

専門家の知識による生地画像検索の提案 Textile Image Retrieval Using Specialists' Knowledge

古川陽介*¹ 佐藤達也*¹ 高木友博*¹
Yosuke Furukawa Tatsuya Sato Tomohiro Takagi

*¹ 明治大学 理工学研究科 基礎理工学専攻
Department of Computer Science, Meiji University

Abstract: We propose a knowledge based image retrieval system. Traditional image retrieval systems measure degrees of similarity based on Color, Texture, and Shape. However, when a specialist and a designer in clothing retrieve textile images, they would rather use pattern, fabric, and usage based image retrieval systems than pixels based image retrieval systems. We propose the image retrieval suit users by estimating attentions and purposes from specialists' knowledge.

1. はじめに

我々は画像検索の機能を用いて、服飾に用いられる生地を検索、参照することで、生地の専門家やデザインの専門家の作業の効率を上げる事を目的としている。

画像の検索方法には、大きく分けてテキストベースとコンテンツベースの二つが存在する。テキストベースは、画像に付与されているキーワードから画像を検索する方法であり、コンテンツベースとは、色・形・テクスチャなどの画像の内容に基づいて検索する方法である[Castelli 01]。

テキストベースによる検索は、画像に付与されているキーワードと、ユーザのクエリーとのギャップが存在する。つまり、画像に不適切なキーワードが付与されていると検索精度に影響する。コンテンツベースによる検索は、色・形・テクスチャによる原始的な画像の特徴と人間が知覚する画像の特徴の間に意味的なギャップが存在し、ユーザの目的に合わない画像を検索してしまうという問題がある。

また、コンテンツベースによる検索は色・形・テクスチャの尺度を求め、最適な組み合わせによって検索を行っている。つまり、ユーザの観点を考慮せず、画像の特徴量同士の類似度で検索を行っている。本来、観点は変化するものであり、単一の観点で検索できるものではない。同一の画像を見ていても、観点が違えば、その類似度は変化する。

テキストベース・コンテンツベースの問題点を改善するために、両方の特性を考慮した検索を実現する必要がある。さらに、単一の観点で検索を行うコンテンツベースの問題を解決するために、観点を定義した画像の知識を構築する必要がある。

そこで、我々は、説明文付きの生地の画像を検索対象として、服飾生地の専門家の知識[鈴木 04]を用いて、ユーザの変化する注意・観点を推定し、ユーザの望む生地の画像を検索する事を提案する。

我々のシステムの検索手法として、まずユーザによる画像のフィードバックを元に、システムは適合画像の説明文と色・テクスチャの画像情報を得る。次に、説明文・画像情報を元に知識から観点を抽出する。観点による検索を行った後、ユーザが選択した画像に類似している順に並び替える。この手法により、上記のテキストベース・コンテンツベースの問題を改善する。

1.1 類似研究

類似研究は、Chu らの医療の知識に基づく画像検索が挙げられる[Chu 97]。この研究では、医療用の脳内 MR 画像を画像コレクションとして、腫瘍の大きさや位置を階層構造の知識として表現している。それによって、単純な類似度を求めるだけでなく、注目している部位を指定した上での検索を可能にしている。しかし、ここで言う知識は、腫瘍の大きさや位置、さらに脳内の前頭葉や側頭室といった部位との外見的位置関係を定義したものであり、ユーザの注意や観点を推定するためのものではない。

これに対し我々のシステムは、ユーザの注意・観点を推定するための知識を構築している。そのため、ユーザへの負担を軽減する事ができる。

2. 専門家の知識

生地の専門家による協力を得て、独自の知識体系を構築した。我々の構築した知識とは、生地の持つ属性を“模様”、“織り方”、“使用用途”の3つに絞り、それぞれの種類をノードとして、階層構造に分類わけした。図1は、本システムで使用した生地の知識体系の一部を表したものである。これらの知識を用い、ユーザが生地のどの部分に注目しているのかを推定している。知識構築に使用したノード数は、約 50 として、画像検索に必要な知識のみを用いた。

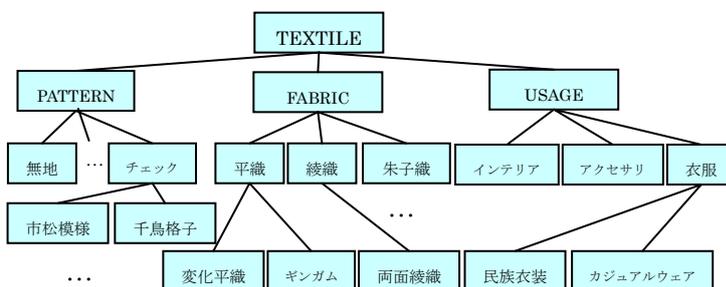


図1 生地の知識体系

2.1 知識の活用

ユーザのフィードバックによって観点の推定と、推定後の検索に知識を利用している。ユーザの選択した画像から、共通の特徴を抽出する(Common Feature Extraction)。共通の特徴から、知識を用いて、注目部位にフォーカスされた単語を得る(Knowledge)。さらに、注目している観点によって、検索方式を変化させる事で、ユーザに特化した検索をしている(Integrated Feature)。図2の知識の活用概念図は知識の活用の流れを表したものである。

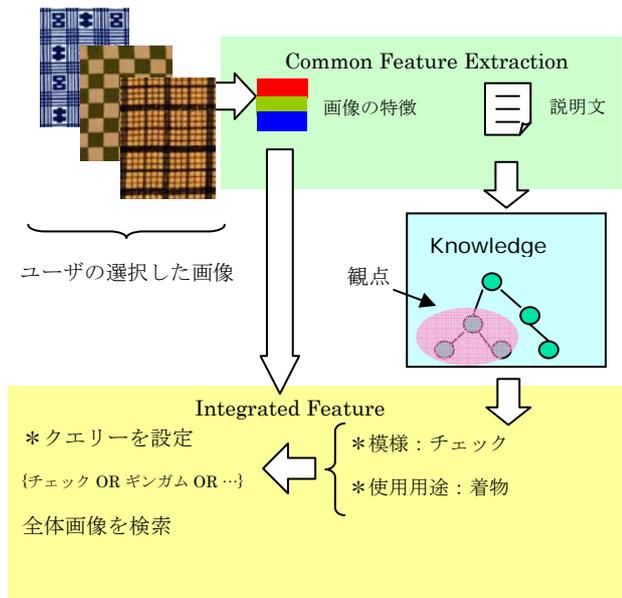


図2 知識の活用概念図

3. システム概要

構築した検索システムの概要を図3に示す。

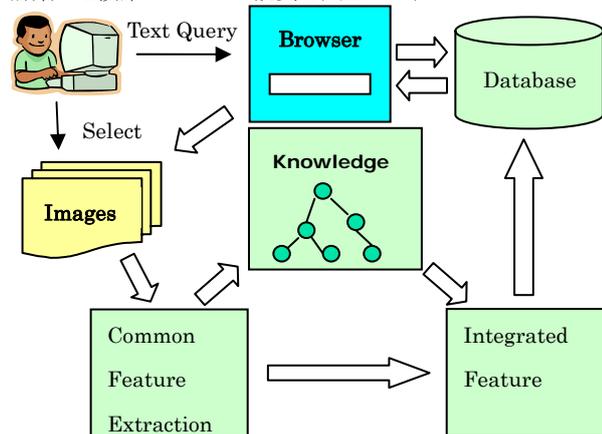


図3 システム概要図

ユーザは、二種類のインターフェースから画像を検索することができる。Text Query によって、説明文に書かれている単語から画像を検索するインターフェースと、画像を選択することで類似している画像を検索するインターフェースである。

両方のインターフェースを併用して用いる事で、ユーザの要求する画像を効果的に検索することができる。Text Query による検索で、ユーザの要求する画像を大まかに検索する事が出来る。画像を選択する事でユーザの詳細な要求を満たす事ができると考えられる。

3.1 Common Feature Extraction

Common Feature Extraction モジュールでは、フィードバック時に、ユーザが選択した画像から、共通の特徴を取り出す作業を行っている。本システムで用いている特徴とは、画像の色・テクスチャの特徴量と説明文の事である。

色の共通特徴は、画像同士の色の平均を求めることで実現している。テクスチャの共通特徴は方向・解像度による2値分類を行い、同一の分類結果になった方向・解像度を特徴量として用いている。つまり、画像の模様・織り方が類似した方向・解像度を用いている。

3.2 Knowledge

Common Feature Extraction モジュールにて、取得したキーワードを知識上のノードにマッピングし、それぞれのノード同士と共通ノードがあった場合、共通の観点として、抽出している。

たとえば、単語として、“市松模様”と“千鳥格子”は、共通の単語ではないが、二つの単語は“チェック”という共通ノードを持つ。そのため、観点として、“チェック”が出力される。

3.3 Integrated Feature

Integrated Feature モジュールは、Knowledge モジュールで出力された多くの観点から画像情報により、ユーザの観点を決定する。

(1) 観点の認識

全体画像のテクスチャの類似度と、拡大画像のテクスチャの類似度を測り、それぞれの類似度から、模様・織り方・使用用途のどれに注目しているのかを決定付けている。

テクスチャは模様や織り方のような一定のパターンを周波数として解析する手法である。そのため、テクスチャの類似度と知識を組み合わせることで観点を決定している。

全体画像のテクスチャの類似度が高く、かつ“模様”のノード(チェック、ストライプ等)が知識上の共通ノードに含まれていた場合、模様に注目していると判断する。

拡大画像についても同じように、テクスチャの類似度が高く、かつ“織り方”(平織、綾織等)が共通ノードに含まれていた場合、織り方に注目していると判断する。

全体画像・拡大画像のどちらも類似しておらず、“使用用途”が共通していた場合、使用用途に注目していると判断している。

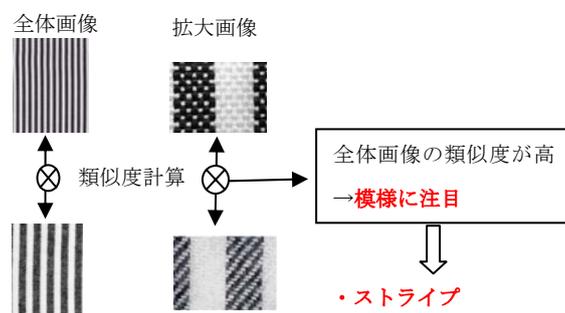


図4 Integrated Feature モジュールの例

図4の例では、全体画像の類似度が高かった場合、模様に注目している。

(2) テキストクエリの決定方法

観点の決定を行った後、クエリを決定する。知識を用いる事で、ユーザの観点到合わせたクエリを設定できる。知識上で観点をマッピングし、下位のノードへ拡張する。図5にその例を示す。

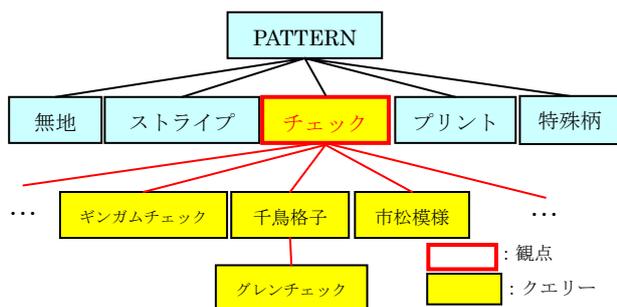


図5:クエリーを決定する例

チェックを観点とした時の例を示した。チェックという観点を認識した後、下位ノードである“ギンガムチェック”・“千鳥格子”といった単語にクエリーを拡張している。これにより説明文の単語が千鳥格子・ギンガムチェックだったとしても、“チェック”として検索することができる。

(3) コンテンツベースの検索方法の決定

次に、コンテンツベースの検索方法を決定する。観点が“模様”の場合、全体画像に対して類似度を求める。観点が“織り方”の場合、拡大した情報が重要になるため拡大画像に対して類似度を求める。観点が“使用用途”の場合、全体・拡大ともに類似度を求める。求めた類似度を元に、ランキングを行う。これにより、ユーザの観点に沿った画像が上位に出力される。

3.4 類似画像検索で用いた尺度

類似画像検索を行うために、我々が用いた尺度について説明する。本システムでは、色とテキストの二つの尺度を用いた。

(1) Earth Mover's Distance

色の類似尺度は、Earth Mover's Distance (EMD)と呼ばれる尺度を用いた[Rubner 97] [谷田川 02]。この手法を用いた理由は、従来の手法より人間の感覚に近い類似尺度として知られている事と画像内の局所的な色の特徴を用いる事が可能なためである。

(2) Wavelet 変換

テキストの特徴抽出方法として、Wavelet 変換による多重解像度解析を用いた。この手法を用いた理由は、画像を複数の解像度において縦横対角方向の高周波成分と低周波成分に分解する事ができるので、画像を巨視的、微視的に分析することが可能になるためである。この手法により、ユーザに合わせた検索が可能であると考えられる[中野 99][榊原 95]。

4. 実験

本システムは服飾のデザイン・制作を目的とするユーザが使用対象となる。そのため、デザイン・制作に関わるようなシナリオを用意して検証を行う必要がある。今回は一般的なユーザにもわかりやすいシナリオを用意して実験を行った。知識を用いたシステムと知識を用いていないシステムの精度比較を行い、本システムの有効性を検証する。

4.1 実験概要

画像データは専門家から提供された生地画像 53 枚を用いた。説明文は、参考文献[鈴木 04]を参考に作成した。シナリオを 3 種類用意し、3 人のユーザに対して実験を行った。

シナリオは、以下のように設定した。

- シナリオ 1. “青い着物を作る。”
- シナリオ 2. “ストライプのシャツを作る。”
- シナリオ 3. “チェックのシャツを作る。”

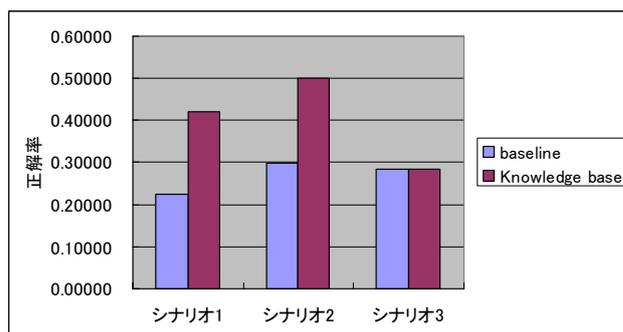
2~3 枚の画像を選択し、検索された画像集合の中で、ユーザの目的に適合する画像の割合を“正解率”とする。この“正解率”を、評価尺度として扱っていく。

4.2 Baseline となるシステム

Knowledge モジュールで行われている、ユーザの観点認識を行わないシステムを Baseline とした。ユーザの選択した画像の説明文に書かれている共通単語をクエリーとして検索するシステムを構築し、同様のテストを行った。

4.3 実験結果

グラフ 1 は、それぞれのシナリオにおける正解率を表している。シナリオ 1, シナリオ 2 の正解率は baseline よりも向上しているが、シナリオ 3 の正解率は baseline とまったく同じ結果になった。



グラフ 1: 正解率

それぞれのシナリオにおいて、システムがどの程度、正確にユーザの意図を反映しているかを出力し、表にまとめた。

表1:各シナリオの観点

	知識上の観点	画像情報の観点	知識と画像情報を組み合わせた観点
シナリオ 1: “青い着物を作る。”	民族衣装, 着物, 日本, 歌舞伎, インド, アフリカ	全体的な色	知識と画像情報を組み合わせた観点 使用用途・色に注目
シナリオ 2: “ストライプのシャツを作る。”	ストライプ, カジュアルウェア, シャツ, ブラウス, スーツ	全体的なテキスト	模様(ストライプ)に注目
シナリオ 3: “チェックのシャツを作る。”	チェック, 平織, カジュアルウェア, シャツ, ブラウス, スーツ	全体的な色とテキスト	模様(チェック)・色に注目

表中、知識上の観点は、専門家の知識を用いて共通ノードを取得したものである。画像情報の観点とは、色・テキストなどの原始的な観点を取得したものである。知識と画像情報を組み合わせた観点とは、知識上の観点と画像情報の観点を組み合わせて、知識上のどの部分に注目しているのかを最終的に決定している観点である。

表2:テキストクエリー比較

	Baseline のクエリー	知識を用いたクエリー
シナリオ 1: “青い着物を作る。”	民族衣装 OR 平織	民族衣装 OR 着物 OR 日本 OR 歌舞伎 OR インド OR アフリカ
シナリオ 2: “ストライプのシャツを作る。”	ストライプ OR シャツ	ストライプ OR ブロックストライプ OR ダイアゴナルストライプ OR ...
シナリオ 3: “チェックのシャツを作る。”	チェック	チェック OR 千鳥格子 OR ギンガムチェック OR 市松模様 OR ...

表 1 の知識と画像情報を組み合わせた観点を見ると、どのシナリオの観点もユーザの視点を反映しているといえる。しかし、表 2 を参照すると、シナリオ 1 とシナリオ 2 は、Baseline に正解率を低下させてしまうような単語(平織, シャツなど)がクエリーとして入っていた。そのため、Baseline の正解率が低下してしまったと考えられる。

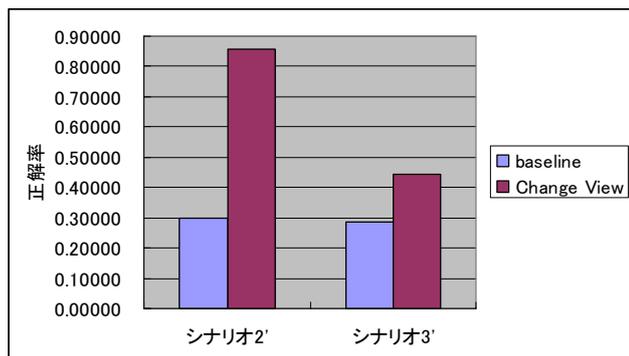
グラフ1において、シナリオ 3 の結果だけ、Baseline とまったく同じ結果になってしまった。表2を参照すると、Baseline の単語は“チェック”となっている。これは、表記ゆれに対応するために、“チェック”という単語を説明文中に多く用いたので、拡張した結果と、Baseline の検索結果が同一になってしまったと考えられる。これらの結果から、以下の事が結論と言える。

- 観点をを用いる事で、クエリーを効果的に設定できる。
- フィードバックによってある程度観点を推定する事が可能である。

4.4 手動で設定した結果

シナリオ2, シナリオ3は、知識上で、模様と使用用途の観点に注目しているが、組み合わせた観点には、模様のみが出力されている。現在、テクスチャの類似度のみで観点を決定しているため、競合してしまい、模様の出力を優先する仕様になっているためである。

観点を手動で変化させることで、知識をより有効に用いる事が可能になる。シナリオ2, シナリオ3に対して、手動で観点を“模様・使用用途”に変化させて検証を行った。結果は以下のようになった。



グラフ2: 観点を变化させた実験の結果

結果から以下の事が言える。

- 手動で変化させる機能によって、システムをさらに有効活用することができる。

5. まとめ

知識を用いた画像検索のシステムを構築し、簡単な実験を行い、システムの利用方法を示した。その結果、Baseline よりもさらに精度の向上を示すことができた。観点認識システムについては、ユーザの観点をある程度反映する事ができるという結果を示した。また、手動で観点を变化させる機能と並行して用いれば、よりユーザの検索要求に合った検索が可能であることを示した。

6. 課題

今回、一般的なユーザにも理解できるようなシナリオでしか実験をしていない。そのため、専門家を交えた実験を行う必要がある。また、画像数が 53 枚ということで、少数の画像数でしか実験をしていないこともあり、未だ有効性は測りきれっていない。

今後の課題として、生地画像検索のみではなく、自然画像などにも合わせて検索できるようにする必要がある。さらに、観点の推定を確率的なアプローチなどを用いて、より精度の向上を目指す必要がある。

7. 謝辞

生地画像データの提供、生地の専門家としての知見からの有意義なご指摘をいただいた杉野服飾大学の鈴木美和子教授に感謝の意を表します。

参考文献

- [鈴木 04] 鈴木美和子, 窪田英男, 徳武正人: アパレルの素材の基本, 織研新聞社, 2004.
- [Chu 97] W. Chu, C. Hsu, A. Cardenas, R. Taira: Knowledge-Based Image Retrieval with Spatial and Temporal Constructs, 1997.
- [Castelli 01] V.Castelli, L. Bergman: Image Databases Search and Retrieval of Digital Imagery, WILEY, 2001.
- [Rubner 97] Y.Rubner, L. Guibas, C.Tomasi: The earth mover's distance, multi-dimensional scaling, and Color-based Image Retrieval, ARPA Image Understanding Workshop, 1997.
- [谷田川 02] 谷田川英治: 色と構図に基づく画像検索ブラウザの提案, 東京大学大学院修士論文, 2002.
- [中野 99] 中野宏毅, 山本鎮男, 吉田靖夫: ウェーブレットによる信号処理と画像処理, 共立出版株式会社, 1999.
- [榊原 95] 榊原進: 数理学ウェーブレットビギナーズガイド, 東京電気大学出版局, 1995.