

# バランスのとれた代表写真群を自動選出するための一手法

塩谷祥加<sup>†</sup> 森下奈保子<sup>†</sup> 伊藤貴之<sup>†</sup> 萩田真理子<sup>†</sup>

近年誰もが手軽に多くの写真を撮影し保存することが可能となった一方、所有する写真データがあまりにも大量すぎて整理や共有することに大変な労力を伴うこともある。そこで本研究では、異なる特徴量を持つ写真を同程度選出することにより大量の写真群からバランスのとれた代表写真群を自動選出し提示する手法を提案する。さらに我々の実装では代表写真群の選出結果を表示するユーザインタフェースも提供している。

## The method for automatic selecting a balanced set of photos

HIROKA SHIOYA<sup>†</sup> NAOKO MORISHITA<sup>†</sup>  
TAKAYUKI ITO<sup>†</sup> MARIKO HAGITA<sup>†</sup>

Recently everyone can easily take photos and retain. But people need time and labor to organize and share photo data if they have too many. In this Paper, we show a method for automatic selecting a balance set of photos from many photo data by selecting same extends photos with different feature values. And We provide a user interface that shows result of selecting photos.

### 1. はじめに

デジタルカメラやスマートフォンの発達に伴い、誰もが多くの写真を気軽に撮影し保存することが可能となった。写真を Web 上で共有する機会も増えているが、大量の写真データを全て共有する場合は、アップロード時間とネットワーク負荷がかかる上にサービスによっては容量制限などを伴う場合もあり困難を伴う。また大量の写真を全て公開しても閲覧者は全て閲覧するのに非常に多くの時間を必要とする。とはいえ保存した大量の写真データを整理する際は、大量の写真を1枚ずつ目視確認して好ましい写真とそうでない写真とに分類する必要がある、非合理的な作業を行わなくてはならない。そこで適切な枚数の代表写真群を自動選別するシステムがあれば、写真を楽しむユーザにとって写真の取捨選択が便利になると考えた。

本報告では大量の写真群から多彩な組み合わせの代表写真群を自動選出するための手法を提案する。また我々の実装では代表写真群の選出結果を表示するユーザインタフェースも提供しており、この画面上で対話的に「選出すべき写真」「選出すべきでない写真」を指定した上で代表写真群の選出を反復することで、ユーザの嗜好を反映して代表写真群の選出結果を調整することもできる。写真の自動選出手法に関して、まず各写真の特徴量を自動取得してそこから評価値を算出し、著しく評価値が低い写真群を処理から省く。残りの写真群から異なる特徴量を有する写真群をバランスよく選出する。本報告では2種類の代表写真群選出手法を提案する。ひとつは行列式を用いた選出手法である。

特徴量のバランスのよさを「多次元空間にて特徴量ベクトルによって形成される多面体の体積」と定義する。そしてユーザ指定の枚数に応じて仮の代表写真群を選出し、その写真群の特徴量ベクトルが成す多面体の体積を計算する。この処理を反復した結果として体積が最大となる写真群の組み合わせを代表写真群とする。もうひとつはグラフ彩色を応用した選出手法である。この手法においては「どの写真ペアをとってきても類似度が低い写真群」がバランスのよい写真群であると定義する。はじめに自動算出した評価値を用いて、評価値合計順に写真を並べ番号を振り、全ての写真のペアが張る面積を計算する。写真を点として、類似度が一定値以上のときにそのペアを辺で結びグラフを生成する。そしてグラフを彩色し、色1グループから写真番号が小さい順にユーザ指定の枚数を選出し代表写真群とする。さらにこれら2種類の手法を用いて結果を比較し考察を行う。

本稿の構成は以下の通りである。2章では関連研究について述べる。続いて3章では提案手法を述べる。4章では本手法の実行結果について述べる。5章では本研究で得られた成果のまとめと今後の展望について述べる。

### 2. 関連研究

写真評価と写真自動選出に関する研究は既に数多く発表されている。写真評価の手法には、写真の色や構図、ピントのぼけ具合といった一般的な特徴量をベースとした手法[1][2][3][4][5][6]が旧来から多く発表されている。これらの手法の多くは、色が鮮やかな写真、被写体が中央に位置している写真、被写体にピントが合っている写真などを高く

<sup>†</sup> お茶の水女子大学  
Ochanomizu University

評価する。これらの手法に加えて最近では、被写体となる各人物の表情や人数、人物間の写真上の位置などを評価要素として加えた手法[7][8][9][10]も提案されている。これらの手法の多くは、人物の表情が良い、複数の人が写っている、人物どうしの距離が近く仲が良さそうな写真などを高く評価する。

また写真評価結果にもとづいて写真に点数を付け、点数が高い順に写真を自動選出する研究[11][12][13][14]も発表されている。また自動選出された写真をアルバムとして編集表示する商用サービス[15]もある。

### 3. 提案手法

関連研究で紹介した写真評価手法に沿って評価が高い写真を選出した場合、撮影状況によっては同じ人物や風景のみが選出されてしまう可能性が高い。そこで我々は、バランスのよい多彩な組み合わせの代表写真群を自動選出する手法を開発している。具体的には、同じ被写体の写真が多数選出されることを防ぎ、さまざまな人物や風景、場面の写真からそれぞれバランスよい枚数で写真を選び出すことを目指す。

本手法では、写真群を大きく分けて以下の3つのグループに分類する。

- 第1群: 好ましい組み合わせの写真群
- 第2群: 第1群にも第3群にも含まれない写真群
- 第3群: 評価が低い写真群

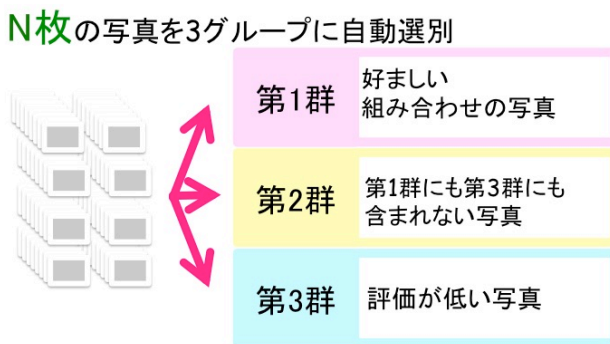


図 1 写真の自動選別  
Figure 1 Automatic selecting of photos

#### 3.1 写真の特徴量と評価値の計算手法

バランスよくさまざまな写真を選ぶための基準の例として、我々は以下のような基準を考えた。

- ・ 写りの悪い写真（例えば暗すぎる、ピントが合っていないなど）は代表写真に選ばない。
- ・ 一人でも多くの人物が代表写真群の最低1枚に、いい表情をして写っているように代表写真を選ぶ。

- ・ 集合写真がある場合には、最低1枚の集合写真を代表写真に選ぶ。
- ・ 多くの場所を撮影している場合には、できるだけその各所の写真を代表写真に選ぶ。

これらの基準に沿って代表写真を選ぶために、現時点での我々の実装では、「人物の数」「人物の表情」「物体認識の情報」「ピント」「鮮やかさ」「撮影日時」の特徴量を自動算出し、それぞれ[0,1]の範囲に正規化する。以上の特徴量をベクトルとしてその長さを算出した際に、長さが著しく短い写真群を第3群に自動分類する。それとは別に「人物識別情報」も特徴量として扱う。これについては実装上の都合により現時点では手動で付与している。

##### 3.1.1 人物の数

我々の実装では画像処理ライブラリ OpenCV の顔検出機能を用いて矩形の顔領域を抽出し、その顔領域数によって人物の数を推定して特徴量とする。現状では特徴量の値を以下の通りとしている。

- 検出された最大人数の 80%以上 : 1
- 検出された最大人数の 60%以上 : 0.8
- それ以下 : 人数 × 0.1

これにより、多くの人物の顔が写っている写真は優先的に代表写真として選出されやすくなる。

##### 3.1.2 人物の表情

我々の実装では Google の画像認識 API である Cloud Vision[16]を用いることで表情の情報を取得している。現状ではその中で笑顔認識情報を採用し、特徴量を以下の通りとしている。

- VELY\_LIKELY : 1
- LIKELY : 0.8
- UNLIKELY : 0.4
- VELY\_UNLIKELY : 0.2
- 取得不可 : 0

なお我々の実装では、3.1.1 項で判別した人物ごとに各写真における特徴量を算出しその平均値を用いる。これにより、同一人物が写る写真の中では笑顔となっている写真が優先的に採用される。

##### 3.1.3 物体認識の情報

我々の実装では、Google の画像認識 API である Cloud Vision[16]を用いることで物体認識の情報を取得している。認識した物体のラベルの数だけ次元を設定し、各写真においてそのラベルに対応する物体を認識した場合には”1”と

し、認識しなかった場合には”0”とする。

### 3.1.4 ピント

我々の実装では、OpenCV を用いて人物の顔がどの程度はっきり写っているのかを計算する。検出された顔領域に対してエッジ検出を適用し、エッジと判断された画素の割合を計算する。ピントが合っていないほど値は小さくなる。ピント特徴量の算出は以下の通りである。

ピント特徴量

$$= \frac{\text{エッジ検出されたピクセル数/}}{\text{顔領域のピクセル数} \times 10.0}$$

この値が 1.0 以上の場合は 1.0、閾値（現状では 0.1）以下の場合は 0.0 とする。これにより、人物がぼやけていて鮮明でない写真は好ましくない写真と判定される。

### 3.1.5 鮮やかさ

我々の実装では、各ピクセルの RGB 値を HSV 値に変換をする。そして彩度（S 値）の写真全体での平均値の 0.01 倍を特徴量として求める。なおこの値が 1.0 以上の場合は 1.0 とする。これにより、色が鮮やかなほど、つまり S の平均値が高いほど評価が高く好ましい写真となる。

### 3.1.6 撮影日時

我々の実装では、写真の exif 情報を取得し、そこから撮影日時の情報を自動取得する。特徴量はヒストグラムを用いることで算出する。ヒストグラムの区間幅は

$$(\text{最初の写真の時刻}-\text{最後の写真の時刻})/(\text{区間数})$$

とする。現状の実装では暫定的に区間数を 10 で固定している。そして度数が大きい区間を重要な時間帯であるとし、下記のように特徴量を求める。

$$\text{特徴量} = \text{度数} / \text{度数の最大値}$$

これにより、より多くの写真が撮影された時間帯に含まれる写真が優先的に選出される。

### 3.1.7 人物識別

写真群に写っている人物の総人数ごとの次元を作成し、各写真においてその人物が写っている場合には”1”とし、写っていない場合は”0”とする。現時点では実装上の都合により、この特徴量手動で付与しているが、将来的には各写真から抽出される顔領域を人物ごとに自動識別する前提で実装を進めている。

## 3.2 好ましい組み合わせ写真群の選出手法

本手法では、M 次元の特徴量が与えられた N 枚の写真の中から、バランスよい組み合わせの N' 枚を選出することを想定する。バランスがよい組み合わせの写真群とはつまり以下の 2 条件

- ・ 評価が高い写真を多く選ぶ
- ・ 特徴が異なる写真を多く選ぶ

をできるだけ満たすように写真を選出することに相当する。ここでは「行列式を用いた手法」と「グラフ彩色を応用した手法」の 2 種類の手法を用いる。

### 3.2.1 行列式を用いた手法

1 枚の写真を 1 本の特徴量ベクトルで表すと、特徴量ベクトルの長さは評価値、向きは特徴量の種類となる。図 2 は代表写真の選び方に関する概念を模式化したものである。図 2(左)のように、短くてなす角も小さい特徴量ベクトル群に対応する写真群が選出された場合、評価が低くかつ似ている写真が多く選出されてしまう。逆に図 2(右)のように、長くてなす角の大きい特徴量ベクトル群に対応する写真群が選出されれば、評価が高く多彩な写真が多く選ばれると考えられる。

図 2(右)のような特徴量ベクトル群を選ぶための手段として本手法では、N' 本の特徴量ベクトルで構成される多次元空間での多面体の体積を計算して、体積を最大にする特徴量ベクトル群を選出し、これに対応する写真群を代表写真群として第 1 群に分類する。

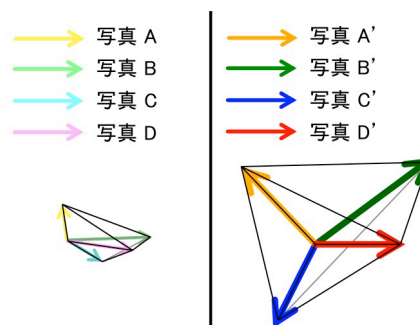


図 2 写真の特徴量をベクトル化

Figure 2 Convert feature values of photos into vectors

N' 本の特徴量ベクトルで構成される多面体の体積を求めるにあたり、特徴量ベクトルが N' 次元であれば、N' × N' 次元の行列を形成してその行列式を求めればよい。そこで M > N' である場合には、特徴量ベクトルを M 次元から N' 次元に次元削減する。現時点での我々の実装では、代表写真の枚数 N' がユーザによって指定された時点で、以下の 2 段階処理

を適用することで各写真の特徴量である  $M$  次元ベクトルを  $N'$  次元ベクトルに変換する。

- 1) まず、 $M$  次元ベクトルにおいて各次元ペア間での Pearson 相関係数を求め、これが一定値以上となる特徴量ベクトルを 1 本に結合する。その結果、 $M$  次元ベクトルが  $N''$  次元ベクトル ( $M > N'' > N'$ ) に削減される。
- 2) 上述の  $N''$  次元ベクトルに対して主成分分析を適用し、上位  $N'$  個の主成分を採用することで、 $N'$  次元ベクトルに変換する。

続いて以下の手順によって何通りかの写真群に対して体積を計算する処理を反復し、体積が最大となる写真群を代表写真群として第 1 群に分類する。

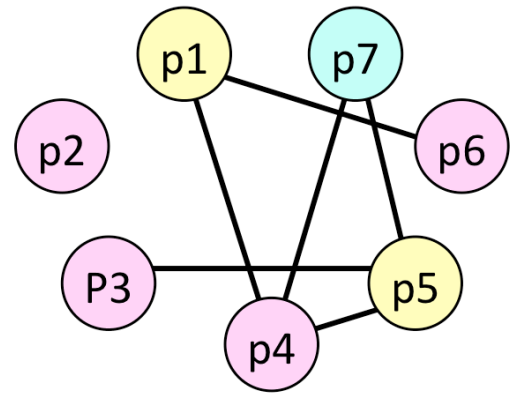
1. ユーザが代表写真の枚数  $N'$  を指定した時点で、全ての写真の特徴量を  $N'$  次元ベクトルに変換する。
2.  $N'$  枚の代表写真を仮に選出して  $N' \times N'$  次元の行列を構成し、掃き出し法によって行列式を算出することで体積を求める。
3. この処理を一定回数反復し、体積が最大となる代表写真群を正式に代表写真として第 1 群に分類する。

現在の我々の実装では、2.においてランダムに  $N'$  枚の代表写真を選出している。将来的には、ここに遺伝的アルゴリズムなどの進化計算手法を適用することで、限られた反復回数の中で最適に近い代表写真選出を実現できると考えている。

### 3.2.2 グラフ彩色を応用した写真選択

ここでは 2 本の特徴量ベクトルがなす面積の大きさを、写真ペアの類似度とする。そうすると面積が大きくなる写真ペアの類似度は低くなり、異なる特徴量を持つといえる。つまりバランスが良い多彩な組み合わせの写真群とは、どの写真ペアをとっても類似度が低いということである。写真選出の手順は以下の通りであり、グラフを彩色した図を図 3 に示す。

- 1) 写真を評価値合計値順に並べ、番号を振る ( $p_1, p_2, \dots$ )。
- 2) 写真すべてのペアについての類似度を計算する。
- 3) 写真を点とし、類似度が一定値以上のときにそのペアを辺で結びグラフを生成。
- 4) グラフを彩色する。
- 5) 色 1 グループから写真番号が小さい順に選出。



●... $p_2, p_3, p_4, p_6 \rightarrow$  3 個選出 ( $p_2, p_3, p_6$ )

図 3 グラフの彩色

Figure 3 Coloring of graph

### 3.3 ユーザによる手動選別

以上の処理によって全自動で写真を第 1 群から第 3 群に振り分けることも可能である。しかし、この結果に全てのユーザが満足するとは限らない。ゆえに、ユーザ自身による対話操作で代表写真選出結果を修正できるように、写真を表示するユーザインタフェースを開発した。図 4 にその概観を示す。この画面上でユーザは、移動する写真と固定する写真を選択することが可能である。例えば、代表写真に選出したくない写真が第 1 群に分類されている場合は、それを第 2 群に移動させることができる。また逆に、代表写真に選出したい写真が第 2 群に分類されている場合は、これを第 1 群に固定することが可能である。ユーザが写真の移動操作を終えるたびに、「行列式を用いた手法」においては、3.2.1 節で示した計算を再実行し、バランスよい組み合わせの代表写真群が再選出される。また「グラフ彩色を応用した手法」においては、ユーザが指定した写真と対応する点を色 1 で彩色した後に残りの点を彩色する。そして色 1 グループからユーザが指定した写真を含む  $N'$  枚が再選出される。

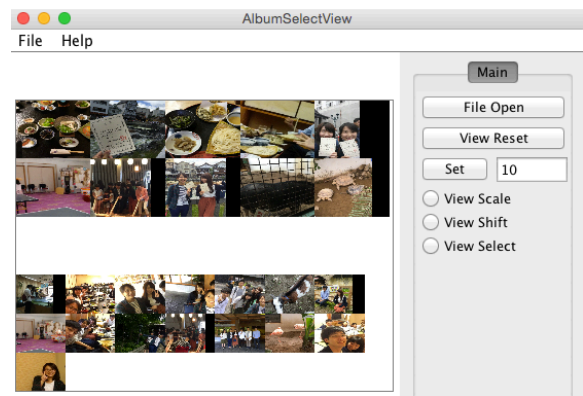


図 4 手動選択のためのユーザインタフェース

Figure 4 User interface for manual selecting

## 4. 実行結果

2種類の手法を用いて旅行写真25枚から5枚を選出した。用いた25枚の写真のうち約2/3が人物にフォーカスした写真であり、それ以外は観光名所や食べ物などにフォーカスした写真であった。

行列式を利用した選出手法による結果では、5枚の写真のうち集合写真が1枚、2人から4人の人物や観光名所が写った写真が3枚、食べ物の写真が1枚となり、比較的バランスがよい結果となった。しかし実行時間に45.8秒かかった。これは計算時間が以下のようになるからである。

・写真の選び方： $O(n^m)$

・掃き出し方による行列式計算： $O(m^3)$

グラフ彩色を応用した選出手法による結果では、5枚全て参加者全員が写った写真となった。観光名所を表すような写真が1枚も含まれず、バランスがよい選出結果とはいえなかった。この原因として、特徴量が0か1かの2値となっている人物識別の特徴量が強く反映されてしまったことが考えられる。一方で実行時間は5.9秒であり高速であった。これは計算時間が以下のようになるからである。

写真の選び方： $O(n)$

掃き出し方による行列式計算： $O(n^2)$

## 5. まとめと今後の課題

本報告では、大量写真群からバランスのとれた代表写真群を自動選出して提示する手法を提案した。本手法では写真を第1群から第3群に自動分類する。まず特徴量の自動算出により写真を評価し、評価値が小さい写真を好ましくない写真として第3群に選別する。そして行列式を利用した手法とグラフ彩色を応用した手法を用いて第1群の写真を選出し残った写真を第2群とする。さらに本研究では、ユーザの対話操作によって第1群と第2群の写真を入れ替えるユーザインタフェースを搭載している。対話操作結果により第1群に属する代表写真群を再度自動選出することで、ユーザの嗜好を反映した代表写真選出を可能とする。

今後の課題として、人物識別と物体認識の情報に関する特徴量算出を改善したい。現状では0または1の2値としているために、特にグラフ彩色を応用した手法を用いた際に、多くの人物や物体が写っている写真の優先度が高すぎる傾向が現れてしまっている。対策として、人物識別と物体認識に関する次元のみを切り取って次元削減を適用してから、3.2節に示した写真選出方法を適用する、と言った方法を試してみたい。

現在我々は提案手法のユーザテストのための準備を進めている。具体的には特定のグループ（現状では我々の所属研究室）の構成員が参加するイベントの写真を多数集め、

その中から代表写真を選出する、というタスクを想定している。このタスクにおいて、ユーザテストに参加する各被験者に自動算出結果を提示し、ユーザインタフェース上で各被験者に写真を手動選別させる。この手動選別操作の回数と内容を記録してデータ化し、平均としてどれくらいの回数が必要となるかを比較したい。そしてその回数が少なくなるように特徴量の算出方法等を見直したい。4章の実行例では写真の枚数が少なかったこともあり、手動選別の回数はせいぜい1回か2回程度であった。そこで現在準備しているユーザテストでは数百枚単位の写真群を用いたタスクを想定している。

## 参考文献

- 1) Ritendra Datta, Dhiraj Joshi, Jia Li, James Z. Wang: Studying Aesthetics in Photographic Images Using a Computational Approach, European Conference on Computer Vision, pp. 288-301 (2006).
- 2) Yan Ke, Xiaoou Tang, Feng Jing: The Design of High-Level Features for Photo Quality Assessment, Conference on Computer Vision and Pattern Recognition, pp. 419-426 (2006).
- 3) Ritendra Datta, Jia Li, James Z. Wang: Learning the Consensus on Visual Quality for Next-Generation Image Management, ACM Multimedia Conference, pp. 533-536 (2007).
- 4) Yiwen Luo and Xiaoou Tang: Photo and Video Quality Evaluation: Focusing on the Subject, European Conference on Computer Vision, pp. 386-399 (2008).
- 5) Congcong Li and Tsuhan Chen: Aesthetic Visual Quality Assessment of Paintings, Sel. Top. In Sig. Proc., IEEE Journal of, vol. 3, no. 2, pp. 236-252 (2009).
- 6) Hsiao-Hang Su, Tse-Wei Chen, Chieh-Chi Kao, Winston H. Hsu, Shao-Yi Chien: Scenic Photo Quality Assessment with Bag of Aesthetics-Preserving Features, ACM international conference on Multimedia, pp.1213-1216 (2011).
- 7) Congcong Li, Andrew Gallagher, Tsuhan Chen: Aesthetic quality assessment of consumer photos with faces, Image Processing, pp. 3221-3224 (2010).
- 8) Shehroz S Khan and Daniel Vogel: Evaluating visual aesthetics in photographic portraiture, Computational Aesthetics in Graphics, Visualization, and Imaging, pp. 1-8 (2012).
- 9) Shao-Fu Xue, Henry Tang, Dan Tretter, Qian Lin, Jan Allebach: Feature Design for Aesthetic Inference on Photos with Faces, Image Processing, pp. 2689-2693 (2013).
- 10) Yufei Wang, Zhe Lin: Event-Specific Image Importance, The IEEE Conference on Computer Vision and Pattern Recognition, pp.4810-4819 (2016).
- 11) Congcong Li, Alexander C. Loui, Tsuhan Chen: Towards Aesthetics: a Photo Quality Assessment and Photo Selection System, Association for Computing Machinery's annual conference on multimedia, pp. 827-830 (2010).
- 12) Lei Huang, Tian Xia, Ji Wan, Yongdong Zhang, Shouxun Lin: Personalized portraits ranking, ACM Multimedia Conference, pp. 1277-1280 (2011).
- 13) Oliver Le Meur, Thierry Baccino: Prediction of the Inter-Observer Visual Congruency (IOCV) and Application to Image Ranking, ACM Multimedia Conference, pp.373-382 (2011).
- 14) Che-Hua Yeh, Yuan-Chen Ho: Personalized Photograph Ranking and Selection System, ACM Multimedia Conference, pp.211-220 (2010).
- 15) 富士フイルム, Year Album, <http://year-album.jp>
- 16) GoogleCloudVision, <https://cloud.google.com/vision/>