

TÍTULO: Efeitos visuais, uma abordagem a partir do processamento digital de imagens

AUTORES: João Victor Boechat Gomide e Arnaldo de Albuquerque Araújo

NÍVEL: elementar

SUMÁRIO

1. Introdução
2. Evolução das trucagens com imagem
3. Imagem digital em movimento
 1. Conceitos gerais
 2. Descrições e formatos de imagens digitais
 3. Codificação e decodificação
4. Efeitos visuais digitais
 1. Introdução
 2. Técnicas e arte dos efeitos visuais digitais
 3. Aplicações e exemplos
5. Perspectivas para a pesquisa e para a produção audiovisual

1. Introdução

O objetivo desse trabalho é introduzir os conceitos mais importantes de imagem digital e demonstrar como eles 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 e o processamento digital de imagens. 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.

Os efeitos visuais fazem parte da pós-produção do audiovisual. É uma área que está no limiar entre a técnica e a arte. Com a utilização predominante da

imagem digital em movimento, ela depende fundamentalmente da pesquisa e desenvolvimento em computação gráfica e tratamento digital de imagens. 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, na maioria dos casos, o pessoal envolvido diretamente na 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. A pesquisa e desenvolvimento de ferramentas utilizadas em efeitos visuais têm grande interesse acadêmico e comercial, sendo realizada de forma interdisciplinar, com estreita colaboração entre pesquisadores da Ciência da Computação, engenheiros e técnicos e artistas da indústria do audiovisual.

O texto está dividido em quatro partes. Na primeira etapa será abordada a evolução das diversas técnicas com trucagem óptica, fornecendo exemplos de cenas clássicas, para que seja introduzido o tema da realização dos efeitos visuais. Em seguida, será apresentada a imagem digital, tanto na sua transformação, quando capturada em película, como nos formatos originalmente digitais, os conceitos de compressão, a estrutura da imagem digital e os resultados possíveis com os diversos tipos de formatos. Assim, se introduzirá o suporte para o tratamento digital dos efeitos visuais. Na terceira parte serão discutidas algumas técnicas de efeitos visuais digitais, como a correção de cor, o rastreamento de imagem, a composição digital e a captura de movimentos, assim como os diferentes hardwares e softwares disponíveis e os conceitos utilizados neles. Estas técnicas serão abordadas juntamente com exemplos tirados do cinema, da publicidade e da teledramaturgia. Na última parte, são apresentadas as perspectivas para a pesquisa e desenvolvimento, assim como para a produção audiovisual.

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 noventa, quando começaram a ser substituídas pelos efeitos realizados em computadores. A maneira em que se produzem efeitos atualmente distanciou-se completamente do que era feito normalmente há vinte anos atrás. Com o advento da imagem e do som digitais, novas técnicas foram inventadas e todas as outras adaptadas para esta nova forma de manipulação do audiovisual, tornando mais amplas as soluções para realizar imagetivamente 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.

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. 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 que são usados massivamente hoje em diversos formatos de imagem em movimento¹.

Nos primeiros anos, os efeitos visuais eram 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. Nele, todas as criaturas da ilha de foram animados com a técnica de *stop motion* e provocou um enorme impacto nas platéias na época. Além da substituição por parada de ação, utilizavam-se miniaturas, tanto em cenários quanto com bonecos. Com Méliès, se começou a empregar recursos óticos

¹ RICKITT, Richard. *Special Effects, the History and Technique*. Billboard Books, 2007, p.14-15.

para obter os efeitos, como superfícies semi-espelhadas, divisor de feixes da projeção para impressão na película, dentre outros.

Um avanço notável foi a invenção da impressora ótica (*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 projeções². Na figura 1 temos uma impressora ótica fabricada pela Oxberry.



Figura 1: impressora ótica Oxberry2.

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, temos:

² Pinteau, P. *Special Effects, an Oral History*, Harry N. Abrams Publishers, 2004, p. 34.

a) o *fade in* e *fade out*, quando a imagem vai ou vem para uma cor, em geral o preto;

b) fusão, quando uma imagem vai desaparecendo enquanto a próxima aparece gradualmente;

c) *wipe*, que são transições feitas com figuras geométricas, como círculos e quadrados, trazendo com ela a próxima cena.

d) virada de página, quando as cenas são mudadas como se estivessem na folha de um livro.

2. quadro congelado, que é feito com o quadro preso na impressora e 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 movimento³.

Assim, durante décadas foram realizados filmes com efeitos visuais óticos, muitos se tornando clássicos e referência na cinematografia. *Guerra nas Estrelas*, *Blade Runner*, *2001, uma Odisséia no Espaço*, *King Kong*, *Contatos Imediatos do Terceiro Grau*, *O Mágico de Oz*, *Roger Rabbit* e *ET, o Extraterrestre* 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 digitais⁴. Essa situação deve reverter rapidamente, já que são mais de cem anos de produção clássica contra o período de evolução da produção digital.

3. Imagens digitais em movimento

3.1. Conceitos gerais:

³ Burum, S. H., *American Cinematographer Manual*, The ASC Press, 2007, p. 270-275.

⁴ <http://www.visualeffectssociety.com>.

A substituição da película e do vídeo gerado eletronicamente pela imagem digital está em plena evolução. A melhoria da qualidade da imagem e dos dispositivos de aquisição tem acelerado esse processo. Atualmente, as imagens em movimento em película são, na maioria das vezes, digitalizadas e passadas novamente para película, depois de finalizado o filme.

A realização do efeito visual digital envolve diversas etapas do processamento digital de imagens. Primeiramente, a aquisição de imagens em movimento pode ser feita por câmeras digitais de vídeo ou de cinema, ou em película e depois digitalizada, ou, então, elas podem ser geradas inteiramente por computador, através da computação gráfica. Em seguida, elas podem passar por uma etapa de pré-processamento para melhorar a sua qualidade e, em seguida, de representação e descrição. Dependendo do tipo de efeito aplicado, as imagens podem passar por processos de segmentação, de transformações geométricas, de realce e de compressão, dentre outros⁵.

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

Para ser representado digitalmente, o sinal da imagem gerado nos dispositivos é amostrado e quantizado. A amostragem é o processo de discretizar o domínio de definição da imagem 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*). 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 ponto da imagem. Na figura II encontra-se

⁵ Gonzalez, R.C. e Woods, R.E. *Digital Image Processing*, Pearson Prentice Hall, 2007.

uma representação da digitalização de um sinal, onde no eixo horizontal temos os intervalos de amostragem e no eixo vertical os níveis de quantização.

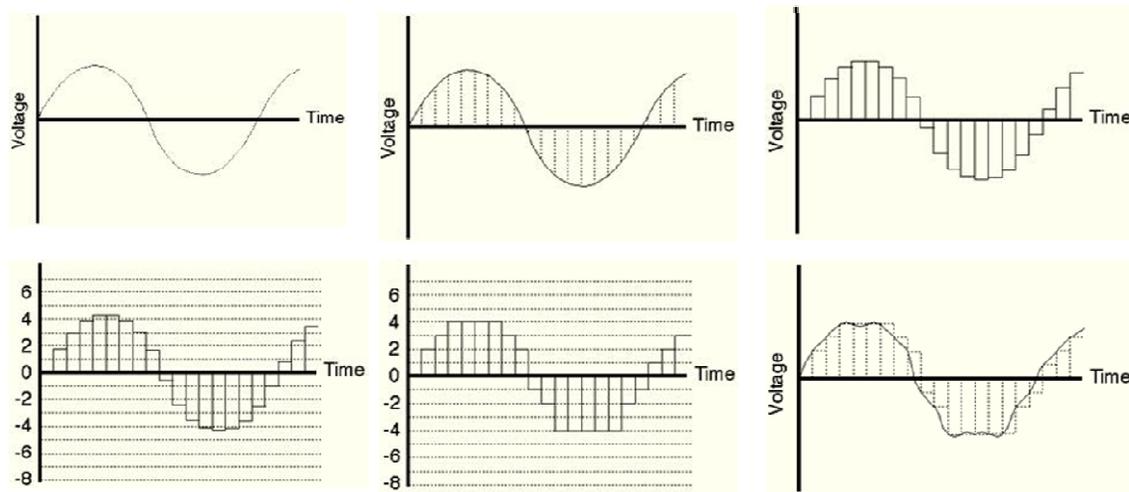


Figura II: digitalização de um sinal contínuo.

A qualidade da imagem digitalizada depende da resolução espacial e da profundidade de cor. A resolução espacial está relacionada com a densidade de pixels da imagem. A profundidade de cor está associada ao número de níveis de quantização da imagem. A cor de uma imagem é representada pelo espaço de cores. Para ser descrita completamente demanda três componentes, que podem ser as cores primárias, isto é, vermelho, verde e azul (RGB), ou as cores secundárias, que são o ciano, a magenta e o amarelo (CMY), ou matiz, luminância e saturação (HLS), dentre outras. Os equipamentos de captura de imagem podem ter um ou três dispositivos, do tipo CCD ou CMOS. Na figura III observa-se a separação das cores com um prisma e a sua captura por 3 CCDs. Para um sensor apenas, utiliza-se um filtro na frente do dispositivo para separar as cores, como a malha de Bayer da figura IV. O número de níveis de quantização é uma potência de 2, dependendo do número de bit, sendo igual a 8 para a televisão digital de definição standard (SDTV) e de 8 a 10 para a televisão digital de alta definição (HDTV)⁶.

⁶ Poynton, C. *Digital Video and HDTV, Algorithms and Interfaces*, Morgan Kaufmann, 2007, p. 16.



Figura III: separação das cores na câmera digital de 3 CCDs⁷.

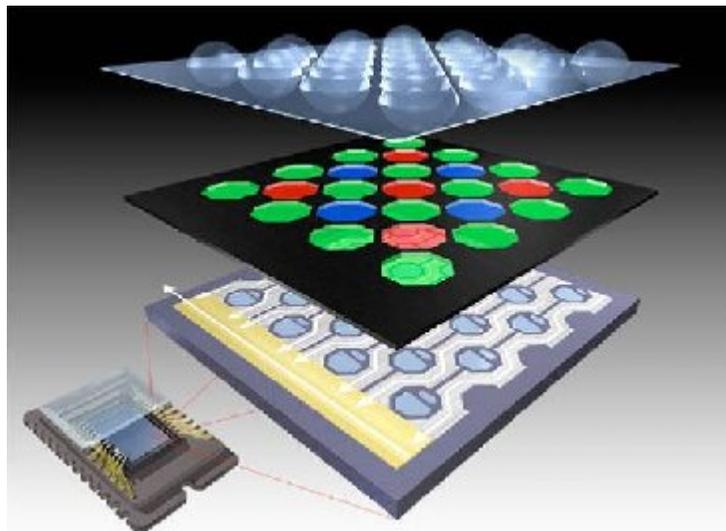


Figura IV: separação das cores na câmera digital com a malha de Bayer.

Para se adequar ao espaço de armazenamento, à quantidade de processamento e ao tempo de transmissão de imagens, surgiram diversas técnicas de compressão de imagens. Com isso se reduz 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

⁷ Ebert, C., Texto do workshop *Cinematografia Digital*, 2008, p. 11.

descompressão são chamados de *codecs*. Eles podem ser classificados em compressão com perda e sem perda.

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 perceptualmente 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 técnicas⁸.

A imagem digital é em geral descrita através dos valores de luminância, ciano e amarelo em cada pixel, o chamado formato Y:U:V. De acordo com a taxa de amostragem em cada pixel, temos comumente 4:4:4, 4:2:2, 4:2:0 e 4:2:1. Além desses conceitos acima, existe a razão de aspecto da imagem, que é a razão entre a largura e a altura da mesma.

A imagem em movimento possui uma cadência de exibição de quadros (*frames*) por segundo. 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, respectivamente. No caso do vídeo, cada quadro pode ser dividido em dois campos (*field*). Cada campo faz a varredura das linhas horizontais pares ou ímpares e eles são exibidos no formato entrelaçado (*interlaced*), para o vídeo analógico. Pode também ser exibido 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.

2.2. Descrição e formatos de imagens digitais

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, de 2003, e outros estão por ser estabelecido

⁸ Pedrini, H. e Schwartz, W.R. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*, Thomson Learning, 2008.

definitivamente, como o intervalo 2K a 4K para o cinema. Eles são definidos a partir de acordos da indústria para a produção de equipamento de aquisição, processamento e exibição de imagens. 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.

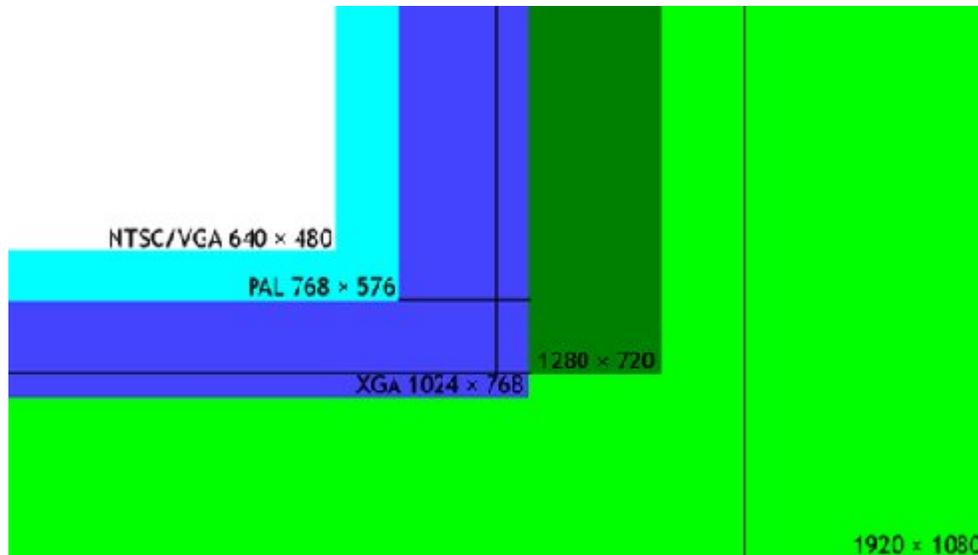


Figura V: formatos de SDTV e HDTV.

A televisão digital de definição padrão (SDTV) possui a razão de aspecto 4:3, profundidade de cor de 8 bit 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. O cinema pode ter a resolução de 2048x1080 pixels, no formato digital 2K, de 4096x2160 pixels, no formato 4K, ou valores intermediários. Na figura V encontram-se as diversas resoluções de SDTV e HDTV, com as duas razões de aspecto.

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

- 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 subamostragem de 4:4:4 para a HDCAM SR e *bitrate* de 880 Mbit/s.
- DVCPPro-HD/ DV100/ D7-HD da Panasonic. O DVCPPro-HD é formatado a 720p ou 1080i para uma variedade de *frame rates*. Para 1080i somente 1280 pixels horizontais são gravados, gerando perda de resolução.

As câmeras mais recentes estão armazenando as imagens em cartões de memória. As fitas devem ser abandonadas ao longo dos próximos anos.

3.3. Codificação e decodificação

Para manipular digitalmente uma imagem, deve-se conhecer como ela está representada e codificada. Fazer um efeito visual digital envolve o processo de rapidamente decodificar a imagem, manipulá-la e codificá-la novamente. O resultado pode ser familiar para o espectador, mas os conceitos técnicos envolvidos na realização do efeito são fundamentais para aquele que o realiza. 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 trabalhar com diferentes níveis de qualidade de imagem, dependentes de seus recursos intrínsecos. A maneira como a imagem em movimento é descrita digitalmente também determina o que se pode fazer com ela. Quem manipula as imagens deve conhecer esses aspectos para realizar artisticamente qualquer tipo de efeito visual, 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 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 isso⁹.

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 diminuindo a quantidade de informação de cores, a chamada subamostragem (*subsampling*) de cores. A compressão baseada em hardware e software permite uma maior quantidade de instruções serem feitas 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 mais importantes de compressão são apoiados em hardware e software, o JPEG (*Joint Photographic Experts Group*) e o MPEG (*Motion Pictures Experts Group*). A proposta do MPEG é utilizada para imagens em movimento e foi

⁹ Pedrini, H. e Schwartz, W.R. *Análise de Imagens Digitais: Princípios, Algoritmos e Aplicações*, Thomson Learning, 2008, p. 207-211.

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 I, 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 colocadas para cada proposta MPEG. O MPEG I camada 3, por exemplo, para o áudio, gerou o MP3. Alguns codecs mais conhecidos, com algumas aplicações, são mostrados na tabela I. Na figura VI tem-se 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 2000

Tabela I: exemplos de codecs, com banda de transmissão e aplicações.

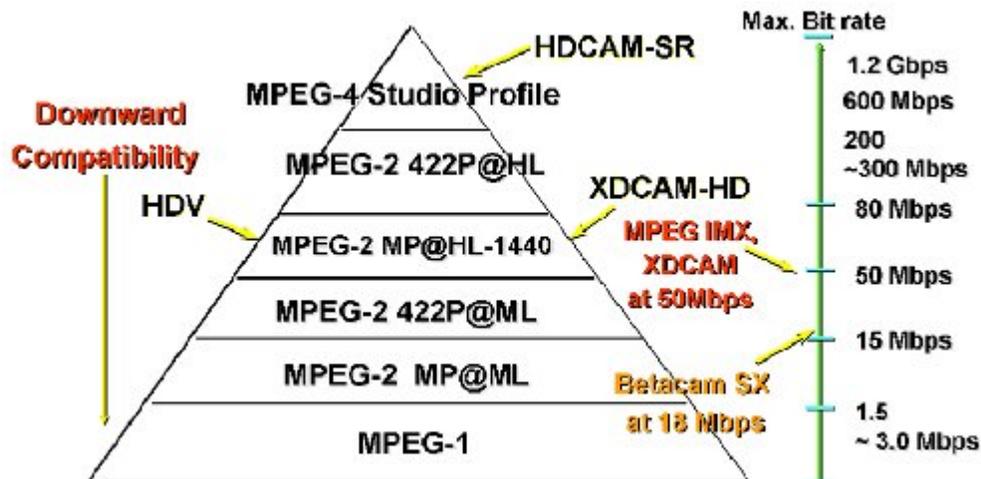


Figura VI: variações do codec MPEG e bitrates.

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 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 próximos¹⁰.

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 figura VII tem-se um esquema dos diferentes frames na compressão MPEG.

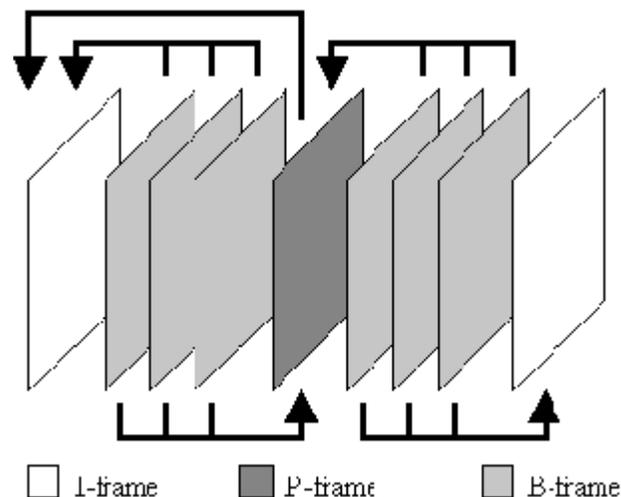


Figura VII: representação esquemática dos frames *I*, *P* e *B* da compressão MPEG.

¹⁰ Poynton, C. *Digital Video and HDTV, Algorithms and Interfaces*, Morgan Kaufmann, 2007.

4. Efeitos visuais digitais

4.1. Introdução

Os programas e equipamentos que estão disponíveis para quem trabalha na pós-produção, em especial de efeitos visuais, apresentam um custo relativamente acessível e qualidade profissional. Uma estação de trabalho bem estruturada, na faixa de R\$ 15000 hoje, realiza o que somente equipamentos de um milhão de dólares faziam 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 variadas mídias, do cinema digital aos aparelhos celulares, a partir de apenas um tipo de instrumento de captura de imagem e de 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 2K, como é feito em diversos filmes. Na tabela II encontra-se uma comparação das taxas de transferência e da resolução para diversas mídias.

Tipo de cenário	Resolução/ Frame rate	Taxas de transferência
Conteúdo de celulares	176X144, 10-15 frames/s	50-60 Kbps
Internet e SD	640X480, 24 fps	1-2 Mbps
HD	1280X720, 24p	5-6 Mbps
HD completa	1920X1080, 24p	7-8 Mbps

Tabela II: cenários de imagens em movimento com compressão MPEG 4 para diversas mídias.

O usuário deve ter em mente o tipo de aplicação e de imagem a ser trabalhada. Precisa saber o que significa trabalhar com vídeo analógico, nos formatos composto, s-vídeo ou componente, ou com o vídeo digital em suas diversas representações. Os processadores e as placas de vídeo podem permitir 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 compatíveis.

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, e possuem muitos outros recursos. Entre tantos programas de edição disponíveis no mercado, podemos 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;
- Edius Pro, da Thomson Grass Valley.

Para composição gráfica, colorização e efeitos visuais, temos:

- 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.

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 dentro 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 se ter cópias dos primeiros lançamentos, que cumprem a sua função durante muito tempo.

4.2. Técnicas 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;
 - Matte;

- Máscaras digitais;
- Correção de cor;
- Rastreamento;
- Personagens e/ou cenários virtuais;
- Captura de movimento;
- *Morphing*;
- *Warping*;
- Mapeamento de camera;
- Uso de técnicas combinadas.

A composição de imagens é o processo de integração de imagens de diferentes fontes em uma única imagem final. 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, combinação das cores definidas em cada pixel, escalamento, translação e/ou máscaras digitais.

Uma técnica amplamente utilizada é produzir máscaras nas fontes de imagem, para aplicá-las umas sobre as outras. A máscara digital é obtida através do canal alfa (*alpha channel*), que descreve tonalidades de cinza entre o branco e o preto, que são incluídos. O canal alfa é um quarto valor, que é acrescentado aos valores de cor de cada pixel. Para uma imagem descrita com 8 bits por componente de cor, o canal alfa terá também 8 bits, que resultará em uma imagem com 32 bits por pixel. Nesse caso, o valor 0 é dado ao preto e o 255 ao branco, com os demais 254 valores descrevendo tonalidades de cinza. O pixel com valor 0 será

completamente transparente e aquele com 255 completamente opaco. As tonalidades de cinza são valores intermediários de transparência. Os valores do canal alfa nos pixels podem ser fixos ou variarem a cada quadro. Os softwares de composição e edição possuem ferramentas, como a aplicação de curvas do tipo bezier, para gerar canais alfa. O mesmo ocorre com os softwares de modelagem 3D e de tratamento de imagem, que produzem imagens com o canal alfa já incorporado.

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 da tonalidade 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 permitem a obtenção de uma máscara com as bordas bem definidas. A forma como a imagem é descrita introduz artefatos nas bordas do elemento que é recortado do fundo com a cor do *chroma-key*. No mercado se encontram inúmeros casos de iluminadores que foram demitidos injustamente sob o argumento de que não sabem iluminar adequadamente a cena gravada com câmeras com subamostragem 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 finalizadoras de cinema e nas produtoras de vídeo realmente 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 sequencias de imagem 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 da vizinhança dessa região, 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 os 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 almeja. 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 podem se integrar a imagens reais ou as imagens podem ser 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.

Os computadores começaram a gerar imagens 3D sintéticas para substituir a câmera no início dos anos 70. Imageamento eletrônico foi primeiramente obtido nos anos 50, quando Ivan Sutherland e sua equipe, no Massachussets Institute of

Technology (MIT), produziram um sistema para controlar ferramentas industriais, o *Sketchpad*. Sutherland utilizou o *Sketchpad* para construir os primeiros objetos 3D gerados por computador. Até então, apenas figuras bidimensionais tinham sido criadas¹¹. Atualmente, imagens geradas por computador podem ser criadas por equipamentos com um custo relativamente baixo, com softwares como o Maya, o Lightwave, o 3DStudio Max e o Blender3D, esse último gratuito.

A captura de movimento, também conhecida como *mocap*, é um conjunto de artifícios usado para mapear e reproduzir deslocamentos em objetos ou seres vivos. Os dispositivos para 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. Ela pode ser dividida em quatro métodos básicos de captura: acústico, magnético, óptico e protético. O sistema óptico, com marcadores passivos, mapeia a posição de refletores colocados no corpo do ator. Um sistema desse tipo está em construção no Núcleo de Processamento Digital de Imagens da UFMG. A entrada do sistema de captura óptico é uma sequência de imagens capturadas pela câmera. A saída são as coordenadas das posições dos marcadores para cada imagem. No computador, a imagem é segmentada para extrair as regiões correspondentes aos marcadores na cena original. A segmentação é baseada em ferramentas de processamento de imagem, principalmente em transformações morfológicas. Através da segmentação, localizam-se os centros dos marcadores e seguem-se as suas posições através da sequência de 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, no cinema ou em vídeo. Uma configuração com 41 marcadores no corpo de um ator está na figura VIII.

¹¹ RICKITT, Richard. *Special Effects, the History and Technique*, pag.154-166.

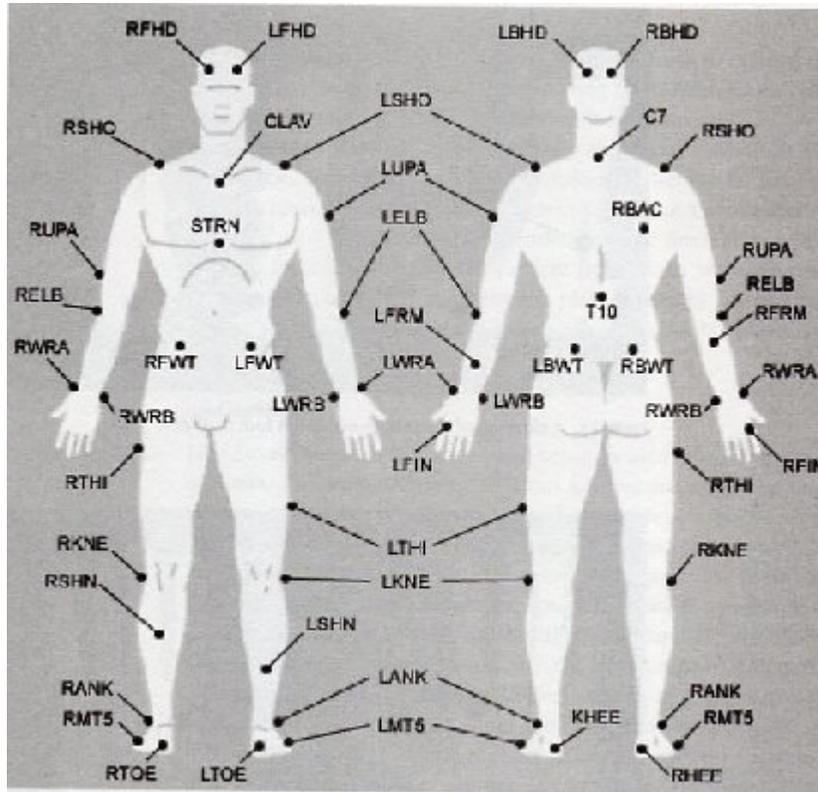


Figura VIII: posições para 41 marcadores¹².

Para seguir os marcadores através dos *frames*, depois da segmentação e da localização dos pontos, se emprega a predição de suas posições 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.

O processo de identificação das posições dos marcadores através do tempo é relativo a cada câmera, com os deslocamentos bidimensionais dos pontos. Através de cálculos de triangulação dos dados de cada câmera obtêm-se os valores tridimensionais dos marcadores no corpo do ator. Na figura IX são apresentados os resultados do mapeamento das posições dos marcadores, como se fossem uma nuvem de pontos.

¹² LIVERMAN, Matt. *The Animator's Motion Capture Guide*. Massachusetts: Charles River Media, 2004, p. 135.

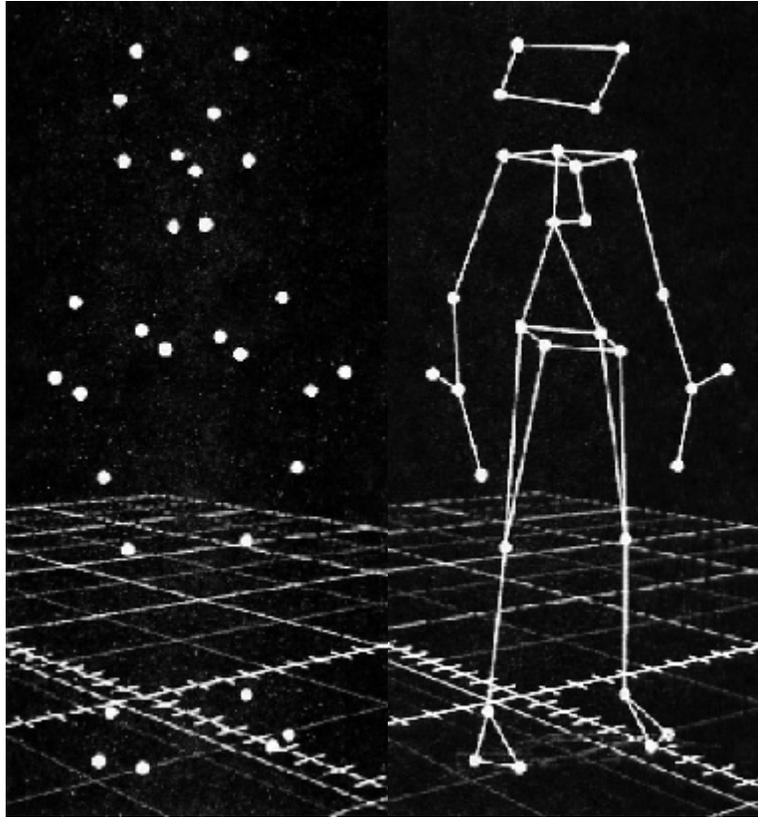


Figura IX: posições dos marcadores como se fossem uma nuvem de pontos, sem conexões, à esquerda, e com conexões¹³.

Essa técnica tem sido utilizada em filmes como *A Lenda de Beowulf*, *King Kong*, *O Senhor dos Anéis*, *Casa Monstro* e *O Expresso Polar*. A *Sony Entertainments* tem feito investimentos na pesquisa e desenvolvimento de sistemas de captura para o cinema e jogos digitais. Os sistemas desenvolvidos pela Sony são chamados de sistemas de captura de performance porque capturam detalhes dos movimentos faciais e corporais devido à alta resolução das câmeras.

O *morphing* e o *warping* pertencem a uma mesma classe de efeitos visuais produzidos por tratamento 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

¹³ LIVERMAN, Matt. *The Animator's Motion Capture Guide*, p. 26.

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 e também 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, se podem fazer, por exemplo, movimentos na boca de animais como se estivessem falando. Na figura X é apresentado um *morphing* de transformação de uma mulher em um homem.

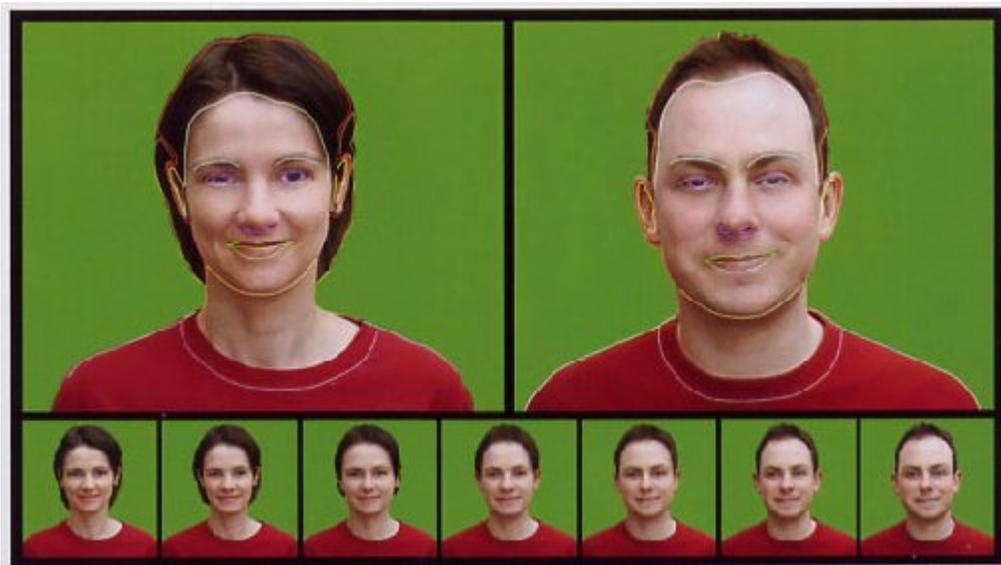


Figura X: transformação morfológica entre dois atores¹⁴

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

¹⁴ RICKITT, Richard. *Special Effects, the History and Technique*, pag.109.

movimentando. A figura XI mostra um esquema de um sistema de controle de movimento, como o utilizado em *Guerra nas Estrelas*.



Figura XI: sistema para movimentação de câmera, controlado por computador¹⁵

Todas essas técnicas descritas 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 tratamento 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, como em filmes de animação, com *Cars*, e de efeitos com modelagens mais aprimoradas de fenômenos físicos, pêlo, folhagem, fumaça, água, fogo e dublês digitais.

4.3. Aplicações e exemplos

A seguir são apresentados quatro exemplos de utilização de efeitos visuais, com a explicação da maneira como foram realizados.

¹⁵ RICKITT, Richard. *Special Effects, the History and Technique*, pag.147.

Como primeiro exemplo, temos 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, teve que ser rastreado o movimento dos pontos para que os elementos introduzidos e compostos na imagem pudessem acompanhar o movimento da câmera. Na figura XII é apresentada a imagem original e na XIII a cena transformada em noite.



Figura XII: imagem diurna de uma filmagem de 1918¹⁶.



Figura XIII: cena modificada, com luzes aplicadas em máscaras digitais, correção de cor, rastreamento de movimento e céu noturno¹⁷.

¹⁶ Imagem original obtida no arquivo da Cinemateca Brasileira.

¹⁷ *Stockshot* de “Cabocla”, Rede Globo (2004-2008).

No segundo exemplo, multiplicou-se a figuração presente em uma cena gravada no Teatro C3lon, em Buenos Aires, que contava com poucos figurantes na plat3eia 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 plat3eia. O enquadramento descia do teto em *tilt* e, ap3s enquadrar o palco por tr3s da plat3eia, fazia um *travelling* lateral. O problema principal foi o da lente grande angular, que introduzia distor33es na imagem, dificultando o rastreamento de pontos. Ao utilizar os softwares *Inferno* ou *After Effects*, a localiza33o dos pontos ficava inst3vel, fazendo com que os elementos da figura33o se movessem diferentemente da imagem. Isso dava um aspecto indesej3vel ao efeito, que tornava vis3vel a sua aplica33o. Ao utilizar o software *Boujou*, foram gerados milhares de eixos em toda a imagem, permitindo escolher aqueles mais adequados para aplicar os elementos. Os dados das posi333es dos eixos ao longo do tempo gerados pelo *Boujou* foram transferidos para o *Inferno*, onde foram aplicados os figurantes virtuais. Nas frisas, os figurantes foram pintados em um quadro, com a ferramenta *Paint* do *Inferno*, e neles foram aplicadas as posi333es dos eixos rastreados. Da plat3eia foram recortados figurantes reais, que foram aplicados nas cadeiras vazias. Essa cena n3o foi gravada com a inten333o de multiplicar a figura33o e os grandes vazios de pessoas na plat3eia s3o foram notados depois, durante a edi333o. Na figura XIV se tem os quadros da cena sem a multiplica333o de figura33o e, depois, com a plat3eia e frisas com os figurantes virtuais aplicados.



*Figura XIV: cenas do Teatro C3lon, antes e depois de modificadas, com figurantes adicionados 3s frisas e 3 plat3eia*¹⁸.

¹⁸ Imagens do primeiro cap3tulo de "Chocolate com Pimenta", Rede Globo (2004, 2007).

No terceiro exemplo, utilizou-se o *warping* para dilatar a pupila do ator, pontuando dramaticamente a cena. Nessa sequência, um dos atores principais levava um tiro. No momento que precede o tiro, em um plano-detalle do olho, foi feita a dilatação 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 maior diâmetro. O software calculou e distorceu as imagens, de forma crível, entre o primeiro e o último quadro do efeito. Na figura XV observa-se o quadro antes e depois do efeito aplicado.



Figura XV: warping aplicado à pupila, antes e depois do efeito aplicado.

O quarto exemplo é uma sequência de tentativa de atropelamento. Nesse caso, foram utilizadas máscaras digitais, rastreamento e composição digital, com atores e dublês atuando em um fundo para *chroma-key*. Em um dos planos, a câmera, movimentando-se sobre um trilho, 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 figura XVI se pode ver dois quadros da sequência, com o plano e o contra-plano.



Figura XVI: frames da cena com composição digital, rastreamento e chroma-key.

5. Perspectivas para a pesquisa e para a produção audiovisual

A pesquisa e o desenvolvimento de soluções de tratamento 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.

Podemos citar como temas atuais 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, 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 nesse período.

A produção e a criação do audiovisual e de jogos têm se beneficiado com a evolução e a ampliação do uso do tratamento digital de imagem e da computação gráfica. 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 talento e 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 ótica 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 ótica – 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 sempre¹⁹.”

¹⁹ RICKITT, Richard. *Special Effects, the History and Technique*, pag.79.