

JOÃO VICTOR BOECHAT GOMIDE

**EFEITOS VISUAIS, DA TRUCAGEM ÓPTICA
À CAPTURA DIGITAL DE MOVIMENTO**

Escola de Belas Artes

Universidade Federal de Minas Gerais

2013

JOÃO VICTOR BOECHAT GOMIDE

**EFEITOS VISUAIS, DA TRUCAGEM ÓPTICA
À CAPTURA DIGITAL DE MOVIMENTO**

**Tese apresentada na Escola de
Belas Artes da UFMG como pré-
requisito à obtenção do título de
Doutor em Artes.**

ORIENTADOR: ARNALDO DE ALBUQUERQUE ARAÚJO

Belo Horizonte

2013

Gomide, João Victor Boechat, 1960-
Efeitos visuais da trucagem óptica: à captura digital de movimento /
João Victor Boechat Gomide. - 2013
212 f: il.

Orientador: Arnaldo de Albuquerque Araújo

Tese (doutorado) - Universidade Federal de Minas Gerais, Escola de
Belas Artes, 2011

1. Animação de personagens – Técnica – Teses 2. Cinematografia –
Efeitos especiais – Teses 3. Animação por computador – Teses 4.
Processamento digital de imagens – Técnica – Teses I. Araújo, Arnaldo
de Albuquerque, - III. Universidade Federal de Minas Gerais. Escola de
Belas Artes IV. Título.

CDD: 778.5347



UNIVERSIDADE FEDERAL DE MINAS GERAIS
ESCOLA DE BELAS ARTES
PROGRAMA DE PÓS-GRADUAÇÃO EM ARTES

Assinatura da Banca Examinadora na Defesa de tese do aluno **JOÃO VICTOR BOECHAT GOMIDE** Número de Registro **2006238068**.

Título:

Efeitos Visuais, da Trucagem Óptica á Captura Digital de Movimento

Prof. Dr. Arnaldo Albuquerque Araújo – Orientador – ICEx/UFMG

Prof. Dr. Milton Terumitsu Sogabe – titular – IA/UNESP

Prof. Dr. Cláudio Márcio Magalhães – titular – UNA/MG

Prof. Dr. Luiz Antonio Cruz Souza – titular- EBA/UFMG

Profa. Dra. Lucia Gouvêa Pimentel – titular- EBA/UFMG

Belo Horizonte, 12 de agosto de 2011

Aos meus filhos, Pedro e Paola, e à Cristina e Pedro.

Agradecimentos

Ao Prof. Arnaldo de Albuquerque Araújo, por todo o apoio, orientação e por disponibilizar a infraestrutura do NPDI para a realização desta tese. Ao Prof. Luiz Roberto Pinto Nazario, pela contribuição no início deste trabalho.

À direção da FACE/FUMEC, pela construção do Laboratório de Captura de Movimento e por todo o apoio logístico e financeiro em todos os momentos importantes. Ao Prof. Emerson Eustáquio Costa, pela amizade e por disponibilizar horários de aula mais amigáveis durante o desenvolvimento deste trabalho. Ao Prof. Júlio Pessoa, pelos mesmos motivos.

Aos alunos e colaboradores, que foram fundamentais para a conclusão de todas as etapas deste empreendimento. Ao David Lunardi Flam, por ter escrito o código do OpenMoCap e por tantas sugestões úteis. Ao Daniel Pacheco de Queiroz, pelo estudo dos filtros para o rastreamento. À Carolina Bigonha, pela paciência e dedicação no início deste trabalho. Aos alunos Samir Rage, Henrique Siqueira e Nathan Santos pela contribuição valiosa na modelagem com o Blender e nas discussões sobre o GDD do jogo e o programa Dango Balango. À Janine Brioude, por ser a nossa atriz para a captura de movimento. Ao João César de Lima, pelo *design* do site e pelos *concepts* do curta-metragem. Ao Glauber da Silva, por ter surgido recentemente, como expert do Blender.

À Cristina e aos meus filhos, pela paciência por esta vida maluca que tenho levado, cheia de compromissos inadiáveis. Isto está acabando, prometo. E por todo o amor, que é recíproco. Dias mais cheios de tempo e com o modo *slow life* estão chegando!

Ao CNPq, FAPEMIG, FUNARTE e Universidade FUMEC, pelo
suporte financeiro.

Resumo

Esta tese descreve o desenvolvimento de um sistema de captura de movimento em tempo real e de código aberto – OpenMoCap – e suas aplicações e seu impacto no audiovisual, dentro da produção de efeitos visuais e da imagem digital. As várias etapas para produzir, do início, dados de captura de movimento para animar personagens para o cinema, a televisão e os jogos digitais são apresentadas. O aplicativo realiza todo o fluxo de trabalho para a aquisição dos dados de movimento e a sua saída em formatos apropriados para os softwares de modelagem e animação. A expertise para se trabalhar com a captura de movimento e os efeitos visuais é discutida, passo a passo, com exemplos de produções apresentadas em rede nacional de televisão aberta. Um portal da web foi construído, centrado na animação de personagens e na realidade aumentada com a captura de movimento, para difundir o seu uso e para atrair contribuições de desenvolvedores para melhorar o seu código-fonte.

A captura de movimento é uma técnica valiosa para a animação de personagens nas imagens em movimento, nos jogos e na realidade aumentada. Infelizmente, a maioria dos sistemas de captura de movimento é muito caro e de propriedade de empresas. Diante desta realidade, uma aplicação de código aberto para captura óptica de movimento foi desenvolvida, baseada em técnicas de análise de imagens digitais. A arquitetura foi desenhada para análise do movimento de marcadores em tempo real, sendo modular e flexível. Ela permite a adição de novos módulos otimizados para funções específicas, como o uso de diferentes conjuntos de câmeras e a captura sem marcadores, utilizando os já existentes.

Três aplicações de animação de personagens com a captura de movimento estão sendo desenvolvidas no presente momento, para demonstrar a viabilidade do aplicativo.

Palavras-chave: efeitos visuais; imagem digital; captura de movimento; animação de personagens; artes visuais.

Abstract

This thesis describes the development of an open source real time motion capture system - the *OpenMoCap* - and its applications and impact to audiovisual, inside visual effects and digital image production. The various steps to produce, from scratch, motion capture data to animate characters for cinema, television and games are related. The software carries out the entire pipeline for acquisition of motion data and its output in appropriate formats to modeling and animation software. The expertise to work with motion capture and visual effects is presented, step by step, with examples from mainstream production. A web portal has been constructed, centered in character animation and augmented reality with motion capture, to disseminate its use and to attract contributions from developers to improve its source code.

Motion capture is a valuable technique for character animation in motion pictures, games and augmented reality. Unfortunately, most of the motion capture systems currently available are expensive and proprietary. In face of this reality, an open source application for optical motion capture was developed based on digital image analysis techniques. The defined architecture is designed for real time motion analysis of markers, being flexible and modular. It allows the addition of new optimized modules for specific functions, as different cameras sets and marker-less motion capture, taking advantage of the existing ones.

Three applications in character animation with motion capture and *OpenMoCap* are being done at the present moment to show the feasibility of the software to different types of uses. The first one is a short movie with the theatrical company Galpão, the second is for character animation in Dango Balango, a telecast from Rede Minas de Televisão, and the latter is a digital game.

Keywords: visual effects; digital image; motion capture; character animation; visual arts.

Lista de Imagens

Imagem 1: representação esquemática da definição mais geral da computação gráfica.	14
Imagem 2: representação das áreas da computação gráfica, segundo [Gomes e Velho 2003].	14
Imagem 3: representação dos dados variando com o tempo e a imagem em movimento na computação gráfica, segundo [Gomes e Velho 2003].	15
Imagem 4: representação do paradigma da abstração, segundo [Gomes e Velho 2003].	16
Imagem 5: representação de uma onda simples.	18
Imagem 6: representação de uma onda eletromagnética.	18
Imagem 7: o espectro eletromagnético, com destaque para a faixa do espectro visível.	19
Imagem 8: espectro visível, do limite do ultravioleta (UV) ao limite do infravermelho (IR).	20
Imagem 9: as cores e seus respectivos comprimentos de onda e frequências correspondentes.	20
Imagem 10: modelos de reflexão da luz por um objeto.	21
Imagem 11: cubo de cores com os seus eixos.	23
Imagem 12: representação das cores com o tom de cor, a luminância e a saturação.	23
Imagem 13: sistema de representação YUV, que será apresentado no capítulo 3.	23
Imagem 14: as cores primárias: vermelho, verde e azul (RGB).	24
Imagem 15: as cores secundárias, ciano, amarelo e magenta (CMY).	25
Imagem 16: fotografia de um pássaro colorido e a decomposição dela no vermelho, verde e azul, e, à direita, nos tons de cinza, da luminância no vermelho.	25
Imagem 17: diagrama esquemático do olho.	26
Imagem 18: os bastonetes e os cones na retina da salamandra tigre, feita com microscópio eletrônico em http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/V/Vision.html .	27

Imagem 19: Regiões de sensibilidade de cada um dos cones e dos bastonetes.	28
Imagem 20: distribuição dos cones e dos bastonetes, em milhares por milímetro quadrado, com relação à distância angular ao centro da fóvea.	29
Imagem 21: corte do olho humano, mostrando o seu interior, com o cristalino e a fóvea em destaque.	30
Imagem 22: diagrama de uma projeção central [Azevedo 2003].	32
Imagem 23: diagrama com os tipos de projeções [Azevedo 2003].	33
Imagem 24: projeção paralela ortográfica [Azevedo 2003].	34
Imagem 25: uma câmera pinhole registrando uma fotografia.	34
Imagem 26: exemplos de ilusão de óptica, com os pontos de fuga e a linha do horizonte.	36
Imagem 27: aparência da cor vermelha com diferentes fundos coloridos.	37
Imagem 28: aparência de um mesmo tom de cinza com diferentes fundos.	38
Imagem 29: tabuleiro com as casas A e B marcadas, um cilindro vermelho e sua sombra.	38
Imagem 30: tabuleiro com as casas A e B, uma régua cinza, um cilindro vermelho e sua sombra.	39
Imagem 31: pincéis com as cores secundárias e o resultado das misturas de cores.	41
Imagem 32: À esquerda, as imagens sendo formadas no olho e em uma câmara fotográfica; à direita, a posição do sensor mostrado através de um corte na câmara.	42
Imagem 33: uma filmadora Arri e um pedaço de película.	42
Imagem 34: corte esquemático de uma ampliadora de fotografia.	43
Imagem 35: fotografias colorida e preto e branco, à esquerda, e os seus negativos à direita.	44
Imagem 36: corte esquemático de um tubo de televisão, mostrando o detalhe do tubo com fósforo, que colore a imagem ao ser atingido por elétrons.	45

Imagem 37: primeira imagem digital, de 1 bit de cor por pixel, de 1957.	47
Imagem 38: primeira forma de interação com a imagem, o <i>Sketchpad</i> . (http://resumbrae.com/ub/dms423_f08/06/sketchpad.jpg).	48
Imagem 39: separação da luz branca nas cores primárias por um prisma.	52
Imagem 40: diagrama esquemático mostrando a separação da luz pelo prisma após passar pela lente.	53
Imagem 41: prisma com os três CCDs acoplados.	54
Imagem 42: diagrama com uma malha de Bayer, à esquerda, e a distribuição de filtros à direita.	54
Imagem 43: digitalização de um sinal contínuo.	56
Imagem 44: Faixas de percepção dos objetos pelo olho e por dois tipos de microscópio.	57
Imagem 45: imagem com os pixels aparentes.	58
Imagem 46: imagem rasterizada depois de ser ampliada, revelando a estrutura de pixels.	58
Imagem 47: estrutura dos pixels de dois monitores LCD, com as cores primárias.	59
Imagem 48: componentes ciano, magenta e amarelo de uma fotografia.	59
Imagem 49: três desenhos construídos por retas curtas, as polilinhas, que dão a impressão de curvas.	61
Imagem 50: detalhe de uma imagem vetorial exibindo as polilinhas.	61
Imagem 51: figuras geradas por retas com preenchimento de cor.	62
Imagem 52: dispositivos de exibição vetoriais: plotter e monitor de um dos primeiros jogos.	62
Imagem 53: frame entrelaçado, com os <i>fields</i> visíveis à esquerda e com movimentos lentos à direita, e os campos menos visíveis para o quadro com menor movimento [Reinhardt 2007].	64
Imagem 54: esquema de pulldown [Lee 2011].	66
Imagem 55: comparações entre os diversos formatos de video.	70
Imagem 56: variações do codec MPEG e bitrates.	70

Imagem 57: representação esquemática dos frames I, P e B da compressão MPEG.	71
Imagem 58: diversas resoluções de televisão, SDTV e HDTV. 1920x1080 pixels é a <i>full</i> HDTV.	73
Imagem 59: cenários de imagens em movimento para as diversas mídias.	78
Imagem 60: a estrutura da pós-produção.	81
Imagem 61: mesa horizontal de montagem Steenbeck.	85
Imagem 62: interface do Adobe Premiere, com destaque para as diversas ferramentas.	86
Imagem 63: interface do Final Cut, com destaque para o waveform e o vectorscope.	86
Imagem 64: uma mesma tomada feita com diferentes lentes e o efeito provocado na imagem [Rickitt 2007].	90
Imagem 65: o efeito trombone, com a câmera se movendo no trilho e alterando a sua distância focal; nos fotogramas na parte inferior se observa a alteração do fundo em relação ao objeto [Rickitt 2007].	91
Imagem 66: representação esquemática do processo Williams [Rickitt 2007].	92
Imagem 67: preparação de cena com fundo azul em <i>Jornada nas Estrelas</i> [Rickitt 2007].	93
Imagem 68: Impressora óptica Oxberry 1600, vencedora do Oscar de 1986.	94
Imagem 69: cenas construídas utilizando o canal alfa fornecido pelo <i>chroma key</i> [Rede Globo 2003].	96-97
Imagem 70: cena construída utilizando o canal alfa fornecida por máscaras para a substituição do céu e colorização da paisagem [Rede Globo 2004].	98
Imagem 71: sequência final da novela Chocolate com Pimenta, com uso intensivo do canal alfa e do processamento de imagem [Rede Globo 2003].	99
Imagem 72: composição de imagens com canal alfa.	100
Imagem 73: transformação morfológica entre dois atores [Rickitt 2007].	104
Imagem 74: sistema para movimentação de câmera, controlado por	105

computador [Rickitt 2007].	
Imagem 75: cena diurna do Teatro Municipal [Cinemateca Brasileira 1918].	106
Imagem 76: cena modificada, com luzes aplicadas em máscaras digitais, correção de cor, rastreamento de movimento e céu noturno [Rede Globo 2004].	106
Imagem 77: cenas do Teatro Cólón, em Buenos Aires, antes e depois de modificadas, com figurantes adicionados às frisas e à platéia [Rede Globo 2003].	108
Imagem 78: <i>warping</i> aplicado à pupila, antes e depois do efeito aplicado [Rede Globo 2004].	108
Imagem 79: frames da cena com composição digital, rastreamento e chroma key [Rede Globo 2002].	109
Imagem 80: cronofotografia, à direita, de um homem com uma roupa preta e marcas brancas, à esquerda [Marey 2011].	118
Imagem 81: fotografias do galope de um cavalo feitas por Muybridge [Muybridge 1984].	119
Imagem 82: patente do processo da rotopscopia [Flescher 2005].	120
Imagem 83: <i>Sexy Robot</i> , primeiro personagem animado por captura de movimento [Abel 2011].	121
Imagem 84: cena do clipe <i>Steam</i> , no qual Gabriel Líquido contracena com duas mulheres de fogo, animados com mocap [Menache 2000].	123
Imagem 85: seqüência dos esqueletos, que poderia ter sido a primeira cena do cinema com captura digital de movimento [Menache 2000].	124
Imagem 86: robô T-1000, primeiro personagem animado por captura de movimento da história do cinema [Menache 2000].	124

Imagem 87: cena do jogo FX Fighter, em sua primeira versão.	125
Imagem 88: personagens de <i>Donkey Kong Country</i> [Nintendo 2011].	126
Imagem 89: diagrama dos diversos componentes de um sistema de mocap [Moeslund 2000].	128
Imagem 90: luva para mocap [Cybergloves 2011].	130
Imagem 91: sistema inercial IGS-190H [Animazoo 2011].	131
Imagem 92: dispositivo magnético Flock of Birds montado para captura de movimento [Ascension 2011].	133
Imagem 93: sistema óptico Impulse [PhaseSpace Tech 2011].	134
Imagem 94: sistema de captura com os LEDs acessos em torno da lente de entrada da câmera, retroiluminada [Vicon 2011].	135
Imagem 95: posições para 41 marcadores [Menache 2000].	141
Imagem 96: 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].	142
Imagem 97: um tipo de hierarquia para o esqueleto [Menache 2000].	143
Imagem 98: fluxo de trabalho do OpenMoCap [Flam 2009] e [Gomide et al 2010].	147
Imagem 99: Diagrama da Arquitetura do OpenMoCap [Flam 2009].	149
Imagem 100: Interface Gráfica Principal do OpenMoCap [Flam 2009].	150
Imagem 101: diagrama da estrutura do portal.	155
Imagem 102: homepage do portal.	162
Imagem 103: cenas do filme Viagem à Lua.	164
Imagem 104: cena original do filme e primeira versão modelada do jogo.	165

Imagem 105: personagem Barbenfoullis, modelado com o Blender.	166
Imagem 106: Selenita, modelado com o Blender.	167
Imagem 107: escritor com o cenário principal ao fundo.	180
Imagem 108: frame da animação com captura de movimento realizada para o programa <i>Dango Balango</i> .	181

Sumário

i	Agradecimentos
iii	Resumo
iv	Abstracts
v	Lista de Imagens
1	Introdução
1	Considerações preliminares
4	Motivação
10	Objetivos
11	Organização do texto
12	Apoios financeiros ao projeto
13	Capítulo 1 – Luz e Imagem
13	1.1. Imagem e computação gráfica
17	1.2. Luz e cores
26	1.3. Visão
31	1.4. Projeções de câmeras. Câmera <i>pinhole</i> .
35	1.5. Ilusões de óptica
40	Capítulo 2 – Geração da Imagem
40	2.1. Imagem pré-digital
46	2.2. Origens da imagem digital
49	2.3. Imagens em movimento
51	2.4. Etapas para a formação da imagem digital
52	2.4.1. Pré-processamento
55	2.4.2. Amostragem e quantização

56	2.5. Imagem raster e imagem vetorial
57	2.5.1. Imagem raster e acuidade visual
60	2.5.2. Imagem vetorial
63	2.6. Resolução espacial, profundidade de cor e fator de forma
64	2.7. Cadência das imagens
66	2.8. Pulldown
67	2.9. Codificação e decodificação da imagem (<i>Codec</i>)
72	2.10. Formatos de imagens digitais em movimento
75	Capítulo 3 – Pós-produção Digital
75	3.1. Pós-produção no audiovisual
77	3.2. As ferramentas para a pós-produção
80	3.3. Tipos de profissionais de pós-produção e mercado de trabalho
81	3.4. Edição e montagem
84	3.5. Ilhas de edição não-lineares
88	Capítulo 4 – Efeitos Visuais
88	4.1. Introdução
89	4.2. Evolução das trucagens com imagem
95	4.3. O canal alfa
100	4.4. Técnica e arte dos efeitos visuais digitais
105	4.5. Exemplos de produções
109	4.6. Perspectivas para a pesquisa e para a produção audiovisual
111	Capítulo 5 – Captura de Movimento e Animação de Personagens
111	5.1. Introdução
113	5.2. Animação de personagens
116	5.3. Captura digital de movimento

127	5.4. Tecnologias de Captura de Movimento
127	5.4.1. Introdução
129	5.4.2. Sistema mecânico
131	5.4.3. Sistema magnético
133	5.4.4. Sistema óptico
136	5.5. Etapas da Captura de Movimento
136	5.5.1. Preparação da sessão de captura de movimentos
139	5.5.2. Marcadores
143	5.5.3. Preparação do personagem para mocap
145	Capítulo 6 – O Software de Código Livre OpenMoCap
145	6.1. Introdução
145	6.2. Metodologia
147	6.3. Visão estereoscópica e triangulação
148	6.4. Arquitetura da Aplicação OpenMoCap
152	6.5. Atualizações mais recentes
153	Capítulo 7 – Portal da Web Openmocap.org
153	7.1. Objetivos
153	7.2. Público-alvo
155	7.3. Sistema de Organização do Site
156	7.4. Sistema de Rotulação e Navegação do Portal
156	7.4.1. Partes comuns a todas as páginas do portal
158	7.4.2. Páginas principais
160	7.4.3. Páginas acessórias
163	Capítulo 8 – Produções audiovisuais com a captura de movimento
163	8.1. Introdução

163	8.2. Jogo inspirado no filme <i>Viagem à Lua</i>
178	8.3. Curta metragem <i>A Primeira Perda da Minha Vida</i>
180	8.4. Programa infantil <i>Dango Balango</i>
182	Conclusões
182	Objetivos alcançados
183	Perspectivas
185	Lista de resultados decorrentes desta dissertação
187	Referências bibliográficas

INTRODUÇÃO

Considerações preliminares

A pesquisa realizada teve o objetivo de desenvolver um sistema de captura de movimento e de abordar de maneira sistemática os efeitos visuais no audiovisual. Ao final do trabalho foi concluído o protótipo para a captura de movimento para a animação de personagens, em tempo real e de código aberto, o *OpenMoCap*, assim como o portal da web para permitir o download do software. É o primeiro sistema de captura de movimento de código aberto desenvolvido no mundo e é o primeiro sistema aplicado à animação de personagens desenvolvido no Brasil. O código-fonte do software será disponibilizado em breve em um portal da web, que está em fase de finalização e praticamente todo em funcionamento. Esse portal terá o público-alvo e o modelo de funcionamento semelhante ao portal do software livre de modelagem [Blender 2011]. O domínio já está registrado em <http://www.openmocap.org>.

Esta tese tem o perfil transdisciplinar e envolve conhecimentos de diversas áreas, como a ciência da computação, a comunicação - através da produção de programas de televisão e produtos audiovisuais -, a física, o design e as artes visuais.

A captura de movimento (*mocap*) é amplamente usada em diversas aplicações importantes em todo o mundo, seja em animação no audiovisual e em jogos eletrônicos, em efeitos especiais para cinema e televisão, na linguagem de sinais, no reconhecimento gestual, na biomecânica ou na medicina de reabilitação. No Brasil, bancos de dados de captura de movimento são raramente utilizados em qualquer tipo de aplicação, tanto com a geração de dados em tempo real como pós-processada. Esse cenário advém do fato que os equipamentos de *mocap* são muito caros, incluindo taxas de importação e impostos, e a sua utilização tem complexidades intrínsecas.

No Brasil existem apenas duas empresas que possuem equipamentos de captura de movimento para animação de personagens, ambos importados. Um dos equipamentos é para aplicação para televisão, na TV Globo do Rio de

Janeiro, e o outro para aplicações no audiovisual, na RPM Produtora Digital, de São Paulo. A RPM se autodenomina em seu site “o único birô na América Latina” a fornecer dados de captura de movimento para empresas, nos mais diversos tipos de animações, dos jogos digitais à web e publicidade na televisão [RPM 2011].

Não existe, até o momento, qualquer equipamento desenvolvido no nosso país com a finalidade de animar personagens para o audiovisual ou os jogos digitais, em tempo real, além do apresentado nessa proposta. Os bancos de dados de movimento, quando utilizados, são alugados, comprados ou conseguidos gratuitamente de empresas fornecedoras estrangeiras. Jogos muito populares e dos mais diversos tipos, como *Fifa Soccer*, *NBA*, *SuperMario*, *Donkey Kong* e tantos outros, utilizam massivamente a captura de movimento para animar os personagens. Isso sem contar a participação cada vez maior de *mocap* no cinema e na televisão de países com intensa produção audiovisual. O filme *Avatar* é um exemplo da utilização madura dessa técnica.

No nosso país, somente o Hospital Sarah Kubitscheck e algumas universidades possuem equipamentos importados para aplicações biomédicas, como a Faculdade de Educação Física da Universidade Federal de Minas Gerais (UFMG). Na Universidade Estadual de Campinas (UNICAMP) foi desenvolvido um sistema pós-processado de captura de movimento para estudos de desempenho atlético, utilizado para estudos acadêmicos [Figuroa et al 2003].

Diante desse cenário, esse trabalho tornou possível a construção do *OpenMoCap*, que teve o apoio de algumas instituições de fomento¹. Como é discutido na tese, o sistema é modular e o software pode ser facilmente modificado para trabalhar com outras aplicações. Testes e melhorias do equipamento foram realizados recentemente.

O sistema utiliza um conjunto de câmeras de vídeo para capturar o movimento, com filtros passa-banda no infravermelho, leds no infravermelho para iluminar a cena e marcadores refletores nessa faixa do espectro. O software é autônomo, ou seja, ele inclui todos os componentes para processar

¹ Projetos aprovados e concluídos nos Editais Universais da FAPEMIG (projetos 11514/2006 e APQ-02239-10/2010 da FAPEMIG), do CNPq (processo 485513/2007-2 e 310965/2010-1), no programa PROPIC (protocolo 52/2008) e PROEXT (protocolo 38/2010) da Universidade FUMEC, e da FUNARTE, através da Bolsa de Produção Cultural para a Internet.

o sinal, e a sua arquitetura é flexível e extensível, pois é possível modificar e adicionar módulos específicos para adaptar o fluxo de trabalho às diversas aplicações. Dessa maneira, extensões que permitam fazer captura de movimento sem marcadores ou mudar o conjunto de câmeras podem ser facilmente implementadas. No presente momento, o *OpenMoCap* faz a captura óptica em todas as suas etapas com diversas câmeras e os dados são gerados no formato adequado para os softwares de modelagem 3D.

Existem duas definições distintas relacionadas à tecnologia da captura de movimento, que geram alguns problemas de taxonomia, como pode ser observado sempre que um novo produto de impacto é lançado, como o filme *Avatar* e *A Lenda de Beowulf*, e o jogo *Fifa Soccer*. Invariavelmente se discute se o trabalho foi feito com captura de movimento ou com captura de atuação, ou performance. Como é abordado adequadamente por [Menache 2000], “captura de movimento é a coleção de dados que representam o movimento, enquanto a animação de performance é o produto final de um personagem guiado por um ator”. Devido à extensão do trabalho apresentado e desenvolvido ao longo desta tese, as discussões mais detalhadas se limitarão à captura de movimento. A captura de performance está sendo realizada neste momento em três produtos audiovisuais que estão sendo desenvolvidos e que serão apresentados ao final do texto.

O texto da tese aborda, passo a passo, a construção do conhecimento necessário para se trabalhar com a captura de movimento, do ponto de vista do artista e do animador, procurando ampliar o uso do sistema. Ao realizar esse objetivo, discute-se também os efeitos visuais de uma maneira geral, desde o seu início, com a trucagem óptica. Desta maneira, a abordagem se inicia com o estudo da visão e da experiência do espectador, passa pela construção da imagem digital, em seus diferentes aspectos, dos efeitos visuais e termina com o estudo da captura de movimento e do *OpenMoCap*. Nesta última parte, aborda-se as diversas etapas para se obter e utilizar os dados do movimento, a arquitetura do software *OpenMoCap*, a estrutura do portal e os trabalhos de animação que estão sendo realizados.

Motivação

A tese é o resultado da busca por soluções para questionamentos surgidos durante o meu trabalho na pós-produção², em emissoras de televisão e produtoras de conteúdo audiovisual, e pela formação em física, especialmente em aplicações da óptica, com luz, espelhos e lentes. Ela seguiu uma metodologia rigorosa para apresentar respostas aos problemas apresentados, como será relatado nos capítulos da tese. Esta seção será na primeira pessoa, para que o tema seja apresentado mais claramente.

Ao longo de quinze anos de atuação em diversos tipos de produção audiovisual, de novelas e longa metragem a peças publicitárias de 30” e jogos, constatei a enorme carência na formação dos profissionais envolvidos na realização dos produtos, assim como na geração de conhecimento pelas soluções encontradas para as diversas demandas da pós-produção, nela incluídos os efeitos visuais. O estudo sistemático sobre esta arte e esta técnica no Brasil é praticamente inexistente. A Associação Brasileira de Cinematografia³ [ABC 2011] tem sido uma exceção neste contexto, trazendo e provocando o debate sobre o trabalho com a imagem em movimento, especialmente a digital, e tem cada vez mais agregado profissionais e estudiosos em seus quadros, que são referências na produção nacional. Esse cenário tem consequências marcantes na produção de cinema, televisão e jogos no país, tanto na qualidade da imagem distribuída quanto nas possibilidades para a narrativa.

Os efeitos visuais são recursos utilizados como apoio à narrativa no audiovisual, tanto na realização artística quanto na técnica. Neles se incluem as trucagens ópticas e digitais, as animações tradicionais ou criadas por computador, que podem interagir com imagens reais, e toda a parafernália mecânica e eletrônica utilizada adicionalmente à imagem capturada, para que

² A pós-produção é uma das etapas da realização de uma obra audiovisual. Ela começa efetivamente após a filmagem ou gravação das imagens e do som, e compreende a edição ou montagem, a composição digital, a sonorização, os efeitos visuais e a computação gráfica, a colorização e a finalização do audiovisual.

³ A Associação Brasileira de Cinematografia surgiu em 2001, com a intenção de agregar os profissionais de cinema e do audiovisual em uma associação e fomentar debates e cursos na área.

um roteiro transforme-se em um produto audiovisual, seja ele um filme, um programa de televisão, um jogo ou outro tipo de aplicação interativa.

O questionamento sobre o uso dos efeitos visuais passa pela formação cultural da equipe de produção, como os diretores, produtores e roteiristas, pelo tipo de narrativa utilizada, pela capacidade técnica e artística das pessoas envolvidas na realização dos mesmos, e pela experiência subjetiva do público, dentre tantos outros fatores. O estudo dos diversos tipos de efeitos visuais e da preparação para a realização dos mesmos, com alguns exemplos mais marcantes na história do audiovisual, traz um suporte teórico para um campo onde em geral se trabalha de maneira empírica e intuitiva, baseando-se na experiência prática e na criatividade. Isso eu pude perceber em diversas situações profissionais, como quando montei o longa-metragem *Aleijadinho*, em 1999, em uma ilha de edição não-linear. A direção e a produção do filme, que não tinham se atualizado sobre as técnicas digitais em uso, que no Brasil à época eram recentes, insistiam em fazer os letreiros com um cartelista⁴ e os efeitos visuais com trucagem óptica, o que implicava maior custo e tempo de produção, para um orçamento limitado e com o tempo de finalização comprometido.

Problemas semelhantes eu tive a oportunidade de vivenciar durante os cinco anos em que trabalhei na TV Globo⁵, no Rio de Janeiro. Este período, de 2000 a 2005, coincidiu com a estruturação da pós-produção digital e do Departamento de Efeitos Visuais da Central Globo de Produções. Em 2000, foi comprado o primeiro sistema para correção de cor da programação. Isto era uma exigência mercadológica, pois vários produtos da Rede Globo tinham sido recusados para compra em feiras internacionais, devido à qualidade insatisfatória da imagem apresentada. A correção de cor⁶, que será abordada nesta tese, era necessária para melhorar a marcação de luz e a imagem capturada por câmeras de vídeo profissionais. Houve, à época, uma grande resistência ao uso da técnica pelos diretores de fotografia de diversos programas, como novelas e mini-séries. Isto mesmo diante de incontáveis exemplos do seu uso, com sucesso, em produções internacionais, tanto no

⁴ Cartelista era o profissional responsável por fazer os letreiros de abertura e encerramento do filme. Em geral eram criadas cartelas em fundo preto, que eram filmadas e editadas para serem exibidas.

⁵ TV Globo é a maior rede de televisão do Brasil e a terceira maior rede comercial aberta do mundo.

⁶ A técnica de correção de cor será abordada no capítulo 5.

cinema como na televisão. Depois de ser quebrada essa resistência pela obrigatoriedade do uso da correção de cor na programação, exigidas pela direção artística e de engenharia da emissora, e pela demonstração de que a correção de cor efetivamente melhorava a qualidade final do produto, a Rede Globo atualmente conta com dois equipamentos de correção de cor que funcionam vinte e quatro horas por dia. A pouca utilização de ajuste de cores é percebida nas outras emissoras de televisão e na produção de cinema nacional, salvo raras exceções, como o filme *Cidade de Deus*, dirigido por Fernando Meirelles e Katia Lund.

Dentre muitos outros exemplos, em quase todos os lugares em que desenvolvi trabalhos, quando cheguei à TV Globo me deparei com uma situação inusitada. Um programa infantil, o *Bambulúá*, tinha uma grande quantidade de cenas gravadas com fundo verde ou azul - *chroma-key*⁷, que será discutido ao longo da tese – que apresentavam sérios problemas para serem recortados, isto é, para que o fundo com cor fosse substituído por algum outro tipo de fundo com imagens (*background*). Insisti que a imagem que chegava para ser trabalhada na pós-produção era de sinal composto⁸. Isto foi fortemente rejeitado por toda a equipe, que argumentava que as câmeras eram do tipo betacam digital da melhor qualidade e o recorte era testado durante a gravação. Esse mesmo problema se repetiu quando o especial do final do ano de 2000 do Roberto Carlos foi feito. Todo o programa foi gravado em um fundo verde, para ser substituído por outras cenas de background depois. Para dar conta de todo o trabalho, já que todas as cenas apresentavam problemas de recorte, o equipamento Inferno⁹, que é o mais sofisticado e caro existente, foi utilizado vinte e quatro horas por dia durante um mês para preparar o programa.

Quando fui chamado para fazer a pré-produção do programa *Sítio do Picapau Amarelo*, que retornaria à programação em 2001 e usaria o mesmo equipamento de *Bambulúá* e do especial de Roberto Carlos, pedi para ir ao

⁷ A técnica do *chroma-key* também será abordada no capítulo 5.

⁸ O vídeo composto é o processo que mais sofre deterioração na transmissão do sinal. Nele se juntam todos os três componentes da imagem em um único condutor. Isto será mais detalhado no capítulo 4.

⁹ Inferno é um equipamento de composição da Discreet Logic, que era executado em computadores da Silicon Graphics Incorporation (SGI), como uma solução completa e fechada de hardware e software. Atualmente a Discreet foi adquirida pela Autodesk Incorporation e o Inferno é executado no sistema Linux, após a falência da SGI.

estúdio para verificar como a imagem era gravada. Nessa época, a TV Globo alugava os estúdios do Renato Aragão, que fica a alguns quilômetros do Projac¹⁰. Atualmente esses estúdios foram comprados pela TV Record. Eram três estúdios que, por falta de espaço, utilizavam três caminhões onde ficava todo o equipamento para gravação da imagem e do som. Os sinais das quatro câmeras em cada um dos três estúdios iam para cada um desses três caminhões por cabo e, lá chegando, eram pré-editados em uma mesa de corte. Aproveitava-se para se fazer o recorte do *chroma-key* para teste, já que estas mesas possuíam este recurso. O resultado ali era excelente, sem qualquer problema na borda do recorte. Depois de passar pela mesa, a cena era gravada em um sistema betacam digital. As quatro câmeras de cada estúdio eram do tipo betacam digital da Sony, de altíssima qualidade e custando mais de 50 mil dólares cada uma. Descobri que, na passagem pela mesa de corte, o sinal era transformado de componente¹¹ em composto, o que tornava impossível fazer um bom recorte de cor, como será discutido nesta tese. Pedi então para que não se usasse a mesa para testar o *chroma-key*. O sinal da câmera tinha que ser gravado diretamente no equipamento de videotape. A resistência da equipe para fazer esse teste foi muito grande, em especial do produtor de engenharia. O teste foi então ordenado pelo diretor de núcleo Roberto Talma¹² e o problema foi finalmente identificado e solucionado. Esse problema gerou, durante alguns anos, uma imagem deteriorada e aumento significativo nos custos de produção para os programas que utilizavam os equipamentos dos estúdios do Renato Aragão. Este é um dos exemplos mais contundentes, mas existem dezenas deles, com os quais tive que lidar ao longo destes anos, ocorridos tanto na Rede Globo quanto em outras produtoras e emissoras de televisão, que ilustram bem o problema gerado pela falta de formação e de conhecimento das equipes na qualidade artística do produto final e nos custos de produção.

¹⁰ Projac é centro de produção da Central Globo de Produção (CGP) e fica em Jacarepaguá, no Rio de Janeiro. É a abreviatura de Projeto Jacarepaguá.

¹¹ No vídeo componente, todas as três componentes do sinal da imagem são trabalhadas separadamente. Isto será mais detalhado no capítulo 3.

¹² Roberto Talma é diretor de núcleo da Rede Globo. Começou a sua carreira na televisão como editor de imagens, e foi assumindo novas funções à medida que demonstrava seu talento e profissionalismo. Tem inúmeros exemplos de produções complexas, como a novela *Que Rei Sou Eu?* e o seriado *Sítio do Picapau Amarelo*.

Na pré-produção do *Sítio do Picapau Amarelo*, me deparei com uma outra situação, que gerou o sistema de captura de movimento que é apresentado nesta tese. O diretor Roberto Talma e o diretor de novas tecnologias e inovação da TV Globo, José Dias¹³, me escalaram para fazer estudos com a captura de movimento para o programa. Nessa época, eu trabalhava na captura de movimento dos personagens Dubem e Dumal do programa *Bambuluá*. O equipamento utilizado era um *Flock of Birds*, da [Ascension 2011], que utiliza campos magnéticos para gerar os dados de movimento. Para o *Sítio*, se pensava em fazer os personagens Emília e Visconde de Sabugosa com modelagem 3D e utilizar a captura de movimento para animá-los e integrá-los às cenas gravadas com os outros personagens, que eram reais. Ao final dos estudos de viabilidade, José Dias propôs enviar o projeto para que o Medialab o realizasse. O Medialab é um estúdio de Nova Iorque, que produziu as animações do Mario Bros. A argumentação, que era verdadeira, era de que o Brasil não possuía expertise para realizar esse tipo de trabalho.

Foi nessa época que tive a ideia de desenvolver um sistema de captura de movimento, que poderia gerar conhecimento e expertise na área. Esse projeto teve que esperar quase dez anos para ser finalizado. Mas o cenário da produção de animações com a captura de movimento praticamente não se alterou no Brasil nestes anos, como discutido na seção anterior.

Em 2005, retornei a Belo Horizonte, para trabalhar com pós-produção na produtora Alterosa Cinevideo, e decidi retornar às atividades acadêmicas, com o intuito de aprofundar os estudos com a pós-produção e desenvolver o sistema de captura de movimento, assim como de divulgar e ensinar o que eu havia aprendido nestes anos de mercado. Para adquirir um melhor conhecimento em captura de movimento, fiz o Mestrado na Escola de Belas Artes com esse tema. E me tornei orientando do prof. Arnaldo de Albuquerque Araújo no Doutorado, procurando aprofundar a busca de soluções para as questões levantadas nesta introdução. Este período do doutoramento foi muito profícuo, com a produção de trabalhos específicos sobre o processamento de

¹³ José Dias é um engenheiro e o inventor que introduziu o maior número de inovações tecnológicas na televisão brasileira, muitas delas inventadas por ele. Um exemplo é o tira-teima dos gols, no futebol, que permite reconstruir a cena em 3D em tempo real e fornecer todos os dados de distância e velocidade do lance.

imagem e a realização de efeitos visuais e o desenvolvimento do sistema de captura de movimento livre e de código aberto, o *OpenMoCap*. Criei a pós-graduação em Pós-Produção para Cinema, Televisão e Novas Mídias, que teve quatro turmas cheias no Centro Universitário UNA. Este curso teve alunos de Portugal, do Equador e de diversos outros estados brasileiros e foi o primeiro e único proposto com esse tema no Brasil até o momento. Atualmente, ele está sendo transferido para a Universidade FUMEC, já que as condições oferecidas pela UNA não são adequadas para os objetivos do curso.

O sistema de captura de movimento foi desenvolvido no Núcleo de Processamento Digital de Imagens (NPDI), coordenado pelo prof. Arnaldo de Albuquerque Araújo. O código foi escrito pelo nosso ex-orientando de mestrado, David Lunardi Flam, que continua colaborando com o seu desenvolvimento, agora através da Universidade FUMEC. O trabalho também contou com a colaboração de outro nosso ex-orientando de mestrado, Daniel Pacheco de Queiroz. Como será mostrado, atualmente o sistema realiza todo o processo para gerar os dados de movimento em formatos adequados para os softwares de modelagem e animação. Durante o desenvolvimento do aplicativo, se decidiu abrir o seu código, utilizando ferramentas de desenvolvimento que permitissem isso. Essa decisão foi tomada devido à carência de apoio financeiro e de pessoal para desenvolver uma solução comercial.

O software está sendo disponibilizado para download em um portal da web. Este portal é sobre animação de personagens e realidade aumentada com a captura de movimento. Ele terá uma seção especial para desenvolvedores, com o intuito de atrair colaboradores para o seu desenvolvimento. O modelo adotado é semelhante ao do software *Blender* de modelagem e animação 3D [Blender 2011]. Esse aplicativo foi totalmente desenvolvido de forma coletiva, a partir do momento em que ele foi disponibilizado em um portal da web para esse fim. Atualmente o Blender é o único software de modelagem e animação que faz frente ao monopólio da Autodesk Incorporation [Autodesk 2011], com os seus produtos Maya e 3DSMax.

Objetivos

A seguir são relacionados o objetivo geral e os objetivos específicos deste trabalho.

1. Objetivo Geral

O objetivo geral desta tese foi o de desenvolver um sistema de captura de movimento e aprofundar e disseminar o conhecimento do seu uso no contexto dos efeitos visuais. Inicialmente se pensou em produzir um sistema comercial, mas, diante do exposto na seção anterior e nos capítulos finais do texto, se decidiu utilizar ferramentas de código aberto para o seu desenvolvimento e disponibilizar o código-fonte em um portal da web. Esta foi a primeira vez no mundo que isto foi realizado para a captura de movimento e a primeira vez no Brasil que um aplicativo de mocap foi desenvolvido para a animação de personagens.

2. Objetivos Específicos

- Organizar e buscar recursos materiais e humanos para realizar o objetivo geral.
- Desenvolver a aplicação de código livre e arquitetura flexível e permitir que outros desenvolvedores colaborem com o amadurecimento do código-fonte.
- Disponibilizar o software para download em um portal da web para atrair a colaboração de desenvolvedores e para disseminar o uso da técnica para a animação de personagens.
- Disponibilizar as informações sobre a captura de movimento, dentro do contexto dos efeitos visuais, de maneira sistemática, para que contribua para a formação dos artistas e animadores envolvidos com a técnica dos efeitos visuais e da captura de movimento.
- Disseminar o debate sobre os efeitos visuais e a imagem digital no audiovisual, motivado pela captura de movimento.

Organização do texto

Na elaboração do texto desta tese, foi necessário demarcar os limites da abordagem que seria dada ao tema. A pós-produção, os efeitos visuais e a captura de movimento, com a utilização de recursos digitais, envolvem diversas áreas do conhecimento humano, como a matemática, a física, a engenharia, a computação, a comunicação e as artes. Com o foco na motivação que levou a este trabalho e por ser uma tese desenvolvida na Escola de Belas Artes da UFMG, decidiu-se que o texto seria dirigido para o artista e o animador que se utilizam da pós-produção e da captura de movimento. Desta maneira, se procurou apresentar as questões relacionadas às ciências exatas de uma forma sucinta, porém inteligível para o profissional que trabalha no audiovisual. Poderia ser apresentado o texto com todo o rigor matemático, mas isto desvirtuaria o propósito inicial do projeto. O conhecimento para se trabalhar e entender os temas da tese é construído ao longo do texto, tendo em vista essa abordagem. Alguns temas que são conexos a este trabalho, mas que não são fundamentais para ele, como o detalhamento da representação de cores e das chamadas *LookUp Tables*, não foram aprofundadas, para não se perder o foco no assunto principal.

O texto apresentado nesta tese resultou em um livro, que está sendo formatado para ser publicado pelas Edições Cine-Clube Avanca, de Portugal. Todas as ilustrações do texto que estão sem crédito são deste livro e foram feitas pela designer Marcela Scarpelli. O livro tem o objetivo de preencher uma lacuna, pois não existe qualquer obra sobre efeitos visuais em língua portuguesa.

A tese é composta de oito capítulos. No Capítulo 1, é abordado o processo de visão e de percepção das cores e da projeção da imagem, com as suas consequências à experiência subjetiva do espectador. No Capítulo 2 são discutidos os diversos processos de geração da imagem, com especial ênfase à imagem digital, seu processamento e formatos disponíveis. No Capítulo 3 é apresentada e discutida a pós-produção digital. Os efeitos visuais ópticos e digitais são discutidos no Capítulo 4. O Capítulo 5 é dedicado à animação de personagens e à captura digital de movimento. O Capítulo 6 discute os

diversos aspectos do software OpenMoCap. O Capítulo 7 apresenta o portal da web. No capítulo 8 são apresentados os três trabalhos de animação que estão sendo realizados com a captura de movimento. Toda a modelagem dos personagens e dos cenários está concluída e já foram realizados alguns testes preliminares. Na Conclusão, são discutidos os objetivos alcançados e as perspectivas do presente trabalho.

Apoios financeiros ao projeto

Para a realização deste projeto, busquei recursos financeiros junto às agências de fomento para viabilizar a compra de equipamentos e bolsas para os colaboradores. Estes recursos estão listados abaixo:

1. Edital Universal do CNPq, processo 483267/2012-0, de dezembro/2012 a dezembro/2014, para compra de equipamentos;
2. Fundo Estadual de Cultura, da Secretaria de Estado de Cultura de Minas Gerais, projeto 125/01/2010/FEC, para o desenvolvimento do curta metragem de animação, com início em março de 2012 e duração de um ano;
3. Dois projetos pelo ProPIC/FUMEC, para quatro bolsas de iniciação científica e gastos com equipamentos, em agosto de 2009 e agosto de 2012, com duração de um ano cada;
4. Edital Universal da FAPEMIG, processo APQ-02239-10, de dezembro/2010 a dezembro/2012, para bolsas e equipamentos;
5. Edital Universal do CNPq, processo 477168/2010-8, de dezembro/2010 a dezembro/2012, para compra de equipamentos;
6. Dois projetos pelo ProEXT/FUMEC, para quatro bolsas de iniciação científica e gastos com equipamentos, em agosto de 2010 e agosto de 2011, com duração de um ano cada;
- 7 Bolsa FUNARTE de Produção Cultural para a Internet, em 2010;
8. Edital Universal do CNPq, processo 485513/2007-2, para compra de equipamentos, de dezembro/2007 a novembro/2009;
9. Edital Universal da FAPEMIG, processo SHA 251/06, para bolsas e compra de equipamentos, de janeiro/2007 a dezembro/2009.

CAPÍTULO 1

LUZ E IMAGEM

1.1. Imagem e computação gráfica

Neste capítulo são discutidos os conceitos de luz, cor, projeção e visão, e como a imagem é formada e percebida. Eles nos levam às questões sobre como é representada a projeção do espaço tridimensional no fundo da retina, em perspectiva, e como são implementadas as combinações de cores e a projeção no universo dos bits. Esta é a base para tudo o que será abordado nos próximos capítulos, seja na imagem digital em movimento, nos efeitos visuais e, mais especificamente, na captura de movimento. Estes conceitos são fundamentais na representação do mundo pela computação gráfica e pela imagem gerada por uma câmera, assim como nos trabalhos de composição e colorização.

Para uma câmera real, a projeção da luz que vem do mundo físico é feita diretamente pela lente e a luz incide no sensor que gera a imagem. Na imagem digital, isto gera um arranjo matricial de pontos, com cor de um só valor definida para cada ponto. Essa imagem digital, que é do tipo raster, será discutida no próximo capítulo. Na computação gráfica, o processo é um pouco mais complexo. A representação do mundo real é feita por vetores, vértices, polígonos, cores e luz, em um modelo matemático. A imagem é obtida por uma matriz de projeção, que define uma câmera sintética no espaço de representação e de implementação computacional. O resultado dessa projeção é a saída do aplicativo, que é uma imagem em um dos formatos existentes e é gerada no processo conhecido como render. O render, ou renderização, nada mais é do que o processamento computacional de todos os cálculos matemáticos da modelagem da cena, que gerará a imagem. No render se escolhe a configuração do que vai ser levado em conta nesses cálculos matemáticos, tendo-se em conta o que deve ser representado na imagem. Pode-se, por exemplo, não se levar em consideração as sombras ou efeitos de refração dos objetos semitransparentes. Isso é uma questão de custo/benefício

para cada tipo de trabalho, pois o render é uma operação que requer muitos recursos sofisticados de hardware e de software para gerar a imagem.

De acordo com [Gomes e Velho 2003], a computação gráfica é o conjunto de técnicas e métodos que tratam da manipulação de dados ou imagens no computador, com os dados geométricos gerando a imagem e a imagem, por outro lado, gerando os dados geométricos, como no esquema da Imagem 1.

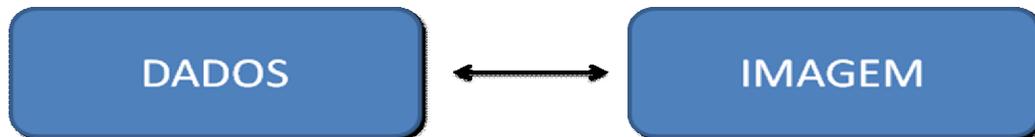


Imagem 1: representação esquemática da definição mais geral da computação gráfica.

Nesta definição, tanto o processamento digital de imagem quanto a visão computacional estão incluídos dentro da grande área da computação gráfica. Detalhando a representação da Imagem 1, é mostrado o funcionamento esquemático da computação gráfica na Imagem 2.

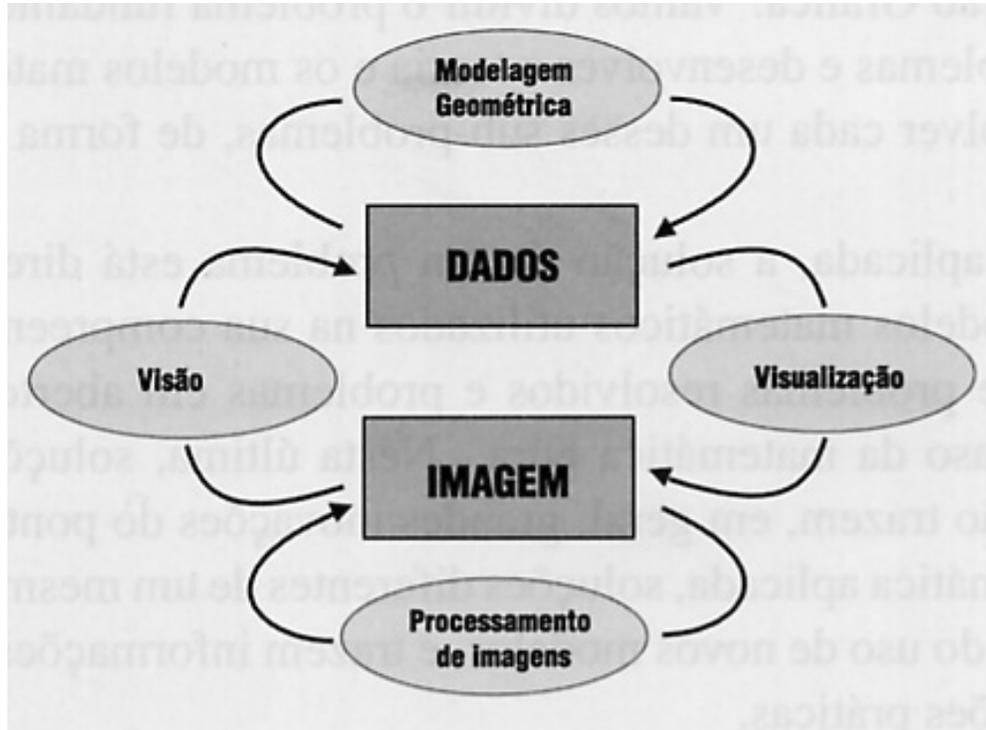


Imagem 2: representação das áreas da computação gráfica, segundo [Gomes e Velho 2003].

A descrição e estruturação de dados em dados geométricos no computador é o objetivo da modelagem geométrica. Nesta área se tem o exemplo do desenho assistido por computador, ou *computer-assisted design* (CAD). Na visualização, também conhecida como síntese de imagens, os dados gerados por um sistema de modelagem geométrica são transformados e o produto final é uma imagem que pode ser exibida. A construção da imagem é feita em geral através de uma câmera sintética, que está dentro do espaço onde se encontram os modelos geométricos. No processamento de imagens, o sistema admite como entrada uma imagem que, após processada, produz outra imagem de saída. Esta é a área da correção de cor e da composição de imagens. A visão computacional, também conhecida como análise de imagens, tem por finalidade obter, a partir de uma ou várias imagens, informações geométricas, topológicas ou físicas sobre os dados que as originaram.

Se a imagem em movimento é levada em consideração, com os dados variando com o tempo, se obtém a representação da Imagem 3.

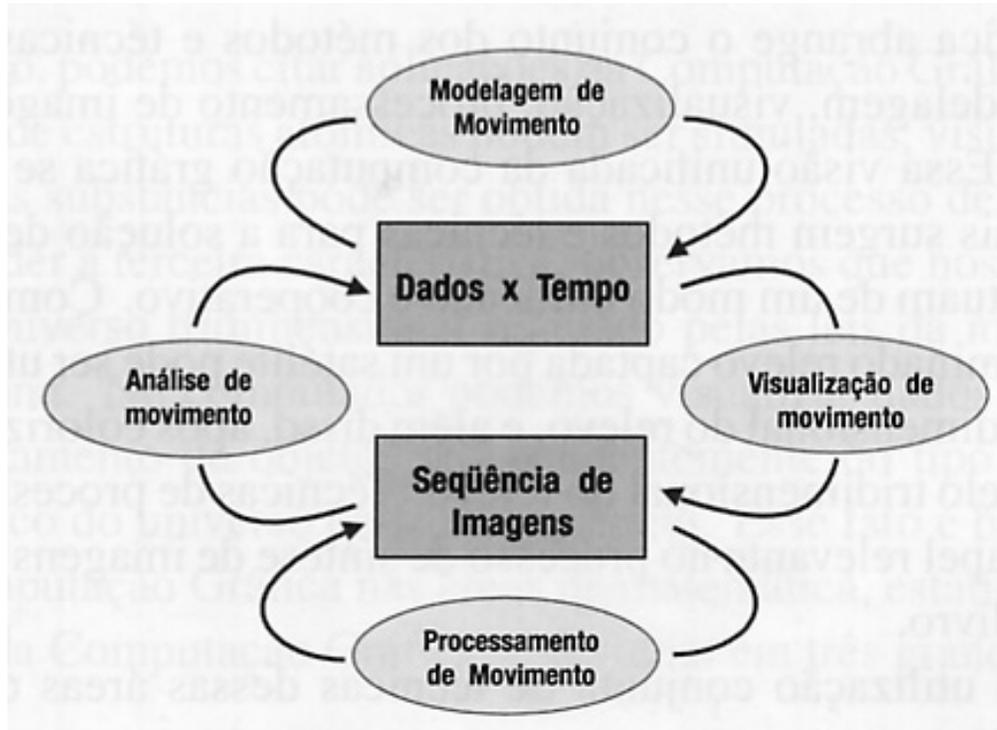


Imagem 3: representação dos dados variando com o tempo e a imagem em movimento na computação gráfica, segundo [Gomes e Velho 2003].

Neste caso, os dados variam com o tempo. Além de modelar a geometria e a topologia dos dados, é necessário descrever o movimento dos objetos em cena. A visualização de movimento é a animação, e o resultado é uma sequência de imagens, ou *frames*, conhecido como vídeo. O processamento de imagem é o processamento de uma sequência de frames, ou vídeo. E a análise de movimento envolve a análise de uma sequência de imagens, através do fluxo óptico ou outras técnicas de marcação. Essa é a área do desenvolvimento de um sistema de captura de movimento. Os dados gerados pela captura de movimento são retornados para que a modelagem adequada dos personagens possa realizar as animações planejadas. E o processamento de movimento pode auxiliar nos planos da produção do audiovisual, corrigindo cor e alterando o tempo das ações.

De acordo com o paradigma da abstração, apresentado por [Gomes e Velho 2003], a representação computacional de mundo real passa por quatro transformações, como na Imagem 4.

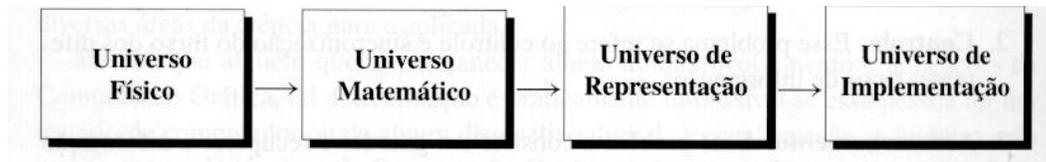


Imagem 4: representação do paradigma da abstração, segundo [Gomes e Velho 2003].

O universo físico contém os objetos do mundo real que se pretende modelar. O universo matemático contém a descrição abstrata desses objetos do mundo físico e das leis físicas que atuam mundo real. O universo da representação trabalha com as descrições simbólicas e finitas associadas a objetos do universo matemático, que é, por sua vez, descrição do universo físico. O universo de implementação associa as descrições do universo de representação às estruturas de dados, para se obter uma representação do objeto no computador. A forma como isto é feito, através de malhas de polígonos e da geometria diferencial, e modelos matemáticos de representação

da luz e da cor, permite que os objetos sejam definidos por uma lista de vértices da malha, valores de cor e de luz associadas a cada polígono da malha. A imagem gerada é uma transformação por uma matriz de projeção, que leva esses vértices no espaço tridimensional ao espaço bidimensional da imagem. Da mesma forma como funciona uma câmera real e o nosso olho, imersos no espaço real.

Por outro lado, o conhecimento sobre a combinação das cores, aditiva e subtrativamente, como será apresentada neste capítulo, auxilia de forma determinante o trabalho de correção de cor e de composição da imagem. Um profissional que sabe lidar com as misturas das cores pode, por exemplo, fazer o efeito da “noite americana”¹⁴ em um tempo significativamente menor do que outro profissional que faça o efeito de forma empírica.

Neste capítulo serão abordadas algumas definições sobre combinação de cores que estão presentes no cinema convencional e nas artes visuais. A definição das cores na imagem digital será discutida no próximo capítulo.

1.2. Luz e cores

Toda a experiência visual que se tem do mundo se dá através da luz que chega aos olhos e de como ela é percebida e interpretada. A luz visível é um tipo de onda que se propaga no espaço em uma velocidade que é a maior que qualquer coisa pode atingir no universo, pelo menos até o momento, segundo o nosso conhecimento, que é de 300 mil quilômetros por segundo. Como é uma onda, ela tem os seus picos e seus vales, como no gráfico da Imagem 5.

A distância entre dois picos ou dois vales da onda é chamada de comprimento de onda, e é representado pela letra grega lambda (λ). O tempo que demora passar dois picos ou dois vales da onda é chamado de período, e é representado pela letra grega tau (Γ). O inverso do período, isto é, $1/\Gamma$, é a frequência ou ciclo da luz, e é representada pela letra f e é medida em Hertz = Hz = 1/segundo.

¹⁴ Noite americana é a transformação de uma cena que foi gravada durante o dia em uma cena noturna, com todas as características de iluminação para luz noturna.

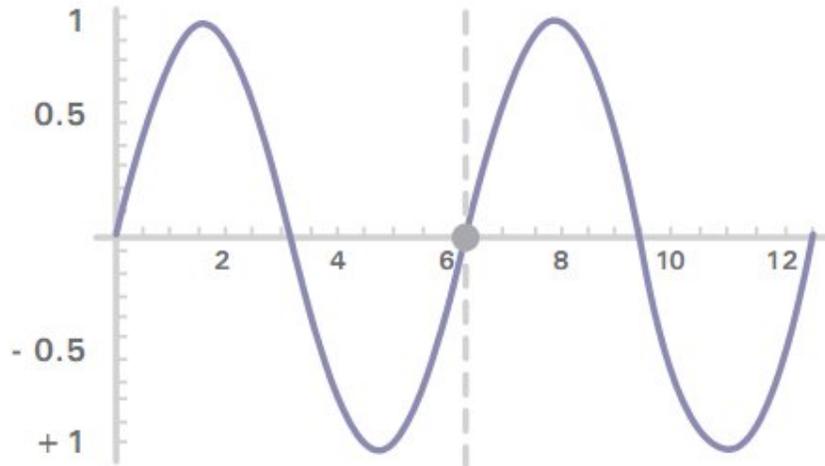


Imagem 5: representação de uma onda simples.

Cada cor tem o seu próprio comprimento de onda ou frequência da luz. A luz visível faz parte de um conjunto de ondas, que são conhecidas como ondas eletromagnéticas. Esse nome foi adotado porque elas são constituídas por ondas elétricas e magnéticas, que se alternam sincronamente, a variação de uma gerando a outra, e vice-versa. No gráfico da Imagem 6 é observada uma representação desse fenômeno e a indicação da seta mostra o comprimento de uma onda.

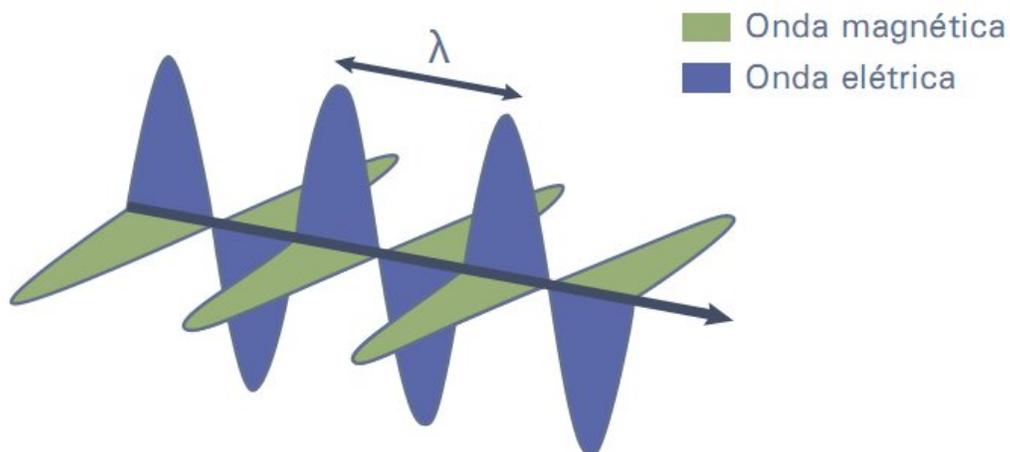


Imagem 6: representação de uma onda eletromagnética.

O conjunto de todas as ondas eletromagnéticas formam o espectro eletromagnético. Estas ondas são estudadas pelo eletromagnetismo e o seu comportamento e efeitos são muito bem conhecidos. A luz visível é uma parte muito pequena de todo o conjunto de ondas eletromagnéticas. E é essa pequena parte que é observada pelos olhos. O comprimento de onda é medido por um bilionésimo do metro, que é o nanômetro (*nm*). Na Imagem 7, é representado o espectro eletromagnético, com destaque para a faixa do espectro visível. Na parte superior da imagem, as várias faixas de ondas estão divididas de acordo com o seu comprimento de onda e é dado destaque ao espectro visível, que é expandido na parte inferior.

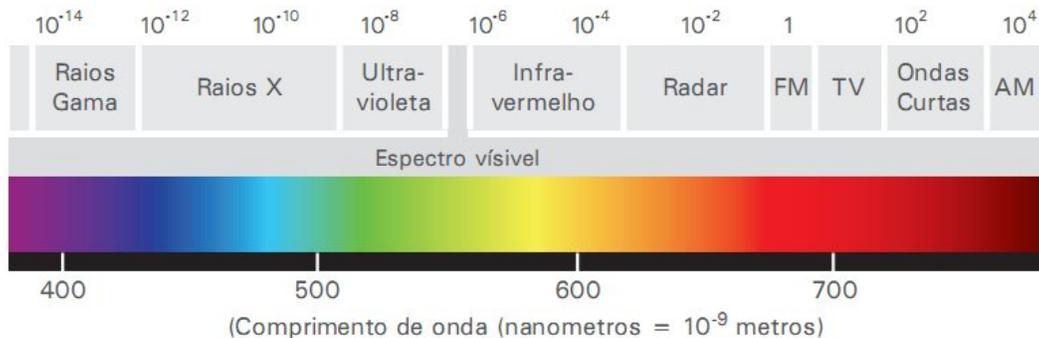


Imagem 7: o espectro eletromagnético, com destaque para a faixa do espectro visível.

Cada espécie de animal enxerga trechos diferentes do espectro eletromagnético, em torno da luz visível, de acordo com o seu aparelho de visão. As abelhas, por exemplo, conseguem perceber um pequeno pedaço do ultravioleta, além do visível, e algumas serpentes enxergam também uma pequena parte do infravermelho, vizinha ao visível. Os cachorros enxergam a luz visível em preto e branco.

Os seres humanos enxergam a luz visível com os seus diversos comprimentos de onda, que são as suas diferentes cores, se não tiverem problemas com os seus sensores de cor. Na Imagem 8, é detalhado o espectro visível, do limite do ultravioleta (UV) ao limite do infravermelho (IR) e se observam as diferentes cores de acordo com o seu comprimento de onda.

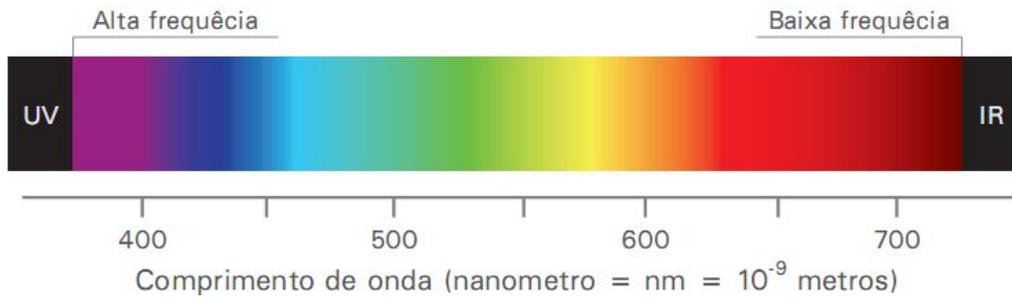


Imagem 8: espectro visível, do limite do ultravioleta (UV) ao limite do infravermelho (IR).

Na tabela da Imagem 9, se tem uma ideia dos valores envolvidos, com as cores e seus respectivos comprimentos de onda e frequências correspondentes.

COR	COMPRIMENTO DE ONDA (NM)	FREQUÊNCIA (10^{12} Hz)
VERMELHO	780 – 622	384 – 482
LARANJA	622 – 597	482 – 503
AMARELO	597 – 577	503 – 520
VERDE	577 – 492	520 – 610
AZUL	492 – 455	610 – 659
VIOLETA	455 - 390	659 – 769

Imagem 9: as cores e seus respectivos comprimentos de onda e frequências correspondentes.

Os objetos são vistos através da luz que enviam para os olhos. Essa luz define as características de cada objeto. A forma dos objetos, através de suas linhas de contorno, as suas superfícies e o seu volume, e a sua cor, é revelada pela luz que eles emitem, e são percebidas pelos diversos sensores do mundo real, ou são modelados computacionalmente.

A luz pode ser irradiada pelos objetos, pode ser refletida, pode ser transmitida, como no caso dos objetos transparentes ou semitransparentes, ou então pode penetrar nos objetos translúcidos, retornando de camadas mais profundas, sem atravessá-los, como no tomate, no leite e na pele. Nos

exemplos de objetos que irradiam se têm o sol, as estrelas, as lâmpadas, os projetores, os televisores e os monitores, que produzem luz por algum processo. A luz que está viajando no espaço se combina como uma soma de luzes, com as suas cores. Nesse caso, a cor é a soma das diversas cores que se combinam. Na televisão, por exemplo, cada ponto da tela emite luz nas cores vermelha, verde e azul. A cor resultante em cada ponto é a soma das intensidades de cada uma dessas cores primárias.

A luz também pode ser refletida pelos objetos, como um espelho ou de forma difusa, tal como pode ser observada na Imagem 10 (a) *reflexão difusa*; (b) *reflexão especular*.

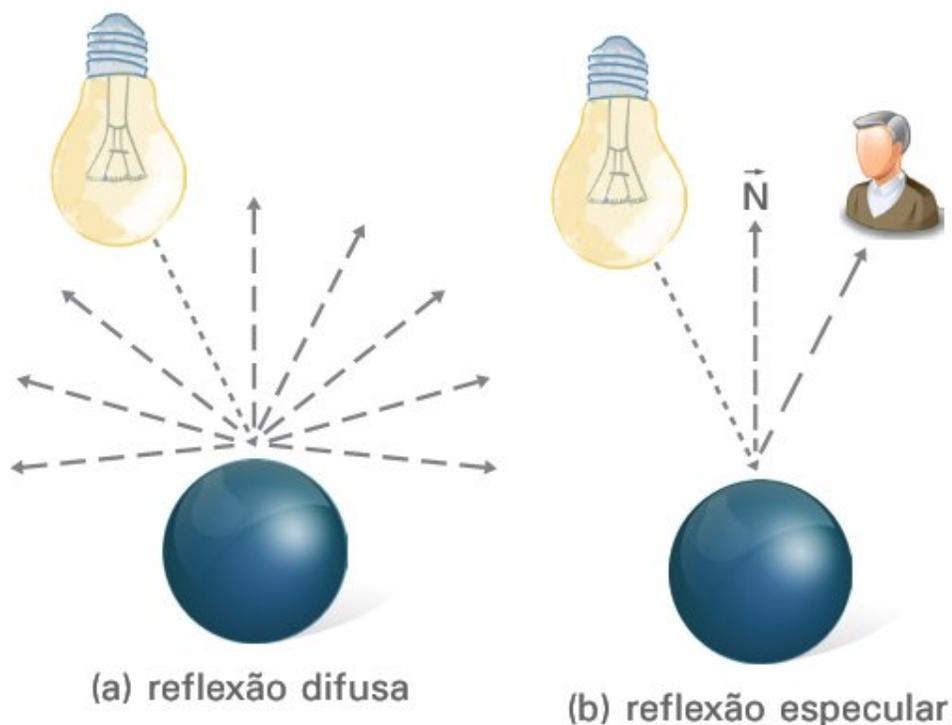


Imagem 10: modelos de reflexão da luz por um objeto.

A cor da superfície desses objetos que chega aos olhos é o resultado da luz que incide neles e que não é absorvida ou transmitida. A luz especular é responsável pelo brilho dos objetos e a difusa pela sua textura. Nesse caso, o objeto subtrai da luz incidente todas as cores que absorve. A luz refletida é o

que sobra da luz incidente, é aquela não é absorvida. A cor do objeto é a soma da luz refletida e da luz transmitida na mesma direção da refletida.

Existe a terceira classe de objetos, transparentes ou semi-transparentes, que desviam e alteram a luz, deixando-a passar por eles. E, por último, os objetos translúcidos, que não deixam passar a luz, mas alteram a luz que retorna de camadas mais profundas, abaixo da superfície refletora. O que enxergamos é o resultado da iluminação e de sua interação com a superfície do objeto. E isso é representado em uma imagem, seja ela gerada por computação gráfica, por uma máquina fotográfica ou desenhada e pintada. Este modelo é o mesmo adotado na implementação computacional e matemática dos programas de modelagem 3D e das bibliotecas gráficas, como a [OpenGL 2011] e a [DirectX 2011]. Este processo de descrição da luz e de sua interação com os objetos é conhecido como *shading*, que pode ser traduzido como tonalização ou sombreamento.

Todas as cores que são visíveis podem ser obtidas pela combinação de três cores escolhidas apropriadamente, gerando o espaço de cores [Feynman 2011]. Este espaço de cores é também tridimensional e cada uma das três componentes da cor forma cada um dos três eixos ortogonais. Ele pode ter diferentes tipos de componentes básicas. Na Imagem 11, é representado o cubo de cor, com as cores secundárias definindo o cubo, à esquerda, e as primárias, à direita. O espaço também pode ser definido com outros valores e sistema de coordenadas, como o matiz de cor, a luminância e a saturação, em coordenadas cilíndricas (Imagem 12), no chamado sistema HLS, ou com a luminância, as cores menos o vermelho e as cores menos o azul, na Imagem 13, no sistema chamado como YUV. Estes conceitos serão estudados ao longo do texto.

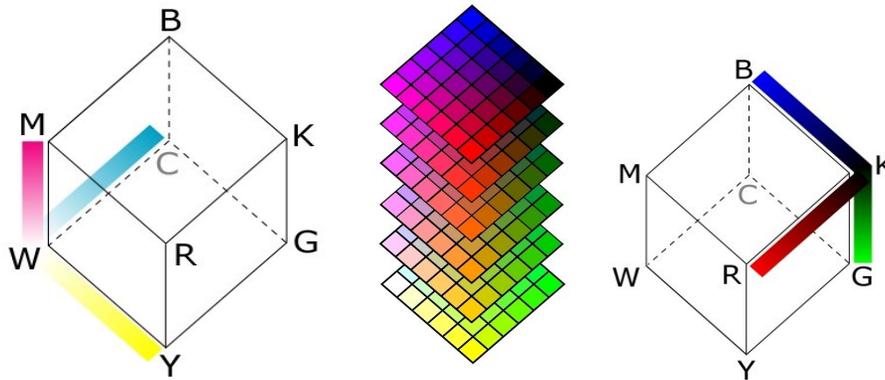


Imagem 11: cubo de cores com os seus eixos.

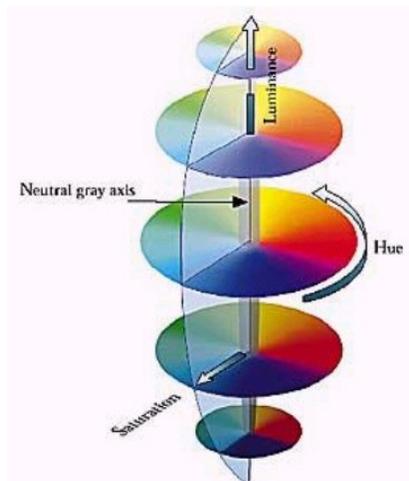


Imagem 12: representação das cores com o tom de cor, a luminância e a saturação.



Imagem 13: sistema de representação YUV, que será apresentado no capítulo 3.

As luzes que viajam no espaço se combinam aditivamente. A cor resultante é a soma das diversas luzes que se encontram. As três cores mais

apropriadas para representar esse tipo de luz são o vermelho, o verde e o azul. Elas são chamadas de cores primárias e estão representadas na Imagem 14.

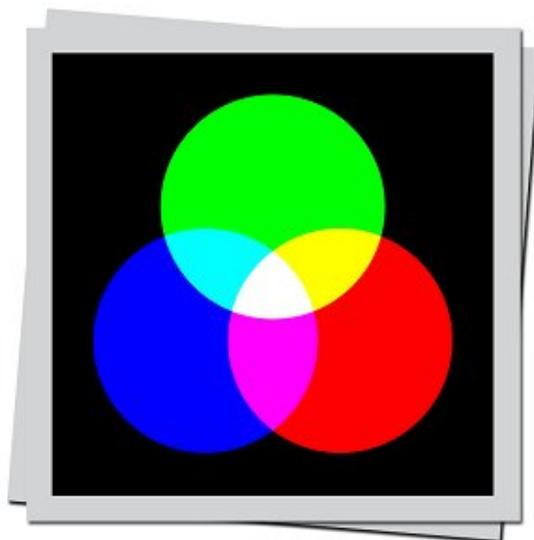


Imagem 14: as cores primárias: vermelho, verde e azul (RGB).

A soma do vermelho com o verde gera o amarelo. Do verde com o azul gera o ciano e do azul com o vermelho gera a cor magenta, e a mistura de cores se comporta diferentemente de uma combinação dessas cores com pigmentos. Essas cores geradas formam as cores secundárias e elas são complementares às cores primárias. O amarelo não tem qualquer componente do azul. O mesmo ocorre com o ciano com relação ao vermelho e com a cor magenta com relação ao verde. A soma das três cores primárias gera o branco, assim como de uma cor e de sua complementar.

As cores secundárias, isto é, ciano, magenta e amarelo, ou CMY, são usadas para representar as cores dos objetos e os pigmentos de suas superfícies. Elas se combinam de forma subtrativa, pois a luz refletida é aquela que sobrou da luz incidente na superfície, depois de ser absorvida. As cores secundárias são utilizadas no design gráfico porque as cores das peças gráficas são de luz refletida.

Na Imagem 15, são observadas as cores secundárias. Pode-se notar que elas se combinam diferentemente das cores aditivas e a união das três cores gera o preto. Na combinação de uma cor com a sua complementar, subtrativamente, é gerado o preto. Na imagem, tem-se também as cores

complementares das cores secundárias. Então, uma superfície amarela absorve toda a cor azul, refletindo amarelo, uma superfície ciano toda a vermelha, e uma superfície magenta absorve todo o verde da luz incidente.

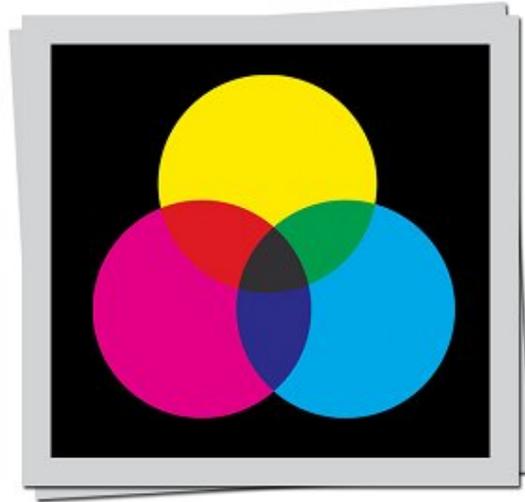


Imagem 15: as cores secundárias, ciano, amarelo e magenta (CMY).

Para exemplificar esta composição de cores, na Imagem 16 se visualiza as fotografias de um pássaro colorido, à esquerda, e a decomposição de cores nas componentes vermelha, verde e azul. Sob a componente vermelha, se tem um mapa de intensidade da luz, em tons de cinza, na última fotografia à direita.



Imagem 16: fotografia de um pássaro colorido e a sua decomposição nas cores vermelha, verde e azul, e, à direita, nos tons de cinza, da luminância no vermelho.

No site da [Brown University 2011], existem vários aplicativos Java para exercitar os conceitos de cor e de luz e de sua manipulação. Esses aplicativos são gratuitos e eles podem ser baixados no site da universidade. Uma série de

artigos sobre o tema pode ser encontrada no site [ABC 2011], em especial o artigo *Cor e Cinematografia*, [Ebert 2009].

1.3. Visão

O mundo é percebido visualmente através dos instrumentos humanos de visão. Eles são compostos dos dois olhos, das conexões que levam os sinais ao cérebro e da parte do cérebro que interpreta a visão. O mundo então pode ser representado da mesma maneira que o enxergamos e podemos falsear na imagem essa experiência visual. Isto é realizado, por exemplo, quando combinamos várias imagens gravadas em situações diferentes em uma única imagem, parecendo que toda ela foi capturada ao mesmo tempo.

Um diagrama esquemático do olho humano pode ser visto na Imagem 17. A luz incide na córnea e atravessa o humor aquoso, antes de atingir o cristalino, que é a lente do nosso sistema de visão. A íris funciona como um diafragma, que controla a intensidade da luz que passa para dentro do olho, através da lente do cristalino.

A luz atravessa o humor vítreo e a imagem é projetada na retina, no fundo do globo ocular. Na retina estão os sensores que percebem e capturam o sinal da luz, com a sua intensidade e a sua cor. Os nervos que levam o sinal elétrico gerado pela luz até o cérebro têm suas terminações nesses sensores.

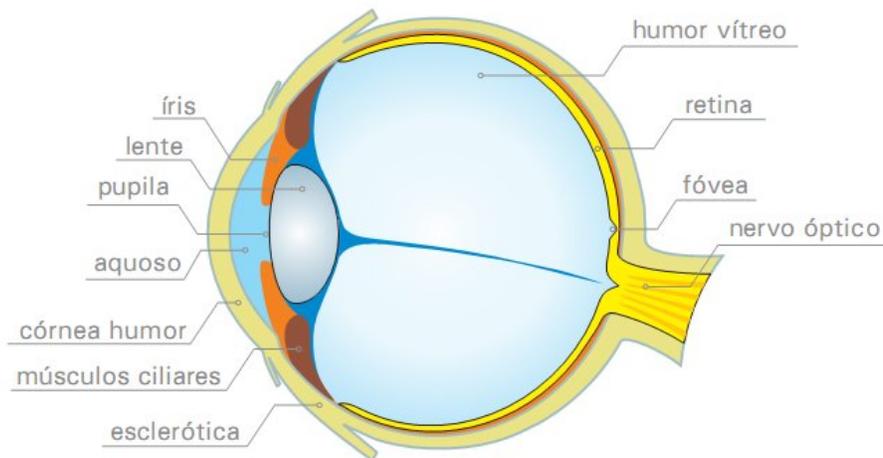


Imagem 17: diagrama esquemático do olho.

Os sensores que estão na retina são de dois tipos: os bastonetes e os cones. Eles são chamados assim devido ao seu formato. Os bastonetes são sensíveis à intensidade da luz e não distinguem cores, apenas tons de cinza. Os cones são sensíveis à cor.

Na Imagem 18, se tem uma imagem obtida com microscópio eletrônico dos cones e bastonetes em uma salamandra tigre, tiradas por Scott Mittman e David R. Copenhagen [Biology Pages 2011].

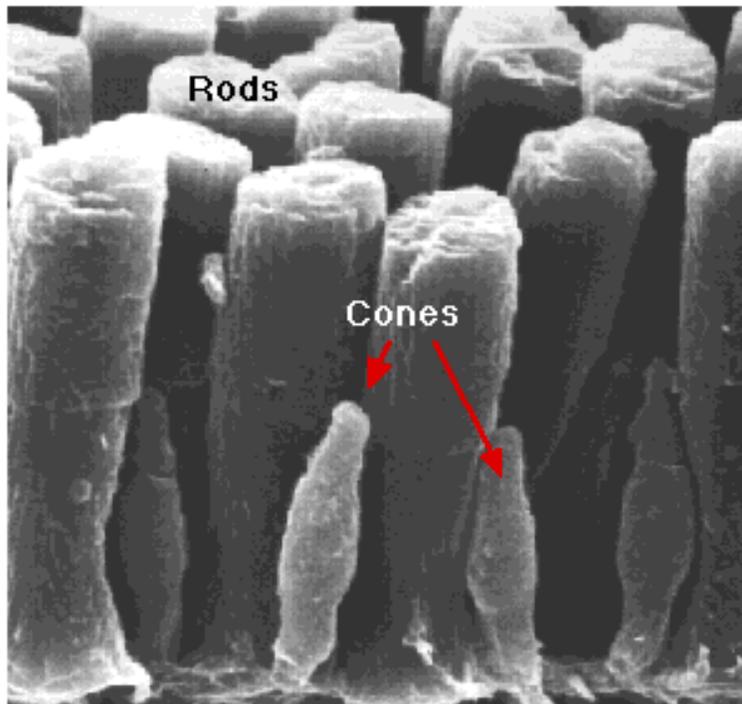


Imagem 18: os bastonetes e os cones na retina da salamandra tigre, feita com microscópio eletrônico em <http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/V/Vision.html>.

Os cones estão concentrados na região da fóvea, que está na retina, no eixo do olho. A concentração de cones na fóvea é responsável por uma maior nitidez e acuidade visual, como será discutido mais à frente. Eles são por volta de seis milhões e são divididos em três tipos diferentes, de acordo com a região do espectro visível em que funcionam.

Na Imagem 19, estão representados gráficos que indicam a sensibilidade de cada um desses três tipos de cones com relação à cor. O pico de intensidade de sensibilidade de cada tipo de cone está no azul, no verde e

no vermelho. Os cones então são chamados de cones azuis, verdes e vermelhos. Na imagem também se observa que o bastonete é mais sensível na faixa do ciano, apesar de ser sensível apenas à intensidade da luz. O sinal de saída dos bastonetes são tons de cinza, de acordo com essa intensidade. Os cones azuis representam apenas 5% do total de cones.

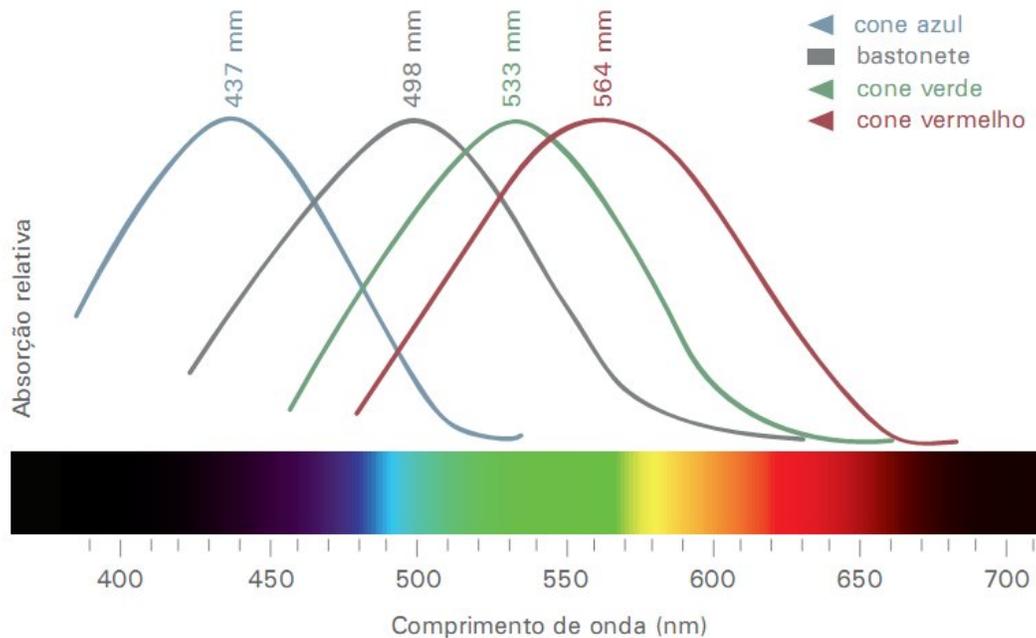


Imagem 19: Regiões de sensibilidade de cada um dos cones e dos bastonetes.

Os bastonetes estão em um número muito maior que os cones, sendo aproximadamente 125 milhões, contra os seis milhões de cones. Eles estão distribuídos ao redor da fóvea, mas um pouco distante dela, como pode ser observado na Imagem 20. Nessa imagem se observa a distribuição dos bastonetes e dos cones na retina, partindo do centro da fóvea.

Os bastonetes são responsáveis pela visão periférica, por estarem na periferia da retina, e, por serem cem vezes mais sensíveis à intensidade da luz que os cones, por isso são responsáveis pela visão noturna. Em uma intensidade de luz baixa, os cones não funcionam. Como a saída de sinal para o cérebro em baixa intensidade de luz são tons de cinza, tem sentido dizer que *à noite todos os gatos são pardos*.

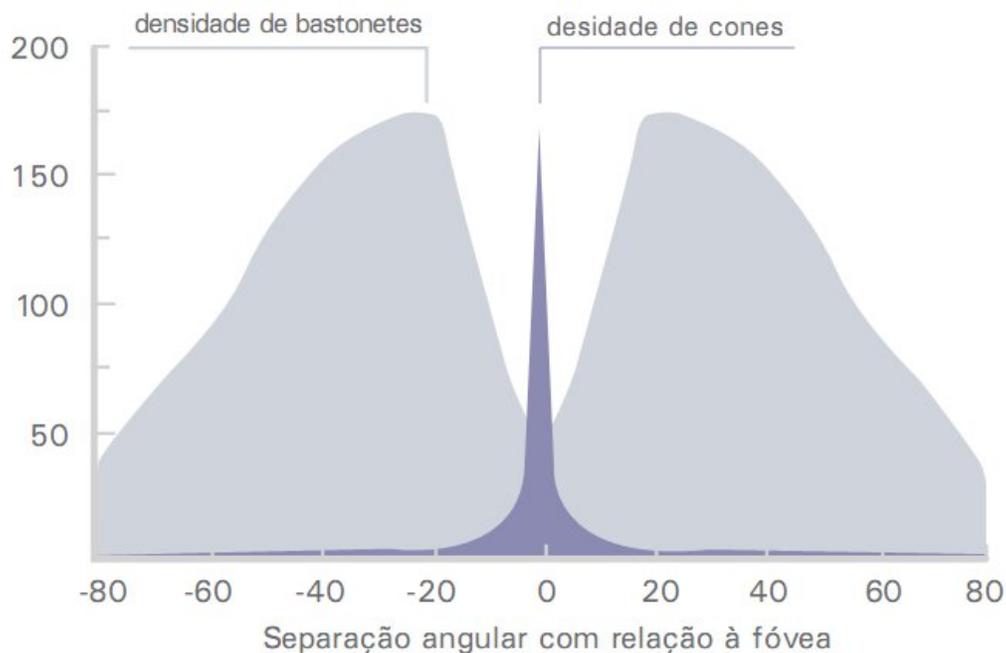


Imagem 20: distribuição dos cones e dos bastonetes, em milhares por milímetro quadrado, com relação à distância angular ao centro da fóvea.

A combinação dos cones e dos bastonetes dá a característica de como o mundo é registrado visualmente. Alguns tipos de cones, por exemplo, podem não funcionar adequadamente. Os azuis nunca apresentam problemas, apesar de existirem em um número muito menor. No entanto, os vermelhos ou os verdes, ou ambos, podem não funcionar, ocasionando o daltonismo, que é comum entre os homens.

A sensibilidade às cores entre a verde e a laranja é maior devido a uma maior concentração de cones nessa faixa do espectro. Essas características da sensibilidade às cores possivelmente vieram da seleção natural, pois existem pouquíssimos alimentos azuis e a maioria está entre o verde e o vermelho. Também é observado que a imagem na periferia do nosso campo visual não é bem definida em cores, porque essa imagem é capturada principalmente pelos bastonetes. Na Imagem 21, se tem um corte do olho humano para que se possa observar o seu interior, onde estão destacados o cristalino e a fóvea.

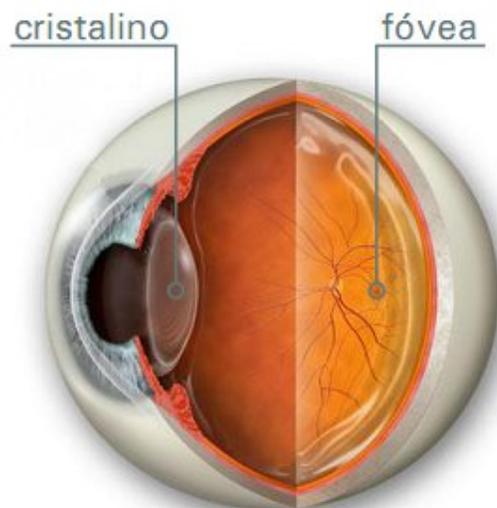


Imagem 21: corte do olho humano, mostrando o seu interior, com o cristalino e a fóvea em destaque.

A representação da imagem digital leva em consideração essas informações. Como cada imagem tem um tamanho em bits, quanto mais precisa a representação da imagem, maior o número de bits de informação. Isto tem um valor significativo quando se quer transmitir, armazenar ou exibir a imagem, como no caso de uma banda de transmissão de dados limitada ou quando se trabalha com vídeo, com uma taxa grande de quadros de imagem por segundo. Então, através dos estudos de percepção visual, se pode reduzir a quantidade de informação da imagem, para se adequar às demandas de transmissão e armazenamento. Isto pode ser feito, por exemplo, reduzindo a precisão da representação das cores. E isso acontece com relação a cada tom de cor. O mesmo não pode ocorrer com relação à intensidade da luz, pois a quantidade de bastonetes é muito maior que a de cones. Isto leva a soluções adequadas para cada caso, com os diferentes codificadores e decodificadores digitais de imagem, que podem gerar imagens para a web mais limitada, para a televisão de alta definição ou para o cinema digital, com diferenças significativas na qualidade da imagem e na quantidade de dados transmitidos.

1.4. Projeções de câmeras. Câmera *pinhole*.

Um dos problemas principais, que se apresenta em quase toda a tese, está relacionado ao processo de formação da imagem. Ela tem diversas implicações em várias aplicações dos efeitos visuais, como na captura de movimento e no rastreamento de imagem, para citar alguns. Nestes dois casos, se parte da análise da imagem para gerar dados geométricos, em aplicações da visão computacional. Mas a câmera é fundamental na geração da imagem.

Uma câmera é um mapeamento entre o mundo 3D, que pode ser o universo físico ou um espaço abstrato, e uma imagem bidimensional, que pode ser, por exemplo, uma fotografia ou a saída de um render. Neste processo se perde uma dimensão. Em todos os casos desta tese, a projeção da câmera é central, isto é, um raio de luz de um ponto no espaço passa por um ponto fixo no espaço, que é o centro de projeção. Esse raio vai interceptar um plano específico, que é o plano da imagem, ou plano de projeção. A intersecção do raio com o plano da imagem representa a imagem do ponto. Uma representação desta situação encontra-se na Imagem 22.

[Hartlley e Zisserman 2006] pontuam que esse modelo está de acordo com um modelo simples de câmera, no qual um raio de luz de um ponto no mundo passa através da lente de uma câmera e incide no filme ou no dispositivo digital, produzindo uma imagem do ponto. Analogamente, o raio de luz pode vir de um ponto em um espaço de representação computacional 3D e o resultado é o render a partir de uma câmera sintética. Ignorando os efeitos gerados pela largura da lente e pela precisão do foco, uma aproximação razoável é de que os raios passam através de um simples ponto, que é o centro da lente. Este é o caso da projeção perspectiva. O olho pode ser, a grosso modo, aproximado por esse modelo, com o cristalino funcionando como uma lente de distância focal variada, de acordo com o trabalho dos músculos da córnea. As máquinas fotográficas, as filmadoras e os programas de computação gráfica produzem imagens baseadas nessa projeção em perspectiva, porque é o tipo de projeção de nossos olhos.

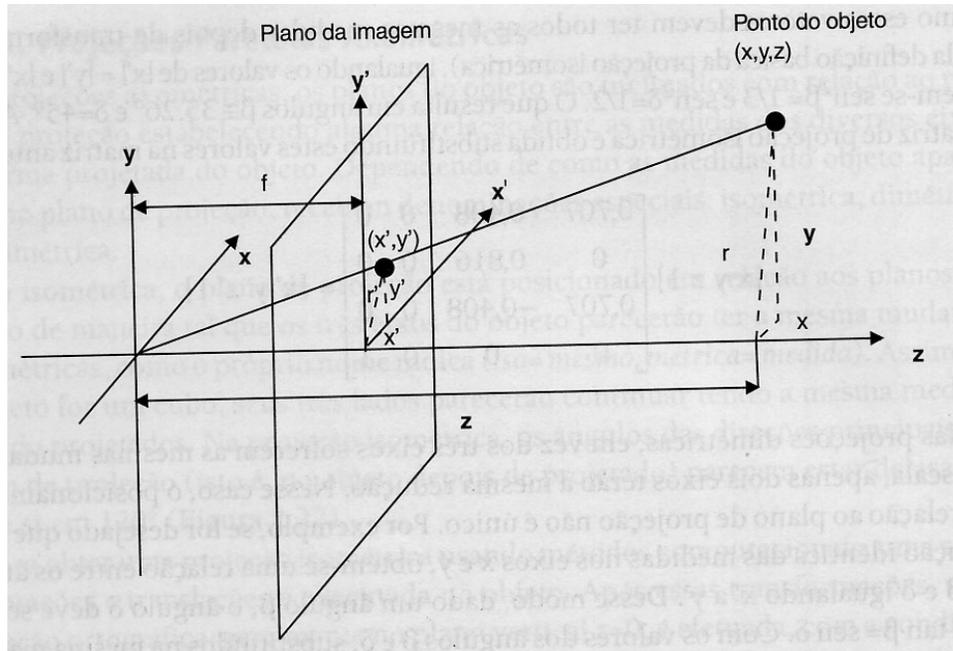


Imagem 22: diagrama de uma projeção central [Azevedo 2003].

Todo o tipo de câmera com projeção central é um caso específico da câmera projetiva geral [Hartley e Zisserman 2006]. A sua estrutura é estudada utilizando-se as ferramentas da geometria projetiva. Os entes geométricos mais importantes da câmera, como o centro de projeção e o plano de imagem, podem ser computados de forma simples por suas representações matriciais. Os modelos são de duas classes principais – as câmeras com centro finito e os modelos de câmera com o centro “no infinito”.

Existem diversas modalidades de transformação do espaço 3D em um plano 2D, através da projeção. Isto é amplamente tratado pela geometria projetiva. Em um diagrama encontrado no livro de [Azevedo 2003], que está na Imagem 23, são exibidos todos os tipos de projeção planificada. A projeção de interesse neste trabalho é a perspectiva, que é o processo que ocorre no olho e nas câmeras utilizadas no audiovisual.

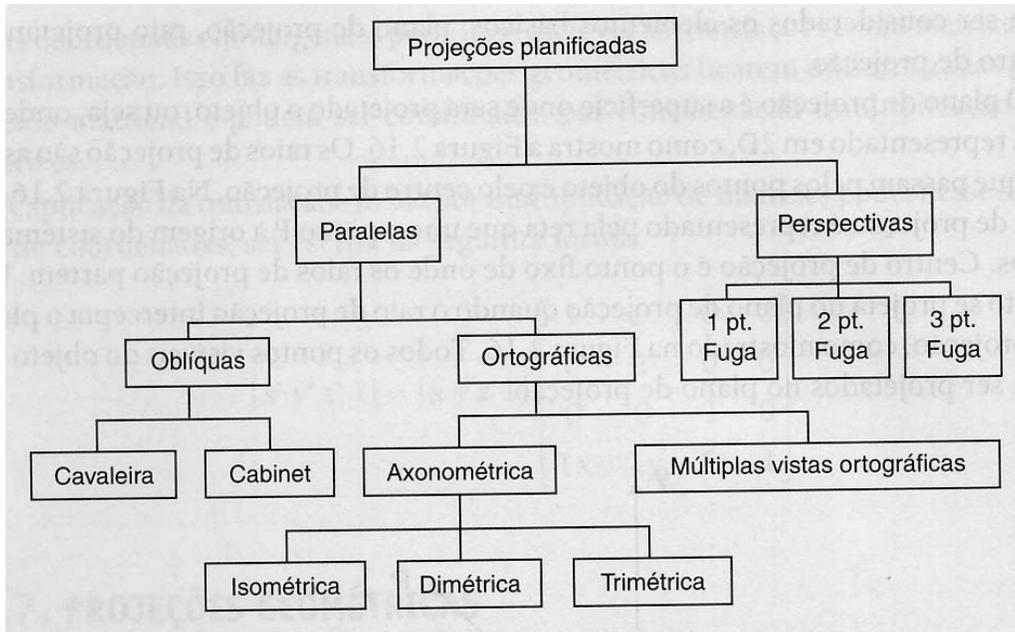


Imagem 23: diagrama com os tipos de projeções [Azevedo 2003].

A projeção paralela ortográfica também é importante para auxiliar no processo de pré-visualização da modelagem tridimensional nos softwares de computação gráfica. Nela, o centro de projeção está representado no infinito e todas as linhas de projeção são paralelas entre si. No caso da projeção ortográfica, as linhas de projeção são perpendiculares ao plano de projeção, como na Imagem 24. Na projeção paralela oblíqua, que é muito utilizada em cristalografia e em outras aplicações, o conjunto de linhas de projeção são inclinadas em relação ao plano de projeção por algum ângulo diferente de 90° . Existem lentes que projetam os raios paralelamente ao plano de imagem, as lentes telecêntricas, e são utilizadas para medidas específicas de distâncias. Elas são raramente utilizadas no audiovisual.

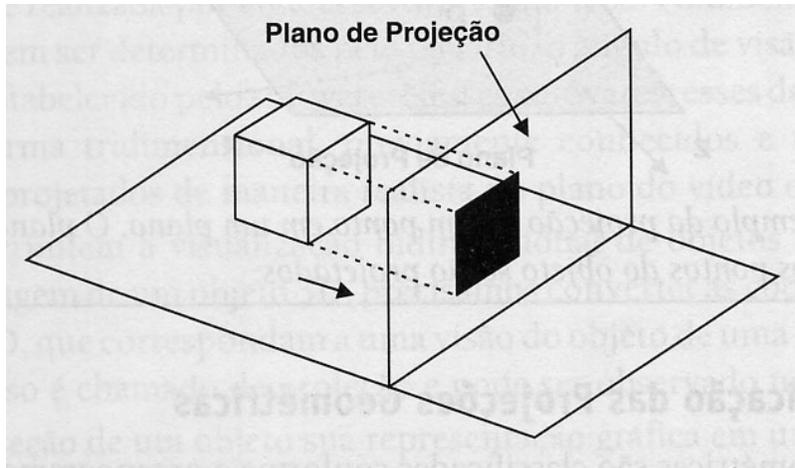


Imagem 24: projeção paralela ortográfica [Azevedo 2003].

O modelo mais simples para descrever uma câmera é a câmera do tipo *pinhole*. A câmera *pinhole* é apenas uma caixa com um furo (*hole*) feito com alfinete (*pin*). A luz entra pelo furo e forma a imagem no fundo da caixa, como se vê na Imagem 25. Nessa representação é mostrada, à direita, a caixa e a imagem formada no fundo da caixa.

A imagem tem a forma invertida com relação ao objeto do mundo real. Isso acontece porque os raios de luz de cada ponto do mundo visível vão em linha reta até atingirem a parede da caixa, passando pelo furo. Para capturar a imagem, no plano de projeção é colocado um sensor que a registra durante o tempo em que se deixa a luz passar através do furo, daí o nome câmera fotográfica *pinhole*.

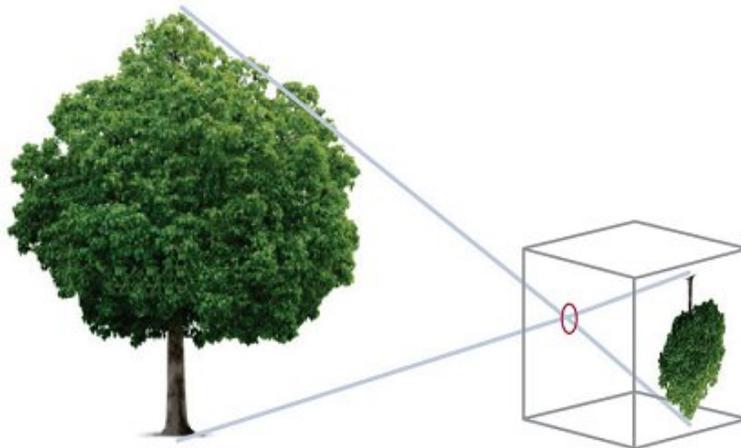


Imagem 25: uma câmera pinhole registrando uma fotografia.

O bloqueio da luz é feito por um obturador, que pode ser o dedo, que deixa a luz passar durante um intervalo de tempo. Nas máquinas fotográficas, com conjuntos de lentes e ótica adequada, o tempo médio de abertura do obturador em um dia ensolarado é de 1/125 segundos.

O modelo matemático da câmera *pinhole* será discutido ao longo da tese, quando forem apresentadas aplicações específicas desse modelo, no rastreamento de imagem e na captura de movimento. Esse modelo é especialmente adequado para câmeras digitais com sensores CCD ou CMOS, como será discutido no capítulo 3.

1.5. Ilusões de óptica

O olho funciona como uma câmera de *pinhole*. O mundo real, que é tridimensional, isto é, tem largura, altura e profundidade, é projetado na retina ou na placa fotográfica, que possuem apenas duas dimensões, a largura e a altura. Os raios de luz passam por um furo – que no olho é a pupila - e viajam na forma de um cone depois de passarem pelo furo. Esse tipo de projeção faz a imagem ficar invertida quando é projetada.

Devido a essa característica, algumas ilusões ópticas básicas ocorrem, que são intrínsecas à maneira em que os nossos olhos funcionam. Uma primeira ilusão é a de que, à distância, duas linhas paralelas se encontram. Isso pode ser observado quando se olha os trilhos do trem ou do metrô e eles parecem se encontrar no horizonte. Ou quando se olha para cima, da quina de um edifício alto, e parece que ele vai se afinando nas suas laterais, à medida que sobem.

É sabido que não é assim que acontece, pois as linhas paralelas não se encontram. O ponto onde parece que as linhas vão se encontrar é chamado de *ponto de fuga* e ele está no horizonte da visão. Pode haver um ponto de fuga, como no caso dos trilhos, dois pontos de fuga, como no caso do edifício alto, ou três, como no caso de um edifício alto e largo. A reta onde as paralelas parecem convergir é chamada de *linha do horizonte*. Na Imagem 26, são mostrados exemplos de alguns efeitos da projeção em perspectiva.

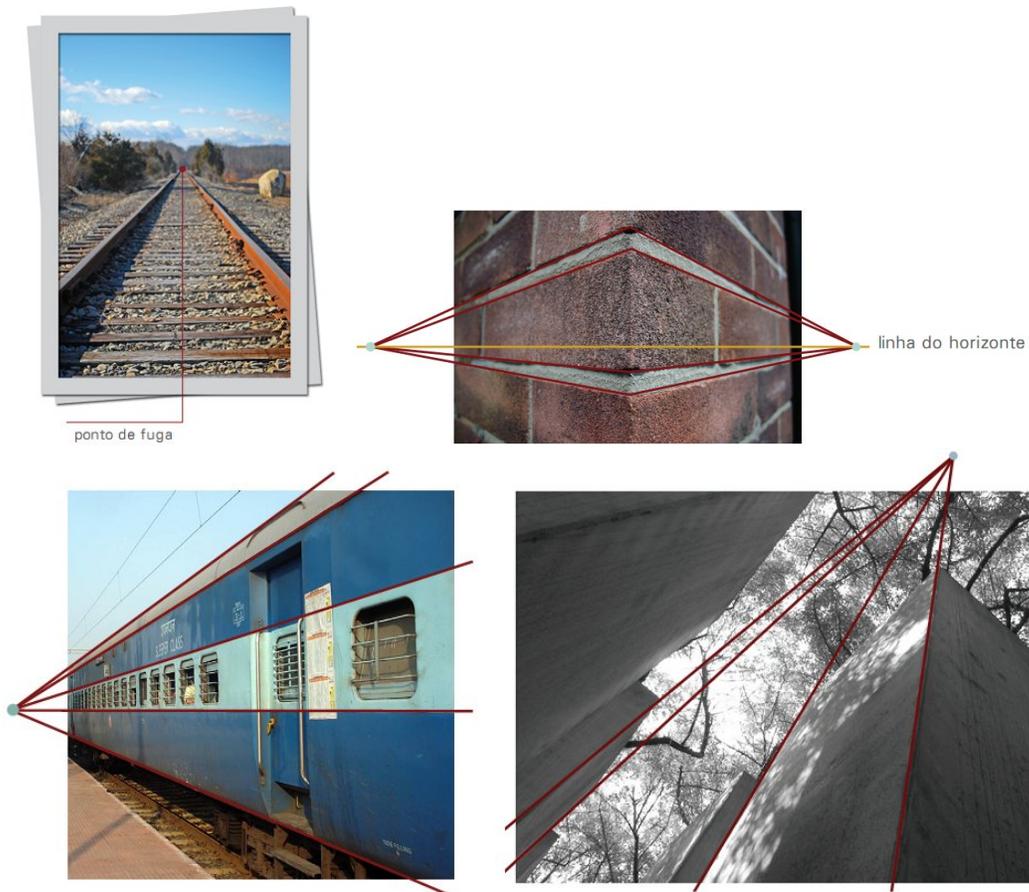


Imagem 26: exemplos de ilusão de óptica, com os pontos de fuga e a linha do horizonte.

Outra ilusão, que é consequência desse tipo de projeção, é que os objetos parecem menores quando se distanciam do observador. Ocorre a ilusão de que os comprimentos dos objetos se alteram.

Os mecanismos pelos quais são percebidas as cores e a intensidade da luz não são de medidas absolutas. A cor de um objeto, em geral, é a combinação de muitas cores diferentes, espalhadas pelo espectro visível. Porém é registrado apenas o resultado médio dessas cores, que gera a cor que é percebida.

A percepção visual aos tons de cores também não é absoluta e depende de como os sensores da retina respondem aos estímulos da luz e em como eles são interpretados pelo cérebro. A percepção é baseada em comparações e isso altera a aparência dos tons de acordo com os contrastes simultâneos do

objeto com o ambiente. Também ocorre a formação de pós-imagens, quando se fixa o olhar por muito tempo em um cenário e depois se desviam os olhos, persistindo a imagem do cenário. Deficiências da visão colorida também alteram significativamente a percepção da luz.

Existem algumas teorias sobre o mecanismo de visão das cores, com abordagens equivalentes, de acordo com a utilização da cor [Brinkman 2008, Ebert 2009]. Na teoria de Hering, ou teoria tetracromática, a percepção das cores se dá por oposição entre pares mutuamente excludentes. Esses pares são o vermelho, verde e o azul - amarelo. Foram realizados experimentos com objetos cinza colocados em diferentes tonalidades de fundo para a observação dos contrastes simultâneos.

Um objeto cinza colocado sobre um fundo vermelho parece esverdeado; sobre um fundo verde, avermelhado; sobre um fundo amarelo, mostra-se azulado e sobre um fundo azul, mais amarelado. Nas três placas apresentadas a seguir, na Imagem 27, pode-se observar o fenômeno do contraste de cor, onde o tom do círculo vermelho é percebido diferentemente de acordo com o fundo.

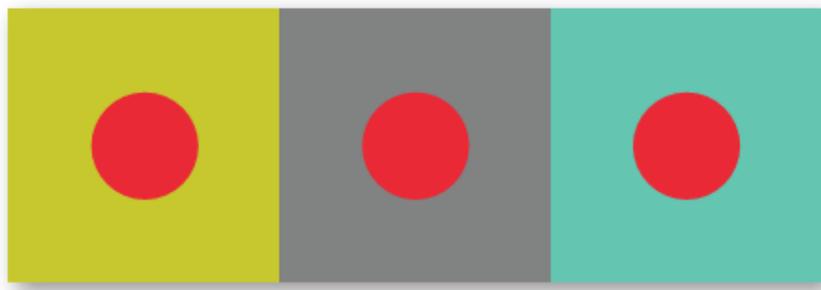


Imagem 27: aparência da cor vermelha com diferentes fundos coloridos.

Da mesma maneira, os tons de cinza são percebidos, quando comparados ao ambiente, através do contraste de luminância. As diferenças entre os tons de cinza é que são notadas e muitas vezes levam a interpretações equivocadas, devido a essa comparação. Na Imagem 28, um círculo cinza é exibido com diferentes fundos, e aparenta ter diferentes tonalidades de acordo com o tom de cinza do fundo.

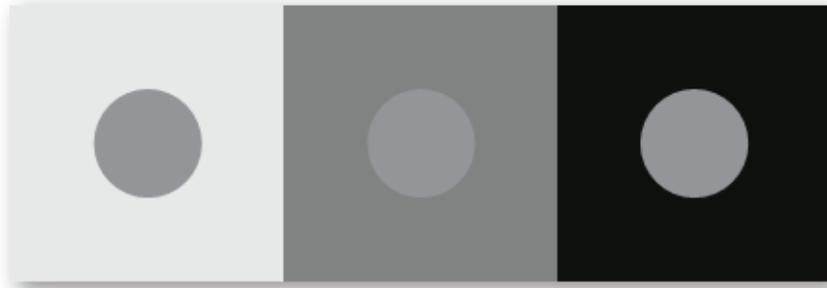


Imagem 28: aparência de um mesmo tom de cinza com diferentes fundos.

Um teste impressionante para essa ilusão, que resulta do contraste de luminância, é a ilusão da sombra. O cérebro faz correções na tonalidade do cinza. No caso do tabuleiro apresentado na Imagem 29, é mostrada uma sombra provocada pelo cilindro. As tonalidades de cinza da casa com a letra A parece completamente diferente da tonalidade da letra B.

Mas, ao se colocar sobre o tabuleiro uma régua com o tom de cinza igual ao da casa com a letra A, se percebe que o tom de cinza de A é o mesmo da casa B, como apresentado na Imagem 30. E isso é completamente diferente do que se percebe na imagem apresentada anteriormente.

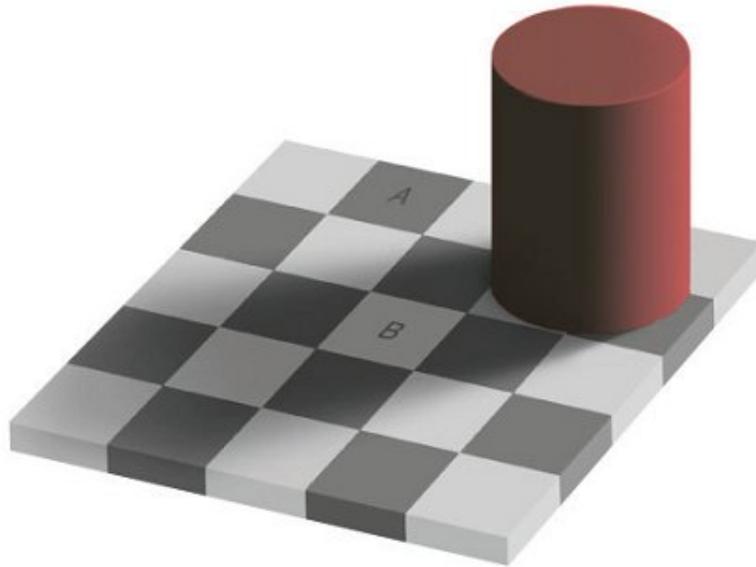


Imagem 29: tabuleiro com as casas A e B marcadas, um cilindro vermelho e sua sombra.

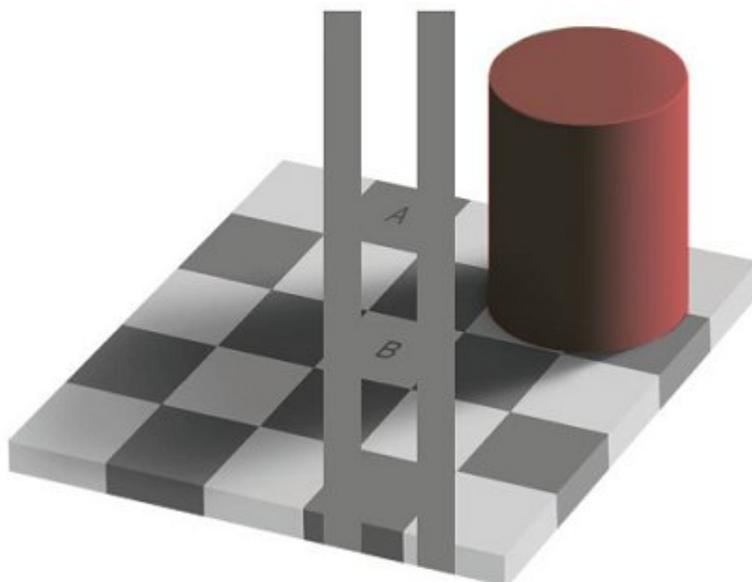


Imagem 30: tabuleiro com as casas A e B, uma régua cinza, um cilindro vermelho e sua sombra.

Tendo como base estas observações, desenvolveram-se diversos esquemas de composição de cores e imagens nas artes plásticas e no audiovisual. Eles são abordados em [Ebert 2009] e [Heller 2007] e não serão tratados aqui, apesar do interesse e da curiosidade que geram, porque fogem do escopo deste trabalho.

CAPÍTULO 2

GERAÇÃO DA IMAGEM

2.1. A imagem pré-digital

A humanidade produz e constrói representações por imagens do mundo real e de mundos abstratos desde os primórdios da história. Os primeiros registros de imagens em pinturas remontam a mais de 40.000 anos atrás [Enciclopédia Itaú Cultural 2012], com as pinturas rupestres. Esta representação tem evoluído e assumido diferentes aspectos ao longo da evolução humana. Novos suportes foram introduzidos e aprimorados ao longo deste tempo, como as telas de desenho e pintura e seus materiais e pigmentos, as projeções, a fotografia e os dispositivos de exibição de raios catódicos, cristal líquido e plasma.

Até a primeira fotografia digital [Wired 2011], capturada em 1957, a produção e a exibição de imagens só era possível através da pintura, do desenho, da fotografia, do cinema em película ou da televisão analógica. Para se produzir imagens, se deveriam usar pincéis, lápis, pigmentos, câmeras fotográficas com filme, câmeras de cinema ou câmeras de televisão e vídeo analógicas. Quase não havia qualquer tipo de interação, devido à maneira como essas imagens eram criadas. A única interação em tempo real possível era obtida com o vídeo, através de modificações eletrônicas da imagem, em equipamentos complicados e pouco produtivos.

Com o aparecimento e a evolução da tecnologia digital, esses processos de produção de imagens foram sendo modificados e, muitas vezes, substituídos. Foi também possível fazer com que o espectador interagisse com a imagem, abrindo o caminho para os jogos eletrônicos, a realidade virtual e a realidade aumentada. A produção de imagens antes da imagem digital se dava pelos métodos tradicionais, muitos desenvolvidos por séculos, como a pintura e o desenho.

A película é o filme utilizado nas câmeras de cinema antes da era digital. O sinal analógico é o sinal elétrico que representa o contínuo de uma grandeza

física. A intensidade da luz e a sua cor foram durante muitos anos representados assim. Apesar do uso dessas imagens serem cada vez mais reduzido, é importante entender como ela funciona, porque a qualidade dessas imagens é ainda comparável ou melhor que a da imagem digital. Muitas vezes o abandono do uso da imagem analógica e da película se deu pela maior simplicidade do trabalho com a imagem digital. Atualmente ela é representada pelos valores digitais, que são discretos e não contínuos.

No caso da pintura e do desenho, as cores resultantes são combinações das cores dos pigmentos da tinta, misturados de forma subtrativa, conforme discutido no capítulo anterior. Na Imagem 31, se observam as cores secundárias, amarelo, ciano e magenta, sendo combinados com pincéis, em uma representação artística.



Imagem 31: pincéis com as cores secundárias e o resultado das misturas de cores.

Para a fotografia e o cinema, as imagens eram capturadas em um filme, através de um processo fotoquímico. No olho, a imagem é formada na posição da retina. Na câmera fotográfica ela é formada no lugar onde o filme é colocado. O filme é enrolado em uma bobina e um mecanismo faz com que ele se desloque da sua posição, trocando o fotograma que já foi impressionado pela luz por um fotograma virgem. No caso da câmera de cinema, esse processo acontece 24 vezes por segundo, para que haja a ilusão do movimento quando o filme é revelado e exibido. Na Imagem 32, é mostrado o

desenho esquemático da formação da imagem no olho e na câmara fotográfica.

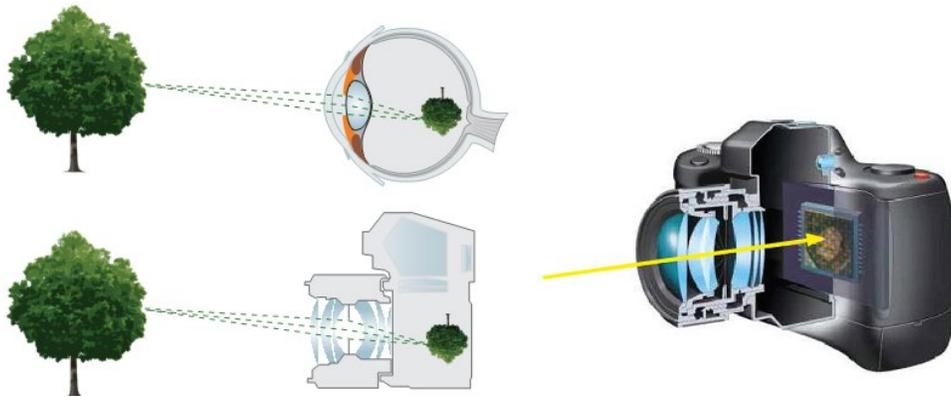


Imagem 32: À esquerda, as imagens sendo formadas no olho e em uma câmara fotográfica; à direita, a posição do sensor mostrado através de um corte na câmara.

O filme é feito de uma emulsão de sais de prata, que sofrem alterações químicas quando a luz incide sobre eles. Depois de um banho em uma solução que reage com os sais de prata modificados pela luz, a imagem é revelada no filme. A imagem é o negativo das cores que incidiram na película, isto é, a imagem é composta das cores complementares. Na Imagem 33 são exibidos uma filmadora para cinema e um pedaço de película.



Imagem 33: uma filmadora Arri e um pedaço de película.

Assim, se a luz branca incide sobre o filme, ele projeta em uma parede branca a imagem com as cores complementares à imagem capturada. Se essa imagem for ampliada em uma folha de papel fotográfico, o mesmo processo químico se repete. A imagem é ampliada passando-se luz branca através do filme e incidindo essa luz por um tempo determinado, no papel fotográfico. Depois disso, o papel passa por um banho e sofre as mesmas modificações químicas do filme, gerando uma imagem com as cores complementares às que estão registradas no filme. Dessa maneira, uma imagem positiva é gerada, recuperando as cores originais.

Na Imagem 34, é exibida uma ampliadora de fotografias. A luz gerada pela lâmpada (2) passa pelas lentes (3) e (7). Ela então atravessa o filme posicionado em (9) e a imagem é formada no papel fotográfico que está em (14). Para fixar a imagem, o papel passa por um banho químico.

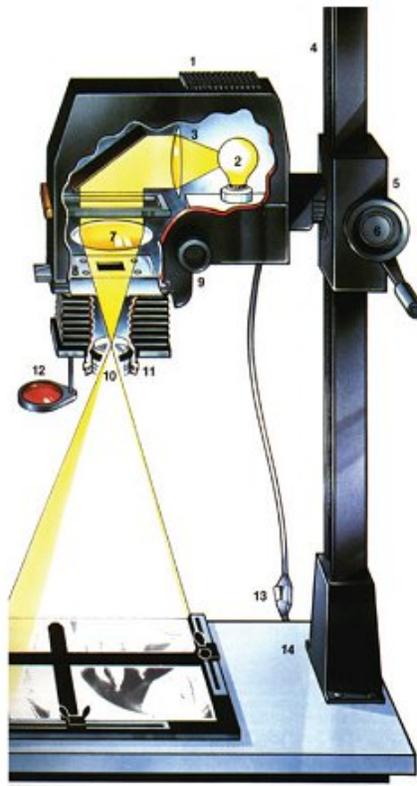


Imagem 34: corte esquemático de uma ampliadora de fotografia.

O primeiro processo químico pelo qual o filme passa gera o negativo. O negativo do negativo, no segundo processo, gera um positivo. Como a

fotografia é impressa, a imagem que é observada são das luzes refletidas no papel fotográfico. Para o cinema o processo é semelhante, mas a imagem é projetada por uma luz que o atravessa. O filme que será exibido é o positivo da imagem.

Na Imagem 35, se observa o negativo, em B e D, e o positivo, em A e C. No caso da fotografia preto e branco, têm-se os tons de cinza complementares.

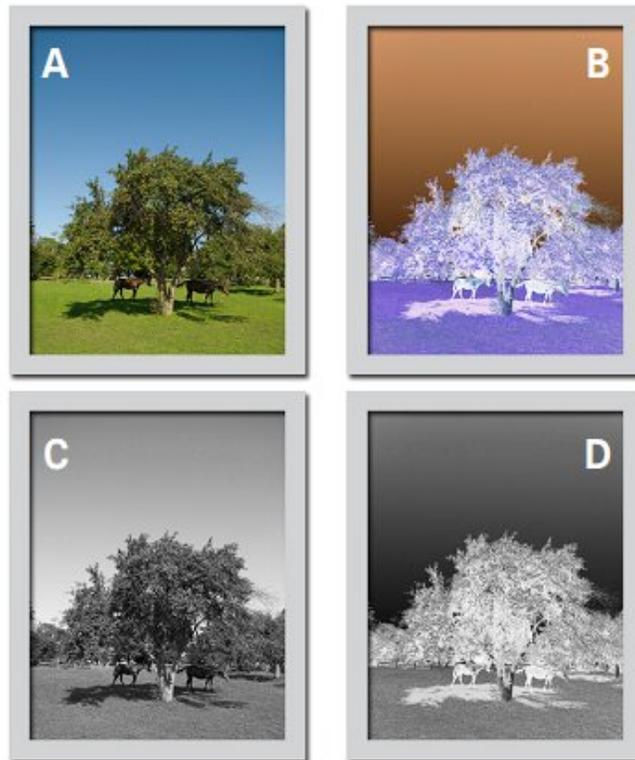


Imagem 35: fotografias colorida e preto e branco, à esquerda, e os seus negativos à direita.

A substituição da película e do vídeo gerado eletronicamente pela imagem digital está em plena evolução acelerada. A melhoria da qualidade da imagem digital e dos dispositivos de aquisição de imagem tem contribuído para que isso aconteça muito rapidamente. Atualmente, as imagens em movimento em película são, na maioria das vezes, digitalizadas e passadas novamente para película para distribuição, depois de finalizado o filme.

Uma discussão sobre essa evolução para o digital está no livro de [Murch 2004]. Ele coloca o seu ponto de vista de montador de cinema vivendo

essa transição, com várias histórias reais. Ele trabalha há mais de quarenta anos na área, é dos mais atuantes ainda hoje e ganhou vários prêmios, inclusive dois Oscars de edição.

No caso da televisão e do vídeo analógicos, a combinação das cores é aditiva. A imagem é capturada através das câmeras. A luz entra pela lente, atravessa um prisma e incide nos sensores. A função do prisma é separar a luz em suas cores primárias, vermelha, verde e azul.

Em uma câmera com três sensores, cada sensor recebe uma componente da luz e transforma o sinal luminoso em sinal elétrico, modificando também o espaço de cor para a luminância, todas as cores menos o vermelho e todas as cores menos o azul, o chamado sistema componente, ou YUV. Esse sinal é então gravado nas fitas magnéticas, que registram a imagem. A câmera também pode ter um sensor apenas mas perde-se muito na qualidade da imagem.

Depois, ao exibir as imagens, cada componente de luz primária é somada produzindo a cor resultante, como se pode ver na Imagem 36. A imagem é formada em televisores de tubos. No fundo desse tubo tem um emissor que produz um feixe de elétrons, a letra A da figura. Esse emissor produz três feixes de elétrons que são colimados e direcionados, como na letra C.

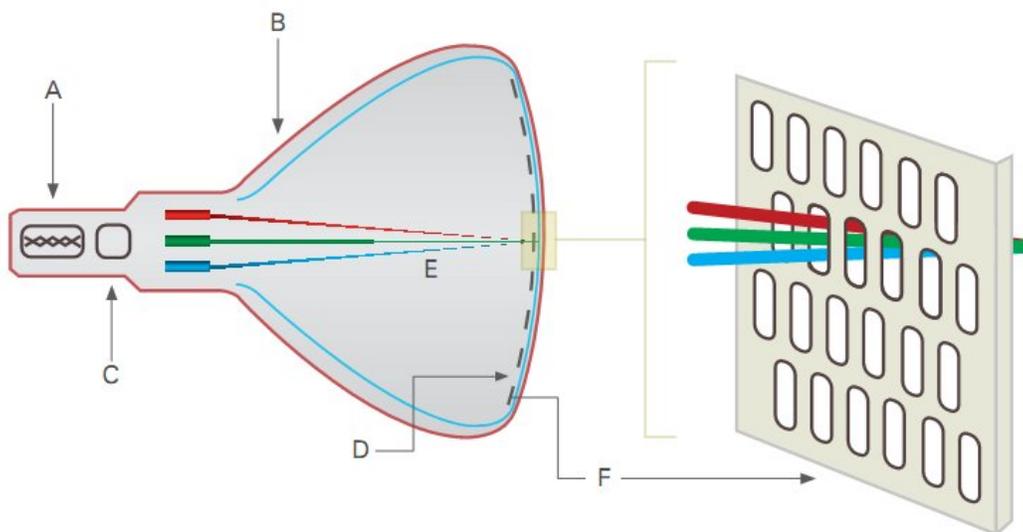


Imagem 36: corte esquemático de um tubo de televisão, mostrando o detalhe do tubo com fósforo, que colore a imagem ao ser atingido por elétrons.

Na Imagem 36, os feixes de elétrons são coloridos artisticamente nas três cores primárias. Isso é para frisar que cada um deles, que não tem cor, é responsável por produzir cada uma dessas cores no tubo. Os feixes incidem no tubo do monitor, que é coberto com compostos de fósforo, que se colorem quando os elétrons incidem neles, como na letra F. Existem compostos de fósforo que emitem no vermelho, no verde ou no azul, e cada feixe incide em um dos tipos. A cor resultante da imagem é a combinação aditiva das três luzes primárias.

2.2. Origens da imagem digital

A tecnologia digital se baseia na transformação de toda informação em bits. O bit é o *binary digit*, isto é, um dígito com dois valores, que, por facilidades da matemática, são o 0 e o 1. O bit forma a base de toda a informação e a combinação adequada dos bits pode representar qualquer coisa, como a imagem, por exemplo. De maneira mais geral, pode-se imaginar o bit como a representação de dois estados opostos, como o verdadeiro e o falso, o claro e o escuro, o frio e o quente e assim por diante.

Para a imagem, tem-se que reproduzir todas as informações possíveis de cor, iluminação, linhas, superfícies e volumes em um encadeamento de bits. Ao se observar o mundo à volta ou uma fotografia, se constata que a quantidade de informação é imensa. A fotografia do busto de uma pessoa em um fundo com uma única cor tem muito menos informação que a fotografia de uma paisagem, com todos os seus detalhes, como as folhas e as plantas, por exemplo.

O início da imagem digital aconteceu em 1957, quando Russell Kirsch, um pesquisador do Massachusetts Institute of Technology (MIT), nos Estados Unidos, usou a fotografia de seu filho recém-nascido para produzir a primeira imagem digital [Kirsch and Kirsch 1988]. Para fazer isso, ele usou um equipamento chamado fotomultiplicadora. A fotomultiplicadora é um aparelho que capta a intensidade da luz. A luz incide nos sensores e faz gerar uma cascata de elétrons, que gera uma corrente elétrica. Ele então fez um arranjo em linha de 176 fotomultiplicadoras apontadas para a fotografia e capturando,

cada uma, um pedaço da imagem na largura da fotografia. Ele moveu essa linha de fotomultiplicadores ao longo da altura da fotografia, de ponta a ponta, em 176 posições diferentes.

Cada uma das fotomultiplicadoras gerou sua corrente elétrica da seguinte maneira: onde a fotografia era preta não havia sinal e onde ela tinha algum tom de cinza, era gerada a corrente. Assim, a saída das fotomultiplicadoras, em cada uma das suas posições ao longo da altura da fotografia, tem dois valores e é representada por um bit. Ele fez um arranjo de pequenos retângulos que compunham a fotografia, como se fosse uma matriz, com 176 linhas e 176 colunas.

Na saída de cada fotomultiplicadora ou havia sinal ou não, reproduzindo uma fotografia com um bit em cada quadrado, preto ou branco. Na Imagem 37 se tem a reprodução dessa primeira imagem digital.



Imagem 37: primeira imagem digital, de 1 bit de cor por pixel, de 1957 [Wired 2011].

Cada retângulo é um elemento da imagem. No caso dessa primeira imagem digital, cada retângulo é o que cada fotomultiplicadora enxerga. A imagem digital é então organizada desde o seu início como um arranjo de elementos de imagem, que é chamada de *pixel*, acronismo de *picture elements*. Cada fotomultiplicadora gera a representação de um pixel. Esse

arranjo pode ser do tipo matricial, como nesse caso, com linhas e colunas de pixels.

De forma geral, cada pixel tem um valor único de cor e de intensidade de luz. Esse valor é dado pelo arranjo de bits. No caso da primeira imagem digital, com apenas dois valores, cada pixel tinha um bit. A quantidade de bits é chamada de profundidade de tons da imagem e a quantidade de pixels dá a resolução da imagem.

Em 1962, a possibilidade de se manipular a imagem na sua forma digital levou à invenção da primeira ferramenta que utilizava uma caneta ótica, o *sketchpad*. Ela permitia desenhar diretamente na tela de um computador ainda rudimentar, com diversos botões para controlar o sinal da caneta. Dessa maneira, surgiu a primeira forma de interatividade com a imagem, que é exibida na Imagem 38.

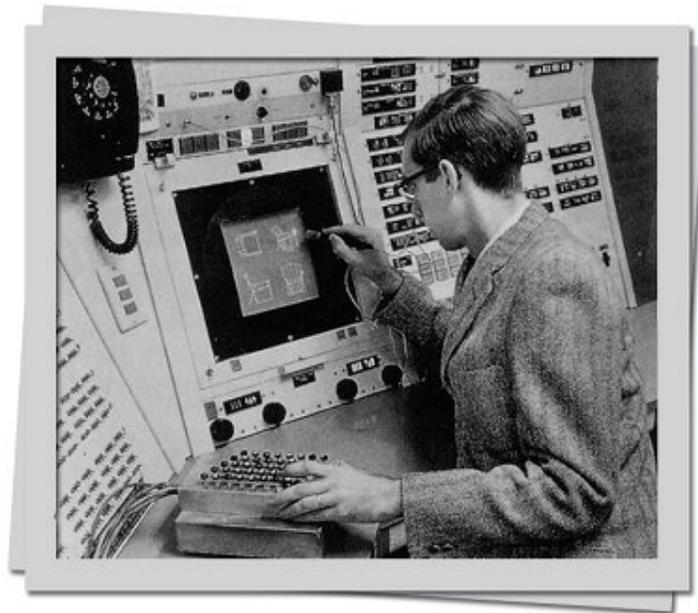


Imagem 38: primeira forma de interação com a imagem, o *Sketchpad* (http://resumbrae.com/ub/dms423_f08/06/sketchpad.jpg).

A partir daí, surgiram novos artefatos de interação, como o *joystick* e o *mouse*, e a resolução e a qualidade da imagem aumentaram. Isso foi possível com o aumento da velocidade dos processadores, da memória dos computadores, da capacidade de armazenamento e acesso aos dados e através da evolução das formas de exibição das imagens.

Os programas que permitem a interatividade e a manipulação da imagem também evoluíram e puderam desempenhar cada vez tarefas mais complexas. Os componentes dessas máquinas também se tornaram cada vez menores e mais poderosos, permitindo a redução considerável do tamanho dos equipamentos.

O impacto representado pelo surgimento da imagem digital e da computação gráfica é muito marcante hoje, acompanhando a evolução dos hardwares e dos softwares. A área dos jogos pode ser usada como um grande exemplo de transformação, pois os primeiros jogos eletrônicos surgiram com os rudimentos da tecnologia digital e eles evoluíram com ela.

2.3. Imagens em movimento

As exibições das imagens em movimento foram evoluindo através dos tempos, desde os primeiros experimentos no século XVII. Abaixo são mostrados os diversos dispositivos para imagem em movimento. Os dispositivos digitais vão ser discutidos a partir da próxima seção.

- Dispositivos pré-cinema:
 - Lanterna mágica, 1645;
 - Traumatoscópio, 1824;
 - Fenaquistoscópio 1828;
 - Estroboscópio 1828;
 - Zootrópio, 1834.
- Cinema mudo, em 1895, com a cadência de 18 quadros por segundo, no formato acadêmico, isto é, a largura da tela dividida pela sua altura é igual a $4/3 = 1,33$.
- Cinema sonoro, em 1927, a 24 quadros por segundo, com o fator de forma de 1,33, como no cinema mudo. A mudança, de 18 para 24 quadros por segundo, foi necessária para melhorar a sincronia do som com a imagem.
- Cinema colorido, em 1932. Na década de 50 começam os primeiros formatos widescreen, como o Cinemascope e o Cinerama.

- Televisão preto&branco, em 1934, nos Estados Unidos e na Inglaterra, e em 1950 no Brasil. Ela exibia as imagens nas cadências de 30 quadros por segundo ou 25 quadros por segundo, dependente da energia elétrica das tomadas, no formato acadêmico.
- TV a cabo, em 1948 nos Estados Unidos e em 1991 no Brasil.
- Televisão colorida, em 1954 nos Estados Unidos e em 1972 no Brasil. Foram definidos os formatos *National Television System Committee* (NTSC). Na brincadeira da época do lançamento, *Never the Same Color* e PAL-M (*PAL: Phase Alternate Line*), na cadência de 30 quadros por segundo e com 525 linhas de exibição. Na cadência de 25 quadros por segundo e 625 linhas de exibição, existem os sistemas PAL e *Séquentiel couleur à mémoire* (SECAM), sistema de tv francês.
- Primeiras imagens digitais para computador, em 1957. Primeiro filme com computação gráfica em 1982 (*Tron*, da Disney).
- Equipamentos de vídeo analógico portátil, surgindo com o primeiro equipamento da Sony, o *Portapak*, lançado em 1967.
- TV a cabo digital, no início dos anos 90 nos Estados Unidos e em 2004 no Brasil.
- World Wide Web, disponível para o grande público a partir de 1993 [W3C 2011].
- Vídeo digital, com a definição do formato DV, em 1995, e do DVD, em 1998.
- Televisão digital de alta definição, lançada em 1997 nos Estados Unidos e em 2006 no Brasil. Ela exibe com 1080 linhas, no sistema completo (*full*), ou 720 linhas.
- Cinema digital, com os formatos 2K (2048 por 1080 pixels) e 4K (4096 por 2160 pixels). O primeiro filme totalmente capturado com câmeras digitais foi o *Guerra nas Estrelas*, em 2003.
- Dispositivos móveis, que estão em plena evolução com os celulares e os aparelhos PDA (*personal digital assistant*). A Apple definiu o nome PDA e lançou o primeiro aparelho do tipo em 1992, o *Newton*.

2.4. Etapas para a formação da imagem digital

A construção da imagem digital envolve diversas etapas. A aquisição de imagens em movimento pode ser feita por câmeras digitais de vídeo ou de cinema, ou podem ser capturadas em película ou em vídeo analógico e depois digitalizadas, ou, então, elas podem ser geradas inteiramente por computador, através da computação gráfica. Essas maneiras de gerar a imagem digital podem ser resumidas da seguinte forma:

- Filmadoras digitais de vídeo ou cinema;
- Película e depois digitalizadas, pelo escaneamento das imagens ou pela telecinagem;
- Imagens em vídeo analógico e depois digitalizadas, através das placas de captura;
- Imagens escaneadas;
- Geradas totalmente em computador, pela computação gráfica.

Em uma primeira etapa, quando o sinal ainda é analógico, as imagens passam por uma etapa de pré-processamento, para amplificar e filtrar o sinal. Em seguida, esse sinal analógico passa por um conversor analógico-digital, passando pelos processos de amostragem e quantização, para fazer a sua representação e descrição digitais. Dependendo do tipo de aplicação da imagem digital, ela então passa por processos segmentação, de transformações geométricas, de realce e de compressão, dentre outros. Essas etapas podem ser resumidas como está a seguir [Pedrini e Schwartz 2008]:

- Pré-processamento;
- Representação e descrição;
- Dependente do que vai ser feito com a imagem:
 - Segmentação;
 - Transformações geométricas;
 - Realce;
- Compressão.

2.4.1. Pré-processamento

Na aquisição da imagem, ela é capturada através de um sensor, que a converte em uma representação digital adequada. Os tipos de dispositivos utilizados são o *Charged Coupled Device* (CCD) ou o *Complementary Metal Oxide Semiconductor* (CMOS), que são dispositivos de estado sólido. Eles são circuitos integrados compostos por elementos que geram uma corrente elétrica quando a luz incide sobre eles. Existem diferenças na maneira como a energia luminosa é transformada em energia elétrica e em como esse sinal é processado em cada um dos dispositivos.

O CMOS tem a corrente elétrica amplificada em cada uma das células fotoelétricas, enquanto o CCD faz esse processo em uma linha de células fotoelétricas. O sinal do CMOS é muito mais ruidoso que o sinal do CCD, mas avanços tecnológicos recentes permitiram reduzir esse ruído e surgir uma nova geração de câmeras e filmadoras [APTINA 2011]. O CCD pode trabalhar em intensidade de luz muito mais baixa que a do CMOS. A descrição da imagem dependerá do tipo de dispositivo, da iluminação, do número de células fotoelétricas e do número de níveis de cinza ou de cores, dentre outros fatores. A luz branca entra nos prismas e é separada nas componentes vermelha, verde e azul.

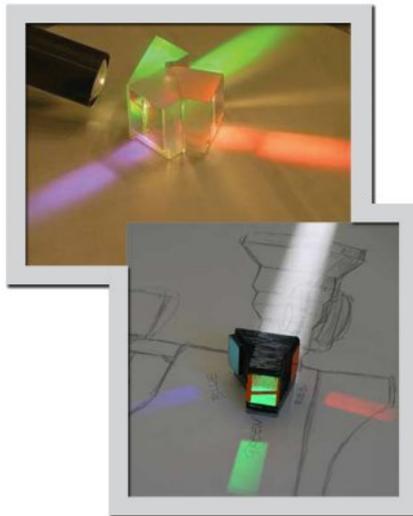


Imagem 39: separação da luz branca nas cores primárias por um prisma.

Os equipamentos de captura de imagem podem ter um ou três dispositivos, do tipo CCD ou CMOS. Na Imagem 39, se pode observar a separação das cores com um conjunto de prismas. Na Imagem 40, se tem um diagrama esquemático da separação das componentes de cores para uma câmera com três CCDs ou três CMOS. Toda a luz de cada uma das componentes de cor é capturada por cada um dos sensores.

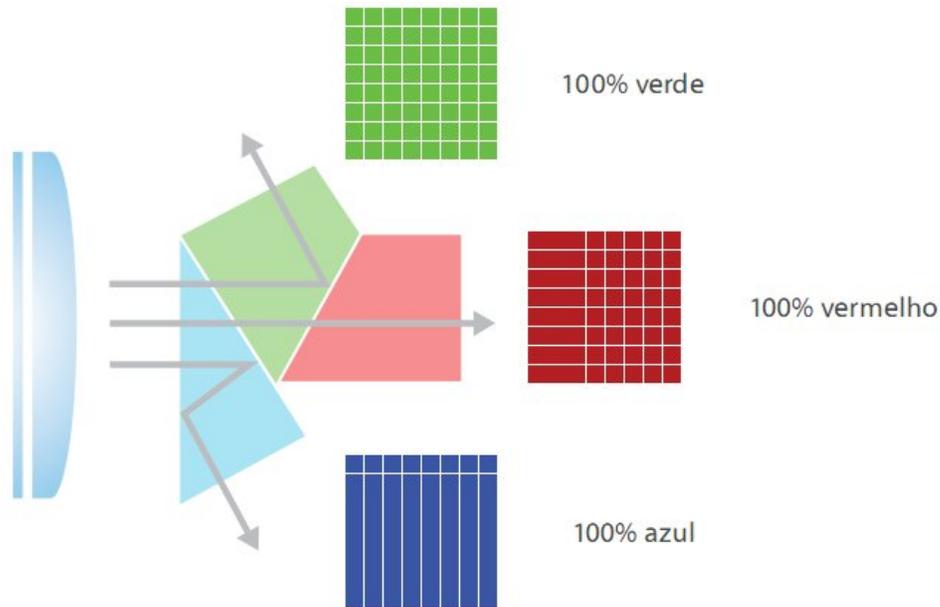


Imagem 40: diagrama esquemático mostrando a separação da luz pelo prisma após passar pela lente.

Na Imagem 41, se vê um prisma com os três CCDs acoplados. Na frente do prisma se observa a lente, que está nesse objeto em forma de anel. O sinal que sai de cada um dos CCDs é analógico, pois é uma corrente elétrica. Em seguida, a corrente é amplificada e filtrada, para reduzir o ruído. Somente depois disso o sinal analógico é convertido em sinal digital.

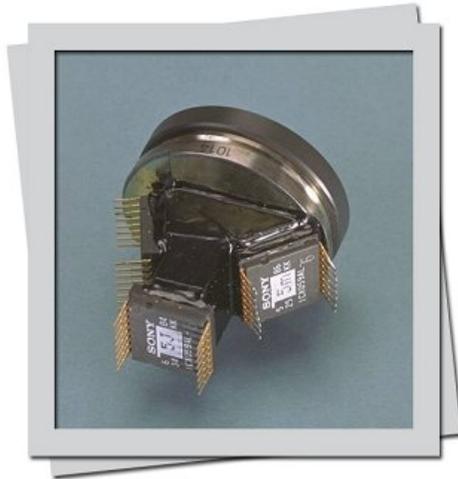


Imagem 41: prisma com os três CCDs acoplados.

Para um sensor apenas, a câmera utiliza um filtro na frente do dispositivo para separar as cores, como a malha de Bayer da Imagem 42. Ela é uma malha de filtros verdes, azuis e vermelhos, em geral na proporção de 50, 25 e 25%, respectivamente. Uma camada com lentes, que está na parte superior da Imagem 42, associadas à malha, completa o conjunto. Câmeras com apenas um dispositivo do tipo CMOS e o filtro de Bayer têm sido utilizadas com sucesso nos últimos anos. No ano de 2009, a maior parte dos Oscars foi vencido pela primeira vez por filmes que utilizaram esse tipo de captura de imagem, como *Quero Ser Milionário* e *O Curioso Caso de Benjamin Button*. Os Oscars de fotografia dos últimos anos foram para filmes que utilizaram filmadoras digitais com essa configuração [Academy of Motion Picture Arts and Sciences 2011].

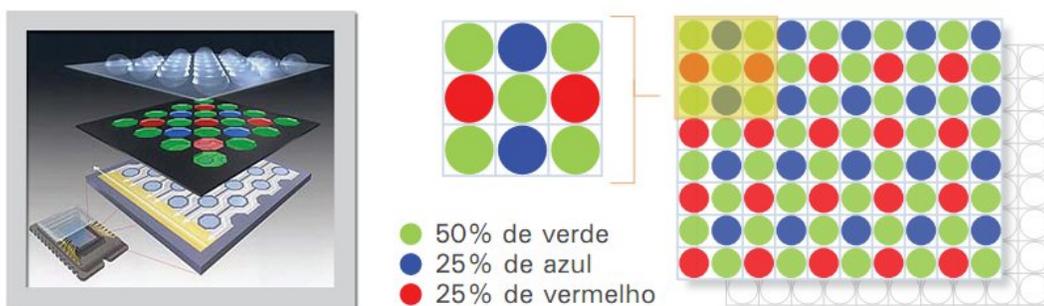


Imagem 42: diagrama com uma malha de Bayer, à esquerda, e a distribuição de filtros à direita.

2.4.2. Amostragem e quantização

Para ser representado digitalmente, o sinal da imagem gerado nos dispositivos é amostrado e quantizado [Pedrini e Schwartz 2008]. A amostragem é o processo de discretização do domínio de definição da imagem bidimensional nas duas direções ortogonais, x e y , para gerar uma matriz de $M \times N$ amostras. Cada elemento dessa matriz é chamado de *pixel* (*picture element*). Existe um teorema fundamental da teoria da informação, chamado de teorema de Nyquist–Shannon, que dá a frequência mínima para uma boa amostragem. Esse teorema diz que se um sinal tem uma frequência máxima, ele é completamente determinado se a frequência de amostragem é pelo menos o dobro dessa frequência máxima do sinal. Então no caso da imagem, o sinal deve ser amostrado quatro vezes, isto é, duas vezes na direção x e duas vezes na direção y .

Um exemplo prático e direto pode ser obtido no som. Os seres humanos captam os sons entre a frequência de 20 Hz (grave) e 20 kHz (agudo). O som é uma onda e o Hz é uma medida da frequência, representando 1 Hz = 1 ciclo por segundo. As mulheres podem perceber sons um pouco mais agudos que os homens, indo um pouco acima de 20 kHz, isto é, 20 mil ciclos de onda por segundo. Assim, a amostragem do som deve ser maior que 40 kHz e é o que ocorre. O som de CD é amostrado a 44,1 kHz e o áudio também pode ser amostrado a 48 kHz. Outra opção, para reduzir o tamanho do arquivo de som, é amostrá-lo a 32 kHz, perdendo assim muita informação na região dos sons agudos.

A quantização é o número inteiro de níveis de cada componente de cor, em uma imagem colorida, ou de níveis de cinza, em uma imagem monocromática, permitidos para cada pixel. Esse número inteiro de níveis é representado pelo número de bits. As imagens mais comumente encontradas, nas componentes RGB, têm 8 bits por cada uma das componentes, ou canais, de cor. Isso dá 256 valores de cor para o vermelho, 256 para o verde e 256 para o azul. A imagem resultante então tem 16 milhões e 700 mil valores diferentes de cor, ou 24 bits. Na Imagem 43, se encontra uma representação

da digitalização de um sinal, onde no eixo horizontal são mostrados os intervalos de amostragem e no eixo vertical os níveis de quantização.

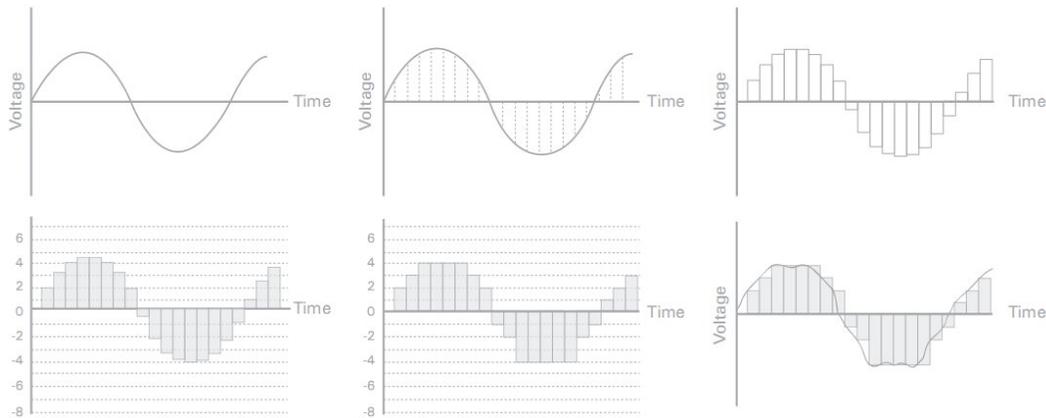


Imagem 43: digitalização de um sinal contínuo.

2.5. Imagem raster e imagem vetorial

A imagem digital pode ser exibida nos monitores, projetores, óculos de realidade virtual, revistas, jornais, outdoors e muitos outros dispositivos. Ela também pode ser gerada de diversas maneiras, como por câmeras digitais (de vídeo, de fotografia ou de cinema), scanners, programas de computação gráfica ou por placas de vídeo e programas que convertem a imagem analógica em digital. Existem, no entanto, apenas duas maneiras para construir e gerar a imagem digital computacionalmente. Elas são chamadas de imagem raster e imagem vetorial.

A imagem raster é a mais utilizada por causa do seu tipo de construção, que demanda menos tempo de processamento. Ela apresenta a sua estrutura de dados em um arranjo de pixels. Cada pixel tem um valor da imagem, de cor e intensidade de luz. Esse arranjo de pixels é chamado de *bit map* ou *pixel map*. Praticamente todos os equipamentos de captura e exibição de imagem são rasterizados, isto é, são estruturados como um arranjo de pixels, em geral na forma de uma matriz com pixels em linhas e colunas. A imagem vetorial é gerada por programas de modelagem, por exemplo, e ela tem que ser rasterizada antes de ser exibida. Isto será discutido nas próximas seções.

2.5.1. Imagem raster e acuidade visual

A imagem raster é formada como um mosaico de pixels, que podem ter diversos formatos, mas são em geral quadrados ou retangulares. Para que se possa perceber a imagem como se fosse contínua e não feita de elementos discretos, o tamanho do pixel deve enganar a percepção dos olhos. Isso é possível fazendo o seu tamanho ser pequeno o suficiente para que o olho não o possa distinguir. O olho humano tem um ângulo mínimo em que consegue separar dois pontos e isso determina a acuidade visual.

A acuidade visual é uma característica do olho de reconhecer dois pontos muito próximos. Ela é determinada pela menor imagem retiniana percebida pelo indivíduo. Sua medida é dada pela relação entre o tamanho do menor objeto visualizado e a distância entre observador e objeto. Na Imagem 44, é apresentado um gráfico com a faixa de percepção do olho e dos microscópios óptico e eletrônico.

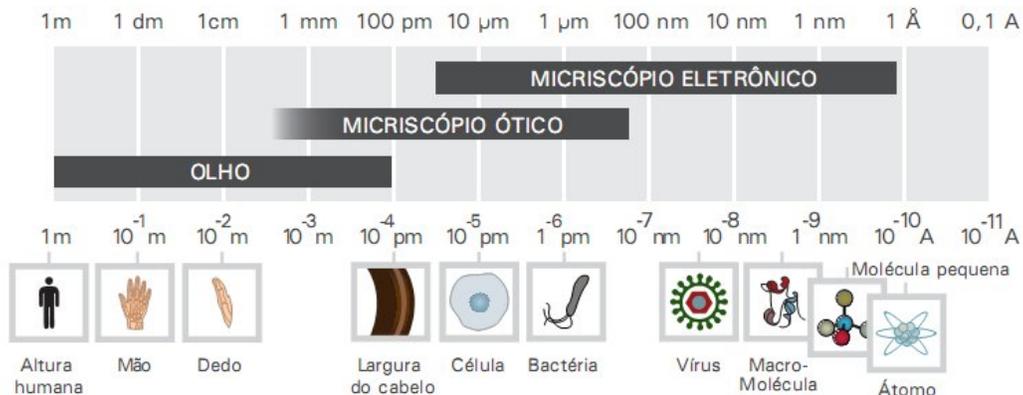


Imagem 44: Faixas de percepção dos objetos pelo olho e por dois tipos de microscópio.

A imagem raster é construída levando-se em consideração essa característica do olho. Quanto maior o número de pixels e menor o seu tamanho, maior pode ser a imagem sem que se perceba a estrutura de pixels. Quando uma imagem vai sendo ampliada, a sua estrutura de pixels começa a aparecer e a qualidade da imagem é deteriorada. O número de pixels de uma imagem é chamado de *resolução da imagem*. Na Imagem 45, se observa a

estrutura de pixels de uma imagem de um olho e, na Imagem 46, se tem uma imagem rasterizada preta e branca ampliada. Observa-se como a estrutura dos pixels é revelada com a ampliação da imagem.

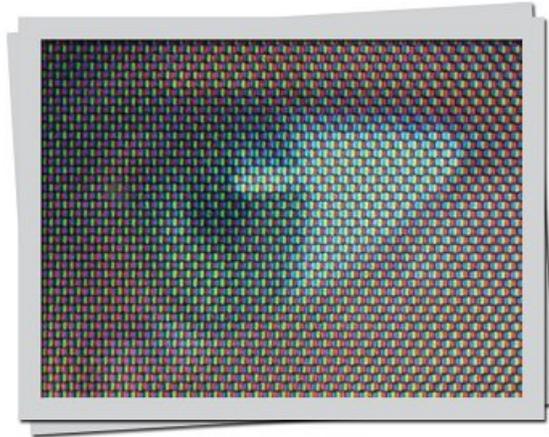


Imagem 45: imagem com os pixels aparentes.

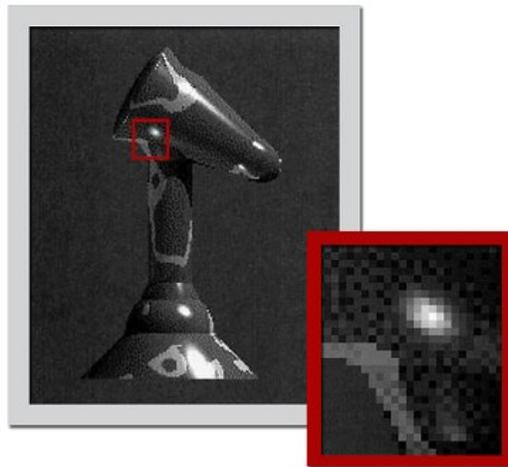


Imagem 46: imagem rasterizada depois de ser ampliada, revelando a estrutura de pixels.

Na imagem colorida, cada pixel é uma mistura das cores básicas. Quando a imagem é gerada em um monitor ou projetor, ela é a combinação das cores primárias, que se misturam aditivamente. Na Imagem 47, se observa a ampliação da estrutura de pixels de dois monitores de LCD. A cor resultante em cada pixel é a combinação da intensidade do vermelho, do verde e do azul.

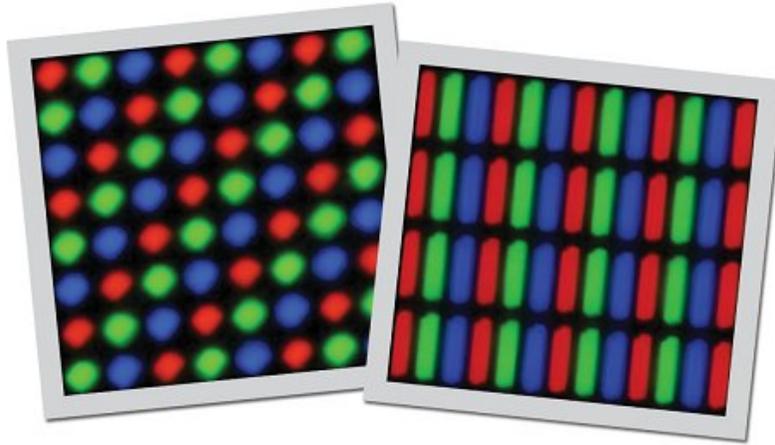


Imagem 47: estrutura dos pixels de dois monitores LCD, com as cores primárias.

Para as imagens impressas, a combinação é feita a partir das cores secundárias, já que a mistura de cores acontece subtrativamente. Cada pixel da imagem tem três pontos, cada um com a cor ciano, magenta e amarelo. A cor resultante é resultado da quantidade de cada cor secundária que entra na mistura. É dessa maneira que são impressas as imagens coloridas nas revistas, jornais, banners e outdoors. Na Imagem 48, se tem uma imagem sendo decomposta nas componentes CMY.



Imagem 48: componentes ciano, magenta e amarelo de uma fotografia.

2.5.2. Imagem vetorial

A imagem vetorial apresenta características distintas da imagem raster. Ela é construída através de curvas e superfícies definidas por equações matemáticas. Todos os softwares de modelagem tridimensional, como o 3DSMAX, o Blender3D, o Maya e o Lightwave, trabalham com a imagem vetorial o tempo todo. Qualquer imagem pode ser gerada vetorialmente e a facilidade em que isso é feito é cada vez maior, porque os programas de computação gráfica estão evoluindo e a cada dia fornecem ferramentas mais precisas e mais amigáveis para o usuário.

A imagem vetorial não perde qualidade quando é ampliada porque todos os seus dados são recalculados e exibidos enquanto o seu tamanho é alterado. Essa é uma grande vantagem da imagem vetorial com relação à rasterizada. A imagem vetorial, porém, para ser exibida demanda uma quantidade de processamento maior que a imagem raster. E existem poucos dispositivos vetoriais de exibição. Isso é devido à dificuldade de descrever as cores nesses dispositivos. A imagem vetorial é então rasterizada para ser exibida, através do render.

A imagem vetorial é gerada por meio de primitivas geométricas, que são os pontos, as linhas, as polilinhas, os polígonos e as curvas. A polilinha, por exemplo, é uma sequência de linhas retas conectadas através dos seus pontos extremos. Na Imagem 49, se observam algumas imagens construídas por polilinhas. Em cada um dos desenhos, a curva não parece ser gerada por retas. A extensão dessas retas é pequena e engana os olhos, pois o comprimento das linhas está no limite da acuidade visual. O desenho da direita é uma representação comum na computação gráfica, que é chamada de estrutura aramada ou *wireframe*. Ela é composta por retas conectadas e o objeto gráfico pode ser observado sem muito comprometimento de processamento. Esse processamento seria maior para exibir o volume do objeto com suas cores e iluminação. Na Imagem 50, se observa o detalhe que revela as retas que compõem a imagem.

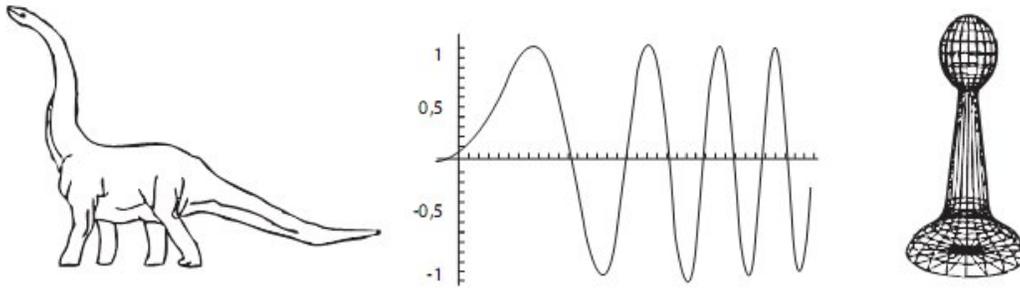


Imagem 49: três desenhos construídos por retas curtas, as polilinhas, que dão a impressão de curvas.

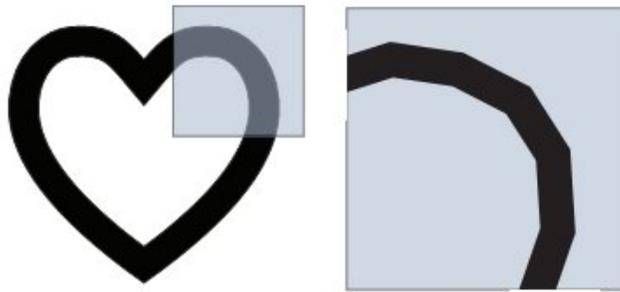


Imagem 50: detalhe de uma imagem vetorial exibindo as polilinhas.

Essas primitivas podem ter o seu preenchimento feito com uma cor ou uma textura. Os arquivos das imagens vetoriais armazenam as primitivas, as formas e as cores que as compõem como fórmulas matemáticas. No entanto, os dispositivos vetoriais de exibição apresentam problemas para mostrar as cores e as texturas e o tempo de processamento ainda é grande, comprometendo a qualidade da imagem exibida.

Na Imagem 51, são exibidas algumas formas geradas vetorialmente: com contorno (A), sem contorno (B) e com diversos tipos de preenchimento. Em (D) o preenchimento é feito através de operações de união e de intersecção dos polígonos gerados pelas retas.

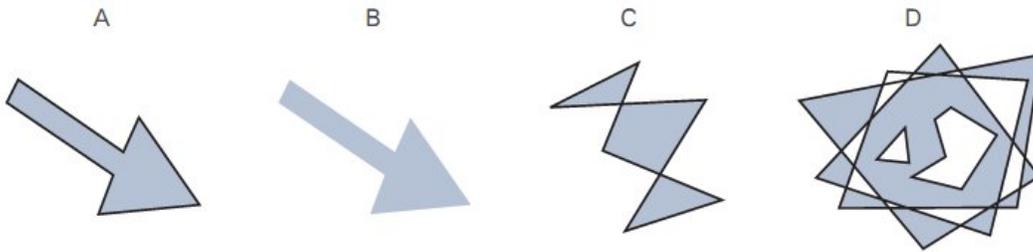


Imagem 51: figuras geradas por retas com preenchimento de cor.

Os dispositivos de exibição vetoriais constroem as imagens desenhando linhas e preenchendo as superfícies delimitadas por essas linhas. Existem poucos exemplos desse tipo de dispositivo em uso. O plotter, para impressão de peças gráficas, e que funciona muito bem para grandes impressões, como o outdoor, funciona vetorialmente, com os pigmentos sendo misturados por linhas.

No caso de um monitor de vídeo, a operação é mais complexa para fazer a combinação de cores juntamente com a construção da imagem. No entanto, temos alguns exemplos de monitores vetoriais, como nos aparelhos de controle de voo em aeroportos. Na Imagem 52, se tem as fotografias de dois dispositivos de exibição vetoriais, um plotter, à esquerda, e o monitor do jogo Asteroids, à direita. As telas de exibição de muitos dos primeiros jogos eram vetoriais.

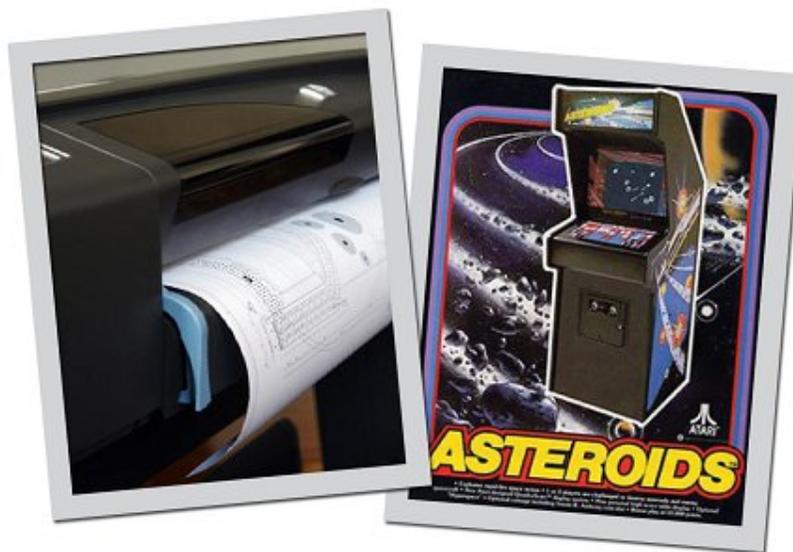


Imagem 52: dispositivos de exibição vetoriais: plotter e monitor de um dos primeiros jogos.

2.6. Resolução espacial, profundidade de cor e fator de forma

A qualidade da imagem digital depende da resolução espacial e da profundidade de cor. A resolução espacial está relacionada com a densidade de pixels da imagem, ou seja, com a amostragem. A profundidade de cor está associada ao número de níveis de quantização da imagem. O número de níveis de quantização é uma potência de 2, dependendo do número de bits. Para cada canal de cor ela é igual a 8 para a televisão digital de definição *standard* (SDTV) e 8 ou 10 para a televisão digital de alta definição (HDTV). O cinema digital tem oferecido a possibilidade de se trabalhar com 12 bits por canal de cor. Vale aqui acrescentar que devido ao número de cores possíveis e às limitações de processamento das filmadoras e dos equipamentos que trabalham com a imagem atualmente, esquemas que já haviam sido usadas em situações semelhantes - na década de 1980, quando as imagens tinham 256 cores no total apenas, mas com processamento limitado – estão sendo usados por aqueles que trabalham com o cinema digital novamente, que são os *LookUp Table* (LUT). Nesses esquemas, o número total de cores é aproximado por valores dados em uma matriz. O Cineon é um exemplo de LUT. A discussão sobre este tema foge do escopo desta tese, apesar do interesse crescente que ele vem tendo [Arri 2011].

A cor de uma imagem é representada pelo espaço de cores, como foi discutido anteriormente. A imagem digital em movimento é mais comumente descrita através dos valores de luminância, ciano (ou todas as cores menos o vermelho) e amarelo (ou todas as cores menos o azul) em cada pixel, o chamado formato Y:U:V. De acordo com a taxa de amostragem em cada pixel, temos, em geral, formatos com 4:4:4, 4:2:2, 4:2:0 e 4:2:1. Nesses casos, o primeiro valor é a taxa de amostragem da luminância, ou brilho, o segundo do ciano e o terceiro do amarelo. As informações de cor são chamadas de sub-amostragem, por diminuir a quantidade de informações de cores [DV 2011].

Além dos conceitos acima descritos, existe a razão de aspecto, ou fator de forma, da imagem, que é a razão entre sua largura e sua altura. Esse fator de forma é de 1,33 para o SDTV e 1,78 para o HDTV. Para o cinema em

película, as razões de aspecto mais comuns são o cinemascope, 2,35, cinerama, 2,59, super 35, 1,74, 70 mm, 2,21, super 16, 1,67, e super 8, 1,36.

2.7. Cadência das imagens

A imagem em movimento possui uma cadência de exibição de quadros (*frames*) por segundo (*frame rate*). O cinema trabalha com 24 quadros por segundo e a televisão e o vídeo analógicos com 25 ou 30, se estiver no sistema PAL ou NTSC e PAL-M, respectivamente. No caso do vídeo, cada quadro pode ser dividido em dois campos (*fields*). Cada campo faz a varredura das linhas horizontais pares ou ímpares e elas são exibidas no formato entrelaçado (*interlaced*), para o vídeo analógico, onde os dois campos formam o quadro, ou frame, da imagem. A imagem pode também ser exibida com o quadro sem divisão por campos, no caso do vídeo progressivo. Esse é o caso do cinema digital e da exibição nos monitores de computador. O vídeo digital pode utilizar quadros entrelaçados (i) ou progressivos (p), nas cadências 24p, 25p, 30p, 50i ou 60i.

As imagens entrelaçadas exibem um fantasma quando são paradas, isto é, quando se exhibe apenas um *frame*. Quando elas estão sendo tocadas, em *play*, isso desaparece. Isso acontece quando o que está sendo capturado se move com muita velocidade ou a câmera se move muito rápido. Isso é natural, pois primeiro se grava um *field* e depois outro. Na Imagem 53 se pode observar esse fenômeno: à esquerda se tem um *zoom* rápido da imagem e à direita ela está sem movimento de câmera ou de lente [Reinhardt 2007].

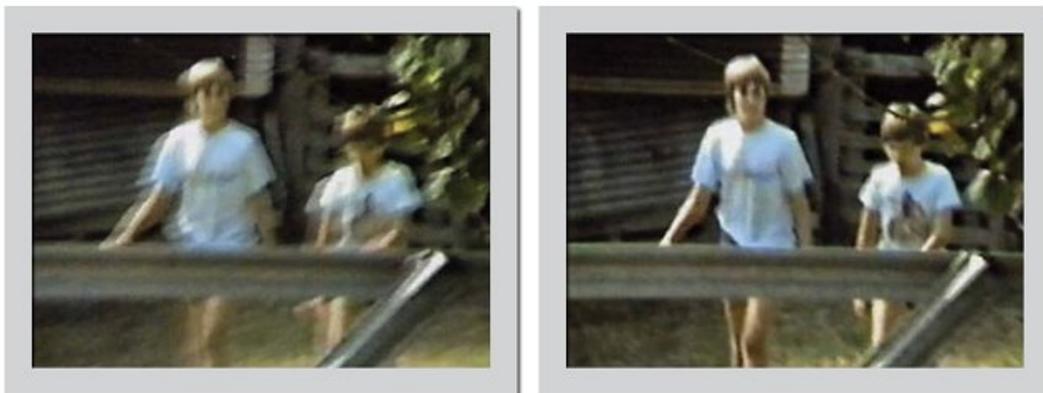


Imagem 53: frame entrelaçado, com os *fields* visíveis à esquerda e com movimentos lentos à direita, e os campos menos visíveis para o quadro com menor movimento [Reinhardt 2007].

A inversão da ordem dos campos gera o efeito de serrilhamento das bordas da imagem. As câmeras e equipamentos de pós-produção permitem configurar qual o campo que será gravado primeiro. Se houver a troca dessa ordem nos equipamentos, o campo é invertido. Em cada quadro, a sequência temporal dos fields está trocada, provocando o efeito de serrilhamento. Este problema é muito comum de ser observado nos comerciais de televisão ou nas vinhetas da programação. Esse defeito pode ser observado visualmente. Se o problema está ocorrendo, os campos devem ser invertidos. Artefatos também podem ser introduzidos ao se confundir e trabalhar com vídeo entrelaçado como se fossem progressivos e vice-versa. Os equipamentos permitem escolher qual o tipo de vídeo e muitas vezes esta escolha é feita equivocadamente. Um teste com diferentes configurações permite determinar qual o tipo de vídeo, caso não se tenha certeza dessa informação.

O vídeo vem com a informação da posição no tempo de cada quadro de imagem. Esta informação é chamada de *timecode*. Ela é representada por uma sequência de números, com os dados de horas, minutos, segundos e frames. O timecode permite localizar precisamente qualquer quadro do vídeo. O vídeo do tipo NTSC colorido tem uma particularidade devido a sua evolução tecnológica. Ele tem uma taxa de frames de 29,97 frames por segundo (fps), e não 30. Sendo assim, ocorre uma discrepância entre o número de frames e o timecode. Como a taxa de quadros por segundo não é 30, após uma hora, por exemplo, a contagem de quadros é de 107.892 e não 108.000 quadros. A solução encontrada para igualar as duas contagens é chamada drop frame ou non drop frame. Na primeira, a cada minuto são tirados dois números do timecode, exceto no décimo minuto. No segundo caso, para cada quadro do vídeo existe um timecode sequencialmente. Essas diferenças podem gerar problemas de sincronização entre o áudio e o vídeo e se deve entender precisamente com qual tipo de vídeo se está trabalhando.

2.8. Pulldown

A televisão sempre exibiu filmes na sua programação. No entanto, as cadências do cinema e da televisão são diferentes. No caso do Brasil, se deve transformar os 24 fps do cinema em 30 fps da televisão. Para a televisão entrelaçada, o problema é ainda maior, pois são 60 campos de imagem. A solução para este problema foi encontrado, após diversas outras propostas, através do telecine e o processo de *pull-down*. No pull-down, a velocidade do filme é reduzida de 1/1000, o que é imperceptível para o espectador. Desta maneira, a razão entre a cadência do filme, que fica em 24,976 fps, e do vídeo, que está a 29,97 fps, é igual a 4/5. Assim, os frames do filme devem ser distribuídos pelos campos do vídeo.

Para fazer isso, utiliza-se o esquema ilustrado na Imagem 54.

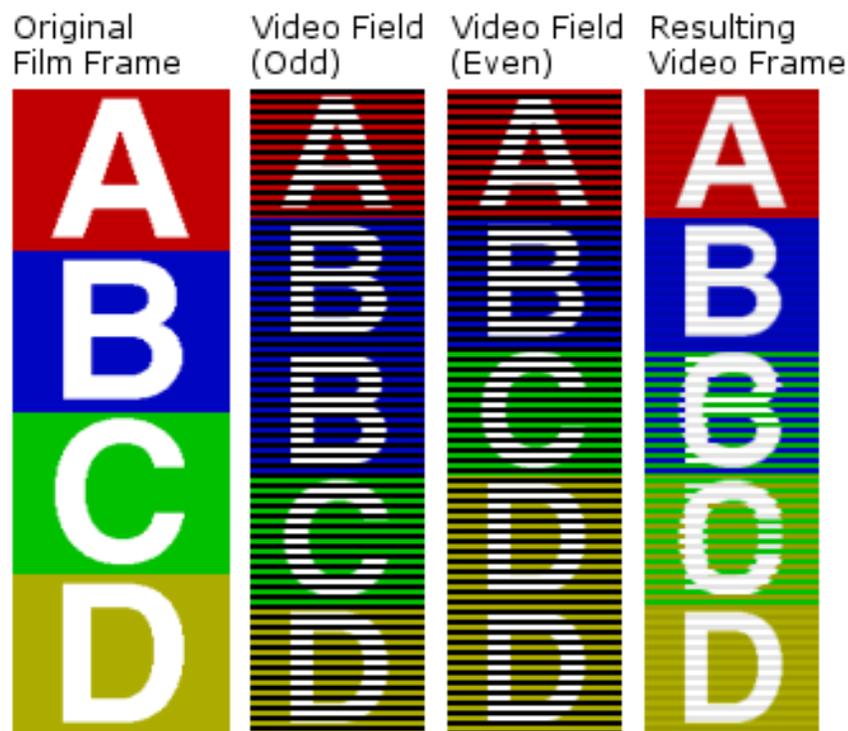


Imagem 54: esquema de pull-down [Lee 2011].

Nesse esquema, se nota que existem quadros com frames “puros” do filme e quadros com frames misturados. Em um dos processos de montagem de cinema, o filme é telecinado e é editado offline em uma ilha não-linear,

gerando uma lista de corte com as marcações dos quadros originais do filme, apesar de ser cortado como vídeo. Deve-se ter atenção com relação a esses frames do vídeo que correspondem a dois frames do filme.

As filmadoras e câmeras de vídeo digitais também permitem gravar em diferentes cadências. Muitas vezes, tem-se optado por gravar a 24 fps, para depois haver o processo de pulldown para ser exibido na televisão. Isto tem gerado uma série de artefatos, quando o processo não é feito de forma cuidadosa e metódica. Nunca se viu tantos problemas com artefatos como atualmente na televisão brasileira, seja por problemas de sequenciamento dos fields, seja por problemas de pulldown ou de compressão da imagem, que será discutida na próxima seção.

2.9. Codificação e decodificação da imagem (Codec)

Para se adequar ao espaço de armazenamento em discos e fitas, à quantidade de processamento de dados e ao tempo de transmissão de imagens, surgiram diversas técnicas de compressão de imagens. Com isso se reduziu a quantidade de bytes para descrever a imagem e se pode transferir uma quantidade maior de dados com a mesma taxa de transferência de bits (*bitrate*). Os formatos de compressão e descompressão são chamados de *codecs*. Eles podem ser classificados em compressão com perda e sem perda [Gonzalez e Woods 2010].

A compressão sem perdas baseia-se unicamente na redução de redundância estatística. A imagem resultante, após a descompressão, é exatamente igual à imagem original. A compressão com perdas baseia-se na redução de redundância estatística e na eliminação de informação perceptivelmente pouco relevante. O sistema visual humano é mais sensível a variações de luminância que de cor. Dessa maneira, a luminância é representada de forma mais completa que a crominância na compressão com perdas. A compressão pode ser do tipo espacial ou temporal. A compressão pode ser feita utilizando a transformação discreta de cosseno (DCT), *wavelets* ou outras técnicas.

Para manipular digitalmente uma imagem, deve-se conhecer como ela está representada e codificada. Trabalhar com a imagem digital envolve o processo de rapidamente decodificar a imagem, manipulá-la e codificá-la novamente. O resultado desse processo pode ser familiar para o espectador, mas os conceitos técnicos envolvidos na sua construção são fundamentais para quem trabalha com ela. De acordo com o tipo de imagem e o processo de codificação e decodificação, pode-se prever até que ponto o resultado será aquele que se planejou.

Os hardwares e softwares que lidam com a imagem em movimento oferecem a possibilidade de se trabalhar com diferentes níveis de qualidade de imagem, dependentes de seus recursos intrínsecos. A maneira como o vídeo é descrito digitalmente também determina o que se pode fazer com ele. Quem manipula as imagens deve conhecer esses aspectos para realizar artisticamente qualquer tipo de transformação da imagem, seja uma correção de cor, um *morphing*, uma mudança temporal, ou uma composição de imagens, através de rastreamento, *mattes*, máscaras ou outras técnicas. Em situações corriqueiras, podem-se gerar artefatos na imagem que não estavam previstos e inviabilizar o trabalho.

Na maioria das situações, as imagens são comprimidas para que os computadores e suas placas de vídeo possam manipulá-las. A compressão dos dados gerados pela imagem em movimento pode ser obtida por métodos baseados apenas em software ou em uma combinação de hardware e software. A vantagem de se construir uma compressão baseada apenas no software é a de que não se precisa planejar e implementar o hardware para essa função, mas as desvantagens são economicamente maiores, devido ao tempo de execução do trabalho e à qualidade do resultado, por exemplo. Os métodos mais poderosos de compressão são obtidos por uma combinação de hardware e software. Diversos formatos de imagem em movimento foram criados para lidar com isso.

As técnicas de compressão por software são aplicadas à imagem na medida em que elas são digitalizadas. Dois tipos principais de compressão são empregados: diminuir o tamanho da matriz de pixels da imagem, reduzindo a resolução espacial, ou diminuir a quantidade de informação de cores, a chamada sub-amostragem (*subsampling*) de cores. A compressão baseada em

hardware e software permite uma maior quantidade de instruções serem realizadas por segundo. Isso resulta em mais tempo para analisar um quadro de imagem, examinando suas redundâncias e atribuindo valores numéricos, antes que o próximo quadro seja processado. Além disso, decisões são tomadas baseadas em como o sistema visual humano processa a informação.

Dois sistemas importantes de compressão são apoiados em hardware e software, o *Joint Photographic Experts Group* (JPEG) e o *Motion Pictures Experts Group* (MPEG) [Gonzalez e Woods 2010]. A proposta do MPEG, diferentemente do JPEG, é utilizada apenas para imagens em movimento e áudio e foi criada por um comitê de indústrias em 1988. Ela começou a definir formatos em 1990, baseada em larguras de banda de transmissão. A primeira proposta, MPEG-1, atendeu inicialmente às necessidades do suporte do tipo CD-ROM, com uma limitação de largura de banda de 150 kB/seg. Várias camadas de padrões são definidas para cada proposta MPEG. O MPEG-1 camada 3, por exemplo, para o áudio, gerou o MP3. O MPEG-1 camada 1 viabilizou o VCD. O MPEG-2 tornou o DVD um formato comercial. O MPEG-4 tem sido uma evolução importante do MPEG 1, permitindo uma qualidade superior em arquivos de imagem e som com menos bytes por segundo, através do MP4. Alguns codecs mais conhecidos, com algumas aplicações, são mostrados na tabela da Imagem 55. Na Imagem 56, se tem um quadro comparativo dos bitrates de diversas variações do MPEG.

	type	parameters	used for	examples
DCT based	MPEG-1	< 1,5 Mbps	CDs, Video-CD Internet	Audio: MP2 (MUSICAM, DAB) Audio: MP3 (Internet)
	MPEG-2	1 ... 100 Mbps scalable	DVD, Digital TV (DVB, ATSC)	4:2:2 HP@ML
	MPEG-4	Audio Visual Objects (AVOs) higher compression ratio	Internet - TV	Media Player, Real Player Virtual Studios
	MPEG-7	Content Management Semantic of Scenes	archives retrieval systems (no compression)	Meta Data
	DVCPPro	25, 50, 100 mbps (HD)	Panasonic	Electronic News Gathering (ENG)
	H.261	64 .. 368 Kbps	video conferencing	ISDN nx64K mode
non DCT	Wavelets	0 ... 100 Mbps	video surveillance space missions research	JPEG 200

Imagem 55: comparações entre os diversos formatos de vídeo.



Imagem 56: variações do codec MPEG e bitrates.

O método MPEG se baseia em uma compressão *interframe* e *intraframe*, em que os dados são armazenados. Dessa maneira, alguns quadros MPEG precisam de quadros co-dependentes para serem exibidos. O método JPEG utiliza apenas a compressão *intraframe*. Os quadros (*frames*) MPEG são uma série de quadros chamados de *I*, *P* e *B*. O quadro *I* (*intra coded image*) é codificado *intraframe*, do tipo JPEG, e seus dados são independentes dos

outros quadros. *P* (*predicted image*) é um quadro predito, isto é, os dados são preditos de um quadro anterior *intraframe* ou de um predito. *B* é o frame bi-direcional (*bi-directionally interpolated image*) onde os dados são interpolados dos quadros *I* e *P* mais próximos.

Quando o processo de compressão MPEG tem início, um quadro é codificado. Este quadro do tipo *I* é do mesmo modelo usado na compressão JPEG. Os frames *P* são criados baseados no código preditivo. Um frame *I* é criado e, imediatamente, um frame *P* é predito. Mais frames *P* são construídos até ser o momento para criar um novo frame *I*. Entre os frames *I* e *P* existem frames do tipo *B*. O quadro *I* é criado, em geral, quando há mudança de movimento nos pixels nos frames que serão comprimidos. No entanto, no MPEG aproximadamente a cada meio segundo há um novo frame *I*. Com os frames *P* e *B* já há uma diminuição dos dados utilizados, já que somente o frame *I* possui os dados completos. Na Imagem 57, se vê um diagrama dos diferentes tipos de frames na compressão MPEG.

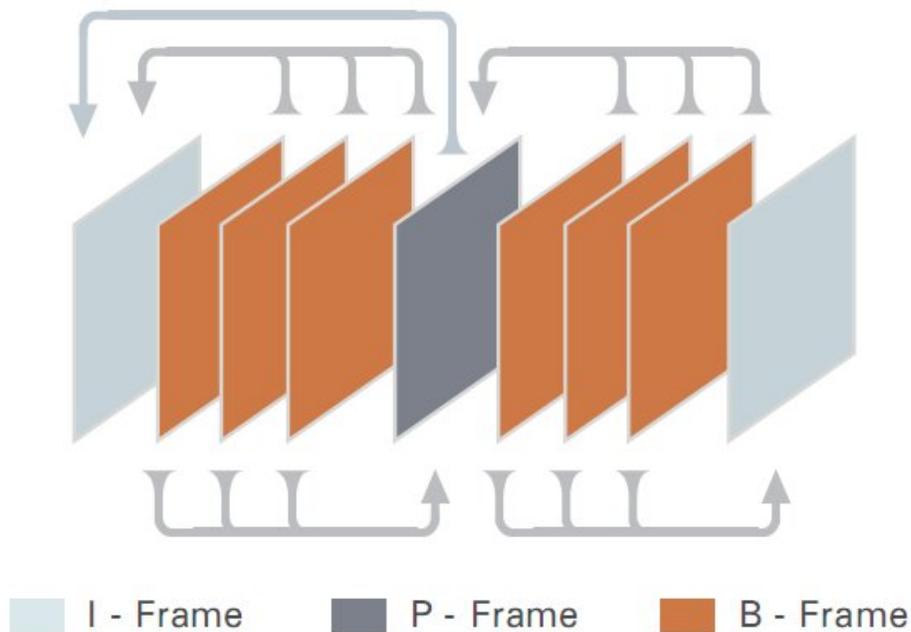


Imagem 57: representação esquemática dos frames I, P e B da compressão MPEG.

2.10. Formatos de imagens digitais em movimento

A imagem digital em movimento possui uma ampla gama de formatos, seja para a televisão, o cinema ou as novas mídias. Novos formatos têm se somado aos existentes, como o HDV (*high-definition vídeo*), de 2003, e outros estão por ser estabelecido definitivamente, como o intervalo 2K a 4K para o cinema. Eles são definidos a partir de acordos da indústria, para padronizar a produção de equipamentos de aquisição, processamento e exibição de imagens, gerando a melhor relação custo/benefício. Cada formato é baseado em uma definição de taxa de compressão, resolução espacial, razão de aspecto, amostragem, profundidade de cor e *bitrate*. O formato determina os tipos de efeitos visuais e tratamento digital de imagens possíveis.

A televisão digital de definição padrão (SDTV) possui a razão de aspecto 4:3, profundidade de cor de 8 bits e diversas resoluções espaciais e de amostragem. A televisão digital de alta definição (HDTV) tem a razão de aspecto 16:9 e algumas resoluções espaciais, com profundidade de cor de 8 ou 10 bits por canal de cor. O cinema digital pode ter a resolução de 2048x1080 pixels, no formato digital 2K, de 4096x2160 pixels, no formato 4K, ou valores intermediários, com 8, 10 ou 12 bits de profundidade por canal de cor. Na Imagem 58, se pode comparar as diferentes resoluções e fatores de forma dos padrões de televisão digital. Para a imagem HDT *full*, há uma linha cortando verticalmente, à esquerda, mostrando como a imagem seria exibida se o fator de forma, ou a razão de aspecto, fosse igual ao formato acadêmico, isto é, quatro larguras para três alturas.

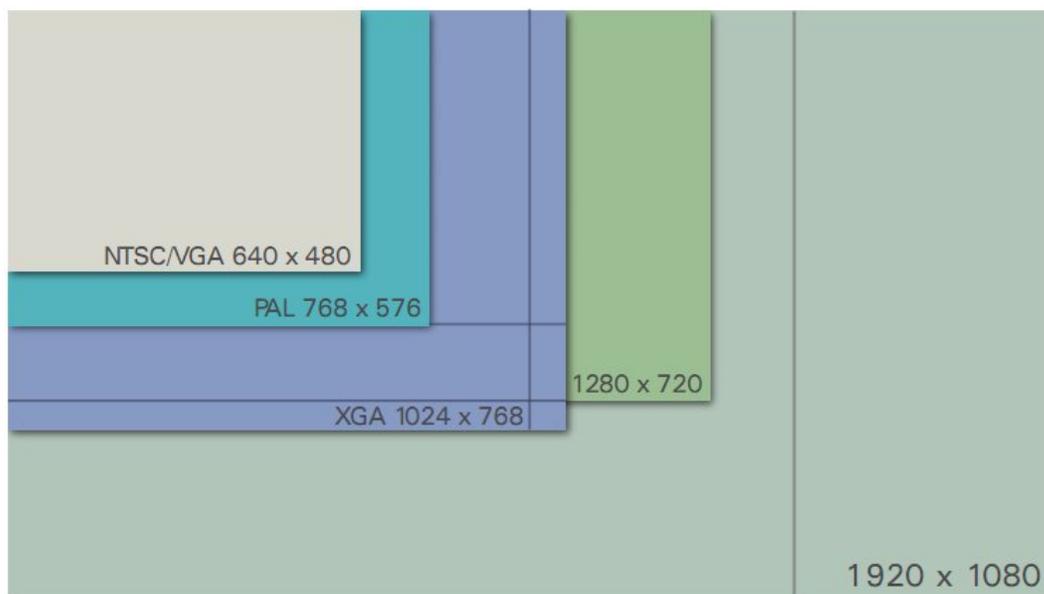


Imagem 58: diversas resoluções de televisão, SDTV e HDTV. 1920x1080 pixels é a *full* HDTV.

É importante frisar que, em 2006, o Brasil adotou o sistema de televisão digital japonês, o *Integrated Services Digital Broadcasting* (ISDB). No mundo atual existem o sistema norte-americano, chamado de *Advanced Television Systems Committee* (ATSC), e o europeu, o *Digital Video Broadcasting* (DVB).

O SDTV, para o NTSC e o PAL-M, tem a resolução de 640x480 pixels ou 720x480 pixels, com frames progressivos ou entrelaçados. A seguir tem-se alguns formatos de vídeo SD, com a respectiva sub-amostragem e fabricantes:

- XDCAM (4:2:2) (Sony);
- DVCAM (720x576), 4:2:0 (PAL), 720x480, 4:1:1 (NTSC) (Sony);
- DVCPRO (4:1:1) (Panasonic);
- DVCPRO50 (4:2:2) (Panasonic);
- D-9 ou Digital-S (JVC);
- DV e mini-DV (4:1:1 e compressão de 5:1).

O formato HD tem a resolução de 1280 por 720 pixels (720p) ou de 1920 por 1080 pixels (1080i ou 1080p) no HDTV *full*. No caso da HDTV norte-americana, as redes de esporte adotaram 720 p (progressivo), pois os espectadores gostam de pausar e fazer *slow motion*, e um frame tem toda a

resolução vertical. Como novelas ou filmes não são vistos em *slow motion*, as redes exibem em geral a 1080i (entrelaçado). Alguns formatos de vídeo HD estão a seguir:

- HDCAM e HDCAM SR da Sony, que foi um dos primeiros formatos comerciais de HD, com amostragem de 4:4:4 para a HDCAM SR e *bitrate* de 880 Mbit/s;
- DVCPPro-HD/ DV100/ D7-HD da Panasonic. O DVCPPro-HD tem a resolução de 720p ou 1080i para uma variedade de *frame rates*. Para 1080i somente 1280 pixels horizontais são gravados, gerando perda de resolução.

Um número cada vez maior de modelos de câmeras está armazenando as imagens em cartões de memória ou em discos rígidos. As fitas magnéticas devem provavelmente ser abandonadas ao longo dos próximos anos.

CAPÍTULO 3

PÓS-PRODUÇÃO DIGITAL

3.1. Pós-produção no audiovisual

A pós-produção é uma das etapas da realização de uma obra audiovisual e existe desde o início do cinema. Ela começa efetivamente após a filmagem ou gravação das imagens e do som, e compreende a edição ou montagem, a composição digital, a sonorização, os efeitos visuais e a computação gráfica, a colorização e a finalização do audiovisual. Porém, em muitos casos, o pessoal de pós-produção participa e é consultado em todas as etapas da realização do audiovisual, seja no momento em que está sendo escrito o roteiro, na pré-produção e na filmagem ou gravação das cenas. Em geral, no Brasil, ela se restringe exclusivamente à edição de imagens e do som, mas deveria ir muito além disso, sabendo-se dos recursos disponíveis atualmente em qualquer ilha de edição. Essa ausência de um cuidado maior com a pós-produção pode ser notada na produção cinematográfica, na teledramaturgia, na publicidade e na própria internet. A situação tem se agravado nos últimos anos, com a disponibilização de diversos formatos digitais, com diferentes cadências de imagem e com os recursos disponíveis para o processamento das imagens.

Existe uma confusão generalizada com relação ao tema “pós-produção e efeitos visuais”, em especial na cultura brasileira, e sobre a sua prática e as suas possibilidades no auxílio à narrativa. A razão dessa situação deve ser analisada também para apresentar soluções através do debate sobre os diversos tipos de recursos e de sua utilização nas mais variadas circunstâncias. Muitas vezes esse tema assume o papel de um assunto tabu, onde as posições são antagônicas e apaixonadas, seja considerando sua simples utilização como pertencente a uma filmografia e videografia mais tecnológica e, segundo os defensores desse ponto de vista, com menos requinte artístico, seja encarando-a com completo deslumbramento e sem qualquer tipo de crítica com relação à qualidade artística e técnica do que está sendo realizado.

Como bem observa [Lucena Júnior 2008],

A arte se fundamenta na técnica. A operação da arte é uma operação da técnica. Mas a arte também opera uma linguagem – que é lícito pensar deva ter surgido com os instrumentos. O trabalho de arte, portanto, envolve uma seção operacional e uma seção expressiva, de tal maneira interligadas a ponto de existir uma fusão dessas instâncias numa complementaridade de interesses indissociável em que a arte é enriquecida pela sutil exploração da técnica.

Na realização de um trabalho, também deve ser estudada a experiência subjetiva da percepção do trabalho com a imagem e o som. Em narrações lúdicas, parece existir um acordo implícito com o espectador de tal maneira que, por mais irreal que o efeito visual pareça, ele é incorporado naturalmente no contexto da história e a imagem modificada é assimilada sem qualquer problema. Por outro lado, em certos tipos de filme a plateia exige que a imagem trabalhada seja o mais próxima de um conceito do que seria real. O efeito, assim, é bom quando não é percebido como tal.

A pós-produção funciona como instrumento de apoio para viabilizar narrativas desde as primeiras obras cinematográficas. Utilizava-se então exclusivamente trucagens ópticas para imprimir o efeito na película. Com o advento da imagem e do som digitais, novas técnicas foram inventadas e todas as outras adaptadas para essa nova forma de manipulação do audiovisual, tornando mais amplas as soluções para realizar imageticamente aquilo que é criado. O efeito visual, por exemplo, permite a transformação visual da cena, ao integrar cenas gravadas com trucagens, introduzindo novas texturas na imagem, utilizando os mattes, personagens e cenários virtuais. Pode-se também corrigir com rapidez erros de gravação de diversos tipos.

O aumento da capacidade de processamento dos computadores pessoais, a disponibilização de ferramentas sofisticadas em softwares de baixo custo e as câmeras de vídeo e cinema digital, com qualidade profissional e preços mais acessíveis, permitiram que o universo da utilização da pós-produção e dos efeitos visuais se expandisse muito além dos limites dos grandes estúdios de cinema, casas de finalização e redes de televisão. Contar uma história com o auxílio desses recursos possibilita reduções drásticas de custo e aumento dos recursos disponíveis para a criação, sendo então

necessário o conhecimento das técnicas e das ferramentas disponíveis para solucionar os desafios. Isso se torna mais importante ainda atualmente, com o surgimento de novas mídias, como a internet, os aparelhos celulares e os assistentes pessoais digitais, e das câmeras de cinema digital, como a RED, a SI-2K, a Arri D-21 e a Thomsom Viper. A explosão de veiculação de vídeos na internet e os Oscars dos últimos anos comprovam isso. A importância do uso de recursos de pós-produção será ainda maior com o aumento da velocidade das conexões, trazendo a possibilidade de transmissão de imagens e som de maior resolução e melhor qualidade, e o estabelecimento do protocolo de televisão para internet (IPTV) no Brasil.

3.2. As ferramentas para a pós-produção

Os programas e equipamentos que estão disponíveis para aquele que vai fazer um trabalho de pós-produção apresentam um custo relativamente acessível e qualidade profissional. Uma estação de trabalho bem estruturada, na faixa de quinze mil reais hoje, realiza o que somente equipamentos de um milhão de dólares há dez anos atrás.

Para que sejam utilizados plenamente os recursos disponíveis, é importante que o usuário conheça bem os conceitos de imagem e som digitais e as diversas aplicações dos recursos. Atualmente se pode utilizar o mesmo equipamento para produzir e finalizar o audiovisual para as diversas mídias, do cinema digital aos aparelhos celulares, a partir de apenas um tipo de instrumento de captura de imagem e som. Sabendo, por exemplo, que o cinema digital 2K tem o formato de 2048X1080 pixels, uma câmera de HDTV pode capturar a imagem e, na pós-produção, ser transformada para o formato de cinema digital 2K, como é feito em diversos filmes, como *Jornada nas Estrelas*, de 2003, e *Os Normais*.

Essa imagem pode também ser convertida para qualquer formato de aparelho pessoal digital ou móvel. Assim, o finalizador deve conhecer como migrar para os diversos formatos, como na tabela da Imagem 59, que mostra o cenário em 2007. Atualmente, diversos celulares já vêm com filmadoras de HD.

Tipo de cenário	Resolução/ Frame rate	Taxas de transferência
Conteúdo de celulares	nHD (640x360) ou VGA (640x480), 30 frames/s	1-2 Mbps
Internet e SD	640X480, 24-30 fps	1-2 Mbps
HD	1280X720, 24-30 fps	5-6 Mbps
HD completa	1920X1080, 24p-30 fps	7-8 Mbps

Imagem 59: cenários de imagens em movimento para as diversas mídias.

O usuário deve ter em mente o tipo de aplicação e a imagem e o som a serem trabalhados. Precisa saber o que significa trabalhar com vídeo analógico composto, s-vídeo ou componente, com o vídeo digital em seus diversos formatos, com os diferentes formatos de áudio. Os processadores e as placas de vídeo e de áudio permitem que se utilizem todos os formatos, dependendo da configuração do equipamento. Placas de vídeo profissionais são fornecidas pelos fabricantes mais conhecidos, como Aja, Blackmagic Design, Matrox e Pinnacle. Os processadores e os sistemas operacionais determinam os tipos de placa de vídeo que são compatíveis. Entre os principais sistemas operacionais temos as diversas versões do Windows, para PC, o MacOSX, para os computadores da Apple, o Linux, para todos os tipos de computador, e o Irix, da Silicon Graphics.

Os softwares de edição não-linear também fazem composição digital, aplicação de efeitos visuais e sonoros sofisticados, como o rastreamento de pontos na imagem, equalização de áudio e colorização, dentre muitos outros recursos. Entre tantos programas de edição disponíveis no mercado, se pode citar os seguintes:

- Smoke e Fire, da Discreet Logic, hoje Autodesk;
- iQ, eQ e sQ, que substituíram o Editbox, da Quantel;
- a família Avid, fundadora do padrão para quase todas as ilhas de edição não-lineares, da Avid Technologies;
- Final Cut Pro HD, da Apple;

- Premiere Pro, da Adobe;
- Vegas Pro e XPRi NS, da Sony;
- .Lightworks, que ganhou alguns Oscars, e que agora abriu o seu código para download gratuito.

Para composição gráfica, colorização e efeitos visuais, se tem:

- a família Inferno, Flame e Flint, e o popular Combustion, da Discreet Logic, hoje Autodesk;
- Symphony, da Avid;
- Paintbox, da Quantel;
- o versátil After Effects, que tem a maior quantidade de plug-ins do mercado, da Adobe;
- Flash, da antiga Macromedia, hoje Adobe;
- Shake e Motion, da Apple;
- Fusion, da Eyeon;
- Nuke, da Forgeon.

Para modelagem e animação tridimensional, existem alguns softwares que já são padrões da indústria:

- Maya e 3DSMax, da Autodesk;
- Lightwave, da Newtek;
- o software livre Blender, que pode ser baixado gratuitamente no endereço www.blender.org , e tem todos os recursos profissionais de um programa 3D.

Existem diversos outros programas de aplicações específicas, lançados por empresas menores, e que acabam comprados pelas indústrias de maior porte e são incorporados em de sua linha de produtos, como o *Shake*, o *Elastic Reality* e outros. Para rastreamento de imagem, uma aplicação imbatível ainda é o *Boujou*. Muitas novidades aparecem e são incorporadas por outras aplicações e se deve estar atento para guardar cópias dos primeiros lançamentos, que cumprem a sua função durante muito tempo.

3.3. Tipos de profissionais de pós-produção e mercado de trabalho

Um departamento completo de pós-produção em uma empresa de grande porte envolve um grande número de profissionais nos diversos setores. Os setores da pós-produção podem ser divididos como a seguir:

- Edição;
- Efeitos visuais;
- Sonorização;
- Miniaturas e cenários virtuais;
- Videografismo;
- Colorização;
- Finalização.
-

Na Imagem 60, se tem um diagrama dos diversos setores da pós-produção e do fluxograma de trabalho.

Os profissionais envolvidos na pós-produção podem ser classificados da seguinte maneira:

- Diretor de finalização;
- Editor ou montador;
- Produtor de efeitos visuais;
- Produtor de som;
- Maestro;
- Videografista;
- Colorista;
- Operador de videotape;
- Operador de computação gráfica;
- Operador de videografismo;
- Assistentes de todas as funções em todos os setores.

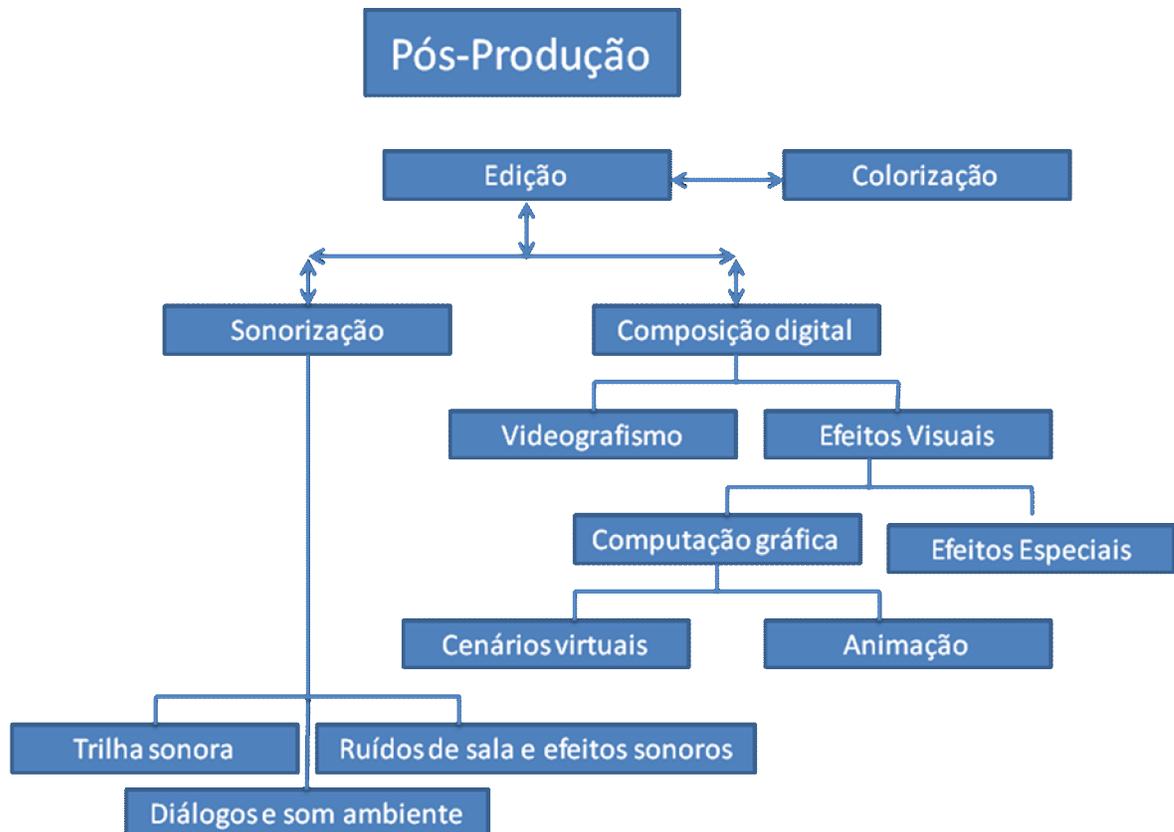


Imagem 60: a estrutura da pós-produção.

O mercado de trabalho para esses profissionais está principalmente em empresas do tipo:

- Produtoras e finalizadoras para publicidade e vídeos institucionais;
- Emissoras de televisão;
- Produtoras de conteúdo para tv por assinatura e internet;
- Operadoras de telefonia e emissoras que possuem canais de IPTV;
- Pós-produção para cinema.

3.4. Edição e montagem

Os termos *edição* e *montagem* significam a mesma coisa, mas, no português, se tornaram distintos, sendo a edição usada para o vídeo e a montagem para o cinema. Isso se deve à influência do francês no cinema, no seu início no Brasil, e do inglês na televisão. Em inglês existem os termos *video*

editing e *film editing* e, no francês, *montage cinématographique* e *montage vidéographique*. Neste texto se dará preferência então ao termo *edição*, que é mais popular.

Grosso modo, se pode definir esses termos como o trabalho de conexão de uma ou mais tomadas para formar uma sequência e a subsequente conexão das sequências para formar a peça audiovisual. É uma atividade que está no limiar entre a arte e a técnica. É, então, a arte que traz ritmo à obra, continuidade narrativa para a imagem e o som, refinamento dos planos visuais e que faz a inserção do som, da música, dos videografismos e dos efeitos visuais e especiais. É uma das atividades-chave do audiovisual e o profissional da área tem grande ascendência na construção do resultado final do produto audiovisual.

Na edição, são escolhidos os pontos de entrada e de saída do material bruto. O material bruto é o material gravado ou filmado, na sua forma original, seja em fitas ou em rolos, depois da revelação. Os intervalos entre os pontos de entrada e os de saída são os planos ou tomadas (*clip*, como no inglês das ilhas). Eles são inseridos na sequência editada e assim vai sendo montado o material. Pode-se colocar efeitos de transição entre esses planos, como as fusões (*dissolve*), os *fades in* e *out*, o *dip to color* e os *wipes*. Ou simplesmente se ter um corte seco. Faz-se também a edição do áudio. O processo de edição de imagem e de som é semelhante à edição de texto e essa semelhança é ainda maior com as ilhas digitais. O editor pode sobrepor, inserir ou substituir os planos na sequência como as palavras em uma frase ou as frases em um parágrafo de um texto de um programa de edição de texto.

Existem dois tipos de equipamento para edição, que dependem do tipo de ilha: linear e não-linear. A edição linear acontece quando se tem que correr as fitas de vídeo ou os rolos de filme. Para se encontrar uma tomada, se deve correr a fita ou o rolo até o ponto em que ela se encontra. Na transição da imagem analógica para a digital surgiram algumas ilhas digitais lineares. Nelas se simulavam, no programa, as fitas e rolos. Assim, para se usar uma tomada, tinha que se correr uma fita ou rolo virtual no próprio computador. Isso foi feito para que os editores e montadores, acostumados com a edição linear analógica, pudessem se adaptar à nova linguagem digital. Tanto nas ilhas digitais lineares quanto nas não-lineares as imagens e o som são passados

para dentro dos discos rígidos, no processo chamado de *digitalização*. É fundamental que a cada tomada seja dado um nome significativo e que todo o material seja organizado rigorosamente no programa de edição, para que elas sejam encontradas facilmente durante o trabalho de edição. Nas ilhas não-lineares, as tomadas são acessadas imediatamente com um clique do mouse. O nome não-linear deve-se a esse acesso randômico aos dados gravados. Ainda existe a necessidade de se correr as fitas para que se faça a digitalização. Com a substituição das fitas por cartões de memória, esse processo é feito de forma não-linear em todas as etapas.

Além desses processos, existe também a edição *offline* e a edição *online*. Elas existem para melhorar o fluxo de trabalho e otimizar os custos envolvidos. A ilha *offline* é, em geral, usada para definir a edição e o custo da sua hora de trabalho é quase dez vezes menor que a da edição *online*. Na *offline*, coloca-se todo o material gravado nos discos rígidos em baixa resolução e se faz a edição utilizando esse material. Ao final, com o trabalho concluído, gera-se um arquivo de texto chamado lista de decisões de edição (EDL – *edit decision list*), que é levado para a edição *online*. Esse arquivo contém todos os pontos de entrada e saída do material, tanto na fita quanto na sequência editada, o número da fita e os timecodes. Esse arquivo é reconhecido pela ilha de edição *online*, que então passa a digitalizar automaticamente somente o material editado, deixando uma sobra de alguns segundos antes e depois de cada plano (*head e tail*). Na ilha de edição *online* o material é digitalizado com a resolução máxima e o trabalho é finalizado, inserindo o som entregue pelo produtor de som ou maestro, os efeitos visuais e os videografismos.

Na TV Globo, por exemplo, todas as novelas e minisséries são editadas dessa maneira. A edição *offline* ainda é feita com ilhas Avid que utilizam o antigo computador Quadra, da Apple. Esses sistemas trabalham desde meados da década de 1990, vinte e quatro horas por dia, seis dias por semana. Os editores se revezam em três turnos de oito horas, sendo um deles o responsável também pela edição *online*. As grandes produtoras de audiovisual para publicidade também trabalham dessa maneira, com as edições *offline* e *online*.

A maneira de se montar filmes mudou radicalmente nos últimos quinze anos. Até então, a montagem era feita exclusivamente em Moviolas ou em mesas planas, como o Steenbeck e o KEM, cortando-se os rolos de filme nos pontos de edição e criando as sequências. 1995 foi o ano em que o número de montagens feitas em ilhas não-lineares se igualou ao número realizado em mesas como a Moviola [Murch 2004]. Nesses últimos anos, as ilhas digitais têm sido utilizadas intensamente, tendo se tornado um padrão na montagem, em três circunstâncias distintas. Na primeira, passam-se todos os rolos de filme para o formato de vídeo, com as marcações do negativo e o *timecode* presentes nas partes superior e inferior do vídeo. Esse é o processo de *telecinagem*. Em seguida, se edita todo o filme numa ilha não-linear *offline*, com os programas Avid ou Final Cut, e se gera, ao final, uma lista de corte (*cut list*), que é um arquivo de texto com todas as marcações do filme. Utiliza-se então essa lista de corte para montar o filme em uma mesa de montagem. A vantagem é que se podem testar diversas alternativas com agilidade na ilha digital, o que é trabalhoso na mesa de montagem. Uma imagem de uma ilha Steenbeck está na Imagem 61, onde se pode ver os rolos de filmes e os dois monitores de visualização da imagem, o da esquerda para o material bruto e o da direita para o material editado.

Na segunda alternativa, os rolos de filme são escaneados, quadro a quadro, e passados para a forma de imagem digital. Na terceira, a imagem já é gravada digitalmente. Nessas duas últimas formas, o filme é montado digitalmente e o produto final pode ser passado novamente para película, através do transfer para ser distribuído nas salas de cinema, no processo denominado *transfer*. O material bruto, digitalizado para um longa metragem, ocupa, em geral, mais de 10 terabytes (1 TB = 1 terabyte = mil gigabytes = um milhão de megabytes) de espaço nos discos rígidos.

3.5. Ilhas de edição não-lineares

Para se editar nos formatos DV ou HDV, basta uma placa e uma conexão firewire para dar entrada e saída do material gravado. Os

processadores acima do Pentium 4 ou Power PC têm capacidade suficiente para lidar com estes formatos. Assim, basta um software de edição para poder



Imagem 61: mesa horizontal de montagem Steenbeck.

trabalhar nesses formatos. Para formatos digitais mais sofisticados, como os de cinema 2K ou 4K, ou para formatos analógicos, placas de vídeo, processadores mais rápidos e discos rígidos especiais têm que ser utilizados, dependendo do caso. No caso do cinema digital, existe fluxos de trabalho (*workflow*) distintos, que dependem da câmera utilizada e do tipo de representação da imagem, como o LogC e o LUT Cineon.

As interfaces dos softwares de edição mais populares, como o Premiere, o Final Cut ou o Avid, são parecidas, seguindo uma arquitetura semelhante às das mesas planas Steenbeck, adotada inicialmente pelo Avid. Elas possuem duas janelas para visualizar a mídia e as ferramentas de edição são configuradas baseando-se nas mesas planas. Já as interfaces das ilhas Fire e Smoke e da família da Quantel são diferentes e com uma arquitetura que requer uma abordagem diferente.

A interface do Premiere está na Imagem 62 e a do Final Cut na Imagem 63. Pode se perceber as semelhanças das duas interfaces, com as ferramentas ocupando as mesmas posições. É possível configurar a interface de acordo com as preferências do usuário.

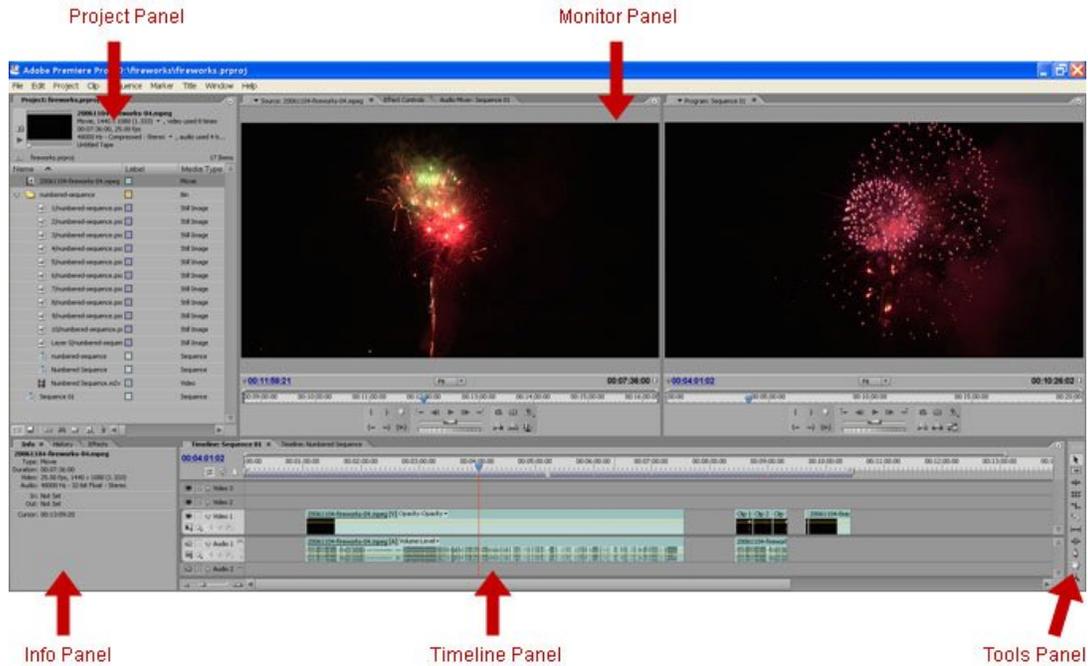


Imagem 62: interface do Adobe Premiere, com destaque para as diversas ferramentas.



Imagem 63: interface do Final Cut, com destaque para o waveform e o vectorscope.

As ilhas de edição podem trabalhar com um ou dois monitores para visualizar a interface do software. É aconselhável ainda ter um outro monitor, de televisão, para visualizar a imagem que está sendo trabalhada. A ilha também pode contar com um vectorscope e um waveform para controlar a qualidade da imagem em cor e brilho. Estas ferramentas também se encontram incorporadas nos softwares de edição, mas o controle externo é mais eficiente. O vectorscope é uma ferramenta que mede a intensidade da crominância, indicando os níveis máximos aceitáveis para as cores primárias e secundárias. O waveform mede a luminância para a imagem.

A parte superior da interface tem a barra de menus, com todas as ferramentas e utilitários do software. No canto superior esquerdo encontra-se toda a mídia envolvida na edição, com a lista de clipes, de imagens escaneadas ou geradas por outros softwares, de áudio e as sequências editadas. Mais à direita temos os dois monitores para visualizar, em baixa resolução, o que está sendo editado. O monitor da esquerda mostra o clipe do material bruto e o da direita a sequência que está sendo editada. Logo abaixo dos monitores encontram-se os controladores do vídeo, para tocá-lo, andar quadro a quadro, retornar, ir ao início e ao final e para marcar os pontos de entrada e saída. Existem alguns outros controles, dependendo do software.

No canto inferior direito se encontram diversas ferramentas para manipular a imagem editada e o visualizador do nível de áudio. Mais à esquerda, encontra-se a *timeline*, a linha de tempo, com o material editado. Ela contém diversas pistas de vídeo e de áudio, cujas saídas podem ser feitas simultaneamente. No canto inferior esquerdo, estão os controles para cada uma dessas pistas.

Além dessas interfaces, existem outras opções, de acordo com o que está sendo feito na edição, como uma interface para composição gráfica e para efeitos visuais, para equalização do som, ou para controle da imagem, com o *vectorscope* e o *waveform*.

CAPÍTULO 4

EFEITOS VISUAIS

4.1. Introdução

A produção atual de efeitos visuais para o cinema e o vídeo distanciou-se completamente do que era feito normalmente até a década de 1990. Com o advento da imagem e do som digitais, novas técnicas foram inventadas e todas as outras adaptadas para essa nova forma de manipulação do audiovisual, tornando mais amplas as soluções para realizar imagneticamente aquilo que é criado. Os resultados inovadores obtidos a cada ano ampliam as fronteiras do que pode ser realizado criativamente. Hoje em dia, pode-se contar qualquer história imaginada, com as imagens produzidas virtualmente sendo completamente semelhantes às gravadas de cenas reais.

O objetivo deste capítulo é discutir como os recursos ópticos e digitais são utilizados para se realizar efeitos visuais e, ao mesmo tempo, fornecer um painel das diversas técnicas de computação gráfica e de tratamento de imagem para o cinema e o vídeo. A produção de efeitos visuais usa intensivamente a modelagem geométrica, a visualização computacional, o processamento digital de imagens e, mais recentemente, a visão computacional. Para que seja então possível planejar e obter resultados satisfatórios, é fundamental o conhecimento dos conceitos teóricos, das técnicas e das ferramentas disponíveis, incluindo as definições dos diversos tipos de compressão e descrição digital da imagem em movimento (*motion picture*) introduzidas no capítulo anterior. Os efeitos visuais fazem parte da pós-produção do audiovisual, um dos setores do audiovisual que mais foi impactado e transformado pela imagem digital, como foi discutido no capítulo anterior.

O capítulo está dividido em cinco seções. Após a introdução, é abordada a evolução das diversas técnicas com trucagem óptica. Na terceira seção, é introduzido o conceito de canal alfa e de composição digital. Na quarta seção, são discutidas algumas técnicas de efeitos visuais digitais, como a correção de cor, o rastreamento de imagem e a controle de câmera, assim como os

diferentes hardwares e softwares disponíveis e os conceitos utilizados neles. Na quinta seção, são apresentados também efeitos visuais realizados na TV Globo pelo autor da tese, em cada etapa de sua produção. Na última seção, são discutidas as perspectivas de pesquisa acadêmica e em produções de menor porte.

4.2. Evolução das trucagens com imagem

Os efeitos visuais vêm sendo empregados como instrumentos de apoio para viabilizar narrativas desde as primeiras obras cinematográficas. Dos primórdios do cinema até o início da era digital utilizavam-se exclusivamente trucagens ópticas para imprimir o efeito na película. Essas trucagens evoluíram significativamente nos primeiros cem anos do cinema, começando no cinema mudo e chegando ao início dos anos 1990, quando começaram a ser substituídas pelos efeitos realizados em computadores.

Georges Méliès inaugurou o uso dos efeitos visuais e foi dos mais importantes inovadores da história do cinema. Ele trabalhava em teatro e trouxe para esta nova linguagem diversos recursos visuais. Com Méliès, se começou a empregar recursos ópticos para obter os efeitos, como superfícies semi-espelhadas, divisor de feixes da projeção para impressão na película, dentre outros. Dessa maneira, iniciou uma nova forma de fazer cinema, não documental, diferentemente dos filmes dos irmãos Lumière. Ele introduziu as visões, os sonhos, os pesadelos, desejos, medos, o inconsciente e a vida interior. Foi o primeiro a utilizar a exposição dupla (*La Caverne Maudite*, 1898), a primeira tomada com um divisor de imagens e atores em dois ambientes diferentes (*Un Homme di Tête*, 1898) e a primeira fusão (*Cendrillon*, 1899), além de tomadas de miniaturas, efeitos de replicação e transparência. Esses recursos são hoje usados massivamente em diversos formatos de imagem em movimento [Rickitt 2007].

Nos primeiros anos, os efeitos visuais eram também obtidos através da técnica do *stop action*, ou *stop motion*. Ela também é conhecida como *substituição por parada de ação*. Consiste em parar a filmagem, alterar a cena e continuar a filmar. No caso de uma animação, isso é feito quadro a quadro.

Um exemplo clássico dessa técnica é o filme *King Kong*, de 1933, da RKO. Nele, todas as criaturas da ilha foram animadas com a técnica de *stop motion* e o filme provocou um enorme impacto nas plateias da época. Além da substituição por parada de ação, utilizavam-se miniaturas, tanto em cenários quanto com bonecos e outras técnicas de composição com projeção de imagens misturadas com efeitos reais, nos chamados efeitos *on camera*. Os procedimentos de efeitos visuais requerem e requeriam o posicionamento preciso das imagens no filme, para que múltiplos elementos possam ser efetivamente combinados para atingir o resultado desejado. Isto era obtido com câmeras especiais para efeitos visuais, que eram chamadas de *camêras de processo*.

Uma série de técnicas foi desenvolvida para fazer a composição dos diversos elementos de imagem opticamente, resultando em um fotograma final como se todos os elementos tivessem sido filmados simultaneamente. Elas não serão apresentadas detalhadamente neste trabalho, mas o livro [Rickitt 2007] é uma excelente referência para o estudo detalhado destas técnicas. Como efeito *on camera* se tem a *tomada trombone*, na qual a câmera se movimenta em um trilho, no movimento conhecido como *travelling*, indo em direção a um objeto ou personagem, ao mesmo tempo em que a distância focal da lente é alterada. Um exemplo clássico desse tipo de tomada é a cena no filme *Psicose*, de Alfred Hitchcock, quando o detetive sobe a escada da casa antes de ser assassinado. Nesta tomada, o objeto ou o personagem permanecem com o mesmo tamanho relativo enquanto o tamanho do background parece se alterar. Na Imagem 64, está representada esquematicamente a tomada com diferentes distâncias focais e, na imagem 65, o efeito da tomada trombone.

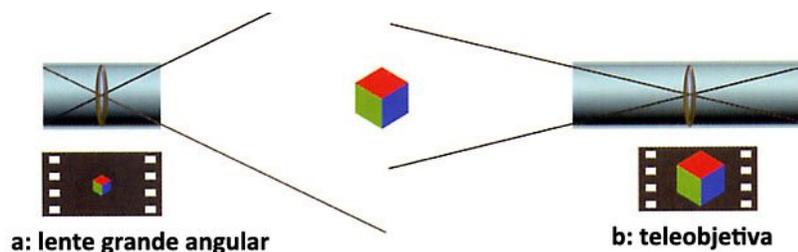


Imagem 64: uma mesma tomada feita com diferentes lentes e o efeito provocado na imagem [Rickitt 2007].

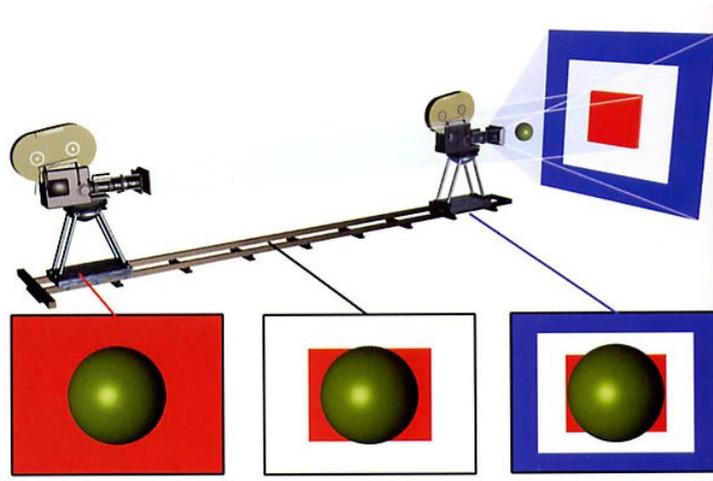


Imagem 65: o efeito trombone, com a câmera se movendo no trilho e alterando a sua distância focal; nos fotogramas na parte inferior se observa a alteração do fundo em relação ao objeto [Rickitt 2007].

Na sobreposição de duas imagens para gerar a composição, se não são realizados truques, a imagem que é aplicada sobre o background aparecerá como um fantasma, devido à transparência que nela deve ser inserida. Para contornar este problema, foi criado o *travelling matte*. Nesta técnica, a imagem de background tem um buraco exatamente onde vai ser inserido o elemento da outra imagem, que é impressa no background posteriormente. É uma técnica complexa de ser realizada, pois o elemento a ser inserido pode estar em movimento, mas ela foi realizada inúmeras vezes com sucesso no cinema. Essa técnica se tornou inteiramente redundante com a imagem digital. Uma variação desta técnica é o processo Williams, patenteado em 1918, no qual o elemento a ser inserido é filmado com fundo negro e é depois composto com o background, através de revelações sucessivas das duas imagens com filme de alto contraste. A técnica é ilustrada na Imagem 66.

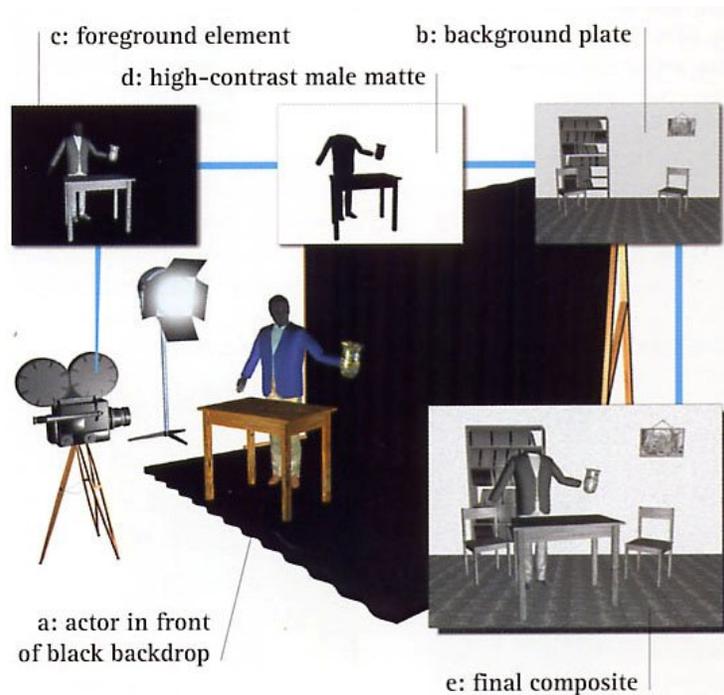


Imagem 66: representação esquemática do processo Williams [Rickitt 2007].

Variações da técnica do *travelling matte* foram desenvolvidas e utilizadas com sucesso no cinema pré-digital. A mais importante delas é a técnica do fundo azul, ou *blue-screen*. Nela, o elemento a ser inserido no background é filmado com o fundo azul. Ao utilizar filtro azul na revelação, cria-se um matte de alto contraste, que é utilizado para compor o fundo com a imagem da frente. Esse efeito continua a ser utilizado com a televisão analógica do tipo componente e com todos os tipos de imagem digital com amostragem adequada de cor. O fundo mais utilizado atualmente é o verde, pois esta cor é representada mais adequadamente na imagem digital. Na Imagem 67, é mostrada a filmagem de um clássico, *Jornada nas Estrelas: o Filme*, de 1979 [Rickitt 2007].



Imagem 67: preparação de cena com fundo azul em *Jornada nas Estrelas* [Rickitt 2007].

Um avanço notável foi a invenção da impressora óptica (*optical printer*). A função mais básica desse equipamento é retrofotografar uma imagem em um novo pedaço de filme. É como se uma câmera estivesse apontada para um projetor. O projetor é carregado com uma imagem positiva e a câmera captura essa imagem, quadro a quadro. Pode-se colocar filtros, lentes e máscaras no caminho da projeção para obter-se os efeitos desejados, assim como misturar duas ou mais projeções [Pintea 2004]. Na Imagem 68, é mostrada uma impressora óptica fabricada pela Oxberry, a 1600. O equipamento ganhou o Oscar de 1986, de tecnologia. Ela utiliza quatro projetores controlados por cinco computadores. As imagens dos projetores são compostas e filmadas quadro a quadro para produzir efeitos visuais complexos.



Imagem 68: Impressora óptica Oxberry 1600, vencedora do Oscar de 1986.

Os efeitos mais comuns obtidos com essa técnica são classificados como:

1. efeitos de transição: empregados para criar uma mudança no tempo e na localização entre as cenas. Dentre diversos tipos de transições entre planos de imagem, temos:

- a) o *fade in* e *fade out*, quando a imagem desaparece ou surge de uma cor, em geral o preto;
- b) fusão, quando um plano de imagem vai desaparecendo enquanto o próximo aparece gradualmente;
- c) *wipe*, que são transições feitas com figuras geométricas, como círculos e quadrados, trazendo com elas a próxima cena;
- d) virada de página, quando as imagens são mudadas como se estivessem na folha de um livro.

2. quadro congelado, que é feito com o quadro (*frame*) de imagem preso na impressora e ele é continuamente repetido;

- 3. mudança da sequência de quadros;
- 4. composição de imagens de diferentes projeções;
- 5. efeito de *zoom*;
- 6. aceleração ou desaceleração do movimento;

7. reversão do movimento [Burum 2007].

Durante décadas, foram realizados filmes com efeitos visuais ópticos, muitos se tornando clássicos e referência na cinematografia. *Guerra nas Estrelas* (1977), *Blade Runner* (1985), *2001, uma Odisséia no Espaço* (1968), *King Kong* (1933), *Contatos Imediatos do Terceiro Grau* (1977), *O Mágico de Oz* (1939), *Roger Rabbit* (1988) e *ET* (1982) são exemplos de filmes realizados com a impressora óptica. Sete dos dez melhores filmes de todos os tempos, segundo a Visual Effects Society, com uso intensivo de efeitos visuais, foram realizados sem recursos digitais. Essa situação deve ser revertida rapidamente, já que são mais de cem anos de produção clássica contra o período de menos de duas décadas de evolução da produção digital.

4.3. O canal alfa

A imagem digital pode ser exibida como uma composição de camadas de imagens. Nesse caso, a imagem é construída a partir de diferentes fontes de imagens, que são compostas para que o resultado pareça como se tudo tivesse sido capturado ao mesmo tempo, de uma única vez. Muitas vezes o fundo (*background*) é separado da imagem da frente (*foreground*), como no caso dos telejornais, onde o apresentador da previsão de tempo em geral está na frente de mapas que formam o *background*.

Para que as camadas de imagens possam ser colocadas umas sobre as outras, deve ser definida a transparência de cada imagem. Se a transparência não existisse, somente a imagem que é colocada por cima seria exibida. Então é necessário criar a informação com a transparência de cada pixel na imagem.

A forma encontrada para se definir a transparência é colocar essa informação em um quarto canal, além dos três canais de cor de cada pixel. Esse quarto canal é chamado de canal alfa e ele contém o mesmo número de bits de que cada um dos outros canais. No caso da imagem com oito bits por canal para cada pixel, então o canal alfa tem 256 valores de transparência por pixel, indo do completamente transparente ao completamente opaco. Nesse caso, a imagem resultante terá 32 bits de informação por cada pixel da imagem, com 8 bits, ou 256 valores, para cada um dos canais.

O canal alfa pode ser criado de diversas maneiras. Pode-se, por exemplo, gravar uma pessoa ou objeto com um fundo verde ou azul. Depois, ao compor essa imagem sobre outra, toma-se essa cor como informação para o canal alfa e ela fica transparente. Esse processo é conhecido por *chroma key* ou *color key*, ou chave de cor. Na Imagem 69, estão dois exemplos de *chroma key* realizados para as novelas *Chocolate com Pimenta* e *Desejos de Mulher*, da Rede Globo.





Imagem 69: cenas construídas utilizando o canal alfa fornecido pelo *chroma key* [Rede Globo 2003].

Outra forma de criar a transparência é contornar regiões da imagem e definir essas regiões como transparentes ou opacas. Isso pode ser feito com ferramentas de desenho, em softwares como o Photoshop, o After Effects, o Illustrator e o Gimp. Essa técnica é conhecida como máscara (*mask*) e as informações dessas transparências e opacidades são incorporadas ao canal alfa. Na Imagem 70, é mostrado um exemplo da utilização de máscara em uma cena trabalhada para a novela *Cabocla*, da Rede Globo. Nessa cena, se contornou com uma máscara o céu acima das montanhas e também utilizei máscaras na paisagem para mudar a cor da cena. Por trás da máscara do céu,

foram feitas alterações, colocando céus diferentes, de outras fontes de imagem. Nas outras máscaras, se pode alterar localmente a cor. Dessa maneira, foi feita uma transição do entardecer para a noite com apenas uma fotografia. Na imagem 70, se vê dois quadros dessa transição.



Imagem 70: cena construída utilizando o canal alfa fornecida por máscaras para a substituição do céu e colorização da paisagem [Rede Globo 2004].

Na Imagem 71, se pode observar outra aplicação de máscaras e filtros de processamento. Essa é a sequência de encerramento da novela *Chocolate com Pimenta*. A transformação do frame com a atriz foi feita no Photoshop e composta depois na caixa de chocolate, com movimento de câmera, utilizando os programas After Effects e Inferno. A imagem final é transformada para

parecer uma gravura. A câmera, à medida que se movimenta, vai abrindo a cena e a imagem aparece como a gravura de uma caixa de chocolate.



Imagem 71: sequência final da novela *Chocolate com Pimenta*, com uso intensivo do canal alfa e do processamento de imagem [Rede Globo 2003].

O canal alfa, em geral, é representado como tons de cinza, onde o branco é o opaco e o preto é o transparente. Os valores intermediários de transparência são representados pelos demais tons de cinza. Mas pode-se também definir qualquer outra cor como base para o canal alfa, que é então

pré-multiplicado por aquela cor com relação ao preto, como no caso do *chroma key*.

Na Imagem 72 se tem a imagem de fundo (*background*), com um tabuleiro de xadrez (4a), e a imagem do primeiro plano (*foreground*), da arara (4b). Na figura também se observa a imagem do canal alfa (4c) e a imagem composta resultante (4d).



Imagem 72: composição de imagens com canal alfa.

4.4. Técnica e arte dos efeitos visuais digitais

As técnicas de efeitos visuais podem ser classificadas de diversas maneiras, dependentes dos recursos artísticos e técnicos empregados. Uma classificação baseada nos métodos computacionais utilizados está a seguir:

- Composição digital;
- Correção de cor;
- Rastreamento;
- Personagens e/ou cenários virtuais;
- Captura de movimento;
- *Morphing*;
- *Warping*;
- Controle de movimento de câmera;
- *Matchmoving*.

A composição de imagens é o processo de integração de imagens de diferentes fontes em uma única imagem final, como foi apresentado

anteriormente neste capítulo. É um processo amplamente conhecido e discutido nas artes plásticas e na fotografia. Para o vídeo, são utilizadas diversas ferramentas e técnicas para obter uma sequência de imagens em movimento a partir da integração de diferentes imagens. O maior desafio é produzir a imagem integrada, sem parecer que foi feita de diferentes fontes de imagens. As imagens originais são aplicadas em diferentes camadas e digitalmente recortadas e compostas para resultar em uma imagem final. A camada superior é aquela que aparece como resultado final, se nenhum efeito for aplicado. Para se obter a imagem final, se aplica nas camadas níveis de transparência e combinação das cores, definidas em cada pixel, escalamento, translação e/ou máscaras digitais. Isto foi discutido na seção anterior.

Como foi apresentado, pode-se também produzir máscaras digitais através do *chroma key* (chave de cor) ou do *luma key* (chave de luminância). No *chroma key*, escolhe-se um matiz de cor, com um \square de variação em torno desse valor, que é a tolerância. O canal alfa é gerado ao se substituir o matiz pelo preto, as diferentes tonalidades dentro do \square por cinzas e as demais cores pelo branco. Em geral utiliza-se a cor azul ou verde para se gerar o *chroma key*, pois são cores diferentes das tonalidades da pele. No *luma key* o processo é semelhante, empregando valores de luminância no lugar da cor.

Na realização do *chroma key* é fundamental o conhecimento do tipo de imagem digital que está sendo utilizada. Uma câmera do tipo DV, DVCAM ou similares, com sub-amostragem 4:1:1, não permite a obtenção de uma máscara com as bordas bem definidas. A forma como a imagem é descrita introduz artefatos nas bordas do elemento cuja imagem é recortada do fundo de *chroma key*. No mercado há inúmeros casos de iluminadores que foram demitidos injustamente sob o argumento de que não sabem iluminar adequadamente a cena gravada com câmeras com sub-amostragem 4:1:1 para se fazer o *chroma key*.

A correção de cor é feita por ferramentas que permitem alterar os valores de cada canal de cor. Essas alterações podem ser feitas em diferentes intervalos de luminância, que são as sombras (*shadows*), os meio-tons (*midtone*s) e os mais luminosos (*highlights*). Elas podem ser realizadas no histograma da imagem, ou através de alteração das curvas de cor nos diferentes intervalos, ou numericamente, atribuindo novos valores às cores, se

alterando o seu matiz, a saturação e a luminância. As alterações podem ser aplicadas em regiões da imagem, através de máscaras digitais, ou em toda ela.

A profissão do colorista é muito valorizada nas finalizadas de cinema e nas produtoras de vídeo profissionais, pois é um técnico com qualidades artísticas especiais. Essa técnica é utilizada na composição de imagem para que as diferentes camadas se integrem adequadamente. Um exemplo de utilização dessa técnica é na chamada *noite americana*. Nesse caso, as sequências de imagens são gravadas com a luz do dia e transformadas em noite, através da correção de cor. Além da correção da cor, as camadas sobrepostas de imagem podem ter os valores das componentes de cor em cada pixel submetidas a operações aritméticas, resultando em uma imagem com um novo padrão de cor. Pode-se somar, subtrair, multiplicar, dividir, tornar os valores negativos ou submeter os canais de cor a diversas outras operações especiais, como a solarização e a posterização.

As ferramentas de rastreamento utilizam diferentes princípios, dependendo do software de composição. Pode-se, por exemplo, fazer a seleção de uma região de um quadro de imagem. Essa região vai sendo mapeada quadro a quadro, comparando-se os valores de cor e luminância de sua vizinhança, o que permite identificar a posição da região em cada quadro. Ou pode-se fazer o rastreamento de toda a imagem, através de segmentação, gerando vetores que indicam a variação da posição de cada ponto da imagem ao longo dos quadros. As duas soluções monitoram alterações de escala, posição e direção dos pontos, mas o segundo processo dá resultados mais satisfatórios. A primeira técnica é utilizada por softwares como o *Inferno*, o *Fire* e o *After Effects*, e a segunda pelo *Boujou*. A vantagem da segunda técnica é que ela incorpora nos dados as aberrações e distorções introduzidas pela lente da câmera. O rastreamento de imagem é utilizado para apagar erros de gravação ou para adicionar elementos na imagem, através da composição digital, entre diversas outras aplicações.

Os personagens e cenários virtuais são gerados por softwares de modelagem tridimensional. Envolve diretamente a computação gráfica, com a modelagem digital e o mapeamento de texturas. Pode-se obter o nível de realismo que se desejar. O nível do realismo da imagem virtual depende da linguagem adotada no produto audiovisual que está sendo feito. Em muitos

casos, ele é dispensável e até indesejável. Esses elementos gerados por computador podem se integrar a imagens reais ou a imagens totalmente virtuais. Os softwares produzem movimentos de câmera, iluminação, texturas e vínculos físicos, como a gravidade, que dão realismo às cenas produzidas. Também podem gerar textos e vários elementos artísticos que são utilizados na composição final.

O *morphing* e o *warping* pertencem a uma mesma classe de efeitos visuais produzidos por processamento digital de imagens. Nela, uma imagem é transformada em outra utilizando regiões da imagem com características semelhantes, que são contornadas por curvas, e sofrem transformações morfológicas. O software calcula as mudanças que devem ser feitas nas cores e formas nas regiões limitadas pelas curvas, em um intervalo de quadros pré-estabelecido pelo operador. Este processo foi desenvolvido pelos pesquisadores do MIT, D. Smyth e A. Brigham, que receberam o Oscar em 1992 pela criação e desenvolvimento do *morphing*. Essa técnica é utilizada tanto para a transformação de personagens como para a justaposição de elementos virtuais por objetos cenográficos ou a criação de planos-sequência a partir de planos de câmera com cortes. O *warping*, por sua vez, provoca distorções na mesma imagem através de contornos pré-definidos. Dessa maneira, é possível fazer, por exemplo, movimentos na boca de animais, como se estivessem falando. Na Imagem 73 é apresentado um *morphing* de transformação de uma mulher em um homem quadro a quadro. As curvas contornando regiões nas faces dos atores são observadas nas figuras.

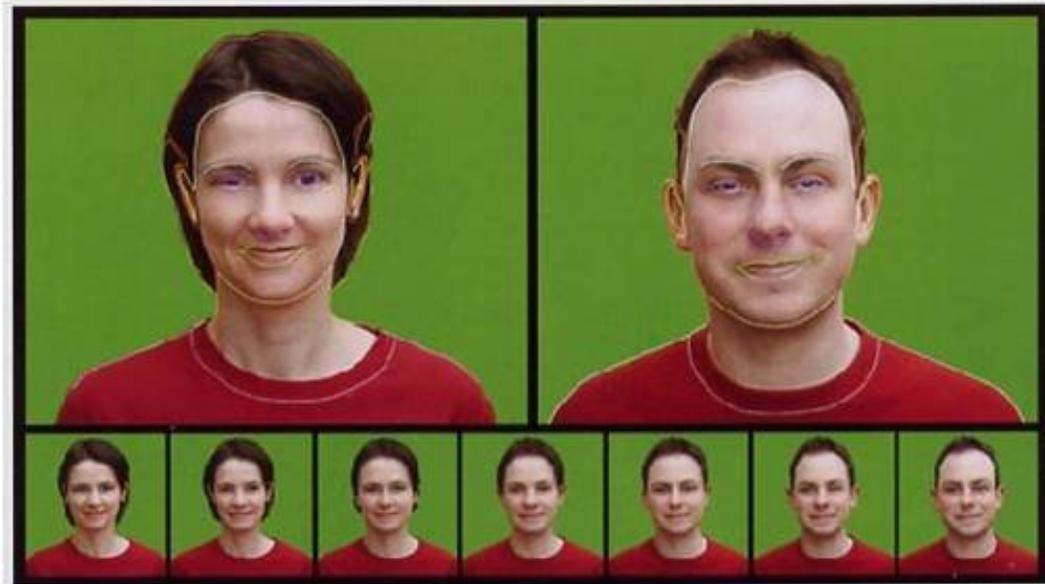


Imagem 73: transformação morfológica entre dois atores [Rickitt 2007].

No controle de movimento, a movimentação da câmera é realizada por dispositivos monitorados e controlados pelo computador. Dessa maneira, os movimentos podem ser pré-programados e repetidos e se obtém planos para serem utilizados em composições sofisticadas. Utilizando efeitos visuais ópticos, uma câmera com controle de movimento filmou a nave *Discovery* em *2001, uma Odisséia no Espaço* (1968). O filme *Guerra nas Estrelas* (1977) empregou o controle de movimento de câmeras para simular movimentação de naves e de atores no espaço. O controle de movimento também é utilizado para multiplicar um mesmo personagem em cena, contracenando consigo mesmo e com a câmera se movimentando. A Imagem 74 mostra um esquema de um sistema de controle de movimento, como o utilizado em *Guerra nas Estrelas*.



Imagem 74: sistema para movimentação de câmera, controlado por computador [Rickitt 2007].

Todas essas técnicas podem ser empregadas de forma combinada. Outras técnicas, como a interpolação de imagens, a pintura digital e a multiplicação de figuração, também utilizam a computação gráfica e o processamento digital de imagens na sua realização. Muitos filmes lançados recentemente têm como motivação principal apresentar novas técnicas de computação gráfica e de tratamento digital de imagens e de efeitos com modelagens mais aprimoradas de fenômenos físicos, pelos, folhagens, fumaça, água, fogo e dublês digitais, como em filmes de animação.

4.5. Exemplos de produções

A seguir são apresentados quatro exemplos de utilização de efeitos visuais, com uma discussão sobre a maneira como foram realizados, além dos que foram discutidos na seção 3. Como primeiro exemplo, se tem a utilização de correção de cor, rastreamento e composição digital para construir tomadas de passagem (*stockshot*) entre duas locações diferentes. Uma cena diurna,

filmada em 1918, foi transformada em uma cena noturna, ao se alterar a cor e colocar luzes na imagem, através de mudanças na luminância aplicadas em máscaras digitais. Como a câmera se movimentava, o movimento da imagem teve que ser rastreado para que os elementos introduzidos e compostos pudessem acompanhar o movimento da câmera. Na Imagem 75, é apresentada a imagem original e, na Imagem 76, a cena transformada em noite.



Imagem 75: cena diurna do Teatro Municipal [Cinemateca Brasileira 1918].



Imagem 76: cena modificada, com luzes aplicadas em máscaras digitais, correção de cor, rastreamento de movimento e céu noturno [Rede Globo 2004].

No segundo exemplo, multiplicou-se a figuração presente em uma cena gravada no Teatro C3lon, em Buenos Aires, que contava com poucos figurantes na plateia e nas frisas do teatro. O movimento da c3mera, com uma lente grande angular, iniciava enquadrando o teto do teatro. A c3mera estava sobre um trilho atr3s da plateia. O movimento da c3mera descia do teto em *tilt* e, ap3s enquadrar o palco por tr3s da plateia, fazia um *travelling* lateral. O problema principal foi introduzido pela lente grande angular, que provocava distorç3es na imagem, dificultando o rastreamento de pontos. Ao utilizar os softwares Inferno ou After Effects, que utilizam rastreamento de pontos escolhidos da imagem, a localizaç3o dos pontos ficava inst3vel, fazendo com que os elementos da figuraç3o se movessem diferentemente do movimento da imagem. Isso dava um aspecto indesej3vel ao efeito, que tornava vis3vel a sua aplicaç3o.

Foi feito ent3o o teste com o software *Boujou*, que tinha sido lançado h3 pouco tempo na 3poca, em 2003. Ele utiliza o fluxo 3ptico e gerava milhares de eixos em toda a imagem, ao longo dos *frames*, em pontos definidos pelo software, permitindo escolher aqueles mais adequados para aplicar os elementos adicionais. Os dados das posiç3es dos eixos ao longo do tempo gerados pelo *Boujou* foram transferidos para o Inferno. Dessa maneira foram aplicados os figurantes virtuais, seguindo o movimento do fluxo 3ptico localmente na imagem. Nas frisas, os figurantes foram pintados, com a ferramenta *Paint* do Inferno, e neles foram aplicadas as posiç3es dos eixos rastreados. Da plateia, foram recortados figurantes reais, que foram aplicados nas cadeiras vazias. Essa cena n3o foi gravada com a intenç3o de multiplicar a figuraç3o e os grandes vazios de pessoas na plateia s3o foram notados depois, durante a ediç3o. Na Imagem 77, se tem os quadros da cena sem a multiplicação de figuraç3o e, depois, com a plateia e frisas com os figurantes virtuais aplicados.



Imagem 77: cenas do Teatro C3lon, em Buenos Aires, antes e depois de modificadas, com figurantes adicionados 3s frisas e 3 plateia [Rede Globo 2003].

No terceiro exemplo, utilizou-se o *warping* para contrair a pupila do ator, pontuando dramaticamente a cena. Nessa sequ4ncia, um dos atores principais levava um tiro. No momento que precede o tiro, em um plano-detanche do olho, foi feita a contra4o da pupila do ator. Para isso, utilizou-se a ferramenta de *warping* do Inferno, criando-se uma curva em torno da pupila e, no final do efeito, uma nova curva com menor di4metro. O software calculou e distorceu as imagens, de forma crível, entre o primeiro e o 3ltimo quadro do efeito. Na Imagem 78, observa-se o quadro antes e depois do efeito aplicado.



Imagem 78: *warping* aplicado 3 pupila, antes e depois do efeito aplicado [Rede Globo 2004].

O quarto exemplo 4 uma sequ4ncia de tentativa de atropelamento. Nesse caso, foram utilizadas m3scaras digitais, rastreamento e composi4o digital, com atores e dubl4s atuando em um fundo para *chroma key*. Em um dos planos, a c3mera movimentou-se sobre um trilho e gravou a atriz

atravessando a avenida. Em seguida, foi feita uma máscara do capô da caminhonete, que foi aplicada sobre o trilho, seguindo o seu movimento. No contra-plano, se gravou a cena com os dublês atuando contra um fundo de *chroma key*. Com a câmera fixa, gravou-se depois a caminhonete se aproximando e fez-se a composição digital das duas tomadas. Na Imagem 79, se pode ver dois quadros da sequência, com o plano e o contra-plano.



Imagem 79: frames da cena com composição digital, rastreamento e chroma key [Rede Globo 2002].

4.6. Perspectivas para a pesquisa e para a produção audiovisual

A pesquisa e o desenvolvimento de soluções de processamento digital de imagem e de computação gráfica para os efeitos visuais digitais é uma área muito ativa na Ciência da Computação. Na maioria das vezes, procura-se a colaboração de artistas e técnicos de efeitos visuais para o desenvolvimento de produtos, de acordo com a demanda do mercado de produção para cinema e jogos digitais. Novas soluções, utilizando conceitos de visão computacional, têm introduzido produtos de nova geração no mercado.

Pode-se citar, como temas atuais de pesquisa, o desenvolvimento de codecs e de sistemas de render mais eficazes, a captura digital de movimento sem marcadores, o rastreamento de movimento em tempo real, o *morphing* através de modelos escaneados em 3D, os *engines* para a descrição de radiância e de fenômenos físicos, os softwares de mapeamento de texturas para aplicação em programas de modelagem 3D e métodos mais eficazes e

menos destrutivos para a transformação, entre si, das diferentes cadências da imagem digital em movimento, 60i, 24p, 50i, dentre outros temas. No mapeamento de texturas, por exemplo, estuda-se a representação de imagens por funções polinomiais em cada pixel, descrevendo os valores de cores, ao invés de valores de cores fixas. Isso tem aplicações, dentre outras, na melhoria da visualização de objetos em diferentes condições de iluminação e na focalização mais eficiente de objetos. Outro exemplo é o desenvolvimento de *engines* para a modelagem de fenômenos físicos e texturas complexas de objetos, que tem evoluído rapidamente nos últimos dez anos, como pode ser comprovado pela melhoria da qualidade das animações em filmes e em jogos digitais.

A produção e a criação do audiovisual e de jogos têm se beneficiado com a evolução e a ampliação do uso do processamento digital de imagem, da computação gráfica e da visão computacional [Hartley e Zisserman 2006]. As ferramentas disponíveis para os efeitos visuais digitais apresentam alto grau de sofisticação a um custo relativamente baixo. Além disso, novas mídias, como a web, os celulares e os assistentes pessoais digitais, e a atualização das mídias tradicionais, como o cinema e a televisão, têm aumentado as possibilidades de emprego de diferentes técnicas e expandido o mercado para o profissional da área. Isso possibilita a popularização do seu uso, por artistas e técnicos os mais variados e para as mais diversificadas aplicações, sendo fundamental então o conhecimento teórico para se obter resultados de qualidade.

Gene Warren, diretor da Fantasy II Film Effects, que produziu, dentre outros filmes, o *Drácula de Bram Stoker* (1992), *Underworld* (2003) e *Lost* (2004-2008), comenta essas mudanças: “Nós finalmente paramos de usar a impressora óptica em 2002. Até o fim podíamos criar composições incríveis com nossas velhas máquinas. Mas não conseguíamos competir com os garotos que podiam gastar alguns poucos milhares de dólares em uma *workstation* e alguns softwares. Qualquer um pode agora juntar algumas camadas de imagem apertando poucas teclas, mas não há dúvida de que saber como realmente uma tomada funciona – seja ela digital ou óptica – é uma arte. Mas a honorável técnica de obter uma grande imagem daquelas velhas máquinas imensas é atualmente redundante e esse talento tradicional será logo perdido para sempre [Rickitt 2007].”

CAPÍTULO 5

CAPTURA DE MOVIMENTO E ANIMAÇÃO DE PERSONAGENS

5.1. Introdução

A animação de personagens em jogos eletrônicos e no audiovisual é uma área em contínuo desenvolvimento acelerado. No caso dos jogos, ela passou por diversas etapas de sofisticação, desde os primeiros games 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 na vida contemporânea, 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 do audiovisual e 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 nesse momento. A técnica evoluiu com as demandas de interatividade para os jogos e de inovações para o cinema, 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 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, 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.

Neste capítulo, a utilização da captura digital de movimento e os conceitos relacionados à animação de personagens em geral são abordados. A captura de movimento (*mocap*) é amplamente usada para animar personagens em jogos e no audiovisual 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 ou no audiovisual, em geral, os dados são alugados, comprados ou adquiridos gratuitamente de fornecedores, já que apenas uma empresa, pelo que se conhece, faz esse trabalho específico no Brasil [RPM 2009], e cobra R\$ 11.000,00 por uma diária de uso do seu equipamento e pelos dados de movimento. 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, o autor desta tese trabalhou com a equipe do prof. Arnaldo de Albuquerque Araújo para construir um sistema de captura de movimento de código aberto.

Esse capítulo discute, tanto do ponto de vista prático quanto teórico, essas questões e também a animação de personagens e os usos da captura de movimento para animar personagens. Neste capítulo, não será abordada mais longamente a evolução histórica da animação de personagens e da captura de movimento. Essa abordagem pode ser encontrada em [Gomide 2006] e nos livros de [Menache 2000] e [Thomas e Johnston 1984].

O capítulo 6 apresentará o OpenMoCap, o software de código livre desenvolvido durante a tese. 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, que teve seu início em 2006 [Gomide 2006]. Atualmente ele é apoiado financeiramente pela FAPEMIG e pelo CNPq. O software atualmente permite a realização, em tempo

real, de todo o fluxo de trabalho e fornece os dados de movimento na saída. 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.

5.2. 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 pela evolução técnica para despontar como uma forma de expressão humana, no início do século XX. A arte da animação é expressa 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.

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, essa 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, [Thomas e Johnston 2004], escreveram o clássico *The Illusion of Life: Disney Animation*. Nesse livro estão descritos os doze princípios de animação de personagens, que podem ser considerados como princípios básicos para a animação de personagens em geral. Tendo em vista esses princípios se pode 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 2004]:

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 ideia 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 sequência 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 intermediários 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 é linear; refere-se aos objetos se movendo no espaço em arcos em vez 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 plateia 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 às 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 1980 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 [Gomide 2010a].

5.3. Captura digital de movimento

A captura de movimento, também conhecida como *motion capture* (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 as séries *Fifa Soccer* e *NBA Live* [Gomide 2010a].

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 2000], “captura de movimento é o processo de gravar um evento de movimento ao vivo, e traduzí-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*. A 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 é adotada uma definição mais vaga, sendo a captura de movimento o “ato de capturar os movimentos em geral”, pode-se localizar sua origem no final do século dezenove, nos trabalhos dos fotógrafos [Eadweard Muybridge 1984], com o zootrópio, e [Etienne-Jules Marey 2011], com a cronofotografia. 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. Marey realizou estudos sobre o

movimento humano e animal. Ele utilizou uma roupa negra com marcações brancas, correspondentes aos ossos mais importantes para a caminhada, e realizou séries de fotografia utilizando uma placa fotográfica, que era exposta diversas vezes para registrar o movimento. Na Imagem 80, é apresentado um homem com a roupa, à esquerda, e a cronofotografia, à direita.

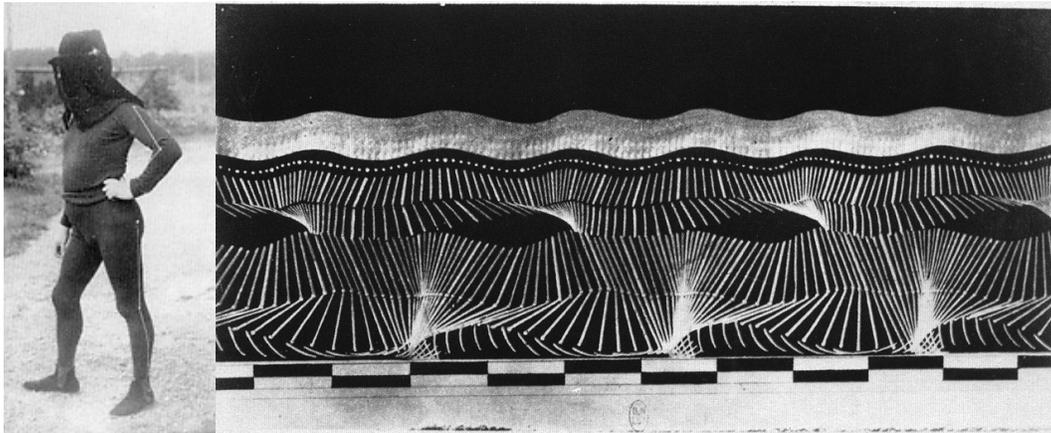


Imagem 80: cronofotografia, à direita, de um homem com uma roupa preta e marcas brancas, à esquerda [Marey 2010].

Muybridge utilizou uma técnica diferente, com uma série de câmeras que eram disparadas em sincronismo, gerando uma sequência de imagens. Depois ele revelava e ampliava as fotografias e as colocava no zootrópio, que era uma espécie de praxinoscópio [Gomide 2006]. Na Imagem 81 é mostrado o galope de um cavalo, que venceu um prêmio por ter resolvido um desafio. O desafio era provar que o cavalo galopando levantava as quatro patas simultaneamente. Isto é observado na Imagem.

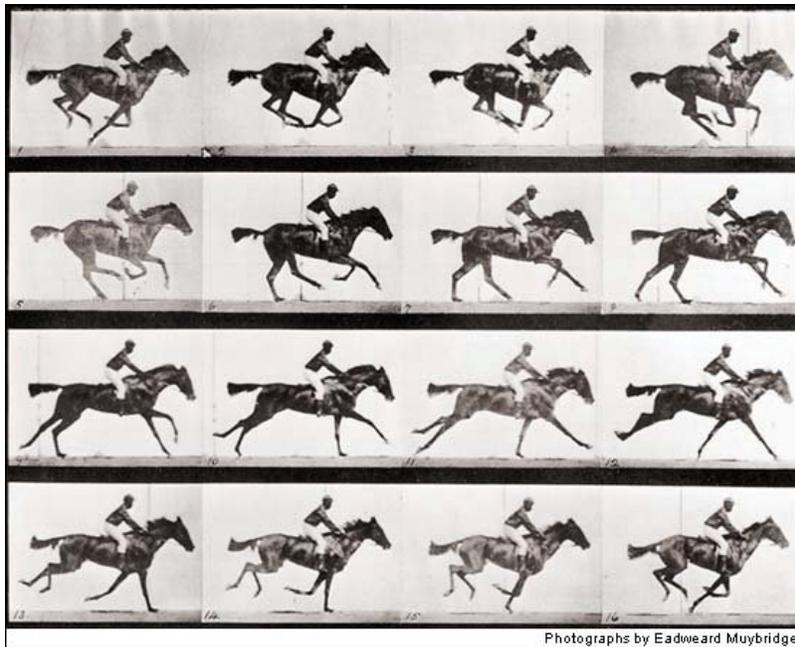


Imagem 81: fotografias do galope de um cavalo feitas por Muybridge [Muybridge 1984].

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 de captura digital de movimento como é conhecida atualmente. 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 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. A imagem 82 exibe a folha do registro da patente pelos irmãos Fleischer.

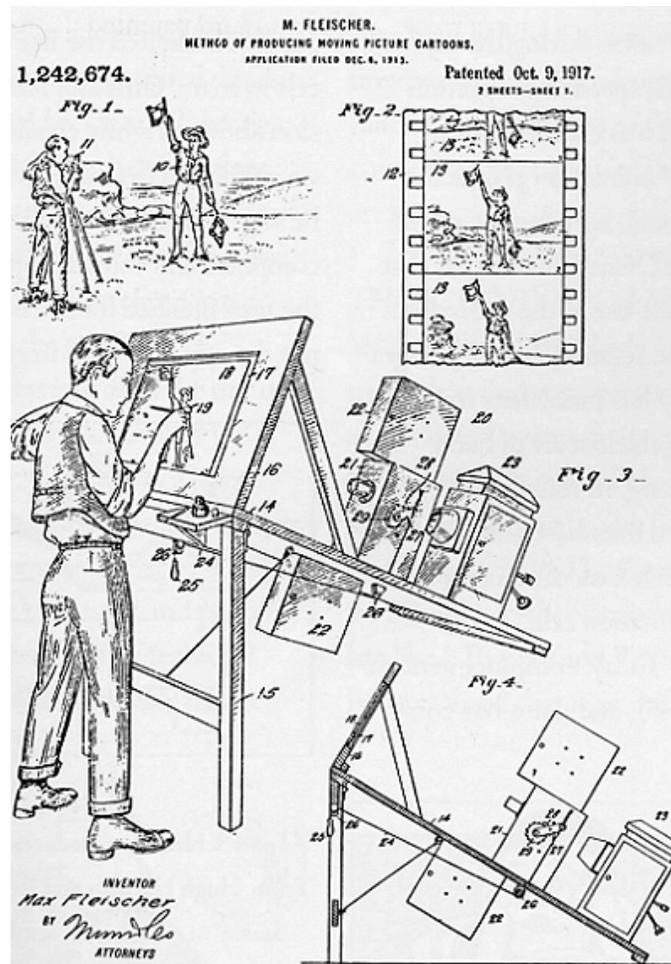


Imagem 82: patente do processo da rotoscopia [Flescher 2005].

Algumas das tecnologias usadas atualmente para a captura de movimento existem desde a década de 1980, utilizadas em aplicações com propósitos médicos e militares. [Sturman 1994] e [Menache 2000] fazem revisões históricas dessa evolução nos primeiros anos da captura digital de movimento. De todos os eventos, serão citados 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. Havia o desejo de 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]. O *making of* da produção do comercial está em [Abel 2011] Na Imagem 83, é exibido um fotograma deste comercial.



Imagem 83: *Sexy Robot*, primeiro personagem animado por captura de movimento [Abel 2011].

No começo dos anos 1990, 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 das empresas que trabalhavam com mocap para animação e seu desenvolvimento no seu início. Como nos anos 1990 os movimentos não precisavam ser muito precisos e os personagens eram 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 interagia 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 1990, 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 interagia 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. Um frame desta animação está na Imagem 84.



Imagem 84: cena do clipe *Steam*, no qual Gabriel Líquido contracena 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 sequência 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 faria 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 sequência 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 Imagem 85, que é um fotograma da sequência. A captura de movimento não forneceu os dados pretendidos e se abandonou o projeto, optando-se pela animação por *keyframe*.

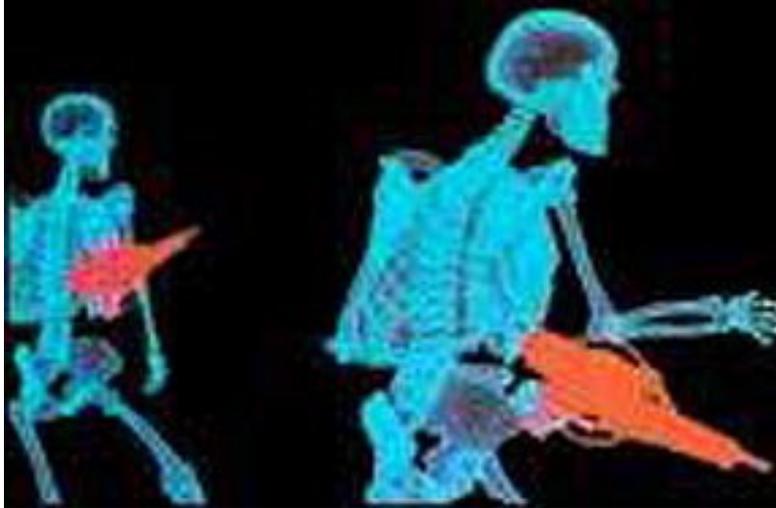


Imagem 85: seqüência dos esqueletos, que poderia ter sido a primeira cena do cinema com captura digital de movimento [Menache 2000].

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. Um fotograma do filme é mostrada na Imagem 86.



Imagem 86: robô T-1000, primeiro personagem animado por captura de movimento da história do cinema [Menache 2000].

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 e 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]. A interface do jogo está na Imagem 87.



Imagem 87: 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 ideias 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 *País 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. A Imagem 88 mostra os personagens do jogo e da animação.



Imagem 88: personagens de *Donkey Kong Country* [Nintendo 2011].

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.

A mocap é uma técnica cada vez mais utilizada nos dias de hoje. Dentre diversos exemplos bem sucedidos do seu uso, o mais marcante é o do filme *Avatar* (2009), que é o maior sucesso de bilheteria da história do cinema até agora.

5.4. Tecnologias de Captura de Movimento

5.4.1. 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, Gomide et al 2010b]. 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 Imagem 89. A complexidade do módulo de análise depende do nível dos dados enviados pelo módulo de percepção do movimento.

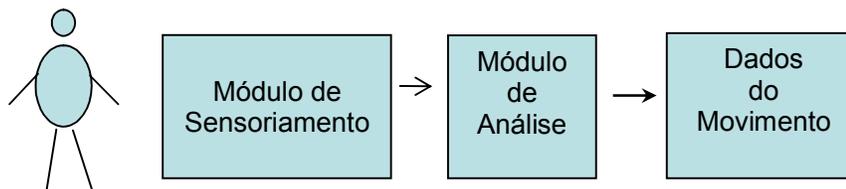


Imagem 89: diagrama dos diversos componentes de um sistema de mocap [Moeslund 2000].

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 ideia 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 capítulo. Serão abordadas brevemente as capturas mecânica e a magnética.

5.4.2. 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 Imagem 90, vê-se uma luva desenvolvida para capturar movimentos da mão, chamada *CyberGlove* [Cybergloves 2011]. 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 1990, 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.



Imagem 90: luva para mocap [Cybergloves 2011].

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 2011]. 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 [Animazoo 2011]. O sistema é apresentado na Imagem 91. É ideal para a captura de corpo inteiro de um indivíduo e seu preço é de 75 mil dólares hoje. É possível também comprar somente a parte superior ou inferior do torso, custando aproximadamente 52 e 43 mil dólares [Inition 2011].



Imagem 91: sistema inercial IGS-190H [Animazoo 2011].

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;
- é complexo devido à quantidade de detalhes mecânicos dos equipamentos;
- traz limitações ao movimento das juntas humanas;
- a maior parte dos sistemas não calcula deslocamentos globais sem a ajuda de sensores magnéticos.

5.4.3. 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].

Esses 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 seção 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 Imagem 92, 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.



Imagem 92: dispositivo magnético Flock of Birds montado para captura de movimento [Ascension 2011].

5.4.4. Sistema óptico

Os sistemas ópticos podem ser divididos em ativos e passivos, com ou sem marcadores [Gomide 2006, Flam 2009]. No sistema ativo os marcadores são fontes de luz, em geral LEDs colocados nas junções do corpo. O sistema passivo usa refletores como marcadores, como na Imagem 93. 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.



Imagem 93: sistema óptico Impulse [PhaseSpace Tech 2011].

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, 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 Imagem 94 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.



Imagem 94: sistema de captura com os LEDs acessos em torno da lente de entrada da câmera, retroiluminada [Vicon 2011].

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. O sistema utilizado no *O Expresso Polar*, por exemplo, 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 feita em um ambiente controlado, sem muitos pontos de reflexão para evitar interferência nos resultados.

5.5. Etapas da Captura de Movimento

Nesta seção serão discutidas as diversas etapas que devem ser trabalhadas para se obter os dados de movimento para animar o personagem. A seção serve como um guia para o artista e o animador para que ele realize, de forma correta, a captura de movimento e preparação do personagem.

5.5.1. Preparação da sessão de captura de movimentos

Como a captura de movimento é utilizada para animar um personagem virtual, se deve 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: compressão e esticamento, 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 sequência 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 sequências, 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 permitam agilizar 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 tanto na pós-produção quanto 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 possa 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. À medida que as tomadas vão sendo feitas, os arquivos devem seguir a nomeação 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].

5.5.2. 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.

A colocação dos marcadores deve levar em conta três questões básicas para: 1. definir a posição ou orientação de uma junção; 2. se fazer distinção entre direita e esquerda e entre os diferentes performáticos; 3. 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 Imagem 95.

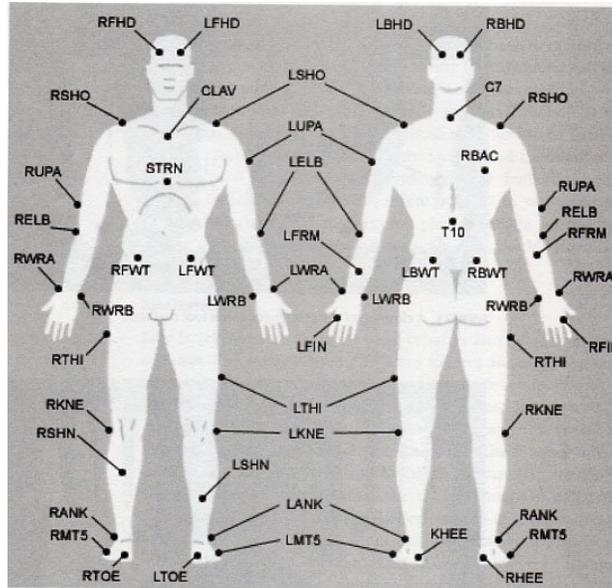


Imagem 95: posições para 41 marcadores [Menache 2000].

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 sequência de imagens registradas pela câmera, com uma nuvem de pontos, como na Imagem 96. 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.

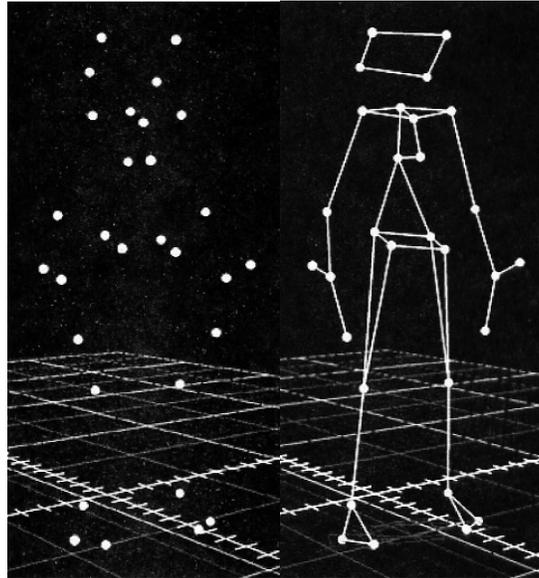


Imagem 96: dados sem grande quantidade de processamento, com os marcadores como se fosse uma nuvem, sem conexões e conectados por um boneco de palitos [Menache 2000].

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.

5.5.3. Preparação do personagem para mocap

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 sequência das transformações das articulações anteriores, como na Imagem 97. 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.

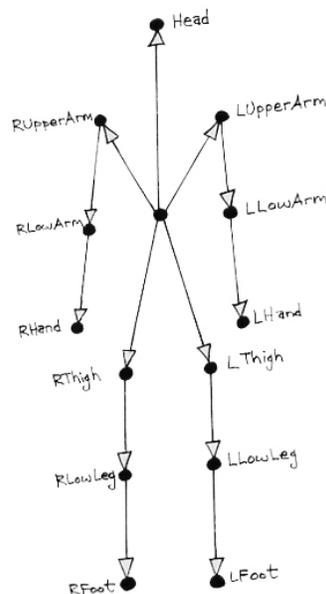


Imagem 97: um tipo de hierarquia para o esqueleto [Menache 2000].

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 esses 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.

CAPÍTULO 6

O SOFTWARE DE CÓDIGO LIVRE OPENMOCAP

6.1. Introdução

Neste capítulo será detalhado o funcionamento do software OpenMoCap. Ele já foi apresentado na introdução e no capítulo anterior, em seus aspectos inovadores e nas perspectivas de sua utilização. Primeiramente, se discutirá as etapas envolvidas no processo de captura de movimento, que foram implementadas no software. Em seguida, um dos aspectos mais importantes do fluxo de trabalho de captura de movimento é apresentado, que é a visão estereoscópica e a triangulação. Na quarta seção é apresentada a arquitetura do software e, por último, as últimas atualizações do OpenMoCap para que ele possa ser disponibilizado no portal sobre animação e realidade aumentada com a captura de movimento. O desenvolvimento do software ocorreu durante o doutorado e através das orientações e dissertações de [Flam 2009] e [Queiroz 2011], e recebeu apoio financeiro de instituições de fomento, como discutido anteriormente.

6.2. Metodologia

O OpenMoCap realiza todo o fluxo de captura de movimento apresentado na Seção 6.5. Este fluxo de trabalho deve ser realizado em tempo real, para que se possa pré-visualizar o que está acontecendo com a animação do personagem. O software, inicialmente, identifica os marcadores, que são os pontos de interesse, através do conjunto de câmeras. Os marcadores são refletores que enviam de volta para as câmeras a luz no infravermelho que incide sobre eles. Na frente de cada câmera, existe um filtro que deixa apenas passar a radiação no infravermelho.

Cada câmera então captura a imagem da cena, com os marcadores brilhantes. Através de uma operação de limiar de luminância, isto é, que determina a região que tem um valor acima de um valor definido da escala de

cinza, são detectados os marcadores. O software determina as posições dos centróides de cada um dos marcadores, através da geometria da região dos pixels contíguos que estão acima do valor de limiar. Ao conjunto de marcadores é dada uma estrutura semântica, correspondente ao esqueleto humano, como descrito na seção 6.5. Atualmente isto é feito manualmente, mas uma versão para determinar automaticamente a semântica está sendo finalizada.

A partir dessa etapa de inicialização, o software prevê, através de filtros de busca, qual a região em que serão procurados os centróides e os pontos de interesse no próximo quadro da imagem, na etapa de rastreamento. Inicialmente foi implementado o filtro Alfa-Beta, descrito por [Flam 2009]. A dissertação de [Queiroz 2011] determinou qual o melhor filtro para ser aplicado nesta etapa, entre três opções: Alfa-Beta, Kalman e CONDENSATION. O filtro de Kalman apresentou a melhor performance e atualmente está implementado no aplicativo. Isto otimiza o processamento para achar os marcadores no próximo frame [Queiroz et al 2011].

Os dados da posição dos centróides em cada câmera, que estão em duas dimensões, são comparados para, através do processo de triangulação [Flam 2009], encontrar-se a terceira coordenada da posição. Essa é a etapa mais complexa do fluxo de trabalho e, para isso, as câmeras devem ser preliminarmente calibradas. Isto será discutido na próxima seção. Com os dados da posição tridimensional dos marcadores em cada quadro da imagem, a saída de dados é dada no formato .bvh, que é lido pelos softwares de modelagem. Esses dados estão prontos para serem aplicados no personagem digital. Alguns softwares, como o Blender, geram automaticamente o esqueleto do personagem a partir do arquivo de dados bvh.

Todo o fluxo de trabalho realizado pelo OpenMoCap é mostrado no diagrama da Imagem 98. Nela se observa cada etapa do fluxo de trabalho discutido nesta seção e na Seção 6.5.

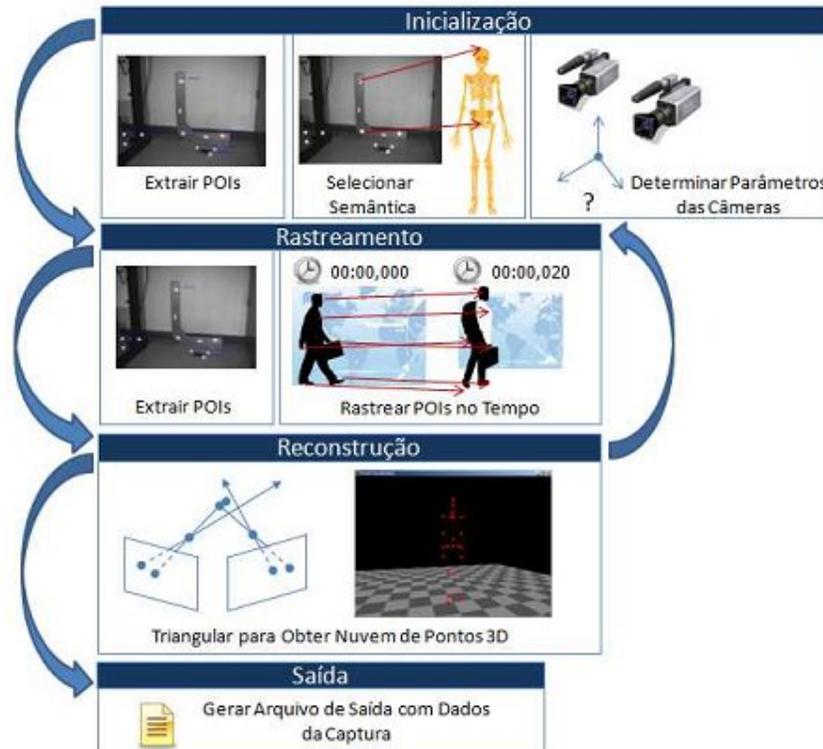


Imagem 98: fluxo de trabalho do OpenMoCap [Flam 2009] e [Gomide et al 2010].

6.3. Visão estereoscópica e triangulação

Segundo [Forsyth e Ponce 2002], a disparidade entre duas imagens produzidas pelos olhos de um homem permite um ganho significativo na noção de profundidade de objetos em uma cena. Este é princípio utilizado para se criar o cinema 3D, onde duas filmadoras posicionadas lado a lado capturam a imagem e, ao serem exibidas, provocam a ilusão de profundidade [Mendiburu 2009]. Nesse processo, a imagem é projetada com dois projetores e, por um processo de separação, a imagem gravada pela filmadora da direita vai para o olho direito e a imagem da filmadora esquerda vai para o olho esquerdo. Para separar as imagens, se utilizam filtros, com as cores complementares, ou polarizadores, na frente dos projetores e nos óculos que o espectador deve usar. Desta maneira, somente a imagem gravada por cada uma das câmeras chega ao olho, produzindo o efeito 3D.

Esse processo visual é conhecido como estereopsis e é um importante artifício binocular usado pelo cérebro para ajudar na percepção da

profundidade, além de vários outros existentes, como o contraste, as sombras, o conhecimento prévio dos objetos e a paralaxe, discutidos no Capítulo 1. A visão estéreo, do inglês *stereo vision*, é a técnica capaz de reproduzir artificialmente esse processo, utilizando duas ou mais câmeras. Ela é imprescindível para aplicações como a captura de movimento, a navegação de robôs, a cartografia e o reconhecimento aéreo. A fusão de características e a reconstrução são os dois passos que devem ser realizados na visão estéreo. A fusão de características consiste em determinar a correspondência entre pontos de imagens produzidas com diferentes origens, mas de uma mesma cena. Já a reconstrução é responsável por obter as coordenadas tridimensionais de pontos na cena por meio das correspondências adquiridas no primeiro passo e algumas propriedades matemáticas.

A triangulação é o processo responsável por estimar a posição tridimensional de um POI por meio de suas projeções nos planos de imagem das câmeras e suas respectivas matrizes de projeção, discutidas na Seção 2.4. Esta etapa é realizada de forma contínua, como o rastreamento, e é processada quadro a quadro pela aplicação de captura de movimento. Mas, diferentemente do rastreamento, o processamento não é local, mas sim global. Logo, informações de todas as imagens obtidas pelas câmeras em um mesmo instante de tempo precisam ser reunidas para obter o resultado desejado, que são as coordenadas tridimensionais dos POIs. Depois de obtidas as coordenadas tridimensionais, é gerado o arquivo de saída para ser utilizado no aplicativo de modelagem e animação 3D, como o Blender e o Maya.

6.4. 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 Imagem 99 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 pontos de interesse (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.

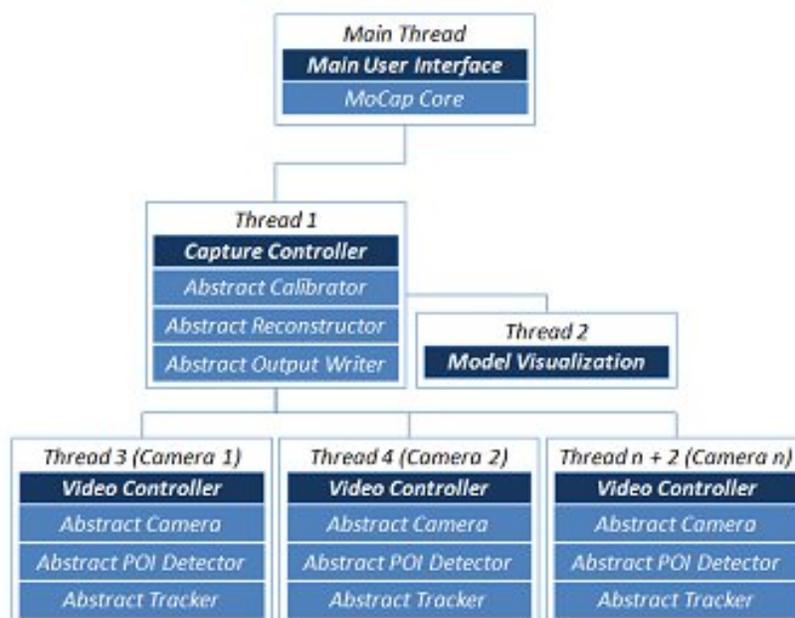


Imagem 99: Diagrama da Arquitetura do OpenMoCap [Flam 2009].

A interface gráfica principal é ilustrada pela Imagem 100, 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.

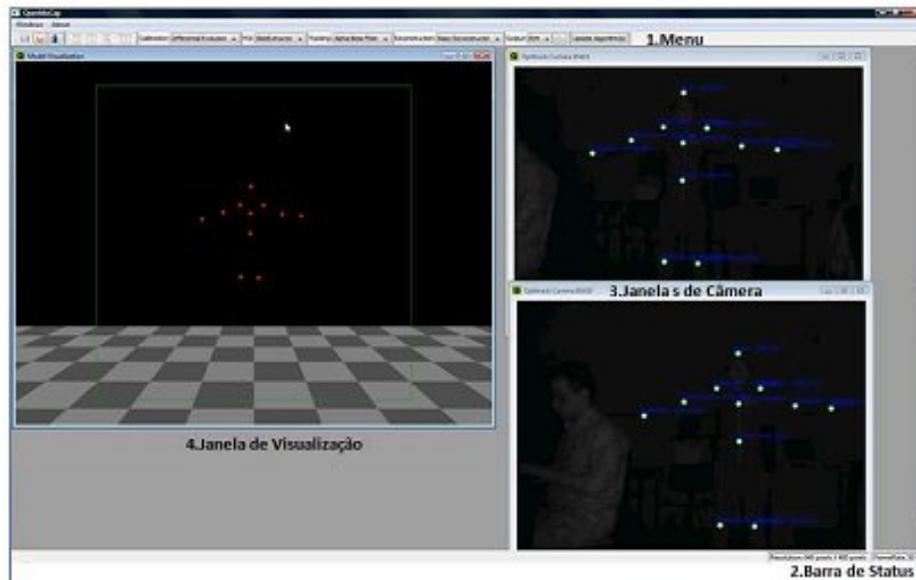


Imagem 100: Interface Gráfica Principal do OpenMoCap [Flam 2009].

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. 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 Imagem 99 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 a 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 Imagem 99, é 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 existente, 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 [2011], 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.

6.5. Atualizações mais recentes

Recentemente foram feitas as atualizações do núcleo do programa, utilizando as últimas versões disponíveis das bibliotecas OpenCV e QT. Estas atualizações fornecem algoritmos mais eficazes, novas propriedades e resolução de alguns *bugs* que haviam nas versões anteriores. O sistema de log de erros também foi melhorado, o que ajuda os desenvolvedores debugar o código de maneira mais eficiente [Gomide et al 2011].

A melhoria mais importante é uma versão final do OpenMoCap com um método de calibração métrica estéreo mais fácil de utilizar e que está incorporada ao software. Este era o principal defeito das versões descritas em [Gomide et al 2010a] e [Flam et al 2009]. Na versão anterior do software, era possível capturar o movimento apenas com valores proporcionais, sem uma escala métrica.

Este problema foi resolvido, e as coordenadas de saída agora têm um valor métrico. O algoritmo de calibração, baseado na evolução diferencial [Storn e Price 1997], foi substituído, na nova versão do OpenCV, por um algoritmo mais tradicional e eficiente. Desta maneira, pode-se usar um padrão de calibração, do tipo tabuleiro de xadrez, e o sistema pode ser calibrado em poucos minutos. O padrão do tabuleiro precisa ser exibido para as câmeras em algumas posições e o sistema está pronto para capturar, depois de se designar os pontos manualmente.

CAPÍTULO 7

PORTAL DA WEB OPENMOCAP.ORG

7.1. Objetivos

Este portal se estrutura nos temas relacionados à animação e à realidade aumentada utilizando a captura digital de movimento [Gomide et al 2011]. No portal será disponibilizado, a partir de setembro, o primeiro software de captura de movimento de código aberto do mundo, o *OpenMoCap*, que funciona em tempo real. O portal também irá fomentar discussões sobre as suas aplicações em animação de personagens e na realidade aumentada. Ele vai disponibilizar gratuitamente o software para *download*, assim como os manuais do usuário e os tutoriais. Em suas diversas seções, se poderá entrar em fóruns de discussão e haverá uma galeria de trabalhos realizados com o software e com a técnica da captura de movimento, permitindo o aprimoramento das produções culturais realizadas com a captura de movimento. Uma das seções do portal será para desenvolvedores, na busca de obter novas contribuições para tornar o software uma solução competitiva, diante da imensa carência para produtos desse tipo, ainda mais sendo ele de código aberto e gratuito. O portal também contará com uma página para a oferta de serviços e trabalhos na área de produção cultural relacionada à captura de movimento.

7.2. Público-alvo

O público-alvo para o qual o conteúdo do portal é dedicado compreende todos os indivíduos no mundo que se interessam e/ou estão envolvidos com a arte da animação e a captura de movimento e suas aplicações no cinema, na televisão, nas novas mídias, em jogos digitais e na realidade aumentada. Esse público envolve animadores, artistas gráficos, produtores culturais, roteiristas, desenvolvedores de softwares, estudantes e interessados nesses temas.

Identificando em tópicos:

7.2.1. Animadores: o animador poderá baixar o software e utilizá-lo como recurso para animar seus personagens, tanto para aplicações profissionais quanto amadoras. Praticamente todas as seções fornecerão conteúdo que interessa o animador, seja nos fóruns de discussão, na página de serviços e trabalhos, na galeria de imagens, nos tutoriais e artigos sobre animação e captura de movimento. Esse público pode ser de qualquer faixa etária e ter qualquer escolaridade.

7.2.2. Artistas gráficos: o artista gráfico terá interesse principalmente nos fóruns de discussão, na página de serviços e trabalhos, na galeria de imagens e nos artigos sobre animação e captura de movimento, para ampliar o seu horizonte de aplicações. Esse público pode ser de qualquer faixa etária e ter qualquer escolaridade.

7.2.3. Produtores culturais: o interesse será principalmente em uma solução prática e de baixo custo para realizar os seus projetos. Também terá interesse nos fóruns de discussão, na galeria de imagens, nos tutoriais e artigos sobre animação e captura de movimento.

7.2.4. Roteiristas: as ferramentas e os recursos disponíveis com a captura de movimento e a animação devem despertar interesse nos roteiristas, pois abrem novas possibilidades para contar histórias.

7.2.5. Desenvolvedores de software: esse público estará interessado nessa produção coletiva em seus diversos aspectos, que envolvem conceitos multidisciplinares de visão computacional, interatividade, transmissão e armazenamento de dados e tem aplicações em diversas áreas com grande impacto social. O público corresponde a uma faixa etária acima dos dezoito anos e com um nível de formação mais elevado.

7.2.6. Estudantes e interessados: todo o conteúdo do portal interessará a esse público. Esse público pode ser de qualquer faixa etária e ter qualquer escolaridade.

7.3. Sistema de Organização do Site

O portal será estruturado para o seu público-alvo, de acordo com os seus objetivos. No esquema da Imagem 101, está representado o mapa do site e a navegabilidade entre as diversas seções.

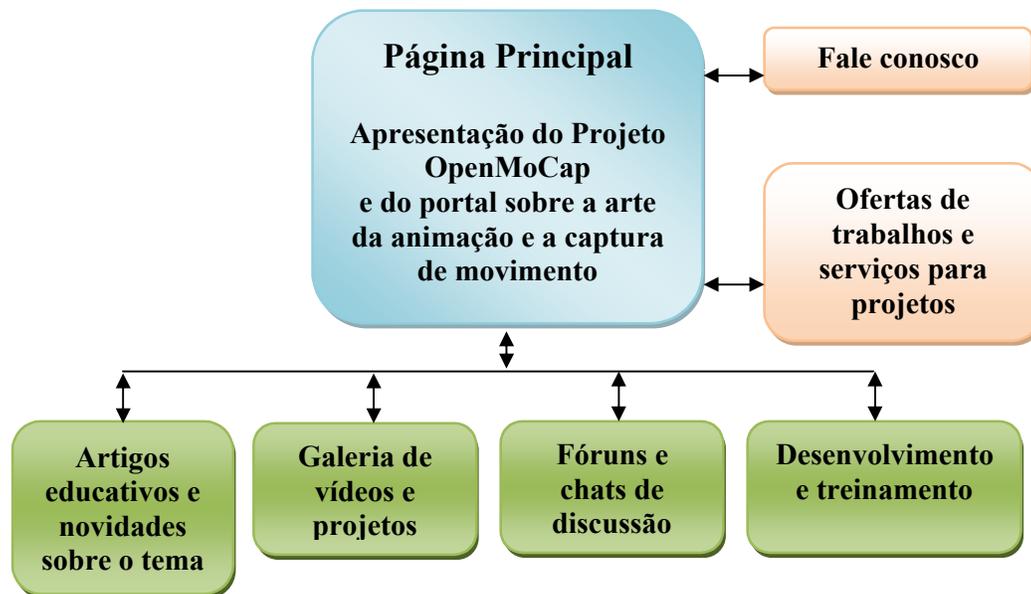


Imagem 101: diagrama da estrutura do portal.

A seguir, é detalhada cada uma das seções:

7.3.1. Página principal: apresentará o projeto coletivo OpenMoCap e o conteúdo do portal.

7.3.2. Artigos educativos e novidades sobre o tema: essa seção será informativa e educativa, abordando os diversos aspectos da arte da animação e das utilizações da captura de movimento. O usuário poderá fazer download ou acessar os artigos de uma lista para ler ou ouvir online.

7.3.3. Galeria de vídeos e projetos: seção dedicada a contribuições dos artistas e produtores culturais, com seus vídeos e projetos envolvendo a captura de movimento e a animação de personagens. O material poderá ser baixado ou assistido por *streaming*.

7.3.4. Fóruns e chats de discussão: seção das mais interativas do portal, com fóruns e chats dedicados aos diversos freqüentadores, sejam animadores, artistas gráficos, produtores culturais, roteiristas, desenvolvedores de softwares, curiosos, estudantes e interessados no tema.

7.3.5. Desenvolvimento e treinamento: seção dedicada às colaborações e treinamento de desenvolvedores que eventualmente poderão colaborar na melhoria do código e nos projetos envolvidos nessa criação coletiva.

7.3.6. Fale conosco: página que disponibiliza uma possibilidade do usuário se comunicar com os mantenedores do portal.

7.3.7. Ofertas de trabalhos e serviços para projetos: página onde artistas poderão oferecer seus serviços e habilidades e os produtores poderão oferecer trabalho na área de abrangência do portal.

O portal será disponibilizado em português, inglês e em arquivos sonoros, com possibilidade de mudança no tamanho das fontes, imagens e contraste. Aplicações para melhorar a acessibilidade para pessoas portadoras de deficiência sensorial e dificuldade de comunicação devem ser esperadas, com a utilização da captura de movimento associada à realidade aumentada.

7.4. Sistema de Rotulação e Navegação do Portal

Nesta seção são apresentadas as funcionalidades e o sistema de navegação de cada página do portal. Inicialmente, são descritas as partes fixas de todas as páginas para, em seguida, serem abordadas as estruturas de cada uma das páginas principais e das páginas acessórias às páginas principais.

7.4.1. Partes comuns a todas as páginas do portal

7.4.1.1. Na parte superior da página, haverá um frame horizontal, ocupando toda a sua extensão, com o nome (OpenMoCap), a logo e a frase “PORTAL PARA PRODUÇÃO DE ANIMAÇÕES DE PERSONAGENS E REALIDADE AUMENTADA COM A CAPTURA DE MOVIMENTO”. Abaixo da frase, um botão para login e um menu de popup para acesso à língua e as opções “Português”, com as bandeiras do Brasil e de Portugal, e *English*, com as bandeiras dos Estados Unidos e da Inglaterra.

7.4.1.2. Logo abaixo deste frame, na horizontal, teremos os botões com acesso às páginas do portal, com um clique sobre o botão. A ordem é a seguinte:

- a. página principal. Este botão será representado pela logo do OpenMoCap.
- b. artigos e novidades sobre o tema. Este botão terá o termo *Artigos e Novidades*.
- c. galeria de fotos e imagens relacionadas à animação de personagens e realidade aumentada com captura de movimento. Este botão terá a palavra *Galeria*.
- d. fóruns de discussão. O botão terá a palavra *Comunidade*.
- e. desenvolvimento e treinamento. O botão terá a palavra *Desenvolvimento e Treinamento*.
- f. ofertas de trabalhos e serviços. O botão terá a palavra *Mercado de Trabalho*.
- g. fale conosco. O botão terá a palavra *Fale Conosco*.

7.4.1.3. Na parte inferior de cada página, teremos sete menus organizados verticalmente, cada um dos sete menus posicionados ao lado do outro horizontalmente. Cada um dos menus terá palavras ou frase sublinhadas, que serão acessados por um clique do mouse. Com isso, se poderá acessar partes específicas de cada portal.

- a. "OpenMoCap.org"
 - a.1. "História do OpenMoCap"
 - a.2. "Finalidades"
 - a.3. "Créditos do Software e do Portal"
- b. "Artigos e Novidades"
 - b.1. "Artigos"
 - b.2. "Novidades"
- c. "Galeria"
 - c.1. "Vídeos"
 - c.2. "Imagens"
- d. "Comunidade"

- d.1. “Fóruns”
- d.2. “Chats”
- e. “Desenvolvimento e Treinamento”
 - e.1. “Seção de desenvolvedores”
 - e.2. “Tutoriais”
 - e.3. “Manual do usuário”
- f. “Mercado de Trabalho”
 - f.1. “Oferta de serviços”
 - f.2. “Oferta de trabalhos”
- g. “Download”
- h. Botões com os ícones de sites, como o Facebook, Tweeter, Vimeo, acesso a RSS etc.

7.4.1.4. No canto inferior da página, três termos, separadas por barras verticais, para acesso com um clique de mouse:

- a. “Mapa do site”
- b. “Termos de uso e privacidade”
- c. “Créditos”

7.4.1.5. Na última seção horizontal inferior da página, centralizadas lado a lado, as logos da Universidade FUMEC, da FUNARTE e do Ministério da Cultura. O da FAPEMIG entrará caso o projeto de extensão seja aprovado.

7.4.2. Páginas principais

7.4.2.1. Homepage

a. Na parte superior, abaixo dos frames fixos, uma imagem alusiva à captura de movimento.

a.1. Na parte inferior da imagem, a três-quartos de altura, a sentença *“OpenMoCap é o primeiro software livre do mundo em funcionamento e o primeiro software de captura de movimento para animação de personagens desenvolvido no Brasil”*.

a.2. No canto inferior direito da imagem, a frase *“Baixe e contribua com o seu desenvolvimento”*. À direita da frase, uma

seta, com o link para download, acessível por um clique de mouse.

b. Abaixo da imagem, ícones com um frame de cada um dos cinco vídeos mais recentes da galeria, arranjados horizontalmente, que podem ser acessados por um clique de mouse sobre o frame.

c. Duas colunas horizontais, com textos fixos:

c.1. à esquerda, com o título “Objetivos”, texto com as finalidades do site.

c.2. à direita, com o título “História”, texto com a história de desenvolvimento do software e do site.

7.4.2.2. Artigos e novidades:

a. Repetição de 1.1. e 1.2. da homepage.

b. Duas colunas horizontais, com links de um clique de mouse ao conteúdo, em HTML, de textos:

b.1. à esquerda, com o título “Artigos”, títulos dos artigos sublinhados, organizados verticalmente, com acesso com um clique, e, abaixo de cada título, o resumo do artigo.

b.2. à direita, com o título “Novidades”, com o título de cada novidade, organizadas verticalmente, acessíveis com um clique, com o resumo da novidade.

7.4.2.3. Galeria:

Nesta seção, os vídeos serão carregados no Vimeo, com links a partir desta página ou da homepage, e as imagens estarão nos servidores da FUMEC. Cada título terá um thumbnail da mesma, com o título e o resumo do conteúdo, acessível com um clique no frame ou no título. Serão organizados em duas colunas.

a. à esquerda, os vídeos.

b. à direita, as imagens.

7.4.2.4. Comunidade:

a. O acesso ao conteúdo será feito através de login, Na parte superior, abaixo de 1.1. e 1.2., o botão de login, centralizado, e um texto curto sobre os fóruns e o chat. Ao clicar no botão de login, o usuário será dirigido à página de *login* ou de *signup*.

b. Abaixo de a., duas colunas:

b.1. à esquerda, sobre os fóruns.

b.2. à direita, sobre o chat.

7.4.2.5. Desenvolvimento e treinamento:

a. Com o título “DESENVOLVIMENTO DO OPENMOCAP”, texto sobre a importância do software de código aberto e o seu desenvolvimento coletivo pela comunidade de programadores voluntários, dividido em duas colunas. Abaixo, em um botão com a frase “Acesso ao Ambiente de Desenvolvedores”, se acessa o código-fonte e o fórum de desenvolvedores.

b. Abaixo do botão “Acesso ao Ambiente de Desenvolvedores”, o título “TUTORIAIS” e texto em duas colunas sobre os tutoriais.

7.4.2.6. Ofertas de serviços e trabalhos:

a. Com o título de “Oportunidades”, listagem com ofertas de trabalhos por anunciantes.

b. Com o título de “Profissionais”, listagem com ofertas de serviços por anunciantes, com seus mini-currículos.

7.4.2.7. Fale Conosco:

a. Página com espaços com preenchimento sobre os dados do usuário e o texto que ele quer dirigir aos administradores do site.

7.4.3. Páginas acessórias

7.4.3.1. À seção “Artigos e novidades”: páginas HTML com os textos completos linkados na página principal.

7.4.3.2. À seção “Galeria”: páginas HTML com as fotos e o cabeçalho da mesmas.

7.4.3.3. À seção “Comunidade”: página de login ou signup. Ambiente de chat e fórum.

7.4.3.4. À seção “Desenvolvimento e treinamento”: página de login ou signup para desenvolvedores. Ambiente de fórum para desenvolvedores.

7.4.3.5. Página específica para download do programa, com informações gerais sobre a construção do programa e o sistema operacional adotado.

Na Imagem 102 é mostrada a homepage do portal.

OPENMOCAP Portal para Produção de Animações de Personagens e Realidade Aumentada com a Captura de Movimento

Login Idioma:

Galeria Serviços Artigos Desenvolvimento Fóruns Fale Conosco

OPENMOCAP

OpenMoCap é o primeiro software no mundo de código livre em funcionamento e o primeiro software de captura de movimento para animação de personagens desenvolvido no Brasil.

Baixe agora e contribua com o desenvolvimento do código.

Últimas notícias do Laboratório de Mocap

- | Lorem ipsum dolor sit amet, consectetur adipiscing elit. 25 de Janeiro de 2011
- | Vestibulum a est vel tellus convallis porttitor. 23 de Janeiro de 2011
- | Donec a metus nec est lacinia dapibus. 20 de Janeiro de 2011
- | Mais...

Artigos da Comunidade OpenMoCap

- | Etiam at felis augue, vel feugiat magna. 25 de Janeiro de 2011
- | Nulla condimentum laoreet tempus. 21 de Janeiro de 2011
- | Nunc posuere nulla nec eros vulputate aliquam. 15 de Janeiro de 2011
- | Mais...

Fóruns da Comunidade OpenMoCap

- | Ut sed nisi risus, in cursus urna. 25 de Janeiro de 2011
- | Nulla congue vehicula turpis mattis malesuada. 21 de Janeiro de 2011
- | Maecenas a ante mi, in facilis mauris. 15 de Janeiro de 2011
- | Mais...

OpenMocCap.org | Galeria | Serviços | Artigos | Desenvolvimento | Fóruns | Fale Conosco | Fique conectado

| Quisque | Venenatis | Fusce | Morbi | Maecenas | Suspendisse | Aliquam |

| Rhoncus | Aliquet | Praesent | Phasellus | Aenean | Etiam | Duis

| Nunc | Aenean | Vestibulum | Proin | Mauris | Mauris | Pellentesque

| Dictum | Curae | Class

Mapa do site | Termos de uso e privacidade | Créditos

FUNDAÇÃO NACIONAL DE ARTES **funarte** Ministério da Cultura

Imagem 102: homepage do portal.

CAPÍTULO 8

PRODUÇÕES AUDIOVISUAIS COM A CAPTURA DE MOVIMENTO

8.1. Introdução

Três produtos audiovisuais foram desenvolvidos para testar o OpenMoCap e para utilizar a captura de movimento. Todos eles estão com os cenários e os personagens desenvolvidos, modelados e com a estrutura pronta para animar. Em um deles já foi feito um vídeo com a captura de movimento.

O primeiro produto é um jogo, baseado no filme *Viagem à Lua*, de Georges Méliès. O segundo é um curta-metragem, desenvolvido em parceria com o Grupo Teatral Galpão. Este trabalho permitirá aprofundar os estudos de captura de performance. O terceiro trabalho são inserções de trinta segundos nos episódios do programa infantil *Dango Balango*, da Rede Minas de Televisão. Este programa é veiculado nacionalmente pelo Canal Brasil. A princípio, ele será o único produto que utilizará um sistema comercial, da [Natural Point 2011], porque as exigências de prazo demandam um sistema mais estável que o OpenMoCap. A seguir, são detalhados cada um destes produtos.

Neste capítulo são apresentados o documento de desenvolvimento do jogo, o primeiro roteiro para o programa *Dango Balango* e a estrutura do curta-metragem. Também são mostrados alguns dos conceitos de arte dos personagens e cenários desses produtos.

8.2. Jogo inspirado no filme *Viagem à Lua*

Este projeto é de um jogo inspirado no primeiro filme de ficção científica do cinema, *Viagem à Lua*, de Georges Méliès. Méliès (1861-1938) talvez seja o cineasta mais inovador da história do cinema. A narrativa do filme tem todos os

elementos para a construção do jogo: personagens, cenários e *game levels*. A transformação desse filme em um game é uma ideia original, pois, apesar desse filme ser um ícone na história do cinema e ser sempre referenciado e reverenciado, nenhuma iniciativa para transformá-lo em um jogo digital ocorreu até hoje. Na Imagem 103 são mostradas algumas imagens de trechos do filme.

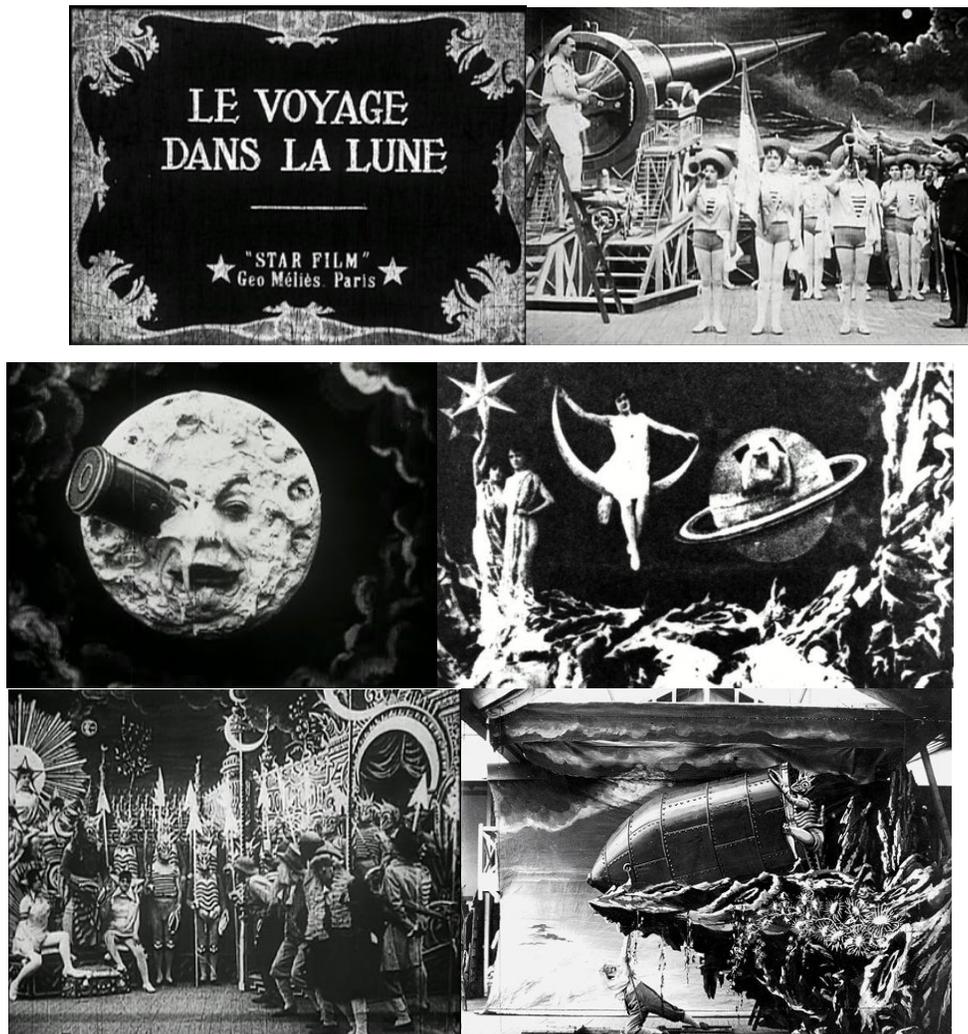


Imagem 103: cenas do filme Viagem à Lua.

A *cut scene*, que é a sequência animada explicando o funcionamento do jogo, antes dele ser iniciado, está sendo animada no momento. Na Imagem 104 é exibido um primeiro render da cena em que os astrônomos discutem o projeto de lançar o foguete com um canhão, juntamente com um frame da cena

original. Todos os personagens já possuem o esqueleto para receber os dados de movimento.



Imagem 104: cena original do filme e primeira versão modelada do jogo.

O jogo começa a partir da *cut scene*, e a sua ação acontece com a chegada à lua e com o astrônomo principal, Barbenfoullis, se encontrando com os seres da lua, os Selenitas. O jogo está sendo desenvolvido para a plataforma iOS, que é o sistema operacional do iPhone e iPad. O game foi modelado com o programa [Blender 2011] e sua programação está sendo feita com o engine SIO2, que faz a integração entre o software de modelagem e o

sistema operacional [SIO2 2011]. Nos próximos parágrafos é apresentado o documento de desenho do jogo (*game design document*, GDD).

Na Imagem 105 é exibido o personagem principal do jogo, Barbenfoullis e, na imagem 106, o Selenita, ambos inspirados no filme de Méliès.



Imagem 105: personagem Barbenfoullis, modelado com o Blender.

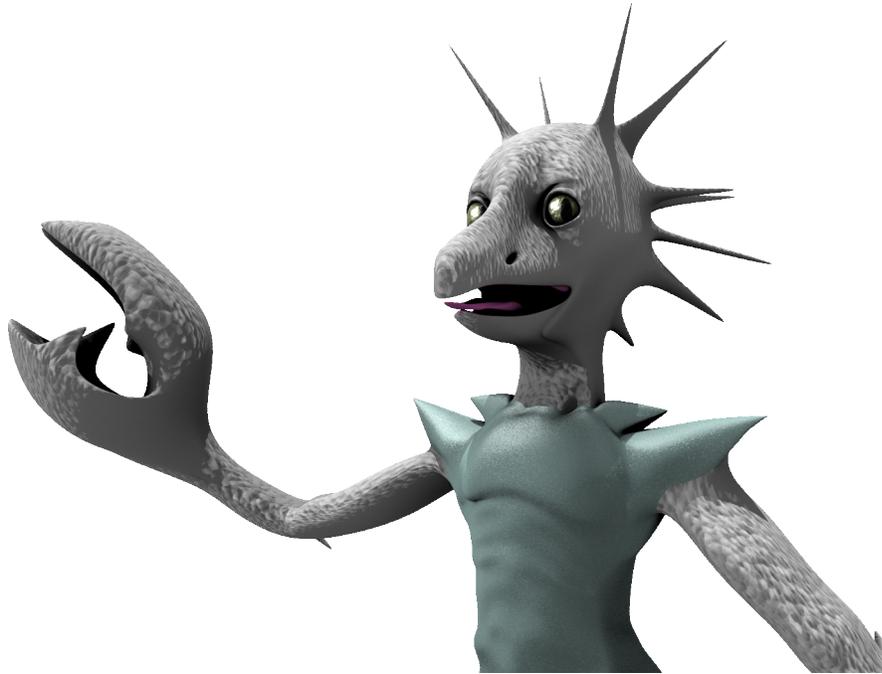


Imagem 106: Selenita, modelado com o Blender.

8.2.1 - Visão Geral

Viagem à Lua é um jogo casual inspirado nos primórdios da ficção científica. É uma homenagem à visão romântica de futuro sugerida por nomes como Georges Méliès, ilusionista francês e um dos percussores do cinema, Júlio Verne, autor francês, e H.G. Wells, autor inglês, expoentes que marcaram a história do gênero.

Em seu filme *Le Voyage dans la Lune*, Méliès narra as aventuras do Professor Barbenfoullis e seus colegas da Liga dos Astrônomos, que se lançam à Lua a bordo de uma cápsula espacial, e lá se deparam com seres muito peculiares e um tanto hostis, os Selenitas. A premissa é muito similar à proposta por Verne, em seu romance *From the Earth to the Moon*, e por Wells, em *The First Men in the Moon*.

O jogo é concebido para dispositivos móveis dotados de interface *touchscreen* e acelerômetro, com o sistema operacional iOS (ex.: iPhone, iPod Touch, iPad), buscando inspiração na surreal temática sugerida para realizar uma releitura divertida do que foi o surgimento da ficção científica. O jogo, que

pode também se chamar *Barbenfoullis Vs. Selenitas*, é um game de ação frenética, em um ambiente com estética inspirada no filme de Méliès.

8.2.1.1 - Conceito do Jogo

O jogador assume o papel do Professor Barbenfoullis e deve sobreviver ao ataque dos Selenitas, progredindo através de vários ambientes lunares até o confronto final com o Rei dos Selenitas.

8.2.1.2 - Escopo do Projeto

a - Locações

O jogo se passa na Lua, mas a Lua de Méliès, surreal e um tanto psicodélica. Sua superfície é montanhosa e acidentada. As estrelas no céu parecem ter vida e passam a sensação de uma espécie de plateia brilhante e observadora.

b - Estilo de Jogo

O jogo será de plataforma (do tipo *side scrolling*, com câmera lateral e rolagem horizontal do cenário) e terá três modos distintos:

- *Adventure*: jogador deve progredir pelas quatro fases e, assim, tentar voltar para a Terra.
- *Time Attack*: jogador deve enfrentar desafios baseados no tempo para prosseguir. Cada desafio será especificado no início de cada fase contida neste modo.
- *Survival*: o jogador deve enfrentar enormes quantidades de Selenitas simultaneamente, para sobreviver em vários níveis diferentes de dificuldades.

8.2.2 - Jogabilidade e Mecânica

A seguir são descritos os vários elementos do jogo, que permitem a sua jogabilidade e as suas ações.

8.2.2.1 - Medidores

a - Vitalidade

O medidor de vitalidade tem 03 pontos, e indica o quanto Barbenfouillis consegue resistir aos ataques dos Selenitas. Cada vez que Barbenfouillis é atingido por um dos Selenitas, seu medidor de Vitalidade é reduzido em um ponto. Quando o medidor é completamente esvaziado, Barbenfouillis é vencido e levado ao líder dos Selenitas e o jogo termina. Aleatoriamente, ao abater um dos Selenitas, este deixará cair um item que irá repor em um ponto a energia de Barbenfouillis.

b - Força de Vontade

O medidor de Força de Vontade reflete a adrenalina de Barbenfouillis, e possui 05 pontos. Sempre que o jogador abater um Selenita, o medidor de confiança será acrescido de 01 ponto. Quando estiver repleto, o medidor piscará para informar ao jogador que ele pode realizar um ataque especial, que poderá ser acionado com determinado comando.

8.2.2.2 - Tela de pontuação

A tela de pontuação pode ser acessada através do menu principal do jogo, e contém informações relativas às estatísticas do jogador, os desafios concluídos e a pontuação máxima obtida em cada modalidade de jogo.

- Cristais coletados #x 40
- Selenitas abatidos #
- Desafios completados # x 5
- Pontuação máxima em modo Time Attack #
- Pontuação máxima em modo Survival #

8.2.2.3 - Desafios

Os desafios serão concluídos quando o jogador finalizar as fases, cumprindo determinados requisitos. Por exemplo, ele pode passar por toda a primeira fase sem levar qualquer dano. Cada desafio será citado no início de

cada fase e, realizando com sucesso, o jogador será recompensado com uma maior pontuação ao fim de cada fase.

- Superfície Lunar – completar a fase sem levar qualquer tipo de dano.
- Cavernas Ocultas – completar a fase sem usar o poder especial da *Força de Vontade*.
- O Lago Branco – completar a fase em menos de 05 minutos.
- O Castelo – completar a fase sem levar qualquer dado e sem usar o poder especial
- A Batalha Final - derrotar o Rei dos Selenitas sem levar qualquer tipo de dano e sem usar o poder especial da *Força de Vontade*.

8.2.2.4 - Tutorial de Jogabilidade

Se o jogador optar por passar pelo tutorial, o jogo será carregado, e, nos momentos designados, as mensagens instrucionais serão transmitidas ao jogador. As mensagens são transmitidas por uma garota acomodada despojadamente sobre um corpo celeste, uma Plêiade, que servirá de guia a Barbenfouillis durante sua jornada.

a - Lista de mensagens

- "Utilize as setas para movimentar Barbenfouillis ": aparece assim que o jogo se inicia. Os botões de movimento irão piscar e a mensagem desaparecerá quando o jogador se movimentar para os dois lados.
- "Utilize a seta ABAIXO para se defender": após a mensagem acima, o botão-seta ABAIXO irá piscar,e a mensagem desaparecerá quando o jogador se agachar.
- "Toque neste botão para atacar": surge juntamente com o primeiro Selenita a surgir no cenário. O botão de ataque irá piscar e a mensagem some assim que o jogador abater o primeiro selenita.
- "Toque neste botão para saltar": assim que o jogador se deparar com um obstáculo a ser saltado, o botão de salto irá piscar e a mensagem desaparecerá quando o jogador vencer o obstáculo.

- "Fique atento ao seu medidor de Vitalidade!": surge quando o jogador for atacado e perder seu primeiro ponto de vitalidade.
- "Seu medidor de Força de Vontade está cheio. Toque o botão de ataque rapidamente duas vezes para um ataque especial": quando o jogador abater cinco Selenitas e tão logo o medidor comece a piscar.
- "Toque no botão para saltar duas vezes rapidamente, e mantenha pressionado para planar": na primeira vez que o jogador se deparar com uma vala a ser transposta, o botão de salto irá piscar e a mensagem desaparece quando o jogador vencer o obstáculo.

8.2.2.5 - Ações de Barbenfouillis

a - Caminhar e Correr

A movimentação do personagem se dará pelo acionamento de botões de toque presentes no HUD. Haverá controles de movimentação em cruz: para avançar ou retroceder no cenário, utiliza-se as setas DIREITA e ESQUERDA, como ocorre tipicamente em jogos para console com este tipo de jogabilidade (ex.: Super Mario, Trine). Opcionalmente, o personagem poderá ter duas velocidades de movimentação, Caminhar e Correr. Para correr, o jogador pressionará duas vezes rapidamente o botão-seta correspondente à direção desejada.

b - Ataques

Os ataques do personagem serão ativados através de toque no botão ATAQUE na tela do dispositivo móvel.

b - Ataque normal

Em seu ataque básico Barbenfouillis golpeia uma vez com o guarda-chuva, na direção em que o personagem estiver apontando no cenário.

c - Ataque especial

Com o medidor de Força de Vontade repleto, tocar duas vezes rapidamente o botão ATAQUE faz com que Barbenfouillis comece a girar sobre

seu próprio eixo como um peão, com o guarda-chuva apontado para frente, atingindo todos os inimigos à sua volta.

d - Defesa

Pressionando o botão-seta PARA BAIXO, Barbenfouillis se agacha, abrindo seu guarda-chuva e se escondendo sob o mesmo. Desta forma ele utiliza o guarda-chuva como uma espécie de escudo, defletindo os ataques dos inimigos. Enquanto estiver defendendo, Barbenfouillis não pode atacar ou se movimentar.

e - Saltar e Planar

Pressionando o botão-seta SALTAR, Barbenfouillis salta para cima. O botão SALTAR é combinado com as setas direcionais ESQUERDA e DIREITA, para que Barbenfouillis consiga transpor obstáculos simples. Acrescentando à jogabilidade, alguns obstáculos podem exigir que Barbenfouillis esteja correndo, para ter impulso suficiente para transpor um obstáculo.

Pressionando o botão SALTAR uma segunda vez, enquanto no ar, faz com que Barbenfouillis abra seu guarda-chuva, utilizando-o como uma espécie de planador e permitindo que seu salto alcance uma distância maior.

8.2.2.6 - Derrota

Quando o medidor de Vitalidade de Barbenfouillis estiver vazio, e ele for atingido por um ataque inimigo, Barbenfouillis dá uma pirueta no mesmo lugar e cai de costas no chão, com as pernas para cima. Uma curta animação é reproduzida, mostrando Barbenfouillis sendo levado até o Rei dos Selenitas, seguida de um menu com as opções *Tentar de novo* e *Sair*.

- Tentar de novo - O jogo é recarregado a partir do último ponto de checagem.
- Sair - A partida é finalizada e o aplicativo retorna ao menu inicial.

8.2.3 - Personagens

8.2.3.1 - Personagens Não-Jogáveis (NPCs)

a - Cientistas

Os colegas de Barbenfouillis, da Liga dos Astrônomos, são seqüestrados pelos Selenitas logo no início do jogo. São eles *Nostradamus*, *Alcofrisbas*, *Omega*, *Micromegas* e *Parafaragamus*.

- Aparência:

Possuem uma caracterização similar à do protagonista. Aparecem amarrados na sala do trono do Rei dos Selenitas.

- Comportamento:

São passivos e não possuem ataques.

b - Selenitas

Os alienígenas hostis que deverão ser enfrentados por Barbenfouillis.

- Aparência:

São seres de aspecto estranho, como uma mistura bípede de ave e réptil, que aparentam serem feitos de *poeira cósmica*, pois se esvaem em uma nuvem de pó ao serem abatidos. Seus movimentos são peculiares e acrobáticos, o que denota ainda mais a sua estranheza.

- Comportamento:

Quando não estão em estado de alerta, os Selenitas patrulham o seu território atentamente e demonstram uma série de comportamentos estranhos, dando cambalhotas ou andando sobre suas garras, plantando bananeiras.

- Ataques:

Existem duas classes de Selenitas, que determinam seus possíveis ataques:

1. Soldados: quando alertados da presença de Barbenfouillis, os Soldados atacam com suas enormes garras e não usam armas. Eles adotam estratégias de ataque, como emboscadas ou, até mesmo dependendo da situação, um ataque frenético em massa, desorganizado mas letal. Podem surpreender Barbenfouillis saindo de buracos escondidos no solo.

2. Elite: quando alertados, esta classe de Selenitas age de maneira mais organizada e utiliza uma afiada lança para atacar à distância e se defender.

c - O Rei dos Selenitas

Inimigo final do jogo, o Rei dos Selenitas aguarda por Barbenfouillis na sala do trono, onde é protegido pela elite de sua guarda.

- Aparência:

O Rei dos Selenitas se parece com os alienígenas que lidera, mas é uma enorme entidade de poeira brilhante, sentado em um imponente trono, cercado pelas belas Plêiades.

- Comportamento:

A princípio, o Rei dos Selenitas permanece sentado em seu trono, observando apaticamente à luta entre Barbenfouillis e a elite de sua guarda. Quando todos os Selenitas são vencidos, no entanto, ele se levanta para o confronto final.

- Ataques:

O Rei dos Selenitas tem habilidades especiais: pode se teleportar (desaparecer em um lugar e aparecer em outro) e esticar seus membros para atacar Barbenfouillis.

8.2.4 - Cenários

O jogador irá se aventurar por quatro ambientes. Cada um deles representa um Ato no modo *Adventure*.

8.2.4.1 - Ato 1 - A Superfície Lunar

A superfície Lunar é acidentada e cinza, repleta de reentrâncias e crateras de impacto, e recoberta por uma camada de "poeira cósmica", que parece nevar periodicamente. Gêiseres de fumaça se manifestam com frequência, entre o que parecem ser estalagmites de rocha brilhante. No céu, estrelado e colorido em tons violáceos, o planeta Terra alterna aparições com as brilhantes Plêiades. É na Superfície Lunar que a cápsula dos cientistas aterrissa, e onde os colegas de Barbenfouillis são seqüestrados pelos Selenitas

8.2.4.2 - Ato 2 - As Cavernas Ocultas

Abaixo de sua superfície há amplas galerias de cavernas perigosas, que parecem se aprofundar até o centro da Lua. A rocha brilha à luz de aglomerados de cristais em cores frias, com estalactites e estalagmites, intercalando-se com coloridos cogumelos gigantes. Os caminhos são estreitos e recortados de perigos, com pontes de pedra, plataformas e partes escorregadias. Um rio de poeira cósmica deságua em uma enorme e colorida cascata, que some na escuridão.

8.2.4.3 - Ato 3 - O Lago Branco

No centro da Lua, uma imensa galeria abriga um belo lago de poeira cósmica, refletindo magicamente o ambiente ao redor. É um ambiente mais escuro e a luz vem do topo, por fachos entre as reentrâncias, e de minérios incrustados nas rochas.

8.2.4.4 -Ato 4 - O Castelo

Às margens do Lago Branco encontra-se o Castelo do Rei Selenita, um imponente monumento incrustado de pedras preciosas. Dentro dele, amplos salões exageradamente adornados abrigam a realeza Selenita.

8.2.4.5 - Ato Final - A Sala do Trono

Na sala do trono, o Rei dos Selenitas é cercado de uma pompa pitoresca, cercado de belas Plêiades e de sua Guarda de Elite. O amplo salão é exageradamente adornado em pedras preciosas e detalhes arquitetônicos.

8.2.5 - Passo-a-passo do modo *Adventure*

8.2.5.1 - Cena de Introdução

Os cientistas entram em uma cápsula em forma de projétil, ao som de música e de muita festa. Eles são atirados por um enorme canhão em direção à Lua, que, ao se aproximar, revela ter uma face e expressão. O projétil acerta a Lua no olho, e, em um corte de cena, vemos o projétil já aterrissado na superfície lunar. Os cientistas descem da cápsula, um a um, e o último a sair é Barbenfouillis. Ele esquece seu guarda-chuva na cápsula e volta para buscá-lo. Enquanto ele está lá dentro, vemos o grupo de astrônomos serem levados por

uma turba de Selenitas. Quando Barbenfouillis sai da cápsula novamente, fica sem entender para onde foram seus colegas.

8.2.5.2 - Objetivos

a - Ato 1 - A Superfície Lunar

O jogador assume o controle de Barbenfouillis e deve avançar pela superfície lunar, superando vários Selenitas e obstáculos, até encontrar a entrada para as Cavernas.

b - Ato 2 - As Cavernas Ocultas

O jogador progride pela galeria de cavernas, enfrentando os inimigos e obstáculos, até encontrar o Lago Branco.

c - Ato 3 - O Lago Branco

O jogador deve vencer diversos desafios até alcançar o Castelo.

d - Ato 4 - O Castelo

O jogador avança pelos salões do Castelo, vencendo obstáculos e inimigos, até alcançar o Rei dos Selenitas.

e - Ato 5 - A Batalha Final

Barbenfouillis deve enfrentar e vencer o Rei dos Selenitas, libertar seus colegas e, assim, conseguir voltar para a Terra.

8.2.5.3 - Cena de Conclusão

Barbenfouillis e seus colegas são vistos correndo desvairadamente através dos 4 cenários, seguidos por uma turba de Selenitas furiosos. Eles conseguem enfim chegar à cápsula que os trouxe à Lua, e Barbenfouillis se pendura na cápsula para que ela possa cair em direção a Terra. A cápsula cai no mar, sonoramente, e um corte leva à cena em que Barbenfouillis e os cientistas são homenageados pela conclusão de sua aventura com uma grande festa.

8.2.6 - Interface

8.2.6.1 - H.U.D.

O conjunto de medidores será exibido graficamente na tela durante todo o jogo, informando sobre pontos de *Vitalidade*, *Força de Vontade*, a pontuação do jogador e o tempo transcorrido, na parte superior da tela.

Na parte inferior da tela encontram-se os botões de movimentação, salto e ataque, acionados por toque.

8.2.6.2 Menus

As opções do jogador serão registradas em telas acessadas por menus contextuais conforme abaixo.

1 - Jogar

1.1 - Campanha

1.1.1 - Novo jogo

1.1.2 - Continuar

1.2 - Time Attack

1.2.1 - Escolha do cenário

1.3 - Survival

1.3.1 - Escolha do cenário

2 - Tutorial

3 - Configurações

3.1 - Sensibilidade dos controles

3.2 - Volume da música

3.3 - Volume dos efeitos sonoros

4 - Pontuação

5 - Créditos e versão do jogo

6 – Sair

8.3. Curta-metragem *A Primeira Perda de Minha Vida*

8.3.1. Conceito do projeto

Essa história será contada em animação. A linguagem da animação é muito apropriada para o roteiro sensível criado por Eduardo Moreira, diretor e ator do grupo teatral Galpão e também diretor desse curta. O roteiro é uma adaptação livre de uma história acontecida com o escritor tcheco Franz Kafka, narrada em diversas ocasiões por terceiros, como por Paul Auster no livro *The Brooklyn Follies*.

Kafka, nos seus passeios ao lado de sua companheira por uma praça de Berlim, nos últimos meses de sua vida, se depara com uma menina aos prantos. Ao interrogá-la, a criança conta que perdeu sua boneca. O escritor diz então que a boneca viajou de férias pelo mundo e começa a escrever cartas imaginárias, como se fossem dela para a menina, narrando as suas viagens. Passa então a encontrar-se com a menina na praça para ler essas cartas. Isso dura duas ou três semanas. Até que a boneca se despede e diz que não voltará mais. Como observa Eduardo Moreira, “a essa altura, claro, a menina não sente mais falta da boneca. Kafka lhe deu em troca uma história e essa história curou sua infelicidade.” [Moreira 2007]

A ação se desenvolve em cinco locações diferentes, Berlim, uma praia, Paris, Olinda e uma igreja. O figurino e as locações foram escolhidos para que a história possa se passar tanto em 1924 como agora. Os cenários serão modelados digitalmente, assim como os personagens. As texturas e os materiais dos cenários são inspirados nas obras do artista plástico e arquiteto austríaco Friedensreich Hundertwasser.

A animação dos personagens será feita através da captura digital de movimento (*mocap*) de atores. É importante ressaltar que o sistema de captura de movimento para animação foi desenvolvido, de ponta a ponta, por parte da equipe técnica desse curta-metragem. Será o primeiro filme realizado no Brasil com captura de movimento para animação.

O formato de curta-metragem é muito adequado a essa proposta. O roteiro e a técnica de animação desenvolvida demandam, nesse momento,

uma obra audiovisual de curta duração para se concretizarem. O projeto será realizado com a resolução de 1920 por 1080 pixels, a 23.96 quadros por segundo, e poderá ser finalizado e transferido para película 35 mm ou hdcam.

“Quando uma pessoa tem a sorte suficiente de viver dentro de uma história, de viver dentro de um mundo imaginário, as dores desse mundo somem. Pelo tempo que durar a história, a realidade deixa de existir.” *Paul Auster* [Moreira 2007]

8.3.2. Apresentação dos personagens

Os personagens principais são três, o escritor (Franz Kafka), a menina e a boneca, e a história tem um personagem coadjuvante, o boneco gigante de Olinda. Os atores Julio Maciel, Bárbara, Inês Peixoto e Beto Franco, respectivamente, do Grupo Teatral Galpão, serão os atores que darão vida aos personagens. Os personagens serão modelados a partir dos atores e os seus movimentos para a animação também serão capturados dos atores. Os atores, com marcadores em seus corpos, encenarão seus papéis em um set com diversas câmeras. As posições ao longo dos frames dos marcadores serão aplicadas em pontos dos corpos dos personagens modelados. Assim, os movimentos dos atores animarão os personagens. A caracterização dos personagens será feita de tal maneira que possa tanto se passar em 1924, data provável do ocorrido, ou agora.

O escritor em seus últimos meses de vida, é um homem maduro de quarenta anos, cabelos negros, levemente grisalhos, e estatura mediana. Está em Berlim desde 1923, com tuberculose. Em seus passeios pelo Jardim Botânico de Berlim depara-se com uma menina aos prantos. A criança tem no máximo cinco anos. É curiosa, inteligente e tem um olhar arrebatador. A partir daí, se desenvolve um relacionamento paternal entre o escritor e a menina, nos seus encontros para a leitura das cartas imaginárias da boneca. Isso acontece durante três semanas. Os encontros são em torno de um banco da praça.

A menina perdera sua boneca, a terceira personagem dessa história. É uma boneca de pano irresistivelmente charmosa e espevitada. Ela aparecerá em diversas situações diferentes, nas narrações das cartas do escritor, *ghost-writer* da boneca nas suas viagens pelo mundo. Viajará para uma praia, para

Paris e Olinda, onde também se casará em uma igreja, voará de balão. Em sua viagem a Olinda, encontra o boneco gigante, por quem se apaixona perdidamente e se casa. O boneco gigante será modelado a partir do ator Beto Franco e terá os movimentos baseados nas procissões carnavalescas. No decorrer das aventuras e peripécias relatadas nas cartas, a menina supera a perda da sua amiga predileta e já não se sente infeliz.

O realismo dos movimentos animados por captura de movimento será moderado pela estrutura física dos personagens e seu figurino, baseados em esculturas de cerâmica. Os cenários, com suas texturas e cores inspiradas no artista e arquiteto Hundertwasser, se harmoniza, com essa proposta. A combinação dos elementos deverá resultar em um filme curta-metragem de animação sensível e plasticamente bonito.

Toda a concepção artística dos personagens e a modelagem dos cenários estão prontas. Na Imagem 107 é mostrado o escritor com o Jardim Botânico ao fundo.



Imagem 107: escritor com o cenário principal ao fundo.

8.4. Programa infantil *Dango Balango*

A Rede Minas de Televisão solicitou inserções de animação de trinta segundos para cada um dos episódios da temporada de 2012 do programa infantil *Dango Balango*. Este pedido veio depois que foi apresentado, à direção

do programa, um vídeo da música-tema do infantil, com a animação dos personagens feita com a captura de movimento. São três personagens modelados: Joduca, Sdrufs e Ziizi. Na Imagem 108 é mostrado um frame desta animação, modelada com o programa de código aberto [Blender 2011]. Está ainda em fase de discussão as condições em que será feito este trabalho.



Imagem 108: frame da animação com captura de movimento realizada para o programa *Dango Balango*.

CONCLUSÕES

Objetivos alcançados

Todos os objetivos propostos para esta tese foram alcançados. O desenvolvimento do sistema de captura de movimento para a animação de personagens, idealizado em 2001, foi concluído com sucesso. O portal da web sobre animação e realidade aumentada, que permitirá o download do OpenMoCap e o acesso ao seu código-fonte, está em fase de conclusão, com lançamento previsto para setembro. A sistematização do conhecimento da utilização e do funcionamento da captura de movimento, dentro do âmbito dos efeitos visuais e da imagem digital, gerou um livro, que será publicado pela editora Edições Cine-Clube de Avanca, de Portugal.

Cada uma das etapas para se chegar a este resultado foi percorrida de forma sistemática e bem sucedida. Primeiramente, a preparação para o início do trabalho foi realizada com a minha dissertação de mestrado e a apresentação de propostas de projetos para os editais universais do CNPq e da FAPEMIG. A defesa da dissertação coincidiu com a aprovação dos editais universais, que tiveram duração de dois anos e iniciaram em 2007, no Núcleo de Processamento Digital de Imagens, sob a coordenação do prof. Arnaldo de Albuquerque Araújo. Com isto, foi possível comprar equipamentos e ter a colaboração de alunos de mestrado e de iniciação científica, para colaborar no desenvolvimento do sistema de captura de movimento. Ao longo destes dois anos, foram apresentados trabalhos para congressos e revistas, que permitiram o aprofundamento da abordagem relacionada aos efeitos visuais e à captura de movimento.

O software de captura de movimento, que inicialmente teria uma finalidade comercial, foi construído com ferramentas de código aberto, o que permitiu a sua disponibilização gratuita. A dissertação de mestrado de David Lunardi Flam, orientado por mim e pelo prof. Arnaldo de Albuquerque Araújo, defendida em 2009, foi fundamental para o desenvolvimento do código-fonte. O código é modular e de arquitetura flexível, facilitando a contribuição de outros

desenvolvedores para o seu amadurecimento. Os seus módulos podem ser modificados e aprimorados, permitindo a sua expansão para outras aplicações, além da animação de personagens. Isto também gera a expectativa de que a captura de movimento poderá um dia ser feita sem marcadores com o OpenMoCap.

O software livre e de código aberto atende a uma imensa demanda por um sistema de captura de movimento de baixo custo e eficiente. O desejo de expandir o seu uso e disseminar o conhecimento relacionado a este tema, no contexto dos efeitos visuais e da pós-produção, provocou a construção do portal da web. Para isto, novamente foram buscados recursos de órgãos de fomento, devido à amplitude do trabalho que deveria ser realizado. A FAPEMIG, novamente, e a FUNARTE apoiaram financeiramente esta nova demanda, apresentada em propostas aprovadas e que estão em curso. O CNPq também deu continuidade a seu apoio ao projeto de aprimoramento do código-fonte. Desta maneira, foi possível concluir esta etapa do trabalho, com a apresentação desta tese, que deve ainda ter muitos desdobramentos.

Perspectivas e desdobramentos

Para dar continuidade a este trabalho, existe uma demanda pela manutenção do portal da web. Em suas várias seções, o portal é dinâmico e é fruto da contribuição de diversos indivíduos, que certamente participarão dos seus fóruns, galeria e desenvolvimento do código. Esta manutenção está garantida através do apoio da Universidade FUMEC e da FAPEMIG. A primeira disponibilizou os seus servidores, para hospedar o portal, e a FAPEMIG está fornecendo bolsas para colaboradores, que ajudarão a desempenhar a tarefa de manutenção do site.

Os três trabalhos que estão sendo desenvolvidos com o objetivo de utilizar o OpenMoCap, devem ser finalizados ao longo destes próximos doze meses. O jogo digital tem o apoio do programa de extensão da Universidade FUMEC e está em fase de integração e programação da sua jogabilidade, detalhada no GDD apresentado nesta tese. As inserções no programa *Dango Balango* serão possíveis através do apoio da Rede Minas de Televisão. O

desenvolvimento do curta-metragem está sendo realizado com a colaboração voluntária de alunos do curso de Jogos Digitais da Universidade FUMEC. Apesar de ter sido apresentado em dois editais, ele não conseguiu ainda o apoio necessário para ser desenvolvido de forma sustentada.

O livro sobre os efeitos visuais e imagem digital está em fase adiantada de formatação. Ele deve também trazer uma contribuição significativa no cenário da produção audiovisual, já que não existe qualquer obra do gênero em língua portuguesa.

Assim, espera-se que esta tese conduza a novos desdobramentos concretos, já que é um trabalho com um viés prático e aplicado dentro das Artes e da Comunicação.

LISTA DE TRABALHOS PUBLICADOS DECORRENTES DESTA TESE

Este trabalho foi um dos três finalistas do Prêmio Santander Ciência e Inovação, na categoria Tecnologia da Informação, Comunicação e Educação, em 2012.

1. GOMIDE, J.V.B, FLAM, D.L., e ARAÚJO, A. de A. (2011) *Recent improvements to OpenMoCap and its release to download and to development as free open source software*. Artigo convidado da International Conference Cinema Art, Technology, Communication, 2011, Avanca - Portugal. Edições Cine Clube de Avanca,, 12 páginas.
2. QUEIROZ, D. P., FLAM, D.L., GOMIDE, J.V.B. and ARAÚJO, A. de A. (2011) *Evaluation of real time tracking methods of motion capture markers*, submetido para a SBGAMES 2011.
3. QUEIROZ, Daniel Pacheco de. (2011) *Avaliação de Métodos de Rastreamento de Marcadores para um Sistema Óptico de Captura de Movimento*. Dissertação de Mestrado, Departamento de Ciência da Computação, UFMG.
4. GOMIDE, J.V.B, FLAM, D.L., NAZARIO, L.R.P e ARAÚJO, A. de A. (2010) *Development of an Open Source Motion Capture System and its Applications to Character Animation*. In: International Conference Cinema Art, Technology, Communication, 2010, Avanca - Portugal. Edições Cine Clube de Avanca, 2010, pp 157-168, julho de 2010.
5. GOMIDE, J.V.B. (2010) *Captura digital de movimento no cinema de animação*, in Diálogos entre Linguagens, pag. 267-268, editora C/ Arte, Belo Horizonte.
6. GOMIDE, J.V.B., FLAM, D.L., QUEIROZ, D.P., e ARAÚJO, A. de A. (2010) *An open source motion capture system and its applications in arts and communication*, artigo completo apresentado e publicado nos anais do World Congress on Communication and Arts, WCCA'2010, Guimarães, Portugal, abril de 2010.
7. GOMIDE, J.V.B e ARAÚJO, A. de A. (2009) *Efeitos Visuais, uma Abordagem a Partir do Tratamento Digital de Imagens*, Revista de Informática Teórica e Aplicada, volume 15, no. 1, pag. 89-116.
8. Projetos aprovado nos Editais Universais da FAPEMIG (2006 e 2010), do CNPq (2007 e 2010), no ProPIC/FUMEC (2009), no ProEXT/FUMEC (2010 e 2011) e Bolsa FUNARTE de Produção Cultural para a Internet, 2010
9. GOMIDE, J.V.B., FLAM, D.L., QUEIROZ, D.P., e ARAÚJO, A. de A. (2009) *Captura de Movimento e Animação de Personagens em Jogos*, tutorial completo do VIII Brazilian Symposium on Digital Games and Entertainment, SBGAMES 2009, Rio de Janeiro.

10. FLAM, D.L., QUEIROZ, D.P., GOMIDE, J.V.B. e ARAÚJO, A. de A. (2009) *OpenMoCap: An Open Source Software for Optical Motion Capture*, artigo completo do VIII Brazilian Symposium on Digital Games and Entertainment, SBGAMES 2009, Rio de Janeiro.
11. FLAM, D.L. (2009) *OpenMoCap: uma aplicação de código livre para a captura óptica de movimento*. Dissertação de Mestrado, Departamento de Ciência da Computação, UFMG.
12. BIGONHA, C., FLAM, D. L. ; GOMIDE, J. V. B. ; ARAUJO, A. de A. (2008) *Software de Aquisição de Dados para um Sistema de Captura de Movimentos*, anais do XXI Brazilian Symposium on Computer Graphics and Image Processing, como trabalho de iniciação científica, no Workshop for Undergraduate Works (WUW-SIBGRAPI2008).
13. GOMIDE, J.V.B. e ARAÚJO, A. de A. (2008) *Efeitos Visuais, uma Abordagem a Partir do Tratamento Digital de Imagens*, tutorial do XXI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI2008).
14. GOMIDE, J.V.B. e ARAÚJO, A. de A. (2008) *Video Effects Design*, no festival de vídeos do XXI Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAPI2008).
15. GOMIDE, J.V.B. (2006) *A captura de movimento para o cinema de animação*. Dissertação de Mestrado, Escola de Belas Artes, UFMG.

REFERÊNCIAS BIBLIOGRÁFICAS

- ABEL 2011, *making of* do comercial *Brilliance*, de *Robert Abel and Associates*, no endereço <http://www.youtube.com/watch?v=eedXpclrKCCc>, acessado em 05/06/2011.
- ACADEMY OF MOTION PICTURE ARTS AND SCIENCES 2011, vencedores do Oscar de fotografia, <http://www.oscars.org/awards/academyawards/legacy/best-pictures.html>, acessado in 28/05/2011.
- ACADEMY OF MOTION PICTURE ARTS AND SCIENCES 2010 awards winning, <http://www.oscars.org/awards/scitech/winners/2009.html>, acessado in 28/05/2011.
- ANIMAZOO 2011, endereço do sistema inercial em <http://www.animazoo.com/motion-capture-systems/igs-190-h-motion-capture-system/>, acessado em 30/06/2011.
- 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, acessado em 25/05/2011.
- APTINA 2011, história do desenvolvimento dos sensores de imagem com CMOS em http://www.aplina.com/about/history_legacy.jsp, acessado em 27/05/2011.
- ARRI 2011, documento sobre 3DLUT da Arri, um dos mais importantes fabricantes de filmadoras para cinema, em http://www.arri.de/digital_intermediate_systems/arricube.html, acessado em 04/07/2011.
- ASCENSION 2011, a página do sistema *Flock of Birds* está no endereço <http://www.ascension-tech.com/realtime/RTflockofBIRDS.php>, acessado em 05/07/2011.
- ASSOCIAÇÃO BRASILEIRA DE CINEMATOGRAFIA, com o site oficial em <http://www.abcine.org.br/>, acessado em 30/06/2011.
- AUTODESK INCORPORATION, 2011 <http://www.autodesk.com>, accessed in 28/05/2011.
- AZEVEDO, Eduardo (2003). *Computação Gráfica – Teoria e Prática*. Editora Campus, São Paulo.
- BIOLOGY PRESS, 2011, <http://users.rcn.com/jkimball.ma.ultranet/BiologyPages/V/Vision.html>, acessado em 28/05/2011.
- BLENDER FOUNDATION, 2011, <http://www.blender.org>, acessado em 28/05/2011.
- BSD license website, 2011, <http://opensource.org/licenses/bsd-license.php>, acessado em 20/05/2011.

- BRINKMANN, Ron. *The art and science of digital compositing*. (2008) New York, Elsevier, USA.
- BROWN UNIVERSITY, 2011, aplicativos sobre combinação de cores estudos com a imagem se encontram no endereço acessado em 28/06/2011, http://www.cs.brown.edu/exploratories/freeSoftware/catalogs/color_theory.html.
- Burum, S. H. (2007), *American Cinematographer Manual*, The ASC Press, USA.
- CASTRO, J., MEDINA-CARNICER, R., AND GALISTEO, A. M. (2006). Design and evaluation of a new three-dimensional motion capture system based on video. *Gait and Posture* 24, 1, 126 –129. <http://dx.doi.org/10.1016/j.gaitpost.2005.08.001>.
- CYBERGLOVES 2011, endereço da empresa desenvolvedora de luvas para mocap em <http://www.metamotion.com/hardware/motion-capture-hardware-gloves-Cybergloves.htm>, acessado em 01/07/2011.
- CRAINE, Craig. *Review: iPi Desktop Motion Capture*, document available at <http://www.3dworldmag.com/2011/01/05/review-ipi-desktop-motion-capture/>, accessed in 29/05/2011.
- DIRECTX 2011, no portal da Microsoft, no endereço <http://www.microsoft.com/download/en/details.aspx?id=35#overview>, acessado em 05/07/2011.
- DV 2011, site da revista Digital Video, que tem discussões aprofundadas sobre todos os aspectos do vídeo digital, em <http://www.dv.com>, acessado em 05/07/2011.
- EBERT, Carlos. *Cor e Cinematografia* (2009) documento eletrônico acessado em 28/06/2011, disponível em <http://www.abcine.org.br/artigos/>.
- ENCICLOPÉDIA ITAÚ CULTURAL 2012, ARTES VISUAIS, http://www.itaucultural.org.br/aplicexternas/enciclopedia_ic/index.cfm?fuseaction=termos_texto&cd_verbete=5354, acessado em 01/03/2012.
- FEYNMAN, R. P. *The Feynman Lectures on Physics* (2011) The New Millenium Edition, USA.
- FIGUEROA, P. J., LEITE, N. J., AND BARROS, R. M. L. (2003). A flexible software for tracking of markers used in human motion analysis. *Computer Methods and Programs in Biomedicine* 72, 2, 155 – 165. [http://dx.doi.org/10.1016/S0169-2607\(02\)00122-0](http://dx.doi.org/10.1016/S0169-2607(02)00122-0).
- FLAM, David Lunardi (2009). *OpenMoCap: uma aplicação de código livre para a captura óptica de movimento*. Dissertação de Mestrado, Departamento de Ciência da Computação, UFMG.
- FLAM, D.L., QUEIROZ, D.P., GOMIDE, J.V.B. e ARAÚJO, A. de A. (2009). OpenMoCap: An Open Source Software for Optical Motion Capture, Proceedings of the VIII Brazilian Symposium on Digital Games and Entertainment, SBGames, pp 153-

- 161, Rio de Janeiro, RJ, Brazil.
http://www.sbgames.org/papers/sbgames09/computing/full/cp20_09.pdf.
- FLEISCHER, Richard (2005). *Out Of The Inkwell: Max Fleischer And The Animation Revolution*. University Press of Kentucky, USA.
- FORSYTH, D. A. and PONCE, J. (2002). *Computer Vision: A Modern Approach*. Prentice Hall Professional Technical Reference.
- GOMES, Jonas e VELHO, Luiz. *Fundamentos de Computação Gráfica*. Rio de Janeiro: Editora IMPA. 2003.
- GOMIDE, J.V.B. (2006) *A captura de movimento para o cinema de animação*. Dissertação de Mestrado, Escola de Belas Artes, UFMG.
- GOMIDE, J. V. B. ; FLAM, D. L. ; NAZARIO, L. R. P. ; ARAUJO, A. A. (2010a). *Development of an Open Source Motion Capture System and its Applications to Character Animation*. In: International Conference Cinema Art, Technology, Communication, 2010, Avanca - Portugal. Edições Cine Clube de Avanca, 2010, pp 157-168.
- GOMIDE, J. V. B. ; FLAM, D. L. ; NAZARIO, L. R. P. ; ARAUJO, A. A. (2010b). *An open source motion capture system and its applications in arts and communication*. In: World Congress of Communication and Arts, 2010, Guimarães - Portugal. (vou colocar os dados de pag até terça-feira)
- GOMIDE, J.V.B., FLAM, D.L., QUEIROZ, D.P., e ARAÚJO, A. de A. (2009) *Captura de Movimento e Animação de Personagens em Jogos*, tutorial apresentado na VIII Brazilian Symposium on Digital Games and Entertainment, SBGAMES 2009, Rio de Janeiro, 2009, disponível em http://wwwusers.rdc.puc-rio.br/sbgames/09/_proceedings/dat/index_eng.html, accessed in 25/05/2011.
- HELLER, Eva (2007). *A Psicologia das Cores*. Editora Gustavo Gili, Portugal.
- INITIATION, 2011. *Motion Capture / Tracking from Inition*. This is an eletronic document available at: <http://www.inition.co.uk/inition/products.php?CatID =11>. accessed in 20/05/2011.
- GONZALEZ, Rafael C. e WOODS, Richard E. *Processamento Digital de Imagens*. São Paulo: Pearson Education do Brasil. 2010.
- INTEL CORPORATION, 2011. *OpenCV – Intel Software Network Blogs* at <http://software.intel.com/en-us/blogs/tag/opencv/?wapkw=%28opencv%29>, accessed in 25/05/2011.
- iPi SOFT website, <http://www.ipisoft.com>, accessed in 28/05/2011.
- KERNIGHAN, B. W., AND PIKE, R. (1999). *The practice of programming*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA.

- KIRSH, J.L. and KIRSH R.A. *Storing Art Images in Intelligent Computers*; documento eletrônico acessível em http://74.125.155.132/scholar?q=cache:vKCVGb78MdYJ:scholar.google.com/+Russell+Kirsch+first+photography&hl=en&as_sdt=0,5, acessado em 27/06/2011.
- KITAGAWA, M. e WINDSOR, B. (2008) *MoCap for artists: workflow and techniques for motion capture*. Focal Press, USA.
- LEE, Eric (2011). Imagem de domínio público em http://commons.wikimedia.org/wiki/File:Three-two_pulldown_diagram_%28telecine%29.png, acessado em 30/06/2011.
- LIVERMAN, Matt (2004). *The Animator's Motion Capture Guide*. Hingham, Massachusetts: Charles River Media, USA.
- LUCENA JÚNIOR, Alberto (2002), *Arte da animação, técnica e estética através da história*. Editora Senac, São Paulo.
- MAREY, E. *Animal Mechanism: A Treatise on Terrestrial and Aerial Locomotion, 1873* (2010), republished by Nabu Press, 2010.
- MCCONNELL, S. (2004). *Code Complete*, Second Edition. Microsoft Press, Redmond, WA, USA.
- MENACHE, Alberto. (2000) *Understanding Motion Capture for Computer Animation and Video Games*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA.
- MENDIBURU, Bernard (2009) *3D Movie Making*. Focal Press. USA.
- MOESLUND, T. B. & GRANUM, E. (2001). *A survey of computer vision-based human motion capture*. *Computer Vision and Image Understanding*, 81(3):231--268. <http://dx.doi.org/10.1006/cviu.2000.0897>. 1, 2, 17
- MOESLUND, T. B.; HILTON, A. & KRÜGER, V. (2006). *A survey of advances in vision-based human motion capture and analysis*. *Computer Vision and Image Understanding*, 104(2):90--126. <http://dx.doi.org/10.1016/j.cviu.2006.08.002>. 1, 17
- MOLET, T. et al. (1997) *An Animation Interface Designed for Motion Capture*, *Computer Animation 1997*.
- MOREIRA, E. L. (2007). Registro do roteiro *A Primeira Perda da Minha Vida* no Escritório de Direitos Autorais da Fundação Biblioteca Nacional, com o número 418.260.
- MOTION ANALYSIS CORPORATION, 2011. The industry leader fo3d passive optical motion capture. Documento eletrônico disponível em <http://www.motionanalysis.com/>, acessado em 26/05/2011.
- MURCH, Walter (2004). *Num Piscar de Olhos*. Jorge Zahar Editor, Brasil.
- MUYBRIDGE, Eadweard (1984). *The Male and Female Figure in Motion : 60 Classic Photographic Sequences*. New York, Dover Publications.

- NATURAL POINT, 2011. Os sistemas Optitrack systems são apresentados em <http://www.naturalpoint.com/optitrack/products/motion-capture/>, acessado em 25/05/2011.
- NINTENDO, 2011, imagens do jogo e do seriado *Donkey Kong Country* são encontrados no endereço <http://donkeykong.nintendo.com/>, acessado em 02/07/2011.
- NOKIA/Qt software website, <http://qt.nokia.com/>, accessed in 20/05/2011.
- NOKIA/Qt license, <http://qt.nokia.com/products/licensing>, accessed in 20/05/2011.
- OPENGL 2011, toda a documentação da biblioteca gráfica OpenGL está no portal <http://www.opengl.org/>, acessado em 05/07/2011.
- PEDRINI, H. e SCHWARTZ, W.R. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*, Thomson Learning, 2008.
- PHASESPACE INC., 2011. *PhaseSpace Inc — Optical Motion Capture*. Este é um documento eletrônico, no endereço <http://www.phasespace.com>, acessado em 25/05/2011.
- Pintea, P (2004). *Special Effects, an Oral History*, Harry N. Abrams Publishers, USA.
- PINTO, F., BUAES, A., FRANCO, D., BINOTTO, A., AND SANTOS, P. (2008). Bratrack: a low-cost marker-based optical stereo tracking system. In *SIGGRAPH '08: ACM SIGGRAPH 2008 posters*, ACM, New York, NY, USA, 1–1.
- POYNTON, C. (2007) *Digital Video and HDTV, Algorithms and Interfaces*, Morgan Kaufmann, USA.
- QUEIROZ, Daniel Pacheco de (2011). *Avaliação de Métodos de Rastreamento de Marcadores para um Sistema Óptico de Captura de Movimento*. Dissertação de Mestrado, Departamento de Ciência da Computação, UFMG.
- QUEIROZ, D. P., FLAM, D.L., GOMIDE, J.V.B. and ARAÚJO, A. de A. (2011). *Evaluation of real time tracking methods of motion capture markers*, submetido para a SBGAMES 2011.
- RASKAR, R., NII, H., DEDECKER, B., HASHIMOTO, Y., SUMMET, J., MOORE, D., ZHAO, Y., WESTHUES, J., DIETZ, P., BARNWELL, J., NAYAR, S., INAMI, M., BEKAERT, P., NOLAND, M., BRANZOI, V., AND BRUNS, E. (2007). Prakash: lighting aware motion capture using photosensing markers and multiplexed illuminators. *ACM Transactions on Graphics* 26, 3, 36. <http://doi.acm.org/10.1145/1276377.1276422>.
- REDE GLOBO (2002), imagens da telenovela “Desejos de Mulher”, Rede Globo (2002 - 2006).
- REDE GLOBO (2003), imagens da telenovela “Chocolate com Pimenta”, Rede Globo (2003-2004, 2007).
- REDE GLOBO (2004), imagens da telenovela “Cabocla”, Rede Globo (2004, 2008).

- REINHARDT, Robert (2007). *Adobe Flash CS3 Professional Video Studio Techniques*, Adobe Press, USA.
- RICKITT, Richard (2007). *Special Effects, the History and Technique*. Billboard Books, USA.
- RPM website, <http://www.rpm.com.br/web/Credenciais.asp>, accessed in 28/05/2011.
- SIO2 2011, portal do *engine* que integra os software Blender e Maya ao sistema operacional iOS, no endereço <http://www.sio2interactive.com/>, acessado em 06/07/2011.
- SOLOMON, Charles (1994) *The History of Animation*, Wing Books, NY, USA.
- STAGE, 2011. *Organic motion: Solutions*. This is an electronic document available at <http://www.organicmotion.com/solutions/stage>, accessed in 25/05/2011.
- STORN, R. and price, K. (1997). Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. *Journal of Global Optimization*, 11(4):341–359. <http://dx.doi.org/10.1023/A:1008202821328>.
- STURMAN, David J (1994). *A Brief History of Motion Capture for Computer Character Animation*. In “Character Motion Systems”, ACM SIGGRAPH 94 Proceedings, Florida, USA.
- THOMAS, Frank and OLLIE, Johnston (2004). *Disney animation: the illusion of life*. Popular ed. New York: Abbeville Press, USA.
- UCHINOUMI, M., TAN, J. K., AND ISHIKAWA, S. (2004). *A simple-structured real-time motion capture system employing silhouette images*. In *Proceedings of the IEEE International Conference on Systems, Man & Cybernetics*, 3094–3098. <http://dx.doi.org/10.1109/ICSMC.2004.1400814>.
- VICON MOTION SYSTEMS, 2011. Motion Capture Systems from Vicon. Este é um documento eletrônico encontrado no endereço at <http://www.vicon.com>, acessado em 25/05/2011.
- W3C 2011, breve histórico da World Wide Web no documento <http://www.w3.org/History.html>, acessado em 03/06/2011.
- WILLOW GARAGE, 2011, <http://opencv.willowgarage.com/wiki/>, accessed in 28/05/2011.
- WIRED 2011, artigo de Rachel Ehrenberg, *Square Pixel Inventor Tries to Smooth Things Out* em <http://www.wired.com/wiredscience/2010/06/smoothing-square-pixels/>