

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 Face Image Annotation using Three Latent Semantic Spaces

川井優司
Yuji Kawai

伊藤秀昭
Hideaki Ito

奥水大和
Hiroyasu Koshimizu

1. はじめに

顔イメージを感性に基づくキーワードを用いて適切に検索するためには、そのイメージに対して適切なキーワードが付与されている必要がある。しかし、大量のイメージに対して人が個々にキーワードを付与することは負担が大きい。我々は、この問題を解決するために潜在的意味索引付け[1]を用いたシステムの開発をこれまで進めてきた[4]。これまで開発を進めてきた方法では、潜在的意味空間という空間を構成する。我々が開発を進めてきたシステムでは、顔イメージから得られる顔部品の大きさや長さの数値で表される視覚特徴と事前に付与したキーワードとに基づき、視覚特徴のみを用いた視覚空間、視覚特徴とキーワードを統合した統合空間の2つの空間を構成していた[3]。しかし、新たにキーワードのみより成る空間を構成することにより、キーワード間の共起関係をキーワードの付与に反映することができると考えられる。本論文では、3つの異なる潜在的意味空間を用いて顔イメージに対してキーワードを付与する方法について述べる。

2. 顔記述

一つの顔記述は、顔イメージ、視覚特徴、キーワードの3つから成る。視覚特徴は、顔イメージの顔部品の大きさや長さに相当する24ヶ所の距離を計測した数値である。顔部品における計測箇所を以下に示す。

[顔部品における計測箇所]

瞳の間,左目の瞳の大きさ,右目の瞳の大きさ,目の間,左目の長さ,右目の長さ,左目の高さ,右目の高さ,鼻の縦の長さ,鼻の横の長さ,鼻と上唇の間,口の横の長さ,口の高さ,上唇の厚さ,下唇の厚さ,左眉の長さ,左眉の高さ,右眉の長さ,右眉の高さ,顔の幅,顔の高さ,顔の長さ,顎と瞳の間,顎の長さ。

キーワードは顔イメージを事前に人が見て付与した顔部品の印象的な大きさや、長さ、広さ等を表現するための印象語である。例えば、大きい口、長い耳、広い額などがある。本研究で事前に付与したキーワードは43種類であった。顔記述の視覚特徴とキーワードに対して、潜在的意味索引付けを適用するために、視覚特徴とキーワードから成る顔記述ベクトルを構成した。顔記述ベクトルは視覚特徴部とキーワード部に分けられる。視覚特徴部は視覚特徴の数値が記入されるが、キーワード部はそのキーワードが顔イメージに付与されていれば1、付与されていなければ0が記入される。顔記述ベクトル F_i の記述形式を図1に示す。本研究では、約250件の顔記述から顔記述ベクトルを構成した。顔記述ベクトルの集まりを顔記述行列として潜在的意味空間を構成するために利用する。

中京大学情報科学研究科

$$F_i^t = (\underbrace{k_{i1}, k_{i2}, k_{i3}, \dots, k_{i43}}_{\text{キーワード部}}, \underbrace{v_{i1}, v_{i2}, \dots, v_{i24}}_{\text{視覚特徴部}})$$

図1. 顔記述ベクトルの記述形式

3. 潜在的意味空間

3.1 特異値分解

顔記述行列を F とする。 F は潜在的意味索引付けにおいて実行される特異値分解により、 $F=USV^t$ の3つの行列に分解される[1]。 U は属性行列、 S は特異値行列、 V^t は顔行列である。顔記述行列の属性はキーワードおよび視覚特徴である。特異値分解によって得られた USV^t の3つの行列を用いて潜在的意味空間と呼ぶ空間を構成する。空間を構成するときに、用いる特異値の数を制限することで、空間の次元を縮小することができる。用いる特異値の数を k とすると、顔記述行列 F は $F=USV^t \approx U_k S_k V_k^t$ で近似され、次元は k である。

3.2 3つの潜在的意味空間

特異値分解を用いて潜在的意味空間を構成するとき、顔記述ベクトルの構成に用いる要素を選択することによって、異なる3つの潜在的意味空間を構成することができる。顔記述ベクトルの視覚特徴部だけを空間の構成に用いれば、視覚空間と呼ぶ潜在的意味空間が構成される。また、キーワード部のみを用いれば記号空間と呼ぶ潜在的意味空間が構成される。さらに、視覚特徴部とキーワード部の両方を用いれば視覚特徴とキーワードが統合された統合空間と呼ぶ潜在的意味空間が構成される。

3.3 キーワード付与

顔イメージにキーワードを付与するためには、潜在的意味空間を用いた問い合わせ処理を行う。問い合わせは図1に示した顔記述ベクトル F_i と同じ形式である。キーワードを付与したい顔イメージの視覚特徴が問い合わせベクトルの視覚特徴部に記入される。また、問い合わせベクトルのキーワード部はすべて0が記入される。問い合わせベクトルを q とすると、 q は潜在的意味空間を構成するために用いた属性行列と特異値行列から $q^t = q^t U_k S_k^{-1}$ と変換される。 q^t は q の転置行列であり S_k^{-1} は特異値行列の逆行列である。また、 k は行列の次元である。 q^t を用いてそれぞれの潜在的意味空間に対して問い合わせ処理を行うことで、顔イメージに付与するキーワードが得られる。問い合わせでは、 q^t 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は、余弦距離を用いて定義している。また、類似度を判定するためのしきい値には角度を用いている。

3. 4 潜在的意味空間の組み合わせ

3つの異なった潜在的意味空間を用いたキーワード付与の方法において、本研究では以下のi~iiiに示した潜在的意味空間の組み合わせをキーワード付与に用いた。

- i. 統合空間のみを用いる
 - ii. 視覚空間と統合空間を用いる
 - iii. 視覚空間、統合空間、および記号空間を用いる
- iは統合空間に対して、第3.3節で示した問い合わせベクトルを用いて問い合わせを行う方法である。統合空間はキーワードと視覚特徴とが混在しているため、問い合わせの結果として顔イメージに付与するキーワードが得られる。iiでは視覚空間で問い合わせベクトルと似た顔を検索する。似た顔が検索されると、統合空間で似た顔の重心ベクトルが作成される。重心ベクトルを用いた統合空間における問い合わせによって、キーワードが得られる[3]。iiiはiiの方法に加えて、統合空間から得られたキーワードを用いて、記号空間でキーワード群の重心ベクトルを作成して、記号空間の問い合わせに用いる方法である。それぞれの方法の概要を図2に表す。

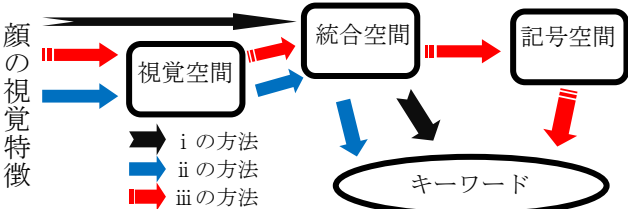


図2. 潜在的意味空間を用いたキーワード付与の概要

4. 問い合わせ実験

第3.4節で述べた3つの方法を用いて問い合わせ実験を行った。それぞれ10件の問い合わせイメージを用いて得られるキーワードの精度と再現率を調べた。10件の問い合わせ顔イメージには事前に、評価を行うために目視によってキーワードを付与した。精度と再現率は式4-1、4-2により求めた。

$$\text{精度} = \frac{\text{事前に顔イメージに付与したキーワード}}{\text{検索で得られたすべてのキーワード}} \quad (\text{式 4-1})$$

$$\text{再現率} = \frac{\text{検索で得られたすべてのキーワード}}{\text{事前に顔イメージに付与したキーワード}} \quad (\text{式 4-2})$$

また、問い合わせにおいて、潜在的意味空間の特異値の累積寄与率を考慮して、問い合わせに用いる次元を変えた。それぞれ、全ての次元、特異値の累積寄与率が0.9を超えた次元、0.8を超えた次元、表示に用いられる3次元の4種である。これら4種の次元を用いて、i~iiiの方法で問い合わせ実験を行った。実験では、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を10°から90°まで変えた。そして、4種の方法のキーワード付与の平均精度と平均再現率を調べた。平均精度と平均再現率を図3および図4に示す。ただし、しきい値を変更したとき、キーワードが求められなければ平均再現率、平均精度の計算には含まないことにした。

実験結果では、平均精度はiの方法が低い次元では高い。しかし次元が高くなると精度が低下する。他の2つの方法

は、次元が高くなるにつれ、精度が向上した。平均再現率は、実験を行った4種の次元におけるすべての結果において、iiiの方法が最も高い。また、iiの方法と比べて、iiiの方法が実験を行った4種の次元におけるすべての結果において、平均精度と平均再現率については高い。このことから、顔イメージに対するキーワードの付与には、iiの方法よりも、iiiの方法が適していると見受けられる。

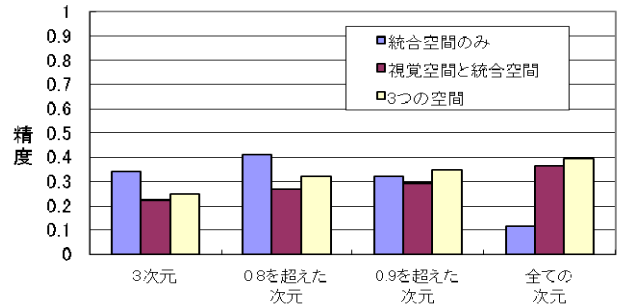


図3. 平均精度 結果グラフ

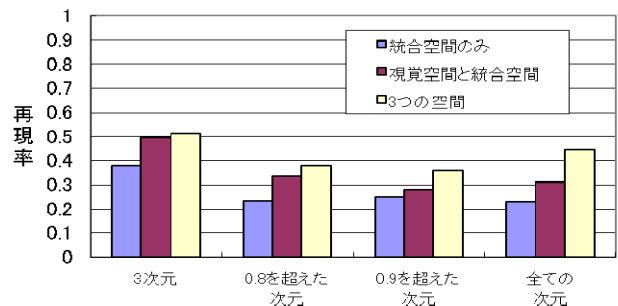


図4. 平均再現率 結果グラフ

5. まとめ

本論文では、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の付与について述べた。しかし、現段階において、十分な精度と再現率を得られてはいないと考えている。3つの潜在的意味空間を用いた方法において、精度と再現率を高めることが今後の課題である。

謝辞

本システムで用いた顔画像データは(財)ソフト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, (to appear).
 [3]Hideaki Ito, Hiroyasu Koshimizui: Face Image Retrieval and Annotation based on Two Latent Semantic Spaces in FIARS, Proc. 8th IEEE International Symposium on Multimedia 2006.
 [4]川井,伊藤,興水: 顔イメージに対するルールに基づくキーワード付与の検討,第12回日本顔学会大会(フォーラム顔学),2007.