



CINTHIA O. DE A. FREITAS



PERCEPÇÃO VISUAL E RECONHECIMENTO DE PALAVRAS MANUSCRITAS

Trabalho apresentado ao Departamento de Ciência da Computação da Pontifícia Universidade Católica do Paraná como requisito para o Concurso de Promoção na Carreira Docente à classe de Professor Titular.

CURITIBA

2002



Sumário

Sumário	iii
Lista de Figuras	v
Lista de Tabelas	vii
Resumo	viii
Abstract	ix
Capítulo 1	1
Introdução	
1.1. Desafio	2
1.2. Motivação	3
1.3. Contribuição	3
1.4. Organização	4
Capítulo 2	5
Percepção Visual	
2.1. Olho Humano	5
2.1.1. Transformações Ópticas	6
2.1.2. Transformações Químicas	7
2.1.3. Transformações Nervosas	7
2.2. Elementos da Percepção Visual	8
2.2.1. Percepção da Luminosidade: A Intensidade da Luz	8
2.2.2. Percepção da Cor: O Comprimento de Onda	9
2.2.3. Percepção das Bordas Visuais: A Distribuição Espacial da Luz	10
2.2.4. Percepção do Contraste: Luminosidade x Bordas	11
2.3. Teoria da Gestalt e a Percepção das Formas	12
2.3.1. Os Princípios da Gestalt	13
2.3.2. Constância Perceptiva	17
2.3.3. Ilusões Ópticas	18
2.4. Captura do Essencial	22
2.5. Abordagens da Percepção Visual	24
2.5.1. A Abordagem Analítica	24
2.5.2. A Abordagem Sintética	25
2.5.3. Qual é a Melhor Abordagem?	25
2.6. Princípio da Maior Probabilidade ou o Direito da Incerteza?	26
2.7. Discussão	28
Capítulo 3	31
Escrita e Leitura	
3.1. Escrita	32
3.2. Leitura	38
3.3. Conhecimento Prévio	40
3.4. Rememoração e Reconhecimento	42
3.5. Leitura e Reconhecimento	45
3.6. Discussão	48

Capítulo 4	55
Percepção e Reconhecimento de Palavras Manuscritas	
4.1. Níveis de Abstração	55
4.1.1. Objeto: A Palavra e seu Entorno	56
4.1.2. Global: Palavra	58
4.1.3. Local: Letra	59
4.1.4. Pixel: Palavra Digital	60
4.2. Percepção e Reconhecimento	61
4.2.1. Vogais e Consoantes, Palavras Curtas e Longas, Ascendentes e Descendentes	62
4.2.2. Velocidade da Escrita para Diferentes Estilos	63
4.2.3. Movimento Horizontal da Escrita	65
4.2.4. O Grafema no Espaço e no Tempo	68
4.2.5. Variabilidade da Escrita Cursiva em Função da Posição da Letra na Palavra	69
4.3. Português e Francês: Percepção e Reconhecimento	70
4.3.1. Considerações Iniciais	70
4.3.2. Bases de Dados: LUCI/EXTENSO/PALAVRAS e IRONOFF	71
4.3.3. Resultados Alcançados por Freitas em [FRE01]	75
4.3.4. Hipótese e Experimentos	76
4.3.5. Análise dos Resultados	78
Conclusão	85
Referências Bibliográficas	89

Lista de Figuras

Figura 1.1	Percepção de características de palavras manuscritas	1
Figura 2.1	Fisiologia do olho humano	6
Figura 2.2	Transformações ópticas	7
Figura 2.3	Parâmetros de influência na cor: matiz, saturação e claridade	9
Figura 2.4	Círculo das cores	10
Figura 2.5	Cores do espectro visível pelo olho humano	10
Figura 2.6	Percepção das bordas visuais: a) esquema das estruturas acumuladoras e inibidoras b) fendas, linhas e bordas	11
Figura 2.7	Percepção do contraste	12
Figura 2.8	Placa de sinalização	12
Figura 2.9	Princípios da Gestalt: a) proximidade, b) similaridade e c) continuidade	14
Figura 2.10	Princípio do Fechamento: a) triângulo de Kanizsa e b) palavras incompletas	14
Figura 2.11	Princípio da Gestalt: fechamento	15
Figura 2.12	Princípio da Gestalt: figura-fundo	15
Figura 2.13	Exemplos do princípio da Gestalt: figura-fundo	16
Figura 2.14	Princípio da Gestalt: a) área e b) simetria	16
Figura 2.15	Princípio da Gestalt: pregnância	17
Figura 2.16	Princípio da Gestalt: pregnância	17
Figura 2.17	Constância perceptiva: a) posição original b) posição deslocada na imagem	18
Figura 2.18	Ilusão de Miller-Lyer	19
Figura 2.19	Ilusão de Hering	19
Figura 2.20	Ilusão de Ponzo	19
Figura 2.21	Ilusão de Poggendorf	20
Figura 2.22	Exemplos de ilusões ópticas: a) azulejos, b) movimento em espiral, c) inibição lateral d) cubos	21
Figura 2.23	Exemplos de ilusões ópticas: a) pernas do elefante, b) ponto e neblina, c) linhas paralelas d) pontos pretos	21
Figura 2.24	Movimentos oculares	22
Figura 2.25	Movimentos oculares durante a percepção visual	23
Figura 2.26	Princípio da maior probabilidade: a) pato, b) coelho, c) pato ou coelho?	27
Figura 3.1	Percepção visual e leitura: manchas e palavras	31
Figura 3.2	Percepção visual e leitura: cores e palavras	32
Figura 3.3	Exemplos de escrita pictórica: a) cantos Ojibwa; b) catecismo asteca; c) inscrição antiga	33
Figura 3.4	Exemplos de escrita ideográfica: escrita egípcia	34
Figura 3.5	Evolução da escrita alfabética	34
Figura 3.6	Estilo mistral: enlace perfeito entre as diversas letras	35
Figura 3.7	Tipos de escrita	35
Figura 3.8	Exemplos da letra “A” em diferentes estilos	36
Figura 3.9	Variações da escrita com base no traçado	37
Figura 3.10	Estilos de escrita cursiva	37

Figura 3.11	Alfabeto e estilo de escrita cursiva para alfabetização no Brasil	38
Figura 3.12	Níveis de leitura	40
Figura 3.13	Rememoração e esquemas: a) afrescos da tumba de Senefer, b) placas de sinalização, c) logotipos	43
Figura 3.14	Caricatura e invariantes da visão	44
Figura 3.15	A faxineira e o seu balde	45
Figura 3.16	Mecanismos e capacidades envolvidos no processamento de um texto	51
Figura 4.1	Exemplos de palavras manuscritas e o contexto: a) cheque bancário e b) envelope postal	56
Figura 4.2	Exemplo de segmentação Figura-Fundo: a) imagem original, b) “bloco-endereço manuscrito e carimbo”, c) “fundo”, d) “selo”	57
Figura 4.3	Exemplo de segmentação Figura-Fundo: a) imagem original, b) “bloco-endereço manuscrito” com <i>endereço=nome de rua + bairro</i> segmentado	58
Figura 4.4	Exemplos de análise global de palavras manuscritas de cheques bancários brasileiros: a) estilo cursivo e b) estilo caixa-alta	59
Figura 4.5	Exemplos de segmentação de palavras manuscritas em letras ou pseudoletas: a) imagem original, b) hipótese de segmentação descartada, c) hipótese de segmentação aceita, d) palavra segmentada..	60
Figura 4.6	Nível de abstração: pixel	60
Figura 4.7	Exemplos de processo de binarização: a) imagem original em 256 níveis de cinza, b) $threshold = 217$ (método anisotropia) e c) $threshold = 110$ (método OTSU)	61
Figura 4.8	Estilos de escrita estudados por [SUE83]	63
Figura 4.9	Exemplos de modelos para escrita em letra de forma	64
Figura 4.10	Velocidade de escrita dos modelos de seqüências para caracteres no estilo <i>block print</i>	65
Figura 4.11	Regiões de inclinação vertical (slant) dos ascendentes e descendentes em palavras manuscritas	66
Figura 4.12	Exemplos do manuscrito do mesmo escritor para 3 espaçamentos horizontais: a) normal, b) estreito e c) largo	67
Figura 4.13	Padrões analisados: a) laços, b) ondas e c) zig-zag	68
Figura 4.14	Classificação do grafema <i>f</i>	70

Lista de Tabelas

Tabela 4.1	Léxico em Português e Francês	72
Tabela 4.2	Distribuição das Imagens para as Bases de Treinamento	74
Tabela 4.3	Taxas de Reconhecimento das Palavras para os Conjuntos de Primitivas	75
Tabela 4.4	Taxas de Reconhecimento das Palavras em Francês	76
Tabela 4.5	Taxas de Reconhecimento das Palavras em Português com Treinamento dos Modelos MEM em Português + Francês	78
Tabela 4.6	Comparação de Resultados: Taxas de Reconhecimento	79
Tabela 4.7	Matriz de Confusão para o Conjunto de Primitivas PPCCR [FRE01]	81
Tabela 4.8	Matriz de Confusão para os Modelos Modificados	82
Tabela 4.9	Acréscimo da Taxa de Reconhecimento para as Meta-Classes	80
Tabela 4.10	Taxa de Reconhecimento dos Conjuntos de Primitivas Perceptivas para Equação (4.4)	83

Resumo

O presente trabalho desenvolve o estudo da relação entre a percepção humana e o processo de leitura para extrair subsídios ao reconhecimento automático de palavras manuscritas. Apresenta-se uma revisão sobre a percepção visual humana, uma vez que o sentido utilizado pelo ser humano para efetuar o reconhecimento de imagens e a leitura de palavras é a visão. Este trabalho desenvolve a hipótese de melhorar o reconhecimento de palavras do extenso manuscrito de cheques bancários brasileiros, utilizando-se para isto uma base de palavras francesas, para o mesmo contexto. As bases utilizadas nos experimentos foram as seguintes: a base LUCI/EXTENSO/PALAVRAS pertencente ao LUCI/PPGIA/PUCPR e a base IRONOFF pertencente ao IRESTE Institute, respectivamente para palavras em português e em francês. Os experimentos realizados confirmam esta hipótese, alcançando-se uma taxa de reconhecimento médio para as 39 palavras de 69,09% (Exp05) e 71,96% (Exp05 + ProbClasse). Destaca-se a melhora na taxa de reconhecimento para a meta-classe da *dezena*, na qual obteve-se uma melhora de 17,33%, comparados aos resultados obtidos nos estudos anteriores. Finalmente, apresenta-se a conclusão juntamente com a proposição de trabalhos futuros.

Palavras-Chave: Percepção Visual, Reconhecimento de Palavras Manuscritas, Escrita, Leitura.

Resumo

O presente trabalho desenvolve o estudo da relação entre a percepção humana e o processo de leitura para extrair subsídios ao reconhecimento automático de palavras manuscritas. Apresenta-se uma revisão sobre a percepção visual humana, uma vez que o sentido utilizado pelo ser humano para efetuar o reconhecimento de imagens e a leitura de palavras é a visão. Este trabalho desenvolve a hipótese de melhorar o reconhecimento de palavras do extenso manuscrito de cheques bancários brasileiros, utilizando-se para isto uma base de palavras francesas, para o mesmo contexto. As bases utilizadas nos experimentos foram as seguintes: a base LUCI/EXTENSO/PALAVRAS pertencente ao LUCI/PPGIA/PUCPR e a base IRONOFF pertencente ao IRESTE Institute, respectivamente para palavras em português e em francês. Os experimentos realizados confirmam esta hipótese, alcançando-se uma taxa de reconhecimento médio para as 39 palavras de 69,09% (Exp05) e 71,96% (Exp05 + ProbClasse). Destaca-se a melhora na taxa de reconhecimento para a meta-classe da *dezena*, na qual obteve-se uma melhora de 17,33%, comparados aos resultados obtidos nos estudos anteriores. Finalmente, apresenta-se a conclusão juntamente com a proposição de trabalhos futuros.

Palavras-Chave: Percepção Visual, Reconhecimento de Palavras Manuscritas, Escrita, Leitura.

Abstract

This work presents a study about the visual perception and the reading process for handwritten word recognition system. A review about human visual perception is presented, because the vision is the sense used by human being for recognizing images and reading words. This work analyses the hypothesis for improving the handwritten word recognition from Brazilian bank checks using a French database into the same context. The databases used in the experiments were the LUCI/EXTENSO/PALAVRAS database (LUCI/PPGIA/PUCPR) and IRONOFF database (IRESTE Institute), including respectively Portuguese and French words. The experiments held with Portuguese lexicon (39 words) confirm the hypothesis and reached a recognition rate of 69,09% (Exp05) and 71,96% (Exp05 + ProbClasse). A significant increase of 17,33% in the recognition rate is observed with the meta-class *dezena*. Finally, the work presents the conclusion and the future works that can be done in continuity to this one.

Keywords: Visual Perception, Handwritten Word Recognition, Writing and Reading.

Abstract

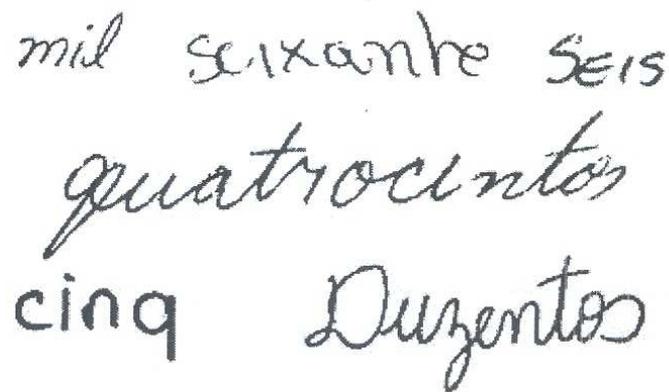
This work presents a study about the visual perception and the reading process for handwritten word recognition system. A review about human visual perception is presented, because the vision is the sense used by human being for recognizing images and reading words. This work analyses the hypothesis for improving the handwritten word recognition from Brazilian bank checks using a French database into the same context. The databases used in the experiments were the LUCI/EXTENSO/PALAVRAS database (LUCI/PPGIA/PUCPR) and IRONOFF database (IRESTE Institute), including respectively Portuguese and French words. The experiments held with Portuguese lexicon (39 words) confirm the hypothesis and reached a recognition rate of 69,09% (Exp05) and 71,96% (Exp05 + ProbClasse). A significant increase of 17,33% in the recognition rate is observed with the meta-class *dezena*. Finally, the work presents the conclusion and the future works that can be done in continuity to this one.

Keywords: Visual Perception, Handwritten Word Recognition, Writing and Reading.

Capítulo 1

Introdução

Retire algumas palavras de um texto e coloque-as sobre uma folha em branco de papel, como na Figura 1.1. Algumas características (localização, inclinação, tamanhos e tipos de letra) podem ser determinadas e descritas por meio de medidas. Pode-se concluir que a palavra “quatrocentos” possui uma inclinação em relação ao eixo vertical, que a palavra “seis” está escrita em letra de forma ou caixa-alta enquanto a palavra “duzentos” está em letra cursiva, que as palavras “cinq” e “soixante” não pertencem ao vocabulário de palavras da língua portuguesa e, ainda, que a palavra “mil” possui menos caracteres que a palavra “quatrocentos”.



mil soixante SEIS
quatrocentos
cinq Duzentos

Figura 1.1: Percepção de características de palavras manuscritas.

Este resultado não é nenhuma surpresa. Não se precisa medir ou efetuar cálculos – percebe-se de relance estas e outras características das palavras. Como, então, ocorre este “ato de perceber”? Será que o comportamento humano assemelha-se ao de um instrumento de medida, extraindo ângulos, alturas e larguras das palavras e das letras, para compará-las entre

si? Provavelmente não. Este não seria o procedimento mais indicado, quando se pensa em seres humanos processando a informação.

Deve-se mencionar também que não foi pedido ao leitor para efetuar a leitura das palavras, mas ao observá-las, com certeza, o leitor leu cada uma delas. Pois, posso garantir que quando as mesmas foram mencionadas no primeiro parágrafo o leitor não teve dificuldade de localizar a palavra mencionada no texto em sua respectiva posição na Figura 01. Portanto, como o “ato de perceber” e o “ato de ler” estão associados? Que características da percepção visual auxiliam na leitura de palavras manuscritas?

Também não foi pedido ao leitor que analise o contexto das palavras, mas em uma única olhada o leitor será capaz de perceber que as palavras relacionam-se a grafia por extenso de valores numéricos. Então, como é possível fazer isto?

O presente trabalho realiza uma revisão dos aspectos da percepção visual e, portanto, das abordagens da percepção visual (analítica e sintética), do princípio da maior probabilidade, do reconhecimento de imagens, do processo de leitura de palavras manuscritas e do seu reconhecimento. Todos estes aspectos estão relacionados de forma a buscar as respostas aos questionamentos apresentados e, portanto, auxiliar no entendimento do processo humano de reconhecimento de imagens e de palavras manuscritas. Objetiva-se subsidiar o desenvolvimento de sistemas computacionais de reconhecimento de palavras manuscritas.

1.1. Desafio

O desenvolvimento de sistemas de reconhecimento automático de palavras manuscritas tem desafiado os pesquisadores devido ao alto grau de dificuldade em reproduzir a capacidade humana de ler. Como escrito por El Yacoubi em [YAC96]: “L’objectif de la reconnaissance est de développer un système qui se rapproche le plus de l’être humain dans sa capacité de lire un mot quelconque (écriture manuscrite sans contrainte)“.

Côte [CÔT96] [CÔT98] foi buscar na percepção humana os aspectos relevantes que pudessem contribuir para o reconhecimento de palavras manuscritas. A autora, pergunta: “What features are detected while reading? How do human access information concerning the meaning of a word? Does perception of a word build up from the perception of its letters, or by its shapes?”. Portanto, através dos anos os pesquisadores das áreas de biologia, neurofisiologia, psicologia cognitiva e lingüística têm se empenhado em estudar estas questões e propor sistemas baseados no processo humano de ler.

Assim, quando se entende sistema de reconhecimento de palavras manuscritas como a implementação de técnicas e métodos que permitem o computador ler o que está escrito, os pesquisadores deparam-se com um grande desafio. Entre estes desafios, pode-se considerar como importante a grande variabilidade da escrita, ou seja, de tipos de letras e criações individuais de cada escritor. Por outro lado, a capacidade humana de ler, praticamente qualquer tipo de letra, desafia os pesquisadores a entender cada vez mais o ser humano para, somente assim, entender porque seus sistemas computacionais falham.

1.2. Motivação

Em Suen [SUE98] encontram-se as seguintes perguntas que são compartilhadas pelo presente trabalho:

- Como o ser humano reconhece os diferentes caracteres?
- Quais são as diferentes partes de um caracter/palavra que o ser humano observa?
- Quais são as características que permitem a leitura de uma palavra?
- Existe uma maneira de transpor para o computador as boas qualidades de extração e reconhecimento do ser humano?
- Pode-se reconhecer os caracteres e as palavras se um destes estiver faltando?
- Como resolver o problema de formas ambíguas?

Não se pretende aqui encontrar todas as respostas a todas estas perguntas, porém a motivação do presente trabalho está no aprofundamento do estudo da percepção humana voltada aos sistemas de reconhecimento de palavras manuscritas.

1.3. Contribuição

O presente trabalho aprofunda o estudo da relação entre a percepção humana e o processo de leitura para extrair subsídios ao reconhecimento automático de palavras manuscritas. Relacionam-se a percepção humana e a leitura com a extração de primitivas para os processos automáticos de reconhecimento.

Analisa-se a possibilidade de melhorar o reconhecimento de palavras manuscritas para um léxico em português através do uso de um léxico em francês. Para isto, estuda-se a aplicação de primitivas perceptivas associadas com concavidades e convexidades [FRE01]. Além, de considerar-se trabalhos de autores que também compartilham dos mesmos questionamentos citados anteriormente, apresentando-se as visões do problema de cada autor e as constatações obtidas por estes autores.

O presente trabalho coloca-se como mais uma contribuição na área de Análise e Reconhecimento de Documentos, podendo vir auxiliar outras aplicações, 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.

1.4. Organização

Nesse capítulo foram apresentados os desafios, a motivação e as contribuições desse trabalho. O Capítulo 2 apresenta uma revisão sobre percepção visual humana, uma vez que o sentido utilizado pelo ser humano para efetuar a leitura de palavras é a visão.

O Capítulo 3 apresenta os aspectos relevantes dos processos de escrita e leitura, os quais devem ser considerados no desenvolvimento de sistemas de reconhecimento automático de palavras.

O Capítulo 4 estabelece a relação entre os capítulos anteriores, analisando trabalhos de autores que compartilham da mesma linha de pensamento e, portanto, investigam aspectos da percepção humana nos processos de escrita e leitura, de modo a transpô-los aos sistemas de reconhecimento. Este capítulo analisa a utilização de uma base de dados de palavras manuscritas em idioma francês de forma a investigar e validar aspectos da percepção humana no processo de leitura/reconhecimento de palavras em uma base de dados em português. Os experimentos foram realizados através do sistema para reconhecimento automático de palavras manuscritas, implementado durante o desenvolvimento da Tese de Doutorado: “Uso de Modelos Escondidos de Markov para Reconhecimento de Palavras Manuscritas”, desenvolvida por Freitas [FRE01].

Finalmente, o Capítulo Conclusão apresenta os pontos relevantes do trabalho e descreve as possibilidades de trabalhos futuros visando dar continuidade às pesquisas na área de reconhecimento de palavras manuscritas na PUCPR.

Capítulo 2

Percepção Visual

O estudo das teorias sobre a Percepção Visual iniciou-se no século XIX com Helmholtz e Fechner [AUM93]. Desde então, muitos estudos têm sido desenvolvidos para entender e relacionar a percepção visual com diversas áreas do conhecimento, tais como, a biologia, a psicologia, a neurofisiologia e, ainda, para o contexto deste trabalho, a lingüística.

Ver não é só olhar. Para ver tem-se que observar com atenção e compreender o que os olhos vêem. A este ato de ver e compreender denomina-se: percepção visual [AUM93]. Pois, a percepção visual permite que informações sejam transmitidas simultaneamente ao cérebro num "abrir e fechar de olhos" e, mesmo que nem todas sejam analisadas e decodificadas, são numerosas as que são registradas [CLO75].

Sabe-se que o ser humano tem a capacidade de reconhecer as coisas que estão ao seu redor, e sabe-se também, que duas pessoas podem olhar para uma forma natural ou um objeto e vê-lo de maneira diferente; o que é importante para um, poderá não ter sido para o outro. Pode-se dizer que a percepção visual de cada ser humano influencia, com certeza, o ato de reconhecer.

O presente capítulo apresenta uma revisão sobre percepção visual humana, uma vez que o sentido utilizado pelo ser humano para efetuar o reconhecimento de imagens e a leitura de palavras é a visão.

2.1. Olho Humano

O olho humano é o instrumento principal da visão [AUM93]. A Figura 2.1 mostra a fisiologia do olho humano, destacando as partes importantes [GRE79]. O olho humano é um globo aproximadamente esférico, de diâmetro em torno de dois centímetros, revestido por

uma camada em parte opaca, a *esclerótica*, em parte transparente, a *córnea*. A *córnea* garante a maior parte de convergência dos raios luminosos que atingem o olho formando a imagem no fundo do olho. Atrás da *córnea* encontra-se a *íris*, músculo esfíncter comandado de modo reflexo, a qual delimita em seu centro uma abertura denominada de *pupila*, cujo diâmetro varia entre 2 a 8 milímetros aproximadamente [GRE79].

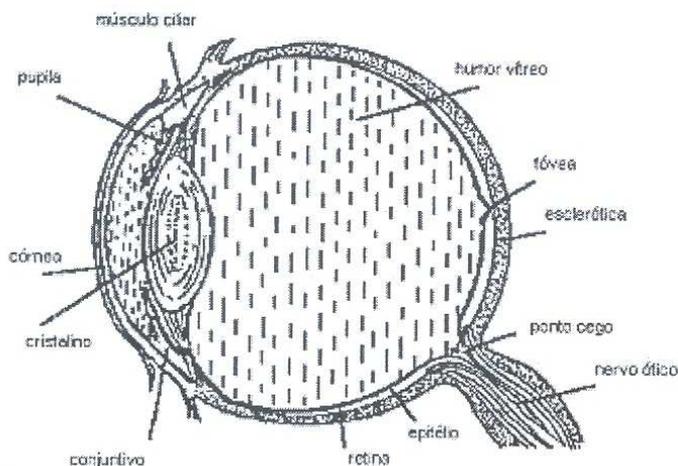


Figura 2.1: Fisiologia do olho humano. (Fonte: [GRE79])

A visão é um processo que utiliza diversos órgãos especializados. Desta forma, pode-se considerar a visão como resultante de três operações distintas e sucessivas, apresentadas a seguir: ópticas, químicas e nervosas [AUM93].

2.1.1. Transformações Ópticas

O olho humano funciona através do princípio da captura de uma grande quantidade de luminosidade (raios solares) sobre uma superfície e da concentração desta em um ponto. Assim, a luz atravessa a pupila e o cristalino que a fará convergir. O cristalino é uma lente biconvexa de convergência variável. O fato de a convergência ser variável dá-se o nome de *acomodação*. Portanto, acomodar é fazer variar a convergência do cristalino ao torná-lo mais ou menos abaulado, em função da distância da fonte de luz. A Figura 2.2 apresenta um esquema das transformações ópticas.

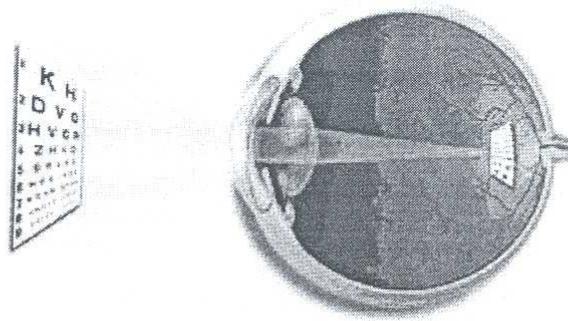


Figura 2.2: Transformações ópticas.
(Fonte: <http://portfolio.med.up.pt/marujo/patologia/Anatomia.htm>)

2.1.2. Transformações Químicas

O fundo do olho é revestido por uma membrana, a *retina*, na qual se encontram inúmeros receptores de luz. Estes receptores são de dois tipos, a saber: os bastonetes e os cones.

Os bastonetes e cones comportam moléculas de pigmento que contém uma substância, a *rodopsina*, a qual absorve a luminosidade e decompõe-se, por reação química, em outras substâncias. Os bastonetes são sensíveis aos detalhes e aos movimentos e são responsáveis por detectar a luminância. Em outras palavras, os bastonetes definem uma imagem em branco e preto.

Os cones são responsáveis pela detecção da cor. Existem 3 variedades de cones em função da cor que os mesmos detectam: verde, vermelho e azul. São menos sensíveis ao detalhe e ao movimento. Isto explica, o porquê de se ter dificuldade de se perceber a cor em um objeto que se movimenta rapidamente.

Assim, os receptores retinianos (bastonetes e cones) passarão a informação para as células nervosas para que transformações nervosas permitam que a imagem chegue ao cérebro através do *nervo óptico*.

2.1.3. Transformações Nervosas

O *nervo óptico* entra no cérebro e dirige-se à região responsável pela visão onde se processa o fenômeno de "formação das imagens". Cada olho recebe e envia ao cérebro uma imagem; no entanto, vê-se os objetos como um só, devido à capacidade de fusão das imagens em uma só. A visão binocular (com os dois olhos) dá ao seres um maior campo visual e a noção de profundidade.

As células nervosas estão ligadas aos receptores retinianos através de uma *synapse*, sendo que as sinapses estão ligadas as células que constituem a fibra do *nervo óptico*. O nervo óptico parte do olho e chega a uma região lateral do cérebro, denominada de *articulação*, de onde novas conexões nervosas saem em direção à parte posterior do cérebro, para chegarem ao *córtex*.

Nota-se que o processo é complexo e que não existe correspondência ponto-a-ponto, mas pelo contrário, ocorre a multiplicação de correspondências transversais. O sistema visual não se contenta em copiar a informação. Portanto, processa-a em cada etapa.

Este terceiro estágio do sistema perceptivo é o mais importante e, também, o mais desconhecido. Para o presente estudo, não interessa aprofundar os conhecimentos sobre as transformações realizadas no olho humano, mas entender como através de um processo de tratamento da informação a imagem pode tornar-se mais próximo dos modelos informatizados.

2.2. Elementos da Percepção Visual

A informação que entra pelos olhos está codificada, ou seja, segue regras de transformações naturais que determinam a atividade nervosa em função da informação contida na luz. A decodificação da informação visual somente é possível devido a capacidade de localizar e interpretar regularidades nos fenômenos luminosos que atingem os olhos. Estas regularidades estão relacionadas com três características da luz: intensidade, comprimento de onda e distribuição no espaço, como apresentado a seguir.

2.2.1. Percepção da Luminosidade: A Intensidade da Luz

A percepção da luminosidade é, na verdade, a interpretação dada pelo ser humano a maior ou menor quantidade de luz emitida por um objeto, se for uma fonte luminosa (exemplos: o sol e uma lâmpada elétrica), ou refletida por este, em todos os demais casos. A ordem de grandeza de luminância de um objeto varia de 10^{10} (sol) a 10^{-6} (limiar absoluto de percepção). Entende-se por luminância a intensidade luminosa por unidade de superfície aparente do objeto luminoso. Esta grandeza não depende do observador, mas apenas da fonte.

Quando a luminância de um objeto está entre 10^7 e 10^2 pode-se distinguir a visão fotópica, ou seja, objetos iluminados por uma luz diurna. Esta visão aciona, principalmente, os cones que são os responsáveis pela visão cromática. Portanto, a visão fotópica é cromática.

Quando a luminância de um objeto está entre 10^{-2} e 10^{-6} pode-se distinguir a visão escotópica, ou seja, objetos iluminados por uma luz noturna. Esta visão apresenta a característica da predominância pelos bastonetes, ou seja, a visão acromática, de fraca acuidade e relativa à periferia da retina.

2.2.2. Percepção da Cor: O Comprimento de Onda

A percepção da cor advém de reações ao comprimento de onda das luzes emitidas ou refletidas pelos objetos. Assim, a cor não está no objeto, mas sim na percepção humana.

A cor existe a partir da combinação de três parâmetros:

- matiz: dependente do comprimento de onda da luz emitida ou refletida pelo objeto,
- saturação: dependente da pureza da cor (exemplo: o rosa é o vermelho “menos saturado”),
- luminosidade ou claridade: dependente da luminância, ou seja, quanto maior for a luminosidade do objeto, mais a cor parecerá luminosa e próxima do branco.

Aumont em [AUM93] explica que essas três variáveis podem ser representadas por dois cones justapostos pela base, como na Figura 2.3 colocando-se a claridade variando no sentido do eixo, de baixo para cima; o matiz no círculo central e a saturação no raio deste círculo. O círculo central, como mostrado na Figura 2.4, é também chamado de círculo de cores. Esta figura apresenta uma síntese aditiva, mostrando as 12 cores provenientes das cores primárias: vermelho, verde e azul e das secundárias: magenta, amarelo e ciano.

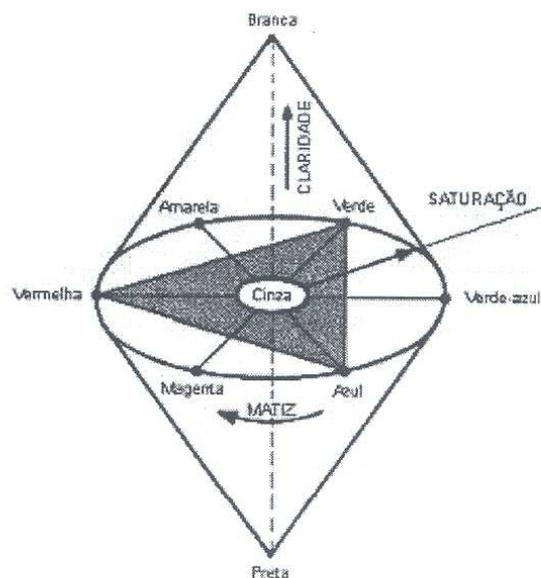


Figura 2.3: Parâmetros de influência na cor: matiz, saturação e claridade. (Fonte: [GRE79])

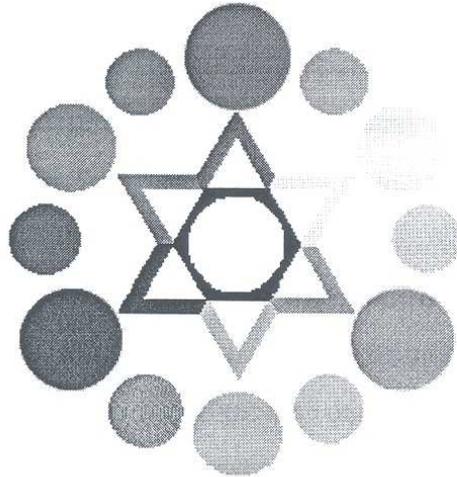


Figura 2.4: Círculo das cores. (Fonte: [GRE79])

A percepção da cor existe devido à atividade de três variedades de cones retinianos, sendo cada um sensível a um comprimento de onda diferente. Em uma pessoa não daltônica, esses comprimentos são os seguintes:

- 0,440 μ correspondendo ao azul-violeta,
- 0,535 μ correspondendo ao verde-azul,
- 0,565 μ correspondendo ao verde-amarelo.

Do ponto de vista físico, as cores do espectro visível podem ser consideradas como ondas eletromagnéticas na faixa aproximada de 400 (ultravioleta) a 700 (infravermelho) nanômetros e compreendem as cores: violeta, azul, anil, verde, amarelo, alaranjado e vermelho; como apresentado na Figura 2.5.

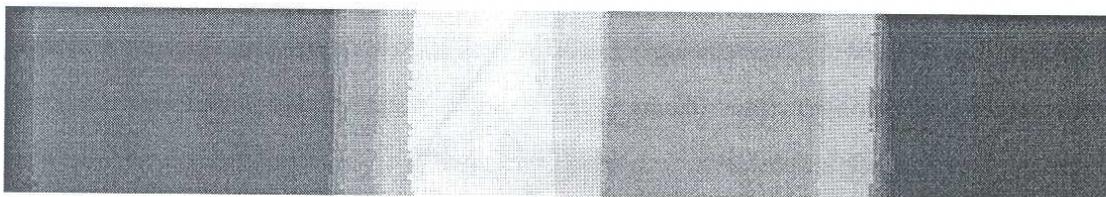


Figura 2.5: Cores do espectro visível pelo olho humano. (Fonte: [GRE79])

2.2.3. Percepção das Bordas Visuais: A Distribuição Espacial da Luz

O olho humano está equipado para ver os limites espaciais dos objetos, ou sejam, as bordas. A definição de borda visual designa a fronteira entre duas superfícies de luminância diferente para um dado ponto de vista. Considere que o olho humano dispõe de estruturas acumuladoras para os receptores situados no centro do campo de uma célula do nervo ótico e estruturas inibidoras para os receptores situados no contorno, como esquematizado na Figura

2.6-a. Assim, estas células reagem à presença de uma borda luminosa, sendo possível detectar: fendas, linhas e bordas, respectivamente como na Figura 2.6.

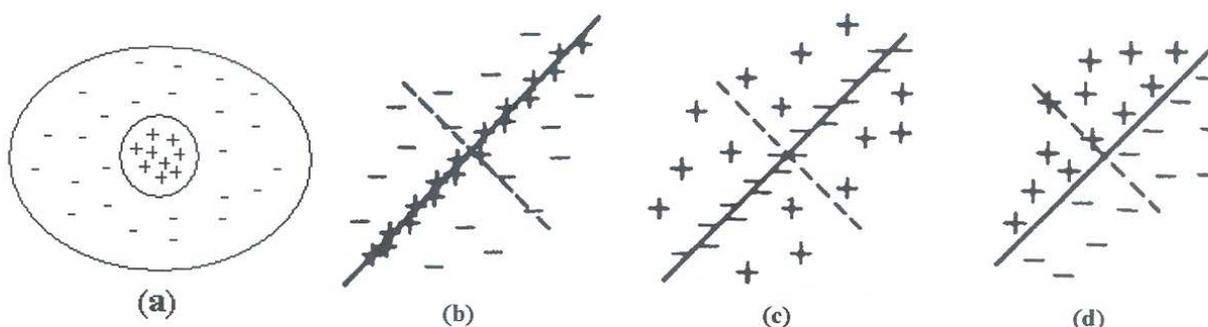


Figura 2.6: Percepção das bordas visuais: a) esquema das estruturas acumuladoras e inibidoras, b) fendas, c) linhas, d) bordas. (Fonte: [AUM93])

Deve-se, então, através de uma simplificação explicar que o córtex (região do cérebro que é a sede da percepção visual) é composto de células organizadas em colunas, cada uma correspondendo a uma região da retina. Essas colunas, por sua vez, agrupam-se em hipercolunas. Cada célula fornece uma informação que por si só é ambígua, e precisa ser combinada à das outras células da mesma hipercoluna, sem que se tenha um verdadeiro modelo dos mecanismos de ponderação que essa combinação implica. Cada coluna está “ajustada” para um padrão de direção do espaço, ou seja, é especializada na detecção de estímulos luminosos na referida direção. Como complementação, as hipercolunas são ainda especializadas, podendo avaliar o comprimento de uma linha e até mesmo um ângulo.

Pode-se concluir, que o sistema visual trabalha “por construção” através de instrumentos capazes de reconhecer uma borda visual e sua orientação, uma fenda, uma linha, um ângulo e um segmento.

2.2.4. Percepção do Contraste: Luminosidade x Bordas

O sistema visual está apto a detectar não somente a luminância, mas também a mudança de luminância. Ou seja, a luminância de uma superfície é determinada pela relação desta com o meio ambiente luminoso, sendo função principalmente do seu fundo. A Figura 2.7 exemplifica que objetos iluminados de modo idêntico serão julgados mais ou menos luminosos diante do fundo ser mais ou menos escuro. Os círculos centrais têm a mesma luminância, porém sua luminosidade aparente é diferente. O círculo sobre o fundo escuro aparece mais luminoso que o círculo sobre o fundo claro.

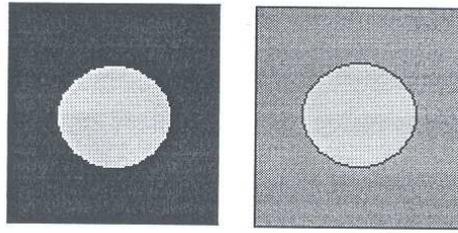


Figura 2.7: Percepção do contraste.

2.3. Teoria da Gestalt e a Percepção das Formas

Se você leu "NÃO PARE NA PISTA" na Figura 2.8, é melhor olhar de novo. Tem gente que olha várias vezes e não vê onde está o erro. Esse fenômeno deve-se a um fato muito comum de ver os padrões globalmente, sem notar os detalhes. Os psicólogos chamam isso de Gestalt.

Gestalt é uma Escola de Psicologia Experimental alemã que teve sua origem com Max Wertheimer (1880-1943), Wolfgang Köhler (1887-1967) e Kurt Koffka (1886-1941), por volta de 1910, atuando principalmente na área da teoria da forma, com contribuições significativas aos estudos da percepção, linguagem, memória, inteligência, entre outros. A teoria da Gestalt sugere uma resposta ao porque de certas formas agradarem mais que outras, mas não baseada no subjetivismo do "feio x bonito", mas sim apoiada na fisiologia do sistema nervoso e na psicologia, sempre através de rigorosos experimentos e pesquisas.



Figura 2.8: Placa de sinalização. (Fonte: <http://omnis.if.ufrj.br/~coelho/DI/olho.html>)

A palavra alemã *Gestalt* não é facilmente traduzível para o português. "boa forma" ou "forma regular" são as expressões que mais se aproximam do significado original, que todavia constitui uma combinação de vários elementos para formar um todo. O princípio enunciado por Wertheimer sobre a organização perceptiva demonstra que o olho humano tende a agrupar

as várias unidades de um campo visual para formar um todo. Este princípio conceitua a visão como uma experiência criativa, não como um simples ato de ver [GOM02].

Segundo essa teoria, o organismo percebe um conjunto de elementos como uma forma completa em que os componentes estão integrados entre si, de um modo que não é possível decompô-los sem destruir o conjunto. Ou seja, um conjunto passa a ser uma nova entidade, que não é simplesmente a soma dos seus componentes, mesmo porque esses componentes podem ser dispostos de diferentes maneiras, formando diferentes conjuntos [GOM02].

Segundo a Gestalt, a percepção da forma pelo cérebro é sempre uma percepção global dos estímulos, ou seja, o cérebro não enxerga elementos isolados, e sim as relações entre eles. Portanto, enxergamos o todo e não partes dele. A hipótese da Gestalt para explicar estas forças integradoras é uma estruturação natural do sistema nervoso, que tende a organizar as formas em todos coerentes e unificados, em busca de sua própria estabilidade. Para a percepção humana, não existe qualidade absoluta de cor ou forma, existem apenas relações.

Através dos estudos, os gestaltistas perceberam a presença de certas constantes nessas forças integradoras, que tentam explicar porque o ser humano vê as coisas. Apresenta-se a seguir os princípios da Gestalt com os quais pode-se detalhar estas forças de percepção. Deve-se ressaltar, que não se pretende analisar a validade dos princípios da Gestalt, e sim utilizá-los no presente estudo buscando relacionar a percepção humana e o reconhecimento de palavras manuscritas.

2.3.1. Os Princípios da Gestalt

O princípio da proximidade descreve a tendência dos elementos individuais de serem intensamente associados com os elementos mais próximos do que com aqueles que estão distantes. Esse fenômeno pode ser observado em dois níveis diferentes na Figura 2.9-a. Os olhos organizam primeiro os pontos em quatro colunas porque a separação horizontal é maior que a separação vertical. Então, devido a separação entre as duas "colunas" do meio ser maior que as brechas externas, a figura toda é vista como dois grupos de duas colunas cada.

O princípio da similaridade observa que alguns elementos são associados com mais intensidade quando eles compartilham de características visuais básicas (como é o caso das variáveis visuais de forma, tamanho, cor, textura, valor e orientação) do que quando eles diferem nessas dimensões. Assim, a Figura 2.9-b novamente aparece como dois grupos de duas colunas cada, apesar do espaçamento interelementos e intercolunas terem sido igualados.

O princípio da continuidade descreve a preferência pelos contornos contínuos e sem quebra ao invés de outras combinações mais complexas, mas igualmente plausíveis de figuras

mais irregulares. A forma da Figura 2.9-c é então percebida como duas linhas que se cruzam ao invés de quatro linhas que se tocam ou dois (ou mesmo quatro) ângulos opostos.

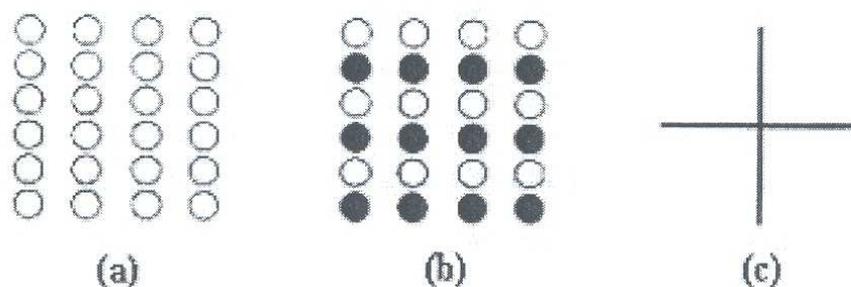


Figura 2.9: Princípios da Gestalt: a) proximidade, b) similaridade e c) continuidade.

O princípio do fechamento descreve a poderosa tendência humana de interpretar o estímulo visual como completo, como figuras fechadas, até quando algumas das informações de contorno estão ausentes. A Figura 2.10-a é facilmente vista como um triângulo sobreposto em três círculos completos mesmo que nenhuma destas formas esteja tecnicamente presente. Este exemplo clássico é conhecido como Triângulo de Kanizsa. A Figura 2.10-b exemplifica o princípio do fechamento para a leitura de palavras incompletas [MAC96].

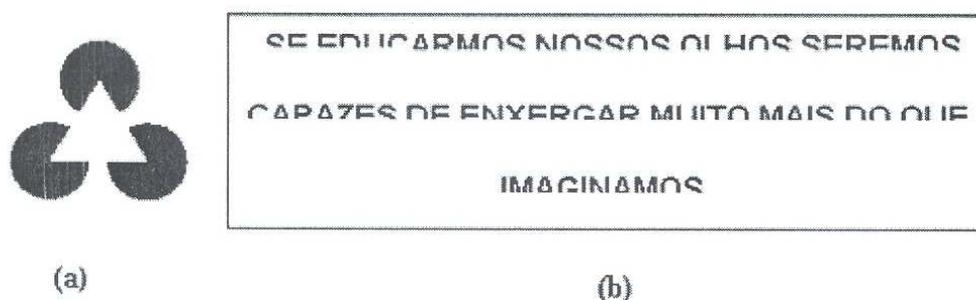


Figura 2.10: Princípio do Fechamento: a) triângulo de Kanizsa e b) palavras incompletas. (Fontes: <http://www.psicologia.freesevers.com/gestalt/> e [MAC96])

O princípio do fechamento é importante para a formação de unidades, uma vez que a Gestalt parte da teoria que a primeira sensação já é de forma, já é global e unificada. Não se vê partes isoladas, mas relações. A Figura 2.11 exemplifica este princípio, pois se percebe facilmente um cavalo, um triângulo e uma mulher de vestido preto. Comprova-se assim, que existe a tendência psicológica de se unir intervalos e estabelecer ligações.

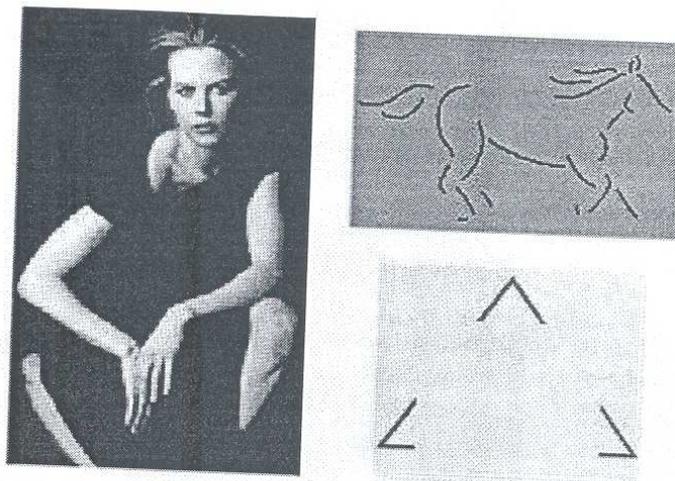


Figura 2.11: Princípio da Gestalt: fechamento .

(Fonte: http://www.belasartes.br/aulas_virtuais/joagomes/gestalt/leitura-visual/)

O princípio de *figura-fundo* tem por base que a atenção visual do observador alterna *entre* o as partes brancas e pretas de uma figura. Por exemplo, na Figura 2.12 o observador pode ver inicialmente um vaso branco em fundo preto e se alternar seu objeto de interesse poderá ver duas faces pretas uma de frente para a outra. Este exemplo clássico foi apresentado por Rubin em 1921 e é conhecido como *Face-Vase*. Ambos podem ser vistos como figura (objeto de interesse) ou como fundo sobre o qual está apoiada a figura. A Figura 2.13 apresenta outros exemplos onde a figura e o fundo podem ser confundidos entre si, podendo ser aceitos como figura ou fundo sem distinção alguma.

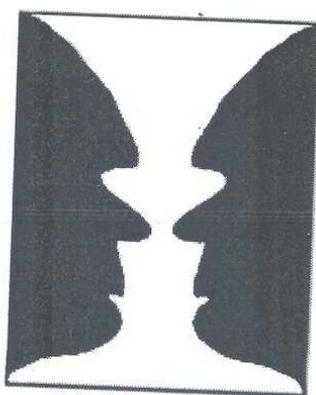


Figura 2.12: Princípio da Gestalt: figura-fundo.

(Fonte: <http://www.psicologia.freeservers.com/gestalt/>)

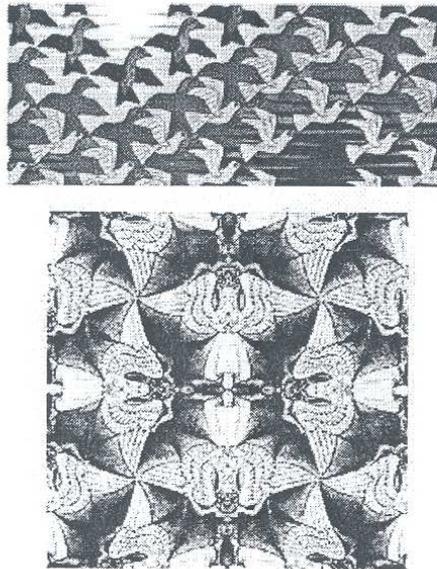


Figura 2.13: Exemplos do princípio da Gestalt: figura-fundo.
(Fontes: <http://www.prof2000.pt/users/essg/olhar/index.htm> e
<http://www.psicologia.freesevers.com/gestalt/>)

O princípio de *área* relata que a menor de duas figuras sobrepostas tenderá a ser interpretada como figura, enquanto que a maior será interpretada como fundo. Na Figura 2.14-a, o quadrado interno é percebido como uma forma distinta na frente de um quadrado maior, em vez de um buraco em uma forma maior. Sabe-se ainda, que o princípio da *simetria* descreve o agrupamento baseado nas propriedades emergentes da forma, ao invés das características das partes que a constituem, ou seja, a Figura 2.14-b é vista como dois objetos sobrepostos ao invés de três.

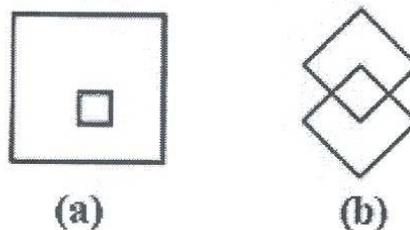


Figura 2.14: Princípio da Gestalt: a) área e b) simetria.

Outro ponto relevante é que a Gestalt afirma que as forças de organização tendem a se dirigir sempre à melhor forma possível, no sentido da clareza, unidade e equilíbrio, tanto quanto permitirem as condições dadas pela imagem, ou seja, quanto melhor a forma, mais pregnância ela terá, e melhor será sua relação com o cérebro. Este é o princípio da pregnância, o qual é geral e abrange todos os demais princípios. Uma imagem de boa Gestalt é enxergada com muito mais clareza pelo cérebro, e conseqüentemente de forma harmoniosa.

A Figura 2.15 apresenta diferentes formas de um mesmo objeto: "a letra A". Percebe-se que a imagem da esquerda possui maior pregnância, uma vez que, pode-se dizer que sua relação com o nosso cérebro é melhor. O mesmo princípio pode ser constatado na Figura 2.16 ao se perceber que alguns estilos de letras são mais facilmente compreendidos, e no presente estudo, mais facilmente lidos pelo observador.

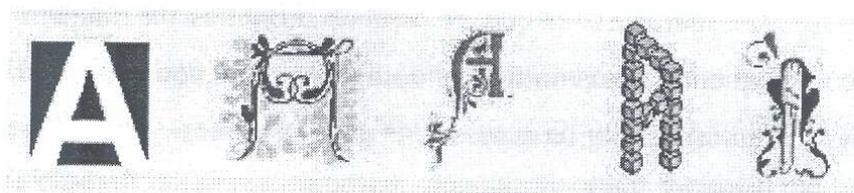


Figura 2.15: Princípio da Gestalt: pregnância.
(Fonte: http://www.belasartes.br/aulas_virtuais/joaogomes/gestalt/leitura-visual)

Finalmente, relembra-se que para a Gestalt a noção de unidade é primordial, pois para a percepção as partes são inseparáveis do todo e, ainda, que a imagem percebida é o resultado da relação entre sujeito-objeto, ou seja, da interação das forças externas (luz na retina) com as forças internas (a tendência de organizar da melhor forma possível os estímulos externos). Assim, afirmou Wertheimer em 1910: "O todo é mais que a soma das partes".

Pregnância da forma

Pregnância da forma

Pregnância da forma

Pregnância da forma

Pregnância da forma

Pregnância da forma

Pregnância da forma

Pregnância da forma

Figura 2.16: Princípio da Gestalt: pregnância.
(Fonte: http://www.belasartes.br/aulas_virtuais/joaogomes/gestalt/leitura-visual)

2.3.2. Constância Perceptiva

Espera-se do mundo que o mesmo tenha sempre a mesma aparência ou espera-se que se encontre uma certa quantidade de elementos invariáveis. É a percepção desses aspectos invariantes do mundo (tamanho dos objetos, forma, localização, orientação, propriedade das

superfícies, etc) que se designa pela noção de constância perceptiva. Ou seja, apesar da variedade de percepções, localiza-se as constantes [AUM93].

Ao olhar uma imagem pode-se entender este princípio, proposto por Boring em 1964. Um experimento simples, realizado por John H. Krantz no site <http://psychlab1.hanover.edu/Classes/Sensation> demonstra o poder do tamanho constante em imagens. Observa-se na Figura 2.17-a que as pessoas presentes na imagem são percebidas como tendo aproximadamente o mesmo tamanho, independentemente da distância ou posição ocupada na imagem. Krantz selecionou a pessoa de camisa branca e deslocou-a até as pessoas que se encontram em primeiro plano na imagem, como mostrado na Figura 2.17-b. Não existe nenhum truque ou modificação na imagem. A percepção de tamanho entre as diferentes pessoas na imagem, na realidade não existe. Para o observador todas as pessoas possuem aproximadamente o mesmo tamanho.

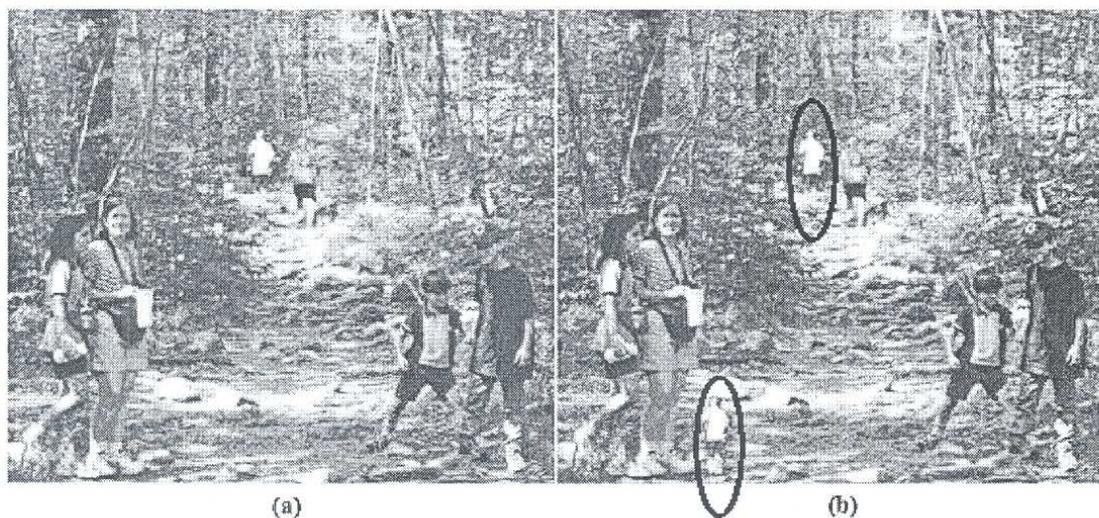


Figura 2.17: Constância perceptiva: a) posição original b) posição deslocada na imagem.
(Fonte: <http://psychlab1.hanover.edu/Classes/Sensation>)

2.3.3. Ilusões Ópticas

Uma imagem produz uma ilusão quando seu observador descreve uma percepção que não concilia com determinado atributo físico do estímulo visual. Alguns exemplos clássicos de ilusão podem ser citados [AUM93]:

- ilusão de Miller-Lyer: Qual a linha mais comprida na Figura 2.18? As duas linhas possuem o mesmo comprimento, mas as setas nas extremidades da linha criam a ilusão que a linha inferior é mais comprida. A distorção da perspectiva é criada no cérebro e não no olho;



Figura 2.18: Ilusão de Miller-Lyer.

- ilusão de Hering: Na Figura 2.19 as linhas horizontais são retas e paralelas? A percepção diz que não, mas são retas e paralelas. O cérebro realiza uma interpretação tridimensional da imagem fazendo parecer que estas linhas são curvas e não paralelas;

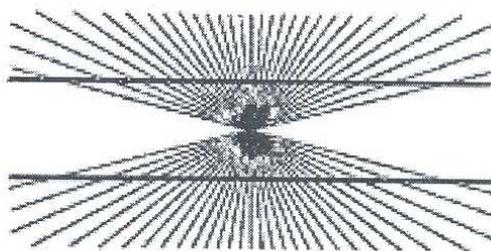


Figura 2.19: Ilusão de Hering.

(Fonte: http://www.daniloap.hpg.ig.com.br/ciencia_e_educacao/8/index_int_2.html)

- ilusão de Ponzo: Qual das linhas horizontais na Figura 2.20 é maior? A superior ou a inferior? Nenhuma. Ambas têm o mesmo comprimento. O que ocorre é um efeito perceptivo induzido pela convergência das linhas oblíquas. Assim, a linha superior é vista como mais distante e tendo projeção tão grande quanto a linha inferior é, portanto, vista como maior;

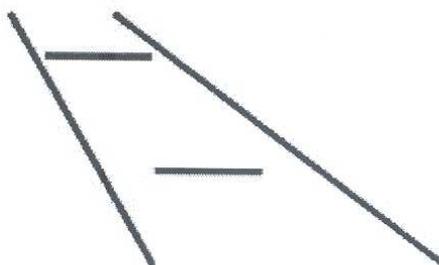


Figura 2.20: Ilusão de Ponzo.

- ilusão de Poggendorf: As Figuras 2.21-a e -b são exatamente as mesmas, exceto pela área escura no centro da Figura 2.21-b. Na figura b percebe-se duas linhas paralelas, uma azul e uma vermelha. Porém, sabe-se que a imagem original é a Figura 2.21-a, criando uma ilusão de continuidade onde não existe.

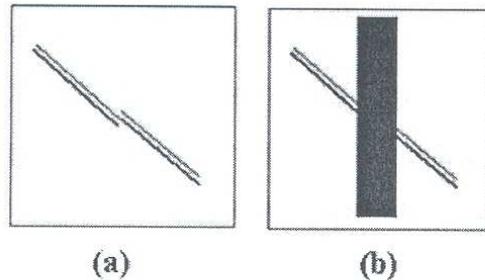


Figura 2.21: Ilusão de Poggendorf.

Outros exemplos de ilusões estão apresentados na Figura 2.22. Os azulejos da Figura 2.22-a compõem-se por quadrados com faces pintadas com tons diferentes, possibilitando a leitura de cubos sobrepostos e criando um jogo de ilusão óptica. A Figura 2.22-b dá ao observador a sensação de que a imagem está mudando constantemente, pois a cada instante que isto acontece, a nova imagem se sobrepõe sobre a anterior, causando uma ilusão de óptica.

Na Figura 2.22-c é impossível não começar a ver coisas. A configuração formada pelas linhas e colunas brancas entre os quadrados pretos, faz com que o observador comece a ver pequenos círculos pretos que se movimentam nas esquinas das linhas e colunas brancas. A Figura 2.22-d é outro exemplo de ilusão, pois o fato da figura superior não estar completa dá ao observador a percepção de um cubo, enquanto na figura inferior percebem-se partes seccionadas de um desenho geométrico.

A Figura 2.23 reúne outros exemplos que colocam nosso cérebro à prova. Tente contar quantas pernas o elefante da Figura 2.23-a possui? Difícil, não. Agora, fixe atentamente o ponto preto da Figura 2.23-b. Com o passar do tempo, você terá a sensação de que a neblina ao redor do ponto preto parecerá encolher. Tente mais esta: as linhas da Figura 2.23-c são inclinadas ou paralelas? Para sua surpresa, são paralelas. E, finalmente, tente contar os pontos pretos na Figura 2.23-d. Quantos são? Bem, se você conseguiu contar algum ponto preto, tente explicar como fez isto, pois não existe nenhum ponto preto nesta figura. Todos estes exemplos mostram que muitas vezes o que se vê não está lá e, ainda, que muitas vezes não se sabe explicar o que se está vendo.

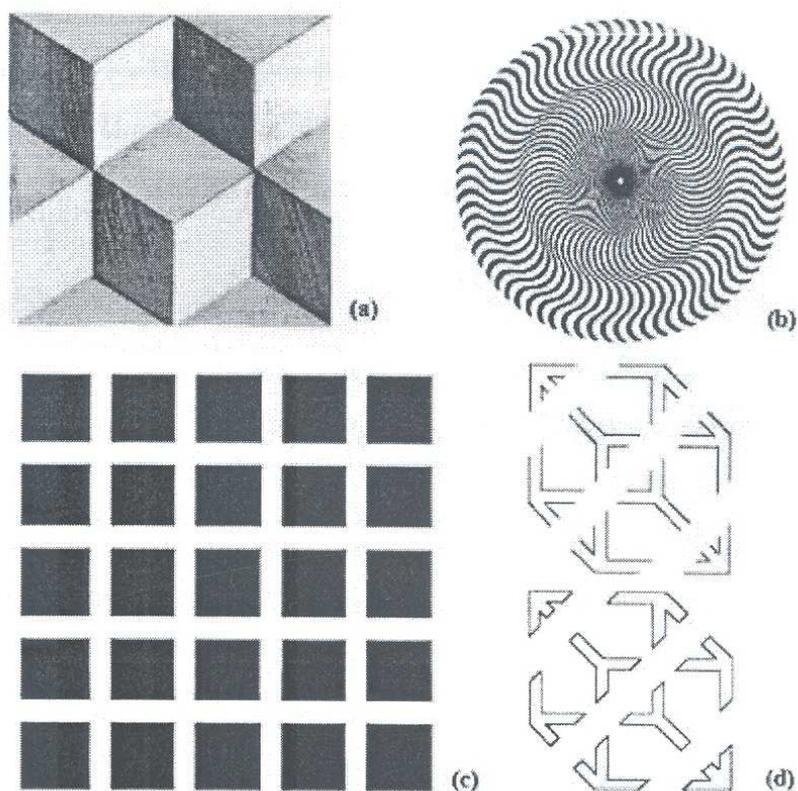


Figura 2.22: Exemplos de ilusões ópticas: a) azulejos, b) movimento em espiral, c) inibição lateral d) cubos. (Fontes: <http://www.prof2000.pt/users/essg/olhar/index.htm>, <http://educar.sc.usp.br/optica/curiosid.htm#cego>, http://www.daniloap.hpg.ig.com.br/ciencia_e_educacao/)

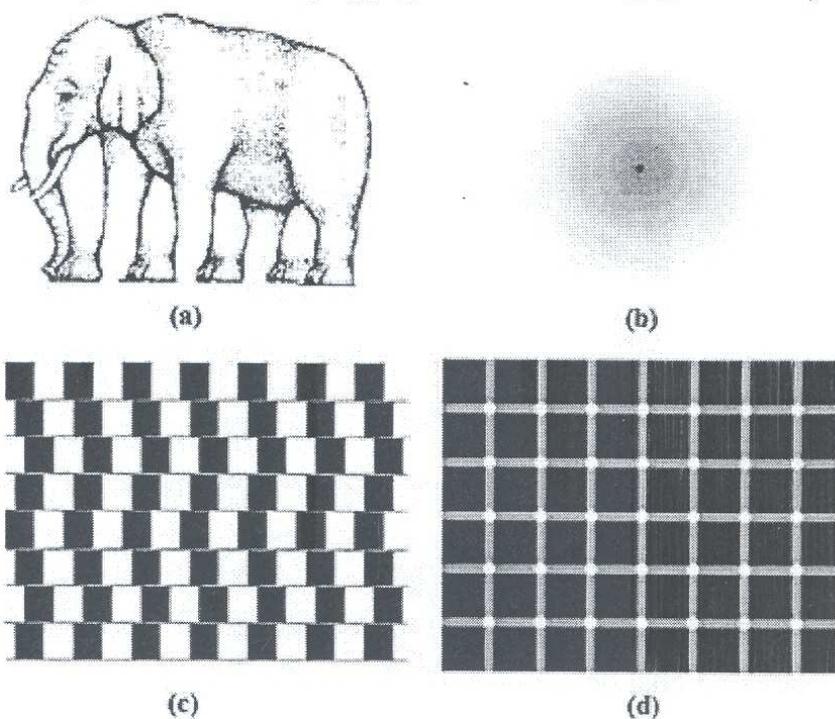


Figura 2.23: Exemplos de ilusões ópticas: a) pernas do elefante, b) ponto e neblina, c) linhas paralelas d) pontos pretos.

Deste modo, as ilusões permitem que se entenda a capacidade que o sistema visual tem de associar às imagens características não-visuais, em particular características espaciais. Deve-se levar, então, em consideração que a imagem é percebida, de modo automático, por uma interpretação em termos espaciais e tridimensionais [AUM93].

2.4. Captura do Essencial

Pode-se, então, perguntar, tal qual Arnheim [ARN97]: Se a visão é um processo de captura ativa da informação, o que a visão realmente apreende? Todos os inúmeros elementos de informação? Ou alguns deles?

Sabe-se, por experiência própria, que os olhos estão equipados para ver pequenos detalhes. Alguns traços relevantes podem identificar um objeto ou uma pessoa. Por exemplo, quando um cartunista cria uma semelhança expressiva de uma pessoa utilizando somente algumas linhas bem escolhidas. Ou ainda, quando se reconhece um amigo a grande distância unicamente pelas proporções e movimentos mais elementares.

Durante o processo de percepção de um objeto, os olhos humanos se movimentam sucessivamente e se fixam nas partes mais relevantes da imagem. Um exemplo clássico é a Figura 2.24 apresentada em [AUM93], ilustrando os movimentos dos olhos quando da busca visual.

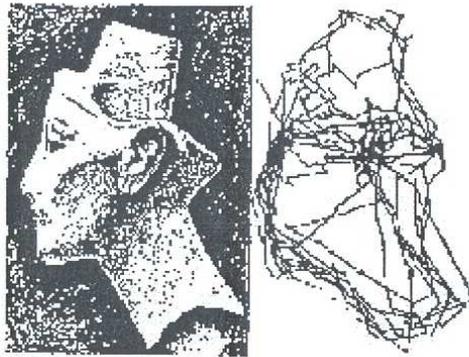


Figura 2.24: Movimentos oculares. (Fonte: <http://www.rybak-et-al.net/vnc.html>)

Para [ARN97] capta-se um rosto humano, exatamente como todo o corpo é captado, ou seja, como um padrão global de componentes essenciais (olhos, nariz, boca) aos quais se pode agregar detalhes. E ao decidir se concentrar nos olhos, também, a percepção funcionará de modo global. Assim, um olho é um olho (íris circular com pupila central escura, rodeada ela moldura acanoada das pálpebras ciliadas [ARN97]).

Os pesquisadores do A. B. Kogan Research Institute for Neurocybernetics da Rostov State University - Rússia [RYB02] demonstram através de um procedimento que utiliza uma câmera para descrever os movimentos oculares realizados pelo ser humano quando da percepção e reconhecimento de faces. Este procedimento comprova que uma imagem (no caso, um rosto ou face) é processada pelo cérebro humano através da fixação seqüencial em pontos relevantes da imagem, como mostrado na Figura 2.25. Algumas etapas do processamento foram suprimidas, de modo a simplificar a figura apresentada.

A idéia principal não é afirmar que a visão negligência detalhes, e sim, que a percepção visual ocorre, de fato, no cérebro humano. E ainda, que quando falta no objeto observado o sentido do todo, de integridade, ou seja, quando se vê o objeto como um aglomerado de partes, os detalhes perdem o significado e o todo se torna irreconhecível [ARN97].

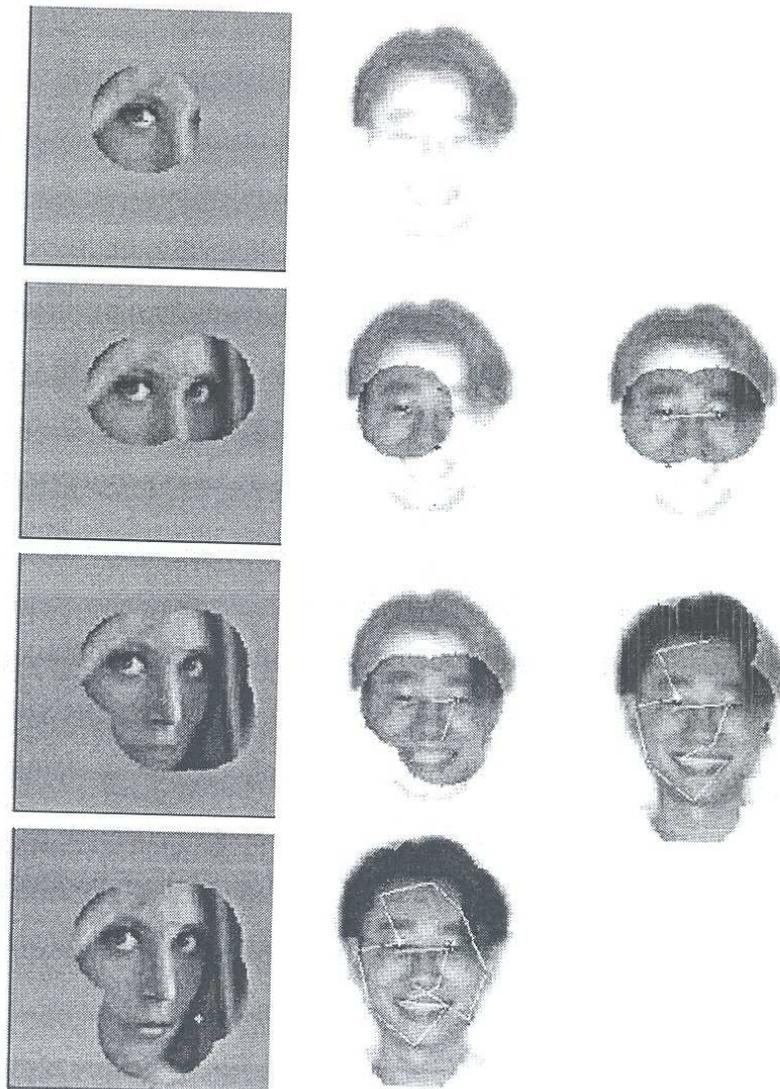


Figura 2.25: Movimentos oculares durante a percepção visual.
(Fonte: <http://www.rybak-et-al.net/vnc.html>)

2.5. Abordagens da Percepção Visual

A discussão e os estudos da percepção visual ocorrem sob duas abordagens [AUM93]: a abordagem analítica e a abordagem sintética. A seguir apresenta-se cada uma das abordagens de modo a favorecer o entendimento da percepção visual e das abordagens que a estudam.

2.5.1. A Abordagem Analítica

Esta abordagem consiste em partir de uma análise da estimulação do sistema visual pela luz, buscando fazer com que os componentes assim isolados correspondam a diversos aspectos da experiência perceptiva dela [AUM93]. Esta abordagem vem ao encontro da idéia de que o cérebro possui células especializadas nas funções elementares (percepção das bordas, linhas, dos movimentos direcionais, entre outros).

A abordagem analítica encontra sua representação através das teorias “combinatórias” ou “algorítmicas”, que acreditam que combinações suficientemente complexas podem analisar os fenômenos naturais. Para estas teorias, de acordo com Aumont [AUM93], o sistema perceptivo gera verdades perceptivas, de acordo com a realidade do mundo circundante, que permitem a predição, ao combinarem variáveis segundo certas regras.

Além disto, estas teorias consideram a informação contida na projeção retiniana como insuficiente para a percepção exata dos objetos no espaço, e sim, que esta necessita do recurso de outras fontes, tais como: as variáveis intrínsecas, deduzidas da análise da informação retiniana e as variáveis extrínsecas, ligadas a outros acontecimentos (por exemplo: movimentos dos olhos e memória). Portanto, estas teorias insistem muito na *aprendizagem* que leva a associar e integrar informações heterogêneas.

Esta hipótese consiste em supor que, entre uma família de objetos, o observador escolhe uma e só uma opção. Assim, em seguida aplica-se repetidamente esta hipótese na base da “tentativa e erro”. Portanto, se uma opção se revela errônea, o sistema visual revê suas opções de invariância e emite outras opções, de forma a fazer coincidir todas as opções com uma configuração possível (aqui o sistema leva em conta a experiência adquirida e as associações possíveis). Perceba que a hipótese de invariância está ligada ao fenômeno de constância perceptiva.

2.5.2. A Abordagem Sintética

Para esta abordagem a imagem óptica na retina contém toda a informação necessária à percepção dos objetos no espaço. Assim, a abordagem sintética se opõe as teorias que supõem uma aprendizagem da visão [AUM93]. Vale ressaltar que se entende como aprendizagem da visão a capacidade do sistema visual como um todo aprender, ou seja, inclui-se aí o cérebro.

Esta abordagem tem por base a teoria de Gibson [AUM93], a qual considera as transformações da projeção retiniana como um todo indissociável, não analisável. Para esta teoria, cada imagem provoca uma percepção global única. Sendo assim, não cabe ao sistema visual decodificar as informações, nem tão pouco, construir percepções, mas sim extrair informações. A percepção é, então, uma atividade direta.

2.5.3. Qual é a Melhor Abordagem?

A resposta desta questão pode parecer óbvia. Não existe uma melhor abordagem. Ambas também não podem ser consideradas contraditórias, como mencionado em [AUM93]. Concorda-se com [AUM93] que o estágio dos estudos atuais levam a pensar que a compreensão do funcionamento do cérebro não se baseia unicamente na seriação de grande quantidade de unidades especializadas, mas em determinados processos globais, ou não-localizados, que implicam paralelismo e/ou transversalidade.

Esta afirmação pode ser constatada quando no desenvolvimento de sistemas de reconhecimento de palavras manuscritas, busca-se a integração de diferentes fontes de conhecimento para melhorar as taxas de acerto dos sistemas [OLI02-a] [OLI02-b].

Para palavras manuscritas, o reconhecimento consiste em a partir de uma forma (uma palavra) desconhecida estabelecer sobre um conjunto de formas conhecidas, quais entre as formas desse conjunto mais se assemelha a forma desconhecida, e sobre tudo isso tomar uma decisão da melhor hipótese. Essa decisão é efetuada geralmente medindo-se a semelhança da forma desconhecida com um conjunto de referência (ou modelos) armazenados na memória e descritos em uma representação análoga. As referências ou modelos são obtidos através da etapa denominada de treinamento.

Nos sistemas de reconhecimento, cada produção manuscrita de uma palavra dada é ligada a uma forma particular. Após a fase de aquisição do sinal, a extração de primitivas permite transformar a forma em uma representação, mais fácil de manipular que a forma original. O reconhecimento consiste em decodificar essa representação atribuindo-se uma das classes conhecidas ou referências do sistema.

Geralmente o reconhecimento de palavras manuscritas - HWR - *Handwritten Word Recognition* baseia-se em duas abordagens do problema: *local* ou *analytical approach* realizada em nível dos caracteres [LEC90], [FAV92], [GIL95] e [KIM96] ou *global approach* em nível das palavras [GUI95] e [CÔT97].

Na primeira abordagem, o sistema se depara com a necessidade de segmentação das palavras em caracteres/letras ou pseudo-letras. Esta abordagem caracteriza-se pela dificuldade em se definir a fronteira entre os caracteres. E, portanto, o sistema de reconhecimento dependerá também do sucesso do processo de segmentação. Assim, os sistemas serão constituídos por etapas distintas para segmentação e reconhecimento, ou associando segmentação e reconhecimento em uma mesma etapa.

Por outro lado, a abordagem global (*global approach*) permite evitar a etapa de segmentação das palavras, extraíndo-se primitivas globais das palavras sem necessidade de segmentação explícita das mesmas. Esta abordagem procura explorar as informações do contexto das palavras, sendo que aspectos baseados em modelos psicológicos podem ser inseridos [GUI95] e [CÔT97]. Porém, é uma abordagem restrita às aplicações de léxico com pequenas dimensões.

As duas abordagens, local e global, não são estanques podendo-se trabalhar com uma abordagem híbrida [CHE93]. Deste modo, é possível combinar as vantagens e diminuir as desvantagens.

Assim, considerando-se as possíveis abordagens do problema, a etapa de extração de primitivas caracteriza-se por uma das tarefas importantes no sucesso do sistema de reconhecimento de palavras manuscritas. A função de um conjunto selecionado de primitivas é diminuir a variabilidade de formas intra-classes e aumentar a discriminação inter-classes. Diversos tipos de primitivas podem ser combinados visando obter sistemas mais robustos através de um conjunto pequeno de primitivas, porém, com alta discriminação das classes consideradas [TRI96] e [YAC99]. É aqui, que o estudo da percepção visual pode contribuir com os sistemas de reconhecimento.

2.6. Princípio da Maior Probabilidade ou o Direito da Incerteza?

De tudo o que foi apresentado, pode-se concluir que a percepção visual é composta por dois atos: o ato de ver e o ato de compreender. O ver ocorre através do olho e suas estruturas e o compreender ocorre no cérebro. Sabe-se, também, que o ser humano pode

localizar e interpretar regularidades nos fenômenos luminosos que atingem os olhos: percepção da luminosidade, da cor, de bordas visuais e do contraste.

E ainda, que para a Gestalt a percepção da forma pelo cérebro é sempre uma percepção global dos estímulos, ou seja, o cérebro não enxerga elementos isolados, e sim as relações entre eles. Constata-se que as imagens são objetos visuais paradoxais. Possuem duas dimensões, mas permitem que sejam vistos objetos em três dimensões.

Deste modo, a idéia de [AUM93] afirmando que a percepção visual implica em mecanismos de paralelismo e/ou transversalidade das informações extraídas da imagem, fornecem a abertura suficiente do assunto para que se possa incluir neste processo (ver e compreender) o princípio da maior probabilidade ou o direito da incerteza.

Para isto, as abordagens da percepção: analítica e sintética se fundem para que o cérebro possa escolher não uma e só uma opção entre as possíveis, e sim, a opção mais provável com base em um processo de extração de informações da imagem observada e um banco de dados aprendido. Tal qual, é feito para os sistemas de reconhecimento automático de palavras manuscritas.

Neste caso, a imagem somente trará problemas ao observador se a mesma oferecer características contraditórias ou não for suficientemente informativa. O exemplo clássico *Duke-Rabbit* apresentado em [AUM93], como mostrado na Figura 2.26, demonstram o princípio da maior probabilidade. As figuras à esquerda e à direita representam com maior probabilidade um pato e um coelho, respectivamente. A figura ao centro, para algumas pessoas pode parecer um pato ou um coelho, com maior probabilidade. Ou será o direito da incerteza?

Assim, para [AUM93] deixa-se o domínio da percepção para entrar no domínio da cognição, ou seja, a seleção do objeto mais provável ocorre através da verificação de uma série de objetos, simbolicamente representados no córtex, já conhecidos e reconhecidos.

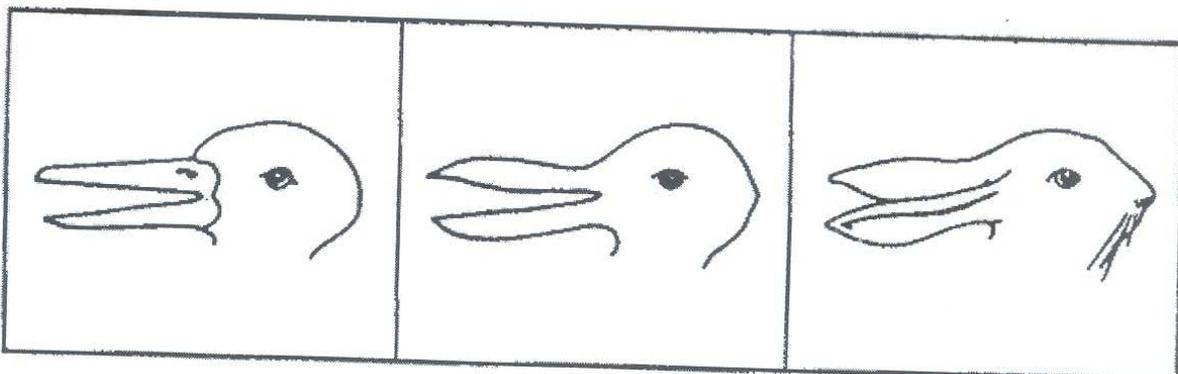


Figura 2.26: Princípio da maior probabilidade: a) pato, b) pato ou coelho? c) coelho.
(Fonte: [AUM93])

2.7. Discussão

A percepção é um processo ativo.

A Gestalt estuda o ato de ver como um processo dinâmico envolvendo o observador e o que é observado. As pessoas quando olham, organizam o que vêem. Elas analisam uma estrutura, resolvem ambigüidades e fazem conexões.

As pessoas organizam o que vêem.

Da Gestalt, se traz a idéia de que o campo visual é normalmente dividido em duas partes: a figura e o fundo. A figura é o todo que se observa e o fundo circunda a figura observada.

As pessoas agrupam as figuras dependendo das propriedades visuais das figuras.

Quando as pessoas percebem um campo visual, o padrão que emerge como figura, e não como fundo, depende das características dos elementos do campo e do relacionamento entre os elementos dentro deste campo.

As pessoas interpretam o estímulo visual como completo.

O princípio do *fechamento* descreve a poderosa tendência humana de interpretar o estímulo visual como completo, como figuras fechadas, até quando algumas das informações de contorno estão ausentes.

As pessoas podem adicionar uma dimensão: de 2-D para 3-D.

As pessoas tendem a ver os desenhos bidimensionais como tridimensionais, ou seja, movendo o plano de representação gráfica bidimensional de 2-D, para uma representação espacial sólida de 3-D. As ilusões são exemplos de inserção da terceira dimensão em figuras observadas.

Finalmente, tendo-se em mente estes pontos relevantes e sabendo-se que a percepção visual compreende os atos de ver e compreender e, ainda, que “o todo é mais que a soma das partes” (Wertheimer, 1910), pode-se concluir, tal qual [RYB02], que a percepção visual e o reconhecimento devem ser considerados como processos comportamentais e, provavelmente, não podem ser completamente entendidos pelos limitados esquemas dos sistemas

computacionais sem que se leve em consideração aspectos comportamentais e cognitivos destes processos. Portanto, para o presente estudo, busca-se no Capítulo 2 responder os seguintes questionamentos: Como o “ato de perceber” e o “ato de ler” estão associados? Que características da percepção visual auxiliam na leitura de palavras manuscritas?

Capítulo 3

Escrita e Leitura

O que você vê na Figura 3.1? Inicialmente, deve ter visto manchas pretas em um fundo branco. Ao olhar, em detalhes as manchas pretas, você pode notar que quatro palavras surgem entre as manchas pretas (“*seeing, illusion, mail box*”). Mesmo que não compreenda o significado das palavras (pois, estão em idioma inglês), sabe que são palavras.

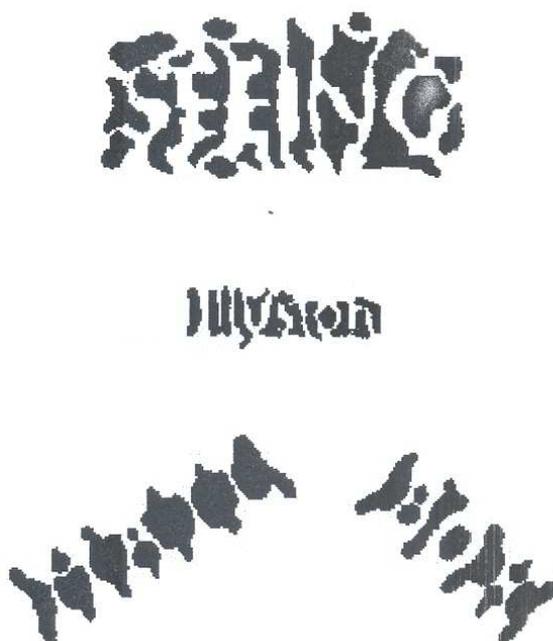


Figura 3.1: Percepção visual e leitura: manchas e palavras.
(Fonte: <http://www.psicologia.freeservers.com/ilusoes/organize.html>)

Observe a Figura 3.2 e tente ler a cor e não as palavras. Algum conflito? Você achou esta tarefa difícil? Bem, tudo isto é normal. O que está ocorrendo é um conflito no seu cérebro. Pois, o lado direito do seu cérebro tenta informar a cor e o lado esquerdo tenta ler as palavras.

AMARELO AZUL LARANJA
 PRETO VERMELHO VERDE
ROXO AMARELO VERMELHO
LARANJA VERDE PRETO
AZUL VERMELHO ROXO
VERDE AZUL LARANJA

Figura 3.2: Percepção visual e leitura: cores e palavras.

Estas figuras permitem que se considere o que foi apresentado no Capítulo 1 sobre percepção visual e, então, neste capítulo se resgate os aspectos relevantes dos processos de escrita e leitura humana, associados ao conhecimento prévio do ser humano. Desta forma, será possível fornecer subsídios ao estudo das relações existentes entre o processo como um todo realizado pelo ser humano e a automatização desses processos através do reconhecimento automático de palavras manuscritas.

3.1. Escrita

O que é escrever? Como concluir que algo é escrita? Qual a evolução da escrita? Como a leitura se processa? O que é conhecimento prévio? Como se pode utilizar este conhecimento? O que é ler? Ler e escrever são processos análogos? [SUE97]

Cada uma destas perguntas possui resposta e auxiliará na compreensão do problema do reconhecimento da escrita humana, mais especificamente, a escrita cursiva.

O que é escrever? A etimologia da palavra (em grego, *gráfein*, em latim, *scribere*) atesta a origem física da ação de escrever. No Aurélio encontra-se que escrever é representar por letras, redigir, compor, fixar, gravar, dirigir (cartas) e dirigir carta (a alguém). Segundo Cagliari [CAG91], escrever é também uma forma de expressão artística e até um passatempo. Ou ainda, citando Teberosky em [TEB96], escrever é uma atividade intelectual em busca de uma certa eficácia e perfeição, que se realiza por meio de um artefato gráfico-manual, impresso ou eletrônico, para registrar, comunicar, controlar ou influir sobre a conduta dos

outros, que possibilita a produção e não só a reprodução, e que supõe tanto um efeito de distanciamento como uma intenção estética.

A escrita, por sua vez, seja ela qual for, conforme Cagliari em [CAG91], tem como objetivo primeiro a leitura. A leitura é uma interpretação da escrita que consiste em traduzir os símbolos escritos em fala. Portanto, como concluir que algo é escrito? Teberosky em [TEB96], responde esta questão explicando que a escrita, ou sejam marcas gráficas produzidas “no lugar de” algo, é formada, no caso da escrita alfabética, por unidades de notação e pela combinação dessas unidades de forma a constituir um sistema. As unidades de notação, por sua vez, denotam a linguagem. Então, são usadas como marcas “no lugar de” alguma unidade de linguagem. Desta forma, conclui-se que algo é escrito por estar representado por unidades de linguagem aceitas e compreendidas pelo sistema em questão.

A história da escrita delinea sua evolução a partir das escritas pictóricas, das escritas ideográficas e das escritas alfabéticas. A escrita pictórica caracteriza-se por uma escrita através de desenhos ou pictogramas. Este tipo de escrita aparece em inscrições antigas, sendo exemplos os cantos Ojibwa da América do Norte, o catecismo asteca e mais recentemente a estória em quadrinhos, conforme Figura 3.3 extraída de Cagliari [CAG91].

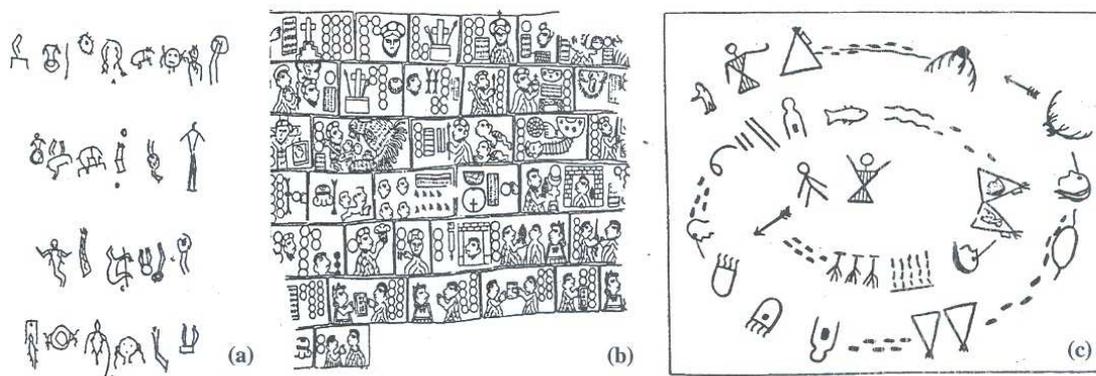


Figura 3.3: Exemplos de escrita pictórica:

a) cantos Ojibwa; b) catecismo asteca; c) inscrição antiga. (Fonte: [CAG91])

A escrita ideográfica baseia-se no uso de desenhos especiais chamados ideogramas, tem por princípio representar só o significado, referindo-se a um pensamento não necessariamente de natureza linguística, ver exemplos da Figura 3.4.

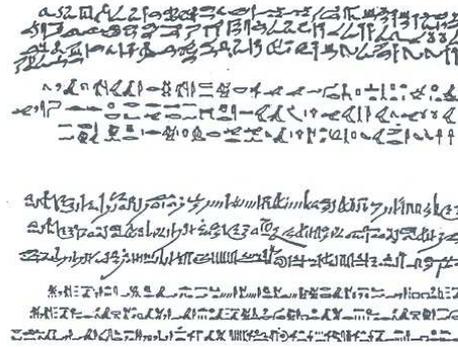


Figura 3.4: Exemplos de escrita ideográfica: escrita egípcia. (Fonte: [CAG91])

Ao longo dos tempos os ideogramas foram perdendo alguns dos traços mais representativos das figuras retratadas e tornou-se uma simples convenção de escrita. A origem das letras está nos ideogramas, mas estes perderam o valor ideográfico, assumindo uma nova função de escrita: a representação puramente fonográfica, ou seja, a escrita baseada no significante e não mais no significado. O ideograma perdeu seu valor pictórico e passou a ser simplesmente uma representação fonética [CAG91] [TEB96]. Atualmente, as letras do alfabeto utilizado derivam dessa evolução. E, portanto, a escrita alfabética se caracteriza pelo uso de letras. A Figura 3.5 apresenta um esquema da evolução da escrita alfabética, onde se mostra que os gregos, escrevendo consoantes e vogais, criaram a escrita alfabética.

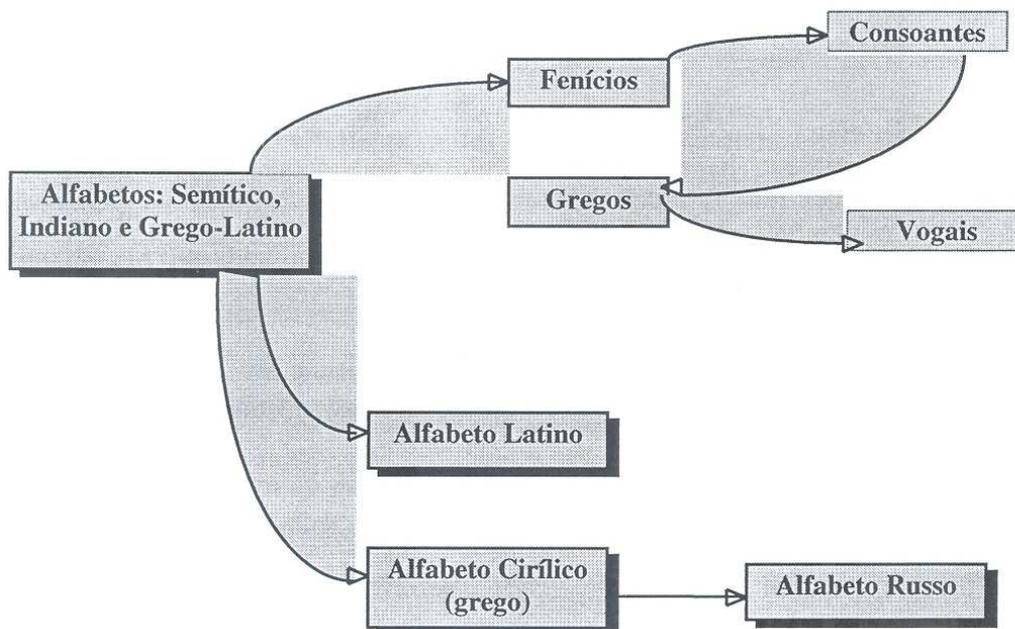


Figura 3.5: Evolução da escrita alfabética.

O uso da letra manuscrita iniciou-se na França, segundo Parramón [PAR91], durante o reinado de Luis XIV, época em que a França dominava quase toda a Europa. A letra

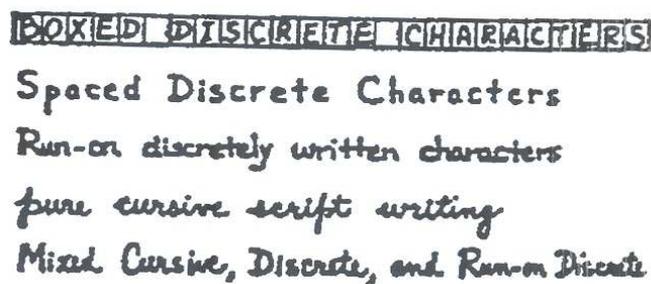
manuscrita francesa chamava-se *rondé*, era vertical e semelhante à letra gótica no seu início. Haviam ainda outros estilos: a letra italiana, denominada *italienne*, e a tipicamente *inglesa*, desenvolvida a partir da italiana com um desenho que perdura até os dias de hoje por seu estilo clássico. O estilo Mistral, membro da família dos tipos manuscritos ou caligráficos, é o estilo que permite a realização de um enlace perfeito entre todas as letras durante a escrita [PAR91], conforme Figura 3.6.



abcdefghijklmnopqrstuvwxyz

Figura 3.6: Estilo mistral: enlace perfeito entre as diversas letras. (Fonte: [SOL89])

De maneira geral, atualmente faz-se uso de um sistema alfabético, porém na verdade, segundo Cagliari em [CAG91], esse sistema não possui uma única forma e nem é completamente alfabético, pois se utilizam outros caracteres de natureza ideográfica, tais como os sinais de pontuação e os números. Segundo o autor, pode parecer paradoxal, mas pouco percebe-se ou explica-se que se faz uso de várias formas de representação gráfica, ou seja, vários tipos de alfabeto em uso, e em geral, de maneira misturada. A escrita cursiva é diferente da escrita de forma (letra tipo imprensa), e como mencionado entre um extremo e outro a escrita cursiva pode ser classificada em 5 categorias distintas [TAP90], conforme exemplificado na Figura 3.7.



BOXED DISCRETE CHARACTERS
Spaced Discrete Characters
Run-on discretely written characters
pure cursive script writing
Mixed Cursive, Discrete, and Run-on Discrete

Figura 3.7: Tipos de escrita. (Fonte: [TAP90])

À medida que um sistema alfabético é utilizado por um número grande de pessoas em diferentes lugares, para usos diversos, a forma das letras do alfabeto, que inicialmente era única, passa a admitir variantes [CAG91]. No mundo antigo, as variantes se restringiam a uns poucos casos. O latim, por exemplo, não tinha as letras maiúsculas. Hoje, encontramos uma enorme variedade de tipos de alfabetos, um exemplo é o Livro: *Lettering and Alphabets*, de

Cavanagh [CAV74], que apresenta 85 tipos de alfabetos completos, e ainda muitos outros [SOL89] [PAR91] [ROS91].

A primeira letra do alfabeto, a letra a, pode aparecer escrita de diferentes maneiras conforme Figura 3.8, sendo cada um dos exemplos pertencente a um tipo de alfabeto diferente. Segundo Cagliari, a letra A é tão diferente de a, quanto p é de m, por exemplo: p, b, d e q são muito mais semelhantes entre si do que b e B, g e G e assim por diante. Realmente, não se percebe a cada instante da escrita ou da leitura estas diferenças, para uma pessoa adulta, um A é um A, seja ele escrito como for [CAG91].



Figura 3.8: Exemplos da letra “A” em diferentes estilos.
(Fontes: [CAG91] e base de dados: LUCI/EXTENSO/PALAVRAS)

A escrita cursiva é a mais complicada das escritas existentes no mundo [CAG91][PAR91][ROS91], porque varia enormemente, seguindo as idiossincrasias de cada usuário. A escrita cursiva só é menos difícil para quem está acostumado com o escrever e com o modo de traçar as letras de quem escreveu, caso contrário, sabe-se muito bem, é difícil ler o que os outros escrevem e, às vezes, até mesmo o que nós próprios escrevemos [CAG91].

A Figura 3.9 apresenta as variações geradas na escrita cursiva com base em três tipos básicos de traçados apresentados por [CAV74]. O primeiro traçado origina uma escrita cursiva mais arredondada e com um laço extra na letra t, o segundo traçado é o oposto do primeiro originando uma escrita cursiva mais acentuada no que se refere as letras h, s e m. Por outro lado, o terceiro traçado origina uma escrita cursiva onde as letras minúsculas (corpo da palavra) é mais alto, apresentando um resultado intermediário entre o primeiro e segundo traçados. Cavanagh [CAV74] apresenta diferentes tipos de escrita cursiva e cita que muitas pessoas geram derivações próprias (diferentes inclinações, condensamento ou expansão dos caracteres) dos dois tipos de escrita cursiva: Formal e Informal. O objetivo de Cavanagh é ensinar a escrever, de forma manual, diferentes tipos de alfabetos para diferentes ocasiões ou necessidades, ou seja, tem o objetivo de tornar o leitor um calígrafo. Entre os 85 alfabetos apresentados por Cavanagh [CAV74] e os 100 alfabetos apresentados por Solo [SOL89]

apresentam-se na Figura 3.10 os alfabetos que mais se aproximam da nossa forma de escrever.

mmmm Christmas
 uuuu Christmas
 mmm Christmas

Figura 3.9: Variações da escrita com base no traçado. (Fonte: [CAV74])

A B C D E F G a b c d e f g h i
 H I J K L M j k l m n o p q r
 N O P Q R S T s t u v w x y z
 U V W X Y Z 1 2 3 4 5 6 7 8 9

A B C D E F G U V W X Y Z
 H I J K L M a b c d e f g h i j
 k l m n o p q r s
 N O P Q R S T t u v v w w x y z

<i>Mercury (Bold)</i>	<i>Mandate</i>	<i>Graphik</i>
A B C D E F G H I	A B C D E F G H	A B C D E F G H I J
J K L M N O P Q R S	J K L M N O P Q R S	T U V W X Y Z
T U V W X Y Z	T U V W X Y Z	
&	(E,.,!?\$)	
abcdefghijklmnop	abcdefghijklmnop	abcdefghijklmnopq
rstuvwxyz	opqrstuvwxyz	rstuvwxyz
1234567890	1234567890	1234567890

Figura 3.10: Estilos de escrita cursiva. (Fontes: [CAV74] e [SOL89])

Deve-se ressaltar que ao se alfabetizar uma criança é apresentado à ela um conjunto de letras que formam o alfabeto utilizado no Brasil, conforme Figura 3.11, o qual permitirá a criança concatenar perfeitamente letra a letra, e portanto, produzir palavras. Ao longo de sua vida cada criança irá gerar um estilo próprio de escrita, alternando várias vezes a forma de desenhar as letras até que na vida adulta defina um estilo pessoal de escrever.

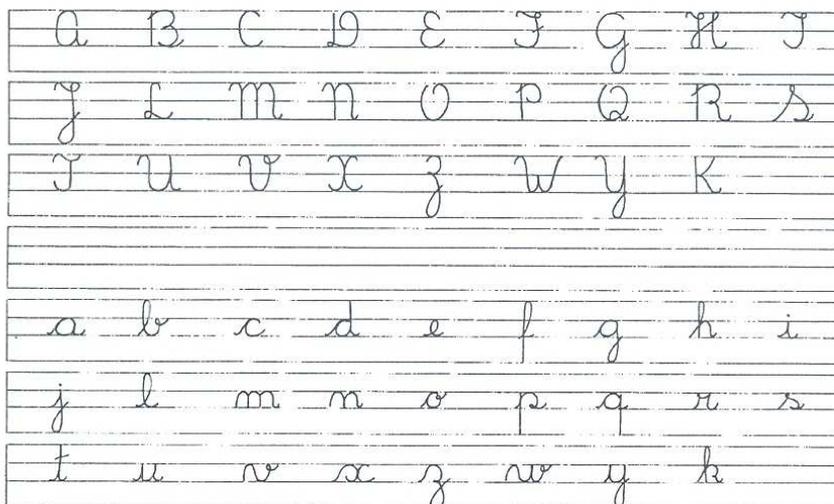


Figura 3.11: Alfabeto e estilo de escrita cursiva para alfabetização no Brasil.

3.2. Leitura

Agora, é possível analisar uma nova atividade, ou seja, a leitura. Para Kato [KAT90], as pessoas e escolas supõem que à produção segue-se automaticamente a recepção. Em outras palavras, se for ensinada a escrita, o aluno aprenderá automaticamente a ler. Contudo, cita Kato, que a leitura pode ser adquirida independentemente da escrita. O exemplo está em leitores que são proficientes leitores de uma língua estrangeira, porém, nada escrevem nessa língua. O autor não quer mostrar que a prática da produção não interfere favoravelmente na capacidade de recepção, mas sim, que existe uma interferência recíproca entre os dois processos.

Pergunta-se então: O que é ler? Por leitura se entende toda manifestação linguística que uma pessoa realiza para recuperar um pensamento formulado por outra e colocado em forma de escrita [CAG91]. A leitura é a realização do objetivo da escrita. No Aurélio encontra-se que ler é percorrer com a vista o que está escrito, proferindo ou não, mas conhecendo as respectivas palavras, conhecer, interpretar por meio da leitura, pronunciar em

voz alta, recitar, ver e estudar (coisas escritas), decifrar, interpretar o sentido de, explicar, adivinhar, predizer, prelecionar, ver e interpretar o que está escrito.

Segundo Machado e Teixeira em [MAC96], ao ler, associa-se às informações lidas com os conhecimentos que se têm armazenado em nosso cérebro e, então, se tem a capacidade de criar, imaginar e sonhar. Os autores citam que ler com velocidade torna a leitura mais proveitosa, estimula a mente e desenvolve o senso de antecipação e expectativa, torna-a mais participativa e não possibilita a fragmentação do assunto. Contudo, solicita do usuário o bom senso de ler com velocidade variada, dependendo da situação.

O processo de alfabetização faz com que as pessoas sejam submetidas a processos com enfoques fônicos, globais ou lingüísticos que, de uma forma geral, estimulam uma área muito pequena da leitura, visando as sílabas ou as palavras. Normalmente, a leitura é realizada letra a letra, sílaba a sílaba ou palavra a palavra, padrão mantido desde o período de alfabetização. A leitura acelerada propõe uma maior captação de elementos - um bloco de palavras que formam uma idéia, dinamizando a leitura.

Como apresentado no Capítulo 1, o olho humano é o instrumento do sentido da visão que possibilita a leitura, além dos dedos que, através do tato, permitem aos deficientes visuais decifram o texto em *braille*. Os olhos funcionam como uma máquina fotográfica sofisticada, levando o conteúdo captado para o cérebro, onde será interpretado. Então a visão, como os outros sentidos, é responsável por informar ao cérebro sobre o estímulo recebido. Um estímulo visual pode conter características como as apresentadas no Capítulo 1: luminosidade, cor, contraste, proximidade, continuidade, fechamento, entre outros. Assim, a percepção traduz as mensagens sensoriais que ela informou, para que ocorra a compreensão. Dois fatores interferem na percepção: os estímulos (com toda a sua gama de variedades) e os fatores pessoais. Desta forma, cada indivíduo reagirá ao meio à sua maneira.

Sabe-se que a maioria dos idiomas propõe uma leitura horizontalizada, da esquerda para a direita. E ainda, que para uma melhor concentração, não se pode estar preso a pequenos detalhes do texto. Pois, para escrever utilizam-se muitas palavras como adereço. Deve-se procurar ler o conjunto e perceber o seu significado. A Figura 3.12 apresenta os diferentes níveis de leitura pelos quais os autores Machado e Teixeira em [MAC96], julgam ser as etapas necessárias para que as pessoas possam ter condições de aproveitar totalmente o assunto lido, e ainda, citam que estas etapas são cumulativas ao longo da vida do leitor.

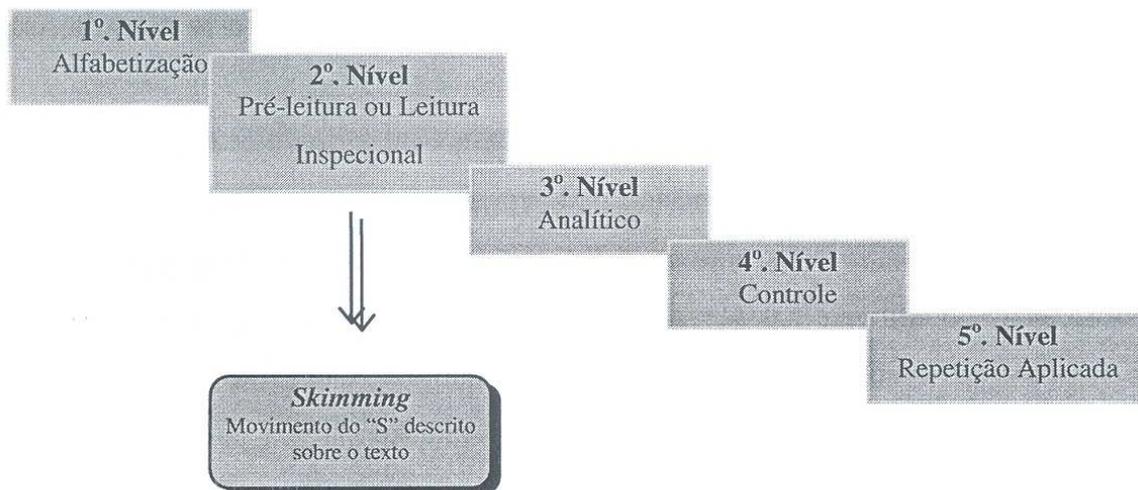


Figura 3.12: Níveis de leitura.

Durante os níveis 2 e 3 o cérebro levará em consideração o “universo de conhecimento do usuário ou leitor”, termo este utilizado como referência ao conjunto de fatores cognitivos que interferem na atividade de interpretação, segundo Trevisan em [TRE92]. Para a autora um texto não é coerente ou incoerente enquanto texto, mas é coerente para alguém numa situação específica. A construção da coerência depende fundamentalmente da bagagem cognitiva do usuário. De acordo com Trevisan [TRE92], citando Beaugrande e Dressler (1981), um texto “faz sentido” porque existe uma continuidade de sentidos entre os conhecimentos ativados pelas expressões do texto. Portanto, a coerência se estabelece a partir da interação entre a configuração de conceitos expressa no texto e o conhecimento do mundo armazenado pelos usuários.

3.3. Conhecimento Prévio

Tendo-se em mente, os aspectos relevantes dos processos de escrita e leitura apresentados, deve-se considerar como relevante o conhecimento prévio que o usuário carrega com si próprio e não se pode deixar de lado os aspectos do conhecimento linguístico, o conhecimento do mundo e o conhecimento partilhado, e ainda, a armazenagem e ativação do conhecimento na memória.

O conhecimento linguístico diz respeito aos elementos linguísticos, tais como os itens lexicais e as estruturas sintáticas. Com base neste conhecimento, ativado em memória, o usuário utilizará os elementos linguísticos como “pistas”, a fim de estabelecer uma relação entre o linguístico e o conceitual-cognitivo. Deve-se lembrar que o contexto linguístico está

associado também a este conhecimento, de modo que a interação entre estes níveis permitirá ao usuário apreender o sentido veiculado pelo texto.

A Inteligência Artificial fez com que a Linguística evolui-se no que se refere ao uso do conhecimento do mundo no estabelecimento da coerência e, conseqüentemente, na apreensão do sentido do texto. Tais conhecimentos consistem em uma espécie de dicionário enciclopédico do mundo e da cultura arquivado na memória, em estruturas de blocos denominadas de *frames*, esquemas, *scripts*, cenários. O conhecimento do mundo faz com que o usuário possa construir o mundo textual, sendo esse diferente do mundo real e formado a partir de conhecimento partilhado, ou seja, nem sempre o usuário possui previamente todo o conhecimento do assunto. Portanto, a partir de um certo grau de conhecimentos partilhados, o usuário é capaz de captar o sentido de informações novas, através do processo denominado de inferenciação.

As questões da armazenagem e da ativação dos conhecimentos na memória são importantes uma vez que a interpretação de um fato, cheiro, gosto ou texto qualquer, pode ser explicado segundo Brown e Yule, citados por Trevisan [TRE92], baseia-se, em grande parte, num simples princípio de analogia com experiências que o usuário já teve no passado. E, ainda, segundo Xavier em [XAVsd], em estudos dos mecanismos da memória, cita que: quanto mais “dicas” você tem em relação à informação original pretendida, melhor. Isso estimulará a atividade nos circuitos de informações relacionadas.

A psicologia tradicional caracteriza a memória como memória de curto termo e memória de longo termo, sendo que alguns autores [TRE92] e [KAT90] consideram um terceiro tipo, a memória operacional, denominada também de memória de médio termo.

A memória de curto termo, ou temporária, possibilita a armazenagem de seqüências de números ou de palavras, com uma capacidade limitada de até sete itens. Na produção e compreensão de textos, esse é o nível em que se integram letras e palavras.

A memória de médio termo, ou operacional, consiste no “lugar” onde o conteúdo proposicional é armazenado, integrando significados provenientes de elementos do texto. Não possui limitação quantitativa e pode operar com significados e não com formas superficiais, como palavras.

A memória de longo termo, ou permanente, é o “lugar” em que a informação mais permanente é armazenada, onde são estocados tanto os conhecimentos originados das experiências particulares quanto os conhecimentos convencionalizados a respeito do mundo.

Esse é o nível onde ocorre a integração entre o significado do texto e as informações oriundas do conhecimento prévio do usuário.

Na memória de longo termo, distinguem-se a memória semântica e a memória episódica. A memória semântica armazena estruturas cognitivas com o caráter de generalidade e contém o conhecimento sistematizado. Essa memória reflete os padrões inerentes à organização do conhecimento, como as estruturas de eventos e situações, de tal forma que o conhecimento semântico apresenta características comuns à maioria dos indivíduos. Por outro lado, a memória episódica armazena eventos lembrados como fatos particulares, fazendo, assim, com que o conhecimento episódico manifeste traços acidentais. Na verdade, segundo Trevisan [TRE92], é como se as pessoas registrassem ao lado da regra geral alguns exemplos concretos. De acordo com Xavier [XAVsd], a informação trafega pelo sistema nervoso elétrica e quimicamente, como mostrado no Capítulo 1, e quando uma nova informação chega, ela segue certos circuitos da rede nervosa, tornando determinados caminhos “facilitados”, e desta forma, pode-se memorizar e usar posteriormente um conhecimento recebido previamente.

3.4. Rememoração e Reconhecimento

Neste ponto, deve-se refletir sobre os processos de rememoração e reconhecimento para, futuramente, auxiliar na discussão sobre o reconhecimento de palavras manuscritas e, ainda, sobre o desenvolvimento de sistemas que se aproximem o máximo possível da capacidade do ser humano de ler uma palavra qualquer.

O entendimento dos aspectos referentes aos processos de rememoração e reconhecimento faz-se importante ao presente estudo diante da afirmação de Aumont em [AUM93] de que o primeiro tem função simbólica, e o segundo, tem função representativa. Ou seja, o primeiro processo, leva para a apreensão do visível, para as funções mais diretamente sensoriais. O segundo, para a memória, logo para o intelecto e para as funções do raciocínio.

A rememoração está ligada a um valor simbólico da imagem, ou seja, representa coisas abstratas em um mundo real. Esta representação conta com a aceitação social do símbolo representado [AUM93]. Assim, o processo de rememoração pela imagem tem por base um esquema. Entende-se por esquema, tal qual Aumont [AUM93], uma estrutura relativamente simples, memorizável como tal além de suas diversas atualizações. O esquema é “econômico”, ou seja, deve ser mais simples e mais legível do que aquilo que representa ou

também imbecil, tolo; rendeu ao autor da caricatura uma pesada multa. A Figura 3.14 mostra a caricatura do rei, a qual a despeito da alteração de cada traço individual, sente-se que o conjunto permanece notavelmente parecido. Assim, aceita-se a caricatura como uma alternativa possível de ver o rosto do rei [GOM95].

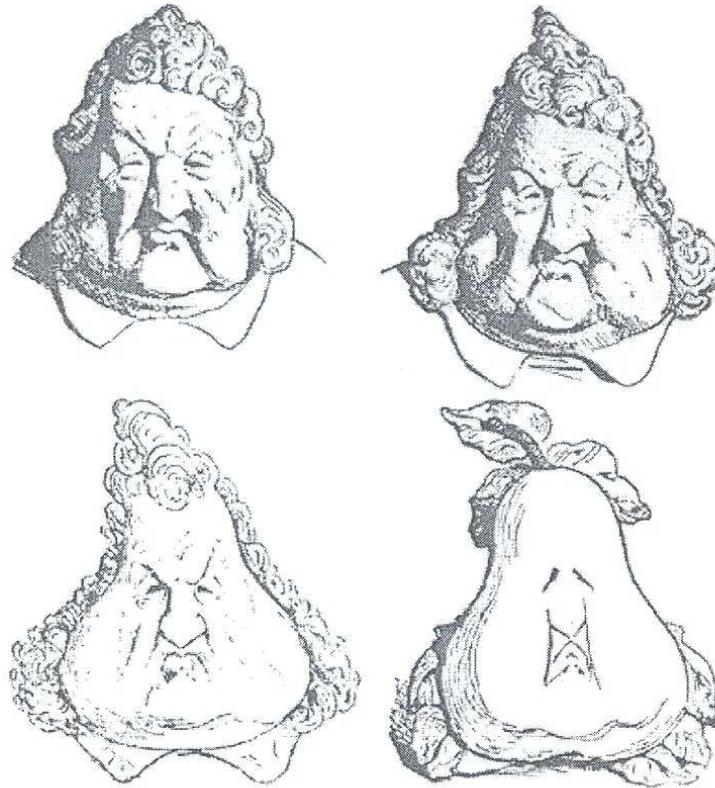


Figura 3.14: Caricatura e invariantes da visão. (Fonte: [GOM95])

Deste modo, ver e reconhecer objetos envolve muitas fontes de informação. Pois, o que o caricaturista tem a fazer é fornecer ao observador um punhado de linhas, bem escolhidas, e o cérebro do observador fará o resto: buscar objetos e encontrá-los, sempre que possível [GRE79]. Cabe, então, apresentar o exemplo da Figura 3.15, um conjunto de linhas sem sentido algum? Como mostrado em [GRE79], nada disso: é uma faxineira lavando o chão com o seu balde. Olhe novamente a figura, as linhas parecerão levemente diferentes, quase sólidas, representando objetos. A partir deste momento, você sempre reconhecerá nesta figura a faxineira e o seu balde, pois os mesmos já fazem parte do seu conhecimento prévio.

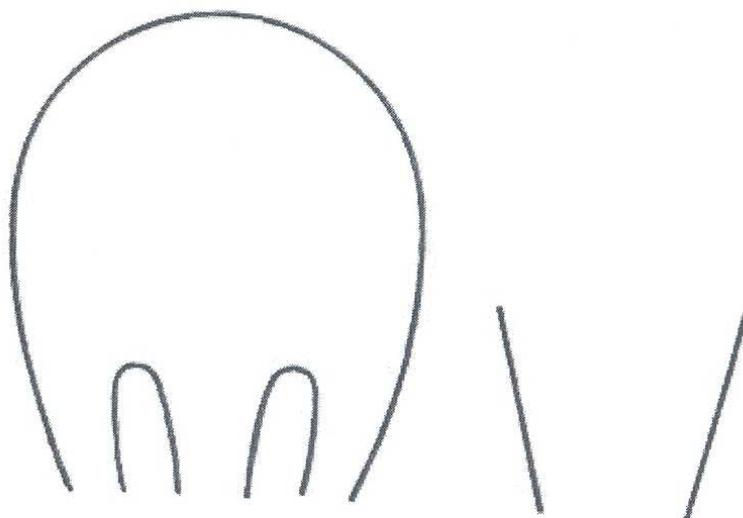


Figura 3.15: A faxineira e o seu balde. (Fonte: [GRE79])

3.5. Leitura e Reconhecimento

A revisão realizada sobre as atividades de escrita e leitura, bem como a associação destes processos com o conhecimento prévio do indivíduo em questão, e ainda, considerando-se os aspectos relevantes dos processos de rememoração e reconhecimento apresentados, permitem que o presente trabalho passe a considerar o indivíduo em questão como sendo o computador. Desta forma, fazendo um paralelo com uma criança, o desenvolvimento de um sistema de reconhecimento de palavras manuscritas assemelha-se ao desejo de iniciar o processo de leitura. Nesta seção, encontram-se apresentados os métodos utilizados para ensino da leitura de palavras manuscritas, sendo que a estes métodos estão associadas também as abordagens provenientes da percepção visual, da alfabetização e dos sistemas automáticos de reconhecimento de palavras. Estes métodos estão ligados as formas de segmentação do texto, a saber: da seqüência em palavras, das palavras em sílabas, e ainda, das palavras em letras.

Com base em todos estes métodos, técnicas e conhecimentos apresentados, pergunta-se tal qual Kato em [KAT90]: em que medida, porém, a percepção visual precisa e um processamento analítico ao nível de unidades menores que a palavra são necessários para a leitura? Segundo o autor, na medida que o aprendiz desenvolve sua capacidade de se apoiar em estruturas cada vez maiores (em seu conhecimento do mundo), esse tipo de operação será cada vez menos em nível de unidade gráficas e silábicas e cada vez mais em nível de fatias informacionais significativas.

O reconhecimento das formas será em grande parte determinado por processo inferenciais e de predições ditadas pelo conhecimento lingüístico e extralingüístico do leitor.

Um leitor proficiente é capaz de compreender um texto escrito em letra cursiva bem pouco legível, exatamente porque ele faz uso dessa habilidade para ler sem decodificar cada grafema [KAT90]. Afinal, sabe-se com base na experiência e nos autores [CAG91] e [PAR91], que aquele que escreve fluentemente, em letra cursiva, também não tem o cuidado de desenhar todos os traços distintivos de cada letra, contando com a capacidade do leitor de ler pelo significado.

Alguns resultados obtidos por Ferreiro e Teberosky (1979) e Lavine (1972), e ainda, confirmados por Rego em (1982) no contexto brasileiro e apresentados por Kato em [KAT90], mostram que as diferentes fases da leitura, em um processo de alfabetização, são importantes para o entendimento das variáveis inerentes ao problema. Estes resultados estão apresentados de maneira resumida fazendo-se uma comparação com os sistemas de computação desenvolvidos e em desenvolvimento para o reconhecimento de palavras manuscritas.

Inicialmente, segundo Ferreiro e Teberosky, a criança não diferencia a função do texto e da figura, achando que esta última também é lida. A figura é vista com função pictográfica pela criança. Lavine mostra que aos três anos de idade, as crianças rejeitam figuras e desenhos como escrita, porém segundo Ferreiro e Teberosky, ainda associam um valor icônico à escrita. É interessante ressaltar, que os sujeitos desses estudos acreditam que objetos grandes são representados por palavras mais longas. Assim, ao escrever-se *casa* e *casinha*, o entendimento ocorrerá achando que é a primeira que representa o conceito da segunda. Nesta fase, o sujeito ainda não atingiu a fase fonográfica, pois ele não tem consciência de que *casinha* é mais longa porque a palavra tem mais sons do que *casa*. Esta fase é chamada de pré-silábica, segundo Ferreiro e Teberosky. Alguns sistemas reconhecedores utilizam como informação o tamanho da palavra, porém este tipo de informação não é muito recomendado pelos pesquisadores, uma vez que recomenda-se o uso de características invariantes, fato este que vem ao encontro das propriedades elementares da percepção visual. Isto devido ao fato de que, ao comparar-se diferentes palavras de mesmo tamanho pode-se cometer erros no processo de avaliação [TRI96].

Após a fase pré-silábica, Ferreiro e Teberosky distinguem três outras fases: a silábica, a silábico-alfabética e a alfabética, a medida que a criança atribui a cada símbolo o valor de uma sílaba, de uma sílaba ou de um segmento fonético e, finalmente, apenas de um segmento fonético. Neste ponto a comparação com os sistemas reconhecedores cabe aos que se baseiam na segmentação das palavras em sílabas ou letras, tais como: [CON96] [KNE96] [GIL94]

[YAC96] [CHE95] [CHO95] [MOH95] [HEU94] [KIM93] [FAV92] [LER91] [LEC90] [BOZ89].

Para Fromkin e Rodman (1974) em [KAT90], a criança só poderá aprender o sistema alfabético quando cada segmento sonoro de sua língua tiver para ela uma realidade psicológica. Em outras palavras, a alfabetização se tornará possível quando a criança tiver consciência da relação símbolo gráfico e som oral de sua língua. Novamente, fazendo um paralelo com os trabalhos de reconhecimento de manuscritos os autores [CÔT98][KIM96][GUI95] procuram inserir, no desenvolvimento de seus sistemas, aspectos baseados em modelos psicológicos.

Kato em [KAT90] apresenta as diferentes hipóteses de aprendizagem, levando-se em conta as características do sistema ortográfico da língua portuguesa, a saber:

- hipótese fonética,
- hipótese de que há arbitrariedades no sistema ortográfico,
- hipótese de que há regularidades contextuais entre fala e escrita e
- hipótese de que há regularidades oriundas da morfologia.

A cada hipótese de aprendizagem, temos os métodos correspondentes para a efetivação do processo de aprendizagem:

- **Método Global Puro:** supõe que a apreensão do estímulo visual se dá de forma ideográfica, sem análise das partes que o compõem,
- **Método Global Analítico-Silábico:** supõe que a concepção da criança sobre a palavra escrita é que os estímulos são decomponíveis em unidades menores, silábicas, ou que é possível introduzir a ela esta noção,
- **Método Global Analítico-Fonêmico:** pressupõe que essa decomposição é possível de ser feita a nível fonêmico-grafêmico,
- **Método Silábico-Sintético:** supõe que a criança seja capaz de perceber uma entidade mais abstrata que a palavra, a sílaba, e a partir de sua representação grafêmica chegar em unidades significativas como a palavra e a frase,
- **Método Fônico-Sintético:** supõe que a criança seja capaz de captar unidades sonoras físicas, menores que a sílaba – o fone – para, a partir de sua representação grafêmica, chegar às unidades significativas.

Assim, o método global analítico, simula melhor o que ocorre naturalmente quando uma criança ou um adulto se depara com o universo visual que o cerca. Portanto, Kato em [KAT90], a percepção não parte de segmentos isolados, mas sim da discriminação dos traços

distintivos dos estímulos que estão sendo observados. Por exemplo, ao observar-se o anúncio da *Coca-Cola*, uma pessoa poderá vir a segmentar a sílaba *co*, dada a sua dupla ocorrência, e na comparação do *co* com o *ca*, poderá vir a diferenciar o *a* do *o* e isolar o *c*, e assim por diante. Uma vez percebidas essas unidades, as quais em si mesmas não tem significado, as pessoas podem combinar as unidades e chegar a unidades significativas. Considerando-se este mesmo exemplo do anúncio da *Coca-Cola*, pode-se dizer também que o observador fará um processo de rememoração do esquema que lembra o produto. Este fato pode ser confirmado, ao se perceber que mesmo uma criança que não sabe ler associa o esquema ao produto veiculado. Portanto, um observador não decompõe a palavra em unidades menores, e sim, realiza um processo global combinando conhecimento prévio, rememoração de esquemas e interpretação de mensagens.

Assim, para Kleiman [KLE00] os principais mecanismos engajados no processamento de um texto pelo ser humano fazem uso das memórias e de capacidades que variam de leitor para leitor.

Os mecanismos iniciais da leitura partem do Kleiman [KLE00] afirma: “durante a leitura os olhos vão para frente, num movimento progressivo, mas também retrocedem, num movimento regressivo”. E ainda, que se realiza muito mais movimentos progressivos quando o material é mais difícil.

Entende-se por mais difícil, não somente a pregnância da forma, mas também, a compreensão do texto. Deste modo, os olhos se fixam num lugar do texto (*fixação*) para depois pular um trecho (*a sacada*), e fixar-se num outro ponto mais adiante.

Durante a fixação, o olho percebe claramente o material focalizado. Durante a sacada, entretanto, a visão fica muito reduzida, ou seja, o leitor não percebe claramente as unidades do texto, mas mantém uma visão periférica. Estes mecanismos (*fixação* e *sacada*) apontam para o fato de que grande parte do que se lê é adivinhado ou inferido, não sendo diretamente percebido. Daí a leitura ser considerada, do ponto de vista cognitivo, um “jogo de adivinhação” [KLE00].

3.6. Discussão

As pessoas lêem como um todo analisado.

Com base nos tópicos apresentados neste capítulo pode-se citar que Kato em [KAT90] e Kleiman [KLE00] apresentam que as pesquisas em leitura, principalmente na área de psicologia e psicolinguística, são unânimes em afirmar que, na leitura proficiente, as palavras

são lidas não letra por letra ou sílaba por sílaba, mas sim como um todo analisado, isto é, por reconhecimento instantâneo e não por reconhecimento analítico-sintético. Citando, Smith (1978) em [KAT90], o reconhecimento de palavras se dá como o reconhecimento de outro objeto qualquer, como por exemplo: árvore, carro, casa e, da mesma forma que se identifica um objeto através de sua configuração geral, pode-se reconhecer uma palavra através do todo (seu contorno e extensão, entre outros) sem análise de suas partes. Conforme já mencionado por [MAC96] e apresentado anteriormente, Kato afirma que a identificação de uma árvore pode ser feita enxergando-se apenas uma parte de sua copa, a palavra pode ser reconhecida ou adivinhada sem enxergarmos a sua totalidade, ver Figura 2.10 (princípio do fechamento). A leitura de uma palavra por um leitor competente é feita, pois, de maneira ideográfica.

A percepção visual pode auxiliar os sistemas automáticos de reconhecimento.

Neste ponto, questiona-se: Como um sistema de reconhecimento automático de palavras manuscritas pode ser construído? Que abordagem deve ser levada em consideração? Como a percepção visual pode contribuir para o desenvolvimento destes sistemas? Para Yacoubi em [YAC96] os sistemas reconhecedores devem ser o mais próximo possível da habilidade humana de ler qualquer palavra escrita. Assim, pode-se entender que os sistemas deveriam considerar uma abordagem global, ou seja, por objetos, no caso por palavras. Mas, então surge a seguinte pergunta: Como desenvolver sistemas globais que possam ler qualquer palavra? Neste caso, considerando-se léxicos de grande dimensão, pois a maioria dos sistemas tenta atingir este objetivo trabalhando com as palavras de modo analítico-silábico, analítico-grafêmico, silábico-sintético e fônico-sintético, conforme sistemas e modelos desenvolvidos em [CON96] [KNE96] [GIL94] [YAC96] [CHO95] [MOH95] [HEU94] [KIM93] [FAV92] [LEC90][BOZ89]. Estes sistemas atendem a léxicos de grande dimensão, permitindo que as dificuldades de implementação de métodos de reconhecimento sejam superadas.

Outros autores, baseiam-se no método global, ou na abordagem holística, como nos trabalhos de [CÔT98][KIM96][GUI95][SEN98][MAD92]. Estes trabalhos estão voltados para léxicos de pequena dimensão, como no conjunto de palavras utilizados em cheques bancários ou meses do ano. Portanto, pergunta-se ainda: Será somente uma limitação de implementação dos sistemas ditos globais? Ou não se sabe suficientemente sobre a percepção visual e como aplicá-la neste problema? E ainda, porque os sistemas falham? Muitas vezes consegue-se ler o que está escrito e o sistema implementado não o faz (análise visual de bases de dados).

O leitor iniciante e o leitor eficiente.

Segundo Smith (1969 e 1978), em [KAT90], é reconhecido na literatura que, quanto mais eficiente o leitor, maior o seu vocabulário visual, cerca de 50.000 itens, e que o processo de análise e síntese da palavra em unidades menores só será usado para itens estranhos a esse universo. Smith acredita ainda que raramente o leitor competente chega a analisar tais palavras, pois o contexto normalmente fornece “pistas” suficientes para seu entendimento.

Para um leitor iniciante, porém, cujo vocabulário visual ainda é muito limitado, mesmo para aqueles alfabetizados pelo método global, o processo de leitura envolve muito pouco reconhecimento visual instantâneo, consistindo a leitura, mais freqüentemente, em operações de análise e síntese, sendo apreensão do significado mediada quase sempre pela decodificação em palavras auditivamente familiares [KAT90].

As pessoas analisam e sintetizam palavras.

As operações de análise e síntese, depois de certo tempo, podem basear-se em segmentos maiores que a letra: sílabas, morfemas e, acredita Kato em [KAT90], também em partes maiores da palavra que não constituem unidades lingüísticas, mas que tem uma incidência muito grande na língua, como é o caso do segmento -ola que aparece em palavras como bola, mola, cebola e muitas outras. Trata-se, pois de uma operação de decomposição e composição, que envolve também o reconhecimento instantâneo, porém parcial, da palavra. Esta característica está presente em outros vocabulários, como por exemplo, em cheques bancários (quar-enta e cinqu-enta, trez-entos e quatroc-entos) e meses do ano (jan-eiro e fever-eiro).

Para Kato em [KAT90], se uma palavra for apenas reconhecida parcialmente, a leitura da parte estranha exigirá decomposições sucessivas até a identificação formal total da palavra. O conhecimento prévio do leitor quanto às restrições fonotático/ortográficas e seu vocabulário auditivo podem atuar para minimizar esse trabalho de análise e síntese. Resumidamente, é como se cada pessoa trabalhasse ora como um interpretador, ora como um compilador.

As pessoas lêem rapidamente graças aos mecanismos e habilidades do cérebro.

Considerando-se todos estes tópicos, e ainda, se ler é um processo tão complexo (pois exige do leitor mecanismos de *fixação*, *sacadas*, *inferências*, entre outros) pergunta-se: como se lê tão rapidamente?

A resposta está em Kleiman [KLE00] que afirma que a leitura é realizada rapidamente para que se possa organizar os traços no papel em material significativo. Após a apreensão do material visual inicia-se a interpretação. A autora propõe inicialmente que o material será interpretado na seguinte seqüência: letras em sílabas e palavras, destas em frases, destas em proposições com significado. Afinal, é assim que aprendemos a ler. Por outro lado, Kleiman [KLE00] demonstra que na verdade esta não é a seqüência de processamento. A seqüência de processamento de um texto utiliza os mecanismos expostos na Figura 3.16.

Assim, o material visual apreendido vai sendo estocado na *memória de trabalho* que permite a organização em unidades sintáticas, segundo regras e princípios da gramática implícita (gramática do idioma do texto que é conhecida pelo leitor). O processo de agrupamento e análise das unidades sintáticas é conhecido como *fatiamento*. A *memória de trabalho* é auxiliada pela *memória intermediária* e pela *memória semântica*, que possibilitam ativar o repositório de conhecimento e, ainda, usar e organizar através de regras o conhecimento a ser aplicado, respectivamente.

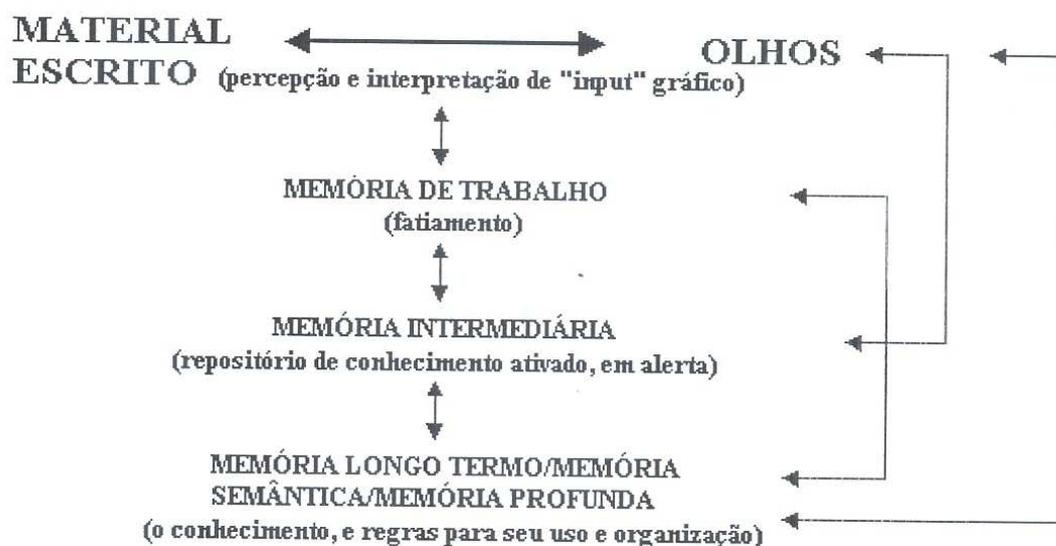


Figura 3.16: Mecanismos e capacidades envolvidos no processamento de um texto.
(Fonte:[KLE00])

Kleiman [KLE00] explica que a memória de trabalho é uma capacidade finita e limitada, uma vez que não pode trabalhar com mais de 7 unidades ao mesmo tempo, aproximadamente (podendo trabalhar com mais ou menos duas unidades, isto é, entre 5 e 9 unidades). O importante dessa capacidade é que não faz diferença, para o seu funcionamento, o tipo de unidade usada para o fatiamento, precisando apenas ser uma unidade significativa. Entende-se como uma unidade significativa àquela que pode ser reconhecida como unidade, seja esta uma letra, sílaba ou palavra.

Para exemplificar estas habilidades, Kleiman [KLE00] mostra um exemplo simples: suponha que a maior unidade significativa para o leitor é a letra, numa seqüência como a seguinte:

a—s—e—s—t—r—u—t—u—r—a—s

Se o leitor estiver lendo letra por letra, não conseguirá manter na memória de trabalho todas as unidades, pois precisará esvaziar as primeiras letras à medida que outras letras vão sendo lidas. Em relação à compreensão, o leitor não será capaz de apreender a seqüência apresentada, muito menos terá lido as duas palavras que se encontram nela.

O mesmo problema acontecerá se o leitor estiver lendo por sílabas. Pois, quando estiver na Quinta sílaba a memória de trabalho deverá ser esvaziada para poder receber novas unidades.

as—es—tru—tu—ras

Suponha, então, que a unidade mínima significativa é a palavra. Neste caso, o leitor poderá ler mais palavras, pois a memória de trabalho não estará sobrecarregada. Assim, sabe-se através de Kleiman [KLE00] e por experiência própria (através da fase de alfabetização de meus filhos) que o leitor iniciante lê vagarosamente, sílaba por sílaba, tendo dificuldade de lembrar o que estava no início da linha quando chega ao final desta. É comum, a criança “silabar” as palavras e depois ler de uma só vez a palavra para, então, compreender o seu significado.

as—estruturas

Deste modo, o leitor rápido é capaz de reconhecer instantaneamente as palavras, pois se a palavra for a unidade reconhecida, ele poderá ler mais rapidamente, conseguindo assim lembrar unidades possíveis de interpretação semântica (ou sejam que possuem um significado, considerando-se o idioma e o conhecimento prévio da gramática).

Ler não é o inverso de escrever.

Em Nemirovsky [TEB96] encontra-se que ler e escrever não são atividades inversas, mas sim diferentes, que exigem das pessoas conhecimento e recursos distintos. Deste modo, acredita-se que o desenvolvimento de sistemas automáticos de reconhecimento de palavras manuscritas deve ter por base o estudo dos conhecimentos que permitem ao ser humano ler

uma palavra qualquer, de forma a implementar métodos que possam representar os distintos recursos utilizados, bem como, a cooperação da informação processada pelos diferentes métodos.

Deste modo, no Capítulo 4, estabelece-se uma relação entre os capítulos anteriores, analisando trabalhos de autores que compartilham da mesma linha de pensamento e, portanto, investigam aspectos da percepção humana nos processo de escrita e leitura, de modo a transpô-los aos sistemas de reconhecimento.

Capítulo 4

Percepção e Reconhecimento de Palavras

Manuscritas

Este Capítulo estabelece a relação entre os capítulos anteriores, analisando trabalhos de autores que compartilham da mesma linha de pensamento e, portanto, investigam aspectos da percepção humana nos processo de escrita e leitura, de modo a transpô-los aos sistemas de reconhecimento. Este capítulo analisa bases de dados de palavras manuscritas de forma a investigar e validar aspectos da percepção humana no processo de leitura através de um sistema de reconhecimento automático de palavras manuscritas implementado durante o desenvolvimento da Tese de Doutorado desenvolvida por Freitas [FRE01]: “Uso de Modelos Escondidos de Markov para Reconhecimento de Palavras Manuscritas”.

4.1. Níveis de Abstração

Uma imagem pode conter vários objetos. Assim, a área de RPM - Reconhecimento de Palavras Manuscritas (HWR – *Handwritten Word Recognition*) dedica-se ao do objeto “paalavra” envolvendo a determinação de parâmetros e características deste objeto, e não necessariamente de toda a imagem.

Para determinar parâmetros e características das palavras precisa-se inicialmente identificar as palavras no documento ou contexto onde estas se encontram, efetuar a segmentação das mesmas (em relação ao fundo, em relação a outros objetos manuscritos ou pré-impresos, em relação a objetos decorativos ou objetos característicos do documento – selo e carimbos em envelopes postais, entre outros) e, somente, então efetuar os cálculos.

A seguir, apresentam-se os níveis de abstração aplicáveis às imagens das palavras manuscritas, das quais se pretende obter características para serem utilizadas como *input* dos

métodos de reconhecimento. Os níveis de abstração auxiliam na compreensão da complexidade do processamento de imagens de palavras manuscritas, não somente do ponto computacional, mas também das limitações inseridas nos sistemas automáticos quando da passagem do processo humano de leitura para o processo computacional de reconhecimento de palavras. Deste modo, as imagens serão analisadas por 4 níveis de abstração, a saber:

- Nível objeto: a palavra e seu entorno,
- Nível global: a palavra,
- Nível local: as letras que compõem uma palavra,
- Nível pixel: a palavra em meio digital.

4.1.1. Objeto: A Palavra e seu Entorno

A palavra manuscrita vem acompanhada do seu entorno, ou seja, do contexto onde ela se encontra. Muitos estudos na área de RPM referem-se aos mais variados contextos, tais como: cheques bancários: [HEU94] [GUI95] [AVI96] [KNE96] [CÔT97] [OLL99] [FRE02] [MOR02]; envelopes postais: [YAC99]; formulários diversos de cartão de crédito e fichas de cadastramento, entre outros. A Figura 4.1 exemplifica imagens de palavras manuscritas nos contextos de cheques bancários e envelopes postais.



Figura 4.1: Exemplos de palavras manuscritas e o contexto:
a) cheque bancário e b) envelope postal.

O contexto influencia as palavras manuscritas, uma vez que pode existir um vocabulário finito de palavras a ser aplicado, como em extensos ou datas de cheques bancários. Para envelopes postais, quando se trata dos nomes de ruas e de cidades, o léxico é de grande dimensão, pois envolve todos os nomes de ruas de uma cidade ou, ainda, de várias cidades.

O contexto também influencia a extração da área que contém as palavras manuscritas. Na Figura 4.1-b pode-se observar que o selo e carimbo sobrepõem-se as palavras manuscritas,

e na Figura 4.1-a pode-se ressaltar a influência do padrão de fundo sobre as informações manuscritas no cheque bancário. Estes fatos fazem com que etapas do processamento dos objetos em questão sejam prejudicadas ou dificultadas.

Assim, muitos esforços computacionais podem ser realizados com o intuito de resolver o princípio da Gestalt conhecido como Figura-Fundo (ver Seção 2.3.1). Muitos algoritmos podem ser estudados e aplicados, tais como o proposto por Yonekura em [YON02] onde uma metodologia de segmentação automática de envelopes postais com fundo complexo busca separar o conteúdo dos envelopes em 3 classes: “bloco-endereço manuscrito e carimbo”, “fundo” e “selo”. O autor utiliza pouco conhecimento à priori dos envelopes, e a metodologia faz uma mescla da abordagem Watershed morfológico com o conceito de matriz de co-ocorrência objetivando resgatar as muitas acumulações geradas em matriz de co-ocorrência geralmente desperdiçadas em metodologias clássicas como, por exemplo, segmentação por limiarização. A Figura 4.2 exemplifica os resultados obtidos com a metodologia proposta.

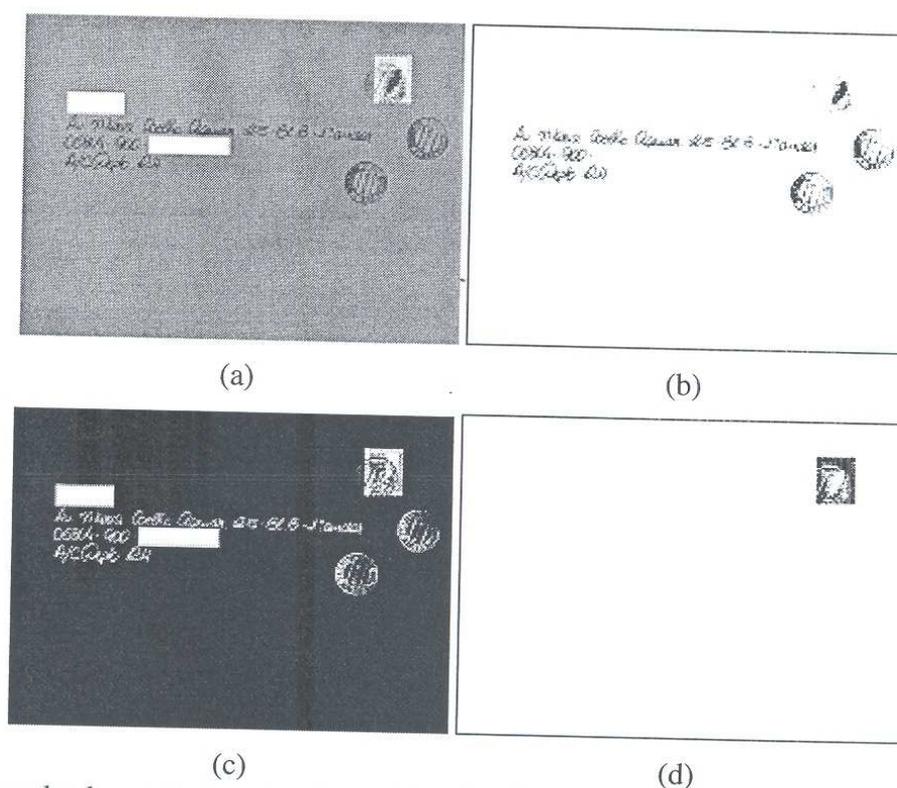


Figura 4.2: Exemplo de segmentação Figura-Fundo: a) imagem original, b) “bloco-endereço manuscrito e carimbo”, c) “fundo” d) “selo”.

No que se refere, a segmentação das palavras manuscritas em detrimento das demais informações muitos estudos já foram realizados aplicando-se diferentes técnicas, tais como:

- segmentação por análise dos componentes conectados da imagem do documento que permite associar aos elementos um valor de confiança [YU 97],
- segmentação por contexto que permite buscar regiões onde a informação manuscrita é esperada *a priori* [LII93],
- segmentação por multi-binarização que permite avaliar melhor as faixas de níveis de cinza presentes na imagem do objeto [TSA85] [YAN96],
- segmentação por textura que permite livrar-se do fundo complexo de certos documentos [RUZ97],
- segmentação morfológica Watershed que permite extrair informações em imagens complexas [BEU92] [FAC96].



Figura 4.3: Exemplo de segmentação Figura-Fundo: a) imagem original, b) “bloco-endereço manuscrito” com o campo *endereço=nome de rua+ bairro* segmentado.

4.1.2. Global: Palavra

Geralmente o reconhecimento de palavras manuscritas baseia-se em duas abordagens do problema: Global (*global approach*) realizado em nível das palavras [GUI95] e [CÔT97] ou Local (*analytical approach*) realizada em nível dos caracteres [LEC90], [FAV92], [GIL95], [KIM96].

A abordagem global permite evitar a etapa de segmentação das palavras, extraíndo-se primitivas globais das palavras sem necessidade de segmentação explícita das mesmas. Esta abordagem procura explorar as informações do contexto das palavras, sendo que aspectos baseados em modelos psicológicos podem ser inseridos [GUI95] e [CÔT97]. Porém, é uma abordagem restrita às aplicações de léxico com pequenas dimensões. Isto devido ao fato da

dificuldade de se obter modelos de reconhecimento individuais (para cada uma das palavras do léxico estudado) e, ainda, pela dificuldade de obtenção de uma base de dados de dimensão satisfatória ao treinamento destes modelos.

O interesse em trabalhar com as palavras globalmente está em não depender do sucesso do método de segmentação das palavras em letras ou pseudo-letras mas, de maneira mais desafiadora, entender como o processo de leitura de palavras realizado pelo ser humano pode ser, ou se tentar ser, transportado para o computador. A Figura 4.4 exemplifica imagens de palavras da base de dados de cheques bancários da PUCPR/PPGIA/LUCI, demonstrando a complexidade do problema. Nesta figura, tal qual Madhvanath et al. em [MAD01] apresenta-se a forma da palavra como informação global suficiente ao reconhecimento para o estilo cursivo [TAP90]. Fato este, que não ocorre quando a palavra está no estilo letra de forma ou caixa-alta, dificultando, também a análise global de palavras somente através da forma do traçado como um todo.

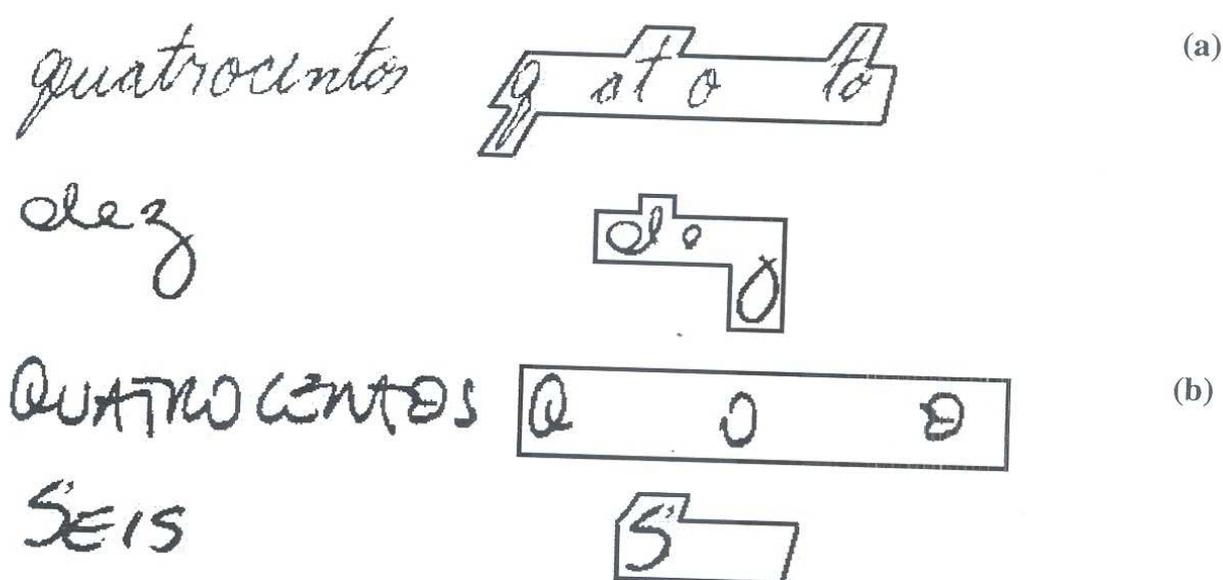


Figura 4.4: Exemplos de análise global de palavras manuscritas de cheques bancários brasileiros: a) estilo cursivo e b) estilo caixa-alta.

4.1.3. Local: Letra

Na abordagem local o método se depara com a necessidade de segmentação das palavras em caracteres/letras ou pseudo-letras. Esta abordagem caracteriza-se pela dificuldade em se definir a fronteira entre os caracteres. E, portanto, o método de reconhecimento dependerá também do sucesso do processo de segmentação. Assim, estes métodos são constituídos por etapas distintas para segmentação e reconhecimento, ou associando segmentação e reconhecimento em uma mesma etapa.

Morita et al. em [MOR01] descreve um método de segmentação de palavras manuscritas e exemplifica a dificuldade da definição da fronteira entre os caracteres de uma mesma palavra. No método proposto busca-se a minimização da largura vertical no provável ponto de segmentação, como mostrado na Figura 4.5.

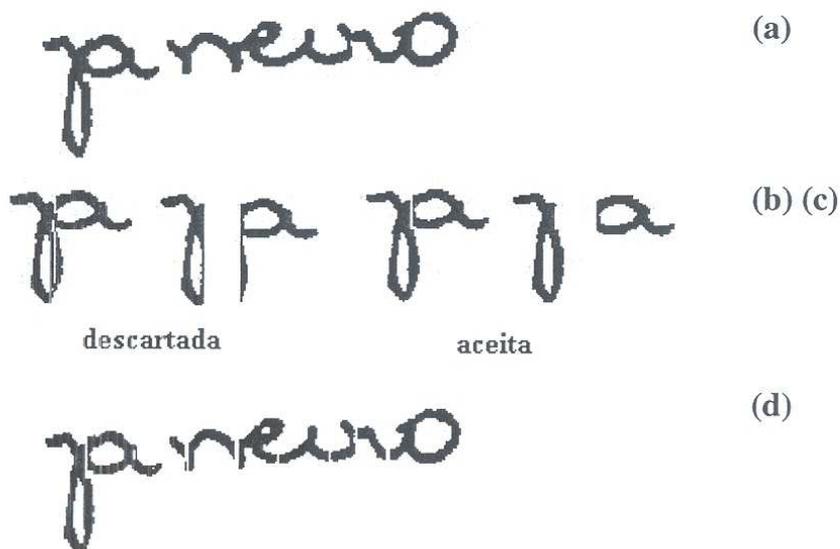


Figura 4.5: Exemplos de segmentação de palavras manuscritas em letras ou pseudoletas: a) imagem original, b) hipótese de segmentação descartada, c) hipótese de segmentação aceita, d) palavra segmentada.

4.1.4. Pixel: Palavra Digital

O presente trabalho considera o pixel o nível mais baixo de abstração de uma imagem e, portanto, é neste nível que muitos dos métodos e técnicas de processamento de imagens são aplicados. Ressalta-se que neste nível de abstração, todos os demais níveis se tornam desconhecidos, ou seja, quanto mais em baixo nível se tratam as imagens menos informação sobre o todo se tem. Neste nível, para cada pixel X na imagem obtém-se no máximo a informação sobre o conjunto dos 8 vizinhos mais próximos $\{v_1, v_2, \dots, v_8\}$, como mostrado na Figura 4.6.

v1	v2	v3
v4	X	v5
v6	v7	v8

Figura 4.6: Nível de abstração: pixel.

Um método clássico de processamento de imagem ao nível pixel são os algoritmos de limiarização ou *thresholding*. Através destes algoritmos procura-se obter uma nova imagem que contenha as informações da original, porém em uma outra composição de níveis de cinza

ou, até mesmo, uma imagem binária (somente 2 níveis, branco de preto). A Figura 4.7 exemplifica diferentes algoritmos de limiarização para a imagem de uma palavra manuscrita originalmente em níveis de cinza. Todos os métodos exemplificados constam do programa PDIImagem, desenvolvido por Alceu de Souza Britto Jr e disponível em [FAC02].

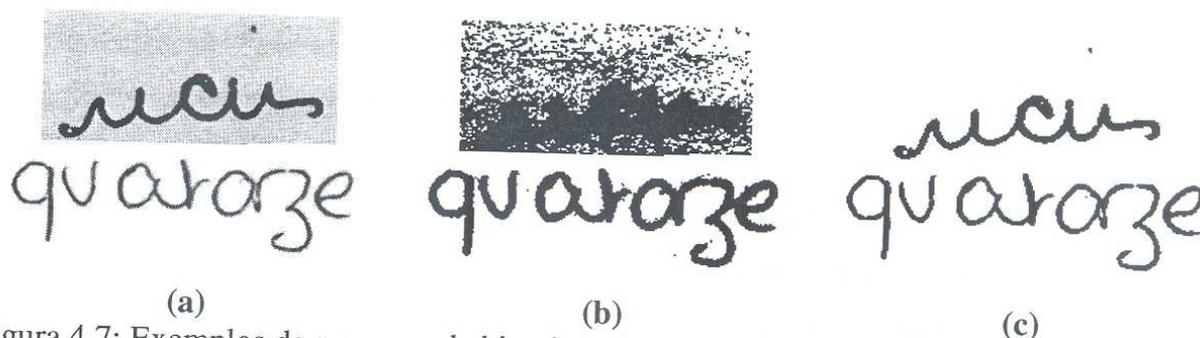


Figura 4.7: Exemplos de processo de binarização: a) imagens originais em níveis de cinza, b) $threshold = 217$ e 255 (método anisotropia) e c) $threshold = 110$ e 179 (método OTSU).

A decomposição do problema de reconhecimento de palavras manuscritas em 4 níveis de abstração permite que se observe, de maneira hierárquica, a complexidade inerente ao problema, bem como, as dificuldades de processamento de imagens em contraposição ao objetivo do reconhecimento. Isto, devido ao fato da imagem estar em formato digital e, portanto, na representação ao nível dos pixels e a visão do observador ao nível dos objetos. Constata-se aqui a distância existente entre o que os olhos vêem e o que o computador processa.

4.2. Percepção e Reconhecimento

Todos os aspectos apresentados até então ressaltam que a capacidade humana de ler, praticamente qualquer tipo de letra, desafia os pesquisadores a entender cada vez mais o ser humano para, somente assim, entender porque seus sistemas computacionais falham. Busca-se, portanto, a resposta para o seguinte questionamento: Como relacionar a percepção visual de palavras manuscritas aos sistemas de reconhecimento?

A resposta está em entender quais as características da percepção visual que contribuem no desenvolvimento de sistemas de reconhecimento. Parece fácil! Mas, não é.

Muitos trabalhos têm sido realizados com a intenção de buscar mais informações sobre a percepção visual e transpô-las aos sistemas de reconhecimento [LEC90] [GUI95] [AVI96] [CÔT97] [DZU98] [MAD98] [SCH98] [SEN98]. Alguns destes trabalhos já foram citados e analisados por Freitas em [FRE01], porém outros ainda devem ser destacados [SUE83] [MAA83] [THO83] [WIN83], pois contribuem na busca de soluções para o

problema do reconhecimento de palavras manuscritas. A seguir apresenta-se um resumo das conclusões destes trabalhos, sempre relacionando a percepção com o objetivo de reconhecer.

4.2.1. Vogais e Consoantes, Palavras Curtas e Longas, Ascendentes e Descendentes

O trabalho de Schomaker & Segers [SCH98] descreve um experimento realizado com o objetivo de determinar as primitivas relevantes no processo de leitura de palavras cursivas realizado pelo ser humano. Os autores propõem um método para inserir mais detalhes sobre primitivas geométricas as quais os seres humanos utilizam para ler. O método consiste em computar o número de *clicks* que uma pessoa realiza com o *mouse* sobre a imagem de uma palavra (idioma inglês) em condições de luminosidade precária, apresentada na tela do computador. O trabalho confirma o alto grau de utilização dos ascendentes, descendentes, cruzamentos e pontos de alta curvatura no reconhecimento do padrão manuscrito. As principais conclusões dos autores são as seguintes:

- em um primeiro nível as pessoas utilizam os ascendentes (d,k,l,h,t,b) e descendentes (q,y,j,g,p), sendo a letra *f* um caso especial, pois possui ambas as características,
- as consoantes possuem uma maior importância no processo de leitura do que as vogais, sendo possível ler ou reconhecer uma palavra sem a presença dessas letras (handwriting = hndwrtnng),
- as vogais (a,e,i,o) geram um mesmo comportamento do leitor, porém a letra *u* necessita de mais informações, isto porque deve ser diferenciada de uma letra *w* ou *m*,
- a primeira e a última letras são muito importante no processo de reconhecimento (estas letras fornecem “pistas” sobre as palavras e permitem uma ligação com o conhecimento prévio do leitor sobre o padrão a ser analisado e reconhecido – Seções 3.2 e 3.3),
- a última letra é muito utilizado no processo de reconhecimento,
- palavras curtas necessitam de um maior número de informações para serem reconhecidas,
- palavras longas necessitam de mais informações no final das mesmas,
- o final das palavras, a barra de corte da letra *t* e o ponto da letra *i* deterioram o processo de reconhecimento quando são mal interpretados.

O ponto forte deste trabalho está em enfatizar a habilidade humana de ler palavras em condições de luminosidade precária, de maneira a exportar características do processo humano para o processo automatizado, permitindo assim, um melhor entendimento de ambos os processos.

4.2.2. Velocidade da Escrita para Diferentes Estilos

O trabalho de Suen [SUE83] descreve diversos experimentos realizados para estudar a velocidade da escrita para 3 diferentes estilos, como apresentados na Figura 4.8, a saber:

- block print: letra de forma ou caixa-alta ou, ainda, boxed discrete characters de acordo com a classificação de [TAP90],
- manuscript: letra manuscrita com caracteres espaçados ou separados ou, letra de imprensa, ou, ainda, spaced discrete characters de acordo com a classificação de [TAP90],
- cursive writing: letra cursiva ou pure cursive script writing de acordo com a classificação de [TAP90].

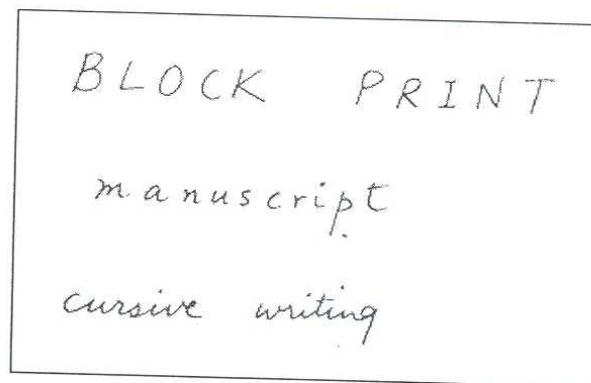


Figura 4.8: Estilos de escrita estudados por [SUE83].

Suen relata uma interessante característica no aprendizado da escrita das crianças norte-americanas: inicialmente as crianças são apresentadas ao estilo *manuscript* para depois migrarem ao estilo *cursive writing*. No Brasil, como já mencionado no Capítulo 3, as crianças iniciam pela letra de forma ou *block print*, migrando posteriormente para o estilo cursivo ou *cursive writing*.

O primeiro conjunto de experimentos realizados demonstra que escrever cursivamente com a mão direita (2,12 caracteres/seg) permite uma maior rapidez entre os estilos testados entre escritores destros ou canhotos.

Um segundo experimento demonstra que o estilo *block print* ou letra de forma permite uma maior legibilidade dos caracteres (14,86 erros/assunto) em relação aos demais estilos: *manuscript* com 30,80 erros/assunto e *cursive writing* com 61,66 erros/assunto. Estes valores são referentes aos escritores canhotos.

Este experimento também mostra que à medida que o escritor escreve diversas vezes a mesma letra, esta sofre deformações. Por exemplo, com o estilo *block print*: a letra “T” passa a ser degenerada para a letra “L”. O mesmo ocorre com as letras “U” e “V”. A letra cursiva é o estilo que mais erros produz. As confusões predominantes são entre as seguintes letras: c-e, c-i, e-i, g-q, g-y, h-b, k-h, n-w, o-a, q-g, r-i, v-u e w-u. Outra constatação é que a escrita gerada por escritores canhotos é menos legível que a gerada por escritores destros.

O terceiro experimento relata a velocidade de escrita de caracteres em letra de forma (*block print*) em formulários ou espaços formatados como na Figura 4.9. O autor observou que, em média, leva-se 2,4 segundos para escrever um caracter respeitando-se o modelo de caixas (*boxes*). Este tempo é três vezes maior do que o necessário para escrever o mesmo estilo de letra sem restrições. Este fato se explica diante da necessidade do escritor prestar mais atenção ao tamanho dos traços, sequência e formação dos caracteres.

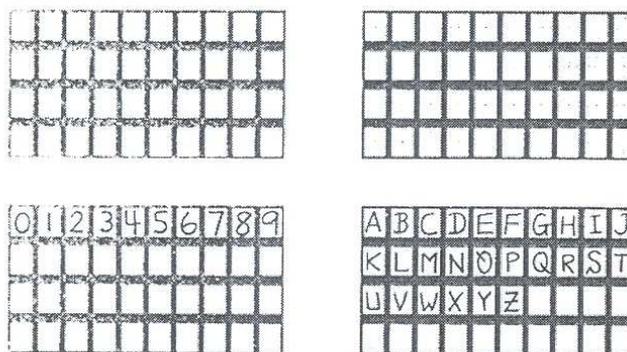


Figura 4.9: Exemplos de modelos para escrita em letra de forma. (Fonte: [SUE83])

Além disto, caracteres formados por traços simples e curvas, sem que o lápis ou caneta precise ser levantado do papel, são mais rápidos de serem escritos, como mostrado na Figura 4.10. Esta figura exemplifica as sequências de traçado dos caracteres escritas mais rapidamente por escritores destros e canhotos. Os valores numéricos representam o tempo em segundos necessário para escrever 20 exemplares.

	A	B	C	D	E	F	G	H	I	J
R	A ₁ 36.40	B ₂ 49.72	C ₁ 34.20	D ₂ 43.69	E ₁ 48.27	F ₂ 38.95	G ₁ 44.37	H ₂ 39.65	I ₁ 40.30	J ₁ 36.15
L	A ₂ 36.98	" 46.70	" 37.27	" 42.47	E ₂ 44.09	" 35.62	" 44.51	" 38.49	" 40.81	" 35.25

	K	L	M	N	O	P	Q	R	S	T
R	K ₂ 38.58	L ₁ 24.65	M ₁ 38.99	N ₁ 36.37	O ₁ 34.51	P ₂ 40.28	Q ₂ 48.14	R ₂ 47.27	S ₁ 40.90	T ₂ 28.39
L	" 38.97	" 24.96	" 41.91	N ₂ 36.13	O ₂ 28.83	" 38.62	" 44.65	R ₃ 48.61	" 39.45	" 29.53

	U	V	W	X	Y	Z
R	U ₁ 33.64	V ₁ 26.69	W ₁ 39.62	X ₂ 33.74	Y ₂ 29.87	Z ₁ 31.27
L	" 35.59	" 26.51	" 39.34	" 31.44	" 29.89	" 33.21

	1	2	3	4	5	6	7	8	9	0
R	1 ₁ 19.92	2 ₁ 37.32	3 ₁ 42.87	4 ₁ 33.40	5 ₁ 42.89	6 ₁ 36.23	7 ₁ 24.87	8 ₁ 35.89	9 ₁ 33.89	0 ₁ 34.51
L	" 16.35	" 36.67	3 ₂ 40.60	4 ₂ 30.71	" 41.33	" 36.05	" 25.01	" 37.39	" 35.67	0 ₂ 28.83

	!	/	\	-	0
R	! ₁ 19.9	/ ₁ 20.32	\ ₁ 20.43	- ₁ 16.37	0 ₁ 34.51
L	" 16.35	/ ₂ 17.04	" 17.01	- ₂ 14.03	0 ₂ 28.83

Figura 4.10: Velocidade de escrita dos modelos de seqüências para caracteres no estilo *block print*. (Fonte:[SUE83])

Este trabalho traz à discussão a diferença entre escritores destros e canhotos e o reconhecimento de suas respectivas letras. E ainda, o autor menciona que os efeitos das restrições espaciais no reconhecimento de caracteres através do computador podem ser mais bem explorados levando-se em conta a velocidade de escrita e a seqüência do traçado.

4.2.3. Movimento Horizontal da Escrita

Os estudos de palavras manuscritas assumem freqüentemente que a influência do movimento horizontal do traçado é igual sobre todas as partes das palavras. O trabalho desenvolvido por Maarse & Thomassen [MAA83] mostra que isto não é verdade e depende de onde os traços são feitos durante a escrita. Os autores mostram que a maioria dos ascendentes está entre 20° e 70° e que os descendentes se encontram entre 220° e 280°. O que confirma a não igualdade de inclinação entre traços destas regiões, como mostrado na Figura 4.11.

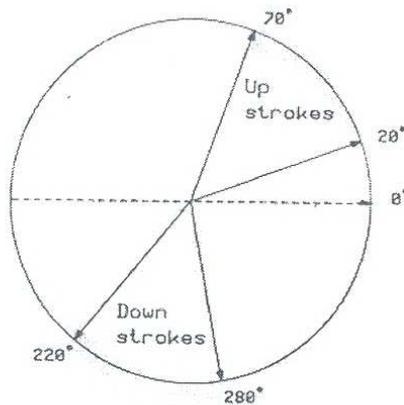


Figura 4.11: Regiões de inclinação vertical (*slant*) dos ascendentes e descendentes em palavras manuscritas. (Fonte:[MAA83])

Os autores observaram que existe uma relação entre a mudança de inclinação dos ascendentes e descendentes em função do espaçamento horizontal disponível para a escrita, como exemplificado na Figura 4.12. Foram testados três tipos de espaçamento horizontal: normal, estreito e largo, considerando-se as características do texto a ser copiado pelos escritores.

Assim, os resultados obtidos mostraram que a mudança de inclinação nos ascendentes é duas vezes maior que nos descendentes, para espaçamento do estreito para o largo. Isto significa que, quando o escritor precisa escrever em um espaço largo ele tende a escrever mais alto e em um espaço estreito ele tende a escrever menor (no que diz respeito à altura). E, ainda, que o tamanho total dos descendentes é aproximadamente de 10% a 40% menor que o tamanho dos ascendentes.

A importância deste trabalho está na contribuição aos procedimentos de pré-processamento das imagens de palavras manuscritas, dentre os quais a correção da inclinação vertical dos caracteres (*slant*).

Het boek is een feitelijk objectieve studie
 van de methoden, de computeren hebben
 uitgewerkt op mensen en organisatie
 en is gebaseerd van strikte feitelijke
 en wetenschappelijke kennis. Het boek is
 een wetenschappelijke en praktische methode van
 personen in alle rangen en standen, van
 de beschouwing genomen. De behandeling is
 niet theoretisch. Het is een zeer
 uitvoerig en uitgebreid boek van de computer.
 De geschiedenis van de computer leeft worden
 beschreven en de toekomstige trend en de

Het boek is een feitelijk objectieve studie
 van de methoden, de computeren hebben
 uitgewerkt op mensen en organisatie
 en is gebaseerd van strikte feitelijke
 en wetenschappelijke kennis. Het boek is
 een wetenschappelijke en praktische methode van
 personen in alle rangen en standen, van
 de beschouwing genomen. De behandeling is
 niet theoretisch. Het is een zeer
 uitvoerig en uitgebreid boek van de computer.
 De geschiedenis van de computer leeft worden
 beschreven en de toekomstige trend en de

Het boek is een feitelijk objectieve studie
 van de methoden, de computeren hebben
 uitgewerkt op mensen en organisatie
 en is gebaseerd van strikte feitelijke
 en wetenschappelijke kennis. Het boek is
 een wetenschappelijke en praktische methode van
 personen in alle rangen en standen, van
 de beschouwing genomen. De behandeling is
 niet theoretisch. Het is een zeer
 uitvoerig en uitgebreid boek van de computer.
 De geschiedenis van de computer leeft worden
 beschreven en de toekomstige trend en de

Figura 4.12: Exemplos do manuscrito do mesmo escritor para 3 espaçamentos horizontais:
 a) normal, b) estreito e c) largo. (Fonte:[MAA83])

4.2.4. O Grafema no Espaço e no Tempo

Thomassen & Teulings em [THO83] questionam se um grafema traçado estacionariamente é independente do contexto ou o fato da integração entre repetições do mesmo grafema no sentido do movimento da esquerda para a direita afeta a produção do mesmo?

Para realizar as análises necessárias, os autores selecionaram 3 padrões, a saber: laços (*loops*), ondas (*waves*) e zig-zag com inclinação aproximada de 40° (*zigzags*), como mostrado na Figura 4.13. A repetição destes padrões foi solicitada para crianças (com idade entre 7 e 11 anos) e adultos (com idade entre 18 e 48 anos). Inicialmente, os padrões deviam ser realizados estacionariamente, ou seja, sem deslocamento horizontal. Posteriormente, o escritor devia realizar o padrão com deslocamento. Foi solicitado que a altura e ângulo de inclinação fossem mantidos o mais constante possível. Os dados foram coletados sob mesa digitalizadora.

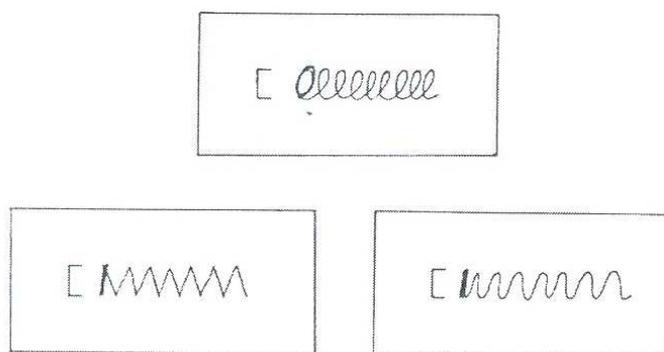


Figura 4.13: Padrões analisados: a) laços, b) ondas e c)zig-zag. (Fonte:[THO83])

Os autores observaram que a realização dos grafemas com deslocamento gera uma diminuição da velocidade de produção dos mesmos. Porém, para maioria dos grafemas utilizados existe a possibilidade destes adquirirem um caráter autônomo em relação ao tempo e espaço. Por exemplo, as letras *a*, *b*, *c*, *d*, *e*, *g*, *l*, *o* contribuem para o estabelecimento de mesmo padrão de movimento: “contorno no sentido do relógio e para baixo, seguido de uma indeterminada curva e uma indeterminada inclinação”. Assim, os laços foram considerados como o padrão mais constante no tempo e espaço.

Este experimento está relacionado com a Figura 3.9 (Capítulo 3) que apresenta as variações geradas na escrita cursiva com base em três tipos básicos de traçados apresentados por [CAV74]. O primeiro traçado origina uma escrita cursiva mais arredondada e com um laço extra na letra *t*, o segundo traçado é o oposto do primeiro originando uma escrita cursiva mais acentuada no que se refere as letras *h*, *s* e *m*. Por outro lado, o terceiro traçado origina

uma escrita cursiva onde as letras minúsculas (corpo da palavra) é mais alto, apresentando um resultado intermediário entre o primeiro e segundo traçados.

Reunindo-se estes autores [CAV74] e [THO83] pode-se dizer que a diferença entre os padrões de traçado não está simplesmente na forma, mas, também na velocidade com que o escritor conseguirá escrever. E ainda, que variações no tempo e espaço do traçado da palavra, decorrentes da velocidade da escrita, auxiliarão ou não na pregnância da forma escrita pelo leitor (ver Figura 2.16).

4.2.5. Variabilidade da Escrita Cursiva em Função da Posição da Letra na Palavra

Wing et al. em [WIN83] afirmam que para for estilo *printed text* (entende-se por *printed text* o mesmo que letra manuscrita com caracteres espaçados ou separados ou, letra de imprensa, ou, ainda, *spaced discrete characters* de acordo com a classificação de [TAP90]) cada letra sucessiva é independente da sua predecessora. O que não ocorre com o estilo *cursive script* (entende-se *cursive script* o mesmo que *curisive writing*, letra cursiva ou *pure cursive script writing* de acordo com a classificação de [TAP90]). Assim, os autores realizaram um estudo para analisar a variabilidade das letras em função da posição da mesma na palavra. A posição da letra na palavra foi classificada, pelos autores, em: no início, no meio ou final.

O estudo tem por base o trabalho de Ellis (1982), o qual é suportado pela hipótese de que a probabilidade de se encontrar a mesma forma de uma letra é maior quando esta posiciona-se no meio ou final da palavra.

Os autores prepararam um texto para que os escritores fossem levados a produzir palavras que contivessem os seguintes grafemas: *f*, *h*, *k*, *p*, *t*. Assim, as palavras continham grafemas no início, meio e final das palavras, distribuídas da seguinte forma:

- Início: *feel*, *head*, *kept*, *person*, *take*,
- Meio: *safe*, *aches*, *banker*, *camped*, *rotate*
- Final: *loaf*, *reach*, *drank*, *tap*, *not*.

A Figura 4.14 exemplifica as variações extraídas das palavras com o grafema *f*. O valor indicado nas células corresponde à parcela em porcentagem (%) que o grafema ocorre para um total de 1065 exemplares. As células me branco(X) indicam que não foram encontrados exemplares desta forma para o grafema analisado.

<i>f</i> 0.1%	<i>f</i> 1.7%	<i>f</i> 0.1%	
--	--	--	
--	<i>f</i> 3.9%	<i>f</i> 22.7%	
<i>f</i> 48.1%	<i>f</i> 15.2%	--	
--	<i>f</i> 0.8%	<i>f</i> 0.1%	
--	<i>f</i> 1.8%	--	
	<i>f</i> 1.6%	<i>f</i> 0.6%	
			<i>f</i> 3.3%

Figura 4.14: Classificação do grafema *f*. (Fonte:[WIN83])

Este trabalho evidencia a diferença na formação de determinados grafemas quando os mesmos se encontram em diferentes partes das palavras. Os autores possibilitam através do estudo que a variabilidade das formas, durante o processo de escrita, venha auxiliar no melhor entendimento da complexidade da letra cursiva.

4.3. Português e Francês: Percepção e Reconhecimento

4.3.1. Considerações Iniciais

Neste ponto deve-se relembrar alguns tópicos já apresentados e discutidos no presente trabalho (Capítulo 2), a saber:

- o todo é mais que a soma das partes,
- a percepção visual é um processo ativo,
- as pessoas organizam o que vêem e lêem como um todo analisado,

- as pessoas interpretam o estímulo visual como completo,
- as pessoas analisam e sintetizam palavras.

E ainda, que ler não é o inverso de escrever. Desta forma, pergunta-se: É possível melhor o reconhecimento de palavras manuscritas em um determinado idioma, utilizando-se palavras de outro idioma?

A idéia é melhorar o reconhecimento das palavras do extenso manuscrito de cheques bancários brasileiros, utilizando-se para isto uma base de palavras francesas, para o mesmo contexto.

Esta hipótese está sustentada em alguns pontos, a saber:

- a origem dos idiomas, português e francês, é a mesma: latina;
- o léxico em questão possui palavras grafêmica e morfologicamente idênticas: *onze*, *quatorze* e *quinze*,
- o léxico em questão possui palavras grafemicamente idênticas: *quatro* e *quatre*, *quarenta* e *quarente*, *cinquenta* e *cinquante*,
- o léxico em questão possui palavras simbolicamente próximas (refere-se ao processo de extração de primitivas e alfabeto de símbolos): *três* e *trois*, *doze* e *douze*, *treze* e *treize*, *trinta* e *trente*, *sessenta* e *soixante*.

Assim, foram realizados experimentos visando analisar esta hipótese. Os experimentos consideram as seguintes bases de dados:

- Léxico em português: base de palavras manuscritas do LUCI/EXTENSO/PALAVRAS, contendo 39 palavras e as características descritas em [FRE01],
- Léxico em francês: base de palavras manuscritas do IRESTE/University of Nantes (França), denominada IRONOFF (IReste ON/OFF Dual Database), para o formulário B, nos campos de B64 até B93, contendo 26 palavras e as características descritas em [VIA99a-b].

4.3.2. Bases de Dados: LUCI/EXTENSO/PALAVRAS e IRONOFF

A Tabela 4.1 apresenta ambos os léxicos para que se possa entender o contexto da hipótese a ser analisada. Observa-se que o léxico em português apresenta uma maior complexidade devido ao seu próprio tamanho (39 palavras).

Nota-se, também, 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 francês, pois os

valores numéricos entre 17 a 19, 70 a 90, e também, os valores entre 200 e 900, os mesmos são grafados por extenso utilizando-se uma composição de palavras. Deste modo, 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 em questão.

Tabela 4.1: Léxico em Português e Francês

Valor Numérico	Português	Francês
1	Um, Hum	Un
2	Dois	Deux
3	Três	Trois
4	Quatro	Quatre
5	Cinco	Cinq
6	Seis	Six
7	Sete	Sept
8	Oito	Huit
9	Nove	Neuf
10	Dez	Dix
11	Onze	Onze
12	Doze	Douze
13	Treze	Treize
14	Quatorze, Catorze	Quatorze
15	Quinze	Quinze
16	Dezesseis	Seize
17	Dezessete	-
18	Dezoito	-
19	Dezenove	-
20	Vinte	Vingt
30	Trinta	Trente
40	Quarenta	Quarante
50	Cinquenta, Cincoenta	Cinquante
60	Sessenta	Soixante
70	Setenta	-
80	Oitenta	-
90	Noventa	-
100	Cem, Cento	Cent
200	Duzentos	-
300	Trezentos	-
400	Quatrocentos	-
500	Quinhentos	-
600	Seiscentos	-
700	Setecentos	-
800	Oitocentos	-
900	Novecentos	-
1000	Mil	Mille
Moeda Corrente	Reais, Real	Francs, Euros
Fração	Centavos, Centavo	Centimes

Outra análise que demonstra a complexidade dos léxicos é o cálculo da entropia $H(P)$, dada por [COV91] e [HUA90]:

$$H(P) = -\sum_{i=1}^N P_i \log_2(P_i) \quad (4.1)$$

O léxico para a aplicação dos extensos manuscritos de cheques bancários brasileiros é composto por 39 palavras diferentes. Assim, o valor da entropia $H(P)$ para o presente estudo, calculado sobre a frequência relativa das palavras base de treinamento, resulta:

$$H(P) = -P_0 \log_2(P_0) - \dots - P_{39} \log_2(P_{39}) = 4,77 \text{ bits} \quad (4.2)$$

Sendo, a unidade de $H(P)$ o bit. O bit é o valor de incerteza associado à seleção de um evento entre dois eventos equiprováveis (probabilidade = 0,5). Portanto, o valor de entropia quantifica a variabilidade do problema. Este valor é a quantidade de informação mínima necessária a fim de poder resolver o problema de classificação de palavras. O valor da entropia $H(P)$ permite comparar o problema do presente trabalho com um problema de 27 classes equiprováveis, pois $\log_2(27,28) = 4,77$.

Para o léxico em francês, considerando-se a frequência das palavras da base IRONOFF para treinamento, o valor da entropia $H(P)$ resulta:

$$H(P) = -P_0 \log_2(P_0) - \dots - P_{27} \log_2(P_{26}) = 4,69 \text{ bits} \quad (4.3)$$

Assim sendo, o valor da entropia $H(P)$ para o léxico em francês/IRONOFF permite comparar o problema do reconhecimento deste léxico com um problema de 26 classes equiprováveis, pois $\log_2(25,98) = 4,69$. Isto ocorre, pois a base IRONOFF possui uma distribuição equiprovável entre as 27 classes de palavras. Portanto, os léxicos em português e francês para as bases de dados consideradas apresentam o mesmo grau de complexidade, no que se refere à entropia.

A Tabela 4.2 apresenta a distribuição das imagens das bases de treinamento para os léxicos em português [FRE01] e francês/IRONOFF [VIA99a-b]. Observa-se a grande diferença de homogeneidade entre as quantidades de exemplares para treinamento na base em português, destacando-se a classe da Dezena com somente 3,16% de exemplares na base de treinamento.

Tabela 4.2: Distribuição das Imagens para as Bases de Treinamento (¹ PC = palavras-chaves)

	Palavras	Português	Palavras	Francês
U N I D A D E	Um	199	Un	243
	Dois	209	Deux	240
	Três	201	Trois	240
	Quatro	193	Quatre	241
	Cinco	199	Cinq	244
	Seis	181	Six	243
	Sete	160	Sept	243
	Oito	182	Huit	242
	Nove	190	Neuf	242
	%	23,98	%	35,28
D E Z E N A	Dez	28	Dix	243
	Onze	29	Onze	244
	Doze	17	Douze	242
	Treze	22	Treize	244
	Quatorze	20	Quatorze	240
	Quinze	26	Quinze	238
	Dezesseis	17	Seize	238
	Dezessete	21	-	-
	Dezoito	20	-	-
	Dezenove	26	-	-
%	3,16	%	27,36	
E N T A	Vinte	229	Vingt	241
	Trinta	223	Trente	237
	Quarenta	202	Quarante	234
	Cinquenta	225	Cinquante	227
	Sessenta	213	Soixante	232
	Setenta	221	-	-
	Oitenta	217	-	-
	Noventa	189	-	-
%	24,06	%	18,97	
E N T O S	Cem	4	-	-
	Cento	161	Cent	236
	Duzentos	149	-	-
	Trezentos	141	-	-
	Quatrocentos	130	-	-
	Quinhentos	143	-	-
	Seiscentos	131	-	-
	Setecentos	148	-	-
	Oitocentos	151	-	-
	Novocentos	141	-	-
%	18,23	%	3,82	
P C ¹	Mil	771	Mille	233
	Reais/Real	757	Francs/Euros	225/221
	Centavos/Centavo	656	Centimes	220
	%	30,56	%	14,56
	Total	7.146	Total	6.173

4.3.3. Resultados Alcançados por Freitas em [FRE01]

Os resultados dos experimentos realizados com 3 conjuntos de primitivas para o léxico em português estão apresentados na Tabela 4.3. Entende-se por PP – primitivas perceptivas, PPCC – Primitivas Perceptivas, Concavidades e Convexidades e PPCCR – Primitivas Perceptivas, Concavidades e Convexidades Rotuladas.

Tabela 4.3: Taxas de Reconhecimento das Palavras para os Conjuntos de Primitivas

	Palavras	PP (%)	PPCC (%)	PPCCR (%)
U N I D A D E	Um	80,30	83,33	92,42
	Dois	64,29	65,71	72,86
	Três	47,68	55,22	59,70
	Quatro	70,31	78,13	81,25
	Cinco	63,64	74,98	77,27
	Seis	35,00	50,00	51,67
	Sete	43,40	51,77	43,40
	Oito	73,77	73,77	73,77
	Nove	61,90	75,37	76,19
	D E Z E N A	Dez	66,67	55,56
Onze		70,00	40,00	50,00
Doze		40,00	40,00	20,00
Treze		85,71	57,14	71,43
Quatorze		28,57	42,86	28,57
Quinze		55,56	77,78	55,56
Dezesseis		50,00	33,33	16,67
Dezessete		28,57	57,14	28,57
Dezoito		12,24	48,98	66,67
Dezenove		55,56	70,31	88,89
E N T A	Vinte	52,63	72,37	63,16
	Trinta	63,51	35,14	47,30
	Quarenta	59,70	62,69	65,67
	Cinquenta	78,67	84,92	78,67
	Sessenta	20,29	37,67	36,62
	Setenta	32,43	48,65	37,84
	Oitenta	50,00	47,22	55,56
	Noventa	52,38	71,43	65,08
E N T O S	Cem/Cento	47,35	58,18	58,18
	Duzentos	67,35	69,39	69,39
	Trezentos	31,91	46,81	53,19
	Quatrocentos	58,14	69,07	67,44
	Quinhentos	81,25	78,23	79,17
	Seiscentos	28,87	39,98	44,19
	Setecentos	55,10	57,67	57,14
	Oitocentos	54,00	70,80	66,00
	Novecentos	76,60	68,09	80,85
P C	Mil	59,92	83,27	84,82
	Reais/Real	44,14	68,65	67,46
	Centavos/Centavo	80,83	86,30	83,03
	Média	57,22	67,34	67,66

4.3.4. Hipótese e Experimentos

A hipótese avaliada com os experimentos realizados é a seguinte: É possível melhorar o reconhecimento de palavras da base em português (LUCI/EXTENSO/PALAVRAS) utilizando-se para isto palavras da base em francês (IRONOFF)?

A idéia é poder melhorar o reconhecimento das palavras em português que possuem uma quantidade pequena de exemplares para treinamento dos modelos MEM – Modelos Escondidos de Markov (HMM – Hidden Markov Models) [RAB93]. Isto significa, melhorar o treinamento dos modelos com a inserção de bases de treinamento em francês.

Para tal, foi realizado o Experimento 01 considerando somente a base de palavras em francês, para avaliar a taxa de reconhecimento global (para todo o léxico) e as taxas individuais (para as palavras). Este experimento considerou a obtenção dos modelos para as 26 palavras do léxico, sendo realizados 3 testes de reconhecimento considerando-se as seguintes situações: léxico com *francs*, léxico com *euros*, léxico com *francs* e *euros*. A Tabela 4.4 apresenta os resultados obtidos. Todos os experimentos foram realizados extraindo-se o conjunto de primitivas PPCCR.

Tabela 4.4: Taxas de Reconhecimento das Palavras em Francês

Palavras	Exp01-a	Exp01-b	Exp01-c
un	54,3	54,3	54,3
deux	35,4	35,4	35,4
trois	43,8	42,5	42,5
quatre	85,0	85,0	85,0
cinq	58,0	55,6	55,6
six	51,9	51,9	51,9
sept	86,3	86,3	86,3
huit	70,0	70,0	70,0
neuf	86,3	86,3	86,3
dix	43,2	42,0	42,0
onze	87,7	87,7	87,7
douze	85,0	85,0	85,0
treize	59,3	59,3	59,3
quatorze	70,0	70,0	70,0
quinze	49,4	49,4	49,4
seize	65,8	65,8	65,8
vingt	45,6	45,6	45,6
trente	82,1	83,3	82,1
quarante	63,6	64,9	63,6
cinquant	68,0	68,0	68,0
soixante	35,1	35,1	35,1
cent	66,7	66,7	66,7
mille	75,3	75,3	75,3
francs	50,0	-	50,0
centimes	65,8	56,2	65,8
euros	-	65,8	54,8
Média	63,4	63,5	62,9

Observa-se na Tabela 4.4 que as palavras *quatre, onze, douze, quatorze, trente e cent* apresentam uma taxa de reconhecimento elevado, principalmente quando comparadas com as taxas obtidas com o léxico em português (Tabela 4.3).

Portanto, foram realizados experimentos seguindo-se as etapas descritas a seguir:

- Treinamento dos modelos MEM de interesse utilizando-se as seqüências de observações das bases de dados em português e em francês,
- Teste de reconhecimento utilizando-se as seqüências de observações das bases de dados em português.

O treinamento dos modelos segue a mesma topologia, os mesmos algoritmos e procedimentos apresentados por Freitas em [FRE01]. Deste modo, os experimentos realizados permitiram obter os resultados apresentados na Tabela 4.5, considerando-se para cada teste os seguintes modelos modificados pela inserção da base de treinamento de palavras francesas:

- Experimento 02: quatro, onze, doze, quatorze, trinta,
- Experimento 03: quatro, onze, doze, quatorze,
- Experimento 04: quatro, onze, doze, quatorze, trinta, um, dois, três, treze, quinze, quarenta, cinquenta, sessenta,
- Experimento 05: quatro, onze, doze, quatorze, trinta, três, treze, quinze, quarenta, cinquenta, sessenta,
- Experimento 06: quatro, onze, doze, quatorze, trinta, três, treze, quinze, quarenta, cinquenta, sessenta, cento, mil.

Na Tabela 4.5, os valores apresentados em **negrito** representam as palavras com taxas que obtiveram aumento, os valores em *itálico* indicam as palavras que obtiveram uma diminuição da taxa de reconhecimento e os demais valores mantiveram-se iguais. Portanto, o melhor resultado no que concerne a taxa média de reconhecimento foi alcançada pelo Experimento 05, envolvendo a modificação dos modelos: quatro, onze, doze, quatorze, trinta, três, treze, quinze, quarenta, cinquenta, sessenta.

O Experimento 05 representa um ganho em 22 palavras da base de dados em português e uma perda em 7 palavras, mantendo-se inalteradas as taxas de 10 palavras. Com base nestes resultados pode-se realizar uma análise mais detalhada, principalmente, comparando-se as confusões entre as palavras.

Tabela 4.5: Taxas de Reconhecimento das Palavras em Português com Treinamento dos Modelos MEM em Português + Francês

Palavras	Exp02	Exp03	Exp04	Exp05	Exp06
Um	92,42	65,15	91,16	90,91	89,39
Dois	75,71	54,29	53,69	74,29	74,29
Três	59,67	49,23	61,19	61,19	63,24
Quatro	86,33	79,31	82,81	82,81	82,81
Cinco	77,27	40,91	77,27	77,27	78,79
Seis	51,67	23,33	53,33	50,00	50,00
Sete	47,17	50,94	45,28	47,17	47,17
Oito	75,41	72,13	77,05	75,41	77,05
Nove	76,19	46,03	77,78	77,78	77,78
Dez	55,56	88,89	66,67	66,67	66,67
Onze	78,92	70,27	80,00	80,00	80,00
Doze	82,25	75,49	76,62	80,00	80,00
Treze	55,25	55,25	85,71	85,71	85,71
Quatorze	85,71	68,57	85,71	85,71	85,71
Quinze	43,45	32,59	88,89	88,89	88,89
Dezesseis	16,67	0,00	16,67	16,67	16,67
Dezessete	28,57	14,29	28,57	28,57	28,57
Dezoito	66,67	66,67	66,67	66,67	66,67
Dezenove	88,89	77,78	77,78	77,78	77,78
Vinte	63,16	38,16	65,79	64,47	64,47
Trinta	37,10	35,38	33,78	33,78	37,76
Quarenta	65,53	61,06	77,61	77,61	77,61
Cinquenta	77,33	64,00	86,67	86,67	86,67
Sessenta	38,44	28,47	45,07	45,07	45,07
Setenta	39,19	28,38	39,19	39,19	39,19
Oitenta	55,56	51,39	54,17	54,17	54,17
Noventa	65,08	58,73	66,67	66,67	66,67
Cem/Cento	60,00	60,00	60,00	60,00	49,62
Duzentos	69,39	28,57	71,43	71,43	71,43
Trezentos	53,19	59,57	53,19	53,19	53,19
Quatrocentos	67,44	39,53	67,44	67,44	67,44
Quinhentos	79,17	39,58	77,08	77,08	77,08
Seiscentos	46,51	9,30	41,86	41,86	41,86
Setecentos	57,14	24,49	57,14	57,14	57,14
Oitocentos	66,00	44,00	68,00	68,00	68,00
Novocentos	80,85	21,28	80,85	80,85	80,85
Mil	85,21	77,82	84,82	85,21	82,04
Reais/Real	68,25	36,51	68,65	68,25	68,25
Centavos/Centavo	83,03	46,79	83,03	83,03	83,94
Média	67,84	50,86	68,78	69,09	68,83

4.3.5. Análise dos Resultados

A análise dos resultados obtidos no Experimento 05 permite uma comparação com os resultados obtidos por Freitas em [FRE01], como mostrado na Tabela 4.6. Nesta tabela apresenta-se também, a indicação das palavras onde houve ganho ou perda na taxa de reconhecimento.

Nota-se, não somente ganhos relativos aos modelos modificados (*três, quatro, onze, doze, treze, quatorze, quinze, quarenta, cinquenta, sessenta*), mas sim ganhos indiretos (*dois,*

sete, oito, nove, vinte, setenta, noventa, cem/cento, duzentos, oitocentos, mil e reais) com uma diminuição da confusão entre determinadas palavras. Com modificação de 11 modelos gerou-se uma melhora em 22 palavras. Com exceção da palavra *trinta*, todas as demais palavras com modelos modificados obtiveram ganho na taxa de reconhecimento. Deve-se ressaltar os ganhos obtidos com as palavras: *onze, doze, treze, quatorze e quinze*.

Tabela 4.6: Comparação de Resultados: Taxas de Reconhecimento

	Palavras	[FRE01] (%)	Exp05 (%)	Análise
U N I D A D E	Um	92,42	90,91	Perda
	Dois	72,86	74,29	Ganho indireto
	Três	59,70	61,19	Ganho
	Quatro	81,25	82,81	Ganho
	Cinco	77,27	77,27	-
	Seis	51,67	50,00	Perda
	Sete	43,40	47,17	Ganho indireto
	Oito	73,77	75,41	Ganho indireto
	Nove	76,19	77,78	Ganho indireto
	Média	70,53	71,40	-
D E Z E N A	Dez	66,67	66,67	-
	Onze	50,00	80,00	Ganho
	Doze	20,00	80,00	Ganho
	Treze	71,43	85,71	Ganho
	Quatorze	28,57	85,71	Ganho
	Quinze	55,56	88,89	Ganho
	Dezesseis	16,67	16,67	-
	Dezessete	28,57	28,57	-
	Dezoito	66,67	66,67	-
	Dezenove	88,89	77,78	Perda
Média	52,00	69,33	-	
E N T A	Vinte	63,16	64,47	Ganho indireto
	Trinta	47,30	33,78	Perda
	Quarenta	65,67	77,61	Ganho
	Cinqüenta	78,67	86,67	Ganho
	Sessenta	36,62	45,07	Ganho
	Setenta	37,84	39,19	Ganho indireto
	Oitenta	55,56	54,17	Perda
	Noventa	65,08	66,67	Ganho indireto
	Média	56,12	58,22	-
E N T O S	Cem/Cento	58,18	60,00	Ganho indireto
	Duzentos	69,39	71,43	Ganho indireto
	Trezentos	53,19	53,19	-
	Quatrocentos	67,44	67,44	-
	Quinhentos	79,17	77,08	Perda
	Seiscentos	44,19	41,86	Perda
	Setecentos	57,14	57,14	-
	Oitocentos	66,00	68,00	Ganho indireto
	Novencentos	80,85	80,85	-
	Média	64,05	64,27	-
P C	Mil	84,82	85,21	Ganho indireto
	Reais/Real	67,46	68,25	Ganho indireto
	Centavos/Centavo	83,03	83,03	-
	Média	78,26	78,86	-
Média	67,66	69,09	-	

As Tabelas 4.7 e 4.8 apresentam as matrizes de confusão obtidas por Freitas em [FRE01] e com os modelos modificados, respectivamente. Observa-se na Tabela 4.8 a melhora no reconhecimento das palavras da meta-classe da *dezena*, tanto no que se refere às confusões intra-classe e inter-classe. Obtiveram-se 16,0% para confusões intra-classe, 14,7% para inter-classe e 69,3% de reconhecimento em relação ao total de imagens para testes (75 imagens). Freitas em [FRE01] estas taxas eram: 28,1%, 19,9% e 52,0%, respectivamente. A Tabela 4.9 apresenta o ganho obtido para as meta-classes, destacando-se um ganho de 17,33% para a meta-classe da *dezena*.

Tabela 4.9: Acréscimo da Taxa de Reconhecimento para as Meta-Classes

Meta-Classe	[FRE01] (%)	Exp05 (%)	Análise
Unidade	70,53	71,40	+ 0,87
Dezena	52,00	69,33	+ 17,33
Enta	56,12	58,22	+ 2,10
Entos	64,05	64,27	+ 0,22
Palavras-Chaves	78,26	78,86	+ 0,60
Média	67,66	69,09	+ 1,43

Outro experimento realizado com as palavras considera a probabilidade de ocorrência de cada uma das palavras na base de dados de treinamento através da Equação 4.4:

$$k = \arg \max_i [\text{pr}(O / \lambda_i) \cdot p(\lambda_i)] \quad (4.4)$$

Sendo que k é o índice i que maximiza a função ($i = 1, \dots, N_c$), N_c é o número de palavras do léxico (39 palavras), O é a seqüência de observações da palavra e $p(\lambda_i)$ é a freqüência relativa da palavra i na base de dados de treinamento. E ainda, λ_i é o modelo da palavra i .

Realiza-se este experimento para considerar da freqüência de ocorrência na base de dados de treinamento de cada palavra sobre a probabilidade calculada para a palavra desconhecida e, posteriormente, computar as hipóteses de reconhecimento.

Assim, pode-se verificar um acréscimo da taxa de reconhecimento em relação a todos os demais experimentos realizados neste trabalho e por Freitas em [FRE01], como mostrado na Tabela 4.10. Confirmando, que a freqüência de ocorrência das palavras na base de treinamento deve ser levada em conta no momento da formulação das hipóteses de reconhecimento.

Tabela 4.10: Taxa de Reconhecimento dos Conjuntos de Primitivas Perceptivas para Equação (4.4)

Experimentos	TOP1	TOP2	TOP3	TOP4	TOP5
Conjunto - PP	57,22	73,16	80,32	84,78	88,10
Conjunto - PPCC	67,34	79,97	85,43	88,87	91,12
Conjunto - PPCCR	67,66	80,42	86,65	89,94	92,21
PPCCR + ProbClasse	70,61	82,44	88,08	90,53	92,84
Exp05	69,09	82,48	88,00	91,05	93,52
Exp05 + ProbClasse	71,96	84,40	89,45	92,69	94,17

Finalmente, conclui-se que os resultados obtidos são promissores, principalmente, quando comparados aos de outros autores e trabalhos já realizados, conforme apresentado em Freitas [FRE01] - Apêndice B. Portanto, pode-se considerar a hipótese de que é possível melhorar o reconhecimento de palavras da base em português (LUCI/EXTENSO/PALAVRAS) utilizando-se para isto palavras da base em francês (IRONOFF).

Assim, propõe-se um acréscimo da base de dados de testes da classe de palavras da *dezena*, de forma a atingir a mesma proporção que as demais classes, para que se possa confirmar os resultados obtidos.

A discussão referente a este capítulo está inserida nas Conclusões deste trabalho, juntamente com a proposição de trabalhos futuros.

Conclusão

O entendimento mais detalhado dos processos de escrita e leitura humana permite estudar as características relevantes pelo ser humano quando da realização da leitura. Pois para Suen em [SUE98] a escrita é o meio mais corrente de comunicação visual para o ser humano.

Deste modo, o presente trabalho aprofunda o estudo da relação entre a percepção humana e o processo de leitura para extrair subsídios ao reconhecimento automático de palavras manuscritas. Relacionam-se a percepção humana e a leitura com a extração de primitivas para os processos automáticos de reconhecimento.

Confirma-se a possibilidade de melhorar o reconhecimento de palavras manuscritas para um léxico em português através do uso de um léxico em francês. Pois, alcançou-se uma taxa de reconhecimento médio para as 39 palavras de 69,09% (Exp05) e 71,96% (Exp05 + ProbClasse).

O presente trabalho coloca-se, ainda, como mais uma contribuição na área de Análise e Reconhecimento de Documentos, podendo vir auxiliar outras aplicações, 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.

Assim, deve-se lembrar alguns pontos relevantes deste trabalho:

- “o todo é mais que a soma das partes” (Wertheimer, 1910),
- a percepção é um processo ativo,
- as pessoas organizam o que vêem,

- as pessoas agrupam as figuras dependendo das propriedades visuais das figuras,
- as pessoas interpretam o estímulo visual como completo.
- as pessoas podem adicionar uma dimensão: de 2-D para 3-D,
- As pessoas lêem como um todo analisado,
- A percepção visual pode auxiliar os sistemas automáticos de reconhecimento,
- O leitor iniciante e o leitor eficiente,
- As pessoas analisam e sintetizam palavras,
- As pessoas lêem rapidamente graças aos mecanismos e habilidades do cérebro,
- Ler não é o inverso de escrever.

E ainda, em relação a percepção e o reconhecimento de palavras manuscritas pode-se concluir que:

a) **As imagens e os níveis de abstração:** Os níveis de abstração auxiliam na compreensão da complexidade do processamento de imagens de palavras manuscritas, não somente do ponto computacional, mas também das limitações inseridas nos sistemas automáticos quando da passagem do processo humano de leitura para o processo computacional de reconhecimento de palavras,

b) Vogais e Consoantes, Palavras Curtas e Longas, Ascendentes e Descendentes:

A hipótese formulada para realização dos experimentos leva em consideração que a modificação dos modelos pela inserção das bases de palavras francesas no treinamento dos modelos das palavras em português utiliza palavras que possuem primitivas perceptivas (ascendentes e descendentes) originalmente nas mesmas posições. Isto representa que a informação posicional dos ascendentes/descendentes (referentes as consoantes: d, q, t, z) são mais significativas do que a das vogais, como afirmado por [SCH98]. Observe:

três	trois
quatro	quatre
onze	onze
doze	douze
treze	treize
quatorze	quatorze
quinze	quinze
trinta	trente
quarenta	quarante
cinquenta	cinquante
sessenta	soixante

Além disso, comprova-se que as vogais (a,e,i,o) geram um mesmo comportamento do leitor para Scomaker & Segers [SCH98], o que pode ser confirmado pelos experimentos realizados (por exemplo: trinta e trente),

c) Variabilidade da Escrita Cursiva em Função da Posição da Letra na Palavra:

Constata-se que os grafemas decorrentes das consoantes d, q, t, z, não sofrem alteração no espaço e tempo, uma vez que originalmente os mesmos se encontram nas mesmas posições, tanto no léxico em português como no em francês. Deste modo, a variabilidade entre as bases de dados é menor, permitindo assim a melhora do treinamento dos modelos.

As questões abordadas neste trabalho permitiram o estudo da similaridade das formas das palavras encontradas em bases de dados contendo léxicos diferentes: português e francês. Busca-se uma nova chance de ler o que está escrito.

Esta capacidade de complementação das bases de dados deve, portanto, ser melhor explorada, permitindo que outros trabalhos possam ser realizados. Isto devido ao fato de que os mesmos contribuirão 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 necessidade de continuidade de coleta da base de dados de cheques bancários brasileiros de laboratório, para que os demais estudos realizados possam contar com um número de imagens, se não o ideal, ao menos equiprovavelmente distribuídas.

Com base nos tópicos apresentados e discutidos neste trabalho, pode-se propor a continuidade dos estudos da modelagem de palavras manuscritas através dos seguintes trabalhos futuros:

- Estudo dos pontos dominantes nas palavras, tendo por base o trabalho de Li & Yeung em [LI 97]: a idéia é considerar como pontos dominantes os *end-points*, os *branch-points* e os *crossing-points* encontrados nas palavras manuscritas,
- Estudo da forma da palavra (ver Figura 4.4), tendo por base o trabalho de Powalka et al. [POW97]: a idéia é explorar informações topológicas e geométricas a serem extraídas das zonas das palavras: ascendente, corpo e descendente.

Referências Bibliográficas¹

- [ARN97] ARNHEIM, R. *Arte e percepção visual: uma psicologia da visão criadora*. Tradução Ivonne Terezinha de Faria. São Paulo, SP: Pioneira:Editora da Universidade de São Paulo, 1997. 503p.
- [AUM93] AUMONT, J. *A imagem*. Tradução Estela dos Santos Abreu. Campinas, SP: Papirus, 1993. 317p. [Online]. Available: <http://www3.lmi.ub.es/documents/textos/nav1/tema6/aumont/aumont.htm> [03 jul 2002]
- [AVI96] ÄVILA, M. *Optimisation de Modeles Markoviens pour la Reconnaissance de L'ecrit*. Thèse de doctorat, Université de Rouen, France, 1994. 215p.
- [BEU92] BEUCHER, S. & MEYER, F. *The morfological approach to segmentation: the watershed transformation*. Mathematical Morphology in Image Processing, Marcel Dekker Inc, 1992, p.433-481.
- [BOZ89] BOZINOVIC, R.M. & SRIHARI, S.N. *Off-line cursive script recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.11, No. 1, 1989, 68-83p.
- [CAG91] CAGLIARI, L.C. *Alfabetização & lingüística*. São Paulo: Editora Scipione. 1991. 96-166p.
- [CAV74] CAVANAGH, J.A. *Lettering and alphabets. 85 complete alphabets*. New York: Dover Publications, Inc., 1974. 121p.
- [CHE93] CHEN, M.Y. & KUNDU, A. *An alternative to variable duration HMM in handwritten word recognition*. IWFHR'93. 1993, 82-91p.
- [CHE95] CHEN, S.; HARALICK, R.M.; PHILLIPS, I.T. *Simultaneous Word Segmentation from Document Images Using Recursive Morphological Closing Transform*. Montreal, ICDAR'95, IEEE 1995, 761-764p.
- [CHO95] 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, 1941-1953p.

¹ Este modelo segue o estilo "alpha" do sistema LaTeX, que permite memorizar melhor as referências durante a leitura do texto e é bem mais compacto que o proposto pela ABNT.

- [CON96] CONNELL, S. *A comparison of hidden markov model features for the recognition of cursive handwriting*. Master Dissertation, Michigan State University, USA, 1996. 154p.
- [CÔT97] 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.
- [CÔT98] CÔTE, M.; LECOLINET, E.; CHERIET, M.; SUEN, C.Y. *Automatic reading of cursive scripts using a reading model and perceptual concepts. The PERCEPTO system*. International Journal on Document Analysis and Recognition -IJ DAR, 1998. 3-17p.
- [COV91] COVER, T.M., THOMAS, J.A. *Elements of Information Theory*. Wiley Series in Telecommunications, 1991.
- [CLO75] CLOUTIER, J. *A era de EMEREC*. Ministério da Educação e Investigação Científica - Instituto de Tecnologia Educativa, 1975. [Online]. Available: <http://www.univ-ab.pt/~bidarra/hyperscapes/video-grafias-326.htm> [21 ago 2002]
- [DZU98] DZUBA, G.; FILATOV, A.; GERSHUNY, D.; KIL, I. *Handwritten word recognition – the approach proved by practice*. IWFHR'98, Taejon, Korea, august, 1998, 99-122p.
- [FAC02] FACON, J. (2002) *Pdimagem modelo*. [Online]. Available: <http://www.ppgia.pucpr.br/~facon/IndexPrincipalBrMaterialGraduacao.htm> [03 set 2002]
- [FAC96] FACON J. *A morfologia matemática: teoria e exemplos*. CEFET-PR, outubro, 1996.
- [FAV92] 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.
- [FRE01] FREITAS, C. O. A. *O Uso de modelos escondidos de markov para reconhecimento de palavras manuscritas*. Tese de Doutorado, PUCPR/PPGIA, Curitiba, PR. 2001. 188p.
- [GIL94] GILLOUX, M. *Hidden markov models in handwriting recognition*. IWFHR'94, 1994, 264-288p.
- [GIL95] 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.

- [GOM95] GOMBRICH, E. H. *Arte e ilusão: um estudo da psicologia da representação pictórica*. Tradução: Raul de Sá Barbosa. São Paulo: Martins Fontes, 1995. 473p.
- [GOM02] GOMES FILHO, J. *Gestalt do objeto - sistema de leitura visual da forma*. São Paulo: Escrituras, 2000. 127 p. [Online]. Available: <http://www.3donline.com.br/artigos-gestalt.asp> [28 jun 2002]
- [GUI95] 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 University, Canada. 1995.
- [GRE79] GREGORY, R.L. *Olho e cérebro: psicologia da visão*. Tradução: Álvaro Cabral. Rio de Janeiro: Zahar Editores, 1979. 251p.
- [HEU94] 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.
- [HUA90] HUANG, X.D.; ARIKI, Y.; JACK, M.A. *Hidden markov models for speech recognition*. Information Technology Series; 7. Edinburgh University Press. 1990, 276p.
- [KAT90] KATO, M. *O aprendizado da leitura*. São Paulo: Martins Fontes, 1990. 121p.
- [KIM96] 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.
- [KIM93] 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, 24-27p.
- [KLE00] KLEIMAM, A. *Oficina de leitura, teoria & prática*. Editora Unicamp, 2000.
- [KNE96] 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.
- [LEC90] 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, Université Pierre et Marie Curie (Paris VI), France. 1990. 283p.
- [LER91] LEROUX, M.; SALOME, J.C.; BADARD, J. *Stratégie de contrôle pour la reconnaissance du montant littéral des chèques postaux*. 1991, 695-706p.

- [LI 97] LI, X. & YUENG, D-Y. *On-line handwritten alphanumeric character recognition using dominant points in strokes*. Pattern Recognition, Vol 30., No. 1, 1997. 31-44p.
- [LII93] LII, J.; PALUMBO, P.W.; SRIHARI, S.N. *Address block location using character recognition and address syntax*. Proceedings of the Second International Conference on Document Analysis and Recognition, 1993. 330-334p.
- [MAA83] MAARSE, F.J. & THOMASSEN, A.J.W.N. *Produced and perceived writing slant: difference between up and down strokes*. Acta Psychologica, 54, 1983. 31-147p.
- [MAC96] MACHADO, A.M.B. & TEIXEIRA, E.A. *Aprendendo leitura dinâmica*. São Paulo: Makron Books, 1996, 87p.
- [MAD01] MADHAVANATH, S. & GOVINDARAJU, V. *The role of holistic paradigms in handwritten word recognition*. IEEE Transactions on Pattern Analysis and machine Intelligence, Vol. 23, No. 2, February, 2001. 149-164p.
- [MAD98] 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, 524-531p.
- [MAD92] 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, 183-199p.
- [MOH95] MOHAMED, M.A. *Handwritten word recognition using generalized hidden markov models*. PhD Thesis, University of Missouri-Columbia, USA, 1995.
- [MOR02] MORITA, M.; OLIVEIRA, L.S.; SABOURIN, R.; BORTOLOZZI, F.; SUEN, C.Y. *An HMM-MLP hybrid system to recognize handwritten dates*. IJCNN'02 - Proceedings of the 2002 International Joint Conference on Neural Networks, 2002.2002. Vol.1, 867-872p.
- [MOR01] MORITA, M.; EL YACOUBI, A.; SABOURIN, R.; BORTOLOZZI, F.; SUEN, C.Y. *Handwritten month word recognition on Brazilian bank cheques*. ICDAR'01 - Proceedings Sixth International Conference on Document Analysis and Recognition, 2001. 972-976p.
- [OLI02-a] OLIVEIRA JR, J. J.; CARVALHO, J. M. DE C.; FREITAS, C. O. DE A.; SABOURIN R. *Feature sets evaluation for handwritten word recognition*. 8th International Workshop on Frontiers in Handwriting Recognition (IWFHR'2002), Niagara-on-the-Lake, Ontario, Canada, ago, 2002. 446-451p.

- [OLI02-b] OLIVEIRA JR, J. J.; CARVALHO, J. M. DE C.; FREITAS, C. O. DE A.; SABOURIN R. *Evaluating NN and HMM classifiers for handwritten word recognition*. 15th Brazilian Symposium on Computer Graphics and Image Processing (SIBGRAP'2002), Fortaleza-CE, out, 2002. p.
- [PAR91] PARRAMÓN, J.M. *Como dibujar letras y logotipos*. Barcelona: Parramón Ediciones, 1991, 112p.
- [POW97] POWALKA, R.K.; SHERKAT, N.; WHITROW, R.J. *Word shape analysis for a hybrid recognition system*. Pattern Recognition, Vol 30., No. 3, 1997. 421-445p.
- [RAB93] RABINER, L. JUANG, B.H. *Fundamentals of speech recognition*. Prentice Hall Inc., London, UK. 1993, 506p.
- [RAM95] RAMESH, J.; KASTURI, R.; SCHUNCK, B.G. *Machine vision*. United State of America, McGraw-Hill, Inc. 1995. 549p.
- [ROS91] ROSA, V.S. *Letras & cartazes*. Porto Alegre: Mercado Aberto, 1991, 79p.
- [RUZ97] Ruzon M. A. (1997) *Texture segmentation. an introduction primer*. [Online] Available: http://robotics.stanford.edu/~ruzon/tex_seg/
- [RYB02] RYBAK, I. A.; GUSAKOVA, V.I.; GOLOVAN, A. V.; PODLADCHIKOVA, L. N.; SHEVTSOVA, N. A. (2002) *BMV: Behavioral Model of Visual Perception and Recognition*. [Online]. Available: <http://www.rybak-et-al.net/vnc.html> [03 jul 2002]
- [SCH98] SCHOMAKER, L. & SEGERS, E. *A method for the determination of features used in human reading of cursive handwriting*. IWFHR'98, 1998, 157-168p.
- [SEN98] SENIOR, A.W. & ROBINSON, A.J. *An off-line cursive handwriting recognition system*. IEEE Transactions on Pattern Analysis and Machine Intelligence. Vol.20, No.3, mar.1998, 309-321p.
- [SOL89] SOLO, D.X. *Bold script alphabets. 100 complete fonts*. New York: Dover Publications Inc., 1989, 100p.
- [SUE83] SUEN, C.Y. *Handwriting generation, perception and recognition*. Acta Psychologica, 54, 1983. 295-312p.
- [SUE97] SUEN, C.Y. & LI, Z.C. *Crucial combinations for the recognition of handwritten letters*. BSDIA'97, Curitiba-PR, nov. 1997, p.19-39.
- [SUE98] SUEN, C.Y. *Réflexions sur la reconnaissance d'écriture cursive*. CIFED'98. Quebec, Canada, may, 1998. 1-8p.

- [TAP90] TAPPERT, C.C.; SUEN, C.Y.; WAKAHARA, T. *The state of art in on-line handwriting recognition*. IEEE Trans. on Pattern Analysis and Machine Intelligence. aug. 1990, p.787-808.
- [TEB96] TEBEROSKY, A. & TOLCHINSKY, L. *Além da alfabetização. a aprendizagem fonológica, ortográfica, textual e matemática*. trad: Stela Oliveira. São Paulo: Editora Ática, 1996. 295p: Cap. 2,9,10.
- [THO83] THOMASSEN, A.J.W.N. & TEULINGS, H.L. *Constancy in stationary and progressive handwriting*. Acta Psychologica, 54, 1983. 179-196p.
- [TRE92] TREVISAN, E.M.C. *Leitura: coerência e conhecimento prévio*. Editora da UFSM, Santa Maria. 1992, 117p.
- [TRI96] 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.
- [TSA85] TSAI, W-H. *Moment-preserving thresholding: a new approach*. Computer Vision, Graphics and Image Processing, Vol. 29, 1985. 377-393p.
- [VIA99a] VIARD-GAUDIN, C. *The ironoff user manual*. IRESTE, University of Nantes, France, 1999.
- [VIA99b] VIARD-GAUDIN, C.; LALLICAN, P.M.; KNERR, S.; BINTER, P. *The ireste on/off (ironoff) dual handwriting database*. IEEE International Conference on Document Analysis and Recognition - ICDAR, 1999. 455-458p.
- [WIN83] WING, A.M.; NIMMO-SMITH, I.; ELDRIDGE, M. A. *The consistency of cursive letter formation as a function of position in the word*. Acta Psychologica, 54, 1983. 295-312p.
- [XAVsd] XAVIER, G. *Mecanismos da memória*. Revista Diálogo Médico;
- [YAC96] EL YACOUBI, A. *Modélisation Markovienne de l'écriture manuscrite. Application à la reconnaissance des adresses postales*. Thèse de doctorat, Université de Rennes, France, 1996.
- [YAC99] EL YACOUBI, A.; GILLOUX, M.; SABOURIN, R.; SUEN, C.Y. *An HMM-based approach for off-line unconstrained handwritten word modeling and recognition*. IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol: 21, Issue: 8, Aug. 1999. 752-760p.
- [YAN96] YAN H., *Unified formulation of a class of image thresholding techniques*. Pattern Recognition, Vol. 29, N° 12, 1996. 2025-2031p.

[YON02] YONEKURA, E.A. *Segmentação de envelopes postais complexos por abordagem morfológica utilizando matriz de co-ocorrência*, Dissertação de mestrado, PUCPR, maio 2002.

[YU 97] YU, B.; JAIN, A.K.; MOHIUDDIN, M. *Address block location on complex mail pieces*. Proceedings of the Fourth International Conference on Document Analysis and Recognition, 1997, Vol. 2,1997. 897-901p.