

# Sistema para Captura do Movimento Humano 2D sem Utilização de Marcações Especiais

Paulo César Rodacki Gomes e Leandro Augusto Frata Fernandes

**Abstract--** The capture of human movement through an optic system is extremely efficient, however the use of special markings in the body of the actor can be a limiting factor. The present paper presents a proposal for implementation of an optic system for motion capture of human actors without the use of special markings. After a description of the process, the proposal's viability is presented through the implementation of the first three stages of the process. The next step of the ongoing research includes the implementation of the remaining stages.

**Index Terms--** Computer graphics, Image motion analysis, Image region analysis, Machine vision.

## I. INTRODUÇÃO

O termo Captura de Movimento refere-se a técnicas de reconhecimento do movimento de objetos. Esse não é o único termo utilizado para referenciar tal tecnologia, talvez o termo mais aplicado seja *Motion Capture*, ou simplesmente sua abreviação MoCap. De acordo com a tecnologia empregada [1], os sistemas de MoCap podem ser divididos em quatro grupos [2]: mecânico, óptico, eletromagnético e acústico. Usualmente, sistemas de captura de movimento que utilizam dispositivos mecânicos, eletromagnéticos ou acústicos apresentam uma série de problemas, tais como a necessidade do uso de equipamentos especiais e a ocorrência de perda de dados de movimento. O presente trabalho apresenta uma alternativa óptica de captura do movimento humano sem utilização de marcações especiais sobre o corpo do ator. Na seção 2 deste artigo são apresentadas as aplicações dos sistemas de MoCap; a seção 3 aborda cada uma das etapas necessárias para o processo completo de captura; na seção 4 é exposto um protótipo de sistema óptico de MoCap que implementa as etapas iniciais do processo de captura; finalmente, na seção 5 são exibidas as conclusões provenientes deste trabalho.

## II. APLICAÇÕES DOS SISTEMAS DE CAPTURA DE MOVIMENTO

Na animação, MoCap é comumente aplicado em situações onde é necessário se ter uma movimentação com bom grau de

realismo ou quando muitos movimentos distintos são atribuídos a um mesmo personagem [3]. O uso de MoCap na indústria cinematográfica é crescente, pois torna o trabalho de animação mais fácil. A rapidez de produção e o realismo da animação gerada fazem desse tipo de processo uma ferramenta extremamente interessante para as produtoras de efeitos especiais. Emissoras de televisão também utilizam dessa tecnologia, e já foram criados programas apresentados por “atores virtuais”.

## III. SISTEMA ÓPTICO DE MoCAP EM TEMPO REAL

Normalmente, os sistemas ópticos fazem uso de marcações especiais sobre o corpo do ator em cena para facilitar o processo de captura. No entanto, os autores deste trabalho abordam o tipo de sistema óptico que não utilize marcações artificiais sobre o corpo do ator. Ao invés disso, utilizam-se técnicas especiais para estimação da postura humana, que podem ser subdivididas nas seguintes etapas: (A) captura dos dados de origem vindos de câmeras de vídeo; (B) remoção do fundo da cena e detecção de objetos em movimento; (C) análise da silhueta dos objetos detectados para localização da provável figura humana; (D) segmentação da silhueta humana para identificação do posicionamento 2D das partes do corpo; (E) identificação do posicionamento 3D das partes do corpo; e (F) reconstituição dos movimentos do ator em uma marionete virtual tridimensional.

Para a execução destas etapas, os autores acreditam que será necessária a divisão do sistema em dois módulos: o módulo 2D, que deve executar as etapas A a D, e o módulo 3D, para a execução das etapas E e F. Cada câmera fica então responsável por capturar um diferente ângulo do ator em cena e é ligada a um módulo 2D exclusivo, sendo os resultados dos processamentos desses módulos 2D enviados ao módulo 3D centralizador.

O desenvolvimento do protótipo distribuído, composto por estes dois módulos, é uma das atividades atualmente em andamento neste trabalho de pesquisa. A seguir, é apresentada uma discussão sobre cada uma das etapas, em especial sobre as etapas A a C.

### A. Captura dos Dados de Entrada

Os dados de entrada são obtidos com o uso de câmeras de vídeo, que captam imagens coloridas a uma determinada frequência, medida em quadros por segundo (fps). As câmeras de vídeo medem a intensidade da luz refletida pelos objetos no ambiente sobre uma matriz de espaçamentos regulares [4].

---

P. C. R. Gomes é professor do Departamento de Sistemas e Computação da FURB – Universidade Regional de Blumenau, Campus IV, Rua Braz Wanka, 238, Bairro Vila Nova, CEP 89.035-160, Blumenau, SC, Brasil. (e-mail: rodacki@inf.furb.br).

L. A. F. Fernandes é analista/programador do Departamento de Desenvolvimento de Sistemas da JJW Informática Ltda., Av. Getúlio Vargas, 370 Centro CEP: 89120-000 Timbó, SC, Brasil. (e-mail: laffernandes@hotmail.com).

### B. Remoção do Fundo da Cena e Detecção de Objetos em Movimento

Uma das habilidades fundamentais da visão humana é a fidelidade à cor. Os seres humanos tendem a associar uma cor constante a um objeto mesmo que esse esteja sob a influência de variações de luminosidade [5]-[7]. A fidelidade à cor é importante para a remoção do fundo da cena, pois permite um método correto para classificação dos objetos na cena independentemente da projeção de sombras sobre eles. Com isso, sombras de objetos em movimento são ignoradas. A remoção, ou subtração do fundo consiste em subtrair a imagem atual de uma imagem usada como referência, contendo apenas o fundo da cena e construída a partir de uma seqüência de imagens durante um período de treinamento. A subtração deixa visível apenas objetos não estáticos e objetos que não estavam na cena original. A subtração faz uso de um modelo de cor capaz de distinguir a variação da cor da variação da luminosidade, a partir de uma imagem de referência e da imagem atual. Esse modelo de cor é mostrado na Fig. 1, onde os três eixos do sistema de coordenadas representam os valores R, G e B de um espaço de cores;  $i$  é o índice da imagem atual;  $O$  é a origem, onde as coordenadas para R, G e B são iguais a zero;  $E = [ER, EG, EB]$  representa o conjunto de valores RGB esperados para o pixel na imagem de referência;  $I_i = [IR(i), IG(i), IB(i)]$  representa o conjunto de valores RGB de um pixel na imagem  $i$ , da qual se deseja subtrair o fundo; e o segmento de reta  $OE$ , que passa pelos pontos  $O$  e  $E$ , é a “linha cromática esperada”, composta por pontos que representam o fundo da cena.

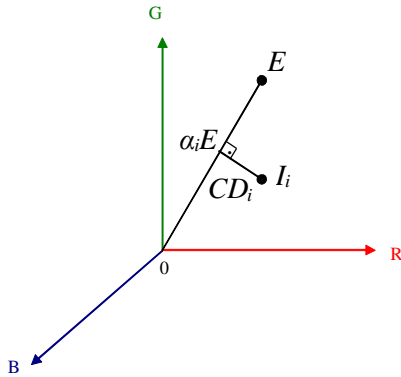


Fig. 1. Modelo de cor proposto no espaço RGB. Os eixos representam os valores das componentes “Red”, “Green” e “Blue” no espaço de cores.

Usando-se a relação entre  $I_i$  e  $OE$ , são encontrados os valores da distorção da luminosidade ( $a_i$ ) e da distorção da cor ( $CD_i$ ), sendo  $CD_i$  a distância ortogonal entre a cor observada em  $I_i$  e a linha cromática esperada. A grandeza  $a_i$  é um escalar, caso seu valor seja igual a um, indica a luminosidade idêntica ao fundo original, um valor maior que um indica o fundo mais iluminado e um valor menor que um indica o fundo menos iluminado. O vídeo é constituído por uma seqüência de imagens *raster* contendo o fundo da cena, objetos em movimento e variações na iluminação. Portanto, o processo

passa pelos seguintes passos para que seja realizada a subtração do fundo da cena:

- 1) **Modelagem do fundo:** a imagem de referência  $E$  é construída a partir de uma série de imagens contendo apenas do fundo da cena;
- 2) **Seleção do limiar:** onde são escolhidos de forma automática os valores mínimo e máximo aceitos para a distorção de luminosidade  $a$  e de cor  $CD$ ;
- 3) **Operação de subtração ou classificação do pixel:** onde os pixels das novas imagens  $I$  informadas são classificados como (i) em movimento, (ii) pertencente ao fundo original, (iii) pertencentes ao fundo com sombra ou (iv) ao fundo mais iluminado.

A classificação dos pixels gera uma imagem binária que indica se cada pixel faz parte do fundo (seja ele original, sombreado ou iluminado) ou em movimento. As formas de cálculo apresentadas para a modelagem do fundo, seleção do limiar e operação de subtração são contribuições originais do presente trabalho. Baseado no modelo de cor da Fig. 1, este artigo propõe quatro métodos de remoção do fundo, um não estatístico (método do mínimo e máximo valor RGB) e três estatísticos (distribuição Normal Padronizada, a distribuição de Student e Empírico). Para os quatro métodos, a modelagem do fundo obedece a mesma regra: para cada pixel da imagem é mapeado um ponto médio  $E$  calculado a partir das cores atribuídas a esse pixel nos  $n$  quadros armazenados como amostra, esse ponto médio representa a cor esperada para o fundo da cena naquele pixel. O resultado do cálculo de  $E$  pode ser visualizado nas Fig. 2 e 3.

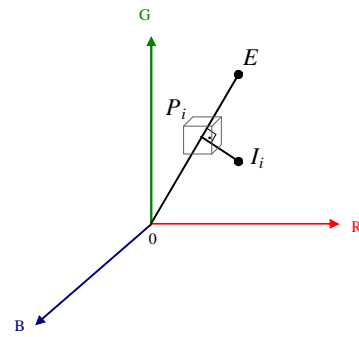


Fig. 2. Representação gráfica da modelagem do fundo da cena e seleção do limiar para remoção do fundo pelo método não estatístico.

Na seleção do limiar, o método de mínimo e máximo valor RGB cria três intervalos de confiança (um para R, um para G e um para B) com base no deslocamento entre as cores atribuídas ao pixel nos  $n$  quadros armazenados como amostra, e suas projeções sobre a linha cromática esperada. Os métodos estatísticos, por sua vez, criam um único intervalo de confiança com base na média das distâncias entre a linha cromática esperada e as cores atribuídas ao pixel nos  $n$  quadros armazenados como amostra, a essa média é adicionado um erro de amostragem calculado com a aplicação

da distribuição Normal Padronizada, distribuição de Student ou pelo método Empírico. No modelo de cor utilizado, o método de mínimo e máximo valor RGB gera um paralelepípedo em torno da projeção  $P_i$ , conforme a Fig. 2.

Os métodos estatísticos, por sua vez, geram um cilindro ao redor da linha cromática esperada cujo raio determina a distorção de cor mínima aceita (zero) e a distorção de cor máxima aceita (comprimento do raio), conforme a Fig. 3.

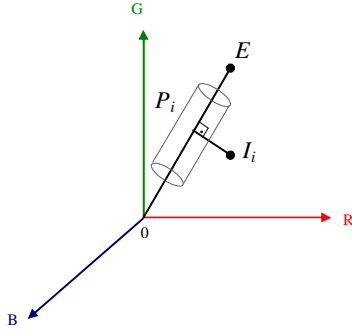


Fig. 3. Representação gráfica da modelagem do fundo da cena e seleção do limiar para remoção do fundo pelo método estatístico.

O terceiro passo da remoção do fundo da cena é a operação de subtração ou a classificação do pixel em fundo da cena ou objeto em movimento. Nessa etapa é testado cada pixel de uma imagem  $I_i$ . O pixel é classificado como fundo da cena caso fique dentro dos limiares pré-calculados, caso contrário o pixel é classificado como objeto em movimento.

O cálculo da cor esperada como fundo da cena para um determinado pixel é feito em (1), onde:  $E$  é a cor de coordenadas RGB esperada para o fundo;  $n$  é o número de quadros armazenados como amostra e que contém apenas o fundo da cena;  $i$  é o índice do quadro na amostra; e  $I$  é a cor de coordenadas RGB do pixel no  $i$ -ésimo quadro da amostra.

$$E = \frac{1}{n} \sum_{i=0}^{n-1} I_i \quad (1)$$

O cálculo do ponto de projeção ortogonal do pixel de  $I$  sobre a linha cromática esperada é feito em (2) e (3), onde:  $aux$  é uma variável auxiliar de cálculo;  $E$  é a cor de coordenadas RGB esperada para o fundo da cena;  $I$  é a cor de coordenadas RGB que está sendo projetada; e  $P$  é o ponto sobre a linha cromática esperada resultante dessa projeção.

$$aux = \frac{(E_R \times I_R) + (E_G \times I_G) + (E_B \times I_B)}{I_R^2 + I_G^2 + I_B^2} \quad (2)$$

$$P = aux \times (I_R, I_G, I_B) \quad (3)$$

No método não estatístico, a partir da relação entre as coordenadas dos pontos  $P$  e  $I$ , já podem ser traçados os três intervalos de confiança para a classificação dos pixels.

Nos métodos estatísticos, é necessário que seja calculada para cada pixel: a distância  $d$  entre os pontos  $P$  e  $I$  através de (4).

$$d = \sqrt{(P_R - I_R)^2 + (P_G - I_G)^2 + (P_B - I_B)^2} \quad (4)$$

Após o cálculo de todos os valores de distância  $d$  para os pixels das imagens da amostra, é calculada a média amostral dessas distâncias  $\bar{d}$  (5) e o desvio padrão amostral  $s_d$  (6), e aplicada a distribuição Normal Padronizada, a distribuição de Student ou o método Empírico.

$$\bar{d} = \frac{1}{n} \sum_{i=0}^{n-1} d_i \quad (5)$$

$$s_d = \sqrt{\frac{1}{n-1} \sum_{i=0}^{n-1} (d_i - \bar{d})^2} \quad (6)$$

O método Empírico foi desenvolvido originalmente neste trabalho e mantém a forma de cálculo da distribuição de Student, no entanto a variável  $t$  de Student é substituída por um valor ajustado manualmente para cada vídeo. Nos vídeos utilizados na pesquisa valores entre 100 e 180 mostraram-se adequados.

O resultado obtido na remoção do fundo é exibido na Fig. 4, onde a região demarcada pela cor preta indica áreas da imagem de vídeo classificadas como cenário e as regiões demarcadas pela cor branca indicam os objetos em movimento. A imagem da Fig. 4 foi gerada com a aplicação do método Empírico. Os outros métodos retornam resultado semelhante, porém com maior quantidade de ruído.



Fig. 4. Resultado da remoção do fundo da cena e detecção de objetos em movimento.

No presente trabalho os autores optaram por não utilizar a identificação dos objetos em cena através da técnica tradicional conhecida como *Chroma-Keying* [8] onde a cena é filmada sobre um fundo azul ou verde e, basicamente, os elementos não azuis ou verdes da imagem são considerados objetos em cena. A justificativa para isso é a intenção dos autores de também gravar imagens fora de um estúdio de gravação, em um cenário qualquer onde a constante azul ou verde em geral não está presente. Como pode ser observado na Fig 5, a maior parte das cenas gravadas para o presente trabalho foram preparadas sobre cenário predominante azul, no entanto elementos extras, tais como uma poltrona e uma escada de metal foram adicionados propositalmente para o teste dos algoritmos.

### C. Análise das Silhuetas e Localização da Provável Figura Humana

A imagem binária (em preto e branco) gerada pela etapa anterior mostra quais pixels foram classificados como fundo da cena ou em movimento. No entanto, os pixels classificados como em movimento nem sempre fazem parte da figura humana, eles podem ser o resultado de falhas no algoritmo de remoção do fundo ou do dispositivo de captura. Os conjuntos de pixels obtidos a partir dessa falha são chamados de “ruído”. Portanto, é preciso desenvolver um mecanismo que distinga o conjunto de pixels que fazem ou não fazem parte da figura humana.

O critério pelo qual é feita a localização da provável figura humana necessita de três parâmetros: largura e altura mínima para que a silhueta seja considerada uma pessoa e a matriz resultante da remoção do fundo da cena, informada a todo novo quadro do vídeo que está sendo tratado e que contém a classificação de cada pixel da imagem, podendo ser fundo da cena (valor igual a zero) ou pixel em movimento (valor igual a um). Como resultado, é retornado um polígono contendo os pontos do contorno interno da pessoa identificada, ou um valor nulo, caso nenhuma pessoa tenha sido identificada. Esse método pode ser dividido em três etapas:

- 1) É utilizado um algoritmo denominado *Contour Following* [9], para identificar os pixels pertencentes ao contorno de um conjunto de pixels vizinhos;
- 2) O contorno que melhor atende aos pré-requisitos (altura e largura mínima para ser considerada humana) é identificado como “figura humana”, os demais passam a ser considerados como sendo parte do fundo da cena;
- 3) Caso mais de uma silhueta atenda os pré-requisitos de altura e largura mínima, então a silhueta de maior área é classificada como humana.

Um exemplo do resultado obtido na análise das silhuetas é exibido na Fig. 5, onde o polígono ao redor da figura humana representa os pontos que fazem parte do contorno do ator.



Fig. 5. Resultado da análise das silhuetas encontradas.

### D. Segmentação da Silhueta para Identificação 2D das Partes do Corpo

Para a segmentação da silhueta, devem ser levados em consideração a topologia humana, a postura do ator e o ângulo

de visão da câmera. O objetivo dessa etapa é de demarcar, sobre a silhueta encontrada, as partes primárias (cabeça, mãos, pés, e torso) e secundárias (cotovelos, joelhos, ombros, axilas, quadril, e parte superior das costas) do corpo do ator. As partes primárias são utilizadas como referência para a animação da marionete virtual 3D e as secundárias são úteis na correta identificação das partes primárias, visto que nem sempre as partes primárias fazem parte do contorno da silhueta, ao contrário das partes secundárias.

### E. Identificação do Posicionamento 3D das Partes do Corpo

A utilização de uma única câmera não é suficiente para a obtenção do posicionamento 3D das partes do corpo do ator. Contudo, ao utilizar mais de uma câmera, é possível capturar as coordenadas 2D e estimar o posicionamento 3D dessas partes [10]. O uso de mais de uma câmera também reduz a ocorrência de “oclusão”, onde uma parte do corpo não é filmada, pois outra parte ou outro elemento em cena fica em seu caminho na obtenção das imagens.

### F. Reconstituição dos Movimentos do Ator

A reconstituição dos movimentos do ator pode ser feita com a utilização de uma marionete virtual 3D cuja topologia seja semelhante à do corpo humano. Devido à grande quantidade de cálculos executados nas etapas de análise das imagens e a necessidade de se manter a execução do vídeo em tempo real, é inevitável a perda de alguns dos quadros do vídeo. Pressupõe-se então que o resultado obtido da triangulação quadrática não mantém a amostragem de trinta quadros por segundo.

Para solucionar o problema da perda de quadros, pode ser utilizada a técnica de animação por quadros-chave (keyframing), onde os quadros analisados seriam os quadros-chave para a geração de quadros intermediários. É importante lembrar que, quando o objetivo é gerar a animação da marionete sincronizada com os movimentos do ator, a interpolação de quadros torna-se inviável, pois esse processo atrasaria a reconstituição dos movimentos do ator na marionete.

## IV. IMPLEMENTAÇÃO DE UM PROTÓTIPO

Na implementação do protótipo foi utilizada a linguagem de programação Java, em seu pacote Java 2 SDK Standard Edition, versão 1.3.0\_02, e a API Java Media Framework (JMF), versão 2.1.1. A JMF proporciona um alto nível de abstração no que diz respeito à captura, conferência e reprodução de mídia baseada em tempo real [11].

A abstração proporcionada por essa API foi o principal motivo da escolha da linguagem Java para a confecção do protótipo. O escopo do protótipo apresentado aqui é restrito ao módulo 2D. A Fig. 6 ilustra as três primeiras etapas do processo, onde: (A) são imagens do fundo da cena utilizadas na construção da imagem de referência; (B) é a imagem de referência; (C) é a imagem da qual deseja-se extrair o fundo da cena; (D) é o resultado da subtração do fundo da cena; (E) é a silhueta humana identificada.

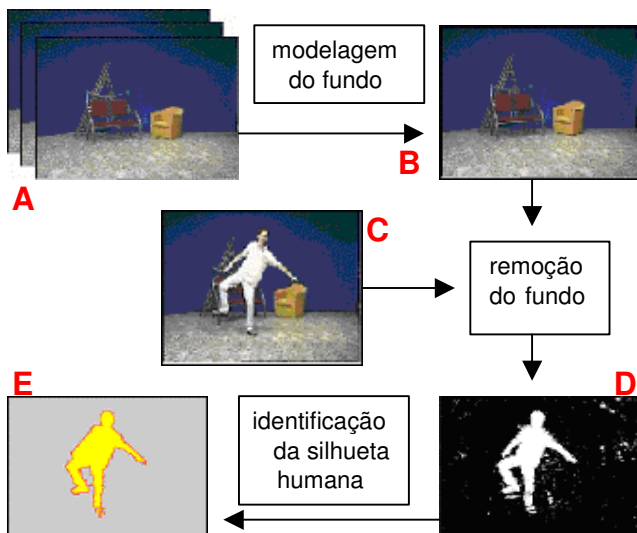


Fig. 6. Ilustração das etapas da captura do movimento.

Para a construção do protótipo foram considerados alguns requisitos básicos, são eles: executar tratamento sobre uma trilha de vídeo obtida de um arquivo ou dispositivo de captura; analisar a trilha de vídeo e implementar as etapas de remoção do fundo da cena e análise das silhuetas e identificação da figura humana; permitir a implementação de mais que um método para cada uma das etapas.

Para o bom funcionamento do protótipo, a câmera deve ter posicionamento estático e não pode sofrer alterações no foco. O cenário pode ser qualquer ambiente estático e não refletivo, sob qualquer tipo de iluminação, desde que essa iluminação não sofra variações de intensidade nem de posicionamento da fonte; o único objeto em movimento deve ser uma figura humana (o ator). A trilha de vídeo deve conter apenas o fundo da cena durante o tempo de aprendizagem, que varia de computador para computador, após esse tempo de aprendizagem o ator está habilitado a entrar em cena.

Os dados de entrada são obtidos diretamente das câmeras de vídeo de alta resolução ou de arquivos Audio Video Interleaved (AVI), que utilizem o codec Indeo® video 5.04, com 320x240 pixels, 24bits de profundidade de cor, a uma frequência de 30 fps. No protótipo, a obtenção e serialização dos dados de entrada são feitas exclusivamente pelo JMF. Como entrada, a etapa 2 recebe um vetor de bytes com os dados do quadro atual do vídeo.

Visualmente, a etapa de remoção do fundo das cenas e detecção dos objetos em movimento demonstrou melhores resultados no método empírico e no método de mínimo e máximo valor RGB. Já os métodos estatísticos que utilizam a distribuição Normal Padronizada e a distribuição de Student apresentaram um resultado com grande quantidade de ruído.

Os resultados dessa análise visuais são comprovados numericamente, onde em amostras de vídeos diferentes o método Empírico teve um melhor aproveitamento (96,32% e 97,50% de acerto), seguido pelo mínimo e máximo valor RGB (92,17% e 92,73% de acerto), distribuição Normal (86,21% e

88,36% de acerto) e distribuição de Student (64,80% e 67,76% de acerto), conforme indicado na Fig 7.

Em relação à quantidade de quadros aproveitados, tanto o método Empírico quanto o método que utiliza a distribuição de Student aproveitam melhor os quadros do vídeo. Em seguida, estão os métodos de mínimo e máximo valor RGB e distribuição Normal. Na análise das silhuetas para localização da figura humana, o método empírico obteve um melhor aproveitamento dos quadros, seguido pela distribuição Normal, mínimo e máximo valor RGB e distribuição de Student. O bom desempenho da análise da silhueta dos objetos em movimento para identificação da figura humana mostrou-se diretamente proporcional ao desempenho na classificação dos pixels da etapa de remoção do fundo da cena, sendo confiável quando obedecidos os pré-requisitos traçados na introdução da seção 4.

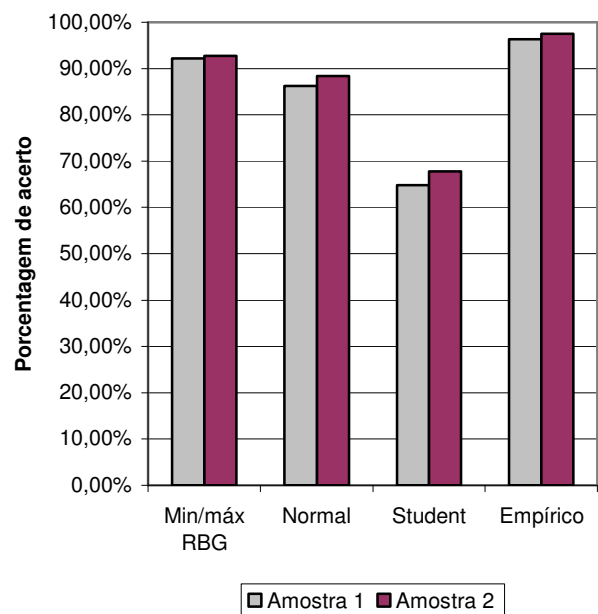


Fig. 7. Gráfico comparativo da porcentagem de acerto dos quatro métodos implementados para remoção do fundo das cenas e detecção dos objetos em movimento.

A etapa de remoção do fundo das cenas mostrou-se pesada em computadores de baixa capacidade de processamento e que dispõem de pouca memória. Melhores resultados foram obtidos em um Pentium III 900Mhz com 128Mb de memória. O algoritmo utilizado na etapa de análise da silhueta dos objetos em movimento para identificação da figura humana mostrou-se pouco exigente nos quesitos de processamento e memória.

Informações detalhadas a respeito da implementação e comparação de desempenho entre os métodos empregados na remoção do fundo da cena são descritas em [12].

## V. CONSIDERAÇÕES FINAIS

Este trabalho propõe um novo método para implementação de sistemas ópticos de captura de movimento humano sem marcações especiais. Neste trabalho, a viabilidade da proposta



é apresentada através da implementação de um protótipo utilizando o Java Media Framework. Não se trata de um protótipo completo para captura do movimento humano, visto que as etapas de detecção das partes do corpo do ator, cálculo das coordenadas 3D de seu movimento e aplicação destas coordenadas em uma marionete virtual ainda estão em desenvolvimento.

Para que seja possível a obtenção de resultados coerentes no protótipo, observa-se a relevância na qualidade de definição das imagens obtidas pelas câmeras e na qualidade no processamento dessas imagens. De modo geral, o sistema implementado é de fácil configuração, fácil calibragem, não restringe os movimentos do ator e executa a captura tanto movimentos sutis quanto movimentos amplos. Para movimentos rápidos, é necessária a utilização de uma plataforma de maior poder de processamento.

Como proposta de continuação da presente pesquisa, encontra-se o desenvolvimento de técnicas para a identificação das partes da figura humana (segmentação da silhueta e identificação do posicionamento 2D das partes do corpo do ator; identificação do posicionamento 3D das partes do corpo do ator e animação da marionete virtual 3D que reproduz os movimentos do ator). Outros objetivos propostos para a continuidade do trabalho são o aprimoramento da remoção do fundo da cena, permitindo o uso de câmeras de posicionamento não estático, e a expansão do método de identificação da figura humana para que esse suporte mais que um ator por vez.

## VI. REFERÊNCIAS BIBLIOGRÁFICAS

- [1] FURNISS, Maureen. Motion capture. In: MEDIA IN TRANSACTION CONFERENCE, 1., 1999, Cambridge. Anais eletrônicos... Cambridge: Massachusetts Institute of Technology, 1999. Disponível em: <<http://media-in-transition.mit.edu/articles/furniss.html>>. Acesso: 7 nov. 2001.
- [2] TRAGER, Wes. *A practical approach to motion capture*: acclaim's optical motion capture system, Paris, maio 1997. Disponível em: <[http://www.css.tayloru.edu/instrmat/graphics/hypgraph/animation/motion\\_capture/motion\\_optical.htm](http://www.css.tayloru.edu/instrmat/graphics/hypgraph/animation/motion_capture/motion_optical.htm)>. Acesso em: 23 out. 2001.
- [3] PYROS PICTURES INC. *Motion capture FAQ*, California, mar. 2001. Disponível em: <[http://www.pyros.com/html/mocap\\_faq.html](http://www.pyros.com/html/mocap_faq.html)>. Acesso em: 10 nov. 2001.
- [4] DAVIS, Larry. *Machine vision systems*, Germantown, 1999. Disponível em: <<http://www.umiacs.umd.edu/~lsd/426/426.html>>. Acesso em: 16 out. 2001.
- [5] DAVIS, Larry et al. *Multi-perspective analysis of human action*, Germantown, [2000?]. Disponível em: <<http://www.cs.umd.edu/users/thanarat/Research.html>>. Acesso: 23 out. 2001.
- [6] HORPRASERT, Thanarat; HARWOOD, David; DAVIS, Larry S. *A robust background subtraction and shadow detection*, Germantown, [2000a?]. Disponível em: <<http://www.cs.umd.edu/users/thanarat/Research/bgs.html>>. Acesso em: 23 out. 2001.
- [7] HORPRASERT, Thanarat; HARWOOD, David; DAVIS, Larry S. *A static approach for real-time robust background subtraction and shadow detection*, Germantown, [2000b?]. Disponível em: <<http://www.cs.umd.edu/users/thanarat/Research/bgs.html>>. Acesso em: 23 out. 2001.

- [8] SMITH, A. R.; BLINN, J. F. Blue screen matting. In: ACM SIGGRAPH 96, 1996. Anais... Addison Wesley: ACM, 1996. p. 259-268.
- [9] KANSAS UNIVERSITY. *The KUIM image processing system*, Lawrence, mai. 1999. Disponível em: <<http://www.cim.mcgill.ca/~sbouix/kuim/>>. Acesso em: 01 mai. 2002.
- [10] HORPRASERT, Thanarat et al. *Real-time 3D motion capture*, Germantown, [2000c?]. Disponível em: <<http://www.cs.umd.edu/users/thanarat/Research.html>>. Acesso: 23 out. 2001.
- [11] SUN MICROSYSTEMS INC. *Java™ media framework API guide*. Mountain View: Sun, 1999.
- [12] FERNANDES, Leandro Augusto Frata. *Protótipo de sistema óptico de captura do movimento humano, sem a utilização de marcações especiais*. 2002. 99 f. Trabalho de Conclusão de Curso (Bacharelado em Ciências da Computação) – Centro de Ciências Exatas e Naturais, Universidade Regional de Blumenau, Blumenau.



**Paulo César Rodacki Gomes** nasceu em Blumenau, SC, em 14 de outubro de 1969. Concluiu sua graduação em Engenharia Civil na UFPR – Universidade Federal do Paraná em 1992. Recebeu títulos de Msc e Dr. em Engenharia (Engenharia Civil) e Informática (Computação Gráfica) pela PUC-Rio, Pontifícia Universidade Católica do Rio de Janeiro, em 1995 e 1999, respectivamente. Sua experiência profissional inclui participação em projetos de pesquisa e desenvolvimento no TeCGraf – Grupo de Tecnologia em Computação Gráfica e no ICAD/IGames – Laboratório de CAD e Jogos Inteligentes da PUC-Rio. Trabalhou na área de Computação Gráfica e Publicidade Virtual na Rede Globo de Televisão, além de ter prestado consultoria na área de CAD para empresas como Embraer, CBPO e Construtora Odebrecht. Desde agosto de 2000, atua como professor e pesquisador do quadro da FURB – Universidade Regional de Blumenau.



**Leandro Augusto Frata Fernandes** nasceu em São Paulo, SP, em 27 de fevereiro de 1981. Concluiu o curso de Bacharelado em Ciência da Computação na FURB – Universidade Regional de Blumenau em 2002. Atuou na área de desenvolvimento de sistemas comerciais e de automação. Atualmente é analista/programador da JJW Informática Ltda., em Timbó, SC.