

インデックス画像を用いた複数距離画像の高速同時位置合わせ

大石岳史[†] 中澤篤志[‡] 池内克史[†]

[†] 東京大学生産技術研究所 〒153-8505 東京都目黒区駒場 4-6-1

[‡] 大阪大学サイバーメディアセンター 〒567-0047 大阪府茨木市美穂ヶ丘 5-1

E-mail: [†] {oishi, ki}@cvl.iis.u-tokyo.ac.jp, [‡] nakazawa@ime.cmc.osaka-u.ac.jp

あらまし 本論文では複数方向から測定された多数の距離画像を高速に同時位置合わせする手法を提案する。位置合わせ計算の中で最も計算コストが高いのは対応点探索である。そこで我々は、インデックス画像を用いて頂点数 n に対して計算量 $O(n)$ で対応点探索を行う手法を開発した。この手法はグラフィックスハードウェアを利用することで、より高速な計算が可能となる。また、少ないメモリ使用量で位置合わせを行う事ができる。変換行列の計算には線形化した誤差評価式を用い、不完全コレスキー分解共役勾配法 (ICCG) を適用する事で、多数の距離画像に対しても短時間で計算させる事が可能となった。実験では、レーザレンジセンサによって測定された複数の距離画像を位置合わせする事で本手法の有効性を示した。

キーワード 距離画像, 位置合わせ, 高速化, インデックス画像

Fast Simultaneous Alignment of Multiple Range Images Using Index Images

Takeshi Oishi[†] Atsushi Nakazawa[‡] Katsushi Ikeuchi[†]

[†] Institute of Industrial Science, The University of Tokyo 4-6-1 Komaba, Meguro-ku, Tokyo, 153-8505 Japan

[‡] Cybermedia Center, Osaka University 5-1 Mihogaoka, Ibaraki, Osaka, 567-0047 Japan

E-mail: [†] {oishi, ki}@cvl.iis.u-tokyo.ac.jp, [‡] nakazawa@ime.cmc.osaka-u.ac.jp

Abstract This paper describes a fast simultaneous alignment method of multiple range images. The most time consuming part in alignment process is searching correspondences. Therefore, we developed the method that performs searching correspondences process in complexity $O(n)$, where n is the number of vertices, by using index images. This method is accelerated by graphics hardware, and can reduce memory space used for computation. By applying linear least square method and incomplete Cholesky conjugate gradient (ICCG) method in order to obtain the rigid transformation matrix, computation time of aligning a large number of range images can be shortened. The experimental results of aligning multiple range images measured with laser range sensor show the effectiveness of our method.

Keyword Range Image, Alignment, Acceleration, Index Image

1. はじめに

近年,レーザレンジセンサと Modeling from reality[1] 技術の発達により,文化遺産や美術品の形状を3次元モデルとして保存する試みが世界各地で行われている[2,3]。これらの計測に用いられるレーザレンジセンサの測定範囲は可視領域に限られるため,一回の測定で物体全体をモデル化することはできない。そこで,完全な全周3次元モデルを得るためには複数方向から多数の距離画像を取得しなければならない。この複数方向から取得された距離画像は,それぞれの測定位置に対する異なった座標系で記述されているため,各々の相対位置姿勢を求め座標系を統一する位置合わせ処理が必要となる。

この距離画像の位置合わせ問題を解決するために,

これまでに多くの手法が提案されている。特によく知られている手法は Besl が提案した ICP (Iterative Closest Point)法[4]や Chen の手法[5]である。ICP では2つの距離画像中の最近傍点を対応点として,この対応点間距離を最小化するような各距離画像の変換行列を求め,繰返し計算によって相対位置を求めていく。一方,Chen の手法では面の法線を計算し,点と面の距離を最小化するように相対姿勢を求めていく。この他にも視線方向に点を投影して他方の距離画像との対応点を求める方法などもある[6]。また ICP は誤対応やノイズによる影響を受け易いため,ランダムサンプリングと LMedS (Least Median Squares Estimation)法によってロバスト性を高める手法も提案されている[7]。

これらの位置合わせ手法は2枚の距離画像の相対位

置姿勢を求めるものであるが、距離画像枚数が多い場合はこれを順次繰り返していきと誤差の蓄積が問題となる。これを避けるためには、全ての距離画像間の相対位置姿勢を同時に推定する同時位置合わせ手法が有効である。Neugebauer は視線方向に対応点を探索し、全ての距離画像の組合せにおける点と面の距離の二乗和を誤差として、最小二乗問題を線形化して解くことによって全体位置合わせを行う手法を提案している[8]。Benjamaa[9]は法線方向による頂点の分割と *multi-z-buffer* を使ってペアワイズな位置合わせを高速化するとともに、Bergevin の手法[10]を拡張して重なり合う距離画像間で複数枚の同時位置合わせする手法を提案している。また Nishino は M 推定を用いたロバストな同時位置合わせ手法を提案している[11]。

このように、これまでに様々な位置合わせ手法が提案されているが、いずれの手法でも問題となるのが対応点探索の計算量である。基本的な ICP では最近傍点を全ての頂点に対して計算するため、対応する 2 枚の距離画像の頂点数を等しく N とした場合、対応点探索の計算量は $O(N^2)$ である。そこで探索を高速化する手法としては探索木 (*kd-tree*) を使う手法や[12], *kd-tree* に加えて近傍点をキャッシュしておく事により、探索範囲を狭める手法なども提案されている[13]。しかし *kd-tree* による探索の計算量は $O(N \log N)$ である。多重解像度と近傍探索を組合せた $O(N)$ の手法もあるが[14], 収束が遅く、複数の異なる解像度を持つ距離画像間で多重解像度を計画するのは難しい。 $O(N)$ の視線投影型手法として Blais らの手法があるが[6], センサ毎或は測定毎にパラメータを求める必要がある点が問題である。

そこで我々は、センサに依存せずに多数の異なる解像度を持つ距離画像を高速に同時位置合わせ手法を提案する。まず、2 章では同時位置合わせ手法とその高速化手法について説明し、3 章では実験結果を示す。4 章はまとめである。

2. 同時位置合わせ手法

本章では、我々が提案する高速同時位置合わせ手法について説明する。位置合わせ手順は ICP と同様であるが、全ての距離画像に対して誤差関数が一つ定義され、全ての距離画像の位置が同時に推定される。手順の概略を以下に示す。

1. 全ての距離画像間で以下の計算を行う
 - 1.1 全頂点の対応点を探索する
 - 1.2 全対応点間の誤差を計算する
2. 計算された全ての誤差を最小化するように各距離画像の変換行列を求める
3. 終了条件を満たすまで 1, 2 を繰り返す

以下、これらの手順についてそれぞれ説明する。

2.1. 対応点探索

まず、全距離画像から 2 つの画像を選択しそれぞれモデル画像、シーン画像と呼ぶことにする。モデル画

像上の頂点とそれに対応するシーン画像のメッシュ上の点に対応付けられる。図 1 に示す で示されたデータがモデル画像であり、 で示されたデータがシーン画像である。対応点はモデル画像上の頂点からシーン画像取得時のレンジセンサの視線方向に直線を延ばし、シーン画像のメッシュとの交点である。このときモデル画像の視線方向に対して、対応するシーン画像のメッシュが裏面方向になる場合や、対応点間距離が与えられた閾値 l_{th} 以上の場合は誤対応として除去する。

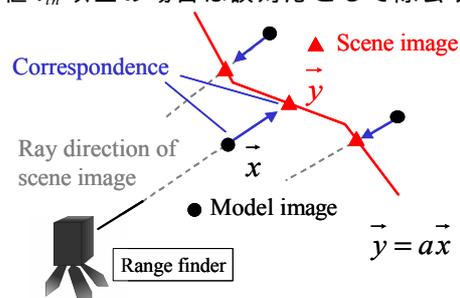
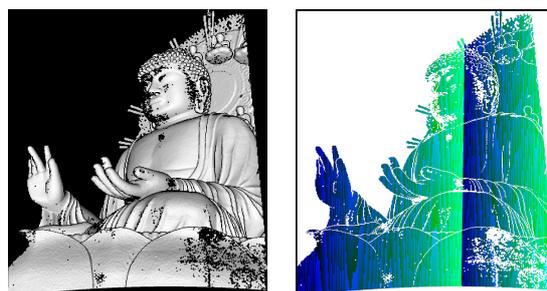


図 1 対応点探索

Fig.1 Search correspondences

2.2. インデックス画像を用いた対応点探索

通常、対応点探索には非常に時間が掛かるため、インデックス画像を用いて高速化する手法を提案する。まず、シーン画像の各メッシュに重複しないインデックス値を割り当てる。そして、同様にこのインデックス値を重複しない色に変換してそれぞれのメッシュの色とする。(ただし、描画色とインデックス値が同じ精度で表されるならば、インデックス値からインデックス色への変換は必要ない。)そして、シーン画像の測定方向から、透視変換で割当てた色によってメッシュを描画する(図 2)。このメッシュの描画処理は、グラフィックスハードウェアを利用して高速に行う事ができる。また、角度等間隔で測定するセンサでは、直交座標から極座標に変換して投影した方が良い場合もある。



(a)陰影画像 (b)インデックス画像

図 2 インデックス色による描画

Fig.2 Rendering with index color

次にモデル画像の各頂点を、シーン画像の視線方向から透視変換によって画像面に投影する。そして得られた座標値から、対応する画像上の色を得る。この色をシーン画像のメッシュインデックス値に逆変換することによって、対応するメッシュを得る事が出来る(図 3)。ただし、メッシュ描画の際にエイリアシングが起

こるため、正しい対応が得られない場合がある。そこで、得られた座標を中心とした周囲 3×3 画素を探索する。正しい対応メッシュは、対応するシーン画像のメッシュが張る平面上に、モデル画像上の頂点を投影する事で判定する。また同時に正確な対応点も計算する。メッシュ数が頂点数 N に比例するとして、このインデックス画像を用いた対応点探索の計算量は O(N)である。また、インデックス画像以外に特にメモリ空間も必要としない。

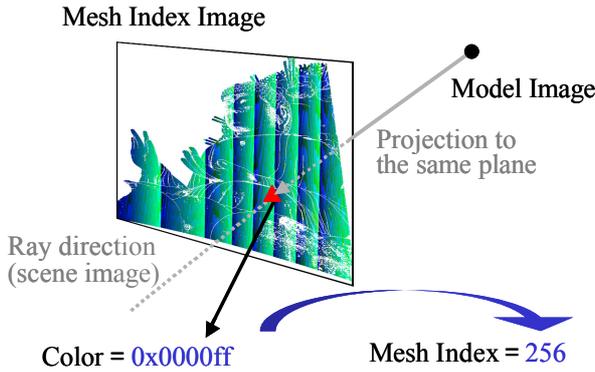


図 3 対応メッシュの取得
Fig.3 Obtaining corresponding meshes

モデル画像の視線方向に対して対応点探索を行う場合は、頂点インデックス画像を用いる。メッシュインデックス画像の場合と同様に、各頂点に重複しない値を割当て、それを重複しない色に変換する。描画はモデル画像の視線方向に対して行う。次に、シーン画像のメッシュに含まれる頂点を、同様にモデル画像の視線方向から画像面に投影して矩形領域を得る。そして、この矩形領域に含まれる頂点に、シーン画像のメッシュ番号を登録していく(図 4)。

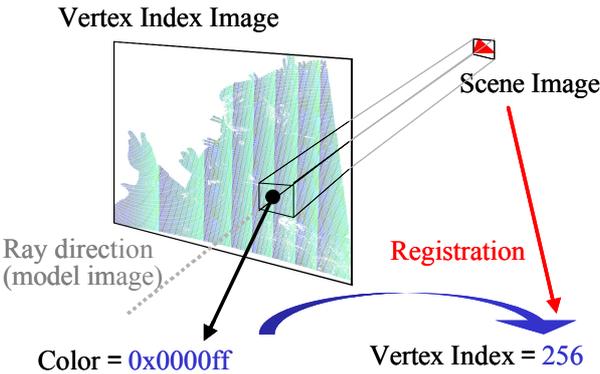


図 4 対応頂点の取得
Fig.4 Obtaining corresponding vertices

このとき、一つの頂点に複数のメッシュが対応する場合がありますので、対応メッシュとの距離を記録しておき、投影されたメッシュとの距離が既に記録されたメッシュの距離より小さい場合のみ更新を行う。頂点インデックス画像を用いる手法は、メッシュインデックス画像を用いる手法より計算量が多いが、より正確に対応点を得る事ができる。本手法では、計算時間が短いメ

ッシュインデックス画像を用いる手法を使用した。

2.3. 誤差評価

対応する 2 点間の距離には頂点法線を考慮した点と面の距離を用いる。モデル画像上の頂点 x とそれに対応するシーン画像メッシュ上の点 y とする。また、 n を頂点 x の法線 n_x と対応点 y の法線 n_y の平均とすると、誤差評価式は以下ようになる。

$$R_M n \cdot \{(R_S y + t_S) - (R_M x + t_M)\} \quad (1)$$

$$n = \frac{n_x + n_y}{\|n_x + n_y\|} \quad (2)$$

ここで R_M, t_M はモデル画像の回転、平行移動行列であり R_S, t_S はシーン画像の回転・平行移動行列である。簡単化のために n に掛かる回転行列は R_M とした。これを全ての距離画像の組合せ (i, j) が含む全頂点 k に対して計算し、以下の式で定義される二乗誤差 e^2 を求める。

$$e^2 = \min_{R, t} \sum_{i \neq j, k} (R_M n \cdot \{(R_S y + t_S) - (R_M x + t_M)\})^2 \quad (3)$$

式(3)に示した誤差評価式はこのままでは非線形であり最小化計算が難しいため、これを線形化する[7]。求める変換の回転角を微小角と仮定すると、回転行列 R は以下のように書くことができる。

$$R = \begin{pmatrix} 1 & -c_3 & c_2 \\ c_3 & 1 & -c_1 \\ -c_2 & c_1 & 1 \end{pmatrix} \quad (4)$$

また、平行移動行列を、

$$t = (t_x \quad t_y \quad t_z) \quad (5)$$

とすると、全ての距離画像枚数を n として式(2)は以下のように変形する事ができる。

$$e^2 = \min_d \sum_{i \neq j, k} \|A_{ijk} d - s_{ijk}\|^2 \quad (6)$$

$$s_{ijk} = n_{ik} \cdot (x_{ik} - y_{ijk}) \quad (7)$$

$$A_{ijk} = \begin{pmatrix} 0 \dots 0 & C_{ijk} & 0 \dots 0 \\ 6i \times 1 & 6(i-1) \times 1 & 6(j-1) \times 1 \end{pmatrix} + \begin{pmatrix} 0 \dots 0 & -C_{ijk} & 0 \dots 0 \\ 6j \times 1 & 6(j-1) \times 1 & 6(l-1) \times 1 \end{pmatrix} \quad (8)$$

$$C_{ijk} = \begin{pmatrix} n_{ik} \times y_{ijk} \\ -n_{ik} \end{pmatrix} \quad (9)$$

$$d = (m_0 \dots m_{n-1}) \quad (10)$$

$$m_i = (c_{1i} \quad c_{2i} \quad c_{3i} \quad t_{xi} \quad t_{yi} \quad t_{zi}) \quad (11)$$

2.4. 変換行列の計算

各距離画像の変換行列は、式(6)より最小二乗法によって d を求める事によって得られる。

$$\left(\sum_{i \neq j, k} A_{ijk}^T A_{ijk} \right) d = \sum_{i \neq j, k} A_{ijk}^T s_{ijk} \quad (12)$$

実際には 0 番目の距離画像を移動させない事で一意な解を求めるため、 d は $((n-1) \times 6)$ 個の変数を持つ連立一次方程式の解として得られる。全ての距離画像が少なくとも一つ以上の他の距離画像と重なりを持つとき、

この連立一次方程式の係数行列は $((n-1) \times 6) \times ((n-1) \times 6)$ の正値対称行列となる。この係数行列は距離画像枚数が多くなるにつれて非常に大規模な行列となると共に、重なり枚数は相対的に少なくなるので疎な行列となる。また、この行列は 6×6 の非零パターンを持つ事も分かる(図5)。

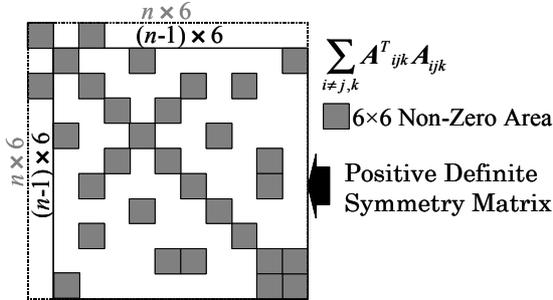


図5 係数行列

Fig.5 Coefficient matrix

係数行列が正値対称行列である場合、連立一次方程式の直接解法としては Cholesky 分解を用いる方法が一般的である。しかし Cholesky 分解の計算量は $O(n^3)$ であるため、大規模行列の扱いには適していない。そこで、反復法の前処理付き共役勾配法 (PCG) を用いる事にする。特にこの係数行列は 6×6 の非零パターンを持つ疎行列である事が分かっているため、これに特化した不完全 Cholesky 分解共役勾配法 (ICCG) を用いる事で高速な計算が可能である。

3. 実験

本章では、実際に測定された複数距離画像の位置合わせを行うことで我々の手法の有効性を示す。計算には ArhlonMP2400+ プロセッサ、2GByte メモリと GeForceTi4600 グラフィックカードを搭載した PC を使用した。また、位置合わせを行ったデータは以下の2組の距離画像である。

モデル1. デーヴァ尊顔 (図6(a))

Sensor	VIVID900
Images	45
Vertices	Max:76612, Min:38190, Ave:67674
Meshes	Max:150786, Min:71526, Ave:130437

モデル2. 奈良大仏 (図6(b))

Sensor	Cyrax2500
Images	114
Vertices	Max:155434, Min:11231, Ave:81107
Meshes	Max:300920, Min:18927, Ave:148281

VIVID900 は近距離用のセンサであり、CCD カメラとレーザを用いた光切断法によって形状を測定する[15]。測定解像度は CCD の画素数によって定められている。Cyrax2500 は time-of-flight 方式のスキャン機構を持つ長距離用レンジセンサであり、モーター制御により測定間隔(角度)や測定解像度を自由に定める事ができる[16]。また、測定範囲内でも測定されなかったデータ点は用いないため、これらの距離画像は

異なった測定解像度を持っている。

実際に測定された距離画像は非常にデータ量が多く、メモリ空間の制約から1台の PC で扱う事ができないため、1/4に間引いたデータを使用した。



(a)

(b)

図6 対象物: (a)デーヴァ尊顔, (b)奈良大仏

Fig. 6 Target objects: (a) Face of Deva,

(b) Nara Great Buddha

3.1. インデックス画像の解像度と対応点数

まず、探索に用いるインデックス画像と、我々の手法によって得られる対応点数の関係について検証する。前述のように、インデックス画像の描画の際には、エイリアシングが起こる場合や、解像度が低いために全てのメッシュが描画されない場合がある。つまり、十分な対応点数を得るためには、インデックス画像に十分な解像度を割当てる必要がある。

図7にモデル1,2の各最小,最大メッシュ数を持つ距離画像をシーン画像とした場合の、インデックス画像解像度と得られる対応点数との関係を示す。これらのシーン画像に対するモデル画像は任意である。図中の対応点数は、実際の対応点数 v_c と本手法を用いて得られた対応点数 v'_c の比で表されている。この図から分かるように、インデックス画像がある一定以上の解像度であれば、ほぼ全ての対応点が得られる。また、メッシュ数によって必ずしも画像解像度が定まるわけでは無いが、他の距離画像の結果もほぼこれらの範囲に収まる。この場合では、 800×800 程度の解像度があれば十分な数の対応点が得られる事が分かる。

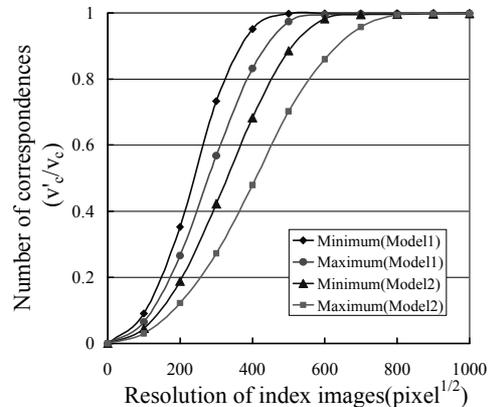


図7 インデックス画像解像度と対応点数

Fig.7 Relation between the resolutions of index image and the number of correspondences

実際に十分な対応点を得るために必要なインデックス画像の解像度は，センサの測定解像度によって決まる．ただし，ある程度以上の解像度が割当てられていれば十分な対応点が得られるため，最も解像度を必要とするデータに合わせて解像度を決めれば良い．つまり，全ての距離画像に対して同じ大きさの画像を用いることができる．そこで我々は，経験的に十分であると思われる値（ $1200^2 \sim 1600^2$ ）を用いて位置合わせ処理を行っている．

3.2. 頂点数に対する計算時間

次に，距離画像に含まれる頂点数に対する計算時間を評価する．この評価のために，モデル1及び2の全ての距離画像を用いて，同じ距離画像同士の1対1の位置合わせを行った．計算時間は対応点数にも影響されるので，互いの位置は移動させず頂点数と対応点数がほぼ等しくなるようにした．計算時間は1回の反復計算に要する時間とし，複数回の反復計算を行ったときの平均時間を評価した．この実験の結果を図8に示す．この図から分かるように，頂点数に対する計算時間はほぼ線形に増加する．また，モデル1と2の実験ではそれぞれ距離閾値 l_{th} 以外のパラメータは変更していない．つまり，本手法の効率性はセンサや測定状況には依存しないと言える．

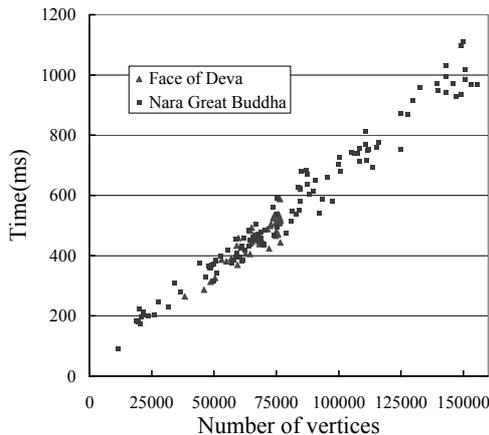


図8 頂点数に対する計算時間

Fig.8 Computation time related to the number of vertices

3.3. 変換行列の計算時間

本節では変換行列の計算時間について評価する．前述のように，行列計算の時間は距離画像の枚数に大きく影響される．そこで，距離画像枚数と行列計算時間の関係について検証する．実験では，距離画像枚数を変えて同時位置合わせを行い，変換行列の計算時間のみを測定した．この実験にはモデル2の距離画像を使用した．図9に距離画像枚数と行列計算時間の関係を示す．比較のために本手法とCholesky分解を用いた場合の二つの結果を示す．このときのICCGの誤差閾値は 10^{-6} とした．

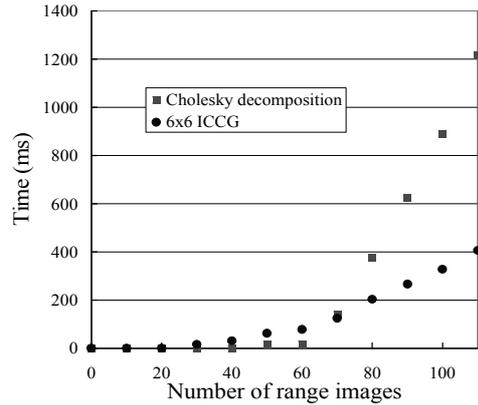


図9 変換行列の計算時間

Fig.9 Time of transformation matrix computation

この図から，距離画像枚数が60枚以下の場合Cholesky分解を用いた方が計算時間は短い事が分かる．しかしCholesky分解の計算量は $O(n^3)$ であり急激に計算時間が増加するため，距離画像枚数が70枚以上になるとICCGを用いた場合の方が計算時間は短くなる．ただし，ICCGを用いた場合の行列計算の計算量はCholesky分解を用いた場合よりも少ないが， $O(n)$ 以下にはならない．そのため，距離画像枚数が更に増加した場合は大きな問題になると考えられる．

3.4. 位置合わせ結果

最後にこれらの距離画像を同時位置合わせした結果を示す．全ての距離画像は，予め1対1で順次位置合わせを行う事で初期位置を求めてある．モデル1の同時位置合わせ結果を図10に示す．実験では20回の反復計算を行い，このときの計算時間は1738秒であった．また，モデル2の位置合わせ結果を図11に示す．モデル1と同様に20回の反復計算を行い，計算時間は7832秒であった．

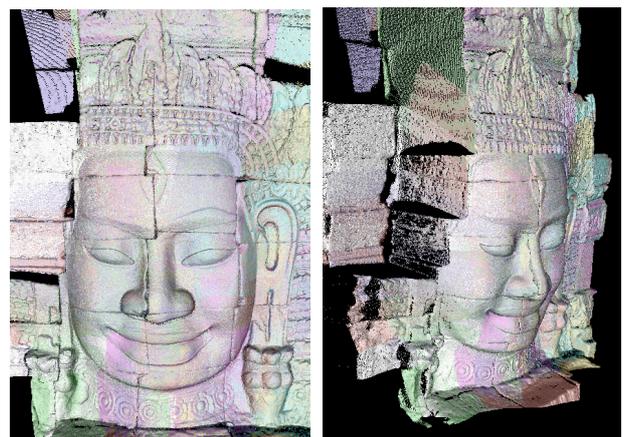


図10 位置合わせ結果（デーヴァ尊顔）

Fig.10 Alignment results(Face of deva)

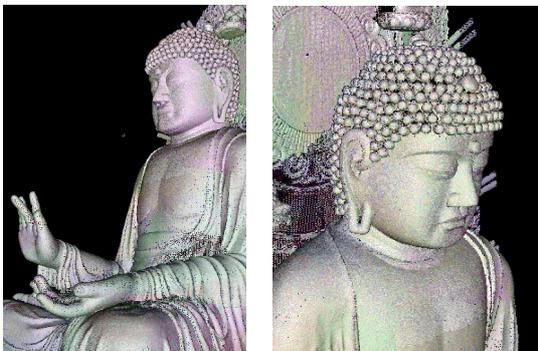


図 11 位置合わせ結果 (奈良大仏)
Fig.11 Alignment results(Nara Great Buddha)

文 献

- [1] K. Ikeuchi and Y. Sato, *Modeling from Reality*, Kluwer Academic Press, 2001.
- [2] K. Ikeuchi, Modeling from Reality, In *Proc. Third International Conference on 3D Digital Imaging and Modeling (3DIM)*, Quebec City, Canada. May 28-June 1, 2001.
- [3] M. Levoy et. al. The Digital Michelangelo Project. In *Proc. SIGGRAPH 2000*, pp.131-144.
- [4] P. J. Besl and N. D. McKay, A method for registration of 3-D shapes, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14(2) 1992, 239-256.
- [5] Y. Chen and G. Medioni, Object modeling by registration of multiple range images, *Image and Vision Computing* 10(3), 1992, 145-155.
- [6] G. Blais and M. Levine. "Registering Multiview Range Data to Create 3D Computer Objects," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 17, No. 8, 1995.
- [7] T. Masuda, K. Sakaue, and N. Yokoya, Registration and Integration of Multiple Range Images for 3-D Model Construction. In *Proc. Computer Society Conference on Computer Vision and Pattern Recognition*, 1996.
- [8] P. J. Neugebauer. Reconstruction of Real-World Objects via Simultaneous Registration and Robust Combination of Multiple Range Images. *International Journal of Shape Modeling*, 3(1&2):71-90, 1997.
- [9] R. Benjema and F. Schmitt. Fast global registration of 3d sampled surfaces using a multi-z-buffer technique. In *Proc. Int. Conf. on Recent Advances in 3-D Digital Imaging and Modeling*, pages 113-120, May 1997.
- [10] R. Bergevin, M. Soucy, H. Gagnon, and D. Laurendeau. To-towards a general multi-view registration technique. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(5):540-547, May 1996.
- [11] K. Nishino and K. Ikeuchi, Robust Simultaneous Registration of Multiple Range Images, In *Proc. Fifth Asian Conference on Computer Vision*, pp454-461, Jan., 2002.
- [12] Z. Zhang. Iterative point matching for registration of free-form curves and surfaces. *International Journal of Computer Vision*, 13(2):119-152, 1994.
- [13] D. Simon, M. Hebert and T. Kanade. Realtime 3-D pose estimation using a high-speed range sensor. In *IEEE Intl. Conf. Robotics and Automation*, pages 2235-2241, San Diego, California, May 8-13 1994.
- [14] Jost, T. and Hugli, H. A multi-resolution ICP with heuristic closest point search for fast and robust 3D registration of range images, In *Proc. Fourth International Conference on 3D Digital Imaging and Modeling (3DIM2003)*, pages 427-433, Oct 2003.
- [15] <http://www.minoltausa.com/vivid/>
- [16] <http://www.cyra.com/>

4. おわりに

本論文では複数枚の距離画像を高速に同時位置あわせする手法を提案した。位置合わせ計算の中で最も計算コストが高い対応点探索を、インデックス画像を用いた手法によって高速化した。また、変換行列の計算には 6×6 の非零パターンを持つ行列に特化した ICCG 法を用いる事によって、大規模な距離画像に対しても短時間で行列を計算する事が可能となった。実験では、レーザレンジセンサによって実際に測定された距離画像を位置合わせする事で、本手法の有効性を示した。今後の課題としては、本手法を拡張して高速かつ高精度な位置合わせ手法を開発する必要があると考えられる。

謝辞

本研究は、科学技術振興機構戦略的創造研究推進事業 (CREST) の支援を受けて行われました。また、測定にご協力頂いた奈良東大寺の方々及び、日本国政府アンコール遺跡救済チーム (JSA) の方々に感謝の意を申し上げます。