

マーカースレスで屋内外を問わず利用できるカメラ位置推定システム

ー 画像処理を利用した 3 次元ポジション(位置・姿勢)推定技術 ー

1. 背景

実世界の中でコンピュータや情報機器が人間の行動を支援するためには、コンピュータや情報機器が自分の位置を知る必要がある。この自分の位置を知ることは、位置推定技術(ポジショニング技術)と呼ばれる。位置推定技術としては、GPS に代表されるように数 m の誤差で数百 km 移動しても位置測定ができる技術から、マーカースを利用した画像処理により位置推定技術のように数 mm の誤差で、数十 cm 程度しか位置測定できない技術まで多々ある。しかし、屋内と屋外を行き来する範囲や、工場の建物内のサイズ程度の範囲である数十 cm から数十 m の測定距離をカバーする、屋内外をシームレスに移動可能な位置測定技術は、測定環境側に過大な設備投資を行う方法をのぞくと有力な方法がなかった。屋内外での利用を考えると、数 cm 程度の精度は必要であり、無線 LAN などの他の方法では不十分であった。そこで本プロジェクトでは、画像処理技術を利用して精度が保証できる仕組みを利用し、広範囲に展開可能な屋内外で利用できる位置推定技術の開発を目的とする。また、環境側へ赤外線センサー、RFID、マーカースなどを設置なしに位置測定ができることを条件とする。

2. 目的

本プロジェクトでは、画像処理技術を用いて、室内外を問わずマーカースを用いずに広範囲に対応したカメラ位置推定システムの開発を目的とする。具体的には入力情報として画像を入れ、出力情報として 3 次元の位置と姿勢を返すポジショニングシステムである。その前準備として、測定環境を複数枚で撮影し画像群として入力し、画像の特徴点、画像同士のマッチング、3 次元再構成をして、リファレンスデータベースとして保存し、位置測定のための基盤とさせる。

3. 開発の内容

本プロジェクトでは画像処理技術を利用して、マーカースレスで、画像を入力情報とし、3 次元の位置・姿勢を出力情報とする位置推定システムの開発を行った。画像を撮影した位置、つまり、カメラ位置を推定するシステムである。

カメラ位置推定システムとしては、2 つの部分から構成される。一つ目はマーカースレスで画像特徴点を用いて 3 次元情報に復元したジオメトリのデータベース(リファレンスデータベース)作成部分である。二つ目は、カメラ位置を知りたい画像を用意し、画像特徴点を作り、作成した特徴点データベースを検索し、パターンマッチングをすることで、3 次元でカメラの位置と姿勢を推定することを行う部分である。

アプローチのコンセプトとしては、丸覚えである。画像特徴点と 3 次元再構成したジオメトリ情報をリファレンスデータベースに現実世界の情報を覚えさせてしまい、次に問い合わせの来た画像の画像特徴点とマッチングを行うことでカメラ位置推定を効果的に行う、というものである。この画像特徴点を画像マーカースや画像同士のマッチングなどにしか利用

されていなかったところを、カメラ位置推定まで行おうという点において新規性がある。また、本アプローチを実現しようとする際には、マッチング等において組み合わせ爆発等の計算コストの問題が発生する。それに対応して並列化に対応した開発を行うことにも独自性がある。

システムの全体の概略を図 1 に示す。

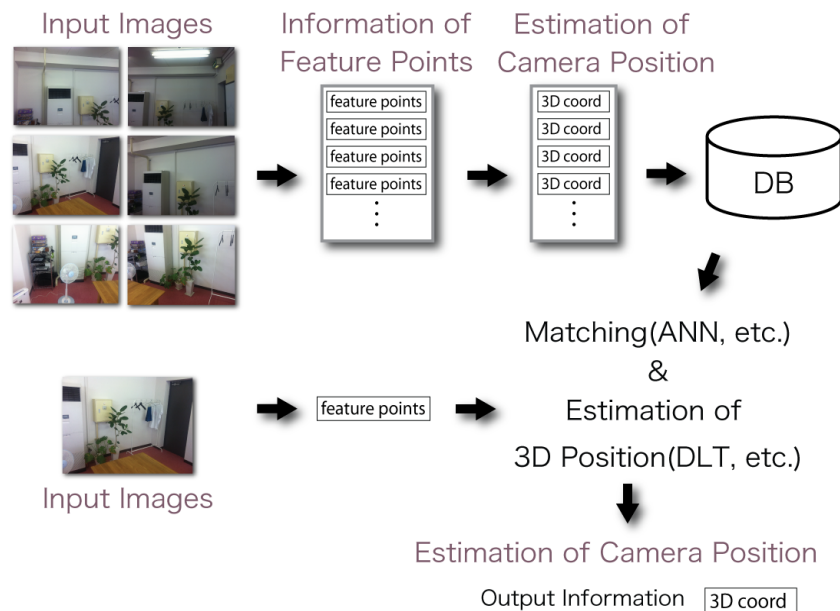


図 1. システム全体の概略図

アルゴリズムについて述べる。まず始めに入力画像それぞれに対して、特徴点を抽出する。特徴点の抽出は、精度の面を考え、SIFT を利用することとした。特徴点の抽出はパラメータに依存するが、数千個程度取得できる。

その次に、入力画像から任意の2枚の画像を選び、それぞれの画像の特徴点が同じ物理的場所をさしているかどうか、マッチングを行う。このマッチングの際には、数千個×数千個のマッチングを、画像数 n とした際、組み合わせ nC_2 回分、実行を行う。そのため、計算量は入力画像数 n のとき $O(n^2)$ となる。そのため、入力画像数が増えると組み合わせ爆発がおこるため、並列化は必須となる。

2つの画像のマッチング後、エピポーラ幾何学を用いて、画像を取得したカメラ間の関係を求めることができる。そのために、Homography 行列 (H : 平面射影変換行列) と基礎行列 (F) を計算で求める。そこから画像間の回転 R (Rotation)、並進 t (Translation) を行列計算から求める。この結果、2つのカメラ間の相対的位置関係を定めることができる。

この計算のみでは2枚の画像のマッチングにおいて様々なノイズが入っている。これらを、計算した行列を用いてフィルタリング処理を行い、ノイズを減らすことを行う。処理としては、カメラの撮影位置がある一定の距離以上離れていることを仮定してフィルタリングを行う threshold filter、エピポーラ幾何学による3次元の奥行き関係を計算して正しい奥行き関係のマッチングのみを抽出する depth filter、Homography 行列、基礎行列から計算される奥行き関係が合っているものを抽出する matrix filter の3種類のフィルターをかけることを行う。

その次に、2 枚の画像同士からマッチングした点同士を照らし合わせ、どのマッチングした点同士が同じ物理的場所を指しているか、さらにマッチングを行う。そうすることで、複数の画像に同じ物理的場所が映っている場合、それぞれの画像のどの点と同じ場所をさしているのか、ということが整理できる。本プロジェクトではこれを track と呼ぶ。複数の画像に同じ点が映っていることのつじつまを合わせて行くことによって、3 次元的关系を計算することができる。

最後にこれらの計算した関係から 3 次元ジオメトリをおこし、3 次元再構成を行う。2 枚の写真からは、ステレオ(立体視)の要領で 3 次元再構成を行うことができる。数珠つなぎに 3 次元再構成した結果に、3 次元再構成した結果を突き合わせ、相対座標系同士のため、スケールや回転、位置について、つじつまが合うように調整していくと、複数枚の 3 次元再構成が可能となって行く。本プロジェクトではこれを bundler と呼んでいる。

この 3 次元再構成した結果に対して、入力画像から特徴点抽出した結果などを計算し、近似最近傍探索(Approximate Nearest Neighbor Searching)などを用いてマッチングし、DLT(Direct Linear Transformation)法などを用いて、カメラ位置を推定し、その結果を位置推定結果として返す。

システムに関しては、マルチプロセスのシステムを実装し、メッセージキュー(RabbitMQ を用いて実装)を活用して、写真それぞれの特徴点の抽出部分、写真同士のマッチングからエピソード幾何学によりカメラ位置関係の推定計算部分、フィルタリング部分などを分散化することを行った。また計算結果については、状態をリレーショナルデータベースの MySQL に、結果などのデータを Key/Value ストア型のデータベースである Riak に格納するようにしている。

また、システムの計算の進行状況や計算結果を確認するために Web ベースのデータマネジメントツールを制作した。これにより、計算の結果の確認、入力写真と特徴点抽出の結果、マッチングの結果、フィルタリングの効果、カメラ同士の計算結果などの把握を容易にする仕組みを構築した。

4. 従来の技術(または機能)との相違

従来の位置測定技術においては 2 次元の地図上の座標を返すことが主な目的であったが、本プロジェクトの方法では 3 次元の位置と姿勢を返すことが目的である。

リファレンスデータベースを構築する際、複数の入力画像から 3 次元再構成を行う。この 3 次元再構成まではマイクロソフトリサーチ社の行っている研究と似ているが、位置情報を返す部分には新規性がある。

リファレンスデータベースを作成する際に、マッチングにより組み合わせ爆発が起こるため、マルチプロセスでの実装を行っており、この部分が従来のシステムとは異なる。マーカレス AR などではマッチングの対象が少ないためマルチプロセスなどの並列化をしなくても対処可能であるが、本プロジェクトのアプローチではマルチプロセスは必須である。

表 1. 位置測定技術の比較

Method of Positioning	Products	Environment setups	Both Indoor and Outdoor environments	Accuracy	Measurable Distance
Electric Wave	GPS	◎	×	×	◎
Wireless LAN	Place Engine (Koozyt, Inc.) etc	○	○	×	◎
Infrared (IR) Light	iGPS (Nikon Metrology NV), etc.	×	×	○	○
Visible Light	Sumitomo Mitsui Construction, etc.	×	○	○	○
RFID	NEC, Hitachi, etc.	×	○	○	×
Computer Vision with Makers	ARToolkit, etc	○	○	◎	×
Computer Vision without Markers	SmartAR (Sony Coop.), etc.	◎	○	◎	×
Proposed System		◎	○	◎	○

5. 期待される効果

アプリケーションの一つとして、産業用途が考えられる。具体的には工場内などの点検作業の応用を考える。例えば、ヘルメットにライトを設置するかのようにカメラを設置し、ある場所に来たら、チェック表とマニュアルが表示されるようにすることや、配管の工事履歴をその場で表示するようにするというアプリケーションを作成し、波及させる。

高機能のヘッドマウントディスプレイ(HMD)が普及すると、小型カメラと組み合わせることで、Augmented Reality（拡張現実感）技術の基盤となる。現実空間に対してコンピュータの情報を重ねて表示する際に、カメラから取得し推定したカメラ位置や姿勢が重要であり、本技術が基盤技術となる。

6. 普及（または活用）の見通し

屋内外をシームレスに利用でき、かつ、3次元の位置と姿勢がわかるポジショニング技術であるため、産業用途への普及が考えられる。普及や活用についてはクリエイターの設立した株式会社テンクーで実施することを考えている。

7. クリエータ名（所属）

西村 邦裕（株式会社テンクー）

坂田 里美（フリーランス）

（参考）関連 URL

プロジェクトサイト

<http://www.kuni24.net/mitoh/>

株式会社テンクー（Xcoo, Inc.）

<http://www.xcoo.jp/>