

クロスメディア適合性フィードバック：特徴空間を横断する 適合性フィードバックとそのレシピデータ検索への応用

杉山 裕樹[†] 加藤 誠[†] 大島 裕明[†] 田中 克己[†]

[†] 京都大学大学院情報学研究科社会情報学専攻 〒606-8501 京都府京都市左京区吉田本町

E-mail: †{sugiyama,kato,ohshima,tanaka}@dl.kuis.kyoto-u.ac.jp

あらまし 本稿では、テキストや画像など様々なメディアを基にした複数の特徴量を持つオブジェクトを検索する際の、特徴空間を横断した適合性フィードバック手法を提案する。このようなオブジェクト検索のクエリとしてはそれぞれの特徴量を対象としたものが考えられるが、検索対象に関して十分な知識を持っていないユーザが全ての特徴量を適切に指定するのは非常に困難である。本研究では、ある特徴空間における正例・負例ベクトルを他空間の特徴量へと写像することで、異なる特徴空間のクエリへのフィードバックを可能とする検索モデルを提案する。

キーワード 情報検索、適合性フィードバック

1 はじめに

近年、ブロードバンドインターネットの進展に伴い、Web上ではテキストのみならず、画像や動画、音声など様々なメディアが利用されるようになった。例えばニュース記事やレシピといったデータにも文章とともに、1枚以上の写真が添えられていることも珍しくなく、このようなデータを検索する機会も増えてきている。このように、Web情報検索の対象は様々なメディアを含んだマルチメディアデータとなっている。

情報検索とは、主にユーザが知らない情報を求めるために行うものである。しかし一般的に求めたい結果を適切に得ることのできるクエリを指定することは非常に困難である。例えば、レシピデータはタイトル、カテゴリ、材料、画像のように様々な要素を含んでいる。このようなデータを検索する場合、1つの要素を指定するだけではユーザの検索意図に沿ったデータを取得できない可能性がある。なぜなら、タイトルはそのデータの作成者次第でシンプルに記述されている場合もあれば、具体的に細かく記述されている場合もあり、材料に関してもそのレシピ独自のアレンジが存在する可能性があるためである。従って、ユーザが求めているレシピを適確に取得するためには、全ての要素に対してユーザの検索意図を表すクエリを指定するのが望ましい。ところが、検索を行うユーザには料理名は分からないが見た目や材料は知っているという状況や、料理名だけしか知らないといった状況など、様々な状況が考えられる。このように、検索対象に対して十分な知識を持っていない場合、ユーザは自分の知っている範囲でしかクエリを指定することができできない。このような状況では、ユーザが入力するクエリは不完全なものである可能性が高くなってしまい、もちろんその不完全なクエリではユーザの検索意図に沿った結果を得ることは困難である。

ユーザがクエリ修正をするのを助ける手法として代表的なものに適合性フィードバック [1] がある。これは、得られた検索

結果に対してユーザが適切な結果であったか否かの正誤判定を行うというインタラクションを利用してクエリベクトルを修正するものである。しかし、従来の適合性フィードバックでは、単一の特徴量を持ったデータの検索を対象としている。そのため、ユーザが正誤判定を行った対象と同一の特徴空間に存在するクエリにしかフィードバックすることができない。

しかし、レシピデータのようにテキスト、画像など様々なメディアを含んだマルチメディアデータにはメディア毎に異なる特徴量を持つ。また、テキストという一つのメディアの中でも、文章による特徴の表現方法や、キーワードによる表現方法など様々な特徴量が存在する。このような場合、クエリの作成が困難であったのと同様に、正誤判定を行うことができる対象もユーザにとって既知の特徴量に限られてしまう。つまり、通常の適合性フィードバックでは、結局ユーザにとって既知な特徴空間のクエリしか修正することができず、未知の特徴空間のクエリベクトルは不完全なままになってしまう。

そこで本研究では、複数の特徴量を持つマルチメディアデータを検索する際に、ユーザの検索意図を適確に表現するため、各特徴量にクエリベクトルを導入し、これらの複合クエリを用いてデータの検索を行う。また、ユーザにとって既知の空間の特徴量を用いて、未知の空間のクエリベクトルを修正する、特徴空間横断型のクロスメディア適合性フィードバックを提案する。

図1に、本研究のイメージをレシピデータ検索を用いて示す。ユーザが検索を行うと、タイトル・カテゴリ・画像などの特徴量を含んだ検索結果が返ってくる。このとき得られた結果に対し、ユーザがあるカテゴリ名に関して適切か不適切かの正誤判定を行った場合、通常の適合性フィードバックではカテゴリの正誤判定の結果はカテゴリクエリにしかフィードバックされない。一方、提案手法では選択されたレシピの特徴量や類似レシピの特徴量を用いて、タイトルクエリや画像クエリといった異なる特徴空間のクエリへフィードバックを伝播させる。これに

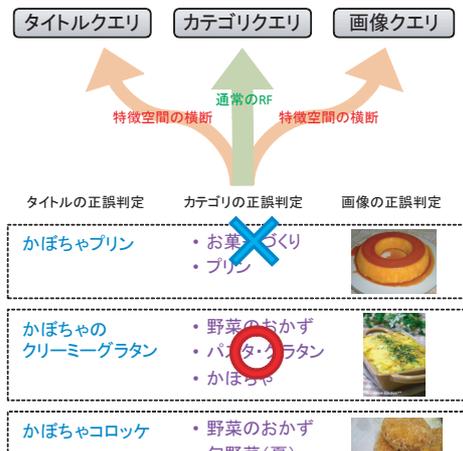


図 1 研究のイメージ

よって、ユーザが指定できなかった特徴空間のクエリベクトルの修正も同時に行うことが可能となる。

本論文の構成は以下の通りである。2 節では本研究の関連研究について述べる。3 節ではクロスメディア適合性フィードバックの検索モデルについて述べ、4 節ではアプリケーションを実装するに当たって設定した事項などについて述べる。5 節では実装したアプリケーションを用いて行った評価実験について述べ、最後に 6 節で本論文のまとめと今後の課題について述べる。

2 関連研究

2.1 Rocchio の適合性フィードバック

適合性フィードバック [1] とはベクトル空間モデルの情報検索において、適合文書や不適合文書の特徴ベクトルをもとにクエリベクトルを再定式化する手法である。Rocchio は以下の式により、クエリベクトルの修正を行っている。

$$\mathbf{q}(t+1) = \alpha \mathbf{q}(t) + \frac{\beta}{|D_r|} \sum_{\mathbf{d} \in D_r} \mathbf{d} - \frac{\gamma}{|D_{nr}|} \sum_{\mathbf{d} \in D_{nr}} \mathbf{d} \quad (1)$$

ただし、 $\mathbf{q}(t)$ は時刻 t におけるクエリベクトル、 \mathbf{d} は文書の特徴ベクトルで、 D_r と D_{nr} はそれぞれユーザが選択した適合・不適合文書の集合である。また α, β, γ は元のクエリベクトルや正のフィードバック、負のフィードバックにどの程度の重みを置くかを定める定数である。

この Rocchio の適合性フィードバックは、ベクトル空間モデルの情報検索において非常に有効であることが示されている [2]。また Rui ら [3] は画像検索において画像特徴量に関して適合性フィードバックを行うことで検索精度が向上することを示している。しかし、何れの研究も扱う特徴量は 1 種類であり、複数の特徴量をもつデータに対して適合性フィードバックを適用するには特徴空間を跨いでフィードバックを行った方がよい場合もあると考えられる。

適合性フィードバックを行う上で負のフィードバックの扱いは非常に重要となる。あまり大きくフィードバックさせすぎるとクエリベクトル中で多くの値が負の値をとってしまい、類似度の計算などに大きな影響を及ぼしてしまう。Müller ら [4] や

Wang ら [5] は適合性フィードバックを行う上で、負のフィードバックを効果的に行う手法に関する研究を行なっている。

2.2 並列検索における適合性フィードバック

Nakamura ら [6] は文書検索と画像検索や日本語検索と英語検索のように異なる検索を並列に行う際、片方の検索結果に対して行った操作を他方に伝播させ、リランキングを行う手法を提案している。また、Halpin ら [7] はハイパーテキスト検索とセマンティック検索において、相互補完しながら検索精度の改善を行う手法を提案している。本研究では、1 つの検索の中で複数の特徴量を取り扱う場合の適合性フィードバック手法を提案する。

2.3 特徴空間の変換

Lu ら [8] は画像特徴量に加えて、画像に付加されているセマンティックな情報を利用する適合性フィードバック手法を提案している。本研究とは正誤判定を画像に対してのみ行うのか、それともテキスト情報に対しても行うのかという点で異なる。

Jeon ら [9] は確率的生成モデルを用い、画像に自動でアノテーション付けを行う研究をしている。Blei ら [10] は潜在的ディリクレ配分法により、アノテーション付けされたデータに対する強力なモデルを提案している。

3 クロスメディア適合性フィードバック

本節では、特徴空間を跨いでフィードバックを伝播させるクロスメディア適合性フィードバックの検索モデルについて述べる。

3.1 検索対象のオブジェクト

ニュース記事やレシピデータなど、Web 上に存在するデータはテキスト、画像、動画など様々なメディアで構成されている。またテキストという一つのメディアの中でも、タイトルや見出しのように短い文によって表現する方法や、カテゴリやタグのようにキーワードで特徴を表現する方法もある。これらの特徴量は表現方法によって使用されやすい単語などに差があるため、異なる特徴量と考えることもできる。また、画像の特徴量に関しても色合いなどから特徴量を計算する手法や、局所特徴量を抽出するといった手法も考えられ、一つの画像から多種の特徴量を計算することが可能である。従って、ニュース記事やレシピ一つに対して、メディア数以上の特徴量を得ることもできる。

本研究では、図 2 の左側に示すように、このような検索対象となるニュース記事やレシピデータといったマルチメディアデータをオブジェクト o_k ($k = 1, \dots, n$) と呼び、各オブジェクトは m 種類の特徴量 $f_i(o_k)$ ($i = 1, \dots, m$) の組で表されているものとする。

$$o_k = (f_1(o_k), f_2(o_k), \dots, f_m(o_k)) \quad (2)$$

$$f_i(o_k) \in V_i \quad (3)$$

ただし V_i は各特徴量のベクトル空間である。例えば、特徴ベクトルとしてはタイトルを形態素解析し Term Frequency を計算したものや、画像からヒストグラムを計算したものなど様々なものが考えられるが、検索対象のオブジェクトに応じて適切なものを選択する。

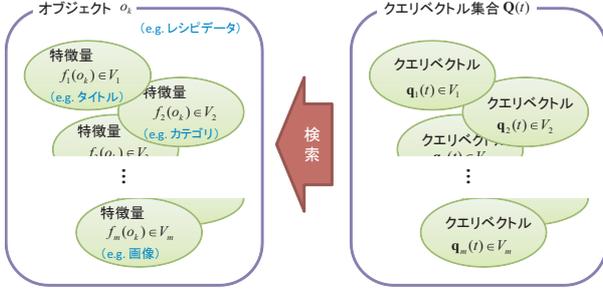


図2 オブジェクトとクエリベクトル集合

3.2 クエリベクトル

前項で定義したようにオブジェクトには m 種類の特徴量が存在する．これらの中にはユーザにとって既知な特徴量と未知な特徴量があるはずである．そのため提案する検索モデルでは，図2に示すように，ユーザがどの特徴量でも検索できるよう，各特徴空間に一つずつクエリベクトル $\mathbf{q}_i(t)$ ($i = 1, \dots, m$) を導入した．

適合性フィードバックではユーザとのインタラクションをもとにクエリベクトルを随時修正していくため，時刻 t における各特徴空間のクエリベクトル $\mathbf{q}_i(t)$ の組み合わせを全体のクエリベクトル集合 $Q(t)$ とする．つまり，

$$Q(t) = (\mathbf{q}_1(t), \mathbf{q}_2(t), \dots, \mathbf{q}_m(t)) \quad (4)$$

$$\mathbf{q}_i(t) \in V_i \quad (5)$$

となる．

3.3 クエリ修正

ユーザは提示された検索結果に対し，得られた結果が自身の検索意図に沿っているか，それとも意図には沿わないものであるかという正誤判定を行う．本検索モデルで検索対象としているオブジェクトは複数の特徴量を持ち，その特徴量の中にはユーザにとって判別しやすいものとそうでないものがあると考えられるため，ユーザはそれぞれの特徴量ごとに正誤判定を行えるようにする．

また，例えばあるテキストが与えられた場合，ユーザはそのテキスト全体に対して正誤判定を行いたいこともあれば，1つの単語に対してのみ行いたいこともある．これはある特徴量に対し，その特徴量全体に対して行うか部分特徴量に対して行うかと考えることができる．そこで，ユーザからのフィードバックの入力は特徴ベクトルに対して行われるものとし，正例と負例の入力を以下のように表記する．

- V_i^+ : 特徴空間 V_i における正解ベクトル集合
- V_i^- : 特徴空間 V_i における不正解ベクトル集合

このユーザからの入力と時刻 t におけるクエリベクトルから，時刻 $t+1$ におけるクエリベクトルの計算を行う．2節で示した Rocchio の適合性フィードバック式 (1) を用いると，次式で示すように，特徴空間 V_i の正解ベクトルは同じ特徴空間のクエリベクトルにしかフィードバックすることができない．

$$\mathbf{q}_j(t+1) = \alpha \mathbf{q}_j(t) + \frac{\beta}{|V_i^+|} \sum_{\mathbf{v}_i \in V_i^+} \mathbf{v}_i - \frac{\gamma}{|V_i^-|} \sum_{\mathbf{v}_i \in V_i^-} \mathbf{v}_i \quad (6)$$

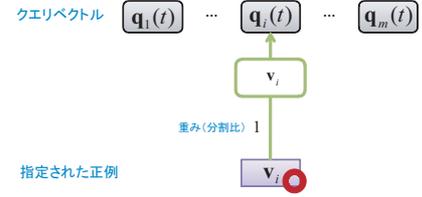


図3 従来手法のフィードバック先

本研究のクロスメディア適合性フィードバックでは， V_i の正解ベクトルを用いて特徴空間 V_j にも重み w_{ij} にもフィードバックを行う．

ところが，ベクトル空間の異なる2ベクトルは足し引きすることができないため，ある写像 M_{ij} を用いて特徴空間 V_j への変換を行う必要がある．

$$\begin{aligned} \mathbf{q}_j(t+1) &= \alpha \mathbf{q}_j(t) \\ &+ \sum_i w_{ij} \left(\frac{\beta}{|V_i^+|} \sum_{\mathbf{v}_i \in V_i^+} M_{ij}(\mathbf{v}_i) - \frac{\gamma}{|V_i^-|} \sum_{\mathbf{v}_i \in V_i^-} M_{ij}(\mathbf{v}_i) \right) \end{aligned} \quad (7)$$

ただし，

$$0 \leq w_{ij} \leq 1, \quad \sum_j w_{ij} = 1$$

$$M_{ij} : V_i \rightarrow V_j$$

α, β, γ : 元クエリベクトル，正のフィードバック，負のフィードバックの重み係数

特に w_{ij} を要素とする重み行列 \mathbf{W} が，

$$\mathbf{W} = (w_{ij}) = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \end{pmatrix} \quad (8)$$

という単位行列であり，写像 M_{ij} の内，自身の特徴空間への写像である M_{ii} が

$$M_{ii}(\mathbf{v}) = \mathbf{v} \quad (9)$$

という恒等写像である場合，このクエリ修正式は Rocchio の適合性フィードバック式 (6) と同一になる．つまり，図3のように通常の適合性フィードバックでは特徴空間 V_i に関するフィードバックの全てを同一の特徴空間のクエリベクトル $\mathbf{q}_i(t)$ に行っていたものを，図4のように w_{ij} という割合で分割して特徴空間 V_j のクエリベクトル $\mathbf{q}_j(t)$ にも伝播させるというのが本モデルのクエリ修正式が示すところである．

重み行列 \mathbf{W} と写像 M_{ij} の具体的な定義に関して，どの特徴空間に大きく伝播させるべきかという問題やどのように伝播させるべきかという問題は，扱うオブジェクトや特徴量によって適切なものが異なると考えられる．そのため，検索モデルでは一般的に記述するのみとし，実装の節にてその詳細を述べる．

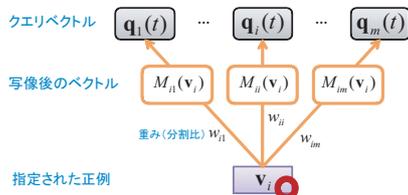


図 4 提案手法のフィードバック先

3.4 ランキング関数

あるクエリに対する検索結果の表示順位はランキング関数によって決定する。ランキング関数とは検索対象オブジェクトとクエリとの関連度を数値化するものである。

今回提案した検索モデルでは、検索対象のオブジェクトは特徴量の組み合わせで検索対象オブジェクトは各特徴空間に特徴ベクトルを持ち、クエリベクトルも各特徴空間に一つずつ存在する。そこで、各々の特徴空間でオブジェクトとクエリの関連度を計算し、それを組み合わせることでランキング計算を行う。

以下に定義するランキング関数の値が大きいオブジェクトほど上位に順位付ける。

$$\begin{aligned} \text{Rank}(o_k, \mathbf{Q}(t)) \\ = \Phi(\text{Sim}_1(f_1(o_k), \mathbf{q}_1(t)), \dots, \text{Sim}_m(f_m(o_k), \mathbf{q}_m(t))) \end{aligned} \quad (10)$$

ただし、 Sim_i はベクトル空間 V_i 上の 2 ベクトルの類似度を計算する関数であり、同一のベクトルの場合に 1, 2 ベクトルが異なるにつれ 0 に近い値をとる。また、関数 Φ は各特徴空間における類似度を組み合わせるものである。

$$\text{Sim}_i : V_i \times V_i \rightarrow [0, 1] \quad (11)$$

$$\Phi : [0, 1]^m \rightarrow [0, 1] \quad (12)$$

類似度 Sim_i としてはコサイン類似度やユークリッド距離などの値を、 $[0, 1]$ の範囲に変換する手法などが考えられる。また、類似度の組み合わせ関数 Φ としては相加平均や相乗平均をとるといった手法などが考えられる。しかし、適切な類似度や組み合わせ関数は用いる特徴量に依存するため、ここでは一般化のために上のように定義しておく。

4 レシピ検索システムの実装

4.1 検索対象のオブジェクト

Web 上には COOKPAD^(注1) や Yahoo! レシピ^(注2) のようなレシピサイトが多く存在する。今回、このようなレシピサイトに登録されているレシピデータのうち、画像やカテゴリがつけられているもの約 1 万 3 千件をテストセットとしてレシピ検索を行うアプリケーションを実装した。

4.2 特徴量と類似度

レシピに含まれるデータのうち、タイトル特徴量 f_{title} 、カテゴリ特徴量 f_{category} 、材料特徴量 $f_{\text{ingredient}}$ 、画像特徴量 f_{image}

の 4 つをレシポオブジェクトの特徴量として利用した。各データから特徴ベクトルの作成方法と、その特徴ベクトル空間における類似度の計算方法を以下に述べる。

4.2.1 タイトル特徴量

レシピのタイトルは、単語を次元としたベクトルにより特徴ベクトル化を行う。形態素解析器 MeCab^(注3) を用いて形態素解析を行い、得られた結果から記号などのストップワードを取り除いた単語集合の Term Frequency (TF) をタイトル特徴ベクトルとする。

この TF で表現された特徴ベクトルの類似度はコサイン類似度を用いて計算を行う。ただし、コサイン類似度は -1 以上 1 以下の値を取り得るので、今回使用する類似度は 0 以上 1 以下となるよう、次のように定義する。

$$\text{Sim}_{\text{title}}(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \left(1 + \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right) \in [0, 1] \quad (13)$$

ただし、 $\mathbf{a} \cdot \mathbf{b}$ は 2 ベクトル \mathbf{a} , \mathbf{b} の内積であり、 $\|\mathbf{a}\|$ と $\|\mathbf{b}\|$ はそれぞれ \mathbf{a} と \mathbf{b} のユークリッドノルムである。

4.2.2 カテゴリ特徴量

今回利用したレシピサイトではそれぞれのレシピにいくつかのカテゴリが設定されている。そこでカテゴリの特徴ベクトルとしては、カテゴリ名を次元としたベクトルにおいて、あるカテゴリが設定されていれば対応する次元の値を 1 を、そうでなければ 0 としたベクトルを用いる。

この特徴ベクトルも一種の TF として考えることができるため、タイトル特徴量と同様にコサイン類似度を用いて類似度を計算する。

$$\text{Sim}_{\text{category}}(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \left(1 + \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right) \in [0, 1] \quad (14)$$

4.2.3 材料特徴量

材料に関してもカテゴリと同様に、材料名を次元としたベクトルにおいて、ある材料が使用されていれば対応する次元の値を 1、そうでなければ 0 としたベクトルを用い、コサイン類似度を用いて類似度を計算する。

$$\text{Sim}_{\text{ingredient}}(\mathbf{a}, \mathbf{b}) = \frac{1}{2} \left(1 + \frac{\mathbf{a} \cdot \mathbf{b}}{\|\mathbf{a}\| \|\mathbf{b}\|} \right) \in [0, 1] \quad (15)$$

4.2.4 画像特徴量

画像の特徴量としては Color Coherence Vector (CCV) [11] を用いた。CCV は色ヒストグラムに色の粗密情報を加えたもので、同じ色の集まり具合を捉えることができる。

CCV の類似度は、2 ベクトル \mathbf{a} , \mathbf{b} のマンハッタン距離に基づいて計算を行う。ただし画像のサイズに対する正規化を行うため、ベクトルの各次元の値の総和、すなわち画像のピクセル数の逆数でベクトルを定数倍してからマンハッタン距離を求める。

$$\text{dist}(\mathbf{a}, \mathbf{b}) = \sum_i \left| \frac{a_i}{s_a} - \frac{b_i}{s_b} \right| \in [0, 2] \quad (16)$$

$$s_a = \sum_i a_i, \quad s_b = \sum_i b_i \quad (17)$$

(注1): <http://cookpad.com/>

(注2): <http://recipe.gourmet.yahoo.co.jp/>

(注3): <http://mecab.sourceforge.net/>

このマンハッタン距離は 0 以上 2 以下の値を取る．そこで、類似度が 0 以上 1 以下の値となるよう、下式により 2 画像の類似度 $\text{Sim}_{\text{image}}(\mathbf{a}, \mathbf{b})$ を計算する．

$$\text{Sim}_{\text{image}}(\mathbf{a}, \mathbf{b}) = 1 - \frac{\text{dist}(\mathbf{a}, \mathbf{b})}{2} \in [0, 1] \quad (18)$$

4.3 検 索

本モデルにおけるクエリは、4 つの特徴空間のクエリベクトルの組み合わせで定義されている．そのため、各クエリベクトルに関してそれぞれレシピオブジェクトの検索をし、得られた結果の和集合をランキング関数によってランキング付けを行う．その結果上位にランキングされたレシピオブジェクトを結果として返すこととした．

タイトル、カテゴリ、材料に関する検索は、クエリベクトルの中で値の大きい 3 つのキーワードを含むレシピオブジェクトを結果として返し、画像に関する検索は、画像特徴量の類似する上位 50 件を結果として返す．

また、最初の検索はいくつかのキーワードによって行うものとし、タイトル、カテゴリ、材料の 3 つの特徴ベクトル空間において、与えられたキーワードに対応する次元の値を 1 としたベクトルを初期クエリベクトルとして検索を開始する．

4.4 フィードバック操作

検索によって得られた結果のうち、ユーザは次の項目に対して正誤判定を行うことにより、フィードバック操作を行うことができる．

- 検索結果のレシピのタイトル全体
- 検索結果のレシピに付けられている各カテゴリ
- 検索結果のレシピの各材料
- 検索結果のレシピの画像

タイトルや画像が選択された場合には、その選択されたタイトル・画像の特徴ベクトルを正解ベクトル集合 $V_{\text{title}}^+, V_{\text{image}}^+$ や不正解ベクトル集合 $V_{\text{title}}^-, V_{\text{image}}^-$ に加える．一方、カテゴリや材料が選択された場合には、選択されたキーワードの次元の値のみ 1 で、その他の次元の値は 0 となるようなベクトルを正解ベクトル集合 $V_{\text{category}}^+, V_{\text{ingredient}}^+$ や不正解ベクトル集合 $V_{\text{category}}^-, V_{\text{ingredient}}^-$ に加える．

4.5 写 像

ユーザの操作によって得られた正解・不正解ベクトル集合を他の特徴空間のクエリベクトルにフィードバックするためには、その正解・不正解ベクトルをクエリベクトルと同一の空間のベクトルに写像しなければならない．

ここで、ある特徴空間 V_j に関して、 n 個のオブジェクトの特徴ベクトルを縦ベクトルとして並べた次のような行列 \mathbf{A}_j を考える．

$$\mathbf{A}_j = (f_j(o_1) \ f_j(o_2) \ \cdots \ f_j(o_n)) \quad (19)$$

この行列を用いて、特徴空間 V_i から V_j への写像 $M_{ij}(\mathbf{v})$ を定義する．

4.5.1 タイトル・画像特徴量から他空間への写像

4.4 項で定義したように、ユーザがタイトルや画像を選択し

た場合はあるレシピオブジェクトのタイトル特徴量や画像特徴量が正解・不正解ベクトル集合に入っている．

そこで、選択された特徴空間において類似するレシピオブジェクトを取得し、それらのオブジェクトの特徴ベクトルの類似度の重みつき平均を求めることで特徴ベクトルの写像を行う．また、同一の特徴空間へフィードバックを与える場合、すなわち写像を必要としない場合は、ユーザが選択したベクトルをそのまま用いることができるため、写像 M_{ij} を次式のように定義する．

$$M_{ij}(\mathbf{v}) = \begin{cases} \frac{1}{s_1} \mathbf{A}_j \begin{pmatrix} \text{Sim}(\mathbf{v}, f_i(o_1)) \\ \vdots \\ \text{Sim}(\mathbf{v}, f_i(o_n)) \end{pmatrix} & \text{if } i \neq j \\ \mathbf{v} & \text{otherwise} \end{cases} \quad (20)$$

ただし、 s_1 は類似度の総和であり、次のように定義される．

$$s_1 = \sum_{k=1}^n \text{Sim}(\mathbf{v}, f_i(o_k)) \quad (21)$$

4.5.2 カテゴリ・材料特徴量から他空間への変換

一方で、ユーザがカテゴリや材料を選択した場合、正解・不正解ベクトル集合にはある 1 つの次元の値のみ 1 であるようなベクトルが入っている．そこで、同じ次元に値をもつレシピオブジェクトを取得し、それらのオブジェクトの平均特徴量を求めることで特徴ベクトルの写像を行う．同じ次元に値をもつオブジェクトは、行列 \mathbf{A}_j の転置行列を利用することで取得できる．また、先程と同様に同一の特徴空間へのフィードバックでは写像を行う必要がないため、写像 M_{ij} は次式のように行う．

$$M_{ij}(\mathbf{v}) = \begin{cases} \frac{1}{s_2} \mathbf{A}_j \mathbf{A}_i^T \mathbf{v} & \text{if } i \neq j \\ \mathbf{v} & \text{otherwise} \end{cases} \quad (22)$$

ただし、 s_2 は選択されたベクトルと同じ次元に値を持つオブジェクトの数である．

4.6 重み係数

今回、フィードバックを伝播させる重み行列を定めるにあたって、次の 4 つの行列を考えた．

(1) 相関係数

1 つ目の行列にはテストセットのレシピデータにおける類似度の相関行列を利用した．テストセット中の 2 レシピの組み合わせに対して各特徴量での類似度を計算し、類似度 X と類似度 Y の相関係数を次式によって求める．

$$w'_{ij} = \frac{\sum_k (x_k - \bar{x})(y_k - \bar{y})}{\sqrt{\sum_k (x_k - \bar{x})^2} \sqrt{\sum_k (y_k - \bar{y})^2}} \quad (23)$$

ただし、 \bar{x} や \bar{y} は X, Y の平均である．

こうして得られた相関行列を表 1 に示す．この相関行列に対して、各行で値の総和が 1 となるように正規化を行った行列 $\mathbf{W} = (w_{ij})$ を伝播の重み行列とした．

表 1 類似度の相関係数 w'_{ij}

$i \setminus j$	title	category	ingredient	image
title	1	0.183	0.103	0.024
category	0.183	1	0.062	0.043
ingredient	0.103	0.062	1	0.035
image	0.024	0.043	0.035	1

$$w_{ij} = \frac{w'_{ij}}{\sum_j w'_{ij}} \quad (24)$$

これは、相関の高い特徴空間には伝播を強く行うべきであるという考えに基づいた重みの設定となっている。

(2) 相関係数の逆数

次に、表 1 の相関行列の全ての成分の逆数を取り、先程と同様に正規化を行った行列を重み行列とした。

$$w_{ij} = \frac{(w'_{ij})^{-1}}{\sum_j (w'_{ij})^{-1}} \quad (25)$$

これは、相関の低い特徴空間にフィードバックを行うことで、今までのクエリでは不完全であった検索の絞り込みを強めるという考えに基づいた重みの設定となっている。

(3) 均等な行列

3 つ目は、全ての成分の値が 0.25 の行列である。これは、どの特徴空間にも均等にフィードバックを行うものであり、重みに差をつけた場合と結果がどのように変わるか比較を行う。

(4) 単位行列

最後は、対角成分の値が 1 で、それ以外の成分の値が全て 0 である単位行列である。これは前項で定義した同一の特徴ベクトル空間への写像が恒等写像である $M_{ij}(\mathbf{v})$ と合わせて、全く他空間への伝播を行わないフィードバックであり、従来の適合性フィードバック手法と考えることができる。

また、元クエリ、正のフィードバック、負のフィードバックは今回同様に扱おうとし、 α, β, γ は全て 1 とした。

4.7 ランキング関数

今回実装したアプリケーションでは検索を行う際に、特徴量ごとに分けてレシピオブジェクトを取得しその和集合をとっているため、1 つの特徴空間でしか類似していないオブジェクトも多く取得してしまっている。そこで、ランキング関数における類似度の組み合わせ関数 Φ は、全ての特徴空間において類似するレシピデータのみを上位にランキングさせるよう、各類似度の積で定義する。

$$\begin{aligned} \Phi(\text{Sim}_{\text{title}}, \text{Sim}_{\text{category}}, \text{Sim}_{\text{ingredient}}, \text{Sim}_{\text{image}}) \\ = \text{Sim}_{\text{title}} \cdot \text{Sim}_{\text{category}} \cdot \text{Sim}_{\text{ingredient}} \cdot \text{Sim}_{\text{image}} \\ \in [0, 1] \end{aligned} \quad (26)$$

各類似度は全て 0 以上 1 以下の値を取るよう定義されている。そのため、この組み合わせ関数 Φ を用いて定義されてい



図 5 アプリケーションの実行例

るランキング関数 $\text{Rank}(o_k, \mathbf{Q}(t))$ も 0 以上 1 以下の値を取ることになる。

このランキング関数の値の上位 30 件をこのアプリケーションの検索結果とする。

4.8 実行例

以上の設定に基づいて作成したアプリケーションの実行例を図 5 に示す。この例では「カレー」というクエリから、タイ風のグリーンカレーを絞り込む検索を行なっている。実装したアプリケーションの実行手順は以下のようになっている。

- (1) 初期クエリ「カレー」をテキストボックスに入力して最初の検索を行い、検索結果を得る。
- (2) 出力された結果のタイトル・カテゴリ・材料・画像からグリーンカレーという検索意図に合っているものと合っていないものを選択する。この例ではグリーンカレーに使用される「ココナッツミルク」という材料を正解とし、グリーンカレーらしからぬ茶色のカレーの画像を不正解とした。
- (3) 正誤判定をもとに自動的に修正されたクエリで再検索を行い、新たな検索結果を得る。
- (4) 手順 (2) に戻る。

5 実験

5.1 実験の概要

実装したアプリケーションを用いて検索精度がどの程度向上したかの評価を行った。ユーザが最初に指定するクエリは曖昧なキーワードから成るものとし、得られた検索結果に対し正誤判定を行って再検索することで、特定のレシピを検索するといったようなタスクを設定した。

与えられたタスクに対し、タスクの目標となるレシピに合っている特徴を正解、合っていない特徴を不正解とする正誤判定を行う。このフィードバックを行わなかった場合と 1~3 回行った場合とでタスクの検索意図に沿った結果がどの程度得られたかという検索精度を測った。

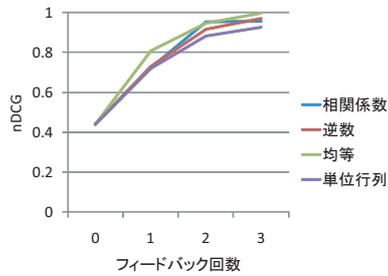


図 6 タイトルに対してフィードバックを行った場合の nDCG の推移

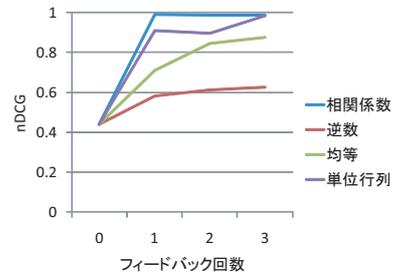


図 8 材料に対してフィードバックを行った場合の nDCG の推移

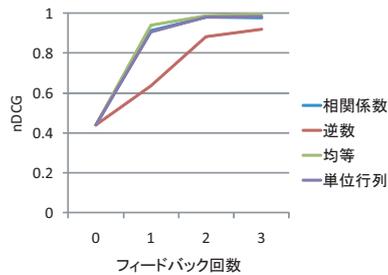


図 7 カテゴリに対してフィードバックを行った場合の nDCG の推移

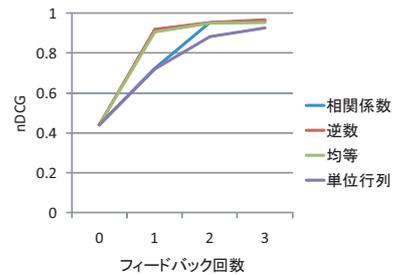


図 9 画像に対してフィードバックを行った場合の nDCG の推移

ユーザの正誤判定はタイトルのみに対して行われた場合、カテゴリのみの場合、材料のみの場合、画像のみの場合の4通りに分けて実験を行った。

5.2 タスク

実験を行うにあたって、初期クエリ「パスタ」からホワイトソースのクリームパスタの検索を行うといったように、曖昧な初期クエリからある特定のレシピを検索するタスクを設定した。

5.3 ベースライン

本手法に対する比較手法として、特徴空間の伝播を行わないフィードバックをベースラインとした。今回実装したアプリケーション上では、4.6項の(4)で示した単位行列を用いた重み行列がベースライン手法のフィードバックを表している。この重み行列を用いた場合と、(1)~(3)の重み行列を用いた場合とで検索精度の比較を行う。

5.4 評価尺度

検索精度の評価尺度として順位付き検索の評価指標である nDCG を用いた。検索結果として得られた 30 件のレシピに対し、ユーザの検索意図に合っているものには 10 点を、少し合っているものには 5 点を、全く合っていないものには 0 点を与え、30 件目のレシピの nDCG の値をベースラインと比較した。nDCG はランキング付き検索結果が理想の出力順序にどの程度近いかを表す評価尺度であり、理想の出力順通りの場合に 1、理想から離れるにつれ 0 に近い値をとる。

5.5 実験結果

実験を行った結果、nDCG の値がどのように変化したかを図 6~9 に示す。このグラフは横軸がフィードバック回数を、縦軸が検索結果 30 件目の nDCG 値を示している。

5.6 考察

カテゴリに正誤判定を行った場合は、そのユーザが選択したカテゴリ自体がそもそも求めているレシピを絞り込むのに十分効果的であった。そのため、同一の特徴空間へ大きくフィードバックを行ったものはかなり良い検索精度を得た。

材料に正誤判定を与えた場合は、タイトルへ伝播させて探索空間を広げることができたことや、画像への伝播が精度の低下を引き起こしていたことから、相関係数を用いた重み行列が最も精度を向上させた。

タイトルに正誤判定を行った場合は、どの特徴空間へも均等に伝播させたときの方が僅かながら良い結果が得られたが、どの重み行列でもあまり変わらない結果となった。

一方で、画像に正誤判定を行った場合は、積極的に他の特徴空間へフィードバックを行った方が効果的であり、他では良い結果が得られていなかった相関係数の逆数を用いた重み行列が、1 回のフィードバックで精度を大きく向上させることが可能であった。

6 まとめと今後の課題

本稿ではテキスト・画像・動画など複数の特徴量を持つマルチメディアデータの検索を行う際に、特徴ベクトルの写像を行うことで特徴空間を横断した適合性フィードバックを行う検索モデルを提案した。また、提案検索モデルの適用例としてレシピデータを取り上げ、アプリケーションの実装を行った。特徴ベクトルの写像の際、類似オブジェクトの特徴量を用い重み付き平均を取ることで、ユーザの正誤判定の意図に合わせたフィードバックを行った。フィードバックの伝播に関してはテストセットのデータにおける類似度の相関係数を利用し、相関

の高い特徴空間へ大きく伝播を行うものや、関連の低い特徴空間へ伝播を行うものなどの比較を行った。

そして、実装したアプリケーションを用いて提案モデルの評価実験を行った結果、伝播の重みは、大きな効果が得られる場合やあまり効果が得られない場合など検索精度に大きく影響を与えることが分かった。ベースライン手法に比べてあまり良い結果が得られなかった場合の問題点として、重み行列が適切に設定できていなかった点や、特徴ベクトルの写像によってノイズとして現れてしまった細かな値の取扱いといった点が挙げられる。今後、どの特徴空間からどの特徴空間へ伝播を行った場合に精度をより向上させることができるかを検討し、フィードバックの振り分け方を定める重み行列としてどのような値が適切なのか、また、特徴ベクトルの写像としてより良い定義はどのようなものなのかを明らかにしていく予定である。

謝辞 本研究の一部は、京都大学 GCOE プログラム「知識循環社会のための情報学教育研究拠点」、および、文部科学省科学研究費補助金特定領域研究「情報爆発時代に向けた新しい IT 基盤技術の研究」、計画研究「情報爆発時代に対応するコンテンツ融合と操作環境融合に関する研究」(研究代表者: 田中克己, A01-00-02, 課題番号: 18049041) によるものです。ここに記して謝意を表します。

文 献

- [1] J. Rocchio: “Relevance Feedback in Information Retrieval”, The SMART retrieval system: experiments in automatic document processing, pp. 313–323 (1971).
- [2] G. Salton and C. Buckley: “Improving Retrieval Performance by Relevance Feedback”, Readings in information retrieval, pp. 355–364 (1997).
- [3] Y. Rui and T. Huang: “A Novel Relevance Feedback Technique in Image Retrieval”, Proceedings of the 7th ACM international conference on Multimedia (Part 2), pp. 67–70 (1999).
- [4] H. Müller, W. Müller, S. Marchand-Maillet, T. Pun and D. Squire: “Strategies for positive and negative relevance feedback in image retrieval”, *wecwis*, p. 5043 (2000).
- [5] X. Wang, H. Fang and C. Zhai: “A study of methods for negative relevance feedback”, Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval, pp. 219–226 (2008).
- [6] S. Nakamura, T. Yamamoto and K. Tanaka: “SyncRerank: Reranking Multi Search Results Based on Vertical and Horizontal Propagation of User Intention”, Proceedings of the 9th international conference on Web Information Systems Engineering, pp. 120–135 (2008).
- [7] H. Halpin and V. Lavrenko: “Relevance Feedback Between Hypertext and Semantic Search”, Proceedings of Semantic Search Workshop at the World Wide Web Conference, pp. 31–40 (2009).
- [8] Y. Lu, C. Hu, X. Zhu, H. Zhang and Q. Yang: “A Unified Framework for Semantics and Feature Based Relevance Feedback in Image Retrieval Systems”, Proceedings of the 8th ACM international conference on Multimedia, pp. 31–37 (2000).
- [9] J. Jeon, V. Lavrenko and R. Manmatha: “Automatic image annotation and retrieval using cross-media relevance models”, Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion

retrieval, pp. 119–126 (2003).

- [10] D. Blei and M. Jordan: “Modeling annotated data”, Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval, pp. 127–134 (2003).
- [11] G. Pass, R. Zabih and J. Miller: “Comparing Images Using Color Coherence Vectors”, Proceedings of the 4th ACM international conference on Multimedia, pp. 65–73 (1997).