

# 画像特徴点の数え上げに基づくマルチベースラインステレオ

佐藤 智和<sup>†1</sup> 横矢 直和<sup>†1</sup>

本論文では、マルチベースラインステレオ法における高速な奥行き探索を実現するための新たな評価尺度 TNIP (Total Number of Interest Points) を提案する。提案手法では、画像上の輝度値情報を直接用いず、画像特徴点の座標の幾何学的な整合性のみから奥行き値を算出することで、高速かつ省メモリな奥行き探索を実現する。また、TNIP は遮蔽による影響を受けにくいいため、従来の一般的な SSSD (Sum of Sum of Squared Differences) を用いる手法に対して遮蔽にロバストな推定を実現できる。ただし、TNIP による奥行き推定は、その推定原理から注目画素に対する正しい奥行き値よりもむしろ三次元空間中の特徴点位置に対応する奥行き値を算出するため、奥行き推定精度が SSSD に対して若干劣る場合がある。そこで本論文では更に、TNIP によって高速に推定された奥行き値を利用し、大幅に限定された探索範囲を SSSD によって再探索することで、効率的かつ高精度な奥行き推定を実現できることを示す。なお、本論文では、動画撮影時のカメラの位置・姿勢パラメータおよびカメラ内部パラメータは推定済みで既知とする。

## Multi-baseline Stereo by Counting Interest Points

TOMOKAZU SATO<sup>†1</sup> and NAOKAZU YOKOYA<sup>†1</sup>

This paper proposes a novel method for estimating depth without similarity measures such as SSD and NCC. Our idea for estimating a depth map is very simple; only counting interest points in images is integrated with the framework of multi-baseline stereo. Even by a simple algorithm, depth can be determined without computing similarity measures such as SSD and NCC that have been used for conventional stereo matching. The proposed method realizes robust depth estimation against occlusions with lower computational cost. Though a naive TNIP based method can realize fast and robust depth estimation, the accuracy of estimated depth is lower than one by SSSD based method because TNIP uses sparse data. In this paper, we also show that accuracy of depth estimation can be increased by combining TNIP based method and SSSD based method. In experiments, the validity and feasibility of our algorithm are demonstrated for both synthetic and real outdoor scenes.

### 1. はじめに

ステレオ画像計測に基づく奥行き推定は、三次元形状復元、物体認識、拡張現実感等に用いられており、コンピュータビジョンの分野において最も重要な課題の一つである。このため現在においても、ステレオ画像計測に基づく奥行き推定法は広く研究されている。本論文では、これらステレオ画像計測に基づく奥行き推定手法の中でも、複数の画像を統合的に扱うことで、比較的精度の高い奥行き推定を実現できるマルチベースラインステレオ法について、新しい手法の提案を行う。

Okutomi らによって提案されたマルチベースラインステレオ法<sup>1)</sup> は、複数の画像を一度に扱うことでス

テレオ法における対応点探索のあいまい性を低減し、奥行き推定精度を向上させるという特長を持ち、現在広く利用されている<sup>2)-6)</sup>。また、近年のカメラキャリブレーション手法の発展により、自由な移動を伴って撮影された動画像や、広範囲に配置された複数のカメラによって撮影された画像群に対しても、マルチベースラインステレオ法の適用が行われている<sup>3)-5)</sup>。なかでも、移動を伴って撮影した動画像を用いる手法<sup>4),5)</sup>では、長いベースラインを容易に確保することができるため、カメラ間の距離に対して撮影対象までの距離が相対的に長くなる屋外環境の三次元モデリングに適している。しかし、このような屋外環境を対象とする場合、従来手法で用いられてきたマルチベースラインステレオの評価尺度 SSSD (Sum of SSD) では、以下のような問題が発生する。

(1) 遮蔽物の影響による誤推定の発生: あるフレームに撮像されている対象物体が、他のフレームにおいて

<sup>†1</sup> 奈良先端科学技術大学院大学 情報科学研究科  
Nara Institute of Science and Technology

遮蔽されている場合, SSSD に遮蔽物と対象物体に関する類似度評価値が加えられてしまうため, 対象物体の正しい奥行き値を算出することが困難となる.

(2) 計算コストの増大: 動画像を入力とした場合には, 一度に多くの画像を用いて奥行きを推定することで推定精度の向上を図ることが可能となるが, その反面, 計算コストや必要メモリ容量の増大といった問題が発生する. また, 上記の問題 (1) に対する改善策を推定手法に加えると, 奥行き推定における計算コストの問題はさらに深刻化する.

このような問題を回避するため, 本研究では, 図 1 に示すような画像特徴点を用いる新しい奥行き推定手法を提案する. 本論文で提案する奥行き推定手法の基本的な枠組みは, 一般的なマルチベースラインステレオ法と同一であるが, 従来用いられてきた SSSD の代わりに, 画像特徴点の数上げに基づく新しい評価尺度 TNIP (Total Number of Interest Points) を用いることで, 画像の変形や遮蔽にロバストな奥行き推定を実現する. 本手法の基本的なアイデアは, “三次元空間中に特徴点が存在する場合には各画像への投影座標周辺に特徴点が存在する” という仮定に基づいている. これにより, マルチベースラインステレオにおける奥行き探索時に, 各画像上への投影座標に最も多くの画像特徴点が存在する点を探索することで奥行き値を決定する. 図 2 に, 提案手法による奥行き画像推定の処理の流れを示す. 本手法では, まず全ての入力画像上で画像特徴点を検出する (A). 次に, 新しい評価尺度 TNIP を用いたマルチベースラインステレオ法によって画像特徴点の奥行き推定を行う (B). 最後に, 推定結果の信頼度を算出することで誤推定結果を排除する (C).

提案手法を用いることで, 従来手法で問題となる二つの問題を解決することができる. すなわち, (1) 新たな評価尺度 TNIP は遮蔽物による影響を受けにくくロバストな奥行き推定を実現できる, (2) 特徴点の数を数え上げるだけで奥行きを決定できるため省メモリで計算コストが小さい. ただし, 本手法においては画像上で特徴点が存在する位置においてしか奥行きを算出できないという手法上の制約が新たに発生するが, 一般的な屋外環境の三次元モデリングなどの分野においては, 特徴点間の奥行き値は補間で十分な場合も多い. また, 本研究で提案する TNIP は高速かつロバストな奥行き推定を実現できるという特長を持つ反面, 注目画素に対する正しい奥行き値よりもむしろ三次元特徴点に対応する奥行き値に近い値を算出するという特性を持つため, 精密な奥行き決定能力に乏しいとい



図 1 画像特徴点の検出例

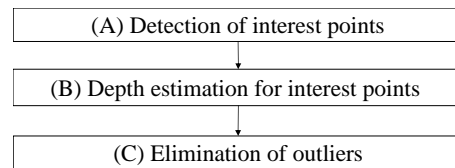


図 2 奥行き画像推定の処理の流れ

う問題があるが, TNIP を用いて算出される奥行き値を用いて探索範囲を限定した上で SSSD による精密な探索を行なうことで, 低い計算コストを維持したまま高精度な奥行き推定を実現できることを示す.

以下, 本論文では, まず 2 章において本研究に関連が深い特徴点を利用したステレオ法と三次元復元手法に関する従来手法, およびマルチベースラインステレオ法に関する従来手法を概観し, 本研究の位置づけについて述べる. 続いて, 3 章において本手法の基礎となる SSSD を用いる従来のマルチベースラインステレオ法について解説する. 4 章では, マルチベースラインステレオに用いる新たな評価尺度 TNIP を提案する. また, 5 章では, TNIP を用いた実際の奥行き推定処理であるステージ (A) からステージ (C) について詳述する. 6 章では, シミュレーションおよび実環境を用いた実験結果を示し, 提案手法の有用性を示す. 最後に, 7 章でまとめと今後の課題について述べる.

## 2. 関連研究

### 2.1 特徴点を用いたステレオ法と三次元復元手法に関する従来研究

奥行き画像推定に関しては, 古くから 2 眼や 3 眼ステレオ画像に関する様々な手法が提案されてきた. これらの手法は, 画素ごとの輝度値の差と滑らかさの制約によって定義される誤差関数を最小化することで密な視差マップを生成する pixel-based stereo 法, 注目画素周辺のパターン類似度を用いて視差を決定する area-based stereo 法, エッジ形状等の特徴情報を用

いて対応点を決定する feature-based stereo 法, に大別できる. これら 2 眼・3 眼ステレオ法のなかでも本研究に深く関連するのは, 画像上のエッジ等の特徴点を利用することで, 頑健な奥行き推定を行う feature-based stereo 法<sup>7)-9)</sup> である. feature-based stereo 法では, まず濃淡エッジ等の特徴点を各画像から検出し, 検出された特徴点同士の対応関係を特徴点周辺におけるテクスチャやエッジ形状を手がかりに決定する. 本研究では, feature-based stereo 法と同様に特徴点を用いた探索を行なうことで高速で頑健な奥行き推定を目指す, 特徴点の画像上での位置以外の情報(テクスチャやエッジ形状)を用いず, 特徴点位置の幾何学的な整合性のみから奥行き推定を行うという点で, これらの手法とは一線を画するものである.

また, 同様にエッジ特徴を用い, 多数の画像から三次元復元を行う手法として, EPI(Epipolar Plane Image) 解析による手法<sup>10),11)</sup> が挙げられる. この手法では, カメラの光軸に対して垂直方向に等速にカメラを移動させて得られる画像系列に対して, 画像面の横軸と時間軸の 2 軸で展開される EPI 画像を作成し, EPI 画像から Hough 変換によって直線検出を行うことで複数画像におけるエッジの対応点を探索し, 奥行きを推定する. また, 類似したアイデアとして, 等速回転する回転台上の物体を撮影した動画から EPI 画像を作成し, 輝度値が一定の値をとる sin カーブを投票によって検出することで三次元復元を行う手法<sup>12)</sup> が提案されている. これらの手法<sup>10)-12)</sup> では, 撮影対象に対するカメラ運動の規則性を利用して三次元復元を行っており, カメラの動きに制約が生じる. これに対して, 画像上の特徴点の動きからカメラの運動パラメータを推定し, EPI による手法が前提としている規則的なカメラの動きからの逸脱を補正する手法<sup>13)</sup> も提案されているが, 光軸方向に関するカメラ移動の補正は原理的に難しく, 本研究が扱う自由なカメラ移動に対して適用できるものではない.

このような Hough 変換による三次元復元の枠組みを, より一般的なカメラ運動に対応させた手法としては, 川戸<sup>14)</sup> による空間への投票に基づく手法が挙げられる. この手法では, カメラ位置・姿勢をキャリブレーション済みの画像系列に対して, エッジ情報をボクセル空間に投票し, 投票数を基準とした閾値処理によって三次元形状を推定している. また, ボクセル空間をオクトツリーとして表現することで, 計算量の削減を図っている. しかし, 広域な屋外環境を対象とした場合には, ボクセル表現による三次元形状の復元は計算コストが膨大になるという問題があり現実的な

い. また, 一般的にはカメラのキャリブレーション誤差によって投票結果がボクセル空間内で分散するため, ボクセル空間における物体の存在判定に用いる閾値を適切に決定することが難しく, 推定誤差の抑制と形状欠損のトレードオフの問題が残されている. これに対して, 本研究では特徴点数が最大となる点を探索することで奥行きを一意に決定するため, 奥行き探索において特別な閾値が必要ないという特長がある.

他方, structure from motion による手法<sup>15)</sup> では, 複数画像上の画像特徴点を対応付けることで, カメラに対する特徴点の三次元位置だけでなくカメラの位置関係をも推定することができる. 提案手法においても, カメラキャリブレーション手段の一つとして structure from motion を想定しているが, 画像上での追跡が難しい特徴点については三次元位置が推定されず, 得られる三次元点群は比較的疎なものとなる. このため, structure from motion で得られる三次元点群を奥行き画像推定や屋外三次元モデリングにそのまま利用することは難しい.

## 2.2 マルチベースラインステレオ法の改良に関する従来研究

前章で述べたように, 本研究の基礎となる従来のマルチベースラインステレオ法では, 遮蔽による誤推定の問題が発生するが, 現在までに誤差関数 SSSD の算出方法を改良することでこの問題の回避を試みる手法が複数提案されている. Okutomi ら<sup>6)</sup> は, マルチベースラインステレオ法にアダプティブウインドウ<sup>16)</sup> を用いることで遮蔽境界付近における奥行き推定精度の向上を図っている. Sato ら<sup>5)</sup> や Kang ら<sup>17)</sup> は, SSSD 算出時において算出される SSD から類似度の高い半数のみを選択し, 新たに SSSD として用いることで, 遮蔽による影響を軽減する手法を提案している. Sanfourche ら<sup>18)</sup> は, 一般に photo consistency と呼ばれるような, 対応点における画素の輝度値の分散を最小化する指標を SSSD に代わる評価尺度として用い, Kang ら<sup>17)</sup> の遮蔽の判定手法と組み合わせる手法を提案している. また, この手法ではカメラキャリブレーションの誤差を考慮し, 評価関数の算出に用いる投影位置を画像毎に 4 近傍にシフトすることで奥行き推定誤差の低減を図っている. このように, 従来提案されているマルチベースラインステレオ法の拡張手法は, 画像上での輝度パターンを評価尺度として直接比較する area-based stereo 法に近い枠組みで構成されており, 本論文で提案する feature-based な手法とは全く異なるものである. 本研究では, feature-based stereo 法で用いられてきた特徴点位置の情報をマルチ

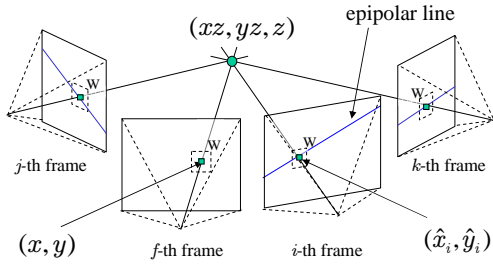


図3 注目フレームにおける画素  $(x, y)$  の三次元位置と各画像上への投影直線

ベースラインステレオ法の枠組みで利用することで、効率的な奥行き推定を実現する。

### 3. 類似度計算に基づくマルチベースラインステレオ

本章ではまず、マルチベースラインステレオ法で用いる座標系に関する定義を行う。次に、通常の類似度評価尺度 SSSD を用いたマルチベースラインステレオ法<sup>1)</sup> の基本原理について概説する。

#### 3.1 座標系の定義

移動を伴って撮影した動画像に対するマルチベースラインステレオ法では、図3に示すように、第  $f$  フレームにおける画素  $(x, y)$  の奥行き値  $z$  を第  $j$  フレームから第  $k$  フレームまでの画像を用いて推定する ( $j \leq f \leq k$ )。以下では、記述の簡単化のために、カメラの焦点距離を1とし、レンズ歪み等はカメラ内部パラメータを用いて補正済みであるものとする。このとき、第  $f$  フレームにおける画素  $(x, y)$  の奥行き値  $z$  に対応する三次元位置は、第  $f$  フレームのカメラ座標系において  $(xz, yz, z)$  と表現される。この三次元座標  $(xz, yz, z)$  は、以下の式によって第  $i$  フレーム上の画像座標  $(\hat{x}_i, \hat{y}_i)$  に投影される。

$$\begin{pmatrix} a\hat{x}_i \\ a\hat{y}_i \\ a \\ 1 \end{pmatrix} = M_{fi} \begin{pmatrix} xz \\ yz \\ z \\ 1 \end{pmatrix} \quad (1)$$

ただし、 $a$  は媒介変数である。また、 $M_{fi}$  は第  $f$  フレームのカメラ座標系から第  $i$  フレームのカメラ座標系への  $4 \times 4$  の変換行列を表す。本研究において、カメラ外部パラメータ  $M_{fi}$  は、何らかの手法によって推定済みで既知としている。図3に示すように、マルチベースラインステレオ法においては、三次元座標  $(xz, yz, z)$  の第  $i$  フレームへの投影座標  $(\hat{x}_i, \hat{y}_i)$  は、各画像上においてエピポーラ線上に拘束される。

#### 3.2 SSSD を用いた奥行き推定

本節では、従来のマルチベースラインステレオ法の基本原理を概説し、移動撮影された動画像を対象とした場合の問題点を明らかにする。従来のマルチベースラインステレオ法では、SSD によって2枚の画像上における一定サイズのウィンドウ  $W$  内のパターン間の類似度を評価し、これらの総和 SSSD を最小化することで画素  $(x, y)$  の奥行き値  $z$  を決定する。第  $f$  フレームにおける画素  $(x, y)$  と、第  $i$  フレームにおける画素  $(\hat{x}_i, \hat{y}_i)$  周辺のパターンの非類似度 SSD は、以下のように定義される。

$$SSD_{fxy}(z) = \sum_{(u,v) \subseteq W} \{I_f(x+u, y+v) - I_i(\hat{x}_i+u, \hat{y}_i+v)\}^2 \quad (2)$$

ただし、 $I_f(x, y)$  は第  $f$  フレームにおける画素  $(x, y)$  の輝度値を、 $(\hat{x}_i, \hat{y}_i)$  は式(1)によって算出される三次元座標  $(xz, yz, z)$  の第  $i$  フレーム画像上への投影座標を表す。また、SSD の総和 SSSD (Sum of SSD) は以下のように定義される。

$$SSSD_{fxy}(z) = \sum_{i=j}^k SSD_{fxy}(z) \quad (3)$$

一般的なマルチベースラインステレオ法では、SSSD を最小にする奥行き値を次元探索することで、第  $f$  フレームにおける画素  $(x, y)$  の奥行き値  $z$  を決定する。ただし、SSSD の大域最小解を得るためには、対象物体が存在すると考えられる範囲内で奥行き値を網羅的に探索する必要がある。

ここで、第  $f$  フレームにおける画素  $(x, y)$  上に撮像されている物体が、第  $i$  フレームの画像上において他の物体に遮蔽されている場合には、画素  $(x, y)$  における真の奥行き値  $z$  に対応する第  $f, i$  フレーム間の非類似度  $SSD_{fxy}(z)$  は大きな値をとる。これに伴って、 $SSSD_{fxy}(z)$  も大きな値となるため、真の奥行き値  $z$  において SSSD が大域最適解とはならず、遮蔽が発生する場合には正しい奥行き値を推定することが困難となる。このような問題を解決するためには、各フレームに対して算出される SSD の中から類似度の高い半分のみを SSSD として再定義すること<sup>5), 17)</sup> が有効であるが、計算コストが大幅に増加するという問題が残されている。

#### 4. 特徴点の数え上げに基づくマルチベースラインステレオ

本章では、一般的なマルチベースラインステレオ法の枠組みを基礎として、画像上の特徴点の数え上げを行うことで高速に奥行きを推定する手法について詳述する。一般に、三次元空間における物体の角や交点および物体表面のテクスチャにおける特徴的な点(三次元特徴点)は、撮影画像上においても輝度エッジのコーナや交点として観測される。このような画像特徴点は、Harris オペレータ<sup>19)</sup> や Moravec オペレータ<sup>20)</sup> を用いることで容易に検出することができる。

本研究では、三次元特徴点の各画像への投影座標において、高い頻度で画像特徴点検出されることに着目し、以下の新しい評価尺度 TNIP を最大化する奥行き値  $z$  を一次元探索する。

$$TNIP_{f_{xy}}(z) = \sum_{i=j}^k \sum_{(u,v) \subseteq W} H_i(\hat{x}_i + u, \hat{y}_i + v) \quad (4)$$

$$H_i(\hat{x}_i + u, \hat{y}_i + v) = \begin{cases} 1; & \text{interest point exists} \\ & \text{at } (\hat{x}_i + u, \hat{y}_i + v) \\ & \text{in } i\text{-th frame} \\ 0; & \text{otherwise} \end{cases} \quad (5)$$

TNIP は、画素  $(x, y)$  の奥行き値  $z$  に対応する各画像上での探索位置  $(\hat{x}_i, \hat{y}_i)$  付近に存在する画像特徴点の総数を表している。ただし、一般に画像特徴点の検出には検出誤差が生じるため、本手法では特徴点検出の誤差を考慮した一定サイズのウィンドウ  $W$  内に存在する特徴点の数え上げる。

SSSD の代わりに TNIP を用いて奥行き値  $z$  を探索することで、計算コストのかかる類似度評価を省略することができる。加えて、遮蔽が起こった場合においても、SSSD の場合とは対照的に、それが直接 TNIP の値に対する大きなペナルティーとはならないため、正しく奥行きを推定することが可能となる。

#### 5. 動画画像からの密な奥行き画像推定

本章では、図 2 に示した密な奥行き画像推定の各ステージについて詳述する。本研究では、まず全ての入力画像上で画像特徴点を検出する (A)。次に TNIP を用いた奥行き推定により、各画像の特徴点位置での奥行き値を算出する (B)。最後に、複数フレームでの奥行き推定結果を用い、奥行き推定の整合性によって定義される信頼度を用いることで、誤推定結果を排除する (C)。

##### 5.1 画像特徴点の検出

ステージ (A) では、画像の拡大・縮小や回転に対して頑健に画像特徴点を抽出可能な Harris オペレータ<sup>19)</sup> を用いることで、各入力画像上の輝度エッジの交点やコーナなどの画像特徴点を検出する。Schmidら<sup>21)</sup> によれば、Harris オペレータは、画像の幾何学的な変形操作を行った場合に、他のインタレストオペレータに比べて、同じ位置に画像特徴点検出される再現度が最も高いオペレータであるとされている<sup>\*1</sup>。

Harris オペレータでは、入力画像上の座標  $\mathbf{x} = (x, y)$  の特徴量  $F(\mathbf{x})$  算出のために、まずガウシアンオペレータによる入力画像の平滑化処理を行う。次に一定の大きさの正方形窓  $W$  において、画像上の輝度  $I$  の勾配  $I_x, I_y$  を用いて以下に示す行列  $\mathbf{A}$  を算出する。

$$\mathbf{A} = \sum_{\mathbf{x} \in W} \begin{pmatrix} I_x(\mathbf{x})^2 & I_x(\mathbf{x})I_y(\mathbf{x}) \\ I_x(\mathbf{x})I_y(\mathbf{x}) & I_y(\mathbf{x})^2 \end{pmatrix} \quad (6)$$

この行列  $\mathbf{A}$  を用いて、特徴量  $F(\mathbf{x})$  を以下の式により算出する。

$$F(\mathbf{x}) = \det(\mathbf{A}) - \alpha \text{trace}(\mathbf{A})^2 \quad (7)$$

ただし、 $\alpha$  はオペレータの感度を表し、本研究では経験的に  $\alpha = 0.06$  を用いる。本手法では、画像内の全ての画素  $\mathbf{x}$  において特徴量  $F(\mathbf{x})$  を算出した後に、一定サイズのウィンドウ内で特徴量  $F(\mathbf{x})$  が極大値となる点を画像特徴点として検出する。本ステージでは、動画画像の全てのフレームの画像に対して画像特徴点の座標を求める。

##### 5.2 TNIP による奥行き推定

ステージ (B) では、検出された全ての画像特徴点に対して 4 章で定義した TNIP を最大化する奥行きをそれぞれ探索することで、全ての画像上で粗な奥行き画像を算出する。ここで、奥行き値  $z$  を決定するためには、あらかじめ設定した範囲内の網羅的な探索が必要となるが、本研究では効率的な探索を実現するために、奥行き値  $z$  の探索ステップ  $l$  を可変とし、探索位置に応じて変化させる。具体的には、三次元位置  $(xz, yz, z)$  と  $(x(z+l), y(z+l), z+l)$  を結ぶ線分  $L$  に対し、各画像上で観測される線分  $L$  の長さの最大値が一定の画素幅となるように探索ステップ  $l$  を適宜決定する。TNIP に基づく奥行きの探索では、画像の輝

\*1 近年、数倍程度のスケール変化を伴う画像変換に対しても頑健に画像特徴点を抽出可能な SIFT 特徴 (Scale Invariant Feature Transform)<sup>22)</sup> が提案されているが、SIFT によって検出される画像特徴点は輝度エッジのコーナや交点から比較的離れた位置に検出される特性を持つため、本研究では Harris オペレータを採用している。



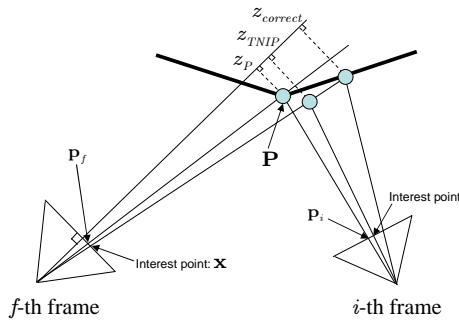


図4 TNIPによる奥行き推定の誤差要因

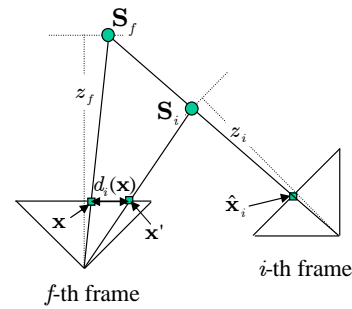


図5 誤推定結果の排除

度情報は必要なく、代わりに特徴点の画像座標が必要となる。これにより、計算機に同時に保持すべきデータの必要メモリ容量は、例えば8bit グレースケール画像に対して1/8となり、省メモリかつ高速な奥行き推定を実現できる。ただし、TNIPを用いた奥行き推定では、その推定原理からSSSDに対して奥行き推定精度が劣る場合がある。この理由について以下に詳述する。

図4に示すように、本論文で提案するTNIPは、三次元空間中に存在する特徴点Pの各画像上への投影位置 $p_i (j \leq i \leq k)$ 周辺に検出される画像特徴点を数え上げることで奥行き推定を行う。TNIPでは、奥行き推定を行う第 $f$ フレームの画像上の特徴点を注目点 $x$ とし、注目点 $x$ と第 $f$ フレームの投影中心を結ぶ直線において奥行きを探索するが、画像特徴点の検出位置には誤差が含まれるため、注目点 $x$ とPの投影位置 $p_f$ は必ずしも一致しない。このとき、図4に示すように、TNIPによって推定される奥行き値 $z_{TNIP}$ は、注目点 $x$ に対する真の奥行き値 $z_{correct}$ ではなく、特徴点の三次元位置Pに対応する奥行き値 $z_P$ に近い値になってしまう。これは、三次元空間中の特徴点位置に対応する奥行き値付近で評価値が大きくなるというTNIPの特性に起因しており、 $p_f$ と注目点 $x$ が特徴点の検出誤差によって一致しない場合には、TNIPを用いて注目点 $x$ に対する正しい奥行き値を算出することが難しい。一方、SSSDによる奥行き推定では、注目点を中心とする画像パターンを用いて奥行き推定を行うため、このような問題は発生しない。

ただし、一般に特徴点位置の検出誤差は最大でも数画素程度であるため、TNIPによって得られる奥行き値 $z_{TNIP}$ は、真の奥行き値 $z_{correct}$ に対して大きく異なるものではない。また、奥行き値 $z_{TNIP}$ は注目点 $x$ 付近に写る三次元空間中の特徴点までの奥行き値 $z_P$ に近い値であると考えられる。よって、画素単位での厳密な奥行き推定精度よりも処理コストの低さ

やロバスト性が重視される衝突回避等の一部のアプリケーションでは、TNIPによって得られる奥行き値をそのまま利用することができる。一方、高精度な奥行き推定が要求される場合には、TNIPによって推定された奥行き値 $z_{TNIP}$ を用いて、限定された探索範囲( $z_{TNIP} - C < z < z_{TNIP} + C$ )に対して再度SSSDを用いた探索を行なうことで、推定結果を高精度化することが可能である。ただし、 $C$ はSSSDによる探索範囲を決定するための定数であり、ここでは奥行き値 $z_{TNIP}$ に対応する探索ステップ幅 $l$ の定数倍の範囲内をSSSDにより再探索するものとする。SSSDによる探索をTNIPの後処理として行う場合には、画像の輝度値情報をメモリに保持しておく必要があるため本手法の省メモリ性は失われるが、SSSDにおいて探索すべき範囲は大幅に限定されており、手法の効率性が損なわれることはない。またTNIPによる奥行き推定では、特別な処理なしに遮蔽に起因する局所解を回避できることから、SSSDによる再探索処理によって高精度かつロバストな奥行き推定を実現できる。これらについては、後述するシミュレーション実験によっても有効性を示す。

### 5.3 誤推定結果の排除

従来の二眼ステレオ法においては、左眼・右眼画像をそれぞれ基準として得られる視差を相互に検証することで誤推定結果を排除するアプローチが一般的に用いられてきたが、本研究においても、各入力画像を基準として算出された奥行き推定結果の整合性を検証することによって誤推定結果の排除を行う。ここでは、まず2枚の画像における奥行き推定結果の整合性について述べる。図5に示すように、第 $f$ フレームの画素 $x = (x, y)$ において推定された奥行き値を $z_f$ 、これに対応する三次元位置を $S_f = (xz_f, yz_f, z_f)$ 、また三次元位置 $S_f$ を第 $i$ フレームに投影した画像上の座標を $\hat{x}_i = (\hat{x}_i, \hat{y}_i)$ とする。このとき、画素 $\hat{x}_i$ における奥行き推定値 $z_i$ から算出される三次元位置

$S_i = M_{if}(\hat{x}_i z_i, \hat{y}_i z_i, z_i, 1)^T$  を用いて、元の画像上への投影位置  $\mathbf{x}' = (x', y')$  を算出することで、2枚の画像  $f, i$  における奥行き推定結果の整合性を画像上での距離  $d_i(\mathbf{x}) = |\mathbf{x} - \mathbf{x}'|$  によって評価する．本研究では、第  $f$  フレームの奥行き推定に用いた全ての画像 ( $j \leq i \leq k$ ) を用い、距離  $d_i(\mathbf{x})$  が一定の閾値  $T$  以下となる割合  $R(\mathbf{x})$  を以下の式によって算出することで、奥行き推定値  $z_f$  の信頼度を評価する．

$$R(\mathbf{x}) = \frac{\sum_{i=j}^k \{0; d_i(\mathbf{x}) > T, 1; d_i(\mathbf{x}) \leq T\}}{k - j + 1} \quad (8)$$

本ステージでは、 $R(\mathbf{x})$  が一定の閾値  $U$  を下回る奥行き推定結果は信頼性が低いと判断し、削除する．なお、TNIP を用いた奥行き推定では、図5における  $S_f$  の投影位置  $\hat{x}_i$  の大半において特徴点が存在するため、 $\hat{x}_i$  において推定済みの奥行き値  $z_i$  を直接利用可能であるが、投影位置  $\hat{x}_i$  に特徴点が存在しないフレームや、ステージ (B) において SSSD を用いた推定精度の高精度化を行う場合には、 $z_i$  が既知ではないため  $d_i(\mathbf{x})$  を算出できない．そこで、本手法では  $z_i$  が直接利用できない場合には、周辺の特徴点に対する奥行き推定結果を線形補間し、整合性の判定に用いる．

## 6. 実験

本章では、SSSD を用いた推定手法、TNIP を用いた推定手法、SSSD と TNIP の組み合わせによる推定手法 (以下、HYBRID 手法) を計算機シミュレーションを用いて比較することで、それぞれの手法の特性を明らかにする．また、広域な屋外環境において奥行き画像を生成できることを示すために、屋外を撮影した実画像を用いた実験結果について示す．

### 6.1 計算機シミュレーションによる定量的評価実験

以下では、まず本シミュレーション実験に用いる実験条件について述べる．次に、SSSD, TNIP および HYBRID 手法で用いる最適なウィンドウサイズを予備実験によって決定し、最後に各手法を詳細に比較する．

#### 6.1.1 実験条件

本実験では、2枚の平面を仮想空間内に配置し、仮想ビデオカメラ (解像度:640 × 480 画素) でそれらを撮影することで、入力画像を作成した．2枚の平面のテクスチャとしては、図6に示すように、自然物である草の模様 (plane 1) と人工物である繰り返しパターンのタイル模様 (plane 2) を用いた．実験に用いた平面の配置とカメラの動きを図7に示す．仮想カメラは円弧を描くよう1度刻みで動き、91枚の画像を撮影し

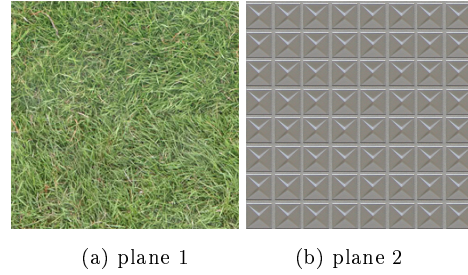


図6 対象物体として用いた平面のテクスチャ

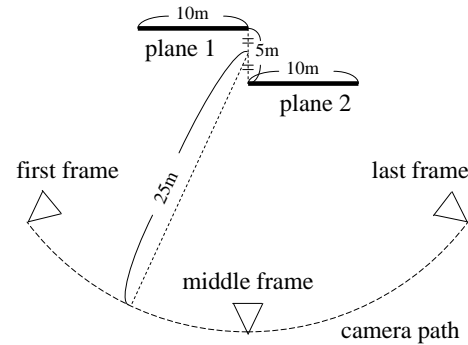


図7 シミュレーションにおける平面の配置とカメラの動き

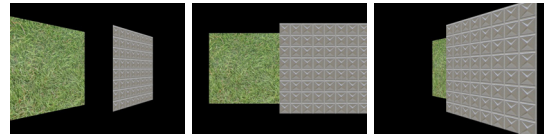


図8 仮想カメラで撮影された91枚のシミュレーション画像の一部

た．仮想カメラによって撮影された91枚の画像の一部を図8に示す．同図から分かるように、カメラの移動によって画像上でのテクスチャの見え方が大きく変化し、また動画後半の画像上では plane 1 が plane 2 によって遮蔽されている．本実験では、カメラの内部および外部パラメータのキャリブレーション誤差を考慮するために、三次元特徴点の投影位置に対して標準偏差  $\sigma$  を持つガウスノイズを付加し、これに加えて投影座標を画素単位に量子化した．また、奥行き値は、全ての入力画像上の画像特徴点に対して基準フレームを変えながら算出した．その際、探索フレームは基準フレーム以外の全ての入力画像とした．その他、本実験に用いた閾値を表1に示す．

#### 6.1.2 ウィンドウサイズの決定

式(2)および式(5)で用いる最適なウィンドウ  $W$  を決定するための予備実験について述べる．この実験では、前項に示した条件を用いてウィンドウサイズを変えながら奥行き推定を行い、出力に占める誤推定の

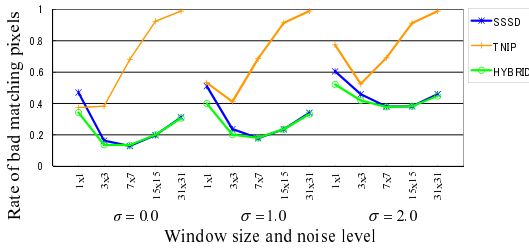


図9 全ての特徴点の奥行き推定結果に対する誤推定の発生率 ( $E \geq 1.0$  画素)

割合を用いて最適なウィンドウサイズを決定する。ここでは、特徴点  $p$  に対して出力された奥行き値に関して、対応する三次元座標を第  $f$  フレームの画像上に投影した座標  $\hat{x}_{fp}$  と、真値に対応する三次元座標を第  $f$  フレームの画像上に投影した座標  $\bar{x}_{fp}$  の差の平均  $E_p$  を算出し、画像上での平均誤差  $E_p$  が一定の画素を超える推定結果の発生率を検証する。

$$E_p = \frac{1}{N} \sum_{f=1}^N |\hat{x}_{fp} - \bar{x}_{fp}| \quad (9)$$

ただし、 $N$  は推定に用いた画像の枚数であり、本実験では  $N = 91$  である。また、この実験では 5.3 節で述べた誤推定の自動排除は行わない。

図9に、様々なウィンドウサイズ  $W$  を用いて得られた特徴点の奥行き推定結果に対して、 $E_p$  が 1.0 画素以上となる推定結果を誤推定とみなした場合の誤推定の発生率を示す。同図から、TNIP ではウィンドウサイズ  $3 \times 3$  画素を選択した場合に最も良い結果が得られるが、SSSD で誤推定率が最小となるウィンドウサイズ  $7 \times 7$  画素の結果と比較して  $E \geq 1.0$  画素となる誤推定の発生率が高いことが分かる。5.2 節で述べたように、TNIP 単体による手法では、奥行き推定を行う注目画像上での特徴点の検出誤差が奥行き推定精度を低下させてしまうという問題があるため、比較的精密な奥行き推定精度を要求する  $E \geq 1.0$  画素の場合において、誤推定の発生率が SSSD よりも高い。

表1 シミュレーション実験に用いた閾値

(a) 奥行き推定の閾値	
奥行き探索範囲	3,000mm ~ 35,000mm
探索ステップ $l$ の画像上での最大幅	1.0 画素
HYBRID 手法における再探索範囲決定係数 $C$	10

(b) 誤推定排除の閾値	
投影誤差の閾値 $T$	1.0 画素
信頼度の閾値 $R$	0.4

これに対して、TNIP で最良の結果となる  $3 \times 3$  ウィンドウによって推定された結果を用いて探索範囲を限定した上で、SSSD を用いて再探索を行なう HYBRID 手法についても図9に結果を示した。なお、HYBRID 手法に対する同図の横軸は SSSD を用いた再探索時のウィンドウサイズを表している。同図から、HYBRID 手法を用いた場合には、SSSD で用いるウィンドウサイズを  $7 \times 7$  に設定することが適切であることが分かる。また、全体での平均的な誤推定の発生率に関して、HYBRID 手法は SSSD 単体での推定結果と同程度の推定精度である。

表2に、異なるウィンドウサイズを用いた場合の、1画素の奥行き算出に要した平均時間を示す。ここでは入力として全91枚の画像を用いており、計算時間は PC (CPU: Pentium-4 Xeon 3.20GHz dual, Memory: 2GB) を用いた場合の値である。表から、TNIP ( $3 \times 3$  画素) 単体を用いた場合には、SSSD ( $7 \times 7$  画素) を用いる手法に対して約9倍程度高速に奥行き推定を実行できることが分かる。また、HYBRID ( $7 \times 7$  画素) は、SSSD ( $7 \times 7$  画素) に対して5倍程度高速である。各手法の詳細な比較については次項で述べるが、HYBRID 手法を用いることで、平均的には SSSD と同程度の精度で効率的に奥行き推定を実現することができる。なお、本実験においては、評価関数以外の奥行き探索の枠組みおよび実装は同一のものを用いているため、粗密探索法等の高速化手法を奥行き探索の枠組みに組み込んだ場合においても、計算コストに関する優劣が逆転することはないと考えられる。

表2 1画素の奥行き算出に要した平均時間 [ミリ秒]

window size	1x1	3x3	7x7	15x15	31x31
SSSD	13.6	25.2	<b>86.3</b>	353.9	1530
TNIP	9.0	<b>9.8</b>	11.2	13.0	21.2
HYBRID	10.3	11.5	<b>16.7</b>	40.1	141.3

### 6.1.3 SSSD, TNIP, HYBRID 手法による推定精度の比較

前項の予備実験において決定したウィンドウサイズ (TNIP:  $3 \times 3$  画素, SSSD:  $7 \times 7$  画素, HYBRID:  $7 \times 7$  画素) を用いて、各手法の出力を詳細に比較する。ここでは、奥行き推定対象となる全領域 (領域 ALL) を、遮蔽が発生する領域 OCC とそれ以外の領域 NOR に分け、それぞれの領域に対する画像上での誤差  $E$  の割合を、誤推定の排除前・排除後に分けて検証する。なお、本実験では奥行き推定に用いる半数以上の画像上で遮蔽される領域を遮蔽領域 OCC とする。

図10(a)に、領域 ALL に対して推定された奥行き



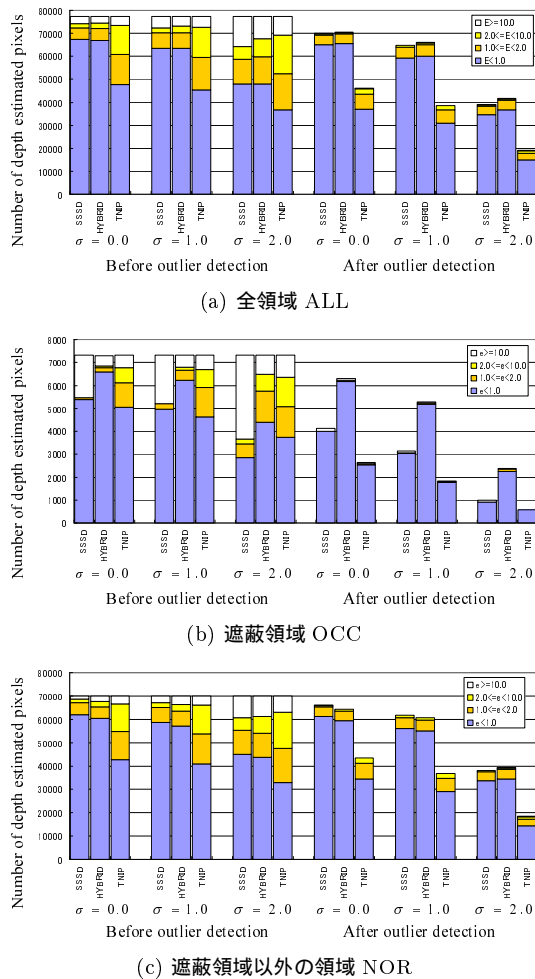


図 10 領域ごとの推定誤差の内訳

値に対する誤差  $E$  の内訳を積み上げ棒グラフで示す。縦軸は、推定された奥行き値の個数を表し、各項目は前述の推定誤差  $E$  の大きさによって分類されている。また、横軸には、5.3 節で述べた誤推定排除手法の有無、ノイズレベル  $\sigma$ 、用いた手法の別を示してある。同図 (a) 左から、領域 ALL に対する誤推定結果の排除前において、高精度な推定結果 ( $E < 1.0$  画素) の割合は SSSD と HYBRID で同程度であるが、これに対して TNIP 単体ではどのノイズレベルに対しても高精度な出力の割合が低い。しかし、 $E \geq 10.0$  画素となる大きな推定誤差の発生率については TNIP が大きく劣るといったことはなく、逆にノイズレベルが大きい場合には TNIP が SSSD に勝る。このため、TNIP による推定結果を初期値として一定範囲内を SSSD によって再探索する HYBRID 手法では、SSSD と同程度またはそれ以上の高精度な推定を実現している。

図 10(a) 右の誤推定の排除後のグラフからは、 $E \geq$

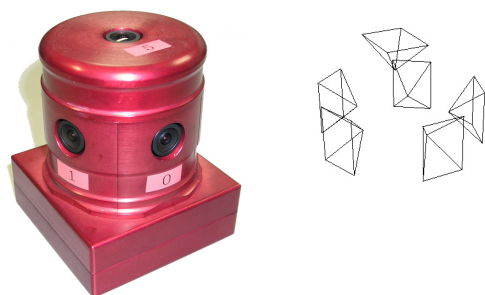
2.0 画素となるような大きな誤推定結果がどの手法においてもおおむね排除されており、5.3 節で述べた誤推定の排除手法が有効であることが分かる。ただし、ノイズレベルが大きい場合 ( $\sigma = 2.0$  画素) においては、どの手法においても誤推定排除前の推定精度が悪いため、正しく推定結果の信頼度を判定することが難しく、結果として高精度な出力まで誤推定と判定して排除されている。特に TNIP 単体では、最終的に利用できる奥行き推定結果が少ないが、SSSD による再探索を利用する HYBRID 手法では、SSSD を若干上回る数の高精度な出力を最終的に利用できる。

次に、図 10(b)(c) に、遮蔽領域 (OCC) および遮蔽領域を除く領域 (NOR) に対して推定された奥行き値に対する誤差  $E$  の内訳を示す。先に述べた領域 ALL に占める領域 OCC の割合は 1 割程度であるため、OCC に対するグラフの縦軸が他のグラフに対して 10 倍の表示スケールとなっていることに注意が必要であるが、同図 (b)(c) の比較から、領域 OCC では SSSD による大きな推定誤差の発生率が高いことが確認できる。これに対して TNIP 単体または TNIP を基礎とする HYBRID 手法では、領域 OCC においても他の領域と同程度の推定精度を維持しており、本論文で提案した TNIP による評価尺度では、遮蔽による影響に対して比較的ロバストに奥行き推定を実現できていることが分かる。

以上のことから、本論文で提案した TNIP による手法は、効率的かつ遮蔽に対してロバストな奥行き探索を実現できるが、単体での利用においては精密な奥行き推定には向かない。しかし、TNIP と SSSD による探索手法を組み合わせることで、少ない計算コストを維持したまま従来手法と同程度の推定精度を実現可能である。

## 6.2 屋外環境における奥行き画像推定

本実験では、図 11(a) に示す、PointGrey Research 社製の全方位型マルチカメラシステム Ladybug<sup>23)</sup> を用い、大学キャンパスを動きながら撮影した。図 11(b) に、マルチカメラシステムのキャリブレーション結果に基づく、各カメラユニットの位置・姿勢の関係を示す。同図から分かるように、Ladybug には水平方向に 5 つ、上向きに 1 つの合計 6 つのカメラユニットが放射状に配置されており、各カメラユニットはそれぞれ  $768 \times 1024$  画素の画像を 15fps の動画像として同期撮影できる。入力として用いた画像は、図 12 に示す 6 枚を含め 3000 枚 (500 フレーム) である。本実験において、各カメラの内部パラメータおよびマルチカメラシステムの位置・姿勢パラメータは、それぞ



(a) appearance (b) view volume

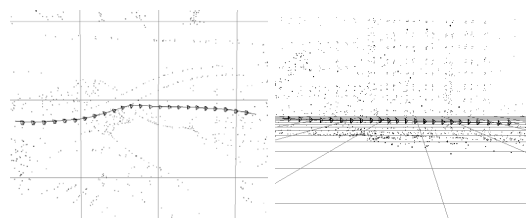
図 11 全方位型マルチカメラシステム Ladybug



図 12 入力動画画像系列の 1 フレーム

れトータルステーションとマーカボードによるキャリブレーション手法<sup>24)</sup> および基準点と自然特徴点の追跡による手法<sup>25)</sup> によってあらかじめ推定した。図 13 に、奥行き画像推定に利用したカメラシステムの移動パラメータを示す。図中の曲線および錘台はそれぞれ、図 12 左上の画像に対応するカメラユニットの移動の軌跡および 20 フレーム毎の姿勢を表している。事前に行った評価実験から、得られたカメラパスの推定精度は、位置に関して平均誤差 50mm、姿勢に関して平均誤差 0.07 度であり、利用したカメラパスの長さは 29m である<sup>25)</sup>。

このような入力データを用い、5 章で述べた手順 (SSSD による再探索を含む HYBRID 手法) で奥行き画像を生成した。まずステージ (A) において、各入力画像上において画像特徴点を検出し、続いて、ステージ (B) では、TNIP によるマルチベースラインステレオ法によって、各画像特徴点の奥行き値を算出した。本実験では、第  $f$  フレームに存在する画像特徴点の奥行きを算出するために、第  $(f - 100)$  フレームから第



(a) top view (b) side view

図 13 入力として用いたカメラパス (29m)

$(f + 100)$  フレームまで 2 フレームおき (606 枚, 101 フレーム) の画像内に存在する特徴点を利用した。また、TNIP での奥行き探索および SSSD による再探索に用いたウインドウ  $W$  のサイズは、前節の実験結果に基づきそれぞれ  $3 \times 3$  画素,  $7 \times 7$  画素に設定し、奥行き探索範囲は 1,000mm から 80,000mm とした。ステージ (C) では、5.3 節で述べた信頼度の閾値  $T, U$  をそれぞれ 2.0 画素, 0.3 と設定し、奥行き推定結果の整合性に基づいて信頼度の低い結果を削除した。その結果、全体の約 5 割の奥行き推定結果が誤推定として排除され、最終的に 1 画像あたり平均約 700 点 (1 フレームでの合計平均約 4,200 点) の奥行き推定結果が得られた。これは、奥行き推定の前段で用いた特徴点追跡によるカメラパラメータ推定<sup>25)</sup> によって得られた特徴点数の画像あたり平均 88 点 (1 フレームでの合計平均約 530 点) を大幅に上回っており、structure from motion に基づく手法の一つである手法<sup>25)</sup> に対して比較的多くの特徴点の奥行きを推定可能であることが分かる。

図 14 に、図 12 に対応する画像特徴点の位置および、奥行き推定結果を輝度値に変換したものを示す。また、各画像中から 1 点ずつ無作為に選択された 6 つの特徴点 (図 14 参照) に対する、奥行き値算出時の TNIP および再探索時の SSSD の値を、図 15 に示す。ただし、再探索時の SSSD の値を表すグラフにおいて、縦軸に平行な実線は TNIP によって算出された奥行き値を、破線は再探索によって算出された奥行き値を表す。同図中 TNIP のグラフから、いずれの画像特徴点においても、正しいと思われる奥行き値付近で TNIP が最大値をとっており、それ以外の奥行き値では TNIP に明白なピークは見られない。よって、これらの点においては TNIP を用いることで比較的ロバストに奥行きを決定できている。また、SSSD による再探索では、point4 を除き TNIP で得られた奥行き値付近に明白な局所最小値が見られるため、TNIP により SSSD の探索範囲が適切に限定されていることが分かる。point4 では SSSD 値の変化が比較的小さい

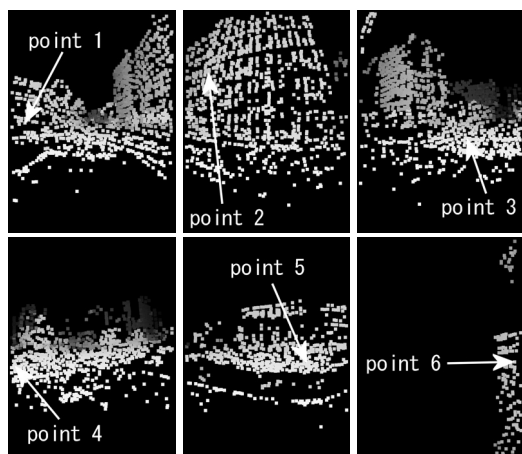


図 14 図 12 に対する奥行き推定結果



図 16 パノラマ展開した全方位動画の 1 フレーム



図 17 生成された全方位奥行き画像

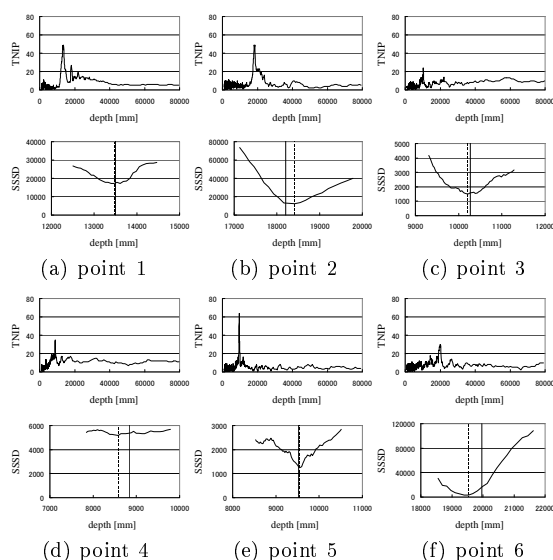


図 15 奥行き探索における TNIP 値と再探索時の SSSD 値

が、これは point4 がコントラストが低く広範囲に分布する地面のタイルのテクスチャを指しているためであり、point4 の TNIP 値には明確なピークが見られることから、奥行き探索範囲の限定自体は適切に行われていると考えられる。

最後に、推定された疎な奥行き画像をパノラマ展開し、奥行きに関する補間処理を行うことで、密な全方位奥行き画像を生成した。ここでは、まず Delaunay の三角分割法<sup>26)</sup>を用いることで、入力画像を画像特徴点を頂点とする多数の三角形に分割し、次に三角形の内部を画像特徴点の奥行き情報を用いて線形補間することで密な奥行き画像を生成した。図 12 から得られた全方位画像を極座標展開したパノラマ画像を図 16

に、奥行き値の内挿処理によって生成された奥行き画像を極座標展開したものを図 17 に示す。図 16 と図 17 の目視による比較から、大半の領域で奥行きの前後関係が良好に復元されていることを確認した。ただし、ここでは画像特徴点を頂点とする三角分割を行っているため、本来とは異なる奥行き値が算出されている箇所が見られる。これを解決するためには、テクスチャの整合性やエッジ情報等を用いた三角分割手法<sup>27),28)</sup>の導入を検討する必要がある。

## 7. まとめ

本論文では、画像特徴点の数え上げに基づく新たな評価尺度 TNIP を用いたマルチベースラインステレオ法を提案した。評価尺度 TNIP は、従来から用いられてきた SSD や NCC のような画像間の類似度評価を必要としないため、多数の画像を用いた場合にも比較的高速に奥行き画像を生成可能である。また、誤差関数が遮蔽による影響を受けにくいいため、遮蔽に頑健な奥行き推定を実現できる。ただし、TNIP による奥行き推定手法は、画像上への平均投影誤差が 1 画素以内となるような精密な位置決め精度で SSSD を用いた手法に劣るため、本論文では TNIP と SSSD を組み合わせることで両者の問題を解決する手法についても提案した。提案手法は特に、長いベースラインを必要とする屋外を対象とした動画や、全方位型のカメラを用いる場合のように、画像上での見え方や照明条件がフレーム毎に異なる入力に対して大きな効果があり、実環境を対象とした実験では、実際の屋外を対象とした全方位画像系列からの奥行き推定が可能であることを示した。また、シミュレーション実験においては、

TNIP の特性について明らかにし, TNIP と SSSD を組み合わせることで, 従来手法と同等の推定結果を 5 倍程度高速に算出可能であることを示した. 今後は, 推定された多数の奥行き画像を統合することで屋外環境の三次元モデル化を行う.

### 参 考 文 献

- 1) M. Okutomi and T. Kanade: "A Multiple-baseline Stereo," IEEE Trans. Pattern Analysis and Machine Intelligence, Vol.15, No.4, pp.353-363, 1993.
- 2) S. B. Kang, J. A. Webb, C. Zitnick and T. Kanade: "A Multibaseline Stereo System with Active Illumination and Real-time Image Acquisition," Proc. Int. Conf. on Computer Vision, pp.88-93, 1995.
- 3) S. B. Kang and R. Szeliski: "3-D Scene Data Recovery using Omnidirectional Multi-baseline Stereo," Int. Journal of Computer Vision, Vol.25, No.2, pp.167-183, 1997.
- 4) W. Zheng, Y. Kanatsugu, Y. Shishikui and Y. Tanaka: "Robust Depth-map Estimation from Image Sequences with Precise Camera Operation Parameters," Proc. Int. Conf. on Image Processing, Vol.II, pp.764-767, 2000.
- 5) T. Sato, M. Kanbara, N. Yokoya and H. Take-mura: "Dense 3-D Reconstruction of an Outdoor Scene by Hundreds-baseline Stereo Using a Hand-held Video Camera," Int. Journal of Computer Vision, Vol.47, Nos.1-3, pp.119-129, 2002.
- 6) M. Okutomi, Y. Katayama and S. Oka: "A Simple Stereo Algorithm to Recover Precise Object Boundaries and Smooth Surface," Int. Journal of Computer Vision, Vol.47, Nos.1-3, pp.261-273, 2002.
- 7) H. H. Baker: "Edge Based Stereo Correlation," Proc. Image Understanding Workshop, pp.168-175, 1980.
- 8) W. E. L. Grimson: "Computational Experiments with a Feature-based Stereo Algorithm," IEEE Trans. on Pattern Analysis and Machine Intelligence, Vol.7, No.1, pp.17-34, 1985.
- 9) S. Pollard, M. Pilu, S. Hayes and A. Lorusso: "View Synthesis by Trinocular Edge Matching and Transfer," Image and Vision Computing, Vol.18, No.9, pp.739-748, 2000.
- 10) 山本正信: "連続ステレオ画像からの 3次元情報の抽出", 電子情報通信学会論文誌 (D), Vol.J69-D, No.5, pp.143-162, 1992.
- 11) R. C. Bolles, H. H. Baker and D. H. Marimont: "Epipolar-plane Image Analysis: An Approach to Determining Structure from Motion," Int. Journal of Computer Vision, Vol.1, No.1, pp.7-55, 1987.
- 12) M. Okutomi and S. Sugimoto: "Shape Recovery of Rotating Object Using Weighted Voting of Spacio-temporal Images," Proc. Int. Conf. on Pattern Recognition, Vol.1, pp.790-793, 2000.
- 13) 中川雅朗, 斎藤英雄, 小沢慎治: "ハンディカメラを用いて撮影された画像列からの EPI の補正に基づく室内環境三次元再構築", 電子情報通信学会論文誌 (D-II), Vol.J84-DII, No.2, pp.266-275, 2001.
- 14) 川戸慎二郎: "空間への 2 段階投票による 3次元情報の抽出", 電子情報通信学会論文誌 (D-II), Vol.J77-D-II, No.12, pp.2334-2341, 1994.
- 15) B. Triggs, P. McLauchlan, R. Hartley and A. Fitzgibbon: "Bundle Adjustment - a Modern Synthesis," Proc. Int. Workshop on Vision Algorithms, pp.298-372, 1999.
- 16) M. Okutomi and T. Kanade: "A Locally Adaptive Window for Signal Matching," Int. Journal of Computer Vision, Vol.7, No.2, pp.143-162, 1992.
- 17) S. B. Kang, R. Szeliski and J. Chai: "Handling Occlusions in Dense Multi-view Stereo," IEEE Conf. on Computer Vision and Pattern Recognition, Vol.1, pp.103-110, 2001.
- 18) M. Sanfourche, G. L. Benerais and F. Champagnat: "On the Choice of the Correlation Term for Multi-baseline Stereo-vision," Proc. British Machine Vision Conference, Vol.II, pp.697-706, 2004.
- 19) C. Harris and M. Stephens: "A Combined Corner and Edge Detector," Proc. Alvey Vision Conf., pp.147-151, 1988.
- 20) H. Moravec: "Towards Automatic Visual Obstacle Avoidance," Proc. Int. Joint Conf. on Artificial Intelligence, p.584, 1977.
- 21) C. Schmid, R. Mohr and C. Bauckhage: "Evaluation of Interest Point Detectors," Int. Journal of Computer Vision, Vol. 37, No. 2, pp.151-172, 2000.
- 22) D. G. Lowe: "Distinctive Image Features from Scale Invariant Keypoints," Int. Journal of Computer Vision, Vol.60, No.2, pp. 91-110, 2004.
- 23) Point Grey Research Inc.: "Ladybug," <http://www.ptgrey.com/>.
- 24) S. Ikeda, T. Sato and N. Yokoya: "High-resolution Panoramic Movie Generation from Video Streams Acquired by an Omnidirectional Multi-camera System," Proc. IEEE Int. Conf. on Multisensor Fusion and Integration for Intelligent System, pp.155-160, 2003.

- 25) 佐藤智和, 池田聖, 横矢直和: “複数動画像からの全方位型マルチカメラシステムの位置・姿勢パラメータの推定”, 電子情報通信学会論文誌 (D-II), Vol.J88-D-II, No.2, pp.347-357, 2005.
- 26) P. Heckbert Ed.: Graphics Gems IV, pp.47-59, Academic Press, 1994.
- 27) D.D. Morris and T. Kanade: “Image-consistent surface triangulation,” Proc. IEEE Conf. on Computer Vision and Pattern Recognition, Vol.1, pp.332-338, 2000.
- 28) A. Nakatsuji, Y. Sugaya and K. Kanatani: “Mesh optimization using an inconsistency detection template,” Proc. Int. Conf. on Computer Vision, Vol.2, pp.1148-1153, 2005.  
(平成 1 年 1 月 1 日受付)  
(平成 1 年 1 月 1 日採録)



佐藤 智和 (正会員)

1999 年阪府大・工・情報工卒 . 2003 年奈良先端科学技術大学院大学情報科学研究科博士後期課程修了 . 現在 , 同大情報科学研究科助教 . コンピュータビジョンの研究に従事 .

2001 年電子情報通信学会学術奨励賞受賞 . 電子情報通信学会 , IEEE 各会員 .



横矢 直和 (正会員)

1974 年阪大・基礎工・情報工卒 . 1979 年同大大学院博士後期課程了 . 同年電子技術総合研究所入所 . 以来 , 画像処理ソフトウェア , 画像データベース , コンピュータビジョンの研究に従事 . 1986 ~ 87 年マッギル大・知能機械研究センター客員教授 . 1992 年奈良先端科学技術大学院大学・情報科学センター教授 . 現在 , 同大情報科学研究科教授 . 1990 年情報処理学会論文賞受賞 . 工博 . 電子情報通信学会 , 日本バーチャルリアリティ学会 , 人工知能学会 , 日本認知科学会 , 映像情報メディア学会 , IEEE 各会員 .

2001 年電子情報通信学会学術奨励賞受賞 . 電子情報通信学会 , IEEE 各会員 .