

## RECONHECIMENTO AUTOMÁTICO DE CARACTERES IMPRESSOS

Antonio José Ferreira Machado e Silva

Centro Científico Rio - IBM Brasil

Caixa Postal 4624, 20001, Rio de Janeiro, RJ, Brasil

**RESUMO** - O reconhecimento automático de caracteres tem como um dos objetivos principais possibilitar a conversão de textos impressos em arquivos eletrônicos no formato ASCII. Esta conversão facilita a manipulação, bem como otimiza o armazenamento destes arquivos. Neste trabalho são definidas algumas propriedades que permitem discriminar diversos caracteres, apresentando alguns resultados práticos.

### 1. INTRODUÇÃO

O crescente aumento na capacidade de armazenamento de dados dos discos magnéticos e, principalmente, óticos, bem como na velocidade de processamento dos computadores estimulam a investigação de algoritmos que possibilitem o reconhecimento automático de caracteres escritos. Estes algoritmos permitem a conversão de textos (documentos, livros, revistas, manuscritos, etc.) impressos em papel para arquivos eletrônicos no formato ASCII. Esta conversão se dá em dois passos. No primeiro, o texto é digitalizado gerando uma imagem digital binária, que representa um retrato fiel do original. Esta imagem necessita de grande espaço para armazenamento. No segundo passo, os caracteres são reconhecidos, gerando arquivos no formato ASCII, representando uma grande economia de espaço em disco.

A transferência de textos impressos para arquivos eletrônicos permite maior facilidade para manipulação desses textos. Alteração, inclusão ou exclusão de partes do texto são facilmente executadas nos arquivos eletrônicos. Esse recurso pode ser utilizado para preparar novas edições de livros, alterar a diagramação de textos, modificar campos de um contrato, etc. A geração de arquivos eletrônicos pode visar apenas diminuir o espaço físico necessário para armazenar textos e garantir a melhor conservação destes. Estes são os casos de cartórios de ofícios e notas e de acervos de obras históricas. É claro, que nestes casos o arquivo eletrônico funciona como um complemento ao documento original.

Finalmente, a transferência do arquivo em formato ASCII para texto em papel não representa nenhum problema, e pode ser feita através de impressoras, inclusive a laser, ou de máquinas de escrever acopladas a computadores.

Neste trabalho, o estudo se concentrou em definir propriedades que permitissem separar um caráter do outro. O campo de provas se resumiu a caracteres impressos, sem teste em caracteres manuscritos, englobando os dez algarismos e as vinte e seis letras do alfabeto maiúsculo, no total de trinta e seis caracteres.

Os caracteres podem ser impressos em diversos padrões, onde em cada um destes a imagem representativa de um caráter é diferente. Nas máquinas de escrever, este fato pode ser visto pelos vários tipos de esferas, ou margaridas, que a máquina apresenta. Neste trabalho elegeu-se um padrão de impressão para gerar o dicionário onde as propriedades seriam definidas.

Uma vez definido o tipo padrão do caráter, inicia-se a primeira fase do processo, que consiste em gerar os esqueletos dos caracteres. Em cima dos esqueletos, define-se as propriedades que vão representar a assinatura de cada caráter. As propriedades definidas foram vizinhança, projeção (Maier et al., 1985), simetria e densidade. As propriedades não precisam ser empregadas todas simultaneamente. Quanto maior for a assinatura do caráter maior é o tempo de processamento para o reconhecimento automático.

## **2. DEFINIÇÃO DAS PROPRIEDADES**

As propriedades são aqui definidas de uma forma global. Existem diversas maneiras de empregá-las para o processo de reconhecimento de caracteres. No Tópico 3, referente à metodologia empregada no trabalho, serão apresentadas as maneiras como foram utilizadas estas propriedades.

### **2.1 VIZINHANÇA**

A vizinhança de um caráter é definida através de um vetor de cinco posições onde cada posição representa um tipo de vizinho que um elemento de imagem tem, e uma das posições representa o número de elementos de imagem do caráter. Os tipos de vizinhança testados são: à noroeste, ao norte, à nordeste e à leste (Figura 1). Toda vez que um elemento de imagem do caráter tiver um desses vizinhos, o valor relativo à posição do vetor correspondente a este tipo de vizinho é incrementada de uma unidade.

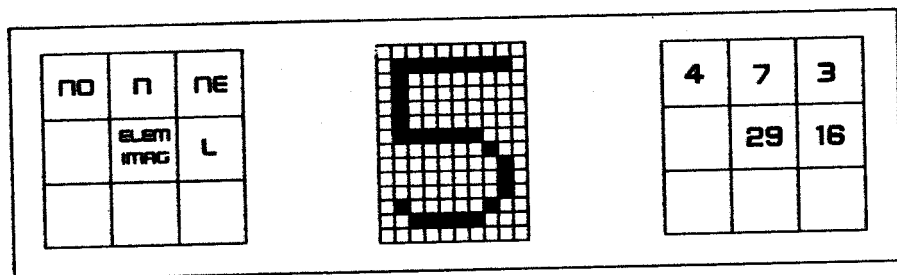


Fig. 1 - Propriedade de vizinhança.

É importante destacar que se um elemento de imagem do caráter tem um vizinho à noroeste, este tem aquele como vizinho à sudeste. O mesmo ocorre para as duplas norte-sul, nordeste-sudoeste e leste-oeste. Esta é a razão para que o vetor de vizinhança tenha apenas cinco posições.

## 2.2 PROJEÇÃO

A projeção de um caráter numa dada direção é dada pelo histograma dos elementos de imagem projetados sobre uma reta normal à citada direção. Por exemplo, a projeção horizontal de um caráter é o vetor representativo do histograma das projeções dos elementos desse caráter sobre uma reta vertical (Figura 2). Neste trabalho, utilizou-se as projeções vertical, horizontal, positiva (sobre uma reta que faz  $135^\circ$  com a horizontal) e negativa (sobre uma reta que faz  $45^\circ$  com a horizontal).

Os vetores de projeção têm dimensões variadas, em função do desenho do caráter. Além disso, as projeções positiva e negativa apresentam mais elementos que a projeção horizontal, que, por sua vez, apresenta mais elementos que a projeção vertical.

Visando equilibrar estas diferenças, e diminuir o tamanho das assinaturas dos caracteres, os vetores representativos das projeções são normalizados para uma dimensão padrão. Esta dimensão tem, no entanto, o compromisso com a manutenção do padrão do caráter. A Figura 2 apresenta uma normalização de um vetor de projeção horizontal de dimensão doze para quatro.

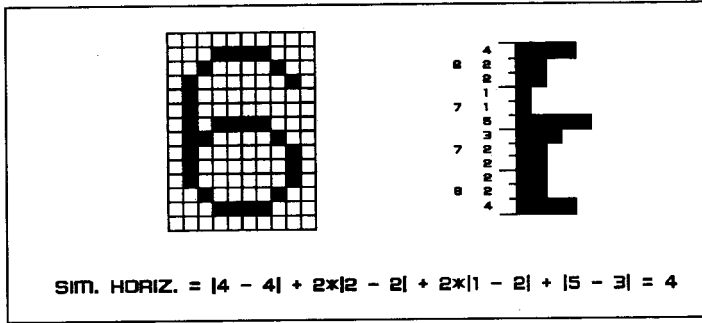


Fig. 2 - Propriedades de projeção e simetria.

### 2.3 SIMETRIA

Esta propriedade sintetiza o grau de simetria que um caráter apresenta em relação a uma reta. As simetrias observadas neste trabalho foram em relação às retas horizontal, vertical, inclinadas a  $45^\circ$  e  $135^\circ$  com a horizontal.

Os valores de simetria são calculados a partir dos vetores primários das projeções (antes da normalização). A simetria vertical é calculada a partir do vetor projeção vertical. A simetria corresponde ao somatório dos módulos das diferenças entre os elementos do vetor projeção equidistantes das extremidades, conforme a Equação 1. Os demais parâmetros de simetria seguem raciocínio análogo. A Figura 2 apresenta o cálculo da simetria horizontal de um caráter.

$$S = \sum_{i=1}^{N/2} |v(i) - v(N+1-i)| \quad (1)$$

É importante destacar que o parâmetro simetria na realidade está medindo a simetria dos vetores de projeção. Nada impede, no entanto, que se meça a própria simetria do caráter, utilizando para isto o retângulo mínimo que envolve o caráter, conforme citado, a seguir, na definição da propriedade de densidade.

### 2.4 DENSIDADE

A propriedade de densidade representa o número de elementos de imagem que o caráter tem por setor. A partir das projeções horizontal e vertical é possível definir o retângulo de dimensões mínimas que envolve o caráter. Este retângulo é uma imagem binária deste caráter. Dividindo este retângulo em setores

iguais e calculando o número de elementos do caráter em cada setor tem-se uma imagem não mais binária, porém de dimensões mais reduzidas. A Figura 3 apresenta a redução de um caráter de dimensões 12x8 para uma matriz de densidade 3x3. Nesta redução, pode ocorrer de um elemento de imagem ser fracionado entre dois ou mais setores.

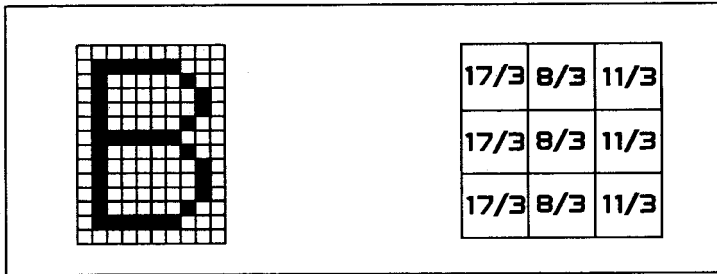


Fig. 3 - Propriedade de densidade.

### 3. METODOLOGIA EMPREGADA

O primeiro passo deste trabalho foi definir o dicionário dos trinta e seis caracteres (esqueletos), que seria o conjunto base para classificar os demais caracteres. Em seguida, foram criados três conjuntos para teste: um com os caracteres ampliados por um fator próximo a 1,25; outro com os caracteres reduzidos (fator 0,75); e um terceiro em que se distorceu, propositadamente, os caracteres. Estes três conjuntos foram confrontados com o dicionário padrão, e para cada elemento destes conjuntos, foram ordenados os caracteres do dicionário (do mais semelhante ao mais diferente).

Os critérios para definir a semelhança (ou dissemelhança) deveriam levar em conta as variações nas dimensões dos caracteres. Desta forma, os parâmetros relativos às diversas propriedades foram normalizados. Para a normalização, obrigou-se que todos os caracteres tivessem o mesmo número de elementos de imagem. Assim sendo, cada posição de um vetor passou a representar valores relativos.

No vetor de vizinhança, a posição que indicava o número de elementos de imagem perdeu o sentido, já que teria um valor fixo. Neste caso, o valor estipulado foi 100. Este vetor teve sua dimensão reduzida para quatro.

Os vetores de projeção foram normalizados para dimensão quatro e, em seguida, cada valor foi multiplicado pelo razão R, entre o valor fixado para o total de elementos de imagem e o número real destes elementos. A soma dos

elementos destes vetores passou a ser constante e a dimensão conjunta destes vetores é dezesseis (quatro vezes quatro).

O vetor de simetria representa a simetria dos vetores primários de projeção (antes da redução da dimensão), porém normalizados pela razão R. A dimensão deste vetor é quatro.

A matriz de densidade após a normalização passou a representar a percentagem de elementos de imagem em cada setor. Utilizou-se matrizes 3x3 o que levou a dimensão desta propriedade a nove. A soma dos elementos da matriz de densidade é constante, e igual a soma dos elementos de cada vetor de projeção.

Para definir as assinaturas dos caracteres utilizou-se a matriz de correlação dos trinta e seis caracteres do dicionário, para cada propriedade. A propriedade que apresentou maior poder de discriminação foi a da vizinhança (foi considerada nesta análise cada vetor de projeção de forma independente). A partir deste resultado, definiu-se três tipos de assinatura: vizinhança com os vetores de projeção (dimensão da assinatura = 20); vizinhança com matriz de densidade (dimensão da assinatura = 13) e vizinhança com vetor de simetria (dimensão da assinatura = 8).

Uma vez definido o tipo de assinatura, confrontou-se cada caráter dos conjuntos de teste com todos os caracteres do dicionário padrão, buscando ordenar estes últimos em função do grau de semelhança.

#### 4. RESULTADOS

Em relação à assinatura vizinhança + projeção o conjunto de caracteres reduzidos mostrou bom desempenho, com apenas um caráter não tendo sido reconhecido corretamente. O algarismo 5 foi reconhecido como 6, apresentando um fator de dissemelhança = 68. Em relação ao algarismo 5 do dicionário padrão, o algarismo 5 reduzido apresentou fator de dissemelhança = 70. Observando a matriz de correlação dos caracteres padrões, em função das propriedades desta assinatura, notou-se que o ordenamento dos caracteres frente ao grau de semelhança no processo de reconhecimento dos caracteres seguiu, em grande parte, o ordenamento do grau de correlação.

O conjunto de caracteres ampliados mostrou resultados bem mais satisfatórios. O reconhecimento foi total, e a matriz de correlação foi praticamente respeitada no ordenamento resultante do reconhecimento.

O conjunto de caracteres deformados merece uma análise subjetiva. Quanto mais deformado parecia o caráter, mais difícil era o reconhecimento. Os resultados encontrados, no entanto, pareceram satisfatórios.

Com a assinatura composta pelas propriedades de vizinhança e densidade os resultados não foram tão bons. O conjunto de caracteres ampliados foi integralmente reconhecido, mas o conjunto de caracteres reduzidos apresentou nove caracteres reconhecidos erroneamente (3, 5, 6, 9, B, G, H, Q e S). Nestes casos, a diferença entre os graus de dissemelhança entre o caráter reconhecido e o correto foi inferior a 7. O conjunto de caracteres deformados teve menos elementos identificados corretamente do que com a assinatura vizinhança + projeção.

A última assinatura testada foi a vizinhança + simetria, que já iniciou com um problema: não fazia distinção entre o algarismo 6 e o algarismo 9. Afora este problema, os resultados ainda foram piores do que com as duas outras assinaturas. Novamente não houve problemas com o conjunto de caracteres ampliados. No entanto, dez caracteres reduzidos não foram identificados corretamente (5, 6, 9, B, G, H, Q, R, S e Z), com a diferença do grau de dissemelhança entre o caráter reconhecido e o correto não excedendo a 10. O conjunto de caracteres deformados teve vários elementos identificados erroneamente, pois perdiam suas características de simetria.

A Figura 4 apresenta a letra R (erre) conforme consta no dicionário padrão e nos conjuntos de caracteres reduzidos e ampliados. Esta letra é um exemplo de um caráter que foi reconhecido pelas três assinaturas quando ampliado, porém, que não teve o mesmo desempenho quando reduzido. Neste caso, foi facilmente identificado pela assinatura vizinhança + projeção; com dificuldades pela assinatura vizinhança + densidade; e não foi identificado pela assinatura vizinhança + simetria.

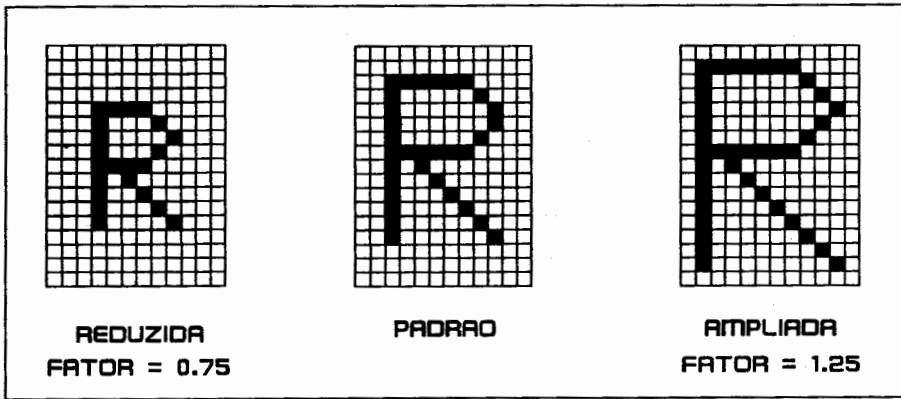


Fig. 4 - Letra maiúscula R reduzida, padrão e ampliada.

A Figura 5 apresenta o algarismo 0 (zero) conforme consta no dicionário padrão e no conjunto de caracteres deformados. Corretamente identificado pelas assinaturas vizinhança + projeção e vizinhança + densidade, o caráter padrão 0 apareceu apenas em 14º lugar quando reconhecido pela assinatura vizinhança + simetria.

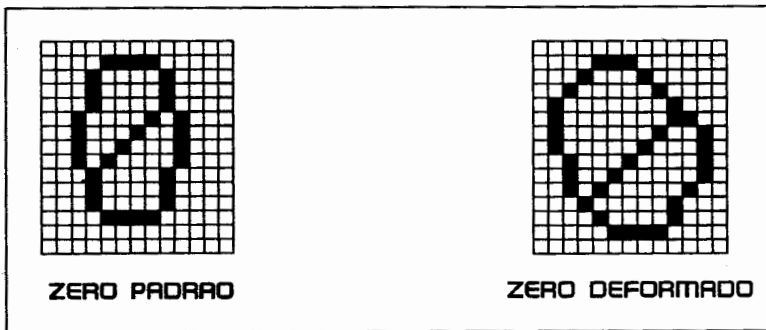


Fig. 5 - Algarismo 0 padrão e deformado.

## 5. CONSIDERAÇÕES FINAIS

Os resultados encontrados foram plenamente satisfatórios. Já era esperado o melhor desempenho do processo de reconhecimento dos caracteres ampliados frente aos reduzidos. Principalmente pelas dimensões das matrizes dos caracteres padrões (12x8). Se as matrizes fossem maiores, os resultados em relação à redução seriam melhores. Já em relação ao conjunto de caracteres deformados, o



critério de análise tem muito de subjetivo, pois o grau de deformação de cada caráter foi diferente.

A propriedade que apresentou melhor poder de discriminação dos caracteres foi a da vizinhança, só perdendo para a das projeções quando consideradas em conjunto. No entanto, a propriedade de vizinhança não separa o algarismo 6 do algarismo 9, já que um nada mais é que o outro após uma rotação de 180°.

As assinaturas tiveram seus desempenhos diretamente associados às suas dimensões. A de melhor desempenho foi a que combinava vizinhança com projeção (dimensão = 20), seguida da vizinhança + densidade (dimensão = 13), vindo em último a vizinhança + simetria (dimensão = 8).

O critério para avaliar o desempenho foi comparar os graus de dissemelhança entre cada um dos elementos dos conjuntos auxiliares com os dois elementos mais semelhantes do dicionário padrão. Quanto maior a diferença entre estes graus de dissemelhança, melhor o poder de discriminação. Outro critério utilizado foi verificar a lista ordenada resultante do reconhecimento com a matriz de correlação dos caracteres.

Dando seqüência a este trabalho, pretende-se estudar novas combinações de propriedades, bem como novas propriedades, que apresentem grande poder de discriminação. As letras minúsculas deverão ser incluídas no dicionário padrão, e este deverá apresentar esqueletos de caracteres com dimensões superiores a 12x8.

Em relação à normalização, serão testadas novas dimensões para os vetores de projeção, e novas formas para normalizar os parâmetros das diversas propriedades. A propriedade de simetria será testada diretamente do caráter, e não apenas dos vetores de projeção. Para isto será utilizado o menor retângulo que englobe o caráter em questão.

O objetivo de todas estas modificações é diminuir ao máximo as dimensões da assinatura sem comprometimento do poder de discriminação dos caracteres.

## BIBLIOGRAFIA

CHIARIGLIONE, L. ATRAS: Un Sistema di Lettura Automatica di Testi. Elettronica e Telecomunicazioni, N.2, pp. 54-62, March 1983.

**MAIER, M.; PORINELLI, R.** Separating Graphic Objects in Written Text. International Conference on Advances in Image Processing and Pattern Recognition, Pisa, Italy, December 10-12, 1985.

**LANGDON Jr., G.G.** An Introduction to Arithmetic Coding. IBM Journal of Research and Development, (28), pp.135-149, 1984.