高次元配置を用いたグラフの視覚化

細部 博史[†]
国立情報学研究所[†]

1 はじめに

インタラクティブなアプリケーションやシステムで必要とされる情報視覚化のために，**グラフ配置**技術が盛んに研究されている．本研究では，**固有ベクトル**を用いる新しい一般無向グラフ配置法を提案する．本手法は，(1)グラフ配置の局所的品質を向上し，(2)インタラクションに応じてグラフ配置を高速に更新するという特徴を持つ．これらの実現のために，本手法では，**高次元空間**におけるグラフ配置を求めた上で，2次元平面上へ射影するというアプローチを採用する．

2 固有ベクトルを用いたグラフ配置法

本節では，固有ベクトルを用いたグラフ配置の従来手法を紹介する．

2.1 Torgerson の方法

固有ベクトルを用いたグラフ配置法の基礎として，**Torgerson の方法**[1]を述べる[1]．この方法は，統計学の分野において代表的な計量的**多次元尺度法**として知られており，すべてのオブジェクトの組合せに対して距離が与えられているときに，それらの距離を満たすようなオブジェクトの配置を求めるものである．

n 個のオブジェクトに対して，任意のオブジェクト i, j の間の距離 d_{ij} が与えられ，距離の公理が満たされていると仮定する．まず，以下のように a_{ij} を定める．

$$a_{ij} = \frac{1}{2}\left(\frac{1}{n}\sum_k d_{ik}{}^2 + \frac{1}{n}\sum_k d_{kj}{}^2 - \frac{1}{n^2}\sum_k\sum_l d_{kl}{}^2 - d_{ij}{}^2\right)$$

次にn次正方行列$A = (a_{ij})$を考えると，これは実対称行列となる．従って，Aは直交行列Xによって$X^TAX = \Lambda$のように対角化可能である．ただし，Aの固有値をλ_kとし，$\boldsymbol{x}_k$をλ_kに対応する固有ベクトルとすると，Λは各(k, k)要素をλ_kとする対角行列であり，$X = (\boldsymbol{x}_1\ \boldsymbol{x}_2 \ldots \boldsymbol{x}_n)$である．

さらに，各(k, k)要素を$\sqrt{\lambda_k}$とする対角行列$\Lambda^{1/2}$を用いて $P = X\Lambda^{1/2}$とおくと，理想的な d_{ij}の組が与えられたとき，各固有値λ_kは非負となり，P の第 i 行$(p_{i1}\ p_{i2} \ldots p_{in})$は，オブジェクト i の n 次元実 Euclid 空間における座標と見なせる．ただし，現実のデータに対しては，負の値を取るような固有値が現れることが多い．

通常の多次元尺度法では，最大の固有値を少数個だけ用いて，それに対応する座標成分のみを利用する．例えば，λ_1 とλ_2 をそれぞれ 1, 2 番目に大きい固有値であるとし，それらのみを用いるとする場合，各オブジェクト i の座標を(p_{i1}, p_{i2})とするような 2 次元平面上の配置が得られる．

2.2 Torgerson の方法に基づくグラフ配置法

Kruskal らは，Torgerson の方法に基づくグラフ配置法を提案した[3]．これは以下のように実現されている．

1. 最初に，グラフの任意のノード間のグラフ理論的距離を求める．
2. 次に，グラフ理論的距離を用いて Torgerson の方法を実行し，2 次元平面上のノードの配置を計算する．

3 提案手法

本節では，固有ベクトルを用いる新しいグラフ配置法を提案する．

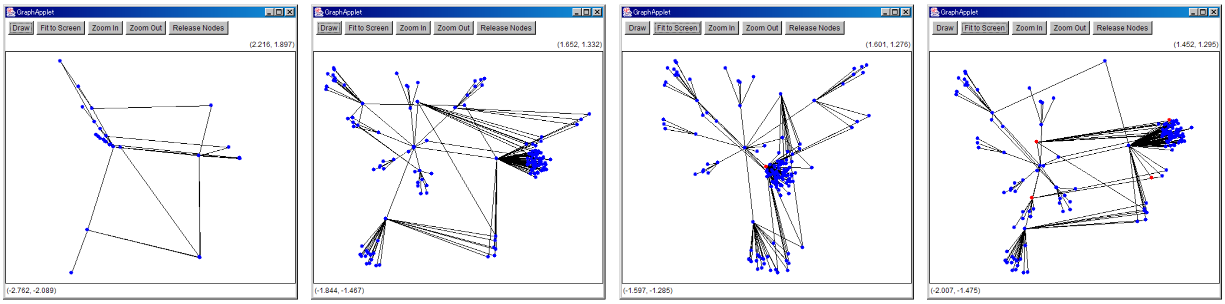
3.1 2 次元グラフ配置の計算

固有ベクトルを用いたグラフ配置の従来手法は，グラフの局所的な情報を表現できないという問題がある[2]．この問題点に対処するため，本研究では，高次元空間におけるグラフ配置の情報を活用することを提案する．具体的には，固有ベクトルを用いる手法を用いて，必要な次元数の空間におけるグラフ配置を求めた後，それを適切な方向から 2 次元平面に射影することで，2 次元グラフ配置を計算する．

最初に，グラフのノード間のグラフ理論的距離を用いて，Torgerson の方法を実行する．固有値$\lambda_1, \lambda_2, \ldots$を大きさの順に並べ，正の固有値の個数を d とし，$d \geq 2$ であると仮定する．従って，$\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_d > 0$ となる．これによって得られる d 次元空間における各ノード i の位置を $\boldsymbol{p}_i = (p_{i1}\ p_{i2} \ldots p_{id})$とする．

Visualization of Graphs by Using Their High-Dimensional Layouts
[†] Hiroshi Hosobe, National Institute of Informatics

[1] 準計量的多次元尺度法である数量化 IV 類によっても，固有ベクトルを用いたグラフ配置が可能であり，例えば文献[2]などで用いられている．



(a) (b) (c) (d)

図 1 (a) Torgerson の方法に基づくグラフ配置；(b) 本研究の提案手法によるグラフ配置；(c) 1 つのノードの選択によって更新されたグラフ配置；(d) 4 つのノードの選択によって更新されたグラフ配置

次に，d 次元空間におけるグラフ配置を射影する 2 次元平面として，2 つの d 次元ベクトル $\boldsymbol{e}_1 = \boldsymbol{f}_1/|\boldsymbol{f}_1|$, $\boldsymbol{e}_2 = \boldsymbol{f}_2/|\boldsymbol{f}_2|$によって張られる平面を考える．ただし，$\boldsymbol{f}_1$, $\boldsymbol{f}_2$ は，以下のように定められる d 次元ベクトルである．

$$\boldsymbol{f}_1 = \frac{1}{\sqrt{\lambda_1}}\left(\sqrt{\lambda_1}, 0, \sqrt{\lambda_3}, 0, \ldots\right),\quad \boldsymbol{f}_2 = \frac{1}{\sqrt{\lambda_2}}\left(0, \sqrt{\lambda_2}, 0, \sqrt{\lambda_4}, \ldots\right)$$

このとき，$\boldsymbol{e}_1$ と $\boldsymbol{e}_2$ は線形独立であり，直交する．

最後に，$\boldsymbol{e}_1$ と $\boldsymbol{e}_2$ を用いて，各ノードを 2 次元平面上へ射影する．これにより，ノード i の 2 次元座標は$(\boldsymbol{p}_i{\cdot}\boldsymbol{e}_1, \boldsymbol{p}_i{\cdot}\boldsymbol{e}_2)$として求められる．

3.2 2 次元グラフ配置の更新

次に，射影すべき 2 次元平面を移動することで，2 次元グラフ配置をインタラクティブに更新できるようにする方法を提案する．この方法の特徴は，高次元空間におけるグラフ配置を変更する必要がないため，極めて高速に 2 次元グラフ配置を更新できる点である．

具体的な方法として，射影すべき 2 次元平面の新しい法線ベクトルを指定する方法を示す．これは，以下のように Gram-Schmidt の直交化を用いて実現される．まず，これまでの 2 次元平面を張っていたベクトルを $\boldsymbol{e}_1$, $\boldsymbol{e}_2$, 新しい法線ベクトルを $\boldsymbol{v}$ とし，$\boldsymbol{v}$ は $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ と線形独立であると仮定する．また，これらのベクトルは正規化されているものとする．このとき，新しい 2 次元平面を張るベクトル $\boldsymbol{e}_1'$, $\boldsymbol{e}_2'$は $\boldsymbol{e}_1' = \boldsymbol{f}_1'/|\boldsymbol{f}_1'|$, $\boldsymbol{e}_2' = \boldsymbol{f}_2'/|\boldsymbol{f}_2'|$として求められる．ただし，$\boldsymbol{f}_1'$, $\boldsymbol{f}_2'$は，以下のように定められるベクトルである．

$$\boldsymbol{f}_1' = \boldsymbol{e}_1 - (\boldsymbol{e}_1 \cdot \boldsymbol{v})\boldsymbol{v},\quad \boldsymbol{f}_2' = \boldsymbol{e}_2 - (\boldsymbol{e}_2 \cdot \boldsymbol{v})\boldsymbol{v} - (\boldsymbol{e}_2 \cdot \boldsymbol{e}_1')\boldsymbol{e}_1'$$

新しい法線ベクトル $\boldsymbol{v}$ の決め方として，例えば以下の方法が考えられる．

1. ユーザに適当なノードを選択させ，そのノードの位置ベクトルを用いる．これによって，選択されたノードを中心とするようなグラフの 2 次元配置が得られる．
2. ユーザに複数のノードを選択させ，それらの位置ベクトルの和を用いる．これによって，選択されたノード群の重心を中心とするようなグラフの 2 次元配置が求められる．

4 実験

本研究で提案したグラフ配置法を Java で実装し，実験を行った．固有ベクトルの計算には行列計算パッケージ Jama を用いた．プログラムのコンパイルと実行には J2SE 1.4.0_01 を用い，実行の環境としては，1.13 GHz の Pentium III 上で動作する Windows XP を使用した．

図 1(a)は，2.2 項で述べた Torgerson の方法に基づくグラフ配置法を用いて，141 個のノードからなる一般無向グラフである AT&T グラフ ug_263 を配置した結果である．一方，図 1(b)は，本研究の提案手法を同じグラフに適用して得られた結果である．この配置の内部的な次元数は 125 であり，計算時間は，Torgerson の方法を含めて 170 ミリ秒であった．

次に，本提案手法を用いて，グラフの 2 次元配置をインタラクティブに更新した例を示す．図 1(c)は，図 1(b)の右側に見られる，多数の密集するノード群に接続されたノードを選択し，それを中心とする 2 次元グラフ配置を求めた例である．また図 1(d)は，図 1(b)の上側に見られる 4 つのノードを選択した場合の配置である．計算時間はいずれの場合も 10 ミリ秒未満であった．

5 おわりに

本研究では，固有ベクトルを用いる新しい一般無向グラフ配置法を提案した．

今後の研究として，インタラクション機能を拡充して本手法の利用性を向上することを計画している．

参考文献

[1] 林知己夫, 飽戸弘(編): 多次元尺度解析法—その有効性と問題点—, サイエンス社(1976).

[2] Koren, Y., Carmel, L. and Harel, D.: ACE: A Fast Multiscale Eigenvectors Computation for Drawing Huge Graphs, *Proc. IEEE InfoVis*, pp.137–144 (2002).

[3] Kruskal, J.B. and Seery, J.B.: Designing Network Diagrams, *Proc. 1st General Conf. on Social Graphics*, pp.22–50, U.S. Dept. of the Census (1980).