

## 2 Matting e Chroma Key

Matting e keying são conceitos relacionados. Porém, tanto a literatura como os profissionais da área de composição não esclarecem completamente esta relação. Este capítulo apresenta os conceitos de matting, chroma key e keying, acompanhados de uma classificação dos seus principais métodos.

### 2.1. O Problema de Matting

Matting é o processo de estimar o primeiro plano (foreground) de uma imagem para posterior composição sobre outra imagem de fundo (background). Este processo parte do princípio de que qualquer imagem  $I$  pode ser representada pela combinação convexa de uma imagem foreground  $F$  e uma imagem background  $B$ :

$$I_p = \alpha_p F_p + (1 - \alpha_p) B_p, \quad \alpha_p \in [0, 1] \quad \text{Eq. 2.1}$$

onde  $p$  se refere à localização do pixel e  $\alpha_p$  é o coeficiente de interpolação denominado de *matte* (Figura 2.1). A Eq. 2.1 é denominada de “equação de composição”.

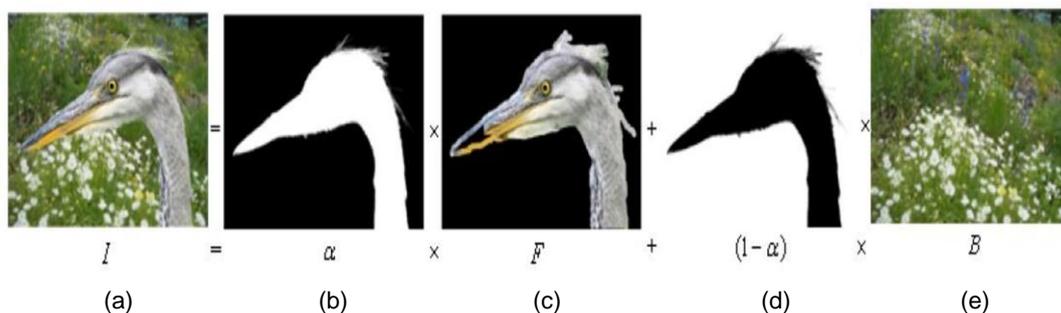


Figura 2.1 Imagem  $I$  como uma combinação convexa da imagem foreground  $F$  e a imagem background  $B$  usando o coeficiente de interpolação  $\alpha$ . O coeficiente  $\alpha$  é chamado de “matte”. [imagem extraída de Wang (2007)].

O coeficiente  $\alpha_p$  pode ser interpretado sob vários pontos de vista: como um problema de interpolação (Rother *et al.* 2004), como probabilidade do pixel

pertencer ao objeto foreground (Chuang *et al.* 2001), como uma solução de equação diferencial parcial (Sun *et al.* 2004), ou até mesmo como resultado de mais de um ponto de vista (Grady *et al.* 2005). O coeficiente  $\alpha_p$  também pode ser interpretado como opacidade de foreground do pixel, isto é:

- $\alpha_p = 1$  determina que o pixel pertence ao foreground (dizemos, neste caso que o pixel é definitivamente foreground);
- $\alpha_p = 0$  determina a pertinência ao background (pixel definitivamente background); e
- $\alpha_p$  com valores intermediários ( $0 < \alpha_p < 1$ ) representa um pixel misto que está na fronteira do foreground (*i.e.* o pixel na fronteira de um objeto foreground contém usualmente alguma percentagem do background).

Esta última interpretação reduz o problema de matting ao de determinar a cobertura parcial ou completa do pixel – um problema apresentado matematicamente pela primeira vez por Porter & Duff (1984). Estes dois autores introduziram o conceito de canal alfa (*alpha channel*) como um meio de controlar a interpolação linear de cores foreground e background para diminuir o *aliasing* durante a renderização de um foreground sobre um background qualquer. Por esta razão, também denominamos o *matte* de *alpha matte*.

A equação de composição (Eq. 2.1) mostra que matting é um problema subespecificado, porque temos que, para cada pixel  $p$ , resolver 7 variáveis

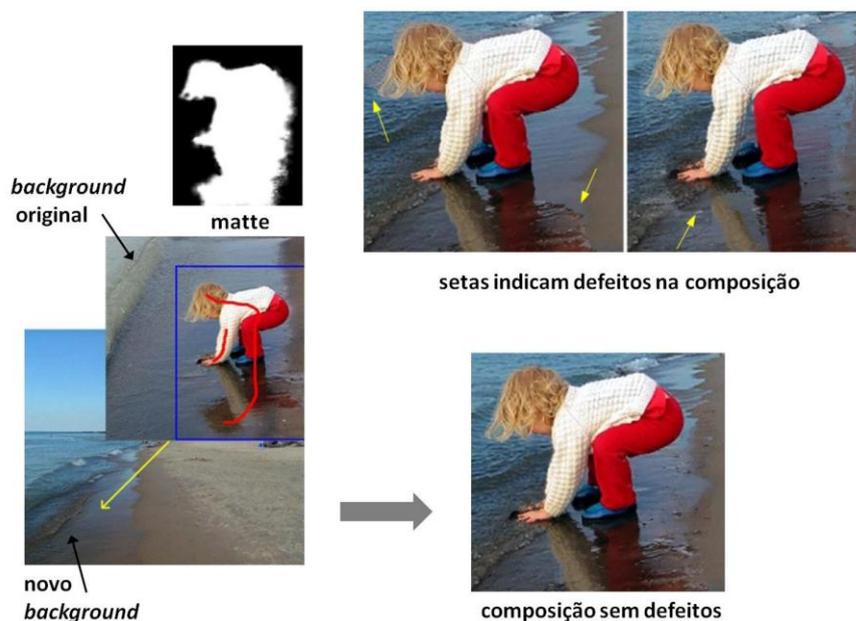


Figura 2.2 Exemplo de defeitos de composição após a extração do foreground. A composição sem defeitos aparentes é obtida com o algoritmo de Wang & Cohen (2007a). [Figura composta a partir de Wang & Cohen 2007a]

desconhecidas (3 componentes do vetor de cor  $F_p$ , 3 do  $B_p$  e um valor escalar  $\alpha_p$ ) a partir de 3 valores conhecidos (3 componentes do vetor de cor  $I_p$ ).

Em outras palavras, matting é um problema mal-posto, no sentido de que não tem unicidade de solução. Portanto, as soluções são sempre estimativas que dependem da orientação do usuário e de hipóteses sobre estatísticas da imagem. Objetos que possuem cabelos e pelos espalhados no seu contorno ou que apresentam transparências irregulares causam grandes dificuldades aos algoritmos propostos. Porém, mesmo após obtermos bons valores de  $F$ ,  $B$  e  $\alpha$ , precisamos tratar do problema de compor o foreground sobre um novo background. Esta não é uma etapa trivial, como explicamos a seguir.

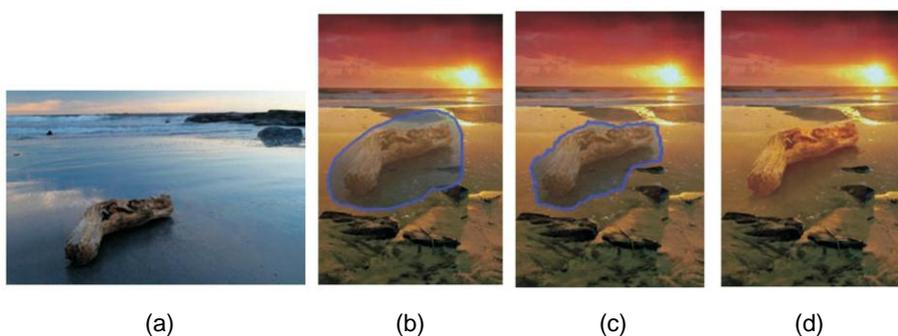


Figura 2.3 O sistema Drag-and-Drop Pasting (Jia et al. 2006), onde a região selecionada pelo usuário é colocada na imagem alvo (b), a região é otimizada (c) e a composição final é obtida (d) [Figura composta a partir de Jia et al. 2006].

A maioria dos métodos de matting tratam matting e composição (*compositing*) como tarefas separadas, ao assumir que o novo background é desconhecido. Geralmente, neste caso, a composição gera regiões com distorções, se distanciando do ideal de uma composição sem costuras aparentes (*i.e. seamless*) (Figura 2.2). Às vezes, estas distorções surgem quando há um zoom na imagem composta. Como o principal objetivo de matting é a composição posterior, uma possível linha de ação é tratar a questão diretamente como composição. Um bom exemplo desta linha é o sistema “Drag-and-Drop Pasting” (Jia et al. 2006), baseado em edição de imagens através de equações de Poisson, onde o usuário desenha uma fronteira que circunda o objeto de interesse, arrasta a região e a coloca sobre a imagem alvo (Figura 2.3). Este sistema altera a cor de todo o foreground para misturá-lo ao novo background, o que nem sempre é um efeito desejado. Uma outra linha de ação propõe integrar matting e composição em um único processo de otimização, gerando um *matte* que considera o novo background (Wang & Cohen 2007a). Estes autores propõem que este tipo de algoritmo seja classificado como “compositional

matting”. O sistema de Wang & Cohen (2007a) não causa a alteração de cor de todo o foreground como em (Jia *et al.* 2006), mas tem outras limitações, tais como: o novo background deve ser grande e ter regiões similares ao background original e o sistema precisa ser chamado cada vez que uma nova imagem de background é selecionada. Este sistema é útil para mover e aumentar a escala de um foreground sobre um mesmo background panorâmico.

O foreground extraído pelo processo de matting (Figura 2.1(c)) pode apresentar pixels da fronteira com tons do background. Este defeito pode ficar mais acentuado quando o fundo é uma tela azul (ou verde). Neste caso, dizemos que a cor do fundo “vazou” ou “derramou” (*spill*) sobre o foreground (fenômeno também chamado de “color spill”). Na realidade, mesmo que este defeito seja minimizado, o foreground extraído pode estar, como um todo, influenciado pelo ambiente. As várias propostas de matting na literatura não tratam explicitamente deste assunto. Muitas vezes, para evitar esta “contaminação” do fundo, o diretor de fotografia tenta iluminar mais fortemente o objeto de foreground durante a captura. Entretanto, na prática, a intervenção do editor termina sempre sendo necessária ser feita quadro-a-quadro, manualmente. No caso de backgrounds de uma única cor, Dupont & Deschênes (2006) propõem uma outra equação de composição, diferente da Eq. 2.1, que explicitamente considera a cor do objeto sob a iluminação da fonte de luz e a iluminação refletida pelo background.

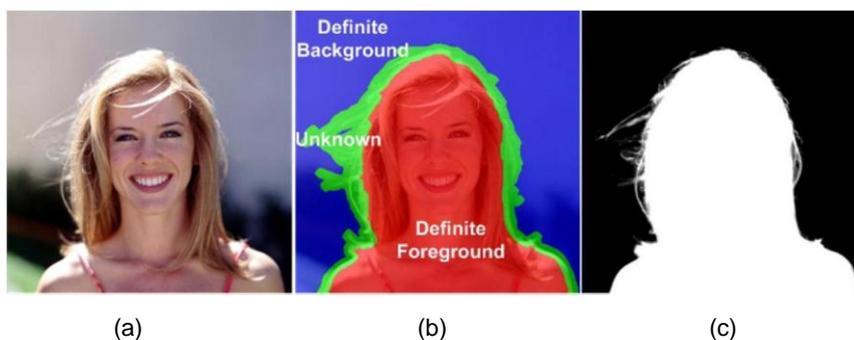


Figura 2.4 Exemplo de trimap (b) e matte obtido a partir deste trimap (c) [Extraído de Wang & Cohen (2008)].

Matting é um problema complexo de segmentação de imagem que não pode ser resolvido de maneira automática. Para evitar esta segmentação direta, a maioria dos algoritmos de matting simplificam o problema partindo de uma segmentação grosseira da imagem em 3 regiões (Figura 2.4):

- definitivamente foreground,
- definitivamente background e

- região desconhecida.

Este mapa com três regiões é chamado de trimap. O trimap pode ser manualmente especificado pelo usuário ou produzido por um algoritmo de segmentação.

## 2.2. Categorias de Algoritmos de Matting

Como matting é um problema mal-posto, temos duas maneiras básicas de aliviar a situação:

1. diminuir o espaço de solução usando informações extras para matting;
2. aliviar o processo considerando as correlações fortes que existem entre pixels próximos, isto é, usando estatísticas locais de imagens.

Na primeira maneira, temos soluções que dependem do ambiente ou equipamentos especialmente preparados para a captura. Por exemplo:

- **Chroma Keying** (também chamado de Matting de Tela Azul) que requer um fundo azul ou verde;
- **Difference Keying** que se baseia em obter duas imagens e explorar algum tipo de diferença entre elas. As técnicas mais conhecidas são as seguintes: o **matting de flash** (ou *flash keying*) que se baseia no fato de que a diferença mais notada entre as imagem com e sem flash está no objeto de foreground (Sun *et al.* 2006); e o **background subtraction** que usa a diferença de cor entre uma imagem com o objeto de interesse presente e uma semelhante sem o objeto de interesse.
- **Depth Keying** que se baseia em processos de estimativa de profundidade. Este método tende a ser monopolizado por câmeras de vídeo que estimam distâncias em tempo real, como a ZCam usada no Projeto Natal<sup>2</sup>;

<sup>2</sup> A empresa israelense 3DV Systems, comprada pela Microsoft (Shelah 2009), desenvolveu inicialmente a câmera ZCam para fins de composição. A descrição de *depth keying* usando a ZCam pode ser encontrada em Gvili *et al.* (2003). Câmeras do tipo ZCam também viabilizam a técnica de TV 3D baseada na transmissão conjunta de vídeo colorido mono e informação associada de profundidade por pixel, ao invés de *stream* de vídeo estereoscópico (Fehn 2004).

- **Thermo Keying** que se baseia no fato de que a temperatura do corpo humano se diferencia da temperatura do background. Esta técnica usa a medida de radiação da luz infravermelha (Yasuda *et al.* 2004);
- **Compositional matting** que requer backgrounds específicos e similares (Wang & Cohen 2007a);
- **Matting com Fileiras de Câmeras** para matting de vídeo em tempo real que explora a paralaxe entre os objetos. Entre algumas de suas limitações estão a exigência de que objetos não devem ter auto-occlusão e o fato de que o sistema é projetado para cenas de interiores (*indoors*) com backgrounds controlados.

Na segunda maneira básica, que usa correlações entre pixels próximos, adotamos a classificação proposta por Wang & Cohen (2008), exceto que propomos chamar a terceira categoria de Métodos Híbridos:

- **Métodos de Amostragem de Cor** que estimam diretamente o valor de  $\alpha$  em cada pixel, baseados no fato de que pixels vizinhos que têm cores similares geralmente têm valores de  $\alpha$  similares. Alguns dos principais exemplos são o método de Ruzon & Tomazi (2000) e o Matting Bayesiano (Chuang *et al.* 2001). O algoritmo de chroma key patenteado por Mishima (1993) foi o primeiro desta categoria e ainda é um dos mais eficiente do mercado. Os métodos baseados em amostragem não funcionam muito bem quando a hipótese de forte correlação não é observada, isto é, quando as imagens tem regiões de pouca suavidade e o *trimap* não está bem definido.
- **Métodos baseados em Afinidades** que definem afinidades entre pixels numa vizinhança pequena (e.g. um peso associado a dois pixels vizinhos) e usa gradientes ao longo da grade de imagem, ao invés de estimar  $\alpha$  diretamente em cada pixel. Alguns métodos estão listados a seguir, devido às suas possibilidades práticas: Matting de Poisson (Sun *et al.* 2004); *Random Walk Matting* (Grady *et al.* 2005); Matting Geodésico para imagem estática e vídeo que usa a distância geodésica de um pixel desconhecido para regiões definitivamente  $F$  ou  $B$  (Bai & Sapiro 2009); Matting usando *Fuzzy Connectedness*<sup>3</sup> que é similar ao Matting Geodésico (Zheng *et al.* 2008); *Closed-form Matting* (Levin *et al.*

<sup>3</sup> *Fuzzy Connectedness* captura adjacência e similaridade entre elementos de imagem (Udupa & Samarasekera 1996).

2006). Os métodos baseados em afinidades, embora sejam mais robustos para lidar com imagens complexas, podem apresentar imprecisão em relação aos métodos baseados em amostragem. Em particular há duas fontes de imprecisão: (1) as estimativas em separado dos valores de alpha e das cores de foreground; (2) propagação e acumulação de erros no processo de ir de pixels conhecidos para pixels desconhecidos.

- **Métodos Híbridos** que combinam amostragem de cor e afinidades em um único processo de otimização. Os exemplos desta categoria são os que estão mais alinhados com os interesses da presente dissertação. Por exemplo, o sistema de matting iterativo (Wang & Cohen 2005), *Easy Matting* (Guan *et al.* 2006) e *Robust Matting* (Wang & Cohen 2007b). Estes métodos tendem a resolver as deficiências dos métodos das categorias anteriores.

Em muitos dos métodos que usam afinidades, as funções de energia para serem minimizadas são definidas em uma forma quadrática, o que leva a sistemas lineares grandes e requer resolvidores de sistema lineares diretos (Duff *et al.* 1986) ou iterativos. Os métodos iterativos (tais como SOR – *Successive Over-Relaxation* e *Conjugate Gradient Descent*) são computacionalmente mais eficientes do que os métodos diretos. O pacote Matlab também pode ser usado para testar protótipos de sistemas de matting, mas deve ser substituído por resolvidores mais eficientes integrados ao sistema final.

### 2.3. Chroma Key

Chroma key é uma técnica de matting de imagem, onde o fundo tem uma cor única conhecida (geralmente azul). Por esta razão, chroma key também é referido como sendo a técnica de Blue Screen Matting.

O problema de color spill (também denominado de blue-spill) é particularmente crítico no chroma key, por causa da iluminação que geralmente fazemos no objeto durante a captura. Praticamente quase todos os softwares de edição e composição no mercado usam um algoritmo de chroma key chamado Primatte, criado por Mishima (1994) da IMAGICA Technology Corp e comercializado pela The Foundry. No algoritmo Primatte, como o background

tem somente uma única cor, todos os pixels de background podem ser cobertos por uma pequena esfera aproximada por um poliedro (uma malha de triângulos) no espaço de cor RGB (Figura 2.5). Todos os pixels de foreground formam um outro poliedro fora do poliedro de background. O valor  $\alpha$  de um pixel desconhecido (i.e. misto) é estimado pelo cálculo de sua posição relativa aos dois poliedros.

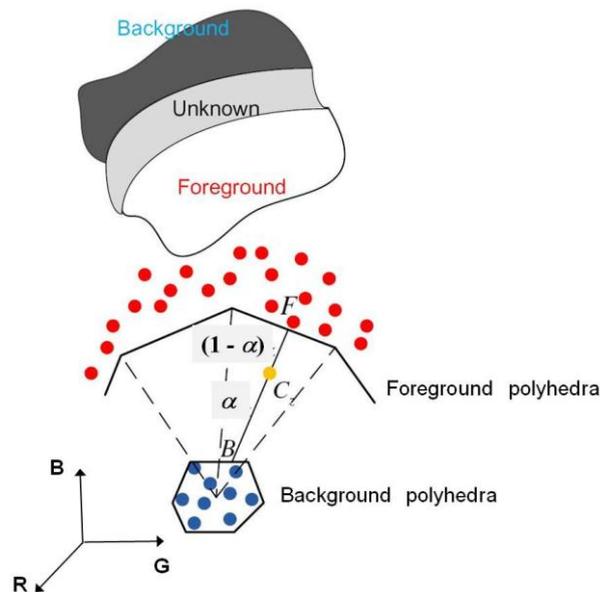


Figura 2.5 No algoritmo Primatte,  $\alpha$  é dado pela distância do pixel C para o poliedro B no espaço de cor RGB [figura adaptada de Wang & Cohen (2008)].

O chroma key também pode ser entendido como uma técnica de *keying*, que é uma forma de superpor duas camadas de vídeo, de maneira que uma camada gera uma espécie de “keyhole” (buraco de fechadura) através do qual a outra camada pode ser vista. Em um chroma key, o “keyhole” é feito por toda a parte da imagem que é de uma certa cor (usualmente azul). Outros parâmetros podem ser usados para construir “keyholes”. Por exemplo, a técnica de luma key usa luminância como chave (key), isto é, o “keyhole” é definido com base nos níveis de brilho ao invés de cor. As técnicas de keying têm uma finalidade muito clara e precisa de fazer composição. Keying e matting são, portanto, conceitos relacionados.

O Processo de chroma key foi executado pela primeira vez em 1930 por Larry Butler, em um processo demorado e complicado, baseado na anulação de

cores através de filtros coloridos seguido de uma serie de refilmagens. O primeiro filme a utilizar esse processo foi “The thief of Bagda”, em 1940.

É preciso salientar que, no período em questão, o cinema ainda era completamente analógico, isto é, usava filme em rolo e reagentes químicos para revelação da imagem e tratamento do negativo. O processo consistia em filmar duas tomadas, a primeira sendo o fundo no qual o ator deveria ser inserido e a segunda vindo da filmagem do ator contra o fundo azul (Figura 2.6).

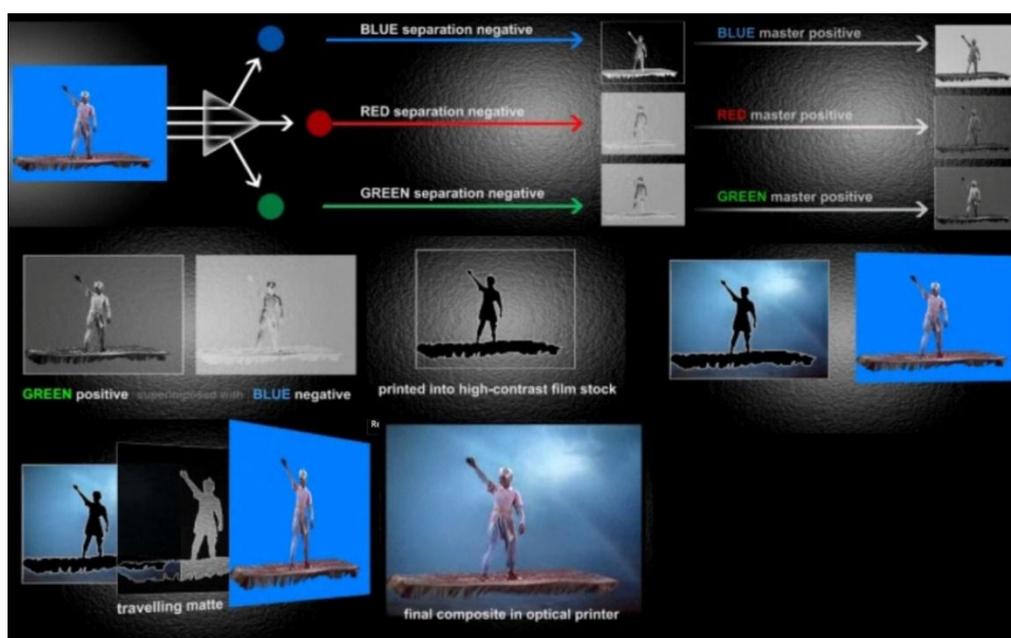


Figura 2.6 Processo de chroma key analógico

Na técnica de chroma key analógico, para que fosse possível colocar o ator no fundo previamente filmado, o ator precisava ser separado do seu fundo azul por truques fotográficos. A cena do ator no fundo azul era então refotografada, isto é: ela era projetada em outro rolo de filme, mas com o uso de um filtro azul de forma que apenas o fundo era exposto. Um filme especial era usado neste processo e o resultado era uma imagem de negativo preto e branco, onde o fundo era preto e a pessoa no meio transparente - isto era chamado de “matte fêmea”. Matte pode ser interpretado como uma máscara, isto é, tudo que nesta máscara for preto não deve aparecer e o que for branco deve ser projetado.

A cena do ator no fundo azul era novamente fotografada, mas dessa vez com um filtro vermelho e verde, fazendo com que apenas o ator fosse exposto e gravado no filme, gerando uma silhueta preta sobre uma área não exposta de filme - isto era chamado de “matte macho”. Finalmente a imagem do fundo era refotografada através do matte macho e este resultado era novamente fotografado através do matte fêmea. Era usada uma impressora ótica, com dois projetores, uma câmera de filmar e um “beam splitter”<sup>4</sup>, com isto as imagens eram combinadas para gerar a composição final.

Em 1950, a Warner Brothers contratou Arthur Widmer, um ex-pesquisador da Kodak, que passou a trabalhar com um processo que empregava luz ultravioleta para o processo de matting. No mesmo período, Arthur Widmer também começou a desenvolver técnicas de *bluescreen*.

O processo de chroma key tornou-se extremamente fácil de ser produzido a partir do processamento digital de imagem. Com este tipo de processamento, a velocidade de produção se tornou incomparavelmente maior do que as obtidas por processos analógicos. O *workflow* desse processo digital pode ser resumido como está apresentado na Figura 2.7.



Figura 2.7 Fluxograma do processo de chroma key digital

<sup>4</sup> Beam splitter é um dispositivo ótico que divide o raio de luz em dois raios.

Os principais problemas encontrados atualmente quando se produz cenas em chroma devem-se a ruídos no CCD ou CMOS dos equipamentos de captura das imagens (Figura 2.8), bem como do tipo de compactação dessas imagens. Estes sensores captam apenas informações monocromaticamente.

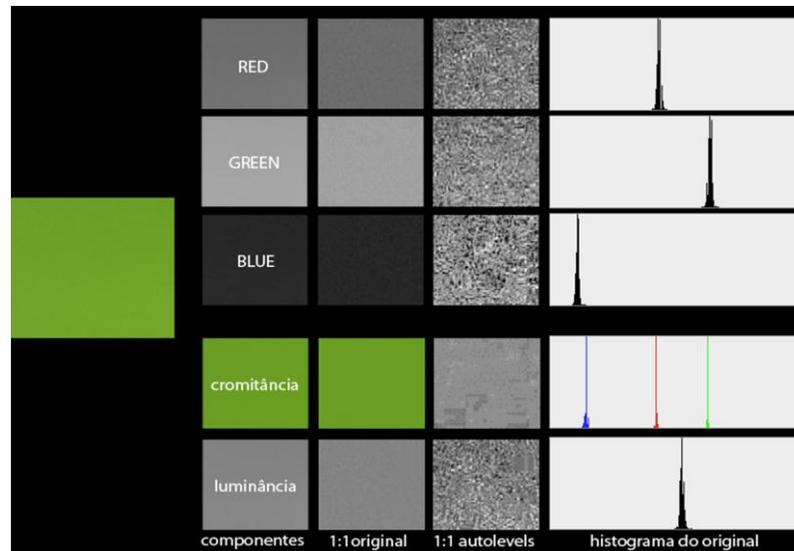


Figura 2.8 Ruídos nos canais

Desta maneira, precisamos filtrar a luz recebida em um prisma e, portando, precisamos de um sensor para cada cor primária (RGB).

Em 1976, Bryce Bayer patenteou um filtro que permite ter duas vezes mais verde do que os outros dois sensores (Figura 2.9), obtendo cores mais precisas e com menos ruído. A idéia deste filtro é imitar a percepção humana. Dessa

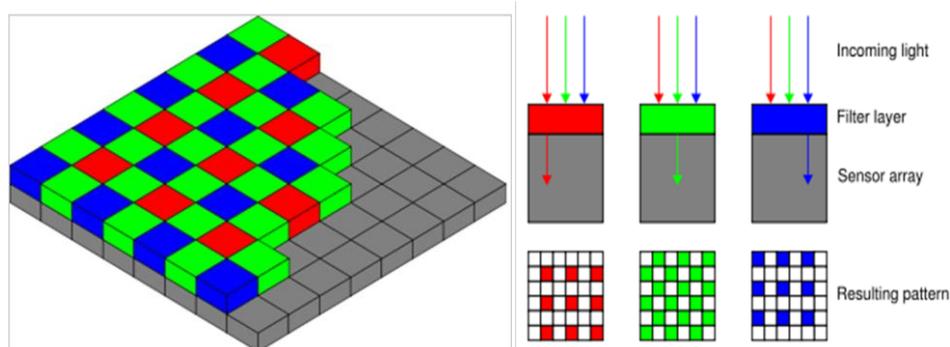


Figura 2.9 Padrão de Bayer para disposição de sensores usado pela maioria das câmeras digitais [Imagens sob GNU FD License, Wikimedia Commons, 28 Dezembro 2006]

forma, apesar da cor Azul ser complementar à cor da pele humana, o fundo azul passou a ceder lugar ao verde nas técnicas de chroma key, devido à

capacidade dos equipamentos digitais captarem maior luminância e informação neste canal.