

Sistema de Processamento Digital de Imagem para WINDOWS

Anabela Godinho, Osvaldo Santos, A. Sousa Pereira

Resumo- Este artigo descreve um sistema genérico de Processamento Digital de Imagem em ambiente Microsoft WINDOWSTM. São usadas várias técnicas, entre as quais Transformadas de Imagem (Fourier, Walsh, Hadamard), Processamento sobre o Histograma, Filtragem Espacial e Processamento sobre a paleta de cores. Também é descrita uma técnica de reconhecimento automático de caracteres.

Abstract- The goal of this paper is to describe a generic Digital Image Processing Package in Microsoft WINDOWSTM environment. Several techniques are used, such as Image Transforms (Fourier, Walsh, Hadamard), Histogram processing, Spatial Filtering and Color Map processing. An automatic character recognition technique is also described.

I. INTRODUÇÃO

A área de Processamento Digital de Imagem tem crescido consideravelmente de importância, com a sua crescente utilização em inúmeras aplicações, permitido pela melhoria das performances e redução de preços dos computadores e *hardware* gráfico.

O processamento de imagem desempenha hoje um papel importante em áreas tão diversas como a Medicina, Astronomia, Geografia, Indústria, Ambiente, Exploração espacial, etc.

A lista de aplicações desta área é tão vasta que só é limitada pela imaginação:

- Transmissão de imagens de sondas espaciais.
- Videofone.
- Aquisição de imagens por meios não ópticos.
- Melhoria da qualidade de imagem.
- Classificação automática de terreno, e detecção de recursos disponíveis por análise de imagens de satélite.
- Detecção de agressões ambientais por análise de imagens de satélite.
- Aquisição e reconstrução de imagens médicas.
- Realização automática de mapas a partir de imagens aéreas.
- Detecção de irregularidades em peças de máquinas, a partir de radiografias industriais.

- Reconhecimento automático de formas e objectos.
- etc.

A versão IDIPW 3.0 (*Interactive Digital Image Processing for Windows*) surge como uma versão para Windows 3.1 ou NT do sistema existente para DOS, o IDIPP 2.0.

A grande diferença para a versão DOS é que não é necessário *hardware* dedicado para o sistema funcionar. A utilização do sistema operativo Windows permite uma independência do *hardware* quase total, bem como *multitasking* [2].

Este sistema implementa de uma forma articulada, um conjunto de técnicas de processamento digital de imagem, de tal forma que estas ferramentas podem ter uma utilização genérica, independentemente do tipo de imagens sobre as quais elas são aplicadas.

Foi dado um ênfase especial na facilidade e possibilidade de interacção com o utilizador, criando um ambiente intuitivo e *user-friendly*, procurando sempre que possível um alinhamento com outras interfaces de aplicações Windows conhecidas.

Na última década, houve um interesse crescente em sistemas de reconhecimento de padrões. Dentro deste contexto, foi implementado um sistema adaptativo de aprendizagem e reconhecimento automático de caracteres, que, devido à falta de tempo para o desenvolver, não passa de um protótipo de uma ferramenta, que poderia ter imensas utilizações.

Das técnicas de processamento digital de imagem inicialmente previstas foram implementadas as seguintes:

- *Load/save* de imagens em diversos formatos.
- Definição de zonas de interesse.
- *Hardcopy*;
- *Undo/Redo*.
- Operações com o *clipboard*.
- Conversão de cores para *gray scale*.
- Segmentação por *threshold*.
- Segmentação por crescimento de regiões.
- Transformadas de imagem.
- Filtragem no domínio da frequência.
- Filtragem no domínio espacial.
- Melhoria de imagem.

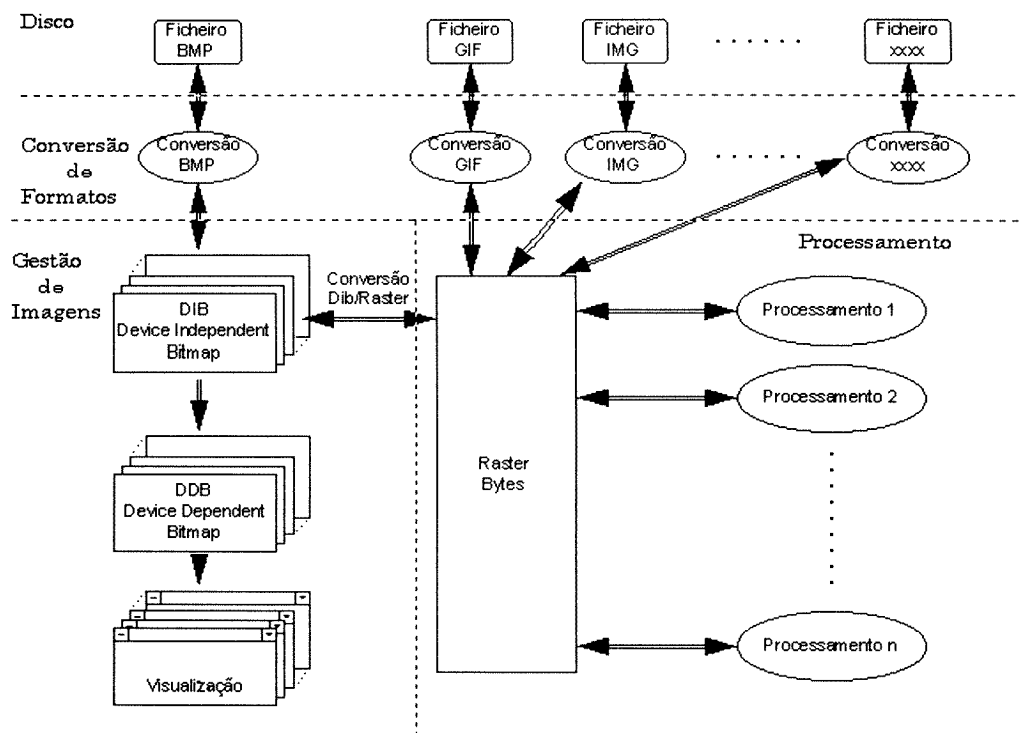


Fig. 1 - Estrutura lógica do programa.

- Visualização do histograma da imagem.
- Operações com o histograma.
- Aprendizagem de caracteres.
- Reconhecimento automático de caracteres.

II. ESTRUTURA DO SISTEMA

O sistema é composto por:

- PC 386 ou superior.
- Sistema Operativo WINDOWS 3.1 ou NT.
- Software IDIPW 3.0.

O sistema não necessita de *hardware* dedicado, no entanto é aconselhável um mínimo de 4 Mbytes de memória RAM, placa gráfica com possibilidade de visualização de 256 cores simultâneas com resolução elevada (1024 x 768 é razoável) e monitor compatível com a placa gráfica, de preferência > 14".

A estrutura lógica do programa (ilustrada na Fig. 1) foi desenvolvida de maneira a facilitar a introdução de novos módulos e ferramentas, independentemente do formato original da imagem, criando um ambiente de suporte às ferramentas de processamento. Este ambiente é responsável pela gestão de múltiplas imagens, fornecendo aos módulos de processamento a visão de uma única imagem, sempre no formato RASTER. Este formato é independente do número de bits por pixel da imagem, bem como do formato original desta. Os módulos de processamento não têm assim que se preocupar com

aspectos como por exemplo a visualização, *load*, *save*, etc.

III. INTERACÇÃO COM O UTILIZADOR

A interacção com o utilizador é feita através de uma interface *WINDOWS user friendly*, utilizando *menu* e *tool bar*, pelo que para um utilizador com alguma experiência neste ambiente não é necessário muito tempo de aprendizagem. A Fig. 2 ilustra um possível cenário de utilização do programa.

A janela principal serve de suporte a toda a interface gráfica do programa. É nesta janela que se encontram o *menu* e a *tool bar*. A *tool bar* permite um rápido acesso às ferramentas mais utilizadas, podendo ser posicionada em quatro posições diferentes, ao gosto do utilizador. É implementada a filosofia *WYSIWYG* (*What You See Is What You Get*), inibindo as ferramentas que não podem ser utilizadas em determinadas situações. Este tipo de filosofia é também aplicado ao *menu*.

IV. ALGUMAS FERRAMENTAS IMPLEMENTADAS

A. Melhoria de Imagem

Este módulo implementa rotinas de melhoria de imagem que permitem alterar dinamicamente a imagem a processar, jogando com cada uma das componentes RGB,

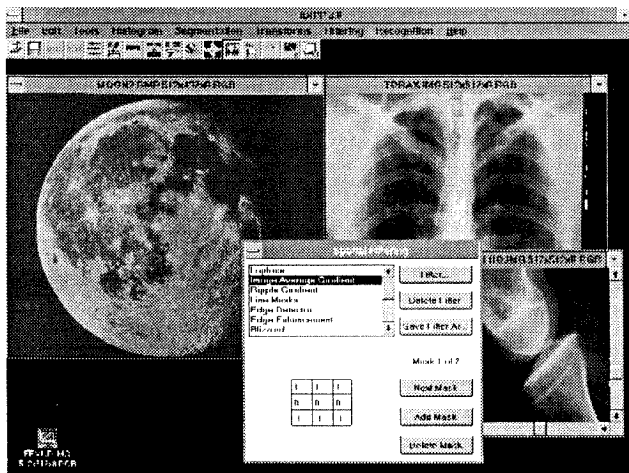


Fig. 2 - Possível cenário de utilização do programa.

ou características como o brilho, contraste e saturação. É utilizada animação dinâmica da *palette* na fase de ajuste, permitindo a visualização em tempo real das alterações efectuadas.

A interacção com o utilizador é feita através de uma caixa de diálogo com seis "potenciômetros", ilustrada na Fig. 3. Cada um destes "potenciômetros" controla uma determinada característica da imagem.

A Fig. 5 representa o resultado da aplicação desta ferramenta na melhoria do brilho e do contraste da imagem representada na Fig. 4.

B. Visualização do histograma.

A visualização do histograma de uma imagem é uma das ferramentas básicas de qualquer sistema de processamento digital de imagem. A ferramenta implementada permite visualizar o histograma de imagens de 8 bits/pixel, *gray scale* ou a cores. Como complemento, permite duas escalas: linear e logarítmica. Permite também a visualização do número de *pixels* de determinada entrada da *palette* e as suas componentes RGB. A Fig. 6 ilustra o aspecto da janela de visualização do histograma.

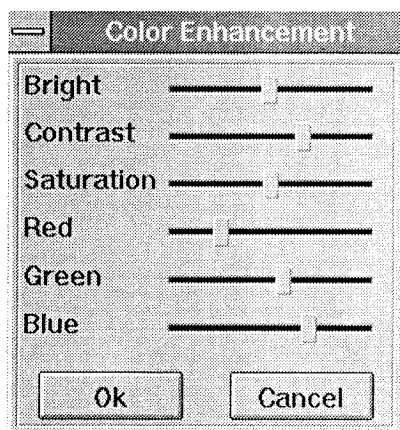


Fig. 3 - Janela da melhoria de imagem.

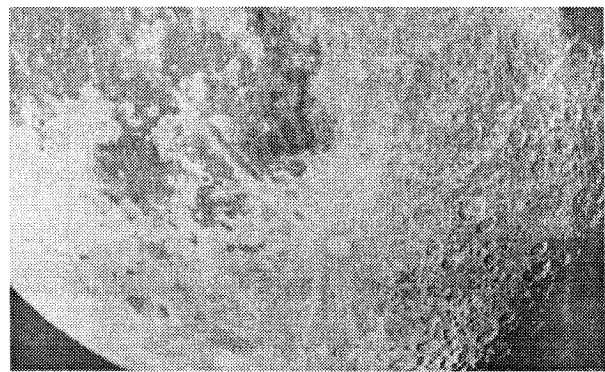


Fig. 4 - Imagem original.

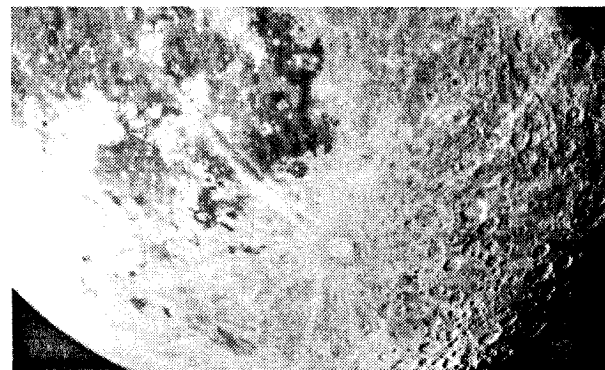


Fig. 5 - Imagem processada com melhoria de contraste.

C. Operações sobre o histograma.

Esta ferramenta permite a definição da curva de distribuição do histograma. Esta técnica permite melhorar o contraste da imagem, definindo as zonas de concentração do histograma. Para obter a lei de distribuição pretendida o utilizador pode optar por uma das quatro leis pré-definidas. Se desejar uma lei de distribuição não contemplada, ou retocar uma das existentes, poderá editá-la, desenhando a curva com o rato. A Fig. 7 ilustra o aspecto da janela de interacção.

D. Segmentação por Threshold.

A segmentação por *Threshold* é utilizada em imagens em *gray scale*, e permite redefinir o nível de cinzento dos *pixels* dentro de uma determinada gama de níveis de cinzento. A cor dos *pixels* pode ser definida pelo

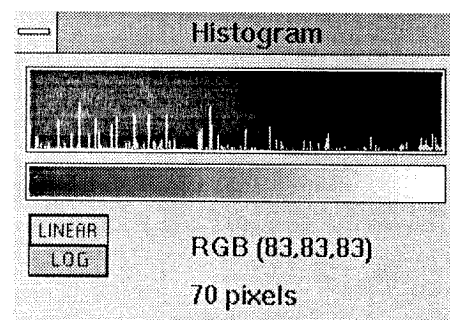


Fig. 6 - Janela de visualização do histograma.

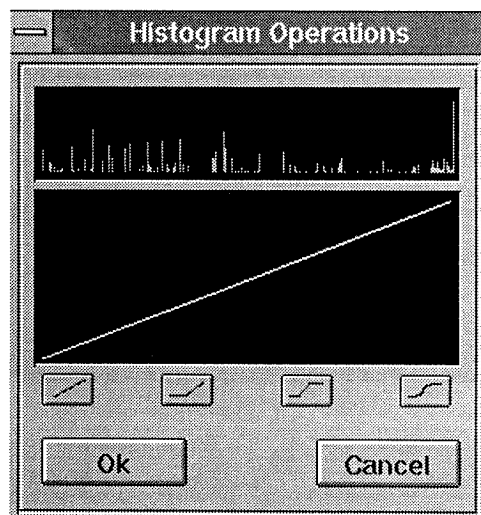


Fig. 7 - Janela das operações com o histograma.

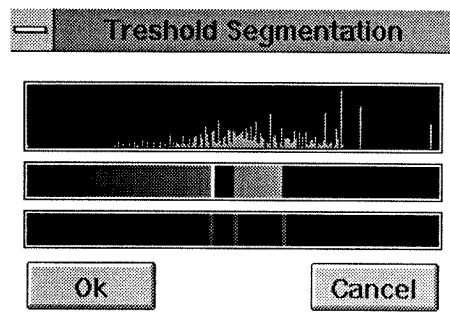
utilizador ou pode ser a escala normal (*gray scale*). Podem ser criadas até vinte gamas de segmentação diferentes.

As alterações definidas pelo utilizador (que podem ser: criação de novas gamas, movimentação das fronteiras das gamas, movimentação de toda uma gama, alteração dos níveis de cinzento dentro de cada gama, etc) reflectem-se na imagem a segmentar em tempo real, devido à utilização de animação dinâmica da *palette*. A Fig. 8 ilustra a janela de interacção com o utilizador.

E. Segmentação por crescimento de regiões.

A segmentação por crescimento de regiões agrega um conjunto de *pixels* em torno de um ponto semente com características tonais próximas. Pode-se assim utilizar esta ferramenta para isolar objectos de uma imagem, desde que estes tenham características tonais diferentes do que os rodeia.

Todo o processo começa no ponto semente. São analisados os pontos vizinhos deste, sendo incluídos ou não na imagem final, de acordo com o critério de proximidade tonal ao ponto semente. O critério de proximidade tonal baseia-se na comparação das componentes RGB do ponto a comparar com as do ponto semente. Pode ser tolerável uma variação percentual relativamente a cada componente, ou então, impôr explicitamente os limites de variação para cada

Fig. 8 - Janela de segmentação por *threshold*.

	R	G	B	%
Upper	255	255	255	51
Lower	114	95	49	

Fig. 9 - Janela de Segmentação por Crescimento de Regiões.

componente.

A Fig. 9 representa a janela de interacção com o utilizador. O número de *pixels* segmentados é visualizado, permitindo assim o cálculo de áreas relativas de objectos irregulares.

As Figuras 10 e 11 ilustram uma aplicação desta ferramenta, neste caso, a segmentação de objectos.

F. Transformadas de imagem.

Uma transformada de imagem é uma representação dessa imagem no domínio da frequência [3]. O cálculo de uma transformada é um processo lento, pelo que se geralmente se utilizam as *Fast Transforms*. A sua utilização, no entanto, não é genérica, requerendo uma área de imagem quadrada, cujo lado seja uma potência de 2. Foram

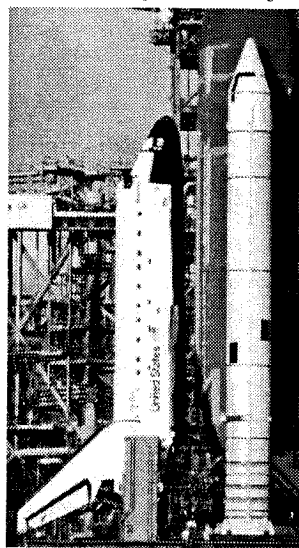


Fig. 10 - Imagem original.



Fig. 11 - Imagem segmentada.

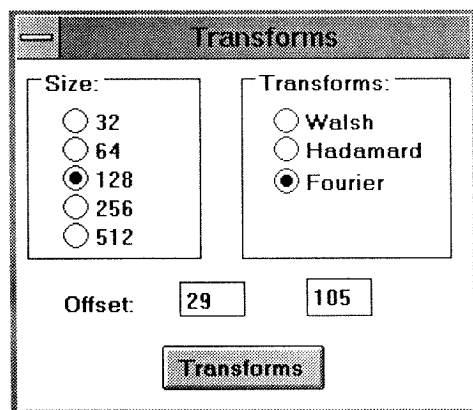


Fig. 12 - Janela das transformadas.

implementados os tamanhos de 32 a 512 pixels de lado do quadrado. As transformadas disponíveis são Transformada de Walsh, Hadamard e Fourier. A janela de interacção com o utilizador está ilustrada na Fig. 12.

A visualização da transformada calculada está a cargo duma classe de janelas (Fig. 13) que permite a visualização do módulo dos coeficientes, do módulo logarítmico, da fase e da transformada inversa. Permite também a execução de filtragem na frequência, através do desenho directo do filtro a aplicar.

Existe a possibilidade de gravação em ficheiro da transformada bem como um posterior *Load* desse ficheiro.

G. Filtragem Espacial.

A filtragem espacial é um tipo de filtragem que se baseia em operações no domínio espacial. Cada pixel da imagem filtrada é função, não só do pixel correspondente na imagem original, mas também dos seus vizinhos.

A lei de transformação, e portanto a topologia do filtro, é estabelecida na máscara de filtragem [1], que define a maneira como vão contribuir os pixels abrangidos no cálculo do pixel filtrado.

Os filtros podem conter uma ou várias máscaras. No caso de várias máscaras é feita uma filtragem aditiva, em que o nível de cinzento de cada pixel é o resultado da soma das várias filtrações (com as várias máscaras).

Devido à grande quantidade de cálculos necessários,

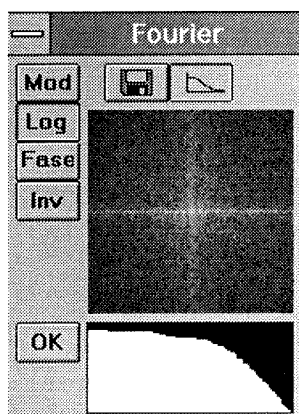


Fig. 13 - Janela de visualização das transformadas.

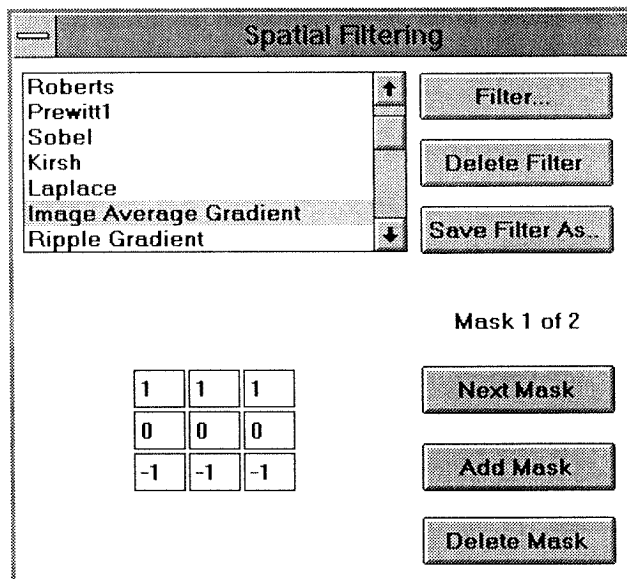


Fig. 14 - Janela da filtragem espacial.

principalmente quando os filtros são bastante elaborados, decidimos implementar as funções de baixo nível em *assembler*, tirando partido das instruções de 32 bits. De facto, os compiladores usuais geram apenas código de 16 bits, não aproveitando assim as capacidades dos processadores 386 e superiores (486, *pentium*, ...).

O código *assembler* de 32 bits, em conjunção com um algoritmo optimizado, permitiram alcançar tempos de filtragem muito curtos. A título de curiosidade, a velocidade de filtragem obtida numa máquina 486 a 33 MHz foi de 146000 pixels/s por elemento de máscara.

A Fig. 14 representa a janela de interacção com o utilizador. A Fig. 16 representa a imagem filtrada da Fig. 15, com o filtro "Image Average Gradient".

H. Reconhecimento automático de caracteres.

O reconhecimento automático de caracteres identifica o texto contido numa imagem através da comparação com

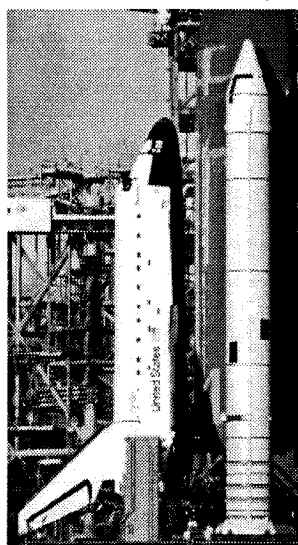


Fig. 15 - Imagem original.

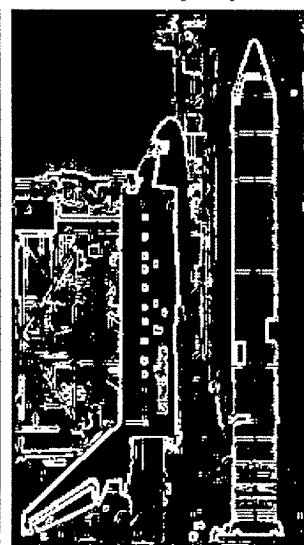


Fig. 16 - Imagem filtrada.

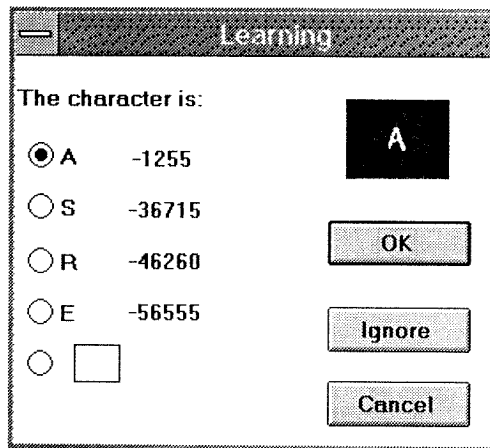


Fig. 17 - Janela da aprendizagem de caracteres.

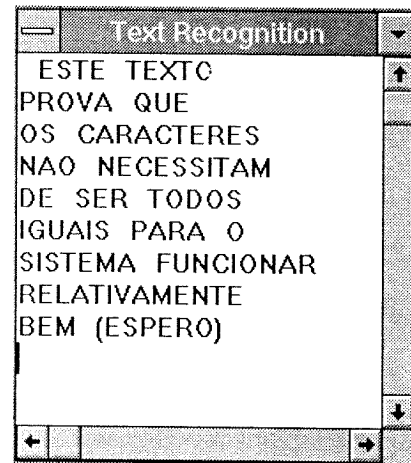


Fig. 19 - Texto reconhecido pelo programa.

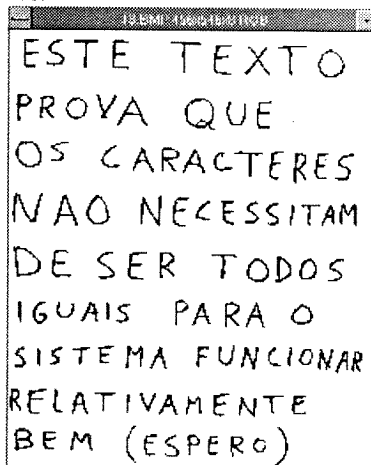
padrões de caracteres previamente carregados em memória. A imagem pode ter sido obtida, por exemplo, através de um *scanner*.

O sistema requer uma aprendizagem prévia para definir os padrões dos caracteres. A aprendizagem é feita interactivamente, ou seja, o programa tenta reconhecer cada caracter e o utilizador deve corrigir ou confirmar a hipótese calculada pelo programa. Obviamente, quanto mais apurada for a fase de aprendizagem, maiores serão as probabilidades de sucesso no reconhecimento de um determinado texto. A estrutura do padrão de cada caracter consiste numa imagem normalizada resultante de uma média pesada das imagens normalizadas do caracter, possivelmente escrito de várias formas diferentes. A formação dos padrões é, assim, um processo de formação de uma imagem padrão da cada caracter.

É possível gravar os padrões para uso futuro e também continuar a aprendizagem em padrões previamente gravados.

Para iniciar o processo de reconhecimento é então necessário ter carregados em memória padrões que devem ser mais ou menos compatíveis com o tipo de caracteres do texto a reconhecer.

O texto, depois de reconhecido, é visualizado numa janela de texto onde pode ser editado, para correcção de eventuais erros.

Fig. 18 - Imagem adquirida com *scanner*.

A Fig. 17 ilustra a janela de interacção com o utilizador na fase de aprendizagem. A Fig. 18 representa uma imagem de um texto obtida através de um *scanner*. A Fig. 19 mostra esse texto depois de reconhecido pelo programa.

V - CONCLUSÕES.

Foi implementada a estrutura base de um sistema genérico de processamento digital de imagem, bem como algumas ferramentas. Os formatos de imagem reconhecidos pelo programa são o BMP, GIF [4] e IMG. A estrutura modular do sistema permite a fácil introdução de novos formatos de imagem, bem como novas ferramentas de processamento e/ou controle. A interacção com o utilizador é fácil e intuitiva. Um utilizador com experiência no ambiente WINDOWS™ não necessita de aprendizagem do sistema, pois são seguidos os princípios comuns à maior parte das aplicações desse ambiente.

Foram utilizados algoritmos e técnicas, que permitiram obter performances assinaláveis, nomeadamente no caso da filtragem espacial. A utilização de animação dinâmica da palette em algumas ferramentas, proporciona uma "afinação" da imagem em tempo real, ao contrário da técnica de *preview* largamente utilizada nas aplicações comerciais mais conhecidas.

VI - REFERÊNCIAS.

- [1] Jorge Saro, Sandra Jardim, "IDIPP 2.0 Manual do Utilizador".
- [2] Microsoft Press, "WINDOWS Programmer's Reference Version 3".
- [3] A. Sousa Pereira, "Apontamentos de Processamento Digital de Imagem".
- [4] Compuserve, "Graphics Interchange Format".