

ステレオカメラを用いた視差情報による屋内床面抽出の研究

Study of Floor Surface Extraction Indoors with Disparity Information Using Stereo Cameras

情報工学専攻 内藤 康太
NAITOU Kouta

概要

本研究の目的は、壁で囲まれた屋内においてステレオカメラで撮影された画像から床面の領域を抽出することである。ステレオカメラを用いた3次元環境認識の手法として、ステレオカメラから得られた視差情報をもとにV-disparity画像を求め、平面を抽出する手法がある。V-disparity画像とは、視差画像の高さごとの視差値の出現頻度を表した画像である。また、本研究で対象とする屋内は、側面の壁のようにカメラに大きく写る平面が多く存在する。そこで本研究では、このV-disparity画像を用いた手法を拡張し水平方向Uの視差の変化に注目したU-disparity画像を用いる。U-disparity画像を用いることで側面の壁を示す視差値は、視差値と画像の横座標で表せる直線となり、この直線を検出することで側面の壁を検出する。本研究では、これら床面の視差値を示す直線と側面の壁を示す直線を検出し、これらの直線にあたる視差値から床面を抽出する手法を提案する。実験で理想となる視差画像とノイズをもった視差画像に対して提案した手法を用いて床面の抽出実験をおこない、それらの実験から本手法の有効性を示した。

キーワード：ステレオ視、視差画像、平面抽出、直線検出、V-disparity、U-disparity。

1 序論

本研究の目的は、壁などの遮蔽物が多い屋内においてステレオカメラで撮影された画像から床面の領域を抽出することである。ステレオカメラとは、対象物を異なる位置から同時に撮影し、カメラと対象物体までの距離を求めることができるカメラである[3]。本研究では、ステレオカメラにより求められた3次元情報を表した画像を視差画像と呼ぶ。Labayrade[2]はステレオカメラから得られた視差画像をもとにV-disparity画像を求め、平面を抽出する手法を提案している。V-disparity画像とは、視差画像の高さごとの視差値の頻出度を表した画像である。抽出したい床面の視差は、V-disparity画像において視差値と画素の高さで表せる直線として表れる。このV-disparity画像上に表れる直線を検出することで、床面の抽出をおこなう。直線の検出には、ノイズに対して頑健であるHough変換を用いる。Labayradeの手法は、視差画像から3次元空間の幾何学的構造を解析し、平面を抽出しているため、ステレオカメラと平面の姿勢の変化に頑健である。しかし、この手法では不正確な対応付けをおこなってしまったときなど視差値を正しく求められない場合、その部分を平面として抽出できない問題がある。

本研究で対象とする屋内は、側面の壁のようにカメラに大きく映る平面が存在する。本研究はこれらの側面の壁が床面に対して垂直であると仮定することで、視差画像からこれらの側面の壁を示す視差を推定する。しかし、側面の壁は、縦方向への視差値の変化はなく、V-disparity画像に特徴となる直線として現れない。そこで、このV-disparity画像を用いた手法を拡張し水平方向Uの視差の変化に注目したU-disparity画像を用いる手法を提案する。U-disparity画像を用いることで、側面の壁を示す視差を視差値と画像上の横座標で表せる直線として検出することが可能になる。具体的には、V-disparity画像から床面の領域を示す視差値を推定し、U-disparity画像から壁の領域を示す視差値を推定する。さらに壁と床面の境界の3次元空間上の座標は同じであり、V-disparity画像とU-disparity画像から推定された2つの推定した視差値は、壁と床面の境界において一致す

る。この一致する視差値が現れる座標を壁と床面の境界として検出し、その境界に囲まれた領域を床面として抽出をおこなう。床面と壁の境界を検出することにより、視差画像の正確度が良くない場合でも、床面の抽出が可能になる。本研究では、手動で求めた理想的な視差画像とノイズのある視差画像について実験をおこない本手法の有効性の確認と検証をする。本研究で対象となる屋内を以下のように仮定する。

- 床面と壁は平らである
- 床面の両端に床面に対して垂直な壁が存在する

2 ステレオカメラを用いた3次元計測

2.1 ステレオカメラモデル

本研究では2台のカメラの位置に制約を設けたモデルを定義することで、視差値と画像座標上での位置から3次元空間上の平面を推定する。本研究で用いるステレオカメラのモデルを図1に示す。カメラはピンホールカメラモデルを採用する。視点 C_L 、視点 C_R とそれらの視点が観測する画像平面との距離を焦点距離といい、視点から画像平面に対して垂線を下ろしたとき、その垂線とその延長した直線を光軸という。また、 C_L と C_R を結ぶ直線を基線といい、基線の長さを基線長と呼ぶ。そして、このモデルでは以下の条件を満たす。

- 左右のカメラで光軸が平行
- 視点の高さが等しい
- 2つの焦点距離が等しい

この条件を満たすステレオカメラにおいて、3次元空間上にある点Pが、左画像において (u_l, v_l) 、右画像において (u_r, v_r) の座標に映ったとする。このとき、点Pの3次元空間上での座標は、

$$P = \frac{b}{u_l - u_r} \begin{pmatrix} u_l \\ v_l \\ f \end{pmatrix} \quad (1)$$

で求まる。 $u_l - u_r$ を視差と呼び、本研究ではその値を視差値と呼ぶ。このステレオカメラのモデルを定義することにより、3次元の空間上の点は、式(1)で示したように視差値を求めることで算出できる。

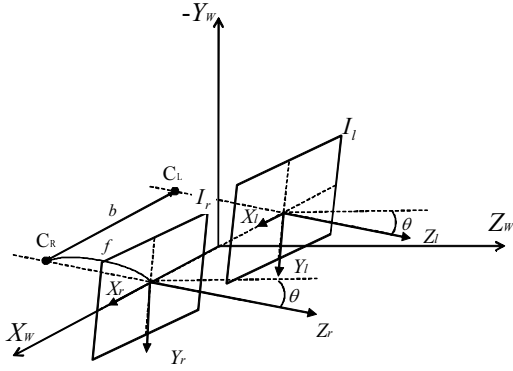


図1 ステレオカメラのモデル

2.2 3次元座標から画像座標への変換

世界座標系を $WCS(X_w, Y_w, Z_w)$, 左カメラ座標系を $LCCS(X_l, Y_l, Z_l)$, 右カメラ座標系を $RCCS(X_r, Y_r, Z_r)$ で表す. 左画像の座標を $I_l(u_l, v_l)$, 右画像の座標 $I_r(u_r, v_r)$ とし, 各画像の中心 $I_{r,l}(u_0, v_0)$ とする. 各カメラ座標系の原点は (u_0, v_0) である. 世界座標系の原点を $I_l(u_0, v_0)$ と $I_r(u_0, v_0)$ を結んだ線分の中点とし, 基線長を b , カメラの光軸とのなす角を θ とすると, 世界座標系から左右のカメラ座標系への平行移動行列 T_l, T_r と回転行列 R は,

$$T_* = \begin{pmatrix} 1 & 0 & 0 & \epsilon_* \frac{b}{2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad (2)$$

$$R = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (3)$$

で表すことができる. ここで, $*$ は l または r であり, $\epsilon_l = -1$, $\epsilon_r = 1$ である. 式 (2) と式 (3) から世界座標系から左右のカメラ座標系への変換行列 D_l, D_r は,

$$D_* = T_* R = \begin{pmatrix} 1 & 0 & 0 & \epsilon_* \frac{b}{2} \\ 0 & \cos \theta & -\sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (4)$$

で表すことができる. この変換行列により, 世界座標系で表された座標からカメラの光軸を基準としたカメラ座標系の座標に変換することができる.

カメラ座標系から画像座標系への透視変換行列 P は,

$$P = \begin{pmatrix} f & 0 & u_0 & 0 \\ 0 & f & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad (5)$$

である. 世界座標から左右の画像座標への変換行列は, 世界座標の同次行列 $(X_w, Y_w, Z_w, 1)^T$ と左右の画像座標の同次行列 $(u, v, 1)^T$ を用いて,

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = PD_* \begin{pmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{pmatrix} \quad (6)$$

と表すことができ, 式 (4) と式 (5) から,

$$PD_* = \begin{pmatrix} f & u_0 \sin \theta & u_0 \cos \theta & \epsilon_* \frac{b}{2} \\ 0 & v_0 \sin \theta + f \cos \theta & v_0 \cos \theta - f \sin \theta & 0 \\ 0 & \sin \theta & \cos \theta & 0 \end{pmatrix} \quad (7)$$

である. 式 (6) と式 (7) から, 世界座標である点 (X_w, Y_w, Z_w) の左右のカメラ画像上での座標 (u_*, v) は,

$$\begin{cases} u_* = u_0 + f \frac{X_w \epsilon_* \frac{b}{2}}{Y_w \sin \theta + Z_w \cos \theta}, \\ v = v_0 + f \frac{Y_w \cos \theta - Z_w \sin \theta}{Y_w \sin \theta + Z_w \cos \theta} \end{cases} \quad (8)$$

と求まる. ここで, 式 (8) を変形し, 左右のカメラの光軸に関して新たに画像座標 (U, V) を定義すると,

$$\begin{cases} U_* = u_* - u_0 = f \frac{X_w \epsilon_* \frac{b}{2}}{Y_w \sin \theta + Z_w \cos \theta}, \\ V = v - v_0 = f \frac{Y_w \cos \theta - Z_w \sin \theta}{Y_w \sin \theta + Z_w \cos \theta} \end{cases} \quad (9)$$

と表すことができる. また, 式 (9) から視差 d は3次元空間上の座標を用いて,

$$d = u_l - u_r = f \frac{b}{Y_w \sin \theta + Z_w \cos \theta} \quad (10)$$

で表すことができる.

2.3 視差画像への投影

ステレオカメラで撮影したときに3次元空間上の平面の視差値 d が視差画像の座標 (U, V) とどのように関係になるかについて述べる. 抽出したい床面は水平な平面であり, 屋内に多く存在する壁は垂直な平面や斜平面である. 本節では, これら水平な平面, 垂直な平面, 斜平面の3つの平面の視差値 d と視差画像の座標の関係を述べる.

水平な平面

高さ h の X_w - Z_w 平面と平行な平面は, 世界座標系を用いて, $Y_w = h$ と表すことができ, 式 (9) と式 (10) から,

$$\frac{h}{b} d = f \sin \theta + V \cos \theta \quad (11)$$

となる. つまり, 3次元空間上において水平な平面は, 画像座標 (U, V) における V と視差値 d を用いて表すことができる.

垂直な平面

基線から距離 p にある X_w - Z_w 平面に対して垂直で基線と平行な平面は, 世界座標系を用いて, $Z_w = p$ と表すことができ, 式 (9) と式 (10) から,

$$\frac{p}{b} d = f \cos \theta - V \sin \theta \quad (12)$$

と表すことができる.

斜平面

X_w - Z_w 平面に対して垂直かつ, 基線と平行でない斜平面は, σ と τ を用いて, $Z_w = \sigma X_w + \tau$ と表すことができ, 式 (9) と式 (10) から, 画像座標 (U, V) と視差 d を用いて,

$$\sigma U + V \sin \theta + \frac{2\tau - \sigma b}{2b} d = f \cos \theta \quad (13)$$

となる. このとき, θ の値が無視できるほど十分小さい場合,

$$\sigma U = \frac{\sigma b - 2\tau}{2b} d + f \quad (14)$$

と表すことができ、 θ の値が十分に小さければ画像の水平方向である U と視差値 d で表せることがわかる。また、光軸に対して平行な $X_w = r$ の平面は式 (9) と式 (10) から、

$$U = \frac{2r + b}{2b}d \quad (15)$$

となり、画像の水平方向である U と視差値 d で表せることがわかる。

3 提案手法

3.1 V-disparity 画像

V-disparity 画像は、縦軸に視差画像の高さ V 、横軸に視差値 d をとった画像であり、視差画像中の高さごとの視差値の頻出度を画素の濃度値で表したものである。V-disparity 画像を求めることにより、視差画像中のある高さ i における視差値の分布を調べることができる。

V-disparity 画像を $I_{v\Delta}$ 、視差画像を I_Δ とし、 I_Δ において画像座標 (v, u) における視差値を $d = I_\Delta(v, u)$ と表す。 $I_{v\Delta}$ の高さ i 、視差値 d の頻出度を $I_{v\Delta}(i, d)$ と表したとき、

$$I_{v\Delta}(i, d) = \sum_{u=0}^W \delta_{I_\Delta(i, u), d} \quad (16)$$

となる。ここで W は I_Δ の幅であり、 $\delta_{a, b}$ は Kronecker delta である。すべての i について式 (16) を用いて $I_{v\Delta}(i, d)$ を求めることで、V-disparity 画像が求まる。

3.2 U-disparity 画像

U-disparity 画像は、縦軸を視差値 d 、横軸を画像の幅 U をとった画像であり、視差画像中のある垂線ごとの視差の頻出度を画素の濃度値で表したものである。U-disparity 画像を求めることにより、視差画像中のある横座標 j 上における視差値の分布を調べることができる。

U-disparity 画像を $I_{u\Delta}$ とし、 $I_{u\Delta}$ において横座標 j 上における視差値 d の頻出度を $I_{u\Delta}(d, j)$ と表したとき、 $I_{u\Delta}(d, j)$ の値は、

$$I_{u\Delta}(d, j) = \sum_{v=0}^H \delta_{I_\Delta(v, j), d} \quad (17)$$

となる。ここで H は I_Δ の高さである。すべての j について式 (17) を用いて $I_{u\Delta}(d, j)$ を求めることで、U-disparity 画像が求まる。

3.3 Hough 変換

直線検出に用いる Hough 変換 [1] について述べる。Hough 変換はノイズに強く、検出の対象となる直線が点線状になっていても検出できるため、V-disparity 画像や U-disparity 画像のような特徴点が連続して出ていない画像からでも頑健に直線が検出可能な手法である。本研究では、Hough 変換の投票の時に V-disparity 画像、及び U-disparity 画像における視差値の頻出度を重みとして投票値に加える。

点 (v, d) を通る直線 $\rho = v \cos \theta + d \sin \theta$ を示す θ と ρ の組み合わせのパラメタ空間に投票される投票値 n_v は、V-disparity 画像において高さ v 、視差値 d の頻出度を $I_{v\Delta}(v, d)$ としたとき、

$$n_v = I_{v\Delta}(v, d) \quad (18)$$

となる。この重みをつけることにより水平方向に大きく映る平面を示す直線をより頑健に検出することができる。

また、U-disparity 画像についても同様に、点 (d, u) を通る直線 $\rho = d \cos \theta + u \sin \theta$ を示す θ と ρ の組み合わせのパラメタ空間に投票される投票値 n_u は、 u 番目の垂線において視差値 d の頻出度を $I_{u\Delta}(d, u)$ としたとき、

$$n_u = I_{u\Delta}(d, u) \quad (19)$$

となる。この重みをつけることにより垂直方向に大きく映る平面を示す直線をより頑健に検出することができる。

3.4 床面の抽出

本節では、3.3 節で述べた Hough 変換を用いて 3.1 節と 3.2 節で述べた V-disparity 画像と U-disparity 画像から床面を抽出する方法について述べる。抽出する床面は、カメラの光軸に対してほぼ平行な平面を考える。

3.4.1 床面の視差を示す直線検出

対象となる平面は 2.3 節でも述べたように、視差値 d と高さ V を用いて、式 (11) で求めることができる。床面を撮影した場合、式 (11) からわかるように、画像の下部のほうが視差値が大きくなり、画像の上部に向かうにしたがって視差値は小さくなる。さらに、床面を示す直線が $d = 0$ のとき、 $Z_w = \infty$ であり、床面の 3 次元空間上で仮定の無限遠点を示す。2.2 節で述べたカメラ座標系の Z_l または、 Z_r と Z_w がなす角度が小さい場合、この仮定の無限遠点は画像座標で $0 < v$ となる。つまり、V-disparity 画像において、床面を示す直線は V 軸上で交わる。これらのことから床面を示す直線を Hough 変換によって V-disparity 画像から検出する際に、直線のパラメタである θ のとりうる値域を、 $\frac{\pi}{2} \leq \theta \leq \pi$ と制限することができる。

3.4.2 側壁の視差を示す直線検出

U-disparity 画像を用いることにより、V-disparity 画像では求められなかった式 (14) や式 (15) で示されるような平面を求めることができる。また、本研究で対象とする側面の壁は、ステレオカメラの両側に存在する。世界座標系を用いて $X_w = l$ で表す左側の壁は、 $l < 0$ となり基線長 b に対して $b < |l|$ が成り立つとき、式 (15) における d の係数は、負の値となる。つまり、左側の壁を表す直線のパラメタは、 $0 \leq \theta \leq \frac{\pi}{2}$ の値をとる。また、世界座標系を用いて $X_w = r$ で表す右側の壁は、 $r > 0$ となり基線長 b に対して $b < |r|$ が成り立つとき、式 (15) における d の係数は、正の値となる。つまり、右側の壁を示す直線のパラメタは、 $\frac{3\pi}{2} \leq \theta \leq 2\pi$ の値をとる。

3.4.3 境界線の検出

前節までに述べた手法を用いて、床面の視差値を示す直線と側面の壁の視差値を示す直線から床と側面の境界線を推定する。床面と壁の境界は 3 次元空間上で同じ位置にあたるので、境界における視差値は床面と壁で等しくなる。このことを利用し、検出した各直線から視差値が等しくなる座標を床面と壁の境界として検出する。床面の視差値を示す直線のパラメタが、 θ_v, ρ_v であり、壁の視差値を示す直線のパラメタが θ_u, ρ_u となったとき、

$$\frac{\rho_v - v \cos \theta_v}{\sin \theta_v} = \frac{\rho_u - u \sin \theta_u}{\cos \theta_u} \quad (20)$$

を満たす (u, v) が床面と壁の境界である。この境界を左側の壁と右側の壁で検出し、その境界線で囲まれた画素を床面として抽出をおこなう。

4 抽出実験

本節では、前節までに述べた手法を用いて床面の抽出をおこなう。実験に使用するカメラは Logicool 製 Webcam C200 を 2 台使用する。左右の画像は横 581[pixel] × 幅 410[pixel] である。また、基線長 b は、55[mm] である。本実験では、図 2 と図 3 に示したステレオ画像に対し、正確度の異なる視差画像を入力として平面の抽出をおこなう。



図 2 左画像



図 3 右画像

4.1 理想的な視差画像に対する実験

本手法の正当性を確認するために、理想的な視差画像に対して実験をおこなう。図 2 と図 3 に対し、手動で視差画像に対して本手法を用いて床面と側面の壁の視差値を直線のパラメタを求めた結果が表 1 である。

表 1 理想的な視差画像に対する検出された直線のパラメタ

検出した平面	θ [度]	ρ [pixel]
床面	171	17
左側の壁	81	64
右側の壁	279	53

4.2 ノイズの入った視差画像に対する実験

この実験ではノイズの入った視差画像を入力として抽出をおこなう。図 4 で示した図は、手動で視差画像と動的計画法 [4] により左右の画素の対応付けをおこなった視差値の平均を表した画像である。このように、対応付けのアルゴリズムの正確度により正しくない視差値が表れることがある。このノイズのある視差画像を入力として、V-disparity 画像、U-disparity 画像を求め、Hough 変換をおこない平面の視差を表す直線を検出した結果が、図 5 と図 6 である。さらに、図 7 で検出された直線から壁と床の境界を検出した結果を示し、図 8 で、V-disparity 画像から、 $\theta = 0$ の直線で表せるカメラの基線と平行で床面に対して垂直な平面を検出し、それを床の終端点としたときの抽出結果を示す。表 2 で検出された直線のパラメタを示す。



図 4 視差画像



図 5 検出した直線

4.3 抽出結果の考察

実験では、手動で求めた理想的な視差画像と軽度のノイズを含んだ視差画像に対して提案した手法を試みた。4.1 節でおこなった理想的な視差画像の抽出結果を基準の床面領域



図 6 側面の壁の視差値を示す直線



図 7 壁と床面の境界



図 8 抽出したされた床面

表 2 ノイズのある視差画像に対する検出された直線のパラメタ

検出した平面	θ [度]	ρ [pixel]
床面	173	14
左側の壁	84	44
右側の壁	276	36

としたときに、図 4 を入力とした結果、本手法により床面領域の 97% の画素を抽出することができた。このことにより軽度のノイズを含んだ視差画像に対し、提案した手法を用いることでノイズの影響を抑え、床面を抽出することができたといえる。また、視差画像の正確度により、V-disparity 画像および U-disparity 画像の値が変化し、Hough 変換による直線検出に影響が出ることがわかった。

5 結論

本研究では、ステレオカメラから得られた視差画像から V-disparity 画像と U-disparity 画像を作成し、床面視差を示す直線と側面の壁の視差を示す直線を検出することで床面と壁の境界を求め、床面を抽出する手法を提案した。本研究の成果として、U-disparity 画像から側面の壁の視差を示す直線を検出することができ、この直線を用いて壁と床面の境界を推定することができた。また、ノイズをもった視差画像からでも推定された境界を用いて床面を抽出することができ、本手法の有用性を示した。

参考文献

- [1] Duda R.O. and Hart, P. E., "Use of the Hough transformation to detect lines and curves in pictures," *Comm. of the ACM*, Vol. 15, no. 1, pp. 11–15, 1972.
- [2] Raphael Labayrade, Didier Aubert, Jean-Philippe Tarel, "Real Time Obstacle Detection in Stereovision on Non Flat Road Geometry Through "V-disparity" Representation," *Proceedings of IEEE Intelligent Vehicle Symposium*, June, 2002.
- [3] 井口征士, 佐藤宏介, 三次元画像計測, 昭晃堂, 東京, 1990.
- [4] 鈴木高彦, "ステレオカメラによる 3 次元計測システムの開発," 中央大学大学院理工学研究科修士論文, 2004.