

## 構図情報を用いた画像クラスタリング

松田莉奈<sup>†1</sup> 秋岡明香<sup>†2</sup>  
 明治大学大学院<sup>†1</sup> 明治大学<sup>†2</sup>

## 1. はじめに

近年、スマートフォンや SNS の普及により画像への需要が高まっている。それに伴い、いくつかの画像検索の方法が提案されている。画像共有サイトである Flickr[1]のキーワード検索と色による検索、Google 画像検索[2]の類似画像検索等は高い精度で検索が可能であり、広く利用されている。しかし、似た構図の画像検索は一般的ではない。そこで本研究ではテキストチャ特徴を用いて画像を構図で分類し、似た構図の画像を提示することを目指す。本研究において構図の定義は、画像内の主要な物体の空間的な位置を単純な図で示したものとする。今回分類する構図は以下の 6 つに設定した。

- (1) 一点透視図法: 画面上に一点の奥行き(消失点)がある。領域は中心から横方向へ広がっている(図 1)。
- (2) 二点透視図法: 画面上に二点の奥行きがあり、2つの消失点がある。領域は中心から両端へと収束する(図 2)。
- (3) アオリ: 下から物を見上げる構図。下から上方向に領域が収束している(図 3(a))。
- (4) 俯瞰: 上から物を見下げる構図。下から上方向に領域が広がる(図 3(b))。
- (5) 水平: 画面を横に二分する構図。画面を二分するように一様に領域が広がっている(図 3(c))。
- (6) 日の丸構図: 画面中央に主要な物体が配置されている。中心から縦横に領域が広がっている(図 3(d))。

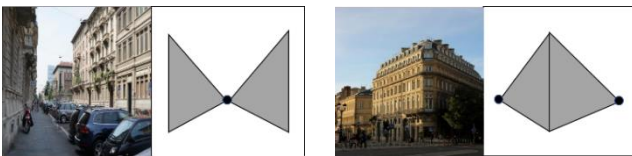


図 1 一点透視図法の例

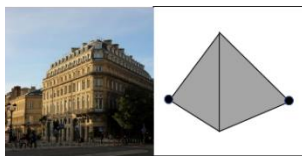
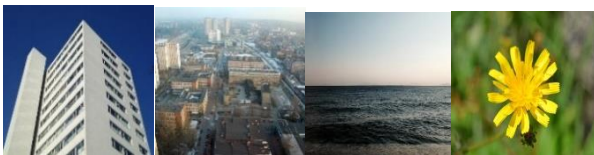


図 2 二点透視図法の例



(a) (b) (c) (d)

図 3 構図(2)~(6)の例

## 2. 提案手法

## 2.1 概要

入力画像を縦横 8 等分にし、64 個の領域ごとにテキストチャ特徴を求め、この値を基に特徴量を計算しブロックをクラスタリングする。同一クラスターのブロックが縦方向、または横方向に存在する数をブロック数とする。生成したクラスターは、物体の領域の広がり方を示す。主要な物体の領域の広がり方と構図の性質から、領域がどの構図の要素を持つかを定める。前処理として、画像サイズを一定にした上で 8bit グレースケール画像にする。

## 2.2 テキスチャ特徴計算

今回、テキストチャ特徴の表現方法として濃度共起行列[3]を使用した。濃度共起行列とは、濃度  $i$  の点から一定の位置関係  $\delta = (D_x, D_y)$  だけ離れた濃度が  $j$  である確率  $P_\delta(i, j)$  の行列である。本研究では  $\delta = (1, 0), (0, 1), (1, 1), (1, -1)$  に絞って共起行列を求める。

## 2.3 特徴量計算を基にしたクラスタリング

上で求めたテキストチャ特徴からブロックごとの特徴量を計算する。本提案手法ではエネルギー、慣性、エントロピーの 3 つを用いる。ブロックの一边の画素数を  $n$  として、エネルギー  $q_1$ 、慣性  $q_2$ 、エントロピー  $q_3$  は下式となる。

$$q_1 = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} \{P_\delta(i, j)\}^2$$

$$q_2 = \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} (i-j)^2 P_\delta(i, j)$$

$$q_3 = - \sum_{i=0}^{n-1} \sum_{j=0}^{n-1} P_\delta(i, j) \log P_\delta(i, j)$$

次に、似たブロックをウォード法によりクラスタリングする(図 4)。



図 4 画像のブロック分けとクラスタリング結果の例

## 2.4 構図への分類

まず 3.3 で求めた特徴量を用いて、主要な物体の領域と背景を切り分ける。多くの場合、背景は空が写っていたりぼかしかかっている等の理由で色の変化が小さいと考えられる。そこで、領域ごとの特徴量の平均を求め、平均値が最も小さい、つまり色の変化が小さい領域を背景とし、それ以外は主要な物体とする。

次に主要な物体の領域に対して、どの構図に当たるかを判断する。1つの領域は複数の構図を持てるが、日の丸構図は他の構図を持たないとする。

### 二点透視図法と一点透視図法

二点透視図法の特徴は消失点が2点存在する事である。しかし、それらが画面内に収まっているとは限らない。そこで、縦軸方向の最大値から左右のブロック数の変化量に注目した。領域の中でy軸方向にブロックの数が最大の位置をmaxとし、そのブロック数を $B_{max}$ と表すとする(図5左)。maxの左右にブロック数が0の座標があれば、そこを消失点とみなす。この消失点が2つ画面内に存在すれば二点透視図法である。消失点が2点存在しない場合は、x軸方向に走査する際の位置を $i(0 \leq i < 8)$ として、maxから左右のブロック数の変化量 $\Delta B$ を求める。次にmax左側( $i < max$ )の変化量の総数と、max右側( $i > max$ )の変化量の総数を求める。これらを用いてmaxから消失点までの距離 $l_L, l_R$ (図5右)を以下の式を用いて計算する。

$$\Delta B = \begin{cases} B_i - B_{i-1} & (i < max) \\ B_i - B_{i+1} & (i > max) \end{cases}$$

$$l_L = (max - i)B_{max} / \sum_{i=0}^{max} \Delta B_i \quad (i < max)$$

$$l_R = (i - max)B_{max} / \sum_{i=max}^n \Delta B_i \quad (i > max)$$

以上の式から出された $l_L, l_R$ が共に閾値以下、つまり適切な距離で領域が収束すれば二点透視図法とする。

一点透視図法は消失点が1点存在する。よって、上で求めた $l_L, l_R$ のどちらかが閾値以下であれば一点透視図法とした。

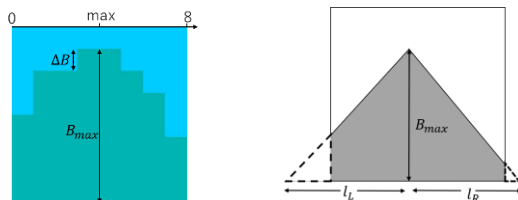


図5 二点透視図法の分類方法

### アオリと俯瞰

アオリは領域が下から上方向に収束している。そこで、x軸方向の領域の広がり最大となる位置を探す。この位置が画像の下部に位置していればアオリとする。上部に位置する場合は俯瞰とする。

### 水平

水平は上部、または下部に様に領域が広がっている。そのため、x軸方向のブロック数が最大である8がy軸方向に連続しているものを水平とした。しかし、この制約だけでは消失点が画面外に存在する二点透視図法等と誤認が多いため、領域の始まりのブロック数が8の場合に限定する。

### 日の丸構図

中央からx軸、y軸方向に広がる構図である。しかし、主要な物体の大きさや形が様々である。そこで、背景の領域に注目した。日の丸構図の多くの場合、主要な物体に焦点が合

わせられ背景には強いぼかしが入る。これを利用し、背景の領域で特徴量である慣性とエネルギーが閾値未満であれば日の丸構図とする。

### 3. 評価と考察

1つの構図につき10枚の画像を用意し、合計60枚の画像で分類をした。結果として、60枚中40枚が正しい構図の要素を含んだクラスタリング結果となった。具体的には、一点透視図法が7枚、二点透視図法8枚、アオリが9枚、俯瞰が2枚、水平が7枚、日の丸構図が7枚である。各構図ごとの判別例を図6に示す。



図6 出力結果の例

一点透視図法: 正しく分類できなかった原因として、画像をブロックで表した際に上手く主要な物体を切り出せなかったことが挙げられる。主な原因は影によるものが多い。

俯瞰: 精度が著しく低かった理由として、画像内に主要な物体が1つではなく、複数の集まり大きな領域として捉えられていたことが挙げられる。

水平: 多くは正しく判別されていた。一方で、雲がはっきりと映し出されている画像については二分割する境界が上手く取れず、正しくクラスタリング出来ていなかった。

日の丸構図: 正しく判別されていない画像は背景のぼかしが強すぎるため、他の画像での空の部分に近い特徴量になってしまい、日の丸構図に判定されていなかった。

あらかじめ6つの構図に分類したが、複数の構図の要素を持っていると見なせる画像もいくつか存在する。複数の構図を挙げられていたものもあったが、主要な物体が重なってしまい、正しく判別できていないものもあった。

### 4. まとめ

概ね正しい構図の要素が分類されていた。俯瞰と複数の構図の要素が含まれている場合については精度を上げるために別のアプローチが必要である。

謝辞 本研究はJSPS 科研費16KK0008の助成を受けたものである。

### 参考文献

[1] "Flickr". <https://www.flickr.com>  
 [2] "Google 画像検索". <https://www.google.com.jp/imhp>  
 [3] 内村圭一, 岩崎洋一郎, 松島宏典. 画像処理入門. 培風館, 2010