

未分類

第9回名古屋CV・PRML勉強会に参加しました

by belltail

2011年7月18日

今月も、名古屋CV・PRML勉強会が開催され、参加してまいりました！

0

いつもの乱雑なメモ形式で報告いたします。

と言ってもスライドをずっと追いかけてたときとか、あと後半ほど自分の脳みその疲労がアレでメモのクオリティ下がってます><；；
そこんところはおゆるし願います…

今回の勉強会について

- 日時: 2011年7月16日
- 場所: 中部大学名古屋キャンパス(いつもの場所)
- イベントページ: <http://partake.in/events/ca2f9aaf-e529-4fce-ad91-7975ae67ec5e>
- テーマ: 先月開かれた学会CVPRの論文紹介
- 参考ページ: CVPR2011のページ <http://www.cvpapers.com/cvpr2011.html>
- together: <http://together.com/li/162170>

CVPRについて

- アメリカの国内会議
 - だけど普通に世界からくる
 - IEEE
 - 開催地はアメリカ
 - 毎年開催
 - ICCV ICPR (国際会議)は2年に1回
- こんどはつくばであるよ！
- 今年は6/20～6/25
- 投稿数1677件！！26%くらいしか
- 採択の分野をみると流行りがわかっているよ！！
- 最近はwebで公開されてたりしてすごい便利！

- 玉木先生のCVPR参加レポートもぜひ
 - <http://ir.lib.hiroshima-u.ac.jp/00031494>

miyabiartsさんより

Real-time Human Pose Recognition in Parts from Single Depth Images

- Best Paper
- Kinectによる姿勢推定の論文
- Microsoft Researchの方

背景

- 関節の3次元位置を取得して頑健でインタラクティブな姿勢推定
- ゲームとかHCIとかセキュリティとかTelePresenceとかヘルスケア
- 単一の距離画像から的人体姿勢認識が凄い盛ん
- 距離画像ベースの研究が盛ん
 - 通常カメラよりずっと計算コストが低い
 - 通常画像でもがんばる研究もそっちはそっちで盛ん

論文の目的

- Kinectによる単一距離画像で高速・高精度・低計算コストな人体姿勢認識

手法の流れ

- 人体を背景差分で検出
- 人体をパーツ分割
- 関節の三次元位置を取得

アプローチ

- 人体パーツを予め学習
 - 膨大な姿勢データが必要になるがこれは自動生成
 - 人体パーツの識別・姿勢推定
 - Random Forest識別器

学習データの生成

- モーションキャプチャデータmocap (CMU)を利用。50万データ
- 似ている形状を間引いた10000データを利用
- 形状データ
- motionbuilderで
- モーションキャプチャデータを元に変形した3Dメッシュをレンダリングしてから深度バッファを首藤
- それをもとに学習
- 結構リアル

RandomForest識別器

- 人体パーツは31種類。腕とか頭とか～
- 特徴量は、深度画像の注目点周りの2点間の距離の差
 - 計算コストが非常に低い！！
- 人体パーツ分割
 - random forest識別器によって各画素の所属する人体パーツを識別
- RandomForestは決定木の集合で、各ノードにおいて特徴を識別
 - これ自体の解説は省略で
- 生成した正解人体パーツつきの画像で学習(RandomForestの生成)

関節の3次元位置推定

- 各パーツごとに関節の3次元位置を推定
 - 密度推定を行って最瀬值を推定
 - Mean-Shiftベースの手法を利用

実験！！

- 人工データ30万枚
 - 学習データに実データは含まれない
 - 実データを学習したものはない
 - テストデータ

- 人工データセット5000枚

- 実データ8808フレーム

- 15個の異なる対象

- かなりいいかんじに出てる

- パーツ推定がうまくいけば92%で推定可能

まとめ

- 高速・高精度・低計算コスト

- 第14回関東CV勉強会(CVPR祭り)でさらに詳しく紹介する

質問

- Kinect用SDKは非常に高精度だけど、

加藤さんより

超解像について2つ！

- CVPRでは1枚超解像、1動画からの時空間ビデオ超解像、ビデオ超解像(モーション・ブラー・ノイズを推定)
 - 最初の2つを紹介

まず超解像について

- 複数フレーム

- 同じものを撮影した複数画像から

- 1枚超解像

- 別にデータベースを用意(低解像度画像と高解像度画像のペア)

- 入力自身からデータベースを生成するような手法も

Single Image Super-Resolution using Gaussian Process Regression

まず実際の画像

- お、 できる

研究概要

- 1枚超解像
- 学習とかない
- ガウス過程回帰による高解像度画像の推定
- エッジとかの局所的構造を

全体の流れ

- 入力画像を大きい画像にアップサンプル(これは当然ボケる)
- それをもう1回ダウンサンプル(こうするとボケのある低解像度画像)
- この元画像と、ボケのある低解像度画像との関係を回帰モデルで学習
- その学習結果を、最初にアップサンプルしたでっかい画像に適用する

回帰をもうちょい

- 入力ベクトル
 - ボケのある画像の周辺画素
- 出力値
 - ボケのない画像の周辺画素
- 全ての画素に対して回帰を行う

質問

1. 小さい画像で求めた回帰をそのまま大きい画像にあてはめて大丈夫？
 - 論文では普通にそのままやってた
2. その他いろいろ><

Space-Time Super-Resolution from a Single Video

- 一つの動画から時間方向と空間方向両方に解像度をあげる(つまり解像度とフレームレートが上がる)

概要

- 一本の動画を入力して時空間超解像をする

- 自己相似性を利用
 - Space-Time Patchers(5x5x3) (空間方向5x5, 前後で計3フレーム)
 - 似たようなパッチが1動画中にはたくさんあるから超解像的な原理ができる

1枚事例ベース超解像

- 入力画像からピラミッド作成(低解像度化)
- 入力画像のあるパッチの類似パッチをピラミッドのいちばん小さい画像中から見つける
- ピラミッドの大きい方に親パッチ(同一パッチ)をさたどって探していく
- そのときの子と親の関係を, 入力画像に適用することで

動画への応用

- 時間方向にもピラミッド化する
- あとは同じ

林さんより(名大M1)

Enforcing topological constraints in random field image segmentation

概要

- トポロジを保存した画像セグメンテーション
- 領域の連結性と穴の数を保存する
- グラフカットとかだと, ひとつの物体がふたつに分かれたりする
- 余計な穴も出したくない

提案手法

- ふつうのセグメンテーションは, 確率場を利用する
 - これはエネルギー関数の最小化問題
 - 画素値とか隣接関係などからエネルギー関数を定義する
 - これをグラフカットに
 - これに制約を導入すると?
 - トポロジによる制約を導入する
 - これを満たす領域でエネルギーが最小の領域のものを求める

質問

1. トポロジに関する情報は既知とするかんじ

- ただ「ひとつの物体を認識したいのでひとつの領域」というかんじの制約設定でわりといい

Natural Image Denoising : Optimally and Inherent Bounds

背景・目的

- 画像のノイズ除去
 - 近年まで飛躍的に進歩
 - でもここそこ大きな改善が無い…
 - じゃあ、ノイズ除去の限界ってどうなの？
- 理論上可能なノイズ除去はどのくらいなの？
 - パッチベースの手法で理論を

概要

- ノイズ除去とはノイズ画像からノイズのない画像を推定すること
- 評価値はPSNR
 - MSEの逆数の対数
- こうしてプロットしてみると、既存の手法がすでに理論値に近い…！？

全体のまとめ

久徳さんから

- 自己紹介！！
- ustのひと！
- fararrow9
- 車載カメラの研究
- 発表資料はこちら: http://www.murase.nuie.nagoya-u.ac.jp/~kyutoku/nagoyacv/kyutoku_nagoyacv_110716.pdf

Glare Encoding of High Dynamic Range Images

HDRとは

- 広いダイナミックレンジをもつ画像を生成
- 通常は複数の露出で撮影した画像を合成している
 - 動いてる画像には向き

研究の目的

- 1枚の画像からHDRしたい！！
- だたLDRから拡張するだけでは限界がある
 - 飽和部分つまり得られてない情報をもっともらしく推測するしかできない
 - そうではなく、フィルタをかけることで能動的に情報を取得
 - レンズに物理的なフィルタをかける
 - きらきらするかんじのフィルタ.

提案手法の流れ

- グレアフィルタのかかった画像から既存手法によってグレアを除去
- これをもとに飽和画素からグレア除去のパラメータをつかって
- うおおおお

実験

- 多くの光源があってもロバストに推定

本手法の制約

- 飽和領域の面積が大きすぎるとむり
 - グレアがわからないほど非飽和領域が残っていないレベル
 - つまり点光源でないとしんどい><
 - 面光源は難しい、つまり部屋の中から窓も写してるかんじだとしんどいかも

まとめ

- 単一画像からHDR画像取得手法の提案

- 既存手法とは大きく異なる
- 特別な器材・センサは不要
- 今後
 - 検出の用意なパターンを持つフィルタの設計
 - ビデオ撮影への適用

皆川さんより

- 関東CVの幹事
- @takminさん
- 慶應大の共同研究員
- CVのアプリケーションを開発するような仕事

A Coarse-to-fine approach for fast deformable object detection

- Deformable Part Modelという物体検出手法を高速化

すみませんずっとスライドを追いかけていました。

- <http://d.hatena.ne.jp/takmin/20110716/1310831901>
- <http://www.slideshare.net/takmin/pedersoli-objectdetection>

山田さんより

- 立命館大学のD3
- 手の形状推定
- japancvもLTするよ！

2D nonrigid partial shape matching using mcmc and contour subdivision

- 変形や隠蔽を含む場合の部分輪郭の対応付けを行う
 - テンプレートと、部分隠蔽のある輪郭情報との対応付け

結果例

論文の概要

- 輪郭の相違度に、回転・並進・スケールに関してロバストなProcrustes Distanceを利用
- 入力形状となる輪郭点列とテンプレートとなる輪郭点列の対応関係を記述す行列Mを導入

輪郭の相違度

- Procrustes Distance
 - 重心を基準とした座標系
 - スケール・並進・回転に不偏

部分輪郭の対応付けと相違度

- Matching Matrix(M)で表現される対応点Z1,Z2だけを用いたProcrustes Distanceを利用
- Mは隣接行列っぽいかんじ？

対応付けの尤度

- 尤度関数

Matching Matrixに求められる性質

- テンプレート点列と類似した点を対応付ける
 - 尤度関数の値を大きくすること
 - できるだけ多くの点を対応付ける
 - 隠蔽による輪郭は連続的に出現するはずだ
- 事前分布でこれを反映させる

事前分布

- 事後分布=尤度関数*事前分布
- この事前分布はギャップGによって決まる(Gが大きいと事前分布が小さく)
- ギャップは対応点が無いとか

中部大・後藤さんより

Monocular 3D Scene Understanding with Explicit Occlusion Reasoning

- 単眼カメラによる3次元シーンの理解と遮蔽の推定
- 全身検出器
 - 人に隠れが生じると検出不可
- 部位の検出器
 - 遮蔽への対応が可能だけど誤対応も

本論文で使用する人検出器

- 人の全身検出器
- 6種類の人の部位の検出器
- 特徴量はHoG, 学習はSVM, 学習データはINRIA Person Dataset

検出対象シーンの条件

- 地平面と三次元座標上のy軸が平行
- 内部パラメータと外部パラメータは既知
- 対象物体は地平面上に

三次元シーンの理解

- 対象物体が地平面に経つと仮定
- 観測モデルと事前知識でやる
- 検出器の定義
- 地平面上の対象物体のスコア

遮蔽の推定と複数識別器の統合

- 検出結果の前後関係の理解
- 前景物体との重なり具合を用いて複数識別器を統合
 - 前景との重なりが大きい時は期待値を低く

実験

- deformable part modelより精度が良好！！

藤吉研のB4の方、三品さん

Recognition Using Visual Phrases

- 複数の物体とその関係を考慮した物体認識
 - Visual Phrasesの導入
 - オブジェクトを複合すると見え方の変化が大きい

Decoding Multiple Detections

Deformable Part Model

- さっきもあった！

検出器

- Latent SVM
- パーツモデルを導入して、目的関数を劣勾配降下法で

評価実験

- 単一クラス識別器と比較すると、ほとんどのPhraseにおいて精度向上

まとめ

- 複数のオブジェクトを持つ物体の認識に置いて精度向上
- 構造を理解することで単一物体に関しても精度向上

各務さんより(藤吉研)

Cross-View Action Recognition via View Knowledge Transfer

研究の目的

- 見え方の異なる2つ以上の映像から人の動作を認識
- どの見え方でも有効であるような動きの特徴を得る
- 2つの映像による転移学習

転移学習

- ・「新規タスクの効果的な仮説を効率的に見つけ出すために一つ以上の別のタスクで学習された知識を得てそれを運用する問題」
- ・チェスの戦略のためにチェッカーの戦略を転移学習する
- ・ある言語を理解するために別の言語を転移学習する

手法の流れ

- ・2つ以上のViewからそれぞれ特徴としてBoVW(Bug of Visual Words)を得る
- ・訓練データベクトルから2項グラフをつくる
- ・これをグラフカットでガッ
- ・グラフの重みは類似度w
- ・つまりもっとも関係の弱いところでカットするかんじ

比較実験

- ・他の同様な手法と比べると全体的に精度向上！！

後藤さん(藤吉研)のもう一人のほう

Discrete-Continuous Optimization for Large-Scale Structure from Motion

- ・オンライン画像から3次元モデルを復元

従来の復元の手順

- ・特徴点検出
- ・特徴点の対応付け
- ・初期化
- ・バンドル調整

逐次バンドル調整

- ・多くのシーンに対応可能だけど、計算コストがすごい
- ・入力画像の順序によって結果が変わる

提案手法

- ・データセットをグラフの幾何学的制約で表現
- ・MRFでモデル化

- カメラペア間での制約
- カメラ単体の制約
- カメラパラメータの最適化をする
 - 回転と並進の誤差が最小になるように

Discrete BP

提案手法の流れ

- $\backslash(^o^)\backslash$

実験

- 計算時間が6倍とかに高速化！

中部大・縣さん

Markerless Motion Capture of Interacting Characters Using Multi-view Image Segmentation

- マークをなくしたモーションキャプチャしたい
 - そういう手法は結構精度いい方法が既にある
- 従来手法では対象人物が一人なので、それを二人にしたい
- また、接触部分も正しくキャプチャしたい

アプローチ

- MRFに基づいて2人をセグメンテーションしてそれぞれでキャプチャ
 - 色の尤度と形状情報を組み合わせることにちうい
- 表面形状とスケルトンを推定

セグメンテーションの方法

- グラフカット
 - それに姿勢と形状情報を表す項を追加して精度向上
 - 色の尤度だけつかってると、うまくセグメンテーションできないことがある

従来法との比較

- 接触部分とかも改善できた！
- マーカーありによる技術との誤差も30mm以内に入るなど十分実用的！

懇親会

- 金山の山ちゃんに
 - 1週間前にQt勉強会で名古屋駅の近くの山ちゃんに行ってたりw
- 名大関係者に周囲を固められたかんじでしたw
- いわゆる純粋な数学と工学数学がどうのっていう話とかありました
 - 僕は工学数学の方しかぜんぜんわからんです
 - 数学者は統計学を異端と呼ぶらしいですね
- 懇親会の話は結構飲んだのであんまし覚えてませんww

感想

- 最先端研究の論文を日本語で紹介してもらうって機会もあまりないです
- twitterでも発言しましたがいくつかは僕も自分の所属する研究室で紹介したいと思います
- 藤吉先生の研究室からB4の方が2人担当されていました
 - 僕もB4なんですよね
 - すごいなあ...
- 実は前回勉強会のあとmiyabiartsさんから何かしゃべらないかとお誘いを頂いて僕も非常にノリ気だったのですが...
 - 院試が一院試が一
 - 甘えですね
 - そういうしてるうちに声すらかからなくなったらおしまいですよっと

ところで

すぐ次の週、というより明後日からMIRU2011があり、その前日には全日本CV勉強会があります
LTで私が5分しゃべることになりました

まだ何をしゃべるか考えてないですがww

備考

来月はお盆なのでお休みです！！

僕もOSC名古屋2011の日まで地元だしちょうどよかったですww

This entry was tagged **Computer Vision**, **イベント**. Bookmark the [permalink](#).

← 第8回名古屋CV・PRML勉強会
に参加しました

全日本CV勉強会に参加しました。
ついでに少ししゃべりました. →