

## フレーミングサポートシステムのための構図特徴を用いた写真の評価

Photo Quality Assessment Using Composition Features  
for Framing Support System濱田 隆平\*  
Ryuhei Hamada堀 磨伊也\*  
Maiya Hori吉村 宏紀\*  
Hiroki Yoshimura岩井 儀雄\*  
Yoshio Iwai

## 1. まえがき

本研究では構図情報を用いた写真の評価により、写真撮影のフレーミングをサポートするシステムを目的とする。近年、デジタルカメラの普及により手軽に写真を撮影、保存、編集できる環境が整備されている。さらに、写真共有サイト [1][2][3] の流行などにより、評価の高い写真を撮影したいというニーズが撮影者に存在する。

撮影したい対象が決定している状況下で品質の高い写真を撮ることができない場合の多くの理由は、構図の問題と言われている [4][5][6]。写真の品質や表現力に関して重要な役割を示す構図を考慮して写真を取得するシステムはいくつか存在する。従来では撮影された写真に対して後処理で自動的にトリミングを行い構図の修正を行う研究 [7][8] があるが、ユーザの意図に反する場合やユーザの撮影時の満足度が満たされない可能性がある。

提案システムでは、あらかじめスコア付けされた学習済み写真データベースとその構図情報を用いて、撮影された写真に対してリアルタイムに自動でスコア付けを行う。さらに、より高いスコア付けが可能な撮影位置・カメラ姿勢を提示することで撮影者のフレーミングをサポートする。本報告では、全体のシステムの中で、写真の構図情報から撮影された写真に対して自動的にスコア付けを行う処理に焦点を当て述べる。

## 2. 写真の審美的品質識別に関する関連研究

撮影された写真の評価の要素として審美的品質が存在する。審美的品質とは、ノイズなどの画像品質とは異なり、人が写真を感覚的に好むかどうかを表す指標である。この審美的品質を識別するために様々な手法が提案されている。それらの多くは、審美的品質のスコア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ものである。審美的品質には写真の構図や色彩調和が重要であることが分かっている。

まず、審美的品質の識別として構図情報を用いる研究を挙げる。写真における代表的な構図として三分割構図、四分分割構図、日の丸構図、対角構図が挙げられる。写真のエッジ情報を用いてこれらの構図を認識し、元画像の印象を最も保ちつつ、より品質が高くなるように写真のトリミングを行い、構図を修正する研究 [7] が存在する。さらに西山ら [8] はエッジ情報のみならず画像の顕著性 [9]、色特徴、ぼけ情報に注目し、写真の品質識別を行っている。これらの研究では、撮影後の後

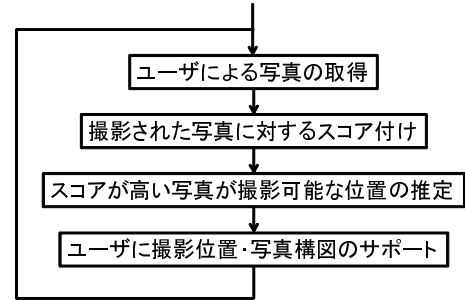


図 1: システム全体の流れ

処理でトリミングを行うことにより構図修正をし、審美的品質の向上を行っている。一方、審美的品質識別の手法として色彩調和を考慮する研究 [10] も存在する。この手法では、エッジ情報、ぼけ情報、顕著性情報に加えて色彩調和情報を審美的品質を決定する局所記述子として用いる。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、この手法で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。

本研究では、これらの中で審美的品質に関して特に重要な役割をなす構図情報に注目する。従来手法 [7][8] では、審美的品質を高めるために撮影後にトリミングを行い構図修正するが、ユーザの意図に反する場合やユーザの撮影時の満足度が満たされない可能性がある。提案手法はこれらに対し、撮影時にリアルタイムで写真にスコア付けを行いフレーミングをサポートすることによって、ユーザの意思を尊重した撮影が可能である。

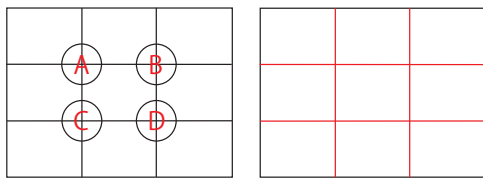
## 3. フレーミングサポートシステムの概要

図 1 に本システムの全体の流れを示す。まず、ユーザはカメラにより対象の撮影を行う。次に撮影された写真に対してシステムが自動的にスコア付けを行う。それと同時に対象の 3 次元形状の取得を行う。取得された対象の 3 次元形状を用いて、擬似的に視点変更を行い、より審美的品質スコアが高い写真が撮影できる地点の探索を行う。最後に推定された撮影推奨地点へユーザを地図で誘導し、撮影推奨視点にてフレーミングサポートをユーザに対して行う。再び撮影した写真に対してスコア付けを行い、同様の処理をユーザが満足するまで行う。以下にそれぞれの詳細を示す。

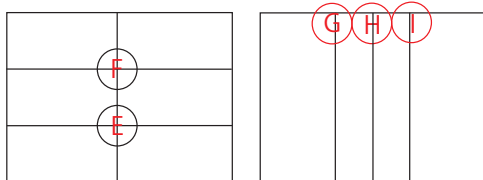
## 3.1. 写真の構図特徴によるスコア付け

ユーザにより撮影された写真に対して審美的品質に関してスコア付けを行う処理を詳述する。本研究では審美的品質に対してスコア付けを行うために写真の構図特徴を用いる。以下に、本研究で使用する代表的

\*鳥取大学大学院工学研究科, Graduate School of Engineering, Tottori University



(a) 三分割線の交点 (b) 三分割線上の点



(c) 三分割線と中央線との交点 (d) 三分割線, 中央線の最上点

図 2: 三分割構図

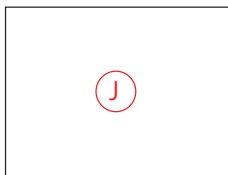


図 3: 日の丸構図

な構図を示す.

(i) 三分割構図

画面を構成する要素が複数ある時に配置を決める経験則の 1 つである. 図 2 各構図のように画面を縦横に三分割した罫線とその交点に構成要素を配置することにより, バランスがとれ安定した構図が得られる.

(ii) 日の丸構図

図 3 のように画面の中心に構成要素を配置した構図である. 人の視線移動に関して非常に強い印象を受け安定する側面があるが, 背景領域の空間処理が難しい. 写真撮影において, 初心者が陥りやすい構図として一般的に避けられる傾向がある. ただし, 消失点が中央に配置された一点透視法などにおいてバランスがとれ安定する場合がある.

(iii) 対角構図

図 4 のように画面を斜めに分割した構図である. 対角線上に構成要素を配置することでリズム感やスピード感といった動作を表現する.

本研究では, 以上の三分割構図, 日の丸構図, 対角構図に注目し, 以下の構図構成要素の抽出を自動的に行い, 特徴量として利用する.

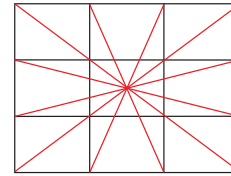
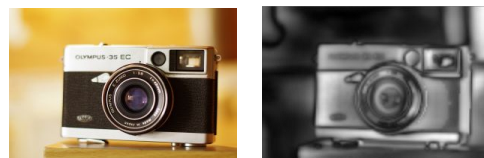


図 4: 対角構図



(a) 元画像 (b) 顕著性マップ

図 5: 顕著領域の抽出例

● 顔領域

顔領域の検出には顔検出の代表的な手法である Haar-like 特徴 [11] を用いる. ブースティング技法を用いて弱分類器を複合させてカスケードを構築することにより高速に物体検出が可能である. 識別に用いる特徴量として, 図 2(a) に示す三分割線の交点上, 図 2(b) に示す三分割線上, 図 2(c) に示す三分割線と中央線との交点上のそれぞれ領域に顔領域があるかどうかの 2 値を用いる.

● 顕著領域

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行っている Itti らの手法 [12] を用いる. この手法は, 入力画像に対してガウシアン・ピラミッドを生成し, スケールごとに輝度・色相成分・方向成分の算出を行う. それらの差分の線形和を顕著度として算出し, 顕著度の高い領域の抽出を行う. 顕著領域の抽出例を図 5 に示す. 識別に用いる特徴量として, 図 2(a) に示す三分割線の交点上, 図 2(b) に示す三分割線上, 図 2(c) に示す三分割線と中央線との交点上のそれぞれ領域に顕著度が高い領域があるかどうかの 2 値を用いる.

● 三角部分

写真の中で三角をなしている領域の抽出を行う. 三角をなしている領域の抽出にはテンプレートマッチングを用いる. 入力画像に対してガウシアンフィルタによってノイズ除去を行う. ノイズ除去された平滑画像に対して Canny アルゴリズム [13] を用いてエッジを検出する. エッジ画像に対してしてモルフォロジ処理を行うことによってエッジが拡大される. このエッジ画像に対してテンプレートマッチングを行う. この処理は図 6 に示される形状, スケールと形状を変化させた三角形テンプレートをエッジ画像に対しマッチングさせるものである. テンプレートの形状は各頂点と各裾との 2

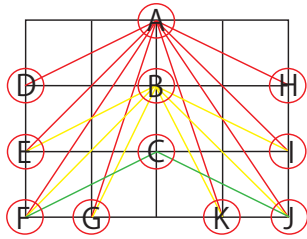


図 6: 三角形テンプレート

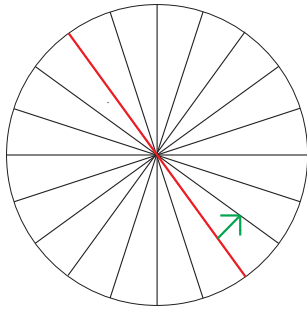


図 7: 対角線抽出のためのテンプレート

点を結ぶものである。A, B, Cを頂点とする三角形はそれぞれ16, 9, 4個あり、29通りが全形状となる。テンプレートのスケールを可変させ、マッチングしたピクセル数が最大となる部分を三角部分として登録する。識別に用いる特徴量として、図2(a)に示す三分割線の交点上、図2(b)に示す三分割線上、図2(c)に示す三分割線と中央線との交点上、図2(d)に示す三分割線・中央線の最上点上のそれぞれ領域に顕著度が高い領域があるかどうかの2値を用いる。

#### ● 水平線

一定の長さを持つ水平に近い線の抽出を行う。まず入力された画像はCannyアルゴリズムを用いてエッジを検出する。エッジ画像に対してハフ変換により直線検出が行われる。ハフ変換のパラメータは、角度分解能が $\pi$ 、直線の長さの閾値が画像の横幅の1/2である。これにより抽出された水平線の中心の座標を構成要素の水平線として登録される。識別に用いる特徴量として、図2(b)に示す三分割線上に水平線があるかどうかの2値を用いる。

#### ● 対角線・消失点

画面の端と端とを結ぶ斜めの線を対角線として、現実世界での平行線が遠近法において交わる点を消失点として抽出する。対角線の抽出ではまずエッジの拡大の処理までは三角部分の抽出と同様の処理が行われる。その後、エッジ画像に対してテンプレートマッチングを行う。テンプレートを図7に示す。テンプレートは18度ずつ変化させ、テンプレートとエッジ画像のマッチングしたピクセル数が最大な直線が対角線となる。抽出された対角

線の端点の2つ座標が構成要素の対角線として登録される。

消失点の抽出では、まずCannyアルゴリズムによってエッジを抽出する。次にエッジ画像に対してハフ変換を行う。ハフ変換のパラメータは角度分解能が $\pi/180$ 、直線の長さの閾値が横幅の1/2である。これにより抽出された2直線の交点が画像内に存在する場合、その交点を構成要素の消失点として登録する。識別に用いる特徴量として、図4に示す対角線上にエッジが存在するかどうか、図2(a)に示す三分割線の交点上、図2(c)に示す三分割線と中央線との交点上、図3に示す中央線上に消失点があるかどうかの2値を用いる。

以上の構成要素が図2の三分割構図、図3の日の丸構図、図4の対角構図の罫線および交点(赤色の領域)上に存在するかどうかを $\{0, 1\}$ の15次元の特徴量として入力とし、学習および識別を行う。学習はあらかじめ人手によって審美的品質のスコアが与えられた訓練サンプルに対して、Support Vector Regression (SVR)[14]により行う。学習後にSVRにより生成された識別器を用いて、ユーザが撮影した写真に対して構図特徴をもとにスコア付けを行う。

### 3.2. 対象の3次元形状を利用したフレーミングサポート

ユーザが撮影した写真に対して、より審美的品質の高い写真を撮影するためにフレーミングサポートを行う手法について詳述する。まず、ユーザが撮影した写真にスコア付けを行うと同時に撮影対象を3次元復元する。3次元形状を復元する手法として撮影された写真群に対してStructure from Motion法を適用する手法[15]や奥行センサを用いる手法[16]などが挙げられる。これらの手法を用いてユーザが写真撮影を行うたびに対象の3次元形状の取得を行う。生成された3次元モデルを用いて、擬似的に視点位置を変更した際の視点画像の生成を行う。生成した視点画像に対して再び構図特徴の抽出を行い、自動的にスコア付けを行うことにより、よりスコアが高い視点画像が取得できる撮影位置を推定する。

ユーザには、よりスコアが高い視点画像が取得可能な撮影位置を地図上にマッピングすることによって提示する。ユーザが提示された位置周辺に移動した後にCGの重畳などによりフレーミングサポートを行う。ユーザ位置はGPSなどのセンサ、カメラの位置・姿勢情報は加速度センサやジャイロセンサやコンパスなどによって取得可能とする。

### 4. 審美的品質識別に関する評価実験

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。写真にスコア付された画像データベースとしてDPChallenge[2]を用いる。DPChallenge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。これらの大量の写真とそれらに付加されたスコアを利用することで人間が共通で持つ主観を統計的に学習する。実際には評価にDPChallengeの上位500枚、下位500枚の計1000枚の写真を用いる。上位500枚

表 1: SVR で使用したパラメータ

内容	パラメータ
SVM のタイプ	$\nu$ -SVM regression
カーネル関数のタイプ	放射基底関数 (RBF)
カーネルのキャッシュサイズ	1
constraints violation の値	10
$\nu$ -SVR 用のパラメータ	0.5
クラスのための重み	無し

のスコアの平均は 7.726 (標準偏差 0.635), 下位 500 枚のスコアの平均は 2.691 (標準偏差 0.238) であった. 実験で用いた SVR に関するパラメータを表 1 に示す. ここでは SVR による学習の識別性能を評価するために計 1000 枚に対して 10 分割交差検証を行った. テストサンプル 1000 枚に対する結果として, 審美的品質のスコアの誤差は平均 2.548 (標準偏差 0.299) となった. 実験の結果, 自動的に付けた審美的スコアがユーザの与えたスコアと大きく異なる問題が生じた. これは構図特徴として用いた 15 次元の特徴量が審美的スコア付けを行う特徴として不十分な次元であるためだと考えられる. 次元を増やすために, 構成要素の位置関係の特徴量やその他の構成要素として色彩調和やぼけ等を追加していく必要があると考えられる.

## 5. むすび

本報告では構図情報を用いた写真の評価により, 写真撮影のフレーミングをサポートするシステムを提案した. 本研究では撮影された写真に対して自動的にスコア付けを行う処理に焦点を当て実験を行った. 実験では, 撮影された写真から構図特徴を抽出し, あらかじめ審美的品質のスコアが与えられた訓練サンプルを用いて学習したデータを用いて自動的に審美的スコア付けを行う実験を行った. 実験の結果, 自動的に付けた審美的スコアがユーザが与えたスコアと大きく異なる問題が生じた. これは構図特徴として用いた 15 次元の特徴量が審美的スコア付けを行う特徴として不十分な次元であるためだと考えられる. 今後の課題として構図特徴量の見直しと構図以外の特徴の検討が挙げられる.

**謝辞** 本研究の一部は, 新学術領域研究 (No. 24118705) の助成を受けたものである.

## 参考文献

- [1] flickr. <http://www.flickr.com/>.
- [2] DPChallenge. <http://www.dpchallenge.com>.
- [3] Photo.net. <http://photo.net>.
- [4] ○と×比べてわかるデジタル写真 (NCフォトシリーズ). 日本カメラ社, 2008.

- [5] 上手に撮れるツボとワザ! デジタル一眼レフ. 学研マーケティング, 2008.
- [6] 杉本恭子. 一眼レフカメラ基本撮影テクニック. 永岡書店, 2003.
- [7] 家田暁, 琴智秀, 萩原将文. 感性を反映した構図修正による写真品質向上システム. 芸術科学会論文誌, Vol. 9, No. 4, pp. 154–162, 2010.
- [8] 西山正志, 岡部孝弘, 佐藤洋一, 佐藤いまり. 複数の注目領域を用いた写真の主観的品質の識別. 電子情報通信学会論文誌. D, 情報・システム, Vol. J93-D, No. 8, pp. 1364–1374, 2010.
- [9] A. Borji and L. Itti. State-of-the-art in visual attention modeling. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 35, No. 1, pp. 185–207, 2013.
- [10] 西山正志, 岡部孝弘, 佐藤いまり, 佐藤洋一. 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. 電子情報通信学会論文誌. D, 情報・システム, Vol. J94-D, No. 8, pp. 1324–1334, 2011.
- [11] R. Lienhart and J. Maydt. An extended set of haar-like features for rapid object detection. *Int'l Conf. Image Processing*, Vol. 1, pp. 900–903, 2002.
- [12] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 20, No. 11, pp. 1254–1259, 1998.
- [13] J. Canny. A computational approach to edge detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, Vol. 8, No. 6, pp. 679–698, 1986.
- [14] C.-C. Chang and C.-J. Lin. Libsvm: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, Vol. 2, No. 27, 2011.
- [15] N. Snavely, S. M. Seitz, and R. Szeliski. Modeling the world from internet photo collections. *Int'l Journal of Computer Vision*, Vol. 80, No. 2, pp. 189–210, 2008.
- [16] S. Izadi, R. A. Newcombe, D. Kim, O. Hilliges, D. Molyneaux, S. Hodges, P. Kohli, J. Shotton, A. J. Davison, and A. Fitzgibbon. Kinectfusion: Real-time dynamic 3D surface reconstruction and interaction. *ACM SIGGRAPH*, No. 23, 2011.