カメラレジストレーション技術の仕組みと複合現実感における応用

亀田 能成
筑波大学

2014/10/30
サイエンティフィック・システム研究会
合同分科会 2014年度会合

概要

- カメラレジストレーション基盤技術の解説

  (3次元)カメラレジストレーションとは:

  カメラ位置姿勢を正確に求めること

  – 方法と原理
  – 長所と短所
  – 得意分野と不得意分野

- 様々な研究における応用事例

情報を世界に溶かし込む技術

- 視覚的に情報を溶かすための技術
  - 人間の視覚系（頭脳）に直接投影
  - 人間の視覚系（網膜）に直接投影
  - 人間が肉眼で見ている世界に投影
  - 人間が見るであろう映像内に投影 ★
    - 映像は人間の置いたカメラから。

人間の視覚系（頭脳）に直接投影

- 研究レベル
  - ○　認知として完璧（になりうる）
  - ×　現時点では視覚を失った人を対象
  - ×　要外科手術

人間の視覚系（網膜）に直接投影

- 網膜ディスプレイ
  - ○　装置装着のみで実現可能
  - ○　消費電力小
  - ？　網膜への影響
  - ×　発色
  - ×　価格

人間が肉眼で見ている世界に投影

- 光学透過ヘッドマウントディスプレイ
  - ○　装置として十分に実現可能
  - △　外界の視認性
  - △　情報の視認性
  - △　発色（レーザー光源）
  - ×　奥行き感のない表示（ハーフミラー）

人間が見るであろう映像内に投影★

- ビデオ透過ヘッドマウントディスプレイ
  - ○　装置として普及（PC・スマートフォン・タブレット）
  - △　外界と情報の視認性
  - ×　肉眼による視界との不一致、一体感の欠如

情報を溶かし込むメディア：映像

- カメラで撮影

- 様々な付加情報を映像中に正確に重畳

  拡張現実（Augmented Reality），複合現実(Mixed Reality)

- 映像からカメラ位置姿勢を正確に推定
  - ３次元世界に対して
  - 映像に対して

映像によるカメラキャリブレーションが必要な理由

- カメラの典型的スペック
  - 水平1920画素・垂直1080画素(FullHD)・30fps
  - 水平画角45°程度

  ⇒ 0.0234°/画素

  ⇒ GPS、加速度センサ、電子コンパスでは
  実現困難な精度

  ※高解像度、ズームアップなどでさらにシビア

↓

◆ 映像解析によってのみ実現可能な精度

カメラレジストレーション

- 問題の根幹：透視投影
  - 三次元世界 vs 二次元画像
    - 不良設定問題
      （画像から三次元世界には一意に戻れない）

- 目標
  - 三次元世界におけるカメラの位置姿勢
    - 根本的な位置合わせの実現
  - 映像内での位置合わせ
    - カメラが映している「三次元世界の見かけ」での一致

カメラレジストレーションの手掛かり

手掛かりになる物体（基準物体）を用意

- ３次元形状（準備無し）
- ３次元形状（準備有り）
- ２次元形状（準備無し）
- ２次元形状（準備有り）

基準物体

- ２次元形状（準備有り）
  - ARToolKit
- ２次元形状（準備無し）
  - 平面を仮定したテキスチャ解析
- ３次元形状（準備有り）
  - 形状既知のランドマーク物体
- ３次元形状（準備無し）
  - 静止世界の自然特徴 / PTAM

3次元物体の“3次元形状”

- 3次元物体
  - 形状
  - 色
- 3次元物体の観測
  - 頂点　★ ⇒ 点群
  - 辺
  - 面（領域）

透視投影と点群による幾何拘束

- 前提
  - 全ての点群は静止
  - 点群は3次元状に分布
  - 透視投影（カメラ内部パラメータは既知とする）
  - 歪み無し

- 幾何拘束
  - ３次元点群の見え方はただ一通り

幾何拘束を解く手がかり

ある点について、

- ３次元世界での座標
- ２次元映像での座標

の対応付けをN点用意

- 世界の中のどの点を使うのか？
- ３次元世界での座標値を誰が与えるか？
- ２次元世界での座標値を誰が与えるか？

点群の配置

A) 配置が2次元状でその三次元座標が既知
B) 配置が2次元状でその三次元座標が未知
C) 配置が3次元状でその三次元座標が既知
D) 配置が3次元状でその三次元座標が未知

A. ２次元状で既知

- N点について $(X_i, Y_i, Z_i) \Leftrightarrow (u_i, v_i)$
- N点は $aX_i + bY_i + cZ_i + d = 0$（基準平面）を満たす

- 「ホモグラフィー変換」の拘束条件が適用可能
  - ホモグラフィー変換自体は平面から平面への対応付け
  - ホモグラフィー変換からカメラの三次元位置姿勢を算出可能
  - 解法は様々なところ実装済
  - 数値解析的に安定させやすい

- 点の選択：
  - 画像処理によって抽出しやすいパターン中の数点
- ３次元座標：
  - パターン上で事前に計測
- ２次元座標：
  - 画像処理により推定

A. ２次元状で既知・Tips

- 平面で記述できる局所特徴の推定は画像処理の進化によって高精度化

- 三次元空間中の基準平面に沿った移動による位置での合成は精度がよい
- 三次元空間中の基準平面に垂直に離れる位置での合成は精度が急激に悪化

B. ２次元状で未知

- N点について $(X_i, Y_i, Z_i) \Leftrightarrow (u_i, v_i)$
- 但し $aX_i + bY_i + cZ_i + d = 0$（基準平面）の係数が未知

- 基準平面を実行時の最初に登録
  - 当該映像フレーム（基準フレーム）中の一部（平面部分）を指定（正対して撮影した面ならそのまま「平面」と同じ扱い）

- 点の選択：
  - 画像処理によって抽出しやすいパターン中の数点
- ３次元座標：
  - 基準フレーム中の指定領域内の局所特徴(*)を登録
- ２次元座標：
  - 基準特徴(*)と同じ見えの点を探索してその座標で対応付け

B. ２次元状で未知・Tips

- 基準フレームでの範囲指定は、三次元世界で平面であることが必須
- 局所特徴(*)を自動抽出に頼るのでその範囲はテキスチャが豊富であることが望ましい
- その範囲を撮影するときには、その面の法線とカメラ光軸の向きを一致させることが望ましい

C. ３次元状で既知

- 基準物体上のN点について $(X_i, Y_i, Z_i) \Leftrightarrow (u_i, v_i)$

- 「PnP問題」
  - 解法は様々なところで実装済
  - 数値解析的に安定させやすい

- 点の選択：基準物体上の点（多いほどよい）
  - $N \geq 5$（透視投影でカメラ内部パラメータ既知の場合）
  - $N \geq 8$（透視投影でカメラ内部パラメータ未知の場合）
- ３次元座標：基準物体上で事前に計測
- ２次元座標：画像処理ないし手作業で推定

C. ３次元状で既知・Tips

- 後で重畳させる仮想物体の想定範囲を必ずこの N 点で囲われる部分空間に収めること
  - 広い空間を覆うのは手間
- 通常 N は数十～数百点
- 基準点に画像処理で抽出しやすい模様を施すと自動処理しやすくなる

- 応用例:「２次元状で既知」で用いる基準物体を3軸直交させて配置

D. ３次元状で未知

- ３次元世界中のN点について $(X_i, Y_i, Z_i) \Leftrightarrow (u_i, v_i)$

- 映像フレームからN点を選んでも3次元座標は不明
- N点が互いに静止していることを利用すれば、フレーム間でN点の対応を取ることで各点の三次元座標を算出することが可能
- よいN点の選択は最近の画像処理技術の進化により実現

- 点の選択：
  - ３次元世界中の点（画像処理による「よりユニークに見える」点の選択）
- ３次元座標：
  - 上述アルゴリズムから算出
- ２次元座標：
  - 画像処理の過程で獲得

D. ３次元状で未知・Tips

- 使う点数は多いほどよい
- フレームを越えて追跡し続けなければならない
- 人手には頼れないので画像処理のみで実現
  - 各点の見かけはユニークであるほうが有利
  - 各点の見かけを記述する形式は詳細であることが必要
  - ただし詳細化しすぎると今度は次フレームで「同じ」と判断できなくなる怖れがあるのでバランスが必要

画像処理による点の抽出

- あるフレーム中での「ユニークな点」
  - 似た点がないこと
  - その点の近傍の微分値を上手に組み合わせ
  - 映像フレーム上で:
    - 微分する方向
    - 微分する単位（スケール）
    - 粒度
  - 大域的に観察する必要はない

※得られた局所特徴はしばしば「キー」と呼ばれる

.

画像処理による点の記述

- フレーム間で対応付け
  - （2次元形状で未知）
  - （3次元形状で未知）
- 間違った対応をしない程度には詳細に記述
- 正しい対応を見逃さない程度に概要を把握
  - トレードオフ

※「キー」の表現とそれによる追跡

画像処理による点の扱いの限界

- 理論のベースは2次元配列の画素の解析

- ３次元空間中での物体･カメラの移動による見えの変化にはそれほど追従できない
  - 最近の手法でも ±30°程度

「点」の探し方と扱い方

- 最近の画像処理技術の成果
  - SIFT(1999)、その祖先、その子孫
  - 点(近傍)を高次元ベクトルで記述することが一般化

- 高次元データ空間でのPRMLの進歩
  - 探索技術（あるデータベクトルに似ているデータベクトルを大量のデータベクトルの中から探し出す技術）の高速化
  - 次元の呪いの回避

画像処理における「点」

- 最近の画像処理技術の成果
  - SIFT(1999)、その祖先、その子孫

局所特徴量としての「点」

- 気持ち良く高次元！
  - 記述力高し
  - 実は人間の視覚系にも近い？
    - ガボールフィルタ
    - エッジベースでの単純な差分（微分）の組み合わせ
- 結果として
  - ユニーク
    - 取り違えなし
    - スケール・回転・移動不変
  - 計算は大変
    - そこは創意工夫で！

「点」特徴量

- ユニーク
  - Bag of words にすら使えるほどの一意性
  - 異なる画像間での点の対応付けの精度を保証

- 1つの点（とその近傍を）の表現
  - 高次元ベクトル
  - 数十～百次元程度の要素
  - 各要素は実数値
    - 高速化を意識した手法の場合は整数値や二値も

対応点探索

- 高次元ベクトル空間！
  - 各点が高次元ベクトルで表現
  - データ処理
    - 最近傍探索
    - １対N
    - 本来は大変な処理（要全解探索）
    - Nearest neighbors (NN)法の近似解法についての進歩

デモ・研究紹介

- デモ
  - Casper Cartridge（USBブートによるAR向けLinux環境）
    - OpenCV
    - ARToolKit
    - PTAM

- 研究紹介
  - 「３次元状・既知」による例（サッカースタジアム）
  - 「２次元状・既知」による例（自動車運転支援）

連絡先

- 画像処理や拡張現実などに関する相談
（具体的な共同研究でなくとも構いません）

- 筑波大学 画像情報研究室 亀田能成
（計算科学研究センター）
kameda@iit.tsukuba.ac.jp