

平成 27 年度博士前期課程学位論文

カメラの微小不規則運動を利用する 3D 復  
元手法の高精度化とその性能評価

首都大学東京大学院シス  
テムデザイン研究科情報  
通信システム学域

13890523	塚田 翔太
指導教員	田川 憲男 教授

# 目次

第1章	序論	1
1.1	本研究の背景	1
1.2	本研究の目的	1
1.3	本論文の構成	3
第2章	固視微動を模擬したカメラモデルの基礎原理	4
2.1	設定したカメラモデル	4
2.2	トレモアを模擬したオプティカルフロー	5
第3章	勾配方程式の選択的利用法	7
3.1	勾配方程式に基づく奥行復元手法	7
3.1.1	剛体運動に関する勾配方程式	7
3.1.2	確率モデル	9
3.1.3	計算アルゴリズム	11
3.2	エイリアス現象と高次項に基づく勾配方程式の選択的利用法	15
3.2.1	エイリアス現象と高次項による影響	15
3.2.2	最適な画像変位の決定	16
3.2.3	閾値による勾配の選択的利用法	18
3.2.4	勾配方程式における高次項を考慮した空間微分値の利用法	20
第4章	カメラシステムの実装	21
4.1	カメラシステムの構築と実験	21
4.2	カメラキャリブレーション	24
第5章	奥行復元の性能評価	28
5.1	実画像による奥行復元の性能評価	28
5.1.1	$\sigma_o^2$ の収束	31
5.1.2	画像ペア使用率と $\sigma_r^2$ の関係	32
5.1.3	$\sigma_r^2=6.13 \times 10^{-6}$ での復元結果と正解の比較	33
5.1.4	$\sigma_r^2=2.33 \times 10^{-5}$ での復元結果と正解の比較	39
5.1.5	$\sigma_r^2=9.39 \times 10^{-5}$ での復元結果と正解の比較	45
5.1.6	$\sigma_d$ を変化させた場合の復元結果と正解の比較	50
5.1.7	RMSEによる奥行精度評価	53
5.1.8	画像特徴点ごとのカメラ回転 $(r_x, r_y)$ の使用傾向	54
5.2	奥行復元の性能評価に対する考察	55

第 6 章 結論	58
謝辭	59
参考文献	60
研究業績	62
付録 1	63
付録 2	65

# 目 次

1.1	固視微動のモデル.....	2
2.1	カメラ投影モデル.....	4
2.2	図 2.1 を上から見た図.....	5
3.1	一次元方向における勾配方程式の概略図.....	7
3.2	エイリアス現象が発生している画像対での輝度値の例.....	16
3.3	$\lambda/4$ の画像変位での画像対における輝度値の例.....	17
3.4	理想的な勾配方向の例.....	18
3.5	理想的な差ベクトルの大きさの例.....	19
4.1	カメラシステム.....	21
4.2	パソコン画面上的カメラ制御用ソフトウェア.....	21
4.3	カメラ駆動装置 (QT-ADM2).....	22
4.4	手動コントローラー (QT-AK).....	22
4.5	実験環境.....	23
4.6	奥行復元の流れ.....	23
4.7	チェスボードパターン.....	27
4.8	ボードを動かして撮影.....	27
4.9	コーナーの検出.....	27
5.1	実験画像.....	30
5.2	平行ステレオによる正解奥行マップ.....	30
5.3	画像ペア使用率を変更した時の $\sigma_o^2$ .....	31
5.4	$\sigma_r^2$ の大きさを変更した時の $\sigma_o^2$ .....	31
5.5	$\sigma_r^2$ の大きさによる画像ペア使用率の変化.....	32
5.6	既提案法の奥行マップ.....	33
5.7	断面図による正解奥行との比較マップ.....	33
5.8	画像ペア使用率 83% 時の奥行マップ.....	34
5.9	断面図による正解奥行との比較マップ.....	34
5.10	画像ペア使用率 71% 時の奥行マップ.....	35
5.11	断面図による正解奥行との比較マップ.....	35

5.12	画像ペア使用率 64%時の奥行マップ .....	36
5.13	断面図による正解奥行との比較マップ .....	36
5.14	画像ペア使用率 48%時の奥行マップ .....	37
5.15	断面図による正解奥行との比較マップ .....	37
5.16	画像ペア使用率 32%時の奥行マップ .....	38
5.17	断面図による正解奥行との比較マップ .....	38
5.18	既提案法の奥行マップ .....	39
5.19	断面図による正解奥行との比較マップ .....	39
5.20	画像ペア使用率 83%時の奥行マップ .....	40
5.21	断面図による正解奥行との比較マップ .....	40
5.22	画像ペア使用率 76%時の奥行マップ .....	41
5.23	断面図による正解奥行との比較マップ .....	41
5.24	画像ペア使用率 68%時の奥行マップ .....	42
5.25	断面図による正解奥行との比較マップ .....	42
5.26	画像ペア使用率 49%時の奥行マップ .....	43
5.27	断面図による正解奥行との比較マップ .....	43
5.28	画像ペア使用率 31%時の奥行マップ .....	44
5.29	断面図による正解奥行との比較マップ .....	44
5.30	画像ペア使用率 80%時の奥行マップ .....	45
5.31	断面図による正解奥行との比較マップ .....	45
5.32	画像ペア使用率 74%時の奥行マップ .....	46
5.33	断面図による正解奥行との比較マップ .....	46
5.34	画像ペア使用率 56%時の奥行マップ .....	47
5.35	断面図による正解奥行との比較マップ .....	47
5.36	画像ペア使用率 37%時の奥行マップ .....	48
5.37	断面図による正解奥行との比較マップ .....	48
5.38	画像ペア使用率 25%時の奥行マップ .....	49
5.39	断面図による正解奥行との比較マップ .....	49



5.40	$\sigma_d = 1.0 \cdot 10^{-1}$ 時の奥行マップ .....	50
5.41	断面図による奥行マップ .....	50
5.42	$\sigma_d = 1.0 \cdot 10^{-4}$ 時の奥行マップ .....	51
5.43	断面図による奥行マップ .....	51
5.44	$\sigma_d = 1.0 \cdot 10^{-6}$ 時の奥行マップ .....	52
5.45	断面図による奥行マップ .....	52
5.46	正解奥行に対する実験による復元奥行の RMSE .....	53
5.47	カメラ回転 $(r_x, r_y)$ の分布と特徴点ごとの使用傾向 .....	53

# 第1章 序論

## 1.1 本研究の背景

近年、ロボットの「眼」を作るコンピュータビジョン技術の進展が著しい。その中、三次元形状復元の分野では、多数の視点から撮像された画像を用い、三角測量の原理で対象の奥行を計算するステレオ立体視[1]が主流となってきた。ステレオ立体視では比較的大きな視差が得られることから、奥行復元精度が高い反面、視点の変化による隠蔽（オクルージョン）や見え方の変化が生じやすく、その解決のために数多くの研究が行われている。そこで本研究では、ステレオ立体視と比較して上記問題の生じにくい単眼での運動立体視による奥行復元手法に着目した。

運動立体視では、画像上で密な奥行を低計算量で復元するために、輝度値の時空間微分値を用いる解析的な手法（勾配法）が注目されてきた。勾配法では連続する二枚の画像間の運動視差が小さい必要があり、ゆえに一般に復元精度が低い。そこで、各画素に対応する奥行の情報を増加させることで、復元精度の向上が期待できる。一方で、人間の視覚システムは、固視微動と呼ばれる無意識の不規則眼球運動を有する[2], [3]。網膜上の視細胞は、変化する光に対してのみ感度を持つため、継続的に像を得るためには固視微動は必須である。固視微動は、確率共鳴と呼ばれる現象の一例と解釈される。これは、外的に加えられたノイズが逆に微弱信号の強調を可能にする非線形現象を指し、神経系ダイナミクスにおいて古くから知られる現象である。この固視微動を模擬するカメラ運動によって得られる複数の画像を入力とすることで、各画素における復元に多数の連続画像ペアを利用可能となる。

## 1.2 本研究の目的

固視微動は、図 1.1 に示すようにある方向に徐々に変位するドリフト成分と瞬間的に大きく動くサッケード成分に加え、小刻みな近傍運動であるトレモア成分からなる。追跡処理なしに同一点の奥行情報を得ることができることから、本研究ではトレモア成分に着目する。これを模擬するカメラ運動として、視軸上でレンズ中心から後方に回転中心をおく微小不規則回転を採用する。これは、眼球運動そのものに対応し、陰に

レンズ中心の並進運動が生じ、運動立体視が可能となる。我々の研究室ではこのようなカメラ運動によって得られる、連続した複数の画像ペアを観測とし、それに対する勾配法を一括して処理する方法[9], [12], [13]を提案している。ただし、勾配法を適応する際、複数の画像ペアを共通の奥行復元に利用するために、オプティカルフローを陽に検出しない直接法の枠組みを採用している[4], [5], [6], [7], [8]。勾配法で定義される勾配方程式とは、運動前後での画像ペアでの輝度値不変性を一次近似したものであり、輝度値の時間差分値と空間微分値に関する制約式となる。勾配法の精度は、この一次近似の当てはまりの程度に大きく依存する。特に輝度値の空間的な波長に対して画像変位が大きい場合は、エイリアス現象による大きな復元誤差が懸念される。本手法では、各画素あたり複数の画像ペアを利用できることから、勾配法の精度向上のための一つの工夫として、近似誤差が大きな勾配方程式、したがってそのような観測は奥行復元に利用しないという方策が考えられる。

そこで本研究では、上述の近似誤差を見積もる量を提案し、それに基づく勾配方程式の選択的利用によって、奥行復元精度の向上を試みる。加えて、精密回転ステージを計算機で制御するカメラシステムを構築し、カメラキャリブレーション[10]を含めた実験環境の整備を行う。このシステムを用いて提案手法を評価し、今後の研究に繋がる情報を獲得することを目的とする。

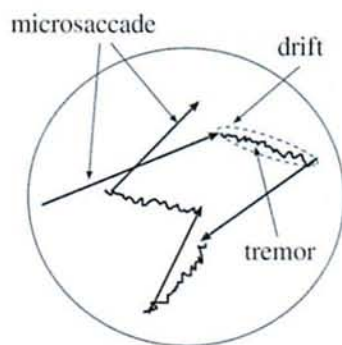


図 1.1 固視微動のモデル



## 1.3 本論文の構成

### 第1章：序論

近年のコンピュータビジョンの技術について述べ、人間の視覚系である固視微動と呼ばれる不規則な眼球運動について説明し、固視微動を用いたカメラ運動の従来研究における有用性を説明し、本研究の目的を示す。

### 第2章：固視微動を模擬したカメラモデルの基礎原理

動画像からの三次元形状復元の基本原理を説明した後、トレモアを模擬したカメラモデルとそれに基づく勾配法による奥行き復元について述べる。

### 第3章：勾配方程式の選択的利用法

勾配方程式を利用した奥行き復元手法の概要を説明した後、これに新たな工夫として提案した、勾配方程式の選択的利用法を詳述する。

### 第4章：カメラシステムの構築と実験

カメラシステムによって撮影された、実画像を用いた実験を通して、提案手法の有効性を評価する。

### 第5章：奥行き復元の性能評価

カメラシステムによって撮影された、実画像を用いた実験を通して、提案手法の有効性を評価する。

### 第6章：結論

本研究で得られた成果を整理するとともに、今後の課題を述べる。



## 2.2 トレモアを模擬したオプティカルフロー

固視微動の中のトレモアを模擬するカメラ運動について説明する。図 2.1 に示すように  $\vec{u} = [u_x, u_y, u_z]^T$ ,  $\vec{r} = [r_x, r_y, r_z]^T$  は並進及び回転速度ベクトルである。このとき画像面上の点  $(x, y, f)^T$  におけるオプティカルフロー  $\vec{v} = [v_x, v_y]^T$  は次式を満足する。

$$v_x = xy r_x - (1 + x^2) r_y + y r_z - (u_x - x u_z) d \quad (2.1)$$

$$v_y = (1 + y^2) r_x - x y r_y - x r_z - (u_y - y u_z) d \quad (2.2)$$

$d$  は奥行き  $Z$  の逆数であり、画像上の各画素における未知数で、 $(\vec{r}, \vec{u})$  は画像全体に対する未知数である。人間の眼球運動では、図 2.2 に示すようにカメラの回転中心が光軸に沿ってレンズ中心の後方  $Z_0$  に位置する。この回転が引き起こすレンズ中心での回転速度ベクトルは回転中心周りの回転速度ベクトルと同じ成分で表現することができる。

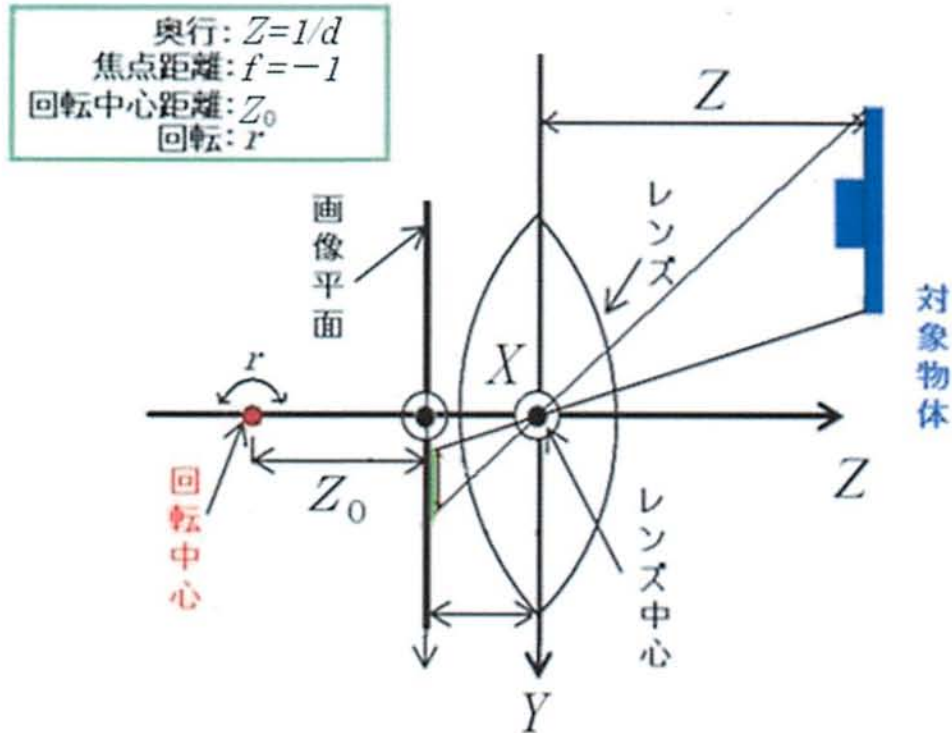


図 2.2 図 2.1 を横から見た図

一方、この回転中心のずれによって発生するレンズ中心の並進運動は、その並進運動ベクトルを  $\bar{\mathbf{u}}$  とすると、 $X$  及び  $Y$  軸周りの回転速度ベクトル  $\vec{\mathbf{r}} = [r_x, r_y]^T$  を用いて、 $\bar{\mathbf{u}}$  を以下の式で表すことができる。

$$\bar{\mathbf{u}} = \vec{\mathbf{r}} \times \begin{bmatrix} 0 \\ 0 \\ Z_0 \end{bmatrix} = Z_0 \begin{bmatrix} r_y \\ -r_x \\ 0 \end{bmatrix} \quad (2.3)$$

この並進速度ベクトル  $\bar{\mathbf{u}}$  を用いると、本研究の画像平面におけるオプティカルフロー  $\bar{\mathbf{v}} = [v_x, v_y]^T$  は次式のように定式化できる。

$$v_x = xy r_x - (1 + x^2) r_y - Z_0 r_y d \equiv v_x^r - r_y Z_0 d \quad (2.4)$$

$$v_y = (1 + y^2) r_x - xy r_y + Z_0 r_x d \equiv v_y^r - r_x Z_0 d \quad (2.5)$$

このカメラ運動モデルは、直接的な並進運動が不要であり、またカメラ回転は不規則で微小なものを想定しているため、カメラ制御が容易となる。加えて、この回転はフレームの数だけ存在し、この値を正確に知ることはシステムを複雑にするので、本手法ではこれも推定することとする。一方で、実カメラにおける  $Z_0$  は一般的に奥行の相対値として求まる。絶対量を知りたい場合は、奥行が既知の形状が必要となる。また、このような形状を求める場合は平行ステレオ [10] 等で奥行を計測する方法が考えられる。平行ステレオを行う際に、ピクセル単位の焦点距離  $f$  を用いるので、事前にカメラキャリブレーションを行う必要がある。

## 第3章 勾配方程式の選択的利用法

### 3.1 勾配方程式に基づく奥行復元手法

#### 3.1.1 剛体運動に関する勾配方程式

画像上の輝度値を  $f(x, y, t)$  で表すと、運動前後での輝度値不変性の一次近似式（勾配方程式）は、1次元に限れば図 3.1 のように解釈され次式のように与えられる。ただし、輝度値の運動前後での不変性を厳密に考慮するならば、物体の光反射特性は完全な拡散反射であり、物体と光源が固定された状況でカメラが運動していると考える必要がある。

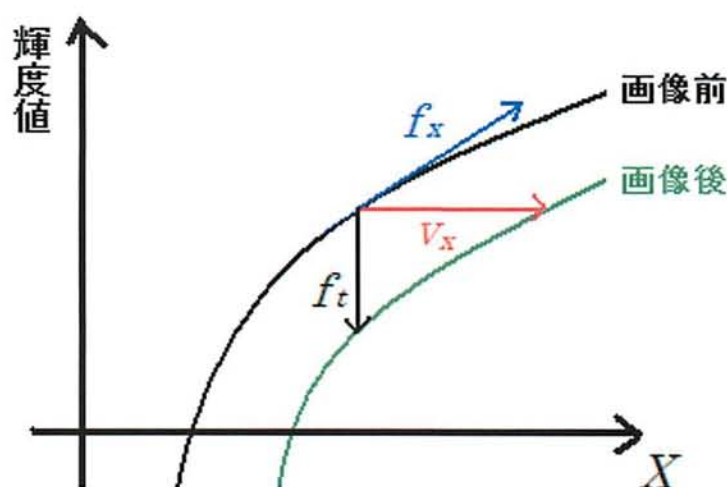


図 3.1 一次元方向における勾配方程式の概略図

$$f_t = -f_x v_x - f_y v_y \quad (3.1)$$

ここで、 $f_t, f_x, f_y$  はそれぞれ時間及び空間の偏導関数である。本研究では差分頻度を高めるために、 $f_x, f_y$  を 5 タップ差分によって求める。5 タップ差分とは  $[-0.108415 \ -0.280353 \ 0.0 \ 0.280353 \ 0.108415]$  をフィルタ係数とする畳み込みであり、画像の横方向に施せば



$f_x$ が、縦方向に施せば $f_y$ が求められる。式(3.1)に式(2.4), (2.5)を代入すると、今回のカメラ運動に対する剛体運動を表現する勾配方程式は

$$\begin{aligned} f_i &= -(f_x v_x^r + f_y v_y^r) - (-f_x r_y + f_y r_x) Z_0 d \\ &\equiv -f^r - f'' d \end{aligned} \quad (3.2)$$

と表すことができる。 $d$ を計算する際、精度を上げるよう複数の画像ペアを用いるが、計算の効率を上げるために、式(3.2)のように $d$ を共通の未知数として定式化することで、複数の画像ペアに対して一括した計算が可能となる。

$f_i, f_x, f_y$ の全ては観測量であるため、誤差を含む一方で、方程式そのものに近似誤差が存在し、これが多くを占めることから、これらの誤差を統計して $f_i$ にのみ誤差が含まれるものとして扱う。全ての観測量に基づいて $\bar{\mathbf{r}}, \bar{\mathbf{u}}, d$ を決定することになる。また、 $f_x, f_y$ は上記の差分により誤差が少ないため、 $f_i$ の観測誤差と方程式の誤差を統計して $f_i$ の誤差を考える必要がある。

### 3.1.2 確率モデル

使用するフレーム総数を  $M$ ，画素数を  $N$  とする． $f_t^{(i,j)}_{i=1\dots N, j=1\dots M}, \bar{r}^{(j)}_{j=1\dots M}, d^{(i)}_{i=1\dots N}$  を確率変数として扱う．まず，観測量  $f_t^{(i,j)}$  に加わる誤差を 0，分散  $\sigma_o^2$  の正規分布に従うものとする． $f_t^{(i,j)}$  の条件付き確率密度関数は次式で表される．

$$p(f_t^{(i,j)} | d^{(i)}, \bar{r}^{(j)}, \sigma_o^2) = \frac{1}{\sqrt{2\pi}\sigma_o} \times \exp\left\{-\frac{(f_t^{(i,j)} + f^{r(i,j)} + f^{u(i,j)}d^{(i)})^2}{2\sigma_o^2}\right\} \quad (3.3)$$

次に，眼球の回転速度  $r^{(j)}$  において，本研究では固視微動のドリフト成分は無視することから，各軸の成分は独立であるとし，平均 0 の次式の正規分布に従うものとする．

$$p(\bar{r}^{(j)} | \sigma_r^2) = \frac{1}{2\pi\sigma_r^2} \exp\left\{-\frac{\bar{r}^{(j)\top}\bar{r}^{(j)}}{2\sigma_r^2}\right\} \quad (3.4)$$

以上のモデルより  $\Theta = \{\sigma_o^2, \sigma_r^2\}$  を未知のパラメータとすると， $\{f_t^{(i,j)}\}, \{\bar{r}^{(j)}\}$  の同時確率は次式で与えられる．

$$\begin{aligned} & p(\{f_t^{(i,j)}\}, \{\bar{r}^{(j)}\} | \Theta) \\ &= \sum_{i=1}^N \sum_{j=1}^M p(f_t^{(i,j)} | d^{(i)}, \bar{r}^{(j)}, \sigma_o^2) \sum_{j=1}^M p(\bar{r}^{(j)} | \sigma_r^2) \\ &= \frac{1}{(2\pi)^{M(N+1)} \sigma_o^{MN} \sigma_r^{2M}} \times \exp\left\{-\frac{(f_t^{(i,j)} + f^{r(i,j)} + f^{u(i,j)}d^{(i)})^2}{2\sigma_o^2} - \frac{\bar{r}^{(j)\top}\bar{r}^{(j)}}{2\sigma_r^2}\right\} \end{aligned} \quad (3.5)$$

ここで，本研究では不規則な微小回転運動  $r^{(j)}$  により振動する複数のフレームを利用することから，各画素位置で得られる  $d^{(i)}$  は純粹にその位置に対応する値ではなく，画像上での振れ幅に相当する近傍領域の平均的な値となる．すなわち，結果的に  $d^{(i)}$  は周囲と相関を持つものとして得られる．この相関の空間的広がりには奥行きにも依存するものであり，本来ならば相関に関する変数を持ち得なければならないのだが，今回は簡単のため次式のように扱う．

$$p(\vec{d} | \sigma_d^2) = \frac{1}{(\sqrt{2\pi}\sigma_d)^N} \exp\left\{-\frac{\vec{d}^\top L \vec{d}}{2\sigma_d^2}\right\} \quad (3.6)$$

ここで、 $\vec{d}$  は  $d^{(i)}$  を成分とする  $N$  次元ベクトルであり、共分散行列を  $\sigma_d^2 L^{-1}$  とする。  $L$  は自由端境界条件を満足する 2 次元ラプラシアン演算子である。このモデルにより、空間的に滑らかな奥行きマップが得られることになる。本研究では、分散  $\sigma_d^2$  は奥行きの滑らかさを制御する既知の量として扱う。以上の全ての確率モデルに基づき、入力画像から、 $\{\sigma_o^2, \sigma_r^2\}$  を最尤推定[14]、 $d^{(i)}$  と  $r^{(i)}$  を MAP 推定により決定することができる。

### 3.1.3 計算アルゴリズム

本研究では効率的で安定な計算のために OSL-MAP(One Step Late-Max A Posteriori)-EM アルゴリズム[11]を採用する. EM アルゴリズムとは観測されていない未知確率変数を観測量（事前分布）に加えて想定したときに，両者（完全データ）が観測されれば問題が解きやすくなる場合に有効な推定手段であり，その未知確率変数の事後分布とパラメータの最尤推定量を同時に得ることのできる反復手法である. また，EM アルゴリズムの拡張である MAP-EM アルゴリズムでは，得られた事後分布を最大化する際に，パラメータに事前確率を導入して，MAP 推定によって決定する. 各反復では以下のステップを実行する.

1. E-step では， $Q$  関数と呼ばれるパラメータ更新のための評価関数を導出する.
2. M-step では， $Q$  関数を最大化するようにパラメータを更新する.

上記の 2 ステップを収束するまで交互に繰り返す.

本研究においては，上記のアルゴリズムを実行するにあたり， $\Theta = \{\vec{d}, \sigma_o^2, \sigma_r^2\}$  を未知のパラメータとすると， $\hat{\Theta}$  を  $p(\Theta | f_i^{(i,j)})$  に基づく MAP 推定量として決定することができる. 加えて， $\hat{r}^{(j)}$  は先ほど推定したパラメータ  $\hat{\Theta}$  を利用し， $p(\{r^{(j)}\} | \{f_i^{(i,j)}\}, \hat{\Theta})$  に基づく MAP 推定量として得ることができる. この 2 段階の MAP 推定により最終的な奥行  $\vec{d}$  を推定することになる. ただし， $\sigma_o^2, \sigma_r^2$  には事前知識を定義していないので，形式的には一様分布を仮定することになる. また，上述の  $\hat{\bullet}$  は推定量を意味する. 以下では EM のスキームについて詳述する.

EM スキームでは， $\{f_i^{(i,j)}, \vec{r}^{(j)}\}$  を完全データ， $\vec{r}^{(j)}$  を欠損データ， $\Theta$  を未知量として扱い，E-step と M-step を収束するまで交互に反復する.

Estep では，完全データの対数尤度関数に対し， $f_i^{(i,j)}$  を観測したときの条件付き期待値操作を，前の反復において得られている MAP 推定値  $\Theta$  を固定して施す. 得られる関数は通常， $Q$  関数と呼ばれる.  $Q$  関数を以下に示す.

$$Q(\Theta, \hat{\Theta}) \equiv E[\ln p(\{f_i^{(i,j)}\}, \{\vec{r}^{(j)}\} | \Theta) | f_i^{(i,j)}, \hat{\Theta}] \quad (3.7)$$

続いての M-step であるが、特に MAP-EM アルゴリズムにおいては、本来の EM アルゴリズムの枠組みではパラメータとして扱う変数、つまり本研究における  $\Theta$  の事前確率の対数をとったものを、上述の  $Q$  関数に加えて、その関数を  $\Theta$  に関して最大化して、 $\Theta$  の更新を行う。前節での確率モデルを用いることで、M-step で最小化すべき評価関数は、事後確率  $p(\{r^{(j)}\}|\{f_t^{(i,j)}\}, \Theta)$  を具体的に記述することによって、次式のように書き下すことができる。

$$p(\{\bar{r}^{(j)}\}|\{f_t^{(i,j)}\}, \Theta) = \frac{p(\{\bar{r}^{(j)}\}, \{f_t^{(i,j)}\}|\Theta)}{p(\{f_t^{(i,j)}\}|\Theta)} \quad (3.8)$$

式(3.3), (3.5), (3.8)から  $p(\{r^{(j)}\}|\{f_t^{(i,j)}\}, \Theta)$  は次式となる。

$$p(\{\bar{r}^{(j)}\}|\{f_t^{(i,j)}\}, \Theta) = \frac{1}{\sqrt{(2\pi)^{2M} \sum_{j=1}^M \det \bar{V}_r^{(j)}}} \times \exp \left\{ -\frac{1}{2} \sum_{j=1}^M (\bar{r}^{(j)} - \bar{r}_m^{(j)})^T V_r^{(j)-1} (\bar{r}^{(j)} - \bar{r}_m^{(j)}) \right\} \quad (3.9)$$

$$\bar{V}_r^{(j)} = \left( \frac{1}{\sigma_o^2} \sum_{i=1}^N \bar{w}^{(i,j)T} \bar{w}^{(i,j)} + \frac{1}{\sigma_r^2} I \right)^{-1} \quad (3.10)$$

$$\begin{aligned} w^{(i,j)} &= \begin{pmatrix} f_x^{(i,j)} x^{(i)} y^{(i)} + f_y^{(i,j)} (1 + y^{(i)2}) \\ -f_x^{(i,j)} (1 + x^{(i)2}) - f_y^{(i,j)} x^{(i)} y^{(i)} \end{pmatrix} + Z_0 d \begin{pmatrix} f_y^{(i,j)} \\ -f_x^{(i,j)} \end{pmatrix} \\ &\equiv \bar{w}_0^{(i,j)} + Z_0 d^{(i)} \bar{w}_d^{(i,j)} \end{aligned} \quad (3.11)$$

式(3.9)に式(3.6)を加え、対数をとったものを式(3.12)に示す。

$$\begin{aligned} lc(\Theta) &= -\frac{MN}{2} \ln \sigma_o^2 - M \ln \sigma_r^2 \\ &\quad - \frac{1}{2\sigma_o^2} \sum_{j=1}^M \left\{ \sum_{i=1}^N f_t^{(i,j)2} + 2 \left( \sum_{i=1}^N w^{(i,j)T} f_t^{(i,j)} \right) \bar{r}^{(j)} + \text{tr} \left[ \left( \sum_{i=1}^N w^{(i,j)T} w^{(i,j)} \right) \bar{r}^{(j)T} \bar{r}^{(j)} \right] \right\} \\ &\quad - \frac{1}{2\sigma_r^2} \sum_{j=1}^M \text{tr}(\bar{r}^{(j)T} \bar{r}^{(j)}) - \frac{\bar{d}^T L \bar{d}}{2\sigma_d^2} \end{aligned} \quad (3.12)$$



$$\mathbb{E}[\tilde{\mathbf{r}}^{(j)}] \equiv \tilde{\mathbf{r}}_m^{(j)} = -\frac{1}{\sigma_o^2} V_r^{(j)} \sum_{i=j}^N f_t^{(i,j)} \tilde{\mathbf{w}}^{(i,j)} \quad (3.13)$$

$$\mathbb{E}[\tilde{\mathbf{r}}^{(j)\top} \tilde{\mathbf{r}}^{(j)}] \equiv \hat{\mathbf{R}}^{(j)} = \hat{\mathbf{V}}_r^{(j)} + \hat{\mathbf{r}}_m^{(j)\top} \hat{\mathbf{r}}_m^{(j)} \quad (3.14)$$

上記の2式による期待値操作により，式(3.12)から次式の $Q$ 関数が導かれる．

$$\begin{aligned} Q(\Theta) = & -\frac{MN}{2} \ln \sigma_o^2 - M \ln \sigma_r^2 \\ & -\frac{1}{2\sigma_o^2} \sum_{j=1}^M \left\{ \sum_{i=1}^N f_t^{(i,j)^2} + 2 \left( \sum_{i=1}^N \tilde{\mathbf{w}}^{(i,j)\top} f_t^{(i,j)} \right) \hat{\mathbf{r}}_m^{(j)} + \text{tr} \left[ \left( \sum_{i=1}^N \tilde{\mathbf{w}}^{(i,j)\top} \tilde{\mathbf{w}}^{(i,j)} \right) \hat{\mathbf{R}}^{(j)} \right] \right\} \\ & -\frac{1}{2\sigma_r^2} \sum_{j=1}^M \text{tr} \hat{\mathbf{R}}^{(j)} - \frac{\vec{d}^\top L \vec{d}}{2\sigma_d^2} \end{aligned} \quad (3.15)$$

M-step では，式(3.15)を最大にするよう $\Theta$ を更新する．定数項を省略すると式(3.15)は以下のように書き直すことができる．

$$Q(\Theta) = -\frac{MN}{2} \ln \sigma_o^2 - M \ln \sigma_r^2 - \frac{1}{2\sigma_o^2} \hat{F}(\{\vec{d}\}) - \frac{1}{2\sigma_r^2} \sum_{j=1}^M \hat{G} - \frac{\vec{d}^\top L \vec{d}}{2\sigma_d^2} \quad (3.16)$$

これにより $\sigma_o^2$ と $\sigma_r^2$ に関しては，次のように更新式を得られる．

$$\sigma_o^2 = \frac{\hat{F}(\{\vec{d}\})}{MN} \quad \sigma_r^2 = \frac{\hat{G}}{2M} \quad (3.17)$$

$\vec{d}$  に関しては，式(3.15)の最終項の各 $d^{(i)}$ に関する偏導関数が，近傍 $d$ の関数となるので， $\vec{d}$ の更新には連立方程式を解く必要がある．これを避けるため，本研究では，One-Step-Late(OSL) テクニックを適用する．

これは、近傍の  $d$  は得られている値を固定した定数とみなし、偏導関数を 0 とおいた方程式を各  $d^{(i)}$  のみの方程式と考える方法である。解は次式のように得られ、これを更新式とする。

$$\hat{d}^{(i)} = \frac{\hat{\sigma}_o^2 \hat{d}^{(i)}}{\sigma_d^2 Z_0 \sum_{j=1}^M \text{tr}(A^{(i,j)} \hat{R}^{(j)}) + \hat{\sigma}_o^2} - \frac{\sigma_d^2 Z_0 \sum_{j=1}^M \left\{ \bar{w}_d^{(i,j)T} f_i^{(i,j)} \hat{r}_m^{(j)} + \text{tr}(B^{(i,j)} \hat{R}^{(j)}) \right\}}{\sigma_d^2 Z_0^2 \sum_{j=1}^M \text{tr}(A^{(i,j)} \hat{R}^{(j)}) + \hat{\sigma}_o^2} \quad (3.18)$$

式(3.18)における  $\bar{d}^{(i)}$  は、 $d^{(i)}$  の 4 近傍系の  $d^{(i)}$  を除いた算術平均を示す。ここで、 $\sigma_o^2$  も現在の値に固定しており、こうして得られた更新式  $\hat{d}$  を式 (3.17) に代入することで  $\sigma_o^2$  の更新が行われる。また  $A^{(i,j)}$ ,  $B^{(i,j)}$  は、以下のように定義される。

$$A^{(i,j)} \equiv \bar{w}_d^{(i,j)T} \bar{w}_d^{(i,j)} \quad (3.19)$$

$$B^{(i,j)} \equiv \frac{\bar{w}_0^{(i,j)T} \bar{w}_d^{(i,j)} + \bar{w}_d^{(i,j)T} \bar{w}_0^{(i,j)}}{2} \quad (3.20)$$

## 3.2 エイリアス現象と高次項に基づく勾配方程式の選択的利用法

### 3.2.1 エイリアス現象と高次項による影響

我々の研究室では、2 章で述べたカメラ運動を採用し、勾配方程式に基づく 3.1 での手法を提案している。勾配方程式とは、3.1.1 で詳述したように運動前後での輝度値不変性を一次近似したものであり、観測量である輝度値の空間差分値とオプティカルフローに関する制約式となる。勾配法の精度は、この一次近似の当てはまりの程度に大きく依存する。しかし、厳密には高次項による誤差が存在し、それが固有の問題として精度に影響している。一方で、画像変位による奥行き復元の原理的な問題としてエイリアス現象による大きな復元誤差が懸念される。逆に画像変位が小さい場合は、信号対雑音比の低下によって復元値がノイズに隠れてしまう可能性もある。この 2 つの現象は画像変位による問題であるので、勾配法とは関係がない。しかし、勾配法は、このエイリアス現象を含め高次項の存在に問題がある。既提案法では、このような勾配方程式の性質を考慮に入れておらず、十分な奥行き復元精度が得られていなかった。本研究では、各画素あたり複数の画像ペアを利用できることから、勾配法の精度向上のための一つの工夫として、近似誤差が大きな勾配方程式、したがってそのような観測は奥行き復元に利用しないという方策を検討する。

本研究では、勾配方程式の性質を踏まえて画像の動きを制限し、そのうえで、上述の近似誤差を見積もる量を提案し、それに基づく勾配方程式の選択的利用によって、奥行き復元精度の向上を図る手法を提案する。加えて、精密回転ステージを計算機で制御する実カメラシステムを構築し、撮影した実画像を利用した評価を行う。本章では、提案手法について詳述する。

### 3.2.2 最適な画像変位の決定

先に述べたように、既提案法ではエイリアス現象による影響を無視していた。エイリアス現象とは、異なる連続信号が標本化によって区別できなくなることをいい、折り返し雑音とも呼ばれる。また、信号対雑音比の低下による問題も存在する。信号対雑音比(SN比)とは、雑音に対する送信信号の割合のことをいい、低下すると雑音の割合が信号の大半を占めるために信号の取得が困難になる。輝度値の空間的な波長に対して画像変位が大きい場合は、エイリアス現象による大きな復元誤差が生じる。逆に画像変位が小さい場合はSN比の低下によって復元精度が下がる。そのため、上記の2つを回避することが望ましい。SN比の低下に関しては、複数の連続する画像ペアを用いて、空間差分値  $f_i$  の情報量を増やすことで回避可能となる。ここで、以下にエイリアス現象が起きている場合、連続する画像対それぞれの一次方向における輝度値の波長を図3.2に示す。

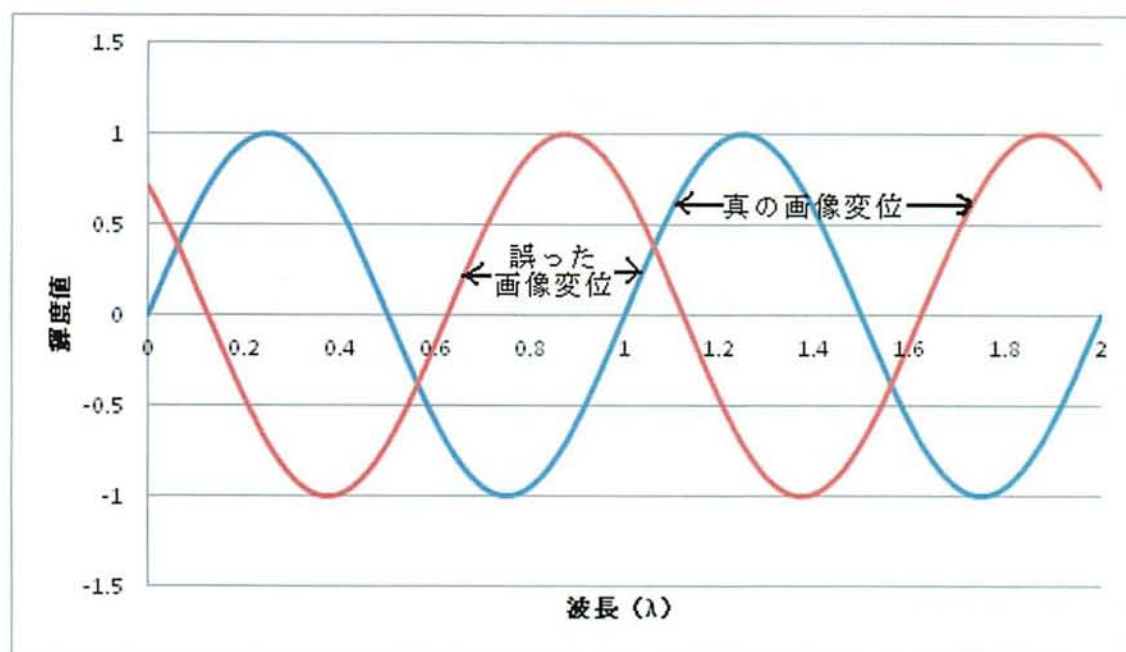


図 3.2 エイリアス現象が発生している画像対における輝度値の例

以上の図は、青い波に対し、赤い波が真の画像変位分動いているものであり、画像変位が  $\lambda/2$  を超えた場合その超過分、エイリアスにより誤った画像変位が検出されてしまうことを表している。

以上を考慮し、輝度値の空間的な波長に対する  $\lambda/2$  の画像変位よりも小さく、一様な差分を多く取れる画像変位を設定することで、問題を解決するための最適な画像変位を決定できる。本研究では以上を考慮した、輝度値の空間的な波長に対する  $\lambda/4$  の画像変位を最適値として得た。また、画像変位が  $\lambda/4$  の場合での一次方向における輝度値の例を図 3.4 に示す。

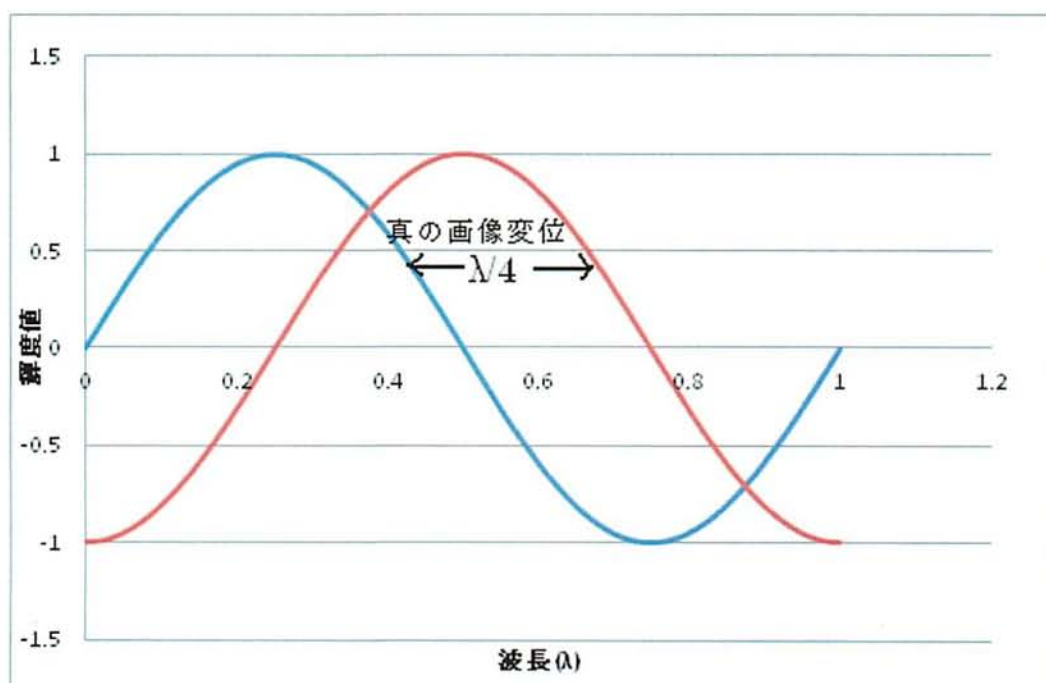


図 3.3  $\lambda/4$  の画像変位での画像対における輝度値の例

以上の図は、青い波に対し、赤い波が真の画像変位である  $\lambda/4$  分動いているものであり、エイリアスが発生していないことを表している。



### 3.2.3 閾値による勾配の選択的利用法

3.2.2 で説明した、適切な画像変位の議論に基づき、輝度勾配ベクトル  $(f_x, f_y)$  の選定によっても精度向上が望める。画素あたり複数の画像ペアを利用できることから、近似誤差が大きな勾配方程式、すなわち、勾配方程式の高次項が多い部分は奥行復元の過程から除くという方策が考えられるからである。本研究では、カメラ運動によって得られた  $M$  枚の連続した画像ペアのうち、各画像ペア 2 枚分の空間微分値から各画素位置の輝度勾配ベクトル  $(f_x, f_y)$  の内積、つまり空間勾配の向きを求め、それが大きく異なる場合は高次項が多くなり近似誤差を生むと類推する。連続する  $M$  枚の画像のうち、隣り合う画像ペアそれぞれの空間微分値を  $(f_{x0...M-1}, f_{y0...M-1})$ 、もう一方を  $(f_{x1...M}, f_{y1...M})$  とすると内積は以下の式で表現できる。

$$\cos \theta = \frac{f_{x0...M-1}f_{x1...M} + f_{y0...M-1}f_{y1...M}}{\sqrt{f_{x0...M-1}^2 + f_{y0...M-1}^2} \times \sqrt{f_{x1...M}^2 + f_{y1...M}^2}} \quad (3.21)$$

例えば、空間勾配の向きが負になった場合は  $90^\circ$  以上となり勾配が大きく異なるため、そのような画素の勾配方程式は選択しないとできる。このような部分は、主に図 3.2、図 3.3 で同じ位置での波同士の角度が  $90^\circ$  以上の部分のことを指す。また、図 3.4 は空間勾配の向きが、正負でどのようになるのかを図示したもので、2 つのベクトルが平行に近ければ理想的で、離れるほど高次項が増加し近似誤差を生むことを示している。こうした理由から、閾値による勾配方程式の取捨選択が必要不可欠となる。本研究では、空間勾配の向きに対し、正の場合は勾配方程式を使用し、負の場合は向きに対する閾値を用いて除くとする。

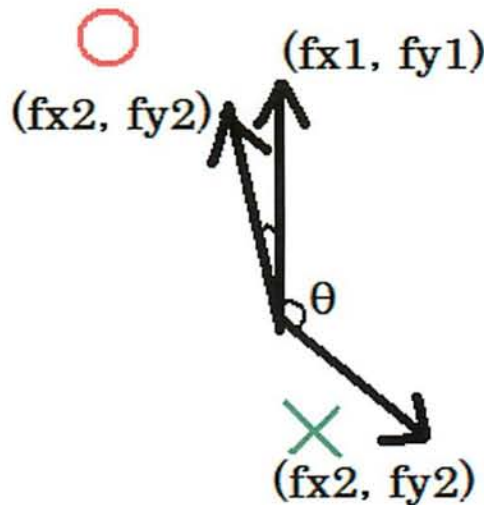


図 3.4 理想的な勾配方向の例

ここで注意しなければならない点がある．上記の式(3.21)が画像ペアの勾配の向きのみ  
にしか考慮されていない点である．これは、向きによる判定が相対的なものでしかないた  
め、空間勾配の差ベクトルの大きさに着目し絶対的な判定を行う必要があることを意味  
する．そのため、次点として式(3.21)で利用できると判断された空間勾配に対し、勾配の  
差ベクトルの大きさと、その基準となるベクトルの大きさ双方の比を採用する必要がある  
(図 3.5 参照)．

理由として、勾配の差ベクトルは、次節で説明する勾配方程式の二次項の成分であり、  
この値が大きいと勾配方程式の誤差量が増えてしまう．言い換えれば、二次項は一次項  
に対する雑音でもあるので、SN 比が増加することになる．そこで基準となるベクトルに  
対する差ベクトルの比が小さなものを見ることで、より誤差の少ない勾配を選定できる  
ということになる．さらに、この勾配がオプティカルフローに対して垂直に近いほど、  
差ベクトルの大きさに関わらず誤差を低減できる．以上の条件から考えた判定基準を  $J$  と  
すると、オプティカルフロー  $(v_x, v_y)$ 、基準となるベクトル  $(f_{x0...M-1}, f_{y0...M-1})$ 、もう一方の  
ベクトル  $(f_{x1...M}, f_{y1...M})$  を用いて以下の式で表現できる．本研究では、全画素での空間勾  
配の大きさの平均を算出し、それを閾値とする．この閾値より  $J$  が小さい場合は勾配方程式  
に利用し、大きい場合は除くとする．

$$J = \frac{|(f_{x0...M-1} - f_{x1...M})v_x + (f_{y0...M-1} - f_{y1...M})v_y|}{2|f_{x0...M-1}v_x + f_{y0...M-1}v_y|} \quad (3.22)$$

以上のような、向きと大きさを考慮した 2 段階での空間勾配の決定法により、閾値によ  
って選択的に、利用する勾配方程式を決めることができる．

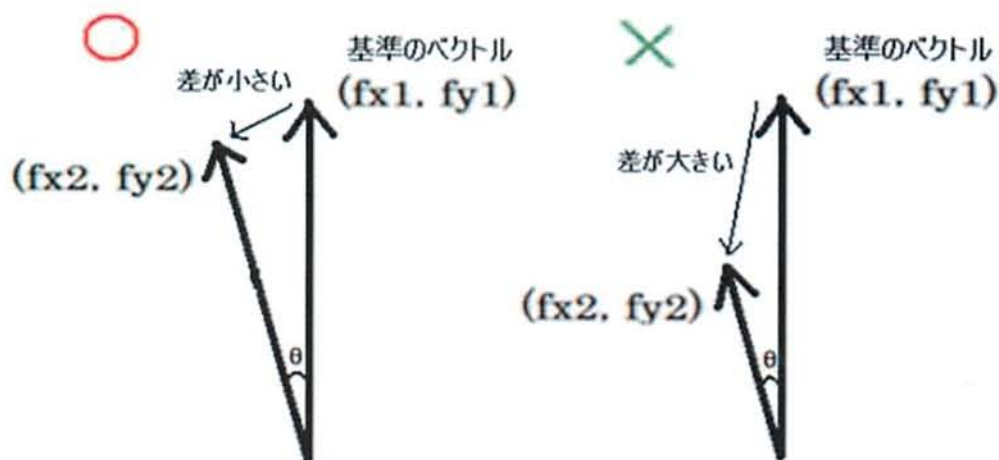


図 3.5 理想的な差ベクトルの大きさの例

### 3.2.4 勾配方程式における高次項を考慮した空間微分値の利用法

本研究での勾配方程式は 1 次項のみだとしているが、厳密には高次項も存在する。高次項は 1 次項に対し割合が少なく、雑音として扱うことが妥当である。逆に、高次項が大きい場合は近似誤差の原因となるので、基本的に高次項が小さい場合のみ有効となる。また前節で内積の判定により、誤差の大きい画像ペアが除かれ、非線形項が小さくなっており 2 次項が一次導関数のように近似可能と考えられるので、2 次項のみに着目した場合を考える。この 2 次項は、カメラ運動によって得られた  $M$  枚の連続画像ペア同士の空間微分値のズレと考えられるので、差分を取ることでより求められる。以下に導出過程を示す。上記の式(3.1)で扱われた空間微分値を採用し  $x$  方向、 $y$  方向にその差分を取ると以下の式で表現できる。

$$f_{x0\dots M-1} - f_{x1\dots M} = f_{xy}v_y + f_{xx}v_x \quad (3.23)$$

$$f_{y0\dots M-1} - f_{y1\dots M} = f_{yx}v_x + f_{yy}v_y \quad (3.24)$$

この式(3.23)に  $v_x$  を、式(3.24)に  $v_y$  をかけ、2つの式の和を取ると以下のようになる。

$$\begin{aligned} & (f_{x0\dots M-1} - f_{x1\dots M})v_x + (f_{y0\dots M-1} - f_{y1\dots M})v_y \\ &= f_{xy}v_xv_y + f_{xx}v_x^2 + f_{yx}v_xv_y + f_{yy}v_y^2 \end{aligned} \quad (3.25)$$

式(3.25)が勾配方程式の 2 次項となり、これを考慮した勾配方程式は以下の式となる。

$$\begin{aligned} f_t &= -f_xv_x - f_yv_y \\ &\quad - (f_{xx}v_x^2 + 2f_{yx}v_xv_y + f_{yy}v_y^2) \end{aligned} \quad (3.26)$$

式(3.26)を元に  $f_x$ 、 $f_y$  を 2 次項を考慮した数値に更新する式を以下に示す。更新したものを  $F_{x0\dots M-1}$ 、 $F_{y0\dots M-1}$  とする。

$$F_{x0\dots M-1} = \frac{3f_{x0\dots M-1} - f_{x1\dots M}}{2} \quad (3.27)$$

$$F_{y0\dots M-1} = \frac{3f_{y0\dots M-1} - f_{y1\dots M}}{2} \quad (3.28)$$

以上より、2 次項を考慮した数値を勾配方程式に用いることで、より実際的な時間差分値を求めることができる。これは、実画像から求められる空間差分値にノイズが含まれていることを利用し、空間差分値の誤差を求めることで評価が可能となる。



## 第4章 カメラシステムの実装

### 4.1 カメラシステムの構築と実験

我々研究室では、従来から固視微動を模擬したカメラ運動に基づく奥行復元に、勾配法を用いたシミュレーションを重ね、様々な改良を行ってきた。しかし、実画像を利用した評価は今まで行なわれておらず、今後に向け実際的な実験が必要不可欠となっている。そこで我々は、図 2.1 のカメラ写像形をハードウェアで実装したカメラを用いて実画像に対する奥行復元実験を行う。実際に実装したカメラシステムを図 4.1、カメラを動作させるためのパソコン画面の制御ソフトウェアを図 4.2 に示す。



図 4.1 カメラシステム

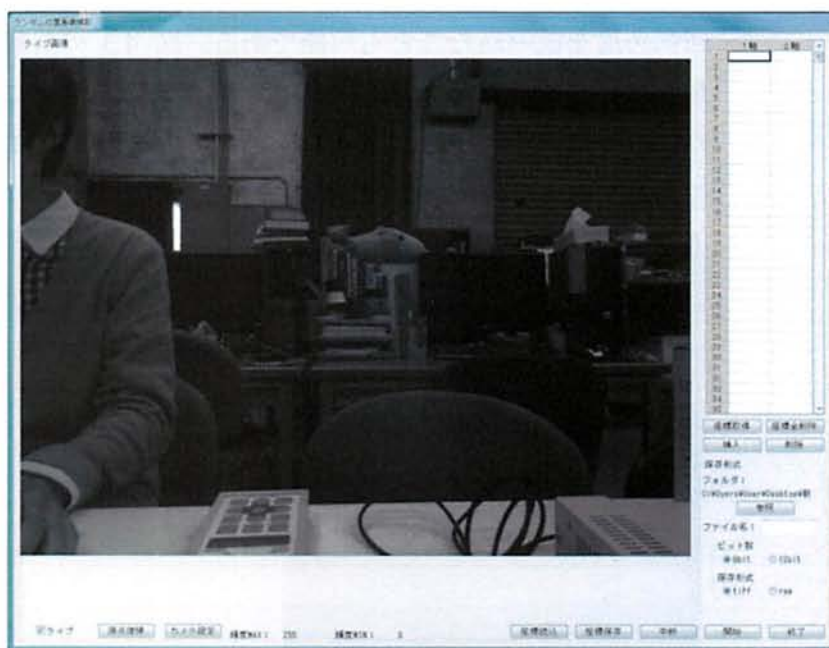


図 4.2 パソコン画面上のカメラ制御用ソフトウェア

このカメラシステムは、ハードウェアを中央精機(株)に、ソフトウェアを MITANI CORPORATION(株)に発注した、オーダーメイド型ハードウェアカメラシステムである。X 軸周りの横回転ステージ、Y 軸周りの縦回転ステージの二つを連動した二次元回転が可能である。また、視軸周りの回転は、奥行復元に不要なため今回の設計では採用していない。カメラの性能に関しては以下の通りである。

- 焦点距離：5 (mm.)
- 画素センサー：200 万 (1200×1600 (pix.))
- 可動域：x 軸 360°, y 軸 -10°～+10°
- 駆動最小単位 1 (pulse)
  - x 軸：1 (pulse) = 0.01 (rad.)
  - y 軸：1 (pulse) = 0.00067 (rad.)
- 撮影画像の種類：8, 12 ビットグレイスケール画像
- 画像フォーマット：tiff, raw

上記の性能に加え、シャッタースピードやゲイン値の操作による撮影明度の調節の他、平行ステレオ視等も可能となっている。次にカメラの動作について説明する。図 4.1 に 1 軸、2 軸とあるが、それぞれ X 軸周りの回転角度、Y 軸周りの回転角度を入力する部分であり、この数値によって、カメラの回転角度が決まる。この数値は、テキストファイルで複数の回転角度を読み込むことも可能である。図 4.3 の装置 (QT-ADM2) はカメラと計算機を同期しながら操作するためのもので、角度を入力することで自動的にカメラが動作する仕組みになっている。また、図 4.4 のコントローラー (QT-AK) を利用することで駆動最小単位である 1pulse での手動操作が可能である。



図 4.3 カメラ駆動装置 (QT-ADM2)



図 4.4 手動コントローラー (QT-AK)



実際に実験を行う場合の準備として、図 2.1 のカメラモデルを元に、水平な台に対象物とカメラを配置し（図 4.5 参照）、それぞれの位置関係をあらかじめ測定する必要がある。その後、上述の動作方法の通りに角度を入力し、指定回数カメラを動かすことで画像を複数枚撮影することができる。なお、画像撮影時の形式はその後の数値計算による処理の事を考え、生データとして扱える RAW 形式を選択した。



図 4.5 実験環境

最後に、次章で数値評価を行うに当たって、ここまでの奥行復元過程を図 4.6 に示す。

#### 実験の流れ

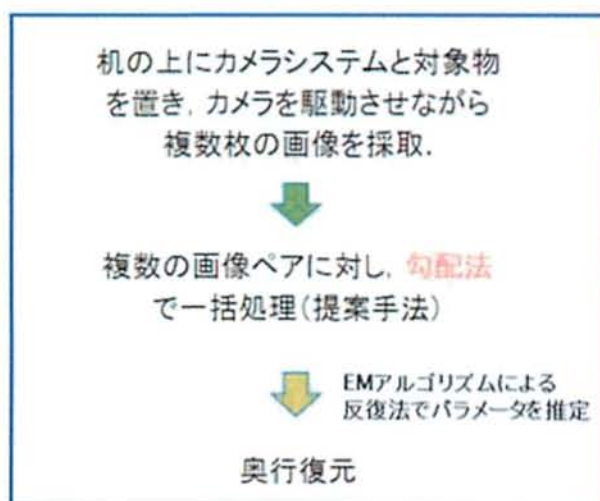


図 4.6 奥行復元の流れ

## 4.2 カメラキャリブレーション

### 4.2.1 キャリブレーションの原理

ここまでカメラのシステムと動作について説明してきたが、実際に画像を取得する際には、より正確な画像を取得するために、事前にカメラキャリブレーションを行う必要がある。本研究で行うカメラキャリブレーションは有名な Zhang の方法 [15] を利用する。特徴として、キャリブレーションを行うための画像を取得する際に制限なく撮影できる利点がある。カメラキャリブレーションとは、主にカメラの内部特性（内部パラメータ）と外部特性（外部パラメータ）を求めることを言う。内部パラメータとは以下の式で表され、 $x$  軸方向のピクセル単位の焦点距離  $f_w$ 、 $y$  軸方向のピクセル単位の焦点距離  $f_h$ 、光軸の中心  $x_0, y_0$  で表される  $3 \times 3$  行列  $\mathbf{A}$  であり、画像の補正に利用される。本研究では、画像の画素輝度値をメインで扱うため、内部パラメータを中心に導出過程を説明し、外部パラメータについては省略する。また、カメラの回転を  $\mathbf{R} = [\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z]^T$ 、並進を  $\mathbf{t}$  とする。

$$\mathbf{A} = \begin{bmatrix} f_w & 0 & x_0 \\ 0 & f_h & y_0 \\ 0 & 0 & 1 \end{bmatrix} \quad (4.1)$$

画像上の点  $\mathbf{m} = [x, y]^T$  と三次元空間上の点  $\mathbf{M} = [X, Y, Z]^T$  のそれぞれの同次座標  $\tilde{\mathbf{m}} = [x, y, 1]^T$ 、 $\tilde{\mathbf{M}} = [X, Y, Z, 1]^T$  とすると、 $\tilde{\mathbf{M}}$  と  $\tilde{\mathbf{M}}$  を画像平面上に投影した  $\tilde{\mathbf{m}}$  について、次式の関連式で表せる。

$$s\tilde{\mathbf{m}} = \mathbf{A}[\mathbf{R} \quad \mathbf{t}]\tilde{\mathbf{M}} \quad (4.2)$$

ここで、 $s$  はスケールである。

次に、ワールド座標において、三次元モデルのある位置を  $Z = 0$  と設定することで、式(4.2)を次式のように簡単化できる。

$$s \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_x & \mathbf{r}_y & \mathbf{r}_z & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{r}_x & \mathbf{r}_y & \mathbf{t} \end{bmatrix} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \quad (4.3)$$

以上から  $\mathbf{M} = [X, Y]^T$ 、 $\tilde{\mathbf{M}} = [X, Y, 1]^T$  と変更できる。

三次元点 $\tilde{\mathbf{M}}$ を変更したことで、画像上の点 $\tilde{\mathbf{m}}$ と同じ次元なったため、 $3 \times 3$ のホモグラフィ行列 $\mathbf{H}$ を用いて、次式の関連式で表せる。

$$s\tilde{\mathbf{m}} = \mathbf{H}\tilde{\mathbf{M}} \quad (4.4)$$

$$\mathbf{H} = \mathbf{A}[\mathbf{r}_x \quad \mathbf{r}_y \quad \mathbf{t}] \quad (4.5)$$

ここで、モデル平面の画像からホモグラフィ行列 $\mathbf{H} = [\mathbf{h}_1 \quad \mathbf{h}_2 \quad \mathbf{h}_3]$ を推定する[10]。推定には最尤推定[14]を用い、画像上の点 $\tilde{\mathbf{m}}_i$ に発生する画像上のノイズを平均0としたときの共分散行列を $\Lambda_{m_i}$ とすると、推定のために最小化すべき式は次式となる。（ $i$ は対応点の数）

$$\sum_i (\mathbf{m}_i - \hat{\mathbf{m}}_i)^T \Lambda_{m_i}^{-1} (\mathbf{m}_i - \hat{\mathbf{m}}_i) \quad (4.6)$$

このとき、 $\hat{\mathbf{m}}_i$ に対して、対応する $\hat{\mathbf{M}}_i$ は次式の関連式で表せる。

$$\hat{\mathbf{m}}_i = \frac{1}{\mathbf{h}_3^T \mathbf{M}_i} \begin{bmatrix} \mathbf{h}_1^T \mathbf{M}_i \\ \mathbf{h}_2^T \mathbf{M}_i \end{bmatrix} \quad (4.7)$$

簡単のために $\Lambda_{m_i} = \sigma^2 \mathbf{I}$ とすると、最小化すべき式は、次式のようにになる。

$$\sum_i \|\mathbf{m}_i - \hat{\mathbf{m}}_i\|^2 \quad (4.8)$$

この場合、非線形最小化問題となるため、Levenberg-Marquardt法（以降LM法）[16]を用いる。式(4.7)と式(4.8)から、式(4.4)、式(4.5)を書き換えると、次の方程式で表せる。

$$\hat{\mathbf{m}}_i = \begin{bmatrix} \tilde{\mathbf{M}}^T & 0 & -x\tilde{\mathbf{M}}^T \\ 0 & \tilde{\mathbf{M}}^T & -y\tilde{\mathbf{M}}^T \end{bmatrix} \begin{bmatrix} \mathbf{h}_1^T \\ \mathbf{h}_2^T \\ \mathbf{h}_3^T \end{bmatrix} = 0 \quad (4.9)$$

上記の式に対しLM法を用いて解くことで、 $\mathbf{H}$ の推定が完了する。また、 $\tilde{\mathbf{M}}^T$ の行列の列数は9あるため、最低5個の平面上の点对応が必要となる。

以上の過程から求めたホモグラフィー行列  $\mathbf{H}$  と内部パラメータ  $\mathbf{A}$  には,  $\mathbf{r}_x, \mathbf{r}_y$  が既知という条件の下, 以下の2つの制約式が成り立つ.

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 = 0 \quad (4.10)$$

$$\mathbf{h}_1^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_1 = \mathbf{h}_2^T \mathbf{A}^{-T} \mathbf{A}^{-1} \mathbf{h}_2 \quad (4.11)$$

上記の式が意味することは,  $\mathbf{H}$  と  $\mathbf{A}$  がお互いに正規直交であることを示しており, 内部パラメータに対し2つの制約を課すことができる. 2のみの制約である理由としてホモグラフィー行列  $\mathbf{H}$  の自由度8に対し, 回転  $\mathbf{R} = [\mathbf{r}_x, \mathbf{r}_y, \mathbf{r}_z]^T$  の自由度3, 並進  $\mathbf{t}$  の自由度3の外部パラメータが含まれており, 残りの自由度が2だからである.

以上の条件から, 実際に内部パラメータ  $\mathbf{A}$  を算出していく. 使用する方法は線形法で, まず内部パラメータ  $\mathbf{A}$  を用いて  $\mathbf{B}$  に変形する. その関係式は次式で表される.

$$\mathbf{B} = \mathbf{A}^{-T} \mathbf{A}^{-1} = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \mathbf{B}_{13} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \mathbf{B}_{23} \\ \mathbf{B}_{31} & \mathbf{B}_{32} & \mathbf{B}_{33} \end{bmatrix} \quad (4.12)$$

ホモグラフィー行列  $\mathbf{H}$  を推定した時と同様に,  $\mathbf{B}$  に対しても最尤推定[14]を行い, 解を推定する. 求めた  $\mathbf{B}$  の値から式(4.12)より, 4つの内部パラメータ  $f_w, f_h, x_0, y_0$  を特定することができる.

以上の方法から求めた内部パラメータは, 相対的なものであるため, 絶対距離としては物理的な意味をなさない. よってこれも非線形最小化問題として推定する. 画像上の点  $\tilde{\mathbf{m}}_j$  に対し, 独立で一様な分布を持つノイズを想定し, 以下の式を最小化することで推定を行う.

$$\sum_{i=1}^n \sum_{j=1}^m \left\| \mathbf{m}_{ij} - \hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j) \right\|^2 \quad (4.13)$$

$\mathbf{m}_{ij}$  は, 実際に観測したときの,  $i$  枚目の画像での  $j$  点目の部分であり,  $\hat{\mathbf{m}}(\mathbf{A}, \mathbf{R}_i, \mathbf{t}_i, \mathbf{M}_j)$  は, キャリブレーション法で  $\mathbf{M}_j$  から推定した画像上の点を表している. 式(4.13)の最小化の際も LM 法を用いて最適化を行うことで, より正しい内部パラメータを推定することができる.



### 4.2.2 キャリブレーションの手順

キャリブレーションを行う準備として、まず、図 4.7 のような白黒のチェスボードパターンを用意する。次に図 4.8 のように、そのボードの向きや位置を変えながら複数回撮影を行う（白黒のコーナーが全て画面内に入るように・保存形式は TIF）。ランダムにできるだけ多く撮影してもキャリブレーションは可能だが、動かし方によっても求められる焦点距離の精度があがるパターンが存在するので、[17]を参考にしてみても良い。

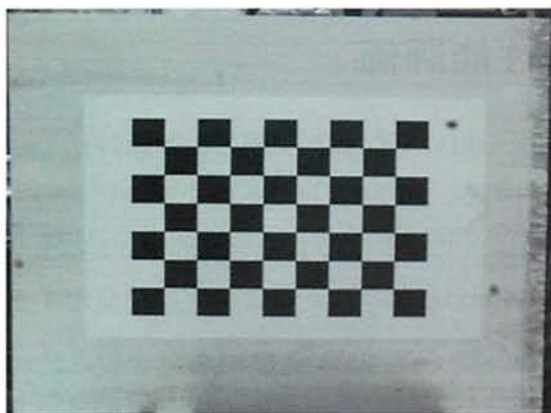


図 4.7 チェスボードパターン

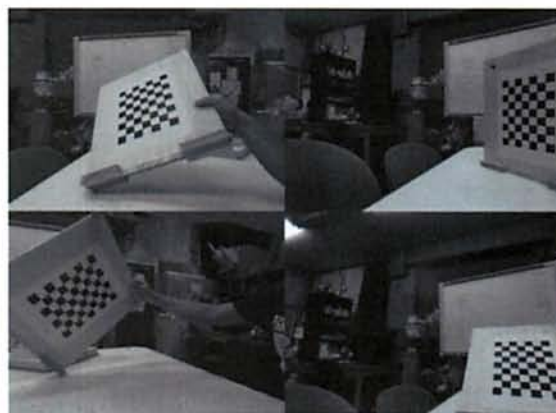


図 4.8 ボードを動かして撮影

撮影した画像 1 枚 1 枚に対し、輝度値の変化が激しいコーナーを特徴点として抽出する。図 4.7 の場合、画像 1 枚における特徴点の数は  $8 \times 6 = 48$  となる。

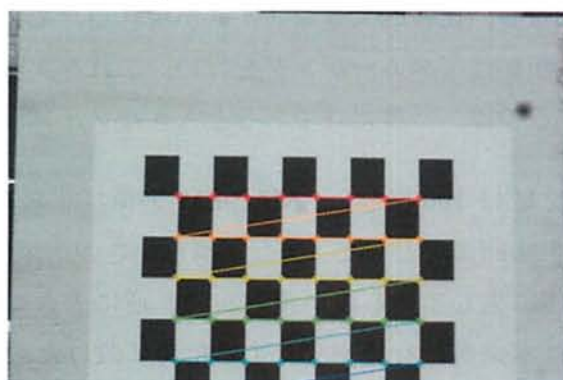


図 4.9 コーナーの検出

図 4.9 のように段階ごとに色分けを行いながら、特徴点を抽出する。その後、抽出された特徴点をもとに、4.2.1 の原理に沿って内部パラメータを求める。本来ならば、この過程で得られた内部パラメータは画像補正に利用されるのだが、本研究ではこれに含まれるピクセル単位の焦点距離  $f$  が、実験での参考値となるので、校正を通して求めた。

なお、この一連の動作を行うためのプログラムは Tsai のキャリブレーション手法の原理をベースとしたサンプルがあり、参考にそのアドレスを記載する。

[http://opencv.jp/sample/camera\\_calibration.html](http://opencv.jp/sample/camera_calibration.html)



## 第 5 章 奥行復元の性能評価

### 5.1 実画像による奥行復元の性能評価

以下にカメラシステムを用いて撮影した実画像を用いて行った実験での数値評価の結果を示す。まずその前に実験での設定について説明する。図 2.1, 図 2.2 のモデルにおいて、カメラ座標原点  $O$  から画像面までの焦点距離  $f$  のピクセル単位のを前節 4.2.2 のキャリブレーションを通してあらかじめ求め、 $f=1141.4(\text{pix})$  となった。ここで実際に使用する画像サイズは  $f=1$  のときのをカメラモデルで定義しているの、例えば実験で扱う画像の画素数が  $256 \times 256(\text{pix})$  であるときは、縦横それぞれの画素数を  $f(\text{pix})$  で割ったものつまり、 $256(\text{pix}) / 1141.4(\text{pix}) = 0.2243$  が設定すべき画像サイズとなる。また、実際の撮影画像は  $1200 \times 1600(\text{pix})$  の 8bit のグレースケール画像であるため、リサンプリングを施し実験時に扱いやすい  $256 \times 256(\text{pix})$  に縮小する。加えて実画像の場合、撮影した時点で輝度値の空間的な波長に対し、比較的大きなノイズが現れるため、前もって画像に対しガウシアンフィルタを複数回かけてスムージングを行い、これを実験画像とした。ガウシアンフィルタを複数回かけるのは、一度だけであると大きなノイズが残ってしまうからである。

使用する画像の枚数  $M$  は 100 枚とし、MAP-EM アルゴリズムの反復は奥行の形が基本的に変化しなくなる 500 回までとした。また、滑らかさの拘束度合いを表す  $\sigma_d$  は、値を小さくすると奥行全体に生じている波が滑らかに、逆に大きくなると波が粗くなるように作用する。この値は本来奥行を見やすくするために設けた設定値であり、過剰に利かせると本来復元すべき奥行が出ない可能性も否めない。なので、ある程度利かせた  $\sigma_d = 1.0 \times 10^{-4}$  を基準とし、条件別に比較するための復元奥行すべてに適用している。

また閾値については基準となる単位が存在しないため、実験比較の基準としてしまうと複雑になってしまう。そのため、3.2.3 節で説明した空間勾配の向きと差ベクトルの大きさを考慮した 2 段階の勾配決定法に基づいて閾値を設けた。まず第一段階として勾配の向きが 90 度である時を閾値の基準とし、次に第二段階として図 5.1 における中央立体付近全画素での勾配の差ベクトルの算術平均をもう一つの閾値の基準とした。向きと差ベクトル共に閾値の基準を下回った場合に採択する。この過程で結果的に残った画像ペアの割合（画像ペア使用率(%)）を参照する。残った全ての画像ペアに関しては 2 次項が小さいた

め、あらかじめ高次項を考慮した式(3.27), 式(3.28)の勾配を使用している。

この画像ペア使用率(%)を調整しながら、既提案手法及び提案手法の奥行精度に関して比較評価を行う。比較方法としては、実験画像の奥行に対し、正解奥行を同時に表示することで形状の比較を行う。より分かりやすくするため、奥行の断面図を通して確認する。表示する断面は全て、3D グラフの中央  $y=64$  を  $x$  軸に沿って切り出した部分としている。この場合、正解奥行に近いほど本手法が有効であることを示しており、実験画像における正解奥行との RMSE を様々な条件から検討することでも、有効性を確認できる。

使用画像については、実験画像を図 5.1 に、平行ステレオ[付録 1]で復元された正解奥行マップを図 5.2 に示す。図 5.1 において、白黒のマス 2 つ分は輝度値の空間的な波長  $\lambda$  の 1 波長分を表しており、それに対する画像変位を最適値  $\lambda/4$  及び  $\lambda/8$ ,  $\lambda/2$  の 3 つの異なる動きの大きさで比較実験を行う。参考に、それぞれの画像変位で設定されたカメラ回転の大きさ  $\sigma_r^2$  の数値について記載する。

画像変位  $\lambda/8 \Rightarrow \sigma_r^2 = 6.13 \times 10^{-6}(\text{rad})$ , 画像変位  $\lambda/4 \Rightarrow \sigma_r^2 = 2.33 \times 10^{-5}(\text{rad})$ ,

画像変位  $\lambda/2 \Rightarrow \sigma_r^2 = 9.39 \times 10^{-5}(\text{rad})$ . (以降記載される  $\sigma_r^2$  は、すべて単位を radian で統一)

図 5.2 では、横軸の面は画像上の位置  $(x, y)(\text{pix})$ , 縦軸は奥行  $Z(\text{mm})$  を示しており、奥行値は奥板が  $Z_b = 191.3(\text{mm})$ , 中央立体部分が  $Z_f = 141.5(\text{mm})$  と算出された。また、焦点距離  $f$  のキャリブレーション以外の導出方法として、この平行ステレオが挙げられる[付録 1]。この場合は  $f = 1168.9(\text{pix})$  と算出された。前にキャリブレーションで求めた  $f$  とどちらを選択するかは自由だが、本研究では奥板と中央立体部分の差分に着目しているため、その差分を正確に測定した値を利用できる後者で求めた  $f$  を採用した。それに合わせて画像サイズも変更しなければならない。

その他必要なパラメータとして、奥行のスケールを左右する、回転中心距離  $Z_0$  の導出[付録 2]も行わなければならない。このパラメータを正確に算出できないと、反復による奥行の収束値が絶対的な数値として見当はずれなものになってしまうからである。この場合は  $Z_0 = 38.8(\text{mm})$  と算出された。

以上から決定した設定値を実験に適応し、実画像による奥行復元結果の数値評価を行った。なお、実際に復元される奥行の縦軸の単位は全て mm 単位としている。

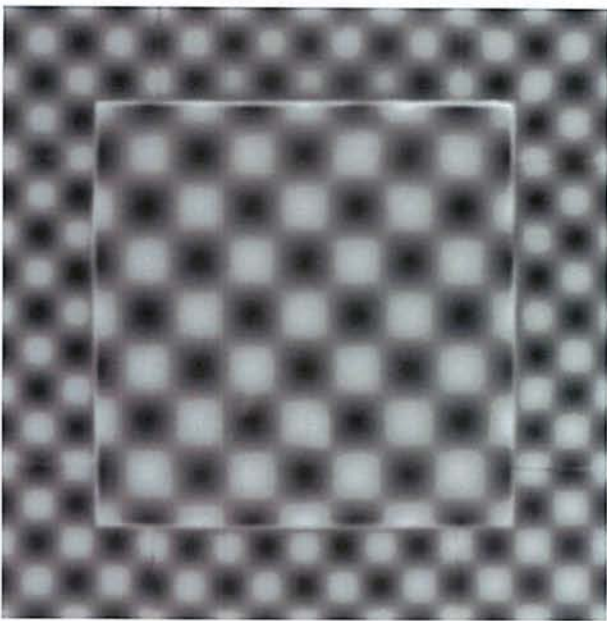


図 5.1 実験画像

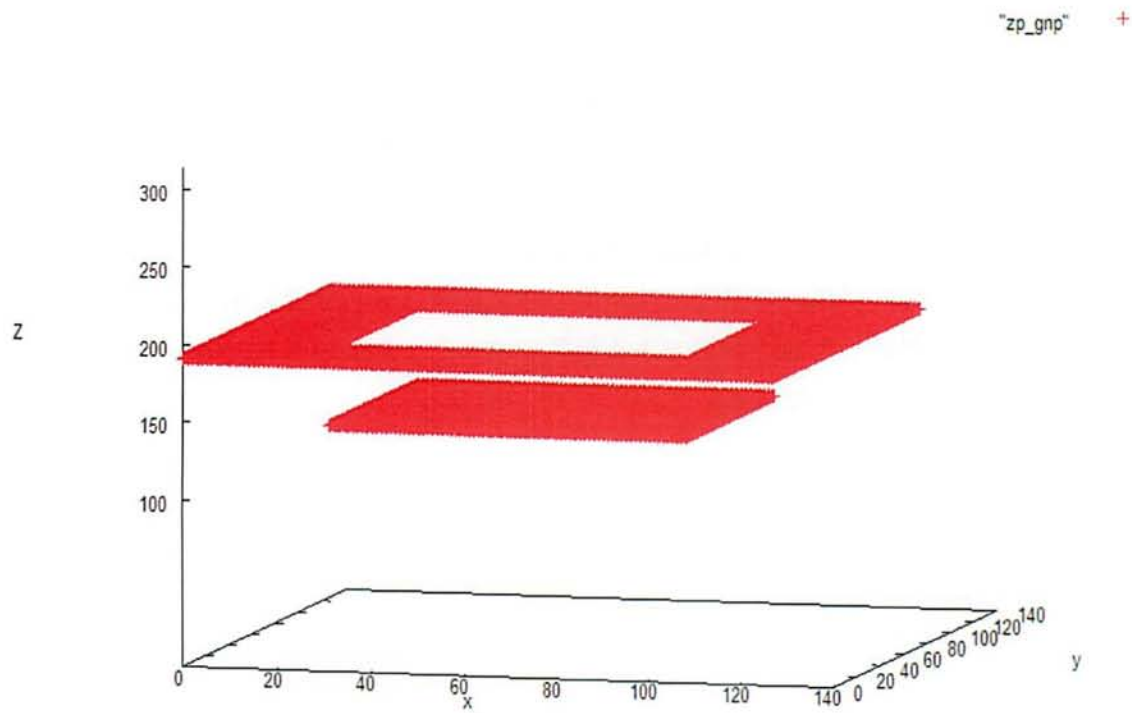
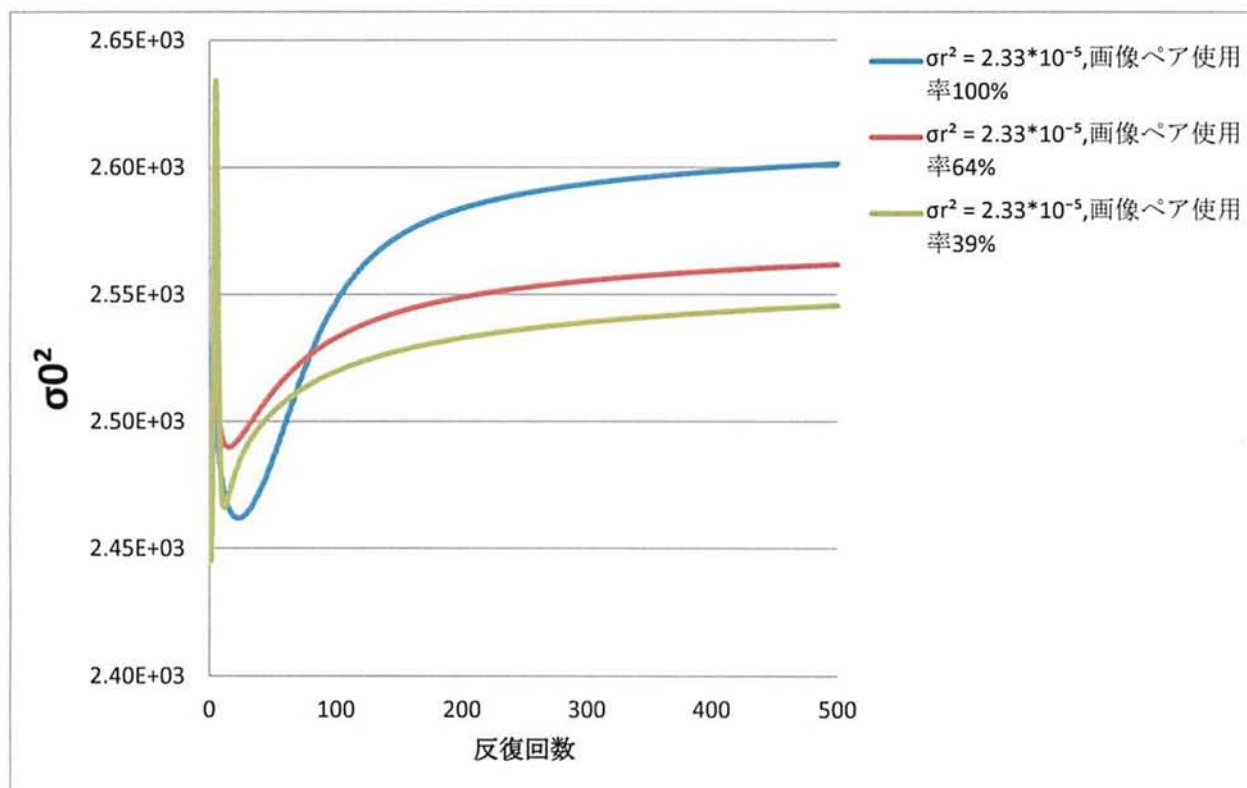
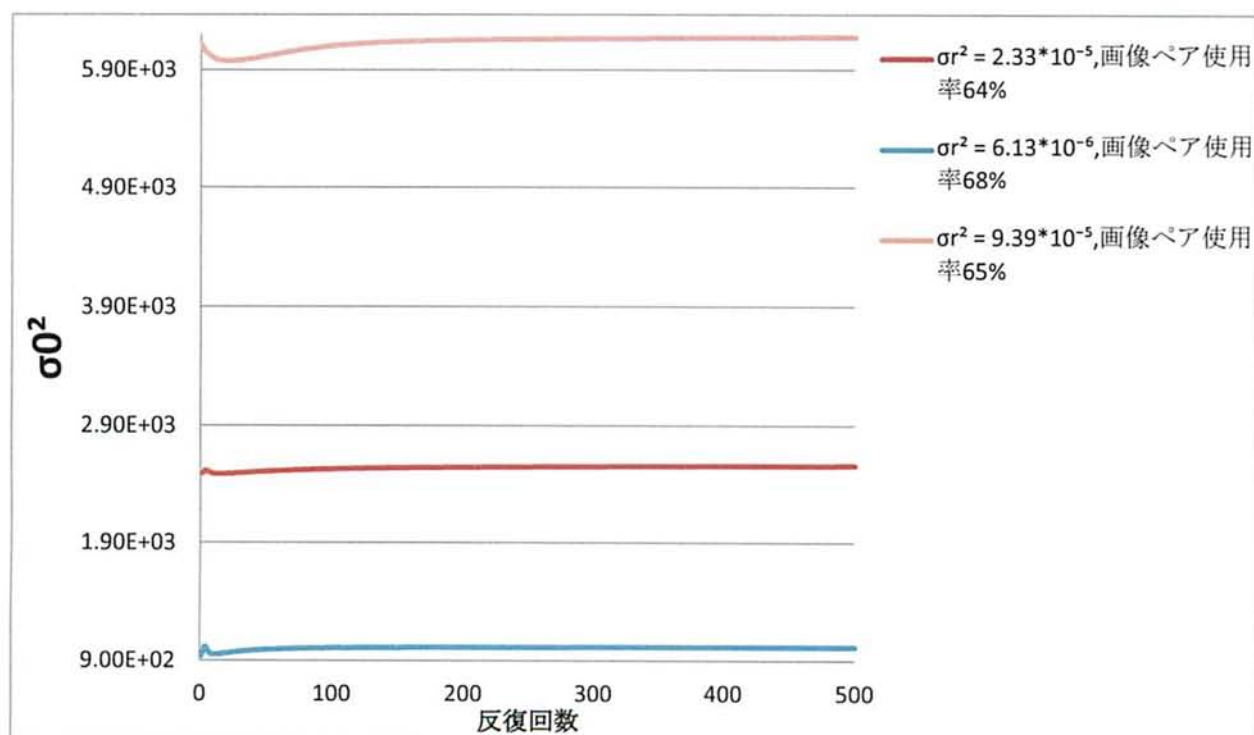


図 5.2 平行ステレオによる正解奥行マップ



5.1.1  $\sigma_o^2$  の収束

ここでは、滑らかさ拘束パラメータが  $\sigma_d^2 = 1.0 \times 10^{-4}$  の時、 $\sigma_r^2$  と画像ペア使用率をそれぞれ変更した際の収束回数による  $\sigma_o^2$  の変化を示す。

図 5.3 画像ペア使用率を変更した時の  $\sigma_o^2$ 図 5.4  $\sigma_r^2$  の大きさを変更した時の  $\sigma_o^2$

### 5.1.2 画像ペア使用率と $\sigma_r^2$ の関係

ここでは、滑らかさ拘束パラメータが  $\sigma_d^2 = 1.0 \times 10^{-4}$  の条件下で、ある同じ閾値を設けた時、 $\sigma_r^2$  の大きさによる画像ペア使用率の変化を示す。

なお、この画像の横軸は閾値の倍率を表しており、5.1 節 (p28 の最終段落) で説明した閾値の基準を 1 として定義している。

例えば、1.2, 1.4 のように横軸が上昇すれば閾値による制限が緩化され多くの画像ペアが選択される。逆に 0.6, 0.4 のように減少すれば条件を満たす画像ペアが少なくなり、使用率が下がることになる。

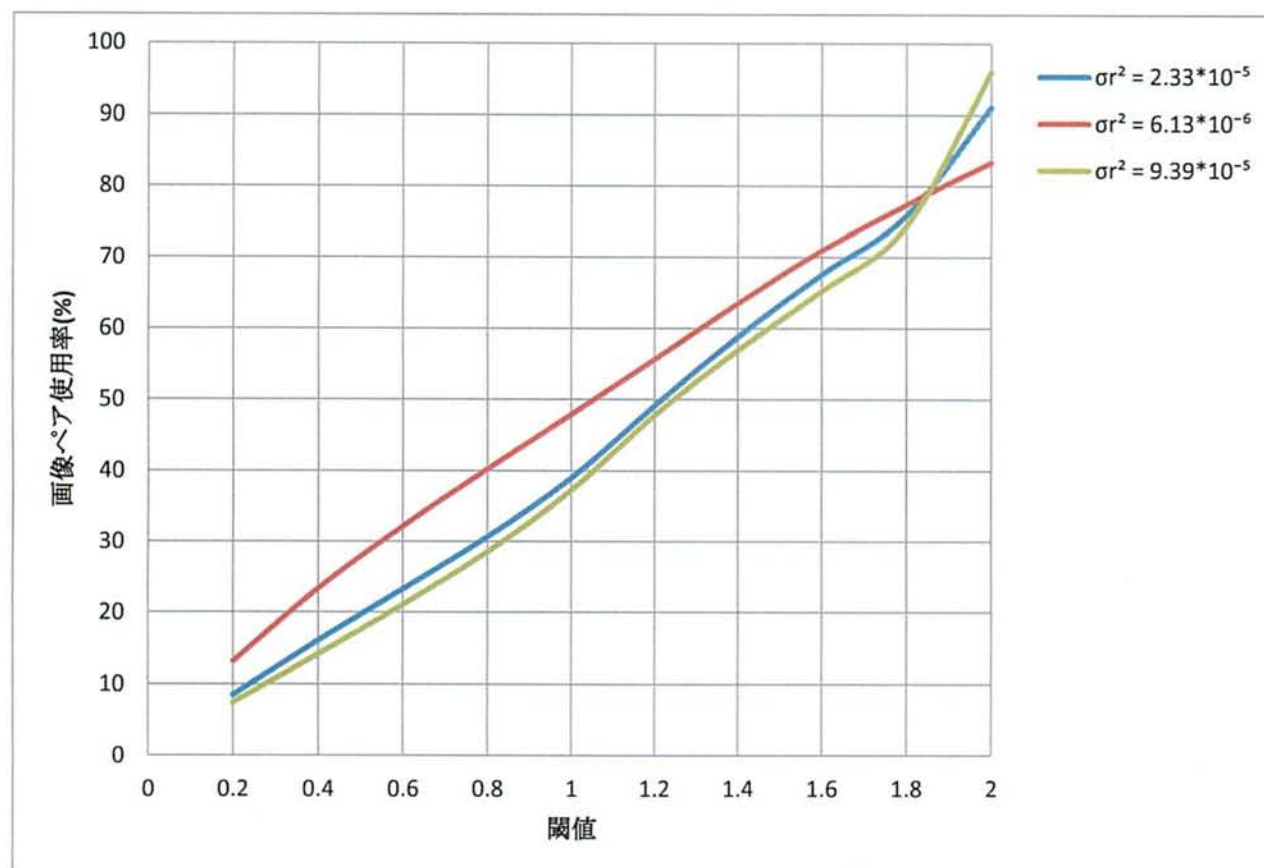


図 5.5  $\sigma_r^2$  の大きさによる画像ペア使用率の変化



### 5.1.3 $\sigma_r^2=6.13 \times 10^{-6}$ での復元結果と正解の比較

ここでは、 $\sigma_r^2=6.13 \times 10^{-6}$  の時、既提案法の奥行と、画像ペア使用率を変化させた時の奥行と、正解奥行と比較した時の断面図を示す。

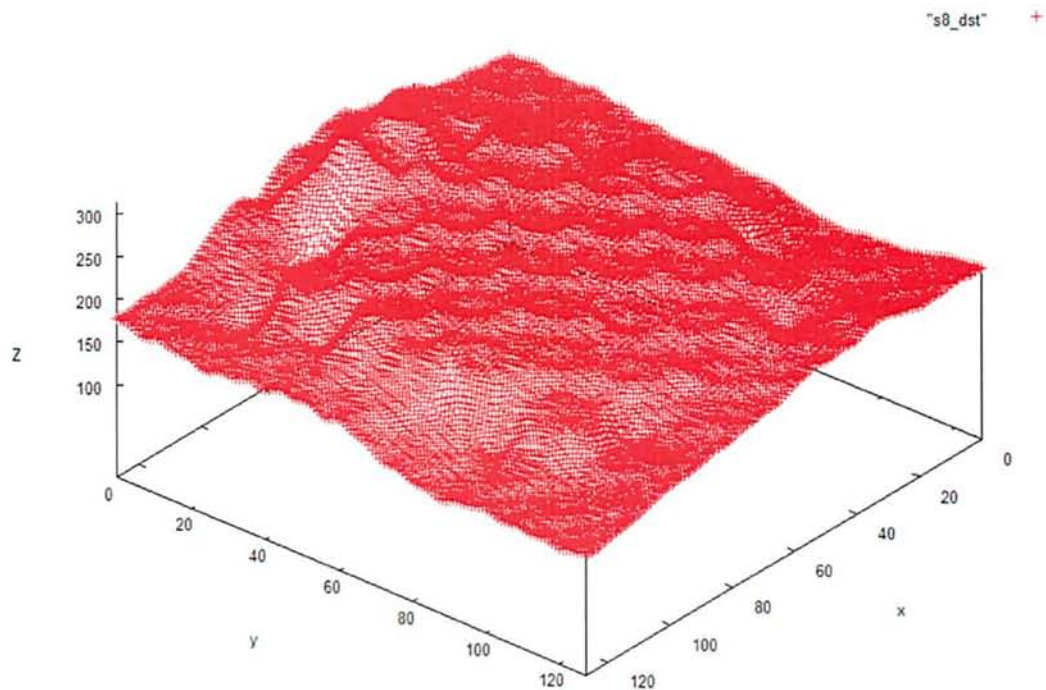


図 5.6 既提案法の奥行マップ

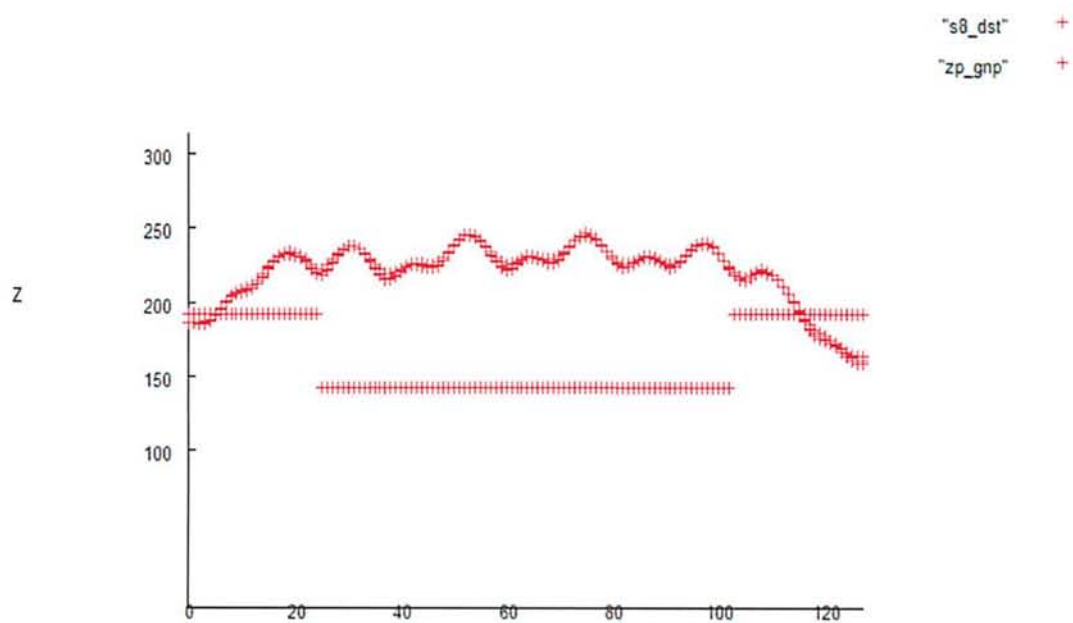


図 5.7 断面図による正解奥行との比較マップ

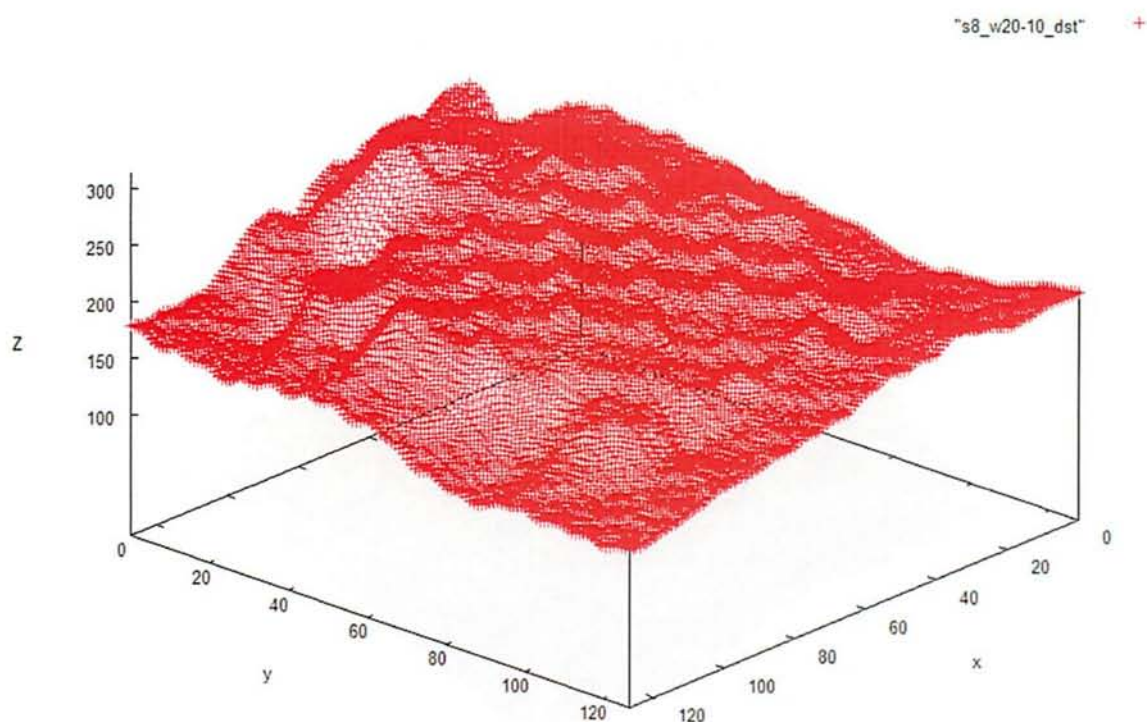


図 5.8 画像ペア使用率 83%時の奥行マップ

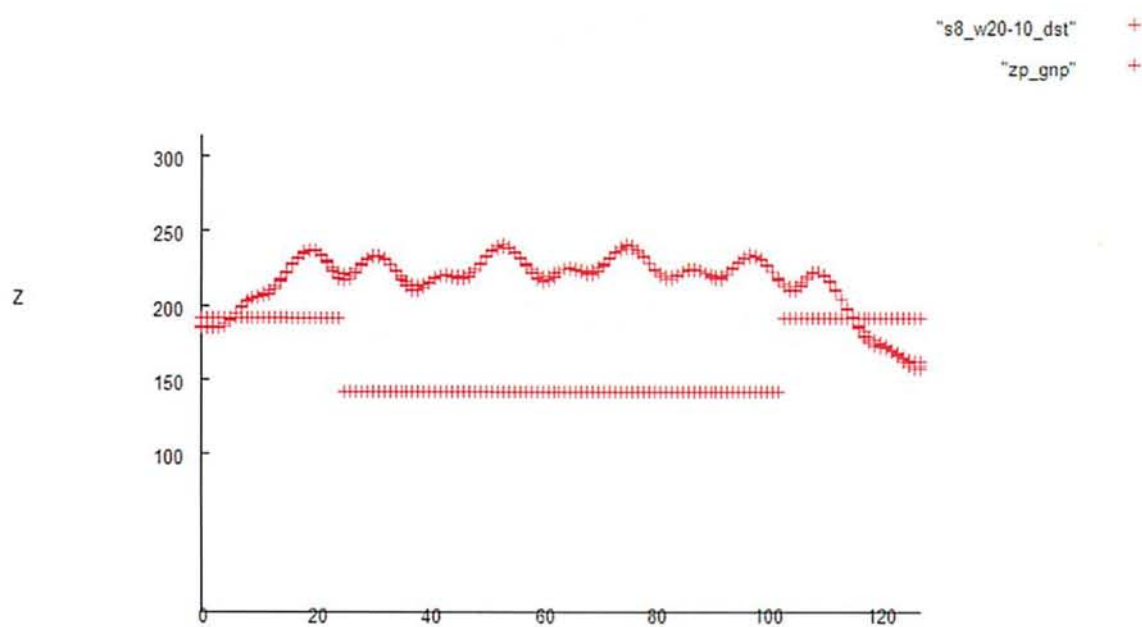


図 5.9 断面図による正解奥行との比較マップ

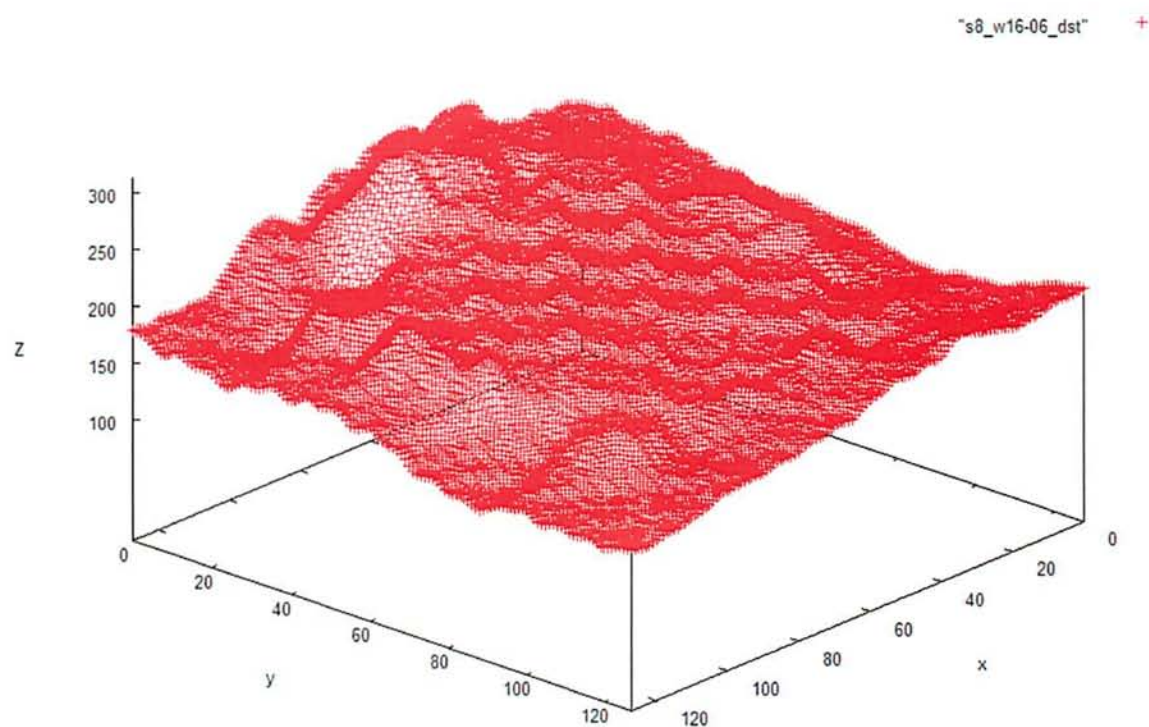


図 5.10 画像ペア使用率 71%時の奥行マップ

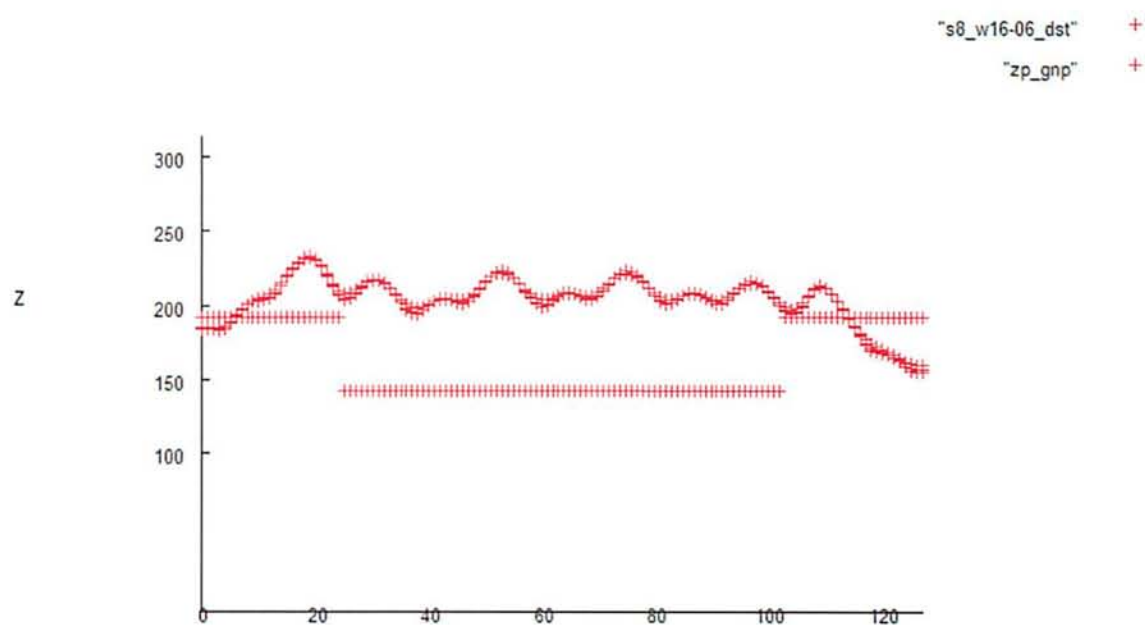


図 5.11 断面図による正解奥行との比較マップ

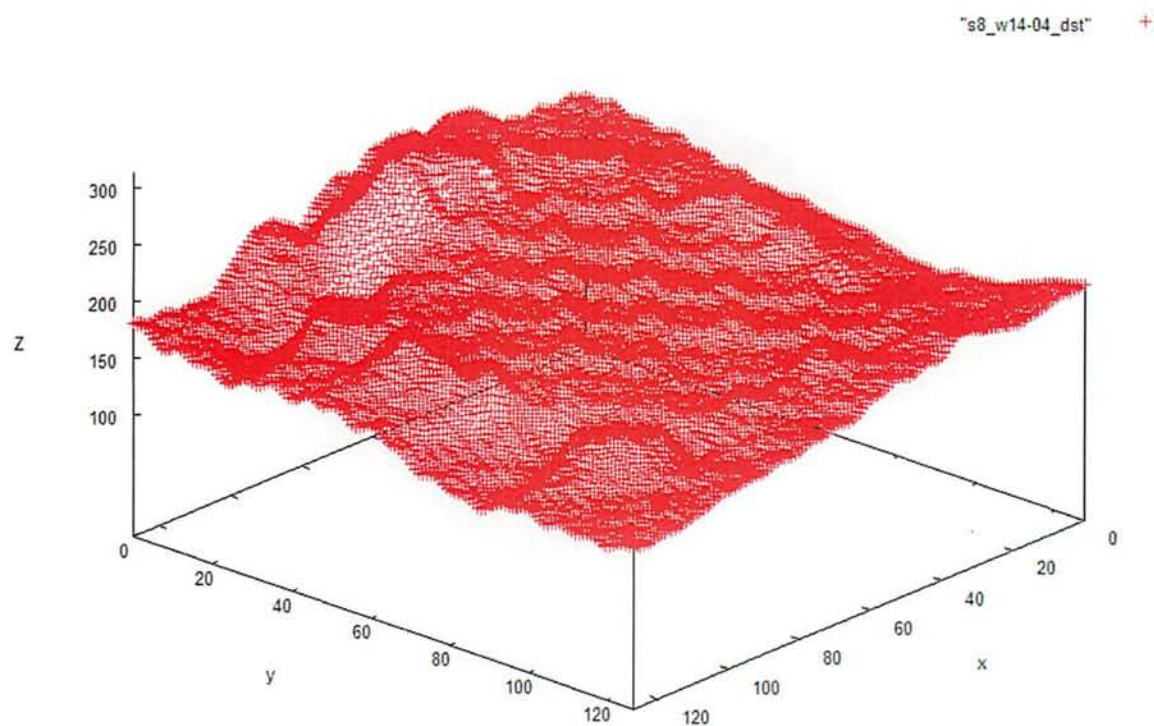


図 5.12 画像ペア使用率 64%時の奥行マップ

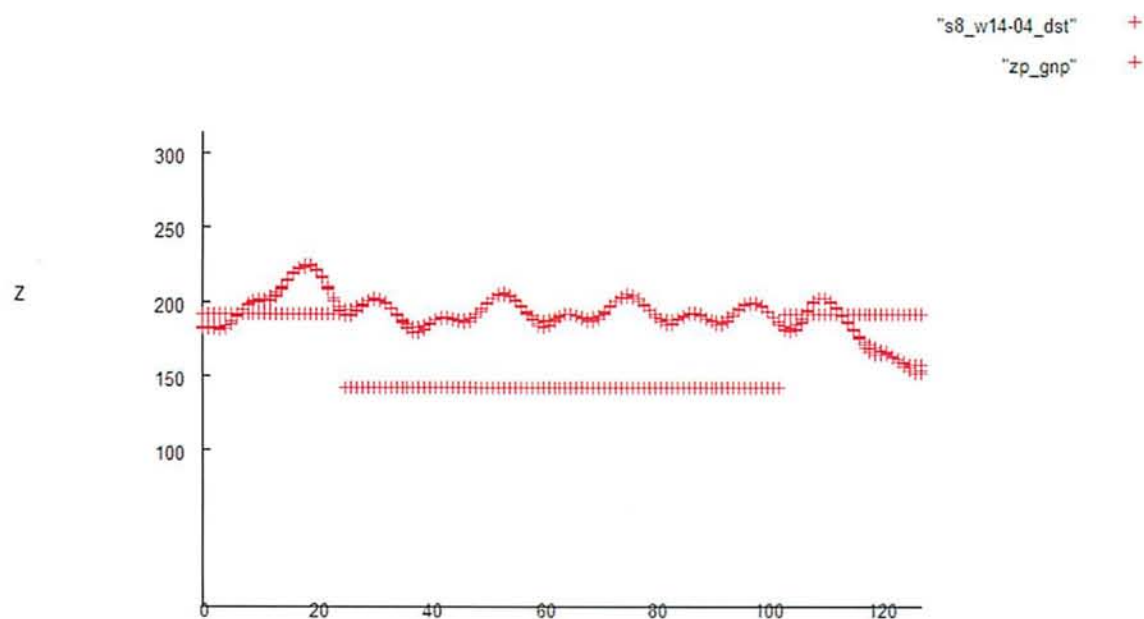


図 5.13 断面図による正解奥行との比較マップ



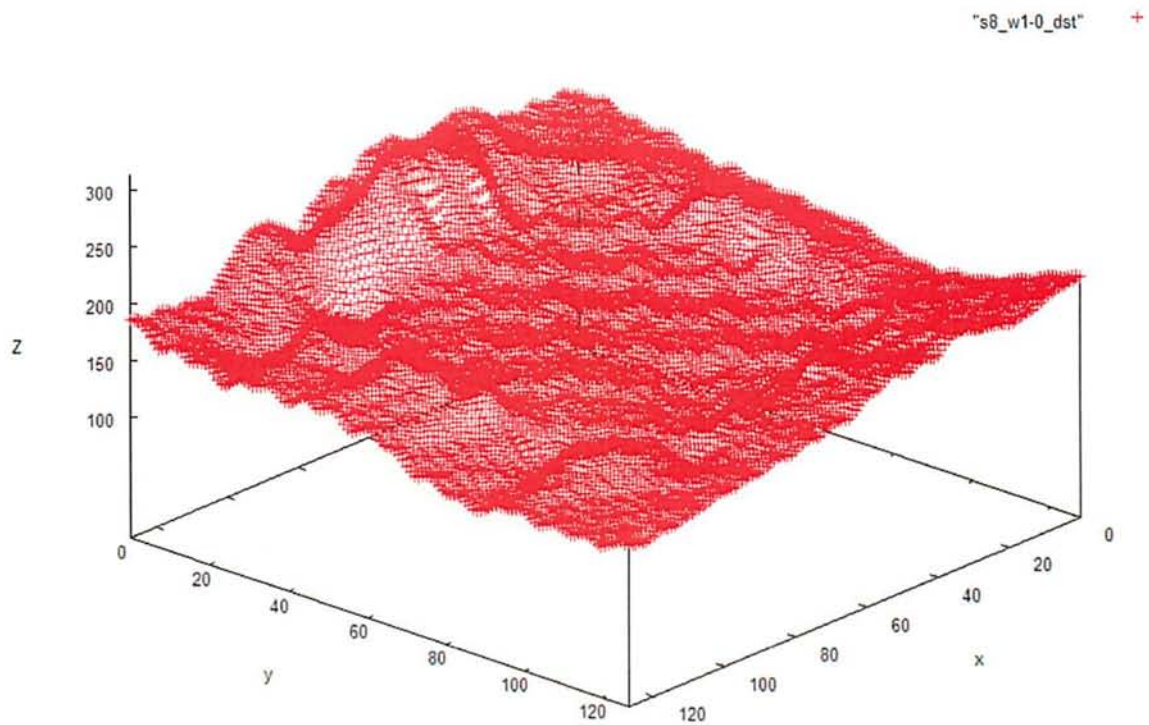


図 5.14 画像ペア使用率 48%時の奥行マップ

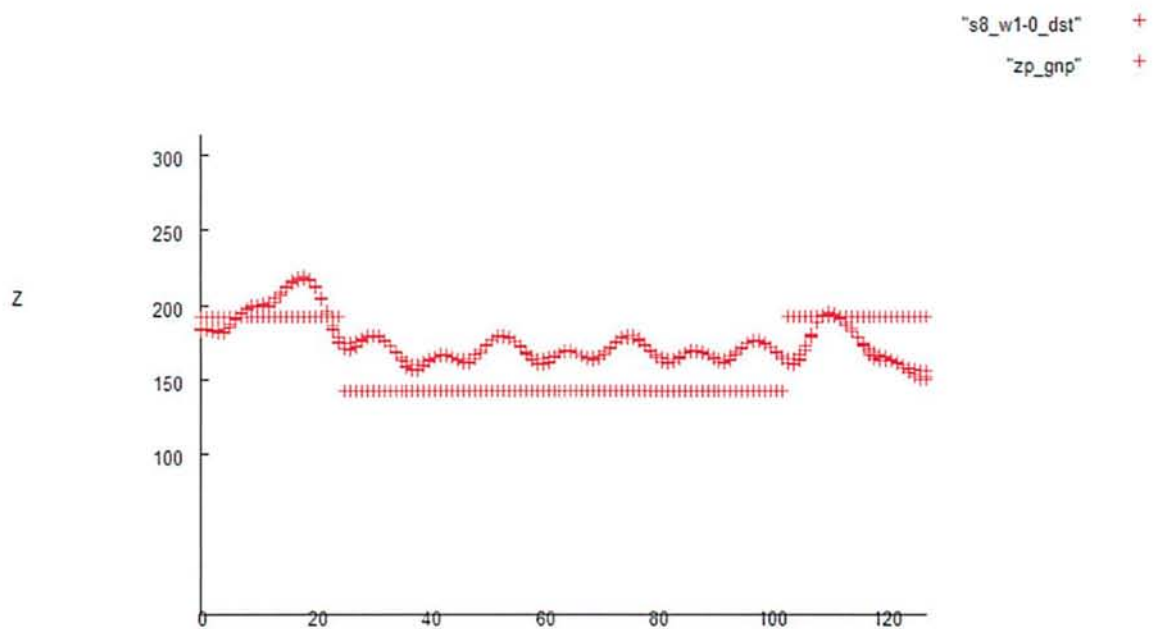


図 5.15 断面図による正解奥行との比較マップ



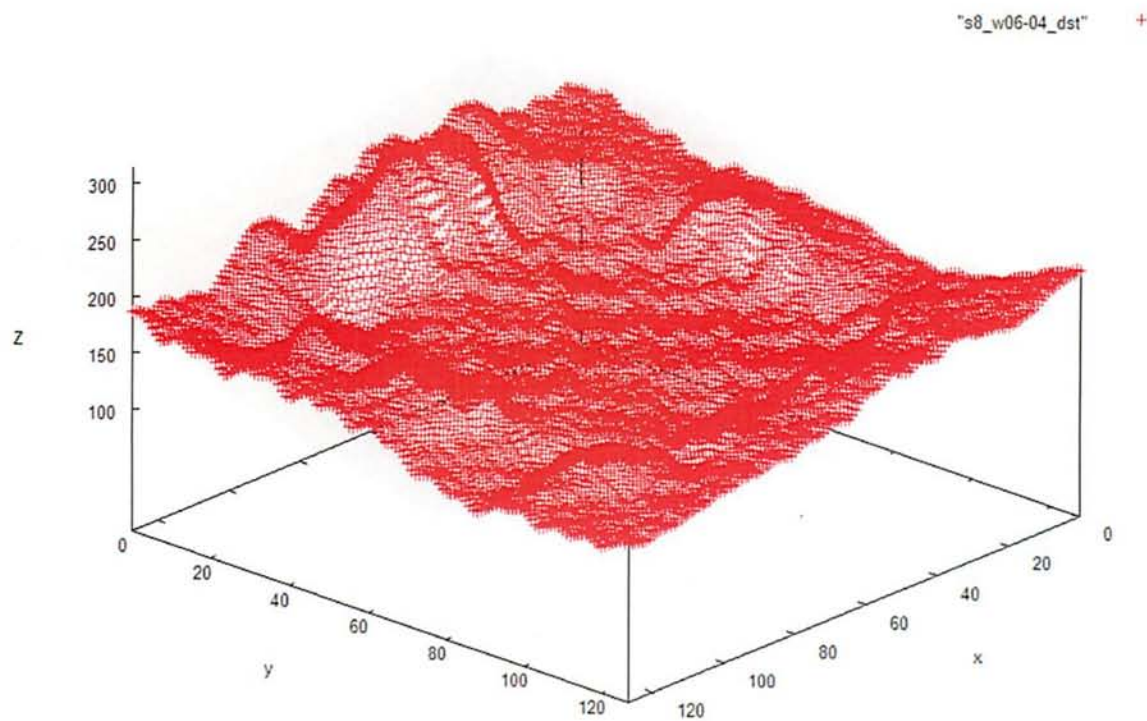


図 5.16 画像ペア使用率 32%時の奥行マップ

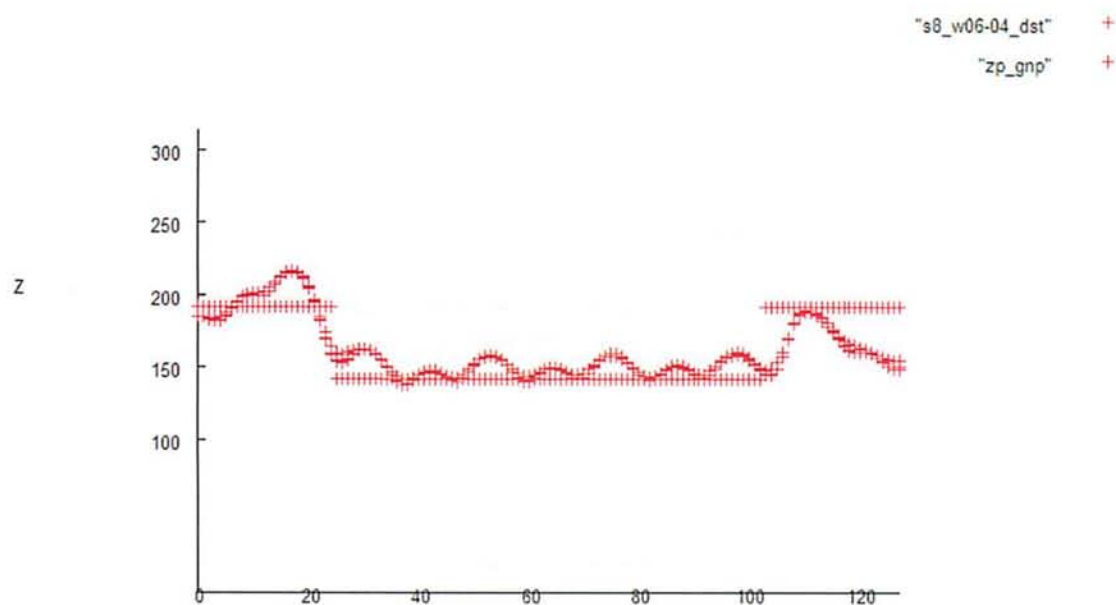


図 5.17 断面図による正解奥行との比較マップ

#### 5.1.4 $\sigma_r^2=2.33*10^{-5}$ での復元結果と正解の比較

ここでは、 $\sigma_r^2=2.33*10^{-5}$  の時、既提案法の奥行と、画像ペア使用率変化させた時の奥行と、正解奥行と比較した時の断面図を示す。

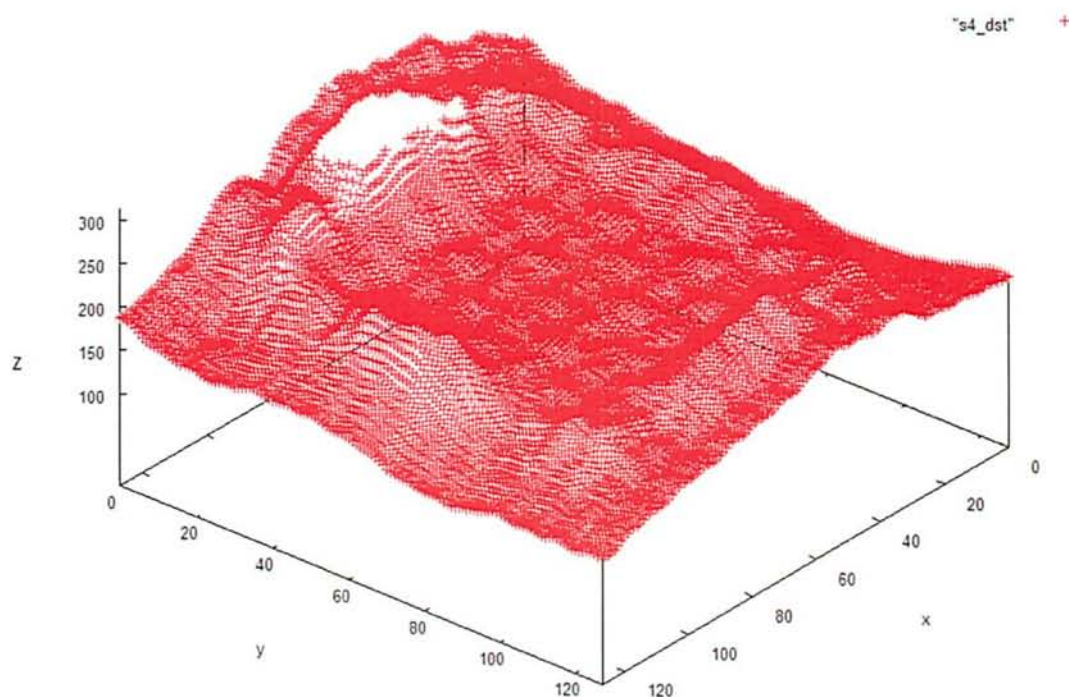


図 5.18 既提案法の奥行マップ

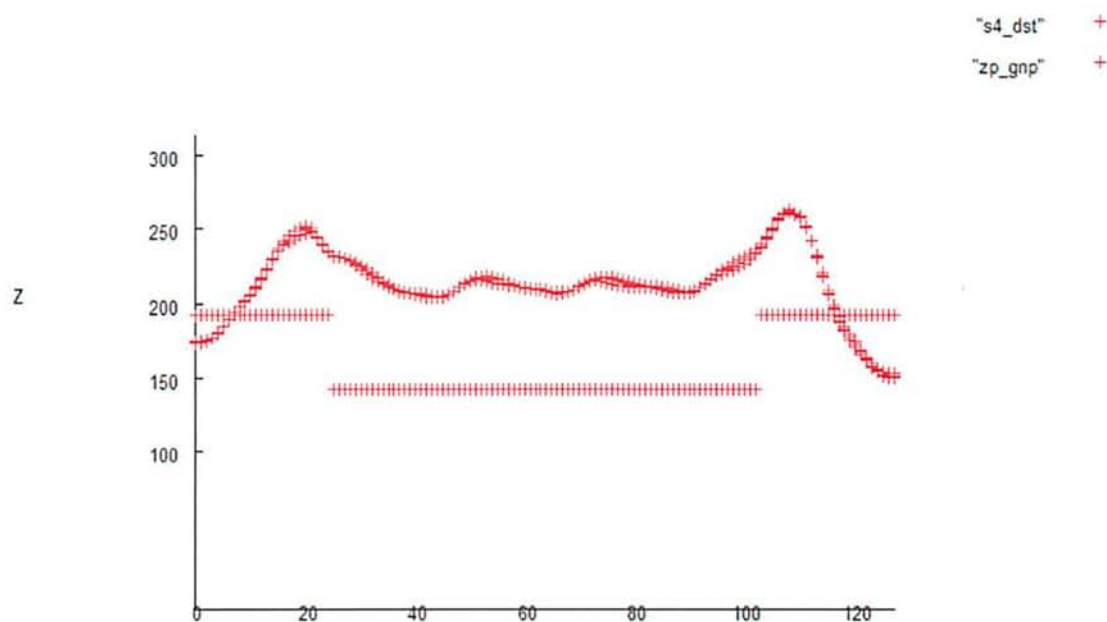


図 5.19 断面図による正解奥行との比較マップ

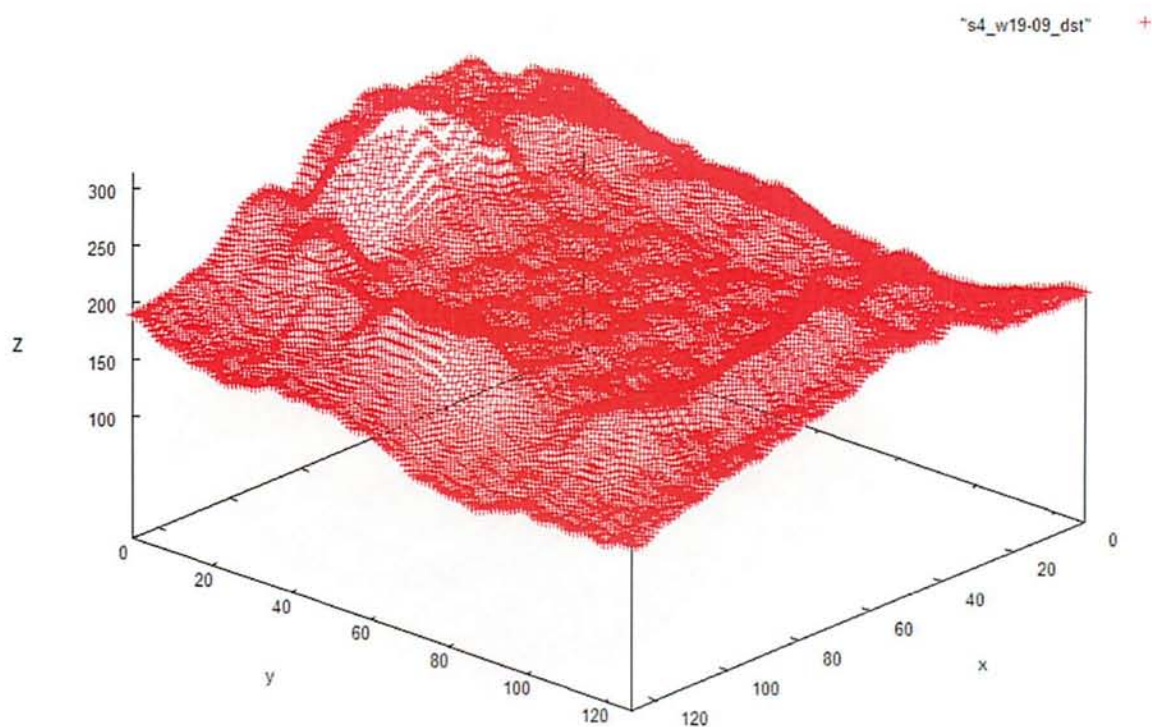


図 5.20 画像ペア使用率 83%時の奥行マップ

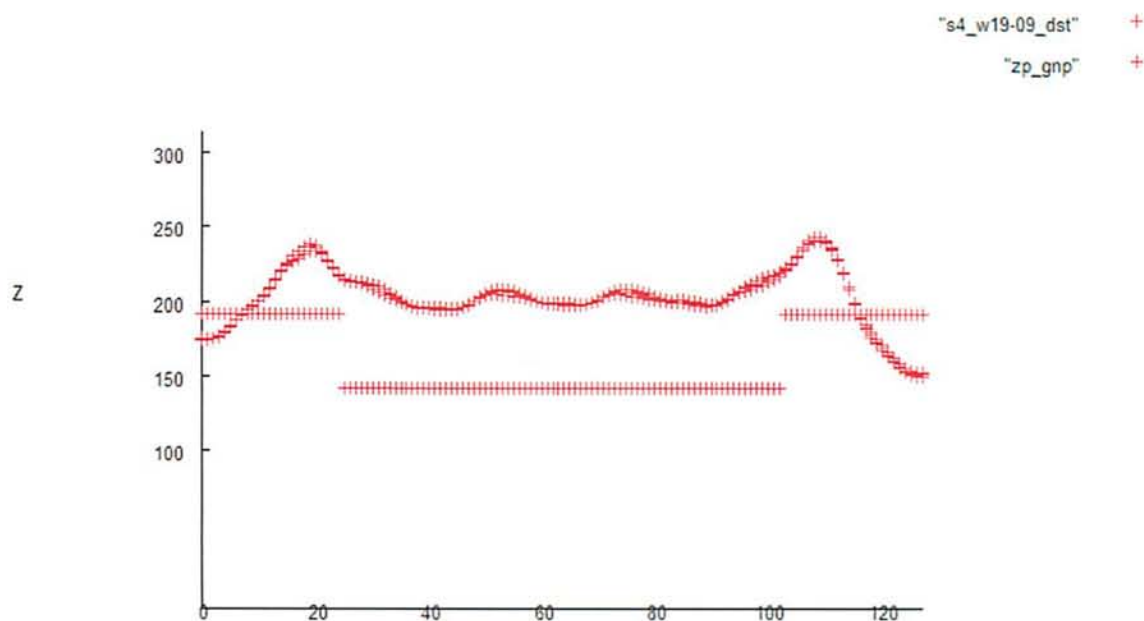


図 5.21 断面図による正解奥行との比較マップ

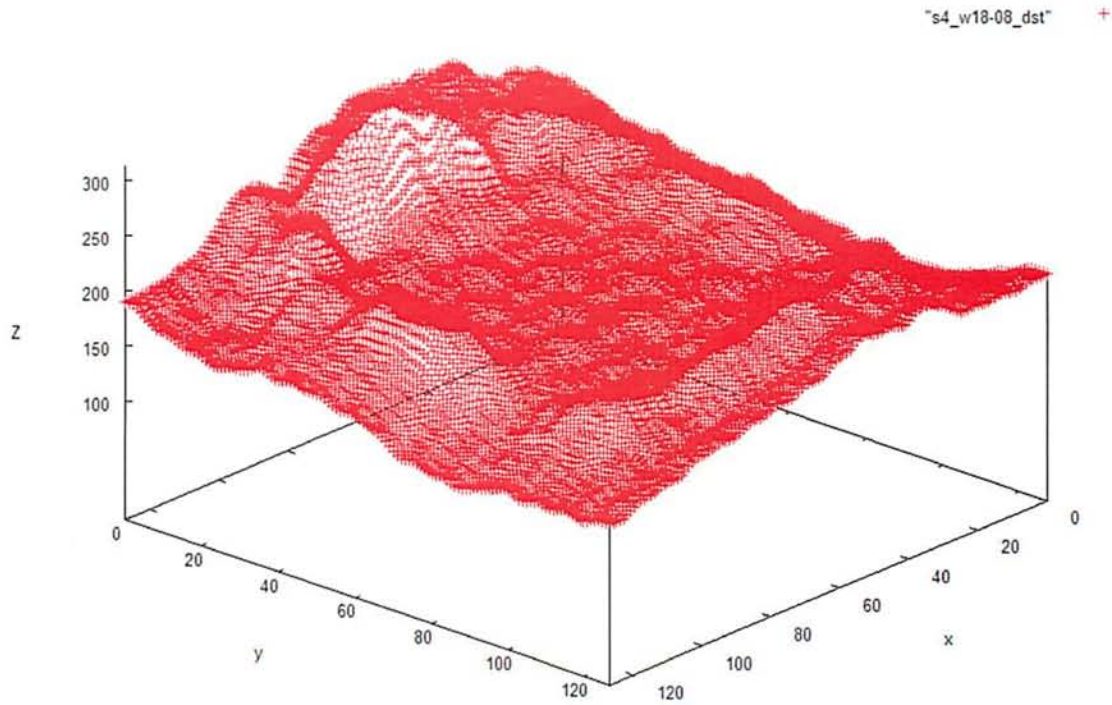


図 5.22 画像ペア使用率 76%時の奥行マップ

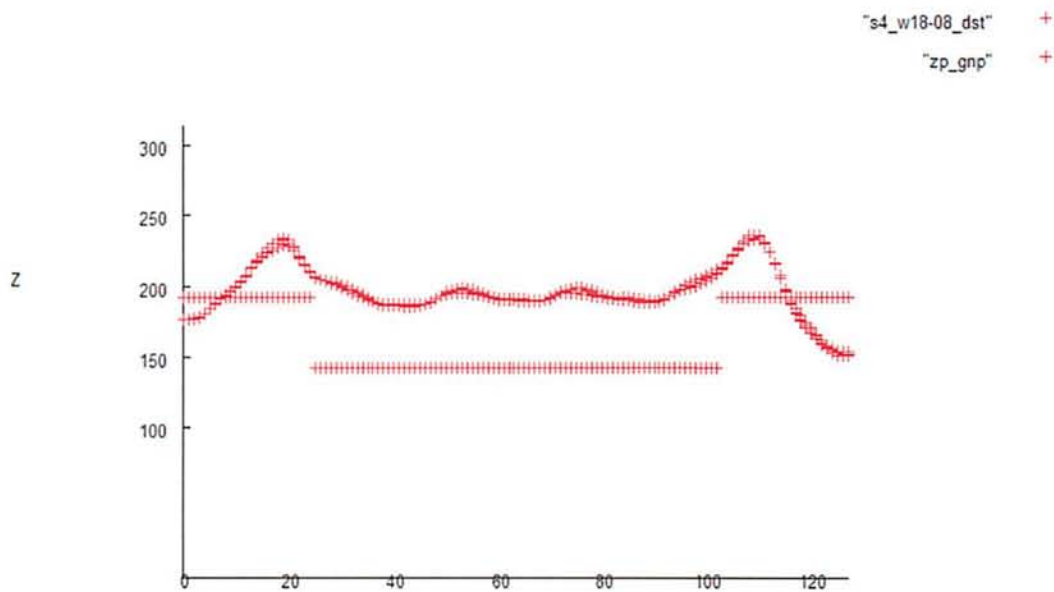


図 5.23 断面図による正解奥行との比較マップ



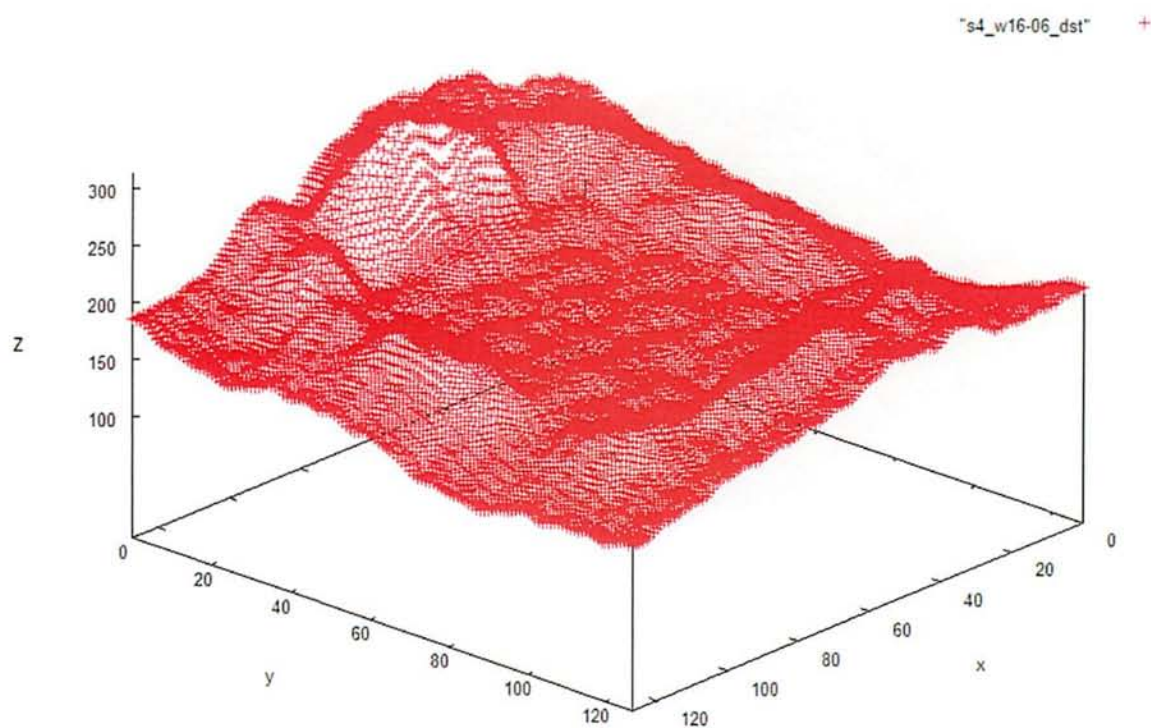


図 5.24 画像ペア使用率 68%時の奥行マップ

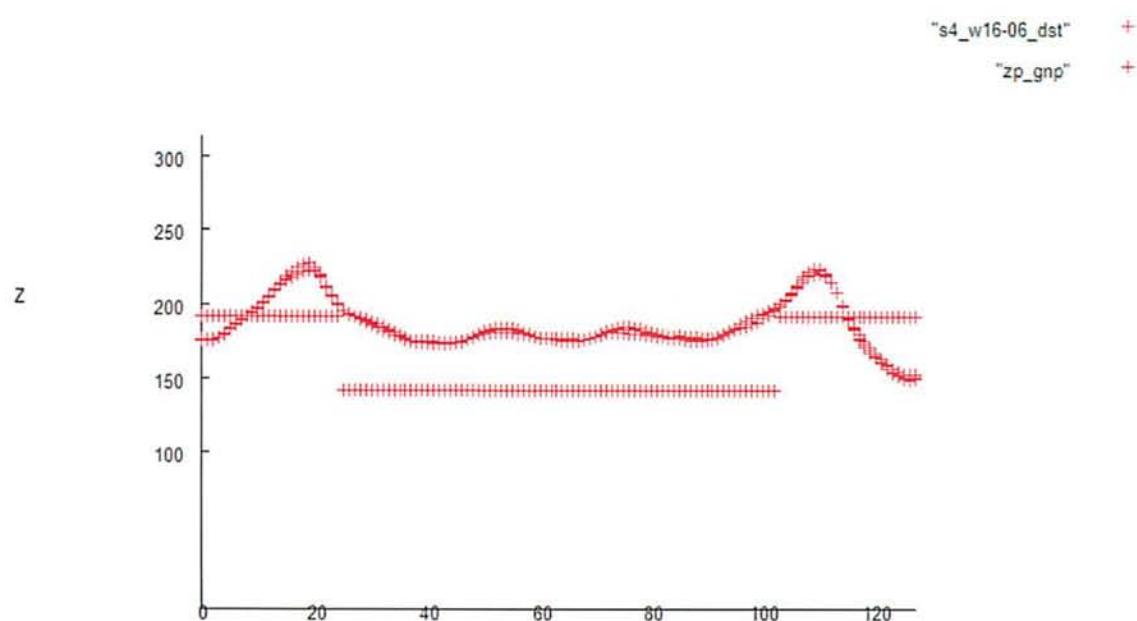


図 5.25 断面図による正解奥行との比較マップ

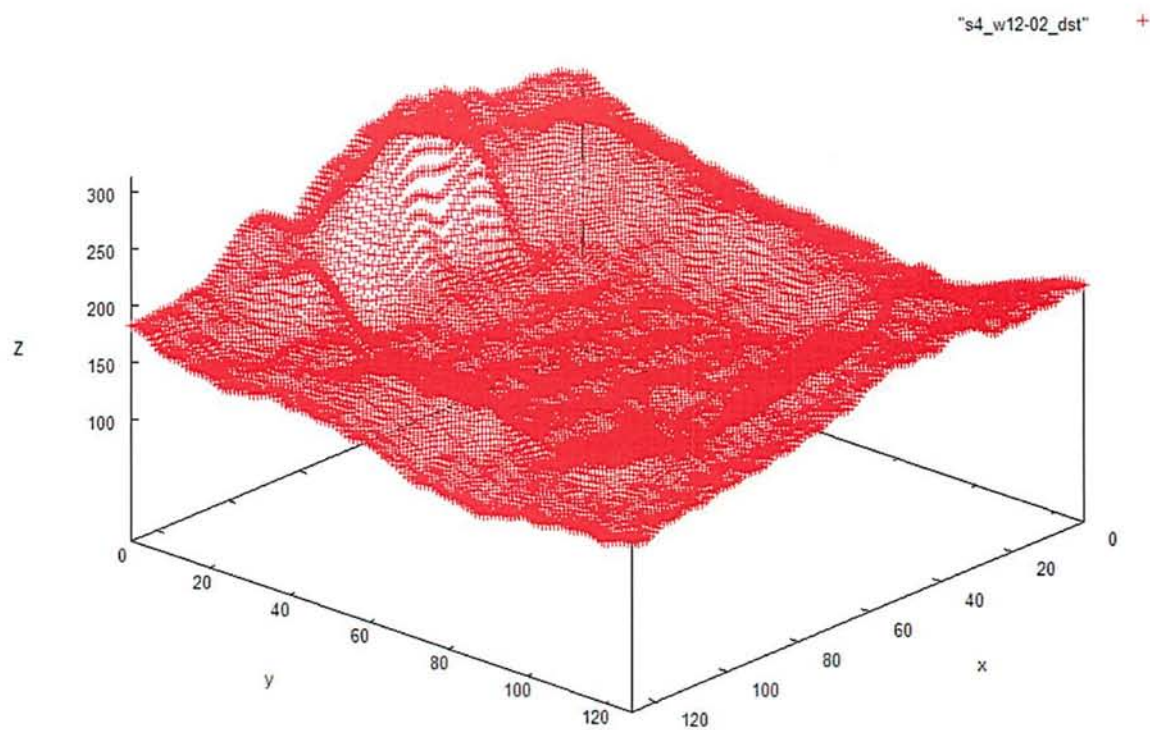


図 5.26 画像ペア使用率 49%時の奥行マップ

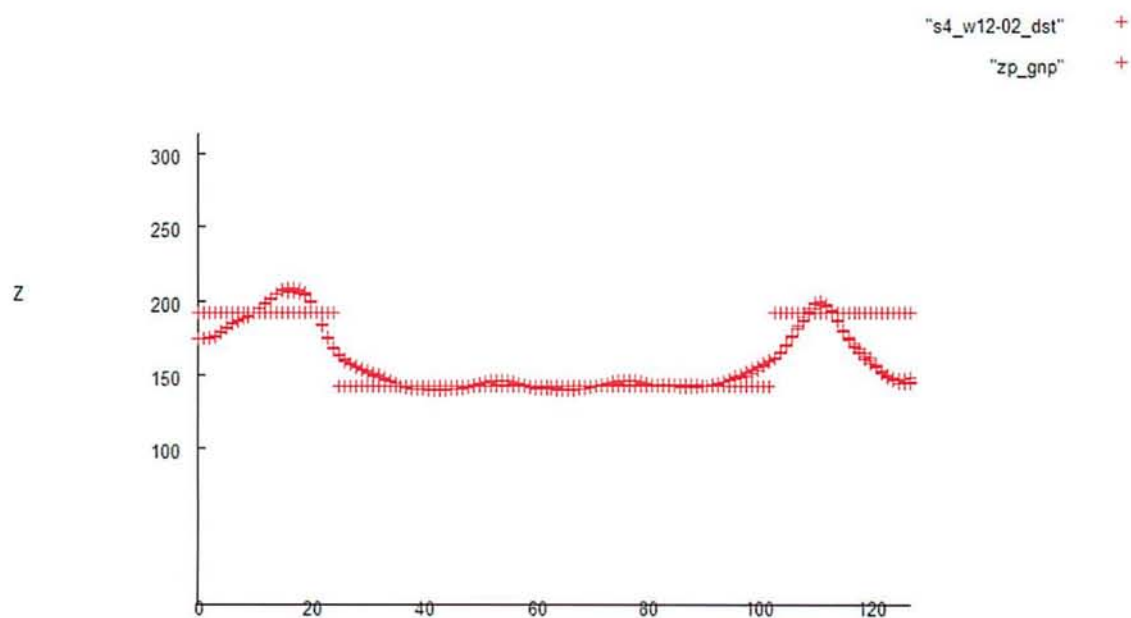


図 5.27 断面図による正解奥行との比較マップ

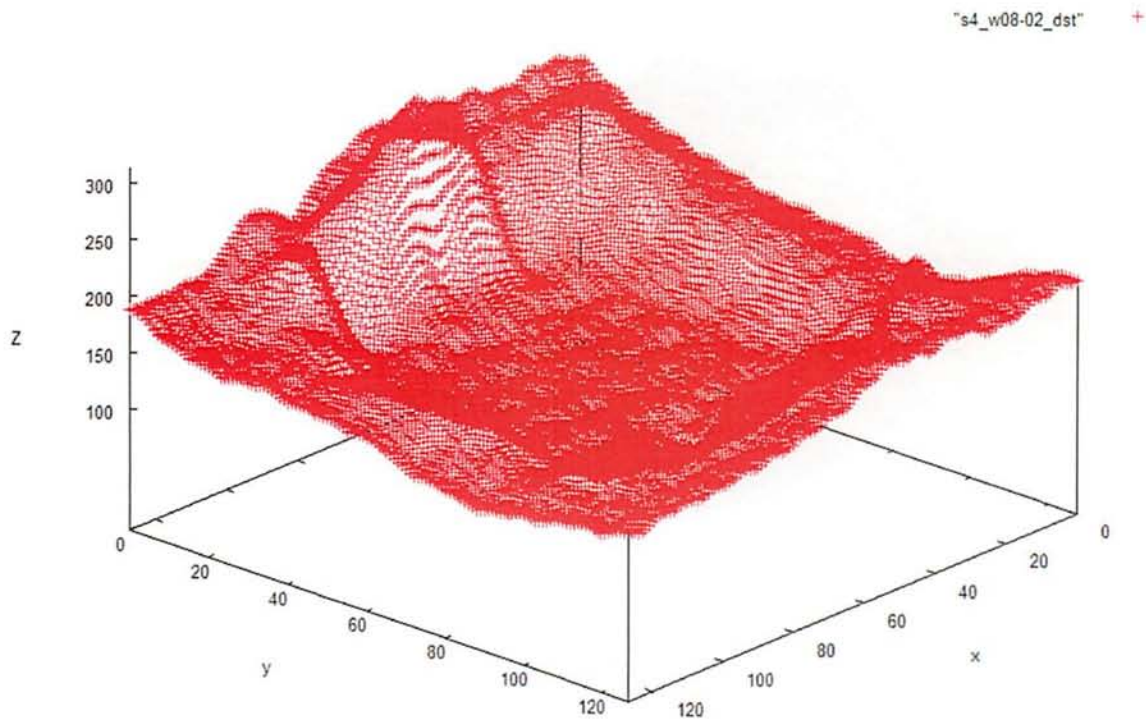


図 5.28 画像ペア使用率 31%時の奥行マップ

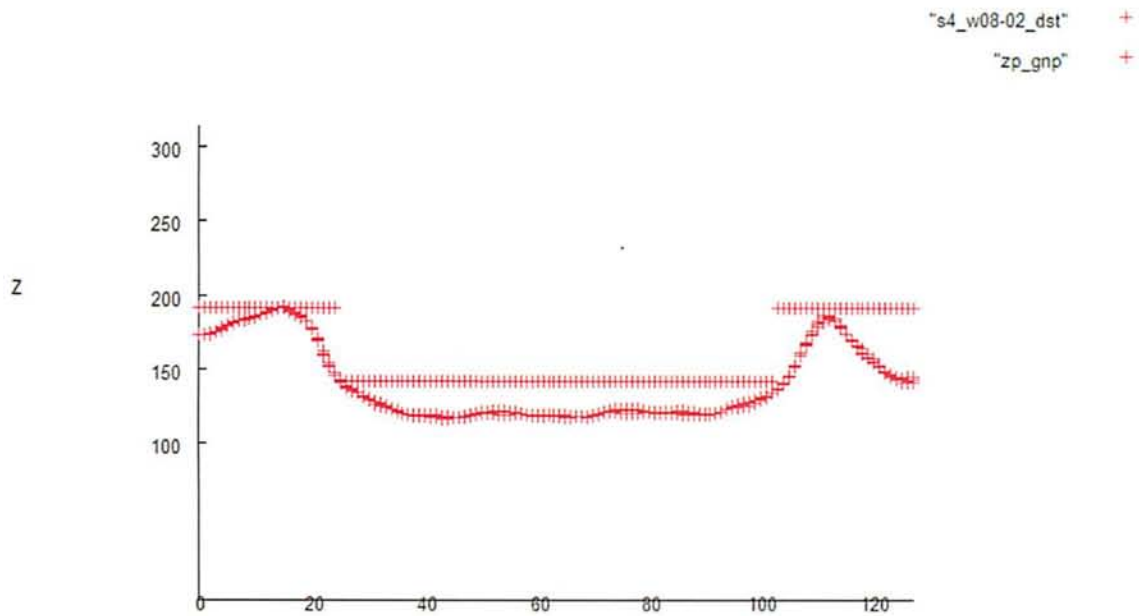


図 5.29 断面図による正解奥行との比較マップ

### 5.1.5 $\sigma_r^2=9.39 \times 10^{-5}$ での復元結果と正解の比較

ここでは、 $\sigma_r^2=9.39 \times 10^{-5}$  の時、既提案法の奥行と、画像ペア使用率を変化させた時の奥行と、正解奥行と比較した時の断面図を示す。

またこの条件での既提案法の奥行は、悪い成分を持つ画像ペアの影響を受け、発散してしまうため本項には記載していない。

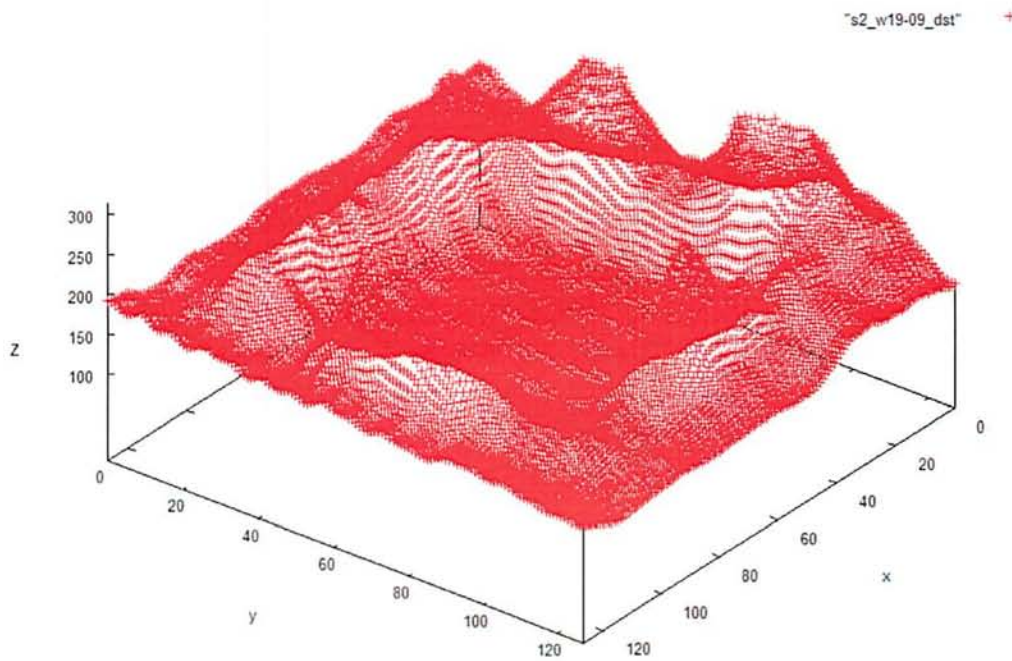


図 5.30 画像ペア使用率 80%時の奥行マップ

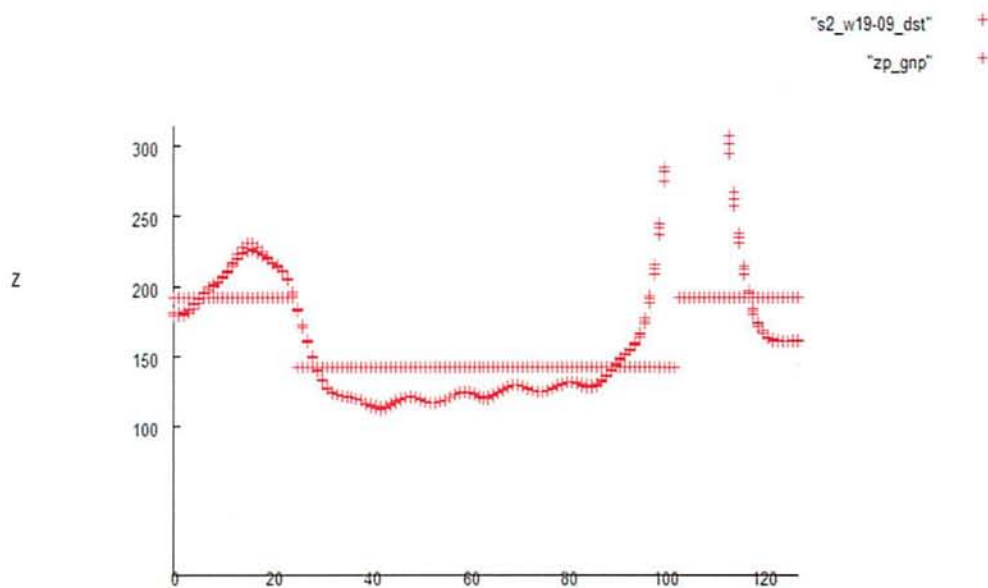


図 5.31 断面図による正解奥行との比較マップ



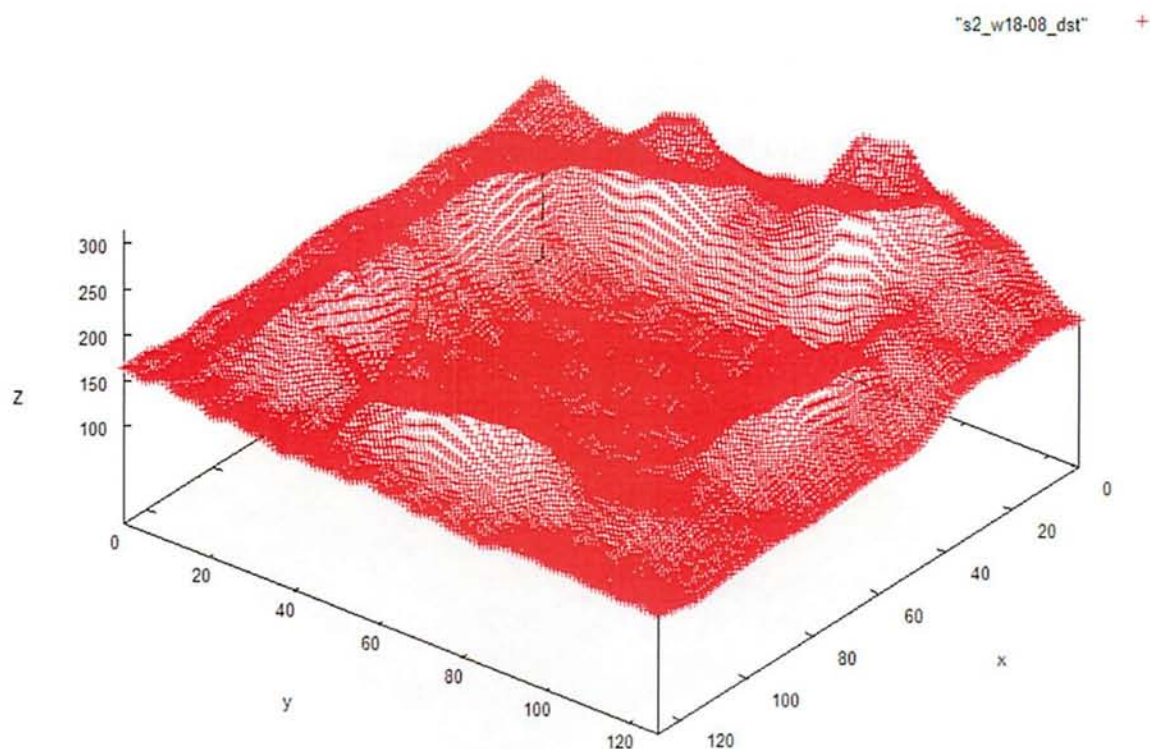


図 5.32 画像ペア使用率 74%時の奥行マップ

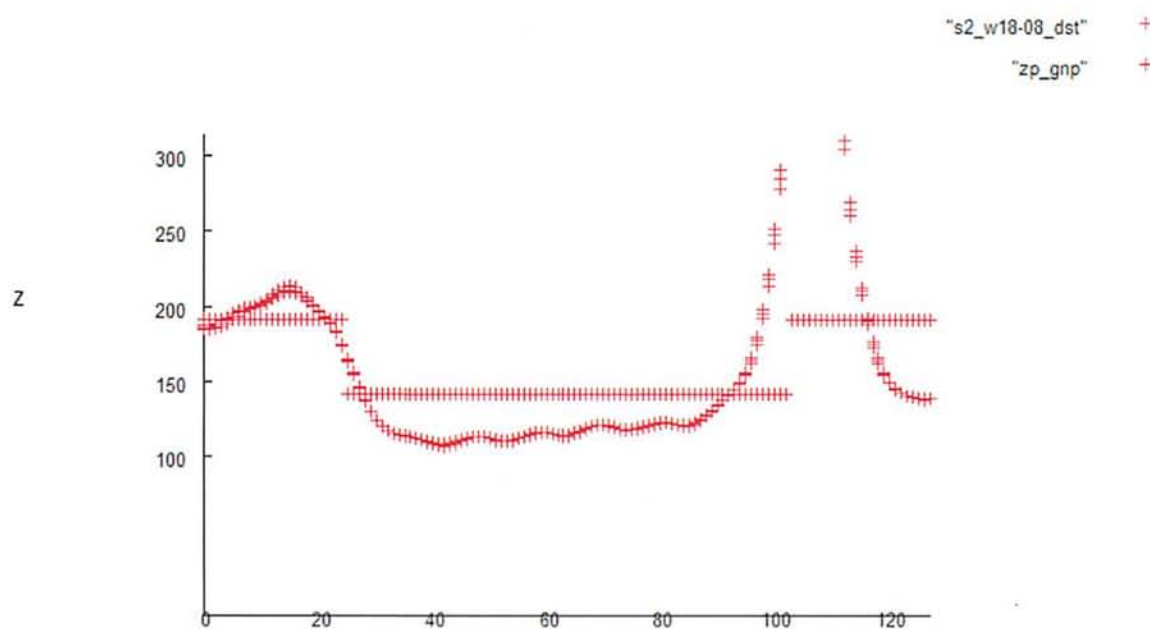


図 5.33 断面図による正解奥行との比較マップ

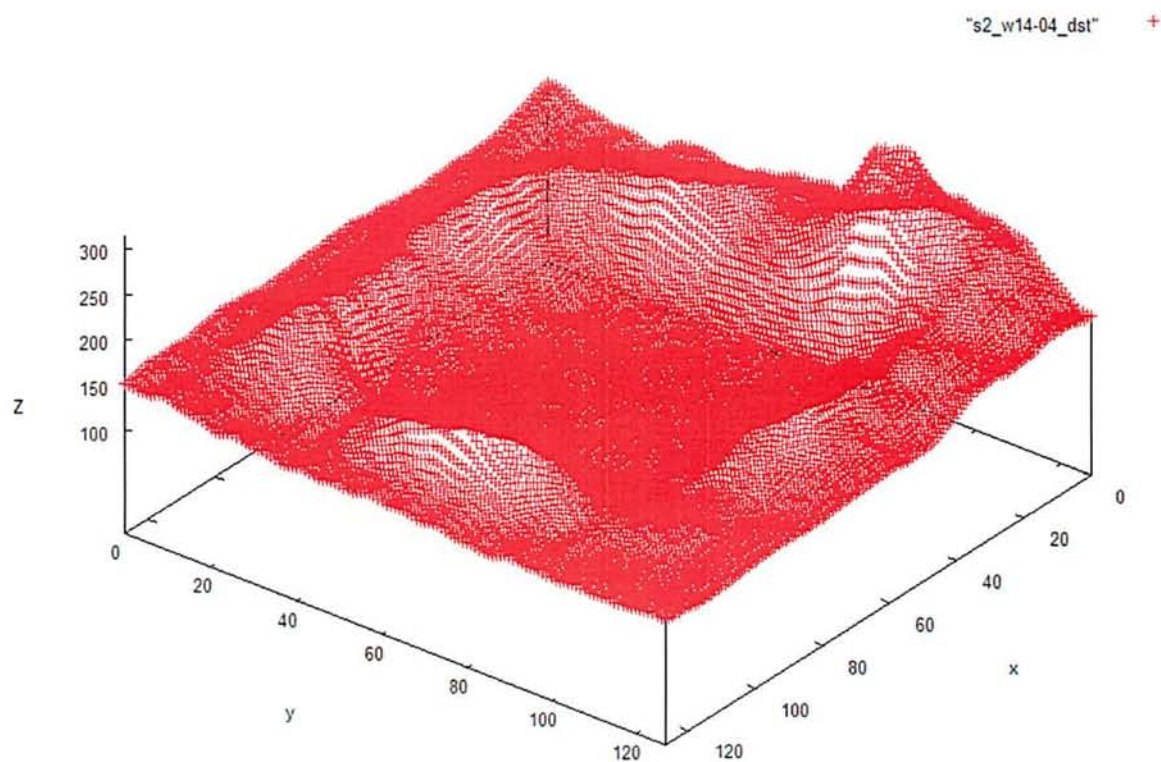


図 5.34 画像ペア使用率 56%時の奥行マップ

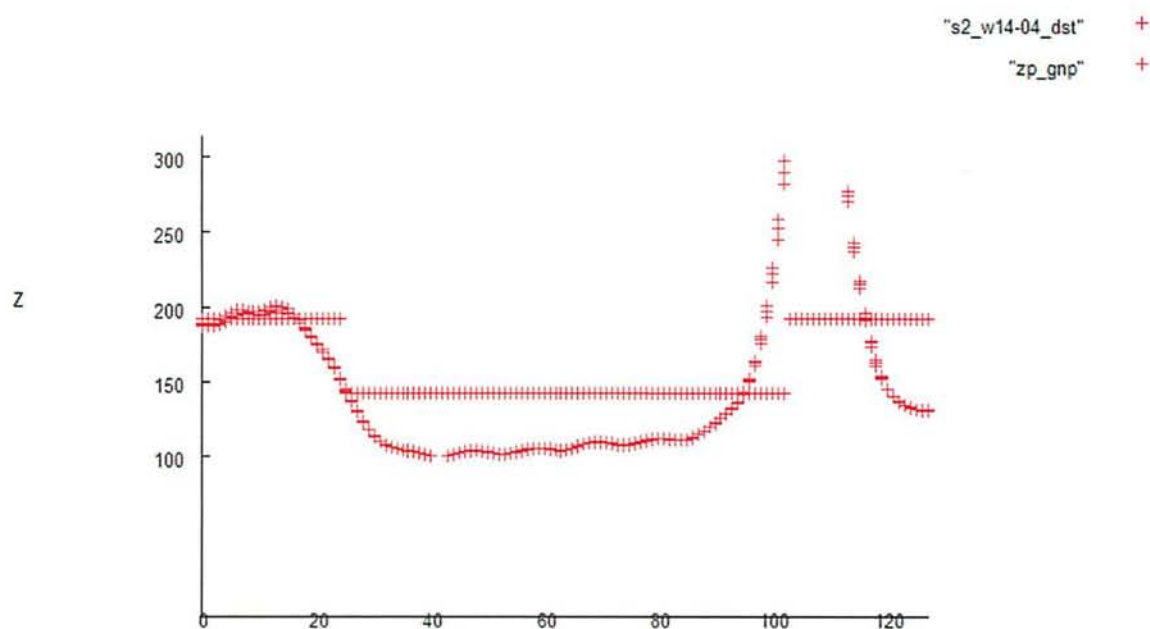


図 5.35 断面図による正解奥行との比較マップ

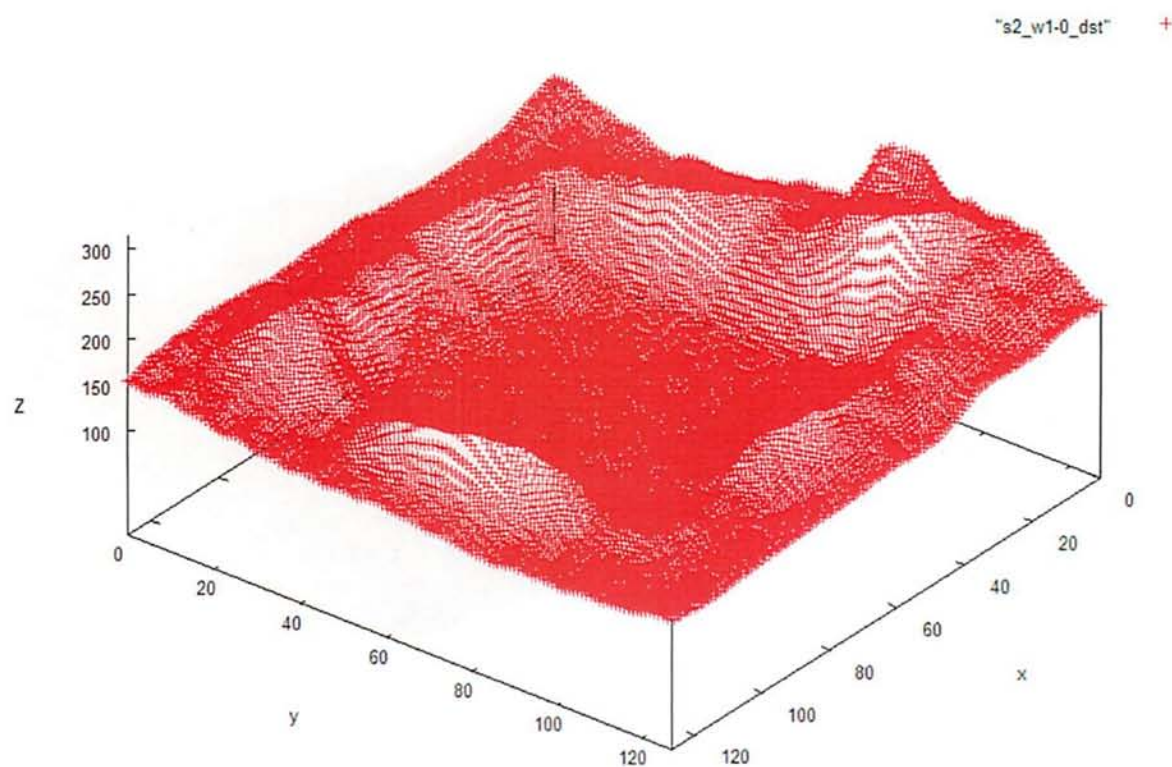


図 5.36 画像ペア使用率 37%時の奥行マップ

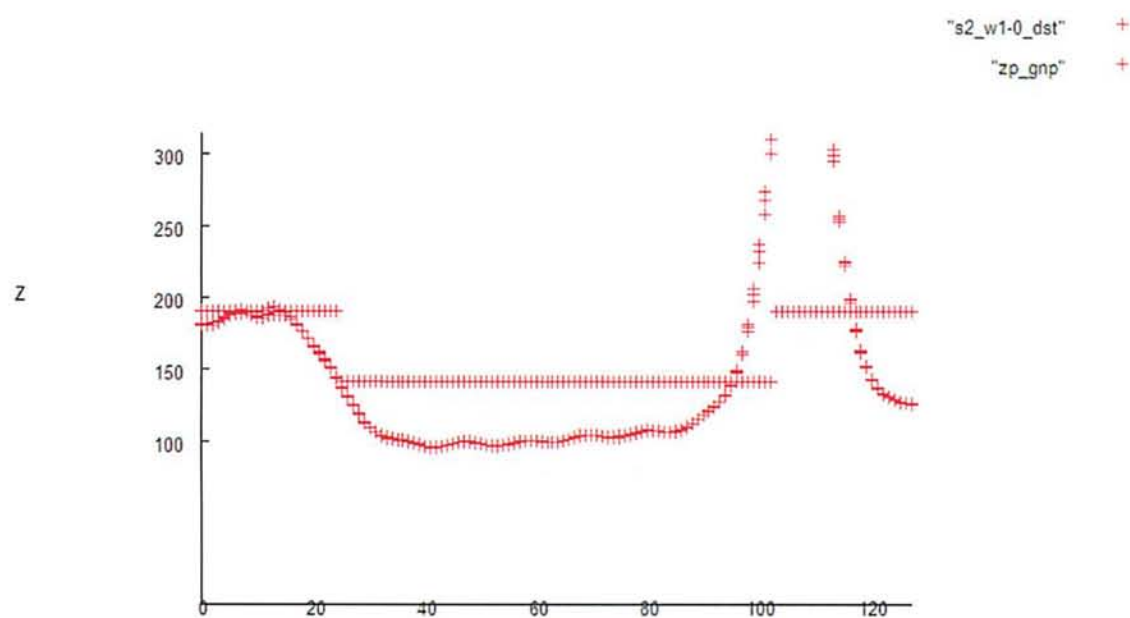


図 5.37 断面図による正解奥行との比較マップ

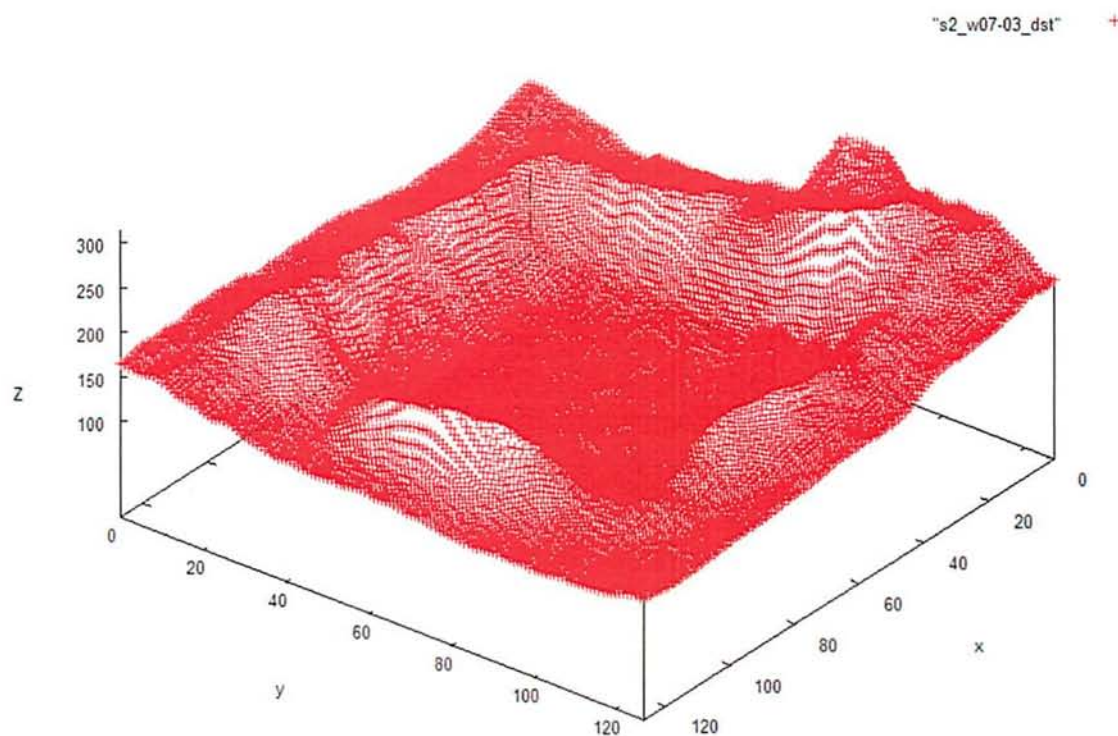


図 5.38 画像ペア使用率 25%時の奥行マップ

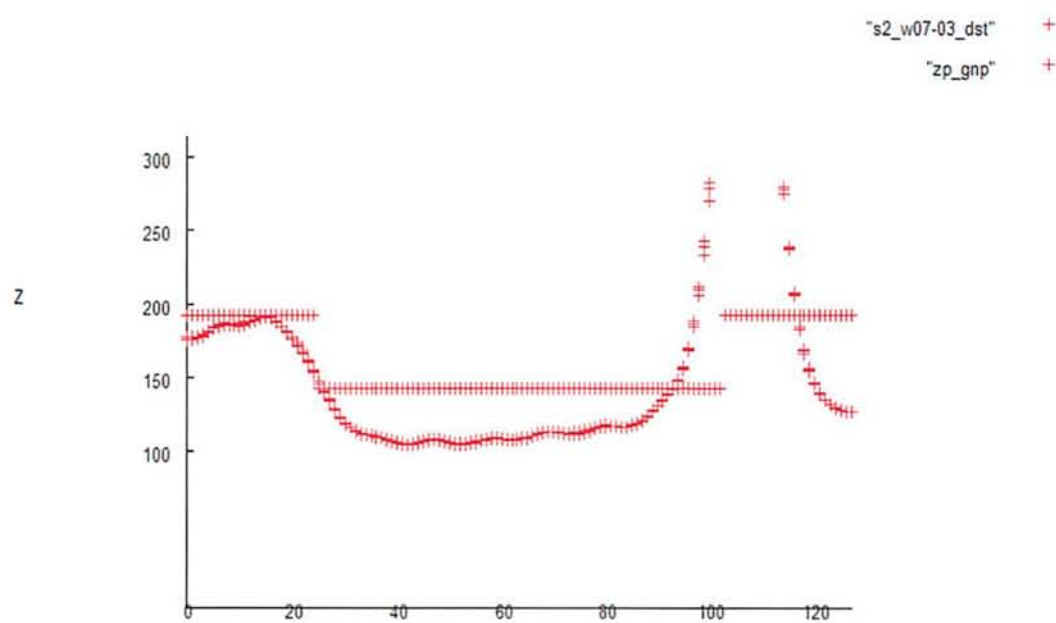


図 5.39 断面図による正解奥行との比較マップ



### 5.1.6 $\sigma_d$ を変化させた場合の復元結果と正解の比較

ここでは、図 5.24 の奥行に対し  $\sigma_r^2=2.33*10^{-5}$  の時、滑らかさの拘束度合  $\sigma_d$  を変化させた場合の奥行と、正解奥行と比較した時の断面図を示す。

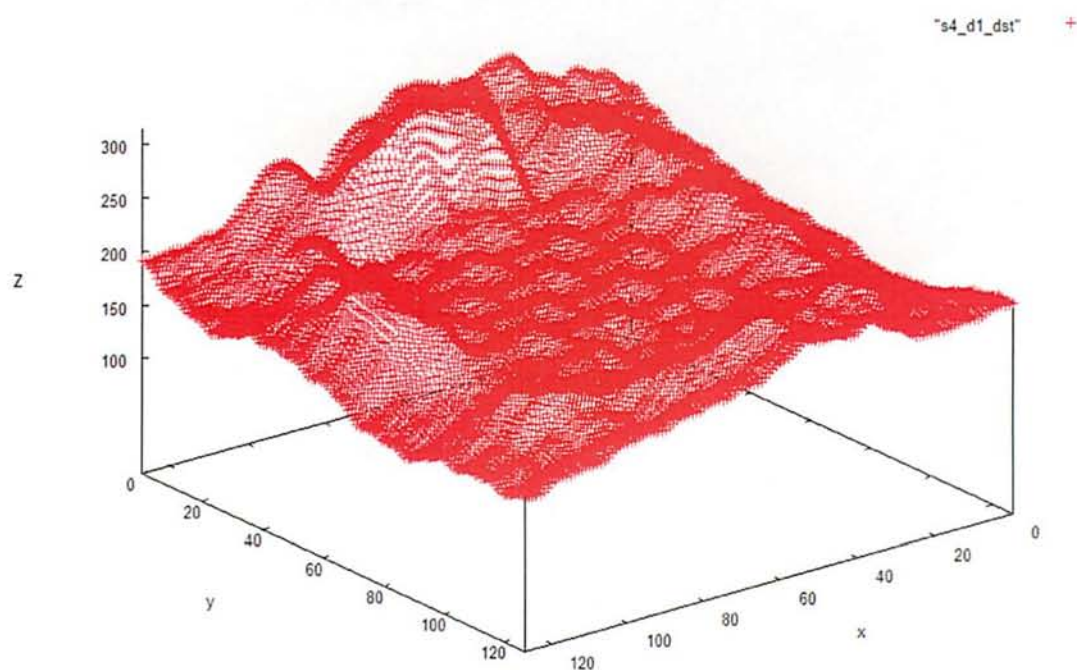


図 5.40  $\sigma_d = 1.0*10^{-1}$  時の奥行マップ

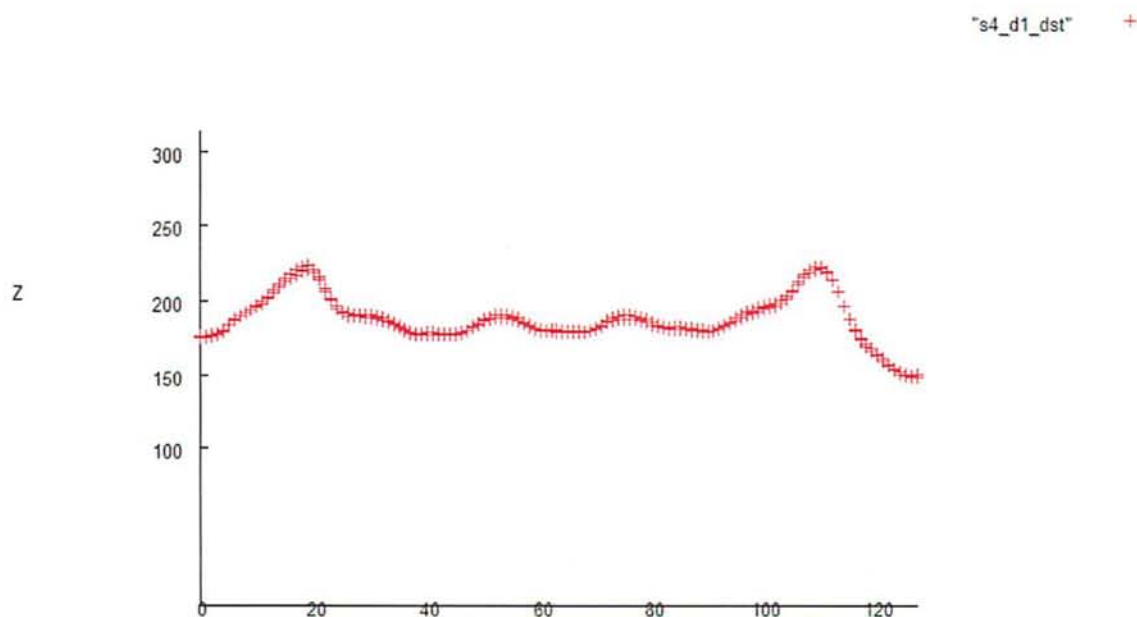


図 5.41 断面図による奥行マップ

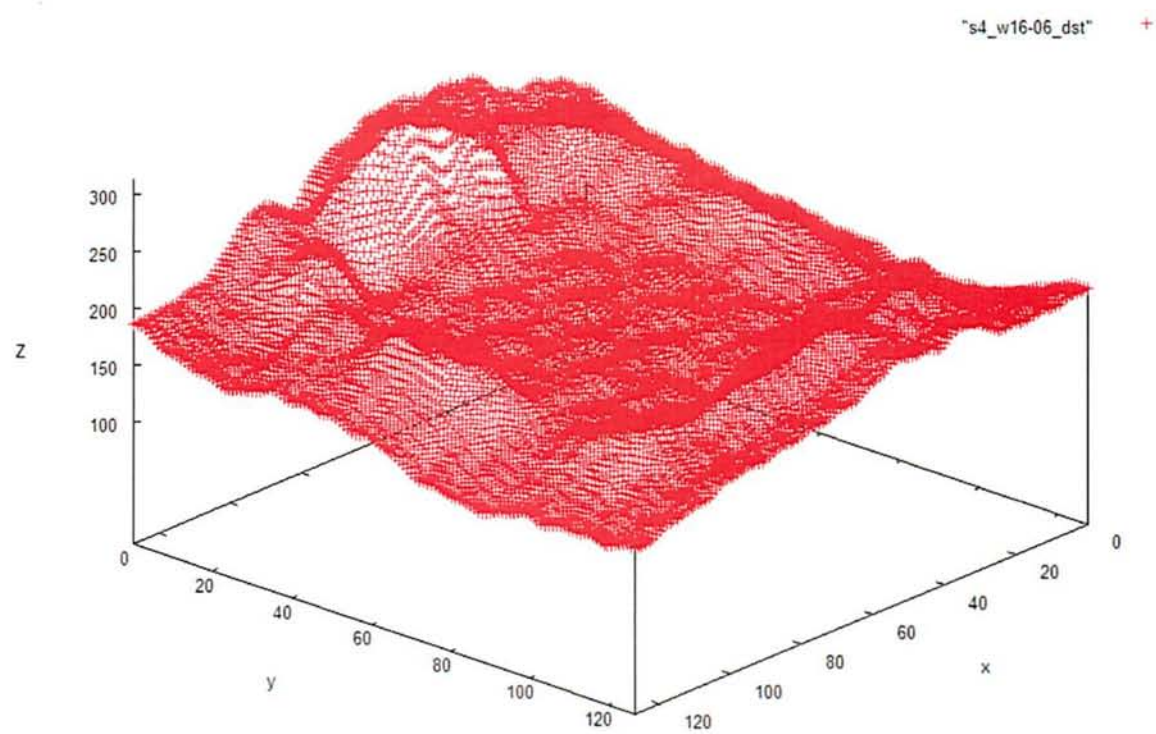


図 5.42  $\sigma_d = 1.0 \cdot 10^{-4}$  時の奥行マップ

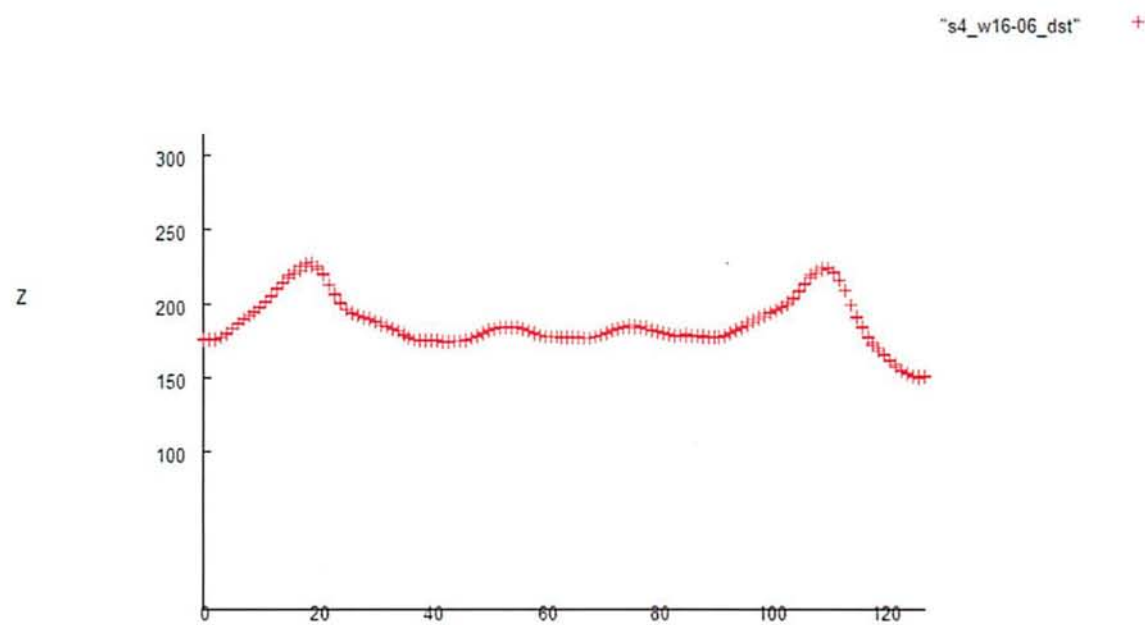


図 5.43 断面図による奥行マップ

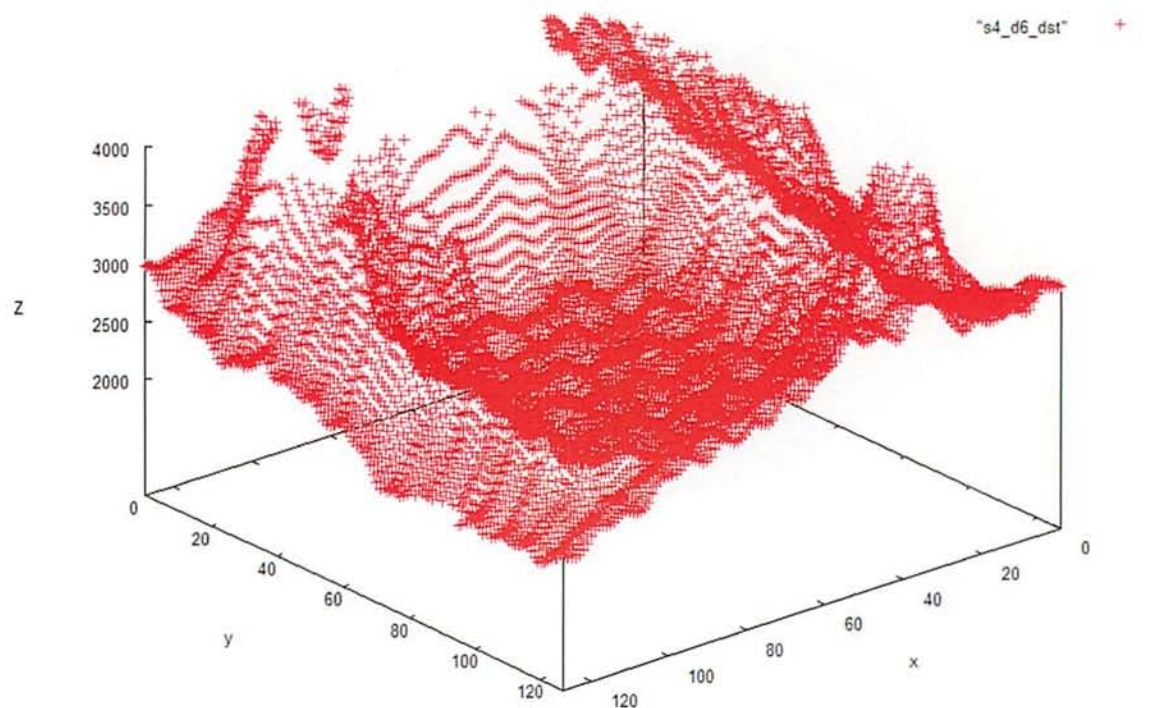
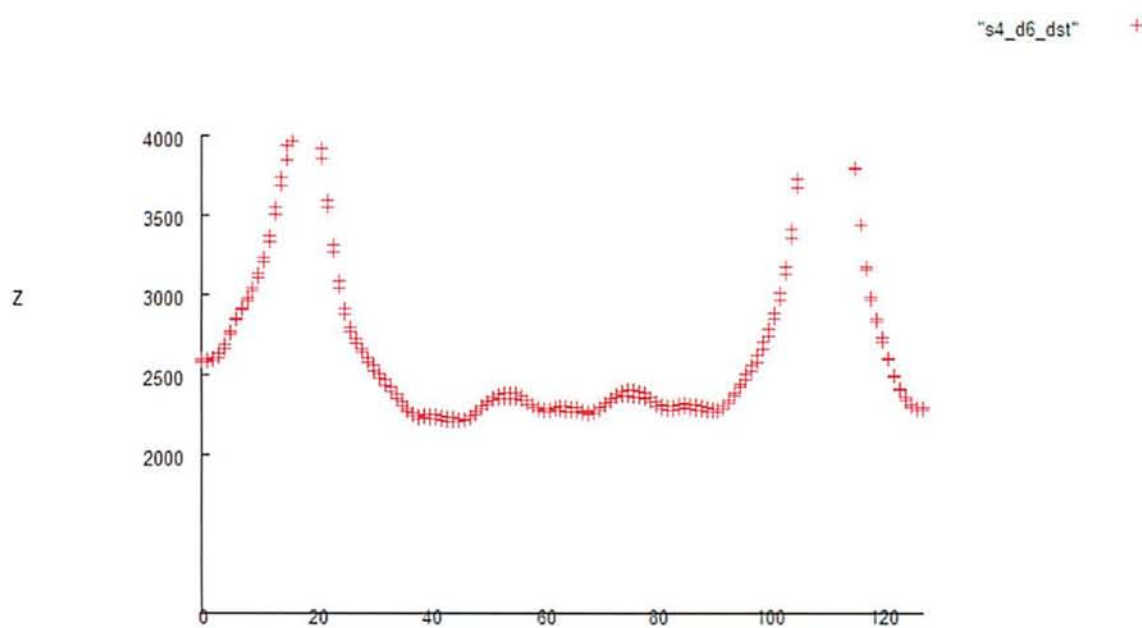
図 5.44  $\sigma_d = 1.0 \cdot 10^{-6}$  時の奥行マップ

図 5.45 断面図による奥行マップ

### 5.1.7 RMSE による奥行精度評価

ここでは、 $\sigma_d^2 = 1.0 \cdot 10^{-4}$  の時、正解奥行に対する復元奥行の RMSE（最小 2 乗誤差）を画像ペア使用率ごとにグラフと示す。グラフの種類は、異なる 3 つの  $\sigma_r^2$  から復元した奥行の計 3 通りとし、また、画像ペア使用率 100% の部分については、既提案法の数値として見る。

なお、 $\sigma_r^2 = 9.39 \cdot 10^{-5}$  時の既提案法の奥行は、5.1.6 節で説明した通り発散してしまうため、本研究での RMSE グラフには記載していない。

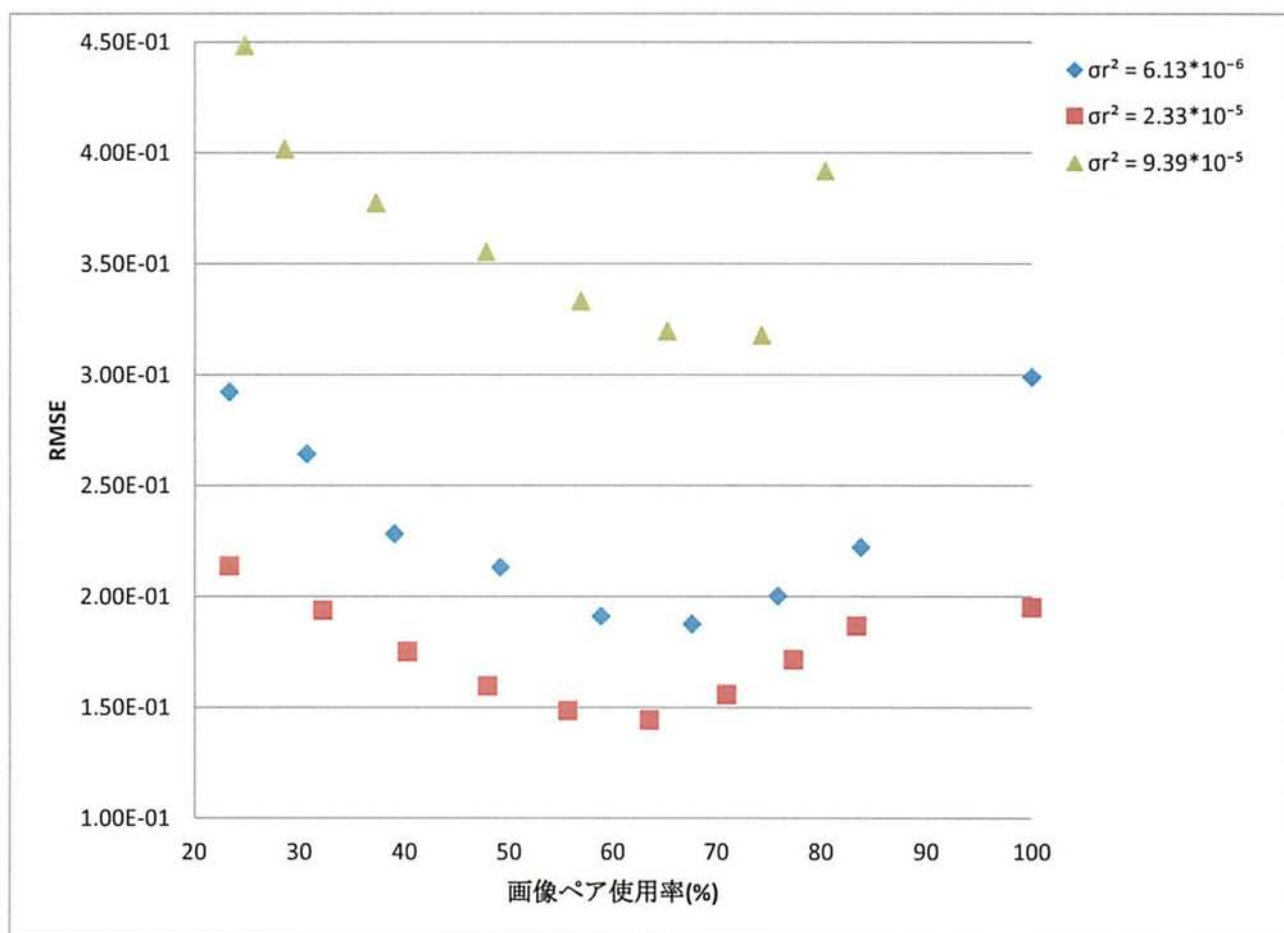


図 5.46 正解奥行に対する実験による復元奥行の RMSE



### 5.1.8 画像特徴点ごとのカメラ回転( $r_x, r_y$ )の使用傾向

ここでは、 $\sigma_r^2=2.33*10^{-5}$ の時、カメラ回転 $\mathbf{r}$ の x 軸 y 軸方向の動きに相当する( $r_x, r_y$ )に着目し、その画像ペアごとの分布をラジアン単位でグラフに示す。

なおグラフの特徴点については 3 点設けており、それぞれ Image center が図 5.1 実験画像における画像中心の画素を参照した場合、Corner が同じく図 5.1 での中央立体部分の左下角付近の画素を参照した場合、Middle がその両者の中間点の画素を参照した場合となっている。また、All pattern used は上記全てで利用された( $r_x, r_y$ )を表しており、逆に No used はいずれでも利用されなかった( $r_x, r_y$ )を表している。

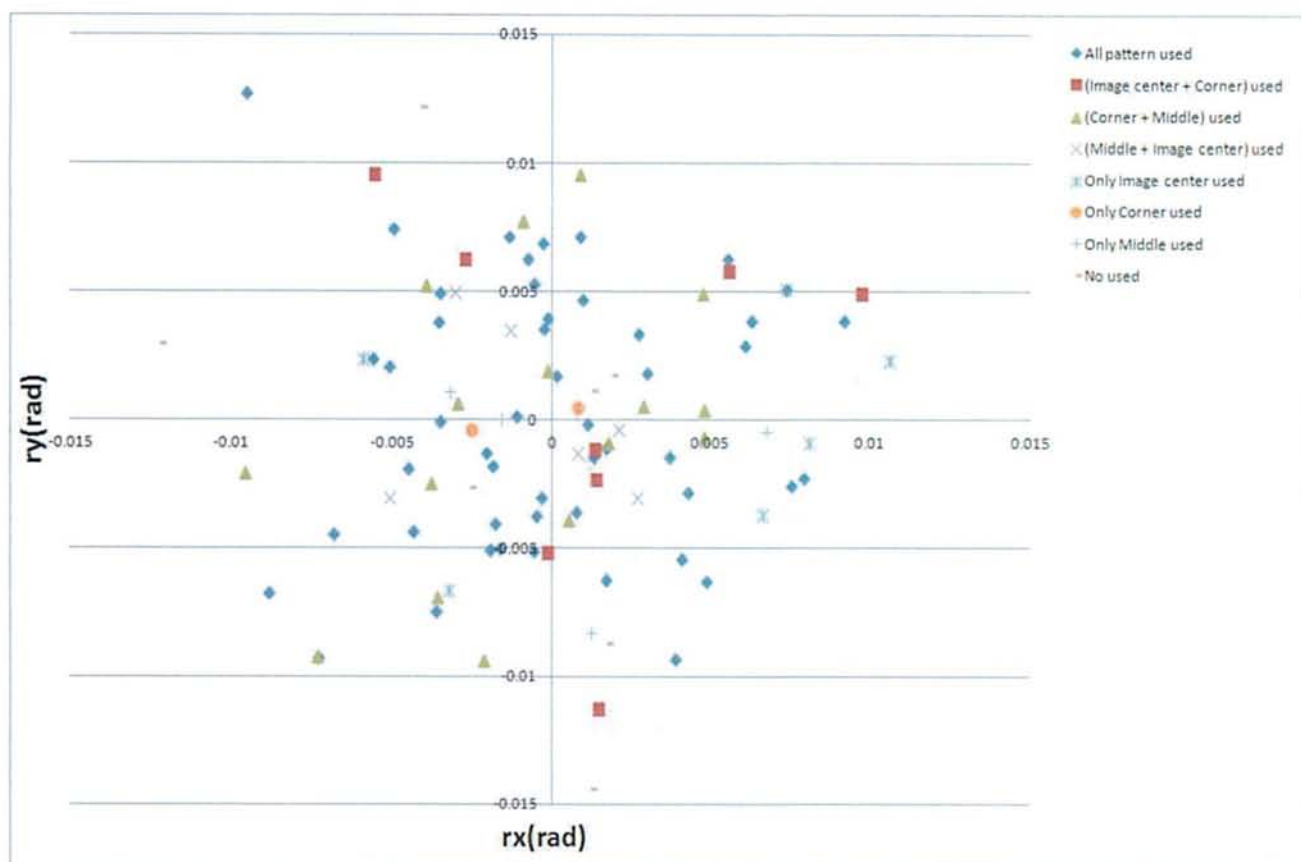


図 5.47 カメラ回転( $r_x, r_y$ )の分布と特徴点ごとの使用傾向

## 5.2 数値評価に対する考察

$\sigma_d^2 = 1.0 \times 10^{-4}$  の時、画像ペア使用率を変化させると  $\sigma_r^2 = 2.33 \times 10^{-5}, 6.13 \times 10^{-6}$  の場合は使用率が下がるに連れ中央立体部分にあたる奥行が徐々に正解に近付き、 $\sigma_r^2 = 2.33 \times 10^{-5}$  では 49%、 $\sigma_r^2 = 6.13 \times 10^{-6}$  では 32%になると正解奥行に一番近くなることが復元奥行の断面図から確認できる。これは、エイリアス現象や近似誤差の多い画素成分が本来与えるべき復元値を邪魔しており、画像使用率を変えることによってそのような成分を多く除いたことが理由の一つとなる。逆にそれ以上使用率を下げても奥行の変化はわずかであり、加えて必要以上に成分を除いてしまっているため、本来利用可能であった情報量が少なくなり、誤った奥行を復元してしまう可能性も有り得る。その証拠に図 5.46 のグラフを見ると、ある一定の画像ペア使用率までは RMSE が下がり続け、それ以降は上昇していることがわかる。その中で、49%、32%というパーセンテージはそれぞれ  $\sigma_r^2 = 2.33 \times 10^{-5}, 6.13 \times 10^{-6}$  での一番 RMSE が低くなる画像ペア使用率を示しており、復元結果の妥当性を確認できる。

なお  $\sigma_r^2 = 9.39 \times 10^{-5}$  の場合については、復元奥行と RMSE グラフ両方を見ても正解から離れてしまっているため、精度が一番高くなる画像ペア使用率は検討できないが、他の  $\sigma_r^2$  と同様に画像ペア使用率を 50%付近まで下げていくと RMSE が低くなることが確認できる。

上述をもとに  $\sigma_r^2$  の大きさごとの復元奥行の比較を行う。

比較対象の画像ペア使用率は上記で示した、復元精度が最高になるものを基準とする。 $\sigma_r^2 = 2.33 \times 10^{-5}$  の場合、約半分画像ペアを除けば復元奥行が高精度になるのに対し、 $\sigma_r^2 = 6.13 \times 10^{-6}$  では、約 70%もの画像ペアを取り除かなければ高精度にならないことがわかる。これは 3.2.2 節に基づいて決定した最適な動きである、空間波長の  $\lambda/4$  分に相当するカメラ回転  $\sigma_r^2 = 2.33 \times 10^{-5}$  で撮影した画像を利用していることが理由の 1 つといえる。もう 1 つの理由として、その最適なカメラ回転の半分  $\sigma_r^2 = 6.13 \times 10^{-6}$  では当然として動きが絶対的に小さくなるので、結果として 1 つ 1 つの成分に含まれている差分の情報量も低下してしまう。このことにより、誤差の大きい画像ペアに、多くの使用すべき画像ペアが埋もれてしまうため、必要以上に画像ペアを取り除かなければならなくなる。

また、 $\sigma_r^2 = 9.39 \times 10^{-5}$  の場合、最適なカメラ回転の 2 倍となるので、 $\lambda/2$  分の動きが平均となる。したがって、エイリアス現象が発生する条件である、 $\lambda/2$  以上の動きを持つ画像ペアが多く存在することになるため、これも画像ペア使用率を変更することになる。

しかし上記で説明した通り、使用率を変更しても RMSE は低くならず、復元奥行も正解に近付かないことから、カメラ回転が  $\lambda/2$  以下という条件は実験前の設定で絶対に順守すべき事項であることがわかる。

次に、波の拘束度合  $\sigma_d$  を変化した時の奥行の形状の変化について説明する。まず  $\sigma_d = 1.0 \times 10^{-1}$  とした場合の奥行の断面図と、元々利用していた  $\sigma_d = 1.0 \times 10^{-4}$  時の奥行の断面図を見比べてみると、 $\sigma_d^2 = 1.0 \times 10^{-1}$  での奥行の場合、若干波が粗くなっていることがわかる。逆に  $\sigma_d^2 = 1.0 \times 10^{-6}$  の場合であると、極端に値が小さくなっているため、波が滑らかになるだけでなく、形状そのものやスケール自体を変えてしまう危険性が伴う。このように  $\sigma_d$  を少しずつ小さくしていくと奥行で発生している波が徐々に低減されていると推測できる。以上から、 $\sigma_d$  の扱い方としては、復元された奥行の波を一度確認し、必要に応じて数値を変更する方法が適当といえる。しかし注意すべきように、極端に数値を上下させると形状に影響を与えてしまうことを考慮しなければならない。

加えて本研究では、実験で使用されたカメラ回転  $\mathbf{r}$  についての向きと大きさを確認することで、画像上で選択した複数の特徴点それぞれで利用されている回転がどのような傾向を持つのかを調査した。図 5.47 を見ると、基本的にはどの特徴点でも全体的に多くの回転が利用されていることがわかる。しかし、使われていないもしくは 1 つの特徴点のみで扱われている回転に着目すると、比較的大きな回転もしくは 0 に近い小さな回転がほとんどである。これは、本考察で上述したエイリアス現象や情報量の低下の影響を受け、利用が限定的になっていると考えられる。今回の確認で明らかになった部分は少ないが、今後の研究で発展させ、特徴点ごとに最適なカメラ回転を想定できれば、奥行の復元精度向上に生かせると期待できる。

最後に問題点が二つ挙げられる。まず一つ目は、復元された奥行の中で形状の一部分が突出し、正解から著しく外れている部分が存在することである。加えて、この部分に引っ張られて多くの復元奥行が斜めに傾いてしまっている。その一方で、この部分に関しては共通点が存在する。図 5.1 を見ると、奥板の左下付近に照明による中央立体部分の影ができていることがわかる。その部分に相当する奥行が、以上の問題点として挙げられているように突出している部分であることがわかった。つまり輝度値に影響を与えないように、奥行の視軸方向に同等の強さを持つ光を当て続け、影ができない状態を保ちつつ撮影を行えばよいことになる。

本研究ではこれを理解し、光量を調節しながら撮影を行った。しかし実験結果を見ると、正面から均一に光を当てるための工夫やそのために必要な器具が十分ではなかったと言わざるを得なく、今後に向けて再考する必要がある。



二つ目については、本実験での閾値の決め方が画像全画素に対する平均を基準として決めており、各画素ごとの特徴を考慮していないことである。このため、平均から大きく外れた画素については閾値が適当でない可能性が高く、一意に全ての画像ペアが正確に選択されているとは考えにくい。したがって解決策としては、個々の画素についてそれぞれの勾配や輝度値から特徴を見極め、その特徴に沿った閾値を個別に設定していかなければならない。しかし、扱っている画像のサイズは  $256 \times 256$  のため、一つ一つ閾値を設定するのは膨大な時間が必要となってしまう。なので、似た特徴を持つ画素ごとにグループに分けを行い、そのグループ単位で最適な閾値を設定することで効率の良い実験に繋がると考えられる。また閾値の基準については  $\sigma_r^2$  の大きさごとに別々の平均が算出されるので、図 5.5、図 5.46 については、同じ基準での画像ペア使用率のグラフとなっていない。なので閾値の基準を統一し、実験結果をより比較しやすくする必要がある。



## 第6章 結論

固視微動という運動の特にトレモアに相当するカメラ運動に基づく形状復元において近似誤差を見積もり、それに基づく勾配方程式の選択的利用法を提案した。加えて、固視微動を模擬するカメラ撮影システムを構成し、提案している奥行復元手法の有効性を実験的に確認した。既提案手法では、輝度値の空間的な波長に対して画像変位が大きくなる場合、エイリアス現象が起きる。それに対し、提案手法により誤差低減を確立し、奥行復元精度の向上を確認した。

また、実画像で奥行復元することは難しく、実空間での正確な測定を求められる。今後の課題として、より正確に測定できる用具を購入し、照明の光による輝度値の影響を考慮した実験を行うことで、奥行復元の精度を上げていく必要がある。

## 謝辞

本研究を遂行するにあたり，御指導御鞭撻を頂いた田川憲男教授，大久保寛准教授に心より深く感謝いたします。また，日頃よりお世話になっている本研究室の学生の皆様に感謝いたします。最後に，システムデザイン研究科情報通信システム学域の教員各位，学生の皆様に心より御礼を申し上げます。

## 参考文献

- [1] N. Lazaros, G. C. Sirakoulis, A. Gasteratos, "Review of stereo vision algorithm: from software to hardware," *Int. J. Optomechatronics*, vol. 5, no. 4, pp. 435-462, 2008.
- [2] S. Martinez-Conde, S. L. Macknik, D. Hubel, "The role of fixational eye movements in visual perception," *Nature Reviews*, vol. 5, pp. 229-240, 2004.
- [3] Martinez-Conde, S., Macknik, S. L., and Hubel, D. (2004). The role of fixational eye movements in visual perception. *Nature Reviews*, 5:229-240.
- [4] A. Bruhn and J. Weickert, "Locas/Kanade meets Horn/Schunck: combining local and global optic flow methods," *Int. J. Comput. Vision*, vol. 61, no. 3, pp. 211-231, 2005.
- [5] B. K. P. Horn and B. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185-203, 1981.
- [6] E. P. Simoncelli, "Bayesian multi-scale differential optical flow," *Handbook of Computer Vision and Applications*, vol. 2, pp. 397-422, 1999.
- [7] A. Bruhn and J. Weickert, "Locas/Kanade meets Horn/Schunck: combining local and global optic flow methods," *Int. J. Comput. Vision*, vol. 61, no. 3, pp. 211-231, 2005.
- [8] Horn, B. K. P. and Schunck, B. (1981). Determining optical flow. *Artif. Intell.*, 17:185-203 Dempster, A.P., Laird, N.M., and Rubin, D.B. (1997).
- [9] N. Tagawa, J. Kawaguchi, S. Naganuma, K. Okubo, "Direct 3-D shape recovery from image sequence based on multi-scale Bayesian network," *Proc. ICPR08, CD-ROM*, 2008.
- [10] ロボットビジョンの基礎 出口光一郎 著 コロナ社 2000年
- [11] P. J. Green, "On use of the Em algorithm for penalized likelihood estimation," *J. Roy. Statist. Soc. B*, vol. 52, pp. 443-452, 1990.

- [12] Tagawa,N. (2010). Depth Perception model based on fixational eye movements using bayesian staticstical inference. In proc.ICPR2010, pages 1662-1665
- [13] Tagawa,N.and Naganuma,S.(2009).pattern Recignition, Structure and motion from image sequence based on multi-scale Bayesian network. In-Tech.Croatia.
- [14] Maximum likelihood from incomplete data. J. Roy. Statist. Soc. B, 39:1-38
- [15] Z., zhang, and Senior Member, IEEE. (2000). A Flexible New Technique for Camera Calibration. Pattern Analysis and Machine Intelligence, IEEE. Transactions on (Volume:22, Isuee:11 ), pages 1330—1334
- [16] J.More,“TheLevenbergMarquardtAlgorithm,Implementation,andTheory,”Numerical Analysis,G.A.Watson,ed.,Springer-Verlag,1977.
- [17] カメラキャリブレーション手法の改良 田島勝一 大分大学工学部技術部  
[www.tech.iwate-u.ac.jp/ihatov2014/houkoku/data/pdf/A-7.pdf](http://www.tech.iwate-u.ac.jp/ihatov2014/houkoku/data/pdf/A-7.pdf)



## 本研究に関する研究業績

### 学会発表

- (1) 固視微動型カメラ運動に基づく微分型奥行き復元アルゴリズムの実装と評価 塚田翔太  
田川憲男, 何 宜欣, 大久保 寛 映像情報メディア情報学会(冬季大会 2014)  
12月 2014年
  
- (2) S. Tsukada, N. Tagawa, Y. Ho and K. Okubo: Accuracy improvement for  
depth from small irregular camera motions and its performance evaluation,  
International Conference on Image Analysis and Recognition (ICIAR2015),  
Paper No. 202, Sep. 2015, Niagara Falls, Canada.
  
- (3) カメラの不規則微小回転を利用する微分型奥行き復元手法の高精度化とその性能評価  
塚田翔太, 田川憲男, 何 宜欣, 大久保 寛  
映像情報メディア情報学会(年次大会 2015) 8月 2015年

## 付録 1

### A. 平行ステレオによる正解奥行値 $Z$ の導出

最初に，必要なパラメータについて，モデルを通して説明を行う．[図 A 参照]  
まずベースライン  $b$  とは，2つのカメラのレンズ中心同士もしくは平行移動前後の  
レンズ中心同士の距離を mm 単位で計ったもので， $b$  (mm) とする．

次に視差  $x_l - x_r$  について説明する．異なる位置にある 2つのカメラが同じ対象物  
体を撮影すると，それぞれの画像平面上に対象物体が写りこむ．画像中心を基準の  
0(pix)とした時，この対象物体の写像位置は  $x$  方向のピクセル単位(カメラ 1 を  $x_l$   
(pix)，カメラ 2 を  $x_r$  (pix)) で計測可能である．そして，その双方の差分  $x_l - x_r$  (pix)  
が視差となる．(位置の計測がスムーズに行えるように，第 5 章図 5.1 の実験画像に  
シール等のマーカーを付けた．)

これらのパラメータを基にすると奥行  $Z$  は以下の式で表現できる．(焦点距離  $f$   
(pix)はキャリブレーションで算出されたものを使用．)

$$Z(\text{mm}) = \frac{f(\text{pix})}{(x_l - x_r)(\text{pix})} b(\text{mm}) \quad (\text{A.1})$$

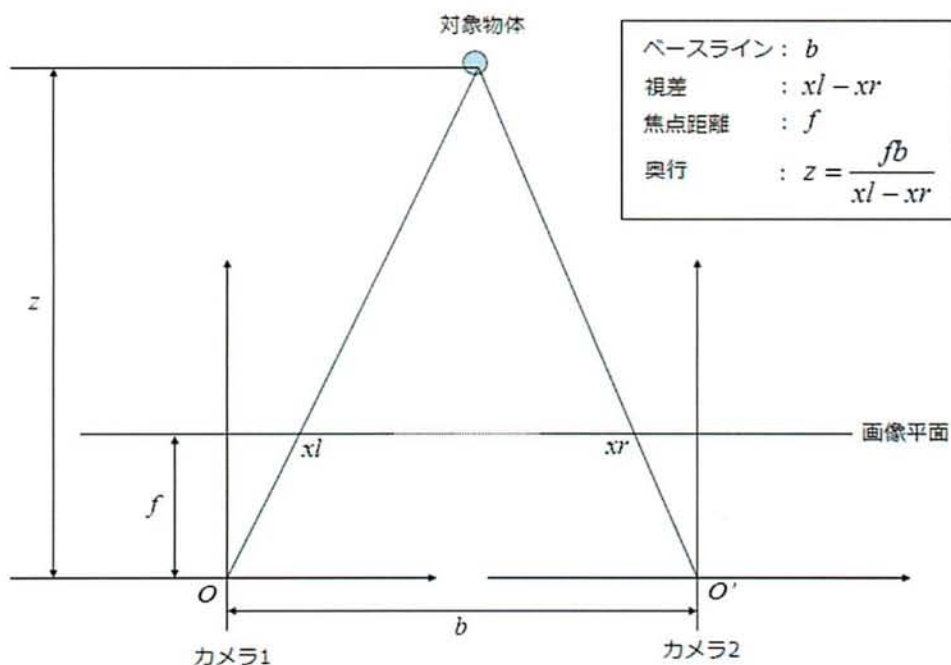


図 A: 平行ステレオモデル

以上から平行ステレオの原理を用いて、実際に奥行  $Z$  を求める。

求める奥行は、図 5.1 のように中央立体部分と奥板の 2 種類あり、両者共通のパラメータとして、ベースライン  $b=30(\text{mm})$ 、焦点距離  $f=1141.4(\text{pix})$  と設定した。

視差については、中央立体部分と奥板それぞれに任意数の特徴点を設け、その板ごとの点から求められた視差の平均を算出値とした。この場合、中央立体部分の視差を  $d_f$ 、奥板の視差を  $d_b$  として計算すると、 $d_f=242(\text{pix})$ 、 $d_b=179(\text{pix})$  と算出され、これを設定値とした。これらのパラメータから、中央立体部分の奥行を  $Z_f$ 、奥板の奥行を  $Z_b$  として、式(A.1)を参考に計算を行った。

$$Z_f = \frac{f}{d_f} b = \frac{1141.4(\text{pix})}{242(\text{pix})} 30(\text{mm}) = 141.5(\text{mm}).$$

$$Z_b = \frac{f}{d_b} b = \frac{1141.4(\text{pix})}{179(\text{pix})} 30(\text{mm}) = 191.3(\text{mm}).$$

このような過程から正解の奥行値  $Z$  を導出することができる。

## B. 平行ステレオによる焦点距離 $Z$ の導出

仮に、対象物体における中央立体部分と奥板の差分  $Z_b - Z_f$  が既知だった場合、式(A.1)から逆算して焦点距離  $f$  を算出することが可能である。

$$Z_b - Z_f = \frac{f}{d_b} b - \frac{f}{d_f} b = bf \left( \frac{1}{d_b} - \frac{1}{d_f} \right) \quad (\text{A.2})$$

$$f = \frac{Z_b - Z_f}{b \left( \frac{1}{d_b} - \frac{1}{d_f} \right)} \quad (\text{A.3})$$

本研究で使用している対象物体の、中央立体部分と奥板の差分は  $51(\text{mm})$  となっている。ベースライン  $b$  と視差  $d_f$ 、 $d_b$  は奥行値  $Z$  を求めたときのものを利用するとして、式(A.3)を参考に焦点距離  $f$  の計算を行った。

$$f = \frac{51}{30 \left( \frac{1}{179} - \frac{1}{242} \right)} = 1168.9(\text{pix}).$$

このように、キャリブレーション以外の焦点距離の導出方法も存在する。

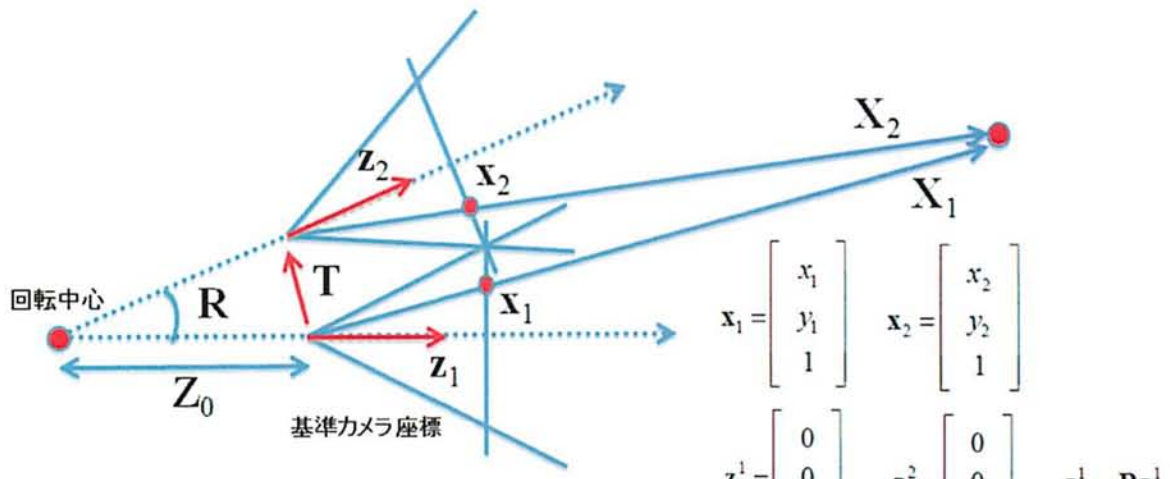
## 付録 2

### 回転中心距離 $Z_0$ の導出

初めに、図 B のモデルで扱われているパラメータについて説明を行う。  
基本的には、カメラの動かし方を平行移動から回転移動に変えたもので、視差の求め方や奥行の関係は変わらないが新たな成分として回転  $R$  が加わり、式がより複雑になる。

第 1 に基本的な情報として、回転は  $x$  軸周りを想定しており、よって視差は画像平面上で  $y$  方向に発生する。さらに焦点距離  $f=1$  とモデルで定義しているので、求めたピクセル単位の視差を  $f$  で割らなければならない。また各ベクトルに付いている下側の数字は、1 が基準のカメラ座標、2 が回転後のカメラ座標の数値をそれぞれ表している。上側の数字は、1 と 2 どちらのカメラ座標の基底から数値を見るかを表している。

第 2 に、各ベクトルの説明を行う。  $x$  は画像上の位置、  $z$  は画像平面の位置、  $X$  は実空間での対象物体の位置、  $Z$  は実空間での対象物体の奥行、  $T$  は回転  $R$  に付随して発生するレンズ中心の平行移動をそれぞれ表している。なお  $Z_2$  に関しては、  $X_2^2$  の計測が複雑であるので、未知としている。



$$x_1 = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}, \quad x_2 = \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$$

$$z_1^1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad z_2^2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad z_2^1 = R z_1^1$$

$$X_2^2 = R^T (X_1^1 - T^1) \Rightarrow R X_2^2 = X_1^1 - T^1$$

$$x_1 = \frac{X_1^1}{Z_1}, \quad x_2 = \frac{X_2^2}{Z_2}, \quad R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix}$$

$$T^1 = Z_0 z_2^1 - Z_0 z_1^1 = Z_0 (R - I) z_1^1 = Z_0 S z_1^1$$

$$Z_2 R x_2 = X_1^1 - Z_0 S z_1^1 \quad Z_0 \text{ と } Z_2 \text{ は未知}$$

図 B: カメラ回転モデル



次に、各ベクトル同士の関係式から、 $Z_0$ の導出を行う。

まず、図 B に記載されている関係式から最終的に  $Z_0$  と  $Z_2$ 、2つの未知数を含んだ方程式が成り立つ。

$$Z_2 \mathbf{R} \mathbf{x}_2 = \mathbf{X}_1^1 - Z_0 \mathbf{S} \mathbf{z}_1^1 \quad (\text{B.1})$$

この式(B.1)に具体的な数値を当てはめていく。

$$Z_2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 \\ Z_1 \end{bmatrix} - Z_0 \begin{bmatrix} 0 & 0 & 0 \\ 0 & \cos \theta - 1 & -\sin \theta \\ 0 & \sin \theta & \cos \theta - 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad (\text{B.2})$$

$$Z_2 \begin{bmatrix} x_2 \\ y_2 \cos \theta - \sin \theta \\ y_2 \sin \theta + \cos \theta \end{bmatrix} = \begin{bmatrix} X_1 \\ Y_1 + Z_0 \sin \theta \\ Z_1 - Z_0 (\cos \theta - 1) \end{bmatrix} \quad (\text{B.3})$$

2行目と3行目に着目。

$$Z_2 (y_2 \cos \theta - \sin \theta) = Y_1 + Z_0 \sin \theta \quad (\text{B.4})$$

$$Z_2 (y_2 \sin \theta + \cos \theta) = Z_1 - Z_0 (\cos \theta - 1) \quad (\text{B.5})$$

式(B.4)を  $Z_2$  のみの項に。

$$Z_2 = \frac{Y_1 + Z_0 \sin \theta}{y_2 \cos \theta - \sin \theta} \quad (\text{B.6})$$

ここで、式(B.5)に式(B.6)を代入する。

$$\frac{Y_1 + Z_0 \sin \theta}{y_2 \cos \theta - \sin \theta} (y_2 \sin \theta + \cos \theta) = Z_1 - Z_0 (\cos \theta - 1) \quad (\text{B.7})$$

$Z_0$  のみの項にする。

$$Z_0 = \frac{Z_1 (y_2 \cos \theta - \sin \theta) - Y_1 (y_2 \sin \theta + \cos \theta)}{\sin \theta (y_2 \sin \theta + \cos \theta) + (\cos \theta - 1) (y_2 \cos \theta - \sin \theta)} \quad (\text{B.8})$$

整理して、

$$Z_0 = Z_1 \frac{(y_2 - y_1) \cos \theta - (y_1 y_2 + 1) \sin \theta}{y_2 (1 - \cos \theta) + \sin \theta} \quad (\text{B.9})$$

以上の導出方法を用いて、回転中心距離  $Z_0$  を求める。

式(B.9)から、必要なパラメータは計4つであり、基準カメラ座標から見た対象物体の奥行  $Z_1(\text{mm})$ 、回転角度  $\theta(\text{rad})$ 、移動前の  $y$  方向の画像座標  $y_1(\text{pix})$ 、移動後の  $y$  方向の画像座標  $y_2(\text{pix})$  である。ここで注意しなければならないのが、焦点距離を  $f=1$  と規格化しているので、画像座標については焦点距離で割る必要がある。(焦点距離は平行ステレオで求めた  $f=1168.9(\text{pix})$  を利用。)

実測値はそれぞれ、 $Z_1 = Z_f = 141.5(\text{mm})$ 、 $\theta = 1(\text{rad})$  と設定してカメラを回転させた。平行移動の時と同様に、カメラの回転移動による視差  $y_2 - y_1$  を13の特徴点から平均を求め26(pix)と算出した。また  $Z_0$  の計算を簡単にするため、移動前の画像座標  $y_1$  を中心の0 (pix)とした。これより  $y_1 = 0(\text{pix})$ ,  $y_2 = 26(\text{pix})$  と決定でき、さらに  $y_1, y_2$  を  $f$  で割って規格化すると、それぞれ  $y_1 = 0$ 、 $y_2 = 0.022243$  となる。

これらのパラメータから式(B.9)を基に回転中心距離  $Z_0$  の計算を行った。

$$Z_0 = 141.5(\text{mm}) \frac{(0.022243 - 0)\cos 1^\circ - (0 * 0.022243 + 1)\sin 1^\circ}{0.022243(1 - \cos 1^\circ) + \sin 1^\circ} = 38.8(\text{mm})$$

このような過程から正確な回転中心距離  $Z_0$  を導出することができる。