

Estudo sobre Detecção de Bordas em Imagens de Profundidade

LUCIANO SILVA
DANIEL WEINGAERTNER
OLGA REGINA PEREIRA BELLON (*orientadora*)
UFPR - Departamento de Informática
Caixa Postal 19081, 81531-990 Curitiba, PR, Brasil
{danielw,luciano}@lobo.inf.ufpr.br
olga@inf.ufpr.br

Abstract. Edge detection is a very important stage in the processing of image information [Marr--Hildreth (1980)]. When the edge detection is performed on range images it provides much more information than the edge detection based only on intensity images because range images have explicit 3-D information. In this paper, we present a study of how to detect step and roof edges on range images, that, when combined, can be used, for example, to give an approximate number of image's regions. We also developed a mathematical morphology based algorithm to perform thinning operation on the detected edges.

Keywords: range images, edge detection, step edge, roof edge, thinning.

1. Introdução

As imagens de intensidade luminosa são as mais conhecidas e utilizadas em processamento de imagens. Estas imagens são obtidas por meios ópticos [Foley et al. (1996)] como uma câmera ou scanner óptico, e o valor de cada ponto da imagem (*pixel*) é uma função da intensidade de luz refletida pelo ponto correspondente na cena.

A maior deficiência das imagens de intensidade luminosa é a ausência da informação de profundidade, perdida na projeção 3D-2D [Gonzalez--Woods (1993)].

Por outro lado, as imagens de profundidade (*range images*) [Bellon (1996); Besl (1988); Jarvis (1983)] tem sido cada vez mais utilizadas em processamento de imagens e visão computacional. Estas imagens podem ser obtidas, por exemplo, através de métodos como a triangulação e métodos baseados no tempo de reflexão, utilizando ultrassom ou raios *laser* [Besl (1988); Jarvis (1983)]. Elas são, na verdade, um conjunto de medidas de distância da origem de um sistema de coordenadas até a superfície do objeto na cena observada. Sua grande vantagem em relação às imagens de intensidade luminosa é possuir informações tridimensionais (3D) explícitas sobre a cena observada, permitindo-nos assim extrair diversas características importantes para a interpretação da imagem como, por exemplo, classificação de bordas.

Nas imagens de profundidade, podemos identificar diversos tipos de bordas [Hoffman--Jain (1987)], sendo as de maior aplicação as bordas de profundidade (*step edge*) e as bordas de orientação (*roof edge*). As bordas de profundidade são compostas por pontos que estão na interseção entre regiões com valores de profundidade significativamente distintos, segundo um valor de limiar pré-definido. As bordas de orientação estão localizadas no encontro de superfícies com orientações distintas (por exemplo no encontro das faces de um cubo). Elas delimitam diferentes faces de um mesmo objeto, ou de objetos distintos que se tocam a uma mesma profundidade.

A detecção de bordas é uma etapa fundamental no processamento de imagens pois as bordas delimitam o tamanho das superfícies dos objetos e fornecem informações importantes na extração de diversas características da cena. Porém os métodos clássicos de detecção como os de Canny [Canny (1986)] não apresentam bons resultados em imagens que apresentam tipos diferentes de borda. Por exemplo, uma imagem pode conter tanto bordas de profundidade quanto de orientação, e estes detectores não localizarão corretamente ambos os tipos.

Neste artigo, apresentamos uma técnica para detectar bordas de profundidade e de orientação em imagens de profundidade e, a partir da combinação e fechamento destas bordas, construir o mapa de bordas da imagem, que pode

então ser utilizado, por exemplo, para determinar o número de regiões que compõem a imagem.

Inicialmente, no tópico 2, discutimos alguns trabalhos relacionados ao tema em questão. No tópico 3, fazemos uma breve apresentação sobre as imagens de profundidade. No tópico 4 falamos sobre detecção de bordas de maneira geral. Na sequência, no tópico 5, apresentamos uma técnica para cálculo do gradiente de profundidade. No tópico 6, mostramos como calcular as normais de uma imagem e, a partir delas calcular o gradiente de orientação. No tópico 7, apresentamos algumas metodologias para definição de limiar para detecção de bordas e, no tópico 8, uma técnica de afinamento baseada em morfologia matemática. Finalmente, apresentamos os resultados experimentais e a conclusão.

2. Trabalhos Relacionados

O problema de detetar bordas em imagens de profundidade tem sido bastante estudado nos últimos anos, e diversas abordagens foram feitas.

Mintz [Mintz (1993)] sugere um método para detetar bordas de profundidade e de orientação em que um *pixel* é declarado de borda somente se há um consenso entre vários processos, que tentam determinar se o *pixel* está numa região de descontinuidade, e não baseado no uso de filtros ou valores de limiar. Para tanto, são utilizados métodos de estimativa robustos para estimar um ajuste local à janelas na vizinhança do *pixel* e acumular votos para cada janela. Mitiche and Aggarwal [Mitiche--Aggarwal (1983)] utilizam modelos probabilísticos para identificar bordas em imagens de profundidade. O objetivo é que método desenvolvido seja imune à existência de ruídos na imagem. Nitzan et al. também utiliza uma abordagem estatística para construir um mapa de bordas para a combinação de duas imagens, uma de profundidade e outra de orientação. O trabalho de Haddon [Haddon (1988)] é uma boa referência para o estudo de abordagens para a determinação de valores de limiar em técnicas de detecção de bordas. Como referências básicas para o estudo do problema de detecção de bordas, podem ser citados [Marr--Hildreth (1980)], [Canny (1986)] e [Pratt (1991)].

3. Imagens de Profundidade

As imagens de profundidade (*range images*) tem sido cada vez mais utilizadas em processamento de imagens e visão computacional. Estas imagens podem ser obtidas, por exemplo, através de métodos como a triangulação e

métodos baseados no tempo de reflexão, utilizando ultrassom ou raios *laser* [Besl (1988)], [Jarvis (1983)]. Elas são na verdade um conjunto de medidas de distância da origem de um sistema de coordenadas até a superfície do objeto na cena observada. Sua grande vantagem em relação às imagens de intensidade luminosa é que elas possuem informações tridimensionais (3D) explícitas sobre a cena observada, permitindo-nos assim extrair diversas características importantes para a interpretação da imagem como, por exemplo, classificação de bordas.

3. Detecção de Bordas

Mudanças ou descontinuidades na amplitude de uma determinada característica da imagem são de fundamental importância no processamento de imagens, pois elas nos fornecem uma idéia da extensão e divisão dos objetos da mesma. Aos *pixels* que estão na fronteira entre duas regiões com características relativamente distintas, aonde ocorrem estas descontinuidades, chamamos de bordas.

Para determinar se um *pixel* da imagem é ou não de borda, calculamos o gradiente (que quantifica a intensidade da variação entre o *pixel* e seus vizinhos) deste *pixel* e, caso o gradiente seja maior do que um valor de limiar pré-definido, dizemos que o *pixel* é de borda.

Em imagens de intensidade luminosa, as bordas são diferenças bruscas na intensidade ou na cor da imagem. Já em imagens de profundidade, podemos encontrar vários tipos de bordas, sendo as mais importantes: bordas de profundidade e bordas de orientação.

4. Bordas de Profundidade

Bordas de profundidade são variações bruscas no valor de profundidade da imagem. Elas são compostas por pontos que estão na interseção entre regiões com valores de profundidade significativamente distintos, e podem ser detetadas utilizando-se operadores de gradiente, como por exemplo os operadores Sobel [Gonzalez--Woods (1996)].

Neste trabalho, utilizamos como gradiente para detetar bordas de profundidade a máxima diferença entre cada *pixel* da imagem e seus vizinhos numa máscara 3x3, ou seja, o gradiente de profundidade $G_p(i,j)$ do *pixel* (i,j) é dado por:

$$G_p(i, j) = \max(|w_{ij} - W|) \quad (1)$$

onde:

w_{ij} é a profundidade do *pixel* (i, j) .

W são os valores de profundidade dos pontos vizinhos a w_{ij} .

Embora esta abordagem seja relativamente sensível ao ruído, ela tem mostrado bons resultados, mesmo em imagens reais.

5. Bordas de Orientação

Bordas de orientação estão localizadas no encontro de superfícies com orientações distintas (por exemplo no encontro das faces de um cubo). Elas delimitam diferentes faces de um mesmo objeto, ou de objetos distintos que se tocam a uma mesma profundidade.

Para detetar bordas de orientação em uma imagem de profundidade, inicialmente calculamos a orientação da normal que passa por cada *pixel* da imagem (utilizando as coordenadas tridimensionais (x,y,z) de cada *pixel*) e o valor de gradiente é dado pelo ângulo formado entre as normais.

O cálculo dos coeficientes da normal $\vec{N}_{ij} = \{a_{ij}, b_{ij}, c_{ij}\}$ que passa pelo *pixel* (i,j) de coordenadas 3-D $\{x_{ij}, y_{ij}, z_{ij}\}$, é feito ajustando-se um plano a uma máscara $n \times n$ centrada no *pixel* (i,j) .

Sejam $\vec{V}_N, \vec{V}_L, \vec{V}_S, \vec{V}_O$, vetores do *pixel* (i,j) até os pontos que estão nas extremidades da máscara nas direções “Norte”, “Leste”, “Sul”, e “Oeste” (vizinhança-4), respectivamente, então:

$$\vec{N}_{ij} = \frac{\sum (\vec{V}_D \wedge \vec{V}_{next(D)})}{4} \quad (2)$$

onde:

$$D = \{N, L, S, O\}.$$

$next(D)$ é a próxima direção no sentido horário.

Desta forma, a normal \vec{N}_{ij} é a média do produto vetorial de quatro vetores tomados dois a dois dentro de uma máscara $n \times n$.

Para evitar as distorções ou cálculo de normais inexistentes, que podem ocorrer caso haja uma grande variação de profundidade dentro da máscara à qual tentamos adaptar um plano, evitamos que esta máscara seja cortada por bordas de profundidade. Isto é feito tomando-se o cuidado de não deixar nenhum vetor \vec{V}_D cruzar uma borda de profundidade. Caso não seja possível calcular um ou mais dos vetores devido à proximidade das bordas de profundidade, a normal é calculada em função apenas dos vetores encontrados. Para os pontos cuja normal não pôde ser calculada, atribuímos o valor da normal do último *pixel* vizinho cuja normal foi calculada com êxito.

Com a imagem das normais (composta pelos coeficientes $\{a,b,c\}$ da reta normal a cada *pixel* da cena), podemos tomar o gradiente de orientação $Go(i,j)$ de um

pixel (i,j) como sendo o maior ângulo formado entre a normal ao *pixel* (i,j) e as normais a seus vizinhos numa máscara 3×3 , ou seja:

$$\theta_{kl} = \arccos \left(\frac{\|\vec{N}_{ij}\| \cdot \|\vec{N}_{kl}\|}{\vec{N}_{ij} \bullet \vec{N}_{kl}} \right) \quad (3)$$

$$Go(i, j) = \max(\theta_{kl}) \quad (4)$$

onde:

\vec{N}_{kl} são os vetores normais aos pontos vizinhos de (i,j) .

θ_{kl} é o ângulo formado entre os vetores \vec{N}_{ij} e \vec{N}_{kl} .

Em muitas situações, algumas bordas de profundidade também são detetadas como sendo de orientação. Isto ocorre devido ao fato de que regiões com variação de profundidade também podem ter uma variação de orientação. Para evitar que bordas de profundidade sejam detetadas como de orientação, sempre que um *pixel* fizer parte de uma borda de profundidade seu gradiente de orientação será setado como zero, evitando a detecção repetida. Existem outras técnicas [Mintz (1994)], [Mitiche--Aggarwal (1983)], [Nitzan et al. (1977)] mais complexas que podem ser utilizadas para eliminar o efeito dos *pixels* das borda de profundidade no cálculo das bordas de orientação. Estas técnicas encontram-se atualmente em estudo.

6. Detecção de Bordas em Imagens de Gradiente

A detecção de bordas em imagens de gradiente consiste basicamente em verificar o valor de gradiente $Gx(i,j)$ de todos os pontos da imagem e, aqueles que tiverem este valor acima de um determinado limiar L pré-definido, são marcados como sendo de borda.

Conforme [Haddon (1988)], a seleção do limiar é tradicionalmente feita em uma de quatro maneiras:

1. O limiar é escolhido iterativamente pelo usuário com base na observação da imagem.
2. O valor de limiar é $\pm n$ vezes o desvio padrão mais a média do gradiente da imagem.
3. Escolher o limiar de forma que uma determinada porcentagem dos pontos é marcada como sendo de borda (normalmente 5% dos pontos da imagem).
4. Em imagens em que a distribuição das bordas é multimodal, o limiar é escolhido como sendo um valor entre os picos.

Neste trabalho, utilizamos tanto a primeira quanto a segunda maneira de selecionar o limiar, sendo, no último caso, $n=1.5$.

7. Afinamento em Imagens de Bordas

O afinamento (*thinning*) [Gonzalez--Woods (1993); Pratt (1991)] é um processo que pode ser utilizado para definir mais precisamente o formato das regiões de uma imagem e a localização dos *pixels* de borda. No caso em estudo, o afinamento é utilizado para eliminar os *pixels* detetados erroneamente como sendo de borda.

A técnica de afinamento de bordas baseada em Morfologia Matemática [Facon (1996)], que apresenta um conjunto de ferramentas para extração de componentes específicos da forma dos objetos da imagem, vem sendo amplamente utilizada em processamento de imagens. Ferramentas como dilatação (*dilation*), erosão (*erosion*), abertura (*opening*), fechamento (*closing*), entre outras, são úteis para melhorar a representação e descrição do formato das regiões em uma imagem.

As máscaras para afinamento utilizadas neste trabalho foram desenvolvidas baseadas em Morfologia Matemática e na observação e estudo das diversas características obtidas pelas combinações de *pixels* em máscaras de tamanho 3x3 e 5x5. Com as máscaras de tamanho 3x3 Figura 1(a)-(b) conseguimos abranger todas as situações nas quais não podemos eliminar o *pixel* central sem que surjam descontinuidades na borda da imagem. Na Figura 1(c) temos uma “máscara reversa” que, centrada num *pixel* que não é de borda, indica a situação em que devemos marcá-lo como sendo de borda, para mais tarde corroer em volta dele, mantendo a direção da borda. As máscaras da Figura 1(d) e Figura 2(a)-(b) evitam que o afinamento corra linhas na imagem.

As demais máscaras representadas na Figura 2, foram baseadas no método *pruning* [Gonzalez--Woods (1993)] que complementa o afinamento, e desenvolvidas com o objetivo de eliminar ruídos junto as bordas e evitar que uma borda representada por uma reta na imagem seja totalmente removida. Caso o *pixel* não esteja em uma situação representada nas Figuras 1-2, ele será eliminado.

Nosso algoritmo de afinamento de bordas, convoluciona as máscaras representadas nas Figuras (1)-(2) e suas variantes (rotações de 90°, 180° e 270° da máscara original) sobre a imagem de borda, iniciando em um dos quatro cantos da imagem. A opção de fazer a convolução em várias direções (iniciando no canto superior esquerdo,

superior direito, inferior esquerdo, ou inferior direito) foi adicionada com base nos experimentos, pois observamos que dependendo da orientação das bordas na imagem, uma ou outra direção de convolução produz melhores resultados.

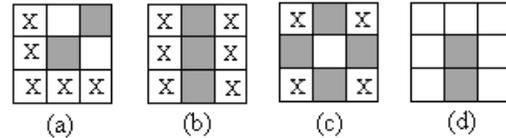


Figura A: Máscaras 3x3

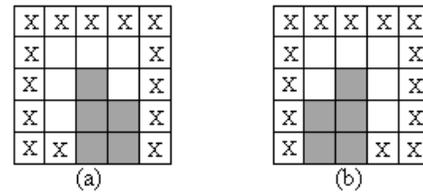


Figura B: Máscaras 5x5

Os quadrados em branco representam *pixels* que não são de borda, os pintados representam *pixels* de borda e os marcados com (X) não são analisados e, portanto, não influenciam no resultado.

O algoritmo executa as convoluções até que não haja mais mudanças na configuração de *pixels* de borda. Desta forma obtemos uma imagem com bordas de 1 (um) *pixel* de largura, embora ainda possam restar “parasitas” [Gonzales--Woods (1993)].

8. Resultados Experimentais

Os experimentos foram realizados utilizando ferramentas desenvolvidas e inseridas no aplicativo Khoros, que é um pacote aberto (permite a inserção de ferramentas próprias) desenvolvido na Universidade do Novo México (EUA).

Inicialmente, desenvolvemos um conversor para que pudessemos utilizar as imagens *perceptron*¹ encontradas pela internet através da indicação de [Hoover et al. (1996)], para o formato padrão do *Khoros*. Este conversor lê o arquivo contendo a imagem de profundidade e, com o auxílio dos parâmetros de calibração fornecidos, calcula as imagens com as coordenadas 3-D.

Após esta fase inicial de obtenção de imagens, iniciamos o desenvolvimento das diversas ferramentas necessárias à implementação deste trabalho. A primeira ferramenta serviu para o cálculo do gradiente de profundidade, que recebe uma imagem de profundidade e

¹ (<http://marathon.csee.usf.edu/seg-comp/SegComp.html>)

calcula o gradiente para cada ponto. A segunda ferramenta foi o detetor de borda, que recebe como entrada uma imagem de gradiente e, baseado num limiar, calculado automática ou manualmente, marca os pontos como sendo ou não de borda. Além disso, este detetor também implementa uma opção para afinamento de bordas, conforme descrito no artigo.

Em seguida desenvolvemos uma ferramenta que calcula os coeficientes da normal a cada ponto da imagem, utilizando as imagens de coordenadas (x,y,z) . O usuário pode escolher o tamanho da máscara a ser utilizada para o cálculo, bem como o número de vetores a serem calculados (4 ou 8), de acordo com a vizinhança definida.

Finalmente, implementamos o cálculo do gradiente de orientação, que recebe as imagens dos coeficientes das normais e, opcionalmente, a imagem de borda de profundidade (para evitar detecção repetida), para calcular o gradiente de cada ponto.

Utilizamos nos experimentos imagens reais e imagens sintéticas com objetos de formas diferentes. As imagens sintéticas apresentaram excelentes resultados na aplicação das ferramentas de detecção de bordas. Com imagens reais não conseguimos obter resultados satisfatórios na detecção de bordas de orientação, pois as imagens reais são muito ruidosas, fazendo com que o cálculo da normal não seja preciso e ocasionando uma má detecção errônea.

A detecção de bordas de profundidade e o afinamento mostraram-se satisfatórios tanto em imagens reais quanto em imagens sintéticas, embora o afinamento deixe alguns “parasitas”.

A Figura 3(a) mostra a imagem de profundidade de um paralelepípedo da qual detetamos as bordas de profundidade e orientação (Figuras 3(b), (c)) e, após somarmos as duas imagens de borda, fizemos o afinamento (Figura 3(d)). Observamos que os resultados foram muito bons, pois as bordas foram detetadas e afinadas corretamente.

Com Figura 4(a), do objeto cilíndrico, fizemos os mesmos procedimentos: detecção de bordas de profundidade e orientação (Figuras 4(b), (c)) além de soma e afinamento das bordas (Figura 4(d)). Nesta imagem os resultados já não foram tão bons quanto na anterior. Podemos observar que alguns trechos de borda de profundidade não foram detetados e, por isso, foram classificados como sendo de orientação. Também notamos que o afinamento deixou alguns ruídos ou “parasitas” presos às bordas.

A Figura 4(a) com vários objetos poliédricos foi a que apresentou os maiores problemas. Embora a detecção e afinamento (Figuras 4(b),(d)) das bordas de profundidade tenham apresentado resultados satisfatórios, não obtivemos êxito na detecção de bordas de orientação, devido à sensibilidade ao ruído.

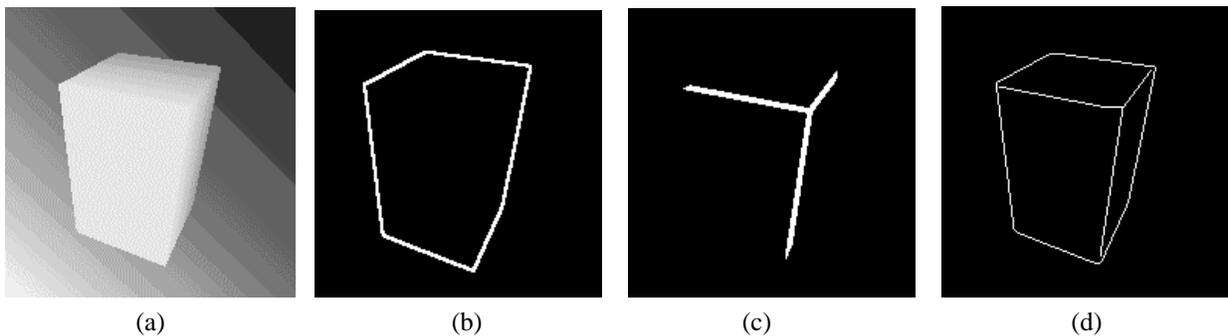


Figura 3. Detecção de bordas na imagem de um paralelepípedo: (a) imagem de profundidade sintética; (b) bordas de profundidade; (c) bordas de orientação; (d) bordas de profundidade e orientação somadas e afinadas.

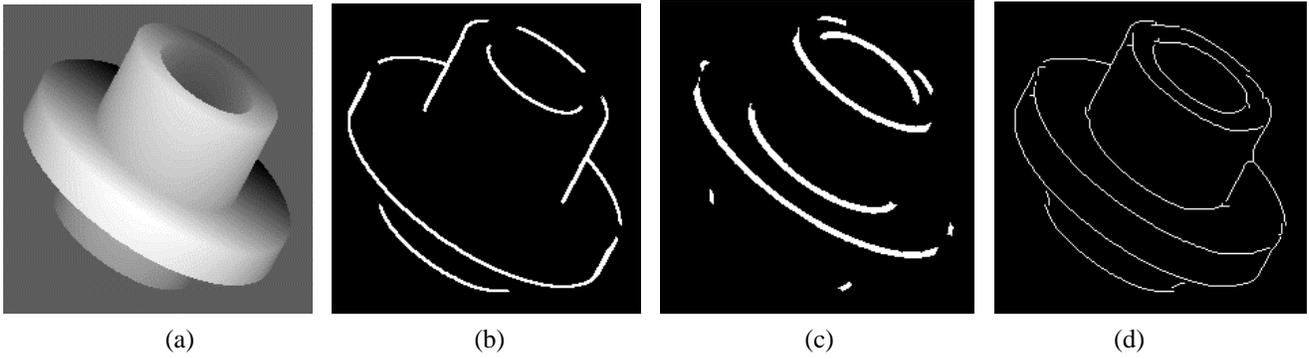


Figura 4. Detecção de bordas numa imagem cilíndrica: (a) imagem de profundidade sintética; (b) bordas de profundidade; (c) bordas de orientação; (d) bordas de profundidade e orientação somadas e afinadas.

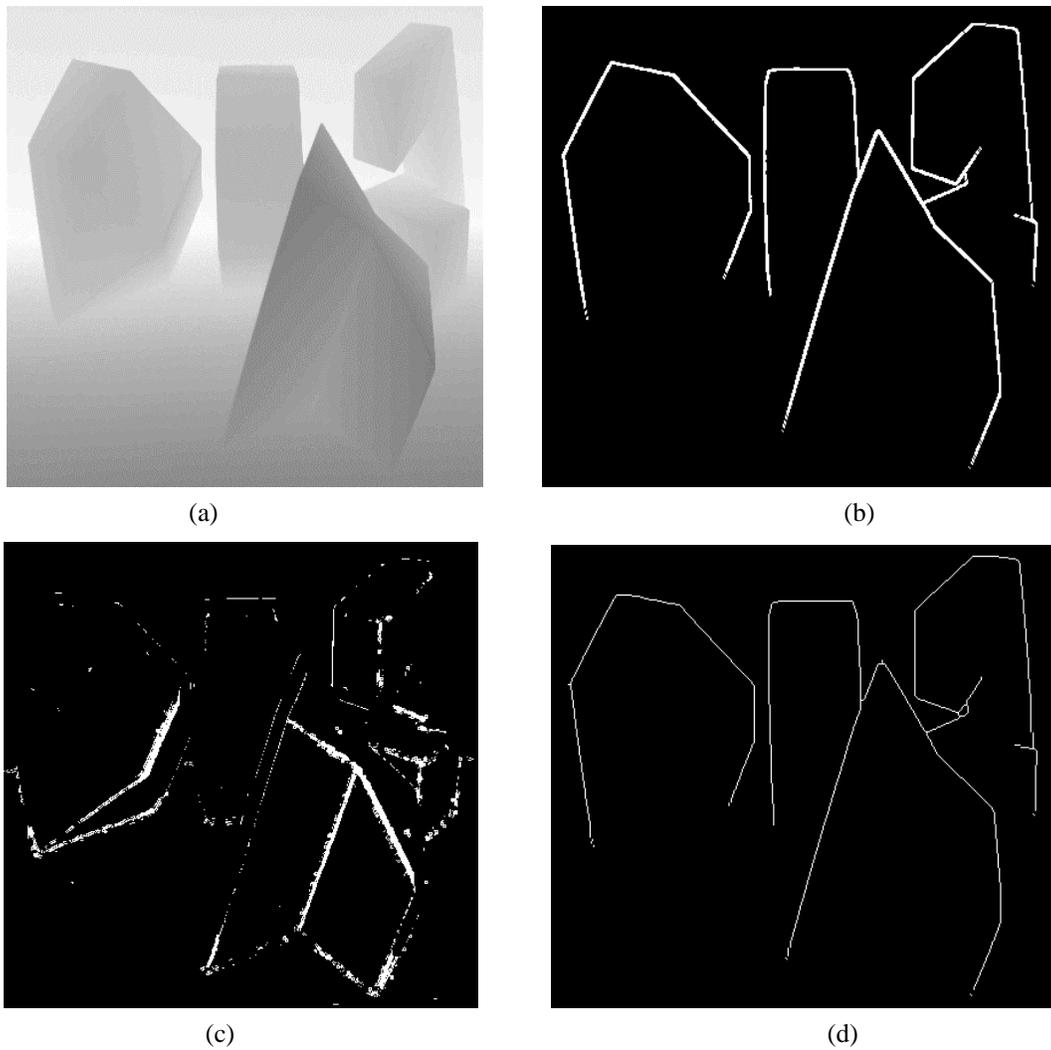


Figura 4. Detecção de bordas em um imagem com objetos poliédricos: (a) imagem de profundidade real; (b) bordas de profundidade; (c) bordas de orientação; (d) bordas de profundidade e orientação somadas e afinadas.

9. Conclusão

Neste trabalho apresentamos um estudo sobre detecção de bordas de profundidade e orientação em imagens de profundidade, e seu afinamento baseado em Morfologia Matemática.

As técnicas implementadas tem mostrado resultados satisfatórios em imagens com pouco ruído, e, embora a detecção de bordas de profundidade e o afinamento tem se mostrado bastante robustos (insensíveis a ruído), a detecção de bordas de orientação é muito sensível ao ruído.

Como continuação deste trabalho pretendemos tornar estas técnicas mais robustas, principalmente a detecção de bordas de orientação. O ponto a ser atacado será o cálculo dos coeficientes da normal a cada *pixel*, que é o fator determinante para a qualidade na detecção destas bordas. Também serão estudadas técnicas para realizar o fechamento de mapas de borda, com o objetivo de avaliar a melhor abordagem a ser aplicada em imagens de profundidade.

As ferramentas implementadas para este trabalho estão disponíveis para os usuários do aplicativo Khoros, dentro de um conjunto de ferramentas para processamento de imagens de profundidade que estamos desenvolvendo. Uma explicação mais detalhada da nossa implementação pode ser encontrada na nossa página na internet².

10. Referências

- O.R.P. Bellon, "Tutorial: Segmentação de imagens de profundidade". *IX SIBGRAPI* (1996), Caxambu-MG.
- P. J. Besl, "Active Optical Range Imaging Sensors", *Advances in Machine Vision*, Springer-Verlag, 1988.
- J. Canny, "A computational approach to edge detection", *IEEE Trans. Pattern Anal. Machine Intel.* 8 (1986), 679-698.
- J. Facon, *Morfologia Matemática: Teoria e Exemplos*, Editora Universitária Champagnat da PUC, 1996.
- J. D. Foley, A. Van Dam, S. K. Feiner, J. F. Hughes, *Computer Graphics*, Addison Wesley, 1996.
- R. C. Gonzales, R. E. Woods, "Digital Image Processing", Addison Wesley, 1993.
- J. F. Haddon, "Generalised threshold selection for edge detection", *Pattern Recognition* 3 (1988), 195-203.

R. M. Haralick, L. G. Shapiro, *Computer and Robot Vision*, volumes I,II, Addison-Wesley, 1993.

R. Hoffman, A. K. Jain, "Segmentation and classification of range images", *IEEE Trans. Pattern Anal. Machine Intel.* 9 (1987), 608-620.

A. Hoover, G. Jean-Baptiste, X. Jiang, P. J. Pflynn, H. Bunke, D. B. Goldgof, K. Bowyer, D. W. Eggert, A. Fitzgibbon, R. B. Fisher, "An experimental comparison of range image segmentation algorithms", *IEEE Trans. Pattern Anal. Machine Intel.* 7 (1996), 673-689.

R. A. Jarvis, "A perspective on range finding techniques for computer vision", *IEEE Trans. Pattern Anal. Machine Intel.* 5(1983), 122-139.

D. Marr, E. Hildreth, "Theory of edge detection", *Proc. of Royal Society of London* 207 (1980), 187-217.

D. Mintz, "Robust Consensus Based Edge Detection", *CVGIP: Image Understanding* 2 (1994), 137-153.

A. Mitiche, J. K. Aggarwal, "Detection of edges using range information", *IEEE Trans. Pattern Anal. Machine Intel.* 5 (1983), 174-178.

D. Nitzan, A. E. Brain, R. O. Duda, "The Measurement and Use of Registered Reflectance and Range Data in Scene Analysis", *Proceedings of the IEEE* 2 (1977), 206-219.

W. K. Pratt, *Digital Image Processing*, Wiley-Interscience, (1991).

² <http://tordo.inf.ufpr.br/~visao/EdgeDetect.html>