原著論文

測域情報を用いた パートナーロボットの環境知覚†

増田 寛之*・久保田 直行*

本論文は、パートナーロボットが状況に適した知覚を実現するために、生態心理学における知覚-行為循環の概念に基づいた知覚システムを提案する。パートナーロボットのように自律動作可能なロボットは、環境を計測するために多種多様なセンサを搭載しており膨大な情報を計測することが可能である。しかし、人間のように柔軟な行動を実現するためには、膨大な情報の中から瞬時に直面する状況に適した情報を知覚しなければならない。本研究は、ロボットが人間のように柔軟な知覚-行為を実現するために、人間の知覚システムに着目し、生態心理学の分野で提唱されている知覚-行為循環の概念に基づいたロボットのための知覚システムを構築する。具体的には、網膜構造に基づく中心視モデルと知覚-行為循環に基づくスパイクニューラルネットワークを用いた手法を提案し、測域センサを用いたパートナーロボットの人間追従実験を通してロボットの知覚システムについて検証した。実機実験を通して、知覚のための構造を用意することで必要な情報を直接的に知覚することが可能であり、スパイクニューラルネットワークを用いることで予測に基づく知覚の実現とノイズに対する耐性向上を示した。

キーワード：パートナーロボット、知覚-行為循環、スパイクニューラルネットワーク、ファジィ制御

1. はじめに

近年、ロボットは生活や公共の場のような一般の環境で自律的に動作することが求められている。その背景には、現在日本の社会が直面する課題として、安心・安全な社会の実現、便利でゆとりある生活の実現がある。これらの課題解決のために、パートナーロボットに関する研究が盛んに行われ、実際に人間と触れあう環境での実証実験も行われている[1, 2]。これらのパートナーロボットは、多種多様なセンサを搭載しており膨大な環境の情報を計測することが可能である。しかし、ロボットが人間のように柔軟な行動を実現するためには、膨大な情報の中からなるべく早く直面する状況に適した情報を知覚し行為を継続しなければならない。

様々なセンサを用いて外界を知覚する研究として、コンピュータビジョン[3, 4]や能動知覚に基づくロボットビジョン[5, 6]、複数の異なるセンサを統融合して新たな情報を計測するセンサフュージョンに関する研究[7]など盛んに行われている。ロボットの知覚と行為に関する従来研究としては、刺激反応に基づく手法[8]をはじめ、環境に適応して行動を獲得する手

法[9, 10]、対象の関連性に基づいて必要な情報に注意を払う手法[11]が提案されている。特に、人間のように柔軟な行為の実現や人間とロボットのコミュニケーションに関する研究は、人間が行う知的な機能の一部をモデル化してロボットで実践している。

本研究では、人間が行う知覚に着目し、生理学的観点から網膜に基づく構造と、心理学の観点から生態心理学の分野で提唱されている知覚-行為循環の概念に基づいたロボットのための知覚システムを構築し、状況に応じて膨大な情報から必要な情報を知覚する手法を提案する。

ロボットの視知覚に関する研究では、人間検出、物体認識、マップビルディング等様々な研究が行われている。本研究では、基本能力の1つである人間追従タスクを対象とする。具体的には、自律移動可能なパートナーロボット『MOBiMac』(以下、MOBiMac)を用いて、歩行する人間を追従するための知覚システムを構築する。この人間追従機能を実現する手法として、既存研究では画像情報を用いて人間を認識する手法[12-14]を提案してきた。しかし、実際にロボットに搭載して走行する際、光環境に対するロバスト性、フレームアウト時の処理、奥行き情報の取り扱いなど問題点がある。そこで、新たに測域センサとしてレーザーレンジファインダ(Laser Range Finder: 以下、LRF)を搭載し広範囲にわたる膨大な距離情報から歩行する人間を認識する手法を提案する。

† Environmental Perception of a Partner Robot based on Range Distance Information

Hiroyuki MASUTA and Naoyuki KUBOTA

* 首都大学東京
Tokyo Metropolitan University

まず第2章では、本研究室で開発したパートナーロボットMOBiMacの概要とLRFの仕様について説明する。次に第3章では、人間の視知覚について、生理学的な観点と生態心理学的な観点から述べる。第4章では、人間の視知覚に基づいたパートナーロボットのための知覚モデルを提案する。第5章で実験結果を示し、パートナーロボットのための知覚システムについて論じる。

2. パートナーロボットMOBiMacの概要

本研究では、使い慣れたパソコンのように愛着を持ち、人間と自然なインタラクションができるロボットの実現を目指して、自律移動可能なパソコン型パートナーロボットMOBiMac(図1)を開発した。MOBiMacは、左右にDCモータを搭載し独立2輪駆動で移動可能である。センサとしてレーザーレンジファインダおよびCCD(Charge Coupled Device)センサを搭載しており環境の距離情報および画像情報を得ることができ、左右にロータリーエンコーダを搭載しているため自己位置認識も可能である。さらに、人間とのインタラクションを行うために、タッチパネル操作が可能なパソコン(VAIO U: CPU Core Solo 1.06GHz)を搭載している。同時に主要な計算はパソコン上でを行い、RS232C通信で駆動要求を内蔵マイコン(V25)に送信しモータ駆動を行う構成となっている。しかし、小型軽量を優先したパソコンを選定したため、演算負荷の大きい処理は困難である。そこで、演算負荷の小さいアルゴリズムが必要となる。また、MOBiMacの制御ソフトはロボットソフトウェアプラットフォームの考え方に基づいた設計を行っている[15, 16]。

MOBiMacに搭載したLRF(北陽電機製 URG-04LX)は、正面240度の範囲を0.35[deg]刻みで計測でき、検出可能距離は4.0[m]である。このため、最大で682点の距離情報を計測することができる。LRFで歩行中の

図2 LRFによる計測

人間を計測したときのスナップショットを図2に示す。矢印で示したところが人間の足の部分であり、この足の部分を状況に応じて効率的に認識することが課題となる。

3. 人間の視知覚

LRFの距離情報から歩行する人間を認識するという知覚処理は、人間に当てはめれば視覚を通して歩く人間を知覚することと同様であるため、人間の視知覚に着目した。本章では、人間の視知覚について述べる。

人間の視知覚に関するモデルは、脳や神経系の構造を議論する生理学的な側面と、人の行為を観察することから内部の構造を議論する心理学的な側面の両面から様々な議論が行われている。まず、生理学的な面において、外界情報は光受容器である網膜で受容し、視床を経て大脳皮質へ伝えられる。大脳皮質では階層的に結ばれたいくつかのモジュールを経て、次第に高度な処理が行われていると考えられている[17, 18]。心理学的な面において、特に生態心理学の分野では知覚と行為の相補性についての議論が行われている[19-22]。

これら生理学的知見と心理学的知見を統融合することで、人間らしい知覚をロボットでも実現することが可能であると考えられる。以下、生理学的な視点から網膜の構造について述べ、心理学的な視点から知覚-行為循環と呼ばれる概念について述べる。

3.1 網膜の構造

人間の視知覚において、光を受容するセンサである網膜の細胞分布を図3に示す[18]。破線で示す部分は杆体細胞の分布であり、この細胞は網膜の周辺に分布し、色や形の弁別能力は低い但し速い動きを知覚する能力に優れており、周辺視と呼ばれる。実線で示す部分は錐体細胞の分布であり、この細胞は網膜中心に極端

図1 パートナーロボットMOBiMac

図3 網膜上の細胞分布

に多く分布し、色や形の弁別能力が高いが約5 [Hz]を超える速い動きの知覚は困難であり、中心視と呼ばれる、このように異なる機能を持つ細胞が偏って網膜状に配置されている構造が、中心視や周辺視の能力を効率よく発揮するために重要なのではないかと考える。ここで、人間が先行者を追従する時を考えてみると、追従者は前方を歩く人間を注視しており速い動きを追う必要はない。すなわち、中心視の能力を使って歩行する人間を知覚していると推察する。

以上より、本論文では錐体細胞の分布に示すような構造が注目すべき情報を顕著にしているのではないかと仮説をたて、網膜の構造に基づいたモデルを用いた検証を行う。

3.2 知覚-行為循環

生態心理学の分野で知覚-行為循環と呼ばれる概念を図4に示す。環境からの入力情報に対して、ある特定の知覚がそれぞれ作用しあって知覚システムを構成している。知覚システムからの入力を受けて、ある特定の行為が作用しあって行為システムを構成している。知覚-行為循環は、この知覚システムと行為システムが互いに相補的で循環的な過程を基本的な単位として考え、内部のダイナミクスと外部のダイナミクスにコヒーレントな関係が生じると主張している[19]。すなわち、知覚が行為を制約し、行為が知覚を制約するという動的な立場での議論である。また、生態心理学における知覚とは、対象がどのような状態にあらうと不変的な情報を直接知覚していることを主張してお

り、膨大な環境情報を取得して特定の情報を探索して同定しているわけではないことを意味している[19-22]。

この直接知覚の考え方に基づき、人間の歩行をLRFで計測したときの不変的な情報について予備実験を行った。実験概要としては、図2に示したような位置関係となるよう、8の字状に歩行する人間の後方1[m]を手動で追従させた。このときLRFで計測した682方向の距離情報に、それぞれFFT(Discrete Fourier transform)解析した結果を重ね合わせたグラフを図5に示す。図中(A)で示した1 [Hz]以下のピークは、ロボットが動くことにより変化する壁などの動きを表しており、図中(B)で示した1~2 [Hz]のピークが人間が歩行する足の動きに相当していると解釈できる。これは、人間が0.6[m/s]程度でゆっくり歩いたときの足の動きが1.5[Hz]、2.2[m/s]程度の早歩きの際に2.4[Hz]という報告とも一致している[23]。

一方、知覚と行為の動的な議論は状況の知覚とも関係する。状況とは、図6で示すように過去から現在までの環境状態や意味や価値の推移だけではなく、現在取り得る行為によって変化するであろう未来の状態や

図4 知覚-行為循環の概念図

図5 LRFデータFFT解析結果

図6 状況の知覚

意味や価値が現在にフィードバックされ、現在の環境状態や意味や価値を見いだすと考えられる[24]。したがって、時々刻々と変化する環境状態の知覚とは、現在取り得る行為の探索であると同時に、行為と関係する未来の環境状態の予測である。

このため、本研究では時間的・空間的な情報の推移を考慮し、現在取り得る行為から未来の情報を予測可能な手法としてスパイクニューラルネットワーク (Spiking Neural Network: 以下, SNN) を適用した手法を提案し検証を行う。

4. MOBiMac の知覚-行為循環に基づく制御

第3章の概念に基づいて提案するモデルについて説明する。まず、歩行する人間を認識するための全体の処理の流れを図7に示す。まず、LRFからの入力を網

図7 処理の流れ(*P-AC: 知覚-行為循環)

膜構造に基づく中心視のモデルで計測数を縮減する、次に人間の歩行を示す特徴である1~2 [Hz]帯を通過するバンドパスフィルタ (Band Pass Filter: 以下 BPF) の出力と直前の行為出力をSNNに入力し、SNNの発火パターンから人間の方向と距離を推定する。その情報を用いて、簡略化ファジィ推論に基づいた行為を生成してMOBiMacは動作をする。本手法の特徴は、知覚のための構造を与えるものであり、一連の流れの中で探索・同定するようなアルゴリズムを用いていないことである。

以下、4.1節で網膜構造に基づく中心視のモデルについて、4.2節でSNNを用いた知覚モデルについて、4.3節でMOBiMacの行動制御について説明を行う。

4.1 網膜構造に基づく中心視モデル

3.1節で示した人間の網膜構造に基づいた中心視モデルとして、図8 (a)に示すように視線方向を中心(0 [deg])として中央を細かく、周辺を粗く計測する。この計測の粗密に関して議論の余地はあるが、本研究では最初の取り組みとして、前方1.0[m]にいる標準的な人間が全て見える視角として密な範囲を20[deg]の範囲とし、一本の足につき5~10点計測できればその動きを十分に識別できるという観点から中心を1.0[step/deg]とした。このグラフに示す計測密度に基づく、計測する点は図8 (b)に示すような分布となり、計測点数はわずか36点となる。LRFの仕様では682点計測できるため、全能力の5.3%である。情報処理の観点からみても計測点が少なくなることは計算負荷の低減などメリットは大きいと考える。

図8 網膜構造に基づく中心視モデル

4.2 知覚-行為循環に基づくモデル

本研究では時間的・空間的な情報の推移を考慮し、現在取り得る行動から未来の情報を予測可能な手法としてSNNを用いる[25-27].

知覚-行為循環に基づくモデルの実現において、SNN以外にも、階層型ニューラルネットワーク (Multi Layer Neural Network: 以下、階層型NN) やリカレントニューラルネットワーク (Recurrent Neural Network: 以下、RNN) 等の選択肢もある[28-32]. その中からSNNを選定した理由は主に以下の2点である. 1点目は、過去の入力情報を考慮できる利点である. 階層型NNの場合、時系列情報に基づく入出力関係の学習はできない. 一方、自らの出力を次の入力に用いるRNNを用いた場合、どの程度まで過去の情報を入力に返すか問題となる. SNNは、特徴の1つとして内部状態を持っており、入力の積分という形で過去の情報を反映することができるため発火の閾値と減衰率の調整により過去の時系列情報の影響度合いを容易に調整可能である. また、入力値の積分により内部状態を表現しているため入力情報の平滑化処理機能の一部も担っており、入力情報の前処理を簡素化することができるメリットがある. 2点目は、サンプリング周期に対する利点である. 人間の脳は、常に一定周期で処理が行われておらず、余裕のある状況では時間をかけて精密な処理をすることも可能であり、危険な状況ではより高速な処理を必要とされる. SNNは、出力がパルスであり最初から離散時間で扱われているため、サンプリング周期が可変であっても出力されるパルス列そのものが意味を持つ. 一方、連続値を扱う階層型NNやRNNでは、サンプリング周期が変わると出力値の持つ意味も変わってくるため出力値の取り扱いも複雑になる. 以上の理由からSNNが本提案において最も適していると判断した.

基本的にSNNは、Hodgkin-Huxley modelをはじめ、integrate-and-fire model、さらに簡単化されたspike response model等が提案されている. 本論文では、実時間で制御を行うために、spike response modelを単純化したモデルを使用する. SNNの入力ニューロンの配置を図9に示す. 計測範囲を、10[deg]毎に分割した範囲と各ニューロンが対応している. 計測した点に対応するニューロンへ1~2[Hz]の通過帯域を持つBPF通過後の出力和を入力とする. また、行為情報の入力として直前の目標駆動トルク指令値の各速度域に応じた度合い(0~1.0)と目標角速度指令値の各方向に応じた度合い(0~1.0)を入力とする. なお、全てのニューロンは相互に結合されている.

次に、 i 番目のニューロンの離散時間 t における内

図9 SNNニューロンの配置

部状態 $h_i(t)$ は次のように計算する.

$$h_i(t) = \tanh(h_i^{syn}(t) + h_i^{ref}(t)) \quad (1)$$

$$h_i^{syn}(t) = \gamma^{syn} \cdot h_i(t-1) + \sum_{j=1}^N w_{j,i} \cdot p_j(t-1) + q_i(t) \quad (2)$$

ここで、 γ^{syn} : 内部状態の減衰率、 $w_{j,i}$: j 番目のニューロンから i 番目のニューロンへの結合強度、 p_j : j 番目のニューロンからの出力、 $q_i(t)$: 環境から i 番目へのニューロンへの入力である. また、不応期を表現するために $h_i^{ref}(t)$ を用い、 γ^{ref} をニューロンが発火した場合の減衰率とすると、下記のように連続発火を抑制するための抑制項として働く.

$$h_i^{ref}(t) = \begin{cases} \gamma^{ref} \cdot h_i(t-1) - R & \text{if } p_i(t-1) = 1 \\ \gamma^{ref} \cdot h_i(t-1) & \text{otherwise} \end{cases} \quad (3)$$

ここで、 R の値を適切に定めることで不応期を調整している. i 番目のニューロンの内部状態 $h_i(t)$ が事前に決められた閾値よりも大きくなった場合、以下のようなパルスとして出力する.

$$p_i(t) = \begin{cases} 1 & \text{if } h_i(t) \geq \theta_i \\ 0 & \text{otherwise} \end{cases} \quad (4)$$

ここで θ_i は発火の閾値 ($0 < \theta_i < 1.0$) であり、スパイキングニューロンの発火を正の領域のみに限定する.

また、結合強度の学習はHebb則に基づいて式(5)によって更新する.

$$w_{j,i} \leftarrow \tanh(\gamma^{wgt} \cdot w_{j,i} + \xi^{wgt} \cdot \eta_i(t) \cdot \eta_j(t-1)) \quad (5)$$

ここで、 γ^{wgt} : 重みの減衰係数、 ξ^{wgt} : ヘップ項の係数、 η : 発火頻度を表している。このように時間発展的に学習することで一連の時系列パターンを学習している。

以上の処理により、1～2 [Hz]で変化する人間の歩行に相当する入力が大きく、かつ直前に発火した他ニューロンとの結合関係が強いとき発火しやすくなる。すなわち、発火頻度の最も高いニューロンに対応する方向を人間がいる方向であると判定することができ、その方向の計測距離が人間までの距離となる。

以上の手法をまとめると、網膜に基づく中心視モデルにより情報を縮約し、BPFで人間らしい特徴量を抽出し、SNNで情報を構造化している。このように、一連の流れの中で特定の情報を探索・同定するようなアルゴリズムは一切用いていないことが本手法の特徴である。

4.3 MOBiMacの行動制御

知覚システムで認識した人間の方向と距離を入力として、MOBiMacがとるべき行動を簡略化ファジィ推論により記述する[33]。前件部への入力 x_j は、人間の方向と距離とし、図10に示す三角型メンバーシップ関数を用いて表現する。後件部の出力 y_k は、目標駆動トルク [Nm]と目標ヨーレート [deg/s]とする。本論文では、行為目標を人間の真後ろ (0 [deg]) を1.0 [m]の距離において追従するとしてルールを構成した。ここで、 $\mu_{A_{i,j}}(x_j)$: ルール i の j 番目のメンバーシップ関数における適合度、 $a_{i,j}$: メンバーシップ関数の中

図10 三角型メンバーシップ関数

央の値、 $b_{i,j}$: メンバーシップ関数の幅とすると式(6)のように表すことが出来る。

$$\mu_{A_{i,j}}(x_j) = \begin{cases} 1 - \frac{|x_j - a_{i,j}|}{b_{i,j}} & |x_j - a_{i,j}| \leq b_{i,j} \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

ここで求めた入力数 n 個とする前件部の適合度 μ_i を以下のように表す。

$$\mu_i = \prod_{j=1}^n \mu_{A_{i,j}}(x_j) \quad (7)$$

最後に、 k 番目の出力 y_k は式(8)のように表される。

$$y_k = \frac{\sum_{i=1}^r \mu_i \cdot w_{i,k}}{\sum_{i=1}^r \mu_i} \quad (8)$$

r : ルール数、 $w_{i,j}$: 各ルールの後件部出力とする。

以上の処理によって推論した目標駆動トルクと目標ヨーレートを入力として、MOBiMacモデルが左右DCモーターへの駆動力配分および目標値を実現するための補償を行い、DCモーターを駆動する。

5. 実験結果

提案する知覚システム検証のため、実機実験結果を示す。人間追従中のスナップショットを図11に示す。実験は、図11のように障害物のない広い場所で、8の字巡回で歩く人間(図中矢印に沿って歩く)を追従する。

なお、予備実験として(i)提案手法を全て搭載した場合、(ii)中心視モデルを使わず全計測点を使った場

図11 人間追従実験

合, (iii)SNNの入力に行為情報を使わず知覚情報だけの場合について, それぞれ1試行を30秒間として5試行ずつ人間追従実験を実施した. その結果, (i) (iii)は全て人間追従に成功し(ii)は一度も人間追従を成功することがなかった.

以下, 提案手法が有効に機能していることを示すため, 網膜構造に基づく中心視のモデルについて検証を行い, 次にSNNを用いた知覚モデルの検証を行う. この時, 同じ条件で比較するために(i)の実験で計測した1試行分のログデータを使って各手法をシミュレートして比較を行う. 最後に, 追従する対象となる人間以外の人間が入ってきたときの結果について述べる.

5.1 網膜構造に基づく中心視モデルの検証

図12に人間の方向を認識した結果を示す. 図12(a)は網膜構造に基づく中心視モデルを用いた結果, 図12(b)は全計測点を用いた結果である. グラフ縦軸は図2に示しているとおりMOBiMacの正面を0[deg]として右側を正としたときの相対方向, 横軸は時間である. 図中実線は知覚システムにおいて推定された人間の方向を示しており, 破線はログデータからオペレータが人間の位置を3秒ごとに記録してスプライン補完した実測値を示した.

この結果より, 網膜構造に基づく中心視モデルが人

(a)網膜構造に基づく中心視モデル

(b)全計測点使用の一様分布

図12 人間の方向の認識結果

間の方向を安定して認識していることが明らかである. この要因は, 全計測点を用い全ての範囲を等価に見ることでノイズの影響も全て加算してしまった結果, 本来欲しい情報がノイズに埋もれてしまい認識を困難にしたと推察する. 一方, 中心視モデルでは人間が存在するであろう方向に注意を払う構造となっているため, 実際に注意の向いている方向に人間がいればその範囲の入力は相対的に大きくなる. つまり, 中央ほど高解像度であり周辺ほど低解像度とする構造そのものが特定の情報を顕著にしていることが分かる.

また, 全点入力の場合中心視モデルと比較して19倍の入力情報を処理しなければならず, MOBiMacの仕様では, 基準となるサンプリング周期10[ms]以内の処理が不可能であった. このため, 不要な情報を最初から入力しない構造を作るとはロボット制御上でも有効であることが分かる.

5.2 SNNを用いた知覚モデルの検証

図13に-30~20[deg]に対応するニューロンの出力波形を示す. 図中の矢印Aで示す部分は人間が正面から右側へ移動するときであり, パルス列が右側のニューロンへ移っている. 逆に, 図中の矢印Bで示す部分では人間が左側へ移動しているときであり, パルス列が左側のニューロンに移っている. ここで, 11秒の時と21秒の瞬間的な発火パターンだけを見れば, (b)と(c)のニューロンが同時発火しているという点で同じ状況と見なすことができる. しかし, 時系列的に見れば次の瞬間には11秒の時は(a)のニューロンの発火を促し, 21秒の時は(d)のニューロンの発火を促している. これは, 過去の発火パターンからニューロン間の結合強度を学習し続けた結果である. つまり, 過去の発火パターンが現在の知覚を制約していることが

図13 -30~20[deg]のニューロン出力波形
(図9 (a)~(e)のニューロンと対応)

図14 行為入力の有無による比較

分かる。

次に、SNNへの入力として目標駆動トルクと目標角速度の行為出力を入力した場合と行為出力を入力しない場合で比較を行う。図14は、人間の方向を認識した結果のうち1～7秒の部分拡大した結果である。濃い実線がSNNへの行為入力あり、薄い実線が行為入力なしの時の結果である。ここで、図中Aで示した箇所で位相進みを確認でき、予測に基づくような知覚が行われている。この予測に基づくような知覚は、行為が入力されるニューロンとの結合関係から知覚ニューロンの発火に制約を与えている。これは、例えば右へ向く行為出力の時はより右側を見やすくなるように行為が知覚を制約していることを示している。また、図中Bで示した箇所で行為入力がない場合ノイズによる誤認識をしている。この耐ノイズ性は、現在の行為からは生じにくいと予測できる知覚情報を抑制するように知覚を制約していることを示している。

以上の結果より、自らの行為が環境の予測に基づく知覚、及び安定した知覚のために必要不可欠であることが分かった。

5.3 複数の人間が存在する環境での検証

本提案手法では、人間を追従する際どの人間を追従すべきかという制御は組み込まれていない。当然、追従中に他の人間が近くを歩いたとき誤認識をする可能性がある。そこで、追従中に他の人間が近くを歩く場合について実験を行った。図15に追従する対象となる人間以外の人間が入ってきたときのスナップショットと、その時のLRFの全点計測結果と、網膜に基づく中心視モデルを用いた計測結果をあわせて示す。図中矢印で示している箇所が追従対象となる人間および、環境内に介在する別の人間、星印はMOBiMacが認識した追従対象となる人間の方向である。このように、特

図15 複数の人が存在する環境での実験結果

定の人間追従中に他の人間が計測範囲に入っても誤認識することなく追従し続けることができた。

この理由としては、2点挙げることができる。1点目は、網膜に基づく中心視モデルによって追従対象以外の領域は入力が粗いため、図15右下の人間認識結果のように他の人間の形状が不明瞭となっている。つまり、当該ニューロンへの入力が相対的に小さくなるため追従対象以外の人間を誤認識しにくくなる。2点目は、自らの行為と過去の発火パターンにおける他ニューロンとの結合関係から知覚が制約されることで、他の人間を誤認識しにくいと推察できる。

ただし、追従すべき人間を意図的に特定しているわけではないため、例えば2人の人間が並んで歩いているところから二手に分かれた場合どちらを追従するかは成り行きとなる。このような意図に基づいた知覚-行為の制御は今後の課題である。

5.4 SNNの考察

SNNを適用する意味について考察する。人間の方向を推定する際、発火頻度すなわち一定時間内の発火回数に注目している。このため、サンプリング周期が変わっても連続値を扱う階層型SNNと異なりそのまま適用可能であろうと推考する。

そこで、SNNの性能を検証するために、パラメータ値はそのままにサンプリング周期を変更して補足実験を行った。本システムでは、LRFのハード的制約からLRFモジュールの読み込みサンプリングは150[ms]

とし、それ以外のモジュールは人間の筋肉系の反応速度が約10[ms]という観点からサンプリングは10[ms]で処理を行っている。そこで、LRFモジュールを除く知覚システムのサンプリング周期を2[ms]、5[ms]、10[ms]、20[ms]、30[ms]としてシミュレーションを行った結果を図16に示す。破線で示した実測値との相関係数をとってみたところ、20[ms]までサンプリングを遅くしても高い相関を持っていた。実際に人間追従を行ったところ、タスク成功率は3回に1回程度となりMOBiMacがゆっくり動く状況ではタスクを達成することができた。これは、サンプリング周期が長くなると入力情報の一部が欠落するためであると推察することができ、MOBiMacがゆっくりと動作する状況でタスクの成功につながったことを説明できる。

一方、サンプリングが速い場合、誤認識をした後すぐに正しい人間の位置を認識できていない。これは、外部からの入力よりもSNNの内部状態の入力(式(2)の第2項)の影響が大きくなることが原因であると推察する。実際に人間追従を行うと見失ったときに誤認方向に動作し続けるため人間追従タスクを達成できなかった。

以上の実験結果よりSNNが示した挙動として、サンプリングを速くすると誤った情報に注意を払いやすくなる。これは、人間が非常に限られた時間内で行う

意思決定と類似している。一方、サンプリングを遅くすると、サンプリング定理[34]により必要な情報が欠落する。これは、人間が不十分な情報の中での意思決定と通ずるところがあり興味深い結果となった。今後、心理学の議論と関連づけて詳細な議論を行う必要がある。

6. おわりに

本論文では、測域センサを搭載したパートナーロボットの間人追従実験を通してパートナーロボットにおける知覚システムの構築と検証を行った。その結果、人間の視知覚をベースとした知覚モデルを構築することで特定の情報を探索・同定するようなアルゴリズムは一切用いずに、歩く人間を認識することができるとを示した。また、錐体細胞の分布に示すような構造が注目すべき情報を顕著にしているのではないかとという仮説に対して、網膜に基づく中心視の細胞分布を考慮したモデルが、不要な情報を削減し、特定の情報を顕著にしていることを示した。すなわち、仮説に基づくモデルが有効であることを示すことができた。さらに、SNNを適用することで時空的な文脈を考慮した出力を得ることができ、ロボットの行為が知覚を制約することで予測に基づく知覚の実現とノイズに対する耐性向上につなげることができることを示した。

一方、本手法は特定の知覚と特定の行為を直接的に結びつけているだけである。このため、意図的に特定の人物を同定して追従し続けることができず、例えば並行して歩く複数の人間が二手に分かれた場合どちらを追従するか意図的に選択することができない。今後の課題として、より上位の意図に基づき制約される知覚-行為循環の実現として、意図的な探索方向の制限、サンプリング周期を動的に可変する数理モデルの構築を目指す。

参考文献

- [1] パートナーロボット資料集成, エヌ・ティー・エス, 2005.
- [2] 鴨志田英樹, ロボット業界最前線の28人が語る!ロボットの現在と未来, エクスメディア, 2005.
- [3] 出口光一郎, コンピュータビジョン-ロボットの「目」をつくる, 丸善, 1989.
- [4] C.H.Chen, L.F.Pau, and P.SP.Wang, Handbook of Pattern Recognition & Computer Vision, World Scientific Publication Co. Pte. Ltd., 1993.
- [5] 出口光一郎, ロボットビジョンの基礎, コロナ社, 2000.
- [6] M.E.Roberts, E.Claridge, "An Artificially Evolved Vision System for Segmenting Skin Lesion Images", Proc. of 6th International Conference of Medical Image Computing and Computer-Assisted Intervention

図16 サンプリング周期毎の人間方向推定結果

- (MICCAI 2003) (1), pp. 655-662, 2003.
- [7] 山崎, 石川, センサフュージョン, コロナ社, 1992.
- [8] R.Brooks, A Robust Layered Control System for a Mobile Robot, IEEE Journal of Robotics and Automation, Vol.2, No.1, pp.14-23, 1986.
- [9] 港, 浅田, 環境の変化に適応する移動ロボットの行動獲得, 日本ロボット学会誌, Vol.18, No.5, pp.706-712, 2000.
- [10] 久保田, 増田, 小島, 福田, 知覚-行為循環に基づく移動ロボットの知覚システムと行為システム, 日本ロボット学会誌, Vol.24, No.4, pp. 473-481, 2006.
- [11] E.Sato, T.Yamaguchi, and F.Harashima, Natural Interface Using Pointing Behavior for Human-Robot Gestural Interaction, IEEE transactions on Industrial Electronics, Vol.54, No.2, 2007.
- [12] N.Kubota, K.Nishida, Cooperative Perceptual Systems for Partner Robots Based on Sensor Network, International Journal of Computer Science and Network Security (IJCSNS), Vol.6, No.11, pp.19-28, 2006.
- [13] N.Kubota, K.Nishida, H.Kojima, Perceptual System of a Partner Robot for Natural Communication Restricted by Environments, Proc. of 2006 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2006), IROS06-170, 2006.
- [14] N.Kubota, K.Nishida, Perceptual Control Based on Prediction for Natural Communication of A Partner Robot, IEEE Transactions on Industrial Electronics, Vol.54, No.2, pp.866-877, 2007.
- [15] 谷江, 松日榮, 次世代ロボット共通プラットフォーム技術-情報構造化環境プラットフォームの構築-, 日本ロボット学会誌, vol.25, No.4, pp.501-504, 2007.
- [16] N.Ando, T.Suehiro, K.Kitagaki, T.Kotoku, W.K.Yoon, RT-Middleware: Distributed Component Middleware for RT (Robot Technology), 2005 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS 2005), pp.3555-3560, 2005.
- [17] 松本元, 大津展之, 脳・神経系が行う情報処理とそのモデル, 培風館, 1994.
- [18] 日本視覚学会編, 視覚情報処理ハンドブック, 朝倉書店, 2000
- [19] J.J.Gibson, The senses considered as perceptual systems, Boston: Houghton Mifflin, 1966.
- [20] 三嶋博之, エコロジカル・マインド: 知性と環境をつなぐ心理学, 日本放送出版協会, 2000.
- [21] 佐々木正人, 三島博之, アフォーダンスの構想 知覚研究の生態心理学的デザイン, 東京大学出版会, 2001.
- [22] D.N.Lee, Guiding movement by coupling taus, Ecological psychology, Vol.10, pp.221-250, 1998.
- [23] H.G.MacDougall, S.T.Moore, Marching to the beat of the same drummer: the spontaneous tempo of human locomotion, Journal of Applied Physiology, 2005.
- [24] 久保田直行, 知能ロボットの行動学習, システム制御情報学会, システム/制御/情報, Vol.48, No.2, pp. 51-56, 2004.
- [25] 黒江, スパイクングニューラルネットワーク, システム/制御/情報, Vol.48, No.2, pp.57-62, 2004.
- [26] W.Maass, and C.M.Bishop, Pulsed Neural Networks, The MIT Press, 1999.
- [27] 渡辺, 田中, 栗田, 三島, シグモイド型パルスジェネレータを取り入れた Pulse-Coupled Neural Network による注視点探索法, 計測自動制御学会論文集, Vol.38, No.8, 726-732, 2002.
- [28] 武藤, 斎藤, ニューラルコンピューティング, 共立出版株式会社, 2001.
- [29] 馬場, 小島, ニューラルネットの基礎と応用, 共立出版株式会社, 1994.
- [30] 福田敏男, インテリジェントシステム, 昭晃堂, 2000.
- [31] E.Mizutani, S.E.Dreyfus, On complexity analysis of supervised MLP-learning for algorithmic comparisons, Proceedings of INNS-IEEE International Joint Conference on Neural Networks, 2001.
- [32] 川人光男, 脳の計算理論, 産業図書, 1996.
- [33] 田中一男, アドバンストファジィ制御, 共立出版, 1994.
- [34] 計測自動制御学会編, 小畑, 浜田, 田村, 信号処理入門, コロナ社, 2007.

(2007年10月4日 受付)

(2008年7月14日 採録)

[問い合わせ先]

〒191-0065 東京都日野市旭が丘6-6
首都大学東京 システムデザイン研究科
増田 寛之
TEL: 042-585-8441
FAX: 042-585-8441
E-mail: masuta-hiroyuki@sd.tmu.ac.jp

著者紹介

ますた ひろゆき
増田 寛之 [学生会員]

2001年福井大学工学部機械工学科卒業。2003年福井大学大学院工学研究科機械工学専攻・博士前期課程修了。同年株式会社デンソー入社。2007年首都大学東京システムデザイン研究科ヒューマンメカトロニクスシステム専修・博士後期課程入学。同年、首都大学東京リサーチアシスタント。現在に至る。

くぼた なおゆき
久保田 直行 [正会員]

1992年大阪教育大学教育学部教養学科卒業。1994年北海道大学大学院工学研究科情報工学専攻修士課程修了。1997年名古屋大学大学院工学研究科マイクロシステム工学専攻博士後期課程修了。1997年大阪工業大学工学部機械工学科助手。1999年同講師。2000年福井大学工学部知能システム工学科助教授。2004年東京都立大学大学院工学研究科機械工学専攻助教授となり。2005年首都大学東京システムデザイン学部ヒューマンメカトロニクスシステムコース准教授。現在に至る。2002年から科学技術振興機構(JST) 戦略的創造研究推進事業「相互作用と賢さ」領域研究員。2006年から同事業発展研究研究員兼任。英国王立協会客員教授(ポーツマス大学。2007年)。博士(工学)。

Environmental Perception of a Partner Robot based on Range Distance Information by

Hiroyuki MASUTA and Naoyuki KUBOTA

Abstract :

This paper proposes a perceptual system based on the concept of perceiving-acting cycle to realize the situation-dependent perception for a partner robot. Recently, various sensors enable an autonomous robot to get much information on the environment easily. However, the robot must extract the information suitable for the situation from such much information to take a flexible action like a human. In this study, we emphasize the importance of human vision in order that the robot flexibility realizes perception and action. Especially we focus on the perceptual system based on perceiving-acting cycle discussed in ecological psychology. First, we propose a retinal model based on human retinal structure, and a method for the information extraction using a spiking neural network based on perceiving-acting cycle. Next, we apply the proposed method of a partner robot with a laser range finder for a human tracking task. As experimental results, we show that the proposed method can directly perceive the necessary information based on the retinal model. Furthermore we show that the spiking neural network can realize the prediction-based perception and can improve the noise tolerance in the perception.

Keywords : Partner Robots, Perceiving-Acting Cycle, Spiking Neural Networks, Fuzzy Control

Contact Address : **Hiroyuki MASUTA**

Tokyo Metropolitan University Dept. of System Design
6-6, Asahigaoka, Hino, Tokyo 191-0065, JAPAN
TEL : 042-585-8441
FAX : 042-585-8441
E-mail : masuta-hiroyuki@sd.tmu.ac.jp