

D-021

決定木を用いた顔データのキーワード付けに関する検討

On Face Image Annotation using Decision Trees

川井 優司
Yuji Kawai三澤 直樹
Naoki Misawa伊藤 秀昭
Hideaki Ito奥水 大和
Hiroyasu Koshimizu

1. はじめに

画像を有効に利用するためには、画像に対して画像を表現する適切なキーワードが付与されている必要がある。これまでに、画像に対して概念的なキーワードを付与する試みが行われており、画像中の対象や部分画像に対するラベル付けが試みられてきた[4]。

我々は顔画像から得られる顔部品の大きさや長さの数値で表される数値特徴とキーワードが事前に付与されている画像に基づき、付与対象となる顔画像に対して印象に基づくキーワードを付与する試みを進めてきた[1]。数値特徴は画像の視覚特徴に相当し、キーワードは画像の記号特徴に相当する。キーワード付与のためには、視覚特徴と記号特徴を統合する必要がある。このために、2つの特徴を統合的に扱うことが可能な隠れ意味空間[2]を用いてキーワード付与システムの開発を進めてきた[1]。しかし、この空間では2つの特徴が混在しているので、視覚特徴と記号特徴の関連が不明瞭である。この問題を解決するために視覚特徴と記号特徴の関連を、決定木[3]を用いて記述することを試みた。決定木は記号特徴に関連付けられる視覚特徴の組み合わせを表現する。

本論文では、顔画像に対して適切なキーワードを付与するために決定木を構築することを検討する。

2. 顔特徴の記述

1つの顔記述は部位データと、付与されたキーワードから成る。部位データは顔部品の大きさや長さである視覚特徴である。決定木を構築するために約250枚の顔画像に対して部位データを計測して、同時にキーワードを付与した。キーワードは、顔部品の印象的な大きさ、長さ、広さを表現するための語彙である。例えば、目が大きい、目が小さい、目が丸いなどがある。我々の実験ではキーワードには43種類がある。顔画像の記述例を示す(図1.a~c)。図1.(b)および(c)における最初の6桁の数字は顔画像IDである。各顔部品の部位データは顔の24箇所の距離を計測して得た。部位データの計測箇所は以下の通りである。

[部位データの計測箇所]

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。

3. 決定木とルール



(a). ID015218の顔画像

(b). キーワード

015218,短い髪,細身顔,丸い鼻

(c). 部位データ

015218:6.2,1.2,1.2,3.5,3.0,2.9,0.9,0.9,4.7,3.5,1.8,4.0,1.65,0.75,
0.9,3.6,0.8,3.3,0.85,13.4,23.6,16.2,11.6,3.45

図1.顔画像と記述例

3.1 決定木の作成

顔画像*i*における部位データに相当する属性*j*の値 v_{ij} は、 $v_{ij}^* = (v_{ij} - \bar{v}_j) / \sigma_j$ により正規化される。 \bar{v}_j および σ_j^2 は部位データ*j*の平均および分散である。次に、正規化された値に基づき、分割点を定めて、その値によって定められた区間の値を記号化した。各々の部位データに関して25%の対象に高位、50%の対象に中位、25%の対象に低位であることを示す記号が割り当てられている。それぞれ、'a', 'b' および 'c' と表されている。また、決定木の作成対象とするキーワードを顔画像が付与されているかを調べて、対象とするキーワードが付与されている画像には○、付与されていない画像には×が付与されている。図1(a)に示した顔画像における部位データの離散化後の顔画像記述例を以下に示す。この例ではキーワード“丸い鼻”を決定木の作成対象としており、図1(a)の顔画像にはキーワード“丸い鼻”が付与されている(図1(b))。

例. 図1(c)の離散化後の顔記述

015218:b,b,a,b,a,b,b,b,b,b,c,b,c,b,b,b,b,b,c,b,○

上記のように作成された顔記述をトレーニングデータとして、各々のキーワードに対して決定木を構築した。ただし顔画像のIDは決定木の作成には用いられていない。決定木の作成には情報利得[3]を利用している。情報利得を各属性毎に計算して、値が最大になる属性を選択して決定木のノードとする。属性により顔画像データがすべて同一のクラスに分けられるまで木の成長を続ける。決定木を作成した後、どのような視覚特徴を持つ顔画像がキーワードを得るかを表現するルールを構成した。

3.2 ルール記述

得られたルールは条件部と結論部から成る。条件部の要

素は部位データの名前とその値から成り、複数の要素がANDで結合されている。結論部はキーワードを付与するか否かである。得られたルールの例を示す。括弧内は離散化したデータに対応した大きさや長さを表す記号である。

例。「離れ目」を付与する視覚特徴のルール
 瞳の間(a)∧左目の長さ(b)∧鼻の横の長さ(a)∧右目の瞳の大きさ(b)∧左目の瞳の大きさ(b)→○

43個のキーワードに対して作成した決定木から、約800個のルールが得られた。個々のルールの条件部の要素数は平均4つであった。また、ルールの作成に利用された各顔部品の頻度を図2に示す。キーワードの付与の判断には、目や目の周りの大きさや長さが重要な役割を果たしていると考えられる。

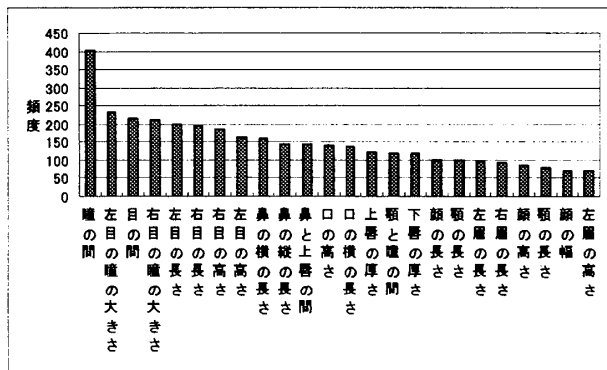


図2. ルールに用いられた部位の頻度

3.3 問い合わせ処理

新たにキーワード付与対象となる顔画像は問い合わせに相当する。ルールを用いて顔画像にキーワードを付与するために、問い合わせに使用する顔画像の部位データを第3.1節に示した方法により離散化した。離散化した部位データが得られたルールの条件部を満足すればキーワードを付与する。

4. 隠れ意味索引付けによるキーワード付与

これまでに、隠れ意味検索を利用したキーワード付与システムの開発を進めてきた[1]。個々の顔記述はキーワードと部位データから成るベクトルで表現されており、ベクトルの集まりを顔記述行列という。図1(a)の顔記述ベクトルを図3に示す。

$$A^T = \begin{bmatrix} \text{【キーワード部】} & \text{【部位データ部】} \\ 1, 1, 1, 0, \dots, & 6.2, 1.2, 1.2, \dots \end{bmatrix}$$

短	細	丸	大	瞳	右	左
い	身	い	き	の	目	目
髪	顔	鼻	い	間	の	の
		目			大	大
					き	き
					さ	さ

図3. 図1(a)の顔記述ベクトル

顔記述行列は隠れ意味索引付けにより分解され、2つの隠れ意味空間が構築される。まず顔部品の視覚特徴のみを

利用した空間が作成される。問い合わせの視覚特徴を基に類似する顔画像が検索される。次に、顔画像のキーワードも含めた空間が構築される。ここでは、先の検索で得られたいくつかの顔記述ベクトルの重心ベクトルを求めて、キーワード検索のための問い合わせとする。このような検索を通してキーワードを得る。

5. ルールと隠れ意味空間による実験

隠れ意味空間手法と、決定木のルールから得たキーワードの精度と再現率を求めた。再現率は事前に付与されたキーワードがどれだけ得られたキーワードに含まれているのか、精度は得られたキーワードに事前に付与されていたキーワードと同じキーワードがどれだけ含まれているのかを示す比率である。

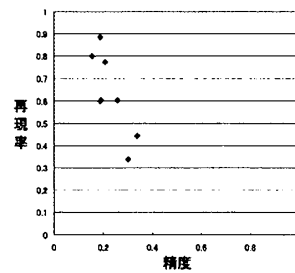


図4. 隠れ意味空間を用いた結果

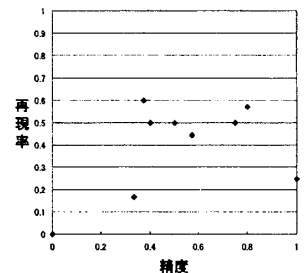


図5. 決定木を用いた結果

10件の顔画像を対象としてキーワードを付与する実験を行った。図4と図5に隠れ意味空間を用いたとき、および決定木を用いたときの精度と再現率を示す。実験結果では、隠れ意味空間を用いた手法では再現率が高く、決定木を用いた手法では精度が高い。

6. まとめ

本論文では、決定木を用いた顔画像に対するキーワード付けについて述べた。現在、決定木を構築するために必要な枝狩り機構の改良を進めている。また、隠れ意味空間手法と決定木手法を統合して、より適切なキーワード付けを行う手法を開発す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。また、本研究の一部は、日比科学技術研究助成金を受けて実施されたものです。

参考文献

[1] H.Ito, H.Koshimizu: Face Image Retrieval and Annotation based on Two Latent Semantic Spaces in FIARS, Proc. 8th IEEE International Symposium on Multimedia, 2006.
 [2] M.W.Berry, S.T.Dumais, T.Letsche: Computational Methods for Intelligent Information Access. Supercomputing, Proc of the IEEE/ACM SC Conference, 1995.
 [3] 元田,津本,山口,沼尾: データマイニングの基礎,オーム社,2006.
 [4] X.Zhou, L.Chen, J.Ye, Q.Zhang, B.Shi: Automatic Image Semantic Annotation Based on Image Keyword Document Model. CIVR 2005, Leow et al.(Eds.), LNCS 3568, 2005.