

ウェーブレット解析手法を用いた建築内部空間画像と知覚イメージの相関関係の分析 CORRELATION ANALYSIS BETWEEN PHOTOS OF INTERNAL SPACE AND SUBJECTIVE IMPRESSION USING TWO-DIMENSIONAL WAVELET TRANSFORM

宮高泰匡*, 加藤直樹**, 藤沢克樹***

Yasumasa MIYATAKA, Naoki KATOH and Katsuki FUJISAWA

When one experiences an architectural space, he/she perceives various impressions. The purpose of this paper is to quantitatively clarify the relationship between the impression perceived on a photo of an architectural internal space and the physical features of its color image. For fifty sample color images of internal space, we have performed a questionnaire concerning what impression he/she acquires for each image by asking him/her to choose one of the impression words from a pair of antonyms. Also, we have computed color and texture features of photos. Here we used two-dimensional wavelet transform to obtain texture features while Lab-color space is used to extract color features. We then applied a decision-tree algorithm in order to derive interpretable and meaningful correlation of the impression words and image features. As a result, for images for which a majority of people had the same impression, we have found an interesting, interpretable common feature among the images.

Keywords: Photos of Internal Space, Subjective Impression, Wavelet Analysis, Correlation Analysis, Data Mining
建築内部空間画像、知覚イメージ、ウェーブレット解析、相関分析、データマイニング

1 はじめに

1.1 研究の背景・動機

日常生活を営む建築物や旅行・観光で訪問する建築物に対して、我々は無意識のうちに何らかの印象を持っているがその印象は対象物により様々なものがあるであろう。言葉で表現できるものもあれば、できないものもあろう。また、抱く印象には個人差がある場合も多く、さらに周囲の状況、時間帯、季節によって同じ個人でもしばしば異なった印象を受ける場合も存在すると考えられる。しかし、一枚の建築物の写真を見た時に一般に多くの人々が同じ印象を持つ場合があることも否定できない。なぜ、多くの人々がそのような共通した印象を持つのであろうか？ それまでに人々が獲得してきた知識の中に普遍的なものがあるためであろうか。このような問いに真の解答を見つけるのは困難であろうが、写真から得られる物理的特徴と印象との間に何らかの関連があると推測される。

一方、近年の画像認識技術の応用として画像特徴情報によるデータベース検索技術の研究開発が盛んに行われている^{1,2)}。これらの研究は「お互いに似ている画像は似た特徴量を持っている」という前提を元に成り立っており、この技術はあらかじめ索引を作って検索するのではなく、画像特徴の類似度によって画像検索を行うものである。ここで用いられている画像特徴とは、一口に言って背景知識を必要としない色彩、テクスチャ¹⁾、形（直線、線分、円などの基本図形）といったローレベル特徴³⁾である。

また一方で建築物においても、その内部空間に雰囲気や印象を与える色彩・テクスチャなどの要素は重要な特徴情報であると考えられる。そのような観点から、本研究では個人を問わず一般的に多くの人々の間に共通する認識が存在するという仮定のもとで、「人が建築空間を体験する際にその空間のどのような特徴に注目して

印象を知覚するのであろうか」という問いを明らかにすることを目標とし、建築内部空間画像が観察者に与える知覚イメージと画像から得られるローレベル特徴の相関関係の分析を行う。また、相関分析を通して得られた情報をもとに「人々がどのような特徴情報に注目して知覚イメージを得ているのか」という点について推論を行う。

1.2 既往研究

都市空間や建築空間に対する画像特徴と知覚イメージとの関連を分析した研究に、柴田等による「街路の景観画像データベースのイメージ語による検索」というタイトルの論文がある⁴⁾。彼らは街路景観画像と知覚イメージの対応関係により、自動的に未知の景観画像を評価する画像検索システムの構築を目指して研究を行った。彼らは多変量解析の手法を用いて、人間が知覚する際に感じる物理的特性（色、明るさ、線分の多さなど）を定量的に解析し、様々な知覚イメージとの相関関係を導出したが、その相関関係を明確に人間に理解できる規則として表すという視点は無かった。また、これまで建築空間の知覚イメージとその物理的特性の対応関係を定量的に分析した研究も筆者の知る限り存在しない。それゆえ、その対応関係をより定量的にかつ人間が解釈可能な規則によって説明しようとする本研究の試みは建築空間の印象を分析する新しい有効な方法であると考えられる。

1.3 研究の目的

本研究では建築内部空間画像の印象とその画像特徴との相関関係を規則として導出し、人々がどのような特徴に着目して画像に対する共通の印象を得ているのかについて推論・解釈を試みる。相関関係を抽出する方法として、ルールを木構造で表現できる決定木手法⁵⁾を採用する。決定木はデータマイニング⁶⁾の一手法であり、分類や予測を行う際に広く使われており、分類規則が自然言語によって表現できるため、人間にとって理解可能なルールが得られるとい

*京都大学工学研究科建築学専攻 修士課程 大学院生

**京都大学工学研究科建築学専攻 教授・工博

***東京電気大学理工学部数理科学科 助教授・博士（理学）

Graduate Student, Graduate School of Architecture, Kyoto Univ.
Prof., Dept. of Archi. and Architectural Systems, Kyoto Univ., Dr. Eng.
Associate Prof., Dept. of Mathematical Sciences, Tokyo Denki Univ.,
Dr. Sci.

う利点がある。なお、一つの建築内部空間についての情報はある一視点の一枚の画像に限定し、その画像から得られるローレベル特徴を画像特徴情報とする。しかし、建築の内部空間画像が対象であるため、天井や壁など5つの窓領域に区切り、建築空間であるという背景知識を利用して画像特徴の定量化を行う。背景知識を必要としない画像特徴（ローレベル特徴）の抽出には、ウェーブレット変換^{7,8)}という空間周波数解析に比較的有効な手法を用いて得られるテクスチャ特徴とともに、Lab色座標空間(CIE-Lab)を用いることによって色彩特徴を抽出する方法を用いている。

1.4 研究の概要

研究の手順は、まず準備として建築空間のサンプル画像の収集を行う。次にサンプル画像から画像の特徴量抽出と印象量測定を行い、ローレベル特徴と知覚イメージを取り出す。この時、印象量の測定は柴田等の研究⁴⁾と同様に、印象語対を用いて建築内部空間画像についてのアンケートを行う。また、画像特徴の抽出に関してはテクスチャ特徴はウェーブレット変換、色彩特徴はLabという色座標空間を用いて定量化する(図1参照)。そして決定木手法を用いて画像の特徴量から印象量を予測するモデルを生成し、画像の印象量と特徴量の対応関係を解釈可能なルールとして見つけ出す。最後に得られた決定木ルールからその知覚イメージに対して相関の高い画像特徴を取り出し、建築空間の画像の物理的特徴と印象量との対応関係を分析する。

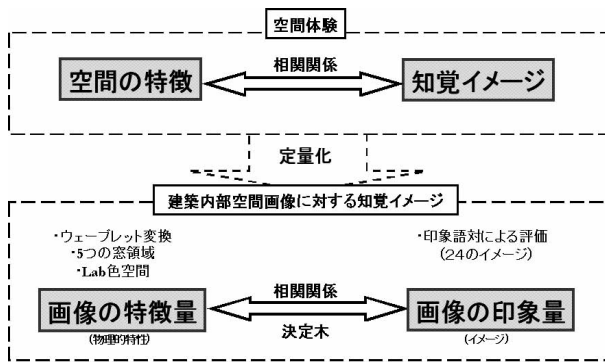


図1: 相関関係分析の過程

2 対象の画像と知覚イメージ

2.1 対象画像の選択

空間のイメージ評価を目的とするため、分析のサンプルとしてある程度奥行き感がある、比較的平らな天井・左右の壁・床(地面)に囲まれた建築内部空間の画像を50枚選択した。紙面の都合上、画像の全てを載せることは省略した。あるイメージ語について特徴的な画像については図16,17に載せているので参照されたい。

2.2 画像の知覚イメージの測定

知覚イメージを測定するため、建築画像に関するアンケートに答えてもらう。その際、画像に対するイメージ評価は対義語となる形容詞ペアを用いる。ここではこのペアを印象語対と呼び、柴田等⁴⁾の使用した23印象語対に「西洋的な*東洋的な」を含め、表1の合計24印象語対を用いる。アンケート対象者の内訳は建築に関わる人10名、関わらない人11名で、合計21名であった。

アンケートの集計方法として、図2のように一つの印象語対(ここでは「暖かい*冷たい」という温度感の形容詞ペア)の四段階評価とし、その知覚イメージの評価量として下図のように-0.75、-0.25、0.25、0.75を与える。評価を行ったアンケート回答者の評価量の

平均値をその画像の評価値とする。その結果、画像によらず評価値が一方に偏っていた「美しさ」、「一体感」、「曖昧性」、「秩序感」、「快適性」の5つの知覚イメージを削除することにした。

表1: 知覚イメージと印象語対

知覚イメージ	印象語対	知覚イメージ	印象語対
地域性	西洋的な*東洋的な	時代性	現代的な*歴史的な
美しさ	美しい*醜い	温度感	暖かい*冷たい
形式性	斬新な*典型的な	安心感	安心できる*不安な
一体感	一体感がある*一体感のない	開放感	狭い*広い
性別感	女性的な*男性的な	密度感	疎な*密な
奥行き	奥行きがある*奥行きのない	秩序感	秩序のある*無秩序な
やわらかさ	柔らかい*硬い	単純性	単純な*複雑な
陽気さ	陽気な*陰気な	伏魔性	伏い*不伏な
曖昧感	和やかな*緊張感のある	曖昧性	明確な*曖昧な
印象度	印象的な*無印象な	大きさ	小さい*大きい
重厚感	軽やかな*重厚しい	好き嫌い	好き*嫌い
落ち着き	落ち着いた*落ち着きのない	新しさ	新しい*古い

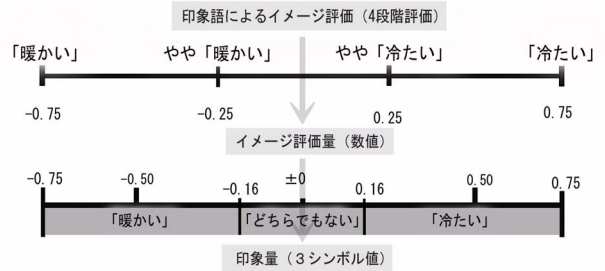


図2: 印象量への集計方法(温度感の印象語ペア)

ここまではイメージ評価量を数値で表していたが、相関分析を行う際に決定木アルゴリズムを用いるため、数値をシンボル値に変換し直す必要がある。評価量を集計すると実際に値の分布する区間は-0.75~0.75よりも狭まるため、図2のように負の評価量平均(-0.5)と正の評価量平均(0.5)を三等分した-0.16と0.16を閾値として三種のシンボル値への変換を行った。つまり、図2の場合、「暖かい」、「どちらでもない」、「冷たい」の3つの値に変換した。

この結果、「どちらでもない」という評価が過半数を超えていた「性別感」、「やわらかさ」、「安心感」の3つの知覚イメージを分析対象から外し、残りの16の印象語対を用いて分析を行う。

3 建築空間の画像特徴量抽出

今回、画像特徴として「色彩」特徴と「テクスチャ」特徴を取り出す。我々が建築空間の雰囲気を感じ取るときには、その空間の一部や壁などの面要素、またその組み合わせ、全体のバランスなどに影響されると考えられる。本研究においては空間の面的要素を考慮するため図15のように5つの窓領域に分割し、空間の面的要素を含んだ特徴量抽出を行う。なお、この際領域ごとに面積が異なるため、領域の特徴量は単位面積当りで計算する。

3.1 テクスチャ特徴

画像のローレベル特徴の一つであるテクスチャ特徴の解析には、空間周波数に基づいた技術の代表的手法であるウェーブレット変換を用いる^{7,8)}。まず、ウェーブレット変換について概説する。

ウェーブレットは”小さな波”のことで、その定義は「局在性のある波」である。この手法は、無限に続く局在性のない三角関数を用いるフーリエ変換とは異なり、多重解像度解析を行えるという利点がある。また、空間信号を異なった周波数(方向)だけではなく、異なった解像度(スケール)ごとに分解することができる。

横軸を(一次元)空間での位置(変数 x とする)、縦軸を空間信号の大きさとしてとれば、ウェーブレットは図3のような波となる。その左の波を基本ウェーブレットといい、 $\psi(x)$ で表す。ウェーブレット関数は2つの実数パラメーター a 、 b を用いて、

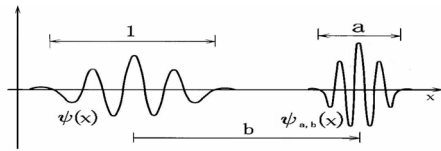


図 3：基本ウェーブレット

$$\psi_{a,b}(x) = \frac{1}{\sqrt{a}} \psi\left(\frac{x-b}{a}\right) \quad (1)$$

と表される。これは図 3 の右側のウェーブレットであり、この関数はマザーウェーブレットを横方向に a 倍に伸ばし、横方向に b だけ平行移動したものである。なお、 a のことをスケールといい、ウェーブレットの空間幅を表す。なお、 $\frac{1}{a}$ は周波数に相当する。

デジタル画像は離散値で色度の階調が表されているため、離散ウェーブレット変換を用いる。よって次にこの手法の説明を行う。

適当なマザーウェーブレットを選び、スケール a および平行移動量 b を適当に離散化すればウェーブレット関数を正規直交基底とすることが可能になる。パラメータ a, b を離散化し、

$$a = 2^{-j}, b = k2^{-j}, j = 0, \pm 1, \pm 2, \dots, k = 0, \pm 1, \pm 2, \dots \quad (2)$$

とすると、基本ウェーブレットは次のように書き換えられ、

$$\psi_{j,k}(x) = \frac{1}{\sqrt{2^{-j}}} \psi\left(\frac{x - k2^{-j}}{2^{-j}}\right) = \sqrt{2^j} \psi(2^j x - k) \quad (3)$$

これを離散ウェーブレット関数という。ここで、 $f[n]$ を N 個のサンプリングデータからなる離散信号とし、基本ウェーブレットの持続区間をサンプリング周期に一致させると、(3) 式の変数 x は整数値 n で置き換えることができ、

$$\psi_{j,k}[n] = \sqrt{2^j} \psi(2^j n - k) \quad (4)$$

となり、離散信号 $f[n]$ に対する離散ウェーブレット変換は、

$$f[n] = \sum_{j,k} W[j,k] \psi_{j,k}[n] \quad (5)$$

と表現され、 $W[j,k]$ をウェーブレット展開係数と呼ぶ。

また、ウェーブレット変換は、低域通過フィルタと高域通過フィルタをペアとして用いて、周波数成分の低域側の信号を次々と 2 分割する作業を繰り返すことによって行われるサブバンド分解であり、低域通過フィルタによって近似波形（低周波成分）と高域通過フィルタによって誤差波形（高周波成分）に分解することによって画像（空間信号）から特徴（ウェーブレット展開係数）を抽出する。

なお、画像は 2 次元の信号であるから 2 次元ウェーブレット変換を行う。2 次元のウェーブレット関数は水平方向と垂直方向の（一次元）ウェーブレット変換を交互に繰り返すことによって作られる。図 4 のようにまず各行について垂直方向の一次元ウェーブレット変換を行い、画像の水平方向の低周波成分と高周波成分に分割する。その後、同様に垂直方向の変換を行う。つまり、一度の 2 次元ウェーブレット変換によって原信号は 4 分割される。さらに、LL 成分に対して以上の操作を繰り返すことにより、合計 7 分割される。一般に N 回の 2 次元ウェーブレット変換によって $3N + 1$ 個に分割されるが、ここでは $N = 3$ とした。つまり、3 回の 2 次元ウェーブレット変換を繰り返し、合計 10 個に分割することによって、高・中・低、三段階の周波数成分（解像度）のテクスチャー特徴を取り出すことにした。すなわち画像のテクスチャー情報はウェーブレット変換によって数値化し、一つの濃淡画像から 3 方向 × 3 解像度に相当する合計 9 次元のベクトルを取り出した。分析に用

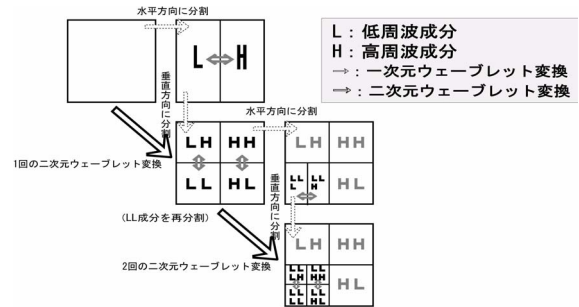


図 4：二次元ウェーブレット変換のプロセス

いる特徴量は、直線の方向性や解像度の大きさも重要な関係があると考えられるため⁸⁾、先ほどの 9 に加えて、方向ごと、つまり垂直（HL, LLHL, LLLLHL）、水平（LH, LLLH, LLLLLH）とその対角方向である斜め（HH, LLHH, LLLLHH）ごとの和や高、中、低の解像度ごとの和も計算し、合計 15 のテクスチャー情報を用いる。

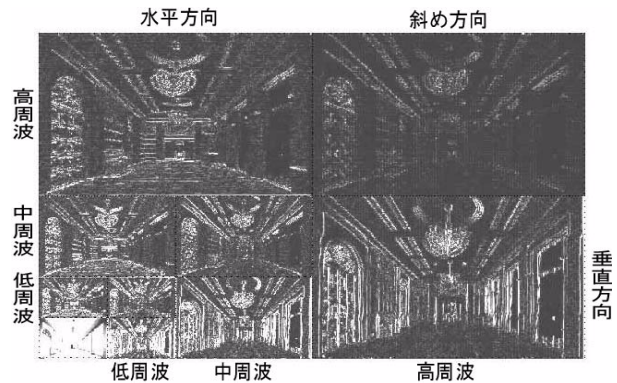


図 5：図 15 で使用された画像の二次元ウェーブレット変換

なお、テクスチャー情報の特徴量として「ウェーブレットエネルギー」を用いる。これは画像の二次元ウェーブレット変換によって得られたウェーブレット展開係数の二乗和を、その画像の全ピクセル数で除した値である。ウェーブレットエネルギーと画像の物理的特性との対応関係について概説しておく、例えば画像において（図 5 参照）、もし横方向の細かい縞が多く見られたならば、水平方向高周波成分のウェーブレットエネルギーが大きいことを示す。また、もし斜め低周波成分のウェーブレットエネルギーが大きかったならば、画像において斜め方向に太目の直線が多いとか、斜めの大きな縞模様が顕著に見られたことを示す。

3.2 色彩特徴

色彩は建築空間に限らず、身の回りのモノや衣服、インテリア、さらには景観などにおいて、視対象に対する人間の心理や印象に大きな影響を与える。本研究では「空間色彩がどのように心理評価に働くのか」ということを考慮に入れて色彩特徴を抽出する。色彩は異なる 3 変数で表される量であり、色を表す 3 次元座標系のことを表色系と言う。例えば、RGB 表色系でカラー画像を表すと、光の三原色である赤、緑、青という 3 つのカラーチャンネルにわけてそれぞれの濃淡画像の重ね合わせで表現される。表色系には RGB 以外に色の 3 属性（色相、明度、彩度）に基づく HSI や Lab など様々なものがあるが、本研究では抽出されるカラー特徴の評価は人間が行うので、人の知覚とよく整合しているとされる、CIE-Lab⁹⁾、^{注 2)} という色座標空間を用いる。Lab 色空間は RGB を線形変換した表色系であるが、この 3 次元色空間における変化量とそれによる視覚の色変化の印象とが比例するよう補正されており、デバイスに依存せず色彩を決定できるという特徴がある。L は「黒～白の階調

数」、つまり明度を表し、aは「赤～緑の階調数」、bは「青～黄の階調数」を表す。

色彩の特徴量としては色の階調数の平均値、標準偏差（色のバラつき度合）を採用し、5つの窓領域ごとにL,a,bの3チャンネルずつ計算を行う。以上に加え、L,a,bのチャンネルに分解した後、各チャンネルにおけるウェーブレットエネルギーを3.1節と同様に求めてゆく。

次にLabの平均値及び標準偏差の値と画像特徴との対応関係を概説する。図6～8は図15で用いた画像（全ピクセル数：460800、階調数：0～255の256）のL,a,b各チャンネルのヒストグラムである。図6では、明度の平均値は128より大きく、その階調のばらつきも大きい。これにより、画像としては全体的に明るいものの、場所によって差がかなりあることが分かる。図7のヒストグラムでは、aの平均値は128より大きく、そのばらつきは小さい。これにより画像は全体的にやや緑っぽい配色であり、色彩において統一感がある、もしくは配色が単調であるということが分かる。図8では、bの平均値は128よりも大きく、ばらつきが小さい。このことから、画像は黄色っぽい配色でその色彩の統一感がある、もしくは配色が単調であることを示す。

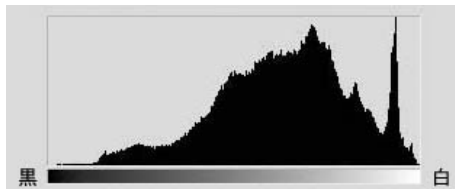


図6：Lのヒストグラム：黒～白（明度）
：平均値 162.50、標準偏差 46.27、中間値 166

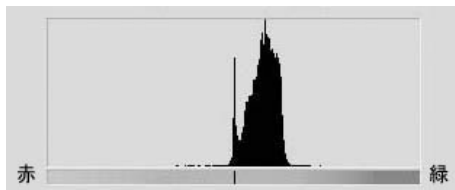


図7：aのヒストグラム：赤（マゼンタ）～緑（エメラルドグリーン）
：平均値 146.84、標準偏差 9.39、中間値 148

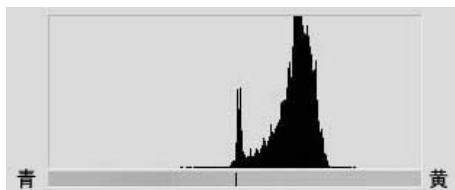


図8：bのヒストグラム青（マリンブルー）～黄（イエロー）
：平均値 166.32、標準偏差 14.92、中間値 170

3.3 画像特徴量の概要

本研究では大きく分けて3つの画像特徴量を用いる。一つ目は「画像全体のウェーブレットエネルギー」であり、L,a,bの各カラーチャンネルごとに15のウェーブレットエネルギーを取り出す。二つ目は「5つの窓領域ごとの色彩の情報量」であり、5つの窓領域ごとにL,a,bの各カラーチャンネルの階調数の平均値・標準偏差の値を取り出す。最後の一つは「領域分割を考慮したウェーブレットエネルギー」であり、5つの窓領域に分割したグレースケール画像から15のウェーブレットエネルギーを取り出す。以上の合計150の画像特徴量を抽出する。

4 相関関係

今回、150もの画像特徴量を取り出したが、空間のイメージにはどのような特徴因子が影響するのかを分析するのは容易ではなく、言われるまで気づかないような画像特徴と知覚イメージとの相関も多くあると考えられる。そのような相関を発見するという観点からデータマイニングという手法を用いる。データマイニングとは大量のデータを解析し、その中から相関関係や規則性などの有用な情報を発掘（mining）する技法で、近年注目を集めている技術である⁶⁾。

本研究では、分析にデータマイニングの代表的手法である「決定木」^{5,6)}という手法を用いる。決定木は、近年のデータマイニング技術の進展とともに注目を集めており、ニューラルネットワークや判別分析と並んで代表的な予測手法である。この手法の利点は理解可能なルールを表現し、予測や分類においてどの項目がもっとも重要であるかを明確に示すことができる点である。もう一つの利点は、計算時間が極めて少ないという点である。

4.1 決定木

決定木とは事例を分類する木構造の形をした多層の分岐ルールのことであり、以下に少し概説する。いま、複数属性から成るレコードの集まり（以下では事例集合という）があり、属性には複数もしくは一つの説明属性と一つの結果属性（二つ以上の離散値）があるものとする、決定木は説明属性を用いて結果属性の取る値を予測するモデルにあたる。その構成は結果属性の値が既知である学習用の事例集合を用いて次のように行われる。結果属性は「正」または「負」という値を取るものとする、様々な分岐ルールの中で、そのルールを満たす事例集合と満たさない集合に分類することにより、事例の「正、負」を判定する能力が最も高くなる分岐ルールを作成し、それを元に分類を再帰的に行ってゆく。こうすることにより事例集合の結果属性の値を説明属性の値によって判別できる決定木が生成される。なお、その判別能力は分類後の情報エントロピー^{12,3)}によって評価する。本研究での事例集合は建築内部空間画像についてのレコードの集まりであり、そのレコードは画像特徴という説明属性と知覚イメージという結果属性を持っている。

図9は「商品Aの購買の有無」に関する決定木の概念を示したものであり、決定木は最上部に「根ノード」、中間部に「中間ノード」、最下部に「葉ノード」があり、根ノード、中間ノードには分岐ルールが対応し、葉ノードでは正負のいずれかの分類結果が対応する。なお、本研究では決定木生成アルゴリズムとしてC5.0¹²⁾を用いる。

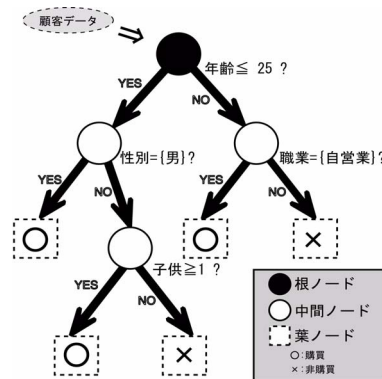


図9：「商品Aの購買の有無」に関する決定木の概念図

C5.0では、上記の決定木ルール以外に、ルールセットとよばれる分類ルールが生成される^{5,6)}。決定木ルールの場合、if-then節

で分岐ごとに分割の条件が説明されており、分類対象のデータが再帰的に分割されていき、理解しやすいのだが、結果属性ごとに集計されておらず全体として結果を分析する際には困難がある。一方ルールセットは決定木ルールの単純化や重要な部分の抽出を行って結果属性ごとにまとめたルール集合であり、ルールの全容をつかみやすいモデルである。そのため、本研究ではルールセットを用いる。例えば、図 10 は「時代性」の知覚イメージについてのルールセットで決定木ルールをそれぞれ「歴史的」や「現代的」という結果属性ごとにまとめている。

ルールセットの精度については予測精度^{注5)}と検証精度を求めた。十分に大きいデータセットがあれば、学習用データと検定用データに分割し、学習用データで生成された予測モデル(決定木ルール)を検定用データで検証するということが可能であるが、今回はサンプルデータの数が50であり、非常に少ないため、その手法ではモデルを検証することができない。そこで交差検証^{注6)}を行って、決定木の検証精度を求めた。

ルールセット:1 (予測精度 92.3%, 検証精度 58.9%)	
● 結果属性が「歴史的」となるルール	
ルール:1	If 天井窓領域垂直高周波成分 > 33.766 then → 歴史的 (10, 0.833)
ルール:2	If b チャンネル水平中周波成分 > 1144.5 then → 歴史的 (7, 0.778)
ルール:3	If 床面窓領域全垂直成分 > 530.547 and 左側窓領域全水平成分 > 1366.06 then → 歴史的 (6, 0.75)
● 結果属性が「現代的」となるルール	
ルール:1	If b チャンネル水平中周波成分 ≤ 1144.5 and 左側窓領域全水平成分 ≤ 1366.06 and 天井窓領域斜め高周波成分 ≤ 33.766 then → 現代的 (13, 0.933)
ルール:2	If 床面窓領域全斜め成分 ≤ 530.547 and b チャンネル水平中周波成分 ≤ 1144.5 and 天井窓領域斜め高周波成分 ≤ 33.766 then → 現代的 (9, 0.909)

図 10 : ルールセット

なおルールセットの then 節の後のカッコ内の二つの数字は左がサポート数で、右が確信度を表す。サポート数はこのルールが適用できるケース数、つまり if の条件に当てはまるデータの数を示す。一方、確信度は予測が正しい場合、つまり if の条件が当てはまるデータのなかで、予測が正しかったデータの割合を示す。

5 実験

これまでの作業によって、50枚の画像各々に対する16の印象語対に対する印象量のデータ、50枚の各画像に対するテクスチャ・色彩情報である合計150種類の画像特徴量のデータが揃ったので、決定木を用いて相関関係の抽出を行う。

5.1 実験の手順

16の印象語対の各々に対して以下の1~4を行う。

1. 初期ルールセットの生成: 決定木生成アルゴリズム(C5.0)を用いて印象量と画像特徴量間の具体的な関連を示すルールセットの作成を行う。この際、入力特徴変数として全ての特徴量を用いて初期ルールセットを作る。
2. 相関の強い特徴量の抽出: 初期ルールセットからその知覚イメージに対して相関の強い画像特徴量を抽出する。生成されたルールセットで使用された画像特徴は印象に対して強い影響がある属性と考え、各々のルールセットで使われた画像特徴量を相関の強い特徴量として取り出す。
3. 優秀なルールセットの生成: より優れたルールセットの生成をおこなうために、先ほど取り出した相関の強い画像特徴のみ

を入力変数として用いて、C5.0を適用する。このことによって、相関の低い画像特徴を使用せずに決定木を生成できるため、相関の強い画像特徴だけを用了優秀な決定木ルールが作成できると考えられる。

4. 精度の検証: 決定木から得られたルールの信憑性、つまり印象量と画像特徴量の対応関係の信頼性を検証する。まず、ルールセットの検証精度、それぞれのルールの確信度・サポートを確認しルールの信憑性を確認する。

5.2 実験結果

C5.0により16の知覚イメージの各々について画像特徴との相関関係を表す優秀なルールセットが得られた。以下の表2にそれぞれの印象と相関の高かった画像特徴についての概要を表記する。最終的に得られたルールは一つではないが全ては書ききれないため、ここでは、16の印象語対それぞれについて、印象語対の結果属性のうち一方の条件を表すルールだけを示す。(「どちらでもない」を示すルールは対応関係の分析には使えないため、省いている。)なお、表には得られた相関関係の確信度、サポートなどの精度も表記してある。

表 2: 知覚イメージと得られた相関ルール (優秀なルールセット)

知覚イメージ (予測精度, 検証精度)	相関ルール (サポート数, 確信度)
地域性 (82.1%, 66.7%)	If Lab-a の斜め高周波成分 > 80.921 and 中央窓領域全高周波成分 > 11.052 → 西洋的 (22, 0.875) If 天井窓領域垂直高周波成分 > 33.766 → 歴史的 (10, 0.833), If b チャンネル水平中周波成分 > 1144.5 → 歴史的 (7, 0.778), 床面窓領域全垂直成分 > 530.547 and 左側窓領域全水平成分 > 1366.06 → 歴史的 (6, 0.75)
時代性 (92.3%, 58.9%)	If 天井窓領域垂直高周波成分 > 33.766 → 歴史的 (10, 0.833), If b チャンネル水平中周波成分 > 1144.5 → 歴史的 (7, 0.778), 床面窓領域全垂直成分 > 530.547 and 左側窓領域全水平成分 > 1366.06 → 歴史的 (6, 0.75)
温度感 (82.1%, 59.2%)	If 中央窓領域 a チャンネル平均値 ≤ 131 → 冷たい (11, 0.692)
形式性 (84.6%, 59.2%)	If 天井窓領域垂直高周波成分 ≤ 1575.58 and 天井窓領域全垂直成分 > 612.835 and 右側窓領域 L チャンネル平均値 ≤ 125 → 典型的 (11, 0.923), If 床面窓領域垂直高周波成分 ≤ 27.478 and 右側窓領域 L チャンネル平均値 > 125 → 典型的 (9, 0.636)
単純性 (84.6%, 51.7%)	If 右側窓領域 b チャンネルの標準偏差 ≤ 7.1 and 左側窓領域全垂直成分 ≤ 1784.78 → 単純な (9, 0.909), If a チャンネル全垂直成分 > 4711.68 and 左側窓領域全垂直成分 ≤ 1784.78 → 単純な (14, 0.813)
開放感 (89.7%, 66.7%)	If 右側窓領域水平中周波成分 ≤ 0.925 → 開放感のある (7, 0.889), If 右側窓領域 a チャンネル平均値 ≤ 130 → 開放感のある (9, 0.636)
奥行き (82.1%, 44.2%)	If 中央窓領域水平高周波成分 ≤ 1348.72 → 奥行きある (34, 0.667)
陽気さ (84.6%, 30.8%)	If 床面窓領域 L チャンネルの平均値 > 129 and 床面窓領域 a チャンネル標準偏差 > 23 → 陽気な (12, 0.786)
密度感 (87.2%, 51.7%)	If 右側窓領域 L チャンネル平均値 > 139 and 中央窓領域 a チャンネル平均値 ≤ 133 → 凝在した、疎な (8, 0.9), If L チャンネル斜め中周波成分 ≤ 610.567 → 凝在した、疎な (9, 0.636)
印象度 (76.9%, 32.5%)	If a チャンネル斜め中周波成分 > 387.73 → 印象的な (9, 0.818), If a チャンネル水平高周波 ≤ 224.831 → 印象的な (12, 0.429)
緊張感 (82.1%, 42.5%)	If b チャンネル全高周波成分 ≤ 236.288 and 天井領域 L チャンネル平均値 > 115 → ゆとりある (9, 0.909), If L チャンネル水平低周波成分 > 5334.83 and 天井領域 L チャンネル平均値 > 115 → ゆとりある (7, 0.667)
大きさ (84.6%, 59.2%)	If a チャンネル斜め低周波成分 > 2219.16 → 大きな (6, 0.875), If 床面窓領域 L チャンネル平均値 > 94 and 床面窓領域垂直中周波成分 ≤ 104.593 and 床面窓領域 a チャンネル標準偏差 ≤ 6.6 → 大きな (10, 0.75)
重量感 (76.9%, 38.3%)	If 中央窓領域 L チャンネル平均値 ≤ 137 and 左側窓領域 a チャンネル標準偏差 ≤ 13 → 重たい (21, 0.522)
好き嫌い (92.3%, 35.0%)	If 天井窓領域 L チャンネル標準偏差 ≤ 24 → 好きな (7, 0.889)
落ち着き (89.7%, 56.7%)	If 天井窓領域 L チャンネルの標準偏差 ≤ 24 → 落ち着きある (7, 0.889), If 左側窓領域 b チャンネル標準偏差 < 8.7 and 左側窓領域斜め高周波成分 ≤ 32.667 → 落ち着きある (9, 0.727)
新しさ (79.5%, 55.0%)	If 天井窓領域斜め低周波成分 ≤ 385.759 → 新しい (12, 0.714), If 左側窓領域全低周波成分 ≤ 1723.94 → 新しい (11, 0.615)

表 2 では予測精度と検証精度を記載しているが、二つの値の間には大きな差が存在する場合も見受けられる。これは、今回使用する画像数が少ないため過学習^{註7)}が起きていることが考えられる。しかし、検証精度に関して言えば本研究で生成したルールセットは3種の値(例えば、知覚イメージが「地域性」の場合、東洋的、西洋的、どちらでもない)を予測するモデルであるため、50%以上であれば、ルールの分類精度が高いと考えてよいと思われる。

6 相関関係の分析

前節では建築内部空間画像の画像特徴とその知覚イメージの相関関係をルールセットとして求めた。このルールセットから、どの画像特徴が各々の知覚イメージと関連があるかが明らかになった。

本節では、対象の画像が建築内部空間であることを念頭において、実際の画像を詳細に検討し、各々の知覚イメージと相関の高い画像特徴との関係についての解釈を試みる。

そのために上で抽出された画像特徴が実際どのような建築空間の特徴に由来するものかを推測した。これは、アンケートによる評価値を参考にしながら、一つの画像特徴ごとにサンプル画像を実際に比較することと、各知覚イメージを表す印象語对各々の代表的画像を見比べ、その印象の差と、その知覚イメージと関連の深い画像特徴量の違いとの関係を推測する作業により行なった。さらに、画像特徴と知覚イメージとの関連の理解を深めるために、二つの特徴量を横軸、縦軸とする平面に各画像の対応する特徴量を座標値として持つ点をプロットして、二つの画像特徴とその知覚イメージとの相関を表す、二次元矩形領域ルール⁶⁾を導出した(図 11-14 参照)。例えば、図 11 において網掛け部分が二次元矩形領域であり、(画像特徴の分布がその領域内であれば)表 3 の画像特徴 1 と 2 の両方が「西洋的」と判断するルールを表している。

紙面の都合上、ここでは特に知覚イメージと特徴量の対応関係がわかりやすかった 2 つの知覚イメージ、「地域性」と「時代性」についての相関関係を分析した結果のみを示す。

6.1 「西洋的」対「東洋的」: 地域性

実験によりルールセットが得られたが、交差検証をしたところ検証精度は 66.7%で決定木ルールの分類精度も比較的高い値であることが確認できた。

表 3: 「地域性」に対して相関の強かった画像特徴

	相関の強い画像特徴	閾値	東洋的	西洋的
1	a チャンネル斜め高周波成分	80.921	小	大
2	中央窓領域全高周波成分	11.052	小	大
3	右側窓領域全垂直成分	370.886	小	大
4	左側窓領域 b チャンネルの標準偏差	67	大	小
5	天井窓領域 L チャンネルの標準偏差	30	大	小

表 3 は「地域性」という知覚イメージと相関の強かった画像特徴とその関係を簡単に表したもので、例えば、画像特徴 1 については、「a チャンネル斜め高周波成分」が閾値 80.921 よりも小さければその画像は「東洋的」、大きければ「西洋的」と判断されるルールであったことを示している。これら 5 つの画像特徴の傾向から「高周波成分」と「色彩の標準偏差」に関わる特徴量の影響が大きかったことが分かる。

表 3 の画像特徴 1,2 に関するルールをみると、「a チャンネルの斜め高周波成分」や「中央窓領域全高周波成分」の値が大きい、つまり「斜め方向の細かい直線成分」が多い、又は「全体として細かい模様」が多いと「西洋的」であり、逆にそういった成分が少ないと「東洋的」空間であることが解釈できる。これは実際の画像(図 16 参照)を見てみると木造軸組みなどの太い直線成分の多い「東洋的」空間に比べ、「西洋的」空間には左右の面や天井面に比較的細

かなテクスチャが多いというような空間特性があるために 2 つの値が増大すれば西洋的であるという結果が得られたと推測される。

また、表 3 から画像特徴 3 の「右側窓領域全垂直成分」について見てみると、「右側の全垂直方向の直線成分の和」という値が少ない方が「東洋的」で、多い方が「西洋的」であるというルールが読み取れる。このルールは縦じま模様が少ないなら、「東洋的」空間、多ければ「西洋的」空間であることを示している。また、画像特徴 4 の「左側窓領域 b チャンネルの標準偏差」から「青から黄色の散らばり具合」の大きい方が「東洋的」で、小さいと「西洋的」であるというルールも理解できる。これら二つのルールを実際の画像(図 16 参照)で確認すると、少し理解しづらいが、左右の面において柱と壁の色彩が異なることが多いという特徴が多く見られたために「東洋的」と判断され、逆に「西洋的」空間では左右の壁全体として色彩の統一感がある場合が多かったため、色のバラつき度合いが少なくなり、上記のルールが得られたと考えられる。

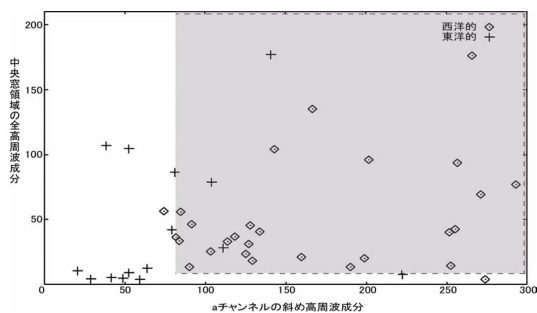


図 11: 「地域性」に関する画像特徴 1,2 の分布図
(網掛け部分は画像特徴 1 > 80.921 かつ 画像特徴 2 > 11.052 となる領域を表す)

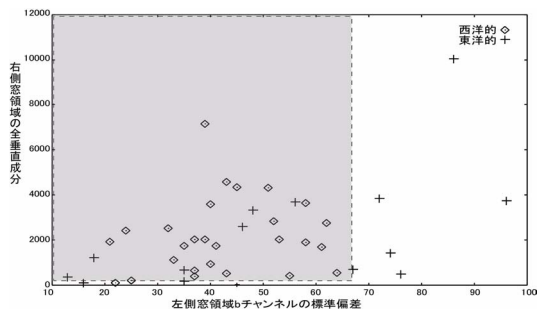


図 12: 「地域性」に関する画像特徴 3,4 の分布図
(網掛け部分は画像特徴 3 > 370.886 かつ 画像特徴 4 < 67 となる領域を表す)

図 11,12 はそれぞれ、上記の二つの画像特徴 1,2 と 3,4 の分布を表し、網掛け部分はどちらもルールが西洋的と判断する領域を示している。領域内全体の点の数に対する「西洋的」である画像に対応する点の数の割合(確信度という)はそれぞれ 89.3%、82.7%であり、比較的高い精度で「地域性」という知覚イメージを分類できることが確かめられた。

6.2 「現代的」対「歴史的」: 時代性

表 4: 「時代性」に対し相関の強い画像特徴

	相関の強い画像特徴	閾値	歴史的	現代的
1	左側窓領域 b チャンネルの標準偏差	67	大	小
2	天井窓領域 b チャンネルの平均値	188	大	小
3	天井窓領域の垂直高周波成分	37.543	大	小
4	b チャンネルの水平中周波成分	1144.5	大	小
5	b チャンネルの水平低周波成分	1815.96	小	大

表 4 にある相関の強い画像特徴を見比べると「時代性」の印象は「b チャンネル」、「水平成分」、「垂直成分」に関わる要素の影響が大きいたことが伺える。

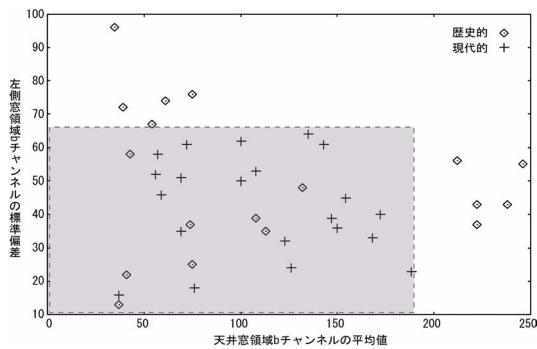


図 13: 「時代性」に関する画像特徴 1,2 の分布図と
(網掛け部分は画像特徴 1<188 かつ 画像特徴 2<67 となる領域を表す)

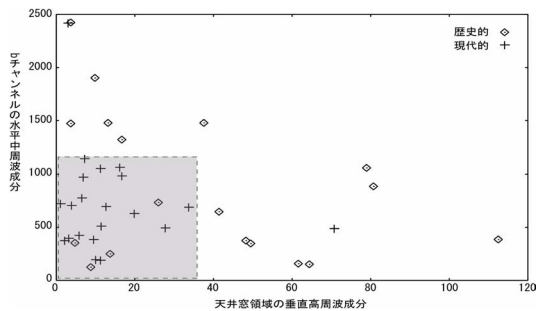


図 14: 「時代性」に関する画像特徴 3,4 の分布図
(網掛け部分は画像特徴 3<37.543 かつ 画像特徴 4<1144.5 となる領域を表す)

表 4 から画像特徴量の関係を見てみると、画像特徴 1 の値である「青～黄の色度のバラつき具合」が少なければ、「現代的」であり、逆に青～黄色のばらつきが多いと「歴史的」であるというルールが読み取れる。画像特徴 2”天井窓領域 b チャンネルの平均値”のルールは、空間内に黄よりも青色成分が多いと「現代的」であり、逆に黄の成分が多いと「歴史的」であるという印象を持つことを示す。アンケートの結果に基づいて実際の画像 (図 17 参照) をみると、実際に「現代的」なサンプルは青みがかっていることが多く、コンクリート打ちっぱなしの建物も多いため、壁の配色では色のバラつきが少ないという空間的特性が見られた。逆に「歴史的」空間の方はやや黄や茶褐色の配色が多く、(木造の建物が多く見られたため) 柱とそれ以外の部分の配色が異なることから色のバラつきが多くなり、上記のルールが得られたと推測される。

また、表 4 の画像特徴量 3,4 の関係を見てみると、“天井窓領域の垂直高周波成分”、“b チャンネルの水平中周波成分”の値が大きい、つまり「天井面の領域に細かい縦じま」が多いとか「青～黄の色度の横方向の縞模様」が多いと「歴史的」な空間であり、これらの値が小さいと、「天井領域には細かい縦じま模様」が少ない又は「青～黄の横縞模様」が少なく、「現代的」空間であるというルールが読み取れる。これは、実際の写真 (図 17 参照) を見れば「現代的」空間では、天井面がやや太いスパンで拭かれていて、その他の部分は平坦な面であるため、細かいテクスチャが少なくなっていると解釈できる。一方「歴史的」な空間では、天井や襖の棧、鳥居の桁などの水平・垂直方向の材の重なりが多いという空間特性があるため画像特徴量において水平・垂直方向の縞模様が多くなり、上記の関係になったのではないかと推測される。

図 13,14 はそれぞれ「時代性」に関する二つの画像特徴 1,2 と 3,4 の分布を表し、網掛け部分はどちらもルールが「現代的」と判断する領域を示している。領域内全体の点の数に対するその領域内の「現代的」な画像数の割合はそれぞれ、68%、83.7%と比較的高く、「時代性」という知覚イメージとの相関の強い画像特徴とし

て、空間印象を判断するのに十分使える特徴成分とみなすことができると考えられる。

7 まとめ

建築内部空間画像を対象として画像の印象と特徴量との相関関係を決定木アルゴリズムによって導出し、数多くの解釈可能なルールを得ることができた。たとえば、「地域性」のように1つか2つの画像特徴量を用いるだけで、比較的良い精度で西洋的か東洋的かという判断をすることができ、知覚イメージとの定量的な相関関係も明確に示すことが可能であった。

今回の実験では、垂直方向や斜め方向の直線成分が多いと「狭い」「奥行きのある」、逆に少ないと「広い」「奥行きのない」というルールがあり、実際の画像を見ても確認できたことから、人々は空間を知覚する際に、無意識に縦や斜め方向の線分の多さなどの特徴に注目し「開放感」や「奥行き感」といった印象を判断していた可能性が見出された。このように、ルールという立場でみることにより改めて特徴と印象の相関関係に気づいたことも多かった。

今後の課題として、多様な特徴をもつ十分な量のサンプル画像を使って学習の柔軟性・普遍性を上げること、色彩・テクスチャも含めて新しい画像特徴を見つけて行くことなどが挙げられる。

参考文献

- 1) 大津展之、栗田多喜夫、関田巖:”パターン認識”, 朝倉書店, 1996 年
- 2) Kinh Tieu and Paul Viola. "Boosting Image Retrieval", Proc. IEEE Conference on Computer Vision and Pattern Recognition, CVPR2000, Vol.1:228-235,2000.
- 3) W.Forstner. "A Framework for Low Level Feature Extraction", Computer Vision - ECCV '94, Vol. II, pp. 383-394, Ed.: J. O. Eklundh, LNCS 802, Springer, 1994
- 4) 柴田滝也、加藤俊一:”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 年 1 月
- 5) J.R. キンラン:”AI によるデータ解析 (Programs for Machine Learning)”, トッパン,1995 年
- 6) 福田剛志、森本康彦、徳山豪:”データマイニング (データサイエンスシリーズ③)”, 共立出版, 2002 年
- 7) 酒井幸市:”デジタル画像処理入門”, コロナ社, 2001 年
- 8) 中野宏毅、山本鏡男、吉田靖夫:”ウェブレットによる信号処理と画像処理”, 共立出版, 2000 年
- 9) 谷内田正彦:”コンピュータビジョン”, 丸善株式会社, 1995 年 3 月

注

- 注 1. 画像処理で言うテクスチャ(texture)とは、もともとの織物模様、布目にみられるように (たとえば、位相のずれた格子模様の境界など)、局所的な特徴だけでは識別や判別できない画像パターンを総称してテクスチャパターンという¹⁾。
- 注 2. Lab 色空間とは 1976 年 CIE より均等知覚色空間として提案された色座標であり、L は明度、a は赤から緑への色度、b は青から黄色への色度を表わす⁹⁾。
Lab は、以下のように、RGB を軸とする線形変換した XYZ 表色系によって定義される。
ここで、RGB 表色系から XYZ 表色系 (XYZ 三刺激値による色空間) への変換式は以下の通り。

$$X = 0.412391R + 0.357584G + 0.180481B \quad (6)$$

$$Y = 0.212639R + 0.715169G + 0.072192B \quad (7)$$

$$Z = 0.019331R + 0.119195G + 0.950532B \quad (8)$$

$$L = 116 \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - 16 \quad \left(\frac{Y}{Y_0} > 0.008856 \right) \quad (9)$$

$$a = 500 \left[\left(\frac{X}{X_0} \right)^{\frac{1}{3}} - \left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} \right] \quad (10)$$

$$b = 200 \left[\left(\frac{Y}{Y_0} \right)^{\frac{1}{3}} - \left(\frac{Z}{Z_0} \right)^{\frac{1}{3}} \right] \quad (11)$$

(但し、 $X_0 = 0.95045$ 、 $Y_0 = 1.0$ 、 $Z_0 = 1.08892$ であり、標準照明 D65(6500 度の色温度照明) における白基準値をあらわす。)

注 3. 情報エントロピー

正事例集合を P 、負事例集合を N とし、各々の事例の数を $|P|$ 、 $|N|$ とする。いまある分岐条件に対して、 P および N の中でその条件を満たす事例の数を P_P 、 N_P 、また P および N の中でその条件を満たさない事例の数を $P_{\bar{P}}$ 、 $N_{\bar{P}}$ とすると、分岐後のエントロピーは

$$ENT(x, y) = \begin{cases} 0 (x = 0 \text{ or } y = 0) \\ -x \log x - y \log y (x, y \neq 0) \end{cases} \quad (12)$$

と定義すると

$$\frac{P_T + N_T}{|P| + |N|} ENT\left(\frac{P_T}{|P| + |N|}, \frac{N_T}{|P| + |N|}\right) + \frac{P_F + N_F}{|P| + |N|} ENT\left(\frac{P_F}{|P| + |N|}, \frac{N_F}{|P| + |N|}\right)$$

によって定義される。この値が最小になるような分岐ルールを求める操作を再帰的に行っていく。

注 4. データマイニングツールとして、C5.0 については Clementine6.0 (SPSS 株式会社) を使用した。

注 5. ルールの精度

ルールセットの精度を計算するに当たり、一致マトリックスを使用する。これは実際のイメージ評価と学習モデルの予測の一致を行列の形で表される。例えば、「新しさ」についての例を示すと、

「新しさ」学習モデルの予測結果

	どちらでもない*	新しい*	古い*
どちらでもない	9	0	1
新しい	3	14	2
古い	0	4	17

行方向は実際の評価で、*の付いている列方向は予測結果を表す。つまり、9, 14, 17 と斜めに並ぶ部分は予測が正確であり、その他の部分は実際の評価と不一致であることを表している。例えば、一番下の古い行 (0 4 17) は実際はその中の 21 枚の画像が「古い」と評価された画像なのだが、予測モデルでは誤って「新しい」と判断された枚数が 4 枚であることを示す。なお、この学習モデルの予測精度は正しかったケース (9+14+17=40) の全体 (50) に対する割合であり、80 % を示している。

注 6. 交差検証

決定木などで生成された本 (学習モデル) の精度を検証する方法の一つで、検証用の事例集合が無い場合 (学習させる事例集合が小さい時など) に決定木の汎用性を検証する技法。「順としては、先ず学習用の事例集合をランダムに大ききの等しい N 個のブロック (データ集合) に分割し、その中の 1 ブロックを検証用データ、それ以外の N-1 個のブロックを学習用データとして二つのデータ集合に分け、次に

N-1 個の学習用データを用いて決定木を作成し、検証用データを用いてその決定木の予測精度を計算する。その処理を N ブロックすべてに対して行い、すべての予測精度の平均を未知のデータに対する木の精度の推定値とする (5, 6)。なお、本研究では N = 5 として検証精度を求めている。

注 7. 過学習

オーバーフィッティングとも言われ、ニューラルネットワークや決定木などの教師付き学習によって生成された学習モデルの汎用性が低下することを指す。学習用データに含まれる本来は意味の無いわずかなノイズ (データの偏り) を拾ったためにそのデータに対してはよい予測結果を示すものの、未知のデータに対しては予測結果が著しく悪くなる可能性がある。未知のデータに対しても汎用性のある決定木であるかどうか調べるために、交差検証などの手法が用いられる (5, 6)。

注 8. 使用画像の出典: コーロッパの家4 (講談社)、世界の建築5 (学研)、建築 20 世紀 part1,2, JA36 (新建築社)、現代建築集成/図書館 (明星出版)

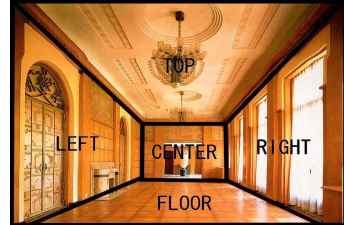
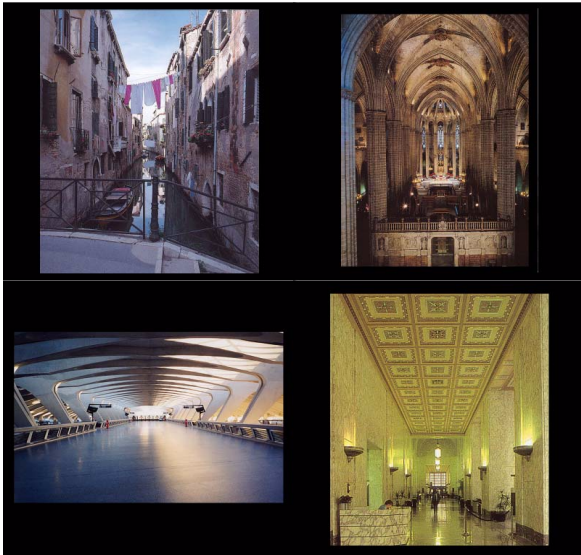


図 15: 5 つの窓領域

「西洋的」



「東洋的」

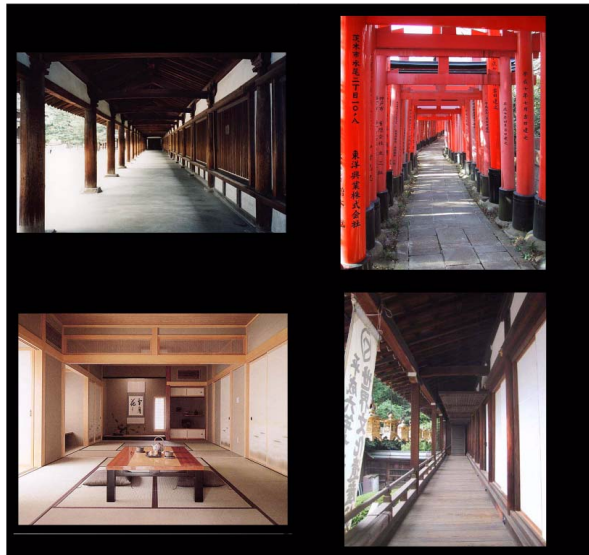
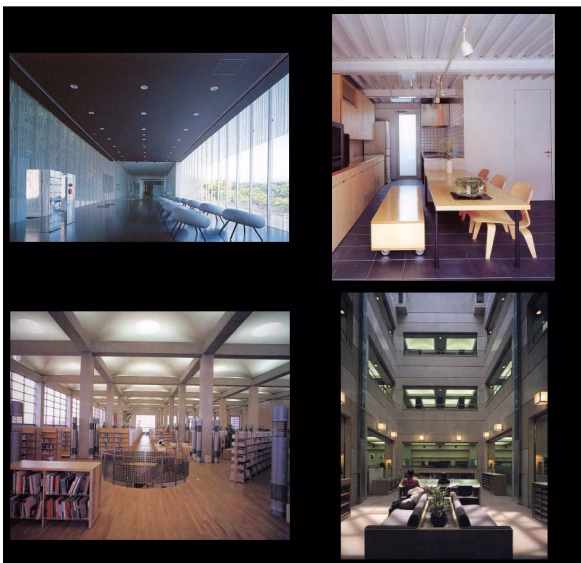


図 16: 知覚イメージ「地域性」の印象語対に対応するサンプル画像 (顕著な例) 注 8)

「現代的」



「歴史的」

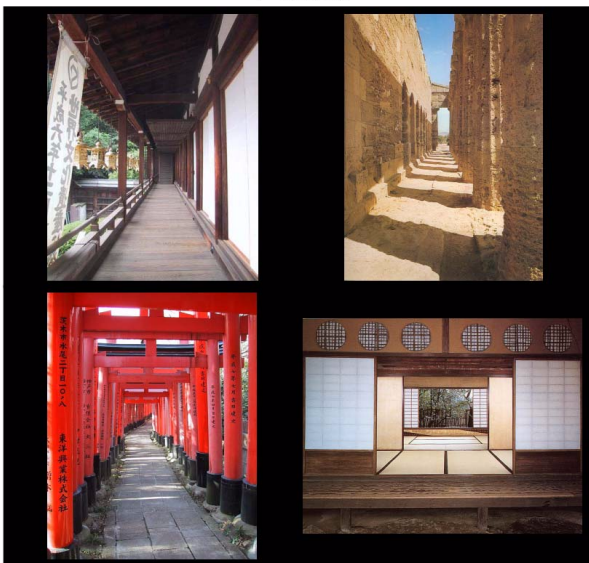


図 17: 知覚イメージ「時代性」の印象語対に対応するサンプル画像 (顕著な例) 注 8)