National College of Ireland

Personalised Multimedia Educational Content for M-learning Environments

by

Arghir-Nicolae Moldovan

B.Eng.

Thesis

M.Sc. by Research

School of Computing

National College of Ireland

Dublin, Ireland

Research Supervisor:

Dr. Cristina Hava Muntean

Submitted to the Higher Education and Training Awards Council

September, 2010

Declaration

I, Arghir-Nicolae Moldovan, declare that this thesis, submitted by me in partial fulfilment of the requirement for the degree of Masters of Science by Research, is entirely my own work except where otherwise accredited. It has not at any time either whole or in part, been submitted for any other educational award.

Signed: ___________________________________

Date: ____________________________________

Acknowledgements

First of all I want to thank my supervisor, Dr. Cristina Hava Muntean for her continuous support, guidance and patience. Second, I would like to thank Dr. Gabriel Miro Muntean for his support over the past two years. By allowing me to conduct part of my research in the PEL@DCU lab, he facilitated my access to useful resources, and he gave me the opportunity to meet new people, to make friends and connections. Both Christina and Gabriel are great examples of supervisors. From them, I learned not only how to become a better researcher, but also how to become a better person. They also gave me the opportunity to collaborate on two summer school projects, which were great experiences.

I would also like to thank here all the lecturers from School of Computing, especially Dr. Stephan Weibelzahl and Dr. Dr. Dietmar Janetzko for giving me the opportunity, to learn so much from them, during the seminars that I attended, or during the discussions that we had.

Nevertheless I would like to thank Ioana, who understood me like no one over the past few months. Without her continuous support, advices and encouragement, I would not be here today.

Abstract

Thanks to the latest technological advances of mobile devices and web technologies, mobile learning (m-learning) has started to be adopted by an increasing community as an educational platform. There are several challenges in m-learning due to the high variety of mobile devices with different characteristics, different user profiles and various and variable network types and conditions, including the need to provide content suitable to user expectations. Solutions focus on adapting the educational material to suit user interests, goals and expectations, particularities of different user devices, or existing network conditions. As multimedia content usage in m-learning has seen an exponential growth in the recent years, and as delivering multimedia content to learners is a high resource intensive task, adaptation of multimedia-based educational content has become a very interesting research topic. Very few researchers in adaptive m-learning have addressed multimedia content adaptation based on mobile device characteristics or network connectivity, and to the knowledge of the author, none has studied the impact of video quality on the m-learning process. The latest is of much importance as most mechanisms for video quality adaptation involve content quality decrease.

In this context, the research presented in this thesis, complements current research on adaptive mechanisms for multimedia educational content delivery with applicability in m-learning. The thesis proposes a strategy for grouping mobile learning devices in classes with similar characteristics. Each class was associated to a video profile meant to support an optimum level of quality on the target devices. This research also presents a study conducted on a significant number of educational multimedia clips which makes recommendations in terms of optimum quality levels for each of the proposed video profiles. Experiments with different types of educational clips were conducted in order to determine how much the quality of the proposed video profiles can be decreased, while still maintaining good user perceived quality level. Results from a subjective study conducted on a number of participants, have validated the results from the experimental studies, and have confirmed that the learners ability to acquire knowledge is not impacted by a controlled decrease of the video quality.

Keywords: adaptive m-learning, educational multimedia clips, video profiles, quality assessment

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

Currently mobile devices are fast becoming the primary means of accessing the Internet. When compared to Desktop Internet, Mobile Internet not only that is growing significantly faster, but this is expected to become ten times as big in terms of device numbers (Meeker 2009). Among the factors that contribute to this growth, one will note the technological advancements that see more powerful and affordable mobile devices being launched with every year.

As the mobile devices have improved significantly in terms of features and capabilities, their potential for being used as powerful learning tools was quickly realised by educators (Rosman 2008). Latest embedded devices not only that are able to render full webpages but they can also retrieve and display high-quality multimedia content. On one side, learners can finally access with ease rich educational content while on the move, without the need of carrying weighty portable devices such as Laptops. On another side, educators can finally concentrate on creating more engaging, interactive and media rich ***mobile learning*** applications in order to enhance learners' experience and their satisfaction. In particular, learners' satisfaction has been shown to be one of the key factors contributing to the acceptance of m-learning systems (Chen et al. 2008; Yi et al. 2009; Liaw et al. 2010).

More recently, multimedia content has started to be used increasingly in online learning. However, as compared to other type of educational content, delivering multimedia educational content to mobile devices is a very resource intensive task. Significant more bandwidth is required to retrieve the packets of the multimedia clip, streamed over the wireless network, as well as significant more computational power is needed to decode and display the video. Furthermore, with the fast increase in mobile traffic, and especially video traffic (Cisco 2010), networks congestion is expected to continue over the next few years, despite the continuous improvements in networks speed and capacity (Fleischman 2010). These factors in turn, may affect the perceived quality of the multimedia clips and the learners' satisfaction.

When considering also the high variety of mobile devices with different characteristics and capabilities, one can note that there is an increasing need for new and innovative solutions for adapting the multimedia content based on the available resources in order to provide the users in general and learners in particular with a good Quality of Experience (QoE). As opposed to the Quality of Service (QoS), which objectively measures the service provided, QoE is a subjective measure of the users' experience, their expectations and their satisfaction with the service.

QoE has been shown to be an important factor not only in a traditional learning environment (Bassi et al. 2007), but also in an e-learning context (Muntean & McMannis 2006). By adapting the educational content based on learners QoE, significant improvements can be achieved in terms of learning performance and learning outcome (Muntean & Muntean 2007). At the same time, delivering a video that cannot be viewed smoothly and entirely by a learner, because it is not appropriate for the device characteristics, or the available network conditions, will directly and negatively impact the learning process and the learning outcome.

1.2 Motivation

Over the past years, much emphasis has been made on adapting the educational content based on the learners' device characteristics. The majority of this research however, has concentrated on adapting educational content, predominantly static media types such as text and images.

Given the significant research in the area of adaptive e-learning, it has become clear that a "one size fits all" approach cannot be the most efficient solution for delivering the educational content to the learners. Therefore, it is important for m-learning systems to integrate adaptive mechanisms that would tailor the educational content to learners based on their particular device characteristics.

This is especially important when delivering multimedia type educational content to mobile devices connected to Internet via wireless networks. This is because multimedia content as opposed to traditional text and image based content, requires significant resources both in terms of network bandwidth and processing capabilities. Therefore, a learner has to be provided with educational content encoded with the appropriate parameters and quality level, so that it can be delivered over the available wireless network and it can be displayed properly by his/her mobile device.

On another side, multimedia content adaptation has become increasingly popular over the recent years, and a number of adaptive solutions are already available. However, despite having the potential to significantly improve learners' QoE, their implementation in m-learning systems is reduced if none.

The existing multimedia adaptation solutions, can usually adapt the quality of a multimedia clip spatially (by changing the resolution), temporally (by changing the frame rate), or in terms of fidelity (by changing the bitrate). The adaptation may be triggered by different factors such as the difference between various screen sizes or the variable network conditions.

The size of a multimedia clip is dependent on parameters such as length and bitrate. For a smooth playback of the streamed multimedia clip the bitrate has to be adapted to the available network bandwidth. However, the bitrate of a multimedia clip depends on other parameters such as the content (motion, colour levels, etc.), audio and video compression formats, frame rate, resolution, etc.

Considering the video resolution and making abstraction of the other parameters, the optimum user perceived quality level, as well as the efficient usage of the device resources (e.g. CPU load, battery power, etc.) each mobile device should receive a version of the multimedia educational clip that fits its particular characteristics. However, given the multitude of device screen resolutions and their aspect ratios, in practice this is difficult even if resource hungry transcoding of educational multimedia clips is used in real time.

In this context the research presented in this thesis, comes to address the adaptation of educational multimedia clips based both on the device screen resolution, and on the available network conditions. A solution to group the multitude of devices in a reduced number of classes that is feasible to be addressed, as well as maximum and minimum reference quality levels for multimedia educational clips are proposed. The proposed solutions and recommendations can be used for actually implementing mechanisms for multimedia content adaptation in m-learning systems.

1.3 Problems and Goals

A first problem that can be noticed is that many of the existing solutions for adapting the multimedia content require multiple versions of the multimedia clips to be created in advance. In practice it is unfeasible to store multiple versions for all possible resolutions, when, e.g. device screen resolution based adaptation is made. Choosing the target references in this case requires information about the learners' device characteristics. Furthermore, creating these versions, not only that requires the content creator or the m-learning system administrator to know how to use video encoding software, but they also need to know what values to be set for the encoding parameters.

The research presented in this thesis addresses the problem by proposing a solution for grouping multiple devices in classes of devices. Video profiles are associated to each of these classes, and recommendations with regard to the encoding settings for optimum quality levels on the targeted devices are made, starting from an analysis conducted on a significant number of educational multimedia clips. Starting from the same analysis, different types of educational multimedia clips were identified.

A second problem that can be seen is for the case when bitrate-based adaptation is required. In such situations it is very important to know how much the bitrate can be decreased for an educational purposed multimedia clip having little impact on the user perceived quality. Considering that the main purpose of an educational multimedia clip is to spread knowledge, the bitrate reduction should have no impact on learner's capacity to acquire knowledge presented in the clip.

This research work addresses this problem, and looks at determining minimum reference bitrates for the video profiles that have been proposed. Since the video bitrate is also dependant on the content type, among other factors, the minimum bitrate thresholds are proposed for different categories of multimedia educational content that have been identified while addressing the first problem.

Results of a subjective experiment, have confirmed that decreasing the bitrate to the proposed thresholds, does not affect the learning process, and learner's capability to acquire knowledge from the educational clips.

1.4 Contributions

The main contributions of this thesis include:

- An analysis of the most common mobile devices that can be used for m-learning activities, and the characteristics of these devices.
- An analysis of a very large collection of educational multimedia clips currently used in education. Based on the type of information presented in the clips, eight different categories of educational multimedia clips, with different characteristics have been identified. These categories are: screencasts, slideshows, animations, games / virtual learning environments recordings, interviews, lecture recordings / presentations, lab demos, and documentaries.
- An analysis of the most common encoding schemes used for encoding educational multimedia clips
- A novel solution for enabling device screen resolution based adaptation of educational multimedia clips. Various video profiles were defined and associated to different classes of mobile devices. Recommended values for the encoding parameters, corresponding to optimum levels of quality were also proposed for each video profile.
- Recommendations on the minimum acceptable bitrate thresholds to be used for the identified types of educational clips, that have little impact on the user perceived multimedia quality. The bitrate thresholds were first estimated using objective video quality assessment metrics. Objective video quality metrics aim to determine the level of user perceived quality.

A subjective study was then conducted on a number of participants. The goals of this study were to validate the estimated values, to confirm that learner's perceived quality was not significantly affected and to demonstrate that learner's capacity to accumulate the knowledge presented in the clip was not affected by a controlled reduction in the video quality.

1.5 Thesis Outline

The thesis is structured as follows:

Chapter 1 introduces the problem and the goals of this research.

Chapter 2 presents the research areas related to this project, as well as the technical and theoretical background required to understand where the research presented in this thesis fits in the area of adaptive mobile learning.

Chapter 3 presents preliminary investigations that were conducted in order to assess common device screen resolutions and video encoding schemes for multimedia educational content. Based on the preliminary findings, video profiles addressing different classes of devices are proposed.

Chapter 4 presents the work conducted with the goal to determine the recommended minimum bitrate levels for the proposed video profiles. The general characteristics of the different categories of the educational multimedia clips, that have been identified and addressed during the testing phase, are also presented in this chapter.

Chapter 5 summarises the results, concludes the thesis and presents future work directions.

Chapter 2

Literature Review

This chapter presents the technical and theoretical background as well as the research areas related to this research project. For the beginning Section 2.1, looks at the evolution of e-learning towards m-learning, as mobile devices are increasingly used for accessing educational content over the wireless networks. Further on, Section 2.2 presents an overview of the main personalisation and adaptation techniques used to overcome many of the issues and challenges characteristic to m-learning in particular, and e-learning in general.

Since, this research has primarily focused on enabling multimedia educational content adaptation in m-learning systems, it is important to have a look at the current state of multimedia content usage in e-learning and m-learning (Section 2.3), as well as at some of the existing approaches for multimedia content adaptation (Section 2.4) and classification (Section 2.5).

In order to provide the learners with a good learning experience, their perceived quality of the educational multimedia content has to be considered. Therefore, Section 2.6 finishes this chapter by providing an overview of the objective and subjective methods for assessing the multimedia quality.

2.1 From E-learning to M-learning

Over the past decades, there have been significant changes on how people learn and train. Characteristic to this change is the increasing usage of technology for facilitating learning both in the classroom and outside of it. New paradigms defining various forms of ***technology enhanced learning*** such as ***electronic learning (e-learning)*** and ***mobile learning (m-learning)*** have appeared, and various definitions have been proposed for each of these over the years.

One of the first and most comprehensive definitions of e-learning, was proposed by Mary White in 1983. In her opinion e-learning represents:

"...learning from any device dependent upon the actions of electronics, such as television, computers, microcomputers, videodiscs, video games, cable, radio interactive cable, videotexts, teletext, and all the other devices in the process of being invented that are electronic in nature" (White 1983, p.51)

More recently, Marc Rosenberg has suggested the following definition for e-learning:

..."the use of Internet technologies to deliver a broad array of solutions that enhance knowledge and performance." (Rosenberg 2001, p.28)

Two years after this definition was proposed, e-learning came to include also the learning delivery by the means of mobile technologies. According to Derek Stockley, e-learning refers to:

"...the delivery of learning, training or educational program by electronic means. E-learning involves the use of a computer or electronic device (e.g. a mobile phone) in some way to provide training, educational or learning material." (Stockley 2003)

However, as the wireless networks and the mobile devices have improved in terms of speed and capabilities, and as they have become more accessible and affordable to an increasing number of people, they have started to be increasingly used for accessing educational content over the Internet.

In this context, m-learning appeared as a new paradigm of e-learning (Leung & Chan 2003), coined as the *"convergence of mobile technologies with e-learning"* (Yi 2009). M-learning was seen as being the natural evolution of e-learning in the context of the mobile revolution, and many researchers have marked this transition (e.g. Nyiri 2002; Georgiev et al. 2004; Sharma & Kitchens 2004; Laouris & Eteokleous 2005). For example, Georgiev et al. (2004) pointed out that m-learning represents:

"...a new stage of the progress of distance learning (d-learning) and e-learning" (see Figure 2.1)

M-learning has evolved from e-learning as the learners became increasingly mobile, while e-learning has previously evolved from d-learning, as the last one came to rely more heavily on technology (Cherian & Williams 2008).

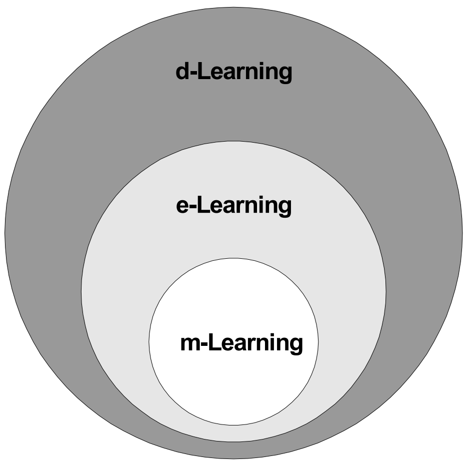

Figure 2.1: The Place of M-learning as Part of E-learning and D-learning (Georgiev et al. 2004)

When it comes to the learning process and how this takes place, there are small differences between m-learning and e-learning from a psychological perspective. Instead m-learning differs more of e-learning from a technological perspective, in terms of the tools and materials that are used during the learning process.

Often, m-learning also differs of e-learning in terms of the learning context where the learning takes place. If characteristic to e-learning was the usage of technology in the classroom, laboratory or at home, m-learning has enabled learning to occur *"in the field, or at any location where the mobile device is fully functional"* (Sharma & Kitchens 2004 as cited in Cobcroft 2006).

Therefore, m-learning has increased learners' independence and their control on the learning process, by enabling them to *"physically move their own learning environments as they move"* (Barbosa and Geyer, 2005). In this sense, m-learning came to represent *"the acquisition of any knowledge and skill through using mobile technology, anywhere and anytime"* (Geddes 2004).

As John Traxler (2007) pointed out, mobile devices are increasingly changing the nature of knowledge in modern societies, fact that in turn alters the nature of learning and also the way in which learning can be delivered. *"Learning that used to be delivered 'just-in-case,' can now be delivered 'just-in-time,' 'just enough,' and 'just-for-me.' Finding information rather than possessing it or knowing it becomes the defining characteristic of learning generally and of mobile learning especially"* (Traxler 2007).

As the mobile technologies are continually improving and new learning strategies and practices are developed, m-learning is gradually evolving to realize the promise of *"ubiquitous, persuasive, personal and connected learning"* (Wagner 2005).

2.2 Personalisation and Adaptation of Educational Content

M-learning is an interdisciplinary field that covers computer science and wireless communications as well as education (Papanikolaou 2006). The transfer of traditional learning into the electronic environment is a complex task and it requires an interdisciplinary approach. This is further complicated by the multitude of technologies involved, by the differences in learners' characteristics, and by different learning and teaching practices.

Many challenges and issues have been identified in m-learning over the years, and many solutions to overcome these have been proposed both by the research community and by the industry. A number of these technological and educational issues can be addressed from the development stage of the m-learning applications. Therefore, it is important for these to be deployed around the targeted audiences, in order to meet their expectations (Papanikolaou 2006).

As it has become clear that a "just-put-it-on-the-Web" approach (Brusilovsky & Peylo 2003) cannot satisfy the specific requirements of all the participants, being them people or technologies, adaptation and personalisation have gradually been brought to the forefront of the research in the area of m-learning in particular and e-learning in general. Various solutions for personalising or adapting the educational content and adapting the learning process have been proposed, that have considered the differences in terms of:

- learner characteristics (learner profile);
- learning styles;
- learning context;
- mobile device characteristics.

2.2.1 Learner Profile Driven Personalisation & Adaptation

Most of the adaptive approaches in the area of e-learning and in particular m-learning, have concentrated on delivering personalised educational content tailored to individuals or groups based on learner characteristics. Various aspects were investigated as important input in the personalisation of the course material and learning process. Content personalisation may be driven by the needs of individual learners, their preferences, goals (Clifford 2000), knowledge level (De Bra et al. 2003), skills or cognitive preferences (Chen & Macredie 2002). Adaptation of the learning process may be driven by the learner's abilities, motivation and their previous interaction with the e-learning environment, as well as by the learner's concentration level and frequency of disruptions (Bomsdorf 2005).

Adaptive e-learning/m-learning systems build a model (named User Model or Learner Model) based on the goals, the preferences and the knowledge of each individual learner, which is used throughout the interaction of learner with the system (De Bra et al 1999). Various navigational, layout and content adaptation techniques are applied in order to adapt to each particular learner’s needs (Brusilovsky 2001). Such techniques include link hiding, annotation or disabling. These have the role to guide the learner towards the information relevant for him/her, while at the same time hiding the information that is inappropriate or non-relevant for him/her (Brusilovsky & Millan 2007).

The information used to build the user model is usually collected in two ways:

- ***explicit*** – the information is provided by the user (e.g. when registering in the system);
- ***implicit*** – the information is collected automatically by the system, by monitoring the learner’s observable behaviour (e.g. mouse and keyboard actions) (Hanani et al. 2001 as cited in Berkovsky et al. 2007).

2.2.2 Learning Styles Driven Personalisation & Adaptation

Since different learners have different learning particularities to create equity among them, the personalisation process has to consider the theories of learning and cognition (Dede 2008; Koohang et al. 2009), as well as the various learning styles.

Sampson and Karagiannidis (2002) have argued that the concepts of adaptive and personalised learning are built mainly on the cognitive and constructivist theories of learning. According to the cognitive theory, adaptive e-learning/m-learning systems support active participation of the learners in order to help them organise and structure their knowledge, and to link the new knowledge to prior structures (Sampson & Karagianidis 2002). In contrast, the adaptive systems based on the constructivist theory, personalise the learning process in a way that encourages the learners to actively explore the concepts and to improve their thinking by creating "new ideas or concepts based both on their past and current knowledge" (Naismith et al. 2004).

Adaptation based on the various learning styles (e.g. visual, auditory, and kinesthetic), includes presenting the information using multiple media channels, such as text, images, audio and video. In particular, combining images with audio has been shown to be very efficient because "pictorial and verbal representations" are held in the working memory at the same time (Mayor & Moreno 1998).

Most of the research on learning styles based adaptation, has concentrated on providing users having particular learning styles, with content that is most appropriate to that style (Brusilovsky & Millán 2007). The adaptation can be: ***navigation-level*** adaptation (e.g. direct guidance, link annotation), ***content-level*** adaptation (e.g. content hiding, additional explanations, specific media hiding, specific items filtering), or ***presentation-level*** adaptation (e.g. inserting/removing fragments, altering fragments, sorting fragments) (Brusilovsky 2004 as cited in Popescu et al. 2007).

Examples of adaptive systems that have integrated learning styles based adaptation include iWeaver (Wolf 2003), INSPIRE (Papanikolaou et al. 2003), AES-CS (Triantafillou et al. 2003) and AHA! (Stash et al. 2006).

Various aspects of these approaches have been criticised in the literature. For example Popescu et al. (2007) have stated that most of the implementations are based on a single learning style model, and have questioned the validity and reliability of using explicit methods (e.g. dedicated psychological questionnaires that need to be filled by learners in the beginning), for detecting the learning styles. Therefore the authors have proposed to use a unified learning style model that integrates relevant characteristics from multiple models in the literature, and the learning style to be determined implicit, by monitoring learner's interactions with the system.

Brown et al. (2009) have questioned the methods used for assessing the impact that such systems have upon the learners, noting that most of them are not backed up with scientific evidence. Further, by presenting two qualitative evaluation studies that have shown no statistically significant benefits of using learning styles based adaptation, the authors argue for the need of more rigorous approaches to be carried out. Given the limitations of online environments, adaptive e-learning systems have mainly favoured the auditory and visual learners. More recently, solutions to increase the number of sensory channels for capturing information have started to be proposed. Hamza-Lup and Adams (2008) have proposed a system that incorporates a multimodal haptic simulator. This allows the students to better understand difficult concepts such as physics concepts, by simulating sensations that normally can be felt in lab experiments.

2.2.3 Learning Context Driven Personalisation & Adaptation

Learning does not depend only on the learner's characteristics, but it also depends on the learning context. The context is described by a multitude of factors, such as learners' intrinsic and psychological properties (e.g. emotional states, confusion, focus of attention, etc.), their location in time and space, their identity, but also by the technology used and its properties (Wang 2004).

M-learning systems have to adapt to the context and often to change it, so that the learner can achieve the learning outcome. E.g. a teacher can adapt his teaching techniques if the students show signs of confusion or frustration. Intelligent feedback mechanisms have appeared to be viable solutions for simulating teacher-specific actions, thus increasing the flexibility and improving the evaluation of m-learning systems (Sharples 2007).

Learner confusion or tiredness may be sensed using eye-tracking systems (Gütl et al. 2005). However, using such sophisticated devices for enabling real-time adaptation of the educational content (Wei et al. 2009), is not feasible given their high prices, as well as their dimensions that makes them still difficult to be implemented on mobile devices.

As opposed to sensing the learner's psychological traits, determining their location and physical environment can be done using technologies that are significantly cheaper and readily available. Since GPS (Global Positioning System) receivers are becoming a norm for the latest embedded devices, in particular smartphones (van Diggelen 2009), these can be used to determine learners' location, within few meters of accuracy.

Examples of adaptive m-learning systems that make use of GPS for enabling personalised location-aware learning include Ambient Wood (Rogers et al. 2005), LOCH – Language learning Outside the Classroom with Handhelds (Ogata et al. 2008) and JAPELAS2 (Yin et al. 2010). An example of integration with LMS was proposed by Rodrigues et al. (2010).

However, GPS is not that useful when it comes to automatically determining learners' indoor location. The latest can be detected by tagging the surrounding objects using technologies such as RFID (Radio Frequency Identification), barcodes or Infrared (Naismith & Smith 2009). Learners' indoor location can also be computed from the signal patterns of the wireless network (Schwabe & Goth 2005).

2.2.4 Mobile Device Driven Personalisation & Adaptation

The transition from e-learning to m-learning did not come without challenges. Delivering learning content to mobile devices requires the educational applications to be able to run efficiently on a multitude of mobile devices that differ broadly in terms of characteristics such as:

- Operating system;
- Screen: size, resolution, pixel density (number of pixels per inch), colour depth (number of bits per pixel), mode (portrait / landscape);
- Processor speed, memory capacity;
- I/O interfaces (e.g. touch screen, keyboard, touchpad, etc.);
- Wireless connectivity (e.g. WiFi, 3G, WiMAX, Bluetooth, etc.)
- Features and capabilities (e.g. the possibility to display e-mail and web pages, to play audio and video (and what are the supported formats), Flash, Java support, etc.).

Given the high diversity of mobile devices, in practice it is difficult and often unfeasible to target a broad audience of devices. In this context, it is very important to select in advance the technologies that will help to best reach the targeted learning audience.

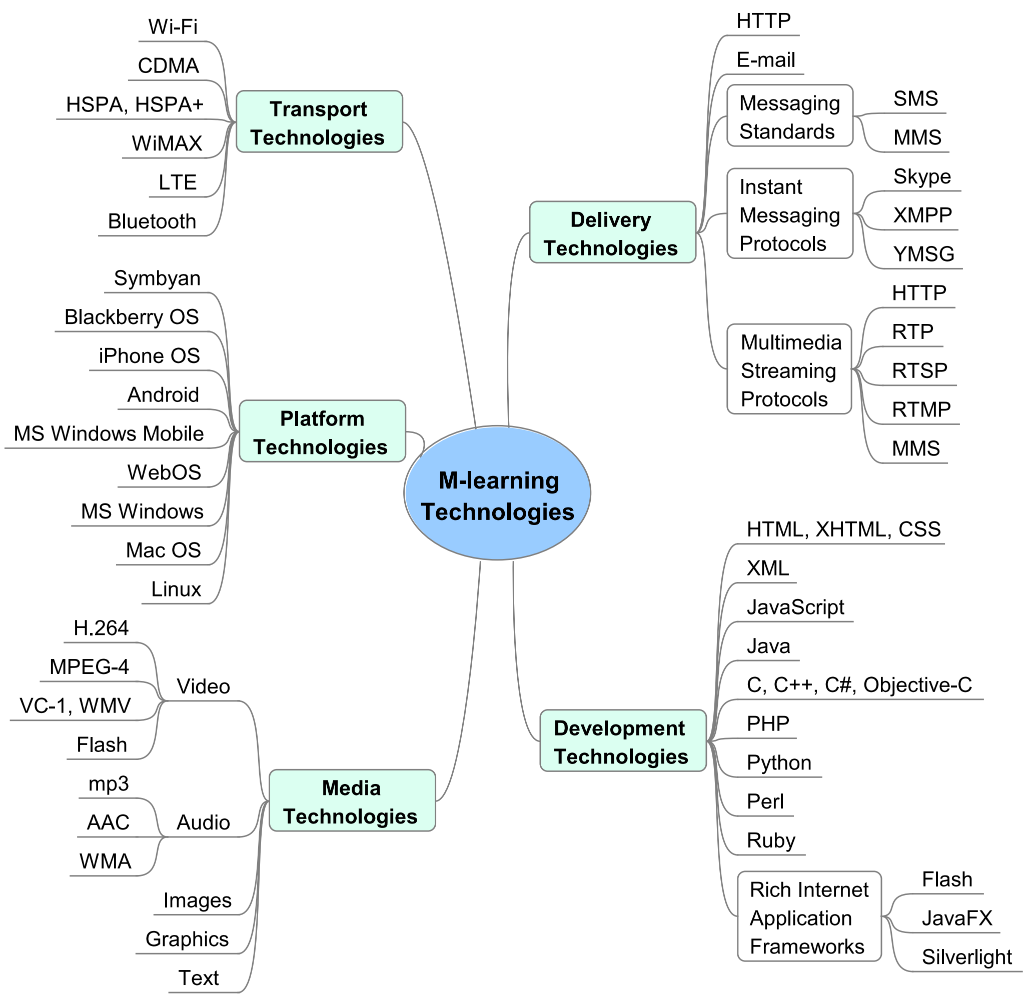

Figure 2.2: Mobile Learning Technologies (based on Attewell 2003)

Attewell (2003) has identified five broad categories of technologies within which specific technologies need to be selected: ***transport***, ***platform***, ***delivery***, ***media*** and ***development*** technologies. Following the same template as in (Attewell 2003),

Figure 2.2 presents the five categories with relevant examples as of 2010. Market research data from GSA (2010), Gartner (2010) and Tiobe (2010) was used for selecting some of the most relevant transport, platform and respectively development technologies.

According to Atwell (2003) the selection of these technologies should consider factors such as:

- the usefulness of the features offered;
- cost;
- availability and patterns of availability;
- reliability;
- ease of use (for both users and developers);
- standardisation;
- likely longevity in the market place;
- likely popularity (in particular with applications that target young audiences, that are concerned about trends, and might refuse to adopt some devices).

Full cross platform solutions are often limited by proprietary solutions, device incompatibilities or licence agreements. For example, Apple's iPhone OS does not offer support for technologies with cross-platform capabilities such as Java or Flash, and through licence conditions, native applications can be developed using only the C, C++ and Objective-C programming languages (Bright 2010).

In such context, when it comes to porting existing e-learning systems to mobile devices, the developers have either to make a compromise and to concentrate on specific devices, or to find standard solutions that will cover a multitude of devices. An example is given by the various projects that have started to enable mobile learning through Moodle[1] Learning Management System (LMS).

[1] Moodle.org: open-source community-based tools for learning, http://moodle.org/

- **mTouch**[2]: native application for iPhone;
- **Moodle4iPhone**[3]: native application for iPhone;
- **MOMO** – Mobile Moodle Project[4]: native application for Java enabled mobile devices;
- **MLE-Moodle**[5] an extension to Moodle LMS that offers two possibilities for accessing the content from mobile devices:
  a) Through the mobile browser of the device;
  b) Through a Java application installed on the mobile device.

The development and adoption of standards and best practices in m-learning, and e-learning generally, has appeared as a necessity for supporting platform independent, open technologies and user-centric systems (Varlamis & Apostolakis 2006) and reducing the barriers to accessing m-learning content (Low 2007).

Although currently there are no standards developed particularly for mobile learning, m-learning can refer to e-learning standards which are developed by organisations such as IEEE LTSC[6] (Learning Technology Standards Committee), AICC[7] (Aviation Industry CBT (Computer-Based Training) Committee), IMS Global Learning Consortium[8], and ADL[9] (Advanced Distributed Learning) (Anani et al. 2008).

---

[2] mTouch- Moodle for IPhone, http://www.pragmasql.com/home/moodletouch.aspx
[3] Moodle4iPhones Project, http://iphone.moodle.com.au/
[4] MOMO (Mobile Moodle) Project, http://www.mobilemoodle.org/momo18/
[5] MLE - Mobile Learning Engine, http://mle.sourceforge.net/mlemoodle/
[6] IEEE Learning Technology Standards Committee (LTSC) — IEEE Learning Technologies Standards Committee http://www.ieeeltsc.org:8080/Plone
[7] Aviation Industry CBT Committee, http://www.aicc.org/dev/
[8] IMS Global Learning Consortium, http://www.imsglobal.org/
[9] ADLnet - Advanced Distributed Learning: Home Page, http://www.adlnet.gov/

E-learning standards have several merits and abilities that help protect the investment in e-learning systems, enabling them to grow, to be sustainable and maintainable, and facilitate the delivery to the learners (The Learning CONSORTIUM 2002). The abilities defined by the Masie Center e-Learning Consortium[10] were summarised by O'Connell and Smith (2007) as follows:

- *interoperability* (ability of two or more systems to share information);
- *reusability* (ability to reuse or modify existing systems, data or code);
- *manageability* (ability to monitor and maintain systems, data or code);
- *accessibility* (ability of many users to access a system and its data or code);
- *durability* (ability of a system to endure over time);
- *scalability* (ability of a system to handle growing amounts of information and work);
- *affordability* (ability of systems and data to remain in financial reach of users).

Personalisation based on device characteristics has mainly focused on the screen resolution. A general framework to deploy device independent user interfaces has the following characteristics (Glavinic et al. 2008):

- the UI of a mobile application is separated from the application logic (as in any client/server architecture) and adapted to the particular mobile device the learner is using to access it;
- the UI is described at a high level of abstraction, using device-independent

[10] The Learning CONSORTIUM | Center Information, http://www.masieweb.com/the-learning-consortium.html

languages (most often XML-based);

- the characteristics of the mobile devices are stored in device profiles;
- the UI is generated by a interface renderer based on the mobile device profile according to the UI rules.

Mobile device and its characteristics can be detected by asking the user to fill a simple form the first time s/he accesses the system, and by matching the answers to one of the devices available in the database (Ally et al. 2005). Another proposed approach was to get this information from the WURFL (Wireless Universal Resource File)[11] online device repository, based on the user agents that can be retrieved automatically when the learners access the server (Zhao & Okamoto 2008).

Existing e-learning standards are also used for enabling content reusability and interoperability, between various e-learning and m-learning systems. For example MEAT (Mobile E-learning Authoring Tool) (Kuo & Huang 2009), is an authoring tool that can produce adaptable learning content and test items conforming to ADL SCORM (Sharable Content Object Reference Model).

Mobile devices may connect to Internet through a variety of wireless networks that differ in terms of capacity and coverage. Learners may encounter connection problems due to:

- reduced capacity;
- variations in network conditions (e.g. available bandwidth, delay, loss, etc.);
- loss of coverage;

[11] Wireless Universal Resource File, http://wurfl.sourceforge.net/

- cost or other reasons that may determine the users to limit their online connection time.

Personalisation of the educational content based on mobile device network capacity and variable network conditions (Muntean 2008), that takes into consideration learners Quality of Experience (Muntean & Muntean 2007) has been shown to improve the learning performance, by decreasing parameters such as time of study, time spent per page and number of page revisits.

Chang et al. (2008) have proposed an adaptive course catching strategy based on the SCORM Sequencing and Navigation (S&N) specification, reducing the latency induced by low speed wireless networks, when displaying multimedia learning resources on mobile devices.

Trifanova et al. (2004) have proposed a personalisation solution to assist the learners that have to limit their online connection times from various reasons. Applied to a language learning application, this solution predicts the texts the user would prefer to study during the next offline session(s) and uploads them on the mobile device. The prediction is based on learner characteristics such as: knowledge level (used to exclude the words that are never viewed by users with similar language skills as the current observed user) or field of interests (e.g. economic, tourism, art, etc.).

2.2.5 Summary

This section has presented an overview of various approaches for personalising and adapting the educational content delivered by e/m-learning systems.

Literature review shows that over the last decade, there has been much emphasis on adaptation and personalisation in the area of e-learning. State of the art research has contributed significantly towards creating more flexible online learning environments that nowadays are capable to sense and adapt to the specific requirements of each particular learner. Adaptive e-learning systems can improve, complement or replace traditional learning with success in many situations.

As the learners have increasingly started to use mobile devices for accessing the educational content, adaptive solutions were necessary to overcome the issues posed by the high variety of characteristics between mobile devices and their limited resources, which initially contributed against the wide acceptance of m-learning. Existing standards also helped overcome the problems posed by the multitude of platforms and the incompatibilities between them, further enabling content reusability across various e/m-learning systems.

2.3 Multimedia Content in E/M-learning

Although ***multimedia content***, and in particular how this can enhance the learning process has been at the forefront of research for more than a decade (Najjar 1996; Rogers & Scaife 1997), it was only in the recent years when thanks to technological advancements and the emergence of a new paradigm, E-learning 2.0 (Downes 2005), the usage of multimedia in online courses has started to increase considerably (Safran et al. 2007).

Multimedia was defined by Neo and Neo (2001) as representing:

"...the combination of various digital media types such as text, images, sound and video, into an integrated multi-sensory interactive application or presentation to convey a message or information to an audience".

Apart of the rich display of information created with the help of multimedia technologies, there are many other ways in which the learners can benefit. By reviewing a number of previous research studies that have looked to assess the advantages of using various channels for displaying information, Shepard (2003) came with the following list of general benefits that multimedia content has for adult learners:

- ***Alternative perspectives*** – clarification needed to understand concepts presented in text;
- ***Higher interactivity*** – learners become active participants rather than passive observers;
- ***Accelerated learning*** – less time needed to complete a set of learning objectives as compared to more traditional approaches;
- ***More effective learning*** – improved retention and application of knowledge;
- ***Acquire computer-related skills.***

There are also a number of more specific benefits for the particular cases represented by ***simulations*** (e.g. development of problem solving skills, increased understanding of systems, processes or concepts, etc.), ***authoring tools*** (e.g. develop collaborative and time management skills, increased retention of new learning, develop ability to effectively use resources to present ideas, etc.) and ***interactive multimedia computer based training*** (e.g. facilitates individualised learning and immediate feedback, increased engagement and learning rate, immediate access to support information, etc.) (Shepard 2003).

Designing a multimedia based course, is a crucial step in order for the learners to benefit from the use of multimedia. As “multimedia adds complexity both to the screen and to the tasks that learners need to perform”, the multimedia elements have to be selected only if they support learning needs (Shank 2006). If too many elements are presented simultaneously, learners may shift their attention in between them reducing the learning performance. Mayor and Moreno (1998) have shown that in multimedia-based learning environments, the *split-attention effect* may occur when for example, an image and the description of the image (text in this case) are both presented visually, overloading the visual *working memory*. Learning performance can be significantly improved if the description is provided on a different channel (e.g. audio description). By reducing the *cognitive load* in this way, additional working memory capacity is freed, allowing the learner to acquire more information and to use it in achieving more advanced schemas in the *long-term memory* (Paas et al. 2003).

From all types of media being used in e-learning, educational multimedia clips have probably seen the most spectacular evolution over the past few years. A recent report that interviewed 57 faculty and librarians from 20 Higher Education institutions from USA, has shown that the video production and consumption for educational purposes is expected to grow exponentially over the next few years, across all disciplines (Kaufman & Mohan 2009).

Currently, an increasing number of educational institutions are making their educational videos available through channels such as iTunes U[12], YouTube[13] and AcademicEarth[14]. Furthermore, freely accessible video sharing services as well as the availability of recording equipments that are getting cheaper and better by day, have created the premises for anybody to contribute and possibly make an impact. A good example in this sense is provided by Khan Academy[15], a not-for-profit organisation started by Salman Khan with the goal “to use technology to provide a free, world-class education to anyone, anywhere” (Khan 2009). With more than 70.000 daily views on its YouTube channel consisting of more than 1600 mini-lectures covering “everything from basic arithmetic to advanced calculus,

[12] Apple - iTunes U - Learn anything, anywhere, anytime., http://www.apple.com/education/itunes-u/
[13] YouTube – Education, http://www.youtube.com/education
[14] Academic Earth | Online Courses | Academic Video Lectures, http://academicearth.org/
[15] Khan Academy, http://www.khanacademy.org/

chemistry, economics and biology" (Khan 2009), Khan Academy has clearly succeeded to make a difference for many students (Krieger 2010).

Multimedia clips have various applications in e-learning, the following representing only few of them:

- instructional screencasts that combine video capture of the screen with audio explanations;
- lecture and lab sessions recordings that can be accessed in real-time or can be viewed at a later time;
- graphics and animations that explain and enforce the understanding of various concepts, physical processes, etc.

2.3.1 Screencasts

Screencasts have the advantage "that it gives the presenter the ability to show his ideas and flow of thoughts rather than simply explain them, which may be more confusing when delivered via simple text instructions" (Wikipedia n.d.). Being provided with clear instructions, learners can better feel the connection with the instructor, and the one-on-one experience characteristic to classroom learning. Furthermore, learners have the advantage of being able to go through the explanations as many times as they want (Wikipedia n.d.).

Based on previous research in the area of cognitive psychology, Oud (2009) has proposed a number of guidelines for creating screencasts that are effective for instruction. Her guidelines cover topics such as:

- ***Cognitive load*** - it has to be minimised through, e.g. removing unnecessary content, using consistent styles, making activities easy to complete without additional help, etc.
- ***Interactivity*** – learners need to have the control, and to be provided with

activities that simulate realistic scenarios, etc.

- ***Critical thinking*** – screencasts should support critical thinking by, e.g. linking the small steps to a broader framework, and by providing activities that help learners develop both higher (evaluation, analysis) and lower levels (understanding, applying) of learning.
- ***Target audience*** **–** learners have to be provided with the appropriate level of difficulty, based on their characteristics and pre-existing knowledge.

2.3.2 Lecture Recordings

Various studies have shown that not only students enjoy being provided with lecture recordings but they can also benefit from these (Soong et al. 2006; Copley 2007; Flores & Savage 2007; Evans 2008). Lecture recordings enable students to re-access parts of the lectures that they do not understand well and help them better preparing for examinations (Soong et al. 2006), as "students believe that podcasts are more effective revision tools than their textbooks and they are more efficient than their own notes in helping them to learn" (Evans 2008).

There are multiple solutions for creating lecture recordings. On a low scale, instructors can record their speech and add it to a slide show using dedicated software such as Camtasia Studio[16]. At a university scale, various solutions with different level of automatisation, cost and complexity have been proposed. A simpler solution is to assign student staff to record the front view of the classroom (including lecturer, whiteboard and projector screen) using a high definition camera (Nagai 2009). A more complex solution is provided by the PAOL (Presentations Automatically Organized from Lectures) system (Dickson et al.

[16] TechSmith | Camtasia Studio, Screen Recording SW, http://www.techsmith.com/camtasia.asp

2009), which records individually the different channels of information (lecturer, whiteboard and lecturer's computer screen) and combines them in a final Flash based presentation. The main advantages of this solution are the complete automatisation of the entire process and the usage of relatively inexpensive technologies that are used, which enable a cost effective automatic lecture recording system. Integrating a lecture recording system with a LMS and CMS (Content Management System) that most universities have in place, has the potential to reduce the workload of lecture recording personnel, while for the students it provides a direct and personalised access to the recordings specific to their subjects (Martens et al. 2008).

2.4 Multimedia Content Adaptation

Section 2.2 has shown much emphasis has been put in e-learning on personalising and adapting the educational content. This research however, has concentrated either on tailoring the educational content to learners, based on their profiles or the context in which the learning occurs, or on adapting the learning material and the presentation layout to the particular requirements of the mobile device used (e.g. screen size, processing power, network connectivity, etc.). In these cases, educational content may be represented by any type of media or combination of multiple media elements (e.g. text, image, video, audio, etc.).

This section provides an overview of the existing approaches for adapting the multimedia content, specifically the multimedia clips. Although these approaches are characteristic to multimedia in general, they can also be applied for educational multimedia content.

Broadly speaking, there have been a multitude of solutions for adapting the multimedia content, and this area still presents a lot of interest both for the research community and for the industry. These adaptive solutions were proposed with various goals, such as:

- Providing the users with the best Quality of Experience (QoE);
- Providing the users with the content they are most interested in viewing;
- Overcoming the issues posed by the multitude of mobile devices and their characteristics (e.g. screen size, memory, processing power, video and audio formats supported, battery life, etc.);
- Overcoming issues with the wireless networks, e.g. different networks capacity (e.g. HSDPA up to 14Mbps, IEEE 802.11g 54Mbps, etc.), variable network conditions (available bandwidth, loss, delay, etc.) caused by environmental factors, by the number of clients connected and their mobility patterns, etc.

Technically, these solutions address various aspects, such as:

- Reshaping the network traffic flow (e.g. Adams & Muntean 2007);
- Improving existing transport protocols (e.g. Padhye et al. 1999);
- Improving existing compression techniques;
- adjusting the quality of the whole clip or of the regions that present less interest for the viewers (e.g. Muntean et al. 2008), etc.

However, presenting such technical solutions is definitely out of the purpose of this section. Instead, this overview will look at the adaptive multimedia from a broader perspective, presenting the main approaches without going into details.

From an architectural point of view, multimedia content adaptation can be done (Marri 2007):

- S***erver side*** – the server is responsible with the adaptation of the multimedia content, based on, e.g. network conditions, device specifications, etc.; the adaptation can be done ***off-line*** or ***on-the-fly***;
- C***lient side*** – the client adapts the content to its characteristics, or selects

the most suitable version form several versions that are provided;

- At a ***proxy node*** placed between the client and the server; the proxy node decides if and when adaptation is needed, and it also adapts the content retrieved from the server, before passing it to the client.

Furthermore, there are a number of so called ***hybrid*** approaches (Tusch et al. 2004), which split the adaptation job between two or more differently located components.

Depending on how the information is stored and/or delivered to the end user, the main approaches for adapting the multimedia can also be classified in:

- Single-layer approaches;
- Multi-layer approaches.

2.4.1 Single-Layer Approaches

Chang and Vetro (2005) have classified the solutions for adapting the multimedia content in the following four categories: ***transcoding***, ***selection/reducing***, ***replacement***, ***synthesis***. Multimedia transcoding especially when is done in real time, requires significant processing power, therefore this is usually performed on the server side or at an intermediate proxy node. The other three adaptive solutions require a more elaborate analysis and pre-processing of the multimedia content, therefore they are usually performed on the server side.

2.4.1.1 Multimedia Transcoding

Multimedia transcoding usually refers to converting the file from one format to another (The Free Dictionary n.d.). Changing the video compression from MPEG-2 to MPEG-4, and the audio compression from mp3 to AAC, represent examples of format transcoding.

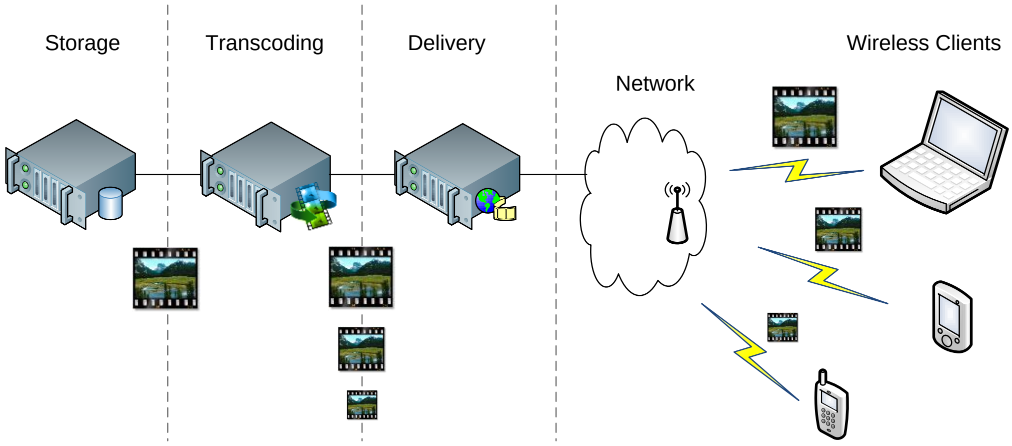

Figure 2.3: Real-Time Transcoding-Based Adaptive Multimedia

Furthermore, a number of other transformations are also often considered as representing forms of transcoding (Chang & Vetro 2005). For example, in case of video content, decreasing the spatial resolution (horizontal and vertical dimensions in number of pixels), temporal resolution (frame rate) and/or the fidelity are such examples.

Multimedia transcoding is a resource intensive task. Therefore it is usually performed on the server or on a proxy node and can be done either in real-time or off-line.

Real-Time Transcoding

In this case, only the version with the highest quality is stored on the server, for each multimedia clip. When a client device requests a multimedia clip, the transcoding process starts to generate a new version of the clip with characteristics that are appropriate for that particular device, and to transmit it to the client. In often situations, the transcoding parameters are also changed dynamically, while the video is transmitted to, e.g. enable a smooth playback on a device with variable network connection. Figure 2.3 presents the conceptual architecture of an adaptive multimedia system based on real-time transcoding.

The main advantage of this solution is the high granularity (the transcoding parameters can be changed in very small steps), which enables an increased level of adaptation. Multimedia transcoding can better overcome the issues posed by the high diversity of mobile devices and the multitude of factors that may impact the transmission of the multimedia clip.

However, transcoding is often not feasible given that it requires significant processing power, and this needs to increase with the number of clients being served at a time. For this purpose, in practice real-time transcoding is mostly used in case of live transmissions and only a limited number of parallel streams are created.

Off-line Transcoding

In this case, multiple versions with different characteristics are created in advance. Usually this is done automatically every time a new clip is uploaded on the server. Depending on the purpose on which the adaptation is done, this can consist in, e.g. choosing the most suitable version for a particular device and/or switching in between different versions while the clip is played. Figure 2.4 presents the conceptual architecture of an adaptive multimedia system based on real-time transcoding. In the figure, the different quality versions of a multimedia clip are symbolised by the images of different sizes.

Despite that storage equipments have increased significantly in capacity, while their prices have continuously dropped, in practice it is feasible to store only a limited number of versions. Therefore the adaptation cannot be done with the same granularity as in case of real-time transcoding.

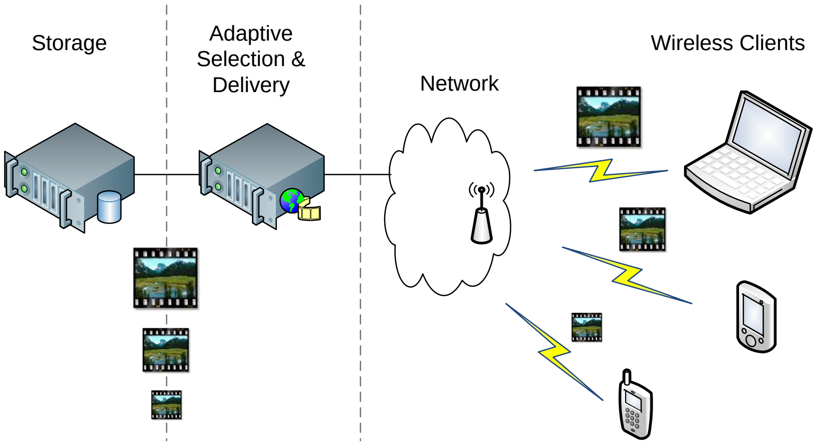

Figure 2.4: Selection Based Adaptive Multimedia

However, given that the selection process is not that resource hungry, in practice adaptive selection streaming proved to be more feasible than real-time transcoding, and currently a number of commercial implementations based on this technique are available. Microsoft IIS Smooth Streaming[17] and Adobe FMS Dynamic Streaming[18] are such streaming solutions that dynamically switch between several versions of the same clip based on viewers' available bandwidth. Both solutions have the drawback, of needing dedicated software for content publishing, as well as dedicated players which are currently supported only by a limited number of devices and platforms.

2.4.1.2 Multimedia Selection/Reduction

This method consists in selecting and reducing or deleting completely various elements from the multimedia clip. These elements can be for example, groups of frames, groups of pixels from particular frames, bit planes in pixels, etc.

---

[17] Microsoft IIS Smooth Streaming, http://www.iis.net/download/smoothstreaming
[18] Adobe Flash Media Streaming Server, http://www.adobe.com/products/flashmediastreaming/

Although the selection process, which is necessary in order to determine which elements to be deleted, can be as simple as uniform decimation in some cases, usually this is more sophisticated and often considers psychophysical factors such as, user interest and preferences, user domain knowledge, human eye perception, etc. (Chang & Vetro 2005).

Multimedia selection and reduction techniques are often applied in case of video summarisation. Video summaries are usually created by extracting "high priority entities and events" from the original videos, and presenting them with a "reasonable degree of continuity", "free of repetition" (Fayzullin et al. 2003). However, quite often, still images, as well as graphical representations and textual descriptions are added to video summaries (Money & Agius 2008), thus additional replacement or synthesis techniques being used.

2.4.1.3 Multimedia Replacement

Multimedia replacement consists in replacing selected elements in the multimedia clip with less expensive counterparts (Chang & Vetro 2005). A good example is replacing consecutive video sequences with representative still images (which can be extracted from the video or not), while keeping the audio, thus creating a slideshow. Replacement techniques can be used with very good efficiency when there are resource constraints (e.g. bandwidth limitations) or for video summarisation.

2.4.1.4 Multimedia Synthesis

Multimedia synthesis consists in analysing and extracting various elements from the original multimedia clip (e.g. sequences, frames, region of a frame image, movement patterns, etc.), and synthesising new content presentations by combining these elements (Chang & Vetro 2005).

The new presentations may be created with various goals, and these can be related with the original content by following the same flow, or can consist in completely unrelated presentations. An example from the first category would be converting a video from 2D to 3D (Ideses 2007), thus offering a new level of user experience. An example from the second category would be to generate new textures by extracting various elements from images or videos, such as groups of people, flowers, fruits, etc. (Kwatra et al. 2003).

2.4.2 Multi-Layer Approaches

The approaches presented below use multiple layers of information, which combined offer a better overall quality of multimedia for the end-user. There are two main techniques in this category: ***Scalable Coding*** and ***Multiple Description Coding***.

2.4.2.1 Scalable Coding

Scalability refers to the fact that the video is compressed in such a way that allows removing various parts from the original stream, thus obtaining valid substreams of lower quality (Schwarz et al. 2007).

The most common forms of video scalability are (Schwarz et al. 2007):

- ***temporal scalability*** – a substream represents the source content with a reduced frame rate;
- ***spatial scalability*** – a substream represents the source content with a reduced frame image size (resolution);
- ***quality scalability*** – a substream has the same spatio-temporal resolution as the complete bitstream, but a lower fidelity.

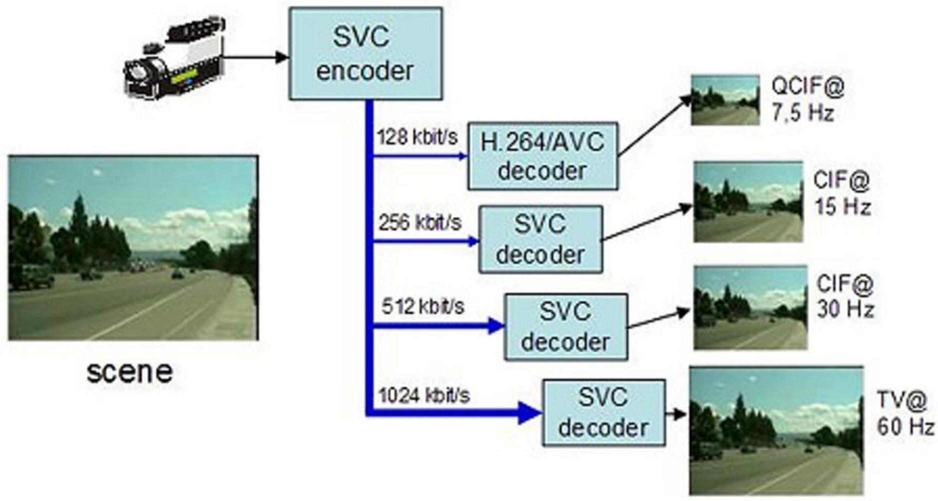

Figure 2.5: Scalable Video Coding (HHI n.d.)

Other less common types of scalability include, *region-of-interest* (ROI) and *object-based* scalability, in which cases, the substreams represent continuous regions from the original frames (Schwarz et al. 2007).

The most advanced implementation of scalable coding to date is the Scalable Video Coding (SVC), an extension of the H.264/MPEG-4 AVC video compression standard (ITU-T 2009), developed by the Joint Video Team (JVT)[19], a group of experts from ITU-T Video Coding Experts Group (VCEG) and ISO/IEC Moving Picture Experts Group (MPEG).

Figure 2.5 presents the principle used by SVC. The recorded material is encoded using a specialised encoder, in a single file that contains several layers of information, from which one is the ***base layer*** while the rest are ***enhancement layers***. The minimum quality is dictated by the base layer, and this can be increased by adding enhancement layers to the base layer. The standard provides

[19] Joint Video Team (JVT), http://www.itu.int/ITU-T/studygroups/com16/jvt/

backward compatibility, in the sense that the base layer can be decoded by existing H.264/AVC decoders. However, a SVC decoder is needed for enabling higher quality levels. Apart of the main scalability types (temporal, spatial and quality scalability), SVC supports ROI scalability, but this is restricted to regions that can be represented as a collection of macroblocks (groups of 16x16 pixels) (Schwarz et al. 2007).

Some of the benefits of SVC are: high granularity, the need for storing a single file from which multiple streams can be generated based on the device characteristics or transmission conditions, possibility for switching between different quality layers in each access unit (a video sequence that can be decoded independently), etc. However, currently very few implementations of SVC encoders and decoders are available.

2.4.2.2 Multiple Description Coding

Multiple Description Coding (MDC), is a technique that consists in splitting the original stream in two or more substreams known as descriptions, so that the first can be approximated from any subset of its descriptions (Goyal 2001).

For the particular case of multimedia streaming delivery, the goal of MDC is to improve the error resilience without using complex channel coding schemes (Vitali 2007). By delivering the descriptions of a multimedia clip independently, the end-user has the big advantage that, if a description is lost, s/he will be able to continue viewing the clip, with the cost of a decrease in the overall quality.

The principle of MDC is illustrated in Figure 2.6. In the figure, the original stream is divided in two descriptions, and each of them is sent to the client on an alternative route. In the first situation (see Figure 2.6a), both descriptions reach the client, thus maximum quality is perceived. In the second case (see Figure 2.6b) one description is lost due to network problems. As only the second description

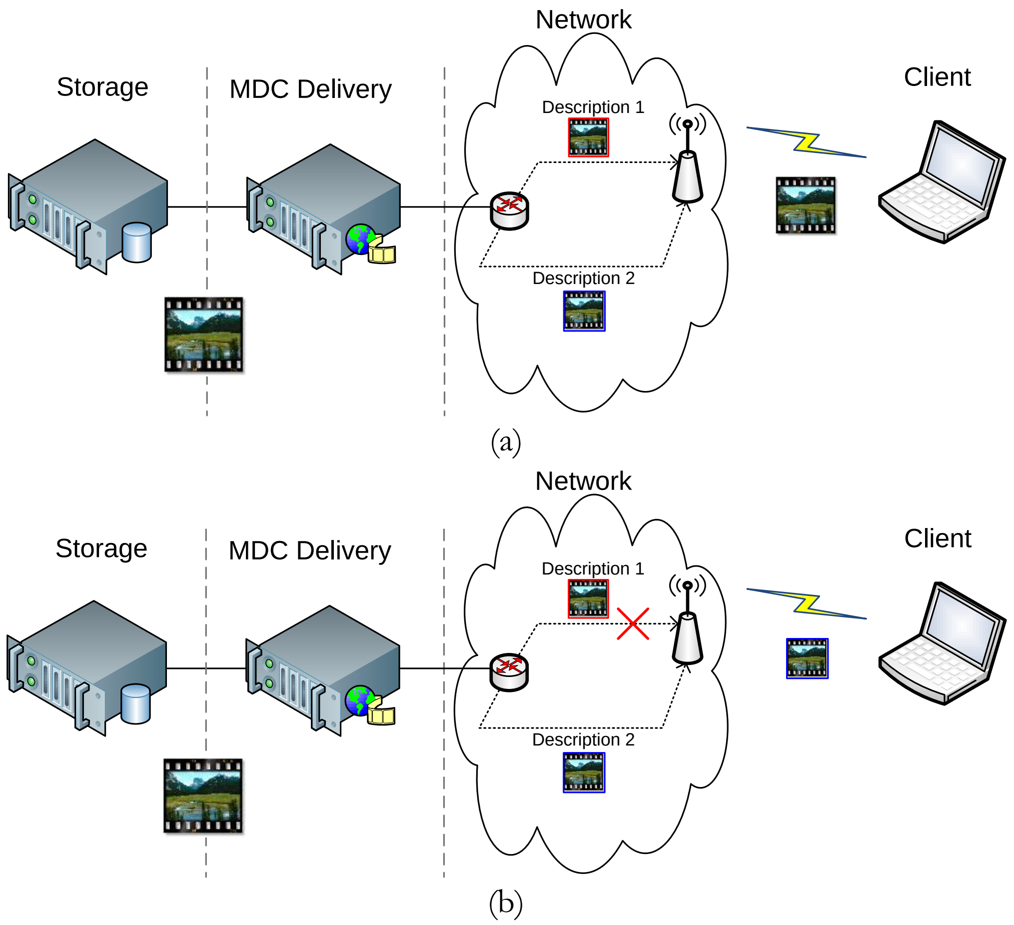

Figure 2.6: Multiple Description Coding
(a) Both Descriptions Reach the Destination; (b) One Description is Lost.

reaches the destination, the client is able to see the clip, but at a lower quality level.

2.4.3 Summary

This section presented an overview of the main approaches for multimedia content adaptation. Various reasons could be behind the adaptation such as: different devices with different characteristics and capabilities, different users have different network conditions, etc. There are two main approaches for adapting the multimedia content in such low resource situations. First is to reduce the perceived quality (e.g. transcoding, SVC, MDC), whereas the second is to make changes in the information presented (e.g. selection/replacement, reduce, synthesis).

2.5 Multimedia Content Classification

Given the high amount of multimedia content available nowadays, significant work has been put in classifying this content. Alike with *multimedia indexing*, the primary goal of *multimedia classification* is to help users navigate more easily through the content and find the one they are interested in. The similarities between video indexing and classification make difficult a complete separation between the two concepts. More often classification is seen as representing a part, a necessary step of multimedia indexing. The second is primarily concerned with the organisation of content from a database perspective, with the aim of retrieving more efficiently and accurately the content requested by the users (Smoliar & Zhang 1994). This is enabled by sorting the content into different categories by attaching meaningful labels that define their *genre* (e.g. ‘sports’, ‘news’, ‘commercials’, ‘movies’, etc.), *subgenre* (e.g. in case of sports content this could be ‘football’, ‘tennis’, ‘basketball’, etc.), etc. (Snoek & Worring 2005).

The classification can be done manually, in which case labels are associated to each clip when this is uploaded or at a later time, either by the content creator or by the content administrator. However this method requires extra work, which is often skipped, and also is prone to human subjectivity. Furthermore, since it is completely unfeasible to classify large collections of videos manually, various automatic techniques have been proposed in the literature. All of these techniques extract some sort of information from the multimedia clips, thus being called *content-based classification techniques*.

There are three main types of information that can be extracted, namely *textual*, *audio* and *visual* information (Brezeale & Cook 2008). Many of the classification algorithms that have been proposed combine more than one source of information. Furthermore, some of them concentrate on classifying the videos as a whole, whereas others concentrate on classifying independently different sections from a multimedia clip. In the second case *video segmentation* (Hampapur et al. 1994) techniques are used for partitioning the clip into homogenous spatial, temporal or spatio-temporal regions (Tekalp 2005).

Since proposing an automatic solution for educational multimedia content classification is out of the scope of this thesis, only an overview of the three main approaches for classifying the multimedia content, ***text-based***, ***audio-based*** and ***video-based***, is presented in this section. However, this overview will help the reader understand, how the different categories of educational multimedia clips that were used during the experimental and subjective studies, were selected looking at their characteristics in terms of content presented, as well as temporal and spatial movement.

2.5.1 Text-Based Approaches

Text-based techniques for classifying the multimedia content, extract textual information which is visible in the video or represents the transcription of a dialog (Brezeale & Cook 2008).

Examples from the first category include scene text (text that is embedded in the video itself, such as the text in a slideshow recording) and graphic text (text placed on the screen such as university logo, speaker name, subtitles, etc.). After the text elements are identified and separated from the rest of visual information, these are converted into text using Optical Character Recognition (OCR). An example, of OCR usage in e-learning is REPLAY, a system developed at ETH Zurich for “recording, handling, indexing, referencing, archiving and distributing” lecture recordings (Schulte et al. 2008). Using OCR, REPLAY extracts textual information from the recorded slides and indexes it using MPEG-7 standard, thus enabling fast access to slides of interest. The main drawback of OCR is that this it is limited to optical characters of fixed fonts and clear text lines, thus handwritten characters that often occur in classroom lecture recordings cannot be extracted (Choudary & Liu 2007). Therefore more sophisticated handwritten recognition algorithms have to be applied to extract such a text.

The second type of textual information being extracted is represented by the transcript of the dialogs in the clip. A possibility is to use speech recognition software for transcribing lectures recordings in particular (Repp & Meinel 2006), and any other videos in general. A simpler way is to extract the transcript from the closed captions if the clips are provided with them (Thompson 2007). As opposed to open captions and subtitles which are embedded in the video, closed captions are separated from the video and can be read using specialised decoders (WebAIM n.d.).

The big advantage of text-based approaches is that once having the text extracted, same techniques as for the case of document classification can be used, for assigning labels to multimedia clips or their sequences (Brezeale & Cook 2008).

2.5.2 Audio-Based Approaches

Audio-based techniques, classify the multimedia clips by analysing various features that can be extracted from the audio signal, represented either in the time domain (signal amplitude vs. time), or in the frequency domain (signal amplitude vs. frequency) (Brezeale & Cook 2008).

An example of feature that can be extracted in the time domain is z*ero crossing rate* (ZCR), which represents the number of times the audio waveform crosses the zero axis each second (Rouvier et al. 2010). By mapping the ZCR to a set of predefined values and patterns, a multimedia clip can be classified as belonging to a particular genre (e.g. ZCR varies higher in speech than in music).

An example of feature that that can be extracted in the frequency domain is the *bandwidth*, or the signal frequency range. Because different sounds have different frequency ranges, a clip can be associated to a predefined class (e.g. speech has a narrow bandwidth than music).

Audio-based techniques have the advantage of requiring significantly less computational resources than the video-based techniques.

2.5.3 Video-Based Approaches

Most of the approaches for classifying multimedia content fall in this category. Visual features can be extracted on a *per-frame* or on a *per-shot* basis, where a shot represents the collection of frames belonging to the same camera action (Brezeale & Cook 2008). Many of the video based techniques make use of features that correspond to cinematic principles such as: using shots length for detecting the flow of the video, colours for detecting light levels, motion for detecting action, etc. (Brezeale & Cook 2008).

Shots are a good way for segmenting the videos and can be represented through a single frame called keyframe. Some information regarding the type of a movie clip can be extracted from the length of the shot. For example, shots tend to be shorter in case of advertisements than in case of other genres such as cartoons, movies, news or sports (Chen et al. 2010). However detecting the shots can be often very difficult due to the multitude of transitions that can be applied during the editing process (e.g. hard cuts, fade in / out, mix / dissolve / crossfade, wipe, etc.)

The colour histogram, which represents the distribution of colours in a frame, can be used for detecting the light level which can be associated with a specific genre (e.g. horror movies tend to be low lit whereas the comedies tend to be well lit) (Brezeale & Cook 2008). The boundary between shots can also be detected by comparing the colour histograms of consecutive frames. If the difference is above a predefined threshold, a transition is assumed (Mittal et al. 2006). Various histogram parameters such as mean of histogram, standard deviation of histogram, mean of absolute difference between histograms, etc. can be combined for better identifying clips belonging to a specific genre such as cartoons, news, movies or sports (Chen et al. 2010).

Motion-based features can also be explored for video classification. There are two types of motion that can be analysed: *motion of the objects being filmed* and *motion of the camera* (Brezeale & Cook 2008). For video classification purposes, various researchers have explored object movement patterns (e.g. Fleischman et al. 2006) and human actions, such as walking, running, jogging, boxing or hand waving (e.g. Niebles & Fei-Fei 2007; Dhillon et al. 2009). Wu et al. (2010) have explored the idea that the video genre is inheritably linked with the video capturing scenario (e.g. multiple cameras used to film a professional video, whereas a single camera is used to film amateur videos), and have introduced two new classification features: the number of cameras and the distance of the subject from the camera.

2.5.4 Summary

This section presented an overview of the main approaches for classifying multimedia content. Classification features can be extracted based on three types of information that can be presented in the multimedia clip: textual, audio and video information. Most of the techniques in the literature have concentrated on classifying the videos in broad domains (e.g. news, movies, commercials, etc.), or on classifying videos from the same domain according to their genre (e.g. movies in drama, horror, action, cartoon, etc.). Other approaches have concentrated on identifying specific scenes such as explosions in movies or goals in sport videos. What is characteristic for all of them are the need to predefine the categories.

Some of the techniques broadly presented here have also been used for the particular case of educational videos with goals such as summarisation of whiteboard content in lecture recordings (Choudary & Liu 2007) or indexing the content of lecture recordings for separating fragments corresponding to various elements such as definitions, theorems, discussions, reviews, etc. (Mittal et al. 2006).

2.6 Assessment of Multimedia Quality

Learner QoE has been shown to be an important factor in e-learning (Muntean & McManis 2006). As opposed to the Quality of Service (QoS), which objectively measures the service provided, QoE is a subjective measure of the users' experience, their expectations and their satisfaction with the service. Therefore, QoS is a technology-cantered measurement, whereas QoE is a user-cantered measurement (Hestnes et al. 2008).

For the particular case of multimedia clips that contain both audio and video content, there are a multitude of factors that contribute to the overall QoE. Jumisko-Pyykkö et al. (2007) have identified the following five major categories of quality related factors:

- ***Content factors***, e.g. news, animation, sport, music videos, etc.
- ***Usage factors***, e.g. purpose of use, device used, etc.
- ***Visual quality factors***, e.g. details, blurriness, motion, colours, text readability, etc.
- ***Audio quality factors***, e.g. background sound, clarity, echo, speech, disturbing inferiority, etc.

- ***Audiovisual quality factors***, e.g. audio quality more important than video, synchronism.

Furthermore, the impact of each of these factors can vary from one clip to another depending on the type of content being presented (Jumisko-Pyykkö et al. 2007). For example, video quality is more important than audio quality in case of sport type clips, whereas audio quality is more important than video quality in case of news type clips.

This section looks at the existing approaches for assessing the video quality of the multimedia clips, which can be classified in:

- subjective methods;
- objective methods.

2.6.1 Subjective Assessment of Video Quality

Subjective methods are the most common methods for assessing the video quality and are considered to be the most accurate ones (Winkler & Mohandas 2008). During subjective testing, a group of subjects are asked to view a set of video sequences and to rate their quality on a specific scale. Numeric scales, such as the five-point scale, are the most often used ones. The scores assigned by all the observers for a specific clip, are averaged in order to obtain the Mean Opinion Score (MOS).

The subjective methods differ in general through the way the video sequences are presented to the observers. Subjective tests are affected by the fact that human observers perceive the quality in different ways. In order to achieve relevant MOS values, the subjectivity and variability of the observer ratings, have to be minimised through a careful selection of the video sequences. Also the evaluation has to follow strict procedures and to take place in controlled environments.

Procedures for subjectively assessing the video quality have been standardised and are thoroughly described in ITU-R recommendation BT500-11 (ITU-R 2002). These methods are divided in two categories:

- Double Stimulus (DS) – viewers rate the difference in quality between an impaired clip and an unimpaired clip, considered as reference;
- Single Stimulus (SS) - viewers are presented with a sequence of impaired videos and they have to rate their overall quality on each of them.

The main advantage of DS methods is that they reduce the context effects on the observer ratings (Pinson & Wolf 2003). However, given the limited duration a subjective test can take place, only a reduced number of video sequences can be evaluated. SS methods were proposed to allow for a higher number of video clips to be rated in the same amount of time.

Examples of DS methods include the Double Stimulus Continuous Quality Scale (DSCQS) and the Double Stimulus Impairment Scale (DSIS), while an example of SS method is the Single Stimulus Continuous Quality Evaluation (SSCQE).

An important advantage of the SSCQE method it that allows continuous rating of the video material presented to the observers for the entire duration of the test (Corriveau 2005b). This is done with the help of a slider that the observers can adjust immediately they notice a variation in the video quality. A computer program is continuously recording the slider position throughout the test, creating a viewer’s rating trace.

Since the test conditions are very similar to those corresponding to a real viewing situation, SSCQE presents a higher potential for assessing the user’s QoE. Although questions have been raised regarding the accuracy of SSCQE method as compared to DSCQS method, Pinson and Wolf (2003) have shown that the results achieved using the two methods, are statistically indistinguishable.

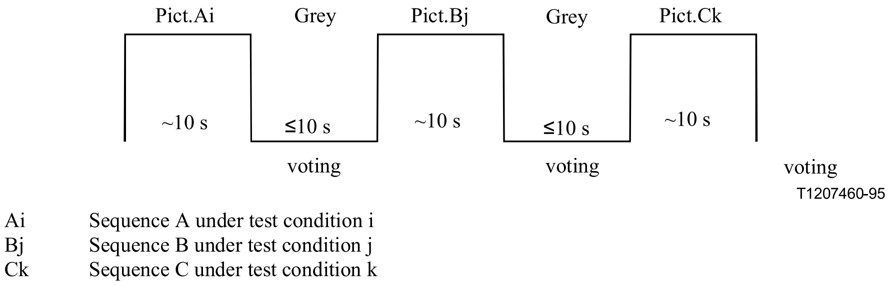

Figure 2.7: Stimulus Presentation for ACR Method (ITU-T 2008)

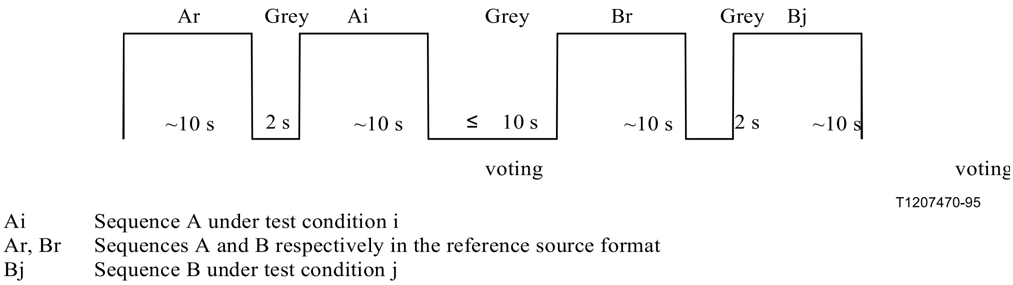

Figure 2.8: Stimulus Presentation for DCR Method (ITU-T 2008)

Subjective methods have also been proposed and standardised for assessing the video quality of multimedia applications (ITU-T 2008). These methods also address mobile devices special requirements, such as reduced size and short viewing distance. Examples include ***Absolute Category Rating (ACR)***, which is a SS method, and ***Degradation Category Rating (DCR)***, which is a DS method.

ACR is a subjective assessment method where the test sequences are displayed one at a time and are rated independently on a on a category scale. The stimulus presentation for ACR method is shown in Figure 2.7. Each sequence is normally displayed for 10 seconds or less, but this time can be decreased or increased depending on the content of the test material. If a constant voting time is used than this should be no more than 10 seconds. The rating is done on a 5-level quality scale, where 1 corresponds to Bad, and 5 to Excellent as seen in Table 2.1. If a higher granularity is necessary, a 10-level quality scale can also be used.

Table 2.1: Five-level Quality Scale for ACR method

Grade	***Estimated Quality***
5	Excellent
4	Good
3	Fair
2	Poor
1	Bad

Table 2.2: Five-level Impairment Scale for DCR method

Grade	***Estimated Quality***
5	Imperceptible
4	Perceptible but not annoying
3	Slightly annoying
2	Annoying
1	Very annoying

DCR is a subjective method where the test sequences are displayed in pairs. The first sequence in each pair is the source reference, while the second is the one under test (for which the quality is measured). When possible, the two sequences can be also displayed simultaneously on the same monitor. The stimulus presentation for the DCR method is presented in Figure 2.8. The presentation times for each sequence, as well as the voting times, are the same like for the case of ACR method. The rating is also done on a fife-level quality scale, but this time the subjects are asked to rate the quality of the second sequence from each pair, as compared to the quality of the first sequence (see Table 2.2).

To limit the time of the subjective testing bellow the recommended duration of less than 30 minutes, the ACR method was used for the subjective study presented in this thesis.

The standardised methods for subjectively assessing the video quality are widely accepted and used both by the industry and the research communities. Therefore little research has concentrated on proposing new methods or improving the existing ones. Richardson and Kannangara (2004) have proposed a solution to increase the speed of the subjective assessment methods for situations such as finding the optimum choice from a set of alternative versions of a video clip. By incorporating a feedback mechanism that allows the observers to select their preferred version, their solution has achieved the same results as a well established subjective method, but in a fraction of the time required by the last.

Since the subjective methods require a large number of subjects to achieve good results, their preparation and running are very time consuming. The evaluation is even more difficult if the tests have to be repeated several times. Various objective methods have been proposed for assessing the video quality faster, and without the need for special resources.

2.6.2 Objective Assessment of Video Quality

Objective methods predict the viewer MOS by using different algorithms to represent the video quality as a function of measurable parameters. The main advantages of these methods are that they are automatic and readily repeatable.

The majority of the proposed methods have focused on measuring the distortions introduced by various processing steps, in particular the lossy compression (Winkler 2007). Winkler and Mohandas (2008) have classified the possibilities for measuring the distortions introduced in a video delivered over an IP packet network in:

- *Data metrics* – they measure the fidelity of the signal without considering its content;
- *Picture metrics* – the video data is treated as the visual information it

contains;

- *Bitstream-based metrics* – they look at the packet header information without fully decoding the stream.

Another possibility to classify the objective video quality metrics is based on the availability of original distortion-free information, considered as reference (Wang et al. 2003). According to this approach the following categories exist:

- Full-Reference metrics (FR);
- No-Reference metrics (NR);
- Reduced-Reference metrics (RR).

The most widely used FR data metrics are ***Peak Signal-to-Noise Ratio (PSNR)*** and ***Mean Squared Error (MSE)***. PSNR is a mathematical metric that measures the ratio between the maximum possible power of a signal with the power of the noise signal. PSNR is computed as a logarithmic expression based on MSE, and it is expressed in dB. The main advantages of these metrics are the fact that are fast to compute, and are easy to be understood and implemented (Wang et al.2003).

However PSNR has been found to not reflect accurately the subjective perceptual evaluation (Nemethova at al. 2004). This is because PSNR just compares the data at byte level, ignoring completely how human vision works (Winkler 2005).

Due to the limitations of data metrics, much research has been put into creating more accurate picture metrics that concentrate on the effect of distortions and content on the perceived quality (Winkler & Mohandas 2008). In particular, MSE has been shown to be one of the best measures for additive noise (Avcıbaş et al. 2002). However vision-based metrics that incorporate factors such as colour perception, contrast sensitivity or pattern masking, reflect better the subjective quality perception than MSE. Examples of objective metrics that consider the human perception are ***Structural Similarity Index (SSIM)*** and ***Video Quality Metric (VQM).***

SSIM, which was introduced by Wang et al. (2004a), is based on the idea that the human visual perception is adapted to extract structural information. The SSIM index is measured based on three different components which are measured independently: luminance similarity, contrast similarity and structural similarity. The result can take values between 0 to 1, 1 being the maximum value (no difference in quality between the two videos compared). SSIM has been shown to have a high correlation with subjective scores both in case of image (Wang, Bovik et al. 2004a) and video quality assessment (Wang, Lu et al. 2004b).

VQM (Pinson & Wolf 2004) is an objective metric that has been shown to also have a high correlation with human perception. The VQM index is normally measured on a scale from 0 to 1, when 0 corresponds to the highest quality level (no difference in quality between the two videos compared). However, sometimes the values can go above 1, if one of the videos is much degraded as compared to the other one. The method was standardised and adopted as internationally as ITU-T Recommendation J.144 (ITU-T 2004).

Apart of the vision characteristics, picture metrics may predict the quality based on parameters extracted from the video, such as blockiness, blur and jerkiness (Susstrunk & Winkler 2004), or motion patterns (Ries et al. 2007).

More recently, technological advancements have contributed to the appearance and the growth of an increasing number of services that stream multimedia content over the IP networks. In this context, assessing the impact of the network on the perceived quality has increased in importance. Bitstream-based metrics rely mostly on indicators that can be extracted from the packet stream, without decoding it (Winkler & Mohandas 2008). Using pertinent factors such as motion, Kanumuri et al. (2006) have proposed a solution to predict if a lost packet will be visible in the MPEG-2 video and if yes with what probability.

With the increasing number of approaches for objectively assessing the video quality, the need for standardised methods has appeared. Video Quality Experts Group[20] was created in 1997 with the goal to narrow down the selection of objective methods when it comes to monitoring the video quality (Corriveau 2005a).

2.6.3 Summary

This section has outlined various approaches for assessing the video quality of multimedia clips. Subjective methods are widely considered a better method for determining the video quality in particular, and the viewers QoE in general. Objective methods have the benefit of speed and repeatability, and since they usually have good correlation with subjective methods, in often cases they can be used to complement or replace the subjective methods.

2.7 Chapter Summary

This chapter has presented the research areas related to this project, as well as the technical and theoretical background that is required to understand where exactly the research presented in this thesis, fits in the context of adaptive mobile learning.

[20] The Video Quality Experts Group Web Site, http://www.its.bldrdoc.gov/vqeg/

In the beginning, the current state of the art in adaptive m-learning, as an evolution of adaptive e-learning, was presented. Following, the current state of multimedia content usage in e-learning and m-learning was presented.

Since the research presented in this thesis is intended to support educational multimedia content adaptation in m-learning systems, the main approaches and techniques for adapting multimedia content in general were also overviewed. Adaptive approaches that select dynamically between multiple versions having different levels of quality are mainly targeted by this research.

Multimedia content classification was also overviewed, as this is necessary to understand why different categories of educational clips were addressed by the experiments presented in this thesis. Much research has been conducted in this area, and various approaches to group multimedia content in genres such as news, sports or cartoons, have been proposed. However, few of this research have addressed multimedia educational clips in particular. The research presented in this thesis is the first to address multiple genres of multimedia educational clips, which differ by the type of the visual information being presented, as well as by the level of dynamicity.

As the most important contribution of this thesis addresses the impact of multimedia educational content quality on the learning process, the methods for assessing the video quality were also overviewed.

Chapter 3

Video Profiling of Educational Multimedia Content

3.1 Introduction

This chapter presents some of the main contributions of this thesis. In order to enable mechanisms for adapting multimedia clips to be used with m-learning systems, this research proposes to group mobile devices in several classes depending on their screen resolution. Video profiles offering excellent quality levels, are defined and associated to each class of devices. Recommendations regarding the optimum values for the encoding parameters of the multimedia clips such as bitrate, frame rate and resolution are also made.

The chapter is structured as follows. At the beginning, Section 3.2 presents an analysis on various categories of mobile devices that can be used for conducting mobile learning activities, briefly overviewing their particularities in terms of form factor, screen size and screen resolution. Following, the results of a study assessing the most common encoding schemes used for educational multimedia clips are presented in Section 3.3. The proposed novel solution for video is presented in Section 3.4, whereas Section 3.5 concludes the chapter.

3.2 Categories of M-learning Devices

To further understand the need for device-specific adaptation of multimedia educational content, this section overviews various categories of mobile devices that are most appropriate for conducting mobile learning activities. The main characteristics of these devices are also presented.

3.2.1 Form-Factors and Screen Sizes

Currently there are a multitude of mobile devices that can be used for conducting e-learning activities. All of these so called ***"m-learning devices"*** (Quinn 2008), have in common the Internet connectivity. However, there are many differences in terms of other characteristics, such as the device size and its form-factor, the screen size and its resolution, the CPU speed, the memory capacity, the I/O components, etc.

After excluding Feature Phones with Web connectivity and Laptops, the remaining of the mobile devices can be classified roughly, in two big categories: ***Smartphones*** and ***Ultra Mobile Devices (UMDs)***. As compared to Feature Phones, these devices are more appropriate for m-learning, due to their advanced web-browsing and multimedia capabilities. At the same time their reduced weight and size, enables increased mobility as compared to Laptops. Furthermore, it is estimated that by 2014-2015, combined early global sales of such devices will top 1 billion (Abi Research 2009, Coda Research Consultancy 2010). Apart of these devices there are other categories of devices with more dedicated functionalities that have wireless connectivity and thus can be used for mobile learning. Such devices include Portable Media Players (PMPs) or Handheld Game Consoles (HGCs).

To classify these devices even further, a traditional approach is to consider their size and form-factor. Figure 3.1 presents the latest categories of mobile

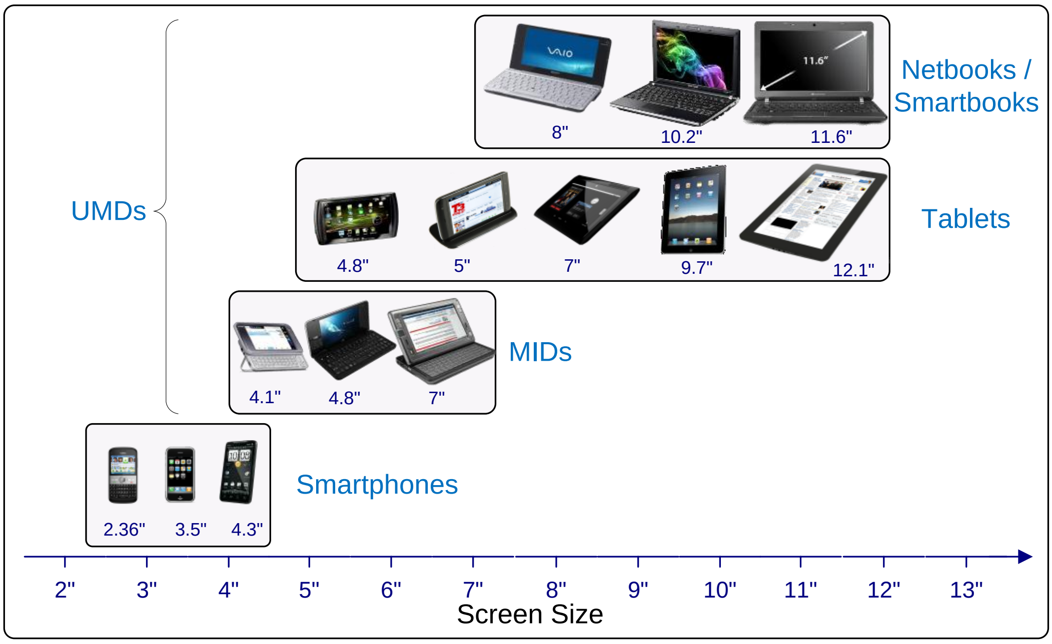

Figure 3.1: Latest Categories of M-Learning Devices

devices and how these partially overlap in terms of screen size. While smartphones usually have the screen size between 2” and 5”, and can present a physical keyboard or a touch-screen as input device, UMDs cover all the multitude of mobile devices between Smartphones and Laptops known under a variety of names such as Mobile Internet Devices (MIDs), Tablets, Ultra Mobile PCs, Netbooks, Smartbooks, etc. In terms of screen size UMDs can go as low as 4” and as high as 12”.

In different situations however, a particular device can belong to more than one category. This is caused by the technological convergence, and in particular mobile device convergence, that sees different devices performing similar tasks. For example, a 4”-5” tablet that is able to make calls thanks to its 3G connectivity, can be easily considered as being a smartphone. Furthermore, there is an increasing trend that sees different devices sharing hardware components as well as operating systems (e.g Apple’s iPhone 4 smartphone and iPad tablet, share the iPhone OS and the A4 CPU).

3.2.2 Device Screen Resolutions

A summary of the most common video resolutions across different classes of m-learning devices is presented in Table 3.1 (the list was compiled using information available on the Cartoonized[1] and UMPC Portal[2] websites).

There are several trends regarding the device screen resolutions. A first one is the fact that high resolutions are becoming increasingly popular on small screen devices such as smartphones. Also up to now, the pixel density had continuously increased, so that it crossed 300ppi (pixels/inch), when Apple introduced the Retina Display[3] on its iPhone 4 smartphone, with a pixel density of 326 (960x640 resolution, 3.5" screen size). A third trend is the one that sees wide resolution screens with aspect ratios such as 16:9 or 16:10, becoming increasingly popular.

3.3 Assessment of Multimedia Educational Clips Encoding Schemes

This section presents one of the main contributions of this thesis, an investigation aimed to assess the most common encoding schemes that are used for converting the educational multimedia clips, before making them available to learners. For this purpose, a significant number of educational clips (904), of different duration and file size, were downloaded from the Internet, and their encoding characteristics were analysed. The next subsections present, the data collection procedure, as well as the results of the analysis. In the end conclusions are drawn.

---

[1] Cell Phone Screen Resolution by Brand and Model, http://cartoonized.net/cellphone-screen-resolution.php
[2] Netbook, Tablet, UMPC and MID Comparison, http://www.umpcportal.com/products/
[3] Apple - iPhone 4 - Learn about the high-resolution Retina display, http://www.apple.com/iphone/features/retina-display.html

Table 3.1: Common Display Resolutions for M-learning Devices

Display Resolution	***Aspect Ratio***	***Standard Name***	***Occurrence***
1600 x 768	25:12	UWXGA	Netbooks
1366 x 768	16:9	WXGA	Netbooks, Tablets
1280 x 800	16:10	WXGA	Netbooks, MIDs
1280 x 768	5:3	WXGA	Netbooks
1024 x 768	4:3	XGA	Netbooks, Tablets
1024 x 600	16:10	WSVGA	Netbooks, Tablets, MIDs
1024 x 576	16:9	WSVGA	Netbooks
1024 x 480		UWVGA	Smartphones, MIDs
960 x 640	3:2		Smartphones
960 x 480	2:1	UWVGA	Smartphones, MIDs
800 x 600	4:3	SVGA	Netbooks, Tablets
854 x 480	16:9	FWVGA	Smartphones, MIDs
800 x 480	5:3	WVGA	Smartphones, PMPs, MIDs, Netbooks, Tablets
640 x 480	4:3	VGA	Smartphones, PDAs,
640 x 360	16:9	nHD	Smartphones
480 x 360	4:3	HVGA+	Smartphones
480 x 320	3:2	HVGA	Smartphones, PMPs, MIDs
480 x 272	16:9	HD1080/16	Smartphones, PMPs, HGCs
480 x 240	2:1		Smartphones
432 x 240	9:5	WQVGA	Smartphones
427 x 240	16:9	WQVGA	Smartphones
400 x 240	5:3	WQVGA	Smartphones
320 x 240	4:3	QVGA	Smartphones, Feature Phones, PMPs, PDAs, MIDs

3.3.1 Data Collection

A number of 904 multimedia clips, tagged as educational videos were downloaded from the Internet using two different media players, iTunes[4] and Miro[5]. Out of the total number, 80% were downloaded from iTunes U, whereas the rest of 20% were downloaded from Miro Guide[6].

The difference in size between the two samples can be explained by the difference between the two sources in terms of the total number of educational clips available on each one of them. On Miro Guide there are 841 collections of educational videos as of August 2010. As opposed, although an exact number for the total number of collections available on iTunes U, was not found, the fact that currently there are more than 600 universities with active iTunes U sites, is a clear indicator that much more videos are available on iTunes than on Miro Guide. While some of the universities on iTunes may have no collections of resources available, others have tens or hundreds collections. For example, on the Open University's iTunes page, there are 326 albums containing 2,778 tracks, from which 1,253 are audio clips, and 1,525 are video clips as of August 2010 (Open University 2010). To date there are more than 350.000 free lectures, videos, films and other resources on iTunes U (Apple 2010).

The reason for downloading educational clips from Miro Guide also, was motivated by the fact that these clips come from multiple websites on the Internet and not from a single source as for the case of iTunes, in which case the videos are stored by Apple. Therefore, a better idea can be made on what are the most

[4] Apple - iTunes - Download iTunes Now, http://www.apple.com/itunes/
[5] Miro | Video Player | Free video and audio podcast player, .http://www.getmiro.com/
[6] Miro Guide - Video Podcast Directory, https://miroguide.com/

common encoding characteristics of the educational clips available on the Internet.

3.3.2 Data Analysis

After the educational multimedia clips were downloaded metadata was extracted using MediaInfo[7] open source software. From the multitude of encoding parameters presented in the extracted metadata, video compression, video resolution, video frame rate and video bitrate were used in this research for extracting information on the most common video encoding schemes currently used.

3.3.2.1 Audio-Video Compression

Regarding the video compression, the results presented in Figure 3.2 show that two codecs from the MPEG-4 family were used in case of more than 99% of the educational multimedia clips being studied. The first one, H.264/MPEG-4 AVC was accounts for 89.60% of the videos studied, whereas the MPEG 4 Visual also known as MPEG-4 Part 2 accounts for 9.85% of the clips.

H.264/MPEG-4 AVC (ITU-T 2009), is the latest version of MPEG-4 codecs, and it is considered to provide approximately 50% bit rate saving as compared to previous generation, MPEG-4 Part 2, for the same quality level (Wiegand 2003). The format was developed with a wide range of applications in mind (Richardson 2003), and over the past years has become increasingly supported by digital television standards such as DVB[8], as well as by providers of Internet video

[7] MediaInfo supplies technical and tag information about a video or audio file. http://mediainfo.sourceforge.net/en

[8] DVB - Digital Video Broadcasting – Home, http://www.dvb.org/index.xml

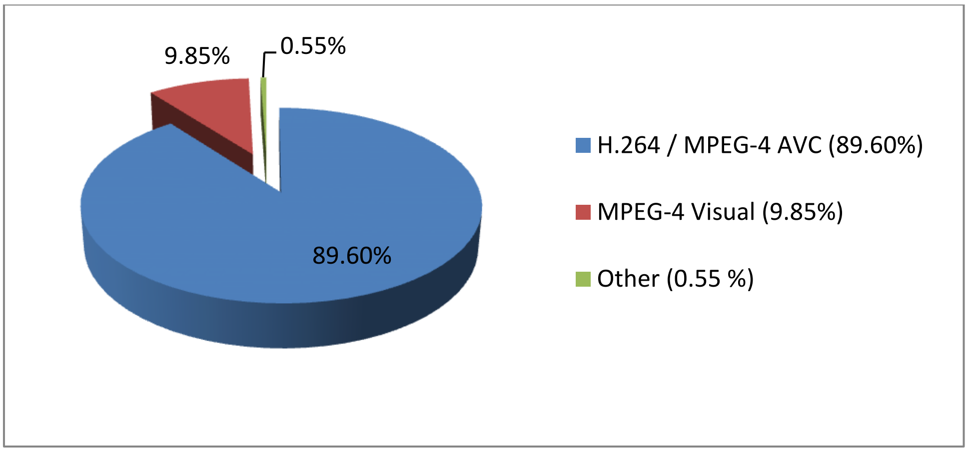

Figure 3.2: Most Common Video Codecs

services such as IPTV, Video on Demand, video conferencing or video sharing (e.g. YouTube, Vimeo). At the same time, an increasing number of mobile devices come with H.264 decoding capabilities.

H.264 / MPEG-4 AVC video compression is usually used in combination with the AAC (Advanced Audio Codec) audio compression. The two data streams, audio and video can be multiplexed in a multitude of multimedia containers, from which common used ones include .mp4, .mov, .m4v, etc. For more than 97% of the videos being studied, AAC audio codec was used.

3.3.2.2 Video Resolution

For the studied sample of 904 educational multimedia clips, 58 unique resolutions were identified, where 320x176 is the lowest one and 1440x1080 was the highest one. A distribution of these resolutions in a X-Y scatter diagram, is presented in Figure 3.3, where on the X axis is represented the video resolution width and on the Y axis is represented the video resolution height, in number of pixels. Since a single video had the highest resolution (1440x1080), this value was left out for a better visualisation. Figure 3.4 presents the percentage of videos having different resolutions. The individual resolutions are represented on the X axis and their occurrence on the Y axis. All the resolutions having less than 0.5% percent of the

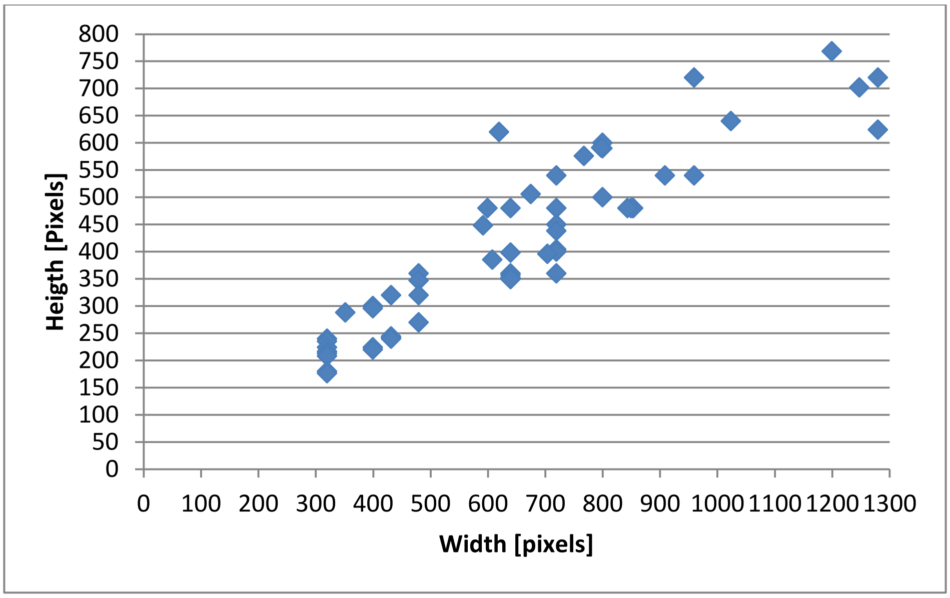

Figure 3.3: Distribution of Video Resolutions

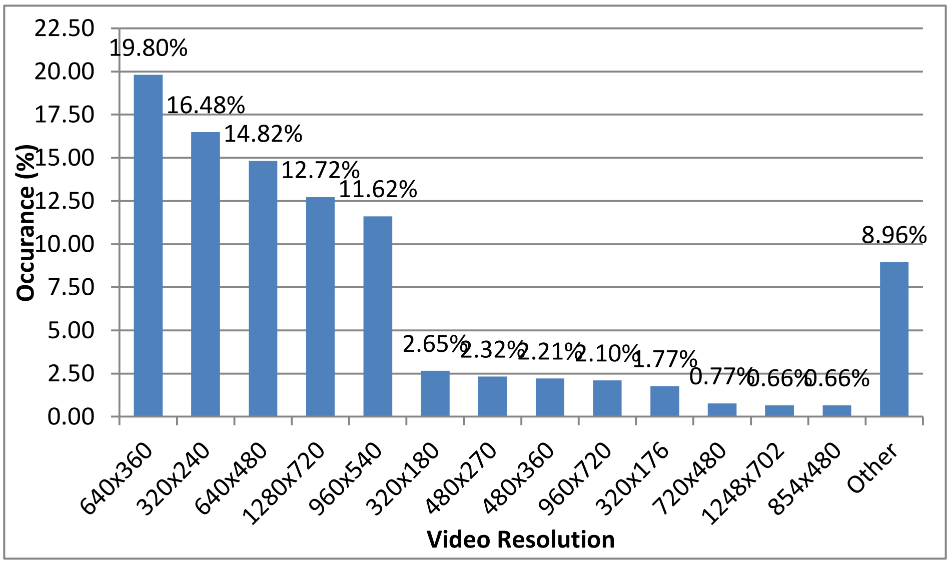

Figure 3.4: Most Common Video Resolutions

total number of videos were grouped as “Other”. As it can be seen in figure, the first five of them account for more than 75% of the multimedia clips being studied. The most common is 640x360 with 19.80%, followed of 320x240 with 6.48%, 640x480 with 14.82%, 1280x720 with 12.72%, and respectively 960x540 with 11.63%. The resolution’s category “other” standing with a percentage of 8.96 % contained a number of 81 educational clips.

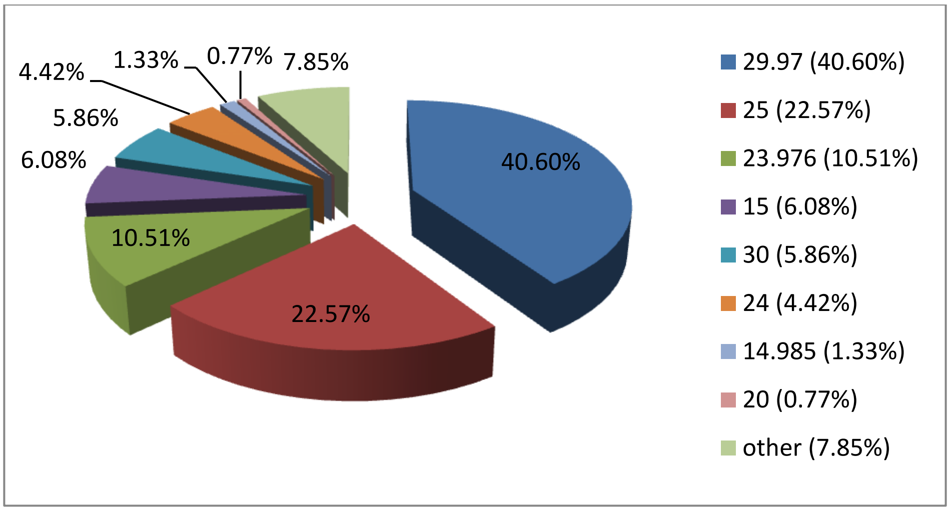

Figure 3.5: Most Common Video Frame Rates

3.3.2.3 Video Frame Rate

A total number of 56 individual frame rates were observed across the 904 educational clips. The lowest one was 1.908 fps (frames per second), whereas the highest one was 30.273 fps.

A distribution of the number of videos for the first 8 frame rates is presented in Figure 3.5. The frame rates with less than 0.5% were grouped as “Other”. The most used frame rate is 29.97 fps accounting for 40.60% of the educational clips being studied. This is followed by the 25 fps with 22.57% and 23.976 fps with 10.51%.

3.3.2.4 Video Bitrate

Since five resolutions, 320x240, 640x360, 640x480, 960x540, 1280x720, out of which two are very close together, the educational clips were grouped in 4 clusters around the representative resolutions having the highest number of videos, as seen in Figure 3.6.

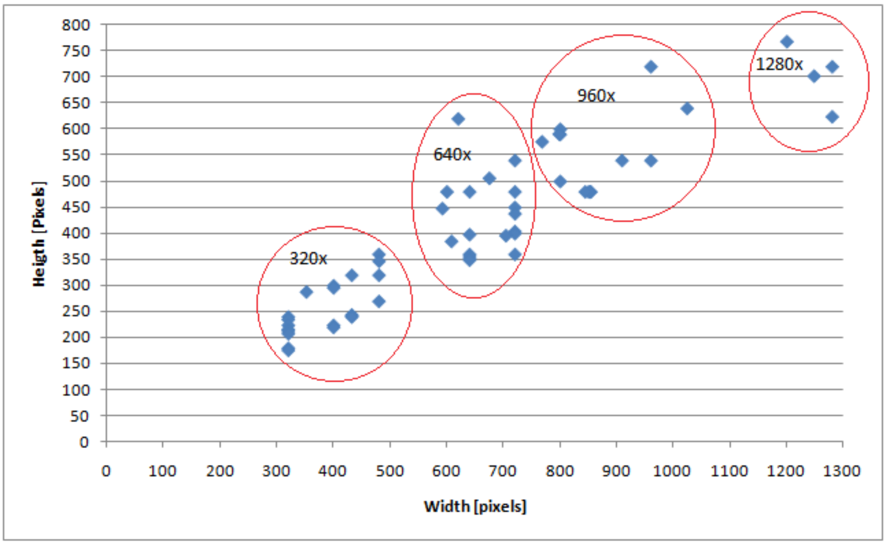

Figure 3.6: Clustering of Video Resolutions

For the first cluster, **320x** the lowest resolution is 320x176, and the highest is 480x360. Out of the 256 clips belonging to this cluster, 196 clips (~76%) have the resolution width 320 pixels, 46 clips (~18%) have the width 480 pixels, and for the rest of 14 (~0.05%) the resolution width is somewhere in between. The video height is variable.

For the first second cluster, **640x** the lowest resolution is 592x448 and the highest is 720x540. Out of the 377 clips belonging to this cluster, the video resolutions with 640 pixels width account for 345 clips or ~91%.

For the first third cluster, **960x** the lowest resolution is 768x576 and the highest is 1024x768. Out of the 147 clips belonging to this cluster, the video resolutions with 960 pixels width account for 132 clips or ~83%.

For the first third cluster, **1280x** the lowest resolution is 1200x640 and the highest is 1280x720. Out of the 124 clips belonging to this cluster, the video resolutions with 1280x720 accounts for 116 clips or ~94%.

Table 3.2: Distributions of Bitrates for the Four Clusters

Video Bitrate [Kbps]				
	Min	**Max**	**Avg**	**STDEV**
320x	54	1466	511	259
640x	6	8885	1142	823
960x	79	9478	2490	1404
1280x	200	16778	3391	2861

The minimum, maximum and average video bitrates for the 4 clusters, along with the standard deviations are presented in Table 3.2. As it can be seen in the table, the values for the video bitrate vary a lot in between the minimum and the maximum levels.

3.3.3 Summary

This section has presented the results of a study assessing the most common encoding schemes used for educational multimedia content. As it can be seen from the results, currently there are large variations in terms of encoding characteristics such as video resolution, frame rate or bitrate. However, some values for the video resolution and the frame rates are significantly more popular than the majority ones. As opposed there are a lot more variations for the video bitrate, since it is dependent on other encoding characteristics, such as compression used, resolution, frame rate, colour level, etc.

3.4 Proposed Solution for Video Profiling

This section presents the proposed solution for video profiling of the educational multimedia content. At the beginning video profiles are proposed targeting different device screen resolutions, and reference encoding settings for the optimum quality level are proposed. An example on how mobile devices can be grouped in different classes associated to the video profiles is also presented.

3.4.1 Video Profiles

In the context of this thesis, a ***video profile*** is a set of reference values for the video encoding parameters such as compression, resolution, bitrate and frame rate, that are used for encoding educational multimedia clips, for a target group of devices. The encoding parameter that defines each video profile is the video resolution, whereas the reference bitrate and frame rate are selected in such a way that the overall combination offers an optimum level of perceived quality, when the clip is displayed on a device from the target class. The proposed video profiles are defined for the H.264/MPEG-4 AVC video compression because currently this is the most advanced and used one.

3.4.1.1 Selecting the Representative Video Resolutions

Given the multitude of device screen resolutions that are common among m-learning devices, as well as the differences in their aspect ratios, in practice it is very difficult for adaptive mechanisms to target all of them. This is even more complicated when considering the different aspect ratios at which the educational clip may be recorded. Because of these facts, a selection of the target resolutions usually has to be made, for the adaptive mechanisms to be feasible in practice.

In order to decide which aspect ratio (e.g. 3:2, 4:3, 16:9, etc) to be considered from the most popular ones two current trends have been considered. The first one is that the mobile devices are increasingly adopting wide screen variations such as WVGA or WXGA. The second one is that, high-definition recording at 1280x720 and 1920x1080 resolutions has become increasingly popular, from smartphone cameras to professional ones. Following these trends, the decision to concentrate on the 16:9 wide aspect ratio was made.

In order to select representative resolutions an initial decision was made to limit the number of profiles to 4. Having mobile devices with screen resolutions as high as UWXGA 1600x768 (e.g. Sony Vaio P with 8” screen), the HD resolution 1280x720 was considered as the reference one corresponding to the highest profile. In naming the proposed video profiles the widely used convention describing the number of horizontal lines in a video was used. Therefore the name associated to the highest video profile was **720p**.

The height and width of the 1280x720 resolution were halved resulting in a resolution of 640x360, which is four times smaller than the one corresponding to the 720p profile in total number of pixels. This was associated to the second profile, **360p.**

The resolution corresponding to the third profile was selected to be as close as possible as half the one corresponding to the 720p profile, and double the one corresponding to the 360p profile. The 854x480 resolution was found as being very close to this threshold and this was associated to the **480p** profile.

For the fourth profile the height and width of the 854x480 resolution were halved resulting in a resolution equal with 427x240, which was associated to the **240p** profile. This resolution is a wide variant of QVGA, which is considered the baseline resolution for m-learning devices such as Smartphones and PDA’s (Australian Flexible Learning Framework 2010).

Figure 3.7 presents the resolutions corresponding to the four profiles that have been proposed, indicating also the differences in size between consecutive profiles.

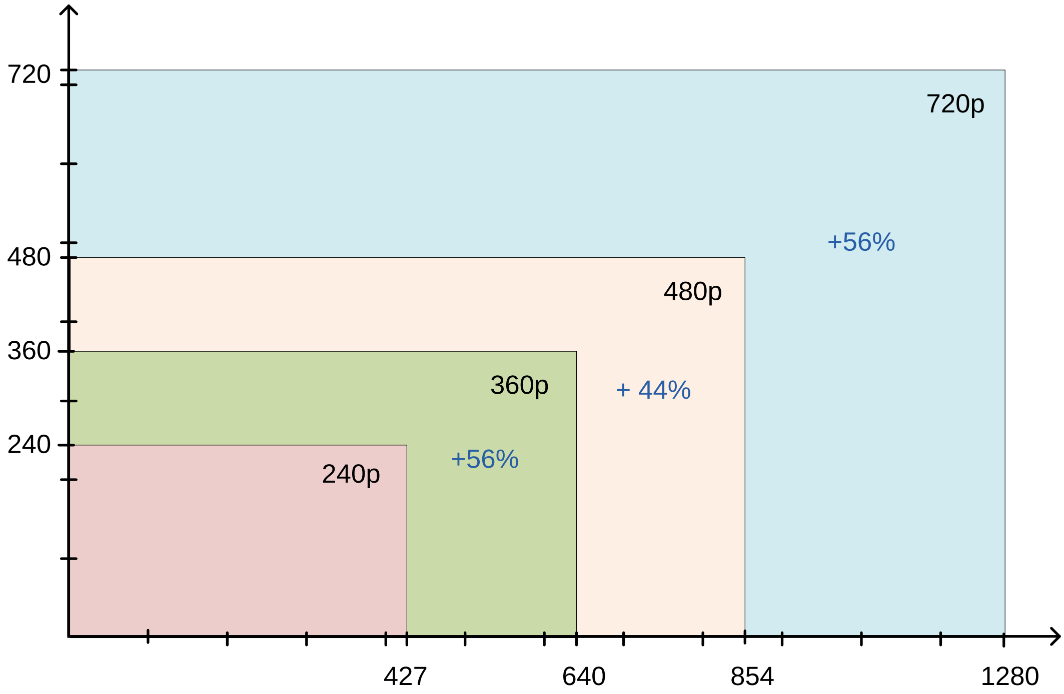

Figure 3.7: Proposed Video Profiles

3.4.1.2 Defining the Reference Bitrates and Frame Rates

No standard recommendations have been found to specify reference bitrate values that have to be used for e-learning multimedia applications, given specific resolution, frame-rate and compression techniques. Therefore the reference bitrates for the proposed video profiles were chosen starting from the average bitrates for multimedia content assessed in Section 3.3.2.4, and following the guidelines and recommendations to create multimedia content for e-learning applications (O'Connell & Smith 2007). The bitrates were chosen at the top range of the recommendations, assuming that this will offer an excellent quality level, and the videos encoded with these bitrates can be considered as reference videos. Following the same results, guidelines and recommendations, the minimum reference frame rate was selected as being at least 24fps.

Table 3.3: Recommended Encoding Setting for the Four Video Profiles

Profile	***Compression***	***Resolution [pixels]***	***Bitrate [Kbps]***	***Frame Rate***
720p	H.264 / MPEG-4 AVC	1280x720	4500	24-30
480p		854x480	2000	
360p		640x360	1200	
240p		427x240	600	

3.4.1.3 Characteristics of Reference Videos

Table 3.3 summarises the encoding parameters characteristic to the four profiles. The bitrate values were selected as presented in the previous section to the values of 4500 Kbps, 2000 Kbps, 1200 Kbps and 600 Kbps, for the four profiles 720p, 480p, 360p, and 240p respectively.

3.4.2 Classes of M-learning Devices

Section 3.2.2 has shown that a multitude of display resolutions exist among the mobile devices. Ideally, for an optimum quality, and for maximum resource consumption efficiency, a mobile device should receive a version of the multimedia clip adapted to fit its particular screen resolution. However, in practice this is not feasible due to the very high number of versions that would have to be created and stored, for each multimedia clip.

To overcome the issue posed by the multitude of screen resolutions a learner's mobile device may have, this section proposes to group the mobile devices in four different classes based on their screen resolutions. Each of these classes has associated a video profile, from the ones that have been proposed in Section 3.4.1. Table 3.4 presents a possible way for grouping the most common resolutions in four classes of devices, associated to the proposed video profiles.

Table 3.4: Possible Classes of M-learning Devices

Device Class	***Display Resolution***	***Video Profile***
Class 1: Large Screen devices	1600 x 768	1280x720
	1366 x 768	
	1280 x 800	
	1280 x 768	
Class 2: Medium Screen Devices	1024 x 768	854 x 480
	1024 x 600	
	1024 x 576	
	1024 x 480	
	960 x 640	
	960 x 480	
	800 x 600	
	854 x 480	
	800 x 480	
Class 3: Low Medium	640 x 480	640 x 360
	640 x 360	
Class 4: Small Screen Devices	480 x 360	427 x 240
	480 x 320	
	480 x 272	
	480 x 240	
	432 x 240	
	427 x 240	
	400 x 240	
	320 x 240	

3.5 Chapter Summary

The research presented in this chapter aimed to assess the characteristics of m-learning devices, as well as the most common encoding schemes that are used for multimedia educational content. Starting from these investigations, video profiles addressing multiple mobile devices having different resolutions were proposed, along with recommended values for encoding the educational clips for optimum quality on the respective devices.

Chapter 4

Assessment of Educational Multimedia Content Quality

4.1 Introduction

Adaptive strategies for delivering multimedia content are necessary to overcome situations when learners have limited or variable resources such as network bandwidth, mobile device battery power, or to overcome the differences between various devices in terms or screen size, processing capabilities, memory, etc.

Most of the existing techniques that can be used for adapting the educational multimedia content involve either providing the learner with a choice between multiple versions of the same content, each with a different quality level, or adapting the content in real-time through either transcoding or scalable encoding. In the first case one version will be selected manually by the learner or automatically by the system depending on the specific situation.

The final goal of the multimedia content adaptation process is to provide learners with good QoE levels, thus supporting learning process in variable conditions determined by factors such as network and device characteristics. As presented in Chapter 3, this is difficult to achieve in practice because of the multitude of factors and conditions that would have to be considered, and often compromises need to be made.

In order to cope with the high number of possible mobile device screen resolutions, the devices were grouped in four classes in the previous chapter. Furthermore, multimedia profiles offering excellent quality levels for most of the viewers, have been associated to each of the four classes. The characteristics of the four multimedia profiles have been chosen starting from a preliminary study that analysed the characteristics of more than 900 educational multimedia clips downloaded from the Internet, as well as other guidelines and recommendations that were found.

The high bitrate associated to the proposed multimedia profiles enables high levels of QoE. However, when a multimedia clip is streamed over the wireless network, and not played locally on the device, the QoE level may be significantly affected by the transmission conditions (e.g. network capacity, available bandwidth, delays, etc.). In such situations, adapting the bitrate of the streamed clip to match available network conditions will result in fewer interruptions and loss of data. The solution has the potential to significantly improve the overall QoE as compared to the situation when the clip is streamed at the reference bitrate value.

However, when considering the final purpose of educational multimedia clips, it is important not to reduce their quality below certain threshold, expressed in terms of a bitrate threshold. In comparison with the reference clip, the video sequence encoded at the bitrate threshold should still offer a good quality level, without affecting learners' capacity to acquire knowledge from the information presented.

This chapter presents the results of both objective and subjective studies performed to identify this bitrate threshold for different categories of educational clips that have been identified in Chapter 3.

These studies have concentrated on visual information, since video usually accounts for most of the overall bitrate of a multimedia clip. By reducing the video bitrate, significantly more bandwidth can be saved as opposed to reducing the audio bitrate. Furthermore from the mobile terminal point of view, video reception, decoding and playback requires significantly more processing and battery power resources as compared to the audio reception, decoding and playback.

This chapter is structured as follows. Section 4.2 presents the test methodology and setup. The particularities of the educational clips selected for this testing, as well as the objective and subjective testing methodologies are detailed. Section 4.3 presents the test results and their analysis, whereas Section 4.4 concludes the chapter.

4.2 Test Methodology

4.2.1 Test Sequences

For this study, eight educational multimedia clips were selected from the ones analysed in Section 3.3. From each of these clips one minute long test sequences were later selected and extracted. The whole process looked at identifying the most suitable clips based on the following criteria:

- The original clips need to be of a high quality, in order to be able to create very good reference clips for all the proposed video profiles.
- The test sequences should have different characteristics in terms of content type, presentation, dynamicity, etc.

Table 4.1 presents the main characteristics of the test sequences that were selected. Each of these corresponds to a different category of educational clips from the eight that have been identified. The following subsections, present the

Table 4.1: Characteristics of the High-Quality Test Sequences

Clip	***Type***	***Resolution [pixels]***	***Video Bitrate [Kbps]***	***Frame Rate [fps]***
Hulu	Screencast	1280x720	2000	30
Arts	Slideshow	1280x720	3700	24
Sol	Animation	960x540	3400	30
Languagelab	Graphics	1280x720	2000	30
Dubus	Interview	1280x720	5000	30
Obesity	Presentation	1280x720	2000	30
Hotness	Lab Demo	1280x720	4800	24
Sleep	Documentary	1280x720	5200	30

general characteristics that are common for educational clips belonging to each of these categories, as well as the particularities of the selected test sequences.

Additional details on how the test sequences were extracted, and further processed, are provided in Section 4.2.2.

4.2.1.1 Screencasts

Screencasts are educational clips that provide learners with very specific guidelines on how to perform different tasks. These are created by recording the computer screen, and thus the actions that have to be made, together with audio explanations.

General characteristic of screencasts include high level of details, in particular text of different sizes. Screencasts are also usually static clips, consisting in a sequence of scenes during which only small parts of the image are changing (e.g. mouse cursor movement, menu open, etc.).

Figure 4.1: Representative Frame for Test Sequence "Hulu"

The test sequence "***Hulu"***, was extracted from the multimedia clip "Super Cool Hulu Tricks"[1] created by Tinkernut[2], and available for download on Miro Guide. Characteristic to this particular sequence is the fact that the flow of ideas, and thus the speed with which the scenes are changing is very fast. Visual effects such as zoom in, window fly and window rotation are also present. Figure 4.1 presents a representative frame for the test sequence "Hulu".

4.2.1.2 Slideshows

Slideshows are another type of educational clips that mainly consist of a sequence of images accompanied by audio narration. Their type and level of details may vary depending on the information presented (image slideshow, slides with text, slides with text and images, etc.).

[1] Super Cool Hulu Tricks, http://www.miroguide.com/items/3269491
[2] Tinkernut.com | Watch. Learn. Share, http://tinkernut.com/

Figure 4.2: Representative Frame for Test Sequence “Arts”

In general slideshows are also static clips. In terms of dynamicity, the difference between various educational clips from this category is made by the speed with which the slides are changing or the various effects applied (e.g. zoom in, fade in/out, rotate, etc.). Sometimes the video frames are divided in regions having different levels of dynamicity. An example is represented by the clips when a recording of the speaker/presenter is present in a part of the video frames.

The test sequence “***Arts”*** is a slideshow, which was extracted from the multimedia clip “Michael McMillen - Multimedia”[3], part of the On Networks’ “For Arts Sake”[4] show. Apart from the very good quality of the original video, this sequence was selected for the high level of details present in the images, as well as for the visual effects applied during the transitions from one image to another. Figure 4.2 presents a representative frame for the test sequence “Arts”.

---

[3] Michael McMillen – Multimedia, http://www.miroguide.com/items/438159
[4] For Art's Sake | ON Networks, http://www.onnetworks.com/videos/for-arts-sake

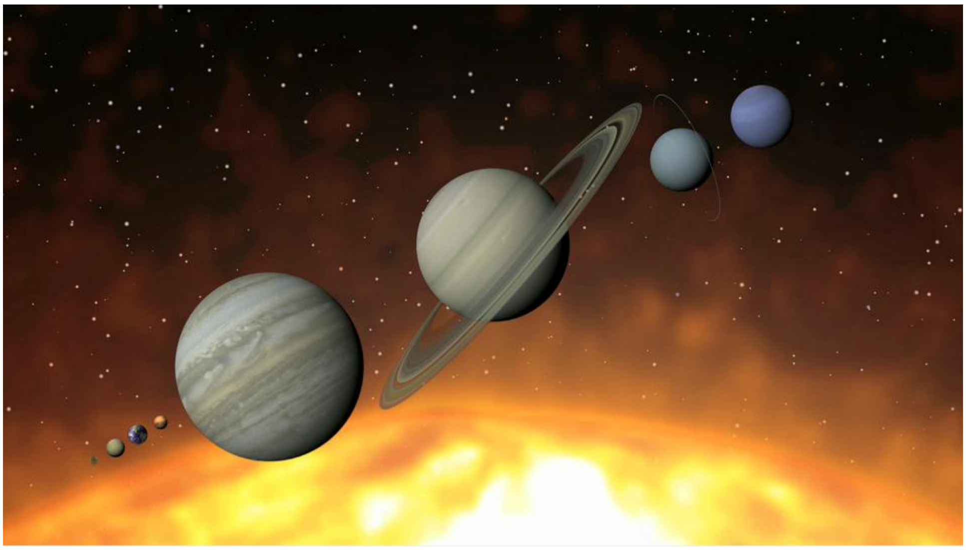

Figure 4.3: Representative Frame for Test Sequence "Sol"

4.2.1.3 Animations

Educational clips from this category consist of computer generated animations that have the role to describe and help learners better understand various concepts. Examples include physics concepts such as gravitation and electromagnetism, chemistry concepts such as the reactions between various chemical elements, etc. The dynamicity as well as the level of details of the various clips belonging to this category, including scenes of the same clip may vary broadly.

The test sequence "***Sol"***, was extracted from the multimedia clip "Bang to Sol"[5] created by the Cassiopeia Project[6]. The sequence has a medium level of dynamicity, describing concepts such as planets orbiting around the Sun and the evolution of the Sun. Figure 4.3 presents a representative frame for the test sequence "Sol".

[5] Bang to Sol, http://www.cassiopeiaproject.com/vid_courses3.php?Tape_Name=Bang
[6] Welcome to Cassiopeia Project, http://www.cassiopeiaproject.com/

Figure 4.4: Representative Frame for Test Sequence "Languagelab"

4.2.1.4 Games & Virtual World Recordings

Educational clips from this category consist of sequences presenting various computer generated learning environments such as educational games, 3D virtual learning environments (VLE), etc. This category can be mapped to the broader genre of multimedia clips known as "***cartoons***".

The test sequence "***Languagelab***", represents a recording of a 3D VLE, and was extracted from the multimedia clip "Languagelab in Second Life"[7] directed by Al Peretz[8]. The sequence has a low dynamicity presenting mostly static characters. It contains low-contrast scenes, with almost constant level of luminosity. Figure 4.4 presents a representative frame for the test sequence "Languagelab".

[7] YouTube - Languagelab In Second Life, http://www.youtube.com/watch?v=8hJZ2bre_FI
[8] Languagelab.com, http://www.languagelab.com/from/alkohn

Figure 4.5: Representative Frame for Test Sequence "Dubus"

4.2.1.5 Interviews

Educational clips from this category can be mapped to the broader category of multimedia clips known as "***news***". The background is usually static, whereas the foreground contains very little movement (e.g. peoples heads, mimics, etc.), making these clips have low levels of dynamicity. Depending on the number of cameras used for filming the interview, the various scenes comprising the clip may be filmed from the same angle or from different angles. Furthermore, a particular scene may present the interviewer, the interviewed, or both of them.

The test sequence "***Dubus"***, was extracted from the educational multimedia clip "Conversations from the Iowa Writers' Workshop: Andre Dubus III"[9] available for download on iTunes. Characteristic to this sequence is the constant, low level luminosity. Also the sequence contains scenes that present both the interviewer and the interviewed, and scenes presenting only one of them. Figure 4.5 presents a representative frame for the test sequence "Dubus".

[9] Iowa Writers' Workshop - Andre Dubus III , http://media.uiowa.edu/btn/dubus.html

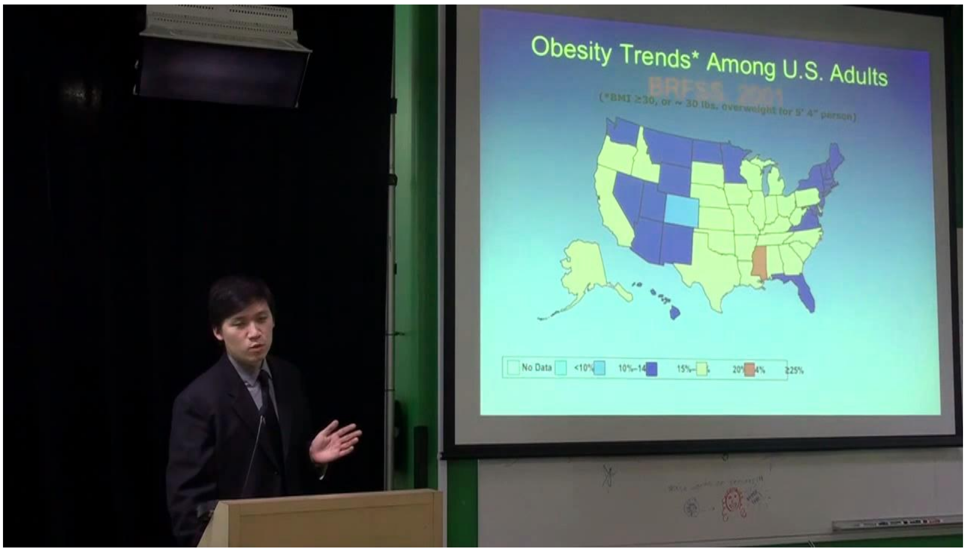

Figure 4.6: Representative Frame for Test Sequence "Obesity"

4.2.1.6 Lecture Recordings & Presentations

Educational clips from this category are characterised by the presence of a lecturer/presenter, accompanied by a display of information such as projected slides, information written on a whiteboard/blackboard, etc. The presenter may be mostly static, or s/he can move inside the camera view.

The test sequence "O***besity"***, was extracted from the multimedia clip "Obesity and Mortality in 130 years of Major League Baseball"[10], part of the Google Tech Talks[11] series, and available for download on Miro Guide. Characteristic to this sequence is the usage of a single, fixed camera, and the static presenter. The information medium is represented by a slideshow projection, whereas the information on the slides is represented both by images and textual information. Figure 4.6 presents a representative frame for the test sequence "Obesity".

---

[10] Obesity and Mortality in 130 years of MLB, http://www.miroguide.com/items/3311641
[11] YouTube - GoogleTechTalks's Channel, http://www.youtube.com/user/GoogleTechTalks

Figure 4.7: Representative Frame for Test Sequence "Hotness"

4.2.1.7 Lab Demos

Educational clips from this category are characteristics to live demonstrations, in which a person explains different concepts and actions, while performing some of them at the same time. The camera may be fixed, in which case a broader view of the demonstrator and the objects of interest are presented, or this can focus closely on the actions performed during some scenes.

The test sequence "***Hotness"***, was extracted from the multimedia clip "Hot and Spicy"[12], part of the On Networks' "Food Science"[13] show. Characteristic to this particular sequence are the static background, and the fast changing between scenes focusing on the demonstrator and respectively on the actions performed. Figure 4.7 presents a representative frame for the test sequence "Hotness".

[12] Food Science, Hot and Spicy, http://www.miroguide.com/items/996746
[13] Hot and Spicy | ON Networks, http://www.onnetworks.com/videos/food-science/

Figure 4.8: Representative Frame for Test Sequence "Sleep"

4.2.1.8 Documentaries

Characteristic to educational documentaries are the higher number content types that may occur across the different scenes comprising the clip. The dynamicity and the level of details may also differ a lot between different scenes. A particular scene may fall in any of the seven previously mentioned categories but is not limited to these.

The test sequence, "***Sleep***", was extracted from the multimedia clip "QUEST Quiz: Sleep"[14]. This short clip is part of the KQED's multimedia series QUEST, and serves as a quiz for viewers to test their knowledge about sleep and sleep disorders. The clip was selected for the fact that it presents questions and answers embedded in the video, while the background consists of various scenes with different level of details and dynamicity. A representative frame for the test sequence "Sleep" is presented in Figure 4.8.

---

[14] KQED QUEST, QUEST Quiz: Sleep, http://www.kqed.org/quest/television/quest-quiz-sleep

4.2.2 Preparing the Test Sequences

As previously mentioned, 1 minute long test sequences corresponding to different categories of educational multimedia clips were selected from eight different educational multimedia clips. The 1 minute long duration of these sequences, was chosen having in mind an acceptable subjective testing time of less than 25 minutes.

The test sequences were extracted from the original clips using the open source video editor Avidemux[15]. In order not to degrade their original video quality, they were saved using raw H.264/MPEG-4 AVC video compression.

Further on each sequence was divided in two, shorter sequences having equal duration of 30 seconds. This procedure was motivated by the desire of conducting the whole testing for multiple video profiles from the ones proposed in Chapter 3. Only the 480p and 360p profiles were tested. This was because no sufficient educational clips of very high quality were found, to be able to create reference clips corresponding to the 720p profile, for all the eight categories of educational clips, previously described.

There were two reasons behind deciding not to test the lowest 240p profile. First, high resolutions such as WVGA are already common among mobile devices with screen sizes as low as 3”-4”. Secondly, low video resolutions such as the one corresponding to the 240p profile, is not suitable for viewing educational clips presenting small textual information, such as screencasts, since this information becomes impossible to be read.

[15] Avidemux - Main Page, http://fixounet.free.fr/avidemux/

The decision to use different sequences for each profile being tested was motivated by the conditions imposed by the subjective testing. To assess the impact of video quality on the learning process, a controlled environment had to be created. Therefore different questions related to the visual information presented in the multimedia sequences were asked. Showing the same sequence two times, for the two profiles, meant that the viewer had to answer the first question after seeing the sequence one time, whereas answering the second question would had happened after viewing the sequence two times.

Once the sequences were extracted at their original quality, reference sequences corresponding to the two video profiles being tested, were created. The first 30 seconds from each of the eight sequences were used to create the references for the profile 360p, whereas the last 30 seconds were used for creating the references for the profile 480p. Apart of the video bitrate and the resolution, which are different for different video profiles, all the other encoding parameters were the same for all the 16 reference sequences created. A summary of the encoding setting are presented in Table 4.2. To convert the reference sequences, XMedia Recode[16] media converter software, with x264[17] H.264/MPEG-4 AVC video encoder and FAAC[18] AAC (Advanced Audio Codec) audio encoder was used. 2-Pass Average Bitrate video encoding mode was chosen, because of the better quality offered as compared to 1-Pass encoding mode. Using this mode, the compressed video sequences have a bitrate that is identical to the target bitrate in all the cases.

[16] XMedia Recode, http://www.xmedia-recode.de/
[17] VideoLAN - VideoLAN - x264, http://www.videolan.org/developers/x264.html
[18] AudioCoding.com – FAAC, http://www.audiocoding.com/faac.html

Table 4.2: Encoding Settings for the Reference Sequences

Parameter	***Value***
Container	.mp4
Video Settings	
Compression	H.264 / MPEG-4 AVC
Resolution [Pixels]	640x360 (360p Profile) 856x480 (480p Profile)
Bitrate [Kbps]	1200 (360p Profile) 2000 (480p Profile)
Bitrate Mode	2-Pass Average Bitrate
Frame Rate	Original
Audio Settings	
Compression	AAC
Channels	2 (Stereo)
Bitrate [Kbps]	128
Sample Rate [KHz]	44.1

4.2.3 Objective Quality Estimation Methodology

Subjective methods for assessing the video quality are widely accepted as the best methods for assessing the video quality. However, by requiring a controlled environment and large enough groups of subjects for the results to be significant, these methods are very time-consuming and expensive to be conducted.

Since in this study minimum bitrate values that still offer good quality levels as compared to the reference bitrates, had to be determined for multiple profiles and for multiple content types, an automatic method was needed for initial estimation of the video quality. This fact has encouraged the use of objective video quality assessment metrics as a starting point. Objective metrics have the advantage of being readily repeatable, since they can be measured using computer software.

Full-reference objective metrics such as SSIM, PSNR or VQM, estimate the video quality by comparing different versions of the same sequence. Having the reference sequences, created for different multimedia profiles as presented in Section 4.2.2, full-reference metrics could be used for estimating the video quality of degraded, lower-quality versions, by comparing them with the references. The bitrate of the reference sequences, was selected so to offer an excellent quality level according to the objective metrics.

The procedure used for determining the minimum reference bitrate of a test sequence is presented in Figure 4.9. This consists in creating multiple versions of that sequence, having lower bitrates as compared to the reference one, and estimating their quality using the SSIM and PSNR objective metrics. These versions are created following the same procedure as for the case of their references described in Section 4.2.2.

The bitrate of the first degraded version being created (Start_Bitrate), and the step between two successive versions (Step_Bitrate) have to be predefined for each of the profiles being tested. Specific values used during the objective quality estimation tests, ate provided in Section 4.3.1. PSNR and SSIM are the full-reference objective metrics used for conducting the measurements. PSNR was selected because, thanks to its simplicity, it is the most widely used objective metric. SSIM was selected as it was shown to have a good correlation with the subjective video quality (Wang, Lu et al. 2004). Since video quality is usually expressed through the Mean Opinion Scores (MOS), the measured values for PSNR and SSIM are mapped to the MOS_{PSNR} and MOS_{SSIM} scores. The mapping proposed by Zinner et al. (2010), and presented in Table 4.3, was used for this purpose.

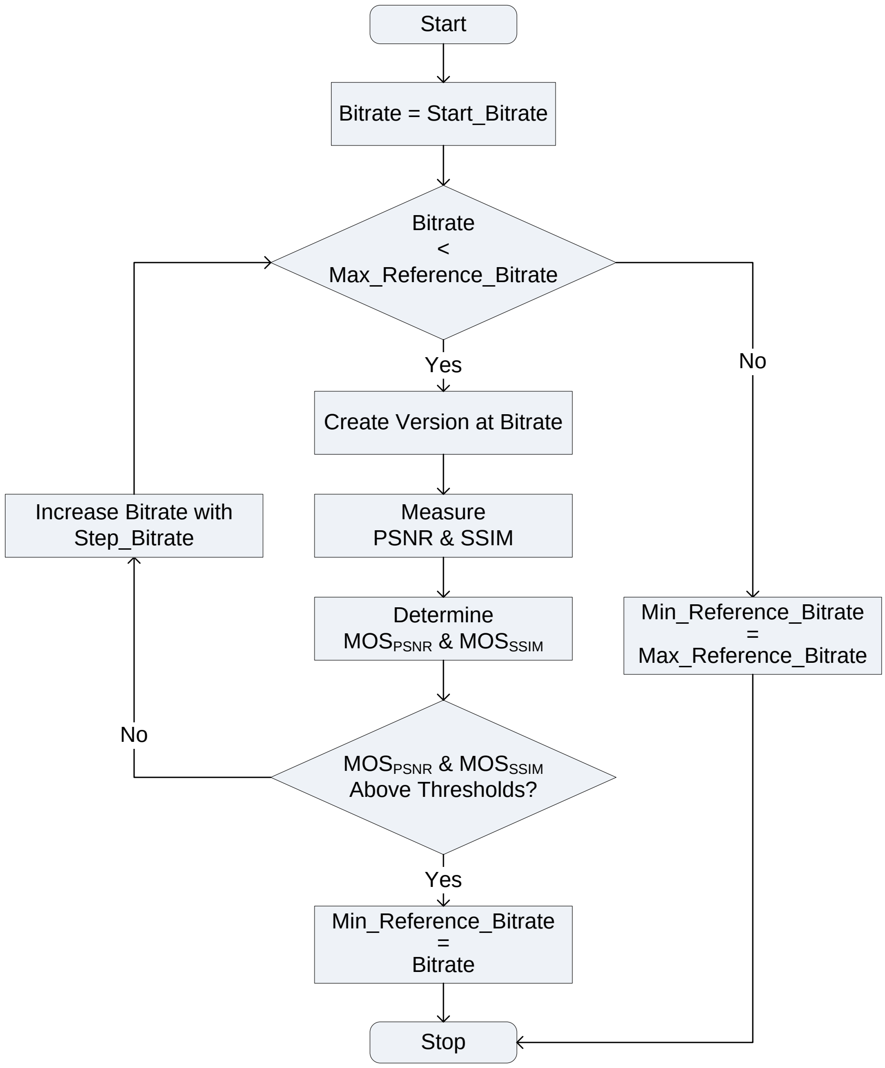

Figure 4.9: Minimum Quality Estimation Procedure

Table 4.3: Mapping of Objective QoE to Subjective QoE (Zinner et al. 2010)

MOS	***PSNR [dB]***	***SSIM***
5 (Excellent)	≥ 45	≥ 0.99
4 (Good)	≥ 33 & < 45	≥ 0.95 & < 0.99
3 (Fair)	≥ 27.4 & < 33	≥ 0.88 & < 0.95
2 (Poor)	≥ 18.7 & < 27.4	≥ 0.5 & < 0.88
1 (Bad)	< 18.7	< 0.5

Having the MOS_{PSNR} and MOS_{SSIM} scores, a checking procedure is made to see if these are above the predefined thresholds. For a bitrate to be considered as the minimum at which the quality can be decreased (Min_Reference_Bitrate), at least one of the MOS_{PSNR} and MOS_{SSIM} scores has to be 5, whereas the other can be 4, but the corresponding PSNR or SSIM value needs to be very close to the threshold (SSIM $\geq$ 0.98, PSNR $\geq$ 42 dB). If the condition is not true, than another version is created, having the bitrate increased with one step. The whole procedure is repeated until a minimum reference bitrate is found, or until the bitrate of the next version to be created is above the one corresponding to the reference sequence (Max_Reference_Bitrate).

For measuring the average PSNR and SSIM values for each multimedia sequence, the MSU Video Quality Measurement Tool[19] was used. The tool provides comparative results visualisation as shown in Figure 4.10. Additionally the results can also be exported to .csv files on a frame by frame basis, for further analysis. Individual frames of original and processed files being compared can also be visualised (see Figure 4.11).

[19] MSU Video Quality Measurement Tool (PSNR, MSE, VQM, SSIM), http://compression.ru/video/quality_measure/video_measurement_tool_en.html

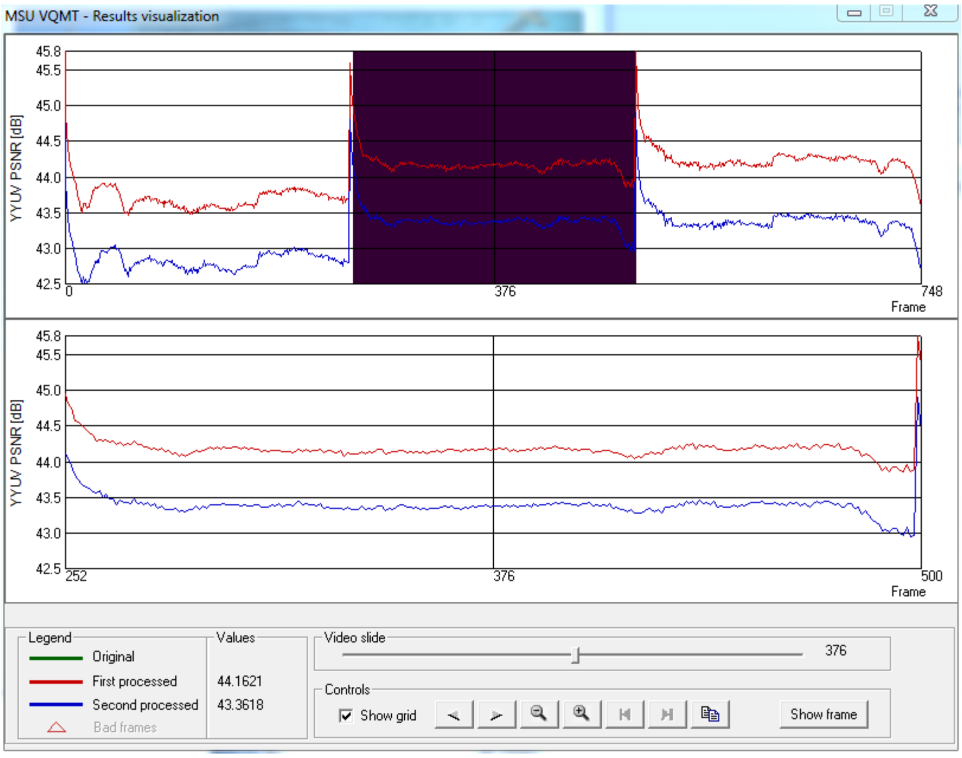

Figure 4.10: MSU VQMT – Results Visualisation Window

Figure 4.11: MSU VQMT – Frames Visualisation Window

4.2.4 Subjective Validation Methodology

The subjective testing had two goals:

- To assess if the minimum reference bitrates selected using the objective quality estimation method proposed in Section 4.2.3, offer a good quality level;
- To assess if learners can achieve knowledge from educational multimedia clips whose visual quality was reduced to the minimum reference bitrates.

4.2.4.1 Test Setup

Two mobile devices were used for this testing. Each of them corresponds to a different class of m-learning devices from the four classes proposed in Section 3.4. The first device, a HP iPAQ 214 PDA, was used for displaying the test sequences corresponding to the 360p profile. The second device, a Dell Inspiron Mini 10 Netbook was used for displaying the test sequences corresponding to the second profile being tested, 480p. The characteristics of the two devices are summarised in Table 4.4.

The test room is shielded from natural light, in order to maintain a constant luminance level across all the participants. In order to offer a maximum visibility for the subjects the screen brightness was set to 100% on both devices. To further control the noise level, the subjects were provided with headphones. Only one participant can attend the test at one moment in time.

The test sequences were stored and played locally on the mobile devices. This was done in order to reduce to minimum the influence of other factors, such as the variable network conditions, on the perceived quality.

Table 4.4: Characteristics of the Mobile Devices Used for the Subjective Tests

Device	*Characteristics*
HP iPAQ 214	**Device Type:** PDA
	OS: MS Windows Mobile 6
	Memory: 128 MB SDRAM
	CPU: Marvell PXA310, 624MHz
	Screen Size: 4”
	Screen Resolution: 640x480 (VGA)
Dell Inspiron Mini 10	**Device Type:** Netbook
	OS: MS Windows XP Home Edition
	Memory: 1 GB DDR2
	CPU: Intel Atom N270, 1.6 GHz
	Screen Size: 10.1”
	Screen Resolution: 1024x576 (WSVGA)

CoreCodec TCPMP[20] open source media player was used for displaying the test sequences on the PDA device, and Google Chrome[21] web browser with HTML5 video support for the H.264/MPEG-4 AVC codec, was used for displaying the test sequences on the netbook device.

On the PDA device the sequences were displayed in full-screen, the screen width being equal with the video resolution width of the test sequences corresponding to the 360p profile (640 pixels). On the netbook device the test sequences were displayed in the middle of the screen at their actual video resolution (856x480 pixels), and not scaled up in full-screen. HTML pages were created for displaying the sequences in the second case. The browser was set to

[20] CoreCodec TCPMP (Open Source) - HPC Factor, http://www.hpcfactor.com/downloads/tcpmp/
[21] Google Chrome - A fast new browser. For PC, Mac, and Linux, http://www.google.com/chrome

Figure 4.12: Test Sequence Display on the Netbook Device

display the pages in full-screen, so that only the video window and the test phase number to be visible on the screen (see Figure 4.12). The background of the pages was set to a level of 50% grey (Y=U=V=128, colour code #808080) following the recommendations from ITU-T's Recommendation P.910 (ITU-T 2008).

4.2.4.2 Test Methodology

Subjective video quality assessment methods for multimedia applications are standardised by ITU-T in the P.910 Recommendation (ITU-T 2008). To maintain the overall testing duration to an adequate duration of less than 30 minutes, the Absolute Category Rating (ACR) method was selected. For this case the test sequences are presented one at a time and are rated independently on a five level quality scale (1- Bad, 2 – Poor, 3 – Fair, 4 – Good, 5 - Excellent) . If a constant voting time is used, this should be less or equal with 10 seconds The ACR method recommends using 10 seconds long sequences, but their length can be reduced or increased according to the content of the test material.

Since 10 seconds duration is too short for presenting educational concepts, 30 seconds long sequences were selected for this testing (see Section 4.2.2).

4.2.4.3 Test Session Description

Before starting, each subject was explained the testing procedure and was asked if s/he has any questions. Personal information and preferences were collected using a questionnaire as presented in Appendix A.

The test session started with a training session for familiarising the subjects with the assessment phase. After the initial training phase, the test session continued with 16 test phases, from which the first 8 were conducted on the PDA device, and the last 8 on the netbook device. For the first 30 seconds of a phase, the subjects have to watch a test sequences. After watching the sequence, the subjects are asked to rate the overall video quality on the provided scale, and to answer a question whose answer was strictly related to the visual information present in the sequence they just saw.

To maintain similar testing conditions for all the subjects, they were not allowed to pause, stop, or review the sequences. Furthermore, the questions corresponding to particular sequences were printed on an individual sheet, and the subjects were instructed to turn the sheet and read the questions only after watching the sequence. In order to provide the subjects with sufficient time for answering the questions, the voting (answering) time was not limited to a fixed duration. Instead, the subjects had to confirm themselves when they were ready to view the next sequence. Assuming the a subject will need up to 60 seconds for answering the questions corresponding to a sequence, the overall duration of testing, should last less than 30 minutes. The test material is available in Appendix A of this thesis.

4.2.4.4 Subjects

21 (13 male, 8 female) subjects participated in the subjective assessment. Their age ranged from 21 to 37 years old, where the average age was 27.14 years old (STDEV=4.74). The percentage of the age range is presented in Table 4.5.

Table 4.5: Subjects' Age Range

Age Range	***Percentage***
<18	0%
18-24	33%
25-35	57%
>35	9.5%

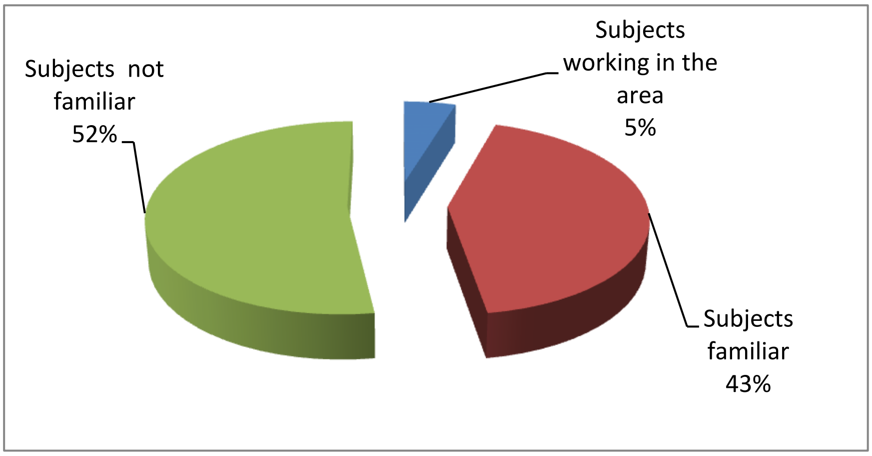

Figure 4.13: Subjects Familiarity with Subjective Video Quality Evaluation

All subjects reported that they had normal vision or have corrected to normal vision (they were wearing glasses). Both English native speakers and non-native speakers have participates in the subjective testing. 76 % of the subjects were English non-native speakers and 24 % were English native speakers. They ranged from cross different domains: undergraduate students, postgraduate students and professionals.

52 % of the subjects have reported not being familiar and 43 % being familiar with subjective video quality evaluation (see Figure 4.13). Only 5 % have reported working in the area. The conclusion that can be drawn is that most of the subjects are non-expert, meeting the requirements specified in the ITU-T BT.500-11 Recommendation for video quality assessment. This is as preliminary findings

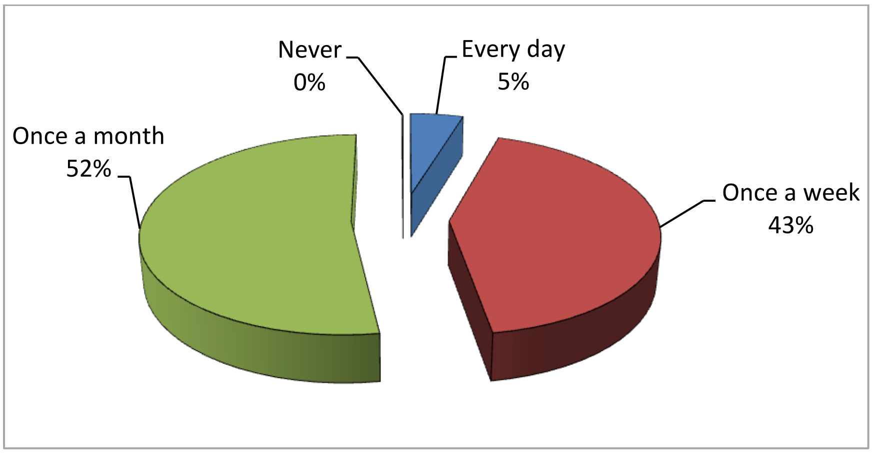

Figure 4.14: Frequency of Watching DVD/Blue Ray Movies

have shown that using non-expert subjects may yield to more critical results (ITU-R 2002).

As shown in Figure 4.14, 5 % of the subjects have reported watching DVD/Blue Ray every day, 43 % watching once a week, 52 % watching once a month. None of the subjects have reported never watching DVD/Blue Ray movies. A conclusion that can be drawn is that the subjects are familiar in a way with high quality multimedia content, being expected to provide more critical grading of the subjective quality assessment.

Looking at subjects' frequency of watching videos on the Internet using mobiles devices (e.g. Smartphone, iPod, PDA, Netbook) the **Error! Reference source not found.** shows that approximate equal percentage of the subjects watch videos every day, once a week or never. 28 % watch every day, 29 % once a week, 14 % once a month, and 29 % never watch video on Internet using mobile devices. Therefore not all the users were watching videos on mobile devices.

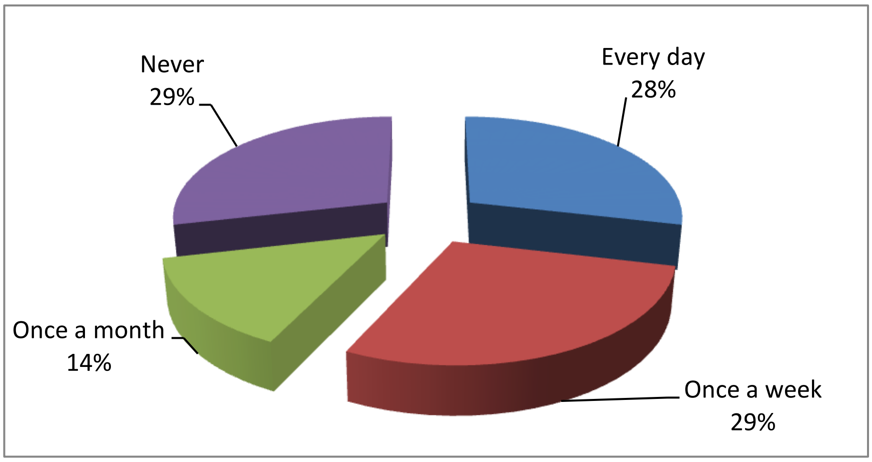

Figure 4.15: Frequency of Watching Videos on the Internet Using Mobile Devices

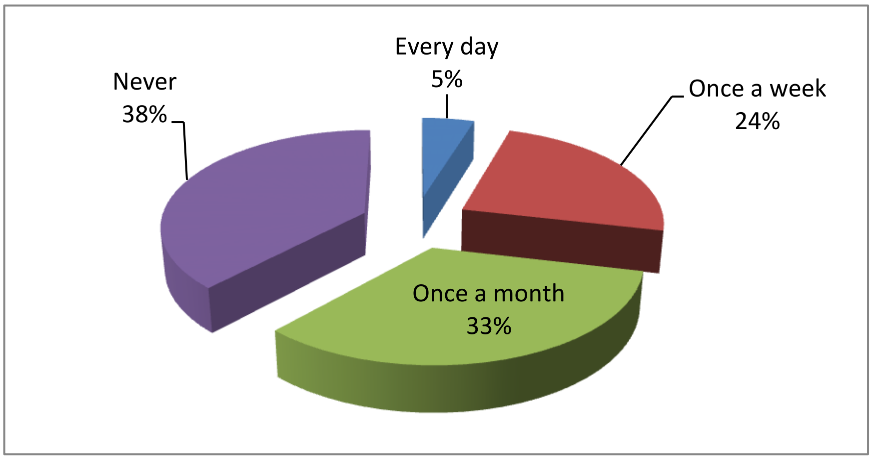

Figure 4.16: Frequency of Watching Educational Content on Mobile Devices

The majority of the subject reported as watching educational content on mobile devices once a week or once a month (24 % watch once a week and 33 % once a month). The percentage of the subjects that reported never watching educational content on mobile devices is 38 % (see Figure 4.16).

When asked what they would prefer, if they were watching educational multimedia clips streamed over the Internet and their quality is impacted by transmission problems, 76% of the respondents have preferred watching a lower

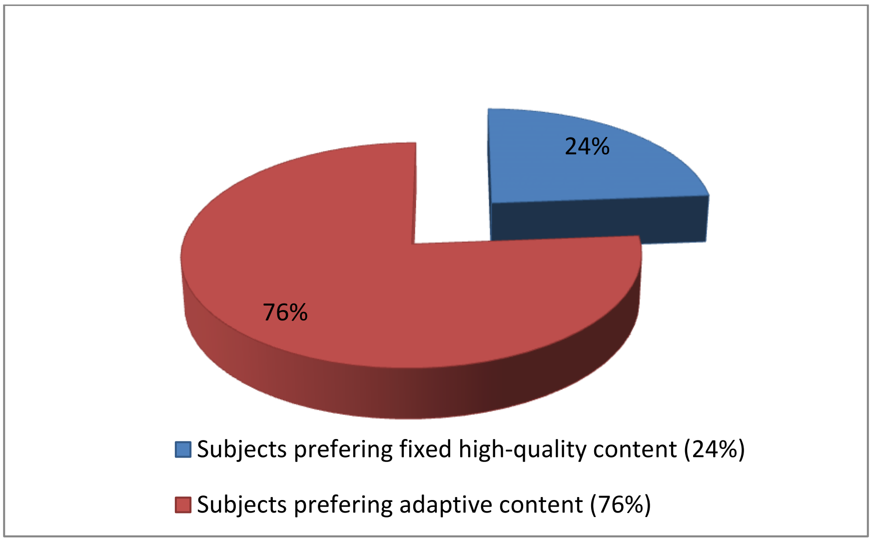

Figure 4.17: Subjects Preferences on Adaptive Multimedia Streaming

but acceptable quality version, if this will offer a smoother viewing experience (see Figure 4.17). The other 24% have preferred to continue watching the high-quality version, even if interruptions and loss of data will occur.

In conclusion in situations when multimedia content adaptivity can be used for overcoming the problems caused by variable network conditions, the majority of the learners will accept being provided with a lower but acceptable quality level.

4.3 Results

4.3.1 Objective Quality Estimation

The goal of the objective quality estimation tests was to determine minimum reference bitrates that offer good quality levels as compared to the reference high-quality/high-bitrate clips, for different categories of multimedia educational clips. Therefore for each of the 16 test sequences, corresponding to one of the eight categories (Screencasts, Slideshows, Animation, Games/VLE Recordings, Interview, Presentation, Lab Demos and Documentaries), and to one of the two multimedia profiles considered for testing (360p and 480p), the minimum reference bitrate was determined.

4.3.1.1 Profile 360p

Starting from a video bitrate equal with the audio bitrate (128 Kbps), multiple versions were created for each of the 8 test sequences corresponding to the 360p profile, and their quality was estimated by measuring the objective PSNR and SSIM metrics and mapping them to the corresponding MOS_{PSNR} or MOS_{SSIM} values on the 1-5 quality scale (see Section 4.2.3 for a description of the procedure). The step between any two consecutive versions was 64 Kbps.

The results for the 360p profile are summarised in Table 4.6 for all the test sequences except Documentary A and in Table 4.7, for the this one. The reason for this is the fact that only the results for the bitrate selected as the minimum reference one (underlined in the tables), and one value below are presented.

For each test sequence, the PSNR and SSIM measured values and respectively the MOS_{PSNR} and MOS_{SSIM} estimated values are presented. For the first seven test sequences (see Table 4.6) the values corresponding to the 256 Kbps and 192 Kbps are presented. For example Interview A has the PSNR 45db that corresponds to a MOS_{PSNR} of 5, and the SSIM 0.98 which corresponds to a MOS_{SSIM} of 4, at 256 Kbps. The same sequence (Interview A) has a PSNR of 42db (MOS_{PSNR} = 4), and a SSIM of 0.97 (MOS_{SSIM} = 4), at 192 Kbps.

The Interview A version at 256 Kbps meets the first requirement of the selection procedure, which is to have at least one of two MOS_{PSNR} and MOS_{SSIM} scores equal to 5 (Excellent). It also meets the second requirement, which is to have the SSIM at least equal with the 0.98 threshold. Therefore for this particular sequence 256 Kbps was selected as the minimum reference value, below which the video bitrate cannot be decreased without significantly reducing users' perceived quality of the multimedia clip.

Table 4.6: Objective Scores for the 360p Profile

Sequence	*PSNR*	*MOS_{PSNR}*	*SSIM*	*MOS_{SSIM}*
		256 Kbps		
Screencast A	44	4	0.99	5
Slideshow A	43	4	0.99	5
Animation A	45	5	0.99	5
VLE A	43	4	0.99	5
Interview A	45	5	0.98	4
Presentation A	46	5	0.99	5
Lab Demo A	43	4	0.99	5
		192 Kbps		
Screencast A	41	4	0.98	4
Slideshow A	39	4	0.96	4
Animation A	44	4	0.98	4
VLE A	40	4	0.97	4
Interview A	42	4	0.97	4
Presentation A	44	4	0.98	4
Lab Demo A	40	4	0.98	4

Table 4.7: Objective Scores Values for Test Sequence Documentary A

Bitrate	*PSNR*	*MOS_{PSNR}*	*SSIM*	*MOS_{SSIM}*
512 Kbps	45	5	0.98	4
448 Kbps	44	4	0.97	4

The same procedure is applied for selecting the minimum reference bitrates for all the other test sequences. As shown in Table 4.6, for the first 7 sequences, the selection process has stopped at 265 Kbps. For the case of Documentary A sequence, the procedure had to be repeated 7 times before finding a bitrate to meet both requirements (128 k, 192k, 256k, 320k, 384k, 448, 512k).

4.3.1.2 Profile 480p

For the second profile tested, the starting value for the video bitrate of the degraded versions of the test sequences was 256 Kbps. The step between any two consecutive versions was 128 Kbps. The reason for doubling these parameters, was motivated by the ~50% difference between the 640x360 and 856x480 video resolutions corresponding to the two profiles being tested , in terms of the overall number of pixels.

Table 4.8 summarises the objective results for the first seven sequences corresponding to the second profile tested, 480p (Screencast B, Slideshow B, Animation B, VLE B, Interview B, Presentation B, Lab Demo B). PSNR and SSIM measured values and respectively the MOS_{PSNR} and MOS_{SSIM} estimated values are presented for each test sequence. Like for the 360p profile the minimum reference selection procedure has stopped at the same bitrate for all of these sequences, specifically 384 Kbps.

Table 4.9 presents the results for the Documentary B sequence. As shown in the table, the degraded version at 768 Kbps has the MOS_{SSIM} equal to 5 (Excellent) and the PSNR above the 42 dB threshold. Therefore the version meets both requirements for being selected as the minimum reference value.

Table 4.8: Objective Scores for the 480p Profile

Sequence	***PSNR***	***MOS_{PSNR}***	***SSIM***	***MOS_{SSIM}***
		384 Kbps		
Screencast B	46	5	0.99	5
Slideshow B	45	5	0.98	4
Animation B	47	5	0.99	5
VLE B	47	5	0.99	5
Interview B	44	4	0.99	5
Presentation B	44	4	0.99	5
Lab Demo B	43	4	0.99	5
		256 Kbps		
Screencast B	44	4	0.97	4
Slideshow B	43	4	0.96	4
Animation B	44	4	0.98	4
VLE B	44	4	0.98	4
Interview B	43	4	0.98	4
Presentation B	42	4	0.98	4
Lab Demo B	40	4	0.98	4

Table 4.9: Objective Scores Values for Test Sequence Documentary B

Bitrate	***PSNR***	***MOS_{PSNR}***	***SSIM***	***MOS_{SSIM}***
768 Kbps	44	4	0.99	5
640 Kbps	42	4	0.98	4

4.3.2 Subjective Quality Assessment

In order to validate the proposed approach for video quality estimation of the minimum reference bitrates, a subjective study was conducted following the procedure described in Section 4.2.4.

The participants were asked to view the sixteen minimum bitrates references that were selected during the objective quality estimation testing and for each one of them to rate their perceived overall video quality. A 5 level quality scale, where 1 – Bad and 5- Excellent, was used for grading purposes. For each sequence, the mean value represented as the MOS score and the standard deviation of the statistical distribution of the assessment grades were computed.

The results are summarised in Table 4.10and Figure 4.18. As it can be seen all the 16 sequences have scored above 3.6 in average. For the 360p profile, only two sequences, VLE A and respectively Presentation A have scored less than 4 (Good), whereas for the 480p profile, three have scored below this threshold, Slideshow A, VLE A and Presentation A. A possible explanation why the VLE and Presentation sequences have scored below 4 on both devices, can be the higher compression rate of the original clips from which these were extracted (2000 Kbps original bitrate).

The average MOS scores are equal with 4.1 for profiles that were tested, 360p and 480p. A two sample T-Test analysis on these means values, indicate with a 99% confidence level, that statistically there is no difference in the final scores for the two video profiles being tested ($\alpha = 0.01$, $t = 0.16$, t-critical $= 2.99$, $p(t) = 0.43$). The Pearson correlation between the MOS scores corresponding to the two profiles is 0.81.

Table 4.10: Subjective Quality Results

Profile	***360p***		***480p***	
Sequence	***MOS***	***STDEV***	***MOS***	***STDEV***
Screencast	4.0	0.95	4.2	0.89
Slideshow	4.0	0.89	3.7	1.10
Animation	4.5	0.68	4.5	0.60
VLE	3.6	1.12	3.9	1.01
Interview	4.5	0.68	4.5	0.75
Presentation	3.9	0.62	3.8	0.77
Lab Demo	4.5	0.75	4.3	0.72
Documentary	4.0	0.84	4.0	0.86
Average	**4.1**	**0.82**	**4.1**	**0.84**

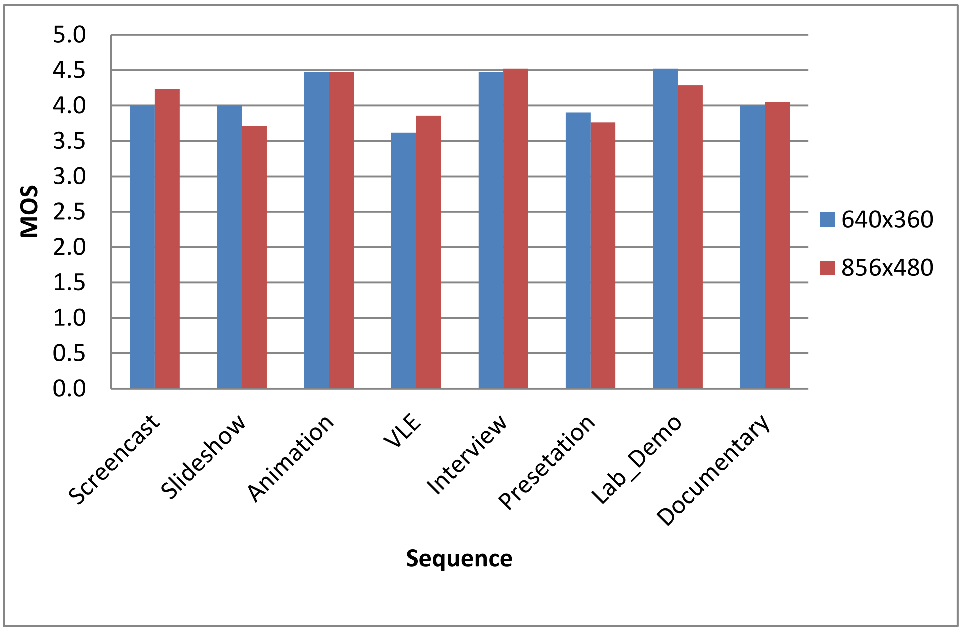

Figure 4.18: Comparative Mean Opinion Scores for the Two Profiles

4.3.3 Learning Assessment

The second goal of the subjective testing was to assess if learners can achieve knowledge from educational multimedia clips whose visual quality was reduced to the minimum reference bitrates. For this purpose, questions related strictly to the visual information presented, were extracted for each of the 16 test sequences. Out of these 12 questions were of “multiple choice” type with only one answer correct, 2 questions were “yes/no” type, and two questions were “short answer” type (see Appendix A).

The average answers for the questions corresponding to the 16 test sequences are presented in Table 4.11and Figure 4.19. Out of the sixteen questions, in six situations all the 21 subjects have answered correctly (Q4, Q5, Q9, Q10, Q12, and Q16), and only in 4 situations more than 10% of the subjects have provided wrong answers (Q1, Q3, Q11, and Q14). When looking at the individual answers for each subject, Figure 4.20 shows that these range from 11 to16, where the average score is 14 (STDEV = 1.24).

Table 4.11: Average Right and Wrong Answers for the 16 Questions

		360p			***480p***	
Sequence	***Question***	***Right (%)***	***Wrong (%)***	***Question***	***Right (%)***	***Wrong (%)***
Screencast	**Q1**	86 %	14 %	**Q9**	100 %	0 %
Slideshow	**Q2**	95 %	5 %	**Q10**	100 %	0 %
Animation	**Q3**	67 %	33 %	**Q11**	62 %	33 %
VLE	**Q4**	100 %	0 %	**Q12**	100 %	0%
Interview	**Q5**	100 %	0 %	**Q13**	95 %	5 %
Presentation	**Q6**	95%	5 %	**Q14**	24 %	77 %
Lab Demo	**Q7**	95 %	5 %	**Q15**	95 %	5 %
Documentary	**Q8**	90 %	10 %	**Q16**	100%	0 %

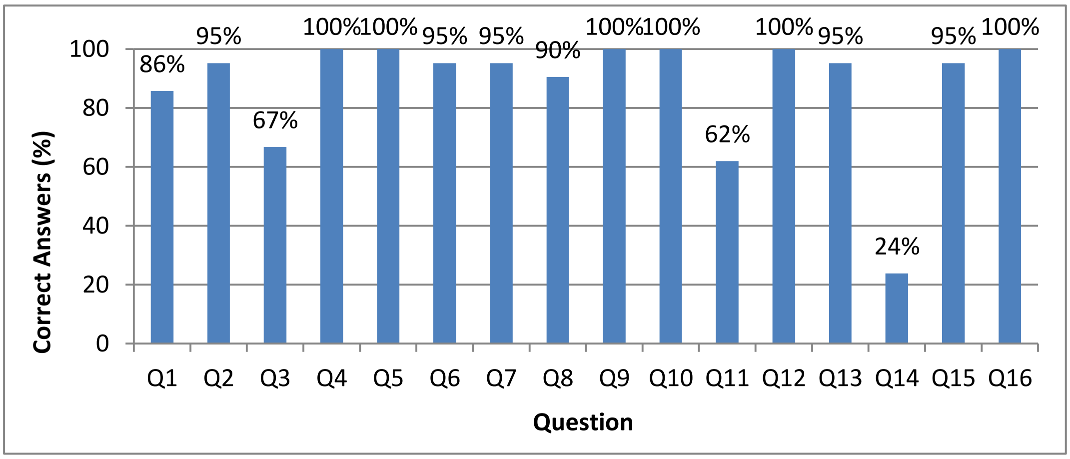

Figure 4.19: Percentage of Correct Answers for the 16 questions

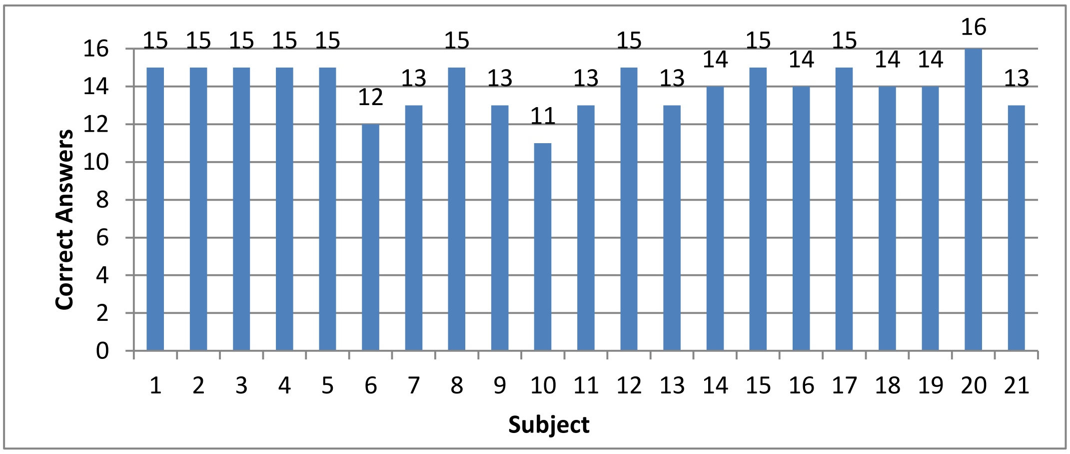

Figure 4.20: Individual Scores per Subject

A first conclusion that can be drawn is that although the bitrate was significantly reduced for all the sequences (78% decrease for the sequences corresponding to Q1-Q7, 57% decrease for the sequence corresponding to Q8, 80% decrease for the sequences corresponding to Q9-Q15, and 61% decrease for the sequence corresponding to Q16), the wide majority of the subjects were able to identify the correct answers for all the questions except Q14 (see Figure 4.20). The average percentage of correct answers across the 16 questions was 87%

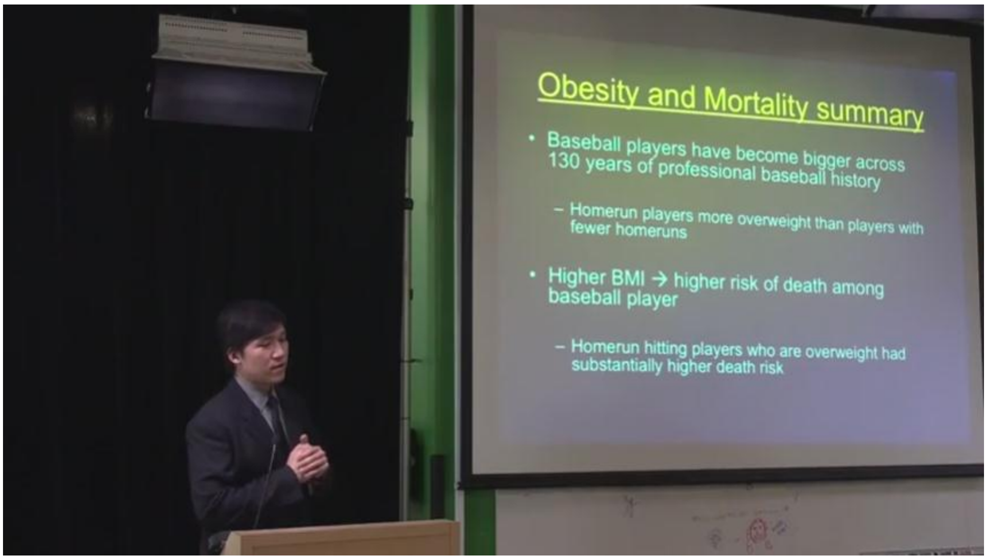

Figure 4.21: Frame Presenting the Answer for Q14

(STDEV = 20.64). In numbers the overall average translates as 18.43 (STDEV = 4.33) correct answers out of the maximum 21 possible.

As it can be seen in Figure 4.21, only 24% of the respondents were able to provide the right answer for Q14. For this particular question the subjects were asked to make their choice on what was the title of the slide presented in the test sequence. A frame containing the answer, which was extracted from the sequence displayed, having the video bitrate equal with 384 Kbps is presented in Figure 4.20. As it can be seen the video quality is good enough so that the text can be read. However despite that the slide was unchanged for the entire 30 seconds duration of the test sequence, only 5 of the 21subjects have selected the right answer. A possible explanation could be that the sequence was too short for the subjects to have time to read the title. However, only one of the 21 participants has indicated of not reading it. Another explanation could be that the other three choices were very similar, confusing the respondents.

To conclude this section, one can note that the video bitrate can be significantly reduced while still allowing learners to achieve knowledge to perceive easily visual elements that could be of interest for acquiring knowledge.

4.4 Chapter Summary

The research presented in this chapter, has investigated low bitrate thresholds for different categories of educational multimedia clips that still offer good perceived quality while maintaining the learners capacity to acquire knowledge from the visual information presented in these clips. Objective video quality assessment methods were used in order to determine the low / minimum reference bitrates.

Subjective tests conducted on 21 participants have confirmed that the video bitrate can be reduced with up to 80% from the reference high-quality value, depending on the content of the educational clips while maintaining the learners' ability to identify elements of interest in the visual information.

The following chapter, the final one, concludes the thesis and indicates possible future directions for extending the work presented in the previous chapters.

Chapter 5

Conclusions

This chapter summarises the research presented in this thesis highlighting the main findings on multimedia educational quality assessment.

5.1 Summary of Research

The research presented in this thesis has explored the idea of multimedia educational content adaptation in mobile learning systems.

The main investigations presented in this thesis were as follows:

- An assessment of the particularities of mobile devices commonly used for conducting m-learning activities; a novel classification of the mobile learning devices in classes of devices was proposed.
- Analysis of a collection of 904 educational multimedia clips currently used in education and their classification in different categories such as: screencasts, slideshows, animations, games / virtual learning environments recordings, interviews, lecture recordings / presentations, lab demos, and documentaries. The classification was performed based on the type of information presented in the clips.

- An investigation aimed to assess common encoding schemes used for educational multimedia content;
- An experimental study aimed to identify low bitrate thresholds for different categories of educational clips, which have been identified during the assessment of multimedia educational clips;
- A subjective experiment aimed to validate the thresholds identified during the objective study.

The first investigation, on mobile learning devices, was conducted with the main goal to assess current trends and particularities in terms of device sizes and form-factors, as well as device screen sizes, resolutions and aspect ratios. Several trends were identified and the need for device specific adaptation was confirmed.

The second investigation was conducted with the goal to identify, different categories of educational multimedia clips, based on the content type of the information presented. The particularities of the proposed categories, in terms of dynamicity or temporal-spatial movement were also identified.

The main purpose of the third investigation was to assess what are the most common encoding parameters of educational multimedia content. A significant number of educational clips from different sources were used for this investigation. Despite high variations of some of the encoding parameters studied, several preferred values were identified.

Starting from these investigations, ***a novel video profiling of multimedia educational clips, in the context of adaptive m-learning systems, and targeting different classes of device screen resolutions, has been proposed.*** Reference values for the encoding parameters were also proposed, as guidelines for achieving optimum quality on the target devices.

Motivated by the need of bitrate-based adaptive multimedia solutions, that have the potential to improve significantly learner quality of experience when watching educational clips delivered over poor network connections, two further studies were conducted in order to identify how much the reference bitrate values, associated to specific device profiles, can be decreased. The decrease should not significantly affect learner perceived quality, and should not impact learning process. Several categories of educational clips that have been identified were addressed during these studies. Objective video quality estimation metrics were explored as a first step for detecting the minimum thresholds, during an experimental study. The results of the first study were validated by a subjective experiment conducted on a significant number of subjects. The subjective study has also confirmed that learner potential to achieve knowledge from the visual information was not impacted.

5.2 Conclusions

The most significant conclusions of the investigations and studies that were conducted are presented.

From the first investigation the conclusion drawn was that adapting multimedia educational content based on the variety of learners' device screen resolutions would be difficult without grouping the devices in classes with similar properties. Therefore four classes of devices were proposed and a video profile was associated to each of these. The resolution of the proposed video profiles were selected targeting an approximately 50% difference between consecutive profiles.

The conclusion drawn from the second investigation was that a classification of multimedia educational clips based on the content type and on temporal-spatial features, can be made. Following eight categories of educational clips were identified: screencasts, slideshows, animations, games & virtual learning environments recordings, interviews, lecture recordings & presentations, lab demos, and documentaries.

From the third investigation the main conclusions that can be drawn are:

- H.264/MPEG-4 AVC, is the most popular video compression format;
- 320x240, 640x360, 640x480, 960x540, 1280x720 are the most common resolutions for educational multimedia clips;
- The most common frame rates are 29.97 fps, 25 fps and 23.976 fps.

The main conclusions drawn from the objective experiments, is that documentaries need at least a double minimum bitrate as compared to the other seven categories of educational clips that were tested, in order to achieve same level of objective quality.

The most important conclusions drawn from the subjective experiment are:

- Subjective MOS scores have confirmed that SSIM and PSNR objective metrics for assessing video quality, can be used for estimating low bitrate thresholds, that offer good quality levels, with good accuracy; out of the 16 test sequences, all have scored above 3.6 in average and only five above 4 (which corresponds to “Good”);
- The individual MOS scores corresponding to the two profiles that were tested were highly correlated, and the average scores were equal with 4.1 for both profiles. T-test analysis has shown that under the conditions studied statistically there was no difference between the scores for the two profiles. This fact suggests that each the proposed profiles offer an optimum level of quality across different devices.
- The reference bitrate can be reduced with up to 80% can be made, depending on the type of the educational clip, and depending on the profile, without significantly affecting the learners perceived video quality,

or their ability to answer questions related to the visual information.

- The wide majority of the subjects were able to answer the questions related to the information being presented, for all the questions except one;
- The average percentage of correct answers across the 16 questions was 87%, showing that the learners capacity to acquire knowledge from the visual information is not impacted.

5.3 Future Work

There are a number of avenues for further exploration of this research.

Since the majority of the participants to the subjective experiment did not answer a question for which the answer could be clearly read in the video, further research questions have appeared. One would be to identify the regions of interest in different types of educational clips such as the ones identified in this thesis. A correlation between the length of the test sequences and the learning outcome could be further explored.

Since very often the audio and video in a multimedia clip are correlated, it is very difficult to isolate learning concepts and test the actual impact of quality reduction on the learning outcome. More rigorous and controlled experiments, with longer test sequences would be needed to find if there really is a relationship between the two.

This work can be a good starting point in actually proposing a novel adaptive mechanism that takes into account the type of educational clips. Also the potential of the minimum bitrate thresholds for categorising the educational clips can be explored.

Appendix A

Subjective Test Material

Test Instructions

Welcome Message

Welcome to the subjective testing session organised by National College of Ireland.

Test Objectives

We are trying to assess the quality of different educational multimedia clips and the effect of the video quality on the learning process.

Disclaimer

Please fill in the personal information page. The information collected will be utilized as reference for analysing the perceptual test results and will never be made public in any form**.**

Test Directions

The test consists of **16** phases on two different mobile devices (8 phases per device). In each phase you will be shown a short multimedia clip and you will be asked to:

- Grade the overall quality on the indicated 1-5 scale, where 1 is the worst quality (“bad”) and 5 - the best (“excellent”).
- Answer a question related to the material presented in the clip.

You are kindly asked not to read the questions on the answer sheet before watching the corresponding clip.
Once a clip has started, you are not allowed to pause, stop, or review it.

Once the test has started you are not allowed to ask questions. The test was designed to take less than 30 minutes.

Table 1: Five-level quality scale for subjective testing (ITU-T Rec. P.911)

Rating	**Quality**
5	**Excellent**
4	**Good**
3	**Fair**
2	**Poor**
1	**Bad**

Personal Information Page

Please mark with an **X** your answer to the following questions.

Record No:	

Gender:

Male	**Female**

Age:

Are you a native English speaker?

Yes	**No**

Do you use glasses or contact lenses?

Yes	**No**

Do you have other visual conditions that may affect your perception of movies (e.g. colour blindness)?

Yes	**No**

How familiar are you with subjective video quality evaluation?

I work in the area	**I am familiar**	**I am not familiar**

Do you watch DVD/Blue Ray movies?

Every day	**Once a week**	**Once a month**	**Never**

Do you watch videos on the Internet using mobile devices (e.g. Smartphone, iPod, PDA, Notebook)?

Every day	**Once a week**	**Once a month**	**Never**

Do you watch educational content on mobile devices?

Every day	**Once a week**	**Once a month**	**Never**

You are watching an educational multimedia clip on your mobile device and there are transmission problems. What would you prefer?

a) To keep watching the video at its current quality even though you will perceive interruptions and loss of data. ☐

b) To be able to watch smoothly the entire clip at a lower, but acceptable quality level. ☐

Test Training Phase

Directions

Now we will show a short practice session to familiarise you with the test methodology. You will be given an opportunity after the practice session to ask any questions that you might have.

Following you are going to see a short multimedia clip.

<u>After</u> you have seen the clip, could you kindly answer the following questions about the sequence shown?
<u>Select one answer only.</u>

Phase No: 0

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q1. What were presented in the multimedia clip?

a) Cars
b) Clocks
c) Pianos
d) Thermometers

Questionnaire

Phase No: 1

How would you rate the overall quality of the multimedia clip?

5. Excellent	**4. Good**	**3. Fair**	**2. Poor**	**1. Bad**

Q1. What were displayed on the right panel of the “Modify Headers” window?

a) circles
b) images
c) nothing
d) buttons

Phase No: 2

How would you rate the overall quality of the multimedia clip?

5. Excellent	**4. Good**	**3. Fair**	**2. Poor**	**1. Bad**

Q2. What type of room is presented in this multimedia clip?

a) garage
b) bedroom
c) kitchen
d) bathroom

Phase No: 3

How would you rate the overall quality of the multimedia clip?

5. Excellent	**4. Good**	**3. Fair**	**2. Poor**	**1. Bad**

Q3. What was the big planet in the center doing?

a) moving
b) jumping

c) spinning
d) nothing

Phase No: 4

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q4. The speaker in this multimedia clip is:

a) sitting down?
b) standing up?

Phase No: 5

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q5. What is the origin of the woman?

a) Asian
b) Eschimo
c) Afro-American
d) European

Phase No: 6

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q6. The map of which country was presented on the slide?

a) Ireland
b) United States
c) Brazil
d) India

Phase No: 7

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q7. What colour was the liquid the woman was drinking from the glass?

a) had no color
b) green
c) blue
d) orange

Phase No: 8

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q8. Which of the animals presented in the video sleeps with one eye open (please fill in the answer)?

……………………… sleep with one eye open.

Phase No: 9

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q9. What was presented in this multimedia clip?

a) How to print a document.
b) How to download and install an application.
c) How to convert a video clip.
d) How to create a database.

Phase No: 10

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q10. Are there any pictures hanging on the walls?

a) yes
b) no

Phase No: 11

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q11. What is presented in this multimedia clip (please fill in the answer)?

The clip presents ……………………………..of the Sun.

Phase No: 12

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q12. What objects were displayed on the background in the multimedia clip?

a) benches
b) bicycles
c) flags
d) cars

Phase No: 13

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q13. The interview takes place:

a) in a study room
b) in a restaurant
c) in the garden
d) on a terrace

Phase No: 14

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q14. What was the title of the slide presented?

a) Obesity and mortality
b) I haven't read
c) Obesity and mortality summary
d) Obesity and mortality in MLB

Phase No: 15

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q15. What were the objects present on the table?

a) bowls
b) toys
c) knives
d) fruits

Phase No: 16

How would you rate the overall quality of the multimedia clip?

5. Excellent	4. Good	3. Fair	2. Poor	1. Bad

Q16. On average, how much sleep do new parents lose during the first year of the baby's life?

a) 20-40 hours
b) 300-600 hours
c) 700-900 hours

Bibliography

ABI Research. (2009). *385 Million Ultra Mobile Devices to Ship in 2014*. Allied Business Intelligence, Inc. Retrieved from http://www.abiresearch.com/press/1530-385+Million+Ultra+Mobile+Devices+to+Ship+in+2014

Adams, J., & Muntean, G. (2007). Adaptive-Buffer Power Save Mechanism for Mobile Multimedia Streaming. In *Proceedings of IEEE International Conference onCommunications 2007 (ICC '07)* (pp. 4548-4553). Glasgow, Scotland, UK: IEEE. doi:10.1109/ICC.2007.751

Ally, M., Lin, F., McGreal, R., & Woo, B. (2005). An Intelligent Agent for Adapting and Delivering Electronic Course Materials to Mobile Learners. Presented at the 4th World Conference on Mobile Learning (mLearn 2005), Cape Town, South Africa. Retrieved from http://www.mlearn.org.za/CD/papers/Ally.pdf

Anani, A., Zhang, D., & Li, H. (2008). M-learning in review: Technology, standard and evaluation. *Journal of Communication and Computer,*, *5*(11). Retrieved from http://www.informatics.org.cn/doc/ucit200811/ucit20081101.pdf

Apple. (2010, August 30). Apple - iTunes U - A wealth of knowledge from top institutions. *Apple*. Retrieved August 30, 2010, from http://www.apple.com/education/itunes-u/whats-on.html

Attewell, J. (2005). From Research and Development to Mobile Learning: Tools for Education and Training Providers and their Learners. Presented at the 4th World Conference on Mobile Learning, mLearn 2005, Cape Town, South Africa. Retrieved from http://www.mlearn.org.za/CD/papers/Attewell.pdf

Australian Flexible Learning Framework. (2010). *2010 E-standards for Training* (No. v1.0) (p. 17). Australian Flexible Learning Framework. Retrieved from http://e-standards.flexiblelearning.net.au/docs/2010-recommended-estandards-v1.0.pdf

Avcıbaş, İ., Sankur, B., & Sayood, K. (2002). Statistical Evaluation of Image Quality Measures. *Journal of Electronic Imaging*, *11*(2), 206-223. doi:10.1117/1.1455011

Barbosa, D., & Geyer, C. (2005). Pervasive Personal Pedagogical Agent: A Mobile AgentShall Always be with a Learner. In *Proceedings of IADIS International Conference Mobile Learning 2005* (pp. 281–285). Qawra, Malta: IADIS. Retrieved from http://www.iadis.net/dl/final_uploads/200506C022.pdf

Bassi, M., Steca, P., Delle Fave, A., & Caprara, G. V. (2006). Academic Self-Efficacy Beliefs and Quality of Experience in Learning. *Journal of Youth and Adolescence*, *36*(3), 301-312. doi:10.1007/s10964-006-9069-y

Berkovsky, S., Kuflik, T., & Ricci, F. (2007). Mediation of User Models for Enhanced Personalization in Recommender Systems. *User Modeling and User-Adapted Interaction*, *18*(3), 245-286. doi:10.1007/s11257-007-9042-9

Bomsdorf, B. (2005). Adaptation of Learning Spaces: Supporting Ubiquitous Learning in Higher Distance Education. In N. Davies, T. Kirste, & H. Schumann (Eds.), *Mobile Computing and Ambient Intelligence: The Challenge of Multimedia*, Dagstuhl Seminar Proceedings (Vol. 5181). Dagstuhl, Germany: Internationales Begegnungs- und Forschungszentrum für Informatik (IBFI), Schloss Dagstuhl, Germany. Retrieved from http://drops.dagstuhl.de/opus/volltexte/2005/371

Brezeale, D., & Cook, D. (2008). Automatic Video Classification: A Survey of the Literature. *IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews*, *38*(3), 416-430. doi:10.1109/TSMCC.2008.919173

Bright, P. (2010, April). Apple takes aim at Adobe... or Android? *Ars Technica*. Retrieved September 3, 2010, from http://arstechnica.com/apple/news/2010/04/apple-takes-aim-at-adobe-or-android.ars

Brown, E. J., Brailsford, T. J., Fisher, T., & Moore, A. (2009). Evaluating Learning Style Personalization in Adaptive Systems: Quantitative Methods and Approaches. *IEEE Transactions on Learning Technologies*, *2*(1), 10-22. doi:10.1109/TLT.2009.11

Brusilovsky, P. (2004). Adaptive navigation support: From adaptive hypermedia to the adaptive web and beyond. *PsychNology Journal*, *2*(1), 7–23.

Brusilovsky, P., & Peylo, C. (2003). Adaptive and Intelligent Web-based Educational Systems. *International Journal of Artificial Intelligence in Education*, *13*(2-4), 159–172.

Brusilovsky, P. (2001). Adaptive Hypermedia. *User Modeling and User-Adapted Interaction*, *11*(1), 87-110. doi:10.1023/A:1011143116306

Brusilovsky, P., & Millán, E. (2007). User Models for Adaptive Hypermedia and Adaptive Educational Systems. In P. Brusilovsky, A. Kobsa, & W. Nejdl (Eds.), *The Adaptive Web*, Lecture Notes in Computer Science (Vol. 4321, pp. 3-53). Berlin, Heidelberg: Springer-Verlag. Retrieved from http://www.dcs.warwick.ac.uk/~acristea/courses/CS411/2009/Book%20-%20The%20Adaptive%20Web/UM.pdf

Chang, H. P., Shih, T. K., Li, Q., Wang, C., Wang, T., & Chao, L. R. (2008). An Adaptive Caching Strategy for m-Learning Based on SCORM Sequencing and Navigation. *World Wide Web*, *11*(3), 387-406. doi:10.1007/s11280-008-0044-2

Chang, S., & Vetro, A. (2005). Video Adaptation: Concepts, Technologies, and Open Issues. *Proceedings of the IEEE*, *93*(1), 148-158. doi:10.1109/JPROC.2004.839600

Chen, J., Lu, H., Wei, R., Jin, C., & Xue, X. (2010). An effective method for video genre classification. In *Proceedings of the ACM International Conference on Image and Video Retrieval* (pp. 97-104). Xi'an, China: ACM New York, NY, USA. doi:10.1145/1816041.1816058

Chen, N., Lin, K., & Kinshuk. (2008). Analysing Users' Satisfaction with E-Learning Using a Negative Critical Incidents Approach. *Innovations in Education and Teaching International*, *45*(2), 115-126. doi:10.1080/14703290801950286

Chen, S. Y., & Macredie, R. D. (2002). Cognitive styles and hypermedia navigation: Development of a learning model. *Journal of the American Society for Information Science and Technology*, *53*(1), 3-15. doi:10.1002/asi.10023

Cherian, E. J., & Williams, P. (2008). Mobile Learning: The Beginning of the End of Classroom Learning. In *Proceedings of the World Congress on Engineering and Computer Science 2008 (WCECS 2008)*. San Francisco, CA: IAENG. Retrieved from http://www.iaeng.org/publication/WCECS2008/WCECS2008_pp508-514.pdf

Choudary, C., & Liu, T. (2007). Summarization of Visual Content in Instructional Videos. *IEEE Transactions on Multimedia*, *9*(7), 1443-1455. doi:10.1109/TMM.2007.906602

Cisco. (2010). *Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update, 2009-2014* (p. 15). Retrieved from http://www.cisco.com/en/US/solutions/collateral/ns341/ns525/ns537/ns705/ns827/white_paper_c11-520862.html

Clifford, R. (2000). Adaptive Hypermedia for Music Instruction. In *Proceedings of the 7th International Technological Directions in Music Learning Conference*. Institute for Music Research, University of Texas at San Antonio. Retrieved from http://music.utsa.edu/tdml/conf-VII/VII-Clifford/VII-Clifford.html

Clough, G., Jones, A., McAndrew, P., & Scanlon, E. (2008). Informal Learning with PDAs and Smartphones. *Journal of Computer Assisted Learning*, *24*(5), 359-371. doi:10.1111/j.1365-2729.2007.00268.x

Cobcroft, R. S. (2006). *Literature Review into Mobile Learning in the University Context* (p. 138). Brisbane, Australia: Queensland University of Technology, Creative Industries Faculty. Retrieved from http://eprints.qut.edu.au/4805/

Coda Reasearch Consultancy. (2010). *Worldwide Smartphone Sales Forecast to 2015* (p. 36). Guildford, England: Coda Research Consultancy Ltd. Retrieved from http://www.codarc.co.uk/smartphoneworld/index.htm

Copley, J. (2007). Audio and Video Podcasts of Lectures for Campus-Based Students: Production and Evaluation of Student Use. *Innovations in Education and Teaching International*, *44*(4), 387-399. doi:10.1080/14703290701602805

Corriveau, P. (2005a). Video Quality Experts Group. In H. Wu & K. Rao (Eds.), *Digital Video Image Quality and Perceptual Coding* (1st ed., pp. 325-334). CRC Press. Retrieved from http://video.ee.ntu.edu.tw/~sernc/books/Digital%20Video%20Image%20Quality%20and%20Perceptual%20Coding/

Corriveau, P. (2005b). Video Quality Testing. In H. Wu & K. Rao (Eds.), *Digital Video Image Quality and Perceptual Coding* (1st ed., pp. 125-155). CRC Press. Retrieved from http://video.ee.ntu.edu.tw/~sernc/books/Digital%20Video%20Image%20Quality%20and%20Perceptual%20Coding/

De Bra, P., Aerts, A., Berden, B., de Lange, B., Rousseau, B., Santic, T., Smits, D., et al. (2003). AHA! The Adaptive Hypermedia Architecture. In *Proceedings of the 14th ACM Conference on Hypertext and Hypermedia* (pp. 81-84). Nottingham, UK: ACM Press New York, NY, USA. doi:10.1145/900051.900068

De Bra, P., Brusilovsky, P., & Houben, G. J. (1999). Adaptive Hypermedia: From Systems to Framework. *ACM Computing Surveys (CSUR)*, *31*(4es). doi:10.1145/345966.345996

Dede, C. (2008). Theoretical Perspectives Influencing the Use of Information Technology in Teaching and Learning. (J. Voogt & G. Knezek, Eds.)*International Handbook of Information Technology in Primary and Secondary Education*, Springer International Handbooks of Education, *20*(1), 43–62. doi:10.1007/978-0-387-73315-9_3

Dhillon, P., Nowozin, S., & Lampert, C. (2009). Combining Appearance and Motion for Human Action Classification in Videos. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern*

Recognition Workshops, 2009 (CVPR Workshops 2009) (pp. 22-29). Miami, Florida, USA: IEEE Explore. doi:10.1109/CVPRW.2009.5204237

Dickson, P. E., Adrion, W. R., Hanson, A. R., & Arbour, D. T. (2009). First Experiences with a Classroom Recording System. In *Proceedings of the 14th Annual Conference on Innovation and Technology in Computer Science Education 2009 (ITiCSE '09)* (pp. 298-302). Paris, France. doi:10.1145/1562877.1562968

van Diggelen, F. (2009, December 1). The Smartphone Revolution: Seven technologies that put GPS in mobile phones around the world - the how and why of location's entry into modern consumer mobile communications. *GPS World*. Retrieved from http://www.gpsworld.com/wireless/smartphone-revolution-9183

Downes, S. (2005). E-learning 2.0. *eLearn Magazine*, *2005*(10). Retrieved from http://www.elearnmag.org/subpage.cfm?section=articles&article=29-1

Evans, C. (2008). The Effectiveness of M-Learning in the Form of Podcast Revision Lectures in Higher Education. *Computers & Education*, *50*(2), 491-498. doi:10.1016/j.compedu.2007.09.016

Fayzullin, M., Subrahmanian, V. S., Picariello, A., & Sapino, M. L. (2005). The CPR Model for Summarizing Video. *Multimedia Tools and Applications*, *26*(2), 153–173. doi:10.1007/s11042-005-0451-7

Fleischman, G. (2010). The state of 4G: it's all about congestion, not speed. *Ars Technica*. Technology News & Info, . Retrieved June 3, 2010, from http://arstechnica.com/telecom/news/2010/03/faster-mobile-broadband-driven-by-congestion-not-speed.ars

Fleischman, M., Decamp, P., & Roy, D. (2006). Mining temporal patterns of movement for video content classification. In *Proceedings of the 8th ACM International Workshop on Multimedia Information Retrieval 2006* (pp. 183-192). Santa Barbara, CA, USA: ACM New York, NY, USA. doi:10.1145/1178677.1178704

Flores, N., & Savage, S. J. (2007). Student Demand for Streaming Lecture Video: Empirical Evidence from Undergraduate Economics Classes. *International Review of Economics Education*, *6*(2), 57–78.

Gartner, I. (2010, May 19). Worldwide Smartphone Sales to End Users by Operating System in 1Q10. *Gartner Newsroom*. Press Release, . Retrieved July 15, 2010, from http://www.gartner.com/it/page.jsp?id=1372013

Geddes, S. J. (2004). Mobile learning in the 21st century: Benefit for learners. *The Knowledge Tree e-Journal*, *30*(3), 214-228.

Georgiev, T., Georgieva, E., & Smrikarov, A. (2004). M-Learning - a New Stage of E-Learning. In *Proceedings of the 5th International Conference on Computer Systems and Technologies (CompSysTech2004)*. Rousse, Bulgaria: ACM New York, NY, USA. doi:10.1145/1050330.1050437

Glavinic, V., Ljubic, S., & Kukec, M. (2008). A Holistic Approach to Enhance Universal Usability in m-Learning. In *Proceedings of the 2nd International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2008)* (pp. 305-310). Valencia, Spain: IEEE. doi:10.1109/UBICOMM.2008.54

Goyal, V. K. (2001). Multiple Description Coding: Compression Meets the Network. *IEEE Signal Processing Magazine*, *18*(5), 74–93. doi:10.1109/79.952806

GSA. (2010, July). GSM/3G Market Update. The Global mobile Suppliers Association. Retrieved from http://www.gsacom.com/gsm_3g/market_update.php4#GSM_3G_Market_Update

Gütl, C., Pivec, M., Trummer, C., García-Barrio, V. M., Mödritscher, F., Pripfl, J., & Umgeher, M. (2005). AdeLE (Adaptive e-Learning with Eye-Tracking): Theoretical Background, System Architecture and Application Scenarios. *European Journal of Open, Distance and E-Learning (EURODL)*, (2005/II), 12.

Hampapur, A., Weymouth, T., & Jain, R. (1994). Digital Video Segmentation. In *Proceedings of the 2nd ACM International Conference on Multimedia* (pp. 357-364). San Francisco, CA, United States: ACM New York, NY, USA. doi:10.1145/192593.192699

Hamza-Lup, F., & Adams, M. (2008). Feel the Pressure: E-learning Systems with Haptic Feedback. In *Proceedings of the 16th International Symposium on Haptic Interfaces for Virtual Environment and Teleoperator Systems 2008* (pp. 445-450). Reno, Nevada, USA: IEEE Computer Society Washington, DC, USA. doi:10.1109/HAPTICS.2008.4479991

Hanani, U., Shapira, B., & Shoval, P. (2001). Information Filtering: Overview of Issues, Research and Systems. *User Modeling and User-Adapted Interaction*, *11*(3), 203-259. doi:10.1023/A:1011196000674

Hestnes, B., Brooks, P., & Heiestad, S. (2008). *QoE (Quality of Experience) – measuring QoE for improving the usage of telecommunication services* (No. R 21/2009) (p. 31). Telenor R&I. Retrieved from http://www.telenor.com/en/resources/images/r21_09_tcm28-47617.pdf

HHI. (n.d.). The Scalable Video Coding Amendment of the H.264/AVC Standard. *Fraunhofer Institute for Telecommunications, Heinrich Hertz Institute*. Retrieved June 26, 2010, from http://www.hhi.fraunhofer.de/en/departments/image-processing/image-communication/video-coding/svc-scalable-extension-of-h264avc/

Ideses, I., Yaroslavsky, L., & Fishbain, B. (2007). Real-time 2D to 3D video conversion. *Journal of Real-Time Image Processing*, *2*(1), 3-9. doi:10.1007/s11554-007-0038-9

ITU-R. (2002). *BT.500-11: Methodology for the subjective assessment of the quality of television pictures* (p. 48). Geneva, Switzerland: International Telecommunication Union. Retrieved from http://www.itu.int/rec/R-REC-BT.500/en

ITU-T. (2004). *J.144 : Objective perceptual video quality measurement techniques for digital cable television in the presence of a full reference* (p. 156). Geneva, Switzerland: International Telecommunication Union. Retrieved from http://www.itu.int/rec/T-REC-J.144-200403-I/en

ITU-T. (2008). *P.910: Subjective video quality assessment methods for multimedia applications* (p. 42). Geneva, Switzerland: International Telecommunication Union. Retrieved from http://www.itu.int/rec/T-REC-P.910-200804-I/en

ITU-T. (2009). *H.264 : Advanced video coding for generic audiovisual services* (p. 670). Geneva, Switzerland: International Telecommunication Union. Retrieved from http://www.itu.int/rec/T-REC-H.264-200903-S/en

Jumisko-Pyykkö, S., Häkkinen, J., & Nyman, G. (2007). Experienced Quality Factors-Qualitative Evaluation Approach to Audiovisual Quality. In *Proceedings of IST/SPIE Conference on Electronic Imaging, Multimedia on Mobile Devices* (Vol. 6507, pp. 28–31). San Jose, CA, USA. doi:10.1117/12.699797

Kanumuri, S., Cosman, P., Reibman, A., & Vaishampayan, V. (2006). Modeling packet-loss visibility in MPEG-2 video. *IEEE Transactions on Multimedia*, *8*(2), 341-355. doi:10.1109/TMM.2005.864343

Kaufman, P., & Mohan, J. (2009). *Video use and higher education: Options for the future* (p. 14). New York, USA: Intelligent Television, New York University. Retrieved from http://library.nyu.edu/about/Video_Use_in_Higher_Education.pdf

Khan, S. (2009). Khan Academy Vision. Khan Academy. Retrieved from http://www.khanacademy.org/

Koohang, A., Riley, L., Smith, T., & Schreurs, J. (2009). E-Learning and Constructivism: From Theory to Application. *Interdisciplinary Journal of E-Learning and Learning Objects*, *5*, 91–109.

Krieger, L. (2010, June 27). A global teacher of 1,516 lessons and counting. *PHYSORG.com*. Science & Tech News, . Retrieved from http://www.physorg.com/news196868176.html

Kuo, Y. H., & Huang, Y. M. (2009). MEAT: An Authoring Tool for Generating Adaptable Learning Resources. *Journal of Educational Technology & Society*, *12*(2), 51–68.

Kwatra, V., Schodl, A., Essa, I., Turk, G., & Bobick, A. (2003). Graphcut Textures: Image and Video Synthesis Using Graph Cuts. *ACM Transactions on Graphics*, *22*(3), 277–286. doi:10.1145/882262.882264

Laouris, Y., & Eteokleous, N. (2005). We Need an Educationally Relevant Definition of Mobile Learning. In *Proceedings of The 4th World Conference on Mobile Learning (mLearn 2005)* (pp. 290-294). Cape Town, South Africa. Retrieved from http://www.mlearn.org.za/CD/papers/Laouris%20&%20Eteokleous.pdf

Leung, C. H., & Chan, Y. Y. (2003). Mobile Learning: A New Paradigm in Electronic Learning. In *Proceedings of The 3rd IEEE International*

Conference on Advanced Learning Technologies (ICALT'03) (pp. 76-80). Athens, Greece: IEEE Computer Society. Retrieved from http://www.computer.org/portal/web/csdl/doi/10.1109/ICALT.2003.1215030

Liaw, S., Hatala, M., & Huang, H. (2010). Investigating Acceptance toward Mobile Learning to Assist Individual Knowledge Management: Based on Activity Theory Approach. *Computers & Education*, *54*(2), 446-454. doi:10.1016/j.compedu.2009.08.029

Low, L. (2007). *M-learning standards report - Background, discussion and recommendations for usable and accessible m-learning* (No. v1.0) (p. 41). Australian Flexible Learning Framework. Retrieved from http://e-standards.flexiblelearning.net.au/docs/m-standards-report-v1-0.pdf

Marri, S. (2007). *Rich Media Content Adaptation in E-Learning Systems* (Ph.D.). Università di Bologna, Bologna, Italy. Retrieved from http://amsdottorato.cib.unibo.it/375/

Mayer, R., & Moreno, R. (1998). A Split-Attention Effect in Multimedia Learning: Evidence for Dual Processing Systems in Working Memory. *Journal of Educational Psychology*, *90*(2), 312-320. doi:10.1037/0022-0663.90.2.312

Meeker, M. (2009, October 20). *Economy + Internet Trends*. Presented at the 2009 Web 2.0 Summit, San Francisco, CA. Retrieved from http://www.morganstanley.com/institutional/techresearch/pdfs/MS_Economy_Internet_Trends_102009_FINAL.pdf

Mertens, R., Birnbaum, N., Ketterl, M., & Rolf, R. (2008). Integrating Lecture Recording with an LMS: An Implementation Report. In *Proceedings of World Conference on E-Learning, in Corporate, Government, Healthcare & Higher Education (E-Learn 2008)* (pp. 1067-1074). Las Vegas, Nevada, USA: AACE, Chesapeake, VA, USA. Retrieved from http://www.inf.uos.de/papers_pdf/2008_06.pdf

Mittal, A., Krishnan, P. V., & Altman, E. (2006). Content Classification and Context-Based Retrieval System for E-Learning. *Educational Technology & Society*, *9*(1), 349-358.

Money, A. G., & Agius, H. (2008). Video Summarisation: A Conceptual Framework and Survey of the State of the Art. *Journal of Visual Communication and Image Representation*, *19*(2), 121-143. doi:10.1016/j.jvcir.2007.04.002

Muntean, C. H., & McManis, J. (2006). End-User Quality of Experience Oriented Adaptive E-learning System. *Journal of Digital Information*, *7*(1). Retrieved from https://journals.tdl.org/jodi/article/view/179/159

Muntean, C., & Muntean, G. (2007). End-User Quality of Experience Aware Personalised E-Learning. In C. Pahl (Ed.), *Architecture Solutions for E-Learning Systems* (1st ed., Vols. 1-1, pp. 154-174). Hershey, PA: IGI Global. Retrieved from http://www.igi-global.com/bookstore/TitleDetails.aspx?TitleId=5234

Muntean, C. H. (2008). Improving learner quality of experience by content adaptation based on network conditions. *Computers in Human Behavior*, *24*(4), 1452-1472. doi:10.1016/j.chb.2007.07.016

Muntean, G., Ghinea, G., & Sheehan, T. N. (2008). Region of Interest-Based Adaptive Multimedia Streaming Scheme. *IEEE Transactions on Broadcasting*, *54*(2), 296-303. doi:10.1109/TBC.2008.919012

Nagai, T. (2009). Automated Lecture Recording System with AVCHD Camcorder and Microserver. In *Proceedings of the 37th International Conference on University and College Computing Services (SIGUCCS '09)* (pp. 47-54). St. Louis, Missouri, USA: ACM New York, NY, USA. doi:10.1145/1629501.1629512

Naismith, L., Lonsdale, P., Vavoula, G., & Sharples, M. (2004). *Literature Review in Mobile Technologies and Learning* (Literature Reviews No. 11). Futurelab Series (p. 48). Bristol, UK: Futurelab. Retrieved from http://www.futurelab.org.uk/resources/publications-reports-articles/literature-reviews/Literature-Review203

Naismith, L., & Smith, M. P. (2009). Using Mobile Technologies for Multimedia Tours in a Traditional Museum Setting. In M. Ally (Ed.), *Mobile Learning: Transforming the Delivery of Education and Training*, Distance Education Series (pp. 247–264). Edmonton, AB, Canada: Athabasca University Press. Retrieved from http://www.aupress.ca/books/120155/ebook/12_Mohamed_Ally_2009-Article12.pdf

Najjar, L. J. (1996). Multimedia Information and Learning. *Journal of Educational Multimedia and Hypermedia*, *5*(2), 129-150.

Nemethova, O., Ries, M., Siffel, E., & Rupp, M. (2004). Quality Assessment for H.264 Coded Low-Rate and Low-Resolution Video Sequences. In *Proceedings of 3rd Conference on Communications, Internet, and Information Technology (CIIT 2004)* (pp. 136–140). St. Thomas, US Virgin Islands. Retrieved from http://publik.tuwien.ac.at/files/pub-et_8787.pdf

Neo, M., & Neo, K. T. (2001). Innovative teaching: Using multimedia in a problem-based learning environment. *Educational Technology & Society*, *4*(4), 19–31.

Niebles, J. C., & Fei-Fei, N. L. (2007). A Hierarchical Model of Shape and Appearance for Human Action Classification. In *Proceedings of IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2007 (CVPR '07)* (Vol. 0, pp. 1-8). Minneapolis, MN, USA: IEEE Computer Society. doi:10.1109/CVPR.2007.383132

Nyiri, K. (2002). Towards a Philosophy of M-Learning. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education (WMTE'02)* (pp. 121-124). Växjö, Sweden: IEEE Explore. doi:10.1109/WMTE.2002.1039233

O'Connell, M., & Smith, J. (2007). *A guide to working with m-learning standards* (No. v1.0) (p. 38). Australian Flexible Learning Framework. Retrieved from http://e-standards.flexiblelearning.net.au/docs/m-standards-guide-v1-0.pdf

Ogata, H., Hui, G. L., Yin, C., Ueda, T., Oishi, Y., & Yano, Y. (2008). Supporting Mobile Language Learning outside Classrooms. *International Journal of Mobile Learning and Organisation*, *2*(3), 271-282. doi:10.1504/IJMLO.2008.020319

Open University. (2010, August 23). iTunes U » Impact. *The Open University on iTunes U.* Retrieved August 30, 2010, from http://projects.kmi.open.ac.uk/itunesu/impact/

Oud, J. (2009). Guidelines for effective online instruction using multimedia screencasts. *Reference Services Review*, *37*(2), 164-177. doi:10.1108/00907320910957206

Paas, F., Renkl, A., & Sweller, J. (2003). Cognitive Load Theory and Instructional Design: Recent Developments. *Educational Psychologist*, *38*(1), 1-4. doi:10.1207/S15326985EP3801_1

Padhye, J., Kurose, J., Towsley, D., & Koodli, R. (1999). A Model Based TCP-Friendly Rate Control Protocol. In *Proceedings of the 9th ACM International Workshop Network and Operating System Support for Digital Audio and Video (NOSSDAV'99)*. Basking Ridge, NJ, USA: ACM Press New York, NY, USA. Retrieved from http://www.nossdav.org/1999/papers/49-248339128.ps.gz

Papanikolaou, K. A., Grigoriadou, M., Kornilakis, H., & Magoulas, G. D. (2003). Personalizing the Interaction in a Web-based Educational Hypermedia System: the case of INSPIRE. *User Modeling and User-Adapted Interaction*, *13*(3), 213–267. doi:10.1023/A:1024746731130

Papanikolaou, K., & Mavromoustakos, S. (2006). Critical Success Factors for the Development of Mobile Learning Applications. In A. C. Boucouvalas (Ed.), *Proceedings of the 2nd IASTED European Conference on Internet and Multimedia Systems and Applications (EuroIMSA 2006)* (pp. 19-24). Innsbruck, Austria: ACTA Press.

Pinson, M., & Wolf, S. (2003). Comparing subjective video quality testing methodologies. In *Proceedings of SPIE Video Communications and Image Processing Conference 2003* (Vol. 5150, pp. 573-582). Lugano, Switzerland: SPIE. doi:10.1117/12.509908

Pinson, M. H., & Wolf, S. (2004). A New Standardized Method for Objectively Measuring Video Quality. *IEEE Transactions on Broadcasting*, *50*(3), 312–322. doi:10.1109/TBC.2004.834028

Popescu, E., Trigano, P., & Badica, C. (2007). Towards a Unified Learning Style Model in Adaptive Educational Systems. In *Proceedings of the 7th IEEE International Conference on Advanced Learning Technologies 2007 (ICALT 2007)* (pp. 804-808). Niigata, Japan: IEEE Explore. doi:10.1109/ICALT.2007.263

Quinn, C. (2008). *mLearning Devices*. Quinnovation. Retrieved from http://quinnovation.com:8000/MobileDevices.pdf

Repp, S., & Meinel, M. (2006). Semantic Indexing for Recorded Educational Lecture Videos. In *Proceedings of the 4th Annual IEEE International Conference on Pervasive Computing and Communications Workshops, 2006 (PERCOMW'06)* (pp. 240-245). Pisa, Italy: IEEE Explore. doi:10.1109/PERCOMW.2006.122

Richardson, I., & Kannangara, C. (2004). Fast subjective video quality measurement with user feedback. *Electronics Letters*, *40*(13), 799-801. doi:10.1049/el:20040553

Richardson, I. E. G. (2003). *H.264 and MPEG-4 Video Compression: Video Coding for Next Generation Multimedia* (1st ed.). Chichester, West Sussex, England: John Wiley and Sons. Retrieved from http://guohanwei.51.net/book/H.264_And_MPEG-4_Video_Compression.pdf

Ries, M., Nemethova, O., & Rupp, M. (2007). Motion Based Reference-Free Quality Estimation for H.264/AVC Video Streaming. In *Proceedings of the 2nd International Symposium on Wireless Pervasive Computing 2007 (ISWPC'07)* (pp. 355–359). San Juan, Puerto Rico: IEEE Explore. doi:10.1109/ISWPC.2007.342629

Rodrigues, J. J. P. C., Veiga, I. D. D. C., & Vaidya, B. (2010). Implementation of Location-Aware M-Learning System. In *Proceedings of the 2nd International Conference on Mobile, Hybrid, and On-Line Learning 2010 (eL&mL 2010)* (pp. 82-86). St. Maarten, Netherlands Antilles. doi:10.1109/eLmL.2010.16

Rogers, Y., Price, S., Randell, C., Fraser, D. S., Weal, M., & Fitzpatrick, G. (2005). Ubi-learning: Integrating Indoor and Outdoor Learning Experiences. *Communications of the ACM*, *48*(1), 55–59. doi:10.1145/1039539.1039570

Rogers, Y., & Scaife, M. (1997). How can interactive multimedia facilitate learning. In J. Lee (Ed.), *Intelligence and Multimodality in Multimedia Interfaces: Research & Applications*. Menlo Park, CA, USA: Association for the Advancement of Artificial Intelligence. Retrieved from http://mcs.open.ac.uk/yr258/papers/Rogers_Scaife98.pdf

Rosenberg, M. J. (2001). *E-Learning: Strategies for Delivering Knowledge in the Digital Age* (1st ed.). New York: The McGraw Hill Companies.

Rosman, P. (2008). M-learning-as a paradigm of new forms in education. *Ekonomie + Management*, *1*, 119–125.

Rouvier, M., Linares, G., & Matrouf, D. (2009). Robust Audio-based Classification of Video Genre. In *Proceedings of the 10th Annual Conference of the International Speech Communication Association (InterSpeech 2009)* (pp. 1159–1162). Brighton, U.K. Retrieved from http://sd-8971.dedibox.fr/files/interspeech_09.pdf

Safran, C., Guetl, C., & Helic, D. (2007). The Impact of Web 2.0 on Learning at a Technical University - A Usage Survey. In T. Bastiaens & S. Carliner (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2007 (E-Learn 2007)* (pp. 436-443). Quebec City, Canada: AACE, Chesapeake, VA, USA. Retrieved from http://www.iicm.tugraz.at/iicm_papers/e-learn2007_csafran.pdf

Sampson, D., & Karagiannidis, C. (2002). Personalised Learning: Educational, Technological and Standardisation Perspective. *Interactive Educational Multimedia*, *4*, 24-39.

Schulte, O. A., Wunden, T., & Brunner, A. (2008). REPLAY: an integrated and open solution to produce, handle, and distributeaudio-visual (lecture) recordings. In *Proceedings of the 36th Annual ACM SIGUCCS Conference on User Services 2008* (pp. 195-198). Portland, OR, USA: ACM. doi:10.1145/1449956.1450016

Schwabe, G., & Goth, C. (2005). Navigating and Interacting Indoors with a Mobile Learning Game. In *Proceedings of the IEEE International Workshop on Wireless and Mobile Technologies in Education, 2005 (WMTE 2005)* (pp. 192-199). Tokushima, Japan: IEEE Computer Society. doi:10.1109/WMTE.2005.47

Schwarz, H., Marpe, D., & Wiegand, T. (2007). Overview of the Scalable Video Coding Extension of the H.264/AVC Standard. *IEEE Transactions on*

Circuits and Systems for Video Technology, *17*(9), 1103–1120. doi:10.1109/TCSVT.2007.905532

Shank, P. (2006). The Value of Multimedia in Learning. *Adobe Motion Design Center*, *10*, 12.

Sharma, S. K., & Kitchens, F. L. (2004). Web Services Architecture for M-Learning. *Electronic Journal on e-Learning Volume*, *2*(1), 203–216. doi:10.1.1.114.5158

Sharples, M. (2007). *Big Issues in Mobile Learning: Report of a workshop by the Kaleidoscope Network of Excellence Mobile Learning Initiative* (p. 40). Nottingham, UK: Learning Sciences Research Institute, University of Nottingham. Retrieved from http://mlearning.noe-kaleidoscope.org/public/news/KALEIDOSCOPE%20REPORT_07_Big_Issues_In_Mobile_Learning.pdf

Shepard, A. (2003). A Case for Computer-based Multimedia in Adult Literacy Classrooms. *Encyclopedia of Educational Technology*, *San Diego State University*. Retrieved from http://edweb.sdsu.edu/EET/articles/adultliteracy/index.htm

Smoliar, S., & Zhang, H. J. (1994). Content-Based Video Indexing and Retrieval. *IEEE Multimedia*, *1*(2), 62-72. doi:10.1109/93.311653

Snoek, C. G., & Worring, M. (2005). Multimodal Video Indexing: A Review of the State-of-the-art. *Multimedia Tools and Applications*, *25*(1), 5–35. doi:10.1023/B:MTAP.0000046380.27575.a5

Soong, S. K., Chan, L. K., Cheers, C., & Hu, C. (2006). Impact of video recorded lectures among students. In *Proceedings of the The 23rd Annual Conference of the Australasian Society for Computers in Learning in Tertiary Education (ascilite 2006)* (pp. 789–793). Sydney, Australia: Sydney University Press. Retrieved from http://www.ascilite.org.au/conferences/sydney06/proceeding/pdf_papers/p179.pdf

Stash, N., Cristea, A., & De Bra, P. (2006). Adaptation to Learning Styles in E-Learning: Approach Evaluation. In T. Reeves & S. Yamashita (Eds.), *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2006 (E-LEARN 2006)* (pp. 284–291). Honolulu, Hawaii, USA: AACE, Chesapeake, VA, USA. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.107.8619&rep=rep1&type=pdf

Stockley, D. (2003). E-learning Definition and Explanation (Elearning, Online Training, Online Learning). *Derek Stockley*. Retrieved from http://derekstockley.com.au/elearning-definition.html

Susstrunk, S., & Winkler, S. (2004). Color image quality on the internet. In *Proceedings of IST/SPIE Conference on Electronic Imaging, Multimedia on Mobile Devices* (Vol. 5304, pp. 118–131). San Jose, CA, USA. doi:10.1117/12.537804

Tekalp, A. M. (2005). Video segmentation. In A. C. Bovik (Ed.), *Handbook of Image and Video Processing* (2nd ed., pp. 471-489). Amsterdam; Boston, MA, USA: Elsevier Academic Press. Retrieved from http://books.google.ie/books?id=OYFYt5C4N94C

The Free Dictionary. (n.d.). Transcode Definition. In *TheFreeDictionary's Encyclopedia*. Online: Farlex, Inc. Retrieved from http://encyclopedia2.thefreedictionary.com/transcode

The Learning CONSORTIUM. (2002). *Making Sense of Learning Specifications & Standards: A Decision Maker's Guide to their Adoption*. New York, USA: The Masie Center. Retrieved from http://www.staffs.ac.uk/COSE/cosenew/s3_guide.pdf

Thompson, T. (2007). Using Closed Captions for Accessibility and Full-text Video Search. In *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2007, EDMEDIA 2007* (Vol. 2007, pp. 4475-4478). Retrieved from http://editlib.org/p/26024

Tiobe. (2010, July). TIOBE Programming Community Index for July 2010. *TIOBE Software - The Coding Standards Company*. Retrieved July 15, 2010, from http://www.tiobe.com/index.php/content/paperinfo/tpci/index.html

Traxler, J. (2007). Defining, Discussing and Evaluating Mobile Learning: The moving finger writes and having writ.... *International Review of Research* *in Open and Distance Learning*, *8*(2). Retrieved from http://www.irrodl.org/index.php/irrodl/article/viewArticle/346/875

Triantafillou, E., Pomportsis, A., & Demetriadis, S. (2003). The Design and the Formative Evaluation of an Adaptive Educational System Based on Cognitive Styles. *Computers & Education*, *41*(1), 87–103. doi:10.1016/S0360-1315(03)00031-9

Trifonova, A., Knapp, J., Ronchetti, M., & Gamper, J. (2004). Mobile ELDIT: Transition from an e-Learning to an m-Learning System. In L. Cantoni & C. McLoughlin (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2004 (ED-MEDIA 2004)* (pp. 188-193). Lugano, Switzerland: AACE, Chesapeake, VA, USA. Retrieved from http://www.trifonova.net/docs/Mobile%20Eldit%20%28Ed-Media04%29%20poster.pdf

Tusch, R., Böszörményi, L., Goldschmidt, B., Hellwagner, H., & Schojer, P. (2004). Offensive and Defensive Adaptation in Distributed Multimedia Systems. *Computer Science and Information Systems, ComSIS*, *1*(1), 45–73.

Varlamis, I., & Apostolakis, I. (2006). The Present and Future of Standards for E-Learning Technologies. *Interdisciplinary Journal of Knowledge and Learning Objects*, *2*, 59–76.

Vitali, A. (2007). *Multiple Description Coding - EBU Technical Review*. Geneva, Switzerland: European Broadcasting Union. Retrieved from http://www.ebu.ch/en/technical/trev/trev_312-vitali_streaming.pdf

Wagner, E. D. (2005, June). Enabling Mobile Learning. *EDUCAUSE Review*, *40*(3), 40–52.

Wang, Y. (2004). Context Awareness and Adaptation in Mobile Learning. In *Proceedings of the The 2nd IEEE International Workshop on Wireless and Mobile Technologies in Education, 2004 (WMTE 2004)* (pp. 154-158). IEEE. doi:10.1109/WMTE.2004.1281370

Wang, Z., Sheikh, H. R., & Bovik, A. C. (2003). Objective video quality assessment. In B. Furth & O. Marques (Eds.), *The Handbook of Video Databases: Design and Applications* (pp. 1041–1078). CRC Press.

Retrieved from http://www.ece.uwaterloo.ca/~z70wang/publications/QA_hvd_bookchapter.pdf

Wang, Z., Bovik, A., Sheikh, H., & Simoncelli, E. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transactions on Image Processing*, *13*(4), 600-612. doi:10.1109/TIP.2003.819861

Wang, Z., Lu, L., & Bovik, A. C. (2004). Video Quality Assessment Based on Structural Distortion Measurement. *Signal Processing: Image Communication, Special Issue on "Objective Video Quality Metrics"*, *19*(2), 121–132. doi:10.1016/S0923-5965(03)00076-6

WebAIM. (n.d.). Web Captioning Overview. *WebAIM - Web Accessibility in Mind*. Retrieved July 27, 2010, from http://webaim.org/techniques/captions/

Wei, H., Moldovan, A., & Muntean, C. H. (2009). Sensing Learner Interest through Eye Tracking. In *Proceedings of the 9th Information Technology & Telecommunications Conference (IT&T 2009)* (p. 8). Dublin, Ireland: ARROW@DIT. Retrieved from http://arrow.dit.ie/cgi/viewcontent.cgi?article=1005&context=ittpapnin

White, M. A. (1983). Toward a Psychology of Electronic Learning. In M. A. White (Ed.), *The Future of Electronic Learning* (1st ed., pp. 51-62). Hillsdale, New Jersey, USA: Lawrence Erlbaum Asociates, Inc. Retrieved from http://books.google.ie/books?id=qfG5xgE6aScC

Wiegand, T., Sullivan, G. J., Bjontegaard, G., & Luthra, A. (2003). Overview of the H. 264/AVC video coding standard. *IEEE Transactions on Circuits and Systems for Video Technology*, *13*(7), 560–576. doi:10.1109/TCSVT.2003.815165

Wikipedia. (n.d.). E-learning - Wikipedia, the free encyclopedia. *Wikipedia*. Wiki, . Retrieved from http://en.wikipedia.org/wiki/E-learning

Winkler, S. (2005). Perceptual Video Quality Metrics — A Review. In H. Wu & K. Rao (Eds.), *Digital Video Image Quality and Perceptual Coding* (1st ed., pp. 155-181). CRC Press. Retrieved from http://video.ee.ntu.edu.tw/~sernc/books/Digital%20Video%20Image%20Quality%20and%20Perceptual%20Coding/

Winkler, S., & Mohandas, P. (2008). The Evolution of Video Quality Measurement: From PSNR to Hybrid Metrics. *IEEE Transactions on Broadcasting*, *54*(3), 660-668. doi:10.1109/TBC.2008.2000733

Winkler, S. (2007). Video Quality and Beyond. In *Proceedings of the 15th European Signal Processing Conference (EUSIPCO 2007)*. Poznań, Poland: EURASIP. Retrieved from http://stefan.winkler.net/Publications/eusipco2007.pdf

Wolf, C. (2003). iWeaver: Towards 'Learning Style'-based e-Learning in Computer Science Education. In T. Greening & R. Lister (Eds.), *Proceedings of the 5th Australasian Conference on Computing Education (ACE2003) - Volume 20*, Conferences in Research and Practice in Information Technology Series; Vol. 140 (pp. 273-279). Adelaide, Australia: Australian Computer Society, Inc. Retrieved from http://crpit.com/confpapers/CRPITV20Wolf.pdf

Wu, P., Purushotham, S., & Kuo, C. (2010). Video Genre Inference Based on Camera Capturing Models. In D. Schonfeld, C. Shan, D. Tao, & L. Wang (Eds.), *Video Search and Mining*, Studies in Computational Intelligence (Vol. 287, pp. 177-201). Berlin, Germany: Springer-Verlag. Retrieved from http://dx.doi.org/10.1007/978-3-642-12900-1_7

Yi, C., Liao, P., Huang, C., & Hwang, I. (2009). Acceptance of Mobile Learning: a Respecification and Validation of Information System Success. *World Academy of Science, Engineering and Technology*, *53*, 726-730.

Yin, C., Ogata, H., Tabata, Y., & Yano, Y. (2010). Supporting the Acquisition of Japanese Polite Expressions in Context-Aware Ubiquitous Learning. *International Journal of Mobile Learning and Organisation (IJMLO)*, *4*(2), 214–234. doi:10.1504/IJMLO.2010.032637

Zhao, X., & Okamoto, T. (2008). A Device-Independent System Architecture for Adaptive Mobile Learning. In *Proceedings of the 8th IEEE International Conference on Advanced Learning Technologies 2008 (ICALT '08)* (pp. 23-25). Santander, Spain: IEEE Explore. doi:10.1109/ICALT.2008.21

Zinner, T., Abboud, O., Hohlfeld, O., Hossfeld, T., & Tran-Gia, P. (2010). Towards QoE Management for Scalable Video Streaming. In *Proceedings of the 21th ITC Specialist Seminar on Multimedia Applications - Traffic, Performance and QoE*. Miyazaki, Japan. Retrieved from ftp://ftp.kom.e-technik.tu-darmstadt.de/pub/papers/ZAHT10.pdf