Para uma correta análise de erros deve-se abstrair das formas das palavras e ter em mente a representação utilizada, ou seja, os conjuntos de primitivas selecionados. Caso contrário, torna-se difícil entender a similaridade de formas e as palavras que são fontes de confusão.

As matrizes de confusão mostram que a similaridade das formas pode ser maior do que a esperada. [OLLIVIER, 1999] afirma que a similaridade de formas deve ser simétrica em relação a matriz de confusão. Ao buscar-se esta simetria na matriz obtida para o conjunto de primitivas perceptivas percebe-se que para alguns casos isto é verdadeiro, como por exemplo: “seis”, “nove” e “um”, “setenta” e “oitenta”, “trezentos” e “duzentos”. Por outro lado, existem casos onde a similaridade simétrica de formas não é verdadeira, como nos seguintes exemplos: “três” e “um”, “dezoito” e “duzentos”, “sessenta” e “noventa”, “novecentos” e “seiscentos”, “seiscentos” e “setecentos”.

As palavras-chaves devem ser analisadas em separado das demais, pois são palavras diferentes entre si na forma agrupadas somente por interesse e, portanto, o grau de confusão entre as mesmas é muito baixo (apenas alguns exemplares) independente do conjunto de primitivas globais utilizado. Observa-se, no entanto, que a palavra “mil” tem sua fonte de confusão na palavra “vinte” (PP e PPCC) e “seis” (PPCCR). Para a palavra “reais/real” a confusão está na palavra “nove” (PP e PPCC) e “seis” (PPCCR). Para a palavra "centavos/centavo" a confusão torna-se evidente com a palavra "seiscentos" (PP), com as palavras "sessenta", "setenta", "oitenta" e "seiscentos" (PPCC) e com as palavras "oitenta", "quinhentos" e "seiscentos" (PPCCR).

Para auxiliar a análise de erros as Tabela 4.15 a 4.17 resumem as matrizes de confusão, para os 3 conjuntos de primitivas (PP, PPCC e PPCCR), de forma a classificar os erros cometidos pelo método de reconhecimento intra-classe e inter-classes e, ainda, especificar em que classes estão sendo reconhecidas as palavras encontradas na matriz de confusão como inter-classes.

As matrizes de confusão e a classificação dos erros permite analisar três fatores de influência na taxa de reconhecimento das palavras manuscritas no contexto do léxico em português para cheques bancários, sendo estes os seguintes:

- a similaridade das formas,
- a freqüência da palavra na base de treinamento,
- os conjuntos de primitivas selecionados.

Tabela 4.15: Análise de Erros - PP

	Palavras	Tipo de Erro		Erro Inter-Classes de Palavras				
		Intra	***Inter***	**Unidade**	**Dezena**	**Enta**	**Entos**	**PC**
UNIDADE	Um	12	1	-	1	0	0	0
	Dois	13	12	-	10	1	0	1
	Três	22	13	-	5	4	3	1
	Quatro	2	17	-	11	5	0	1
	Cinco	16	8	-	2	2	2	2
	Seis	30	9	-	3	1	0	5
	Sete	18	12	-	7	0	2	3
	Oito	9	7	-	5	0	1	1
	Nove	10	14	-	4	2	0	8
	TOTAL	**58,67**	**41,33**	**-**	**51,61**	**16,13**	**8,60**	**23,66**
DEZENA	Dez	2	1	1	-	0	0	0
	Onze	2	1	0	-	0	1	0
	Doze	3	0	0	-	0	0	0
	Treze	0	1	1	-	0	0	0
	Quatorze	2	3	1	-	1	0	1
	Quinze	1	3	3	-	0	0	0
	Dezesseis	3	0	0	-	0	0	0
	Dezessete	2	3	0	-	1	2	0
	Dezoito	1	4	1	-	0	3	0
	Dezenove	1	3	1	-	0	2	0
	TOTAL	**47,22**	**52,78**	**42,11**	**-**	**10,53**	**42,11**	**5,26**
ENTA	Vinte	10	26	15	2	-	5	4
	Trinta	3	24	3	2	-	14	5
	Quarenta	4	23	5	6	-	12	0
	Cinqüenta	3	13	2	1	-	10	0
	Sessenta	22	35	6	2	-	25	2
	Setenta	16	34	10	4	-	17	3
	Oitenta	10	26	6	9	-	7	4
	Noventa	8	22	0	0	-	18	4
	TOTAL	**27,24**	**72,76**	**23,15**	**12,81**	**-**	**53,20**	**10,84**
ENTOS	Cem/Cento	0	29	10	1	17	-	1
	Duzentos	6	10	2	2	3	-	3
	Trezentos	26	6	1	2	1	-	2
	Quatrocentos	10	8	1	2	4	-	1
	Quinhentos	5	4	0	1	3	-	0
	Seiscentos	29	13	1	0	9	-	3
	Setecentos	12	10	0	2	8	-	0
	Oitocentos	10	13	1	4	6	-	2
	Novecentos	9	2	0	0	2	-	0
	TOTAL	**52,97**	**47,03**	**16,84**	**14,74**	**55,79**	**-**	**12,63**
PC	Mil	0	103	58	3	37	5	-
	Reais/Real	1	140	116	9	10	5	-
	Centavos/Centavo	0	42	4	4	15	19	-
	TOTAL	**0,35**	**99,65**	**62,46**	**5,61**	**21,75**	**10,18**	**-**

Tabela 4.16: Análise de Erros - PPCC

	Palavras	Tipo de Erro		Erro Inter-Classes de Palavras				
		Intra	***Inter***	**Unidade**	**Dezena**	**Enta**	**Entos**	**PC**
UNIDADE	Um	8	3	-	1	1	1	0
	Dois	11	13	-	8	2	0	3
	Três	19	11	-	4	3	2	2
	Quatro	1	13	-	9	4	0	0
	Cinco	3	14	-	1	7	4	2
	Seis	19	11	-	6	2	0	3
	Sete	16	10	-	4	3	0	3
	Oito	10	6	-	4	1	1	0
	Nove	3	13	-	1	3	0	9
	TOTAL	**48,91**	**51,09**	**-**	**40,43**	**27,66**	**8,51**	**23,40**
DEZENA	Dez	4	0	0	-	0	0	0
	Onze	2	4	3	-	0	1	0
	Doze	3	0	0	-	0	0	0
	Treze	2	1	1	-	0	0	0
	Quatorze	1	3	1	-	0	2	0
	Quinze	0	2	1	-	1	0	0
	Dezesseis	2	2	0	-	0	2	0
	Dezessete	3	0	0	-	0	0	0
	Dezoito	0	3	2	-	0	1	0
	Dezenove	3	0	0	-	0	0	0
	TOTAL	**57,14**	**42,86**	**53,33**	**-**	**6,67**	**40,00**	**0,00**
ENTA	Vinte	5	16	9	0	-	6	1
	Trinta	18	30	5	2	-	21	2
	Quarenta	6	19	3	6	-	10	0
	Cinqüenta	4	9	0	2	-	7	0
	Sessenta	21	24	0	4	-	19	1
	Setenta	14	24	1	4	-	18	1
	Oitenta	17	21	2	5	-	10	4
	Noventa	7	11	0	2	-	8	1
	TOTAL	**37,40**	**62,60**	**12,99**	**16,23**		**64,29**	**6,49**
ENTOS	Cem/Cento	1	22	6	2	13	-	1
	Duzentos	6	9	2	1	4	-	2
	Trezentos	22	3	0	1	2	-	0
	Quatrocentos	5	8	1	2	5	-	0
	Quinhentos	6	4	0	0	4	-	0
	Seiscentos	32	10	1	1	7	-	1
	Setecentos	13	8	0	2	5	-	1
	Oitocentos	9	6	0	0	6	-	0
	Novecentos	9	6	0	1	5	-	0
	TOTAL	**57,54**	**42,46**	**13,16**	**13,16**	**67,11**	**-**	**6,58**
PC	Mil	1	42	26	1	13	2	-
	Reais/Real	2	77	57	7	8	5	-
	Centavos/Centavo	1	29	2	5	12	10	-
	TOTAL	**2,63**	**97,37**	**57,43**	**8,78**	**22,30**	**11,49**	**-**

Tabela 4.17: Análise de Erros - PPCCR

	Palavras	Tipo de Erros		Erro Inter-Classes de Palavras				
		Intra	*Inter*	Unidade	Dezena	Enta	Entos	PC
UNIDADE	Um	4	1	-	1	0	0	0
	Dois	7	12	-	7	1	1	3
	Três	18	9	-	3	0	3	3
	Quatro	1	11	-	5	4	0	2
	Cinco	5	10	-	1	6	2	1
	Seis	17	12	-	4	1	1	6
	Sete	16	14	-	6	6	0	2
	Oito	11	5	-	3	0	2	0
	Nove	6	9	-	1	4	0	4
	TOTAL	**50,60**	**49,40**	**-**	**37,35**	**26,51**	**10,84**	**25,30**
DEZENA	Dez	3	0	0	-	0	0	0
	Onze	2	3	2	-	0	1	0
	Doze	4	0	0	-	0	0	0
	Treze	1	1	0	-	0	0	1
	Quatorze	1	4	0	-	1	1	2
	Quinze	1	3	2	-	1	0	0
	Dezesseis	3	2	0	-	0	2	0
	Dezessete	2	3	0	-	1	2	0
	Dezoito	1	1	1	-	0	0	0
	Dezenove	1	0	0	-	0	0	0
	TOTAL	**52,78**	**47,22**	**29,41**	**-**	**17,65**	**35,29**	**17,65**
ENTA	Vinte	9	19	11	0	-	5	3
	Trinta	19	20	2	5	-	12	1
	Quarenta	4	19	3	3	-	11	2
	Cinqüenta	5	11	1	0	-	10	0
	Sessenta	18	27	1	3	-	22	1
	Setenta	25	21	1	4	-	15	1
	Oitenta	14	18	3	3	-	9	3
	Noventa	10	12	0	0	-	12	0
	TOTAL	**37,40**	**62,60**	**12,99**	**16,23**		**64,29**	**6,49**
ENTOS	Cem/Cento	2	21	4	1	14	-	2
	Duzentos	6	9	1	3	4	-	1
	Trezentos	17	5	0	1	4	-	0
	Quatrocentos	6	8	0	4	3	-	1
	Quinhentos	7	3	0	0	3	-	0
	Seiscentos	35	7	1	0	5	-	1
	Setecentos	11	10	0	1	9	-	0
	Oitocentos	10	7	0	0	6	-	1
	Novecentos	6	3	0	0	3	-	0
	TOTAL	**57,80**	**42,20**	**8,22**	**13,70**	**69,86**	**-**	**8,22**
PC	Mil	1	38	25	1	12	0	-
	Reais/Real	3	79	58	10	6	5	-
	Centavos/Centavo	1	36	5	5	10	16	-
	TOTAL	**3,16**	**96,84**	**57,52**	**10,46**	**18,30**	**13,73**	**-**

Alguns aspectos da similaridade das formas das palavras já foram mencionados anteriormente, deseja-se neste momento discutir a similaridade de formas entre palavras de uma mesma classes e entre as classes de palavras.

Na Figura 1.3, Capítulo 1, apresenta-se o conjunto de palavras do léxico o qual tem sua origem no valor numérico grafado no cheque bancário. Agora, pode-se observar o mesmo conjunto de palavras levando-se em consideração a similaridade de formas entre as palavras de uma mesma classes. Esta similaridade torna-se evidente com os sufixos e prefixos destacados na Figura 1.3, a saber: “ dez-”, “-ze”, “-enta” e “-entos”.

As Tabelas 4.15 a 4.17 mostram que a similaridade de formas entre as classes de palavras ocorre da seguinte forma:

- unidade com dezena e dezena com unidade,
- "enta" com "entos" e vice-versa,
- palavras-chaves com unidade, porém deve-se observar que esta similaridade é válida para as palavras "mil" e "reais/real". No caso da palavra "centavos/centavo" a similaridade está com as classes "enta" e "entos".

Outro fator de influência é a freqüência da palavra na base de dados de treinamento. A Tabela 4.8 demonstra que a classe da dezena representa aproximadamente 3% dos conjuntos de treinamento e validação. Este fato, não é diferente com outros autores, tais como: [AVILA, 1996] e [OLLIVIER, 1999]. Sabe-se, então, que o processo de treinamento destas palavras não será adequado, confirmando suas deficiências nos experimentos de reconhecimento.

As Tabelas 4.15 a 4.17 permitem também uma análise dos conjuntos de primitivas selecionados. Observa-se que a % de erros de reconhecimento cometidos entre as classes de palavras diminui para as classes da unidade, da dezena e das palavras-chaves, quando compara-se os 3 conjuntos de primitivas PP, PPCC e PPCCR.

As classes "enta" e "entos" apresentam um acréscimo desta % de erro (inter-classes diferentes). Deste modo, pode-se concluir que o conjunto de primitivas perceptivas PP trabalha bem com as palavras da classe da dezena. Isto, explica-se uma vez que esta classe possui baixa freqüência na base de treinamento/validação sendo, portanto, melhor o treinamento quando realizado com um conjunto menor de símbolos.

O conjunto PPCC (primitivas perceptivas, concavidades e convexidades) trabalha bem com a classe "enta" e palavras-chaves. A exceção entre as palavras-chaves é a palavra "mil" que obtém os melhores resultados com o conjunto PPCCR.

Finalmente, o conjunto PPCCR (primitivas perceptivas, concavidades e convexidades rotuladas) trabalha bem para as classes de palavras da unidade e "entos". Para a classe "entos" pode-se ter o direito da incerteza entre os conjuntos PPCC e PPCCR. Esta classe gera confusões não somente pela similaridade com a classe "enta", mas também, por ser o conjunto de palavras mais longas do léxico em português e possuir como sufixos outras palavras do léxico, por exemplo: “quatro”, “seis”, “sete”, “oito”, “nove”. A Figura 4.16 apresenta alguns erros de reconhecimento de palavras ocorridos com os experimentos realizados.

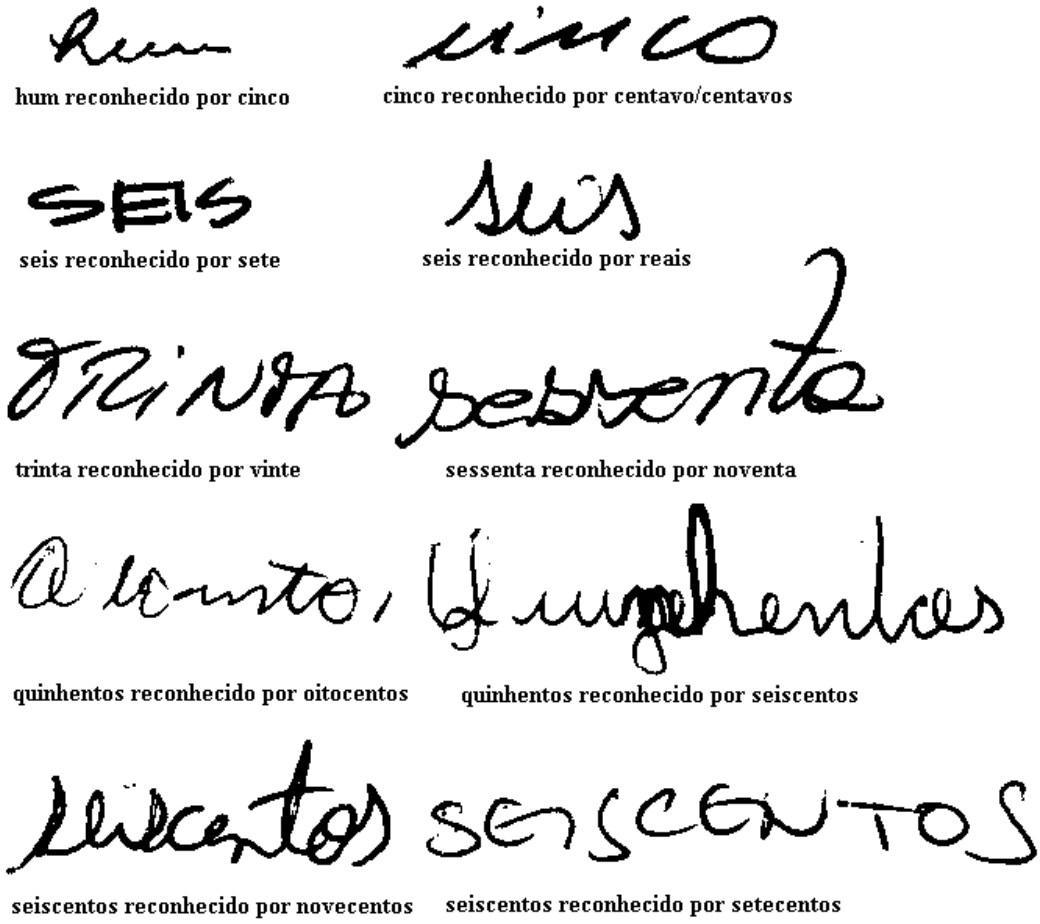

Figura 4.16: Exemplos de confusão no reconhecimento de palavras - conjunto PPCCR

Finalmente, conclui-se que os resultados obtidos são promissores, principalmente, quando comparados aos de outros autores e trabalhos já realizados, conforme apresentado no Apêndice B.

Como resultado da análise de erros percebe-se a necessidade de uma melhor representação das classes com ausência de primitivas perceptivas (ascendentes, descendentes e laços no corpo) e das palavras curtas, tais como: “cinco”, “seis”, “sete”,

“oito”, “nove”, “dez”, “onze”, “doze”, “cem/cento” e “mil”. Deve-se melhorar também a representação das palavras longas e com sufixos similares ("enta" e "entos"). E, ainda, das palavras que estão escritas em caixa alta não possuidoras de primitivas perceptivas. Assim, propõe-se um conjunto de primitivas geométricas apresentado a seguir neste capítulo.

4.3. Primitivas Geométricas

A implementação de um conjunto de primitivas geométricas tem por base a melhoria da representação das palavras curtas e das palavras sem presença de primitivas perceptivas (ascendentes, descendentes e laços no corpo). Pois, para [SCHOMAKER & SEGERS, 1998] as palavras curtas necessitam de um maior número de informações para serem reconhecidas e as palavras longas necessitam de mais informações no final das mesmas (Capítulo 2).

A idéia principal é permitir um enfoque mais local à extração de primitivas, com o objetivo de enriquecer as informações dos elementos lingüísticos que servem de “pistas” ao estabelecimento de uma relação entre o lingüistico e o conceitual-cognitivo (Capítulo2). Em outras palavras, busca-se uma nova chance de ler o que está escrito. Sendo que o interesse é melhorar a representação das regularidades de uma palavra, tal qual [CÔTE, 1997].

Assim, a utilização de primitivas geométricas em métodos de reconhecimento de palavras manuscritas explica-se diante do fato que estas apresentam tolerância às distorções e às variações de estilo e que estas podem descrever as propriedades globais ou locais de palavras e caracteres [HEUTTE, 1994].

Uma análise das partes côncavas e convexas das palavras permite evidenciar as propriedades topológicas e geométricas das formas a reconhecer. Para [HEUTTE, 1994] existem dois tipo de métodos que permitem extrair as concavidades, a saber:

- métodos com base na rotulação das áreas brancas das imagens ou do fundo ou *background* das imagens,
- métodos com base no tratamento do contorno das formas.

Opta-se pelo método composto pela rotulação dos pixels do fundo da imagem do corpo das palavras, de maneira a extrair para cada coluna da imagem um vetor de características ou *feature vector*. Métodos baseados no contorno das formas foram

descartados, visto que a extração de primitivas globais trabalha com a forma em si propriamente dita.

Outros autores desenvolveram conjuntos de primitivas com base no estudo das concavidades, rotulando os pixels do fundo da imagem, tais como: [LECOLINET, 1990], [GILLIES, 1992]. O primeiro autor trabalha com a extração de primitivas geométricas, rotulando os pixels do fundo das imagens e extraindo medidas estatísticas para valores numéricos. O segundo autor trabalha rotulando as concavidades e convexidades no corpo de imagens de palavras manuscritas de endereços postais.

4.3.1. Vetor de Características

A composição do vetor de características tem por base a rotulação de cada pixel do fundo da imagem mapeando-se 13 situações diferentes. Deste modo, este vetor armazena a quantidade de cada um dos 13 rótulos encontrados para cada uma das colunas da imagem no corpo das palavras, considerando-se dois tipos de verificação sobre a imagem, a saber:

- verificação dos 4 pixels vizinhos, norte-sul-leste-oeste (NSLO), para cada pixel branco do corpo da palavra e
- verificação de falso laço com possibilidade de escape em 4 direções de sondagem para os pixels determinados como fechado para as 4 direções principais (NSLO).

A Figura 4.17 exemplifica a rotulação dos pixels brancos do corpo da palavra "cinco" retirada da base de dados de treinamento, seguindo-se a rotulação especificada na Tabela 4.18.

Apresenta-se na Figura 4.17 a imagem original da palavra "cinco", o resultado da correção da inclinação vertical dos caracteres e a detecção do corpo da palavra. Um esquema da rotulação dos pixels do fundo da imagem mostra as direções fechadas (com seta), abertas (sem seta) e com escape (seta ⌟) para alguns pixels da imagem. Mostra-se a extração do vetor de características para uma das colunas da palavra "cinco" exemplificada. Cada posição do vetor contabiliza o total de cada tipo de rótulo encontrado na coluna analisada.

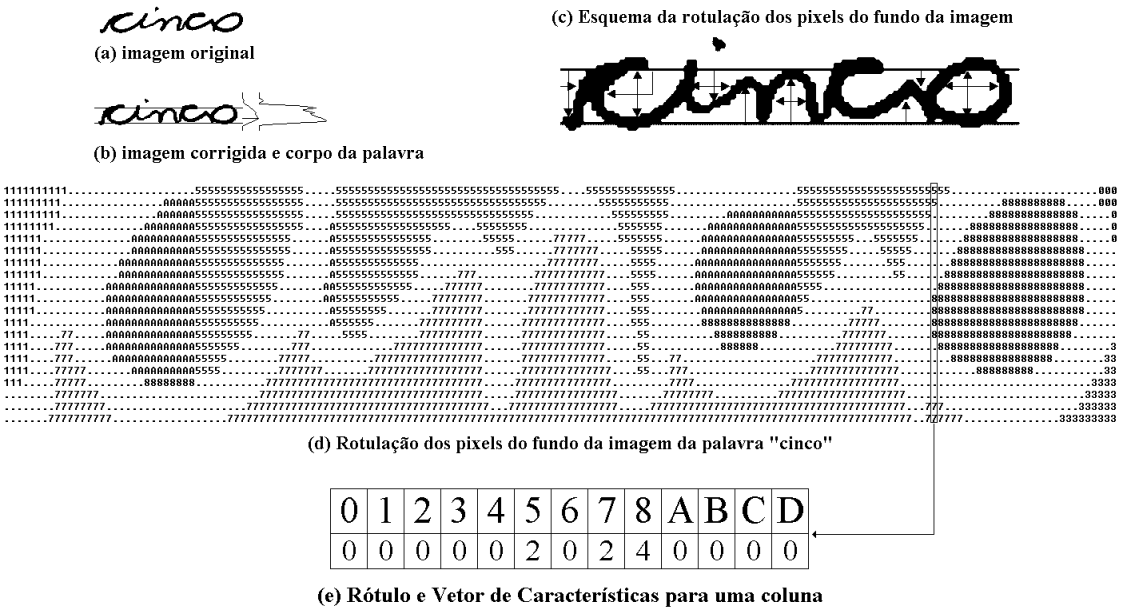

Figura 4.17: Exemplo da rotulação do corpo da palavra para primitivas geométricas

Tabela 4.18: Rótulos do Vetor de Características

Rótulo	Tipo	Busca
0	Aberto à direita e acima	(NSLO)
1	Aberto à esquerda e acima	(NSLO)
2	Aberto à esquerda a abaxo	(NSLO)
3	Aberto à direita e abaixo	(NSLO)
4	Aberto à direita	(NSLO)
5	Aberto acima	(NSLO)
6	Aberto à esquerda	(NSLO)
7	Aberto abaixo	(NSLO)
8	Laço fechado	(NSLO)
A	Escape para direita e acima	Sondagem
B	Escape para esquerda e acima	Sondagem
C	Escape para direita e abaixo	Sondagem
D	Escape para esquerda e abaixo	Sondagem

A distribuição dos rótulos extraídos para a base de dados de treinamento mostram uma representação adequada das formas a reconhecer, conforme Figura 4.18. Observa-se que a palavra “um” é a que possui menor quantidade de pixels rotulados por **8**, visto que a mesma não possui laços fechados no corpo como parte da morfologia básica. Por outro, lado a palavra “dois” é a palavra que mais apresenta laços fechados, sendo a proporção de rótulos **A** a **D** inferior as demais palavras. O rótulo **A** apresenta a

maior proporção (20% a 34%) entre os rótulos **A** a **D**. Este rótulo caracteriza o escape para direita e acima, tal qual as letras **c** da palavra "cinco" na Figura 4.17.

Os rótulos **5** e **7** apresentam uma a maior proporção entre os demais rótulos (de 25% a 45%). Estes rótulos são justamente os que permitem descrever as regularidades das palavras, como especificado na Tabela 4.18.

Assim, os vetores de características passam a ser a entrada de dados para o método de classificação que permitirá estabelecer os símbolos correspondentes a cada coluna do corpo da palavra.

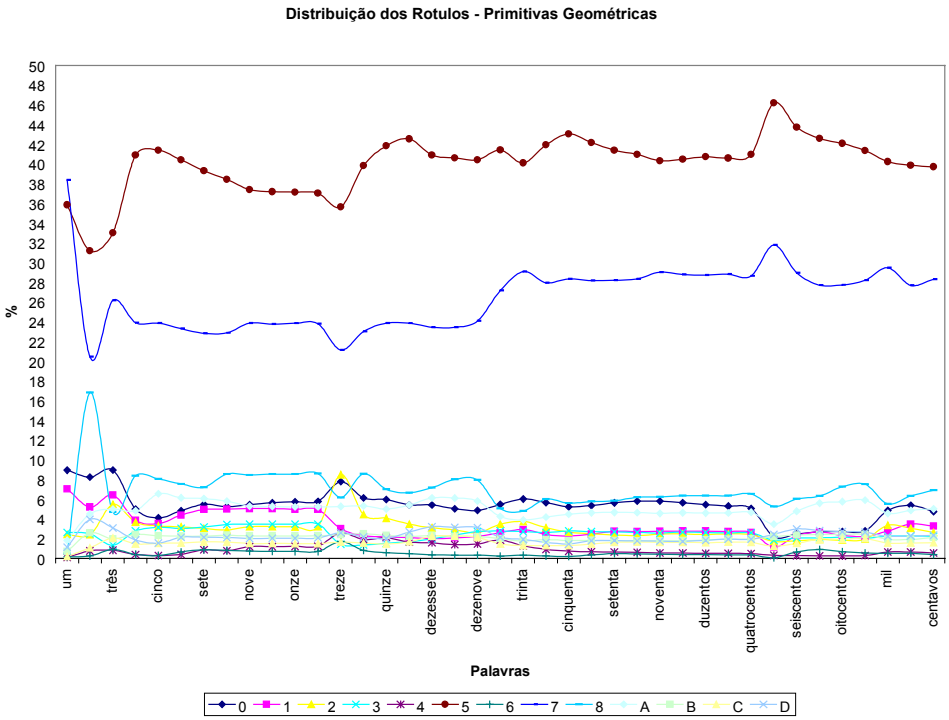

Figura 4.18: Distribuição de freqüência dos rótulos

4.3.2. Método de Classificação: Quantização Vetorial

O Capítulo 2 fornece os subsídios necessários a escolha de um método de classificação visto que para cada coluna da imagem do corpo da palavra deseja-se emitir um símbolo correspondente a descrição das concavidades e convexidades rotuladas anteriormente.

Percebe-se, portanto, uma diferença em relação ao conjunto de primitivas globais denominado de Conjunto 03 – PPCCR – primitivas perceptivas, concavidades e convexidades rotuladas. Para este conjunto efetua-se a classificação juntamente com o processo de extração das primitivas, rotulando-se cada pseudo-segmento. Agora, estes

processos estão separados, visando-se uma classificação mais adequada e discriminativa das formas a reconhecer.

O método escolhido é denominado k-Vizinhos Mais Próximos *(k-Nearest-Neighbor* - kNN) que utiliza o conceito de classificação através de funções de distância e caracteriza-se por ser um método não-paramétrico (Capítulo 2). A motivação para aplicação deste método está no estabelecimento de uma medida de similaridade entre vetores de características, considerando-se os vetores em um espaço Euclidiano e determinando-se a proximidade entre estes vetores.

O método dos k-Vizinhos Mais Próximos trabalha calculando as distâncias entre uma entrada de dados , vetor de características, $x \in V$ de um espaço vetorial e um número de símbolos ou *codebooks* $w_i \in V$. A distância Euclidiana mínima entre o vetor x e os i-ésimos símbolos é dada por [NESCHEN, 1995]:

$$D_i = \min_l \| x - w_i^l \| \quad l = 1,2,...,N_i \tag{4.30}$$

Entende-se por *codebook* o conjunto de símbolos que permitem a classificação de cada vetor de características em um símbolo, o qual possui a menor distância em relação aos centróides dos símbolos que compõem o próprio *codebook*.

O problema da classificação inicia-se pela definição das fronteiras ou *clustering* entre os símbolos do *codebook* e, posteriormente, pela efetiva classificação dos vetores de primitivas nos símbolos do *codebook* correspondentes aos de mínima distância. Usualmente, a decisão é tomada considerando-se os símbolos dos k-vizinhos mais próximos. Esta aproximação é denominada de Quantização Vetorial ou *Vector Quantization* [NESCHEN, 1995].

O problema da definição das fronteiras entre os símbolos do *codebook* pode ser tratado por critérios heurísticos ou por critérios de minimização ou de maximização de parâmetros. As fronteiras são determinadas a partir do momento que se conhece o centróide (x,y) de cada símbolo do *codebook* w_i e a área de influência de cada um no espaço, conforme Figura 4.19.

Para que as regiões não resultem sobrepostas uma pesquisa da "representação ótima" dos centróides dos símbolos deve ser efetuada. Para isto, existem vários algoritmos que auxiliam na definição dos centróides dos símbolos [NESCHEN, 1995]: *Simple Cluster-Seeking Algorithm*, *Maxmin-Distance Algorithm*, *K-Means Algorithm* e *Isodata Algorithm*.

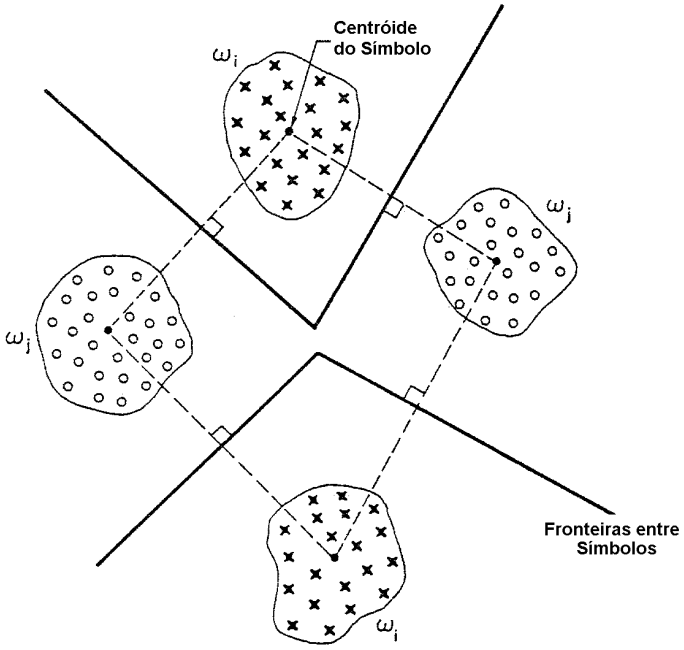

Figura 4.19: Fronteiras para dois símbolos compostas cada qual por duas amostras

O Algoritmo *K-Means* é um procedimento iterativo de classificação que trabalha com a minimização de uma medida de distância dada pela soma dos quadrados das distâncias de todos os vetores de características no domínio da aplicação ao centróide do símbolo. O algoritmo *K-Means* é composto pelos seguintes passos de acordo com [RABINER & JUANG, 1993]:

Passo 1: Escolher K iniciais símbolos com centróides $z_1(1)$, $z_2(1)$,..., $z_k(1)$. Estes centóides são arbitrários e normalmente utiliza-se as primeiras K amostras do conjunto de vetores de características da base de treinamento. Para evitar a associação inicial dos centróides dentro de uma mesma classe de palavras, realiza-se um procedimento randômico para escolha do valor inicial dos centróides, evitando-se também que diferentes centróides recebem valores iguais;

Passo 2: Distribuir as amostras {x} entre o domínio dos K símbolos, utilizando a seguinte relação:

$$x \in S_j(k) \text{ se } \| x - z_j(k) \| < \| x - z_i(k) \| \tag{4.31}$$

Para todo i = 1,2,..,K, i ≠ j, onde $S_j(k)$ denota o conjunto de amostras de vetores de características cujo os centróides dos símbolos são $z_j(k)$.

Passo 3: Com base no resultado do Passo 2, calcular os novos centróides $z_j(k+1)$, j = 1,2,...,K, tal que a soma do quadrado das distâncias de todos os vetores de

características em $S_j(k)$ para os novos centróides dos símbolos seja minimizada. Em outras palavras, os novos centróides $z_j(k+1)$ são calculados até que a taxa de erro seja minimizada. Adota-se um critério de erro < 10^{-3}. O centróide $z_j(k+1)$ que minimiza esta taxa de erro é a média dos $S_j(k)$. Então, o novo centróide é dado por:

$$z_j(k+1) = \frac{1}{N_j} \sum_{x \in Sj(k)} x \qquad j = 1,2,...K \tag{4.32}$$

Onde N_j é o número de amostras de $S_j(k)$. O nome "*K*-means" é obviamente derivado da maneira na qual os centróides dos símbolos são sequencialmente recalculados.

Passo 4: Se $z_j(k+1) = z_j(k)$ para j = 1,2,...,*K*, o algoritmo converge e o procedimento de definição dos centróides dos símbolos pode ser finalizado. Caso contrário, deve-se retornar ao Passo 2.

O comportamento do algoritmo *K-Means* é influenciado pelo número de centróides especificados, pela escolha inicial dos valores dos centróides, pela ordem em que estes são escolhidos e, também, pelas propriedades geométricas dos dados, ou melhor, das palavras. Na prática, realiza-se diversos experimentos com diferentes valores de *K*.

A definição do melhor conjunto de símbolos somente é possível após a execução dos experimentos de reconhecimento com a obtenção das respectivas taxas e matrizes de confusão. Portanto, apresenta-se a seguir as características das etapas de treinamento-validação e reconhecimento dos modelos MEM.

4.3.3. Reconhecimento das Palavras Manuscritas com MEM

As seqüências de observações que servem de entrada para a etapa de treinamento e validação dos modelos MEM advém da conversão, para cada conjunto de símbolos em teste, de cada vetor de primitivas em um símbolo.

A etapa de treinamento dos modelos segue os mesmos algoritmos já apresentados no Capítulo 2, Baum-Welch e Validação Cruzada. Estes procedimentos geram, portanto, para cada palavra dois modelos: λ_{PE} (perceptivo) e λ_{GE} (geométrico). Inicialmente, compara-se as taxas de reconhecimento isoladamente, para depois compor as hipóteses de reconhecimento através da combinação das probabilidades de reconhecimento da palavra desconhecida por ambos os modelos (log $pr(\lambda_{PE})$ + log $pr(\lambda_{GE})$). Esta estratégia está representada na Figura 4.20.

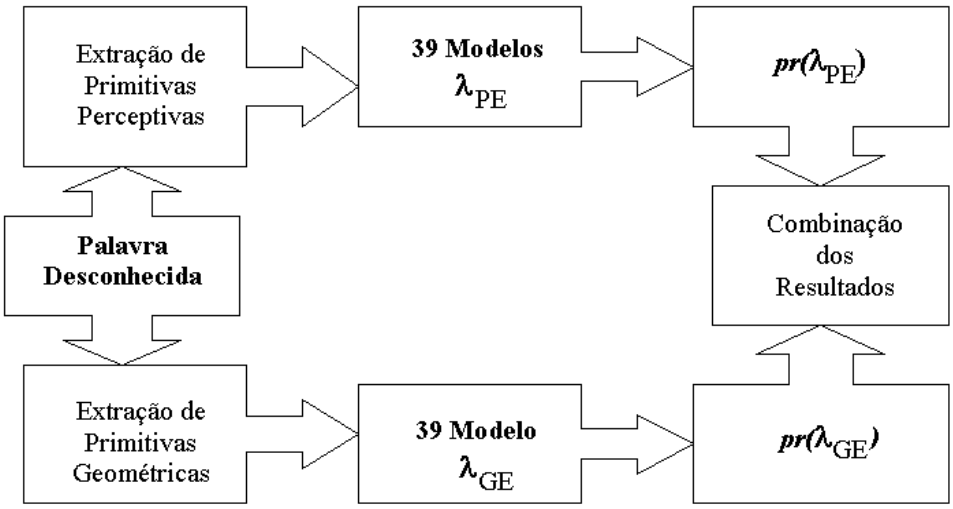

Figura 4.20: Esquema de reconhecimento: primitivas perceptivas e geométricas

A utilização de Múltiplos Conjuntos de Símbolos ou *Multiple Codebooks* para os MEM não é possível, uma vez que os processos de extração das primitivas perceptivas e geométricas são diferentes. Este fato impossibilita a combinação de observações obtidas em instantes *t* diferentes.

O processo paralelo de definição do conjunto de símbolos e obtenção dos modelos está ainda em realização. Porém, alguns problemas podem ser adiantados referentes ao tamanho das seqüências de observações, ao número de símbolos de cada conjunto testado e ao número de amostras para o treinamento.

Deve-se lembrar que para cada coluna da imagem é emitido um símbolo e, portanto, as seqüências de observações tornam-se longas. A média do tamanho das seqüência para as 39 palavras é de 244 colunas, sendo o menor valor 130 para a palavra "dez" e o maior valor 425 para a palavra "quatrocentos". Sabe-se da teoria dos MEM que para seqüências longas os cálculos dos algoritmos *Forward* e *Backward* podem fazer com que a probabilidade tenda a zero. Deste modo, deve-se recorrer ao procedimento de Reescalonamento ou *Scaling* [RABINER & JUANG, 1993].

Sabe-se também, que um conjunto composto por um número pequeno de símbolos não é discriminativo das formas a reconhecer. Por outro lado, um número elevado de símbolos associado com a existência de poucas amostras de treinamento pode fazer com que o modelo não possa ser treinado adequadamente.

Mesmo tendo-se consciência de todos estes prováveis problemas acredita-se que as primitivas geométricas contribuam positivamente no processo de reconhecimento das palavras.

Constata-se que os resultados obtidos neste trabalho para os diferentes experimentos realizados vão de encontro aos objetivos estabelecidos para o reconhecimento de palavras isoladas do extenso manuscrito de cheques bancários brasileiros. Desta forma, apresenta-se no Apêndice B um comparativo dos resultados obtidos por outros trabalhos realizados no mesmo escopo.

A seguir, no Capítulo 5 apresenta-se a conclusão deste trabalho, bem como as possibilidades de trabalhos futuros visando a continuidade das pesquisas na área de reconhecimento de palavras manuscritas na PUCPR.

Capítulo 5

Conclusão

O presente trabalho modela e avalia computacionalmente o reconhecimento de palavras manuscritas no contexto de cheques bancários brasileiros. Para isso, foram consideradas características perceptivas relacionadas à escrita e a leitura, e ao uso de Modelos Escondidos de Makov – MEM para processamento da informação.

Considerando-se todos estes aspectos, as primitivas selecionadas respondem as seguintes questões:

- **Quais são as primitivas perceptivas relevantes no processo de reconhecimento de palavras manuscritas?**
- **Como representar palavras cursivas sem a presença de primitivas perceptivas?**

Assim, explora-se os ascendentes, os descendentes, os laços, as concavidades e as convexidades existentes nas formas a reconhecer. Estuda-se 3 conjuntos de primitivas denominados de PP - primitivas perceptivas, PPCC - primitivas perceptivas, côncavo, convexo e PPCCR - primitivas perceptivas, concavidades e convexidades rotuladas.

Pode-se concluir que as concavidades e convexidades exercem um importante papel no desempenho do método, uma vez que permitem representar palavras cursivas sem a presença de primitivas perceptivas, bem como ressaltar as regularidades existentes nas mesmas. Este fato, torna-se evidente diante da modelagem das palavras cursivas através dos MEM.

Os MEM possuem a capacidade de oferecer um modelo probabilístico ao enfoque estrutural e permitem um treinamento automático dos parâmetros a estimar dos modelos. Além do que, possibilitam modelar eficazmente diferentes fontes de conhecimento, a saber: morfológico – a forma a reconhecer, léxico – o vocabulário de

palavras e sintático – o modelo para cada palavra. Estes fatos vem ao encontro dos aspectos de escrita e leitura apresentados neste trabalho.

Desta forma, o desempenho alcançado com o método desenvolvido é promissor, visto que os resultados obtidos com os 3 conjuntos de primitivas são comparáveis aos obtidos por outros autores (Apêndice B). As taxas de reconhecimento obtidas são 56,4% (PP), 64,8% (PPCC) e 70,6% (PPCCR-FR), respectivamente. Além disto, os principais problemas são apontados pela análise de erros. O presente trabalho não inclui observações ou descrição da performance do método uma vez que não se tem preocupação com o tempo de execução.

As questões abordadas na análise de erros referem-se à similaridade das formas das palavras encontradas na matriz de confusão reconhecidas como intra-classe e inter-classes de palavras, à freqüência das palavras na base de treinamento e ao desempenho dos conjuntos de primitivas selecionados. Todas estas questões foram exploradas apontando os pontos favoráveis e as dificuldades encontradas no processo de reconhecimento.

Sendo assim, os estudos realizados e os resultados obtidos possibilitam que outros trabalhos possam ser realizados, uma vez que o mesmo contribui para aplicação dos Modelos Escondidos de Markov no reconhecimento de palavras manuscritas provenientes de cheques bancários. Além, da contribuição de sua aplicação em problemas de reconhecimento de palavras do idioma português.

Ressalta-se a contribuição do presente trabalho através da criação da base de dados de cheques bancários brasileiros de laboratório.

Para que as bases de dados não tenham proporções infinitas, reduz-se a variabilidade da escrita cursiva através das etapas pré-processamento, a saber: aquisição das imagens, correção da linha de base, correção da inclinação vertical dos caracteres, detecção e normalização do corpo das palavras. Sabe-se que muitos outros procedimentos podem vir a ser desenvolvidos e aplicados visando-se melhorar os resultados obtidos.

Os trabalhos futuros envolvem a continuidade da modelagem das palavras através das primitivas geométricas melhorando a representação das regularidades de uma palavra, tal qual [CÔTE, 1997]. Busca-se uma nova chance de ler o que está escrito com um enfoque mais local. Para [HEUTTE, 1994] a utilização de primitivas geométricas em métodos de reconhecimento de palavras manuscritas explica-se diante

do fato que estas apresentam tolerância às distorções e às variações de estilo e que estas podem descrever as propriedades globais ou locais de palavras e caracteres.

Outro trabalho é o tratamento do extenso completo, incluindo-se a gramática de grafia dos valores numéricos por extenso para o idioma português no processo de reconhecimento. Uma abordagem para o extenso é considerar as palavras "mil", "reais/real", "centavos/centavo" como palavras-chaves no extenso, como apresentado em [FREITAS et al., 2000b], [FREITAS et al., 2000c], [FREITAS et al., 2001].

A modelagem dos prefixos e sufixos das classes das palavras caracteriza-se por outro trabalho a ser realizado, levando-se em consideração as classes da Figura 1.3. Os estudos podem envolver, ainda, a modelagem das palavras de maneira que o reconhecimento seja realizado através da concatenação de modelos, como por exemplo: palavra "novecentos" = modelo "nove" + modelo "entos", palavra "oitenta" = modelo "oito" + modelo "enta" e palavra "dezesseis" = modelo "dez" + modelo "seis".

O presente trabalho coloca-se, ainda, como mais uma contribuição na área de Análise e Reconhecimento de Documentos, podendo vir auxiliar outras pesquisas, tais como: reconhecimento de palavras manuscritas em formulários, projetos e anotações de campo, identificação dos procedimentos de recebimento e despacho de documentos para recuperação e consultas, conversão de desenhos e anotações manuscritas para recuperação ou união com Projetos Auxiliados por Computador – PAC.

Referências Bibliográficas

[AGARWAL et al, 1995] AGARWAL, A.; GRANOWETTER, L.; HUSSEIN, K.; GUPTA, A. *Detection of Courtesy Amount Block on Bank Checks*. Montreal, ICDAR'95,IEEE 1995, p.748-751.

[AGAZZI & KUO, 1993] AGAZZI, O. E. & KUO, S.S. *Hidden Markov models based optical character* recognition *in the presence of deterministic transformations*. Pattern Recognition, Vol.26, No.12, 1993, p.1813-1826.

[AVILA, 1996] ÄVILA, M. *Optimisation de Modeles Markoviens pour la Reconnaissance de L'ecrit*. Thèse de doctorat, Université de Rouen, France, 1994. 215p.

[BANCO DO BRASIL, 1983] Banco Central do Brasil. *Manual e Normas de Instruções*. Brasília, DF. 1983.

[BATITTI, 1994] BATITTI, R. *Using Mutual Information for Selecting Features in Supervised Neural Net Learning*. IEEE Transactions on Neural Networks, Vol.5, no.4,1994, p.537-550.

[BERCU & LORETTE, 1993] BERCU, S.; LORETTE, G. *On-line handwritten word recognition: an approach based on hidden markov models.* IWFHR-3, 1993, p.385-390.

[BERTILLE et al, 1994] BERTILLE, J.M.; GILLOUX, M.; YACOUBI, A. *Localisation et reconnaissance conjointes de noms de voies dans les lignes distribution des adresses postales.* SRTP/RD/ Treitement automatique ligne distribution. TRANSITION No. 7, 1994, p.16-25.

[BOZINOVIC & SRIHARI, 1989] BOZINOVIC, R.M. & SRIHARI, S.N. *Off-line cursive script recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.11, No. 1, 1989, p.68-83.

[BUNKE et al, 1995] BUNKE, H.; ROTH, M.; SCHUKAT-TALAMAZZINI, E.G. *Off-line cursive handwriting recognition using hidden markov models.* Pattern Recognition, Vol.28, No. 9, 1995, p.1399-1413.

[BURR, 1983] BURR, D.J. *Designing a handwriting reader.* IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-5(5). 1983, p.554-559.

[BUSE et al., 1997] BUSE,R.; LIU,Z.; CAELLI, T. *A structural and relational approach to handwritten word recognition*. IEEE Transations on Systems, Man. *a*nd Cybernetics, Part B, Cybernetics, Vol.27, No. 5, 1997.

[CAGLIARI, 1991] CAGLIARI, L.C. *Alfabetização & lingüística*. São Paulo: Editora Scipione. 1991, p.96-166.

[CAVANAGH ,1974] CAVANAGH, J.A. *Lettering and alphabets. 85 complete alphabets*. New York: Dover Publications, Inc., 1974. 121p.

[CHEN & KUNDU, 1993] CHEN, M.Y. & KUNDU, A. *An alternative to variable duration HMM in handwritten word recognition.* IWFHR'93. 1993, p.82-91.

[CHEN et al, 1995a] CHEN, S.; HARALICK, R.M.; PHILLIPS, I.T. *Simultaneous Word Segmentation from Document Images Using Recursive Morphological Closing Transform*. Montreal, ICDAR'95,IEEE 1995, p.761-764.

[CHEN et al, 1995b] CHEN, M.Y. & KUNDU, A.; SRIHARI, S.N. *Variable duration hidden* Markov *model and morphological segmentation for handwritten word recognition*. IEEE, Vol.4, No.12, 1995, p.1675-1686.

[CHO et al., 1995] CHO, W.; LEE, S.W.; KIM, J.H. *Modeling and recognition of cursive words with hidden Markov models*. Pattern Recognition, Vol.28, No.12, 1995, p.1941-1953.

[CONNELL, 1996] CONNELL, S. *A comparison of hidden markov model features for the recognition of cursive handwriting*. Master Dissertation, Michigan State University, USA, 1996. 154p.

[CONGEDO et al., 1995] CONGEDO, G.; DIMAURO, G.; IMPEDOVO, S.; PIRLO, G. A. *Structural* Method *with Local Refinig for Handwritten Character Recognition*. Montreal, ICDAR'95, IEEE 1995, p.853-856.

[CÔTE, 1997] CÔTÉ, M. *Utilisation d'un modèle d'accès lexical et de concepts perceptifs pour la reconnaissance d'images de mots cursifs*. Thèse de doctorat, École Nationale Supérieure des Télécommunications, France,1997. 171p.

[COVER & THOMAS, 1991] COVER, T.M., THOMAS, J.A. *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.

[DOLFING, 1998] DOLFING, J.G.A. *Handwritten recognition and verification – a hidden markov approach*. 1998.

[DZUBA et al., 1998] DZUBA, G.; FILATOV, A.; GERSHUNY, D.; KIL, I. *Handwritten word recognition – the approach proved by practice.* IWFHR´98, Taejon, Korea, august, 1998, p.99-122.

[EL YACOUBI et al., 1994] EL YACOUBI,A.; BERTILLE, J.M.; GILLOUX, M. *Towards a more effective handwritten word recognition system.* IWFRH'94, 1994, p.378-385.

[EL YACOUBI, 1996] EL YACOUBI, A . *Modélisation Markovienne de l'écriture manuscrite. Application à la reconnaissance des adresses postales.* Thèse de doctorat, Universite de Rennes, France,1996. 307p.

[EL YACOUBI et al., 1999] EL YACOUBI, A., GILLOUX, M., SABOURIN, R. SUEN, C.Y. *Unconstrained handwritten word recognition using hidden markov models*. IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol.2, no.8,1999, p.752-760.

[FACON, 1995] FACON, J. *A morfologia matemática através de exemplos.* 2ª. SBAI-Simpósio Brasileiro de Automação Inteligente. CEFET-PR. Curitiba, set. 1995.

[FARAG, 1979] FARAG, R. *Word*-level *recognition of cursive script.* IEEE Transactions on Computers. C-28(2). 1979, p.172-175.

[FAVATA & SRIHARI, 1992] FAVATA, J. & SRIHARI, S. *Recognition of general handwritten words using a hypothesis generation and reduction methodology.* Proc. United State Postal Service Advanced Technology Conference. Washington DC, nov. 1992, p.237-253.

[FILATOV et al, 1995] FILATOV, A.; GITIS, A.; KIL, I. *Graph-based Handwritten Digit String Recognition*. Montreal, ICDAR'95, IEEE 1995, p.845-848.

[FREITAS et al, 2000a] FREITAS, C. O. A.; MORITA, M.; SOARES, L. E. O.; JUSTINO, E.; EL YACOUBI, A.; LETHELIER, E.; BORTOLOZZI, F.; SABOURIN, R. *Brazilian Bank Check Databases*. XXVI Conferência Latino-Americana de Informática. CLEI'2000, Mexico, 2000, p.134-152.

[FREITAS et al., 2000b] FREITAS, C. O. A.; EL YACOUBI, A.; BORTOLOZZI, F.; SABOURIN, R. *Brazilian Bank Check Handwritten Legal Amount Recognition*. 13° Brazilian Symposium on Computer Graphics and Image Processing. SIBGRAPI2000, Brasil, 2000, p.97-104.

[FREITAS et al., 2000c] FREITAS, C. O. A.; EL YACOUBI, A.; BORTOLOZZI, F.; SABOURIN, R. *Isolated Word Recognition in Brazilian Bank Check Legal Amounts*. 4° Workshop on Document Analysis and Systems. DAS2000, Brasil, 2000, p.279-290.

[FREITAS et al., 2001] FREITAS, C. O. A.; BORTOLOZZI, F.; SABOURIN, R. *Handwritten word recognition: an approach based on mutual information for feature set validation*. Accepted to Sixth International Conference on Document Analysis and Recognition, ICDAR`2001, september, Seattle, Washington, USA, 2001.

[GILLIES, 1992] GILLIES, A.M. Cursive word recognition using hidden markov models. Proc. of 5th USPS Advanced Technology Conference, 1992, p.557-562.

[GILLOUX et al., 1993] GILLOUX, M.; LEROUX, M.; BERTILLE, J.M. *Strategies for handwritten words recognition using hidden markov models*. Proc. of the Second International Conf. On Document Analysis and Recognition. ICDAR'93, 1993, p.299-304.

[GILLOUX, 1994] GILLOUX, M. *Hidden markov models in handwriting recognition.* IWFHR'94, 1994, p.264-288.

[GILLOUX et al, 1995] GILLOUX, M.; LEROUX, M.; BERTILLE, J.M. *Strategies for cursive script recognition using hidden markov models*. Machine Vision and Applications, Springer-Verlog, Vol.8, 1995, p.197-205.

[GRANDIDIER et al., 2000] GRANDIDIER, F., SABOURIN, R., SUEN, C.Y., GILLOUX, M. *Une nouvelle stratégie pour l´amélioration des jeux de primitives d´un système de reconnaissance de l´écriture*. CIFED2000, França, may, 2000, p.210-228.

[GUILLEVIC, 1995] GUILLEVIC D. *Unconstrained Handwriting Recognition Applied to the Processing of Bank Cheques*. Thesis of Doctor's Degree in the Department of Computer Science at Concordia Uiversity.Canada.1995. 183p.

[HAN & SETHI, 1997] HAN, K. & SETHI, I.K. *An off-line cursive handwritten word recognition system and its application to legal amount interpretation*. International Journal of Pattern Recognition and Artificial Intelligence. Vol.11, No. 5, 1997, p.757-770.

[HARMON, 1962] HARMON, L.D. *Handwriting reader recognizes whole words.* Eletronics, 35. 1962, p.29-31.

[HEUTTE, 1994] HEUTTE, L. *Reconnaissance de caractères manuscrits: application à la lecture automatique des chèques et des enveloppes postales*. Docteur Thèse de L'Université de Rouen. Rouen, France, dec., 1994. 239p.

[HEUTTE et al., 1998] HEUTTE, L.; PAQUET, T.; MOREAU, J.V.; LECOUTIER, Y.; OLIVER, C. *A strutural / statistical feature based vector for handwritten character recognition.* Pattern Recognition Letters 19, 1998, p.629-641.

[HULL & FERINCH, 1993] HULL,J.J. & FERINCH, R.K. *Large database organization for document images*, In: Fundamentals in Handwriting Recognition, Edited by Sabastiano Impedovo, Nato ASI Series, Vol. 124, 1993, p.397-414.

[HULL, 1994] HULL, J.J. *A database for handwritten text recognition research,* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.16, No. 5, 1994, p.550-554.

[HULL, 1988] HULL, J.J. *A computational theory of visual word recognition.* PhD Thesis, University of New York at Buffalo, USA,1988.

[HUANG et al., 1990] HUANG, X.D.; ARIKI, Y.; JACK, M.A. Hidden Markov models for speech recognition. Information Technology Series; 7. Edinburgh University Press. 1990, 276p.

[IMPEDOVO et al., 1997] IMPEDOVO, S.; WANG, P.S.P.; BUNKE, H. *Automatic bankcheck processing*. Series in Machine Perception and Artificial Intelligence. World Scientific Publishing Co. Pte. Ltd., 1997, 381p.

[KATO, 1990] KATO,M. *O* aprendizado *da leitura*. São Paulo: Martins Fontes, 1990, 121p.

[KIM, 1996] KIM, G. *Recognition of offline handwritten words and its extension to phrase* recognition. PhD Thesis, University of New York at Buffalo, USA, mar., 1996. 138p.

[KIM & GOVINDARAJU, 1995] KIM, G.; GOVINDARAJU, V. *Handwritten word recognition for real-time applications.* International Conference on Document Analysis and Recognition (ICDAR), Vol.1, 1995, p.24-27.

[KIM & GOVINDARAJU, 1997] KIM, G.; GOVINDARAJU, V. *A lexicon driven approach to* handwritten *word recognition for real-time applications.* IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.9, No.4, 1997.

[KIMURA et al., 1993] KIMURA, F.; SHRIDHAR, M.; CHEN, Z. *Improvements of a lexicon directed algorithm for recognition of unconstrained handwritten words*. International Conference on Document Analysis and Recognition. ICDAR'95. vol 1, aug., 1995, p.24-27.

[KNERR et al., 1996] KNERR, S.; ANISIMOV, V.; BARET, O.; PRICE, D. & SIMON, J.C. *The A2iA recognition system for handwritten checks*. A2iA Technical Report, No.11, 1996. 49p.

[KOERICH, 1997] KOERICH, A . L. *Uma proposta para o processamento de imagens de cheques bancários: extração e armazenamento de informações*. Dissertação de mestrado na Universidade Estadual de Campinas. Campinas, SP, fev., 1997, 190p.

[KOSKINEN et al, 1994] KOSKINEN, L.; HUTTUNEN, H.; ASTOLA, J. *Text enhacement method based on soft morphological filters.* SPIE - Document Recognition, Vol. 2181, 1994, p.243-253.

[LECOLINET, 1990] LECOLINET, E. *Segmentation d'images de mots manuscrits: application à la lecture de chaînes de caractères majuscules alphanumériques et à la lecture de l`écriture cursive*. Thèse de doctorat, Universite Pierre et Marie Curie (Paris VI), France. 1990. 283p.

[LEPLUMEY et al., 1995] LEPLUMEY, I.; CAMILLERAPP, J.; QUEGUINER, C. *Kalman Filter Contributions Towards Document Segmentation*. Montreal, ICDAR'95,IEEE, 1995, p.765-769.

[LEROUX et al., 1991] LEROUX, M.; SALOME, J.C.; BADARD, J. *Strategie de controle pour la reconnaissance du montant litteral des cheques postaux.* 1991, p.695-706.

[LI & YEUNG, 1997] LI, X. & YEUNG, D.Y. *On-line hadwritten alphanumeric character recognition using dominant points in strokes.* Pattern Recognition, Vol. 30, No. 1, 1997, p.31-44.

[LIU & SRIHARI, 1994] LIU, Y.; SRIHARI, S.N. *Document image binarization based on texture analysis.* SPIE, Vol. 2181, 1994.

[LU & SHRIDHAR, 1996] LU, Y. & SHRIDHAR, M. *Character segmentation in handwritten words – an overview.* Pattern Recognition, Vol. 29, No. 1, 1996, p.77-96.

[MACHADO & TEIXEIRA, 1996] MACHADO, A.M.B. & TEIXEIRA, E.A. *Aprendendo leitura dinâmica*. São Paulo: Makron *Books*, 1996, 87p.

[MADHVANATH & GOVINDARAJU, 1992] MADHVANATH,S.; GOVINDARAJU, V. *Using holistic features in handwritten word recognition.* Proc. United State Postal Service Advanced Technology Conference. Washington DC, Vol.1, nov. 1992, p.183-199.

[MADHVANATH & GOVINDARAJU, 1998] MADHAVANATH, S. & GOVINDARAJU, V. *Preceptual features for off-line handwritten word recognition: a framework for prediction, representation and matching.* Advances in Pattern Recognition, august, 1998, p.524-531.

[MARTI & BUNKE, 1998] MARTI, U.V. & BUNKE, H. Towards general cursive script recognition. Proc. of Sixth International Workshop on Frontiers in handwriting Recognition - IWFHR, Korea, 1998, p.379-388.

[MOHAMED, 1995] MOHAMED, M.A. *Handwritten word recognition using generalized hidden markov models*. PhD Thesis, University of Missouri-Columbia, USA, 1995.

[MONTOLIU, 1995] MONTOLIU, L. *Architecture multi-agents et réseaux connexionnistes. Application à la lecture de chèques manuscrits*. Thèse de doctorat, Ecole Polytechnique, France, 1995. 233p.

[MOREAU et al., 1991] MOREAU, J.V.; PLESSIS, B.; BOUGEOIS, J.L.P. *A postal check reading system*. Proc. of ICDAR´91, France, 1991, p.758-766.

[MORITA, 1998] MORITA, M.E. *Estudo para a melhoria da correção da inclinação da linha de base de palavras manuscritas.* Dissertação de Mestrado, Curitiba: CPGEI/CEFET-PR, 1998, 137p.

[NESCHEN, 1995] NESCHEN, M. *Vector quantization classifiers for handwritten character recognition.* Proc. of the ZEUS-95 Worshop, Sweden, may, 1995, p.1-14.

[OLIVEIRA, 1998] OLIVEIRA, L.E.S. *Estudo sobre a extração das estruturas lógica e física a partir de imagens de cheques bancários brasileiros.* Dissertação de Mestrado, Curitiba: CPGEI/CEFET-PR, 1998, 120p.

[OLLIVIER, 1999] OLLIVIER, D. *Une appproche économisant les traitements pour reconnâitre l'écriture manuscrite: application à la reconnaissance des montants littéraux de chéques bancaires.* Thèse de doctorat, Université de Paris XI Orsay, France, 1999. 251p.

[OTSU, 1979] OTSU, N. *A threshold selection method from gray-level histograms.* IEEE Transations Systems, Man. and Cybernetics. SMC 9, Vol.1, 1979, p.63-66.

[PAQUET & LECOURTIER, 1993] PAQUET, T. & LECOURTIER, Y. *Automatic reading of the literal amount of bank checks*. Machine Vision and Application, No.3, 1993, p.151-162.

[PARKER, 1997] PARKER, J.R. *Algorithms for Image Processing and Computer Vision*. Ed. John Wiley & Sons, Inc. 1997, p.310-315.

[PARRAMÓN, 1991] PARRAMÓN,J.M. *Como dibujar letras y logotipos*. Barcelona: Parramón Ediciones, 1991, 112p.

[RABINER & JUANG, 1993] RABINER, L. JUANG, B.H. *Fundamentals of speech recognition. P*rentice Hall Inc., London, UK. 1993, 506p.

[ROSA, 1991] ROSA, V.S. *Letras & cartazes*. Porto Alegre: Mercado Aberto, 1991, 79p.

[SANTOS, 1997] SANTOS, J.E.B. *Estudo sobre métodos e técnicas para a segmentação de imagens de cheques bancários.* Dissertação de Mestrado, Curitiba: CPGEI/CEFET-PR, 1997, 74p.

[SENIOR & ROBINSON, 1998] SENIOR. A.W. & ROBINSON, A.J. *An off-line cursive handwritting recognition system.* IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.20, No.3, mar.1998, p.309-321.

[SAION & BELAÏD, 1997] SAION, G. & BELAÏD, A. High performance unconstrained word recognition system combining HMMs and markov Random Fields. International Journal of Pattern Recognition and Artificial Intelligence, Vol. 11, No. 5, 1997, p.771-787.

[SCHOMAKER & SEGERS, 1998] SCHOMAKER, L. & SEGERS, E. *A method for the determination of features used in human reading of cursive handwriting.* IWFHR`98, 1998, p.157-168.

[SOLO, 1989] SOLO, D.X. *Bold script alphabets. 100 complete fonts*. New York: Dover Publications Inc., 1989, 100p.

[STRATHY, 1993] STRATHY, N.W. *A method for segmentation of touching handwritten numerals*. Master's thesis, Concordia University, Montreal-Canada, 1993, 120p.

[SUEN & LI, 1997] SUEN,C.Y. & LI, Z.C. *Crucial combinations for the recognition of handwritten letters.* BSDIA'97, Curitiba-PR, nov. 1997, p.19-39.

[SUEN, 1998] SUEN, C.Y. *Réflexions sur la reconnaissance d´écriture cursive*. CIFED´98. Quebec, Canada, may, 1998, p.1-8.

[TAPPERT et al., 1990] TAPPERT, C.C.; SUEN, C.Y.; WAKAHARA, T. *The state of art in on-line hadwritting recognition.* IEEE Trans. on Pattern Analysis and Machine Intelligence. aug. 1990, p.787-808.

[TEBEROSKY & TOLCHINSKY, 1996] TEBEROSKY, A. & TOLCHINSKY, L. *Além da alfabetização. a aprendizagem fonológica, ortográfica, textual e matemática*. trad: Stela Oliveira. São Paulo: Editora Ática, 1996, 295p.

[TREVISAN, 1992] TREVISAN, E.M.C. *Leitura: coerência e conhecimento prévio*. Editora da UFSM, Santa Maria. 1992, 117p.

[TRIER et al., 1996] TRIER, O.D.; JAIN, A.K.; TAXT, T. *Feature extraction methods for character recognition - a survey.* Pattern Recognition, Vol.29, No. 4, 1996, p.641-662.

[VIARD-GAUDIN et al., 1999] VIARD-GAUDIN,C.; LALLICAN, P.M.; KNERR, S.; BINTER, P. *The IRESTE On/OFF (IRONOFF) dual handwriting database.* International Conference on Document Analysis and Recognition. ICDAR'99, 1999, p.455-458.

Apêndice A

Estrutura dos Cheques Bancários Brasileiros

O cheque bancário brasileiro é um instrumento que os bancos colocam a disposição dos clientes para movimentar fundos em seu poder e efetuar pagamentos sem a necessidade do manuseio do papel moeda. Um cheque representa uma ordem incondicional em direito que: é emitida por uma pessoa em favor de outra; é assinado pela pessoa que emite; e exige que o banco pague na apresentação uma determinada quantia em dinheiro [BANCO DO BRASIL, 1983].

Este Apêndice apresenta a estrutura de um cheque, suas características em relação ao conteúdo e a disposição dos seus elementos componentes.

A.1. Modelos para Cheques: Distribuição das Informações e Preenchimento dos Campos

O Conselho Monetário Nacional instituiu um modelo único de cheque visando uma maior segurança e eficiência nos serviços que os bancos prestam à coletividade. A Figura A1 mostra a localização dos elementos dentro da área de um cheque bancário. A Figura A2 apresenta o modelo padrão de preenchimento dos campos e áreas do cheque, segundo o padrão estabelecido pelas normas do Banco Central do Brasil.

A.2. Dimensões do Cheque

A resolução Nº 855 editada pelo Banco Central [BAN83] estabelece que os cheques devem ter como dimensões:

- comprimento de 175 mm, com tolerância de mais ou menos 1mm;

- largura de 80 mm, com tolerância de 1mm para mais e de até 4 mm para menos.

Desta forma, um cheque digitalizado em 300 dpi pode assumir uma dimensão de no mínimo 2057 pixels de comprimento e 897 pixels de largura; e uma dimensão de no máximo 2079 pixels de comprimento e 957 pixels de largura, como representado na Figura A3.

Número da Compensação, Número do Banco, Número da Agência, Número da Conta, Número do Cheque e Dígitos Verificadores	Valor Numérico
Valor por Extenso	
Nome do Favorecido	
Logotipo do Banco	Local e Data de Emissão
Nome do Banco (Nome, Número, Endereço e Telefone da Agência)	Assinatura, Nome do Cliente, Número do CPP/CGC e outras informações
Linha MICR CMC-7	

Figura A1: Modelo para distribuição das informações [KOERICH, 1997]

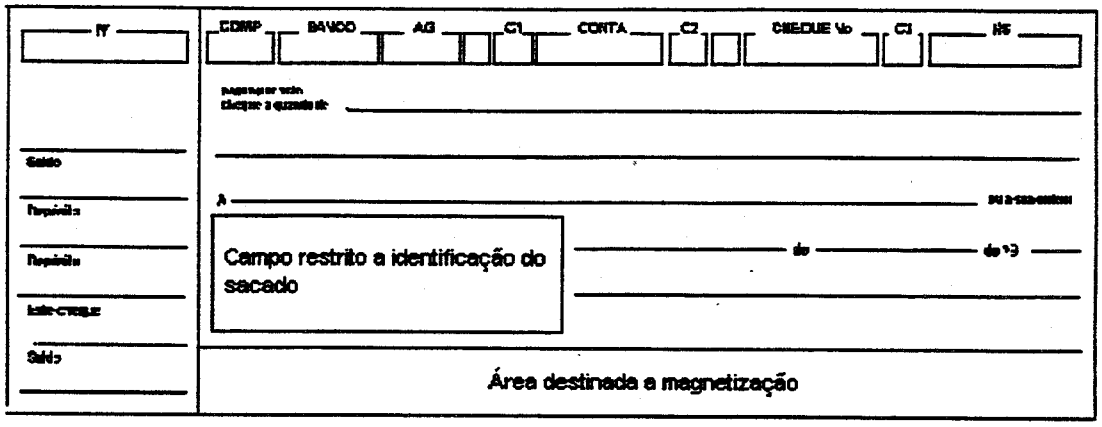

Figura A2: Modelo padrão de preenchimento dos campos e áreas de um cheque [OLIVEIRA, 1998]

A.3. Elementos Componentes

Os cheques podem ser divididos em quatro faixas compreendendo os dados descritos como segue e apresentadas na Figura A4:

a) Primeira Faixa:

- número da Câmara de Compensação (COMP);
- número da inscrição do Banco no Serviço de Compensação (BANCO);
- código da agência sacada (AG);
- posição facultativa: dígito verificador correspondente ao código da agência sacada();

- dígito verificador correspondente aos campos COMP, BANCO e AG (C1);
- número da conta do emitente (CONTA);
- dígito verificador relativo ao número da conta (C2);
- posição facultativa: registro da série de numeração do cheque ();
- número do cheque;
- dígito verificador do número do cheque (C3);
- valor do cheque em algarismos (R$);

b) Segunda Faixa:

- valor do cheque por extenso e
- nome do beneficiário;

c) Terceira Faixa:

- identificação do banco e agência sacados;
- local e data de emissão do cheque;
- assinatura e identificação do emitente;

d) Quarta Faixa:

- impressão de caracteres magnéticos, ou linha MICR (*Magnetic Ink Character Recognition)*.

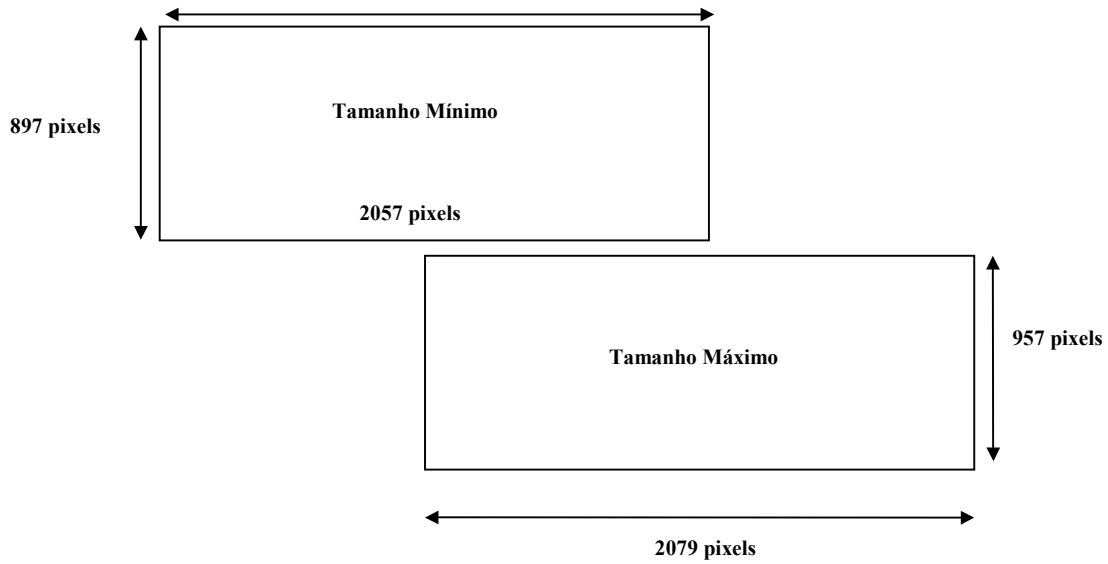

Figura A3: Dimensões mínima e máxima de cheques bancários

A.4. Formulários dos Cheques

Os formulários dos cheques podem ser impressos em qualquer cor, com tinta indelével e de tonalidade firme, sendo permitido o uso de impressões por perfuração ou filigrama, desde que não estejam localizados na faixa de magnetização e não produzam mossas ou rebarbas que prejudiquem a leitura mecânica do cheque.

<table>
<tr><td colspan="2">Número da Compensação, Número do Banco, Número da Agência, Número da Conta, Número do Cheque e Dígitos Verificadores</td><td>Valor Numérico</td></tr>
<tr><td colspan="3">Valor por Extenso

Nome do Favorecido</td></tr>
<tr><td>Logotipo do Banco

Nome do Banco (Nome, Número, Endereço e Telefone da Agência)</td><td colspan="2">Local e Data de Emissão

Assinatura, Nome do Cliente, Número do CPP/CGC e outras informações</td></tr>
<tr><td colspan="3">Linha MICR CMC-7</td></tr>
</table>

Figura A4: Faixas de um cheque bancário brasileiro [OLIVEIRA, 1998]

Os cheques podem ser confeccionados em folhas planas ou em formulários contínuos, devendo, nesse último caso, serem destacados por guilhotina de modo que se obtenham bordas retas e lisas nas extremidades inferior e direita, utilizadas como referência no processo de leitura mecânica. A confecção dos cheques bancários brasileiros devem seguir rigorosamente as especificações do documento "Modelo-Padrão do Cheque", constante do CADOC (Cadastro de Documentos) como modelo nº. 38058-0 (Res. 885), do Banco Central do Brasil.

Os cheques bancários brasileiros apresentam grande diversidade de cores, desenhos e estilos. A estrutura de *layout* de um cheque é considerada complexa, uma vez que caracteriza-se pela presença de desenhos e figuras ilustrativas que fazem parte do padrão de fundo do cheque. O padrão de fundo do cheque pode assumir várias cores e padrões, porém, os elementos que são impressos sobre o fundo, como linhas e seqüência de caracteres, apresentam um posicionamento ordenado, sendo impressos em uma cor única, geralmente, em preto.

Alguns trabalhos podem ser citados no que se refere ao estudo da estrutura do cheque bancário brasileiro [OLIVEIRA, 1998], bem como na extração das informações pré-impressas e padrão de fundo [KOERICH, 1997].

A.5. Características do Valor Numérico por Extenso no Cheque

A apresentação do valor numérico por extenso em um cheque bancário brasileiro é obrigatória [BANCO DO BRASIL, 1993]. Esta informação corresponde a segunda faixa de um cheque e deve ser apresentada logo após a expressão "Pague por este cheque a quantia de". Seguem, então, duas linhas reservadas para a especificação do valor numérico do cheque por extenso, não se admitindo a impressão de quaisquer outras palavras ou símbolos. Existe, ainda nesta faixa, uma terceira linha com a preposição "a" e terminada com a expressão "à ordem", ou se for o caso, "não à ordem", reservada para indicação do nome do beneficiário do cheque. Caso, não seja indicado o nome do beneficiário, o cheque será considerado ao portador.

É facultativo na emissão de um cheque, segundo Circular nº. 1825 (Artigo 1º.) do Banco Central do Brasil, a grafia por extenso dos valores correspondentes aos centavos, sendo entretanto, obrigatória a especificação, no campo próprio do formulário de cheque, dos respectivos centavos, em algarismos.

Ressalta-se o Artigo 12 da Lei nº. 7357, no que se refere ao fato de feita a indicação da quantia em algarismos e por extenso, prevalece esta no caso de divergência. E ainda, que indicada a quantia mais de uma vez, quer por extenso, quer por algarismos, prevalece, no caso de divergência, a indicação da menor quantia.

Apêndice B

Comparativo de Métodos de Reconhecimento de Palavras Manuscritas

A comparação de resultados não é uma tarefa simples, porém é de suma importância para posicionar o trabalho realizado diante da comunidade científica que pesquisa na área de reconhecimento de palavras manuscritas em cheques bancários.

A dificuldade na realização da comparação de resultados está, em primeiro plano, no fato que não se dispõe de trabalhos em idioma português, somente inglês e francês. Em segundo plano, no fato de que os trabalhos são realizados com bases de dados diferentes, sendo a comparação de difícil realização até mesmo entre trabalhos sobre a mesma base de dados. Quando os métodos de reconhecimento são os mesmos, por exemplo HMM, existem diferenças no treinamento, nos algoritmos aplicados, na topologia dos modelos, nos critérios de aceitação e rejeição, entre outros. Existe, portanto, uma grande variedade de parâmetros nas condições dos testes de reconhecimento realizados por diferentes autores.

O presente Capítulo apresenta os resultados obtidos com o reconhecimento de palavras isoladas do extenso manuscrito de cheques bancários brasileiros e realiza uma comparação com outros trabalhos no mesmo contexto.

A seleção de outros trabalhos foi realizada com base no contexto de cheques bancários, na abordagem Global adotada para o reconhecimento das palavras, nas primitivas extraídas e nos métodos utilizados para o reconhecimento. Apresenta-se, inicialmente, as características das bases de dados sobre as quais os trabalhos foram realizados, os procedimentos de pré-processamento aplicados, as primitivas extraídas, os métodos de treinamento e reconhecimento adotados e os resultados obtidos.

Os trabalhos selecionados para efetuar-se a comparação dos resultados já foram muitas vezes mencionados no presente trabalho e são os seguintes:

[GUILLEVIC, 1995]: Unconstrained Handwriting Recognition Applied to the Processing of Bank Checks, Tese de Doutorado, Canadá, 1995,

[AVILA, 1996]: Optimisation de Modèles Markoviens pour la Reconnaissance de L'ecrit, Tese de Doutorado, França, 1996,

[CÔTE, 1997]: Utilisation d'un Modèle D'Accès et de Concepts Perceptifs pour la Reconnaissance D'images de Mots Cursifs, Tese de Doutorado, França, 1997,

[OLLIVIER, 1999]: Une Approche Économisant les Traitements pour Reconnaître L'ecriture Manuscrite: Application à la Reconnaissance des Montants Littéraux de Chèques Bancaires, Tese de Doutorado, França, 1999.

B.1. Bases de Dados

O Capítulo 3 reconhece a existência das diversas bases de dados de cheques bancários em nível internacional e nacional e, ainda, justifica a criação da base de dados de laboratório utilizada no presente trabalho. Descreve-se a especificação, a geração dos valores numéricos, o processo de coleta, armazenamento e recuperação das imagens dos extensos bancários.

As principais características das bases de dados utilizadas pelos diferentes autores estão apresentadas de forma resumida na Tabela B.1. Divide-se a apresentação desta Tabela em três partes: a) para as características das bases de dados de palavras em idioma inglês [GUILLEVIC, 1995] e [CÔTE, 1997], b) para as palavras em idioma francês [GUILLEVIC, 1995], [AVILA, 1996] e [OLLIVIER, 1999] e c) para as palavras em idioma português [FREITAS et al, 2000a].

O número reduzido de exemplares da base de treinamento, 184 imagens, no trabalho realizado por [CÔTE, 1997] deve-se ao fato de que a autora realizou o treinamento das primitivas extraídas das palavras com uma rede neural distribuída, portanto, não foi necessário realizar o treinamento dos pesos da rede neural, reduzindo-se assim o número de exemplares para treinamento da rede neural.

Tabela B.1-a: Características das Bases de Dados - Palavras do Idioma Inglês

BD Inglês	**Côte**	**Guillevic**
Origem da base de dados	CENPARMI	CENPARMI
N°. de cheques	2.500	2.500
N°. de palavras	7.837	7.837
N°. de escritores	800	800
Base de treinamento	184	5.322
Base de validação	Não utiliza	Não utiliza
Base de testes	2.929	2.515
Tamanho do léxico	32	32
Número de letras por palavra	3 a 9	

Tabela B.1-b: Características das Bases de Dados - Palavras do Idioma Francês

BD Francês	**Avila**	**Guillevic**	**Ollivier**
Origem da base de dados	SRTP	CENPARMI	LIX / SRTP
N°. de cheques	Não indicado	1.861	1.705 / 11.066
N°. de palavras	6.771	6.135	16.497 / 44.901
N°. de escritores	Omni-escritor	600	Omni-escritor
Base de treinamento	3.623	1.345	8.113 / 40.400
Base de validação	Não utiliza	Não utiliza	Não utiliza
Base de testes	3.148	516	8.384 / 4.490
Tamanho do léxico	27	25	29
Número de letras por palavra	2 a 9		

Tabela B.1-c: Características das Bases de Dados - Palavras do Idioma Português

BD Português	**Freitas**
Origem da base de dados	PUCPR/PPGIA/LARDOC
N°. de cheques	2.016
N°. de palavras	11.948
N°. de escritores	2.016
Base de treinamento	7.146
Base de validação	2.427
Base de testes	2.375
Tamanho do léxico	39
Número de letras por palavra	2 a 12

Observa-se que o léxico em português apresenta como fator de maior complexidade o seu próprio tamanho (39 palavras). Isto, torna-se evidente através da Tabela B.2 que apresenta os léxicos nos três diferentes idiomas a serem consideradas no presente Capítulo. Nota-se que o léxico em português apresenta uma palavra para cada valor numérico a ser grafado por extenso. Este fato não ocorre com as palavras em inglês para valores entre 200 a 900, pois estes valores são grafados por extenso

utilizando-se uma composição de palavras ("two hundred" – "nine hundred"). Por outro lado, observa-se no léxico em francês que este mesmo tipo de composição ocorre para os valores numéricos entre 17 a 19, 70 a 90, e também, para os valores entre 200 e 900. Em francês, a grafia do valor 90 é expressa como: "quatre vingt dix". Estas observações explicam o fato da diferença entre os tamanhos dos léxicos considerados pelos diferentes autores.

Tabela B.2: Léxico em Português, Inglês e Francês

Valor Numérico	**Português**	**Inglês**	**Francês**
0	-	-	Zéro
1	Um, Hum	One	Un
2	Dois	Two	Deux
3	Três	Three	Trois
4	Quatro	Four	Quatre
5	Cinco	Five	Cinq
6	Seis	Six	Six
7	Sete	Seven	Sept
8	Oito	Eight	Huit
9	Nove	Nine	Neuf
10	Dez	Ten	Dix
11	Onze	Eleven	Onze
12	Doze	Twelve	Douze
13	Treze	Thirteen	Treize
14	Quatorze, Catorze	Fourteen	Quatorze
15	Quinze	Fifteen	Quinze
16	Dezesseis	Sixteen	Seize
17	Dezessete	Seventeen	-
18	Dezoito	Eighteen	-
19	Dezenove	Nineteen	-
20	Vinte	Twenty	Vingt
30	Trinta	Thirty	Trente
40	Quarenta	Forty	Quarante
50	Cinquenta, Cincoenta	Fifty	Cinquante
60	Sessenta	Sixty	Soixante
70	Setenta	Seventy	-
80	Oitenta	Eighty	-
90	Noventa	Ninety	-
100	Cem, Cento	Hundred	Cent
200	Duzentos	-	-
300	Trezentos	-	-
400	Quatrocentos	-	-
500	Quinhentos	-	-
600	Seiscentos	-	-
700	Setecentos	-	-
800	Oitocentos	-	-
900	Novecentos	-	-
1000	Mil	Thousand	Mille
Moeda Corrente	Reais, Real	Dollars	Francs, Frs
Fração	Centavos, Centavo	-	Centimes, Cts
Outras palavras	-	And, Only	et

Deve-se, também, levar em consideração a distribuição das palavras nas bases de dados utilizadas pelos diferentes autores na etapa de treinamento como mostrado na Tabela B.3. Observa-se que as palavras correspondentes aos valores numéricos de 10 a 19 apresentam uma baixa freqüência independentemente da base de dados, com exceção da base de dados utilizada por [CÔTE, 1997]. Verifica-se também, uma alta freqüência das palavras "dollars" – "francs" – "reais / real". Isto, deve-se ao fato da obrigatoriedade da grafia da moeda corrente no país de origem dos cheques.

Na base de dados de treinamento do presente trabalho as palavras "mil", "reais/real" e "centavos/centavo" representam 30,56% do total da base de dados de treinamento. Este fato pode ser explicado, uma vez que, as bases de dados utilizadas contam com extensos que foram grafados com base em valore numéricos que necessitam destas palavras (Capítulo 3).

Deve-se levar em consideração as características inerentes ao presente trabalho descritas no Capítulo 2, referentes aos seguintes fatores: tipo de aquisição das imagens, tipo de escrita, número de escritores, tamanho do léxico, qualidade das imagens, inclinação vertical, traços contíguos/desconectados, ortografia diferenciada e linhas guia e linhas de separação. Estes fatores sofrem influência da população de escritores que geram as amostras para as bases de dados, sobre as quais os métodos são desenvolvidos e testados. Como mostrado no Capítulo 2 a escrita cursiva é diferente para cada pessoa e, ainda, sofre influência do processo de alfabetização utilizado no país de origem.

Com base nestas características apresenta-se a seguir os procedimentos de pré-processamento aplicados nas imagens das palavras pelos diferentes autores.

B.2. Pré-Processamento das Palavras

Entende-se por pré-processamento as etapas que antecedem o reconhecimento de palavras manuscritas, ou ainda, as etapas que preparam as imagens para a extração de primitivas. Portanto, descreve-se as técnicas adotadas e/ou desenvolvidas para realização das etapas necessárias ao objetivo pretendido, estando organizado pelos seguintes itens: Procedimentos preliminares, Correção da linha de base, Correção da inclinação vertical dos caracteres e Normalização do corpo das palavras.

Tabela B.3: Distribuição das Palavras nas Bases de Dados de Treinamento

Léxico Inglês	**Côte**	**Guillevic**	**Léxico Francês**	**Avila**	**Ollivier**	**Léxico Português**	**Freitas**
-	-	-	Zéro	0,14	0,10	-	-
One	3,80	4,41	Un	0,55	0,80	Um, Hum	2,78
Two	4,35	3,67	Deux	5,74	5,20	Dois	2,92
Three	4,35	4,04	Trois	3,48	3,30	Três	2,81
Four	3,26	3,95	Quatre	7,26	7,20	Quatro	2,70
Five	2,72	3,95	Cinq	3,42	3,40	Cinco	2,78
Six	1,09	3,79	Six	1,41	2,20	Seis	2,53
Seven	2,17	4,05	Sept	1,49	2,10	Sete	2,24
Eight	2,17	3,62	Huit	1,63	2,30	Oito	2,55
Nine	2,72	3,67	Neuf	1,79	2,60	Nove	2,66
Ten	4,35	1,43	Dix	3,23	3,30	Dez	0,39
Eleven	2,72	1,41	Onze	0,30	0,30	Onze	0,41
Twelve	4,35	1,03	Douze	0,55	0,60	Doze	0,24
Thirteen	4,35	1,21	Treize	0,30	0,40	Treze	0,31
Fourteen	2,72	1,19	Quatorze	0,28	0,40	Quatorze	0,28
Fifteen	2,17	1,24	Quinze	0,77	0,90	Quinze	0,36
Sixteen	3,80	1,05	Seize	0,25	0,50	Dezesseis	0,24
Seventeen	1,63	1,23	-	-		Dezessete	0,29
Eighteen	3,26	1,20	-	-		Dezoito	0,28
Nineteen	1,63	1,13	-	-		Dezenove	0,36
Twenty	4,35	1,88	Vingt	7,26	6,70	Vinte	3,20
Thirty	4,35	1,72	Trente	2,24	2,20	Trinta	3,12
Forty	1,63	1,72	Quarante	2,01	1,90	Quarenta	2,83
Fifty	2,72	1,69	Cinquante	2,15	2,50	Cinquenta	3,15
Sixty	2,72	1,61	Soixante	3,12	3,90	Sessenta	2,98
Seventy	3,26	1,99	-	-		Setenta	3,09
Eighty	2,17	1,79	-	-		Oitenta	3,04
Ninety	2,72	1,43	-	-		Noventa	2,64
Hundred	2,17	15,41	Cent	20,62	17,10	Cem, Cento	2,36
-	-	-	-	-	-	Duzentos	2,09
-	-	-	-	-	-	Trezentos	1,97
-	-	-	-	-	-	Quatrocentos	1,82
-	-	-	-	-	-	Quinhentos	2,00
-	-	-	-	-	-	Seiscentos	1,83
-	-	-	-	-	-	Setecentos	2,07
-	-	-	-	-	-	Oitocentos	2,11
-	-	-	-	-	-	Novecentos	1,97
Thousand	4,35	6,21	Mille	3,51	4,70	Mil	10,79
Dollars	4,35	6,06	Francs, Frs	24,68	19,80	Reais, Real	10,59
And	4,35	10,17	et	0,80	2,30	-	-
Only	3,26	1,07	-	-		-	-
-	-	-	Centimes, Cts	1,02	3,30	Centavos, Centavo	9,18

A etapa de segmentação do extenso manuscrito em palavras não está apresentada no presente estudo, uma vez que esta segmentação está relacionada ao tratamento do extenso como um todo.

O procedimento de normalização das imagens não é normalmente utilizado pela maioria dos autores, tendo sido aplicado somente por [OLLIVIER, 1999]. O autor justifica a aplicação deste procedimento a normalização através de duas vantagens do procedimento: redução da quantidade de informação a tratar e redução da variabilidade das amostras. Portanto, a redução do número de pixels na imagem implica, para o trabalho de [OLLIVIER, 1999] com redes neurais, em:

- redução do tamanho do vetor a ser comparado pelo método de reconhecimento adotado e
- redução do número de neurônios e de conexões em uma rede neural.

Desde modo, [OLLIVIER, 1999] afirma ser possível reduzir os tempos de treinamento e reconhecimento das palavras manuscritas.

A Tabela B.4 apresenta os procedimentos que foram adotados pelos autores selecionados. A composição dos diferentes procedimentos está relacionada ao direcionamento do trabalho de cada um dos autores, bem como, as características intrínsecas das bases de dados utilizadas.

Para obtenção de todos os detalhes de cada um dos procedimentos aplicados pelos autores recomenda-se consultar os respectivos trabalhos.

Tabela B.4: Pré-Processamento de Palavras Manuscritas

Procedimentos	**Avila**	**Côte**	**Guillevic**	**Ollivier**	**Freitas**
Pré-processamento	Sim	Sim	Sim	Sim	Sim
Correção da linha de base	Sim	Não	Sim	Não	Não
Correção da inclinação vertical dos caracteres	Sim	Não	Sim	Não	Sim
Detecção do corpo das palavras	Sim	Sim	Sim	Não	Sim
Normalização do corpo das palavras	Sim	Não	Não	Não	Não
Normalização do tamanho das imagens	Não	Não	Não	Sim	Não
Esqueletização	Sim	Não	Não	Não	Sim[1]

[1] Somente para a extração das primitivas globais do tipo côncavo e convexo no corpo das palavras

B.3. Extração de Primitivas

Apresenta-se os conjuntos de primitivas selecionados pelos diferentes autores para serem extraídas das palavras manuscritas. Assim, pode-se considerar que diferentes tipos de formas precisam de diferentes tipos de primitivas, que por sua vez precisam de diferentes tipos de classificadores. Apesar de todos os trabalhos processarem palavras manuscritas no contexto de cheques bancários, diferentes primitivas foram selecionadas buscando-se conjuntos invariantes e que permitam a reconstrução, tal qual descrito no Capítulo 2 [TRIER et al., 1996].

As aplicações na área de reconhecimento de palavras manuscritas combinam os diferentes tipos de primitivas visando obter métodos mais robustos através de um conjuntos de primitivas com alta representatividade das formas das palavras.

A Tabela B.5 apresenta as primitivas selecionadas pelos diferentes autores, sendo que a escolha das mesmas está diretamente relacionada com os métodos de treinamento e de reconhecimento adotados.

Ressalta-se que [AVILA, 1996] utiliza primitivas primárias para definição de um alfabeto de 12 traços. [OLLIVIER, 1999] utiliza como método de treinamento e reconhecimento as redes neurais, portanto seu trabalho descreve a arquitetura da rede neural adotada, não enfatizando a extração de primitivas.

Deve-se relacionar ainda, a abordagem adotada por cada autor no que se refere a aplicação ou não da segmentação das palavras em letras ou pseudo-letras. Todos os resultados mostrados no presente estudo dos autores selecionados adotam a abordagem global, ou seja, sem segmentação explícita das palavras em letras ou pseudo-letras. Alguns autores trabalharam também com a abordagem analítica, porém estes resultados não se encontram apresentados neste trabalho.

B.4. Metodologias de Reconhecimento

Os autores selecionados trabalham com os diferentes métodos de treinamento e reconhecimento das palavras manuscritas. Estes métodos encontram-se descritos no Capítulo 2 e foram empregados pelos diferentes conforme Tabela B.6.

Tabela B.5: Extração de Primitivas

Primitivas	Avila	Côte	Guillevic	Freitas
1) Perceptivas: ascendente, descendente e laço	X	X	X	X
2) Secundárias ou condicionais: barras do "t", laços associados com ascendentes ou descendentes ("d" ou "p")		X		
3) Concavidades e convexidades		X		X
4) Número de transições preto-branco sobre as linhas do corpo das palavras		X		
5) Comprimento da palavra			X	
6) Traços horizontais e verticais			X	
7) Diagonais leste e oeste			X	

Tabela B.6: Metodologias de Reconhecimento

Métodos	Avila	Côte	Guillevic	Ollivier	Freitas
Rede Neural		X		X	
HMM	X				X
Não-Paramétricos (kPPV)[1]			X	X	

[1] kPPV significa Métodos dos k-Vizinhos mais Próximos

B.5. Resultados Obtidos

Os resultados obtidos pelos autores consideram todas as etapas e métodos anteriormente apresentados. Os resultados apresentados são referentes aos experimentos realizados com as bases de dados de testes e com a escolha da melhor opção de reconhecimento, ou seja, para o modelo que apresenta a maior probabilidade de reconhecimento.

As Tabelas B.7 e B.8 mostram os resultados obtidos pelos autores para o reconhecimento de palavras isoladas no contexto de cheques bancários. Na Tabela B.7 estão expostos os resultados obtidos pelos autores que trabalharam com as bases de dados de palavras do idioma inglês. As colunas denominadas *Guillevic ADS* e *Guillevic AD* correspondem aos resultados para dois conjuntos distintos de primitivas, denominados de ADS – conjunto formado por todas as primitivas e AD - conjunto formado somente por ascendentes e descendentes. Os resultados obtidos por [GUILLEVIC, 1995] utilizando o conjunto de primitivas ADS é melhor que quando empregado o conjunto AD. Observa-se também, que os resultados de [CÔTE, 1997] e

[GUILLEVIC, 1995] para o conjunto ADS são muito próximos. Isto, deve-se ao fato das primitivas globais selecionados pelos autores serem de mesma conceituação e formalismo de extração.

As palavras do léxico em inglês que apresentam maior grau de confusão para [GUILLEVIC, 1995] são as seguintes: “one” e “two”, “one” e “five”, “one” e “six”, “one” e “nine”. Para o conjunto de primitivas ADS as palavras com taxa de reconhecimento inferior a 50% são as seguintes: “sixteen”, “sixty”, “seven”, “seventeen”, “nineteen”, “ninety”. Observa-se que estas palavras possuem os mesmos prefixos entre si (“six”, “seven” e “nine”) mostrando que as primitivas extraídas não são suficientes para a representação dos sufixos (“teen” e “ty”).

Tabela B.7: Resultados Obtidos – Idioma Inglês

Palavras	**Côte**	**Guillevic ADS**	**Guillevic AD**
One	51,0	62,5	86,5
Two	81,0	77,5	23,6
Three	79,1	62,4	49,5
Four	82,2	65,5	60,0
Five	85,6	60,2	30,6
Six	30,1	70,5	36,8
Seven	42,2	52,6	54,6
Eight	76,6	87,6	72,4
Nine	46,8	63,0	43,5
Ten	58,5	52,9	52,9
Eleven	71,0	55,6	57,8
Twelve	73,5	57,1	42,9
Thirteen	59,6	58,8	35,3
Fourteen	74,2	56,7	43,3
Fifteen	72,0	59,5	40,5
Sixteen	81,0	39,1	25,0
Seventeen	69,0	45,5	45,5
Eighteen	79,2	74,1	63,0
Nineteen	81,8	36,7	48,3
Twenty	85,9	70,0	45,0
Thirty	63,5	73,3	55,6
Forty	76,9	58,7	58,7
Fifty	71,9	60,5	47,4
Sixty	66,7	37,5	37,5
Seventy	84,2	59,2	59,2
Eighty	84,1	86,0	81,4
Ninety	86,4	47,5	50,0
Hundred	78,2	95,1	95,1
Thousand	85,6	80,9	79,6
Dollars	84,6	72,5	73,2
And	71,6	90,4	81,6
Only	53,4	72,7	36,4
Média	**73,6**	**72,6**	**63,9**

Para o conjunto AD as palavras com taxa de reconhecimento inferior a 50% são as seguintes: “two”, “three”, “five”, “six”, “nine”, “twelve”, “thirteen”, “fourteen”, “fifteen”, “sixteen”, “seventeen”, “nineteen”, “twenty”, “fifty”, “sixty”, “ninety” e “only”. Este maior número de palavras com taxas inferiores a 50 % se traduz em uma taxa média para este conjunto AD inferior a obtida com o conjunto ADS. Estas palavras representam isoladamente parcelas aproximadamente de 1% da base de dados de treinamento, sendo este fato mais evidente nas palavras da classe “teen”.

Para [CÔTE, 1997] as palavras que apresentam maiores problemas de reconhecimento, com taxa inferior a 50%, são as seguintes: “six”, “seven” e “nine”.

Percebe-se que as dificuldades inerentes ao problema do reconhecimento de classes de palavras com prefixos e sufixos semelhantes independe do idioma do léxico estudado.

A Tabela B.8 apresenta os resultados obtidos com as bases de dados de palavras em idioma francês. [GUILLEVIC, 1995] apresenta somente os resultados das palavras isoladas obtidos com o conjunto AD (ascendentes e descendentes), o qual apresenta uma taxa média de reconhecimento de 78,3%, uma vez que, o mesmo revela uma significativa melhora quando comparados com os resultados obtidos com o conjunto ADS (todas as primitivas), média de 71,9%. [GUILLEVIC, 1995] comenta que o fato do conjunto AD ser mais representativo deve-se a dois fatores: primeiro que o tamanho do léxico de palavras francesas (25 palavras) é menor do que o léxico de palavras inglesas (32 palavras), e segundo, que as palavras em francês possuem mais ascendentes e descendentes distintos entre todas as palavras do léxico.

[GUILLEVIC, 1995] mostra que o conjunto AD resulta em taxas de reconhecimento para as palavras “huit” e “dix” inferiores a 30%, taxas estas muito abaixo das taxas das demais palavras. [GUILLEVIC, 1995] explica que a palavra “huit” é confundida com a palavra “cent” e que a palavra “dix” é confundida com as palavras “six” e “deux”.

[AVILA, 1996] não apresenta a matriz de confusão para as palavras estudadas, porém, observa-se que as palavras “zéro”, “six”, “sept”, “huit”, “treize”, “quatorze”, “quinze”, “seize”, “trente”, “centimes” e “et” apresentam taxas de reconhecimento inferior a 30%. As palavras com maior grau de confusão entre si são: “deux” confundida com “dix” e com “trois”, “quarante” confundida com “quatre”, e ainda, “six” confundida com a palavra “dix”.

Os resultados obtidos por [OLLIVIER, 1999] mostram que os dois métodos aplicados, apesar de serem diferentes, traduzem resultados semelhantes, com uma grande diferença de tempo de processamento: as redes neurais são mais rápidas. As redes neurais apresentam grande grau de confusão entre as seguintes palavras: “seize” e “treize”, “quatorze” e “quinze”, “six” e “dix”, “quarante” e “soixante”, “cinquante” e “soixante”, “deux” e “francs”. O método dos kPPV apresentam grande grau de confusão entre as seguintes palavras: “cinquante” e “quarante”, “soixante” e “quarante”, “cent” e “et”, “deux” e “francs”.

Tabela B.8: Resultados Obtidos – Idioma Francês

Palavras	**Avila**	**Guillevic AD**	**Ollivier Redes Neurais**	**Ollivier kPPV**
Zéro	20,0	-	9,0	3,0
Un	33,3	86,7	43,0	60,0
Deux	50,2	73,1	67,0	61,0
Trois	36,2	76,0	52,0	40,0
Quatre	75,2	95,2	78,0	83,0
Cinq	51,6	92,2	67,0	69,0
Six	14,6	64,4	42,0	36,0
Sept	23,5	94,0	40,0	24,0
Huit	13,7	27,4	58,0	32,0
Neuf	41,7	71,9	64,0	66,0
Dix	67,3	18,2	67,0	66,0
Onze	25,0	79,5	30,0	24,0
Douze	26,7	82,9	29,0	29,0
Treize	11,1	73,0	18,0	8,0
Quatorze	9,1	97,8	37,0	32,0
Quinze	17,1	86,7	51,0	21,0
Seize	11,1	51,4	29,0	13,0
Vingt	68,3	85,2	80,0	68,0
Trente	21,3	64,4	50,0	36,0
Quarante	45,3	91,2	64,0	77,0
Cinquante	56,6	95,0	77,0	83,0
Soixante	59,3	84,5	68,0	67,0
Cent	81,7	87,3	87,0	87,0
Mille	59,8	80,3	74,0	66,0
Francs	78,0	89,2	88,0	88,0
Centimes	9,6	-	45,0	51,0
Et	21,4	95,2	63,0	89,0
Média	**62,2**	**78,3**	**75,0**	**75,0**

A seguir apresenta-se uma discussão dos resultados obtidos evidenciando-se as principais diferenças entre os experimentos apresentados.

B.6. Análise dos Resultados

Menciona-se no início deste Capítulo que a comparação de resultados não é uma tarefa simples devido a diversos fatores. Apresenta-se os resultados obtidos pelos diferentes autores selecionados e pode-se perceber que diferentes características de cada estudo em particular.

A principal diferença são as origens das palavras: inglesa, francesa e portuguesa. Pode-se dizer, num primeiro momento que o francês e português possuem a mesma origem, ou seja, são idiomas latinos. Desde modo, pode-se esperar que exista uma semelhança entre as palavras, o que de fato ocorre. Por outro lado, o francês expressa alguns valores numéricos por escrito utilizando mais de uma palavras (90 – "quatre vingt dix"), o que torna o léxico a ser reconhecido em cada uma das línguas diferente, considerando-se principalmente o tamanho do léxico. Por sua vez, o inglês é um idioma de origem não latina tornando a comparação de resultados ainda mais difícil.

As palavras da classe "dezena" apresentam taxas de reconhecimento baixas independente do idioma do léxico, os resultados obtidos estão relacionados com a freqüência destas palavras nas bases de dados e, especialmente, na base de dados de treinamento.

O reconhecimento das palavras curtas enfrenta problemas independentemente do idioma do léxico, revelando que as observações de [SCHOMAKER & SEGER, 1998] são verdadeiras. Assim, o reconhecimento de palavras curtas exige um maior número de informações a serem extraídas.

O reconhecimento de palavras longas da classe " entos", para o caso do léxico em idioma português, revela não somente a similaridade com a classe de palavras " enta" como também confirma as observações de [SCHOMAKER & SEGER, 1998]. Para os autores, palavras longas necessitam de maiores informações no final das mesmas para serem reconhecidas. Ressalta-se que esta classe de palavras não existe nos léxicos dos idiomas em francês e inglês, uma vez que os valores numéricos são grafados pela combinação de outras palavras existentes no léxico.

Assim, levando-se em consideração todos estes fatores intrínsecos a cada idioma, a cada léxico e a cada trabalho realizado, apresenta-se na Tabela B.9 as médias dos resultados obtidos pelos diferentes autores. Neste Tabela **%** é a taxa de

reconhecimento, **TL** é o tamanho do léxico e **BT** é a dimensão da base de dados de treinamento utilizada nos experimentos realizados.

Sabe-se, ainda, que muitos outros trabalhos poderiam ter sido selecionados para realização da comparação apresentada por este estudo. Ressalta-se a importância de alguns sistemas desenvolvidos no mesmo contexto do presente trabalho, descritos em [IMPEDOVO et al., 1997], sendo estes os seguintes:

- O sistema A2iA apresentado por Knerr et al.,
- O sistema proposto por Dimauro et al.,
- O sistema proposto por Leroux et al. e
- O sistema proposto por Saion & Belaïd.

A Tabela B.10 mostra os resultados obtidos com os experimentos realizados pelos respectivos sistemas envolvendo o reconhecimento de palavras isoladas do extenso manuscrito de cheques bancários em idioma francês e italiano.

Tabela B.9: Resumo dos Resultados Obtidos

Autor	**Inglês**			**Francês**			**Português**		
	%	**TL**	**BT**	**%**	**TL**	**BT**	**%**	**TL**	**BT**
Côte	73,6	32	184	-			-		
Guillevic ADS	72,6	32	5.322	-			-		
Guillevic AD	63,9	32	5.322	-			-		
Avila	-			62,2	27	3.623	-		
Guillevic AD	-			78,3	25	1.345	-		
Ollivier - Redes Neurais	-			75,0	29	40.400	-		
Ollivier - kPPV	-			75,0	29	40.400	-		
Freitas – PP	-			-			56,4	39	7.136
Freitas – PPCC	-			-			64,8	39	7.136
Freitas – PPCCR	-			-			67,7	39	7.136
Freitas – PPCCR-FR	-			-			70,6	39	7.136

Tabela B.10: Resultados de Sistemas de Reconhecimento

Autores	**Idioma**	**%**	**TL**	**BT**
Sistema A2iA	Francês	77,1	26	10.000
Dimauro et al.	Italiano	69,0	50	*
Leroux et al.	Francês	74,0	26	23.066
Saion & Belaïd	Francês	72,4	26	4.653

* apresenta somente o tamanho da base de dados de testes