

CameraPhone

Interacção baseada em Câmara

Tiago Trindade
 Dep. Eng^a. Informática, IST
 Av. Rovisco Pais, 1000 Lisboa
 tiago.trindade@ist.utl.pt

Tiago Guerreiro
 Dep. Eng^a. Informática, IST
 Av. Rovisco Pais, 1000 Lisboa
 tjvg@immi.inesc-id.pt

Joaquim A. Jorge
 Dep. Eng^a. Informática, IST
 Av. Rovisco Pais, 1000 Lisboa
 jorgej@acm.org

Sumário

Nos últimos anos os dispositivos móveis têm evoluído a uma velocidade enorme. No entanto, o modo de interagir com os mesmos tem-se mantido constante, continuando a apresentar problemas em diversas aplicações e contextos, revelando uma interacção muito restritiva. Estas restrições apresentam-se como mais relevantes no contexto de pessoas com deficiências a nível motor. O trabalho aqui apresentado expõe uma possível solução para este problema através da utilização da câmara como um método de interacção com os próprios dispositivos ou mesmo com sistemas externos. Através da detecção de movimento e da identificação de pontos de referência nas imagens captadas pela câmara é possível associar acções específicas aos padrões detectados, permitindo assim, por exemplo, a utilização da câmara de um telemóvel para aceitar ou rejeitar chamadas, escrever mensagens ou mesmo para navegar num sistema de controlo de ambiente.

Palavras-chave

Câmara, Processamento de Imagem, Tecnologias Assistivas, Dispositivos móveis, Dificuldades Motoras

1. INTRODUÇÃO

No espaço de duas décadas, a maneira como interagimos no dia-a-dia com os dispositivos móveis mudou radicalmente. Aparelhos como o telemóvel, leitores de música portáteis, *Personal Digital Assistants* (PDA) ou mesmo câmaras fotográficas digitais passaram a estar presentes no nosso quotidiano, saindo diariamente para o mercado novos modelos renovados com inúmeros melhoramentos e inovações. Os telemóveis, por exemplo, passaram de simples aparelhos de luxo a bens indispensáveis na nossa sociedade, contribuindo para isso o baixo preço dos mesmos e a evolução que estes tiveram ao nível das comunicações e da panóplia de operações que estes nos permitem executar.

No entanto, e apesar da evolução que se tem sentido ao nível das características técnicas destes aparelhos, o mesmo não se pode dizer no que toca ao modo como interagimos com eles. A interacção com estes dispositivos manteve-se praticamente constante ao longo dos anos, sendo baseada em grande parte no modo de interacção presente nos primeiros modelos destes aparelhos. O facto de o seu tamanho ser pequeno limitou também os métodos de interacção disponíveis, sendo o teclado e a utilização de botões/joysticks os métodos de interacção mais utilizados, sendo assim todo o sistema desenhado em função destes. Isto levanta problemas como a navegação lenta em aplicações com vários menus e submenus, tornando-se esta confusa, o desenvolvimento de aplicações pouco orientadas ao utilizador, ou mesmo a utilização em situações do

quotidiano em que não seja possível por exemplo estar a olhar para o teclado ou para o ecrã. Os dispositivos móveis com ecrãs tácteis têm ganho alguma relevância nos últimos anos mas apresentam-se como uma opção ainda mais restritiva e que obriga a maior precisão. Todas as operações efectuadas nestes dispositivos tornam-se ainda mais complicadas para pessoas com dificuldades motoras. Esta população está fisicamente limitada o que leva a que, em indivíduos com lesões mais graves como a tetraplegia, o simples carregar com sucesso num dado botão do telemóvel para atender uma chamada se torne algo extremamente difícil ou mesmo impossível, dependendo do nível da lesão. Os telemóveis, apesar de cada vez mais evoluídos não estão ainda dotados de mecanismos de interacção para pessoas com este tipo de problemáticas. Para estas pessoas, a ideia de poder e conseguir comunicar facilmente com alguém ou com o ambiente que as rodeia tem uma importância enorme no seu dia-a-dia e no seu desenvolvimento. Utilizadores deste grupo podem utilizar, no seu dia-a-dia, diferentes tipos de dispositivos para interacção com o ambiente envolvente, nomeadamente switches, trackballs, joysticks, teclados virtuais, ponteiros, entre outros (revisto em [Guerreiro08]). No entanto, estas tecnologias são ainda demasiado restritas ao ambiente e acção para o qual são desenvolvidas, não permitindo uma adaptação fácil por parte dos utilizadores a novos cenários de utilização. São ainda, na sua maioria, muito dependentes de um posicionamento e acomodação exactos e específicos, que limitam os utilizadores a um horário diminuto no qual

têm ajuda de terceiros para reunir as condições necessárias para operar o dispositivo.

Os dispositivos móveis no entanto, com um custo baixo e uma mobilidade elevada que permite que sejam levados com o utilizador para onde este se deslocar, apresentam-se como uma alternativa interessante para substituírem alguns dos aparelhos referidos acima, caso seja possível dotá-los de algumas capacidades/características destes últimos. É no entanto fulcral fazer a ponte entre o utilizador e o dispositivo, visto que a interacção actual apresenta lacunas e não permite um controlo efectivo.

Recentemente têm sido desenvolvidas abordagens para interacção alternativa com dispositivos móveis, das quais se destacam as baseadas em acelerómetros, ecrãs tácteis, câmaras e identificação por radiofrequência (RFID). Estas representam abordagens com foco na realização de um dado movimento ou através da aproximação, sendo menos exigentes ao nível da precisão de movimentos que as interfaces de controlo tradicionais. Depois de analisadas e comparadas [Gamboia07], rapidamente nos apercebemos que as duas hipóteses que se mostram aplicáveis neste caso são o uso de acelerómetros e o uso da câmara. No entanto, e depois de um confronto directo entre estas duas ultimas a que se parece mais viável é a utilização da câmara. Principalmente por estar hoje presente na maioria dos telemóveis, para além de ser relativamente barata e da qualidade das mesmas ter vindo a melhorar, o que permite que no futuro a utilização desta tecnologia demonstre ainda melhores resultados. A nossa abordagem caracteriza-se então pelo uso de uma câmara para potenciar a interacção de utilizadores com deficiências motoras (mas com algumas capacidades residuais dos membros superiores) com o dispositivo móvel e, através deste, com o ambiente envolvente.

2. TRABALHO RELACIONADO

Existem alguns trabalhos baseados na utilização da câmara como método de interacção alternativo ou complementar, tendo a evolução dos telemóveis com câmara levado a que mais estudos se desenrolem nesse sentido. Dos vários trabalhos, existem duas áreas que se apresentam importantes na compreensão do funcionamento destas soluções: detecção de movimentos e pontos de referência.

2.1 Detecção de Movimentos

A detecção de movimento apresenta-se como uma das bases para toda a interacção baseada em câmara. Neste caso as imagens adquiridas são processadas relativamente às imagens capturadas anteriormente ou a imagens pré-definidas.

Na área dos jogos, a Nokia através de Haro et al. [Haro05] apresentou algoritmos de detecção de movimentos com câmara em diversos ambientes virtuais, como por exemplo um puzzle 2D e um jogo de acção. Aprofundaram também o facto deste tipo de interacção facilitar a utilização de outro tipo de aplicações, como navegar na Web ou visualizar imagens. Vendo que se tornava complicado e cansativo navegar numa aplicação

móvel através de um joystick ou de uma caneta, propuseram utilizar a câmara do telemóvel para tal efeito. Utilizaram o vídeo que a câmara capturava para detectar não só a direcção, como também a força com que esse gesto tinha sido feito. De referir que ainda ao nível dos videojogos e da interacção nos mesmos através da câmara, a Sony utilizou este tipo de tecnologia, através do aparelho EyeToy para desenvolver e lançar para o mercado vários jogos baseados apenas no vídeo captado por essa câmara. Wang et al. [Wang06] utilizando, técnicas simples como diferença de imagens e correlação de blocos para estimativa de movimento, conseguiu que o algoritmo TinyMotion, com apenas 4 etapas, permitisse controlar 3 tipos de aplicações no telemóvel com graus de satisfação bastante elevados.

Este tipo de técnica enfrenta no entanto alguns problemas difíceis de resolver, nomeadamente as diferenças de luminosidade e de brilhos em diferentes zonas enquanto a pessoa se movimenta (ou movimenta a câmara do telemóvel), e também a análise de movimentos muito rápidos, visto que cada imagem tem de ser analisada antes da chegada [Çapin06] da nova próxima imagem capturada pela câmara, de modo a que a interacção possa ser real.

2.2 Pontos de Referência / Etiquetas Visuais

A utilização de códigos ou etiquetas visuais aparece como uma alternativa ao simples processamento de imagem proveniente da câmara do telemóvel [Rohs04]. Depois de realizada a identificação do código através da câmara, o dispositivo passa a ter uma base de processamento relativo a esse código. Pode processar tanto a informação presente nesse código, como utiliza-lo como ponto de referência para interacções relativas ao mesmo.

Os códigos visuais, nomeadamente os códigos de barras, apareceram bastante cedo, sendo que depois do sucesso do CyberCode da Sony [Falas07] vários investigadores começaram a desenvolver projectos nesta área. Semacode, Widgor [Falas07] e QR-Code, são apenas três exemplos de empresas que criaram soluções deste tipo, sendo o último específico para uso em dispositivos móveis. Rohs et al. [Rohs07] aprofunda as metáforas de interacção com jogos de realidade aumentada através da sua técnica baseada nestes códigos visuais ou “marcadores”. Através da utilização desses marcadores em objectos reais permite-se que os mesmos forneçam a informação necessária ao jogo ou aplicação sobre a posição do jogador. Ao comparar a imagem do marcador passada pela câmara com a imagem original desse marcador é possível determinar qual a inclinação e distância a que o utilizador está num determinado momento. Ao nível de pessoas com dificuldades cognitivas e motoras, [Granollers06] num dos seus estudos utilizou uma webcam como um dispositivo de controlo do rato para pessoas com paralisia cerebral. Os resultados mostraram-se positivos para interacções mais simples em que eram utilizados como pontos de referencia a identificação de marcadores de cor.

3. ABORDAGEM

Este trabalho foca-se em, através da utilização das câmaras presentes nos telemóveis e outros dispositivos móveis, permitir criar uma solução de interacção com os próprios dispositivos ou com sistemas externos, para pessoas com dificuldades a nível motor, nomeadamente tetraplegia. De entre este grupo dos utilizadores as capacidades motoras divergem dependendo do nível de lesão. Utilizadores com lesões superiores à vértebra C6 têm dificuldades ou incapacidade nesta interacção. Mesmo quando existem capacidades motoras para interagir com estes aparelhos, esta interacção é feita com bastantes erros, aumentando a frustração por parte do utilizador. Assim, e visto que a criação de soluções para este problema poderá ter um grande impacto no quotidiano destes indivíduos, definimos dois cenários nos quais esta tecnologia pode ser aplicável em telemóveis.

3.1 Controlo do telemóvel

Através da identificação de movimentos básicos por parte do utilizador, visto que muitos dos indivíduos desta população têm dificuldade em realizar movimentos precisos e repetidos, é possível utilizar a câmara do telemóvel como um *switch*. Depois de definida qual a função a associar a cada detecção de movimento podemos utilizar a câmara como interface de controlo. Temos o caso de se atender uma chamada passando duas vezes o braço/mão em frente da câmara. Ou rejeitar a chamada colocando a mão a tapar a câmara. Utilizando objectos de cor, podemos associar a identificação dos mesmos, através da câmara, a uma acção. Por exemplo associar a cor rosa à acção “ligar para a mãe”. Assim, sempre que o utilizador apontar a câmara para um objecto dessa cor (ou passar um objecto de cor na linha de visão da câmara), o telemóvel liga para a sua mãe e coloca o telefone em alta voz.

3.2 Controlo de sistemas externos

Visto que os telemóveis são sistemas com uma capacidade de comunicação cada vez maior (*Bluetooth*, *infravermelhos*, *Wifi*) é possível utilizá-los para controlar outros sistemas. Utilizando a detecção de movimento o telemóvel pode, por exemplo, navegar e controlar um menu de um sistema de controlo de ambiente. Pode ainda, mais uma vez, associar objectos de cor a determinadas funções desse, ou de outros sistemas.

4. INTERACÇÃO BASEADA EM CÂMARA

Surge assim a necessidade de criar uma solução para dispositivos móveis que suporte estas tarefas. De referir que as soluções apresentadas anteriormente se basearam em duas técnicas: detecção de movimento e identificação de cores. Assim foi desenvolvida uma biblioteca de processamento de imagem genérica para telemóveis e outros dispositivos móveis (*Windows Mobile*), funcionando esta também em computadores pessoais com webcam. A utilização desta assenta em filtros de imagem. Os filtros de imagem são a base para todo o processamento de imagem. É através destes que é possível alterar as imagens captadas de modo a obtermos informação de contexto útil contida nas mesmas. Existem

diferentes filtros, desde filtros de pré-processamento (correção de luminosidade, etc.), aos filtros de pós-processamento (eliminação de imperfeições, etc.), passando pelos filtros centrais de processamento (*threshold*, diferença de imagens, etc.). A conjugação dos diversos filtros permite criar mecanismos de processamento específicos para cada tipo de necessidade.

4.1 Filtros de Detecção de Movimento

A detecção de movimento pode ser realizada de diferentes formas, mas uma das mais usuais é a da utilização de filtros de diferença de imagem para detectar diferenças entre as várias imagens capturadas pela câmara. Neste caso a detecção de movimentos é efectuada através da utilização de um filtro de diferença de imagens conjugado com filtros de pré e pós-processamento. A imagem capturada é comparada com uma imagem anterior pré-definida, sendo detectadas as diferenças entre as mesmas. Depois de aplicados filtros de limpeza da imagem, é analisada a percentagem de mudança entre as imagens, pelo que se esta atingir um determinado valor podemos concluir que houve movimento durante aquele intervalo de tempo.

4.2 Filtros de Cor e Detecção de Objectos

Os filtros de cor apresentam-se como um filtro que apenas torna visível uma determinada cor já definida. Este percorre a imagem, e analisando-a coloca a branco apenas as partes da imagem que sejam dessa cor, deixando a preto as outras. No final obtém-se assim uma imagem binarizada (apenas com pixéis brancos ou pretos), onde é possível identificar as partes da imagem relativas à cor pretendida. A detecção de objectos é feita depois através de um filtro especial de “*blob-counter*”. Este percorre a imagem à procura dessas tais zonas brancas que tenham um determinado tamanho mínimo e máximo, identificando cada uma dessas como objectos da cor inicialmente definida.

5. APLICAÇÃO

Através de algumas características é possível definir um “alfabeto” para controlo, usando o telemóvel. Este alfabeto pode depender de gestos, cores, formas.

5.1 Detecção de movimentos

Utilizando o filtro descrito na secção 4.2 criou-se um mecanismo que utiliza essa detecção de movimento como uma interacção com o sistema. De referir que, neste caso, a imagem pré-definida utilizada para comparação é actualizada regularmente de modo a que a deslocação da câmara não influencie o resultado (Figura 1).

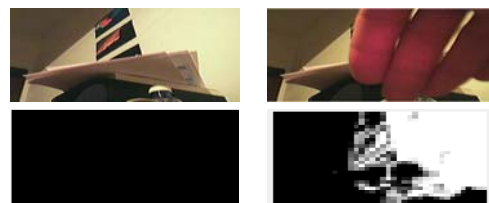


Figura 1 – Detecção de movimento

5.2 Identificação de Cores

Utilizando uma conversão da imagem capturada pela câmara do espaço de cor RGB para o espaço de cor HSL, é possível decompor a imagem em apenas cor, saturação e luminância. Ao termos a componente da cor separada da componente da luz é possível identificar determinadas cores sem que a variação de luz num determinado local ao longo do tempo influencie em demasia essa detecção.



Figura 2 – Detecção de objectos de cor

Assim a identificação de objectos de cores é realizada, depois de feita a conversão para HSL, aplicando um filtro que detecta os objectos de uma dada cor (Figura 2).

6. RESULTADOS

Para validar a abordagem, efectuámos testes com dois utilizadores tetraplégicos de nível C5, ou seja, que apresentam controlo dos ombros e algum controlo impreciso ao nível dos braços (bíceps). Para tal era apresentado ao utilizador num ecrã um sistema que simula um menu de uma aplicação de controlo de ambiente. Utilizando uma técnica de varrimento, as linhas de botões vão sendo percorridas uma a uma automaticamente pelo sistema (linhas a vermelho). Quando o utilizador pretende seleccionar uma das linhas em questão efectua uma passagem do braço/mão por cima da câmara, sendo essa passagem detectada pelo sistema e a linha seleccionada. O procedimento é depois o mesmo para a selecção dos botões presentes nessa linha já seleccionada.



Figura 3 – Controlo de ambiente através de varrimento

Neste tipo de utilizadores ficou patente que a utilização desta tecnologia como um *switch* se apresenta como uma boa solução, simples e fácil de utilizar, permitindo ao utilizador acompanhar todo o processo de selecção, obtendo-se melhores resultados que a utilização do rato ou teclado por exemplo. Os resultados obtidos mostram uma taxa de sucesso na selecção de 98,9%, tendo realizado com sucesso todas as tarefas pedidas, e demorando por exemplo 5,5 segundos para seleccionar o botão “Quarto”. A utilização da técnica de identificação

de cores não teve no entanto resultados tão bons com estes utilizadores, tendo-se verificado que esta técnica está bastante dependente da capacidade de cada utilizador conseguir agarrar e movimentar objectos (neste caso, o telemóvel).

7. CONCLUSÃO E TRABALHO FUTURO

Podemos assim concluir que dada a tecnologia existente actualmente no mercado, nomeadamente a quantidade de telemóveis com câmara, e o baixo custo, mobilidade e capacidade de comunicação destes dispositivos comparativamente a outros aparelhos existentes, tornam a hipótese de utiliza-lo como um novo método de interacção bastante interessante. Estes poderão desempenhar o papel ainda mais importante em pessoas com dificuldades motoras e/ou cognitivas, pois estas dependem ainda bastante de tecnologias assistivas demasiado estáticas e que necessitam de ajuda externa para serem utilizadas/montadas. O trabalho futuro foca-se na realização de testes com utilizadores para aferir a sua utilidade e melhorias a realizar. Com a evolução da capacidade de processamento destes dispositivos móveis e do aumento da qualidade das câmaras, fica aberta a utilização deste tipo de interacção para outro tipo de actividades no quotidiano, mesmo para populações que não enfrentam dificuldades a nível motor.

8. AGRADECIMENTOS

Tiago Guerreiro foi financiado pela Fundação para a Ciência e a Tecnologia, bolsa SFRH/BD/28110/2006.

9. REFERÊNCIAS

- [Çapin06] Çapin, T., A. Haro et al. Camera-Based Virtual Environment Interaction on Mobile Devices, ISCI 2006
- [Falas07] Falas, T., Two-Dimensional Bar-Code Decoding with Camera-Equipped Mobile Phones, PerComW'07
- [Gamboa07] Gamboa, Ricardo., Mnemonical Body Shortcuts Gestural Interface for Mobile Devices (2007), Master Thesis, IST, Technical University of Lisbon.
- [Granollers06] Granollers, T., Mauri, C., et al., Computer Vision Interaction for People with Severe Movement Restrictions. In Human Technology, 2, 38-54 (April 2006)
- [Guerreiro08] Guerreiro, Tiago. Myographic Mobile Accessibility for Tetraplegics (2008), Master Thesis, IST, Technical University of Lisbon.
- [Haro05] A. Haro, K. Mori, et al., Mobile camera-based user interaction. ICCV 2005 Workshop on HCI (2005)
- [Rohs04] Rohs, M., Real-World Interaction with Camera-Phones, UCS (2004)
- [Rohs07] Rohs, M., Marker-Based Embodied Interaction for Handheld Augmented Reality Games. In Journal of Virtual Reality and Broadcasting, 4(5) (Mar. 2007)
- [Wang06] Wang, J. and Canny, J. 2006. TinyMotion: camera phone based interaction methods. CHI '06 (2006)