

Motion Capture

Introdução à Tecnologia

Por:

Fernando Wagner Serpa Vieira da Silva
nando@lcg.ufrj.br

LCG – Laboratório de Computação Gráfica - COPPE/SISTEMAS
Universidade Federal do Rio de Janeiro

1 Introdução

O conceito de Animação é antigo, tendo sido pela primeira vez utilizado pelo cartunista Winsor McKey, em 1911. Winsor conseguiu obter o movimento de um personagem desenhando-o em múltiplos pedaços de papel, que eram visualizados segundo uma certa taxa de amostragem.

A evolução das técnicas de animação se deu vagarosamente até a utilização de computadores como ferramenta de desenho. Para produzir o desenho animado “Branca de Neve e os Sete Anões”, os estúdios Walt Disney precisaram desenhar cerca 250.000 quadros, o que tornava impraticável produções desse tipo. Foi nos estúdios Disney que surgiram técnicas como *keyframing*, que ainda são utilizadas nos modernos sistemas de animação.

O uso de computadores facilita bastante o processo de criação de animações pois todo o trabalho de interpolação entre os quadros-chave de uma animação é realizado automaticamente. Os animadores profissionais têm à disposição todo um conjunto de técnicas de animação como *keyframing*, simulação e cinemática. No entanto, certos tipos de movimentos encontrados no mundo real são extremamente complexos para serem representados por meio destas técnicas, devido ao seu alto grau de complexidade. Técnicas como *keyframing* e cinemática são impraticáveis para representar movimentos como o caminhar humano, devido à enorme quantidade de articulações existentes; e o uso de simulação numérica ainda é computacionalmente caro. A solução encontrada foi capturar os movimentos dos objetos reais e então inserí-los no modelo tridimensional criado no computador.

Antes de aparecer pela primeira vez em produtos comerciais, em 1993, uma das poucas formas de capturar movimentos para gerar animações era através de um processo chamado *rotoscoping*, que consistia em desenhar o movimento dos personagens animados sobre imagens de atores reais realizando o movimento desejado, gerando assim animações bidimensionais. O uso de *rotoscoping* para a geração de animações 3D era limitado, uma vez que os animadores trabalhavam apenas com imagens 2D e não possuíam nenhuma informação tridimensional.

Uma outra forma criativa e inteligente de representar movimentos humanos no computador foi desenvolvida no laboratório de computação gráfica do Instituto de Tecnologia da Universidade de Nova Iorque. A técnica consistia em refletir na tela do computador imagens de um dançarino real, que serviriam de guia para a animação de um ator sintético no computador, utilizando técnicas de cinemática tradicional e *keyframing*. A reflexão da imagem na tela era feita através de um espelho especial.

Evolução da técnica de *rotoscoping*, Motion Capture, também conhecida como Performance Animation, consiste em capturar a posição e/ou orientação de objetos reais através de processos óticos ou magnéticos. O conjunto de dados capturados contendo a informação sobre os movimentos é inserido e mapeado nos modelos 3D dos objetos no computador. Esta tecnologia é recente e ainda possui algumas limitações, porém atualmente sua utilização é ampla, e vai desde filmes como o *Toy Story* da Pixar, até jogos de computador como o *Virtua Fighter*, da Sega.

Neste texto usaremos a palavra Motion Capture ou sua abreviação MC, para representar o processo de captura em geral. O termo Performance Animation é geralmente utilizado para designar o processo onde os dados capturados do ator são

utilizados diretamente (em tempo real) na animação como, por exemplo, o personagem animado *Moxy*, da Cartoon Network.

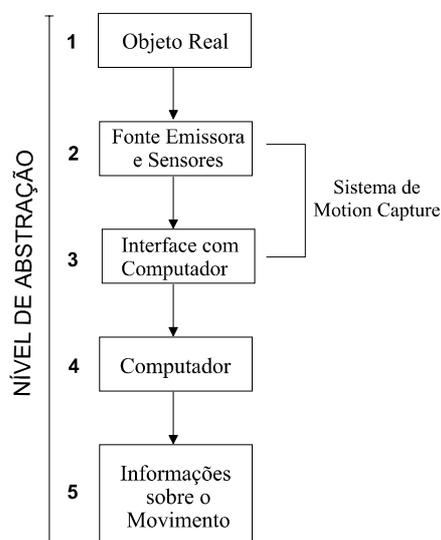
É importante ressaltar que, apesar de ser amplamente utilizada para a captura de movimentos humanos, Motion Capture pode ser utilizada para capturar o movimento de praticamente qualquer objeto existente no nosso mundo. Neste trabalho, daremos ênfase a movimentos capturados de seres humanos.

2 *Sistemas de Motion Capture*

Nesta seção faremos uma discussão sobre o estágio atual da tecnologia e dos equipamentos utilizados no processo de captura dos movimentos. Podemos dividir os sistemas de Motion Capture em quatro categorias: acústicos, mecânicos, óticos e magnéticos. Atualmente, os sistemas mais utilizados são os baseados nas tecnologias ótica e magnética, e serão o alvo principal de nosso estudo.

2.1 *Estrutura*

Sistemas de Motion Capture são, por definição, sistemas que geram para o computador informações que representam as medidas físicas do movimento capturado. O processo de captura consiste, de uma forma geral, em vestir um ator com uma roupa especial, onde serão posicionados refletores (sistemas óticos) ou transmissores (sistemas magnéticos). Tais marcadores são geralmente posicionados nas chamadas “articulações universais”, que são as 19 posições que fornecem o mínimo de precisão para a representação de um movimento humano [Blinn 87]. Em sistemas mecânicos como o *Animatton*, da Analogus, a roupa especial contém diversos potenciômetros, também posicionados estrategicamente, que fornecerão as posições e orientações das principais articulações do performer. Em geral, os sistemas de captura seguem o seguinte pipeline:



Algumas das etapas acima podem ser omitidas, dependendo da tecnologia utilizada no processo. Em 1993, Kalawsky [Kalawsky 93] propôs o seguinte conjunto de medidas para a especificação de sistemas de captura de movimento humano:

Medidas descritivas

- Processo de captura utilizado (acústico, eletromagnético, ótico ou mecânico)
- Localização dos sensores e marcadores (sensores e marcadores no corpo / fora do corpo)
- Parte do corpo processada (espinha, cabeça, ombros, braços, pernas, dedos, íris etc.)

Medidas estáticas

- Resolução espacial, alcance
- Acurácia espacial
- Linearidade, histeresis, calibragem

Medidas dinâmicas

- Acurácia dinâmica (= estática se sistema é linear)
- Resolução temporal, bandwidth, alcance da frequência
- Taxa de amostragem, acurácia temporal, latência

Medidas de precisão

- Estabilidade, durabilidade
- Ruído, interferência externa (materiais metálicos, oclusão etc.), filtragem, suavidade

Medidas de interface

- Interface com o sensor (câmera, transmissores) (voltagem AC/DC, resistência etc.)
- Suprimento de energia
- Interface com o usuário e modos de operação
- Interface de comunicação com o *host* (RS232, RS422, IEEE488 etc.)
- Conexão com o ator (com ou sem cabos)

Medidas computacionais

- Dados de referência (*joint-based, muscle-based, world-based*)
- Formato dos dados (binário ou ASCII, especificação do formato etc.)
- Representação dos dados (posição como um vetor de tamanho 3, tensão muscular e angular como um escalar, orientações como ângulos de Euler, matrizes de rotação 3x3 ou quaternions etc.)

Medidas operacionais

- Fator de forma (tamanho e peso dos marcadores/sensores/cabos)
- Espaço de trabalho
- Conforto, obtrusividade
- Tempo de preparação para a captura
- Compatibilidade com outros sistemas
- Ambiente de operação (humidade, temperatura)

2.2 Classificação

Podemos utilizar diversas classificações para os sistemas de Motion Capture atuais. A seguir, analisaremos três maneiras distintas de classificação: quanto à tecnologia utilizada, quanto à posição da fonte emissora e dos sensores/marcadores, e quanto à forma de obtenção dos dados capturados. Utilizando o pipeline descrito na seção anterior, a primeira classificação compreenderia os níveis **2** e **3**, a segunda o nível **2**, e a última se situaria no nível **5**.

2.2.1 Tecnologia

Esta classificação é muito importante pois define explicitamente o processo utilizado para a aquisição dos movimentos. Existem sistemas específicos, de tecnologias diferentes, que são mais indicados para certos tipos de aplicações. Nesta classificação veremos o estado atual da tecnologia dos sistemas de Motion Capture, juntamente com os prós e contras de cada tipo de sistema. Podemos dividi-los em quatro categorias: acústicos, mecânicos, óticos e magnéticos. Os dois últimos são os mais utilizados hoje em dia, e por isso serão o alvo principal de estudo neste texto.

Sistemas Acústicos

Neste tipo de sistema, um conjunto de emissores sonoros é colocado nas principais articulações do ator, enquanto que três receptores sensíveis são posicionados no local de captura. Os transmissores são então sequencialmente acionados para produzir um ruído característico, que será captado pelos transmissores, que então calcularão as posições no espaço.

O cálculo da posição de cada transmissor é feita da seguinte forma: utilizando como dados o tempo decorrido entre a emissão do ruído pelo transmissor e o seu recebimento pelo receptor, e a velocidade do som no ambiente, consegue-se calcular a distância percorrida pelo som, do transmissor até cada um dos receptores. Para

calcular a posição 3D de cada transmissor, é feita uma triangulação das distâncias deles em relação aos três receptores.

Um dos problemas deste método é a dificuldade de obter uma descrição correta dos dados num instante desejado, devido ao caráter sequencial do disparo dos transmissores no corpo do ator. Além disso, os sistemas acústicos sofrem do mesmo tipo de problema dos sistemas magnéticos: os incômodos cabos, que prejudicam a movimentação do ator, reduzindo assim o escopo de movimentos que podem ser executados. O número de transmissores que podem ser utilizados simultaneamente também é limitado, o que pode não fornecer uma descrição suficientemente correta do movimento capturado.

Sistemas acústicos estão sujeitos a problemas causados pela própria tecnologia - reflexões do som emitido pelos transmissores ou ruídos externos podem afetar o processo de captura e destruir os dados obtidos. Em compensação, este tipo de sistema não possui problemas de oclusão, típico de sistemas óticos; e interferência por objetos metálicos, percebido em sistemas magnéticos.

Somando vantagens e desvantagens, os sistemas acústicos devem ser considerados como a “última alternativa” na hora da aquisição de um equipamento de captura de movimento.

Sistemas Mecânicos

Os sistemas mecânicos de captura não contam com uma tecnologia muito avançada, porém possuem algumas vantagens que os tornam extremamente atraentes, especialmente na indústria cinematográfica. Geralmente, esse tipo de sistema é bem obtrusivo, talvez até mais do que os sistemas magnéticos com cabos (analisados adiante).

Sistemas mecânicos são compostos de potenciômetros e/ou *sliders* (ver figura 1) que, posicionados nas articulações desejadas, fornecem suas posições e orientações em altas taxas de amostragem (tempo real). Uma das vantagens desse tipo de sistema é que possuem uma interface parecida com as utilizadas em sistemas de *stop-motion*, muito utilizados na produção de filmes. Com isso, uma grande quantidade de animadores que utilizavam *stop-motion* puderam migrar para essa nova tecnologia sem problemas, aumentando a popularidade do sistema.

Os sistemas mecânicos são equipamentos de medida absoluta, não sendo afetados por campos magnéticos ou reflexões indesejadas, problemas típicos de sistemas magnéticos e óticos. Por isso, não necessitam de um processo longo de calibragem, o que torna a sua utilização mais fácil e produtiva.

Como exemplos de sistemas mecânicos de captura, podemos citar o The Monkey - Digital Poseable Mannequin e o Animatton. O primeiro pode ser considerado como um “*stop-motion* assistido por computador”. Nesse sistema, o “manequim” pode ser qualquer objeto desejado - desde humanóides até dinossauros. Um esqueleto articulado composto por potenciômetros é então fixado no modelo, fornecendo assim as posições e orientações impostas ao modelo pelo animador. Fica claro que a taxa de amostragem da captura pode ser tão alta quanto for a paciência do

animador. Uma desvantagem deste sistema é o realismo, que depende exclusivamente da habilidade do animador em posicionar cada articulação do modelo corretamente, num processo que requer, além de técnica, um pouco de arte. Geralmente, esse tipo de sistema mecânico é utilizado em conjunto com técnicas de técnicas de *keyframing* e cinemática.

Uma nova tecnologia em sistemas mecânicos de captura foi desenvolvida pela Analogus com o sistema Animatton. Este sistema, desenvolvido exclusivamente para a captura de movimentos humanos, é composto por articulações e potenciômetros que são posicionados diretamente no corpo do ator, formando uma verdadeira “armadura”. A captura é feita em tempo real e o sistema requer pouca ou nenhuma calibragem. Além disso, permite a captura de movimento de múltiplos atores numa mesma sessão. Apesar de ser extremamente obtrusivo (figura 1), este sistema possui vantagens como o preço (em torno de US\$ 35,000), o que o torna acessível para a maioria das produtoras que se utilizam de Motion Capture; e a possibilidade da captura de múltiplos atores numa mesma sessão. Veja na tabela a seguir, algumas características de sistemas mecânicos da captura:

	Sistemas Mecânicos
Resolução espacial	Independente do alcance, ~ 0.5 - 1 graus
Acurácia espacial	<= 5 graus
Precisão	de média a elevada
Taxa de amostragem	> 120 FPS ¹
Campo de trabalho	Ilimitado (depende apenas do comprimento dos cabos de transmissão)
Exemplos de sistemas	Animatton, Digital Monkey, Power Glove
Softwares compatíveis	Alias Wavefront, Jack, 3D Studio Max, SoftImage
Preço	US\$ 1,000 a US\$ 35,000

Tabela 1 - Informações técnicas sobre sistemas mecânicos

¹ Frames per Second

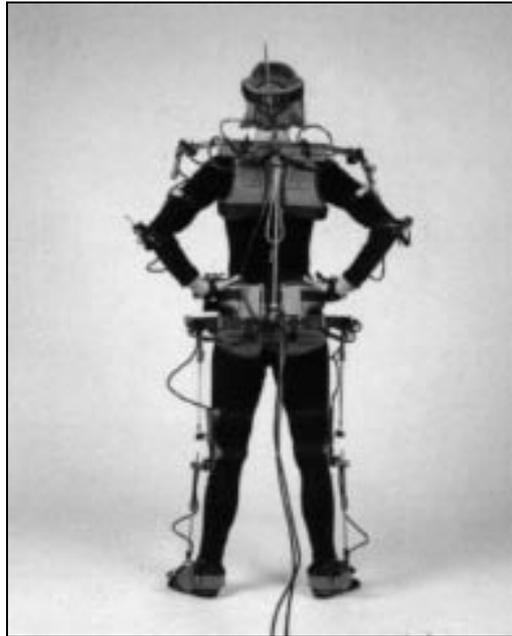


Figura 1 - O sistema Animatton

Sistemas Óticos

Neste tipo de sistema, o ator veste uma roupa especial coberta com refletores (em geral emissores LED) posicionados nas suas principais articulações (ver figura 2). Câmeras especiais são então posicionadas estrategicamente para fazer o *tracking* desses refletores durante o movimento do ator. Cada câmera gera as coordenadas 2D para cada refletor (obtidas via processo de segmentação). O conjunto dos dados 2D capturados pelas câmeras independentes é então analisado por um software, que fornecerá as coordenadas 3D dos refletores.



Figura 2 - Posicionamento dos marcadores em sistemas óticos de captura

Existem algumas variações entre diferentes sistemas óticos de captura. No sistema MultiTrax, da Adaptive Optics Associates, os marcadores posicionados nas articulações são refletores que, iluminados por flashes sincronizados de luz infravermelho, são capturados por uma ou mais câmeras. Também podem ser utilizados marcadores de outros tipos (focos de laser e retro-prismas).

O sistema Optotrak, da Northern Digital, utiliza marcadores LED sincronizados, que são rastreados por três câmeras sensíveis ao espectro infravermelho. O software fornecido no sistema utiliza os dados fornecidos pelas câmeras para gerar as coordenadas 3D, orientações e momentos dos marcadores. Um problema gerado pelos marcadores LED é o aquecimento, o que não recomenda a sua utilização por longos períodos de tempo.

O sistema ELITE, da Bioengineering Tech., utiliza marcadores refletivos passivos (hemisférios de plástico cobertos com material refletivo) e câmeras eletrônicas CCD com LED's infravermelhos em volta das lentes. O software fornecido no sistema utiliza algoritmos de reconhecimento de formas (do tipo *shape from shading*) para auxiliar o processo de tracking dos marcadores. Existem ainda muitos outros sistemas óticos no mercado. Entre eles, podemos citar o HiRES ExpertVision 3D (Motion Analysis Corp.) e o MacReflex (Optikon Corp.).

Os sistemas óticos de captura são os mais caros existentes atualmente, podendo chegar a até U\$ 250,000. Seu preço elevado é resultado da alta tecnologia utilizada no processo. As câmeras de alta resolução e os softwares de última geração são os principais responsáveis pelo preço elevado.

Uma das vantagens da utilização de sistemas óticos é a alta taxa de amostragem, que permite a captura de movimentos rápidos como os utilizados em artes marciais e esportes olímpicos. A taxa de amostragem depende basicamente da capacidade de definição das câmeras utilizadas no processo. Quanto maior resolução tiverem as câmeras, maior será a taxa de amostragem. Taxas de amostragem de até 200 quadros/segundo são conseguidas neste tipo processo quando se utilizam câmeras especiais de alta resolução e velocidade. Veja, na tabela 2, algumas informações técnicas complementares sobre os sistemas óticos.

Outra vantagem dos sistemas óticos é a liberdade que oferecem ao ator durante a execução do movimento. Ao contrário dos sistemas magnéticos, onde o ator é coberto de transmissores e fios, os refletores não oferecem resistência aos movimentos do ator. Além disso, não existe limite para o número de refletores posicionados no ator, o que permite alcançar um maior nível de detalhe na representação dos movimentos. O *workspace* do processo de captura para sistemas óticos é muito superior ao dos sistemas magnéticos e depende, basicamente, do campo de visão (FOV) das câmeras utilizadas no processo. Um cuidado, porém, deve ser tomado: a iluminação global do ambiente deve ser calibrada de modo a não causar interferências no processo de *tracking* dos refletores.

A grande desvantagem dos sistemas óticos é a oclusão de um ou mais refletores durante o processo de captura. Este tipo de problema é mais frequente durante a captura de movimentos de objetos pequenos (dedos das mãos) ou de vários atores interagindo muito próximos uns dos outros. Neste caso, a recuperação da posição 3D do refletor pode ser impossível, o que compromete todo o processo. A utilização de um maior número de câmeras reduz este tipo de problema.

O problema de oclusão pode ser minimizada com a utilização de um número maior de câmeras e de refletores. Porém existe um *trade-off* desses fatores: um maior número de câmeras acarreta um maior tempo de processamento para a CPU, durante o *tracking* dos refletores. Ao se aumentar o número de refletores, surge o problema de "*tracking confusion*", ou seja, a dificuldade de indentificar os refletores que estão muito próximos. Este problema é diretamente influenciado pela resolução das câmeras - câmeras de maior resolução conseguem identificar melhor refletores próximos.

Existe ainda a necessidade de um processamento via software dos dados obtidos pelas câmeras para a obtenção das informações 3D, o que não permite a interatividade, que é muito importante para a velocidade de produção e redução de custos operacionais. A utilização de softwares para o processamento dos dados obtidos pelas câmeras introduz erros de precisão (*offset*) no processo de obtenção dos dados 3D. Em alguns casos, é necessário um processo de filtragem dos dados finais para a eliminação dos erros de precisão e dos ruídos introduzidos pelo *tracking* das câmeras (o SNR é uma limitação em certos sistemas óticos).

	Sistemas Óticos
Resolução espacial	~ 0.0015 a 0.2 % do FOV ²
Acurácia espacial	~ 0.004 a 0.5 % do FOV
Precisão	~ 0.0055 a 0.02 % do FOV
Taxa de amostragem	> 200 FPS ³ (geralmente)
Campo de trabalho	raio de 1 a 4 m
Exemplos de sistemas	ExpertVision HiRES 3-D System, Multitrax Motion Capture System, OPTOTRAK
Softwares compatíveis	ZoeTrax, The Creative Motion Editor, PowerAnimator V7
Preço	US\$ 20,000 a US\$150,000

Tabela 2 - Informações técnicas sobre sistemas óticos

Sistemas Magnéticos

Os sistemas magnéticos de captura se caracterizam pela velocidade de processamento dos dados capturados (tempo real). Neste tipo de sistema, emprega-se um conjunto de receptores que são posicionados nas articulações do ator. Tais receptores medem a posição 3D e orientação das articulações em relação a uma antena transmissora, que emite um sinal de pulso. Cada receptor necessita de um cabo para se conectar à antena.

Os sistemas magnéticos são mais acessíveis em termos de preço. Sua tecnologia também é menos sofisticada que a utilizada nos sistemas óticos. Existem sistemas magnéticos simples e sofisticados, que vão desde US\$ 5,000 (Polhemus InsideTrak) a US\$ 33,000 (Ascension Flock of Birds). O preço desse tipo de sistema

² Field of Vision

³ Frames per Second

depende basicamente do número de receptores que conseguem ser processados simultaneamente. Pode-se, por exemplo, adquirir um *Flock of Birds* com poucos receptores por até US\$ 3,000.

Os sistemas magnéticos mais conhecidos são o Ultratrak Pro (figura 3), da Polhemus e o Flock of Birds, da Ascension. Neste último, por exemplo, um conjunto de até 30 receptores é conectado ao ator, fornecendo assim a orientação e posição de suas principais articulações.

Algumas vantagens dos sistemas magnéticos são o baixo custo computacional para o processamento dos dados, maior precisão dos dados (não existem problemas de oclusão) e o baixo custo do equipamento. Com uma taxa de amostragem que gira em torno de 100 FPS, os sistemas magnéticos são ideais para a captura de movimentos mais simples.

A maior desvantagem deste tipo de sistema são os diversos cabos que conectam os receptores à antena. Tais cabos restringem o movimento do ator, não permitindo deste modo que movimentos complexos e rápidos possam ser representados com naturalidade. Felizmente, algumas empresas já estão desenvolvendo sistemas magnéticos que não possuem cabos. Podemos citar, por exemplo, o *Motion Star*, da Ascension.

Uma outra desvantagem do processo magnético é a interferência causada por objetos de metal próximos ao local de captura. Campos magnéticos são extremamente sensíveis a objetos metálicos. Até mesmo a estrutura do prédio pode causar alguma interferência.

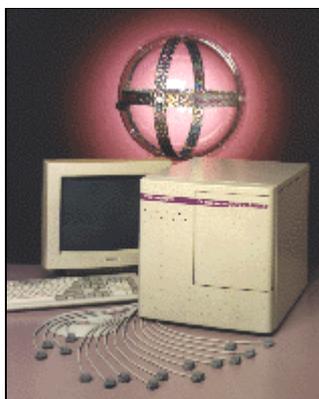


Figura 3 - O sistema UltraTrak Pro

Veja a seguir uma tabela contendo algumas características de sistemas magnéticos.

	Sistemas Magnéticos
Resolução espacial	~ 0.005 a 8 mm; ~ 0.025 a 0.1 graus
Acurácia espacial	~ 0.8 a 5 mm; 0.1 a 3 graus
Precisão	elevada
Taxa de amostragem	~ 100 FPS
Campo de trabalho	raio de 1 a 2 m
Exemplos de sistemas	A Flock of Birds, ULTRATRAK, FASTRAK, INSIDETRAK
Softwares compatíveis	Alias Wavefront, 4Dvision, SoftImage, Kinemation 3.0, 3D Studio MAX, PowerAnimator V7, Photo4D
Preço	US\$ 5,000 a US\$ 70,000

Tabela 3 - Características dos sistemas magnéticos

2.2.2 Fonte Emissora x Sensores/Marcadores

Uma classificação interessante, proposta em [Mulder 94], divide os sistemas de captura em três categorias: *inside-in*, *inside-out* e *outside-in*. A semântica desta classificação é simples: a primeira palavra refere-se à localização do sensor ou marcador no sistema (*inside*, se o sensor está no corpo do ator; e *outside*, caso esteja fora). Já a segunda palavra indica a posição da fonte emissora (por exemplo, a câmera dos sistemas óticos), em relação ao corpo do ator. Veja, na figura 4, uma imagem ilustrando os componentes típicos de sistemas de captura de movimento.

- Sistemas *Inside-in*

Neste tipo de sistema, os sensores e a fonte transmissora estão localizados no próprio corpo do performer. Um exemplo de sistemas deste tipo é a *Power Glove* (Nintendo-Mattel), que utiliza sensores flexíveis para captar orientações. Os sensores deste tipo de sistema geralmente possuem um baixo “fator de forma” e, conseqüentemente, são os mais indicados para a captura de movimentos de objetos pequenos (dedos das mãos, olhos etc.). Seu *workspace* é ilimitado devido à sua restrição espacial ao corpo do ator, porém são considerados extremamente obtrusivos. As coordenadas obtidas por sistemas deste tipo são, geralmente, desprovidas de informação 3D, fornecendo somente orientações relativas. Sistemas mecânicos de captura como o *Digital Monkey* estão inseridos nesta categoria.

- Sistemas *Inside-out*

São sistemas onde os sensores estão conectados ao corpo do ator, e que respondem a sinais emitidos por uma fonte emissora externa. Sistemas magnéticos e acústicos estão nesta categoria. O *workspace* e a precisão de sistemas *inside-out* são limitados devido ao uso de fontes externas, mas os dados obtidos fornecem uma boa descrição 3D dos sensores/marcadores. Além disso, o seu “fator de forma” mais alto restringe sua utilização à captura de partes maiores do corpo.

- Sistemas *Outside-in*

Em sistemas deste tipo, a fonte emissora está localizada no corpo do ator, enquanto que os sensores externos capturam seus sinais. Em sistemas óticos, a fonte emissora são os marcadores cobertos com material refletivo, que refletem a luz emitida por flashes. Tais reflexões são captadas pelos sensores, que são as câmeras. Sistemas deste tipo sofrem do problema de oclusão e possuem um workspace limitado, porém possuem o menor grau de obtrusividade. Outro exemplo de sistemas *outside-in* são os utilizados para o tracking da pupila e da córnea.

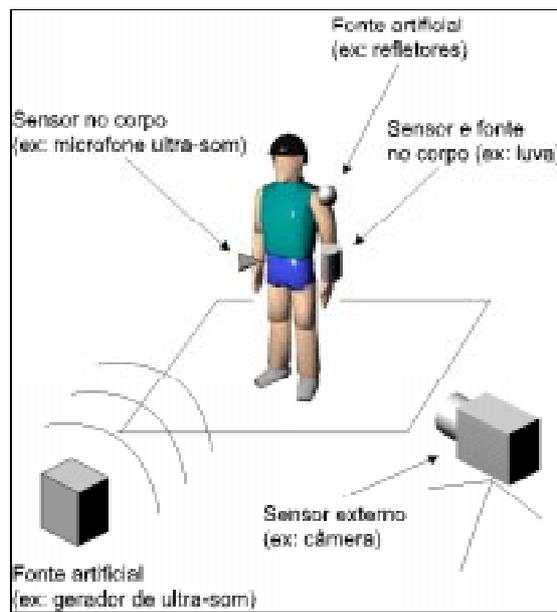


Figura 4 - Componentes típicos de sistemas de Motion Capture

2.2.3 Aquisição Direta x Indireta

Essa é uma classificação importante pois fornece uma indicação de como são obtidos os dados capturados: direta ou indiretamente.

Sistemas de aquisição direta são aqueles cujos dados obtidos não necessitam de nenhum processamento posterior à captura. Em geral, sistemas de aquisição direta são os mais obtrusivos e possuem uma baixa taxa de amostragem. Estão incluídos nessa categoria os sistemas acústicos, mecânicos e magnéticos. Note que, apesar de não haver a necessidade de um pós-processamento dos dados, geralmente é aplicado um procedimento de filtragem para melhorar a qualidade e eliminar ruídos.

Os sistemas de aquisição indireta permitem uma maior liberdade ao performer e possuem uma alta taxa de amostragem. Porém, devido a sua alta tecnologia, são extremamente caros. Nesse tipo de sistema, os dados obtidos durante o processo são agrupados e analisados, geralmente através de um software poderoso, de modo a gerar o produto final. Este fato inviabiliza a utilização de sistemas deste tipo em aplicações que necessitam de interatividade, porém os tornam mais precisos e mais indicados

para a capura de movimentos complexos e rápidos. Sistemas óticos se encaixam nesta classificação.

2.2.4 Comparação 1: Ótico vs. Magnético

Dependendo da aplicação desejada, diversos fatores devem ser levados em conta na hora da aquisição ou utilização de um equipamento de captura. A tabela apresentada a seguir apresenta uma comparação entre as características dos sistemas de Motion Capture disponíveis atualmente.

	Sistemas Óticos	Sistemas Magnéticos
Características principais dos sistemas	Conjunto de câmeras, fazendo o tracking 2D dos refletores para posterior conversão para 3D	Transmissores magnéticos passam posição XYZ e orientação dos marcadores
Dados obtidos	Posição XYZ dos refletores	Posição XYZ + orientação dos marcadores
Taxa de amostragem	> 200 FPS	Cerca de 100 FPS
Processamento dos dados	Tracking via software dos conjuntos de posições 2D para posterior geração de coordenadas 3D	Tempo real (direto do transmissor para o visualizador)
Principais vantagens	Alta taxa de amostragem (captura de movimentos mais complexos), número de refletores ilimitado, liberdade dos atores durante o processo	Processamento em tempo real, baixo custo do equipamento, obtenção de coordenadas 3D + orientações das articulações, maior precisão dos dados
Principais desvantagens	Oclusão de refletores, necessita tracking via software para geração de coordenadas 3D, alto custo do equipamento, menor precisão dos dados	Restrição do movimento dos atores, limitação do número de marcadores, objetos de metal interferem nos resultados
Preço médio	US\$ 150,000 a US\$ 250,000	US\$ 5,000 a US\$ 70,000
Principais fornecedores de hardware	Motion Analysis Corp., Adaptive Optics Associates, Northern Digital Inc.	Polhemus, Ascension Technology, Acclaim Inc., SimGraphics
Exemplos de sistemas de aquisição	ExpertVision HiRES 3-D System, Multitrax Motion Capture System, OPTOTRAK	A Flock of Birds, UltraTrak, FastTrak, InsideTrak
Produtoras atualmente utilizando esta tecnologia	Tsi, Acclaim, OZ, Pixar, Boss Films, Pacific Data Images	Motek, Magnet Studios, Colossal Pictures, R/Greenberg Associates, Lamb & Company
Softwares compatíveis	ZoeTrax, The Creative Motion Editor, PowerAnimator V7	Alias Wavefront, 4Dvision, SoftImage, Kinemation 3.0, 3D Studio MAX, PowerAnimator V7, Photo4D

Tabela 4 - Comparação entre sistemas óticos e magnéticos

3 O Formato dos Dados Capturados

Existem atualmente diversas empresas especializadas em captura de movimentos. Tais empresas fornecem serviços de captura e/ou aluguel de seus equipamentos. O formato de dados gerado pelas empresas depende, basicamente, do tipo de sistema utilizado no processo de captura. No geral, um arquivo de dados MC contém o número total de quadros gravados, uma descrição dos marcadores (nomes, posição inicial etc.) e suas posições ao longo dos quadros da animação. Informações adicionais como fatores de escala (para o mapeamento dos dados nos modelos do computador) e dados sobre a câmera (FOV, posicionamento etc.) podem ser adicionados pelos fornecedores. Neste texto serão analisados dois tipos de arquivos ASCII: o VPM (Viewpoint Motion File), fornecido pela Viewpoint utilizando os serviços da empresa BioVision, especializada em Motion Capture; e o arquivo XTK (XIST TRACKING), produzido pela empresa Vierte Art.

3.1 O formato VPM

Os arquivos VPM contêm as informações necessárias para a reprodução da performance de um ator real no universo 3D do computador. As informações são relativas a um corpo segmentado, que é equivalente à divisão do corpo do ator em segmentos independentes, dependendo da posição dos marcadores utilizados no processo de captura. Estes segmentos são fornecidos em arquivos 3DS, também da Viewpoint, e que contêm um modelo 3D e um *bounding box* para cada segmento descrito no arquivo VPM. Estes objetos ajudam no posicionamento adequado dos pontos de pivot para a rotação em cada segmento, além dos fatores de escala e orientação. É possível mapear os dados de um arquivo VPM em outras estruturas humanas diferentes das fornecidas pela Viewpoint, porém elas devem possuir os mesmos fatores de escala e as mesmas posições de pivot nos segmentos, para produzir uma animação perfeita.

Os arquivos VPM contêm informações para cada segmento do corpo, e são formadas por um *header* e nove canais de dados. O *header* de cada segmento tem a seguinte sintaxe:

```
Segment: name
Frames: number
Frame Time: seconds
channel1 label      channel2 label      ...
      units                units
```

Os nove canais contêm as posições XYZ, rotações e fatores de escala do segmento durante a animação. Estes dados estão ordenados em *number* linhas, onde *number* representa o número de quadros da animação. Os segmentos são independentes uns dos outros e são referenciados segundo um sistema de coordenadas *right handed*, com o os eixos X e Z no chão e eixo Y apontando para cima (ver figura

5). A ordem correta de aplicação das transformações no modelo é a seguinte: a translação XYZ deve ser aplicada primeiro, de modo a mover os segmentos para a posição correta no espaço; então os segmentos são orientados aplicando-se as rotações na ordem X-Z-Y e, por último, os fatores de escala devem ser adicionados, de modo a fazer com que os segmentos possuam o mesmo tamanho dos utilizados pelo ator. Veja a seguir um trecho de código de um arquivo VPM (zwalk.vpm), que representa o movimento de caminhar de um ator.

```
Segment:    Hips
Frames:     86
Frame Time: 0.033333
XTRAN  YTRAN  ZTRAN  XROT  YROT  ZROT  XSCALE YSCALE ZSCALE
INCHES INCHES INCHES DEGREES DEGREES DEGREES PERCENT PERCENT PERCENT
0.34   38.43   4.03   11.67  -4.68   2.80   423.06  423.06  423.06
0.42   38.57   5.26   11.64  -4.06   2.72   423.06  423.06  423.06
0.46   38.64   6.45   11.58  -2.99   2.18   423.06  423.06  423.06
0.51   38.67   7.62   11.75  -1.06   1.90   423.06  423.06  423.06
0.58   38.74   8.81   11.75   0.80   2.00   423.06  423.06  423.06
0.60   38.82   9.97   11.88   2.14   1.77   423.06  423.06  423.06
0.53   38.88  11.08  12.01   2.91   0.80   423.06  423.06  423.06
0.43   38.91  12.21  11.83   3.61  -0.48   423.06  423.06  423.06
0.38   38.92  13.37  11.44   4.65  -1.43   423.06  423.06  423.06
0.36   38.87  14.54  11.03   5.78  -1.92   423.06  423.06  423.06
0.37   38.78  15.74  10.84   6.82  -2.07   423.06  423.06  423.06
```

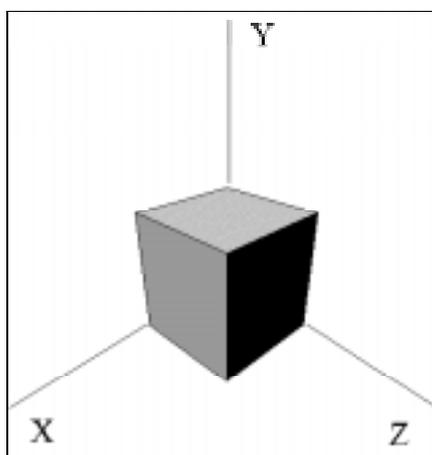


Figura 5 - O sistema de coordenadas dos arquivos VPM

3.2 O formato XTK

Os arquivos do tipo XTK são semelhantes aos do tipo VPM, porém possuem algumas informações adicionais como, por exemplo, a hora e data da gravação do movimento. Além disso, permite a utilização de símbolos especiais para comentários. Os arquivos XTK foram originalmente desenvolvidos para a captura de expressões

faciais, mas podem ser utilizados para armazenar dados de outras partes do corpo. As seguintes observações são válidas para arquivos XTK:

- A primeira linha deve contém o *header* X-IST
- Comentários são iniciados pelos caracteres # ou //
- Valores são assinalados na forma KEYWORD=value (separados pelo sinal =)
- Blocos de dados são iniciados por um rótulo LABEL: (variável) e finalizados por um rótulo END. O rótulo inicial LABEL deve ser seguido pelo caracter :
- Os blocos de dados são únicos, não podendo existir nomes de blocos iguais em um mesmo arquivo XTK

Veja a seguir o exemplo de um arquivo no formato XTK.

```
X-IST:
# primeira linha deve ser sempre X-IST:
// isto também é um comentário

// versão do formato X-IST (VERSION=version(*100)
VERSION=101

// identificação do programa
CREATED=XIST_FTR 1.01 © VIERTE ART, Munich

// data e hora da gravação (dd.mm.yy hh.mm.ss)
TIME=10.5.95 10:37:01

// bloco para comentários diversos
TEXT:
Preste atenção nos movimentos suaves dos olhos!
END

// frame rate de gravação/segundo
// PAL rate = 25, NTSC rate = 30
RATE=25

// descrição das áreas de busca
SEARCHAREAS:
// número do marcador = x1 y1 x2 y2 xa ya
// xa e ya descrevem a posição do marcador em uma face
1=5 6 100 200 50 102
2=6 8 100 200 48 78
END
```

```

// descrição dos marcadores (bloco opcional)
MARKERS:
1=Pupila esquerda
2=Sombrancelha direita
END

// posições de todos os marcadores adicionais na cabeça.
// bloco necessário somente se foi feita a utilização
// de marcadores adicionais
// número_do_marcador=x y
ADDMARKERPOS:
18=700 60
19=750 60
END

// canais ativos no arquivo
CHANNELS:
// número_do_canal, opcionalmente seguido por =
0
1
8=marcador 5.x
9=marcador 5.y
// Obs: não ha verificação da correspondência entre
// marcadores e relações entre os canais
END

// número total de quadros gravados (inteiro >0)
FRAMES=2

// bloco de dados dos quadros e canais
// cada linha contém a informação de um quadro
// valores normalizados entre 0.0 e 1.0, de acordo
// com o tamanho do bounding box

DATA:
0.0 0.0 0.0
// quadro 1: canais 0 1 8 9
// canais 0 e 1 são relativos aos marcadores 1.x e 1.y
// canais 8 e 9 são relativos aos marcadores 5.x e 5.y
0.1 0.2 -0.1 -0.3
// quadro 2: canais 0 1 8 9
END

// fim do arquivo XTK

```

END

4 Aplicações de Motion Capture

4.1 Cinema e Televisão

O uso de Motion Capture facilita o trabalho de animadores, que estão gradualmente migrando do tradicional *keyframing* para esta nova tecnologia. A rapidez de produção e o realismo da animação gerada torna este tipo de processo extremamente interessante para as produtoras de efeitos especiais e de televisão. Filmes recentes como *Species* [Boss 96], *Toy Story*, *Independence Day* e *Batman Forever* já utilizam esta tecnologia.

Alguns canais de televisão também já criaram programas que são apresentados por “atores virtuais”, cujos movimentos foram capturados via Motion Capture. Um exemplo é *Moxy*, que apresenta um programa de desenhos no *Cartoon Network*. *Moxy* utiliza uma mistura de animação convencional com Motion Capture, pois os dados capturados de um ator real são mapeados em um personagem animado dos desenhos, necessitando assim de algumas expressões características deste tipo de personagem.

Outro exemplo do uso de Motion Capture em programas de televisão é o personagem *Weldon* (figura 6), uma simpática ovelha que contracena com atores reais, em um programa de muito sucesso na TV.



Figura 6 - A “ovelha” Weldon

4.2 Jogos Interativos

Motion Capture provocou uma revolução na indústria dos jogos de computador. Movimentos que antes eram complicados de serem reproduzidos agora são fáceis de serem aplicados aos personagens dos jogos, criando uma nova dimensão

de realismo. Em jogos de luta, a qualidade dos movimentos é visivelmente melhor naqueles que utilizam a técnica de Motion Capture. Como exemplo de jogos que utilizam esta técnica, podemos citar o *Virtua Fighter* (figura 7) e o *FX Fighter* [FX 95].



Figura 7 - Utilização de MC em jogos

4.5 Pesquisa

Sendo uma tecnologia nova, Motion Capture ainda possui muitas limitações e abre espaço para diversas frentes de pesquisa. A seguir discutiremos alguns problemas e rumos de pesquisa inerentes à esta tecnologia.

1. **Melhoramento de Técnicas de Tracking.** Como foi estudado anteriormente, o problema de oclusão dos marcadores nos sistemas óticos é crítico. Técnicas de *tracking* mais sofisticadas e apropriadas para Motion Capture podem ser desenvolvidas.
2. **Análise e Modificação de Movimentos Capturados.** Cada arquivo de dados capturados define um conjunto limitado e específico de ações. Além disso, as informações que estão no arquivo definem um *script* de movimento do personagem na cena, inviabilizando portanto um controle interno destes dados. Uma linha de pesquisa interessante seria analisar certos tipos de movimentos capturados para então criar novos tipos de movimentos a partir deles.
3. **Concatenação e Fusão de Diversos Movimentos Capturados.** Como foi citado anteriormente, nem sempre os movimentos capturados estão num mesmo arquivo MC. Uma idéia interessante [Wagner, Roma 96] é concatenar diversos movimentos capturados, de arquivos diferentes, e desenvolver um algoritmo que permita a fusão desses movimentos, respeitando as restrições de um ser humano. Poderia-se então criar uma base de movimentos que seria utilizada para dar uma maior liberdade ao animador e gerar movimentos mais complexos.
4. **Combinação de Motion Capture com outras Técnicas de Animação.** Sistemas de animação podem combinar as vantagens de Motion Capture com técnicas como *keyframing* [Watt 92], animação procedural [Perlin 95] e simulação numérica, proporcionando assim um maior poder de criação de animações realistas.

5. **Técnicas de Mapeamento de Dados Capturados.** Nem sempre se deseja animar realisticamente seres humanos. No filme *Jurassic Park*, por exemplo, o movimento de um elefante foi capturado para dar vida aos dinossauros virtuais no computador. Técnicas para mapear movimentos de seres reais em objetos ou seres que não existem são um interessante campo de pesquisa. Na figura 8, podemos observar um exemplo de mapeamento de dados obtidos via MC de atores reais em um objeto não-humano no computador.



Figura 8 - Mapeamento de dados MC

5 Referências

[IMPA 96] - Velho, L., Jonas, M. G., Roma, P. C., Wagner, F. S. V. S., *Motion Capture*, Apostila de referências para consulta interna do IMPA, brevemente disponível via WWW.

[Wagner, Roma 96] - Wagner, F. S. V. S., Roma, P. C., “Animações Complexas em Tempo Real utilizando Movimentos Capturados”, Anais do SIBGRAPI’96, Outubro 1996.

[Watt 92] - A. Watt, A. Watt, *Advanced Animation and Rendering Techniques - Theory and Practice*, Addison-Wesley, 1992.

[Perlin 95] - K. Perlin, “Real Time Responsive Animation with Personality”, IEEE Transactions on Visualization and Computer Graphics, 5-15, Vol. 1, no. 1, March 1995.

[Blinn 87] - J. Blinn, “Nested Transformations and Blobby Man”, IEEE Transactions on Graphics, 59-65, October 1987.

[Boss 96] - Boss Film Studios, <http://www.boss.com>

[FX 95] - Motion Capture vs Keyframing, <http://?>

[Kalawsky 93] - *The Science of Virtual Reality and Virtual Environments*, pp 277-291, 135-164, 187-197. Addison-Wesley. Wokingham, England, 1993.

[Mulder 94] - Mulder, A., "Human Movement Tracking Technology", Technical Report 94-1, School of Kinesiology of the Simon Fraser University, July 1994.