距離アスペクト画像照合による高効率なロバスト探索

非会員 武口 智行* 正員 金子 俊一*

An Efficient Robust Search by Depth Aspect Image Matching

Tomoyuki Takeguchi*, Non-member, Shun'ichi Kaneko*, Member

A novel approach of object recognition is proposed, by which multiple three dimensional objects of range contours with occlusion and/or observation defects are efficiently and robustly recognized in cluttered scene. In our approach, in order to make the object recognition insensitive to peculiar features, such as edges and surface color, a uniform voxel framing has been utilized to define local coordination and to generate local depth structures on the objects, which are called *depth aspect images DAI*. In this paper, some techniques for efficient object recognition using DAI matching are designed. The first one is hashing of DAI images based on two dimensional hash table by use of local geometric key features, which enables fast search in the DAI database. The second one in recognition is random sampling for the voxel coordination, followed by parameter design with probabilistic modeling of sampling. By use of those two techniques, many redundant candidates for objects can be prevented from matching computation, realizing efficient and robust recognition of multiple objects in the complicated scene. The third one is introduction of support spaces for DAI calculation which cooperates with a fast statistical evaluation of the similarity. We demonstrate recognition examinations to clutter scenes with a large DAI database involving eight models and verify the effectiveness and efficiency of these techniques through analysis of recognition on 100 scenes.

キーワード：距離データ，物体認識，距離アスペクト画像，ランダムサンプリング，ハッシング

Keywords: Range data, Object recognition, Depth aspect image, Random sampling, Hashing

1. はじめに

本論文の目的は，距離アスペクト照合法にいくつかの効率化手法を新たに導入することによって，複雑情景内におけるロバストでかつ高効率な物体探索法を実現し，その実応用における可能性について検討することである。複雑情景内におけるロバストな物体認識手法は，ロボットなどの知的機械の視覚系を実現するために重要であり，FAにおけるマシンビジョン，情景理解，ピンピッキング，マンマシンインターフェース，視覚フィードバック制御などの高い付加価値を有する技術の基礎ともなる。これまでさまざまな物体認識の手法が提案されているが，それらは入力データにより大きく2つに分けることができる。まず，2次元画像のみに基づく手法は，従来からある多くの実用的な画像処理技術を利用することが可能であり，更に入力デ

バイスが安価であるため，従来から多くの手法が提案されている⁽⁹⁾⁽¹⁰⁾。一方，距離データや距離画像という3次元データに基づく認識手法が提案されている。3次元データをそのまま認識処理に用いるためには一般に高い処理コストを要求するので，それらの多くは，距離データをどのように認識に効果的かつ効率的処理が可能であるような形式に変換するかが焦点となっている⁽¹⁾⁻⁽⁷⁾。

そこで著者らは，2次元の濃淡画像である距離アスペクト画像 (Depth Aspect Image, 以下 DAI) と呼ぶ，DAI 作成用に設定した座標系に基づく3次元データの局所的な表現形式を提案した⁽⁶⁾⁽⁷⁾。また距離アスペクト画像を用いることにより，3次元物体探索問題を2次元画像の照合問題に帰着させ，DAI 照合により対象物の種別・位置・姿勢を精度よく高信頼で推定するモデル駆動認識法を提案してきた。そこでは，遮へいやデータ欠損に対する高いロバスト性を実現し，複雑情景における複数物体の認識を可能とした。また，距離アスペクト画像を作成する際の基準として，一様なボクセル構造に基づく手法を提案し，エッジなどの幾何的な特徴量を持たない対象物の認識をも可能としている。しかしながら実時間処理などの要求に応えるため

* 北海道大学大学院工学研究科

〒060-8628 札幌市北区北13条西8丁目

Faculty of Engineering, Hokkaido University

Kita 13, Nishi 8, Kita-ku, Sapporo, 060-0813 Japan

{take@mee, kaneko}@coin.eng.hokudai.ac.jp

には、高い認識率を保持しつつ更なる高速化或いは効率化を必要とする。

照合・認識における高速化の方法は、生産ラインへの導入、マルチメディアなどのリアルタイム処理など実応用を指向した場合、非常に重要な技術である⁽¹³⁾。照合・認識における高速化手法にはさまざまな種類がある。逐次残差比較法(SSDA)⁽¹⁴⁾やヒストグラムによる打ち切り法⁽¹⁵⁾などは、基本的に性能劣化がない優れた方法である。統計的推定を用いた計算打ち切りとしては、ランダムサンプリング⁽¹¹⁾があり、統計モデルにおいて信頼度などに基づく識別精度設計が合理的に行え、高速化が可能である。Multi-Resolution⁽¹⁷⁾、Scale Space⁽¹⁸⁾と呼ばれる粗照合による照合処理全体の効率化手法もあり、全探索を回避する有効な手段のひとつである。また一般的にモデルベースの認識手法の場合、リスト、ハッシュ、木などのデータ構造⁽¹⁶⁾を用いてモデルから作成されるデータベースを工夫することにより、認識段階で高い高速化効果が期待でき、物体認識への応用例⁽¹⁰⁾も多い。

そこで本研究では、これらの高速化手法を踏まえ、次の4つの効率化処理法をDAI照合法に導入した。(1)DAIデータベースのハッシング、(2)情景DAI作成におけるランダムサンプリングの導入、(3)DAI作成におけるボクセルを利用した高速化、(4)DAI照合処理の効率化である。(1)は、DAI照合においてデータベースの全探索を避ける方法である。(2)は、統計モデルを用いたランダムサンプリングによって、照合処理における全数検査を避けることができ、特に複雑で大きな情景に対して効果を発揮する。(3)は、DAI作成に用いるボクセルを利用し、幾何学的に冗長な処理を排除する。(4)は、ヒストグラム計算において最適解を得るために必要なしきい値を利用した高速化手法であり、(3)、(4)の手法は導入以前との結果に劣化がない高速化方法である。

さらに、これらの手法を導入した物体認識手法の有効性を確認するため、実データを用いた複数モデルによる実験を行い、本手法の有効性を示す。また全体としての高速化効果や、認識精度を検証するために、遮へいや複雑さの指標を与えた情景を100枚用意し実験を行うことで、認識率や高速化に関する定量的評価を行った。

2. 距離アスペクト画像

2.1 作成処理の概要

距離アスペクト画像作成処理を図1に示す。距離アスペクト画像は、距離データに与えたアスペクト座標系(Aspect Coordinate Frame, 以下ACF)によって定義される z 方向から観測した仮想的な正射影距離画像である。DAIは、(1)距離データのボクセル化、(2)ボクセルの3つ組選択、(3)3つ組によるACFの定義と距離データの座標変換、(4)距離データ点群の離散化と投影、の手順により作成される。これらの手法を順に説明する。

2.2 空間分割とボクセルデータ

距離データは測定

図1 距離アスペクト画像作成の流れ

Fig. 1. Procedure of making depth aspect image.

点群 $P = \{p\}$ と、それらを結ぶパッチ R とで表現される。パッチは単位法線ベクトル n_R をもつ。辺長 Δ_v の立方体ボクセルによって、距離データの空間を規則的に分割し、測定点群を内包する有効なボクセル群を固有の整数インデックスによって区別し $V = \{v = (\hat{x}, \hat{y}, \hat{z})\}$ で表す ($|V|$ はボクセル数)。各ボクセルは重心点 u 、代表法線 n (内包されるパッチの法線 n_R の合成単位ベクトル) をもつ。

2.3 3つ組ボクセルの選択 ACFを設定するための基準点組: 3つの u (各ボクセルに対応する) の組を選択する。可能な基準点組の全数は $O(|V|^3)$ と膨大になるため、次の5つの拘束条件を設定してこれを制限する。拘束条件の詳細は文献⁽⁷⁾に譲るが、これらに期待されるのは、まず(1)なるべく直交系に近い組合せとし、(2)近接している、(3)共線上に配列する、ノイズの影響を受けやすいボクセルの組合せを排除する効果である。次に全周囲データなどの閉じたデータをもつモデルを用いる場合、(4)同時に観測できない、(5)ACFの z 方向(DAIの仮想的な観測方向)が遠心的でない、組合せをも排除する。これらの条件により3つ組の数を有効に低減することができる。

2.4 アスペクト座標系 選択された3つ組ボクセルの各重心点を基準点組と呼び、図1に示すように、右手直交座標系であるACF $c = (o, e_x, e_y, e_z)$ 、及び基準点組によって張られる基準面(Base plane, 以下BP)を設定する。ACF原点 o に u_i 、 x 軸単位ベクトル e_x は原点と u_j を結び、 z 軸単位ベクトル e_z としてBPの単位法線ととり、これらに直交するように y 軸単位ベクトル e_y を算出する。

図 2 認識処理の概要

Fig. 2. Overview of recognition.

ACF c の単位ベクトル e_x, e_y, e_z は正規直交基底を構成し、この基底による点群の座標変換行列を $A(c) = [e_x, e_y, e_z]$ と表す。ACF c によって座標変換された点 p' は次式より求める。

$$p' = [e_x, e_y, e_z]^T [p - o] = A(c)^T [p - o] \cdots \cdots (1)$$

これにより点群 P は、ACF c で構成されるローカル座標系で表現された点群 P' に変換される。

2.5 距離アスペクト画像の作成 BP 上に格子間隔 Δ_x 及び Δ_y をもつ $L \times M$ 個のアスペクト格子 $A = \{a_{lm}\}$ を設定し、 A 上に投影される大きさ $L\Delta_x \times M\Delta_y$ の明度配列として DAI を算出する。このとき ACF 原点を中心として $\pm \frac{1}{2}L\Delta_x$ 及び $\pm \frac{1}{2}M\Delta_y$ の範囲を切り取る。DAI の明度 (z 方向) 分解能を Δ_z によって指定する。上記の範囲内にある P' の相当部分に対して、点群の z 座標値 (奥行き値) を $\bar{z} = \lceil z'/\Delta_z + 128 \rceil$ のように量子化する (単一格子内に複数点が投影されるときには $\max\{\bar{z}\}$ を選ぶ。[1, 255] を超えるときは $\bar{z} = 0$)。この処理によって BP 上の点 ($z' = 0$) を中間明度 128 とし、BP から $\pm 127 \times \Delta_z$ 以内にある点が投影され、BP 上に 1 枚の DAI が作成される。

このように DAI はモデルの 3 次元形状を可視化するものであり、基準点組の配置に依存した様々な物体姿勢において観測される距離画像を仮想的に表現している。

2.6 DAI を用いた認識処理の概要 DAI を用いた物体認識処理の概要を図 2 に示す。全体は登録処理と認識処理に分かれる。登録処理では、登録物体 (モデル) の距

離データをボクセル分割し、拘束条件を満たす全ての 3 つ組ボクセルから ACF を構成して各 ACF に従って DAI を作成するため、モデル距離データの様々な見え方を DAI で登録することになる。よって、データベースは各モデルの距離データ、ボクセル情報、ACF を含む DAI 群によって構成される。

認識処理では、情景内の距離データから同様に DAI を作成し、データベース内の登録 DAI と照合を行う。DAI 画像照合によって得られる最適照合が照合条件を満たす場合に認識候補が得られ、照合 DAI の ACF による座標変換から、情景における認識候補の位置姿勢が得られる。 c_r, c_s をそれぞれ照合されたデータベース DAI の ACF、情景 DAI の ACF とすると、認識候補モデルの情景内における位置・姿勢を表す点群 $P_r^s = \{p_r^s\}$ は、次式を用いて得られる。

$$p_r^s = A(c_s)A(c_r)^T [p_r - o_r] + o_s \cdots \cdots (2)$$

認識候補が見つかった段階で、情景中に投影された認識候補に対して、ICP アルゴリズム⁽⁸⁾ によって位置・姿勢の微調整を行った後、情景へのフィッティングに基づく検証を行う。検証条件 (近接性) を満たした認識候補が認識物体となり、認識物体近傍にある情景中のボクセルを以後の処理から除く。DAI 画像照合における照合条件或いは検証条件を満たさない場合は、別のボクセル組によって情景 DAI の再作成を行い、再度データベースとの照合を行う。複数物体を認識するため、認識段階ではこれらの処理を繰り返す。

3. DAI データベースのハッシング

登録段階では、各モデルから複数枚の DAI を作成し DAI データベースを構成する。このデータベースの第 1 の目的は DAI 画像照合を通じて登録モデル中から認識候補を推定することであるが、認識精度を保持しながらこの照合計算の回数を如何に低減するかが本節の課題である。そこで、DAI 照合を行う前に照合候補 DAI を絞り込むことを考える。図 3 は、この絞り込みを行うためのオープンハッシュ化されたデータベースを示す⁽¹⁶⁾。ハッシュキーとして重要な特性は計算が容易であることとその再現性である。そのために ACF と基準点組から計算できる次の 2 つの特徴量を整数ハッシュキーとする。

$$\text{key}_1 = \lceil |u_i - u_j| / \Delta_l \rceil \cdots \cdots (3)$$

$$\text{key}_2 = \lceil \cos^{-1}(n_i \cdot e_z) / \Delta_\theta \rceil \cdots \cdots (4)$$

key_1 は ACF の x 軸を構成するボクセル重心点間距離、 key_2 は原点ボクセル法線 n_i と、ACF の z 軸単位ベクトル e_z との成す角である。いずれもそれぞれ量子化幅 Δ_l, Δ_θ を用いて整数化されている。図 3 に示すように、登録段階では、すべての ACF に対応するハッシュキーを計算し、列幅 Δ_l 、行幅 Δ_θ で区分される 2 次元ハッシュテーブルに登録する。認識段階では、情景 ACF から算出されるハッシュキーによる表引きに基づいて DAI 照合すべき登録 DAI の範囲を

図3 DAIデータベースのハッシング

Fig. 3. Hashing of DAI database.

制限することにより、効率的な探索を可能としている。また選択されたボクセルによる登録モデルにおける測定点群の切り出しと、情景内からの切り出しにはずれがあるため、上に述べたハッシュキーは誤差を含んでいる。この影響を避けるために認識段階において、 $key_1 \pm 1$, $key_2 \pm 1$ に対応するハッシュキーに登録されているDAIとの照合を行う。データベースに登録されていないハッシュキーを持つACFはDAI画像照合以前に破棄することが可能となるため、不要な情景DAIの作成及び照合処理の手間をなくすることができる。

4. ボクセルのランダムサンプリング

4.1 目的と概要 2. において述べたように、認識段階においてはすべての可能なDAIを作成する必要はなく、DAIデータベースに登録されている1つのDAIを介して認識候補モデルにたどり着くことが期待されている。そこで、認識段階ではボクセル組の選択にランダムサンプリング⁽¹¹⁾を導入し、情景DAI作成の回数低減を目指す。

ランダムサンプリングは様々な場面で利用される効果的な手法であり、複数要素の組合せをサンプリングを通じてそろえる場合には、冗長な組合せをなるべく排除するための工夫が必要である。2.3 において述べたように、有効なACFを構成するためにはその基となるボクセル組はいくつかの拘束条件を満たす必要がある。従って、ここではACF原点に相当する単一のボクセルをランダムに情景内から選択し、その他の2ボクセルは拘束条件を満たすように設定する手法を採用する。これによって明らかに条件を満たさないACFを作成するオーバーヘッドを減らすことが期待できる。

4.2 サンプリング試行数の決定 ランダムサンプリングにおいて試行回数の見積りは基本的なものである。妥当なサンプリング試行回数を決定するために、認識処理における有効ボクセル選択率を $R_s \cdot R_v$ と定める。ここで、 R_s は物体占有率、 R_v は有効ボクセル率を表す。これらの具体的な算出方法は次節以下で説明するが、物体占有率は

情景距離データ内で1つの登録モデルが占める割合を示し、有効ボクセル率は登録段階および認識段階で別々に設定されるボクセル群において、十分に大きな重なりをもつボクセルを選択する確率とする。このとき、原点ボクセルをランダムに N 回選択するとき少なくとも1回は認識に有効なボクセルを選択する確率は、

$$\phi(N) = 1 - (1 - R_s \cdot R_v)^N \dots\dots\dots (5)$$

与えられる。すなわち、 $1 - R_s \cdot R_v$ が1回の試行において失敗する確率であり、その N 乗は N 回の試行において一度も成功しない確率を表す。従って、 R_t を利用者が設定する期待認識率とすると、試行数の期待下限値 N_o は次式によって決定できる。

$$N_o = \arg \min_N \{ \phi(N) \geq R_t \} \dots\dots\dots (6)$$

これによって、 R_s , R_v の値を与えることにより、ランダムサンプリングの試行回数を見積もることができる。

更にACFを構成するため、ランダムに選択した原点ボクセルを固定し、拘束条件を満たすように残りの2ボクセルを選択する。このときも、組合せの試行数を N_r に制限する。すなわち、原点ボクセルが有効であれば少なくとも N_r 回の試行によって認識が成功すると仮定している。以上、情景から作成されるDAIの数は最悪でも $N_o \cdot N_r$ となる。これは一見して、全登録モデル数 N_m に比例していないように思われ、不都合に見える。しかしながら、本論文においては3. で述べたようにDAIデータベース内を個々のモデルごとに区分せずハッシュキーのみによって整列しており、ランダムサンプリングにおいても、1つの3組ボクセルに対して、単一の登録モデルではなく N_m 個の全登録モデルを探索する方法となっている。従って計算時間は実質的には $N_m \cdot N_r \cdot N_o$ に比例することになる。

4.3 物体占有率 R_s 物体占有率 R_s は、情景全体の測定点群に対する登録モデルの点群が占める割合であり、一般にモデルが大きいと占有率が大きくなり、情景の撮像空間が広いと占有率は低くなる。本来は測定点数を比較して求めるべきではあるが計数の手間が大きいので、ここでは登録モデルを覆うボクセルと情景ボクセルを用いて物体占有率を算出する。登録モデルは全周囲データを用いるため、情景中に撮像される面積は多くても半分(センサに向けた半面)であると仮定し、物体占有率 R_s を次のように定める。

$$R_s = \frac{1}{2} \frac{\text{登録モデルのボクセル数}}{\text{情景のボクセル数}} \dots\dots\dots (7)$$

複数物体を登録する場合、登録モデルごとに物体占有率は異なるので、登録モデル中で最も小さい物体占有率を用いてランダムサンプリングの試行回数を見積もる。

4.4 有効ボクセル率 R_v 2.2 で説明したボクセル分割はいわゆるセンサ中心座標系で行われるため、登録モデルと情景に対する分割は一般に重複していない。この

図 4 試行回数と認識率

Fig. 4. Number of trials and recognition rate.

ような場合には、重なり大きいボクセルと小さいものが存在することになる。重なり大きいボクセルの切り取る物体曲面は類似するので、重心点の距離が近くなることが予想され、これらのボクセルを用いた ACF に基づく DAI は類似する。この重なり分布をモデル化し、認識時の有効ボクセル率 R_v の決定に用いる。

まず、単一ボクセル内の 2 点間距離の分布を求める。ボクセル内にランダムに置いた 2 点間の距離はボクセル幅に依存し、モンテカルロ法によるシミュレーションの結果では、平均 $\mu = 0.67\Delta_v$ 、標準偏差 $\sigma = 0.26\Delta_v$ の正規分布 $N(\mu, \sigma^2)$ に近似できることがわかった。一方、立方体ボクセルの重複度合いはそれらの位置と姿勢（回転）に依存するため複雑であるが、ここではボクセル中心間距離が 2 点間距離分布において十分に小さければ、照合可能な DAI が作成されると考えることとする。2 点間の距離が X_T 以下になる確率を有効ボクセル率 R_v とし、次の式から計算する。

$$R_v = 2 \int_{-\infty}^{X_T} \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} dx \dots (8)$$

ここで、ボクセルは密着して整列しているため、照合可能な DAI は折り返して出現することを考慮している。2 点間の距離 X_T は、DAI の量子化幅などにも依存することが予想され、これを厳密に求めることは難しいため、本論文においては実験的に適当な値を求めることとする。

4.5 有効ボクセル選択率の検証 モデルと情景の距離データを用いて、ランダムサンプリングに関する有効ボクセル選択率の検証を行う。モデルは $70 \times 80 \times 140\text{mm}$ 程度の大きさであり、予備実験において、認識が十分成功するボクセル幅を割り出し、その中から $\Delta_v = 15\text{mm}, 12\text{mm}, 9\text{mm}$ を用いて 3 種類の DAI データベースを作成した。各ボクセル幅において、それぞれ 181, 296, 500 個のボクセルが作成され、1217, 4067, 11601 枚の DAI によってデータベースが構成されている。各ボクセル幅でモデルの初期位置・姿勢を変更してそれぞれ 100 回の認識実験を行い、認識されるまでのランダムサンプリング試行回数を調べた。実験に用いた登録モデルと情景距離データ、及び結果を図 4 に示す。破線で示した近似曲線は、各ボクセル幅に対する $\phi(N)$ であり、 X_T を $2.4 \sim 3.4\text{mm}$ の間で近似精度の高

図 5 ボクセル幅による処理時間の比較

Fig. 5. Processing time with some voxel widths.

い値を選んでいる。各ボクセル幅による試行回数と認識率の関係が有効ボクセル選択率によってよく近似できていることがわかる。 X_T は DAI の xy 方向離散化幅 3mm と同程度の値になっている。ボクセル幅が小さいほど試行回数は少なくなるため認識処理を早く打ち切ることができる。一方、ボクセル幅が小さいほど登録 DAI 数はボクセル数の 3 乗に比例して増加するため、情景 DAI 1 枚あたりの処理時間が大きくなる。

モデルの発見までにかかった実際の処理時間を図 5 に示す。ボクセル幅 12mm において、モデル発見までの実際の処理時間が最も少なくなっていることがわかる。ボクセル幅によるデータベースの大きさと、有効ボクセル選択までの試行回数はトレードオフの関係があり、データベースが大きすぎず、試行回数になるべく少ないボクセル幅を選択することが望ましい。一般的にはモデルとする物体の大きさに依存して適当なボクセル幅を決定できるが、より効率的なデータベースを作成するために、実際の認識実験においては、予備実験によりボクセル幅を決めることとする。

5. DAI 作成・照合の高速化

本章では、提案手法において最も基本的な処理である DAI の作成および照合計算における高速化について述べる。

5.1 DAI 作成の高速化 測定点群を ACF 座標系に変換し DAI を作成する際、ACF 座標系で原点 o を中心とした直方体空間内の点群が DAI 作成に使われる。図 6 に示すこの空間を DAI サポート空間と呼ぶ。サポート空間外にある点群は DAI 作成に使われないため、DAI サポート空間に関する内外判定により、無駄な座標変換計算を省くことができる。

ボクセルインデックスを用いて内外判定を行う。 v_o を原点ボクセル、 v を判定対象ボクセルとすると、ACF に座標変換され Δ_v により実寸化されたインデックスベクトル $A(c)^T[v - v_o] \times \Delta_v$ がサポート空間に含まれるか否かによって内外判定を行う。なお、図 6 に示すように、単一ボクセルの容積を考慮したボクセル幅によるサポート空間の補正が必要である。すなわち、DAI サポート空間を各方向 $\sqrt{3}\Delta_v$ だけ広げた空間外に判定ベクトルがあれば、そのボクセルは DAI サポート空間内に入ることではない。この内外

図 6 DAI サポート空間と内外判定

Fig. 6. DAI support space and inside-outside check.

図 7 DAI 作成時間高速化の効果

Fig. 7. Efficiency of making DAI with support space check.

判定による高速化処理は撮像した情景の空間が広いほど効果が高い。

$\Delta_v = 15\text{mm}$ とし, 20 枚の情景を用いて DAI 作成時間を比較することにより内外判定の効果を調べた。実験結果を図 7 に示す。内外判定を導入することによって, DAI 作成処理は最大 2 倍程度高速化される。情景の大きさを近似的に示すボクセル数と比較してみると, 高速化の割合がボクセル数の増加につれて増大する傾向にあることがわかった。

5.2 DAI 照合の高速化 DAI 画像照合と認識候補モデルの検証には, Least Quantile of Residuals (以下, LQR) と呼ぶ統計的推定手法を用いている。これは, 大きさに順にソートされた全数 N の残差値の小さいほうから $Q \cdot N$ 番目の値を評価基準にするものであり, $Q = 0.5$ とすればよく知られている最小メディアン評価⁽¹²⁾ となる。大きな例外値を含む評価対象データに対してもロバストな照合が期待でき, さらに評価率 Q を可変にすることにより, 評価対象データの特性を考慮できる, Q を高くすることにより, メディアン評価よりも信頼性のある評価を下すことができる, などの利点がある。本論文の DAI 画像照合においては, 互いの明度 0 の画素を除いた明度残差列を評価対象データ

図 8 しきい値に基づく LQR 計算

Fig. 8. LQR evaluation based on threshold.

とする。そのヒストグラム $H = (h_0, h_1, \dots, h_{|H|-1})$ ($|H|$ は階級数) を用いて, 評価率 $Q (0 \leq Q \leq 1)$ における H の評価量 $f_Q(H)$ は次式で表す。

$$f_Q(H) = \arg \min_q \left\{ \frac{\sum_{i=0}^q h_i}{\sum_{i=0}^{|H|-1} h_i} \geq Q \right\} \dots\dots\dots (9)$$

情景 DAI と登録 DAI について $f_Q(H)$ を評価し最小値を求める。その値がしきい値 F_T より小さくなる DAI データベース内の DAI を探索する。ハッシュ表においてヒットするすべての登録 DAI との評価値がしきい値より大きい場合, 情景内から別の ACF を選択して DAI を更新し, ふたたび照合を行うため, 評価量 $f_Q(H)$ がしきい値 F_T 以上の場合にのみ, 正確な評価量が得られれば十分である。よって, 図 8 に示すように, LQR の計算は q の値を増加させ, しきい値を越えた段階で処理を打ち切ることができるため, 無駄な計算を省くことが可能である。

認識候補の検証では, 情景に投影した認識候補の各点から情景点群中の最近傍点への最小距離を評価データ列としてヒストグラムを作成し, LQR によって 2 つの点群の重なり程度を見る。DAI 照合時の LQR 評価率を Q_1 とし, 検証段階の LQR 評価率を Q_2 として区別する。このヒストグラムと評価率 Q_2 を用いた LQR による評価量 $f_{Q_2}(H)$ がしきい値 D_T 以下となる場合, 認識候補は情景の対応する部分に十分合っていると判断する。よって DAI 照合時と同様, しきい値 D_T による LQR の打ち切り計算を用いることができる。本研究ではモデルに統合距離データを用い, 情景データは 1 方向から観測した距離データを用いるため, モデル距離データの点群は最大で約半分しか情景距離データにフィットしないため, 検証段階の LQR 評価率 Q_2 は 0.5 以下に設定する。

DAI 照合を行うための残差ヒストグラム作成は, マスク処理, 残差計算などの単純な計算で出来るため, ハードウェアなどによる高速化も可能である。実験システムではマルチメディア命令 (Intel MMX 命令) を利用することにより, DAI 評価ヒストグラム作成関数の高速化を実現している。

LQR の打ち切り処理とマルチメディア処理による高速化により従来に比較して約 3 倍の高速化が得られた。

図 9 登録モデル

Fig. 9. Models.

表 1 DAI 作成時の各パラメータ

Table 1. Parameters for DAI.

Voxel width	Δ_v	12mm
Pixel size of DAI	$\Delta_x \times \Delta_y$	3×3 mm
Resolution of z	Δ_z	0.3 mm
Pixel number of DAI	$L \times M$	64×64

表 2 モデルの Voxel 数と DAI 枚数

Table 2. Number of voxel and DAI for each model.

Model	a	b	c	d	e	f	g	h
Voxel	296	342	286	306	335	268	315	299
DAI	4067	5254	3421	4388	5961	3620	3068	3380

6. 実 験

複数の物体を登録し、それらを含む複雑情景における認識実験を行った。すべての実験は Pentium4, 2.0GHz の CPU を搭載した PC を用いて行った。

6.1 登録モデルと DAI データベース 8 つの登録モデルを図 9 に示す。それぞれの登録モデルは, LRF (ミノルタ製. VIVID300) からの測定距離を 600mm 程度にして対象物体の周囲 45°毎に得た 8 枚の点群データを, VIVID 付属のポリゴン編集ツールによって位置合わせ及び統合処理を行った全周囲データから構成される。穴や不良パッチの排除は同ツールを用いて手作業で行った。底部或いは凹部の点群データは補完していない。DAI データベースの作成に必要なパラメータを決定する際は, モデルの大きさ, レンジセンサの分解能を考慮して定める。各パラメータを表 1 のように与えた。各モデル距離データから登録 DAI を作成し, データベースを構成する。それぞれのボクセル数, 及び作成された DAI 数などを表 2 に示す。33159 枚の全登録 DAI をハッシュテーブルに保存する。ハッシュテーブルの幅は $\Delta_l = 5\text{mm}$, $\Delta_\theta = 5^\circ$ としている。2 次元ハッシュテーブル中の最大登録 DAI 数は key(6,6) で 842 枚である。この前後 ± 1 を含めた登録 DAI 数は 6240 枚であり, この

表 3 認識時の各パラメータ

Table 3. Parameters in recognition.

Effective voxel rate	R_v	0.108 ($X_T=3$ mm)
Expected recognition rate	R_t	0.99
Number of DAI at an origin voxel	N_r	12
LQR rate in DAI matching	Q_1	0.6
LQR threshold in DAI matching	F_T	2.1 mm
LQR rate in verification	Q_2	0.25
LQR threshold in verification	D_T	3.5 mm

表 4 各情景の認識結果

Table 4. Recognition result of each scene.

Scene	#1	#2	#3	#4
Recognition object	abcdefgh	cdefgh	bcdefgh	acdegh
Voxel	1201	1756	1622	1230
Sampling num. N_o	380	557	514	390
Scene DAI	1837	4370	3103	2405
DAI matching success	12	13	58	7
Time (sec.)	44.7	116.3	181.1	51.2

データベースにおいて最大値であるので, このデータベースを用いた認識段階における DAI 照合は, 最大でも登録数の約 20% を対象とすればよいこととなった。

6.2 認識実験 認識時の照合パラメータを表 3 に示す。情景におけるボクセル幅や情景 DAI の仕様は登録段階のものと同じである。認識実験の結果を図 10 に示す。認識実験に用いた情景と認識結果を図 10 に示す。図中で Scene は情景をパッチで表現したもので, Result は情景をワイヤフレームにし, 認識したモデルを情景中に投影してパッチで表現したものである。各情景には, 遮へいによる点の欠損が多く発生しているため, 情景中のすべての物体が認識できるわけではない。また各情景は認識処理時において無駄な照合時間を増大させ, 偽照合を起こす危険を増加させる登録物体以外の物体も同時に撮影している。このような条件の下でも本手法は的確に登録モデルを検出していることがわかる。各情景における認識結果を表 4 にまとめる。

ランダムサンプリングにより, 情景 DAI 作成の対象として選ぶ原点ボクセル数は約 $\frac{1}{3}$ となる。1 モデルあたりの探索時間は, 数秒から 26 秒程度であり, モデルあたりの処理速度が情景データサイズによって大きく異なる。情景 DAI 1 枚を作成し, DAI データベースとの照合を行うためにかかる処理時間は約 0.02 秒であり, DAI 照合が成功して得られる認識候補モデル 1 個に対する ICP アルゴリズムの適用と検証処理にかかる時間は約 1.5 秒である。このため, 同程度の DAI を作成した情景において, DAI 照合の成功回数により大きく処理時間が異なることとなる。ICP を含めた検証処理が情景によって何回発生するかは不明であるため, 認識処理以前に正確な処理時間を予想するのは困難であるが, DAI 照合における偽照合 (検証処理で棄却されるような照合) の回数を減少させることにより, 検証処理の回数も減らすことができるので, さらなる高速化が

図 10 実験結果の例

Fig. 10. Examples of recognition.

期待できる。また検証処理が DAI 照合処理よりも非常に大きな時間を要しているため、検証処理自体の高速化も全体の速度向上に大きく寄与できると考えられる。これらを実現する具体的な手法については現在検討中であり今後の課題とする。

6・3 効率化手法の効果の検証 効率化手法の効果

及び遮へいや情景の複雑さに対する定量的評価をするための実験を行った。旧手法との比較を行うために、旧手法の実験で用いたボクセル幅 $\Delta_v = 15\text{mm}$ を用いた。その他認識に関するすべてのパラメタも、旧手法で用いた値と一致させた⁽⁷⁾。ランダムサンプリングのパラメタはボクセル幅にあわせ $R_v = 0.072$ とした。

図 11 100 情景の遮へいに対する認識率

Fig. 11. Recognition rate versus occlusion for 100 scenes.

モデル (猫) は a~d の 4 個とし、情景データを 100 枚用意した⁽⁷⁾。それぞれ 2~4 個の登録物体が存在する。また多くの情景には床や他の物体 (未登録) を同時に含んでいる。

各情景には遮へい率 (occlusion) と情景の複雑さ (clutter) を評価する値が付加されている。登録モデルに対応する領域を情景中から手作業により抜き出し、登録モデルのパッチ面積と情景中の対応領域のパッチ面積との面積比を用いて、遮へい率を $1 - \text{パッチの面積比}$ として計算した。登録モデルは全周囲データであり、情景中には任意の側面しか写らないため、最小でも約 50% となる。一般に遮へい率が高いほど、情景中の登録モデルに対応する曲面が少ないため認識が困難となる。

情景の複雑さは、登録モデル以外の測定点群データの含有率を表し、情景内の測定点数と全登録モデルに対応する点数の比を用いて、 $1 - \text{点数比}$ として計算した。複雑さ 0% の場合、情景中の点群はすべて登録モデルに対応する。一般に複雑さが高いほど登録モデル以外の曲面データが増加するため、処理速度の低下、DAI 照合の失敗や誤認識を引き起こす可能性が高まる。

情景はモデルと異なりデータには手を加えていないため、穴や不良パッチなどの不良条件はそのまま存在する。これらの含有量を直接数えてはいないが、情景中の認識したい物体上の穴を隠れ (occlusion) とし、不良パッチを登録モデル以外の測定点群データ (clutter) として考えることができる。

100 枚の情景データにおける遮へいに対する認識率を図 11 に示す。約 65% までの隠れに対しては、非常に良い認識率を維持できる。また、各情景の複雑さを 10% ごとの階級に分けた場合、情景の複雑さに対する認識率は $87 \pm 5\%$ であった。この実験から、隠れが大きくなると認識率が低下し最終的には認識不可能となるが、情景の複雑さに対しては、認識率に大きな変化がないことがわかる。100 情景中には総計 356 個の登録モデルが含まれ、正しく認識された登録モデルは 304 個 (true positive)、検出されなかった登録モデルは 52 個であり (false negative)、誤認識 (false positive) は 0 回であった。

図 12 情景のパッチ面積に対する処理時間

Fig. 12. Processing time for patch area.

図 13 情景の複雑さに対する処理時間比

Fig. 13. Efficiency for clutter rate.

検証段階で用いるしきい値 Q_2 は遮へい率約 75% までの物体を認識できるように設定されている。100 情景中で遮へい率 75% 未満の物体は 328 個であり、そのうち認識が成功したのは 301 個である。従って、認識の成功が期待される情景中の物体に関する認識率は 91.8% となる。一方、旧手法においては 310 個が認識成功し、認識率は 94.5% であるため認識率の低下は 2.7% であった。提案する効率化を用いても誤認識が増加することなく、認識率の低下を低く抑えることができた。

次に、図 12 に情景のパッチ総面積による 5 階級に対する平均処理時間を示す。平行投影型 LRF の場合は測定点数が撮像空間の大きさを表すが、実験に用いた LRF は放射投影型のため、パッチ総面積が撮像した情景の大きさを表す。情景の大きさにつれて処理時間がほぼ線形に増えることがわかった。100 情景中で、4 個の登録モデルと登録以外の物体を多く含む (複雑さ 47.2%) パッチ総面積が最も大きい情景では、約 40 秒の処理時間を要した。

次に、図 13 に情景の複雑さによる 6 階級に対する平均処理時間比 (効率化率) を示す。効率化の効果は情景の複雑さが増えるにつれて増加している。約 8 倍 (複雑さは 0% に近い) から約 23 倍 (複雑さは 50% 以上) の効率化が得られた。この傾向は次のように説明される。複雑さが小さいときには情景中のかなり点が登録モデルの一部であると考えられ、認識が成功した段階で認識モデル近傍のボクセルを以後の処理から除くことができるため、認識が進むにつれ最終的に処理対象となるボクセルがほとんどなくなり、処理を打ち切ることができる。すなわち、このような場合はランダムサンプリングにおける処理の削減効果が小さい。一方、情景の複雑さが増すときには、ランダムサンプリン

グによる効果で，全探索と比べて登録モデルに対応しないボクセルを用いた無駄な情景 DAI の作成回数が制限されるため，効率化効果が相対的に大きくなる。ちなみに，100 情景中で最も効率化効果が大きい情景における複雑さは約 58% で，認識時間は約 33 秒（登録モデルあたり約 8 秒）であり，旧手法と比べて約 45 倍の効率化効果が得られた。このように，本論文で提案する効率化手法は情景の複雑さが増してくる場合に，より効果的であることがわかった。

7. む す び

DAI 照合による物体認識手法における効率探索手法を新たに提案した。提案手法は DAI データベースのハッシング，ランダムサンプリング，照合のための基本特徴量である DAI の作成手順の高速化，その照合評価における打ち切り処理，によって構成される。遮へい率及び複雑さの定量化を可能とするデータベースを作成し，それらに対する認識率を統計的に評価することを可能とした。それを用いて実物体を測定した距離データに対する認識実験を行った。結果として高い認識率を保持しながら平均で約 15 倍の効率化を実現することができた。情景内に未登録物体が多く写り込むような状況においてより高い効率化効果が得られることが実験的に示された。このことから，提案した効率化手法を含む認識法が未知物体を多く含むような複雑な実情景を扱う場合に効果的に利用可能である。

提案手法に関する今後と課題として，全周囲データ以外の局所的登録モデルを用いるための改良，物体表面の反射率データ（色や濃淡）を融合することによる更なる高信頼化などを予定している。

（平成 15 年 7 月 1 日受付，平成 15 年 10 月 15 日再受付）

文 献

- (1) G. Hetzel, B. Leibe, P. Levi, and B. Schiele: "3D Object Recognition from Range Images using Local Feature Histograms", *Proc. CVPR*, II, pp.394-399 (2001)
- (2) F. Stein and G. Medioni: "Structural indexing: Efficient 3-D object recognition", *IEEE Trans. on PAMI*, Vol.14, No.2, pp.125-145 (1992)
- (3) C.S. Chua and R. Jarvis: "Point Signatures: A New Representation for 3D Object Recognition", *International Journal of Computer Vision*, Vol.25, No.1, pp.63-85 (1997)
- (4) A.E. Johnson and M. Hebert: "Using Spin Images for Efficient Object Recognition in Cluttered 3D Scenes", *IEEE Trans. on PAMI*, Vol.21, No.5, pp.433-449 (1999)
- (5) S. Ruiz-Correa, L.G. Shapiro, and M. Melia: "A New Signature-Based Method for Efficient 3-D Object Recognition", *Proc. CVPR*, I, pp.769-776 (2001)
- (6) T. Takeguchi, S. Kaneko, T. Kondo, and S. Igarashi: "Robust Object Recognition based on Depth Aspect Image Matching", *IEICE Japan*, Vol.J-84-D-II, No.8, pp.1710-1721 (2001-8) (in Japanese)
武口智行・金子俊一・近藤 司・五十嵐 悟:「距離アスペクト画像の 2 次元照合によるロバスト物体認識」, 信学論, J-84-D-II, 8, pp.1710-1721 (2001-8)
- (7) T. Takeguchi and S. Kaneko: "Robust Object Recognition through Depth Aspect Image by Regular Voxels", *IEICE Japan*, Vol.J-86-D-II, No.9, pp.1308-1319 (2003-9) (in Japanese)
武口智行・金子俊一:「規則的ボクセル基準による距離アスペクト画像

を用いたロバスト物体認識」, 信学論, J-86-D-II, 9, pp.1308-1319 (2003)

- (8) P.J. Besl and N.D. McKay: "A Method for Registration of 3-D Shapes", *IEEE Trans. on PAMI*, Vol.14, No.2 (1992)
- (9) H. Murase and S.K. Nayar: "3D Object Recognition from Appearance - Parametric Eigenspace Method - ", *IEICE Japan*, J77-D-II, No.11, pp.2179-2187 (1994-11) (in Japanese)
村瀬 洋・シュリーナイヤー:「2 次元照合による 3 次元物体認識 - パラメトリック固有空間法 - 」, 信学論, J77-D-II, 11, pp.2179-2187 (1994-11)
- (10) Y. Lamdan and H.J. Wolfson: "Geometric Hashing: A General and Efficient Model-Based Recognition Scheme", *Proc. ICCV*, pp.238-249 (1988)
- (11) M.A. Fishler, R.C. Bolles: "Random Sampling Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography", *Communications of the ACM*, Vol.24, No.6, pp.381-395 (1981)
- (12) K. Inui, S. Kaneko, and S. Igarashi: "Robust Line Fitting Using LMedS Clustering", *IEICE Japan*, Vol.84-D-II, No.1, pp.48-55 (2001-1) (in Japanese)
乾健太郎・金子俊一・五十嵐 悟:「LMedS クラスタリングによるロバスト直線あてはめ」, 信学論, 84-D-II, 1, pp.48-55 (2001-1)
- (13) S. Hata: "Trends of Automatic Vision Inspection", *JSPE*, Vol.56, No.8, pp.1366-1375 (1990-8) (in Japanese)
秦 清治:「外観検査の自動化の現状と展望」, 精密工学誌, 56, 8, pp.1366-1375 (1990-8)
- (14) D.I. Barnea and H.F. Silverman: "A Class of Algorithms for Fast Digital Image Registration", *IEEE Trans. Computer*, Vol.21, No.2, pp.170-186 (1972)
- (15) H. Murase and V.V. Vinod: "Fast Visual Search Using Focused Color Matching - Active Search - ", *IEICE Japan*, J81-D-II, No.9, pp.2035-2042 (1998-9) (in Japanese)
村瀬 洋・V.V. Vinod:「局所色情報を用いた高速物体探索 - アクティブ探索法 - 」, 信学論, J81-D-II, 9, pp.2035-2042 (1998-9)
- (16) A.V. エイホ他・大野義夫訳:「データ構造とアルゴリズム」, 培風館 (1987)
- (17) G.S. van der Wal and P.J. Burt: "A VLSI Pyramid Chip for Multi-resolution Image Analysis", *IJCV*, Vol.8, No.3, pp.177-189 (1992)
- (18) J. Babaud, A.P. Witkin, M. Baudin, and R.O. Duda: "Uniqueness of the Gaussian Kernel for Scale-Space Filtering", *IEEE Trans. on PAMI*, Vol.8, No.1 (1986)

武 口 智 行（非会員）平 11 年北海道大学・工・システム工学科卒業。平 13 同大・工・博士前期過程（システム情報工学専攻）修了。現在，同大・工・博士後期過程在学中。コンピュータビジョン及びマシンビジョンの研究に従事。特にテンプレートマッチングや物体認識手法に興味を持つ。電子情報通信学会学生員。

金 子 俊 一（正員）昭 53 年北大・工・精密工学科卒。昭 55 年同大学院修士課程了（情報工学）。同年東京農工大学・工・数理情報工学科助手，同大学院生物システム応用科学研究科助教授を経て，平 8 年北大・大学院・システム情報工学専攻助教授。工博。画像理解・計測，ロバストパターン認識の研究に従事。平 2 年度精密工学会論文賞，平 10 年度精密工学会賞，平 12 年度小田原論文賞（精密工学会外観検査の自動化 WS）受賞。精密工学会，電子情報通信学会，IEEE 各会員。