

Abordagem Morfológica para a avaliação da inclinação de documentos contendo linhas

Luiz Antônio Pereira Neves –
IPPUC - Instituto de Pesquisa e Planejamento Urbano de Curitiba,
Rua Bom Jesus 669, Cabral, 80035-010 Curitiba-Pr, Brasil
Email: neves@ippuc.curitiba.pr.gov.br

Jacques Facon
PUCPR – Pontifícia Universidade Católica do Paraná
PPGIA - Programa de Pós-Graduação em Informática Aplicada
Rua Imaculada Conceição 1155, Prado Velho, 80215-901 Curitiba-Pr, Brasil
Email: facon@ppgia.pucpr.br

Resumo: Neste artigo, apresenta-se uma abordagem para detectar e avaliar o grau de inclinação de imagens binárias de documentos contendo linhas (tabelas, gráficos etc...). O método baseia-se em operadores morfológicos e numa família específica de elementos estruturantes. O método proposto não requer nenhuma informação a priori e nenhuma segmentação prévia do documento. Resultados experimentais sobre vários tipos de tabelas e gráficos são apresentados. Uma possível extensão dessa abordagem para documentos que não contêm linhas é introduzida.

Palavras-chave: Documento binário, Inclinação, Projection profile, Transformada de Hough, Morfologia matemática binária.

1. Introdução

Há uma série de pesquisas na área de análise e reconhecimento de documentos pre-impresos e manuscritos em busca de soluções que ainda não foram abrangidas. As pesquisas nessa área não são uma tarefa trivial devido a uma série de fatores. Dentre esses fatores podemos citar a própria complexidade dos documentos pre-impresos e manuscritos. Para reduzir essa variabilidade, muitos sistemas de análise e reconhecimento de documentos corrigem a possível inclinação desses documentos, de forma a minimizar erros na fase de reconhecimento.

Esse estudo tem como propósito apresentar uma abordagem morfológica de detecção e avaliação do grau de inclinação de documentos apresentando linhas. Na seção 2 será apresentado o estado da arte em relação à detecção da inclinação da documentos. A seção 3 fará uma breve descrição de ferramentas morfológicas, descreverá a abordagem de avaliação da inclinação de documentos desenvolvida e ilustrará alguns resultados obtidos. A seção 5 ilustrará possíveis extensões do método.

2. Estado da Arte

Na literatura encontram-se várias abordagens de detecção da inclinação de documentos [ISH93][LE94] [CHE95] [YU96][CHA97-a] [CHA97-b] [GAT97] [JIA97]. [ISH93] sugere calcular um parâmetro de complexidade baseado no número de transições preto-brancas em imagens binárias (formato A4 em 300 DPI) para avaliar o grau de inclinação. [ISH93] aprimora essa primeira avaliação buscando definir corretamente o grau de inclinação em regiões complexas contendo linhas de texto. [ISH93] demonstra que essa abordagem permite tratar imagens com elementos não textuais como figuras, tabelas, fotografias ou ainda ruído. A precisão obtida é de 0,12 graus e o tempo de processamento é aproximadamente de 1,38 segundos CPU numa workstation Sparc 2.

A transformada de Hough é muito usada para a detecção da inclinação. Nesse caso, a transformada consiste em extrair as características das linhas. [HIN90] aprimorou essa abordagem usando o conceito de imagem "bursty" onde o valor total de pixels de cada direção da linha é colocado no extremo direito ou inferior da linha. A partir

desses valores, faz-se a mudança clássica do espaço de representação para coordenadas polares. Acha-se a inclinação através da tabela de acumulação da transformada de Hough. Apesar dessa melhoria, esta técnica possui a desvantagem de ser lenta. [YU96] pela associação da transformada de Hough hierárquica com uma estrutura de componentes conexos BAG demonstra que é possível avaliar de forma rápida e robusta o grau de inclinação de documentos genéricos de baixa resolução. A partir de imagens (413x575 em 50 DPI) com elementos não textuais como figuras, tabelas, fotografias, partes manuscritas ou códigos de barra, a precisão obtida é de 0,10 graus com um tempo de processamento de aproximadamente de 0,45 segundos CPU numa workstation Sparc 20. De maneira análoga [JIA97] emprega a transformada de Hough hierárquica para aprimorar uma primeira avaliação da inclinação de documentos obtida através dos cantos esquerda superior de blocos de tamanho empírico de 100 colunas. O tempo de processamento é aproximadamente de 4,16 segundos para documentos inclinados de -31 a -45 graus num IBM 486 PC.

Uma outra técnica muito usada é a de "Projection Profile" [OGO98]. A técnica de "Projection Profile" consiste em efetuar rotações sucessivas da imagem em estudo e criar para cada imagem rotacionada uma projeção horizontal ou vertical, contando o acúmulo de pixels pretos de cada linha. Monta-se um histograma e verifica-se a maior altura da projeção para calcular o ângulo. A rotação que corresponde a esse maior altura representa o grau de inclinação do documento. Os autores comentam que a sua eficácia no acerto da inclinação está limitada para ângulos menores que 10 graus. Esta técnica possui a desvantagem de requerer um tempo de processamento exaustivo. [LE94] aprimora essa técnica empregando uma estrutura de dados piramidal sobre porções textuais de imagens associada a transformada de Hough. Com essa abordagem, [LE94] consegue processar imagens binárias (formato A4 em 200 DPI) com uma precisão de 99,9 % num tempo de processamento aproximadamente de 4,15 segundos para definir a orientação de uma página e de 3,85 segundos para avaliar a inclinação de um documento num DELL 486D/50 PC.

[CHE95] calcula o grau de inclinação de imagens binárias (3300x2550 pixels) pelo estimador de Bayes sobre subconjuntos de imagens adequadamente escolhidos. Para isso é preciso empregar aberturas e fechamentos morfológicas recursivos para preencher espaços inter-caracteres e para remover ascendentes e descendentes. O tempo de processamento é aproximadamente de 10 segundos CPU numa workstation Sparc 10.

[CHA97-a] apresenta um método de detecção de inclinação de imagens binárias (300 DPI) inclinadas de -15 a +15 graus que associa a correlação cruzada entre linhas com a técnica de Monte Carlo. Nenhuma segmentação textual não textual é necessária. O tempo de processamento é aproximadamente de 1 segundo numa DEC Alpha 400 station.

[CHA97-b] emprega a rotulação de componentes conexos e o contorno superior desses componentes conexos para avaliar o grau de inclinação de imagens binárias (300 DPI) de documentos indianos inclinados entre 0 e + 45 graus. Numa Sun machine 3, o tempo de processamento é aproximadamente de 17,8 segundos para uma precisão de 1 grau.

3. Abordagem morfológica de avaliação de inclinação

3.1 Introdução sobre morfologia matemática

A morfologia matemática representa uma abordagem não linear de processamento de imagens. O seu princípio básico consiste em extrair informações relativas à geometria e à topologia de um conjunto desconhecido de uma imagem a partir de um conjunto perfeitamente definido, chamado de Elemento Estruturante [FAC96].

A morfologia matemática apoia-se em dois operadores básicos que são a erosão e a dilatação. Erodir um conjunto X consiste em reduzir de forma isotrópica o tamanho desse conjunto. Essa redução é proporcional ao tamanho do elemento estruturante empregado. Dilatar o conjunto X consiste em adicionar de forma isotrópica informações proporcionais ao tamanho do elemento estruturante. As definições das erosão e dilatação envolvendo o formalismo de Minkowski [SER82] são:

A erosão ε de um conjunto X pelo elemento estruturante B é:

$$\varepsilon^B(X) = X \ominus B = \bigcap_{b \in B} X_b$$

A dilatação δ de um conjunto X pelo elemento estruturante B é:

$$\delta^B(X) = X \oplus \tilde{B} = \bigcup_{b \in B} X_b$$

A grande força e ao mesmo tempo o grande desafio da morfologia matemática reside na associação da escolha adequada e judiciosa do elemento estruturante com as propriedades dos operadores de erosão e de dilatação. A Fig.1 ilustra os efeitos da erosão e da dilatação com um elemento estruturante quadrado de tamanho 7x7 pixels.

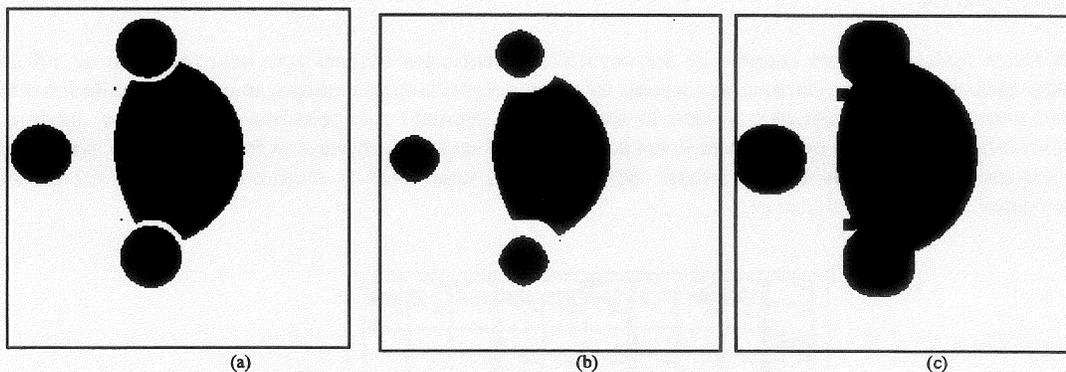


Fig 1 – Exemplos de erosão e de dilatação: (a) Imagem original; (b) Imagem erodida; (c) Imagem dilatada.

3.2 Metodologia de detecção e de avaliação da inclinação de documentos contendo linhas

Do ponto de vista morfológico, sabe-se que, qualquer que seja o elemento estruturante empregado, o uso da erosão tem como consequência reduzir os conjuntos processados enquanto que a dilatação os aumenta. Porém o grau de mudança no resultado do operador de erosão e de dilatação depende do elemento estruturante usado. Sabe-se ainda que um elemento estruturante parecido em formato e em direção ao conjunto estudado produz menores mudanças que um elemento estruturante diferente em formato ou em direção. Por exemplo, linhas horizontais sofrem menores desgastes no caso de erosão com um elemento estruturante linha horizontal tipo $B_{hor} = \{ \bullet \bullet \bullet \}$ que no caso de erosão com um elemento

estruturante linha vertical $B_{ver} = \left\{ \begin{matrix} \bullet \\ \bullet \\ \bullet \end{matrix} \right\}$. Isto vale também para a dilatação.

A abordagem adotada na detecção e na avaliação da inclinação de documentos contendo linhas consiste em avaliar o resultado dos operadores de erosão e de dilatação medindo o grau de mudanças ocorridos nos resultados em função do tipo de elemento estruturante empregado. No caso de documentos que contêm linhas, a idéia reside no uso de vários elementos estruturantes de tipo linha apresentando várias inclinações possíveis e medir o grau de mudança ocorrida. Que sejam empregadas a erosão ou a dilatação, a menor mudança ocorrerá para o elemento estruturante apresentando a mesma inclinação ou a mais parecida com a do documento. No caso da erosão, a menor mudança corresponde à menor redução de pixels e no caso da dilatação à menor adição de pixels. A Fig.2 ilustra esses efeitos no caso de uma tabela sem inclinação erodida e dilatada com dois elementos estruturantes de tipo linha de comprimento 11 pixels, o primeiro horizontal e o segundo inclinado de 9 graus. Percebe-se que, no caso da erosão, o desgaste é menor usando o elemento estruturante horizontal e, que, no caso da dilatação, a adição de pixel é também menor com o elemento estruturante horizontal.

A metodologia empregada vai consistir em erodir (ou dilatar) a imagem a partir de uma família de elementos estruturantes inclinados adequadamente definidos para atender à realidade, e em contar o número de pixels sobrando (ou adicionados no caso da abordagem por dilatação).

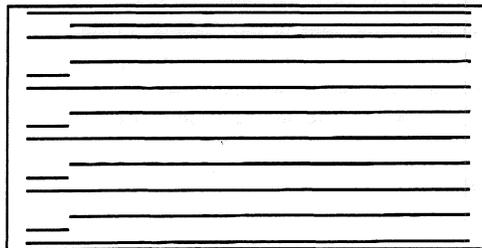
Devido ao problema de digitalização, um estudo preliminar foi realizado de maneira a definir o número ideal de pixels necessários nos elementos estruturantes de tipo linha para obter uma boa precisão na detecção da inclinação. O grau de precisão estudado é 1 grau. Desse estudo foi deduzido que os elementos estruturantes devem ser constituídos de 57 pixels. Com esse tamanho, a medida de abertura para cada ângulo é distinta, proporcionando uma detecção confiável (Fig 3).

3.3 Resultados Obtidos

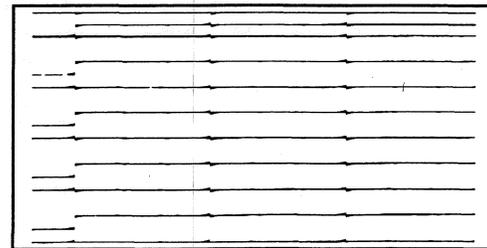
Testes foram realizados sobre imagens de documentos adquiridas por scanner com uma resolução de 300 DPI. Inicialmente adquiridas em níveis de cinza, essas imagens sofreram um processo de binarização pelo método de Johannsen [JOH82]. Como muitos desses documentos tinham basicamente um formato A4, percebeu-se que o erro máximo de posicionamento no scanner cometido por um usuário era de 10 graus. Portanto, a abordagem foi testada para avaliar graus de inclinação variando de -10 graus até +10 graus. Seguem abaixo uma parte da família de elementos estruturantes definidos para atender a essa realidade, na Fig 4.

ângulo	FLUXO			FLUXO			FLUXO		
	comprimento	espessura	abertura	comprimento	espessura	abertura	comprimento	espessura	abertura
-90	57	1	57	57	1	57	57	1	57
-75	57	1	57	57	1	57	57	1	57
-60	57	1	57	57	1	57	57	1	57
-45	57	1	57	57	1	57	57	1	57
-30	57	1	57	57	1	57	57	1	57
-15	57	1	57	57	1	57	57	1	57
0	57	1	57	57	1	57	57	1	57
15	57	1	57	57	1	57	57	1	57
30	57	1	57	57	1	57	57	1	57
45	57	1	57	57	1	57	57	1	57
60	57	1	57	57	1	57	57	1	57
75	57	1	57	57	1	57	57	1	57
90	57	1	57	57	1	57	57	1	57

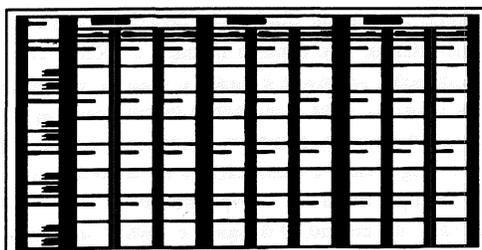
(a)



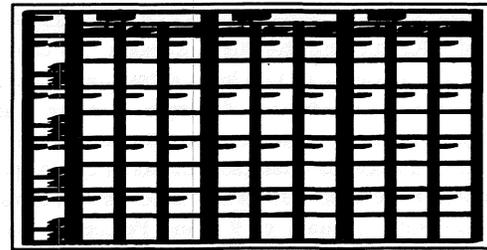
(b)



(c)



(d)



(e)

Fig 2 – Exemplo de efeitos da erosão e da dilatação com os elementos estruturantes EE1 (linha horizontal) e EE2 (linha inclinada de 9 graus): (a) Imagem original; (b) Imagem erodida por EE1; (c) Imagem erodida por EE2; (d) Imagem dilatada por EE1; (e) Imagem dilatada por EE2.

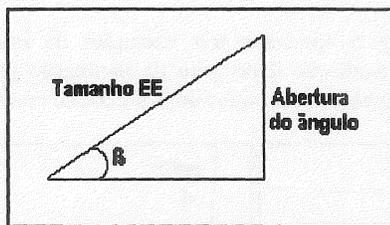


Fig 3 – Tamanho Elemento Estruturante versus abertura do ângulo.

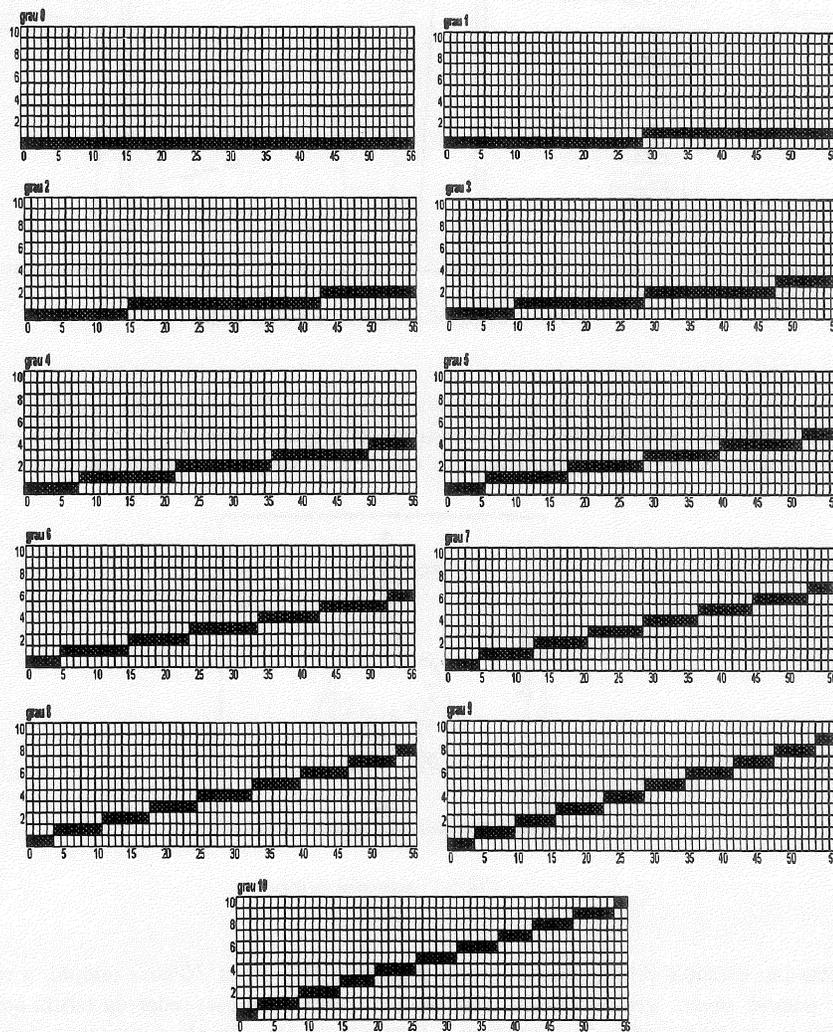


Fig.4 – Família de elementos estruturantes entre 0 e +10 graus.

Nos testes realizados, comparou-se a abordagem morfológica com as técnicas de “Projection Profile” onde só a projeção horizontal foi empregada e da transformada de Hough otimizada com o conceito de imagem “bursty”

unicamente no sentido horizontal. A Fig 5 apresenta três exemplos de imagens de documentos proposadamente inclinadas de 3 graus. Os resultados da avaliação dado grau de inclinação pelas três foram idênticos. Os tempos de processamento respectivos mostraram que a abordagem morfológica é muito mais veloz (Tabela 1).

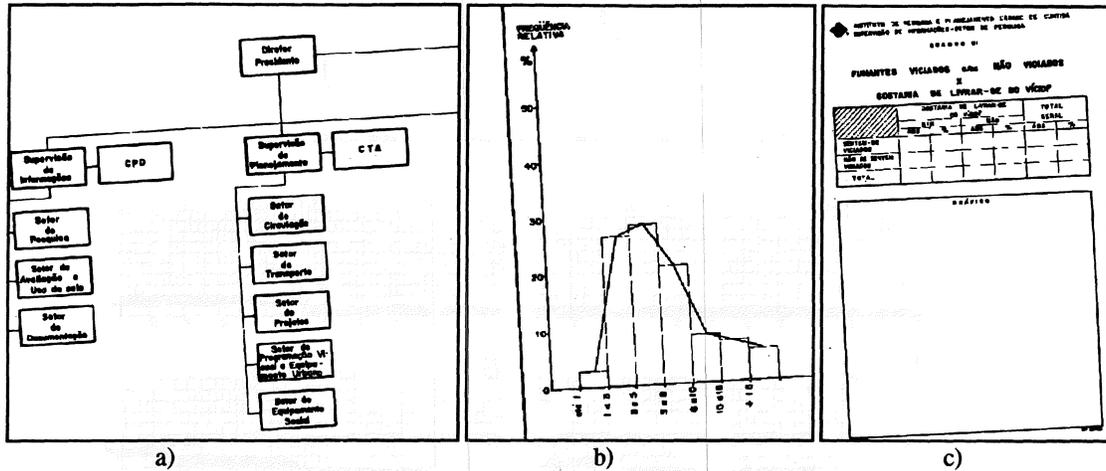


Fig 5 : Exemplos de documentos contendo linhas.

A Fig 6 ilustra um exemplo de imagem de logotipo não inclinado. Nesse exemplo, percebeu-se que a abordagem por "Projection Profile" detectou erradamente uma inclinação de 2 graus enquanto a "Transformada de Hough" e a abordagem morfológica avaliaram corretamente a inclinação com tempos de processamento parecidos (Tabela 1).



Fig 6 : Exemplo de logotipo

A Fig 7 ilustra um exemplo de imagem de tabela inclinada de 3 graus. Nesse exemplo, a tabela foi adquirida deixando a tampa do scanner aberta, gerando assim uma faixa cinzenta escura ao redor da tabela que foi transformada como preta pelo processo de binarização de Johannsen. Percebeu-se que, devida à presença dessa faixa preta, as abordagens por "Projection Profile" e pela "Transformada de Hough" não conseguiram determinar a inclinação da tabela. Pela abordagem morfológica, tanto por erosão como por dilatação, a avaliação da inclinação foi de 3 graus com um tempo de processamento de 15 segundos por erosão (Tabela 1).

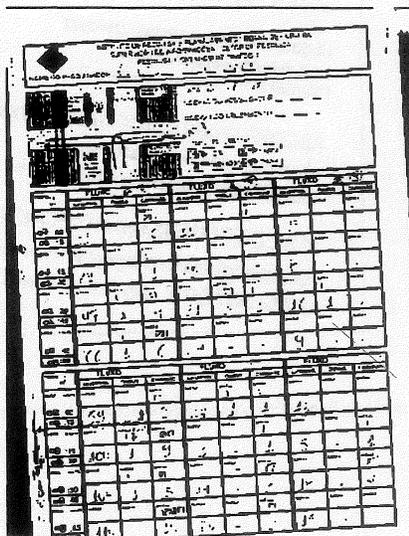


Fig 7 : Imagem de tabela "mal adquirida"

A Tabela 1 agrupa os resultados numéricos realizados num Personal Compatible Pentium 133, 32 bits, das abordagens por "Projection Profile", pela Transformada de Hough e pela abordagem morfológica proposta a partir as imagens apresentadas nesse artigo.

Métodos		Projection Profile.		Transformada de Hough		Abordagem Morfológica	
Imagem	Tamanho	ângulo	tempo(s)	ângulo	tempo(s)	ângulo	tempo(s)
Fig 5-a	854x835	3	10	3	6	3	3,26
Fig 5-b	430x541	3	3,5	3	2	3	1,23
Fig 5-c	642x976	3	9,5	3	4,5	3	2,96
Fig 6	327x333	2	2	0	1	0	0,42
Fig 7	3502x2680	-	-	-	-	3	15

Tabela 1: Resultados numéricos das três abordagens testadas

4. Extensão do Método

A abordagem descrita anteriormente visa a processar imagens de documentos contendo linhas. Nesse caso específico, tanto o processo de avaliação do grau de inclinação por erosão ou por dilatação fornece resultados precisos. No caso de outros tipos de documentos sem linhas (texto etc.), essa abordagem nova mostrou-se eficiente e precisa somente por dilatação. A abordagem por erosão, por eliminar excessivamente informações contidas em documentos outros que tabelas, revelou-se ineficiente. A Fig 8 ilustra uma imagem de texto onde a abordagem morfológica por dilatação mostrou-se eficiente e precisa (inclinação de 6 graus).

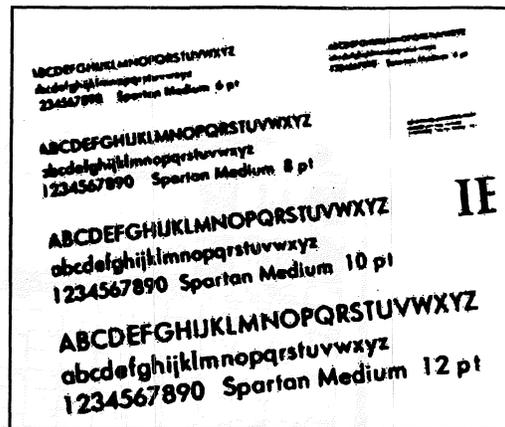


Fig 8 – Documento sem linha eficientemente processado pela abordagem morfológica por dilatação

5. Conclusão

A meta desse estudo foi apresentar uma abordagem morfológica para detectar e avaliar o grau de inclinação de documentos contendo linhas. Uma família específica de elementos estruturantes foi definida para corrigir esse tipo de documentos cuja inclinação pode variar entre -10 e $+10$ graus. A abordagem morfológica mostrou-se eficiente para qualquer tipo de documento testado. No caso de outros tipos de documentos (texto por exemplo), essa abordagem nova por dilatação mostrou-se também eficiente e precisa. Além da precisão, pelas medições em tempo efetuadas, essa abordagem nova comprovou-se mais rápida em alguns casos que as abordagens clássicas. No futuro, pretende-se reduzir esse tempo decompondo os elementos estruturantes empregados pelo método de [BOO92] ou ainda usando a técnica recursiva de [SOI96].

6. Referências

- [BOO92] Boomgaard R.V.D. and Van Balen R., "Methods for Fast Morphological Image Transforms Using Bimapped Binary Images", *CVGIP Graphical Models and Image Processing*, vol. 54, No 3, pp 252-258, march 1992.
- [CHA97-a] Chaudhuri A and Chaudhuri S., "Robust Detection of skew in Document Images", *IEEE Transactions on Image Processing*, vol. 6, No 2, pp 344-349, february 1997.
- [CHA97-b] Chaudhuri B.B. and Pal U., "Skew angle detection of digitized Indian script documents", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 19, No 2, pp 182-186, february 1997.
- [CHE95] Chen S. and Haralick R.M. and Phillips I. T., "Automatic text skew estimation in document images", *IEEE*, pp 1153-1556, 1995.
- [FAC96] Facon J., *Morfologia Matemática: Teoria e Exemplos*, Gráfica Champagnat, Curitiba, Brasil, october 1996.
- [GAT97] Gatos B. and Papamarkos N. and Chamzas C., "Skew Detection and text line position determination in digitized documents", *Pattern Recognition*, vol. 30, No 9, pp 1505-1519, 1997.

- [HIN90] Hinda S.C. and Fisher J.L. and D'Amato D.P. , "A document skew detection method using run-length encoding and the Hough transform", *Proceedings of 10th Int. Conf. Pattern Recognition*, 1, pp.464-468, 1990.
- [ISH93] Ishitani Y., "Document skew detection based on local region complexity", *Proceedings of 2rd Int. Conf. on Document Analysis and Recognition*, Tsukuba Science City, Japon, pp.49-52, october 1993.
- [JIA97] Jiang H.F. and Huan C.C. and Fan K.C., "A fast approach to the detection and correction of skew documents", *Pattern Recognition Letters*, vol. 18, pp 675-686, 1997.
- [JOH82] Johansen G. and Bille J., "A Threshold Selection Method using Information Measures", *Proceedings of 6th Int. Conf. Pattern Recognition*, Munich, Germany, pp.140-143, 1982.
- [LE94] Le D.S. and Thoma G. R. and Wechsler H., "Automated Page Orientation and Skew Angle Detection for binary document images", *Pattern Recognition*, vol. 27, No 10, pp 1325-1344, 1994.
- [OGO98] O'Gorman, L. and Kasturi R., *Document image analysis*. IEEE Computer Society Press, USA, 1998.
- [SER82] Serra J., *Image Analysis and Mathematical Morphology*, Academic Press, London, 1982.
- [SOI96] Soille P. and Breen E.J. and Jones R., "Recursive Implementation of Erosions and Dilations along Discrete Lines at Arbitrary Angles", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18, No 5, pp 562-568, may 1996.
- [YU96] Yu B. and Jain A. K., "A Robust and fast skew detection algorithm for generic documents", *Pattern Recognition*, vol. 29, No 10, pp 1599-1629, 1996.

