

多視点画像マッチング法を用いた立体物の形状計測

中西 良成[†] 小林 希一[†] 蓼沼 眞^{††}
三ッ峰 秀樹^{††} 斎藤 豪^{†††} 中嶋 正之^{†††}

[†]NHK エンジニアリングサービス ^{††}NHK 放送技術研究所 ^{†††}東京工業大学

概要： 高精細な立体映像部品の作成を可能とするモデリング手法について検討、実験を行った。今回、モデリングの第1ステップである形状計測において、異なる視点からの画像を多数使用する2段階ブロックマッチング法により、高精細な形状計測を行う方法について提案する。提案手法はロータリーテーブル上のオブジェクトを撮像した連続画像の中央縦1ラインに注目し、注目点がそれらの画像内を等速に平行移動すると仮定できる画像複数枚から、ブロック画像を切り出し、2段階ブロックマッチング法を行う。これにより目標精度0.4画素での形状計測を可能としたので報告する。

Shape measurement of 3D object by block matching method using multiple viewpoint images

Yoshinari Nakanishi[†] Kiuchi Kobayashi[†] Makoto Tadenuma^{††}
Hideki Mitsumine^{††} Suguru Saito^{†††} Masayuki Nakajima^{†††}

[†]NHK Engineering Services Inc. ^{††}NHK Science & Technical Res. Labs.

^{†††}Tokyo Institute of Technology

Abstract: We propose a new block matching method for surface measurement of 3D objects, which is useful for precisely modeling and making high quality 3D image parts for virtual TV studio application. Multiple viewpoint images of 3D object photoed by HDTV camera within $\pm \theta$ degree are used, and each image block in these images is assumed to move parallelly to central image with the same velocity. Two-step block matching algorithm is adopted to improve measurement accuracy. In our experiments, ± 1 pixel accuracy was obtained in coarse matching, and final RMS error was less than 0.4 pixel with 4 times scaled images.

1. はじめに

近年、コンテンツ制作の効率化を目指した研究開発として、背景映像の部品化、任意視点からの画像の作成などの技術開発が進められ、異なる視点からの簡便な奥行き計測システムが実用化の域に達しつつある。しかしながら、ハイビジョン放送のコンテンツ制作に利用可能な立体映像部品を作成する高精細モデリング手法は未だ報告されていない。そこで我々は高精細な立体映像部品の作成を目標とする高精細モデリング手法について検討を進めてきた[1]。今回、従来のステレオ法を工夫し、オブジェクトをその周囲の多視点から撮像した多数の画像(多視点画像)を使用する2段階ブロックマッチング法により、立体物の形状を高精細に計測する方法を考察し、実験の結果、その有効性が確認されたので報告する。また本提案手法の精度向上についても検討し、その有効性が確認されたので併せて報告する。

2. オブジェクトの撮像条件と目標精度

本提案手法では、オブジェクト上のある一点の奥行きを計測するために、多視点画像を使用する。今回扱う多視点画像は、高さ 50cm のオブジェクトが撮像距離 1.3m で画面(1920x1080)一杯に撮像されていることを想定している。したがって、各視点はオブジェクトを中心とし、半径 1.3m の円周上に一定の角度間隔で位置することになる。またカメラの光軸はロータリーテーブルの台と平行であり、各視点から成る円の中心を通る。計測されるオブジェクトは、奥行き方向の大きさが撮像距離に比較して十分に小さいものを前提とする。

形状計測の目標精度については、立体画像における奥行き情報の擾乱と画質の関係についての主観評価実験の結果[2]を参考にした。この実験結果に拠れば、オブジェクトの輪郭における奥行き情報の擾乱に対する検知限は視覚 0.2 分、許容限が 0.8 分程度とされる。今回、形状計測の目標精度を、上記結果に若干のズームアップマージンを加味して、0.6 分すなわち 0.4(式(1),(2)を参照)画素と設定した。

$$500(mm)/1080(pixel) \approx 0.5(mm / pixel) \quad (1)$$

$$\frac{1300(mm) \tan 0.6(^{\circ})}{0.5(mm / pixel)} \approx 0.4(pixel) \quad (2)$$

ただし、上記精度 0.4 画素はオブジェクトの輪郭の精度であり、計測された奥行き誤差がオブジェクトの輪郭に最も大きく影響を与える向き、すなわちカメラの光軸と計測された点の奥行き方向が 90° の角度をなす向きで撮像されている画像における精度である。本提案手法では、奥行きを計測する点を正面画像の中央縦 1 ラインとし、前後 $\pm \theta^{\circ}$ 内に位置する視点からの画像を使用する。そのため、上記精度 0.4 画素を満足するためのマッチングの精度は $0.4 \sin \theta$ 画素となり、これを本提案手法の目標精度とする。

3. 多視点画像マッチング法

上記目標精度を満足する高精細な奥行き計測方法として、従来のステレオ法を工夫し、多視点画像を使用する多視点画像マッチング法を提案する。

3-1. 多視点画像の使用方法

ステレオ法において、奥行き計測のための対応点探索は、2 枚の画像が撮像された視点の間隔が狭いほど容易に、精度良く行える。一方、対応点探索における誤差が計測結果に及ぼす影響は、視点間隔に反比例する。また近接する画像間に対応を取りつつ、徐々に視点間隔を広げていき奥行きを計測する方法も考えられるが、誤差の伝搬による精度の低下が考えられる。そこでこれら相反する 2 つの性質を両立し、かつ誤差伝搬のない方法で計測精度を向上することを目指し、多視点画像を利用する手法を検討した。

本提案手法では、各画像の中央縦 1 ラインに注目し、そのライン上の各点の奥行き計測に、前後数点の視点から撮像した画像を利用する(例えば、注目画像とその前後各 5 点からの画像、計 11 枚を利用)。また注目画像の中央縦 1 ライン上の各画素は、撮像面と平行に等速運動をしていると仮定し、多視点画像はその等速運動の時系列画像であ

るとみなす。したがって、注目画像と最端画像での移動距離(シフト量)を決定することで、注目画像の中央縦1ライン上の各画素に各画像のどの点に対応するかが一意に決まる。そこで各画素のシフト量を仮定し、各画像からそのシフト量に応じた点を中心に加重平均によりブロック画像を作成する。このように作成したブロック画像を評価関数(3-2節参照)により総合的に評価し、最適なシフト量を求め、オブジェクト上の各点の奥行きを計測する。このように多数の画像を利用することにより、N枚の画像に対して計測誤差を $1/\sqrt{N}$ 倍に改善することができる。

3-2. 2段階マッチング法

ブロックマッチングは2段階(1次マッチング/2次マッチング)に分けて、階層的に行われる。1次マッチングでは、多視点画像から上記方法でブロック画像を作成し、任意の2ブロック画像間の相関係数を求める。その組み合わせ総和を評価関数とし、その値が最大となるものをシフト量とする。1次マッチングでは1画素刻みでシフト量を仮定し、 ± 1 画素の精度でシフト量を求める。

2次マッチングは、加重平均による内挿により多視点画像を縦横各4倍に拡大した画像に対して行う。マッチング範囲は、1次マッチングの結果 ± 2 画素(原画像の尺度での)シフト量とし、評価関数は1次マッチングと同じものを用いる。これにより ± 0.25 画素の精度でシフト量が求まるが、これでは目標精度を達成できない。そこでシフト量に対する評価関数値のグラフにおいて、評価関数値が最大となるシフト量の近傍5点の微分値を求め、直線近似し、そのゼロクロス点を2次マッチングの結果とする。上記結果は統計的には多数の画像を用いて計測したことに同意であり、これにより目標精度を満足する高精度なマッチングが可能となる。

4. 多視点画像マッチング法の改良

上記多視点画像マッチング法の精度向上のため、重み付けと窓関数の適用について検討した。

4-1. 重み付けの導入

先に述べたように、ステレオ法では視点間隔が大きいかほど奥行き計測の誤差は小さい。すなわち、奥行き計測誤差は基線長に反比例する。また多視点画像マッチング法は、多数の画像を総合的に評価し、対応点を決定するが、実際は任意の2ブロック間の相関係数の組み合わせ総和により評価している。そこでこの任意の2ブロック間の相関係数に、それらブロックを作成した2画像間の距離に比例した重み付けをし、その組み合わせ総和を評価関数とすることにより、より少ない画像数で誤差少なく計測することを可能とした。

両端の画像(使用する画像のうち最も視点間距離が大きくなる2枚の画像)で計測した場合の誤差を σ と仮定し、他の2枚の組み合わせによる計測に関しては視点間距離に反比例するように誤差を仮定する。例えば視点間距離が両端の画像間距離の1/2になる2枚の画像で計測する場合の誤差は 2σ と仮定する。この仮定の下で、多視点画像マッチング法の奥行き計測誤差を計算すると、N枚の画像に対して、重み付けの導入以前が $1/\sqrt{N}$ 倍で改善されるのに対し、重み付け導入後は $1/N$ 倍で改善される(図1を参照)。したがって、任意の2ブロック画像間の相関係数にそれらを作成した2画像間の距離に比例した重み付けを行うことは有効であり、画像数が多いほど効果が大きい。

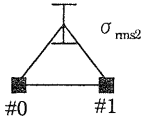
なお、重み付けによる改良は2次マッチングに対してのみ導入する。

4-2. 窓関数の導入

さらに相関の誤検出の影響を低減する試みとして、窓関数を用いることを検討する。評価関数において、任意の2ブロック間の相関係数が小さいものについてはノイズとみなし、相関係数を0とするような窓関数を用いる。

これによりオクルージョンやスペキュラーに起因する非対応点の影響を排除でき、誤った点を対応点とすることによる奥行き計測の低減が可能となる。

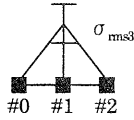
【N=2 の場合】



画像	誤差 σ_i	重み w_i
#0#1	1σ	1

$$\sigma_{rms2} = 1\sigma$$

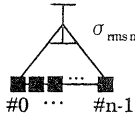
【N=3 の場合】



画像	誤差 σ_i	重み w_i
#0#2	1σ	1
#0#1 #1#2	2σ	$1/2$

$$\sigma_{rms3} = \frac{\sqrt{1 + (2 \times \frac{1}{2})^2 \times 2}}{1 + \frac{1}{2} \times 2} \sigma = \frac{\sqrt{3}}{2} \sigma$$

【N=n の場合】



画像	誤差 σ_i	重み w_i
1個	1σ	1
2個	$\frac{n-1}{n-2}\sigma$	$\frac{n-2}{n-1}$
3個	$\frac{n-1}{n-3}\sigma$	$\frac{n-3}{n-1}$
⋮		
n-1個	$(n-1)\sigma$	$\frac{1}{n-1}$

$$\sigma_{rmsn} = \frac{\sqrt{1 + 1 \times 2 + 1 \times 3 + \dots + 1 \times (n-1)}}{\frac{n-1}{n-1} \times 1 + \frac{n-2}{n-1} \times 2 + \frac{n-3}{n-1} \times 3 + \dots + \frac{n-(n-1)}{n-1} \times (n-1)} \sigma \approx \frac{\sqrt{18}}{n} \sigma$$

N: 画像数

σ : 両端画像による計測誤差

図1 画像間距離の重み付けと奥行き計測誤差の関係

5. 検証実験

上記提案手法の有効性を検証するために、実験システムを構築し、検証実験を行った。実験システムと結果、考察について以下に記す。

5-1. 実験システム

今回構築した実験システムは、HDTV(High Definition TV)カメラ、カーブブーム装置、ロータリーテーブル、光源(フラッドライト)、VDR(Video Disk Recorder)、WS(Work Station)から成る。ロータリーテーブル上にオブジェクト

を置き、ロータリーテーブルを回転させながら撮像することにより、1台のHDTVカメラでオブジェクトの多視点画像を容易に取得できる。使用する光源は2台であり、本提案手法で注目する縦中央部にスペキュラーが生じないように配置する。撮像された多視点画像を1度VDRに蓄積し、WSに転送後、本提案手法により奥行き計測を行う。なお本実験システムのロータリーテーブルの角度位置の精度は 0.1° であり、今回の実験ではカメラの光軸とロータリーテーブルの回転軸が直交するように固定した。

今回の実験において、多視点画像マッチング法で使用する画像は、注目画像と回転角 $\pm 10^\circ$ の範囲で 2° 毎に撮像された画像の計11枚とした。したがってマッチングの目標精度は約 $0.07(=0.4\sin 10^\circ)$ 画素となる。

5-2. 結果と考察

最大直径約8cmの一輪挿の形状計測を行った。図2に原画像、図3および4に1次マッチングおよび2次マッチングの結果を示す。1次マッチングは ± 1 画素の精度で、2次マッチングはサブピクセルの精度でそれぞれ滑らかに計測できている。また2次マッチング結果を曲線近似し、その近似曲線と2次マッチング結果の水平方向の差から誤差を検証したところ、平均誤差は約0.05画素であり目標精度を満足していると考えられる。

次にある点におけるシフト量と評価関数値のグラフを示す。図5は窓関数の有無による比較(1次マッチング)を、図6は重み付けの有無による比較(2次マッチング)を行っている。いずれもグラフ形

状の比較のため、ノーマライズしている。グラフから分かるように、窓関数あるいは重み付けの導入によりピークが急峻になり、選択性が向上していることが確認できる。

最後に、 4° 毎の2次マッチング結果をオブジェクトの奥行きに変換してボクセルに投影し、2次元に射影したものを、図7に示す。今回、多視点画像は平行投影されたものと仮定し、2次マッチング結果を $\sin 10^\circ$ で除算することにより、簡易的にオブジェクトの奥行きを求めた。

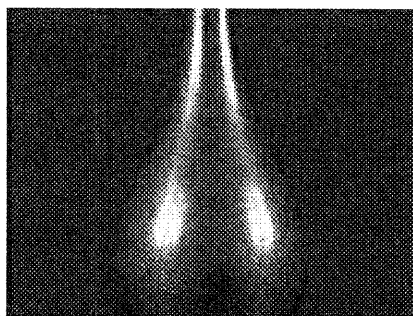


図2 原画像

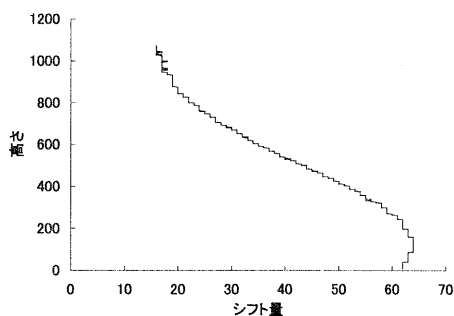


図3 1次マッチングの結果

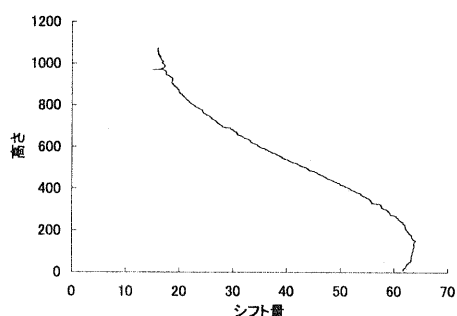


図4 2次マッチングの結果

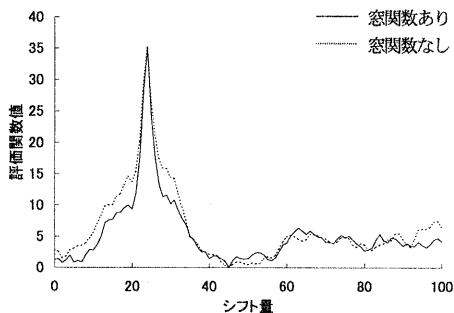


図5 窓関数の有無の比較

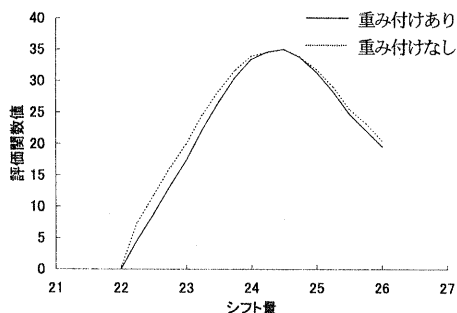


図6 重み付けの有無の比較

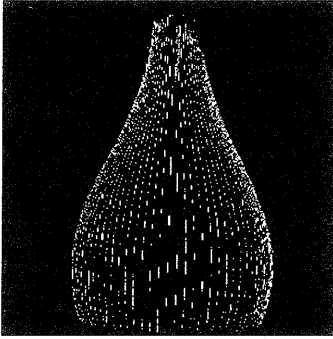


図7 形状の投影図

[3] 蓼沼他, “画像内オブジェクトの統計量解析に基づいた多眼立体カメラの最適配置法の検討”, 2000年映像情報メディア学会技術報告, Vol.24, No.19, pp7-12

6. まとめ

今回、高精細な形状計測を目標とし、多視点画像マッチング法の検討・実験を行った。実験の結果、 $\pm 10^\circ$ 内の画像を用いて平均誤差約0.05画素の精度で対応点探索ができた。この値は正対時の形状計測では0.4画素以下となる。しかしオブジェクトの形状が小さく、回転による画像のシフト量が小さい部分では、マッチングの誤検出が生じることも分かっており、ズームアップ画像などによる計測も必要である。この際、各部分毎に計測した結果をマージする技術が求められる。

今後の課題としては、上記マージ技術の検討のほか、計測が不能であった点の補間、真値と計測値の比較による計測精度の検証、オクルージョンへの対応が挙げられる。

7. 謝辞

本研究開発は通信・放送機構の委託研究「高精細・立体・臨場感コンテンツ技術の研究開発」の一環であり、厚く感謝します。

参考文献

- [1] 小林他, “高精細立体映像部品作成のための多視点画像マッチング法”, 2000年電子情報通信学会総合大会予稿, p132
- [2] 蓼沼他, “立体画像における奥行き情報の擾乱と画質の関係”, 1998年映像情報メディア学会年次大会予稿, 14-6