

高精度GPS視覚障害者歩行支援システムの開発
画像によるシーン検出

亀田 能成
筑波大学

“What You See Tells Where You Are”
「一回行ったことあるのなら見覚えあるでしょ？」

屋内外を問わない定位

- 歩行者ナビゲーションへの究極基盤技術
 - いつでも (24時間)
 - どこでも (屋内・屋外・人が行けるところ全て)
(世界の側に設備不要)
 - だれでも (普及可能なコンシューマ技術)
- 視覚障害者への支援
 - Visionをその手に (First Person Vision)
 - 外出への需要と期待

シーンベースの位置推定

- 空間中の特定の目標物に頼らない
- カメラで撮影した写真そのものを利用

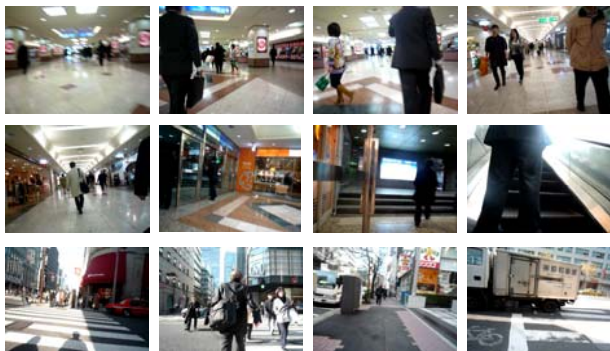


シーンベースの位置推定

- 空間中の特定の目標物に頼らない
- カメラで撮影した写真そのものを利用
 - 全く同じ構図の写真を取ろうとすると同じところに行くことに
(ズーム固定の場合)
- 手順
 1. 予めあちこちでたくさん撮影しておく
 2. 実際にお出かけ時に1枚撮る
 3. 記憶を手繰って一番近い構図をどこで取ったか思い出す
- 難点
 - そんなに記憶力のいい人いるの? ⇒ 画像認識

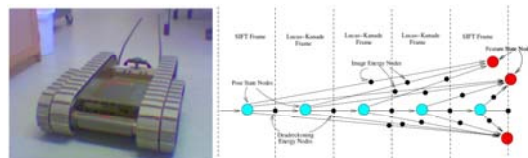
シーンベースの位置推定

- 空間中の特定の目標物に頼らない…とはいえ。



周辺研究

- Visual SLAM
 - Robotics (特に車両)
 - (例) Outdoor SLAM using stereo vision and SIFT features
 - (例) SIFT Based Graphical SLAM on a Packbot



【想定】 環境には「障害物が比較的少ない」
最終目標は「正確な定位」(自動運転のため)

First Person Visionの実際

- 外出計画の例
 - 東京駅八重洲地下中央口⇒高島屋⇒ノースタワー
 - 地上・地下・階段・ビル・街路樹・人々・etc

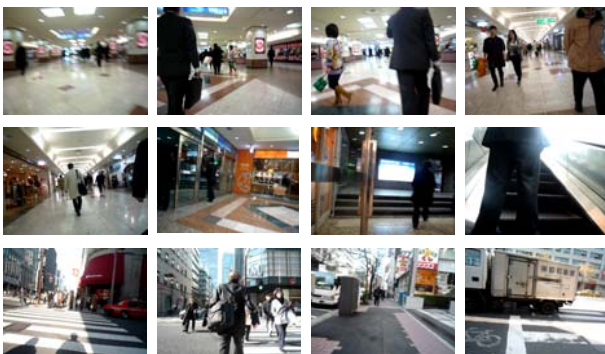


非Visualな屋外定位技術

- Differential GPS
 - Satellite Based Augmentation System (SBAS)
 - 準天頂衛星
- 多数のビーコンを対象地区に埋込
 - 屋内GPS
 - 電波ビーコン
- 自立航法 (Dead Reckoning)
 - IMU、電子コンパス

First Person Visionの画像例

やっぱり難しそう？



研究目的

- 歩行者カメラに基づく歩行者位置推定
 - ロバスト性
 - 精度(希望は1m程度、できれば10cm)
 - リアルタイム
- 視覚障害者に対する歩行支援
 - Precision rateを重視 (Recall rateは低めでもよい)
 - 歩行計画 (Mappingは不要)

歩行者ナビゲーション問題の定義

- 視覚障害者の外出:歩行計画
 - 事前に経路を指定
 - 経路内のどこを歩くかは状況次第
 - 経路上の移動距離(現在地点)が歩行中での重要情報
- シーンに対する知識を事前に蓄積
 - 経路に沿って第三者が映像撮影
 - 位置情報と合わせてデータベースに記録

歩行者視点カメラ

- Computer Vision First
 - ・とはいかない
 - 頭部(帽子型) ×
 - メガネ型 ×
 - 白杖
 - 点字ディスプレイ
 - 手や腕が映像に入る
- 歩行進行方向
- 胸部



画像特徴量(SIFT)

- 輝度分布の幾何的な分布から自動的に決定
- 画像ごとに概ね数十～数百箇所検出
- 1箇所ごとの記述能力は非常に高い
 - 100万個～ほどあっても重ならない



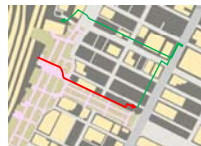
画像特徴量(SIFT)

- 輝度分布の幾何的な分布から自動的に決定



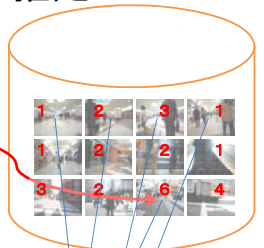
(1) 特徴量抽出

- 参照画像集合の1枚ずつで事前に画像特徴量(SIFT)を算出
- 入力画像ごとに画像特徴量を算出(1枚について数十～数百ms程度)



(2) 類似画像推定

- 画像特徴量(SIFT)は画像ごとに数百個
- ペア数第一位の画像を採用

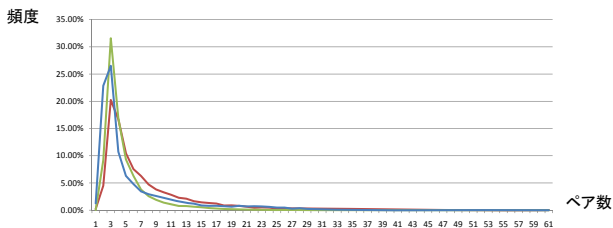


特徴記述の128次元ベクトルのみでペアを決定
KD木により高速化
(～1秒)



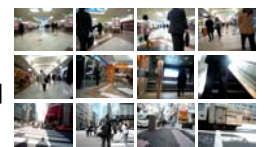
第一画像候補のペア数分布例

- 東京駅周辺経路(約10分)
 - 参照画像集合(1月7日撮影)
 - 赤:1月7日, 緑:11月20日, 青:1月7日(別時刻別カメラ)



実験

- 東京駅八重洲地下中央口
 - ⇒ 高島屋
 - ⇒ 東京駅ノースタワー入口
 - 約300mが地下街
 - 約600mが地上
 - 片側2車線道路(大通り)、片側1車線道路(狭い道)
 - 高層ビルが林立
 - 時間帯によって順光、逆光、日陰
 - 晴天・曇天
 - 人を避けずに歩くことは不可能



実験

- 左側: 入力画像(4/21に歩行)
- 右側: 参照画像集合(2/20と3/4撮影分)からの検索結果第一位



おわりに

- 歩行者視点カメラの画像による歩行者位置推定手法の提案
 - 参照画像集合に対する検索により実現
 - SIFT特徴量のペア数により検索
- 実際の環境で有用性を検証

謝辞

厚生労働省による「画像・GPS等のセンサ統合による日常利用可能な屋内外視覚障害者歩行支援システムの開発」(研究代表者: 静岡県立大学 石川准)

今後の課題

- 時間軸整合性の考慮
- Sensor Fusion
- 参照画像集合の改善
 - 類似フレームの自動排除
 - 複数映像系列の同時使用
 - マッチングにほとんど表れないSIFTキーの削除
 - 時間密度向上
- アフィン不変な局所特徴量の導入
- ペア数が多い場合のエピソード拘束導入
- 高速化・メモリ効率化
- Etc, etc ...