

複数の注目領域を用いた写真の主観的品質の識別

西山 正志^{†a)} 岡部 孝弘[†] 佐藤 洋一[†] 佐藤いまり^{††}

Visual Quality Classification Using Multiple Subject Regions

Masashi NISHIYAMA^{†a)}, Takahiro OKABE[†], Yoichi SATO[†], and Imari SATO^{††}

あらまし 本論文では、人間が写真に対して持つ主観的な品質を識別する手法を提案すると共に、その写真の整理と提示への応用について述べる。提案手法は、品質識別のために、視覚注目の強度を表す顕著度により複数の注目領域を検出する。検出された複数の注目領域と背景領域から、品質識別のための特徴量を抽出する。これにより提案手法は、ぼけを用いて単一の注目領域を検出していた従来手法では取り扱うことのできなかった写真から、品質識別に有効な特徴を抽出することができ、識別率のさらなる向上を狙うことができる。実験により、品質識別の精度が従来手法と比べて改善することを確認した。大量に収集された写真を品質識別を用いて自動で整理し提示する応用として、写真トリミング、写真整理を紹介する。

キーワード 写真, 品質, 注目領域

1. はじめに

ユーザがカメラやインターネットで簡単に収集できる写真が大量に増加していることに伴い、収集された写真を自動で整理し提示することが求められている。この要求を満たすためにインターネット上の大量の写真から整理・提示の知識を抽出する手法が近年注目されている。

Snively ら [1] は、シーンの 3 次元形状と見えのモデルを構築し、写真提示に利用する手法を提案している。一方、シーンをモデル化するのではなく、写真の良し悪しという人間が写真に対して持つ主観を統計的に学習し、写真提示に利用する手法 [2] ~ [7] も提案されている。これらの手法では、図 1 に示すような高品質画像と低品質画像とを識別することを目的としている。品質識別を用いて、キーワードで検索した写真を人間が感覚的に好む順に再整理などの応用が提案されている。

品質識別を行うこれらの従来研究では、識別に用い

る特徴量を画像全体、もしくは、単一の注目領域から抽出していた (詳細は 2. で議論)。しかしながら、撮影者が写真を通じて伝えたいことは、一つの注目領域のみによって表現されるとは限らない。したがって、品質識別の精度を高めるためには、複数の注目領域と背景領域との関係を考慮することが有効であると考えられる。

そこで本論文では、複数の注目領域と背景領域から経験則に基づいた特徴量を抽出すると共に、それらを効果的に組み合わせることにより、主観的品質の識別を行う。さらに、応用の具体例として、複数注目領域を用いた品質識別により、写真トリミングや写真整理を行う方法を紹介する。

ここで、本論文の主要な貢献は以下のとおりである。

- 品質識別のために、複数の注目領域と背景領域から特徴量を抽出し統合する手法を新たに提案する。写真撮影の経験則を考慮した品質識別を行うことができ、従来手法と比べて認識性能を高めることができる。
- これまでにない新しい応用として、提案手法の品質識別を用いた写真トリミングを提案する。
- 実用的な応用の写真整理において、提案手法の品質識別の有効性を確認した。提案手法は従来手法と比べて、人が行った主観評価の傾向と近い結果が得られることを確認した。

以下、2. で品質識別の概要、3. で提案手法の詳細に

[†] 東京大学生産技術研究所 〒 153-8505 東京都目黒区駒場 4-6-1
Institute of Industrial Science, The University of Tokyo 4-6-1
Komaba, Meguro-ku, Tokyo, 153-8505 Japan

^{††} 国立情報学研究所 〒 101-8430 東京都千代田区一ツ橋 2-1-2
National Institute of Informatics 2-1-2 Hitotsubashi,
Chiyoda-ku, Tokyo 101-8430, Japan

a) E-mail: masashi1@iis.u-tokyo.ac.jp

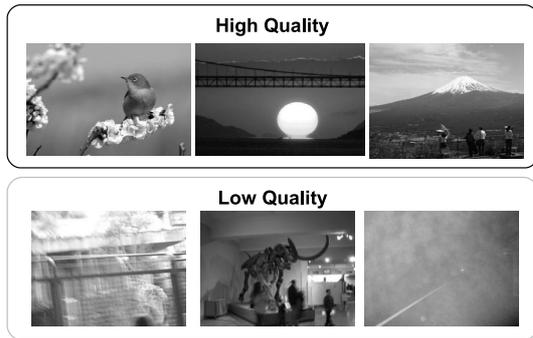


図 1 高/低品質の写真の例．本論文では，主観的な品質に対する 2 クラス識別問題を取り扱う．大勢の人間が共通に高品質と感ずる写真を自動で判定できる手法の実現を目指す．

Fig. 1 Examples of high/low quality photos.

について述べる．次に，4. で品質識別の基本性能，5. で 2 つの応用を評価する．最後に，6. でまとめる．

2. 複数領域を用いた品質識別

2.1 主観的な品質を決める要因

主観的な品質識別の性能を高めるためには，撮影技術，注目領域の変動要因を考慮した特徴量の設計が重要である．本論文では，品質識別の特徴量を経験則に基づきトップダウンに決め，その特徴量をインターネット上の大量の写真から抽出し識別器を学習する．ここで，主観的な品質を決める変動要因について写真の教則本 [10] を参考にし列挙する．

- 撮影技術: 手ぶれをしていない，コントラストが高い，輝度ヒストグラムに偏りが無いなど，撮影の基本的な条件が満たされている．

- 注目領域: 撮影者が伝えようとする被写体を含む注目領域を，写真を観賞する人間が直ちに理解することができる．なお，注目領域は一つと限らない．

- 好み: 写真を観賞する人間が，その写真の注目領域に共感できる．

個々のユーザの好みも主観に大きく影響する．注目領域が決まらなければ，写真からの受け取り方がユーザによって大きくばらつく．これにより好みが決まらなくなり，ユーザが低品質と判定する可能性が高くなる．よって，品質を識別するためには，注目領域の検出は重要である．

2.2 従来手法

ここでは品質識別の従来手法について述べる．Tong ら [2] と Datta ら [4] は，2.1 の変動要因を考慮せず

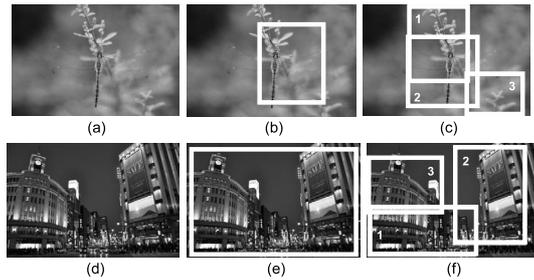


図 2 注目領域と背景領域の例．(a),(d) は原画像，(b),(e) の枠は手法 [5] で検出された注目領域，(c),(f) の枠は提案手法で検出された注目領域を表す．各画像で 3 個の注目領域 R_i と背景領域を検出した．枠内の数字は領域番号 i を表す．注目領域以外の部分を背景領域とする．

Fig. 2 Examples of subjective and background regions.

品質識別を行う手法を提案している．識別に用いる特徴量を，画像検索で一般的に使用される低レベル特徴量を組合せ，その中から選択する．この特徴量は，訓練サンプルに強く依存するため品質識別に最適であるとは言えない．訓練サンプルの分布から外れる写真に対しては識別性能を維持できない．

一方，Ke ら [3] は，写真撮影の経験則に基づき，2.1 の撮影技術を考慮した特徴量を品質識別に用いる手法を提案している．これにより，訓練サンプルへの依存度が減る．しかしこの手法は，画像全体から特徴量を算出しているため，2.1 の注目領域が考慮されていなかった．写真を撮影する際に注目領域を設定することは，経験則上重要である．注目領域を検出することで，品質識別に有効な特徴量の抽出を狙うことができる．

Luo ら [5] と Loui ら [6] は，注目領域と背景領域を検出し，特徴量の抽出に利用する手法を提案している．Luo ら [5] の手法は，注目領域を際立たせるために背景をぼかす撮影技術を仮定し，各画素のぼけ度合いから注目領域を決定する．例えば，図 2(a) のように昆虫など小さな被写体はこの仮定が成り立つが，(d) のように風景などの大きな被写体にはこの仮定が成立しない．このため (e) のように注目領域と背景領域の違いがなくなり，注目領域の特定による特徴量の抽出に意味がなくなる．さらに，これらの手法 [5], [6] は，注目領域は単一と仮定している．実際に写真を撮影する際は，注目領域が単一とは限らず複数あることが多い．

2.3 改良の方針

本論文では，複数の注目領域を検出し，複数の注目領域と背景領域との間の関係を考慮した特徴量を品質

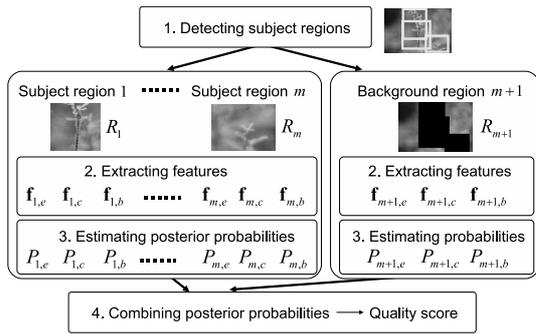


図3 提案手法の流れ．4つのステップをもつ．
Fig. 3 Flow of our method.

識別に用いる．注目領域の検出には，従来のぼけ具合ではなく，視覚注目の強度を表す顕著度を用いる．図2(c),(f)に，提案手法を用いて複数の注目領域を検出した例を示す．提案手法は注目領域間の重なりを許す．(c)は(b)と比べて，主題である昆虫と植物が検出されている．同様に(f)も(e)と比べて，それぞれの建築物が検出されている．複数の注目領域と背景領域から，撮影技術を考慮した特徴量を抽出し統合することで品質を決定する．提案手法は，2.1で述べた変動要因の注目領域を従来手法より詳細に取り扱うことを狙う．残る変動要因の好みは大量の訓練サンプルから統計的に学習することを狙う．以上により，品質識別の精度向上を実現する．

3. 提案手法のアルゴリズム

提案手法は図3の4つのステップにより構成される．第1ステップでは，画像が与えられると注目領域と背景領域を検出する(3.1)．第2ステップでは，各領域に対して，コントラストやぼけなど写真技術を反映した基本特徴量を求める(3.2)．第3ステップでは，各特徴量が与えられた時に高品質と分類される割合である事後確率を求める．(3.3)．第4ステップでは，基本特徴量に対する事後確率を領域間で統合し品質識別の判定を行う(3.4)．

上記のアプローチ以外にも，例えば，複数の注目領域に対する基本特徴量を単純に連結する特徴量を用いることが考えられる．しかし，この方法では実験結果(図5のLinking)で示すように品質識別の精度が得られなかった．このため，提案手法は，基本特徴量を個別に評価し統合するアプローチをとる．

訓練サンプルには，インターネット上の写真デー

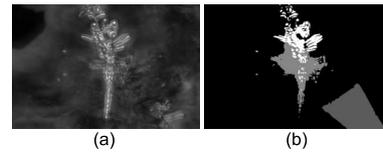


図4 クラスタリング処理の例．図2(a)に対して顕著度を求めた結果を(a)に示す．各画素において白に近い方が顕著度が高い．この顕著度を用いてクラスタリングを行った結果を(b)に示す．画素値がクラスタのラベルを表す．

Fig. 4 Example of clustering process.

タベース(DPChallenge[8], Photo.net[9])と写真の教則本[10]を用いる．DPChallenge, Photo.netでは様々な写真に対して様々な人間が主観的なスコア付けを行っている．スコアの高い写真を高品質，スコアの低い写真を低品質とする．教則本では写真家が品質を定義している．提案手法は，教師ラベルを高品質と低品質とし，2クラス分類を行う．これら大量の写真を利用することで，大勢の人間が共通でもつ主観を統計的に学習する．芸術的に優れているが平均的な品質からは大きくはずれている写真は本論文の対象外とする．以下，各ステップの詳細を説明する．

3.1 複数の注目領域の検出

提案手法は，顕著度を用いて画像中で注目を集める部分を検出し，複数の注目領域としてそれぞれの領域を割り当てる．

顕著度の算出には，人間の視覚機能の低レベル部のモデル化を狙ったIttiらの手法[11]を用いる．この手法は，スケール毎に色，輝度，方向について局所的な顕著度を算出し，ボトムアップに統合された結果を画像の顕著度とする．色は赤値と緑値との差，青値と黄値との差，輝度はDifference of Gaussianの出力，方向はGaborフィルタの出力とする．エッジなど周辺領域と比較して異なる特性をもつ領域の顕著度は高くなり，平坦な領域の顕著度は低くなる．顕著度を求めた例を図4(a)に示す．

次に，顕著度を用いてクラスタリングを行い注目領域を決定する．提案手法は，次式のベクトル v に対しk-Means法[12]を適用する．

$$v = (n(x), n(y), n(s_{x,y})) \quad (1)$$

ここで， $s_{x,y}$ は画像中の座標 (x,y) の顕著度， $n(\cdot)$ は値がとり得る範囲を正規化する関数である．この関数は，とり得る値の最小から最大が $[0, 1]$ となるよう値を変換する．k-Means法で獲得した k 個のクラス

タの中で平均顕著度が高い m 個のクラスタを注目領域 R_i とする．領域番号 i は，平均顕著度が高い順に $1, 2, \dots, m$ とする．クラスタリングした例を図 4 (b) に示す．以後の処理での取り扱いを簡単にするために，注目領域 R_i は各クラスタに含まれる座標 (x, y) の外接矩形とし，画像中の注目領域以外の部分を背景領域 R_{m+1} としている．クラスタ数の決め方については 4.2 で考察する．Mean-shift mode detector [13]，混合ガウス分布に対する EM アルゴリズム [12] など，他のクラスタリング手法の適用も検討したが，計算量とメモリ量が増加するのみで品質識別の精度には影響しなかった．そのため，提案手法では単純な k-Means 法を用いている．

3.2 各領域に対する基本特徴量の抽出

各領域 $R_i (i = 1, \dots, m+1)$ に対して，コントラストやぼけなど写真技術を反映した基本特徴量を算出する．写真の教則本 [10] と Ke らの文献 [3] で述べられている撮影の経験則を参考にして，基本特徴量に以下のエッジ，色，ぼけを用いる．

- エッジ $f_{i,e}$: 縦方向，横方向の Sobel フィルタの出力値から 256 ビンのヒストグラムを生成する．
- 色 $f_{i,c}$: RGB 空間を $8 \times 8 \times 8$ に区切り 512 ビンのヒストグラムを生成する．
- ぼけ $f_{i,b}$: 離散フーリエ変換で各周波数成分に対する振幅値を求め，周波数空間内で 32×32 にリサンプリングし，1024 次元のベクトルを生成する．

3.3 事後確率への変換

各基本特徴量が与えられたときの事後確率を求める．提案手法は，Support Vector Machines (SVM) [14] の出力を事後確率へ変換する．SVM の訓練サンプルには，人が与えた主観スコアに基づき品質ラベルが教示された写真を用いる．高・低の品質ラベルを教師とした 2 クラス分類を行う．

注目領域 $R_i (i = 1, \dots, m)$ における基本特徴量 $f_{i,j} (j = e, c, b)$ が与えられた時，SVM の出力 $s_{i,j}$ を次式で表す．

$$s_{i,j} = SVM_{subject,j}(f_{i,j}) \quad (2)$$

背景領域に対しても同様に表す．

$$s_{m+1,j} = SVM_{background,j}(f_{m+1,j}) \quad (3)$$

SVM の出力は識別超平面からの正規化された距離であるため，文献 [15], [16] の方法で次式のシグモイド関数に当てはめ事後確率を求める．

$$P_{i,j}(high|s_{i,j}) = \frac{1}{1 + \exp(As_{i,j} + B)} \quad (4)$$

ここで， A, B は訓練サンプルから決まる定数である．高品質の場合には $P_{i,j}$ が 1，低品質の場合には $P_{i,j}$ が 0 へ近づくように，クロスエントロピーをニュートン法で最小化することで A, B が求まる．詳細は文献 [15], [16] を参照のこと．

3.4 領域間の事後確率の統合

各基本特徴量に対する事後確率を領域間で統合し，品質識別の出力を決定する．統合には Meynet ら [17] のように事後確率間の積，和，中央値などを用いることが考えられるが，単純な方法では性能改善を得ることができなかった．そこで本論文では，事後確率を要素とした次式の特徴量を用いる．

$$f_{all} = (P_{1,e}, P_{1,c}, P_{1,b}, P_{2,e}, P_{2,c}, P_{2,b}, \dots, P_{m+1,e}, P_{m+1,c}, P_{m+1,b}) \quad (5)$$

統合された特徴量による SVM の出力 s_{all} を次式で表す．

$$s_{all} = SVM_{all}(f_{all}) \quad (6)$$

ここで SVM は，3.3 と同様に，教師ラベルは高品質と低品質とする．この出力より品質を最終的に決定する．

さらに，領域間の関係を事後確率間の積で表した要素を f_{all} の要素に加えた特徴量を次式で定義する．

$$f_{all'} = (P_{1,e}, \dots, P_{m+1,b}, P_{1,e} \cdot P_{2,e}, P_{2,e} \cdot P_{3,e}, \dots, P_{m-1,b} \cdot P_{m,b}, P_{m,b} \cdot P_{m+1,b}) \quad (7)$$

この特徴量も同様に SVM で識別する．実験では f_{all} と $f_{all'}$ との識別性能の比較を行う．

4. 品質識別の基本性能の評価

4.1 従来手法との識別性能の比較

提案手法の識別性能を確認するために基本実験を行った．評価には，写真品質がスコア付けされデータベース (DPChallenge [8], Photo.net [9], 教則本 [10]) を用いる．文献 [7] でも述べられているようにスコア付けの人数，評価項目，傾向がデータベース毎で異なる．本論文では，主観を決める変動要因の写真技術，注目領域，好みに多様性をもたすため，3 つのデータベースを混合し，合計 15544 枚の写真で評価を行った．DPChallenge からは，スコア分布の上位 10 % と下位

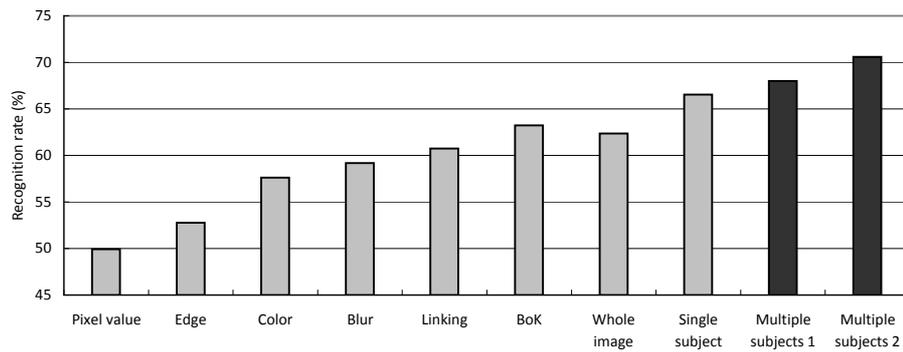


図 5 品質識別の性能比較 .

Fig. 5 Comparison of performance of quality classification.

10 % の 13420 枚を用いた . Luo ら [5] が WEB 上で写真を公開している . Photo.net からは , スコア分布の上位 20 % と下位 20 % の 1800 枚を用いた . Datta ら [4] が WEB 上でリンク先を公開している . 写真の教則本からは 324 枚を用いた . 各データベースを半数に分けた後に混合し , 一方を訓練サンプル , 残りをテストサンプルとした .

本節の実験では , 品質識別に用いる特徴量について性能比較を行う . 特徴量を入力する識別器は SVM とし , 識別時に計算量の少ない線形カーネルを用いた . 性能向上のためには非線形のカーネルを用い , そのパラメータを交差確認で決めることも考えられるが , 学習時と識別時の計算コストが増すため今後の課題とする . 注目領域を決定するパラメータは $k = 12, m = 5$ とした . このパラメータと写真のカテゴリに関する実験は 4.2 で述べる . 以下 , 性能比較を行う 10 個の特徴量を列挙する .

Pixel value: 単純な基本特徴量である . 画像の大きさを 32×32 にダウンサンプリングし , RGB 値をラスタースキャンして特徴量を生成する .

Edge: 画像全体から求めたエッジの基本特徴量 $f_{whole,e}$ である .

Color: 画像全体から求めた色の基本特徴量 $f_{whole,c}$ である .

Blur: 画像全体から求めたぼけの基本特徴量 $f_{whole,b}$ である .

Linking: 複数の注目領域と背景領域から求めた基本特徴量を連結した特徴量 $f_{all'} = (f_{1,e}^t, f_{1,c}^t, f_{1,b}^t, \dots, f_{m+1,e}^t, f_{m+1,c}^t, f_{m+1,b}^t)$ である .

BoK: 物体認識や画像検索で用いられる Bags-of-

keypoints [19] が特徴量である .

Whole image: Ke ら [3] のように , 画像全体から求めた基本特徴量を用いる . この実験では注目領域の取り扱いについて比較を行いたいため , この文献の特徴量ではなく , 本論文で述べた基本特徴量 $f_{whole,e}, f_{whole,c}, f_{whole,b}$ を用いた . 提案手法と同様に各基本特徴量の事後確率を求め統合した .

Single subject: Luo ら [5] のように , 一つの注目領域と背景領域とから求めた基本特徴量を求める . この文献の注目領域の検出方法を用いたが , 基本特徴量は各領域のエッジ , 色 , ぼけとした . 提案手法と同様に各基本特徴量の事後確率を求め統合した .

Multiple subjects 1: 提案手法の特徴量 f_{all} である .

Multiple subjects 2: 提案手法の特徴量 $f_{all'}$ である .

図 5 に各特徴量の識別性能を示す . 指標には , 正しい品質に分類された割合を表す認識率を用いる . 単純な Pixel value と比べると , 写真技術を考慮した基本特徴量 Color, Edge, Blur が識別性能を得られる . 基本特徴量を単純に連結した Linking よりも , 各基本特徴量を個別に評価し統合する Whole image, Single subject, Multiple subjects 1, Multiple subjects 2 は性能が高い . 提案手法の Multiple subjects 1, Multiple subjects 2 は , 物体認識や画像検索で有効な特徴量の BoK より性能が得られる . さらに提案手法は , 単一の注目領域の Single subject より優れている . 領域間の関係を考慮する Multiple subjects 2 が , 考慮しない Multiple subjects 1 より優れている . 以上 , 複数注目領域が単一注目領域より品質識別の性能を得ることができることを確認した . 今後はデータベースを充実させ , さらに認識率の向上を目指す .

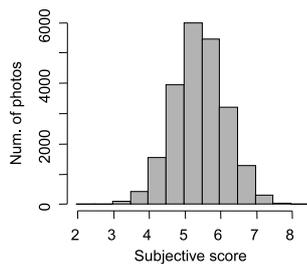


図 6 カテゴリ'Nature'において、各写真に与えられた主観スコアのヒストグラム。

Fig.6 Histogram of subjective score given to each photo in the category 'Nature'.

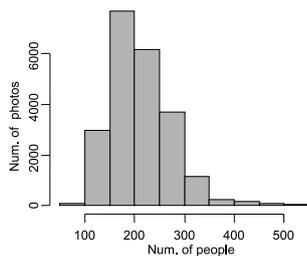


図 7 カテゴリ'Nature'において、各写真に対して主観スコアを与えた人数のヒストグラム。

Fig.7 Histogram of people who give subjective scores for photos in the category 'Nature'.

4.2 写真のカテゴリ毎の識別性能

提案手法の品質識別の性能が、写真の被写体のカテゴリに対して依存するかどうかを確認する。前節では様々なカテゴリに属している写真を一つにまとめ実験を行っていたが、ここでは写真をカテゴリ毎に分けて評価した。カテゴリとして Animal, Cityscape, Landscape, Nature, Rural, Seascape, Sky, Still-life を用いた。これらのカテゴリは、DPChallenge においてユーザが写真を投稿する時に指定するタグである。このタグは一枚の写真に対して複数指定できる。この実験では、ポートレートなど人を含むカテゴリは除外した。各カテゴリに属する写真の枚数は、13412 枚, 5117 枚, 16134 枚, 22359 枚, 4928 枚, 3733 枚, 4264 枚, 13864 枚であった。最も枚数の多いカテゴリ Nature において、主観スコアのヒストグラムを図 6 に、主観スコアを与えた人数のヒストグラムを図 7 に示す。他のカテゴリも枚数は違うが同様の分布であった。主観スコア分布の上位 10% の写真を高品質、下位 10% の写真を低品質のラベルとし、品質識別の学習とテストに用いた。

図 8 に、写真をカテゴリ毎に分けた場合の認識率を

示す。ここでは、従来手法 Single subject と提案手法 Multiple subjects 2 を比較する。まず、提案手法のみの認識率をカテゴリ間で比較すると、提案手法はカテゴリ間で得手不得手があることが分かる。認識率は、カテゴリ Nature, Landscape で高く、Seascape, Sky で低い。次に、カテゴリ毎に提案手法と従来手法との認識率を比べると、提案手法は全てのカテゴリにおいて改善していることが分かる。カテゴリ Animal, Seascape では改善の幅が小さいが、カテゴリ Cityscape, Rural, Sky では大きく改善していた。同一カテゴリでも様々な写真が含まれるため一概には言えないが、写真中に複数の被写体が配置されることが多いカテゴリで提案手法が有効に働くと考えられる。

次に、注目領域の個数について考察を行う。提案手法では、その個数は全カテゴリで一定と仮定しているが、カテゴリ毎に最適な個数は異なると考えられる。これを確認するため、図 9 に、注目領域の個数 m を増やした場合の品質識別の性能の変化を示す。比較のため、従来手法である単一の注目領域の結果も $m = 1$ で示す。識別性能と m との関係について、8 カテゴリ中で大きく分けて 4 つの傾向が見られた。図中に 4 つの傾向に対する代表カテゴリの結果を示す。Animals, Landscape, Stilllife では m が増加するにつれて性能が改善した。Nature では m に関係なく性能がほぼ一定であった。Rural では $m = 6$ 以降で性能が改善しなかった。Cityscape, Seascapes, Sky では $m = 9$ で性能が最大になりその後は性能が低下した。ここでは注目領域を決定するもう一方のパラメータを $k = 12$ としたが、これを変更した場合も同様の傾向が見られた。実験結果より、カテゴリ毎に最適な領域数は異なることが分かった。一方で、提案手法は領域数が増加しても劇的な識別性能の変化はあまり見られないことが分かった。さらなる識別性能の改善には、 m を決定するためのカテゴリ認識を提案手法の前処理として導入することが考えられる。

以上、複数の注目領域を用いる提案手法が従来手法と比べて有効であることを確認した。

5. 応用

提案手法の品質識別を応用した 2 つの例を述べる。

5.1 写真トリミング

一枚の写真から高品質の部分画像を、提案手法を用いてトリミングする新しい応用について述べる。ここでは、図 10(a) の写真を対象とした。前提条件として、

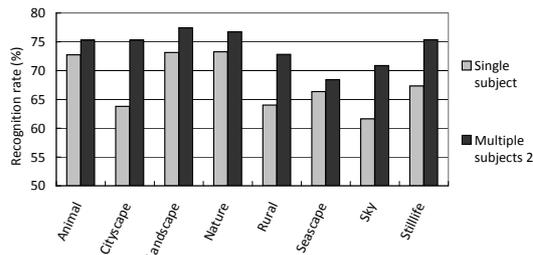


図 8 写真をカテゴリ毎に分けた場合の品質識別の性能 .
Fig.8 Performance of quality classification in the case that photos are categorized.

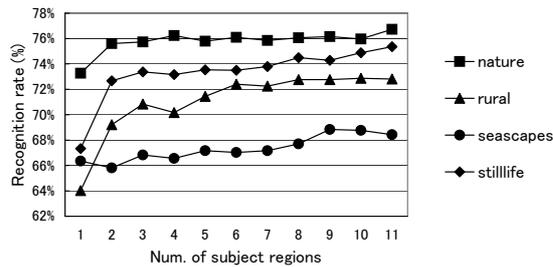


図 9 注目領域の個数を増やした場合の品質識別の性能 .
Fig.9 Performance of quality classification in the case that the number of subject regions is increased.

ユーザが指定した写真中の白枠で表す領域はトリミング結果に含まれるとする .

まず、指定領域を含む部分画像を、位置 (x, y) 、サイズ (w, h) を様々に変え生成する . ただし、アスペクト比は 4:3 または 3:4 と固定した . 生成した部分画像それぞれに対して、3. で述べた提案手法により、品質識別の出力 $s_{x,y,w,h}$ を求める . この出力は、位置、サイズが近い部分画像同士で近い値になることが多い . そのため、単純に $s_{x,y,w,h}$ で品質をソートすると、トリミングされた部分画像の見えが非常に似る . 複数の部分画像をトリミングしユーザに選択してもらおうといったインターフェースを考えた場合、複数の部分画像の見えがほとんど同じでは意味が無い . そこで、ユーザに対して違いが分かるように結果を提示するために、k-Means 法でクラスタリングを行う . クラスタリングのためのベクトルには $(n(x), n(y), n(w), n(h), n(s_{x,y,w,h}))$ とした .

図 10(b)-(e) に写真トリミングの例を示す . ここでは、品質識別に 4. で述べた提案手法 (Multiple subjects 2) を用いた . 上位の部分画像は下位のものとは比べて、山と花のバランスがうまく配置されている . 同

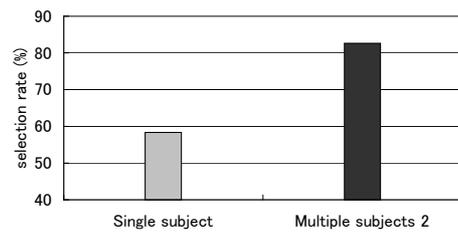


図 11 写真トリミングの結果 (図 10) に対する主観評価 . 被験者に高品質と低品質の部分画像のペアを提示し、高品質の部分画像が選択された割合を示す .
Fig.11 Subjective assessment against the results of photo cropping.

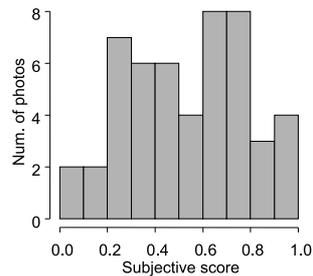


図 12 一対比較による主観スコアのヒストグラム .
Fig.12 Histogram of subjective score using subjective assessment.

様のことが、葉に対する花の配置にも言える . ただし、提案手法で低品質となった部分画像の一部については、見え方が高品質のものと同じく必ずしも低品質とはいえない .

トリミングされた部分画像に対するユーザからの主観評価の結果を図 11 に示す . 24 名の被験者に対して高品質と低品質の部分画像のペアを提示し、高品質の部分画像が選択された割合を評価に用いた . 被験者は、図 10 で示した各品質の部分画像の全組合せ $(3 \times 3 = 9 \text{ 種類})$ のペアについて評価した . ここでは、Multiple subjects 2 だけでなく、Single subject についても評価した . この結果から、従来手法と比べて提案手法は被験者の同意を得られていることが分かる .

5.2 写真整理

デジタルカメラで撮影した写真コレクションの整理を行う . プロフェッショナルが撮影した写真ではなく、素人が撮影した写真を対象とする . 例えば、高品質の写真はスライドショーとして提示し、低品質の写真は削除候補として提示することを想定する .

この実験では筆者らが撮影した 50 枚の写真を用いた . それの写真に対し、Thurstone の一対比較法 [18]

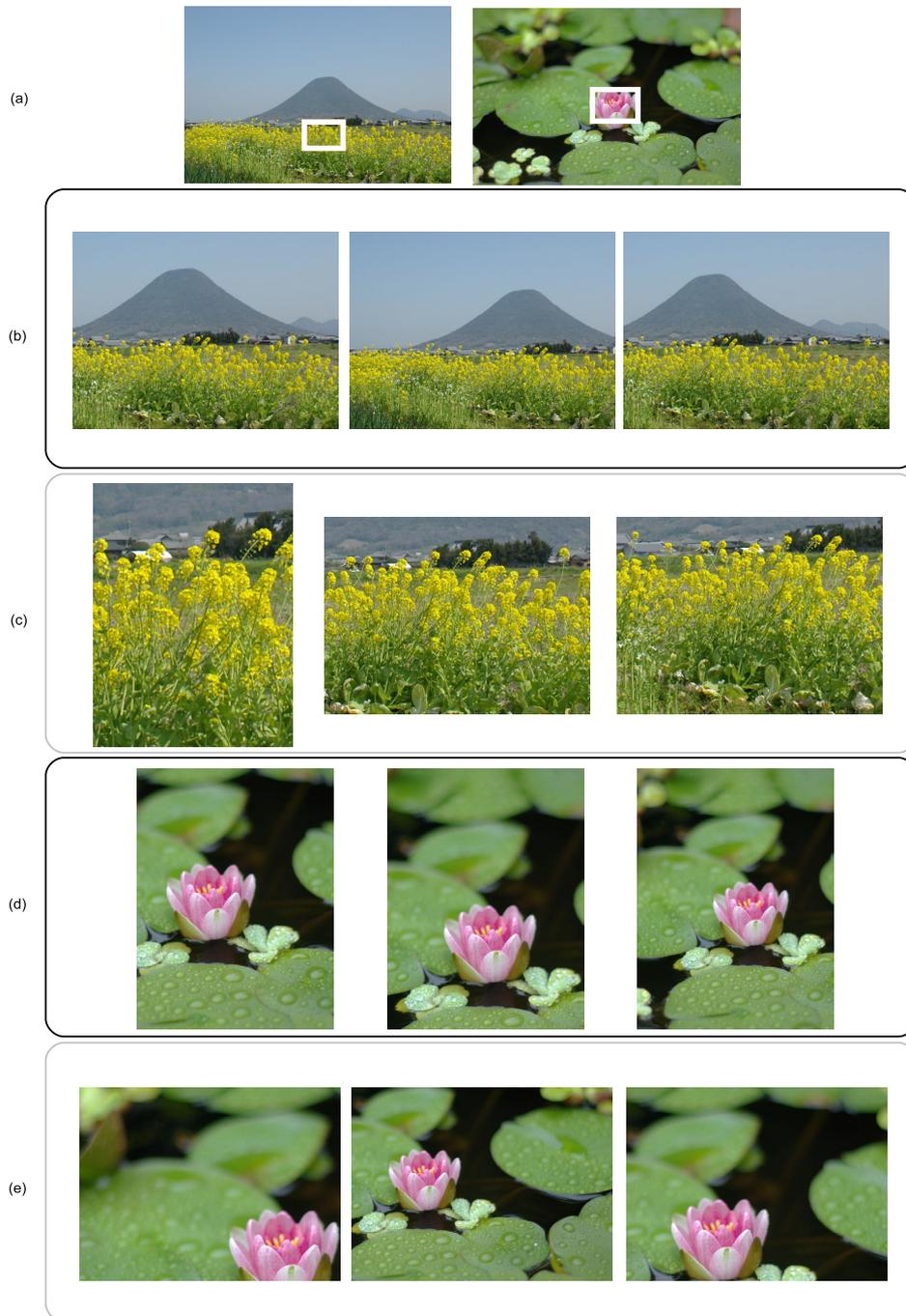


図 10 品質識別を用いた写真トリミングの応用例．広角の写真 (a) から高品質の部分画像を抽出する．白枠は部分画像に必ず含める指定領域である．(b),(d) が品質上位の部分画像，(c),(e) が品質下位の部分画像である．

Fig. 10 Examples of photo cropping using quality classification.

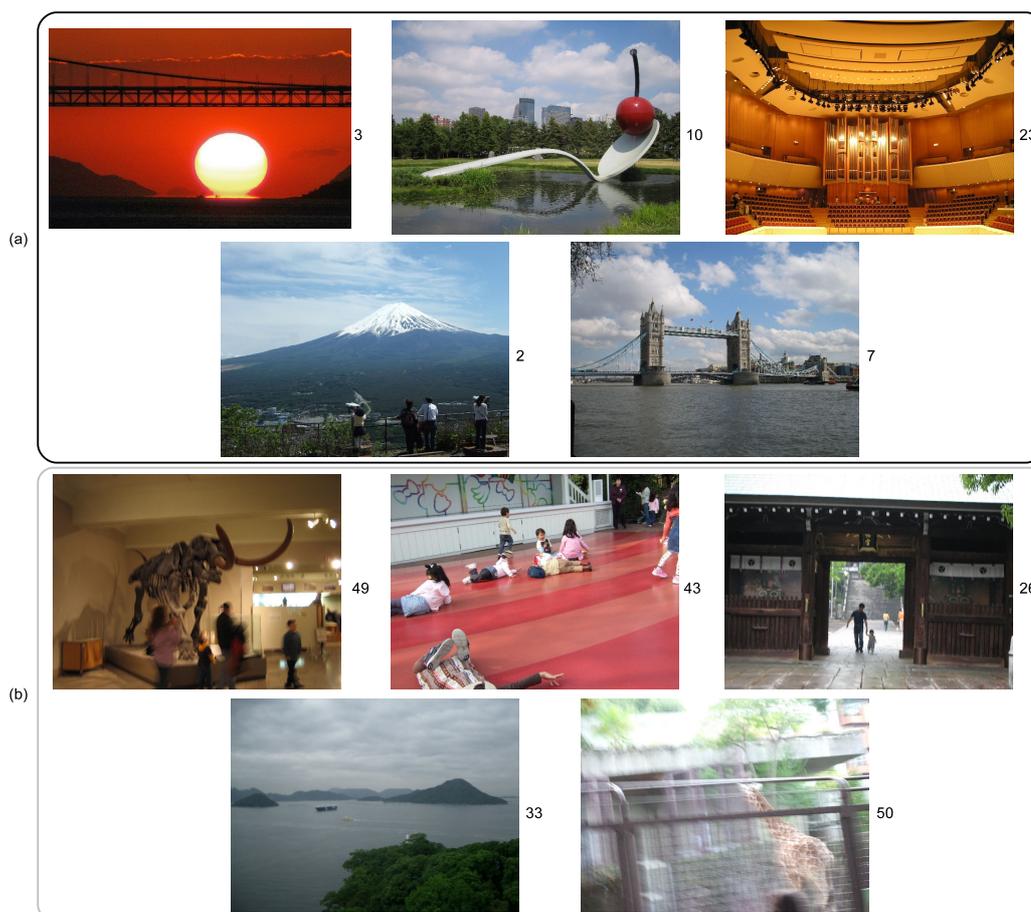


図 13 応用例の品質識別を用いた写真整理の結果。(a) は品質が上位の写真, (b) は品質が下位の写真である。各画像の右側に示す数値は人間が与えた主観スコアの順位である。

Fig. 13 Application of photo arrangement using quality classification.

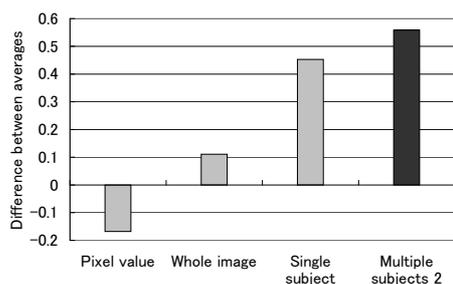


図 14 写真整理の結果に対する平均主観スコアの差分値。値が大きいほど主観評価の結果と近い。

Fig. 14 Difference of averages of subjective scores against photo arrangement.

を用いた主観評価でスコア付けを行った。各被験者に対し、2つの写真のペアを ${}_{50}C_2 = 1225$ 回提示し、より高い品質と感じた一方の写真を選択させた。被験者がどちらとも言えない写真のペアに対してはランダムに選択させた。この一対比較法により、50枚の写真の中で相対的な主観スコアが各写真に対して与えられる。一対比較法による主観スコアは、相対的な比較をすることなく各写真に対して直接与えられた4.の主観スコアと比べて、ぶれが少なく信頼度が高い。本実験は、4.とは異なる方法で与えられた主観スコアを評価に用いても、提案手法の結果が人間の主観と近くなることを確認するために行う。被験者は24名とした。一対比較による主観スコアの分布を図12に示す。図1は、この主観評価で上位と下位になった写真であ

る。全員が共通で高品質と感じた写真は高い主観スコアに、その逆は低い主観スコアとなる。

写真整理を適用した結果を図 13 に示す。品質識別には 4. で述べた提案手法 Multiple subjects 2 を用いた。主観的な品質が高いものが上位に選ばれていることが分かる。これらの写真整理の結果を定量的に評価するために、上記の対比較による人が与えた主観スコアを用いて、従来手法との比較を行った。指標として、上位/下位の写真に対する主観スコアの平均を求め、それら平均間の差分値を求める。差分値が大きいほど、主観評価の結果に近づく。図 14 に、各手法 (Pixel value, Whole image, Single subject, Multiple subjects 2) に対する平均主観スコアの差分値を示す。提案手法が、他の手法と比べて優れていることが確認できる。

6. おわりに

本論文では、顕著度を用いて複数の注目領域を検出し、複数の注目領域と背景領域との間の関係を考慮した品質識別を行う手法を提案した。提案手法は、従来の単一の注目領域に比べて、より詳細な特徴を抽出することができる。実験により、品質識別の性能が向上することを確認した。提案手法の品質識別を利用した写真提示の応用例として、写真トリミングと写真整理とを紹介した。

今後の課題として、識別性能の向上、注目領域の個数に対するさらなる評価考察、注目領域の位置関係の導入、主観スコアに対する回帰の適用などが挙げられる。

文 献

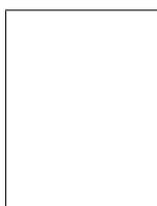
- [1] N. Snavely, S. M. Seitz, and R. Szeliski, Photo tourism: exploring photo collections in 3D, SIGGRAPH Conf. Proc., pp. 835 – 846, 2006.
- [2] H. Tong, M. Li, H. Zhang, J. He, and C. Zhang, Classification of Digital Photos Taken by Photographers or Home Users, Proc. Pacific-Rim Conference on Multimedia, 2004.
- [3] Y. Ke, X. Tang, and F. Jing, The Design of High-Level Features for Photo Quality Assessment, Proc. CVPR, vol. 1, pp. 419 – 426, 2006.
- [4] R. Datta, D. Joshi, J. Li, and J. Z. Wang, Studying Aesthetics in Photographic Images Using a Computational Approach, Proc. ECCV, part III, pp. 288 – 301, 2006.
- [5] Y. Luo and X. Tang, Photo and Video Quality Evaluation: Focusing on the Subject, Proc. ECCV, part III, pp. 386 – 399, 2008.

- [6] A. Loui, M. D. Wood, A. Scalise, and J. Birkelund, Multidimensional Image Value Assessment and Rating for Automated Albuming and Retrieval, Proc. ICIP, pp. 97 – 100, 2008.
- [7] R. Datta, J. Li, and J. Z. Wang, Algorithmic Inferring of Aesthetics and Emotion in Natural Images: an Exposition, Proc. ICIP, pp. 105 – 108, 2008.
- [8] DPChallenge, <http://www.dpchallenge.com>.
- [9] Photo.net, <http://photo.net>.
- [10] Photography Books, ISBN : 4817950951, 4056051232, 4522421257.
- [11] L. Itti, C. Koch, and E. Niebur, A Model of Saliency-Based Visual Attention for Rapid Scene Analysis, Trans. PAMI, vol. 20, no. 11, pp. 1254 – 1259, 1998.
- [12] R. O. Duda, P. E. Hart, and D. G. Stork. Pattern classification second edition. John Wiley & Sons, 2000.
- [13] D. Comaniciu and P. Meer, Mean shift: A robust approach toward feature space analysis, Trans. PAMI, vol. 24, no. 5, pp. 603 – 619, 2002.
- [14] V. Vapnik, The Nature of Statistical Learning Theory. New York, Springer-Verlag, 1995.
- [15] J. Platt, Probabilistic outputs for support vector machines and comparison to regularized likelihood methods, Advances in Large Margin Classifiers, MIT Press, pp. 61 – 74, 2000.
- [16] H. T. Lin, C. J. Lin, and R. C. Weng, A note on Platt's probabilistic outputs for support vector machines, Technical report, National Taiwan University, 2003.
- [17] J. Meynet, V. Popovici, M. Sorci, and J. P. Thiran, Combining SVMs for Face Class Modeling, Proc. 13th European Signal Processing Conference, 2005.
- [18] L. L. Thurstone, A law of comparative judgement, Psychological Review, Vol. 34, pp. 273 – 286, 1927.
- [19] G. Csurka, C. R. Dance, L. Fan, J. Willamowski, and C. Bray, Visual Categorization with Bags of Keypoints, Proc. Workshop on Statistical Learning in Computer Vision, 2004.

(平成 xx 年 xx 月 xx 日受付)

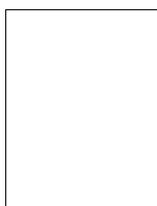
西山 正志 (学生員)

2002 岡山大学大学院博士前期課程了。同年 (株) 東芝入社。研究開発センターマルチメディアラボラトリー勤務。研究主務。2008 東京大学大学院学際情報府博士課程入学。コンピュータビジョンの研究に従事。2006 本会パターン認識・メディア理解研究会研究奨励賞受賞, 2007 画像センシングシンポジウム優秀論文賞, 2009 山下記念研究賞受賞。情報処理学会会員



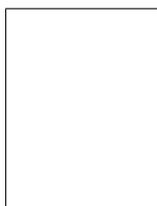
岡部 孝弘 (正員)

1997年東京大学理学部物理学科卒業。1999年同大学大学院理学系研究科物理学専攻修士課程修了。2000年同博士課程中退。2001年より東京大学生産技術研究所技官、助手を経て、現在同研究所助教。コンピュータビジョン、画像パターン認識、コンピュータグラフィックスに関する研究に従事。平成17年度電子情報通信学会論文賞、MIRU2004・MIRU2009優秀論文賞、平成16年度PRMU研究奨励賞、平成19年度情報処理学会山下記念研究賞などを受賞。情報処理学会、IEEE各会員。



佐藤 洋一 (正員)

平2東大・工・機械卒。平9カーネギーメロン大計算機科学部博士課程了。東大・生産技術研究所助教授を経て、同大学院情報学環准教授。コンピュータビジョンに関する研究に従事。平20及び平18本学会論文賞、平11情報処理学会山下記念研究賞、平11日本VR学会論文賞、ICPR2008 Best Industry Related Award等を受賞。



佐藤いまり

1994年慶應義塾大学総合政策学部卒業、2005年東京大学大学院学際情報学府博士課程修了。学際情報学博士。国立情報学研究所助手を経て、現在同研究所准教授、イメージベースモデリング&レンダリングに関する研究に従事。1992-1993 Carnegie Mellon University (CMU), Center for Machine Translation, Research Assistant, 1994-1996 CMU, The Robotics Institute, Visiting Scholar. 2005-2009 科学技術振興機構さきがけ研究員