

ソーシャルキュレーションデータを用いた 画像コンテキストマイニング

木村 昭悟^{1,a)} 石黒 勝彦¹ Alejandro Marcos Alvarez^{1,†1}
山田 誠^{1,†2} 片岡 香織² 村崎 和彦²

受付日 2015年5月25日, 再受付日 2015年7月15日,
採録日 2015年7月23日

概要: 本論文では, 画像を扱うソーシャルキュレーションサービスの1つであるPinterestの画像データから, 共通するコンテキストを持つ画像群を自動的に発見する手法を提案する. ソーシャルキュレーションは, 日本語では一般に「まとめサイト」とも呼ばれ, キュレーターと呼ばれる人間が既存コンテンツを人手で収集・選択した結果を編集コンテンツ群として共有する仕組みである. その結果として得られた編集コンテンツ群は, 通常のSNS上のコンテンツとは異なり, キュレーターにとって不要なコンテンツが除去され, キュレーターの意図や意見を反映した一貫した指針に基づいて生成されている. このことから, 編集コンテンツ群を構成するコンテンツは, 共通のコンテキストを保持していることが期待される. 本論文では, このソーシャルキュレーションのプロセスに着目することで, 自然言語との直接的な対応付けが必ずしも容易であるとは限らない概念やコンテキストなどを共有する画像群を, 非常にシンプルな手法で大量に発見できることを示す. さらに本論文では, 発見した画像群を画像認識・検索のための類似性を自動的に獲得するための重要なステップとして, コンテキストの共通性を考慮した画像特徴量の低次元埋め込み手法を提案する. この手法を利用することで, 類似するコンテキストを持つと思われる画像が類似する低次元特徴量を持つような埋め込みを実現するとともに, 画像分類のタスクで同様の機能を実現する従来手法と比較して高い分類性能を実現できることを実験的に示す.

キーワード: ソーシャルキュレーション, グラフクラスタリング, 画像分類, 補助情報, 次元削減

Discovering and Describing Image Contexts from Socially Curated Contents as Corpora

AKISATO KIMURA^{1,a)} KATSUHIKO ISHIGURO¹ ALEJANDRO MARCOS ALVAREZ^{1,†1}
MAKOTO YAMADA^{1,†2} KAORI KATAOKA² KAZUHIKO MURASAKI²

Received: May 25, 2015, Revised: July 15, 2015,
Accepted: July 23, 2015

Abstract: This paper proposes a novel method for discovering a set of image *contexts* sharing a specific *context* with the help of image collections obtained from content curation platforms. Socially curated contents are promising to analyze various kinds of multimedia information, since they are manually filtered and organized based on specific individual preferences, interests or perspectives. Our method fully exploits the process of social curation: How images are manually grouped together by users, and how images are distributed in the platform. Our method reveals that images with a specific context are naturally grouped together and every image includes really various contexts that cannot necessarily be verbalized by texts. In addition, we show how the image contexts obtained from socially curated contents can be used for several tasks such as image classification, data visualization and image retrieval. The key idea is to incorporate the contexts as side information to derive an embedding transformation so that images with similar contexts are close in the embedding space. Through experiments with images obtained from Pinterest, we show that our proposed method benefits several image-related tasks such as visualization, image classification and image retrieval.

Keywords: social curation, graph clustering, image classification, side information, dimensionality reduction

1. Introduction

任意の画像を柔軟に認識・理解する知能の構築は、パターン認識・メディア理解の研究分野における究極目標の1つでありながら、いまだに挑戦的な研究課題の1つである。その理由として、画像の見た目と内容との間に存在するいわゆる“セマンティックギャップ”を埋める特徴量・モデル化などの技術が十分ではない点がある。特に、自然言語との直接的な対応付けが必ずしも容易ではあるとは限らない概念やコンテキストなどを獲得し表現することにおいては、計算機の能力は人間にはるか及ばないのが現状である。一方で人間は、言語・音・外部知識など各種の補助情報を利用することで、画像の内容を形成する概念を獲得していると考えられ、計算機においても同様の戦略が有効であると期待される。

上記の問題を解決しうる可能性を持つ新しいタイプの情報資源として、ソーシャルメディアが非常に注目を集めている。ソーシャルメディアが注目を集めている1つの理由は、コンテンツ間の関連性・ユーザネットワーク・位置情報など、従来のマルチメディアコンテンツにはあまり含まれなかった多様な付加情報が大量に獲得できる点にある。ソーシャルメディアの基盤をなすソーシャルネットワーキングサービス (SNS) において、画像コンテンツはテキストと並ぶ重要な位置を占めるようになり、現在では、Twitter・Facebookなどの主要なSNSが、Flickr・YouTube・ニコニコ動画などのコンテンツ共有サイトと同等に大量の画像コンテンツを保有するようになり、その量は増加の一途をたどっている。また、InstagramなどSNS上での画像をよりリッチにすることを意識したサービスや、Pixiv・Snapchatなど画像に特化したSNSも普及し、SNS上での画像の重要性は日に日に増している。さらに、SNSの発展によって、誰もが情報を発信し消費できるようになったがゆえに、大量の画像コンテンツが次々と生み出されるようになり、人間ですらも、日々拡大する膨大な量のメディアコンテンツを扱いきれなくなっている。そのため、まずは、あまりに膨大すぎるデータの中から、価値あるメディアや情報源を効率的に抽出することが必要となる。SNS上のメディアコンテンツを解析する様々な先行研究の多くは、コンテンツ

から情報や価値を自動的に抽出し可視化することを目指している。

その一方で、自動的なソーシャルメディア解析とは異なる方向性で注目を集める一大潮流として、ソーシャルキュレーション (social curation) がある (図 1)。これは、ユーザである人間が自発的に人手でメディアの収集・評価・選択および編集を行い、その結果として得られた編集コンテンツ群を共有するプロセスとして規定され、日本では一般に「まとめサイト」の名で広く知られている。ソーシャルキュレーションを支えるのは、キュレーター (curator) と呼ばれる、メディアの取捨選択と再構成を行うユーザ層である。このキュレーターの介在により、ソーシャルキュレーションは、一般的な自動解析・可視化技術とは異なった独自の再編集コンテンツを生み出すことが可能となる。すなわち、取捨選択された再編集コンテンツは、キュレーターの目と手によって人為的にかつ一貫した方針によって収集され、同一のコンテキストを共有していることが期待される。

上記の背景に基づき、本論文では、ソーシャルキュレーションの特性に着目した、新しい画像コンテンツ解析の方法を提案する*1。特に本論文では、画像を対象コンテンツとするソーシャルキュレーションサービスの代表格であるPinterestを利用する。Pinterestには、画像コンテンツをまとめるグループである“board”を自由に作成できる機能がある。Pinterestでは、このboardを、ユーザ自身がアップロードした画像ではなく、他のユーザが持つ別のboardに含まれる画像で構成することが一般的である。すなわち、boardにユーザが画像コンテンツを集めていく作業は、画像コンテンツを対象としたソーシャルキュレーションと見なすことができる。また、多くのユーザは、自分自身が所望の画像を見つけやすいように、各boardに特定のコンテキストを持たせる。各boardが特定のコンテキストを持ち、かつ画像がboardからboardへコピーされていくことを考慮すると、数多くの画像を共有しているboardの対は、類似したコンテキストを持つことが期待される。本論文では、この知見を利用することで、類似したコンテキストを持つ画像群を発見する問題を、Pinterest上での画像の共有行動を表現するグラフから強連結構造を抽出する問題として定式化できることを示すとともに、独自に収集した大規模データセットを用いて、グラフ構造の近接関係がコンテキストの類似関係と密接にかかわることを実験的に示す (図 2 左)。

さらに本論文では、このように獲得したコンテキストの類似関係を、未知の画像の認識や検索のために利用する手法についても提案する。具体的には、グラフ表現されたコンテキストの類似関係を補助情報として利用する、画像特徴量の低次元埋め込みの新しい手法を提案する。この手法

¹ 日本電信電話 (株) コミュニケーション科学基礎研究所
Communication Science Laboratories, NTT Corporation,
Atsugi, Kanagawa 243-0198, Japan

² 日本電信電話 (株) メディアインテリジェンス研究所
Media Intelligence Laboratories, NTT Corporation, Yokosuka,
Kanagawa 239-0847, Japan

^{f1} 現在, リエージュ大学
Presently with Systems and Modeling Research Unit, University of Liege

^{f2} 現在, 京都大学化学研究所
Presently with Institute of Chemical Research, Kyoto University

a) akisato@ieee.org

*1 本論文の一部は、ACMMM2013にて報告した内容に基づく [1], [2].

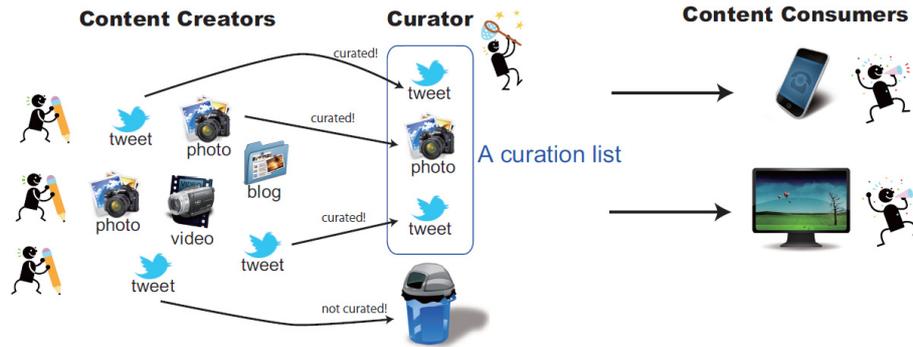


図 1 ソーシャルキュレーションのプロセス
Fig. 1 Process of social curation.

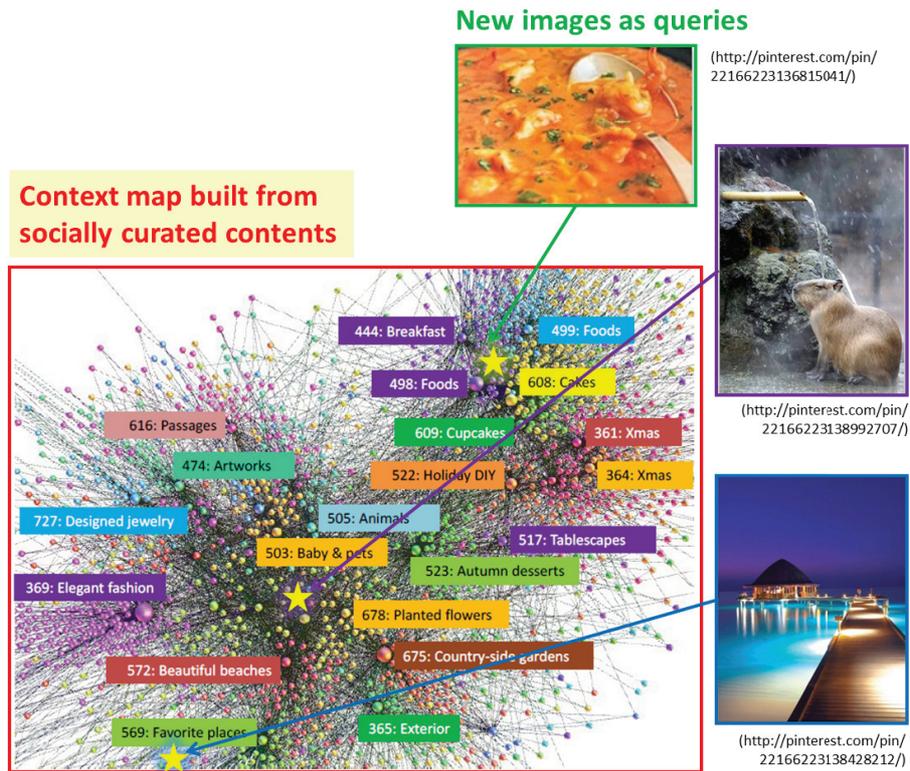


図 2 提案手法の概略
Fig. 2 Outline of the proposed method.

を利用することで、言語化が必ずしも容易とは限らないコンテキストの類似関係を、低次元画像特徴量の類似関係に変換できるようになるとともに、より高次の意味や概念を考慮した画像分類・認識・検索が可能となる (図 2 右)。

本論文の以降の構成は以下のとおりである。まず、2章で本論文に関連する既存研究を紹介するとともに、それら既存研究に対する本論文の独自性について言及する。続く2つの章では、本論文で提案する2つの手法について詳しく述べる。まず、3章では、ソーシャルキュレーションデータから類似したコンテキストを持つ画像群を自動的に発見する手法を説明する。続く4章で、前章で発見したコンテキストの類似関係を利用した画像特徴量の低次元埋め込み手法を具体的に示す。5章でクロールしたソーシャルキュ

レーションデータを用いた実験により、これら提案手法の有効性を検証し、6章で本論文を総括し、今後の展望について述べる。

2. 関連研究

ソーシャルメディア上の画像コンテンツの解析については、これまでに様々な研究がなされている。その際に最も数多く用いられているデータ資源が Flickr である。Flickr は画像共有サイトではあるが、コミュニティやグループなど SNS に類する機能も存在しており、このようなソーシャル機能に注目した研究、たとえば、ユーザに好みの Flickr グループを推薦する方法 [3], [4], Flickr 内のコミュニティ検出 [5], 画像検索のランキングの修正 [6] などに関する研

究が知られている。また、Flickr にアップロードされた画像には、テキストタグやジオタグなど、様々なメタデータが利用可能であることも特筆すべき点としてあげられる。位置情報が類似する画像に付与されたテキストタグを集約することで、その地域で人々の興味関心を集める対象を示すテキストタグや、そのタグを端的に表現する画像を抽出する研究 [7] は、このメタデータを有効に活用した研究の代表例である。また、テキストタグやメタデータの共通性を画像間の関係性を表現する特徴量とすることで画像分類や画像説明文生成の性能を向上させる手法 [8], [9], [10] も提案されている。さらに、YouTube や Chictopia など、他の SNS プラットフォームのデータを用いることで、映像コンテンツの推薦 [11], 画像アノテーション [12], ファッションアイテムの人気度予測 [13] などを行う研究が知られている。

一方、本論文では、一般的なソーシャルメディアではなく、自発的に人手でメディアの収集・評価・選択を行い、その結果を共有するプロセスに独自性があるソーシャルキュレーションを対象としており、そのプロセスに着目した解析を行う点に特長がある。

ソーシャルキュレーションは比較的新しく登場した概念であり、この特性に着目した学術的な研究はそれほど多くはない [14], [15]。また、それら先行研究の多くは、ソーシャルキュレーションというプロセスそのものがどのようなものであるか、それがもたらしたユーザマインドの変化は何か、など、社会学的な見地からソーシャルキュレーションという現象そのものあるいはソーシャルキュレーションサイトにおけるユーザ行動を解明しようとする試みである [16], [17], [18], [19], [20], [21], [22], [23], [24], [25], [26], [27]。これらの研究は、本論文で取り扱うソーシャルキュレーションデータを用いた画像コンテンツの解析において、重要な知見として活かされているが、本論文と直接関連するものではない。

ソーシャルキュレーションデータを用いたコンテンツ解析の初期の研究例として、Duh ら [28], [29] は、Twitter まとめサイトの先駆的存在である Together を対象として、

大量に存在する Twitter 投稿から Together まとめを半自動的に構成する方法を提案している。Greene ら [30] は、Twitter 投稿を映画、Together まとめを映画再生リストに置き換えて同様のアイデアを用いることで、映画の推薦が実現できることを示した。Ishiguro ら [31] は、Together のまとめに数多くの画像コンテンツが含まれていることに着目し、これらの画像コンテンツの潜在的な商業価値を示す閲覧回数を、ソーシャルキュレーションから獲得できる非画像特徴量から高精度に予測可能であることを示した。Chen ら [32] は、Ishiguro らと同様のアプローチで、Twitter 投稿のテキストと画像の関連性を高精度に判別できることを示した。Pinterest を利用した研究例としては、Pinterest の各画像コンテンツを EC サイトのページにリンクさせる方法が知られている [33]。

しかしながら、我々の知る限り、ソーシャルキュレーションデータを用いて画像のコンテキストを解析し、それを画像の認識や検索に活用する研究は、本論文が初出である。

3. 画像コンテキストの自動発見

3.1 Pinterest の仕組み

まず、本論文で対象とするソーシャルキュレーションサービスである Pinterest の概要について、図 3 および図 4 を用いて説明する。

Pinterest は、スクラップブックに写真や雑誌の切り抜きを貼り付けていくような感覚で、お気に入りの画像を集めたり共有できる、画像共有サービスと見ることができる。Pinterest のユーザは、“board” と呼ばれる画像のグループをいくつか所有し、一般的な SNS 同様に、フォローしている友人が共有したコンテンツが、“wall” と呼ばれるタイムラインに表示される。この wall でお気に入りの画像を見つけた際に、ほぼ 1 クリックでその画像を自分が所有するいずれかの board に“pin” (画像へのリンクを作成) することができる。ユーザが pin する手段は、主に以下の 2 通りが提供されている。

- (1) Pin: Web 上にある画像へのリンクを作成し、それをいずれかの board に紐付ける。



図 3 Pinterest の概要

Fig. 3 Outline of Pinterest.

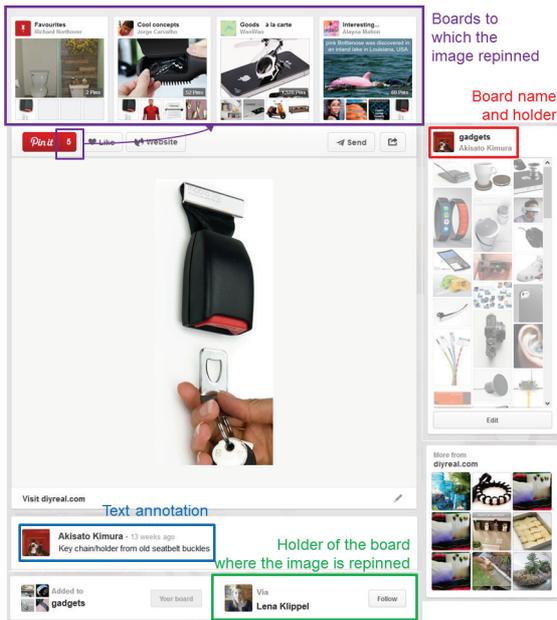


図 4 Pinterest における pin の一例
 (See <http://pinterest.com/pin/22166223140004983/>)
 Fig. 4 Example of a pin in Pinterest.

(2) “Repin”: Twitter におけるリツイートと同様の概念であり，他のユーザのいずれかの board に存在する pin をコピーして，それを自分のいずれかの board に紐付ける。

この 2 通りの pin する手段が提供されることにより，
 (1) まず，あるユーザが興味を持つ画像を web で見つけて pin し，
 (2) その画像を別のユーザが wall で見つけて repin し，
 (3) また他のユーザが repin する，
 という，画像コンテンツの伝搬のサイクルが形成される。

このように作成されたそれぞれの pin には，図 4 に示すように，画像コンテンツだけではなく，画像出展元の URL，画像が pin された board，repin 元の board の所有者，画像を最初に pin したユーザや board，その pin が repin された board やユーザなど，様々な補助方法が含まれる。多くのユーザは，自分自身もしくは他のユーザが所望の画像を素早く見つけられるように，各 board に自身が設定した特定の概念や内容を対応させ，その概念や内容に適合する画像のみをその board に pin する傾向が強く見られる。

本研究の遂行のために，Pinterest から総計約 100 万 boards，1 億 pins のデータをクロールしている。本論文の以降では，その部分集合である 19 万 boards，104 万 pins を利用する*2。104 万 pins のうち repin は 90 万，ユニーク画像数は 78 万枚であった。すべての pin の repin 元を完全にクロールできなかったため，repin であることが確認できた pin であっても，その repin 元がデータセットに含

*2 “Image Context Discovery” dataset <http://www.brl.ntt.co.jp/people/akisato/socialweb4.html>.

Algorithm 1 Compute pin diffusion graphs

```

Require: allPins: a set of all the pins in the dataset
for pin in allPins do
    pin.node ← new node
    Append pin.node to allNodes
end for
for pin in allPins do
    parent ← find the parent of pin from allPins
    if parent not exists then
        parent ← NULL
    end if
    edge ← new edge
    edge.root = parent.node
    edge.dest = pin.node
    Append edge to allEdges
end for
Return: allNodes, allEdges
    
```

まれていないケースも多数存在する。ユニーク画像数が，repin ではない pin の数よりも多くなっているのは，このためである。

3.2 類似コンテキストを持つ画像群の発見

各 board が特定のコンテキストを持ち，かつ repin によって複数の board を次々と伝搬していくことを考慮すると，伝搬を通じて数多くの画像を共有した board の対は，類似したコンテキストを持つことが期待される。この知見を利用すると，類似したコンテキストを持つ画像群を発見する問題は，伝搬を通じて数多くの画像を共有した board の集合を発見する問題に置き換えることができる。本節では，このような board の集合を効率的に発見する手法について説明する。

前節に示したように，各 pin から，pin された画像の出展元 URL，その pin が repin であるかどうか，repin 元の board を所有するユーザを知ることができる。これらの情報から，当該 pin の repin 元である pin をただ 1 つに特定することができる。図 5 において，Pin 2 が Pin 1 の repin 元として特定された pin である。同様にして，すべての pin に対してその repin 元をたどることにより，各 pin をノード，repin 元 pin に対応するノードから repin 先 pin に対応するノードへの経路を有向エッジとする，木形状の有向グラフを多数得ることができる。この有向グラフを pin diffusion graph と呼ぶ。Pin diffusion graph の例を図 5 に示す。また，pin diffusion graph を得るアルゴリズムを Algorithm 1 に示す。

Repin 元 pin と repin 先 pin が異なる pin ID を持つことから，各 pin が属する board は一意に特定できる。つまり，すべての pin はいずれか 1 つの board に属することになる。このことから，Algorithm 2 に示すように，pin diffusion graph の各ノードを，所属する board ごとにと

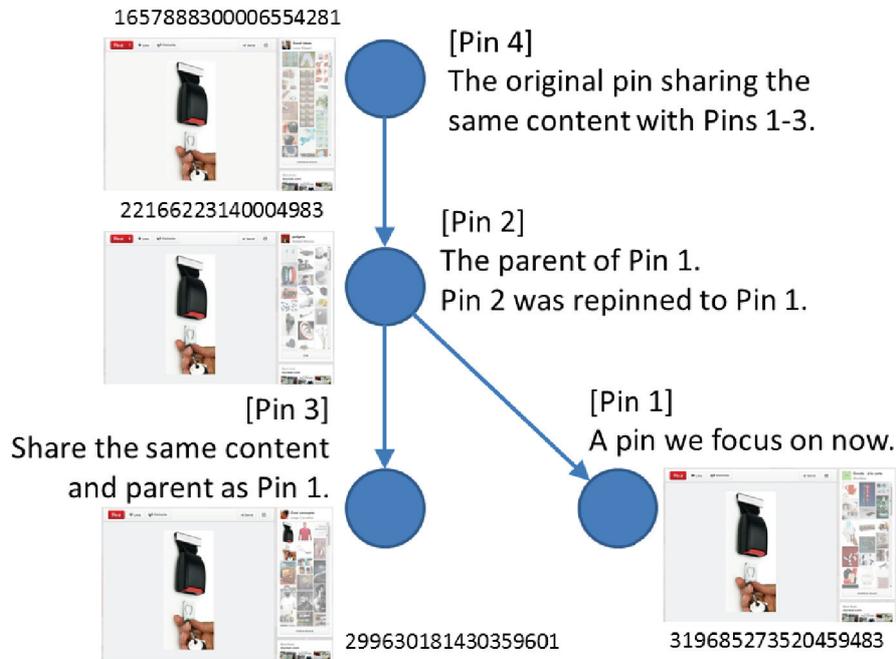


図 5 Pin diffusion graph

Fig. 5 Pin diffusion graph.

Algorithm 2 Compute pin diffusion graphs

Require: *allBoards*: a set of all the boards in the dataset

Require: *allNodes*: a set of all the nodes in the pin diffusion graph

Require: *allEdges*: a set of all the edges in the pin diffusion graph

for *board* in *allBoards* **do**

board.node ← new node

 Append *board.node* to *allNewNodes*

end for

for *board* in *allBoards* **do**

for *pin* in *board* **do**

parentBoard ← Find the board that contains *pin.parent*

if *parentBoard* exists **then**

nowEdge ← find the edge from *parentBoard.node* to *board.node* in *allEdges*

if *nowEdge* not exists **then**

edge ← new edge

edge.root ← *parentBoard.node*; *edge.dest* ← *board.node*

edge.weight ← 1

 Append *edge* to *allNewEdges*

else

nowEdge.weight ← *nowEdge.weight* + 1

end if

end if

end for

end for

Return: *allNewNodes*, *allNewEdges*

めることができ、その結果は一意に定まる。その過程で、*repin* 元・*repin* 先双方の board を共有するエッジも 1 つにまとめられ、新しいエッジの重みはまとめられたエッジの数となる。このように構成された、board をノードとする新たな有向グラフを、board diffusion graph と呼ぶ。図 6

に、board diffusion graph の構成例を示す。また、board diffusion graph を得るアルゴリズムを Algorithm 2 に示す。前節 (3.1 節) に示したデータに上記の手順を実行することにより、18 万ノード、32 万エッジを持つ board diffusion graph が得られた。

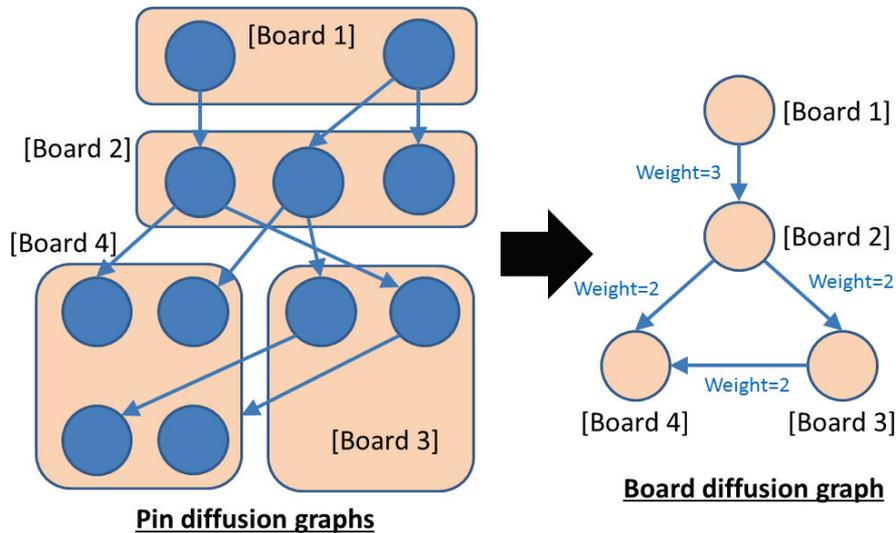


図 6 Board diffusion graph
Fig. 6 Board diffusion graph.



図 7 主要クラスタに所属する board 名の word cloud (1) “Holidays”
Fig. 7 Word cloud of board names in clusters (1) “Holidays”.

Board diffusion graph は、board から board への画像の伝搬の様子をグラフ表現したものである。このことから、本論文で解くべき問題である、類似したコンテキストを持つ画像群を発見する問題は、board diffusion graph で相互連結の強いノード (board) の集合を検出する問題と読み替えることができる。この問題は、グラフクラスタリングとして広く知られており、モジュラリティ最大化として定式化する方法が標準的である。本論文では、モジュラリティ最大化基準のグラフクラスタリング手法として Louvain 法 [34], [35] を採用するとともに、多数のノードを含む巨大

クラスタが生成されないように、一定数以上のノードを含むクラスタに再帰的にクラスタリングを実行する階層クラスタリングを行った。

本節前半に示した 18 万ノード、32 万エッジを持つ board diffusion graph に対して上記のクラスタリングを実行した結果、約 1,000 個のクラスタが得られた。得られたクラスタのうちサイズの大きな上位 15 クラスタを取り出し、各クラスタに含まれる board の名前で word cloud をそれぞれ構成した結果を、図 7, 図 8, 図 9 に示す。この word cloud では、より高頻度に出現した単語ほどフォントが大き



図 8 主要クラスタに所属する board 名の word cloud (2) “Art colors”
Fig. 8 Word cloud of board names in clusters (2) “Art colors”.



図 9 主要クラスタに所属する board 名の word cloud (3) “Foods & cute”
Fig. 9 Word cloud of board names in clusters (3) “Foods & cute”.

く表示される。また、board diffusion graph を 2次元平面上に可視化した結果を図 10 に示す。このとき、ノードの大きさが次数を、ノードの色が所属するクラスタを表現し、エッジの重みが大きいノードをより近くに配置する可視化

を行った。これらの可視化結果から、以下の 3 点が明らかとなった。(1) 各クラスタが共通のコンテキストを持った board によって構成されている。たとえば、図 7 を見ると、Christmas, Halloween, Thanksgiving, holiday など祝日

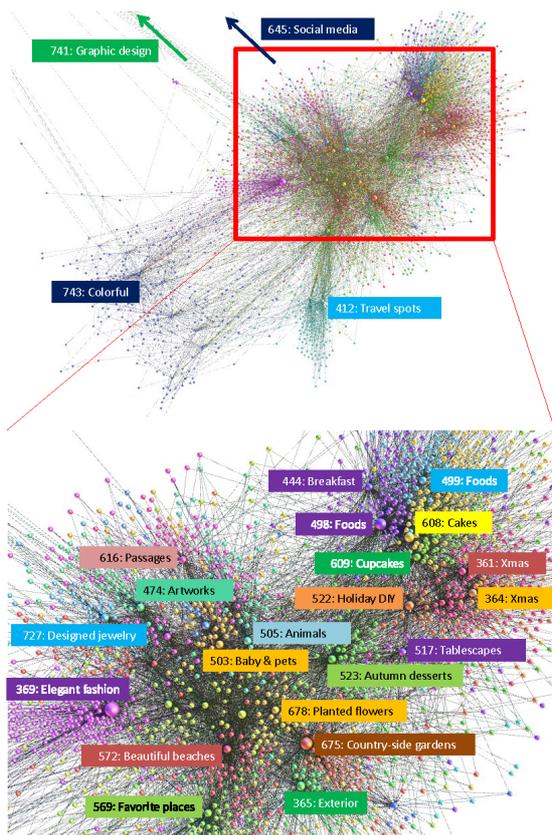


図 10 Board diffusion graph およびそのクラスタリング結果の可視化

Fig. 10 Visualization of a board diffusion graph and board clustering.

に関連する単語や, craft, diy, idea など手芸に関連する単語が共通して多く含まれることが見てとれる. このことは, 図 7 に示された 5 つの board クラスタが, 休日を彩る手芸品の画像によって構成されていることを示唆する. 一方で, (2) わずかに異なるコンテキストが適切に別のクラスタに分離されていることも見てとれる. 図 7 に示された 5 つのクラスタは, 主要な単語が少しずつ異なっている. すなわち, それぞれのクラスタは, Christmas, Halloween, Thanksgiving など個別の休日に焦点を絞って構成されている. さらに, (3) 類似したコンテキストがグラフ上で隣接した位置にあることも, 図 10 から見てとれる. Christmas に関連するクラスタが右上, その近隣にはケーキや食事に関するクラスタ, さらにその左下に休日の手芸に関するクラスタが見られる可視化の結果から, 関連したクラスタで多くの画像コンテンツが共有されていることが示唆される. 各クラスタが類似したコンテキストを持つかどうかについては, 後述の 5.1 節で実験的に検証を試みる.

4. コンテキストを考慮した画像特徴量の低次元埋め込み

ここまで, ソーシャルキュレーションデータから類似したコンテキストを持つ画像群を自動的に発見する方法につ

いて説明してきた. すなわち, board diffusion graph をクラスタリングした結果, 同じクラスタに所属した board が, 類似したコンテキストを共有する画像群と見なせることを示してきた. さらに本論文では, このように獲得したコンテキストの類似関係を, 未知の画像の認識や検索のために利用する手法についても提案する. 具体的には, board diffusion graph という形でグラフ表現されたコンテキストの類似関係を補助情報として利用する, 画像特徴量の低次元埋め込みの新しい手法を提案する. この手法を利用することで, コンテキストの類似関係を低次元画像特徴量の類似関係に変換することを可能とする.

本章で提案する低次元埋め込み手法は, learning using privileged information (LUPI) と呼ばれる, 学習時のみに利用可能な補助情報を用いた機械学習手法の枠組を利用している [36], [37]. すなわち, 特徴ベクトル $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^d$ からクラスラベル $\mathbf{y} \in \mathcal{Y} \subset \mathbb{N}$ を予測する予測器 $f: \mathcal{X} \rightarrow \mathcal{Y}$ の学習に, 特徴ベクトルとは別に用意された補助情報 (privileged information) $\mathbf{v} \in \mathcal{V} \subset \mathbb{R}^p$ を利用する枠組である. ここで注意すべきは, 予測器 f は特徴ベクトルのみからクラスラベルを予測することから, 補助情報は予測器の学習にしか用いることができない点というである.

本論文では, LUPI の枠組で取り扱われてこなかった教師付き次元削減 [2] に言及し, コンテキストの類似関係を考慮した補助情報を具体的に構成する方法を提案する.

4.1 問題設定

まず, 本節で説明する教師付き次元削減の問題設定について記述する.

画像特徴量は d 次元ベクトル $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^d$ で表現され, それぞれクラスラベル $\mathbf{y} \in \mathcal{Y} = \{1, 2, \dots, C\}$ が与えられている. さらに, 各画像特徴量に対して, 別の p 次元ベクトル $\mathbf{v} \in \mathcal{V} \subset \mathbb{R}^p$ が補助情報として与えられているものとする. 補助情報の具体的な構成方法については, 別途後述する. 本節で説明する教師付き次元削減は, LUPI の一般的な設定と同様, この画像特徴量・補助情報・クラスラベルの三つ組み n 個

$$D = \{(\mathbf{x}_i, \mathbf{v}_i, y_i) | \mathbf{x}_i \in \mathcal{X}, \mathbf{v}_i \in \mathcal{V}, y_i \in \mathcal{Y}, i = 1, \dots, n\}.$$

が学習データとして与えられたときに, 各クラスの特徴量をよりよく分離する低次元表現 $\mathbf{z} \in \mathcal{Z} \subset \mathbb{R}^m$ ($m \leq d$) を得るための変換 $f: \mathcal{X} \rightarrow \mathcal{Z}$ を獲得することが目的である. ここでは簡単のため, 線形変換を対象とする.

$$\mathbf{z} = \mathbf{T}^T \mathbf{x}, \tag{1}$$

ここで, $\mathbf{T} \in \mathbb{R}^{d \times m}$ は変換行列, T は行列の転置をそれぞれ表現する. 低次元表現を得るための変換 f は特徴量 \mathbf{x} のみを入力として用いることに注意する. すなわち, 補助情報 \mathbf{v} は, 変換 f を学習するためだけに用いられる.

4.2 局所線形判別分析

本節で採用する次元削減手法は、局所線形判別分析 (Local Fisher Discriminant Analysis: LFDA) [38] と呼ばれる線形判別分析 (FDA: Fisher Discriminant Analysis) の改良手法に基づいて構成されている。まずは、この局所線形判別分析について説明する。

局所線形判別分析 LFDA は、特徴量の局所的な構造を考慮した教師付き次元削減手法であり、線形判別分析 FDA と同様、クラス間分散を大きくし、クラス内分散を小さくするような線形変換 \mathbf{T}^* を一般化固有値問題によって求める。LFDA では、このクラス内・クラス間分散の計算で、特徴量の局所構造を考慮する。クラス間共分散行列 \mathbf{S}_b およびクラス内共分散行列 \mathbf{S}_w はそれぞれ以下のように計算される。

$$\begin{aligned} \mathbf{S}_b &= \frac{1}{2} \sum_{i,j=1}^n D_{i,j}^b (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\top, \\ \mathbf{S}_w &= \frac{1}{2} \sum_{i,j=1}^n D_{i,j}^w (\mathbf{x}_i - \mathbf{x}_j) (\mathbf{x}_i - \mathbf{x}_j)^\top, \\ D_{i,j}^b &= \begin{cases} A_{i,j}(1/n - 1/n_c) & \text{if } y_i = y_j = c, \\ 1/n & \text{if } y_i \neq y_j, \end{cases} \\ D_{i,j}^w &= \begin{cases} A_{i,j}/n_c & \text{if } y_i = y_j = c, \\ 0 & \text{if } y_i \neq y_j. \end{cases} \end{aligned}$$

ここで、 n_c はクラス c に所属する特徴ベクトルの数、 $\mathbf{A} = \{A_{i,j}\}$ は隣接行列であり、その i 行 j 列の要素 $A_{i,j}$ に、 i 番目の特徴ベクトルと j 番目の特徴ベクトルとの間の類似性を表現する値を持つ。一般には、局所スケール法 [39] で特徴ベクトルからアフィン行列を計算すると良い性能が得られることが知られている [38]。

上記 2 つの共分散行列を用いることで、変換行列 \mathbf{T}^* は、以下のトレースノルム最大化問題の解として得られる。

$$\mathbf{T}^* = \operatorname{argmax}_{\mathbf{T} \in \mathbb{R}^{d \times m}} \operatorname{tr} \left[(\mathbf{T}^\top (\mathbf{S}_w + \eta \mathbf{I}_d) \mathbf{T})^{-1} \mathbf{T}^\top \mathbf{S}_b \mathbf{T} \right], \quad (2)$$

ここで、 $\mathbf{I}_d \in \mathbb{R}^{d \times d}$ は $d \times d$ 単位行列、 $\eta > 0$ は正則化パラメータ、 $\operatorname{tr}[\cdot]$ はトレースノルムである。この最適化問題は、以下のような一般化固有値問題と等価である [40]。

$$\mathbf{S}_b \boldsymbol{\varphi} = \lambda (\mathbf{S}_w + \eta \mathbf{I}_d) \boldsymbol{\varphi}.$$

よって、変換行列 \mathbf{T}^* は、上記の一般化固有値問題の解のうち、固有値 λ が大きい順に m 本の固有ベクトル $\boldsymbol{\varphi}_1, \dots, \boldsymbol{\varphi}_m$ を取り出し、それを列ベクトルとして並べた、以下の行列として得られる。

$$\mathbf{T}^* = (\boldsymbol{\varphi}_1 | \boldsymbol{\varphi}_2 | \dots | \boldsymbol{\varphi}_m),$$

4.3 補助情報を用いた局所線形判別分析

LFDA では、アフィン行列 \mathbf{A} に特徴ベクトル間の類似

関係を持たせ、この類似関係に応じてクラス内・クラス間共分散行列の計算に重み付けを与えていた。より具体的には、アフィン行列の要素が大きい値を持つ特徴ベクトルの対についての共分散を重視し、アフィン行列の要素が小さな値を持つ特徴ベクトルの対についての共分散は無視する演算となっている。本論文で採用する補助情報付き局所線形判別分析 [2] では、この点に着目し、アフィン行列の計算に補助情報を用いる。より具体的には、補助情報が類似する場合にアフィン行列の要素が大きく、補助情報が大きく異なる場合にアフィン行列の要素が小さくなるように設定することで、補助情報が類似する特徴ベクトルが共通に持つ特徴成分を強調する。

$$A_{i,j} = \exp \left(-\frac{\|\mathbf{v}_i - \mathbf{v}_j\|^2}{2\sigma^2} \right),$$

ここで σ は分散を制御するパラメータである。

4.4 補助情報の構成方法

本節では、補助情報の具体的な構成方法を提案する。補助情報は、前節で述べたとおり、各画像 \mathbf{x}_i ($i = 1, 2, \dots, n$) について 1 本の多次元ベクトル \mathbf{v}_i が構成され、提案手法が有効に機能するように、コンテキストが類似すると思われる画像についての補助情報が、互いに類似するように構成する必要がある。特に本節では、3 章で獲得したコンテキストの類似関係を考慮した以下の 3 通りの構成方法を提案する。

User: この構成方法では、所定の画像にどのユーザが興味を持ったかについての情報をベクトル化することにより、関心を持つユーザ層の類似性から画像の類似性を表現することを狙う。具体的には、Pinterest のユーザ数と同じ次元のベクトルを用意し、 j 番目のユーザが i 番目の画像を pin もしくは repin したときに、 i 番目のベクトル \mathbf{v}_i の j 次元目の要素 $v_{i,j}$ を 1 に、それ以外の場合には 0 とする。

Board: この構成方法では、所定の画像がどの board に属しているかについての情報をベクトル化することにより、所属 board の類似性から画像の類似性を表現することを狙う。具体的には、board 数を同じ次元数のベクトルを用意し、 j 番目の board が i 番目の画像を含むとき、 i 番目のベクトル \mathbf{v}_i の j 次元目の要素 $v_{i,j}$ を 1 に、それ以外の場合には 0 とする。

Cluster: この構成方法は、board diffusion graph から抽出した board クラスタに関する情報をベクトル化するもので、最もコンテキストに関する情報が多く含まれると期待される。具体的には、クラスタ数と同じ次元数のベクトルを用意し、 i 番目のベクトル \mathbf{v}_i の j 次元目の要素 $v_{i,j}$ を以下のように設定する。

$$v_{i,j} = n_{i,j}/n_i,$$

$n_{i,j}$ は i 番目の画像が j 番目のクラスタに含まれる board

に pin されたのべ回数, n_i は i 番目の画像が pin されたのべ回数の合計値である.

上記 3 通りの構成いずれも, Pinterest を含むソーシャルキュレーションデータに依存している. すなわち, これらの構成による補助情報は, Pinterest に含まれない未知の画像に対しては計算することができない, という点に注意する. 本節で提案する画像特徴量の低次元埋め込み手法は, あらかじめ十分な量のソーシャルキュレーションデータで埋め込みのための変換を学習しておけば, 未知の画像に対しても適用が可能である.

5. 実験

本章では, 提案手法の有効性を検証するために行った, Pinterest から収集した実際のソーシャルキュレーションデータを用いた実験の結果について説明する. 収集したデータは 3 章に示したものと同一で, 19 万 boards, 104 万 pins によって構成される. 構成された board diffusion graph は 18 万ノード, 32 万エッジを持つスパースなグラフとなり, グラフクラスタリングによって約 1,000 個のクラスが抽出された.

5.1 自動検出した画像群のコンテキストの共通性

本節ではまず, Pinterest の board に含まれる画像が共通のコンテキストを持つかどうか, 3 章で説明した手法が類似コンテキストを適切に抽出できたかどうか, この 2 点を検証するために, 4 章で提案した補助情報を画像の特徴量と見なして画像分類を実行する予備実験を行った. あらかじめ定めた 10 クラス^{*3}に関連する単語を複数用意し, その単語を board 名として含む board と, その board に属する 30 万 pins およびユニーク画像 20 万枚をデータセットとして用いた. データセット内の各画像について, 所属する board の名前に対応するクラスラベルを付与するとともに, 4.4 節に説明した 3 通りの手法 (User, Board, Cluster) で補助情報を抽出し, これを画像特徴量とした. 20 万枚の画像を学習用 15 万枚とテスト用 5 万枚にランダムに分割し, 学習用画像の特徴量を用いて 10 クラスの分類器を学習し, テスト用画像で分類器の性能を測定した.

表 1 に, 3 通りの補助情報の構成方法および 4 通りの分類器についての分類性能を示す. クラスラベルの付与の方法から, board 補助情報がきわめて高い分類性能を示す^{*4}ことが予想される. すなわち, cluster 補助情報と board 補助情報との分類性能の相対的な差が, 本節における議論の主

^{*3} Pinterest 上で数多くの画像を確保でき, クラスが相互に関係性を持つよう, 以下の 10 クラスを選択した. Architecture, fashion, cupcake, animals, chocolate, flowers, blue, sea, Christmas, green.

^{*4} 1 枚の画像が同時に複数のクラスに所属することが可能であるマルチラベル分類の設定であるため, board 補助情報を画像特徴量として用いたとしても, 100%の分類率を達成することを保証はできないことに注意する.

表 1 補助情報の特徴量として画像分類を行った際の分類性能 (in %). 行: 分類器の種類, 列: 補助情報の種類

Table 1 Performance of image classification with side information as image features (in %).

	User	Board	Cluster
Linear regression	87.09	98.94	97.36
kNN $k = 1$	85.46	98.70	97.13
$k = 5$	88.08	98.88	97.64
$k = 10$	88.78	98.84	97.62

眼となる. 表 1 の結果は, cluster 補助情報が board 補助情報とほぼ同等の分類性能であることを示している. すなわち, 3 章で説明した, board diffusion graph のクラスタリングに基づく提案手法が, キュレーションデータに内在する画像のコンテキストを適切に抽出できていることを示唆する結果であるといえる.

5.2 コンテキストを考慮した特徴量の低次元埋め込みの効果

続いて, 抽出したコンテキストが画像の認識・検索に有用であるか, 提案の画像特徴量の低次元埋め込み手法が有効に機能するか, この 2 点について検証する.

本節では, 前節に示したデータから無作為に抽出した画像 12,500 枚および関連する pin や board の情報をデータセットとして利用する. このデータセットに含まれる Pinterest ユーザ数は 170, board 数は 630, board クラス数は 390 であった. 画像分類のための画像特徴として, 本節では, 大域的特徴量の 1 つである GIST [41] 960 次元を用いる. データセットに含まれる 12,500 枚の画像から, 5,000 枚を学習用として無作為に抽出し, 画像特徴量の低次元埋め込みの学習および画像分類器の学習に利用した. これとは別に, 5,000 枚の画像をテスト用として無作為に抽出し, このテスト用画像を用いて分類器の性能を測定した. LFDA および提案手法における正則化パラメータ η は, 予備実験により 0.1 と設定した. また, アフィン行列の分散パラメータ σ は, 以下の方法により決定した [42].

$$\sigma = 2^{-0.5} \text{median}(\{\|v_i - v_j\|\}_{i,j=1}^n) \quad (3)$$

図 11 に, 960 次元の GIST 特徴量を 2 次元に埋め込んだ結果を示す. ここでは, 提案手法 (cluster 補助情報) を主成分分析 (PCA)・局所線形判別分析 (LFDA) と比較する. この図から, 提案手法が既存手法と比較してより適切にクラスを分離できていることが分かる.

続いて, 表 2 に, 提案手法を画像分類のタスクに利用した際の実験結果を示す. 本実験では, PCA・LFDA の他に, 低次元埋め込みをしない GIST 960 次元 (Orig.), およびグラフ構造を分類学習の正則化に利用する従来手法 [43] (Graph Reg.) と, 提案手法 (cluster 補助情報) の性能を比較する. 画像特徴量の低次元埋め込み手法の効果を適切

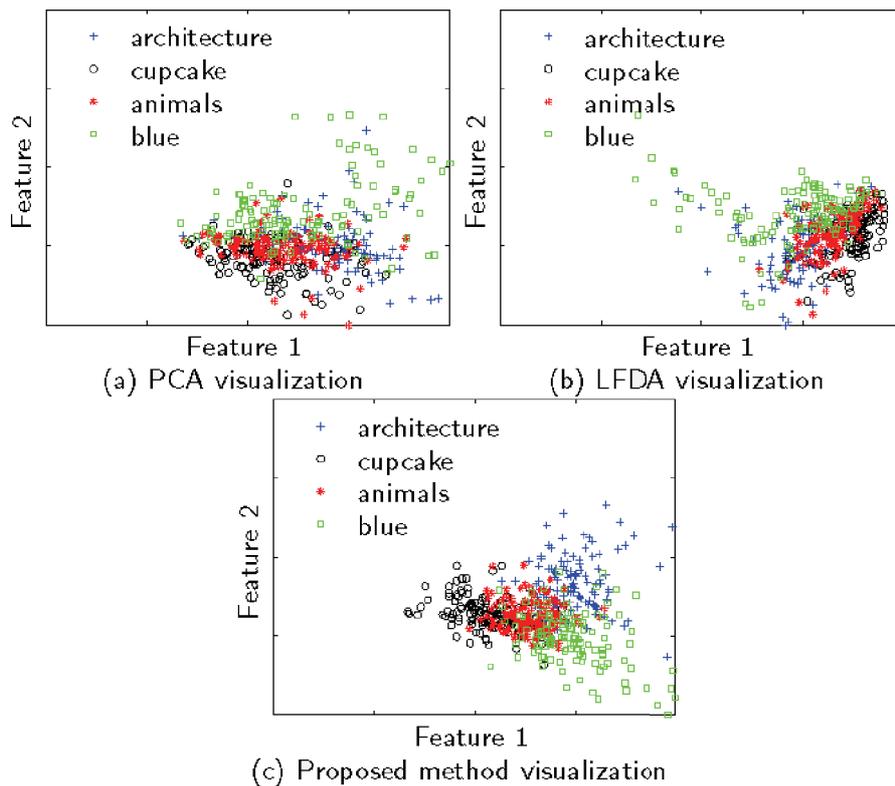


図 11 画像特徴量の低次元埋め込み手法の比較. 各埋め込み手法における主要な基底 2 本を抽出し, 各特徴量を 2 次元埋め込みしている

Fig. 11 Comparison of low-dimensional embedding of image features.

表 2 画像分類を行った際の分類性能 (in %). 行: 分類器 (k-NN) の種類, 列: 低次元埋め込み手法の種類. Graph Reg. = グラフ正則化を用いた従来手法 [43]

Table 2 Performance of image classification.

	Orig.	Graph Reg.	PCA	LFDA	Prop.
	960 dim	960 dim	10 dim	10 dim	10 dim
Linear regression	39.7	39.4	33.6	38.7	42.5
kNN $k = 15$	37.9	-	35.7	42.5	45.1
$k = 20$	38.1	-	36.2	42.7	45.4
$k = 25$	38.2	-	36.4	42.8	45.6

に評価することを目的とする本実験では, 分類器として, 手法 Graph Reg. を除き, 比較的単純な k 近傍識別器および線形回帰を採用し, 低次元埋め込み特徴量の分布や構造がより直接的に分類性能に反映されるように実験を設計した. また, マルチラベル分類の設定を採用する本実験では, 各クラスで独立に (2 値) 分類器を設計した. 手法 Graph Reg. では, board diffusion graph を正則化のためのグラフ構造として利用した.

表 2 における低次元埋め込みを行わない Orig. と提案手法 Prop. との比較から, コンテキストを考慮した補助情報を利用することにより, 特徴量次元数が大幅に減っているにもかかわらず, より優れた分類性能を示すことが分かる. また, 従来法 Graph Reg. と提案手法 Prop. との比較から, 提案する低次元埋め込みがグラフ構造の利用方法としてよ

り適切であることが分かる. また, 提案手法と他の埋め込み手法との比較から, コンテキストを考慮した補助情報が分類性能向上に大きく寄与することが分かる.

図 12 および図 13 に, 埋め込み次元数を変化させたときの分類性能の変化の様子を示す. 図 12 では異なる特徴量埋め込み手法の比較を, 図 13 では異なる補助情報の構成方法の比較を行っている. 図 12 から, 提案手法がいずれの埋め込み次元においても他の手法よりも高い性能を示し, 非常に低い次元でほぼ最高の性能に到達していることが分かる. また, 図 13 から, わずかではあるが cluster information が他の構成方法よりも高い性能を示すことが見てとれる.

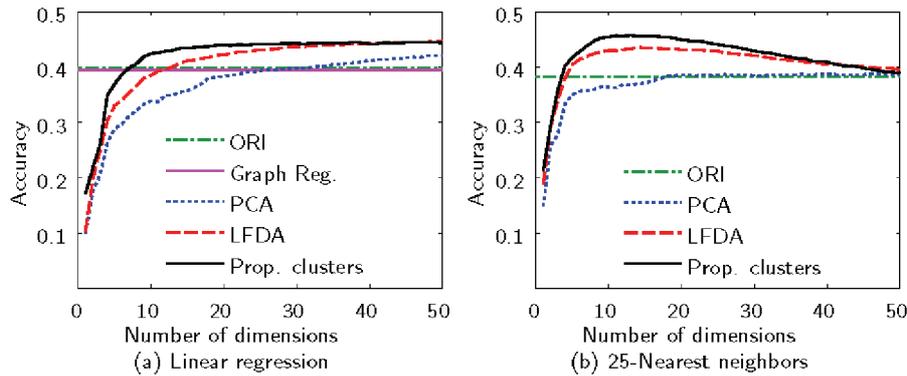


図 12 特徴量埋め込み手法の比較. 横軸：埋め込み次元数, 縦軸：画像分類の分類率
 Fig. 12 Comparison of feature embedding.

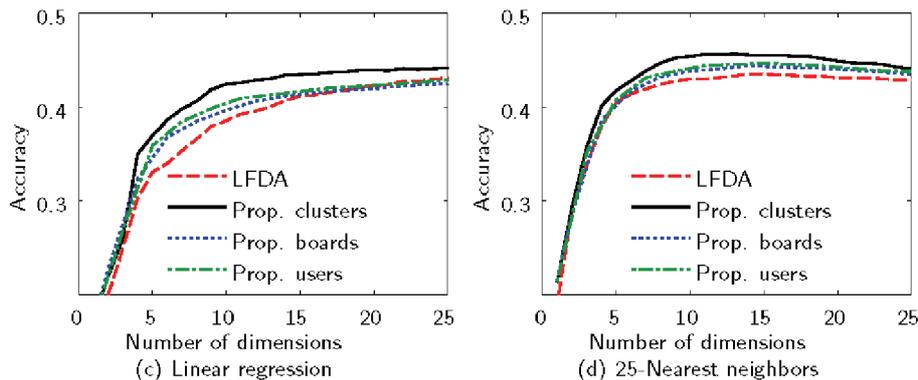


図 13 補助情報の比較. 横軸：埋め込み次元数, 縦軸：画像分類の分類率
 Fig. 13 Comparison of side information.

6. まとめ

本論文では、ソーシャルキュレーションのプロセスに着目することで、共通のコンテキストを持つ画像群を非常にシンプルにかつ大量に発見する手法を提案するとともに、発見した画像群を画像認識・検索のための類似性を自動的に獲得するための重要なステップとして、コンテキストの共通性を考慮した画像特徴量の低次元埋め込み手法を提案した。提案手法により、言語化が必ずしも容易であるとは限らないコンテキストの類似関係を、低次元画像特徴量の類似関係に変換できるようになるとともに、より高次の意味や概念を考慮した画像分類・認識・検索が可能となることを示した。画像分類をタスクとする実験により、提案手法の有効性を示した。

ソーシャルキュレーションのプロセスに着目する提案手法は、Pinterest 以外のキュレーションデータを用いた解析にも応用可能である。たとえば、クラウド型音楽ストリーミングサービスにおいて、ユーザが自分の好みに応じてプレイリストを構成するプロセスも、音楽を対象としたキュレーションと考えることができる。このことを利用すると、共通の嗜好性を持つ楽曲の集合を提案手法とほぼ同じような手法で大量に獲得可能であり、それに基づくプレイリストの自動生成や音楽推薦が可能になる [44]。この例に

代表される、他のキュレーションデータやメディアコンテンツへの適用可能性を広く探るとともに、獲得したコンテンツを表現するメディアコンテンツの特徴量成分についてより深い検討を進めることが、今後の研究の方向性として重要になるであろう。

謝辞 本研究の推進に多大なる御支援・ご協力をいただきくとともに有意義な研究議論およびコメントを頂戴した、NTT コミュニケーション科学基礎研究所藤本裕文氏、大和淳司主幹研究員、竹内孝研究員に感謝する。

参考文献

- [1] Kimura, A., Ishiguro, K., Marcos Alvarez, A., Kataoka, K., Murasaki, K. and Yamada, M.: Image context discovery from social curated contents, *Proc. ACM International Conference on Multimedia (ACMMM)*, pp.565-568 (2013).
- [2] Marcos Alvarez, A., Yamada, M. and Kimura, A.: Exploiting socially-generated side information in dimensionality reduction, *Proc. International Workshop on Socially-Aware Multimedia (SAM)*, pp.9-12 (2013).
- [3] Zheng, N., Li, Q., Liao, S. and Zhang, L.: Flickr group recommendation based on tensor decomposition, *ACM SIGIR*, pp.737-738 (2010).
- [4] Wang, J., Zhao, Z., Zhou, J., Wang, H., Cui, B. and Qi, G.: Recommending Flickr groups with social topic model, *Information Retrieval*, Vol.15, No.3-4, pp.278-295 (2012).

- [5] Negoescu, R.-A. and Gatica-Perez, D.: Modeling Flickr communities through probabilistic topic-based analysis, *IEEE Trans. MM*, Vol.12, No.5, pp.399–416 (2010).
- [6] Trevisiol, M., Chiarandini, L., Aiello, L.M. and Jaimes, A.: Image ranking based on user browsing behavior, *ACM SIGIR*, pp.445–454 (2012).
- [7] Ahern, S., Naaman, M., Nair, R. and Yang, J.: World explorer: Visualizing aggregate data from unstructured text in geo-referenced collections, *Proc. ACM/IEEE Joint Conference on Digital Library (JCDL)*, pp.1–10 (2007).
- [8] McAuley, J. and Leskovec, J.: Image labeling on a network: Using social-network metadata for image classification, *Proc. European Conference on Computer Vision (ECCV)*, pp.828–841 (2012).
- [9] Qi, G.-J., Aggarwal, C., Tian, Q., Ji, H. and Huang, T.: Exploring context and content links in social media: A latent space method, *IEEE Transactions on Pattern Analysis and Machine Intelligence (TMAP)*, Vol.34, No.5, pp.850–862 (2012).
- [10] Gong, Y., Wang, L., Hodosh, M., Hockenmaier, J. and Lazebnik, S.: Improving image-sentence embedding using large weakly annotated photo collections, *Proc. European Conference on Computer Vision (ECCV)*, pp.529–545 (2014).
- [11] Baluja, S., Seth, R., Sivakumar, D., Jing, Y., Yagnik, J., Kumar, S., Ravichandran, D. and Aly, M.: Video suggestion and discovery for YouTube: taking random walks through the view graph, *Proc. ACM International Conference on World Wide Web (WWW)*, pp.895–904 (2008).
- [12] Xu, T., Liu, D., Cao, H., Chen, E. and Tian, J.: Towards annotating media contents through social diffusion analysis, *Proc. IEEE International Conference on Data Mining (ICDM)*, pp.1158–1163 (2012).
- [13] Yamaguchi, K., Berg, T.L. and Ortiz, L.E.: Chic or social: Visual popularity analysis in online fashion networks, *Proc. ACM International Conference on Multimedia (ACMMM)*, pp.773–776 (2014).
- [14] Kimura, A.: Social curation as corpora for large-scale multimedia content analysis, *Proc. ACM International Conference on Multimedia Retrieval (ICMR)* (2013). available from <http://bit.ly/ICMR2013Curation>.
- [15] Kimura, A.: Large-scale cross-media analysis and mining from socially curated contents, *Progress in Informatics*, Vol.11, No.1, pp.19–30 (2014).
- [16] Dare, E. and Weinberg, L.: Algorithms for social curation: Designing and evaluating an embodied and subjectively situated visual art Interpretation and navigation system (VAINS), *Body, Space & Technology*, Vol.10, No.1, pp.1–29 (2010).
- [17] Gaskill, K.: In search of the Social - Toward an understanding of the Social Curator, PhD Thesis, University of Huddersfield (2010).
- [18] Ammann, R.: Reciprocity, social curation and the emergence of blogging: A study in community formation, *Procedia - Social and Behavioral Sciences*, Vol.22, No.0, pp.26–36 (2011).
- [19] Rhue, L.: The pins that bind: Preference affirmation, social norms, and networks on Pinterest, *Proc. International Conference on Information Systems (ICIS)*, p.68 (2012).
- [20] Zarro, M. and Hall, C.: Exploring social curation, *D-Lib Magazine*, Vol.18, No.11/12, p.1 (2012).
- [21] Gilbert, E., Bakhshi, S., Chang, S. and Terveen, L.: I need to try this?: A statistical overview of Pinterest, *Proc. ACM International Conference on Human Factors in Computing Systems (CHI)*, pp.2427–2436 (2013).
- [22] Hemphill, L., Otterbacher, J. and Shapiro, M.: What’s congress doing on Twitter?, *Proc. ACM International Conference on Computer Supported Cooperative Work (CWCW)*, pp.877–886 (2013).
- [23] Zhong, C., Shah, S., Sundaravadivelan, K. and Sastry, N.: Sharing the loves: Understanding the how and why of online content curation, *Proc. International AAAI Conference on Weblogs and Social Media (ICWSM)*, pp.1–4 (2013).
- [24] Ottoni, R., Pesce, J.P., Casas, D.L., Jr., G.F., Jr., W.M., Kumaraguru, P. and Almeida, V.: Ladies first: Analyzing gender roles and behaviors in Pinterest, *Proc. International AAAI Conference on Weblogs and Social Media (ICWSM)*, pp.457–465 (2013).
- [25] Zarro, M., Hall, C. and Forte, A.: Wedding dresses and wanted criminals: Pinterest.com as an infrastructure for repository building, *Proc. International AAAI Conference on Weblogs and Social Media (ICWSM)*, pp.650–658 (2013).
- [26] Greene, D. and Cunningham, P.: Discovering latent patterns from the analysis of user-curated movie lists, *ArXiv*, Vol.abs/1308.5125, pp.1–13 (2013).
- [27] Chang, S., Kumar, V., Gilbert, E. and Terveen, L.: Specialization, homophily, and gender in a social curation site: Finding from Pinterest, *Proc. ACM International Conference on Computer Supported Cooperative Work (CWCW)*, pp.674–686 (2014).
- [28] Duh, K., Hirao, T., Kimura, A., Ishiguro, K., Iwata, T. and Au Yeung, C.-M.: Creating stories: Social curation of Twitter messages, *Proc. International AAAI Conference on Weblogs and Social Media (ICWSM)*, pp.447–450 (2012).
- [29] Kimura, A., Duh, K., Hirao, T., Ishiguro, K., Iwata, T. and Au Yeung, A.: Creating stories from socially curated microblog messages, *IEICE Transactions on Information and Systems*, Vol.E97-D, No.6, pp.1557–1566 (2014).
- [30] Greene, D., Sheridan, G., Smyth, B. and Cunningham, P.: Aggregating content and network information to curate Twitter user lists, *Proc. ACM RecSys Workshop on Recommender Systems and the Social Web (RSWeb)*, pp.29–36 (2012).
- [31] Ishiguro, K., Kimura, A. and Takeuchi, K.: Towards automatic image understanding and mining via social curation, *ICDM* (2012).
- [32] Chen, T., Lu, D., Kan, M.-Y. and Cui, P.: Understanding and classifying image tweets, *Proc. ACM International Conference on Multimedia (ACMMM)*, pp.781–784 (2013).
- [33] Zoghbi, S., Vulić, I. and Moens, M.-F.: I pinned it. Where can I buy one like it?: Automatically linking Pinterest pins to online webshops, *Proc. CIKM Workshop on Data-driven User Behavioral Modelling and Mining from Social Media (DUBMOD)*, pp.9–12 (2013).
- [34] Blondel, V., Guillaume, J.-L., Lambiotte, R. and Lefebvre, E.: Fast unfolding of communities in large networks, *Journal of Statistical Mechanics*, Vol.2008, No.10, p.P10008 (2008).
- [35] Shiokawa, H., Fujiwara, Y. and Onizuka, M.: Fast algorithm for modularity-based graph clustering, *AAAI*, pp.1170–1176 (2013).

- [36] Vapnik, V. and Vashist, A.: A new learning paradigm: Learning using privileged information, *Neural Networks*, Vol.22, No.5-6, pp.544-557 (2009).
- [37] Sharmanska, V., Quadrianto, N. and Lampert, C.: Learning to rank using privileged information, *IEEE International Conference on Computer Vision (ICCV)*, pp.825-832 (2013).
- [38] Sugiyama, M.: Dimensionality reduction of multimodal labeled data by local Fisher discriminant analysis, *Journal of Machine Learning Research (JMLR)*, Vol.8, No.5, pp.1027-1061 (2007).
- [39] Zelnik-Manor, L. and Perona, P.: Self-tuning spectral clustering, *Proc. Conference on Neural Information Processing (NIPS)*, pp.1-8 (2004).
- [40] Kimura, A., Sugiyama, M., Sakano, H. and Kameoka, H.: Designing various multivariate analysis at will via generalized pairwise expression, *IPSJ Transactions on Mathematical Modeling and its Applications (TOM)*, Vol.6, No.1, pp.136-145 (2013).
- [41] Oliva, A. and Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope, *International Journal of Computer Vision (IJCV)*, Vol.42, No.3, pp.145-175 (2001).
- [42] Schölkopf, B. and Smola, A.J.: *Learning with kernels: Support vector machines, regularization, optimization, and beyond*, MIT Press (2001).
- [43] Mahajan, D.K. and Slaney, M.: Image classification using the web graph, *Proc. ACM International Conference on Multimedia (ACMMM)*, pp.991-994 (2010).
- [44] 木村昭悟, 柏野邦夫, 川西隆仁, 平松 薫: クラウドプレイリストネットワークを用いた音楽ストリーミング再生リストの自動生成, 電子情報通信学会技術報告 PRMU, pp.1-6 (2014).



木村 昭悟 (正会員)

平成 10 年東京工業大学工学部電気電子工学科卒業, 平成 12 年同大学大学院理工学研究科電気電子工学専攻修士課程修了. 同年日本電信電話 (株) 入社. 平成 19 年東京工業大学大学院理工学研究科集積システム専攻博士課程修了, 博士 (工学). 現在, 日本電信電話 (株) コミュニケーション科学基礎研究所メディア情報研究部主任研究員. パターン認識およびデータマイニングの研究開発に従事. IEEE, ACM SIGMM, SIGKDD, 電子情報通信学会各会員.



石黒 勝彦 (正会員)

平成 16 年東京大学工学部機械情報工学科卒業. 平成 18 年東京大学大学院情報理工系研究科知能機械情報学専攻修士課程修了. 同年, 日本電信電話 (株) 入社. 現在, 日本電信電話 (株) コミュニケーション科学基礎研究所研究主任. 博士 (工学). 統計的機械学習, 時系列データ解析, 関係データ解析の研究に従事. 電子情報通信学会, IEEE 各会員.

Alejandro Marcos Alvarez

2009 年リエージュ大学電気電子工学科卒業. 2011 年リエージュ大学大学院電気電子工学専攻修士課程修了. 現在, リエージュ大学大学院電気工学・計算機科学専攻博士課程在学中. 機械学習, 最適化, コンピュータビジョンの研究に従事. IEEE 学生会員.



山田 誠

平成 15 年会津大学コンピュータ理工学部ソフトウェア学科卒業, 平成 17 年コロラド州立大学工学部電気電子工学科修士課程修了. 平成 22 年総合研究大学院大学複合科学研究科統計数理専攻博士後期課程修了, 博士 (統計科学) 取得. 平成 22~24 年東京工業大学大学院情報理工学研究科産学官連携研究員, 平成 24~25 年 NTT コミュニケーション科学基礎研究所リサーチアソシエイト, 平成 25~27 年, Yahoo Labs, Research Scientist. 現在, 京都大学化学研究所助教. 機械学習およびデータマイニングの研究開発に従事. IEEE 会員.



片岡 香織

平成 10 年早稲田大学理工学部物理学科卒業, 平成 12 年同大学院理工学研究科物理および応用物理学専攻修士課程修了. 同年, 日本電信電話 (株) 入社. 現在, 日本電信電話 (株) NTT メディアインテリジェンス研究所画像メディアプロジェクト所属. パターン認識の研究開発に従事. 電子情報通信学会会員.



村崎 和彦

平成 21 年東京大学工学部機械情報工
学科卒業。平成 23 年同大学院情報
理工学系研究科知能機械情報学専攻修
士課程修了。同年，日本電信電話（株）
入社。現在，NTT メディアインテリ
ジェンス研究所画像メディアプロジェ

クト映像メディア処理技術グループ研究員。電子情報通信
学会，画像電子学会各会員。