

画像における主観的解釈の差のモデル化 - 画像データベースへの応用 -

柴田 滝也
産業技術総合研究所
知能システム研究部門

加藤 俊一
中央大学理工学部
経営システム工学科

本研究では、主成分分析と重回帰分析によるオフィス空間画像における主観的解釈(形容詞)の推定法と提案した。この手法で、個人差はあるがおよそ80%の正答率を得ることができた。また、ユーザによって同じ画像の解釈が異なること、また、対象あるいはカテゴリーによって、解釈が変化することを示した。よって、画像におけるユーザの解釈を理解するアルゴリズムとして、任意の画像を、まず、自動的に対象カテゴリーに分類し、そして、そのカテゴリー内の主観的解釈を推定する平均あるいは個人モデルにより、形容詞のインデックスを自動的に張る仕組みを提案した。

Modeling subjective response to depicted space -application for image database-

Tatsuya Shibata
National Institute of Advanced
Industrial Science and Technology

Toshikazu Kato
Faculty of Science and Engineering,
Chuo University

We show a context-based indexing method that: (1) a key image is selected, (2) the image is categorized into a category which has a model of the relationship by the image categorizing method, and (3) adjectives are estimated by an adjective-text indexing method. The context-based method is based on the concept that the relationship between images and adjectives changes due to not only individual experience but also image categories because interpretation for an adjective varies in individual experience and image categories. However, visual functions seem less varied. So, we hypothesize that parameters for the visual functions such as color, direction and frequent pattern can be fixed while the relationship between adjectives and images is not fixed.

1. はじめに

インターネットはユーザにテキスト、画像、動画、音楽などのマルチメディアデータを提供する一方、データ洪水によりユーザが欲しい情報を短時間に検索するのは困難となっている。その問題として、マルチメディアデータの形式がそれぞれ異なり、異なる情報を言葉などの単一のメディアで検索することが困難であることが挙げられる。また、ユーザ、データ対象によって、同じ言葉でも解釈が異なったり、言葉と言葉との関係が変化する。例えば、同じ画像でも経験・知識によってユーザの対象への文脈、あるいは、解釈が異なり、その解釈がテキストとしてインデックス化可能なら、ユーザ、データ対象に文脈の観点から対応でき、さらに、異なるマルチメディアデータをテキストをもとに検索可能になる。

そこで、本研究は、ユーザによって同じ画像の解釈が異なること、また、対象あるいはカテゴリーによって、主観的な解釈が変化することを示す。そして、それらの問題を解決するために、例として、

オフィス画像における画像特徴量と解釈を表現するメディアとして形容詞(テキスト)との関係を統計的に分析し、その2つの対応関係をもとにユーザモデルを構築する方法を示す。

ここで、画像の解釈の差は、対象画像への文脈的な解釈の違いからおこると考え、個人かつ対象画像カテゴリーに依存するユーザモデルを考える。よって、同じ対象画像カテゴリーでも、個人によって解釈が異なるとともに、同じ個人でも対象画像カテゴリーによって解釈や言葉間の関係が異なると考える。それらの各対象におけるユーザモデルは、明暗・色、方向、位置などの物理的特性と形容詞の関係モデルから構築され、形容詞による街路景観やオフィスの画像検索システムに応用してきた1)。ここで、画像特徴量とは、画像内の明暗・色、方向、繰返しパターン、位置の物理的特性を組合せ数量化したものと定義する。ここで、対象画像カテゴリーによりモデルが異なる点が問題となるが、本研究では、上記の画像特徴量を利用して、任意の画像を自動的にカテゴリー化する手法を提

案する。

順序として、任意の画像は、まず、対象カテゴリーに自動的に分類され、そして、そのカテゴリー内の平均あるいは個人ユーザモデルにより、形容詞のインデックスが自動的に張られることになる。

さて、オフィスなどの室内空間の解釈(主観的解釈)についての研究として、快適性評価項目と空間構成要素(レイアウト、パーティション、机の密度、植栽、照明、環境色、アクセント・カラー、木製家具)の質的関係を分析2)、輝度分布画像から各周波数ごとの輝度変化画像から平均輝度変化量を計算し、印象と輝度分布特性との関係を分析3)4)、行為としての「集中・緊張」、「活動的・対人的+リラックス・気軽」、「静的・個人的」と室内照明の明るさや均一さとの関係を定量的分析5)などが行われている。ただ、オフィス空間においてどのような視覚機能(視覚で感知する物理的特性)6)7)が複合して主観的解釈に影響を及ぼしているかという低レベルの視覚機能に関する研究は少ない。また、オフィス空間の主観的解釈の推定の研究として、快適度、満足度、印象度と物理的なオフィス空間構成要素との関係を予測し、空間構成要素を操作する手法8)があるが、手動で空間構成要素データを作らなければいけない点など、自動化が困難である。

そのような中、感性情報処理の分野では、人間の異なる感性(主観的な解釈)を分析・モデル化を行っている。具体的には、ニューラルネットや多変量解析9)10)を用いて、人間の主観的解釈と客観的な画像特徴量との関係をモデル化している。例えば、「あたたかい」という言葉について、ある人は赤い色と対応関係をもち、別な人は明るさと対応関係をもつとすれば、「あたたかい」という言葉に対して、着眼点(物理的特性)が異なることになる。これは、個人の過去の経験や知識によって、言葉と対応する物理的特性(この場合、色と明暗)が変化するためにおこる。

そこで、経験・知識の違いによりすべての人に共有できない解釈もあるため、個人個人の学習により、解釈と画像特徴量との関係を構築する必要がある。この関係をモデル化することにより、未知のオフィス画像に対してその画像特徴量をもとに、このオフィス画像から感じる言葉、「清楚な」、「軽快な」などと自動的に推定することが可能になる。よって、オフィス画像に対して個人個人の主観的解釈を自動的に推測することが可能になる。

モデルは、刺激(入力画像)への心理的反応データに対する画像特徴量(変数)の重み(対応関係の強度)から構成されると仮定する。心理反応デー

タのばらつき(分散)が説明可能な変数には重みを大きくし、ばらつきに寄与しない変数には小さくする。また、同じ変数に対しても、個人や形容詞によって変数の重みは変化する。このような重みが変わるモデルを構築するために主成分分析および重回帰分析11)、12)を用いる。また、任意の画像を自動的にカテゴリー化する手法として、主成分分析および判別分析を用いる。

2. 主観的解釈と画像特徴量との関係

2-1. オフィス画像の主観的解釈の数量化

分析するオフィス画像として、人間の影響をなくすため、1997年から2000年の間に出版されたオフィス空間の特集を組んだ雑誌の中から新しいオフィス空間写真をスキャナでとりこみ画像サンプルを作成し、およそ200サンプル(640x480ピクセル)のオフィス画像を得た。そして、2人の人間の手で、約200サンプルを40種類に分類し、その中から代表サンプル1枚を選択した。その40サンプルを心理実験で利用するとともに、分析するオフィス画像とした。

分析する形容詞については、過去のオフィス空間の主観的解釈・評価についての研究をもとに、186語の形容詞を選定した。そして、筆者が、似たような形容詞を除き、対になっていない形容詞を対にし、46形容詞対に絞った。最後に、2人の被験者によって、上記の40オフィスサンプル画像に対して、7段階・46形容詞対の予備心理実験を行った。形容詞対の段階ごとに1から7の数字を割りあて、SD法やクラスター分析により、似た意味をもつ形容詞対にグループ化し、形容詞対の片側に偏らないよう平均が4で、形容詞対データに幅があるように分散が大きい形容詞対をその中から選択した。さらに街路景観画像で分析した画像特徴量を一部分析するので、街路景観の評価において代表的な形容詞対1)も含めた。最終的に、28形容詞対になった。

オフィス画像の主観的解釈を分析するために、下記の実験1を行った。ただし、サンプル、形容詞対は、ランダムに被験者に提示した。また、後ほど説明するが、このデータは学習用データとしても利用する。

[実験1]

サンプル:最近のオフィス空間画像40サンプル

被験者:20代の男性8人、女性11人、計19人(その内、デザイナー2人)

スケール:7段階の28形容詞対(表1)

そして、形容詞対の段階ごとに1から7の数字を割りあて、19人の平均値を主観的解釈を分析す

表1 形容詞対と繰返しパターンの分析結果

| 形容詞対 | | R2 |
|------------|------------|------|
| 上品な | 下品な | 0.91 |
| 洗練された | 野暮な | 0.80 |
| 清潔な | 不潔な | 0.84 |
| 調和のとれた | 不調和な | 0.83 |
| 秩序のある | 無秩序な | 1.00 |
| まとまった | 散漫な | 1.00 |
| 清楚な | 清楚でない | 0.85 |
| 規則的な | 不規則な | 0.84 |
| 落ち着いたものがある | 落ち着いたものがない | 0.74 |
| 複雑な | 単調な | 0.99 |
| 透明感のある | 透明感のない | 0.98 |
| 躍動的な | 静的な | 0.90 |
| 軽快な | 重厚な | 1.00 |
| 自由な | 束縛された | 1.00 |
| 開放的な | 閉鎖的な | 1.00 |
| 明るい | 暗い | 1.00 |
| カジュアルな | フォーマルな | 0.59 |
| 活気のある | 活気がない | 0.91 |
| あたたかい | つめたい | 0.98 |
| 自然な | 人工的な | 0.97 |
| やわらかい | かたい | 1.00 |
| ゆったりした | 緊張感ある | 0.91 |
| 歴史的な | 現代的な | 0.76 |
| 親しみやすい | 親しみにくい | 0.85 |
| インパクトのある | インパクトのない | 1.00 |
| 派手な | 地味な | 1.00 |
| 変化のある | 変化のない | 0.37 |
| はっきりした | ぼんやりした | 0.73 |
| 平均 | | 0.88 |

るためのデータとした。よって、40サンプル×28形容詞対のデータを被験者の平均データとし分析した。

2-2. 繰返しパターンを含めた画像特徴量

オフィス空間において照明・日照の影響による輝度分布などの明暗の物理的特性は重要なことはわかっている4)。また、屋外の街路景観の画像とその主観的解釈の関係から、明暗・色、方向、位置の物理的特性は重要であると確認した1)。そこで、本研究では、それらの物理的特性に加え、壁・床・天井などのテクスチャやテーブル・椅子に繰返しのパターンが見られることから、繰返しパターンの画像特徴量を新たに加えた。その算出方法については、1次の局所自己相関特徴モデルを適用した。そのモデルによる変数などを含む15のパラメータ変数で、壁・床・天井などのテクスチャ画像を70%の精度で18種類の素材に自動的に分類することができたことを確認した13)。

色空間としては、明暗と色の独立性を考慮して、Lab色空間にした。Lab色空間は、Lは明(+) - 暗(-)、aは赤(+) - 緑(-)、bは黄(+) - 青(-)を表す3軸からなる。その3軸に対して、平均値と分散を算出し、全体的な明暗・色の物理的特性とした。

方向については、領域の境界線(エッジ)や細線などの連続かつ直線の程度を1次の局所自己相関特徴モデルを用いて数値化する(式1参照)。その

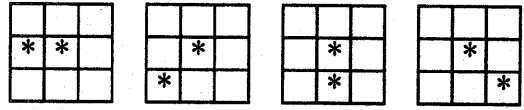


図1 方向パラメータ (ax, ay)

$$\sigma_L(a_x, a_y) = \frac{1}{mn} \sum_{i=1}^m \sum_{j=1}^n \frac{(L(i, j) - \bar{L})(L(i+a_x, j+a_y) - \bar{L})}{\sigma_i \sigma_j} \quad (1)$$

$$Dir_L(a_x, a_y) = \max_{res} \sigma_L(a_x, a_y) \quad (2)$$

$$Pat_L(a_x, a_y) = \min_{res} \sigma_L(a_x, a_y) \quad (3)$$

$\sigma_L(a_x, a_y)$: L (明度) の1次の局所自己相関特徴量 $((a_x, a_y))$ (-1, 0) (-1, 1) (0, 1) (1, 1) (図1参照)

m, n: 画像の幅, 高さ (ピクセル)

L(i, j): ピクセル座標 (i, j) のL (明度) の値

\bar{L}, σ_i^2 : m x n ピクセル画像のLの平均, 分散

res: 解像度

$Dir_L(a_x, a_y), Pat_L(a_x, a_y)$:

Lの方向, 繰返しパターンにおける画像特徴量

性質として、ある方向に連続かつ直線が存在すると1に近づく。ただし、式1において、解像度を決めて算出する方法(h×vのピクセルの平滑化画像を1単位ピクセルとし、低解像度で算出する方法)と、式2のように、h, vを可変として、その中の最大値を算出する方法がある。方向に関しては、式1と式2との違いはそれほど相違がなく、ノイズの関係から原画像(640x480ピクセル)を80x60ピクセルの低解像度(8x8ピクセルの平滑化画像を1単位ピクセルとして算出)にした時の1次の局所自己相関特徴モデル(式1)を用いた。一方、連続かつ直線が存在しないランダムな模様や均一色の場合は、0に近づく。方向のパラメータは、(ax, ay)で決まるが、図1に示したように水平、左斜め下、垂直、右斜め下、の4方向にした。

繰返しパターンに関しても、4方向の1次の局所自己相関特徴モデルを用いる。例えば、パターンの周波数が解像度と一致した場合は-1になる。よって、-1に近づくほど、周期的な繰返しパターンをもち、周期性のないランダムな模様や均一色の場合は、0に近づく。そのような性質を求めするために、h×vのピクセルの平滑化画像を1単位ピクセルとして算出し、h, vを可変として、式3のように、その中の最小値を繰返しパターンの数値とした。ただし、多重な周波数パターンをもつ画像の場合は、強い周期性のパターンに対応し、それ以外の弱い周波数のパターンは影響を無視することになる。

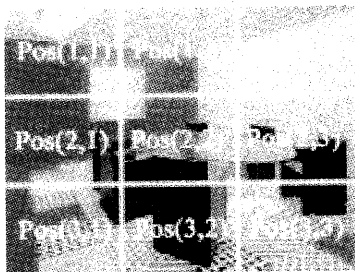


図2 位置のための3x3の分割画像

位置については、分割方法により規定しているが、前稿1)のように1点透視にあわせた三角形による分割より、図2のように3x3分割のほうが形容詞との関連が高いことが分かった10)ので、分割位置を位置のパラメータとする。

各物理的特性のパラメータは下記ようになる。

明暗 (①L)・色 (①ab) :

平均 (②ave), 分散 (②var),

方向:4方向の局所自己相関特微量

(8x8ピクセルの平滑化画像)

②dir(-1,0), dir(-1,1), dir(0,1), dir(1,1)

繰返しパターン:4方向の局所自己相関特微量

(解像度を可変の時の最小値。ただしLのみ)

②pat(-1,0), pat(-1,1), pat(0,1), pat(1,1)

位置:3x3の部分画像(図2)

③pos(1,1), pos(1,2), pos(1,3), pos(2,1),

pos(2,2), pos(2,3), pos(3,1), pos(3,2), pos(3,3)

よって、上記の丸数字のパラメータの組み合わせ(①_②_③)により、198のパラメータを画像特微量とした。例えば、図2の左上(③pos(1,1))の部分画像の明るさ(①L)の平均(②ave)は、L_ave_pos(1,1)と表現する。

2-3. 形容詞と画像特微量との対応関係の分析

実験1の形容詞対によるデータを主観的解釈のデータとし、198変数の画像特微量との関係を分析する。手法は、画像特微量のデータ間での従属性をなくすため、最初に画像特微量の変数に対して、主成分分析を行い、固有値が1以上である主成分変数を新たな変数として記録する。そして、強制投入法による重回帰分析より、画像特微量の主成分変数と各形容詞対データとの関係分析を行った。

その結果、重相関係数はすべて0.8以上になったが、有意性を検証できなかった。強制投入法による重回帰分析を用いた場合、サンプル数に対して変数の数が近いと重相関係数は高くなるためである。

そこで、手法を重回帰分析のステップワイズ法に変更した。この手法の利点は、変数がサンプル数より大きくても分析が可能で、かつ、有意な変数のみを取り込み、少数の変数で回帰式を構築す

位置については、分割方法により規定しているが、前稿1)のように1点透視にあわせた三角形による分割より、図2のように3x3分割のほうが形容詞との関連が高いことが分かった10)ので、分割位置を位置のパラメータとする。

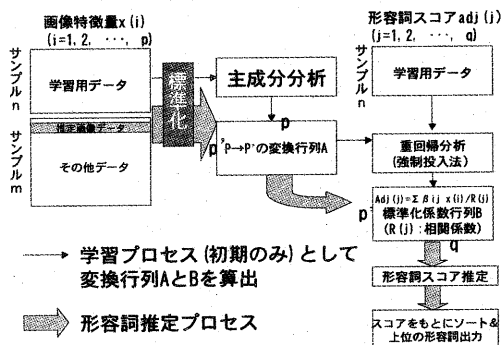


図3 学習プロセスとオフィス画像の形容詞推定プロセス

ることが可能な点にある。

表1は、各形容詞対の重相関係数を示している。28形容詞対における重回帰分析の重相関係数の平均値は、0.88となった。その結果から、繰返しパターンを含む上記4つの物理的特性と関係が強いことがわかる。ただ、「変化のあるーない」と「カジュアルなーフォーマルな」の形容詞対については、低い数値となっているが、式1の局所自己相関特微量モデルの隣接するピクセルを乗算ではなく、引算にして計算を行い、再度、分析すると、重相関係数は0.9に近い値をとった。よって、「変化のあるー変化のない」などの形容詞対は、隣との差が重要であることがわかった。ただ、それ以外の形容詞対とでは局所自己相関モデルよりよい結果が得られなかったため、その画像特微量はモデルから除いた。

強制投入法とステップワイズ法による分析結果より、オフィス画像において、明暗・色、方向、繰返しパターン、位置の物理的特性は形容詞対と強い関係が有ることが分かった。また、形容詞対とそれらの画像特微量との間に強い相関があることから、形容詞対スコアを目的変数とし、画像特微量を独立変数とした線形モデルも可能であることがわかった。この性質を用いて、次章では、主観的解釈の推定法について述べる。

3. オフィス画像の主観的解釈の推定法

3-1. 主観的解釈の推定手法

図3は学習プロセスとオフィス画像の形容詞推定プロセスを示す。まず、学習用サンプル画像および学習プロセスより、変換行列AとBを求める。特に、変換行列Bが、画像データを言葉に変換するモデルとなる。画像の画像特微量(この場合はp=198変数)を計算してデータベース内に保存しておけば、任意の画像を選択したら、その画像特微量と変換行列AとBをもとに、形容詞対(この場

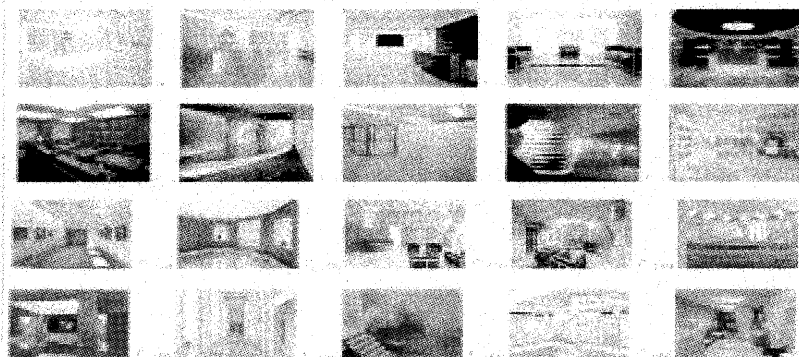


図4 実験1での学習用サンプルを含む実験2で用いたオフィス空間画像とCGで作成され画像化されたオフィス空間の一部

なオフィス空間画像50サンプル、新しくデザインされたCGを画像化した10サンプル(図4参照)の計100オフィス空間画像とした。実験2では、2種類のモデル、実験1での被験者全員のスコアを平均して構築した平均モデルと、被験者の自身のみのスコアを用いて構築した個人モ

デルを評価対象とした。
[実験2]
サンプル:100オフィス空間画像(図4参照)
被験者:実験1に参加した19人のうち、

合は $q=28$ 形容詞対)スコアを計算し、ソートすることによりその画像から受ける主観的解釈を形容詞で推定することができる。以下に簡単な手順を示す。

1. 学習用サンプル画像の標準化された画像特徴量 x に対して主成分分析を行い、固有値が1以上となる独立な主成分変数 X を得る。その変換行列を A とすると、 $X=Ax$ となる。

2. その変数 X に対して、実験1で得られた各形容詞対スコア Adj (標準化データ)に対して重回帰分析をおこなう。その変換行列を B 、 $Adj = BX = BAx$ となる。

3. 任意の画像から画像特徴量 x を計算し、変換行列 BA を掛ければ各形容詞対スコアが求まる。そして、そのスコアの絶対値が大きい形容詞がその画像から感じるであろう言葉になる。

ここで、形容詞対スコアで被験者グループの平均値を用いれば、平均モデル、よって変換行列 B' を、被験者個人のスコアを用いれば、個人モデル、よって変換行列 B'' を構築することになる。つまり、この変換行列の違いが、モデルの違いとなる。

3-2. 主観的解釈の推定法の評価実験

前節の主観的解釈、形容詞の推定法を検証するために、実験1を受けた被験者のうち15人の被験者が次の実験2をおこなった。サンプルは、実験1の学習用40サンプル、新た

デルを評価対象とした。

[実験2]

サンプル:100オフィス空間画像(図4参照)

被験者:実験1に参加した19人のうち、

男性8人、女性7人、計15人(デザイナー2人)

推定形容詞:実験1で使用した28形容詞対計56形容詞
評価モデル:(1)被験者全員の平均モデル(2)各被験者の個人モデル

評価方法:推定された上位5形容詞に対してYESかNO

3-3. 主観的解釈の評価実験結果

表2(a)は、推定された上位5形容詞に対して、実験2に参加した被験者の15人の全員の場合(ケース1)、デザイナー2人を除く13人の場合(ケース2)に分けた結果である。そして、被験者の平均モデルによる正答率と、各個人モデルによる正答率を平均した正答率を示している。

はっきりした差があるのは、平均モデルのほうが個人モデルより評価が高い点である。これは、平均モデルがよかったというより、個人データを用いる際に、誤差が発生して精度の高いモデルが構築できなかった可能性がある。よって、個人モデルを構築する際には、再度、同様な学習を行い、誤差をなくす必要がある。

表2 形容詞推定手法の評価結果

(a) 全体の平均による結果

| | 被験者15人の平均 | | デザイナー除く13人の平均 | |
|------|-----------|-------|---------------|-------|
| | 平均モデル | 個人モデル | 平均モデル | 個人モデル |
| 第1候補 | 75.5% | 66.5% | 78.0% | 68.4% |
| 第2候補 | 77.4% | 71.1% | 79.6% | 72.7% |
| 第3候補 | 78.9% | 71.8% | 82.3% | 73.7% |
| 第4候補 | 74.7% | 72.2% | 77.2% | 73.5% |
| 第5候補 | 74.5% | 71.6% | 76.8% | 73.7% |
| 平均 | 76.2% | 70.6% | 78.8% | 72.4% |

(b) 評価が高い個人の結果

| | 被験者1 | 被験者2 | 被験者3 | 被験者4 |
|------|------|------|------|------|
| 第1候補 | 87% | 86% | 82% | 83% |
| 第2候補 | 92% | 92% | 89% | 89% |
| 第3候補 | 92% | 89% | 90% | 87% |
| 第4候補 | 93% | 90% | 87% | 85% |
| 第5候補 | 86% | 91% | 89% | 91% |
| 平均 | 90% | 90% | 87% | 87% |

また、有意性は現データでは示せないが、デザイナーの被験者の評価は、他の被験者より悪かった。この原因としては、デザイナーが見て

表3 平均モデルと個人(被験者A)モデルによる図4にあるサンプルの推定結果例

| サンプルNo | モデル種別 | 第1候補 | 第2候補 | 第3候補 | 第4候補 | 第5候補 |
|--------|----------|--------|--------|---------|---------|----------|
| サンプル01 | 平均モデル | 清楚な | 洗練された | 清潔な | 落ち着きのある | 調和のとれた |
| サンプル01 | 被験者Aのモデル | 静的な | 清楚な | 洗練された | 不規則な | インパクトのない |
| サンプル02 | 平均モデル | 単調な | かたい | 清潔な | つめたい | 変化のない |
| サンプル02 | 被験者Aのモデル | 地味な | 単調な | 変化のない | つめたい | 暗い |
| サンプル04 | 平均モデル | フォーマルな | 秩序のある | まとまった | 単調な | 調和のとれた |
| サンプル04 | 被験者Aのモデル | 重厚な | 清楚な | 落ち着きのある | 洗練された | 透明感のある |
| サンプル05 | 平均モデル | かたい | 暗い | 重厚な | 束縛された | 活気がない |
| サンプル05 | 被験者Aのモデル | 閉鎖的な | 束縛された | ぼんやりした | 地味な | 変化のない |
| サンプル06 | 平均モデル | 重厚な | フォーマルな | 変化のない | 束縛された | 歴史的な |
| サンプル06 | 被験者Aのモデル | ぼんやりした | 歴史的な | フォーマルな | 閉鎖的な | 野暮な |
| サンプル10 | 平均モデル | 親しみやすい | 自由な | 開放的な | 清楚な | 自然な |
| サンプル10 | 被験者Aのモデル | ゆったりした | 不規則な | 自然な | 親しみやすい | 清楚な |

いる物理的特性が、他の被験者とは異なる可能性があり、その物理的特性をモデルに組み込む必要があると考える。

ただ、表2(a)で、デザイナーを除く13人の平均モデルの上位3位に限れば、80%前後の正答率を得ている事がわかり、利用の許容範囲であろう。

また、表2(b)には、平均モデルにおいて正答率が高い被験者を順に示して。これを見て分かるように、被験者によっては、約90%満足していることがわかる。

平均モデルと個人(被験者A)モデルによる推定結果の例を表2に示す。サンプル番号は、図4のサンプルの左上から順に対応している。見て分かるように、同じ画像でも主観的解釈が違うことが分かる。

4. 画像のカテゴリー化手法

ここでは、画像のカテゴリー、例えば、室内空間や外部空間などにより、主観的解釈や言葉と言葉の関係が変化すると仮定する。理由は、個人の経験・知識により、カテゴリー間の主観的な解釈が異なるからである。一方、画像のカテゴリー化においては、それほど主観的ではなく、個人差が少ないと考える。この仮定から、一カテゴリー内における画像と主観的解釈の関係は部分的かつダイナミックに変化する反面、画像のカテゴリー化は、より全体的かつ静的(あまり、主観的でないあるいは個人差が少ない)と考える。この章では、オフィス空間や街路景観の画像を自動的にカテゴリー化する手法を述べる。

実験1, 2の結果と街路景観の分析結果1) 10) より、明暗・色、方向、繰返しパターン、位置の物理的特性を組合せた198画像特徴量でオフィス空間や街路景観の画像をカテゴリー化する。手法としては、分析モデルにおいて、線形モデルが有効であることが分かったので、判別分析を用いる。カテゴリー化するための学習用サンプル画像として2

セットを用意し、下記の実験を行った。

[実験3]

学習用サンプルセット:

(1) 233 オフィス空間と街路景観画像

(2) 632 オフィス空間、街路景観、絵画、テキストチャ画像

分析方法: 主成分分析、判別分析

目的変数: 上記の198画像特徴量

カテゴリー化の具体的な手法は、学習アルゴリズムとカテゴリー化アルゴリズムに分かれる。学習アルゴリズムは、

1. 学習用サンプル画像から標準化された198画像特徴量 x を計算し、画像特徴量 x に対して主成分分析を行い、固有値が1以上となる独立な主成分変数 Y を得る。その変換行列を C とすると、 $Y=Cx$ となる。

2. 変数 Y に対して、実験3であらかじめカテゴリー化された各画像に対して判別分析をおこない、判別式をえる。

となり、カテゴリー化アルゴリズムでは、任意の画像から画像特徴量 y を計算し、判別式により、画像のカテゴリー化が行われる。

ここでは、学習アルゴリズムにおける分析結果のみを述べる。最初のサンプルセットのオフィス空間と街路景観画像では、98%の正答率で、両者のカテゴリー化が行なわれた。2番目のサンプルセットでは、88%の正答率で4つのカテゴリー化が行なわれた。テキストチャについては100%の正答率を得ることができたが、絵画が入ることによって、正答率が落ちた。ただ、198画像特徴量と判別分析により、かなりの精度で画像の判別が可能になることが分かった。ただ、カテゴリー化のアルゴリズムの検証がまだ、行われていないので今後の課題としたい。

5. 文脈を解釈した画像のインデックス化へ

画像の解釈の差は、対象画像の文脈の違いからおこると考え、個人かつ対象画像カテゴリーに依存するユーザモデルを考える。同じ対象画像カテゴリーでも、個人によって解釈が異なるとともに、同じ個人でも対象画像カテゴリーによって解釈や言葉間の関係が異なると考える。

そのため、4章では、任意の画像をカテゴリー化する手法を提案した。ここで、対象画像の文脈の違いに対応する仕組みとして、次のことを仮定した。

1. 画像内の明暗・色、方向、繰返しパターン、位置の物理的特性による画像特徴量による空間は、画像を判別可能な限り空間を固定可能

2. そして、その空間の中に、個人や対象画像カテゴリーに依存する部分空間が構成される。その部分空間で、個人や対象画像カテゴリーによる文脈の違いを吸収可能

よって、画像のインデックス化は、最初に、判別分析の判別式より、対象カテゴリーを判別・推定し、次に、その対象カテゴリー内で、主観的解釈の推定手法を用いてモデルを構築し、形容詞を推定・インデックス化する。ここで、ユーザが異なれば、そのユーザのモデルを用いて、インデックス化も可能である。

今回は、なぜ、このような部分空間が必要かを分析した。例えば、図5(a)は、文献9)を参考にし、実験1から被験者の平均モデルを構築し、135枚のオフィス画像から「やわらかく、あたたかい」で検索した結果である。一方、図5(b)では、文献10)を参考にし、街路景観画像30枚に対して、106人の被験者に対して心理実験を行い、その平均モデルを構築し、135枚のオフィス画像プラス98枚の街路景観画像から、「やわらかく、あたたかい」で検索した結果である。これから、街路景観画像による学習モデルを用いて、オフィス画像を検索するのは、適当でないことが分かる。

図6は、主観的解釈の推定手法を用いてモデルを構築し、形容詞を推定・インデックス化した結果を示す。図6(a)は、実験1から被験者の平均モデルとある被験者1名の個人モデルを構築し、任意の1枚のオフィス画像を選択し、主観的解釈を推定した結果である。平均モデルでは「かたい、暗い」、個人モデルでは、「かたい、透明感のない」と推定している。

一方、図6(b)では、文献10)の実験データを参考にし、街路景観画像30サンプル、106人の平均

値から、主観的解釈の推定手法を用いて平均モデルを構築し、任意の1枚の街路景観画像を選択し、主観的解釈を推定した結果、「暖かい、心がやすまる、柔らかい」となった。また、図6(c)では、実験1からオフィス画像40枚に対して、被験者の平均モデルと個人モデルを構築し、同じ街路景観画像を選択し、主観的解釈を推定した結果、平均モデルでは「インパクトのある、派手な」、個人モデルでは「親しみにくい、複雑な」となった。よって、同カテゴリー内によるモデルでは、推定に妥当性が見出せるが、オフィス画像による学習モデルを用いて、街路景観画像を推定すると妥当でないことが分かる。

よって、カテゴリーごとにモデルが必要であることが分かるとともに、カテゴリーごとに部分空間を構築する必要がある。そうすることによって、精度が高く、文脈も考慮できる主観的解釈の推定ができると考える。

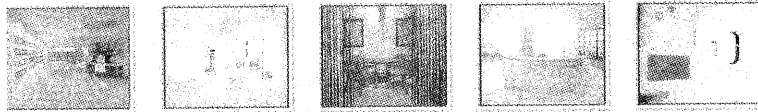
6. まとめ

本研究では、主成分分析と重回帰分析によるオフィス空間画像における主観的解釈(形容詞)の推定法と提案した。個人差はあるがおおよそ80%の正答率を得ることができた。

ユーザによって同じ画像の解釈が異なること、また、対象あるいはカテゴリーによって、解釈が変化することを示した。その原因として、個人の経験・知識からおこる対象画像への文脈の違いからおこると考え、個人かつ対象画像カテゴリーに依存するユーザモデルを考え、つまり、同じ対象画像カテゴリーでも、個人によって解釈が異なるとともに、同じ個人でも対象画像カテゴリーによって解釈が異なることがわかった。それらの各対象へのユーザモデルは、明暗・色、方向、繰返しパターン、位置の物理的特性と形容詞の関係モデルから構築され、本研究では、これらの物理的特性と判別分析を用いた学習アルゴリズムから、オフィス空間や街路景観画像などを自動的にカテゴリー化する手法を提案し、カテゴリー数が少ない場合では、約90%の割合で判別することができた。

よって、画像におけるユーザの文脈を理解するアルゴリズムとして、任意の画像を、まず、自動的に対象カテゴリーに分類し、そして、そのカテゴリー内の主観的解釈を推定する平均あるいは個人モデルにより、形容詞のインデックスを自動的に張る仕組みを提案した。

今後の課題として、この仕組みを評価していく。



第1候補
smp010.jpg
(0.0687)

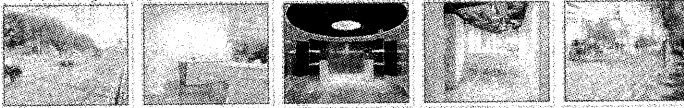
第2候補
smp051.jpg
(0.1757)

第3候補
smp053.jpg
(0.4506)

第4候補
smp041.jpg
(0.9136)

第5候補
9911136.jpg
(0.9261)

(a) オフィス画像における平均学習モデルとオフィス画像のみにおける検索結果



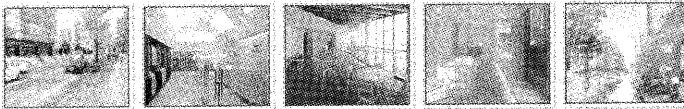
第1候補
SAMP0091.jpg
(0.0000)

第2候補
smp036.jpg
(0.0009)

第3候補
smp005.jpg
(0.0003)

第4候補
smp061.jpg
(0.0005)

第5候補
SAMP0092.jpg
(0.0022)



第6候補
SAMP0035.jpg
(0.0034)

第7候補
9906119.jpg
(0.0078)

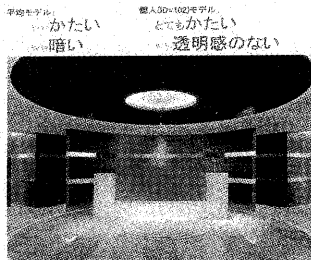
第8候補
9906218.jpg
(0.0063)

第9候補
9711138.jpg
(0.0096)

第10候補
SAMP0049.jpg
(0.0108)

(b) 街路景観画像における平均学習モデルとオフィスと街路景観画像における検索結果

図5 「やわらかく、あたたかい」をキーとした検索結果



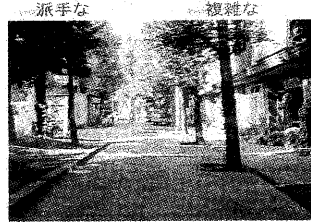
(a) オフィス画像における学習モデル

暖かい、心がやすまる、柔らかな



(b) 街路景観画像における平均学習モデル

インパクトのある、派手な、親しみにくい、複雑な



(c) オフィス画像における学習モデル

図6 主観的解釈の推定手法による形容詞推定・インデックス化結果

参考文献

- 1) 柴田: 街路景観の主観的解釈のモデル化と画像検索システムへの応用 - 形容詞と画像特徴量との対応関係, 日本建築学会計画系論文集 No. 515, pp. 297-304, 1999. 1.
- 2) 乾, 中村, 窪田, 丸山, 李: オフィス環境の快適性評価に関する研究, 日本建築学会計画系論文集 No. 399, pp. 29-36, 1989. 5.
- 3) 中村, 乾: 視環境の輝度分布特性に関する研究, 日本建築学会計画系論文集 No. 438, pp. 1-8, 1992. 8.
- 4) 中村, 乾: オフィスの輝度分布特性とその主観的解釈, 日本建築学会計画系論文集 No. 445, pp. 27-33, 1993. 3.

- 5) 小林, 乾, 中村, 北村: 室内環境照明の明るさ、均一さと生活行為の関係, 日本建築学会計画系論文集 No. 481, pp. 13-22, 1996. 3.
- 6) Zeki, S.: A Vision of the Brain, Blackwell Science, Oxford, 1993.
- 7) 平井有三: 視覚と記憶の情報処理, 培風館, 1995
- 8) 大森, 今川, 小島, 平手: オフィスにおける「居住性の快適性」に関する簡易予測方法の実験的検証とその応用 - オフィス空間の快適性簡易予測システムの開発 (その2), 日本建築学会計画系論文集 No. 479, pp. 57-66, 1996. 1.
- 9) 栗田, 加藤, 都田, 坂倉: 印象語による絵画データベースの検索, 情報処理学会論文誌, Vol. 33, No. 11, pp. 1373-1383, 1992. 11.
- 10) 柴田, 加藤: 街路の景観画像データベースのイメージ語による検索, 電子情報通信学会論文誌 Vol. J82-D-1, No. 1, pp. 174-183, 1999. 1.
- 11) 奥野ほか: 統多変量解析法, 日科技連出版社, 1976.
- 12) 竹内, 柳井: 多変量解析の基礎, 東洋経済新報社, 1972.
- 13) Shibata and Kato: Bilateral functions of direction and pattern by auto-correlation model - Categorizing texture and office images, Proc. of IEEE International Conference on System Man and Cybernetics 2000, 1639-1643, October 2000, Nashville, (Tennessee), USA