

称号及び氏名 博士（工学） 中居 友弘

学位授与の日付 平成 20 年 3 月 31 日

論文名 「Study on Hash-Based Document Image Retrieval  
and Its Application to Information Recycle」

(ハッシュに基づく文書画像検索と

その情報リサイクルへの応用に関する研究)

論文審査委員 主査 黄瀬 浩一

副査 福永 邦雄

副査 大松 繁

## 論文要旨

近年、デジタルカメラの普及により、高性能なデジタルカメラを安価に入手できるようになっている。携帯電話付属のデジタルカメラについても、通常のものに比べて遜色のない解像度を備えるようになり、ユーザが高解像度のデジタルカメラを常に持ち歩くという環境が形成されている。その一方で、現状ではデジタルカメラの用途はスナップ写真の撮影が主であり、検索や認識を伴う高度なサービスには結び付けられていない。そのため、デジタルカメラを利用した便利なサービスを提案することができれば、我々の生活に大きなインパクトを与えることができると考えられる。

デジタルカメラを利用したサービスとして、撮影画像を検索質問とした画像検索が考えられる。これは、ユーザが注目した物体をデジタルカメラで撮影したとき、撮影画像に基づいて画像データベースから同一の物体の画像を検索するというものである。Web サービスとして一般的な、キーワードを検索質問とする画像検索とは異なり、画像を検索質問とする画像検索では、検索対象の画像は事前にユーザが持っている。従って、単に画像を検索して提示するだけではサービスとしての有用性は乏しい。ところが、データベース側で画像に関連情報を付加しておけば、画像を検索質問とする画像検索は有用となる。例えば、データベースの画像に物体の名称に関連情報として定義しておけば、検索時に関連情報を提示することで検索を通じた物体認識が可能となる。また、関連する Web ページの URL を登録しておけば、物体自体をリンクのアンカーとして利用することができる。

デジタルカメラを用いた画像検索の検索対象には、さまざまなものが考えられる。その中でも、印刷文書はその一般性から特に重要である。紙文書は高い可読性と携帯性をもつため、OA 化の進んだ近年においても依然として広く利用されている。むしろ、コンピュータの普及により電子文書の作成が容易となったため、電子文書の閲覧媒体として紙文書の利用量は増加する傾向にある。従って、印刷文書を画像検索の検索対象とし、撮影によって関連情報を取得できるようにすれば、ユーザの身の回りに多く存在する印刷文書がさらなる情報の引き出し口となり、高い

利便性を実現すると考えられる。

デジタルカメラを用いた文書画像検索を実現する上で問題となるのは、カメラで撮影された画像に生じる射影歪みである。デジタルカメラで撮影する場合、撮影方向に応じて得られる画像の見かけが大きく変化するため、データベースに登録されている電子文書との対応付けは困難である。文書画像検索に関する従来研究の多くはスキャナの利用を前提しており、射影歪みが生じる場合には適用できない。カメラで撮影された文書画像の撮影角度を推定し、正対画像への補正を行う手法も提案されているが、精度が十分ではなく、また多くの処理時間を要する。そのため、射影歪みを受けた検索質問画像からでも高速かつ高精度に画像を検索できる手法が必要である。

文書画像検索の入力機器として通常のデジタルカメラを用いる場合、検索を行う際にはシャッターを押すなどのユーザの能動的な操作が必要である。従って、画像検索による情報取得はユーザの自発的な意思によるプル型のものになる。これに対して、Webカメラなどを入力機器として、検索処理をリアルタイム化し、検索対象にカメラを向けるだけで検索するプッシュ型の情報取得を考えると、応用可能なサービスの幅を広げることが可能になる。例えば、ウェアラブルカメラなどと組み合わせ、リアルタイム処理によって撮影された画像からの検索を常に行っておけば、関連情報の定義された対象が視界に入った際に自動的に情報の提示を行うといったサービスを実現することができる。また、撮影画像に関連情報を重畳表示する拡張現実の処理を施すことで、より表現力豊かな情報提供が可能になる。

文書画像検索の他の応用として、書き込み抽出に基づく情報リサイクルの実現が挙げられる。情報リサイクルとは、既存の情報を分解して情報の素材とし、それをユーザの好みに応じた情報へと再生するプロセスである。このような処理を実現するためには、ユーザの好みや興味（ユーザプロファイル）の収集が必要不可欠である。さまざまなユーザプロファイル収集のための研究が行われているが、従来研究の多くは電子メディアにおける情報を対象としており、紙メディアを対象とした研究はあまり行われていない。前述したように、現在でも紙メディアは広く用いられており、そこに表現された情報を活用することは重要である。紙メディアからユーザプロファイルを収集する手法として、ユーザの行った書き込みを利用するものが考えられる。書き込みにはユーザの興味や知識が表現されているため、それらを抽出・解析することでユーザプロファイルを得ることができる。書き込み抽出の手法として、書き込まれた文書の画像と元の電子文書の差分を取るものが考えられる。このような処理を実現する場合、書き込まれた文書からの元の電子文書の検索と、それらの位置合わせが必要となる。これらの処理は文書画像検索によって実現することができる。

以上のように、デジタルカメラを利用した文書画像検索は高い利便性と有用性をもち、情報リサイクルへの応用が可能である。本論文では、デジタルカメラを利用した画像検索法である **Locally Likely Arrangement Hashing(LLAH)** と、LLAH に基づく文書画像検索法を提案する。LLAH は幾何学的不変量を用いて画像を表現するため、射影歪みを受けた画像でも高い精度で検索が可能である。また、ハッシュ技術の利用により、高速な処理を実現している。本論文ではさらに、文書画像検索の応用として情報リサイクルについても述べる。

本論文の構成は、以下の通りである。

第1章では、本研究の背景と目的について述べ、研究成果について概観する。

第2章では、LLAH に基づく文書画像検索について述べる。LLAH は、画像から抽出された特徴点に基づいて画像を検索する手法である。デジタルカメラを用いる際に撮影画像に生じる射影歪みに対処するため、特徴点の配置から計算される幾何学的不変量を用いる。また、計算量の低減のため、各特徴点について近傍点を組み合わせて特徴量を計算する。さらに、効率的な検索を実現するため、ハッシュを通じて特徴量を比較する。LLAH を文書画像検索に適用する場合、文書画像からの特徴点抽出法が必要となる。第2章の手法では、英文文書を対象として、単語の重心を特徴点として抽出する。実験により、大規模な文書画像データベースにおいても高速かつ高精度に検索できることが確認された。

第3章では、LLAHの性能の改善方法について述べる。第2章の手法では幾何学的不変量として射影変換の不変量である複比を用いているが、画像の歪みが大きいと複比の値が安定しないという問題がある。また、LLAHは高速かつ高精度な画像検索である一方、実行には多量のメモリを必要とするという問題もある。さらに、リアルタイム検索などに用いるためには、さらなる高速化も必要である。第3章では、さまざまな幾何学的不変量の利用を検討し、アフィン不変量の導入によってより高い精度が得られることを明らかにする。また、重要性の低いデータのデータベースからの削除と、データベースの構造の単純化によって必要なメモリの量を削減し、アルゴリズムの改善によって処理時間を短縮する手法についても述べる。

第4章では、LLAHに基づく文書画像検索法のリアルタイム文書画像検索への応用について述べる。リアルタイム文書画像検索は、第2章の手法をWebカメラで取得した画像の各フレームに適用し、繰り返し行うことで実現される。これは、入力の変化に対する追従性のためである。以前のフレームの検索結果を利用するものでは、撮影される文書が変化しても反応が数フレーム遅れ、追従性が低下する。そのため、第4章の手法ではフレームごとに独立した検索処理を繰り返し行う。この場合、フレームごとの処理が高速である必要があるため、クライアントサーバシステムによるパイプライン処理を行う。また、リアルタイム文書画像検索の効果的な応用のため、文書単位の検索だけでなく、撮影範囲の特定も行う。これにより、ユーザによる検索結果の確認が容易となる。また、拡張現実処理のように、文書の部分に関連付けられた情報を適切に提示できるようになる。

第5章では、文書画像検索の応用として、情報リサイクルのための紙文書からの書き込み抽出について述べる。紙文書からの書き込み抽出法としては、さまざまなものが提案されているが、これらの手法には書き込みの色や形状に制約があるため適用範囲が十分とはいえない。第5章で述べる手法は、書き込みのある文書をスキャンした画像（書き込み画像）と元となった電子文書から得た画像（元画像）を比較し差分を得ることで、書き込みに対して制約を課すことなく多様な書き込みを抽出するものである。このような処理を実現するためには、書き込み画像と元画像の正確な位置合わせが必要となる。第2章で述べた文書画像検索法では、検索の過程で特徴点の対応関係が得られる。これを利用することで書き込み画像と元画像の位置合わせが実現される。

最後に、第6章で本研究について総括し、今後の課題について述べる。

## 審査結果の要旨

本論文は、デジタルカメラを用いた文書画像検索手法と、その改善や応用について論じたものである。特徴点の局所的配置に基づく画像検索法であるLLAH (Locally Likely Arrangement Hashing)を提案し、その文書画像検索への適用や性能の改善、既存の情報の分解と再構成のプロセスである情報リサイクルへの応用について検討し、以下の成果を得ている。

- (1) LLAHと呼ばれるアルゴリズムを新規に提案するとともに、それに基づく文書画像検索を明らかにした。LLAHは、画像から抽出された特徴点に基づいて画像を検索する手法である。デジタルカメラを用いる際に撮影画像に生じる射影歪みに対処するため、特徴点の配置から計算される幾何学的不変量を用いる。また、計算量の低減のため、各特徴点について近傍点を組み合わせて特徴量を計算する。さらに、効率的な検索を実現するため、ハッシュを通じて特徴量を比較する。実験により、大規模な文書画像データベースにおいても高速かつ高精度に検索できることが確認された。
- (2) LLAHの性能の改善方法を明らかにした。種々の幾何学的不変量の利用を検討し、アフィン不変量の導入により高い精度が得られることを明らかにした。また、重要性の低いデータの削除と、データ構造の単純化により必要なメモリ量を削減し、アルゴリズムの改善に

より処理時間を短縮する手法についても明らかにした。

- (3) **LLAH** の応用として、リアルタイム文書画像検索手法について明らかにした。この手法では、文書単位の検索だけでなく、撮影範囲の特定も行う。これにより、検索結果の確認が容易となる。また、拡張現実処理のように、文書の部分に関連付けられた情報の適切な提示も可能にした。
- (4) 文書画像検索の応用として、情報リサイクルのための紙文書からの書き込み抽出について明らかにした。書き込みのある文書をスキャンした画像（書き込み画像）と元の電子文書から得た画像（元画像）を **LLAH** により位置合わせして差分を得ることで、多様な書き込みを抽出する手法を実現した。

以上の諸成果は、知能情報工学分野におけるメディア検索の方法論に重要な知見を与えるものであり、本分野の学術的な発展に貢献すること大である。また、申請者が自立して研究活動を行うのに必要な能力と学識を有することを証したものである。