

スマートフォン上のスクリーンキャプチャ探索支援手法

前島紘希^{†1} 阿部和樹^{†1} 中村聡史^{†1}

概要: スマートフォンのスクリーンキャプチャは、メモや思い出を残すためといった様々な用途で利用される。そのため、人によってはスクリーンキャプチャ画像の数が膨大になってしまい、必要な画像を探すのが困難になり、結果的に死蔵されてしまうという問題がある。我々はこれまでの研究で、スクリーンキャプチャ画像の特性について調査し、その特性に基づいた分類および、探索実験を行ってきた。その結果、画像特徴量を用いた探索がユーザにとって使い易いという傾向が出ていたものの、探索を支援するうえでは不十分であった。そこで本研究では、画像の領域分割を行うとともに、画像中の文字の量、顔の数といった別の特徴量を用いて分類を行う手法についてユーザベースの実験を行い、有用性を検証した。また、実験により探索にかかる時間は従来のスマートフォンの表示方法の方が短かったが、ユーザの満足度はこうした分類手法を利用した場合の方が高くなることが明らかになった。

キーワード: スクリーンキャプチャ画像、自動分類、画像特徴量

1. はじめに

現在、多くの人がパソコンやスマートフォンなどのデバイスを所持し、仕事やプライベートで利用しており、多くの人にとって欠かせないものとなりつつある。こうしたデバイスにはユーザが提供された情報を手軽に保存するための機能として、デバイス上に表示されている内容をそのまま画像として保存できるスクリーンキャプチャと呼ばれる機能が備わっている。スクリーンキャプチャは手軽に情報の保存を行えるため、様々な用途で利用されている。

ここで、パソコンは一般的にマルチウインドウ環境であり、多様な作業を同時並行して行うことが多いため、全画面に対するスクリーンキャプチャを行うよりは、特定のウインドウやウインドウの一部に対してキャプチャを行う機会が多い。実際、Gyazo[7]などのスクリーンキャプチャサービスは多くのユーザにより日々利用されている。一方、スマートフォンはシングルウインドウ環境であり、画面上には1つの作業に関するものが提示されていることが多いことや、スマートフォンのスクリーンキャプチャ機能の都合で、画面上の情報全てをそのまま保存することが一般的である。本研究では、後者のスマートフォン上でのスクリーンキャプチャ画像を対象とする。

さて、スマートフォンでのスクリーンキャプチャは、ウェブで調べた飲食店の情報や、目的地周辺の地図を画像として残したり、ゲームアプリ内においてなにか目標を達成した時のような嬉しいことがあった時の記念で保存したり、登録番号が必要なイベントに参加する際にすぐに思い出出すためのメモとして利用したりと、利用方法は多様である。ここで、スマートフォンでのスクリーンキャプチャは、パソコンでのスクリーンキャプチャとは異なり、コマ撮りのように何枚も連続して撮影されることも多い。パソコン上でも同様のスクリーンキャプチャが行われることもあるが、動画としてキャプチャされることが多く、スマート

フォンで利用されるようなコマ撮りのものは少ない。これは、キャプチャ環境やキャプチャの手軽さ、そして利用されるアプリケーションの性質の違いから、その内容は大きく異なっているといえる。

スマートフォンでスクリーンキャプチャを多用するユーザの写真フォルダには、膨大な量の画像が保存されている。一般的なスマートフォンの画像フォルダの提示方法は、単に画像を撮影時間順で並べているだけであるため、この中から目的とする画像を探し出すには、膨大な画像群の中から撮影日時を思い出しながら1枚ずつ内容を確認して探すという手間が発生する。このような手間を減らすためには、定期的に写真フォルダ内の不要な画像の消去を行えば良いが、消去してしまってから必要になって後悔することも多いうえ、消去作業自体には手間がかかる。また、アルバム機能などを用いて探索を容易にする方法もあるが、そもそもアルバム化が面倒であり、整理などは行われぬ。そのためスクリーンキャプチャ画像の探索は困難となり、結果的にフォルダ内で死蔵されてしまうことになる。これは、写真を手軽に撮影できるようになった結果、その量が膨大になり、探索が困難になったために結果的に死蔵してしまうことに似ている。

我々はこうした問題を解決するため、膨大な量のスクリーンキャプチャ画像を所有するユーザが、スマートフォン内の目的のスクリーンキャプチャ画像の探索を支援する手法を実現することを目的とした研究を行ってきた[13]。この研究では、スマートフォンのスクリーンキャプチャ画像には短い時間で連続して撮影されることがあることや、似た構図の画像が多く存在するといった固有の特徴に着目し、それらの特徴に基づいて撮影時間と画像特徴量をそれぞれ利用した2種類の画像分類手法を提案した。画像探索実験を行なった結果、画像探索にかかる時間に差はほとんどないものの、ユーザにとって撮影時間を用いた分類手

^{†1} 明治大学
Meiji University

法は探索がしづらく、画像特徴量を利用した分類は探索がしやすくなる傾向があった。しかし、画像特徴量を用いた場合でもクラスタリングの際にクラスタ内の画像の枚数に差が出てしまうという問題があった。

そこで本研究では、画像特徴量を用いた新たなスクリーンキャプチャ画像の分類について、以下の3つの手法についてユーザベースの実験によりその比較検討を行う。

- スクリーンキャプチャ画像をヘッダ部分、メイン部分、フッタ部分に分割し、ヘッダ部分とフッタ部分に重み付けをすることでアプリケーションによる分類を行う手法
- スクリーンキャプチャ画像内のキャラクタの顔部分を抽出し、その類似度を計算し、類似したキャラクタの存在によって分類を行う手法
- スクリーンキャプチャ画像内のテキストの量によって分類を行う手法

また、この3つの分類手法を組み込んだスクリーンキャプチャ画像探索システムの実装を行い、利用実験によってその有用性を検証する。

2. 関連研究

膨大な量のプライベート画像の探索においては、どうしても探索の手掛かりとなる情報が必要となる。こうした手掛かりとなる情報を牛尼ら[1]は、e-mailと撮影画像という異なる種類のメディアを、推論ネットワークを利用して時間的な側面から統合し、それぞれのメディアの特徴を用いて検索を行うことができるライフログ検索システムを提案している。スマートフォン上でのスクリーンキャプチャは、インターネットやアプリケーションを撮影したものが多いため、現実と連動していないことが多い。そのため、スクリーンキャプチャの探索においては、こうした情報は使いにくいと考えられる。

一方、岸野ら[3]は、稀に起こる事象に反応して生じる脳波である「事象関連脳電位」を用いて、自動撮影した大量の写真から非日常的な、その人にとって興味深い写真を選別する手法を提案している。スマートフォンにおける実装を考えた場合、脳波などを利用するには計測に別の機器を要するためコストがかかる点、スクリーンキャプチャは画面上の内容を保存するものであるため、スマートフォンに内蔵されているセンサはあまり有用ではないと考えられる点などから、人に関するセンサを用いるのは難しいと考えられる。

膨大な画像の探索を支援する手法として、Gomiら[4]は、前処理として大量写真を多階層にクラスタリングし、代表写真を抽出したのちに「平安京ビュー」[5]を用いて、各クラスタを表す長方形領域に写真群を配置する写真ブラウザを実現している。その際、クラスタ内の画像の数によって

代表画像のサイズが変化するようにになっている。スクリーンキャプチャ画像でもそうしたグループ化が効果的に働くと考えられるが、スマートフォンという限られたディスプレイ環境では、大画面を想定したビューは適していない。しかし、代表画像のサイズをクラスタ内の画像の枚数によって変化させる方法は応用可能であると考えられる。

また、森ら[6]は、ユーザが提示したオブジェクト画像が有する特徴量の中から、悪影響を及ぼす特徴量を使用しない類似画像検索システムを提案している。スクリーンキャプチャ画像でもこのような特徴量を用いることで目的にあった画像の探索ができる可能性がある。

また、画像認識や文字認識に関する研究は多数行われている。上坂ら[8]は、人物の顔から様々な特徴量を抽出し、それらを用いて写真の自動分類を行っている。この研究で抽出し、利用された特徴量として、色を表すカラーヒストグラムと、形・テクスチャを表す3点間コントラストが挙げられている。我々の以前の研究で、スクリーンキャプチャ画像にも人物の顔があるが、アニメ風の絵であることが多いことがわかっている。そのため、本研究ではアニメ顔のデータセットを用いて顔認識を行い、顔画像の抽出を行っている。

画像からの一般物体認識については、Csurkaら[12]のBag-of-Keypointsを用いた物体認識が有名である。また、上藤ら[11]は、機械学習の手法であるMultiple Kernel Learningを用いて、カテゴリごとに複数の特徴量に最適な重み付けを行い統合する手法を提案した。複数の画像特徴量を組み合わせた分類を行う方法はスクリーンキャプチャ画像に対しても利用でき、効果的な分類ができる可能性がある。

永橋ら[9]は、あるカテゴリにおいて、対象カテゴリに関わる範囲を前景、それ以外の範囲を後景と定義し、画像を前景と後景に分割したうえで、前景、後景のそれぞれについてBag-of-Featuresの特徴を抽出し、前景情報のみからの特徴量と、前景情報と後景情報の両方の特徴量を用いた場合の認識率の比較を行っている。それによって、従来のBag-of-Featuresによる認識手法と比較して3.8%の認識率の向上が見られている。スクリーンキャプチャ画像の中ではゲームアプリが特に前景と後景が分かれているものだと考えられるため、大まかな分類を行なった後にこのような分類が利用できると考えられる。

中島ら[10]は、Bag-of-Featuresで使用される特徴量を従来のSIFT特徴量に対し、色や形状を表す特徴量を用いたGeneralized adaptive modelと呼ばれる新しいフレームワークを提唱している。スクリーンキャプチャ画像でも色や形状を考慮した特徴量を用いることで探索に有用な分類が行える可能性がある。

趙ら[14]は、2枚の画像全体が類似していることを示す画像の類似性と、同じ物体を対象にしているが2枚の画像の

一部分しか類似していないことを示す画像の隣接性の2種類の関係性を同時に利用し、画像検索を行う手法を提案している。この研究では、類似性の計算に局所特徴と Color Coherence Vectors(CCV)を、隣接性の計算に SURF を使用している。

文字認識についての研究としては、森ら[2]の研究などが挙げられる。森らの研究では、テレビなどのテロップ文字に対して認識を行っている。テロップ文字は背景との境界が曖昧になってしまい、うまく識別できないという問題があったため、この研究では背景情報を利用することでより高精度に認識を行なっている。

我々は、画像特徴量として以前の研究では SIFT を用いており、本研究では、領域分割を用いた手法で画素値の度数分布を、文字認識を用いた手法で光学文字認識(OCR)を使用する。

3. スクリーンキャプチャ画像分類手法

現在のスマートフォンの画像提示の手法では、ユーザはスクリーンキャプチャ画像のサムネイルから、目的の画像を見つけなければならない。しかし、スクリーンキャプチャ画像が膨大な量になると、その作業の負担も大きくなる。目的のスクリーンキャプチャ画像を撮影した時期、前後に撮影した他の画像などの背景情報を覚えていれば多少探索の難易度は緩和されるが、そうした情報を覚えていない場合、全ての画像をしらみつぶしに探索していかなければならない。

我々は以前の研究[13]において画像特徴量を用いた分類対象とする画像が、その画像自体に意味があるかについて取り組んでいたが、その時の様子やその後の調査により以下の3つがスマートフォンのスクリーンキャプチャ画像探索において重要であると考えた。

- スマートフォンアプリには、コンテンツとは関係ないそれぞれ特有のヘッダ、フッタが存在するため、それぞれを分類して認識すること
- 探索対象の画像が SNS でのやり取りやメールなどといったテキスト主体であるのか、絵など画像が主体であるかの違い
- スマートフォン上で頻繁にスクリーンキャプチャされるゲームアプリは、ゲームという世界と日常的に関わるライフログのようなものであるため、登場するキャラクターの顔が重要であること

上記の点を踏まえ、本研究では領域分割、顔の数、文字数という3つの探索の手がかりを用意し、スクリーンキャプチャ画像の探索の際の支援を行う。検討する分類手法の具体的な内容は次節の通りである。

3.1 領域分割を用いた分類手法

スマートフォンのスクリーンキャプチャ画像は特定の

アプリケーション使用時に撮影されることが多い。また、同じアプリケーション内では基本的に共通したヘッダやフッタが用いられている。そのような特徴を利用して、図1のようにスクリーンキャプチャ画像をヘッダ部分、メイン部分、フッタ部分に分割し、ヘッダ部分とフッタ部分に重み付けをして画像類似度を計算することで、スクリーンキャプチャ画像を分類する。この手法によって、スクリーンキャプチャ画像をアプリケーションごとに分類することができ、目的のスクリーンキャプチャ画像のアプリケーションがわかれば容易にアクセスできるようになると考えられる。また、本手法により、アプリケーションは異なるものの、類似した絵(同一の絵師による作品)で横断的に類似検索することが可能となる。

この手法の実装では、まずスクリーンキャプチャ画像の上下10分の1の領域をそれぞれヘッダ部分、フッタ部分と定め、残りの部分をメイン部分と定める。(この割合については様々なアプリから判断した)その後、ヘッダ部分、フッタ部分、メイン部分のそれぞれについて2枚の画像を比較して類似度を計算し、それらの類似度に重みをつけて足し合わせることで2枚の画像の類似度を計算している。なお、ヘッダ部分、フッタ部分、メイン部分の重み付けの割合は2:2:1となっている。最後に、2枚の画像の類似度が閾値を超えた画像を同じクラスに分類する。この手順を全ての画像のペアに適用するといった手順で分類を行う。

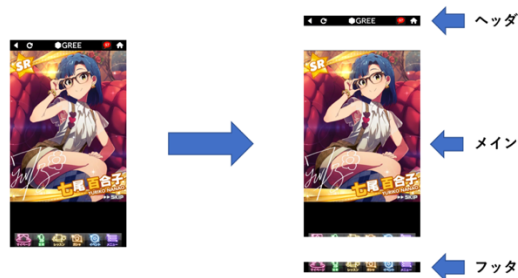


図1: スクリーンキャプチャ画像の領域分割

© GREE, Inc

3.2 顔認識を用いた分類手法

スマートフォンでのスクリーンキャプチャ画像(特にゲームの画像のライフログとして利用する場合)には、キャラクターが写っている画像が多く存在する傾向にある。そこで、図2のようにスクリーンキャプチャ画像内のキャラクターの顔認識を行いキャラクターの顔画像の数を数えることで分類を行なう。この手法によって、キャラクターが多数存在しているスクリーンキャプチャ画像へのアクセスが特に容易になると考えられる。

この手法の実装では、まずスクリーンキャプチャ画像内から顔画像を抽出する。この際にアニメ顔画像のデータセット[15]を利用した。その後、画像から抽出された顔画像をカウントすることでその画像の顔画像の数とした。なお、

この手法を拡張することにより、人物による横断検索を行うことも可能であるが、今回の分類においては単純に顔の数のみを利用するものとした。

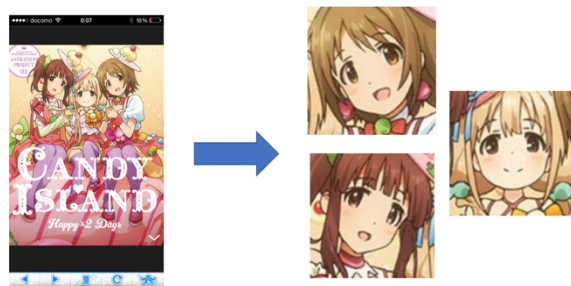


図 2: スクリーンキャプチャ画像からの顔認識

©DeNA Co.,Ltd

3.3 文字数を用いた分類手法

テキストに関するスクリーンキャプチャ画像においてまず考えられるのが、そこに何が書いてあったのかというキーワードベースでの検索である。もちろんこうしたキーワードベースの検索は重要であるが、そのキーワードを思い出せないことも多く、容易ではない。ここで、何が考えてあったかは覚えていないもののどれくらいの文字が書いてあったかは思い出せることがある。実際スマートフォンのスクリーンキャプチャ画像には、画像内に提示されている文字数が少ないものから多いものまで様々なものが存在している。そこで、光学文字認識(OCR)を利用して文字数を計算し、その数によって分類を行う。この手法によって、画像内に満遍なく文字が散らばっている画像や、反対に文字がほとんど存在しない画像の探索が容易になると考えられる。

この手法の実装では、まずスクリーンキャプチャ画像に対して OCR を適用する。それによって画像内に存在する文字列が取得できるため、その文字列の数をカウントすることで画像の文字数とした。

4. プロトタイプシステム

3章で検討した3つの手法を取り入れたスクリーンキャプチャ画像探索のプロトタイプシステムを実装した。

本来はスマートフォン上で実装するべきであるが、今回はその分類の利用のされ方を明らかにすることを重視したため、デスクトップのアプリケーションとして、Processing を使用して実装した。プロトタイプシステムでは、分類手法の結果を格納したデータベースにアクセスし、その中から必要なデータにアクセスすることで結果を出力している。なおデータベースの生成部分は Python を用いて実装している。

画像探索画面では、図3のように各画像のサムネイル画

像が表示されている。画面上部の四角の中をクリックすると各分類手法によって絞り込みを行う画面になる。この画面内の一番左の四角をクリックすると、領域分割を用いた手法による分類の選択画面に遷移し、図4のように各クラスタの代表画像が表示されるようになっていく。この中から探索したい画像に近いものを選ぶことでそのクラスタの画像が表示されるようになる。真ん中の四角をクリックすると文字数による分類の選択画面に遷移し、図5のように画像内の文字の数が表示され、選んだ文字数に対応する画像群が表示される。一番右の四角をクリックすると顔画像による分類の選択画面に遷移し、図6のように画像内の顔画像の数を選択することで、選んだ顔画像の数に対応する画像が表示されるようになる。

いずれかの分類手法で絞り込みを行うと、絞り込みを行った手法に対応する四角は図7のように赤くなり、どの手法を選択しているかがわかるようになっていく。図3または図7の画像探索画面で画像を選択すると、図8のように選択した画像が大きく表示されるようになっていく。

5. 利用実験

スクリーンキャプチャ画像の探索において、今回用意した手法がどのように利用されるかを検証するため、実際にユーザが記録したスクリーンキャプチャ画像を用いて実験を行う。実験協力者はスクリーンキャプチャ画像を500枚以上所持している大学生、社会人で合計9名である。実験に使用した画像は実験協力者が所持しているスクリーンキャプチャ画像とした理由は、他人のスクリーンキャプチャ画像の探索はそもそも困難であるためである。

探索実験にはプロトタイプシステムを利用し、まず実際にプロトタイプシステムを使用してもらって操作に慣れてもらった。その後、目的ベースの探索タスク、特定画像の探索タスクいずれかをランダムに提示し、探索を行ってもらった。なお、探索実験では従来の iPhone のようにただ単純に画像を並べる既存手法と、3種類の手法によって絞り込みを行う提案手法の2種類の手法で行なった。

目的ベースの探索タスクを提示する場合は、まずそのタスクに当てはまるスクリーンキャプチャ画像を思い浮かべてもらい、その画像を探索してもらった。なお、思い浮かべてもらった画像と違う画像を選んでも良いこととなっている。使用したタスクは、既存手法と提案手法で時系列が異なるようにした。具体的には以下の表1の8つのタスクを用いた。

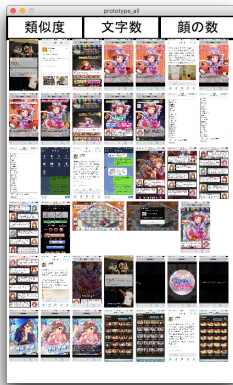


図 3：探索画面

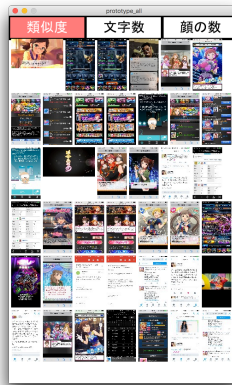


図 4：領域分割による分類時の画面

類似度	文字数	顔の数
文字数		
なし	~50	50~150
150~250		250~

図 5：文字数による分類時の画面

類似度	文字数	顔の数
顔の数		
なし	1	2
3		3以上

図 6：顔の数による分類時の画面

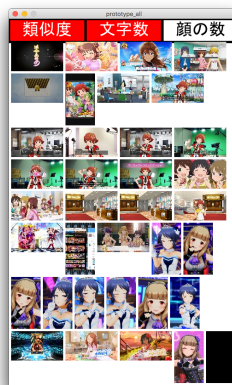


図 7：絞り込み後の画面



図 8：画像選択後の画面

©GREE, Inc., ©DeNA Co.,Ltd., ©窪岡俊之,
 ©BANDAI NAMCO Entertainment Inc., ©Akatsuki Inc., ©TYPE-MOON / FGO PROJECT

表 1：目的ベースの探索タスクで使用した内容

最近の嬉しかった画像	昔の嬉しかった画像
最近の面白い画像	昔の面白い画像
昔の後で使うために撮影した画像	最近の後で使うために撮影した画像
昔の悲しかった画像	最近の悲しかった画像

また、特定の画像を提示する場合（特定画像探索タスク）は、筆者が選んだスクリーンキャプチャ画像を既存手法と提案手法で4枚ずつ、合計で8枚探してもらった。

実験では画像を発見するまでの時間を計測するとともに、3種類の分類手法を利用した順番を記録した。また、実験ごとに下記のアンケートについて-2~2の5段階評価で回答してもらった。

- このシステムでの探索の難しさはどの程度に感じましたか？
- 最初に思いついた画像と見つけた画像との違いはどの程度だと思いますか？
- 自分の想定通りに絞り込みができましたか？

6. 実験結果

6.1 目的ベースの探索タスク

探索タスクについて、既存手法（従来型の探索）と提案手法（プロトタイプシステム）での探索時間の平均をユーザごとに算出したものを表2に示す。

表 2：目的ベースの探索タスクによる探索時間の平均

	既存手法	提案手法
A	1:00	1:07
B	4:12	3:38
C	3:02	0:55
D	2:31	3:41
E	1:00	1:11
F	1:05	4:16
G	0:55	1:37
H	1:01	0:54
I	0:55	0:54
平均	1:58	2:21

表2を見ると、ユーザA, D, E, F, Gは既存手法の方

が早く画像を見つけることができている。ユーザ B, C, H, I は提案手法の方が早く画像を見つけることができている。これは、提案手法を用いて少ない絞り込み回数で探索タスクに該当する画像を見つけることができれば既存手法よりも早く探索を行えるが、提案手法で少ない回数で見つけられなかった場合には、多くの絞り込み方法を試しその結果を見ていく必要があるために既存手法よりもむしろ探索時間が長くなってしまったと考えられる。つまりユーザがよりシステムに慣れることである程度探索時間は短くなると考えられる。また文字や顔の誤認識によってユーザが思った場所と違う場所に分類されてしまっていることが考えられるため、文字認識、顔認識の精度を上げることで容易な探索が可能になると考えられる。

続いて、目的ベースの探索タスクを与えられた際のアンケート結果を図 9 から図 11 に示す。

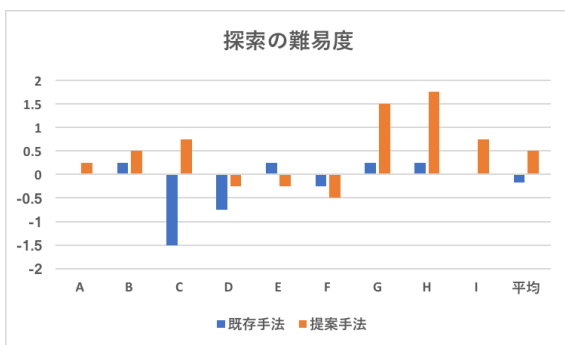


図 9：目的ベースの探索難易度についてのアンケート結果

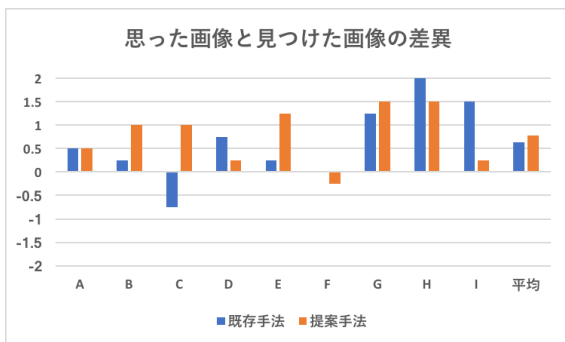


図 10：目的ベースの探索タスクにおける思った画像と見つけた画像の違いのアンケート結果

図 9, 図 10 のグラフは既存手法を青、提案手法を橙色の棒グラフで表現しており、図 9 はスコアが高いほど探索が簡単であること、図 10 はスコアが高いほど思った画像と見つけた画像の差異が小さく、図 11 はスコアが高いほど提案手法で想定通りに絞り込みができたことを表している。これらの図を見ると、提案手法はユーザにとって既存手法よりも高い評価を得ていることがわかる。探索にかかった時間の結果と合わせて見ると、探索タスクを提示した探索

では探索の時間はユーザ毎にばらつきはあるものの、ユーザの評価では提案手法の方が高いという結果が得られた。

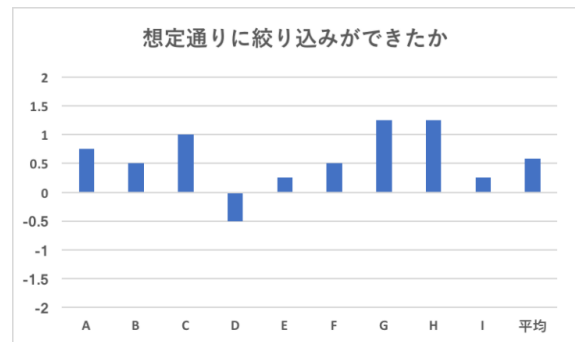


図 11：目的ベースの探索タスクにおける、想定通りに絞り込みができたかのアンケート結果

続いて目的ベースの探索タスクについて、3種類の探索手法を使った回数の合計と複数の手法を併用した回数を表 3 に示す。

表 3 を見ると、ユーザによって利用する手法に大きな違いがあることがわかる。また、複数の手法を併用して探索を行うユーザもあり、多様な探索のパターンを用意することは有効であることがこの結果から示唆される。

表 3：3種類の探索手法の使用回数の合計と併用の回数

	類似	文字	顔	合計数	併用数
A	12	5	4	19	2
B	11	11	15	31	15
C	8	1	0	9	0
D	68	22	42	96	36
E	3	5	3	10	1
F	4	6	5	14	1
G	20	4	0	24	0
H	0	13	5	18	0
I	16	17	16	32	13

6.2 特定画像探索タスク

特定画像について、既存手法と提案手法での探索時間の平均をユーザ毎に算出したものを表 4 に示す。

表 4 を見ると、ユーザ B 以外が既存手法の方が早く画像を見つけることができている。ユーザ B のみが提案手法の方が早く画像を見つけることができている。この結果から、特定の画像を探索する際は既存手法の方が短時間で画像の探索を行うことができることがわかる。

表 4：特定画像における探索時間の平均

	既存手法	提案手法
A	0:49	0:50
B	1:54	1:11
C	0:51	2:57
D	1:44	2:03
E	0:31	0:52
F	0:18	1:32
G	1:24	3:29
H	0:32	1:49
I	0:36	1:38
平均	1:04	2:21

続いて特定画像探索タスクの際のアンケートの結果を図 12、図 13 に示す。

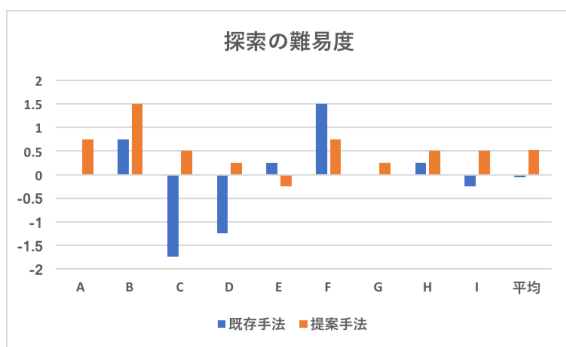


図 12：特定画像探索タスクにおける難易度についてのアンケート結果



図 13：特定画像探索タスクにおける想定通りに絞り込みができたかのアンケート結果

図 12 は図 9 と同様に、スコアが高いほど探索が簡単であること、図 13 は図 11 と同様にスコアが高いほど提案手法で想定通りに絞り込みができたことを表している。この結果と探索時間の結果より、特定画像を探索する際は多少探索に時間がかかっても絞り込みを行えるようにすることでユーザにとって探索が簡単になることがわかった。

続いて特定画像探索タスクについて、3 種類の探索手法

を使った回数の合計と複数の手法を併用した回数を表 5 に示す。

表 5：特定画像探索タスクにおける 3 種類の探索手法の使用回数の合計と併用の回数

	類似	文字	顔	合計数	併用数
A	6	19	5	26	4
B	7	6	5	14	4
C	13	23	26	42	16
D	12	20	14	38	8
E	2	4	3	9	1
F	0	6	6	11	1
G	20	15	26	54	7
H	3	12	5	20	0
I	11	29	32	55	17

表 5 を見ると、目的ベースの探索タスクに比べて絞り込みの合計数は大きく変化しなかった。しかし、それぞれのユーザがどの手法を多く使うか、複数の手法を併用するかどうかは目的ベースの探索タスクと違う傾向が現れた。

6.3 考察

目的ベースの探索タスクと、特定画像の探索タスクにおいて違いが表れた理由は、目的ベースの探索タスクを用いた時はタスクに当てはまるものを思い浮かべてもらうという探索対象が曖昧であり試行錯誤が必要となるが、特定画像の探索タスクの実際に画像を提示するという探索対象が明確であり、情報の確実性が違うために探索行動が変化すると推測される。なお、今回は調査を行っていないが、目的ベースの探索タスクでは、当初予定していたものよりもっと良いものを発見できる可能性も考えられる。そこで、今後はこうしたセレンディピティについても注目して実験を行っていく予定である。

目的ベースの探索タスクでは、ユーザ G、H は、図 10 の思った画像と見つけた画像の差異についてのアンケートで既存手法、提案手法共に小さいと回答している。さらに図 9 の探索の難易度では、従来手法よりも高い値を回答している。また、ユーザ C は、既存手法では図 9 の探索の難易度、図 10 の思った画像と見つけた画像の差異のアンケート結果の平均が共に 1.5 を超えている。この 3 人のユーザに共通する点として、複数の手法を併用していない点が挙げられる。このことから、探索の際に手法の併用を行うと探索の難易度は高くなってしまふことがわかる。

また、特定画像探索タスクでは、ユーザ C、D は図 12 において従来手法ではアンケートの評価が-1を下回っているが、提案手法では 0 を超えている。この 2 人のユーザは 2 つ以

上の手法を併用して探索を行う回数が多いことが表 5 から読み取れる。この結果から、ある特定の画像を探索する場合においては、従来手法に比べ複数種類の手法を併用することでより簡単に探索ができるようになることがわかる。

探索してもらった画像ごとの難易度を算出し、その時に用いられた分類手法、併用数の平均を表 6 に示す。

表 6 : 探索難易度別の探索手法の使用回数と併用の回数

	類似	文字	顔	合計数	併用数
容易	1.45	2.64	2.41	5.32	0.12
困難	3.56	5.67	5.44	11.89	0.67

表 6 より、画像探索の難易度が高かったタスクは容易だったタスクよりも 2 倍以上探索行動を行なっていることがわかる。また、併用数は探索が困難な画像は容易だったタスクに比べて 5 倍以上探索行動を行なっていることがわかる。このことから探索の難易度は探索の回数が多いほど困難になり、さらに複数の手法を併用することで難易度が高くなることが明らかになった。こうした手法の組み合わせにより難易度が高くなる問題は、そもそも組み合わせが悪かったことが原因として考えられる。この問題については、分類をどのように行うのかについて検討を行い、修正していく予定である。

なお、実験時にユーザから「選択した画像からさらに類似度を用いた探索がしたい」や「カテゴリや縦横の画像サイズでの分類などもう少し大きなくくりでの分類が欲しい」などのフィードバックも得られた。今後はこれらの意見を参考に分類手法を実装していく予定である。

7. まとめ

本研究では、スマートフォンの写真フォルダにスクリーンキャプチャ画像が蓄積されることで、目的の画像の探索が困難になる問題を解決するために、領域分割を用いた分類、顔の数をを用いた分類、文字数を用いた分類の 3 種類の分類手法を検討した。また、3 種類の分類手法を併用可能なプロトタイプシステムを実装し、探索実験を行うことにより、目的ベースの探索タスクを与えた際は従来の iPhone の写真フォルダのようにただ時系列に並べた場合と比較して、探索時間が短くなるユーザとそうでないユーザに分かれたが、特定画像探索タスクでは探索時間が長くなる傾向になることが分かった。また、イメージからの探索タスクと特定画像探索タスクのどちらにおいてもユーザにとって提案手法を用いることで探索の難易度は緩和することがアンケート調査から明らかになった。

今後の展開としては、今回利用した手法は分類の際に誤認識がいくつか見られたため、そのような誤認識を減らすためにデータセットの作成、新たな手法の検討を行う予定である。また、6 章でユーザの要望としてあげられた機能

についても実装する予定である。

本研究は、スマートフォンのアプリケーションとしての実装を目指している。スマートフォンのアプリケーションの操作はパソコンの操作とは異なるため、スマートフォン上での動作が、ユーザの負担にならないような写真提示手法の検討や、類似度の計算をサーバで行うといったスマートフォン機器への負担を軽減する方法を模索する必要があると考えられる。

謝辞 本研究の一部は JST ACCEL(グラント番号 JPMJAC1602), 明治大学重点研究 A の支援を受けたものである。

参考文献

- [1] 牛尼剛聡, 渡邊豊英. ライフログ内の記録画像を対象とした撮影コンテキストに基づく検索, 情報処理学会研究報告 pp.285-292(2004).
- [2] 森稔, 倉掛正治, 杉村利明, 塩昭夫, 鈴木章. 背景・文字の形状特徴と動的修正識別関数を用いた映像中テロップ文字認識, 電気電子通信学会論文誌 D Vol.J-83-D2 No.7 pp.1658-1666(2000).
- [3] 岸野亜理沙, 白石陽. 事象関連脳電位を用いた写真ライフログのための写真選別手法, 情報処理学会全国大会公演論文集 pp.451-452(2014).
- [4] A. Gomi, R. Miyazaki, T. Itoh, J. Li. CAT:A Hierarchical Image Browser Using a Rectangle Packing Technique, 12th International Conference on Information Visualization(IV08) pp.82-87(2008).
- [5] T. Itoh, Y. Yamaguchi, Y. Ikehata, Y. Kajinaga. Hierarchical Data Visualization Using a Fast Rectangle-Packing Algorithm, IEEE Transactions on Visualization and Computer Graphics 10(3) pp.302-313(2004).
- [6] 森拓矢, 松本哲也, 大西昇. 画像内のオブジェクトに注目した画像内容検索, 情報処理学会研究報告 pp.167-174(2005).
- [7] Gyazo <https://gyazo.com/ja>.
- [8] 上坂俊輔, 安川和希, 重松利季, 加藤俊一. 人物を含む画像から受ける印象の分析・モデル化と自動分類への応用, 社会法人映像情報メディア学会技術報告 pp.31-34(2010).
- [9] 永橋知行, 伊原有仁, 藤吉弘亘. 前景と背景情報の共起表現を用いた Bag-of-features による画像分類, 画像の認識・理解シンポジウム(2010).
- [10] 中島基輝, 韓先花, 陳延偉. Generalized Super-Vector を用いた一般画像分類, 電子情報通信学会研究報告 pp.113-117(2014).
- [11] 上藤太一, 甫足創, 柳生啓司. Multiple Kernel Learning による 50 種類の食事画像の認識, 電気情報通信学会論文誌 pp.1397-1406(2010).
- [12] G. Csurka, C. Bray, C. Dance, L. Fan. Visual categorization with bags of keypoints, ECCV International Workshop on Statistical Learning in Computer Vision pp.1-22(2004).
- [13] 前島紘希, 中村聡史. スマートフォン上のスクリーンキャプチャ画像探索手法の提案. 研究報告ヒューマンコンピュータインタラクション(2017).
- [14] 趙夢, 大島裕明, 田中克己. 類似度と隣接度に基づく画像検索. 第 4 回データ工学と情報マネジメントに関するフォーラム(2012).
- [15] lbpcascade_animeface.xml(https://github.com/nagadomi/lbpcascade_animeface)