

# 第4章 二つのステレオアルゴリズムを統合した奥行き獲得手法

本章では隠れ領域とそうでない領域に対して適切なステレオアルゴリズムを選択的に適用することによってシーン全体の奥行き情報を高精度に獲得可能な手法を提案する。本手法を実現するために、隠れ処理を行わないステレオ法を用いた対応点探索の結果を再利用した「隠れ」領域の抽出手法、および「隠れ」領域の奥行きを精度良く獲得可能なステレオアルゴリズムを提案する。また、このアルゴリズムのビデオレートステレオマシンによる実現例を示す。

## 4.1 隠れ処理を含まない方式と含む方式を統合したステレオ法

多眼ステレオ法を用いて隠れの影響を回避し、物体輪郭における奥行き推定精度を向上させる手法として、佐藤らによってカメラマトリクスステレオ法SEA[8][9]が提案されており、物体輪郭の鮮明な高精細な奥行きマップが得られることが確認されている。SEAにおける隠れ処理の原理は、隠れの生じているカメラのデータを対応点探索処理から外すことであるが、隠れの有無の判定は対応点探索と表裏一体の問題であるため、SEAでは隠れの無い場合と隠れの有る場合について複数の仮説を立てて、それらを並行して調べることによって、最も確からしいものを選ぶ方式を採用している。このため、対応点探索処理に必要な計算量は、隠れ処理を行わない単純な方式の数倍となり、このアルゴリズムをビデオレートステレオマシンで実現する場合の障害となっている。

一般の情景では、隠れ処理が本当に必要な部分が画面全体に占める割合は比較的小さいのが普通である。また、隠れが生じていない部分では、隠れ処理を行わない方式でも、SEAのように隠れ処理を行う方式でも、基本的には同じ奥行きマップを生成する。むしろ、隠れの無い部分では、隠れ処理を行わない単純な方式の方が良好な結果が得られることが多い。従って、隠れの無い部分には隠れ処理を行わない方式、隠れの有る部分には隠れ処理を行う方式を分けて適用することで、全ての領域で高精度な奥行きが獲得できる。

本研究では、図4.1に示すように、画面全体に隠れ処理を含まない方式を最初に適用し、隠れ処理が必要な部分にのみ、隠れ処理を含む方式を適用して、両方式で得られた奥行きマップを統合する多眼ステレオ法を提案する。この方式を実現するためには、隠れ処理が必要な領域の抽出法と、隠れ部分に有効な対応点探索アルゴリズムを開発することがポイ

ントである。前者に対しては、隠れ処理を行わない対応点探索の過程で得られた評価値を再利用することによって、十分な精度で隠れ領域を抽出する手法を提案する。また、後者については、複雑な隠れ形状にも対応可能な隠れ処理アルゴリズムを提案する。

また、シーン中の領域の持つ性質ごとに異なる奥行き推定方式を採用することで、従来の隠れ処理を含んだ多眼ステレオ法と比較して、ビデオレートのステレオマシンとの親和性が向上する。この方式をビデオレートステレオを用いて実現する場合を考える。まず、画面全体の奥行きをビデオレートのステレオにより獲得する。次に、隠れ領域に対して隠れ処理を行うステレオ法を適用する。最後に、ビデオレートステレオによって得られた奥行きの隠れ領域を隠れ処理を行った方式により得られた奥行きで更新する。このとき、隠れ処理を行わない方式によって生成される奥行きマップが殆んど遅延なく出力されるのに続いて、隠れ処理を行う方式によって得られた結果により更新された奥行きマップが少し遅延して出力される。すなわち、隠れの無い部分については高精度な奥行きマップが遅延なく出力され、物体輪郭部分では、最初は精度の悪い奥行きが出力され、少し遅延して高精度な奥行きマップに更新される。従って、画面中で静止している物体輪郭では高精度な奥行きマップが常に得られているが、物体輪郭が動いた場合には、その部分について高精度な奥行きマップは得られず、それが静止した時点で若干の遅延を含んで高精度な奥行きマップが出力される。実物体と仮想物体の隠れを表現する場合に、物体輪郭が移動し隠れの状況が変化している過程では、物体輪郭の奥行きマップの精度が観察者に与える影響は小さいと予想されるため、提案するこの手法は上記の隠れを表現するための奥行き獲得法として有効であると考えられる。

次節より本方式を実現するための隠れ領域の抽出手法および、隠れ領域の奥行きを高精度に獲得可能なステレオ法について詳しく述べる。また、ビデオレートステレオマシンを用いた本方式の実現方法についても後の節で説明する。

## 4.2 隠れ領域の抽出

本節では、隠れ処理を行わない多眼ステレオ法による対応点探索の過程で得られる各仮定視差ごとの評価値の分布をもとに隠れ領域を抽出する手法について述べる。

### 4.2.1 特徴空間の設定

隠れ処理を行わない多眼ステレオ法において、対応点探索時に算出される各仮定視差での評価値の分布と画像特徴の関係について考察する。ここでいう画像特徴とは、隠れやテクスチャの有無を表す。評価値の算出は、2.3.3 節で説明した方法で行うものとする。

各仮定視差での評価値の振舞いは、以下に示す画像特徴の組合せによって大きく三つのパターンに分類できる。

- テクスチャも隠れも無い領域(図 4.2(a))

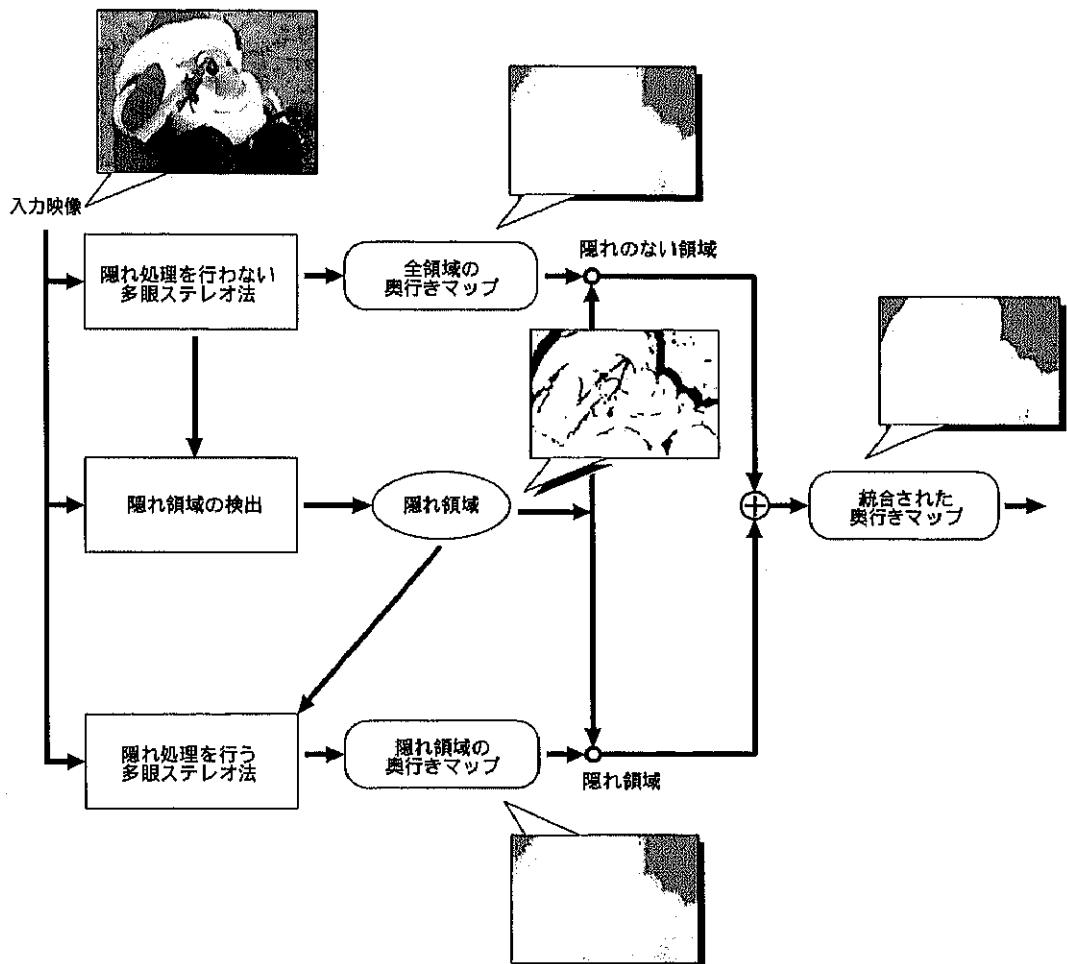


図 4.1: 隠れ処理を含まない方式と含む方式の連結統合による多眼ステレオ法

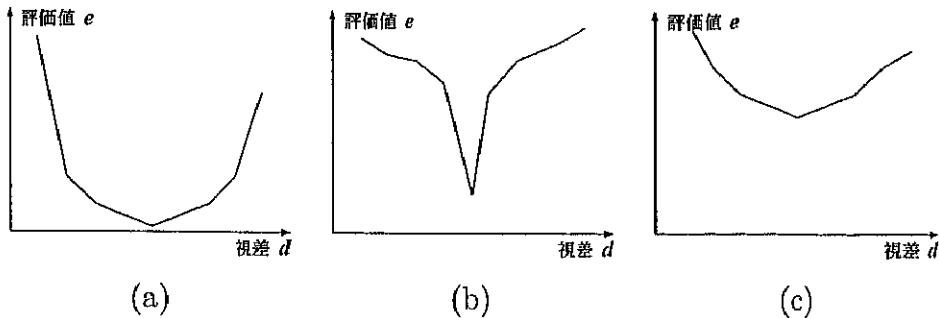


図 4.2: 評価値の分布と画像特徴との関係 (a) テクスチャも隠れも無い領域 (b) テクスチャが有り隠れが無い領域 (c) 隠れの有る領域

- テクスチャが有り隠れが無い領域 (図 4.2(b))
- 隠れが有る領域 (図 4.2(c))

隠れの無い部分で対応点探索を行った場合、テクスチャの少ない点では、図 4.2(a) のように、正解視差付近で評価値が小さくなる。また正解視差の近傍でも、正解視差を仮定したときと同様に均一色領域が存在するため、評価値は比較的小さい値をとる。テクスチャのある点では、図 4.2(b) のように、正解視差で評価値がテクスチャのない部分ほど小さくならず、その近傍の視差値での評価値が急速に大きくなる傾向がある。一方、隠れが有る部分では、隠れの有るステレオ画像対で対応点探索が失敗するため、図 4.2(c) に示すように、評価値が全体的に大きくなる。以上のことから、以下に示す二つの特徴量により隠れの有る点を抽出可能であると考える。

1. 評価値の最小値  $e$
2. 評価値が最小となる視差  $d$  近傍での評価値の変化率  $e'$

ここで  $e'$  は、評価値が最小となる視差  $d$  とそれに隣合った視差の評価値との差のうち小さい方として、式 (4.1) のように定義する。 $e(d)$  は視差  $d$  における評価値を表す。

$$e' = \min(e(d-1) - e(d), e(d+1) - e(d)) \quad (4.1)$$

設定した特徴量を軸とした特徴空間にシーン中の点をマッピングすると、図 4.3 に示すようなクラスタが形成される。この特徴空間にマッピングされた点をクラスタ分割することで隠れ領域の抽出を行う。次節で、直線によるクラスタ分割、距離関数による領域の再分類による隠れ領域の抽出法について説明する。

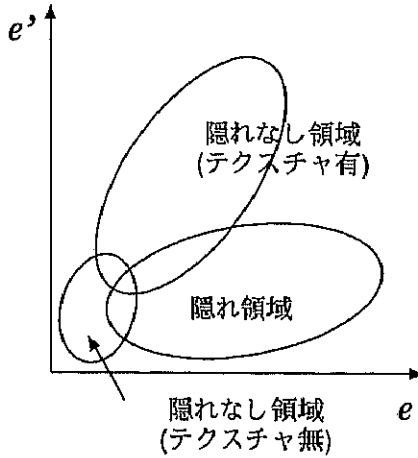


図 4.3: 特徴空間でのクラスタの分布

#### 4.2.2 二直線による初期クラスタの分割

図 4.4 に示した二直線により三つの初期クラスタに分割する。二つの直線は図 4.4 に示すパラメータ  $\theta$ ,  $\phi$  によって決定される。それぞれのパラメータは、図 4.5 に示すように、特徴空間中の点を、各パラメータに沿って投影した 1 次元のヒストグラムについて双峰性があると仮定し、判別分析法 [24] により最適な閾値を決定する。

パラメータ  $\theta$  を例に判別分析法による閾値決定法について以下に述べる。 $\theta$  の範囲は 0 から  $L$  までで、1 刻みで変化するものとする。角度  $i$  の直線上に存在する画素数を  $n_i$ 、全画素数を  $N = n_0 + n_1 + \dots + n_L$  とすると、直線上に存在する画素数の正規化ヒストグラムは、

$$p_i = \frac{n_i}{N} \quad (4.2)$$

で与えられる。このとき、直線上に存在する画素数の平均値および分散は、

$$\mu_T = \sum_{i=0}^L i p_i \quad (4.3)$$

$$\sigma_T^2 = \sum_{i=0}^L (i - \mu_T)^2 p_i \quad (4.4)$$

で与えられる。

ここで、 $k$  を閾値として、区間  $S_1 = [0, \dots, k]$  および  $S_2 = [k+1, \dots, L]$  に属する画素をそれぞれ二つのクラス  $C_1$  および  $C_2$  に分類するものとする。また、正規化ヒストグラムに対して、次の二つの類計量を定義する。

$$\omega(k) = \sum_{i=0}^k p_i \quad (4.5)$$

$$\mu(k) = \sum_{i=0}^k i p_i \quad (4.6)$$

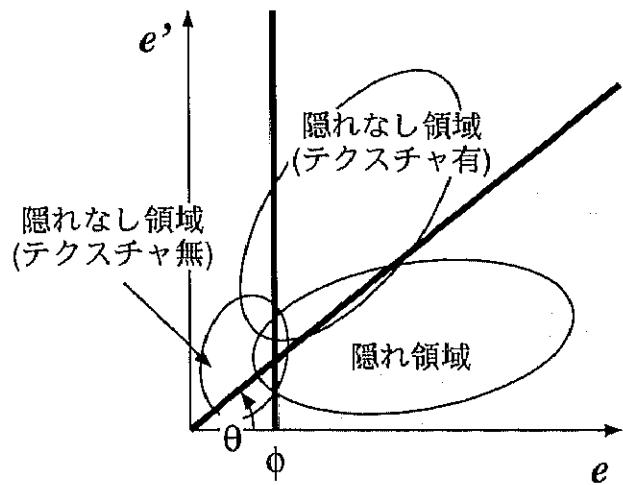


図 4.4: 二直線による初期クラスタ分割

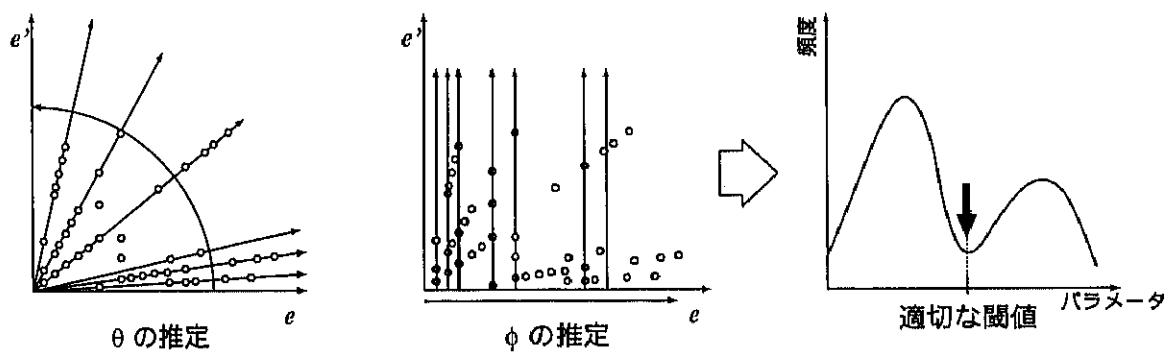


図 4.5: パラメータ  $\theta$ ,  $\phi$  の推定

このとき、閾値が  $k$  の時の各クラスの生起確率は、この累計量を用いて、

$$w_1 = \sum_{i \in S_1} p_i = \omega(k) \quad (4.7)$$

$$w_2 = \sum_{i \in S_2} p_i = 1 - \omega(k) \quad (4.8)$$

と表すことができる。また、各クラスの平均角度は、

$$\mu_1 = \frac{\mu(k)}{\omega(k)} \quad (4.9)$$

$$\mu_2 = \frac{\mu_T - \mu(k)}{1 - \omega(k)} \quad (4.10)$$

となる。これらの量の間には、閾値  $k$  によらず常に、

$$\omega_1 \mu_1 + \omega_2 \mu_2 = \mu_T, \omega_1 + \omega_2 = 1 \quad (4.11)$$

の関係が成立する。また、各クラスの角度の分散は、

$$\sigma_1^2 = \sum_{i \in S_1} \frac{(i - \mu_1)^2 p_i}{\omega_1} \quad (4.12)$$

$$\sigma_2^2 = \sum_{i \in S_2} \frac{(i - \mu_2)^2 p_i}{\omega_2} \quad (4.13)$$

で与えられる。閾値  $k$  で二つのクラスに分類したとき、クラスの分離度が高く、それぞれのクラスでのまとまりが良い程、閾値が適切であると考えられる。従って、最適な閾値  $k$  は、クラスタ内分散  $\sigma_W^2$  とクラスタ間分散  $\sigma_B^2$  の比  $\sigma_B^2/\sigma_W^2$  を最大にするパラメータとして求めることができる。ここで、クラスタ内分散  $\sigma_W^2$  とクラスタ間分散  $\sigma_B^2$  は以下の式で与えられる。

$$\sigma_B^2 = \omega_1 \sigma_1^2 + \omega_2 \sigma_2^2 \quad (4.14)$$

$$\sigma_W^2 = \omega_1 \omega_2 (\mu_1 - \mu_2)^2 \quad (4.15)$$

#### 4.2.3 距離関数によるクラスタの再分類

前節で得られた初期クラスタは二直線によって単純に領域を分割したものであり、実際のクラスタの散らばりを考慮していないため、クラスタの分割精度低い。そこで、初期クラスタのクラスタ中心と各点との距離をもとにクラスタの再分類を行う。これにより、隠れ領域の検出漏れを低減することができる。

再分類の手順を以下に示す。まず、それぞれのクラスタにおけるクラスタ中心  $M_i$ 、分散共分散行列  $V_i$  を求める。

$$M_i = (M_{i1}, M_{i2}) \quad (4.16)$$

$$M_{ik} = \frac{1}{N_i} \sum_{t=1}^{N_i} X_{tk} \quad (4.17)$$

$$V_i = \begin{bmatrix} V_{i11} & V_{i12} \\ V_{i21} & V_{i22} \end{bmatrix} \quad (4.18)$$

$$V_{ikt} = \frac{1}{N_i - 1} \sum_{t=1}^{N_i} (X_{tk} - M_{ik})(X_{ti} - M_{ii}) \quad (4.19)$$

ここで、 $X_t = (X_{t1}, X_{t2})$  であり、 $X_{t1}$  は点  $t$  における評価値の最小値  $e$ 、 $X_{t2}$  は式 (4.1) で定義した点  $t$  における評価値の変化率  $e'$  を表す。また  $N_i$  は、クラスタ  $i$  に含まれる点の数を表す。次に、式 (4.20) に示すマハラノビス距離  $d_i$  を、サンプル点  $X_t$  と各クラスタ中心との間で計算し、距離が最小になるクラスタへ再分類する。距離関数としてマハラノビス距離を採用することで、クラスタの分散を考慮した適切なクラスタ分割が行える。

$$d_i = (X_t - M_i) V_i^{-1} (X_t - M_i)^T \quad (4.20)$$

### 4.3 従来までの隠れ処理手法

隠れは中心画像上では観測できるシーン中の点が一つ以上の周辺画像上で手前の物体に隠されて観測できなくなる現象である。隠れが起こっているステレオ画像対では、シーン中の点が持つ真の視差  $\tilde{d}$  に対して、前述の  $I^{m,n}(x - m\tilde{d}, y - n\tilde{d})$  は手前の物体の像であり、この点における相違度は小さくならない。この結果、 $\tilde{d}$  に対する評価値も大きくなり、式(2.12)を最小にする推定視差として  $\tilde{d}$  とは別のものが選ばれる可能性が大きくなる。これを解決するためには、隠れが発生している周辺画像を排除して式(2.12)を計算すればよいが、あるステレオ画像対における視差  $d$  についての相違度が大きいときに、それが隠れによるのか、それとも、視差  $d$  が正しくないことによるのかを分別することは困難である。

SEA では、隠れ輪郭の形状が局所的には直線であると仮定し、 $3 \times 3$  のカメラマトリクスにおける 8 個の隠れパターンを設定して対応探索を行う。この隠れパターンを用いて隠れが無い場合と有る場合の複数の仮定を立て、最も確からしいものを選択することで「偽の視差」と「隠れが発生している正しい視差」の分別を行っている。

SEA による隠れ処理を行った対応点探索は以下の手順で行われる。

1. 隠れパターンを表現した図 4.6 に示す 8 タイプの隠れマスクを用意する。

網かけ部分は隠れが発生している画像とし、その値を 0、白い部分は隠れがない画像とし、その値を 1 で表す。また、 $M_0$  は隠れが全く生じていない場合を表す。

2. 隠れの発生している周辺画像は利用しないという方針により、タイプ  $t$  の隠れパターンが生じていると仮定した時の視差  $d$  の評価関数を次式のように定義する。

$$e_t(x, y, d) = \frac{w_t}{n_t} \sum_{M_t(m,n)=1} e^{m,n}(x, y, d) \quad (4.21)$$

ここで、 $n_t$  は各マスク  $M_t$  の周辺画像における 1 の数である。 $w_t$  は隠れが生じていない場合に  $t=0$  が選択されやすくする重みで、 $w_0=1$ 、それ以外の時は 1 よりもわずかに大きい値に設定する。このとき、 $e_t(x, y, d)$  は、真の視差  $\tilde{d}$  を仮定し、本来の隠れパターンに対応する隠れマスクを用いて評価した場合に最小となることが期待できる。

3. 式(4.21)に従って、各隠れパターンを仮定して求めた評価値を最小とする  $d$  を  $I^{0,0}(x, y)$  における推定視差  $\hat{d}$  とする。

$$e(x, y, d) = \min_t e_t(x, y, d) \quad (4.22)$$

これにより、隠れの生じているステレオ画像対の影響を受けること無く正しい視差を推定することが可能となる。

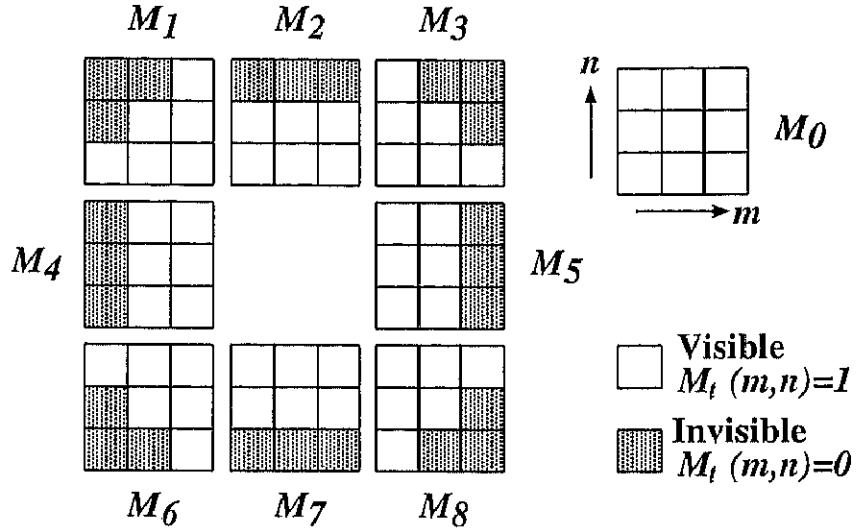


図 4.6: SEA で用いられる隠れマスク

#### 4.4 ソート法による隠れ処理

SEA では隠れの影響を避けるため、隠れ部分において高い頻度で現れる隠れ形状を隠れマスクとして設定したマスク法により対応点探索を行っている。しかし、この方法は式(4.21)で定義される評価値でもわかるように、隠れが無いと仮定したステレオ画像対での相違度の総和として評価値を計算するため、隠れの判断はあくまで「隠れパターン」としての判断しか行えず、各々のステレオ画像対ごとに隠れの有無を判断することは不可能である。そのため、マスク法では設定した隠れマスクから外れた複雑な隠れには対応できないという問題がある。

本節では、隠れ形状を考慮した隠れマスクを用いずに隠れの影響を回避する手法としてソート法を提案する。本方式では、隠れの生じているステレオ画像対での相違度は隠れの無い画像対での相違度よりも大きな値をとると仮定し、図4.7に示すように、カメラマトリクスから得られるステレオ画像対での相違度中から、相違度が十分小さく隠れが生じていないと判断されるものだけを評価の対象とし対応点探索を行う。従って、隠れのパターンを特定することなく隠れの影響を回避できる。その反面、隠れの形状を考慮していないため、ノイズの影響を受けやすく、隠れのない部分での視差推定精度が低下する傾向がある。ソート法による視差推定手順は以下のようである。

1. 画素  $I^{0,0}$  と視差  $d$  と仮定したときの各周辺画像上の対応点  $I^{k,l}(x - kd, y - ld)$  との相違度  $e^{k,l}(x, y, d)$  を求める。相違度は、2画素間(または、各画素を中心とした小ブロック間)のRGB空間におけるユークリッド距離を用いる。
2. 各ステレオ画像対で得られた相違度  $e^{k,l}(x, y, d)$  を昇順にソートする。
3. 昇順にソートされた相違度  $e^{k,l}(x, y, d)$  に対して、小さい方から  $m$  個までの相違度の

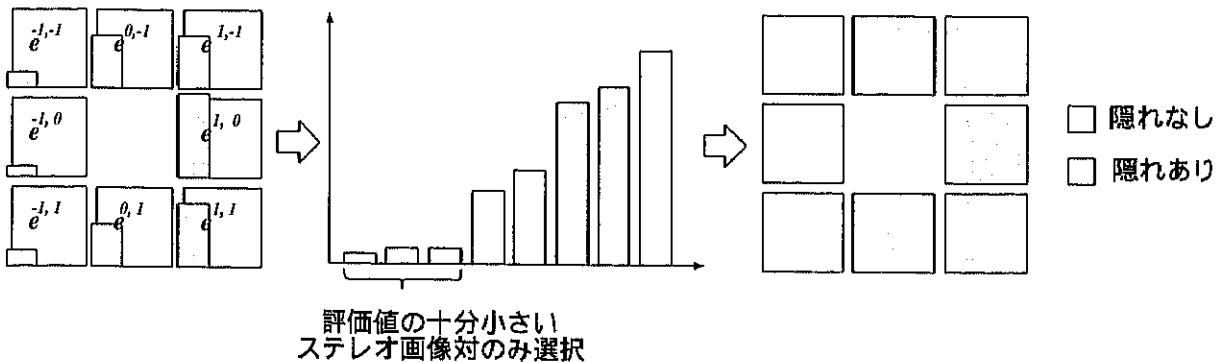


図 4.7: ソート法による隠れの影響の回避

平均値  $ave_m$ , および  $m + 1$  個までの相違度の平均値  $ave_{m+1}$  を計算する.

4. 2つの平均値の差が閾値  $T$  以上となる  $\hat{m}$  を探し, そのときの  $ave_{\hat{m}}$  を視差  $d$  での評価位置  $e(x, y, d)$ (式(4.23)) とする.

$$e(x, y, d) = \frac{1}{\hat{m}} \sum_{i=1}^{\hat{m}} f(i) \quad (4.23)$$

ここで  $f(i)$  は, 昇順にソートされた  $i$  番目の相違度を表す. 十分な視差推定精度を確保するために必要なステレオ画像対数  $L$  は, 対応点が観測されていることが望ましいステレオ画像対数と隠れの生じている点で確保可能な最低ステレオ画像対数の関係 [9] から決定できる. これを用いて  $m$  の探索範囲を  $L \leq m \leq N$  と定める. ここで  $N$  はカメラマトリクスの台数を表す.

## 4.5 実験結果

提案した手法により隠れ抽出を行い, 隠れのない部分では隠れ処理を含まない多眼ステレオ法の結果を採用し, 隠れ部分についてはソート法を適用した結果を求めて統合する視差推定実験を行った. 実験は, 十字配置の5台のカメラマトリクスと格子状に配置した  $3 \times 3$  台のカメラマトリクスを用いて行った. 実験に用いた画像を図 4.8-4.10 に示す. 図 4.11 に 5 眼十字配置を用いた場合, 図 4.12 に 9 眼マトリクス配置を用いた場合の隠れ領域の抽出結果を示す. 図 4.11(a) は, 直線による領域分割によって得た隠れ領域, 図 4.11(b) は直線による領域分割によって得たクラスタを更に距離関数によって再分類することで抽出した隠れ領域である. 比較のために直線による領域分割によって得た各クラスタのクラスタ中心を初期値として k-means 法により領域を再分類した結果を図 4.11(c) に示す. 図 4.11(d) は, 視差の正解データをもとに生成した正解隠れ領域<sup>1</sup>である. 同様に異なるシ

<sup>1</sup> 正解の視差画像をもとに, 物体輪郭付近で生じる隠れ領域で, かつ 1 画素以上の幅を持つ領域を隠れの正解データとして生成した.

ンに対して隠れ領域の抽出を行った結果を図4.13-図4.16に示す。表4.1, 4.2に閾値による分割で得た隠れ領域と、再分類によって抽出した隠れ領域の正答率を示す。k-means法による再分類結果では、誤検出が増加する傾向が見られた。これは初期値として与えた直線による領域分割によって得た各クラスタのクラスタ中心がk-means法を用いることで更新され、隠れ領域ではない点も隠れ領域として判断されたためと考えられる。従って、直線による領域分割によって得た各クラスタはある程度精度良く得られていると判断できる。更に、このクラスタ中心を変更することなく距離関数を用いた再分類を行うことで、隠れ領域の抽出精度を向上させることができた。図4.5に図4.12(c), 4.14(c), 4.16(c)の領域分割結果における各クラスタの分布を示す。

表4.3, 4.4に提案手法による視差推定結果の正答率<sup>2</sup>を示す。比較のために、隠れ処理を含まない多眼ステレオ法、および、SEA、全画面に対して適用したソート法による正答率を示す。各手法を用いて視差推定を行った結果および、正解視差画像を図4.18-図4.23に示す。

図4.12-図4.16の実験結果から、閾値による分割では抽出漏れした隠れ領域が、再分類により、抽出されていることが確認できる。抽出誤りにより若干隠れのないテクスチャ領域も抽出されているが、表4.3, 4.4の正答率からもわかるように、隠れのない点にソート法を適用した場合にも、ほぼ正しい視差が求められるので問題ない。

SEAと比較すると、9眼の場合には隠れ領域において、1~20%の精度の向上が認められる。また、隠れ領域の視差推定精度は、全画面に対してソート法を適用した結果が一番良好なことから、隠れ領域に対するソート法の有効性は明らかである。5眼の場合には、SEA、すなわちマスク法の方がソート法よりも全般に良い結果を得ている。これは、眼数が少ない時には、ソート法はノイズの影響を受けやすく、隠れパターンに制約を加えたマスク法が優位になることを示している。従って、本提案手法でも5眼の場合には、隠れ処理方式としてマスク法を採用することを考慮してもよい。

表4.1: 5眼十字配置を用いたときのカメラ隠れ領域検出の正答率(隠れ領域の検出率/誤検出率)

	データ1(石膏像)	データ2(サンタ)	データ3(植物)
閾値による分割で得た隠れ領域	51.1/2.7	56.1/3.1	74.3/13.7
再分類による隠れ領域	67.4/5.4	75.4/5.7	79.8/18.1
k-means法による隠れ領域	59.5/30.5	52.8/3.5	80.7/25.3

<sup>2</sup>正解の視差画像を手動で作成し、推定視差画像の各画素の値が正解の視差の±1以内にある割合を正答率とした。

表 4.2: 9眼マトリクス配置を用いたときの隠れ領域検出の正答率(隠れ領域の検出率/誤検出率)

	データ 1(石膏像)	データ 2(サンタ)	データ 3(植物)
閾値による分割で得た隠れ領域	51.1/2.7	56.1/3.1	74.3/13.7
再分類による隠れ領域	67.4/5.4	75.4/5.7	79.8/18.1
k-means 法による隠れ領域	57.8/30.6	89.9/31.9	62.8/9.0

表 4.3: 5眼十字配置を用いたときの各手法による視差推定結果の正答率(画像全体/隠れのない領域/隠れ領域)

	データ 1(石膏像)	データ 2(サンタ)	データ 3(植物)
隠れ処理を含まない方式	92.5/97.7/47.6	88.3/93.2/45.5	72.5/91.1/36.3
SEA	95.2/97.8/72.7	92.1/93.4/80.7	77.4/91.5/50.0
ソート法(全画面)	91.6/94.1/70.5	90.9/92.0/80.6	78.8/90.1/56.8
提案手法	94.0/97.5/64.1	92.0/93.3/80.2	78.9/91.3/55.0

表 4.4: 9眼マトリクス配置を用いたときの各手法による視差推定結果の正答率(画像全体/隠れのない領域/隠れ領域)

	データ 1(石膏像)	データ 2(サンタ)	データ 3(植物)
隠れ処理を含まない方式	92.7/98.2/45.0	88.7/94.7/36.1	72.9/91.9/35.9
SEA	95.1/98.3/67.7	93.2/95.2/75.3	76.1/92.4/44.7
ソート法(全画面)	95.6/97.8/77.1	93.7/94.8/83.9	82.9/91.9/65.3
提案手法	95.2/98.2/68.7	93.8/95.1/82.2	82.6/92.1/64.0

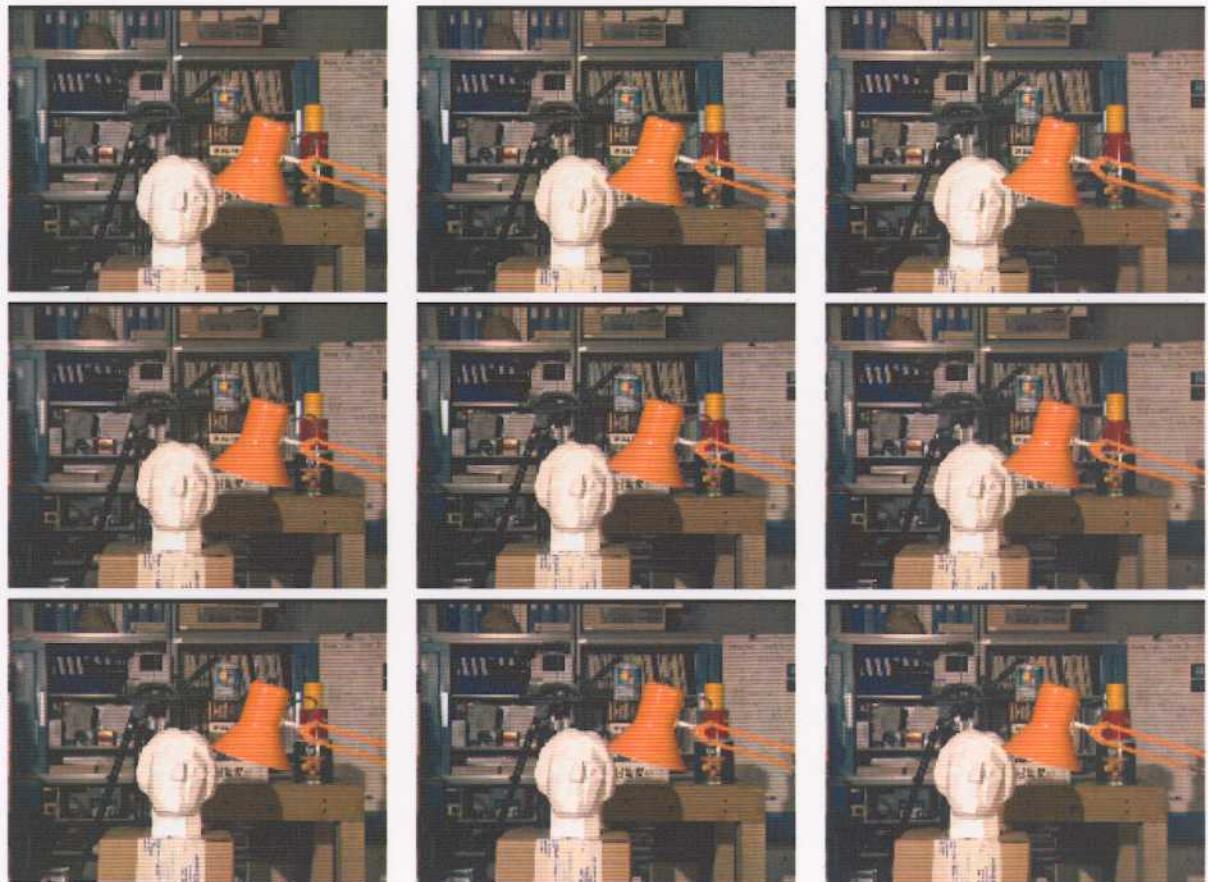


図 4.8: 実験データ 1(石膏像)



図 4.9: 実験データ 2(サンタ)



図 4.10: 実験データ 3(植物)



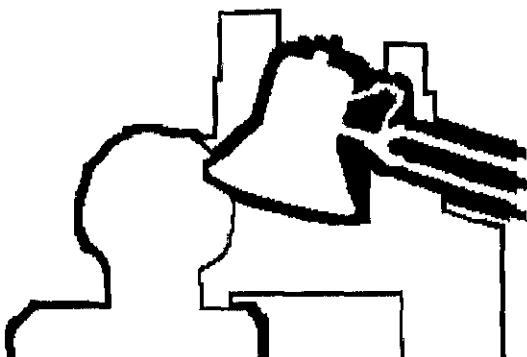
(a) 直線による領域分割



(b) 距離関数による領域の再分類



(c) k-means 法による領域分割



(b) 正解隠れ領域

図 4.11: 隠れ領域の抽出結果 (石膏像, 5 眼十字配置)



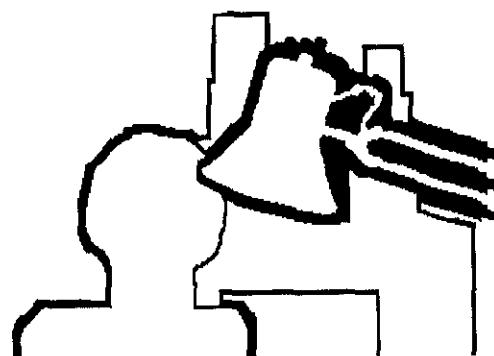
(a) 直線による領域分割



(b) 距離関数による領域の再分類



(c) k-means 法による領域分割

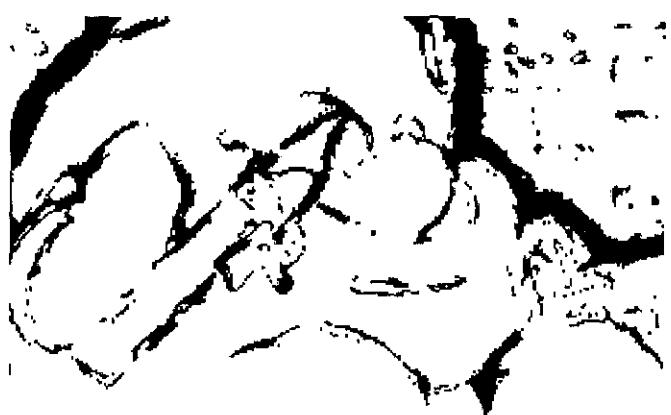


(b) 正解隠れ領域

図 4.12: 隠れ領域の抽出結果 (石膏像,9 眼マトリクス配置)



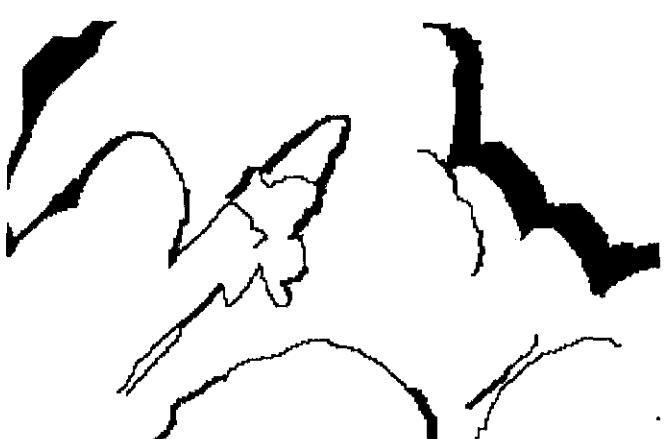
(a) 直線による領域分割



(b) 距離関数による領域の再分類



(c) 別手法による領域分割

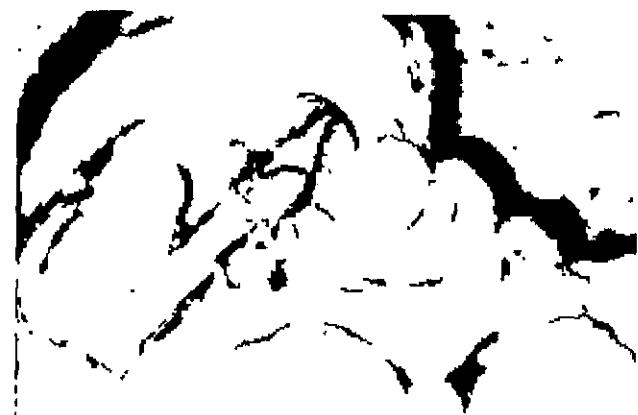


(b) 正解隠れ領域

図 4.13: 隠れ領域の抽出結果 (サンタ,5 眼十字配置)



(a) 直線による領域分割



(b) 距離関数による領域の再分類



(c) k-means 法による領域分割



(b) 正解隠れ領域

図 4.14: 隠れ領域の抽出結果 (サンタ,9眼マトリクス配置)



(a) 直線による領域分割



(b) 距離関数による領域の再分類



(c) 別手法による領域分割



(d) 正解隠れ領域

図 4.15: 隠れ領域の抽出結果 (植物, 5眼十字配置)



(a) 直線による領域分割



(b) 距離関数による領域の再分類

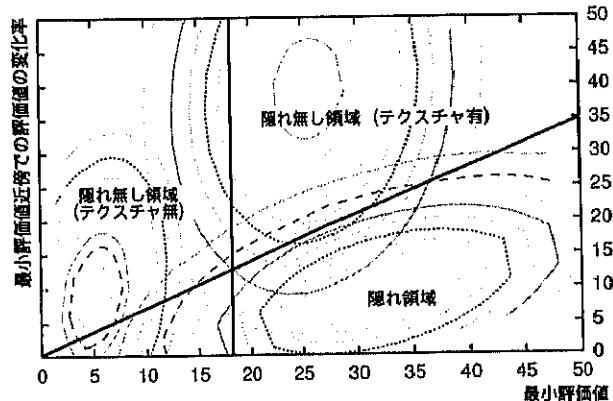


(c) k-means 法による領域分割

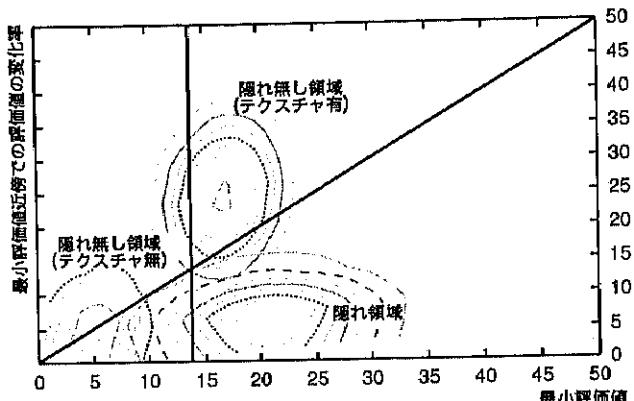


(b) 正解隠れ領域

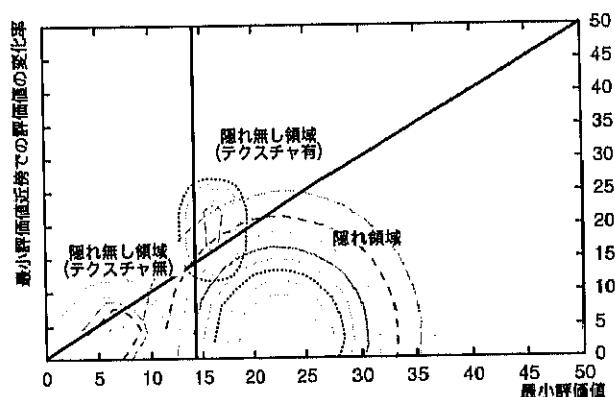
図 4.16: 隠れ領域の抽出結果(植物,9眼マトリクス配置)



(a) 石膏像

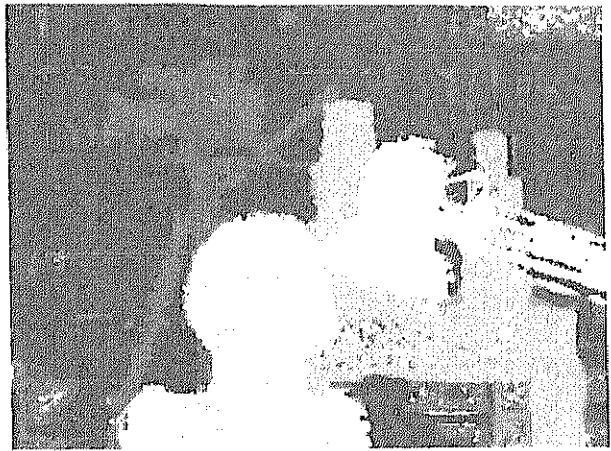


(b) サンタ

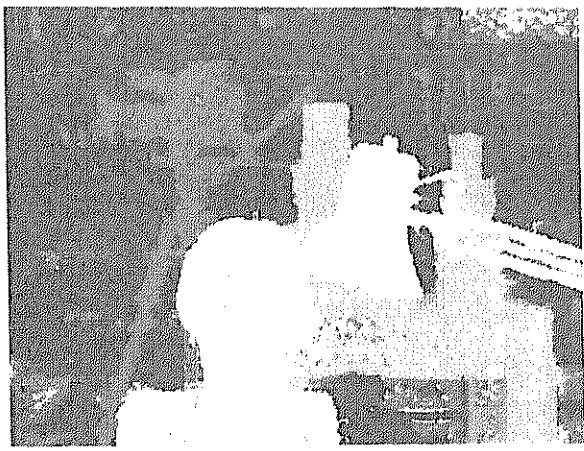


(c) 植物

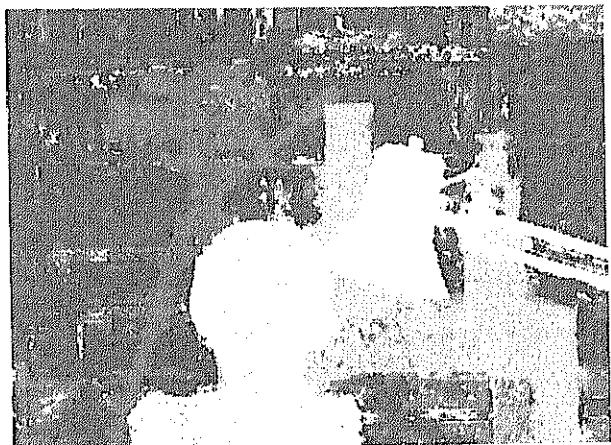
図 4.17: 領域分割によって得た各クラスタの分布



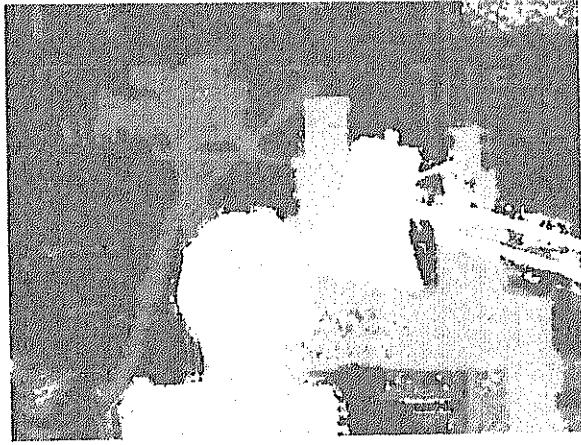
(a) 隠れ処理を含まない多眼ステレオ



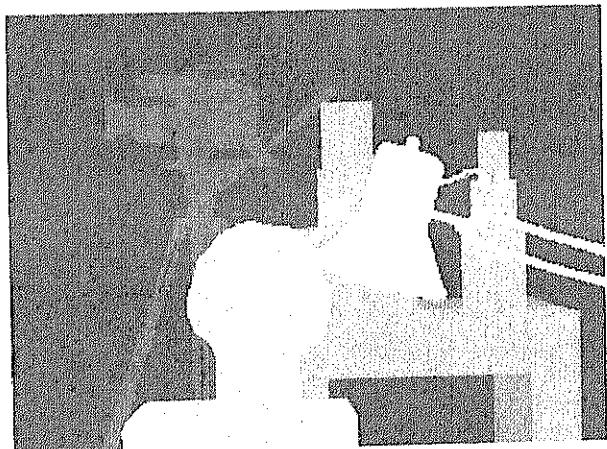
(b) SEA



(c) ソート法



(d) 提案手法(抽出した隠れ領域を使用)



(e) 正解データ

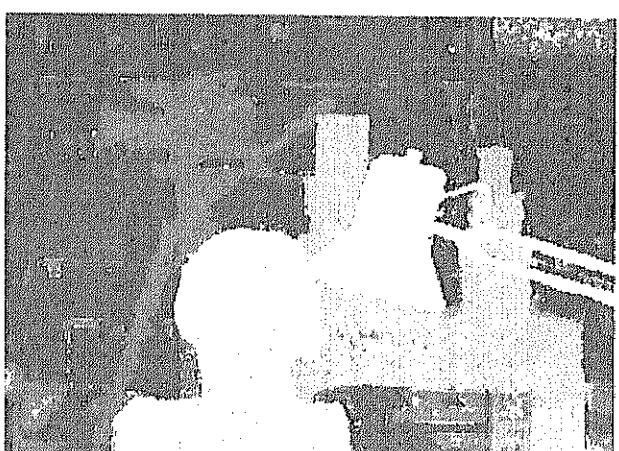
図 4.18: 視差推定結果(石膏像,5眼十字配置)



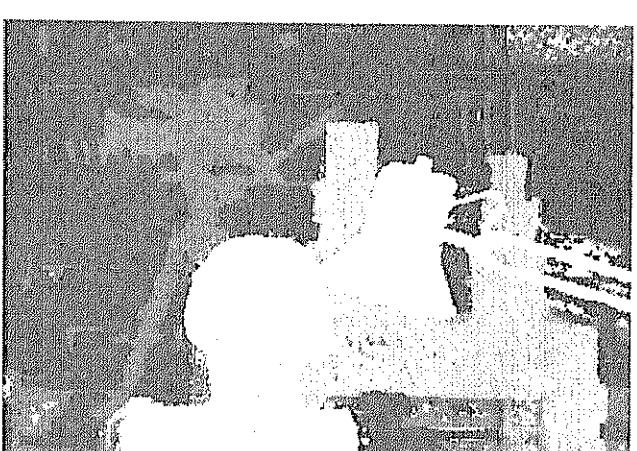
(a) 隠れ処理を含まない多眼ステレオ



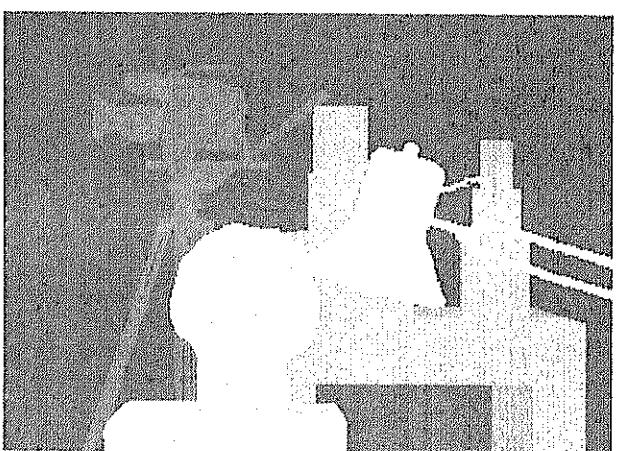
(b) SEA



(c) ソート法



(d) 提案手法(抽出した隠れ領域を使用)



(e) 正解データ

図 4.19: 視差推定結果(石膏像,9眼マトリクス配置)



(a) 隠れ処理を含まない多眼ステレオ



(b) SEA



(c) ソート法



(d) 提案手法(抽出した隠れ領域を使用)



(e) 正解データ

図 4.20: 視差推定結果(サンタ,5眼十字配置)



(a) 隠れ処理を含まない多眼ステレオ



(b) SEA



(c) ソート法

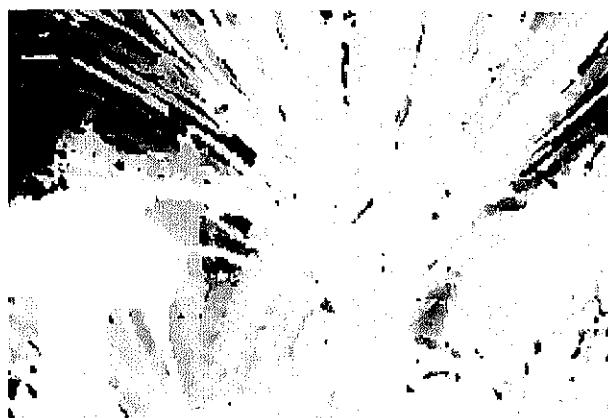


(d) 提案手法(抽出した隠れ領域を使用)

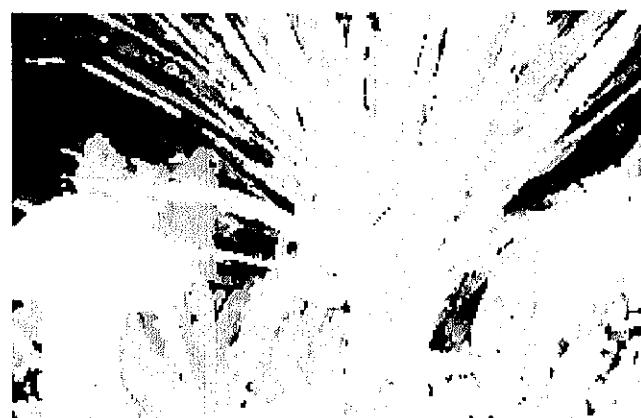


(e) 正解データ

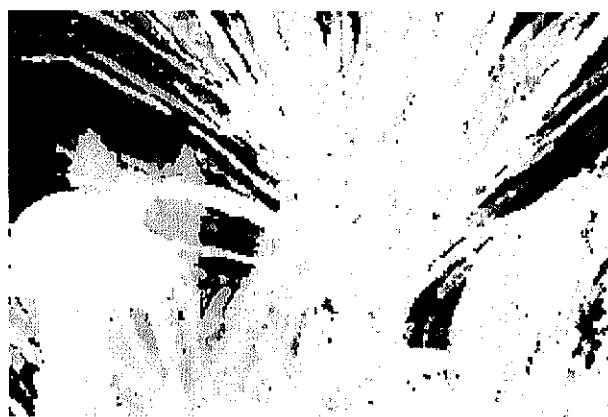
図 4.21: 視差推定結果 (サンタ,9眼マトリクス配置)



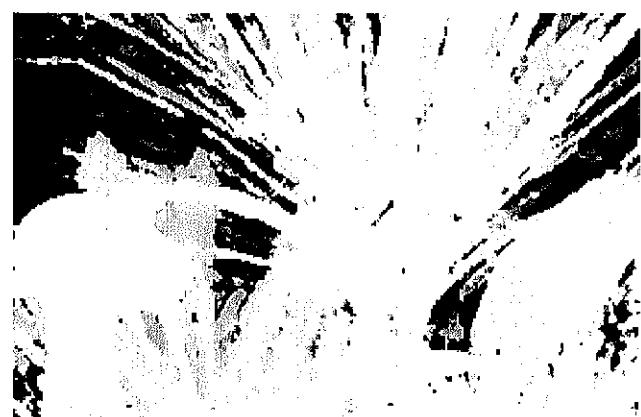
(a) 隠れ処理を含まない多眼ステレオ



(b) SEA



(c) ソート法



(d) 提案手法(抽出した隠れ領域を使用)

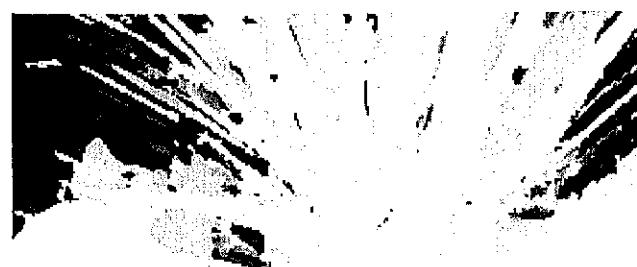


(e) 正解データ

図 4.22: 視差推定結果(植物,5眼十字配置)



(a) 隠れ処理を含まない多眼ステレオ



(b) SEA



(c) ソート法



(d) 提案手法(抽出した隠れ領域を使用)



(e) 正解データ

図 4.23: 視差推定結果(植物,9眼マトリクス配置)

## 4.6 ビデオレートステレオを用いたシステムの構築

本節では、前述の隠れ処理を含まない方式と含む方式を統合したステレオ法をビデオレートのステレオマシンを用いて実現したシステムの構成および、実現方法について述べる。

### 4.6.1 システムの概要

隠れ処理を含まない方式と含む方式を統合したステレオ法をビデオレートステレオマシンを用いて構築する。図 4.24 に構築したステレオシステムの概要を示す。隠れ領域を抽出する代わりにシーン中の動物体を抽出し、動領域の奥行きマップをビデオレートステレオを用いた隠れ処理を含まないステレオ法により獲得する。一方、静止領域の奥行きマップは隠れ処理を含んだステレオ法により獲得する。二つの異なるステレオ法により獲得された奥行きマップを動き情報をもとに統合することで、動きのある領域では、若干精度が落ちるが遅れのない奥行きマップが獲得でき、静止領域では、奥行き獲得に多少の時間がかかるが精度のよい奥行きマップが獲得可能なステレオシステムが実現できる。

### 4.6.2 システムの構成

ビデオレートステレオシステムの構成を図 4.25 に示す。本システムは、動領域抽出モジュール、隠れ処理を含まないステレオモジュール、隠れ処理を含むステレオモジュール、奥行きマップ統合モジュール、およびデータ管理モジュールから構成される。入力映像はビデオレートステレオマシンで利用される 5 眼カメラヘッドから入力される。動領域抽出モジュール、隠れ処理を含まないステレオモジュール、隠れ処理を含むステレオモジュールはそれぞれ独立して並行に動作する。動領域抽出モジュールでは、5 眼カメラヘッドの中心カメラの映像を入力とし、背景差分によりシーン中の動領域を抽出する。抽出された動領域データは圧縮され TCP/IP ソケット通信によりデータ管理モジュールに転送される。隠れ処理を含んだステレオモジュールでは、データ管理モジュールから動領域データを TCP/IP ソケット通信により獲得し、そのデータをもとに静止領域での奥行きマップを推定する。獲得された奥行きマップは TCP/IP ソケット通信により、データ管理モジュールに転送される。隠れ処理を含まないステレオモジュールにはビデオレートステレオマシンを利用し、シーン全体の奥行きマップをビデオレートで獲得し、ディジタル信号で出力する。奥行きマップ統合モジュールでは、データ管理モジュールで管理された動領域データをもとに、隠れ処理を含まないステレオモジュールの結果から動領域の奥行きマップを、隠れ処理を含むステレオモジュールの結果から静止領域の奥行きマップを選択し統合を行う。データ管理モジュールでは、動領域抽出モジュールによって得られた動領域データ、および隠れ処理を含んだステレオモジュールによって得られた奥行きマップをダブルバッファを用いて管理する。ダブルバッファを用いてデータを管理することで、隠れ処理

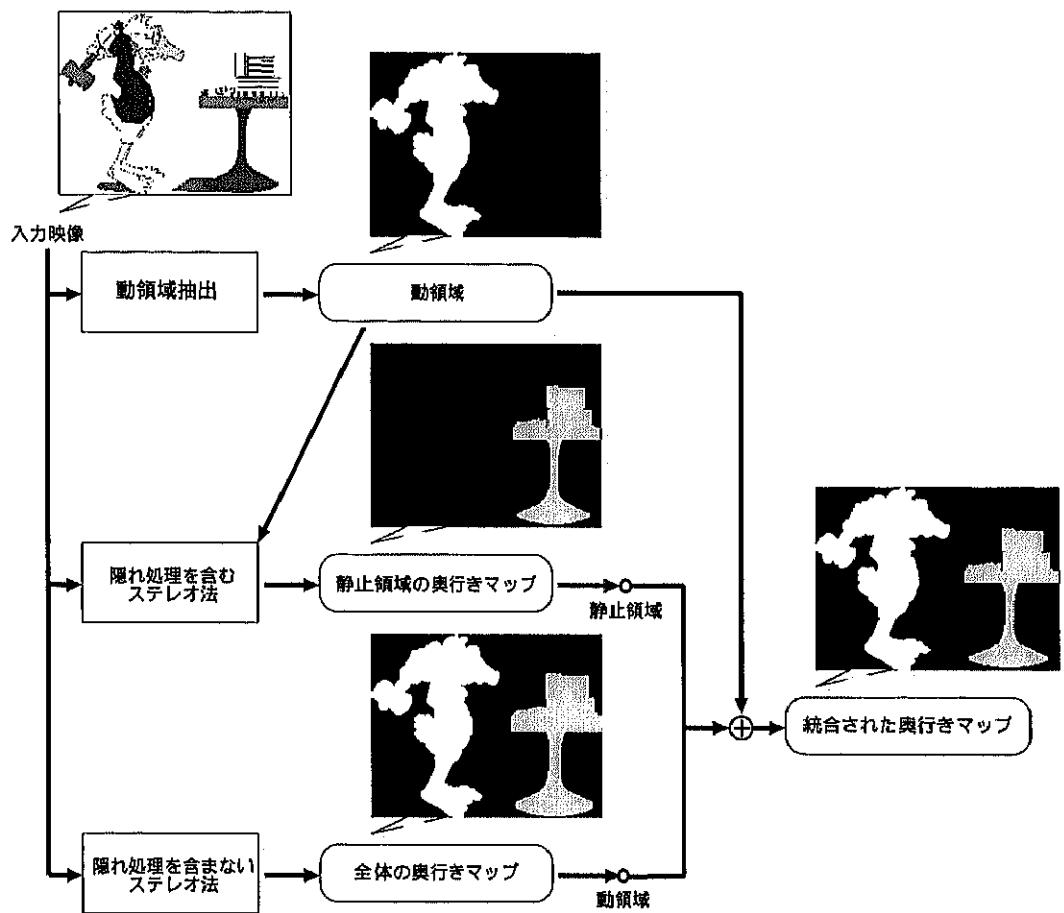


図 4.24: 隠れ処理を含まない方式と含む方式を連結統合したステレオシステムの概念図

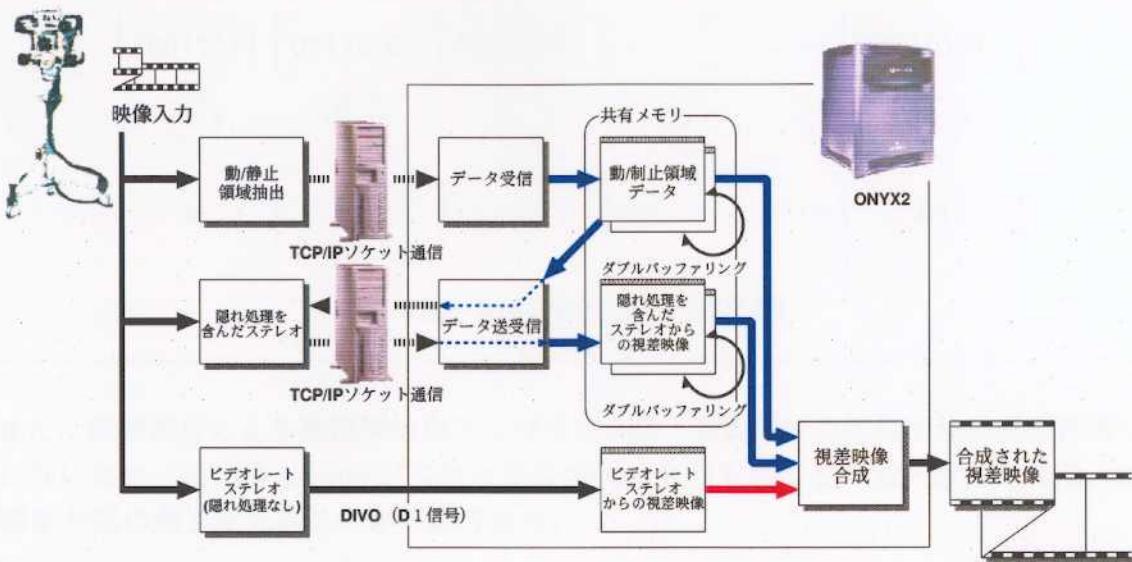


図 4.25: 隠れ処理を含まない方式と含む方式を連結統合した多眼ステレオシステム

を含んだステレオモジュールや奥行きマップ統合モジュール側で必要となるデータの転送待ちを解消できる。以下では、各モジュールの詳細について説明する。

### 動領域抽出モジュール

動領域抽出モジュールでは5眼カメラヘッドの中心カメラから得られた映像中の動領域を背景差分を用いて抽出する。背景画像は動物体がない状態で一番始めに獲得し、その後一定周期で更新を行う。動領域抽出、および背景画像の更新処理の流れを以下に示す。

1. 現在の入力画像と背景画像との差分を取り、差分値が閾値<sup>3</sup>よりも大きい場合、その点を動領域とする。
2. 画像中の全ての点にカウンタを設定し、動領域と判断された点のカウンタを1増やす。これにより、ある一定時間内にその点がどれくらいの割合動領域であったかを調べる目安とする。
3. 上記の処理を繰り返し、カウンタの値が閾値<sup>4</sup>以上であった点を一定時間停留した点、すなわち物体位置の移動などで新しく出現した背景物体と判断し背景とする。

抽出結果は動領域を1、静止領域を0として表現される。データの転送コストを向上させるため、図4.26に示すように8pixelごとにデータをパッキングしてTCP/IPソケット通信でデータ管理モジュールに転送する。これにより送信するデータ量を1/8に抑えることができる。

<sup>3</sup> 使用する環境に応じて最適値が変化するため、経験的に決定している。

<sup>4</sup> システムの動作速度、および使用条件に応じて経験的に決定する。

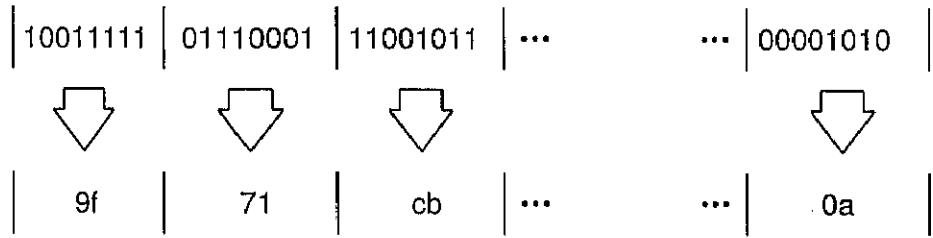


図 4.26: 動領域抽出結果の圧縮

また、背景差分による動領域抽出アルゴリズムは、各画素での処理が他の画素情報に依存しないため、Intel Pentium プロセッサの MMX 命令を利用した並列計算により、処理時間を大幅に向上させることが可能である。

### 隠れ処理を含まないステレオモジュール

隠れ処理を含まないステレオモジュールには、既製のビデオレートステレオマシンを利用する。ビデオレートステレオマシンを用いることで、密で精度の良い奥行きマップを高速に獲得できる。獲得された奥行きマップはデジタル信号で毎秒 30 フレーム出力される。

### 隠れ処理を含むステレオモジュール

動領域抽出モジュールによって抽出された動領域情報をもとに、静止領域に対して隠れ処理を含むステレオモジュールが適用される。この処理は、PC を用いてソフトウェアで実行される。また、カメラに幅轍があり各カメラの光軸が並行でないような場合や異なるカメラを用いる場合には、各カメラの位置姿勢を得るためのキャリブレーションおよび、カメラの個体差によって生じる輝度値の補正といった調整を予め正確に行うことが必要となる。カメラのキャリブレーションについては、Tsai のキャリブレーション法 [25] が有名である。本モジュールでは、処理の高速化のために、カメラキャリブレーションによって得られる中心カメラと周辺カメラの位置関係をもとに、中心カメラで得られる画像上の点における探索範囲内の全ての視差を仮定した時の周辺画像での対応点位置を予め計算しておく。本研究で用いたカメラキャリブレーション法、輝度値の補正法およびカメラキャリブレーション結果を用いた対応点と視差の決定法については付録 C, D, E に記す。

### 奥行きマップ統合モジュール

このモジュールでは、データ管理モジュールで管理された動領域データをもとに、隠れ処理を含まないステレオモジュールから得られる奥行きマップからは動領域部分を、隠れ処理を含んだステレオモジュールから得られる奥行きマップからは静止領域部分を選択し

統合する。隠れ処理を含まないステレオモジュールからは、奥行きマップがビデオレートが出力されるが、動領域データや隠れ処理を含んだステレオモジュールからの奥行きマップは若干遅延して出力される。しかし、データ管理モジュールでこれらのデータをダブルバッファを利用して管理することで、奥行きマップ統合モジュールはデータの転送待ちをすることなく、ビデオレートで奥行きマップを統合して出力することが可能である。

### データ管理モジュール

動領域抽出モジュールや隠れ処理を含んだステレオモジュールは、計算コストが高いため処理結果がビデオレートでは出力されない。ビデオレートで出力される隠れ処理を含まないステレオ結果に合わせて奥行きマップを統合し出力するために、データ管理モジュールはダブルバッファを利用して、動領域データと隠れ処理を含んだステレオモジュールから得られる奥行きマップを管理する。ダブルバッファを用いて一方のバッファをそれぞれのモジュールからの書き出し用、もう一方を奥行き統合モジュールからの読み取り用として使用することで、奥行き統合モジュールからは常にビデオレートでどちらかのバッファにアクセスが可能となる。

### 4.6.3 ビデオレートステレオシステムの構築

前節で述べたビデオレートステレオシステムをビデオレートステレオマシン、SGI ワークステーション Onyx2 および PC を用いて構築した。

動領域抽出モジュールと隠れ処理を含んだステレオモジュールは PC を用いて実装した。動領域抽出モジュールでは、5 眼カメラヘッドから入力された大きさ  $360 \times 243$  の映像中から背景差分により動領域を抽出する。背景差分アルゴリズムは Intel Pentium プロセッサの MMX 命令を用いて実装した。これにより毎秒 18 フレームのスループットを実現した。また、背景映像の更新は 5 秒毎に行う。隠れ処理を含んだステレオモジュールでは、 $360 \times 243$  の映像の静止領域について奥行き獲得を行う。奥行き探索範囲は 32 段階である。全ての領域が静止領域の場合、奥行き獲得時間は 5 から 8 秒である。

隠れ処理を含まないステレオモジュールには既製のビデオレートステレオマシンを利用することで  $360 \times 240$  の映像の奥行きマップを毎秒 60 フレーム獲得することが可能である。使用したビデオレートステレオマシンの詳細は付録 F に示す。

データ管理および奥行き統合処理は SGI ワークステーション Onyx2 によって行われる。動領域データと隠れ処理を含んだステレオから得られる奥行きマップは、TCP/IP ソケット通信により Onyx2 に送信される。ビデオレートステレオマシンで得られる奥行きマップはディジタル D1 信号で Onyx2 に装着された DIVO で取り込まれる。データ管理モジュールでダブルバッファによりデータを管理することで、毎秒 30 フレームでの奥行きマップを統合・出力が実現された。

図 4.27 に実験結果の例を示す。図 4.27(a), (b) はそれぞれ背景映像と入力映像であり、

二つの映像間で差分を取ることで、図4.27(c)の動領域を抽出する。図4.27(d)は隠れ処理を含まないステレオモジュールすなわちビデオレートステレオマシンで獲得したシーン全体の奥行きマップである。また、図4.27(e)は、隠れ処理を含んだステレオモジュールによって静止領域の奥行きマップを獲得した結果である。二つのステレオモジュールから得られた奥行きマップを動領域データをもとに統合したのが図4.27(f)の奥行きマップである。

## 4.7 考察

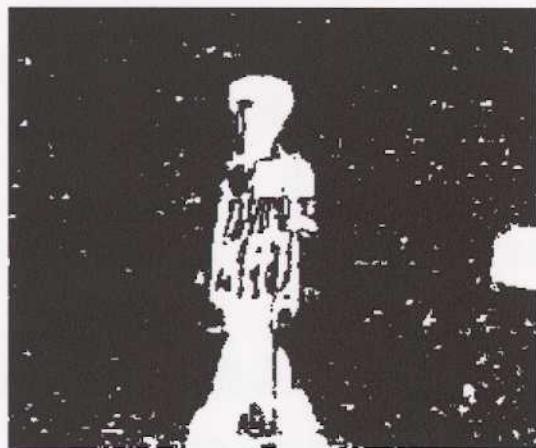
本章では、隠れのない領域と隠れ領域の両方の領域で精度の良い奥行き獲得を実現するため、隠れ領域を抽出し、隠れのない領域と隠れ領域のそれぞれに対して有効な異なるステレオアルゴリズムを適用し、統合する手法を提案した。隠れ領域の抽出手法として、隠れ処理を含まないステレオ処理の際に得られる各仮定視差における評価値を再利用した手法を提案し、静止画像を用いてその有効性を検証した。また、各ステレオ画像対で生じる隠れのパターンに依存しない隠れ処理アルゴリズムを提案することで、複雑な隠れが生じている場合にも精度の良い奥行きマップが得られることを確認した。これらの提案したアルゴリズムをビデオレートステレオマシンとPCを用いて実装することで、提案手法のビデオレートステレオマシンとの親和性を実証した。今後は、隠れ領域抽出アルゴリズムを実装したビデオレートステレオシステムの構築、隠れ処理を含んだステレオアルゴリズムの高精度化、Intel Pentium プロセッサのMMX命令を使用した高速化などの課題がある。



(a) 背景映像



(b) 入力映像



(c) 動領域抽出結果



(d) 隠れ処理を含まないステレオ  
モジュールからの奥行きマップ



(e) 奥行きマップ



(f) 隠れ処理を含んだステレオ  
モジュールからの奥行きマップ

図 4.27: ビデオレートステレオシステムによる奥行き獲得結果