

ラージスケールステレオ

高橋 裕信[†] 富田 文明[‡][†]三洋電機筑波研究所 [‡]電子技術総合研究所

ステレオを用いて3次元計測を行なう場合、屋内での限定された環境に比較して、屋外での3次元計測はその範囲が大きく十分な精度を得ることは難しい。これはステレオカメラの基線長が対象までの距離に比べて短いためである。そこで単眼のカメラを移動して得られた画像を対応付け、基線長の大きなステレオ画像として処理する方法が考えられる。本研究報告では、移動量が未知のカメラから得られた画像を対応付け、その移動と3次元シーンの再構成を行った。移動による対象の変形に対処するため、画像から特徴的なプリミティブを抽出し対応単位とした。

Large Scale Stereo Vision

Hironobu TAKAHASHI[†] Fumiaki TOMITA[‡][†] Tsukuba Research Center, Sanyo Electric Company, Ltd.

2-1 Koyadai Tsukuba Ibaraki 305 Japan

[‡] Electrotechnical Laboratory

1-1-4 Umezono Tsukuba Ibaraki 305 Japan

In the case of stereo measuring in the 3-dimensional world, it is difficult to obtain the sufficient accuracy in the outside world compared to the structured indoor environment, since the length of the base line is not enough for the distance between the objects and cameras. We propose to move a single camera and to match images of which the length of the base line is long enough for measurement. This paper reports methods and experimental results of the primitive based correspondence and the estimation of the motion of the camera.

1. はじめに

ステレオを用いて3次元計測を行なう場合、屋内での限定された環境に比較して、屋外での3次元計測はその範囲が大きく十分な精度を得ることは難しい。これはステレオカメラの基線長が対象までの距離に比べて短いことが大きな原因である。

こうした問題を解決するために、カメラを移動して得られる画像を用いる方法が考えられる。しかし、不整地環境などでは移動量の測定に誤差が含まれたり、移動量が全く不明という場合もあり、3次元計測と同時にカメラの移動量の測定も必要となる。

こうした問題に対する研究としては、単眼の連続したモーション画像を対象とした多くの研究や、局所的に得られたステレオ像を対象とする方法などがある[1]。

この研究報告では離散的に得られた単眼の画像を用いて、基線長の大きな大規模ステレオ (large scale stereo) として扱う方法について検討した。

2. アルゴリズム

図1のような手順で処理を行なった。ステレオ画像間で正確な対応付けができれば、その2枚の画像を撮影したカメラのパラメータは、スケールファクタを除いて決定することができる。これらのうち1台のカメラ位置を固定して適当なスケールファクタを与えれば、相対的に他のカメラの位置が決定できるため、撮像時の移動軌跡を求めることができる。しかし、一方では正確な対応付けを行うためには正確なカメラキャリブレーションが必要であり、これらの要求は矛盾する。

カメラパラメータ (すなわち移動軌跡および回転) の決定には、少数の対応点で十分である。そこで画像から特徴的なプリミティブを抽出しこの対応から移動軌跡を求めた。次にこれらのパラメータを基礎にしてステレオ対応を行なう。

対象としたシーン中にはすでに特徴が明らかな人工物があることを仮定している。例えば、図1のように家、ビルディングなどが環境中にあり、少なくともその一部が画像中に撮像されているものとする。システムにはあらかじめこれらの対象についての知識が与えられる。この知識には、画像

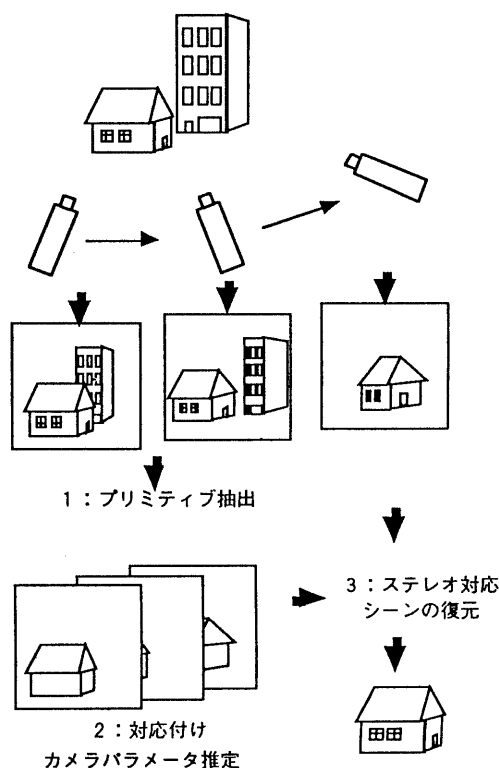


図1 処理手順

中での見え方と3次元空間中での質的な拘束の2つが含まれる。

たとえば「家」の知識のうち、画像中での見え方は、図2に示したようなプリミティブの階層構造を形成する[2]。撮影したカメラは水平方向に向いており、極端な傾きは考えない。対象はまず見え方によって分けられる。次にそれらは構成するプリミティブに分割される。この例では「壁」や「屋根」の部分に分割される。さらにそれぞれは下位のプリミティブに分割される。この例では、壁は地面に対して垂直に立っているため、各辺が画像中で水平と垂直に向いている長方形で近似される。屋根は平行四辺形または2等辺三角形で近似でき、平行四辺形のうちの2辺と三角形の底辺は水平である。これらのプリミティブは平行な線分や左右対称な線分の対から構成され、さらに個々の線分に分割される。逆に画像からこの過程をたどれば対象が抽出できる。

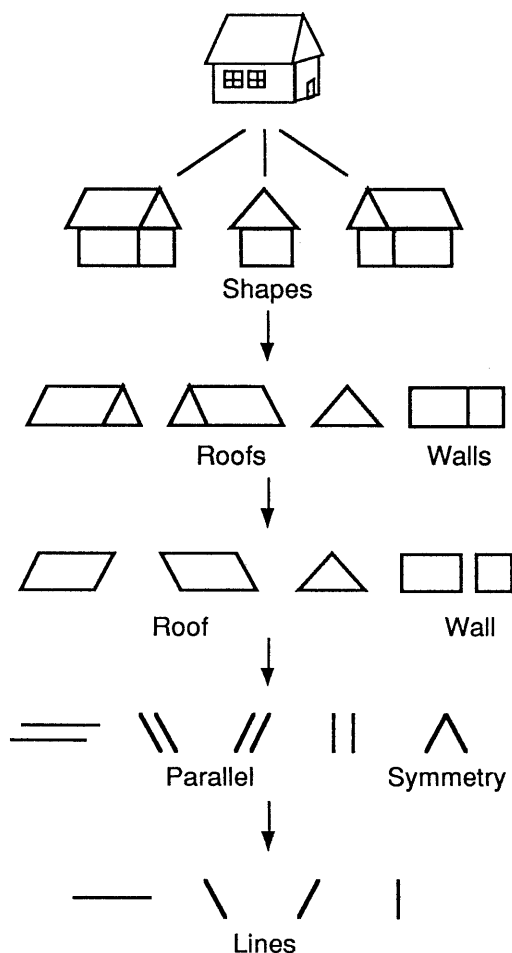


図2 プリミティブ

3次元空間中では「家」を構成する「壁」や「屋根」などのプリミティブの単位ごとの知識を拘束として用いることができる。例えば「屋根」では図3のように、それらの頂点が同一平面上にあること、頂点での角度、面の方向が上方を向いていることなどがある。これらは対応付けや検証に利用される。

3. カメラ外部パラメータの算出

カメラパラメータのうちカメラの内部パラメータは移動により変化しないので、あらかじめ他の測定方法によって求め、画像は理想的なピンホー

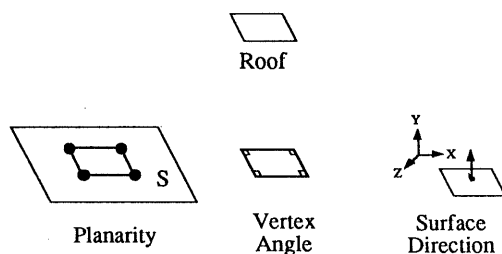


図3 プリミティブの3次元拘束

ルカメラと同等となるように補正する。移動による、焦点の位置と回転の外部パラメータについて検討する。画像からの対象の位置の測定精度を向上させるため、ズームレンズにより焦点距離を変化させて、注視する方法も考えられるので、焦点距離を変化させる場合についても検討する。

・焦点距離固定の場合

N台のカメラ画像上で同時にM点が撮像されており、画像間での対応付けがされているものとする。この時自由度は、個々のカメラについて空間中の移動と回転の6自由度、各点が移動の3自由度ある。この時全自由度は以下の式で表される。

$$6N + 3M$$

画像上の点の対応は拘束をあたえ、画像上の1点あたり自由度が2減る。また絶対位置および回転とスケールファクタは求めることはできないので、あらかじめ既定値として与える。このため自由度を7減じる。

$$\text{自由度} = 6N + 3M - 2NM - 7$$

この値が0以下の場合にカメラパラメータと対応点の位置が決まる。表1にこの結果を示した。これから3台のカメラで同時に4点の対応が得られるか、または2台のカメラで5点以上の対応点が必要であることがわかる。

対応点としてプリミティブ上の各点を与える場合には、プリミティブ固有の拘束条件を与えることができる。例えば、「屋根」ではその4つの頂点が同一平面上にある。

表1 焦点距離固定の場合の自由度

対応点	画像							
	2	3	4	5	6	7	8	
2	3	8	7	9	11	13	15	
3	2	2	2	2	2	2	2	
4	1	-1	-3	-5	-7	-9	-11	
5	0	-4	-8	-12	-16	-20	-24	
6	-1	-7	-13	-19	-25	-31	-37	
7	-2	-10	-18	-26	-34	-42	-50	

表2 対応点が平面上にある場合の自由度

対応点	画像							
	2	3	4	5	6	7	8	
2	3	8	7	9	11	13	15	
3	2	2	2	2	2	2	2	
4	0	-2	-4	-6	-8	-10	-12	
5	-2	-6	-10	-14	-18	-22	-26	
6	-4	-10	-16	-22	-28	-34	-40	
7	-6	-14	-22	-30	-38	-46	-54	

対応点が空間中で同一の平面上にある場合には4点目以降の点の自由度は各点あたり2となる。そのため全自由度は次の式で表わされる。

$$\text{全自由度} = 6N + (2M + 3) - 2NM - 7$$

表2にこの結果を示した。対応点が同一平面上にある場合には2台のカメラで4点の対応点で十分であることがわかる。

さらにプリミティブの各頂点での直交性を拘束として用いることも可能である。この報告では対応後に得られた対応点の空間中での位置関係より対応の妥当性の検証に用いた。

・焦点距離が可変の場合

同様に自由度は、個々のカメラについて空間中の移動と回転と焦点距離の7自由度、各点が移動の3自由度あるので以下の式で表される。

$$\text{自由度} = 7N + 3M$$

表3 焦点距離可変の場合の自由度

対応点	画像							
	2	3	4	5	6	7	8	
2	4	7	10	13	14	17	20	
3	3	4	5	6	7	8	9	
4	2	1	0	-1	-2	-3	-4	
5	1	-2	-5	-8	-11	-14	-17	
6	0	-5	-10	-15	-20	-25	-30	
7	-1	-8	-15	-22	-29	-36	-43	

絶対位置および回転と基準となる焦点距離、スケールファクタは求めることはできないので、あらかじめ既定値として与える。そのため8自由度を減じる。

$$\text{全自由度} = 7N + 3M - 2NM - 8$$

この計算結果を表3に示した。これから2台のカメラでは6点以上の対応点が必要であることがわかる。

焦点距離が固定の場合とは異なり、対応点が同一の平面上にある場合には焦点距離とその平面までの距離を独立に求めることはできない。例えばカメラの光軸が平面に直交している場合では、平面までの距離と焦点距離の比を一定にすると、画像は変化しない。このため焦点距離の変化も求めるには同一平面上にない6対応点が必要となる。

2枚の画像の対応点の位置より各カメラのパラメータを求めるにはステレオカメラのセルフキャリブレーションの手法を用いた[3]。2台のカメラが基準カメラモデル(図4)に一致した配置となっていれば、各画像中で対応点はエピポーラ条件を満たし、垂直方法の座標値は一致する。実際に得られた2枚の画像に対してカメラの回転と等価な変形を行ない、エピポーラ条件を満たすようにする。この回転からカメラ間の相対関係を求める。

対応点が空間中で同一平面上にある場合には、対応点の画像上での位置は一次の分数式で関係付けられる。カメラの位置、回転の各パラメータは解析的に求めることができ[4]、2つの解が得られる。対象として仮定したプリミティブの知識(例え

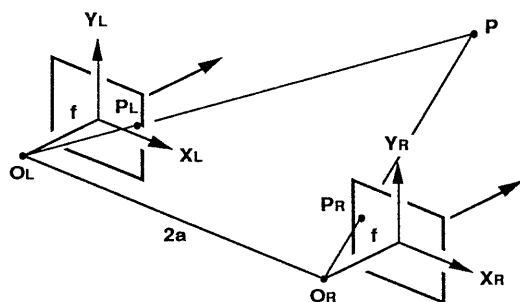


図4 基準カメラモデル

ば平面の方向が上に向いている等)を用いて適当な解を選択する。

2枚の画像の対応ではスケールファクタは未知である。3ヶ所以上の点で撮像した場合には基準となる画像と対象の特定の点を定めて、カメラからの距離を基準値として統合する。

5. 実験結果

対象として一般的な傾斜屋根のある住宅を選び、移動を行ない撮影した。屋根の部分の対応付けにより、移動軌跡と屋根の形状を求める。

画像からエッジを抽出し境界線表現に変換する[5]。次に基準カメラモデルに一致するように、あらかじめ求めたカメラの内部パラメータに従って画像を変換する。

図5に示したような手順で屋根の抽出を行った。視点によっていくつかのタイプに分類することができるので、屋根の見込む方向によって左向きのもつと右向きのもつの2つに分ける。

道路上を移動しながら22地点より撮影した。これらの画像を処理しプリミティブを抽出する。

[線分の抽出]

屋根を構成するセグメントとして「水平」、「左上り」、「右上り」のセグメントを選択する。セグメントは方向を持っているので、始点から終点への変位 dx, dy により分類する。

水平 :	$\text{Absolute } (dx / dy) < 0.3$
左上り :	$dx * dy < 0$
右上り :	$dx * dy > 0$

得られた候補から、長さにより上位200個を選択した。図5(a)の画像を処理すると、それぞれ図5(b),(c),(d)が得られる。

[平行線分の抽出]

次に互いに平行で向いあった位置にある組合せを探す。組合せ爆発を避けるために、線分の画像中での長さを評価値としてソートし、評価値の大きなものから組み合わせを試行する。一定数(200-1000)の組み合わせが得られた時点でそれぞれの組み合わせを中止するようにしている。

[平行四辺形の抽出]

平行な線分を組合わせて平行四辺形を探索する。条件として各頂点で凸角を形成するようなものを選択する(c)。

[統合と評価]

エッジのかすれなどによって線分が分割されたために、同一の平行四辺形を重複して求めている可能性がある(図7a,b)。そこで得られた図形を互いに重ね合せ、構成する4本の線分がすべて平行である場合には同一のものとして統合する(c)。

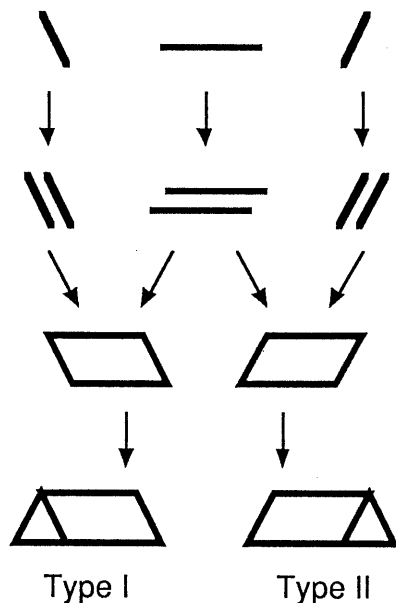


図5 プリミティブの抽出

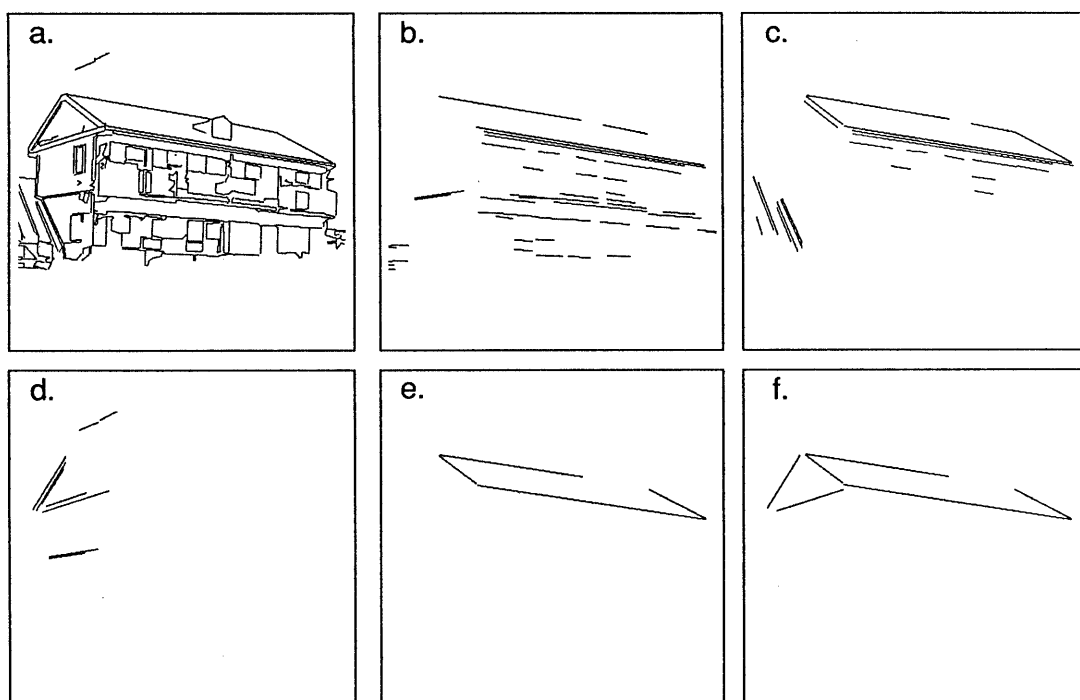


図6 プリミティブの抽出

4つの頂点を結ぶ全ての線分(d)のうち、実際にエッジが存在する部分(c)の比率をこのプリミティブの評価値とした。ただし、極端に小さなものは除いた。

[三角形の抽出]

屋根を構成するもののうち、平行四辺形と隣接して2等辺三角形の部分が見えることが期待できる。平行四辺形のうちいずれかの側の2頂点を通る線分を選び適当な2等辺三角形を形成するものを探す(f)。その結果図8に示したように各画像から屋根を発見された(図8)。

実験ではカメラの焦点距離を一定としたので、屋根を構成する平行四辺形部分の4頂点から撮像位置を解析的に求めることができ、解のうち屋根の向きが上方を示しているものを選択した。

実験の結果21画像のうち9-21の画像からは同一のタイプの屋根が得られたので、画像10を基準としてカメラ位置と方向を推定し、その屋根の位置を求めた(図9)。

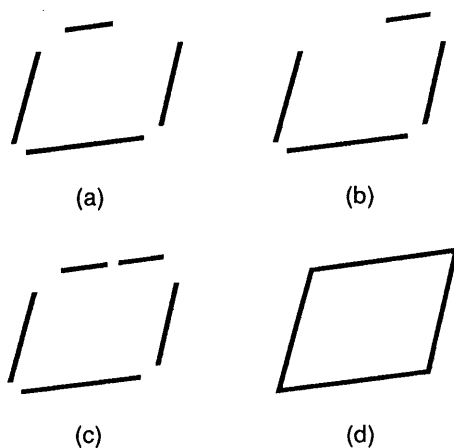


図7 プリミティブの統合

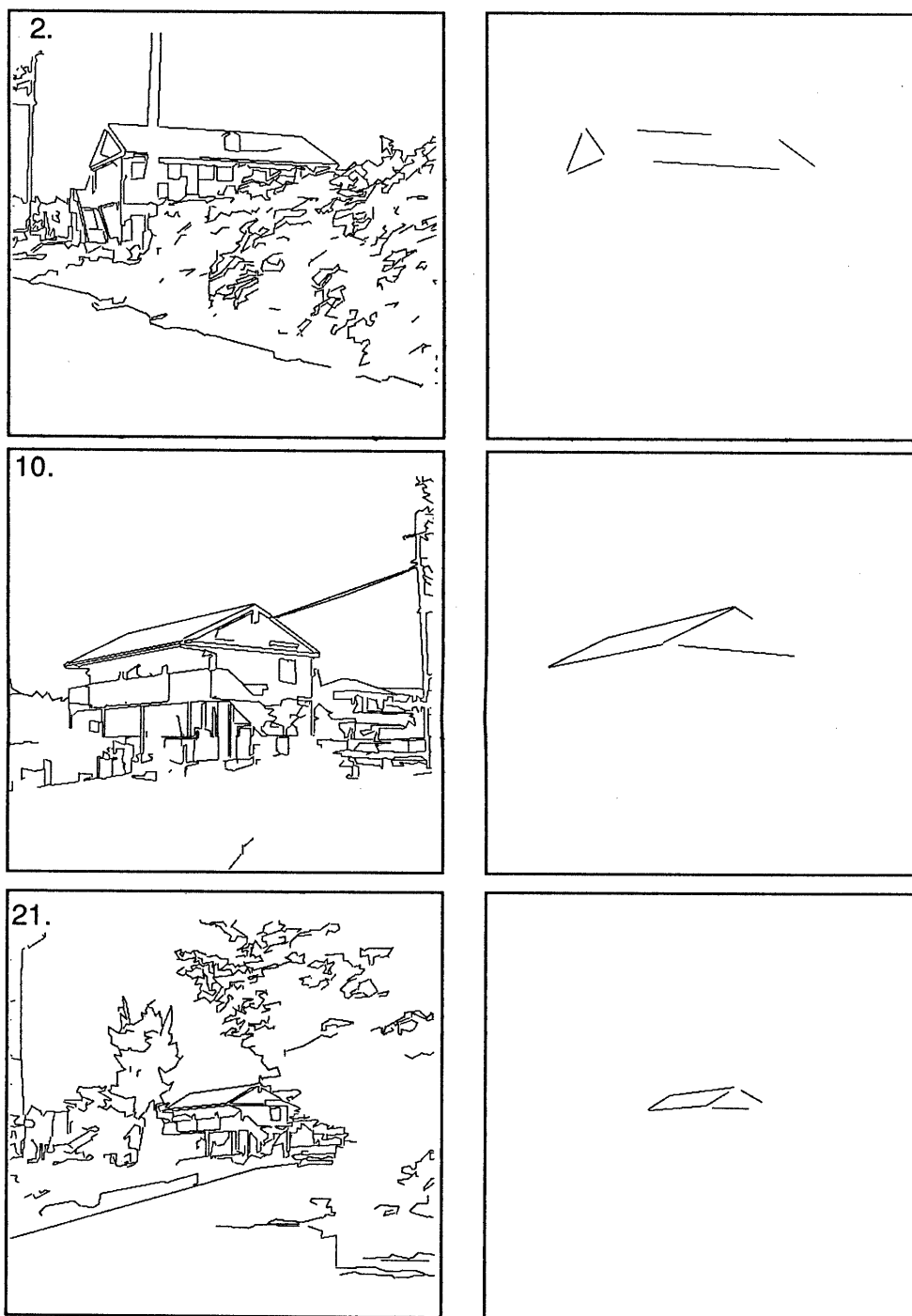


図8 実験結果 (左: B-rep像 右: 抽出結果)

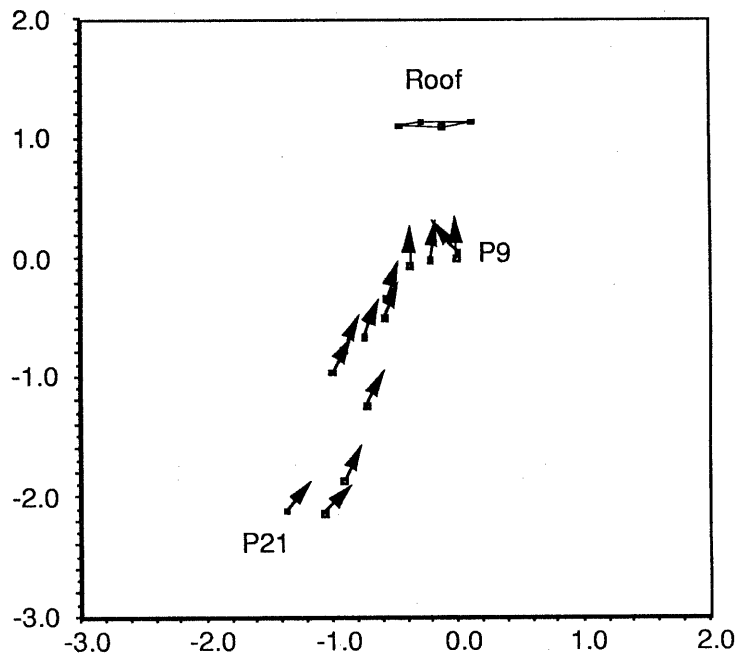


図9 カメラ位置および方向

6. まとめ

移動を行なうシステムで、それ自身の移動量と対象の復元を行なった。この方法の特徴は不連続に得られた画像であっても、プリミティブを対応単位として用いているため、安定した対応が得られることである。

今後はより汎用性を持たせるためにプリミティブの認識木の自動作成などについて検討する。

謝辞

日頃、熱心な御討論と助言を頂く三洋電機筑波研究所知能システム研究室と電子技術総合研究所視覚情報研究室の皆様へ感謝します。

参考文献

- (1) 杉本、高橋、富田: 移動ロボットのためのステレオビジョンによる環境の幾何モデリング、情報処理、CV68-5,1990
- (2) 小泉、富田: 階層的な形状プリミティブの並列的発見、情報処理、CV69-1,1990.
- (3) 高橋、富田: ステレオカメラのセルフキャリブレーション、情報処理、31-5,1990.
- (4) 金谷健一: 画像理解/3次元認識の数理 森北出版 1990.
- (5) 高橋、富田: 画像のB-R-E-Pのためのアルゴリズム、信学会、PRU86-87,1987.