

人物の共起性を代表画像選出に反映した個人写真ブラウザ

安田理紗 (お茶の水女子大学), 五味愛(お茶の水女子大学), 伊藤貴之(お茶の水女子大学)

Personal Photograph Browser Featuring Human Co-occurrence Based Technique

Risa YASUDA Ai GOMI and Takayuki ITOH

ABSTRACT

Recently, photograph browsing techniques became an active research topic. This paper proposes a visualization technique for browsing of large photograph collections arranged by co-occurrence of humans. As a preprocessing, the technique applies hierarchical clustering to the images based on their timestamps, and applies a face recognition technique to list the humans taken in the photographs. It then constructs a network of the clusters by connecting according to the co-occurrence of the face recognition result. The technique places these clusters by a hybrid space-filling and force-directed layout algorithm. Placing photographs which have common faces close, we expect that we can understand communities of the photograph owners intuitively.

Keywords: Hierarchical data visualization, Photograph browser, Force-directed layout, Space-filling layout

1. はじめに

近年のデジタルカメラの普及に伴い、個人が所有する写真のデジタルデータは膨大な量となっている。そのような大量の写真の管理はユーザの大きな負担となり、写真の自動分類は重要な課題である。一方、138人の学生を対象にアンケート調査を実施したところ、主な被写体は何かという問いに対し、人物と回答した人は、デジタルカメラを所有している人の85%にのぼった。よって、個人写真の被写体として、人物は重要な要素であるといえる。

以上を踏まえ我々は、被写体としての人物に基づいて写真を自動分類し、一緒に写ることの多い人物の写真が画面絵で近くに配置される写真ブラウザの一手法を提案する。本手法では、まず写真群を撮影日時に基づいてクラスタリングし、クラスタ毎に代表写真を選出する。多くの場合において、このクラスタリングは撮影者のイベントを自動検出することに相当する。続いて各クラスタ中の写真群に写る人物の共起性に基づいて、各クラスタの画面上の配置を決定し、代表写真を表示する。これにより、ユーザは自身が持つ写真群の全貌を容易に把握でき、興味の対象となる写真を短時間で発見できる。また共起性の高い人物が写る写真が近くに配置されることで、配置結果にはコミュニティが反映される。一般的に写真は時系列に沿って閲覧されることが多いのに対して、本手法を用いた写真閲覧はコミュニティに沿った新しい写真閲覧をもたらす、思いがけない気づきにつながると期待できる。

2. 関連研究

2.1 商用の個人写真ブラウザ

個人写真を管理・閲覧するために商用ソフトウェアとして、Picasa(Google) や iPhoto(Apple) が有名である。これらは撮影場所・日時・人物などのメタ情報から写真を管理・閲覧する機能をサポートしている。しかし、メタ情報の人物に着目してみると、特定の人物が写る写真の検索機能は優れているが、写真を一覧表示する際に人物情報が活用されているとは言い難い。

2.2 CAT

CAT [1]とは、大量画像のクラスタリング手法、およびその一覧表示と詳細度制御を持ち合わせた可視化手法である。前処理として、画像に対してキーワード、画素情報に基づく2段階のクラスタリングを適用し、階層型データを構築する。そして、各クラスタから代表画像を選出する。続いて、この階層型データに、階層型データ可視化手法「平安京ビュー」[2]の配置アルゴリズムを適用することで、画像群を一覧可視化する。このアルゴリズムによって、以下の2条件

【配置条件1】 画像どうし、クラスタどうしの遮蔽回避

【配置条件2】 画像群の配置占有面積の低減

を満たす画面配置を実現する。

CATではズーム率に合わせた詳細度制御を設けており、ズームアウト時は各クラスタの代表画像を、ズームイン時には局所的に各画像を表示する。

2.3 FRUTIS Net

本手法で適用するネットワーク可視化手法FRUTIS Net

(FRamework and User Interface for Tangled Segments Network) [3]は、1個以上のアイテムが各ノードに付加されたネットワークを対象とした可視化手法である。FRUITS Netではノードをアイテム毎に色分けし、ノード間の連結をエッジで表す。そして、力学モデルと空間充填モデルを併用した配置アルゴリズムによりノード配置を決定することにより、複数のアイテム情報を有するネットワークの全体像を一画面上で表すことが可能となる。FRUITS Netでは上述の[配置条件1][配置条件2]に加えて、

[配置条件3] 共通するアイテムを有するノード群を近くに配置

[配置条件4] エッジ長の総計、エッジ交差数の低減も同時に満たす画面配置を実現する。

我々は FRUITS Net の画面配置アルゴリズムを画像ブラウザに適用した手法を発表している[4]。この手法では各画像に1個以上のキーワードが付与されていることを前提として、そのキーワードに基づいて画像をクラスタリングし、キーワードの共有度が高い画像クラスタが画面上で隣接するような画面配置を算出する。

本報告の提案手法はこの手法から派生したものであり、

- 顔認識結果に基づく人物の共起性に基づいて各クラスタを画面配置する
- 人物の出現頻度に基づいて各クラスタの代表画像を選出する

という点がこの手法に対する差異である。

3. 提案手法

本手法で用いる各々の写真は Google Picasa APIs を用いて人物を識別し、各人物に ID を割り当て、この ID をメタ情報として保持しているものとする。

3.1 階層型データの構築

本手法では、撮影日時に基づいて写真群に2段階のクラスタリングを適用する。まず写真群を撮影日時順に並び替え、前後2枚の写真の撮影時刻の間隔が α 以上であれば新しいクラスタを生成する。生成されたクラスタを高階層クラスタとし、1高階層クラスタが1イベントを表す。現時点での我々の実装では $\alpha=24(\text{hours})$ としたが、ユーザに応じて調節可能である。続いて各高階層クラスタを構成する写真群を撮影日毎に区切ることで写真をさらに分類し、これを低階層クラスタとする。

3.2 代表写真の選出

続いて各クラスタに対して代表写真を自動選択する。この代表写真は、ズームアウト時に各クラスタを代表する写真として表示されるものである。

まず、各低階層クラスタについてキーパーソンを特定する。我々はキーパーソンを特定する方法として、自然言語処理の基本的な手法としてよく用いられる tf-idf 法に類似した手法を採用した。我々の実装では、 i 番目の人物が低階層クラスタに出現する回数を a_i とし、写真群全体に出現する回数を b_i としたとき、その低階層クラスタ

における i 番目の人物のキーパーソン度を a_i/b_i と定義し、この値が最大である人物を当該クラスタのキーパーソンであるとした。この方法を採用した根拠は、写真群全体を見渡すとあまり写っていないが、そのクラスタに限っては頻繁に写っている、という人がそのイベントを象徴する存在である可能性が高い、という直感に基づくものである。

続いて各低階層クラスタについて、キーパーソンが写っている写真のうち1枚を代表写真に選出する。まず写真群の中から、キーパーソンが写っている写真を抽出する。続いて抽出された写真の中から、前後2枚との撮影時刻間隔が短い写真を特定し、これを代表写真とする。この方法を採用した根拠は、撮影時刻間隔が短い風景は、撮影者が特に重要だと感じている可能性が高い、という直感に基づくものである。

なお、人物が写っている写真が1枚もない低階層クラスタについては単純に、前後2枚との撮影時刻間隔が短い写真を代表写真とした。

ここまでの処理で各低階層クラスタの代表写真を選んだら、続いて高階層クラスタの代表写真を選択する。この処理では単純に、高階層クラスタの配下にある低階層クラスタのうち、最も写真枚数の多い低階層クラスタの代表写真を、高階層クラスタの代表写真に選出する。

3.3 ネットワークの構成と画面配置

続いて各クラスタを画面配置する。本手法では高階層クラスタ間に人物の共起性に基づいてエッジを生成することで高階層クラスタのネットワークを形成する。この処理では高階層クラスタごとに、配下の写真に写る人物を列挙する。そして任意の高階層クラスタのペアに対して、共通に写っている人物を列挙し、その人数に比例した太さのエッジを生成する。

以上の処理によって構築されたネットワーク構造に対して、FRUITS Net [3]のアルゴリズムを適用することで、各クラスタの画面上の位置を算出する。画像ブラウザにこのアルゴリズムを適用した際の画面配置結果の妥当性は、既に文献[4]でも検証されている。

3.4 ユーザインタフェース

我々の実装におけるウィンドウ構成を Fig. 1 に示す。この例はズームアウトした状態を示しており、写真1枚が1つのイベントに該当する。これを眺めることで、写真群に含まれるイベントを一目で把握できる。本手法では CAT [1]と同様に詳細度制御機能を設けており、ズームイン操作により局所的にクラスタ内の各写真を表示する。ここで、人物タブで人物を選択すると、選択された人物が写るイベントをピンク色にハイライトする。また、日付タブで日付を選択すると、該当する日付のイベントをオレンジ色にハイライトする。以上の機能により、人物と日付の相互操作が可能となる。

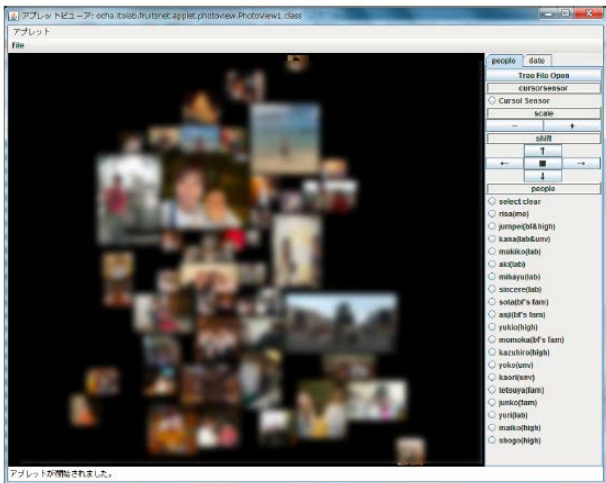


Fig. 1 GUIの概観.

4. 実行結果

個人が所有している写真群，複数の人物で共有している写真群の2種類の写真群に本手法を適用した．

4.1 個人が所有する写真群への適用

特定の個人が撮影した1368枚の写真を対象に，19の人物にIDを割りあてて表示した結果をFig. 2に示す．クラスタリングにより生じた高階層クラスタ数は50，低階層クラスタ数は79であった．1人の人物を選択すると，右下部にイベントが集中していることがみてとれることから，確かに写真の配置に人物が反映されているといえる．この例からは，研究室，大学の友人，高校の友人，家族，恋人，恋人の家族といったコミュニティが，配置結果から観察された．

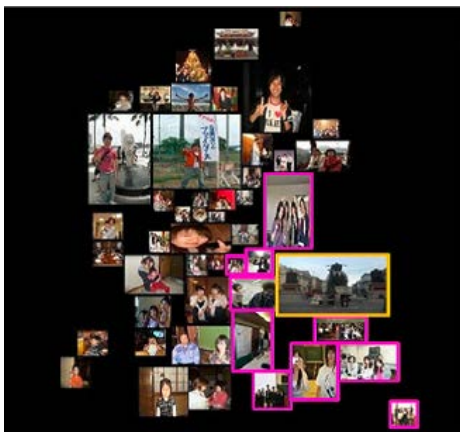


Fig. 2 個人が所有する写真群の表示結果.

4.2 複数の人物で共有する写真群への適用

我々が所属する研究室の複数のメンバーにより撮影されて共有されている2913枚の写真を対象に，30の人物にIDを割りあてて表示した結果をFig. 3に示す．クラスタリングにより生じた高階層クラスタ数は26，低階層クラスタ数は65であった．1人の人物を選択すると，中央部

にイベントが集中していることがみてとれることから，複数人で共有している写真においても写真の配置に人物が反映されているといえる．この例からは，多くのイベントに参加している人物，あまりイベントに参加していない人物といった個々人の特徴，同一のイベントに参加することの多いメンバーといったメンバー間の相関関係を見てとることができた．

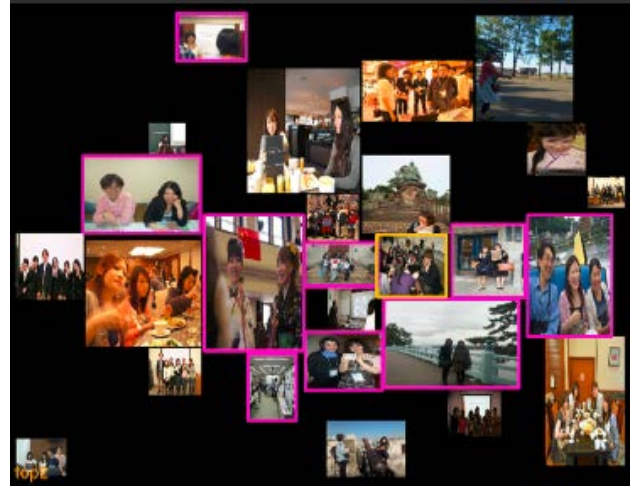


Fig. 3 複数の人物で共有する写真群の表示結果.

5. 代表写真の選出方法に関する考察

3.2節で提案したtf-idf法に基づく代表写真選出手法の妥当性を考察するために，以下の3種類の代表写真選出手法を適用して写真群を一覧表示し，その結果を比較した．

- 写真中の人物を考慮せず，単に撮影時刻間隔の狭い画像を代表写真に選出する．
- tf-idfを考慮せずに，クラスタ内での出現頻度が最も高い人物をキーパーソンとして，キーパーソンが写る写真の中で撮影時刻間隔の狭い画像を代表写真に選出する．
- 3.2節で提案したとおり，tf-idf法に基づいてキーパーソンを選び，キーパーソンが写る写真の中で撮影時刻間隔の狭い画像を代表写真に選出する．

4.1節で示した写真群において，a)b)c)の3種類の代表写真選出方法を採用したところ，a)では所有者自身が写る写真が多く代表写真に選出された．それと比較してb)c)では，a)で選出された写真のうち数枚が，所有者でない人が写る写真に入れ替わっていた．このことからa)よりもb)c)のほうが，各クラスタを象徴する代表写真を選出できていると考える．

ここでb)c)の結果を比較するために，各人物がキーパーソンに選ばれた頻度をFig. 4に示す．この結果において，b)による代表写真の選出では，キーパーソンに一度も選出されない人物，何度もキーパーソンに選出される人物と偏りが見受けられた．それに対してc)では，全ての人物がキーパーソンに選出されており，また人物間の偏りも人物の出現頻度によるものと比較すると少ないといえる．

以上より、個人が所有する写真群においては、tf-idf 法による代表写真の選出が最も適しているのではないかと見える。

続いて、4.2節で示した写真群において、b)c)の方法で各人物がキーパーソンに選ばれる頻度をFig. 5に示す。この結果からは、Fig. 4において表れたような差異は見受けられなかった。

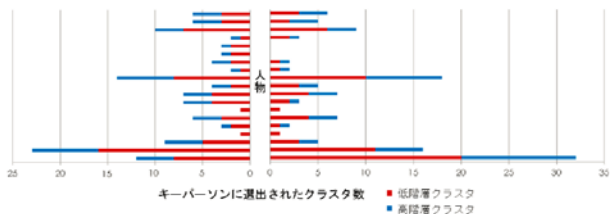


Fig. 4 4.1節で示した写真群において、各人物がキーパーソンに選ばれた頻度を集計した結果。(左)c)の方法による代表写真選出。(右)b)の方法による代表写真選出。

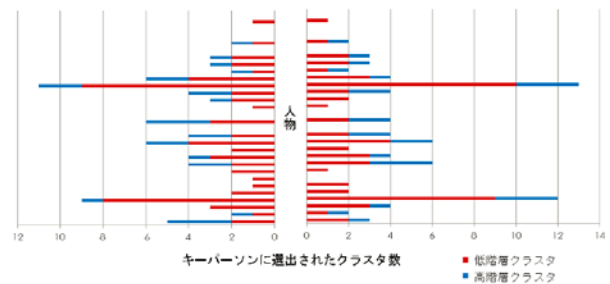


Fig. 5 4.2節で示した写真群において、各人物がキーパーソンに選ばれた頻度を集計した結果。(左)c)の方法による代表写真選出。(右)b)の方法による代表写真選出。

そこで我々は、ユーザアンケートによって代表写真の選出結果を検証した。我々はPicasaが選出した各イベントの代表写真、およびa)b)c)によって選出した各イベントの代表写真を、被験者である17名の女子学生に提示した。そして各代表写真について、各々のイベントらしさを最もよく表している写真を、以下の2つの観点から選択してもらった。

【観点1】 誰が参加したイベントかを最もよく表している。

【観点2】 どのようなイベントかを最もよく表している。

以下に、アンケート集計結果からわかった点を論じる。観点1では、ほぼ全てのイベントにおいてb)c)により選出された代表写真の方がそのイベントらしさをよく表しているとの回答結果が得られた。観点2においても、概ねb)c)により選出された代表写真の方がそのイベントらしさをよく表しているとの回答結果が得られた。また、観点1では回答が一意に定まることが多かったが、観点2では回答に分散がみられた。以上から、人物という着目点を与えることで、多くの人物がそれを代表写真と共通で認識しうる写真を選出することができたのではないかと見える。一方、tf-idf 法と人物の出現頻度の間での顕著な差異は見

受けられなかった。この点については、さらなる検証が必要である。

6. まとめ

本報告では、人物の共通性を配置結果に反映した大量個人写真の一覧可視化手法について提案した。

今後の課題として、クラスタリング手法および代表写真選出方法の再検討、GUIの機能の拡充、Android携帯エミュレータ上での実装を進めたい。

参考文献

- 1) 五味, 宮崎, 伊藤, Li.: CAT: 大量写真の一覧可視化と詳細度制御のためのGUI, 写真電子学会誌, Vol.38, No.4 (2008), pp.1108-1115.
- 2) 伊藤, 山口, 小山田: 長方形の入れ子構造による階層型データ可視化手法の計算時間および画面専有面積の改善, 可視化情報学会論文集, Vol.26, No.6 (2006), pp.51-61.
- 3) T. Itoh, C. Muelder, K.-L. Ma, J. Sese: A Hybrid Space-Filling and Force-Directed Layout Method for Visualizing Multiple-Category Graphs, IEEE Pacific Visualization Symposium (2009), pp.121-128.
- 4) 安田, 五味, 伊藤: キーワードの共通性を配置結果に反映した大量画像の一覧可視化手法, 可視化情報学会第38回可視化情報シンポジウム, P7-002, 2010.