

多次元データの視覚的ブラウジング検索システム

折茂 栄美子

電気通信大学 電子工学科
〒182-8585 調布市調布ヶ丘1-5-1
TEL:(0424)43-5654
emikoo@vogue.is.uec.ac.jp

小池 英樹

電気通信大学 大学院情報システム学研究科
〒182-8585 調布市調布ヶ丘1-5-1
TEL:(0424)43-5651
koike@is.uec.ac.jp

1 はじめに

近年のネットワークの発達により、ユーザの入手できる情報が増加した。これに伴い、膨大な量の情報からユーザの必要とする情報にまで絞り込むためには、困難な作業を要する事になった。そこで、必要な情報を得るために情報検索が用いられる。一般に情報検索には、キーワード検索と、ブラウジング検索があげられる。

WWW等ではキーワード検索が多く用いられている。キーワード検索をするにはあらかじめ明確な目的が必要であるが、情報検索をするにあたり、ユーザがあらかじめ明確な目的を持っているとは限らない。この場合にはユーザが検索を進めるための適当な目的を提示する必要がある。

ブラウジング検索は情報を視覚化し、観察することで情報を絞る検索手法であるので、明確な目的がない場合でも検索を進める事ができる。そこで、本研究では、ユーザのインタラクションに従い類似したデータを提示する事で検索を進める事ができると考え、情報視覚化技術を用いた検索システム「ZASH」を構築した。

2 ブラウジング検索の関連研究

ブラウジング検索を支援するシステムは広く研究されている。DocSpace [1]では、多数の文献を2次元上に視覚的に配置する視覚的分類を用いている。これは算出される各文献の重みはインデックス内のキーワードに依存しているため、キーワードが多数になると情報構造が複雑になり、視覚化情報の把握が困難になる。

CinemaScape[2]は、映画のタイトルと批評者を表すノードが別々のウィンドウに類似性に基づいて配置され、ユーザの評価から類似性を算出し、新たに趣味に合っていると予測される映画を提示する。これは映画のタイトルまたは批評者以外からの映画の検索はキーワードを用いる形式になっているため、目的がはっきりしていない場合の検索は困難な作業になりうる。

納豆ビュー [3]は、大規模なグラフ構造の情報空間の理解を支援するために、WWWの構造を3次元空間上に視覚化する技術である。これはノード配置が外面的な属性に基づいているので、配置から得られる情報は少ない。また、大きさの違うノードや距離の離れたノード同士の属性を比較する事は難しい。

本システム ZASH では多種多様な視覚化技術を取り入れ、上記の研究における問題点を解決する事を試みた。

3 ZASHの概要

図1はZASHの実行画面である。中央の平面で選ばれたタイトルがfocusとなり、付随するデータとはエッジでつながれて表示されている。3次元空間に配置された5つの平面には以下のような機能がある。

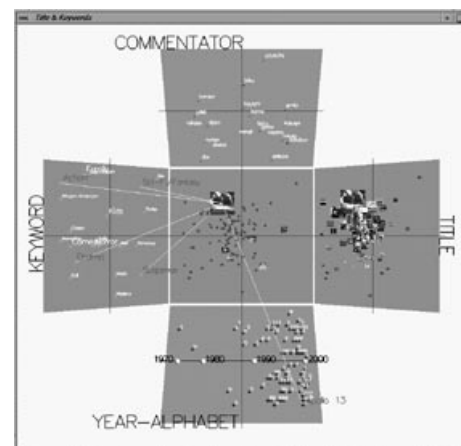


図1: ZASHの実行画面

3.1 中央の平面

3.1.1 多次元尺度法

情報には、それに付随する情報が存在する。本研究で扱うデータは映画のデータである。映画のデータには、タイトル、キーワード(ジャンル)、監督、出演者、製作年代、推薦者などの付随情報が存在する。ここではそれらを多次元データとみなし、多次元データ相互の関係から多次元尺度法を用い類似性を算出することにした。多次元尺度法とは、いくつかの項目の間の関連性を統計的に分析し、現象を要約して簡潔な表現を加える手法であり、回帰分析、判別分析、主成分分析、因子分析、グラフ解析、数量化法など多くの方法がある。本研究では、情報の互いに似たものを近くに、似ていないものを遠くに位置付けるように空間的に配置する手法である数量化法を用いた。

3.1.2 位置による類似検索

映画のタイトルを表すノードが多次元尺度法により配置されている。これを用いて類似検索を行う。似たものを近くに配置してあるので、位置で類似性が把握できる。図2では、focusの周囲にあるものが類似したデータである。

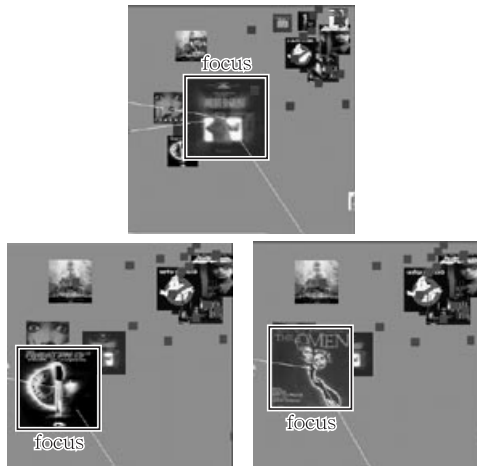
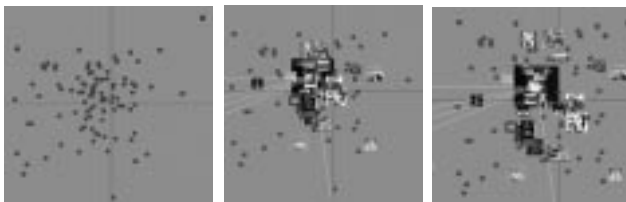


図 2: 位置による類似検索

3.1.3 Graphical Fisheye View

膨大な情報の全てを小さなディスプレイの中で一度に表示する事は難しい。Fisheye Viewとは、着目点(focus)は大きく詳細情報まで表示し、着目点以外は小さく表示する技術である。Graphical Fisheye View[4]とは、Fisheye Viewを図表などに応用する技術である。あらかじめ各ノードにその情報の重要度を与え、focusと重要度から各ノードの大きさや位置を決定する。本研究では、ユーザの選んだノードをfocusとし、各ノードの重要度からノードの大きさだけを変化させるモード [図3 (a) (b)] と、さらに位置も変化させるモード [図3 (a) (b) (c)] の2つに分けている。



(a) 標準座標 (b) ノードサイズの変化 (c) ノード位置の変化

図 3: Fisheye Viewによる表示の変化

3.2 左側の平面

映画のキーワードを表すノードが多次元尺度法により配置され、キーワードでの類似検索を行う。モード変更により、キーワードをアルファベット順に並べ換えての検索も可能である。付随するデータが共通である場合にも類似検索は可能である。ユーザの選択したノードと共通の付随データを持つノードは、エッジとノードサイズの変化によって再表示されるようになっている。

また、映画のタイトル同士のキーワードを介した類似性を知る事も可能である。例えば、ユーザがタイトルを表すノードを選択すると、それに関係するキーワード(または推薦者)

を共有しているタイトルを表すノードは大きく表示される。図4では、focusと同じキーワードを持つA,B,Cが大きく表示されている。

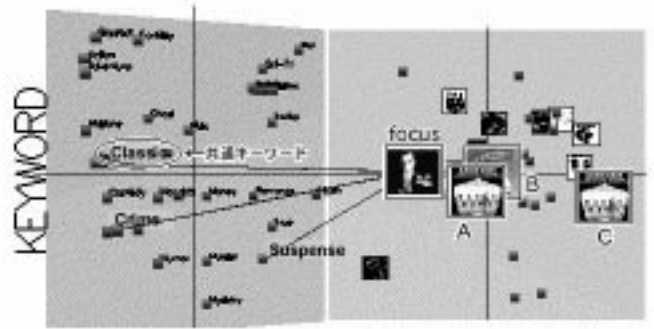


図 4: 共通キーワードによる類似検索

3.3 上側の平面

映画の批評者を表すノードが多次元尺度法により配置されている。信頼できる批評者がいる場合は、この検索を行う。

3.4 下側の平面

縦軸は映画のタイトルをアルファベット順に並べるものであり、横軸は映画の製作年順に並べるものである。キーワード検索的な要素を含むことで検索を始める時の「きっかけ」を作る事が可能になる。

3.5 右側の平面

検索の履歴を表す。映画のタイトルを表すノードが直前の検索結果を表示する。

4 まとめ

本研究では、多変量解析や Fisheye View による配置法により、類似検索が可能なシステム ZASH を構築した。これにより、目的がはっきりしていない場合にもユーザに適切な目的を与える事が可能になった。また、多次元データを3次元空間に配置することにより、情報の相互関係を崩す事なくユーザに視覚的に捉えやすいレイアウトを実現した。

参考文献

- [1] 館村純一 “DocSpace:文献空間のインタラクティブ視覚化”, 日本ソフトウェア学会 WISS'96, インタラクティブシステムとソフトウェア IV, 近代科学社, pp.11-20, 1996.
- [2] <http://www-tate.iis.u-tokyo.ac.jp/~tatemura/CinemaScape/>.
- [3] 塩澤秀和 “切り取り操作による柔軟な情報選択が出来る WWW 視覚化”, 情処研報 97-HI-72-11, pp.61-66, 1997.
- [4] Manojit Sarkar, et al, “Graphical Fisheye views of graphs”, CHI '92 Conference Proceedings, pp.83-91, 1992.