

[特別講演] 印象的な研究を目指して
～ユニバーサル・パターン・プロジェクトと
大阪府立大学 知能メディア処理研究室における最近の研究とエピソード～

岩村 雅一[†]

[†] 大阪府立大学大学院工学研究科 〒599-8531 大阪府堺市中区学園町 1-1
E-mail: †masa@cs.osakafu-u.ac.jp

あらまし 本稿では、筆者が携わっている文字認識・文書処理に関する研究の概要を述べるとともに、それらの研究の背景や研究が印象的になるように工夫したエピソードを紹介する。また、文字・文書分野の現状や進むべき方向について筆者が個人的に感じていることについても言及する。

キーワード 文字認識, 文書処理, ユニバーサル・パターン・プロジェクト, LLAH

[Special Talk] Toward An Impressive Research

—Recent Researches and Their Episodes of Universal Pattern Project and
Intelligent Media Processing Laboratory, Osaka Prefecture University—

Masakazu IWAMURA[†]

[†] Graduate School of Engineering, Osaka Prefecture University
1-1 Gakuencho, Naka, Sakai, Osaka, 599-8531 Japan
E-mail: †masa@cs.osakafu-u.ac.jp

Abstract In this report, I will introduce the overviews, backgrounds and episodes of my recent researches related to character recognition and document processing. I will also mention my personal opinion on the current status and requirements of the research field.

Key words character recognition, document processing, universal pattern project, LLAH

1. ま え が き

本稿のタイトルに「印象的な研究」とある理由に興味を持たれた方がいらっしゃるかもしれないが、その理由を述べるには筆者の学生時代について語る必要がある。筆者は学生時代、統計的パターン認識の少数サンプル問題に従事していた。当時は自他共に認めるプレゼン下手だったこともあり、発表の途中で興味を失っていく聴衆の冷えきった雰囲気と結局ほとんどの人に理解してもらえないまま発表を終えてしまった落胆を毎回のようには味わっていた。この苦痛から開放されたいがために、どうすれば研究内容を理解してもらえるのかと同時に、どうすれば興味の無い人にも研究に目を向けてもらえるのかを考えるようになっていった。

前者に関しては、不慣れな英語で研究内容を伝える機会などを通して、話の抽象化と筋道の立て方が徐々に身に付いてき

た。後者に関しては、人間は感情で動く動物であるため、いかに「心が動く」かが重要であると思いついた。「心が動く」とは、「何となく面白そう」とか「何か凄そう」といった感情のことである。筆者はここ数年、研究発表においてそのような感情を引き起こすことの重要性を感じ、筆者なりに研究が少しでも印象的になるように努力してきたつもりである。

本稿では、筆者が携わっている文字認識・文書処理に関する研究の概要を述べるとともに、それらの研究の背景や研究が印象的になるように工夫したエピソードを紹介する。そして最後に、文字・文書分野の現状や進むべき方向について筆者が個人的に感じていることについて言及したいと思う。

2. ユニバーサル・パターン・プロジェクトの事例

2.1 ユニバーサル・パターン・プロジェクト (u-pat)

ユニバーサル・パターン・プロジェクトは、大阪府立大学の

黄瀬浩一先生，東北大学の町真一郎先生，九州大学の内田誠一先生，筆者の4名で行っている共同研究である^(注1)。略してu-pat (ユー・パット) という。これまでに20回のオフラインミーティングと6,500通以上のメールでの議論を重ね，論文6本，国際会議11本，国内会議23本，解説記事3本，特許申請5件などの成果を挙げている。そして，ICDAR2007ではThe IAPR/ICDAR Best Paper Award [1] を，MIRU2005ではインタラクティブセッション優秀賞 [2] をそれぞれ受賞した。

学会活動においても2005年から隔年で開催しているCamera-Based Document Analysis and Recognition (CBDAR) やThe Eighth IAPR Workshop on Document Analysis Systems (DAS2008) の運営に参加してきた。また，共同研究をテーマとしたMIRU2008若手プログラム「COLLABOってMIRU?」ではメンバー4人が特別講演に招かれ，講演を行った [3]。

2.2 プロジェクトの始まり

ユニバーサル・パターン・プロジェクトは，文字認識・文書解析分野に対する熱い思いと多少の偶然で始まった。きっかけは2002年12月のPRMUであった^(注2)。ATRの萩田紀博氏から黄瀬先生，町真先生に文字・文書分野に対する強い危機感が語られ，分野を盛り上げるための活動をするように強く勧められたという。それではということで，2003年11月にまずdocument-MLが立ち上がり，続いて共同研究プロジェクトが始まることになった。当初は黄瀬先生，町真先生の2人だけの共同研究も視野にあったそうだが，たまたま筆者が町真先生の所属する研究室から黄瀬先生の研究室に異動することになったため，気が付けば筆者も加わるようになっていた。

こうして2004年7月，黄瀬先生，町真先生，筆者の3人はMIRU2004の会場で最初の打ち合わせをすることになった。3人でセッションを抜け出し，打ち合わせに適した場所を探していたところ，たまたま3人の横を早足で通り過ぎていった人がいた。内田先生だった。それに気付いた黄瀬先生が「内田せんせい」と声を掛け，戻ってきた内田先生をお誘いしたところ，「それは是非」ということになり，4人で第1回ミーティングを始めることとなった。

2.3 第1回ミーティング：現状認識と目指すべき方向

第1回ミーティングでは主に文字・文書分野の現状とプロジェクトの方向性，実施可能な研究テーマについて議論した。今から思えば，プロジェクトの大枠はこの議論で決まっていた。時を経たために記憶が定かでないところもあるが，可能な限り書き起こしてみる。

文字・文書分野の現状については，大学，企業を問わず，かつて文字認識の研究をしていた人達がどんどん違う分野に移り，コミュニティから人がいなくなっているという認識だった。その原因としては，文字認識関連で研究予算の獲得が難しくなっていること，OCRの普及によって「文字認識は解決した」と誤解している人が少なからずいることなどが考えられた。また，

簡単に解決可能な問題はほぼやり尽されてしまい，簡単ではない問題ばかりが残ってしまったという話もあった。このような状況では新しい研究テーマがなかなか開拓されないため，分野の閉塞感を生んでいるのかもしれない。

このような議論から，本プロジェクトで行う研究には，無難に論文にできそうな堅実なネタより，突拍子がないもの，独自性が強いもの，とにかく目立つものが求められた。

具体的な研究テーマを設定するにあたって，何か新しい事を始めるのならばカメラベースがいいということになった。スキャナは撮像環境をコントロールできるため，多くの場合ほぼ理想的な文書画像を得られるが，カメラベースの文字認識では様々な撮像環境が考えられるため，スキャナベースで培われが技術が使用できない事が多い。そのため，解くべき課題が山積していた [4]。

もう一つの軸として，何とか文字認識率を100%にできないだろうかという話が出た。例えば，ポスターに興味を持った利用者がホームページに手軽にアクセスしたいとする。そこでポスターに記載されたホームページのURLを文字認識機能を搭載した携帯電話で撮影し，認識されたURLを用いてホームページにアクセスする———ということができればそこで話は終わるのであるが，実際にはポスターに印刷されたQRコードを携帯電話で撮影することが多い。このように文字認識技術ではなくQRコードが使用される最大の理由は頑健さであろう。QRコードは誤り訂正符号の援用により，多少の読み取り誤りであれば正しく認識できてしまうし，何より間違った情報が復元されることがまず無い。このような頑健さを文字認識で実現する方法を求めて議論を重ね，研究テーマが徐々に具体化していった。次節ではユニバーサル・パターン・プロジェクトで行っている研究を簡単に紹介する。

2.4 研究テーマ

2.4.1 付加情報を用いるパターン認識

前述の文字認識率を100%にしたいという話から始まったのがこの研究である。100%でなくとも99.99%では駄目なのかと思われるかもしれないが，99.99%では10,000文字に1文字は誤認識になる。仮にその誤認識を手で修正するならば，最初から全て手で認識し直すのと同程度の大変な作業になるため，100%と99.99%の差は非常に大きい。

だからといって認識率100%を達成するのは容易ではないが，何とか認識率100%を達成できる状況が作り出せないだろうか。このように考えた末に辿り着いたのが認識対象である文字パターンを加工するという方法であった。これはパターン認識の問題設定からすれば反則手であるが，文字認識においてはOCRフォントやMICRフォント(図1参照)といった前例があった。

こうして我々はカメラ版OCRフォントとでも言うべき，カメラベースの文字認識に対応したフォントを作ることになった。具体的には，認識に有用な情報を字種毎に割り当てて，フォントに埋め込む方策を採った。埋め込む情報を「付加情報」と呼ぶ。図2に付加情報の割り当て例を示す。この例ではクラスAからGまで7つあり，楕円で囲まれた識別器が誤りやすいクラス(EとFなど)が既知であるとすると，1または2の(1ピッ

(注1): <http://u-pat.org/>。ちなみに，プロジェクト名が決定したのは最初のミーティングから1年4ヶ月後の第9回ミーティングのときであった。

(注2): 筆者はこのときの様子を知らないため，伝聞である。萩田氏によるdocument-MLの最初のメールにも記載されている。

0 1 2 3 4 5 6 7 8 9

図 1 MICR フォント (MICR E13-B) .

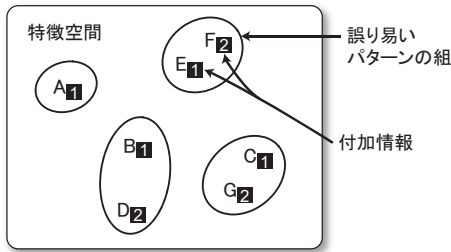


図 2 付加情報の割り当て例 .



図 3 文字への付加情報の埋め込み例 . (左) : 縞模様の複比 [5], [6] . (中) : 文字領域と影領域の面積比 [7] . (右) : 文字領域と文字領域の凸包と影領域の面積複比 [8] .

トの) 付加情報で認識率 100%が達成できる .

実際に付加情報を埋め込んだフォントを図 3 に示す . カメラで撮像すると射影歪みが生じるため , 図 3(左) は射影歪みに対する不変量である複比を利用して付加情報を埋め込んだ . この方法は機械可読性を向上する反面 , バーコード状の MICR CMC-7 フォントのように人間には若干読み難いものになった . そのため , 図 3(中) に示す方法を提案した [7] . この方法では射影歪みをアフィン歪みで近似して , アフィン不変量である面積比を利用して付加情報を埋め込んだ . さらに , 図 3(右) のような面積の複比を利用した方法も提案している [8] .

ここまでは付加情報をフォントに埋め込む方法を紹介してきたが , 付加情報は必ずしも文字パターンに埋め込む必要はない . 代わりに QR コードを利用した場合の付加情報の効果を紹介する . 書籍などの文書に書かれている情報をコンピュータが読み込む状況を想定し , 文字コードをそのまま埋め込む場合と , 認識率 100%に必要な付加情報のみを埋め込む場合で QR コードの大きさを比較する . それぞれの QR コードを図 4 に示す . 両者を比較すると , 図 4(右) は縦 , 横の大きさが図 4(左) の約 1/3 , 面積が約 1/9 である . したがって , 付加情報を用いる場合は文字コードをそのまま埋め込む場合に比べて付加すべき情報量が大幅に削減され , 効率的であることがわかる .

2.4.2 検出容易な文字パターンの生成

前述の付加情報を用いるパターン認識では , 認識性能の向上を目指して認識対象に細工する方法を議論したが , 検出・切り出しが容易となる認識対象 (文字パターン) の生成方法についても検討している [9] .

2.4.3 情報埋め込みペン

前述の付加情報を用いるパターン認識ではフォントを加工することや紙面に QR コードを印刷することが必要なため , 手書

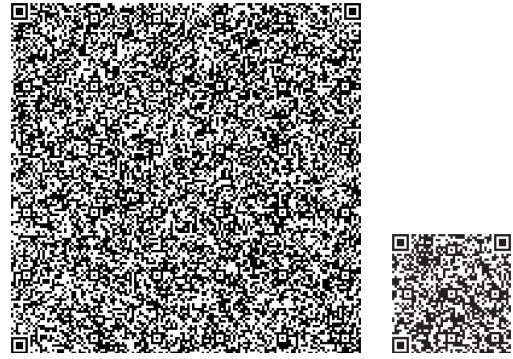


図 4 書籍などの文書に書かれている情報をコンピュータが読み込む状況を想定したときに誤りなく読み込むために必要な QR コードの大きさ [2] . A4 紙面 1 ページに日本語 3,036 字種が 1,000 文字含まれるとし , 付加情報を計算するための文字認識の誤り傾向が既知であるとした . (左) : 1 文字当たり 12 ビットの文字コードをそのまま埋め込む場合 . (右) : 1 文字当たり 1 ビットの付加情報のみを埋め込む場合 .

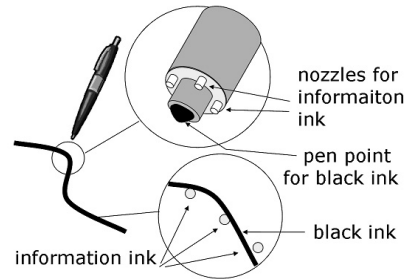


図 5 情報埋め込みペン .

き文字に情報を埋め込むことは困難である . そこで手書き文字に情報を埋め込むことを目指したのが , 情報埋め込みペンである [10] . このペンはペン先にインクを噴出する機構が付いており , 筆記と同時にインクによって付加情報も同時に記録する . 同様のペンデバイスにアノトペンがあるが , アノトペンが特殊な紙に筆記することでコンピュータに筆記データを取り込むことができるのに対して , この情報埋め込みペンは普通の紙に筆記して様々な情報を紙面に埋め込む点が異なる .

2.4.4 変量と不変量を用いた幾何歪みの復元

カメラで撮影した文書画像は様々な歪みを受けたため , そのままでは従来の文書画像処理技術を適用することができない . そのため , カメラで撮影した文書画像をスキャナで撮影した画像のように変換する研究が行われている . 既存の手法は , 文字行が平行である , 紙面の枠が長方形であるという仮定を用いたり , ステレオカメラや動画を用いる場合が多い . 我々は幾何歪みによって変化する値である「変量」と変化しない値である「不変量」に着目して , 紙面が平面である場合について , 可能な限り一般的に問題を解決することを試みた . そして , レイアウトに依存しない方法を提案している (図 6 参照) [11] ~ [13] .

3. 大阪府立大学 知能メディア処理研究室の事例

3.1 大阪府立大学 知能メディア処理研究室

大阪府立大学大学院工学研究科 知能情報工学分野 知能メディ

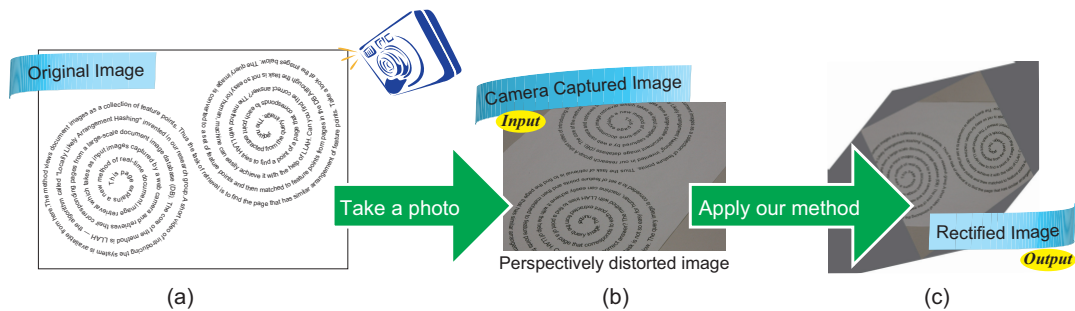


図 6 レイアウトに依らない文書の射影歪みの復元 [12].

ア処理研究室は平成 17 年 4 月に発足した比較的若い研究室である。黄瀬浩一教授、筆者の 2 名の教員に加えて、日本学術振興会特別研究員の中居友弘氏、学生 24 名 (大学院博士前期課程が 13 名、学部生 11 名) の計 27 名が在籍している。主な研究分野は文書画像解析、画像認識、パターン認識、情報検索である。

3.2 デジタルカメラを用いた高速文書画像検索

大阪府立大学 知能メディア処理研究室を代表する研究として、中居氏を中心に行ってきたデジタルカメラを用いた高速文書画像検索がある^(注3)。図 7 に示すように、この技術はカメラなどで撮影した文書画像が、データベース中のどの画像のどの部分であるかを高速に探し当てるものである。画像入力にスキャナを用いる場合と異なり、カメラを用いて取得した画像は解像度が低く、射影変換によって画像が歪んでしまう。そこで我々は単純な画像処理によって得られる特徴点の局所的な配置に着目し、幾何学的不変量とハッシュを用いて高速に照合した。検索の頑健性を保ちつつ、計算量とメモリ使用量を従来手法の数億分の一に減少させたこの手法を筆者らは LLAH (Locally Likely Arrangement Hashing) と名付けた。

この研究は、研究室発足の直前に当たる 2005 年 3 月の PRMU で最初に発表 [18] した後、改良を加えながら発展してきた。まず大幅な高速化 [14] を実現し、その後更なる高速化と省メモリ化を実現した [19]。当初は認識対象が英語などの分かち書きされる言語の文書に限られたが、まず日本語の認識が可能になり [20]、その後多言語文書の認識も可能になった [21]。また、この手法を文書の書き込み抽出を行う際の位置合わせに応用することにより、処理の大幅な高速化を実現した [22]。さらに、理論面では、LLAH が頑健で高速な検索を実現する理由を解明した [23]。

3.3 研究の始まり

この研究のきっかけは、黄瀬先生がカメラで撮影した文書画像から書き込みを抽出するテーマを B4 の学生に与え、M1 になって黄瀬先生の下に移ってきた中居氏にその補助を指示したことだった。詳細は割愛するが、書き込み抽出を実現するためにはカメラで撮影した文書画像の検索手法と位置合わせ手法が求められた。そこで当初は勉強を兼ねて文献調査や追実験などをを行いながら、有効な手法を模索していた。そうこうしている

うちに u-pat の議論で複比を使うアイデアが出たため、それならばと試してみることにになり、ここから研究が動き始めた。

しかし、ここからが問題であった。射影不変量である複比は、同一平面上の 5 点から射影歪みの有無に関わらず同じ値が計算できる理想的な代物であるが、そのためには同一の 5 点を選ばないといけないという制約があった。すなわち、点の対応と複比には、点の対応を求めるために複比を使いたいのに、複比を使うには点の対応が必要という、卵と鶏の関係があった。

この問題は何度かの「中居マジック」によって解決し、LLAH が生まれた。「中居マジック」とは、彼が時折見せる予想以上の進捗や予想外の手法のことで、黄瀬先生と二人でこのように呼んだ。LLAH によって導かれた前述の問題の解決方法は、対応する可能性がある全ての 5 点を検証する一方、1 回の検証に必要な計算量を大幅に削減することだった。これにより、検索精度を保ったまま全体の計算量を大幅に削減できた。ところが、このアルゴリズムがあまりにも変だったため、当初はアルゴリズムの本質が何で、何故うまくいくのかを簡潔に説明できなかった。この日から LLAH のより簡潔な解釈を目指した議論が始まり、仮説を立てては「がっかりする」日々が 1 年以上も続くこととなった。

3.4 その後のエピソード

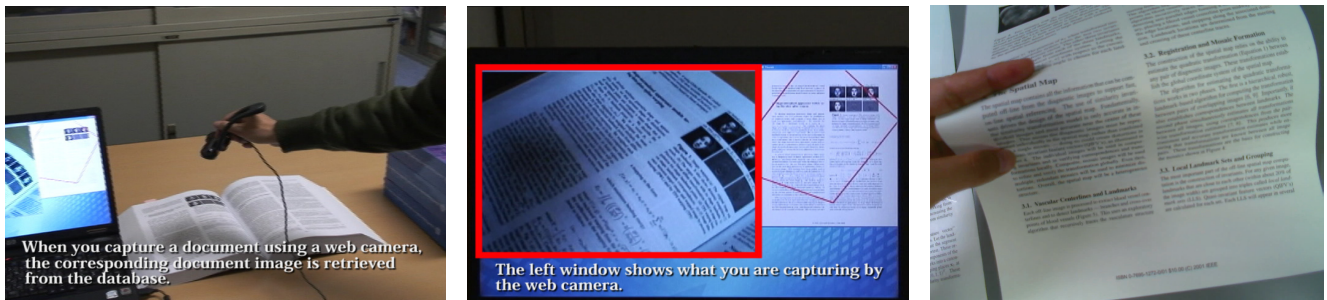
3.4.1 タイトル変更

一連の LLAH 研究の最初の発表のタイトルは、「複比とハッシュに基づく文書画像検索」 [18] という手法を説明するオーソドックスなものであった。しかし、直後の FIT2005 では「デジタルカメラによる文書画像検索 — 1 万ページから 0.1 秒で検索する —」 [14] という手法の効果を強調するタイトルに変わった。この変更には次のようなエピソードがあった。

最初の発表では画像 1 枚の認識に 100 ~ 300 秒程度の処理時間を必要としていたが、その後の FIT までに「中居マジック」によって処理時間が約 0.1 秒にまで短縮できた。実に約 1,000 倍の高速化であった。そこで、どうすればこの成果をより多くの人に知ってもらえるかと思い、頭をひねった。

その当時の実験は、データベースに 200 枚の画像を登録し、クエリ画像と同一の画像を 0.1 秒で検索するものであった。しかし、昨今の計算能力の向上具合からして、この程度の性能は数年後にちょっと頑張れば実現できそうである。これでは大して印象に残らないように思えた。認識性能や計算速度は登録画像の枚数にほぼ依存しない事が分かっていたので、画像の枚数

(注3): FIT2005 ヤングリサーチャー賞 [14], MIRU2006 デモセッション優秀賞 [15], 平成 19 年度情報処理学会山下記念研究賞 [16], 平成 18 年度電子情報通信学会論文賞 [17] の各賞を受賞した。



(a) web カメラで検索の様子

(b) 撮影画像 (画面左) と推定された撮影部分 (画面右の枠)

(c) 検索可能な画像の例

図 7 LLAH を使用したリアルタイム文書画像検索のデモの様子 (デモプログラムとデモビデオは LLAH ホームページ (<http://imlab.jp/LLAH/>) からダウンロード可能)。このデモは、130 万画素程度の web カメラで撮影した文書画像が、登録された約 5,000 枚の文書画像のどの画像のどの部分であるかを 2CPU 搭載の通常のノートパソコン上で高速に (7Hz 以上で) 探し当てるものである。回転、拡大・縮小に不変、射影変換、オクルージョンや紙面の曲がりに頑健である。特徴量を局所的に計算しているため、(c) に示すような画像でも検索可能である。

を増やせば増やすだけ手法の凄みが増すと思い、慌てて文書画像を 1 万枚かき集めることにした。登録画像 1 万枚は計算機の限界ぎりぎりであったが、無事に動いてくれた。そしてタイトルに手を入れることにした。「200 ページから 0.1 秒で検索」はいかにも地味であるが、「1 万ページから 0.1 秒で検索」であれば、聞く人に驚きを与えられる十分魅力的なタイトルに思えた。手法は何も変わらないのだが。

タイトルが良かったのかどうかは不明だが、結果的にはこの論文で中居氏がヤングリサーチ賞を受賞することができた。

3.4.2 手法を信じてもらえない!!

MIRU2005 のインタラクティブセッションで発表することになった。中居氏が会場でコンピュータビジョンの分野で著名な先生に手法を説明したところ、「そんなことはありえない。何かが間違っている。」と言われて信じてもらえなかったようだ。このときの発表はポスターの説明のみだったが、デモがあれば違った結果になったかもしれない。これが後にデモを作成するきっかけの一つになった。

3.4.3 デモセッション

MIRU2006

前述のこともあり、この頃にはデモを作成して多くの人に手法を知ってもらいたいと強く思うようになっていた。そこで MIRU2006 のデモセッションに出すことにした。デモ発表するに当たって考えたことは、特に有名でもない筆者らの手法に興味を持ってもらうにはどうしたらいいかであった。考えた末に辿り着いた答えは「初見で驚いてもらう」ことであった。最初は必ずしも手法で驚いてもらう必要はなく、見た目でも何でもいいのだが、とにかく目立つことが重要だと思った (失笑でも OK)。そうして出した結論は、検索対象の文書画像 20,000 ページを全て印刷して持参し、検索するページを「お客さん」に自由に選んでもらうことだった。そのときの様子が図 8(a) である。この発表では運良くデモセッション優秀賞を受賞でき、多くの人に手法を認知してもらった重要なきっかけになったように

思う。

ICPR2006

国内では徐々に認知してもらえるようになってきたので、次は海外に出ていくことにした。ICPR2006 のデモでは、MIRU で使用した 20,000 ページの文書の代わりに、現地で受け取る予稿集 (冊子) を使用することにした。昨晚受け取ったばかりの予稿集でデモをすれば、データベース作成の高速性もアピールでき、さらに驚いてもらえると考えた。ところが、現地に行っただけで驚いた。今回から冊子の予稿集は別途購入していないと入手できなかったらしいのだ。それならば購入しようと受付に行ったら、既に売り切れであり、今購入すれば会議後発送される旨を伝えられた。それではデモができないので、せめて貸してもらえないかと何度もお願いしたのだが、梨のつぶてであった。そんなとき、ふと岐阜大学の山本和彦先生のお顔が浮かび、大急ぎでお探しして事情をお話したところ、快く事務局に交渉してしていただけることになった。そして、嘘のようにあっさり予稿集を借りることができた。

山本先生のお陰でようやく実現したデモであるが、デモ会場は会場内の辺境の地にあり、たまた訪れた人もデモを見に来たというより、協賛企業の機器展示が目当ての人ばかりだった。しかしこのまま終わるのでは旅費が勿体無いので、通りかかった人を無理矢理呼び止めてデモを見せることにした。最初はみんな迷惑そうにしているも、面白いことに、多くの方は興味津々になって写真を撮って帰っていった。このような現象のことを筆者らは「目から鱗が落ちる音が聞こえる」と表現している。

CVPR2007

LLAH は本来平面文書を検索する手法であるが、局所的に平面があれば図 7(c) のように文書が多少変形していても検索可能である。CVPR2007 のデモでは手法の頑健性を示しつつデモを目立たせるために新たな検索対象を用意した。それは文書をアイロンプリントした図 8(b) の「文書 T シャツ」だった。折



(a) MIRU2006 でのデモの様子。



(b) 文書 T シャツ。

図 8 (a) MIRU2006 のデモでは、データベースに登録されている文書画像 20,000 ページを持参し、自由に選んでもらった。
(b) CVPR2007 のデモではこの T シャツを着てデモをした。

角なのでプリントする文書には CVPR2007 に投稿してリジェクトされた論文を選んだ。

こんな奇妙な T シャツを着た日本人がデモをしていたらきっと滑稽なはずで、からかい半分にデモを見に来てくれるに違いないと予想していたが、この予想は半分当たりで半分外れであった。結局、見に来てくれた人には馬鹿受けであったが、雑然とした会場で興味の無い人を引き付けるには至らなかった。しかし、LLAH の説明の際には、この T シャツを見せるだけでみんな最高のリアクションをしてくれるので、非常に重宝している。

4. む す び

本稿では、筆者が携わっている文字認識・文書処理に関する最近の研究の概要とそれにまつわるエピソードを紹介した。稿を閉じるにあたって、文字・文書分野の現状や進むべき方向について筆者が感じていることを少しだけ述べたいと思う。

萩田氏による document-ML の最初のメールには

どんな研究も波がある、サインカーブを書いているので、時期的に位相が 180 度 ~ 270 度付近にいる時が、研究者にとって一番チャンスである。

と書かれている。文字・文書分野が実際に 180 度 ~ 270 度の位相にあるとすれば、今求められているのは何だろうかと考えてみた。筆者なりの結論は、「何となく面白そう」とか「何か凄そう」と思ってもらえる類の研究だった。心が動くことこそが多くの人を引き付ける力であり、状況を変える糸口になると思う。

最後に、本稿は信学技報に似つかわしくない駄文であるが、何かの形で位相を進める(戻す?)一助になれば幸いである。

謝辞 本稿で紹介した研究の共同研究者である大阪府立大学 黄瀬浩一先生、東北大学 大町真一郎先生、九州大学 内田誠一先生、日本学術振興会/大阪府立大学 中居友弘博士に感謝する。

文 献

[1] S. Uchida, M. Sakai, M. Iwamura, S. Omachi and K. Kise, "Extraction of embedded class information from universal character pattern, Proc. ICDAR2007," Vol. 1, pp.437-441, Sept. 2007.

[2] 岩村雅一, 内田誠一, 大町真一郎, 黄瀬浩一, "情報付加による認識率 100%の実現—一人にも機械にも理解可能な情報伝達のために—," MIRU2005 論文集, pp.901-908, July 2005.

[3] 尾関基行, 神原誠之, 北原格, 高橋友和, 西山正志, 新田直子, 堀田政二, "MIRU2008 若手プログラム報告—COLLABOって MIRU?—," 信学技報 PRMU2008-187, Dec. 2008.

[4] 黄瀬浩一, 大町真一郎, 内田誠一, 岩村雅一, "カメラを用いた文字認識・文書画像解析の現状と課題," 信学技報 PRMU2004-246, March 2005.

[5] 内田誠一, 岩村雅一, 大町真一郎, 黄瀬浩一, "カメラによる文字認識のためのカテゴリ情報の埋込に関する検討," 信学論 D, vol.J89-D, no.2, pp.344-352, Feb. 2006.

[6] 内田誠一, 酒井恵, 岩村雅一, 大町真一郎, 黄瀬浩一, "オートマトン制御付き最適セグメンテーション法とその実環境文字認識への応用," 信学論 D, vol.J90-D, no.8, pp.1966-1976, Aug. 2007.

[7] 大町真一郎, 岩村雅一, 内田誠一, 黄瀬浩一, "実環境文字認識のための面積比による付加情報埋込," 信学論 D, vol.J90-D, no.12, pp.3246-3256, Dec. 2007.

[8] S. Omachi, M. Iwamura, S. Uchida and K. Kise, "Information embedment with cross ratio of areas for accurate camera-based character recognition," Proc. KJPR2008, pp.111-112, Nov. 2008.

[9] 服部亮史, 内田誠一, 岩村雅一, 大町真一郎, 黄瀬浩一, "検出容易な文字パターン生成に関する検討," 信学技報, Feb. 2009.

[10] 田中一弘, 内田誠一, 岩村雅一, 大町真一郎, 黄瀬浩一, "データ埋め込みペンに関する基礎的検討," ヒューマンインタフェース学会論文誌, vol.10, no.4, pp.559-567, Nov. 2008.

[11] 丹羽亮, 岩村雅一, 黄瀬浩一, 内田誠一, 大町真一郎, "変量と不変量に基づく射影歪みからアフィン歪みへの補正," MIRU2007 論文集, pp.720-725, July 2007.

[12] M. Iwamura, R. Niwa, A. Horimatsu, K. Kise, S. Uchida and S. Omachi, "Layout-free dewarping of planar document images," Proc. DRR XVI, Jan. 2009.

[13] 内田誠一, 酒井恵, 岩村雅一, 大町真一郎, 黄瀬浩一, "事例に基づく文書画像の回転角推定," 信学論 D, vol.J91-D, no.1, pp.136-138, Jan. 2008.

[14] 中居友弘, 黄瀬浩一, 岩村雅一, "デジタルカメラによる文書画像検索—1万ページから 0.1秒で検索する—," 情報科学技術レターズ, vol.4, pp.133-136, Sept. 2005.

[15] 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とその拡張現実への応用," MIRU2006 論文集, pp.1379-1380, July 2006.

[16] 中居友弘, 黄瀬浩一, 岩村雅一, "特徴点の局所的配置に基づくリアルタイム文書画像検索とその拡張現実への応用," 信学技報 PRMU2006-66, Sept. 2006.

[17] 中居友弘, 黄瀬浩一, 岩村雅一, "特徴点の局所的配置に基づくデジタルカメラを用いた高速文書画像検索," 信学論 D, vol.J89-D, no.9, pp.2045-2054, Sept. 2006.

[18] 中居友弘, 黄瀬浩一, 岩村雅一, 松本啓之亮, "複比とハッシュに基づく文書画像検索," 信学技報 PRMU2004-249, March 2005.

[19] 中居友弘, 黄瀬浩一, 岩村雅一, "処理速度とメモリ効率の改善された LLAH によるカメラベース文書画像検索," MIRU2008 論文集, pp.1252-1259, July 2008.

[20] 中居友弘, 黄瀬浩一, 岩村雅一, "LLAH によるリアルタイム文書画像検索の日本語文書への拡張," MIRU2007 論文集, pp.471-472, Aug. 2007.

[21] T. Nakai, K. Kise and M. Iwamura, "Real-time retrieval for images of documents in various languages," Handout of DAS2008, p.37, Sept. 2008.

[22] 中居友弘, 黄瀬浩一, 岩村雅一, "特徴点の局所的配置に基づく位置合わせを用いた文書からの書き込み抽出法," 信学技報 PRMU2006-245, March 2007.

[23] M. Iwamura, T. Nakai and K. Kise, "Improvement of retrieval speed and required amount of memory for geometric hashing by combining local invariants, Proc. BMVC2007," Vol. 2, pp.1010-1019, Sept. 2007.