

第9回コンピュータビジョン欧州会議 (ECCV2006) 報告

石川 博 (名古屋市立大), 金谷 健一 (岡山大), 金澤 靖 (豊橋技科大), 牧 淳人 (京都大),
 榎原 靖 (大阪大), 三浦 純 (大阪大), 佐藤 智和 (奈良先端科技大),
 菅谷 保之 (豊橋技科大), 八木 康史 (大阪大)

hi@nsc.nagoya-cu.ac.jp, kanatani@suri.it.okayama-u.ac.jp, kanazawa@tutkie.tut.ac.jp,
 maki@i.kyoto-u.ac.jp, makihara@am.sanken.osaka-u.ac.jp, jun@mech.eng.osaka-u.ac.jp,
 tomoka-s@is.naist.jp, sugaya@iim.ics.tut.ac.jp, yagi@am.sanken.osaka-u.ac.jp

2006年6月8-11日にオーストリア(グラーツ)で開催された第9回コンピュータビジョン欧州会議 (ECCV2006) の概要を報告する。

Report on the 9th European Conference on Computer Vision (ECCV2006)

Hiroshi Ishikawa (Nagoya City University), Kenichi Kanatani (Okayama University),
 Yasushi Kanazawa (Toyohashi University of Technology), Atsuto Maki (Kyoto
 University), Yasushi Makihara (Osaka University), Jun Miura (Osaka University),
 Tomokazu Sato (NAIST), Yasuyuki Sugaya (Toyohashi University of Technology),
 Yasushi Yagi (Osaka University)

This paper reports an overview of the 9th European Conference on Computer Vision (ECCV2006) held in Graz, Austria, from 8-11, May 2006.

1. まえがき

2006年6月8-11日にオーストリアのグラーツで第9回コンピュータビジョン欧州会議 (ECCV2006) が開催された。グラーツはオーストリアでウィーンに次ぐ2番目の都市であり、会議を主催したグラーツ工科大学 (創立1811年) はオーストリアの最も代表的な理工系大学である。ドップラー効果で知られる物理数学者 Christian Doppler (1803-1853) はこの大学の出身者である。

会議には約600人の参加者があった。会場のグラーツ会議堂は音楽の国オーストリアにふさわしく、元々は音楽演奏のために建てられたもので、音響効果が非常に優れているということである。重厚かつ華麗な内装も素晴らしいものであった。

さて、ECCVはもともとは西ヨーロッパの地域的な集会として1990年にスタートしたが、現在ではコンピュータビジョン国際会議 (ICCV) に匹敵するコンピュータビジョンに関する最も中心的な国際会議の一つになっている。

しかし、この成立の由来を反映してか、伝統的に

1. 数理的, 幾何学的な発表が多い,
2. 日本からの発表が著しく少ない,

と思われるようである。第1点は、過去に英国のOxford大学, フランスのINRIA, スウェーデンのLund大学などでコンピュータビジョンの数理的, 幾何学的な研究が精力的に行なわれたという歴史的な背景のためと思われる。しかし、今日では世界中から論文が投稿され, ICCV や米国のCVPR 比べてもそれほどの差異は見られなくなっている。現在では米国やヨーロッパではICCV, ECCV, CVPR をかけもちで発表する研究者が多く, ほぼ“平滑化”されているようである。

表1はECCVのWebに掲載されていた今回の論文のテーマ別の投稿数, 採択数, および採択率である (投稿数順にソート)。ECCVの伝統的なテーマといえる3次元復元, および多視点幾何学は合計では161件で, 確かに少なくはないが, やはり多いのは物体認識・検索, 領域分割・併合, 運動追跡, 画像特徴抽出, 顔・ジェスチャ関係であり, 他の国際会議でも同様の今日的なテーマである。

なお, どのテーマも採択率はほぼ等しく, 全体の24%に近い値であるが, なぜかロボットビジョンの採択率が著しく低い。やはり実践的なシステム開発は何らかの新規的な原理が提示されないと採択され

表 1: テーマ別の投稿数, 採択数, 採択率.

| テーマ | 投稿 | 採択 | 採択率 |
|-----------|-----|-----|------|
| 物体認識・検索 | 190 | 45 | 0.23 |
| 領域分割・併合 | 179 | 51 | 0.28 |
| 運動追跡 | 177 | 45 | 0.25 |
| 画像特徴抽出 | 156 | 38 | 0.24 |
| 顔・ジェスチャ | 112 | 26 | 0.23 |
| マッチング・フロー | 108 | 30 | 0.27 |
| 統計的学習 | 105 | 61 | 0.29 |
| 形状表現・モデル | 101 | 29 | 0.28 |
| 3次元復元 | 91 | 28 | 0.30 |
| 多視点幾何学 | 70 | 22 | 0.25 |
| 色・テクスチャ | 69 | 16 | 0.24 |
| 医用画像解析 | 64 | 11 | 0.17 |
| ロボットビジョン | 40 | 2 | 0.05 |
| 画像取得・センサ | 38 | 10 | 0.26 |
| 認知・生体ビジョン | 34 | 6 | 0.17 |
| 表面反射モデル | 32 | 10 | 0.31 |
| 全体 | 811 | 192 | 0.24 |

ないのかもしれない。これに次いで医用画像解析や認知・生体ビジョン関係も採択率が低い。それらは他にその専門の国際会議があって、この会議の中心的なテーマでないからであろう。

第2点の日本からの発表の少なさの原因は、日本の応用システム研究者が「ECCV は理論」と思い込んでいる先入観のためと思われる。過去の会議でも日本からは毎回1, 2件の発表しかなく、今回も4件 [27, 29, 31, 45] である。これは日本からの発表数が毎回著しく多いパタン認識国際会議 (ICPR) と対照的である。

表2に国別の論文数を示す。米, 英, 仏, 独が多いのは当然で, イスラエル, カナダ, オーストラリアも常連であるが, アジアからは中国が突出して2位である。この結果を見る限り, コンピュータビジョン研究のアジアの中心拠点は中国であるという印象を与える。表1に示すように, ECCV は決して理論中心ではないので (ただし, 原理的な新規性は要求される), 今後は日本からも世界的に注目されるこの会議により多くの発表があることが望まれる。

今回の Longuet-Higgins 最優秀論文賞には Levin (Hebrew 大) ら [39] が, 優秀論文賞には Hasinoff (Toronto 大) ら [22] と Ait-Aider (CNRS) ら [4] が選ばれた。なお, 次回の ECCV は2008年にフランスのマルセイユで, その次は2010年にギリシャのクレタ島で開催されることが決まった。

以下, テーマ別に代表的な論文を取り上げて報告する。(以上, 金谷)

表 2: 国 (= 第1著者の所属機関の所在地) ごとの論文数.

| | |
|------------|----|
| 米国 | 69 |
| 中国 (香港を含む) | 20 |
| 英国 | 19 |
| フランス | 18 |
| ドイツ | 10 |
| イスラエル | 9 |
| カナダ | 7 |
| オーストラリア | 5 |
| イタリア | 4 |
| 日本 | 4 |
| スイス | 4 |
| オーストリア | 3 |
| インド | 3 |
| ベルギー | 3 |
| 台湾 | 3 |
| オランダ | 3 |
| スウェーデン | 2 |
| 韓国 | 1 |
| メキシコ | 1 |
| 南アフリカ | 1 |
| スペイン | 1 |
| ポルトガル | 1 |
| ロシア | 1 |

2. エネルギー最小化関係

今回の ECCV06 で一番目を引いたのは, 「エネルギー最小化」というセッションが設けられていたことである。以前からステレオやセグメンテーションなどの低レベルビジョン問題が「エネルギー最小化」として定式化されてきたが, それをグラフアルゴリズムとして解く手法, 手法間の比較, および実際の応用に適用する論文が目だった。

そのようなグラフ蟻後リズムとして, グラフカット (GC) が標準的な技術の1つになってきている。またより最近では, メッセージ伝播による, 信念伝播法 (BP) やツリー重み再配分メッセージ伝達法 (TRW) なども注目されている。今回の ECCV では, これらの手法を比較する論文が2つ発表されている。

コンピュータビジョンでこれらの手法を使うときのグラフは普通, 例えばピクセルごとに頂点を定義し, 四方の隣接ピクセルのみと辺によって接続するなど, 可能な頂点間の接続数 (~ 頂点数の2乗) に比べてごく少数の辺しか持たない。そのような場合には TRW が最も性能が高いという複数の結果が出ている。ところが, 今回 Kolmogorov (Microsoft 社) ら [37] が, 例えば隠蔽をモデル化したステレオなど, 普通よりやや辺の数の多い, すなわち次数が高いグラフに GC, BP, TRW の3手法を適用して実験的

に比較したところ，この場合には最小化したエネルギーにおいても，また真の解と比較した誤差においても，GC が他の 2 つよりかなり性能がよいという結論が出ている．

一方，Szeliski (Microsoft 社) ら [58] はエネルギー最小化アルゴリズムのためのベンチマークを提案し，やはり GC，BP，TRW をステレオとセグメンテーションについて実験的に比較している．全般にこれらの方法は，比較に使われた従来法 (反復条件付きモデル: ICM) に比べて非常に改善されており，各ベンチマークについて，(最良の方法はベンチマーク毎に変わるものの) 誤差 1 パーセント以下の結果を出している．

グラフカットのバリエーションもまだ出尽してはいないようである．Kohli (Oxford Brookes 大) ら [36] は，グラフの変化が小さいときにグラフカットを効率的に計算する，dynamic graph cut を，ICCV2005 に続いて提案している．ICCV では動画の連続フレーム間の変化が小さいことから，動画セグメンテーションを効率化するという話だったが，今回は MRF の各頂点でのラベル選択の confidence measure を効率的に求めている．そのために，各頂点で選択可能なラベルを一つ固定したときのグローバルな最大確率 maximum marginal を求め，それをこの頂点で他のラベルを選択したときと比較して，特定の選択に他と比べてどれだけ自信があるかの指標としている．これらの最大確率を求めるときには，その頂点の周辺だけが異なるグラフを使うので，dynamic graph cut により非常に効率的に計算することができる．

また，同じグループの Bray (Oxford Brookes 大) ら [10] は，人物の 3 次元姿勢モデルと画像のセグメンテーションを 1 つの MRF でモデル化し，dynamic graph cut によって効率的に最大確率状態を求める人物動画のセグメンテーション法を提案し，pose cut と呼んでいる．

グラフカットの使用例としては，この他に，テレコンファレンスのような 1 つのカメラで撮った背景の安定した人物動画の前景を分離するのにグラフカットを使った，Sun (Microsoft 社) ら [57] の background cut もあった．

(以上，石川)

3. 画像特徴抽出・マッチング関係

ステレオなどで曲面の事前分布をモデル化する場合，コンピュータービジョンではほとんど常に近隣画

素間の高さの変化の何らかの増加関数を最小化する平滑化事前確率分布モデルが使われるが，Ishikawa (Nagoya City 大) ら [27] は，人間の視覚系がそのような平滑化モデルは使っていないばかりか，既存のすべての事前分布モデルとある意味において正反対のモデルを使っていることを示す心理物理学の実験結果を紹介している (国内資料 [25]) ．

Rosten (Cambridge 大) ら [51] は実時間で動作するアプリケーションのための特徴点抽出手法を提案している．彼らは Harris を始めとして Shi, Tomasi らの画像の 2 次微分に基づく方法についてまとめ，一方で画像の 2 次微分に依らない周辺画素との輝度変化を単純に調べる高速な特徴点抽出アルゴリズムを提案している．さらにさまざまな手法との比較実験を行い，動作の高速性の検証だけではなく，シーンを異なる視点から撮影した場合にも同一の特徴点を抽出する性能が高いことも検証している．特徴点抽出は，2 画像間の対応付けや追跡などに用いられることが多く，このような性能を持つ特徴点抽出手法は非常に有用であると思われる．

同様に，Bay (ETH Zurich) ら [9] もまた高速な特徴点抽出手法を提案している．Rosten ら [51] らに対して彼らは画像の 2 次微分に基づく手法を基本に，画像の 2 次微分を単純な形のフィルタに近似することによる高速化手法を提案している．提案手法はスケール変化や回転に不変である．さらに特徴点の抽出だけではなく，特徴点の性質を表す記述子も示している．似た特徴点の抽出・記述子には Lowe [42] の SIFT と呼ばれる手法があり，提案手法はこの手法の影響を受けている．Rosten ら [51] や Bay ら [9] は同様に提案手法のプログラムをインターネット上で公開している．このように研究成果を他の研究者が使用できるように公開して頂けると大変ありがたい．

Favaro (Simens 社) ら [17] は，Defocus Inpainting と呼ばれる静止画像上の焦点ぼけを 1 枚の入力画像のみで解消する手法を提案している．本研究では，画像上に類似する物体が多数撮影されており，それらの一部が被写界深度内に存在するものと仮定した上で，合焦している領域の情報を用いてその他の領域の焦点ぼけを修復する．具体的には，まず画像内からガウシアンによるぼけを考慮した類似パターンの探索を行ない，類似パターン間でのぼけパラメータの変化を算出する．次に，類似パターン全体でぼけ具合に関するエネルギー最小化を行うことで，画像を修正する．

Kaziska (Florida State 大) ら [32] は、歩容シルエットの輪郭を用いた個人認証手法を提案している。最初に、ある二つの輪郭に対して、一方の輪郭上の各点からもう一方の輪郭上への最短距離を考え、その総和が最小になるように位置合わせを行っている。その最短距離の総和を歩容シルエットの類似尺度として、入力と辞書データの照合をはじめとして、周期の抽出や位相の位置合わせ等を行っている。また、異なる周期のシーケンスを正規化するために、最短距離の対応点を利用した任意位相に対する輪郭の補間生成を行っている。

Alter (École Normale Supérieure de Cachan) ら [5] は、暗い照明条件下におけるノイズを考慮した明るさ類似度を提案している。暗い照明条件下では、観測される明るさがポワソン分布に従うことが知られており、著者らはそれに量子化の影響を加えたノイズモデルを導入している。さらに、そのノイズモデルの下で、ある二つの明るさが観測される同時確率を定式化し、その最大対数尤度を明るさの類似度として定義している。実験では、従来のユークリッド距離による類似度を比較対象としてブロックマッチングを行い、ブロックサイズが小さくシャッタースピードの速い(暗い)状況において、大きな性能改善が見られることを報告している。

Platel (Technische Universiteit Eindhoven) ら [49] は、画像の対応付けのための特徴点としてスケールスペースとカタストロフ理論に基づく特異点を用いることを提案している。このような特徴点は濃淡に関するスケールリングやオフセットに対して不変であり、ノイズに対する安定性が正確に予測可能であることを示している。さらに対応付けのための特徴量として3次までの勾配で正規化した微係数を用い、マハラノビス距離による類似性尺度を用いることを提案している。Lowe [42] のSIFTと比較し、特徴点抽出の再現性および相似変換における画像の対応づけにおいては、よい結果が得られたものの、透視的な変化が大きい場合には、SIFTの方がよい結果が得られている。

(以上、石川、菅谷、佐藤、八木、槇原、金澤)

4. 物体認識・検索関係

Shotton (Cambridge 大) ら [54] は、1枚の画像に対して領域分割と物体認識を同時に行う手法を提案している。この手法では、領域分割とラベリング済みの事前画像データベースを用いて conditional random

field (CRF) モデルによる学習を行い、次に学習データを用いて入力画像の領域分割とラベリングを同時に行う。従来手法では、領域分割とラベリングは個々に行われていたが、提案手法では領域分割と認識を同時に行うことで認識精度を向上させている。また、本手法では効率的な学習を行うために texton map を用いる特徴抽出手法についても新たに提案している。なお発表では、著者の Web ページからソースコードが利用可能であることが報告された。

Crandall (Cornell 大) ら [15] は、同じく画像からの物体認識問題において、事前画像データベースからの学習時に手動による物体領域の指定を必要としない手法 (weakly supervised method) を提案している。本手法では、物体を部分領域を表すパッチの集合とし、部分モデルの見え方と位置関係 (接続関係) を無向グラフによって表現することで、位置関係と見え方を同時に考慮した学習を実現する。実験では、一般的な手動による分類を伴う学習手法に対して、手動による分類を行わない提案手法の方が優れた認識精度が得られる場合があることを示している。本発表についても、著者の Web ページからソースコードを利用できるとのことである。

A. Agrawal (GRAVIR-INRIA) ら [2] は、hyperfeature と呼ばれる階層的な特徴量の記述手法を提案し、画像分類・物体認識への応用例を示している。この手法では、まず空間的な局所特徴量を例えば SIFT 記述子を用いて記述し、特徴ベクトル群をベクトル量子化する。次に、生成されたコードから局所ヒストグラムを生成し、上位階層の特徴量とする。これを繰り返せば、上位層では画像全体の特徴が要約され、下位層では局所的な特長が記述されるため、物体認識において階層的な意味づけ・分類が可能となる。なお、マルチスケールの画像ピラミッドに対して hyperfeature の検出を行うことでスケールの変化にも対応できるとしている。

Avidan (Mitsubishi 社) ら [8] は、サーバ・クライアント型の画像処理システムの利用シーンにおけるプライバシー保護を目的として、画像そのものをサーバに送信することなく顔画像検出や OCR などの各種画像解析処理を実現する手法を提案している。本手法は、暗号プロトコルである secure multiparty protocol を画像解析に応用することで、画像そのものを伝送することなく画像処理を実現する。この手法を用いれば、サーバでは何に対して画像処理を行っているかを知ることはできず、また逆にクライアン

トにおいてはサーバで実行されるアルゴリズムを知ることにはできない．現時点では1枚の画像内からの顔検出処理に30分程度かかるなど実用上問題が残っているが、街角に監視カメラが増えつつある現在の状況に対してマッチした手法であり潜在的な需要は高そうである．

Ferrari (ETH Zurich) ら [18] は、実環境を撮影した静止画像から、事前に手書きで作成されたテンプレート形状を探索する手法を提案している．本手法では、まず画像内から直線的な輪郭の候補を多数検出し、輪郭の接続状態を表すネットワークを構築する．次に、構築したネットワークからテンプレート形状を最大限再構築できる箇所を探索することで、オブジェクトの位置を推定する．このような輪郭ネットワークによる探索では、物体の輪郭がテンプレート形状に完全に一致する必要がないため、様々な入力に対して比較的柔軟に対応できるとしている．

(以上、佐藤)

5. 領域分割・併合関係

Levin (Hebrew 大) ら [39] は領域分割の学習について取り組み、Longest-Higgins 最優秀論文賞に選ばれた．領域分割問題は古くから扱われている問題であるが、最近の流れではボトムアップ的なアプローチとトップダウン的なアプローチを組合せ、さらに効果的な学習アルゴリズムを用いるようになってきている．Levin らはボトムアップのアプローチに対する学習とトップダウンのアプローチに対する学習を同時に行う手法を提案している．具体的にはCRFと呼ばれる条件付き確率場を利用してエネルギー関数を最小にするパラメータを効率的に求める問題に定式化している．これによりトップダウンのアプローチでは学習に多くのデータが必要なのに対して、少数のデータでも効率的に学習できることを実験結果からも示している．また、CRFの計算をモンテカルロサンプリング等で近似することにより計算コストの効率化についても取り組んでいる．

Foulonneau (LCPC) ら [19] は、物体の輪郭抽出問題について、事前知識として物体形状のシルエットを複数与えることで、物体の輪郭抽出時のオクルージョン問題を回避する手法を提案している．本手法では、動的輪郭法におけるエネルギー関数の正則化項として、データベースに存在する輪郭データとの一致度に関する指標を追加している．この正則化項は、シルエットに対する二次元アフィン変換に対し

て独立となるように設計されているため、物体の変形にも対応可能である．なお発表では、著者のWebページにさらに詳しいテクニカルレポートがありそちらも参照してほしいとの宣伝があった．

Ren (UC Berkley) ら [50] は、オクルージョンを含む自然景観を撮影した静止画像から、画像に含まれる各領域の前後関係を自動で推定する手法を提案している．本手法では、まずローカルな輪郭の接続関係から各局所領域の前後関係を確率的に表し、次にconditional random fieldを用いて画像全体の大域的な情報から個々の領域の前後関係を判定する．なお、手法前段のローカルな認識において前後関係判定の正解率は64%程度であるが、全体での最適化を行うことで、正解率が78%程度に向上することが報告された．

Kanatani (Okayama 大) [29] は楕円弧をなす点列に楕円を当てはめる最尤推定の誤差を2次の項まで詳細に解析し、得られる2次の偏差項を差し引くことによって最尤推定よりも高精度(超精度)の解が得られることを示している(国内資料[65])．

(以上、菅谷、佐藤、金谷)

6. 運動追跡関係

Kim (IPIX) ら [35] は、複数のカメラで複数人物を同時追跡する問題を扱っている．各人物の地面上の位置についての推定が得られれば、それらから各画像上での各人物位置の事前分布が得られる．さらに、各人物の見えモデルを用いた尤度を用いて各画素が各人物に属する事後確率を計算して、各人物領域を抽出する．そして、各画面で得られた人物領域へのベクトルを地面上に投影し、人物位置を再推定する．この分割と再推定の繰り返しをparticle filterで実装することにより効率的かつロバストな追跡を実現している．

Wang (MIT) ら [63] は、blob追跡によって得た軌道データの集合を解析して、画像内の物体の動きに関するモデルを生成している．軌道の類似度として、対応点間の距離の近さを信頼度として重みづけした特徴ベクトル(物体の大きさや速度など)間距離の総和を用い、軌道集合のクラスタリングを行っている．道路シーンに対して提案手法を適用し、生成された各クラスタから速度情報をもとに車線や歩道といった解釈を導くとともに、どのクラスタからも外れた動きを検出することによる異常検知への応用についても述べている．

Micilotta (Surrey 大) [47] は、単眼視からの人間の上半身の 3 次元姿勢推定を行っている。顔や手などの部品の検出器は複雑背景下では多数の候補領域を出力するため、それらの適切な組み合わせの候補を部品間の位置や大きさの制約を利用した RANSAC により求め、さらに多数のデータから学習した部品の画像内の配置に関する事前分布と各部品の確からしさを考慮して、最適な組み合わせを選択する。次に、人間がさまざまな姿勢をとったときの見えを CG により作成しておき、シルエットとエッジパターンの比較により、最も適合する見えを選択し、3 次元姿勢を推定する。複雑環境下で比較的高速な推定を実現している。

Özuyusal (EPFL) ら [48] は、ゆっくりと運動する対象物体の観察によりモデルを学習し、それを用いてフレームごとに物体を検出・追跡する方法を提案している。学習時には特徴点を中心とする小領域を取り出し、randomized trees を識別器として用いてフレーム間での小領域の対応付けを行うとともに、特徴点が新たな得られたり見えなくなったときには、木の修正をオンラインで行う。学習が終了すると、得られた対応づけから各特徴点の位置を計算する。学習には、対象に関する事前の情報は不要である。追跡時には入力画像ごとに特徴点の対応づけを行い、位置を推定する。運動の連続性をしないので急激な運動変化にも対応できる。

Sullivan (KTH) ら [56] は、サッカーなどのスポーツ映像解析において、人物の交差が頻繁に起こる状況での各人物の対応付け問題を扱っている。まず、背景差分法により人物領域を抽出・追跡する。単独の人物が移動している部分 (player track, PT) と複数人物が交差している部分 (multiple player track, MPT) とに分け、さらにそれらの接続関係を解析しグラフを生成する。単純な交差の場合には、運動の連続性や見えの類似性に基づき MPT 前後の PT 間の対応づけを行う。さらに、同チームの他の選手との位置関係、およびある人物は各時点で 1 か所にしか存在できないという拘束のもとで、すべての PT を 11 のグループに分割することにより、全フレームにおける PT 間の対応付けを決定する。

Cai (British Columbia 大) ら [12] は、mean shift (MS) と boosted particle filter (BPF) を統合した、多数人物の追跡を扱っている。各フレームで検出された人物領域と追跡対象との対応づけは、対応するもの同士の距離の合計を最小化するような組み合わせを

求めることによって決定する。追跡対象ごとに BPF を構成し、追跡を行う。BPF では各パーティクルを運動予測と検出された人物領域の位置の両方を考慮して移動させるが、ここではそこからさらに MS を用いて移動させる。その後、各パーティクルの重みを観測の尤度に従って計算し、重みに基づいたリサンプリングを行う。MS による移動は、より可能性の高い領域にパーティクルを集中させる効果があり、追跡のロバスト性を高めることができる。

Jiang (Duke 大) ら [28] は、level set 法 (LS) と particle filter (PF) を統合した物体追跡法を提案している。LS では物体の境界は、時空間の補助関数 $\phi(x, t)$ の $\phi = 0$ を満す境界として表現される。 ϕ の分布を PF で表現する場合、各パーティクルの表す関数は多くの位置で似た値をとると考え、平均的な形状を表す母パーティクル ϕ_0 とそこから少し変化した形状を表す子パーティクルの集合 $\phi_p (p = 1 \sim P)$ で表す。母形状からの変化分を表す $\Delta\phi_p = \phi_p - \phi_0$ を取り出し、そのうちのいくつかを母形状に適用することを考えれば、 $P+1$ 個のパーティクルにより、潜在的に 2^P 個のパーティクルを表現することができる。多様性と効率を両立することができる。 2^P 個のパーティクルがあるとしたときに、実際の $P+1$ 個のパーティクル集合を移動、重み付け、リサンプリングする方法について述べている。新たな領域が出現した場合には、その周囲に小さな形状要素をランダムに生成し、それを母パーティクルに加えることにより対処する。

Shet (Maryland 大) ら [53] は、複数カメラによる監視において、異なる時間/画像で検出された人物が同一であるかを判定する問題を扱っている。従来のような見えだけを用いた判定では間違える可能性が大きいので、人間が行っているように、環境の構造に依存した状況 (例えば、部屋に入ってまた出てきた) や物体との関係に依存した状況 (例えば、ある人物の所有物である車から降りてきた) を考慮して、判定を行う。さらに、情報の信頼度を考慮しながら、同一性の判定結果の信頼性を評価する。新たに加わった情報によっては、判定の信頼性が低下したり覆ったりする可能性もある。そのような推論の非単調性を実現する手段として、多値デフォルト論理を用いている。

Ait-Aider (CNRS) ら [4] は、近年のデジタルカメラに多く採用され始めた CMOS を撮像素子とするビデオカメラを用いて物体の姿勢変化と移動速度を

3次元復元する手法を提案している．現在市販されている CMOS 素子による安価なビデオカメラには，rolling shutter と呼ばれる画素毎の露光タイミングが異なるものが多く採用されているが，従来の3次元復元手法は全画素が同時に露光していることを前提としているため，このようなビデオ映像の処理には不適であった．本発表では，rolling shutter カメラをモデル化し，同一フレームの撮影中は物体の移動速度が一定であるという仮定の下でバンドル調整の枠組みによって物体の姿勢変化と移動速度を推定する手法を提案している．本発表は優秀論文賞を受賞している．

Briassouli (Illinois 大) ら [11] は，映像中の複数物体の異なる周期運動を推定する手法を提案している．映像に対して時間方向への短時間フーリエ変換 (STFT) を適用することで，時間軸と周波数軸に関する強度分布 (TFD: time-frequency distribution) を作成し，そこから極大となる軌跡を抽出することで周期を推定している．また STFT の位相のシフトを利用して，物体の平均的な並進運動成分を推定している．さらに，推定された周期と並進成分に基づいて，フレーム間での画像の相関を計算することで周期運動領域を抽出している．

Lin (National Chiao-Tung 大) ら [40] は，隠蔽や急激な動きを含む条件下での near-regular texture (NRT) の追跡手法を提案している．従来手法では NRT の追跡時に texton 同士が誤対応することが問題であった．提案手法では，texton を各ノードとする格子を構成し，隣接格子間にばねモデルを導入することで，トポロジーに対する拘束として利用し，誤対応を軽減している．さらに，texton の可視性を考慮した観測モデルを導入して，ノードの変形推定問題をマルコフネットワークで定式化することで，テクスチャの重なり等による隠蔽に対応している．

Dedeoglu (CMU) ら [16] は，active appearance model (AAM) を低解像度画像に適用する際の，より正確な形状や見えのパラメタの推定手法を提案している．従来手法では，入力画像を AAM の解像度に合わせるように補間することでモデル当てはめを行っていたため，入力画像が AAM に比べて低解像度である場合にはパラメタの推定精度が著しく落ちるといった問題があった．これに対して提案手法では，入力画像の解像度で AAM を撮影する場合のカメラモデルを考慮し，そのモデルに基づいて AAM の解像度を入力画像側に合わせてから当てはめを行うこ

とで，AAM の形状及び見えのパラメタの正確な推定を実現している．

Khan (Central Florida 大) ら [34] は，隠蔽を含む雑踏シーンでの人物追跡手法を提案している．従来の色情報や形状情報を利用した人物追跡では，似た色や形状の人物が複数存在する場合に追跡に失敗することがあった．提案手法では，人物の足下の点が地面に対するホモグラフィを満たすことを利用し，各カメラの画像の各点に対する前景確率を，参照カメラの画像にホモグラフィで変換して統合する (掛け合わせる) ことで，人物の存在確率地図を作成する．次に存在確率地図を2値化して時間軸上に並べて時空間ボリュームを作成し，そこから一貫性のある軌跡をグラフカットにより抽出することで，各人物を追跡する．

Makihara (Osaka 大) ら [45] は，方向変化に対応した歩容認証手法を提案している．最初に歩行者の時空間のシルエットボリュームを作成して，そこからフーリエ解析によって歩行周期に基づく周波数領域特徴を抽出している．次に，複数被験者の複数方向の周波数領域特徴を用いて方向変換モデルを学習し，登録人物の特徴を入力特徴と同じ方向の特徴に変換することで照合を行っている．実験では，15度刻みの24方向の歩行に対する個人認証を行い，手法の有効性を示している (国内資料 [44]) ．

Wang (Monash) 大ら [62] は，物体追跡のための見え方のモデルとして spatial-color mixture of Gaussian (SMOG) を定義し，それをベースとした類似性尺度を提案している．さらにそれを用いた粒子型フィルタにおいて，integral Gaussian mixture (IGM) と呼ぶ SMOG のパラメタの高速計算法を提案している．SMOG とは，画像内の2次元座標と RGB 色空間における座標の独立性を考慮して点の確率密度をモデル化したものであり，彼らは，それらに関する統計量を用いて類似性尺度を提案している．この尺度は，通常よく用いられるカラーヒストグラムを用いた Bhattacharyya 距離と比べて，平行移動に関して高い識別能力を持つこと，また IGM を用いた粒子型フィルタでは，通常カラーヒストグラムを用いた CONDENSATION [26] に比べ，計算時間は短く，かつ粒子数や追跡領域サイズの大きさに影響を受けにくいことを示している．

Yan (North Carolina 大) ら [66] は，特徴点軌跡の分離の問題としての motion segmentation を，意欲的なタイトルに示されるように，関節でつながった

物体, 非剛体, 退化のある場合, およびそれらが任意に組み合わせられた状況において行うための方法を提唱している. その手順は, 計測行列の各列を単位ベクトルに正規化することに始まり, 同一物体に属する特徴点の局所性を反映させながら, 軌跡の部分空間の分離を affinity matrix の spectral clustering によって行う流れとなっている. GPCA との比較において優位性が論じられている.

(以上, 三浦, 佐藤, 八木, 榎原, 金澤, 牧)

7. 3次元復元・多視点幾何学関係

Gurdjos (TCI-IRIT, Toulouse) ら [21] は未校正のカメラで撮影したシーン中の平行な円を用いた平面の復元に関する手法を提案している. 彼らはこれまで2個の平行な円を用いて解かれてきた問題を2個以上の平行な円を用いた場合に拡張して, 厳密な定式化を行っている. 円が2個の場合には厳密な解が計算できるが, 円の数がそれ以上の場合には最小二乗法による解を計算し, バンドルアジャストメントによる非線形最適化を行うことで解の精度を向上している. 実験ではシミュレーションと実画像を用いて, 多数の円を使用した場合の精度の向上性を示している.

Furukawa (Illinois 大) ら [20] は, 複数画像からの3次元モデリングにおいて, silhouette constraint と photo consistency を同時に満たす高精度なモデルを自動生成する手法を提案している. 提案手法では, まずシルエット画像から visual hull を算出し, 次に visual hull と物体の境界線を推定する. また, グラフカットによって photo consistency を満たす面を構成する. これにより得られた初期モデルに対して, silhouette constraint と photo consistency から成るエネルギーを最小化することで最終的な出力を得る. 小物体をモデリングするために必要な計算時間は一般的な PC (P4-3GHz) で2時間程度とのことであるが, 生成結果は非常に高精細で細かな凹凸まで再現されている.

同様の枠組みで, 複数視点画像から photo consistency に基づく3次元モデリングを, グラフカットを利用したエネルギー最小化によって行う下記3件のポスター発表 [59, 24, 38] があった. いずれも Vogiatis ら [61] の手法を共通のベースとしたものと考えられる.

Tran (Maryland 大) ら [59] は visual hull を初期値とし, やはりグラフカットアルゴリズムを用いたエ

ネルギー最小化により color consistency を満たす最適な表面を計算している. ただし, グラフカットが通常短いパスを好む傾向にあり, 求まる表面が平滑化されてしまうという問題に対し, 突起部分に constraint points なる通過すべき点を設けた上で最適化を行っている. その上で, 背景差分から求められるシルエットの不確定性を考慮することで凹面を積極的に削る工夫がなされている.

Hornung (Aachen 大) ら [24] は, 形状復元処理のハードウェア実装による高速化の提案を行っている. visual hull の生成, 隠蔽のチェック, photo consistency の計算をパイプライン化し, photo consistency については画像のペアを単位とするかわりに, すべての入力画像を同時に考慮して計算する提案を行っている. CPU (P4-3.2GHz) と市販の GPU (NVIDIA GeForce 6800) 上で3-85倍の高速化が可能となり, 従来は数時間が必要であった計算時間を10分程度までに短縮するものと報告されている.

Lempitsky (Moscow State 大) ら [38] は, 複数視点画像からの形状復元のうちでも, 特に各視点から物体表面上の各パッチが見えているか否かを判断する段階に焦点をあてている. 従来手法では推定形状を利用して隠蔽のチェックが行われているのに対し, 各パッチの位置と法線方向に基づく oriented visibility なる概念を導入している. 突起部分への対応や計算コストの観点ではそれぞれ前記2件に譲ると思われるが, 初期解として visual hull を必要としない, あるいはよい初期解が与えられていない場合の有効性が示されている.

Kemelmacher (Weizmann Institute) ら [33] は, 顔の形状を, 1枚の正面顔画像を手がかりに一般的な顔の3次元モデルを参照・変形させることによって求める手法を提案している. ランバート反射と無限遠光源を仮定し, shape from shading の枠組みの下, 未知の光源, 表面の反射率および法線方向を最適化問題の解として復元する. まず3次元モデルを参照して光源の初期推定を行い, 次いでモデルの反射率も併用して形状を求め, さらに対象の反射率を計算する. このステップの繰り返しを行う. 実験例では性別, 年齢, 表情の異なる様々な入力画像が対象とされている. patent pending とのことであった.

Chen (Academia Sinica) ら [13] は, 4枚の画像から, 未知の遠方光源下で photometric stereo を解く提案をしている. 対象物体にはランバート表面を仮定し, その法線方向, 反射率, 各入力画像における光

源を逐次処理によって推定する．先ず1次の球面調和関数で表現される光源を初期推定し，これに基づいて法線方向を計算する．続いて光源パラメタを2次の項を用いてさら新し，反射率も求め，さらに法線方向を探索によって再計算する．従来に比べて入力画像の枚数が少なく済むこと，attached shadowを取り扱えることを特徴としている．

Agarwal (Maryland 大) ら [3] は，gradient fieldからの物体表面形状の復元問題に関して，gradient field が必ずしも積分可能でない場合においても表面形状を復元できる手法を提案している．本手法では，勾配を積分する際に勾配の信頼度を算出し，積分方向によって重みを変えることで，ノイズが存在する場合にもロバストに表面形状を推定できるとしている．また，本発表では，従来提案されている Poisson solver や Frankot-Chellappa 法が，提案手法の特別な場合として表現可能であることを示している．なお，本研究に関してもソースコードが公開されている．

Agarwal (UC San Diego) ら [1] は，多視点画像からの3次元復元やカメラパラメタ推定における再投影誤差の非線形最小化問題に関して，大域最小解を効率的に求める手法を提案している．この手法では，求めるべき各パラメタに対して上限値，下限値が与えられているものとし，分岐限定法によってその探索範囲を再帰的に限定していく．これにより，大域最小解から微小に離れた一定の範囲内に解が求まるとしている．また，本手法はガウスノイズを想定した L_2 ノルムによる再投影誤差最小化問題だけでなく，誤差関数を L_1 ノルムとした場合にも同一の枠組みで手法を適用できるという特徴がある．

Hasinoff (Toronto 大) ら [22] は，髪の毛のように複雑で精細な3次元形状を持つ物体やシーンを対象に奥行き画像を密に推定する手法を提案している．本研究では，絞り・焦点距離を変化させながら対象となるシーンを高解像カメラ（本発表では12M pixelカメラ）で撮影し， $A \times F$ 枚の aperture-focus image (AFI) を解析することで奥行きを推定する．基本的な奥行き推定のアプローチは shape from focus そのものであるが，本手法では画素が合焦している場合には絞りを変化させても観測される輝度値が変化しないという仮定に基づき，ウィンドウを用いなくても画素ごとに合焦判定ができることに特長があるとしている．また，本発表では AFI において画素ごとの幾何学的な対応関係を算出するためのカメラモデルとその校正手法についても紹介された．なお，本

発表は優秀論文賞を受賞した．

Schindler (Monash 大) ら [52] は，モデル選択を用いた，シーケンス画像からの複数物体の形状と運動復元方法を提案している．これは，まず隣り合う2フレーム間でモンテカルロサンプリングにより，運動モデルとインライアを抽出し，得られた運動をクラスタ化しておく．次に，前後のフレーム間において似た運動を見つけ，運動候補の集合を作る．最後にモデル選択により，全フレームを通した運動を見つけている．個々に使われている手法自体は既出のものであり新しくないが，それら既出の方法をうまく組み合わせることによって，完全な透視投影における複数物体の一般運動の分離に成功している．

Kanatani (Okayama 大) ら [31] はアフィンカメラモデルをより一般化する因子分解法を拡張を示し，未校正カメラによる画像列から3次元復元ができることを示している（国内資料 [7, 30]）．

（以上，菅谷，佐藤，牧，金澤，金谷）

8. 統計的学習関係

Subbarao (Rutgers 大) ら [55] は，グラスマン多様体に projection based M-estimator (pbM) [14] を適用することにより，ロバストに部分空間を推定する方法を提案している．提案法においては，部分空間の次元や運動物体の個数は既知とするものの，インライアが中央に分布しない場合や部分空間が直交していない場合にも分離可能であり，アウトライアの方が多い場合でも可能である．実画像を用いた実験では，運動物体の分離を例にとり，一般化 PCA [60] や RANSAC などに比べて，より正確な分離が可能となることを示している．

Heiler (Mannheim 大) ら [23] は，特徴抽出や画像表現，認識などに用いられる非負テンソル因子分解において，sparseness を制御する分解方法を提案している．これは，非負行列因子分解における sparse 度の尺度をテンソルに拡張したものであり，同時に sparsity maximization algorithm (SMA) と呼ぶ分解方法を提案している．そして MIT CBCL の顔画像データベースを用いた実験により，sparseness を制御しない非負テンソル因子分解に比べ，認識精度が向上したことを示している．

（以上，金澤）

9. 表面反射モデル関係

Mahajan (Columbia 大) ら [43] は，一様な反射

特性を持つ曲面に対して、BRDF・照明・観測色の関係を球面調和係数によって表現し、その等式に基づくBRDF・照明の復元方法等を提案している。本論文では、 r 個の複数照明と s 個の複数物体 (BRDF) に対する r 行 s 列の観測色に対する球面調和係数行列が、ランクが 1 になることが示されている。また、その拘束により、照明や BRDF を直接復元せずに、照明や BRDF の変化による観測色の変換関数が得られることを示している。実験では、変換関数に基づく画像の一貫性判定 (照明変化等で説明が付くかどうか) を行い、画像の不正改竄検出等への応用が可能であることを主張している。

White (UC Berkley) ら [64] は、有限色で構成される非剛体表面に対するテクスチャ置き換えの手法を提案している。最初に、各照度下におけるテクスチャ中の各色と参照色 (白色) の観測値の対応関係を学習し、表面中の各テクスチャの色を対応した参照色に置き換えていくことで、元のテクスチャが消去された陰影効果のある参照色表面を作成している。次に規則的に並べられたテクスチャパターン検出、もしくは正面テクスチャの変形フィッティングにより、新しいテクスチャをマッピングする。これにより、従来の表面形状推定に基づく手法では困難であった布のしわの部分においても、テクスチャ置き換えが可能であることを報告している。

Mallick (UC San Diego) ら [46] は、静止画像及び動画における鏡面・拡散反射成分の分離手法を提案している。最初に、色空間において既知の光源色軸 (S) とその直交平面 (UV) を考え、色ベクトルの UV 平面とのなす角を ϕ 、UV 平面への射影ベクトルの大きさを ρ 、方向を一般化色相 θ とする色空間 (ρ, θ, ϕ) を定義している。次に、同一反射率領域内の鏡面反射成分のない点において ϕ が最小になることに基づいて、 ϕ に対する収縮を行うことで拡散反射成分のみの画像を生成し、反射成分を分離している。また、収縮の停止関数には、領域境界のための一般化色相 θ の勾配や動画に対するオプティカルフロー等の情報が利用されている。

(以上、八木、楨原)

10. 顔・ジェスチャー関係

Zhang (Harbin 大) ら [67] は、顔の発見に、誤り訂正符号 (ECOC) を用いたパターン識別器を適用している。対象として、正面顔、左右の横顔、顔以外の 4 つのクラスを考え、それらを冗長に識別するた

めの 7 つの 2 元分類器を用いている。画像パターンの特徴量として、spatial histogram features (SHF) と呼ばれる特徴ベクトルを用い、2 元分類器としては、識別に対する寄与度の高い SHF の要素をいくつか、順に適用して階層的に分類する手法と SVM とを組み合わせたものを用いている。顔の向きの変化に強く、またいろいろな方向の画像を同時に検出できる。

Lin (Hong Kong 大) ら [41] は、通常画像と赤外線画像、あるいは通常画像と似顔絵、といった異なる空間 (モダリティ) で得られた顔データ間の照合の問題を扱っている。それぞれの空間でのデータを共通の空間へマッピングし、そこで照合を行う。そのためのマッピングを、共通空間での識別性能およびマッピングの連続性に基づく評価関数を用いて決定する。一方の空間を他方の空間へ変換し、そこで照合を行う場合に比べ、よりよい識別性能が示されている。また、カーネルトリックを併用することにより、さらに性能を上げることができる。

Arandjelović (Cambridge 大) ら [6] は、さまざまな条件下で撮像された画像系列を用いた顔認識を扱っている。まず 2 つの画像系列 1, 2 について、系列 1 の各画像に対し、系列 2 中で顔の向きが最も似ているものを選びだし、その近傍の画像を用いて、向きは系列 1、照明条件は系列 2 と同じになるような新たな画像系列を生成する (この処理を reillumination と呼ぶ)。画素値が反射率とその他の要因 (照明や形状など) の積で表されるという弱い仮定だけを用いると、生成した系列と元の系列 1 からそれらの対数の差をとった系列は多様体 (shape-illumination manifold, SIM と呼ぶ) を構成する。照明条件と顔の向きを十分に変動させて画像系列を取得したときの SIM がすべての人に共通である (gSIM と呼ぶ) と仮定すると、入力系列をデータベース中のある人物の系列を用いて reillumination を行って得た差の系列が gSIM にどの程度適合するかどうかで、その人物かどうかを評価できる。既存の手法に比べて、かなり高い認識率を達成している。

(以上、三浦)

参考文献

- [1] S. Agarwal, M. Chandraker, F. Kahl, D. Kriegman, and S. Belongie, Practical global optimization for multiview geometry, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 592–605.
- [2] A. Agarwal and B. Triggs, Hyperfeatures – Multi-level local coding for visual recognition, *Proc. 9th*

- ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 30–43.
- [3] A. Agrawal, R. Raskar, and R. Chellappa, What is the range of surface reconstructions from a gradient field? *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 587–591.
- [4] O. Ait-Aider, N. Andreff, J. M. Lavest, and P. Martinet, Simultaneous object pose and velocity computation using a single view from a rolling shutter camera, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 56–68.
- [5] F. Alter, Y. Matsushita, and X. Tang, An intensity similarity measure in low-light conditions, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 267–280.
- [6] O. Arandjelović and R. Cipolla, Face recognition from video using the generic shape-illumination manifold, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 27–40.
- [7] 浅原 清太郎, 金谷 健一, 菅谷 保之, ハノ・アッカーマン, 未校正因子分解法による 3 次元復元: 比較実験, 情報処理学会研究報告, 2005-CVIM-151-20 (2005-11), 145–152.
- [8] S. Avidan and M. Butman, Blind vision, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 1–13.
- [9] H. Bay, T. Tuytelaars, and L. Van Gool, SURF: Speeded up robust features, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 404–417.
- [10] M. Bray, P. Kohli, and P. Torr, PoseCut: Simultaneous segmentation and 3D pose estimation of humans using dynamic graph-cuts, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 642–655.
- [11] A. Briassouli and N. Ahuja, Estimation of multiple periodic motions from video, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 147–159.
- [12] Y. Cai, N. de Freitas, and J. Little, Robust visual tracking for multiple targets, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 107–118.
- [13] C.-P. Chen and C.-S. Chen, The 4-source photometric stereo under general unknown lighting, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 72–83.
- [14] H. Chen and P. Meer, Robust regression with projection based M-estimators, *Proc. 9th ICCV*, October 2003, Nice, France, Vol.2, pp. 878–885.
- [15] D. Crandall and D. Huttenlocher, Weakly supervised learning of part-based spatial models for visual object recognition, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 16–29.
- [16] G. Dedeoğlu, S. Baker, and T. Kanade, Resolution-aware fitting of active appearance models to low resolution images, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 83–97.
- [17] P. Favaro and, E. Grisan, Defocus inpainting, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol.2 , pp. 349–359.
- [18] V. Ferrari, T. Tuytelaars, and L. Van Gool, Object detection by contour segment networks, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 14–28.
- [19] A. Foulonneau, P. Charbonnier, and F. Heitz, Affine-invariant multi-reference shape priors for active contours, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 601–613.
- [20] Y. Furukawa and J. Ponce, Carved visual hulls for image-based modeling, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 564–577.
- [21] P. Gurdjos, P. Sturm, and Y. Wu, Euclidean structure from $N > 2$ parallel circles: Theory and algorithms, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 238–252.
- [22] S. W. Hasinoff and K. N. Kutulakos, Confocal stereo, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 620–634.
- [23] M. Heiler, and C. Schnoerr, Controlling sparseness in non-negative tensor factorization, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 56–67.
- [24] A. Hornung and L. Kobbelt, Robust and efficient photo-consistency estimation for volumetric 3D reconstruction, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 179–190.
- [25] 石川 博, ヒト視覚系から示唆される高階ステレオ事前分布, 画像の認識・理解シンポジウム (MIRU2006), July 2006, 仙台.
- [26] M. Isard and A. Blake, Condensation—conditional density propagation for visual tracking, *Int. J. Comput. Vision*, **29**-1 (1998), 5–28.
- [27] H. Ishikawa and D. Geiger, Rethinking the prior model for stereo, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 526–537.
- [28] T. Jiang and C. Tomasi, Level-set curve particles, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 633–644.
- [29] K. Kanatani, Ellipse fitting with hyperaccuracy, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 484–495.
- [30] 金谷 健一, 浅原 清太郎, 菅谷 保之, ハノ・アッカーマン, 未校正因子分解法: カメラモデルを指定しないユークリッド復元, 情報処理学会研究報告, 2005-CVIM-150-16 (2005-9), pp. 131–138.
- [31] K. Kanatani, Y. Sugaya, and H. Ackermann, Uncalibrated factorization using a variable symmetric affine camera, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 147–158.
- [32] D. Kaziska and A. Srivastava, Cyclostationary processes on shape spaces for gait-based recognition, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 442–453.
- [33] I. Kemelmacher and R. Basri, Molding face shapes by example, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 277–288.
- [34] S. Khan and M. Shah, A Multiview approach to tracking people in crowded scenes using a planar homography constraint, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 133–146.
- [35] K. Kim and L. Davis, Multi-camera tracking and segmentation of occluded people on ground plane using search-guided particle filtering, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 98–109.
- [36] P. Kohli and P. H. S. Torr, Measuring uncertainty in graph cut solutions—Efficiently computing min-marginal energies using dynamic graph cuts *Proc.*

- 9th ECCV, May 2006, Graz, Austria, Vol. 2, pp. 20–43.
- [37] V. Kolmogorov, C. Rother, Comparison of energy minimization algorithms for highly connected graphs, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 1–15.
- [38] V. Lempitsky, Y. Boykov, and D. Ivanov, Oriented visibility for multiview reconstruction, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 226–238.
- [39] A. Levin and Y. Weiss, Learning to combine bottom-up and top-down segmentation, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 581–594.
- [40] W.-C. Lin and Y. Liu, Tracking dynamic near-regular textures under occlusions and rapid movements, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 44–55.
- [41] D. Lin and X. Tang, Inter-modality face recognition, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 13–26.
- [42] D. Lowe, Distinctive image features from scale-invariant keypoints, cascade filtering approach, *Int. J. Comput. Vision*, **60**-11 (2004), 91–110.
- [43] D. Mahajan, R. Ramamoorthi, and B. Curless, A theory of spherical harmonic identities for BRDF/lighting transfer and image consistency, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 41–55.
- [44] 榎原 靖, 佐川 立昌, 向川 康博, 越後 富夫, 八木 康史, 周波数領域における方向変換モデルを用いた歩容認証, 情報処理学会研究報告, 2006-CVIM-152 (2006-1), pp.117–124.
- [45] Y. Makihara, R. Sagawa, Y. Mukaigawa, T. Echigo, and Y. Yagi, Gait recognition using a view transformation model in the frequency domain, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 151–163.
- [46] S. Mallik, T. Zickler, P. Belhumeur, and D. Kriegman, Specularity removal in images and videos: A PDE approach, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 550–563.
- [47] A. Micilotta, E.-J. Ong, and R. Bowden, Real-time upper body detection and 3D pose estimation in monoscopic images, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 139–150.
- [48] M. Özuysal, F. Fleuret, V. Lepetit, and P. Fua, Feature harvesting for tracking-by-detection, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 592–506.
- [49] B. Platel, E. Balmachnova, L. Florack, and B. ter Haar Romeny, Top-points as interest points for image matching, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 418–429.
- [50] X. Ren, C. Fowlkes, and J. Malik, Figure/ground assignment in natural images, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 614–627.
- [51] E. Rosten and D. Tom, Machine learning for very high-speed corner detection, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 430–443.
- [52] K. Schindler and J. U, H. Wang, Perspective n-view multibody structure-and-motion through model selection, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 606–619.
- [53] V. Shet, D. Harwood, and L. Davis, Multivalued default logic for identity maintenance in visual surveillance, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 119–132.
- [54] J. Shotton, J. Winn, C. Rother, and A. Criminisi, TextonBoost: Joint appearance, shape and context modeling for multi-class object recognition and segmentation, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 1–15.
- [55] R. Subbarao and P. Meer, Subspace estimation using projection based M-estimators over Grassmann manifolds, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 1, pp. 301–312.
- [56] J. Sullivan and S. Carlsson, Tracking and labelling of interacting multiple targets, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 619–632.
- [57] J. Sun, W. Zhang, X. Tang and H.-Y. Shum, Background cut, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 628–641.
- [58] R. Szeliski, R. Zabih, D. Scharstein, O. Veksler, V. Kolmogorov, A. Agarwala, M. Tappen, and C. Rother, A comparative study of energy minimization methods for Markov random fields, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol.2 , pp. 16–29.
- [59] S. Tran and L. Davis, 3D surface reconstruction using graph cuts with surface constraints, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 2, pp. 219–231.
- [60] R. Vidal, Yi Ma and S. Sastry, Generalized principal component analysis (GPCA), *IEEE Trans. Patt. Anal. Mach. Intell.* **27**-12 (2005), 1–15.
- [61] G. Vogiatzis, P. H. S. Torr, and R. Cipolla, Multi-view stereo via volumetric graph-cuts, *Proc. IEEE CVPR*, Jun 2005, San Diego, CA, U.S.A., Vol. 2, pp. 391–398.
- [62] H. Wang, D. Suter, and K. Schindler, Effective appearance model and similarity measure for particle filtering and visual tracking, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 606–618.
- [63] X. Wang, K. Tieu and E. Grimson, Learning semantic scene models by trajectory analysis, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 3, pp. 110–123.
- [64] R. White and D. A. Forsyth, Retexturing single views using texture and shading, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 70–81.
- [65] 山田 純平, 金谷 健一, 超精度の楕円当てはめ, 情報処理学会研究報告, 2005-CVIM-151-15 (2005-11), 107–114.
- [66] J. Yan and M. Pollefeys, A general framework for motion segmentation: Independent, articulated, rigid, non-rigid, degenerate and non-degenerate, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 94–106.
- [67] H. Zhang, W. Gao, X. Chen, S. Shan, and D. Zhao, Robust multi-view face detection using error correcting output codes, *Proc. 9th ECCV*, May 2006, Graz, Austria, Vol. 4, pp. 1–12.