Towards a Multidimensional Approach for the Evaluation of Multimodal Application User Interfaces

José E.R. de Queiroz, Joseana M. Fechine, Ana E.V. Barbosa, and Danilo de S. Ferreira

Federal University of Campina Grande
Electrical Engineering and Computer Science Center – Computer Science Department
Av. Aprígio Veloso, s/n – Bodocongó, Campina Grande,
CEP 58109-970, Paraíba, Brazil
{rangel,joseana,aesther}@dsc.ufcg.edu.br,
{rangeldequeiroz,danilo.sousa}@gmail.com

Abstract. This paper focuses on a multidimensional approach for evaluating multimodal UI based upon a set of well known techniques for usability evaluation. Each technique evaluates the problem from a different perspective which combines user opinion, standard conformity assessment, and user performance measurement. The method's presentation and the evaluation analysis will be supported by a discussion of the results obtained from the method's application to a case study involving a Smartphone. The main objective of the study was to investigate the need for adapting a multidimensional evaluation approach for desktop applications to the context of multimodal devices and applications.

Keywords: Multimodal user interface; usability evaluation techniques; multidimensional approach; multimodal application.

1 Introduction

During the past two decades, a plethora of UI usability evaluation methods have been conceived. In consequence, the importance of focussing on HCI usability assessment is recognized today by academia, industry, and governments around the world. For years, users have utilized keyboards and/or mices as HCI input devices. In more recent years, a number of applications have been changing the interaction focus to multi-modal navigational cues, e.g. text + graphical/visual highlight + auditory instructions + gesture, which support a more natural interaction. Such systems whose interfaces support more than one input modality (e.g., via speech input, handwriting, gesture input, facial expressions) are known as *multimodal dialogue systems* (MDS).

In the context of such variety of I/O channels, efforts have been focused on proposing new techniques for multimodal user interface (MUI) usability evaluation. However the more evolved is the computing market toward MDS, the harder becomes for HCI evaluators to choose among the plethora of approaches for MUI usability evaluation recently proposed in the literature. It has to be taken into account that it is not a trivial task to compare between different studies, based on a number of claims made without solid statistical results. Face to these new UI features, such as dynamic

J.A. Jacko (Ed.): Human-Computer Interaction, Part II, HCII 2009, LNCS 5611, pp. 29–38, 2009.

contexts of use, mobility, and accessibility, HCI practitioners ask how effective methods, techniques and settings well known to them from previous experiences can be. And the consequent important question is if it could be possible to adapt them to MDS, by considering their distinctive features. Another question is related to whether to adopt a field or a lab approach. Although many of these issues are known to many practitioners, very little discussion is found in literature of which technique or combination of them is best suited for a specific modality and its context of use. Beyond this level of detail and for a successful choice of a methodological approach, it seems essential to know about its effectiveness, in what ways and for what purposes as well. Otherwise, he/she will expend significant effort with little payoff result.

2 Usability Assessment of MUI

It is a fact that UI are a determining factor for acceptance of interactive products. As a quality attribute, usability ables one to assess how easy UI are to use, and it can determine the difference between success and failure [1]. According to [2], usability dimensions are taken to include *effectiveness* (accuracy and completeness with which users achieve specified goals in particular environments), *efficiency* (accuracy and completeness goals achieved in relation to resources expanded); and *subjective satisfaction* (comfort and acceptability of using a system).

Despite discussing frequently to whether to adopt a field or a lab approach, it is often neglected in HCI literature effectively talking about which techniques are best suited for specific evaluation targets and their contexts of use. In order to exercise a choice of approach, this polemic subject may represents to the HCI community an equally important concern, since the efforts carried on the evaluation may not payoff if a method, technique and/or setting is not well chosen or well employed [3].

The majority of the usability studies for MUI focus on some user testing, in which user activity is monitored and/or recorded, while users perform predefined tasks. This allows the evaluator to gather quantitative data regarding how users interact with multimodal technology, which provides valuable information about usability and user experience as well. A number of types of user testing have been conducted, both in lab and in field environments, revealing user preferences for interaction modalities based on factors such as acceptance in different social contexts, e.g., noisy and mobile environments [4].

Inspection methods have also been extensively adopted, in which HCI experts rely on ergonomic knowledge to identify usability problems while inspecting the user interface. Well-known methods belonging to this category include cognitive walkthroughs (e.g., [5]), formative evaluation/ heuristic evaluation (e.g., [6]), and benchmarking approaches involving ISO 9241 usability recommendations or conformance to guidelines (e.g., [7]).

Questionnaires have been extensively administered to gather qualitative information from users, such as subjective satisfaction, perceived utility of the system, and user preferences for modality (e.g., [8]) and cognitive workload (e.g., [9]). Quite often, questionnaire use has been combined with user-testing techniques [4].

Simulation and model-based checking of system specifications have been recently used to predict usability problems, e.g. unreachable states of the systems or conflict detection of events required for fusion. A combination of task models based on concur task tree (CTT) notation with multiple data sources, such as eye-tracking data and video records has been used to better understand the user interaction [10].

A plethora of empirical studies have been carried out to determine the contribution of modalities to the user interaction, in terms of (i) understanding how usability and user acceptance is influenced by new devices and novel interaction techniques (e.g., [11][12]); (ii) showing that the perceived usability is impacted according to the kind of tasks performed (e.g., [4][13]), and according to the context of use - indoor versus outdoor conditions, mobile applications (e.g., [14]); or (iii) trying to assess the accuracy of multimodal interaction for given tasks (e.g., [15][16]).

3 The Multidimensional Evaluation Approach

The original version of the present approach was conceived for evaluating desktop application UI [17], and further adapted to evaluate the usability of mobile application UI [18]. With the aim of providing complementary results and more integrated usability information, it is based upon the premises that (i) each evaluation technique provides a different level of information, which will help the evaluator to identify usability problems from a specific point of view; and (ii) data triangulation can be used to review information obtained from multiple data sources and techniques. This resulted in a hybrid strategy which encompasses the best aspects of (i) *standards inspection*; (ii) *user performance measurement*; and (iii*) user inquiry*.

According to [2] *conformity assessment* means checking whether products, services, materials, processes, systems, and personnel measure up to the requirements of standards. For conformity assessment, the desktop version of the multidimensional approach adopts the standard ISO 9241 (*Ergonomic Requirements for Office Work with Visual Display Terminals*). In its multimodal version, and more specifically for the Smartphone case study it was found that only some parts of ISO 9241 could be applied - 14 [19], 16 [20], and 17 [21]. There are also some other standards that apply to this kind of device such as the ISO/IEC 14754 (*Pen-based Interfaces—Common gestures for text editing with pen-based systems*) [22]; ISO 9241-171 (*Ergonomics of human-system interaction — Guidance on software accessibility*) [23]; ISO/IEC 24755 (*Information technology — Screen icons and symbols for personal mobile communication devices*) [24]; ISO/IEC 18021 (*User interfaces for mobile tools for management of database communications in a client-server model*) [25]; and ITU-TE.161 (Arrangement of digits, letters, and symbols on telephones and other devices that can be used for gaining access to a telephone network, also known as ANSI T1.703-1995/1999); and ISO/IEC 9995-8:1994) [26].

In general, *user performance measurement* aims to enable real time monitoring of user activities, providing data on the effectiveness and efficiency of his/her interaction with a product. It also enables comparisons with similar products, or with previous versions of the same product along its development lifecycle, highlighting areas where the product usability can be improved. When combined with the other methods, it can provide a more comprehensive view of the usability of a system. The major change introduced in the original evaluation approach concerns the introduction of field tests as a complement to the original lab tests.

User subjective satisfaction measurement has been widely adopted as an IS success indicator, and has been the subject of a number of researches since the 1980s (e.g., [27][28]). User satisfaction diagnosis provides an insight into the level of user satisfaction with the product, highlighting the relevance of the problems found and their impact on the product acceptance. In this approach, user subjective satisfaction data are gathered from three methods: (i) *automated questionnaires* administered before and after test sessions; (ii) *informal think-aloud trials* performed during test sessions; and (iii) *unstructured interviews* conducted at the end of test sessions.

Going beyond the definition of *usability* in function of *effectiveness*, *efficiency* and *satisfaction* [7], ISO also defines that at least one indicator in each of these aspects should be measured to determine the level of usability achieved. As briefly exposed in this section, the multidimensional approach presented here meets the requirements set by ISO 9241-11 because it is used: (i) the *task execution time* as an efficiency indicator; (ii) the *number of incorrect actions*, the *number of incorrect choices*, the *number of repeated errors*, and the *number of accesses to the online/ printed help* as effectiveness indicators; and (iii) the *think-aloud comments*, the *unstructured interview responses*, and the *questionnaire scores* as subjective satisfaction indicators.

4 The Experiment: Use of Smartphone

The main objective of the study was to investigate the need for adapting a multidimensional evaluation approach for desktop applications to the context of multimodal devices and applications. Thus, the influence of (i) the context (field, mobility use), (ii) the interaction modality, and (iii) the user experience on the evaluation of multimodal devices and applications was investigated.

Mobile devices, e.g. smartphones and PDAs, can be used to implement efficient speech-based and multimodal interfaces. With the aim of investigate the influence of the context of use on usability with regard to environmental noise, three experiments were carried out in the field, for three realistic environments - *quiet*, *normal* and *noisy*. Thirty-six subjects took part in the experiment. The subjects were divided into three groups of twelve. The multimodal device chosen as the target for this study was the Smartphone *HP iPAQ 910c Bussiness Messenger* and some of its native applications.

Experiment Design

An experimental plan was drawn from the objectives of the study, and three independent variables were chosen – (i) *task context*, (ii) *user mobility*, and (iii) *user experience level* - which are not influenced by the context, the test facilitator or external factors, such as lighting, activity flow and interruptions of people. Eight dependent variables were chosen: (i) *task time*, (ii) *number of incorrect choices*, (iii) *number of incorrect actions* (incorrect decisions excluding the number of incorrect choices while performing a task), (iv) *number of repeated errors,* (v) *number of incorrect voice recognition*, (vi) *number of accesses to the help* (online help and printed one), (vii) *perceived usefulness*, and (viii) *perceived ease of use*.

Test Environment, Materials and Participants

In all test environments, all the elements (e.g., tasks, informal think aloud, unstructured interviews) were identical, only the test environment was different with regard to background noise. To minimize moderator bias, the tests were conducted with three experienced usability practitioners with 3 to 12 years of experience in usability testing. The instructions given to participants were predefined. All moderators participated in data gathering and data analysis. Statistical analysis was performed by one moderator, and revised by another one. The multimodal device used in this experiment was the *HP iPAQ 910c Bussiness Messenger* and some of its native applications. A micro-camera connected to a transmitter was coupled to the device to remotely record and transmit user-device interaction data to the lab through a wireless connection (see Fig. 1).

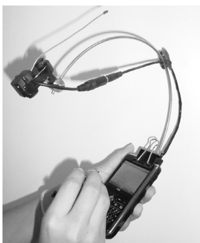

Fig. 1. Kit to support video camera during experiment

A remote screen capture software (*VNC*) was also used to take test session screenshots, and a web tool named *WebQuest* [29] was used as well for supporting the user subjective satisfaction measurement. *WebQuest* supports the specialist during data collection, automatic score computation, performs statistical analysis, and generates graphical results. Currently *WebQuest* supports two questionnaires: (i) a pre test questionnaire, the *USer* (*User Sketcher*), conceived to raise the profile of the system users; and (ii) a post test questionnaire, the *USE* (*User Satisfaction Enquirer*), conceived to raise the user degree of satisfaction with the system. The experiment setting required a wireless network. The experiments in silent and normal environments were carried out in an uncontrolled space (a typical work room). The experiment in noisy environment was carried out in an open place, characterized by the constant flow of people (a mall). In both cases, the session moderators remained with the participants, in case any explanation on the test procedure was required. Participants were selected on the basis of having previous experience with mobile devices (such as mobile phones), computers, and the Internet. They were also required to have some familiarity with the English language, since this is the language adopted in the device's user interface and in its documentation. The recruited users were divided into two groups of 12 for the tests in the quiet, normal and noisy environments (sensitive, normal and accurate sensitivities). According to their experience levels, each group was then subdivided into two subgroups. A ratio of 6 beginners to 6 intermediates was adopted.

Experimental Procedure

Observation and retrospective audio/video analysis for quantitative and qualitative data were employed. Participants were required to provide written consent to be filmed prior to, during and immediately after the test sessions, and to permit the use of their images/sound for research purposes without limitation or additional compensation. On the other hand, the evaluation team was committed to do not disclose the user performance or other personal information. According to the approach basis, the first step consisted in defining the evaluation scope for the product as well as in designing a test task scenario, in which the target problems addressed were related to the: (i) usage and interpretation of modalities; (ii) mechanisms for information input/output;

(iii) processing power; (iv) navigation between functions; and (v) information legibility. Since the test objectives focused on (i) investigating the target problems; and (ii) detecting other problems which affect usability, a basic but representative set of test tasks was selected and implemented. The test tasks consisted of (i) initializing the device; (ii) consulting and scheduling appointments; (iii) entering textual information; (iv) using the e-mail; (v) making a phone call; e (vi) using the audio player. After planning, 3 pilot field tests (*quiet*, *normal* and *noisy* environments) were conducted to verify the adequacy of the experimental procedure, materials, and environments. Aiming to prevent user tiredness, the session time was limited to 60 minutes, and the test scenario was dimensioned to 6 tasks. Thus, each test session consisted of (i) introducing the user to the test environment by explaining the test purpose and procedure; (ii) applying the pre-test questionnaire (*USer*); (iii) performing the six-task script; (iv) applying the post-test questionnaire (*USE*); and (v) performing a non-structured interview. For the participants who declared not having had any previous contact with the smartphone device, a brief explanation of the device I/O modes and its main resources was given, considering that it was not yet widely spread in Brazil at the time of the experiment.

5 Results

For summarizing conformity assessment results (see Table 1), it was computed *Adherence Ratings* (*AR*), which are the percentages of the *Applicable recommendations* (*Ar*) that were *Successfully adhered* to (*Sar*) [19]. All the ARs, excluding that one related to ISO 14745, are higher than 75%, which means successful results. As for ISO 14745, the result indicates the need to improve the input text via write recognition. Those results corroborate with the idea that the efficacy of standards inspection can be considerably improved if it is based upon standards conceived specifically for mobile devices, which could evidence more usability problems.

As for the user subjective satisfaction measurement, both questions and answers of the post test questionnaire (*USE*) were previously configured. The questionnaire was applied soon after the usability test and answered using the mobile device itself, with the purpose to collect information on the user degree of satisfaction with the device by means of 38 questions about menu items, navigation cues, understandability of the messages, ease of use functions, I/O mechanisms, online help and printed manuals,

Table 1. HP iPAQ 910c conformity with standards

Standard	#Sar	#Ar	AR(%)
ISO 9241 Part 14	66,0	72,0	91,67
ISO 9241 Part 16	31,0	34,0	91,18
ISO 9241 Part 17	54,0	57,0	94,74
ISO 14754	7,0*	11,0	63,64
ISO 24755	10,0**	17,0	58,82

* Not adhered recommendations in full, made small changes consistently in 4/7 points (Transcriber), 1/3 points (Letter recognizer) and 2/4 points (Block recognizer).

** Not adhered recommendations in full, made small changes consistently in 6 points.

Table 2. Overlay of results obtained from different techniques described above

PROBLEM CATEGORY	SI	PM	SM
Location and sequence of menu options	✓ (2)		✓ (2)
Menu navigation		✓ (1)	
Presentation of menu options	✓ (1)		✓ (1)
Information feedback		✓ (3)	× (3)
Object manipulation	✓ (4)		
Form filling dialogues	✓ (3)		
Symbols and icons	✓ (7)		× (3)
Text entry via keyboard		✓ (1)	✓ (1)
Text entry via virtual keyboard		✓ (1)	✓ (1)
Text entry via stylus (Transcriber)	✓ (4)	✓ (1)	✓ (5)
Text entry via stylus (Letter recognizer)	✓ (8)		✓ (8)
Text entry via stylus (Block recognizer)	✓ (7)		✓ (7)
Voice recognizer	✓ (2)	✓ (1)	✓ (3)
Processing power		✓ (1)	✓ (1)
Hardware issues		✓ (3)	✓ (3)
Fluent tasks execution		✓ (5)	✓ (5)
Online and offline help		✓ (2)	✓ (2)
Legend: SI – Standards Inspection. PM – Performance Measurement. SM – Satisfaction Measurement.	✓ Consistent findings. × Contradictory findings.		

users' impression, and product acceptance level. With the support of the pre test questionnaire (*USer*), the user sample profile was drawn. It was composed of 20 male and 16 female users, of which 12 were *undergraduate students*, 12 *post-graduate students*, five *graduate level*, and seven *post-graduate level*. The ages varied between *18* and *55* years. They were all *right handed* and mostly (18) used some sort of reading aid (*glasses* or *contact lenses)*. They also had familiarity with the English language. All of them had at least one year of *previous experience of computer systems*, and were currently using computers on a *daily* basis.

The ranges for USE normalized user satisfaction are 0.67 to 1.00 (*Extremely Satisfied*), 0.33 to 0.66 (*Very satisfied*), 0.01 to 0.32 (*Fairly satisfied*), 0.00 (*Neither satisfied nor unsatisfied*), 0.01 to -0.32 (*Fairly dissatisfied*), -0.33 to -0.66 (*Very dissatisfied*), and -0.67 to -1.00 (*Extremely dissatisfied*) [27]. The normalized user satisfaction levels achieved were 0.25 (*Fairly satisfied*), 0.32 (*Fairly satisfied*), and 0.21 (*Fairly satisfied*) for the quiet, normal, and noisy environments, respectively.

Since the multidimensional approach is based upon the triangulation of results, Table 2 summarizes the usability problem categories which were identified during the evaluation process. For each category, the number of problems identified by each technique is given. As can be seen, some of the usability problem categories were more associated to the performance measurement (e.g. hardware aspects, help mechanisms) whereas others (e.g. menu navigation, presentation of menu options) were identified by the conformity assessment.

It was possible to identify 67.6% of the problems found by other methods when combining the results from the post-test questionnaire to the comments made during the test sessions. In addition, the non-structured interviews at the end of each experiment session showed that the user opinion was in agreement (e.g., location and sequence of menu options) or disagreement (e.g., menu navigation) with the results obtained from the other two evaluation techniques. This discrepancy can originate from the users' perception of product quality, and from the perception of their own skills to perform the task.

The statistic analysis consisted of: (1) building a report with univariate statistics; (2) generating the covariance matrices for the predefined objective and subjective indicators; (3) applying the one-way F ANOVA test to the data obtained from the previous step in order to investigate possible differences; and (4) applying the Tukey-Kramer process to the one-way F ANOVA results aiming to investigate if the found differences were statistically significant to support inferences from the selected sample.

According to the results (see Table 3), the series of one-factor ANOVA showed that exists significant differences between the session's test results realized in the silent, normal and noisy environments. The Tukey-Kramer process (see Table 4) shows that the main differences are found in the comparison between the normal and noisy environments and quiet and noisy environments.

Table 3. *Quiet x Normal x Noisy* experiment results

Variable Pair	p-Value (α=0.05)		
	Quiet	Normal	Noisy
Experience x Task Time	0.226	0.007	0.146
Experience x Incorrect Actions	0.078	0.124	0.040
Experience x Incorrect Choices	0.191	0.121	0.080
Experience x Repeated Errors	0.640	0.515	0.078
Experience x Help Accesses	0.016	0.282	0.059

The results obtained from the experiment in which the multidimensional approach was applied support the original assumption that, in spite of the distinctive features of this class of devices, it is possible to adapt from the evaluation experience with conventional and multimodal devices.

Table 4. Tukey-Kramer experiment results

Variable Pair	Quiet x Normal	Quiet x Noisy	Normal x Noisy
Experience x Task Time	Not different	Not different	Different
Experience x Incorrect Actions	Different	Not different	Different
Experience x Incorrect Choices	Not different	Not different	Different
Experience x Repeated Errors	Not different	Not different	Not different
Experience x Help Accesses	Not different	Different	Not different
Experience x Incorrect Voice Recognition	Not different	Different	Different

6 Final Considerations

From the data analysis it became evident that certain problem categories are better found by specific techniques, as shown in Table 2. The analysis of the pre-test and post-test questionnaires and the informal interviews showed that familiarity with the English language, domain knowledge and computer literacy, which characterize the user experience level, have significant influence on the user performance with multimodal devices and applications. Physical characteristics of the device affected the user performance and the data gathering during the experiment. Outdoors, in ambient light and higher noise level conditions, the screen's legibility was reduced as well as the accuracy level of the voice recognizer was affected. According to the user's opinion stated during the informal interview, the kit to support video camera did not interfere with the task execution, but the majority decided to lay the device down during task execution. As for the entry of text information, the users showed a preference for the on-screen keyboard instead of virtual keyboard and handwriting recognition feature such as *Block Recognizer*, *Letter Recognizer*, or *Transcriber* to enter text. Based on their comments, as well as on the informal interview, it was concluded that writing long messages is very cumbersome both using the virtual keyboard and using the handwriting recognition application.

The individual user preference was for the modality touchscreen combined with graphical interaction, aided by the use of the stylus. Handwriting recognition and voice recognition features had influence in the test results, especially regarding to the time of completion of the task.

From the application of the multi-layered approach, the data gathered and analyzed support the initial assumption that minor adaptations in the traditional evaluation techniques and respective settings are adequate to accommodate the evaluation of the category of multimodal devices and application targeted by this study.

References

1. Nielsen, J.: Usability engineering. Academic Press, Boston (1993)
2. ISO 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs)—Part 11: Guidance on usability. International Organization for Standardization, Geneva, Switzerland (1998)
3. Turnell, M.F.Q.V., de Queiroz, J.E.R., Ferreira, D.S.: Multilayered Approach to Evaluate Mobile User Interfaces. In: Lumsden, J. (ed.) Handbook of Research on User Interface Design and Evaluation for Mobile Technology, IGI Global, vol. 1, pp. 847–862 (2008)
4. Jöst, M., Haubler, J., Merdes, M., Malaka, R.: Multimodal interaction for pedestrians: an evaluation study. In: Riedl, J., Jameson, A., Billsus, D., Lau, T. (eds.) Proc. IUI 2005, San Diego, pp. 59–66. ACM Press, New York (2005)
5. Blandford, A., Hyde, J.K., Curzon, P., Papatzanis, G.: EMU in the car: evaluating multimodal usability of a satellite navigation system. In: Graham, N., Palanque, P. (eds.) DSVIS 2008, pp. 1–14. Springer, Heidelberg (2008)
6. Taib, R., Ruiz, N.: Evaluating Tangible Objects for Multimodal Interaction Design. In: Proc. OZCHI 2005, Narrabundah, Australia, pp. 1–4 (2005)
7. Bach, C., Scapin, D.: Ergonomic criteria adapted to human virtual environment interaction. In: Rauterberg, M., Menozzi, M., Wesson, J. (eds.) Proc. INTERACT 2003, Zurich, Switzerland, pp. 880–883. IOS Press, Amsterdam (2003)
8. Kaster, T., Pfeiffer, M., Bauckhage, C.: Combining speech and haptics for intuitive and efficient navigation through image database. In: Oviatt, S. (ed.) Proc. ICMI 2003, pp. 180–187. ACM Press, New York (2003)
9. Trevisan, D.G., Nedel, L.P., Macq, B., Vanderdonckt, J.: Detecting interaction variables in a mixed reality system for maxillofacial-guided surgery. In: SBC Symposium on Virtual Reality (SVR 2006), Belém do Pará, pp. 39–50. SBC Press, Brazil (2006)
10. Paternò, F., Santos, I.: Designing and developing multi-user, multi-device web interfaces. In: Calvary, G., Vanderdonckt, J. (eds.) CADUI 2006. Kluwer Academics Publishers, Dordrecht (2006)

11. Bowman, D., Gabbard, J., Hix, D.: A survey of usability evaluation in virtual environments: Classification and comparison of methods. Presence: Teleoperators and Virtual Environments 11(4), 404–424 (2002)
12. Nedel, L., Freitas, C.M.D.S., Jacob, L., Pimenta, M.: Testing the use of egocentric interactive techniques in immersive virtual environments. In: Rauterberg, M., Menozzi, M., Wesson, J. (eds.) Proc. INTERACT 2003, pp. 471–478. IOS Press, Zurich (2003)
13. Dybkjær, L., Bernsen, N.O., Minker, W.: Evaluation and usability of multimodal spoken language dialogue systems. Speech Communication 43, 33–54 (2004)
14. Baille, L., Schatz, R.: Exploring Multimodality in the Laboratory and the Field. In: Lazzari, G., Pianesi, P. (eds.) Proc. ICMI 2005, pp. 100–107. ACM Press, New York (2005)
15. Balbo, S., Coutaz, J., Salber, D.: Towards automatic evaluation of multimodal user interfaces. Intelligent User Interfaces, Knowledge-Based Systems 6(4), 267–274 (2003)
16. Holzapfel, H., Nickler, K., Stiefelhagen, R.: Implementation and evaluation of a constraint-based multimodal fusion system for speech and 3D pointing gesture. In: Sharma, R., Darrell, T. (eds.) Proc. ICMI 2004, pp. 175–182. ACM Press, New York (2004)
17. De Queiroz, J.E.R.: Abordagem Híbrida para avaliação da usabilidade de interfaces com o usuário. Tese de Doutorado, UFPB, Brazil. 410 p. (in Portuguese) (2001)
18. Turnell, M.F.Q.V., de Queiroz, J.E.R., Ferreira, D.S.: Multilayered Approach to Evaluate Mobile User Interfaces. In: Lumsden, J. (ed.) Handbook of Research on User Interface Design and Evaluation for Mobile Technology, IGI Global, vol. 1, pp. 847–862 (2008)
19. ISO 9241-14, Ergonomic requirements for office work with visual display terminals (VDTs) —Part 14: Menu dialogues. ISO, Geneva, Switzerland (1997)
20. ISO 9241-16, Ergonomic requirements for office work with visual display terminals (VDTs)—Part 16: Direct manipulation dialogues. ISO, Geneva, Switzerland (1999)
21. IS0 9241-17, Ergonomic requirements for office work with visual display terminals (VDTs)—Part 17: Form filling dialogues. ISO, Geneva, Switzerland (1998)
22. ISO/IEC 14754, Information technology—pen-based interfaces—common gestures for text editing with pen-based systems. ISO, Geneva, Switzerland (1999)
23. ISO 9241-171, Ergonomics of human-system interaction – Part 171: Guidance on software accessibility. ISO, Geneva, Switzerland (2008)
24. ISO/IEC 24755, Screen icons and symbols for personal mobile communication device. ISO. Geneva, Switzerland (2007)
25. ISO/IEC 18021, Information technology— user interfaces for mobile tools for management of database communications in a client-server model. ISO, Geneva, Switzerland (2002)
26. ITU-T E.161. Arrangement of digits, letters and symbols on telephones and other devices that can be used for gaining access to a telephone network. International Telecommunications Union-telecommunications, Geneva, Switzerland (2001)
27. Bailey, J.E., Pearson, S.W.: Development of a Tool for Measuring and Analyzing Computer User Satisfaction. Management Science 29(5), 530–545 (1983)
28. Aladwani, A.M., Palvia, P.C.: Developing and validating an instrument for measuring user-perceived Web quality. Information & Management 39, 467–476 (2002)
29. De Oliveira, R.C.L., De Queiroz, J.E.R., Vieira Turnell, M.F.Q.: WebQuest: A configurable Web Tool to prospect the user profile and user subjective satisfaction. In: Salvendy, G. (ed.) Proc. HCI 2005, vol. 2. Lawrence Erlbaum, Mahwah (2005) (U.S. CD-ROM Multi Platform)