

Algoritmos de generalização vetorial em SIG

Mariane Alves Dal Santo¹

Carlos Loch²

Universidade Federal de Santa Catarina
Programa de Pós Graduação em Engenharia Civil

88040-970 – Florianópolis - SC

¹marianedalsanto@udesc.com

²loch@ecv.ufsc.br

RESUMO : Os Sistemas de Informação Geográfica possuem ferramentas para a generalização cartográfica que se baseiam em transformações espaciais que alteram a representação geométrica e topológica dos dados. Como a maioria dos objetos espaciais utiliza o vetor como entidade básica para sua representação, a generalização vetorial tem sido bastante estudada e é a transformação comumente encontrada nos sistemas existentes. Nestes sistemas, os três elementos básicos da representação vetorial, pontos, linhas e áreas, são traduzidos em objetos geográficos. Neste artigo trataremos dos métodos de simplificação de linhas. Estes métodos procuram selecionar e manter os pontos que melhor caracterizam a representação digital de uma linha. Desta forma, este artigo relata os experimentos e a avaliação de dois algoritmos de simplificação de linhas aplicados em SIG: uma adaptação do algoritmo de Douglas-Peucker que usa o quociente área/perímetro, e uma adaptação do algoritmo de Wang, que aplica técnicas de reconhecimento de arcos.

PALAVRAS CHAVES: Sistemas de Informação Geográfica, Generalização Vetorial, precisão cartográfica

ABSTRACT : The Geographic Information Systems have tools to product a cartographical generalization that are based in spatial transformations that alter the data geometrical and topological representation. The major spatial objects use vector as the basic entities for representation, the vectorial generalization has been well studied and is the most common transformation used in the existent systems. The three basic elements for vectorial representation are points, lines and areas and are translated to geographical objects. This article discusses polygonal lines simplifying methods and this paper explains the experiments and evaluation process for two algorithms to improve line simplification applied to GIS: an adaptation of Douglas-Peucker algorithm, an area/perimeter cocient, and an adaptation of Wang, algorithm applying arcs technical recognition.

KEYWORDS: Geographical Information Systems, Vetorial Generalization, cartographical precision

1 Introdução

Objetos e fenômenos espaciais podem ser representados em mapas de diferentes escalas, obedecendo aos objetivos e propósitos da sua aplicação. Atualmente, com o crescimento na utilização de sistemas espaciais georeferenciados, em meio digital, tem-se gerado modelos a partir de métodos para o tratamento dos dados.

Tradicionalmente, em cartografia tem-se utilizado a generalização cartográfica para a elaboração dos processos de representação simplificada de escalas. Através destes processos são estabelecidas regras a fim de preservar a geometria e as características topológicas dos objetos espaciais.

Segundo Tang e Adams(1996), objetos que representam o espaço geográfico, contem em si, quatro elementos de representação: geometria, propriedades não espaciais, relações topológicas e relações não

topológicas. Estes objetos podem ser transformados, através de métodos e processos de generalização.

A partir do momento que são generalizados os objetos geográficos passam por transformações na sua geometria e a partir destas podem ocorrer mudanças nas relações topológicas ou não topológicas entre os elementos. Porém, segundo McMaster e Shea (1992), o processo de generalização deve procurar preservar a consistência dos relacionamentos topológicos.

Paralelamente, através do crescente incremento operacional implementado a partir da informatização de dados geográficos em sistemas espacializados, a Cartografia Cadastral tem se beneficiado com estas aplicações, uma vez que necessita de mapeamentos em escalas variadas que atendam as necessidades do Cadastro Técnico Multifinalitário e da gestão territorial. Dentro das perspectivas da escala cadastral, várias transformações devem ser efetuadas para se obter um produto cartográfico de qualidade, e que esteja em consonância com os princípios da generalização cartográfica convencional.

2 SISTEMAS DE INFORMAÇÃO GEOGRÁFICA

Os Sistemas de informação Geográfica (SIG), acompanhando a recente informatização das prefeituras, instituem-se como uma ferramenta fundamental para atualização cadastral.

Os SIGs atuais apresentam funções especializadas de generalização, que vem a auxiliar o trabalho do cartógrafo e do usuário final do sistema. No desenvolvimento deste artigo apresentamos as funcionalidades de um SIG, na geração automatizada de simplificação de linhas.

Neste artigo trataremos da suavização de linhas como uma maneira útil de se construir uma base de dados nova através de operadores de transformação e generalização. Desta forma os dados podem ser adequados a uma representação em escala menor. Isto significa que a generalização deriva dados novos com a transformação das propriedades espaciais e não espaciais de um atributo.

2.1 Estrutura de dados vetoriais em SIG

Segundo Câmara (2007) as estruturas de dados vetoriais são utilizadas para representar os dados espaciais através das coordenadas geográficas de cada entidade, tendo como formas básicas: pontos, linhas e áreas (polígonos), definidas por suas coordenadas cartesianas, conforme mostrado na Figura 1.

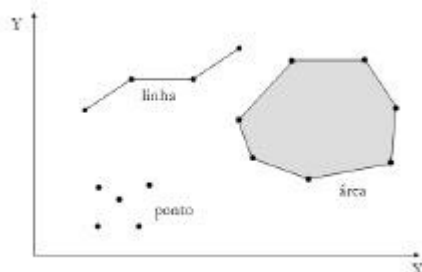


Figura 1 – Representação de estruturas vetoriais
Fonte: Câmara (2007)

Onde:

- Um ponto é um par ordenado (x,y) de coordenadas espaciais;
- Uma linha é um conjunto de pontos conectados;
- Um área (ou polígono) é a região do plano limitada por uma ou mais linhas poligonais, de tal forma que o último ponto da linha seja idêntico e coincida ao primeiro da próxima.

Cada estrutura vetorial possui uma topologia que pode ser descrita sistematicamente da seguinte forma:
Topologia: nó, nó + nó = arco e arco+nó = polígono.

A topologia é a parte da matemática na qual se investigam as propriedades das configurações que permanecem invariantes nas transformações de rotação, translação e escala. No caso de dados geográficos, é útil ser capaz de determinar relações como adjacência ("vizinho de"), pertinência ("vizinho de"), intersecção, e cruzamento. Objetos de área podem ter duas formas diferentes de utilização: como objetos isolados ou objetos adjacentes. O caso de objetos isolados (DAVIS, 2000) é bastante comum em

SIGs urbanos, e ocorre no caso em que os objetos da mesma classe em geral não se tocam. Por exemplo, edificações, piscinas, e mesmo as quadras das aplicações cadastrais ocorrem isoladamente, não existindo segmentos poligonais compartilhados entre os objetos. Finalmente, temos objetos adjacentes, e os exemplos típicos são todas as modalidades de divisão territorial: bairros, setores censitários, redes, municípios e outros. Neste caso, pode-se ter o compartilhamento de fronteiras entre objetos adjacentes, gerando a necessidade por estruturas topológicas.

Objetos de linha podem ter variadas formas de utilização. Analogamente aos objetos de área, podemos ter objetos de linha isolados, em árvore e em rede. Objetos de linha isolados ocorrem, por exemplo, na representação de muros e cercas em mapas urbanos. Objetos de linha organizados em uma árvore podem ser encontrados nas representações de rios e seus afluentes, e também em redes de esgotos e drenagem pluvial. E podem ser organizados em rede, nos casos de redes elétricas, telefônicas, de água ou mesmo na malha viária urbana e nas malhas rodoviária e ferroviária. No caso das redes, é fundamental armazenar explicitamente as relações de adjacência, utilizamos a topologia arco-nó. Um nó pode ser definido como o ponto de intersecção entre duas ou mais linhas, correspondente ao ponto inicial ou final de cada linha. Nenhuma linha poderá estar desconectada das demais para que a topologia da rede possa ficar totalmente definida.

Uma das possibilidades associadas a dados vetoriais é a associação de valores que denotem a variação espacial de uma grandeza numérica (2,5 D). No caso mais simples, associamos a cada localização no espaço um valor numérico de atributo. Neste caso, como os valores de localização estão no plano e o valor adicional descreve uma superfície sobre este plano. Os dados resultantes são chamados de dimensão “dois e meio”, pois não se tratam estritamente de dados tridimensionais, pois o suporte espacial ainda são localizações 2D. Temos três alternativas que usam estruturas vetoriais:

- Conjunto de amostras esparsas 2,5D, constituído de pares ordenados (x,y,z) , onde (x,y) é uma localização no plano e z um valor numérico de atributo;
- Conjunto de isolinhas (curvas de nível), que são linhas às quais estão associados valores numéricos. As isolinhas não se cruzam, e são entendidas como estando “empilhadas” umas sobre as outras;
- A malha triangular ou TIN (do inglês “triangular irregular network”) é uma estrutura do tipo vetorial com topologia do tipo nó-arco e representa uma superfície através de um conjunto de faces triangulares interligadas.

A malha triangular é a estrutura vetorial mais utilizada para armazenar dados 2,5 D. Cada um dos três vértices da face do triângulo armazenados as coordenadas de localização (x, y) e o atributo z , com o valor de elevação ou altitude. Em geral, nos SIGs que possuem pacotes para MNT, os algoritmos para geração da malha triangular baseiam-se na triangulação de Delaunay com restrição de região. Quanto mais equiláteras forem as faces triangulares, maior a exatidão com que se descreve a superfície. O valor de elevação em qualquer ponto dentro da superfície pode ser estimado a partir das faces triangulares, utilizando-se interpoladores.

2.2 Simplificação de Linhas

A simplificação de linhas é uma redução seletiva do número de pontos (pares de coordenadas) requeridos para representar uma feição. Por exemplo, o número de pontos capturados na digitalização são reduzidos pela seleção de pares de coordenadas considerados mais representativos para que as características sejam preservadas; ou remoção de detalhes desnecessários como curvas e flutuações de uma linha ou um limite de área sem destruir sua forma essencial. Simplificação é o operador mais comumente usado em generalização tendo como principal objetivo atingir a melhor geometria possível. Muitas rotinas de simplificação utilizam critérios geométricos complexos (medidas de ângulo e distancia) para selecionar pontos críticos ou significantes.

O operador de simplificação de linhas é um dos mais utilizados. Segundo D'alge e Goodchild (1996), como a maioria dos objetos utiliza a linha como entidade básica para sua representação, a simplificação de linhas tem sido bastante estudada e é a transformação mais comumente encontrada nos sistemas existentes. O autor coloca que alguns exemplos de algoritmos para simplificação de linhas são: o algoritmo original de Douglas-Peucker de 1992, que usa o quociente área/perímetro e o algoritmo de suavização de Li-Openshaw de 1992, que acumula as distâncias percorridas sobre cada linha. Porém, estes métodos de simplificação atuam sobre as linhas sem preocupação com relações topológicas previamente criadas. Por isso, devem sempre ser sucedidos pelas operações de ajuste de nós e poligonalização.

Douglas-Peucker - Trata-se do método mais utilizado pelos sistemas de informação geográfica. Concebido inicialmente para resolver o problema do número excessivo de pontos resultantes da conversão de dados gráficos para o formato digital, o método de Douglas-Peucker baseia-se na seguinte idéia: se nenhum ponto da linha encontra-se mais afastado do que uma certa distância vertical ao segmento de reta que liga os extremos da linha, então esse segmento de reta é suficiente para representar a linha. Este método é considerado uma técnica global de generalização, pois analisa cada linha como um todo. A Figura 7 ilustra a aplicação do algoritmo de Douglas-Peucker. O primeiro ponto da linha é usado como âncora e o último como ponto flutuante. Este varia até que não haja mais pontos a retirar (A, B, C, D e E). O segundo ponto da linha passa a ser o âncora e o procedimento é reiniciado, continuando até que não haja mais candidatos a âncora (F, G, H e I).

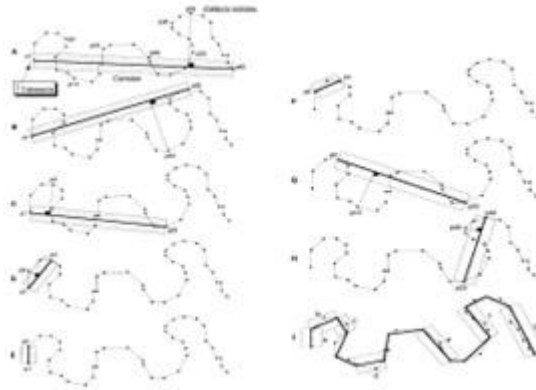


Figura 2 - Algoritmo de Douglas-Peucker
(adaptado de McMaster e Shea, 1992)

Razão Área/Perímetro - Este método utiliza exatamente o mesmo procedimento de análise global de cada linha empregado no método de Douglas-Peucker. A única diferença consiste na adoção da razão área/perímetro calculada em função da tolerância escolhida pelo usuário. O uso da razão área/perímetro permite que triângulos formados por três pontos consecutivos que tenham um ângulo agudo muito pequeno no segundo ponto possam ser detectados de modo mais eficiente que no método de Douglas-Peucker.

Distância Acumulada - O método da distância acumulada é uma adaptação da implementação vetorial do algoritmo de Li-Openshaw que usa como critério o conceito de menor objeto visível. Este método acumula as distâncias a medida em que a linha é percorrida até atingir um certo limiar, removendo todos os pontos acumulados nesse trecho. Trata-se, portanto, de um método bastante simples, mas que, ao contrário dos dois métodos anteriores, não analisa a linha em sua totalidade.

3 Algoritmos de simplificação de linhas em SIG

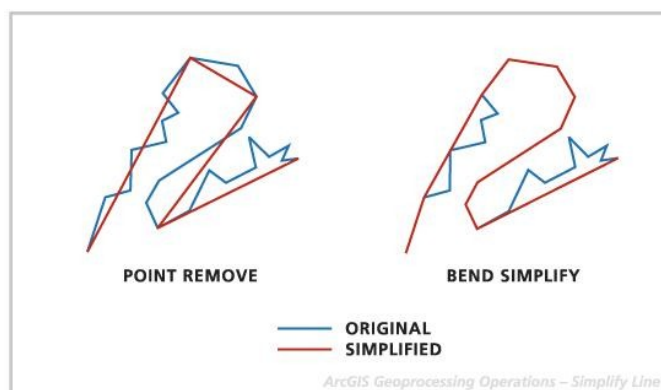
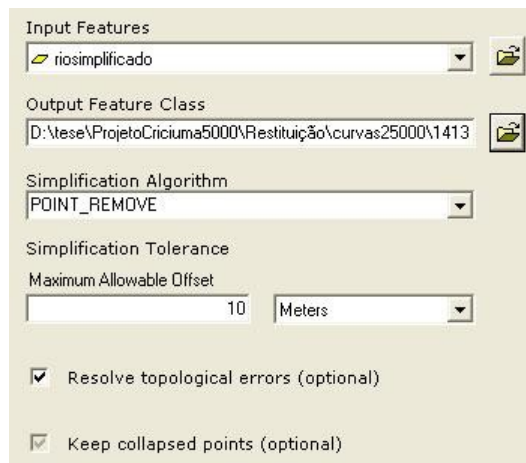


Figura 3 – Algoritmos de simplificação de linhas do SIG ARCGIS

A luz da perspectiva de alcançar uma precisão cartográfica pertinente à escala cadastral, este artigo relata os experimentos e a avaliação de dois algoritmos de simplificação de linhas aplicados no SIG ARCGIS (Figura 3): uma adaptação do algoritmo de Douglas-Peucker – *POINT-REMOVE*, que usa o quociente área/perímetro, e uma adaptação do algoritmo de Wang (Wang, 1999, citado em ESRI, 2007) – *BEND-SIMPLIFY*, que aplica técnicas de reconhecimento de arcos, analisa as suas características e elimina os mais insignificantes. Uma vez que o sistema não é transparente ao usuário, para a aplicação dos algoritmos foram usados critérios subjetivos, que se traduzem em valores de tolerância que foram escolhidos e testados. Desta forma o artigo relata o desenvolvimento, análise e avaliação do impacto das diferentes tolerâncias, em cada método. Outro aspecto avaliado foi a topologia, representada pelo relacionamento espacial entre os objetos, uma vez que estes métodos atuam sobre as linhas sem a preocupação com as relações topológicas previamente existentes entre os dados, como por exemplo proximidade e vizinhança. Desta forma procurou-se chegar a resultados que minimizassem o processo de ajuste de nós e poligonalização.

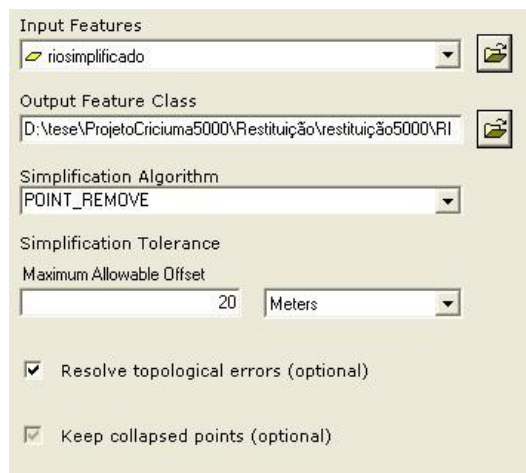
3.1 Aplicação dos Algoritmos

Os algoritmos de simplificação de linhas do software ARCGIS foram aplicados com os seguintes parâmetros, conforme figuras 4, 5, 6 e 7.



The screenshot shows the 'Simplify' tool interface in ArcGIS. The 'Input Features' dropdown is set to 'riosimplificado'. The 'Output Feature Class' is 'D:\tese\ProjetoCriciuma5000\Restituição\curvas25000\1413'. The 'Simplification Algorithm' is 'POINT_REMOVE'. Under 'Simplification Tolerance', the 'Maximum Allowable Offset' is set to '10' and the unit is 'Meters'. Two checkboxes are checked: 'Resolve topological errors (optional)' and 'Keep collapsed points (optional)'.

Figura 4 – Aplicação do Algoritmo *POINT-REMOVE* (Douglas-Peucker) - Parâmetro de 10 m.



The screenshot shows the 'Simplify' tool interface in ArcGIS. The 'Input Features' dropdown is set to 'riosimplificado'. The 'Output Feature Class' is 'D:\tese\ProjetoCriciuma5000\Restituição\restituição5000\RI'. The 'Simplification Algorithm' is 'POINT_REMOVE'. Under 'Simplification Tolerance', the 'Maximum Allowable Offset' is set to '20' and the unit is 'Meters'. Two checkboxes are checked: 'Resolve topological errors (optional)' and 'Keep collapsed points (optional)'.

Figura 5 – Aplicação do Algoritmo *POINT-REMOVE* (Douglas-Peucker) - Parâmetro de 20 m.

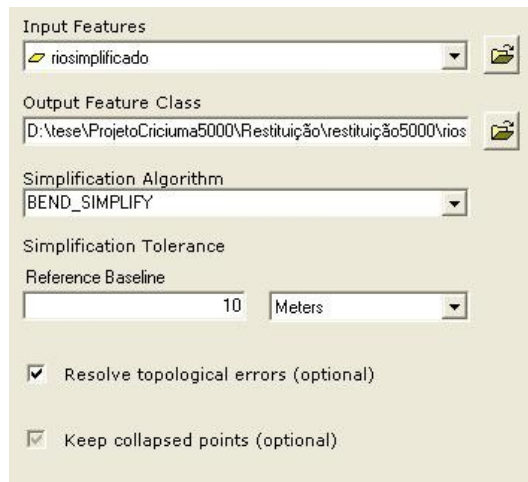


Figura 6 – Aplicação do Algoritmo *BEND-SIMPLIFY* (WANG) - Parâmetro de 10 m.

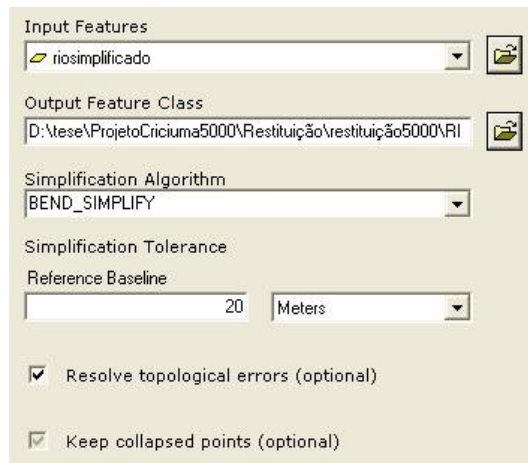


Figura 7 – Aplicação do Algoritmo *BEND-SIMPLIFY* (WANG) - Parâmetro de 20 m.

A seguir apresentamos a linha original (Figura 8) com seu respectivo número de vértices, e o resultado da aplicação dos dois algoritmos para os parâmetros estabelecidos (Figuras 9 e 12).



Figura 8 – Representação da linha original com 263 vértices (pontos)



Figura 9 – Representação da linha simplificada usando o algoritmo *POINT-REMOVE* e parâmetro de 10 metros entre os vértices – Número de vértices resultantes: 16



Figura 10 – Representação da linha simplificada usando o algoritmo *POINT-REMOVE* e parâmetro de 20 metros entre os vértices – Número de vértices resultantes: 9



Figura 11 – Representação da linha simplificada usando o algoritmo *BEND-SIMPLIFY* e parâmetro de 10 metros entre os vértices – Número de vértices resultantes: 171



Figura 12 – Representação da linha simplificada usando o algoritmo *BEND-SIMPLIFY* e parâmetro de 20 metros entre os vértices – Número de vértices resultantes: 147

TABELA 1 – Comparação da Simplificação do linhas

Parâmetro de Simplificação	Algoritmo <i>POINT-REMOVE</i>	Algoritmo <i>BEND-SIMPLIFY</i>
10 metros	16	171
20 metros	9	147

Conforme pode ser observado na tabela 1, de um total de 263 vértices na linha original, com aplicação do algoritmo *POINT-REMOVE*, adaptado de Douglas & Peucker, obteve-se uma diminuição 94% dos vértices quando aplicado o parâmetro de 10 metros entre os vértices e de 97% dos vértices quando aplicado o parâmetro de 20 metros.

Já com a aplicação do algoritmo *BEND-SIMPLIFY*, adaptado de WANG, obteve-se uma diminuição de 35% dos vértices quando aplicado o parâmetro de 10 metros entre os vértices e de 45% dos vértices

quando aplicado o parâmetro de 20 metros.

Com relação as relações topológicas, cruzou-se as linhas simplificadas (Figura 13) como os layer referente as curvas de nível e observou-se que com a aplicação do algoritmo *POINT-REMOVE*, apesar de haver uma maior diminuição de vértices (pontos) da linha, a mesma perde completamente a sua geometria original o que prejudica assuas relações topológicas de vizinhança.

Já com a aplicação do algoritmo *BEND-SIMPLIFY*, não há uma diminuição significava de vértices, mas a linha apresenta uma geometria mais próxima ao original, o que resulta numa melhor relação topológica com os outros elementos do mapa.

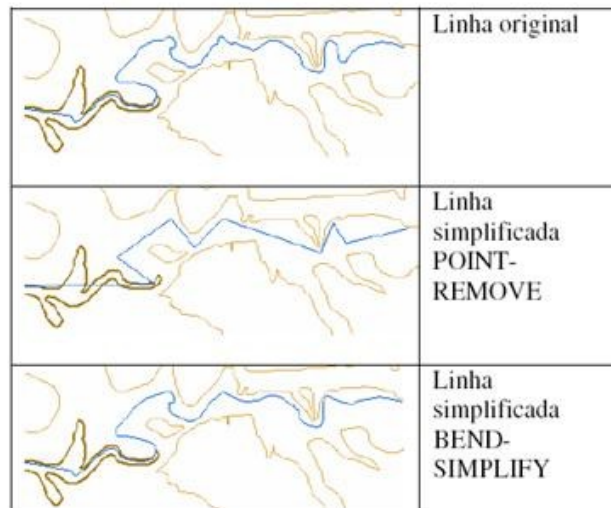


Figura 13 – Relação Topológica entre os elementos

4 Conclusões

Após o experimento, utilizando as mesmas tolerâncias, constatou-se que o algoritmo *POINT-REMOVE*, adaptado de Douglas & Peucker possui a característica de eliminar mais pontos da linha, de forma a transformá-la em um arquivo de dados mais leve e maleável ao usuário, ao mesmo tempo em que perde as características topológicas da linha original. O algoritmo de *BEND-SIMPLIFY*, adaptado de Wang elimina menos pontos, mas mantém uma topologia mais próxima a da linha original . Quando cruzados com os outros elementos topológicos do mapa, para verificar a proximidade e vizinhança, o algoritmo de Wang foi o que apresentou melhor ajuste.

5 Referências Bibliográficas

CÂMARA, Gilberto. Representação computacional de dados geográficos. Em: Banco de Dados Geográficos –Disponível em: <<http://www.dpi.inpe.br/livros/bdados/cap1.pdf>>, acesso: 25 de abril de 2007.

D'ALGE, J. C. L. e GOODCHILD, M. F. Generalização cartográfica, representação do conhecimento e SIG. In Anais VIII Simpósio Brasileiro de Sensoriamento Remoto, Salvador, 1996, p.147-151. INPE-6172-PRE/2261.

DAVIS, C. A. Múltiplas Representações em Sistemas de Informações Geográficas. Tese de Doutorado, Curso de Pós-Graduação em Ciências da Computação, Instituto de Ciências Exatas, Universidade Federal de Minas Gerais, MG, 2000.

ESRI, Map Generalization in GIS: Practical Solutions with Workstation ArcInfo Software, technical paper. Disponível em: <<http://arconline.esri.com>>, acesso: 15 março 2007.

MCMMASTER, R.B., SHEA, K. S. Generalization in Digital Cartography. Resource Publication in Geography. Washington, Association of American Geographers, 1992.

TANG, A.Y., ADAMS T.M. and E.L. Usery, A spatial data model design for feature-based geographical information systems, INT.J.Geographical Information Systems, 1996, Vol.10, No.5, pages 643-659.