

# 4 $\pi$ 計測システム: 複数カメラを用いた動物体の3次元形状計測\*

飯山将晃                      亀田能成                      美濃導彦  
京都大学                      京都大学  
情報学研究科<sup>†</sup>                      総合情報メディアセンター<sup>‡</sup>

## 1 はじめに

本研究では動物体の3次元形状を計測することを目的としたシステムを提案する。

実世界に存在する物体の形状を観測によって獲得し、計算機上で構築する技術は、バーチャル博物館などに代表される電子的な情報蓄積、アミューズメント分野などにおけるCGアニメーションなどさまざまな分野への応用が期待できる技術である。

三次元形状を計測する手法として、レーザレンジファインダを用いた手法と視体積交差法などの画像ベースの手法 [1][2][3] とがあるが、前者は計測に要する時間が後者と比較して長く、動物体の計測には適さない。そこで提案システムでは視体積交差法を用いて3次元形状の計測を行う。

また、一般的に3次元形状の計測能力は観測方向、つまりカメラ台数に比例する。これまで提案されてきたシステムでは、カメラは撮影物体の上方にしか配置されておらず、物体の上面についての計測能力は高いものの、物体の下面の形状を計測することは困難であった。そこで、我々のシステムでは物体を全方位(4 $\pi$ 方向)から計測できるように、20台の同期撮影が可能なカメラ配置した。これによって物体の下面を含む物体全周の三次元形状計測が可能になると期待できる。

以下2節では、3次元形状を再構成するアルゴリズムについて述べ、シミュレーション実験によって計測能力を評価する。3節では我々が提案する計測システムの概要について述べ、三次元計測の空間解像度・時間解像度について考察する。4節では実物体を用いた

実験を行いシステムの性能を評価する。5節はまとめである。

## 2 三次元形状の計測

三次元形状を計測する手法として、ステレオ視 [4] と視体積交差法 [5] がある。この2つ手法を比較した場合、視体積交差法はデータとして物体が投影されたシルエットしか必要でないため画像上で生じるノイズに対してロバストであり、また、テクスチャ特徴に乏しい物体に対しても適用可能な方法である。また、得られる結果がボリュームであるため、ステレオ視のように2.5次元の計測結果を統合しなければならないというデメリットから開放されるという利点がある。

以上のことから本システムでは視体積交差法を用いて3次元形状を復元する。

### 2.1 視体積交差法

視体積交差法について述べる。以下、 $n$ 台のカメラを用いるとし各カメラの位置・向きは既知であるとする。

物体はカメラ  $C_i$  で撮影された画像上に投影され、投影像は二次元画像平面上の領域として得られる。この領域を物体投影領域と呼ぶ。このとき、物体は必ず、カメラ  $C_i$  のレンズ中心を端点とし物体投影領域上の任意の点を通る半直線の集合からなる錐体状の開空間に内接して存在する。この開空間をカメラ  $C_i$  における視体積と呼ぶ。

従って、対象空間内の物体が複数のカメラ  $C_1, \dots, C_n$  によって観測されるとき、物体は必ず各カメラにおける視体積の積空間に内接する(図1)。この積空間を、視体積と呼ぶ。

\*4 $\pi$  Measurement System: A Complete Volume Reconstruction System for Freely-moving Objects

<sup>†</sup>Masaaki IYAMA, Graduate School of Informatics, Kyoto University

<sup>‡</sup>Yoshinari KAMEDA and Michihiko MINOH, Academic Center for Computing and Media Studies, Kyoto University

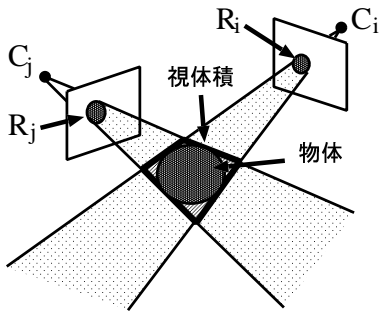


図 1: 視体積と物体形状

## 2.2 視体積交差法の計測能力

視体積交差法によって三次元形状がどのくらい正確に計測できるかについての考察を行う。

カメラ台数が増加するにつれ視体積の大きさは実物体の大きさに近づくが、一般には視体積の形状が実物体の形状と一致することはない。しかし実際的には、十分な数のカメラを用いれば、物体形状を十分近似するに足る視体積を得ることが可能である。

なお、十分な数のカメラを用いるとの仮定の下で、物体表面上の点が視体積交差法によって計測可能であるためには、その点を通る接線の中で、接点以外で物体と交差しないものが存在することが必要十分条件となる [6]。そのため、物体表面のうち窪んだ凹面については、カメラを増やしてもその部分を計測することはできないことに注意する必要がある。

また、原理的には計測可能な形状であっても、現実的には使用可能なカメラ台数には制限がある。そこで、シミュレーション実験を行い、カメラ台数と視体積の大きさとの関係を検証した。

計測する対象の物体として、立方体、球体、恐竜のモデルの三種類を作成し(図 2)、計測に使用するカメラの台数を 4,6,8,12,20,32(=12+20) と変化させ、獲得された視体積の体積を計測した。なお、カメラの配置を正  $n$  面体の頂点に視点があり、観測方向がその多面体の中心を向くように設定した。

カメラ台数と復元率 ((視体積の体積)/(物体の体積) で与える) との関係を図 3 に示す。カメラ台数が増加するにつれ復元率は 1 に近づくが、カメラ台数を 20 台以上増加させても復元率にそれほどの向上がみられなくなる。以上のことから、提案システムで使用するカメラ台数を 20 台と定めた。

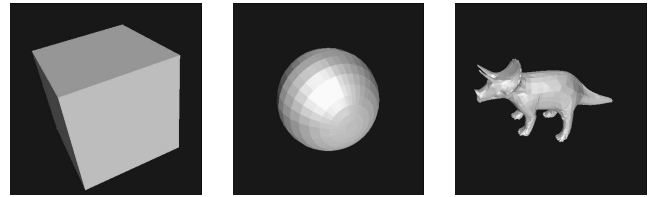


図 2: シミュレーションデータ

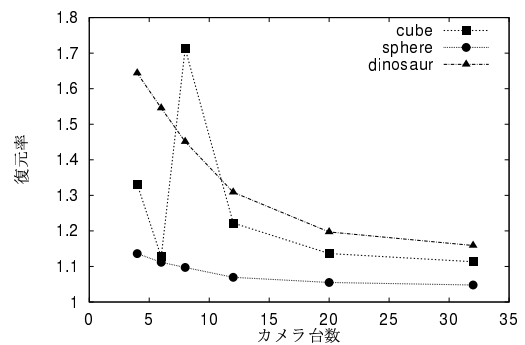


図 3: カメラ台数と復元率との関係

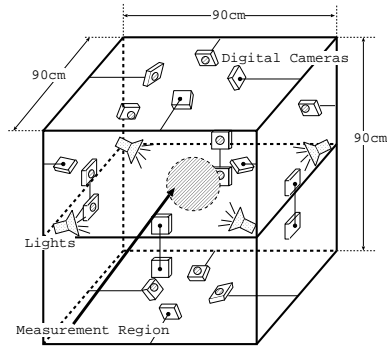
## 3 計測システム

### 3.1 システムの概要

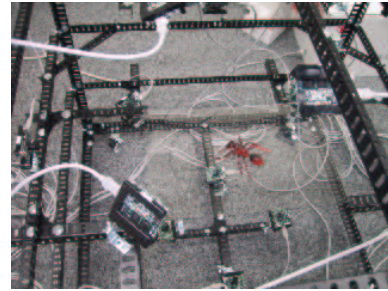
我々が構築した計測システムを図 4 に示す。

撮影カメラとして Point Gray Research 社の Dragonfly を用いた。640 × 480 の解像度で 15fps の撮影を行うことが可能である。また、各カメラは FireWire で同一バス上接続され、撮影信号を FireWire バス上で同期させることによって同期撮影が可能である。1 秒間の撮影によって獲得される画像データはカメラ 1 台当り  $640 \times 480(\text{pixel}) \times 8(\text{bit}) \times 15(\text{fps}) = 4.4\text{Mbytes}$  である。HDD の転送量の制約のため、撮影された画像を 5 台の PC に分散して蓄積する。

なお、視体積交差法においては物体が投影された領域を抽出することが必要である。本システムでは、背景差分によってその領域を抽出する。背景差分の精度を向上させるため、フレームの周囲に単色 (ブルー) のスクリーンを設置した。



(a)



(b)

図 4:  $4\pi$  計測システム

### 3.2 空間解像度と時間解像度

計測された三次元形状はボクセル表現によって表される。各ボクセルの大きさは画像の解像度と計測範囲によって決定される。今回作成したシステムでは各画像の解像度が  $640 \times 480$ 、計測範囲が直径  $24\text{cm}$  の球領域である。そこで、ボクセルの大きさを  $24\text{cm}/480 = 0.5\text{mm}$  とした。現在のところカメラの解像度には技術的な制約があり、計測範囲を広くすればするほどボクセルの大きさは大きくならざるを得ない。この問題を解決する方法として、全てのカメラが同一の領域を計測するのではなく、計測範囲を分割し、分割した計測範囲毎にカメラを割り当てるのが考えられる。これについては今後の検討課題である。

また、運動する物体を撮影する際、1 フレームの撮影を行っている間に物体が運動することによって、画像上でブレが発生する。さらにカメラの撮影動作の同期誤差によって、各カメラが投影した物体の位置にずれが生じる。これらの位置の誤差は空間解像度と同等の範囲に収まる必要がある。今回作成したシステムで使用するカメラのシャッター速度は  $1/16000\text{sec}$  である。また、撮影信号の同期誤差は  $20\mu\text{s}$  未満でありシャッター速度に比べて影響は少ない。1 ピクセルに相当するボクセルの大きさは  $0.5\text{mm}$  であることから、物体が  $8\text{m/s}$  で運動した場合、画像上での誤差が 1 ピクセルになる。以上のことから、物体の運動が  $8\text{m/s}$  までならば、シャッター速度に起因する位置の誤差は空間解像度と同等の範囲に収まると結論づけられる。

## 4 実験

本稿で提案したシステムの能力を検証するため、実物体を用いた実験を行った。

最初の実験として、大きさが既知の球および立方体の模型を用い、3次元形状の計測能力を2節で述べたシミュレーション実験と比較した。その結果を表1および図5に示す。20台のカメラを用いた時の復元率はシミュレーション結果とほぼ同等であるのに対し、上方の12台のカメラを用いた時の復元率はシミュレーション結果よりも大きく離れている。次の実験として、動物体の撮影能力を検証するため、人間の手の形状を計測した。その結果を図6に示す。運動する物体の形状を計測できていることが示されている。

表 1: Statistics of reconstructed volume

カメラ台数 (物体の種類)	体積		再構成率
	実データ	正解	
20(立方体)	$146.62\text{cm}^3$	$125\text{cm}^3$	1.173
12(立方体)	$234.82\text{cm}^3$	$125\text{cm}^3$	1.879
20(球)	$69.29\text{cm}^3$	$65.45\text{cm}^3$	1.054
12(球)	$79.12\text{cm}^3$	$65.45\text{cm}^3$	1.209

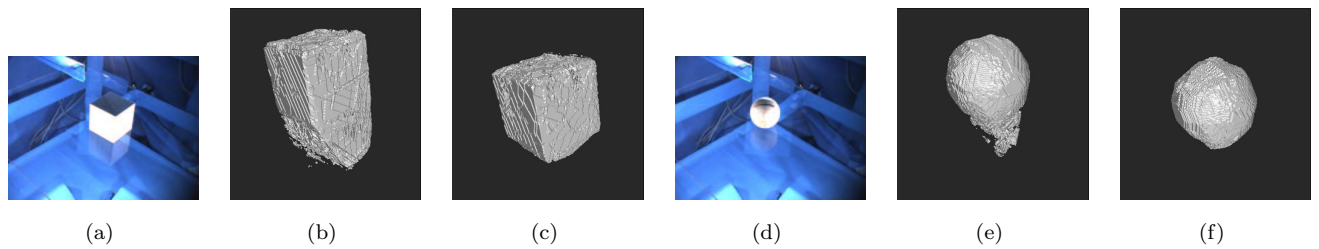


図 5: 計測結果 : (a) 撮影画像 , (b) 上方の 12 台のカメラを用いた計測結果 , (c) 20 台のカメラを用いた計測結果

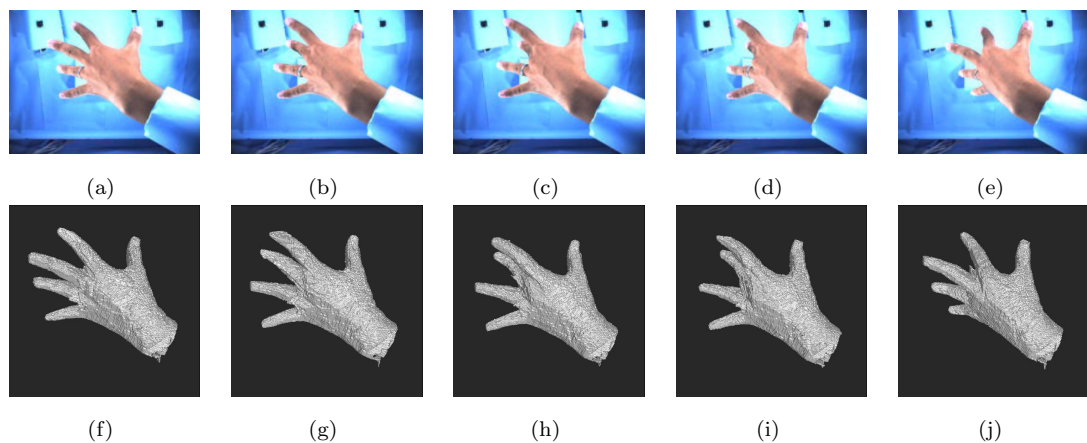


図 6: 動物体の計測結果: (a)-(e) 撮影画像, (f)-(j) 計測結果

## 5 まとめ

本研究では、動物体の 3 次元形状を計測することを目的としたシステムを提案し、その計測能力を評価した。模型を用いた実験の結果、シミュレーション結果によって得られた理想的な結果とほぼ同等の精度で計測されることが確認された。

視体積交差法はテクスチャ特徴に乏しい物体に対しても 3 次元形状の計測が可能という利点がある反面、計測できる形状に制限があるという欠点がある。十分なテクスチャが得られる場合、視体積交差法よりもステレオ視を用いて 3 次元形状を計測したほうがより正確である場合が多い。今後の課題として、得られる画像に応じて動的に最適な計測方法を選択する枠組みを考察することが挙げられる。

## 参考文献

[1] Takeo Kanade, Peter Rander, P.J. Narayanan, "Virtualized Reality: Constructing Virtual Worlds from Real

Scenes," IEEE MultiMedia, vol.4, no.1, pp.34-47, 1997

- [2] R. Bajcsy, R. Enciso, G. Kamberova, L. Nocera, R. Sara, "3D reconstruction of environments for virtual collaboration," Proc. 4th IEEE Workshop on Applications of Computer Vision, p.160-167, 1998
- [3] Wojciech Matusik, Chris Buehler, Ramesh Raskar, "Image-Based Visual Hulls," Proc. SIGGRAPH2000, pp.369-374, 2000
- [4] M. Okutomi, T. Kanade, "A multiple-baseline stereo," IEEE Trans. Pattern Analysis and Machine Intelligence. PAMI-15(4).pp.353-363,1993
- [5] W.N. Martin and J.K. Aggarwal, "Volumetric Descriptions of Objects from Multiple Views," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol. 5, No. 2, pp.150-158, 1983
- [6] Aldo Laurentini, "How Far 3D Shapes Can Be Understood from 2D Silhouettes," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 17, No.2, pp.188-195, 1995