

[フェロー記念講演]

体験談：パターン認識の研究

村瀬 洋

名古屋大学 情報科学研究科 〒464-8603 名古屋市千種区不老町

E-mail: murase@is.nagoya-u.ac.jp

あらまし これまで約30年間行ってきたパターン認識の研究に対して、「画像認識と映像探索に関する先駆的研究」という名称で、IEICEフェローの称号をいただいた。本報告では、これまで研究してきた文字認識、透明物体の認識、3次元物体の認識、放送映像検索とともに、最近の研究である車載カメラ映像の認識などを、研究のノウハウとともに紹介する。

キーワード 体験談, パターン認識, 文字認識, パラメトリック固有空間法, アクティブ探索

My Experiences on Pattern Recognition Research

Hiroshi MURASE

Graduate School of Information Science, Nagoya University Furo-cho, Chikusa-ku, Nagoya, 464-8603 Japan

E-mail: murase@is.nagoya-u.ac.jp

Keyword: Experiences, Pattern recognition, Character recognition, Parametric eigenspace, Active search

1. はじめに

パターン認識は、古くて新しい研究分野である。私自身、1978年、大学4年生のときに文字認識関係の研究室に入り、パターン認識の分野に足を踏み入れ、それから約30年が経過した。一般にコンピュータに関連した研究は、当時も、今も、非常に進歩が早く内容は目まぐるしく変化する。実際に、コンピュータ関連で多くの流行となる技術やキーワードが生まれそして消えていった。その中で、パターン認識という分野がこの30年の間に、着実に進化はしているものの、学問的にはその枠組みは大きく変化しなかったことは驚きである。30年前に私が大学の修士学生のために勉強に使用した「R. Duda and P. Hart の Pattern Classification」や、「K. Fukunaga の Introduction to Statistical Pattern Recognition」は、現在の教科書にしてもまったく遜色の無いものである(改訂版がでている)。私は、初めての研究会として、学生時代にPRL(パターン認識と学習)研究会で発表したが、その当時に、30年後にも同じ研究会が存続し、その研究会で私自身が自分の研究を振り返って発表することになるとは、当時はまったく想像していなかった。ちなみに、PRL研

究会はその後、PRMU研究会に名称変更したが、実質的に同じ研究会である。

パターン認識が学問的にあまり変化しなかったのは、この分野がやさしいようで奥の深い難しい研究分野であったということに加え、ものを分類するという行為は人間にとってかなり本質的な行為であることにも起因するからであると思う。人は物事を分類するのが好きである。例えば、血液型で人の性格を分類する、誕生日の星座で人の運を分類することなどは、とても楽しい。もともと人間が生きていくということは、「安全-危険」、「好き-きらい」、「旨い-まずい」、「快-不快」などの分類行動の連続である。つまり分類が得意で、瞬時に適切な分類をする能力の高かった人類が、結果

表1. 約30年の研究歴

西暦	年齢	印象に残った研究	ノウハウ
1978	22	変動吸収特性核(文字認識)	若い頃の苦勞が一生に影響
1980	24	筆記過程の研究	アナロジーは成功も失敗もある
1983	27	候補文字ラティス法	研究タイミングは早すぎてもだめ メジャーな研究よりマイナーな研究
1989	33	透明物体の形状復元	実用研究の禁止により新しい展開に
1993	36	パラメトリック固有空間法	分野が違えば意外に受ける
2000	42	アクティブ探索法	人との出会いが大きな展開に
2004	50	車載カメラ映像認識	ポジティブマインドこそ重要

として繁栄したということかもしれない。この分類の基本原則を扱う学問であるパターン認識は、数学や物理などと同様に今後も多くの人に興味をもたれながら研究が進化していくものであると信じている。

これまでの30年間の研究体験から、私自身さまざまなノウハウを獲得した。本報告では、研究のノウハウの切り口で、私のこれまでの研究を体験談の形で紹介していくことにする。

2. 若い頃に苦労したことが一生に影響する

【研究：変動吸収特性核(手書文字認識), 1980, 22歳】

最近、大学の学生時代に指導していただいたY先生に出会ったときに、「村瀬君の研究は最近の研究を含めて、どれも学生のときと同じ発想の研究だね」と感想を言われた。たしかにそのように感じるときもある。学生(修士)の時に頑張って貯蓄した発想をいまだに消費(浪費)しているのかもしれない。

学生時代の所属研究室は、良い意味で非常に放任であった。研究は何でも好きなことをやっても良かったが、基本的に自分で考えなければいけない。そこで、画像認識の基礎知識が何もないままに、私がひたすら考え続けたのは手書き文字の変形や変動をいかに吸収するかである。それまで画像処理について勉強していない私は識別器としてはテンプレートマッチングしか思いつかなかった。しかし、考えてみるとテンプレートマッチングと形状変形の吸収は、実は相性の悪い水と油のようなアプローチである。良い発想も出てこないし、少し試してみてもまったくうまくいかない。しかしながら修論研究として仕上げるために無理やりまとめる必要がある。それで次のようにまとめた。テンプレートマッチングを勉強したての部分空間法に改良し、変形吸収は学習段階で学習画像を変形させるということで解決しようとした(図1)。

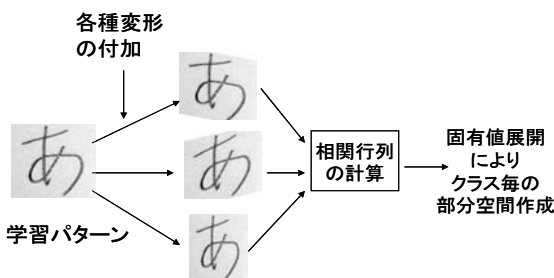


図1. 部分空間法における変形の付加

計算機上での実装段階で、部分空間法のための自己相関行列(特性核)を直接操作したために、変動吸収特性核と名づけた。認識率という観点からは、必ずしも満足できる研究結果ではなかったが、この水と油をつなげるような発想が、その後の私の発想に大きく影

響を与えた研究になっていくことは、当時はまだ気がついていなかった。また学生時代は、この研究は論文にはせずに研究のやり逃げのつもりだったが、卒業直後に先生方の暖かいご指導により、私の初めての通信学会論文誌の論文となった。論文にすることの重要性は後でわかったが、これにはとても感謝している。

3. アナロジーは参考になるが成功も失敗もある

【研究：筆記過程の研究, 1981, 24歳】

大学卒業後に、日本電信電話公社(現NTT)に入社した。当事、パターン認識といえば、音声認識と文字認識が主流であり、データの取り扱いの容易さから音声の研究のほうが先行していたように思う。またその時期、音声分野で優れた研究が出ていたので、その文字認識の研究もそのアナロジーがきくのではないかという発想があったと思う。

私が電電公社に入社して最初に与えられた訓練テーマは「筆記過程の研究」である。当事音声分野で板倉先生(同じ研究部で、隣の研究室)のパーコール音声符号化が成功を収めていた。音声の生成過程の物理モデル(声道モデル)を用いることで、効率のよい音声記述ができるというものである。そのアナロジーで手書き動作を考えると、例えば腕の物理モデルで、手書き文字の良い記述ができるということになる。しかし、調査を進めていくうちにこのアプローチに懐疑的になってきた。次の事例が特にそれに拍車をかけた。

「成人になってから交通事故で腕をなくした人が、口や足に筆記具を付けて文字を書いた場合でも、昔に腕で書いていた文字の特徴がそのまま現れる」というものである。つまり、手書き文字の形状や個人性は腕にあるのではなく脳の中にあると思われる。音声も同じ要素はあるものの、声道の制約ほど腕の制約は大きくなく、手書き動作は音声以上に脳の影響が大きいものと思われる。結局H室長と上司のW氏とさんざん議論した結果、テーマを変更することにした。

今から思えば懐疑的にならなければもう少し別の展開もあったかもしれないが、少なくとも懐疑的になるとその研究はうまくいかなくなることだけは確かである。ただ、その準備段階で以下のような成果が生まれ

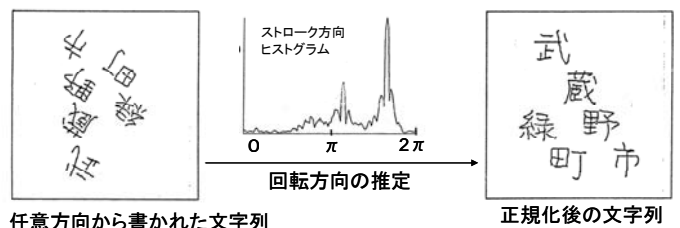


図2. 筆記方向の統計的な文字図形の回転の正規化法

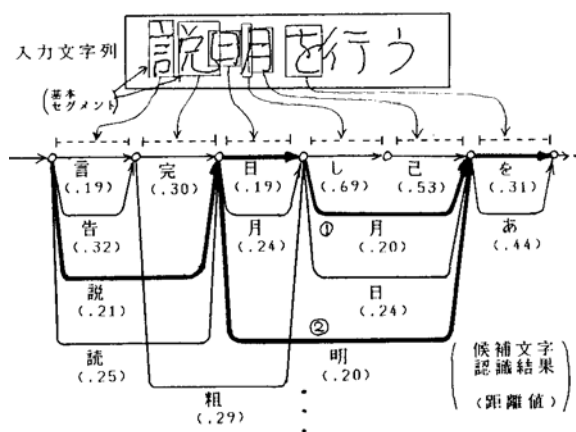
た．全国大会にしか発表していなかったが，この機会に紹介する．タブレット上に書いた文字の向きの正規化である．文字の多くは上から下に，左から右に筆記される．そのときの文字の筆記方向の統計的な性質を利用して任意の方向から書いた文字の向きを正規化するということである．(図2)

4. 研究のタイミングは早すぎても良くない

【研究：候補文字ラティス法，1983, 27歳】

音声とのアナロジーから生まれたもうひとつの研究は候補文字ラティス法である．

1980年の入社時の新人研修で他研究室の研究を聞く機会があり，興味を引いたものに連続音声の認識があった．曖昧な音声信号から言語知識や単語情報を利用しながら，単語や音素を切り出しながら音声認識するというものである．音声で利用されているような考え方で，文字の切り出し，文字認識，言語処理を統合した自由書式文字列認識ができそうに感じた．そこで提案したのが，候補文字ラティス法という自由書式連続文字列認識である(図3)．1980年当時，日本語で手書きの連続文字列認識を行っている研究(認識と切り出しを同時に行っている研究)は，私の知る限りでは無かった．また，それに付随して，偏と旁に分離するような文字の構造についてもいろいろ分析し，分離有意文字という名前を付けた．常用漢字約2000文字の中で266文字は，偏と旁に分かれそれ自身が別の文字になることを明らかにした．ちなみに，分離有意文字の例としては，「暗＝日＋音」「横＝木＋黄」「動＝重＋力」「終＝糸＋冬」などがある．



①「説」「日」「月」「を」 ---> 文節にならない

②「説」「明」「を」 ---> 文節(名詞+格助詞)

図3. 自由書式文字列の認識

ただし残念ながら，時代が早すぎたと思う．ハードウェアについてきていない．メモリ，計算機の手速と

もに，現在より1000倍程度劣っていたのである．ちなみにこの時点で我々グループが愛用していた計算機はDG社 MV/8000 でCPUクロック4.5MHz，メインメモリ512KBであった．これでも当時スーパーミニコンと呼ばれていた．一方，候補文字ラティス法では多数の文字切り出しの候補に関して，候補文字を抽出して候補文字認識する必要がある．2000カテゴリを含む文字との照合に加えてたくさんの文字候補を認識しないといけないために非常に多数の照合が必要になる．結局，1文字の計算負荷の少ないオンライン文字認識で実装したが，それでもリアルタイムには程遠く，遅かった．それからかなりの年月が経過した後にOCRで連続手書き文字列認識の研究がたくさん出てきたが，研究のタイミングを完全にはずしていたと思う．やはり研究にはタイミングが重要で，早すぎても良くない(もちろん遅いよりは良いが)ことを痛感した．

5. メジャーな研究よりマイナーな研究

【研究：オンライン図形，文字認識，1984, 28歳】

当時の所属研究室では，OCR文字認識グループとオンライン文字認識グループに分かれていて，私はオンラインのグループにいた．世の中では，OCR文字認識の研究が活発で，学会でもたくさんの発表があった．一方でオンラインはマイナーな研究分野のようで，学会での発表も少なく，研究予算も少なく寂しい思いをした．また最初は手書き入力ワープロを目指そうと言っていたが，1980年ころからキーボード入力のカナ漢字変換が急速に普及しだして，何に役に立つかも不透明になっていた．しかし，激しい競争にさらされていないことは，基礎研究としてはメリットもあったと思う．認識率を上げるという目の前のミッションがないために，アイデアを自由に広げることができるし，それが新規性にもつながる．上述した「候補文字ラティス法」も「手書き文字列の回転正規化」も，OCRの研究をしていたら試すことも出来なかったと思う．対象も文字列だけでなく「手書き図形に拡張」したり，開拓されていない研究分野をいろいろ試したりで，とても研究がやりやすかったと思う．私の学位論文も結局オンライン文字・図形認識で取得することになった．流行でメジャーな研究の中で激しく競争を繰り返すのも良いが，マイナーな研究分野で自由に発想するのも良い．この頃から研究分野に対して結構楽天的になっていったように思う．むしろ楽天的にならなければいけないと思うようになった．

6. 実用研究を禁止されたことが新しい展開に

【研究：透明物体の認識，1989, 33歳】

企業では，組織全体の戦略を考えて組織整備が行わ

れることがしばしばある。当時の私のいた研究所においても、1986年のこと、基礎研究と実用研究のミッションの明確化を目指して組織整備が行われた。その結果として私のグループでは「文字認識や画像認識などの実用よりの研究は禁止。より基礎的な研究を行え。」ということになった。これまでの画像認識の研究分野が否定されたのである。その際、上司のN氏は錯視の研究を、先輩のH氏は視覚心理の研究を始めた。もともと工学センスの強かった私は何も方向性もないままに2年程度は経過したと思う。ちょうどそのときに、上司のN氏から、「ICCVが1990年に大阪で開催されるが、日本であるから、何でも良いから出したら」と言われた。それまでコンピュータビジョンにはあまり縁の無かった私であり、しかも難関で知られるICCVであるので、自信はまったくなかったが、とにかく出そうということでそれから研究ネタを考え始めた。蓄積がないだけに、人が扱ったことのある対象だと論文を書く自信がなかったので、人が扱ったことの無い対象を扱おうと思った。研究所の中にある池を見ていて考えたのが、透明でグニャグニャした物体である。これなら誰も扱ったことは無いだろうと考え、透明非剛体物体の形状復元に関して論文を出した。屈折と統計的な性質を利用して、水底の見えの揺らぎから水面の形状を復元しようというものである(図4)。実用性もなさそうなので研究所の当時の条件も満たしていた。

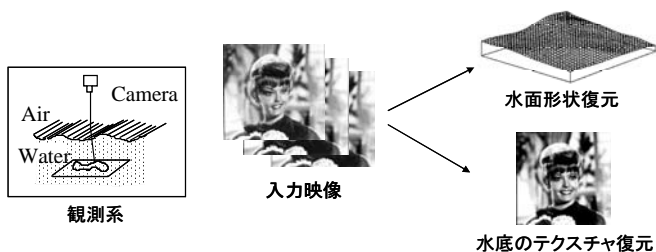


図4. 透明非剛体(水面)の形状復元

あまり深く考えた研究ではなく、極めて短期間に仕上げた研究であったが、運よくICCV1990に採録された。更にこの研究はその後、私の初めてのIEEE-PAMIの論文にもなった。しかし、この突貫工事で行った研究は、私にとって思わぬ展開のきっかけを作った。数人の海外の研究者が、この研究に非常に興味を示してくれた。その一人がNayar先生で、その後頻りにメールが来るようになった。「物体の反射特性だの、鏡面反射だの」私のそれまでの研究分野とは違う内容のメールが多かったが、Nayar先生はICCV1990で最優秀論文賞の受賞者だったので光栄なことだとは思っていた。ちょうどそのときに、NTTから海外研修の話が持ち上がったので、行き先についてはあまり深く考えずにNayarの

いるコロンビア大学に行ってみることにした。

話は違うが、それから約20年たった現在、今度はICCV2009が京都で開催される。若い人は特にチャレンジしてみると良いと思う。思わぬ展開が待っているかもしれない。

7. 分野が違えば意外に受ける

【研究：パラメトリック固有空間法, 1993, 38歳】

結局、コロンビア大学に1年間滞在することになった。渡米時には、1年間ではたいした研究はできないから、研究よりは、動向の調査と欧米流の研究手法でも勉強してくるつもりであった。そこで最低限の研究として簡単にテンプレートマッチングで物体認識でもやり、あとはアメリカをエンジョイするつもりでいた。これが、その後、意外に受ける研究に発展するとは当初まったく想像していなかった。

たしかに最初の段階では、3次元物体は向きによって見かけが大きく変わるのだからテンプレートマッチング(2次元画像照合)は筋が悪いと、周囲の研究者は否定的であったし、物体認識でこのようなアプローチはあまり無かった。ただ研究が進むにつれて、周囲の研究者に面白いとおだてられることになり、結局アメリカをエンジョイするつもりが、ついおだてにのってしまい毎日実験室で実験を行うことになってしまった。当時コロンビア大学のCS学科にはパターン認識関係の研究者が、幸か不幸かいなかったために、なんとなく興味をもたれたのかもしれない。

研究としては、固有空間上で手書き文字の変動を吸収するという大学時代の研究を、3次元物体の向きや照明の変動を吸収するように発展させた手法であるパラメトリック固有空間法を提案した。特徴空間での多様体という考えを導入することにより、物体の向きも同時に検出できる手法に発展させた。これが予想外にコンピュータビジョン研究者に興味を持たれ、この論文は、その後、被引用論文数でかなり上位になった(現

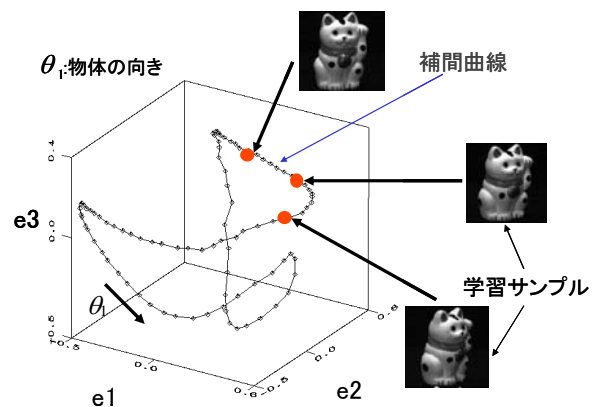


図5. 学習パターンの特徴空間上での補間

在, Google scholar で被引用数は 1100 件を超えている). 分野が異なる人ではあったが, 周囲の人に面白がられたことがこのように発展した理由だと思う(図5).

8. 人々との偶然な出会いで大きな展開に

【研究：アクティブ探索法, 2000, 42 歳】

1993 年に米国からの帰国後に, アクティブ探索法の研究を始めたが, これは度重なる偶然や周囲からの援助により何段階かのステップで進化していった.

私の数少ない体験から, 場所や組織などが変化したときには, 過去の研究の惰性に流されないように, 多少無理をしても新しい研究を始めることは良いことだと思うようになっていた. そのため米国から帰国したときに, 新しいテーマを考えていた. ちょうど隣の研究グループで, 「ウォーリーを探せ」などを例に人間の視覚を心理的に研究(視覚探索の研究)していた Y 氏の話聞く機会があり, 面白かった. それを画像処理で実現してみようと, ちょっと冒険したのがメディア探索の研究であり, アクティブ探索の始まりでもある. 特に計算速度に着目した. 従来, テンプレートマッチングで特定領域を検出するのに, SSDA などの高速アルゴリズムが提案されている. SSDA は照合 1 回にかかる時間を, 精度を保証したまま減らす手法である. しかし探索速度は, 照合回数×照合 1 回の時間であるので, 照合回数のほうを, 精度を保証したまま減らせないか考え, それを定式化したのがアクティブ探索の基本式である. 個人的には面白い手法と思っていたが, その段階では実用上は, 役立つか不明の研究であった.

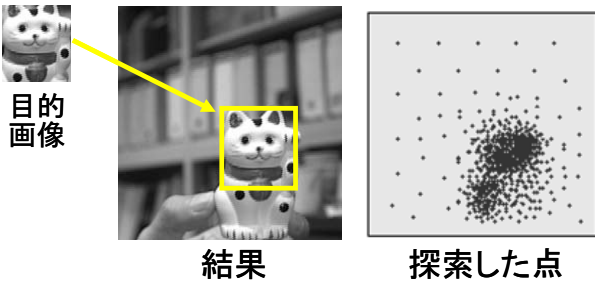


図 6. アクティブ探索法の基本

ちょうどそのとき留学生(ケンブリッジ大学の学部生)が数ヶ月インターンで滞在するので面倒を見てくれという話があった. 彼に何の経験があるかと聞いたところ, 「学生実験で音声をゼロクロス特徴で分類する実験をしたことがある」と言っていた. 期間も短いので手持ち技術の組み合わせとして, 「ゼロクロス特徴」+「アクティブ探索」で何か実験しようとなった. 簡単なゼロクロス特徴では, たいした識別は出来ず, まったく同じ信号くらいしか検出できそうもない. そ

こでテレビCMの音声を使った. CMを対象とした理由は, 出来るのはそれぐらいだということからである. 応用はまったく考えていなかった.

偶然が重なったのであるが, それから少し後に放送局のCM間引きの問題が発生し, ニュースなどで大きく取り上げられた. その際に, 本技術がCMのモニタリングにすぐに使えるのではないかという話が持ち上がった. そこで音響信号処理の経験豊かな K 氏が音響特徴抽出部分を改良して出来たのが時系列アクティブ探索法であった. 当時, 映像検索においては, 内容に基づく検索はよく研究されていたが, まったく同一のものを見つけるといー見役にたたなそうこの技術は, それがゆえにあまり類似研究は無く, 研究もやりやすかった.

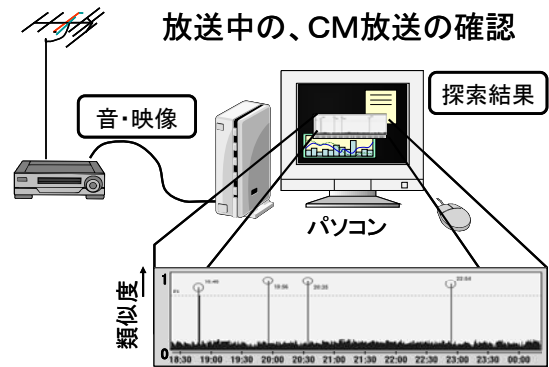


図 7. 時系列アクティブ探索法

また, 当時の上司であった H 氏の積極的な営業活動により, 放送関係, 著作権関係, 楽曲認識関係などに応用されていき, 大きな展開につながっていった. 放送局の人や, 電通や博報堂の広告代理店の人や, ジャスラックなど, 基礎研究を行っていたときにはあまり縁のない人々とたくさんの会議を持ち, いろいろな話を聞くことができ, 楽しかった時期でもある. メディア探索技術の, 具体的な実用化事例も少しずつ出てきた. 最初は実用をまったく考えなかったメディア探索の研究が, 予想外の展開で, 世の中で役にたつようになったのが愉快でもあった.

9. 楽しんで研究を行う

【研究：車載カメラ映像認識, 2004, 50 歳】

2003 年から名古屋大学に異動した. 名古屋といえば, 製造業, 特に自動車関係の企業が集中し, 同分野の研究も盛んである. そこで名古屋に異動してからは, これまでの一般の物体認識に加えて, 車載カメラ映像や車載センサーを使ったパターン認識の研究を自分だったら何がほしいかを考えて楽しんで研究を行っている.

車載カメラを用いて標識や信号機や看板などの認識や、天候の認識などに関する研究を行った。また、自動車に関するパターン認識では、地図情報をうまく利用することができる。例えば、車載カメラを積んで街の中を長期間にわたり自由に走行して収集した映像データを解析することにより、新しい建物を検出するなどの、街並み変化の認識などの研究を行っている。

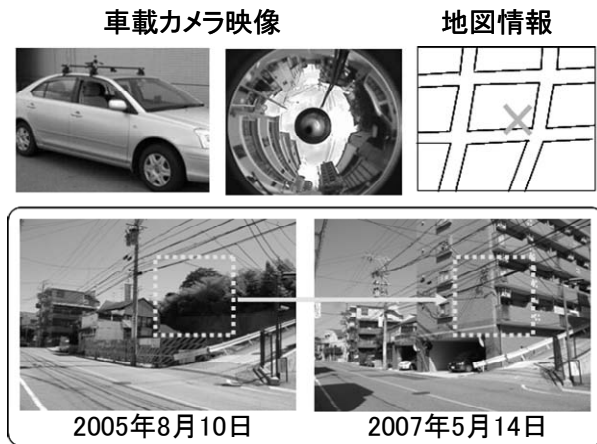


図 8. 車載カメラを用いて街並み変化を検出

1.0. ポジティブマインドこそ重要である

以上に述べたように私の研究の展開は、数多くの偶然に支えられていたと思う。特に30歳前後のときに、「すぐに実用になる研究をやりたい」と強く希望をしながら、「実用になるような研究は禁止」というミッションがあたえられ、そのミスマッチが試練になり、成長させてくれたように思う。現在の時代において、「実用にならないような研究をやれ」という余裕のある企業も大学もあまりないことが少し残念でもある。また一般に自分の境遇に何か変化が起きたときに、悪く考えがちである。しかし振り返ってみると、何か変化が起こったときに、それをポジティブに考えたときは必ずその後の良い偶然が重なり面白く展開して行ったし、ネガティブに考えるとそこで展開は終わっていることに気がつく。私自身のこれまでは、研究に関してはどちらかといえばポジティブマインドであったように思う。研究にとって重要なことはポジティブマインドであると思ふ。

1.1. むすび

本報告では、私の研究体験から得た研究のノウハウを紹介した。多分、研究のノウハウは十人十色でさまざまなものがあると思う。本報告の3でも書いたようにアナロジーは参考になるが成功も失敗もある。つまり私の体験談も、参考になるものもあれば反面教師になるものもあると思う。それを理解した上で、少しで

も若い研究者の参考になればと思う。

参考文献

- [1] K. Fununaga, "Introduction to Statistical Pattern Recognition", Academic Press, New York, 1972.
- [2] R. Duda and P. Hart, Pattern Classification, Wiley-Interscience, 1973.
- [3] 村瀬 洋, 木村 文隆, 吉村 ミツ, 三宅 康二, 変動吸収特性核を用いた手書き平仮名文字認識, 信学技報 PRL, PRL79-87 pp.7-12, 1980
- [4] 村瀬 洋, 木村 文隆, 吉村 ミツ, 三宅 康二, "パターン整合法における特性核の改良とその手書き平仮名文字認識への応用", 信学論(D), J64-D, 3, pp.276-283, 1981
- [5] 村瀬 洋, 若原 徹, 梅田 三千雄, ストロークの方向分布を利用した文字・図形の回転正規化, 信学会情報・システム部門全大, 1983
- [6] 村瀬 洋, 若原 徹, 梅田 三千雄, "候補文字ラティス法による枠無し筆記文字列のオンライン認識", 信学論(D), J68-D, 4, pp.765-772, 1985.
- [7] 村瀬 洋, 新谷 幹夫, 若原 徹, 小高 和己, "言語情報を導入した手書き文字列からの文字の切り出しと認識"信学論(D), J69-D, 9, pp.1292-1301, 1986.
- [8] H.Murase, Surface shape reconstruction of an undulating transparent object, 3rd International Conference on Computer Vision (ICCV), pp.313-317, 1990
- [9] H.Murase, Surface shape reconstruction of a non-rigid transparent object using refraction and motion, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 14, No. 10, October 1992.
- [10] 村瀬 洋, S.K.Nayar, 2次元照合による3次元物体認識 - パラメトリック固有空間法-, 信学論, J77-D-II, 11, pp.2179-2187, 1994.
- [11] H.Murase, S.K.Nayar, Visual learning and recognition of 3-D objects from appearance, International Journal of Computer Vision, Vol. 14, pp.5-24, 1995.
- [12] H.Murase, S.K.Nayar, Illumination Planning for object recognition using parametric eigenspace, IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 16, No. 12, pp.1219-1227, 1994.
- [13] 村瀬 洋, V. V. Vinod, 局所色情報を用いた高速物体探索-アクティブ探索法-, 電子情報通信学会論文誌, Vol. J81-DII, No.9, pp.2035-2042, 1998.
- [14] G.Smith, H.Murase, K.Kashino, Quick audio retrieval using active search, IEEE ICASSP, Vol. VI, pp.3777-3780, 1998
- [15] K.Kashino, T.Kurozumi, and H.Murase, A Quick Search Method for Audio and Video Signals Based on Histogram Pruning, IEEE Trans. on Multimedia, Vol.5, No.3, pp.348-357, 2003/09
- [16] 佐藤准嗣, 高橋友和, 井手一郎, 村瀬洋, GPS座標付き全方位映像群からの市街地映像マップの構築と町並み変化の検出, 電子情報通信学会論文誌, Vol. J90-D No.4 pp.1085-1095 2007/4