

写真からのオブジェクト抽出を利用した物品貸し借りプラットフォームの提案

加藤貴之[†] 中西泰人^{††} 服部隆志^{††} 萩野達也^{††}

[†] 慶應義塾大学大学院政策・メディア研究科
^{††} 慶應義塾大学環境情報学部

あらまし 本論文では物品貸し借り支援システムにおける新しい手法を提案する。我々の研究の目的は貸し借りイベントの情報を広く共有することによって、個人の所有物の管理と再流通を支援することである。そのために本論文では、携帯電話の写真とメールを用いて画像ベースで人と物の情報を記述する手法を提案する。この手法では不特定の物品を対象とすることができ、さらに対象を個体として直接扱うことができるという特長をもつ。本論文ではとくに、カメラ付き携帯電話を用いて簡単かつ安全に貸し借りイベントの情報を登録するための手法を提案する。写真の対象および構図に制約を設けることによって、1枚の画像から人と物の2つのオブジェクトを取得することができる。さらに、この制約の柔軟性について検討する。貸し借りの対象が大きい場合と小さい場合について情報登録および閲覧がうまくいくかをそれぞれ試みる。どちらの場合も写真撮影時のユーザの工夫によってシステムを正しく機能させることを示し、制約は貸し借りを阻害しないことを確認する。

キーワード 物品管理、貸し借り、写真、顔検出、構図

A Borrowing-lending Management Platform using Object Detection from Photographs

Takayuki KATO[†], Yasuto NAKANISHI^{††}, Takashi HATTORI^{††}, and Tatsuya HAGINO^{††}

[†] Keio University Graduate School of Media and Governance
^{††} Keio University Faculty of Environment and Information Studies

Abstract In this paper, we propose a new method for borrowing-lending support system. Our method is to use photographs and e-mails by cellular phones and describe the person and goods by images. The method realizes that the system manages unspecified objects and deals them as entities, which conventional methods cannot achieve. Especially, we propose a method to register the information of a borrowing-lending event easily and safely by cellular phones. By restricting the object and composition of the photographs, the system can obtain the images of the person and goods from one photograph. We also certify the flexibility of the restriction by showing that however extreme the size of the goods is, the form of the system functions correctly by users' devices of photography.

Key words Goods Management, Borrowing and Lending, Photograph, Object Detection, Pose

1. はじめに

1.1 背景

物の流通、保管のための個体管理の手法が近年注目されている。個体管理とは、個々の物について「何がいつの時点でどこにあるのか」を把握することである。RFIDを利用した情報システムは個体管理の手法の一つであり、図書館の本の管理[1]、冷蔵庫内の食品の管理[2]、日常生活の記憶の補助[3]など様々な応用例が提案されている。RFIDを用いたシステムは、情報

の入力作業が自動化できる点の特長である。しかし、このシステムではすべての対象物にあらかじめRFIDを付けることが前提となる。したがってこの手法は対象物の集合が定まっている場合には有効であるが、不特定の対象を管理することはできない。本研究では後者、すなわち対象物の集合が不特定である場合を採りあげる。特にその例として、個人間での物の貸し借りを題材とする。

1.2 BOZAAR: 貸し借り管理プラットフォーム

本研究の目的は、個人間での貸し借りの情報を共有して、所

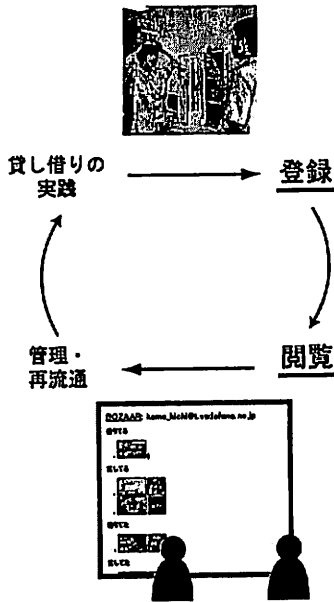


図1 BOZAAR システム全体図

有物の管理と再流通を促進させるための情報システムの構築である。このシステムを“BOZAAR”（ボザール）と名付ける。貸し借りの主体となる人、対象となる物の集合は不特定であるとする。BOZAAR は貸し借り実践のある任意の個体について、その物がいつの時点でどこにあるのかという情報を共有するためのシステムである。

本研究の特徴として、物の管理だけではなくて再流通にも主眼を置いている点が挙げられる。「A が B に g を貸した」という情報からは付随的に「A は g を持っている」という情報や「A と B は友人同士である」という情報を得ることができる。この付随情報が新しい貸し借りイベントを発生させるきっかけになると我々は考えている。このように所有物の情報を共有して再流通を促すメディアとしては BOZAAR は本棚.org [4] のような道具に近いといえる。

BOZAAR の機能は (1) 貸し借りイベントの登録および (1) 貸し借り情報の閲覧である。まず (1) 貸し借りイベントの登録は、貸し借りのイベントの際にその付随情報（誰が誰に何を貸した/返した）をシステムに入力する機能である。これは携帯電話のカメラと電子メールを用いる。次に (2) 貸し借り情報の閲覧機能では、登録されている任意の貸し借りについて、その当事者および対象物の情報を取得することができる。両者の情報は画像で記述される。図1は BOZAAR の機能と目的の関係をまとめたものである。

BOZAAR の手法について2点議論する。まず、物の情報を記述するのに画像を用いる理由は2つある。まず第一には、対象に形がありさえすれば必ず写真に収めることができるという点がある。したがって、不特定の対象を記述することができる。例えば“Strata Drawer” [5] では書類という分類にくい対象を記述するために写真を用いている。第二に、名前や属性

といった言葉による記述はその対象の概念的側面を表すのが得意であるのに対して、画像はその物の実体的側面を記述するのが得意だという点である。例えば「夏目漱石の『三四郎』を借りた」という記述の中の『三四郎』は、世界中に無数に流通している『三四郎』のどれなのか特定できない。一方、誰かの所有する『三四郎』を撮った写真は、その個体だけを指示する。BOZAAR の対象は個体の管理であるので、画像を使用することに意義があると考えられる。

つぎに、入力機器としてなぜ携帯電話を用いる理由を検討する。BOZAAR のユーザおよび対象物は不特定である。したがってすべての物に RFID を付け、すべての人が RFID リーダを持つというのは現実的ではない。物流として貸し借り行為を捉えたとき、物が移動する場所にあたるのは人である。ほとんどの人は常に携帯電話を持っている。携帯電話には通信機能があり、カメラというセンサも付属している。携帯電話による入力には人の労働力が必要であり RFID のシステムのような簡便さはないが、十分に入力の役割を担うことができる。

1.3 本論文のスコープ

BOZAAR のシステムはミクロ的、マクロ的のふたつの視点から捉えることができる。ミクロ的とは、貸す/借りるという単一のイベントに注目し、一つのイベントにおいてどのように入力機能と閲覧機能が連携するのかを考える視点である。一方のマクロ的な視点とは各イベント間の関係に注目して、どのように各対象物および各当事者の状態を経時的に管理するかを考えるものである。本論文ではマクロ的な視点からの検討は今後の課題として、まずミクロ的な視点からの設計および実装について述べる。

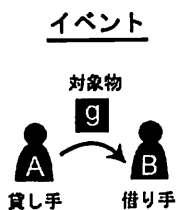
2. システム要件

本論文では貸し借りのイベントの際にまずユーザの携帯電話で入力を行い、そこから取得したイベント情報を画像によって提示する部分の設計および実装について述べる。より具体的には「人 A が人 B に物 g を貸す」というイベントから、「A が B に g を貸した」という情報、借り手 B の画像および対象物 g の画像を取得することが目標となる。図2はこの目標を模式的に表現したものである。とくに本論文では画像の取得の部分を中心に議論する。

また、システムの設計にあたっては次の2点の制約がある。第一に、入力作業はできる限り簡単でなくてはならない。これは入力機器が携帯電話である点、貸し借りの際にその場で行う点に由来する。第二に、入力において偽装が行われてはいけない。すなわち、A が勝手に C を騙って貸したり、g ではなくて h が貸されたことにする、ということがあってはならない。

3. 写真の構図の制約を利用したオブジェクト抽出手法

前章では BOZAAR を実現するために、貸し借りのイベントから人と物の画像を取得することを目標としてあげた。さらに、この目標には簡便さと安全さのふたつの制約が付随することを



↓ 取得

情報

- が に を貸した
- =
- =

図 2 貸し借りイベントからのデータ取得

述べた。本章ではふたつの制約を満たしつつ目標を実現するための方法を検討する。

3.1 複数の画像を使う手法の問題点

最も単純な手法は、ユーザが人・物のふたつの写真をそれぞれを撮影して送る、というものである。しかし、この方法は制約を満たさないので不適合である。まず第一に 2 回の撮影を行うのは面倒であるし、ふたつの画像ファイルを送るためにはそれぞれに名前を付けなければならない、作業が複雑となる。第二にこの方法だと偽装が可能である。別の時間・場所で撮影した人・物の写真をそれぞれ送ることで、実際には行われていない貸し借りが登録されるおそれがある。

3.2 単一の画像を使う手法

以上のような問題点をふまえて、本研究では 1 枚の画像を入力とするという手法を用いる。すなわち、借り手 B と対象物 g が一枚に収まった写真を貸し手 A が撮影して、それをシステムへ投稿するという手法である。

この手法は制約を満たしている。第一に撮影が 1 回になることによって入力とは簡単になる。また、システムへ送るファイルが 1 枚であれば名前付けの手間もかからない。第二に画像が証拠写真の役割を果たすので、入力の偽装は起こりにくくなる。B と g が一枚の写真に収まっているということは、現実には B と g は同じ場所にいるということである。すなわち、B が g を借りるといったイベントが実際に行われたことを証明する。したがって入力において偽装を行うことが難しくなる。

3.3 オブジェクト抽出のための制約付け

次に入力と出力の間を橋渡しする処理について検討する。入力が 1 枚の写真になるということで入力と出力の間にギャップが発生する。そこで、1 枚の写真を入力として、そこから 2 つの画像を出力するための仕組みが必要となる。具体的には、借り手 B と対象物 g が混在して写っているイベント写真から、B の画像および g の画像をそれぞれ抽出するための機構である。

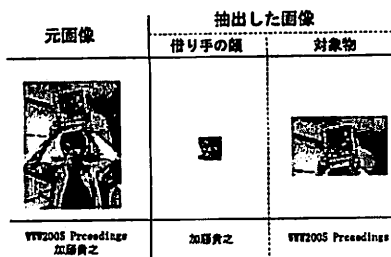


図 3 写真からのからのオブジェクト取得

図 3 はこの手法を模式的に表現したものである。この部分は計算機によって実現することとする。以下、計算機によるオブジェクト抽出の方法を考える。

まず、借り手 B の抽出は簡単に行うことができる。コンピュータ・ビジョンの技術の発展によって、一定の特徴を持つ画像を検知することが可能となった。Haar-like 特徴を用いるアルゴリズム [7] はその手法のひとつである。例えば、「人間の顔」はその見た目によって一定の特徴を持っているので、特徴を記述して計算機に与えることによって計算機は任意の画像の中から「人間の顔」の部分を見つけることができる。BOZAAR における借り手 B は人間であるから、貸し借りのイベント写真から B の画像を抽出することが可能となる。その際、ひとつ制約が必要である。すなわち、イベント写真には借り手 B 以外の人は写ってはいけないこととする。複数の人物が写っていると、そのうちのどれが B であるかを判別することができない。そこでこのように撮影対象に関する制約を設ける。

次に、対象物 g の画像の抽出方法を考える。ここではコンピュータ・ビジョンの手法は有効ではない。なぜなら BOZAAR で貸し借りされる対象物は自由だからである。したがって対象物の視覚的特徴をあらかじめ記述しておくことは不可能である。そこで本論文では、写真内での借り手と対象物との相対的な位置関係から対象物を特定するという手法を提案する。すなわち、写真内における借り手の位置は特定可能であるので、借り手と対象物の位置関係が定まっていれば対象物の位置が特定可能となる。そこで BOZAAR では撮影する写真の構図に次のような制約を設ける。図 4.1 のような「借り手の頭の上に対象物がある」という構図である。この構図で写真を撮ることによって、写真内の借り手と対象物の位置関係が固定化され、計算機が対象物の画像を抽出することが可能となる。

以上のように、撮影する写真に次のふたつの制約を設けることによって、一枚の写真から借り手および対象物の画像を抽出することができる。

- 借り手以外の顔は写ってはいけない
- 借り手の頭の上に貸し借りの対象物がある

4. 設計および実装

前章で提示した手法を実現するための画像処理システムを設計する。本手法で決めた構図の写真を入力とし、人と物それぞれの画像を出力するプログラムを設計する。(1) まず写真の画



1. 画像読み込み



2. 顔検出



4. 両画像の
切り抜き・保存



3. 物の範囲の特定

図4 画像処理のプロセス

像 I を読み込み、(2)I 内に含まれる顔を検出する。顔の範囲を長方形として特定する。この長方形を F とする。(3) 次に物の範囲 G を特定する。F の上辺を延長した直線を f とする。I の上辺、側辺および f で囲まれる長方形を G とする。(4) 最後に F と G をそれぞれ切り取り、顔の画像、物の画像として保存する。以上のプロセスを図示したものが図 4 である。

さらに、この設計に基づいて画像処理プログラムを実装した。画像処理にはインテル社が開発したオープンソースのコンピュータビジョンのライブラリである OpenCV [6] を利用した。とくに顔オブジェクトの検出に当たっては、Haar-like 特徴を用いるオブジェクト検出 [7] を用いた。また、顔オブジェクトの特徴データは OpenCV に付属しているものを使用した。

5. 実験および考察

本章では手法の有効性を検討する。検討すべき点は次の 2 点である。

- 登録機能のユーザビリティ: 貸し借りイベント登録機能において、登録したいイベントをユーザは必ず登録することができるか

- 抽出された画像の視認性: 貸し借り情報閲覧機能において、画像処理によって抽出された画像から、ユーザは実際の貸し借りの対象物・主体者を正しく想起することができるか

登録機能のユーザビリティが問題となるのは、机や車、部屋などの大きな物を貸し借りするときである。対象物を持ち上げることができないと、本研究で用いるポーズの写真の撮れない場合もあるだろう。そうだとすると、こういった物の貸し借りは BOZAAR では管理できなくなってしまう。しかし、ユーザの工夫次第ではうまく写真を撮影して貸し借り情報を登録することができるかと我々は考える。

また、抽出された画像の視認性が問題となるのは文房具や装

身具など小さな物を貸し借りするときである。小さな物は先ほどとは違って登録は容易に行うことができる。しかし閲覧機能を使ったときに、画像の中の対象が小さすぎて判別できない、あるいは背景と紛れてしまってどれが対象物なのか分からない、といった問題が派生するだろう。しかしこの場合も、ユーザが工夫することによって視認性を上げることができる可能性があるかと我々は考える。

以上の仮説を元に、大きい物と小さい物それぞれについて、ユーザの工夫によって両機能の性能を向上させる実験を行った。なお、視認性を測定するための基準として、(1) オブジェクトを抽出できた割合と (2) オブジェクトが画像中に占める割合のふたつを設ける。(1) 抽出できた割合 r は、元画像に含まれる対象物 g の面積を S_g 、 g をオブジェクト抽出した後の画像に含まれる g の面積を D_g として、 $r = D_g/S_g$ という式で求める。また、(2) 画像中に占める割合 p は、オブジェクト抽出した後の画像の面積を D 、この画像に含まれる g の面積を D_g として、 $p = D_g/D$ という計算式で求める。

5.1 大きな物の貸し借り実験

本実験では大きな物の例として、こたつ台 (幅 100cm, 奥行き 75cm, 高さ 40cm) を対象とした。この台を頭上に持ち上げることは不可能だった。第一の手法として、借り手が台の下に潜り、台の天板を持ち上げているようなポーズでの撮影を試みた。その結果、図 5 群の画像を得た。台のオブジェクト抽出の性能を測定すると、抽出できた割合は 67%、画像中に占める割合は 48% であった。

第二の手法では抽出の割合を向上させるための工夫を行った。ここでは借り手が机の上に逆立ち (ブリッジのポーズ) をして、上下逆さまの写真を撮る、という方法を用いた。その結果、図 6 の画像群を得た。この場合、台を抽出できた割合は 100%、画像中に占める割合は 25% であった。ただし物の画像において、台は上下逆さまである。

両者の手法の結果を比較すると、画像中に占める割合は第二の手法のほうが高く、抽出できた割合は第一のほうが高い。したがってどちらが視認性が高いかを定めることはできない。しかし、後者の方法のほうが視認性を高められる可能性が高いと考える。なぜなら、まず撮影者が被写体に近づくことによっていくらかでも画像に含まれる割合は高められる。さらに、抽出できた割合はある一定の値を超えればほとんど視認性に影響を与えないと我々は考える。

ともかくこの実験から、本研究における写真の構図の制約は柔軟性を持っていることが確認できた。すなわち、借り手の頭の上に対象物があるような構図、というのは、必ずしも借り手が対象物を持ち上げるポーズを指定していない。そうではなくて、ユーザが工夫することによってこの制約を満たす構図の写真を作ることができるのである。

5.2 小さな物の貸し借り実験

次の実験では小さな物の例として、USB メモリ (幅 5cm, 奥行き 1.5cm, 高さ 0.5cm) を対象とした。第一の手法では通常の体勢、すなわち頭の上に対象物を載せるポーズである。この写真から図 7 の画像群を得た。USB メモリのオブジェクト抽

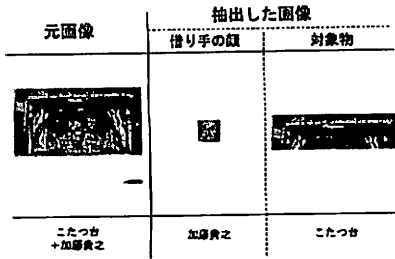


図 5 大きな物の実験 (第一の手法) で取得した画像群

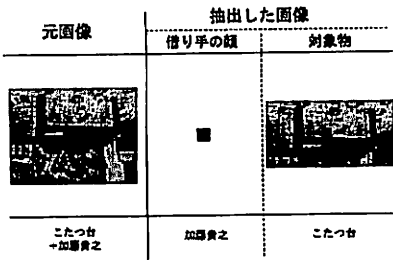


図 6 大きな物の実験 (第二の手法) で取得した画像群

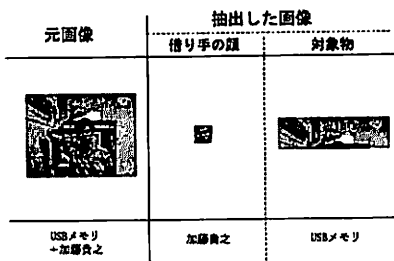


図 7 小さな物の実験 (第一の手法) で取得した画像群

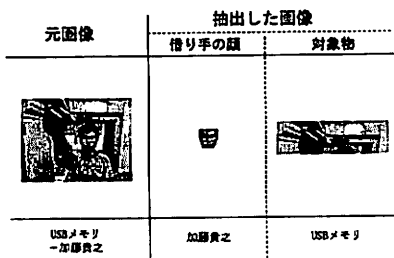


図 8 小さな物の実験 (第二の手法) で取得した画像群

出の性能を測定すると、抽出できた割合は 100%、画像に含まれる割合は 0.2% であった。

第二の手法では画像に含まれる割合を上げるための工夫を試みた。ここでは、借り手が USB メモリを持つ手を前に伸ばして、USB メモリを撮影者に向かって近づける、というポーズを用いた。借り手の顔と USB メモリの位置関係は前回と変わらないようにした。また、撮影者および借り手の顔は位置を変えなかった。その結果、図 8 の画像群を得た。このとき、抽出できた割合は 100%、画像に含まれる割合は 3.5% であった。

両手法の結果を比較すると、抽出できた割合は両者とも同じ

であり、画像に含まれる割合は 2 回目の実験のほうが高い。したがって 2 回目のほうが対象物である USB メモリの視認性は高まっている。また、2 回目の実験では写真撮影にやや特殊なポーズを用いたが、これはユーザへの負担は少ないと思われる。したがって登録機能のユーザビリティという点でも 2 回目の方法は有効である。

以上から、小さな物の貸し借りではユーザの工夫によって入力ユーザビリティを損なわずに出力の視認性を高めることができると言える。

6. おわりに

本章では本論文のまとめおよび今後の課題を述べる。

6.1 まとめ

本論文では物の貸し借りの管理と再流通の支援を行う情報システムである BOZAAR を提示した。BOZAAR では画像を用いて物や人の情報を記述することを述べた。そのための方法として、貸し借りイベントの登録のために携帯電話のカメラと電子メールを用いることにした。それをふまえて、簡単かつ安全に貸し借りイベントを登録するための手法を構築することを本論文では主に述べた。そこで、写真の構図に制約を設け、その制約を利用して一枚の写真から物と人のふたつのオブジェクトを抽出する、という手法を提案し、実装した。さらに、この手法の有効性を検討した。大きな物や小さな物を登録する際に、登録のユーザビリティおよび抽出された画像の視認性が損なわれてしまうのではないかと、という問題を提示した。その問題の確認と解決のために実験を行った。実験を通じて問題の存在を確認するとともに、ユーザの工夫によって 2 つの問題は乗り越えられることを示した。これによって本手法が BOZAAR の実現のために最低限の役割を果たすこと確認した。

6.2 今後の課題

最後に本研究の今後の課題を述べる。まず本手法の一般的な有効性を評価する必要がある。そのためにはまず基礎的な研究として、評価手法をより洗練させる必要がある。例えば本論文の中ではオブジェクトが画像に含まれる割合と視認性との関係を憶測として述べたが、この点は明らかにしたい。

最後に、研究全体の課題は BOZAAR が所有物の管理と再流通が行われるメディアとして機能することを示すことである。そのためにいち早くメールのやりとりのプロトコルおよび情報閲覧のユーザインタフェースを設計、実装したい。また、本論文では写真に収まるような形のあるものを貸し借りの対象物としたが、一般には形がないものも貸し借りされ得る。手伝いや時間など、形のないものの貸し借りをどのように支援すればよいのかも今後考えたい。

文 献

- [1] 池田大輔, 高性能 RFID リーダ付き書架の性能評価と新たな図書館サービスの提案, 第 29 回デジタル図書館ワークショップ, 2005.
- [2] 金野 紋子, 増永 良文, RFID を用いたインテリジェント冷蔵庫システムのプロトタイプング, 電子情報通信学会 第 18 回データ工学ワークショップ, 2007.
- [3] 福原知宏 [他], Ubiquitous Memories: 実世界の物理的オブジェクトを用いた記憶外在化システム, 第 15 回人工知能学会全国大

会, 2001.

- [4] 増井俊之, 本棚通信: 控え目なグループコミュニケーション, 情報処理学会 インタラクション 2005, 2005.
- [5] 椎尾一郎, James Rawan, Elizabeth Mynatt, Digital Decor: 強化された家具によるインタラクション, 情報処理学会 インタラクション 2003, 2003.
- [6] Intel, Open Source Computer Vision Library, <http://www.intel.com/technology/computing/opencv/index.htm>
- [7] R. Lienhart and J Maydt, An extended set of Haar-like features for rapid object detection, in Proceedings of International Conference on Image Processing. 2002.