

論 文

連続ステレオ画像からの3次元情報の抽出

正員 山本 正信[†]

Determining Three-Dimensional Structure from Image Sequences
given by Horizontal and Vertical Moving Camera

Masanobu YAMAMOTO[†], Member

あらまし 両眼ステレオ視は、3次元情報の手軽な獲得法ではあるが、左右画像間の対応付けの際に、(1)多重対応、(2)隠れ、(3)順序の逆転、(4)水平エッジ、などがある場合に於て、対応点探索が困難になる。

本論文では、視点を移動させつつ得られた連続ステレオ画像から、シーンの3次元情報を推定する手法を提案する。まず、連続ステレオ画像から、見かけの運動軌跡が線分として画像化されるような2次元画像を合成する。その時、対応付け問題は、合成画像上の線分検出問題として簡略化される。検出された線分の相互関係を調べることにより、隠れの検出や、見かけの逆転が起きた場合の対応付けも容易になる。更に、視点を一方向だけではなく、それと直交する方向にも移動させることにより、エッジの向きに依存しない一意な対応付けが可能となる。大量の連続ステレオ画像が高速に入力可能な装置を構成し、実際に、複雑なシーンの3次元構造を復元する。

1. まえがき

3次元情報を抽出する数多くの手段のなかでも、ステレオ立体視は画像の入力が手軽なことから最も良く研究されている。

視点の間隔を小さく取る微少視差ステレオ^{(1),(2)}では、勾配法により、距離情報が算出できる。しかし、演算が容易である反面、測定精度に難点がある。

測定精度を向上させるためには、左右の視点の間隔を広げる必要がある。それに伴い、画像間の対応点探索が問題となる。対応の単位として、エッジなどの特徴点あるいはエッジで挟まれた区間を用いることが多い。粗密解析⁽³⁾や弛緩法⁽⁴⁾や動的計画法^{(5),(6)}は、数多くの対応付け可能な組み合わせのなかから、最も良い対応付けを効率的に探索してくれる。ところが、これらの探索アルゴリズムが適用可能であるためには、対応要素の順序が逆転しないことや、原則として隠れは起こらないことが、前提条件である。更に、両眼視の共通の問題として、エピポーラ線に平行なエッジ上では、一意な対応付けが難しい。そのため、対応単位の多様化⁽⁷⁾や画像モデルの利用⁽⁸⁾などが検討されている。

一方、左右の視点間に新たな視点をいくつか置くことにより得られる、多視点ステレオ画像の利用も進められている^{(9)~(13)}。多視点ステレオ画像によれば、左右画像間対応のあいまいさは、他の視点からの観察で、その真偽が試される。また隠れの検出や、順序が逆転している場合の対応付けも可能である。更に、上下にも配置された視点からの画像を利用することにより、水平エッジの対応も一意に定まる。多視点ステレオ画像の処理手法は、探索過程の繰返しを基本としているが、本論文では、視点の数を更に増すことにより、対応点の探索過程が不要となる方法を提案する。

多視点ステレオの視点を、隣接するステレオ対が微小視差ステレオとなるまで殖す。このようにして得られる多視点ステレオ画像を連続ステレオ画像とよぶ。連続ステレオ画像は、視点を視軸に垂直な方向に移動させて得られる動画像である。この動画像では、遠くの対象は緩やかに、近くの対象ほど速く移動する。運動軌跡の画像化手法^{(14)~(16)}によれば、対象上の点の運動軌跡は2次元合成画像上の線分軌跡として画像化される。従って、対応付け問題は、合成画像から線分を検出する問題に簡略化される⁽¹⁷⁾。また、検出された線分の相互関係を解釈することにより、隠れの発見や、順序の逆転が起った場合での対応付け也可能となる。更に、視点を直交方向にも移動させることによ

† 電子技術総合研究所制御部、茨城県

Automatic Control Division, Electrotechnical Laboratory,
Ibaraki-ken, 305 Japan

り、水平エッジの対応付けも一意に定まる。

このように、連続ステレオ画像に運動軌跡の画像化手法を適用すれば、多視点ステレオ画像の長所を保存したまま、対応付けの簡便化と高速化が可能となる。この方法は、物体の準静的回転⁽¹⁸⁾、あるいは視軸方向への視点の移動⁽¹⁹⁾に対しても拡張可能である。しかしながら、ステレオ視の利点の一つである、入力の手軽さが失われる。そこで、視点の精密な移動機構と、それに連動するビデオディスクを組み合わせた、連続ステレオ画像入力装置を構成した。この専用装置で得られた連続ステレオ画像から、実際にいくつかの複雑なシーンの3次元構造を復元し本手法の有効性を実証する。

2. 原 理

2.1 運動軌跡の画像化による3次元情報の抽出

シーンの座標 $O-XYZ$ を図1のように設定する。視点を X 軸上にとり、視軸を Z 軸の正方向とする。シーンは、 XY 平面に平行な面 $Z=F$ 上に、焦点距離 F を中心投影される。投影面の座標 $o-xy$ は、視軸との交点を原点とし、座標軸方向をシーンの座標軸方向と一致させる。今、視点を X 軸上に移動させつつ、微小移動ピッチ、 δx ごとに撮影する。得られた画像列を順番に並べ、図1のように3次元画像を構成する。 n 番目の視点位置の X 座標値を U とし、 $U=n \cdot \delta x$ とする。ただし、 n は整数。そのとき、対象面上の点 $P(X, Y, Z)$ は投影面上の点 $p(X', Y', F)$ に次式で写像される。

$$\frac{X-U}{X'-U} = \frac{Y}{Y'} = \frac{Z}{F} \quad (1)$$

k_x, k_y を投影面座標系のシーン座標系に対する x 方向、 y 方向の縮尺とすれば、

$$\begin{cases} x = k_x(X-U) \\ y = k_y Y' \end{cases} \quad (2)$$

(2)式を(1)式に代入し、 $U=n \cdot \delta x$ であることを考慮して整理すれば、

$$\begin{cases} \frac{Z}{k_x \cdot F} \cdot x = X - n \cdot \delta x \end{cases} \quad (3a)$$

$$\begin{cases} \frac{Z}{k_y \cdot F} \cdot y = Y \end{cases} \quad (3b)$$

(3b)式は、視点の移動につれ、点 P の投影点が描く運動軌跡が、3次元画像の水平断面画像上に拘束されることを示している。また(3a)式はその軌跡が直線であり、その勾配 $-Z/(k_x \cdot F \cdot \delta x)$ が深さ量 Z に比例

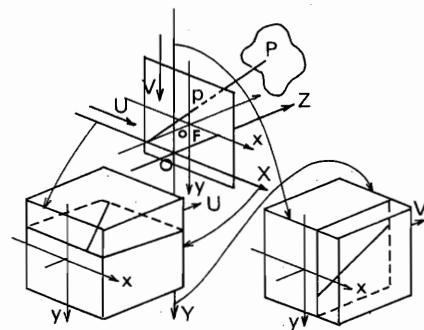


図1 運動軌跡の画像化
Fig. 1 Visualization of apparent motion.

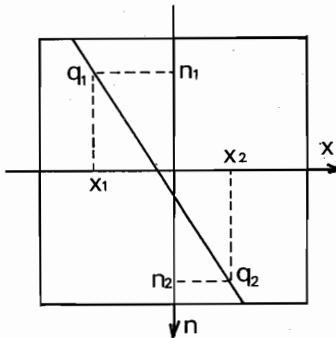


図2 断面画像上の流跡
Fig. 2 A path line on the cross section.

することを示している。この直線軌跡を流跡⁽¹⁶⁾(Path line)とよび、勾配の大きさを流跡の傾きとよぶ。

もし、点 P の投影像がエッジであれば、対応する流跡は、水平断面画像に直線エッジとして画像化される。そのとき、画像間の対応付けの問題は、直線エッジの検出問題に帰着される。この直線エッジ、すなわち、流跡を検出することができれば、対応する点 P の3次元座標値は以下のように復元される。

図2に $y=k_y \cdot F \cdot Y/Z$ での水平断面画像と点 P の流跡を示す。縦軸は、撮影順に順序付けられた視点の番号、横軸は投影面の水平座標値である。流跡上の点 $q_1(x_1, n_1)$ は n_1 番目の視点での点 P の投影点、また $q_2(x_2, n_2)$ は n_2 番目の視点での点 P の投影点に対応している。(3a)式より

$$\frac{Z}{k_x \cdot F} \cdot x_1 = X - n_1 \cdot \delta x \quad (4)$$

$$\frac{Z}{k_x \cdot F} \cdot x_2 = X - n_2 \cdot \delta x \quad (5)$$

(4), (5)式より

$$Z = \frac{n_2 - n_1}{x_1 - x_2} \cdot k_x \cdot F \cdot \delta x \quad (6)$$

となる。流跡の傾き, l_x を直線の最小二乗当てはめにより求める。そのとき, $(n_2 - n_1)/(x_1 - x_2)$ を l_x で置き換えれば、深さ量 Z は

$$Z = l_x \cdot k_x \cdot F \cdot \delta x \quad (7)$$

となる。また Y 座標値は(3b)式より

$$Y = y \cdot l_x \cdot \delta x \cdot k_x / k_y \quad (8)$$

で与えられる。 X 座標値は流跡上の 1 点 (x, n) から(3a)式より、

$$X = (x \cdot l_x + n) \cdot \delta x \quad (9)$$

として得られるが、正確を期すためには、流跡上のすべての点から得られる値を平均する。

$$X = \frac{1}{N} \sum (x_i \cdot l_x + n_i) \cdot \delta x \quad (10)$$

ここで、 (x_i, n_i) は流跡上の点、 N はその総数である。

点 P の投影点が水平、あるいは水平に近いエッジに載っている場合には、流跡が画像化されないか、画像化されたとしてもその信頼性は低い。そこで、視点を Y 軸上に移動させ、垂直移動連続ステレオ画像を得る。図 1 のように、この画像列から 3 次元画像を構成すれば、その垂直断面画像には水平エッジが確実に画像化される。垂直断面画像を利用した場合、 P 点の 3 次元座標値は次の式で与えられる。

$$\begin{cases} X = x \cdot l_y \cdot \delta y \cdot k_y / k_x \\ Y = \frac{1}{M} \sum (y_i \cdot l_y + m_i) \cdot \delta y \\ Z = k_y \cdot F \cdot l_y \cdot \delta y \end{cases} \quad (11)$$

但し、 (y_i, m_i) は流跡上の点、 M はその総数、 l_y は流跡の傾きである。 $k_x \cdot F(k_y \cdot F)$ とカメラのアスペクト比 k_y/k_x は実験的に定める。

2.2 隠れの検出と位置順序の逆転が起こった場合の対応付け

図 3(a)に示されるように、前後に置かれた二つの角柱の 3 次元構造を復元する問題について考えてみよう。このシーンの連続ステレオ画像のうち、視点が左端、中心および右端にあるときの画像を、連続ステレオ画像からサンプルして、図 3(b)に示す。この画像に記入された実線の位置で、合成された断面画像を図 3(c)に示す。角柱の稜上の点 a, b, \dots, h は断面画像の流跡 A, B, \dots, H に、それぞれ対応している。この断面画像に載っている流跡同士の相互関係を解釈すること

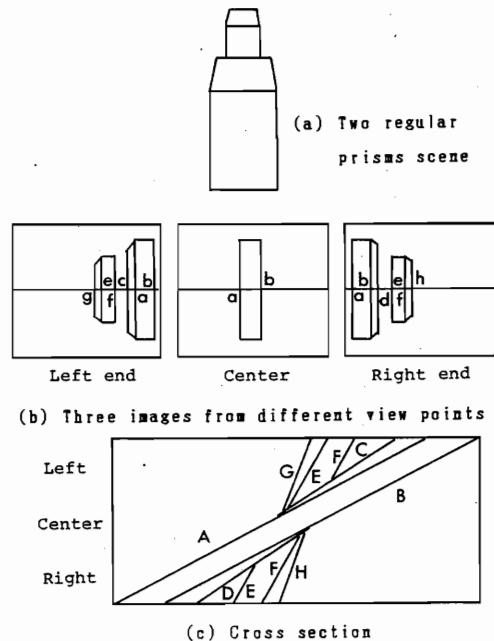


図 3 隠れと位置順序の逆転
Fig. 3 Occlusion and positional reversals.

により、次のことが解る。

(a) 角柱の側面では、視点の位置により、隠れが見られる。例えば、点 a と c の関係のように、遠くの点は近くの点に見かけ上隠される。断面画像上では、遠くの点に対応する傾きの大きな流跡は、近くの点に対応している傾きの小さな流跡に行手を阻まれる。従って、逆に、断面画像上で跡切れた流跡が見つかったならば、その流跡に対応する点の隠れが検出されることになる。

(b) 点 a, e のように、視点の位置によって、見かけの順序が逆転する場合がある。その際、違う点は視点の移動につれ、近くの点に一時的に隠される。断面画像上では、遠くの点に対応する流跡は近くの点に対応する流跡に分断される。従って、逆に、断面画像上で、同一直線上に載る流跡を見つけ出すことができれば、それらの流跡をたどることにより、順序の逆転にも拘らず、対応付けが可能である。

いくつかの流跡が、同一直線上に載るための条件を次のように定める。図 4 に示すように、いくつかの流跡の集まりを視点 n に関する一価関数 $f(n)$ とする。 $f(n)$ の端点を結ぶ線分を $g(n)$ とする。 $f(n)$ と $g(n)$ との最大距離を $d_m = \max_n \|f(n) - g(n)\|$ としたとき、最大距離が適当な許容範囲内に収まっていれば、

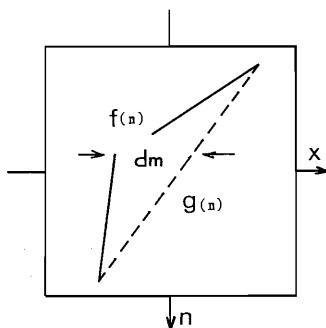


図4 一直線上に載る流跡
Fig. 4 Path lines on a line.

$f(n)$ は同一流跡であるとする。本論文ではこの許容値を2とした。この条件を統合規範とよぶ。

重要なことは、順序が逆転している場合は勿論、隠れが起きている場合でも、流跡が見えている限り、対応する点の3次元座標値が復元可能であることである。

3. 連続ステレオ画像入力システム

連続ステレオ画像の入力システムは、カメラの移動機構、撮影された画像を蓄積するビデオディスク(VM-1000 M, ビクター製)、およびそれらを制御するパーソナルコンピュータ(MBC-225, 三洋電機製)から構成されている。

図5に移動機構を示す。移動機構では、カメラは直角に組み合わされた、スライドパック(THK製)に取り付けられている。カメラの動きは、各々のスライドパックにより水平・垂直方向にガイドされ、その移動量はステップモーター直結のボールネジ送り機構により精密に制御される。最大移動幅は縦横共500mm、移動ピッチは自由に変えられ、最小移動ピッチは縦横共25μmである。カメラは16mm標準TVカメラレンズを付けたCCDカメラ(XC-37, ソニー製)である。パーソナルコンピュータは、移動機構のステップモーターにパルス信号を、ビデオディスクに開始・コマ撮り・コマ送り・停止等の命令をASCIIコードで送っている。

微小移動ピッチ進むごとに撮影された画像は、逐次ビデオディスクに入力される。このディスクは最大600フレームの画像が蓄積可能である。本入力システムは、縦横各々128枚の連続ステレオ画像が約2分で入力可能である。

連続ステレオ画像の全処理システムを図6に示す。

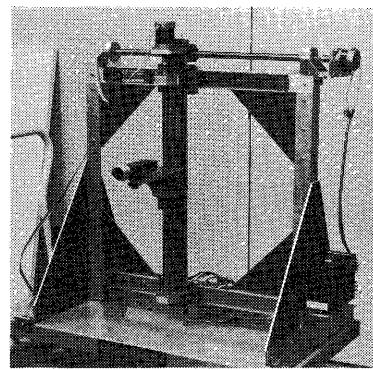


図5 カメラ移動機構
Fig. 5 Camera carrier.

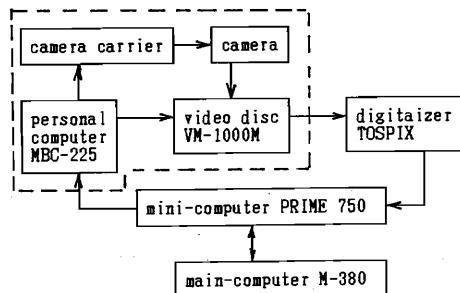


図6 連続ステレオ画像処理システム
Fig. 6 A processing system for position based image sequence.

点線で囲まれた部分が入力システムである。いったん、ビデオディスクに蓄えられた画像は、ディジタイザー(TOSPIX, 東芝製)によって、 512×512 , 256階調に量子化される。更に、 128×128 に縮小スムージングされ、ミニコンピュータ(PRIME 750)に転送される⁽²⁰⁾。全画像量が膨大なことから、断面画像の生成や距離情報の抽出は大型コンピュータ(M-380)に画像を転送して行う。

4. 実験

連続ステレオ画像から断面画像を生成し、流跡の抽出法と3次元構造の復元例について述べる。

視点の移動につれ物体の見かけの運動軌跡が画像化される様子を見てみよう。カメラの三脚やモニター類などが置かれている実験室シーンを対象に、その水平移動連続ステレオ画像を入力した。視点の移動ピッチは3mmで、画像の総数は128枚である。視点が左

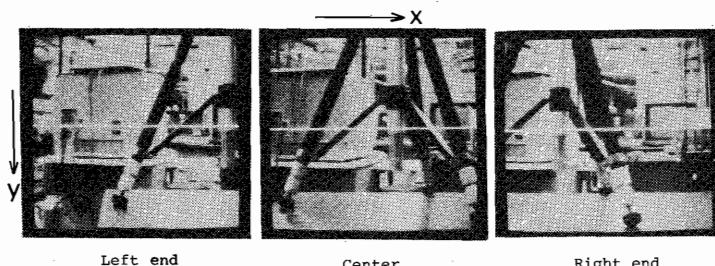


図7 実験室のシーン
Fig. 7 Laboratory scene.

端、中心および右端に位置している時の画像を図7に示す。画像中に白い水平線で示された位置で生成された水平断面画像を図8に示す。

4.1 流跡の抽出

流跡が画像化可能であることが、図8の特徴ある画像により確認できた。この流跡は次のアルゴリズムにより抽出される。

『流跡抽出アルゴリズム』

STEP 1 (流跡候補の抽出)

まず、各画素点の明るさの勾配を求める。勾配の大きさが、ある閾値以上の点を残す。勾配の大きさが、横方向あるいは縦方向で極大となる点を残す。残った画素のうち、8-隣接しているものを互いに連結すれば、いくつかの線の集まりからなる線画が得られる。その際孤立点はノイズと見なして除去する。抽出結果を図9に示す。以下、添え数字は線の番号である。

得られた線画には、流跡の性質と矛盾するものや、流跡として不完全なものが数多く含まれている。その意味で、この段階で抽出された線を流跡候補とよぶ。

STEP 2 (線分集合への分割)

流跡は線分である。従って、流跡候補を線分の集まりとなるように細分化する。

step 2.1 (交差点の除去)

3本以上の線が交わる点で、流跡候補を分断する。

step 2.2 (一価関数化)

視点番号に関して各々一価関数となるように流跡候補を分断する。

step 2.3 (折れ線の分割)

統合規範を満たさない流跡候補は、異なった流跡が誤って繋がった折れ線とみなし、その最大距離を与える流跡候補上の点で流跡候補を分断する。分断点は一般に折れ線の折れ点である。流跡候補の分割結果を図10に示す。

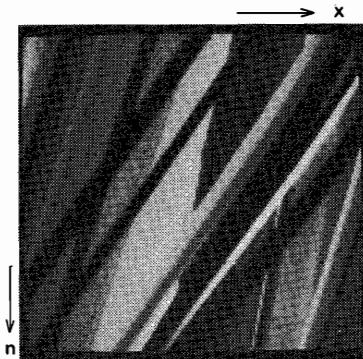


図8 水平断面画像 ($y=64$)
Fig. 8 Cross section at $y=64$.

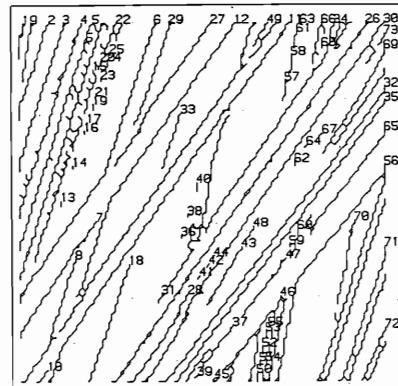


図9 STEP 1 の結果
Fig. 9 Result of STEP 1.

STEP 3 (延長結合)

図10の個々の線分は、それぞれ流跡である。が、本来ならば連続しているはずの流跡が跡切れているものも、なかにはある。流跡を延長し他の流跡と結合することにより、跡切れた流跡を完全な形に復元する。

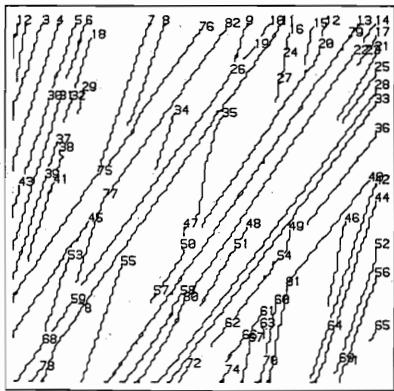
延長させる流跡の順番は流跡の傾きの小さな順、すなわち、より手前にある点に対応する流跡からとする。延長の方向は流跡の両端点を結ぶ直線方向とする。実際に延長する前に、延長方向に他の流跡が存在するのか否かを確かめておく。

他の流跡が存在しないならば、そのまま流跡を延長する。

他の流跡が存在するならば、次の結合処理が試される。

step 3.1 (結合処理)

延長側流跡と相手側流跡を一つの流跡とみなしたと



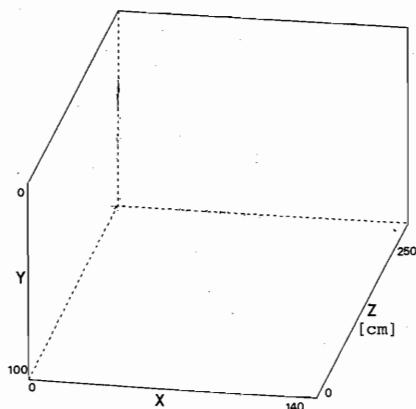


図 14 実験室シーンの3次元構造の推定
Fig. 14 3-D structure of the scene to Fig. 7.

のように、二つの流跡を $f(n)$, $h(n)$ とする。二つの流跡のうち少なくとも一方が定義されている視点の数を N_1 とおく、また $\|f(n)-h(n)\| \leq w_t$ なる視点の数を N_2 とする。そのとき、比 N_2/N_1 がある閾値以上ならば、この二つの流跡は近接しているとする。本実験では閾値を0.9とした。また、 $w_t=2$ とした。一つの流跡 $f(n)$ に対し近接している流跡が複数個存在するならば、そのうち最も比 N_2/N_1 が大きいものを選ぶ。

4.2 3次元構造の復元

図7に示したシーンの3次元構造を復元した結果を図14に示す。3次元の輪郭線を実線で、その投影像を三面図上に点線で表示した。

次に、垂直移動を加えた場合の効用を見てみよう。

机の上に載っているモニターのシーンを対象に直交移動連続ステレオ画像を入力した。視点の移動ピッチ

は1.25 mmで、画像総数は水平方向、垂直方向共に128枚の計256枚である。図15に視点が中心位置にある場合の画像を連続ステレオ画像からサンプルして表示した。また、水平断面画像($y=70$)と垂直断面画像($x=45$)の例が図16(a), (b)に表示してある。図15の白線はそれらの断面位置である。

まず、水平移動連続ステレオ画像からシーンの3次元構造を復元した結果を図17に示す。更に、垂直移動連続ステレオ画像から復元した結果を重ね合わせて、図18に示す。

なお、本実験では $k_x \cdot F = 240$, $k_y/k_x = 1.18$ であった。

5. むすび

視点を移動させることによって得られる、連続ステレオ画像から、シーンの3次元構造を推定する手法を提案した。

具体的には、まず、大量の連続ステレオ画像が高速に入力可能な、専用装置を構成した。本手法によれば、対応付け問題は、連続ステレオ画像から生成される2次元合成画像上の線分検出問題に簡略化される。検出された線分同士の相互関係を調べることにより、隠れの発見や見かけの順序が逆転した場合の対応付けも容易になった。更に、視点を一方向だけではなく、それと直交する方向にも移動させることにより、エッジの向きと関係なく、一意な対応付けが可能になった。最後に、かなり複雑なシーンの3次元構造を復元して本手法の有効性を確認した。

今後の課題として、まず処理の高速化が挙げられる。専用の入力装置により、連続画像の入力は高速に

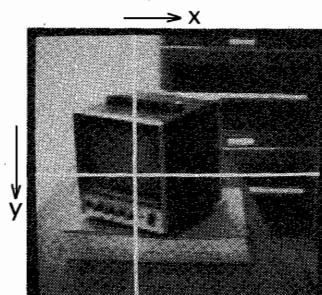
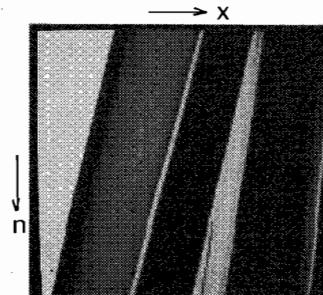
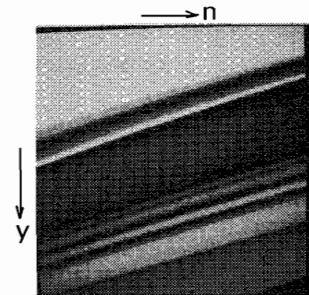


図 15 モニターのシーン
Fig. 15 Monitor scene.



(a)



(b)

図 16 (a) 水平断面画像 ($y=70$)
(b) 垂直断面画像 ($x=45$)

Fig. 16 (a) Horizontal cross section at $y=70$.
(b) Vertical cross section at $x=45$.

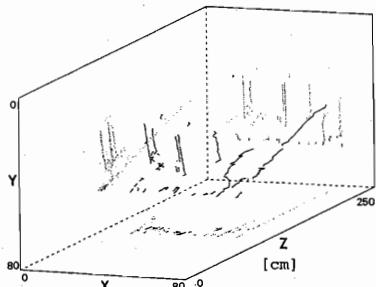


図 17 水平移動連続ステレオ画像からの復元結果
Fig. 17 3-D structure from horizontal moving camera.

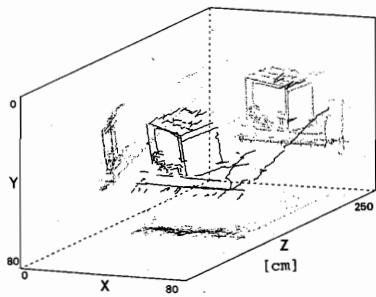


図 18 水平・垂直移動連続ステレオ画像からの復元結果
Fig. 18 3-D structure from horizontal and vertical moving camera.

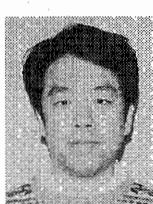
なった。しかし、その後の処理は、現在のところかなりの時間を要する。処理の並列化とハードウェア化は検討すべき問題である。次に、本研究を、環境モデルの構築や物体認識へと進めるにあたっては、本論文で示した復元結果は必ずしも満足のいくものではない。更に精密な移動機構の設計と測定精度の定量的な把握が必要である。

謝辞 本研究を進めるに当り、御指導頂いた白井良明制御部長、並びに御討論頂いた大島正毅室長をはじめとする視覚システム研究室の皆様に感謝致します。また、移動機構の試作に御尽力下さいました棚沢清氏をはじめ、研究品試作課の方々に感謝致します。

文 献

- (1) B. D. Lucas and T. Kanade : "An iterative image registration technique with an application to stereo vision", 7th-IJCAI, pp. 674-679 (Aug. 1981).
- (2) K. F. Prazdny : "A simple method for recovering relative depth map in the case of a translating sensor", 7th-IJCAT, pp. 698-699 (Aug. 1981).
- (3) W. E. L. Grimson : "Computational experiments with a feature based stereo algorithm", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-7, 1, pp. 17-34 (Jan. 1985).
- (4) S. T. Barnard and W. D. Thompson : "Disparity analysis of images", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-2, 4, pp. 333-340 (July 1985).
- (5) 大田、正井、池田：“動的計画法によるステレオ画像の区間対応法”，信学論(D), J68-D, 4, pp.554-561 (昭60-04).
- (6) Y. Ohta and T. Kanade : "Stereo by intra-and inter-scan line search using dynamic programming", IEEE Trans. Pattern Anal. & Mach. Intell., PAMI-7, 2, pp. 139-154 (March 1985).
- (7) 池内克史：“二枚の法線分布図からの領域対応法による距離分布図の生成”，信学論(D), J68-D, 10, pp.1761-1768 (昭60-10).
- (8) 富田文明：“ステレオ画像の領域の輪郭線セグメントに基づく対応”，情処学コンピュータビジョン研資, 38-1 (昭60-09).
- (9) H. P. Moravec : "Visual mapping by a robot rover", Proc. 6th-IJCAI, pp. 598-600 (Aug. 1979).
- (10) 大森、森下：“多視点ステレオ画像を用いた物体検出”，SICE論文集, 18, 7, pp.716-722 (昭57-07).
- (11) 谷内田正彦：“多重画像を用いて3次元物体を認識”，日経メカニカル, 157, pp.82-91 (昭59-01).
- (12) 伊藤、石井：“3眼画像による立体視対応点処理”，情報処理学会第29回全国大会, 2M-3 (昭59-09).
- (13) 大田、池田：“3眼ステレオについて”，ibid, 2M-4.
- (14) J. F. Jarvis: "Automatic visual inspection of glass-metal seals", Proc. 4th-IJCPR, pp. 961-965 (Nov. 1978).
- (15) 佐々木、古川：“多眼画像からの中間像の生成”，信学論(D), J63-D, 9, pp.813-814 (昭55-09).
- (16) 山本正信：“画像化された運動軌跡による動画像処理”，情処学論, 22, 5, pp.442-449 (昭56-09).
- (17) 山本正信：“動画像処理による形状測定”，昭56信学総全大, S 18-12.
- (18) 金子透：“物体の準静的回転運動による3次元形状入力”，信学技報, PRL85-22 (1985-09).
- (19) R. C. Bolles and H. H. Baker : "Epipolar-plane image analysis: A technique for analyzing motion sequences", Proc. 3rd ISRR (Oct. 1985).
- (20) 吉見、築山：“局所並列画像処理装置の開発”，第2回日本ロボット学会学術講演会, 3317 (1984-12).

(昭和61年5月24日受付)



山本 正信

昭48九大工・工・制御卒。昭50東大大学院修士課程了。同年電子技術総合研究所入所。以来、動画像処理、知能ロボット等の研究に従事。現在、制御部視覚システム研究室主任研究官。情報処理学会、計測自動制御学会、日本ロボット学会、AVIRG等各会員。