

## 感性を反映した構図修正による写真品質向上システム

家田 暁 琴智秀 萩原将文

慶應義塾大学 理工学部

## Photo Quality Improvement System by Modification of Composition Reflecting *Kansei*

Akira Ieda Ji-Soo Keum Masafumi Hagiwara

Keio University

{ieda, keum, hagiwara}@soft.ics.keio.ac.jp

### アブストラクト

本論文では感性を反映した構図修正による、デジタル写真の品質向上システムを提案する。提案システムでは入力画像に適した構図を自動選択し、画像の構図を原画像の印象を保ちながら修正することで、画像の品質向上を行うことができる。修正は4段階の処理によって行われる。まず、入力された画像から、顔、際立つ領域、三角部分、水平線、対角線、遠近法消失点の6種類の主となる構成要素が検出される。検出された各構成要素の位置や他構成要素との関係から、それぞれの構成要素に適すると考えられる出力構図案が計算される。次にそれらの構図案から、原画像の印象から大きく変化する案が削除される。すなわち、画像に対する構成要素の位置が大きく移動する案が削除される。最後に、残った出力構図案の中から原画像の印象を最も保つことが可能な構図案、すなわち切り出す面積が最も大きい構図案が選択され、その構図案に従い画像の構図修正が行なわれる。ユーザアンケートによる2種類の評価実験を行った。その結果、提案システムによりユーザにとって好ましい修正が行われること、また既存手法と比較しても好ましい修正が行われることが確認された。

キーワード: 写真, 感性, 品質向上, 構図

### Abstract

This paper proposes a new system for photo quality improvement by modification of composition reflecting *Kansei*. The proposed system can improve photo quality by selecting a proper composition for an input image and modifying the composition while keeping the impression. The modification consists of the following four-step process. First, the proposed system detects main components from the input image. Second, the proposed system calculates proper composition plans for each main component from the detected position and relation with the other components. Then, the proposed system removes some composition plans which significantly change the impression of the original image. Finally, the proposed system selects the most appropriate composition plan which can maintain the impression of original image maximally. In this way, the proposed system modifies the image composition according to the plan. In order to evaluate the proposed system, two kinds of experiments were carried out. It is confirmed that the proposed system can modify the image to the preferable image for user and can make more desirable image for user compared with the previous method.

*Keywords:* Photo, *kansei*, quality improvement, composition

## 1. はじめに

近年、大量のデジタル写真を撮影、保管、編集できる環境が整備されて来ている。その背景にはコンパクトデジタルカメラ・カメラ付き携帯電話の普及、flickr[1], Picasa[2], Photo.net[3]のようなネット上の写真共有の流行があげられる。デジタル写真の普及とともに、Adobe PhotoShop[4], Paint Shop Pro[5], GIMP[6]などのフォトタッチソフトが普及している。また、Picnik[7], flauntR[8], FotoFlexer[9]などの簡易画像修正サイトも出現している。これらソフトやサイトの登場から、近年、所持している写真の品質・表現力を向上させたいというニーズがあることが分かる。

写真の品質や表現力に関して、構図は重要な要素である[10]-[19]。しかし、フォトタッチソフトは写真の構図を変更する等の高度な修正が可能である一方、手動で切り出すため手間がかかり、専門知識も必要である。また簡易画像修正サイトは手軽に使用可能であるが、色やコントラストの変更や特殊効果の付与などの単純な修正しか出来ない。双方の欠点を補うような手軽に構図を修正出来るシステムはデジタル写真を扱うユーザーに大きな利益をもたらすと考えられる。撮りたいものが決まっているが上手く撮ることが出来ない多くの場合は、構図の取り方に問題があると言われている[10]-[19]。

写真構図を考慮したシステムの研究[20][21]はいくつか行われている。Banerjeeら[20]は写真撮影中における構図適用の自動化システムを提案した。このシステムはあくまで撮影中の補助であるので、既に撮影済みの写真に対しては適用できない。また配置については3分割の位置に主要物体を置くことのみが行われている。Changら[21]はパノラマ風景写真から好ましい構図を発見するシステムを提案した。このシステムでは書籍等に載っている体系化された構図法を適用するのでは無く、プロの写真を参照画像として用い、その構図と類似した部分を切り出すものである。しかし入力画像は大きなパノラマ写真でなくてはならず、一般のデジタル写真ユーザーが使用するには問題がある。

撮影後の一般的なデジタル写真を扱った自動画像切り出しシステムについての研究がいくつか行われてきた[22]-[26]。Suhら[22]は主要物体を検出し、自動でサムネイルを切り出すシステムを提案した。これはSaliency map[27][28]を用いて主要物体を検出し、最適なバランスで主要物体を含み、かつ小さく切り出すというシステムである。Santellaら[23]は視線追跡を用いた画像の自動切り出しシステムを提案した。しかし、視線追跡が必要であるため手間がかかり、また構図については重要視していない。Luo[24]は主要物体を検出し、その主要物体をできるだけ含むように切り出すシステムを提案した。Amruthaら[25]は自動でRegions of Interests (ROIs) を抽出して、それを元に切り出すシステムを提案した。Nishiyamaら[26]は、大量の画像から学習によって品質分類器を構築し、次に主要物体を含む切り出し候補をいくつか生成し、候補の中から品質が高いと推定されるものを選択するシステムを提案した。しかし、これら既存の自動切り出しシステムではユーザーにとって主観的に満足いく結果

が得にくい。原因として写真の品質・表現力を向上させる技術として一般的である構図を考慮していないことが挙げられる。

近年の増加しているデジタル写真の一般ユーザは、一般的なサイズの写真を数多く所持している。このような撮影済み写真の構図を自動修正することで写真の品質と表現力を向上させるニーズは大きいと考えられる。

そこで本論文では、感性を反映した構図修正による、デジタル写真の品質向上システムを提案する[37]。提案システムは6種類の構図を組み合わせて扱うことが可能である。また原画像の印象を保った切り出しを行うことができる。

以降、本論文では、第2章で提案システムについて詳しく説明する、第3章では評価実験について説明し、第4章を結論とする。

## 2. 感性を反映した構図修正による写真品質の向上システム

### 2.1 概要

提案システムでは、ユーザが画像を入力するだけで、入力画像に適した構図を自動選択し、原画像の印象を保ちながら画像の構図の修正が行われる。図1に提案システムの流れを示す。この手順は手動で画像の修正を行う際、主題を特定し、いくつかの修正案を作成し、そこから選択するという感性を反映する一般的な流れを参考にした。システムは以下のように動作する。

- Step 1 ユーザによる画像入力
  - Step 2 画像から主となる構成要素を検出
  - Step 3 要素の状況から出力構図案を計算
  - Step 4 切り出し制限によって出力構図案を選定
  - Step 5 原画像との差から最終出力構図案を決定
  - Step 6 最終出力構図案に従って修正された画像を出力
- ユーザはまず、Step1で修正したい画像を入力する。後の処理は自動で行われる。
- Step2では画像から6種類の主となる構成要素が検出される。
- Step3では検出された要素それぞれの状況から出力構図案が決定される。出力構図案は、構成要素に適する形で三分割法やその他の構図[10]-[19]を組み合わせた構図案となる。
- Step4では画像に対する構成要素の位置が大きく移動する出

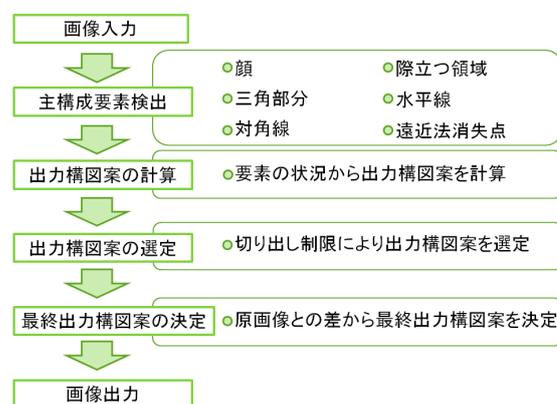
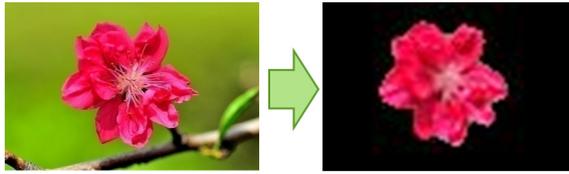


図1 提案システムの概要



入力画像 際立つ領域検出画像

図2 際立つ領域検出例

力構図案を削除する。

Step5 では原画像と出力構図案の切り出し量の差から最終出力構図案が決定される。

Step6 では最終出力構図案に従って修正された画像が出力される。

## 2.2 主構成要素検出

写真構図にはいくつかの重要な構成要素がある[10]-[19]。写真構図としての重要性を第一に、次にソフトウェア上での実現性を考慮した。その観点から以下の6つの構成要素に関する検出を行う。画像はダウンサンプリングされ、各処理が行われる。

- ・顔： 人物の顔
- ・際立つ領域： 特徴的な色を持ち、背景・小さな物体ではない領域
- ・三角部分： 三角をなしている領域
- ・水平線： ある程度の長さを持つ水平に近い線
- ・対角線： 画面の端と端を結ぶ斜めの線
- ・遠近法消失点： 線遠近法で投射線が一定点に集中する点

### (1) 顔検出

顔検出はNilssonら[29]の顔検出手法を用い、顔領域 $C_{face}$ が検出される。次に新たに追加した顔確定処理が行われる。この処理では肌色の確認によって検出された領域が、顔であるかの判定を行う。具体的には検出された四角形の内側の肌色ピクセル数 $S_{skin}$ が、四角形のピクセル数 $S_{quad}$ の $T_{skin}$ 以上の割合である時、顔と判定され $C_{face}$ として構成要素に登録される。肌色色相は文献[30]を参考に $6^\circ \sim 38^\circ$ と定義とした。

$$C_{face} = \begin{cases} C_{face}, & (S_{skin} \geq S_{quad} T_{skin}) \\ None, & (S_{skin} < S_{quad} T_{skin}) \end{cases} \quad (1)$$

### (2) 際立つ領域検出

際立つ領域を、画像中で主要な色を持ち、背景または小さな物体ではない領域と定義する。図2に検出例を示す。式(2)-(4)に示すように、入力された画像 $I$ はノイズ除去のためにガウスフィルタ $F_{gaussian}(H_{dst,blur}$  (フィルタサイズ),  $\sigma_{dst,blur}$  (標準偏差))との畳みこみで平滑化され、平滑化画像 $I_{blur}$ が生成される。なお座標系は画像の左下が原点となっている。

$$h_g(x_g, y_g) = e^{-(x_g^2 + y_g^2) / 2\sigma_{dst,blur}^2} \quad (2)$$

$$F_{gaussian} = \frac{h_g(x_g, y_g)}{\sum_{x_g} \sum_{y_g} h_g} \quad (3)$$

$$I_{blur} = I * F_{gaussian} \quad (4)$$

$x_g, y_g$  : フィルタ内での位置

平滑化画像 $I_{blur}$ は $L*a*b$ 表色系[31][32]の $a$ と $b$ を用いて、 $k$ -means法( $k = k_{dst}$ )によって領域分割される。これにより主要な色で領域が分割される。分割された領域 $r$ のうち小さすぎる領域または背景領域が削除される。

式(5)に示すように、画像4辺に接している面積の割合が、画像総面積 $Z$ の $T_{dst,bg1}$ 以上の領域 $r$ は背景とみなされ削除される。

次に式(6)で示すように各 $k_{dst}$ 個の領域の中で、互いに接していない領域を更に細かい1つの小領域 $r''$ とする。式(7)に示すように $r''$ のうち画像総面積の $T_{dst,small}$ 以下の面積である小領域が削除され、また小領域の総面積の $T_{dst,bg2}$ 以上が画像4辺に接していたらその小領域は削除される。残った領域 $r^*$ を統合させた領域が際立つ領域 $C_{dst}$ として登録される。

$$r' = \{r'_i \in r : \theta(r'_i) < Z \times T_{dst,bg1}\} \quad (5)$$

$$r'' = \{r''_j \in \cup \varphi(r')\} \quad (6)$$

$$r^* = \{r^*_k \in r'' : \delta(r^*_k) > Z \times T_{dst,small} \text{ and } \mu(r^*_k) < Z \times T_{dst,bg2}\} \quad (7)$$

$$C_{dst} = \cup r^* \quad (8)$$

$\theta(\cdot)$ : 領域が画像の4辺に接している面積を計算する関数

$\varphi(\cdot)$ : 領域の中で離れている部分を小領域として分割する関数

$\delta(\cdot)$ : 小領域の総面積を計算する関数

$\mu(\cdot)$ : 小領域が画像4辺に接している面積を計算する関数

### (3) 三角部分検出

三角部分の検出はテンプレートマッチングによって行われる。入力された画像はノイズ除去のために際立つ領域検出と同様にガウスフィルタ $F_{gaussian1}(H_{tri,blur}, \sigma_{tri,blur})$ と原画像 $I$ との畳み込みにより平滑化され、平滑化画像 $I_{blur}$ が生成される。式(9)-(11)に示すようにその平滑化画像 $I_{blur}$ に対しLoG(Laplacian of Gaussian)フィルタ[33] $F_{log}(H_{tri,edge}, \sigma_{tri,edge})$ との畳みこみによるエッジ検出が行われる。

$$h_g(x_g, y_g) = e^{-(x_g^2 + y_g^2) / 2\sigma_{tri,edge}^2} \quad (9)$$

$$F_{log} = \frac{(x_g^2 + y_g^2 - 2\sigma_{tri,edge}^2)h_g(x_g, y_g)}{2\pi\sigma_{tri,edge}^6 \sum_{x_g} \sum_{y_g} h_g} \quad (10)$$

$$I_{edge} = I * F_{log} \quad (11)$$

検出されたエッジは、際立つ領域検出時の平滑化と同様のガウスフィルタ $F_{gaussian2}(H_{tri,thick}, \sigma_{tri,thick})$ を用いて平滑化される。次に式(13)に示すように、0以外の値を持ったピクセルに1が与えられ、エッジが拡大される。

$$I_{edge,blur} = I_{edge} * F_{gaussian2} \quad (12)$$

$$I_{edge,thick}(x, y) = \begin{cases} 0, & (I_{edge,blur}(x, y) = 0) \\ 1, & (I_{edge,blur}(x, y) \neq 0) \end{cases} \quad (13)$$

このエッジ画像に対し、テンプレートマッチングが行われる。この処理は図3に表わされる形状、スケールと位置を変化させた三角形のテンプレートを、エッジ画像に対しマッチングさせる

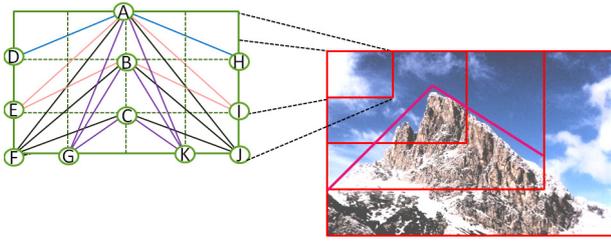


図3 三角部分検出テンプレート

ものである。テンプレートの形状は図3の各頂点と各裾の2点を結ぶ三角形である。例えばAとD, Hを結んだ三角形, CとG, Jを結んだ三角形である。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。テンプレートの初期サイズは画像面積 $Z \times S_{tri}$ とし, 以降4倍, 9倍, ...と大きくなっていく。各サイズにおいてテンプレートの位置は初期テンプレートの縦及び横の半分のピクセル数ずつスライディングさせていく。テンプレート $v_{x,y}$ とエッジ画像 $I_{edge,thick}$ のマッチングしたピクセル数が最も多い部分が三角部分 $V_{max}$ として選択される。

$$V_{max} = v_{x^*, y^*, s^*} \quad (14)$$

$$\{x^*, y^*, s^*\} = \arg \max_{\substack{x, y \in I_{edge,thick} \\ s \in A}} (f(v_{x,y,s}, I_{edge,thick})) \quad (15)$$

$$A = \{Z \times S_{tri}, Z \times 4S_{tri}, Z \times 9S_{tri}, \dots, Z\} \quad (16)$$

$f(\cdot, \cdot)$ :  $v_{x,y,s}$ と $I_{edge,thick}$ のマッチングピクセル数を計算する関数  
 選択された部分の正規化スコア $P_{tri,nor}$ が以下の式で算出され, その値が閾値 $T_{tri}$ 以上ならば三角部分の頂点座標が $C_{tri}$ として登録される。

$$P_{tri,nor} = \frac{P_{tri,match}}{P_{tri,total}} \quad (17)$$

$$C_{tri} = \begin{cases} (x^*, y^*), & (P_{tri,nor} \geq T_{tri}) \\ None, & (P_{tri,nor} < T_{tri}) \end{cases} \quad (18)$$

$P_{tri,match}$ : マッチングピクセル数

$P_{tri,total}$ : テンプレート上で値を持っている総ピクセル数

#### (4) 水平線検出

入力された画像はまずコントラストの強調が行われる。これには式(19)のラプラシアンフィルタ $F_{unsharp}$  ( $H_{hor, cnt} = 3, \alpha_{hor, cnt}$  (形状調節パラメータ)) が用いられる。

$$F_{unsharp} = \frac{1}{\alpha_{hor, cnt} + 1} \begin{bmatrix} -\alpha_{hor, cnt} & \alpha_{hor, cnt} - 1 & -\alpha_{hor, cnt} \\ \alpha_{hor, cnt} - 1 & \alpha_{hor, cnt} + 5 & \alpha_{hor, cnt} - 1 \\ -\alpha_{hor, cnt} & \alpha_{hor, cnt} - 1 & -\alpha_{hor, cnt} \end{bmatrix} \quad (19)$$

このフィルタと原画像 $I$ との畳みこみにより, コントラストの強調された画像 $I_{cnt}$ が計算される。

$$I_{cnt} = I * F_{unsharp} \quad (20)$$

次にこの $I_{cnt}$ に対しハフ変換による直線検出が行われる。検出された直線 $\ell$ のうち, 長さが画像横幅の $T_{hor, len}$ 以上でかつ水平線と為す角度が $-T_{hor, ang}$ から $T_{hor, ang}$ のもののうち中点のY座標が最大のものが水平線 $C_{hor}$ として登録される。

$$C_{hor} = \ell^* \quad (21)$$

$$\ell^* = \arg \max_{\ell' \in L} (h(\ell')) \quad (22)$$

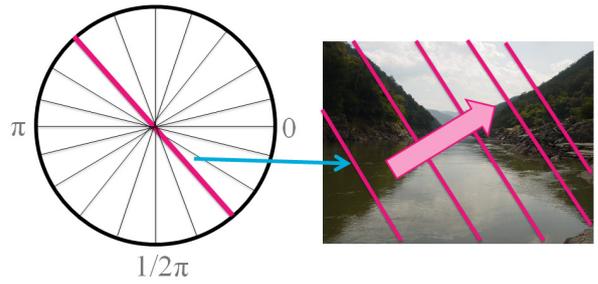


図4 対角線・遠近法消失点検出テンプレート

$$L = \{\ell_i' \in \ell : \nu(\ell_i') \geq X_{size} \times T_{hor, len} \text{ and } -T_{hor, ang} \leq \varepsilon(\ell_i') \leq T_{hor, ang}\} \quad (23)$$

$h(\cdot)$ : 直線の中点のy座標を計算する関数

$\nu(\cdot)$ : 直線の長さを計算する関数

$\varepsilon(\cdot)$ : 直線が水平線と為す角度を計算する関数

$X_{size}$ : 原画像の横幅

#### (5) 対角線・遠近法消失点検出

エッジの拡大までは三角部分検出と同様の処理が行われる。このエッジ画像に対し, テンプレートマッチングが行われる。この処理は図4に表わされるように角度, 位置を変化させた直線のテンプレートを, エッジ画像に対しマッチングさせるものである。テンプレートの角度は $\pi/2$ と $\pi$ を除き $D_{line}$ から $\pi$ まで,  $D_{line}$ ずつ大きくなっていく。テンプレートの位置は1ピクセルずつスライディングさせていく。 $(x,y)$ から角度 $d$ に伸びる直線を持つテンプレート $\ell_{x,y,d}$ とエッジ画像 $I_{edge}$ のマッチングしたピクセル数が最も多い直線部分が式(24)-(26)に示すように $L_{max}$ として選択される。

$$L_{max} = \ell_{x^*, y^*, d^*} \quad (24)$$

$$\{x^*, y^*, d^*\} = \arg \max_{\substack{x, y \in I_{edge} \\ d \in A}} (f(\ell_{x,y,d}, I_{edge})) \quad (25)$$

$$A = \{D_{line}, 2D_{line}, \dots, \frac{1}{2}\pi - D_{line}, \frac{1}{2}\pi + D_{line}, \dots, \pi - D_{line}\} \quad (26)$$

$f(\cdot, \cdot)$ :  $\ell_{x,y,d}$ と $I_{edge}$ のマッチングピクセル数を計算する関数

また式(27)-(29)のように,  $\pi/2$ を挟んで $L_{max}$ と反対側の角度群の直線のうちマッチングしたピクセル数が最も多いものが $L_{second}$ として選択される。

$$L_{second} = \ell_{x^*, y^*, d^*} \quad (27)$$

$$\{x^*, y^*, d^*\} = \arg \max_{\substack{x, y \in I_{edge} \\ d \in A}} (f(\ell_{x,y,d}, I_{edge})) \quad (28)$$

$$A = \begin{cases} \{D_{line}, 2D_{line}, \dots, \frac{1}{2}\pi - D_{line}\}, & (\frac{1}{2}\pi < d^*) \\ \{\frac{1}{2}\pi + D_{line}, \frac{1}{2}\pi + 2D_{line}, \dots, \pi - D_{line}\}, & (d^* < \frac{1}{2}\pi) \end{cases} \quad (29)$$

式(30),(31)に示すように, 2直線の交点が画像内にあり,  $L_{max}$ と $L_{second}$ の正規化スコア $P_{line, no, b}$ が閾値 $T_{line}$ 以上ならば, 2直線の交点が遠近法消失点 $C_{per}$ として登録される。 $L_{max}$ の正規化スコアのみが閾値 $T_{line}$ 以上ならば, その直線が対角線 $C_{dia}$ として登録され

る.

$$C_{per} = \begin{cases} g(L_{max}, L_{second}), & (P_{line,nor,I_{max}} > T_{line} \\ \text{and } P_{line,nor,I_{second}} > T_{line} \text{ and } g(L_{max}, L_{second}) \in I_{edge}) & (30) \\ None, & \text{otherwise} \end{cases}$$

$$C_{dia} = \begin{cases} L_{max}, & (P_{line,nor,I_{max}} > T_{line} \text{ and } P_{line,nor,I_{second}} < T_{line}) \\ None, & \text{otherwise} \end{cases} \quad (31)$$

$g(.,.)$ :直線 $L_{max}$ と直線 $L_{second}$ の交点を計算する関数  
正規化スコア $P_{line,nor,l}$ は以下の式で計算される.

$$P_{line,nor,l} = \frac{P_{tri,match,l}}{P_{tr,total,l}} \quad (32)$$

$P_{line,match,l}$ : 直線 $l$ とエッジ画像のマッチングピクセル数

$P_{tr,total,l}$ : テンプレート上で値を持っている総ピクセル数

### 2.3 出力構図案の計算

検出された各構成要素の位置や他構成要素との関係から、その状況に合わせた出力構図案が決定される。図5に示す(a)-(f)の配置位置に構成要素が配置されるような出力構図案が計算される。表1に示すように、検出された構成要素によってこれらの組み合わせが異なり、構成要素ごとの構図案が計算される。検出された各構成要素における構図案全てを合わせたものを、出力構図案とする。

#### (1) 各配置位置と切り取り方

図5の(a)は、三分割線の交点に構成要素が配置される構図案であり、この位置に構成要素が来るように画像の上下左右が切り取られる。

(b)は水平方向の三分割線と垂直方向の中央線の交点に構成要素が配置される構図案であり、この位置に構成要素が来るように画像の上下左右が切り取られる。

(c)は垂直方向の三分割線と二分割線の最上部に構成要素が配置される構図案であり、この位置に構成要素が来るように画像の上下左右が切り取られる。

(d)は三分割線上のどこかに構成要素が配置される構図案であり、この位置に構成要素が来るように画像の上下または左右が切り取られ、アスペクト比を保つように左側のみ、右側のみ、または上側のみ、下側のみが切り取られる。

(e)は対角線が画像の4隅から対角の隅に向けて配置、または三分割線の起点となる位置同士で斜めのラインが構成されるように配置される構図案である。前者の案は画像の上下左右が切り取られ、後者の案は上下または左右が切り取られる。

(f)は構成要素が画像の中心に来るように配置される構図案であり、この位置に構成要素が来るように画像の上下左右が切り取られる。

#### (2) 三角構図と水平線構図

図6に示すように、三角構図と水平線構図は三角部分と水平線の状況によっていずれかの構図案もしくは両方の構図案が出力される。三角部分が検出され、水平線が検出されなかった場合は三角構図の構図案が出力される。水平線が検出され、三角部分が検出されなかった場合、水平線から三角部分の頂点までの距離が原画像の縦サイズ $T_{tri,hor}$ 以下の場合、水平線が三角部分の

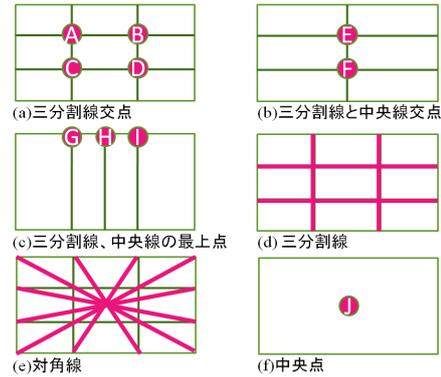


図5 構成要素の配置位置パターン

表1 各構図の配置位置パターン組み合わせ

構図	配置位置
顔構図	(a), (b), (d)
際立つ領域構図	(a), (b), (d)
三角構図	(a), (b), (c), (d)
水平線構図	(d)の水平方向
対角線構図	(e)
遠近法構図	(a), (b), (f)

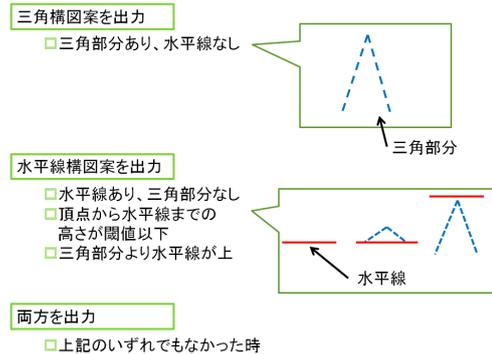


図6 三角構図案と水平線構図案の選択

頂点より上にある場合、水平線構図の構図案が出力される。上記のいずれでもなかった場合には両方の出力構図案が出力される。

#### (3) 構図案の削除

それぞれの構図案を計算する際、いくつかの条件により、構図案から削除されるものがある。いずれの構図においても修正後の画像サイズが元サイズの $T_{cmp}$ 以下になる案は削除される。また顔が検出されていた場合、顔領域を $E_{cmp,face}$ 倍した領域が切り出し後の画像外にはみ出すものは削除される。加えて、足の不自然な切断を防ぐため、顔領域が顔領域サイズのいくつ分の高さにあるかを計算する。高さが顔領域サイズの $E_{cmp,foot}$ 倍以上であるとき、画像の下部分を切り取るものは削除される。その他、構図ごとに削除される条件を以下に示す。

際立つ領域構図では、際立つ領域が切り出し後の画像外にはみ出すものは削除される。三角構図では、三角部分の頂点が切り出し後の画像外にはみ出すものは削除される。また際立つ領域が検出され、その領域内に三角部分の頂点があるとき、際立つ領域が切り出し後の画像外にはみ出すものは削除される。対

角線構図では、着目している3分割線の内側に対角線が検出された場合はその案は削除される。遠近法構図では、遠近法消失点が切り出し後の画像外にはみ出すものは削除される。

### 2.4. 切り出し制限による出力構図案の選定

元の画像の印象を保ち、無理な切り出しが行われないように、主構成要素の位置によって出力構図案が更に絞られる。原画像での画像全体に対する位置から近い位置に移動する出力構図案に制限されることとなる。対角線構図を除く全ての構図にこの切り出し制限が設けられている。

図7に構成要素の元の位置と移動先の関係の全パターンを示す。色の濃い領域に構成要素があるとき、太い線とか、丸で表わされた位置に構成要素が来るような構図案に絞られる。X軸方向、Y軸方向それぞれで、構成要素の位置から出力される出力構図案を絞っていく。

表2に各構図でのパターンの組み合わせを示す。X軸方向は三分割線と垂直線を基準とするため基本的に5パターンである。水平線構図はX軸方向の制限は存在しない。Y軸方向は三分割線を基準とするので対角線構図を除き3パターンである。図7に示すように、範囲の区切り方は三分割線(実線)と、三分割線内を更に三分割した線(点線)の計4本を用いる。5パターンの場合はこの4本、3パターンの場合は点線を境とする。なお、図7の(x-2)と(y-2)はいくつかの構図での切り出し制限を合わせて示しているため、元の構図案には無い点が描かれているが、その点は無視してよい。

### 2.5. 最終出力構図案の決定

各構成要素に合わせた出力構図案の中から最終出力が1つ選択される。選択する基準として、写真の印象を保ちながら品質向上を行うという観点のもと、なるべく変更量が少ない、すなわち式(33)のように出力構図案により切り出された画像 $I_i$ と原画像 $I$ の面積差が最も少ない出力構図案が最終出力構図案 $I_{final}$ として選択される。

$$I_{final} = \arg \min(\kappa(I_i, I)) \quad (33)$$

$\kappa(\dots)$ :2画像の面積の差を計算する関数

## 3. 実験

提案システムの特性を評価するために、以下の項目に関して実験を行った。

- ・効果性：ユーザにとって好ましい修正がなされるか
- ・有効性：既存手法と比較して好ましい画像が出力されるか

### 3.1 実験 1

#### (1) 目的

システムの効果性を評価するために、原画像と比較して提案システムによって修正が行われた画像がユーザにとって好ましい画像になったかの評価を行った。

#### (2) 実験方法

使用画像はフォト蔵[34]、Google 画像検索[35]、Flickr の画像を集めたDB合計460枚を用いた。使用したカテゴリは表3で示した11カテゴリである。ユーザに、DBの中から各カテゴリ15

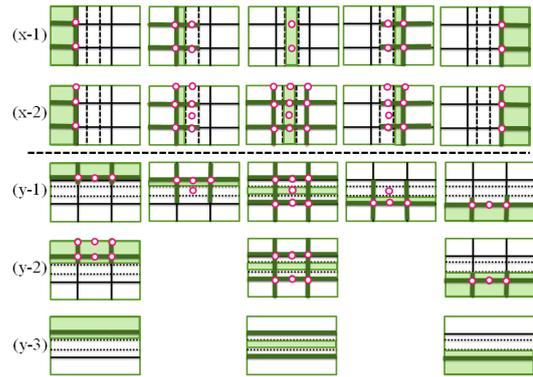


図7 切り出し制限パターン

表2 各構図の切り出し制限パターン組み合わせ

構図	組み合わせ
顔構図	(x-1), (y-2)
際立つ領域構図	(x-2), (y-2)
三角構図	(x-2), (y-2)
水平線構図	(y-3)
遠近法構図	(x-2), (y-1)

表3 評価実験1の結果

(原画像と提案画像を比較、画像総数 460 枚)

写真	悪くなった	変わらない	良くなった
Building	21.3%	33.9%	44.9%
Car	25.8%	22.7%	51.5%
Flower	8.5%	40.2%	51.3%
Forest	14.2%	38.8%	46.0%
Human	10.3%	43.1%	46.6%
Mountain	17.1%	26.5%	56.4%
Nature	8.2%	45.9%	45.9%
River	13.3%	33.3%	53.3%
Sea	17.8%	29.9%	52.3%
Snap	22.1%	31.7%	46.2%
Train	26.9%	23.9%	49.3%
平均	17.1%	33.5%	49.5%

枚ずつの画像を提示した。評価方法としてはThurstone[36]の提案した2枚の写真ペアを提示する方法を用いた。評価実験の手順は次の通りである。ユーザに原画像と提案法による修正画像を並べて提示し

- 悪くなった- 変わらない
- 良くなった

で評価してもらった。被験者は男女計11名である。また実験で用いたパラメータの値は表4の通りである。これらの値は予備実験をもとに経験的に求めたものである。

#### (3) 実験結果

表3にユーザ評価の平均値結果を示す。「悪くなった」が17.1%であるのに対し、「良くなった」は約3倍の49.5%であり、また「変わらない」と「良くなった」を合わせた割合は83%となっている。

図8～図10に出力例を示す。左が原画像、中央が提案システムによる出力画像、右が切り出した部分を網掛けにした画像である。

表4 実験で用いたパラメータ

顔検出パラメータ	
肌色による顔検出閾値: $T_{skin}$	0.6
際立つ領域パラメータ	
平滑化ガウスフィルタサイズ: $H_{dst,bur}$	5
平滑化ガウスフィルタ標準偏差: $\sigma_{dst,bur}$	2
kmeans クラスタ数 $k_{dst}$	3
背景削除閾値 1: $T_{dst,bg1}$	1/3
背景削除閾値 2: $T_{dst,bg2}$	1/50
小さい領域削除閾値: $T_{dst,small}$	1/100
三角部分検出パラメータ	
平滑化ガウスフィルタサイズ: $H_{tri,bur}$	5
平滑化ガウスフィルタ標準偏差: $\sigma_{tri,bur}$	2
エッジ検出 LoG フィルタサイズ: $H_{tri,edge}$	13
エッジ検出 LoG フィルタ標準偏差: $\sigma_{tri,edge}$	2
拡大化ガウスフィルタサイズ: $H_{tri,thick}$	5
拡大化ガウスフィルタ標準偏差: $\sigma_{tri,thick}$	0.5
最少テンプレートサイズ: $S_{tri}$	1/16
正規化スコア閾値: $T_{tri}$	0.6
水平線検出パラメータ	
ラプラシアンフィルタサイズ: $H_{horicut}$	3
ラプラシアンフィルタ 形状調節パラメータ: $O_{horicut}$	0.2
水平線長さ閾値: $T_{horilen}$	1/2
水平線角度閾値: $T_{horilang}$	$\pi/18$
対角線・遠近法消失点検出パラメータ	
テンプレート最小角度: $D_{line}$	$\pi/10$
正規化スコア閾値: $T_{line}$	0.7
出力構図削除パラメータ	
サイズ閾値: $T_{cmp}$	1/2
顔領域はみ出し判断パラメータ	
顔領域拡大倍率: $E_{cmp,face}$	9
足切断制限倍率: $E_{cmp,foot}$	5
三角構図と水平線構図の判断パラメータ	
頂点から水平線への距離閾値: $T_{tri,hor}$	1/3



図8 Flower 出力例

(左:原画像 中央:提案手法画像 右:切り出し部分)



図9 Sea 出力例

(左:原画像 中央:提案手法画像 右:切り出し部分)



図10 River 出力例

(左:原画像 中央:提案手法画像 右:切り出し部分)



図11 Car 出力例

(左:既存手法[26]画像 右:提案手法画像)



図12 Forest 出力例

(左:既存手法[26]画像 右:提案画像)



図13 Mountain 失敗出力例

(左:既存手法[26]画像 右:提案画像)

#### (4) 考察

どのカテゴリにおいても「良くなった」と判断されたものが「悪くなった」より多くなっているが、カテゴリ別に見ると、特にFlower, Human, Nature, Riverは「悪くなった」が10%前後と少なく、「良くなった」は「悪くなった」の4倍から6倍となっている。構図を考慮した点以外の理由として考えられるものを以下で述べる。

Flowerについては際立つ領域検出, Humanについては顔検出によって検出される処理があること, またそれらが不自然に切れることがあり得る構図案を削除する処理があるため, 「悪くなった」が少なくなり, 「良くなった」が多くなる結果になったと考えられる。Natureについては特定の物体が画面に大きく写っているということが少なく, それらが不自然に切れることがなかったという理由が考えられる。Riverについてはどこかが切れてしまったら不自然に感じるという被写体が写っていることが少なかったこと, また画像として平坦な部分が多く, 遠近

表4 評価実験2の結果  
(既存手法と提案手法の出力の好ましい方を選択 画像総数 460 枚)

写真	既存手法[26]	同じ位	提案手法
Building	33.3%	19.3%	47.3%
Car	25.0%	23.5%	51.5%
Flower	24.8%	37.0%	38.2%
Forest	19.7%	46.0%	34.3%
Human	18.8%	23.6%	57.6%
Mountain	30.5%	31.9%	37.6%
Nature	19.7%	40.2%	40.2%
River	15.8%	44.6%	39.6%
Sea	16.7%	39.1%	44.2%
Snap	21.1%	41.4%	37.6%
Train	20.0%	18.7%	61.3%
平均	22.4%	32.8%	44.8%

法消失点や水平線が検出しやすかったことが理由として挙げられる。

逆にBuilding, Car, Trainは「悪くなった」が比較的高い値になった。これは、顔以外の特定物体を検出できる処理としては際立つ領域検出しか備えていないため、周囲もそれ自体にも様々な色を有するこれらをうまく検出できなかったこと、また画像に大きく写されている物体が多く、画像を切り出すとこれらの物体が切れてしまうことが多かったためと考えられる。

各カテゴリを通して、主要な物体や目立つ部分が写っている画像は修正後の変化が大きく感じられ、全体的に平坦な画像や同じ物が繰り返し写り込んだ画像は修正後の変化が少なく感じられるようであった。これは主観評価において変化を感じるためには目印となるものが必要であるためと考えられる。

### 3.2 実験2

#### (1) 目的

既存手法に対する有効性を評価するために、既存手法による修正画像と提案システムによる修正画像のどちらがユーザにとって好ましい画像かを評価した。

#### (2) 実験方法

既存手法としてNishiyamaら[26]の自動切り出しシステムを用いた。画像は評価実験1と同様のDBを用いた。ユーザに各カテゴリにつき評価実験1とは異なる15枚ずつを既存手法画像と提案法画像を並べて提示し、どちらが好ましいかを評価してもらった。被験者は同様に男女計11名である。実験で用いたパラメータは実験1と同様である。

#### (3) 実験結果

表4にユーザ評価の平均値結果を示す。提案法の出力は従来法の2倍のユーザ支持を受けたことが分かる。図11～図13に提示した既存手法と提案手法の出力例を示す。いずれの図も左が既存手法、右が提案手法の出力となっている。

#### (4) 考察

いずれのカテゴリにおいても提案手法の方が好ましいと判断された画像が多かった。特にCar, Human, Nature, River, Sea, Trainについては既存手法と比べ提案手法が約2倍から3倍の支持が得られた。

CarとTrainについては既存手法に特定物体を検出する機構が

無く、図11のようにこれらの物体が不自然に切れてしまったこと、また提案手法で構図案の削除と切り出し制限という印象を保持するための処理があるため、既存手法に比べて無理な切り出しが少なかったことが原因と考えられる。Humanについては既存手法には顔検出が無いが、提案手法には顔検出や、その後の顔、足が不自然に切り出されないような処理があることが理由と考えられる。Nature, River, Seaについては、構図を用いた提案手法と、構図を考慮していない既存手法の差が特に出たのではないかと考えられる。

一方、Building, Flower, Forest, Mountainは既存手法と提案手法で差が少なくなる結果となった。Buildingについては提案手法ではどこを主構成要素すべきかが定めにくく、構図を上手く適用できなかったためではないかと考えられる。Flowerについては原画像の段階で花が画像一杯に写されているものが多く、ユーザは差をつけにくかったためと考えられる。Forestについては、図12のような画像が多く、画面内に多くの木が写っているためどこを基準として良いかが分からず評価が割れたのではないかと考えられる。Mountainについては、提案手法で図13のようなものがいくつか出力されたことで既存手法が支持されたためだと考えられる。これは台形である富士山を始めとして、山は綺麗な三角形をしておらず、三角部分検出でうまく検出できなかったためと考えられる。それにより、他の構成要素、特に山裾に水平線があることが多いのでその構図に合わせた修正が行われてしまい、山の山頂が切り取られてしまい評価が低下したと考えられる。

### 4. まとめ

本論文では感性を反映した構図修正による、デジタル写真の品質向上システムを提案した。提案システムは一般的なサイズの撮影済みデジタル写真を対象として、その構図を自動修正することで写真の品質と表現力を向上させ、ユーザにとってより好ましい写真を出力するシステムである。

提案システムではまず、入力された画像から主となる構成要素が検出される。扱うことのできる構成要素は、顔、際立つ領域、三角部分、水平線、対角線、遠近法消失点の6種類である。検出された各構成要素の位置や他構成要素との関係から、それぞれの構成要素に適すると思われる出力構図案が計算される。この出力構図案では複数の構図を組み合わせて扱うことが可能である。続けて、その構図案の中から、原画像の印象から大きく変わってしまう案が削除される。残った出力構図案の中から原画像の印象を最も保つことが可能な構図案、すなわち切り出す面積が最も大きい構図案が選択される。これらの処理によって無理な切り出しが行われないように制限をかけることで、原画像の印象を変化させすぎない修正が可能となる。最後に、最終的な構図案に従い画像の構図修正が行なわれる。

2種類の主観評価実験を行った。この実験から、提案システムによって画像に対しユーザの好ましい修正が行われることが確認された。また既存システムと比べて、ユーザにとってより好ましい修正が行われることが確認された。

提案システムの目的は、まずは構図修正の自動化であった。特に文献[29]の手法を用いた顔検出の部分で多くの計算量を必要としているため、デジカメに直接インプリメントするのは現状では極めて困難である。今後は、アルゴリズムの各部分において計算量を削減し、デジカメにインプリメントできる程の量にしたい。

**謝辞** 評価実験の際に既存研究として比較させて頂いたシステムを考案された株式会社東芝の西山正志様には実験のご協力と貴重なアドバイスを頂戴しました。この場を借りて深くお礼申し上げます。

## 参考文献

- [1] flickr: <http://www.flickr.com/>
- [2] Picasa: <http://picasa.google.co.jp/>
- [3] Photo.net: <http://photo.net/>
- [4] Adobe PhotoShop:  
<http://www.adobe.com/jp/products/photoshop/family>
- [5] Corel Paint Shop Pro:  
<http://www.corel.com/servlet/Satellite/jp/jp/Product/1184951547051>
- [6] GIMP: <http://www.gimp.org/>
- [7] Picnik: <http://www.picnik.com/>
- [8] flauntR: <http://www.flaustr.com/>
- [9] FotoFlexer: <http://fotoflexer.com/>
- [10] 桜井始ほか: すぐに役立つ被写体別画面構成, 玄光社, 2008.
- [11] 山口高志: 風景写真被写体別構図の決定, 学習研究所, 2008.
- [12] 山口高志: 風景写真フレーミングの決定, 学習研究所, 2009.
- [13] 山口高志: 写真撮影とらの巻フレーミング事例辞典, 学習研究所, 2008
- [14] デジタルフォト編集部: 今さら人には聞けないデジタル一眼レフの疑問構図編, ソフトバンククリエイティブ, 2009.
- [15] 大和田良ら: 三分割構図法を完璧にマスターする. デジタルフォト2009年12月号, pp. 41-76, ソフトバンククリエイティブ.
- [16] 内田広由紀: 構図エッセンス, 視覚デザイン研究所, 2006.
- [17] 萩窪圭: デジカメ撮影のネタ帳シーン別ベストショットの撮り方, エムディエヌコーポレーション, 2009.
- [18] 西川晃太郎: キレイに見せるデジカメ写真術手軽に楽しむ撮影とフォトレタッチ, ソフトバンククリエイティブ, 2008.
- [19] 岡嶋和幸: デジタル一眼レフ初心者のためのレタッチの基本教えます, 玄光社, 2007.
- [20] S. Banerjee and B. L. Evans. In-camera automation of photographic composition rules. IEEE Trans. IP, 16(7), pp. 1807-1820, 2007.
- [21] Y. Chang and H. Chen, Finding Good Composition in Panoramic Scenes. in Proc. ICCV, 2009.
- [22] B. Suh, H. Ling, B. Bederson, and D. Jacobs. Automatic thumbnail cropping and its effectiveness. in Proc. ACM UIST, pp. 95-104, 2003.
- [23] A. Santella, M. Agrawala, D. DeCarlo, D. Salesin, and M. Cohen. Gaze-based interaction for semi-automatic photo cropping. in Proc. SIGCHI Conf. Human Factors in Computing Systems, pp. 771-780, 2006.
- [24] J. Luo. Subject content-based intelligent cropping of digital photos. in Proc. ICME, pp. 2218-2221, 2007.
- [25] S. Amrutha, S. S. Shylaja, S. Natarajan and K. N. Balasubramanya Murthy, A Smart Automatic Thumbnail Cropping Based on Attention Driven Regions of Interest Extraction. in Proc. ACM ICPS, pp. 957-962, 2009.
- [26] M. Nishiyama, T. Okabe, Y. Sato and I. Sato, Sensation-based Photo Cropping. in Proc. ACM MM, 2009.
- [27] L. Itti, C. Koch, and E. Niebur. A model of saliency-based visual attention for rapid scene analysis. IEEE Trans. PAMI, 20(11):1254-1259, 1998.
- [28] L. Itti and C. Koch, A comparison of feature combination strategies for saliency-based visual attention systems. In Human Vision and Electronic Imaging, vol. 3644, pp. 473-482, 1999.
- [29] M. Nilsson, J. Nordberg, and I. Claesson, Face Detection using Local SMQT Features and Split up Snow Classifier. in Proc. ICASSP, vol. 2, pp. 589-592, 2007
- [30] J. Sherrah and S. Gong. Skin Color Analysis.  
[http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL\\_COPIES/GONG1/cvOnline-skinColourAnalysis.html](http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/GONG1/cvOnline-skinColourAnalysis.html), 2001
- [31] Lab 色空間:  
[http://ja.wikipedia.org/wiki/Lab\\_色空間](http://ja.wikipedia.org/wiki/Lab_色空間)
- [32] L\*a\*b\*表色系: <http://www.colordream.net/lab.htm>
- [33] D. Marr and E. Hildreth. Theory of edge detection. Proc. R. Soc. Load. B, 207:187-217, 1980
- [34] 画像共有サイト「フォト蔵」: <http://photozou.jp/>
- [35] Google 画像検索: <http://images.google.co.jp/>
- [36] L. Thurstone. A law of comparative judgment. Psychological Review, 34:273-286, 1927.
- [37] 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,” 第9回NICOGRAPH春季大会論文 & アート部門コンテスト, III-3, pp.1-8, 2010-03.

## 家田暁



2008年慶應義塾大学理工学部情報工学科卒業。2010年慶應義塾大学大学院理工学研究科修士課程修了。同年(株)野村総合研究所入社。感性工学に興味を持つ。

## 琴智秀



2000年慶熙大学校大学院修士課程修了。2008年慶熙大学校大学院博士課程修了。工学博士。2008年より慶應義塾大学理工学部訪問研究員。画像処理, 音声処理, 感性工学に興味を持つ。2009年International Symposium on Advanced Intelligent Systems (ISIS 2009)にてBest paper award受賞。

## 萩原将文



1982年慶應義塾大学工学部電気工学科卒。1987年同大学院博士課程修了。工博。同年同大助手。以来、ニューラルネットワーク, ファジィシステム, 進化計算, 感性工学の研究に従事。現在, 同大教授。1991-92年度スタンフォード大学訪問研究員。1990年IEEE Consumer Electronics Society論文賞, 1996年日本ファジィ学会著述賞, 2003年日本感性工学会技術賞, 2004年同学会論文賞, 2005年日本知能情報ファジィ学会貢献賞受賞。IEEEシニアメンバ。