

## 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

前島紘希<sup>†1</sup> 中村聰史<sup>†1</sup>

**概要：**近年、スマートフォンやタブレット端末、パソコンといった機器でスクリーンキャプチャを行う機会が増えており、その用途もメモのため、思い出に残すためといったように多岐にわたる。ここで、スマートフォンにおけるスクリーンキャプチャ画像の数も人によっては膨大になりつつあり、その膨大な量のスクリーンキャプチャ画像の中から見たいと思ったものを探すことが困難であるがために、結果的に死蔵されてしまうという問題がある。そこで本稿では、まずスマートフォンのスクリーンキャプチャ画像の特徴の調査を行う。次に、スクリーンキャプチャ画像の特性に基づいて画像特微量や撮影時間を用いたクラスタリングを行い、クラスタごとの代表画像を提示することで、階層的にスクリーンキャプチャ画像を表示し、画像探索をする手法を提案する。また、その有用性を実験により検証する。

**キーワード：**スクリーンキャプチャ画像、自動分類、画像特微量

### 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いる。これらのデバイスは様々な情報をディスプレイ上に映し出すことで、ユーザに多くの情報を提供している。また、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、それらのデバイスには、デバイス上に表示された画面やウインドウ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ときに利用されることが多い。そのため、その内容それぞれについて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場合が多いことや、スマートフォンのスクリーンキャプチャ機能の都合で、画面上の情報を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、この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、スマートフォンでスクリーンキャプチャは、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プライベートで何か嬉しいことがあった時の記念のためや、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すためのメモとして利用する時といった、忘れてはいけない情報を保存するリマインダーとして使われることが多い。また、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、スマートフォンで利用するアプリケーションと、パソコンで利用するアプリケーションの違いから、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るといったように後悔することも多い。消去には手間がかかる。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、そもそものアルバム化が面倒であり、整理などは行われない。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、その量だけは膨大になったが、探索が難しいために結果的に死蔵してしまうことに似ている。

そこで本稿では、こうした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する。実際には、ユーザが画像の整理や削除を必要とせず、スクリーンキャプチャ画像の特性を用いた自動分類を行い、ユーザの探索を支援する手法を提案する。なお、自動分類においては、特徴点分析によるクラスタリングによって画像を用途ごとに分類する手法と、スマートフォンでは連続的にスクリーンキャプチャされるという特性を活かし、一定時間内で連続して撮影された画像集合を1つのグループとしてまとめる手法を提案し、その有用性を検証する。また、その分類手法を組み込んだス

†1 明治大学  
Meiji University

クリーンキャプチャ画像探索システムの実装を行い、利用実験により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。こうした手掛かり情報を牛尾ら[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レグ検索システムを提案している。また、上村ら[2]は、ライフレグとして記録した写真データからコンテキストを抽出し、抽出したコンテキストを用いて多様な検索を実現する連想検索を提案している。スマートフォン上でのスクリーンキャプチャは、現実と連動しておらず、アプリケーションなどで閉じていることが多い。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、稀に起こる事象や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。スマートフォン上でのスクリーンキャプチャにおいても同様の方法をとることで、重要な画像を探索できる可能性がある。

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、各クラスターを表す長方形領域に写真群を配置する写真プラウザを実現している。スクリーンキャプチャ画像でもこ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。

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によって抽出する特徴量を変更することで、抽出の際に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。

スクリーンキャプチャに関するサービスとしてGyazo[7]が挙げられる。このサービスはパソコン上で撮ったスクリーンキャプチャ画像をWeb上に保存し、Webページ内でユーザ自身が撮ったスクリーンキャプチャ画像を検索する機能をもつ。一方本研究は、Gyazoではまだ考慮されていないスマートフォンのスクリーンキャプチャ画像の効果的な探索を目指すものである。前章でも述べたように、スマートフォン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撮影の範囲や対象が異なっているため、それに適した画像探索が必要になると考えられる。

一方、画像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、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。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。そのため、こうした手法を適用する場合は、アニメ絵を認識する手法が必要となる。

画像からの一般物体認識については、Csurkaら[13]のBag-of-Keypointsを用いた物体認識が有名である。また、上東ら[9]は、Web上の画像に対してBag-of-Keypoints手法を用いて、画像のクラス分類を行なっており、クラス分類にはテストデータに最も近い学習データを探し、その学習データのクラスに分類する最近傍分類と、複数のSVMを組み合わせることでマルチクラスに対応させたSVM<sup>multiclass</sup>を利用している。また、上藤ら[12]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[10]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[11]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。

本研究では、全てのスクリーンキャプチャ画像に対してSIFT特徴量を使用し、Bag-of-Keypoints表現を行なっているものである。また、それらを用いた特徴点分析を行い、その後k-means法によるクラスタリングを行い、特徴点が類似しているクラスターを生成している。本研究では、SIFTのみを画像特徴量として利用しているが、本研究はスマートフォンのスクリーンキャプチャ画像に関する基礎研究となっており、今回SIFTのみを用いた分類での結果を算出し、今後の研究で他の画像特徴量を利用し、さらに良い

分類手法を考案していく予定である。

### 3. 予備調査

#### 3.1 スクリーンキャプチャ画像の撮影枚数

スクリーンキャプチャ画像を大量に撮影するユーザがどの程度存在するのか、またユーザごとのキャプチャ画像の枚数にどの程度の差があるかを明らかにするため、スクリーンキャプチャ画像の撮影枚数に関する調査を行なった。調査では、33人の大学生から回答を得た。33人のユーザの撮影枚数ごとの分布を図1に示す。表をみると撮影枚数が100枚以下のユーザから1000枚以上のユーザまで幅広く存在していることがわかる。

なお、スクリーンキャプチャ画像を最も多く所有しているユーザは著者自身で、枚数は4324枚であった。著者を除くと最も多くキャプチャ画像を所有しているユーザは3103枚所有していた。また、表の4つのグループのそれれにおいても撮影枚数にばらつきがあり、利用頻度も多様であることがわかった。しかし、スクリーンキャプチャ画像の枚数が100枚以下のユーザの中には定期的に画像の削除を行うユーザが数名おり、スクリーンキャプチャの利用頻度という点では図1で示されている結果よりも多くなると考えられる。

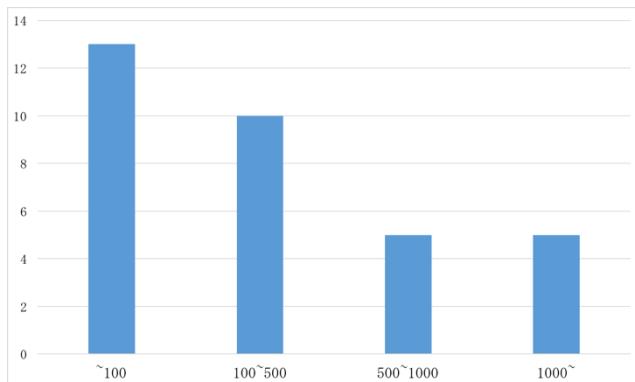


図1 スクリーンキャプチャ画像の枚数ごとのユーザの分布

#### 3.2 スクリーンキャプチャ画像の特徴の分析

スマートフォンにおけるスクリーンキャプチャ画像について分析するため、スクリーンキャプチャ画像の撮影枚数が500枚を超えていた著者を含む学生5人に協力してもらい、合計7438枚の画像について確認するとともに分析を行った。なお、事前にプライベートな画像は省いてもらった。

スマートフォン上でのスクリーンキャプチャ画像の特徴としては、下記の3点が挙げられる。

- 似た構図の画像が多い：**スマートフォンでスクリーンキャプチャを撮影する際、そのユーザが頻繁に利用

するアプリケーションを利用するが多く、起動しているアプリケーションやタイミングがユーザによってある程度固定される傾向にある。また、スマートフォンのアプリケーションでは画面の上部または下部に常にボタンが配置されていることが多い。図2においては画面の下部にボタン群が配置されている。そのため同じアプリケーションでは、シーンが異なっていてもそもそもその画像的な特徴が似ていると考えられる。

- 短い時間内で連続して撮影される：**スクリーンキャプチャ画像の所持枚数が特に多いユーザは、ゲームアプリなどの会話のシーンなどを連続で撮影する傾向が見られた。今回の調査協力者がスマートフォン上で主に利用しているアプリは基本的に静止画の組み合わせで成り立っているものであるため、動画として残ることは求められていないことが多い。そのため、連続的な画像として残すことが多いと考えられる。また、ここでは何らかのアクションの前後にそれぞれ撮影するような行為も多数みられた。
- 実世界で利用するために撮影される：**地図や電車の乗り換え、QRコードといったように実世界に関する情報や実世界で利用するものをスクリーンキャプチャ画像として保存しているものがいくつか見られた。これらの画像は、再度インターネットやアプリを使って検索をすれば手に入れることができる情報であることが多いが、より簡単に情報を見返すためにスクリーンキャプチャ画像として保存していると考えられる。また、映画館で予約した座席などを記録して利用しているケースなどもあった。



図2 似た構図のスクリーンキャプチャ画像の例

(© GREE, Inc ©1999-2016 DeNA Co.,Ltd)

### 4. 提案手法

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

撮影した他の画像といった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。

そこで我々は、前章において整理したスクリーンキャプチャ画像の特徴を活かした画像探索手法を実現する。そうすることで、ユーザがスクリーンキャプチャ画像の探索を行う際の支援ができるのではないかと考えた。本稿では、前章にて調査から明らかにした特徴のうち、「似た構図のスクリーンキャプチャ画像が多い」、「短い時間内で連続して撮影される」という特徴を利用した画像探索のための分類手法を提案する。

- **特徴点分析を用いた分類手法:** スマートフォンの画像は似た構図のものが多いと言う特性をもとに、特徴点分析を行い、クラスタリングすることで類似画像を分類する。この手法では、類似している画像を同じクラスタとして提示することで、求めている画像に類似しているものをまとめる。これにより、ユーザは少ない候補から探索を行うことができるようになると期待される。
- **撮影時間を用いた分類手法:** スマートフォンでのスクリーンキャプチャは短時間に連続撮影されるという特性を利用して、直前に撮影された画像との撮影時間の差から連続撮影された画像群を判断し、それらを同じグループに分類する。この手法では、探索画面で提示する画像の数を減らすことができるため、画面のスクロールを減らすことができる。また、ユーザは同じグループに分類された画像を順に見ていくことで連続撮影を行なった際の背景情報を思い出しやすくなると期待される。

## 5. 分類に関する調査

下記のスクリーンキャプチャ画像の特性を、先述の手法により分析可能かどうかを判定するため、下記の2点の仮説を立てた。

- 似ている構図のスクリーンキャプチャ画像は特徴点分析によるクラスタリングができる
- スクリーンキャプチャ画像を時間的連続性によりグループ化すると、連続撮影時の状況を一括で見ることができ、探索対象の枚数を減らすことができる

本章では、これを立証するために分類に関する調査を行なう。分類に関する調査では、4章で提案した手法について、実際に画像に対して適用し、その精度についての分析を行う。なお、提案手法の実装にはpythonを用いた。以下では、それぞれの手法の実装の詳細について説明する。

特徴点分析を用いた分類にはOpenCVのモジュールであるSIFTを用いた。SIFTは、画像の照明変化や回転、拡大

縮小に不变な128次元の特微量を複数検出することが可能である。SIFT特微量を用いてBag-of-Keypointsで表現をし、それらを用いて画像のクラスタリングを行うことで、撮影時に使用していたアプリケーションごとの分類や、構図の類似している画像ごとの分類が可能になると考えられる。また、画像のクラスタリングについてはk-means法を用いた。

撮影時間を用いた分類では、画像のExif情報や画像タイトルに含まれる撮影時間の情報を利用し、直前に撮られた画像との時間の差が規定の値より小さい場合にそれらの画像を同じグループとしてまとめている。これにより、連続して撮影した画像を一つにまとめ、ユーザの探索のコストを軽減することができる。

また、それぞれの手法にて分類された画像をクラスタやグループごとにテキストファイルで出力を行うことで、後述するプロトタイプシステムに適用する。

実験においては、スクリーンキャプチャ画像の枚数が500枚を超えており、著者を含む大学生8人のスクリーンキャプチャ画像を利用した。なお、実験実施において、プライベートな画像は事前に除いてもらった。

### 5.1 特徴点分析を用いた分類についての調査

SIFTによる特徴点分析を用いて20のクラスタに分類した。実行の結果の一部が図3、図4である。

クラスタ全体で見ると、図3のように同じゲームアプリのスクリーンキャプチャは同一のクラスタ分類できている場合が多かった。ここで、同じゲームアプリ内であっても、異なる場面の画像は、場面ごとにクラスタが分かれる傾向にあった。また、テキスト情報が多い画像は、同じクラスタに属する傾向にあったが、「文字が多い」という画像特微量でクラスタリングされてしまい文字の内容が反映された分類ではないため探索のしやすさが損なわれてしまっている可能性があった。

一方、適切な分類がされていなかったクラスタについては、クラスタの内部に様々な画像が含まれており、クラスタ内でさらに細かいグループに分類できるものが多かった。図3の例では、同じような構図の画像のみが1つのクラスタ画像として出力されているが、図4の例では、アプリの種類、キャラクターの有無、表示されているテキスト情報の量、場所などに統一性が見られないにも関わらず、1つのクラスタとして出力されてしまっている。また、飲食店のウェブページのように、ページ内に写真や地図が入っているものについては、写真や地図の部分に特徴点が集中してしまい、様々なクラスタに分散して提示されてしまっていた。さらに、各クラスタ間の写真の量も、1~2枚のクラスタから100枚以上の写真からなるクラスタといったような偏りが生じてしまっていた。



図 3 適切に分類できた例 (©Donuts)



図 4 適切に分類できなかった例 (© GungHo Online Entertainment, Inc. ©TYPE-MOON / FGO PROJECT)

これらの結果から、ゲームアプリの画像は適切に分類される傾向にあったことがわかる。つまり、「ゲームアプリ」というカテゴリにおいては、「ユーザごとに似ている構図のスクリーンキャプチャ画像は特徴点分析によるクラスタリングが可能である」という仮説が正しいことが示されたと考えられる。しかし、テキストが多いTwitterのタイムラインや電車の乗り換え案内といった画像については、似ている構図のものが多かったにもかかわらずいくつかのクラスタに分かれてしまっていた。そのため、画像特微量を用いる場合には、最初にテキスト情報が多い画像とそれ以外の画像で大まかな分類を行い、テキスト量が多いものと少ないもののそれぞれについて改めて画像特微量を用いた分類をすることで精度がさらに向上すると考えられる。

## 5.2 撮影時間用いた分類の調査

直前の画像との撮影間隔が60秒以内のものを同一のグループとしてまとめたときの、グループ数などの結果を表2に示す。著者の場合、撮影枚数が多いにもかかわらず、グループの数は25%まで減少している。これは、かなりの頻度で連続撮影を行なっており、撮影時間を用いた分類が有効に働くことを示している。一方、著者を除いた7人のユーザについて、提示される画像の数は平均すると減少幅はあまり大きくなかった。

以上のように、撮影時間を用いた分類では、ユーザに大きく依存することが分かった。撮影枚数が多く、また連続撮影枚数が多いユーザにとってはこの時間による分類が効果的に働くと考えられる。

一方、今回の分類では、撮影間隔が60秒以内の画像を同じグループとしていたが、それによって画像の内容的な連續性が見られるにも関わらず別のグループに属してしまっている画像がいくつか見られた。今後はそのような連續性をできるだけ保ち、かつ連續性がない画像はほぼ確実に違うグループに分類できるような撮影間隔の閾値を考える必要がある。

表2 撮影数、グループ数、提示数の割合

	撮影数	グループ数	割合
著者	4324	1090	0.25
ユーザ A	1776	1440	0.81
ユーザ B	1188	990	0.83
ユーザ C	1131	861	0.76
ユーザ D	1112	958	0.86
ユーザ E	631	585	0.92
ユーザ F	571	441	0.77
ユーザ G	522	386	0.74

## 6. プロトタイプシステム

提案手法に基づく画像探索のプロトタイプシステムの実装を行なった。開発にはProcessingを利用した。プロトタイプシステムでは、比較手法として全ての画像を提示する従来手法と、4章で提案した2種類の手法の3種類の画像探索手法を実装している。

全ての手法に共通する特徴として、探索画面では、図5のように最大16枚のサムネイル画像が提示されている。表示するサムネイル画像は、前章で生成したテキストファイルを読み込むことでプロトタイプシステムの動作の負担を軽減している。カーソルキーの上下を押すことで画面がスクロールし、表示されていない範囲の画像を探索することができる。また、提案手法を用いた際には、提示する画像の数を減らすために階層構造を用いてそれぞれの階層の代表画像を提示している。代表画像の選出については、分類手法により切り替えるものとする。それぞれの階層は、分類の際に生成されたクラスタ、グループを1つの階層として用いている。クラスタ、グループの代表画像をクリックすると、その画像が含まれている階層に入り、再び最大16枚のサムネイル画像が提示される。これによってクラスタ、グループ内の探索を行うことができる。また、カーソルキーの左を押すことで階層を1つ戻り、クラスタ、グループの選択を再度行うことができる。最下層に到達した場合、図6のように大きく画像が提示される。その際に画面上部に出るボタンをクリックすることで、探索していた画像、選んだ際の検索方法、選ぶまでにかかった時間といった情報を記録する。

全ての画像を提示する手法では、撮影された日時が古い順に画像を並べている。これは、従来のスマートフォンの写真フォルダにおける画像の提示方法にできるだけ近づけ、他手法との比較をしやすくするためである。

画像特微量を用いた分類手法では、図7のように1つのクラスタにつき4枚の代表画像を横に並べており、黒の枠で囲まれた部分をクリックすることで代表画像がある階層に入ることができる。なお、代表画像の選出は各クラスタ



(© 2013 プロジェクトラブライブ！ ©2016 プロジェクトラブライブ！サンシャイン!! ©KLabGames ©bushiroad  
©BANDAI NAMCO Entertainment Inc.)

の一番新しい画像と、クラスタ内の画像を時間順に並べ、スクリーンキャプチャ画像群を3分割した際のそれぞれの先頭のスクリーンキャプチャ画像の4枚としている。例えば、クラスタ内の画像数が12枚の時、一番新しい画像である12枚目と1枚目、5枚目、9枚目の写真が代表画像として選出される。また、代表画像が4枚未満の場合は、同じ画像を複数回提示することで全てのクラスタから代表画像が選出されるようにしている。これにより、画像特徴量の類似度が高い画像同士が同じクラスタに属しているため、代表画像からクラスタの中を連想しやすく、画像の構図や画像内のキャラクターの状態など、探そうとしているスクリーンキャプチャ画像の背景情報がある程度わかっている場合にのみ検索が容易になると考えられる。

撮影時間を用いた分類手法では、直前に撮影したスクリーンキャプチャ画像との撮影間隔が60秒以内の画像をすべて同じ階層に配置し、その中で最初の1枚を代表画像として提示している。代表画像をクリックすることで、その代表画像が含まれている階層に入り、画像の探索を行うことができる。この手法では、最初の階層で代表画像以外の画像を見ることはできない。しかし、撮影間隔が60秒以内の画像ということは相互に何かしら関連した画像であり、ユーザが探索しようとしているスクリーンキャプチャ画像は、直前もしくは直後に撮影した関連画像の情報とある程度関連があるため、直前の画像から探索しようとしているスクリーンキャプチャ画像の特徴を想起できると考えられ、検索に支障は出ないと予想される。

## 7. 探索実験

我々は、「特徴点分析による分類を用いることで類似している画像から目的のスクリーンキャプチャ画像に容易にたどり着くことができる」、「時系列に並んでいるスクリーンキャプチャ画像は、撮影時間による分類を用いることで

新しく撮影されたスクリーンキャプチャ画像へ容易にアクセスできる」といった2つの仮説を立て、これを検証するために探索実験を行なった。探索実験では、プロトタイプシステムを利用し、検索タスクを提示してそのタスクに当たるスクリーンキャプチャ画像の探索をしてもらった。タスクは以下に示す5つを提示した。

- 一番面白い画像
- 自分が一番気に入っている画像
- メモとして文字をキャプチャした画像
- SNSで共有するために撮った画像
- 一番嬉しかった画像

実験の評価は、各タスクにおいて、タスクにあったスクリーンキャプチャ画像を発見するまでにかかった時間の計測をシステムで行い、探索がうまく出来たか否か、探索の満足度に関する-2~+2の5段階評価と自由記述欄を設けたアンケートを用いて行った。また、全ての探索終了後にシステムを利用して感じたことなどを記入するように指示をした。実験協力者は、5章においてスクリーンキャプチャの分類を行った際の撮影者のうち、著者を除く大学生7名である。

実験の結果だが、まずスクリーンキャプチャ画像の探索にかかった時間の平均と従来手法との時間の差を表3にまとめる。従来の手法と提案した2手法の間に大きな差は見られない結果となった。ユーザごとのスクリーンキャプチャ画像の探索にかかった時間の平均を表4にまとめる。それぞれのユーザで探索時間の平均が一番短かった手法をまとめる、ユーザB, Cは従来手法、ユーザEは撮影時間を用いた分類手法、ユーザA, D, F, Gは画像特徴量を用いた分類手法がそれぞれ一番探索時間の短い分類手法となつた。表4より、今回の実験では半数を超える4人について画像特徴量を用いた探索が一番所要時間の少ない探索手

法であることがわかった。

表3 スクリーンキャプチャ画像の探索にかかった時間

	かかった時間（平均）	従来手法との差
従来手法	1:10	
撮影時間	1:04	-0:06
画像特微量	1:07	-0:03

表4 ユーザごとのスクリーンキャプチャ画像探索にかかった時間の平均

	従来手法	撮影時間	画像特微量
ユーザ A	1:24	0:56	0:44
ユーザ B	0:40	1:28	1:47
ユーザ C	0:20	0:53	0:59
ユーザ D	1:18	0:53	0:28
ユーザ E	2:21	1:06	2:40
ユーザ F	1:13	1:42	0:45
ユーザ G	0:54	0:30	0:27

次に、探索のうまくできたか否か、探索の満足度に関するアンケートの結果を図8にまとめる。いずれも撮影時間は従来手法よりも低く、特微量分析は従来手法よりも高い結果となった。この結果より、特徴点分析による分類手法は実験協力者が従来手法よりも簡単に探索ができ、さらに探索した画像も従来のものよりもより目的にあった画像が探せていることがわかった。逆に、撮影時間を用いた分類手法は従来手法よりも探索がしづらくなっていることがわかった。

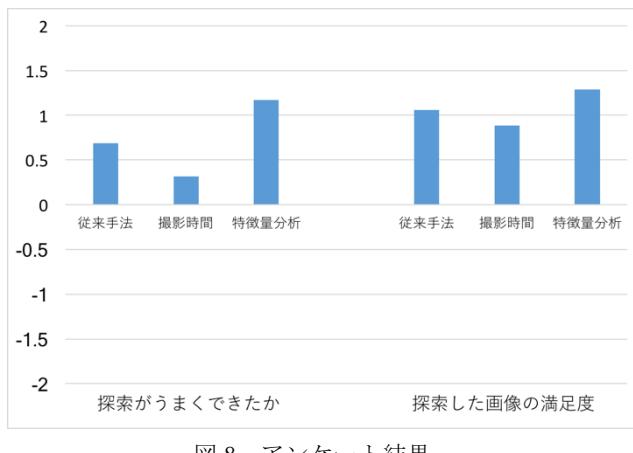


図8 アンケート結果

次に、自由記述のアンケートで得られた意見についてまとめる。撮影時間を用いた分類についての意見には、「新しいスクリーンキャプチャ画像を探すときにスクロール量が少なくて良い」、「従来手法との違いがあまり感じられない」、「1枚の画像で構成されたグループを選んだ時でも2回ク

リックをしなければならなくて面倒」といったものがあつた。「新しい画像を探すときにスクロール量が少なくなる」という意見は「時系列に並んでいるスクリーンキャプチャ画像は、撮影時間による分類を用いることで新しく撮影されたスクリーンキャプチャ画像へ容易にアクセスできる」という仮説が正しいものであることを示していると思われる。「従来手法との違いがあまり感じられない」という意見は、グループに含まれる画像数が1枚のものが多くなってしまったために生まれた意見だと思われる。この問題を解決するためには、撮影間隔の閾値を再度考察し直すが必要だと考えられる。また、今回はすべてのユーザで同じ閾値を用いたが、ユーザのスクリーンキャプチャ画像のExif情報と画像類似度を分析するといった方法で、連続撮影や関連したスクリーンキャプチャ画像かどうかを判断し、ユーザごとの適切な閾値を自動で算出できるようにすることで解決できるのではないかと考えられる。「1枚の画像で構成されたグループを選んだ時も2回クリックをしなければならないのが面倒」という意見は、今回のプロトタイプシステムのUIについての意見のため、1枚の画像で構成されたグループを選んだ際は直接図6のような画面になるようすれば良いと考えられる。

画像特微量を用いた分類についての意見は、「似たような画像がまとめられていて探索がしやすかった」というものが多く挙げられており、「特徴点分析による分類を用いることで類似している画像から目的のスクリーンキャプチャ画像に容易にたどり着くことができる」という仮説が正しいことを示していると考えられる。しかし、「似たような画像が複数のクラスタに分散しており、いくつもクラスタを見る必要があった」という意見も多く見られた。今回の特徴点分析を用いた分類は、すべてのスクリーンキャプチャ画像の特微量を求めた後に類似しているものを同クラスタに分類するといったように教師データを用いない方法で分類を行なっている。スクリーンキャプチャ画像の一部を教師データとして学習を行うことで似た画像がすべて同じクラスタにまとまる可能性があるため、検討していきたい。

全ての探索終了後に行ったシステム全体についてのアンケートでは、「画像特微量の分類は探しやすかった」、「撮影時間を用いた分類はうまく検索できなかった」といったように5段階評価のアンケートと同じような結果が得られた。プロトタイプシステムの改善案として、「それぞれの階層内の画像の枚数を表示して欲しい」、「今どの辺りを見ているかのスクロールバーが欲しい」、「使用しているアプリケーションでの分類が欲しい」といったことが挙げられたため、今後機能として実装していきたい。

## 8. まとめ

本稿では、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

困難になることを問題視した。この問題の解決のために、予備調査としてスマートフォンのスクリーンキャプチャ画像特有の撮影方法や画像の特徴、画像の枚数の調査を行った。その結果、スクリーンキャプチャの撮影枚数にはユーザによるばらつきが多いこと、画像サイズがある程度定まっていること、似た構図のスクリーンキャプチャが多いこと、短時間で連続撮影される場合があることがわかった。

また、画像探索のためにスマートフォンのスクリーンキャプチャ画像の特徴を利用した2種類の分類手法を提案し、それをスクリーンキャプチャ画像に適用した場合の分類結果の傾向分析を行い、ゲームアプリは高い精度で分類することができることがわかった。しかし、テキストが多いスクリーンキャプチャ画像はクラスタリングによって同じクラスタには属するものの、テキストの内容による分類はできていなかった。さらに、プロトタイプシステムを実装し、プロトタイプシステムを用いた探索実験を行った。探索実験により、手法ごとの全ユーザの探索の時間の平均の差には大きな違いはなかったが、ユーザへのアンケート調査によって特徴点分析による分類手法は従来手法よりも探索がうまくできることが明らかになった。

今回、著者が本システムを使用した際に感じた使用感として、撮影時間を用いた分類は、ゲームアプリの会話などを撮影したものがグループ化され、代表画像の提示枚数が大幅に削減されており、目的の画像と代表画像に関連があるかわかっている際に簡単に探索ができた。一方、画像特徴量を用いた分類は、代表画像が4枚提示されるため、目的のスクリーンキャプチャ画像に近いものは比較的見つけやすいという特徴があった。クラスタ内には似たような画像は多かったが、提示される画像の枚数が減っているので探索は容易になっているように感じた。

今後の展開としては、SIFT特徴量での分類の精度を上げる方法の検討、SIFT以外の特徴点分析の際に使うことができる特徴量の検討、連続撮影に基づいて分類を行う際の撮影時間の差についての考察といったことが挙げられる。また、スクリーンキャプチャ画像のテキストに着目して検索を行うことも期待される。そのため、文字認識の適応の可能性を探っていく必要がある。加えて、今回の実験において、プロトタイプシステムでのスクリーンキャプチャ画像の提示方法によって評価が下がった可能性も十分に考えられる。よって、ユーザにとって探索の際にストレスが溜まらないような画像の提示手法についても検討していきたい。

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。本稿で提案したプロトタイプシステムでは、検索の際に階層構造を用いていたが、それにより探索が難化してしまった可能性がある。よって、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が必要であると考えられる。

**謝辞** 本研究の一部は JST ACCEL の支援を受けたものである。

## 参考文献

- [1] 牛尾剛聰, 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クストに基づく検索, 情報処理学会研究報告 pp.285-292(2004).
- [2] 上村真也, 梶克彦, 廣井慧, 竹中光, 竹内重樹, 岡本学, 川口信夫. ライフログ写真データに基づく連想探索に関する研究, マルチメディア, 分散協調とモバイルシンポジウム 2014 論文集(2014).
- [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, 12<sup>th</sup> 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] 上東太一, 柳井啓司. Bag-of-Keypoints 表現を用いた Web 画像分類, 情報処理学会研究報告 pp.201-208(2007).
- [10] 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).
- [11] 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).
- [12] 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).
- [13] 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).