

7

Conclusões e trabalhos futuros

7.1.

Conclusões

A maior contribuição deste trabalho é a introdução de técnicas de Aprendizado Competitivo nos problemas de extração e visualização de superfícies de falha. Os algoritmos de Aprendizado Competitivo agem de forma a gerar uma subamostragem adaptativa dos dados volumétricos, transformando um campo escalar com amostragem regular em um grafo. O processo de competição permite que o grafo gerado capture aspectos mais globais das estruturas presentes originalmente nos dados. O problema de extração de regiões de falha foi abordado pelo algoritmo Growing Neural Gas, enquanto a visualização de superfícies de falha foi implementada com o uso do algoritmo Malhas Neurais Abertas.

A aplicação do GNG permitiu gerar um grafo que representa as principais estruturas do atributo de falha. Para tal, foi proposto um mapeamento linear do atributo em probabilidades que desconsidera as regiões de alta coerência (seção 4.6.1) e o uso do GNG com medida de erro local pelo número de vezes que a unidade foi vencedora (seção 4.6.2). Como visto no teste com dados reais (seção 4.6.3), o grafo obtido representa bem as estruturas presentes nos dados de entrada.

Entretanto, o GNG não é capaz de, por si só, gerar subgrafos independentes para cada uma das superfícies de falha. Testes mostraram que o aumento do número de nós permite, em princípio, isolar estruturas, mas o algoritmo não fornece um mecanismo para correlacionar a segmentação com a evolução do algoritmo (seção 4.7).

A solução proposta para esse problema foi aplicar um algoritmo de segmentação do grafo executado em dois passos: no primeiro (seção 5.3), as arestas frágeis são eliminadas, enquanto no segundo (seção 5.4), restrições geométricas próprias das superfícies de falha são impostas. O teste desta estratégia com dados reais gerou uma solução onde as estruturas de falha foram isoladas corretamente.

A segunda maior contribuição deste trabalho foi a criação do algoritmo Malhas Neurais Abertas. O MNA permite gerar uma malha triangular com poucos triângulos a partir de uma função de probabilidades com topologia de uma superfície aberta. O esquema de suavização de malhas tanto na direção normal como na tangencial torna o algoritmo mais robusto diante de dados ruidosos (como é o caso dos dados sísmicos) do que o esquema original do Neural Meshes (seção 6.5). Isso foi verificado nos testes com dados sintético exibidos na seção 6.6. Com uso do MNA, foi possível gerar uma malha triangular leve representando as principais superfícies de falha dos dados sísmicos de teste (seção 6.7). O MNA pode ser utilizado também como um método de reconstrução de superfícies abertas para outras classes de problema (nuvens de pontos, por exemplo).

Outra contribuição deste trabalho foi a proposta de construção de malhas geológicas com uso do GNG (seção 4.8). Nesse caso, o mapeamento do atributo em probabilidades considera as regiões de alta coerência como um pano de fundo, gerando um valor constante de probabilidade, enquanto a medida de erro local é dada pela distância euclidiana entre a unidade vencedora e o sinal de entrada. A estratégia foi testada em dados bidimensionais gerando um bom resultado.

Como pontos negativos das estratégias propostas, podem-se citar o grande número de parâmetros envolvidos e o desempenho computacional: o método completo de extração e visualização compreende 22 parâmetros (ver lista de parâmetros exibida na seção 6.7), enquanto os algoritmos GNG e MNA têm um comportamento $O(n^2)$ (ver seção 4.4), sendo n o número de nós do grafo.

No caso do dado sísmico real utilizado nesta tese (composto por 64x106x99 amostras), todo o processo de extração e visualização de falhas foi executado em 2min e 30s, aproximadamente (40s de GNG, 8s de segmentação do grafo e 1min 38s de MNA). Estes valores de tempo foram obtidos com a utilização de equipamento com processador Pentium 4 de 3.4GHz. Entretanto, o tamanho deste volume é muito pequeno quando comparado com o tamanho dos dados sísmicos utilizados normalmente na indústria de petróleo.

Quanto ao grande número de parâmetros deve ser dito que os valores utilizados para o GNG e o MNA em vários testes bem sucedidos com dados sintéticos e reais correspondem aos mesmos valores típicos exibidos na Tabela 1 da seção 4.3 e na Tabela 5 da seção 6.5.

7.2. Trabalhos futuros

Duas sugestões naturais para trabalhos futuros são a execução de mais testes com dados reais e o ataque dos problemas identificados ao final da seção anterior.

Testes com uma variedade de dados reais permitirão identificar com maior precisão as qualidades e limites das estratégias propostas. Espera-se, por exemplo, verificar se os valores típicos dos parâmetros do GNG e do MNA são realmente de uso geral ou não. Espera-se também que mais testes possibilitem inferir regras ou estratégias para especificar o número de nós a ser gerado pelo GNG e o MNA.

Também é importante verificar o nível de dificuldade de parametrização e a qualidade dos resultados dos dois passos de segmentação do grafo.

Outro ponto relativo a testes seria a avaliação da qualidade dos resultados por intérpretes.

Neste trabalho foi utilizado um mapeamento linear do atributo de falha em probabilidades. Como comentado na seção 4.6.1, esta não é a única forma possível, embora seja a mais simples. Outras formas de mapeamento podem ser testadas.

Pode-se também avaliar o efeito da utilização de outros algoritmos de segmentação de grafos disponíveis na literatura, além do algoritmo FH (seções 5.3.1 e 5.3.2).

O algoritmo Malhas Neurais Abertas pode receber como entrada dados obtidos de forma diferente da que foi descrita nos Capítulos 4 e 5. Uma possibilidade seria gerar um conjunto de pontos através da supressão de não-máximos de Canny conforme apresentada nas seções 3.2.1, 3.4.1.1 e 6.7. De qualquer forma, uma heurística similar à descrita no Capítulo 5 deve ser empregada para separar os pontos em subconjuntos, cada um associado a uma superfície diferente. Cada subconjunto de pontos obtido poderia, então, ser entregue a uma instância diferente do MNA. Na seleção dos sinais de entrada, todos os pontos seriam considerados como possuindo um mesmo valor de probabilidade (isto é, probabilidade uniforme). Estratégia semelhante pode ser empregada para um volume de dados sísmicos classificado segundo o algoritmo de Figueiredo (2007).

Além disso, como comentado anteriormente, o MNA pode ser utilizado também como um método de reconstrução de superfícies abertas para dados de

outras classes de problema, como nuvens de pontos adquiridos por *scanners* 3D, por exemplo.

Quanto ao desempenho computacional, deve-se considerar a avaliação de uma alteração no GNG e no MNA semelhante à que foi proposta por Saleem (2004) para o Neural Meshes que permitiu diminuir o custo computacional para $O(n \log n)$ (seção 6.4). De qualquer forma, mesmo com melhorias desse tipo, o tamanho dos dados sísmicos possivelmente irá demandar uma abordagem através de paralelização de código.

O processo que guiou o desenvolvimento da estratégia de extração e visualização de falhas a partir de um atributo apresentada nesta tese se baseou em três hipóteses. A primeira correspondia à expectativa de que a construção de um grafo representando as estruturas de falha poderia ser empreendida pelo Growing Neural Gas, no lugar do algoritmo de malhas atômicas de Hale. Essa idéia logo se revelou plenamente aceitável conforme os testes foram sendo empreendidos. A segunda hipótese, entretanto, era muito mais nebulosa: esperava-se descobrir uma seqüência de processos que, aplicados sobre esse grafo, seria capaz de gerar uma malha triangulada para as superfícies de falha. Essa idéia se realizou com a criação do algoritmo de Malhas Neurais Abertas e com as heurísticas que permitiram segmentar o grafo.

Por fim, a terceira hipótese previa que, de posse dessa seqüência de processos, estes poderiam ser combinados com o Growing Neural Gas, gerando um único algoritmo de Aprendizado Competitivo capaz de resolver todo o problema. Ou, dito de outra maneira, o algoritmo de Malhas Neurais Abertas poderia ser alterado de forma a incorporar a inteligência embutida nos passos anteriores e dessa forma ser capaz de identificar as várias superfícies presentes nos dados de entrada. Essa idéia não foi abordada nesta tese e permanece como conjectura.

Finalmente, incluindo também a fase de geração do atributo de falha, deve-se considerar a questão de implementar alguns dos vários atributos disponíveis na literatura, bem como algumas ferramentas de pré e pós-processamento de indicadores de falha (assuntos abordados no Capítulo 2). Isso é importante caso se deseje ter um ambiente completo de extração e visualização de superfícies de falhas a partir de dados de amplitude sísmica.