

## Segmentação de imagens de profundidade por limiarização com morfologia matemática

Jorge Antonio Silva Centeno <sup>1</sup>  
Regina Tiemy Kishi <sup>2</sup>

Universidade Federal do Paraná - UFPR

<sup>1</sup>Departamento de Geomática

<sup>2</sup> Departamento de Hidráulica e Saneamento

CEP 81531-990 Curitiba/PR, Brasil

{centeno; ;rtkishi.dhs}@ufpr.br

**Abstract.** This paper introduces a method to segment depth images that are obtained by a ToF camera. The ToF (Time of flight) device is a camera that captures the distance between the camera and the objects within its field of view. In order to recognize the object in the foreground, it is necessary to segment the range image. The separation between the foreground and the background is performed by histogram thresholding. The threshold is obtained by computing the difference between the histogram and another function that is obtained after dilating the histogram. The method is described and an example is presented.

**Palavras-chave:** Range images; histogram thresholding; image processing.

### 1. Introdução

Tradicionalmente a fotogrametria e os levantamentos topográficos têm sido utilizados para a coleta de dados tridimensionais de objetos. A Fotogrametria terrestre é um método clássico de coleta de imagens para a formação de modelos estereoscópicos que permitem a visão tridimensional. Para isto, pelo menos duas fotografias cobrindo a área de interesse são necessárias. Atualmente, as imagens são capturadas usando sensores digitais CCD ou CMOS, facilitando o processamento digital das mesmas para a derivação de informação tridimensional. Embora grande parte do trabalho possa ser executada usando algoritmos computacionais, o mesmo ainda demanda intervenção humana.

Um dos mais recentes avanços no campo da fotogrametria é conhecido como câmara de distância ou *range imaging* - RIM (Kolb et al (2009)). Estes dispositivos permitem obter imagens usando a tecnologia digital, com a diferença de que cada pixel não armazena a cor refletida pelas superfícies, mas sim a distância que separa o sensor de cada ponto visível na cena. Cada pixel registra a informação associada à distância a um ponto da cena e também uma imagem co-registrada da intensidade com alta taxa de repetição (A intensidade medida depende fortemente do comprimento de onda usado (geralmente é utilizado o infravermelho próximo), da fonte da iluminação e da característica da superfície).

Um dos grandes desafios na atualidade é usar estes dispositivos para capturar movimentos do usuário e assim facilitar a interação humano-computador, suprimindo, em parte, o uso de teclado ou mouse, substituindo a entrada de comando pela captura de movimentos do corpo.

O presente trabalho mostra um algoritmo de segmentação de imagem de profundidade baseado nos princípios da morfologia matemática.

### 2. Metodologia de Trabalho

Os sensores RIM capturam dados dos pontos em uma superfície usando o princípio do tempo de voo (time of flight - TOF). A câmera usa uma fonte eletromagnética para iluminar a cena e a superfície do objeto reflete parte da energia incidente na direção do sensor. Quando a radiação eletromagnética refletida é suficiente para ativar o sensor, é possível calcular a distância entre a câmera e o objeto, segundo a relação mostrada na equação 1.

$$R = (c \cdot dt) / 2 \quad (1)$$

Um grande número pontos é registrado porque a superfície completa de um CMOS é usada como elemento de medição. Detalhes do funcionamento deste tipo de equipamento podem ser encontrados em Oliveira e Centeno (2011) ou Lange (2000).

O sensor de um RIM utiliza a tecnologia CMOS para a captura dos dados. A diferença entre o CMOS de um sensor RIM e um CMOS usado na fotografia digital é que o sensor RIM utiliza detectores (“*smart pixel*” Lange (2000) ) para medir a informação de distância associada a um pixel; O sensor RIM calcula a distância aos objetos medindo o deslocamento de fase ( $\phi$ ) do sinal refletido em relação ao sinal emitido pela fonte luminosa para cada pixel. Consequentemente o modulador da fonte luminosa está conectado diretamente a cada pixel.

Como Dorrington et al. (2009) explica, a distância ( $R$ ) pode ser calculada a partir da diferença de fase ( $\phi$ ) e da frequência de modulação ( $f_m$ ):

$$R = (c \cdot \phi) / (4 \cdot \pi \cdot f_m) \quad (2)$$

Os dados de cada pixel mostram a distância entre o sensor e os objetos na cena. Para separar os elementos mais próximos ao sensor do fundo, é necessário segmentar a imagem de profundidade.

Os métodos de segmentação global assumem que os pixels que pertencem a um objeto são espectralmente similares e diferentes dos pixels de outros objetos. É assumido, ainda, que os pixels de cada objeto são vizinhos, ou seja, existe conectividade nas regiões. Esta suposição é muito restritiva, motivo pelo qual os métodos de limiarização são aplicados para casos mais específicos, quando a separação entre objetos, ou áreas de interesse é bem definida. Na aplicação aqui apresentada, os pixels do objeto a ser detectado são conexos.

A tarefa, neste caso, é detectar os limites entre classes (Objeto e Fundo). No caso mais simples, o limite (Limiar) entre duas classes poderia ser descrito por um único valor digital. Valores superiores a este limite pertenceriam a um objeto e os menores a outro objeto ou o fundo.

$$\begin{aligned} g(x) &= 1 \text{ se } f(x) \geq T \\ g(x) &= 0 \text{ se } f(x) < T \end{aligned} \quad (3)$$

A limiarização é o método mais simples, no qual a variação dos pixels na imagem é analisada de forma global, com ajuda do histograma da imagem. O Histograma representa a frequência relativa dos valores digitais. Considerando apenas duas classes, por exemplo, um objeto e o fundo, como são mostrados na figura 3, é esperado que o histograma apresente duas grandes concentrações de valores, distâncias altas (o fundo) e baixas (o objeto).

Na literatura (por exemplo, Sahoo, Soltani e Wong, 1988) existem diversos métodos propostos para a determinação de um limiar que separe os dois agrupamentos, porém constata-se que não existe ainda uma solução geral e que o sucesso de cada método depende da natureza da imagem. Na revisão de métodos de limiarização de Sahoo, Soltani e Wong (1988) são descritos vários métodos.

O método aqui proposto é aplicar a morfologia matemática ao histograma. A ideia central, neste caso, é preencher os vales do histograma de distâncias aplicando o método de fechamento para detectar os mínimos locais do histograma.

Isto pode ser explicado formalmente da seguinte forma. Sendo  $H(R)$  o histograma da imagem de distância e  $G(R)$  o resultado da sucessiva aplicação de um operador de dilatação

que adota o maior valor da vizinhança para cada pixel, então  $G(r)$  é sempre maior que  $H(R)$ . A diferença  $D(R)=G(R)-H(R)$  é positiva nos vales do histograma.

Inicialmente, o histograma da imagem é calculado. para isto é necessário discretizar as distâncias armazenadas nos pixels. Neste exemplo, 256 valores foram adotados, mas isto pode ser variado. Para remover pequenos picos, um filtro passa baixas é aplicado ao histograma. este filtro reduz também os picos do histograma, mas produz uma curva mais suave. A figura 1 mostra um exemplo de histograma suavizado. O comportamento geral da curva não é significativamente alterado.

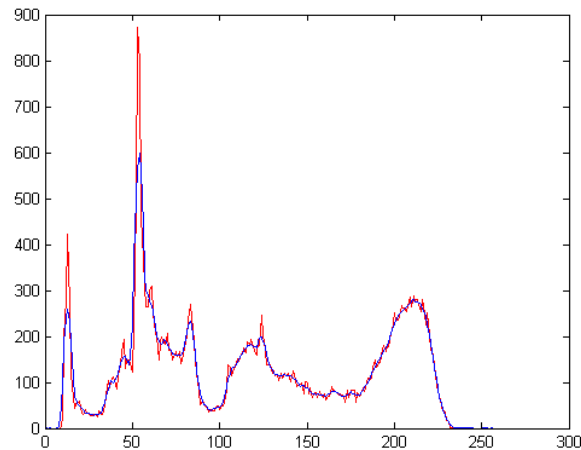


Figura 1 - suavização do histograma. Original em vermelho, suavizado em azul.

O seguinte passo é dilatar a curva, adotando o maior valor da vizinhança para cada elemento do histograma. No exemplo mostrado na figura 2, o tamanho da vizinhança é de 13 elementos. Finalmente, a diferença entre o histograma dilatado e o suavizado é calculada.

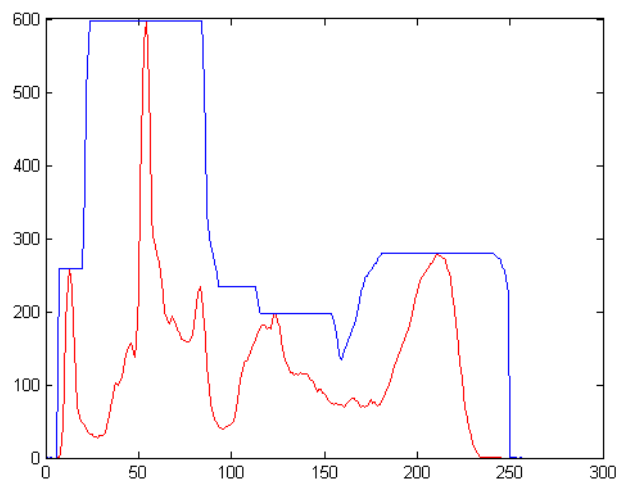


Figura 2 - dilatação (azul) do histograma (vermelho).

### 3. Resultados e Discussão

A seguir é apresentado um exemplo da aplicação do método. A Figura 3 mostra uma imagem de profundidade, onde os pontos mais próximos aparecem escuros e os mais distantes claros.

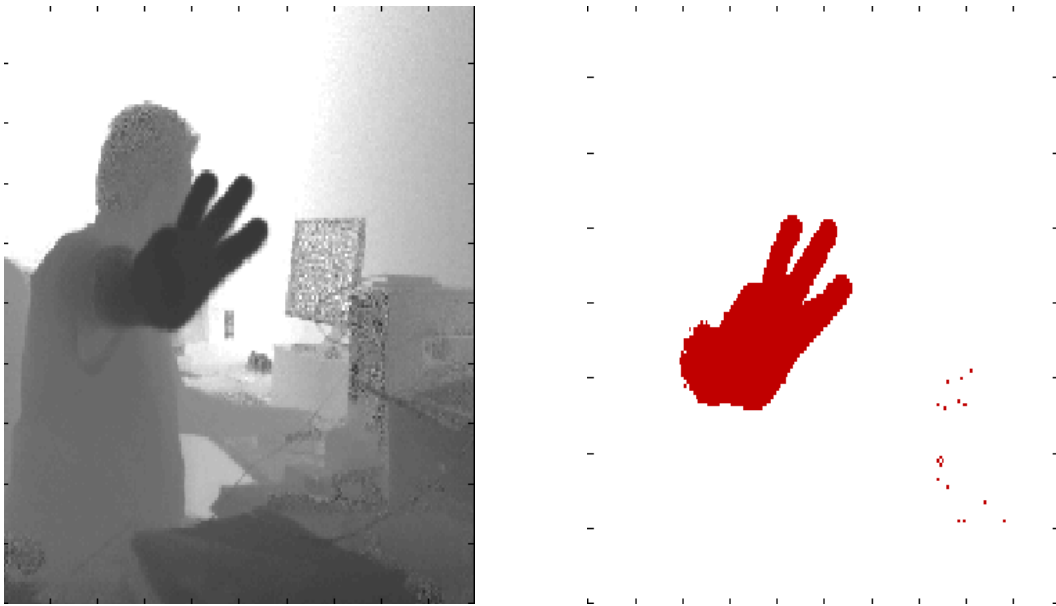


Figura 3 - Imagem de profundidade (a) e imagem Segmentada (b).

A figura 4 mostra o histograma filtrado, em vermelho, e a diferença entre o histograma dilatado e o suavizado, em azul. Para fins de melhor entendimento, o primeiro pico da diferença foi pintado de cinza.

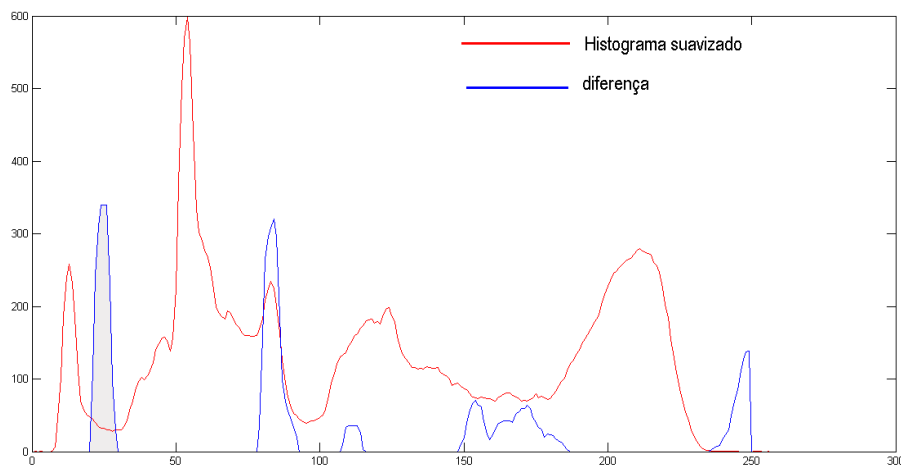


Figura 4 - diferença da dilatação e histograma suavizado.

A determinação do limiar resulta da identificação do máximo local da função diferença. A adoção deste limiar permite segmentar a imagem, como mostra a figura. Finalmente, aplicando o operador de abertura, os pequenos pontos isolados podem ser removidas desta imagem. A figura 3b mostra o resultado da binarização.

#### 4. Conclusões

Foi apresentado um método para segmentação de imagens de profundidade baseado na aplicação do operador de dilatação ao histograma de distâncias. Os resultados mostram que este método é viável, embora deva ser salientado que sua aplicação demanda a que não exista mistura entre o objeto e o fundo em termos de distância. O método, porém, mostra-se

eficiente, pois soluciona um problema que, se fosse tratado com imagens fotográficas convencionais, não permitiria separar o objeto da frentes se este não tem cores diferentes do fundo. No método aqui mostrado, as cores do fundo, ou a ocorrência de outros objetos não prejudicam o desempenho.

### **Agradecimentos**

Os autores agradecem ao CNPq, que permitiu a aquisição do equipamento para a realização desta pesquisa.

### **Referências Bibliográficas**

Dorrington, A.A., Kelly, C.D.B., Shane H. McClure, S.H., Payne, A.D. & Cree, M.J.; **Advantages of 3D Time-of-Flight Range Imaging Cameras in Machine Vision Applications**. The 16th Electronics New Zealand Conference (ENZCon), Dunedin, New Zealand, 18-20 November, 2009

Kolb, A., Barth, E., Koch, R., **Larsen R. Time-of-Flight Sensors in Computer Graphics**. EUROGRAPHICS 2009/ M. Pauly and G. Greiner STAR – State of The Art Report

Lange, R. **3D Time-of-Flight distance measurement with custom solid-state image sensors in CMOS/CCD-technology**. 205 p. (Dissertation submitted to the Department of Electrical Engineering and Computer Science at University of Siegen for the degree of Doctor of Technical Sciences). Univesity of Siegen, Alemanha, 2000.

Oliveira, A.A.A. & Centeno, J.A.S.. 2011. Modelagem tridimensional de superfícies planas a partir de imagem TOF. Bol. Ciênc. Geod., vol.17, n.3, pp. 361-378

Sahoo, P.K., Soltani, S., Wong, A.K.C., **A survey of thresholding techniques**, Comput. Vision, Graphics Image Processing 41 (2) (1988) 233--260.