

# 逆射影変換を用いた対象物位置姿勢認識法の認識精度の検証

## Verification of Estimation Performance of Pose Using Inverse Projection

○山本 太郎 (岡山大)      舟久保 龍希 (岡山大)  
 寇 ギョウ郡 (岡山大)   正 田 宏志 (岡山大)  
 李 想 (岡山大)          正 松野 隆幸 (岡山大)  
 正 見浪 護 (岡山大)

Taro YAMAMOTO, Okayama University, pjmv95a0@s.okayama-u.ac.jp

Ryuki FUNAKUBO, Okayama University

Yejun KOU, Okayama University

Hongzhi TIAN, Okayama University

Xiang LI, Okayama University

Takayuki MATSUNO, Okayama University

Mamoru MINAMI, Okayama University

Requirement for robots has been changed from the work that performed by operators via robots, to the autonomous action took by themselves. Visual servoing is a method to control robots' motion by installing the visual information into the feedback loop. Therefore, it is expected to enable robots to work in the constantly changing environment or unknown environment. However, to recognize target object from images, some degree of predefined knowledge is needed. Therefore, our research group proposed a hand-eye visual servoing system without using them. Models are created firstly from the image projected on the left camera, then after performing inverse projection to 3D space to make search models. Projection of the search models in 3D space are projected to the right camera to determine the position and rotation of target object. In this study, the accuracy of pose estimation of arbitrary-shaped target object by this system is verified.

**Key Words:** Dual-eyes recognition, Genetic Algorithm, Inverse projection

### 1 緒言

ロボットを制御する方式のひとつとしてビジュアルサーボが挙げられる。これは視覚情報をフィードバックループに組み込む方式であり、常に変化する環境や未知環境で働くロボットへの適用が期待されている。現在でもこれを実現するための複数の方法が研究されているが、画像から対象を認識するためには、文献 [1, 2, 3] のようにその対象に関するある程度の事前知識が必要である。文献 [4, 5, 6] のようにエビポラ幾何を用いる場合では、画像中の対応点を探す必要があったりと課題も多い。そこで、我々の研究グループではそれらを用いずにビジュアルサーボを行うシステムを開発している。カメラに映った画像からモデルを作成し、3次元空間へ逆射影を行った後、別のカメラへ再度射影することで対象の位置姿勢を確定する。本研究ではそのシステムによる対象物認識について検証した。

近年のロボットの技術の発展に伴い、その高い正確性からヒューマンエラー対策として、或いは作業効率を求めて、現在では様々な分野においてロボットの導入が進んでいる。また今日の日本では少子高齢化が進行しており、労働力の低下は深刻な問題となっている。減少する労働力を補うひとつの方策としてさらにロボットの需要は高まるものと考えられる。同時にロボットに求められることも、従来のように操縦者がロボットを介して作業を行うものから、変化する環境の中でロボット自身が自律行動をとるものへと変化することが予想される。

そこで著者らの研究グループでは、ロボットが自律行動をとるために必要となる対象の認識および追従について、複眼カメラシステムでの実現を目指し研究を進めている。本報告では事前知識を必要としない対象の認識、追従を行うことの出来るシステムについて述べる。

### 2 システム概要

本研究で述べるロボットシステムに用いている対象物の認識手法は、静止画像認識が基本となる。よって、ここでは、1枚の静止画像に対する認識手法の概要を説明する。

Fig.1 に Projection-based Matching 法の手順の概略図を示す。PA10 に搭載されるカメラは Fig.2 に示すように、2台のカメラを用いる。基準点を手先座標系  $\Sigma_H$  とする。また基準座標系  $\Sigma_W$  とする。左のカメラ座標系を  $\Sigma_{CL}$ 、画像座標系を  $\Sigma_{IL}$ 、右のカメラ座標系を  $\Sigma_{CR}$ 、画像座標系を  $\Sigma_{IR}$  とする。左カメラ画像に映った対象物をモデルとして抽出し、3次元空間における位置/姿勢  $\phi = ({}^{CL}z_{MiC}, {}^{Hx}\theta_M, {}^{Hy}\theta_M)$  を GA(遺伝的アルゴリズム) の遺伝子により決定する。このとき、 ${}^{CL}x_{MiC}$  および  ${}^{CL}y_{MiC}$  は  ${}^{CL}z_{MiC}$  より求められる。このモデルを左カメラ画像から3次元空間へ逆射影し右カメラ画像に再度射影することで、2次元の平面モデルを得る。この平面モデルと右カメラ画像との照合により適合度関数を計算し、評価を行う。そして、空間中の平面モデルの位置/姿勢  $\phi$  (3変数) が対象物の位置/姿勢と一致したとき、適合度関数の値が最大値を示す。そのため、対象物の認識問題は、適合度関数の最大値探索問題に変換できる。適合度関数の最大値を探索・発見する場合、様々な方法があるが、一番単純かつ簡単な方法は全探索法である。これは、全てのとりうる値を計算することで最大値を発見するものであり、必ず最大値を発見できるが効率が悪い、つまり計算時間を多く費やすといった欠点を持っている。目標としているビジュアルサーボは動画像認識であり、動画像認識は計算処理が短いことが重要であるため、本研究では最適解探索に GA を適用することで、効率良く短時間で最大値探索処理を行う。そして、GA が収束した時、GA の遺伝子により決定される立体モデルの位置/姿勢  $\phi$  を対象物の位置/姿勢と考え

ることで、対象物の位置/姿勢が計測できる。

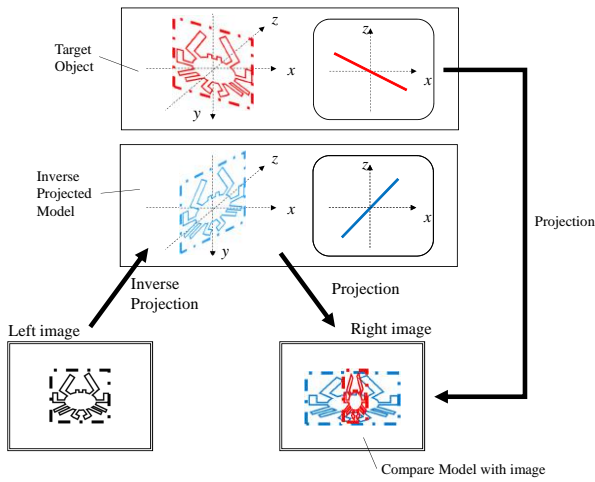


Fig.1 Projection-based Method process

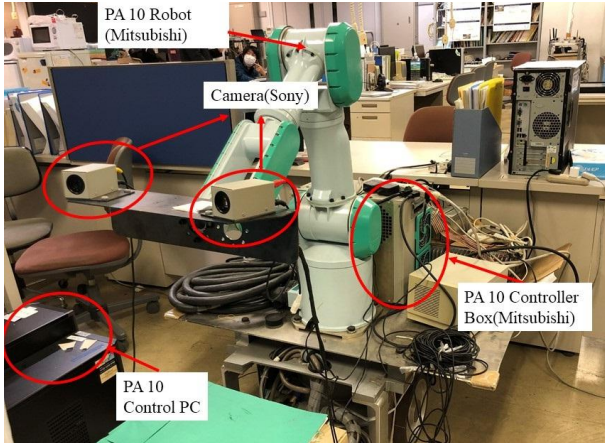


Fig.2 Hand-eye manipulator used in the system

### 3 位置姿勢認識

#### 3.1 射影変換行列

Fig.3(a) に示す中心射影の関係から、カメラの焦点距離を  $f$ 、中心座標を  $({}^I x_0, {}^I y_0)$ 、カメラ座標  $\Sigma_C$  における  $x$  軸、 $y$  軸の距離 [mm] と画像座標系  $\Sigma_I$  における  $x$  軸、 $y$  軸の距離 [pixel] の比を  $\eta_x, \eta_y$  [mm/pixel] とする。また、 $\Sigma_C$  における任意の座標  $({}^C x_i, {}^C y_i, {}^C z_i)$  は、 $\Sigma_I$  で像  $({}^I x_i, {}^I y_i)$  となる。Fig.3(b) より  $\triangle oab$  と  $\triangle oa'b'$  が相似であることより、

$$ab : a'b' = bo : b'o \quad (1)$$

と表され、これは

$$X : Y = A : B \quad (2)$$

$$\frac{X}{Y} = \frac{A}{B} \quad (3)$$

と書ける。よって、任意の点と像の座標関係は Fig.3(a) より次式で表される。

$$\frac{({}^I y_i - {}^I y_0)\eta_y}{c_{z_i}} = \frac{a}{c_{z_i}} \quad (4)$$

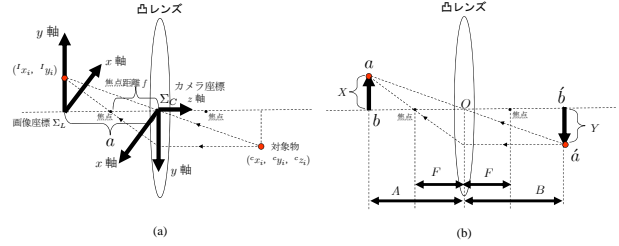


Fig.3 Perspective projection schematic

$$\frac{({}^I x_i - {}^I x_0)\eta_x}{c_{z_i}} = \frac{a}{c_{z_i}} \quad (5)$$

また、対象物からのカメラの距離  $a$  は、焦点距離  $f$  に比べ大きいので、 $a \simeq f$  と近似できると考えられる。また、画像の中心座標  $({}^I x_0, {}^I y_0)$  を画像座標系の原点  $(0, 0)$  とする。以上のことと式 (4), (5) より式 (6) が導出される。

$$\begin{bmatrix} {}^I x_i \\ {}^I y_i \end{bmatrix} = \frac{1}{c_{z_i}} \begin{bmatrix} f/\eta_x & 0 & 0 & 0 \\ 0 & f/\eta_y & 0 & 0 \end{bmatrix} \begin{bmatrix} {}^C x_i \\ {}^C y_i \\ {}^C z_i \\ 1 \end{bmatrix} \quad (6)$$

これより、カメラへの射影変換行列を  $P$  と表し、次のようにまとめる。

$$P = \frac{1}{c_{z_i}} \begin{bmatrix} f/\eta_x & 0 & 0 & 0 \\ 0 & f/\eta_y & 0 & 0 \end{bmatrix} \quad (7)$$

#### 3.2 左右カメラへの射影変換

まず、手先座標系  $\Sigma_H$  とモデル座標系  $\Sigma_M$  の関係は、 $\Sigma_H$  から  $\Sigma_M$  への同次変換行列を  ${}^H T_M$ 、 $\Sigma_H$  及び  $\Sigma_M$  で表した、空間中に逆射影された  $i$  番目のモデル上の任意座標を  ${}^H r_i, {}^M r_i$  とすると、 ${}^H r_i$  は、

$${}^H r_i = {}^H T_M {}^M r_i \quad (8)$$

と表される。左カメラ座標系  $\Sigma_{CL}$  と手先座標系  $\Sigma_H$  の関係は、 $\Sigma_{CL}$  から  $\Sigma_H$  への同次変換行列を  ${}^{CL} T_H$  とすると、 ${}^{CL} r_i$  は、

$${}^{CL} r_i = {}^{CL} T_H {}^H r_i \quad (9)$$

と表される。左画像座標系  $\Sigma_{IL}$  と左カメラ座標系  $\Sigma_{CL}$  の関係は、 ${}^{IR} r_i$  と同様に、射影変換行列  $P$  を使うと、 ${}^{IL} r_i$  は、

$${}^{IL} r_i = P({}^{CL} z_i) {}^{CL} r_i \quad (10)$$

と表される。

また、右カメラ座標系  $\Sigma_{CR}$  と手先座標系  $\Sigma_H$  の関係は、 $\Sigma_{CR}$  から  $\Sigma_H$  への同次変換行列を  ${}^{CR} T_H$  とすると、 ${}^{CR} r_i$  は、

$${}^{CR} r_i = {}^{CR} T_H {}^H r_i \quad (11)$$

と表される。右画像座標系  $\Sigma_{IR}$  と右カメラ座標系  $\Sigma_{CR}$  の関係は、射影行列  $P$  を使い、 ${}^{CR} r_i$  を  $\Sigma_{IR}$  に対して射影することで、 ${}^{IR} r_i$  は、

$${}^{IR} r_i = P({}^{CR} z_i) {}^{CR} r_i \quad (12)$$

と表される。よって、式 (8), (9), (10) より左画像への射影変換が、式 (8), (11), (12) より右画像への射影変換が求まる。

$${}^{IL} r_i = P({}^{CL} z_i) {}^{CL} T_H {}^H T_M {}^M r_i \quad (13)$$

$${}^{IR} r_i = P({}^{CR} z_i) {}^{CR} T_H {}^H T_M {}^M r_i \quad (14)$$

本研究では、手先にカメラを直接固定しているため、 ${}^{CL} T_H, {}^{CR} T_H$  は既知である。そこで、式 (13), (14) を次のように変形する。

$${}^{IL} r_i = P({}^{CL} z_i) {}^{CL} T_M {}^M r_i \quad (15)$$

$${}^{IR} r_i = P({}^{CR} z_i) {}^{CR} T_M {}^M r_i \quad (16)$$

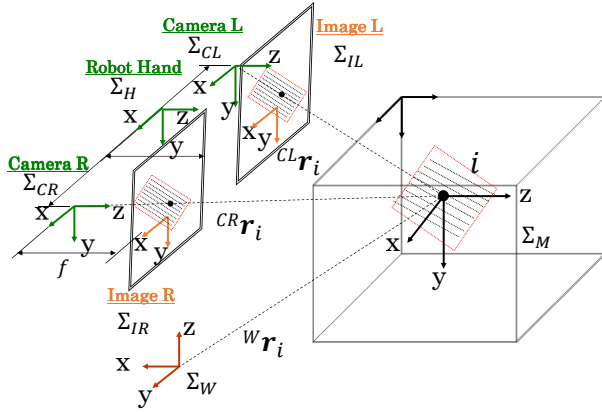


Fig.4 Coordinate system

### 3.3 左カメラからの逆射影変換

Projection-based Matching 法では、左カメラにおいて作成したモデルを逆射影変換を用いて空間中へ投影する。この逆射影変換を行うために、まず射影変換行列式 (7) に対応する逆射影変換行列  $\mathbf{P}^+$  を定義する。

$$\mathbf{P}^+(C_{z_i}) = C_{z_i} \begin{bmatrix} \frac{\eta_x}{f} & 0 & 0 & 0 \\ 0 & \frac{\eta_y}{f} & 0 & 0 \end{bmatrix}^T \quad (17)$$

またこれによる  $\Sigma_I$  から  $\Sigma_C$  への逆射影変換は以下のとおりである。

$$\begin{bmatrix} C_{x_i} & C_{y_i} & 0 & 0 \end{bmatrix}^T = \mathbf{P}^+(C_{z_i}) \begin{bmatrix} I_{x_i} & I_{y_i} \end{bmatrix}^T \quad (18)$$

$C_{z_i}$  は事前に仮定しなければならないが、 $C_{z_i}$  を 1 つ仮定すると、それに対応する  $C_{x_i}$ ,  $C_{y_i}$  がそれぞれ一意に決定される。

$$\begin{cases} C_{x_i} = f(C_{z_i}) = C_{z_i} \frac{\eta_x}{f} I_{x_i} = C_{z_i} k_x I_{x_i} \\ C_{y_i} = f(C_{z_i}) = C_{z_i} \frac{\eta_y}{f} I_{y_i} = C_{z_i} k_y I_{y_i} \end{cases} \quad (19)$$

ただし

$$k_x = \frac{\eta_x}{f}, \quad k_y = \frac{\eta_y}{f} \quad (20)$$

である。

いま、画像座標系上のモデル中心座標  ${}^{IL}\mathbf{r}_{iC}$  からカメラ座標系上のモデル中心座標  ${}^{CL}\mathbf{r}_{MiC}$  への逆射影変換は次式で表される。

$$\begin{bmatrix} {}^{CL}x_{MiC} \\ {}^{CL}y_{MiC} \\ 0 \\ 0 \end{bmatrix} = \mathbf{P}^+({}^{CL}z_{MiC}) \begin{bmatrix} {}^{IL}x_{iC} \\ {}^{IL}y_{iC} \end{bmatrix} \quad (21)$$

${}^{IL}x_{iC}$ ,  ${}^{IL}y_{iC}$  は既知であるので、式 (19), (21) より  ${}^{CL}z_{MiC}$  が求められる。また同時に  ${}^{CL}x_{MiC}$ ,  ${}^{CL}y_{MiC}$  も一意に定まる。

次に、空間中のモデル座標系にある点  ${}^{Mi}\mathbf{r}_j$  の、カメラ座標系への同次変換を考える。この同次変換は式 (22) で表される。

$$\begin{bmatrix} {}^{CL}x_{Mij} \\ {}^{CL}y_{Mij} \\ {}^{CL}z_{Mij} \\ 1 \end{bmatrix} = {}^{CL}\mathbf{T}_{Mi}({}^{CL}\mathbf{r}_{Mi}, {}^{CL}\phi_{Mi}) {}^{Mi}\mathbf{r}_j \quad (22)$$

この式を展開し式 (21) の関係を用いると、 ${}^{CL}x_{Mij}$ ,  ${}^{CL}y_{Mij}$ ,

${}^{CL}z_{Mij}$  が、式 (23) に例示するように  ${}^{CL}z_{MiC}$  および  ${}^{Mi}x_j$ ,  ${}^{Mi}y_j$  の関数として与えられる。ここで  ${}^L\theta_x$ ,  ${}^L\theta_y$  は左カメラから見た対象物の角度である。

$$\begin{aligned} {}^{CL}z_{Mij} &= S_{L\theta_x} {}^{Mi}y_j - C_{L\theta_x} S_{L\theta_y} {}^{Mi}x_j + {}^{CL}z_{MiC} \\ &= [-C_{L\theta_x} S_{L\theta_y}, S_{L\theta_x}, 0, {}^{CL}z_{MiC}] \begin{bmatrix} {}^{Mi}x_j \\ {}^{Mi}y_j \\ 0 \\ 1 \end{bmatrix} \end{aligned} \quad (23)$$

ここで、 $\Sigma_M$  に逆射影したモデルの座標  ${}^{Mi}x_j$ ,  ${}^{Mi}y_j$  を求めるために、式 (15) の逆射影式 (24) を展開する。展開すると  ${}^{Mi}x_j$ ,  ${}^{Mi}y_j$  と  ${}^{CL}z_{MiC}$  との関係式が得られるため、これよりモデルの逆射影後の座標  ${}^{Mi}\mathbf{r}_j$  が求められる。

$${}^{Mi}\mathbf{r}_j = {}^{Mi}\mathbf{T}_{CL}({}^{CL}\mathbf{r}_{Mi}, {}^{CL}\theta_{Mi}) \cdot \mathbf{P}^+({}^{CL}z_{Mij}) \cdot {}^{IL}\mathbf{r}_{ij} \quad (24)$$

以上のことから、空間中に逆射影されたモデルの  $\Sigma_{CL}$  上での座標  ${}^{CL}\mathbf{r}_{Mij} = ({}^{CL}x_{Mij}, {}^{CL}y_{Mij}, {}^{CL}z_{Mij})$  は、 ${}^{CL}z_{MiC}$  及び  ${}^L\theta_x$ ,  ${}^L\theta_y$  が与えられれば一意に定まることが分かる。 $\Sigma_{CL}$  と  $\Sigma_H$  との位置関係は既知であるため、 ${}^L\theta_x$ ,  ${}^L\theta_y$  は手先座標系から見る対象の姿勢  ${}^{Hx}\theta_M$ ,  ${}^{Hy}\theta_M$  より求められる。

### 3.4 適合度関数の定義

左右ビデオカメラから入力される画像は、0 ~ 359 で表される色相値によって構成される。探索モデルは、Fig.5 に示すように対象物の評価とその周囲の色相変化を評価するために、内部領域  $S_{R,in}$ ,  $S_{L,in}$  と帯領域  $S_{R,out}$ ,  $S_{L,out}$  によって構成される。ここで、右画像領域  ${}^{IR}\mathbf{r}_i^j(\phi_M^j)$  における色相値に基づく評価値を  $p_{R,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j))$  とする。

対象物を評価する内部領域において、モデルの色相値  $H_{IR,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j))$  とカメラ画像の色相値  $H_{MR,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j))$  の差が 20 以内であれば評価値を “+2”，20 以上離れた場合は、評価値を “-1” と設定する。同様に、対象物を評価する外部領域において、モデルの色相値  $H_{IR,out}({}^{IR}\mathbf{r}_i^j(\phi_M^j))$  とカメラ画像の色相値  $H_{MR,out}({}^{IR}\mathbf{r}_i^j(\phi_M^j))$  の差が 20 以内であれば評価値を “0.1” と設定し、そうでない場合は評価値を “-0.5” と設定する。

このような正負の評価値をもつ探索モデルと画像との相関を表す適合度関数を次式に示す。

$$F(\phi_M^j) = \left\{ \begin{array}{l} \sum_{\substack{{}^{IR}\mathbf{r}_i^j \in \\ S_{R,in}(\phi_M^j)}} p_{R,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) \\ + \sum_{\substack{{}^{IR}\mathbf{r}_i^j \in \\ S_{R,out}(\phi_M^j)}} p_{R,out}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) \end{array} \right\} / (n_{in} + n_{out}) \quad (25)$$

$$\begin{aligned} p_{R,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) &= \begin{cases} 2, & \text{if } (|H_{IR,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) - H_{MR,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j))| \leq 20); \\ -1, & \text{if } (|H_{IR,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) - H_{MR,in}({}^{IR}\mathbf{r}_i^j(\phi_M^j))| \geq 20); \end{cases} \end{aligned} \quad (26)$$

$$\begin{aligned} p_{R,out}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) &= \begin{cases} 0.1, & \text{if } (|H_{IR,out}({}^{IR}\mathbf{r}_i^j(\phi_M^j)) - H_{MR,out}({}^{IR}\mathbf{r}_i^j(\phi_M^j))| \leq 20); \\ -0.5, & \text{otherwise.} \end{cases} \end{aligned} \quad (27)$$

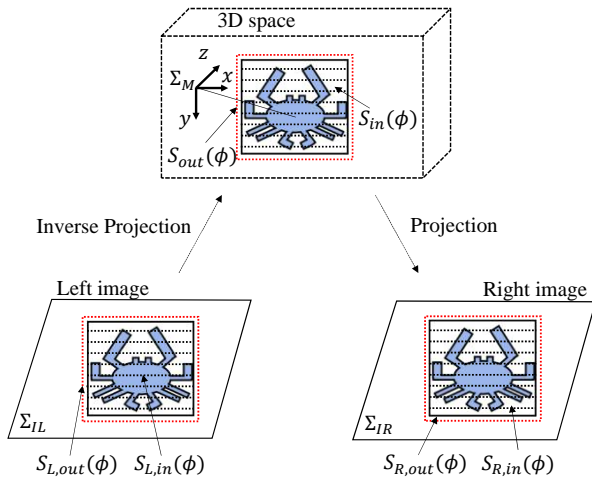


Fig.5 Searching model

この関数は、右画像領域  ${}^{IR}r_i^j(\phi_M^j)$  の中で、 $S_{R,in}(\phi_M^j)$  に重なる画像部分の適合度  $p({}^{IR}r_i)$  の合計値と、 $S_{R,out}(\phi_M^j)$  に重なる部分のそれとを合計し平均を取ることで、適合度  $F(\phi_M^j)$  を得る。これを用いることにより、右の画像と再射影された平面モデルとの一致度を評価する。この適合度関数  $F(\phi_M^j)$  は、空間中に逆射影されたモデルの持つ位置/姿勢情報より一意に決定する関数であり、逆射影されたモデルが対象物と一致すれば、右画像においても対象物と探索モデルが一致して適合度関数は最大値を示すはずである。ただし、 $F(\phi_M^j) \leq 0$  の場合、 $F(\phi_M^j) = 0$  とする。

#### 4 認識実験結果

認識実験の内容について示す。本実験では、式 (25) にて算出された適合度の分布について調査を行った。本実験において使用した対象物は蟹の模型であり、Fig.6 に示すように  $\Sigma_H$  座標系から見て  $({}^Hx_M, {}^Hy_M, {}^Hz_M) = (0, -80, 500)$  [mm]、姿勢はすべて 0[deg] として設置した。照度環境は蛍光灯下の 700lx である。 $\Sigma_H$  からの距離について 1[mm] ごとに変化させ求めた適合度分布を Fig.7 に、各角度について 5[deg] ごとに変化させ求めた姿勢の適合度分布を Fig.8 にそれぞれ示す。

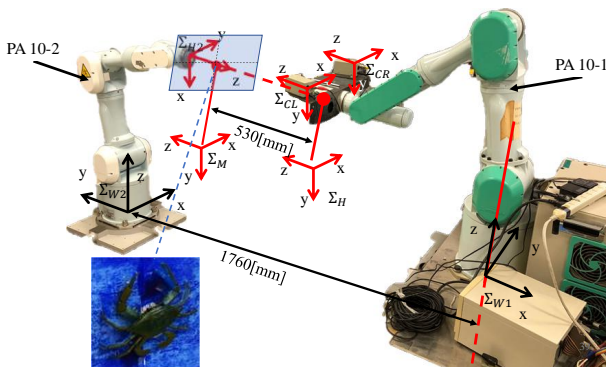


Fig.6 Experiment environment

Fig.7 では、 ${}^Hz_M = 499$  [mm] において、適合度の最大値 0.7188 を示した。また Fig.8 では、 ${}^Hx_{\theta_M} = 0$  [deg]、 ${}^Hy_{\theta_M} = 5$  [deg] において、適合度の最大値 0.6973 を示した。どちらもただ 1 つのピークが確認できたことから、本手法によって対象の位置姿勢認識が可能であるとの結果を得た。

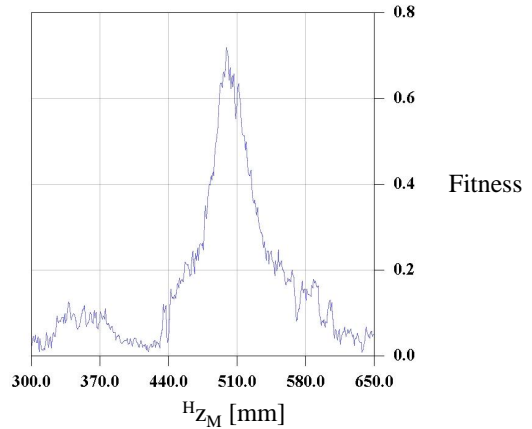


Fig.7 Fitness distribution

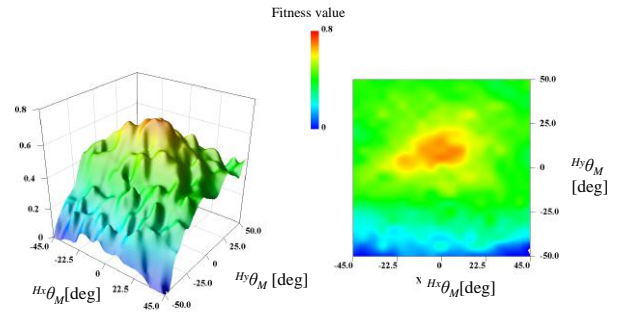


Fig.8 Fitness distribution

#### 5 結言

本報告では、複眼カメラシステムによる逆射影を用いての対象物位置姿勢認識法の理論について示した。また適合度関数の分布を求め、この手法による対象の位置姿勢認識が可能であるとの結果を得た。

#### 参考文献

- [1] Petsch, Susanne, and Darius Burschka. "Estimation of spatio-temporal object properties for manipulation tasks from observation of humans." Robotics and Automation (ICRA), 2010 IEEE International Conference on. IEEE, 2010.
- [2] Myint, Myo, et al. "Visual-servo-based autonomous docking system for underwater vehicle using dual-eyes camera 3D-pose tracking." 2015 IEEE/SICE International Symposium on System Integration (SII). IEEE, 2015.
- [3] Cui, Yu, et al. "Analyses about Trackability of Hand-eye-vergence Visual Servoing in Lateral Direction." Applied Mechanics and Materials. Vol. 772. Trans Tech Publications, 2015.
- [4] 酒井修二, 高橋徹, 伊藤康一, 青木孝文, & 運天弘樹. (2011). 汎用デジタルカメラを用いた 2 視点からの 3 次元復元. 研究報告コンピュータビジョンとイメージメディア (CVIM), 2011(4), 1-8.
- [5] Woodfill, John Iselin, Gaile Gordon, and Ron Buck. "Tyzx deepsea high speed stereo vision system." Computer Vision and Pattern Recognition Workshop, 2004. CVPRW'04. Conference on. IEEE, 2004.
- [6] Wang, Min, and Weiguang Liu. "Advanced algorithm based on SIFT and its application in binocular stereo vision." Jisuanji Gongcheng yu Yingyong(Computer Engineering and Applications) 49.2 (2013): 203-206.
- [7] 吉川恒夫, "ロボット制御基礎論", pp.222-225 (1988), コロナ社