

Manipulação de Artefactos Moveis Multimodais: Avaliação e Estudos Comportamentais em Contextos Reais

Tiago Reis
treis@lasige.di.fc.ul.pt

Luís Carriço
lmc@di.fc.ul.pt

Marco de Sá
marcosa@di.fc.ul.pt

LaSIGE, Departamento de Informática da Faculdade de Ciências da Universidade de Lisboa.

Sumário

Tanto a avaliação de aplicações móveis, como estudos realizados em torno das mesmas devem ser conduzidos tendo em consideração o propósito das aplicações e os contextos nos quais serão utilizadas. Neste artigo apresentamos uma ferramenta de manipulação de artefactos móveis multimodais, a sua avaliação e um estudo comportamental conduzido com a mesma. Tanto a avaliação como o estudo foram conduzidos com um subconjunto de utilizadores finais da aplicação (estudantes). Estes responderam a um questionário, utilizando a ferramenta, em contextos reais (sala de aula, casa, parque, metro e carro) definidos por diferentes variáveis de contexto (dispositivo, ruído, iluminação, movimento, posição, utilização das mãos, tipo de conteúdos, limitações temporais, etc.). A avaliação apontou ao registo dos problemas de usabilidade da ferramenta, ao desempenho das diferentes modalidades nos diferentes contextos considerados e à aceitação da aplicação por parte dos utilizadores finais. O estudo permitiu-nos perceber que a forma como os utilizadores interagem com aplicações móveis varia de acordo com o contexto no qual estão inseridos. Identificamos padrões comportamentais, percebendo o impacto de algumas variáveis de contexto na escolha das modalidades de interacção (voz, gestos, teclado, ecrã táctil).

Palavras-chave

Interacção Multimodal, Dispositivos Moveis, Avaliação, Estudos Comportamentais, Contextos Reais

1. INTRODUÇÃO

Interacção multimodal é um conceito inerente às actividades quotidianas do ser humano. Nestas: falamos, ouvimos, visualizamos, tocamos, apontamos, escrevemos e desenhamos, alternada ou simultaneamente, de forma a atingir a um objectivo. Este tipo de interacção considera os diferentes canais da percepção humana (sentidos), introduzindo elementos naturais de comportamento em interfaces pessoa-máquina [Turk00].

Os interfaces multimodais podem tornar as aplicações informáticas universalmente acessíveis através da consideração de diferentes tipos de utilizador [Nicolle01] e contextos de utilização [Sá08, Hurtig06]. Estes possibilitam interacção pessoa-máquina a utilizadores com deficiências [Boyd90, Burger93, Blenkhorn98] e a utilizadores sem deficiências em situações sub-óptimas [Goodman04]. São particularmente úteis em ambientes moveis, nos quais existe uma maior variação do contexto de utilização, aumentam a acessibilidade, desempenho, estabilidade, robustez e eficiência das actividades suportadas por computadores e/ou dispositivos moveis [Oviat04, Lai04]. Este tipo de interacção já foi estudado para propósitos que não estão directamente associados ao problema da

acessibilidade [Blattner96, Signer06, Santoro07]: aumento de actividades uni modais [Reis08], jogos, aplicações multimédia, etc. Destes, destacamos algumas linhas de investigação que tentaram perceber como ultrapassar as limitações introduzidas pelo tamanho limitado dos ecrãs de dispositivos móveis, através da inclusão de multimodalidades [Brewster02, Brewster03].

As modalidades de interacção (entrada/saída) disponíveis numa aplicação multimodal podem ser utilizadas de três formas distintas [Oviat99]:

1) Complementar: a interacção é conseguida através de um conjunto de modalidades que se complementam representando, em conjunto, uma acção do utilizador (e.g.: input de voz e toque: dizer: “como vou para”; tocar: num local de um mapa apresentado no ecrã”; output: caminho mais curto é desenhado no mapa, e indicações de voz guiam o utilizador enquanto este conduz);

2) Redundante: o mesmo tipo de interacção é possível através de mais que uma modalidade (e.g.: um questionário onde o utilizador pode ver e ouvir as perguntas, e respondê-las através de escrita no teclado, reconhecimento de escrita no ecrã, ou por voz).

3) Alternativa: a interacção é conseguida através de uma modalidade que serve de alternativa a outra, inibida por qualquer razão (e.g.: um questionário onde o utilizador só pode ver as perguntas, e respondê-las através de escrita no teclado, porque o dispositivo não tem periféricos de áudio ou ecrã táctil).

No que respeita a avaliação, de aplicações móveis multimodais, são consideradas duas alternativas: avaliação em laboratório [Beck03] ou no terreno [Kjeldskov04, Make-la00]. Ambas podem considerar o registo das actividades dos utilizadores através de logs ou filmagens, e as suas opiniões através de questionários ou entrevistas. Alguns dos estudos disponíveis na bibliografia actual foram direccionados ao registo de informação sobre dispositivos e aplicações já em utilização, outros à fase de avaliação das aplicações. Realçam-se estudos que consideraram os dois tipos de avaliação [Baillie05], comparando-os e sugerindo que a avaliação de aplicações móveis dever ser feita no terreno, em contextos reais de aplicação do sistema em avaliação, com utilizadores finais da mesma.

Este artigo é iniciado com uma apresentação de uma plataforma de criação, manipulação e análise de artefactos multimodais móveis (MAFra). De seguida, os artefactos, ferramentas de criação e análise são descritos em linhas gerais, ao passo que a ferramenta de manipulação, foco da avaliação e estudos apresentados neste artigo, é descrita em detalhe, com especial atenção às diferentes modalidades de interacção e modos de execução disponíveis. Apresentamos e discutimos uma avaliação e um estudo comportamental conduzidos com um subconjunto de utilizadores finais desta aplicação, em contextos reais (sala de aula, casa, parque, metro, carro), definidos por diferentes variáveis (dispositivo, ruído, iluminação, movimento, posição, utilização das mãos, tipo de conteúdos, limitações temporais). A avaliação apontou principalmente à identificação de problemas de usabilidade, ao registo do desempenho das diferentes modalidades de interacção em diferentes contextos e à aceitação final da aplicação por parte dos utilizadores. O estudo teve como objectivo perceber padrões comportamentais dos utilizadores no que respeita a escolha de modalidades de interacção em diferentes contextos.

2. PLATAFORMA MOVEL DE ARTEFACTOS MULTIMODAIS

Esta secção apresenta, em linhas gerais, uma plataforma móvel para criação, manipulação e análise de artefactos multimodais: MAFra. Os artefactos e as ferramentas de criação e análise são descritos nesta secção. A ferramenta de manipulação, foco deste artigo, é apresentada detalhadamente na próxima.

Todas as ferramentas e bibliotecas desta plataforma foram desenvolvidas em C# e são compatíveis com Microsoft Windows: XP, Vista e Mobile 5. Consequentemente, estão disponíveis versões para computadores pessoais e dispositivos móveis, sendo as últimas, especialmente no caso da ferramenta de criação, versões simplificadas das primeiras.

2.1 Artefactos Multi Modais Proactivos

Nesta plataforma, um artefacto é uma entidade abstracta composta por páginas e regras. As páginas contêm elementos de diferentes tipos: text/áudio/vídeo labels; text/áudio/image multiple choice objects; e, text/áudio/vídeo answer/recording objects. Cada elemento combina uma componente de apresentação gráfica e audível. Ambas consideram o tipo de elemento (e.g.: elemento: múltipla escolha; gráfica: drop down menu; audível: descrição sobre como interagir com o elemento) e o seu conteúdo (e.g.: elemento: múltipla escolha; conteúdo: as diferentes escolhas nas suas formas gráficas e audíveis). O conteúdo de um elemento pode ser estático ou dinâmico. Regra geral, é estático e definido a quando da criação do elemento (e.g.: conteúdo de um text label). Não obstante, pode ser dinâmico e gerado a quando da manipulação do artefacto (e.g.: o texto/áudio inserido numa entrada de texto/áudio). Nestes casos o conteúdo audível do elemento é gerado a partir do texto inserido (através de um motor de sintetização de discurso), ou introduzido pelo utilizador (através do microfone do dispositivo).

As regras (e.g.: if answer is YES goto page Y), componentes que possibilitam a proactividade dos artefactos, permitem a modificação das características dos mesmos: páginas, elementos e, principalmente, navegação. Baseiam-se em eventos (e.g.: if X minutes have passed, if Next is pressed) e condições (USER answered YES or BLUE) que activam comportamentos da aplicação (hide element Y on page X, goto page Z).

2.2 Ferramentas de Criação e Análise

Alem da ferramenta de manipulação, esta plataforma considera duas outras ferramentas:

Uma ferramenta de criação (wizard based), permite a construção de artefactos em três passos simples: 1) criação de páginas; 2) definição da ordem das páginas; 3) definição de regras. Este processo pode ser conduzido por alertas e sugestões automáticas, permite o uso de templates para páginas, sequencias e regras.

Uma ferramenta de análise possibilita a avaliação da manipulação dos artefactos, podendo funcionar como visualizador de resultados, ou como leitor de logs, disponibilizando a reprodução integral da interacção do utilizador com o artefacto. Esta ferramenta teve um papel preponderante na avaliação e estudos descritos neste artigo.

3. FERRAMENTA DE MANIPULAÇÃO

Esta plataforma disponibiliza também uma ferramenta multimodal móvel que permite a manipulação de artefactos (Fig 1). Esta ferramenta regista em ficheiros XML toda a informação introduzida pelo utilizador e todas as suas acções. Cada entrada deste ficheiro contém informação temporal sobre o instante da interacção.

Estão disponíveis 5 modalidades de interacção: Interacção Directa (ID), Barra de Interacção (BI), Teclado do

Dispositivo (T), Reconhecimento de Gestos (RG) e Reconhecimento de Voz (RV).

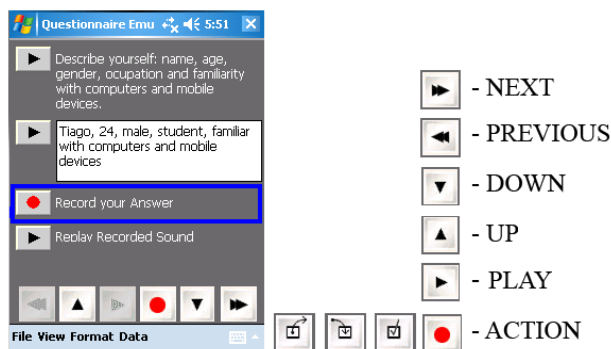


Fig. 1. Ferramenta de Manipulação de Artefactos.

A introdução de dados de conteúdo por parte do utilizador pode ser feita de duas formas: através do teclado do dispositivo, ou do microfone.

Estas modalidades foram introduzidas considerando diversas características intrínsecas aos dispositivos móveis (e.g.: ecrãs tácteis/inactivos, teclados físicos/virtuais), contextos nos quais são utilizados (e.g.: salas de aula, rua, jardins, metro) e formas de interacção com o dispositivo (e.g.: a forma como se agarra o dispositivo e como se interage com o mesmo).

Este artigo foca-se nas cinco modalidades de interacção disponíveis nesta ferramenta, algumas das quais (ID e BI) já avaliadas (através de protótipos de baixa e alta fidelidade) [Reis08] por utilizadores finais nas fases iniciais do desenho desta aplicação.

Esta ferramenta permite, na sua versão actual, a utilização das cinco modalidades de interacção de forma complementar e alternativa. Quanto à introdução de conteúdos textuais ou audíveis por parte dos utilizadores são consideradas todas as formas de utilização das modalidades: complementar, redundante e alternativa.

3.1 Modalidades de Interacção

3.1.1 Interacção Directa (ID):

Esta modalidade é direccionada a dispositivos com ecrã táctil, caneta ou rato: a interacção é efectuada directamente nos elementos disponíveis no ecrã. Estes devem apresentar a sua componente gráfica (composta pelo elemento e por um botão que permite ouvir o seu conteúdo). Esta é uma modalidade de interacção incompleta visto que a navegação entre páginas e elementos é conseguida de forma alternativa, através de uma das modalidades apresentadas de seguida. Não obstante, alguns artefactos podem permitir a navegação entre páginas de forma explícita (e.g. um botão) ou implícita (e.g. através de uma regra).

3.1.2 Barra de Interacção (BI)

Esta barra é composta por seis botões. Geralmente, aparece na parte inferior do artefacto (Fig. 1). Os botões

“Next” e “Previous” permitem a navegação entre páginas. A navegação através dos elementos da mesma página é feita usando os botões “Up” e “Down”. O botão “Play” reproduz a forma audível do conteúdo do elemento seleccionado (dentro de uma caixa azul, Fig. 1) e o botão “Action” muda de acordo com o elemento seleccionado, permitindo alterar o seu conteúdo servindo para: iniciar/parar gravações; seleccionar escolhas em listas; ou, mais complexo, iniciar/terminar navegação através dos componentes de um elemento: depois de iniciada, esta navegação permite: a selecção horizontal de componentes de um elemento através dos botões “Next e Previous” (e.g. numa trackbar), ou a selecção vertical através dos botões “Up” e “Down” (e.g. numa combo box). Uma vez cessada a navegação entre componentes de um elemento, os botões voltam todos a ter a sua função inicial/principal.

O conjunto dos comandos aqui apresentados constitui a lista de funcionalidades disponíveis para interagir com um artefacto (excepto para a introdução de conteúdos textuais).

3.1.3 Teclado (T)

Esta modalidade é construída com base nas funcionalidades supramencionadas, mapeando-as para o teclado do dispositivo. Este mapeamento pode, ou não, ser completo, de acordo com o número de teclas disponíveis no teclado. Na maior parte dos dispositivos, que têm teclados de navegação, este mapeamento é: “Previous” – seta para a esquerda, “Next” – seta para a direita, “Up” – seta para cima, “Down” seta para baixo, “Play” – tecla Enter, “Action” – Aguentar a tecla Enter premida mais que 1 segundo.

3.1.4 Reconhecimento de Gestos (RG)

À semelhança da modalidade anterior, também esta é construída com base nas funcionalidades referidas, mapeando-as para gestos (Fig 2) que são reconhecidos no ecrã táctil do dispositivo. Para disponibilizar este tipo de interacção foi-nos necessário construir um algoritmo de reconhecimento de gestos que permite a interpretação dos 6 gestos apresentados a baixo:

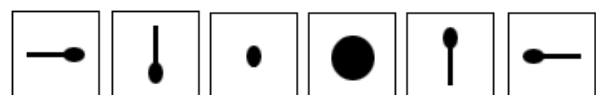


Fig. 2. Gestos reconhecidos pela aplicação (num ecrã táctil)

Da esquerda para a direita: “Previous”, “Up”, “Play”, “Action”, “Down”, “Next”. O círculo pequeno representa o primeiro toque no ecrã e a linha um gesto contínuo que se lhe segue. O círculo grande representa um toque que dura mais do que um segundo (sem gestos).

3.1.5 Reconhecimento de Voz (RV)

Esta modalidade é construída com base nas funcionalidades anteriormente mencionadas, mapeando-as para comandos de voz: “Previous”, “Up”, “Play”, “Action”, “Down”, “Next”.

As modalidades baseadas em reconhecimento, tanto de voz como de gestos foram avaliadas, per-se, em laboratório, demonstrando taxas de sucesso próximas de 100%.

3.2 Modos de Utilização Predefinidos

Esta aplicação pode correr em diferentes modos, que podem ser combinados entre si: visual, eyes-free e hands-free.

3.2.1 Modo Visual:

Por omissão, o modo visual ignora a componente audível dos elementos durante a interacção com o artefacto. No entanto, neste modo, a interacção pode ser suportada por qualquer uma das modalidades disponíveis.

3.2.2 Modo Eyes-Free:

O modo eyes free é direccionado para situações em que, por qualquer razão, o utilizador não pode despendar, muita ou nenhuma atenção visual ao dispositivo. Este modo considera interacção por: RV, RG, T ou cartão táctil. Este cartão é direccionado para dispositivos com ecrã táctil e sem teclado, sendo colocado no topo do ecrã, mapeando um teclado T9 (para input textual e numérico) e uma versão simplificada da BI. A barra foi simplificada devido às limitações espaciais inerentes aos ecrãs dos dispositivos móveis. Estão disponíveis os botões: “Up”, “Down”, “Play”, “Action”, e a navegação entre páginas é conseguida através dos botões “Up” (“Previous”) e “Down” (“Next”), premindo-os um segundo ou mais.



Fig. 3. Cartão Táctil Para Dispositivos sem Teclado.

Durante a navegação entre páginas, os utilizadores são informados, via áudio, do número de sequência da página actual. Aquando da navegação entre elementos o utilizador é informado, via áudio, do tipo de elemento seleccionado e sobre como interagir com ele. Esta informação pode ser saltada através “Play”, “Down”, “Up”, “Next” ou “Previous”.

Geralmente, este modo também considera ID e a parte visual da aplicação está disponível (excepto quando utilizando o cartão táctil) contemplando situações em que o utilizador pode dispensar, esporadicamente, atenção visual ao dispositivo.

3.2.3 Modo Hands-Free:

O modo hands-free considera interacção por voz para entrada (reconhecimento e gravação) e qualquer das modalidades disponíveis, ou combinações das mesmas, para saída.

4. AVALIAÇÃO E ESTUDO COMPORTAMENTAL

Tanto a avaliação da ferramenta de manipulação como o estudo comportamental apresentados nesta secção foram conduzidos em Lisboa (Portugal), envolvendo 12 utilizadores sem deficiências: 6 homens, 6 mulheres, com idades compreendidas entre os 20 e os 35 anos, estudantes, habituados a utilizar computadores e dispositivos móveis.

A avaliação teve como objectivos principais a análise de: 1) possíveis problemas de usabilidade; 2) diferenças de desempenho quando interagindo pelas diferentes modalidades de interacção disponíveis; 3) aceitação dos utilizadores finais.

O estudo comportamental apontou à compreensão do impacto dos diferentes contextos considerados e respectivas variáveis de contexto envolvidas, na escolha das modalidades de interacção por parte do utilizador.



Fig. 4. Avaliação e estudos comportamentais em contextos reais – em cima: casa, parque, metro; em baixo: carro (utilizador a conduzir e Tablet PC no banco ao lado).

Em ambos, os utilizadores manipularam um questionário multimodal formado por 6 questões, 5 questões abertas, que podiam ser respondidas por input textual ou por gravação áudio, e uma questão fechada, de múltipla escolha (6 escolhas). As questões são apresentadas abaixo:

- 1) Descreva-se: nome, idade, género, profissão, familiaridade com computadores e dispositivos moveis.
- 2) Qual é o seu numero de telefone?
- 3) Qual é a sua morada?
- 4) Quais foram as dificuldades que encontrou até chegar a esta pergunta?
- 5) Acha que esta aplicação podia ser usada para fazer trabalhos de casa ou exames escolares?
- 6) Quão difícil lhe foi levar a cabo esta tarefa? (escolhas: impossível, muito difícil, difícil, normal, fácil, muito fácil)

Este questionário teve 2 papéis, foi o artefacto sob avaliação e ao mesmo tempo reuniu informação sobre o perfil e opiniões dos utilizadores.

A tarefa levada a cabo pelos utilizadores consistiu em responder às 6 perguntas do questionário, e, de seguida,

alterar as respostas a 2 destas perguntas. Por cada pergunta respondida/alterada com sucesso, numa só tentativa, foi creditado um ponto; em duas tentativas meio ponto; e, em qualquer outro caso 0 pontos.

Todas as tarefas executadas pelos utilizadores foram filmadas e registadas em logs, permitindo assim a análise de informação acerca das intenções dos utilizadores, do comportamento da aplicação, do tempo consumido em cada tarefa e das dificuldades experienciadas pelos utilizadores.

4.1 Avaliação de Usabilidade

A avaliação da ferramenta de manipulação de artefactos apontou principalmente à identificação de possíveis problemas de usabilidade. Não obstante, outros 2 factores de interesse foram considerados, estendendo a lista de objectivos:

- 1) Problemas de usabilidade da ferramenta;
- 2) Variações no desempenho da ferramenta (tempo e sucesso de execução de uma tarefa), a quando da interacção através das diferentes modalidades disponíveis;
- 3) Aceitação da aplicação por parte de um grupo de utilizadores finais desta aplicação (estudantes);

Os 12 utilizadores envolvidos foram divididos em 2 grupos de 6: um usando Tablet PCs e o outro PDAs. Cada utilizador respondeu a um questionário através de cada uma das modalidades de interacção disponíveis no seu dispositivo, uma vez que estas estão intrinsecamente associadas ao hardware disponível no dispositivo hospedeiro (ID, BI, RG e RV no Tablet PC; e, ID, BI, RG e T no PDA).

Foram considerados 2 contextos, caracterizados por diferentes variáveis, como por exemplo: dispositivo, ruído, iluminação e movimento. Tanto a ordem das perguntas como a ordem de utilização das modalidades de interacção foram definidas aleatoriamente em cada uma das tarefas levadas a cabo, de forma a minimizar o impacto da aprendizagem dos utilizadores nos resultados finais da avaliação. Metade dos utilizadores de cada grupo (3 utilizadores) responderam ao questionário numa sala de aula (sentados, em silêncio num local bem iluminado), a outra metade fê-lo num parque (andando, num ambiente ruidoso e com níveis de iluminação que dificultavam a interpretação da informação visual apresentada no ecrã).

4.1.1 Resultados:

	Modalidades de Interação	Tempo Médio	Classificação Média
Contexto 1: Sala de Aula	ID	0:58 min	100%
	BI	1:30 min	100%
	RG	1:38 min	100%
	RV	1:42 min	100%
Contexto 2:	ID	2:01 min	100%
	BI	1:58 min	100%

Parque (andando)	RG	2:00 min	100%
	RV	4:30 min	55%

Tabela 2. Primeiro grupo de utilizadores (Tablet PC). Resultados da avaliação das diferentes modalidades em diferentes contextos.

	Modalidades de Interação	Tempo Médio	Classificação Média
Contexto 1: Sala de Aula	ID	1:50 min	100%
	BI	2:40 min	100%
	RG	2:48 min	100%
	T	2:45 min	100%
Contexto 2: Parque (andando)	ID	2:55 min	100%
	BI	2:60 min	100%
	RG	2:58 min	100%
	T	2:45 min	100%

Tabela 3. Segundo grupo de utilizadores (PDA). Resultados da avaliação das diferentes modalidades em diferentes contextos.

4.1.2 Discussão:

No desenvolvimento desta aplicação foi empregada uma metodologia de desenho centrada no utilizador [Reis08]. Esta considerou a avaliação de protótipos de baixa e alta fidelidade, por parte de um grupo de utilizadores finais, identificando-se assim a maior parte dos problemas de usabilidade nas fases iniciais do processo de desenvolvimento.

Não obstante, o tempo dispendido na execução das tarefas foi consideravelmente maior (chegando a diferenças de mais de 100%) quando utilizando PDAs, sugerindo que a versão desta aplicação para Windows Mobile 5 deverá ser optimizada. No entanto, as taxas de sucesso das tarefas foram de 100% para a maior parte dos testes. Geralmente, os utilizadores despenderam mais tempo em tarefas que consideravam movimento do que em tarefas estacionárias. No entanto, isto não foi verificado quando os utilizadores usaram o teclado do dispositivo, sugerindo que esta modalidade é a mais apropriada para situações de movimento.

As tarefas levadas a cabo em contextos definidos por níveis baixos de ruído e situações estacionárias (e.g. sala de aula) apresentaram taxas de sucesso de 100%. Não se verificaram diferenças substanciais no tempo dispendido a quando da interacção através de BI, RG, RV e T. No entanto, o contrário foi verificado a quando da utilização de ID, sugerindo que esta é a modalidade mais indicada para situações estacionárias.

As tarefas levadas a cabo em ambientes com ruído e níveis de iluminação que dificultam a visualização nos ecrãs dos dispositivos e situações de movimento (e.g.: parque) apresentaram taxas de sucesso de 100% para ID, BI e RG. Não se verificaram diferenças substanciais no tempo dispendido pelos utilizadores a interagindo através de ID, BI, e RG (no mesmo dispositivo). No entanto, as taxas de sucesso descem 45% quando utilizando reconhecimento de voz. Verificou-se que quando o ruído aumen-

ta o reconhecimento de voz interpreta comandos que não são verbalizados pelo utilizador, sugerindo a inibição desta modalidade em contextos com ruído.

As opiniões dos utilizadores revelam uma excelente aceitação da aplicação. Todos os utilizadores envolvidos nesta avaliação consideraram a aplicação muito fácil de utilizar através de todas as modalidades de interacção disponíveis, na maioria contextos. No entanto, os utilizadores manipulando o questionário por RV no parque, consideraram a aplicação muito difícil de utilizar em tais condições. Todos os utilizadores consideraram que a ferramenta é aplicável a tarefas educacionais como trabalhos de casa e exames escolares.

4.2 Estudo Comportamental

O preenchimento de um questionário é uma tarefa que pode ser levada a cabo em diferentes contextos, contextos estes que podem apresentar características distintas (e.g.: ruído, iluminação, movimento, posição, numero de pessoas presentes, limitações de tempo), que, por sua vez, podem ter impacto na selecção das modalidades de interacção por parte dos utilizadores. Por outro lado, esta selecção pode também ser influenciada pelo tipo de informação considerada no questionário (e.g.: privada ou não).

Depois de conduzida a avaliação apresentada na subsecção anterior, conduzimos, com os mesmos utilizadores, um estudo comportamental de forma a entendermos o impacto que algumas variáveis de contexto e o tipo de informação abordada num questionário têm na selecção das modalidades de interacção por parte dos utilizadores.

Uma vez que os utilizadores envolvidos já haviam aprendido a interagir com a aplicação através de todas as modalidades disponíveis, foi-lhes pedido que respondessem ao questionário mais 2 vezes, interagindo livremente com a ferramenta, no que respeita à escolha das modalidades de interacção. Este estudo permitiu-nos a identificar alguns padrões comportamentais gerados pelas variáveis de contexto e tipos de informação.

Dada a natureza móvel e o propósito da ferramenta e tarefas em causa, decidimos conduzir este estudo no terreno, considerando um conjunto de contextos reais, um pouco mais vasto do que o considerado na avaliação. Estes contextos foram escolhidos pelos utilizadores e pela nossa equipa: casa, parque, metro e carro (a conduzir).

4.2.1 Casa

O estudo conduzido neste cenário envolveu 2 utilizadores que responderam ao mesmo questionário 2 vezes. Os utilizadores encontravam-se sentados utilizando um Tablet PC num ambiente silencioso e bem iluminado.

Na primeira tarefa não foram considerados limites temporais. Os 2 utilizadores escolheram interagir com o questionário sem utilizarem as mãos, através de modalidades de voz, tanto no que respeita a navegação entre elementos e páginas, como a entrada de dados (reconhecimento de voz e gravação áudio). Os dois utilizadores conseguiram utilizar a ferramenta com sucesso e apesar da existência

de perguntas de cariz pessoal, ambos decidiram responder às perguntas abertas através de gravações áudio.

No final da tarefa, os utilizadores explicaram que tentaram explorar a modalidade de interacção que lhes pareceu mais interessante de um ponto de vista inovativo-tecnológico. As condições óptimas deste cenário e a ausência de limitações temporais na execução da tarefa parecem justificar esta decisão, assim sendo, decidimos levar a cabo uma segunda tarefa, desta vez, considerando limitações temporais de forma a que os utilizadores se empenhassem em utilizar a ferramenta da forma mais eficiente possível.

A quando da execução da segunda tarefa os utilizadores preferiram introduzir conteúdos sempre por interacção directa: usando o ecrã táctil para seleccionar escolhas múltiplas e para controlarem as gravações de áudio em respostas abertas. A navegação entre páginas foi feita por reconhecimento de voz e a navegação entre elementos não foi considerada.

Ambos utilizadores consideraram as duas tarefas fáceis.

4.2.2 Parque

O estudo conduzido neste cenário envolveu 4 utilizadores que responderam 2 vezes ao mesmo questionário.

Ambas tarefas consideraram limitações temporais e foram levadas a cabo num ambiente ruidoso, com níveis de iluminação que dificultavam a interpretação da informação visual apresentada no ecrã dos dispositivos.

Na primeira tarefa os utilizadores encontravam-se numa situação estacionária, sentados num banco do parque, a interagir com um TabletPC. Nesta tarefa os utilizadores ignoraram a navegação entre elementos e a modalidade de interacção escolhida para navegação entre páginas variou: 3 usaram RV e 1 RG. Mais uma vez a maior parte os utilizadores justificaram as suas escolhas com base em interesses inovativo-tecnológicos. No entanto, em ocasiões em que o reconhecimento de voz falhou 2 ou 3 vezes os utilizadores recorreram ao reconhecimento de gestos, utilizando a mesma justificação. No que respeita a entrada de dados em questões abertas os utilizadores gravaram as suas respostas às perguntas 1, 4, 5 e utilizaram o teclado virtual para responder às perguntas 2 e 3, justificando estas escolhas com base no facto de não se sentirem confortáveis em dizer o seu numero de telefone e morada em voz alta num local onde se encontravam mais pessoas, especialmente porque não conheciam estas pessoas. Já no que respeita as respostas à questão de múltipla escolha, todos os utilizadores usaram ID. Esta tarefa considerada fácil pelos 4 utilizadores envolvidos.

Na segunda tarefa, os utilizadores encontravam-se numa situação de movimento, andando no parque enquanto respondiam ao questionário utilizando um PDA. Nesta tarefa, as modalidades escolhidas para efectuar a navegação entre páginas e as justificações para estas escolhas foram as mesmas que na primeira tarefa. Não obstante, todos os utilizadores consideraram a navegação entre elementos, utilizando, para tal efeito, a modalidade escolhida para a efectuar a navegação entre páginas. Modali-

dade esta que serviu também para entrada de dados em questões fechadas. A modalidade eleita para entrada de dados em repostas abertas foi gravação áudio. No entanto, todos os utilizadores explicaram que não se sentiram a vontade enquanto respondiam às questões de cariz pessoal, e que numa situação da vida quotidiana o fariam noutra contexto, com menos pessoas envolvidas, ou por escrita no teclado virtual, mas numa situação estacionária. Todos os utilizadores consideraram esta tarefa normal, no que respeita o seu nível de dificuldade.

4.2.3 Metro

Este estudo envolveu 4 utilizadores que responderam a 2 vezes a um questionário usando PDAs.

Ambas tarefas foram levadas a cabo num ambiente muito ruidoso, bem iluminado, e com padrões de movimento muito característicos (travagens, acelerações, pessoas a entrarem e saírem do metro).

Na primeira tarefa os utilizadores encontravam-se sentados, a modalidade adoptada para efectuar a navegação entre páginas variou: 1 utilizador usou o teclado do dispositivo e os outros 3 usaram gestos. Os utilizadores justificaram as suas escolhas com base nas suas preferências tecnológicas. Quando confrontados com a possibilidade de utilizar o reconhecimento de voz neste contexto os utilizadores explicaram que não o fariam pelo facto de se sentirem envergonhados em falarem para um dispositivo dentro do metro, rodeados por desconhecidos. Pela mesma razão a entrada de dados em questões abertas foi feita através do teclado virtual. Nenhum dos utilizadores considerou a navegação entre elementos e todos usaram interacção directa para entrada de dados. Esta tarefa foi considerada muito fácil pelos 4 utilizadores envolvidos.



Fig. 5. Estudo Comportamental no Metro – Utilizadora de pé a responder a um questionário segurando e interagindo com o dispositivo utilizando só uma mão.

Na segunda tarefa os utilizadores encontravam-se de pé, segurando e interagindo com o dispositivo só com uma mão, enquanto se seguravam com a outra para manter o equilíbrio (Fig 5).

Tanto as modalidades eleitas para efectuar a navegação entre páginas e elementos, como as justificações para tais escolhas foram as mesmas que na primeira tarefa. A entrada de dados em perguntas abertas foi feita pelo teclado virtual. No entanto, os utilizadores explicaram

que tiveram muitas dificuldades em introduzir texto neste teclado utilizando os dedos.

Esta tarefa foi considerada difícil pelos 4 utilizadores envolvidos.

4.2.4 Carro (conduzindo)

O estudo conduzido neste cenário envolveu 2 utilizadores que preencheram um questionário sem considerar limitações de tempo. O ambiente envolvente era ruidoso e bem iluminado, os utilizadores estavam sentados a conduzir, interagindo com a ferramenta sem utilizarem as mãos nem os olhos, devido à localização do dispositivo que estava no banco do passageiro. Este foi o único cenário em que os utilizadores utilizaram a parte audível dos questionários. Tanto a navegação entre páginas e elementos como a entrada de dados foram feitas através de modalidades de voz (reconhecimento e gravação). Os utilizadores justificaram estas escolhas com base no facto de não poderem utilizar as mãos nem os olhos, visto que estavam a conduzir e que o dispositivo estava num local que dificultava olhar para ele. Ambos os utilizadores consideraram esta tarefa normal no que respeita a sua dificuldade.

4.2.5 Discussão

Ao longo dos estudos aqui apresentados, houve uma tendência por parte dos utilizadores em utilizarem as modalidades que lhes interessavam mais de um ponto de vista inovativo-tecnológico. A inclusão de limitações temporais fez com que os utilizadores só recorressem às suas modalidades preferidas em casos em que o desempenho destas fosse melhor ou igual ao de todas as outras, tentando otimizar a sua utilização da ferramenta. Começaram a utilizar as modalidades mais eficientes sempre que possível: navegando entre páginas através da sua modalidade preferida; ignorando a navegação entre elementos; interagindo directamente com os elementos em vez de indirectamente através da sua modalidade preferida; e efectuando a entrada de dados em questões abertas através de gravações áudio em vez do teclado virtual (exceptuando situações em que os utilizadores estavam rodeados por mais pessoas).

Em ambientes onde os utilizadores estavam sozinhos ou rodeados por poucas pessoas em locais abertos, o tipo de conteúdo das perguntas não ditou a escolha da modalidade para a resposta, indicando uma clara preferência para as gravações áudio e não para o teclado virtual. No entanto, há medida que o número de pessoas presentes no cenário aumentava, os utilizadores sentiam-se cada vez mais intimidados, inibindo o uso das modalidades de voz (favoritas pela maioria dos utilizadores na maioria dos contextos). Inicialmente houve uma redução na utilização de modalidades de voz quando os utilizadores, rodeados por poucas pessoas, respondiam a perguntas de cariz pessoal (como no exemplo do parque). Quando rodeados por mais pessoas, os utilizadores deixaram de recorrer às modalidades de voz por completo, justificando tal atitude com o facto de se sentirem envergonhados em falar para

um dispositivo à frente de desconhecidos (como no metro).

Mesmo quando o output visual não oferecia condições óptimas (no parque devido ao reflexo da luz solar) os utilizadores decidiram não utilizar a parte audível dos questionários. A única situação em que esta componente foi considerada foi no caso de estudo conduzido no carro, porque os utilizadores não conseguiam dispensar quase nenhuma atenção visual ao dispositivo.

As restrições na utilização das mãos por parte dos utilizadores (e.g.: no metro com os utilizadores de pé) tendem a introduzir dificuldade na interacção directa com os elementos (maioritariamente porque os utilizadores não conseguem alcançar com o polegar alguns dos elementos apresentados no ecrã) e porque escrever num teclado virtual utilizando os dedos foi considerado difícil por todos os utilizadores que tiveram de o fazer. Nestas situações, os utilizadores preferiram interagir indirectamente com os elementos através da sua modalidade preferida e inserir dados de conteúdo através de gravações áudio (excepto quando se encontravam rodeados por mais pessoas).

Ao longo dos estudos apresentados, houve situações em que, além dos utilizadores não conseguirem usar o teclado virtual sem alguma frustração, não se sentiam confortáveis em inserir dados através de gravações áudio (e.g.: andando no parque, e mais acentuadamente, em pé no metro). Nestas situações os utilizadores não conseguiram encontrar uma combinação de modalidades com a qual se sentissem confortáveis. Uma possível solução para este problema é a introdução de um teclado T9 (virtual ou físico) desenhado para interacção só com uma mão (e.g.: com botões maiores que o teclado virtual QWERTY).

5. CONCLUSÃO E TRABALHO FUTURO

Neste artigo apresentamos uma plataforma para criação, manipulação e análise de artefactos multimodais móveis. Explicamos a plataforma, artefactos e ferramentas de criação e análise em linhas gerais. Providenciamos uma descrição detalhada da ferramenta de manipulação, focando as diferentes modalidades de interacção disponíveis, bem como os diferentes modos de execução da aplicação.

Avaliamos a ferramenta de manipulação através de um processo, conduzido em contextos reais (sala de aula, parque) definidos por diferentes variáveis (dispositivo, ruído, iluminação, movimento), no qual um subconjunto de utilizadores finais da aplicação (estudantes) respondeu a um questionário através das diferentes modalidades de interacção disponíveis. Esta avaliação teve como objectivos perceber os problemas de usabilidade da aplicação, o desempenho das diferentes modalidades em diferentes contextos e a aceitação da aplicação por parte dos seus utilizadores finais. Devido à aplicação de uma metodologia de desenho centrada no utilizador [Reis08] a quando do processo de desenvolvimento desta ferramenta não foram encontrados problemas de usabilidade. Não obstante, percebeu-se que: 1) o desempenho da aplicação decresce significativamente a quando da sua utilização num PDA; 2) há modalidades de interacção e dispositivos que são mais eficientes em determinados contextos; 3) os

utilizadores, em geral, consideram esta ferramenta fácil de utilizar; 4) e concordam com a sua utilização para o suporte de actividades educacionais como trabalhos de casa e exames escolares.

Conduzimos também um estudo comportamental de forma a entender quais as modalidades de interacção adoptadas pelos utilizadores finais desta aplicação num conjunto de contextos, um pouco maior do que o considerado na avaliação: casa, parque, metro e carro. Este estudo permitiu-nos perceber alguns padrões comportamentais dos utilizadores no que respeita as modalidades de interacção utilizadas em diferentes contextos do quotidiano, definidos por um grupo de variáveis um pouco maior do que o considerado na avaliação: dispositivo, ruído, iluminação, movimento, posição, conteúdos, número de pessoas presentes, limitações temporais, forma de agarrar o dispositivo.

Os nossos planos para trabalho futuro assentam na condução de uma avaliação e estudo semelhantes direccionado para os modos visual e eyes-free da ferramenta de manipulação, envolvendo somente a utilizadores com diferentes tipos de deficiências, de forma a tentar encontrar padrões comportamentais no que respeita a selecção de modalidades de interacção em contextos reais.

6. AGRADECIMENTOS

Este trabalho foi parcialmente financiado por EU, LaSIGE e FCT, através do projecto JoinTS e do programa Multiannual Funding Program

This work was supported by EU, LASIGE and FCT, through project JoinTS and through the Multiannual Funding Program.

7. REFERÊNCIAS

- [Baillie05] Baillie, L., Schatz, R., Exploring multimodality in the laboratory and the field, Proc. of the 7th international conference on Multimodal interfaces, Toronto, Italy (2005)
- [Beck03] Beck, E.T., Christiansen, M.T., and Kjeldskov, J.: Experimental Evaluation of Techniques for Usability testing of Mobile Systems in a Laboratory Setting. In Proceedings of OzCHI'2003. Brisbane Australia, (2003)
- [Blattner96] Blattner, M.M. and Gliner, E.P.: Multimodal integration. IEEE Multimedia 14–24 (1996)
- [Blenkhorn98] Blenkhorn P., Evans D. G.: Using speech and touch to enable blind people to access schematic diagrams. Journal of Network and Computer Applications, Vol. 21, pp. 17-29 (1998)
- [Boyd90] Boyd, L.H., Boyd, W.L., Vanderheiden, G.C.: The Graphical User Interface: Crisis, Danger and Opportunity, Q. Journal of Visual Impairment and Blindness, Vol. 84, pp. 496-502 (1990)
- [Brewster02] Brewster SA: Overcoming the lack of screen space on mobile computers. Personal and Ubiquitous Computing 6(3):188–205 (2002)
- [Brewster03] Brewster S.A., Lumsden J., Bell M., Hall M., Tasker S.: Multi-modal 'eyes free' interaction techniques for wearable devices. Proc. CHI'03 Conference on Human Factors in Computing Systems, CHI Letters, ACM Press, 5(1): 473–480 (2003)

- [Burger93] Burger, D., Sperandio, J.: (ed.) *Non-Visual Human-Computer Interactions: Prospects for the visually handicapped*. Paris: John Libbey Eurotext. pp. 181-194 (1993)
- [Goodman04] Goodman, J. Gray, P., Khammampad, T., and Brewster, S. Using Landmarks to Support Older People in Navigation. In *Proceedings of Mobile HCI 2004*, Verlag, Berlin, Heidelberg, 38-48 (2004)
- [Hurtig06] Hurtig, T., 2006. A mobile multimodal dialogue system for public transportation navigation evaluated. In *Procs. of HCI'06*. ACM Press pp. 251-254 (2006)
- [Kjeldskov04] Kjeldskov, J., Skov, M.B., Als, B.S., and Hoegh, R.T. Is it Worth the Hassle? Exploring the Added Value of Evaluating the Usability of Context-Aware Mobile Systems in the Field. In *Proceedings of Mobile HCI 2004*, Berlin, Heidelberg, 61-73 (2004)
- [Lai04] Lai, J.: Facilitating Mobile Communication with Multimodal Access to Email Messages on a Cell Phone. In *Procs. of CHI'04*. ACM Press, New York, pp. 1259-1262 (2004)
- [Makela00] Makela, A., Giller, V., Tscheligi, M., and Sefelin, R.: Joking, storytelling, art sharing, expressing affection: A field trial of how children and their social network communicate with digital images in leisure time. In *Proceedings of CHI'00*, ACM Press, New York, 548-555 (2000)
- [Nicolle01] Nicolle, C. and Abascal, J.: Eds. *Inclusive Design Guidelines for HCI*. Taylor and Francis, London (2001)
- [Reis08] Reis, T., Sá, M. Carriço, L. "Designing Mobile Multimodal Artefacts ". In *Procs of ICEIS 2008*, Barcelona - Spain, June, 2008. Pp. 79-85 ISBN: 978-989-8111-48-7.
- [Oviatt99] Oviatt, S: Mutual disambiguation of recognition errors in a multimodal architecture. *Procs. CHI'99*. ACM Press: New York, pp. 576-583 (1999)
- [Oviatt04] Oviatt, S., Darrell, T., and Flickner, M.: Multimodal interfaces that flex, adapt, and persist. *Commun. ACM* 47, 1, 30-33 (2004)
- [Sá08] Sá, M., Carriço, L.: Defining Scenarios for Mobile Design and Evaluation. In *Procs of CHI'08, SIGCHI Conference on Human Factors in Computing Systems*, Florence, Italy, ACM Press (2008)
- [Santoro07] Santoro, C., Paternò, F., Ricci, G., Leporini, B.: A Multimodal Museum Guide for All. In *Mobile interaction with the Real World Workshop, Mobile HCI*, Singapore (2007)
- [Signer06] Signer, B., Norrie, M., Grossniklaus, M., Belotti, R., Decurtins, C., Weibel, N.: Paper Based Mobile Access to Databases. In *Procs. of the ACM SIGMOD* pp: 763-765 (2006)
- [Turk00] Turk, M. and Robertson, G.: Perceptual user interfaces introduction. *Commun. of the ACM* 43, 3, 33-35 (2000)