

L角対を用いた6自由度復元と物体認識

深田陽司

三菱電機産業システム研究所
〒661 尼崎市塚口本町8-1-1

2本のエッジで構成されるL角を特徴として採用し、単眼視による物体の6自由度復元・認識を行う手法を述べる。注目したL角のカド点に視線を移動し、仮想的な画像を作成する。仮想画像上での画像L角と物体L角の対応を仮定することによって、仮想画像L角の角度が物体座標系のカメラ座標系に対する2つの回転角にのみ依存する拘束式を導出する。3つのL角の対応による3つの拘束曲線から、それらの交点として回転角を算出し仮説を生成する。あらゆる可能な仮説を生成した後、互いに適合する仮説を統合しクリークを抽出する。クリークに属する仮説の回転角と頂点に対する解釈から、平行移動を算出する。

Recovery of 6 Degrees of Freedom and Object Recognition using L-corners

Youji FUKADA

Industrial Electronics & Systems Development Lab., Mitsubishi Electric Co.
8-1-1, Tsukaguchi-honmachi, Amagasaki 661, Japan

In this paper, a method that recovers 6 degrees of freedom and recognizes objects using a perspective image is described. L-corners composed of 2 edges are adopted as features. First, a constraint equation regarding an image angle is derived by hypothesizing the correspondence between an image L-corner and an object one. Rotational angles are determined as the intersections of these constraint curves. Every possible hypothesis is generated and hypotheses that are compatible with each other are gathered to cleaks. Hereafter, 3 translational parameters are calculated using those cleaks.

1 まえがき

3次元情報は一枚の画像にも含まれており、単眼視による物体の3次元構造復元・認識に関する研究が数多くなされている[1]~[4]。認識問題に関しては、仮説検証[6]、ハフ変換[5]、両者の併用[2],[4]や最近ではコネクショニスト的アプローチ[3]等の手法が研究されている。このうち仮説検証は逐次処理であるが他は局所並列であり、特にコネクショニスト法は脳の処理メカニズムをとりいれている。

3次元構造のうち文献[1]では3自由度(回転3)、[2]では6自由度(回転3と平行移動3)、[3],[4]では4自由度(回転1と平行移動3)復元を扱っている。復元・認識に使用される特徴は様々であるが、エッジはそのうちで最も典型的な特徴であり、これらの研究でもエッジ単体[4]もしくは2本のエッジで構成されるL角特徴[1]~[3]を利用している。

文献[1]では注目する画像L角のカド点に視線を移動することによって、物体L角と中心投影された画像L角との関係を基本的で強力な拘束式として導出している。しかしこの式は物体L角がのっている面の法線ベクトルに関する拘束であり、複数の面からの手がかり情報を統合するためには、直接的に適用することは困難である。

一方文献[2]では、画像L角が物体座標系の2つの回転角の関数として表せることを明かにしている。この事実は上に述べた困難を解決する。しかしながら、この研究では正射投影像とみなせる前提条件が必要である。ところで、2つの回転角を決定するためには1つの画像頂点で2つ以上の画像L角が必要であるが、3面頂点からなる物体を対象にすると、多くの画像頂点では画像L角は1つしか得られない。そこで、この研究では頂点間に仮想的なエッジを想定し、利用できるL角を増やすことを提案している。

本稿では視線を移動することによって、中心投影のもとで画像L角が物体座標系の2つの回転角の関数として表せることを明かにする。このことを利用して、画像L角と物体L角の対応(解釈)を仮定した時の回転角を算出する。つまり、仮説を生成する。あらゆる可能な仮説を生成した後、互いにその解釈と算出回転角が矛盾しない仮説を統合してクラスタを生成し、平行移動を算出する。大きなクラスタから順に、その算出値を用いて物体像を画像上に投影し、検証することによって復元・認識を行う。

2 座標系と自由度復元問題

2.1 座標系の設定

本稿では4つの座標系を設定する。即ち、カメラ座標系 $O_c X_c Y_c Z_c$, 仮想カメラ座標系 $O_v X_v Y_v Z_v$, モデル座標系 $O_n X_n Y_n Z_n$, 画像座標系 $O_i UV$ である。カメラ座標系及び仮想画像座標系において撮像された画像を、画像及び仮想画像と呼ぶ。本節ではこれらの座標系間の関係を定式化し、自由度復元問題を明確にする。

これらの座標系は全て左手系で定義されており、それぞれの座標系における点やベクトルの表現を区別するために上に述べた座標系の順に $\langle \dots \rangle$, $\langle \langle \dots \rangle \rangle$, $[\dots]$, $\{\dots\}$ の表記を用いる。本稿では平行移動も線型で扱えるように式の展開には同次座標表現を使用し、単純に点やベクトルを表す時には普通座標表現を用いることとする。回転は全て軸の正方向からみて時計方向とする。

2.2 (仮想)カメラ座標系と画像座標系の関係

まず本稿で使用するカメラパラメータを定義する。これらは全て既知とする。

s_x : 1画素あたりの X_c 軸方向の長さ (mm/pixel)

s_v : 1画素あたりの Y_c 軸方向の長さ (mm/pixel)

f : 焦点距離 (mm)

レンズ中心をカメラ座標系の原点とし、画像面は光軸 (Z_c 軸) に直交して前方の焦点距離の位置に設置されているとする。光軸と画像面との交点を画像原点とし、 U 軸と X_c 軸、 V 軸と Y_c 軸は平行とする。

(仮想)カメラ座標系と画像座標系は、点 $P = [X_c, Y_c, Z_c]$ が $p = [x_c, y_c, f] = \{u, v\}$ に中心投影されたとすると、(1)によって関係づけられる。ここで $[P]$ は $\langle P \rangle$ または $\langle P \rangle$ のどちらかを表す。

$$\begin{aligned} x_c &= f X_c / Z_c, & y_c &= f Y_c / Z_c, \\ u &= x_c / s_x, & v &= y_c / s_v. \end{aligned} \quad (1)$$

2.3 カメラ座標系とモデル座標系の関係

図1に示すように、モデル座標系が Y_c 軸, X_c 軸, Z_c 軸まわりに β ($-\pi \leq \beta < \pi$), α ($-\pi/2 \leq \alpha \leq \pi/2$), γ ($-\pi \leq \gamma < \pi$) 回転した後 $T = (T_x, T_y, T_z)$ 平行移動しているとする。この時、点 P 及びベクトル E は回転行列を R 、平行移動行列を T とすると(2)の関係を持つ。

$$\langle P \rangle = [P] R T, \quad \langle E \rangle = [E] R \quad (2)$$

行列の具体的表現は次頁下段に示す。

2. 4 カメラ座標系と仮想カメラ座標系の関係

仮想カメラ座標系は、画像上の注目点 $p = \{u, v\}$ が仮想画像の原点に撮像されるよう、カメラ座標系を Y_c 軸まわりに ν 、 X_c 軸まわりに μ 回転して得られる座標系である (図2参照)。ここで、

$$\begin{aligned} \nu &= \arctan(us_x/f) \\ \mu &= \arctan(-vs_u/\sqrt{f^2+(us_x)^2}) \end{aligned} \quad (3)$$

この時、点P及びベクトルEは回転行列を R_{NN} とすると(4)の関係を持つ。

$$\langle P \rangle = \langle P \rangle R_{NN}^{-1}, \quad \langle E \rangle = \langle E \rangle R_{NN}^{-1} \quad (4)$$

ここで $R_{NN} = \begin{pmatrix} \cos \nu & 0 & -\sin \nu & 0 \\ \sin \mu \sin \nu & \cos \mu & \sin \mu \cos \nu & 0 \\ \cos \mu \sin \nu & -\sin \mu & \cos \mu \cos \nu & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$

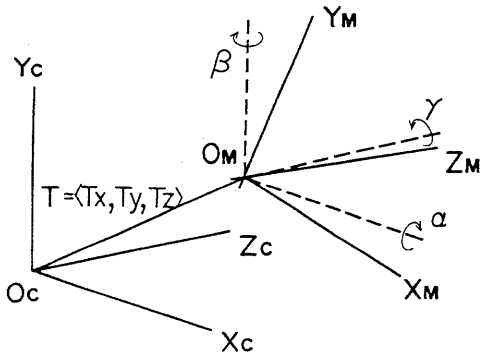


図1. 座標系の設定

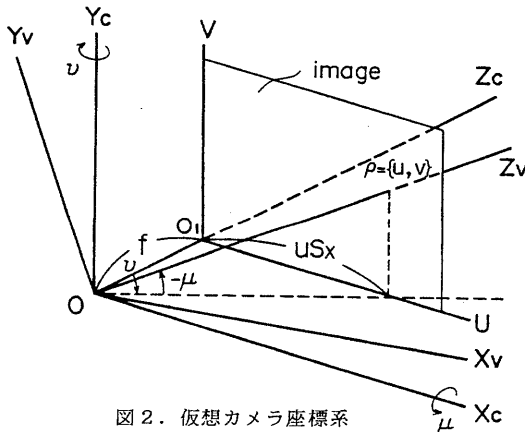


図2. 仮想カメラ座標系

2. 5 仮想カメラ座標系とモデル座標系の関係

モデル座標系が Y_u 軸, X_u 軸, Z_u 軸まわりに ψ ($-\pi \leq \psi < \pi$), ϕ ($-\pi/2 \leq \phi \leq \pi/2$), ξ ($-\pi \leq \xi < \pi$) 回転した後 $T_u = (t_x, t_u, t_z)$ 平行移動しているとする。この時、点P及びベクトルEは回転行列を R_u 、平行移動行列を T_u とすると(5)の関係を持つ。

$$\langle P \rangle = \langle P \rangle R_u T_u, \quad \langle E \rangle = \langle E \rangle R_u \quad (5)$$

ここで、 T_u, R_u は T, R とそれぞれ同型である。

2. 6 自由度復元問題

カメラ座標系に対するモデル座標系の変換を表す変数 $\alpha, \beta, \gamma, T_x, T_u, T_z$ が本稿で算出すべきパラメータであり、(1)~(5)の関係式を用いて解く。

復元手法は多面体を対象としており、物体はモデル座標系における頂点(座標)とエッジ(ベクトル)の集合として表現されている。このような多面体を中心投影した画像を1枚使用し、物体の6自由度を復元する。即ち、単眼視による復元である。

復元に使用する特徴はエッジの対からなるL角である。画像L角と物体L角の対応から導出された拘束を用いて回転角を算出し、その後平行移動を算出する。

3 仮想カメラ座標系における自由度復元

3. 1 ベクトルの中心投影

2点P, Qのモデル座標表現を $X_p = [X_p, Y_p, Z_p]$, $X_q = [X_q, Y_q, Z_q]$ 、それらの仮想画像上の像を $p = \{u, v\}$, $q = \{u_q, v_q\} = \{X_q, Y_q, f\}$ とし、ベクトルPQ, pqをF, Gとする。

変換行列 $R_u T_u$ を以下のように表す。

$$R_u T_u = \begin{pmatrix} r_{u1} & r_{u2} & r_{u3} & 0 \\ t_x & t_u & t_z & 1 \end{pmatrix}$$

ここで、 r_{ui} は R_u の左上 3×3 行列 ($(R_u)_{3 \times 3}$ と表記する) の i 列を要素とするベクトル、0 はゼロベクトル、肩付き T は転置を表す。

すると、(5)と(1)より p は以下の式

$$\begin{aligned} x_o &= f(X_p \cdot r_{u1} + t_x) / (X_p \cdot r_{u3} + t_z) \\ y_o &= f(X_p \cdot r_{u2} + t_u) / (X_p \cdot r_{u3} + t_z) \end{aligned} \quad (6)$$

で算出され、q も同様である。

$$R = \begin{pmatrix} \cos \beta \cos \gamma - \sin \alpha \sin \beta \sin \gamma & \cos \beta \sin \gamma + \sin \alpha \sin \beta \cos \gamma & -\cos \alpha \sin \beta & 0 \\ -\cos \alpha \sin \gamma & \cos \alpha \cos \gamma & \sin \alpha & 0 \\ \sin \beta \cos \gamma + \sin \alpha \cos \beta \sin \gamma & \sin \beta \sin \gamma - \sin \alpha \cos \beta \cos \gamma & \cos \alpha \cos \beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad T = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ T_x & T_u & T_z & 1 \end{pmatrix}$$

ここで図3に示すように、p が仮想画像原点と一致している場合、 $x_0=y_0=0$ より、

$$X_p \cdot r_{u1} + t_x = 0 \quad X_p \cdot r_{u2} + t_u = 0 \quad (7)$$

だから G の仮想カメラ座標表現は k を比例定数とすると

$$\begin{aligned} \langle G \rangle &= \langle x_0 - x_0, y_0 - y_0, 0 \rangle \\ &= \langle f [F] \cdot r_{u1} / (X_0 \cdot r_{u3} + t_z), \\ &\quad f [F] \cdot r_{u2} / (X_0 \cdot r_{u3} + t_z), 0 \rangle \\ &= k [F] \begin{pmatrix} r_{u1}^T & r_{u2}^T & 0^T \end{pmatrix} \end{aligned} \quad (8)$$

即ち、仮想画像上の原点を始点とするベクトルの仮想カメラ座標表現は平行移動によらず、回転行列の第1及び2列にのみ依存する。

3.2 ϕ と ψ の決定

回転行列を $R_0 = R_Y R_Z$ と分解すると、

$$\begin{pmatrix} R_{YX} \end{pmatrix}_{3 \times 3} = \begin{pmatrix} R_Z \end{pmatrix}_{3 \times 3} = \begin{pmatrix} \cos \psi & \sin \phi \sin \psi & -\cos \phi \sin \psi \\ 0 & \cos \phi & \sin \phi \\ \sin \psi & -\sin \phi \cos \psi & \cos \phi \cos \psi \end{pmatrix} \begin{pmatrix} \cos \xi & \sin \xi & 0 \\ -\sin \xi & \cos \xi & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

r_{2i} を $(R_Z)_{3 \times 3}$ の i 列を要素とするベクトルとすると

$$r_{u1}^T = (R_{YX})_{3 \times 3} r_{Z1}^T, \quad r_{u2}^T = (R_{YX})_{3 \times 3} r_{Z2}^T$$

と表せるから(8)は更に(9)と展開される。

$$\langle G \rangle = k [F] \begin{pmatrix} R_{YX} \end{pmatrix}_{3 \times 3} \begin{pmatrix} r_{Z1}^T & r_{Z2}^T & 0^T \end{pmatrix} \quad (9)$$

仮想画像上に存在する、原点から2点 q_1, q_2 へのベクトル G_1 と G_2 の内積は、 r_{YX3} を $(R_{YX})_{3 \times 3}$ の3列を要素とするベクトルとすると付録Aより、

$$G_1 \cdot G_2 / k_1 k_2$$

$$= [F_1] [F_2]^T - ([F_1] r_{YX3}^T) ([F_2] r_{YX3}^T)$$

と展開されるから、 G_1 と G_2 の成す角を θ とし F_1, F_2 上の単位ベクトルを E_1, E_2 とすると、

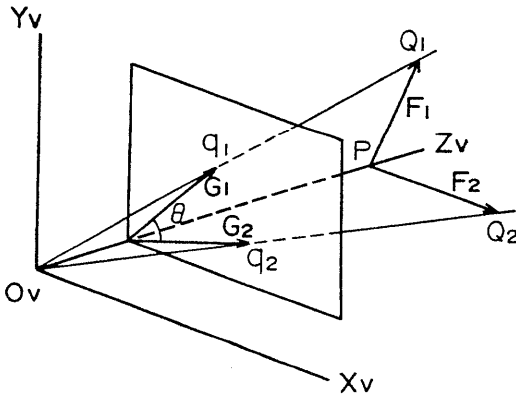


図3. L角の中心投影

$$\cos \theta = G_1 \cdot G_2 / |G_1| |G_2|$$

$$= \frac{[E_1] [E_2]^T - ([E_1] r_{YX3}^T) ([E_2] r_{YX3}^T)}{\sqrt{1 - ([E_1] r_{YX3}^T)^2} \sqrt{1 - ([E_2] r_{YX3}^T)^2}} \quad (10)$$

(10)は仮想画像上で原点を始点とする2つのベクトルの成す角は、 ϕ と ψ の2つの変数にのみ依存していることを示している。

故に、仮想画像原点をカド点とする2つの仮想画像L角と物体L角の対応がとれれば、2つの回転角 ϕ と ψ を決定することができる。ただし、(10)は解析的に解くことはできないので、 ϕ と ψ を与えて得られる拘束曲線の交点として、回転角が決定する。

3.3 ξ の算出

F の ξ 回転する前のベクトルを F' としその像を G' とすると、 G' の仮想座標表現は(9)において $\xi=0$ だから

$$\langle G' \rangle = k [F] \begin{pmatrix} R_{YX} \end{pmatrix}_{3 \times 3} I^T \quad (11)$$

ここで I^T の行ベクトルは $(1,0,0), (0,1,0), (0,0,0)$ 。

前項において、各エッジの対応(解釈)と ϕ, ψ が決定しているから $[F]$ と R_{YX} は既知であり、(11)は k を除いて計算できる。

$$k [F] \begin{pmatrix} R_{YX} \end{pmatrix}_{3 \times 3} = (a_1, a_2, a_3) \text{ とおくと、}$$

$$\langle G' \rangle = \langle a_1, a_2, 0 \rangle$$

$$\langle G \rangle = \langle a_1 \cos \xi - a_2 \sin \xi, a_1 \sin \xi + a_2 \cos \xi, 0 \rangle$$

G と G' の内積をとると、

$$G \cdot G' / |G| |G'| = \cos \xi \quad (12)$$

故に、 ξ が算出できる。 ξ の符号は G と G' のベクトル積より決定できる。

4 カメラ座標系における自由度復元

4.1 回転角の算出

(4)と(5)のベクトルに関する式より、

$$\langle E \rangle = \langle E \rangle R_{NN} = [E] R_0 R_{NN} \quad (13)$$

(2)との比較より、

$$R = R_0 R_{NN} \quad (14)$$

画像上のあるカド点 p に注目する時、 R_{NN} は(3)より決定し、画像上の他の頂点の仮想画像座標値は(4)、(1)より算出できる。そこで、p における2つ以上の仮想画像L角と物体L角の対応がとれれば、3における展開により回転行列 R_0 が決定するので、(14)の右辺が決定する。すると、 α, β, γ はその存在範囲が制限されているので、これらの回転角は R の2行、3列より一意に決定する。

4. 2 平行移動の算出

次に、決定された回転を用いて平行移動を算出する。2点 P, Q の画像上の像を $p = \{u_0, v_0\} = \langle x_0, y_0, f \rangle$, $q = \{u_0, v_0\} = \langle x_0, y_0, f \rangle$ とする。P, Q のモデル座標表現を $X_p = [X_p, Y_p, Z_p]$, $X_q = [X_q, Y_q, Z_q]$ とし、R T を以下に示すように表す。

$$RT = \begin{pmatrix} r_1^T & r_2^T & r_3^T & 0^T \\ T_x & T_y & T_z & 1 \end{pmatrix}$$

ここで、 r_i は (R) 3×3 の i 列を要素とするベクトルを表す。

すると p, q は (6) と同型の式で算出される。これらの式におけるそれぞれ2つの式の比をとることにより、

$$x_0 (X_p \cdot r_2 + T_y) = y_0 (X_p \cdot r_1 + T_x)$$

$$x_0 (X_q \cdot r_2 + T_y) = y_0 (X_q \cdot r_1 + T_x) \quad (15)$$

(15)の2つの式より、

$$T_x = (A1 (X_p \cdot X_q) \cdot r_2 + (A2 X_q \cdot r_1 - A3 X_p \cdot r_1)) / (A3 - A2)$$

$$T_y = y_0 (X_p \cdot r_1 + T_x) / x_0 - X_p \cdot r_2$$

$$T_z = f (X_p \cdot r_1 + T_x) / x_0 - X_p \cdot r_3 \quad (16)$$

$$\text{or } T_y = y_0 (X_q \cdot r_1 + T_x) / x_0 - X_q \cdot r_2$$

$$T_z = f (X_q \cdot r_1 + T_x) / x_0 - X_q \cdot r_3 \quad (16')$$

ここで、 $A1 = x_0 x_q$, $A2 = x_0 y_q$, $A3 = x_q y_0$ 。

5 復元と認識

5. 1 L角特徴

物体の多くは3面頂点で構成されているので、自分自身のオクルージョンのためエッジが2本しか見えない頂点や、本来3本のエッジが見えるはずなのに、画像処理の結果2本しか得られない頂点が画像には数多く出現する可能性がある。これらの頂点では、実在するエッジで構成されるL角だけを考えると1つしかないで、拘束式が不足し回転角を算出することができない。しかし文献[7]におけるエッジと同様、これらの頂点(L角)も復元・認識には大きく寄与しているし、画像がこのような頂点ばかりで構成されているとき自由度の復元はできなくなってしまうが、我々人間には所定の物体が見える例もある。

そこで、図4に示すようにL角の頂点を結んでできる、仮想的なエッジ(以下においては、これらを仮想エッジ、実在するエッジを実エッジと呼ぶ)を考える。すると、実エッジが2本しか存在しない頂点でもL角を多数つくることができ、復元に利用できる。ただし、2本とも仮想エッジのL角まで採用するとあまりにも

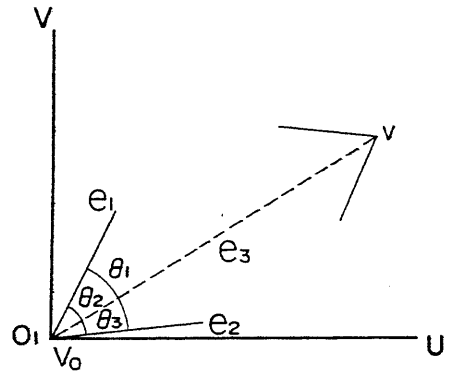
その数が多くなってしまうので、本稿では画像L角も物体L角も少なくとも1本は実エッジで構成されるものをL角として採用する。

5. 2 仮説の生成

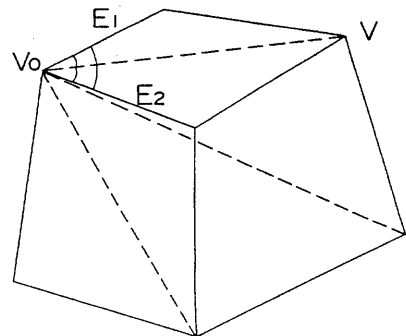
3と4で展開した議論は、仮想画像L角と物体L角の対応が2つ以上とれば、仮想画像座標系に対するモデル座標系の回転が決定した後、画像座標系に対するモデル座標系の変換行列が決定する、つまり6自由度の復元ができることを示している。

しかし、画像に関してなんらの情報も与えられないので、仮想画像L角がどの物体L角に対応しているかは不明である。そこで、仮想画像L角と認識したい物体のL角の対応を仮定し、これら仮定の組に対して回転角を算出する。つまり、エッジ(もしくは頂点)の対応とその時の回転角を仮説としてたてるのである。

仮説の生成は次のようにして行われる。まず、画像において2本以上の実エッジをもつ頂点 v_i に注目し、仮想画像を作成する(生成方法の記述を簡略にするため、仮想画像を単に画像とする)。



a. 画像



b. 物体

図4. 仮想エッジと仮想L角

画像頂点における仮説の生成方法

1) 異なる3本の画像実エッジの組に対して

v_i を物体頂点 V_i の像と仮定する。 V_i におけるあらゆる異なる3本の物体実エッジの組を対応させ、それぞれ(10)による3本の拘束曲線の交点として、 ϕ と ψ を決定する。その後、 ξ を算出する。

これを全ての物体頂点を対象にして実行する。

全ての画像エッジの組に対して以上を実行する。

2) 異なる2本の画像実エッジと他の画像L角頂点へ

の仮想画像エッジの3本のエッジの組に対して

2-1) v_i を物体頂点 V_i の像と仮定する。 V_i における3本の物体実エッジとの対応による回転角の算出。

これを全ての物体頂点を対象にして実行する。

2-2) v_i を物体頂点 V_i の像と仮定する。 V_i における2本の物体実エッジと1本の仮想物体エッジとの対応による回転角の算出。

これを全ての物体頂点を対象にして実行する。

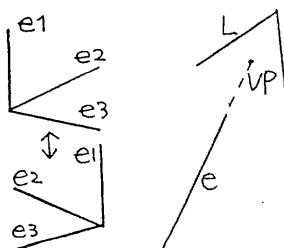
全ての画像エッジの組に対して以上を実行する。

2-1)は、実エッジにも拘らず画像処理の結果抽出できなかったエッジを考慮している。上の手続きによって回転角 ϕ, ψ, ξ が算出できても、これらは必ずしも適切とはいえず、この時点で得られている情報を利用してチェックする。即ち、エッジの対応が分かっているので、算出された回転角 (ϕ, ψ, ξ) による物体エッジの像が仮想画像上に計算できる。この像と対応する仮想画像エッジに矛盾が生じないかどうかを回転角の適切性条件により判定する(図5参照)。

回転角の適切性条件

- 1) 3本のエッジの並びの順序(鏡像性)
- 2) エッジの消失点の存在位置
- 3) エッジや頂点の可視性

適切な回転角 (ϕ, ψ, ξ) が算出できるとき、以下の形式の仮説が生成される。



a. 鏡像

b. 消失点の存在位置

図5. 回転角の適切性

$$H(e_i, e_j, e_k, E_i, E_j, E_k, \alpha, \beta, \gamma)$$

ここで、 e は画像エッジ、 E は物体エッジを表わし、各順に (e_i と $E_i, \dots$) 対応している、即ち解釈されていることを意味している。

5. 3 仮説の統合

あらゆる仮説の集合から、矛盾のない仮説(適合仮説)を統合しクラスタを抽出することによって、物体の認識・復元を行う。仮説の総数は膨大になるが、仮説はその生成方法で述べたように階層的に並んでいるので、統合はこの階層性を利用して行われる。

5. 3. 1 第1レベルの統合

まず、2つの仮説の適合性について述べる。仮説を

$$H_x(e_x^1, e_x^2, e_x^3, E_x^1, E_x^2, E_x^3, \alpha_x, \beta_x, \gamma_x)$$

$$H_y(e_y^1, e_y^2, e_y^3, E_y^1, E_y^2, E_y^3, \alpha_y, \beta_y, \gamma_y)$$

とすると、以下の条件を満たすならば H_x と H_y は適合するという。

仮説の適合条件

$$1) e_x^i = e_y^k \text{ なら } E_x^i = E_y^k \quad (i, k=1-3)$$

$$2) e_x^i \neq e_y^k \text{ なら } E_x^i \neq E_y^k \quad (i, k=1-3)$$

$$3) |\theta_x - \theta_y| \leq d\theta \quad (\theta = \alpha, \beta, \gamma)$$

適合する仮説の統合を画像頂点毎に行う。画像頂点毎に生成の時点で分割されている仮説集合をまず解釈毎に(即ち、物体頂点毎に)集めて分割する。すると、1つの画像頂点に対する物体頂点の解釈は同じでなければならないので、異なる分割に属する2つの仮説は適合することはなく、これらの分割内での極大適合仮説集合(以下、極大クリークと呼ぶ)[5]が当該の画像頂点における仮説集合の極大クリークとなっている。

解釈毎に仮説を集める操作は仮想的に行われており、その結果得られる集合を G とすると、 G における仮説の並びは階層性が保たれている。つまり、 G は画像エッジの組 i に対する仮説群 G_i 毎に並べられている。 G_i は、可能な物体エッジの組に対する仮説から成っており、各物体エッジの組は複数の仮説をもちうる。

G における極大クリークは以下のようにして生成される。まず、1つの仮説 H に注目する。 H が G_i に属しているとする、 H は一意性の拘束より G_i に属している仮説とは適合しないので、 H と適合する仮説は G_i 以外に属していなければならない。そこで、 G_i 以外に属する仮説のうちで H と適合する仮説をすべて集める。この際、1つの物体エッジの組であるとの解釈に対して複数の仮説が存在する時は、角度が H と最

も近似している適合仮説を採用する。この集合に対して、文献[5] のアルゴリズムを改良した極大クリーク生成アルゴリズム（付録B）を施す。

生成された極大クリークは以下の形式をもっている。

$$C(\alpha, \beta, \gamma, v, V)$$

ここで回転角は極大クリークに属する仮説の角度の平均である。 v は画像頂点の組 $(v_1, v_2, \dots)$ を、 V は解釈された物体頂点の組 $(V_1, V_2, \dots)$ を表しており、 v の i 番目の画像頂点が V の i 番目の物体頂点と解釈されていることを意味している。頂点に対する解釈は仮説のエッジに対する解釈から容易に変換できる。

5. 3. 2 第2レベルの統合

2つの極大クリークを

$$C_x(\alpha_x, \beta_x, \gamma_x, v_x, V_x), C_y(\alpha_y, \beta_y, \gamma_y, v_y, V_y)$$

$$v_x = (v_x^1, v_x^2, \dots), V_x = (V_x^1, V_x^2, \dots)$$

$$v_y = (v_y^1, v_y^2, \dots), V_y = (V_y^1, V_y^2, \dots)$$

解釈の個数はそれぞれ m, n 個

とすると、以下の条件を満たすならば C_x と C_y は適合するという。

極大クリークの適合条件

$$1) v_x^i = v_y^k \text{ なら } V_x^i = V_y^k \quad (i=1-m, k=1-n)$$

$$2) v_x^i \neq v_y^k \text{ なら } V_x^i \neq V_y^k \quad (i=1-m, k=1-n)$$

$$3) |\theta_x - \theta_y| \leq d\theta \quad (\theta = \alpha, \beta, \gamma)$$

極大クリークの集合から、互いに適合する極大クリークを統合し、極大クラスタを生成する[5]。極大クラスタは極大クリークと同型であり、回転角は属する極大クリークの平均である。

5. 3. 4 選択と検証

解釈されている頂点の個数の多い極大クラスタの順に平行移動 T を算出し、 RT による物体の投影像を作成し検証する。

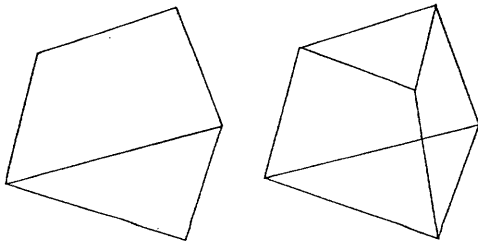


図6. a. 画像1

b. 画像2

6 実験

正三角錐台を対象物体として、提案した手法を検証するための実験を行った。物体を $(\alpha, \beta, \gamma) = (30^\circ, 60^\circ, 0)$, $T = (100, 0, 1000)$ 移動し、中心投影像を作成した。 $f=15.0$, $s_x=0.015$, $s_y=0.014$ である。仮説及び極大クリーク適合のための角度のしきい値は $d\theta = 6^\circ$ である。回転角適切性判定は 1) のみ適用しており、以下での仮説候補とは判定前の仮説の数である。

なお、(10)において (ϕ, ψ) を解とすると、 $(-\phi, \psi + \pi)$ に対して $-r_{v \times s}$ となり、 $\cos \theta$ は同じ値になるので、 ψ ($0 \leq \psi < \pi$), ϕ ($-\pi/2 \leq \phi \leq \pi/2$) のみの仮説を以下では採用している。

実験1：不透明正三角錐台（図6 a）

仮説候補 1911個, 仮説 663個

極大クリーク 47個, 極大クラスタ 2個

(極大クラスタ1)

極大クリーク 5個, 仮説 20個

$$(\alpha, \beta, \gamma) = (29.4^\circ, 60.4^\circ, 0.0)$$

$$T = (99.5, -0.4, 994.6)$$

Tの標準偏差 (3.886, 1.795, 31.922)

(極大クラスタ2)

極大クリーク 5個, 仮説 20個

$$(\alpha, \beta, \gamma) = (29.4^\circ, -178.4^\circ, 0.0)$$

$$T = (202.2, -28.8, 1040.0)$$

Tの標準偏差 (8.607, 3.161, 77.386)

実験2：透明正三角錐台（図6 b）

仮説候補 5198個, 仮説 1649個

極大クリーク 114個, 極大クラスタ 3個

(極大クラスタ1)

極大クリーク 7個, 仮説 35個

$$(\alpha, \beta, \gamma) = (29.7^\circ, 60.3^\circ, 0.0)$$

$$T = (100.9, -0.1, 1006.0)$$

Tの標準偏差 (3.426, 1.248, 23.924)

(極大クラスタ2)

極大クリーク 8個, 仮説 39個

$$(\alpha, \beta, \gamma) = (29.6^\circ, -178.5^\circ, 0.4)$$

$$T = (203.0, -29.6, 1041.1)$$

Tの標準偏差 (9.821, 3.901, 69.120)

(極大クラスタ3)

極大クリーク 3個, 仮説 15個

$$(\alpha, \beta, \gamma) = (-33.3^\circ, 75.7^\circ, 9.3)$$

$$T = (109.7, 0.8, 1074.7)$$

Tの標準偏差 (8.815, 3.905, 62.642)

7 あとがき

単眼視による物体の6自由度復元・認識アルゴリズムについて述べ、検証するための実験を行った。注目している画像特徴に視線を移動することによって、中心投影画像からの回転角復元のための拘束式を導出した。画像特徴と物体特徴の対応を仮定することによって、拘束を用いて仮説を生成する。互いに適合する仮説を統合することによって、復元・認識を行う。

今後は、部分線画における極大クリーク適合条件を拡張し、実画像に対して適用してその効果を検証する予定である。

謝辞

本研究を実施するに際し、機会を与えて頂いた中堀一郎部長に感謝します。なお本研究の基本的アイデアと検討は筆者が ATR視聴覚機構研究所に出向していたときに行ったものであり、同研究所の淀川英司社長に感謝します。

参考文献

- [1] T.Shakunaga and H.Kaneko, "Perspective angle transform : principle of shape from angles," Computer Vision, 3, pp.239-254, 1989.
- [2] D.W.Thompson and J.L.Mundy, "Three-dimensional model matching from an unconstrained viewpoint," Proc. Int. Conf. on Robotics and Automation, pp.208-220, 1987.
- [3] 佐藤, 田村, "パラメータ化モデルを用いた3次元物体認識: Hopfield モデルによる並列探索アプローチ," 情報論, Vol.31, No.1, pp.56-67, 1990.
- [4] 深田, "線特徴からの3次元情報復元と物体認識," PRU90-50, pp.69-76, 1990.
- [5] D.H.Ballard and C.M.Brown, "Computer Vision," Prentice Hall, 1982.
- [6] R.C. Bolles and P. Horaud, "3DPO: a three-dimensional part orientation system," Robotics Research, Vol.5, No.3, pp.3-26, 1986.
- [7] D.G.Lowe, "Perceptual Organization and Visual Recognition," Kluwer Academic Publishers, 1985

付録A: 仮想画像上の角度

$$\langle G \rangle = k [F] (R_{yx})_{3 \times 3} (r_{21}^T \ r_{22}^T \ 0^T) \quad (a1)$$

すると2つのベクトル G_1 と G_2 の内積は、

$$\begin{aligned} G_1 \cdot G_2 / k_1 k_2 &= ([F_1] (R_{yx})_{3 \times 3} (r_{21}^T \ r_{22}^T \ 0^T)) \\ &\quad ([F_2] (R_{yx})_{3 \times 3} (r_{21}^T \ r_{22}^T \ 0^T))^T \\ &= [F_1] (R_{yx})_{3 \times 3} (r_{21}^T \ r_{22}^T \ 0^T) \\ &\quad (r_{21}^T \ r_{22}^T \ 0^T)^T (R_{yx})_{3 \times 3} [F_2]^T \quad (a2) \end{aligned}$$

ここで、 $r_{21} = (\cos \xi, -\sin \xi, 0)$, $r_{22} = (\sin \xi, \cos \xi, 0)$ だから、 $(r_{21}^T \ r_{22}^T \ 0^T) (r_{21}^T \ r_{22}^T \ 0^T)^T = I'$ を (a2) に代入すると、

$$\begin{aligned} G_1 \cdot G_2 / k_1 k_2 &= [F_1] (R_{yx})_{3 \times 3} I' I' (R_{yx})_{3 \times 3}^T [F_2]^T \\ &= [F_1] ((R_{yx})_{3 \times 3} I') ((R_{yx})_{3 \times 3} I')^T [F_2]^T \\ &\quad \text{上式における行列を展開すると、} \\ &\quad ((R_{yx})_{3 \times 3} I') ((R_{yx})_{3 \times 3} I')^T \\ &\quad = (r_{yx1}^T \ r_{yx2}^T \ 0^T) (r_{yx1}^T \ r_{yx2}^T \ 0^T)^T \\ &\quad = I - (0^T \ 0^T \ r_{yx3}^T) (0^T \ 0^T \ r_{yx3}^T)^T \end{aligned}$$

ここで、 I は単位行列である。

故に2つのベクトルの内積は、

$$G_1 \cdot G_2 / k_1 k_2 = [F_1] [F_2]^T - ([F_1] r_{yx3}^T) ([F_2] r_{yx3}^T)^T$$

付録B: 極大クリーク生成アルゴリズム

定理: 仮説集合 S の全ての極大クリークは以下のアルゴリズムによって生成できる (証明は省略)。

ここで、 $S = \{h_0, h_1, \dots, h_N\}$ 、 h_i ($i=0-N$) は仮説、 h_0 と h_i ($i=1-N$) は互いに適合する。

生成アルゴリズム

$S^k = \{h_0, h_1, \dots, h_k\}$ なる仮説集合とする。

step1) $S^1 = \{h_0, h_1\}$ $S_1^1 = \{h_0, h_1\}$ とする。

$k = 2-N$ に対して

stepk) stepk-1 で得られた適合仮説集合を S_m^{k-1} , $m=1-M$ とすると、

操作1) 仮説集合 $S_m^{k-1} + \{h_k\}$, $m=1-M$ の各々に対して

1-1) h_k が S_m^{k-1} の任意の仮説と適合するなら、適合仮説集合 $S_m^{k-1} + \{h_k\}$ を生成する。

1-2) S_m^{k-1} の仮説のうち h_k と適合しない仮説の集合を T_m^{k-1} とすると2つの適合仮説集合 S_m^{k-1} と $S_m^{k-1} - T_m^{k-1} + \{h_k\}$ を生成する。

操作2) 操作1) で生成された適合仮説集合のうちで、部分集合になっている集合を削除し、これらを stepk で生成された適合仮説集合とする。