

3W-02 空間情報符合化のための多視点カメラによるオブジェクト認識方式

栗田 裕二 †* 國枝 和雄 †

†NECソフトウェア中部 *通信・放送機構 ぎふMVLリサーチセンター †NEC ヒューマンメディア研究所

1. はじめに

MPEG-4のオブジェクト符合化^[1]において、実空間の3次元情報を実時間で構造化する技術が求められており、それには実空間における対象物の位置や姿勢を高速に認識することが必要となる。3次元物体認識としては、対象物の見掛けの2次元画像を用いた固有空間法^[2]等が提案されているが、環境の変化に弱い点や、予め多方向の見掛け画像を生成する必要がある等、実用性の面では問題が残る。また、固有空間法では既知の物体を対象とし、位置と姿勢を同時に求められることを特徴とするが、MPEG-4のオブジェクト符合化による実空間伝送について考えると、背景実写とオブジェクト実写のスプライト処理の様に対象物の位置情報のみを必要とする場合や、対象物体が不特定で形状情報をもその場で獲得できなければ姿勢認識を行っても意味がない場合などが多く、固有空間法を用いるのは必ずしも適していないと言える。すなわち、オブジェクト符合化においては、まずは確実な位置認識を実現することが重要であり、次の段階として対象の形状が既知となる応用では、姿勢認識を導入するのが有効である。さらに、実写スプライト処理では、物体映像の切り出し処理が不可欠であり、物体認識法を考える上では、同処理との親和性についても考慮すべきである。

以上の観点から、我々は、対象映像切り出しをベースとし、環境変化に対して頑強な実時間3次元物体認識の実現を目標に研究を進めている。以下、本稿では、その基本部として開発した多視点カメラでの対象映像切り出しと、同映像を用いた位置認識方式について主に報告する。また、対象が既知の場合を想定した姿勢認識についても合わせて述べる。

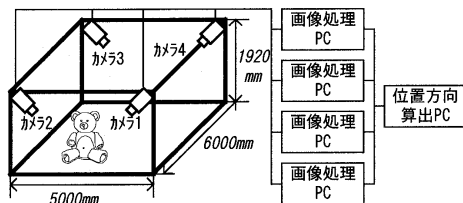


図1: 実験環境

A Method of Object Recognition by Multi Angle Cameras for 3D Space Encoding. KURITA Yuji †* and KUNIEDA Kazuo †
† NEC Software Chubu, Ltd.
* Telecommunications Advancement Organization of Japan
† NEC Corporation

なお、本研究は通信・放送機構ぎふMVLリサーチセンターにおける研究活動の一環である。

2. 多視点カメラシステム

2.1 特徴

本方式では、符合化の対象となる空間に複数のカメラを設置し、各方向からの映像について独立に対象物体を背景差分方式によって切り出し、各映像上での物体の座標を算出する。さらに、座標値と、各カメラの位置関係およびカメラパラメータを統合処理することによって実空間中における対象物体の位置を算出する。ここで、本方式の特徴は

- ・各映像毎に画像処理用ボードを用い、全体をリアルタイム処理。
- ・不特定な物体を対象とした位置認識を実現。
- ・各カメラ映像の処理結果に動的な重み付けを統合することによって認識精度を向上。

である。以下、各部の処理について説明する。

2.2 実験環境

空間符合化の対象である実験室とカメラの配置を図1に、機器構成を表1に示す。各カメラには番号を割り当て、カメラ映像の中央が実験室内床中央となるようカメラ角度を設定する。画像処理PCに実装されるIMAP-VISIONはSIMD型のプロセッサ(エレメント数:320)である。また、IMAP-VISIONの入力解像度は640×480×30fps(インタレース)であり、カメラ映像のフレームレートを変換後、入力する。

カメラ	フォトロン社製 PhotoCamRGB (640×480×60Hz プログレッシブ)
画像処理用 PC	NEC製 PC-98NX(Pen2 450MHz) NEC製 IMAP-VISION(320PE)
位置姿勢算出用 PC	Intergraph社製 TDZ2000GX1 (Pen2 400MHz × 2)

表1: 機器構成

2.3 認識方式

(1)位置情報の算出

各カメラ映像におけるオブジェクト領域切り出し、カメラ視点からオブジェクト重心へ向かうベクトル(直線)の算出、各カメラの重み付けの決定、を順次行うことによって位置情報を算出する。

オブジェクト領域切り出しは、橋本^[3]による背景差分方式を用いた。同方式では照明条件がある程度変化する

環境で任意物体の映像を抽出可能である。

次に、切出したオブジェクト領域の重心のスクリーン座標とカメラパラメータ(位置, 角度, 画角)から、透視法によってオブジェクト重心とカメラ視点を通る直線を求める。

さらに、得られた4つの直線の式から隣り合うカメラ同士の直線の交点(近接点)を4点求め、式(2)、(3)を用いて交点の重み付けを行う。ここで、式(2)、(3)は各カメラでの認識結果の誤差の傾向を分析することにより導出した。分析の過程では、対象物の実際の座標が各認識結果の矩形の内部に含まれることや、矩形の中では対象物から遠い位置にあるカメラの認識結果の誤差が少ない(対象物から遠いカメラの方が精度を得られやすい事にも合致)ことに着目した(図2)。また、各交点に2台のカメラの処理結果が反映されるため、4つの交点への正の重み付けは冗長で、対角をなす2点の組み合わせで十分あると考えられる。本実験ではカメラ配置が長方形であり、交点1,3と交点2,4の組み合わせでは有意な差が見られたため、より精度の得られる交点1,3の組み合わせについてのみ評価を行った。

最後に、位置情報は上記で求めた重み付けの値と各交点座標を式(1)に適用して求める。

$$(O_x, O_z) = \left(\sum_{i=1}^4 (W_{ix} \times P_{ix}), \sum_{i=1}^4 (W_{iz} \times P_{iz}) \right) \quad \dots(1)$$

但し、 (O_x, O_z) : オブジェクト重心のワールド座標
 (P_{ix}, P_{iz}) : i 番目の交点座標
 (W_{ix}, W_{iz}) : i 番目の交点座標に対する重み
 i 番目の交点: カメラ i とカメラ $i+1$ の交点

$$(W_{1x}, W_{2x}, W_{3x}, W_{4x}) = (\alpha, 0, (1-\alpha), 0) \quad \dots(2)$$

$$(W_{1z}, W_{2z}, W_{3z}, W_{4z}) = (\beta, 0, (1-\beta), 0) \quad \dots(3)$$

但し、 α, β は4点の平均座標を x 座標、 z 座標について算出し、認識対象空間での x および z の分割比によって決定する。

(2) 姿勢情報の算出

対象物が剛体かつ形状が既知であることを前提として姿勢検出についても評価を行った。まず、対象物上の特定の領域 3 箇所を予め特徴点として設定し、前述の対象画像切出しの際に特徴点抽出も同時に行い、その結果から姿勢情報(軸周辺の回転角)を算出した。なお、ここでのカメラ選択については、3つの特徴点が写っているカメラ1台ないし2台について、1台の場合はそ

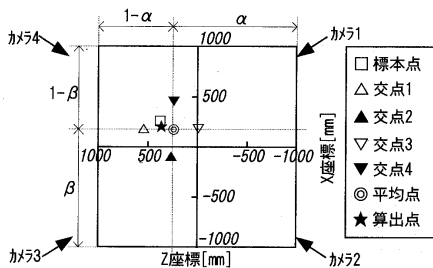


図2: 標準点に対する交点の分布

の値を用い、2台の場合は、オブジェクトの正面に近いカメラの方が精度が得られることに着目し、以下の手順で値を決定する。

- (1) 2台のカメラ映像から算出した姿勢の平均を求める(オブジェクトの仮法線ベクトルが判明)。
- (2) 仮法線ベクトルとカメラ向きとの角度差を2つのカメラについて求める。
- (3) 互いの角度差の差分が10度以内の時は、平均値を認識結果とし、それ以外の場合は角度差が小さい方の算出結果を認識結果とする。

3. 評価実験

対象空間の $2 \times 2\text{m}$ の範囲に15箇所の標本点をランダムに設定し、各点での位置および姿勢の認識結果と実際値との差を測定した。なお、対象物としては人形(480W×650H×310D(mm))を使用し、これを各点で床上に置いた(Y座標は一定)。また、姿勢変化はY軸周りの回転に限定した(回転量はランダム)。表2に測定結果を示す。ここで、位置認識については、前述の重み付け方式(方式1)、4つの交点座標を単純平均した場合(方式2)、仮平均として算出された点から最遠にあるカメラ2台の交点を認識結果とする場合(方式3)である。姿勢認識については、前述の選択方式(方式1)、平均値を用いた場合(方式2)である。

種別	位置算出誤差の 平均値	種別	姿勢算出誤差の 平均値
方式1	54.98(mm)	方式1	5.77(度)
方式2	92.30(mm)	方式2	6.41(度)
方式3	95.35(mm)		

表2: 測定結果

この結果から、いずれも提案方式(方式1)で誤差が最小となっており、その有効性が明らかとなった。

また、処理速度は1回の認識につき約100ミリ秒であり、10回/秒のサンプリングが可能であった。

4. おわりに

多視点カメラを用いたオブジェクト認識について報告した。重み付け処理を行うことによって認識精度を向上することができた。しかし、誤差の絶対的な大きさとしては十分とは言えず(位置認識の場合で $55\text{mm} = \text{約}13$ 画素)、カメラキャリブレーションによる精度向上やカメラ増設による効果についても検証したい。また、実際にスプライト処理を用いた実空間映像の伝送を行い評価を行う予定である。

参考文献

- [1] 三木: MPEG-4のすべて, 工業調査会 (1999)
- [2] 村瀬他: 2次元照合による3次元物体認識, 電子情報通信学会論文誌 D2, Vol. J77, No. 11, pp. 2179-2187 (1994)
- [3] 橋本他: 多地点接続遠隔会議システム「サイバーサークル」における背景分離処理, 信学技法 PRMU99-69, pp. 39-46 (1999)