

平成 22 年 5 月 28 日現在

研究種目：基盤研究（B）
 研究期間：2007～2009
 課題番号：19300062
 研究課題名（和文） 局所特徴群の近似近傍探索に基づく大規模高速画像認識技術の理論的基盤と応用
 研究課題名（英文） Theoretical foundations and applications of a technology for large-scale, efficient recognition of images based on near neighbor search on a set of local features
 研究代表者
 黄瀬 浩一（KISE KOICHI）
 大阪府立大学・工学研究科・教授
 研究者番号：80224939

研究成果の概要（和文）：

あたかも物体そのものにハイパーリンクが張られているように、カメラで物体を撮影(クリック)すると、その物体の関連情報が取得できたり、また撮影した画像を介してその物体の関連情報を登録できれば、高齢者や障害者にも優しい実世界指向インタフェースが実現できる。本研究では、100万枚の平面物体、55個の小型立体物、関東および関西の歴史的な建物、および2万ページの文書を対象として、大規模高速認識・検索技術を構築することにより、この目的を実現した。

研究成果の概要（英文）：

It is convenient not only for ordinary people but also for elderly and disabled people to have a real-world oriented interface that allows us to "click" real-world objects for obtaining their relevant information as well as to establish hyperlinks by taking their pictures. In this research, we have implemented such functionality based on a large-scale, high speed recognition/retrieval method which realizes real-time processing for databases of 1 million pictures, 55 small objects, several historical buildings in Kanto and Kansai areas, and 20,000 document pages.

交付決定額

(金額単位：円)

	直接経費	間接経費	合計
2007年度	6,800,000円	2,040,000円	8,840,000円
2008年度	4,500,000円	1,350,000円	5,850,000円
2009年度	3,300,000円	990,000円	4,290,000円
年度			
年度			
総計	14,600,000円	4,380,000円	18,980,000円

研究分野：情報工学

科研費の分科・細目：情報学・知覚情報補処理・知能ロボティクス

キーワード：画像情報処理，画像認識，画像検索，近似最近傍探索，局所特徴量，大規模画像データベース，ハッシュ表，実世界指向 Web

1. 研究開始当初の背景

(1) 画像認識への要望

あたかも物体そのものにハイパーリンクが張られているように、カメラで物体を撮影

(クリック)すると、その物体の関連情報が取得できたり、また撮影した画像を介してその物体の関連情報を登録できれば、高齢者や障害者にも優しい実世界指向インタフェース

が実現できる。従来の Web が計算機の中の世界（サイバー空間）に閉じているのに対して、この技術で実現できる Web は実世界に開かれた Web である。本研究ではこれを実世界指向 Web と呼ぶ。

このような実世界指向 Web を実現するためには、撮影によって得られた対象物体の画像を、計算機によって認識する必要がある。このため、画像認識への要望が高まっている。ここで要求される画像認識技術は、画像内に写っている物体のクラス（例えば、「車」や「犬」などの、物体の一般的な名称）を得るものではなく、画像内に写っている特定の物体（例えば、車ならば A 社のモデル B など）を識別するものである。前者のタスクが「一般物体認識」と呼ばれるのに対して、後者は「特定物体認識」と呼ばれる。

(2) 従来技術の限界

従来技術によって、このような画像認識を実現しようとする、概ね次の処理となる。まず、(a)画像から物体を切り出し、(b)切り出した画像から特徴を抽出した後、(c)データベースにある物体のモデルと照合する。問題は、(a)どこをどう切り出せばよいか分からない、(b)撮影条件や物体の姿勢によって画像が大幅に変化するため、安定した特徴を抽出できない、(c)得られた特徴を、柔軟性や高速性を保ったまま、大量の物体モデルと照合する方法がない、といった点にある。

もちろん、従来技術に工夫がないわけではない。(a)の切り出しの問題を回避し、(b)のうち姿勢の問題を解決する方法として、幾何学的ハッシング (Geometric Hashing ; GH) が提案されている。この方法は、物体から特徴点を取り出し、その配置によって物体を識別するものである。ところが、この方法を実際に利用すると、特徴点の数 N 、画像数 M に対して、 $O(N^2M)$ という膨大なメモリ量と、 $O(N^2)$ という膨大な計算量が必要となる。

(3) 基礎となる技術

当時、これらの問題を解決する基礎となる技術が出つつあった。

①局所特徴量(local feature)

局所特徴量とは、画像全体ではなく、画像の特徴的な部分から抽出される特徴量である。処理は大きく、部分を定めるための検出 (detection)、部分から特徴量を取り出す記述 (description) の 2 ステップに分かれる。局所特徴量を抽出する手法の多くは、撮影の条件が様々に変化しても、同じ物体の同じ箇所を検出し、それを照明変動に頑健な形で記述できるので、撮影条件にあまり依存しない安定した抽出が可能である。代表的な局所特徴量としては、SIFT (Scale-Invariant Feature Transform)、PCA-SIFT (Principal Component Analysis SIFT) がある。

②文書画像検索法

研究開始当初、我々のグループでは、LLAH (Locally Likely Arrangement Hashing) と呼ぶ新しい文書画像検索法を開発済みであった。この方法は、文書をカメラで撮影した画像をキーとして、対応する文書をデータベースから瞬時に検索するものである。この手法は、GH と同様に特徴点に基づく検索法であるが、メモリ量、計算量共に特徴点の数 N に対して $O(N)$ であるほか、1 万ページのデータベースに対して実時間で動作可能である。

③平面物体認識法

局所特徴量を用いた画像認識法として、ハッシュに基づく近似最近傍探索の手法を開発済みであった。この手法は、10 万枚の平面物体の画像データベースに対して検索質問あたり 20ms で検索でき、精度は 87% という性能であった。

2. 研究の目的

本研究の目的は、上記の技術を発展させると共に、統合的に利用することによって、大規模高速画像認識を実現するものである。具体的には以下の点の実現を目的とする。

(1) 大規模化

より大規模な画像データベースを扱うことのできる認識を実現する。具体的には、100 万画像のデータベースを用いた場合でも高速に動作し、精度も改善することを目標とする。また、大規模化に伴いメモリ量の問題が生じるため、これを解決する手法も検討する。

(2) 異なるメディア

平面物体を対象とした静止画だけではなく、動画を対象とした画像認識も実現する。また、全く異なるメディアとして、物体ではなく文書も対象とする。

(3) 理論的基盤

ハッシュに基づく高速認識法の理論的基盤について明らかにする。

(4) アプリケーションの開発

技術の利用シーンを考慮した、具体的なアプリケーションを開発する。

3. 研究の方法

(1) 大規模化

①高速化

我々が提案しているハッシュに基づく近似最近傍探索の手法では、ハッシュ表を用いることで高速化を達成し、多重探査 (multi-probe) と呼ぶ手法に基づいてハッシュ表を複数回探索することにより精度を確保する。多重探査の回数を増やせばそれだけ精度は確保出来るものの処理時間が犠牲になる。

本研究では、画像によって適切な探査回数が異なることに着目し、その回数を適応的に制御する仕組みを開発する。具体的には、近似最近傍探索を用いた画像認識装置を多段階に接続する方法 (多段階化) と呼ぶ技術と

開発し、有効性を検証する。

②高精度化

データベースの画像の枚数が増えたり、検索質問の画質が悪いと、認識率が下落するという問題がある。この問題に対処するため、本研究では生成型学習と呼ばれる手法を導入する。この手法は、元画像から画像劣化をシミュレートして劣化画像を生成し、認識に利用するものである。

この手法はデータベースに登録する画像の枚数を大幅に増やすため、これまで大規模画像認識に導入されることはなかった。本研究では、処理速度やメモリ量にどの程度、影響があるのかを解析する。

③省メモリ

100万枚の画像でも高速に検索するためには、データをハードディスク上に置くのではなく、メモリ上に置く必要がある。局所特徴量は多次元ベクトルであるため、必要なメモリ量は少なくなく、すべての局所特徴量をそのままの形でメモリ上に置くことは現実的ではない。

従来法では、特徴ベクトルのベクトル量子化を行うことによりこの問題点を解決している。ところが、特定物体認識の場合、ベクトル量子化によって精度が大幅に下落することが知られている。したがって、本研究では全く異なる原理による方法を開発する。

具体的には、(a)特徴ベクトルをスカラ量子化する方法、(b)特徴ベクトルを取捨選択する方法の2つを開発する。後者については、最も単純な方法はランダムサンプリングである。これを精度面で大幅に上回る方法を考案する。

(2)異なるメディア

平面物体を写した静止画だけではなく、他のメディアを認識対象に加える。具体的には、(a)動画で撮影した平面物体、(b)小型立体物、(c)建物、(d)文書の4種類を対象にする。(a)は本研究の目標であるユーザインタフェースの構築に必須となる機能である。すなわち、カメラで動画を撮り続け、認識対象が撮影されたときに、それを即座に認識し、サービスを起動するという処理である。(b)と(c)は対象が平面ではなく立体であるという点で、困難を伴うものである。平面物体認識と同じ技術が通用するかどうかを検証する。(d)については、通常の画像認識の技術では扱えない対象である文書を検索するためのテーマである。(d)は、本研究の元となった文書画像検索を発展させるものである。具体的には、ペン先に付けたカメラで撮影した極めて狭い領域であっても正しく検索するための技術を開発する。

(3)理論的基盤

我々が提案するハッシュに基づく近似最近傍探索が、従来法である GH と比べてなぜ

高速であるのかを明確にするとともに、ハッシュに基づく近似最近傍探索法の改良方法について、理論的に考察する。

(4)アプリケーションの開発

動画入力された環境の映像から、以下の認識を行うアプリケーションを開発する。

①物体認識：データベースに登録済みの平面物体を発見し、それに関連付けられたサービスを提供する。

②カメラペン：ペン先につけたカメラで撮影した文書画像をキーとして、瞬時にデータベース中の対応する文書と撮影箇所を検索する手法を開発する。これにより、筆跡の復元などの新しい応用が可能となる。

4. 研究成果

(1)大規模化

①高速化

ハッシュに基づく提案手法は、従来の代表的な近似最近傍探索を用いた手法と比べて約 10 倍の高速化を達成した。また、前述の多段階化を導入することにより、多段階化がない場合に比べて 5 倍の高速化を達成した。従って、従来法に比べて 50 倍の高速化が実現できた。

100万画像のデータベースに対する認識実験の結果を図 1 に示す。この結果からも分かるように、データベースの登録画像数が 1,000 画像から 100 万画像まで千倍になっても、処理時間は 0.85ms から 60ms へと 70 倍にしかならず、本手法の有効性が示された。

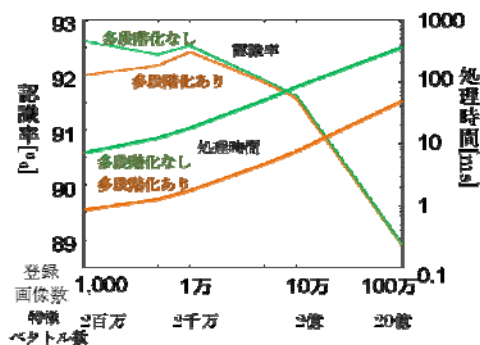


図 1 大規模化の実験結果

②高精度化

精度についても、同様に図 1 に示す。データベースの規模が小さいときには多段階化の有無で少し精度に差はあるが、大規模化に伴ってその差はほぼ無くなっている。この結果から、多段階化は精度に悪影響を及ぼさないことがわかる。

次に、生成型学習の効果について述べる。原画像 1 枚に対してボケやブレをシミュレートした 8 枚の画像を生成してデータベースに加えた。そして、携帯電話で撮影した低品質

な画像を検索質問として認識実験を行った。データベースの画像数が1万枚のとき、生成型学習ありの場合は、特徴ベクトル数が生成型学習なしの5.5倍になった（劣化した画像からは、より少ない局所特徴量しか抽出されない）。認識率については、生成型学習なしの場合の81%が、生成型学習ありの場合に92.5%まで向上した。従って、生成型学習は高精度化の効果が極めて高いといえる。処理時間については、生成型学習なしの場合に2.3msであったものが、生成型学習ありの場合に1.5msにまで短縮した。特徴ベクトルの数が5.5倍になっているにもかかわらず、処理時間を短縮することができる点は、多段階化の大きな特徴である。

③省メモリ

元の特徴ベクトルは、36次元の実数値ベクトルである。ただし、値の分布からshort型(16bit)の整数値で表現可能である。これに対して、スカラ量子化を施し、次元あたり2bitまで圧縮した。その結果、認識率は殆ど変化しなかった。このことから、局所特徴量に対してはスカラ量子化が有効であることがわかる。

次にサンプリングについて述べる。ランダムサンプリングに対して優位性を得るため、本研究では(a)画像平面上で均一にサンプリングする方法、(b)大きな局所領域を優先してサンプリングする方法(より安定した局所特徴量を保存できる)を組み合わせる用いた。結果を図2に示す。この図からも分かるように、提案手法はランダムサンプリングを大幅に改善していることに加え、元の容量の5%程度まで大幅に削減した場合、スカラ量子化よりも高い認識率が得られることがわかった。

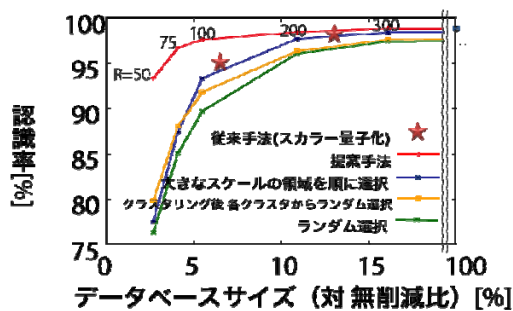


図2 取捨選択による省メモリ化

(2) 異なるメディア

①動画

動画を用いた対象の認識を実現する上で最もシンプルな手法は、動画を静止画の列とみて認識処理を施すことである。ただし、殆ど同じ画像であるにもかかわらず独立に認識処理を施すことになり無駄が多い。そのた

め本手法では、画像から得た特徴点に対してKLTトラックと呼ばれる手法を適用し、フレーム間の対応を取り、新たに画像に加わった部分のみから局所特徴量を取り出すという仕組みを導入する。これにより、撮影した動画のフレーム上に、実時間で認識結果を重畳することができる。

②立体物

フィギュアやぬいぐるみなど、55個の物体を撮影したデータベースを用いて、立体物認識の手法を構築した。この手法は、平面物体を対象とした画像認識の手法を、立体物の認識に適用したものである。すなわち、様々な角度から撮影した物体の画像を、同じ平面物体としてデータベースに格納し、認識に用いる。物体を記述するための特徴ベクトルの量は膨大になるものの、省メモリの技術と併せて用いることにより、元の特徴ベクトルの容量の1/140しか使わなくても、認識率92.9%を達成することができた。

③建物

②の立体物と同様に、神社、仏閣などの観光地の歴史的建造物のデータベースを作成し、それを用いて建物を認識する手法を構築した。

④文書

ペン先カメラで捉えた5行程度の狭い範囲であっても、対応する文書と撮影箇所を正しく検索する手法を構築した。この手法の特徴は、従来のLLAHに面積順位特徴という新しい特徴を導入し、また、データベース拡張、クエリ拡張という新しい検索方式を導入したことにより、識別性を向上させた点にある。例えば、カメラを紙面に対して60°(90°が正対)に傾けたとき、従来のLLAHでは文書画像検索精度が20.9%であったものが、クエリ拡張を行うと81.4%にまで改善することが分かった。

(3) 理論的基盤

以下の2つの成果を得た。

①GHとLLAHの比較

LLAHをGHの進化型の手法ととらえ、GHに対して4つの改良を施すことにより、LLAHを導出し、有効性をシミュレーション実験により検証した。

②近さの多段階表現と隣接バケット参照

現在、我々が用いているハッシュを用いた近似最近傍探索は多重探査を基礎とするものである。多重探査を行う際の工夫として、新たに「近さの多段階表現」および「隣接バケット参照」という方式を提案し、理論式ならびにシミュレーション実験により有効性を検証した。

(4) アプリケーションの開発

①動画に基づく物体認識とリンク

Webカメラを入力機器として、実時間で対象物を識別し、データベースに登録されてい

る場合には、その物体の画像上に矩形を描画するシステムを開発した。この矩形をクリックすると、Web 起動などの各種サービス呼び出せるものである。

②カメラペンシステム

①の手法と同様に、文書画像についてもKLTトラッカを導入し、処理の低減を図る。その上で、検索によって得られたペン先の位置をつなぎ合わせるにより、印刷文書に書いた筆記を、デジタルインクとして復元するカメラペンシステムを構築した。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文] (計 12 件)

■論文 (すべて査読有り)

1. 黄瀬 浩一, 野口 和人, 岩村 雅一, 参照特徴ベクトルの増加による低品質画像の高速・高精度認識, 電子情報通信学会論文誌D, J93-D, 2010 (掲載確定)
2. 井上 勝文, 黄瀬 浩一, 特定物体認識におけるBloomier Filterを用いたメモリ削減法とその実験的評価, 電子情報通信学会論文誌D, J93-D, 2010 (掲載確定)
3. 孫 維瀚, 黄瀬 浩一, 線画の著作権保護のための部分的複製検出法, 電子情報通信学会論文誌D, J93-D, 2010 (掲載確定)
4. 岩村 雅一, 中居 友弘, 黄瀬 浩一, 特徴点の配置に基づく画像検索手法の解析 —Geometric HashingとLLAHの比較と解析—, 電子情報通信学会論文誌D, J93-D, pp. 494-501, 2010
5. 野口 和人, 黄瀬 浩一, 岩村 雅一, 近似最近傍探索の多段階化による高速特定物認識, 電子情報通信学会論文誌D, J92-D, pp. 2238-2248, 2009
6. 井上 勝文, 三宅 弘志, 黄瀬 浩一, 電子情報通信学会論文誌D, J92-D, pp. 1686-1689, 2009
7. 野口 和人, 黄瀬 浩一, 岩村 雅一, 大規模特定物体認識における認識率, 処理時間, メモリ量のバランスに関する実験的検討, 電子情報通信学会論文誌D, J92-D, pp. 1135-1143, 2009
8. 黄瀬 浩一, 岩村 雅一, 中居 友弘, 野口 和人, 局所特徴量のハッシングによる大規模画像検索, 日本データベース学会論文誌, 8, pp. 119-124, 2009
9. 中居 友弘, 黄瀬 浩一, 岩村 雅一, Webカメラを用いたリアルタイム文書画像検索, 電子情報通信学会論文誌D, J90-D, pp. 2262-2265, 2007

■解説・総説 (すべて査読無し)

10. 黄瀬 浩一, 講座: 第2回 マルチメディア検索の最先端 大規模静止画像DB

の検索, 映像情報メディア学会誌, 64, pp. 192-197, 2010

11. 黄瀬 浩一, 特定物体認識, 電子情報通信学会技術研究報告, PRMU2009-104, pp. 79-87, 2009
12. 黄瀬 浩一, 岩村 雅一, 物体認識における照合の高速化, 画像ラボ, pp. 65-71, 2009
13. 黄瀬 浩一, 岩村 雅一, 3日で作る高速特定物体認識システム, 情報処理, 49, pp. 1082-1089, 2008

[学会発表] (計 80 件)

■国際会議 (すべて査読有り)

1. Koichi Kise, Megumi Chikano, Kazumasa Iwata, Masakazu Iwamura, Seiichi Uchida, Shinichiro Omachi, Expansion of Queries and Databases for Improving the Retrieval Accuracy of Document Portions — An Application to a Camera-Pen System, 9th IAPR International Workshop on Document Analysis Systems, June 9, 2010, Boston, USA.
2. Katsufumi Inoue, Koichi Kise, Compressed Representation of Feature Vectors Using a Bloomier Filter and Its Application to Specific Object Recognition, 1st Int'l Workshop on Emergent Issues in Large Amount of Visual Data, Oct. 4, 2009, Kyoto, Japan.
3. Koichi Kise, Kazuto Noguchi, Masakazu Iwamura, Robust and Efficient Recognition of Low-quality Images by Cascaded Recognizers with Massive Local Features, 1st Int'l Workshop on Emergent Issues in Large Amount of Visual Data, Oct. 4, 2009, Kyoto, Japan.
4. Kazumasa Iwata, Koichi Kise, Tomohiro Nakai, Masakazu Iwamura, Seiichi Uchida, Shinichiro Omachi, Capturing Digital Ink as Retrieving Fragments of Document Images, 10th Int'l Conf. on Document Analysis and Recognition, July 29, 2009, Barcelona, Spain.
5. Weihang Sun, Koichi Kise, Detecting Printed and Handwritten Partial Copies of Line Drawings Detecting Printed and Handwritten Partial Copies of Line Drawings, 10th Int'l Conf. on Document Analysis and Recognition, July 27, 2009, Barcelona, Spain.
6. Tomohiro Nakai, Koichi Kise, Masakazu Iwamura, Real-Time Retrieval for Images of Documents in Various Languages using a Web Camera, 10th Int'l Conf. on Document

- Analysis and Recognition, pp.146-150, July 27, 2009, Barcelona, Spain.
7. Koichi Kise, Kazuto Noguchi, Masakazu Iwamura, Memory Efficient Recognition of Specific Objects with Local Features, 19th Int'l Conf. on Pattern Recognition, Dec. 10, 2008, Tampa, USA.
 8. Tomohiro Nakai, Kazumasa Iwata, Koichi Kise, Accuracy Improvement and Objective Evaluation of Annotation Extraction from Printed Documents, 8th IAPR Int'l Workshop on Document Analysis Systems, Sept. 17, 2008, Nara, Japan.
 9. Koichi Kise, Kazuto Noguchi, and Masakazu Iwamura, Simple Representation and Approximate Search of Feature Vectors for Large-Scale Object Recognition, 18th British Machine Vision Conf., Sept. 12, 2007, Warwick, UK.
 10. Masakazu Iwamura, Tomohiro Nakai, and Koichi Kise, Improvement of Retrieval Speed and Required Amount of Memory for Geometric Hashing by Combining Local Invariants, 18th British Machine Vision Conf., Sept. 12, 2007, Warwick, UK.
 11. Tomohiro Nakai, Koichi Kise, Masakazu Iwamura, A Method of Annotation Extraction from Paper Documents Using Alignment Based on Local Arrangements of Feature Points, 9th Int'l Conf. on Document Analysis and Recognition, Sept. 24, 2007, Curitiba, Brazil.

[図書] (計0件)

[産業財産権]

○出願状況 (計23件)

1. 名称: 画像検索方法および画像検索プログラム
発明者: 黄瀬浩一, 岩村雅一, ほか
権利者: 大阪府立大学, ほか
種類: 特許 (PCT出願)
番号: PCT/JP2010/053446
出願年月日: 2010年3月3日
国内外の別: 外国
2. 名称: 画像データベースの作成方法および画像データベースの作成プログラム
発明者: 黄瀬浩一, ほか
権利者: 大阪府立大学, ほか
種類: 特許 (PCT出願)
番号: PCT/JP2010/053448
出願年月日: 2010年3月3日
国内外の別: 外国
3. 名称: パターン認識装置
発明者: 岩村雅一, 黄瀬浩一
権利者: 大阪府立大学
種類: 特許 (PCT出願)

- 番号: PCT/JP2010/51889
出願年月日: 2010年2月9日
国内外の別: 外国
4. 名称: 3次元物体認識用画像データベースの作成方法, 処理装置および処理プログラム
発明者: 黄瀬浩一, ほか
権利者: 大阪府立大学
種類: 特許 (PCT出願)
番号: PCT/JP2009/58284
出願年月日: 2009年4月27日
国内外の別: 外国
 5. 名称: 物体認識用画像データベースの作成方法, 処理装置および処理プログラム
発明者: 黄瀬浩一, 岩村雅一, ほか
権利者: 大阪府立大学
種類: 特許 (PCT出願)
番号: PCT/JP2009/58285
出願年月日: 2009年4月27日
国内外の別: 外国
 6. 名称: 画像認識方法および画像認識装置
発明者: 黄瀬浩一, ほか
権利者: 大阪府立大学
種類: 特許 (PCT出願)
番号: PCT/JP2007/65086
出願年月日: 2007年8月1日
国内外の別: 外国

○取得状況 (計1件)

1. 名称: 文書・画像検索方法とそのプログラム
発明者: 黄瀬浩一, 岩村雅一, ほか
権利者: 大阪府立大学
種類: 特許
番号: 特許第4332556号
取得年月日: 2009年6月26日
国内外の別: 国内

[その他]

<http://imlab.jp/LLAH/>
<http://imlab.jp/camocr/>

6. 研究組織
- (1) 研究代表者
黄瀬 浩一 (KISE KOICHI)
大阪府立大学・工学研究科・教授
研究者番号: 80224939
- (2) 研究分担者
- ◇岩村 雅一 (IWAMURA MASAKAZU)
大阪府立大学・工学研究科・助教
研究者番号: 80361129
- ◇馬場口 登 (BABAGUCHI NOBORU)
大阪大学・大学院工学研究科・教授
研究者番号: 30156541
(H19→H20: 連携研究者)