

NAIST-IS-MT9851119

修士論文

複数の全方位画像センサを用いたステレオ視による 3 次元モデルの作成

山下 敏行

2000年2月14日

奈良先端科学技術大学院大学
情報科学研究科 情報システム学専攻

本論文は奈良先端科学技術大学院大学情報科学研究科に
修士(工学)授与の要件として提出した修士論文である。

山下 敏行

審査委員： 横矢 直和 教授
小笠原 司 教授
竹村 治雄 助教授

複数の全方位画像センサを用いたステレオ視による 3 次元モデルの作成*

山下 敏行

内容梗概

近年, 仮想現実感や複合現実感の研究において実環境の構造を仮想空間へ取り込む要求が高まっている.

実環境から 3 次元情報の獲得のための手法としてステレオ視による方法が知られている. しかし, 通常のカメラによるステレオ視は, カメラの撮影可能領域の問題から一度に広い領域を 3 次元モデル化するには多くの画像を必要とする.

本論文では屋内環境全体のような広い範囲を 3 次元モデル化することを目的とし, 全方位画像センサを用いたステレオ視による 3 次元モデル作成の手法を提案する. 提案手法は全方位画像センサでありながら透視投影の性質を持つ HyperOmni Vision を 3 台使う. まず, 3 台のセンサを直線上等間隔に離して水平に配置し周囲の環境を撮影する. 次に 3 つの入力画像を円柱面へ投影する. 投影された 3 つの円柱画像のうち中央を基準画像として, 基準画像に対してエッジ導出を行ない, このエッジの上でほかの二つの円柱画像からステレオ三眼視により視差を求める. 求めた視差から 3 次元座標を導出し, あり得ない座標値を持つ点を削除した上で三角パッチを作り 3 次元モデルとして表示する.

実験により, 実際に実験室内を 3 次元モデル化して有効性を示した.

キーワード

全方位画像センサ, HyperOmni Vision, ステレオ視, 3 次元モデル, 円柱画像

*奈良先端科学技術大学院大学 情報科学研究科 情報システム学専攻 修士論文, NAIST-IS-MT9851119, 2000 年 2 月 14 日.

A 3D-model reconstruction of a real environment by using a stereo vision with omnidirectional image sensors*

Toshiyuki Yamashita

Abstract

Recently, methods for generating a virtual environment of the real world are required in virtual/mixed reality.

Stereovision is known as one of such methods for obtaining 3-D map. But a lot of images are needed for virtualizing the real environment when conventional stereo camera is used.

In this paper, we propose a method of Stereovision using omnidirectional image sensors for 3-Dmodel reconstruction of the indoor environment. In this proposal, we use the three HyperOmni Visions of which optical feature is perspective. At first, we take a images from three HyperOmni Visions. And we project omnidirectional input images onto cylindrical surface of which axis is on the base line for stereovision. We apply the edge filter to the center cylindrical image, named base image. Then,we detect corresponding points from the two pairs of the cylinder images and obtain 3D information simulataneously in whole of indoor environment. At last, we generate the 3D model with using this information.

In practical experiment, we confirm 3-D model is reconstructed by using our method.

Keywords:

Omnidirectional image, HyperOmni Vision, Stereo vision, 3D-model, cylindrical image

*Master's Thesis, Department of Information Systems, Graduate School of Information Science, Nara Institute of Science and Technology, NAIST-IS-MT9851119, February 14, 2000.

目次

1. はじめに	1
1.1 従来手法	1
1.2 本研究の方針	9
2. 全方位画像センサ HyperOmni Vision	11
2.1 HyperOmni Vision の構成	11
2.2 HyperOmni Vision の特徴	14
3. 3次元モデルの作成	15
3.1 全方位画像から円柱面への投影	16
3.2 エッジの導出	18
3.3 ステレオマッチング	19
3.3.1 距離情報および3次元位置の導出	21
4. 実験	21
4.1 実験環境	21
4.2 実験結果および考察	23
4.2.1 円柱画像	23
4.2.2 視差画像	23
4.2.3 3次元モデル化	25
5. おわりに	26
謝辞	28
参考文献	29

目 次

1	標準ステレオ撮像系	3
2	六角錐ミラーを用いた全方位画像センサ	6
3	全方位画像センサの視野領域	8
4	HyperOmni Vision の外観と視覚系の構成	11
5	点の射影水平方向	12
6	点の射影鉛直方向	13
7	全方位画像	14
8	全方位画像と平面透視投影画像の関係	15
9	円柱面への投影	17
10	Sobel Filter によるエッジ導出	19
11	ステレオマッチング	20
12	入力画像	22
13	円柱画像	23
14	円柱面へ投影したエッジ画像	24
15	視差画像	24
16	視差と色の関係	24
17	3次元座標変換	25
18	作成された3次元モデル	26

表 目 次

1	Sobel エッジオペレータ	19
2	各パラメータの値	22

1. はじめに

近年, 仮想現実感や複合現実感の研究において実環境を仮想空間へ取り込む要求が高まっている. 屋内環境の3次元モデル化は, 動物体の場所を視覚的にわかりやすく表示するなどの警備システムへの応用や, 仮想空間中をウォークスルーしたり, 部屋に仮想的にものをおいたりするシミュレーションなどに利用できる. このとき, 実写画像を用いた広範囲の3次元モデルがよく使われ, 奥行き再現や写実性の特徴を持つ. この3次元モデルを作成するためには, センサの周囲360度について3次元情報の獲得が不可欠である. 全方位の3次元情報の獲得するには, 各方向で順に情報を獲得していくことになる. しかし, この方法では, 計測時のカメラの姿勢や位置を測定し, この位置情報と得た3次元情報から周囲の3次元情報を再構築しなければならない. よって, それぞれの位置の計測誤差の問題や計測に手間がかかるなどの欠点がある. これを解決するため, 本論文は屋内環境全体を3次元モデル化するために周囲360度の3次元情報を一度に取得する手法を提案する. 第1.1節では, 従来の全方位の3次元情報の獲得のための従来手法とその前提となる研究について, また, 第1.2節で本手法の方針を述べる.

1.1 従来手法

実環境の3次元モデル作成する際, 環境中の対象物体の3次元情報獲得が不可欠である. その手法として様々なものが提案されている. 以下, 全方位の3次元情報獲得の従来手法についての関連研究を述べる.

3次元情報獲得の従来手法

全方位に限らず, 3次元情報の獲得のための手法として, 能動的手法と受動的手法がある.

能動的手法とは環境中の観測対象に光などのエネルギーを照射したり, 光源を制御しながら取得した画像から3次元情報を獲得する. この手法では高速に高精度で密度の高いデータが得られるのが特徴であり, 光レーダ法, 光投影法, モアレ法, 照度差ステレオ法に分類できる.

光レーダ法は光を対象物体に照射してその位相差/照度差を計測することにより距離を計測する。光投影法は、スリット光やスポット光を対象物体に照射しそれを別の角度からカメラによって撮影する手法である。モアレ法は、対象物体に格子縞を投影し、それを別の格子縞を通して観測してモアレ縞を作ることによって計測する。照度差ステレオ法は光源の位置を変化させて同一の対象を同一の位置から撮影した複数の画像から物体面上の各点での面の方向を求める。

光投影法以外の手法では環境中の照明条件を変化させたり限定させる必要があるため、屋内環境全体のような環境の計測を行なうには環境中の光をすべて遮るという作業が必要で手間がかかる。また、能動的手法全体にいえるが、対象物体が光を吸収しやすい材質の場合、3次元計測は困難となる。さらに通常のカメラを使う場合、カメラの撮影可能領域の問題から計測範囲が狭い。

受動的手法とは得られるカメラの画像とその画像が撮影された時の条件から制約条件をもとに画像生成仮定の逆問題を解くことで3次元情報を獲得する。この手法では自然光での計測が可能であることが特徴である。この手法には、単眼視やステレオ視がある。

単眼視はひとつの画像か光学的性質、幾何学的性質、撮影環境に関する付加的情報から距離を推測するステレオ視 [1] は対象物体を異なる場所で撮影することによって生じる対象物体が写る場所のずれを検出して三角測量の原理で距離を計算する。

受動的手法の場合、画像から距離を導出するため、通常のカメラでは能動的手法と同様にカメラの撮影可能領域の問題がある。単眼視では対象表面の光学的性質、幾何学的性質、撮影環境に関する付加的情報が必須であり画像単体から距離情報を得ることは困難である。ステレオ視はずれを検出するために計測対象が各画像のどの位置に写っているかの対応問題を解く必要がある。そのために、エッジやテクスチャのような目立った特徴のあるところでしか距離は得られにくい。また、計算時間がかかる。

ステレオ視

能動的手法では環境中の物体の材質によって計測できないものがある. よって, カメラの撮影可能領域の制限がなければステレオ視が屋内環境全体の 3 次元モデルの作成に適しているといえる. ここでは, まず, ステレオ視について詳しく述べる.

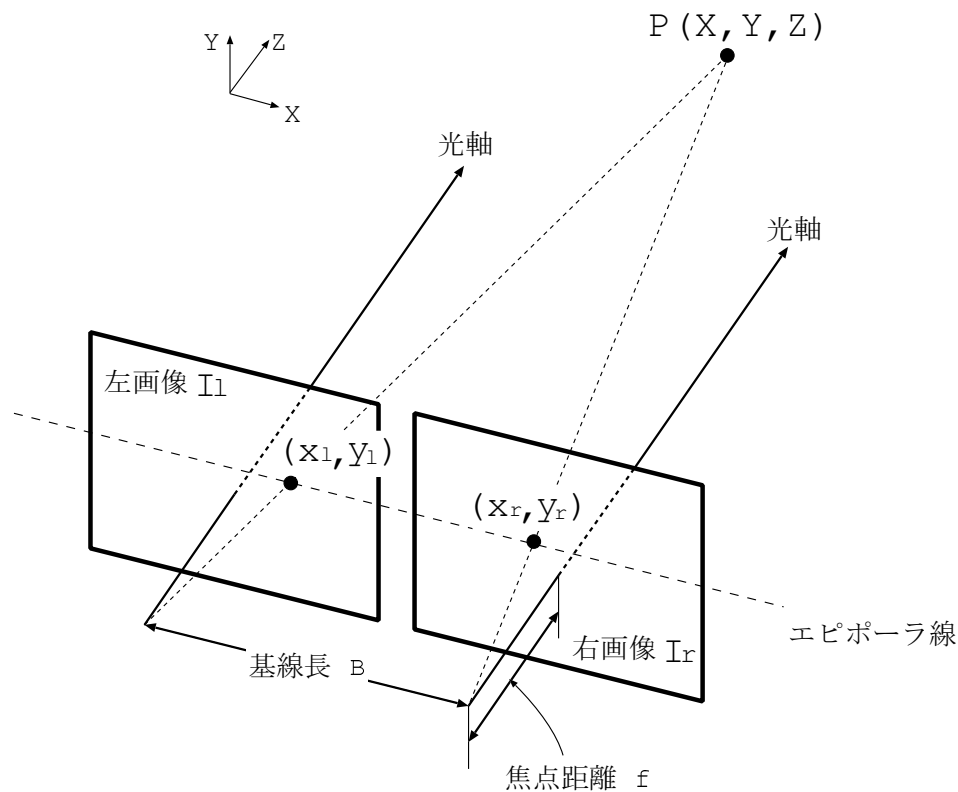


図 1 標準ステレオ撮像系

この手法は二つのカメラの方向, カメラ間の距離がわかれば対象物体までの絶対距離が検出できる. また, 自然光での距離検出ができることから屋内・屋外を問わず利用できるという特徴がある. 図 1 に示すようにふたつのカメラの光軸が平行な標準ステレオ撮像系を考える. 対象となる 3 次元環境は焦点距離距離 f で画像平面に一点透視投影される.

左右のカメラの中心を結ぶ線分の長さ基線長 B を考えるとこの線分を含む任意の平面をエピポーラ面, エピポーラ面と画像平面の交線をエピポーラ線という. ここで 3次元環境座標系の原点を二つのカメラレンズ中心の midpoint におき, X 軸を基線と平行な方向, Y 軸を光軸と基線に垂直な方向, Z 軸を光軸と平行な方向にとると 3次元環境中の点 $P(X, Y, Z)$ が左右画像平面の $(x_l, y_l), (x_r, y_r)$ に現れたとすると, 点 P の座標 (X, Y, Z) は式 (1) で与えられる. このとき, 標準ステレオ撮像系では, 右画像の注目点 (x_r, y_r) に対する左画像の対応点 (x_l, y_l) の位置は, 右画像上で見える位置よりも必ず右方向にあり $(x_l > x_r)$, 対応するエピポーラ線上にある $(y_l = y_r)$. これをエピポーラ拘束と呼ぶ. 注目点と対応点間の位置ずれ $x_l - x_r$ を両眼視差と呼ぶ.

$$\left. \begin{aligned} X &= B \frac{(x_l + x_r)/2}{x_l - x_r} \\ Y &= B \frac{(y_l + y_r)/2}{\frac{x_l - x_r}{y_r}} \\ &= B \frac{y_r}{x_l - x_r} \\ Z &= B \frac{f}{x_l - x_r} \end{aligned} \right\} \quad (1)$$

式 (1) に示したように, 3次元環境中の点 $P(X, Y, Z)$ は, 左右のステレオ画像に対する各写像点の座標から求めることができる. よって, 両眼ステレオ視により 3次元位置を求めるためには, 右の画像中の点が左画像のどこに対応しているかを求める対応点探索を行なわなければならない. 基線長 B ならびに焦点距離 f は既知であるとする. エピポーラ拘束が成立しているため, 右画像の注目点 (x_r, y_r) に対する左画像の対応点 (x_l, y_r) は, エピポーラ線上である矩形小領域 (探索窓と呼ぶ) でテンプレート・マッチングを行なうことにより求めることができる. その主な手法として最小自乗誤差和を用いる方法と正規化相互相関を使う方法がある.

最小自乗誤差和を使う方法 [16] ではエピポーラ線上で下式 (2) を計算し, この最小自乗誤差和 (SSD: Sum of Squared Differences) の値が最小となる部分を対応点とすればよい. ここで, $I_l(i, j), I_r(i, j)$ は左右画像上の点 (i, j) での輝度値とし, 探索窓を W とする.

$$SSD(x_r, y_r; x_l) = \sum_{i,j \in W} \{I_l(x_l + i, y_r + j) - I_r(x_r + i, y_r + j)\}^2 \quad (2)$$

この方法では、二つの画像を撮影したカメラの感度が違う場合、正確なマッチングができない可能性がある。

正規化相互相関 [2] とはふたつの画像上のピクセル (i, j) の輝度を $I(i, j)$, (i, j) を中心とする探索窓内の全ピクセル値を成分とするベクトルを $V(i, j)$ とする。ここで、右画像の点 (a, b) と左画像での点 (c, d) でのこのベクトルを $V(a, b)$, $V'(c, d)$ とし、次の相関値 $Corr$ を考える。

$$Corr = \frac{V(a, b) \cdot V'(c, d)}{|V(a, b)| |V'(c, d)|} \quad (3)$$

この相関値は二つの画像の探索窓 W 内の成分によるベクトルの角度の余弦である。よって、二つのベクトルの方向が同じ方向であるほど $Corr$ の値は 1 に近づく。つまり、 $Corr$ が 1 に近い ($Corr \leq 1$) と、探索窓内のピクセルの相関度が高いということになる。この方法ではふたつのベクトル相関値を比べるため、二つの画像を撮影したカメラの感度が違って精度の高いマッチングができる。

しかし、どちらの方法もピクセル間に相関が見つかりやすい画像中のエッジやテクスチャのはっきりした部分でしか正確な対応を求めることができないことが多い。しかし、実環境の 3 次元モデルでは、環境中にある物体の表面上の点全点で 3 次元情報が必要であるわけではなく、各物体の面の特徴的な部分たとえば物体の角での奥行きがわかるだけで良い。

全方位画像センサ

画角の制限からステレオ視の計測範囲は屋内環境全体を取り込むには狭かった。これを解決するために観測位置において周囲 360 度の画像を同時に取得できる全方位画像センサが利用した、全方位ステレオと呼ばれる手法がある。ここでは、ステレオ視に利用できる全方位画像センサについて考察する。

全方位画像センサはロボットビジョンなどの分野で周囲 360 度の環境を知ることのできる視覚センサとして利用され、さまざまな手法が提案されている。たと

えば、複数のカメラを用いて同時に撮影してつなぎあわせる方法 [2], 球面や双曲面などのミラーを利用する方法 [3, 4, 5, 6] などがある.

複数のカメラを使う方法には川西の手法 [2] のように複数の鉛直方向のカメラをピラミッド型のミラーの各側面に反射させて撮影する方法がある (図 1.1). この方法は複数のカメラからの画像をつなぎ合わせる際に各カメラのパラメータの違いから輝度や焦点が微妙にずれて不連続になることがあり, それを吸収するカメラのキャリブレーション技術の確立が不可欠である.



図 2 六角錐ミラーを用いた全方位画像センサ

球面や双曲面のミラーに反射させる方法 [3, 4, 5, 6] にはいくつか方法があるが一般的に鉛直上向きカメラで鉛直下向きのある形をしたミラーに周囲の環境を反射させて撮影するものである. 代表的なミラーの形には球面・円錐・双曲面ミラーがある. それぞれのミラーの形とその視野領域について以下のような特徴がある.

- 球面ミラー [3]

図 3(a) のように球面ミラーを鉛直下向きに取り付け, 鉛直上向きカメラで撮影する. この方法では球面の下方半球に写り込む部分を同時に撮影する. 視野領域は, 球面の特性上, カメラ自身の写り込みや床面を多く含むので環境を理解する上で重要な側方領域があまり得られない.

- 円錐ミラー [4]

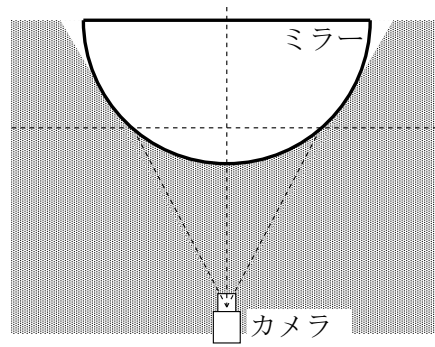
図 3(b) からわかるように鉛直下向きの円錐ミラーと鉛直上向きにとりつけられたカメラで撮影する。視野領域は側方領域の情報を多く含んだ画像が得られるが逆に足元の情報を得るにはミラーとレンズの関係を十分に考慮しなければならない。

- 双曲面ミラー [5]

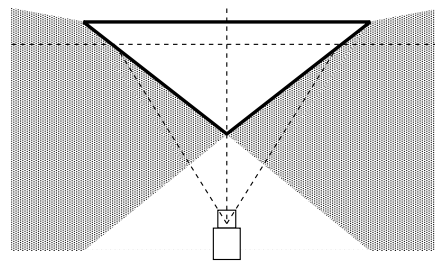
図 3(c) からわかるように鉛直下向きの双曲面ミラーに鉛直上向きに取り付けられたカメラで構成される。この手法の視野領域は側方領域も足元の情報もバランス良く含まれる。

曲面ミラーに反射させる方法では環境全体をひとつのカメラで撮影するため、あまり高い解像度が得られないが、全方位の画像を一度に撮像できる。球面ミラーもしくは円錐ミラーを使う方法は光学系の特性が透視射影でなく、双曲面ミラーを用いる方法では透視投影の光学特性を持つ。

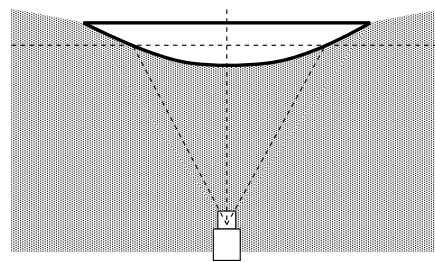
以上のことから、ステレオ視を使うためには、光学特性が透視投影である全方位画像センサが都合が良い。よって、複数のカメラを使う手法か双曲面ミラーに反射させる方法が適しているといえる。



(a) 球面ミラー



(b) 円錐ミラー



(c) 双曲面ミラー

■ : 視野領域

図 3 全方位画像センサの視野領域

全方位ステレオ

前述の全方位画像センサを用いたステレオ視の従来手法として複数のカメラを使う方法による手法 [2] もしくは, 双曲面ミラーに反射させる方法による手法 [5] がある. 複数のカメラを使う方法による手法 [2] では周囲の環境を六角錐ミラーに反射させて 6 つのカメラで撮影するものを上下 2 つ組み合わせて上下のペアでステレオ視を行なうものである.(図 1.1) この手法では高解像度の画像が得られる. しかし, 12 個のカメラの焦点を手動で合わせることは難しい. その他, 同機種のカメラを使ってもカメラ毎に感度が微妙に違うため, 画像間の補正が必要である.

双曲面ミラーを使う方法による手法 [5] では双曲面ミラーを使った全方位画像センサである **HyperOmni Vision** を垂直に 2 台配置してステレオ視を行なう. この場合, エピポーラ線はカメラの中心から放射線状にエピポーラ拘束を成立させることができ対応点探索を直線上孤立的に行なうことができる. 屋内環境では天井の高さの制限から, ある程度広い屋内環境だと基線が短過ぎて部屋全体の取り込みは困難となる. また, 縦方向にセンサの数を増やすことで精度をあげることも同様な理由から難しくなる.

1.2 本研究の方針

全方位ステレオの従来研究では次のような問題がある. 複数のカメラを使った方法 [2] では, 複数の画像センサを使用しているためカメラの各パラメータの手動補正などの問題があった. また, 双曲面ミラーを使った全方位画像センサを垂直に配置する方法 [7] では, 天井の高さの問題から基線長やセンサの数が制限された.

そこで, 本研究では双曲面ミラーを用いた全方位画像センサ **HyperOmni Vision** を水平に配置することによりステレオ視を行なうことを提案する.

本手法では, 全方位画像一枚を得るためのカメラは一つであるので, 画像統合などをする必要がない. また, 水平に配置することにより, 環境中のものの高さに制限されることなく屋内環境の広さに合わせて基線の長さを決定できる. また, 本手法ではセンサの数を 3 つにして 3 眼視を行なうことでさらに精度をあげる.

本論文では以下, 第 2 章では本研究で用いた **HyperOmni Vision** について述べる. 第 3 章では本手法について詳細に説明する. 第 4 章で実際に行なった実験とそ

れについての考察を述べる. 第5章では本論文のまとめと今後の課題について述べる.

2. 全方位画像センサ HyperOmni Vision

本章では, 本研究で使用した全方位画像センサ HyperOmni Vision[8] について述べる. まず最初に第 2.1 節で本センサの構成を示し, 次に 2.2 節で本センサの特徴について簡単に述べる.

2.1 HyperOmni Vision の構成

全方位画像センサ HyperOmni Vision は, 図 4(a) に示すように鉛直下向きに設置した二葉双曲面ミラーとその下に鉛直上向きに設置した CCD カメラから構成される. HyperOmni Vision の視覚系の構成を図 4(b) に示す. ここで, ミ

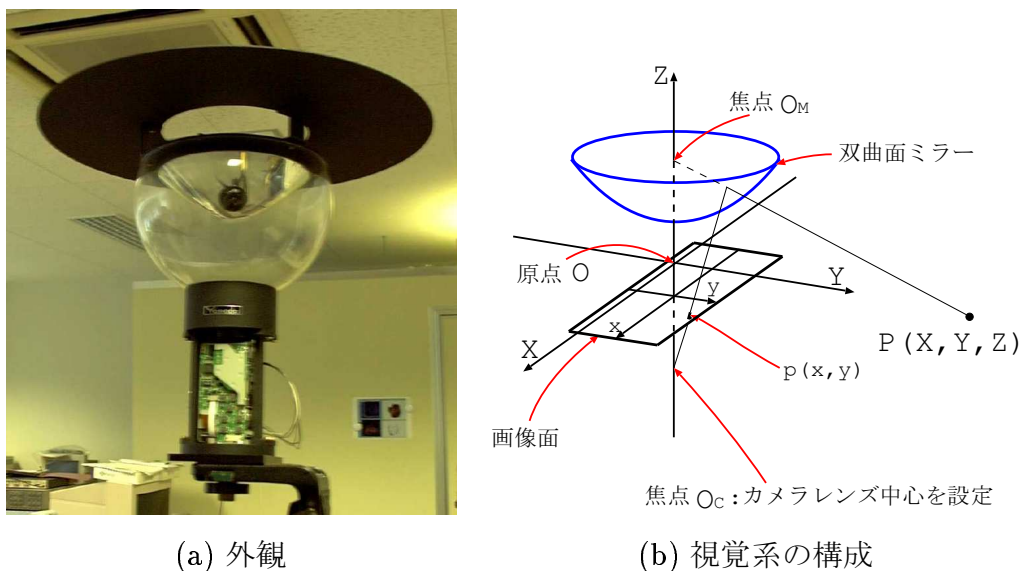


図 4 HyperOmni Vision の外観と視覚系の構成

ラーの焦点 O_M 及びカメラのレンズ中心 O_C は各々二葉双曲面が持つ二つの焦点 $(0, 0, +c), (0, 0, -c)$ に位置し, 画像面 xy は XY 平面に平行でカメラのレンズ中心 O_C からカメラの焦点距離 f だけ離れた平面とする. ミラー面とその焦点 O_M , カメラのレンズ中心 O_C は次式で表わされる.

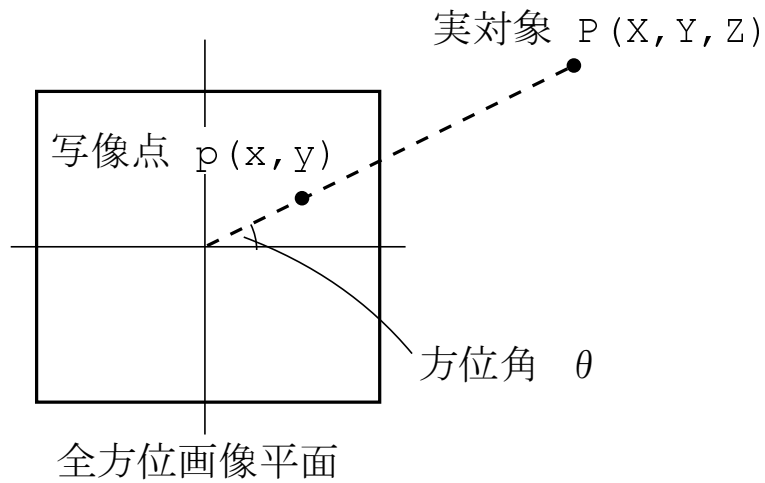


図 5 点の射影水平方向

$$\begin{aligned}
 \text{ミラー面} & \quad \frac{X^2 + Y^2}{a^2} - \frac{Z^2}{b^2} = -1 (Z > 0) \\
 \text{ミラーの焦点 } O_M & \quad (0, 0, +c) \\
 \text{カメラのレンズ中心 } O_C & \quad (0, 0, -c) \\
 (c = \sqrt{a^2 + b^2}) &
 \end{aligned} \tag{4}$$

ここで a , b は双曲面の形状を定義する定数である。

空間内の任意の点 $P(X, Y, Z)$ に対する画像上での写像点を $p(x, y)$ とした時、点 P の方位角 θ は次式で表わされる。

$$\tan \theta = Y/X = y/x \tag{5}$$

即ち Y/X で定まる P の方位角 θ は、図 5 に示されるように、 y/x で定まる写像点 p の方位角 θ を算出することで得られる。この様に 360 度パノラマ状の領域内にある対象物体の方位角 θ が、その物体の画像面上の写像の方位として直接現われる。

図 6 は点 P と Z 軸を含む画像センサの鉛直断面である。ここで、点 P と写像

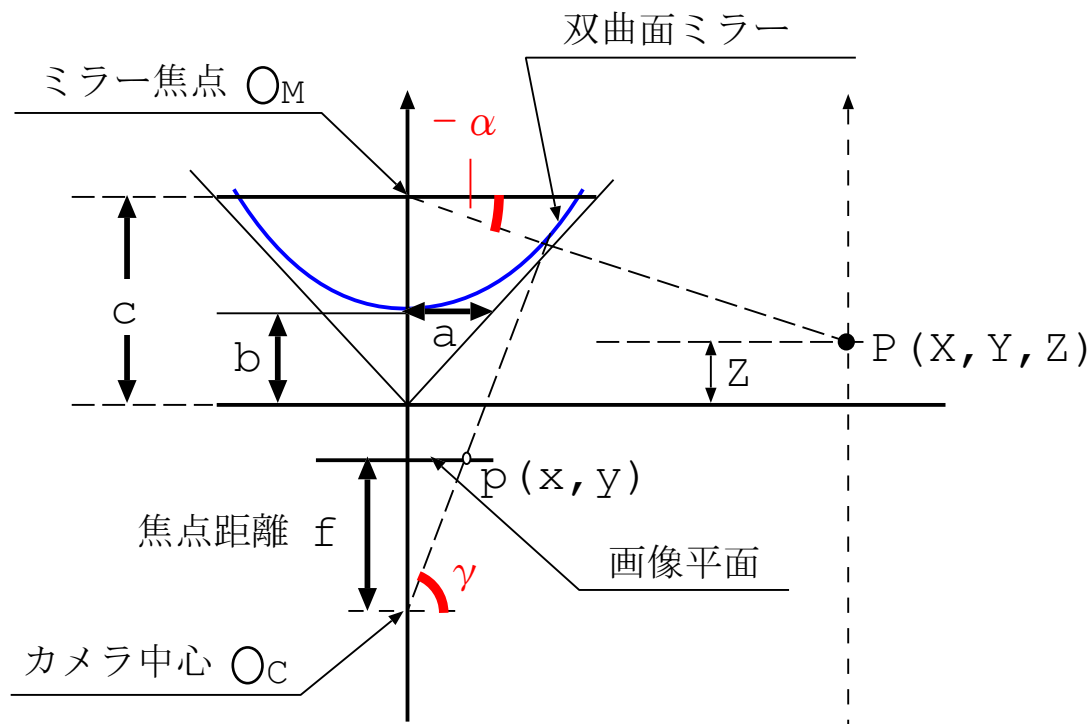


図 6 点の射影鉛直方向

点 p の間には次式の関係が成り立つ.

$$\begin{aligned}
 Z &= \sqrt{X^2 + Y^2} \tan \alpha + c \\
 \alpha &= \tan^{-1} \frac{(b^2 + c^2) \sin \gamma - 2bc}{(b^2 - c^2) \cos \gamma} \\
 \gamma &= \tan^{-1} \left(\frac{f}{\sqrt{x^2 + y^2}} \right)
 \end{aligned} \tag{6}$$

即ちミラーの焦点 O_M からの点 P の方位角 θ 及び仰角 α は、写像点 $p(x, y)$ より一意にもとまる. よって 3 次元環境中の点 $P(X, Y, Z)$ と全方位画像上の写像点 $p(x, y)$ の間には次式の関係が成立する.

$$\begin{aligned}
 x &= X \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}} \\
 y &= Y \times f \times \frac{(b^2 - c^2)}{(b^2 + c^2)Z - 2bc\sqrt{X^2 + Y^2 + Z^2}}
 \end{aligned}
 \tag{7}$$

式 (7) には三角関数が含まれておらず, これを使うことにより拘束に 3 次元環境中の点 $P(X, Y, Z)$ に対応する全方位画像上の点 $p(x, y)$ を求めることができる.

2.2 HyperOmni Vision の特徴

HyperOmni Vision は双曲面の焦点に向かう光はもう一方の焦点に向かって反射するという性質と, カメラをミラーのない方の焦点に置いてあるため, 得られる全方位画像は光学的に透視投影という特徴を持つ. よって, 幾何学的変換によりミラー焦点から見た透視投影画像に変換可能である. HyperOmni Vision を用いて得られる全方位画像の例を図 7 に示す.

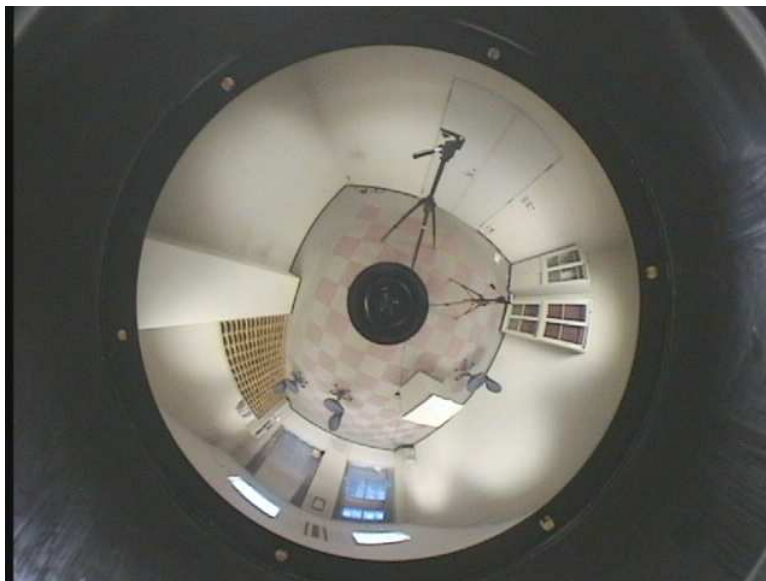
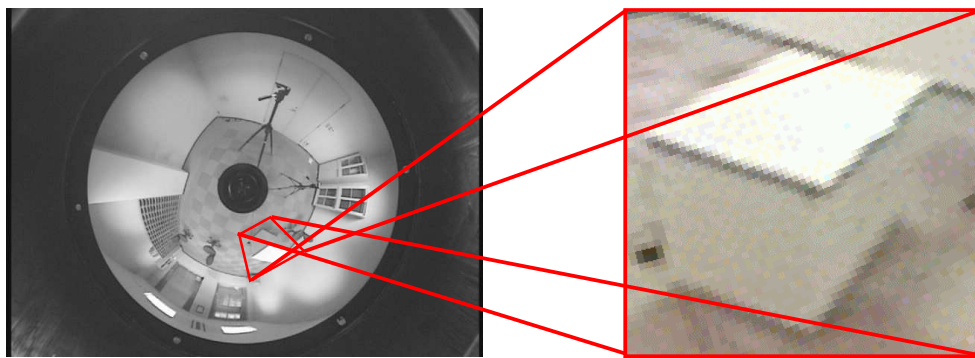


図 7 全方位画像

図 7 から全方位画像には周囲 360 度の情報が含まれていることがわかる. ここ

で図 6 よりミラーの焦点 O_M は固定なため HyperOmni Vision の入力画像は双曲面への 1 点中心投影画像であり, 図 8(a) に示すような全方位画像から図 8(b) に示すようなミラーの焦点 O_M から見た一般の平面透視投影画像に変換することが可能である.

ステレオ視には透視投影の光学特性が必要であるので本研究ではこの特性を用いてステレオマッチングを行なうことによりステレオ視を行なう.



(a) 全方位画像

(b) 平面透視投影画像

図 8 全方位画像と平面透視投影画像の関係

3. 3次元モデルの作成

本章では提案手法である水平に並べた HyperOmni Vision を使った全方位ステレオによる 3 次元情報の獲得および 3 次元モデル化について述べる.

本手法ではまず, 同じ高さに調節した三つの HyperOmni Vision を水平に直線上に等間隔に並べて置き, 各観測地点での画像を撮影する. 中央を基準画像, 両端を左右参照画像とする. それぞれの入力画像を円柱面へ投影し, 基準円柱画像, 左右参照円柱画像を作成する. 次に基準円柱画像に対してエッジ導出を行い, エッジ上で対応点探索を行うことで視差を導出する. 求めた視差から基準画像の視点からの距離を導出し, 不自然に遠い点を削除した上で, 視差画像から三角パッチを作

成して、視差を 3 次元座標系に変換した後、円柱画像からテクスチャを張ることで三次元モデルを作成する。

以下、エッジの導出、円柱面への投影、対応点探索、3 次元情報の獲得および 3 次元モデル化について順に述べる。

3.1 全方位画像から円柱面への投影

2 台の HyperOmni Vision を水平に並べた場合、エピポーラ面は HyperOmni Vision の焦点を結ぶ直線の上にある。この場合、エピポーラ線は双曲面ミラーとエピポーラ面の交線を画像平面に投影したものになるので、曲線を描く。HyperOmni Vision からの画像はその構造から中心からの距離によって解像度がことなる。よって、この曲線上をブロックマッチングによるステレオ視を行なうことを考えると探索窓のサイズを場所によって変えなければならない。また、この曲線上の点を逐一求めることで計算量が多くなる。そこで、本手法では円柱面へ投影してエピポーラ線を直線化することを考える。エピポーラ線はエピポーラ面と画像平面の交線であるので、基線を軸とする円柱面へ投影すればエピポーラ線を直線化することができる。エピポーラ線は軸に平行に存在することから標準ステレオ撮像系と同様な大きさを固定した探索窓での対応点探索を直線上孤立的に行なうことができる。ここで円柱面の上に投影された画像を円柱画像と呼ぶことにする。

円柱画像の作成方法は次の通りである。まず、図 9 のように、円柱画像を解像度縦 $2h$ × 横 $2w$ 矩形画像と見る。円柱画像の対角線の中点を原点とする座標系 (I, J) を考えて J 軸は円柱の軸に平行、 I 軸は J 軸に垂直にとる。円柱画像上の座標 $p(i, j)$ の実環境上の座標 $p(X, Y, Z)$ は HyperOmni Vision の焦点とミラーの一番外側を結ぶ線分と焦点から鉛直に下ろした直線のなす角度を α_{max} 、投影する円柱の高さを cyl_h 、投影する円柱の半径を f_{cyl} とすると

$$\alpha = \frac{(j - h) \times \alpha_{max}}{h} \quad (8)$$

$$t = \frac{(i - w) \times cyl_h}{w} \quad (9)$$

$$X = t \quad (10)$$

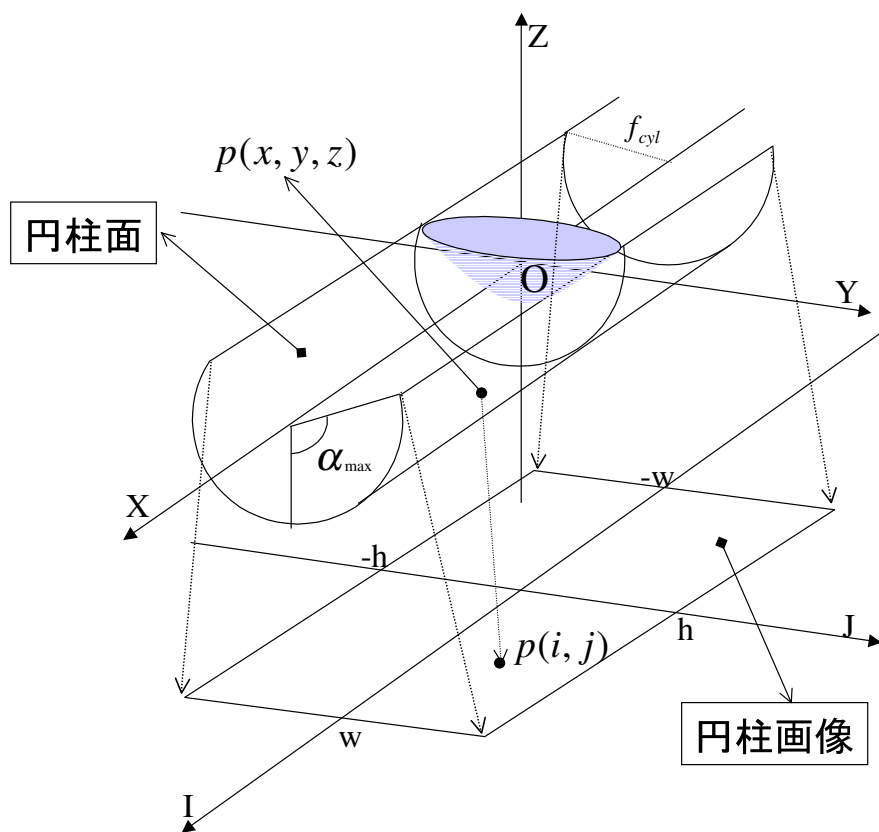


図 9 円柱面への投影

$$Y = f_{cyl} \sin \alpha \quad (11)$$

$$Z = f_{cyl} \cos \alpha \quad (12)$$

$$-h < i < h, -w < j < w$$

となる.これと式(7)により,入力画像上の対応するピクセルを求めることができる.ここで円柱画像の任意の点 $p(i, j)$ について α は, p を鉛直下向きを 0 度とするエピポーラ面と ZX 平面の角度であり, t は円柱画像上の J 成分である.

3.2 エッジの導出

本手法では 3 次元モデル化のための 3 次元情報の獲得が目的であるので,ステレオマッチングを行なう上で比較的精度の高いエッジ部のみで対応点探索を行なうことにする.

1.1章で挙げた対応点探索の手法は明るさが激変するエッジで高い精度でのテンプレートマッチングが行なえるから,画像の不連続点を検出する操作として画像の空間微分である差分型オペレータを使うことにする.画像 $f(x, y)$ の一次微分 ∇f とはベクトル

$$\nabla f(x, y) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right)$$

その大きさは

$$|\nabla f| = \sqrt{\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2}$$

となる.ここで ∇f をエッジ強度とする.

デジタル画像に対する微分は一意には決まらないが,本手法では最も簡単な sobel フィルタと呼ばれる近傍 3×3 の窓をもつ表 1 のような差分オペレータを適用する.

このオペレータを適用した結果の絶対値はエッジの縦方向のエッジ強度が,これを 90 度回転させたものではエッジの横方向のエッジ強度が求められる.縦横両方のエッジを求める時は二つの絶対値和がエッジ強度となる. (x, y) でのエッジ強度 $|\nabla f(x, y)|$ とし,ある閾値を B とするとエッジ画像 $f_e(x, y)$ は

-1	0	1
-2	0	2
-1	0	1

表 1 Sobel エッジオペレータ

$$f_e(x, y) = \begin{cases} 0 & (|\nabla f(x, y)| < B) \\ 1 & (|\nabla f(x, y)| \geq B) \end{cases}$$

となる.

本手法では基準円柱画像に対してエッジ導出を行う. このとき, エピポーラ線は I 軸に平行であるので, I 軸に平行なエッジ上よりより垂直なエッジ上の方がエピポーラ線上で対応点を求めるべき点同士の距離が遠くなる. これによって, それぞれの似たような点を探索することによる誤対応が減る. よって, 本手法では縦方向のエッジのみを導出する.

基準円柱画像に対する結果を図 10 にしめす.



入力画像



エッジ画像

図 10 Sobel Filter によるエッジ導出

3.3 ステレオマッチング

本研究では環境中の画像から 3 次元情報を獲得するために全方位画像に対して三眼ステレオ視を適用する. 前節で述べたように円柱画像に対して円柱の軸に平行な直線上で対応点探索を行なう.

三眼ステレオ視で正規化相互相関によるブロックマッチングを行うことを考える。このとき、三眼ステレオ視の正規化相互相関の相関値の定義を基準円柱画像と左右の参照円柱画像による二つのステレオペアで求めた正規化相互相関値の積であるとする。正規化相互相関による相関値は1に近いほど相関度が高いということになる。二つの相関度の和より積の方が、片方が小さいとき求める相関値はより小さくなる。よって、今回は積を相関度として採用した。

ここで、実際の対応点を求める手順は、基準画像と参照画像は等間隔に離れた場所で撮影していることから視差は二つのステレオペアで同じ値をもつ。このことから、二つのステレオペアで視差を順番に増やしていき相関度の一番高いときの視差を求める視差とする。

ステレオマッチングの手順を図 11 を使って説明する。まず基準円柱画像の抽出されたエッジ点 (i, j) を中心に基準探索窓を切り出す。次に左円柱画像上では $(i + d, j)$ 右円柱画像上では $(i - d, j)$ を中心にそれぞれ左右探索窓を切り出す。基準探索窓と左探索窓、基準探索窓と右探索窓の正規化相互相関値を求める。二つの相関値の積を評価値として d を変更しながら評価値が最大になる d を求める。抽出されたエッジ点に対してこれらの作業を繰り返す。

以上により精度の高いエッジ点のみの視差 d を求めることができる。

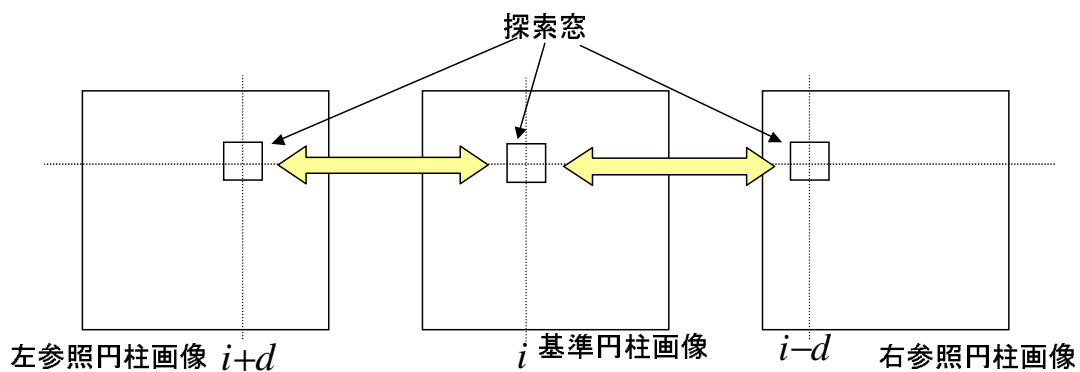


図 11 ステレオマッチング

3.3.1 距離情報および 3 次元位置の導出

第 3.3 節では, 円柱画像上の視差を求め, 対応点探索の手法を述べた. しかし, 実際の 3 次元情報は視差情報そのものではなく, 視差を元に対象物体の 3 次元上座標を計算する必要がある. 円柱画像でのステレオ視によって求めた視差 d と奥行き r , その点の実環境上の座標 $P(X, Y, Z)$ は, l_{base} を基線の距離, f_{cyl} を円柱の半径, c を双曲面ミラーのミラーパラメータとすると

$$r = \frac{l_{base} f_{cyl}}{d} \quad (13)$$

$$X = \frac{rt_1}{f_{cyl}} \quad (14)$$

$$Y = r \sin \alpha \quad (15)$$

$$Z = -r \cos \alpha + c \quad (16)$$

となる. これらの情報を元に 3 次元モデル化を行なう.

ここで, 一度各点の 3 次元座標を求め, 不自然に遠い点を削除する. 残った点に対して Delaunay の三角形分割法 [9] により, 三角パッチを作成する. この三角パッチと 3 次元座標からテクスチャを張った三次元モデルを作成する. その際, テクスチャ画像として基準円柱画像を利用する.

4. 実験

本章では提案した円柱画像によるステレオ視により観測環境中の点の 3 次元情報が獲得できることを確認する. そのために実画像を用いた実験を行ない, 得られた 3 次元情報の獲得結果から提案手法が有効であることを示す.

4.1 実験環境

実験に用いた画像を図 12 に示す. 図 12 は照明条件一定の条件の元で実験室内を撮影したものである. 画像サイズは 640×480 のカラー画像であり, 三脚を用いて水平方向に 500.0mm 離して配置してある. また, 三つとも三脚の高さは同じで床

a	42.1mm
b	42.7
c	59.1
焦点距離 f	435mm

表 2 各パラメータの値

から HyperOmni Vision のミラーの焦点までの高さが 1600mm である。このとき、お互いのカメラの方向さえわかれば良いので三脚を目印として、直線上を移動させて 3 箇所を撮影した。

実験において横 11 × 縦 7 の矩形領域の探索窓を用いた。対応点探索の範囲は直前にマッチしたピクセルから比較画像側の焦点を通るエピポーラ線と平行な直線と比較円柱画像の交点との間のエッジ上のみとする。3 次元情報を計算する際

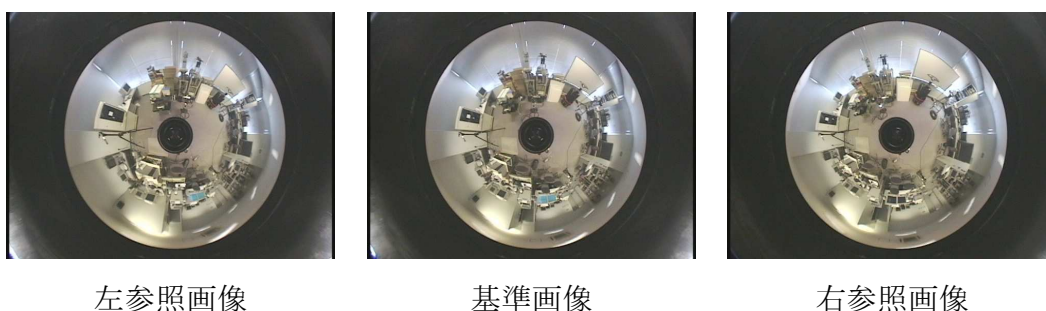


図 12 入力画像

に用いた HyperOmni Vision の双曲面ミラーの各パラメータ a, b, c および焦点距離 f を表に示す。

4.2 実験結果および考察

4.2.1 円柱画像

入力画像から円柱画像へ投影した結果を図 13に示す。円柱画像の解像度は横 551×縦 1801 である。これは、入力画像の縦の解像度が 480 となっているため、縦方向の縮小が起こらないように入力画像よりも多少大きめのピクセル数にしてある。さらに、画像のアスペクト比が変わらないように左右方向の解像度を決定している。円柱面へ投影する時、近傍 4 点で内挿によるピクセルの補間を行なっている。

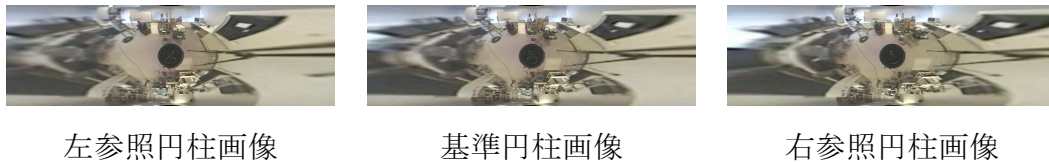


図 13 円柱画像

画像の左右両端でピクセルの拡大が起こるため、エッジがぼやけた感じになっていることがわかる。また、部屋にある同じ高さにあるものが 3 つの画像でほぼおなじ J 座標をとっていることからエピソード線の直線化に実際に行なわれていることがわかる。

4.2.2 視差画像

まず、基準円柱画像に対して Sobel フィルタを適用した結果を図 14に示す。この得られたエッジ上で得られた 3 枚の円柱画像を使って三眼視を行なった結果の視差画像を 15に示す。視差の値は図 16のように茶色が視差が大きく青色になるほど視差は小さくなる。この視差を求める際、実験環境中に存在し得ない点として基準画像の視点より水平に 10m、上下にそれぞれ 2m 以上はなれている点を削除した。

左右両端部分はエッジがぼやけているためエッジがでていない。このため、視差は左右両端約 $\frac{1}{3}$ 部分で導出されていないことがわかる。また、カメラが消えているが、これは視差が 0 もしくはそれに近い値で出ているため、あり得ないほど遠い

点と判定されて視差画像から消去されている. よって, ベースライン方向の視差はほとんど出ていない.

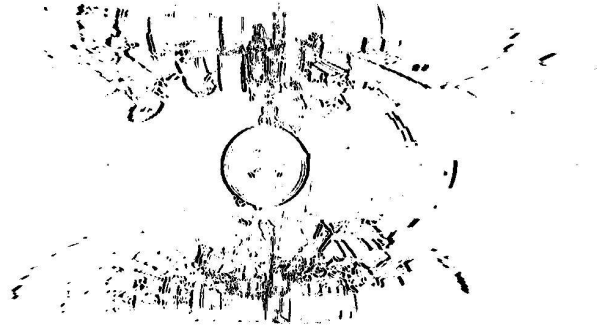


図 14 円柱面へ投影したエッジ画像

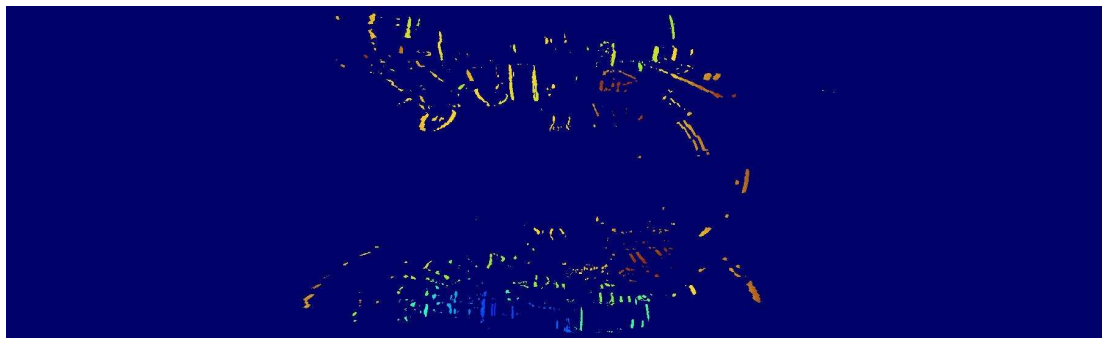


図 15 視差画像

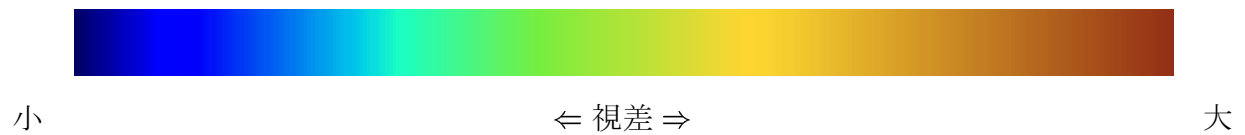


図 16 視差と色の関係

4.2.3 3次元モデル化

第3.3.1節で述べた視差と3次元情報の関係を用いて得られた視差から奥行きと高さを計算した結果を3次元座標形に直し,それをVRMLファイルとして出力して真上から見た図を図17に示す.この結果から実験室の形状にエッジ点が配置されていることがわかる.次にこれらの点を基に3次元モデルを作成する.まず,得られた視差画像に対してDelaunayの三角形分割法[9]を適用して三角パッチを作成する.この三角パッチと視差から得られる3次元座標をもとにテクスチャと点の対応をVRMLファイルとして出力した.この点群からもベースライン方向の座標値が出ていないことがわかる.図18に図17に示す視点位置からのモデルを見た図を示す.撮影された実験室中の奥行き関係を復元し,3次元モデルとして違和感のない表示が行なえている.

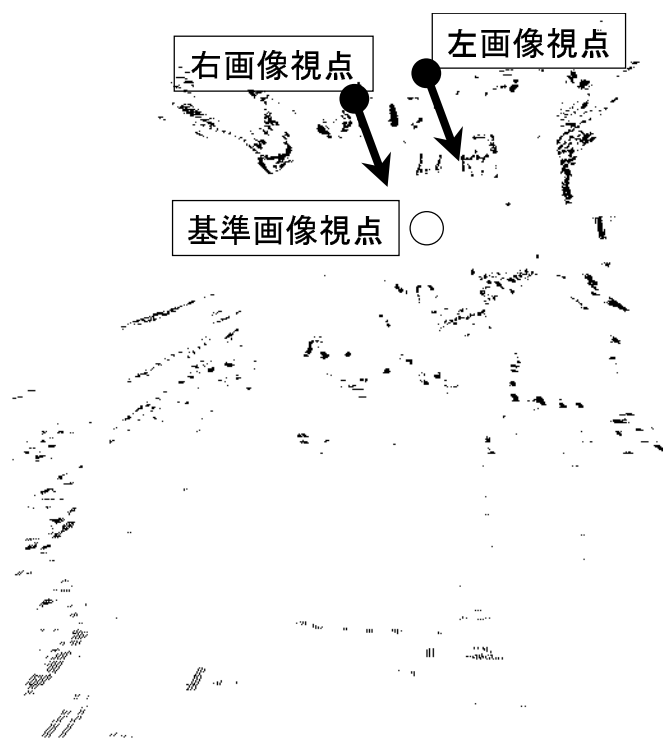
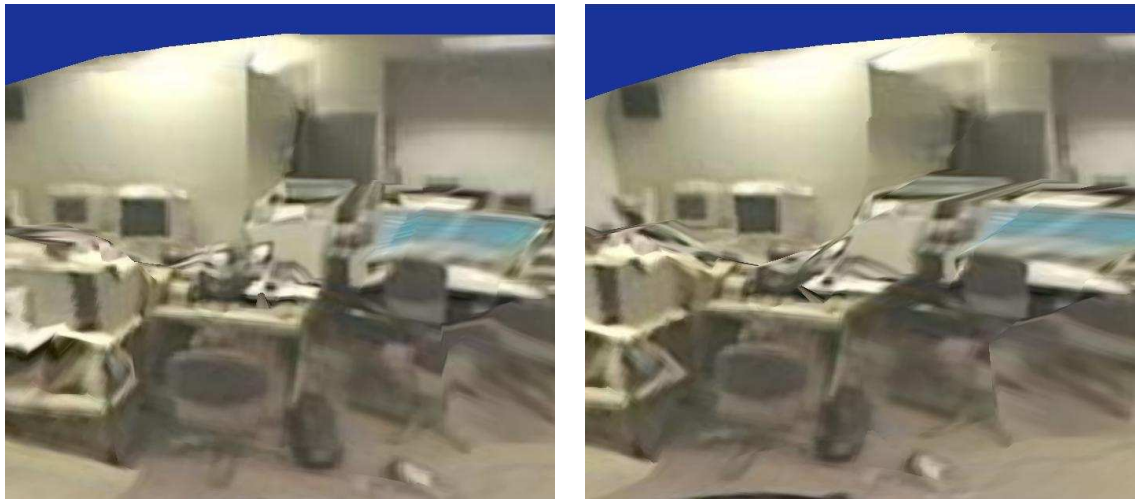


図 17 3次元座標変換



左画像

右画像

図 18 作成された 3 次元モデル

5. おわりに

本論文では, 屋内環境の 3 次元モデル化を行なうため全方位画像センサ Hyper-Omni Vision を用いたステレオ視による 3 次元情報獲得の手法を提案した. また, 実際に得られた 3 次元情報から 3 次元モデルを作成した.

本手法では実環境の水平に配置した全方位画像センサからの入力から円柱画像を作ってステレオマッチングを行い, 視差から周囲の 3 次元情報を獲得することができる. また, 得られた 3 次元情報から 3 次元モデルを作成・表示することができる.

また, この 3 次元モデルによって, 精度の高い点のみを選んで 3 次元情報を獲得できたことを示すことができた. これにより, テレプレゼンスや仮想物体との融合といった複合現実感への応用のために利用可能であると考えられる.

問題点は, 円柱面を作成する時の左右両端にあたる部分では入力画像が拡大されるため正規化相互相関によるブロックマッチングでは精度があげられない. また, ステレオ視の特性からベースライン方向のマッチング精度が上がらない. これらの問題に対応するため, センサの数を増やして基線を直線上だけでなく複数

の方向にとることによりこれらの問題に対応できると考える.

今後は上に述べた課題を解決し,HyperOmni Vision を使った動物体を検出する警備システム [10] と組み合わせて効果的な警備システムの構築を検討する.

謝辞

本研究を進めるにあたり, 終始ご指導いただいたソフトウェア基礎講座 横矢 直和教授に厚く御礼申し上げます。また, 本学での研究におけるあらゆる面で多大なご助言, 御協力賜りましたことを深く感謝します。

副指導教官としてご助言をいただいたロボティクス講座 小笠原 司 教授, ならびにソフトウェア基礎講座 竹村 治雄 助教授に深く感謝致します。

さらに, 本研究や研究環境の整備に御協力をいただいたソフトウェア基礎講座 山澤 一誠 助手, ならびにソフトウェア基礎講座 岩佐 英彦 助手に深く感謝します。

ソフトウェア基礎講座の諸氏には, 本研究を進めるにあたり多大なる協力をいただきました。ここに, 感謝の意を表します。

また, 日々の研究室の活動を支えて下さった, ソフトウェア基礎講座 福永 博美女史に感謝致します。

日頃, 励ましの言葉を与えてくれた高木 幹子女史に深く感謝します。また, その他応援して下さったみなさまに感謝いたします。

参考文献

- [1] 尾上守夫:“情報処理ハンドブック”, 昭晃堂,p.394-p.397,1987.
- [2] T. Kawanishi, K. Yamazawa, H. Iwasa, H. Takemura and N. Yokoya: “Generation of high-resolution stereo panoramic images by omnidirectional imaging sensor using hexagonal pyramidal mirrors”, *Proc of 14th Int. Conf. on Pattern Recognition*, Vol I, pp. 445-489, 1998.
- [3] J.Hong, X.Tan, B.Pinette, R.Weiss and E.M.Riseman: “Image-based Homing”, *Proc. IEEE Int. Conf. on Robotics and Automation*, pp.620-625, 1991.
- [4] 八木, 川戸: “円すいミラーを用いた全方位視覚センサによる位置情報の獲得”, 信学論 (D-II), Vol.J74-D-II, No. 1, pp.19-26, 1991.
- [5] 山澤, 八木, 谷内田: “移動ロボットのナビゲーションのための全方位視覚センサ HyperOmni Vision の提案”, 信学論 (D-II), Vol.J79-D-II, No. 5, pp.698-707,1996.
- [6] D. Southwell, A. Basu, M. Fiala and J. Reyda: “Panoramic Stereo”, *Proc. Int. Conf. on Pattern Recognition*, I, pp.378-382, 1996.
- [7] 茶園, 山澤, 横矢, 竹村: “HyperOmni Vision を用いた全方位ステレオ”, 画電学研報, 96-04-06, pp.21-24, 1996.
- [8] 山澤, 八木, 谷内田: “移動ロボットのナビゲーションのための全方位視覚系 HyperOmni Vision の提案”, 信学論 (D-II), J79-D-II, No.5, pp.698-707, 1996.
- [9] 谷口:“FEMのための要素自動分割-デローニー三角分割法の利用” pp.28-122, 森北出版,1992.
- [10] 三木, 横矢, 山澤, 竹村:“A Real-time Surveillance and Monitoring System Using Multiple Omnidirectional Video Cameras”, *Proc. of the Forth Asian Conf. on Computer Vision*,pp.528-534,2000.

- [11] 井口, 佐藤: 三次元画像計測, 昭晃堂, 1990.
- [12] Y.Ohta and T.Kanade: “Stereo by Intra- and Inter-scanline Search Using Dynamic Programming”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol.PAMI-7, No. 2, pp.139–154, 1985.
- [13] R.March: “Computation of Stereo Disparity Using Regularization”, *Pattern Recognition Letters*, Vol. 8, No. 3, pp.181–187, October 1988.
- [14] R.March: “A Regularization Model for Stereo Vision with Controlled Continuity”, *Pattern Recognition Letters*, Vol. 10, No. 4, pp.259–263, October 1989.
- [15] U.R.Dhond and J.K.Aggarwal: “Structure from Stereo – A Review”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 19, No. 6, pp.1489–1510, 1989.
- [16] J.S.Chen and G.Medioni: “Parallel Multiscale Matching Using Adaptive Smoothing”, *Proc. European Conference on Computer Vision*, pp.99–103, 1990.
- [17] 奥富, 金出: “複数の基線長を利用したステレオマッチング”, 信学論 (D-II), Vol.J75-D-II, No. 8, pp.1317–1327, 1992.
- [18] S.D.Cochran and G.Medioni: “3-D Surface Description from Binocular Stereo”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 14, No. 10, pp.981–994, 1992.
- [19] 横矢: “多重スケールでの正則化によるステレオ画像からの不連続を保存した曲面再構成”, 信学論 (D-II), Vol.J76-D-II, No. 8, pp.1667–1675, 1993.
- [20] 奥富, 吉崎, 富田: “カラーステレオマッチングとその視神経乳頭 3 次元計測への応用”, 信学論 (D-II), Vol.J76-D-II, No. 2, pp.342–349, 1993.

- [21] T.Kanade and M.Okutomi: “A Stereo Matching Algorithm with an Adaptive Window : Theory and Experiment”, *IEEE Trans. Pattern Anal. Machine Intell.*, Vol. 16, No. 9, pp.920–932, 1994.
- [22] J.L.Lotti and G.Giraudon: “Adaptive Window Algorithm for Aerial Image Stereo”, *Proc. 12th IEEE Int. Conf. on Pattern Recognition*, I, pp.701–703, 1994.
- [23] S. S. Intille and A. F. Bobick: “Disparity-Space Images and Large Occlusion Stereo”, *Proc. European Conference on Computer Vision*, pp.179–186, 1994.
- [24] F.Devernay and O.D.Faugeras: “Computing Differential Properties of 3-D Shapes from Stereoscopic Images without 3-D Models”, *Proc. IEEE CS Conf. on Computer Vision and Pattern Recognition*, pp.208–213, 1994.
- [25] L.Robert and R.Deriche: “Dense Depth Map Reconstruction : A Minimization and Regularization Approach which Preserves Discontinuities”, *Proc. European Conference on Computer Vision*, I, pp.439–451, 1996.
- [26] T.Kanade: “A Stereo Machine for Video-Rate Dense Depth Mapping and Its New Applications”, *Proc. Image Understanding Workshop*, II, pp.805–811, 1996.
- [27] 中山, 白井: “正規化した明度情報を利用したオクルージョン境界のステレオ対応付け”, *日本ロボット学会誌*, Vol. 14, No. 7, pp.1026–1032, 1996.
- [28] 茶園, 山澤, 横矢, 竹村: “HyperOmni Vision による全方位ステレオ視”, *信学技報*, IE96-122, pp.99–106, 1997.
- [29] T.Poggio, V.Torre and C.Koch: “Computational Vision and Regularization Theory”, *Nature*, Vol.317, No.6035, pp.314–319, September 1985.
- [30] 坂上, 横矢: “弛緩法と正則化”, *情処論*, Vol. 30, No. 9, pp.1047–1057, September 1989.

- [31] S.B.Kang and R.Szeliski: "3-D Scene Data Recovery using Omnidirectional Multibaseline Stereo", *Proc. IEEE CS Conf. on Computer Vision and Pattern Recognition*, pp.364-370, 1996.
- [32] 横矢:" 多重スケールでの正則化によるステレオ画像からの不連続を保存した曲面再構成", 信学論 (D-II),Vol.J76-D-II,No.8,pp.1667-1675,1993.
- [33] T.Kanade and M.Okutomi:"A Stereo Matching Algorithm with an Adaptive Window :Theory and Experiment",*IEEE Trans. Pattern Anal. Mach. Intell.*,Vol.16,No.9,pp.920-932,1994.
- [34] 井口, 佐藤:三次元画像計測, 昭晃堂,1990.
- [35] 茶園, 山澤, 横矢, 竹村:" 全方位ステレオ視による 3次元情報の獲得", 映像情報メディア学会誌, Vol.51, No.10, pp.1745-1753, 1997.
- [36] 守屋, 大場:" 全方位ステレオ画像から得られる距離画像情報の精度向上に関する研究", 信学技報, PRMU98-237, 1999