., Pernici, B. (Eds.), Proceedings of CAISE'03 Workshops, University of Maribor Press, Slovenia, 136-149.

- Kennedy, R.S., Lanham, D.S., Drexler, J.M., Massey, C.J, Lilienthal, M.G., 1997. A comparison of cybersickness incidences, symptom profiles, measurement techniques, and suggestions for further research, in *Teleoperator & Virtual Environment*, 6(6), 638-644.
- Kim, L., Albers, M.J., 2001. Web Design Issues when Searching for Information in a Small Screen Display. University of Memphis. ACM SIGDOC'01, 193-200.
- Kistler, J.J., Satyanareyanan, M., February 1992. Disconnected Operation in the Coda File System. *ACM Transactions on Computer Systems* 10 (1), 3-25.
- Kolasinki, E.M., 1995. Simulator sickness in virtual environments. Technical Report 1027, U.S. Army Research Institute for the Behaviour and social Sciences, Alexandria, VA. (15 July 2003) http://www.cyberedge.com/3a5_2.html
- Kowler, E. (1990) The role of visual and cognitive processes in the control of eye movement. In *Eye movements and their Role in Visual and Cognitive Processes*. Kowler, E. Ed. Elsevier Science Publishers.
- Lantz, E., 1997. Future Directions in Visual Display Systems. *Computer Graphics*, 31(2), 38-45.
- Mackworth, J. F., Morandi, A.J., 1967. The gaze selects informative details within pictures. *Perception and Psychophysics*, (2) 547-552.
- Mackworth, J.F., and Bruner, J.S., 1970. How adults and children search and recognize pictures. *Human Development*, (13), 149-177.
- Majaranta P., Rähkä K., 2002. Twenty years of eye typing: systems and design issues, in: *Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium*, New Orleans, Louisiana, 15-22.
- Needham, R.M. and Schroeder, M.D., 1978. Using Encryption for Authentication in Large Networks of Computers. *Communications of the ACM* 21 (12).
- Pablo Research Group, 2002. Intelligent Information Spaces. retrieved from Department of Computer Science, University of Illinois at Urbana-Champaign (8 June 2003) <http://www-pablo.cs.uiuc.edu/Project/SmartSpaces/SmartSpaceOverview.htm>.
- Parkhurst, D., Culurciello, E., Niebur, E., 2000. Evaluating Variable Resolution Displays with Visual Search: Task Performance and Eye Movements, in: *Proceedings of the Eye Tracking Research and Applications Symposium*, Palm Beach Gardens, Florida, USA, 105-109.
- Partala, T., Jokiniemi, M., Surakk, V., Pupillary, 2000. Responses to emotionally provocative stimuli, in: *Proceedings of the symposium on Eye tracking research & applications 2000*, Palm Beach Gardens, Florida, United States, 123-128.
- Pelz, J.B., Canosa, R., Babcock, J., 2000. Extended Tasks Elicit Complex Eye Movement Patterns, in: *Proceedings of the Eye Tracking Research and Applications Symposium*, Palm Beach Gardens, Florida, USA, 37-43.
- Pomerleau, D. and Baluja, S, 1993. Non-Intrusive Gaze Tracking Using Artificial Neural Networks, *AAAI Fall Symposium on Machine Learning in Computer Vision*, Raleigh, NC, 153-156.
- Razdan, R., and Kiehl, A., 1998. Eye Tracking for Man/Machine Interfaces. *Sensors*.
- Regan, E.C., 1995. An investigation into nausea and other side-effects of head-coupled immersive virtual reality. *Virtual Reality* 1(1), 17-32.
- Reingold, E., Loschky, L. C., 2002. Reduced saliency of peripheral targets in gaze-contingent multi-resolutional displays: blended versus sharp boundary windows. *Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium 2002*, New Orleans, Louisiana, 89-93.

- Royer, E.M., Toh, C.K., April 1999. A Review of Current Routing Protocols for Ad Hoc Mobile Wireless Networks. *IEEE Personal Communications* 6(2), 46-55.
- Salapatel, P., and Kessen, W., 1966. Visual scanning of triangles in the human newborn. *Journal of Experimental Child Psychology*, (3), 155-167.
- Satyanarayanan, M., 1989. A Survey of Distributed File Systems, in: Traub, J.F, Grosz, B., Lambson, B., Nilsson, N.J. (Eds.), *Annual Review of Computer Science*, Annual Reviews, Inc, 73-104.
- Satyanarayanan, M., 2001. Pervasive Computing: Vision and Challenges. Carnegie Mellon University. *IEEE Personal Communications* 6(8), 10-17.
- Scialfa, C.T., McPhee, L., Ho, G., 2000. The Effects of a Simulated Cellular Phone Conversation on Search for Traffic Signs in an Elderly Sample, in *Proceedings of the Eye Tracking Research and Applications Symposium*, Palm Beach Gardens, Florida, USA, 45-50.
- Sodhi, M., Reimer, B., Cohen, J.L, Vastenburger, E., Kaars, R., Kirschenbaum, S., 2002. On-road driver eye movement tracking using head-mounted devices, in: *Proceedings of the symposium on ETRA 2002: eye tracking research & applications symposium 2002*, New Orleans, Louisiana, 61-68.
- Steinmetz R, Nahrstedt, K., 1995. *Multimedia: Computing, Communications and Applications*. Prentice-Hall, Englewood Cliffs, NJ.
- Tresman, A., 1986. Features and objects in visual processing. *Scientific American*, 255 (5), 106-115.
- TV-Anytime, 2003. TV Anytime Forum Website, retrieved from TV-Anytime.org (15 June 2003) <http://www.tv-anytime.org>
- Ward, A., Jones, A., Hopper, A., October 1997. A New Location Techniques for the Active Office. *IEEE Personal Communications* 4 (5), 42-47.
- Watson, A., Sasse, M. A., 1997. Multimedia Conferencing via Multicast: Determining the Quality of Service Required by the End-User, in: *Proceedings of AVSPN '97 - International Workshop on Audio-Visual Services over Packet Networks*, Aberdeen, Scotland, pp 189-194.
- Weiser, M., 1998. The Future Of Ubiquitous Computing On Campus. *Communications of the ACM* (41), 41-42.
- Wijesekera, D., Srivastava, J, Nerode, A. and Foresti, M., 1999. Experimental Evaluation of Loss Perception in Continuous Media. *Multimedia Systems*, 7 (6), 486-499.
- Xybernaut, 2003. Xybernaut Corporation, retrieved from Xybernaut.com (10 July 2003) <http://www.xybernaut.com>
- Yarbus, A.L., 1967. *Eye movement and vision*, trans. B. Haigh. Plenum Press, New York.
- Young, L.R., Sheena, D.,1975. Survey of eyemovement recording methods, *Behavioral Research Methods and Instrumentation*, 7 (5), 397-429.