

## 2 段階 GrabCut を用いた 注目物体の視認性を向上させたサムネイル生成

新井 啓介<sup>†</sup> 武井 宏将<sup>†</sup> 山名 早人<sup>‡§</sup>

<sup>†</sup> 早稲田大学大学院基幹理工学研究科 〒169-8555 東京都新宿区大久保 3-4-1

<sup>‡</sup> 早稲田大学理工学術院 〒169-8555 東京都新宿区大久保 3-4-1

<sup>§</sup> 国立情報学研究所 〒101-8430 東京都千代田区一ツ橋 2-1-2

E-mail: {k-arai, htakei, yamana}@yama.info.waseda.ac.jp

**あらまし** 本稿では、携帯端末のような小さい表示領域で利用できるサムネイル生成手法を提案する。画像を縮小したサムネイルは、大量の画像を取り扱う場合の一覧性確保に効果的である。しかし、単純な縮小によるサムネイルはサイズが小さくなると何が写っているのか把握できず、視認性が失われてしまう。この問題に対し、本稿では GrabCut を 2 段階に適用することによってユーザが注目している物体と背景とを分離し、これらを別々に縮小することによって注目物体の視認性を向上させる手法を提案する。従来の GrabCut は注目物体の大きな指定が必要であったのに対し、本稿で提案する 2 段階 GrabCut ではこれを自動化すると共に注目物体の抽出精度を高めた。75 枚のサムネイルから指定画像を探索する 20 人による画像探索実験の結果、単純な縮小を用いたサムネイルに比べ、提案手法により、探索時間において最大 44% の時間短縮と、探索精度において最大 19% の精度向上を確認した。

**キーワード** サムネイル, Small size displays, Content-aware image resizing, GrabCut

### 1. はじめに

近年、スマートフォンやタブレット端末の普及から、画面の小さな端末上で大量の画像を取り扱うことが増えた。これらの携帯端末は、インターネットに接続できるために Web 上から、もしくは搭載されたカメラから、大量の画像を容易に集めることができる。しかし、携帯端末のように小さな画面上で、大量の画像の中から目的の画像を探すことは困難である。このため、大量の画像を効率よく取り扱えるような、閲覧性の高い画像ビューワが求められている。本稿で定義する閲覧性とは、多くの画像を同時に見ることができの一覧性と、画像に何が写っているのか、どこで写っているかなどの概要を把握することができる視認性からなる。一般に、一覧性は画像が小さいほど高くなり、反対に、視認性は画像が大きいほど高くなる。

コンピュータにおいて大量の画像を取り扱う際、複数の画像を同時に閲覧する効果的な手法として、サムネイルと呼ばれる縮小画像が利用されている。デスクトップコンピュータのように画面の表示領域が大きい場合、単純な縮小によるサムネイルは十分な視認性を保ちながらも一覧性を図ることができる。このため、今日では、画像ビューワや検索エンジンの結果表示などで利用されている。しかし、デジタルカメラや携帯電話といった携帯端末では画面が小さく、単純な画像縮小(以下、Scaling と呼ぶ)によるサムネイルは適して

いない[1]。Scaling によるサムネイルでは、人物などユーザが注目したい物体が小さくなりすぎて、利用者がどのような画像なのかを把握することができないからである。また、限られた画面サイズのなかで多くの画像を表示するためには、元画像と同一のアスペクト比を保てないことがある。この場合、Scaling では注目している物体が変形してしまう問題がある。注目している物体のアスペクト比を変更することは、視認性を損なう要因となる[2]。

Scaling のような一般的な縮小手法に対し、画像内の注目物体が画像に占める割合を大きくするように縮小することで、サムネイルの視認性を上げることができる。このような物体に注目した画像縮小手法として、Content-aware image cropping(以下、Cropping と呼ぶ)と、Content-aware image resizing(以下、CAIR と呼ぶ)がある。

Cropping とは、画像内の注目領域を含むような矩形を切り出す手法である[3]-[7]。注目領域の決定法には、画素ごとに重要度を定めた Importance Map の値によって決定される。Importance Map の決定法として Saliency Map[8]や Visual Attention Map[9]を利用した人の視覚特性に基づいた手法[3][4]がある。また、顔認識を利用した手法[5]や、顔認識、SIFT 特徴量、画素位置によって決定される手法[6]もある。Importance Map を用いずに注目領域を決定する手法としては、画像間の類似

性計算式を定義し、元画像と切り取った画像との間で類似度が最大になる領域を指定する方法[7]がある。**Cropping** は画像の注目物体が一つしかない場合や、物体同士が集中している場合に有効である。しかし、注目領域以外の背景をすべて削除してしまうために全体像が掴めなくなる。さらに、物体が散在していると物体全てを包含するように矩形を設定するため、**Scaling** とほとんど違いがなくなってしまうという問題がある。

CAIR による縮小手法は、注目度の低い領域を削除する手法[10]-[13]、注目度の低い領域を歪める手法[14]-[16]、部分画像から縮小画像を構築する手法がある。このうち、部分画像から縮小画像を構築する手法は、領域分割を元にした手法[1]と、パッチ画像を元にした手法[7]に分類される。CAIR は **Cropping** と違い、注目物体と背景情報の両方を保持することができるため、視認性の高いサムネイルを得ることができる。一方で CAIR は **Importance Map** やそれに類似した指標を元に縮小を行うが、**Importance Map** を適切に指定しなければ、望み通りの結果を得ることができない。任意の物体を保存したい場合、**Cropping** が大雑把に矩形を指定するのに対し、CAIR は保存したい物体の形を指定する必要がある[17]。

この問題に対し、本稿では 2 段階 **GrabCut** を用いた視認性の高いサムネイル生成手法を提案する。本手法では領域分割を元にした CAIR 手法と同様に、注目物体を前景、それ以外を背景として分離し、別々の縮小率で縮小する。注目物体が画像に占める割合を大きくすることで、視認性を向上させることができる。さらに、**GrabCut** を 2 段階に用いる新たな手法により、手動で物体の形を指定しなければならなかった従来の CAIR と違い、矩形を指定するだけの単純な操作から注目物体の形を推定する。

本稿では以下の構成をとる。2 節において関連研究について述べ、続く 3 節では提案手法について述べる。次に 4 節にて提案手法によるサムネイルの生成結果を表示し、最後に 5 節にてまとめる。

## 2. 関連研究

本節では、画像を縮小する関連研究として、CAIR 手法について述べる。CAIR は **Cropping** と違い、注目物体と背景情報の両方を保持できることから、視認性の高いサムネイルの生成に適している。

### 領域分割による手法

領域分割に基づいた手法は、前景と背景を分離し、それぞれ違う縮小率で縮小後、合成することで縮小画像を作成する。注目物体と背景を別々に縮小するため、注目物体を歪ませることなく大きく見せることができ、

サムネイルの視認性を向上させることができる。Setlur ら[1]は自動的に前景と背景を分離し、画像を縮小する手法を考案した。前景の切り出しには、**Saliency Map**、画素座標、顔認識からなる **Importance Map** を利用している。しかし、Setlur ら[1]の手法はテキスト領域や色変化の影響を受けやすく、目的の物体を抽出できない問題がある。この問題を回避するために **Importance Map** を手動で設定することもできるが、その場合、注目したい物体の形を細かく指定しなければならない。

### 領域を削除する手法

領域を削除する手法は、画像内における注目度の低い個所から漸次削除していくことで画像を縮小する手法である。Avidan らが提案した **Seam Carving**[10]は、画像にエネルギーを定義し、水平、垂直方向にエネルギーが最小となるパスを削除することで画像を縮小していく。注目したい物体に適したエネルギーが定義されていれば、**Seam Carving** は視認性の高い縮小画像を作成することができる。しかしエネルギーが適切に定義されていなかった場合、目的のサイズになるまでに注目物体が変形してしまう問題がある。この問題に対し、Dong ら[11]は **Scaling** を、Rubinstein ら[12]は **Scaling** と **Cropping** を組み合わせることで目的のサイズに落とし込む手法を考案した。また、Avidan ら[10]は顔認識と物体指定を、Hwang ら[13]は顔認識や **Saliency Map** によってエネルギーにバイアスをかける方法で解決を図っている。エネルギーにバイアスをかけることは注目物体の保存に有効である。しかし、画像ごとに適したエネルギーは異なるため、手動で保存したい注目物体の形にバイアスをかける必要がある。また、適切なバイアスをかけた場合であっても、注目物体を避けて低エネルギーのパスを設定できなくなれば、注目物体の変形が行われてしまう問題がある。

### 領域を歪ませる手法

領域を歪ませる手法は、重要度の低い領域を小さくすることで、画像を縮小する手法である。Wolf ら[14]は画素を線型方程式に従って移動させることで画像を縮小した。Wolf らの手法[14]では、画像の勾配、顔認識、動き量から画素に重みづけを行い、その重みと 8 近傍画素との関連性によって画素の移動量を決定した。Wang ら[15]と Guo ら[16]は、画像をメッシュ状に分割し、**Importance Map** に従い、メッシュの face 単位で画像を歪ませることで縮小画像を作成した。Wang ら[15]は画像をグリッド上に分割したメッシュと、画像の勾配、**Saliency Map** からなる **Importance Map** によって縮小画像を作成した。Guo ら[16]は三角形からなるメッシュと、**Saliency Map**、顔認識、顔認識から推定した

胴体情報からなる Importance Map によって縮小画像を作成した。これらの手法では、領域を削除する手法と同様に、Importance Map が縮小画像を生成する上で重要な役割を担っている。このため、注目物体を適切に保存するためには、手動で注目物体の形を指定しなければならない。また、適切なバイアスを指定した場合であっても、画像の縮小率が高くなれば注目物体が変形してしまう問題がある。

### 3. 提案手法

本節では、2 段階 GrabCut を用いた視認性の高いサムネイル生成手法を提案する。関連研究で述べたように、領域を削除する手法と領域を歪める手法は、画像の縮小率が高まるにつれて注目物体を歪めてしまうため、小さいサイズのサムネイル生成には向いていない。また、CAIR 手法では、Importance Map やエネルギーに結果が左右されるため、注目物体を保存するためには、手動で注目物体の形を指定しなければならない問題があった。これに対し、提案手法では、GrabCut を 2 段階に利用することで、従来の GrabCut が必要としていた「物体の大まかな指定」を廃し、「注目物体を包含した矩形」から注目物体を抽出できるようにする。提案手法の全体の流れは、領域分割による CAIR 法[1]に基づいており、図 1 に示す三つのステップから成る。

- I. 入力画像に対して 2 段階 GrabCut を適用し、注目物体を表す前景と背景とに分離する。
- II. 前景を Scaling によって、背景を Seam Carving と Scaling によって、それぞれ縮小する。
- III. 縮小した前景と背景を合成する。

まず、指定された注目物体を囲む矩形から、2 段階 GrabCut によって注目物体を示すマスク画像を自動的に生成する。マスク画像は 0 か 1 の値を持つ 2 値画像であり、注目物体位置の画素を 1、それ以外は 0 の値を持つ。次に、マスク画像から前景と背景を分離し、注目物体が切り出された背景に対して画像補完による穴埋めを行う。穴埋めされた背景は Seam Carving と Scaling を用いて事前に設定したサイズになるまで縮小する。最後に、前景画像同士の位置関係を計算し、元の位置関係を保つように大きさを変更し、背景との合成を行う。

提案手法によって作成されたサムネイル画像は、注目物体が歪んでおらず、画像内で大きく表示されるため、Scaling によるサムネイルよりも視認性が高い。加えて、2 段階 GrabCut を利用することで、Setlur らの手法[1]に比べ、背景の顕著性やテクスチャに頑健な前景切り出しが可能となっている。

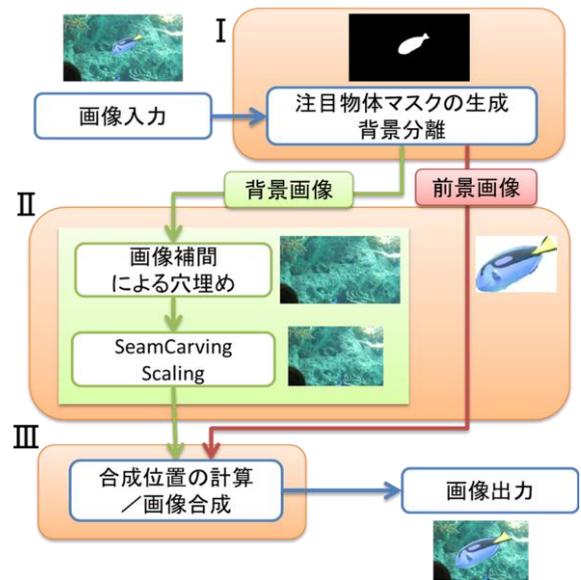


図 1 提案手法概要

#### 3.1. 前景と背景の分離

提案手法では、最初に入力画像における注目物体を前景とし、背景との分離を行う。注目物体とは、画像内においてユーザが重要だと感じる物体であり、例えば人物や特定の物体などを指す。注目物体を包含した矩形は、HOG[18]やテンプレートマッチングなど、注目したい物体の検出器を用いることで自動的に与える。



図 2 GrabCut による領域分割結果

左：元画像，中央：矩形指定のみ，右：提案手法

提案手法では、注目物体を取り囲むように指定された矩形から、注目物体のみの領域を自動で切り出す。物体領域の切り出しには Rother らの GrabCut[19]を利用する。GrabCut は、矩形で囲まれた領域内を、前景と背景に分離する手法である。前景となる物体が含まれている矩形と、画素の色、前景か背景かを示すシードラベルを定義することで、精度の高い前景分離を行うことができる。本稿で利用するラベルは、「前景である(FGD)」「おそらく前景である(PR\_FGD)」「背景である(BGD)」「おそらく背景である(PR\_BGD)」の 4 種類とする。

GrabCut のシードラベルは、ユーザが動的に与えなければならない。シードラベルを与えずに物体を包含する矩形のみを指定した場合、矩形からシードラベル

を決定するため、精度の高い前景分離は難しい（図 2 中央）．この問題に対し、提案手法では GrabCut を 2 段階に用いることで前景分離を行う．まず、Felzenszwalb らの領域分割手法[20]による領域分割の結果から、1 段階目の GrabCut のシードラベルを推定し、GrabCut を行う．次に 1 段階目 GrabCut の領域分割結果で得られた画素ごとの前景らしさを投票空間に投票する．最後に、投票結果から 2 段階目 GrabCut のシードラベルを推定し、GrabCut による前景分離を行う．

以上を踏まえ、提案手法による物体領域の切り出しは次の手順にて行う．以下のフローにおいて、 $i [1 \leq i \leq M]$  は、 $i$  番目の矩形を表しており、全ての矩形において以下に示す 8 つの処理を行う．

1. 背景領域(BGD)情報を得るために、選択された矩形  $R_i$  を上下左右に  $m$  画素ずつ膨らませた矩形  $R'_i$  を作成する．次に、矩形  $R'_i$  を用いて、入力画像から物体画像  $O_i$  を切り出す．拡大された矩形  $R'_i$  内の外側  $m$  画素の領域を、本手法では物体画像  $O_i$  の背景領域(BGD)であると仮定する．矩形  $R_i$  を膨らませる大きさ  $m$  は、本稿では 3.2.1 項で利用するパッチサイズと等しくなるように、 $m=7$  としている．矩形  $R_i$ 、矩形  $R'_i$ 、物体画像  $O_i$  の関係を図 3 として示す．

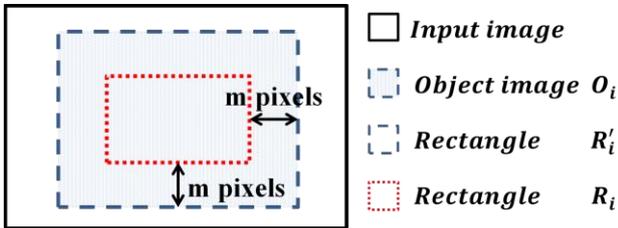


図 3  $R_i, R'_i, O_i$  の関係

2.  $O_i$  を Felzenszwalb らの手法によって領域分割し、領域分割結果を  $S_i = \{s_{ij} | 1 \leq j \leq M, s_{ij} \text{ is } j\text{-th segment in } O_i\}$  とする．Felzenszwalb らの手法[20]による領域分割結果の一例と、領域分割結果からの前景領域の推定結果を図 4 として示す．



図 4 領域分割結果(左)と前景推定結果(右)

前景分離を行うために、式(1)に従って、全ての画素にラベルを割り当てる．割り当てられたラベルを、1 段階目 GrabCut のシードラベル

$SL1_i(x,y)$  とする．図 5 に示すように、 $R_i$  の外側の領域にある画素は BGD のラベルを付与する．次に、領域全てが  $R_i$  内に収まっている領域の画素を PR\_FGD とし、残りの画素を PR\_BGD とする．

$$SL1_i(x,y) = \begin{cases} PR\_FGD & \text{if } S_i(x,y) = PR\_FGD \\ BGD & \text{else if } (x,y) \notin R_i \\ PR\_BGD & \text{otherwise} \end{cases} \quad (1)$$

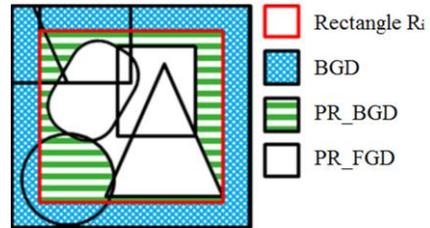


図 5 ラベルの定義

3. 1 段階目の GrabCut を行い、GrabCut の結果ラベル  $RL1_i(x,y)$  から(2)式に従って投票空間  $V_i(x,y)$  に対して投票を行う．投票空間  $V_i$  は  $O_i$  と同じ大きさであり、画素ごとの前景らしさを投票数によってあらわしている．

$$V_i(x,y) = \begin{cases} +1 & \text{if } RL1_i(x,y) = FGD \text{ or } PR\_FGD \\ +0 & \text{otherwise} \end{cases} \quad (2)$$

4. 領域分割のパラメタ  $\sigma, k$  を変更しながら、手順 2 から手順 4 を繰り返し実行する． $\sigma, k$  は論文[20]において定義されているパラメタであり、これらを変更することによって領域分割の粒度を変えることができる．本稿では粗い分割から細かい分割までの多様な結果を得るために、 $\sigma$  を  $[0.1:1.1]$ 、 $k$  を  $[10:300]$  の正規乱数とした．正規乱数の平均は  $\sigma = 0.55$ 、 $k = 150$  である．また、手順 2 から手順 4 の反復回数を 150 とした．

5. 投票空間  $V_i(x,y)$  を閾値  $TH$  によって二値化し  $V'_i(x,y)$  とする． $R_i$  内の画素数を  $AP$  としたとき、 $TH$  は(3)式のように定義する．

$$TH = \frac{1}{2AP} \sum_{x,y \in R_i} V_i(x,y) \quad (3)$$

6. (4)式に基づいて  $V'_i(x,y)$  から 2 段階目の GrabCut で用いるシードラベル  $SL2_i(x,y)$  を決定する．

$$SL2_i(x,y) = \begin{cases} PR\_FGD & \text{if } V'_i(x,y) = 1 \\ PR\_BGD & \text{else if } (x,y) \subset R_i \\ BGD & \text{otherwise} \end{cases} \quad (4)$$

7. 2 段階目 GrabCut の結果ラベルから、同じラベルが割り当てられた画素同士を、8 近傍連結によって統合する．統合された領域ラベルが FGD、もしくは PR\_FGD であった場合、前景を

表す領域の候補とする。

8. 前景領域候補のうち、最大の面積を持つ領域を切り出し、矩形 $R_i$ における注目物体を表した前景領域とする。

図 2 に示したように、提案手法では通常の GrabCut[19]（同じように矩形のみを指定）に比べて、精度の高い前景分離を実現している。最後に、手順 1 から手順 8 によって抽出された注目物体を表す前景領域からマスク画像  $M_s$  を作成し、 $M_s$  を元に入力画像を前景と背景に分離する。

## 3.2. 背景画像の補完と縮小

### 3.2.1. 背景画像の補完

3.1 で生成した背景画像には、注目物体が存在していた場所に穴が空くため、この穴を画像補完により穴埋めする。これは、提案手法では、切り出した前景を縮小された背景画像に合成するとき、別々に縮小するために元の位置に戻るとは限らず、生成したサムネイルに穴が残ってしまう可能性があるからである。補完には PatchMatch [21] を利用する。補完に利用するパッチサイズは、文献[21]に従い、7 とする。補完処理によって得られた補完画像を図 6 として示す。補完により物体が存在した箇所に、違和感の少ない画像が作成されているのが確認できる。



図 6 画像補完による背景の穴埋め結果

### 3.2.2. 背景画像の縮小

補完された背景画像を、目的のサイズまで縮小する。この時、Scaling による縮小のみを行うと、背景画像内にユーザが注目したい物体や領域があった場合に、視認性が失われるという問題がある。そこで提案手法では、顕著性の高い領域は重要度が高いという仮定[13]に従い、顕著性の高い領域を残すように、Seam Carving による縮小を行う。まず、Seam Carving によって背景画像を縮小することで、Scaling とは違い、重要度の高い領域の視認性を向上させる。

提案手法では、Seam Carving に利用するエネルギー  $e_p$  を(5)式によって定義する。背景画像における  $(x,y)$  座標の勾配強度を  $I(x,y)$ 、Saliency Map の強度を  $Sal(x,y)$ 、3.1 で作成したマスクの値を  $M_s(x,y)$  とする。 $M_s$  は 0 と 1 の 2 値で構成されており、0 は背景領域を、1 は前景

領域であることを示している。背景領域におけるペナルティコストの定数  $C$  は 1 とした。

$$\begin{cases} e_p(x,y) = I(x,y) + \frac{Sal(x,y)}{2} + C & \text{if } M_s(x,y) = 0 \\ e_p(x,y) = 0 & \text{else if } M_s(x,y) = 1 \end{cases} \quad (5)$$

(5)式で定義されたエネルギー  $e_p$  を利用することで、3.2.1 において穴埋めを行った領域から優先的に削除されるようになる。穴埋めされた領域とは、注目物体が存在していた領域である。穴埋めのために作成された部分画像は、元の画像には含まれていない画像であることから、ユーザが注目したい領域にはなり得ない。加えて、サムネイルは元画像の概要を把握できる内容である必要があるため、補完によって作成された領域は優先的に削除されるべきである。また、注目物体が存在した領域以外については、顕著性が高い領域は重要度が高いという仮定の下、顕著性の低い領域から削除していく。

Seam Carving は、画像内の物体とそれ以外の領域を区別せずに削除していくため、縮小率が高くなるにつれて、物体の構造を破壊してしまう。そこで、提案手法ではまず Seam Carving によって顕著性の高い領域を残し、画像内の物体が歪んでしまう前に Scaling による縮小手法に切り替えることを行う。これにより、Scaling のみによる縮小よりも重要度の高い領域を残すことができ、Seam Carving のみによる縮小のように物体の構造を破壊してしまうことを防ぐことができる。

## 3.3. 前景・背景画像の合成

前景画像を、縮小された背景画像に合成する。前景画像の位置は、Seam Carving による中心位置の移動量と、縮小による移動量を計算し、縮小する以前の場合に対応する位置とほぼ同じ場所になるように設定する。また、前景画像の大きさは、各前景画像が大きく重ならないように調整する。本稿では、前景画像を背景画像に貼り付けることで合成を行う。合成において、単純な貼り付けの代わりに Poisson Blending[22]などを利用することで、自然な合成結果を望むことが期待できる。

## 4. 実験

本節では、現在広く利用されている Scaling と比較することで、提案手法の有効性を考察する。3.1 で述べたように、提案手法では検出器によって与えられる矩形を元に、注目物体の分離を行う。しかし、本節の実験においては、検出器の精度にサムネイルの生成結果が左右されないように、これらの矩形は手動で与えている。

#### 4.1. 提案手法によるサムネイル生成実験

図 7 に提案手法と Scaling によって作成されたサムネイルを示す。元画像において注目物体が小さかった場合、提案手法では、Scaling よりも注目物体が大きく、視認性が優れていることがわかる。元画像において注目する物体が十分に大きかった場合、提案手法による縮小は優位性が低い。しかし、アスペクト比が変更された場合、アスペクト比の変更が起こっても注目物体が変形しないため、Scaling に比べて視認性が高くなっていること見受けられる。

次に、アスペクト変更に対する評価を図 8 として示す。提案手法では Scaling と違い、アスペクト比が変更されても注目物体が歪まないのが確認できる。

図 7, 図 8 より提案手法では Scaling に比べ、低サイズサムネイルにおいても注目物体の視認性が高いことが確認できた。

#### 4.2. ユーザによる画像探索実験

提案手法のサムネイルによる、画像探索の閲覧性についての評価を行う。サムネイルの閲覧性が高ければ、目的の画像を素早く、正確に探し出すことができる。閲覧性の評価を行うために、本稿では 75 枚の画像の中から目的の画像を探し出す実験を、3 種類行った。それぞれの実験は 10 個の問題で構成されており、ランダムに選出された探索の目的となる画像を、75 枚のサムネイル画像の中から探し出すものである。実験に利用したシステムのスクリーンショットを図 9 として示す。まず、左のペインに問題となる画像が 5 秒間表示され、その間に被験者には画像を覚える。次に、中央のペインに 75 枚のサムネイルが表示されるので、60 秒間でサムネイルから問題の画像を捜し出す。サムネイルをクリックすると、右のペインに元の画像が表示されるので、覚えた画像と同じであれば「確認ボタン」を押し正解、不正解の判定が行われる。探索時間の上

限である 60 秒を超えてしまった場合、不正解として取り扱う。

実験 1 と実験 2 では、ある画像に含まれる注目物体を問題の画像とし、元となる画像をサムネイルから探す。実験 1 では小さいサムネイルとして、画面上で 1.2cm×1.6cm、実験 2 では大きいサムネイルとして、2.6cm×1.6cm の縮小画像を用いた。実験 3 は、問題の画像として Scaling による画像を表示し、サムネイルから同じ画像を探すものである。利用するサムネイルは実験 2 と同様に大きいサムネイルとした。各実験において利用した画像は、131 枚の画像からサブセットとして、ランダム抽出した 75 枚を利用した。また、問題となる画像はサブセット内の画像から対応する画像をランダムに選出した。日常的に携帯電話や PC を利用する 20 代から 30 代の男性と女性、合計 20 名を対象に 3 種類全ての実験を行った。

表 1 は、それぞれの実験における精度、探索時間、クリック回数の平均と標準偏差を表したものである。精度とは各実験における正答率を表し、探索時間は 1 問あたりの解答時間を、クリック回数はサムネイルのクリック回数を示している。実験 1, 実験 2 のように注目物体の記憶を頼りに画像を探す場合、精度、探索時間、クリック数の全てにおいて提案手法が優れていることが確認できる。特に実験 1 においては、提案手法では Scaling に比べ、平均で 44% の時間短縮、19% の精度向上が確認された。しかし実験 3 のようにサムネイルの記憶を頼りにサムネイルから画像を探す場合、Scaling に比べて提案手法の有効性は確認できなかった。これは、注目物体を大きく見せるために、画像全体の構造が保存されていないためであると考えられる。以上の実験により、提案手法は Scaling に比べ、小さいサムネイルの時、特に何が写っていたかの記憶を頼りに探す場合に有効であるといえる。

Table 1. 実験 1-3 における精度、探索時間、クリック数

	method	Precision(%)		Search time(sec.)		Number of Clicks	
		average	SD	average	SD	average	SD
実験 1	scaling	79.0	14.5	22.33	7.28	31.40	18.90
	ours	94.0	8.21	12.44	3.78	15.30	5.53
実験 2	scaling	91.5	11.37	15.47	4.97	14.15	5.31
	ours	94.5	6.86	10.87	3.26	12.35	2.65
実験 3	scaling	99.0	3.08	8.80	3.30	10.50	0.97
	ours	95.0	7.61	10.24	4.20	11.30	2.00

## 5. まとめ

本稿では、注目物体と背景を分離して縮小することで、注目物体に対する視認性の高いサムネイル生成手法について述べた。2段階 GrabCut を用いることで、従来では手動で指定していた注目物体の形状を、既存の検出器から与えられる矩形から、自動で作成することができる。提案手法によって生成されたサムネイルは、小さい画面上で利用できるようなサイズのサムネイルにおいても、Scaling に比べて十分な視認性があることが確認できた。

しかし、提案手法では以下の課題が残されている。

### 複雑な領域分割に対する精度

提案手法では、1段階目の GrabCut によって2段階目の GrabCut の初期ラベルを推定し、領域分割の精度を高めている。しかし、1段階目 GrabCut のラベル推定では、領域分割の結果において、矩形外は背景であるという仮定から背景領域を推定している。このため、色の変化が激しい背景であった場合や、背景領域が前景領域に囲まれていた飛び地になっている場合などに、背景情報がうまく伝播せず、分離がうまくいかないといった問題がある。この問題に対し、本稿で述べているような背景領域を削除していく手法以外に、色類似度などから前景領域を拡大していく手法を検討する必要がある。

### 背景縮小の閾値設定

提案手法では、背景領域の縮小に Seam Carving と Scaling を利用している。しかし、どこまでを Seam Carving で縮小し、どのサイズで Scaling に切り替えるのかの調整は人手で行っている。提案手法の完全な自動化を図るうえで、Dong らの手法[11]や Rubinstein らの手法[12]のように縮小手法の切り替えを適切に行う必要がある。

### 謝辞

本研究の一部は、科研費（基盤（B）21300038）によるものである。

### 参考文献

- [1] V. Setlur, V. S. Takagi, R. Raskar, M. Gleicher, and B. Gooch, "Automatic Image Retargeting," Proc. of the 4th Int'l. Conf. on Mobile and Ubiquitous Multimedia, 2005.
- [2] H. Rushmeier, H. Barrett, P. Rheingans, S. Uselton, and A. Watson, "Perceptual measures for effective visualizations," Proc. of the 8th Conf. on Visualization, pp. 515-517, 1997.
- [3] I. S. Amurutha, S. S. Shylaja, S. Natarjan, and K. N. Balasubramanya Murthy, "A smart automatic thumbnail cropping based on attention driven regions of interest extraction," Proc. of the 2nd Int'l Conf. on Interaction Sciences, pp.957-962, 2009.
- [4] B. Suh, H. Ling, B. B. Bederson, and D.W. Jacobs, "Automatic thumbnail cropping and its effectiveness," Proc. of the 16th Ann. ACM Symp. on User Interface Software and Technology, pp.95-104, 2003.
- [5] X. Fan, X. Xie, W. Y. Ma, H. J. Zhang, and H. Q. Zhou, "Visual attention based image browsing on mobile devices," Proc. of Int'l Conf. on Multimedia and Expo, pp. 53-56, 2003.
- [6] K. Lyndon, V. Z. Roelof, Z. Nicolas, and T. Belle, "Learning crop regions for content-aware generation of thumbnail images," Proc. of Int'l Conf. on Multimedia Retrieval '11, No.30, 2011.
- [7] D. Simakov, Y. Caspi, E. Shechtman, and M. Irani, "Summarizing visual data using bidirectional similarity," Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, pp.1-8, 2008.
- [8] L. Itti, C. Koch, and E. Niebur, "A model of saliency-based visual attention for rapid scene analysis," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 20, No. 11, pp. 1254-1259, 1998.
- [9] O. Marques, L. M. Mayron, G. B. Borba, and H. R. Gamba, "An attention-Driven Model for Grouping Similar Images with Image Retrieval Application," EURASIP J. on Applied Signal, Vol.2007, No.1, 2007.
- [10] S. Avidan, and S. Shamir, "Seam Carving for Content-Aware Image Resizing," ACM Trans. on Graphics, Vol. 26, No. 3, 2007.
- [11] W. Dong, N. Zhou, J. C. Paul, and X. Zhang, "Optimized Image Resizing Using Seam Carving and Scaling," ACM Trans. on Graphics, Vol. 28, No. 5, 2009.
- [12] M. Rubinstein, A. Shamir, S. Avidan, "Multi-operator media retargeting," ACM Trans. on Graphics, Vol. 28, No. 3, 2009.
- [13] D. S. Hwang, and S. Y. Chien, "Content-aware image resizing using perceptual seam carving with human attention model," Proc. of Int'l Conf. on Multimedia and Expo, pp.1029-1302, 2008.
- [14] L. Wolf, M. Guttman, and D. Cohen-Or, "Non-homogeneous content-driven video-retargeting," Proc. of IEEE Int'l Conf. on Computer Vision, 2007.
- [15] Y. S. Wang, C. L. Tai, O. Sorkine and T. Y. Lee, "Optimized scale-and-stretch for image resizing," ACM Trans. on Graphics, Vol. 27, No. 5, 2008.
- [16] Y. Guo, F. Liu, J. Shi, Z. H. Zhou and M. Gleicher, "Image retargeting using mesh parametrization," IEEE Trans. on Multimedia, Vol. 11, No. 5, pp. 856-867, 2009.
- [17] D. Vaquero, M. Turk, K. Pulli, M. Tico, and N. Gelfand, "A survey of image retargeting techniques," Proc. of the SPIE, pp.779-814, 2010.
- [18] N. Dalal, and B. Triggs, "Histograms of Oriented Gradients for Human Detection," Proc. of IEEE Conf. on Computer Vision and Pattern Recognition, Vol. 2, pp. 886-893, 2005.
- [19] C. Rother, A. Blake, and V. Kolmogorov, "GrabCut - interactive foreground extraction using iterated graph cuts," ACM Trans. on Graphics, Vol. 23, No. 3, pp. 309-314, 2004.
- [20] P. F. Felzenszwalb, and D. P. Huttenlocher, "Efficient

Graph-Based Image Segmentation,” Int’l J. of Computer Vision, Vol. 59, No. 2, pp.167-181, 2004.

[21] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, “PatchMatch: a randomized correspondence algorithm for structural image

editing,” ACM Trans. on Graphics, Vol. 28, No. 3, pp. 1-11, 2009.

[22] P. Perez, M. Gangnet and A. Blake, “Poisson Image Editing,” ACM Trans. on Graphics, Vol. 22, No. 3, 2003.

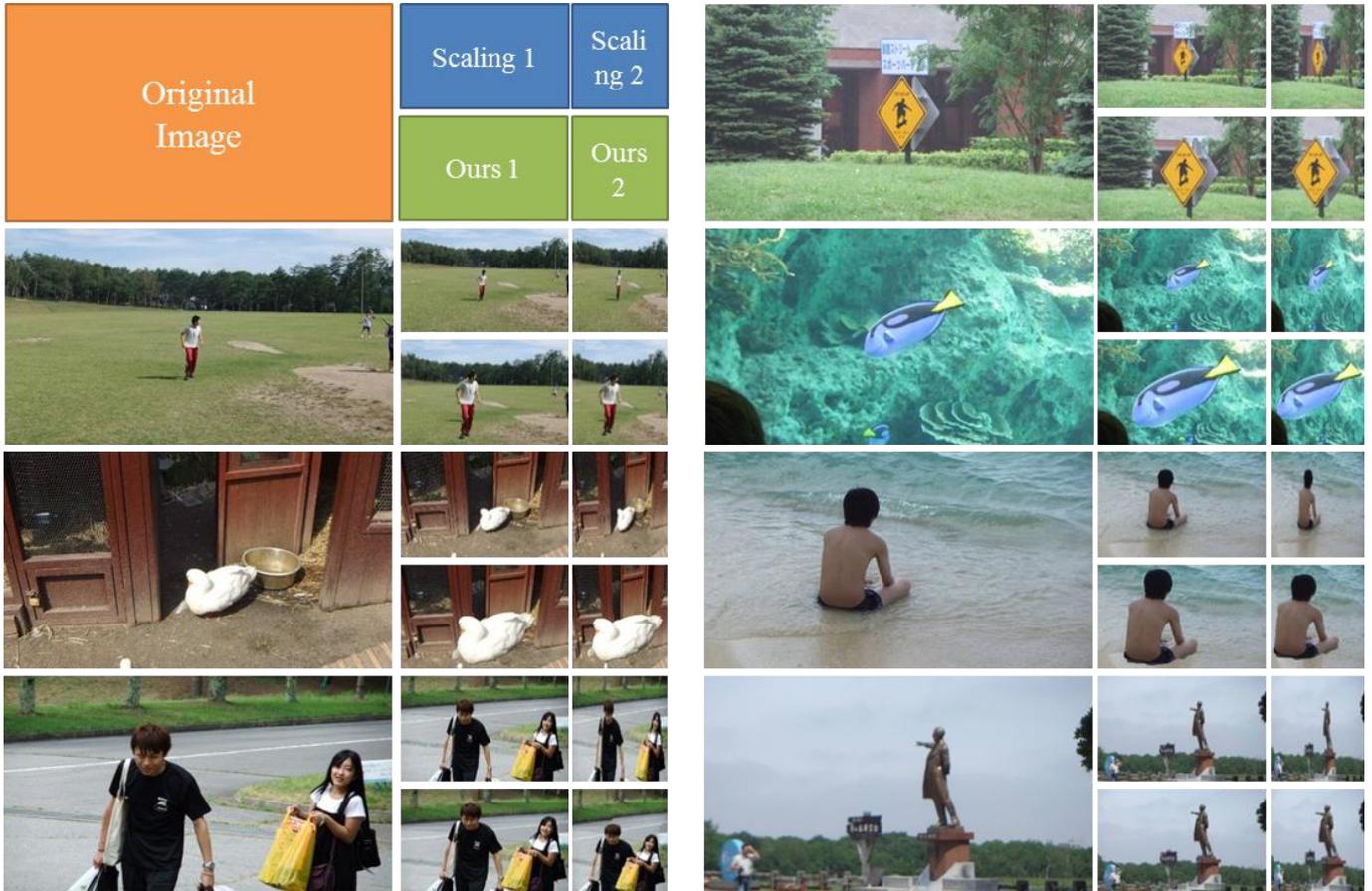


図 7 提案手法によるサムネイルの生成結果

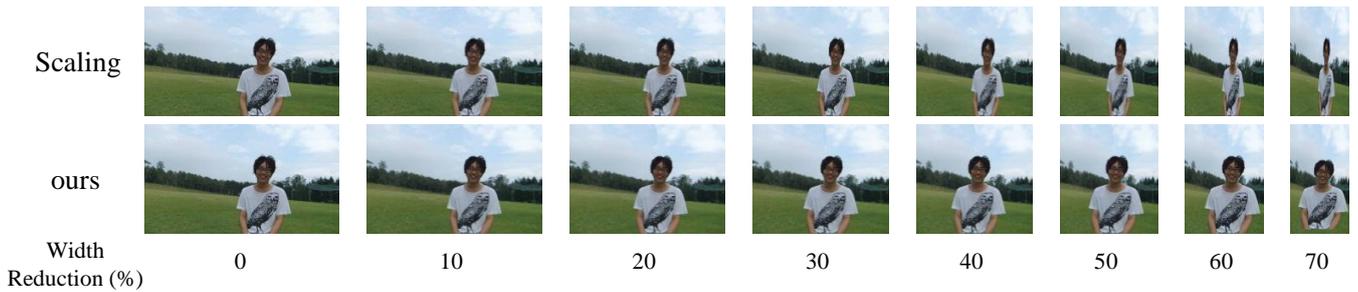


図 8 水平方向の変化への評価



図 9 サムネイル探索実験のシステム画面 問題画像表示画面(左), 探索画面(右)