

SEGMENTAÇÃO DE ESCRITA CURSIVA POR HISTOGRAMA DE PROJEÇÃO VERTICAL

L. R. VELOSO J. M. CARVALHO

Laboratório de Automação e Processamento de Sinais

Departamento de Engenharia Elétrica – Universidade Federal da Paraíba

58.109-970 - Campina Grande, PB, Brasil Phone: +55 83 3101410 Fax: +55 83 3101418

{veloso,carvalho}@dee.ufpb.br

ABSTRACT

Embora o reconhecimento de palavras seja tema de pesquisas há diversos anos, a segmentação de palavras em caracteres continua sendo um dos principais problemas da área. No presente trabalho é apresentado um algoritmo de segmentação de palavras que propõe um alto número de pontos de segmentação, utilizando informações obtidas através do histograma de projeção vertical. Testes realizados com palavras obtidas de doze escritores distintos comprovam a eficiência do método proposto.

1. INTRODUÇÃO

Palavras são essencialmente formadas por uma seqüência de caracteres, de acordo com as restrições imposta pela língua. Desta forma uma técnica natural de reconhecimento de palavras é a segmentação destas palavras em caracteres, seguida pelo posterior reconhecimento destes caracteres como símbolos válidos do alfabeto, e finalmente, pelo casamento entre as seqüências de caracteres identificados e palavras de um dicionário. A Figura 1 ilustra um exemplo da forma intuitiva de reconhecimento de palavras manuscritas. Embora esta seja a forma mais natural de reconhecimento, a diversidade de estilos e a grande ambigüidade encontrada na escrita cursiva, torna o processo de reconhecimento automático de manuscritos bastante complexo, principalmente devido à dificuldade de segmentar as palavras em letras.



Fig. 1 - Forma intuitiva de reconhecimento de palavras manuscritas: (a) palavra manuscrita; (b) separação dos segmentos; (c) reconhecimento do primeiro segmento; (d) análise do segundo segmento; (e) segmentação do segundo segmento.

Devido aos motivos mencionados no parágrafo anterior, muitos sistemas de reconhecimento de manuscritos utilizam a técnica “holística”, onde a palavra é considerada co-

Expressamos nossos agradecimentos ao CNPq pelo apoio financeiro ao trabalho.

mo uma unidade única, indivisível. Este tipo de sistema, também denominado de global, é empregado, exclusivamente, para o reconhecimento de pequenos vocabulários, como aquele presente em cheques bancários. Quando a aplicação em questão necessita de grandes vocabulários, como em aplicações postais, a identificação individual de caracteres é necessária. Caso contrario, seria necessário realizar uma busca exaustiva em todo o conjunto de possíveis palavras do dicionário, o que tornaria o sistema inviável. Como resultado a grande maioria dos sistemas de leitura automática de manuscritos utiliza a segmentação como uma etapa antecedente ao reconhecimento [1].

A Figura 2 apresenta exemplos de palavras manuscritas. Nestes exemplos, embora as letras apresentem-se ambíguas e muitas vezes ilegíveis, o leitor consegue identificar o sentido da palavra. Isto só é possível devido ao complexo sistema de leitura do seres humanos, que utiliza tanto as informações presentes na escrita como informações contextuais. Na Figura 2 é apresentado um clássico problema: ambas imagens (a) e (c) possuem forma similar de representar letra distintas (*m* e *ui*). Na segunda interpretação as letras devem ser segmentadas, o que não é necessário para o primeiro caso. Por conseguinte, é extremamente difícil a segmentação automática de caracteres em palavras manuscritas.

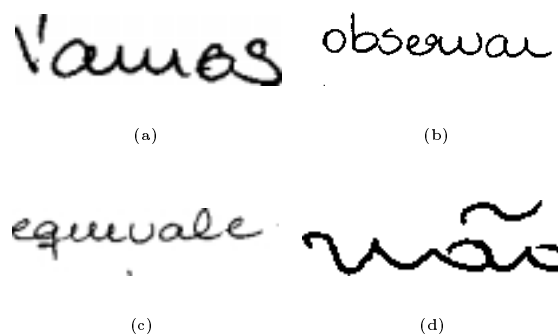


Fig. 2 - Exemplos de palavras manuscritas.

Visando resolver os problemas provocados pela ambigüidade dos caracteres na escrita cursiva, muitos sistemas dividem a imagem em segmentos, ou pedaços da palavra original. Cada segmento idealmente consiste de um caractere ou parte de um caractere. O processo de casamento determina a melhor maneira de concatenar estes segmentos de forma a representar palavras contidas em um vocabulário.

Entretanto, pesquisadores da área de processamento de documentos estão cada vez mais convencidos que a leitura de manuscritos pelo computador é algo mais do que simplesmente isolar e identificar os símbolos individuais que formam as palavras. A informação contextual não é apenas útil, ela é necessária [2].

Neste contexto, este artigo apresenta um algoritmo de segmentação de palavras manuscritas que deliberadamente propõe um alto número inicial de pontos de segmentação. Este algoritmo é baseado em duas hipóteses: 1) existem pontos de segmentação natural e 2) os pontos de segmentação entre caracteres conectados estão localizados nas ligações entre caracteres.

2. SEGMENTAÇÃO NATURAL

Palavras manuscritas podem ser puramente cursivas, escritas em letras de forma ou por uma combinação destes estilos (escrita mista). Na prática a escrita mista é a mais comum de ser encontrada. Mesmo na escrita cursiva, em geral, algumas letras que compõem as palavras não se apresentam conectadas, formando uma pequena descontinuidade na escrita.

Esta característica de descontinuidade na escrita cursiva é aproveitada na primeira etapa do processo de segmentação, onde se determina a existência de pontos de segmentação natural. O procedimento adotado é baseado no sistema de reconhecimento de texto impresso, onde os espaços verticais geralmente são utilizados para separar os caracteres. Uma forma de detectar estes espaços verticais em branco entre letras sucessivas é a análise do perfil de projeção vertical, também chamada de histograma vertical. Este histograma consiste da simples contagem da quantidade de pixels pretos em cada coluna da imagem [3].

No caso de manuscritos, o procedimento utilizado para determinar a existência e a localização dos pontos de segmentação natural utiliza o cálculo do perfil de projeção inclinado (*angled projection profile*), em cinco diferentes ângulos [4]. Durante o pré-processamento a correção da inclinação é utilizada para normalizar as palavras manuscritas, com objetivo de ter palavras com uma inclinação média de zero grau com relação à vertical. Entretanto, uma inclinação média de zero grau não significa que todos os caracteres que compõem a palavra não estejam inclinados. Desta forma, o perfil de projeção inclinado necessita ser calculado. Os ângulos utilizados para o cálculo do perfil de projeção inclinado são -20, -10, 0, 10, 20 graus, com relação à vertical. O perfil de projeção inclinado indica a quantidade de pixels pretos existentes em linhas inclinadas.

O algoritmo proposto para o cálculo do perfil de projeção inclinado, para uma imagem de tamanho $(M \times N)$, é descrito a seguir:

Para cada pixel (i, j) da imagem; $i=1,2,\dots,M$ e $j=1,2,\dots,N$.

Para cada ângulo de inclinação (K)

1. Determine o novo valor da coordenada j na imagem, como sendo:

$$v = j - (M - i) * \tan(K); \quad (1)$$

2. Se o valor do pixel (i, v) for igual a 1, incremente a v -ésima coluna do k -ésimo histograma.

$$Histograma_k(v) = Histograma_k(v) + 1; \quad (2)$$

Para cada ângulo analisado, os pontos de segmentação são detectados utilizando linhas de separação com a mesma inclinação do ângulo analisado. Estas linhas de separação interceptam a linha de base da palavra na coordenada correspondente ao ponto médio em cada seqüência de zeros no histograma analisado.

No final deste processo, tem-se a localização dos pontos de segmentação natural, armazenados em 5 vetores, um para cada ângulo K analisado. Vale salientar, que os cinco vetores que armazenam as localizações dos pontos de segmentação (i, j) armazenam apenas a coordenada j de todos os pontos de segmentação, já que a coordenada i é determinada através do cálculo da linha de base da palavra e esta é a mesma para todo ponto de segmentação. É importante ressaltar que todas as palavras foram normalizadas em inclinação e declive na etapa de pré-processamento.

A figura 3(a)-(f) mostra exemplos de segmentação natural para a palavra *Bem* utilizando linhas de separação com -20, -10, 0, 10 e 20 graus.

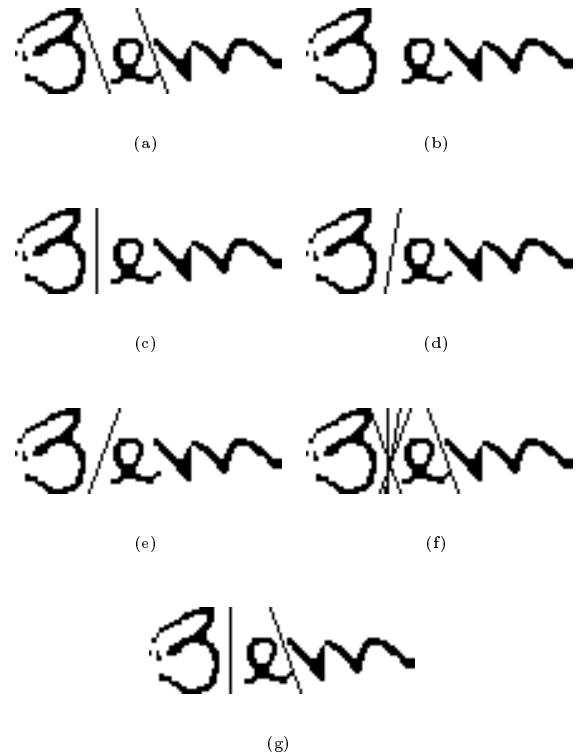


Fig. 3 - Exemplos de segmentação natural nos seguintes ângulos a) -20, b) -10, c) 0, d) 10, e) 20 graus, f) união entre os pontos de segmentação natural g) resultado final do processo de segmentação natural.

3. SEGMENTAÇÃO NAS LIGAÇÕES

Manuscritos são geralmente produzidos pelo movimento do instrumento de escrita (lápis ou caneta) sobre uma superfície, da esquerda para direita, de cima para baixo em uma direção horizontal. Este movimento produz traços horizontais unindo caracteres adjacentes, que são denominados de pontes ou traços de ligações. Conseqüentemente é sobre estes traços que existe a maior probabilidade de serem encontrados os pontos de segmentação nas palavras cursivas.

Nas palavras manuscritas onde os caracteres estão desconectados os pontos de segmentação são encontrados apenas analisando o perfil de projeção inclinado, como foi anteriormente explicado. Entretanto, se a palavra manuscrita é puramente cursiva, este procedimento não faz sentido, pois não existe nenhum ponto no histograma de projeção vertical que tenha valor zero, como exemplificado na Figura 4.

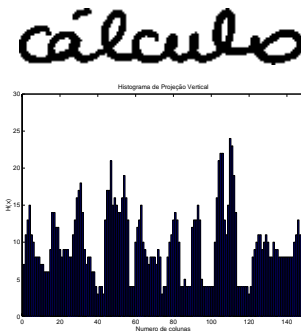


Fig. 4 - histograma de projeção vertical

Uma análise mais detalhada do perfil de projeção, Figura 4, determina que a quantidade de pixels pretos nas colunas onde existem os traços de ligação corresponde à largura média dos traços da palavra, calculada através do algoritmo desenvolvido por Chen [5], [6]. Neste algoritmo o histograma de projeção vertical e a quantidade de transições ao longo das colunas entre pixels pretos e brancos são utilizados para estimar a largura média dos traços que compõem as palavras.

Com base nas constatações feitas anteriormente e no critério de segmentação utilizado (cada caracter pode ser segmentado no máximo em três partes), foi desenvolvido um algoritmo de segmentação baseado no perfil de projeção inclinado e na estimativa da largura média dos traços que compõem a palavra.

A primeira etapa do procedimento de determinação dos pontos de segmentação sobre as ligações é proceder ao fechamento de todas as cavidades centrais da palavra (*fill-hole*). Por cavidades centrais entende-se uma região da imagem composta por pixels brancos e delimitada por pixels pretos. Esta operação é necessária para prevenir que o algoritmo de determinação das ligações detecte pontos de segmentação no meio de letras, tais como: “b”, “o” e “a” entre outras que possuem cavidades centrais.

O perfil de projeção inclinado, em ângulos de -20, -10, 0, 10 e 20 graus com relação à vertical é calculado para a imagem com cavidades preenchidas. A determinação dos traços

de ligação é feita analisando o histograma de projeção inclinado para cada ângulo K , conforme as Equações 1 e 2. Para cada $Histograma_k$, uma projeção menor ou igual a duas vezes a largura média dos traços que compõem a palavra é considerada como pertencente à ligação e assim marcada. O ponto de segmentação corresponde ao ponto médio de cada traço de ligação marcado. A posição dos pontos de segmentação e o valor do ângulo de inclinação K são armazenados para posterior processamento, que realizará a combinação com os pontos de segmentação natural.

3.1 COMBINAÇÃO DOS PONTOS DE SEGMENTAÇÃO

Geralmente, os limites do traçado de uma dada letra não são detectados em todas as segmentações, mas são usualmente detectados em pelo menos uma delas. Desta forma, tomando a união de todas elas, teremos uma melhor segmentação. Por outro lado, a união usualmente pode conter mais de um ponto de segmentação para a mesma ligação entre letras. Estes pontos devem ser identificados e unidos para reduzir o efeito conhecido como super-segmentação (*oversegmentation*) [4]. A Figura 3(f) mostra o resultado da união entre os pontos de segmentação detectados utilizando os cinco ângulos de inclinação. Na Figura 3(g) é mostrado o resultado final do processo de segmentação natural.

A seleção dos pontos de segmentação é realizada através da análise da distância entre eles, segundo uma ordem de prioridade. A mais alta prioridade é dada aos pontos de segmentação natural gerados através da análise do histograma de projeção com inclinação de zero grau; em uma prioridade imediatamente inferior encontram-se os pontos de segmentação posicionados sobre as ligações e que possuem linha de separação com inclinação de zero grau; em seguida, estão os pontos de segmentação natural que possuem linha de inclinação de 10, -10, 20 e -20 graus em relação à vertical, respectivamente, por ultimo, tem-se os pontos de segmentação sobre as ligações com ângulo de inclinação de 10, -10, 20 e -20 graus, nesta ordem de prioridade.

4. RESULTADOS EXPERIMENTAIS

Na realização dos testes com o método de segmentação proposto neste trabalho, foram utilizados amostras de palavras manuscritas, obtidas de doze escritores distintos, digitalizadas por um *scanner* com uma resolução de 200 PPI. Alguns processamentos preliminares foram necessários para as imagens digitalizadas: a normalização da inclinação e do declive da palavra e a suavização.

A Figura 5 mostra alguns exemplos dos resultados obtidos. As linhas de separação indicam os pontos de segmentação. Nestes exemplos, as letras “m”, “n” e “u” são segmentadas em mais de um ponto. Esta é uma característica do sistema que possui como objetivo prevenir problemas devido à ambigüidade da escrita cursiva. Outra característica desta técnica é que nunca são encontrados pontos de segmentação no meio de letras que possuem cavidades centrais. A utilização de linhas de separação inclinadas permite que sejam encontrados pontos de segmentação entre letras que não possuem traços de ligação bem definidos, como nos exemplos (c) e (d) na Figura 5.

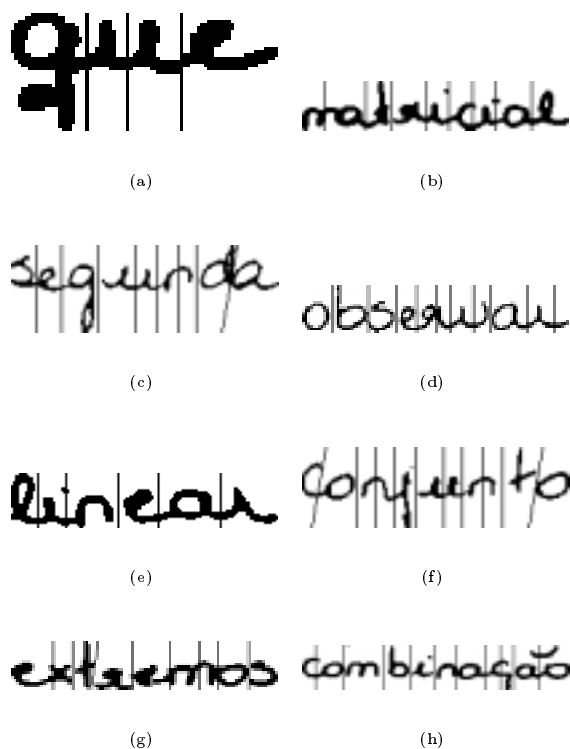


Fig. 5 - Exemplos de segmentação

Estes resultados foram comparados com os resultados obtidos com o algoritmo anteriormente desenvolvido pelos autores [7], em que os pontos de segmentação sobre as ligações são obtidos mediante operações morfológicas de abertura e fechamento, utilizando elementos estruturantes obtidos através de algoritmo genético. Ambos os sistemas possuem performance semelhante, detectando corretamente pontos de segmentação, mesmo em casos em que a palavra é de difícil segmentação. Entretanto o método proposto neste artigo possui uma menor complexidade, não requerendo a realização de nenhuma operação de convolução. Apenas, o cálculo de histogramas de projeção vertical é necessário.

Os resultados obtidos neste trabalho foram semelhantes aos resultados obtidos por Yanikoglu [4]. Naquele sistema a determinação dos pontos de segmentação é realizada através do cálculo de uma função custo sobre todo pixel pertencente à linha base da palavra. O custo em um dado ponto é função de parâmetros de estilo de escrita (largura do traço, ângulo dominante, largura média das letras) e características locais do ponto. Os pesos dos parâmetros da função custo são determinados mediante treinamento com todas as possíveis combinações de letras utilizando programação linear. Estas operações implicam em um método bastante mais complexo do que o proposto neste trabalho.

5. CONCLUSÃO

Este artigo descreve um novo método de segmentação de palavras manuscritas cursivas, que utiliza histogramas de

perfis de projeção inclinados, para determinar pontos de ligação entre caracteres, combinado com pontos de segmentação natural. Este algoritmo é de baixa complexidade, entretanto apresenta desempenho equivalente ao de métodos mais complexos anteriormente desenvolvidos. Experimento para avaliação do desempenho foi realizado com palavras manuscritas reais.

6. REFERÊNCIAS

- [1] El-Yacoubi, A. and Gilloux, M. and Sabourin, R. and Suen, C. Y. . Objective evaluation of the discriminant power of features in an hmm-based word recognition system. In N. Murshed and F. Bortolozzi, editors, *Advances in Document Image Analysis*, pages 60–73. Springer, November 1997.
- [2] Gader, P.D., Mohamed, M. and Chiang, J.-H. Handwritten word recognition with character and inter-character neural networks. *IEEE Transactions on Systems, man, and Cybernetics -Part B*, 27(1):158–164, February 1997.
- [3] Lu, Y. and Shridhar, M. Character Segmentation in Handwritten Words - An Overview. *Pattern Recognition*, 29(1):77–96, 1996.
- [4] Yanikoglu, B. and Sandon, P. A. Segmentation of off-line cursive handwriting using linear programming. *Pattern Recognition*, 31(12):1825–1833, december 1998.
- [5] Chen, M-Y. and Kundu, A. and Zhou, J. Off-line handwritten word recognition using a hidden markov model type stochastic network. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 16(5):481–497, may 1994.
- [6] Chen, M-Y. and Kundu, A. and Srihari, S. N. Variable duration hidden markov model and morphological segmentation for handwritten word recognition. *IEEE Trans. on Image Processing*, 4(12):1675–1687, december 1995.
- [7] Veloso, L.R., Sousa, R. P. and Carvalho, J. M. A new method for cursive word segmentation. In *Proceeding of IEEE International Conference on Image Processing - ICIP'2000*. IEEE Press, Vancouver, Canada, september 2000.