We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

154
Countries delivered to

Our authors are among the

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Understanding User Experience of Mobile Video: Framework, Measurement, and Optimization

Wei Song, Dian Tjondronegoro and Michael Docherty
Queensland University of Technology
Australia

1. Introduction

Since users have become the focus of product/service design in last decade, the term User eXperience (UX) has been frequently used in the field of Human-Computer-Interaction (HCI). Research on UX facilitates a better understanding of the various aspects of the user's interaction with the product or service. Mobile video, as a new and promising service and research field, has attracted great attention. Due to the significance of UX in the success of mobile video (Jordan, 2002), many researchers have centered on this area, examining users' expectations, motivations, requirements, and usage context. As a result, many influencing factors have been explored (Buchinger, Kriglstein, Brandt & Hlavacs, 2011; Buchinger, Kriglstein & Hlavacs, 2009). However, a general framework for specific mobile video service is lacking for structuring such a great number of factors.

To measure user experience of multimedia services such as mobile video, quality of experience (QoE) has recently become a prominent concept. In contrast to the traditionally used concept quality of service (QoS), QoE not only involves objectively measuring the delivered service but also takes into account user's needs and desires when using the service, emphasizing the user's overall acceptability on the service. Many QoE metrics are able to estimate the user perceived quality or acceptability of mobile video, but may be not enough accurate for the overall UX prediction due to the complexity of UX. Only a few frameworks of QoE have addressed more aspects of UX for mobile multimedia applications but need be transformed into practical measures. The challenge of optimizing UX remains adaptations to the resource constrains (e.g., network conditions, mobile device capabilities, and heterogeneous usage contexts) as well as meeting complicated user requirements (e.g., usage purposes and personal preferences).

In this chapter, we investigate the existing important UX frameworks, compare their similarities and discuss some important features that fit in the mobile video service. Based on the previous research, we propose a simple UX framework for mobile video application by mapping a variety of influencing factors of UX upon a typical mobile video delivery system. Each component and its factors are explored with comprehensive literature reviews. The proposed framework may benefit in user-centred design of mobile video through taking a complete consideration of UX influences and in improvement of mobile video

service quality by adjusting the values of certain factors to produce a positive user experience. It may also facilitate relative research in the way of locating important issues to study, clarifying research scopes, and setting up proper study procedures.

We then review a great deal of research on UX measurement, including QoE metrics and QoE frameworks of mobile multimedia. Finally, we discuss how to achieve an optimal quality of user experience by focusing on the issues of various aspects of UX of mobile video. In the conclusion, we suggest some open issues for future study.

2. User experience in mobile video

Though the term user experience (UX) has been frequently used in multimedia services, as of now, there is no common definition for UX. According to a survey on UX (Law, Roto, Hassenzahl, Vermeeren & Kort, 2009), Hassenzahl and Tractinsky's definition is the most preferred by both academics and industry. They define UX as "*a consequence of a user's internal state, the characteristics of the designed system and the context (or the environment) within which the interaction occurs*"(2006, p. 95). A more formal definition for UX is issued in ISO 9240-210 (2010). It states that UX is an individual person's perceptions and responses; is related to usage; and includes consequences from both current use and anticipated use of a product, system or service (Law et al., 2009).

It is a continuous process in understanding what is user experience and/or what are its building blocks (components) (Alben, 1996; Hassenzahl & Tractinsky, 2006; McCarthy & Wright, 2004; Roto, 2006a). To clarify the UX in the particular mobile video service, we firstly get through the overall understanding of UX; then analyze the important features of UX, so that we can identify the essential issues in the mobile video field.

2.1 Comparison of general UX frameworks

It is very hard to distinguish a UX definition from a UX framework, because the definition of UX is usually given in the form of describing various aspects involved in the interaction process of generating UX (Alben, 1996; Hassenzahl & Tractinsky, 2006). The UX framework can be presented as either the building blocks of UX (Hassenzahl & Tractinsky, 2006; Roto, 2006a) or the interaction processing structures of UX (McCarthy & Wright, 2004; Norman, 2004). The interaction process involves the people's senses, behaviors and reflections, which are more abstract and more difficult to measure than the building blocks. To compare the UX frameworks, we transpose the interaction processing frameworks into building blocks based on the relations between the producing process of UX and the involved objects. Table 1 shows the comparison results of a group of seven important UX frameworks or definitions in terms of their related building blocks.

The definition of experience given by Alben (1996) indicates seven attributes of experience in user-product interaction (shown in Table 1). The way to feel a product in one's hands refers to the attributes of overall appearance of the product, the user's first perception of it, and the user's physical resources – hands (even if it is too narrow to just talk about the hands). Understanding how the product works and using it involves the attributes of the product's functionality and usability. How well the product to serve people's purpose and to fit into the entire using context involves users' needs and the usage context.

Components	Attributes	Alben (1996)	Forlizzi & Ford (2000)	Arhippainen & Tähti (2003)	Norman (2004) & Orist et al. (2010)	McCarthy & Wright (2004)	Hassenzahl & Tractinsky (2006)	Roto (2006b)
User	Emotion		√	√	√	√	√	√
	Needs	√		√	√	√	√	√
	Prior experiences		√	√	√	√		√
	Perceptions	√	√		√	√		√
	Expectations			√			√	√
	Motivation			√			√	√
	Profile (age, sex, preference, skill/knowledge)		√	√			√	
	Physical resources	√		√				√
Product/ System/ Service	Product appearance or system complexity	√	√	√	√	√	√	√
	Functionality	√	√	√	√	√	√	√
	Usability	√	√	√	√	√	√	√
	Aesthetic quality		√	√	√		√	
	Interactivity		√	√	√	√		√
Context	Context of use or physical context	√	√	√		√	√	√
	Social context		√	√	√	√	√	√
	Culture context		√	√				
	Temporal and task context						√	√

Table 1. Comparison of UX frameworks

Forlizzi and Ford (2000) deem that experience is influenced by the components of user-product interaction, including the user's emotions, prior experiences, values and cognitive models and product's features, usability, and aesthetic qualities; and the interaction surroundings, such as a context of use and social, culture and organizational behavior patterns. The user's values and cognitive models are relevant to prior experience or knowledge and personality, and the aesthetic quality of the product is associated with the user's pleasure of using the product. Forlizzi and Ford also highlight the interactivity of the product, meaning that cognition dimension of experience enables the product to offer the user a learning experience.

Similarly, Arhippainen and Tähti (2003) also think that user experience forms in the interaction between user and product in a particular context of use and social and cultural environment, but they separate social, culture and context of use into independent components. They list a good amount of attributes for each component; however, some of these attributes were not recognized in their testing with two mobile application prototypes. This indicates that the attributes affecting user experience are variable in different cases.

Regarding the temporal dimension of UX, Donald Norman (2004) states three levels (visceral, behavioral and reflective level) of interaction. At the visceral level, people have the first impression (i.e., perception) of a product through its appearance and their feelings, e.g., like or dislike, occur spontaneously. At the behavioral level, when people start to use a

product, their experience is about how well the product's functions fulfill their needs, and how easily the product can be used. Therefore, this level involves product's functions, usability, and user needs. At the reflective level consciousness takes part in the process; whereby people understand and interpret things, and remember past experiences and may use their current experiences for future actions. The reflection level is relative to the product's interactivity and aesthetic quality, and may also engage the user's prior experience and social context when it affects the user's understandings of the product and its usage for social purposes. Recently, Norman's structure is extended by increasing a pre-experience level prior to the visceral level to indicate people's pre-experiences with similar product/services (Obrist et al., 2010). Prior experience is more important at this level.

Wright and McCarthy's framework (2004) analyses experience with technology, which has four intertwined threads of experience and six sense-making processes. The four threads: "sensual, emotional, spatio-temporal and compositional" represent the visceral character of experience, value judgment ascribed to emotions, place and time effects, and coherent experience, respectively. The sense-making processes are anticipating (expectation associated with prior experience), connecting (immediate and pre-conceptual sense), interpreting (working out what is going on), reflecting (evaluation in the interaction and reflection with feelings), appropriating (making an experience one's own) and recounting (storytelling with others or oneself about the experience). Compared to Norman's framework, this framework emphasizes the effects of physical context and the connections of previous sense with the product.

Hassenzahl and Tractinsky's definition (2006, p. 95) clearly lists attributes for each UX component. The user's internal state includes predispositions, expectations, needs, motivation, mood, etc.; the system has the characteristics of complexity, purpose, usability and functionality; and the context involves physical environment, organisational/social setting, and task context (e.g., meaningfulness of the activity or voluntariness of use). Roto (2006a) has followed the definition and developed UX building blocks which consists of three main components: user, context, and system. Here, "System" is suggested to replace "product" in order to include all involved infrastructures (such as products, objects, and services) in the interaction. Furthermore, based on her study on UX for mobile web browsing, she divides contextual attributes into four categories: physical, social, temporal and task contexts. The physical context refers to physically sensed circumstances and geographical location; social context refers to other people's influence on the user and the user's social contribution goals; temporal context refers to the time available for task execution; and task context refers to the role of the current usage task (which is mobile browsing in her case) related to other tasks (Roto, 2006b).

Concluding the similarity of the above UX frameworks, we can distribute their attributes into three components: user, product/system/service, and context, shown in Table 1. In each component, it can be observed that some attributes are highlighted, such as user's emotions, perceptions and needs, functionality and usability of product or system, and context of use. However, there are a couple of attributes (indicated in blue color in Table 1) are either ambiguous or less mentioned.

Firstly, temporal context and task context are only specified by Roto (2006a). Secondly, while people's visceral or sensual experience has been addressed (McCarthy & Wright, 2004; Norman, 2004), the relevant physical resources and characteristics are not mentioned. In many

situations, these should be considered as important. For instance, Roto mentioned that in the mobile context the user may only have one hand for the device (2006a). Also, characteristics of human eyes and ears can affect the user's perception on videos and audios. Thirdly, user's motivations and expectations are also seldom mentioned. A user may be motivated to use a product/service by his/her expectation to achieve a goal, current need, social influences, or physical context limitations; whereas, motivations can not cover user's expectations and needs. The motivation refers to why a user uses an object (i.e., product/service/system); the expectation refers to what the user expects to gain from using the object; and the need refers to how well the requirements are fulfilled by using the object. Fourthly, user profile may contribute to a more personalized product/service. People at different ages or with different genders and preferences often experience the same thing in distinct ways. Compared to prior experience that refers to the previous experience of using a similar product/service, the user's knowledge or skill background covers more wide areas that indirectly associate with the current usage. For example, a person who has a computer science background usually has a deeper understanding to a brand new digital device than others without the background.

It can also be noticed, in Table 1, that when a specific domain is concerned, more detailed attributes are provided. For instance, using the case of mobile web browsing as the example, the temporal context and the task context are proposed (Roto, 2006a); while in the case of evaluating UX with adaptive mobile application prototypes, the user's personal characteristics (e.g., motivations, personalities, prior experience) are obvious (Arhippainen, 2003). These situations indicate that it is necessary to get a deeper insight into all aspects of UX in order to achieve a good user experience of mobile video applications.

2.2 UX Framework for mobile video

User experience of mobile video is generated when users manipulate it by selecting video content to watch, perceiving service and video quality and evaluating them. There are a large number of factors affecting UX of mobile video. Many players on the technology side directly associate with video coding, network transmission, and device and system performance. On the non-technology side, the users' characteristics, service provisioning modes and use contexts are diverse.

Based on the previous research, an overall UX framework for mobile video emerges by allocating all kinds of the influencing factors to a typical mobile video delivery framework, as shown in Figure 1. This structure summarizes and simplifies previous work (Buchinger et al., 2011; Buchinger, Kriglstein, et al., 2009; Jumisko-Pyykkö & Häkkinen, 2005; Knoche, McCarthy & Sasse, 2005; Orgad, 2006), where a huge number of factors influencing UX of mobile video are not well organized; and it also extends previous frameworks (Forlizzi & Ford, 2000; Hassenzahl & Tractinsky, 2006; McCarthy & Wright, 2004; Norman, 2004) to the specific domain of mobile video. In accordance with the generally accepted UX components (shown in Table 1), the proposed structure organizes the influencing factors of UX into three components: USER, SYSTEM and CONTEXT, and maps their impacts upon four elements of the mobile video delivery framework, namely mobile user, mobile device, mobile network, and mobile video service.

The following sections will introduce the factors of each component and the relevant research, some of which provide better understanding of UX of mobile video and others make progress in optimizing UX by utilizing the impacts of the factors on UX.
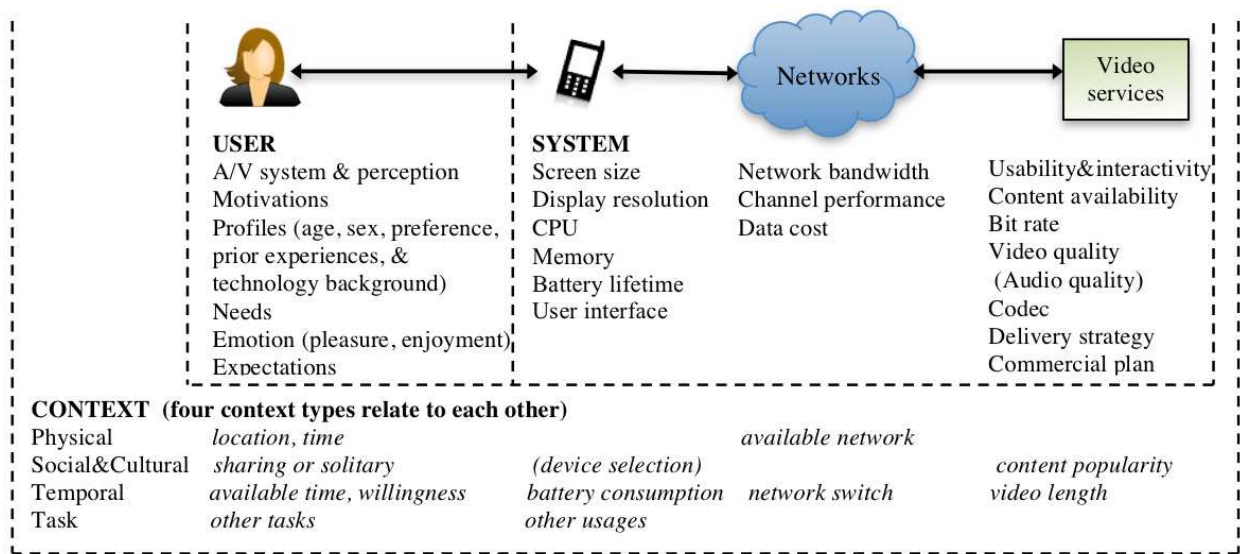
Fig. 1. User experience framework of mobile vide

2.2.1 User

For the mobile user, the factors are human audio-visual system and perception, motivations, user profiles, needs, expectations, and emotions. Mobile video is mainly a visual product, and user's perception of video quality is firstly the result of Human Visual System (HVS) perceiving the video. As a result, the human eyes' features, as physical characteristics, can be utilized to improve user's visual perception. For example, in a resource limited condition (e.g., limited network bandwidth), video coding based on Region-of-Interest (ROI) can increase user perceived video quality by maintaining or enhancing the quality of ROIs, which are detected salient areas in terms of the human eyes' selective sensitivity and visual attention (Buchinger, Nezveda, Robitza, Hummelbrunner & Hlavacs, 2009; Engelke & Zepernick, 2009; Lu et al., 2005). Human auditory system helps the visual system work well, particularly in a situation that the user can not concentrate on the screen of mobile device, e.g., walking, or in a case that the user is viewing a sound-important content such as news and music videos (Jumisko-Pyykkö, Ilvonen & Väänänen-Vainio-Mattila, 2005; Song, Tjondronegoro & Docherty, 2011).

The user profiles consist of several aspects: age, sex, preference for video content type, prior experiences in viewing videos and mobile videos, and technology background (especially in information and computer technology). Although a lot of research has observed the behavior differences of using mobile video (TV) between groups classified by age, gender and technology (Eronen, 2001; Jumisko-Pyykkö, Weitzel & Strohmeier, 2008; Orgad, 2006; Södergård, 2003), the comprehensions in how the differences influence UX is inadequate. For example, are young people (males) easier to satisfy in terms of quality of mobile video service than older people (females)? How does prior experience in viewing videos impact upon current viewing? A few studies have addressed the positive correlation between user's preference (also called interest) for video content and overall user experience (Jumisko-Pyykkö et al., 2005; Song, Tjondronegoro, Wang & Docherty, 2010). Recent studies have

found that people's desired quality of mobile video varies with their preferences for video content, viewing experiences of mobile videos, technical backgrounds, and even their genders. There may also be an interactive impact across these aspects of user profiles (Song, Tjondronegoro & Docherty, 2010; Song et al., 2011). For instance, frequent male viewers of mobile video may request a higher quality than occasional viewers. (Song et al., 2011).

Buchinger, Kriglstein and Hlavacs (2009) have summarized a dozen motivations of watching mobile TV. Simplifying those, the major motivations of viewing mobile videos are: consuming time, being entertained, staying up to date (e.g., with news or popular events), sharing with others or isolating oneself from the surrounding.

These user factors do not only work independently. It is very likely that user profiles and motivations are closely bound up with user needs. When mobile video viewing is for killing time on a bus, people may need short videos with fair quality, while when for an entertainment use at home, they might need a good quality video. Expectations have been found to relate to previous experience. E.g., people who often watch high quality video expect a higher quality of mobile videos than those who do not (Song et al., 2011).

Another factor - emotion has been noticed in many UX frameworks. Hassenzahl and Tractinsky (2006) summarized two ways of dealing with emotions in UX: stressing the importance of emotions as consequences of using a product, and using emotions as important evaluative judgments. For example, satisfaction and entertainment were investigated as emotional consequences of task-directed mobile video use (Jumisko-Pyykkö & Hannuksela, 2008), and pleasantness has been found to create affective responses on judgments (e.g., willing to watch in long-term) (Song, Tjondronegoro & Docherty, 2010). Emotions sometimes also mean user's internal state of feelings and moods (e.g., love, sad, happy). However, this kind of personal emotion is secret and its effect on UX has hardly been reported in the mobile video interaction. Therefore, the emotion, in this proposed framework, refers to user's viewing mood, that is, the enjoyment (or pleasantness) of viewing. Song et al. studies (2010; 2011) have shown that the enjoyable or pleasing emotion is not only an important index of positive UX but also a determining factor of user needs for video quality, where users tend to request a much higher quality when their criteria are based on the pleasantness.

2.2.2 System

The component "SYSTEM" is related to the overall performance of the infrastructure of mobile video delivery, and therefore covers three objects from the sender to receiver: video services, networks and mobile devices. For a mobile device, a bigger screen is preferred but reduces its portability (Knoche & McCarthy, 2004; Knoche & Sasse, 2008). A screen with high display resolution can support high quality video playing but cause big consumptions of CPU resource, buffering memory and battery life, which may negatively affect user's usage behavior (Chipchase, Yanqing & Jung, 2006; Kaasinen, Kulju, Kivinen & Oksman, 2009; Knoche & Sasse, 2008). Apart from these factors, user interface of a media player is also an important influence. A good user interface comes from good design of the media player (e.g., interactivity, flexibility and easy to use), but also from effectively utilizing some advance functionalities of the mobile device, e.g., touch screen and gesture recognition (Huber, Steimle & Mühlhäuser, 2010; MacLean, 2008).

Factors in networks are mainly bandwidth, channel features such as jitter, delay and packet loss, and data cost. Narrow bandwidth and poor channel performance will result in a negative UX due to the distortions of video quality caused by the transmission (Bradeanu, Munteanu, Rincu & Geanta, 2006; Ketyko, De Moor, Joseph, Martens & De Marez, 2010; Tasaka, Yoshimi & Hirashima, 2008). The data cost means not only the spent money on using the network, but also how much of a total available data amount has been used. For example, if a user has a free network or he/she has paid for a huge amount of data flow, the user may watch videos quite often and would like to watch high quality videos. In another scenario, when a user knows the data flow is limited (or shared with other people), even though the network is free, the user may be concerned with the data consumption and not use too much. Therefore, user's affordable cost (money or data amount) for video data consumption affects their watching behaviors.

On the video service side, usability and interactivity are two important factors because they are directly associated with the customer's use. Even if the usability and interactivity are reflected in the user interface of a mobile video player, such as content navigation (Buchinger, Kriglstein, et al., 2009), search (Hussain et al., 2008), and easy to play (Carlsson & Walden, 2007), they must be underpinned by the functionality of the video service. The term "functionality" is too narrow to express the connection between video service and the user. Also it is overlapped by the usability and interactivity in mobile video service. Therefore, we choose the term "usability and interactivity" to represent the influence of the service function on UX. Its importance can be shown in at least two aspects. On the one hand, the information for content navigation and searching must be provided by the video service; on the other hand, the user's interaction requirements, e.g., content selection, quality selection, and rating, must be responded to by the service.

Another factor, content availability refers to what and how much video content the video service can provide to users. Abundant and interesting content can meet more users' requirements (Song & Tjondronegoro, 2010). Bit rate of a video affects the user's data cost and the user's perceived video quality. Given a bit rate constraint, the video can be encoded with different parameters by different video coding codecs; and the variations eventually lead to divergent user-perceived video qualities (Ahmad, 2006b; Cranley, Murphy & Perry, 2004; Kun, Richard & Shih-Ping, 2001; Song, Tjondronegoro & Azad, 2010). Audio quality, including the volume, sampling rate, bit rate of the audio, often takes effects with the usage ambient (e.g., noisy or quiet) and the content type (e.g., music videos) together (Jumisko-Pyykkö, Häkkinen & Nyman, 2007). Delivery strategy is about how the video service is delivered to the user. Under different delivery strategies, a user may watch a video in real-time and can access to an arbitrary time point; the user may have to watch a video after it is fully downloaded into the terminal device; or the user may wait for a shorter or longer buffering time before watch. Commercial plan refers to the providing manner of a video service, such as subscription, online free, or pay for individual video. It is suggested that for the success of mobile TV, the right pricing approach should be to give users a choice of various payment options anyway (Trefzger, 2005).

2.2.3 Context

Based on the study on UX of mobile web browsing, Roto (2006a) has classified context into four types: physical, social, temporal and task context. Due to the similarity of the mobile

context, we also classify CONTEXT into the four types, but replace the social context with social & cultural context. We relate the roles of the four contexts to the four elements of mobile video delivery system (i.e., users, mobile devices, networks, and video services) based on how their impacts are reflected through these elements.

First of all, the physical context is about where and when a user is using the mobile video. Except from light and noise that will have a direct impact on the user's watching and listening, in mobile environment, changes of the physical context often lead to changes of available networks or network conditions, which may cause a significant variation of UX. For example, shifting from a high-speed Wi-Fi network at home to a low-speed 3G network outside, a user may be unhappy with a longer waiting time to load a video. In addition, during network traffic time, one may have difficulty to watch smooth videos. Secondly, the social context refers to how a user is influenced by others and whether the user joins the influences to others. Its impacts are presented in sharing or solitary use of mobile video, selections of video content, and voting popularity. When solitary viewing or video sharing happens, people are using mobile video to manage relationships with others in shared or public settings. They are trying to either cut off the outside setting or enjoy others' attendance (O'Hara, et al., 2007). In addition, social recommendations highly influence what people watch and how they feel; sometimes, also influence people's options for mobile devices and mobile communication companies. The influences of the culture context itself are not explicit, but contribute to users' viewing habits such as preferred video content and viewing situations (Song & Tjondronegoro, 2010). For example, the study in Belgium (Vangenck, Jacobs, Lievens, Vanhengel & Pierson, 2008) found people tended to use mobile TV at home, while the study in Japan (Miyauchi, Sugahara & Oda, 2008) stated that the main consumption of mobile TV was 'on the go'. In Australia, music video is the most popular content type for mobile video (Song & Tjondronegoro, 2010), which conflicts with the result of "news" in other countries' studies (Chipchase et al., 2006; Mäki, 2005; Södergård, 2003). Since it is hard to draw a clear line between the impacts of social and those of culture, it is better to put them together. Thirdly, the temporal context refers to that given the context restrictions how long the dedicated viewing process will last (i.e., the period that a user immerses into the viewing). The restrictions can be the user's available time (e.g., 5 minutes waiting for a bus), and the user's willingness to watch for a long or short time. Also, the user sometimes has to stop viewing due to a low battery warning; or the user's viewing process can be paused by network switches. The viewing period is also restricted by the duration of a video as well. If the available video is only 2 minutes long, the dedicated viewing will not last over 2 minutes. Fourthly, user's viewing task often runs parallel to other tasks or it is motivated by a higher-level task. For instance, a user's viewing with friends has a higher-level purpose of sharing experience and a parallel task of spending time with friends. While, when the user watches videos on a bus, the higher-level task is to kill time and the parallel task is to take the bus. User's viewing can also be interrupted by other usages of mobile device such as a coming call. A study has found that interrupted viewing such as viewing on a bus will result in a relative lower user perception of a good quality video than relaxed viewing (Song, Tjondronegoro & Docherty, 2010).

In spite of being separated, there are correlations between the four context types. For example, the video sharing behavior often happens in a physical crowd context with a specific task context; different cultures determine the most frequent viewing locations and times (Buchinger, Kriglstein, et al., 2009); a short-time and interrupted viewing often takes

place on a bus, accompanying with a higher-level task of taking the bus to the destination (Knoche & McCarthy, 2004).

In the above, we have proposed a UX framework for mobile video and explained each factor in it. It may bring an overall idea of how the UX of mobile video is influenced. Understanding the UX serves a higher-level goal that is to find out a way to optimize the UX under a series of resource constraints of mobile context. Prior to achieve this purpose, there is a central question need to be answered - how to measure the UX? Without measurement of the UX, we are not able to evaluate the holistic system performance in satisfying users and meeting their needs.

3. Measuring Quality of Experience

The term Quality of Experience (QoE), sometimes also known as quality of user experience, has been frequently used to represent the measurement of user experience with a service, especially in web browsing, communication, and TV/video delivery. QoE came after another well-established concept Quality of Service (QoS). QoS is a measure of technological performance, such as network capacity (e.g., throughput, error rate, latency, etc.) and device capabilities and product features (e.g., battery lifetime, video bitrate, frame rate, etc.), but does not deal with user's overall experience. QoE therefore is proposed to involve human dimensions into the measurement of multimedia service performance, together with the objective technical aspects together.

In ITU-T Recommendation of QoE for IPTV service (2007), QoE is defined as overall acceptability of a service/application perceived by a end user; it is influenced by various effects of system (device, network, services infrastructure, etc.), user needs and expectations, and usage context. Wu et al. (2009) proposed a refined definition for QoE based on the study in Distributed Interactive Multimedia Environments (DIME). They defined QoE as "*a multi-dimensional construct of perceptions and behaviors of a user, which represents his/her emotional, cognitive, and behavioral responses, both subjective and objective, while using a system*". Both the definitions indicate a close relationship between QoE and UX as well as the way to measure QoE. That is, QoE can be evaluated based on the end-users' responses, and it should reflect multi-dimensional effects.

To measure QoE, a great number of QoE metrics for perceived video quality have been developed and used for quality management in mobile video service. However, these metrics are limited in taking into consideration only some influencing factors of user experience. From the overall perspective, a few comprehensive QoE frameworks have been proposed, but it is still extremely challenging to apply these frameworks into a practical use.

3.1 QoE metrics

In terms of the QoE definitions (ITU-T Study Group 12, 2007; Wu et al., 2009), it accentuates how the end-user accepts and perceives the received quality of mobile video. Subjective tests are commonly used to evaluate the perceived video quality. In the tests, the subjects are asked to rate the quality of the presented video sequences that are impaired by controlled conditions, such as (simulated) network and device conditions. The subjective quality assessment is regarded as the most reliable way to assess video quality and the most fundamental methodology for evaluating QoE (Tominaga, hayashi, Okamoto & Takahashi, 2010).

The commonly used subjective testing methodologies are proposed by the ITU-T and ITU-R, including the Absolute Category Rating (ACR), the Degraded Category Rating (DCR) (also called DSIS), the Single Stimulus Continuous Quality Evaluation (SSCQE) and the Double-Stimulus Continuous Quality Scale (DSCQS) (ITU-T P.910 Recommendation,ITU-R Recommendation BT. 500-11:, 2004; 1999). The average ratings obtained from the above assessment methods are called the Mean Opinion Score (MOS), which is in a form of 5/11point scales. A study on performance comparison of these methods for mobile video applications (Tominaga et al., 2010) demonstrates that the ACR and DSIS (or DCR) methods with 5 scales perform better than the others.

Notwithstanding that the scaled assessments are widely used, they are subject to overburden participants, who especially struggle to determine a proper score for the quality of a video (Sasse & Knoche, 2006). Furthermore, they can not sufficiently answer the question: which quality level is acceptable to end users (Schatz, Egger & Platzer, 2011). Binary measure is therefore suggested to use when assessing the acceptability of mobile TV (videos) (Agboma & liotta, 2007; 2008; Knoche et al., 2005; McCarthy, Sasse & Miras, 2004). The idea of acceptability is to identify the lowest acceptable quality level or threshold. A psychological method used to determine threshold is known as the Method of Limits created by Gustav Theodor Fechner (cited in Agboma & liotta, 2007). It is often done through asking participants to simply decide whether or not they accept the quality of a displaying video in successive, discrete steps either in ascending or descending series.

As regard to the relation between the acceptability and the MOS, a little research has been done. One study has proposed a set of mapping formula from MOS scores to acceptability values (de Koning, Veldhoven, Knoche & Kooij, 2007). However, another study did not find a reliable mapping relationship (Jumisko-Pyykkö, Vadakital, et al., 2008). A recent study took this issue into the field of mobile broadband data services and conducted a series of lab and field experiments. It turned out that a consistent mapping between the binary acceptance and the ordinal MOS ratings exists across different applications, such as web browsing and file downloads (Schatz et al., 2011).

Since subjective quality assessment is inconvenient, time-consuming and expensive, objective video quality metrics are then developed to predict the perceived video quality automatically. The objective video quality metrics are commonly considered as the computing models of QoE or objective QoE (oQoE) in (Zinner, Hohlfeld, Abboud & Hossfeld, 2010). The performance of objective QoE metric can be evaluated by comparing the prediction results with the scores obtained from the subjective quality assessments.

According to the availability of the original video sequence, the objective video quality metrics can be classified into full-reference (FR), blind or no-reference (NR) and reduced-reference (RR) metrics (Wang, Sheikh & Bovik, 2004). The FR metric needs a distortion-free reference video and performs the quality assessment by comparing the distortion video with the reference. The NR metric assesses the quality of a distorted video without any reference and assumes the video distortions, e.g., blur and blockiness. The RR metric evaluates a test video based on a series of features extracted previously from the reference videos.

The most widely used FR metrics are mean squared error (MSE) and peak signal-to-noise ratio (PSNR). However, PSNR or MSE is thought unable to represent the exact perceptual quality because it is based on pixel-to-pixel difference calculations, thereby neglecting the

effects of viewing conditions and characteristics of the HVS (Masry & Hemami, 2002; Zhenghua & Wu, 2000). To date, many more effective metrics have been developed, such as structural similarity (SSIM) (Wang, Bovik, Sheikh & Simoncelli, 2004), multiscale SSIM (MS-SSIM) (Wang, Simoncelli & Bovik, 2003), video quality metric (VQM) (Pinson & Wolf, 2004), visual information fidelity (VIF) (Sheikh & Bovik, 2006) and motion-based video integrity evaluation (MOVIE) (Seshadrinathan & Bovik, 2010). The performances of these objective video quality metrics has been evaluated by Seshadrinathan et al. (Seshadrinathan, Soundararajan, Bovik & Cormack, 2010) and Chikkerur et al. (Chikkerur, Sundaram, Reisslein & Karam, 2011). The results show that the MS-SSIM, the VQM and the MOVIE metrics outperform than other metrics. However, these metrics seem not to work well for videos playing on mobile devices. According to Eichhorn and Ni (2009), SSIM and VQM perform bad in estimating the scalable video quality on mobile screens. FR metrics are hardly used in many practical video services where the reference video sequences are often inaccessible.

No-reference (NR) metrics estimate QoE though mainly measuring image distortions: blockiness (Leontaris & Reibman, 2005; Saad, Bovik & Charrier, 2010; Zhou, Bovik & Evan, 2000), blur (Marziliano, Dufaux, Winkler & Ebrahimi, 2002; Sadaka, Karam, Ferzli & Abousleman, 2008; Yun-Chung, Jung-Ming, Bailey, Sei-Wang & Shyang-Lih, 2004), and noise (Ghazal, Amer & Ghrayeb, 2007). An overview of existing NR image and video quality estimation studies have been given by Hemai and Reibman (Hemami & Reibman, 2010). These artifactual effects are mostly generated during the process of encoding, decoding and transmission. For example, the blockiness is caused by a block-based video coding such as MPEG-4 and H.264/AVC codec; the blur can be resulted from the spatial scaling and decoding; and the noise may be added due to transmission errors.

Reduced-Reference (RR) metrics are usually developed based on the technical influencing factors of perceptual video quality, such as video coding parameters, video content features and network transmission parameters, which can be known in advance or detected. Therefore, RR metrics have been used in practical QoE predictions or QoE managements. The RR metrics can be further divided into two classes: encoding-parameter-based class and network-parameter-based class.

A well-known encoding-parameter-based model has been given in Recommendation ITU-T G.1070 (2007). In this model, the computing coefficients are determined by codec type, video display format, key frame interval and video display size. Based on this model, a better parametric model is developed, which is able to estimate perceptual MOS values for different codecs (MPEG-4 and H.264/AVC), bitrates and display formats, and video content (distinguished by movement intensity) (Joskowicz & Ardao, 2010). To estimate video quality in mobile video streaming scenarios, two reference-free models have been provided by Ries, Nemethova and Rupp (2008). The first method estimates the video quality using average bitrate and four motion characteristics of the video, while the second model is a content-dependent and low-complexity metric with two objective parameters bitrate and frame rate. However, in the second model, the parameters' coefficients vary with the applied content types such as news, soccer, cartoon, panorama, and the rest, therefore, content classification needs to be performed before using the model.

Many implemented QoE models have considered the important effect of network transmission, which quality can be estimated by QoS measurement. Fielder, Hossfeld and

Tran-Gia (2010) have found that there is a generic exponential relationship between user-perceived QoE and network-caused QoS. Other effects such as video content types and video coding parameters have also been considered together with the network effect. For example, Tasaka, et al. (2008) estimated QoE from the measured application–level QoS. The generated QoE metrics are for three content types: sports, animation, and music, and in the form of nonlinear equations with the indicators of error concealment ratio and MU (which refers to the information unit for transfer between the application layers) loss ratio. Whereas, Bradeanu et al. (2006) used both video coding profiles (based on the encoding bitrate) and network conditions such as transmission error and buffering occurrence to model QoE. While most network-focused QoE metrics were developed under simulated network environment, Ketyko et al. (2010) have focused on measuring the QoE of mobile video streaming under an actual 3G network and real usage context. They conducted subjective assessments under six different usage contexts, including indoor and outdoor at home, at work and on train/bus. Based on the collected data, they modeled a general QoE as a linear function of video packet loss rate, video packet jitter, audio packet jitter, and RSSI (received signal strength indication). This study also found that spatial quality (which is formed by the content, the sound quality, the fit to feeling, and the picture quality) and emotional satisfaction were the most related aspects of the general QoE.

The above QoE metrics are all built using Mean Opinion Score (MOS) as the index. In terms of Schatz et al.'s study (2011), acceptability is a relevant and useful concept for QoE assessment. Agboma and Liotta (2008) have proposed a QoE management methodology with the purpose of maximizing QoE under a constraint network, where binary QoE were employed to predict if a video quality could be acceptable by users. The QoE models were built using statistical discriminant analysis with two parameters video bitrate and frame rate for three different terminals: mobile phone, PDA and laptop. And six content types: news, sports, animation, music, comedy and movie were included in their studies (Agboma & Liotta, 2010). Likewise, another study also focused on acceptable QoE model, but used Machine Learning (ML) classification algorithms to produce more accurate and adaptive QoE predictions, where the spatial and temporal complexity of video content joined the prediction (Menkovski, Oredope, Liotta & Sánchez, 2009).

To sum up, most existing QoE metrics mainly focus on the impacts of network conditions and video encoding on user experience without sufficiently considering other aspects, such as user's personal needs, mobile devices, and context. More comprehensive understandings of QoE are presented in some QoE frameworks.

3.2 QoE frameworks in mobile multimedia

There are a few QoE frameworks in mobile multimedia, which often involve Quality of Service (QoS) into the construction due to the significance of QoS in reflecting the object aspects of multimedia quality.

A taxonomy of QoS and QoE aspects in multimodal human computer interaction have been proposed by Moller et al. (2009). It consists of three layers: 1) QoS influencing factors, which include the characteristics of user, system and context of use, exerting a impact on perceived quality, 2) QoS interaction performance aspects, describing the user and system performance and behavior, and 3) QoE aspects, relating to users' quality perception and

judgment. Extending Moeller et al.'s work, Geerts et al. (2010) have presented a QoE framework based on multidisciplinary research to give more detailed insight into different user aspects. The framework combines the technical aspects and the user aspects with use process (i.e., interaction process with a product on a regular basis), and divides context into three levels: the top level is the socio-cultural context, followed by situational context and interactional context. The authors of the above frameworks have also suggested some metrics and methods to measure the aspects, e.g., using psycho-physiological measurement tools and questionnaires to measure user's motions and system usability. However, current measurements focus on either technical aspects or human's emotions and values aspects, but are unable to provide an integrate consideration.

Nokia has suggested two approaches to measure QoE of mobile services: service level approach and network management system approach (Nokia, 2004). Service level approach relies on statistical measurement using a group of QoE KPIs (Key Performance Indicators). Network management system approach maps network QoS performance, which are measured by QoS KPIs, onto user perceptible QoE performance. And the best option would be use both approaches in a complementary way.

Also based on the relation of QoE and QoS parameters, a specific framework for distributed interactive multimedia environments (DIMEs), given by Wu et al. (2009), includes both cognitive perceptions and behavioral consequences of users. The cognitive perceptions consist of flow (enjoyment and concentration), telepresence and technology acceptance (perceived usefulness and perceived ease of use); and the corresponding behavioral consequences contain performance grains, exploratory behaviors, and technology adoption.

A conceptual framework of QoE, proposed by De Moor et al. (2010), attempts to provide multidimensional evaluation of QoE in a mobile, real-life environment, so as to bridge the gap between technical parameters and human experience factors. In the framework, the core component mobile agent, installing on the end-user device, is able to monitor QoS, context and experience by detecting technical parameters (e.g., terminal screen size, battery level, network conditions, content quality, etc.) and contextual entities (e.g., location, mobility, sensors, and other running applications), and gathering feedbacks with questionnaires and other forms. The authors have conducted experiments to evaluate QoE of mobile video streaming in situ. However, so far, they have not offered a conclusion how to use those detected entries evaluate the overall QoE.

The abovementioned frameworks are comprehensive and theoretical; however, more challenging issues, e.g., the mapping relations between various aspects and the real user experience, have to be further considered in real mobile applications. Although these frameworks have been presented in different ways, they all indicate that QoE measurement is a synthetic result of a set of involved factors. Therefore, to improve QoE, in other words, to optimize UX, the effective way is to appropriately deal with all sorts of UX factors.

4. Optimizing user experience of mobile video

For mobile video service, there are many constraints that make user experience optimization become a tough task, such as unstable wireless networks, heterogeneous mobile devices, diverse usage contexts, and complex user needs and preferences. The way to handle the huge variation of resource limitations and at the same time to meet diverse user

requirements is generally called adaptation (Chang & Vetro, 2005). In line with the UX framework of mobile video (Figure 1), there are many different adaption strategies that work on different UX factors to optimize UX. We classify them into three categories based on their main focuses: video coding adaptations, video transmission adaptations, and other adaptations. These adaptation strategies can be combined to achieve a better performance.

The video coding adaptations aim to reduce video encoding bitrate at a minimum cost of user-perceived video quality. They directly act on the factors on the side of video services, e.g., objective video quality, bitrate and codec efficiency, video content features, and target on the user factor - visual perception. This kind of adaptation can also benefit other factors of "SYSTEM". For example, low encoding bitrate can easily adapt to the bandwidth limitation and reduce CPU, memory and battery consumptions of the mobile device. There are many relevant studies in this area.

Using a video codec with well performance means efficient video compression and sufficient video quality guarantee. Compared to the video coding standards of MPEG-2, MPEG-4 Simple Profile, H.261 and H.263, H.264/MPEG-4 Part 10 AVC (Wiegand, Sullivan, Bjontegaard & Luthra, 2003) has a higher compression ratio and has gained a landslide victory in subjective perceived quality evaluation (Winkler & Faller, 2005). The performance of H.264/AVC codecs has been improving as new algorithms are developed continually (Guixu, Shibao & Jianling, 2008; Kim & Suh, 2009; Kim, Heo & Ho, 2010). In addition, two extended H.264/AVC-based video coding schemes - H.264/SVC (Scalable Video Coding) and H.264/MDC (Multiple Description Coding) - are becoming available for mobile use (Reguant, Prats, de Pozuelo, Margalef & Ubiergo, 2008), which may bring more benefits in resource saving and good video presentation.

Content-based (content-aware) video adaptation schemes utilize the dependency of user-perceived quality on video content to reduce unnecessary bitrate allocation to the contents, for which a high bitrate does not contribute to a high perceptual video quality. For example, in Chang et al.'s adaptive streaming system (2005), based on the result of semantic video analysis, important segments are displayed as high quality, and non-important segments are replaced by static key frames plus audio or textual captions. Claypool and Tripathi (2004; 2002) have designed a content-aware scaling mechanism based on the subjective phenomena that the temporal continuity is more important for the fast-motion shot, while the image quality is more important for the low-motion shot when reducing bandwidth. This content aware system could improve the perceptual quality of video by as much as 50%. By means of characteristics of content spatial and temporal complexity, Cranley, Perry and Murphy (2006) have developed an optimal adaptation trajectory (OAT) to maximize user-perceived quality by dynamically adjusting encoding configurations.

Considering the impacts of the features of human visual system (HVS) and visual attention, ROI(Region-of-Interest)-based video adaptation can bring better perceptual quality and user experience in a restricted resource condition of mobile video, such as low bitrate and small display size. In Engelke and Zepernick's study (2009), the result from image quality assessment showed that subjects were more sensitive to the image quality distortion in the ROI area than that in the background. According to Muntean et al. (2008), by blurring the background and decreasing its saturation but increasing that of the foreground, better visibility could be achieved on mobile devices; and marginal background changes could reduce the video bitrate. Other studies (Knoche, Papaleo, Sasse & Vanelli-Coralli, 2007;

Song, Tjondronegoro, Wang, et al., 2010) revealed that properly zooming the ROI or enhancing the quality of ROI based on the shot type of video content could efficiently improve the overall UX on a mobile device in a low bitrate condition. The above studies indicate two ways to utilize ROI for the adaptation. One is to allocate more of the limited encoding bits to the important areas than the background (Sun, Ahmad & Zhang, 2006). Many different bit allocation schemes for ROI-based video coding have been presented in (Ahmad, 2006a; Ahmad & Lee, 2008; Liu, Li & Soh, 2008; Lu et al., 2005; Shi, Yue & Yin, 2008; Sun et al., 2006; Yang, Zhang, Ma & Zhao, 2009). The other way is to crop and display the ROI area only on the small screen of a mobile device in order to zoom in the small objects, which has been examined for sports videos (Knoche et al., 2007; Seo & Kim, 2006; Shenghong, Yufu & Shenglong, 2010).

The video transmission adaptations mainly focus on the network transmission factors, which are especially important for mobile networks. A numerous of approaches exist, such as channel adaptive video streaming for reducing latency and package error (Girod, Kalman, Liang & Zhang, 2002), rate-control scheme for achieving good temporal-spatial quality trade-offs (2003), and dynamic congestion scheduling through multi-hop networks (Zhang, 2005). Network-focused approaches can perform better in maximizing the user-perceived quality when involving other factors related to the video content and the video coding parameters. For example, knowing that frame rate is more important for action movie whilst bitrate is for news video in terms of the end-user's perception, an acceptable QoE-aware strategy (Agboma & Liotta, 2008) manages the video quality degradation through decreasing bitrate for action movie and frame rate for news to achieve efficient network utilization. Another QoE-driven scheme (Khan, Sun, Jammeh & Ifeachor, 2010) adapts the video sender bitrate and frame rate based on the video content type and the network package loss ratio. In addition, ROI-based coding can also be combined with network transmission control to improve user-perceived quality. Jerbi, Wang and Shirani (2005) have adopted a non-linear transformation to duplicated ROI macroblocks before encoding, and an inverse non-linear transformation after decoding, in which way a high transmission robustness can be achieved in an error-prone channel. Another approach is to encode the ROI region separately from the background and transmit it with a higher priority to ensure the most important information will not be lost via the error-prone channel (Chen, Song, Yang & Zhang, 2007; Lambert et al., 2006).

Moreover, the video transmission adaptations also depend on the delivery strategy adopted by the video server. Principally, efficient transmission protocols are needed to deliver the big amount of video data. Real-time transport/streaming protocols such as RTP and RTSP used to be the common protocols for video streaming, which allows a client to remotely control (e.g., play and pause) a streaming media server and access to the files on the server. Nowadays, adaptive bitrate streaming techniques based on HTTP (Hypertext Transfer Protocol) become more popular to support large distributed HTTP networks such as the Internet. It requires that multiple bitrate-based qualities of a single video source are encoded, segmented and stored in the server; and a player can dynamically and seamlessly switch between these qualities to match the available resources such as network bandwidth and device capability. Apple Inc. has employed HTTP Live Streaming technology for streaming media to Apple products (e.g., iPhone, iPad and Mac) or HTML5-based website (Apple Inc., 2011). Microsoft's Smooth Streaming (IIS, 2011) and Adobe's Flash Player (Adobe Systems, 2011) have also supported the HTTP-based streaming solution. Other

strategies are also proved to be helpful in improving user's viewing experience. For instance, to avoid people waiting for too long to get the first glance of a video, a short and very low start (i.e., low bitrate for the first part of a video) is a good way to shorten the beginning buffering time (Buchinger & Hlavacs, 2008). To reduce the delay between mobile TV channel switches, one solution is to play preloaded video clips, such as short commercial clips and useful information guides (Robitza, Buchinger, Hummelbrunner & Hlavacs, 2010).

Though the understandings of other UX factors such as context and user profiles are limited, there are still some adaptation schemes attempted to tackle these factors. For example, an acceptance threshold adaptive strategy determines the lowest acceptable bitrate and the corresponding encoding parameters based on video content type and user information, including prior experience of mobile video, preference for video content, and age group (Song, Tjondronegoro & Docherty, 2010). Different real usage contexts (Ketyko et al., 2010; Song, Tjondronegoro & Docherty, 2010) and different mobile devices (Agboma & Liotta, 2010) have also been considered when establishing QoE models. Using an emotional index (e.g., satisfaction and pleasantness) to control the video delivery may provide more user-friendly service (Ketyko et al., 2010; Song et al., 2011).

Apart from the above adaptation strategies that deal with the resource limits, there are some non-adaptation strategies to improve UX through exercising control on the factors like commercial plan, user interface, and social context. For charged mobile video services such as Mobile TV, the right pricing approach should be to give users a choice of various payment options anyway according to Trefzger (2005). The options can be a fixed monthly fee model, single content payment, time-based payment, or event-based payment (Buchinger, Kriglstein, et al., 2009; Carlsson, Carlsson, Puhakainen & Walden, 2006; Carlsson & Walden, 2007). Free mode financed by ads may also attract users (The Nielsen Company, 2011; Winder, 2001). Regarding the improvement of usability and interactivity, most current studies focus on effective user interface design, such as content navigation (Jumisko-Pyykkö, Weitzel, et al., 2008), searching (Hussain et al., 2008), and new technology utilization of mobile devices (MacLean, 2008). Huber et al. (2010) have evaluated seven interface concepts for mobile video browsing, including GUI-based, touch-gesture based, and physical interaction; and then derived some design guidelines to improve the usability and user perception of mobile video browsers. The guides include supporting spatio-temporal browsing metaphors and discrete temporal navigation, and carefully placing interface elements. Schatz and Egger (2008) have found that social features could enrich the mobile video experience; they have suggested to use more and flexible interaction means (e.g., audio, text, non-verbal - thumbs up/down display, shaking, and squeezing, etc.) for social mobile video service to address the diverse influences on ratings and user attitudes, and to remove barriers in mobile social activity. Furthermore, allowing users to edit, upload and share user-created contents can provide excellent interactive experience for mobile video users (Orgad, 2006).

Given the truth that user experience is the consequence of interaction of all variety of factors in the entire delivery process, comprehensive considerations of the source, the network environment and the end user is expected to bring great benefits in improving the QoE of the service. However, only a little research has aimed to enable holistic QoE measurement and modeling, but is still at a conceptual stage (De Moor et al., 2010). In conclusion, due to

great difficulties in thoroughly understanding UX and implementing complicated adaptations, it is still a grant challenge to achieve optimal UX for mobile video services.

5. Conclusion

With the increase in demand for mobile video, user experience (UX) of mobile video has attracted a lot of attention from both video service providers and researchers. This chapter aimed to provide a better understanding of UX of mobile video by describing three issues: framework, measurement and optimization. We firstly reviewed a body of well-known research on UX and generalized their similarities and differences. Drawing from these studies and combining the concluded UX attributes with the influencing factors of mobile video, we then proposed a UX framework for mobile video and further elucidated the factors with rich literature. The framework is simple but encompasses many factors and clears each factor's contribution to the overall UX. It may benefit in user-centred design of mobile video delivery and relative research. Mobile video vendors may develop effective strategies to improve UX by taking into consideration the factors in different components of the framework. Researchers may use the framework to determine research direction and target (e.g., what will be focused, a single part technical improvement or a holistic understanding?), or to design a proper user study scheme (e.g., what aspects will be investigated? what kind of participants and contexts should be involved?).

To measure user experience, Quality of Experience (QoE) is commonly used in multimedia service. We reviewed many existing QoE metrics that provide predictions of the perceptual video quality, and introduced some QoE frameworks that provide comprehensive considerations of QoE. In line with the proposed UX framework of mobile video, we summarized the methods to optimize UX into adaptation strategies (such as video coding adaptations, video transmission adaptations and other adaptations) and non-adaptation strategies.

Concluding the research in UX measurement and optimization, there are some open issues needing further study. Firstly, in user-centered research, empirical and field studies are needed to collect natural and real user data, which can represent a more realistic UX. To reduce the study cost, a living lab testbed, like that proposed by De Moor et al. (2010), may bring many benefits for future research. Secondly, to achieve a good prediction of the quality of UX, accurate QoE models need take more UX aspects into consideration. However, how to determine the key aspects of UX and how to choose proper indicators to represent them is difficult. It is more challenging to find out the correlations between these factors. Thirdly, some influencing factors of UX have not been sufficiently studied, such as user profiles and context. Also, how to utilize or adapt to the impacts of these factors for UX optimization is still a big question. Overall, automatic adaptation strategies can take care of general user's requirements. However, the automatic adaptation may not always be optimal in terms of individual user's QoE. It was argued that automatic adaptation may benefit from manual adjustments by the user in some cases, such as incomplete usage environment information; incomplete content characteristics; interest of the user in a specific content; limit of user's mobile data contract; and the viewing environment of the user (Murillo, Ransburg & Graciá, 2010).

6. References

Adobe Systems (September 2011). HTTP Dynamic Streaming, In: *Adobe*, 2.09.2011. Avaialable from http://www.adobe.com/products/httpdynamicstreaming/

Agboma, F., & liotta, A. (2007). Addressing User Expectations in Mobile Content Delivery *Mobile Information Systems,* Vol.3, No.4, (October 2007), pp. 153-164, ISSN 1574-017x

Agboma, F., & Liotta, A. (2008). QoE-aware QoS Management, *Proceedings of 6th International Conference on Advances in Mobile Computing and Multimedia,* pp. 111-116, Linz, Austria, November 24-26, 2008

Agboma, F., & Liotta, A. (2010). Quality of Experience Management in Mobile Content Delivery Systems. *Telecommunication Systems,* pp. 1-14, ISSN 1018-4864

Ahmad, A. M. A. (2006a). Content and Resources Based Mobile and Wireless Video Transcoding. *Proceedings of WASET'06 World Academy of Science, Engineering and Technology,* pp. 49-54, November, 2006

Ahmad, A. M. A. (2006b). Content-based Video Streaming Approaches and Challenges. In: *Handbook of Research on Mobile Multimedia,* I. K. Ibrahim (Ed.), pp. 357-367, Idea Group Reference, London, UK

Ahmad, A. M. A., & Lee, S.-Y. (2008). Fast and Robust Object-extraction Framework for Object-based Streaming System. *International Journal of Virtual Technology and Multimedia,* Vol.1, No.1, (Feburary 2008), pp. 39-60, ISSN 1741-1874

Alben, L. (1996). Quality of Experience. *Interactions,* Vol.3, No.3, (May 1996), pp. 11-15, ISSN 1072-5520

Apple Inc. (September 2011). HTTP Live Streaming, In: *Apple Inc.,* 2.09.2011, Available from http://tools.ietf.org/html/draft-pantos-http-live-streaming-06

Arhippainen, L., Tähti, M. (2003). Empirical Evaluation of User Experience in Two Adaptive Mobile Application Prototypes. *Proceedings of 2nd International Conference on Mobile and Ubiquitous Multimedia.* pp. 27-34, Norrköping, Sweden

Bradeanu, O., Munteanu, D., Rincu, I., & Geanta, F. (2006). Mobile Multimedia End-user Quality of Experience Modeling. *Proceedings of International Conference on Digital Telecommunications, ICDT '06,* pp. 49-52, Cap Esterel, Cote d'Azur, France

Buchinger, S., & Hlavacs, H. (2008). A Low Start for Mobile Video Patching. *Proceedings of Euro FGI IA.7.6 Workshop on Socio-Economoic Aspects of Future Generation,* Karlskrona, Sweden

Buchinger, S., Kriglstein, S., Brandt, S., & Hlavacs, H. (2011). A Survey on User Studies and Technical Aspects of Mobile Multimedia Applications. *Entertainment Computing, In Press,* ISSN 1875-9521

Buchinger, S., Kriglstein, S., & Hlavacs, H. (2009). A Comprehensive View on User Studies: Survey and Open Issues for Mobile TV. *Proceedings of 7th European Conference on Interactive Television.* pp. 179-188, Leuven, Belgium

Buchinger, S., Nezveda, M., Robitza, W., Hummelbrunner, P., & Hlavacs, H. (2009). Mobile TV Coding, *Proceedings of Workshop on IPTV Technology and Multidisciplinary Application.* Zagreb, Croatia.

Carlsson, C., Carlsson, J., Puhakainen, J., & Walden, P. (2006). Nice Mobile Services do not Fly: Observations of Mobile Services and the Finnish Consumers. *Proceedings of 19th Bled eCommerce Conference,* Bled, Slovenia

Carlsson, C., & Walden, P. (2007). Mobile TV-to Live or Die by content, *Proceedings of 40th Annual Hawaii International Conference on System Sciences (HICSS)*, pp. 51-57

Chang, S.-F., & Vetro, A. (2005). Video adapation: Concepts, Technologies, and Open Isses. *Proceedings of The IEEE*, Vol.93, No.1, (27 June 2005), pp. 148-158, ISSN 0018-9219

Chen, Q., Song, L., Yang, X., & Zhang, W. (2007). Robust Region-of-Interest Scalable Coding with Leaky Prediction in H.264/AVC, *2007 IEEE Workshop on Signal Processing Systems*, pp. 357-362

Chikkerur, S., Sundaram, V., Reisslein, M., & Karam, L. J. (2011). Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison. *IEEE Transactions on Broadcasting*, Vol.57, No.2, (June 2011), pp. 165-182, ISSN 0018-9316

Chipchase, J., Yanqing, C., & Jung, Y. (2006). Personal Television: A Qualitative Study of Mobile TV Users in South Korea, *Proceedings of Mobile HCI 2006*, Espoo, Finland

Claypool, M., & Tripathi, A. (2004). Adaptive Video Streaming Using Content-aware Media Scaling, (Janurary 7, 2004), pp. 20, Worcester, Massachusetts, USA: Worcester Polytechnic Institute

Cranley, N., Murphy, L., & Perry, P. (2004). Content-based Adaptation of Streamed Multimedia, *Proceedings of 7th International Conference on Management of Multimedia Networks and Services*. pp. 39-49, ISSN 0302-9743, San Diego, CA, Oct 03-06, 2004

Cranley, N., Perry, P., & Murphy, L. (2006). Dynamic Content-based Adaptation of Streamed Multimedia. *Journal of Network and Computer Applications*, Vol.30, No.3, pp. 983-1006, ISSN 1084-8043

de Koning, T. C. M., Veldhoven, P., Knoche, H., & Kooij, R. E. (2007). Of MOS and Men: Bridging the Gap between Objective and Subjective Quality Measurements in Mobile TV. *Proceedings of Multimedia on Mobile Devices 2007, IS&T/SPIE Symposium on Electronic Imaging*, San Jose, CA, USA, 28 Januray 2007

De Moor, K., Ketyko, I., Joseph, W., Deryckere, T., De Marez, L., Martens, L., et al. (2010). Proposed Framework for Evaluating Quality of Experience in a Mobile, Testbed-oriented Living Lab Setting. *Mobile Networks and Applications*, Vol.15, No.3, (29 January 2010), pp. 378-391, ISSN 1383-469x

Eichhorn, A., & Ni, P. (2009). Pick Your Layers Wisely - A Quality Assessment of H.264 Scalable Video Coding for Mobile Devices. *Proceedings of IEEE International Conference on Communications*, pp.1019-1025, ISBN 978-1-4244-3435, June 14-18, 2009

Engelke, U., & Zepernick, H.-J. (2009). Optimal Region-Of-Interest Based Visual Quality Assessment. *Proceedings of Human Vision and Electronic Imaging XIV*, Vol.7240, No. 11, (Janurary 2009), pp. 1-12,

Eronen, L. (2001). Combining quantitative and qualitative data in user research on digital television. *Proceedings of 1st Panhallenic Conference PC HCI'01*. pp. 51-56, Patras, Greece, December 7-9, 2001

Fiedler, M., Hossfeld, T., & Tran-Gia, P. (2010). A Generic Quantitative Relationship between Quality of Experience and Quality of Service. *IEEE Network: The Magazine of Global Internetworking - Special issue on improving quality of experience for network services*, Vol.24, No.2, (March/April 2010), pp. 36-41, ISSN 0890-8044

Forlizzi, J., & Ford, S. (2000). The Building Blocks of Experience: An Early Framework for Interaction Designer. *Proceedings of Dis 2000*, pp. 419-423, ISBN 1-58113-219-0, Brooklyn, New York, n. d.

Geerts, D., De Moor, K., Ketyko, I., Jacobs, A., Van den Bergh, J., Joseph, W., et al. (2010). Linking an Integrated Framework with Appropriate Methods for Measuring QoE. *Proceedings of second International Workshop on Quality of Multimedia Experience (QoMEX 2010),* pp. 158-163, Trondheim, Norway, June 21-23, 2010

Ghazal, M., Amer, A., & Ghrayeb, A. (2007). A Real-Time Technique for Spatio-Temporal Video Noise Estimation.*IEEE Transactions on Circuits and Systems for Video Technology,* Vol.17, No.12, (December 2007), pp. 1690-1699, ISSN 1051-8215

Girod, B., Kalman, M., Liang, Y. J., & Zhang, R. (2002). Advances in Channel-adaptive Video Streaming. *Proceedings of International Conference on Image Processing*, pp. I9-12, ISSN 1522-4880, Rochester, New York, USA, 22-25 September, 2002

Guixu, L., Shibao, Z., & Jianling, H. (2008). A Two-stage ρ-domain Rate control Scheme for H.264 Encoder. *Proceedings of 2008 IEEE International Conference on Multimedia and Expo,* pp. 713-716, ISBN 978-1-4244-2570-9, Hannover, Germany , June 23-26, 2008

Hassenzahl, M., & Tractinsky, N. (2006). User Experience - A Research Agenda. *Behaviour & Information Technology,* Vol.25, No.2, (March-April 2006), pp. 91-97, ISSN 0144-929x

Hemami, S. S., & Reibman, A. R. (2010). No-reference Image and Video Quality Estimation: Applications and Human-motivated Design. *Signal processing: Image communication,* Vol.25, No.27, (2010), pp. 469-481, ISSN 0923-5965

Huber, J., Steimle, J., & Mühlhäuser, M. (2010). Toward More Efficient User Interfaces for Mobile Video Browsing: An In-depth Exploration of the Design Space. *Proceedings of 16th international conference on Multimedia (ACM MM'10)*, pp. 341-350, Firenze, Italy, October 25-29, 2010

Hussain, Z., Lechner, M., Milchrahm, H., Shahzad, S., Slany, W., Umgeher, M., et al. (2008). User Interface Design for a Mobile Multimedia Application: An Iterative Approach. *Proceedings of International Conference on Advances in Computer-Human Interaction - ACHI'08*, pp. 189-194, Sainte Luce, Martinique, February 10-15, 2008

IIS (2011). Smooth Streaming. In: *ISS.net*, 2.09.2011, Available from http://www.iis.net/download/SmoothStreaming

ISO (2010). Ergonomics of Human-system Interaction -- Part 210: Human-centred Design for Interactive Systems, *9241-210: 2010*. In: *International Organization for Standardization (ISO)*, Switzerland

ITU-R Recommendation BT. 500-11: (2004). Methodology for the Subjective Assessment of Quality for Television Pictures. In: *International Telecommunication Union*

ITU-T (1999). Subjective Video Quality Assessment Methods for Multimedia Applications. In: *International Telecommunication Union*

ITU-T (2007). Quality of Service and Performance – Generic and User-related Aspects: Opinion Model for Video-telephony Applications (ITU-T Recommendation G.1070). In: *International Telecommunication Union*

ITU-T Study Group 12 (2007). Definition of quality of experience, *Reference: TD 109rev2 (PLEN/12)*. In: *International Telecommunication Union*

Jerbi, A., Wang, J., & Shirani, S. (2005). Error-resilient region-of-interest video coding. *IEEE Transactions on Circuits and Systems for Video Technology,* Vol.15, No.9, (September 2005) pp. 1175-1181, ISSN 1051-8215

Jordan, P. W. (2002). *Designing Pleasurable Products: An Introduction to the New Human Factors.* Taylor & Francis, ISBN 0-203-30568-X, London, English

Joskowicz, J., & Ardao, J. C. L. (2010). A Parametric Model for Perceptual Video Quality Estimation. *Telecommunication System,* Vol.46, (June 2010), pp. 14, ISSN 1018-4864

Jumisko-Pyykkö, S., & Häkkinen, J. (2005). Evaluation of Subjective Video Quality on Mobile Devices. *Proceedings of 13th ACM International Conference on Multimedia.* pp. 535-538, Singapore, 6-12 November, 2005

Jumisko-Pyykkö, S., Häkkinen, J., & Nyman, G. (2007). Experienced Quality Factors-Qualitative Evaluation Approach to Audiovisual Quality. *Proceedings of SPIE Multimedia on Mobile Devices*

Jumisko-Pyykkö, S., & Hannuksela, M. M. (2008). Does Context Matter in Quality Evaluation of Mobile Television? *Proceedings of 10th International Conference on Human Computer Interaction with Mobile Devices and Services.* pp. 63-72, ISBN 978-1-59593-952-4, Amsterdam, The Netherlands, September 2–5, 2008

Jumisko-Pyykkö, S., Ilvonen, V. P., & Väänänen-Vainio-Mattila, K. A. (2005). Effect of TV Content in Subjective Assessment of Video Quality on Mobile Devices. *Proceedings of Multimedia on Mobile Devices.* pp. 243-254, San Jose CA, 17-18 January 2005

Jumisko-Pyykkö, S., Vadakital, V. K. M., & Hannuksela, M. M. (2008). Acceptance Threshold: A Bidimensional Research Method for User-oriented Quality Evaluation Studies. *International Journal of Digital Multimedia Broadcasting,* Vol.2008, pp. 20, ISSN 1687-7578

Jumisko-Pyykkö, S., Weitzel, M., & Strohmeier, D. (2008). Designing for User Experience: What to Expect from Mobile 3D TV and Video? *Proceedings of 1st International Conference on Designing Interactive User Experiences for TV and Video,* pp. 183-192, ISBN 978-1-60558-100-2, Silicon Valley, California, USA, October 22–24, 2008

Kaasinen, E., Kulju, M., Kivinen, T., & Oksman, V. (2009). User Acceptance of Mobile TV Services. *Proceedings of Mobile HCI 09.* pp. 1-10, ISBN 978-1-60558-281-8, Bonn, Germany, September 15 - 18, 2009

Ketyko, I. n., De Moor, K., Joseph, W., Martens, L., & De Marez, L. (2010). Performing QoE-Measurements in an Actual 3G Network. *Proceedings of 2010 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting (BMSB)* pp. 1-6. Shanghai, China, March 24-26, 2010

Khan, A., Sun, L., Jammeh, E., & Ifeachor, E. (2010). Quality of Experience-driven Adaptation Scheme for Video Applications over Wireless Networks. *IET Communications,* Vol.4, No.11, (July 2006), pp. 1337-1347, ISSN 1751-8628

Kim, J.-G., & Suh, J.-W. (2009). An Efficient MB Layer Rate Control for H.264/AVC Based on Adaptive QP Decision. *Proceedings of Advances in Multimedia Information Processing-PCM 2009,* pp. 1058-1067 ISBN 978-3-642-10466-4 Bangkok, Tailand, December 15-18, 2009

Kim, S.-H., Heo, J., & Ho, Y.-S. (2010). Efficient Entropy Coding Scheme for H.264/AVC Lossless Video Coding. *Signal Processing: Image Communication* Vol.25, No.9, (October 2010), pp. 687-696

Knoche, H., & McCarthy, J. D. (2004). Mobile Users' Needs and Expectations of Future Multimedia Services. *Proceedings of Wireless World Research Forum (WWRF)12.* pp. 1-8, Toronto, Canada, n.d.

Knoche, H., McCarthy, J. D., & Sasse, M. A. (2005). Can Small be Beautiful?: Assessing Image Resolution Requirements for Mobile TV. *Proceedings of 13th annual ACM international conference on Multimedia.* pp.829-838, ISBN 1-59593-044-2, Hilton, Singapore, November 6–11, 2005

Knoche, H., Papaleo, M., Sasse, M. A., & Vanelli-Coralli, A. (2007). The Kindest Cut: Enhancing the User Experience of Mobile TV through Adequate Zooming. *Proceedings of 15th international conference on Multimedia.* pp. 87-96, ISBN 978-1-59593-701-8, Augsburg, Bavaria, Germany, September 23–28, 2007

Knoche, H., & Sasse, M. A. (2008). Getting the Big Picture on Small Screens: Quality of Experience in Mobile TV. In: *Multimedia Transcoding in Mobile and Wireless Networks,* A. M. A. Ahmad & I. K. Ibrahim (Eds.), pp. 31-46, Information Science Reference, ISBN 1-59904-984-8, New York, USA

Kun, T., Richard, R., & Shih-Ping, L. (2001). Content-sensitive Video Streaming over Low Bitrate and Lossy Wireless Network. *Proceedings of 9th ACM international conference on Multimedia.* pp. 512-515, ISBN 1-58113-394-4, Ottawa, Canada, September 30 - October 5, 2001

Lambert, P., Schrijver, D. D., Deursen, D. V., Neve, W. D., Dhondt, Y., Walle, R. V. d. (2006). A Real-time Content Adaptation Framework for Exploiting ROI Scalability in H.264/AVC, *Proccedings of 8th International Conference of Advanced Concepts for Intelligent Vision Systems,* pp. 442-453, Antwerp, Belgium, September 18-21, 2006

Law, E. L.-C., Roto, V., Hassenzahl, M., Vermeeren, A. P. O. S., & Kort, J. (2009). Understanding, Scoping and Defining User Experience: A Survey Approach. *Proceedings of 27th international conference on Human factors in computing systems.* pp. 719-728, ISBN 978-1-60558-246-7, Boston, MA, USA, April 4-9, 2009

Leontaris, A., & Reibman, A. (2005). Comparison of Blocking and Blurring Metrics for Video Compression. *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP'05),* pp. 585-589, ISSN 1520-6149, Philadelphia, USA, 23-23 March 2005

Liu, S., & Kuo, C. C. J. (2003). Joint Temporal-spatial Rate Control for Adaptive Video Transcoding. *Proceedings of 2003 International Conference on Multimedia and Expo, ICME '03,* pp. II-225-228, ISBN 0-7803-7965-9, July 6-9, 2003

Liu, Y., Li, Z. G., & Soh, Y. C. (2008). Region-of-Interest Based Resource Allocation for Conversational Video Communication of H.264/AVC. *IEEE Transaction on Circuits and Systems for Video Technology,* Vol.18, No.1, (January 2008), pp. 134-139, ISSN 1051-8215

Lu, Z., Lin, W., Li, Z., Lim, K. P., Lin, X., Rahardija, S., et al. (2005). Perceptual Region-of-interest (ROI) Based Scalable Video Coding, In: *Joint Video Team (JVT) of ISO/IEC 15th meeting,* pp. 16-22, Busan, KR, 16-22 April, 2005

MacLean, K. E. (2008). Using Haptics for Mobile Information Display, *Proceedings of Pervasive Mobile Interaction Devices (PERMID'08), Workshop at 6th International Conference on Pervasive Computing,* pp. 175-179, Sydney, Australia, May 19, 2008

Mäki, J. (2005). Finnish Mobile TV Pilot, In: *Research International Finland*, Augst 2005.

Marziliano, P., Dufaux, F., Winkler, S., & Ebrahimi, T. (2002). A No-reference Perceptual Blur Metric. *Proceedings of 2002 International Conference on Image Process*, pp. 57-60, ISSN 1522-4880, Rochester, New York, USA, September 22-25, 2002

Masry, M., & Hemami, S. S. (2002). Perceived Quality Metrics for Low Bit Rate Compressed Video. *Proceedings of 2002 International Conference on Image Processing*. pp. 49-52, ISSN 1522-4880, Rochester, New York, USA, September 22-25, 2002

McCarthy, J., & Wright, P. (2004). *Technology as experience*. MIT Press, Cambridge MA.

McCarthy, J. D., Sasse, M. A., & Miras, D. (2004). Sharp or Smooth?: Comparing the Effects of Quantization vs. Frame Rate for Streamed Video. *Proceedings of SIGCHI Conference on Human Factors in Computing Systems*. pp. 535-542, ISBN 1-58113-702-8, Vienna, Austria, April 24–29, 2004

Menkovski, V., Oredope, A., Liotta, A., & Sánchez, A. C. (2009). Predicting Quality of Experience in Multimedia Streaming. *Proceedings of the 7th International Conference on Advances in Mobile Computing and Multimedia*. pp. 52-59, ISBN 978-1-60558-659-5, Kuala Lumpur, Malaysia, December 14–16, 2009

Miyauchi, K., Sugahara, T., & Oda, H. (2008). Relax or Study?: A Qualitative User Study on the Usage of Mobile TV and Video. *Proceedings of the 6th EuroITV*, pp. 128-132, ISBN 978-3-540-69478-6, Salzburg, Austria, July 3-4, 2008

Moller, S., Engelbrecht, K. P., Kuhnel, C., Wechsung, I., & Weiss, B. (2009). A taxonomy of Quality of Service and Quality of Experience of Multimodal Human-machine Interaction. *Proccedings of International Workshop on Quality of Multimedia Experience, (QoMEx 2009),* pp. 7-12, Trondheim, Norway, 29-31 July 2009

Muntean, G. M., Ghincea, G., & Sheehan, T. N. (2008). Region of Interest-based Adaptive Multimedia Streaming Scheme. *IEEE Transactions on Broadcasting,* Vol.54, No.2, (June 2008) pp. 296-303, ISSN 0018-9316

Murillo, J. O., Ransburg, M., & Graciá, E. M. (2010). Towards User-driven Adaptation of H.264/SVC Streams. P*roceedings of EuroITV2010*. ISBN 978-1-60558-31-5, Tampere, Finland, June 9-11, 2010

Nokia (2004). Quality of Experience (QoE) of Mobile Services: Can It be Measured and Improved? In: Nokia, 30.08.2009, Available from www.nokia.com/NOKIA_COM_1/Operators/Downloads/Nokia_Services/white paper_qoe_net.pdf

Norman, D. A. (2004). *Emotional Design: Why We Love (Or Hate) Everyday Things*. Basic Book, New York

Obrist, M., Meschtscherjakov, A., & Tscheligi, M. (2010). User Experience Evaluation in the Mobile Context. In: *Mobile TV: Customizing Content and Experience,* A. Marcus et al. (Ed.), pp. 195-204, ISBN 978-1-84882-701-1, Springer-Verlag, London, UK

Orgad, S. (2006). This Box was Made for Walking, In: *Department of Media and Communications, London School of Economics and Political Science.*

Pinson, M. H., & Wolf, S. (2004). A New Standardized Method for Objectively Measuring Video Quality. *IEEE Transactions on Broadcasting,* Vol.50, No.3, (September 2004), pp. 312-322, ISSN 0018-9316

Reguant, V. D., Prats, F. E., de Pozuelo, R. M., Margalef, F. P., & Ubiergo, G. F. (2008). Delivery of H264 SVC/MDC Streams over WiMAX and DVB-T Networks. *ISCE 2008. IEEE International Symposium on Consumer Electronics,* pp. 1-4, ISBN 978-1-4244-2422-1, April 14-16, 2008

Ries, M., Nemethova, O., & Rupp, M. (2008). Video Quality Estimation for Mobile H.264/AVC Video Streaming. *Journal of Communications,* Vol.3, No.1, (Juanuary 2008), pp. 41-50, ISSN 1796-2021

Robitza, W., Buchinger, S., Hummelbrunner, P., & Hlavacs, H. (2010). Acceptance of Mobile TV Channel Switching Delays. *Proceedings of Workshop on Quality of Multimedia Experience.* pp. 236-241, Trondheim, Norway, June 21-23, 2010

Roto, V. (2006a). User Experience Building Block. *Proceedings of NordiCHI'06.* pp. 1-4, Oslo, Norway, October 14-18, 2006

Roto, V. (2006b). *Web Browsing on Mobile Phones - Characteristics of User Experience.* Unpublished Doctoral Dissertation, Helsinki University of Technology, Finland.

Saad, M. A., Bovik, A. C., & Charrier, C. (2010). A DCT Statistics-Based Blind Image Quality Index. *IEEE Signal Processing Letters,* Vol.17, No.6, (June 2010), pp. 583-586, ISSN 1070-9908

Sadaka, N. G., Karam, L. J., Ferzli, R., & Abousleman, G. P. (2008). A No-reference Perceptual Image Sharpness Metric based on Saliency-weighted Foveal Pooling, *Proceedings of 15th IEEE International Conference on Image Processing.* pp. 369-372, ISSN 1522-4880, San Francisco, California, USA, October 12-15, 2008

Sasse, M. A., & Knoche, H. (2006). Quality in Context-an Ecological Approach to assessing QoS for mobile TV. *Proceedings of 2nd ISCA/DEGA Tutorial & Research Workshop on Perceptual Quality of System.* Berlin, Germany, September 3-5, 2006

Schatz, R., & Egger, S. (2008). Social Interaction Features for Mobile TV Services. *Proceedings of 2008 IEEE International Symposium on Broadband Multimedia Systems and Broadcasting.* pp. 1-6, Las Vegas, NV, USA, March 31 - April 2 2008

Schatz, R., Egger, S., & Platzer, A. (2011). Poor, Good Enough or Even Better? Bridging the Gap between Acceptability and QoE of Mobile Broadband Data Services. *Proceedings of IEEE International Conference on Communications ICC 2011.* pp. 1-6, ISBN 978-1-61284-232-5, Kyoto, Japan, 5-9 June 2011

Seo, K., & Kim, C. (2006). A Context-aware Video Display Scheme for Mobile Devices *Proceedings of Multimedia on Mobile Devices II.* pp. 278-287, San Jose, CA, USA

Seshadrinathan, K., & Bovik, A. C. (2010). Motion Tuned Spatio-temporal Quality Assessment of Natural Videos. *IEEE transactions on image processing,* Vol.19, No.2, (February 2010), pp. 335-350, ISSN 1057-7149

Seshadrinathan, K., Soundararajan, R., Bovik, A. C., & Cormack, L. K. (2010). Study of Subjective and Objective Quality Assessment of Video. *IEEE Transaction on Image Processing,* Vol.19, No.6, (June 2010), pp. 1427-1441, ISSN 1057-7149

Sheikh, H. R., & Bovik, A. C. (2006). Image Information and Visual Quality, *IEEE Transactions on Image Processing,* Vol.15, No.2, (Februray 2006), pp. 430-444, ISSN 1057-7149

Shenghong, H., Yufu, J., & Shenglong, T. (2010). Content Aware Retargeting of Soccer Video, *Proceedings of 2010 International Conference on Information Science and Engineering (ICISE),* pp. 1-4, ISBN 978-1-4244- 7618-3, Hangzhou, China, December 4-6, 2010

Shi, Y., Yue, S., & Yin, B. (2008). A Novel Rate Control Scheme with Region-of-Interest Concern, *Proceedings of the 5th International Conference on Visual Information Engineering, VIE 2008,* pp. 634-637, ISSN 0537-9989, July 29-Auguset 1, 2008.

Södergård, C. (2003). Mobile Television-technology and User Experiences, In: *Report on the Mobile-TV project. VTT, Finland*, VTT Information Technology.

Song, W., & Tjondronegoro, D. (2010). A Survey on Usage of Mobile Video in Australia, *Proceedings of OZCHI 2010.* pp. 5-8, ISBN 978-1-4503-0502-0, Brisbane, Australia, 22-26 Nov 2010

Song, W., Tjondronegoro, D., & Azad, S. (2010). User-centered video quality assessment for scalable video coding of H.264/AVC standard, *Proceedings of 16th International Multimedia Modeling Conference.* pp. 55-65, Chongqing, China, Januray 6-8, 2010

Song, W., Tjondronegoro, D., & Docherty, M. (2010). Exploration and Optimisation of User Experience in Viewing Videos on A Mobile Phone. *International Journal of Software Engineering and Knowledge Engineering,* Vol.8, No.20, (December 2010), pp. 1045-1075, ISSN 0218-1940

Song, W., Tjondronegoro, D., & Docherty, M. (2011). Saving Bitrate vs. Pleasing Users: Where is the Break-even Point in Mobile Video Quality?, *Proceedings of ACM Multimedia 2011.* pp. 403-412, Scottscale, Arizona, USA.

Song, W., Tjondronegoro, D., Wang, S.-H., & Docherty, M. (2010). Impact of Zooming and Enhancing Region of Interests for Optimizing User Experience on Mobile Sports Video, *Proceedings of ACM Multimedia 2010,* pp. 321-330, ISBN 978-1-60558-933-6, Firenze, Italy, October, 25-29, 2010

Sun, Y., Ahmad, I., & Zhang, Y. Q. (2006). Region-based Rate Control, and Bit Allocation for Wireless Video Trasmission. *IEEE Transactions on Multimedia,* Vol.8, No.1, (Februray 2006), pp. 1-10, ISSN 1520-9210

Tasaka, S., Yoshimi, H., & Hirashima, A. (2008). The Effectiveness of a QoE-based Video Output Scheme for Audio-video IP Transmission, *Proceedings of ACM Multimedia Information System.* pp. 259-268, ISBN 978-1-60558-303-7, Vancouver, Canada, October 26-31, 2008.

The Nielsen Company (2011). Telstra Smartphone Index. In: *The Nielsen Company,* 10.07.2011, Available from
http://sensisdigitalmedia.com.au/Files/Mobile/Nielsen_Telstra_Smartphone_Index_June2011_Presentation.pdf

Tominaga, T., hayashi, T., Okamoto, J., & Takahashi, A. (2010). Performance Comparisions of Subjective Quality Assessment Methods for Mobile Video. *Proceedings of 2nd International Workshop on Quality of Multimedia Experience (QoMEX 2010).* pp. 82-87, Trondheim, Norway, June 21-23, 2010

Trefzger, J. (2005). Mobile TV Launch in Germany: Challenges and Implications, In: *Institute for Broadcasting Economics Cologne University*.

Tripathi, A., & Claypool, M. (2002). Improving Multimedia Streaming with Content-aware Video Scaling, *Proceedings of 2nd International Workshop on Intelligent Multimedia Computing and Networking (IMMCN)*, Durham, North Carolina, USA, March 8-12, 2002

Vangenck, M., Jacobs, A., Lievens, B., Vanhengel, E., & Pierson, J. (2008). Does Mobile Television Challenge the Dimension of Viewing Television? An Explorative Research on Time, Place and Social context of the Use of Mobile Television Content, *Proceedings of EuroITV 2008*, pp. 122-127, Salzburg, Austria, July 3-4, 2008

Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2004). Image Quality Assessment: From Error Visibility to Structural Similarity. *IEEE Transaction on Image Processing,* Vol.13, No.4, (April 2004), pp. 600-612, ISSN 1057-7149

Wang, Z., Sheikh, H. R., & Bovik, A. C. (2004). Objective Video Quality Assessment. In *Handbook of Video Databases Design and Applications,* B. Furht & O. Marques (Eds.), pp. 1041-1078, CRC Press LLC, ISBN 9780849370069

Wang, Z., Simoncelli, E. P., & Bovik, A. C. (2003). Multiscale Structural Similarity for Image Quality Assessment, *Proceedings of 37th Asilomar Conference on Signals, Systems and Computers, 2003.* pp. 1398-1402, Pacific Grove, CA, USA, November 9-12, 2003

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,* Vol.13, No.7, (July 2003), pp. 560-576, ISSN 1051-8215

Winder, J. (2001). Net Content: From Free to Fee. In: *Harvard Business Publishing*, 19.08.2011, Available from http://hbr.harvardbusiness.org/2001/07/net-content-from-free-to-fee/ar/1

Winkler, S., & Faller, C. (2005). Maximizing Audiovisual Quality at Low Bitrates. *Proceedings of Workshop on Video Processing and Qaulity Metrics for Consumer Electronics*, pp. 23-25, Scottsdale, Arizona, USA, Juanuray 2005

Wu, W., Arefin, A., Rivas, R., Nahrstedt, K., Sheppard, R., Yang, Z. (2009). Quality of Experience in Distributed Interactive Multimedia Environments: Toward a Theoretical Framework. *Proceedings of 17th ACM international conference on Multimedia.* pp. 481-490, Beijing, China, October 19–24, 2009

Yang, L., Zhang, L., Ma, S., & Zhao, D. (2009). A ROI Quality Adjustable Rate Control Scheme for Low Bitrate Video Coding. *Proceedings of 27th conference on Picture Coding Symposium*, pp. 441-444, n.d.

Yun-Chung, C., Jung-Ming, W., Bailey, R. R., Sei-Wang, C., & Shyang-Lih, C. (2004). A Non-parametric Blur Measure Based on Edge Analysis for Image Processing Applications, *Proceedings of 2004 IEEE Conference on Cybernetics and Intelligent Systems*, pp. 356-360, December 1-3, 2004

Zhang, Q. (2005). Video Delivery over Wireless Multi-hop Networks. *Proceedings of 2005 International Symposium on Intelligent Signal Processing and Communication Systems.* pp. 793-796, Hong Kong, China, December 13-16, 2005

Zhenghua, Y., & Wu, H. R. (2000). Human Visual System based Objective Digital Video Quality Metrics. *Proceedings of 5th International Conference on Signal Processing (WCCC-ICSP)*. pp. 1088-1095, Beijing, China, n.d.

Zhou, W., Bovik, A. C., & Evan, B. L. (2000). Blind Measurement of Blocking Artifacts in Images, *Proceedings of 2000 International Conference on Image Processing,* pp. 981-984, ISBN 0-7803-6297-7, Vancouver, BC , Canada, September 10-13, 2000

Zinner, T., Hohlfeld, O., Abboud, O., & Hossfeld, T. (2010). Impact of Frame Rate and Resolution on Objective QoE Metrics. *Proceedings of 2nd International Workshop on Quality of Multimedia Experience (QoMEX 2010).* pp. 29-34, Trondheim, Norway, June 21-23, 2010
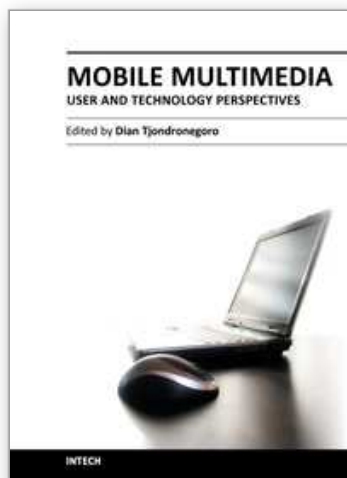
Mobile Multimedia - User and Technology Perspectives

Edited by Dr. Dian Tjondronegoro

As multimedia-enabled mobile devices such as smart phones and tablets are becoming the day-to-day computing device of choice for users of all ages, everyone expects that all mobile multimedia applications and services should be as smooth and as high-quality as the desktop experience. The grand challenge in delivering multimedia to mobile devices using the Internet is to ensure the quality of experience that meets the users' expectations, within reasonable costs, while supporting heterogeneous platforms and wireless network conditions. This book aims to provide a holistic overview of the current and future technologies used for delivering high-quality mobile multimedia applications, while focusing on user experience as the key requirement. The book opens with a section dealing with the challenges in mobile video delivery as one of the most bandwidth-intensive media that requires smooth streaming and a user-centric strategy to ensure quality of experience. The second section addresses this challenge by introducing some important concepts for future mobile multimedia coding and the network technologies to deliver quality services. The last section combines the user and technology perspectives by demonstrating how user experience can be measured using case studies on urban community interfaces and Internet telephones.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Wei Song, Dian Tjondronegoro and Michael Docherty (2012). Understanding User Experience of Mobile Video: Framework, Measurement, and Optimization, Mobile Multimedia - User and Technology Perspectives, Dr. Dian Tjondronegoro (Ed.), ISBN: 978-953-307-908-0, InTech, Available from: http://www.intechopen.com/books/mobile-multimedia-user-and-technology-perspectives/understanding-user-experience-of-mobile-video-framework-measurement-and-optimization

INTECH
open science | open minds