

Aplicações Interactivas em Realidade Aumentada

Paulo Ricca Rui Braga Marta Boas Vasco Duarte Marco Sousa João Coelho

Luís Romero

Escola Superior de Tecnologia e Gestão
Instituto Politécnico de Viana do Castelo
Av. do Atlântico, 4900 Viana do Castelo
Portugal

{paulo.goncalves, rbraga, mboas, vduarte, marcosousa, jcoelho}@ipvc.pt
romero@estg.ipvc.pt

Sumário

Este documento descreve duas aplicações de Realidade Aumentada onde são aplicados conceitos de interacção natural. A primeira utiliza um feixe de luz laser como instrumento principal de interacção para aumentar informação no meio ambiente. A segunda desenvolve um jogo educativo, direccionado para crianças em idade pré-escolar, cuja interacção se baseia em cubos para formar palavras. Em ambas a interacção processa-se de forma natural, sem recurso a dispositivos auxiliares. Ambas foram desenvolvidas por alunos no âmbito duma disciplina do curso de Engenharia da Computação Gráfica e Multimédia.

Palavras-chave

Realidade Aumentada, Computação Ubíqua, Interacção, Computação Gráfica, Multimédia.

1. INTRODUÇÃO

Este artigo apresenta contribuições para a construção de aplicações em Realidade Aumentada onde a interacção se processa de forma natural. Para o efeito descrevem-se duas aplicações; A *eVoSpot* e *Cubinhos*. A primeira aplicação aumenta informação no espaço real dum ambiente consoante a posição dum feixe de luz controlado pelo utilizador. A informação, visual, sonora ou de outra natureza, é adicionada ao espaço dependendo da área para onde se aponta o feixe de luz. Esta aplicação estende o conceito de realidade aumentada, desviando-se do tradicional aumento visual associado à realidade aumentada. O aumento da informação é feito directamente na realidade através da projecção.

A segunda aplicação, *Cubinhos*, utiliza ferramentas de realidade aumentada para oferecer um jogo infantil de descoberta de palavras. A criança dispõe os cubos com letras num tabuleiro que são reconhecidas pela aplicação caso forme uma palavra. Ao reconhecer uma palavra o sistema apresenta num monitor a forma e o som relacionados com a palavra reconhecida.

Ambos os trabalhos foram desenvolvidos por alunos do curso de Engenharia da Computação Gráfica e Multimédia do Instituto Politécnico de Viana do Castelo. O artigo apresenta uma breve panorâmica do trabalho relacionado, descreve a implementação das duas aplicações, apresenta

resultados para ambas e propõe desenvolvimentos futuros.

2. TRABALHO RELACIONADO

A Realidade Aumentada tem sido, desde há uma década, uma área de intensa investigação científica [Azuma01]. A interacção em ambientes de Realidade Aumentada foi, no início, baseada em dispositivos transportados pelos utilizadores como óculos ópticos [Billinghurst01], dispositivos móveis [Schmidt00], e outros. Mais recentemente foi desenvolvido o conceito de Tangible User Interfaces (TUI) [Ullmer00], que consistem em mesas interactivas onde os utilizadores manipulam ferramentas que alteram o conteúdo da interface, incluída ou não no próprio TUI. A computação ubíqua conjuga a disponibilização de informação associada a posições geográficas do utilizador, estendendo o conceito de aumento de informação em espaços reais [Abowd00]. Vários trabalhos têm sido desenvolvidos abordando estes paradigmas que conjuguem os vários modos de interacção em ambientes de realidade aumentada como em DWARF [Sandor05] (Uma *framework* baseada em CORBA para prototipagem rápida de aplicações distribuídas de realidade aumentada).

O facto da aplicação *eVoSpot* se basear em apontar um ponteiro laser para uma projecção coloca a sua natureza próxima de uma aplicação comercial para crianças desenvolvida pela YDreams [HELLO] onde é usada uma

lanterna para escolher cores de um desenho projectado, como num livro para colorir, no entanto, a plataforma aqui apresentada não é apenas uma aplicação de realidade aumentada, mas uma plataforma de criação de aplicações desta natureza e pode ser usada para vários fins, por exemplo, para desenvolver aplicações informativas ou de ajuda à navegação. Foram encontrados também alguns protótipos semelhantes, elaborados por Ashish Derhga-wen [Derhga-wen08], onde era usado um ponteiro laser para controlar aplicações de desenho e de leitores de música, e para controlar o movimento de uma câmara. Recentemente foi também desenvolvida uma aplicação muito semelhante ao DigitalGraffiti para um stand da TMN na conferência OFFF em Maio de 2008 [TMN08]. Análises de utilização de interfaces baseadas em projecções podem ser encontrados em [Bonanni05].

3. APLICAÇÃO EVOSPOT

Este trabalho teve raízes no estudo sobre Realidade Aumentada, mais concretamente, na procura da solução para o problema de interacção com um computador que, do ponto de vista do utilizador, não existe. Quando a Interface e os conteúdos se transformam aparentemente no próprio sistema, a procura de novos paradigmas de interacção torna-se uma necessidade. Teve-se também em mente o crescente interesse global por interfaces cooperativas, multi-utilizador. É neste contexto que se resolveu recorrer ao um gesto natural e multi-cultural, que se faz desde criança: apontar. Como recurso tecnológico é utilizado um ponteiro laser por ser um dispositivo comum, de baixo custo, sem fios e com grande autonomia. A aplicação funciona como um *middleware* que se encarrega de transformar o ponteiro laser num dispositivo de interface humana, que poderá ser usado para interagir com outras aplicações.

A interacção é feita apontando o ponteiro para um espaço físico. O resultado desta interacção varia consoante a aplicação e poderá ser tanto informação adicional ou uma ampliação da zona para a qual se está a apontar ou até o desencadeamento de uma acção pré-programada, por exemplo, o envio de um SMS específico para alguém.

O leque de aplicações possíveis é muito vasto e são sugeridas algumas:

- (a) Assistente de telefonemas para idosos: É comum os idosos terem retratos dos familiares ao seu redor. Seria útil poderem efectuar uma chamada telefónica apenas apontando o laser para o retrato de quem querem contactar.
- (b) Detalhes sobre um objecto (Figura 1): ao apontar o ponteiro para um ponto específico do objecto, são mostrados os detalhes sobre esse ponto, e/ou informações visuais ou auditivas.
- (c) Chamamento de colegas: Num hospital os enfermeiros e os médicos podem ser chamados apontando o ponteiro para as suas fotografias, dispostas numa parede. Este sistema pode funcionar também para chamar assistentes num super-mercado.



Figura 1: Esquema de utilização do eVoSpot numa aplicação onde é mostrada informação adicional sobre uma estátua.

- (d) *Graffiti* virtual: Numa determinada parede (na rua, numa galeria, num museu, etc.), os artistas podem expressar a sua arte livremente através de pinturas digitais. No final da pintura esta é guardada e posteriormente mostrada a sequência de pinturas feitas anteriormente.
- (e) Alternativa económica ao ecrã táctil permitindo ainda a interacção à distância, podendo ser utilizado, por exemplo, na apresentação da Meteorologia em televisão.

Sobre este *middleware* foram elaborados dois trabalhos diferentes: *eVoHotSpot* e *DigitalGraffiti*.

eVoHotSpot

Esta aplicação permite ao administrador do sistema especificar polígonos de acção na imagem captada pela câmara. Os polígonos dependem da calibração descrita à frente, permitindo que a câmara mude de posição sem que estes deixem de referenciar os objectos correctos. O administrador poderá especificar que acções são despoletadas quando o utilizador aponta o ponteiro para uma determinada zona.

O *eVoHotSpot Manager*, onde é configurado o sistema, foi programado em *C#*, devido à rapidez de desenvolvimento, e o *eVoHotSpot*, a aplicação que despoleta as acções determinadas previamente, foi criado em *Adobe Flash*, devido às suas reconhecidas capacidades multimédia. Esta aplicação poderá ser aplicada nos sistemas (b) e (c).

DigitalGraffiti

Esta aplicação, desenvolvida em *C#*, é um programa de desenho onde o utilizador poderá escolher entre “pintar” com uma lata de *spray* ou com um lápis virtuais. Poderá também escolher a cor e a largura do traço. O utilizador

aponta para a superfície e, consoante o movimento do ponteiro, é criado um rasto de “tinta”. Como o fundo projectado é inicialmente preto (ausência de projecção), o utilizador não tem noção de que existe uma imagem projectada na parede, no entanto quando aponta o ponteiro o desenho parece aparecer “magicamente”.

3.1 Sistema

O Sistema consiste no ponteiro laser, num computador que processa a imagem captada por uma *webcam* e, opcionalmente, um projector que é usado para dar um *feedback* visual da actuação do utilizador e/ou altifalantes para o *feedback* sonoro, caso a apresentação assim o exija. A organização física do sistema está apresentada na figura 2.

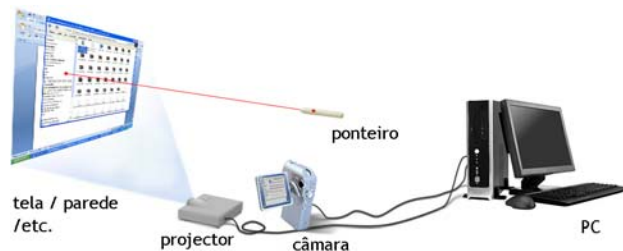


Figura 2 – Esquema do sistema

3.2 Módulos

Quanto ao funcionamento interno do *eVoSpot*, pode-se considerar que existem dois módulos principais, descritos de seguida.

3.2.1 Calibração

A Calibração é o processo que garante uma correlação correcta entre o espaço físico e o espaço virtual. O espaço físico corresponde ao espaço para onde o utilizador pode apontar o ponteiro e o espaço virtual é uma representação esquemática desse espaço, onde se moverá um ponteiro virtual (guiado pelo ponteiro real) que poderá interagir com objectos virtuais representativos de objectos reais. O objectivo da calibração é transformar o sistema de coordenadas distorcido da imagem captada pela *webcam* num sistema de eixos ortogonais, utilizáveis pelo sistema, e também delimitar a área de interacção, no espaço físico.

Existem vários métodos para fazer a calibração. Caso exista um projector pode-se optar entre dois:

- (a) O programa pode fazer a calibração automaticamente, projectando no espaço pontos brancos nos cantos da imagem projectada, que são posteriormente captados pela *webcam* e reconhecidos pelo sistema.
- (b) Esta calibração pode ser feita manualmente. Neste caso o projector projecta uma tela em branco e o utilizador indica ao sistema a localização dos cantos da imagem directamente, clicando na imagem captada pela *webcam*. A imagem pode ser uma área física, como se pode ver na Figura 3.

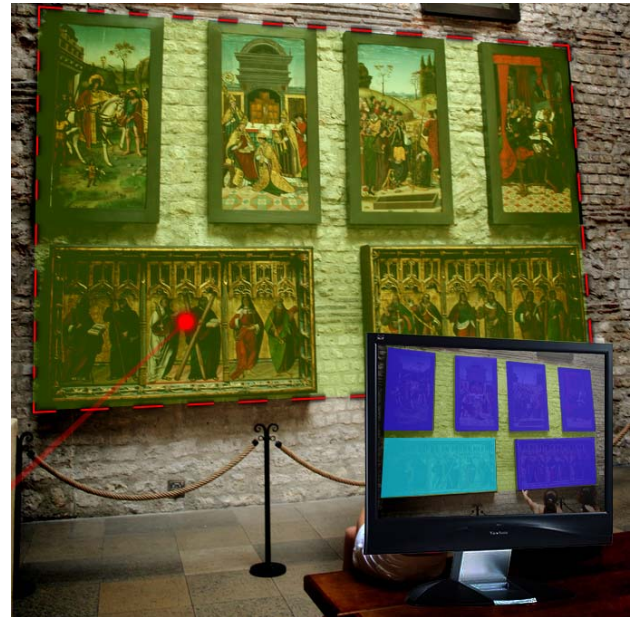


Figura 3: Esquema de Calibração e áreas sensíveis do espaço físico.

Caso não exista um projector, o utilizador deverá indicar ao programa as dimensões físicas da área com que pretende interagir e de seguida indicar os cantos dessa área, clicando nos respectivos pontos, na imagem captada pela câmara.

3.2.2 Reconhecimento

Há a necessidade de procurar o ponteiro laser na imagem captada. Este processo poderá parecer simples, porém, a diversidade de texturas, cores e iluminação presentes num mesmo local dificulta a tarefa ao sistema de reconhecimento. Foram testadas várias técnicas diferentes sendo possível escolhermos entre duas delas:

Ponto Brilhante

Este algoritmo consiste basicamente em encontrar o ponto mais brilhante e, em caso de dúvida, dar mais “importância” aos pontos brilhantes que tenham também uma aura vermelha. Este método funciona bem quando não existem pontos de luz ou objectos reflectores (como vidros, metais, etc.).

Reconhecimento de Padrões

Devido ao mau funcionamento do algoritmo anterior nos casos indicado, foi desenvolvido um outro algoritmo, baseado em reconhecimento de padrões. Este algoritmo funciona de forma genérica e procura um dado padrão na imagem captada pela câmara.

Existem várias opções características de cada algoritmo, como o raio de procura, técnicas de procura, etc., e duas comuns aos dois, nomeadamente, a preferência por pontos em movimento (em dúvida, opta sempre por um ponto que está em movimento; geralmente o ponteiro não fica completamente parado, mesmo que o utilizador tenha essa intenção) e preferência por pontos próximos da

escolha anterior (em dúvida, opta pelo ponto mais próximo do ponto detectado anteriormente).

Depois de reconhecido o ponteiro, a sua posição relativamente à área calibrada da imagem é atribuída ao cursor do rato, para que este possa interagir com as outras aplicações, por exemplo o *eVoHotSpot*.

3.3 Resultados

Os resultados práticos deste trabalho medem-se pelo desempenho dos utilizadores, pelo seu grau de satisfação e pela quantidade de informação que a aplicação ajudou a fornecer.

Foram criadas algumas aplicações com o *eVoHotSpot*, nomeadamente, o Mapa Aumentado que consiste num mapa de Portugal em papel pendurado na parede em que, quando o utilizador aponta o laser para um dado distrito, este ilumina-se e aparecem informações adicionais relativas ao mesmo (nome, pequena descrição e fotografia). Foi ainda implementado um protótipo de um sistema para chamar pessoas por altifalantes apontando o laser para fotografias colocadas numa parede.

Em geral os utilizadores acham interessante poder interagir com uma projecção, sem nenhum dispositivo ligado ao computador. Um pormenor que reduz um pouco a confiança dos utilizadores enquanto utilizam o programa é o pequeno atraso do cursor em relação ao ponteiro laser. Este problema pode ser resolvido com a utilização de melhores câmaras (com menos atraso) e no desenvolvimento de técnicas mais avançadas de captura e análise da imagem, para melhorar a performance global da aplicação.

4. APLICAÇÃO CUBINHOS

A aplicação *Cubinhos* consiste no desenvolvimento de um jogo educacional para crianças utilizando novos paradigmas de interacção. O objectivo do jogo é a criação de palavras através da junção de vários cubos que contêm uma letra em cada face. Cada palavra tem um tema associado, como animais ou cores.

Através da realidade aumentada, é proporcionado à criança uma forma simples de obter conhecimento utilizando o entretenimento na aprendizagem.

4.1 Sistema

O sistema consiste num computador, uma câmara, colunas e um conjunto de cubos com letras. Poderá ser necessário em alguns casos utilizar uma fonte de luz auxiliar, dependendo do tipo de iluminação existente no local. O Computador contém uma estrutura de dados flexível que permite associar a cada palavra modelos 3D, áudio e imagens que são mostradas à criança quando esta forma uma palavra válida com os cubos. Os cubos são distribuídos num tabuleiro, ou mesa, que a criança dispõe com toda a liberdade numa área do tabuleiro, como se pode observar na Figura 4.



Figura 4: Concepção do sistema.

O sistema interpreta o vídeo capturado pela câmara, analisando a existência de possíveis marcas e de possíveis sequências válidas. Caso exista uma sequência válida é apresentado no monitor uma representação tridimensional da palavra constituída e simultaneamente produzido um som.

4.2 Ferramentas

Para o desenvolvimento da aplicação foi utilizado um conjunto de ferramentas de desenvolvimento descritas a seguir:

1. ARToolkit [Billinghurst01]: Uma biblioteca de desenvolvimento de aplicações de Realidade Aumentada. Esta biblioteca apresenta algumas características importantes tais como:
 - Fácil calibração da câmara;
 - Possibilidade de criação de novas marcas
 - Possibilidade de utilizar qualquer marca, desde que seja quadrada;
 - Renderização rápida;
 - Muitos exemplos e documentação disponível.
2. MS Visual Studio C++: Toda a base da aplicação foi desenvolvida no Ambiente Integrado de Desenvolvimento do Visual Studio utilizando a linguagem de programação C++. É responsável pela implementação da comunicação com a livreria ARToolKit e pelo suporte das várias classes criadas para apoiar o funcionamento do jogo, nomeadamente, o acesso aos objectos 3D, texturas e sons.
3. OpenGL [Woo98]: Biblioteca de desenvolvimento de aplicações tridimensionais. É responsável por toda a interface gráfica da aplicação e pelo modo de navegação entre menus.

4.3 Estrutura de Dados

A estruturada da informação é apoiada directamente pelo sistema de ficheiros, fazendo uso da sua característica hierárquica. A cada categoria está associada uma directoria que contém um conjunto de ficheiros onde são arma-

zenados o conjunto de palavras reconhecíveis e respectivas imagens, sons e objectos 3D. A associação entre as palavras e as respectivas apresentações (modelos 3D, imagens, sons) é feita através dum ficheiro de texto onde é definido a respectiva relação.



Figura 5: Sistema de ficheiros hierárquicos.

Actualmente existem 5 categorias: *cores*, *números*, *transportes*, *animais*, e *frutas*. Cada categoria é constituída por um ficheiro de texto que contém a descrição da palavra e as apresentações relacionadas (Figura 5). Por exemplo na categoria *cores*, este ficheiro contém as palavras; *verde*, *azul*, *branco* e *vermelho*. São estes nomes que são considerados pela aplicação aquando da validação da sequência de letras (marcas) existente no tabuleiro. Os nomes contidos neste ficheiro de texto são responsáveis pela associação de uma sequência correcta e à respectiva representação tridimensional, imagem e som associado à palavra detectada.

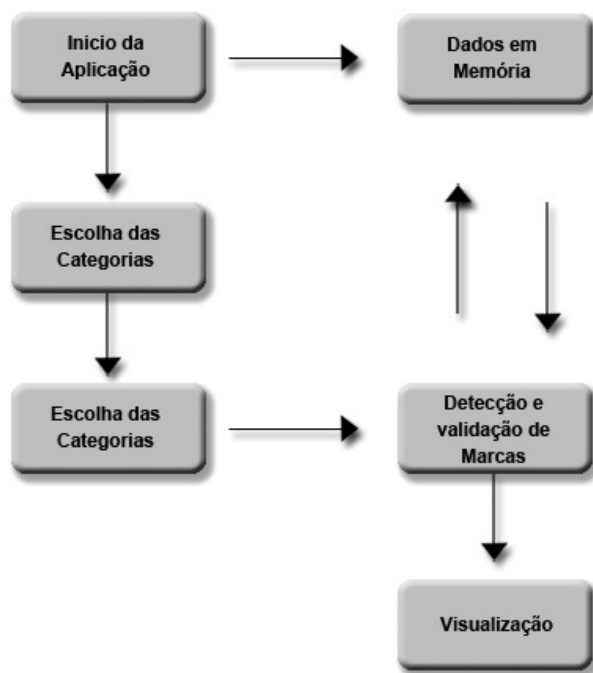


Figura 6: Arquitectura da aplicação.

4.4 Funcionamento

A aplicação inicia com o carregamento para a memória de todos os dados necessários, nomeadamente texturas, modelos 3D, sons e as categorias definidas na estrutura de ficheiros descrita.

É escolhida uma categoria antes do processo de validação de palavras iniciar. De seguida é iniciada a captura de vídeo pela câmara de forma a permitir o reconhecimento de marcas através do ARTToolkit. A cada marca está associado um identificador que representa uma letra de forma a possibilitar a validação das palavras. Após a aferição da posição e orientação das marcas, identificadas na aplicação pelo seu identificador, é analisada uma área específica presente no tabuleiro e verificada a existência de uma sequência válida de letras que forme uma palavra definida na estrutura de informação. Caso seja encontrada uma sequência válida de palavras, é produzido um som associado à palavra formada, assim como é apresentada no monitor uma representação tridimensional dessa mesma palavra.

4.5 Resultados

A aplicação *Cubinhos* proporciona à criança uma interacção simples, divertida e natural, pois não é necessário criar novos comportamentos na criança. A criança não tem de possuir qualquer dispositivo para interagir com o sistema.

A aplicação permite alterar a dificuldade do jogo, bastando para isso que sejam alterados os elementos constituintes de cada categoria e que sejam colocados nas directorias respectivas, as imagens, modelos 3D e sons.

Na Figura 7 podemos verificar as três áreas distintas da interface da aplicação. Do lado direito as miniaturas dos elementos disponíveis para a categoria *animais*, em baixo o menu, onde se poderá escolher outra categoria, activar/desactivar o som ou sair da aplicação. Na parte superior direita é apresentado o vídeo capturado pela câmara onde se pode verificar a área disponível para formar uma palavra e, caso seja validado uma palavra correcta, onde é igualmente apresentado uma representação tridimensional associada à palavra, neste caso um cão.

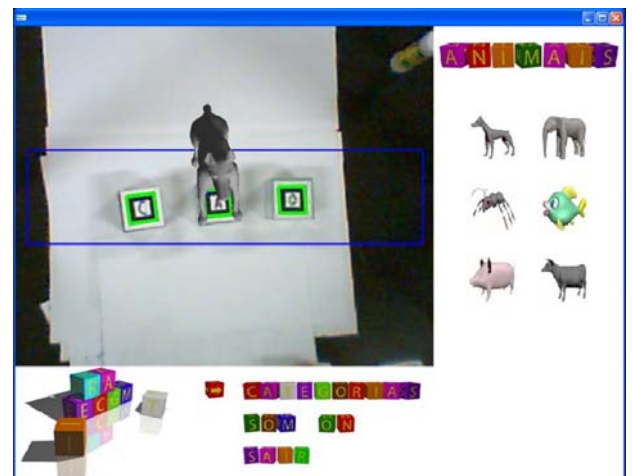


Figura 7: Interface gráfica de Cubinhos.

5. CONCLUSÕES

Os novos meios de interacção que são possibilitados através de aplicações como o *eVoSpot* fazem cada vez mais uso dum interacção natural, onde o utilizador ape-

nas utiliza gestos naturais para interagir com os sistemas. Existem, no entanto, muitas barreiras por ultrapassar até os tradicionais métodos de interacção homem-máquina serem totalmente alterados. Nessa altura a interacção poderá ser, em alguns casos, totalmente privada de auxílios tecnológicos associados ao utilizador. As interfaces tangíveis e hepáticas têm vindo a ganhar uma nova vida porque foi finalmente percebida a importância do *feedback* instantâneo das acções do utilizador na performance do sistema e na confiança que é garantida quando interagimos de forma instintiva e inata.

Aplicações como *Cubinhos* poderão ser utilizadas no ensino da escrita com crianças com maior dificuldade de aprendizagem. A interacção natural da aplicação permite que estas não se apercebam do sistema enquanto estão a brincar, proporcionando um ambiente de brincadeira natural e ao mesmo tempo educativo. Estas aplicações poderão ser utilizadas em vários contextos, não apenas no ensino da escrita, mas também, por exemplo, no auxílio da aprendizagem de outras ciências como a matemática. Desta forma podemos incentivar e motivar as crianças para se empenharem em áreas de maior dificuldade de aprendizagem. Como diz um ditado; a brincar se aprende.

6. AGRADECIMENTOS

Agradecemos a colaboração dos restantes alunos do 5.º ano do curso de Engenharia da Computação Gráfica e Multimédia pela sua ajuda na concepção das aplicações através das várias mesas redondas e dos seus comentários aos vários trabalhos desenvolvidos.

7. REFERÊNCIAS

[Abowd00] Abowd, Gregoy, Mynatt, Elizabeth.. Charting Past, Present, and Future Research in Ubiquitous

Computing. *ACM Transactions on Computer-Human Interaction*, Vol. 7, No. 1, March 2000, Pages 29–58.

- [Azuma01] R. Azuma, Y. Baillot, R. Behringer, S. Feiner, S. Julier e B. MacIntyre. Recent Advances in Augmented Reality. *IEEE Computer Graphics and Applications*, November/December 2001, 34-47.
- [Billinghurst01] Billinghurst, M., Hirkazu, K., Poupyrev, I.: The MagicBook: Moving Seamlessly between Reality and Virtuality. *IEEE Computer Graphics and Applications*, Amy-June 2001 pp 2-4.
- [Bonanni05] Bonanni, Leonardo, Lee, Chia-Hsun, Selker, Ted. Attention-Based Design of Augmented Reality Interfaces. *CHI 2005*, April 2-7, 2005, Portland, Oregon, USA.
- [Derhgawen08] Ashish Derhgawen, *Laser Tracking with webcam for Human-Computer Interaction*. <http://ashishrd.blogspot.com/2008/03/laser-tracking-with-webcam-for-human.html>
- [HELLO] YDreams, Aplicação para a loja Hello Kitty no Bairro Alto, Lisboa. <http://www.ydreams.com>
- [Schmidt00] Schmidt A, Gellersen H-W, Beigl M, Thate O. Developing user interfaces for wearable computers—don't stop to point and click. *Proceedings of IMC 2000*, Rostock, Germany, November, 2000
- [TMN08] TMN, Aplicação interactiva de Graffiti digital, OFFF08.
- [Ullmer00] Ullmer B, Ishii H (2000). Emerging frameworks for tangible user interfaces. *IBM Systems Journal* 39(3–4):915–931.
- [Woo98], Woo, M, Neider, J, Davis, T. OpenGL Programming Guide. *Addison Wesley*, 1998.