

Plataforma SIGUS e o Desenvolvimento de Jogos Eletrônicos com Interação Homem-Computador baseada em Visão Computacional

Vinícius A. Saueia
Bruno Cesar G. de Toledo
Thiago G. Moretto
Roberto H. Viana
Hemerson Pistori

Universidade Católica Dom Bosco

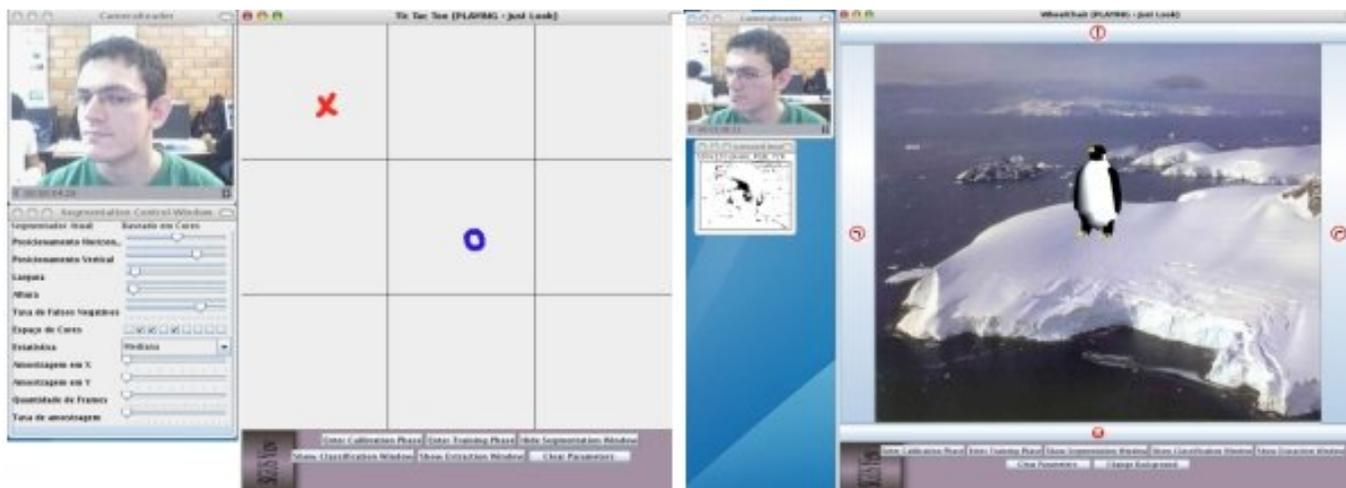


Figura 1: Aplicativos desenvolvidos com a biblioteca SIGUS.

Resumo

Neste artigo será apresentada a plataforma SIGUS que é uma biblioteca de apoio ao desenvolvimento de aplicações para pessoas com necessidades especiais. A biblioteca SIGUS reúne diversas técnicas de Visão Computacional, aprendizagem automática e processamento digital de imagens em um único ambiente multiplataforma e com códigos-fonte livre. Este trabalho mostra como utilizar técnicas de visão computacional no desenvolvimento de jogos eletrônicos para um diferente modo de interação homem-computador, uma opção pouco explorada atualmente. São mostrados também exemplos de jogos desenvolvidos com base na plataforma SIGUS que utilizam interface guiada por sinais visuais.

Keywords:: Visão Computacional, Interação Homem-Computador, Pessoas com Necessidades Especiais

Author's Contact:

{vsilvaucdb,brunoucdb,thiago.moretto}@gmail.com
hdbobos@gmail.com
pistori@ucdb.br

1 Introdução

Teclado, mouse e *joysticks* são as interfaces de interação com a máquina mais utilizadas hoje em dia. A utilização dessas interfaces por pessoas com necessidades especiais pode ser desconfortável e muito difícil. Por não conseguirem interagir de maneira simples com computadores, ou consoles de vídeo game, pessoas com necessidades diferentes acabam sendo excluídas do mundo digital. O número de pessoas com alguma deficiência pelo mundo é muito alto, mais de 600 milhões de acordo com a Organização Mundial de Saúde (OMS)[Passafaro 2002].

Utilizar interfaces baseadas em técnicas de Visão Computacional pode ser uma boa alternativa tanto para pessoas com necessidades especiais, quanto para usuários das interfaces convencionais.

O custo de *webcams*, dispositivo de captura de imagens mais comum hoje em dia, caiu muito nos últimos anos, tornando-se um acessório de fácil acesso para muitos usuários de computador. Com técnicas de Visão Computacional, pode-se fazer a substituição total ou parcial dos dispositivos de interação homem-computador convencionais por um dispositivo de aquisição de imagens.

Os dispositivos convencionais de comunicação com usuário citados podem causar alguns malefícios para quem faz o uso constante e por um longo intervalo de tempo, como por exemplo, a Lesão por Esforço Repetitivo (LER). Utilizando-se a *webcam* como interface de entrada de dados do sistema, os usuários podem interagir com aplicativos, como jogos eletrônicos, através de movimentos com o corpo inteiro e/ou gestos com as mãos, braços e cabeça. Este modelo de interação fornece à pessoa, além da diversão propiciada pelo jogo eletrônico, um exercício físico que pode beneficiar a sua saúde. Portanto, a utilização dessas técnicas se tornam uma alternativa para inclusão digital de pessoas impossibilitadas de interagir com as interfaces homem-computador mais comuns atualmente.

Este trabalho apresenta a biblioteca SIGUS que pode ser utilizada por desenvolvedores de aplicações, inclusive de jogos, para incluir técnicas de Visão Computacional e assim criar um ambiente de interação homem-computador diferente em seus produtos. Por se tratar de um novo tipo de comunicação entre usuário e máquina, a Visão Computacional apresenta um grande potencial para atrair novos consumidores ao mercado de games [de Paula et al. 2006]. Neste trabalho são apresentados jogos eletrônicos desenvolvidos com a plataforma SIGUS, os quais são jogados com sinais visuais (posturas ou gestos capturados pela *webcam*).

Este artigo está estruturado em cinco seções. A Seção 2 apresenta alguns trabalhos correlatos que mostram alguns jogos desenvolvidos utilizando o mesmo tipo de interface homem-máquina proposto no presente trabalho. A Seção 3 mostra a plataforma utilizada para a implementação dos jogos. Em sequência são apresentados os jogos implementados. Por fim, são mostradas as conclusões e propostas para trabalhos futuros.

2 Trabalhos Correlatos

Em [de Paula et al. 2006] é descrito o *Camera Kombat*, um jogo de combate virtual multijogador que dispensa o uso de dispositivos comuns de interação com computador através da combinação de técnicas de visão computacional e uma *webcam*. Durante a partida os jogadores realizam movimentos em frente a uma *webcam* devidamente posicionada e quando uma seqüência de movimentos é detectada o avatar que representa o jogador dispara um objeto virtual em seu adversário. Não há necessidade de contato físico entre os jogadores e as seqüências de movimentos são pré-cadastradas. Na fase de segmentação dos jogadores são utilizadas técnicas de subtração de fundo e os gestos são reconhecidos com auxílio da técnica de fecho convexo. Segundo os autores, este tipo de abordagem apresentou um elevado nível de atividade física e é mais atrativa aos consumidores do que dispositivos convencionais de interação como teclado, mouse e *joysticks*. Os autores também enfatizam a viabilidade da utilização deste modelo de interação tendo em vista o barateamento de *webcams* e a evolução de algoritmos de visão computacional.

Um outro jogo com suporte a gestos é apresentado por Teixeira em [Teixeira et al. 2006], porém com mais preocupação em relação aos tempos de resposta do sistema e no impacto que o modelo de interação causa na renderização de frames. Dessa forma, para análise de desempenho foram utilizados como métrica o tempo de resposta (tempo entre a realização do movimento a exibição do mesmo) e a taxa de frames por segundo (FPS). O jogo é dividido em dois módulos, onde o módulo CIDA é responsável pela interação com o usuário, e segundo os autores apresenta um alto nível de abstração permitindo que o código seja independente do modelo de interação. Ao decorrer da partida, imagens são capturadas através de uma *webcam* e processadas em tempo real. Dois marcadores são utilizados para realização dos movimentos, um contendo a letra G e outro contendo a letra F, no seu centro, que devem estar sempre de frente para o dispositivo de captura de imagens, o que torna a tarefa de classificação muito simples. O jogo requer pouco esforço físico e as várias configurações dos marcadores tornam o jogo complicado. Segundo os autores o sistema apresentou resultados satisfatórios em relação ao FPS, porém o tempo de resposta não se aplicaria a jogos de resposta rápida (corrida, luta entre outros). Os autores não especificam os métodos de segmentação e reconhecimento de padrões utilizados que foram identificados como gargalo do sistema juntamente com a comunicação com a *webcam*.

O desempenho de dois métodos de rastreamento é avaliado em [Christopher Fagiani 2002]. Os métodos são aplicados a um sistema de extração de características em tempo real, o *Camera Mouse*, que interpreta entradas de vídeo para manipulação de um ponteiro de mouse. Foram testados os métodos de Lukas-Kanade e um método baseado em correlação juntamente com filtros de Kalman multidimensionais. Várias combinações dos métodos com filtros em diferentes dimensões foram testadas quanto a acurácia, performance e praticidade. O método de correlação sem filtros de Kalman apresentou melhores resultados nos experimentos. O método de correlação é calculado para cada ponto, em uma área de 50x50 *pixels*, que a cada interação é posicionada com base no frame anterior. O ponto com maior valor de correlação é considerado o novo ponto do mouse. Um valor mínimo deve ser atingido para que todos os pontos não sejam considerados perdidos e o sistema necessite ser reiniciado manualmente. O método de Lucas-Kanade é baseado em fluxo óptico e assume brilho constante para características com pouca variação de movimento. Uma função descrita no artigo é responsável pelo mapeamento de movimentos faciais em movimentos de mouse. A função leva em consideração a resolução da tela do usuário e o tamanho das imagens capturadas. As combinações de métodos foram testadas com uma série de softwares de IHC e avaliados quanto a velocidade, acurácia, capacidade de manipulação de movimentos aleatórios ou errôneos e tempo de processamento. Apesar dos filtros de Kalman terem aumentado significativamente a performance computacional não apresentaram bons resultados em algumas modalidades de teste, como os de movimentos errôneos e aleatórios.

O grupo de pesquisa MERL (Mitsubishi Electric Research Laboratories) desenvolveu pesquisas na área de Visão Computacional para

interação homem-computador. Em [Freeman et al. 1998] os autores mostram como podem ser utilizadas várias técnicas de Visão Computacional, como segmentação de imagens, rastreamento de objetos, entre outras, para interagir com aplicações computacionais através de uma câmera. No mesmo trabalho, são apresentados vários jogos eletrônicos desenvolvidos, como *Jockey Pow* e ainda diversas modalidades de esporte adaptadas para o novo modo de interação. Mostra-se também em [Freeman et al. 1998] como a interação através de câmeras pode ser utilizada no dia-a-dia, onde é apresentado um controle remoto virtual em que as funções da televisão são acionadas através de comandos gerados pela mão e pelo braço do usuário.

3 Plataforma SIGUS

O desenvolvimento de aplicações com suporte à interação homem-computador utilizando técnicas de Visão Computacional é, normalmente, uma tarefa complicada e requer um tempo de desenvolvimento maior. Desenvolver esse tipo de sistema é difícil pois as técnicas são, geralmente, muito complexas.

Para facilitar o desenvolvimento de aplicações nas quais a interação homem-computador é feita através da *webcam* foi desenvolvida a Plataforma de Apoio ao Desenvolvimento de Sistemas para Inclusão Digital de Pessoas com Necessidades Especiais, SIGUS¹. A plataforma SIGUS consiste em uma biblioteca de apoio ao desenvolvimento de aplicações que utilizam um dispositivo de captura de imagens como interface para interação homem-computador. Para isso, técnicas de áreas como Visão Computacional, aprendizagem automática e processamento digital de imagens estão disponíveis para o uso na construção de novas aplicações [Pistori 2004; Pistori et al. 2006].

O desenvolvimento da plataforma SIGUS visa melhorar o bem-estar de pessoas com necessidades especiais e explorar novas formas de interação do ser humano com a máquina. Diversas aplicações onde podem ser utilizadas as interfaces visuais vêm sendo estudadas, inclusive em jogos eletrônicos. A Sony desenvolveu o *EyeToy*², uma câmera que captura os movimentos do jogador de *Play Station 2* e substitui o *joystick* em alguns jogos. Existem alguns outros problemas que tem sido pesquisados na área de Visão Computacional, tais como: “Tradução da Linguagem de Sinais” e “Controle de Cadeira de Rodas”. [Pistori 2006]

A plataforma SIGUS é uma biblioteca inteiramente livre e gratuita³. Os códigos-fonte foram desenvolvidos na linguagem de programação JavaTM, versão 5.0. Escolheu-se essa linguagem de programação por ser facilmente portátil para qualquer sistema operacional.

Estão presentes na plataforma SIGUS diversas técnicas das áreas de Visão Computacional, Processamento Digital de Imagens e Aprendizagem Automática. Para facilitar o desenvolvimento do ambiente optou-se por utilizar bibliotecas bastante desenvolvidas e consolidadas para os módulos de processamento de imagens e aprendizagem automática. A biblioteca escolhida para auxílio no processamento de imagens foi a ImageJ⁴, uma biblioteca livre e toda desenvolvida em Java. Para os módulos de aprendizagem automática escolheu-se o WEKA⁵ [Witten and Frank 2005].

As ferramentas de apoio em conjunto com a plataforma SIGUS, criam um ambiente de programação completo para o desenvolvimento de aplicações com suporte a interação homem-computador através de sinais visuais. Diversas técnicas de segmentação de imagens, extração de atributos, rastreamento de objetos, seleção de atributos, entre outras, estão disponíveis para serem utilizadas pelos programadores. Estão disponíveis técnicas como segmentação por fundo adaptativo, segmentação por subtração de fundo [McIvor 2000], rastreamento baseado em filtro de partículas [Goldenstein

¹ *website* da plataforma SIGUS: <http://www.gpec.ucdb.br/sigus>

² Mais informações em <http://www.eyetoy.com>

³ Os códigos-fonte podem ser obtidos em <http://svn.gpec.ucdb.br/libraries/sigus/src/>

⁴ Disponível em <http://rsb.info.nih.gov/ij/>

⁵ Disponível em <http://www.cs.waikato.ac.nz/~ml/weka/>

2004; Morais et al. 2005] e modelos de Markov oculto [Cunha and Velho 2002].

4 Protótipos Desenvolvidos

Foram desenvolvidos protótipos para demonstração de como podem ser aplicadas as técnicas de visão computacional na construção de jogos eletrônicos com a plataforma SIGUS. Todas as aplicações desenvolvidas são baseadas em interfaces visuais e o dispositivo de aquisição de imagens utilizado é a *webcam*. Os jogos podem ser controlados por diversos sinais visuais diferentes e essa característica os tornam acessíveis a várias pessoas, inclusive as pessoas com necessidades especiais.

Os protótipos desenvolvidos possuem uma estrutura semelhante pois são aplicações desenvolvidas com base na mesma biblioteca de desenvolvimento, o SIGUS. Todas as aplicações possuem fase de pré-processamento, fase de treinamento e fase de classificação, ou fase de execução.

Na fase de pré-processamento é realizado o tratamento das imagens para alimentação do sistema. É nessa fase, pré-processamento, que são feitas as escolhas dos métodos de segmentação de imagens e ajustes de seus parâmetros. Nessa fase é feita a escolha da técnica de extração de atributos e ajuste dos parâmetros dos extratores. Pode-se selecionar também a técnica de aprendizagem automática e as configurações do classificador. Os métodos de segmentação e extração de atributos são facilmente selecionáveis através de painéis de controle existentes na biblioteca de desenvolvimento, onde é feito o mapeamento dos valores desejados pelo usuário com os valores a serem utilizados no processamento.

Na fase de treinamento é feita a aprendizagem supervisionada das posturas com as ações a serem tomadas. Os algoritmos de aprendizagem automática são treinados com base nas técnicas de extração de características selecionadas na fase de pré-processamento. As posturas utilizadas para o treinamento podem ser baseadas em posições dos braços, posição da cabeça, configurações da mão, entre outras. Para cada postura definida pelo usuário, são extraídos os atributos da imagem que servem de entrada para o classificador escolhido.

A fase de classificação, também chamada de fase de execução, é responsável por classificar, com base no classificador usado na fase de treinamento, as posturas de entrada adquiridas pela *webcam*. O classificador vai relacionar a postura utilizada no treinamento que mais se adequa à postura de entrada para gerar uma saída de controle do software.

Todas as aplicações que serão descritas aqui foram implementadas na linguagem de programação JavaTM, mesma linguagem da biblioteca de desenvolvimento, para facilitar as implementações e para garantir a portabilidade. Os aplicativos desenvolvidos possuem interface gráfica modular para facilitar o controle de cada um dos elementos e fases contidas no projeto. Os jogos aqui descritos são apenas protótipos, visto que o alvo dessas implementações é demonstrar como se constrói uma aplicação utilizando o SIGUS.

A seguir serão comentadas as implementações de jogos bastante conhecidos que ganharam uma versão jogável utilizando visão computacional. O jogo da velha e Genius[®] são exemplos de jogos que foram desenvolvidos utilizando a biblioteca.

4.1 Jogo da velha

Foi desenvolvido um jogo da velha guiado por sinais visuais, onde o usuário faz o controle do software utilizando somente expressões faciais com o auxílio de uma *webcam*. O tabuleiro do jogo é uma matriz de 3x3 e por isso o usuário precisa realizar o treinamento para 9 regiões do tabuleiro, uma postura para cada região do tabuleiro. Na fase de execução, o jogador executa uma postura, que é automaticamente classificada pelo módulo de aprendizagem, e produz a marcação de um "X" na região, que durante a fase de treinamento, foi associada àquela expressão facial. A Figura 2 mostra a interface do jogo da velha guiado por expressões faciais.

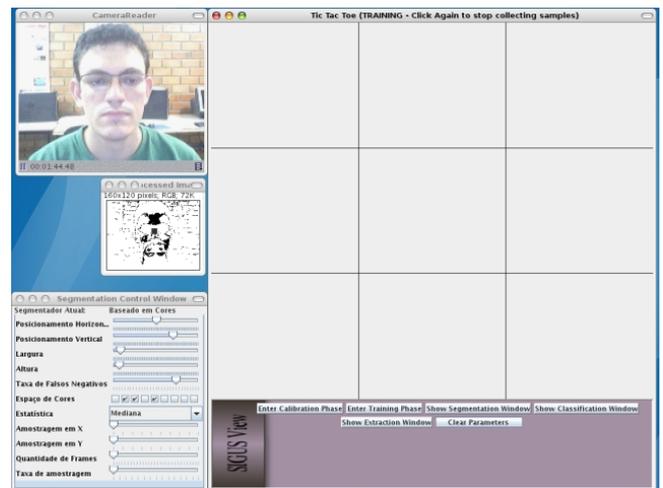


Figura 2: Interface do aplicativo TicTacToe

4.2 MemoSigus

Simon é um jogo de memória, muito famoso na década de 1980, conhecido no Brasil como Genius[®], e que foi lançado nesse país pela Estrela[®] Brinquedos. O tabuleiro do jogo é composto de quatro arcos, cada um de cor diferente, e para cada arco existe um respectivo som. O objetivo do jogo é acompanhar as seqüências que são geradas automaticamente e incrementalmente, isto é, na primeira vez é gerada uma seqüência de tamanho um, se o jogador acertar a primeira seqüência, é gerada uma nova seqüência com tamanho dois e assim por diante.

Foi implementado um jogo de memória com as mesmas características do Simon com a biblioteca SIGUS. Na versão para sinais visuais cada arco colorido do jogo original é representado por uma postura realizada com o corpo do jogador. O jogo segue a mesma estrutura do jogo original mas com alguns ajustes para melhorar a integração com a jogabilidade baseada em sinais visuais. Existe um sinalizador circular no centro da imagem do tabuleiro que quando o computador está gerando a seqüência fica em vermelho. No momento em que o jogador precisa executar as posturas, o sinalizador circular fica verde. A Figura 3 mostra a interface do jogo de memória desenvolvido.

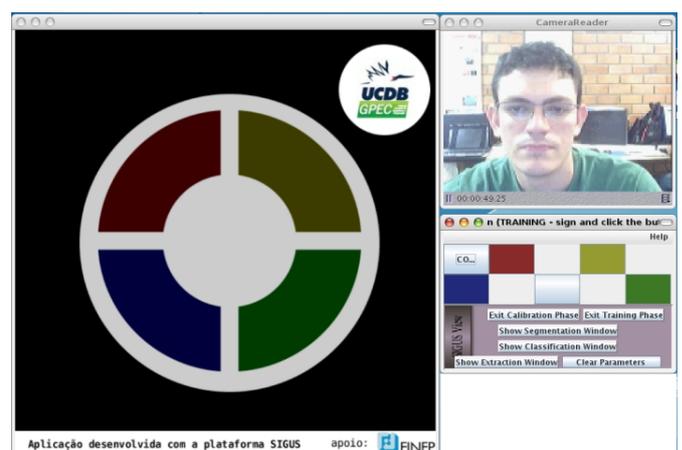


Figura 3: Interface do jogo da memória.

4.3 Wheelchair

Está incluído na plataforma SIGUS um simulador de um controlador de cadeira de rodas elétrica automática guiada por movimentos da cabeça. Este aplicativo mostra como se pode controlar um objeto virtual através de técnicas de visão computacional. Nesta aplicação são associados comandos de movimento do objeto virtual com mo-

vimentos da cabeça realizados pelo usuário. Os comandos gerados pelo usuário são para movimentar o objeto para a esquerda, para a direita, para frente ou para trás.

O aplicativo de controle de cadeira de rodas pode ser utilizado em jogos eletrônicos. O controle de avatares em um jogo serve de modelo para o modo de comandar uma cadeira de rodas utilizando sinais visuais. Essa forma de interação é bastante diferente das utilizadas atualmente. A interação com sinais visuais pode ser mais intuitiva e facilita a utilização do sistema por pessoas atualmente excluídas do meio digital.

Um aplicativo foi desenvolvido para demonstrar como se pode utilizar comando feitos com a cabeça para controlar um avatar em um ambiente virtual. Um pingüim em um ambiente tridimensional é controlado por expressões faciais. A Figura 4 mostra um usuário controlando o avatar utilizando posturas realizadas em frente a uma webcam.

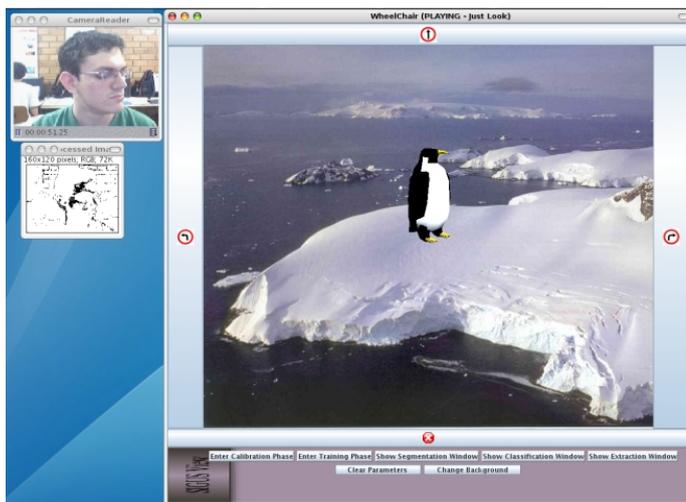


Figura 4: Interface do aplicativo Wheelchair.

5 Conclusões e Trabalhos Futuros

Muitas pessoas acabam sendo excluídas do mundo digital pelas dificuldades de interação com as interfaces mais utilizadas atualmente, mouse, teclado e *joystick*. Pessoas com necessidades especiais, por consequência, tornam-se excluídas do mundo do entretenimento digital pois além das interfaces serem impróprias para essas pessoas, exige-se muita coordenação motora pois os jogos são, em sua maioria, pouco intuitivos. A Visão Computacional pode contribuir para a solução desse problema, fazendo a inclusão de pessoas com necessidades especiais tanto no mundo digital como no acesso a tecnologias de entretenimento eletrônico.

Existem diversas aplicações das interfaces guiadas por sinais visuais para serem utilizadas em jogos: controle do mouse e reconhecimento de gestos e posturas. As interfaces com essas características podem ser mais intuitivas que as atuais e ainda favorecem o acesso a pessoas com necessidades especiais. Para esse fim, a biblioteca de desenvolvimento apresentada mostrou-se bastante versátil e de fácil utilização para o desenvolvimento dos aplicativos descritos anteriormente.

Há muitas aplicações que podem ser desenvolvidas para que a interação homem-computador baseada em sinais visuais se torne comum em jogos eletrônicos. Incluir algoritmos de visão computacional nos motores de desenvolvimento de jogos, para que possa se fazer controles dos elementos dos jogos por sinais visuais é um trabalho futuro que pode apresentar resultados satisfatórios.

Agradecimentos

Este trabalho recebeu apoio financeiro da Universidade Católica Dom Bosco, UCDB, da Agência Financiadora de Estudos e Projetos, FINEP e da Fundação de Apoio ao Desenvolvimento do En-

sino, Ciência e Tecnologia do Estado de Mato Grosso do Sul, FUNDECT. Os autores também contaram com bolsas do Conselho Nacional de Desenvolvimento Científico e Tecnológico, CNPQ, nas modalidades ITI-A, PIBIC e Produtividade em Desenvolvimento Tecnológico e Extensão Inovadora.

Referências

- CHRISTOPHER FAGIANI, MARGRIT BETKE, J. G. 2002. Evaluation of tracking methods for human-computer interaction. *Sixth IEEE Workshop on Applications of Computer Vision*, 121.
- CUNHA, A. M., AND VELHO, L. 2002. Hidden markov models. Tech. Rep. TR-02-02, IMPA, Rio de Janeiro, Brasil.
- DE PAULA, L. R. P., BONINI, R., AND DE MIRANDA, F. R. 2006. Camera combat - interação livre para jogos. *SBGAMES - V Brazilian Symposium on Computer Games and Digital* (November).
- FREEMAN, W. T., ANDERSON, D. B., BEARDSLEY, P. A., DODGE, C. N., ROTH, M., WEISSMAN, C. D., YERAZUNIS, W. S., KAGE, H., KYUMA, K., MIYAKE, Y., AND ICHI TANAKA, K. 1998. Computer vision for interactive computer graphics. *IEEE Computer Graphics and Applications* 18, 3, 42–53.
- GOLDENSTEIN, S. K. 2004. A gentle introduction to predictive filters. *Revista de Informatica Teórica e Aplicada (RITA)* 11, 61–89.
- MCIVOR, A. 2000. Background subtraction techniques. In *Proceedings of Image and Vision Computing*.
- MORAIS, E. F., CAMPOS, M. F. M., PADUA, F. L. C., AND CARCERONI, R. L. 2005. Particle filter-based predictive tracking for robust fish counting. In *SIBGRAPI '05: Proceedings of the XVIII Brazilian Symposium on Computer Graphics and Image Processing*, IEEE Computer Society, Washington, DC, USA, 367.
- PASSAFARO, E. L. 2002. Guia de acessibilidade em edificações. *CPA - Comissão Permanente de Acessibilidade da SEHAB*.
- PISTORI, H., MARTINS, P. S., PEREIRA, M. C., JR., A. A. C., AND NETO, J. J. 2006. Plataforma de apoio ao desenvolvimento de sistemas para inclusão digital de pessoas com necessidades especiais. In *IV Congresso Iberdiscap: Tecnologias de Apoio a Portadores de Deficiência*.
- PISTORI, H. 2004. Sigus - plataforma de apoio ao desenvolvimento de sistemas para inclusão digital de pessoas com necessidades especiais. <http://www.gpec.ucdb.br/sigus>.
- PISTORI, H. 2006. Computer vision and digital inclusion of persons with special needs: Overview and state of art. *CompIMAGE - Computational Modelling of Objects Represented in Images: Fundamentals, Methods and Applications* (October), 20–21.
- TEIXEIRA, J. M., FARIAS, T., MOURA, G., LIMA, J., PESSOA, S., AND TEICHRIB, V. 2006. Gefighters: an experiment for gesture-based interaction analysis in a fighting game. *SBGAMES - V Brazilian Symposium on Computer Games and Digital* (November).
- WITTEN, I. H., AND FRANK, E. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco.