

画像を用いた物体認識と人追跡による 図書管理システム

網島 健太[†] 藤田 茂^{††} 菅原 研次^{††}

[†] 千葉工業大学大学院 情報科学研究科 〒 275-0016 習志野市津田沼 2-17-1

^{††} 千葉工業大学 情報科学部 〒 275-0016 習志野市津田沼 2-17-1

概要: 一般に物体の管理を計算機の支援を得て行う場合には、RFID タグやバーコード、QR コードなどの付加情報を対象物に添付することが多くある。しかしながら、既に多様な物体が大量に存在する管理対象に対して新たに添付することは煩雑な作業である。我々は、日常生活の中で利用が容易である USB カメラのみを用いて、書架に存在する書籍とその書籍を利用する人の情報を管理するシステムを構築した。構築したシステムは、本の表紙の特徴と人の動作追跡を組合せ、図書の管理を実現している。

キーワード: 物体認識, 物体識別, SIFT, 共生コンピューティング

A System for Book Management by Object Recognition and Human Tracking with Camera

Kenta TSUNASHIMA[†] Sigeru FUJITA^{††} Kenji SUGAWARA^{††}

[†] Graduate School of Information and Computer Science, Chiba Institute of Technology 2-17-1, Tsudanuma, Narashino, 275-0016 Japan

^{††} Computer Science, Chiba Institute of Technology 2-17-1, Tsudanuma, Narashino, 275-0016 Japan

abstract In general, when the support of the computer is won and the object is managed, additional information on RFID-tag, the bar code, and the quick response code (QR code), etc. is often appended to the object. However, newly appending it to the management object where a large amount of various objects already exist is troublesome work. We designed a system that managed information on the person who used the book that existed on the bookshelf and the book with only an easy USB camera to use in daily life. The system that constructs it combines the feature of the cover of a book and person's motion tracking, and has achieved the management of books.

keyword: Object recognition, SIFT, Symbiotic Computing

1 はじめに

近年、ユビキタス社会の発展に伴い、非接触のカードによる入出管理、GPS の位置情報を利用したサービスなどのユビキタスデバイスを利用したサービスが注目されている。カメラもユビキタスデバイスの1つであり、ネットワークカメラはインテリジェンスカメラとして活用が期待されている。現在、カメラの利用は、街角・店の監視や施設における顔認証が実現されている。犯罪抑制の面から監視カメラの数は増加しており、様々な施設に設置されている。しかし、現在カメラからの映

像から必要な情報を抽出し、管理を行うものは少ない。例えば、書店においても、人を監視し、犯罪の抑制を行っているため、カメラの映像を人が監視し、カメラの画像を人が処理している。そのため、自動化のためにはカメラの映像から情報を抽出をすることが求められる。

この煩雑さは現実空間とデジタル空間との間のギャップとして示されている [1]。これは、デジタル空間側に人や社会を理解する仕組みが不足しているためであり、そのため、現実空間で求められている支援が十分に実現しにくい。このギャップを解消するためには、現実空間とデジタル空

間における、共認知を定義・利用することで、環境情報をデジタル空間が自律的に獲得し、得た情報をわかりやすく提供することを目指している。

現実空間とデジタル空間のギャップを解消する方法の一つとして、物体に RFID を添付する方法がある。新規の図書館には RFID を用いた図書管理システムが用いられ始めている。既存の図書館など大量に存在する管理対象に対して新たに添付することは難しく、それぞれの本に RFID タグを付けなければならないといった煩雑な作業と費用がかかる。

そこで我々は、共生コンピューティングに基づいた図書管理システムを提案する。このシステムは、実社会から情報を得るためのデバイスとしてカメラを使い、本に RFID タグやバーコードなどの付加情報を添付することなく、本の管理を行う。

職場や学校などの自分の生活空間内において、自分の探している本が本棚にないとき、それを探す必要が出てくる。誰かがその本を持ち出した場合には、誰が持ち出している人を知り、その人物から借りる必要がある。その持ち出された本が生活空間内での移動について考えるとき、本は自ら動くことはないので誰かが所持することで本棚からの移動する。そのため、本の移動を追跡するのではなく、本を所持している人物を追跡することで、本の効果的な管理を行う。

提案するシステムは図書の管理空間を小規模に限定し、本の管理を行うためにカメラを用いた。カメラから画像を取得し、事前にシステムに登録しておいた本の画像情報郡とのマッチングを行い、本棚の管理を行う。この情報だけでは書架を利用した人物を知ることはできない。そこで、もう一つカメラを使い、人物の追跡を行う。本システムは、二つのカメラを用い、利用するユーザが本棚で本を見付けられない場合でも追跡情報を得ることで、誰が本を所持しているか知ることができる。

2 現在ある図書管理方法

現在、図書の管理方法としてバーコードを用いて管理する方法と RFID タグを付与する方法がある。

バーコードを使い管理する方法は、本に固有の ID を持つバーコードラベルを張り付けて本を管理する方法である。バーコードリーダーを使い、本のバーコードを読み取り、貸出し・返却の管理を行う。現在多くの図書館で使用されている方法で

ある。

RFID(Radio Frequency Identification) を使い管理する方法は、それぞれの固有の ID を持つラベル状の RFID タグを図書に付与することで本の管理を行っている。RFID はユビキタス社会実現に向け期待され、RFID の利用により、図書の管理の効率化が進められ、現在多くの場所で図書を管理する方法として利用されている [2]。

RFID の特徴として、近距離から非接触でタグの情報の読み取りが可能である。本の貸出し・返却は専用のリーダー上に複数冊の本をかざすことで手続きを行う。これにより、バーコードを用いた方法では一冊づつ処理していた作業を同時処理が可能となり、貸出しや返却を効果的に行える。また、近距離で情報の読み取りが行えるため、本の点検の管理、誤配架の発見が容易になる。

既存の二つの方法は、RFID やバーコードを新たに添付するには煩雑な作業が伴う。また、RFID やバーコードを用いた方法では本のあるべき位置の情報を得ることができず、整理を行うことが難しい、そのため、カメラを用いることで、本の管理の他に書架における位置情報を取得できる。

3 提案システム

本章ではカメラを用いた図書管理システムについて述べ、次にシステムを成す 2 つの手法について説明をする。

3.1 提案システムの概要

カメラを用いた図書管理システムの全体を図 1 に示す。システムは二台のカメラを使い、一つは本の管理を行い、もう一つは人物の追跡を行う。それぞれのカメラが本の管理を行うシステム、人を追跡するシステムに繋がり、人物追跡情報と書架の管理情報を得る。二つのシステムから得られた情報を用いて図書を管理を行う。

3.2 本の管理の処理

本の管理のために、次の情報を扱う。本を個別に管理するために本の名前とは別に本棚に固有の ID と並べる順を付加した ID を与える。情報を管理して書架に置ける有無と、位置情報を得る。

本情報=<SIFT 特徴, 本の名前, 本の ID, 本棚での有無, 本棚での位置情報>

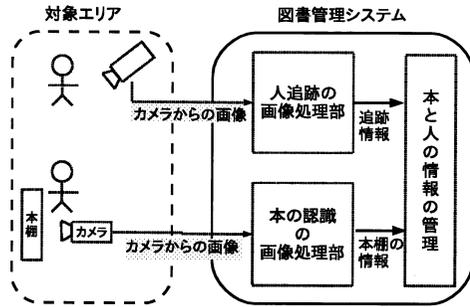


図 1 図書館管理システムの構成

本の管理は事前に本の特徴情報を抽出する。本の特徴の情報として SIFT(Scale-Invariant Feature Transform)[3] を使う。SIFT は特徴点の検出と特徴量を記述を行うアルゴリズムであり、SIFT 特徴量は輝度勾配に基づく局所不変特徴量である。検出した特徴に対して、画像の回転やスケール変化・照明変化に頑健な特徴量を記述するため、画像のマッチングや物体の認識・検出について用いられる。

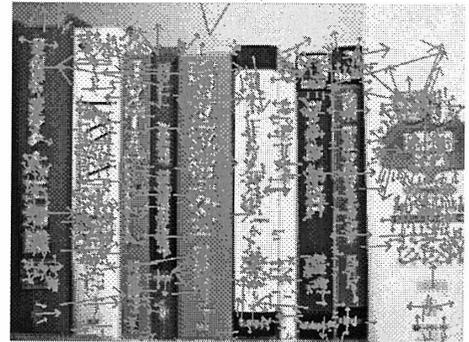
書架の管理を行うカメラの設置位置を本棚の前とした。このとき、カメラから取得できる入力画像は本の背表紙となる。そのため、本から特徴を抽出する際に本の背表紙の画像情報を使い、特徴の抽出した。図 2 に本 (a) と本棚 (b) から特徴を抽出した例を示す。

図 2(a)(b) の特徴点の対応付けした例を図 3 に示す。入力画像から本の背表紙の特徴を抽出し、事前に抽出した特徴と対応付けを行い、対応点を得た。しかし、この対応付けのみでは局所的に類似した部分が背景にある場合、誤対応が生じる。このことは対応点が存在が直接本の存在を示すことにならないことを意味し、類似した特徴があるからそこに本が存在するとは限らない。そのため事前に抽出した特徴とカメラから得た特徴の対応付けが正しいかを確認する必要がある。

一冊の本に対し、対応点を背景と他の本と分離をすることを、外れ値 (outlier) を含んだ対応点に対するモデル当てはめ問題として扱う。ここでモデルとは、1つの認識物の対応点が従う大局的な拘束を意味し、本の背表紙は平面であるため、射影変換を使用することができる。外れ値を含む対応点に対



本 (a)



本棚 (b)

図 2 SIFT 特徴の例

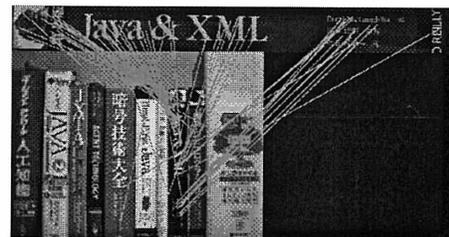


図 3 SIFT 特徴を用いた対応付けの例



図 4 本の検出例

する射影変換の当てはめに、RANSAC(RANdom SAmple Consensus)[4]を使用する。RANSACはロバストな統計量の推定手法であり、ランダムに抽出したサンプルから得られた推定値が正しいと仮定した場合の正対応の数を数え、処理を繰り返し行い、正対応の数が最大になるような推定を正しい推定とする手法である。図2(a)と(b)から対応付けを行い、RANSACを用いた結果が図4となり、本が認識されていることがわかる。射影変換行列を使い、本棚での本の位置情報を得る。

次にSFIT特徴量の求めかたとRANSACのアルゴリズムを記す。

3.2.1 SIFT特徴点の抽出と対応付け

SIFT特徴量は特徴点の位置 \mathbf{p} 、スケール σ 、記述子 \mathbf{d} によって構成される。

エッジ抽出オペレータとして、ガウシアン G の2階微分であるLoG(Laplacian of Gaussian)を近似したDoG(Difference of Gaussian)を用いる。DoGは $G(\cdot)$ はガウス関数、 x, y を注目している画素の座標値とすると次式で表される。

$$DoG(x, y, k^n\sigma) = G(x, y, k^{n+1}\sigma) - G(x, y, k^n\sigma) \quad (1)$$

ここで、 k は定数パラメータである。また、スケールスペースとしてそれぞれ $n = 1, 2, 3, \dots, N - 1$ で計算されたガウシアン画像の階層構造を考える。このスケールスペースにおいて、隣り合うガウシアン画像の差分取り、DoG画像の階層構造を得る。DoG画像の階層構造を3次元的に見たとき、ある画素と隣接する26近傍の画素とそれぞれを比較をして、その画素が極値となるようなときに特徴点の候補とする。この候補の特徴点の内、コントラストや主曲率により、特徴点を絞りこむ。特徴点にはスケール σ の情報を持たせておく。

得られた特徴点について、その近傍領域で輝度勾配の方向ヒストグラムを計算をし、最も頻度が高い方向を探す。そして、特徴点を中心とした画像座標系をその方向に回転した局所極座標系を作る。この極座標系を中心とし、スケール σ に比例した大きさの局所領域を 4×4 のブロックに分割して8方向のヒストグラムを作成する。そのため、128次元のベクトルを持ち、それを正規化したものをSIFT特徴量の記述子 \mathbf{d} とする。

対応付けは次のように求められる。背表紙Aの i 番目の特徴点 \mathbf{p}_i^A に対するSIFT記述子 \mathbf{d}_i^A で

表現し、同様にカメラからの画像Iを $\mathbf{p}_j^I, \mathbf{d}_j^I$ と表現し、特徴量間の距離を記述子間ユークリッド距離 $d_{ij} = \|\mathbf{d}_i - \mathbf{d}_j\|$ で表現する。距離尺度 d_{ij} を用いるとき、テンプレート中の各特徴点 \mathbf{p}_i^A について、入力画像中の最近傍点のインデックスを $NN = \operatorname{argmin}_j d_{ij}$ で表し、2番目に近いものを2-NNとすると次式を満たすようなシーン画像中の点 \mathbf{p}_{NN}^I を点 \mathbf{p}_i^A と対応づける。

$$\frac{d_{iNN}}{d_{i2-NN}} < \text{match_threshold} \quad (2)$$

3.2.2 RANSACによる位置の推定

RANSACアルゴリズムはランダムサンプリングに基づく統計量の推定法である。RANSACはサンプルをランダムに抽出し、得られた推定値が正しいと仮定した場合の正対応(inlier)の数を数えるとう処理を繰り返し行い、inlierの数が最大になるような推定を正しい推定とする手法である。以下にそのアルゴリズムを記す。

1. SIFTを使い求めた対応した点をランダムに4組を選択する。そして射影変換行列 \mathbf{H} を求める。
2. 求めた \mathbf{H} と事前に抽出した特徴点 \mathbf{p}_A に対応するカメラからの入力画像Iの特徴点 \mathbf{p}_I をラベル画像上に変換したときの変換誤差を求める。
3. 変換誤差が閾値以下となるラベル画像上の特徴点をinlierとして、ラベル画像上の全ての特徴点で2.処理を行い、inlierの数を求める。
4. 1.~3.の処理を繰り返し、最大数のinlierを与えるときinlierと判定去れた組のみを使い、射影変換行列を求める。

求めた射影変換行列を使い、ラベル画像の四隅の点のカメラからの画像の位置を求める。

3.3 人の追跡処理

本の貸出しが行われた際に、誰が本を所持しているかを知るために人の追跡を行う。本を所持している人物の情報を得るために本システムでは、本棚付近の部分に対し背景の差分を行い、その差分画像から得られた部分を人領域を定め、Flocks Features track[5]手法を用いて処理を行い、人の追跡を行う。

追跡している人物の画像と情報を情報の管理部に送る。

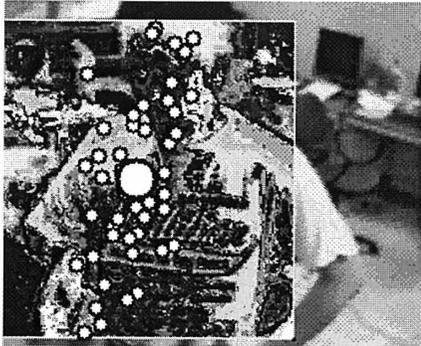


図 5 Flock Features track の例

3.3.1 Flock Features track

“Flocks of Features” は手のような非剛体で関節物体を高速に追跡することができる手法である。二次元位置追跡を容易にするために追跡領域にマスクを用いて、KLT 特徴と前景の色分布の学習結果を結合する。今回、背景差分の画像を利用し、マスクの作成した。追跡での利点は、速度、背景のノイズに対しロバストで自由な回転と膨大で急激な変形に耐え、物体を追跡する能力がある。変化した物体のあとを追跡しつつということを可能にするために flock features (群の特徴) の振舞を用いている。

Flock Features track を行っている例を図5に示す。囲まれている部分が追跡領域であり、画像の濃淡が物体から学習した前景と表した結果である。白い部分ほど、特徴が失われたときに特徴点を補填する可能性が高い。それぞれに小さな点が特徴点であり、大きな点は中心点である。

始めに追跡したい物体のある領域を定め、その領域の色ヒストグラムを学習する。そして、good-features-to-track アルゴリズム [6] を用いて、特徴点を検出する。そして、色とマスクを用いて順位付けを行い、N 個の特徴を選択する。今回、この特徴の数を 50 とした。追跡にはピラミッドに基づく KLT 特徴追跡法を用い、範囲を絞り追跡することで動いている物体で高速に結果を出し、効率的に計算ができる。領域内で特徴点を追跡できなかったとき、それは特徴が「失われた」状態になる。このとき、領域内で「失われた特徴」の代わりにシステムが学習した色分布を使い、前景と判断した部分に新しい特徴を補う。

3.4 本と人の情報管理

本と人の情報管理処理部では「本の管理の処理」と「人の追跡処理」の二つから得られた情報を使い、本の持ち出しが行われたとき、どの人物がその場にいたかという情報を使い、持っている人物の画像とシステムが所持していると判断した情報の管理を行う。

本の管理として ID と位置情報を使い、本が正しい順番で並んでいるかを判断する。正しく並んでいないなら、どの本の位置が間違っているかを出力を行い管理を行う。

4 評価実験

提案したカメラを用いた図書管理システムの動作を検証するために評価実験を行った。評価環境は、表 1 の環境を使用した。

表 1 実験環境

OS	openSUSE Linux 10.2
CPU	Core Duo T2300E(1.66GHz)
メモリ	512M
実装言語	gcc version 4.1.2
カメラ	Elecom UCAM-H1C30M
入力画像サイズ	640 × 480
本の数	10 冊

提案システムの評価のために、以下のようなシナリオを設定して評価実験を行った。

シナリオ 1

本棚に正しい順番に並べて、実験を行う。

本の貸出しをしない状態での管理状況の確認

シナリオ 2

本棚に正しくない順番に並べて、実験を行う。

本の貸出しをしない状態での管理状況の確認

シナリオ 3

本を持ち出しと返却の動作に対する

本の管理と追跡の状況。

シナリオ 4

類似した表紙を持つ本についての

管理状態の確認。

シナリオ 1 では一冊を除き、認識が行えることが確認できた。認識がうまくいかなかった一冊は、

10 回当たりの処理で約 5 回しか、認識できなかった。認識できない原因の一つとしてラベルのサイズが実験で用いた他の本に比べ、小さかった事が考えられる。これは解像度の高いカメラを用いれば認識が行えると考えられ、解像度の高いデジタルカメラの画像を用いたところ、認識が行えた。それ以外の本は、本の位置情報を使い、本の順番は正しいことが認識できた。

シナリオ 2 では正しくない順番に本を並べた。シナリオ 1 と同様に一冊を除き、認識を行えた。そして、その順番の間違った位置にある本を確認できた。

シナリオ 3 では本の持ち出しと返却の動作の確認と本を持っていった人物の追跡を行い、本棚の情報と人物の情報を使うことで本棚の情報を管理できることを確認した。人が物体 (人や物) との重なりがない場合には追跡が行えたが、物体が重なる場合、追跡に失敗した。追跡の失敗理由として、本棚付近で追跡する人物の登録をしておき、また、部屋にある他の物体の情報をシステムが持っていないため、追跡の失敗に繋がったと考えられる。

シナリオ 4 は類似した本を本棚に置き、実験を行った。類似した本には同じフォントで一部に同じ文字を持つ背表紙を用いたが、認識はされなかった。

5 おわりに

今回、カメラを用いた図書管理システムについて提案した。本システムでは、RFID タグやバーコードなど付加することなく、2 台のカメラを使い、本の管理に SIFT、人の追跡に Flocks of Features tracking を用いて、図書の管理システムを構築した。

本システムの有効性を確認するために、生活空間内の本の移動を仮定し、2 台のカメラを用いて実験を行った。その結果から、本棚を監視するカメラからは本の有無情報だけでなく、本の位置情報の抽出も行うことが可能となり、その位置情報から本棚が整理されているかの判断を行えることが確認できた。人の位置情報と本棚の管理状況を組み合わせる事で、本の貸出し状況を把握できることを確認した。

今回の実験で、一部で本の認識率が低いという課題が残った。今後、残った課題と管理できる本の数の増加や検出精度の向上を行っていく予定である。

参考文献

- [1] 白鳥則郎, 菅原研次, 菅沼拓夫, 藤田茂, 小出和秀. Symbiotic computing - ポスト・ユビキタス情報環境へ向けて -. 情報処理学会誌, Vol. 47, No. 8, pp. 811-816, 2006.
- [2] Bibliotheca. Rfid library systems, 2006. <http://www.bibliothecarfid.com/references>.
- [3] David G.Lowe. Distinctive image features from scale-invariant keypoints. *International Journal of Computer Vision*, Vol. 60, pp. 91-110, 2004.
- [4] Martin A. Fischler and Robert C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, Vol. 24, No. 6, pp. 381-395, 1981.
- [5] Mathias Kolsch and Matthew Turk. Fast 2d hand tracking with flocks of features and multi-cue integration. *CVPRW '04: Proceedings of the 2004 Conference on Computer Vision and Pattern Recognition Workshop (CVPRW'04) Volume 10*, p. 158, 2004.
- [6] Jianbo Shi and Carlo Tomasi. Good features to track. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR'94)*, 1994.