

Auto SAGURU: 写真を介したコミュニケーションのための 探し合い支援システムの提案

武田 十季^{1,2,a)} 牛尼 剛聡^{3,b)}

概要: 我々は、2者間の対面コミュニケーションにおいて、両者が保有する写真の中から、写真に付与されるキーワード、撮影日時・場所のタグ情報を用いて、コミュニケーションの観点から高い満足度を得られる写真を探り出して提示するシステム Auto SAGURU を開発中である。本論文では、写真を用いた対面コミュニケーションの実験から、話題として写真を提供する「提供者」が、情報の「受け手」である相手の興味を理解していないことが、両者の満足度が不一致となる要因の一つであることを明らかにする。そして、この問題点を解決するために、対話の開始前に相互の写真情報の関連性に基づいてランキングし、写真が選択されると、その写真との関連性に基づいて写真を再ランキングする手法を提案する。本手法では、Random Walk with Restart 法を利用して、コミュニケーションの話題に応じた動的なランキングを実現する。

Auto SAGURU: Supporting Communication with Photos

TOKI TAKEDA^{1,2,a)} TAKETOSHI USHIAMA^{3,b)}

Abstract: We propose a system for promoting the face-to-face communication between two users, which we call “Auto SAGURU”, by ranking photos the users have. Our proposed system finds the photos that are useful in terms of communication from the users’ photos by using meta-information assigned with the photos, such as keyword, date, and geographical information. We conducted an experimental study about a face-to-face communication using photos and confirmed that both users cannot be satisfied with the communication when an information provider does not understand the interest of a recipient. To solve such problems, our system ranks photos based on the their relevance before communication and re-ranks photos after a user select a photo. We used the Random Walk with Restart (RWR) algorithm for dynamically re-ranking photos based on the relevance regarding the selected photo.

1. はじめに

近年、デジタルカメラの普及と、ソーシャルネットワークサービス (SNS) などの写真の共有の場の増加に伴い、写真を介したコミュニケーションが盛んに行われている。膨

大化する写真の利用に向けて、時間順や位置情報に基づいた写真の整理や、顔認識技術による人ごとにカテゴライズ [1] するサービスも登場してきた。写真は過去の振り返りに有効なコンテンツであり [2][3]、私たちはしばしば写真を介して他人と過去の体験を共有する。写真を介したコミュニケーションは実世界上の日常生活でもしばしば行われる。例えば、人は人と久しぶりに出会った友人と写真を用いて近状について伝え合い、会話が行われることがある。このような対面のコミュニケーションでは、私たちは相手に合わせて写真や話題を選ぶことが多い。しかし、相手の好みが分からない場合や、伝えたい体験が明確でない場合

¹ 九州大学 大学院 芸術工学府
Kyushu University

² NTT サービスエボリューション研究所
NTT Service Evolution Laboratories

³ 九州大学 大学院 芸術工学研究院
Kyushu University

a) tokitokibu@gmail.com

b) ushiama@design.kyushu-u.ac.jp



図 1 Auto SAGURU の利用イメージ

は、写真を検索するための検索キーワード自体を思いつけないため、時間情報や位置情報、またカテゴリ名を効果的に活用することはできない。このように、伝える写真が明確に決定していない時に、話題のきっかけとなる写真を見つけて出すまでには、多くの時間と労力を要する。

本研究では、スマートフォンやタブレット型コンピュータなどを利用して、写真を介した対面コミュニケーションにおいて、伝える写真が明確に決定していない時に、対話者双方がコミュニケーションの観点で高い満足度を得られる写真を推薦する *Auto SAGURU* を提案する (図 1)。人は互いの興味を探り合いながら話題のきっかけとなる写真を選択していることに着目して、この探し合いを *Auto SAGURU* が自動的に行うことを目指す。各対話者が保有する写真には各自の興味が含まれていると考え、*Auto SAGURU* は相互の写真の関連性に基づいて写真のランキング付けを行う。

本稿では、まず、写真と話題を提供する「提供者」は、受け手に合わせた写真の選択を行うという仮説をたて、写真を利用した対話の実験を通じて検証を行う。次に、提供者と話を聞く「受け手」の満足はどのような時に満たされるか、また、どのような時に満足の不一致が生じるかを明らかにし、双方に高い満足を与えられる写真のランキングシステムの検討を行う。最後に、システム実現に向けて *Random Walk with Restart (RWR)* を適用した *Auto SAGURU* の写真のランキング法を提案する。

2. 関連研究

写真の管理に関する研究として、画像処理によってメタデータを付与するアノテーション技術がある [4][5]。これらは、膨大な写真をキーワードによって検索することを可能としているが、ユーザが検索する写真を明確に決めていることを前提としている。Vennelakanti[6]らは、写真を介したコミュニケーション時の映像や音声などを用いて、対話時の様子をメタデータとして写真に付与し、新たな写真の活用方法を提案している。例えば、ある写真に対して、誰がどれくらい閲覧したか、どの部分を指しながらどのようなキーワードを発したかなどのデータを付与している。し

かしこの研究では、コミュニケーション時の状況を取得することを目的としており、コミュニケーションに適した写真を提示することには取り組まれていない。Hilligesら [7]は、テーブルトップのディスプレイを用いた写真の共有時に、対話の進行に合わせて写真を探し出すインターフェース設計を行っている。物理的なつまみコントローラを用いて直感的に写真を選び出すことを可能としているが、この研究では、提供者の操作性に主眼を置いており、受け手側の満足については考慮していない。写真を閲覧する場面を対象として、文脈に基づいた次の写真を推薦する研究がある [8]。これは写真の画像情報を用いて、現在閲覧している写真と類似した写真を推薦するものである。しかしコミュニケーションにおいては、必ずしも次に提示する写真として、画像の特徴が類似したものが良いとは限らない。本研究では、複数人コミュニケーションにおいて、各自が保有する写真の情報を元に、相互が満足なコミュニケーションを行える写真を提示することを目的とする。

3. 写真を介したコミュニケーション

3.1 写真の選択基準の分析

本節では、「提供者が写真の選択を行う時、受け手に合わせた写真の選択を行う」という仮説に対して、写真を利用した対面コミュニケーションの実験を行う。

被験者は4名であり2グループに分けて、互いに仲の良い友人をペアとさせた。各グループの中で、一方を提供者役、もう一方を受け手役とした。提供者には自分のスマートフォンの中から過去の写真を選択して、撮影時の体験について受け手に伝えさせた。5枚以上の写真について対話を行う指示を与えて約10分程度対話を行った後、相互の役割を交代させて同じことを実行した。本実験では、提供者の写真の選択基準を観察するために、全ての写真を公平に選択可能な状況下にする必要があったため、被験者にはあらかじめ、端末内の写真について振り返りを行い、一覧のどのあたりにどの写真があるかを覚えさせておいた。

実験終了後、被験者には写真を選択した理由を、写真1枚ごとに回答させた。得られた回答から提供者の選択理由の傾向を分類したところ、受け手の趣味・趣向などの知識を用いた「相手主体とした選択基準」と、提供者にとって印象的だった体験を共有する「自分主体とした選択基準」が存在した。また、いずれの場合においても、提供者は選択した写真に関する話題について継続するか話題変更を行うかを、受け手の反応を窺って判断する、「相手主体とした話題の継続・変更」の行動が抽出された。以上のことから、提供者は受け手が興味を持ちそうな写真を選択しながら「探す」行為を行っていると考えられる。つまり、本研究から「提供者が写真の選択を行う時、受け手に合わせた写真の選択を行う」仮説に従う傾向があることを確認できた。

3.2 提供者と受け手の満足度を考慮した写真推薦の条件

前節の実験では、ユーザが写真を選択する際の、相手の興味を探る行為を観測することができた。本節では、写真を選択して対話を行った際の、提供者と受け手の満足度に影響を与える要因について、実験を通じて明らかにする。そして、システムが相互の満足度を満たす写真を推薦するために必要な条件をまとめる。

被験者は、計6名を仲の良い友人同士をペアにして3グループに分けた。また、実験内容、及び、指示については、前節と同様に行った。対話終了後、各写真に対して、提供者と受け手の満足度を5段階(1. とても不満, 2. やや不安, 3. どちらでもない, 4. やや満足, 5. とても満足)で主観評価させた。

各写真に対する提供者と受け手の満足度スコアを比較すると、スコアが完全に一致した写真と、不一致となった写真が存在した。満足度の不一致においては、提供者のみが満足している場合と、受け手のみが満足している場合の2種類が存在した。表1に、それぞれの満足度と、インタビューより得られた、提供者の満足度スコアが高い/低い理由と受け手の満足度スコアが高い/低い理由の例を示す。

ここで、提供者と受け手の満足度スコアが一致した際の、相互の評価理由の例として次のことがあげられる。

- 提供者は、あらかじめ受け手の趣味を知っており、それに合わせた話題を提供したため話が盛り上がり相互に満足した
- 提供者が、自分の楽しかった体験や趣味を話したところ、受け手と共通の体験や趣味と一致したため、相互に新たな共通点を見出すことができて満足した

また、提供者と受け手の満足度スコアが不一致になった際の、相互の評価理由を比較したところ、次のような2者間の満足のすれ違いが見られた。

- 提供者は自分が楽しかった体験を伝えられて満足したが、受け手は興味がなかったため満足しなかった
- 提供者は、受け手が話題に興味を持ったか判断できなかったため満足しなかったが、実は受け手は興味があり満足していた

さらに、提供者が話している最中に、受け手も関連する写真を探し出して提示する様子や、提供者と受け手の担当が交代にした際に、交代前の話題に関連する写真を選択する様子も見受けられた。これは、相互の共通点に基づいて写真を選択していると考えられる。実験後のインタビューでも、相手が提示した写真を通して、相手も自分と同じ体験をしていることを知ることができて満足したという意見が得られた。

以上より、相互の満足度を満たす写真の推薦条件として次の2点があげられる。

条件 I 提供者は、受け手の写真に対する興味があらかじめ分かり、受け手が興味を持つ写真を提示できること

条件 II 提供者に選択された写真と関連する写真を両者が提示できること

条件 I の実現は、満足度の不一致を解決すると期待される。つまり、受け手の満足度向上のみでなく、提供者にとって、受け手の興味が分かることは、話が盛り上がり確信を持って写真を選ぶことを可能とすることから、提供者の満足度向上にも繋がると考えられる。条件 II の実現は、相互の共通点の気づきを素早く与えると共に、対話の文脈に基づいた話題提供を可能とすることが期待される。

4. Auto SAGURU の概要

本研究では、対話者双方が高い満足度を得られる写真を探り出して提示するシステム、Auto SAGURU を提案する。本章では、前章で述べた条件 I, II を満たすためのシステムの概要と実現方法について述べる。

4.1 システムの概要

受け手が興味を持つ理由として「自分と関連性・共通点があり興味がある」と「自分と関連性はないが興味がある」という2点が考えられる。前者の場合は、同じ場所を訪問したこと、同じジャンルの食べ物を食べたこと、共通の趣味、などがあげられる。後者の場合は、行ったことはないが自分も行ってみたいと思うこと、初めて見る食べ物だが食べてみたいと思うこと、などがあげられる。本研究では、前者の関連性・共通点があって興味を持つことに焦点を当てる。

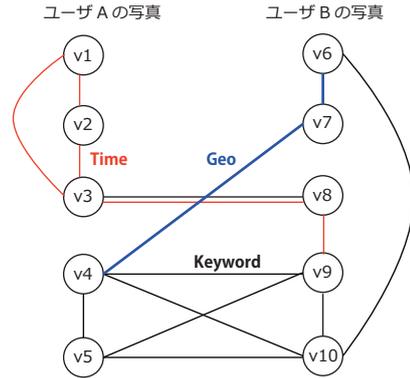
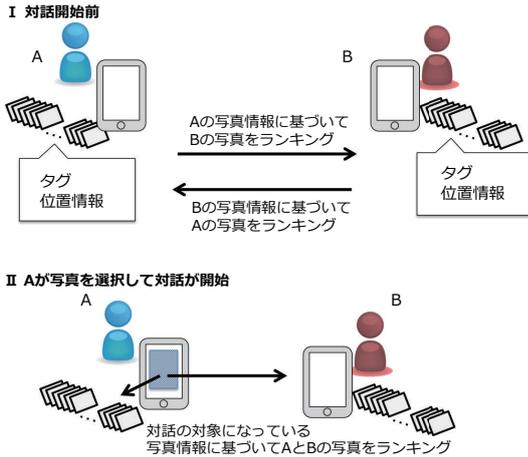
対話者は、各自の端末内に写真を保有しており、各写真にはあらかじめ撮影場所の位置・時刻情報や写真を表すキーワードなどのタグが付与されていることを前提とする。いま、ユーザ A とユーザ B が対話を開始しようとしているとする(図 2.I)。条件 I を満たすために、Auto SAGURU は各ユーザの端末内で、相互の関係性が強い写真を上位に提示する。つまり、ユーザ A とユーザ B の写真のタグ情報に基づいて、ユーザ A とユーザ B の各端末内の写真をランキング付けする。対話者は、図 1 のように、ランキング付けされた写真をスライドしながら探すことができる。提供する写真を見つけたら、写真をクリックするなどの選択行為によって対話開始と判定される。ここで、ユーザ A が写真を選択し対話が始まったとする(図 2.II)。このとき Auto SAGURU は、条件 II を満たすために、選択された写真を元に、ユーザ A とユーザ B の端末内の写真を再ランキングする。

4.2 Random Walk with Restart による写真のランキング

2者間の写真提示コミュニケーションにおいて、対話前の相互の写真のランキングと、選択した写真を元に相互の端末内の写真を再ランキングするために、写真をノード、

表 1 提供者と受け手の満足度、及び、満足度が高い/低い時の評価理由

写真の内容	話し手満足度	話し手の評価の理由	受け手満足度	受け手の評価の理由
うなぎ定食	4	うなぎが美味しくて伝えたかったため	2	特に珍しい話ではなかったため
友人との旅行	4	楽しい旅行の思い出だったため	5	友人の友人は、どのような人なのか興味を持ったため
妹からのプレゼント	5	自分の家族について紹介できたことで満足感があった	5	今後使えるネタと感じながら聞いていたため。また、相手の新しい一面を知れた
サイクリング	5	相手も同じ趣味であることを知っていて共感してくれただと思った。自分ならではの情報も発信できた	5	自分の趣味に合った話題だった。知らなかった情報も得られた
料理	2	相手が興味を持ったか分からなかった	5	最近自炊をしようと考えていたため興味があった
展示会の入り口	3	その先の話題にあまりつながらなかったため	3	詳細な様子が分からなかった
展示会の作品	2	大した情報ではなく、ただの自慢話のように感じた	4	作品の仕組みに興味があった



$$e_{i,j} = \begin{cases} 1 & \text{if } s_{i,j} < \theta_{time} \text{ or } g_{i,j} < \theta_{geo} \\ 0 & \text{otherwise} \end{cases} \quad (2)$$

写真のタグ情報からノード間にエッジが張られるグラフを考え、Random Walk with Restart (RWR) [9] によるランキング手法を提案する。

4.2.1 写真グラフの構成

本章では、ユーザ A, B が保有する写真から構成されるグラフ $G = (V, E)$ について示す。まず、ノード集合 V については、ユーザ A, B が保有する写真の集合をそれぞれ V_A, V_B とし、 $V = V_A \cup V_B$ と定義する。次に、2つの写真 v_i, v_j の間のエッジについて、 $e_{i,j} \in E = 1$ のとき、写真 v_i, v_j の間にエッジが張られていると定義する。

写真 $v_i \in V$ に付与された N_i 個のキーワードタグの集合を $T_i = \{t_k \mid k = 1, \dots, N_i\}$ としたとき、 $e_{i,j}$ の値は式 (2) とする。

$$e_{i,j} = \begin{cases} 1 & \text{if exists } (t_k \in T_i, t_l \in T_j) \text{ such that } t_k = t_l \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

次に、キーワードタグによりエッジが張られなかった ($e_{i,j} = 0$) ノード間を対象として、撮影時刻タグ s_i, s_l の差と位置情報タグ g_i, g_j の差を用いてエッジ $e_{i,j}$ を張る。いま、時刻情報の2点間の時間差を $s_{i,j}$ 、位置情報の2点間の距離を $g_{i,j}$ とした時、 $e_{i,j}$ の値は式 (2) とする。

図 3 に写真グラフ $G = (V, E)$ の例を示す。写真 $V_1 \sim V_5$ はユーザ A が保有する写真、写真 $V_6 \sim V_{10}$ はユーザ B が保有する写真を表す。結ばれた線はエッジを表しており、赤が時刻情報に基づいて張られたエッジ、青が位置情報に基づいて張られたエッジ、黒がキーワードに基づいて張られたエッジである。例えば、写真 V_3 と写真 V_8 は、撮影時刻が近いタグと、同じキーワードであるタグを持つ関係にある。

4.2.2 Random Walk with Restart モデルの応用

Random Walk with Restart (RWR) はグラフのノードの類似度を計算する方法である [9]。RWR では問い合わせノード v_q を始点とするランダムウォークを行い、ノードに到達するたびに一定の確率 c で問い合わせノードに戻る。ここで、 \mathbf{p} を $n \times 1$ 行列とし、要素 p_i をノード v_i における存在確率とする。また、 \mathbf{q} を $n \times 1$ 行列とし、ノード v_q に対応する要素 q_q を 1、その他の要素を 0 とする。そして、 \mathbf{A} を列が正規化されたグラフの隣接行列とする。すなわち、要素 $A_{i,j}$ は式 (3) に示す様にノード v_j からノード v_i へランダムウォークする確率を表す。

$$A_{i,j} = e_{i,j} / \left(\sum_k e_{k,j} \right) \quad (3)$$

このとき、定常状態における各ノードにおける存在確率 \mathbf{p}

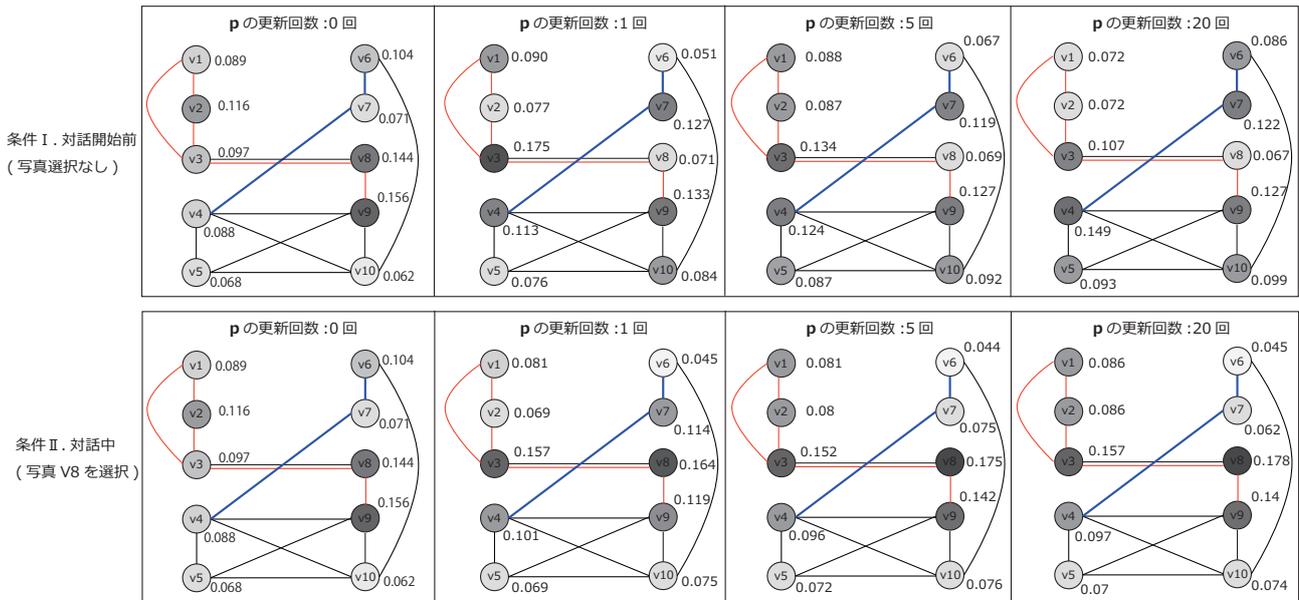


図 4 RWR の実行例

は、式 (4) を再起的に収束するまで繰り返すことで計算できる。

$$p = (1 - c) \cdot Ap + c \cdot q \quad (4)$$

提案手法では、条件 I. 対話開始前 (写真選択無し) と、条件 II. 対話中 (写真選択あり) において、それぞれ最適化を行う必要がある。I. 対話開始前は、問い合わせノードが存在しない状態であるため $c = 0$ としてノードの存在確率 p を計算する。そして、条件 II. 対話中は、問い合わせノード v_q を現在提示されている写真とし、写真提示の度にノードの存在確率 p を再計算する。このノードの存在確率を、2 者の写真の集合 V の中から次に提示すべき写真のスコアと考え、推薦候補の写真のランキングに用いる。

図 3 の写真グラフを用いて p の更新を行った際の、更新回数が 0 回、1 回、5 回、20 回の時のノード存在確率を図 4 に示す。図 4 上は対話開始前 (条件 I) に計算されるノード存在確率の更新過程である。まず、対話開始前はランダムウォーク (リスタートなし) によりノードの存在確率 p が計算されるため、 p の更新回数の増加に伴い、他のノードと多くのエッジで繋がっているノード、すなわち両者の共通の話題となりやすい写真 (V_4, V_9, V_7 など) のスコアが高くなる。次に、図 4 下は、ユーザ B が写真 V_8 を選択して対話が始まった時 (条件 II) に計算されるノード存在確率を示す。対話中は、問い合わせノード (V_8) と類似するノードの存在確率が高くなるため、 V_8 とエッジで繋がっている写真 (V_3, V_9) のスコアが高まり、ネットワーク上で遷移しにくい写真 V_6 などのスコアが下がる。

5. まとめと課題

本研究では、写真を用いた対面コミュニケーションにお

いて、伝える写真が明確に決定していない時に、対話者双方がコミュニケーションの観点で高い満足を得られる写真を優先的に提示する *Auto SAGURU* を提案した。写真を介したコミュニケーションにおいてユーザが互いに興味を探り合って写真を選択する行為について仮説を立て、実験により検証した。また、同実験において写真の提供者と受け手の間で満足度が不一致する状況が観測された。まず、提供者は自らの印象的な体験を伝えられたことに満足した一方で、受け手はその話題に興味をなかったケースが抽出された。次に、提供者は相手に興味がないと思った一方で、受け手は興味があったケースが抽出された。そして、提供者が提示した写真と関連する写真を受け手が提示する行為も観測された。以上の結果を元に、条件 I. 受け手の興味に基づいた写真を提示できること、条件 II. 提供者に選択された写真と関連する写真を両者が提示できること、を *Auto SAGURU* が満たすランキングシステムの条件とした。さらに、I と II の条件を実現するために、Random Walk with Restart (RWR) のモデルを適用することを提案した。

今後は、提案した *Auto SAGURU* の構築と被験者実験による評価を行う。また、写真に付与されているタグ情報以外の情報を取り入れることを考える。例えば、ブログや SNS などからユーザの興味の情報を取得することで、よりユーザの興味に適した写真を探り出すことが可能となると期待される。さらに、対話するのユーザ相互の間柄も考慮する必要があると考えている。本研究では、仲の良い友人を対象とした実験を行ったが、年の離れた上司と部下や先輩と後輩の間では、話題の選択基準が異なってくる可能性が考えられる。この点についても今後の課題とする。

参考文献

- [1] <http://picasa.google.com>
- [2] Steve, H., Lyndsay, W., Emma, B., Shahram, I., James, S., Alex, B., Gavin, S., Narinder, K., and Ken, W. Sense-Cam: A Retrospective Memory Aid UbiComp, LNCS 4206, pp.177-193, 2006.
- [3] Matthew L. L., and Anind K. D. Lifelogging Memory Appliance for People with Episodic Memory Impairment. UbiComp'08, September 21-24, 2008.
- [4] Lemstrom, K., Korte, J., Kuusi, P., Kyheroinen, P., Paivakumpu, and P. PICSearch: A Platform for Image Content-based Searching Algorithm. In The Sixth Int. Conf. in Central Europe on Computer Graphics and Visualization, Plzen, 1998.
- [5] Lieberman, H., Rozenweig, E., and Singh, and P. Aria: An Agent for Annotating and Retrieving Images. IEEE Computer 34, 7 (July 2001), 57-62.
- [6] Ramadevi, V., Prasenjit, D., Ankit, S., and Phanindra P. The Picture Says It All! Multimodal Interactions and Interaction Metadata. ICMI'11, p.14-18, 2011.
- [7] Otmars, H., David S. K. Getting Sidetracked: Display Design and Occasioning Photo-Talk with the Photohe-lix. CHI 2009, p.4-9, 2009.
- [8] D. Liu, S. Yan, and H. Zhang. Next photo please: towards visually consistent sequential photo browsing. MM'11, Proceedings of the 19th ACM international conference on Multimedia, pp.973-976 (2011)
- [9] H. Tong and C. Faloutsos. Center-piece subgraphs: problem definition and fast solutions. In *KDD*, pages 404-413, 2006.