

審美的品質識別のための局所領域の組合せによる色彩調和の評価

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

Assessment of Color Harmony using Combination of Local Regions for Classification of Aesthetic Quality

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

あらまし 本論文では、写真の審美的な品質識別を目的とし、その性能向上に寄与する色彩調和の評価手法について述べる。数色からなる単純な配色が対象であれば、従来の色彩調和モデルで人の感覚と近い評価を行うことができる。しかし、従来モデルは、様々な色が様々な形で同時に表れる写真の色彩調和をどのように評価するかについて考慮していなかった。この問題を解決するために、従来モデルで取り扱っていた単純配色と写真の配色との間に存在するギャップを埋める必要がある。そこで、写真の色彩調和はその写真中の局所領域の組合せから評価できると考え、その審美的品質が高いか低いかを識別する手法を設計する。審美的品質が与えられた大規模データベースを用いた実験を行い、提案手法により識別性能が向上することを確認した。

キーワード 色彩調和, 写真, 審美的品質, 局所領域, 相対的色分布

1. ま え が き

本研究では、写真の良し悪しという人間が写真に対して持つ審美的品質を取り扱う。ここで述べる審美的品質とは、ノイズや圧縮歪みなどの画像品質とは違い、人が写真を感覚的に好むかどうかである。審美的品質を識別するために、いくつかの手法 [1] ~ [7] が提案されている。これらの手法は、一般的に知られている写真撮影の知識を用いて品質を判定する。文献 [1] ~ [3], [5] では、審美的品質のラベルが与えられた訓練サンプルを用いて識別器を学習することで、高品質の写真と低品質の写真とを区別する。その識別性能を高めるためには訓練サンプルを十分に用意する必要があり、これらの文献はインターネット上に存在する大量写真を利用する。これにより、多数の人からコンセンサスの得られる審美的品質を統計的に分析する。この審美的品質を取り扱う上で写真の色彩は重要な要因である。色彩調和を評価するために、色彩工学の分野で、従

来より様々なモデル [8] ~ [11] が提案されている。ところがこれらの従来モデルは、図 1 左のような単純配色を対象としており、少数の色で構成される写真を取り扱うことはできるが、同図右のような一般的な写真の複雑な色彩をどう取り扱うか考慮していなかった。

さらに、従来より提案されている審美的品質識別の手法 [1] ~ [7] は、色彩調和についてほとんど議論がなされていなかった。例えば、文献 [1] ~ [3], [5] は、色分布と審美的品質との相関をみるために、識別する際の特徴量として RGB 値ヒストグラムや色相ヒストグラムを用いるのみであった。これらの単純ヒストグラム特徴量は各色の出現頻度をみるといった写真撮影の基本技術を表すためには有効であるが、人の感性である色彩調和を表すためには十分ではないと考えられる。

そこで、本論文では、審美的品質識別の性能向上を目的とし、写真の色彩調和を評価する手法について述べる。写真の局所領域に注目すると、それは単純配色とみなすことができるため、提案手法は、写真全体を単純配色の集合と考え、写真の色彩調和をその集合から評価する (詳細は 3.1 で議論)。この考え方にに基づき審美的品質の識別手法を設計することで認識性能が向上することを実験で確認した。

以下、2. で従来の色彩調和モデル、3. で色彩の調和度が高いか低いかを識別する手法について述べる。次

[†] 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@m.ieice.org



図 1 本論文の目的. 色彩調和に関して, 従来の単純配色と写真の配色との間に存在するギャップを埋める.

Fig. 1 Our goal is to bridge a gap between simple color patterns and color arrangements of photos for assessing color harmony.

に, 4. でその識別性能を評価し, 5. で色彩調和に加えてエッジ, ぼけ, 顕著度の特徴量を統合することで審美的品質識別の性能を高める. 最後に, 6. でまとめる.

2. 従来の色彩調和モデル

従来の色彩調和モデル [8] ~ [11] の中でも, 工学分野では, 計算機上での取り扱いが容易な Moon-Spencer モデル [8] や松田モデル [11] が用いられることが多い. ここで, Moon-Spencer モデルの概要を図 2 に示す. このモデルは, 2 色からなる単純配色を対象とし, 基準となる色 (Chosen color) からの相対的な色を評価に用いる. マンセル表色系における色相・輝度・彩度の相対値が図 2 のあいまいさ (Ambiguity) の領域に存在しなければ, その 2 色間で調和が保たれているとみなす. 調和が保たれている場合として, 基準色から大きく離れている色の ‘Contrast’, 近い色の ‘Similarity’, 同色の ‘Identity’ の 3 種類が用意されている.

上記の従来モデルは, 図 1 左のような単色カラーパッチの組合せ配色や洋服の単純な配色を刺激とする心理実験に基づいている. そのため, 様々な色が様々な形で表れる自然画像に対しては, 必ずしも適切な色彩調和モデルとはいえない. 例えば, 文献 [12] では, 手入力した画像中の大まかな領域から色相ヒストグラムを求め, 松田モデルを適用することで, 色彩調和を保った色変換を行っている. また, 文献 [13] は, 動画の審美的品質を識別するために, 各フレームの画像全体から求めた色相ヒストグラムと松田モデルとの適合度から色彩調和を評価している.

我々の実験で, これらの文献を参考にして, Moon-Spencer モデルや松田モデルを画像全体から求めた色分布に適用したところ, その結果 (図 9 を参照) が示すように, 従来モデルは写真の審美的識別の性能向上に効果がなかった. よって, 従来モデルが対象として

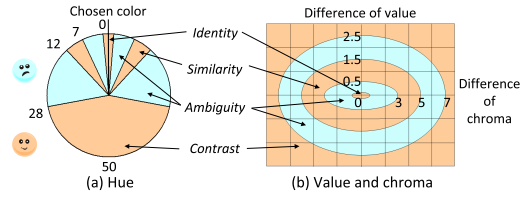


図 2 色彩調和を評価するための Moon-Spencer モデルの概要. 基準となる色 (Chosen color) からの相対的な値を評価に用いる. マンセル表色系における色相・輝度・彩度の相対値が, 図中の水色の領域で示されているあいまいさ (Ambiguity) の領域に存在しなければ, 2 色間の調和が保たれているとみなす.

Fig. 2 The model of color harmony proposed by Moon and Spencer. The model uses relative values from a chosen color, and considers two colors in harmony when not containing the region of ambiguity.

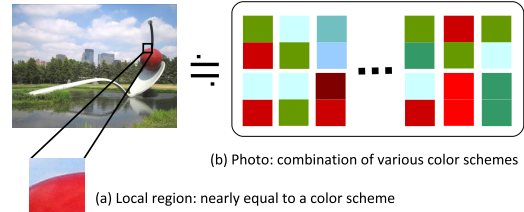


図 3 写真の色彩調和. 図中 (a) のように写真中の局所領域に注目すると, その領域は単純配色と見なすことができるので, 図中 (b) のように写真全体を単純配色の集合と仮定する. 提案手法はこれを利用し, 局所領域の組合せから色彩調和を評価する.

Fig. 3 We take color harmony for photos into consideration. A local region, e.g. shown in (a), is nearly equal to a simple color pattern. We assume that a photo consists of the set of simple color patterns such as Fig.(b). Our method evaluates color harmony of a photo using the combination of local regions.

いた単純配色だけではなく, 写真の複雑な配色も取り扱えるような色彩調和の評価手法を新たに構築する必要がある.

3. 写真の色彩調和の評価アルゴリズム

3.1 考え方

上記で述べたように, 写真の色彩調和の評価は非常に難しい問題である. このためにまず, 従来の色彩調和モデルで使用する単純配色と, 提案手法が取り扱う写真の配色との違いを考える. 色分布の複雑さをみると, 図 1 で示したように, 両者には大きな隔たりが存在する.

ここで, 写真の色彩調和の問題に取り組む手始めと

して、画像の局所領域に注目する。局所領域内の配色は、図 3(a) のように、比較的単純であると言える。よって、局所領域は単純配色に近いとみなすことができ、その色彩調和を従来モデルにより評価できると考えられる。さらに、一枚の画像が単純配色の組合せであるといえることから、写真全体の色彩評価は局所領域の調和度から決めることができると考えられる。実際に、4. で述べる予備実験において、Moon-Spencer モデルで局所領域の調和度を求め画像全体で加算することで写真の審美的品質の性能が向上することを確認している。さらに予備実験で、審美的品質が高い(低い)写真は調和度の高い(低い)局所領域の出現頻度が高いことが分かっている。

ここまでの考察に基づき、本論文において写真の色彩調和の評価手法をさらに発展させる。提案手法は、図 3(b) のように、一枚の写真を大量の単純配色の集合と仮定し、図 4 のように、その単純配色の出現頻度から色彩調和の評価を行う。このために写真から局所領域をサンプリングし、色彩調和モデルの知見に従い局所領域から局所記述子を抽出する。それら抽出された局所記述子の集合から画像全体を表す特徴量を算出し、識別器に入力することで写真の色彩の調和度が高いか低いかを識別する。

提案手法は、一般物体認識や画像検索で有効な手法として知られている Bags-of-features [14], [15] のアプローチと類似点が多い。このアプローチは一枚の画像を局所的な特徴量の集合として捉えることで、物体姿勢や照明条件の変動を抑えて識別することを狙っている。本論文では、一般物体認識のための SIFT に相当する局所記述子を色彩調和モデルの知見から設計する。

3.2 評価アルゴリズムの設計

提案手法は以下の三つのステップを処理する。

- 局所領域のサンプリング
- 局所記述子の生成
- 画像全体を表す特徴量の表現

画像全体を表す特徴量が生成された後は Support Vector Machine (SVM) で識別を行う。SVM の学習には、審美的品質が高いか低いかの 2 値のラベルが付いた訓練サンプルを用いる。以下、各方法について詳しく述べる。

3.2.1 局所領域のサンプリング

まず、与えられた画像から局所領域の集合を抽出する。一般物体認識の手法 [14], [15] では、エッジやコーナーなど画像の低レベル特徴を用いた Difference of

Gaussian や Harris-Laplace などの演算子で局所領域を決定する。しかし、色彩調和を評価するためには、画像中のエッジやコーナー付近の色彩だけでなく平坦な領域の色彩も必要であると考えられる。これは、図 2 の従来モデルで、色の対比を表す 'Contrast' や 'Similarity' だけでなく同色の 'Identity' もまた色彩調和の評価に用いることから類推される。

よって、本論文では、領域の大きさを固定とした格子状のサンプリングで局所領域を抽出する。その際に決めるべきパラメータとして大きさだけでなく、一枚の画像からどれだけの数の局所領域を抽出するかのサンプリング密度も重要である。これらのパラメータについては 4.4 の実験で考察する。

3.2.2 局所記述子の生成

次に、局所領域から局所記述子を生成する方法について述べる。従来の審美的品質の識別手法や一般物体認識の手法では、色を表すための局所記述子として、RGB 色空間でのヒストグラムや CIE LCH 色空間での色相ヒストグラムを用いることが多かった。RGB 値ヒストグラムは色分布の類似性を評価することはできるが、人間の主観である色彩調和の評価に適しているとはいえない。4. の予備実験で示すように、RGB 値ヒストグラムでは識別性能を得ることができなかった。一方の LCH 色空間は人間の色知覚をモデル化したものであるが、この色相ヒストグラムでも識別性能を得ることができなかった。

本論文では、従来の色彩調和モデルを参考にすることで、人の感覚と同じように写真の色彩調和を評価する記述子を設計する。2. で述べたように、従来の色彩調和モデルは色の相対的な値を評価に用いる。この知見に基づき本論文では、マンセル表色系における色相・輝度・彩度の相対値の出現頻度を局所記述子とし、色彩調和の評価を行う。従来の審美的品質識別の手法は色相ヒストグラムのような色の絶対値の分布を用いていたが、提案手法は色の相対値の分布を用いる。

ここで、局所記述子を生成する方法について詳細を述べる。まず、局所領域が与えられると RGB 色空間からマンセル表色系へ変換する^(注1)。次に、局所領域内における色相の最頻値を基準色と定め、各画素から

(注1): Rochester Institute of Technology の R. S. Berns らが公開している表 (<http://www.cis.rit.edu/mcsl/online/munsell.php>) を利用した。この表では、CIE xyY 色空間での離散的な値に対するマンセル表色系での値が示されている。本論文では、表に存在しない値を、その近傍からの線形補間で求めた。

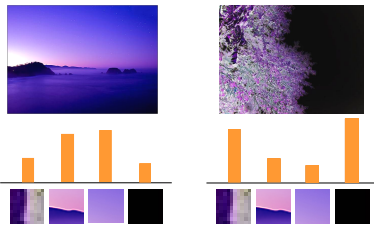


図 4 色彩調和の違いを表すヒストグラム．画像を局所的な特徴量の集合として表現する Bags-of-features のアプローチを写真の色彩調和の評価に用いる．単純配色と見なせる代表的な局所領域の出現頻度の違いから色彩の違いを見分ける．

Fig. 4 Histograms representing the difference of color harmony. We use bags-of-features that represents a photo as a set of local regions to assess color harmony. The method classifies the quality of color harmony using the histogram of representative local regions which are nearly equal to simple color pattern.

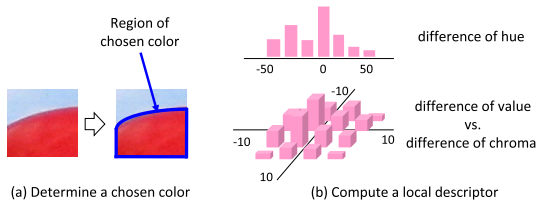


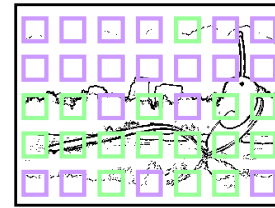
図 5 局所記述子の生成．図中の (a) のように、局所領域に対して基準色を設定する．次に (b) のように、基準色からの相対的な色分布の頻度から局所記述子を算出する．

Fig. 5 We set a chosen color from a local region shown in Fig.(a) to compute a local descriptor. Next, we count two histograms of (b) from distribution of relative color using the chosen color.

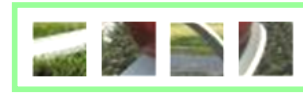
相対値を算出する．選択された基準色の例を図 5(a) に示す．局所領域内の各画素の色相と基準色の色相との差を求め、図 5(b) 上のような、色相差の値をピンとするヒストグラムを生成する．この色相差ヒストグラムは、図 2(a) で示した従来の色彩調和モデルで利用している色相差をもとに設計されている．次に、基準色となった画素から彩度と輝度の平均を求める．それらの平均と各画素の彩度と輝度から求めた差分値から、図 5(b) 下のような、2次元ヒストグラムを生成する．このヒストグラムは図 2(b) で示した従来モデルで利用している輝度差と彩度差をもとに設計されている．最後に、これらの 2 つのヒストグラムを連結することで局所記述子を生成する．

3. 2. 3 画像全体を表す特徴量の表現

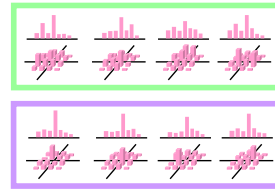
局所記述子の集合を用いて画像全体の色彩を表す特



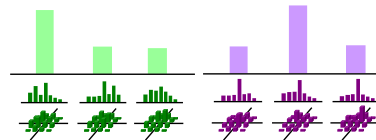
(a) Color boundary



(b) Local regions



(c) Local descriptors



(d) Histograms

図 6 画像全体を表す特徴量の抽出．局所領域へ色境界、または、境界以外のラベル付けを行う．図中 (a) の枠の色がラベルの違いを表す．次に、(b) の各ラベルの局所領域の集合から (c) の局所記述子を求めた後、それぞれの集合から (d) のヒストグラムを生成する．この二つのヒストグラムを連結することで特徴量を決定する．

Fig. 6 To extract a feature representing a whole photo, we label local regions whether it contains color boundary or not. The pixel value in Fig. (a) represents the color boundary. We make two sets of local regions using the labels like (b) and compute the sets of local descriptors of (c) and two histograms of (d) from them. By concatenating the histograms, we compute the feature.

徴量を設計する．本論文では、局所領域が様々な色で構成されるか、または、単色に近いかを考えながら局所記述子の集合を統合する．これは、従来モデルが色空間内で近い色の ‘Similarity’ や離れた色の ‘Contrast’ といった複数色を評価することと、同色の ‘Identity’ といった単色もまた評価することを、特徴量の設計へ

反映させるためである。

画像全体を表す特徴量の算出方法について詳しく述べる。まず、3.2.1 でサンプリングした局所領域の集合から様々な色で構成されている局所領域を検出する。その際に、微小な色変化の影響を除去し色境界を決定するため、Mean shift segmentation [16] を原画像へ適用する。隣接する全ての領域間で RGB 色空間における距離差を求めた後に、大津の二値化 [17] を用いて色境界を決定する。局所領域に含まれている色境界の割合から、図 6(a) のように、色境界が境界以外かのラベルを局所領域へ与える。推定されたそれらのラベルを用いて、図 6(b) のように、色境界の局所領域の集合と境界以外の局所領域の集合に分ける。

次に、図 6(c) のように、局所領域から局所記述子を求め、それらの局所記述子の集合を統合する。このために、代表的な局所記述子である visual words で構成されるコードブックを用いる。このコードブックの生成方法は後述する。上記で求めた色境界と境界以外の局所記述子から、コードブックの visual words と類似する局所記述子の出現頻度を数えることで、図 6(d) のように二つのヒストグラムを求める。

これらのヒストグラムは、色彩について色分布を評価することはできるが、画像上での局所領域間の位置関係を含んでいないため色配置を評価することはできない。そこで、本論文では、大まかな色配置を特徴量に反映させるために、写真をブロックに分割して、ブロック毎の局所領域からヒストグラム特徴量を算出する。これら各ブロックから算出された色境界と境界以外の全てのヒストグラムを連結することで、画像全体を表す特徴量を決定する。実験では分割する際、ブロックの大きさは均等とし、ブロック間で重なりがないようにした。

最後に、コードブックを事前に生成する手順について述べる。まず、高品質と低品質の訓練サンプルから色境界と境界以外の局所記述子を抽出する。それらの局所記述子の集合を k-Means クラスタリングで量子化し、各クラスタの平均ベクトルをコードブックの visual words とする。この手続きで、色境界と境界以外から 2 個のコードブックを生成する。

4. 色彩調和についての識別性能

提案手法の識別性能を確認するために評価実験を行った。まず、4.1 で評価に用いるデータベースの詳細について述べ、4.2 で色彩調和の評価方法の基本的

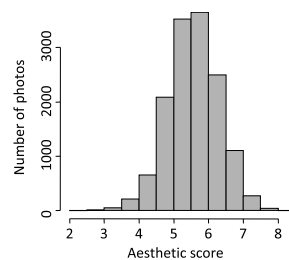


図 7 カテゴリ 'Landscape' において、各写真に与えられた審美的スコアのヒストグラム。

Fig. 7 Histogram of aesthetic score given to each photo in the category 'Landscape'.

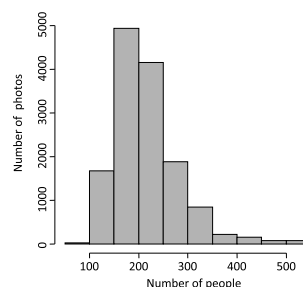


図 8 カテゴリ 'Landscape' において、各写真に対して審美的スコアを与えた人数のヒストグラム。

Fig. 8 Histogram of people who give aesthetic scores for photos in the category 'Landscape'.

能について述べ、4.3 で Visual words の評価について述べる。最後に、4.4 で局所領域について考察する。

4.1 評価に用いるデータベース

写真共有サイト DPChallenge [18] において審美的品質がスコア付けされた写真データベースを用いて実験を行った。14 種類のカテゴリ (Abstract, Animals, Cityscape, Floral, Landscape, Macro, Nature, Portraiture, Rural, Seascapes, Sky, Stilllife, Urban, Water) から、白黒やセピア調の写真を除き合計 124,664 枚のカラー写真を収集した。一例として、Landscape における主観スコアのヒストグラムを図 7 に、審美的スコアを与えた人数のヒストグラムを図 8 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。DPChallenge の写真は、一枚あたり約 200 人から審美的スコアが与えられている。審美的スコア分布の上位 10 % の写真を高品質、下位 10 % の写真を低品質のラベルとし、品質識別器の学習とテストに用いた。

提案手法の有効性を確認するために、色彩調和についてのみスコア付けされたデータベースを用いるこ

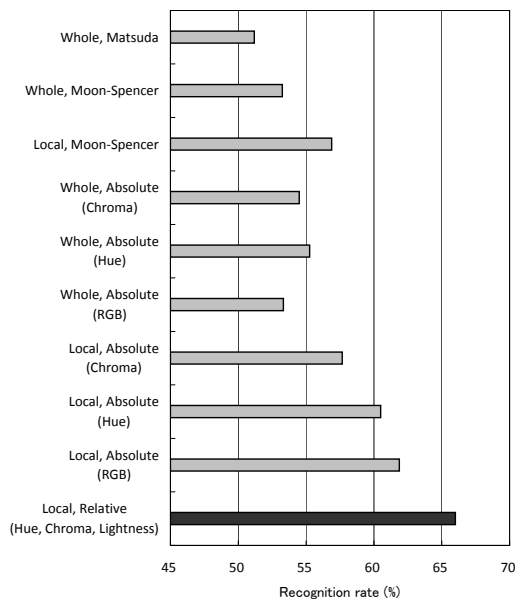


図 9 写真の色彩調和についての審美的品質識別の結果。
Fig. 9 Performance of quality classification of color harmony of photos.

とが望ましい。ところが、そのようなものは存在しないため、審美的品質についてスコア付けされている DPChallenge のデータベースを実験に用いた。審美的スコアは、色彩調和だけでなく、構図、注目領域、ぼけ具合、コントラストなど様々な観点から評価された総合的な値である。そのため、スコア分布上位の写真には色彩の調和度が低いものは含まれず、下位の写真には色彩の調和度が高いものは含まれないと考えられる。例えば、調和度は高いが構図の品質は低い写真は、スコア分布の中間に位置すると考えられる。このような写真は本実験では用いないことにした。よって、本論文では、DPChallenge の上位と下位の写真で色彩調和の評価を行う。

4.2 基本性能

提案手法による色彩調和の評価性能を確認するために、従来モデル [8], [11] との比較、画像全体から求めた特徴量 'Global' と局所領域の集合から求めた特徴量 'Local' の比較、色相ヒストグラムなど色の絶対値を用いる特徴量 'Absolute' と色の相対値を用いる特徴量 'Relative' の比較の実験を行った。特徴量を入力する識別器は線形カーネルの SVM とし、1 枚の画像から大きさ 32×32 の局所領域を $48 \times 48 = 2304$ 個ほど抽出した。領域の大きさとサンプリング個数について

は 4.4 で考察する。以下、各手法について述べる。

Whole, Matsuda: 松田モデル [11] による特徴量である。このモデルは、色彩調和の評価が高い代表的な色相ヒストグラムとして、9 つの調和テンプレート (Harmonic templates) を用意している。このテンプレートは、2. で述べた Moon-Spencer モデルと同様に、色相の相対的な出現範囲を定義している。実験では画像全体から CIE LCH 色空間における色相ヒストグラムを求め、文献 [12], [13] のように調和テンプレートへの類似度を調和度として算出し、その値を特徴量とした。

Whole, Moon-Spencer: 画像全体から算出したヒストグラムへ Moon-Spencer モデル [8] を適用し獲得した調和度を特徴量とした。調和度は、色相差・輝度差・彩度差ヒストグラムにおいて、図 2 の 'Identity' と 'Similarity' と 'Contrast' との領域に対応するピンの出現頻度の合計値から算出した。

Local, Moon-Spencer: 写真からサンプリングした個々の局所領域のヒストグラムへ Moon-Spencer モデル [8] を適用し調和度を求め、それら局所領域の調和度を加算した値を特徴量とした。調和度は 'Whole, Moon-Spencer' で述べた方法で算出した。

Whole, Absolute (Chroma): CIE LCH 色空間における彩度の絶対値ヒストグラムを画像全体から算出した特徴量である。

Whole, Absolute (Hue): 同様に、色相の絶対値ヒストグラムが特徴量である。

Whole, Absolute (RGB): RGB 色空間における絶対値ヒストグラムを画像全体から算出した特徴量である。

Local, Absolute (Chroma): 局所領域から局所記述子として彩度のヒストグラムを算出し、3.2.3 で述べた方法で統合した特徴量である。

Local, Absolute (Hue): 同様に、局所領域から色相のヒストグラムを算出し統合した特徴量である。

Local, Absolute (RGB): 同様に、局所領域から RGB 値のヒストグラムを算出し統合した特徴量である。

Local, Relative (Hue, Chroma, Lightness): 色相・輝度・彩度の相対値ヒストグラムを局所記述子として用いる提案手法の特徴量である。局所記述子の次元数は 200 とし写真の色彩調和を評価した。

これらの特徴量を用いることで 14 種類のカテゴリから求めた平均識別率を図 9 に示す。従来の色彩調

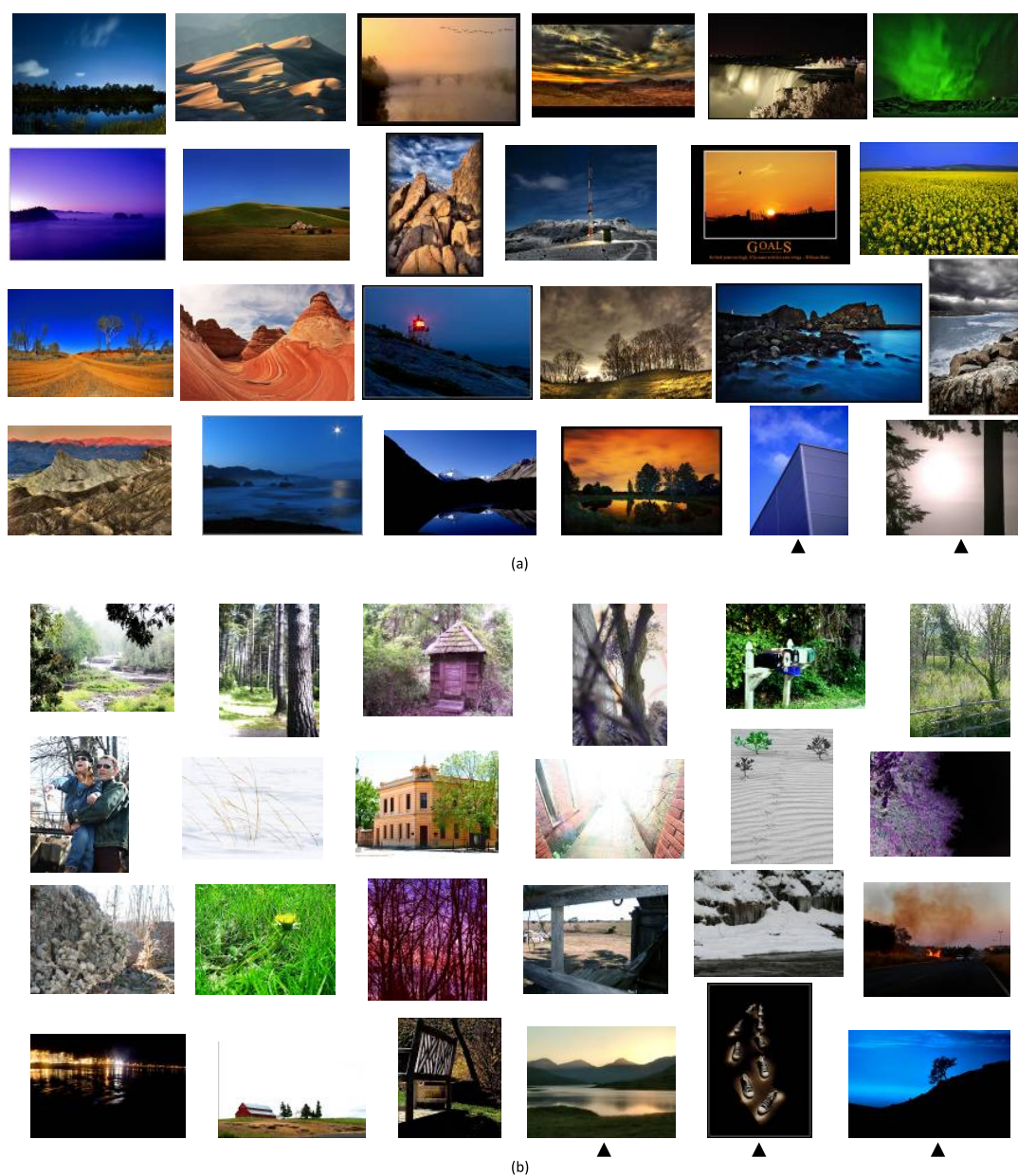


図 10 色彩の調和度が (a) 高い・(b) 低いと識別された写真の例．誤識別された写真の下へ三角印を記している．

Fig. 10 Examples of photos which are classified to high quality in (a) and low quality in (b) by assessing color harmony. We put triangular signs below false classified photos.

和モデルや画像全体から求めた特徴量 ‘Global’ と比べて、局所領域を用いる手法 ‘Local’ は識別性能が高かった。また、色の絶対値を用いる手法 ‘Absolute’ と比べて、色の相対値を用いる手法 ‘Relative’ は性能が

高かった。この結果より、3.1 で述べた考え方にに基づき設計した提案手法が、写真の審美的な色彩調和の識別性能を高めることが分かった。

次に、図 10 に、色彩の調和度が (a) 高い、もしくは

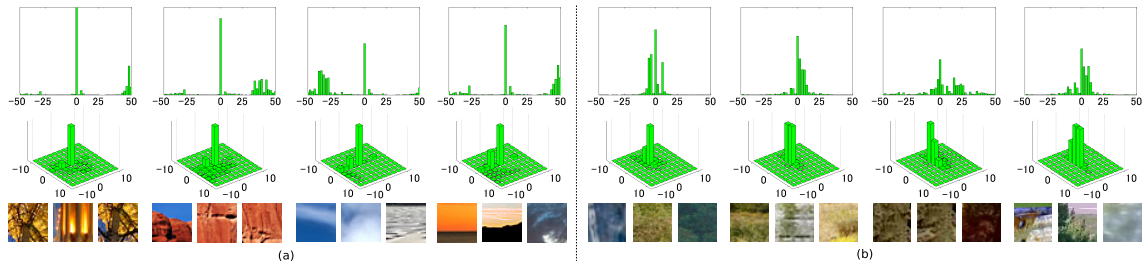


図 11 色彩の調和度が (a) 高い, または, (b) 低い Visual words の例. 各 Visual word について, 色相差ヒストグラムを上段に, 輝度差・彩度差の 2 次元ヒストグラムを中段に, 対応した局所領域を下段に示す.

Fig. 11 Examples of visual words of high harmony (a) and low harmony (b). We show the histogram of hue difference on the top row, the 2D histogram of hue and differences on the middle row, and the local regions correspond to individual visual words on the bottom row.

は, (b) 低いと判定された写真の例を示す. 図中の (b) には色の違いが分かり難い写真が多く見られるが, (a) には色が鮮やかな写真が多いといえる.

4.3 Visual words の評価

提案手法は Visual words の出現頻度を特徴量として用いるが, その有効性を色彩調和モデルに基づき評価する. そのために, 個々の Visual word について色彩の調和度を Moon-Spencer モデルを用いて算出した. 調和度の算出に 4.2 の ‘Whole, Moon-Spencer’ で述べた方法を用いた. 調和度が高い 100 個の Visual words と調和度が低い 100 個の Visual words とに対応した局所領域の出現頻度を比較する. 評価に用いる写真のカテゴリは ‘landscape’ とした.

調和度が高い上位 100 個の Visual words に対応した局所領域の出現頻度を, 審美的品質のラベルが付いたテストサンプルから算出した. 高品質の写真からは 6,055 個の局所領域が, 低品質の写真からは 4,848 個の局所領域が対応した. 調和度が低い下位 100 個の Visual words についても同様に出現頻度を算出した結果, 高品質の写真からは 19,690 個の局所領域が, 低品質の写真からは 24,448 個の局所領域が対応した. これらの結果は, Visual words の出現頻度は審美的品質を識別する能力があることを示している.

次に, 図 11(a) に調和度が高い Visual words とそれに対応した局所領域の例を, (b) に調和度が低い Visual words と局所領域の例を示す. これらの局所領域を比較すると, (b) より (a) の局所領域の方が, 色彩の調和が保たれていると感ずることができる. この結果は, 基準色からの相対値分布を評価する Visual

words の一部が人の主観を反映していることを示唆している.

4.4 特徴量を生成する際のパラメータ設定についての考察

まず, 局所領域を写真からサンプリングする際のパラメータ設定について実験する. 提案手法は, 3.2.1 で述べたように, 格子状のサンプリングで局所領域を抽出する. 実験では, 局所領域同士が互いに部分的に重なることを許し, 格子状サンプリングを行った. そのパラメータとして局所領域の大きさと, 一枚の画像からどれだけの数の局所領域を抽出するかのサンプリング密度がある. 図 12 にサンプリング密度は 48×48 とし局所領域の大きさを変更した場合の識別性能を, 図 13 に局所領域の大きさを 16×16 としサンプリング密度を変更した場合の識別性能を示す. 局所領域の大きさは, 小さすぎても大きすぎても性能が悪化することが分かった. 一方, サンプリング密度は, 増加するほど性能は向上するが, その割合は小さくなっていくことが分かった. なお, それぞれの実験で固定した片方のパラメータを変更しても同様の傾向がみられた. 最適な局所領域の設定方法については今後の課題とする.

次に, 局所記述子の次元数について実験する. この次元数は, 色相差ヒストグラムのビン数と輝度差・彩度差ヒストグラムのビン数とを合計したものである. 例えば 200 次元の場合, 色相差ヒストグラムのビン数が 100 で輝度差・彩度差ヒストグラムのビン数が 10×10 である. 図 14 に局所記述子の次元数を変更した場合の識別性能を示す. 実験結果より, この次元数

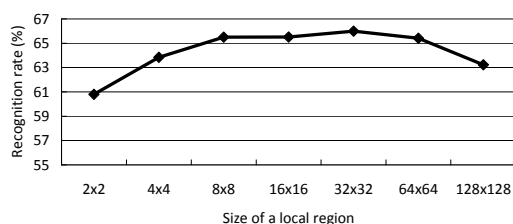


図 12 局所領域の大きさを増やした場合の色彩調和の評価性能。

Fig. 12 Performance of assessing color harmony in the case that the size of local regions is increased.

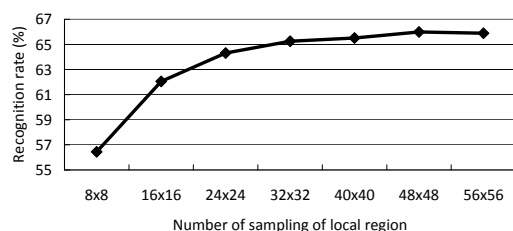


図 13 局所領域のサンプリング密度を増やした場合の色彩調和の評価性能。

Fig. 13 Performance of assessing color harmony in the case that the density of sampling local regions is increased.

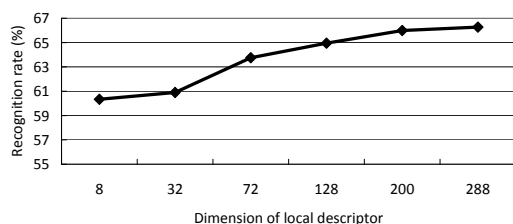


図 14 局所記述子の次元数を増やした場合の色彩調和の評価性能。

Fig. 14 Performance of assessing the color harmony when the dimension of local descriptors is increased.

が増加するほど性能は向上することが分かった。本実験ではヒストグラム生成の際に、最も近い値をもつピンへ単純に投票を行っていた。さらに性能を改善するために、重み付け投票を行い近い値をもつ周辺のピンも考慮することが考えられる。

最後に、画像全体を表す特徴量を生成する際の画像分割数について実験する。提案手法は一枚の写真をブロックへ等分割しており、ここではそのブロック数について評価した。図 15 に分割数を変更した場合の識別性能を示す。分割することで性能改善は見られるも

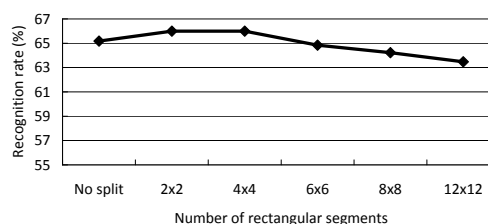


図 15 分割数を増やした場合の色彩調和の評価性能。

Fig. 15 Performance of assessing the color harmony when the number of rectangular segments divided from a photo is varied.

の、分割数が大きすぎると性能が悪化することが分かった。色配置を定量的に評価できる手法の開発は必要であり、色配置の良し悪しの定義と併せて今後の課題とする。

5. 様々な特徴量との統合

5.1 統合アルゴリズム

DPCChallenge でスコア付けされた写真の審美的品質は総合的な値であり、色彩調和の評価だけでは決まらない。識別性能をさらに高めるためには、審美的品質を決める他の要因も評価する必要がある。

ここでは、3.2.2 で述べた色彩調和の局所記述子に加えて、エッジ、ぼけ、顕著度の局所記述子を審美的品質識別に用いる。局所記述子を求めるために、エッジはその方向ヒストグラムである DAISY [23]、ぼけは周波数空間の振幅値 [22]、顕著度は Graph-Based Visual Saliency [24] を用いる。これら 3 つの局所記述子についてそれぞれ Bags-of-features のアプローチで画像全体を表す特徴量を生成する。ただし、3.2.3 で述べた色境界と境界以外のラベル付けは、エッジ、ぼけ、顕著度の特徴量の生成には用いない。

最終的な品質を決定するために、各特徴量の高品質らしさを表す事後確率を求め、それらの積をとる。ここでは、色彩調和、エッジ、ぼけ、顕著度の特徴量は独立であると仮定している。事後確率は、それぞれの特徴量に対し SVM を適用し、文献 [19] の方法でその出力を変換することで算出される。

5.2 統合した場合の識別性能

前節で述べた統合手法、従来の審美的品質識別の手法 [3], [5]、ノイズや圧縮歪みなどの画像品質を測る structural similarity (SSIM) [20] や colorfulness and SSIM [21] を用いる手法の比較実験を行った。以下、それぞれの手法について述べる。

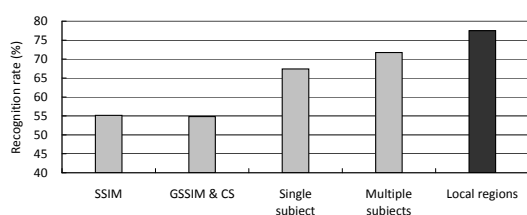


図 16 様々な特徴量の組み合わせによる審美的品質の識別性能の比較。

Fig. 16 Comparison of classification of aesthetic quality in the case that the features of color harmony, edge, blur, and saliency are combined.

SSIM: 2枚の画像間の類似度を SSIM [20] とし、最近傍識別を用いて品質を決定した。ここでは、与えられたテスト画像と全ての訓練画像を比較し、最も類似度の高い訓練画像の品質を返した。SSIM を算出する際に、グレー画像の画素値を用い、等分割された画像のブロックを用いた。

GSSIM & CS: 文献 [21] で提案された colorfulness and SSIM を類似度とし最近傍識別を用いて品質を決定した。この類似度を算出するために、グレー画像のエッジ強度から求めた GSSIM と、カラー画像から求めた colorfulness similarity を用いた。

Single subject: Luo ら [3] のように、写真中の単一注目領域と背景領域から求めた特徴量に対して、SVM を用いて品質を決定した。この実験では、文献 [3] の注目領域の検出方法を用いたが、特徴量は文献 [5] のように各領域のエッジ、色、ぼけを用いた。

Multiple subjects: 文献 [5] で述べられているように、複数の注目領域と背景領域から求めた特徴量に対して、SVM を用いて品質を決定した。特徴量は各領域のエッジ、色、ぼけを用いた。

Local regions: 5.1 で述べたように、局所領域から抽出した色彩調和、エッジ、ぼけ、顕著度の4つの特徴量を用いる統合手法である。

上記の各手法について、図 16 で 14 カテゴリの平均識別率を示す。本節で述べた統合手法は、従来の画像品質を測る手法に比べて、識別性能が大きく改善していることが確認された。また、注目領域から特徴量を算出する従来の審美的品質識別の手法よりも優れていることも確認された。以上の結果より、局所領域毎で特徴量を抽出し、その識別結果を統合する手法は、写真の審美的品質識別に有効であると言える。

6. おわりに

本論文では、審美的品質識別の性能向上のために、局所領域を単純配色とみなし、その組合せから写真の色彩調和を評価する手法を提案した。写真共有サイト DPChallenge 上の大規模データベースを用いた実験において、提案手法により識別性能が向上することを確認した。実験結果で、画像全体ではなく画像中の局所領域の組合せで色彩調和を評価することの有効性を示した。また、単純な色ヒストグラムではなく色の相対的な値の分布で色彩調和を評価することの有効性を示した。さらに、色彩調和に加えて、エッジ、ぼけ、顕著度も用いることで審美的品質識別の性能が高まることを実験で確認した。

今後の課題として、適切な局所領域の設定方法の探求、Color Transfer やホワイトバランス調整などへの応用が挙げられる。さらに、文献 [25] のようにトーンモデルを利用することで、トーンの統一度をはかることが挙げられる。

文 献

- [1] 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.
- [2] 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.
- [3] Y. Luo and X. Tang, Photo and Video Quality Evaluation: Focusing on the Subject, Proc. ECCV, part III, pp. 386 – 399, 2008.
- [4] 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.
- [5] M. Nishiyama, T. Okabe, Y. Sato, and I. Sato, Sensation-based Photo Cropping, Proc. ACM International Conference on Multimedia, pp. 669 – 672, 2009.
- [6] L. Liu, R. Chen, L. Wolf, and D. Cohen-Or, Optimizing Photo Composition, Computer Graphic Forum, vol. 29, no. 2, pp. 469 – 478, 2010.
- [7] S. Bhattacharya, R. Sukthankar, and M. Shah, A Framework for Photo-Quality Assessment and Enhancement based on Visual Aesthetics, Proc. ACM International Conference on Multimedia, pp. 271 – 280, 2010.
- [8] P. Moon, and D. E. Spencer, Geometric Formulation of Classical Color Harmony, Journal of the Optical Society of America, vol. 34, pp. 46 – 50, 1944.
- [9] J. Itten, The Art of Color, Van Nostrand Reinhold

- Company, 1960.
- [10] D. B. Judd, and G. Wyszecki, Color in Business, Science, and Industry, John Wiley & Sons, 1975.
- [11] 松田 豊, 色彩のデザイン, 朝倉書店, 1995.
- [12] D. Cohen-Or, O. Sorkine, R. Gal, T. Leyvand, and Y. Q. Xu, Color Harmonization, ACM Transactions on Graphics, vol. 25, no. 3, pp. 624 – 630, 2006.
- [13] A. K. Moorthy, P. Obrador, and N. Oliver, Towards Computational Models of the Visual Aesthetic Appeal of Consumer Videos, Proc. ECCV, part V, vol. 6315, pp. 1 – 14, 2010.
- [14] C. Dance, J. Willamowski, L. Fan, C. Bray, and G. Csurka, Visual Categorization with Bags of Keypoints, Proc. ECCV International Workshop on Statistical Learning in Computer Vision, pp. 1 – 22, 2004.
- [15] E. Nowak, F. Jurie, and B. Triggs, Sampling Strategies for Bag-of-Features Image Classification, Proc. ECCV, pp. 490 – 503, 2006.
- [16] D. Comaniciu and P. Meer, Mean shift: A robust approach toward feature space analysis, Trans. PAMI, vol. 24, no. 5, pp. 603 – 619, 2002.
- [17] N. Otsu, A Threshold Selection Method from Gray-Level Histograms, IEEE Trans. Systems, Man and Cybernetics, vol. 9, no. 1, pp. 62 – 66, 1979.
- [18] DPChallenge, <http://www.dpchallenge.com>.
- [19] 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.
- [20] Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600 – 612, 2004.
- [21] Y. Xiang, B. Zou, and H. Li, Selective color transfer with multi-source images, Pattern Recognition Letters, Vol. 30, no. 7, pp. 682 – 689, 2009.
- [22] M. Nishiyama, H. Takeshima, J. Shotton, T. Koza-kaya, and O. Yamaguchi, Facial Deblur Inference to Improve Recognition of Blurred Faces, Proc. CVPR, pp. 1115 – 1122, 2009.
- [23] E. Tola, V. Lepetit, and P. Fua, DAISY: An Efficient Dense Descriptor Applied to Wide-Baseline, IEEE Trans. PAMI, vol. 32, no. 5, pp. 815 – 830, 2010.
- [24] J. Harel, C. Koch, and P. Perona, Graph-Based Visual Saliency, Proc. Advances in Neural Information Processing Systems, vol. 19, pp. 545 – 552, 2007.
- [25] 佐藤 洋一, 田島 譲二, トーンモデルを用いた配色支援, 全国大会講演論文集 第 44 回平成 4 年前期 (2), pp. 373-374, 1992.

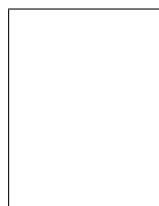
(平成 xx 年 xx 月 xx 日受付)

西山 正志 (正員)



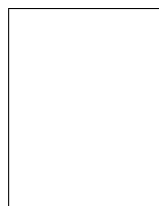
2002 岡山大学大学院博士前期課程了。同年 (株) 東芝入社。現在同社研究開発センターマルチメディアラボラトリー勤務。研究主務。2011 東京大学大学院学際情報学府博士課程修了。学際情報学博士。コンピュータビジョン, パターン認識の研究に従事。2006 本会パターン認識・メディア理解研究会研究奨励賞, 2007 画像センシングシンポジウム優秀論文賞, 2009 山下記念研究賞などを受賞。情報処理学会会員

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

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, 同大学院情報学環准教授を経て現在同大学生産技術研究所教授。コンピュータビジョンに関する研究に従事。平 21, 平 20 及び平 18 電子情報通信学会論文賞, 平 22 情報処理学会 50 周年記念論文賞, 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。