

FLSMを用いた隠れに頑強なモーションキャプチャの開発

岩下 友美(九大) 倉爪 亮(九大) 辻 徳生(九大) 原 健二(九大) 長谷川 勉(九大)

Development of Robust Motion Capture System for Target Occlusion using FLSM

*Yumi IWASHITA[†], Ryo KURAZUME[†], Tokuo TSUJI[†], Kenji HARA[†],
Tsutomu HASEGAWA[†] [†]Kyushu University

Abstract— The multi-view stereo techniques and the visual hull are well-known 3D all-round reconstruction methods of a real object using multiple cameras. In both methods, however, when there are multiple targets, each 3D model is not reconstructed due to mutual occlusion between them. To tackle this problem, we have developed a prototype system of a 3D motion tracking system of multiple targets in stereo range images using Fast Level Set Method. This paper introduces the 2nd prototype system using a PC-cluster and experiments of real-time tracking of multiple objects.

Key Words: Level set method, 3D shape reconstruction, Motion capture, Pc cluster

1. はじめに

実在物体の3次元全周形状を復元する手法として、視体積交差法 [1][2], 多視点ステレオ [3] などが挙げられる。しかし、シーン内に複数物体が存在する場合、物体間のオクルージョンが発生するため、これらの手法で正確な物体形状の復元を行うことは困難である。この問題に対し、我々はこれまでにステレオカメラを用い、オクルージョンに頑強なモーションキャプチャの試作機を提案している [6]。構築したシステムでは、ステレオカメラから得た3次元距離画像に対して、移動体追跡や3次元幾何モデリングで用いられる Level Set Method(LSM)[4] を高速化した Fast Level Set Method(FLSM)[5] を適用し、複数対象物体の3次元形状を復元しており、オクルージョンが発生する場合でもそれぞれの形状を分離して復元できることを確認している。本報告では、Fast Level Set Method の計算を PC クラスタを用いた並列計算により高速に実現することで、これまでに構築した試作機に比べて対象物体のより詳細な3次元形状をより高速に復元できるモーションキャプチャの第2次試作機を紹介し、開発したシステムを用いた人体概形の実時間追跡実験の様子を示す。

2. Fast Level Set Method

2次元 xy 平面上での Level Set Method(LSM)[4] を用いた境界追跡法について説明する。本手法ではまず、平面上の各点 (x,y) において、時刻 t における補助関数 $\psi(x,y,t)$ を導入し、境界位置 $\gamma(x,y,t)$ は $\psi(x,y,t) = 0$ を満たす (x,y) で表すこととする。次に、 ψ についての差分方程式を以下のように定義する。

$$\psi_{i,j}^{n+1} = \psi_{i,j}^n - F(i,j)|\nabla\psi_{i,j}^n|\Delta t \quad (1)$$

ここで、 (i,j) は xy 平面上のグリッド座標、 Δt は時間ステップ幅である。また、 F は ψ の成長速度であり、単位時間あたりに ψ が法線方向に移動する量を示す。Fig.1 に例として、境界の移動の過程で位相が変化する様子を示す。まず、初期境界 $\gamma(t)$ の補助関数 $\psi(t)$ を生成し (Fig.1(b) (a)), 式 (1) に従って $\psi(t)$ を更新する

(Fig.1(a) (c))。次に、更新した補助関数 $\psi(x,y,t) = 0$ が0となる位置を新たな境界とすることで、位相の変化に対応した領域追跡が可能となる (Fig.1(c) (d))。

しかしこの LSM は成長速度 F を決定する処理に多くの計算が必要であり、リアルタイム性が要求されるアプリケーションには不向きであると考えられてきた。この問題に対し、我々はこれまでに成長速度 F の決定をあるルールに基づく単純な数値の上書き処理に置き換えることで、精度を低下させることなく LSM を高速に実現できる Fast Level Set Method(FLSM)[5] を提案している。

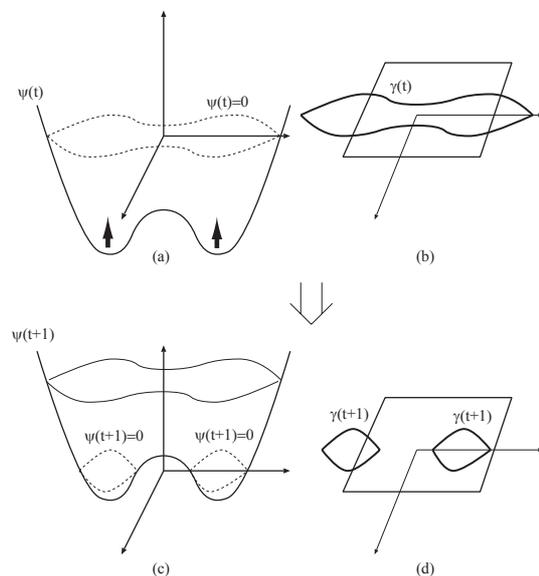


Fig.1 Level Set Method の概略

3. ステレオカメラを用いた複数人体の3次元リアルタイム追跡

対象物体全周の3次元形状をテレビカメラを用いて復元する手法として用いられる視体積交差法は、任意の視点から対象を撮影し、その時得られる2次元シル

エッジを実空間に投影した錐体(視体積)を考えると、対象はその錐体中に含まれるという制約条件に基づいている。従って、対象物体の詳細な形状を求めるには多くのカメラを配置する必要があり、またシルエットが正確に切り出されるように背景も工夫する必要がある。また、凹物体やオクルージョンがある場合には、原理的に形状を復元できない。

そこで、我々はこれまでにより少数のステレオカメラ(PointGrey社製, Bumblebee)を用い、提案したFLSMを適用して凹物体やオクルージョンがある場合でも人体概形のリアルタイム追跡を行うことができるモーションキャプチャの試作機を構築した[6]。本報告では、Fast Level Set Methodの計算をPCクラスタを用いた並列計算により高速化することで、これまでに構築した試作機に比べて対象物体のより詳細な3次元形状をより高速に復元できるモーションキャプチャの第2次試作機を紹介し、開発したシステムを用いた人体概形の長時間追跡実験の様子を示す。

3.1 複数距離画像の統合とFLSMによる人体概形の追跡

Fig.2に構築したシステムの概要を示す。2台のステレオカメラはそれぞれ1.2m立方のボクセル空間中心から2.3m離して設置されている。まず、それぞれのステレオカメラから距離画像を獲得した後、次の手順で3次元のボクセル空間で複数距離画像の統合を行う。まず、各ステレオカメラに対してボクセル空間を1つずつ設定する。また、各ボクセルはIN(人体の内部)、OUT(人体の外部)の投票箱を持つものとし、最初全てのボクセルは初期値としてINが投票されている。次に、Fig.3に示すように、ステレオカメラの距離画像上の各画素をボクセル空間に逆投影し、その画素に対応する距離データ点を含むボクセルAとカメラの投影中心との間に存在するボクセルに対してOUTを投票する。このような投票を全ステレオカメラの全ての画素に対して行い、全てのボクセル空間に共通してINが投票されているボクセルの集合Sを求める。

次に、求めたボクセル集合Sを用い、FLSMによる人体概形の追跡を行う。まず、ボクセル集合Sの中でOUTが投票されているボクセルに隣接するボクセルを停止領域(stopping region)、その他のボクセルを内部領域(inner region)と定義する。全ボクセル空間を内包する初期曲面を考え、FLSMを用いて曲面をある速度で収縮させ、停止領域内にzero level setが到達した場合には、成長速度を小さくすることにより人体概形を復元する。

上記の計算は、Fig.4に示すように処理の高速化を図るため、ギガビットイーサネットで相互に接続された7台の計算機(Pentium Xeon×2, 3.06GHz)により並列処理される。まずそれぞれステレオカメラが接続された2台の計算機(PC1, PC2)により、同期した2枚のステレオ距離画像を取得する。次に得られた距離画像はFLSM計算用計算機に送られる。ただしここではボクセル空間を4つの領域に分割し、それぞれの空間に1台ずつ計算機を割り当て(PC3-PC6)、4台で並列にFLSMの計算を実行した。各FLSM用計算機では、まず複数距離画像の統合を行った後、それぞれの領域に

対してFLSMを適用する。最後に処理結果が表示用計算機(PC7)に送られ、ボクセル空間全体の処理結果として統合、表示される。

Fig.5(a)にボクセル空間内に複数人体が存在する様子の水平断面図と、その時の停止領域と内部領域の様子を示す。このように両方のカメラから複数人体の全ての表面が観測できれば、FLSMを適用して正確な人体概形の復元が可能である。しかし、人体が移動すると相互オクルージョンが発生し、また腕や足などによっても自己オクルージョンが生じることがある。これらの場合、例えば停止領域がFig.5(b)のようになるため、2つが結合した概形が復元される。そこで、オクルージョン領域が発生する場合でもそれぞれの形状を分離して復元するために、オクルージョンの領域内では成長速度を小さくし、zero level setを進行させないことにする。これにより、分離して観測されていた物体は、オクルージョンが生じて、ある程度元の形状を保ち続けることができる(Fig.5(c) Fig.5(d))。

本システムの特徴として、次の3つが挙げられる。まず、本手法では距離データを時空間的に統合するため、距離画像内のノイズに対して頑強である。次に、本手法はシルエット法に比べて背景と対象の分離が容易であり、背景差分や特殊なスタジオは必要ない。また、FLSMにより閉領域境界を追跡するため、一時的にデータ欠損が生じて滑らかな境界の保存が可能である。すなわち、一時的にオクルージョン領域が発生しても、その領域では以前の境界を保つことで物体形状の概形を得ることができ、オクルージョンに強い。

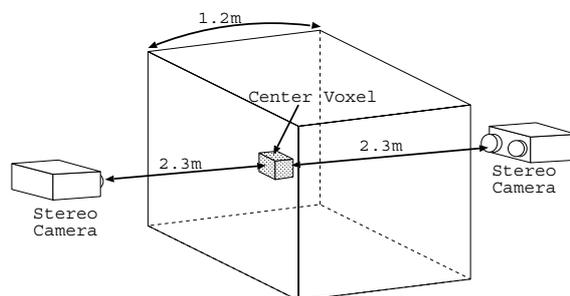


Fig.2 実験配置図

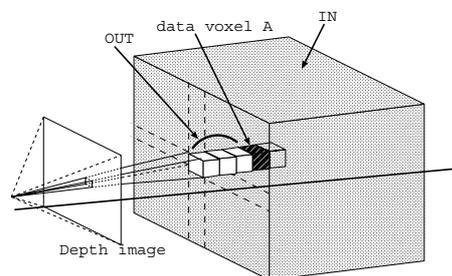


Fig.3 距離データ点によるボクセル空間への投票

3.2 実験

提案したFLSMを適用して人体の追跡実験を行った。使用した距離画像のサイズは640x480pixelである。

図6に人間動作の追跡実験の結果を示す。ただし、図6(a)は実験の様子、図6(b)は人体形状の追跡結果

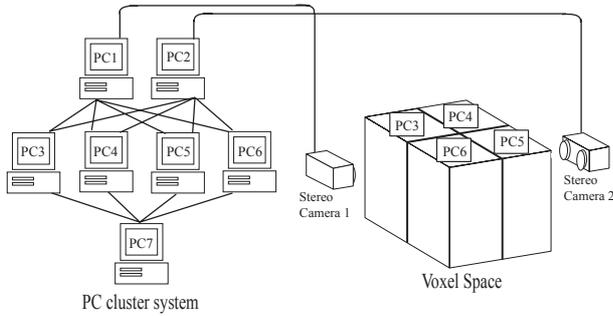


Fig.4 PC クラスタシステムの構成

を示している．この時使用したボクセル空間の解像度は $100 \times 100 \times 100$ (ボクセルの一边は 1.2cm) であり，ステレオカメラから距離画像を取得した後，複数距離画像の統合処理が 31ms ，各計算機間での通信時間は最大 69ms であった．また，Table 1 に FLSM の計算のそれぞれの処理時間を示す．PC1 台では，FLSM の計算時間は 220ms であったことから，PC クラスタを用いた並列計算により，対象物体の 3 次元形状をより高速に復元できることが確認できた．しかし，本システムは FLSM 処理においてボクセル空間を単純に分割しているため，物体の位置や形状，物体数によって特定の計算機に負荷が集中する可能性がある．そこで，今後はこの問題の解決策として対象の位置や対象形状の複雑さ（面の数）に応じた対象分割による並列分散処理による計算最適化手法を検討する予定である．

Table 1 FLSM 処理の計算時間

	PC クラスタ			PC1 台
	平均	最大	最小	
FLSM の処理時間 [msec.]	67	82	56	220
復元した形状のボクセル数	2535	3336	1710	8905

次に，ボクセル空間の解像度を $200 \times 200 \times 200$ (ボクセルの一边は 0.6cm) にし，人間動作の追跡実験を行った．ただし，FLSM の処理は 1 台の計算機のみで行っている．図 7 に，得られた人体形状にマーキングキューブ法を適用し，その後テクスチャマッピングを行った結果を示す．詳細に人体形状を復元できているものの，全体の処理時間は 1190ms であった．

最後に，複数対象物の追跡を行った結果を図 8 に示す．ただし，図 8(a) は実験の様子，図 8(b) はボクセル空間を上方から示したもので，複数対象物は図 5(c) から図 5(d) に示すように移動する．また，図 8(c) はボクセル空間を斜め上方から示したものである．図 8 より，複数対象物が交差し，相互オクルージョン領域を含む場合でも，人体概形を保存しつつ追跡できていることがわかる．これらの実験より，シーンに人体が複数存在し，お互いが交差するような場合でもそれぞれを分離した 3 次元形状の復元と追跡が可能であることが確認できた．しかし，本システムは単純な閾値処理によ

り複数対象物をそれぞれ分離しているため，オクルージョン時間や物体間の距離によって複数対象物を分離することが困難になる場合がある．そこで，今後はこの問題の解決策として，ラベリングした対象物の概形や体積などの情報を保存し，それらの情報を用いて複数対象物の分離を行う手法を検討する予定である．

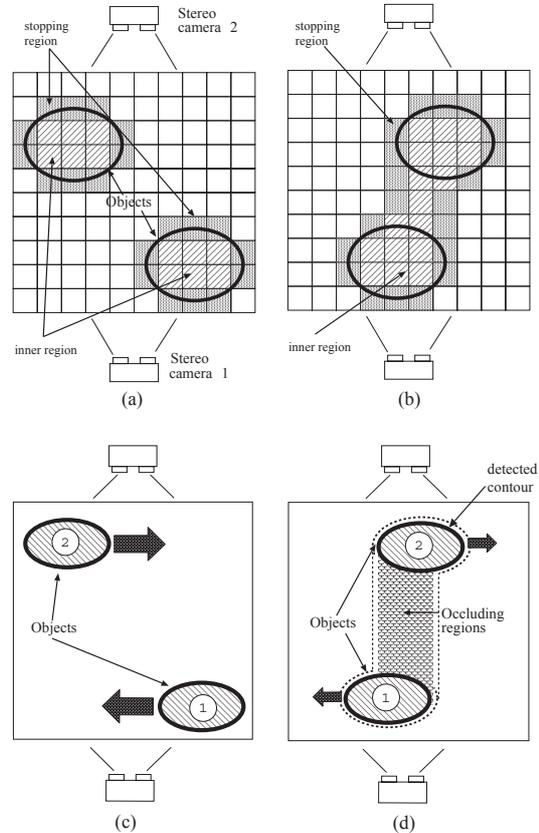


Fig.5 停止領域と内部領域

4. まとめ

本報告では Fast Level Set Method の計算を PC クラスタを用いた並列計算により高速化することで，これまでに構築した試作機に比べて対象物体のより詳細な 3 次元形状をより高速に復元できるモーションキャプチャの第 2 次試作機を紹介した．今後は，ステレオカメラの台数を増やし対象物体の全周に配置し，より詳細な対象物体の 3 次元形状を復元するシステムを開発する．

謝辞

本研究の一部は，21 世紀 COE プログラム「システム情報科学での社会基盤システム形成」，総務省戦略的情報通信研究開発推進制度及び科学研究費補助金（16700185）の支援を受けた．

参考文献

- [1] W. N. Martin and J. K. Aggarwal, "Volumetric description of objects from multiple views", *IEEE Trans. on PAMI*, Vol.5, No.2, pp.150-158, 1983
- [2] A. Laurentini, "The visual hull concept for silhouette-based image understanding", *IEEE Trans. on PAMI*, Vol.16, No.2, pp.155-162, 1994

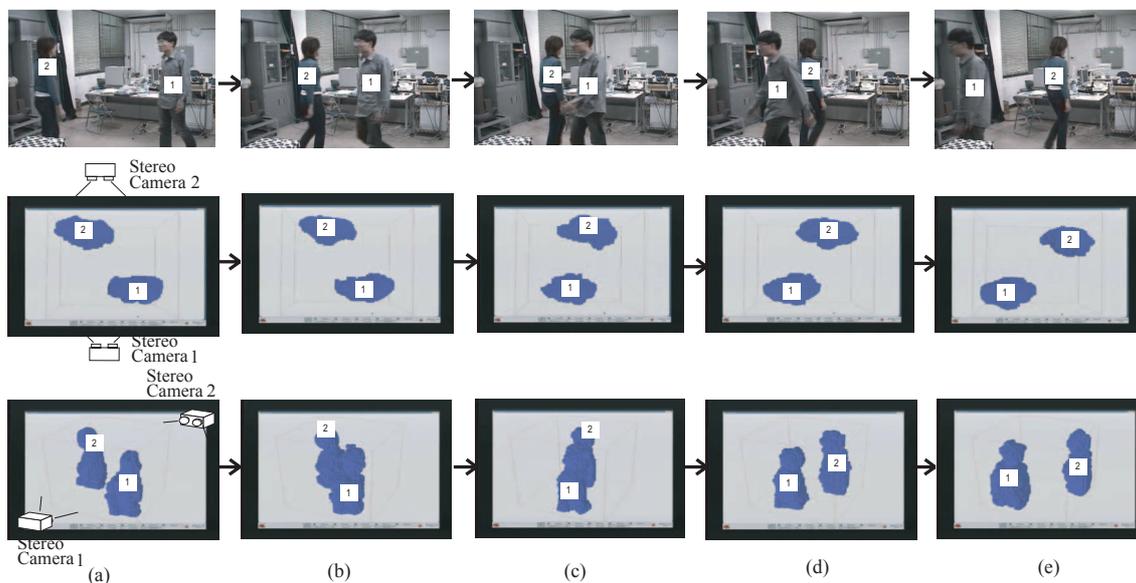


Fig.8 複数人体の3次元形状復元

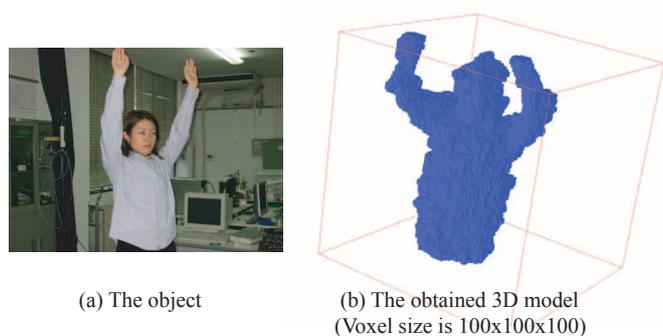


Fig.6 3次元人体形状の復元

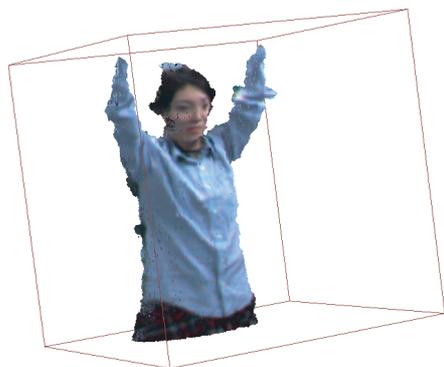


Fig.7 詳細な3次元人体形状の復元 (Voxel size is $200 \times 200 \times 200$)

Physics, Vol.79, pp.12-49, 1988

- [5] 倉爪 亮, 由井 俊太郎, 辻 徳生, 岩下 友美, 原 健二, 長谷川 勉, “Fast Level Set Method の提案とビデオ画像の移動体のリアルタイム追跡”, 情報処理学会論文誌, Vol.44, No.8, pp.2244-2254, 2003
- [6] 岩下 友美, 倉爪 亮, 辻 徳生, 原 健二, 長谷川 勉, “Fast Level Set Method を用いた3次元人体形状の実時間計測システムの構築”, 情報処理学会コンピュータビジョンとイメージメディア (CVIM) 研究会, CVIM-141-18, 2003

- [3] Saito, H., Baba, S., Kimura, M., Vedula, S. and Kanade, T.: *Appearance-based virtual view generation of temporally-varying events from multi-camera images in the 3d Room*, In Proceedings of Second International Conference on 3-D Digital Imaging and Modeling (3DIM99), pp. 516-525 (1999).
- [4] S. Osher and J. A. Sethian, “Fronts propagating with curvature dependent speed: Algorithm based on Hamilton-Jacobi formation”, *J. Computational*