

画像特徴によるパノラマ画像からの サムネイル画像自動生成法

北田博也[†] 村上和人[†]

パノラマ画像のような横長の画像は、単純に縮小してもサムネイル画像には適していない。本論文では、パノラマ画像から重要な部分領域を抽出し、サムネイル画像を生成する一手法について述べる。重要な部分領域の抽出は、(1)特異点の密度が周囲領域より高いこと、(2)サムネイル画像同士の非類似度ができるだけ高くなること、の2つの条件を満たすように抽出した。サムネイル画像の非類似度はヒストグラム間距離によって求めた。手法をいくつかのパノラマ画像に適用して実験した結果、概ねサムネイル画像として利用できることが確認できた。

Automatic Generation Method of Thumbnail Regions from Panorama Images

Hiroya Kitada[†] and Kazuhito Murakami[†]

This paper proposes a method to extract the most essential and characteristic region, thumbnail image, from a panorama image. The most characteristic region is extracted based on the following two conditions, (1) the density of the feature points is higher than neighboring area and (2) the distance between other thumbnail images in color space calculated by their histograms is larger than a threshold. We confirmed that our proposed method could extract a thumbnail image from some experiments.

1. はじめに

画像をはじめとするデジタルコンテンツの管理は、どのような形式のデータであっても、多くのデータを一覧できる視認性の高さを評価基準としている。例えば、サムネイル画像表示もその一例である。ファイルビューアや画像ビューアなどに用意されているサムネイル画像用の領域は、一般的なアスペクト比で用意されている。しかし、パノラマ画像のように横に長い画像をサムネイル表示する際、単純に全体を縮小する方法では画像を長い横方向のサイズに合わせて縮小することになる。そのため、サムネイル画像中の被写体が小さく画像の内容がわかりづらくなり、サムネイル画像表示が持つ利点の一つである視認性を損ねる結果となる。その様子を図1に示す。

サムネイル画像領域とサムネイル表示するコンテンツのアスペクト比が異なる場合、縮小表示による視認性低下を防ぐ一手法として、サムネイル表示するコンテンツの一部の領域サムネイル画像として切り出して表示することが考えられる(以下、「部分表示型サムネイル画像」と呼ぶ)。しかし、どのような息を切り出すかが課題となる。

画像中の重要な領域を抽出するアルゴリズムはこれまで多く提案されている。木村らはエッジ強度や色相に加えて「主要な被写体は中心付近にある」という予想を加えてアルゴリズムを構成し、興味領域の切り出しに成功している[1]。しかし、「主要な被写体は中心付近にある」という予想を利用できない画像も存在する。例えば、水平方向360°を捉えたパノラマ画像(以下、「360°パノラマ画像」と呼ぶ)のように、中心という概念がない画像には適用できない。

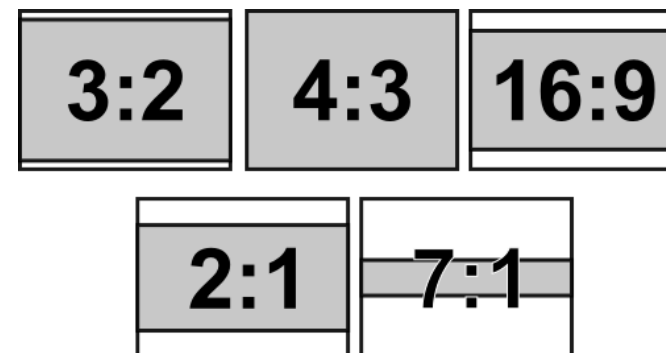


図1 アスペクト比の違う画像をアスペクト比4:3の枠に納めた様子

[†] 愛知県立大学 大学院情報科学研究科
Graduate School of Information Science and Technology, Aichi Prefectural University

この外、重要な領域を抽出する方法として、テンプレートマッチングや学習、識別等の手法を利用した手法も提案されている。しかし、これらの手法では事前にモデルや大規模なデータベースを用意しておく必要があり、被写体の多様さに対応できるように予め準備することは難しい。

そこで本論文では、画像中からサムネイル画像に適した領域を抽出する一手法を提案する。提案手法では、サムネイル画像に適している指標として、特徴点密度が周囲よりも大きいこととした。また、他のサムネイル画像との非類似の指標として、ヒストグラム間の累積差分を基に計算した色空間内距離を利用した。実験画像には 360° パノラマ画像を用い、サムネイル表示したい画像中に重要と考えられる被写体が複数写っている場合や、水平方向の中心という概念がない画像にも対応可能かどうか確認した。以下、2章で基本的な処理の流れについて述べ、次いで3章で実験結果を示し、4章で提案手法の有効性について考察する。

2. 部分表示型サムネイル画像生成手法

2.1 サムネイル画像の種類

「サムネイル画像」は、特別に注釈が加えられていなければ原画像を単純に縮小表示した画像のことを指す（以下、「縮小表示型サムネイル画像」と呼ぶ）。コンテンツ制作において一般的には、原画像から一部領域を切り出す方法は避けるべきとされている。これは、コンテンツを利用するユーザは通常、サムネイル画像と原画像が同じ構図であることを期待するためである。

しかし、原画像から一部領域を切り出してサムネイル画像を生成する手法は効率よく情報を減らすことができる、あるいは情報を圧縮できるという利点がある。さらに、一部領域を切り出すことはサムネイル画像用領域内において被写体を大きく見せる効果もある。横長の画像からサムネイル画像を生成する際には、これらの利点を十分活用できると考えられる。

2.2 部分表示型サムネイル画像の条件

部分表示型サムネイル画像からその元となった原画像のイメージを損なわないようにするため、

- i. サムネイル画像から原画像を連想させる特徴的な被写体が写っていること
- ii. サムネイル画像同士の非類似度が高いこと

といった2つの条件が求められると考えられる。前者は画像の特異性を、後者は画像の非類似性を基に領域を抽出する問題に置き換えることができる。

2.2.1 特異性を示す領域検出

本論文では、画像中に存在する原画像とサムネイル画像を結び付けることができる被写体を含む領域を、「特異点(コーナー)の密度が高い領域」とする。これは、「特徴的な被写体は、複数の特異点を持つテクスチャで構成されている」という経験則によるものである。

本論文では、一般的に用いられることの多いHarrisとStephensの特異点検出法を利用した。画素 (x, y) の輝度を $f(x, y)$ とする。HarrisとStephensの方法は以下のように定義される[2].

$$Harris = \det \hat{C} - k(tr\hat{C})^2 \quad (1)$$

$$\hat{C} = \begin{pmatrix} G_\sigma(f_x^2) & G_\sigma(f_x f_y) \\ G_\sigma(f_x f_y) & G_\sigma(f_y^2) \end{pmatrix} \quad (2)$$

ただし、

k : 定数

$G_\sigma(\bullet)$: ガウス平滑化

f_x : f の x 方向の微分値

f_y : f の y 方向の微分値

$\det \hat{C}$: \hat{C} の行列式

$tr\hat{C}$: \hat{C} のトレース

$$C = \begin{pmatrix} f_x^2 & f_{xy} \\ f_{xy} & f_y^2 \end{pmatrix}$$

である。式(1)の正の極大値を値の大きい順に指定した個数 N だけ、特異点として取り出す。HarrisとStephensの特異点検出法を適用した結果を図2に示す。

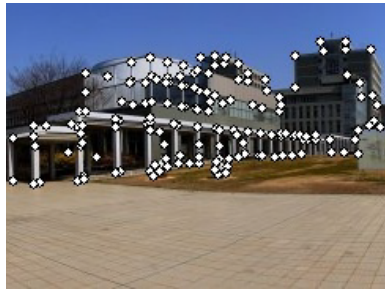


図 2 特異点検出適用例

サムネイル画像の候補領域決定のために、検出された全ての特異点に対して、周囲よりも高い密度で存在する特異点同士をクラスタリングする。本論文においてはこの「周囲よりも密度が高い」という条件を「検出された特異点座標の y 軸方向（垂直方向）の標本分散以下に存在すること」とした。図 3 にクラスタリング結果の一例を示す。検出された特異点を青点で、同一ラベルを赤枠で表示している。このように固まって存在しているように見える特異点を一かたまりとして捉え、特徴的な被写体が写っている領域であるかどうか判断する。その後、同一クラスタに属する全ての特異点について平均座標を相加平均によって計算し、候補領域の中心点とした。2.2.1 節で述べたように、特異点(コーナー)の密度が高い領域を候補領域としたので、より多くの特異点が候補領域の中心に位置するようにするためである。

2.2.2 サムネイル画像間の非類似度評価

本論文では、「サムネイル画像同士の非類似度が高い」という条件を、サムネイル領域の輝度ヒストグラム間距離によって判断した。いくつかの距離尺度が考えられるが、本稿では

$$BD(H_1, H_2) = \sqrt{1 - \sum \sqrt{H_1(I) \cdot H_2(I)}} \quad (3)$$

$$d = \sqrt{BD_R^2 + BD_G^2 + BD_B^2} \quad (4)$$

H_1 : 1番目のヒストグラム

H_2 : 2番目のヒストグラム

I : ヒストグラムの度数 I

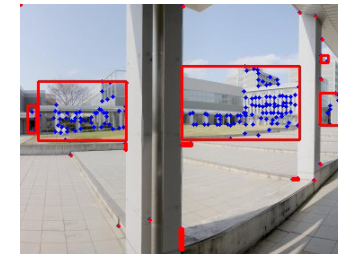


図 3 特異点クラスタリング結果

で定義される Bhattacharyya 距離を用いた。ここで、カラー画像に対応させるため、サムネイル画像間のヒストグラム間距離 d は式(4)に示すように、RGB の各チャンネル間の Bhattacharyya 距離の二乗和とした。これを、違う入力画像で検出された候補領域の全ての組合せに対して計算した。

計算された輝度ヒストグラム間距離を基に、 N 枚の原画像から部分表示型サムネイル画像として候補領域を 1 つずつ決定し、サムネイル画像を作成する。非類似度検査は入力された原画像から検出された候補領域内のみで行う。生成されるサムネイル画像全体の視認性を保つために、サムネイル画像用に切り出す候補領域の組合せは輝度ヒストグラム間距離の最小値が最大となる組合せとする。最小値が最大となる候補領域の組み合わせは、特定の組み合わせのみがきわめて類似する可能性を低くできると期待できるためである。輝度ヒストグラム間距離の最小値が最大となる組合せが複数あったときは、平均値の高い候補領域の組み合わせをサムネイル画像として切り出す。

2.3 提案手法

2.2 で述べた領域をサムネイル画像とするために、以下のようなアルゴリズムを構築した。

<サムネイル画像生成手順>

1. 特異点検出
 N 枚の入力画像全てに対して特異点検出を行い、特異点の座標を求める。
2. 特異点のクラスタリング
検出された全ての特異点に対して、特異点座標の y 軸方向（垂直方向）の標本分散以下の特異点を同一クラスタにクラスタリングする。
3. 候補領域決定

同一クラスタ内の特異点の平均座標を相加平均により求め、作成するサムネイル画像の中心座標とする。

4. 輝度分布間距離を計算

各入力画像で検出された候補領域の全てに対する任意の 2 領域について輝度分布間距離を計算する。

5. 入力画像からサムネイル画像領域を切り出す

作成されたサムネイル画像間の輝度ヒストグラム間距離の最小値が最大となる組み合わせとなるように、 N 枚の入力画像から候補領域を 1 つずつ選びサムネイル画像を生成する。

3. 部分表示型サムネイル画像生成実験

提案手法が「主要な被写体は中心付近にある」という経験則を利用できない画像対しても有効であることを確認するために、 360° パノラマ画像を対象としてサムネイル画像生成実験を行った。実験対象とした画像は、 360° パノラマ画像の中でも一般的なアスペクト比よりも横長になりやすい円筒形に投影する形態の画像を選択した。 360° パノラマ画像の特性から、どの画素も水平方向に連続であるため中心という概念がなく、原画像を連想させる特徴的な被写体が複数写りこんでいる可能性があるため、実験対象として適当であると判断した。

3.1 評価方法

生成された部分表示型サムネイル画像について、

- i. サムネイル画像中から原画像を特定できる被写体が写っているか
- ii. 他のサムネイル画像と区別しやすいか

の 2 点について判断する。i については原画像に写る他の被写体と比較し、生成されたサムネイル画像が原画像を連想しやすいものであるか判断する。原画像には特徴的な被写体が複数写りこんでいる可能性があり、人によるサムネイル画像生成結果も選択する領域にばらつきが出てしまうことが考えられる。従って、原画像中で特徴的な被写体と認められる可能性がある領域が選択できていれば良いものとする。ii は非類似度の判断基準としている輝度ヒストグラム間距離の値によって判断する。原画像全体を対象とした場合よりも、サムネイル画像領域の輝度ヒストグラム間距離値が大きいならば、原画像よりも色空間距離内において非類似であると言える。

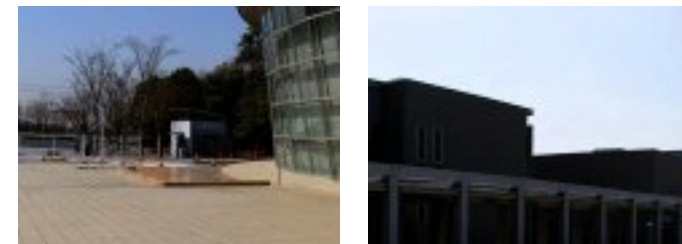
3.2 生成結果

提案手法で生成されたサムネイル画像の一例として、雰囲気の良い 360° パノラマ画像 2 枚を入力しサムネイル画像を生成した結果を 3 組掲載する。図 4 から図 6 に生成された部分表示型サムネイル画像を、また、図 7 から図 9 に生成された

部分表示型と原画像との位置関係を示す。表 1 に提案手法で非類似度を判断している輝度ヒストグラム間距離を示す。数値は原画像全体を対象に計算した値とサムネイル画像領域のみを対象に計算した値である。ヒストグラム間距離は原画像全体を対象に計算した値とサムネイル画像領域のみを対象に計算した値の最大値がどちらも 1 となるように正規化した。原画像よりも互いに非類似であることによりサムネイル画像として視認性が高いことを示す。

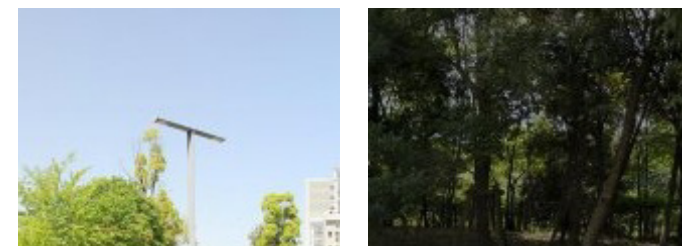
表 1 原画像間及びサムネイル画像間の輝度ヒストグラム間距離

原画像	原画像全体	サムネイル画像領域
1 組目	0.0266	0.5694
2 組目	0.0392	0.8561
3 組目	0.0294	0.4933



(a) 1 組目 1 枚目 (b) 1 組目 2 枚目

図 4 サムネイル画像生成結果：1 組目



(a) 2 組目 1 枚目 (b) 2 組目 2 枚目

図 5 サムネイル画像生成結果：2 組目



(a) 3組目 1枚目 (b) 3組目 2枚目
図 6 サムネイル画像生成結果：3組目

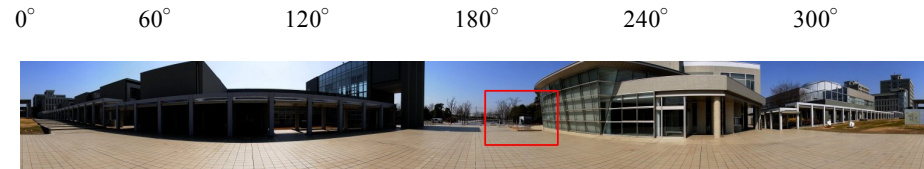


(a) 2組目 1枚目



(b) 2組目 2枚目

図 8 原画像と生成されたパノラマ画像の位置関係：2組目



(a) 1組目 1枚目



(b) 1組目 2枚目

図 7 原画像と生成されたパノラマ画像の位置関係：1組目



(a) 3組目 1枚目



(b) 3組目 2枚目

図 9 原画像と生成されたパノラマ画像の位置関係：3組目

4. 考察

原画像 1 組目(a)からは、撮影したアングルの都合上見えづらくなっているが石像が存在している領域がサムネイル画像として選ばれた。しかし、原画像もサムネイル画像も画像中から石像の存在を見つけ出すことができない。よって、2.2 で述べた条件のうち「i.サムネイル画像から原画像を連想させる特徴的な被写体が写っていること」という条件を満たさない。1 組目(b)は、生成されたサムネイル画像の大部分を逆光によって黒く映ってしまった領域が選ばれている。そのため、非類似度の判断基準とした輝度ヒストグラム間距離は、原画像全体を対象としたときと比較しても大きく離れている。しかし、逆光のため黒く写ってしまった建物の屋根だけでは撮影地点を特定することはできず、「原画像を連想させる」領域とは言えない。サムネイル画像として不適切と考えられる。

原画像 2 組目(a)は 0° から 60° 付近に木立の中に白い建物が存在し、画像中でもとりわけ特徴的な領域といえる。逆に 2 組目(b)は原画像全体に目立った被写体が存在しないという特徴がある。自動生成による結果では、2 組目(b)は 180° から 360° のあたりに存在する木立の中央をサムネイル画像として抽出した。輝度ヒストグラム間距離が離れているため一見して違う画像と認識できるが、ユーザにとっては原画像を連想しづらく不適切な領域と判断できる。また、木立のように自然の風景などを被写体とした場合は特異点が多く検出される傾向にあり、特徴的ではない領域が選択されてしまったと考えられる。

3 組目はそれぞれ屋根のある通路から撮影したパノラマ画像を原画像としている。3 組目の原画像は通路の中からの撮影という特徴的な構図を持っている。しかし、通路の奥行きが表現されている中央付近ではなく、通路の柱の間から見える景色をサムネイル画像に選んできたことは、サムネイル画像から原画像へユーザを誘導することに大きく貢献できると考えられる。輝度ヒストグラム間距離の値から、原画像と比較してもより非類似度を示す領域を抽出できたと言える。

これらの結果から、比較対象が類似した雰囲気を持つパノラマ画像であっても、個々の原画像の雰囲気を損なうことなくサムネイル画像が生成できたといえる。しかし、非類似度を輝度ヒストグラム間距離による判断や、特徴的であるという条件を特異点の密集とただけでは不十分なケースも多く見られ、特徴量の改良が必要である。

5. おわりに

本論文では、パノラマ画像から特徴的な領域を自動抽出する一手法を提案した。特異点という比較的簡易な画像特徴を利用して、比較的原画像を連想しやすい部分領域を部分表示型サムネイル画像として抽出可能であることを確認した。一方で、画像特

徴からは特徴的な領域と考えられるものの、原画像を連想させることが難しいと判断される領域を選択してしまうケースも多数認められた。

今後は、サムネイル画像候補領域のための特徴量を改良することによって主観評価とより合致する生成結果を出せるよう改良を加えていく予定である。

参考文献

- [1] 木村 雅之, 山内 真樹, 岡林 一郎; “注目度に基づく画像からの興味領域抽出”, 電子情報通信学会技術研究報告 ICD 集積回路, No.570(20060120), pp.35-39, Jan. 2006.
- [2] 金澤 靖, 金谷 健一; “コンピュータビジョンのための画像の特徴点の抽出”, 電子情報通信学会誌, No.12(20041201), pp.1043-1048, Dec. 2004.