

構図による画像検索

松田莉奈^{†1} 秋岡明香^{†2}

概要：近年、スマートフォンと SNS の普及によってインターネット利用者の写真、画像への関心が高まっている。それにより、いくつかの検索方法が提案されている。しかし、似た構図を探す画像検索は一般的ではない。そこで、本研究ではエッジを特徴量とし、2つの画像間でエッジの距離を比較する。これを用いることにより、入力画像と似た構図の画像を検索するシステムを提案する。

1. はじめに

近年、スマートフォンと SNS が普及したことによりインターネット利用者の写真、画像への関心が高まっている。これに伴い、多様な検索方法が提案されている。画像検索の現状として一般的なものは、キーワード検索や類似画像検索が挙げられる。ともに技術の発展によって高い精度で画像検索が可能となっている。また、色やテクスチャによる検索も広く利用されるものとなってきた。このように現在、複数の検索方法が存在する。一方で、ユーザーが好みの画像をより高い精度で探そうとした場合、構図情報を考慮することは検索の質を高めることにつながると考えられる。なぜならば、構図は取り方によって作品の印象は大きく変わり、写真等において重要な位置にあるためだ。そのため、好みの色があるように好みの構図があると考えられる。しかし、上で挙げた方法では色やテクスチャ、形状、内容が似ているもの、または入力画像と同一のものを検索することが目的である。このように、構図情報による画像検索は未だに一般的なものとはなっておらず、研究の余地がある。

そこで、本研究では画像の色や、具体的な物体名に依存せずに、構図のみによって似た構図の画像検索を行うシステムを提案する。本研究において構図の定義は、画像上で主要となる物体の空間的な位置を単純な図で示したものとする。今回選んだ構図は次の7つに設定した。

日の丸構図

画面中央に主体となるものが配置されているものとする。

一点透視図法

画面上に一つの奥行きがある画像である。この構図では図 2(b)で画面中央に存在する点の位置を焦点と呼ぶ。

二点透視図法

画面の両端に奥行きがある。この構図は図 3(b)のように画面内、または外に焦点が2つ存在する。

アオリ

下から物を見上げる構図である。

俯瞰

上から物を見下げる構図である。

水平

画面を横に二分割する画像を水平とする。

遠近感のない構図

画面内に存在する物体が、空間的位置に差のない画像を集めた。



図 1 日の丸構図の例

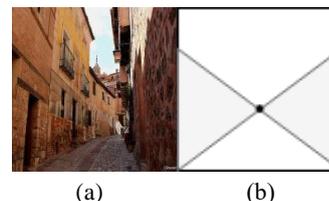


図 2 一点透視図法の例

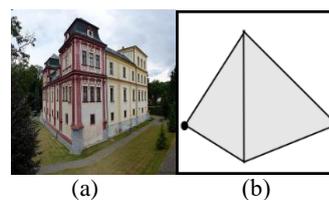


図 3 二点透視図法の例



図 4 アオリの例



図 5 俯瞰の例

^{†1} 明治大学大学院
Meiji University
^{†2} 明治大学
Meiji University



図 6 水平の例



図 7 遠近感のない構図の例

2. 関連研究

構図情報による類似画像検索のための特徴量抽出[1] この研究は言語化されにくい情報を検索したい場合、スケッチ画像をもとに類似画像を検索する手法を示している。検索のために用いる特徴量は構図情報のみを使用している。検索の目的は手元のない画像を検索することである。本研究は構図が似ている画像を検索することを目的としており、特定の画像を検索することが目的ではない。また、入力画像と検索される画像は共に写真という点で異なっている。

3. 提案手法

本研究では、入力画像をもとに用意された複数枚の画像から画像検索を行う。これにより、入力画像と構図が似ているもの上位 5 件を提案するシステムを作成する。本手法の流れを図 8 のフローチャートに示す。

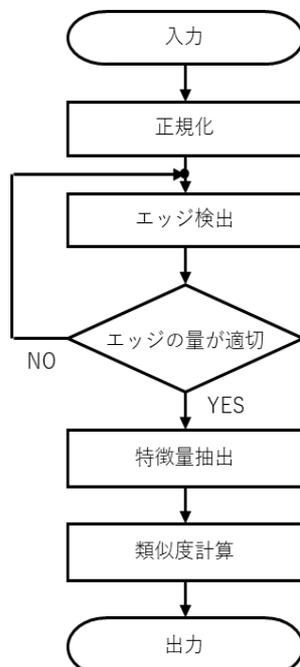


図 4 システムのフローチャート

3.1 正規化

本手法では画像からエッジを検出し、2つの画像間のエッジのユークリッド距離を比較することで類似度を計算す

る。この処理の際に、エッジが存在する座標をもとにして行うため、入力画像と検索される画像のサイズを一定にする必要がある。また、エッジ検出を行うため、画像をグレースケールに変換をする。これらを前処理とした。

3.2 エッジ検出

提案手法ではノイズに強く、検出位置が正確なキャニーエッジ検出器を用いた[2]。キャニーエッジ検出器とは通常の検出に加えて、細線化とヒステリシスしきい値処理を行う。細線化とは現在座標のエッジ強度が勾配方向にある2つの画素のエッジ強度より大きければエッジとする処理である。また、この勾配方向の座標を求めるために最近傍法を利用した。ヒステリシスしきい値処理とはあらかじめ上限値、下限値を設定し、各画素のエッジ強度が上限値以上であればエッジとして採用、下限値以下であれば不採用とする。また、上限値と下限値の間の値であれば現在位置の画素と隣接する8個と比較し、それらより値が大きければエッジとして採用する処理である[3]。これらの処理により、精度の高いものが取れるとともに、しきい値を変化させることでエッジのとれ具合を調整することが可能となる。

検出を終了したのちに、抽出されたエッジが構図として適切かどうかを判定する。今回、エッジである画素の数によって判断を行った。具体的な方法は以下のようにする。まず画像を二値化し、エッジ部分となる画素である、エッジ強度が0の画素数を出す。この値が事前に用意した上限値と下限値の間に収まっていれば適切。収まらなければ適切ではないとする。適切であった場合は次の処理に進む。適切でない場合は、フィルターのぼかし具合、しきい値を変化させた上で再度検出する。量が適切と判定されるまでこの処理を繰り返す。この際に使用した上限値と下限値は、複数の画像でエッジ検出を行い、構図らしくとれる時のエッジの画素数と明らかに構図ではない時の画素数から独自に決めた。提案手法でエッジを検出する手法について、フローチャートとして図9にまとめる。

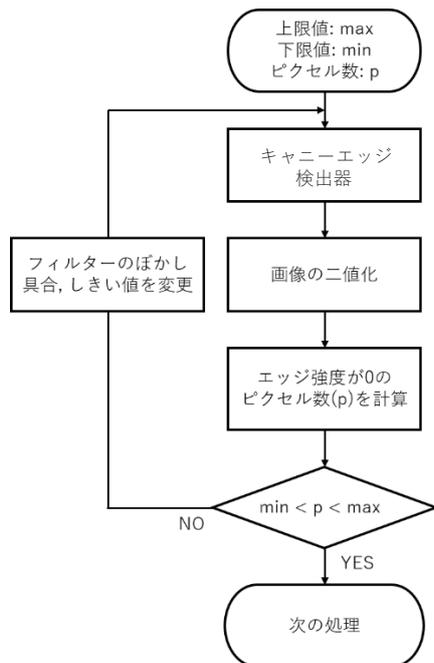


図 5 適切な量のエッジをとるためのフローチャート

3.3 特徴量抽出

本研究では構図を把握するために、画像上でエッジが存在する座標を特徴量とみなし、この特徴量を用いて構図をクエリとした構図検索を可能とする。入力画像と検索される画像でそれらの比較を行い、2つのエッジのユークリッド距離を計算する。これを特徴量として扱うこととする。座標の取り方は以下のようにした。まず画像上で図 10(a)のように、格子状になるように一定間隔で点を配置する。エッジがどこに存在するかを把握するため、これらの点から放射状に画像上を走査する。走査の終了条件はエッジである画素を発見するか、画像の辺にぶつかった時とする。終了した際に、各点は自身の番号、方向ごとの移動距離を保存する。

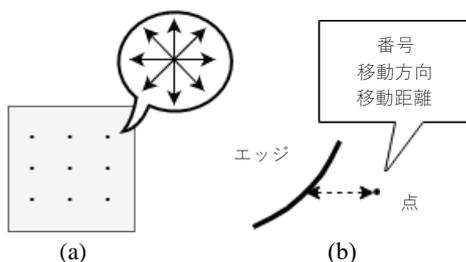


図 6 特徴量抽出のための手法

3.4 類似度計算

計算式は尾内らの研究 [1]を参考に実装を行った。大まかな方法としては、特徴量である点を対応するもの同士で比較を行う。これにより、2つのエッジのユークリッド距離が求められる。算出された距離を類似度として用いる。

具体的な手法は以下のとおりである。まず 3.3 節で求めた特徴量の点を用いて、対応する点同士で同じ方向の移動距離の絶対値をとる。この処理を全ての点、全方向で計算し、全てを足し合わせる。この値が小さいほど似ているとした。つまり、2つのエッジの距離が小さい程類似度が高く、大きい程類似度が小さいとしている。しかし、これだけでは検索の精度が低い。なぜならば、エッジ検出の際にしきい値処理によってノイズを抑え、構図として適切なエッジを検出したが、依然として画像には不必要なエッジが残っている。これらが検索の妨げとなっているためである。エッジが集中している箇所では点の移動距離が小さい可能性が高い。よって、3.3 節での点の移動距離が大きければ特徴量としての価値が高く、移動距離が小さければ価値が低いとすることで不要なエッジによる類似度への影響を抑えた。これを実現するために以下の条件を加える。対応する 2 点の探索範囲がともに δ_{min} 未満であれば重み 0、ともに δ_{max} 以上であれば 1 以上の重み ω 、それ以外は重み 1 を移動距離の絶対値にかける。これにより、各点の探索範囲の大小による点の情報価値を画像間の距離 D に反映させることにした。

以上を踏まえて、距離 D を次式のように定義する。この時、 m は点の数、 n は点からの方向、 d_i は i 番目の点の移動距離とする。

$$D = \frac{\sum_{i=1}^m \sum_{j=1}^n \omega_i d_j}{n \sum_{i=1}^m A_i}$$

$$\alpha_i = \sum_{i=1}^n d_i$$

$$A_i = \begin{cases} 0 & (\alpha_i < \delta_{min}) \\ 1 & (\alpha_i \geq \delta_{min}) \end{cases}$$

$$\omega_i = \begin{cases} 0 & (\alpha_{1i} < \delta_{min} \text{ かつ } \alpha_{2i} < \delta_{min}) \\ \omega & (\delta_{max} \leq \alpha_{1i} \text{ かつ } \delta_{max} \leq \alpha_{2i}) \\ 1 & (\text{それ以外}) \end{cases}$$

4. 評価

4.1 実験概要

20 人を対象にアンケートを実施した。

4.2 実験方法

本研究では 1 章で示した 7 つの構図を使って実験を行う。1 つの構図につき 5 枚の画像を用意した。入力画像はこれら 35 枚の中から順に 1 枚取り出したものとする。また、検索される画像は 35 枚から入力画像 1 枚を引いた 34 枚の画像を使用することとなる。従って、35 個の結果が出力される。最も理想的な出力結果は、上位 5 枚の中に同じ構図として含めた他 4 枚の構図を含むことである。35

枚の画像は以下の図 11~図 17 を使用した。上から日の丸構図、一点透視図法、二点透視図法、俯瞰、アオリ、水平、遠近感のない構図として選んだ。出力画像を提示し、被験者に入力画像と似た構図の画像番号を 35 枚それぞれに対して答えてもらうことにした。これによって、出力結果の精度と画像セットの適切さを調査する。



図 7 日の丸構図(構図 1)



図 8 一点透視図法(構図 2)



図 9 二点透視図法(構図 3)



図 10 アオリ(構図 4)



図 11 俯瞰(構図 5)



図 12 水平(構図 6)



図 13 遠近感のない構図(構図 7)

4.3 実験結果

アンケートとして、出力結果の 35 枚それぞれに対し、入力画像と似た構図だと思う画像の番号を選択してもらった。選択肢として 1~5 番の画像番号と該当なしの 6 つの選択肢を用意した。

図 18~図 24 では、画像ごとに似た構図と回答された画像の枚数をグラフ化した。黒い四角は、出力結果から想定していた似た画像を含む画像枚数の回答数を表している。例えば、画像番号 1_1 の出力結果には期待していた画像が 2 枚含まれていた。従って、2 枚を示すグラフ部分に黒い四角を加えている。また図 25 は、1 つの出力画像に含まれる似た構図の画像の枚数を全ての画像、回答をもとにグラフ化したものである。但し、図 25 においては空間を考慮していない構図 7 の結果を含めていない。

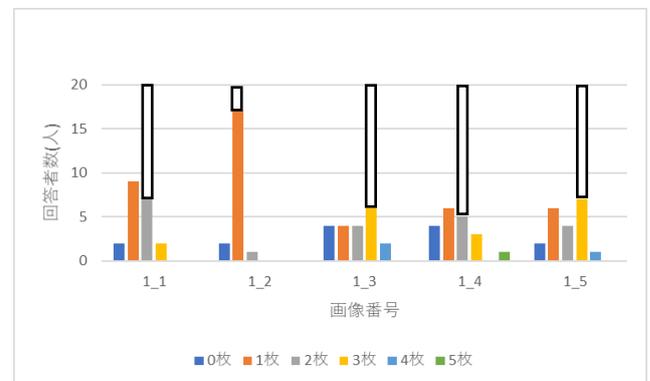


図 14 構図 1 のアンケート結果

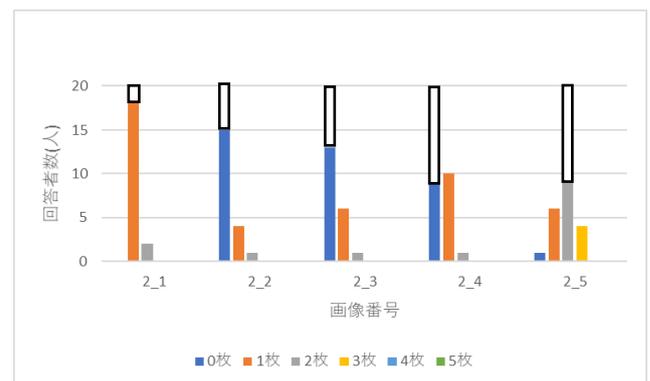


図 15 構図 2 のアンケート結果

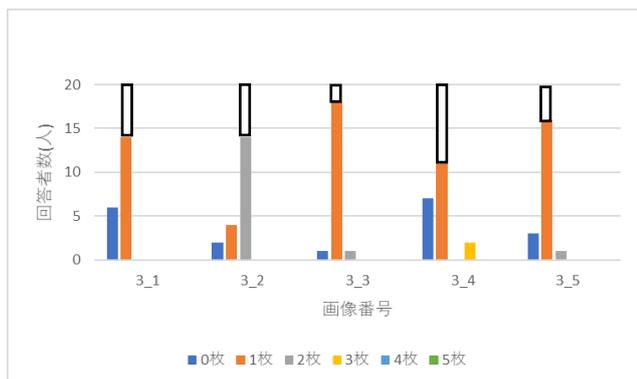


図 16 構図 3 のアンケート結果

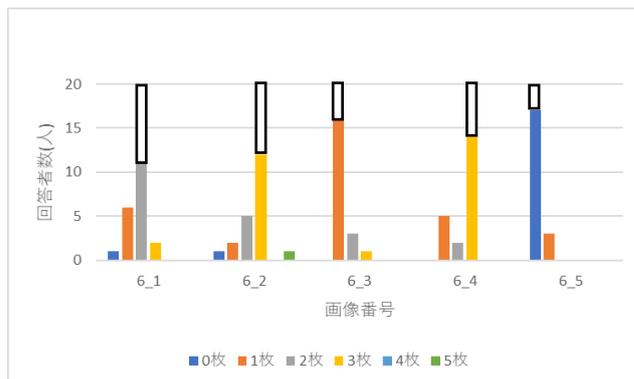


図 19 構図 6 のアンケート結果

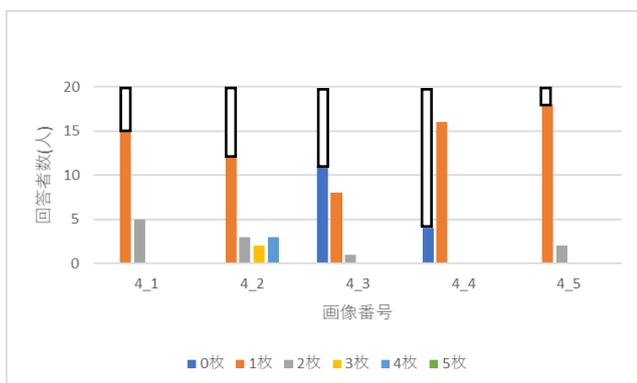


図 17 構図 4 のアンケート結果

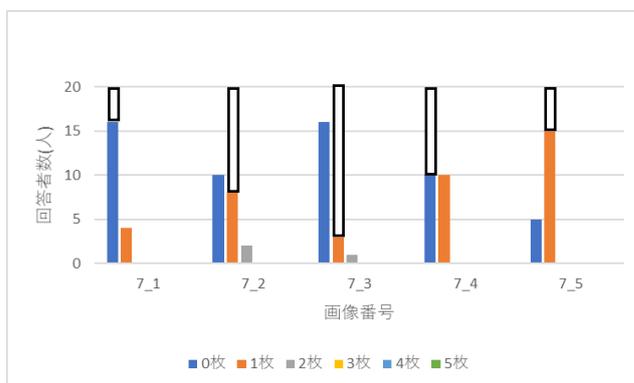


図 20 構図 7 のアンケート結果

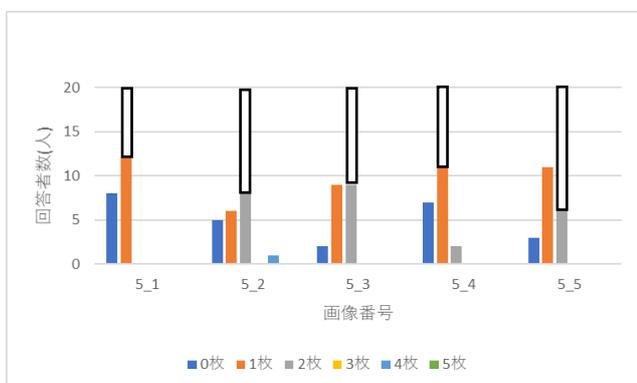


図 18 構図 5 のアンケート結果

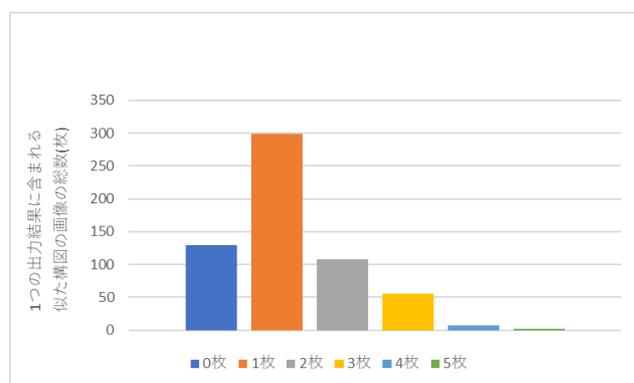


図 21 構図 1~6 の画像でのアンケート結果

結果として、構図 1, 5, 6 は他と比べて精度が良い。一方で、構図 2 や 7 については芳しくない結果となった。また、黒い四角が示している想定していた画像の枚数が、概ね回答者数の最も多い枚数と一致している。従って、今回の画像セットは適切な画像を選べたといえそうである。全体としては、600 枚中 470 枚が少なくとも 1 枚以上似た構図が上位 5 件に含まれるという結果になった。

5. 考察

アンケート結果を見ると、構図によって検索精度が異なることがわかる。構図1や構図6は似た構図の画像が1枚以上含まれていると答えた人がほとんどだったのに対して、構図2では半分近くの人が該当する画像が無いと答えている。このような結果となった理由を構図ごとに以下のように考察した。

5.1 構図1(日の丸構図)

理想的な画像が全体として上位5件に2~3枚ほど入ってきているため、他と比べると良い結果となった。しかし、図3の下段については他の画像と違った特徴があったため、まとめて言及する。他と比べて良い結果になった理由として、主体となる物体のエッジがどの画像でも中心に収まっており、他の構図との競合が起きづらかったと考えられる。



図 22 構図1の検索結果の一部

5.2 構図2(一点透視図法)

上位5件中0~2件が理想的な画像であり、結果はよくない。焦点点の位置によってエッジの形も変化する上、建物以外の人や車が多く入るとそれらのエッジが影響してくる。これにより、エッジが下に固まっている画像と上に固まっているもので類似度に差が出てしまっている。上段の画像では、上部が開けているが、中心部にエッジが入っている。そのために中心と横にエッジが広がっている画像が上位になっていると考えられる。



図 23 構図2の検索結果の一部

5.3 構図3(二点透視図法)

入力画像にもよるが精度はあまり良くはない。上位に想定していた画像が1~2枚という結果となった。一点透視

図法と違い、焦点点が画面内に収まっていないものが多い。そのため、画像上で焦点点までの距離が長く写っている部分が一点透視図法との類似度を高めていると考えられる。また、画面中央に主要な物体が配置されている構図となっているため、日の丸構図が候補の上位に入っていると思われる。



図 24 構図3の検索結果の一部

5.4 構図4(アオリ)

想定していた画像が1枚ほどしか入っていない点で精度はよくない。しかし、候補として直線的な形状で似た雰囲気をもつ建物が多く挙がっている。特に一点透視図法や二点透視図法との混同が多くみられた。

上段の画像はビルの先端より手前で両脇に空間がある。これによって候補の画像が左右に空間があってもなくても縦方向が強いものが上位に取れている。

下段になると右側にしか空間がない、また窓や装飾によって縦だけでなく横方向も取れている。これによって縦に強調されているものだけではなく、画面左側に細かなエッジが取れている画像が上位に入っていると考えられる。また、アオリは画面上部に焦点点をもつ一点透視図法の一つとみなすこともできる。特に、下段2番目の画像は焦点点の位置が入力画像とよく似ており、距離が小さくなったと思われる。



図 25 構図4の検索結果の一部

5.5 構図5(俯瞰)

上段の画像では中央にある丸い建物の形がはっきりとしているため、中央に物体が配置された画像が多く出ている。下段の入力は賽の目上にエッジが取れており、細かい縦横が取れている1番目2番目の距離が近くなった。また、上部に空間があることも1, 2, 3番目の画像が上位

に入った要因と考えられる。



図 26 構図 5 の検索結果の一部

5.6 構図 6(水平)

単純な図となっていることもあり、精度は良かった。分割線が画面を完全に二分し、位置もほぼ同じであるため、良い精度が得たと考えられる。俯瞰の画像が多く含まれる理由は空と地上の境が水平であるということが一つとして挙げられる。もう一つの理由として、上段、下段の入力画像は、ともに下部にエッジが集中している。このエッジと、賽の目状のエッジの距離が小さく計算されたためと考えられる。



図 27 構図 6 の検索結果の一部

5.7 構図 7(遠近感がない構図)

平面的な画像を選んだためエッジは写っている物体の模様、飾りに依存する。したがって、良い精度はでないことがわかる。

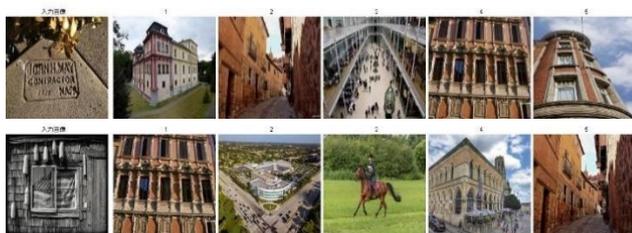


図 28 構図 7 の検索結果の一部

5.8 高頻度で上位 5 件に出現する画像

本実験において、ノイズがほぼなく取れているにも関わらず画像の精度が低い、上位 5 件に高い頻度で出現するといった画像が存在した。それらについて以下のようにまとめる。

5.8.1 図 33 について

下に示す図 33 を見るとわかるように他画像と比べてノ

イズがほぼなく、エッジが取れている。ところが図 33(b) のような結果であった。この画像で特徴量をとると、走査している点が探索を終了するのは中央にあるエッジにぶつかった時のみである。したがって、移動距離は大きくなり、重みも 1 より大きくなる可能性が高くなる。しかし、他のノイズが含まれる画像はノイズにぶつかった時にも探索が終了してしまう。そのため、移動距離は比較的小さくなり、ノイズの多さによっては重みが 0 か 1 になりやすい。この状態で図 21 と他の画像で類似度計算を行うと正しい類似度が得られないと予想される。大半の画像で理想的なエッジ検出が検出できれば精度の向上が望めるが、1 つの画像のみでは逆効果であることがわかった。



(a) (b)

図 29 入力画像とエッジ画像

5.8.2 図 34, 図 35 について

これらの画像は高い頻度で候補として上位 5 件に出現した。その原因を以下にまとめた。

図 34 の特徴は縦横一定間隔にエッジが取れていることである。さらに他にノイズがあまり乗っていない。そのため、どのようなエッジとでも安定して距離が小さくとれると考えられる。本来ならば、距離 D を計算する際に考慮した閾値によって弾かれるべきものであった。しかし、エッジの間隔によってしきい値で上手く弾けず、他の画像との距離 D が小さくなったと思われる。

図 35 も図 34 と同様に頻繁に出現する。図 34 と違い、一定の間隔ではないものの、画面全体に偏りなくエッジが広っており、同様の理由で距離 D が小さくなるものと考えられる。



図 30 高い頻度で上位 5 件に出現する画像



図 31 高い頻度で上位 5 件に出現する画像

6. まとめ

図 13 のアンケート結果から、上位 5 件には似た構図の画像が概ね少なくとも 1 枚以上含まれるということが出来る。よって、似た構図を検索するためにエッジを特徴量として用いることは有用と言えそうである。一方でサービス利用を想定すると、システムの精度としては大きな改善の余地がある。改善方法の一つとして、エッジ検出の精度を高めることが挙げられる。しかし、単に検出の精度を上げるだけではなく、いかに上手く画像の主体のエッジだけを切り出すかが重要となる。この出力結果とは別に、エッジ検出の前にフィルターをかけることでエッジが以前よりもよく取れた画像を用いて検索を行った。ところが、出力結果に含まれる想定していた画像の枚数は改善前と大差はなかった。従って、エッジ検出の精度を単に上げるのではなく、今回用意した構図に適したエッジを抽出することが求められていると考えられる。また、特徴量抽出の手法、類似度計算の改善も必要と思われる。今後は、これらを改善していくことを課題として開発を行っていきたい。

参考文献

- [1] 尾内克郎, 鈴木輝彦, 太原育夫. 構図情報による類似画像検索のための特徴量抽出. 第 10 回情報科学技術フォーラム
- [2] 内村圭一, 岩崎洋一郎, 松島宏典. 画像処理入門. 培風館, 2010, p.114-117.
- [3] 奥富正敏編. デジタル画像処理. CG-ARTS 協会, 2015, 232p.