4
Uma Proposta de Técnica Chromaless

À luz das análises feitas com produtos comerciais e algoritmos propostos na literatura de matting, este capítulo apresenta uma proposta de técnica chromaless capaz de lidar com imagens HD em tempo real.

4.1.
Círculo de Amostragem

Embora o Soft Scissors tenha um desempenho muito alto, encontramos dificuldades para tratar de imagens em alta definição (HD). A principal dificuldade é que as funções de energia a serem minimizadas caracterizam um tratamento mais global da imagem e que depende da sua resolução. Nossa proposta é buscar uma técnica que trate a questão de matting localmente. Por esta razão não vamos gastar tempo calculando clusters de pixels definitivamente foreground e definitivamente background. A idéia é pegar uma vizinhança do ponto a ser classificado e usar um critério mais elaborado para o cálculo do alpha. Sendo um processo local podemos pensar futuramente em partir para uma programação paralela em GPU ou em sistemas multicore com uso intensivo de memória cache.

Neste trabalho não vamos tratar do método que gera o trimap - algo que posteriormente poderemos implementar com a ajuda de um pincel utilizado pelo usuário (como no Soft Scissors). Portanto, partimos do trimap com uma região preta (valor 0) representando os pixels definitivamente background, uma região branca (valor 1) representando os pixels definitivamente foreground e uma região cinza representando os pixels mistos desconhecidos.

Propomos a definição de um círculo de amostragem em torno do pixel pertencente à região desconhecida do trimap para posterior processamento com métricas em mais de um espaço de cor. Para construir este círculo, partimos de uma janela de tamanho s (p.ex. 8 x 8 pixels) centrada no pixel cinza U que queremos classificar. O tamanho s da janela é estabelecido de forma a garantir

PUC-Rio - Certificação Digital Nº 0721340/CA

que haja pixels pretos e brancos. O raio r do círculo de amostragem, centrado no pixel cinza U, é dado por:

$$r = (1+p)\left[\frac{d_F + d_B + 2s}{2}\right] \qquad \text{Eq. 4.1}$$

onde d_F é a distância do pixel U ao pixel branco mais próximo, d_B é a distância do pixel U ao pixel preto mais próximo, s é o tamanho da janela e p é coeficiente de segurança $0 \leq p \leq 1$ (Figura 4.1). A idéia aqui é garantir que o círculo de amostragem é maior do que a janela e contém suficientes pixels mais próximos F e B. O raio r se adapta dinamicamente à largura da faixa de pixels mistos à medida que outros pixels vão sendo classificados. O coeficente de segurança p

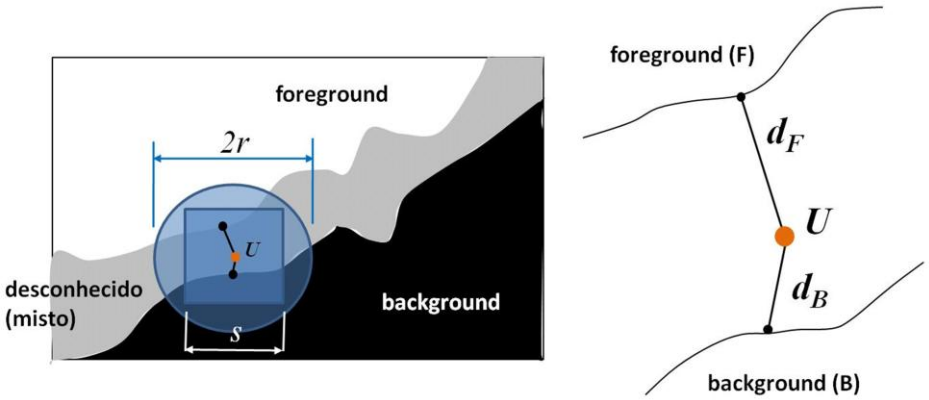

Figura 4.1 Círculo de amostragem de raio r centrado no pixel misto U, janela de tamanho s e distâncias d_F e d_B

é um parâmetro de controle que nos testes do presente trabalho está definido como sendo 20% (*i.e.* p = 0.20).

4.2. Cálculo de Alpha

Os pixels dentro do círculo de amostragem são mapeados no espaço de cor P. A escolha do espaço de cor é um parâmetro de análise e, no presente trabalho, consideramos $P = RGB$ e $P = LAB$.

Com inspiração em modelos físicos, propomos que a classificação do pixel seja equivalente ao cálculo de potencial elétrico ou ao cálculo de energia potencial elétrica. Em quaisquer dos dois casos, pixels são tratados como partículas eletricamente carregadas no espaço de cor P. Numa primeira maneira de classificar, o valor de alpha é dado por uma equação semelhante à do cálculo

PUC-Rio - Certificação Digital Nº 0721340/CA

do valor do potencial no ponto U (que representa o pixel misto a ser classificado):

$$\alpha = V_U = K\sum_{i=1}^{n}\left(\frac{q_i}{d_i}\right) \qquad \text{Eq. 4.2}$$

onde *n* é o número total de pixels F e B no círculo de amostragem, q_i é a carga elétrica (igual a -1 para background e +1 para foreground) da cor i, d_i é a distância do ponto U à partícula i, e K é o equivalente à constante de Coulomb (considerada unitária neste trabalho, *i.e.* K = 1). A distribuição das cores como cargas elétricas está ilustrada na Figura 4.2.

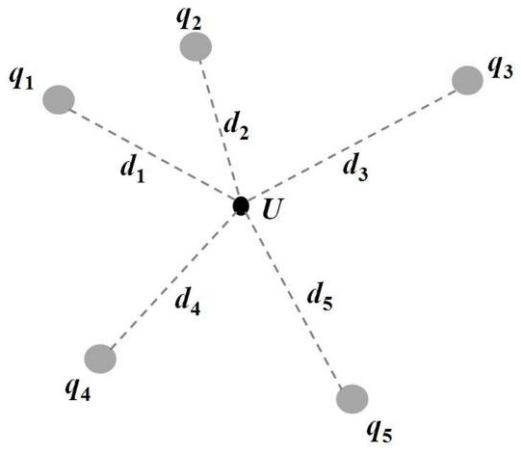

Figura 4.2 Distribuição das cores como cargas elétricas no espaço de cor P (RGB ou LAB, onde q1, q2, e q3 são cores foreground (com carga elétrica +1), q4 e q5 são cores background, di são distâncias e U é o pixel a ser classificado.

A outra maneira de classificar U é considerar o valor de alpha como sendo a porção da energia potencial elétrica das cores foreground em relação à energia total. Neste caso a energia potencial elétrica foreground V_F é dada por:

$$V_F = K\left(\frac{q_1q_2}{d_{12}} + \frac{q_1q_3}{d_{13}} + \frac{q_2q_3}{d_{23}} + \cdots + \frac{q_iq_j}{d_{ij}} + \cdots\right) \quad i = 1, ..., n_F \text{ e } j = 1, ..., n_F \qquad \text{Eq. 4.3}$$

onde d12 é a distância entre as cores i e j, n_F é o número de cores foreground no círculo de amostragem, q_i é o valor de carga elétrica da cor i e K = 1. De maneira semelhante temos a energia potencial background V_B calculada pela seguinte expressão:

$$V_B = K\left(\frac{q_1q_2}{d_{12}} + \frac{q_1q_3}{d_{13}} + \frac{q_2q_3}{d_{23}} + \cdots + \frac{q_iq_j}{d_{ij}} + \cdots\right) \quad i = 1, ..., n_B \text{ e } j = 1, ..., n_B \qquad \text{Eq. 4.4}$$

Neste caso, o valor de alpha é dado por:

$$\alpha = \frac{V_F}{V_F + V_B} \qquad \text{Eq. 4.5}$$

PUC-Rio - Certificação Digital Nº 0721340/CA

4.3.
Tratamento Global com Clusterização por K-Means

Várias técnicas de matting partem da identificação de clusters de pontos amostrados a partir das áreas conhecidas de foreground e background. Neste trabalho foi feita uma tentativa de classificação baseada em clusters, como alternativa ao método anteriormente apresentado.

Para o calculo de alpha, todos os pixels marcados como foreground são dispostos no espaço 3D e, através de uma função randômica, amostras desses pixels são sorteadas. Essas amostras viram candidatas a receber uma esfera de influência, denotada como cluster. O sistema define previamente a quantidade máxima de clusters que deverão ser posicionados ao longo da mancha de pixels tidos como foreground. Uma vez todos posicionados, o mesmo é feito para os pixels marcados como background.

Apos o posicionamentos de todos os clusters, começa o processo de clusterização, que é expandir essas esferas de forma a coletar amostras de pixels próximas. O processo começa com a varredura de todos os pixels de background e foreground e termina após termos todos os pixels posicionados em algum cluster (seja B ou F). Para tanto, os pixels são varridos e associados ao cluster com a superfície mais próxima. O cluster então tem o seu centro deslocado de um fator na direção do centróide dos pixels associados a ele nesta etapa.

O raio de influência do cluster (supondo que seja de B) é dado tomando-se a distância dele ao cluster de F mais próximo e multiplicando pelo fator raio, a ser definido pelo usuário. Quanto maior esse fator, o alpha resultante é mais binarizado, isto é: há menos gradação de cinzas, praticamente sendo um matte composto exclusivamente por brancos e pretos.

Em alguns casos, um cluster pode ficar sendo atraído por direções opostas e resultar numa flutuação. Para resolver isto, uma variável chamada “traveled distance” foi implementada. Quando percebemos que algum cluster está “flutuando”, ele é duplicado; fazendo com que cada cluster siga para uma nova direção, resultando em uma melhor cobertura.

PUC-Rio - Certificação Digital Nº 0721340/CA

A Detecção da flutuação de um cluster ocorre observando o quanto esse cluster, ao longo do processo de treinamento de posição, tem o valor de seu “traveled distance” alterado. Isto é, não importa se ele andou para trás ou para frente, é sempre somado o quanto foi deslocado. Quando este valor excede uma constante arbitrária, o cluster é duplicado, fazendo com que cada um seja atraído para cada lado que antes os disputavam.

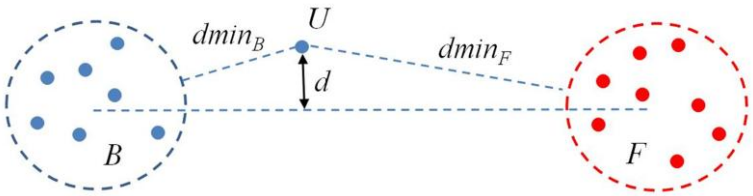

Figura 4.3 Distâncias mínimas de foreground e background no espaço de cor, do ponto U até às superfícies dos clusters, e distância de U até o eixo entre os clusters.

Uma vez que todos os pixels pertencem a um cluster, os pixels (que no trimap foram marcados como desconhecidos) são classificados, recebendo duas informações, o valor do alpha e sua confiabilidade. Para cada um desses pixels, acha-se o cluster de foreground e background mais próximo e a medida é feita em relação à superfície do cluster e não ao seu centro. O valor de alpha então é atribuído pela seguinte fórmula (Figura 4.3) e o grau de confiabilidade é dado por um valor de 1/10 do valor de alpha.

$$\alpha = \frac{d\min_B}{d\min_B + d\min_F} \qquad \text{Eq. 4.6}$$

Uma outra maneira mais global de calcular o alpha seria procuramos cores de foreground e background que minimizassem a distância d que é a menor distância da coordenada de cor do pixel U com o segmento que une os eleitos (Figura 4.3). Esta maneira é, entretanto, muito custosa porque teríamos de percorrer todos os pixels.

PUC-Rio - Certificação Digital Nº 0721340/CA