

PhotoLoop: 写真閲覧時の活動を利用したアノテーションシステム

PhotoLoop: An annotation system using users' activities while watching slideshow

渡邊 恵太 塚田 浩二 安村通晃*

Summary. デジタル写真の楽しみ方として、イベント後に友人達と撮影した写真をスライドショーで閲覧し、写真やイベントの内容について盛り上がることは多い。本研究では、こうしたスライドショーの前でのユーザの活動に着目し、デジタル写真をリッチにするアノテーションを容易に実現するシステム PhotoLoop を提案する。PhotoLoop は、スライドショー閲覧時に鑑賞者の様子をカメラ/マイクなどで記録し、ビデオ/音声、操作履歴などのメタデータを付きスライドショーを生成することができる。閲覧者は特別な処理を必要とせず、友人達と自然に語りながらスライドショーを楽しむだけで、写真の背景や場の雰囲気や記録することができる。本稿では、PhotoLoop のコンセプト、実装について述べ、運用経験を踏まえた議論を行う。

1 はじめに

近年デジタルカメラが普及し、手軽に写真撮影が行なえるようになり、個人が多くの写真を所有するようになってきた。一方で、蓄積された膨大な量の写真をどうやって見ていくかなど、蓄積された写真をうまく活用していくことが課題となっている。

デジタル写真の活用の仕方としては、印刷やスライドショー、Web 上で共有するなどの楽しみ方があり、今後もさまざまなスタイルでの写真の楽しみ方が広がると考えられる。

本研究では、デジタル写真の楽しみ方の1つであるスライドショーに着目し、スライドショーの閲覧時のユーザの活動をシステムが自動的に記録することで、自動的にアノテーション付加される新しいスライドショーシステム「PhotoLoop」を試作した。

1.1 スライドショーとユーザー

写真の楽しみ方として、スライドショーという手法がある。スライドショーは保存された写真を特定の時間間隔で切り替えて提示していく手法である。デジタル写真は、手軽さゆえに明確な意図を持たずにシャッターを切ることも多く、それ1枚で見ても面白さは薄い。スライドショーのような順次写真を提示する映像的な見方では、時間の経過や状況の変化を伝えるような役割を果たす。したがって、スライドショーは見せ方としてもデジタル写真に適している。また自動的に写真が切り替わるため、ユーザの負担が少なく、膨大な量の写真を閲覧する手法としても適している。さらに、スライドショーは大

きいスクリーンに提示して閲覧することが多いため、複数人で写真を楽しむ手法としても適している。

スライドショーは、何らかのイベントが終わったあとに、イベントに参加した仲間同士で見ることが多い。こうした場合、映し出される写真に対して、そこに写っている物の説明をしたり、状況の説明を行うことが推察される。

そこで、本研究ではまず複数人がスライドショーを閲覧する状況においてどのような種類の発言があるか、スライドショーを閲覧する様子をビデオに録画し、その分析を行った。

1.2 スライドショーにおけるユーザの分析

スライドショー中、閲覧者の発言内容の多くは、そのイベントの「感想」であった。

他者に対して「これは何であったか」「なぜ撮ったのか」など疑問を問いかけることも多かった。特にそのイベントに参加しなかった人は、そうした質問が多くなる傾向があり、参加者も、より詳細な説明を行う傾向があった。一人だけが撮影者でその他が非参加者の場合、撮影者は自慢のように語る傾向もあった。

逆に、全員参加者の場合は、「～だったねえ」というような確認(共感)を求めるような発言が目立った。このように発言内容には、その写真に対する説明と、個人の思い出や心境を語るような2種類がみられた。

また、変わった状況の写真が写る場合、「誰が撮った写真」という質問がなされる場合があり、誰が撮影したか自明である場合は、閲覧者は撮影者に対して意見(つつこみ)を述べる場合があった。また写真自体が美しい場合は、いい写真と評価され、撮影者までも評価されることもあった。

このように、スライドショーのプロセスの多くは、

Copyright is held by the author(s).

* Keita Watanabe and Michiaki Yasumura, 慶應義塾大学大学院 政策・メディア研究科, Koji Tsukada, 産業技術総合研究所



図 1. PhotoLoop を利用してスライドショーを閲覧する様子



図 2. PhotoLoop のスクリーン

主に質問と共感を繰り返しながら、進められる。スライドショーを目の前にすると、多くの発言をする傾向が見られ、具体的なコメントがない場合でも、笑ったりすることでその写真に対する考えを表現する。そのほか、写真に対しては「この」「あの」などの指示語を使う傾向が多く見られた。

一般的にスライドショーに対してユーザは、受動的な関わり方をしていると思われがちだが、このような仲間同士でスライドショーを閲覧する場合、積極的に「語り」を行う傾向が見られた。

本研究では、このようなスライドショーの前での語らいなどの積極的な活動を、その写真に対するメタデータとして捉えた。

そこで、スライドショー閲覧時のユーザー間の自然なコミュニケーションをシステムが自動的に記録し、それをスライドショーに対応づけて再生することができるシステム PhotoLoop を試作した。PhotoLoop を使うことによって、自然に語らいながらスライドショーを見るだけで、メタデータを記録したり、ナレーション付きスライドショーを生成できる。

2 試作システム: PhotoLoop

2.1 PhotoLoop とは

仲間内でスライドショーを楽しむだけで、自動的にナレーション付きのリッチな情報を含むスライドショーコンテンツや、メタデータが付加され、再利用価値の高いコンテンツが自動的に生成されるシステムである。

2.2 ナレーション映像付加

ナレーション付きスライドショーは、Apple 社の iPhoto や iMovie など連携することで、作成できる。しかし、このような作業を行う場合、質の高い作品が制作可能である一方、莫大な製作時間が必要になる問題がある。また、制作に関わるのは一人であることが多いため、主観的になりがちである。

PhotoLoop はスライドショーを楽しく仲間と閲覧

するだけでナレーション映像付きのスライドショーが自動生成されるため、ユーザが行うことは、写真についてお互いに語るだけでよい。したがって、コンテンツ制作のために特別に時間を浪費する必要がなく、手軽である。また、複数人の視点があるために、内容もさまざまな視点のナレーションが含まれる。一方、写真のナレーションとして見た場合の質という観点では、時間をかけて一人で制作した方がよい物ができる可能性もある。この点は議論で述べる。

2.3 メタデータ付加

盛り上がり度メタデータ: PhotoLoop では音声入力レベル、映像入力の変化量を記録することによって、スライドショーの盛り上がり度を時間の変化と共に記録することができる。盛り上がり度とは、大きな声を発話したり発話数が多い場合、人の動きが激しい場合である。これにより、スライドショーの中でどの写真がもっとも盛り上がったかがわかるようになる。

注目メタデータ: スライドショー閲覧時ユーザは指示語を使うとともに、指を指して、説明したり質問することがあった。この指さし行為をスライドショー上で実現するために、PhotoLoop では指さし機能を実装した。

指さし機能とは、マウスやジャイロマウスを利用し写真内にポインティングできる機能である。画面内に表示される矢印アイコンを移動させ、クリック操作を行うことで、その位置に矢印が付加できる。なお、再生する際は記録された位置をつつづくように往復運動するアニメーションして表示される。

指さし機能によって「どの写真がより注目を引くものであったか」がわかるようになる。また、スライドショー全体で、何回指さしを行なったかによって、そのイベントの重要度を判定することができる。

切り替えメタデータ: スライドショーは基本的に自動再生されるが、ユーザの操作によって、次の写真へ即時に切り替えたり、また前の写真に戻ることが

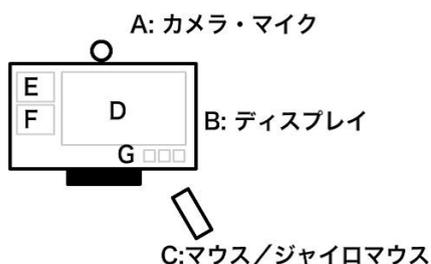


図 3. PhotoLoop のシステム構成: (D) スライドショー提示, (E) カメラ映像表示, (F) ナレーション映像, (G) コントロールボタン

できる。またスライドショーは一時的停止することができる。このようなユーザの切り替えや一時停止によって、その写真を見ていた時間がわかる。したがって、写真を見ていた時間から、その写真の重要度を判別することが可能になるため、切り替え操作をメタデータとして利用することができる。

2.4 システム構成

PhotoLoop は一般的な PC, ディスプレイ, PC カメラとマイクから構成される (図 3)。ディスプレイには PhotoLoop のソフトウェアが全画面で表示される。ディスプレイ内には、図 3(D) に写真が順次クロスフェードのエフェクトで提示される。図 3(E) にはカメラから入力された映像がリアルタイムで提示される。図 3(F) には記録された映像が提示される。図 3(G) は Pause, Prev, Next のボタンが並びスライドショーを操作できる。

3 議論

3.1 試用による考察

ここでは、PhotoLoop を実際に利用した結果から考察を述べる。

3.1.1 PhotoLoop による写真閲覧

PhotoLoop で写真を閲覧する場合、指差し機能を除けば一般的なスライドショーを閲覧することと変わりはない。ただし、画面上に PC カメラから入力された映像がリアルタイムに表示されるため、その様子が気になるという意見があった。こういった撮られているという意識が、そこでの発話内容にどのように影響するかは今後検討していく必要がある。

3.1.2 ナレーション付きスライドショー閲覧

記録された映像および音声と同時に再生 (ナレーション付きスライドショー) し、閲覧者自身が見た場合、そのときの笑っている様子や発言から、再

度笑ってしまったり、発言している自分に対してつこみをいれるなどの様子がみてとれた。自分が何を発言したのかを覚えていないためか、ナレーション付けを行なった直後に見返すだけでも、楽しさがあり、再び盛り上がる傾向があった。

そして、第 3 者が、ナレーション付きのスライドショーを閲覧した場合、そのイベントに参加してなくても、状況がわかるといった意見が得られた。また、矢印がある場合とない場合では、矢印がある際の方が閲覧者の注目がすぐそこに向かうため、音声による会話の内容だけではわかりにくい指示語に対してより明示的な注目を引き付ける。

ナレーション付きスライドショーを閲覧する際も、さらにその様子を記録すべきかどうかは今後の課題である。現状のシステムではナレーションがスピーカーから提示されるため、再度録音しようとするとうハウリングが起る問題がある。エコーキャンセル機能搭載のスピーカーフォン¹を利用したが、新しい音声だけ厳密に付加することはできなかった。解決策としては閲覧者全員に、ヘッドフォンを利用してもらうか、骨伝導マイクロフォンを利用することで、新しい音声だけを抽出することができるが、手軽さに欠ける点が問題である。またさらに、見るたびに音声ナレーションが付加するとすると、音声が多重化されるため、重ねるたびに聞き取りにくいものになる。したがって、再生時は音声を重ねずに最後に付加されたものだけを再生するなどの工夫が必要であるだろう。

あるいは、閲覧する回数 (順番) によって、アノテーションの権限を弱めるルールを用いる手法も考えられる。最初の閲覧者は、音声付加や、矢印付加、切り替えタイミングの制御ができるが、2 回目以降は、音声付加や切り替えタイミングはできないようにし、矢印だけは付加できるようにするなどが考えられる。

3.1.3 ナレーションの性質と可能性

普段の写真は撮影者やそのイベントに参加した人がいなければ、その写真の撮影意図やその前後の状況のストーリーはわからない。PhotoLoop によるナレーション付きスライドショーを Web で公開すれば、その写真の撮影意図や周辺の状態を理解する手がかりを与えてくれる。これは、ナレーションに含まれる内容が複数ユーザの体験、感想、対話というリッチな情報を含むためである。iMovie などで写真にナレーションを意図的に発話する場合に比べて、ナレーションは構造化されていないが、仲間内で見る分には十分楽しめる性質がある。

3.2 応用

ここでは、PhotoLoop の応用について述べる。

¹ Polycom communicator PPVOIP-C100S

3.2.1 写真の閲覧までがイベント・思い出

旅行へ行く際にカメラを持って行くことは一般的である。旅行へ行って帰るとそのイベントは終了とみなされるが、写真の観点から考えると、仲間で写真を見て楽しく語りながら振り返ることも、その旅行のイベントのひとつである。しかしながら、このような会話や雰囲気は記録されることは少なく、またカメラを持ち出してまで写真について語る自分たちを記録しようとする人は少ない。PhotoLoopでは、ユーザが特別に記録のために何かする必要はないため、負担はほとんどない。したがって、PhotoLoopは写真を撮る際のイベント、その後写真仲間同士して閲覧するイベント、総合的なイベント記録ツールとして利用できるであろう。

3.2.2 友人や親戚に見せる

PhotoLoopを利用して、ナレーション付きのスライドショーを生成すれば、そのイベントには参加しなかった友人に対して写真を見せる際に、イベントに参加した人の感想などが含むスライドショーを見せることができるため、写真画像からは得られないような臨場感や、写真の周辺のストーリーなどを伝えることができる。たとえば、家族旅行での写真を親戚に見せるなどが考えられる。この際、さらに友人や親戚らにPhotoLoopを使ってもらうことで、それを友人への返信としたり、友人や親戚ら自身の写真閲覧の思い出となるとも考えられる(図4)。

3.2.3 体験記録の検索

今日、写真の検索は、タグ付けなどをユーザが積極的に行うことによってテキストで検索可能にするアプローチがある。しかし、タグ付けは時間の浪費やタグの選定の難しさの問題がある。PhotoLoopでは、現在音声そのまま記録しているが、将来的には音声認識を採用することで、テキストによるアノテーションも可能となり、検索に活用しやすくなる。また、閲覧時の記録において、総合的盛り上がり度が高いスライドは、そのイベントの重要度が高いと考えることができる。矢印をつけた写真がそのイベント内で多ければ、そのイベントは注目すべきものであることが示唆される。従来のタグ付けと異なるのは、タグづける意識をしなくともユーザはスライドショーを楽しむことだけに専念すればよい点である。

4 関連研究

筆者らは、写真撮影時に着目したアノテーションシステムWillCamを試作した[1]。WillCamは、写真撮影時にコンテキスト情報などを同時に記録することで、閲覧者が写真の撮影意図を理解するのを支援するシステムである。

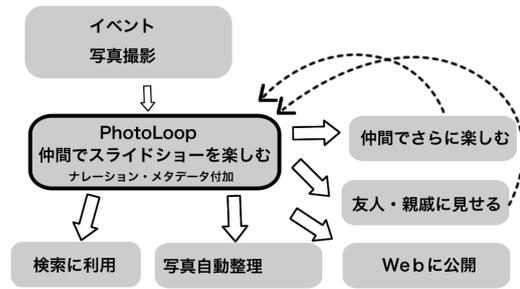


図 4. PhotoLoop の応用

ビデオメッセージにおいて、互いにコミュニケーションする際にビデオの再生を行いながら、以前のビデオにコメントを付加する研究がある[2]。すでに撮影したものに、さらに撮影によって情報を付加するという点で類似する点はあるが、PhotoLoopでは写真のスライドショーによる自然なナレーション付け、メタデータの付加を目的とする点で異なる。

ビデオメールのやり取りにおいて、ビデオ閲覧時に相手の表情を記録することで、それを返信用コンテンツとするVMailがある[3]。着想は近いが、筆者らは写真という蓄積されるコンテンツを対象としている点で異なる。

5 おわりに

本研究では、デジタル写真のスライドショー閲覧時の自然なユーザの語らいや指差し行為に着目し、ナレーションやメタデータとして利用する手法PhotoLoopを提案した。PhotoLoopでは写真を通常のスライドショーのように閲覧し楽しむだけで、その先の楽しみ方や幅広い活用が期待できる。本稿では、スライドショーシステムの試作までであったが、今後はメタデータを活用した検索や分類システムおよび、さまざまな写真の楽しみ方を提案していきたい。

参考文献

- [1] K. Watanabe, K. Tsukada, and M. Yasumura. WillCam: a digital camera visualizing users' interest. In *CHI '07: CHI '07 extended abstracts on Human factors in computing systems*, pp. 2747–2752, New York, NY, USA, 2007. ACM Press.
- [2] 敏弘 高田, 康德 原田. 引用可能なビデオメッセージ・システムの提案と実現 (特集・インタラクティブソフトウェア). *コンピュータソフトウェア*, 16(6):562–570, 1999.
- [3] 木村 博巳, 田中 充, 勅使河原 可海. ビデオメール VMail におけるビデオ主体のインタラクションを簡易化させる ユーザインタフェースの設計と開発. *インタラクション 2003 論文集*, pp. i-ii, 2003.