

特別企画： わが国における IP、CV 研究の軌跡と現状

松山 隆司（京都大学）、久野義徳（大阪大学）
谷口倫一郎（九州大学）、和田俊和（岡山大学）

今回の研究会は、

- CV 研究会発足以来 100 回目の研究会に当たる
- CV 研究会の前身である IP（イメージプロセッシング）研究会発足から 20 年（正確には 21 年）が経つ
- 研究会の名称が「コンピュータビジョンとイメージメディア (CVIM)」に変わる

という様々な意味で大きな節目に当たる。そこで、本報告では、まず IP および CV 研究会のこれまでの活動を振り返り、CVIM 研究会が目指すものについて説明する。次に、CV 研究連絡委員会が現在行っている活動として、CV アルゴリズムに関する技術評論 WG および IUE ソフトウェア・データベース WG からの活動状況報告、そして最後に第 100 回研究会を記念して行ったプログラムコンテストの結果について述べる。

Special Report: Retrospective and Prospective Overviews of SIGs- IP and CV

T. Matsuyama (Kyoto Univ.), Y. Kuno (Osaka Univ.)
R. Taniguchi (Kyushu Univ.), T. Wada (Okayama Univ.)

This memorial meeting celebrates

- 100th regular meeting of SIG-CV
- 20th year (to be accurate, 21st year) anniversary of SIG-IP (antecedent of SIG-CV)
- inauguration of SIG-CVIM (renamed from SIG-CV).

Following a retrospective overview of SIG-IP and SIG-CV activities, this report gives prospective research directions of SIG-CVIM. Then, current research activities by CVCV-WG for CV algorithm survey and IUE-WG for CV/IU software and image database development are reported. Finally, the result of the program contest, which was organized as a memorial event of this meeting, is described.

1 はじめに

今回の研究会は、

- CV 研究会発足以来 100 回目の研究会に当たる
- CV 研究会の前身であるイメージプロセッシング (IP) 研究会発足から 20 年 (正確には 21 年) が経つ
- 研究会の名称が「コンピュータビジョンとイメージメディア (CVIM)」に変わる

という様々な意味で大きな節目に当たる。

そこで、本報告では、まず IP および CV 研究会におけるこれまでの活動を振り返り、CVIM 研究会が目指すものについて説明する。次に、最近の CV 研究連絡委員会における活動として、CV アルゴリズムに関する技術評論 WG および IUE ソフトウェア・データベース WG からの活動状況報告、そして最後に第 100 回研究会を記念して行ったプログラムコンテストの結果について述べる。

[記念資料の販売]

CV 研究連絡委員会では、第 100 回研究会を記念して、この 20 年間の IP、CV 研究会を中心とした研究活動の歴史をまとめた資料集を編集しました。記念資料集の内容は、

- IP および CV 研究会の歴代主査、幹事の方々からのメッセージ
- IP および CV 研究会を中心としたこの 20 年間の研究活動をまとめた年表
- IP および CV 関連の各種国内、国際学会での授賞論文リスト
- IP および CV 研究における主要論文・書籍リスト (コメント付き)
- IP および CV 研究会における全発表論文リスト (著者およびキーワードインデックス付き)

となっており、希望者に有料配布しています。購入を希望される方は、3,000 円 (資料代 + 郵送料) を現金書留で下記に送ってください。(かなり余裕をみて印刷しましたが、売り切れの場合はご容赦下さい。)

〒606

京都市左京区吉田本町

京都大学工学部電子通信工学教室

松山隆司

2 IP、CV 研究会 20 年の軌跡

情報処理学会におけるビジョン研究に関する研究会活動は、1973 年 6 月に活動を始めたイメージプロセッシング研究委員会 (主査: 坂井利之 当時京都大学教授)¹ に起源を持つ。この委員会は非公開で、国内の主な研究機関を見学し、主として画像処理装置に関する議論を行った。

当時国内では、1971 年に通産省が大型プロジェクト「パターン情報処理システムの研究開発」を始め、国際的には、1973 年に第一回パターン認識国際会議が米国ワシントン DC で開催されており、計算機を用いた画像の処理は最先端の研究分野として活発な研究が進められていた。

こうした状況を踏まえ、情報処理学会では、公開の定例的な研究発表・討論の場としてイメージプロセッシング (IP) 研究会を設け、その第一回研究会が 1975 年 6 月 23 日東京機械振興会館で開催された。この時の主査は尾上守夫 (当時東京大学教授)、幹事は高木幹雄 (東京大学教授)、長尾 眞 (京都大学教授)、白井良明 (当時電総研) であった。

IP 研究会では、研究発表・討論の場を提供するだけでなく、

- 標準画像データフォーマット
- ポータブルな画像処理ソフトウェア
- 画像処理文献データベース

に関する調査、検討を行い、これを受けて標準画像データベース SIDBA (東京大学)、ポータブルな画像処理用 FORTRAN サブルーチンライブラリ SPIDER (電総研)、オンライン画像処理文献データベース (CV 研究連絡委員会) が開発され、わが国における画像処理研究の普及、発展に大きな貢献をした。

こうした研究連絡委員会が主体となった積極的な研究活動は、よき伝統として現在に受け継がれ、最近では、

- 米国を中心として開発が進められている画像理解、コンピュータビジョン研究用ソフトウェア開発環境 IUE (Image Understanding Environment) プロジェクトへの日本代表としての参加
- IUE に対する日本独自の貢献としての Calibrated Image Database の構築

¹以下 IP、CV 研究会を指導されてきた先生方のお名前を挙げることになるが、敬称はすべて省略させていただく。

- コンピュータビジョンの各種アルゴリズムに関するサーベイの実施

が行われている。(これらについては、後にその活動状況を述べる。)

IP研究会当時の情報処理学会の規定では、研究会は1期2年で2期が上限であったため、1979年にIP研究会が廃止され、新たにコンピュータビジョン研究会が発足した。第一回CV研究会は、主査：福村晃夫(当時名古屋大学教授)、幹事：高木幹雄、鳥脇純一郎(名古屋大学教授)、田村秀行(当時電総研)によって、1979年5月19日名古屋大学で開催された。

当時、「コンピュータビジョン」という名前はそれほど一般的ではなく、研究会での発表もいわゆる画像処理アルゴリズムとその応用が中心であった。しかし、1980年代に入り、D. Marrの提案したComputational Visionの考え方が広まるにつれて、画像からの3次元距離・運動情報の復元や3次元物体の形状モデルを利用した物体認識、ロボットビジョンへの応用などが多く発表されるようになり、コンピュータビジョンのイメージが定着した。

研究活動の進展に伴い、研究会の登録会員数、研究発表件数とも順調に伸び、1980年代末には、会員数：約500人、年間発表件数：約60件という状態になった。この間、情報処理学会の規定が改正され、研究会継続の回数制限がなくなり、辻 三郎(当時大阪大学教授)、白井良明、鳥脇純一郎、谷内田正彦(大阪大学教授)、杉原厚吉(東京大学教授)の先生方がCV研究会の主査として研究会活動を指導されてきた。

1980年代は世をあげてのAIブームで、1984年には、若手ビジョン研究者が発起人となり、私的研究会としてのエキスパートビジョン(EV)研究会を組織し、知識工学の観点からのビジョン研究を行った。1986年には、EV研究会は、電子通信学会パターン認識・理解(PRU)研究会の下第3種研究会となり、1988年には、情報処理学会論文誌で「画像処理エキスパートシステム」特集が組まれた。こうした研究活動は、わが国独自の視点に立ったビジョン研究として、欧米の研究者にもかなりの影響を与えた。

CV研究会では、定例の研究会以外にシンポジウムの開催も積極的に行っており、1984年にはコンピュータビジョン・フォーラム'82がつくば研究交流センターで行われた。1980年代末からは、PRU研究会と協同してシンポジウムを開催することが始められ、1989年の「画像理解の高度化と高速化」シンポジウム、1990年の「コンピュータビジョン'90」シンポジウム、そして1992年からはMIRU(Meeting on

Image Recognition and Understanding)が隔年で開かれるようになった。また、これと並行して最近では、若手研究者のためのセミナーや、国際的研究交流を深めるために、1991年に日韓コンピュータビジョン会議、1993年には第一回アジアコンピュータビジョン会議を開催してきた。

3 CVIM研究会の目指すもの

以上述べた諸先輩の先生方の努力に支えられ、1990年ごろからのバブル経済の崩壊に伴う不景気、企業のリストラにも拘らず、CV研究会の活動は着実に成果を挙げ、最近では国際学会においてもわが国からの研究に対して各種の賞が与えられるようになっている。

しかし、その一方で、「CVの研究は研究(論文)のための研究になっており、現実世界の問題解決に役立たない」という批判が次第に広がってきたのも事実である。特に、不景気の中で応用指向を強めているわが国の企業においては、そうした見方が強い。

また、最近では、高度情報化社会における基盤技術としてのマルチメディア情報処理への期待の高まりとそれを受けた技術開発の進展を受け、通信や符号化、データベース、ヒューマンインターフェイスのように、従来コンピュータビジョンとは別の研究分野と考えられてきた画像関連分野においても、視覚情報処理に関する研究・技術開発が活発に行われるようになっている。

こうした状況の中でCV研究会の運営を任された我々は、21世紀に向けた今後の研究会活動の方向として、次のような考え方を打ち出し、研究会の名称を「コンピュータビジョンとイメージメディア(CVIM)」に変更することにした。

[1] 「コンピュータビジョン」の再定義

CVが先に述べたような批判を受けることになったのは、逆説的のいうと、

Computer Vision = Computational Vision

すなわち、「コンピュータビジョンとは、Marrの提案した視覚モデルに則った、視覚認識に対する数理的アプローチである」という見方が非常に大きな成果をもたらしたことによる。

しかし、本来、「コンピュータビジョン」は、人間の持つ高度な視覚認識機能を情報科学的観点から解明することおよび、計算機を用いて柔軟な視覚認識機能を実現することを目的とした学問分野であり、Computational

Vision はその目標への1つのアプローチの仕方にしか過ぎない。事実最近では、アクティブビジョンをはじめ、センサフュージョン、タスク指向型ビジョン、定性的ビジョン、分散協調型画像理解、といった新たな視覚認識研究の方向性を示す考え方がいろいろ提案されている。

CV 研究会をより一層活発なものとするには、「コンピュータビジョン」を広く捉え、情報科学の立場から視覚認識の問題を考える研究およびその応用をすべて包含したものとする「コンピュータビジョン」の再定義が必要であろう。

[2] 視覚情報メディア処理への展開

上記のように、

Computer Vision = Visual Recognition

と「コンピュータビジョン」を広義に捉え直したとしても、それだけでは21世紀に向けた今後の研究会活動の方向性を示すものとしては十分とはいえない。すなわち、一般に、視覚認識の目的は、画像データを解析してその「意味」を理解することにある。従来のCVでは、画像の持つ「意味」は3次元の物理世界の幾何学的、光学的、時間的特徴・構造によって与えられると考え、2次元画像データ→3次元シーン情報という情報の変換（意味理解）を行うためのアルゴリズムの研究を行ってきた。しかし、こうした一方の情報変換のみを考える世界（オープンループ情報処理世界）は専門性は高いが、その反面問題領域の幅が狭く早晩行き詰まりが露になる。

こうした問題意識に基づき、我々は、図1に示した視覚情報（イメージ）メディア処理というクローズドループ情報処理世界への展開を目指して、研究会の名称に「イメージメディア」を付け加えることにした。イメージメディア処理の世界では、

- CV と CG という相補的な情報変換プロセスを統合する
- この両者を繋ぐものとして「編集」という処理プロセスを導入する
- 認識、生成される画像・映像を撮る、見る、評価・判断する主体としての人間（あるいはロボット）を想定し、処理の目的（タスク）を明確化する
- 画像やビデオ映像、3次元物体イメージをメディアとして捉え、画像・映像メディアの処理・変換・編集・蓄積・伝送を考える

ことによって、情報の流れが閉じられる。

すなわち、CVIM 研究会では、イメージメディア処理という広い世界の中にコンピュータビジョン（視覚認識）を位置づけ、新たな視点からの研究を進めていくというわけである。もちろん、本研究会の基盤はイメージプロセッシングおよびコンピュータビジョンにあるのは変わらない。こうした問題領域の拡張によって、応用としても従来からある工業応用、ロボットビジョンに加え、インタラクティブ立体テレビ、人間の顔や手足を写した画像からの表情や運動・行動情報の抽出、電子図書館における画像・映像メディアの検索・編集・表示、3次元CTやMRIを用いた外科手術シミュレーション、スポーツや演劇・ダンスを写したビデオ映像からの行動分析や映像編集などが加わる。

イメージメディア処理は、マルチメディア情報処理関連の他の研究分野と重なる部分が多いのは事実であるが、我々は、その中核はあくまでも認識処理にあり、その問題に正面から取り組んできた本研究会における今後の研究活動から大きな成果が生まれるものと期待している。

（松山 隆司 京都大学大学院工学研究科）

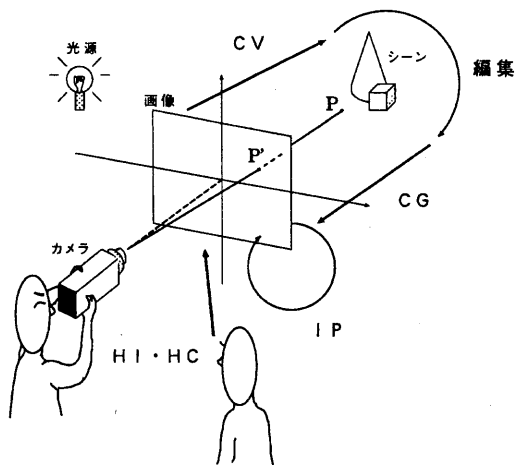


図1: イメージメディアの世界

4 CVCV-WG の活動

4.1 CVCV-WG とは

コンピュータビジョンの研究は各地で盛んに行われ、複雑なアルゴリズムの提案や多様な応用の試みが報告されている。視覚情報は人間にとって、親しみやすいものであり、視覚に関するメディアへの利用という点で、コンピュータビジョンの実際面への応用の期待は大きい。しかし、工業応用のように環境を整備できるような場合なら、コンピュータビジョンのアルゴリズムは動作するが、メディア応用等の場合のように、それが難しい場合には安定に信頼性高く動作するものを実現することは困難なのが現状である。さらに言えば、前者の場合でも、環境を整備して、結局は非常に簡単なアルゴリズムで済むようにできる場合に限ってだけ、実応用が可能なのであって、多くの複雑なアルゴリズムが研究発表されているが、それらが実際に使われるのは、あまり多くないのが現状と言って良い。

このような状況の中で、1980年代からのコンピュータビジョン研究の総括を行い²、「何ができて何ができないのか、またその理由は何か」を具体的アルゴリズムのレベルから詳細に見直すことによって、今後のビジョン研究の方向を見出すことを目的として、コンピュータビジョン研究会(今回よりコンピュータビジョンとイメージメディア研究会)では、1994年よりCVCV(Current and Vision of Computer Vision)と名付けたワーキンググループ活動を行っている。ワーキンググループは久野(阪大)と井宮(千葉大)が幹事を務め、20名程度の若手中心の研究者がメンバーとして参加している。

4.2 これまでの活動

ワーキンググループ活動の最終的な目標にはコンピュータビジョン研究全体の枠組の確立が入ってくるが、まずはコンピュータビジョンの主な個々の技術について議論をすることから活動を進めている。実際の活動の中心になっているのは、次のような形で行う技術評論である。技術評論では、ある技術についてメンバーの一人が問題点を批評するという観点からサーベイパーを作成し、それに対してメンバーの間でメーリングリストで議論を行う。そして、最終的にその議論を反映するように論文を作成者が修正し、研究会で発表する。発表の際には通常の研究発表より時間をいただき、

²70年代までのまとめとしては、コンピュータビジョン研究会の前身のイメージプロセッシング研究会におけるSPIDERの活動がある

十分な討論を行えるようにしている。技術評論はほぼ毎回の研究会で発表し、現在までに以下に示すような8回の報告を行っている。発表題目は「CVCV-WG 特別報告: コンピュータビジョンにおける技術評論と将来展望(回数を示すローマ数字)・サブタイトル」となっているので、以下では回数とサブタイトルの部分、著者、所属(発表当時)、発表年月を示す。

- I. 投票と多数決原理に基づく幾何学的対象の検出と識別
和田俊和(岡山大学), 1994.11.
- II. ビジョンにおける不変量とその応用
杉本晃宏(ATR), 1995.3.
- III. 能動視覚(Active Vision)
石黒 浩(京大), 1995.5.
- IV. 正則化(Regularization)
天野 晃(広島市大), 1995.7.
- V. 多層ニューラルネットワーク
鶴田直之(九大), 1995.9.
- VI. 固有空間法による画像認識
村瀬 洋(NIT), 1995.11
- VII. コンピュータビジョンのための画像センサ
佐藤宏介(奈良先端), 1996.1.
- VIII. 運動からの3次元復元
- 複数運動の扱いを中心に -
志沢雅彦(JRDC/東大), 1996.3.

以上の技術評論の他に、昨年と今年の4月の2回、大阪と日立基礎研で1泊の泊り込みで、コンピュータビジョンの問題について集中的に討論を行った。

4.3 今後の技術評論予定

技術評論については以下のように予定している。

- | | |
|---------|--------------------------------|
| 1996年9月 | スケールスペース(守田, 山口大) |
| | ステレオ(奥富, 東工大) |
| 11月 | フォトメトリック関連のキャリブレーション(浅田, 広島市大) |
| 1997年1月 | reconstructionの立場からの検討(徐, 立命館) |
| | メディア・通信とCV(柴藤, 松下) |
| 3月 | 情報統合(長屋, RWCP) |

その他、認知科学の観点からのCV、トラッキングとハードウェア、デフォーマブルモデルと医用応用、人間とのコミュニケーションとCVといったテーマについても技術評論を検討している。

4.4 CVCV 活動の展開

先にも述べたように、CVCV 活動の目標は個々の技術分野の評論を行うだけに留まるものではない。個々の技術評論を通じ、その技術分野で成し遂げられてきた成果を理解し、その問題点を把握するとともに、コンピュータビジョン研究全体の中での位置付けを考えたい。全体に対する位置付けを考えるためには、個々の分野だけを見るだけでは分からない。今までのところは、まずは個々について分からなければ全体を見ることもできないので、個々の分野に少し閉じて議論していたが、これからは全体の中の個々という観点からも検討し、コンピュータビジョン研究の全体像の確立という方向へもっていきたい。

全体を見るためには、いろいろな切口から検討する必要がある。その点で技術評論のテーマを見てみると、まず、正則化のように技法の点から考えたもの、ステレオのようにある情報を得るための処理という観点からのものがある。それから個々の情報を得た後の各種の情報統合する観点からの議論と、purposive な処理を重視する active vision と 3 次元シーンの reconstruction を重視する立場というような、基本的なアプローチの問題がある。さらにコンピュータビジョンの応用の観点からのものもある。それに、できればコンピュータビジョンを違う分野からとらえたようなものも含められないかと検討している。

このように、技術評論ではコンピュータビジョンのすべての項目を網羅することはできないが、重要と思われる要素を、種々の観点から選んでであると考えている。従って、技術評論とそれに続く議論を通して、全体の中に個々の技術を位置付けるとともに、全体像も確立していけるのではないかと、期待している。そして、できれば今後のコンピュータビジョン研究の展望を示したいと考えている。これらの目標は難問なので、まずはきちんとした技術評論をまとめて、それをもとに目標にできるだけ近付けるように努力していきたい。

4.5 おわりに

CVCV 活動の現状と今後の予定について述べた。研究報告の中の技術評論について質問や意見がある、技術評論でこういうテーマを取り上げたらどうか、あるいは他の点でも CVCV 活動に関して意見があれば、CVCV メンバーのメーリングリスト cvcv@top.it.okayama-u.ac.jp 宛にメールで連絡して下さい。コンピュータビジョンの将来について考えておられる皆様からの厳しい意見を期待しています。

(久野義徳 大阪大学工学部)

5 IUE-WG (J-AIUEO) の活動

Image Understanding Environment プロジェクト [1] は、現在の画像理解研究の抱える問題点、すなわち「アルゴリズムの有効性の検証が不十分である」という点や「対象世界（問題）の複雑さを規定する基準が欠如している」という点を解決すべく、画像理解研究用の標準的なソフトウェアシステムを開発しようというプロジェクトである。このプロジェクトは当初米国の ARPA のプロジェクト（GE の J. Mundy をリーダーとする IUE 技術委員会が中心となって活動）として始められたが、1994 年から欧州諸国、日本を含めた国際協同プロジェクトとして発展してきている。日本では、1994 年以来、情報処理学会コンピュータビジョン研究会の下にワーキンググループ J-AIUEO (Japanese IUE cOmmittee, 委員長：松山隆司) を組織し、IUE プロジェクトに関する活動を行っている [1]。

5.1 IUE の概要

IUE の開発は以下のような目的を達成するために始められた。

- 画像理解の計算モデルの確立
現在の画像理解研究の主流になっている、3 次元幾何光学に基づいた概念の整理・体系化を行い、それをソフトウェアとして実現するためのデータモデル、アルゴリズムを設計・開発する。
- 研究の効率化
最新の研究成果が、どこでも誰でも使えるようにすることにより、迅速で正しい研究成果の評価を可能にする。
- 技術の蓄積、技術移転の促進
最新の技術を系統的に蓄積し、すぐに利用できるようにすることで画像理解研究の成果を迅速に社会に還元できるようにする。
- 教材としての利用
IUE で提供するデータモデルやプログラミングサンプルを用いて、体系的な画像理解の教育を行うことができる。

これらの目的を達成するために具体的な要求仕様として以下のようなものが挙げられている。

- シーンの幾何学的、光学的状況の厳密な記述

- 数学的、物理的な基礎の確立
- 様々なタイプの画像への対応
- 様々なセンサーへの対応
- オブジェクト指向による記述
- データ交換フォーマットの制定
- 標準的なUNIX環境での動作
- 既存の画像処理ソフトウェア (Khoros や PIK) との融合
- ドキュメントオリエンテッドな開発³

IUE では画像理解に必要な概念、すなわちシーンを構成する要素の幾何学的、光学的関係を正確に記述するために必要な概念を数学的に定式化するとともに、それをオブジェクト指向プログラミングにおけるクラス階層として実現している。詳細は省略するが、画像、空間的オブジェクト、センサー、座標系、タスクなどの多数のクラスが定義されており、それに関連するプログラムなどが提供される予定である。

5.2 Real World Image Database の開発

J-AIUEO の活動は、米国で開発されている IUE ソフトウェアの検証・評価を行うことから始まったが、それに加えて日本独自の活動を展開している。それは、画像理解アルゴリズムのベンチマークを行うための評価用画像データベース (Real World Image Database と呼んでいる) を開発することである。すなわち、画像理解アルゴリズムの客観的評価と画像理解に関する問題の複雑さの基準作りを目指した研究を促進するための画像データベースで、次のような考え方に基づいて設計している。

1. 従来の標準画像データベースでは、画像に写されている対象世界の複雑さの度合いや、その画像を解析するというタスクがどの程度難しいのかという評価に基づく画像データの分類といった観点はほとんどなく、単に多様な画像データの集合としての意味しかなかった。これに対し、今回作成する Real World Image Database では、画像理解研究で得られた知見を基に、対象世界や認識対象の持つ複雑さを表す尺度を考案し、それに基づいて

対象世界や画像解析タスクの複雑さの分類を行う。その上で、各クラスに属する具体的画像を撮影し、データベース化する。

2. データベースに格納する画像の撮影に際しては、利用する各種センサーの特性を表すデータ、光源の位置や分光特性、画像に写された物体の詳細な形状データやそれらの配置、運動情報といった正確なキャリブレーション (グランド・ツルース) 情報を求めておく。これによって、画像を解析した結果が良いのか悪いのか、誤差はどの程度なのかを客観的に評価できる。
3. データベースに蓄えられる画像データには様々なものがあり、解析プログラムからデータを読み書きするには、標準的なデータモデル、データ構造が必要になる。IUE では多種多様な画像、物体を表すデータモデルとして C++ のクラスライブラリが用意されており、それをデータベースのソフトウェアモデルとして用いる。これにより、世界共通の画像理解研究用ソフトウェアである IUE とのスムーズなインターフェイスが実現でき、作成したデータベースが国際的にも有用で汎用性のあるものとなる。
4. 画像理解以外の画像関連研究分野においても標準画像の作成が行われており、それらの画像も可能な限り収集蓄積することにより、作成された画像データベースがより普遍的な意義を持つように努める。

このような特徴を備えた Real World Image Database を構築するにあたって、今回行った設計では、格納される画像データの特性や使用目的に応じて Real World Image Database を以下の 3 つの Image Database に分類し、各 Database の仕様およびそれらに格納される画像データの撮影、獲得法を検討した。

Calibrated Image Database

Real World Image Database の中核となるもので、画像を撮影する際の (1) 対象の幾何学的、光学的特性、(2) エネルギー源 (光源) の特性、(3) センサーの特性、(4) 対象、エネルギー源、センサーの配置 を性格に計測したキャリブレーションデータを画像と共に記録する。こうしたキャリブレーションデータを用いることにより、画像を解析した結果の客観的、定量的評価が可能となる。

³LaTeX で書かれた仕様書から自動的に C++ のコードを生成することを試みている

Thematic Image Database

ある目的で作られたシステムやアルゴリズムがどの程度多様な状況で機能するかの検証や、他のシステム、アルゴリズムとの相対的な性能比較を可能にするため、厳密なキャリブレーションにはこだわらず、テーマ（目的）別に撮られた画像を収集、蓄積する。たとえば、いろいろな状況で自動車や人間を写した動画像、特殊なセンサーやセンサー・ワーク（制御方式）で撮った画像を格納する。

Standard Image Database

これは、これまでの画像処理、信号処理研究で広く使われ多くの論文に載っている有名な標準画像を集めたもので、論文に載っているアルゴリズムの追試や新たなアルゴリズムとの性能評価が行える。

5.3 現在までの進捗状況

平成7年度末までに得られた、J-AIUEOの成果は以下の通りである。

1. Calibrated Image Databaseの構築

- 画像理解の「問題」の複雑さを明らかにするために、「問題」が何に依存して決まるのかを明らかにする必要がある。そこで、「問題」を頂点として「問題」を規定する要素間の依存関係を調査し、それを画像理解のためのワールドモデルと定めた。そのうえで、このワールドモデルとIUEで定義されるクラス階層との対応関係を調査した。
- Calibrated Image Databaseを構築する際には、撮像センサーの原理や限界といった特性分析を十分に行う必要がある。そこで、画像処理・計測・理解の視点から見た撮像センサーの特性分析を行った。また、現在利用可能なセンサーの調査を行った。
- Calibrated Image Databaseを作成する際の対象物体は、その幾何学的、光学的特性がきちんと判明している必要がある。そこで、以下の仕様の高精度3次元物体を作成した。
形状 球、円錐、円柱、角錐、角柱、立方体、凹面体
表面特性 鏡面反射特性（ステンレス）と完全拡散面に近い特性（石膏）
精度 ± 0.1mm

- キャリブレーション技法に関して、特にセンサーを中心に、キャリブレーションのためのデータとキャリブレーション法についての調査を行うと共に、いくつかの新しい手法を開発した。

2. Thematic Image Databaseの構築

- 予備的な画像の撮影として、以下のようなものを撮影した。
 - － xy 自動ステージを利用した疑似動画像列
 - － 3次元曲面物体のイメージスキャナによる撮影
 - － 高精細高濃淡分解能カラー画像
2048 × 2048 画素、14 ビット階調／各RGB
- Thematic Image 収集のためのチェックリストを作成した。作成したチェックリストは、基本事項（撮影の日付や場所、撮影者など）、センサー、対象物体、エネルギー源に対応した4部から成っている。

3. Standard Image Databaseの構築

画像処理・画像理解関連研究でよく用いられてきた画像を中心に以下のようなものを収集した。

- SIDBA (Standard Image DataBase)
- SVR (Standard Video for Research)
- Vision List Archive
- その他 (Lena など多くの論文で使われてきた画像)

4. ソフトウェアシステムの整備

- IUE ベータバージョン (V1.1) の Solaris 版をインストールし、いくつかのサンプルプログラムの動作確認を行った。IUE は極めて膨大なクラスライブラリを備えているため、現状ではコンパイルに必要なメモリと時間が極めて大きく、また、実行形式の大きさもかなり大きなものになっている。
- WWWサーバー⁴を開設し、ワーキンググループで行った調査報告やIUEドキュメントなどを公開している。また、簡単な画像検索ツールも公開している。

⁴URL=<http://www.image-lab.or.jp:80/IUE/>

ここでは、紙面の制約上、IUE ワーキンググループの活動状況の大筋しか報告できないが、これらの成果は上記の WWW で公開されているので、是非ご覧いただき、ご意見をお寄せいただければ幸いである。

参考文献

- [1] 松山, 谷口: Image Understanding Environment (IUE) プロジェクト, 情報処理, Vol.36, No.3, pp.203-211 (1995)

(谷口 倫一郎 九州大学大学院システム情報科学研究科)

6 プログラムコンテスト 結果報告

コンピュータビジョン研究会では、第 100 回研究会を記念し、図形検出プログラムのコンテストを行なった。コンテストの課題は、「ノイズが混入した 2 値画像から画像に含まれる直線分を検出するプログラムを作成する」というものである。応募者総数は 12 人、応募プログラム数は 13 であった。

6.1 プログラムの評価法

プログラムの評価は、各応募プログラムによる線分抽出結果の画像と抽出された線分の本数から算定された評価値、および、プログラムの実行速度について行なった。

6.1.1 評価用画像

プログラムの評価用に用いた画像は、理想的なデジタル直線分が複数本描かれた 512×512 の大きさの正解画像にホワイトノイズを加えて作成した。ノイズの混入した画像は、

1. 直線描画時に直線上の各画素において $1/2$ の割合で白点を混入する。
2. さらに、画像上の各画素において $1/10$ の割合で正解画像の白黒を反転させる。

という手続きにより生成している。また、正解画像中のデジタル直線分は、端点座標と線幅を乱数によって発生させたものと、人工的に描いたものを用意した。以下に、各画像生成方法を示す。

test-pattern1: 乱数により端点座標と線幅 (1 ~ 3 画素) を生成し、長さが 30 以上になる線分を 2

0 本描いている。(図 2: 正解画像の白点数の割合: 96.373%)

test-pattern2: test-pattern1 の線幅を全て 1 にしたもの。(図 3: 正解画像の白点数の割合: 98.159%)

test-pattern3: 人工的に生成した四角形と三角形 (線幅は 1 ~ 3 画素) から成る画像。(図 4: 正解画像の白点数の割合: 97.339%)

6.1.2 評価値

このコンテストでは、精度に関してはいかに正しい図形要素を検出することができるかという観点から評価を行なった。具体的には a) 検出結果を画像としてみた場合の正確さ、b) 直線分の本数の正確さ、の 2 つを考慮して評価値を定めた。

a) 出力された線分抽出結果の画像を正解画像と比較することにより、画素あたりの“平均正解率”を求めることができる。この平均正解率は、白点のみから成る画像に対しては前述の“正解画像の白点数の割合”と一致するため、白点のみから成る画像を生成してもかなりの高得点が得られてしまう。したがって、

$$score = \frac{\text{平均正解率} - \text{正解画像の白点数の割合}}{1 - \text{正解画像の白点数の割合}} \quad (1)$$

を線分検出結果を画像として見た場合の精度の評価値とした⁵。但し、score が負になった場合は、score = 0 とする。

b) score に、検出された線分の本数の正確さを表す係数をかけて、以下のように評価値を求めた。

$$\text{評価値} = \frac{score \times \alpha}{|\text{正解の本数} - \text{検出された本数}| + \alpha}$$

定数 α を小さくすると、評価値は検出される直線の本数の正確さに大きく左右されることになり、大きくすると評価値は score とほぼ同じになる。今回は、 α は“正解の本数 $\times 2$ ”として、検出される直線の本数の多少の誤りは許容することにした。

2) 速度に関しては、unix の time コマンドを用い、ユーザプロセス、カーネルプロセスの 2 つで消費された時間の合計で評価した。

⁵false positive, false negative の値から、評価値を計算する方法も考えられるが、これでは直線が存在しない地の部分での誤りの評価が低くなり、目視評価と大きくずれてしまうため、今回の評価では上述の評価式を採用した。

6.1.3 動作条件

募集時の問題記述が曖昧であり、閾値などのパラメータを外部から与えられるようになっているプログラムが多く見受けられた。パラメータを与えなければ動作せず、デフォルト値も指定されていないものや、あまりにも結果が不正確になるものが多かったため、基本的には外部パラメータを1つに制限して、検出される直線の本数が正解の本数にできるだけ近付くように調整した⁶。

複数のパラメータが与えられるプログラムの場合は、直線の本数に影響を与えるパラメータを一つだけ選んで調整し、その他のパラメータはデフォルト値を用いた。

パラメータを外部から与えることができないプログラムのうち、動作に影響を与えない範囲で外部からパラメータを与えるように変更可能なものについては、プログラムを変更した。この変更ができないプログラムについては、評価の際に考慮することとした。

尚、プログラムのコンパイル、実行には SS10-Model41 (CPU: SuperSPARC, CLOCK: 40MHz, SPECint92: 52.6, SPECfp92: 64.7) を用い、コンパイラは gcc2.6.3、コンパイル時のオプションは -O2 -lm を用いた。

6.2 プログラムの紹介

投稿されたプログラムで採用されているアルゴリズムは、次のように分類できる。

[前処理]

a) なし、b) 孤立点除去 (連結性のチェック)、c) ノイズ点除去 (局所パターンの直線性チェック)、d) 細線化

[Hough 変換]

(1) 直線パラメータの表現

a) $\rho-\theta$ 、b) fhlt (区分直線)、c) $\gamma-\theta$ 、d) dual a-b (a-d, $\gamma-\omega$ と同じ)

(2) パラメータ空間のサンプリング

a) パラメータ空間自体を用いない、b) 一様

(3) 画像に対する投票処理の制御 (戦略)

a) 一様処理、b) local-to-global、c) coarse-to-fine、d) その他

(4) 動的なパラメータ調整 (閾値など)

a) あり、b) なし

(5) 投票画素のサンプリング法

a) ランダムサンプリング、b) 一様サンプリング

[線分検出]

(1) 線分セグメンテーション時のチェック項目

⁶ プログラムが非常に低速である場合や、パラメータを変化させても検出される直線の本数が正解の本数に近付かないものについては、細かいパラメータの調整が行えなかった。

a) 密度、b) 間隙、c) 線長、d) 線幅

(2) 線分パラメータの推定

a) 角度、切辺、線長、b) 角度、切辺、線長、線幅

(3) 近接直線の処理

a) なし、b) Non-maximum suppression、c) バックマッピング、d) 隣接線分のリンキング

以下、各応募プログラムの概略を紹介をするが、応募者名は匿名とし、プログラム名のみを用いて紹介する。プログラム名は基本的に投稿者がつけた名前を用いるが、以下の場合については名称を変更した：

- 同じプログラムの名称がある場合について「番号」をプログラム名の後ろにつけて区別する。
- プログラム名から投稿者の名前が分かる場合には、途中でプログラム名を切り、後ろに「i」をつけ、名前を変更したことを示す。
- 異なるコンパイル・オプションを与えることによって、使用するアルゴリズムが大きく変わる場合には、後ろに「i」をつけ、使用するアルゴリズムを表す文字列を付加した。

また、プログラムを変更して外部パラメータを与えられるように変更したものについては「パラメータ数」の欄に「*」を記している。

hough-1 (パラメータ数0)

テンプレートマッチングによる線素検出と、それを用いて、 $\rho-\theta$ Hough 変換を行なっている。投票値の極大点を投票値が高い順に検出しながら、バックマッピングを行なう。

contest (パラメータ数5)

直線を x 軸となす角で4つのクラスに分割し、各クラスごとにパラメータ空間を用いない Hough 変換に似た計算で直線候補を絞り込む。得られた直線候補上で、連結成分を求め線分の抽出を行なう。

findline (パラメータ数0)

画像の分割を行ないながら線分検出を行なうことにより短い線分も正確に検出しようとするプログラム。コンパイル時のオプションでパラメータ空間として a-d と fhlt のいずれかが選択できるため、それぞれに対して findline-ad, findline-fhlt というプログラム名を付けた。

dtht (パラメータ数1)

画像中のあらゆる2点間に一定密度以上の特徴点が存在すれば、線分を検出するというアルゴリズム。

cht (パラメータ数1*)

画像を複数の部分画像に分割し、各々の内部で Combinatorial Hough を計算するアルゴリズムを用いており、動作速度は速い。

TestAtib (パラメータ数0)

ノイズ除去、直線候補の生成 (角度パラメータは、 θ ではなく MUFF のそれに近い)、バックマッピング、線分候補の抽出、線分候補のマージ、という処理を行っている。

hough-2 (パラメータ数1)

ノイズ除去、 ρ - θ Hough 変換による直線候補の生成の後、「ローカル法」と呼ばれる分離尺度を基準としたセグメンテーションを組み合わせている。

kura- (パラメータ数1)

ノイズ除去、 ρ - θ Hough 変換による直線候補の生成の後、(線分上の黒点数) - (線分上の白点数) $\times$ (定数) で表される「適応値」を最大化する線分を探索し、線分を求めるプログラム。

myhough (パラメータ数1)

ノイズ除去、 ρ - θ Hough 変換による直線候補の生成の後、線分検出を行なうが、線分の評価を行なう部分が効率良くコーディングされている。

hough_nisi- (パラメータ数1*)

ノイズ除去、 ρ - θ Hough 変換を行なった後、直線のパラメータを変化させながら画像とのマッチングを行なう。この後、線分間の長さ (白画素数) に応じて線分検出を行ない、バックマッピングを行なう。

fiht (パラメータ数1)

高速インクリメンタル Hough 変換法 (FIHT) を投票部に用いたプログラム。ノイズ除去、FIHT、ピーク抽出を行なった後、ローカル法を用いた線分検出を行っている。

rvht (パラメータ数1)

ランダムに特徴点をサンプリングして投票し、閾値に達したセルを順次抽出する方式を用いたプログラム。この方式では抽出されたピークを通過する投票曲線をクリアしている。

6.3 評価結果

test-pattern1~3 に対する個別の評価の結果を表 2 に示す。この表の正解率、評価値、時間の欄を加算した結果は以下の表 1 のようになる。この結果から、正解率では 'TestAtib' が、評価値では 'hough_nisi-' が、時間では 'myhough' が最も優れていた。

本プログラムコンテストでは、計算速度、図形検出精度のそれぞれについて最優秀プログラムの作者に「高速プログラム賞」、「高精度プログラム賞」を贈る。

● 速度に関しては、線分検出部分を効率良く作成することにより、短い計算時間で、精度の面でも特に問

表 1: 総合評価結果

プログラム	正解率合計	評価値合計	時間合計 [sec.]
hough-1	2.881784	0.0000	763.2
contest	2.920308	0.1739	28.2
findline-ad	2.903329	0.0000	269.5
findline-fiht	2.882400	0.0000	261.9
dtht	2.909149	0.1058	1516.7
cht	2.936416	0.5760	21.2
TestAtib	<u>2.960502</u>	0.3960	457.9
hough-2	2.914699	0.2577	252.2
kura-	2.938763	0.6999	45.4
myhough	2.931496	0.4427	<u>11.4</u>
hough_nisi-fiht	2.951477	<u>1.1891</u>	50.0
fiht	2.920994	0.3247	294.4
rvht	2.912224	0.1332	176.9

題のない直線分の検出を行なった 'myhough' が最も優れていると判定した。

● 図形検出精度に関しては、'hough_nisi-' では直線の本数を指定できるのに対して、'TestAtib' ではパラメータの調整をプログラム内部で行なっているため、単に評価値で評価したのでは公平性の点で問題が残る。この点を考慮するため、パラメータを指定せず 'hough_nisi-' を動かし、得られた結果に基づいて最優秀プログラムを判定することにした。この結果 正解率の合計は、2.975457、評価値の合計は 0.807794 となり、いずれも 'TestAtib' より優れていた⁷。'hough_nisi-' で採用されている基本的な考え方は、Hough 変換で得られた直線パラメータに誤差が含まれていることを考慮し、得られた直線パラメータを変更しながら、与えられた画像に最も良くマッチする直線パラメータを求めるというものである。このように、'hough_nisi-' は Hough 変換の特性を良く理解したプログラムであり、精度に関して最も優れていると判定できる。パラメータを指定せず、'hough_nisi-' を動かして得られた出力結果を図 5~7 に示す。

以上の評価結果に基づき、

「高速プログラム賞」: 'myhough' (作者: 九州大学大学院 総合理工学研究科 牧山 靖君)

「高精度プログラム賞」: 'hough_nisi-' (作者: 九州大学大学院 総合理工学科 西島 成 君)

とした。

(和田 俊和 岡山大学工学部)

⁷ パラメータの省略が可能な他のプログラムについても同様の計算を行なったが、正解率の合計、評価値の合計が、この結果よりも大きくなるものはなかった。

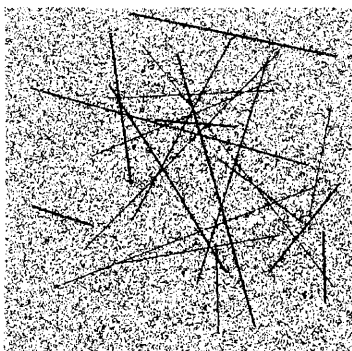


図 2: test-pattern1

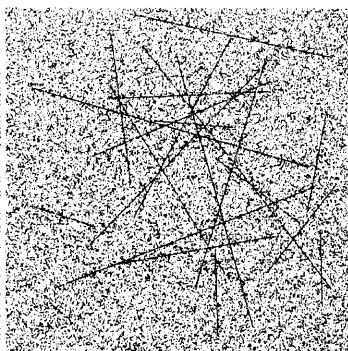


図 3: test-pattern2

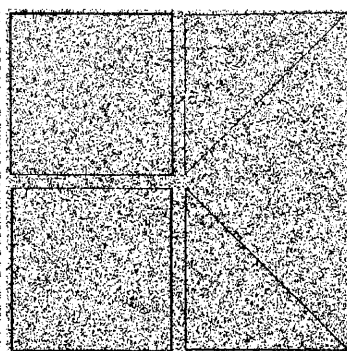


図 4: test-pattern3

表 2: 各テストパターンに対する評価結果

プログラム	test-pattern-1				test-pattern-2				test-pattern-3			
	直線 の数	正解率	評価 値	時間 [sec.]	直線 の数	正解率	評価 値	時間 [sec.]	直線 の数	正解率	評価 値	時間 [sec.]
hough-1	20	0.9569	0.0000	255.1	23	0.9619	0.0000	253.8	12	0.9629	0.0000	254.3
contest	20	0.9634	0.0000	12.6	19	0.9788	0.0000	7.9	14	0.9780	0.1739	7.7
findline-ad	30	0.9574	0.0000	91.4	30	0.9730	0.0000	88.5	23	0.9729	0.0000	89.6
findline-fiht	41	0.9521	0.0000	89.2	29	0.9670	0.0000	85.9	27	0.9633	0.0000	86.8
dtht	365	0.9345	0.1058	283.8	192	0.9942	0.0000	887.3	300	0.9805	0.0000	345.6
cht	20	0.9728	0.2500	8.7	20	0.9823	0.0381	5.8	13	0.9813	0.2879	6.7
TestAtib	71	0.9831	0.0000	203.6	20	0.9889	0.3960	90.7	63	0.9885	0.0000	163.6
hough-2	20	0.9654	0.0464	89.2	19	0.9705	0.0000	75.4	15	0.9788	0.2113	87.6
kura-	20	0.9742	0.2881	11.7	20	0.9846	0.1652	11.8	14	0.9800	0.2466	21.9
myhough	20	0.9697	0.1658	4.3	19	0.9825	0.0488	3.4	15	0.9792	0.2281	3.7
hough_nisi-	20	0.9790	0.4214	17.6	20	0.9882	0.3595	17.5	14	0.9843	0.4082	14.9
fiht	20	0.9695	0.1580	106.8	19	0.9735	0.0000	90.7	13	0.9780	0.1667	96.9
rvht	19	0.9404	0.0000	62.2	19	0.9715	0.0000	55.3	15	0.9768	0.1332	59.4

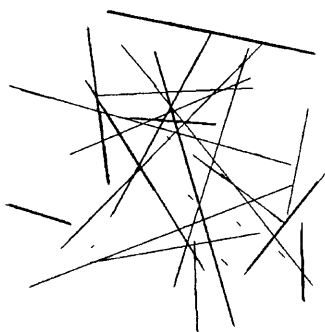


図 5: hough_nisi-の test-pattern1 に対する出力結果

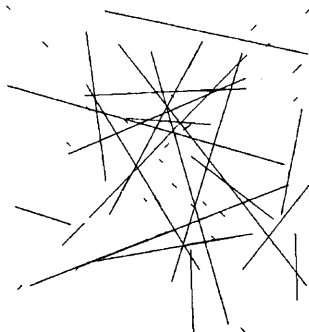


図 6: hough_nisi-の test-pattern2 に対する出力結果

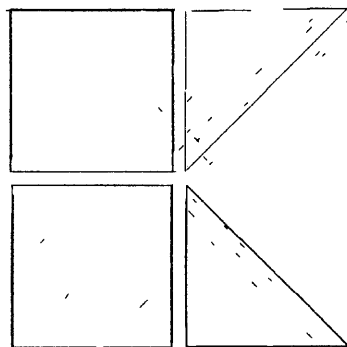


図 7: hough_nisi-の test-pattern3 に対する出力結果