
SEGMENTAÇÃO DE MAPAS AUTO-ORGANIZÁVEIS COM ESPAÇO DE SAÍDA 3-D

José Alfredo Ferreira Costa*
alfredo@dee.ufrn.br

Márcio Luiz de Andrade Netto†
marcio@dca.fee.unicamp.br

*Departamento de Engenharia Elétrica - Centro de Tecnologia
Universidade Federal do Rio Grande do Norte - 59.072-970 - Natal - RN

†Dept. de Engenharia de Computação e Automação - FEEC - UNICAMP - 13083-970 - Campinas - SP

RESUMO

O mapa de Kohonen (SOM) tem sido utilizado como ferramenta para visualização de dados de elevada dimensionalidade. Características importantes da rede SOM incluem a compressão de informação e a tentativa de manutenção da topologia dos dados. Dados similares no espaço de entrada deveriam ser mapeados no mesmo neurônio, ou em neurônios vizinhos. Uma das ferramentas de visualização de um mapa 2-D treinado é *U-matrix*, que apresenta as relações de distância de pesos entre neurônios vizinhos do espaço de saída. A suposição de preservação topológica não é verdade em muitos problemas envolvendo redução de dimensionalidade. Com a automação da detecção de agrupamentos na rede SOM espaços de saída maiores podem ser utilizados em problemas envolvendo a descoberta de classes em dados multidimensionais. Mostra-se a ocorrência de erros topológicos em um exemplo simples de agrupamentos de dados 2-D em um mapa com saída 1-D. Este artigo apresenta uma extensão da *U-matrix*, o *U-array*, para espaços de saída maior que 2-D e sua aplicação em conjunto com o algoritmo SL-SOM, que possibilita a detecção do número e o geometria das classes em mapas treinados. Todo o processo é não-supervisionado.

A vantagem de trabalhar com dimensões mais elevadas no espaço de saída é a melhor preservação da topologia em problemas de análise automática de dados. Apresenta-se um exemplo de uso de descoberta de classes de dados não linearmente separáveis.

PALAVRAS-CHAVE: Classificação não supervisionada de dados, segmentação de volumes, mapa de Kohonen, Redes neurais, mineração de dados.

ABSTRACT

The self-organizing map (SOM) has been widely used as a software tool for visualization of high-dimensional data. Important SOM features include information compression while trying to preserve topological and metric relationship of the primary data items. Similar data in the input space would be mapped to the same neuron or in a nearby unit. The clustering properties of a trained SOM 2-D can be visualized by the U-matrix, which is a neuron's neighborhood distance based image. This assumption of topological preservation is not true for many SOM mappings involving dimension reduction. With the automation of cluster detection in SOM network higher output dimensions can be used in problems involving discovery of classes in multidimensional data. Results of topological errors are shown in a simple 2-D clustering in a 1-D output grid SOM. This paper presents the U-array as an extension of the U-matrix for 3-D output grids. The advantage of the method relies in working with higher dimensions in the output space, which can lead to a better

ARTIGO CONVIDADADO:

Versão completa e revisada de artigo apresentado no SBAI-2005

Artigo submetido em 27/05/2006

1a. Revisão em 31/08/2006

2a. Revisão em 27/11/2006

3a. Revisão em 10/12/2006

Aceito sob recomendação do Editor Convitado

Prof. Osvaldo Ronald Saavedra Mendez

topological preservation in data analysis. Examples of automatic class discovery using U-arrays are also presented.

KEYWORDS: Data Clustering, Volume Segmentation, Self-Organizing Maps, Neural Networks, Data Mining.

1 INTRODUÇÃO

Mineração de Dados (*data mining*) é um campo relativamente novo que tem sido descrito como conjuntos de técnicas para a extração de conhecimento e procura de padrões úteis, e previamente desconhecidos, em grandes volumes de dados multidimensionais (Murtagh, 2000; Cios & Kurgan, 2002; Costa & Netto, 2001b). Aplicações podem ter diferentes objetivos, como a determinação de classes em um conjunto de dados. Pode-se fazer uma síntese do banco de dados observando os objetos representantes de cada subgrupo, para confirmar, ou não, hipóteses a respeito dos dados em questão. Pode-se também formular hipóteses sobre a estrutura dos dados e determinar esquemas de classificação de novas amostras.

Análise de agrupamentos (ou classificação não supervisionada) tem como objetivo descobrir a estrutura (classes de um determinado conjunto de dados não rotulados, $X = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$, onde cada objeto (ou amostra) $\mathbf{x}_i, i = 1, \dots, n$, é descrito por p variáveis (atributos ou características). Apesar do processo ser autônomo, i.e., não supervisionado, na maioria das abordagens o usuário tem grande influência sobre o resultado final do processo através da escolha de parâmetros iniciais. Além disso, a maioria dos métodos, por exemplo estatísticos como o Expectation-Maximization (EM) (Bishop, 1995) impõe geometrias pré-estabelecidas aos agrupamentos, além de possuir custo computacional elevado.

A rede SOM – *self-organizing map* – é o modelo de redes neurais competitivas em maior uso na atualidade, e tem sido utilizado em uma gama de aplicações, incluindo agrupamento e visualização de dados multidimensionais (Kohonen, 2001; Costa, 1999; Costa & Netto, 1999a). O método mais comum de visualização da estrutura de classes na rede SOM 2-D é através da *U-matrix*, que representa distâncias entre neurônios vizinhos na forma de uma superfície, ou imagem $U(x, y)$. As relações entre os neurônios vizinhos são vistas na superfície $U(x, y)$ como vales e montanhas. Porém, a segmentação automática dos agrupamentos de neurônios, representando as classes de padrões na rede SOM treinada é uma tarefa complexa.

Costa e Netto (1999a) apresentaram um método de segmentação da rede SOM usando técnicas de morfologia matemática sobre a *U-matrix*. O resultado do algoritmo são regiões conectadas de neurônios rotuladas, e que definem no espaço de entrada geometrias complexas e não paramétri-

cas. O método foi aplicado a diversos conjuntos de dados com formatos e dimensões variadas (Costa & Netto, 2001b). Recentemente, Vesanto e Alhoniemi (2000) descreveram o uso de técnicas de agrupamentos hierárquicos para aglomerar neurônios, porém há muita iteração do usuário com o resultado final obtido.

Costa e Netto (2003) apresentaram um método de segmentação da rede SOM utilizando a idéia de particionamento de grafos, através da eliminação de conexões (arestas) inconsistentes entre neurônios vizinhos. Tanto o número de agrupamentos quanto os membros das classes são determinados automaticamente pelo algoritmo. Aplicações com sucesso do algoritmo foram também obtidas por Silva *et al* (2004) na área de detecção de fragmentos urbanos utilizando dados geoespaciais.

Uma das premissas principais da qual os métodos supracitados dependem é da preservação topológica efetuada pela rede SOM. Em casos onde a dimensionalidade do espaço de entrada é elevado, geralmente ocorrem distorções no mapa, e o resultado da detecção de clusters pelos algoritmos fica comprometida. Este artigo apresenta o *U-array* como uma extensão da *U-matrix* para mapas com espaço de saída 3-D. A vantagem de se trabalhar com mapas com maiores espaços de saída é exatamente diminuir a interferência da redução da dimensionalidade nas distorções obtidas no processo automático de extração de classes da rede SOM. Como dispomos de métodos automáticos de extração de classes (Costa & Netto, 2001b; 2002, 2003) não há necessidade de visualização. *U-arrays* de dimensão qualquer poderiam ser aplicados. Este trabalho mostra o caso 3-D e ilustra com um conjunto de dados onde duas classes são não linearmente separáveis.

O restante do artigo é organizado da seguinte forma: a seção 2 descreve-se brevemente a rede SOM e a *U-matrix* enquanto que a seção 3 aborda o problema de mapeamento de dados e preservação topológica na rede SOM. O *U-array* é apresentado na seção 4 e resultados de uma segmentação automática utilizando os métodos descritos em (Costa & Netto, 2001b; 2002, 2003) são apresentados na seção 5. Finalmente a seção 6 apresenta as conclusões.

2 VISUALIZAÇÃO DE RELAÇÕES TOPOLÓGICAS NA REDE SOM

A rede SOM é usado para mapear um espaço de entrada p -dimensional, contendo n padrões, para uma grade de neurônios uni- ou bidimensional, com os objetivos de quantizar o espaço de entrada e representar da melhor forma possível a topologia original em um espaço de menor dimensão. Uma vez treinado, podemos rotular neurônios caso tenhamos, além da informação dos padrões, as suas respectivas classes. Padrões são apresentados ao mapa e o neurônio ven-

cedor será o mais similar, ou o mais próximo, de acordo com o critério de similaridade escolhido. Diferentemente de análise de componentes principais (PCA), onde o resultado da projeção é contínua (Jolliffe, 2002), a projeção dos dados na rede SOM é discreta. O resultado de cada apresentação é um índice (i, j) denotando o neurônio vencedor.

Para efetuar análise de agrupamentos, este mapeamento topologicamente ordenado geralmente não é suficiente, pois a informação de distâncias entre os neurônios é perdida. A saída de uma rede SOM para um dado padrão inserido é geralmente o índice do neurônio vencedor c , no caso bidimensional um par de valores (i, j) , e o nível de ativação diretamente relacionado à quantização (no nosso caso a distância do padrão ao neurônio c , computado no espaço de pesos).

A *U-matrix*, proposta por A. Ultsch (1993), tem como objetivo de permitir a detecção visual das relações topológicas dos neurônios. A idéia básica é usar a mesma métrica que foi utilizada durante o treinamento para calcular distâncias entre pesos sinápticos de neurônios adjacentes. O resultado é uma imagem $f(x, y)$, na qual as coordenadas de cada pixel (x, y) são derivadas das coordenadas dos neurônios no *grid* do mapa, ex. $(1, 1), (1, 2) \dots (X, Y) \rightarrow (1, 1), (1, 2) \dots (2*X-1, 2*Y-1)$, e a intensidade de cada *pixel* na imagem $f(x, y)$ corresponde a uma distância calculada. Pode-se pensar uma imagem como uma função tridimensional em que o valor do pixel na coordenada (x, y) é representado por um ponto na coordenada z . Neste caso, teremos uma superfície em 3D cuja topografia revela a configuração dos neurônios obtida pelo treinamento. Vales, neste relevo topográfico, correspondem a regiões de neurônios que são similares, enquanto que montanhas, i.e., valores relativamente elevados na *U-matrix*, refletem a dissimilaridade entre neurônios vizinhos e podem ser associadas a fronteiras (bordas) de agrupamentos de neurônios. Pelo fato da *U-matrix* ser uma imagem relativamente complexa, principalmente em problemas de análise de dados reais, geralmente seu uso é restrito a visualização, sendo uma ferramenta de auxílio na separação manual dos agrupamentos da rede SOM (Kohonen, 2001; Costa, 1999; Costa & Netto, 2001a).

Considere um mapa cujos neurônios possuem vizinhança retangular de tamanho $X \times Y$. Seja $[b_{x,y}]$ a matriz de neurônios e $[w_{i_x,y}]$ a matriz de pesos. Para cada neurônio em b existem três distâncias d_x, d_y e d_{xy} , na *U-matrix* (fig. 1).

Considerando distâncias euclidianas, no caso da topologia da grade de neurônios ser retangular, as distâncias d_x, d_y e d_{xy} podem ser definidas como

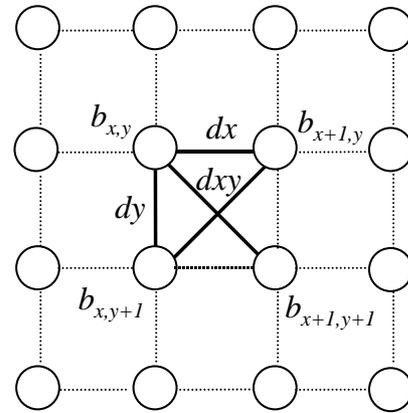


Figura 1: As três distâncias da *U-matrix*

$$d_x(x, y) = \|b_{x,y} - b_{x+1,y}\| = \sqrt{\sum_i (w_{i_x,y} - w_{i_{x+1},y})^2} \quad (1)$$

$$d_y(x, y) = \|b_{x,y} - b_{x,y+1}\| = \sqrt{\sum_i (w_{i_x,y} - w_{i_x,y+1})^2} \quad (2)$$

$$\begin{aligned} d_{xy}(x, y) &= \frac{1}{2} \left(\frac{\|b_{x,y} - b_{x+1,y+1}\|}{\sqrt{2}} + \frac{\|b_{x,y+1} - b_{x+1,y}\|}{\sqrt{2}} \right) \\ &= \frac{1}{2\sqrt{2}} \left[\sqrt{\sum_i (w_{i_x,y} - w_{i_{x+1},y+1})^2} + \sqrt{\sum_i (w_{i_x,y+1} - w_{i_{x+1},y})^2} \right] \quad (3) \end{aligned}$$

3 REDUÇÃO DE DIMENSIONALIDADE E PRESERVAÇÃO TOPOLÓGICA NA REDE SOM

Um algoritmo que efetua redução de dimensionalidade preservando a topologia dos dados é uma transformação $\Phi: \mathbb{R}^p \rightarrow \mathbb{R}^q$ que preserva a ordem de similaridade dos pontos no espaço de entrada \mathbb{R}^p quando estes são mapeados no espaço de saída \mathbb{R}^q (Flexer, 1997). Existem várias técnicas para se encontrar a transformação Φ , como por exemplo técnicas métricas de escalonamento multidimensional (MDS) (Torgerson, 1952), técnicas não métricas de MDS (Shepard, 1962), a projeção de Sammon (1969) e PCA (Jolliffe, 2002).

A rede SOM forma um mapeamento de um espaço de entrada p -dimensional em um arranjo de neurônios, geralmente bidimensional, via aprendizado não supervisionado. Como

geralmente a dimensão dos dados $p \gg 2$ este mapeamento realiza redução de dimensionalidade. Dois aspectos estão envolvidos no mapeamento do espaço de entrada ao espaço de saída. Primeiro, os neurônios tendem a agrupar em regiões de elevada densidade de pontos, i.e., efetuam a tarefa de quantização vetorial. Por outro lado, há uma diferença básica em relação a esquemas convencionais de agrupamentos, como o K -means: a noção de vizinhança dos neurônios e a preservação da topologia entre os espaços. Porém, atingir os dois objetivos de quantização vetorial e preservação da topologia não é uma tarefa trivial.

Pelo fato da rede SOM ter sido descrito de forma algorítmica, e não extremizar explicitamente um funcional conhecido durante o aprendizado, não há, para a rede SOM, uma função de otimização similar a Sammon (1969), i.e., não há conexão formal entre o SOM a outro algoritmo de MDS quanto a redução de dimensionalidade. Outra diferença fundamental entre a rede SOM e outros algoritmos de MDS é que no SOM o número de neurônios é diferente do tamanho do conjunto de dados, n , ocorrendo a quantização dos dados no espaço de saída.

Arranjos de neurônios em redes do tipo SOM de dimensões elevadas raramente são usados na prática por que o objetivo principal do SOM, na atualidade, é a visualização dos dados. Porém, quando usamos a rede SOM com uma dimensão menor que a dimensão natural dos dados quase sempre haverá alguma perda ou distorção da topologia. Por preservação da topologia podemos pensar de forma simples que as relações métricas entre os padrões no espaço original de dados devem ser mantidas no espaço de saída, seja em uma rede do tipo SOM, ou em outro processo de redução de dimensionalidade.

A principal motivação para o uso da rede SOM, com dimensão maior do que 2, é a manutenção da topologia que geralmente é perdida quando diminuímos a dimensionalidade via mapeamento de um espaço p -dimensional para um espaço de menor dimensão. Com a automação da descoberta de conhecimento e relacionamentos entre dados descritas pelos métodos descritos em (Costa, 1999; Costa & Netto, 2001a,b), pode-se usar um arranjo de dimensão igual ou menor que a dimensão do espaço de entrada e fazer com que apenas os resultados finais sejam mostrados, na forma de sub-grupos de dados, o relacionamento entre os subgrupos etc.

Kiviluoto (1996) caracterizou como mapeamento contínuo quando vetores próximos no espaço de entrada são mapeados próximos no espaço de saída. Para caracterizar como uma boa resolução, nenhum par de vetores que estejam distantes do espaço de entrada deveriam ser mapeados em posições próximas do espaço de saída.

A rede SOM tenta aproximar dimensões elevadas contor-

cendo a grade elástica de neurônios na forma de uma curva Peano, resultando em discontinuidades no mapeamento. Kohonen (2001) denominou isto de seleção automática da dimensão dos atributos. Esta propriedade é importante quando uma resolução elevada é desejada. O erro de casamento de dimensões não ocorre entre a dimensão do espaço de saída D_A e a dimensão nominal do espaço de entrada D_X , e sim entre D_A e a dimensão efetiva do espaço D_{eff} (Bauer *et al.*, 1999).

Em algumas aplicações, a preservação da topologia do mapeamento é mais importante do que boa resolução. Seguindo Kiviluoto (1996), deveríamos usar o mapa de forma flexível a fim de encontrar, possivelmente, componentes principais não lineares do espaço de entrada, mas de forma relativamente rígida para que o mapa não se dobre, tentando representar também os componentes menos importantes. A rigidez do mapa pode ser controlada ajustando o tamanho da influência da função de vizinhança, como proposto por Speckmann *et al.* (1994).

O compromisso entre continuidade e resolução não é trivial e alguns critérios foram propostos para quantificar estas propriedades. O erro de quantização é simples de ser avaliado, sendo computado como a distância média das amostras do conjunto de dados X aos neurônios mais próximos. Para quantificar a continuidade vários critérios foram propostos na literatura (Costa, 1999; Kiviluoto, 1996). O mais simples, o Erro Topográfico, ξ_t , é obtido considerando a adjacência dos campos receptivos e a proporção dos objetos que indicam a discontinuidade local do mapeamento. Dado um padrão $\mathbf{x} \in X$, seja \mathbf{m}_i e \mathbf{m}_j o primeiro e o segundo vetores de pesos mais próximos de \mathbf{x} . Caso os neurônios correspondentes n_i e n_j sejam adjacentes o mapeamento preserva, localmente, a topologia. Caso contrário, i.e., se n_i e n_j não forem adjacentes há um erro topográfico local. O erro topográfico para o mapeamento é obtido somando-se todos os erros topográficos locais para todos os padrões.

$$\xi_t = \frac{1}{n} \sum_{k=1}^n u(x_k), \quad (4)$$

onde $u(x_k)$ é igual a 1 caso o primeiro e o segundo neurônios vencedores não sejam adjacentes e 0 caso contrário, e onde n é o número total de padrões. ξ_t corresponde a uma proporção do número de padrões que foram mapeados incorretamente. Porém, não é descrito o tipo de mapeamento incorreto. Por exemplo, não há informação caso dois padrões próximos sejam mapeados em neurônios que distem de uma unidade ou que estejam em cantos opostos do mapa.

3.1 Exemplo de mapeamento de um espaço 2-D em uma rede SOM 1-D

Apresenta-se um exemplo simples de um conjunto de dados no \mathbb{R}^2 com um mapa com espaço de saída unidimensional. A figura 2 mostra um mapa com 40 neurônios, após 1000 iterações do algoritmo tipo lote, onde os pesos foram inicializados linearmente. A figura ilustra também o diagrama de Voronoi, o qual podemos identificar as células de influência dos neurônios, ou campos receptivos induzidos. Note a contorção do *grid* de neurônios no espaço de dimensão maior que a dimensão do mapa para atingir o objetivo de quantização espacial. Os dados foram gerados a partir de médias centradas nos vértices de um triângulo equilátero, com ruído Gaussiano com desvio padrão 0.15. Porém, vemos que apesar das classes serem equidistantes no espaço de entrada, caso utilizássemos a informação do mapa, de forma automática, iríamos pensar que as classes 2 e 3 (parte superior e parte direita da figura) estão em porções distantes do espaço de atributos, pois foram mapeadas em cantos opostos do mapa.

A figura 3 explicita um padrão $x' = (0.2, 0.2)$ que apresenta um erro local topográfico, pois possui como neurônios mais próximos 13 e o 32, respectivamente. A figura 4 ilustra a configuração de neurônios e as regiões no espaço dos atributos onde há erro topográfico local. Apenas os objetos que caírem nas regiões marcadas em preto contribuirão para o erro topográfico. Pode-se definir um erro topográfico global (ξ_{tG}), como a integral no espaço do erro topográfico local. Idealmente ξ_{tG} deveria ser nulo. Um exemplo em que as dimensões do espaço de entrada e de saída são iguais, por exemplo em uma região quadrada bidimensional, o erro topológico global é nulo (Costa, 1999).

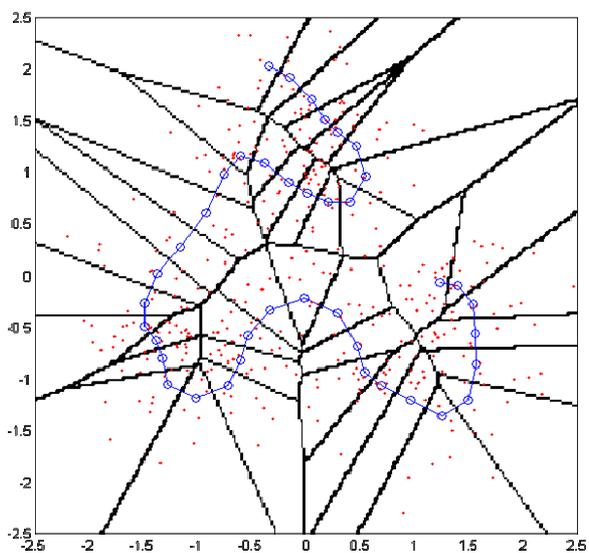


Figura 2: Grid de neurônios e diagrama de Voronoi.

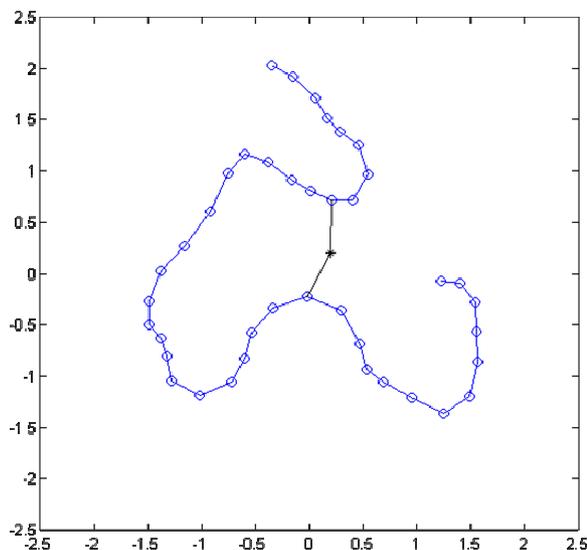


Figura 3: Um exemplo de erro topográfico local.

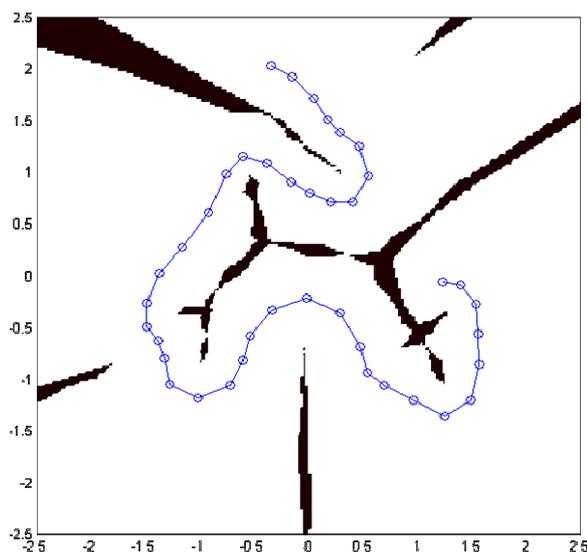


Figura 4: Configuração de neurônios e as regiões no espaço que possuem erro topográfico local.

Tais problemas ocorrem devido ao fato da rede SOM estar efetuando quantização vetorial ao mesmo tempo em que tenta reduzir a dimensionalidade, lembrando algoritmos de escalonamento multidimensional. Porém, aplicando-se (e estendendo-se) os métodos de segmentação de imagens 2D, desenvolvidas para os mapas bidimensionais, para mapas com dimensão 3 ou maior, pode-se extrair, além do correto número das classes, suas informações de proximidades relativas, que é de grande importância em mineração de dados, sem temer estar incorrendo em erros.

4 U-ARRAY - U-MATRIX PARA MAPAS 3-D

O *U-array* é proposto como uma extensão da *U-matrix*. Para o caso da rede SOM com dimensão do espaço de saída 3 podemos ainda visualizar as relações entre os neurônios efetuando-se cortes no *array* tridimensional. Para dimensões maiores do espaço de saída, apesar de não podermos visualizar em uma única imagem as relações entre os neurônios, podemos escolher combinações de pares de componentes e visualizar imagens bidimensionais como projeções do espaço de maior dimensão.

A extensão para casos de dimensão acima de 3 pode ser efetuada. Idealmente deveríamos escolher a dimensão do espaço de saída da rede SOM através de informações provenientes diretamente dos dados, por exemplo, usando alguma informação derivada da dimensão fractal (Bauer et al., 1999).

A estratégia adotada na prática é escolher uma dimensão para o mapa e para os outros parâmetros, e ao final do treinamento pode-se checar índices como o produto topográfico, o erro topográfico ou a função topográfica. Caso os índices estejam dentro de valores aceitáveis pode-se dar início à análise da rede SOM pela *U-matrix* ou pelo *U-array*. Caso contrário, deveríamos reiniciar o treinamento (ou acrescentar uma dimensão ao espaço de saída) na rede SOM com maior ou menor dimensão, de acordo com as informações obtidas a partir dos índices de manutenção topológica (Kiviluoto, 1996; Bauer et al., 1999; Costa, 1999).

No caso do *U-array* onde a rede SOM tem espaço de saída tridimensional, temos, além das distâncias dx , dy e dxy , as distâncias dz , dxz , dyz e $dxyz$. A figura 5 ilustra a rede SOM com tamanho $3 \times 3 \times 3$. Note que podemos pensar em vários mapas bidimensionais contidos neste mapa. Para facilitar a visualização das relações entre os neurônios, apenas o espaço de saída está sendo apresentado, porém, supõe-se que todos os neurônios estejam conectados à camada de entrada por pesos sinápticos.

Note que poderíamos pensar em várias *U-matrizes*, uma para cada um dos nove mapas bidimensionais (após cortes transversais no cubo). Considere um mapa retangular de tamanho $X \times Y \times Z$, no caso tridimensional. Da mesma forma que na *U-matrix*, o *U-array* terá tamanho $(2X-1) \times (2Y-1) \times (2Z-1)$. Por exemplo, para o mapa $3 \times 3 \times 3$ apresentado na figura 5 teremos um *U-array* de tamanho $5 \times 5 \times 5$, e poderíamos usar a informação das *U-matrizes* de cada mapa e interpolar nas posições interiores do *U-array*. As distâncias entre os neurônios podem ser vistas de forma mais simples para um mapa com tamanho $2 \times 2 \times 2$, como o apresentado na figura 6 por questões de visibilidade, onde aparece também o endereço ou índice do neurônio no espaço de saída da rede.

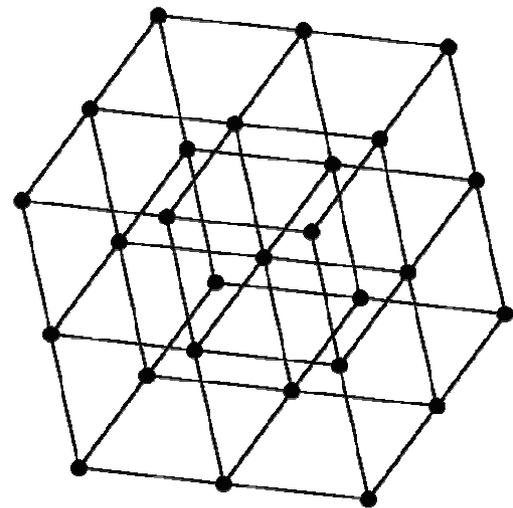


Figura 5: Espaço de saída de uma rede SOM com tamanho $3 \times 3 \times 3$.

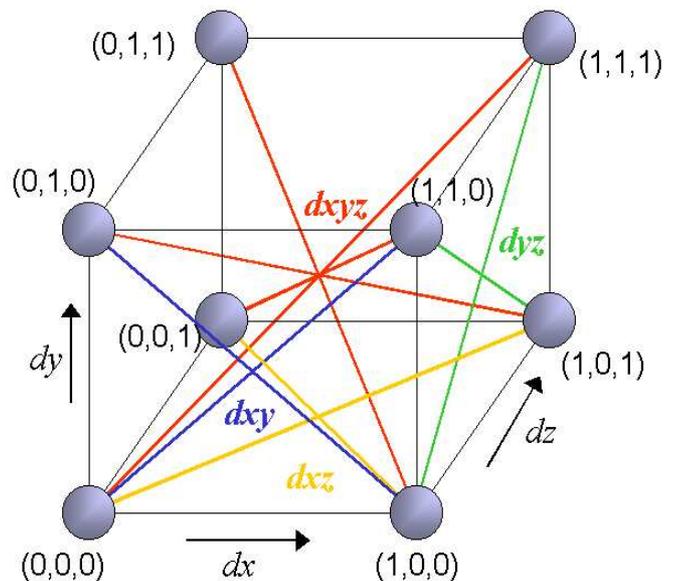


Figura 6: Ilustração do mapa de tamanho $2 \times 2 \times 2$.

As distâncias dx , dy e dz (equações 1-3) também são mostradas na figura 6. A notação usada nesta seção é a mesma da seção 2. Seja $[b_{x,y,z}]$ a matriz de neurônios e $[w_{ix,y,z}]$ a matriz de pesos. As distâncias dxy , dxz e dyz são semelhantes à distância usada no caso bidimensional, dxy , apresentada pela equação 3. As fórmulas para as outras distâncias são apresentadas nas equações 5-11.

$$dx(x, y, z) = \|b_{x,y,z} - b_{x+1,y,z}\| = \sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x+1,y,z}})^2} \quad (5)$$

$$dy(x, y, z) = \|b_{x,y,z} - b_{x,y+1,z}\| = \sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x,y+1,z}})^2} \quad (6)$$

$$dz(x, y, z) = \|b_{x,y,z} - b_{x,y,z+1}\| = \sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x,y,z+1}})^2} \quad (7)$$

$$dxy(x, y, z) = \frac{1}{2} \left(\frac{\|b_{x,y,z} - b_{x+1,y+1,z}\|}{\sqrt{2}} + \frac{\|b_{x,y+1,z} - b_{x+1,y,z}\|}{\sqrt{2}} \right) = \frac{1}{2\sqrt{2}} \left[\sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x+1,y+1,z}})^2} + \sqrt{\sum_i (w_{i_{x,y+1,z}} - w_{i_{x+1,y,z}})^2} \right] \quad (8)$$

$$dxz(x, y, z) = \frac{1}{2} \left(\frac{\|b_{x,y,z} - b_{x+1,y,z+1}\|}{\sqrt{2}} + \frac{\|b_{x,y,z+1} - b_{x+1,y,z}\|}{\sqrt{2}} \right) = \frac{1}{2\sqrt{2}} \left[\sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x+1,y,z+1}})^2} + \sqrt{\sum_i (w_{i_{x,y,z+1}} - w_{i_{x+1,y,z}})^2} \right] \quad (9)$$

$$dyz(x, y, z) = \frac{1}{2} \left(\frac{\|b_{x,y,z} - b_{x,y+1,z+1}\|}{\sqrt{2}} + \frac{\|b_{x,y+1,z} - b_{x,y,z+1}\|}{\sqrt{2}} \right) = \frac{1}{2\sqrt{2}} \left[\sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x,y+1,z+1}})^2} + \sqrt{\sum_i (w_{i_{x,y+1,z}} - w_{i_{x,y,z+1}})^2} \right] \quad (10)$$

A distância $dxyz$ é uma média das distâncias entre os vértices opostos do cubo:

$$dxyz(x, y, z) = \frac{1}{4\sqrt{3}} \left[\sqrt{\sum_i (w_{i_{x,y,z}} - w_{i_{x+1,y+1,z+1}})^2} + \sqrt{\sum_i (w_{i_{x,y,z+1}} - w_{i_{x+1,y+1,z}})^2} + \sqrt{\sum_i (w_{i_{x,y+1,z+1}} - w_{i_{x+1,y,z}})^2} + \sqrt{\sum_i (w_{i_{x,y+1,z}} - w_{i_{x+1,y,z+1}})^2} \right] \quad (11)$$

Considerando os índices da figura 6, o U -array de tamanho $3 \times 3 \times 3$ poderia ser visto como o apresentado na figura 7. Da mesma forma que na U -matrix, o cálculo dos valores $du(x, y, z)$ pode ser feito tomando-se o valor médio ou a mediana dos elementos circunvizinhos.

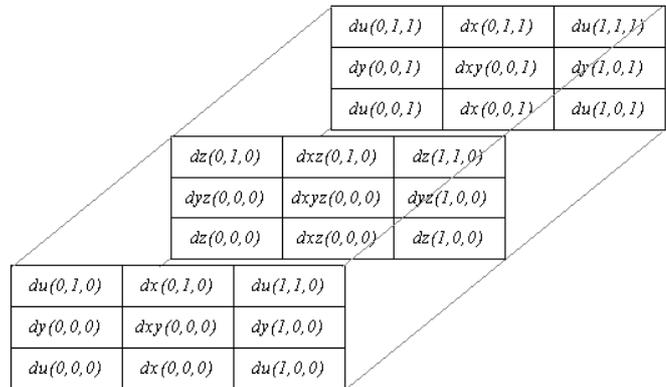


Figura 7: Esquema do U -array para a rede SOM com tamanho $2 \times 2 \times 2$.

5 ADAPTAÇÃO DO ALGORITMO SL-SOM PARA O U -ARRAY

O algoritmo SL-SOM foi implementado utilizando a segmentação por *watersheds* (Costa, 1999). A seguir, descreve-se, de forma bastante sucinta, a idéia básica da watershed, baseado em Meijster e Roerdink (1996).

Considere uma imagem digital em níveis de cinza como uma função $f : D \rightarrow \mathbb{IN}$, onde $D \subseteq \mathbb{Z}^2$ é o domínio da imagem e $f(p)$ o nível de cinza do pixel $p \in D$. Seja $E \subset \mathbb{Z}^2 \times \mathbb{Z}^2$ o conjunto de coordenadas da imagem, P o caminho entre dois pixels p e q e com comprimento $l(P)$. Para um pixel $p \in D$, seja $N_E(p) = \{ q \in D \mid (p, q) \in E \}$ o conjunto de pixels vizinhos de p .

O custo de ir de uma posição p a uma posição vizinha q pode ser definido como

$$custo(p, q) = \begin{cases} LS(p), & \text{caso } f(p) > f(q) \\ LS(q), & \text{caso } f(p) < f(q) \\ \frac{LS(p)+LS(q)}{2}, & \text{caso } f(p) = f(q) \end{cases} \quad (12)$$

onde $LS(p)$ é a rampa máxima ligando um pixel p a qualquer um de seus vizinhos de menor altitude, definida como

$$LS(p) = \max_{q \in \{p\} \cup N_E(p)} [f(p) - f(q)] \quad (13)$$

A distância topográfica entre dois pontos p e q ao longo de um caminho $P = (p_0, \dots, p_{l(P)})$ é definida como

$$T_f^P(p, q) = \sum_{i=0}^{l(P)-1} custo(p_i, p_{i+1}) \quad (14)$$

e a distância topográfica entre dois pontos p e q é definida como a distância topográfica mínima considerando todos os caminhos entre p e q :

$$T_f(p, q) = \text{MIN}_{P: p \rightarrow q} T_f^P(p, q) \quad (15)$$

onde $p \mapsto q$ representa o conjunto de todos os caminhos de p a q . Note que $T_f(p, q)$ é nula caso os pixels p e q sejam do mesmo objeto e possuam o mesmo nível de cinza. A distância topográfica entre um ponto $p \in D$ e um conjunto $A \subseteq D$ é definida como

$$T_f(p, A) = \text{MIN}_{a \in A} T_f(p, a). \quad (16)$$

Seja a função f^* definida como a função f , porém substituindo todos os valores dos mínimos locais por 0, ou seja, caso p seja um mínimo $f^*(p) = 0$, caso contrário, $f^*(p) = f(p)$. Seja $(m_i)_{i \in I}$ o conjunto de mínimos da função f^* . A base de retenção de um mínimo m_i , denotada por $CB(m_i)$, é definida como o conjunto de pontos $p \in D$ que estejam topograficamente mais próximos a m_i do que a qualquer outro mínimo m_j :

$$CB(m_i) = \{p \in D \mid \forall j \in I \setminus \{i\} : T_{f^*}(p, m_i) < T_{f^*}(p, m_j)\} \quad (17)$$

e o *watershed* de uma função f é o conjunto de pontos deste domínio que não pertence a nenhuma bacia de retenção

$$W(f) = D \cap \left[\bigcup_{i \in I} CB(m_i) \right]^c \quad (18)$$

onde c denota o conjunto complemento.

A extensão para 3-D é feita considerando $D \subseteq \mathbb{Z}^3$, $E \subset \mathbb{Z}^2 \times \mathbb{Z}^2 \times \mathbb{Z}^2$, o conjunto de coordenadas do volume, P o caminho entre dois voxels p e q e com comprimento $l(P)$. Para um voxel $p \in D$, seja $N_E(p) = \{q \in D \mid (p, q) \in E\}$ o conjunto de voxels vizinhos de p .

Métodos de computação de *watershed* baseados em algoritmos de grafos paralelos, usando o algoritmo de Dijkstra, foram apresentados em Meijster & Roerdink (1996). A implementação atual, utilizada neste artigo, é feita de forma eficiente através da IFT - Image Foresting Transform (Falcão *et al.*, 2004; Lotufo e Falcão, 2000). A IFT é uma técnica baseada em grafos para projeto de operadores de processamento de imagens baseados em conectividade.

Obviamente, a segmentação de um cubo, ou hiper-cubo, é mais complexa do que de uma imagem. No caso de um cubo, o gráfico de regiões planas conectadas *versus* limiar (nível de cinza) torna-se um gráfico de um volume conectado *versus* o limiar. A conectividade entre *pixels* torna-se conectividade entre *voxels*. O caso mais simples em 2D, i.e., conectividade 4-adjacente, pode ser estendida e computada de forma relativamente simples em uma imagem p -dimensional: para qualquer coordenada caso a distância entre dois *pixels* seja 1 eles estarão conectados.

A segmentação de uma imagem 3D pode também ser vista como a segmentação de uma seqüência de imagens 2D. O algoritmo watershed foi adaptado para efetuar tal operação (em 3D). Um dos principais problemas é a detecção dos marcadores, suavização da imagem e eliminação de voxels isolados. A operação rotulação de regiões conectadas passa a ser uma operação de rotulação de volumes conectados. Os tipos de conectividades utilizados neste trabalho foram os padrões 6 e 26-adjacente. Maiores detalhes em (Costa, 1999).

Um dos passos principais do algoritmo SL-SOM é a escolha dos marcadores para a watershed. Seja o U-array de uma rede SOM treinado dada pela imagem f , de tamanho $2N-1 \times 2M-1 \times 2K-1$, onde $N \times M \times K$ é o tamanho do mapa. Considere que $[f_{min}, f_{max}] = [0, 255]$ e n_i é 1, ou seja, há 256 níveis de cinza na imagem (array) f . Os seguintes passos são efetuados:

Algoritmo de escolha dos marcadores da watershed

1. Filtragem: a imagem f_1 é gerada a partir do fechamento morfológico (dilatação seguido de abertura) na imagem f .
2. Para $k = 1, \dots, f_{max}$, onde f_{max} é o nível de cinza máximo na imagem f_1 , crie as imagens (volumes) binárias f_2^k correspondendo a conversões de f_1 usando k como valor de limiar.
3. Calcule o número de regiões conectadas de f_2^k , para cada valor de k , N_{rc}^k .
4. Procure no gráfico $k \times N_{rc}^k$ a maior seqüência contígua e constante de número de regiões conectadas N_{rc}^k , denotado por S_{max} .

5. A imagem de marcadores será a imagem f_2^j , onde j é o valor inicial da seqüência S_{max} .

O passo 1 suaviza, discretamente, a imagem original. O elemento estruturante utilizado foi uma volume $3 \times 3 \times 3$. O passo 2 pode ser entendido como uma seqüência de limiarizações de níveis da imagem, do mínimo ao máximo. Para cada imagem (ou volume) binária derivada de uma operação de limiarização, a operação de rotulação atribui um código diferente a cada um dos componentes conectados, resultando em um número de objetos na imagem para cada valor limiar k aplicado, N_{rc}^k .

Toda análise e segmentação de um mapa de tamanho $N \times M \times K$ treinado é feita inicialmente sobre o U -array, e posteriormente as informações associadas aos voxels são associadas aos neurônios. Os passos do SL-SOM são:

O algoritmo SL-SOM

1. Obtenção do U -array (seção 4);
2. Encontrar os marcadores para a U -array (seção 5);
3. Aplicar o *watershed* sobre a U -array usando os marcadores obtidos no passo 2;
4. Rotulagem dos volumes conectados; e
5. Cópia dos rótulos obtidos no passo 4 para os neurônios associados a cada voxel do U -array.

6 EXEMPLO DE SEGMENTAÇÃO DE MAPAS ESPAÇO DE SAÍDA 3-D

Um exemplo para comparações de métodos de agrupamentos em dados multidimensionais foi proposto por Ultsch (1995), o *chainlink*, que consiste de 1000 pontos no espaço \mathbb{R}^3 tal que eles possuem a forma de dois anéis tridimensionais entrelaçados. Um dos anéis se estende na direção $x - y$ enquanto o outro se estende na direção de $x - z$. Os dois anéis podem ser pensados como elementos de uma corrente, cada um consistindo de 500 objetos de dados. Este problema ilustra a capacidades da rede SOM em descobrir a estrutura dos dados mesmo para conjuntos de dados com forma complexas e não-esféricas, e não separáveis linearmente.

Efetando análises estatísticas junto ao conjunto de dados, O objetivo é, partindo apenas do conjunto de dados, X , determinar o número de classes presentes nos dados, e prover conjuntos de neurônios no modelo distribuído de protótipos para cada classe, caso sejam descobertas. O resultado da aplicação deste conjunto de dados em um mapa 2D foi apresentada em Costa & Netto (1999b). Apesar do resultado em 2D ser

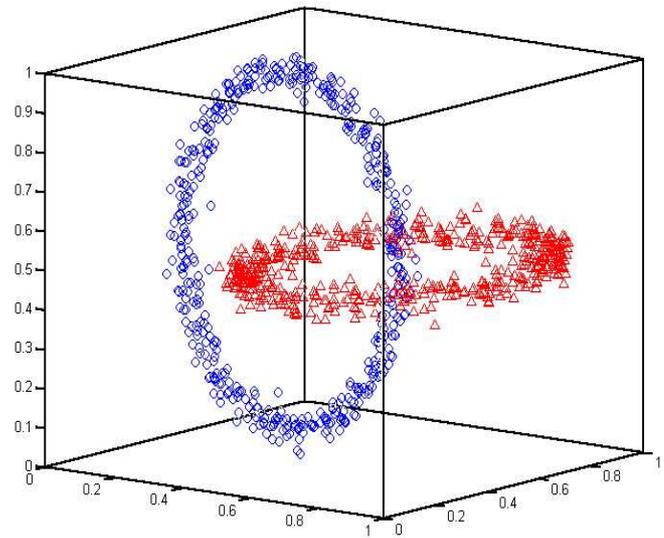


Figura 8: Ilustração do conjunto de dados chainlink.

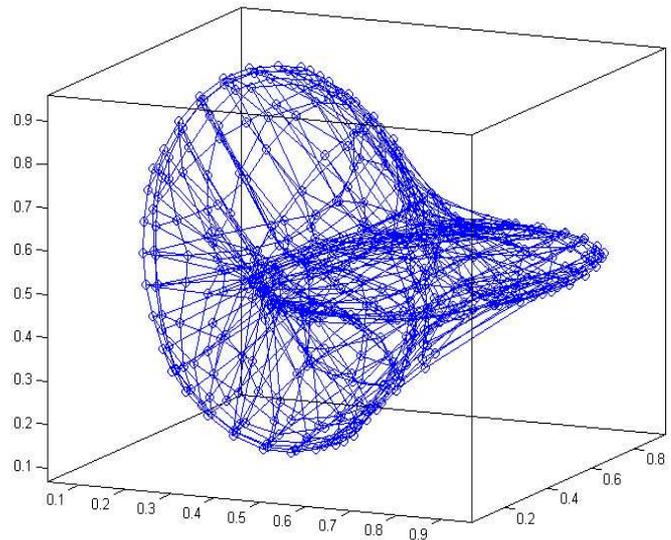


Figura 9: Rede SOM após 200 épocas de treinamento.

suficiente para este problema, onde os dados estão no \mathbb{R}^3 , a motivação para o uso desta base com uma rede SOM com espaço de saída 3D é ilustrativa, no sentido de tentar entender a adequação do mapa a problemas em que não há redução de dimensionalidade, e a topologia pode efetivamente ser preservada.

A estrutura da rede SOM utilizada no experimento continha $8 \times 8 \times 8$ neurônios e o número de épocas (algoritmo em lote) utilizado foi 200. A figura 9 ilustra a configuração da rede após o treinamento enquanto que o U -array correspondente é mostrado na figura 10, na forma de um cubo (3D).

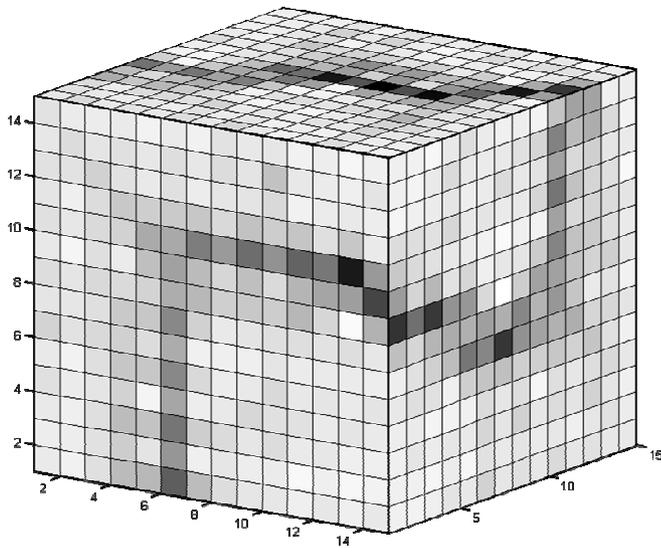


Figura 10: U-array relativo à rede mostrada na figura 9.

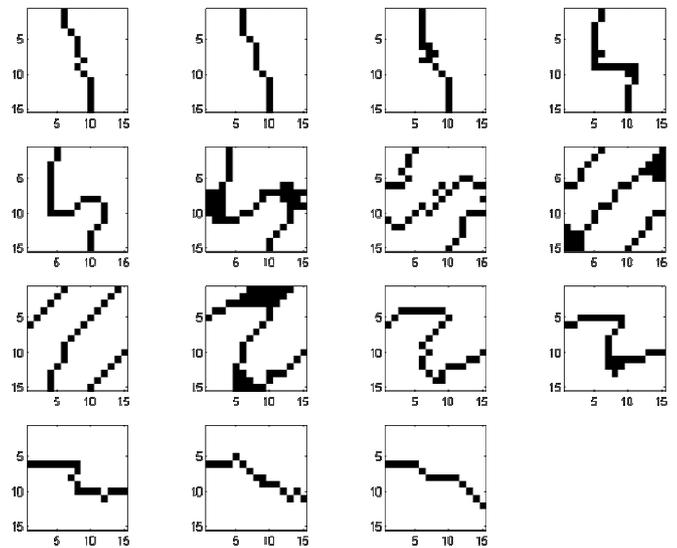


Figura 12: Linhas de watershed obtidas (visão de cortes).

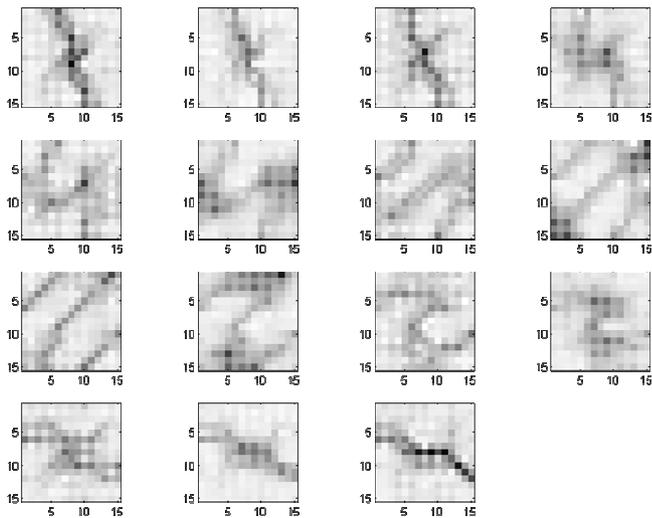


Figura 11: Cortes ortogonais ao eixo z do U-array.

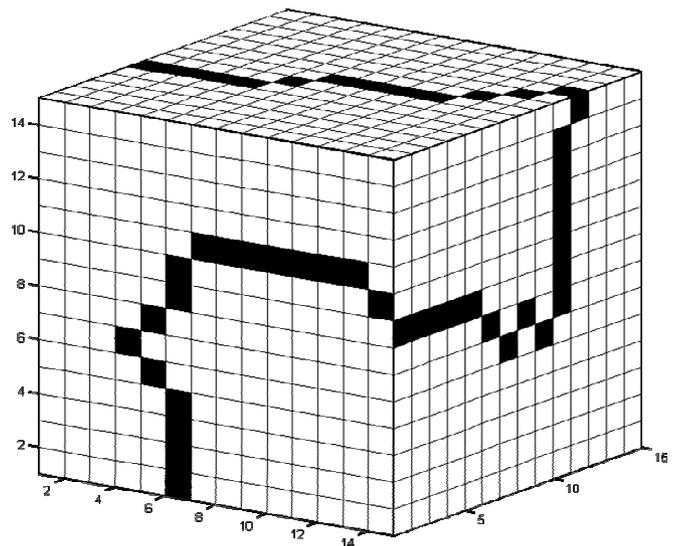


Figura 13: Linhas de watershed obtidas (visão 3D).

A figura 11 apresenta o U-array na forma de cortes ortogonais ao eixo z. Em ambas as figuras, níveis de cinza escuros simbolizam dissimilaridade mais acentuada. Um dos maiores problemas neste tipo de análise é a segmentação correta de imagens no \mathcal{R}^p , $p > 2$. O algoritmo watershed foi estendido para suportar matrizes multidimensionais. A escolha correta dos marcadores, de forma automática, é sem dúvida, um dos elementos chave no processo. Os marcadores utilizados foram obtidos pela binarização do U-array utilizando o valor inicial (40) do platô de estabilidade no gráfico número de volumes conectados *versus* o limiar, cuja faixa foi de 40 a 132, estável para a solução dois agrupamentos, solução obtida automaticamente pelo método SL-SOM (Costa & Netto, 2001a,b; 2002). As linhas da watershed obtidas são

mostradas nas figuras 12 e 13.

Os rótulos obtidos na segmentação do U-array são copiados para as posições respectivas do mapa de Kohonen, no qual vemos na figuras 14. A figura 15 ilustra a configuração dos neurônios eliminando o efeito dos neurônios inativos (Costa, 1999), $H(i) = 0$. A influência espacial dos agrupamentos detectados é ilustrado, limitando o raio das esferas em 0.05. Nesta última figura, buscou-se um ângulo de visualização que enfoca a solução obtida pelo método para este conjunto de dados não linearmente separável. Relações entre neurônios vizinhos também são mostradas.

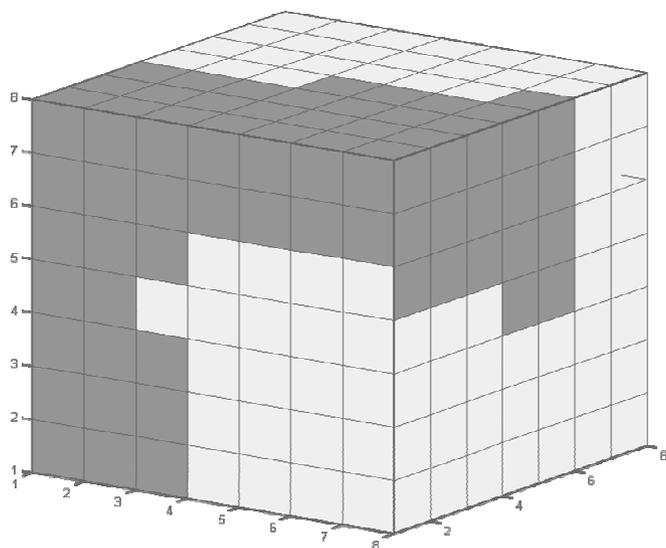


Figura 14: Espaço de saída da rede SOM rotulada (dois volumes segmentados) - representação 3D.

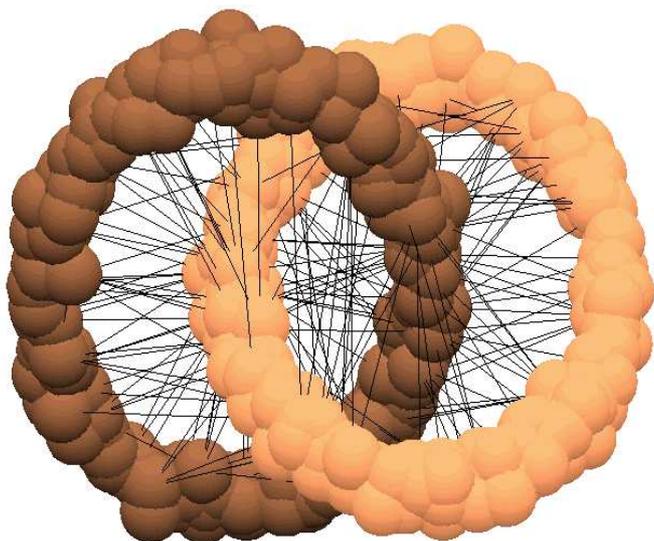


Figura 15: Configuração dos neurônios eliminando o efeito dos neurônios inativos.

Note que ambos os agrupamentos de neurônios fecham os círculos relacionados aos agrupamentos originais dos dados. No caso bidimensional, a deformação da rede 2D apenas simula tal fechamento. Vale ressaltar que métodos como o *K-means* ou o baseado em misturas de densidades Gaussianas foi incapaz de resolver este problema, particionando o conjunto de dados por hiperplanos (no caso do *K-means*) e superfícies quadráticas (maiores detalhes em Costa (1999a,b) e Costa & Netto (2001a,b; 2002).

7 CONCLUSÕES

Atualmente grande parte das aplicações e do uso dos mapas de Kohonen usam espaços de saída bidimensionais porque a principal função é a visualização do mapeamento do espaço original (de elevada dimensão) no display ou conjunto de neurônios da rede. Porém, geralmente há perda de relações topológicas entre mapeamentos de espaços com dimensão mais elevada para espaços com menor dimensão, e muitas vezes tal informação pode ser bastante útil. Um exemplo ilustrado na seção 3.1 mostra que, em um problema relativamente simples (dados em 2-D), classes vizinhas no espaço foram mapeadas em posições relativamente distantes no mapa 1-D.

Em aplicações de mineração de dados e descoberta de conhecimento em bases de dados, pode ser importante saber não só a existência das classes e seus elementos componentes, mas também que classes são semelhantes. Esta última informação motivou o desenvolvimento do *U-array* e da extensão dos métodos para sua segmentação e análise.

As contribuições deste artigo incluem a extensão dos métodos de classificação não supervisionada de dados utilizando a rede SOM para espaços de saída com dimensão maiores que 2, com a definição do *U-array* e sua utilização em conjunto com o algoritmo *SL-SOM*.

A metodologia apresentada neste artigo para segmentação do *U-array* baseou-se no uso de morfologia matemática com a transformada watershed, porém métodos complementares para segmentação de mapas com espaço de saída 2-D e que podem ser estendidos a espaços de maior dimensionalidade tem sido abordados pelos autores, como por exemplo, a segmentação da rede SOM utilizando a idéia de particionamento de grafos (Costa e Netto, 2003) e o uso de métodos de agrupamentos hierárquicos com conectividade restrita (Costa, 2005). Os algoritmos têm sido aplicados com sucesso em vários casos como em segmentação de imagens multi-espectrais (Gonçalves et al., 2005), detecção de fragmentos urbanos utilizando dados geoespaciais (Silva et al., 2004), entre outros. Vantagens em relação a outros métodos como o apresentado em Vesanto e Alhoniemi (2000), que possuem muita interação do usuário com o resultado final obtido, incluem, por exemplo, métodos para estimar o número de agrupamentos baseado na estabilidade das fusões versus a distância entre os agrupamentos.

Trabalhos futuros podem incluir medidas de erro topológico como indicadores para ajuste da dimensão do espaço de saída da rede SOM, de forma que tenhamos algoritmos de treinamento que adaptem automaticamente a dimensão da rede para um dado problema. Também há estudos promissores do uso da transformada watershed juntamente com índices de validação de agrupamentos (Gonçalves et al., 2006) e que

podem ser estendidos para o caso de mapas com espaço de saída 3-D.

REFERÊNCIAS

- Bauer, H.-U., Herrmann, M., & Villmann, T. (1999). Neural maps and topographic vector quantization. *Neural Networks*, v. 12, [4-5], pp. 659-676.
- Bishop, C.M. (1995). *Neural Networks for Pattern Recognition*, Oxford: Oxford University Press.
- Cios, K. J., & Kurgan, L. (2002), Trends in Data Mining and Knowledge Discovery, In: Pal, N.R. et al. (Eds.), *Knowledge Discovery in Advanced Information Systems*, Springer.
- Costa, J. A. F. (1999). *Classificação Automática e Análise de Dados por Redes Neurais Auto-Organizáveis*. Tese de Doutorado, Unicamp, SP.
- Costa, J.A.F. and Netto, M.L.A. (1999a). Estimating the Number of Clusters in Multivariate Data by Self-Organizing Maps. *Intl. Journal of Neural Systems*, vol. 9, pp. 195-202.
- Costa, J.A.F., & Netto, M. L. A. (1999b). Cluster Analysis Using Self-Organizing Maps and Image Processing Techniques. In: *Proc. of the IEEE Conf. on Systems, Man, and Cybernetics*, v. 5, pp. 367-372, Tokyo, Japan.
- Costa, J.A.F. and Netto, M.L.A. (2001a). A new tree-structured self-organizing map for data analysis. In: *Proc. of the Intl. Joint Conf. on Neural Networks*, Washington, DC, pp. 1931-1936.
- Costa, J.A.F., and Netto, M. L. A. (2001b). Clustering of complex shaped data sets via Kohonen maps and mathematical morphology. *Proceedings of the SPIE*, vol. 4384, pp. 16-27.
- Costa, J.A.F., e Netto, M.L.A. (2002). Segmentação automática de mapas de Kohonen. *Congresso Brasileiro de Automática*, Natal, RN, pp. 1607-1613.
- Costa, J.A.F., e Netto, M.L.A. (2003) Segmentação do SOM Baseada em Particionamento de Grafos. *Anais do VI Congresso Brasileiro de Redes Neurais*, São Paulo, pp. 451-456.
- Costa, J.A.F. (2005). Segmentação do SOM por Métodos de Agrupamentos Hierárquicos com Conectividade Restrita. *VII Congresso Brasileiro de Redes Neurais*, Natal, RN, outubro de 2005.
- Falcão, A.X., Stolfi, J., and Lotufo, R.A. (2004). The Image Foresting Transform: Theory, Algorithms and Applications. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, v. 26, 1, pp.19-29.
- Flexer, A. (1997). Limitations of Self-organizing Maps for Vector Quantization and Multidimensional Scaling, in Mozer M.C., et al.(eds.), *Advances in Neural Information Processing Systems 9*, MIT Press, pp. 445-451.
- Gonçalves, M., Netto, M.L.A., Costa, J.A.F., e Zullo, J. (2005). Automatic Remotely Sensed Data Clustering by Tree-Structured Self-Organizing Maps. In: *IGARSS'05 (2005 IEEE International Geoscience and Remote Sensing Symposium)*, July, 25-29 Seoul, Korea.
- Goncalves, M., Netto, M., Costa, J.A.F., and Zullo, J. (2006). Data Clustering using Self-Organizing Maps segmented by Mathematic Morphology and Simplified Cluster Validity Indexes: an application in remotely sensed image. In: *2006 IEEE International Joint Conference on Neural Networks*, Vancouver, BC, Canada. Accepted.
- Jolliffe, I.T. (2002). *Principal Component Analysis, 2nd Ed.*. New York: Springer.
- Kiviluoto, K. (1996). Topology preservation in Self-Organizing Maps. In: *Proc. IEEE Intl. Conf. on Neural Networks*, v. 1, pp. 249-254.
- Kohonen, T. (2001). *Self-Organizing Maps, 3rd Ed.*, Berlim: Springer.
- Lotufo, R. and Falcao, A. (2000). The Ordered Queue and the Optimality of the Watershed Approaches. in: *Computational Imaging and Vision*, vol. 18. Boston, MA: Kluwer. June.
- Meijster, A. & Roerdink, J. B. (1996). Computation of watersheds based on parallel graph algorithms. In: Maragos, P. et al. (Eds). *Mathematical Morphology and its Applications to Image and Signal Processing*, pp. 305-312. Boston, MA: Kluwer.
- Murtagh, F. (2000). Clustering massive data sets, in J. Abello et al. (Eds.), *Handbook of Massive Data Sets*, Kluwer.
- Sammon, J. W. (1969). A Non-Linear Mapping for Data Structure Analysis. *IEEE Trans. in Computers*, v.18, pp. 401-409.
- Shepard, N. R. (1962). The analysis of proximity: multidimensional scaling with an unknown distance function I. *Psychometrika*, v.27, 125-139.
- Silva, M.A.S., Monteiro, A.M.V., and Medeiros, J.S. (2004). Semi-Automatic Geospatial Data Clustering by Self-Organizing Maps. In: *Anais do Simpósio Brasileiro de Redes Neurais*, São Luiz, MA, outubro de 2004.

- Speckmann, H., Raddatz, G. & Rosenstiel, W. (1994). Considerations of geometrical and fractal dimension of SOM to get better learning results. In: *Proc. ICANN'94, Int. Conf. on Artificial Neural Networks*, vol. I, pp. 342-345.
- Torgerson, W. S. (1952). Multidimensional Scaling I. Theory and Method. *Psychometrika*, vol. 17, pp. 401-419.
- Ultsch, A. (1993). Self-Organizing Neural Networks for Visualization and Classification. In: O. Opitz et al. (Eds). *Information and Classification*. Springer, Berlin, pp. 307-313.
- Ultsch, A. (1995). Self-Organizing Neural Networks perform different from statistical k-means clustering. *Gesellschaft für Klassifikation*, Basel.
- Vesanto, J. and Alhoniemi, E. (2000). Clustering of the Self-Organizing Map, *IEEE Trans. on Neural Networks*, v. 11, (3), pp. 586-602.