

国際会議 "ECCV '92" 報告

八木康史* 栄藤稔** 志沢雅彦*** 谷内田正彦*

* 大阪大学基礎工学部 情報工学科
** 松下電器中央研究所 電子機器基礎研究所
*** ATR視聴覚機構研究所 認知機構研究室

〒560 豊中市待兼山町1-1

あらまし

Second European Conference on Computer Vision (ECCV'92)が、5月18日～23日、Santa Margherita Ligure (Italy)において開催された。その概要を、参加者4名が分担して報告する。

和文キーワード 会議報告, ECCV, コンピュータビジョン

Report on ECCV'92

Yasushi YAGI* , Minoru ETOH** ,
Masahiko SHIZAWA*** and Masahiko YACHIDA*

* Depart. of Information & Computer Science, Osaka University
** Central Research Lab., Matsushita Electric Industrial Co., Ltd
*** ATR Auditory and Visual Preception Research Lab.

Toyonaka, Osaka 560, Japan

Abstract

Second European Conference on Computer Vision (ECCV'92) was held in Santa Margherita Ligure, Italy, May 18-23 1992. This is a report on ECCV'92 by four participants.

英文 key words Conference Report, ECCV, Computer Vision

1. 会議の構成

ECCVは、今回が2回目であるが、ヨーロッパ内におけるコンピュータビジョンの中心的会議である。2年に一度開催され、査読はfull paperで行われる。投稿件数308件に対し、long paper 16件,short paper 41件,poster 48件と計105件が採択され、全体で採択率は34%であった。ちなみにlong paperについては、約5%と言う狭き門であった。これらの発表以外に、poster sessionと並列で、videoとexperimental sessionがあった。experimental sessionは、ワークステーションなどによる実演であった。以下、会議の規模、構成を示す。

主催： C.N.R Special Project on Robotics
European Vision Society
期間： 1992年5月18日～23日
(内本会議は4日間)
会場： Grand Hotel Miramare,
Santa Margherita Ligure, Genova, Italy
Conference Chairman： Prof. Giulio Sandini
(Dist University of Genova)
参加人数：約360名
構成： 口頭発表：57件, ポスタ：48件
ビデオ：十数件, 実演：16件
投稿数：308件
セッション構成：シングルセッション、ただし、poster,video,experimental sessionは並列
論文賞：2件
S.W.Lee and R.Bajcsy
R.Cipolla and B.Blake

2. 会議・セッションの概要と感想

4人の報告者により各人の感じた会議全体の概要ならびに印象に残った論文について報告する。

谷内田正彦

会議全体の概要と印象

ECCV'92 (European Conf. on Computer Vision)に参加した。ECCVはその名の示すごとくヨーロッパにおけるコンピュータビジョン関連の会議で今回で2回目という比較的新しい会議である。今回はイタリアのコートダジュールにあるリゾート地の一つであるSanta Margherita Ligureで5月19日-22日の間開催された。ヨーロッパ

の会議であるが、最近の国際化やコンピュータビジョンの研究者人口の増加などにより米国、日本からの参加者も多数見受けられた。また、この会議の前に、IEEE主催のInt. Conf. on Robotics and Automationが開催されていたために、ビジョン関連の研究者は両方の会議に参加する人も多数見られた(実は筆者も両方に参加した)。

会議の全体的な動向はセッション名で分かると思われるので、Posterも含むsessionの名前と発表件数を次にあげておこう。

Features(12), Color(3), Texture and Shading(5), Motion Estimation(13), Calibration and Matching(5), Depth(6), Stereo-motion(4), Tracking(6), Active Vision(6), Binocular Heads(3), Curved Surfaces and Objects(7), Reconstruction and Shape(15), Recognition(14), Applications(6)

こうしてセッション名を並べてみると、コンピュータビジョンの各分野をほぼ満遍なくカバーしているが、特徴抽出、動きの推定、3次元再構成、物体認識の発表が多いことに気づく。全体的な印象としては、実際のシーンをハードウェアで実時間に近い速度で処理するという研究が増加している点が目立った。ESPRITプロジェクトとの関連で実時間に近い速度で目的とする結果を出さないと研究費をもらうのが難しくなってきたという背景事情があるらしい。ESPRITに参加しているグループとそうでないところでは研究内容や発表方法にかなり差ができてきている。ESPRITに参加しているグループでは実際的な問題を対象として発表もビデオをきれいにアトラクティブに発表している。一方、そうでないところは理論的な研究が多くOHPなど文字が小さくて見えないなどプレゼンテーションの初歩も守られていない発表が多く見受けられた。ESPRITに参加しているグループは比較的研究費にも恵まれているようで活気が感じられた。我が国にもESPRITの様な大学も参加できる大型プロジェクトがぜひ欲しいなと感じた次第です。

会場となったSanta Margherita Ligureはリゾート地と聞いて期待して行ったのですが、単に田舎の港町といった感じで(特にRobotics and Automationが開催された世界的なリゾート地ニースの後に行ったためかもしれないが)、しかも天気が悪いため会議に出なければ何も他にすることがないという状況で会議の出席率は非常に良かった。ちなみに、この会議はシングルセッションで大きな会場が常に一杯だった。しかし、部屋が細長いため後半に席を取るとほとんど字が見えないため、良い席を取るのに苦労

した。また、会場が大きいため実りのあるディスカッションを行うのが難しい雰囲気だった。シングルセッションは聞きたいのが重なるといふ点は避けられるが、興味のない話も堪えて聞いていなければならないという欠点の方が多いのではと強く感じた。

帰国して2ヶ月後、MIRU'92に参加した。研究内容のレベルはあまり違いがなく（私個人にとってはMIRU'92の方が興味ある研究が少し多かった）、プレゼンテーションと討論の平均レベルはMIRU'92の方が少し良かったと思うのは私の欲目だろうか。

栄 藤 稔

会議全体の印象

会議の開始時点で、予稿集は参加者の半分に渡る冊数しか用意されておらず、残り半分は印刷中であった。チェアを勤める Prof. Sandiniの「ここはイタリアだから...」というジョークで会議がスタートした。彼のオープニングの基調講演では、近年のCV技術の理論の進展を述べ、ビジョンにおける諸手法の整理だけでなく、リアルタイムビジョン試作の重要性を訴えていた。会議の内容は、まさにこの基調に沿うもので、Binocular heads というセッションが独立してあるくらいである。このセッションとTrackingやApplicationのセッション、毎夕開かれたビデオセッションでは、非常に完成度の高い（スポンサーの納得の行くような）デモシステムが多く報告され、見応えのあるビデオが多かった。

Color

発表は口頭発表のみ3件あった。LeeとBajcsyは「Detection of Specularity Using Color and Multiple Views」で、物体の表面反射領域を視点移動により分離する手法を述べた。Bajcsyによれば、Physics-based visionとActive visionの2つの考えに沿った研究であるとのことである。Physics-basedにおける議論は、過去のKlinkerらのDichromatic Modelと同様である。アルゴリズムは、絶縁体の観測輝度の完全拡散反射成分が観測視点に依存しない「Lambertian constancy」の仮定に基づいている。視点移動に伴い、各画素の値は変化する。このとき、RGBの特徴空間で、各画素における観測値と他の視点で得られた画像の全ての標本値の中で最小となるユークリッド距離を計算する。観測値が完全拡散反射成分のみであれば、そのユークリッド距離は零となるが、反射成分があれば、視点移動によって零

にならない。これを、彼らはMSD(minimum spectral distance)と呼んでいる。アルゴリズムの並列性とGMの実験施設で、車の表面反射成分を抽出して概観検査に利用できるという事例を示し、有用性をアピールした。ただ、視点移動で、表面反射の位置が移動し色の変動(MSD)として観測できるような照明条件と物体はそう多くないように思える。

Funtらは「Recovering Shading from Color Images」で、表面が複色で構成された表面をShape-from-Shadingにより復元する手法を述べた。従来の輝度変化の大きな点を不連続として検出しShadingを補正する手法に比べ、色度(chromaticity)の変化を用いることから、より閾値設定がロバストであると述べている。Fathimaは「Data and Model-Driven Selection Using Color Regions」で、色における際立ち度(saliency)を定義し、色領域を用いた認識系の枠組を提案した。

Motion-estimation

動画処理では、例えばBlack「Combining Intensity and Motion for Incremental Segmentation and Tracking Over Long Image Sequences」(Trackingのセッション)や、Peleg他「Detection of Multiple Moving Objects Using Temporal Integration」など、長い時系列画像(long sequence)を扱ったもの、temporal integrationを強調する発表が多かったように思える。Bergen Anandanのグループ(David Sarnoff Research Center)は「Hierarchical Model-Based Motion Estimation」で、彼らが今まで進めてきた階層的な動き推定に加えて、制約モデルを陽に取り入れた枠組を示した。制約モデルには、「fully-parametric」(画面全体または一部ををアフィン変換または2次形式のパラメータで拘束)、「quasi-parametric」(剛体+カメラの動きパラメータ)、「non-parametric」(2Dフロー)の3段階に分け、上位階層でパラメータ推定と推定パラメータによる参照画像の変形(warping)を行ない、結果を逐次下位階層へ降ろしていく手続きを取っている。どの制約モデルを選択するかは、始めに与えられる。階層的な動き推定と(準)パラメトリック記述の制約を組み合わせた動き補償のデモビデオはその能力の高さをアピールした。他に興味ある発表として、Cipollaは「Surface Orientation and Time to Contact from Image Divergence and Deformation」で、目標となる物体の平面にスネークを置き、そのスネークの変形(制御点の移動)から直接、平面の傾きと観測者と平面との相対運動による平面への到達時間を推定する手法を述べた。

Active Vision

Aloimonos以外では、Culhane&Tsotsos: "An Attentional Prototype for Early Vision", Rimey&Brown: "Where to Look Next using aBayes Net: Incorporating Geometric Relations"など注視点制御の発表があった。Recognitionのセッションはshapedescriptionが話題の中心となっている。

その他

robust統計の応用を陽に述べたのは、Textureのセッションで、Fleckの "Texture: Plus ca Change, ..." (ca のcはsの発音のための仏語文字、theを表すジョーク?) だけであった。M推定での "outlier detection" がtextureboundaryの発見に利用できることを述べた。最小記述長規範 (MDL) を用いた発表はなかったようである。Trackingでは、モデルベースの方法とエネルギー最小化などの最適化問題に帰着させる方法が報告された。Nagelのグループは、自動車の追跡を題材に、自動車の幾何モデル、照明モデルを詳細に与えて追跡するものであった。後者は柔軟物体が対象である(Cohen他, Ueda他)。Cohenらは既に抽出された柔軟物体の輪郭時系列の対応問題を扱ったが、肝心の問題は、その輪郭時系列の抽出そのものにあると感じた。Curved surfaces and objectsのセッションでは曲面物体のAspect modelについて Ponceらは "Computing Exact Aspect Graphs of Curved Objects: Algebraic Surfaces" を発表した。Featureのセッションでは、Rosenthalerらは "Detection of General Edges and Keypoints" を発表した。従来のエッジ検出に加えて、T交差点、端点、角などの Key Pointを抽出する非線形の局所エネルギー計算型のフィルタを提案した点で興味深い。

志沢雅彦

会議全体の印象

口頭発表はすべてシングルセッションで行なわれたこともあって、以前参加したことのあるICCVやCVPRに比較して非常に密度が濃い会議であったというのが第一印象である。会議の開催地が、無名の小さいリゾート地であったこともそのもうひとつの理由であろう。また、ヨーロッパの会議ということで、採録論文もヨーロッパからのものが多いのではないかと予想していたが、ヨーロッパ以外 (特に、アメリカ) の発表が半分近いのは意外であった。

研究動向

個人的には、physics-based visionのアプローチ

チが大きな流れとして感じられた[Lee&Bajcsy]。また、それに関連して、心理物理学のデータとの整合性を主張した論文が興味を引いた (次節で紹介する[Nguyen&Huang]を参照)。初期視覚においては、線形フィルタを用いた任意方向検出の理論[Perona]、および、線形フィルタを用いたステレオ整合[Jones&Malik]の様に、画像から効率的に特徴を取り出すフィルタ列の設計理論とその応用が特徴的である。また、光源方向、表面反射率、影などが既知であることを前提とせずに陰影からの形状復元を行なう試み[Bretton et al.]の様に、従来、前提条件として仮定されてきた属性を未知数として扱おうする流れが顕著である。

特徴的な論文の紹介

1. Image Blurring Effects Due to Depth Discontinuities: Blurring that Creates Emergent Image Details (T.C.Nguyen & T.S.Huang)

この論文は、焦点ぼけと奥行きの不連続によって生じる画像面上の輝度について理論的かつ実験的に研究したものである。一般に、PentlandやSubbaraoが行なった焦点ぼけから奥行きを求める方法では、ぼけの点ひろがり関数は、一つのGaussianによってモデル化されている。しかし、奥行きの不連続が生じている点では、画像面の局所領域に複数の奥行きが存在することになるため、一般に複数の異なる広がりパラメータをもつぼけ関数の和としてモデル化されるべきである。著者らは、これを従来の一ぼけモデルに対して多重ぼけモデル(multi-component blurring, MCB)と名づけその性質を調べている。重要な結論は、この様な多重ぼけが生じている場合、輝度分布に新しい正負の極値が生じるということである。この二つの極値の間の輝度の振幅は比較的コントラストが低く、信号の周波数は高いという特徴をもつ。この影響で、Pentlandの方法はMCBが存在する場合には信頼性がかなり低くなることをシミュレーションで示した。また、実際のカメラを用いた実験系を用いて、MCBの効果が無視できないレベルで存在し、それが、実際に信号として検出できることを示した。

この論文は、さらに、人間のぼけ検出能力に関する最近の心理物理実験を報告したHessらの論文を引用し、MCBの成分に似た信号をぼけエッジに加えるとぼけの識別率が向上するという結果に対して、これは、人間の視覚系が、MCBの物理を「知っている」結果ではないかという推測している。本研究は、報告者 (志沢)

の研究に関連するので特に興味をもった。ビジョンを画像生成過程の逆問題として捉えたとき、この論文で述べられた様な画像生成の物理で、我々がまだ正しく扱わずにいる現象に目を向けることの重要性をあらためて感じた。

2. Steerable-Scalable Kernels for EdgeDetection and Junction Analysis(P.Perona)

この論文は、Freeman&Adelson, Simoncelliらのsteerable filterをさらに発展させ、方向とスケールの両方に関して線形結合による補間可能なフィルタ群とそれらの重み関数を求める一般的な方法を提案している。まず、回転角のパラメータを θ カーネルの座標を x としたとき、実現したい (θ, x) の空間から複素数への写像(フィルタ関数)は、基底関数 $a_i(x)$ と補間関数 $b_i(\theta)$ と重み σ_i によって

$$F_{\theta}^{[n]} = \sum_{i=1}^n \sigma_i a_i(x) b_i(\theta)$$

の様に[n]項の和として近似できる。これは、一般に特異値分解を用いて、得ることができる。さらに、 $a_i(x)$ に関して、 x を極座標表現し、動径 ρ とスケール σ に関して同様の分解を行なって、これらを合成するとSteerableかつScalableなフィルタが構成できる。構成されるSteerable-scalableフィルタは、一般には近似であるが、特異値分解から得られる特異値から近似誤差を見積もることができる。

著者はさらに、この設計理論を用いて任意方向のL,T,Xジャンクションおよび端点を検出、識別するためのフィルタを設計した。この方法では、Freeman&Adelsonの方法の様に窓関数が等方である必要がないので、低次の微分でもエッジ方向に強くtuneされたフィルタを設計できる。そのため、多重方向のエッジに対しても、干渉作用をある程度押えることができる様である。本論文の方法は、Freeman&Adelsonの研究に始まる少数のフィルタの加重線形和を用いて連続するパラメータで変形される方向性フィルタを実現する方法を、最も広い意味に一般化したものになっている。

八木康史

全体会議の印象

ヨーロッパの会議と言うこともあり、参加者(発表者ではない)大半はヨーロッパ諸国からであった。少数ではあるが東欧諸国(旧東ドイツ、旧ソビエトなど)からも参加者がおり、東

西の壁がなくなったことを感じさせる会議であった。

発表は、プロジェクトの影響が実験的な研究(実際にものを作って動かした研究)が多かった様に感じられる。またESPRIT関係の研究(モデルベースでの運動推定、物体認識、飛行場での監視、また交差点での監視)は発表が良かったせいでしょうか、大変印象深く感じた。前週、IEEE主催のRobotics and Automationの国際会議に出席していたため、特に感じるのかもしれないが、発表件数全体に占める移動ロボットなどロボットビジョンに関する研究の割合が低かったように感じる。私自身は、ロボットビジョンに興味を持つだけに、この分野の発表が少なかったのは若干残念だった。全体として、結構いい学会だったかと思う。

論文紹介

1. Model-Based Object Tracking in Traffic Scenes

(D.Koller, K.Daniilidis, T. Thorhallson and H.-H.Nagel)

この論文では、交差点などの自動車が往来する環境下で、静止した観測系(固定カメラ)から個々の自動車を発見、追跡する方法について報告している。自動車の形状の知識として、パラメータ化された自動車モデルを使い、さらに運動に関しては、並進速度、角速度一定で、平面上を運動すると言う前提条件を設け、運動パラメータを再帰的に推定している。

自動車のモデルは、異なった種類に対応できるように、parameterized 3D generic modelで表現されている。残念なことに現システムでは、パラメータは固定されている。モデルと観測データ間のマッチングは、各成分(各線分の位置、長さ、方向)の共分散を求め、そのマハラノビス距離を計算することで評価している。運動の推定は、光源の方向を既知とし影情報を利用することで求めている。ただし、オクルージョンが起きた場合についての考慮はなされていない。

また、別発表でエッジのロバストな検出方法についても報告され(Extraction of Line Drawings from Gray Value Images by Non-Local Analysis of Edge Element Structures (M. Otte and H.-H Nagel))、全体として動くシステムを構築しようと言う感じがした。

2. Using Automatically Constructed View-Independent Relational Model in 3D Object Recognition

(S.Zhang, G.D.Sullivan and K.D. Baker)

この論文は、3次元物体(実際には自動車)

を認識する方法で、姿勢に依存しない属性や関係からなる物体モデルの表現方法 (VIRM(view-independent relational model)) を提案している。VIRMでは、ガウス球面の中心においた対象物体のワイヤフレームモデルをランダムに選ばれた視点から見て、その時の画像面への射影像の見え方の関係を記述している。具体的には、モデルの特徴に対応するノードと、それらの特徴間でのCo-visibilityを示すハイパエッジからなる。2つの特徴間のCo-visibilityは、一方が観測できる時に他方が観測できる確率によって表現される。またVIRMには、特徴間の幾何学的な関係として、平行性、直線性、相対的サイズ、隣接関係についても記述されている。観測データとの照合は、VIRMと部分的に合う時を探索することで行われ、実際には、いくつかの仮説が生成される。これらの仮説から正しいものを選ぶ方法については、他の論文に譲っており本論文では述べられていない。姿勢の推定については、実験結果を見るかぎりあまりよくないが、照合する姿勢の候補削減には利用可能かと思う。しかし処理時間はSun sparc 2で1時間半かかっており実用性に関しては疑問である。

その他にBaker氏の研究グループからは、時系列画像を用い、路面についての拘束条件をおくことで、屋外道路シーンからトラックなどの車の運動および構造を推定する方法についても報告されていた。

また物体を追跡する研究では、F.Meyerらによる領域を特徴とした時系列間での物体追跡方法の研究があった。さらにCar-followingの研究では、T. Zielkeらによる車の左右対称性を利用した先行車両の追跡方法があった。この研究では、対称性を評価する関数を定義し、時系列画像からエッジの対称な領域を実時間で検出している。この研究については、Video Proceedingで実験結果を見ることが出来る。