

Revista Brasileira de Cartografia (2013) N^o 65/4: 775-783
Sociedade Brasileira de Cartografia, Geodésia, Fotogrametria e Sensoriamento Remoto
ISSN: 1808-0936

IDENTIFICAÇÃO DE SINAIS MANUAIS EM IMAGENS DE PROFUNDIDADE APLICANDO MORFOLOGIA MATEMÁTICA

Hand Signals Recognition in Range Images Applying Mathematical Morphology

Jorge Antonio Silva Centeno¹ & Regina Tiemy Kishi²

Universidade Federal do Paraná - UFPR

¹Departamento de Geomática

²Departamento de Hidráulica e Saneamento

CEP 81531-990 Curitiba/PR, Brasil

{centeno, rtkishi.dhs}@ufpr.br

Recebido em 31 de março, 2013/ Aceito em 09 de junho, 2013

Received on March 31, 2013/ Accepted on June 09, 2013

RESUMO

Neste artigo é descrito um método para o reconhecimento de sinais feitos com a mão utilizando dados de imagens de profundidade. A entrada é uma imagem contendo uma cena na qual o usuário apresenta sua mão mostrando números de um a cinco. Esta imagem é inicialmente segmentada utilizando um método de segmentação global baseado na morfologia matemática. A seguir, a imagem binária resultante é analisada, utilizando morfologia matemática binária para separar os dedos do restante da mão e assim contar quantos dedos o usuário apresenta. Testes com imagens reais são mostrados para comprovar a viabilidade do método.

Palavras-chave: Fotogrametria, Imagens de profundidade; morfologia matemática; processamento de imagens.

ABSTRACT

In this paper we introduce a method for the recognition of hand gestures from range images. The input is an image of a scene where the user displays his hand making a sign of different numbers between one and five. This image is first segmented using a global segmentation method based on mathematical morphology. Then, the resultant binary image is analyzed, applying binary mathematical morphology to separate the fingers from the rest of the hand, allowing counting how many fingers the user displays. Tests with real images are shown to prove the viability of the method.

Keywords: Photogrammetry, Range images; mathematical morphology, image processing.

1. INTRODUÇÃO

A miniaturização de equipamentos eletrônicos, como telefones ou *tablets*, e a necessidade de facilitar a integração de pessoas que apresentam algum tipo de restrição ao uso dos dispositivos convencionais de interação com computadores ou máquinas justificam a pesquisa que trata do reconhecimento de gestos e sinais humanos como forma de estabelecer a comunicação entre o usuário e o computador.

Ibraheem e Kahn (2012) apontam diferentes opções para o reconhecimento de gestos e sinais, como o uso de luvas de dados, muito populares em sistemas de realidade virtual, uso de luvas com marcas ou coloridas com padrões, as quais são fotografadas e as marcas são reconhecidas nas imagens para estimar a posição da mão e dedos, e o uso de métodos baseados em imagens das mãos nuas (*vision based approaches*). Neste último caso, a mão é reconhecida em imagens usando métodos de segmentação e processamento de imagens.

Um dos campos de pesquisa se concentra no reconhecimento de sinais e movimentos das mãos, como maneira de facilitar a interação humano-computador. A literatura apresenta diferentes trabalhos de revisão de publicações e métodos para esta finalidade, como em Harshith et al (2010), Ibraheem e Kahn. (2012) ou Khan e Ibraheem (2012). Inicialmente, deve-se esclarecer que são reconhecidos dois campos de pesquisa: o reconhecimento de sinais em uma imagem e o reconhecimento de gestos numa série de imagens. Este artigo trata do reconhecimento de sinais em uma imagem.

As câmaras de distância ou *range imaging* (Kolb et al., 2009) constituem uma inovação no campo da fotogrametria digital. Embora permitam obter imagens de uma restrita faixa do espectro eletromagnético, as câmaras de distância se destacam por registrar uma grande quantidade de medições de distância utilizando um sensor de quadro, com uma alta taxa de repetição. Estas características as tornam diferentes e vantajosas, em alguns aspectos, em relação a dispositivos como o laser scanner ou a fotogrametria estéreo.

A alta taxa de obtenção de imagens e a disponibilidade de imagens de profundidade, mesmo com baixa resolução, são motivos que fundamentam estudos visando usar estas imagens para capturar

movimentos ou sinais do usuário e assim facilitar a interação humano-computador, suprimindo, na medida do possível, o uso de teclado ou mouse. O presente artigo descreve uma metodologia de reconhecimento de sinais feitos com as mãos, mostrando os números 1-5, utilizando imagens de profundidade e operações de morfologia matemática.

2. Revisão de literatura

Para Garg et AL (2009), os métodos de reconhecimento de gestos e sinais podem ser agrupados em métodos baseados na análise temporal e métodos de análise espacial. O reconhecimento de sinais, por não incluir movimentos, se restringe ao escopo espacial, porém também podem ser o passo inicial para a modelagem temporal. A modelagem espacial pode ser abordada no domínio tridimensional ou bidimensional.

De acordo a Kahn e Ibraheem (2012) a Modelagem espacial 2D pode ser baseada na análise da forma, do movimento ou na aplicação de modelos deformáveis (*Deformable templates*). A modelagem de forma, por sua vez, pode ser classificada em modelos geométricos ou modelos não-geométricos. Os modelos geométricos se concentram na localização e posição dos dedos, bem como as características de palma. Já os modelos não-geométricos se baseiam em propriedades como a cor da pele ou a análise da distribuição espacial dos pixels que representam a mão, através do estudo de seus contornos ou a parametrização da forma usando descritores como orientação ou suas componentes principais.

Exemplos de trabalhos a respeito do reconhecimento de sinais em imagens podem ser encontrados na revisão de literatura de Chaudhar et al. (2011) ou Zabulis et al. (2009).

Segundo Harshith (2010), o uso de câmaras de distância (TOF - Time of flight) é uma opção vantajosa, pois permite simplificar a segmentação usando a informação de profundidade e facilita o reconhecimento de certos sinais que não são passíveis de interpretação em duas dimensões apenas.

As câmaras TOF, assim como a varredura a laser (laser scanner) ou a fotogrametria estéreo, permitem obter uma nuvem de pontos para representar a mão no espaço. Os dados tridimensionais podem ser usados para a interpretação de sinais do usuário.

2.1 Imagens TOF

As câmaras de distância capturam dados dos pontos em uma superfície usando o princípio do tempo de propagação (time of flight - TOF). A câmera usa uma fonte artificial para emitir radiação eletromagnética na direção do objeto. Parte desta radiação é refletida de volta ao sensor (retro-espalhada). Quando a radiação eletromagnética refletida é suficiente para ativar o sensor é possível calcular a distância entre a câmara e o objeto.

Detalhes do funcionamento deste tipo de equipamento podem ser encontrados em Lange (2000). A seguir é feita uma breve introdução.

A radiação emitida se propaga com a velocidade da luz c . Dependendo da distância que separa a câmara da superfície, a radiação eletromagnética registrada pelo sensor sofre um atraso dt em relação à radiação emitida. Este atraso permite calcular a distância entre a câmara e a superfície para cada ponto usando a equação 1.

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

O atraso do sinal de retorno é medido com base na diferença de fase (ϕ) do sinal refletido em relação ao sinal emitido pela fonte. Como Baltsavias (1999) explica, e de acordo com figura 1, sendo P o período da onda emitida, dt o atraso e ϕ a diferença de fase:

$$\begin{aligned} P &\cong 2 * \pi \\ dt &\cong \phi \end{aligned}$$

então:

$$dt = \frac{(\phi * P)}{2 * \pi} + n * P \quad (2)$$

Onde n representa o número de períodos completos adicionais que são negligenciados, pois não podem ser determinados.

$$dt = \frac{(\phi * P)}{2 * \pi} \quad (3)$$

ou em termos de frequência de modulação (f_m) da onda emitida:

$$dt = \frac{\phi}{(2 * \pi * f_m)} \quad (4)$$

Substituindo a equação 4 na equação 1 obtém-se uma expressão que permite calcular a distância R usando a frequência de modulação e a diferença de fase (Dorrington et al., 2009).

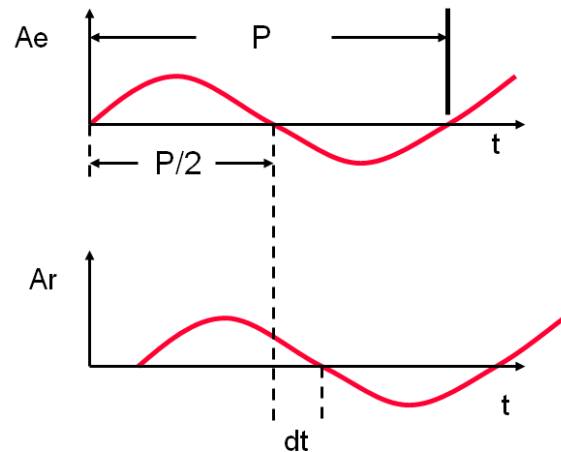


Fig. 1 - Cálculo do atraso da onda em função da diferença de fase.

$$R = \frac{(c * \phi)}{(4 * \pi * f_m)} \quad (5)$$

O sensor da câmara PMD utiliza a tecnologia CMOS para a registrar a radiação refletida pelo objeto e se baseia num novo desenvolvimento, o conceito de *smart pixel* (Lange, 2000). Por usar uma CMOS, este equipamento é capaz de medir um grande número de pontos a cada instante, uma medição de distância para cada pixel. Os dados registrados em cada pixel representam a diferença de fase (e o atraso) de diferentes pontos na cena. Aplicando a equação 5, é possível calcular a distância entre o sensor e os objetos na cena, seguindo um raio que passa pelo centro de projeção da câmara. Os dados são fornecidos com três coordenadas (linha, coluna, Range).

2.1 Morfologia matemática

A Morfologia Matemática surgiu nos anos 60 na França e pode ser descrita como um conjunto de operadores lógicos entre imagens capazes de modificar as formas presentes na imagem (Morfologia) utilizando operadores matemáticos. Conceitos Básicos de morfologia matemática podem ser encontrados na literatura de processamento de imagens, como Serra (1982), Haralick e Shapiro (2002) ou Sonka (1999) e não serão descritos aqui em detalhe.

Foram utilizados os operadores de erosão e dilatação em imagem binária utilizando um elemento estruturante. O elemento estruturante pode ser uma segunda imagem, pequena, também binária.

A erosão é formalmente descrita pela equação:

$$A \ominus B = \{ x \mid x + b \in A \text{ para todo } b \in B \} \quad (6)$$

A dilatação é descrita na equação (7).

$$A \oplus B = \{ c \mid c = a + b, a \in A, b \in B \} \quad (7)$$

Estes operadores não são complementares, logo a combinação destes dois operadores origina novos operadores, como por exemplo, o fechamento e a abertura. A abertura suaviza o contorno dos objetos na imagem, elimina partes estreitas (em comparação ao elemento estruturante) como também proeminências estreitas. Por estas propriedades, ela é muito usada para remover ruídos, pequenos conjuntos de pixels isolados, em imagens. A abertura de um conjunto A por elemento estruturante B e definida como mostra a equação 8:

$$A \circ B = (A \ominus B) \oplus B \quad (8)$$

O fechamento elimina pequenos espaços vazios dentro dos elementos presentes na imagem, preenchendo também pequenas concavidades no contorno das figuras. Este operador é composto por uma dilatação seguida de uma erosão com o mesmo elemento estruturante e pode ser formalmente escrito como mostra a equação 9:

$$A \bullet B = (A \oplus B) \ominus B \quad (9)$$

3. Métodos

A finalidade deste estudo é reconhecer sinais feitos com a mão na frente da câmara para facilitar a interação humano-computador. O método aqui usado segue, de forma geral, os passos apontados por Harshith (2010):

1. Segmentação da mão usando informação de distância;
2. Determinação da área de análise;
3. Extração da mão;
4. Interpretação do sinal.

O dado de entrada no processo é uma imagem de distância, ou seja, uma matriz cujos pixels (linha, coluna) armazenam a distância à superfície dos objetos, incluindo uma mão.

3.1 Cálculo de coordenadas tridimensionais

Os dados da câmara estão num sistema polar, com as componentes descritas em termos de linha/coluna e range (distância). Das coordenadas linha/coluna, e considerando o centro de projeção obtido na etapa de calibração, são calculados dois ângulos

que, junto com a medida de distância, permitem projetar os pontos no espaço 3D desde a posição da câmara. A transformação das coordenadas radiais em ortogonais (X,Y,Z) pode ser feita através da equações(10-12).

$$X = x * Z/f \quad (10)$$

$$Y = y * Z/f \quad (11)$$

$$Z = R * \cos(\text{atan}(rf)) \quad (12)$$

Onde:

R- Distância radial (m);

f- Distância focal da câmara (mm);

x,y- Coordenadas dos pontos no sistema de imagem (mm).

Para calcular as coordenadas (x,y) dos pontos no plano da imagem são necessárias as coordenadas do Ponto Principal (x_0, y_0) no referencial da imagem e o tamanho efetivo do pixel nas direções horizontal e vertical.

3.1. Segmentação

O seguinte passo consiste na segmentação da mão nos dados tridimensionais. Para isto é assumido que o usuário coloca a mão na frente da câmara mostrando o sinal desejado. Logo, a mão se encontra mais próxima da câmara do que os outros objetos, ou seja, está associada à coordenadas Z menores (eq. 10-12).

Os métodos de segmentação global assumem que os valores dos pixels que pertencem a um objeto são similares e diferentes dos valores de pixels de outros objetos. É assumido, ainda, que os pixels de cada objeto são vizinhos, ou seja, existe conectividade nas regiões. Na aplicação aqui apresentada os pixels do objeto a ser detectado são conexos.

A tarefa, neste caso, é detectar o limite entre classes (Objeto e Fundo). Na aplicação aqui descrita, é desejado separar pontos perto de pontos distantes da câmara e o limiar pode ser descrito por um único valor de distância. Valores superiores a este limite pertenceriam ao fundo e os menores à mão (objeto).

A variação da distância armazenada em cada pixel da imagem é analisada com ajuda do histograma. Considerando apenas duas classes, é esperado que o histograma apresentasse duas grandes concentrações de valores, distâncias altas (o fundo) e baixas (o objeto).

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

O método aqui proposto é aplicar a morfologia matemática unidimensional ao histograma para detectar seus vales. Inicialmente, o histograma da imagem é calculado. Para isto é necessário quantizar os valores das distâncias armazenadas nos pixels. Neste exemplo, a quantização foi feita em 8 bits (256 valores).

Por ser unidimensional, a aplicação da morfologia matemática difere parcialmente da comumente aplicada no processamento de imagens (bidimensional), porém se baseia nos mesmos princípios.

Sendo $H(R)$ o histograma da distância, o resultado da aplicação do fechamento nesta função resulta em uma nova função $G(R)$ que é sempre maior ou igual que $H(R)$.

$$G = H \bullet E \quad (13)$$

O elemento estruturante, neste caso, é unidimensional, sendo aqui usado um elemento 1×7 ($E = [1 \ 1 \ 1 \ x \ 1 \ 1 \ 1]$).

Uma das propriedades do fechamento é eliminar concavidades na função original. Por isso, as maiores diferenças entre o resultado do fechamento e o histograma original se localizam nos vales do histograma original.

Como a área de interesse é a mais próxima da câmara, o primeiro vale do histograma corresponde ao valor do limiar a ser aplicado e ele coincide com o primeiro pico da diferença entre o resultado do fechamento e o histograma original (equação 14). Por isto, é necessário identificar o primeiro máximo local, o qual é assumido como sendo o valor ideal para separar os pixels da mão do fundo, o limiar L (eq. 15). Um exemplo é mostrado na figura 3.

$$D(R) = G(R) - H(R) \quad (14)$$

$$L = \text{máx}(D(R)) \quad (15)$$

Finalmente, este limiar é aplicado para separar os objetos próximos ao sensor, ou seja a mão sina-

lizando para a câmara, que serve de entrada ao seguinte passo, a identificação do sinal.

3.2. Identificação de sinais

Após a segmentação da mão em função da distância, o seguinte passo consiste em ler, automaticamente, o gesto mostrado com a mão. Para isto, ferramentas de morfologia matemática foram utilizadas.

A imagem de entrada, nesta fase, é uma imagem binária que contém a área segmentada como mão e o fundo. A abordagem se baseia na hipótese de que o objeto visível em questão, a mão, é composto por um elemento maior, a palma da mão, e outros elementos mais finos, os dedos, que dela se projetam em diferentes direções e cuja quantidade também é desconhecida, como mostra a figura 2.a. O objetivo é contar quantos destes membros existem. No exemplo ilustrativo da figura 2.b o sinal de dois é mostrado e neste caso apenas dois membros se prolongam significativamente da palma mão.

Para contar o número de elementos que se prolongam da mão é necessário separar os dedos, elementos mais finos e alongados, da mancha da mão. Para esta finalidade utiliza-se novamente a morfologia matemática, desta vez bidimensional e binária. A imagem segmentada (A) é modificada aplicando o operador de abertura com um elemento estruturante circular (B).

$$C = A \circ B \quad (16)$$

O resultado esperado da abertura é uma imagem contendo apenas a palma da mão, o que pode não acontecer em função do tamanho do elemento estruturante, que pode deixar a base dos

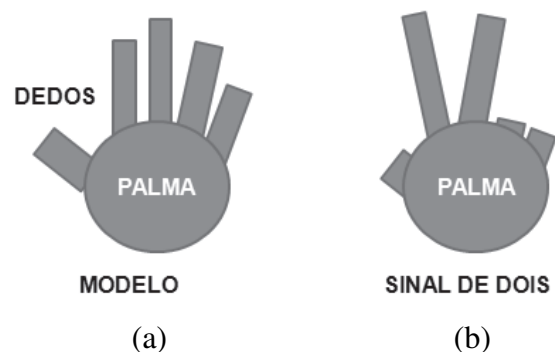


Fig. 2 - Modelo de mão a) mão completa; b) mão mostrando o número 2.

dedos na imagem. A escolha do tamanho do elemento estruturante é crucial nesta fase.

A seguir, a imagem é dilatada, com a finalidade de extrapolar os contornos da palma e garantir que ela esteja totalmente contida na figura. Para isto é usado um elemento estruturante (E2) 50% maior que o primeiro.

$$D = C \text{ \AA } E2 \quad (17)$$

É calculada a diferença entre a imagem segmentada (original) e esta última (C), o que origina uma imagem contendo apenas a ponta dos dedos.

$$E = A - D \quad (18)$$

A presença de ruído na imagem, pixels isolados resultantes da segmentação é eliminada aplicando uma última operação de abertura. Finalmente, o operador de componentes conexos é usado para contar quantos elementos se encontram presentes na imagem, o que fornece o número de dedos mostrados na mão. Este operador numera as regiões isoladas na imagem.

Os operadores aqui empregados identificam o objeto de interesse em função de sua proximidade e separam suas partes com operadores morfológicos. Estes operadores não dependem da rotação da mão, pois a erosão e posterior dilatação eliminam os dedos da imagem em qualquer direção. Esta vantagem, porém, pode se tornar uma limitação quando a identificação do sinal depende da orientação do dedo. Por exemplo no caso de diferenciar o sinal positivo (polegar para cima) de negativo (polegar para baixo). No entanto, a rotação não é um fator limitante na contagem de dedos.

4. 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. Esta imagem foi obtida com a câmara PMD Camcube3, que tem um quadro de 200x200 pixels e alcance entre 0,7 e 7,5 metros.

A figura 4.a mostra o histograma da imagem, em azul. Após o fechamento deste histograma, os vales são preenchidos, como é mostrado na figura 4.a. O efeito do fechamento é a dilatação da função. A diferença entre estas duas funções (Eq. 14) é



Fig. 3 - Imagem de profundidade com usuário mostrando quatro dedos, número 4.

representada na figura 4.b. Nota-se que as maiores diferenças ocorrem nos vales do histograma. O primeiro pico detectado nesta função diferença é mostrado em preto e é assumido como limiar. A adoção deste limiar permite segmentar a imagem em duas classes, elementos próximos (mão) e outros (fundo), como mostra a figura 5.a.

Na figura 5.b pode ser visto o resultado da aplicação da erosão na imagem 5.a. Os dedos, por terem dimensões inferiores ao elemento estruturante, um disco de 10 pixels de raio, são removidos, ficando apenas a área mais larga, correspondente à palma da mão. Esta única região é então dilatada, usando um elemento maior, de 15 pixels, para obter uma aproximação grosseira da palma da mão e eventualmente parte do braço. O resultado intermediário é mostrado na figura 5.c.

A diferença entre a primeira imagem 5.a e o resultado da dilatação (figura 5.c) permite obter a ponta dos dedos em uma imagem binária. Porém, os pixels isolados visíveis na imagem 5.a também aparecem e devem ser removidos. Para isto, é aplicada novamente a operação de abertura, mas desta vez usando um elemento estruturante menor, um disco de raio igual a três pixels, para não prejudicar as pontas dos dedos. O resultado final é mostrado na figura 5.d.

Finalmente, o operador de componentes conexos é aplicado para contar quantas regiões isoladas aparecem na imagem, no exemplo quatro.

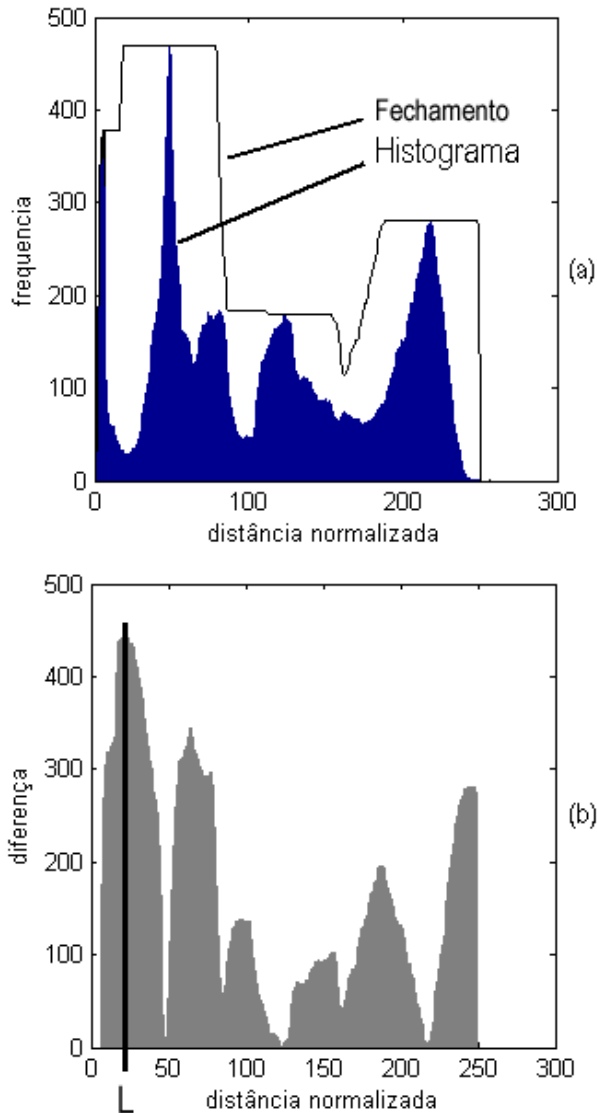


Fig. 4 - Limiarização. a) histograma original e o resultado da operação de fechamento. b) diferença entre o resultado do fechamento e o histograma original.

Este algoritmo foi codificado no ambiente *Matlab*. O programa lê a imagem e como resultado fornece o total de áreas identificadas como dedos, como um número entre um e cinco.

O mesmo experimento foi repetido para outras imagens mostrando outros sinais, os números 1,2,3,4 e 5. O resultado final pode ser visto na figura 6. Para efeito de ilustração, na figura são enfatizadas as áreas detectadas como dedos. A contagem destas áreas pelo método dos componentes conexos corresponde a 1 para a figura 6.a, 2 para a 6.b, 3 para a 6.c, quatro para a 6.d e cinco para a 6.e.

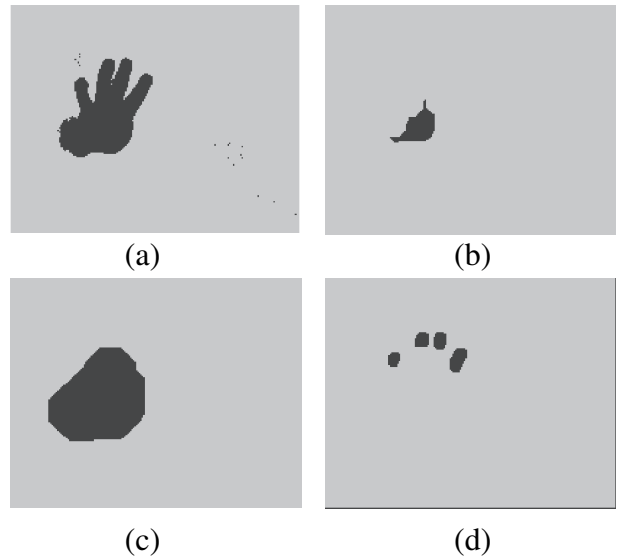


Fig. 5 - Processo de detecção de dedos. a) imagem segmentada; b) resultado da erosão. c) palma da mão dilatada; d) Diferença entre a imagem segmentada e a da palma da mão.

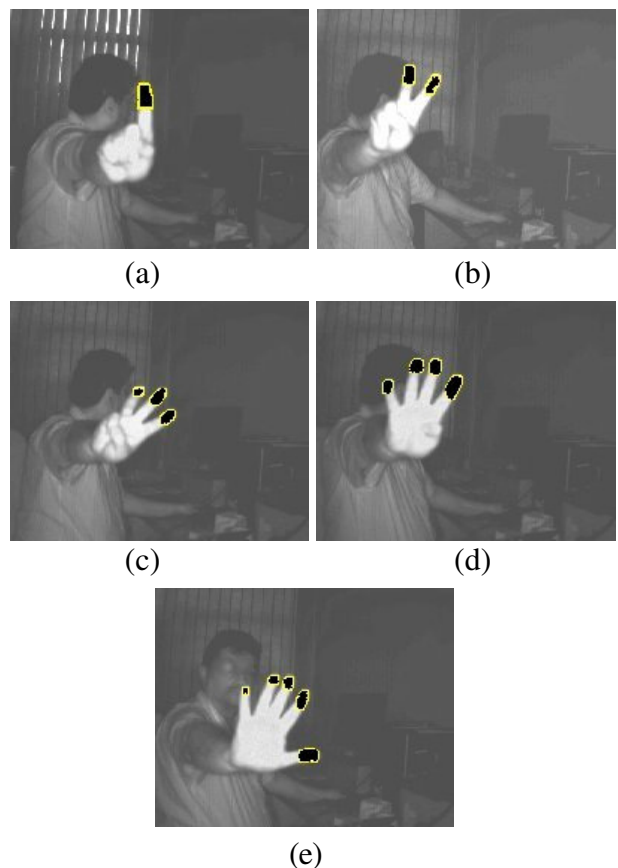


Fig. 6 - Resultado do reconhecimento de mão mostrando 1,2,3,4 e 5 dedos, superposto à imagem de intensidade (P/B).

5. Conclusões

Foi apresentado um novo método para identificar sinais feitos com a mão na frente de uma câmara de distância. O método proposto é capaz de identificar números de 1 a 5, e é baseado na morfologia matemática. O método depende da qualidade da imagem e do tamanho do objeto dentro do quadro da imagem, além da escolha dos operadores morfológicos. Vale a pena ressaltar que os exemplos mostrados utilizam valores que servem para curta distância e que tais valores devem ser modificados quando a mão se encontra distante.

Os experimentos comprovam a viabilidade desta metodologia como ferramenta para facilitar a interação humano-computador e apontam para outras aplicações. Por exemplo, pode se usar este método para detectar a ponta dos dedos em cada quadro de uma série de imagens, um vídeo, e depois seguir estas regiões numa série de imagens, para detectar sinais em movimento. O método oferece uma alternativa viável para a identificação de sinais manuais que pode contornar o problema de falta de contraste entre a mão e o fundo quando a análise é feita apenas considerando valores associados às características espectrais (cor) dos objetos na cena, visto que podem existir objetos com a mesma cor que a mão. Além disso, o método pode ser usado na ausência de luz, pois usa um sensor ativo.

Agradecimentos

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

Referências Bibliográficas

CHAUDHARY, A.; RAHEJA, J. L.; DAS, K. e RAHEJA, S.. **Intelligent approaches to interact with machines using hand gesture recognition in natural way: A Survey**. International Journal of Computer Science & Engineering Survey (IJCSSES) Vol.2, No.1, Feb.

DORRINGTON, A.A., KELLY, C.D.B., SHANE H. MCCLURE, S.H., PAYNE, A.D. e 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.

GARG, P.; AGGARWAL, N. e SOFAT, S.. **Vision Based Hand Gesture Recognition**. World Academy of Science, Engineering and Technology. vol. 49, pp. 972–977, 2009.

HARALICK, R. M.; e SHAPIRO, L. G. (2002). **Computer and Robot Vision** (Volume I). Addison Wesley Longman.

HARSHITH.C.; SHASTRY, K.R.; RAVINDRAN, M.; SRIKANTH, M.V.V.N. e LAKSHMIKANTH, N. **Survey on various gesture recognition techniques for interfacing machines based on ambient intelligence**. International Journal of Computer Science & Engineering Survey (IJCSSES) Vol.1, No.2, November 2010.

IBRAHEEM, N.A. E KHAN, R.Z.. **Survey on Various Gesture Recognition Technologies and Techniques**. International Journal of Computer Applications (0975 – 8887). Volume 50 – No.7, July 2012.

KHAN, R.Z. e IBRAHEEM, N.A.. **Comparative study of hand gesture recognition system**. Computer Science & Information Technology. ISSN :2231 - 5403 [Online]. Volume 2, Number 3, 2012.

KOLB, A., BARTH, E. 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). University of Siegen, Alemanha, 2000.

SAHOO, P.K., SOLTANI, S., e WONG, A.K.C., **A survey of thresholding techniques**, Comput. Vision, Graphics Image Processing 41 (2) (1988) 233—260.

SONKA, M.;HLAVAC,V. e BOYLE, R.. **Image Processing, Analysis and Machine Vision**, 2nd edition, PWS publishing, 1999.

Serra, J. Image Analysis and Mathematical Morphology. Ac. Press, Volume 1 (1982), Volume 2 (1988).

ZABULIS, H. BALTZAKIS, A. e ARGYROS, A.. **Vision-based Hand Gesture Recognition for**

Human Computer Interaction, Chapter 34, in “The Universal Access Handbook”, Lawrence Erlbaum Associates, Inc. (LEA), Series on “Human Factors and Ergonomics”, ISBN: 978-0-8058-6280-5, pp 34.1 - 34.30, Jun 2009.