

Fish View: 文書整理支援における視点情報の活用†

3U-4

高間 康史

石塚 満*

*東京大学工学部電子情報工学科

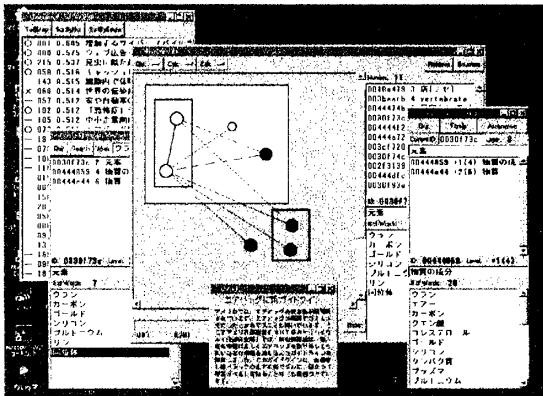


図 1: 文書整理支援システム Fish View

1 まえがき

我々は、インターネット等を通じて大量に収集された文書を熟読し、有効に活用するためには、図解を用いて文書間の関係を整理しつつ、漸進的に読み進めていくことが有効であり、この過程を計算機によって効率的に支援するためには、その時点におけるユーザの興味・視点に基づいた文書間の関係を見出すことが必要であると考えます。

この様な観点から、我々は **Fish Eye マッチング** と呼ぶ、概念体系を利用した新しい動的ベクトル生成・マッチング機構を提案し [1, 2], これを基盤技術として用いた文書整理支援システム **Fish View** を開発している (図 1)。

本稿では Fish Eye マッチングについて簡単に紹介した後、開発した文書整理支援システム Fish View について、文書整理プロセスを順に追いつながりながら紹介する。

†FISH VIEW: Utilization of Viewpoint Information for Document Ordering Support.

*Yasufumi Takama, Mitsuru Ishizuka
University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan
takama@miv.t.u-tokyo.ac.jp

2 Fish Eye マッチング

Fish Eye マッチングでは、ドキュメント中に含まれる単語をそのまま特徴として用いるのではなく、ユーザの視点において同様とみなせる単語は一つの特徴に縮退 (**Shrink**) したり、その視点に関係のある単語のみを選択 (**Magnify**) したりといった操作を加えることにより、その時点に応じた特徴ベクトルを動的に生成する。この様な特徴選択・縮退のプリミティブとして、EDR 電子化辞書中の概念体系中に存在する概念を利用する。すなわち、ある概念に属する単語の集合を **意味グループ** と呼び、特徴生成演算子の引き数として利用する。Fish Eye Matching の詳細については文献 [1, 2] にゆずり、ここでは図 2 に概念体系および特徴選択・縮退の例について示す。

また、ユーザが興味に従って分類した少数の正/負文書集合から、視点を反映した意味グループを求めるアルゴリズムを提案、検索精度に関して有効な結果が得られている [1, 2]。

3 文書整理支援ツール Fish View

図 1 は文書整理支援システム Fish View のクライアントであり、図解作成・視点情報の編集を行う Main Window を中心に、意味グループ検索用などの複数の Window から構成される。サーバプログラ

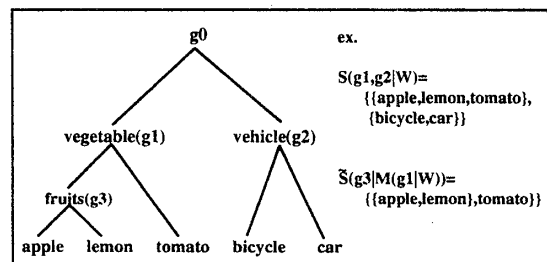
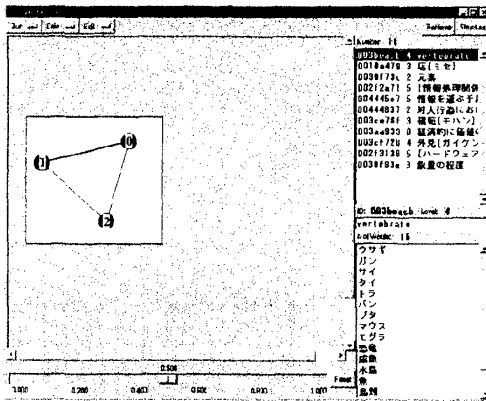
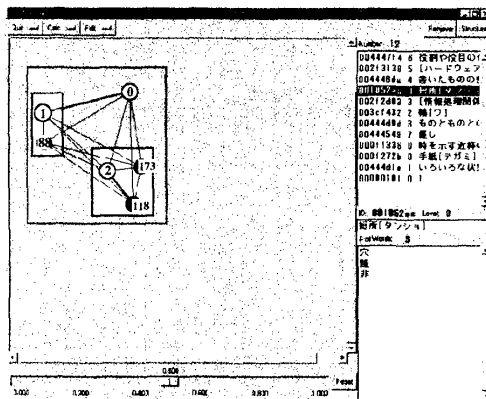


図 2: 概念体系と特徴生成の例



0: ウェブ広告を見るたびにお金が入る「サイバーゴールド」
 1: 増加するサイバースパイの異態に迫る
 2: 情報価値の高いウェブサイトを案内：ブリタニカ・インターネット・ガイド

図 3: 文書検索前の図解



118: ウェブの悪質な情報の対策を検討
 173: アレルギーに関する情報を掲載。ウェブサイト「フレディズ」
 173: コンピュータ・ウィルスの情報を提供

図 4: 文書検索後の図解

ムは UNIX 上に実装されており、EDR 電子化辞書から作成した単語・概念に関するデータベースをもとに、視点抽出や文書検索といった作業を全て行う。ここでは文書整理プロセスを順に追いながら、Fish View の機能について紹介する*。

図解作成

Document List Window から読みたい文書を選択すると、対応するノードが Main Window 内のキャンバスに出現する。このキャンバス上で、自分の視点・興味に従って関連があると思う文書をグループ化することにより図解を作成する。グループは入れ

子構造をとることができる。また、Fish Eye マッチングに基づき、関連のある文書間には自動的にリンクが張られる (図 3)。

視点抽出・編集

グループを FIX 状態にすると、システムにより視点抽出が行われ、Main Window 右側に視点情報リストが表示される。視点情報は、意味グループ毎に見出し情報や所属単語リストなどが提示され、不要な意味グループの削除、ある単語の所属する意味グループの入れ換え、新規グループの追加などの編集作業を行う事ができる。

新規文書検索

視点情報に基づいて Fish Eye マッチングを行うと、その視点に関連するグループから計算されたクエリーベクトルとの類似度に基づいて Document List Window の文書がソートされる。このリストから次に読むべき文書を選択し、「図解作成」ステップを繰り返す (図 4)。

図解の HTML 化

最終的に得られた図解を HTML 化して出力する事ができる。HTML ページは、グループ単位で、入れ子構造を生かして構成する。

4 むすび

本稿では、Fish Eye マッチングを要素技術として用いた情報整理支援システム Fish View を紹介した。実際に使用してもらったところ、良好な評価が得られている。今後は、概念体系・単語各辞書の整備および、ツールの完成度を高める事が課題である。

参考文献

- [1] 高間 康史, 石塚 満: Fish Eye マッチング: 概念体系を利用した視点抽出に基づく文書整理支援機能, 人工知能学会誌, Vol. 14, No. 1, pp. 93-101, 1999.
- [2] 高間 康史, 石塚 満: 概念体系を用いた Fish Eye マッチングによる視点を考慮した文書整理支援機能の実現, 信学技報 OFS98-17, AI98-26, pp. 37-44, 1998.

* マニュアル: http://www.miv.t.u-tokyo.ac.jp/~takama/fview_manual.html