

# Captura de Movimento e Animação de Personagens em Jogos

João Victor Boechat Gomide 1, David Lunardi Flam 2, Daniel Pacheco 2, Arnaldo de Albuquerque Araújo 2

1 Faculdade de Ciências Empresariais, Universidade FUMEC

2 Departamento de Ciência da Computação, UFMG

## Resumo

Esse tutorial discute, tanto do ponto de vista prático quanto teórico, a animação de personagens e a utilização da captura de movimento para esse fim. Assim, ele apresenta brevemente a história da animação e da captura de movimento. Em seguida, aborda de maneira mais detalhada todo o fluxo de trabalho para realizar a captura de movimento e animar personagens com essa técnica. Ele também apresenta uma solução de código aberto para a captura de movimento proposta pelos autores e o estado atual desse desenvolvimento.

## Abstract

This tutorial discusses character animation and the uses of motion capture for animation. It presents a brief history of animation and motion capture. It discusses more deeply the workflow to realize motion capture and to animate characters with this technique. At the end, it presents an open source code solution to capture motion, proposed by the authors, and the actual stage of development.

## Contatos dos autores:

jvbg@face.fumec.br  
{arnaldo, david, pacheco}@dcc.ufmg.br

## 1. Introdução

A animação de personagens em jogos eletrônicos é uma área em contínuo desenvolvimento acelerado. Ela passou por diversas etapas de sofisticação, desde os primeiros jogos com poucos pixels e poses, como no jogo Pac-Man, até o estágio atual, caracterizado por uma interatividade cada vez maior do jogador com o software e seus personagens. As perspectivas de aonde se pode chegar são bastante claras, porém todo o caminho tem que ser percorrido se pautando no desenvolvimento de hardware, primeiramente, e de algoritmos.

Apesar de ter a sua plenitude de realização ancorada em desenvolvimentos tecnológicos a serem ainda realizados, a animação de personagens é intensamente baseada na arte de animar. Essa arte se iniciou em projeções com a lanterna mágica, no século 17, e evoluiu até os dias de hoje, passando pela

introdução do cinema nas nossas vidas. Sempre limitada pelo tipo de suporte midiático para realizá-la. A evolução da animação de personagens, a partir de 1906, com o primeiro filme de Blackton, para o estúdio de Thomas Edson, às animações com computador, dialoga com o desenvolvimento dos jogos eletrônicos desde o seu primeiro momento. Os ciclos de caminhada, a aceleração e desaceleração de movimentos, os movimentos secundários, a animação por poses-chave ou direta e outros princípios foram formalizados de maneira sistemática nos Estúdios Disney, em torno de 1940, e são utilizados plenamente na construção de animações em jogos nesse momento. A técnica evoluiu com as demandas de interatividade para os jogos e tem apresentado novas soluções, baseadas em ciclos de animação, fusão entre ciclos de movimentos e na busca de maior realismo com os recursos disponíveis.

Existem duas maneiras de se animar personagens: por pose-chave, ou *keyframe*, e a direta, que é construída do início ao fim, momento a momento, ou quadro a quadro. A captura de movimento é uma animação do tipo direta, diferentemente da animação por *keyframe*, que determina a posição em momentos extremos no tempo e interpola o movimento e a forma entre essas poses. A animação de jogos, atualmente, é uma mistura desses dois modos, utilizando a animação direta para ciclos de animação e a animação por *keyframe* para a mudança entre os ciclos.

Nesse tutorial abordamos a utilização da captura digital de movimento e os conceitos relacionados à animação de personagens em geral. A captura de movimento (*mocap*) é amplamente usada para animar personagens em jogos em todo o mundo. No Brasil, bancos de dados de captura de movimento são pouco utilizados para animação de personagens virtuais em qualquer tipo de aplicação. Quando usados em jogos, em geral, os dados são alugados, comprados ou adquiridos gratuitamente de fornecedores, já que apenas uma empresa, pelo que conhecemos, faz esse trabalho específico no Brasil no presente momento. Essa situação advém do fato que os equipamentos de *mocap* são muito caros, incluindo as taxas de importação e os impostos, e a sua utilização tem complexidades intrínsecas. Diante desse cenário, os autores desse tutorial estão trabalhando para construir um sistema de captura de movimento de código aberto.

Esse tutorial discute, tanto do ponto de vista prático quanto teórico, essas questões e também a animação de personagens, os usos da captura de

movimento em jogos, o estágio atual da nossa solução e próximos passos a serem dados. Devido à quantidade de informação contida nesse tutorial, não abordaremos mais longamente a evolução histórica da animação de personagens e da captura de movimento. Essa abordagem pode ser encontrada na dissertação de mestrado de um dos autores [Gomide 2006] e nos livros de Alberto Menache [2000] e Thomas e Johnston [1984].

Ele também apresenta o OpenMoCap, o software de código livre que está sendo desenvolvido pelos autores, e será apresentado em seu atual estágio [Flam et al 2009]. Esse software faz parte de um projeto desenvolvido pelo Núcleo de Processamento Digital de Imagens, do Departamento de Ciência da Computação da UFMG em colaboração com o Laboratório de Computação Gráfica da Universidade FUMEC, e teve seu início em 2006 [Gomide 2006]. Atualmente ele é apoiado financeiramente pela FAPEMIG e pelo CNPq. O software, nesse momento, permite a realização, em tempo real, de todo o fluxo de trabalho e fornece os dados de movimento na saída. Estão sendo implementadas melhorias no rastreamento de pontos e na velocidade dos processos. Futuros desenvolvimentos para essa solução e as possibilidades, como uma opção de animação de personagens com os dados gratuitos gerados pelo OpenMoCap, serão discutidos ao final.

## 2. Animação de personagens

A arte da animação sempre esteve à procura de técnicas que diminuíssem o trabalho envolvido na construção da história quadro a quadro e que contribuíssem para expandir o seu universo de expressão. Antes de se elevar à categoria de manifestação artística, a animação esperou por milênios de evolução técnica para despontar como uma forma de expressão humana, no início do século XX. A arte da animação é expressada nos desenhos animados, nas animações abstratas, nos jogos eletrônicos, nos efeitos visuais e em outras modalidades. A utilização da animação se torna ainda mais ampla com a imagem digital em movimento. Podem-se empregar os mesmos equipamentos para produzir imagens para jogos, cinema, televisão, internet, telefones celulares, painéis eletrônicos e o que mais possa reproduzi-las. As imagens são, em última análise, luz que atinge um dispositivo eletrônico de estado sólido, chamado de ccd (*charged-coupled device*) ou cmos (*complementary metal oxide semiconductor*), ou são produzidas inteiramente por computador, e são usadas para contar histórias, interativas ou não.

Nos primórdios da produção cinematográfica existia o desenho animado e o *trickfilm*, ou filme de efeito, que se utilizava da técnica de substituição por parada da ação (*stop motion*), isto é, parava-se a encenação, alterava-se o que se desejava dentro do enquadramento da câmera, e depois se continuava a filmar, mantendo a posição do resto do quadro como estava antes da modificação. Isso era feito também

para introduzir elementos de animação nas cenas filmadas. Utilizavam-se então exclusivamente desenhos quadro a quadro e trucagens ópticas para imprimir a animação na película.

Ao longo dos tempos, desde o primeiro curta-metragem todo feito com desenho animado em 1908, *Fantasmagorie*, de Emile Cohl, foi-se procurando aperfeiçoar as técnicas para animar e dar personalidade aos personagens e encontrar caminhos para a expressão dessa nova arte. Apesar de já existirem as histórias em quadrinho, que indicavam uma maneira de realizar a decupagem e confeccionar os desenhos animados, esta arte permaneceu em um segundo plano por muito tempo. Isso diante das dificuldades para construir um número tão grande de quadros para contar a história, ao trabalho mecânico e tedioso envolvido e à construção da própria linguagem a ser utilizada.

A grande mudança de paradigma ocorreu com o aparecimento dos Estúdios Disney no final da década de vinte, em meio a um cenário dominado pelo Gato Felix, de Otto Messmer, e esforços pouco eficazes para aumentar a eficiência da produção dos quadros da animação. Walt Disney proporcionou uma nova forma de abordar a animação em todos os seus aspectos, dos movimentos aos tipos de personagens, ao sistema de produção e testes. Pode-se dizer que nessa época se estabeleceram os conceitos mais gerais e eficazes para fazer animação, e esses conceitos são aplicáveis a qualquer técnica de animação de personagens, seja a tradicional, a feita por computação gráfica com keyframes ou a captura de movimento [Barbosa Júnior 2004].

Dois dos participantes da equipe dos Estúdios Disney em seus tempos de maior riqueza criativa, Frank Thomas e Ollie Johnston, escreveram o clássico *The Illusion of Life: Disney Animation*. Nesse livro estão descritos os doze princípios de animação de personagens, que podemos considerar como princípios básicos para a animação de personagens em geral. Tendo em vista esses princípios podemos analisar as possibilidades da captura de movimento para as diferentes demandas de uso da técnica. Os doze princípios básicos estão listados a seguir, utilizando algumas observações adicionais de Livermann [2007]:

1. Comprimir e esticar: pode ser criado para criar uma sensação de peso;
2. Antecipação: é o movimento na direção oposta antes que a ação principal comece;
3. Encenação: refere-se a apresentar uma idéia ou ação claramente: como os personagens interagem entre si, como se movem, como a cena é vista;
4. Animação direta e posição-chave: são dois métodos de animação. No primeiro caso, a ação é construída em uma seqüência de eventos quadro a quadro, começando no início até que a animação esteja completa. Na posição-chave as poses mais importantes da ação são construídas primeiramente. A seguir, os quadros entre elas vão sendo preenchidos

- até que a animação esteja completa (*in-between*);
5. Continuidade e sobreposição da ação: refere-se ao movimento secundário, como passar do ponto de parada e depois retornar, por exemplo; é o oposto da antecipação;
  6. Aceleração e desaceleração: mudanças na velocidade dos objetos animados;
  7. Movimento em arco: a maior parte dos movimentos não são lineares; refere-se aos objetos se movendo no espaço em arcos ao invés de retas;
  8. Ação secundária: movimento de outras partes ou objetos que reagem à ação primária dos membros e face, como cabelos e roupas;
  9. Temporização: refere-se ao ritmo em que a ação ocorre;
  10. Exageração: implica em aproximar ou ultrapassar os limites da realidade física para aumentar ou dramatizar a performance do personagem;
  11. Personalidade: dois personagens idênticos podem aparecer totalmente diferentes com personalidades diferentes em cada um deles;
  12. Apelo: a platéia deve achar os personagens interessantes e com apelo, tanto em animação quanto em cenas captadas ao vivo.

Os princípios de sobreposição de ação, animação direta, aceleração e desaceleração, arcos e movimento secundário são naturais para uma atuação ao vivo. Assim, têm-se os princípios naturalmente no processo de captura de movimento. A temporização, o apelo, a personalidade e a encenação requerem trabalho qualquer que seja a técnica utilizada.

A computação gráfica veio atender em grande parte as demandas para minimizar e racionalizar os esforços para produzir cenas de animação, através dos programas de modelagem e animação. Esses softwares começaram a surgir nos anos oitenta e hoje são populares. Desde o primeiro filme a utilizar cenários gerados por computador, *Tron* (1982), uma produção dos Estúdios Disney, seu uso só vem crescendo. Juntamente com o computador e a imagem digital, surgiu a captura digital de movimento, que vem evoluindo e está introduzindo novos paradigmas na forma de animar os personagens.

Existem duas maneiras de se fazer animação com computação gráfica, como na animação tradicional, que são a animação por *keyframes* e a captura de movimento. Na animação por *keyframes* são utilizados quadros-chave (*keyframe*), ou poses-chave, onde se dão valores a parâmetros envolvidos no movimento, como posição, cor, ou o que quer que possa variar com o tempo. Faz-se isso nos quadros principais, e o computador calcula a evolução dos parâmetros entre os quadros-chave, utilizando interpolação e ajustes pré-determinados, habilitados

pelo animador, como o *easy-in* e o *easy-out* – que podem ser traduzidos por aceleração e desaceleração. Nos trabalhos feitos com captura de movimento, usa-se a animação direta, animando-se cada quadro, do início ao fim. As animações, em geral, adotam um método híbrido entre essas duas maneiras, utilizando uma ou outra quando é mais adequado.

### 3. Captura digital de movimento

A captura de movimento, também conhecida como mocap, é um conjunto de artifícios usado para mapear e reproduzir deslocamentos em objetos ou seres vivos. Ela é atualmente realizada usando recursos digitais, e é um campo do conhecimento relativamente recente, ainda à procura de se definir e de aperfeiçoar sua tecnologia, que evolui em diversas direções. A mocap foi primeiramente utilizada e desenvolvida para aplicações médicas, mas as produções cinematográficas e de jogos eletrônicos se apropriaram dela e expandiram suas aplicações. Atualmente ela está nos planejamentos de produção em diversos tipos jogos bem sucedidos, como o *FX Fighter* e as séries *Fifa Soccer* e *NBA Live*.

De maneira resumida, nos dias de hoje a captura de movimento é o processo que permite traduzir uma atuação ao vivo em uma atuação digital. Ou, como definido por Menache, “captura de movimento é o processo de gravar um evento de movimento ao vivo, e traduzi-lo em termos matemáticos utilizáveis ao rastrear um número de pontos-chave no espaço através do tempo, e combiná-los para obter uma representação tridimensional única da performance”. Em geral, capturam-se separadamente os movimentos corporais dos movimentos faciais, devido às características e sutilezas desses últimos. São colocados marcadores ou sensores em pontos-chave do corpo, como as junções das articulações, e suas posições são monitoradas ao longo do tempo. Depois essas posições são transferidas para as articulações do personagem a ser animado e dessa maneira ele ganha vida.

A captura de movimento envolve um conhecimento multidisciplinar, utilizando técnicas de computação e de engenharia associadas à criação e realização de animações, tanto bidimensionais como tridimensionais, sobre imagens puramente animadas virtualmente ou aplicando animações sobre cenas gravadas. Além disso, deve-se dirigir a performance do ator que dará vida ao personagem.

A captura de movimento é uma técnica em pleno desenvolvimento e ainda existem diversas polêmicas de taxonomia, referentes às possíveis denominações e divisões de trabalho e de responsabilidades na produção da animação com captura de movimentos. Alguns dos nomes adotados são captura de movimento, captura de performance, animação de performance, marionete digital e animação em tempo real.

No presente trabalho todos esses nomes e definições serão reunidos sob o nome de captura de movimento. Essa discussão sobre esses termos e suas

nuances não traz efetivamente um aprofundamento da teoria e da práxis. Pode-se tê-los todos sob essa denominação mais geral, que é captura de movimento, sem prejuízo de entendimento ou de cobrir as diferenças entre as aplicações, nele incluindo a captura da performance de um ator, que será usada em um personagem virtual de animação, ou a captura da performance de um atleta em estudos de biomecânica.

Se adotarmos uma definição mais vaga, sendo a captura de movimento o “ato de capturar os movimentos em geral”, podemos localizar sua origem no final do século dezanove, nos trabalhos dos fotógrafos Eadweard Muybridge [1984] e Etienne-Jules Marey [1873]. Eles desenvolveram independentemente técnicas de fotografia para estudar a locomoção, que foram amplamente utilizadas por artistas plásticos, animadores e cientistas. São considerados os precursores, com suas técnicas fotográficas, do cinema, que estava para ser inventado com o cinematógrafo dos irmãos Lumière e o cinetoscópio de Thomas Edison.

A rotoscopia, que é uma forma manual de capturar o movimento em cenas filmadas, e foi desenvolvida por Max Fleischer e seus irmãos em torno de 1915, é precursora da técnica digital como a conhecemos hoje. A intenção de Fleischer era fornecer um método de mecanização da animação, que reduziria custos e tempo de produção. Fleischer utilizou a rotoscopia para animar as séries *Out of the Inkwell*, começando em 1915 com o palhaço em *Koko, the Clown*, usando cenas filmadas como base para fazer as animações [Fleischer 2005].

A rotoscopia pode ser considerada um ancestral direto da captura de movimento como a conhecemos hoje. Ela é uma técnica na qual os animadores copiam o movimento gravado, quadro a quadro, para usá-lo em desenhos animados. As imagens do filme são projetadas sobre uma placa de vidro e redesenhadas pelo animador, quadro a quadro, de trás para frente. Esse equipamento é chamado de rotoscópio e essa técnica é ainda usada em estúdios de animação tradicionais para copiar movimentos reais filmados e aplicar em personagens de desenho animado.

Algumas das tecnologias usadas atualmente para a captura de movimento existem desde a década de oitenta, utilizadas em aplicações com propósitos médicos e militares. David Sturman [1994] e Menache [2000] fazem revisões históricas dessa evolução nos primeiros anos da captura digital de movimento. De toda essa evolução citaremos alguns momentos marcantes.

Em 1984 foi feita a primeira animação para publicidade utilizando a captura de movimento em computadores. Uma associação formada pelos maiores fabricantes de comida enlatada encomendou uma animação feita dessa maneira para ser veiculada durante o Super Bowl, em janeiro de 1985. Eles queriam associar a imagem de modernidade aos enlatados, e a Apple havia lançado o comercial *1984* durante o campeonato de basquete de 1984 e tinha sido um grande impacto. Esse desafio foi feito à empresa

Robert Abel and Associates. Bob Abel tinha um software que, com melhoramentos, poderia animar o primeiro personagem virtual. No grupo de Abel participavam Bill Kovaks e Roy Hall, co-fundadores da futura Wavefront Technologies, Con Pederson, co-fundador do futuro Metrolight Studios, Charlie Gibson, vencedor do Oscar como supervisor de efeitos especiais em *Babe* (1995), dentre tantos outros colaboradores. Foram utilizadas três câmeras de 35 mm para triangular as posições 3D dos pontos na atriz real. Eles trabalharam utilizando o primeiro computador da Silicon Graphics, o SGI Iris 1000, para desenvolver o personagem. Finalmente conseguiram animar o personagem e o desafio final foi processar o comercial de 30 segundos nas duas semanas que faltavam para terminar o prazo de entrega. Para isso foram utilizados 60 computadores VAX 750 em vários lugares diferentes dos Estados Unidos. Tudo foi finalizado dois dias antes do prazo final. O nome do comercial é *Brilliance*, ou *Sexy Robot*, e foi apresentado no campeonato de Super Bowl em janeiro de 1985 [Menache 2000].



Figura 1: *Sexy Robot*, primeiro personagem animado por captura de movimento

No começo dos anos 90 a captura de movimento começava a se apresentar como parte confiável de projetos de computação gráfica. Várias empresas buscavam aplicações em tempo real, como Medialab, Mr. Film, SimGraphics, Brad deGraf, Windlight Studios, e outros, como Tsi, Biovision e Acclaim, desenvolviam aplicações não em tempo real, com ênfase no mercado de videogames. É importante ressaltar que a indústria de videogames foi a grande responsável pela sobrevivência durante anos das empresas que trabalhavam com mocap para animação. Como os movimentos não precisavam ser muito precisos e os personagens eram à época mais rudimentares, ela foi uma ferramenta amplamente utilizada nesse mercado.

Uma produtora francesa de computação gráfica, a Medialab - a mesma que criou e realizou o show de televisão *Donkey Kong Country*, impedido de disputar o Emmy de 1999 por utilizar a captura de movimento - desenvolveu, em 1991, um projeto para

criar personagens utilizando a captura de movimento. O primeiro personagem foi *Mat, o Fantasma*, que foi usado em um programa infantil com aparições diárias de um minuto. Ele interagiu com cenas gravadas previamente, e ficou no ar durante três anos e meio. *Mat* era controlado por vários animadores. Dois controlavam os movimentos faciais e o sincronismo labial, um ator encenava os movimentos corporais do tronco e braços, e assim por diante. O Medialab continuou a desenvolver esse sistema, e criou dezenas de personagens para televisão.

Dessa fase inicial de desenvolvimento da captura de movimento, vale citar os esforços da SimGraphics no desenvolvimento da captura de expressões faciais. No início dos anos noventa, utilizando sensores mecânicos presos a partes importantes de movimento do rosto, e sensores eletromagnéticos na estrutura de sustentação, podia-se capturar os movimentos mais importantes da face e mapeá-los em personagens de computador em tempo real. Um só ator manipulava todas as expressões faciais do personagem simplesmente fazendo a mímica em si mesmo. O exemplo mais famoso de personagem usando esse sistema é Mario, de 1992, do jogo da Nintendo, e que tinha também um programa de televisão onde interagiu em tempo real com o público.

Em 1992 foi produzido o vídeo para a música *Steam*, de Peter Gabriel, que utilizava captura de movimento em diversas cenas. Este foi o primeiro vídeo de música utilizando essa técnica e ganhou o Grammy de Vídeo do Ano de 1993.



Figura 2: cena do clipe *Steam*, no qual Gabriel líquido contracenava com duas mulheres de fogo, animados com mocap [Menache 2000].

O primeiro filme a utilizar a captura de movimento com sucesso foi o *Exterminador do Futuro 2* (*Terminator 2: Judgement Day*), de 1991, dirigido por James Cameron. Antes desse filme, a técnica tinha sido empregada e depois descartada em uma seqüência do filme *Total Recall*, de 1990, dirigido por Paul Verhoeven, também estrelado por Arnold Schwarzenegger, e que no Brasil recebeu o nome de *O Vingador do Futuro*.

Para o *Total Recall*, o Metrolight Studio fazia a captura de movimento com um sistema óptico que

tinha sido lançado há pouco tempo. O filme estava sendo rodado na Cidade do México, e decidiu-se mandar para lá o sistema de captura e um operador fornecido pelo fabricante do equipamento. A seqüência a ser trabalhada era aquela em que o personagem de Schwarzenegger, com soldados e vários extras, atravessava um aparelho de raio-X, e suas imagens apareciam como esqueletos caminhando, como na figura 3, que é um fotograma da seqüência. A captura de movimento não forneceu os dados pretendidos e se abandonou o projeto, optando-se pela animação por *keyframe*.



Figura 3: seqüência dos esqueletos, que poderia ter sido a primeira cena do cinema com captura digital de movimento.

No *Exterminador do Futuro 2* a técnica de captura de movimento foi utilizada para animar o robô T-1000, encenado por Robert Patrick. O robô era de metal líquido, e veio do futuro para enfrentar o robô T-800, encenado por Schwarzenegger, e também mandado do futuro pela Skynet. Dennis Muren e equipe levaram o Oscar de Melhor Efeito A captura de movimento foi feita pela Pacific Data Images, PDI, empresa que trabalha com captura de movimento desde os primeiros protótipos, e a Industrial Light & Magic.



Figura 4: robô T-1000, primeiro personagem animado por captura de movimento da história do cinema.

Lançado em 1995, o game *FX Fighter* é o primeiro jogo de luta em tempo real, com personagens e cenários 3D. Ele também foi um dos primeiros jogos a utilizar mocap para animar os personagens. Nos games com captura de movimento, os personagens são

animados com o usuário dando entrada em um conjunto de ações animadas por mocap, como corrida, pulo, chutes, socos. Os pedaços de ações aparecem de tal maneira que o usuário não percebe a mudança entre elas, parecendo que ele tem total controle dos movimentos dos personagens. Cada conjunto de ação é arquivada em bancos de dados e a mudança entre esses bancos de dados são provocados pelo usuário. O sucesso desse jogo incentivou outras empresas a utilizarem essa tecnologia em seus games [Kitagawa e Windsor 2008].



Figura 5: cena do jogo FX Fighter, em sua primeira versão.

Esses projetos pioneiros indicaram as vantagens e desvantagens da utilização de diferentes princípios físicos para a captura de movimento e as limitações relacionadas à aquisição de dados. Atualmente muitas das idéias originalmente utilizadas permanecem, com o emprego de hardwares com capacidade de processamento que permite fazer a animação em tempo real. Esses primeiros esforços ajudaram a pavimentar a estrada por onde viria toda a indústria de animação, com o lançamento de plug-ins para todos os principais programas de modelagem de personagens 3D que permitem uma interface de animação mais amigável e economia de tempo na realização das produções para cinema e jogos eletrônicos. As colaborações entre os realizadores desses primeiros projetos de pesquisa e produtoras de efeitos visuais para cinema consolidaram o uso dessa técnica em diversos filmes que se destacaram ao abrir novas perspectivas de realização cinematográfica.

A discussão sobre a mocap servir ou não à arte e à animação atingiu seu ápice em 1999, quando o show animado de televisão *Pais de Donkey Kong (Donkey Kong Country)*, que usava captura de movimento juntamente com animação por *keyframes*, foi rejeitado para disputar o Emmy sob a argumentação de que não era animação. Esse programa de televisão foi transposto para o formato de jogo eletrônico, sendo aí também um grande sucesso.

Em 2003 ocorreu um cenário oposto ao da recusa de *Donkey Kong Country* para concorrer ao Emmy. A mocap foi utilizada para criar o personagem Gollum na trilogia dirigida por Peter Jackson, *O Senhor dos Anéis (Lord of the Rings)*, no segundo e terceiro filmes. O personagem teve uma participação

central na história e contracenou diretamente com atores reais, como se todos estivessem sendo filmados ao mesmo tempo. Houve então, à época, uma discussão intensa para se criar uma nova nomeação para o Oscar, para o ator que dava vida ao personagem animado por captura de movimento.



Figura 6: personagens de *Donkey Kong Country*.

## 4. Tecnologias de Captura de Movimento

### 4.1. Introdução

Os dispositivos para se fazer a captura de movimento podem ser classificados em ativos ou passivos, síncronos ou assíncronos, com marcadores ou sem marcadores, e/ou de acordo com os princípios físicos empregados [Moeslund et al 2006]. A taxonomia aqui adotada é a partir dos princípios físicos usados, mas as divisões de acordo com as outras definições serão estudadas. Podem-se dividir os diferentes tipos de sistemas de captura de movimento em três princípios físicos básicos: mecânico, magnético e óptico. O sistema mecânico, por sua vez, pode ser inercial, acústico ou se basear em próteses.

Na captura de movimento acústico transmissores que emitem som são colocados no sujeito. Receptores de áudio em torno dele medem o tempo que leva para que o som vá dos transmissores ao receptor. Por trilateração chega-se à localização dos transmissores, que, em geral, são colocados nas juntas do sujeito. No sistema protético, uma estrutura externa é colocada em algumas partes do corpo. Nessa estrutura estão sensores que medem a angulação e orientação da estrutura, através de piezelétricos, por exemplo, e dessa maneira o movimento pode ser analisado. No sistema inercial, giroscópios e acelerômetros são posicionados nas articulações para capturar o movimento.

Na captura magnética os transmissores emitem campos magnéticos e os receptores medem a sua orientação com relação a um campo do qual se sabe a intensidade e sentido. Na captura óptica os transmissores são refletores ou fontes de luz e os receptores são câmeras, chegando por triangulação à posição dos transmissores.

Um sistema de captura de movimento de acordo com o que foi exposto acima pode ser esquematizado como no diagrama da figura 7. A complexidade do módulo de análise depende do nível dos dados enviados pelo módulo de percepção do movimento.

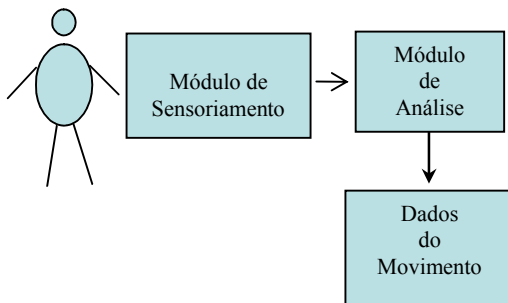


Figura 7: diagrama dos diversos componentes de um sistema de mocap [Moeslund 2000].

O sistema ativo usa dispositivos no sujeito que transmite ou recebe sinais. Quando o dispositivo funciona como transmissor, ele gera um sinal que pode ser medido por outro dispositivo nas vizinhanças. Quando ele funciona como receptor, o sinal é geralmente gerado por uma fonte em suas vizinhanças. Marey foi o primeiro a utilizar uma mocap ativa, em 1873, quando usou sensores pneumáticos e câmeras de pressão sob os pés para medir a pressão no pé ao caminhar. O sensor magnético é um exemplo de dispositivo usado em um sistema ativo, como discutido mais à frente.

No sistema passivo os dispositivos não afetam as suas vizinhanças. Eles simplesmente observam o que já está no universo, como por exemplo as ondas eletromagnéticas, e não geram novos sinais. A idéia é usar uma imagem obtida de uma câmera e capturar o movimento baseado naquela imagem. Muybridge foi o primeiro a utilizar essa técnica, em 1878, quando demonstrou que o cavalo tira as quatro patas do chão.

Por interação síncrona entende-se aquela na qual o movimento capturado é imediatamente usado para controlar algo no mundo real ou virtual. A interação assíncrona é assim chamada por ser utilizada depois de gravada e elaborada.

Aqui será dada ênfase à captura passiva óptica, que é a que vem tendo maior receptividade em aplicações de animação e é o princípio empregado no protótipo que está sendo construído pelos autores desse tutorial. Serão abordadas brevemente as capturas mecânica e a magnética. Um dos autores desse tutorial também aborda esses métodos em sua dissertação de mestrado [Flam 2009].

#### 4.2. Sistema mecânico

O sistema com próteses é menos prático de ser usado para captura de movimento do corpo inteiro, porque demanda uma série de conexões para medir as variáveis mecânicas utilizadas, como exemplo a

pressão. As conexões ficam pendentes no corpo do ator ou de onde se quer medir o movimento. O primeiro aparato desse tipo que se tem notícia foi o utilizado por Étienne-Jules Marey, descrito no livro *Animal mechanism: a treatise on terrestrial and aerial locomotion* de 1873.

Para movimentos da mão, o dispositivo mecânico é muito útil. No artigo de Molet [1997], os movimentos dos dedos e da mão são capturados por uma luva desenhada para esse fim, e os dados são mandados para um software. Ele simula o posicionamento da câmera de acordo com a orientação da luva com relação à cabeça da pessoa. Essa orientação é captada com sensores magnéticos, que dão a orientação dela com relação ao campo magnético da Terra, no caso. Foram desenvolvidos diversos protótipos de luvas para capturar os movimentos sutis da mão e dedos. Na figura 8 vê-se uma luva desenvolvida para capturar movimentos da mão, chamada *CyberGlove*. Neste caso, mede-se a inclinação dos dedos através da posição de suas pontas, usando cinemática inversa. O sistema é do tipo protético, e medem-se os ângulos através de uma liga metálica que transmite uma tensão à medida que é dobrada.

Os sistemas acústicos de captura de movimento foram testados até o início da década de noventa, mas demonstraram ser soluções menos versáteis que os sistemas magnéticos e ópticos. Esses últimos são mais precisos na localização dos pontos e mais compactos quanto ao tamanho do aparato.

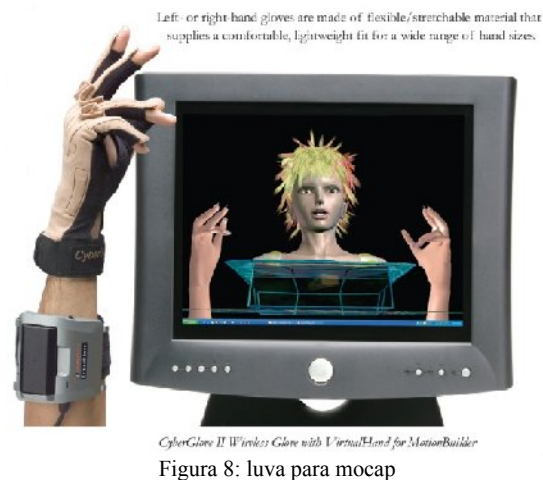


Figura 8: luva para mocap

Nos sistemas inerciais, marcadores ativos contendo giroscópios e acelerômetros são posicionados estrategicamente nas articulações de uma roupa vestida pela pessoa alvo da captura. Após uma etapa de calibração, é possível usar as informações de rotação dos giroscópios e de posicionamento dos acelerômetros para aquisição direta dos dados.

Os sistemas inerciais são portáteis e quase não possuem restrições com relação ao local de captura. Alguns até permitem obter dados em ambientes subaquáticos [Animazoo, 2008]. Além disso, não estão sujeitos a efeitos de oclusão e interferência nos dados adquiridos. Outra vantagem é que grande parte deles

usa tecnologia de transmissão de dados sem fio. Isso garante boa liberdade de movimento e um volume extenso de gravação. Porém, é preciso usar baterias, o que limita o tempo de captura sem recarga ou implica na necessidade de possuir baterias reservas.

O custo de sistemas inerciais é elevado, comparável aos melhores sistemas ópticos. A principal razão é o nível de tecnologia empregado na miniaturização dos sensores. Portanto, o valor final de um sistema desse tipo depende essencialmente do número de marcadores utilizados. Um exemplo é o sistema *Animazoo IGS 190* que possui 18 sensores e captura movimento a uma taxa máxima de 120Hz. É ideal para a captura de corpo inteiro de um indivíduo e seu preço é de 75 mil dólares. É possível também comprar somente a parte superior ou inferior do torso, custando aproximadamente 52 e 43 mil dólares [Inition 2009].



Figura 9: sistema inercial IGS-190 (Animazoo).

#### Vantagens:

- o intervalo de captura pode ser grande;
- custa menos que os sistemas magnéticos e ópticos;
- o sistema é em geral portátil;
- captura em tempo real é possível;
- os sensores nunca sofrem oclusão;
- é possível capturar o movimentos de vários atores ao mesmo tempo com vários sistemas.

#### Desvantagens:

- tem uma taxa de amostragem muito baixa;
- ele é complexo devido à quantidade de detalhes mecânicos dos equipamentos;
- traz limitações ao movimento das juntas humanas;
- a maior parte dos sistemas não calculam deslocamentos globais sem a ajuda de sensores magnéticos.

### 4.3. Sistema magnético

Os sistemas magnéticos continuam em uso e existem diversos fabricantes desses dispositivos, como a empresa Ascension. Os equipamentos têm evoluído procurando solucionar problemas relacionados ao princípio físico utilizado, e estão atualmente

apresentando equipamentos chamados de terceira geração [Anisfeld 2006].

Estes sistemas utilizam campos magnéticos gerados em emissores localizados nas junções do corpo. Esses campos tênues são comparados ao campo magnético da Terra no local onde está sendo medido. Para isso, antes de começar a sessão de captura tem-se sempre que calibrar o equipamento. O sistema é um dispositivo ativo, como definido anteriormente.

Os principais problemas relacionados a esse sistema são os cabos de alimentação que ficam presos ao corpo do ator, limitando os seus movimentos, e a interferência externa no campo magnético. Essa interferência limita a atuação de atores cujos movimentos estão sendo capturados e estão muito próximos, e também quanto à escolha de locais apropriados para a atuação, onde não existam interferências externas no campo gerado pelos emissores.

Os sistemas usam marcadores ativos presos nas articulações de uma roupa vestida por um ator. Eles são capazes de medir o fluxo magnético entre bobinas presentes no próprio marcador e bobinas estacionárias em posições conhecidas. Assim, é possível determinar a translação e orientação de cada marcador. É, portanto, um tipo de sistema de aquisição direta.

A liberdade de movimento em sistemas magnéticos é menor se comparada aos sistemas ópticos. Necessariamente, os marcadores utilizados possuem baterias ou fios que limitam o movimento do ator e o volume de captura. Além disso, os dados produzidos podem ser corrompidos e distorcidos devido a interferências elétricas ou magnéticas. Ou seja, existem também restrições em relação ao local de gravação. Preferencialmente, o ambiente escolhido não deve possuir objetos metálicos e aparelhos eletrônicos sem blindagem eletromagnética.

Uma vantagem de sistemas magnéticos é a ausência de oclusão e o custo reduzido em relação aos sistemas ópticos. Um exemplo é o sistema da *Ascension*, o *Flock of Birds* da Figura 10, com 12 sensores, capaz de gravar movimento a 144Hz e adequado para a captura de movimento humano. Pode ser comprado por aproximadamente 25 mil dólares [Inition 2009].

#### Vantagens:

- os dados podem ser fornecidos para os personagens em tempo real;
- dados de posição e orientação ficam disponíveis sem necessidade de processamento adicional;
- custam menos que os sistemas ópticos, na faixa de 5000 a 15000 dólares;
- os sensores nunca sofrem oclusão;
- é possível capturar vários atores simultaneamente utilizando vários sistemas, estando atento às interferências.

#### Desvantagens:



- a sensibilidade dos rastreadores magnéticos a metais pode levar a uma saída de sinal com ruído;
- os movimentos dos atores são limitados pelos cabos em geral;
- tem uma taxa de amostragem menor que dos sistemas ópticos;
- a área de captura, isto é, o set de captura é menor que o que é possível com os sistemas ópticos;
- é difícil mudar a configuração dos marcadores.



Figura 10: dispositivo magnético *Flock of Birds* montado para captura de movimento (Ascension).

#### 4.4. Sistema óptico

Os sistemas ópticos podem ser divididos em ativos e passivos, com ou sem marcadores. No sistema ativo os marcadores são fontes de luz, em geral LEDs colocados nas junções do corpo. O sistema passivo usa refletores como marcadores. Em ambos os casos os detectores são câmeras de vídeo, em geral com dispositivos cmos, que permitem uma maior precisão na localização dos marcadores com relação aos dispositivos com ccd.



Figura 11: sistema óptico *Impulse* (PhaseSpace Tech)

Se marcadores não forem usados, os pontos de interesse podem ser determinados por meio de detecção de silhueta ou componentes estruturais (cabeça, ombro, perna, mãos e etc.). Uma técnica recente simplifica esse processo por meio de uma maquiagem feita por tintas fosforescentes, usada no produto *Mova Contour Reality Capture* (Mova LLC, 2008).

A ênfase será dada ao equipamento de captura passiva óptica. Nesse caso as câmeras de vídeo são retro-iluminadas com LEDs e os marcadores ópticos são retro-refletores. Na figura 12 vê-se a câmera com os LEDs acesos formando um círculo ao redor da lente que captura a imagem. A luz dos LEDs ilumina o set de gravação. O princípio físico utilizado apresenta maior flexibilidade, podendo ser deslocado e instalado nos maiores ambientes possíveis para a técnica, e não tem fios ou acessórios que prejudiquem a performance do ator, possibilitando maior liberdade de expressão.



Figura 12: sistema de captura com os LEDs acesos em torno da lente de entrada da câmera, retroiluminada (Vicon).

Os movimentos do ator são monitorados utilizando-se marcadores ao longo do seu corpo. Eles são colocados em articulações e pontos-chave do corpo para a movimentação. A performance do ator é capturada por câmeras de cmos, nas quais as posições dos marcadores são bidimensionais. Utilizando um arranjo de câmeras podemos recuperar as coordenadas espaciais dos marcadores, por triangulação, ao mandar os sinais das câmeras para um computador e compará-los através de um software. Estas coordenadas são aplicadas em pontos pré-determinados do personagem virtual e, com isso, o movimento da pessoa real é utilizado para animar o personagem virtual. Assim como se usa uma pessoa como modelo real para capturar o movimento, poderia se utilizar objetos ou animais, modificando-se a modelagem 3D e a localização dos marcadores.

#### Vantagens:

- os dados ópticos são muito precisos na maior parte dos casos;

- um número maior de marcadores pode ser usado. Lembrar do sistema utilizado no *O Expresso Polar*, que tem 600 marcadores;
- é fácil mudar a configuração dos marcadores;
- é possível obter aproximações com esqueletos internos utilizando grupos de marcadores;
- os atores não são limitados por cabos;
- permite uma área de captura maior que em todos os outros sistemas;
- possuem uma maior frequência de captura, permitindo assim uma taxa de amostragem maior.

#### **Desvantagens:**

- requer uma pós-produção, isto é, um posterior processamento dos dados mais intenso;
- o sistema é mais caro, variando entre 80.000 e 250.000 dólares;
- não pode capturar movimentos quando os marcadores ficam oclusos por períodos longos de tempo;
- a captura de movimento tem que ser feito em um ambiente controlado, sem muitos pontos de reflexão para evitar interferência nos resultados.

## **5. Capturando o Movimento**

### **5.1. Preparando a sessão de captura de movimentos**

Como a captura de movimento é utilizada para animar um personagem virtual, devemos então focar a atenção nele, em como ele é criado e em como ela lhe dá vida. A construção e conceituação do personagem, enfim, sua personalidade, é fundamental para definir como a captura será utilizada. Por personalidade se entende a sua aparência física, os seus movimentos básicos, suas ações e reações a acontecimentos e a outros personagens.

Os seguintes princípios são difíceis de serem trabalhados com a captura de movimento: comprimir e esticar, antecipação, continuidade e exageração além dos limites físicos. Existem métodos manuais e procedimentos para adicionar alguns desses princípios à captura de movimento, depois que os dados são colhidos. Deve-se analisar a relação custo/benefício caso se devam utilizar esses métodos, pois muitas vezes é mais eficaz utilizar a animação por *keyframes*. Existem diversos exemplos de produções que desistiram de usar a mocap por causa dessas limitações, algumas mostradas anteriormente.

Na pré-produção de uma animação que se queira utilizar a captura de movimento deve-se primeiro estudar as características do personagem virtual. A animação por mocap é bastante realista, e o ideal é fazer alguns testes prévios de como vai se portar o personagem com os movimentos captados. Se for então decidido por esse processo de animação, deve-se fazer um planejamento minucioso de cada etapa da produção, isto é, das sessões de captura, para

que os dados coletados cheguem com menos modificações para se fazer na fase de animação, ou seja, para a pós-produção.

O planejamento das sessões leva em conta como os personagens virtuais estarão em cena e em que tipo de formato. Pode ser um jogo eletrônico, um filme, um desenho animado, uma telenovela ou uma interação com um cliente pela internet. Deve-se saber se o personagem cumprirá um roteiro preciso ou se será construída uma biblioteca de movimentos que serão escolhidos de acordo com a demanda do jogo ou da interação com o cliente. Em qualquer dos casos, leva-se em conta o cenário em que o personagem desenvolverá a ação e quais tipos de interação com o ambiente e com outros personagens.

No caso de se desejar construir uma biblioteca de movimentos para um jogo eletrônico, tem que se fazer uma listagem dos movimentos, com o tipo de enquadramento. Se houver continuidade entre os movimentos, isto é, o início de um começar na seqüência do final de outro, isso deve entrar no detalhamento da lista. Com a lista concluída, organizam-se as sessões de captura, agendando-as de acordo com o tipo de movimento e de enquadramento para que seja o mais eficiente possível.

O ideal é que se tenha um roteiro detalhado em tomadas e seqüências, com as falas e situações em que o personagem se envolve. A partir disso, organizam-se as sessões de captura levando em conta os tipos de cenários, personagens e enquadramento, para que sejam capturados os movimentos utilizando critérios que agilizem o fluxo de trabalho. Tudo isso deve ser discriminado no plano de gravação. Na decupagem das cenas, caso haja interação do personagem com outros personagens ou com objetos, elas devem ser divididas por partes, de acordo com a interação e movimento. Isso facilita muito o trabalho de pós-produção, na etapa de elaboração dos dados.

Em todas as situações, deve-se adotar um critério minucioso para nomear os arquivos, pois em geral eles são em grande número. E isso ajuda na pós-produção e depois para arquivá-los para posterior uso.

Tendo em mãos o roteiro e o plano de gravação, podem-se iniciar as sessões de captura. Nesse momento deve-se estar atento às limitações do equipamento, baseado nos princípios físicos utilizados, e do software de captura. No caso de sensores magnéticos, interferências devido a metais próximos, e para os marcadores ópticos, oclusão, são as limitações principais. A melhor situação é quando se tem já o modelo do personagem virtual a ser animado, ou mesmo uma figura de palitos, para que se possam aplicar os movimentos onde haja dúvidas quanto ao resultado. Dessa maneira, pode-se corrigir a tempo imperfeições introduzidas pelo equipamento, evitando problemas futuros na pós-produção.

É importante estar atento ao número de personagens que estão contracenando e à posição dos marcadores. Antes do início das sessões deve-se definir onde ficarão os marcadores e não se deve mudá-los de lugar. Se isso ocorrer, a nova configuração tem que ser mapeada detalhadamente. A medida que as tomadas

vão sendo feitas, os arquivos devem seguir a nomenclatura de acordo com os critérios estabelecidos pela equipe de captura, para que possam ser acessados com facilidade. O número de arquivos pode chegar a algumas centenas e até milhares. Esses cuidados precisam ser tomados com certo rigor para que não se inviabilize o trabalho por questões corriqueiras, mas que podem significar uma grande dor de cabeça na pós-produção. Com isso, o trabalho fica mais ágil e pode-se prever com mais facilidade o tempo para finalizar a animação dos personagens.

Os dados devem chegar à etapa de finalização sem muitas alterações por fazer. A performance do ator tem que ser a mais próxima do que se pretende da personalidade do personagem. Alterar manualmente a intensidade dos movimentos é muito trabalhoso na pós-produção, e muitas vezes impossível. Para que não haja aumento no tempo de trabalho e custo além do previsto é que o planejamento tem que ser o mais minucioso possível. Muitos dados deverão ser trabalhados na fase de finalização, devido à oclusão dos marcadores ou ruídos introduzidos na captura, mas devem estar dentro de níveis aceitáveis, isto é, que não afetem o cronograma significativamente, como ocorre quando a produção não é bem organizada.

Seguindo essas regras básicas e utilizando-se a técnica em um contexto em que ela não comprometerá o enredo da história, pode-se esperar um resultado bem sucedido, podendo ao final reduzir o tempo e o custo com relação à animação tradicional ou computação gráfica por *keyframe* [Liverman 2004].

## 5.2. Marcadores

Os pontos monitorados no corpo do ator recebem duas denominações diferentes. Eles são chamados de sensores quando o sistema é magnético e de marcadores quando é óptico, ou de marcadores em ambos os casos. Será adotada essa última convenção, de denominá-los sempre como marcadores. Para o sistema magnético, a roupa vem com os sensores no sistema comercializado, e tem-se pouca liberdade de modificar essas posições. Para os marcadores ópticos têm-se mais possibilidades de modificar suas localizações ao longo do corpo assim como o tamanho dos mesmos.

Para seguir o movimento do corpo nas situações em que o ator estará fazendo a performance, os marcadores devem estar fixos e presos aos pontos escolhidos no corpo. Para se conseguir isso podem usar roupas de lycra justas ao corpo, com os marcadores presos a ela. Uma outra alternativa é prendê-los diretamente no corpo, com o ator usando calção e deixando os braços descobertos. Nesse caso, os marcadores são posicionados usando faixas de lycra com velcro, adesivo de dupla face ou cola especial no corpo. O que é importante em todas as situações é que os marcadores não se soltem ou se movimentem.

Para o sistema óptico, os marcadores para o corpo podem ser pequenos *leds*, bolas ou discos refletivos. É conveniente usar a radiação de iluminação ou dos *leds*

no infravermelho, com filtros, para deixar passar somente essa frequência, posicionados na frente das câmeras. O diâmetro desses marcadores pode chegar a 5 mm, dependendo da resolução das imagens.

As posições dos marcadores devem permanecer as mesmas durante as sessões de captura, pois elas têm uma equivalência com as articulações do personagem virtual. Caso se mudem essas localizações, devem ser bem especificadas e mapeadas para que as mesmas sejam alteradas na pós-produção. O número de marcadores depende da capacidade de processamento do computador que está sendo usado na sessão de captura, e do tempo que se tem disponível para fazer o processamento dos dados. Com o aumento dessa capacidade, tem-se chegado a números que permitem um nível de detalhamento e realismo cada vez maiores e em tempo real. No caso do sistema utilizado no O Expresso Polar, utilizaram-se 600 marcadores nos corpos dos atores, levando quase duas horas para posicioná-los antes de cada sessão de captura.

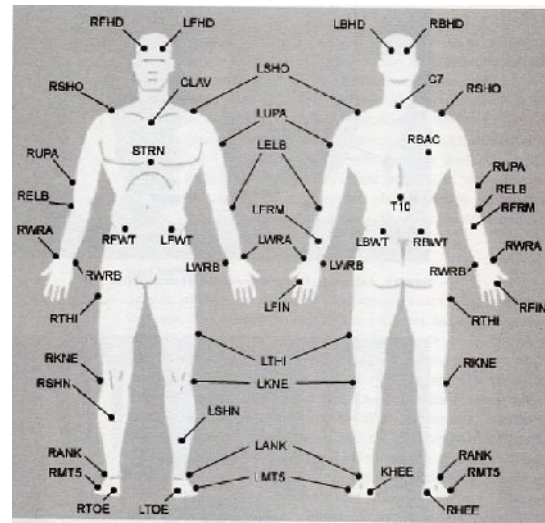


Figura 13: posições para 41 marcadores [Menache 2000].

A colocação dos marcadores deve levar em conta três questões básicas: 1. definir a posição ou orientação de uma junção; 2. para se fazer distinção entre direita e esquerda e entre os diferentes performáticos; 3. para ajudar a reconstruir os dados de outros marcadores, que se perderam ou estiverem faltando. Então há marcadores para ajudar a definir o esqueleto e seus movimentos, outros para distinguir a esquerda da direita e para identificar os performáticos, e outros para achar-se marcadores que ficarem ocultos. Baseado nesses três tipos tem-se a proposta de posicionamento de 41 marcadores na figura 13.

Para os movimentos faciais pode-se fazer uma sessão separada para esse objetivo, ou não, dependendo do software de captura, da capacidade de processamento e da resolução das câmeras. Os marcadores faciais são menores, entre 3 e 4 mm, e são colocados diretamente na pele, com fita de dupla face

ou cola especial. O número deles vai de 12 a 150, como no sistema da Sony Pictures Imageworks.

Na captura do movimento das mãos usam-se os marcadores de 3 e 4 mm também. Eles são empregados para dar a orientação do antebraço e para movimentos gerais da mão. Os movimentos mais detalhados dos dedos são capturados ainda separadamente, e utiliza-se um sistema com 25 marcadores, em geral.

A entrada do sistema de captura óptico é então uma seqüência de imagens registradas pela câmera. A saída são as coordenadas das posições dos marcadores para cada imagem. No computador a imagem é segmentada para extrair as regiões correspondentes aos marcadores na cena original. A segmentação é baseada em ferramentas de processamento de imagem, principalmente transformações morfológicas, que são representadas por algumas operações básicas. Através desse processo, localizam-se os centros dos marcadores e seguem-se as suas posições através das imagens, isto é, através do tempo. A quantidade de imagens por segundo é a taxa de amostragem das mesmas. Essa taxa deve ser igual ou maior que a taxa de quadros do formato com o qual se está trabalhando, cinema ou vídeo.

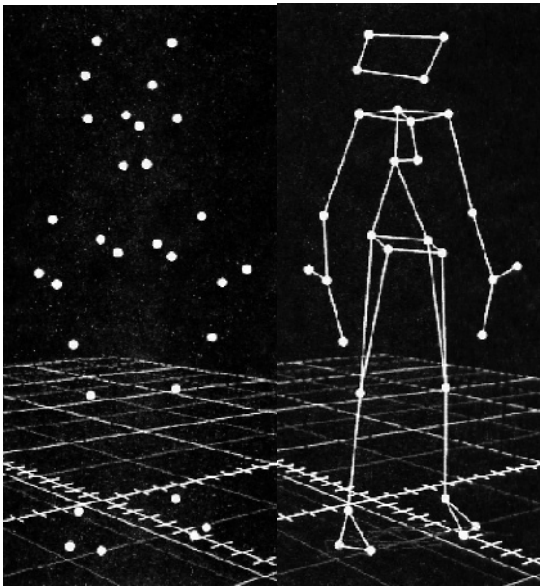


Figura 14: dados sem grande quantidade de processamento, com os marcadores como se fosse uma nuvem, sem conexões e conectados por um boneco de palitos [Menache 2000].

Para seguir os marcadores através dos fotogramas, depois da segmentação e da localização dos pontos, vem a predição de onde eles estão nos próximos quadros. Esse processo é necessário para reduzir a região onde vão ser procurados os marcadores nas próximas imagens. Assim, tem-se a evolução temporal dos pontos no corpo de quem faz a performance, que corresponderão a pontos no personagem virtual.

Terminado o processo de identificação das posições dos marcadores através do tempo, os dados disponíveis são relativos a cada câmera, com os deslocamentos

bidimensionais dos pontos. Através de cálculos de triangulação dos dados de cada câmera feitos a seguir, baseados na posição das mesmas, chega-se aos valores tridimensionais dos marcadores no corpo do ator.

### 5.3. Preparando o personagem

Os personagens podem ser criados em diversos programas de modelagem 3D, como o Maya, o Softimage, o 3D Studio Max e o Blender3D. Todos esses programas possuem plug-ins que convertem os dados gerados no processo de captura de movimento, como descrito anteriormente, e os aplicam em pontos pré-determinados, equivalentes às posições dos marcadores no ator.

No software de modelagem 3D, o personagem virtual é construído com estruturas articuladas. Elas são conjuntos de objetos rígidos, que seriam os ossos, por exemplo, conectados por articulações. As articulações formam o vínculo geométrico entre os objetos rígidos, permitindo o movimento relativo entre eles. As estruturas articuladas são representadas por estruturas hierárquicas (árvores), onde a posição de cada articulação é definida através da composição em seqüência das transformações das articulações anteriores. Com isto, apenas a primeira articulação da estrutura precisa ser posicionada no espaço, enquanto o resto da estrutura é posicionado apenas pelos ângulos entre as articulações, chamados de ângulos relativos. São estas informações, a posição da articulação e o ângulo relativo entre o restante das articulações, os parâmetros da animação.

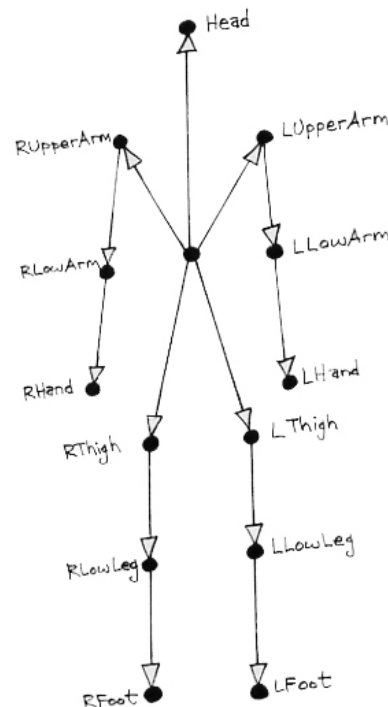


Figura 15: um tipo de hierarquia para o esqueleto [Menache 2000].

Os dados gerados no processo de captura vêm em formatos diversos, de acordo com as maneiras em que são organizadas as informações. As extensões para estes diversos formatos são .amc, .bva, .bvh, .trc. O formato a ser adotado no OpenMocap é o .bvh, desenvolvido pela empresa Biovision, especializada em análise de esportes e a animação. Esse formato é aceito por vários aplicativos de modelagem 3D, como o 3DSMax. O arquivo é dividido em duas sessões principais: hierarquia e movimento. A sessão de hierarquia inclui todas as definições necessárias para criar o esqueleto, e a sessão de movimento contém o fluxo de dados.

Com esse nível de tratamento técnico conclui-se a abordagem das sessões de captura e de processamento de dados. O nível de sofisticação técnica é o suficiente para que o artista e aquele que for utilizar-se do recurso disponível possa compreender o comportamento da captura digital de movimento nas diversas etapas do processo. Para a construção do protótipo, já está sendo aprofundada a abordagem técnica para escrever o aplicativo que fará a aquisição dos dados coletados pelas câmeras, a identificação dos pontos no corpo do ator, a evolução temporal de suas posições e a saída dos valores tridimensionais no formato apropriado para a utilização no personagem virtual pelo programa de modelagem 3D.

## 6. Arquitetura da Aplicação OpenMoCap

O OpenMoCap foi construído visando alcançar todos os objetivos de um sistema de captura de movimento de código aberto, empregando bons princípios de construção de software, descritos em McConnell [2004] e em Kernighan & Pike [1999]. Portanto, diversas decisões de arquitetura e implementação foram tomadas e são apresentadas a seguir.

A Figura 15 ilustra a arquitetura definida para a aplicação por meio de um diagrama de módulos. O software e o código foram desenvolvidos totalmente em inglês, almejando universalizar seu uso e o desenvolvimento de extensões. A separação dos módulos por fluxos de execução (do inglês *threads*) foi feita para aproveitar a tendência atual da indústria de processadores de produzir chips com cada vez mais núcleos. Tarefas empreendidas em tempo real, como a detecção de POI, o rastreamento, a triangulação e a visualização, puderam ser paralelizadas.

O fluxo principal de execução (Main Thread) é composto pelo núcleo da aplicação (MoCap Core) e pela interface gráfica principal (Main User Interface). O núcleo é responsável por inicializar corretamente todos os outros fluxos de execução e seus módulos e por disponibilizar também uma ligação entre eles por meio dos controladores de captura e de vídeo. Além disso, é um repositório central de informações, como qual é a configuração das câmeras conectadas, quais são os algoritmos disponíveis para realização do fluxo

de captura e quais são as semânticas possíveis para os POIs.

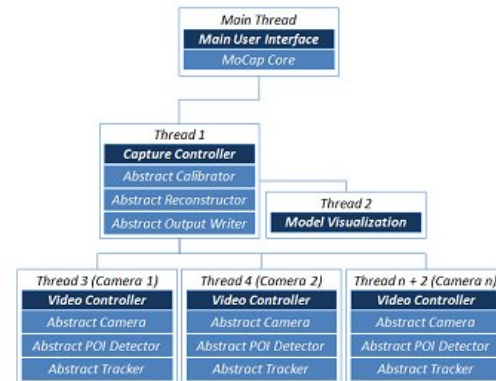


Figura 15. Diagrama da Arquitetura do OpenMoCap [Flam 2009].

A interface gráfica principal é ilustrada pela Figura 16, obtida do OpenMoCap em execução. Ela é responsável por exibir dados de captura e receber solicitações do usuário, acionando funções específicas dos controladores, necessárias à gravação de movimento. Existem basicamente quatro componentes de interação com o usuário: o menu, a barra de status, as janelas de câmera e a janela de visualização.

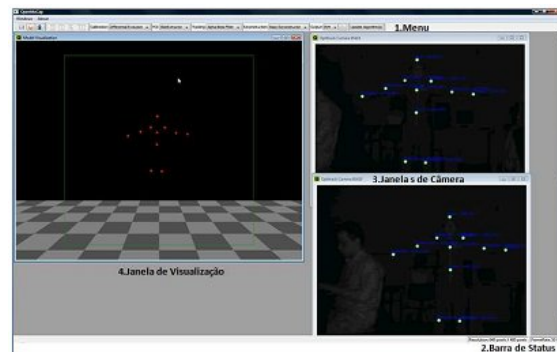


Figura 16. Interface Gráfica Principal do OpenMoCap [Flam 2009].

O menu permite, por meio de botões, o começo e o término de uma seção de captura e a estimação dos parâmetros das câmeras. Ele informa também o tempo de captura total e os algoritmos utilizados. Por outro lado, a barra de status exibe apenas a resolução das câmeras ligadas ao software e a taxa de quadros sendo processada por segundo. As janelas de câmera exibem as imagens obtidas por cada dispositivo existente no sistema e permitem a seleção de semântica dos POIs detectados. A janela de visualização exibe uma prévia em tempo real do movimento que está sendo gravado, por meio de um fluxo separado composto pelo módulo de visualização (Model Visualization). O modelo de interface de múltiplos documentos (do inglês Multiple Document Interface) foi adotado por proporcionar um

espaço comum e flexível para o uso das janelas existentes.

Cada uma das janelas de câmera é processada exclusivamente por um fluxo de execução, coordenado por um controlador de vídeo (Video Controller). Por sua vez, cada um desses controladores possui uma instância de câmera, do algoritmo de detecção de POI e do algoritmo de rastreamento, que são executados sequencialmente. Na Figura 15 isso é representado pelos módulos Abstract Camera, Abstract POI Detector e Abstract Tracker. Eles definem uma interface comum e a mais genérica possível para as câmeras e os algoritmos de processamento bidimensionais, citados anteriormente.

Esse conceito de módulos abstratos que definem interfaces é muito importante para garantir extensibilidade da aplicação, um dos objetivos propostos neste trabalho. Caso seja usado um novo modelo de câmera e esse equipamento não for compatível com a implementação atual, basta desenvolver funcionalidades específicas, como a obtenção de quadros e a mudança de resolução seguindo o padrão do módulo abstrato de câmera para aproveitar todo o fluxo já existente. Outro exemplo vantajoso é a possibilidade de substituir o algoritmo de detecção de POI por um baseado em reconhecimento de partes do corpo ao invés de intensidades de regiões, isto é, fazer com que o sistema possa capturar movimento sem marcadores. Portanto, é possível realizar alterações de otimização e novas funcionalidades, pontualmente e incrementalmente.

O último fluxo de execução, presente na Figura 15, é o composto pelo controlador de captura (Capture Controller). Ele é o principal responsável pelo funcionamento correto da aplicação e executa efetivamente uma seção de captura, gerenciando e processando as informações produzidas pelos controladores de vídeo e atualizando o fluxo do visualizador. Baseado no mesmo princípio de módulos abstratos, o controlador de captura possui uma instância de um algoritmo de estimação de parâmetros de câmeras (Abstract Calibrator), uma de um algoritmo de reconstrução (Abstract Reconstructor) e uma de um algoritmo para gerar saída de dados (Abstract Output Writer).

Até o momento foram discutidas decisões de mais alto nível de arquitetura, relacionadas principalmente aos conceitos de abstração, generalidade, separação de interesses e desenvolvimento incremental. Porém, decisões de mais baixo nível, ligadas à implementação em si, também foram feitas para garantir o aproveitamento de bibliotecas de código livre já existentes, o padrão de codificação e a modularidade.

Dada a natureza da aplicação construída, uma biblioteca eficiente e madura de visão computacional foi especialmente útil para seu desenvolvimento. A OpenCV, atualmente mantida pela Willow Garage [2009], foi a biblioteca selecionada desse tipo. Ela foi escrita inicialmente na linguagem de programação C e C++ para exibir a performance dos processadores da Intel Corporation, sua criadora.

## Agradecimentos

Esse trabalho teve o apoio financeiro do CNPq e da FAPEMIG.

## Bibliografia

ANISFELD, N. *The Rise of a New Art Form – The Birth of Mocap*, Ascension Technology Corporation technical report, que pode ser encontrado no endereço [www.ascension-tech.com/applications/pdf/birth\\_of\\_mocap.pdf](http://www.ascension-tech.com/applications/pdf/birth_of_mocap.pdf).

BARBOSA JÚNIOR, Alberto Lacerda, *Arte da animação, técnica e estética através da história*. Editora Senac, 2002.

FLAM, D.L., QUEIROZ, D.P., GOMIDE, J.V.B. e ARAÚJO, A. de A. *OpenMoCap: An Open Source Software for Optical Motion Capture*, artigo completo do SBGAMES 2009.

FLAM, D.L. *OpenMoCap: uma aplicação de código livre para a captura óptica de movimento*. Dissertação de Mestrado, Departamento de Ciência da Computação, UFMG, 2009.

FLEISCHER, Richard. *Out Of The Inkwell: Max Fleischer And The Animation Revolution*. University Press of Kentucky, 2005.

GOMIDE, J.V.B. *A captura de movimento para o cinema de animação*. Dissertação de Mestrado, Escola de Belas Artes, UFMG, 2006.

INATION. Motion Capture / Tracking from Inition. Documento disponível em <http://www.inition.co.uk/inition/products.php?CatID=11>.

KERNIGHAN, B. W., AND PIKE, R. 1999. *The practice of programming*. Addison-Wesley Longman Publishing Co., USA.

KITAGAWA, M. e WINDSOR, B. *MoCap for artists: workflow and techniques for motion capture*. Focal Press, 2008.

LIVERMAN, Matt. *The animator's motion capture guide: organizing, managing and editing*. Charles River Media, 2004.

MAREY, E. *Animal Mechanism: A Treatise on Terrestrial and Aerial Locomotion, 1873*, New York, Appleton, Republished as Vol. XI of the International Scientific Series.

MCCONNELL, S. *Code Complete, Second Edition*. Microsoft Press, Redmond, WA, USA, 2004.

MENACHE, Alberto. *Understanding motion capture for computer animation and video games*. Morgan Kaufmann, 2000.

MOESLUND, T. B, HILTON, A. e KRUGER, V. *A survey of advances in vision-based human motion capture and analysis*, in *Interaction in Virtual Inhabited 3D Worlds*, cap. 11. Springer-Verlag, 2000.

MOESLUND, T. B. *Interacting with a virtual world through motion capture*, in *Interaction in Virtual Inhabited 3D Worlds*”, cap. 11. Springer-Verlag, 2000.

MOLET, T. et al. *An Animation Interface Designed for Motion Capture*, *Computer Animation* 1997, 1997.

MOTION ANALYSIS CORPORATION, 2009. The industry leader fo3d passive optical motion capture. Documento eletrônico disponível em <http://www.motionanalysis.com/>.

MUYBRIDGE, Eadweard. *The Male and Female Figure in Motion : 60 Classic Photographic Sequences*. New York, Dover Publications, 1984.

STURMAN, David J. *A Brief History of Motion Capture for Computer Character Animation*. In “Character Motion Systems”, ACM SIGGRAPH 94 Proceedings, Florida.

THOMAS, Frank e JOHNSTON, Ollie. *Disney animation: the illusion of life*. Abbeville Press, 1984.

WILLOW GARAGE,. *OpenCV - Wiki*. Documento eletrônico disponível em <http://pr.willowgarage.com/wiki/OpenCV>.