



9º Congresso de Pós-Graduação

INCLUSÃO DE UM MÓDULO DE INTELIGÊNCIA NA ARQUITETURA MMIF DA W3C

Autor(es)

FABIO MARQUES ROQUE

Orientador(es)

PROFA.DRA. CECILIA SOSA ARIAS PEIXOTO

1. Introdução

INTERFACE MULTIMODAL

Desde o advento da microinformática, há trinta anos, a forma do homem se relacionar com o computador tem evoluído significativamente. No princípio, o contato do homem com o computador se dava apenas através do teclado, do cartão perfurado, do monitor e da impressora, que ganharam um novo aliado no final da década de oitenta: o mouse.

Isso quer dizer que toda informação que o homem precisava ler do computador era através do monitor e da impressora e todo comando se dava exclusivamente através do teclado e do mouse, que são chamados de dispositivos de entrada.

Com a implantação desses dispositivos de entrada, criou-se durante anos um paradigma de utilização exclusiva de ícones, ponteiros, menus e janelas. Seja clicando ou digitando, o processo de busca exaustiva em vários menus de consulta ou pastas de armazenamento de arquivos digitais tornou-se desgastante, além de restringir a produtividade e a acessibilidade e requerer uma aprendizagem específica.

Uma alternativa para melhorar a interação entre homem e computador é aproveitar os sentidos e movimentos naturais do homem e fazer com que o computador possa interpretá-los e transformá-los em comandos, textos e outras funções, que até então eram desempenhadas com a utilização do mouse e do teclado.

A computação moderna deve agregar os cinco sentidos fundamentais que propiciam o nosso relacionamento com o ambiente. Essa proposta de interação entre uma pessoa e um computador advém da definição utilizada em psicologia que se refere às modalidades sensoriais das pessoas, como visão, audição, tato, olfato e paladar.

WORLD WIDE WEB CONSORTIUM / MULTIMODAL INTERACTION FRAMEWORK (W3C/MMIF)

Para viabilizar a utilização de interfaces multimodais dentro da internet foi criado um padrão, o W3C/MMIF. O World Wide Web Consortium (W3C) é um consórcio internacional que congrega mais de 300 membros, dentre eles empresas, órgãos governamentais e organizações independentes de âmbito internacional, que visam desenvolver padrões para a criação e a interpretação de conteúdos para a Internet.

O W3C possui diversos comitês que estudam as diferentes tecnologias existentes para a apresentação do conteúdo da Internet e criam padrões de recomendação para utilizar essas tecnologias. Com isso, os programas dos diferentes segmentos do mercado conseguem acessar facilmente os códigos e entender onde deve ser aplicado cada conhecimento expresso no documento.

Dentro dos diversos comitês do W3C foi criado um grupo de trabalho chamado Multimodal Interaction (MMI). O grupo de trabalho MMI desenvolve uma arquitetura ou padronização para a estrutura de Interface Multimodal (MMIF). Essa padronização é um formato que deve ser seguido por todas as empresas ou pessoas que desenvolverem aplicações na Internet, que visam criar uma interação multimodal.

A Arquitetura Multimodal descreve um quadro geral e flexível de interoperabilidade dos diferentes componentes do quadro multimodal de uma forma abstrata. Por isso não pode ser considerada uma arquitetura propriamente dita. Pois, uma arquitetura indica como os componentes são atribuídos aos dispositivos de hardware e do sistema de comunicação que permitam os dispositivos de hardware para se comunicar uns com os outros.

O W3C MMIF não descreve de que forma os componentes estão atribuídos a dispositivos de hardware ou de como o sistema de comunicação permite que os dispositivos de hardware para se comunicar. Cada componente representa um conjunto de funções relacionadas.

O W3C MMIF pode ser visto como uma base para o desenvolvimento de aplicações multimodais. O framework é construído de uma gama de linguagens de acordo com o padrão do W3C chamado de DOM (Document Object Model). DOM define interfaces em que programas e scripts dinamicamente podem acessar e atualizar o seu conteúdo.

W3C – APLICATIVOS MÓVEIS

Visando melhorar a experiência dos usuários na Internet o World Wide Web Consortium (W3C) criou um documento de melhores práticas para a entrega de conteúdos na internet. O conteúdo desse material é dirigido principalmente a criadores, mantenedores e operadores de sites.

Esse guia de melhores práticas tem a intenção de criar uma base comum de entendimento mesmo que isso possa parecer óbvio ou trivial para aqueles que possuem experiência nessa área.

A qualidade da experiência dos usuários Web através da utilização dos dispositivos móveis depende significativamente da usabilidade, dos navegadores e do próprio dispositivo. Apesar disso esse documento destaca as melhores práticas para melhorar a usabilidade do site.

O documento é formado por diversos tópicos que destacam dentre outras questões a forma de apresentação, limitações das tecnologias, vantagens, formas de entrada e saída e descreveremos abaixo alguns tópicos que julgamos merecer destaque.

Dentro das questões de apresentação o documento apresenta que existem experiências ruins nos usuários na utilização dos dispositivos móveis por conta ao tamanho limitado da tela e da quantidade de material limitado visível para o usuário, exigindo dessa forma que o usuário utilize muitas vezes o recurso de rolagem da tela.

REVISÃO BIBLIOGRÁFICA

As publicações obtidas no contexto deste trabalhos são:

- Authoring Applications for the Multimodal Architecture, W3C Working Group, Note 2 July 2008, disponível na internet no site da W3C (www.w3.org).
- Mobile Web Best Practices 1.0, Basic Guidelines, W3C Recommendation 29 July 2008, disponível na internet no site da W3C (www.w3.org).
- Voice XML - Voice Extensible Markup Language, disponível na internet no site da W3C (www.w3.org).
- Software Engineering and Human-Computer Interaction, Richard N. Taylor, Joëlle Coutaz, ICSE '94 Workshop on SE-HCI, Sorrento, Italy, may, 1994
- Dix, A. J., Finlay, J. E., Abowd, G. D. and Beale, R.(1998), Human-Computer Interaction, Prentice Hall, London.

2. Objetivos

A arquitetura MMIF proposta pela W3C tem dentre seus objetivos definir uma forma de interação dos dispositivos de entrada e saída entre os usuários e o sistema.

Verifica-se que não há nenhum componente responsável em gerenciar uma inteligência que possa auxiliar o usuário durante a utilização do sistema.

Na especificação MMIF não há indicação de que algum componente seja responsável por: recuperar e filtrar informações, reduzir o tempo na localização da informação e acesso a serviços, considerar perfil dos usuários, ajudar os usuários a realizar tarefas. Segundo Ehtel, estas características são atribuídas as interfaces inteligentes que auxiliam o usuário durante a realização da tarefa.

Essa proposta apresenta a possibilidade de incluir um módulo de inteligência dentro da arquitetura definida pela W3C/MMIF que é hoje um padrão de interfaces multimodais.

Esse módulo de inteligência seria um quarto componente na arquitetura W3C/MMIF para ampliar a relação entre o sistema e o usuário.

Esse módulo de inteligência pode reunir informações que contribuiriam com o usuário durante a utilização do sistema. Dentre as muitas contribuições que um módulo de inteligência poderia realizar relatamos alguns itens que são: recuperar informações relevantes, filtrar informações irrelevantes, reduzir tempo na localização da informação e acesso ao serviço, considerar características particulares dos usuários, ajudar os usuários em sistemas mais complexos e até realizar tarefas para os usuários

Desenvolvimento

Para a avaliação desse trabalho será desenvolvida uma interface multimodal à ferramenta Kira, que é uma ferramenta de ensino de mineração de dados (MENDES; VIEIRA, 2009). Nessa proposta será desenvolvida uma interface que sugere o uso de voz e toque. Além de permitir sua utilização em dispositivos móveis, nela será implementada o módulo de inteligência que apresentará guias de auxílio de aprendizagem a ferramenta, controle de nível de dificuldade na interação do usuário, recuperação de informações relevantes e filtros de informações irrelevantes. O importante desse trabalho é que esse componente irá se apresentar ao usuário, pois, existem muitos componentes que ficam ocultos ao usuário.

Utilizaremos para a aplicação o desenvolvimento da aplicação em ferramenta Java (Java Sun Microsystems) com a utilização de um SDK (Software Development Kit, ou kit para desenvolvimento de aplicativos) para compilação em Android, que é um sistema operacional para dispositivos móveis.

3. Desenvolvimento

Para a avaliação desse trabalho aplicaremos uma interface à ferramenta Kira, uma ferramenta de ensino de mineração de dados (MENDES; VIEIRA, 2009), que será desenvolvida com voz e toque e para dispositivos móveis, nela será implementada o módulo de inteligência que apresentará guias de auxílio de aprendizagem a ferramenta, controle de nível de dificuldade na interação do usuário, recuperação de informações relevantes e filtros de informações irrelevantes. O importante desse trabalho é que esse componente irá se apresentar ao usuário, pois, existem muitos componentes que ficam ocultos ao usuário.

Utilizaremos para a aplicação o desenvolvimento da aplicação em ferramenta Java (Java Sun Microsystems) com a utilização de um SDK (Software Development Kit, ou kit para desenvolvimento de aplicativos) para compilação em Android, que é um sistema operacional para dispositivos móveis.

TRABALHOS RELACIONADOS

A multimodalidade vem ganhando análise de pesquisadores e profissionais de todo o mundo, isso pode ser comprovado pelas inúmeras pesquisas e tecnologias disponíveis.

As pesquisas apresentadas exploram muito a utilização de comando de voz e toques em telas sensíveis chamadas também chamado de touch screen.

No artigo Speak4it: Multimodal Interaction for Local Search (Copyright 2010 ACM), de Patrick Ehlen, AT&T e Michael Johnston, AT&T Labs Research vemos um aplicativo denominado Speak4itSM que realiza pesquisas orientadas para o consumidor de aplicativos móveis, ele permite aos usuários procurarem informações sobre negócios.

Esse aplicativo suporta uma verdadeira integração multimodal, onde as entradas do usuário podem ser distribuídas através de múltiplos modos de entrada. Além de especificar consultas através da voz, os usuários podem combinar voz e gestos.

O aplicativo permite que ao mesmo tempo seja ditado o texto "postos de gasolina" e desenhado a rota no display do dispositivo e ele retornará os postos de gasolina ao longo do especificado rota traçada no visor.

Esse artigo soma a um montante de outros artigos que demonstram que a busca de locais tem sido frequentemente utilizada como uma aplicação de teste para investigação de técnicas de processamento e interface multimodal modelos.

No artigo: A Multimodal Interactive Text Generation System (Copyright 2010 ACM 978), de Luis Rodríguez, Ismael García-Varea, A. Revuelta-Martínez e Enrique Vidal é apresentado um sistema de geração interativa de texto destinado a fornecer assistência para a digitação de textos em diferentes ambientes. Este sistema funciona através da previsão de que o usuário vai digitar com base no texto que ele ou ela digitou anteriormente.

4. Resultado e Discussão

O que se pretende discutir são os resultados obtidos com o desenvolvimento da aplicação levando em conta o novo módulo incluído na arquitetura da W3C.

Pretende-se também discutir o avanço de utilizar um módulo de inteligência com a utilização de interfaces multimodais em aplicativos móveis.

Essa proposta de trabalho quer mostrar a inclusão de um módulo inteligente na especificação da arquitetura MMIF, é que está inclusão contribui de modo significativo durante a interação multimodal da interface. Existem diversos sistemas que utilizam filtros de informações irrelevantes e recuperação de informações relevantes, dentre outras inteligências sistêmicas expressivas, contudo, muitas vezes o próprio usuário não sabe que a interface possui estas habilidades e não tem acesso ao tipo de filtro que é usado.

Já essa proposta é a utilização de técnicas de IA (Inteligência Artificial) no desenvolvimento do módulo de inteligência, permitindo assim que ele se apresente ao usuário, possa mostrar a sua linha de raciocínio na busca de informações e na realização das tarefas. Desta forma possibilitará que o usuário esteja consciente do que a interface pode ou não fazer e como este módulo contribui na realização de tarefas. Ao incluir o módulo de inteligência no estudo de caso da ferramenta KIRA, testes de usabilidade serão feitos e avaliações dos testes serão analisados em relação à inclusão do módulo inteligente e também em relação à interação multimodal. Ainda sendo uma ferramenta específica de mineração de dados, o que se pretende levantar é como o módulo auxilia especificamente durante o aprendizado e durante as diversas etapas de mineração.

5. Considerações Finais

Considera-se importante a entrada de um módulo de inteligência na interação homem-máquina, pois, há uma quantidade significativa de ferramentas tecnológicas introduzindo a multimodalidade na interação, todavia, é interessante que não se apóie somente nela para aproximar o sistema das pessoas. Pode-se avançar muito mais nessa interação.

Os recursos disponíveis nos equipamentos hoje em dia devem estimular o desenvolvedor de aplicações a aprimorar ainda mais o seu

trabalho no momento de criar recursos sistêmicos que armazenem informações para apoiar o usuário no momento do uso da ferramenta, criando principalmente possibilidades de recuperação de informações e apoio a navegação no sistema. Conclui-se que hoje é possível utilizar de recursos multimodais e ainda poder utilizar um módulo de inteligência que representará um grande avanço na interação homem-máquina, abrindo assim novas discussões sobre a melhoria na usabilidade dos sistemas.

Referências Bibliográficas

- Authoring Applications for the Multimodal Architecture, W3C Working Group, Note 2 July 2008, , disponível na internet no site da W3C (www.w3.org).
- Mobile Web Best Practices 1.0, Basic Guidelines, W3C Recommendation 29 July 2008, , disponível na internet no site da W3C (www.w3.org).
- Voice XML - Voice Extensible Markup Language, disponível na internet no site da W3C (www.w3.org).
- Software engineering and human-computer interaction, Richard N. Taylor, Joëlle Coutaz, ICSE '94 Workshop on SE-HCI, Sorrento, Italy, may, 1994
- Dix, A. J., Finlay, J. E., Abowd, G. D. and Beale, R.(1998), Human-Computer Interaction, Prentice Hall, London.
- Toward multimodal human-computer interface, R Sharma, VI Pavlovic - Proceedings of the IEEE, 1998 - ieeexplore.ieee.org
- WAZLAWICK, Raul Sidnei; Metodologia de Pesquisa para Ciência da Computação; Rio de Janeiro: Elsevier, 2008.
- Multimodal Interaction for Local Search (Copyright 2010 ACM), de Patrick Ehlen, AT&T e Michael Johnston, AT&T Labs Research
- A Multimodal Interactive Text Generation System (Copyright 2010 ACM 978), Luis Rodríguez, Ismael García-Varea, A. Revuelta-Martínez e Enrique Vidal
- Kira: Uma Ferramenta Instrucional para Apoiar a Aplicação do Processo de Mineração de Dados, Eduardo Fernando Mendes, Marina Teresa Pires Vieira, XX Simpósio Brasileiro de Informática na educação (2009)
- EHLERT, P. Intelligent User Interfaces. Technical Report DKS03-01/ICE01, Data and Knowledge Systems Group, Department of Information Technology and Systems, The Netherlands: Delft University of Technology, 2003.