

REGRAF: Um Sistema para o Reconhecimento de Documentos Gráficos

Luiz Eduardo Seabra Varella ^{1,2}
Luis Paulo Vicira Braga ³

¹ PETROBRÁS-Petróleo Brasileiro S.A.
Depex-Departamento de Exploração
Av. Chile 65, sala 1402
20035 Rio de Janeiro, RJ, Brasil.
g073@c53000.petrobras.anrj.br

² IME-Instituto Militar de Engenharia
Seção de Sistemas e Computação/Cartografia
Praça General Tibúrcio, 80
22290 Rio de Janeiro, RJ, Brasil.
pass@lncc.bitnet

³ IM/UFRJ-Instituto de Matemática
Caixa Postal 68530
21945 Rio de Janeiro, RJ, Brasil.
mad02005@ufrj.bitnet

Abstract: This paper presents a system designed to make semi-automatic recognition of graphic documents digitized by Scanners. The following steps are implemented in the system: image pre-processing and thinning; raster to vector conversion; separation of graphic elements into classes, geometric (lines, curves) and non-geometric (texts, symbols); recognition of the geometric and non-geometric elements. The most important data transformations and the algorithms for the semi-automatic tasks are described, becoming the main subject of this communication.

1. Introdução

Um grande número de organizações, tais como o *Environment Canada*, o *US National Parks Service* e o *US Forest Service*, utilizam *Scanners* como dispositivo principal para aquisição de grandes volumes de dados espaciais. Estima-se que este dispositivo reduza em aproximadamente dez vezes o tempo dispendido para realização desta tarefa, [Aronoff(1989)], se comparada com a digitalização feita através de Mesas Digitalizadoras.

No presente trabalho é descrito um sistema computacional para o reconhecimento de informações gráficas digitalizadas via *Scanners* (REGRAF), que combina soluções automáticas com semi-automáticas. O texto desta comunicação foi redigido de maneira a enfatizar, predominantemente, as etapas semi-automáticas do reconhecimento. Para maiores detalhes acerca dos demais procedimentos ver [Varella (1992 a,b,c)].

2. Arquitetura Básica do REGRAF

O REGRAF foi implementado com seis módulos: Pré-Processamento, Afinamento, Codificação,

Classificação, Reconhecimento Geométrico e Reconhecimento não Geométrico. O Pré-Processamento consiste na eliminação dos ruídos da imagem. O Afinamento tem por objetivo gerar o esqueleto da imagem. A Codificação executa a conversão raster-vector, representando a imagem por códigos de direção. A Classificação separa os elementos gráficos em duas categorias: geométricos (retas, poligonais, curvas e curvas de isovalores) e não geométricos (caracteres e símbolos). O Reconhecimento Geométrico é responsável pela criação dos elementos gráficos do tipo curvas de isovalores, retas e poligonais. O Reconhecimento não Geométrico é definido por um método de reconhecimento de caracteres e símbolos.

Do ponto de vista de transformação de dados, o REGRAF utiliza 3 conjuntos de arquivos (Figura 1): RLE, COD e ELE. A imagem binária é inicialmente armazenada no arquivo RLE (Run Length Encode), para em seguida ser convertida em uma representação, ainda matricial, mas compactada, de nome COD (Códigos de Direção). Sobre o arquivo COD as informações são reconhecidas e armazenadas em ELE (Elementos),

com a descrição dos elementos gráficos próximos a um Sistema de Computação Gráfica.

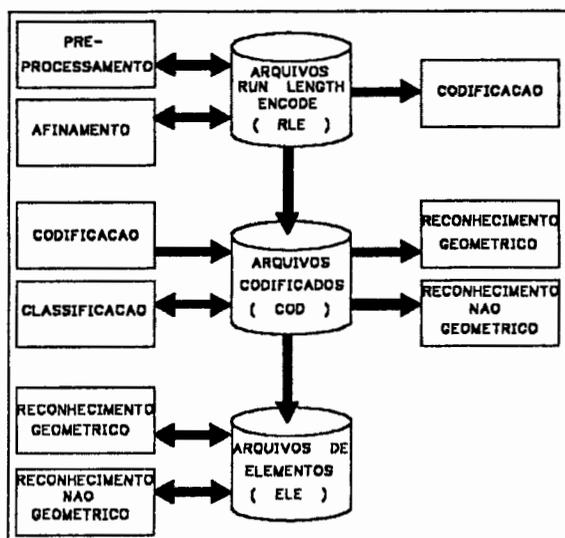


Figura 1 - Arquivos do Sistema.

3. Descrição dos Módulos

Com exceção do Reconhecimento não Geométrico, os módulos do REGRAF são aqui descritos, dando-se ênfase aos procedimentos semi-automáticos.

3.1 Pré-Processamento

O Pré-Processamento visa remover os ruídos e realçar a imagem de entrada. O módulo dispõe de duas funções básicas: Manipulação e Filtragem. A Manipulação se constitui de um editor de imagens. A Filtragem é composta de algoritmos para Remoção de Ruídos ; Preenchimento de Vazios; Remoção de Pixels Isolados; Suavização Lógica Direcional e Suavização Lógica pela Maioria com Teste de Contorno.

3.2 Afinamento

O Afinamento é responsável pela geração do esqueleto da imagem. Após sua execução, os elementos gráficos passam a ter espessuras unitárias localizadas sobre os eixos médio. Foi adotado o método clássico de múltiplas varreduras.

3.3 Codificação

A codificação foi decomposta em três processos: Vetorização, Edição e Segmentação. A Vetorização

é responsável pela descrição da imagem segundo códigos de direção (conversão raster-vector). A Edição destina-se à manipulação dos elementos topológicos da imagem (caminhos, nós e componentes conectados). A Segmentação separa os temas da imagem que se encontram, do ponto de vista lógico, conectados erradamente. Em um mapa de isolinhas, um caso bastante típico, que degrada a conectividade dos elementos, é a malha do sistema de coordenadas. A solução para este problema encontra-se em detalhe, a seguir.

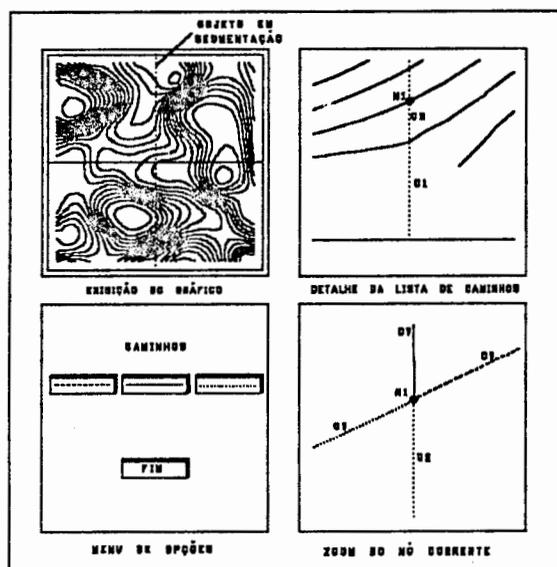


Figura 2 - Interface da Segmentação.

3.3.1 Segmentação

O problema da separação de temas pode ser formulado como: dado um componente conectado, que possui pelo menos dois elementos logicamente distintos, deseja-se transformar este componente em vários outros. A separação é feita da seguinte maneira: a-identificação do componente a ser separado; b-obtenção da lista de caminhos conexos que definem um único elemento; c-modelagem deste novo elemento; d-retirada do elemento no componente de origem; e-recodificação do componente de origem. A maior parte dos passos do algoritmo se constitui em interações com operador, sendo a aquisição da lista de caminhos seu ponto crítico. A solução, portanto, recai na definição de uma rotina de busca a esses caminhos.

Foram projetados dois tipos de buscas: Busca Incondicionada e Busca com Regressão Linear. No primeiro caso se define um caminho inicial. Em

seguida, é apresentada uma saída gráfica contendo quatro janelas distintas (Figura 2). Nos cantos superiores esquerdo e direito, são exibidos o desenho total e um detalhe da lista de caminhos até então definidos. No canto inferior direito é mostrado, em *zoom*, o nó associado ao último caminho que, portanto, deve ser resolvido. Os caminhos que partem deste nó são desenhados com atributos diferentes. Na última janela, canto inferior esquerdo, encontra-se o menu de opções, onde é feita a escolha do caminho a ser trilhado.

O segundo tipo, a Busca com Regressão Linear, explora a forma retilínea dos elementos gráficos. Esta busca atua como um extrator de propriedades geométricas dos caminhos, e foi assim implementado: Para um dado conjunto de pontos que descrevem um caminho da imagem, calcula-se o coeficiente de correlação, $Corr(X,Y)$. Se $-0.7 \leq Corr(X,Y) \leq 0.7$, o caminho é dito não retilíneo, caso contrário, ele é retilíneo. A escolha do caminho a ser incorporado à lista é tomada da seguinte maneira: Dado um nó, classificam-se todos os caminhos a ele associados. Os não retilíneos são descartados, enquanto os retilíneos têm suas inclinações calculadas via método dos mínimos quadrados. O caminho a ser incorporado à lista é aquele cuja inclinação mais se aproxima do caminho imediatamente anterior. Se todos os caminhos foram classificados como não retilíneos, a decisão é tomada pelo operador, com a utilização de uma interface idêntica a do método anterior.

3.4 Classificação

Grande parte dos gráficos possui textos e símbolos com diversas formas, tamanhos e inclinações. Adotou-se para a classificação dos elementos um método que combina atributos de área, densidade de pixels acesos e razão altura/largura dos componentes conectados da imagem. Em função de sua natureza robusta, alguns elementos são classificados erradamente. A fim de se obter um maior controle, é aconselhável se incorporar funções que permitam a manipulação interativa dos limiares dos atributos e que reclassifiquem, individualmente, cada elemento.

A classificação, então, deve obedecer a uma seqüência de passos, sendo iniciada com uma inspeção visual do gráfico, a fim de se eliminar

pequenos componentes da imagem. Em seguida, aplica-se a separação com os limiares definidos acima e observa-se o resultado. Se um grande número de textos forem classificados de maneira errada, deve-se solicitar a manipulação dos limiares.

A Figura 3 mostra a interface implementada, composta de quatro janelas, três para exibição gráfica e uma para informar e/ou obter os limiares. No canto superior esquerdo é apresentado todo o desenho. Ao seu lado, encontram-se os elementos geométricos, e abaixo destes são exibidos os elementos não geométricos. A interação com o operador é feita modificando-se os limiares.

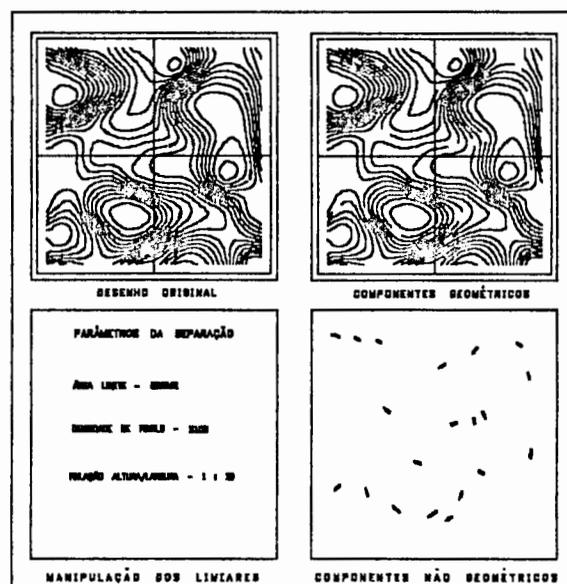


Figura 3 - Interface da Classificação.

3.5 Reconhecimento Geométrico

O Reconhecimento Geométrico foi implementado com dois processos: Geração e Manipulação. A Geração inicia a vetorização definitiva da imagem, convertendo cada componente conectado em um único elemento gráfico, restringindo-se, apenas, a reduzir o número de pontos dos componentes. A Manipulação completa o reconhecimento, identificando os diferentes tipos de primitivas. Este último processo é descrito a seguir.

3.5.1 Manipulação

A manipulação reconstrói formas, reconhece primitivas e organiza hierarquicamente as informações armazenadas nos arquivos ELE. O processo foi implementado com três funções:

Ajustamento Geométrico, Reconhecimento de Curvas de Isovalores e Edição Geométrica.

A primeira, o Ajustamento, visa a reconstruir feições com formas bem definidas. Para se corrigir estes elementos, foram implementados interpoladores de retas, retângulos, círculos, arcos e elipses, via método dos mínimos quadrados.

Quanto às primitivas possíveis, foram previstas três tipos: poligonais; curvas e curvas de isovalores. Como o trabalho encontra-se, de certa maneira, envolvido com um documento altimétrico, apenas o reconhecimento de curvas de isovalores foi implementado. Para tal desenvolveram-se dois métodos de identificação: Busca Individual e via Definição de Feixes. No primeiro, a Busca Individual, o operador posiciona o dispositivo de interação sobre a poligonal. No segundo, a identificação via Definição de Feixes, explora a regularidade de variação numérica dos níveis. A identificação consiste na definição de um feixe (uma reta) que intercepta um grande número de poligonais que devem ser transformadas em primitivas do tipo curvas de nível. Definido o feixe, o Sistema responde realçando as poligonais que interceptam a reta, e solicita a entrada do nível inicial, o intervalo de variação dos níveis e a especificação do acidente geográfico.

A Figura 4 mostra a interface do reconhecimento de curvas via feixes. Definido o feixe, o Sistema apresenta um *display* com a exibição integral do gráfico no canto superior esquerdo da tela, um detalhe ampliado da região do feixe no canto superior direito, uma ampliação do primeiro ponto do feixe no canto inferior direito e um menu de opções no canto inferior esquerdo. Sobre o menu, o operador define as entradas numéricas, solicita alterações na superfície de exibição (*zoom* ou panorâmica da vista mostrada no canto superior esquerdo), especifica o acidente geográfico (aclive ou declive) e informa se o reconhecimento é automático ou nível a nível. No reconhecimento automático todos os níveis são batizados segundo a entrada numérica, sem que o operador acompanhe os valores atribuídos.

Finalizando, a Edição Geométrica organiza as informações segundo a estrutura de dados dos arquivos ELE, oferecendo um conjunto de ferramentas que permitem alterações, quer na localização espacial das primitivas, quer na especificação de atributos de visualização gráfica.

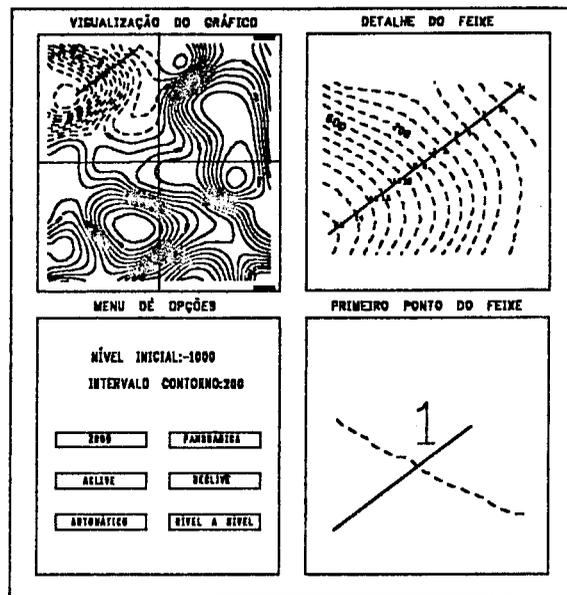


Figura 4 - Reconhecimento de Curvas.

4. Conclusões

Um Sistema para o reconhecimento de documentos gráficos foi apresentado. Nele utilizaram-se técnicas automáticas (que não requerem a intervenção do usuário) para as etapas de filtragem da imagem, afinamento e conversão raster-vector. Para o reconhecimento dos elementos gráficos, que é iniciado com a classificação, foram sugeridos métodos que exploraram a interface com usuário.

O Sistema apresentado encontra-se em fase de implementação e testes, mostrando, nesta etapa, ganhos significativos de tempo e precisão.

5. Referências

- S. Aronoff, Geographic Information Systems: a Management Perspective, WDL Publications, Canada, 1989.
- L.E.S. Varella, Reconhecedor de Elementos Gráficos Digitalizados via Scanners, Dissertação de Mestrado, IME, RJ, 1992 a.
- L.E.S. Varella & L.P.V. Braga, Reconhecimento de Mapas de Isovalores Digitalizados via Scanners, SIBGRAPI'92, São Paulo, SP, Nov. 1992, pp. 203-212, 1992 b.
- L.E.S. Varella, E.P.L. Passos, M.A. Santos & R.L. Araujo, Reconhecimento de Caracteres com Variações de Escala, Rotação e Translação, SIBGRAPI'92, São Paulo, SP, Nov. 1992, pp. 301-310, 1992 c.