

コンピュータビジョンとパターン認識会議 CVPR'94報告

村瀬 洋, 加藤晃市 (NTT), 浅田 稔, 佐藤嘉伸, 顧 海松(大阪大学),
黄瀬浩一 (大阪府立大学), 川嶋稔夫 (北海道大学),
赤松 茂, 大谷 淳, 志沢雅彦(ATR), 宮島耕治 (NTTデータ),
杉本和英 (RWCP), 藤原浩次 (ミノルタ) (順不同)

murase@apollo3.ntt.jp, kato@nttcvg.ntt.jp, asada@robotics.ccm.eng.osaka-u.ac.jp,
yoshi@medsch.osaka-u.ac.jp, ko@am.sanken.osaka-u.ac.jp, kise@ss.cs.osakafu-u.ac.jp,
kawasima@huie.hokudai.ac.jp, akamatsu@hip.atr.co.jp, ohya@atr-sw.atr.co.jp, shizawa@hip.atr.co.jp,
miya@rd.nttdata.jp, sugimoto@trc.rwcp.or.jp, fujiwara@eie.minolta.co.jp

1994年6月, 米国シアトルで開催されたIEEEのコンピュータビジョンとパターン認識の会議
(CVPR'94) に出席したので, その概要を報告する。

A Report on IEEE CVPR'94

Hiroshi Murase, Koichi Kato (NTT), Minoru Asada, Yoshinobu Sato,
Kaisyo Ko (Osaka University), Koichi Kise (Osaka Furitsu University),
Toshio Kawashima (Hokkaido University), Shigeru Akamatsu, Jun Ohya,
Masahiko Shizawa (ATR), Koji Miyajima (NTT Data),
Kazuhide Sugimoto (RWC), Koji Fujiwara (Minolta)

This report describes an overview of the 1994 IEEE Computer Vision and Pattern Recognition Conference (CVPR'94), which was held at Seattle, USA in June.

1. はじめに

1994年6月19日から26日にかけてシアトルのウェスティンホテルにてIEEEのコンピュータビジョンとパターン認識に関する会議(IEEE Conference on Computer Vision and Pattern Recognition, CVPR94)および関連ワークショップが開催された。報告者らはこれらの会議に参加したのでその概要を紹介する。CVPRは1977年にスタートして以来ほぼ毎年のペースでアメリカで開催され(但し、米国でCV関係の大規模な会議が開催される年は除く)、CV分野ではICCVやECCVと並んで注目度の高い会議の

表1. 項目一覧 (@印は今回報告の項目)

パネルディスカッション (3件)	
@PNL1. Computer Vision: Where do we stand?	
@PNL2. The role of Computer Vision in Multimedia	
PNL3. Object Representation for Object Recognition	
一般講演 (20項目)	
@S1. Low Level 1 (Surfaces)	4件
@S2. Object Recognition 1	5
@S3. Pattern Recognition/ System 1	4
@S4. Low Level 2 (Curves and Contours)	4
@S5. Motion 1	4
S6. 3D Vision	5
S7. Motion 2	5
@S8. Low level 3 (Filtering)	4
@S9. Object Recognition 2	5
S10. System and Applications	5
@S11. Active Vision 1	4
S12. Physics-Based Vision 1	4
@S13. Document Processing	5
S14. Pattern Recognition/ System 2	4
@S15. Motion 3	5
S16. Object Recognition 3 (Proj. Geometry & Invariance)	4
S17. Physics-Based Vision 2	4
@S18. Active Vision 2	5
@S19. 3D Models	5
S20. Low-Level Vision 4 (Feature Extraction)	4
ポスターセッション (7項目)	
P1. Low Level Posters	13
P2. Pattern Recognition/ System Posters	17
P3. Active Vision Posters	14
@P4. 3D Vision Posters	5
P5. Object Recognition Posters	6
@P6. Motion Posters	19
P7. Physics-Based Posters	4
ワークショップ (6項目)	
@W1. Workshop on Visual Behaviors	
@W3. Workshop on the Role of Functionality in Object Recognition	
@W4. Workshop on Visualization and Machine Vision	
@W5. Workshop on Biomedical Image Analysis	
W6. Workshop on Performance Versus Methodology in Computer Vision	

一つである。CVPRの本会議の規模はプロシーディング一冊に収まる範囲に例年抑えられているが、それに付随して開催されるワークショップは毎年増加傾向にあり会議全体の規模は大きくなっている。ワークショップはその時代の動向を反映したものであり、その存在の重要性は増している。今回の会議では表1に示す20項目の講演セッション、7項目のポスターセッション、および6種類のワークショップが開催された。今回のCVPRの採択件数は投稿472件中講演発表が87件(採択率18%)、ポスターが77件(講演とポスターと合わせた採択率34%)であった。今回のプログラム委員長は地元ワシントン大学のLinda ShapiroとSteven Tanimoto、また全体委員長はフロリダ大学のKevin Bowyerであった。

今回の会議ではベストペーパー賞としてNTTの村瀬らの'Illumination planning for object recognition in structured environments' (物体認識関係)が、また学生優秀賞としてウイスコンシン大学のKutulakosらの'Occcluding contour Detection using affine invariants and purposive viewpoint control' (アクティブビジョン関係)が選ばれ、バンケット会場で受賞式が開催された。日本からのこの賞の受賞は初めてだそうである。

本報告は会議に参加した有志がその会議の様子をCV研究会で紹介するためにまとめたものであり、各人が自分の興味のある項目について互いに重複することなく概要を記述したものである。誰からも申し出のなかった項目の紹介は省略されている。本報告は会議全体の概要の紹介であるため、もし個々の論文の技術内容に興味がある方は直接プロシーディングを参照したり、本報告のその部分の執筆者(括弧内に記入)に議論を持ちかけたりして欲しい。以降、パネルディスカッション、講演、ポスター、ワークショップの順に報告する。(以上、村瀬)

2. パネルディスカッション

PNL1. Computer Vision: Where do we stand? (村瀬)

CVの過去と現在を反省し将来を予測するという興味深いパネルディスカッションが、J. Aggarwal (U. Texas), E. Grimson (MIT), J. Mundy (GE), E. Riseman (U. Mass.)をパネリストとして開催された。各人が、過去にコンピュータビジョンで成功したものと、失敗したものと、および将来展望の各項目についてリストアップし報告した。独断や偏見はあるものの以下のような項目が報告された。

応用として成功したものは、目的の明確なもので、例えば、1次元処理ではバーコードリーダーなど、2次元ではインスペクション(回路基板の検査等)、文書処理、医用画像処理、3次元では例えば野菜などの品質分類などである。またハードウェアの進歩によりリアルタイム処理も成功した。手法的に成功したものは、ステレ

オ、デフォーダブルモデル、不偏特徴量などがあげられた。

失敗したものとしては、必ずしも全員の合意はとれていないと思われるが、以下のものが報告された。まず商業的には多くが失敗した。例えば画像処理関係のベンチャー企業も多くは倒産した。一般室外シーンの解析、アナログVLSI、ニューラルプログラミング、マルコフランダムフィールドモデル、物理ベースビジョンなども、実際のインパクトがない、あるいは実現が困難過ぎるなどの理由で失敗例として報告された。幾何学的なトポロジーに基づく手法も実データへの適用が容易ではなく成功したとは言えない。また研究体制の問題として、研究成果が共有されない、評価が不十分であり科学的でないとする意見もあった。

将来展望としては、まず身近で深刻な問題としてCVでPhDを取った学生の就職先が非常に少ない、新しくCVの会議ばかり増えて内容的に新しい研究結果が少ない、実際の画像処理の問題は最初に感じるよりずっと困難である等の暗い面も存在するものの、今後のCVのキーワードとしては、バーチャルリアリティ、マルチメディア、インタラクティブ処理、柔軟な学習、目的の明確な処理、医用画像処理、データベース、低価格リアルタイム、文脈の利用、アクティブビジョンなどがあげられた。

PNL2. The role of Computer Vision in Multimedia (藤原)

「マルチメディアにおけるコンピュータビジョンの役割」と題されたこのパネルセッションは、以下のパネラーを迎えて行なわれた。E. Delp (Purdue University), E. Hannah (Intel), A. Pentland (MIT), R. Gove (Texas Instruments), R. Jain (University of California, San Diego), D. Petkovic, B. Dom (IBM)。話は、主にマルチメディアのどの分野にコンピュータビジョンのテクニックを応用できるかを、画像圧縮や画像データベースの鍵となる技術を紹介しながら進められた。特に、パネラーであるDr. Pentlandが、コンピュータビジョン研究者にマルチメディアへの積極的な関与を促しているのが印象的であった。また、Texas Instrumentsから、Multimedia Video Processor(MVP)とDigital Micro-Mirror Device(DMD)が紹介され、研究者の注目を集めていた。なお、このパネルセッションのハンドアウトは、以下のanonymous ftpで手に入れることができる。FTP SITE: skynet.ecn.purdue.edu (128.46.154.48)。DIRECTORY: /pub/dist/delp/cvpr1994

3. 講演発表

S1. Low Level 1 (Surfaces) (志沢)

表面復元とセグメンテーションに関する4件の発表があった。C.V.Stewart(Rensselaer Polytechnic Institute)は、

outlierが50%以上存在しても良い結果が得られるロバストな表面推定法MINPRANを提案した。outlierが一様に分布するという仮定だけを用いた。この方法では、局所的に平面パッチを仮定した場合、最適解を得るために $O(N^4)$ の時間計算量を要する。S個のランダムサンプルを用いて準最適解を $O(SN \log N + N^2)$ の時間計算量で求めるアルゴリズムを提案した。彼は、別の論文でこの方法を多重表面の復元にも応用した。S.H.Lai & B.C.Vemuri (Univ.Florida)は、表面復元と初期視覚に表れるポアソン方程式を $O(N)$ で解く新しいアルゴリズムを提案した。従来、multi-grid法も同じオーダの計算量を必要としていたが、定義域の形状に制限があった。

S2 & S9. Object Recognition 1 & 2 (宮島)

10件の発表のうち、3件が航空写真中の建築物の抽出に関するものであり、これらはエッジ情報から建築物を構成する線分を仮説・検証することにより建築物を抽出している。また、3件が画像から得られた特徴点とモデルとの対応を効率良く行う手法に関するものであった。

NTTの村瀬らは、認識システムにとって最適な照明の方向や色などの照明条件をプランニングする問題を提起し、それを実現するためにパラメトリック固有空間表現を用いる手法を提案した。光源の方向、RGBのフィルタの種類をパラメータとしてどのような照明条件が対象物の認識率にとって良いかを示している。尚、これは今回のCVPRのBest Paper受賞論文である。Weizmann大のHel-Orらは、カルマンフィルタに基づいたモデルベースの姿勢推定方法を提案した。従来の手法では、認識対象を多面体のような剛体とし、その稜線の位置に関する拘束条件を考慮した認識手法が提案されていたが、これは、関節を持つ対象物や非剛体にも適用できる手法である。CMUのRouxらは、複数の視点から撮影された航空写真中の建築物を切り出す手法を提案した。これは、画像から抽出された画像特徴から3次元建築物の頂点を求め、異なる視点から撮影された画像間に対応するように頂点を結んで建築物の稜線を求め、建築物の屋根の面を切り出している。CMUのMcGioneらは、航空写真中の建築物の屋根から得られる水平・垂直方向のエッジ成分が空間的な制約を満足するように、エッジの組み合わせを求めることより建築物を抽出する手法を提案した。USCのLinらは、単眼で得られた航空写真中の建築物を抽出する手法を提案した。これは、エッジ画像から得られた線分から構成される平行四辺形を抽出し、エッジの途切れ、接続状態、及び太陽光線の入射方向と建築物の影を基に、仮説・検証によって、建築物の屋根の領域を抽出する手法を提案している。UCBのOlsonは、物体モデルと画像から得られた特徴点の対応において、木探索に基づいた計算コストが少なくなるような3次元物体の姿勢推定方

法を提案した。イスラエル工科大学のRivlinらは、物体を構成するパーツの機能を推論するfunction-basedの物体認識手法を提案した。これは、物体を構成する円柱などのプリミティブな物体を超二次関数で表し、これらの位置関係、接続条件の制約によって物体を認識している。British Columbia大学のBeisらは、3Dモデルベースの物体認識に使う学習インデックス関数を提案した。これは隠蔽などで特徴が欠落した場合に、仮説の曖昧性を解析することによって、画像からの特徴をモデルと対応付ける物である。ペンシルバニア州大のGandhiらは、画像から得られる不確かな特徴点とモデルとの対応を効率良く行う手法を提案し、物体の姿勢推定に適用した結果を示した。

S3. Pattern Recognition/ System 1 (赤松)

本セッションでは、米国におけるCV研究の代表的存在であるMIT, Maryland大学から、顔による個人識別、表情動作のパラメータ抽出、表情の認識という顔画像を対象とした3件の論文が発表され、顔認識という応用課題への米国での関心の高まりが感じられた。

この内MITのPentlandらは、Turkらによる固有顔(Eigenface)を用いた顔の個人識別法の検討の延長として、(1)約3千名の大規模顔画像データベースに対する識別性能、(2)入力顔画像の姿勢変化に対するView-based Eigenspace法による識別のロバスト性の評価、(3)顔の識別において顕著性をもつ目や口等の部分画像別に固有空間を求めることによって顔画像の記述を行なうModular Eigenspace法による識別性能、に関する実験結果を示した。(2)に関しては、顔画像の観測方向に応じた固有顔を留意し、まず部分空間類別法によって入力顔画像の姿勢判定を行なった後、観測方向を特定して個人間の識別を行なう方法を提案し、顔の姿勢の変動に対する識別のロバスト性に関して、NTT村瀬らのパラメトリック固有空間法による識別方法との比較を行なった。これらの結果は、その実験データの規模において従来の研究をはるかに凌駕するものであり、参考資料としての価値は評価できる。しかし、認識手法に関する新しい研究の展開が示されているとは考えられない。

同グループでは、顔表情の動作を表現する記述システムとしてFACS+の提案を発表している。こちらのほうはまだ有効性の評価が十分とは言いが、オプティカルフローによる動き情報と顔の筋肉モデルを組み合わせることで、従来のFACSと比較して、顔の表情動作によって生じる時空間パターンのより精密な分析・生成に、今後威力を発揮するものと期待される。

S4 & S8. Low Level 2&3 (川嶋)

Low-Level 2(Curves and Contours)では非線形の拡散に関するものが2件(NiessenらとKimiaら)、認識における

ベイジアンネットワークの自動生成に関するもの(Sarkarら)と主観的輪郭に関するもの(Williamsら)がそれぞれ1件報告された。Niessenらは各種の非線形拡散方程式が座標不変形式に書き直せるということを示した上で、それを差分方程式(Euler法)に変換するための一般的な枠組みを明らかにした。このなかで彼らは、最終的な差分計算の更新量が、スケールsをもつGauss関数とその偏微分をフーリエ変換したものと画像のフーリエ変換の積の逆変換の簡単な演算に還元されることを示している。Kimiaらは輪郭線に関するcurvature deformationとそれを濃淡画像に拡張したものについて、トポロジーの保存性や、収束形状、極や変曲点の単調性などの性質を明らかにするとともに、curvature deformationと他の線形、非線形の拡散方程式との比較を行っている。

Low-Level 3(Filtering)ではsteerable filterに関するもの2件(GreenspanらとShyら)と、motion推定に関するもの(Jaehne)、特異点の検出に関するもの(Falzonら)が1件づつあった。steerable filterに関する発表はいづれもPeronaらのグループのもので、ひとつは回転に対して不変なテクスチャ認識を行うためのピラミッドカーネルに関するもの、もうひとつは、steerable filterの計算コストを低減するためにカーネル関数をx-yについて変数分離しようという試みについてであった。特異点解析に関するFalzonらの発表ではderivative scale-space(DSS)が提案された。彼らは原波形をn回微分したときに初めて不連続となる特異点の性質をnで特徴づけられることに着目し、このnを実数へ拡張したfractional derivativeの次数をscaleとみなしてDSSを構成する方法を示した。

S5. Motion 1 (志沢)

対応付けと運動推定に関して4件の発表があった。P.Fua & Y.G.Leclerc(SRI)は、正則化手法を取り入れたエネルギー最小化によって、複数の画像からメッシュ表現された表面形状とカメラパラメータを求める手法を開発した。各画像でvisibleな部分を表現するFacet-ID画像を用いて自己オクルージョンの問題も解決した。Z.Wang & A.Jepson(Univ. Toronto)は、3次元剛体の3次元特徴点対応から、3次元並進パラメータと3次元回転パラメータを部分空間法を用いてそれぞれ独立に閉形式で求める方法を提案した。この問題の解は、四元数を用いて回転を先に求めるHornのアルゴリズムが有名であった。提案されたアルゴリズムは並進運動が先に求められる。そのため、並進運動に関する誤差を低く押さえられる。R.Manmatha(U Mass)は、2枚の、画像輝度、点、直線の分布のガウシアンフィルタリングの結果から線形計算で比較的大きな有限アフィン運動を計算する統一的方法を提案した。点、直線の分布に関しては対応は不要である。この方法は、アフィン変換におけるスケールと歪み成分を補正する項が基本方程式に含まれるのが特長である。この補正

項は、ガウス関数が満たす拡散方程式を用いて導かれた。

S11 & S18. Active Vision 1 & 2 (川嶋)

合計8件の発表があった。Kutulakosらはpurposive controlについて、遮蔽輪郭(occluding contour)の検出と表面形状の復元という点から2件の発表をおこなった。最初の遮蔽輪郭の検出に関する論文では、まず物体表面の同一平面上にない4点に対する3Dアフィン不変量が異なる2つの視線方向からの観測によって決定できることを示したうえで、この不変量と観測した特徴の座標値から予測した画像の特徴位置が観測値と一致するかどうかによって、遮蔽輪郭と物体表面上の模様とを識別している。このとき、視線を識別に適した方向にコントロールするための手順を示している。つぎのグローバルな形状の復元のための実時間の視点制御に関する論文では、遮蔽輪郭を構成する物体表面上の曲線(visible rim)が視点の変化をずらすことによって表面上を移動することから、visible rimが可能な限り多くの領域上を通過するようにトポロジーを考慮しながら視点を移動させる方法について報告している。

視覚センサのコントロールに関しては、uncertaintyに基づくものが2件(Whaiteら, Chaumetteら), vergence controlに関するものが1件(Taylorら)あった。このほかhand-eye coordinationのための実時間トラッキングに関するもの(Hager), 定性的障害物検出に関する報告(Zhangら)があった。

趣の異なる研究としてStarkらの"3次元形状と物理的特性に基づく物体の機能の認識"に関する報告があった。この研究では、物体の機能を形状サイズ、部品間の距離、物体表面のなす角度、表面によけいな別の物体が接触していないこと、安定性、包容能力)と、能動的な対象物に対する操作(姿勢の変更と、力の付加、変形量の測定)の両方から推論するGRUFF-TWシステムを開発している。このシステムでは、最初は形に関する情報から、対象物が例えば椅子の機能を持っていると判別し(GRUFF system), さらにその椅子が物体を支えられるだけの強度のある材質でできているかどうかを調べるために「座ることができる面 (sittable surface)」に荷重分布を与えて変形の具合をみて、物理的にも椅子として機能するかどうかを判定している(THING WORLD system)。

S13. Document Processing (黄瀬)

本セッションでは、文書画像のレイアウト解析に関する発表2件、文字認識に関する発表3件があった。

R.Haralickは、レイアウト解析のサーベイ論文を発表した。このサーベイでは、1991年、1993年のICDAR(文書解析と認識の国際会議)からの引用が全体の6割以上を占めることから、最近の手法に重点が置かれている。また、引用された手法の大半が、比較的少数の文書

画像(多いものでも数百ページ)で評価されている点を問題として挙げ、評価用の大規模文書画像データベースの必要性を強調した。その具体例として、ワシントン大で作成された英文文書画像データベース(CD-ROM2枚組:千ページを越える画像、レイアウト解析、文字認識の正解データを含む)を紹介した。

S.Randriamasyらは、種々のレイアウト解析アルゴリズムを比較評価する手法を提案した。従来、この種の評価法はほとんど提案されていないが、近年、様々なアルゴリズムが提案されるに伴って、必要性が高まってきている。提案手法では、各アルゴリズムの領域表現形式に依存せずに処理結果を評価するため、画素レベルの比較法が用いられている。

W.Kimらは、Zernikeモーメントを用い、文字の位置、大きさ、回転に依存しない印刷英字認識法を提案した。また、プリンタで作成した評価用文書や地図などを対象に認識実験を行った。この論文で示された考えは、フリーフォーマットの文書認識を実現する上で重要と考えられるが、マルチフォントに対応していない、文字の接触が考慮されていないなどの問題点もある。

B.Yanikogluら, G.Seniらは、それぞれ、筆記体の書き英単語を対象としたオフライン認識法、オンライン認識法を提案した。前者は、単語画像から文字を切り出した後に認識するものであり、切り出しに筆記者の癖などを表す6種類のパラメータを用いる点に特徴がある。一方、後者は、タブレット入力された単語を文字を切り出さず認識するものであり、単語候補の生成にはTDNN(Time Delay Neural Network)を用いる。得られた単語候補は、単語辞書とのマッチングにより選択されるが、この処理のコスト削減法も併せて提案された。しかし、両者共に制限を設けた部分的な実験結果が示されただけであり、この点が残念である。

S15. Motion 3 (顧)

本セッションでは5件の発表があった。動画像の運動解析に関する四件の発表を紹介する。

阪大の顧らは複数の動物体を含むシーンに対して、MDL(最小記述長原理)に基づく運動解析法を提案した。これは、CVPR'93で発表した二枚の画像を対象とした空間分割の話と異なり、長い時系列画像を対象としたものである。運動変化に対して、時間軸での運動分割と、二枚以上の画像による空間分割の最適化を行って、シーン全体の運動記述の最小化により各物体の運動を最適にモデリングした。

Iraniらは、静止環境におけるカメラの3D運動パラメータを推定する新しい方法を提案した。推定の手順は二段階に分けられている。まず、カメラの運動により表われた画像面での2D運動を消去することで並進成分を求めるとして、並進成分を除いた運動連立方程式で回

転成分を推定する。対応付けせずに、しかもロバストな結果を得られることが印象的だった。

Burlinaらは複数のエージェントを含む動的な環境で無人走行車と各エージェントの距離を多項式でモデリングし、TTC (Time to Collision)とTTS (Time to Synchronization)により定性的な運動記述を求める方法を提案した。

MITのNiyogiらはXYTデータを用いて歩調 (gait)から人物を識別する方法を提案した。カメラの光軸に垂直な方向に歩いている人に対して足の高さでのXT空間に特有な編みパターンが表されている。これをスネックとして検出する。さらに高さ方向に沿って、各XTスライスでのスネックを順番に得る。これにより、各時刻の人体輪郭を求め、歩調の識別を行う。実画像に対して割と高い認識率が得られた点は注目される。

S19. 3D Models (大谷)

5件の論文のうち4件までが、多視点からのレンジイメージを統合して、物体の3次元モデルを得ようとする手法に関連するものであった。

CMUのShumらは、複数視点からのレンジイメージからの3次元モデリングの問題を、欠落データのある場合の主成分分析の問題、即ち、重み付き最小自乗法として定式化している。スイス連邦工科大学のRutishauserらは、任意形状の物体の複数の距離画像を統合するために、localな情報を限定的に用いるアプローチを検討している。まず、複数視点からの距離画像の独立な三角パッチ表現を行う。ここで各ノードは、3次元位置に関する精度の属性をもつ。次に、誤差を軽減するために、重なりをもつ領域の相互近似を行なう。最後に、再度三角パッチ表現を行い、統合する。Laval大学のGagnonらは、多視点からの距離画像を得、重ね合わせ、統合する、という手法を提案している。ポイントは、前述の重ね合わせにおいて、各視点のネットワークを考え、全ての視点における重ね合わせ誤差を最小化することである。INRIAのThirionは、3次元画像の重ね合わせと認識のための信頼度の高い目印として用いることのできる、幾何学的に不変な極値点 (Extremal Points) の抽出について述べている。

一方、ペンシルヴァニア大学のDeCarloらは、2つのパラメータ表現された3次元物体の間を線形的に内挿し、球とドーナツのような極端に異なる形状同士でもブレンド可能な行なう新しいモデルBlended DeformableModelを提案している。

3. ポスターセッション

P4. 3D Vision Posters (大谷)

本ポスターは、移動ロボットのNavigation、距離画像のレンダリング、3次元画像からのゼロ交差点の抽出、

3次元形状表現、時系列の距離画像獲得、というようにそれぞれ異なった論文発表があった。

Cornell大学のHuttenlocherらは、移動ロボットのNavigationを視覚情報を用いて行うために、2次元のエッジ画像を比較する手法を提案している。ある目印を常に画像の中心にもってくるようにロボットを制御する。3次元復元が不要であることを強調しているものの、それほど斬新なアイデアとは言えない。

ワシントン州立大学のFlynnは、モデルベースの物体認識システムにおける仮説の生成と検定の改良を検討した。即ち、距離画像中の物体の姿勢の検定のため、物体モデルをもとの距離画像に投影し、比較して検定する際、距離画像に陰影付けを行い、検定処理性能の向上を目指している。

INRIAのMongaらは既に、3次元Volumic imageにおける表面形状の微分的特性が偏微分により復元できることを示しているが、ここでは偏微分の計算のマルチスケールによる改良、尾根線の抽出への応用等について述べている。

INRIAのDelingetteは、Simplex Mesh (三角パッチ表現において、隣接した3個の三角パッチ内部の点をトポロジカルに連結するメッシュ)による3次元形状復元について述べている。物体の弾性的な挙動や、簡単な複数の形状を連結して、複雑な形状を復元できることを示している。

ATRのOhyaらは、熱画像と可視光画像のステレオペアを統合的に用いて、時系列距離画像を獲得する手法の検討を進め、連続する時刻における相関を用いてステレオマッチングの精度向上を目指している。

P6. Motion Posters (加藤)

18件のポスターの内、3次元情報の復元に関するもの8件、Motion解析に関するもの5件、Motionからの画像セグメントに関するもの3件、カメラキャリブレーションに関するもの1件、ロボットの姿勢推定に関するもの1件であった。

Stanford大のTomasiは、Tomasi&Kanadeの提案した、Factorization法に対し、2Step法を採用する手法を提案している。2Step法はカメラキャリブレーションにおいては古典的計算手法であるが、この採用でカメラの動きと対象物の3次元構造が透視投影画像からも取得可能でありさらにノイズに強いことを示している。ただし、カメラキャリブレーションでは、既知の3次元構造、あるいはカメラモーションから、カメラパラメータを決定するのに対し、提案手法では、その逆問題を解く形になっている。よって前提条件として、高精度なカメラキャリブレーションが必要である。

阪大のChenらは、モーションベースで画像セグメントを行う際に、ローカルなエッジポイントの動きを観測

し、結合性の強いものをグループ化していくことで、グローバルなセグメントを行う手法の提案をしている。実験では歩行者の腿と膝下の部分が明確にグループ分けできることを示した。

NTTのKatoらは、超広角単眼視によって撮影した画像から形状および表面テクスチャを取得する手法を提案している。実験により、走行車両の車体形状および表面特徴の取得が可能であることを示した。

4. ワークショップ

W1. Workshop on Visual Behaviors (浅田)

Visual BehaviorのWS(6月19日(日)開催)主催者はVirginia大のMartinで、趣旨は、タスクを遂行するためにbehaviorと知覚が密に結合したシステムに内在するkey issueを討論することであった。但し、発表論文が多過ぎて、全体を通しての方向性や課題などを討論する時間がなかったこと、また、ほとんどの論文がbehaviorが知覚自身に閉じており、ロボットのbehaviorを考えている人が少なかった。

Rochester大のBallardらは、人間のハンドアイシステムにおいて、実行時に利用されるメモリ容量が、眼球運動の助けをかりて、少なくともよいことの実験結果を提示した。メモリの少なさには賛同できるが、機械にとって、どのような情報が本質的かが議論されるべきであろう。シカゴ大のSwainらは、雑貨店などの商品展示の規則性を事前知識として与え、その規則性を利用したvisual routineをシミュレータの世界で構築した。ペン大のBajcsyらのグループは、球面射影におけるフロー解析を、移動注視者の仮定で行なったが、Aloimonosらの研究との類似性などの質問があり、新規性に問題があった。MIT(Media Lab)のBobickは、ダンサーの運動記述を例に、ダイナミックシーンを記述する有効な表現の枠組、座標系などを考察した。同じくMIT(Media Lab)のPentlandのグループは、実世界とシミュレータを結びつけるALIVE(Artificial Life Interactive Video Env.)を発表した。人のシルエットから意図を解析し、それに呼応する人工生命(ハムスター)をビデオプロジェクターの中に、人と写し出すものである。この後開催されたA-Life IVでは、Genetic Programmingを用いて、人間の手の部分を検出する手法を発表していた。実システムとして動いていたものは、IRISA/INRIA(仏)のChaumetteによる画像特徴量によるビジュアルサーボ、Asadaら(阪大)による、視覚に基づく行動学習によるサッカーロボットの実演、MIT(AI Lab)のHorswillによる廉価なアクティブビジョンシステムなどがあった。特に、Asadaの発表は、強化学習を用いて視覚情報から直接行動への写像を行なうもので、これまでの計測や再構成主体のアプローチとは異なり、新しい方向性を示唆しているものとみなせる。

W3. Workshop on the Role of Functionality in Object

Recognition(黄瀬)

本ワークショップは、物体を認識する上で、物体の持つ機能に着目することが、どのような利点をもたらすかを議論する目的で開催され、CV, Robotics; AI, 認知科学などの分野から、合計11件の発表があった。以下では、このうち5件の概要について述べる。

K.Bowyer, L.Starkらは、ジョイントを持つ物体の認識法を提案した。この手法は、物体の可能な動きを表す一連の3D形状データから、まず物体の部品形状、部品間のジョイントの位置などを推定し、その後、機能に基づいて物体を分類するものである。認識実験には様々な形のハサミを用いた。K.Kise, T.Kitahashiらは、テコの機能と作用対象により道具をモデル化し、認識する手法を提案した。また、種々の栓抜き、スパナ、ドライバに対する実験結果を示した。以上の2手法は、単一の形状モデルではカバーできない程、様々な形を持つ物体を認識可能であるが、入力に完全な3D形状データを必要とするという問題が残されている。

E.Rivlin, S.J.Dickinson, A.Rosenfeldは、"Recognition by parts"の考えを拡張した"Recognition by functional parts"に基づく認識法を提案した。彼らの手法は、部品の形状と機能を一対一対応に制限した点に問題があるものの、superquadricsの利用等により実画像を対象とした点が評価される。

L.Bogoni, R.Bajcsyは、穴空け(piercing)タスクに適した道具の機能的な特徴を、能動的に調べる手法を提案した。この手法では、ロボットに付けられた力センサーとカメラ、腕の位置情報を統合することにより、タスクが成功したかどうかを判断する。E.Krotkovは、ロボットにより物体の材質を知覚する手法を提案した。知覚には、叩いて音を聞く、突いて動きを見る、与えた力と変形の関係を測る、などの方法が用いられる。以上の2手法は、物体を認識するものではないが、物体の機能を推定する上で重要な特徴を、認識者(ロボット)が能動的に知覚するという点が興味深い。

最後に行われたディスカッションの際、T.Binfordは、「物体認識において、物体の持つ機能に着目する重要性は、かなり以前から指摘されており、問題としては古い。」と述べた。ところが、機能の捉え方や記述法、利用法が研究者により様々であり、有効性も限定されていることから、依然として未解決な問題とも言える。今後の発展を期待したい。なお、本ワークショップでは、会場で一部の論文のコピーが配られたが、論文集は発行されていない。詳しい内容に興味のある方は、kise@ss.cs.osakafu-u.ac.jpまで照会されたい。

W4. Workshop on Visualization and Machine Vision

(杉本)

Workshop on Visualization and Machine Visionは、IEEE PAMI および CG の Technical Committeeの後援により、1994年6月24日に、Seattle WestinHotelにて開催された。十件の口頭発表と、四人のパネラー L.Hesselink (Stanford University), T.S.Newman(National Institute of Health), I.Carlbom(Digital Equipment Co.), B.Kamgar-Parsi (Naval Research Lab) が各一件づつ話題を提供するという形のパネル討論が行なわれた。また、一般講演に関する査読は、匿名の fullpaper による。米国内のみならず、ヨーロッパからの発表もあり、日本からは、K.Sugimoto (RWCP)らによる、境界線のセグメンテーションに関するものを含め2件の発表が行なわれた。

Visualizationは、近年急速に成長した分野といえる。その目的は、映像化や画像を用いた対話的な操作を通じて、大量のデータから有意義な情報を如何に効率良く抽出するかにあるといえる。画像を扱い、また膨大な計算量を要するという点で、コンピュータビジョンと共通する問題を抱えている。本Workshopは、両分野の接点を明らかにすると共に、新たな方向性と応用分野を確認する狙いから、今回初めての開催を迎えることとなった。

可視化にコンピュータビジョンを応用した例として、選択的な可視化があげられる。この場合、データ中の処理対象とすべき領域のみを抽出するために、画像解析が画像表示に先立って行なわれる。関連する発表として、Z.Zhu(Univ.Mississippi State)らによる、流体に関する観測データから、渦領域のみを特徴抽出し可視化する技術に関するものがあつた。また、E.Mesrobian(UCLA)らによる、膨大な気象観測データから時空間パターンを如何に効率良く抽出するかに関する発表も、注目を集めていた。

コンピュータビジョンに可視化の技術を利用した例として、画像の多重表示や、S.B.Marapane (Univ.Tennessee)らによる、ロボットや撮像装置の動作や動作環境のシミュレーションに関する発表があつた。これらを含めた多くの場合、解析やデバッグ時に、データを可視化するためのツールとして利用されている。

W5. Workshop on Biomedical Image Analysis (佐藤)

最近、欧米のCVやRobotics研究者の間で、にわかに医学応用(Medical CV, Medical Robotics)が注目され始め、新たなコミュニティを形成しつつある(この原因は、冷戦構造の崩壊に伴う軍予算削減の影響が少なからずあるようだ)。従来から、臨床と密接に結びついた形で、この分野の研究が行われてきたが、発表の場(会議や雑誌)がバラバラであつた。本ワークショップの目的は、このコミュニティに共通の発表・討論の場を持つことである。来年には、さらに、国際会議(CVRMed, Conference on

Computer Vision, Virtual Reality, and Robotics in Medicine)が開催される予定であり、CV研究のBig Nameもプログラム委員に名を連ねている。

発表論文に関しては、deformable modelや3次元非剛体の位置合わせに関するものが目を引いた。deformable modelは、単なる流行とみられる向きもあるが、心臓などの非剛体物体の解析に関して、臨床に近い研究者の関心も大きいようである。その中でも、A.A.Young, D.L.Kraitchman, and L.Axel (Penn大,放射線科)の発表が注目されていた(L.Axelは放射線科医の間で著名)。この発表は、tagged MRIという心筋に格子上の印(tag)をつけて撮像したMRIを用い、tagをdeformable modelの変形版で追跡して、心筋の動きを求めめるものである。tagは、心筋に直接つけられた印であるので、心臓壁だけでなく心筋内部の動きもわかる点が大きな利点である。この発表は、CVの技法とtagged MRIという撮像技法をうまく組合せ、臨床的に有効な心臓の動き計測法を開発した点で評価できる(なお、tagged MRIを用いた動き計測は、阪大や三重大でも試みられている)。CVの医学応用が、1つのコミュニティとして認知される鍵は、単に、医用データをCVの手法に適用するだけでなく、臨床における有用性をいかにアピールするかにあると思われる。

また、特筆すべきことは、鳥脇(名大)が招待講演を行ったことである。欧米の研究者も、鳥脇グループの成果には大きな関心を持っている。その他、日本からは、清水(名大)、長谷川(中京大)、仁木(徳島大)らの参加があつた。

5. むすび

本報告は、CVPR94の概要を有志によりまとめたものである。コンピュータビジョンおよびパターン認識に関する国内の技術レベルは近年随分向上してきた。しかし、まだまだ米国の流行に追随している部分も多い。国内に、既存の枠組みにとられない発想を自由に出せる雰囲気、又それを育てられる風土を作り上げ、日本独自の新しい研究の展開を目指したいものである。最後に、本報告のきっかけ及び機会を作っていただいた岡山大学の松山教授に感謝する。(以上、村瀬)

参考文献

[1] Proceedings of Computer Vision and Pattern Recognition, IEEE Computer Society Press, Seattle, June, 1994