

三次元情報の整合性を考慮したマルチカメラによる形状復元法*

川上 耕平[†] 小田倉 聡司[†]
 東北大学大学院情報科学研究科[†]

阿部 亨[‡] 木下 哲男[‡]
 東北大学情報シナジー機構[‡]

1 はじめに

異なる視点から撮影された2枚の画像を用い三次元形状の復元を行うステレオ法には、画像間での対応付けの曖昧さや遮蔽などに起因するいくつかの問題が存在する。これらの問題の解決を図るために、3枚以上の画像を用い形状復元を行う多眼ステレオ法が提案されている。

Marrらは、立体視の基本仮定（適合性、一意性、連続性）により2枚の画像間で唯一の対応が得られることを示し、これを用いて立体照合を行うアルゴリズムを提案した[1]。この基本仮定は、2枚の画像を用いるステレオ法では効果的に働くものの[2]、3枚以上の画像を用いる多眼ステレオ法には、そのままでは適用できない場合がある。

本稿では、3枚以上の画像を用いる場合へも適用できるよう拡張された適合性と一意性を用い、処理の過程で得られる三次元情報の整合性を考慮することで、三次元形状の復元をより安定に行う新たな多眼ステレオ法を提案する。

2 提案手法

ここでは、 N 台のカメラを用い、各視点 $o_n (n = 1, 2, \dots, N)$ から撮影された画像 I_n から、三次元形状を復元するものとする。提案手法では、計測範囲である三次元空間をボクセル $v(x, y, z)$ に分割し、全ボクセルの集合を V で表す。ボクセル v が物体境界に対応する確からしさを $e(v)$ とし、各ボクセル

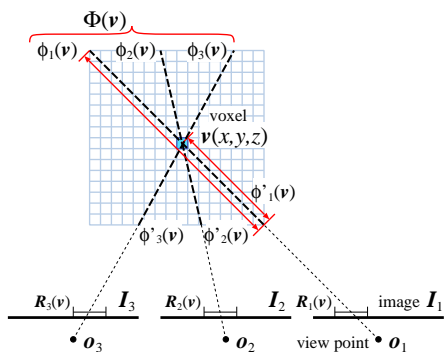


図1 提案手法における視点とボクセルとの関係

ル毎に三次元空間全体で整合が取れるように e の修正を繰り返し、修正された確からしさをもとに物体境界に対応するボクセルを判定することで三次元形状の復元を行う。

提案手法における視点とボクセルとの関係を図1に示す。 v と各 o_n とを結ぶ直線上にあるボクセルの集合を $\phi_n(v)$ 、 v を始点とし各 o_n を通る半直線上にあるボクセルの集合を $\phi'_n(v)$ とし、各 v に対する $\cup_{n=1}^N \phi_n(v)$ を $\Phi(v)$ と表す。

2.1 ボクセル v が画像 I_n で可視となる確からしさ $w_n(v)$

適合性を拡張するために、各ボクセル v について、 v が各画像 I_n で可視となる確からしさ $w_n(v)$ を導入する。

$w_n(v)$ は $\phi_n(v)$ 内のボクセルの状態から計算する。 e が高いボクセルが $\phi'_n(v)$ 内に存在する場合、 v はそのボクセルに遮蔽されて I_n から不可視である可能性が高い。そこで、提案手法では、式(1)により $w_n(v)$ を計算する。

$$w_n(v) = 1 - \frac{A_n(v) - \min_{u \in \phi_n(v)} e(u)}{\max_{u \in \phi_n(v)} e(u) - \min_{u \in \phi_n(v)} e(u)} \quad (1)$$

e が極大となるボクセルの周辺で w_n が低くなることを防ぐため、向きを o_n から v へ設定した軸を考え、 $\phi'_n(v)$ 内で e が増加しない箇所での e の最大値を $A_n(v)$ として用いる。

2.2 ボクセル v が物体境界に対応する確からしさ $e(v)$

提案手法では、各ボクセル v を入力画像へ投影した箇所周辺の画素値の一致度 $S(v)$ で適合性を評価し、この $S(v)$ に基づいて、ボクセルが物体境界となる確からしさを計算する。

ボクセル v の重心を I_n に投影した点を中心とした矩形領域 $R_n(v)$ を設定し、全ての画像対 I_n, I_m について $R_n(v), R_m(v)$ 間の画素値の差の2乗和 $SSD_{m,n}(v)$ から画素の一致度 $S(v)$ を式(2)により計算する。

$$S(v) = \frac{\sum_{1 \leq m < n \leq N} w_{m,n}(v) \times SSD_{m,n}(v)}{\sum_{1 \leq m < n \leq N} w_{m,n}(v)} \quad (2)$$

式(2)では、各画像対のSSDに重み $w_{m,n}(v)$ ($w_m(v), w_n(v)$ で値の小さい方)を与えている。これにより、 v が可視である可能性が低い画像の影響により $S(v)$ が低くなることを防ぎ、3枚以上の画像を用いる場合でも適合性が正しく評価できるようにしている。

$S(v)$ が高いボクセルは、物体境界に対応する可能性が高いと考えられるため、 $e(v)$ を式(3)により計算する。

$$e(v) = 1 - \frac{S(v) - \min_{u \in V} S(u)}{\max_{u \in V} S(u) - \min_{u \in V} S(u)} \quad (3)$$

2.3 三次元情報の整合性を考慮した $e(v)$ の修正

三次元空間全体で整合が取れるよう、ボクセル v が物体境界に対応する確からしさ $e(v)$ を式(4)により修正する。

* Shape reconstruction method considering consistency of 3D information

[†] Kouhei KAWAKAMI and Satoshi ODAKURA, Graduate School of Information Sciences, Tohoku University

[‡] Toru ABE and Tetsuo KINOSHITA, Information Synergy Center, Tohoku University

$$e_{t+1}(v) = \begin{cases} e_0(v) \times B_t(v)^\alpha, & B_t(v) > \beta \\ e_0(v) \times \beta^\alpha, & \text{otherwise} \end{cases} \quad (4)$$

ここで、 $e_0(v)$ は、更新される初期値を表し、式 (3) で求めた $e(v)$ を用いる。 α, β は定数 ($\alpha > 1, 0 < \beta < 1$) である。 $B_t(v)$ は、 t 回目の修正における $\Phi(v)$ 内での v の状態 $e_t(v)$ から決定する。2枚の画像を用いるステレオ法では、 $B_t(v)$ には v の一意性を評価した値が用いられ、この場合、物体境界に対応しかつ画像対で可視となるボクセルが、1つの $\phi_n(v)$ 上に1つ以下しか存在しないことが仮定される [2]。しかし、実際には、物体境界に対応するボクセルが $\phi_n(v)$ 上に複数存在する場合があります、3枚以上の画像を用いる場合には、それらが異なる画像対で各々可視となり得る。そこで、提案手法では、このようなボクセルでも確からしさが高くなるよう、 e の修正を行う。まず、式 (5) により、 $\phi_n(v)$ 内での $e(v)$ の相対的な大きさを b_n として求め、式 (6) により、 $\phi_n(v)$ ごとに w_n で重み付けした b_n の総和を $B_t(v)$ として用いる。

$$b_n(v) = \frac{e_t(v) - \min_{u \in \phi_n(v)} e_t(u)}{\max_{u \in \phi_n(v)} e_t(u) - \min_{u \in \phi_n(v)} e_t(u)} \quad (5)$$

$$B_t(v) = \frac{\sum_{n=1}^N w_n(v) \times b_n(v)}{\sum_{n=1}^N w_n(v)} \quad (6)$$

2.4 物体境界に対応するボクセルの判定

更新された e 及び w_n に基づき、各 v が物体境界に対応するか否かの判定を行う。提案手法では、雑音などによる影響を除くため、着目する v の $e(v)$ が各 $\phi_n(v)$ 上で極大となるかを調べ、極大となる $\phi_n(v)$ における $w_n(v)$ の総和 W が閾値以上であれば、 v が物体境界に対応すると判定する。これにより、物体境界に対応する可能性が高く (e が極大)、かつ、複数の画像で可視となっている可能性も高い (W が閾値以上) のボクセルを物体境界として選択する。

2.5 提案手法による処理の流れ

以下に提案手法による処理の流れを示す。

Step 0. 各 v において、 v が物体境界に対応する確からしさ $e(v)$ を 0 に設定。

Step 1. 各 v において以下の値を計算。

- $\phi_n(v)$ 内のボクセルの e から、 v が I_n で可視となる確からしさ $w_n(v)$ を計算。
- v を各 I_n へ投影した箇所近傍 $R_n(v)$ の画素と $w_n(v)$ とを用い画素の一致度 $S(v)$ を計算。
- $S(v)$ に基づき $e(v)$ を再計算。

Step 2.

- 各 v において、 $\Phi(v)$ 内のボクセルの e に基づき $e(v)$ を修正し、この処理を V 全体に対し繰り返す。
- Step 1, 2 の反復を続ける場合は Step 1 へ、反復を終了する場合は Step 3 へ。

Step 3. 各 v における $e(v)$ と $w_n(v)$ とに基づき、 v が物体境界に対応するか否かを判定。

3 実験

提案手法の有効性を検証するために、実画像を対象として、三次元形状を復元する実験を行った。実験では4台のカメラを使用し、図2のような画像を撮影して入力画像とした。

図3は、提案手法により物体境界と判定されたボクセルを三次元空間にプロットし、異なる方向から表示したものである。再現率 $R = |C \cap D|/|C|$ と適合率 $P = |C \cap D|/|D|$ により復元精度を定量的に評価した結果、 $R = 0.86, P = 0.89$ となった。ここで、 C は実際に物体境界に対応するボクセルの集合を、 D は提案手法より物体境界であると判定されたボクセルの集合を表し、 $|E|$ は集合 E に含まれるボクセルの個数を表す。

4 おわりに

本稿では、処理の過程で得られる三次元情報の整合性を考慮することで、三次元形状の復元をより安定に行う新たな多眼ステレオ法を提案した。提案手法を用いて実画像を対象とした実験を行い、その結果の評価を行った。

今後、提案手法の有効性をさらに高めるために、物体境界に関する連続性を適切に導入するなど、各ボクセルにおける確からしさをより効果的・効率的に更新する手法について検討を進める予定である。

参考文献

- [1] D.Marr. “ビジョン -視覚の計算理論と脳内表現-”. 産業図書, 1983. 乾敏郎, 安藤広志 訳.
- [2] C. L. Zitnick and T. Kanade. “A cooperative algorithm for stereo matching and occlusion detection,”. *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 22, No. 7, pp. 675–684, 2000.

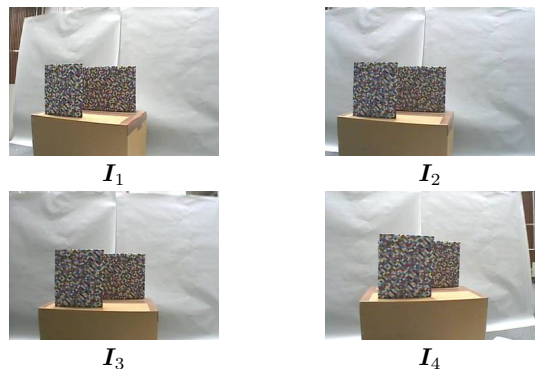


図2 入力画像



図3 物体境界と判定されたボクセル