

2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2007) 報告

岡谷 貴之 (東北大学)
杉本 茂樹 (東京工業大学)
日浦 慎作 (大阪大学)
向川 康博 (大阪大学)

木下 航一 (オムロン株式会社)
西山 正志 (株式会社東芝)
右田 剛史 (岡山大学)
奥富 正敏 (東京工業大学)

okatani@fractal.is.tohoku.ac.jp
shige@ok.ctrl.titech.ac.jp
shinsaku@sys.es.osaka-u.ac.jp
mukaigaw@am.sanken.osaka-u.ac.jp

kino@ari.ncl.omron.co.jp
masashi.nishiyama@toshiba.co.jp
migita@it.okayama-u.ac.jp
mxo@ctrl.titech.ac.jp

あらまし 2007 年 6 月 19 日～21 日に米国ミネソタ州ミネアポリスで開催された CVPR2007 の概要を報告する。

Report on 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2007)

Takayuki Okatani (Tohoku University)
Shigeki Sugimoto (Tokyo Inst. of Tech.)
Shinsaku Hiura (Osaka University)
Yasuhiro Mukaigawa (Osaka University)

Koichi Kinoshita (Omron Corporation)
Masashi Nishiyama (Toshiba Corporation)
Tsuyoshi Migita (Okayama University)
Masatoshi Okutomi (Tokyo Inst. of Tech.)

Abstract This report gives an overview of 2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2007), held in Minneapolis, Minnesota, USA, from June 19th to 21st, 2007.

1. はじめに

2007 年 6 月 19 日から 21 日の 3 日間にわたり, IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2007) が米国ミネソタ州ミネアポリスの Hyatt Regency Hotel にて開催された。今回は第 25 回目にあたり, General Chairs は Takeo Kanade (Carnegie Mellon University) と Gerard Medioni (University of Southern California), Program Chairs は Simon Baker (Microsoft Research), Jiri Matas (Czech Technical University), Ramin Zabih (Cornell University) が務めた。

投稿論文数は, 引き続き増加傾向にあり, 1,256 件と新記録を更新した。その中からオーラル発表が 60

件, ポスター発表が 292 件採択された。採択率は昨年とほぼ同様で, オーラルで 4.8%, ポスターを含めた全体でも 28.0%とかなり厳しい。表 1 に, 第 1 著者の所属に基づいて分類した, 国別発表件数を示す。開催国のアメリカが相変わらず圧倒的に多いが, 中国とフランスが前回に比べて 10 件以上増えているのが目を引く。

表 2 はセッションの一覧を示している。オーラルセッションは全 15 セッションで構成され, 1 日目と 3 日目は 2 セッションパラレル, 2 日目はシングルトラックで, 受賞論文はこの日に集められていた。

ポスターセッションは毎日午前と午後の 2 セッションで全 6 セッションだが, 表 2 に示すように, 分野ご

表 1 国別発表件数

国名	オーラル	ポスター	合計
アメリカ	30	155	185
中国	6	31	37
フランス	3	19	22
イギリス	2	14	16
イスラエル	7	8	15
カナダ	1	10	11
スイス	2	8	10
ドイツ	1	9	10
オーストラリア	1	7	8
シンガポール	2	4	6
日本	0	6	6
オーストリア	0	6	6
チェコ	1	2	3
韓国	0	3	3
スウェーデン	1	1	2
ロシア	0	2	2
インド	0	2	2
台湾	1	0	1
ギリシャ	1	0	1
オランダ	1	0	1
ブラジル	0	1	1
ウルグアイ	0	1	1
トルコ	0	1	1
フィンランド	0	1	1
スロベニア	0	1	1
合計	60	292	352

とにさらに細分化されていた。今回の CVPR で特徴的なこととして、全ポスターが開催期間中ずっと貼り出されており、参加者はいつでも全てのポスターを見ることができた。もちろん、各発表者は原則割り当てられたセッションの時間のみ説明をすればよいのだが、中には期間中ずっと熱心に説明をしている発表者もいた。そこまでではなくとも、気が向いたとき、あるいは興味を示した人が現れたときにはいつでも、自分のポスターを前にして説明をすることができるため、報告者はなかなか良いシステム（スペースが許せばであるが）であると思った。

賞に関しては、以下のようである。Best Paper Prize は、B. Leibe, N. Cornelis, K. Cornelis, and L. van Gool (ETH Zurich, KU Leuven) による "Dynamic 3D Scene Analysis from a Moving Vehicle" が、次点の Best Paper Prize Runner Ups は、A. Levin, A. Rav Acha, and D. Lischinski (Hebrew Univ., MIT) による "Spectral Matting" と、O. Tuzel, F. Porikli, and P. Meer (Rutgers Univ., Mitsubishi Electric Research Labs.) による "Human Detection via Classification on Riemannian Manifolds" がそれぞれ受賞

表 2 セッション一覧

オーラルセッション		
7/19	1A	Matching and Features
7/19	1B	Motion Segmentation and Tracking
7/19	2A	Enhancement 1: Blur and Resolution
7/19	2B	Medical
7/19	3A	Learning and Shape
7/19	3B	3D and Geometry
7/20	1	Recognition, Learning, and Optimization
7/20	2	3D/Graphics
7/20	3	Detection/Matching/Tracking
7/21	1A	Search and Optimization
7/21	1B	Physics
7/21	2A	Enhancement 2: Noise
7/21	2B	Handwriting and Faces
7/21	3A	Shape and Boundaries
7/21	3B	Multimodal and Sign Language
ポスターセッション		
Learning and Pattern Recognition (1,2,3)		
Optical Flow and Tracking (1,2,3)		
Sensing, Photometrics, and Image Processing (1,2)		
Segmentation (1,2)		
Shape (1,2)		
Navigation and SLAM		
Recognition and Detection (1,2,3)		
Faces and Biometrics (1,2,3)		
Geometry and Structure-From-Motion (1,2)		
Surveillance and Change Detection		
Graphics and Computational Photography		
Features, Regions, and Boundaries		
Body Tracking, Gait, and Gesture (1,2)		
3D Reconstruction and Processing (1,2)		
Retrieval and Search (1,2)		
Stereo (1,2)		
Document Processing		
Miscellaneous Applications		
Medical Imaging (1,2) and Biologically Motivated		

した。

さらに学生が第1著者の論文に与えられる Best Student Paper Prize は、Y. Li, H. Ai, T. Yamashita, S. Lao, and M. Kawade (Tsinghua Univ., Omron Corp.) による、"Tracking in Low Frame Rate Video: A Cascade Particle Filter with Discriminative Observers of Different Lifespans" が受賞した。

また CVPR1997 で発表された論文の中から、"10 年の歳月のテストに耐えた貢献"に対して与えられる Longuet-Higgins Prize は、J. Shi and J. Malik (CMU, UCB) による "Normalized Cuts and Image Segmentation" と、E. Osuna, R. Freund, and F. Girosi (MIT) による "Training Support Vector Ma-

chines: An Application to Face Detection” が受賞した。

加えて、デモ発表者を対象にした Best Demo Prize が, Jan-Michael Frahm, A. Akbarzadeh, P. Mordohai, D. Nister and M. Pollefeys (UNC Chapel Hill) による “Real-time Urban Reconstruction” に、査読者を対象にした Outstanding Reviewer Awards が合計 23 人の査読者に対して贈られた。

また本会議前後の合計 3 日間に、15 のワークショップと 9 つのチュートリアルが行われた。本会議と合わせ拡大傾向が続いているようである。

(以上、奥富)

2. オーラルセッション

Matching and Features

本セッションは、関連分野の 5 件の発表があった。うち 2 件は画像セグメンテーションとの関連も深い。

Nowak と Jurie (INRIA) は、未知の物体を撮影した 2 枚の画像が同一の物体のものか否かを識別する際に、従来のように画像ごとに特徴量 (descriptor) を抽出するのではなく、画像ペア (すなわち 2 枚) ごとに特徴量を抽出して学習する方法を提案した。ここでは、画像ペア中の複数の類似パッチペアを randomized tree に通すことで特徴量を算出し、LDA (線形判別分析) による判定を行っている。

Jegou ら (INRIA) は、bag-of-features と最近傍探索に基づいた画像データベース検索方法として、特徴空間における距離の正規化による高精度化と、inverted file system に基づいた検索効率化手法を提案した。画像特徴として、PCA-SIFT 特徴空間を量子化した ‘visual word’ の頻度を利用しており、提案する距離の正規化方法を、文書検索における索引語頻度の偏りによる重み付け方法と対比させている。

Winder と Brown (Microsoft Research) は、SIFT や GLOH 等の既存の特徴量算出手法を、特徴変換 (transformation) やサンプリングなどの複数の処理ブロックの組み合わせとし、各処理ブロックにおけるパラメータを ROC 曲線面積最大化によって最適化した。各ブロックは既存手法では未使用な処理にも置き換えられ、組み合わせにより従来よりも良いマッチング結果を得ている。

Deng (Oregon State Univ.) らは、MSER のように輝度変化の少ない注目領域を検出する方法として、principal curvature 画像による領域分割を利用する手法を提案した。ここでは、ヘッセ行列固有値を画像化する方法 (C.Steger, PAMI 1998) に基づいて principal curvature 画像を生成した後に、固有ベクトルの方向を利用して境界線をつなげる。そしてスケールの違

いによる分割結果のロバスト性を利用して同一セグメントの注目領域を検出する。

Toshev ら (Univ. of Pennsylvania) は、2 枚の画像間の領域分割と領域対応を同時に求める手法として、Joint Image Graph (JIG) を用いた手法を提案した。JIG の重み行列には、領域分割によって定まる画像内の画素間類似度と、特徴点対応によって定まる画像間の画素間類似度が保持され、そこから intra/inter image similarity がそれぞれ算出される。そこで、それらの合算を最大にする領域分割と、その結果による特徴点対応の更新処理を繰返し行う。

(以上、杉本)

Motion Segmentation and Tracking

このセッションでは以下の 5 件の発表があった。

Tron と Vidal (Johns Hopkins Univ.) は、画像系列における複数の剛体運動物体の特徴点軌跡を入力として、アフィンカメラの仮定の下に物体ごとに軌跡を分離する問題について、これまでに提案されている 4 つの手法 (GPCA, LSA, MSL, RANSAC) を評価・比較している。速度と精度のバランスから LSA にベストの評価を与えている。

Li (Australian National Univ.) は、上の研究と同種の、ただし透視投影のカメラモデルの下で 2 枚の画像間の軌跡を運動物体ごとに分類する問題に対し、アルゴリズムを提案している。最尤推定では、各運動 k を記述する基礎行列 F_k と、各軌跡 i がどの物体 k に属するかを示す 2 値の変数 z_{ik} が未知数となるが、 F_k の「候補」の集合を用意しこれを定数とすることで、問題を z_{ik} のみの最小化問題に書き換え、最終的に (整数) 線形計画問題へと帰着している。

Lu (Siemens Corp. Research) と Hager (Johns Hopkins Univ.) は、時々刻々背景が (そして前景も) 変化するような動画像系列上において前景を分離抽出する問題に対し、画像小領域ごとの特徴量により背景と前景の 2 クラスの分類をするようにした上で、画像の変化に追従して両クラスのモデルを更新する方法を提案している。

Ali と Shah (Univ. of Central Florida) は、動画像系列上の人や車の流れのセグメンテーションを計算する手法を提案している。流れの異常の検出への応用が示唆されている。

Yang (UNC Chapel Hill) らは、4 つの固定カメラの集合体を用いた、画像のアピアランスのみに基づくトラッキング法を提案している。4 つの固定カメラからの画像集合にパラメトリック固有空間法を適用し、画像変化からカメラの運動パラメータを逆算する。特徴点の抽出を要さず、透明物体や曲面鏡などの物体で構成されたシーンにも適用できるという。なお著

者は、ビジュアルサーボを扱った類似研究 (Deguchi, K, IJCV37(1), 7-20) の存在を指摘しておきたい。

(以上, 岡谷)

Enhancement 1: Blur and Resolution

このセッションの発表は、いずれも低解像度の画像やぶれ・ぼけを含んだ画像から、より高解像度または鮮明な画像を得る超解像処理に関するものであった。この分野では複数の画像を用いる手法が実用的段階に達しているが、このセッションの3件は全て単一の画像をもとにする手法であった。この場合、原画像に関するなんらかの先験的知識またはセンサにおける工夫が必要とされ、これが研究の特徴を決定づける。

Dai(Northwestern Univ.) らは先験的知識として、物体と背景の間の境界が滑らかな曲線により構成されるという edge smoothness prior を仮定している。しかし原画像のエッジは必ずしもはっきりしているとは限らず、グラデーション状のソフトエッジからなる場合がある。この場合について従来提案されている Geocuts 法を一般化し、また alpha matting 法と組み合わせることで自然な超解像結果が得られることを示した。

Jia(Chinese Univ. of Hong Kong) は物体の動きによるぶれに対し、背景と物体の境界部分に半透明な領域が観察されることを利用してぶれのカーネルを推定する手法を提案している。ただし物体と背景の色彩が二値的であることを仮定した手法である。

Agrawal(MERL) らは Flutter Shutter Camera という断続的に露光を繰り返すカメラについて、2006年 SIGGRAPH で発表した動きぶれの除去法に超解像処理を加えた結果を発表した。

(以上, 日浦)

Medical

このセッションでは医用画像処理について3件の発表がなされた。

A. Barbu (Siemens Corp. Research) らの "Hierarchical Learning of Curves; Application to Guidewire Localization in Fluoroscopy" では、X線透視診断装置で獲得した画像から、ガイドワイヤーの位置を検出する手法を提案している。ガイドワイヤーは1次元の自由曲線であり、これを表現するためのパラメータ数は膨大である。全パラメータ同時推定は困難であるため、断片化された形と見えによる階層的なモデルを用いて自由曲線の検出を行う。各階層で特徴量の異なる Probabilistic Boosting Tree を用いて自由曲線の断片を検出し、最終的に統合する。ガイドワイヤーの全体像を自動検出できる世界初のシステムである。

Y. Li (National Univ. of Singapore) らの "Model-

Guided Segmentation of 3D Neuroradiological Image Using Statistical Surface Wavelet Model" では、MRI で獲得した画像から脳の尾状核の領域を事前情報を用いて分割する手法について述べている。事前情報として3次元形状から学習した統計的表面ウェーブレットモデルを新たに提案している。このモデルを画像への当てはめることで領域分割を行う。

J. Schlecht(Univ. of Arizona) らの "Inferring Grammar-based Structure Models from 3D Microscopy Data" では、異なる焦点で撮影した複数の顕微鏡画像から、生物の構造を推定する手法を提案している。自然物の構造を再帰的に記述する L-System と類似した考え方をを用いており、確率論に基づき定義した構造の文法を再帰的に呼び出すことで、最終的な構造を決定する。

(以上, 西山)

Learning and Shape

このセッションでは物体認識と機械学習について4件の発表がなされた。

Y. Lin (National Taiwan Univ.) らの "Local Ensemble Kernel Learning for Object Category Recognition" では、物体認識に用いる特徴量を検討している。特徴量は各カテゴリに依存するがカテゴリは多数存在するため、単一の特徴量のみでは識別が難しい。そこで、特徴量がアンサンブルされた弱識別器の組み合わせで認識を行う。訓練サンプル毎にその近傍では正しく識別できる弱識別器を学習する考え方を新規に導入した。

B. Epshtein(Weizmann Inst. of Sci.) らの "Semantic Hierarchies for Recognizing Objects and Parts" では、物体認識のために、(I) 識別に用いる特徴量をカテゴリ毎に階層構造で表現し、(II) 各階層内に同カテゴリ複数物体の見えを格納する手法を提案している。従来より (I),(II) は個別に提案されていたが、これらを統合する方法について述べている。

M. Marszalek (INRIA) らの "Accurate Object Localization with Shape Masks" では、画像中の物体から輪郭の検出を行い、その輪郭から物体の位置を決定する手法を提案している。局所特徴から輪郭の仮説を複数生成し、統合された仮説の尤度で全体輪郭を決定する。統合する際に、一般化ハフ空間のように高次元空間へ投票するのではなく二次元画像上に投票する。

M. Pham (Nanyang Tech. Univ.) らの "Online Learning Asymmetric Boosted Classifiers for Object Detection" では、オンライン学習におけるブースティングの一手法を提案している。複数の弱識別器を学習するために、ある弱識別器から他の弱識別器へ訓

練データの重みを伝播させる。この際に、各弱識別器の識別誤りを等しくする制約条件を加えた点が新しい。

(以上、西山)

3D and Geometry

Oskiper (Sarnoff Corp.) らは、前向きと後ろ向きの2組のステレオカメラと慣性運動センサを使って、ユーザの位置姿勢を推定するウェアラブルシステムを提案した。15Hzで動作し、室内・屋外を500m以上移動したときの誤差は、0.5-1%程度であり、リアルタイムシステムでは従来得られなかった精度を実現した。

Schindler (Georgia Inst. of Tech.) らは、数十年に渡る異なる時点に撮影された都市の写真を、時間順に並べ替える手法を提案した。この手法は、時間に応じて変化するシーンに対する SfM であり、建物がある時点で建設され、ある期間存続した後、別の建物に建て替えられることを仮定し、写真の並べ替えを行う。

Nistér (Microsoft Research) らは、5点の対応からの Essential 行列の復元と2画像からの L2 最適三角測量について、それぞれ10次方程式と6次方程式より低い次数の定式化が存在しないことを示した。Galois 理論に基づくこの証明法は、上記の問題に限定されず、より一般の問題にも適用できる。

Li (Johns Hopkins Univ.) らは、透視投影画像列における複数物体の因子分解法の手法を提案した。特徴点の距離が既知であれば軌跡が4次元部分空間に属するので、GPCA や LSA によって特徴点を物体毎に分けることができる。実際には距離が未知であるから反復推定を行う。即ち、適当な分割に基づいて距離を計算し、距離を使って再分割を行う。

(以上、右田)

Recognition, Learning, and Optimization

このセッションでは以下の5件の発表があった。

Leordeanu(CMU) らは、画像1枚からそこに写る物体のカテゴリを答える認識問題において、単純な種類の画像特徴を物体の複数の位置で抽出し、それらのペアが作る幾何学的関係の集合(クリーク)によって、カテゴリのモデルを表現し、認識を行う方法を提案している。実験では、物体のエッジが特徴として利用されている。学習用サンプルは、画像のどこかに対象物体が写っていさえすれば良い。

Weiss(Hebrew Univ.) と Freeman(MIT) は、自然画像のモデル—例えば自然画像の空間微分の分布が一般に裾の長い非ガウス性の分布になるなどの知見—に関する研究を示している。画像(のフィルタ出力)の確率モデルにおいて評価が必要となる高次元の積

分計算を、サンプリングなどの近似に頼らず実行するアルゴリズムを提案している。

Dundar と Bi(Siemens Medical Solutions Inc.) は、カスケード型分類器(cascaded classifier)の最適化方法を提案している。全識別器を最終的な目的関数に対して最適化することとし、そのための適当な構造を持たせることで、計算時間とオンライン学習時間が短縮できるという。識別性能と実行スピード両面で、カスケード AdaBoost を勝るとする結果が報告されている。

Potetz(CMU) は、信念伝播法(belief propagation)の計算を提案している。あるポテンシャル関数のクラスに対して、クリークのサイズに比例した時間でメッセージを計算できるという。これにより高次のインタラクションを伴う複雑なグラフを対象とすることが可能となる。陰影からの形状復元を実例に、計算結果が示されている。

Komodakis(Univ. of Crete) らは、MRF の最適化のための新しいアルゴリズムを提案している。従来の α -expansion を一般化したこのアルゴリズムは、元の MRF 問題だけでなくその双対からの情報も活用する点が特徴で、精度を犠牲にすることなく従来法より3から9倍も高速に実行可能である上、NP-hard の MRF 問題のずっと広いクラスに対して、ほぼ最適な解を与えることが保証されるという。

(以上、岡谷)

3D/Graphics

本セッションではベストペーパー賞と準ベストペーパー賞を含む3件の発表があった。

Li Guan (UNC Chapel Hill) らによる“3D Occlusion Inference from Silhouette Cues”は、背景差分によって得られるシルエットを手がかりとして、シーン中に存在する遮蔽物の3次元形状を推定した。一般的な Shape-from-Silhouette (SfS) では動物体が存在する領域に着目するが、本研究では逆に遮蔽物体が存在しない領域に着目した。この領域を時間的に累積し、遮蔽物が存在する確率をベイズ推定の枠組みで定式化することで、動物体と遮蔽物体を区別できることを示している。

Bastian Leibe (ETH Zurich) らによる“Dynamic 3D Scene Analysis from a Moving Vehicle”は、車載カメラの映像から、車両と歩行者の3次元的位置と軌跡を推定するという難しい問題に取り組んだ研究であり、ベストペーパー賞に選ばれた。基本的な考え方は、Structure-from-Motion、2次元物体検出、3次元軌跡推定の各処理をうまく統合したこと、及び、様々な仮説を立てて、観測された情報をもっとうまく説明できる仮説を選択したことにある。例えば、

特徴点追跡による SfM によって路面の傾きと位置を推定する方法や、車両の複数方向からの見えを記憶しておいてアピアランススペースで物体検出する方法など、個々の手法はシンプルで新しさはないものの、先に路面の位置を推定しておくことで車両の位置が限定され、誤検出が大幅に減少する。このように、単純なアルゴリズムでも、他のアルゴリズムとうまく組み合わせることで、互いにメリットとなりうる。そのようなまい組合せを探すことが重要となるというメッセージで発表は締めくくられた。

Anat Levin (MIT) らによる “Spectral Matting” は、2 件の準ベストペーパー賞のうちの一つである。領域分割法であるスペクトルセグメンテーションを、マッティングに応用した研究である。入力画像をラブラシアン行列の固有ベクトルから算出される複数の成分に分解し、この成分を組み合わせることで自動的にマッティングを行う方法について述べている。基本的な考え方は、既に CVPR2006 で著者らによって発表されているが、分解された成分ができるだけ 0 か 1 に近くなるように固有ベクトルを線形変換することで、より精度の高い分解を実現している。また、最終的なマッティングは、分解された各成分のうちから、前景となる成分を選択することで実現できるが、この選択についてもラブラシアンから計算されるコストを評価することで、教師なしで自動的に行う方法を述べている。

(以上、向川)

Detection/Matching/Tracking

本セッションでは 4 件の発表があり、準ベストペーパー賞と学生論文賞が含まれている。

Tuzel (Rutgers Univ.) らは、リーマン多様体数学と LogitBoost を利用した人物検出法を提案して、準ベストペーパー賞を受賞した。ここでは、対称行列である共分散行列がリーマン多様体上の点として数学的に定義できることを応用し、画像領域中の勾配等から算出される共分散行列を特徴量として利用する。そして、多様体上の各点が人物に属するか否かを学習する際に、各弱識別器が多様体上の異なる接平面を受け持つように識別器を構成した。この方法は、対数写像などの多様体数学をうまく活用した方法であり、輝度変化にロバストな特徴を利用して精度のよい人物検出を実現している。

Shechtman と Irani (Weizmann Inst. of Sci.) は、’local self-similarity descriptor’ を提案し、テンプレートと類似した模様 (geometric layout) を他の画像から検出したり、動画画像テンプレート中の人物動作を他の動画画像中から検出したりする方法を提示した。この descriptor は、画素間の SSD 値を低解像

度な log-polar 座標系 (動画画像では log-log-polar 座標系) に格納したものであり、双方の画像から得られる descriptor と ’ensemble matching’ (ICCV 2005) を利用して対応を求めている。

Li (Tsinghua Univ.) らは、低フレームレートビデオ中の物体をトラッキングする方法として、異なる経過時間をサポートする 3 つの観測モデル (observer) をカスケードする手法を提案した。3 つの observer は、入力フレームに対して、particle filter の更新規則に利用される尤度をそれぞれ出力する。そして、observer のサポート時間が短い順に更新をカスケードすることで、’cascade particle filter’ を構成した。この論文は、既存の tracking と detection の枠組みを巧みに組み合わせており、学生論文賞を受賞した。

Zhu と Lyu (Chinese Univ. of Hong Kong) は、テンプレート上の三角メッシュの変形を特徴点対応と平滑化拘束の下で求めるという非剛体変形の問題が、対応コストが特徴点対応の座標差分二乗和を底に持つ単純なロバスト誤差関数で表現される場合、単に 2 次多項式を最小化する問題に帰着できることを明らかにした。また、これを finite Newton 法の下で高速に解く事ができることを示し、実時間の非剛体トラッキングを行った。

(以上、杉本)

Search and Optimization

Basri (Weizmann Inst. of Sci.) らは、光源や視点の変化による画像の変化を表す部分空間が複数あるとき、与えられた画像に最も近い部分空間を効率的に探索する手法を提案した。部分空間を高次元空間の点に写像した後に、木構造やハッシュを用いた (近似的) 最近傍点探索の手法を用いる。写像の妥当性と誤差の上限も理論的に示された。

Olsson (Lund Univ.) らは、2 値変数を未知数とする 2 次関数の最適化として定式化されたセグメンテーションやクラスタリングに対して、スペクトル弛緩法と SDP を組み合わせた解法を提案した。一般に、スペクトル法は計算量が少なく、SDP は高精度であるが、提案手法は、これらの特徴を併せ持つ。

Rother (Microsoft Research Cambridge) らは、Boros らが提案したグラフカットの ’probing’ の効率的な実装を行い、セグメンテーション、画像合成、超解像等に適用した。グラフカット、TRW、BP、ICM、アニーリング等の既存手法と比較して常に良い結果が得られた。

Kohli (Oxford Brookes Univ.) らは、エネルギー最適化の解法として知られる α -expansion と $\alpha\beta$ -swap を拡張し、より広範な目的関数に適用できる手法を提案し、テクスチャに基づく動画画像の領域分割に適

用した。

Shekhovtsov (Czech Tech. Univ.) らは、2 画像を最適に一致させる変形場を求める際に、線形計画法を利用して計算量を削減した、手法は準最適であるが、従来法より大きな変形に対応することができる。

(以上、右田)

Physics

近年、フォトメトリや物理ベースビジョン研究は相対的に減少傾向にあるようで、CVPR では物理ベースビジョンに関する研究 5 件がこのセッションに集められていた。

Gijsenji(Univ. of Amsterdam) の研究は自然画像に対する色の恒常性に関する研究である。これについては従来より様々な手法が提案されているが、画像によってうまく動作するアルゴリズムが異なる。そこで画像の空間周波数分布の解析によりシーンの種類を判断し、適用するアルゴリズムを切り替えたり組み合わせることで良好な結果を得ている。

Tan(Hong Kong Univ. of Sci. and Tech.) は、陰影からの形状復元における形状と光源位置の間の曖昧性に対し、BRDF の等方性とヘルムホルツの相反性を適用することで解決できることを示した。Alldrin(UCSD) も同じ Bas-Relief Ambiguity 問題に取り組んでいるが、こちらは物体表面の拡散反射率の分布から計算したエントロピーを最小化する手法を提案した。

Chen(MPI Informatik) らは半透明物体など光の散乱を生じる物体の形状を光パターン投影により計測する際に、プロジェクタとカメラに装着した偏光板の角度を変えながら撮影することで高精度化する手法について従来手法と比較検討した。

Lou(UCLA) らは画像のぼけから形状復元する Shape from Defocus について、カメラのパラメータを画像から推定する手法を示した。

(以上、日浦)

Enhancement 2: Noise

このセッションでは画像からのノイズ除去や欠損部分の修復に関する 3 件の研究が発表された。Zhou(Fudan Univ.) はレンズ交換時に埃が侵入しやすいデジタル一眼レフにおいて問題視されているセンサダストについて、光学的なモデリングにより高精度に検出する手法、またダストの半影で得られる原画像の情報を用いて画像修復する手法を提案した。Chen(Hong Kong Univ. of Sci. and Tech.) らは時空間マルコフ確率場モデルによる映像シーケンスからのノイズ除去法を提案し、良好な結果を得ている。Kaftory(Israel Inst. of Tech.) らは、大気中の塵などにより遠方の映像がかすんだように観察されること

に着目し、その画像を修復する手法を提案している。遠景を撮影すると、通常画面の下方面では画面の情報よりも近くの物体が写る。そのため画像の劣化度は画面全体で均一ではなく、画面中の位置により大きく異なるため、そのような不均等性を考慮した手法により良好な画像修復が可能であることを示した。

(以上、日浦)

Handwriting and Face

Face Alignment の分野では、Active Appearance Model(AAM) に基礎を置く手法が広く活用されている。AAM は形状とアピアランスの両者について Generic なモデルにより、与えられた顔画像へのフィッティングを行う手法である。しかしながら、モデルの表現の多様性を増すために大量のデータで学習し、学習に使用されなかった顔にフィッティングを行う場合、フィッティングの性能が急激に劣化する課題が指摘されている。Xiaoming Liura(General Electric Global Research Center) は、アピアランスについては Discriminative な手法を導入し、これと Generic な形状のフィッティングを融合することにより、フィッティングのロバスト化を図っている。具体的には、正解パラメータに摂動を加えた形状モデルから特徴量をサンプリングし、これをネガティブサンプルとして GentleBoost による学習を行う。検出の際はこのスコアが最大となるようにパラメータを更新する。従来の一般的な手法と比べて高速、かつ精度のよいフィッティング結果が得られたと報告されている。

(以上、木下)

Shape and Boundaries

人の歩行動作を動画から復元する問題において、モデルの持つさまざまな制約から、復元された動作が不自然な動きを示す課題が指摘されている。本セッションの発表 4 件中 3 件が、歩行者についての動画もしくは静止画からの形状復元問題に関する報告であった。Marcus A. Brubaker ら (Univ. of Toronto) はこの問題に対して、下肢部分だけからなるシンプルな運動学的モデルをトラッキング過程に導入することにより、より自然な動作復元とトラッキングの安定化を同時に図ることが可能であることを示した。彼らが用いたモデルは非常にシンプルなものながら、これまでの一般的なものとは異なり、足が接地した際の衝撃による不連続な動作を取り扱うことが可能なものとなっている。このモデルを、パーティクルフィルタにおける時系列予測の確率モデルに組み込むことで、トラッキング処理の安定化が図られ、足の一部が隠れたような場合でも安定してトラッキングを行うことに成功している。また復元される動作もより自然なものとなることが示された。

(以上, 木下)

Multimodal and Sign Language

このセッションでは、音声と映像の統合に関して2件、手話認識に関して2件の発表が行われた。

Zohar Barzelay (Israel Inst. of Tech.) らの“Harmony in Motion”は、音声と映像の時間的な共起性に基づいて、混合された複数音声を分離する方法を提案している。音声だけでは分離が難しくても、映像中で追跡された特徴点を利用することで、分離に成功している。男性と女性の声、バイオリンとギターの音を分離する実験では、品質は完璧とまでは言えないものの、もう一方の音声はほぼ完全に消えていた。

Adam O'Donovan (Univ. of Maryland) らの“Microphone Arrays as Generalized Cameras for Integrated Audio Visual Processing”は、マイクロフォンアレイをカメラと見なして、音声処理と画像処理の統合に関する研究である。具体的には、球面状に配置されたマイクロホンアレイは、全方位カメラと見なせることから、CVにおける幾何の概念を音声処理に導入することができる。カメラとマイクロホンアレイの間で、エピソード拘束が成り立ち、カメラとマイクを用いた三角測量という考え方はおもしろい。

Ali Farhadi (Univ. of Illinois) らの“Transfer Learning in Sign language”は、手話を解析するために、アバターのアニメーションから学習された単語モデルが、人物や撮影方向が異なる場合にも変換できる Transfer Learning と呼ぶ技術の報告である。CGで描画されたアバターのラベル付き手話アニメーションは十分な数を用意できるが、人間の手話映像は限られた数しか用意できないという場合でも、共通に存在する単語の関係から単語モデルを変換できる手法である。

Ruiduo Yang (Univ. of Illinois) らの“Enhanced Level Building Algorithm to the Movement Epenthesis Problem in Sign Language”は、手話認識において連続した二つの単語をつなぐための特に意味のない動作は認識の妨げとなるが、これを動的計画法の枠組みで解決する方法を報告した。この単語間の動作はモデル化が難しく、積極的に取り除くことは難しいが、提案手法では手話単語としてうまく認識できないときに、単語間動作と認めている。

(以上, 向川)

3. ポスターセッション

Learning and Pattern Recognition (1,2,3)

Kim(Univ. of Cambridge) らは線形判別分析(LDA)のための逐次学習法を提案している。従来法に比べてより一般的で効率の良い解を出力すると

いい、バッチ LDA の結果とかなり近い解でありながらかなり計算時間を短縮できるという。

Grabner(Graz Univ. of Tech.) らは、物体認識において generative model (PCA 等) が持つロバスト性と、識別器(線形判別分析等) が持つ識別精度の両方を、同時に実現するという eigenboosting と呼ぶ手法を提案している。

Lasserre(Univ. of Cambridge) らは、1枚の画像を、jigsaw と呼ばれる画像から切り出した様々な形状のバッチの組み合わせとして表現する画像の jigsaw 生成モデルを対象に、計算量の問題から従来法では扱えなかった大きな jigsaw 画像を効率よく計算可能な方法を提案している。

(以上, 岡谷)

Optical Flow and Tracking (1,2,3)

本セッションでは、29件の発表があった。うち14件は、基本的に単一物体のトラッキングに関するものであったが、実時間推定を行うための効率化手法がいくつかみられた。Benhimane(Technical Univ. of Munich) らは、直接法ベースのテンプレートマッチングの収束を速くするための画素領域選択方法を検討した。Leung & Gong (Queen Mary Univ. of London) は、ランダムサンプリングを用いて mean sift を効率化する手法を提案した。Sakagaito & Wada (Wakayama Univ.) は、データベース検索による人物認識とその姿勢の同時推定を目的とし、NN(nearest neighbor) 法を高速に行うための Nearest First Traversing Graph を提案した。

複数物体のトラッキングに関する発表は7件あった。Yu(USC) らは、複数物体のトラッキング問題を、複数フレームに渡る軌跡とエラーによって生じる孤立領域を同時に推定する問題と考え、MCMC(Markov Chain Monte Carlo) に基いた手法を提案した。Hess & Fern(Oregon State Univ.) は、膨大な数の動物をトラッキングして数をカウントするために、消失や発生を許容した event driven な手法を提案した。

その他は、特徴点追跡関連が2件、非剛体変形推定関連が2件、テクスチャフロー推定関連が1件、dynamic texture 関連が1件、共有視野のない複数カメラによる軌跡の対応づけ1件、トラッキングを応用した物体モデリングが1件であった。

(以上, 杉本)

Sensing, Photometrics, and Image Processing(1,2)

このセッションではハードウェアや物理モデルと関連が深い研究が19件発表されていた。特にセンサにおいて従来は邪魔者扱いされてきたぼけやノイズの特性を有効に利用する手法の提案が目立つ。Matsushita

(MSRA)らはセンサの熱雑音の特性を解析し、これに基づく輝度の類似度の定義に関する研究と、カメラのガンマ補正パラメータを推定する手法に関する研究の2件を発表した。また Okatani(Tohoku Univ.)らは単眼では推定の難しい被写体のスケールを、画像のぼけから推定する手法を示した。また Kang(Microsoft Research)はレンズ色収差のために画像のエッジ部分に発生する色のにじみを除去する手法について、単に各色の画像をずらすだけでなく、カメラ内部の画像先鋭化処理の影響などを明示的に扱うことで結果が大幅に良好になることを示した。他には超解像処理やノイズ除去に関する研究が目立つ。Heo(Seoul National Univ.)らは多くのノイズを含むステレオ画像対からの距離推定とノイズ除去法を、Subbarao(Rutgers Univ.)らは画像の不連続性を保持しつつ色ノイズを除去する方法を示した。Azzabou(DxO Labs)らも画像のノイズモデルを利用したノイズ除去法を発表している。超解像分野では Yang(Univ. of Kentucky)らが距離画像の超解像処理手法を示している。

(以上、日浦)

Segmentation (1,2)

合計 19 件の発表のうち、常に人ばかりができていたポスター発表を、それぞれのセッションから 1 件ずつ紹介する。

Sharon Alpert (Weizmann Inst. of Sci.) らによる “Image Segmentation by Probabilistic Bottom-Up Aggregation and Cue Integration” は、最初はピクセルレベルから始めて徐々に領域を統合していくボトムアップ処理による領域分割法を提案している。輝度とテクスチャ情報を統合した確率モデルにより、階層的な領域分割を実現した。著者らは、パラメータ設定が不要であることを特に強調している。

Omer Rotem (Tel-Aviv Univ.) らによる “Combining Region and Edge Cues for Image Segmentation in a Probabilistic Gaussian Mixture Framework” は、色と位置情報の統計量に基づいて領域分割する確率的 GMM 手法を拡張し、画素間の勾配に基づくエッジベース手法と、画素間の類似性に基づく領域ベース手法を組み合わせる方法を提案している。ある注目画素とガウス分布の中心との間を横切るエッジの有無により重みを変化させることで、EM アルゴリズムにエッジ情報を組み込んだ。

(以上、向川)

Shape (1,2)

多角形や閉曲線で表現した物体の輪郭形状を用いてマッチングや認識を行う際、2つの形状間の距離の定義や、特徴点の安定な抽出が重要である。このとき、変形や欠損等に対処できることが望ましい。

Felzenszwalb(Univ. of Chicago)らは、境界形状の階層的な木構造表現と、それに基づく弾性マッチングを提案した。Ren(Toyota Technological Inst.)は、見え方が大きく変わるフィギュアスケートの画像を対象に、輪郭形状のクラスタリングを行った。Joshi(Florida State Univ.)らは、2つの n 次元曲線間の測地線の計算法を提案し、2つの形状を滑らかに繋ぐモーフィングを行った。Liu(HuaZhong Univ. of Sci. and Tech.)らは、ノイズに影響されにくい曲率の計算法を提案し、特徴点を抽出した。

また、輪郭表現は医用画像にも応用される。Pons(INRIA)らは、ドローネー可変形状モデルによる 3D 医用画像の領域分割を行った。Abd (Univ. of Louisville)らは、従来の符号付き距離にかわるベクトル距離関数を用いて、2つの形状の位置合わせを行った。

(以上、右田)

Navigation and SLAM

Koch and Teller(MIT)は、環境を移動するカメラのエゴモーションを、カメラで得た長い画像系列(数千フレーム)から復元する手法を提案している。環境の荒いモデル(主に室内を扱っている)と全方位カメラによる画像系列を用いている。画像の2次元エッジと環境モデルの3次元エッジのロバストなマッチング手法が新規性のポイントで、数種類の実画像を用いた性能評価が示されている。

(以上、岡谷)

Recognition and Detection (1,2,3)

合計 30 件の発表のうち、常に人ばかりができていたポスター発表を、それぞれのセッションから 1 件ずつ紹介する。

Vittorio Ferrari (INRIA) らによる “Accurate Object Detection with Deformable Shape Models Learnt from Images” は、「カップ」や「瓶」などのクラス毎に学習された輪郭モデルを用いて、シーン中からクラスに属する物体を検出する方法を提案している。クラス毎の輪郭モデルは、いくつかの事例から得られた平均的な輪郭と、そのクラス内変形のモデルとして学習される。物体の検出時には、外接矩形ではなく、モデルの変形によって正確な輪郭を見つけることが特徴である。

Grant Schindler (Georgia Inst. of Tech.) らによる “City-Scale Location Recognition” は、ボキャブラリツリーと呼ばれるデータ構造に基づいて、画像をもとに現在地を推定する手法を提案している。このような問題は従来から取り扱われているが、本研究では 20km もの市街地を対象とし、3万枚もの画像から 1億もの SIFT 特徴量を抽出するという、従来法では

破綻してしまうような規模を対象としているところが興味深い。

Deva Ramanan (Toyota Technological Inst.) らによる“Using Segmentation to Verify Object Hypotheses”は、仮説検証の枠組みで物体検出と領域分割を実現する手法を提案している。ウィンドウベースの検出器で多発する誤検出に着目し、これらの多数の仮説をトレーニングデータから学習した形状に関する知識で検証するという単純ではあるが効果的な手法である。

(以上、向川)

Faces and Biometrics (1,2,3)

多くの顔認識の手法において、画像はベクトルとして扱われ、PCA や LDA 等の手法が適用される。ここでは各ピクセルは独立したものとして扱われ、ピクセル間の関係性は一切考慮されない（つまり、必要以上に自由度の高い状態で問題を扱っている）。それに対して、近年画像を行列として扱ったまま解析を行うテンソルベースの手法への注目が高まっている。これは画像の 2 次元的な構造を保持することにより、元画像のもつ空間的な情報が考慮された解析が行われると期待されていることである。しかしながら Deng Cai ら (UIUC, Yahoo! Research Labs) は報告の中で、従来のテンソルベースの手法では実は、同一の行内（もしくは列内）の情報のみに関連性しか考慮されていないことを指摘している。彼らは顔画像のもつ 2 次方向への相関に配慮するため Laplacian penalty を拘束条件に導入したテンソルベースの部分空間学習を提案した。これにより、2 次元的な連続性が考慮された処理が可能となっている。

また、Tae-Kyun Kim ら (Univ. of Cambridge) は、LDA によるインクリメンタルラーニングの手法を提案した。特徴量次元の固有値問題を毎回解く必要がないため高速であり、必要とされるメモリ量も非常に小さなものとなっている。さらに、EM アルゴリズムによるクラスタリング推定を処理に組み込むことにより、オンラインでの準教師あり学習を実現している。

顔認証技術の実用化は急速に勢いを増しつつあるが、そこで大きな課題となっているのが照明環境の問題である。Zhifeng Li ら (The Chinese Univ. of Hong Kong) は、屋内で撮影した顔画像と屋外で撮影した顔画像が大きく異なった特性を持つことから、それぞれを別々にモデル化する必要があることを指摘する一方、両者とも同一人物の画像であれば共通した特徴をもつため、その個人固有の特徴を反映した部分空間も学習されるべきとの考えに基づき、Discriminant Mutual Subspace Learning(DMSL) と呼ばれる部分

空間学習手法を提案している。同一人物に対する屋内、屋外それぞれの環境から変換される特徴量の類似度を最大化する一方、別人物に対する各環境内での特徴量の類似度を最小化するように、各環境における変換行列を学習する。さらに、複数の識別器を別々の特徴量ごとに学習させ、それぞれの信頼度を考慮し結果を融合する Adaptive Informative Fusion と呼ばれる手法を用いることによって、通常の PCA, LDA を利用した手法に比べ、良好な認識性能を実現している。

(以上、木下)

Geometry and Structure-From-Motion(1,2)

近年、エピポーラ幾何の復元では、最小限の数の特徴点を用いて非線形連立方程式を構成し、行列の固有値問題に帰着させて解く手法が研究されている。Kukelova(Czech Tech. Univ.) らは、基礎行列と放射歪の同時推定において、ランク 2 拘束を使うことで対応点数を 9 点から 8 点に削減した。Geyer(CMU) らは、放物鏡を用いた全方位カメラを用いて、最少の 9 点で基礎行列推定を求める方法を提案した。Stewénius(Univ. of Kentucky) らは、5 点の運動ベクトルから、カメラの並進ベクトルと回転角速度を求める手法を提案した。

また、従来、行列の線形操作のみで解くために無視されることが多かったパラメータに関する制約を導入することで、精度を向上させる研究も見られた。Chandraker(Czech Tech. Univ.) らは、メトリック形状復元のための dual quadric を推定する際にランク制約を導入し、LMI(linear matrix inequality) 法によって解いた。Bocquillon(IRIT) らは、画像列からのメトリック形状復元に必要な 4 変数の非線形方程式を interval analysis によって解いた。Migita(Okayama Univ.) らは、基礎行列の推定において、探索空間をランク 2 の行列に限定することによる推定精度の向上を解析した。

(以上、右田)

Surveillance and Change Detection

Qureshi(Univ. of Toronto) と Terzopoulos(UCLA) は、監視システムを設計・評価するための仮想環境シミュレータ Virtual Vision を提案している。シミュレータでは、能動的なパンチルトズームカメラを含むマルチカメラネットワークを構成でき、実世界での実験と比べて大幅にコストが削減可能であるとしている。その他、画像列から行動ルールを自動生成する方法の研究などがあつた。

(以上、岡谷)

Graphics and Computational Photography

Computational Photography とは、計算機でデー

タを処理することを前提にセンサ等を構築し、従来困難であった撮影を可能にしようとする分野であり、2007年のSIGGRAPHでも course および paper セッションが組まれるなど近年急速に注目を集めている。Sun (Xi'an Jiaotong Univ.) らはフラッシュを照射した写真としていない写真の2枚から前景物体を切り出す手法を示しているが、前景物体はその間に移動しても良い。他には画像からの影の除去やパノラマ接続、マッピングなどの研究が発表された。

(以上、日浦)

Features, Regions, and Boundaries

特徴抽出については、従来多く用いられてきたエッジやコーナーのような小さい領域であるがありふれた特徴の検出にかわり、SIFT オペレータに代表されるようにより大きな領域から豊富な情報を取り出したり、スケール不変性を得たりする手法がトレンドになっている。Gao(Univ. of California) はこのような手法の中で Discriminant method と呼ばれる手法の安定性が高いことを示している。また Forssen(Univ. of British Columbia) は安定な色領域の検出法を提案した。

(以上、日浦)

Body Tracking, Gait, and Gesture (1,2)

人体の検出およびトラッキング、あるいは歩行動作の解析技術は、ここ数年で急速に注目を集めている研究領域であり、この CVPR2007 においても2つのセッションに別れ計22本の論文が発表され、活発な議論がなされた。この急速な関心の高まりの背景として、近年のテロや凶悪犯罪の増加に伴う、監視カメラ、監視システム等に対する社会的ニーズの増加が原因のひとつとして考えられるであろう。また技術的な側面からは、特徴量抽出技術や識別・回帰分析技術等の発展により、このような問題に対して実用的な精度で、実時間レベルでの処理の実現が視野に入ってきていることも、関心の高まりを後押ししているものと考えられる。

Alessandro Bissacco ら (Google, Inc.) は、顔検出等で一般的に使われる Haar 特徴量を10度刻みで回転させた特徴量セットを用意し、この特徴量出力と人体モデル姿勢との相関を、学習により直接求める試みについて報告を行った。人体モデルは頭部、胴体、上腕、下腕等10のパーツで表現され、これらが26の自由度を持って組み合わせられたものが使用されている。対応関係の学習には、J. H. Friedman らが提案した決定木の枠組みである Gradient TreeBoost を多次元に拡張した技術が用いられている。人体の各パーツの姿勢には強い相関関係があるため、一次元の回帰木を複数用いるよりも効率的な学習が可能に

なると報告されている。さらに、入力情報として静止画だけでなく、隣り合うフレームの差分画像を用いることによって動的な情報も付加することにより、これまでの同様の手法に比べ、良好な推定結果が得られることが示された。

(以上、木下)

3D Reconstruction and Processing (1,2)

3次元形状復元を行うには、プロジェクトとカメラを用いたレンジファインダが有効な手段である。従来は複数のパターンを投影する構造光が用いられたが、この原理に基づき復元をリアルタイムで行う研究が見られた。Weise(ETH Zurich) らは、ステレオカメラと構造光によるリアルタイム3D スキャナを提案した。Young(UCSC) らは、単一パターンを複数カメラで撮影する構造光レンジファインダを提案した。

拡散反射面のみを対象とした形状復元は比較的容易であるが、より複雑な反射モデルを対象とする研究が行われている。Liao(Univ. of Kentucky) らは、距離による光の減衰を用いたシーンの距離復元を行った。Rozenfeld(Israel Inst. of Tech.) らは、鏡・ガラス・金属等の3D復元を行った。これには、計算機のディスプレイに表示された白黒の帯の反射像を用いる。Ahmed(Univ. of Louisville) らは、Shape from Shading を、鏡面反射・透視投影カメラ・近接光源に拡張した。Chandraker(UCSD) らは、Photometric stereo において影を利用した。

計測された複数の3次元形状の位置合わせに関して、Amberg らは、ICP法を拡張して変形を扱った。Sofka(Rensselaer Polytechnic Inst.) らは、ICP法に代わるCDC法を提案した。また、レンジデータへの曲面の当てはめ(Lempitsky ら, Moscow State Univ.) や、3Dモデルと画像の位置合わせ(Liebelt ら, Technical Univ. of Munich) の発表もあった。

画像や画像列からのシーンの復元に関して、Li(The Chinese Univ. of Hong Kong) らは、1枚の画像からの多面体シーンの復元を行った。Seki(Tokyo Inst. of Tech.) らは、非同期ステレオカメラによる動的シーンにおける運動と形状の復元を行った。Lhuillier(Université Blaise Pascal) は、全方位カメラ画像列からの3Dモデリングを行った。

(以上、右田)

Retrieval and Search (1,2)

このセッションでは画像検索について15件の発表がなされた。

F. Fraundorfer(Univ. of Kentucky) らの "A Binning Scheme for Fast Hard Drive Based Image Search" では、大規模データベースの検索を高速に行うために、記憶装置へ特徴量を効率的に格納す

る方法を提案している。類似する画像をビンに対応付けることで記憶装置へのアクセス回数を大幅に減らした。

(以上, 西山)

Stereo (1,2)

近年のステレオ計測の研究は、視差や奥行きを推定するだけでなく陽にサーフェスを生成する手法に関するものが多く、サーフェス生成では、常に局所的な平面姿勢を求めることによって高精度化がなされている。本セッションでは、12件の発表があったが、半分の6件が平面姿勢推定およびその結果を利用したサーフェス生成に関連している。

Ohら(USC)の研究は、多眼ステレオと graph cuts から求めた密な奥行きを修正しつつ平面パッチを生成する方法に関するものであり、Furukawa(Univ. of Illinois)らは、ローカルな対応推定から輝度変化を考慮した平面推定を行い、エッジ検出を利用した多眼ステレオによるサーフェス生成を提案した。また、Habbecke & Kobbelt(RWTH Aachen Univ.)らは、ローカルな平面推定を周囲に拡張しており、それぞれ高精度な推定手法によるフォトリアルな結果が注目を集めていた。これらの方法は、計算速度についてはほとんど考慮されていないものの、Sugimoto & Okutomi (Tokyo Inst. of Tech.)らは、ローカルな平面推定とグローバルなサーフェス生成との両方の大幅な高速化を提案した。

また、ステレオ計測では奥行きのエッジを求めて、滑らかな視差変化をもつ局所領域を得ることが重要であるが、そのための方法として、Pock(Graz Univ. of Tech.)らはMumford-Shahのコスト関数を応用しており、Xiong & Jia(The Chinese Univ. of Hong Kong)は alpha matting を応用するなど、既存の画像セグメンテーション手法を導入する提案もみられた。その他、輝度変化を考慮したさまざまなステレオマッチングコストに関して、ローカルな手法とグローバルな手法を適用した際の精度の比較がHirschmüller(German Aerospace Center)らによってなされていた。

(以上, 杉本)

Document Processing

文書画像処理分野に分類はされているが、実際にはフォトメトリに関連する研究が2件あった。1つはTsoi(Nanyang Tech. Univ.)による、書籍のページの曲がりに起因する形状の歪みと陰影を除去する研究であり、もう1つはLandon(Univ. of Kentucky)による、しわのある紙面の陰影除去である。他には比較的劣化した文書に対する前処理、文書構造解析、文書からの署名部分の検出に関する研究発表があった。

(以上, 日浦)

Miscellaneous Applications

このセッションでは画像認識の応用について4件の発表がなされた。指の爪認識、火星画像の岩認識、定点監視画像の統計的な解析、車載カメラの霧対策が提案されていた。

(以上, 西山)

Medical Imaging (1,2) and Biologically Motivated

このセッションでは、医用画像処理について12件の発表がなされた。

X. Song(OGI School of Sci. and Eng.)らの"Speckle Tracking in 3D Echocardiography with Motion Coherence"では、3次元超音波診断装置で撮影した心臓の画像において特徴点追跡を精度よく行う手法を提案している。各特徴点を独立に追跡するのではなく、速度ベクトルの変化が特徴点間で滑らかとなる制約条件の下で最適化問題を解く点が新しい。

(以上, 西山)

4. ま と め

本稿では、2007年6月19日から21日までの3日間にわたり開催された the 25th IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR2007) について概要をまとめた。

なお、次回 CVPR2008 はアラスカ州アンカレッジで、次々回 CVPR2009 はフロリダ州マイアミで開催されることが決まっている。日本からも多くの論文が発表されることを期待したい。

(以上, 奥富)