

Google Glass で動作するポスターアノテーションシステム

田中 克磨^{1,a)} カイ クンツェ^{2,b)} 岩田 基^{2,c)} 岩村 雅一^{2,d)} 黄瀬 浩一^{2,e)}

1. はじめに

学会におけるポスター発表は、研究者が自身の研究に対する意見やアイデアを得るための重要な場である。しかし、ポスター発表の性質上、そこでの議論は少人数の間でのやり取りであることが多く、交わされた意見やアイデアが他の参加者と共有されないという問題がある。この問題を解決するため、本研究では、眼鏡型ウェアラブルデバイスである Google Glass で動作するポスターアノテーションシステムを提案する。これは、ユーザがポスターに対してコメントや音声などのアノテーションをデジタル情報として付加し、それを他の参加者と共有できるというものである。本システムでは、まずユーザに Google Glass でポスターを撮影してもらい、Locally Likely Arrangement Hashing (LLAH)[1] と呼ばれる文書画像検索を用いて撮影画像から対応するポスターを特定することで、紙のポスターとデジタル情報の関連付けを実現している。本稿では、本システムが想定しているユースケースシナリオや、システムの実装・性能について述べる。

2. 関連研究

“PaperUP”[2] は、紙をディスプレイ、携帯端末をマウスのように使用するインターフェイスである。PaperUI では、文書を特定するために QR コードや SIFT[3] などの局所特徴量を用いている。QR コードは高速に認識できるという利点があるが、埋め込める情報量には限界があり、また文書上にプリントする必要があるため視覚的に目立ちすぎるといった問題がある。SIFT 特徴量は文書のレイアウト等によらず使用することができ、文書にマーカーをプリントする必要はないが、特徴量の計算に時間がかかるためハードウェア資源の乏しいウェアラブルデバイスでは十分な速度を出せないという問題がある。

3. シナリオ

ここでは、本システムの各機能がどのような場合に有用であるかについて考察する。

3.1 コメントの追加

図 1 に示すように、ユーザはポスターに対してコメントを追加することができる。これにより、例えばポスターを後で読むときのためにコメントを残したり、ポスターの発表者が不在の場合に質問を残しておいたりするという使い方が考えられる。Google Glass は文字入力 of インターフェイスを持たないため、コメントの入力には Google Glass の持つ音声認識の機能を用いる。

3.2 音声の追加

ポスターに長いコメントを追加したい場合、音声入力だけで完璧に入力するのは困難である。そのため、本システムでは図 2 のような、ユーザが音声を録音して追加できる機能を実装している。これにより、例えば発表者との議論を録音しポスターに追加しておくことで、後から誰でもその議論を聞き直すことが可能になる。

3.3 関連情報へのアクセス

本システムでは、発表者も自らのポスターにコンテンツを追加することができる。例えば、デモ動画や論文のファイルをポスターに登録しておくことで、図 3 のようにユーザがポスターを読みながらデモ動画を同時に再生したり、気に入った研究の論文を自分のメールアドレス宛に送信するなどといったことが可能になる。

3.4 アノテーションを後から参照する

ポスター発表でのやり取りをすべて記憶することは困難だが、本システムでコメントや音声として記録しておけば、後から Web インターフェイスを通して閲覧することができる(図 4)。これにより、例えば発表者が参加者から得られたフィードバックを残しておき、後から好きなときに参照することができるようになる。

¹ 大阪府立大学工学部 〒599-8531 大阪府堺市中区学園町 1-1

² 大阪府立大学大学院工学研究科 〒599-8531 大阪府堺市中区学園町 1-1

a) katsuma@m.cs.osakafu-u.ac.jp

b) kai.kunze@gmail.com

c) iwata@cs.osakafu-u.ac.jp

d) masa@cs.osakafu-u.ac.jp

e) kise@cs.osakafu-u.ac.jp

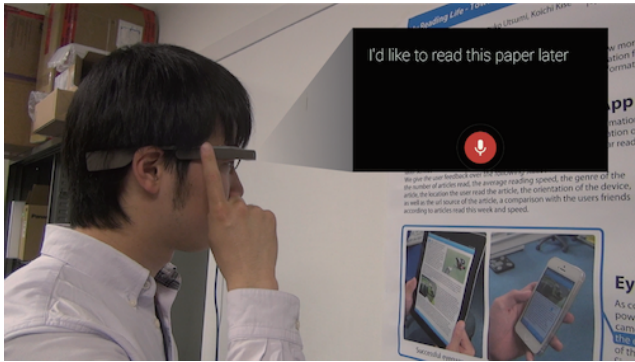


図 1 音声入力によるコメントの追加

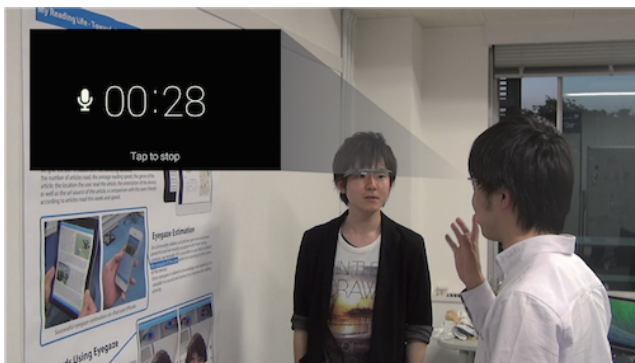


図 2 発表者との議論を録音する



図 3 ポスターを見ながらデモ動画を再生する

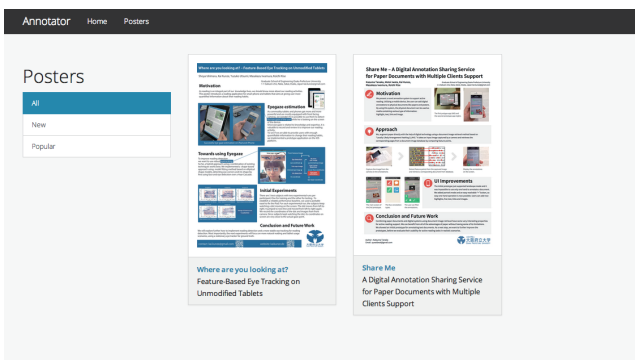


図 4 Web インターフェース

4. ソフトウェア

4.1 実装

本システムでは文書画像検索として LLAH を利用するが、ハードウェア資源の少ない Google Glass 上で検索処理をすべて行うことは困難である。そこで、Google Glass では撮影画像からの特徴点抽出のみを行い、検索処理はあらかじめ用意されたサーバ上で行うこととする。このサーバは、ポスターの文書画像データベースと前述の Web インターフェースを保持している。

4.2 検索性能

本システムで文書画像検索の手法として用いる LLAH については、1 億ページの文書画像を登録したデータベースを使った実験において、必要メモリ量 236[GB]、検索精度 98.7[%]、処理時間 26.8[ms] での検索が可能であることが確認されている [5]。そのため、ポスター発表の規模（ポスター 30～50 枚）においても十分に実用的な速度と精度で検索可能である。

5. まとめ

本研究では Google Glass 上で動作するポスターアノテーションシステムを提案した。本システムを利用することで、従来のポスター発表の形式を壊すことなく、ポスター発表の参加者同士が議論を簡単に共有できるようになる。また本稿では、本システムがどのようなケースで有用であるかについて考察した。今後の課題としては、実際のポスター発表でソフトウェアを配布してユーザからフィードバックを集め、アプリケーションのユーザインタフェースとユーザビリティを向上させることが挙げられる。

参考文献

- [1] T. Nakai, K. Kise, and M. Iwamura, "Use of Affine Invariants in Locally Likely Arrangement Hashing for Camera-Based Document Image Retrieval," Proc. DAS 2006, 3872, pp. 541–552, Feb. 2006.
- [2] Q. Liu and Chunyuan Liao, "PaperUI," Proc. 4th International Workshop on Camera-Based Document Analysis and Recognition, pp. 3–10, Sep. 2011.
- [3] D.G. Lowe, "Distinctive image features from scale-invariant keypoints," Int. J. Comput. Vision, vol. 60, pp. 91–110, Nov. 2004.
- [4] K. Takeda, K. Kise, and M. Iwamura, "Real-time document image retrieval on a smartphone," Proc. IAPR, pp. 225–229, 2012.
- [5] 竹田 一貴, 黄瀬 浩一, 岩村 雅一, "1 億ページのデータベースを対象とした大規模文書画像検索," パターン認識・メディア理解研究会 (PRMU2012), pp. 131–136, 2012 年.