

# 潜在的意味索引付けとルールを統合した顔注釈システムについて

川井 優司<sup>†</sup> 伊藤 秀昭<sup>†</sup> 奥水 大和<sup>†</sup>

中京大学<sup>†</sup>

## 1. はじめに

印象語に基づいた検索方法の開発が求められている。印象語は画像から受ける印象を表現する単語である。例えば、画像データベースにおいて、感性に基づく画像検索を実現するためには印象語による検索を可能とする必要がある[4]。また、この機能を実現するためには、事前に画像に印象語が付与されている必要がある。

我々の研究室では、顔画像に対してキーワードを付与する顔注釈付けシステムの開発を進めてきた[2][3]。このシステムでは潜在的意味索引付け[1]を用いて潜在的意味空間と呼ぶ空間を構成する[3]。潜在的意味空間を用いて顔に付与するキーワードを得る。しかし、この方法は得られるキーワードの数が少なく、精度、再現率が十分ではない。連想規則と決定木を従来の方法と統合することで得られるキーワードの数を増加させ、精度、再現率を向上できると考えた。本論文では潜在的意味索引付けとルールを統合した顔注釈システムと、その実験結果について述べる。

## 2. 顔記述

個々の顔記述は、顔画像、顔画像の視覚特徴、キーワードの 3 つから成る。視覚特徴は、顔画像の顔部品の大きさや長さに相当する 24 ヶ所の顔部品の距離を計測した数値で表わされている。キーワードは顔部品の印象的な大きさや長さを表す印象語である。例えば、大きい目、小さい口、長い眉などがある。約 240 件の顔画像に対して顔部品の距離を計測して、印象語を付与した。印象語として 41 種のキーワードを得ている。

視覚特徴とキーワードを用いて個々の顔を表現する顔記述ベクトル[3]を構成する。顔記述ベクトルの集まりは顔記述行列となる。

## 3. 潜在的意味空間

顔記述行列に潜在的意味索引付けを用いて潜在的意味空間と呼ぶ空間を構成する。本研究では 3 つの異なる潜在的意味空間を構成している。それぞれ視覚特徴と顔から成る視覚空間、顔と視覚特徴およびキーワードから成る統合空間、顔とキーワードから成る記号空間である。

潜在的意味空間を用いて顔にキーワードを付与するには、付与対象となる顔を用いて空間に対して重心検索[2]と呼ぶ問い合わせ処理を施す。最初に、重心検索では視覚空間で問い合わせの顔によく似た顔を検索する。次に、似た顔の集まりを用いて統合空間で問い合わせ処理を行いキーワードを得る。最後に、得られたキーワードを用

いて記号空間で問い合わせ処理を行う。記号空間を用いて問い合わせ処理を行うことで、顔に対してキーワード間の共起関係を反映するキーワードの付与を行うことができる。以下に統合空間、記号空間の重心検索で得られたキーワードの例を示す。

[統合空間で得られたキーワードの例]

大きい顔、長い耳、大きい鼻

[記号空間で得られたキーワードの例]

大きい顔、長い耳、大きい鼻、大きい口

記号空間での問い合わせ処理により、統合空間で得られたキーワードと関連するキーワード“大きい口”が得られた。

## 4. 連想規則

顔記述を用いて連想規則[5]を構成した。支持度は A/B、確信度は C/D と表わされ、A は 1 つまたは同時に付与されているキーワードの組の総数、B は訓練集合の顔の数、C は条件部および結論部のキーワードが付与されている顔の数、D は条件部のキーワードが付与されている顔の数である。

以下に構成されたルールの例を示す。

[連想規則で得られたルール例]

大きい鼻 大きい口 → 大きい顔

確信度 45.7% 支持度 18.5%

## 5. 決定木

顔記述の視覚特徴を用いて決定木を構築するために、視覚特徴の数値を離散化した。離散化では分割点を定めて、高位を large、低位を small、それ以外を middle となるように数値を記号化する。離散化した視覚特徴を用いて決定木を構築する。なお、決定木の構築には情報利得[5]および誤り率を利用する。誤り率は A/R と表され、A は木のノードに含まれるキーワードを付与されていない顔の数、R は木のノードに含まれる顔の総数である。

事前に誤り率をしきい値として指定して、木のすべてのノードの誤り率がしきい値以下になるまで木の成長を進める。

決定木の葉がキーワードを付与するクラスに相当する場合、木の根からその葉までの視覚特徴の組み合わせをルールとして得る。以下に構成されたルールの例を示す。

[キーワード“長い鼻”の決定木から構成されたルール例]

(鼻の縦の長さ,large)^(口の横の長さ,large)^(鼻の横の長さ,small)^(下唇の厚さ,middle)

ルールの視覚特徴の組み合わせを全て満たす顔には、キーワードを付与する。

## 6. 潜在的意味索引付けとルールの統合

[3]では第 3 節で述べた潜在的意味空間のみを用いて顔

A Face Annotation System based on Integration of Latent Semantic Spaces and Rules

<sup>†</sup> Yuji Kawai, Hideaki Ito, Hiroyasu Koshimizu: Graduate School of Computer and Cognitive Sciences, Chukyo University

にキーワードを付与していた。この方法では、得られるキーワードの数が少なく再現率が低い。

これを解決するために第4節および第5節で述べた連想規則および決定木から得られるルールを従来の方法に統合することで精度、再現率が向上できると考えた。連想規則を潜在的意味空間で得られたキーワードに適用してキーワードを展開する。大量のキーワードを人が確認することは労力を要するので、展開されたキーワードの出力順位を決定木のルールを用いて変更する。訓練集合の顔記述に付与されているキーワードの平均個数が8.77個であるので、得られたキーワードの内、上位9個を得られるキーワードとする。図1にこの顔注釈付け方法の流れを示す。

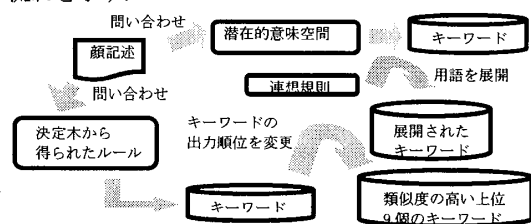


図1. システムの処理の流れ

潜在的意味空間では問い合わせベクトルとキーワードベクトル間の余弦類似度を、決定木では決定木の正当率を問い合わせに対するキーワードの類似度とする。正当率は「1-得られているルールの誤り率の最大値」とした。連想規則は、連想規則の条件部に含まれるキーワードの類似度の内、最大の類似度に適用する連想規則の確信度を乗じて類似度とする。

キーワードの妥当性を考慮するために、決定木を用いて出力順位を変更する。まず、問い合わせの顔に対して決定木から得られるルールを適用してキーワードを得る。次に、連想規則によって展開されたキーワードと決定木のルールから得られたキーワードとを比較する。もし同一のキーワードが双方に含まれていた場合、キーワードの類似度を比較して高いほうをキーワードの類似度として、上位に出力する。

## 7. 実験

問い合わせ実験には30件の顔画像を用いた。なお、実験に用いた顔画像には評価のために、事前にキーワードを付与している。

実験に用いた方法はIからVの5つの方法である。

- I. 潜在的意味空間のみを用いた方法
- II. 潜在的意味空間で得られたキーワードを連想規則のルールを用いて展開する方法
- III. 決定木のみを用いた方法
- IV. 決定木のルールから得られたキーワードを連想規則のルールを用いて展開する方法
- V. 潜在的意味空間、連想規則、決定木を用いた方法(上位9個のキーワード)

Iは[3]で用いている方法である。Vは第6節で提案した方法である。IIは潜在的意味空間で得られたキーワードに連想規則を適用してキーワードを展開した方法である。IIIは決定木のルールを用いてキーワードを得る方法である。IVは決定木のルールから得られたキーワードに連想規則を適用してキーワードを展開した方法である。

IからVのそれぞれの方法で得られるキーワードの精

度、再現率を調べた。精度はA/B、再現率はA/Cと表わされ、ここでAは事前に顔に付与されたキーワードと一致する問い合わせで得られたキーワードの数、Bは問い合わせで得られたキーワードの数、Cは事前に顔に付与されたキーワードの数である。

潜在的意味空間の次元を定めるために特異値の累積寄与率を用いた。特異値の累積寄与率が80%を超えた次元を用いる。問い合わせベクトルと空間の要素である顔イメージやキーワードのベクトルとの類似度を判定するためのしきい値は視覚空間では $10^\circ$ 、統合空間では $80^\circ$ 、記号空間では $70^\circ$ である。連想規則に用いる支持度、確信度のしきい値はそれぞれ10%、40%である。決定木に用いる誤り率のしきい値は20%である。実験結果を図2に、平均精度、平均再現率を表1に示す。

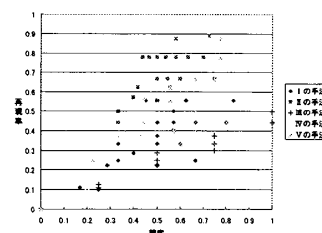


図2. 精度、再現率

表1. 平均精度、平均再現率

	平均精度	平均再現率
Iの方法	0.52	0.36
IIの方法	0.53	0.68
IIIの方法	0.69	0.31
IVの方法	0.63	0.35
Vの方法	0.57	0.57

IIの方法およびVの方法が精度、再現率が高い。このことから、潜在的意味空間で得られたキーワードに連想規則を適用してキーワードを展開する方法は顔注釈に有効であると思われる。

## 8. まとめ

本論文では潜在的意味索引付けとルールを統合した顔注釈システムと、その実験結果について述べた。今後の課題はVの方法における類似度の変更方法を改良して、さらに精度、再現率を向上することである。

## 謝辞

本システムで用いた顔画像データは(財)ソフピアジャパン研究開発グループ地域結合型共同研究推進室から使用許諾を受けたものです。

## 参考文献

- [1]Michael W. Berry, Susan T. Dumais, Todd A. Letsche: Computational Methods for Intelligent Information Access, Proc. IEEE/ACM SC Conference 1995.
- [2]Hideaki Ito, Yuji Kawai, Hiroyasu Koshimizui: Face Image Annotation based on Latent Semantic Space and Rules, Proc. KES 2008.
- [3]川井, 伊藤, 奥水: 3つの潜在的意味空間を用いた顔イメージに対するキーワード付与, FIT 2008.
- [4]Michael S. Lew, et al: Content-Based Multimedia Information Retrieval: State of the Art and Challenges, ACM TOMCCAP, vol.2, No1, 2006.
- [5]元田, 津本, 山口, 沼尾: データマイニングの基礎, オーム社, 2006.