

12th SICE Symposium on Computational Intelligence

December 15-16, 2017, Narashino

第12回コンピューターショナル・インテリジェンス研究会
講演論文集

期 日：2017年12月15日(金)，16日(土)

会 場：千葉工業大学 津田沼キャンパス

SICE[®]

主 催：計測自動制御学会 システム・情報部門

企 画：コンピューターショナル・インテリジェンス部会，知能工学部会

協 賛：情報処理学会，システム制御情報学会，電子情報通信学会，電気学会，日本
神経回路学会，日本知能情報ファジィ学会，日本機械学会，人工知能学会，
ヒューマンインタフェース学会，進化計算学会，IEEE Computational
Intelligence Society Japan Chapter (CISJ)，IEEE Systems, Man, and Cybernetics
Society Japan Chapter

著作権 © 2017

公益社団法人計測自動制御学会(SICE)
〒101-0052 東京都千代田区神田小川町 1-11-9 金子ビル 4 階

カタログ番号 17 PG 0011

著作権は、計測自動制御学会がもっている
ので、個人の使用のための複製以外の
目的で掲載の記事の一部または全文を
複製する場合には、著作権者に許可を求
め規定の複製料を支払うこと。

発行日：2017年12月15日

発行者：公益社団法人計測自動制御学会 システム・情報部門
コンピューターショナル・インテリジェンス部会， 知能工学部会

第12回コンピュータショナル・インテリジェンス研究会プログラム

発表 25 分（発表 20 分、質疑応答 5 分）

12月15日（金）

13:05~14:45 (100 分 4 件) 司会 畠中利治 (大阪大学)

- 1 スパイキングニューラルネットワークにおける自発的発火活動の同期解析 1
信川創, 相浦弘 (千葉工業大学), 西村治彦 (兵庫県立大学), 山西輝也 (福井工業大学)
- 2 四元数ニューラルネットを用いた 9 軸モーションセンサによる前腕姿勢の推定 5
ハフィズディン フィルダウス, 小川毅彦 (拓殖大学)
- 3 並列 RBF ネットワークに基づく紙の汚れ計測 9
名取直毅, 長谷部光威, 古賀敏之, 齊藤廣大 (株式会社東芝)
- 4 不精密ルールを考慮したラフ集合による知識獲得及びファジィ判別分析 13
関宏理 (大阪大学), 山下亮 (伊藤忠商事), 乾口雅弘 (大阪大学)

15:00~16:15 (75 分 3 件) 司会 関宏理 (大阪大学)

- 5 重み付き投票法による離散化とArtificial Bee Colonyを用いた設計最適化法 19
横谷元, 畠中利治 (大阪大学)
- 6 罰を受ける状況に関する情報の抽象化と強化学習効率化への利用 23
坂下悠太, 村田純一 (九州大学)
- 7 報酬分配を用いた Deep Q-Network の実現 28
中矢裕太, 長名優子 (東京工科大学)

16:30~17:30 Invitation to Computational Intelligence Fons Scientiae Project

12月16日（土）

9:00~10:15 (75 分 3 件) 司会 近藤伸彦 (首都大学東京)

- 8 構造適応型 Deep Belief Network 学習法による検診データの学習速度の向上 36
鎌田真 (広島市立大学), 市村匠 (県立広島大学)
- 9 深層生成モデルとエンコーダを用いた自動修正法および住宅設計への適用 42
植田考哉, 瀬尾昌孝, 西川郁子 (立命館大学)
- 10 複雑な環境下における可変型測域センサアレイの知的制御 48
北井瑳佳, 戸田雄一郎, 武居直行, 和田一義, 久保田直行 (首都大学東京)

10:30~12:00 特別講演 (企画: SICE 部門連携・活性化委員会 応用クリフォード代数 WG)

講演 1 尚方先生 (電気通信大学) 「四元数ニューラルネットワークに基づく全偏波合成開口レーダのデータ解析」

講演 2 本間経康先生 (東北大学) 「A stable gradient-descent adaptation of higher order neural units」

13:15~14:30 (75分3件) 司会 市村匠 (県立広島大学)

- 11 MAを用いた絵画的迷路の自動生成 52
今宮明則, 小野典彦, 永田裕一 (徳島大学)
- 12 車両配送問題における追加注文を考慮した配送計画システムの提案 60
織田雄大, 小野典彦, 永田裕一 (徳島大学)
- 13 遺伝子ネットワークにおける周期発現パターン遷移列のアイソクロンの計算法 68
森禎弘, 黒江康明 (京都工芸繊維大学)

14:45~16:00 (75分3件) 司会 信川創 (千葉工業大学)

- 14 量子ダイナミクスによる群ロボットの知的タスク処理 74
伊丹哲郎 (ロボット産業振興会議), 松井伸之 (兵庫県立大学)
- 15 肘付き制御によるマニピュレータの作業精度とエネルギー効率の改善 80
李想, 見浪護 (岡山大学)
- 16 PHDフィルタを用いたバス乗客の行動把握 88
南哲志 (工学院大学), ファン ミン トウン (ダナン工科大学), 橘完太 (工学院大学)

スパイクングニューラルネットワークにおける 自発的発火活動の同期解析

○信川創 相浦弘 (千葉工業大学) 西村治彦 (兵庫県立大学) 山西輝也 (福井工業大学)

Analysis of Synchronization in Spontaneous Activity by Simulation of Spiking Neural Networks

*S. Nobukawa, H. Aiura (Chiba Institute of Technology), H. Nishimura (University of Hyogo),
and T. Yamanishi (Fukui University of Technology)

Abstract— Cortical neural networks sustain some degree of electrical activity representing the brain's noisy internal state even in the absence of sensory stimulation, which is called spontaneous activity or ongoing activity. This neural activity exhibits typically irregular neural spiking with a low firing rate (≈ 1 Hz), but high coherent spike transmission between specific neurons. Teramae *et al.* proposed that the spontaneous activity is produced by log-normal distribution of excitatory synaptic weights. Recently, it has been also reported that network, which is constructed by the synapses with large amplitude of excitatory synaptic weights, exhibits high clustering coefficient and low short path length, i.e., the characteristic of small world network. However, there are few studies regarding relationship between the characteristic of small world network and complexity of spontaneous activity. Therefore, in this study, we construct the spiking neural network composed of the small world network with large amplitude of excitatory synaptic weights and random network with the other synapses. We evaluate the fluctuation of neural activity in the spiking neural network by multi-fractal analysis.

Key Words: spiking neural network, spontaneous activity, synchronization, small-world network, multi-fractal analysis

1 はじめに

確率共鳴理論において、脳・神経系における適度なゆらぎは信号伝達や情報処理機構の機能性を促進するとされている¹⁾。現在、この確率共鳴理論に基づいたゆらぎと脳機能との関連が活発に研究されている^{2, 3, 4, 5, 6, 7)}。特に、脳波や脳磁図、機能的核磁気共鳴などの脳機能画像法による研究から、脳活動のダイナミクスは大きな変動性を持ち、それはエントロピーやカオス/フラクタル解析によって特徴付けられ、そのゆらぎの程度は脳機能と相関することが明らかとなっている^{6, 7)}。

そして近年、この脳におけるゆらぎ源の一つである自発的発火活動に注目が集まっている^{8, 9, 10)}。大脳皮質におけるニューロン集団では、外界からの刺激を遮断された環境においても、自発的発火活動を起こす。この発火活動の特徴として、発火率が数 [Hz] 程度という低い頻度で、ポアソン過程程度の高い不規則性を持ち^{11, 12)}、かつ大多数のニューロン間において同期性が見られない¹³⁾ というノイズ的な性質が挙げられる。一方、自発的発火活動時の興奮性細胞集団の活動度と抑制性細胞集団の活動度には強い相関が見られること¹⁴⁾、また持続発火中の細胞間スパイク列から、数ミリ程度の高い時間制度を持つスパイク時系列構造が発見される¹⁵⁾ など、ノイズ的性質とは相反する精密な特性も報告されている。

大脳皮質における化学シナプスにおいては、シナプス前ニューロン (pre ニューロン) の発火によって、シナプス後ニューロン (post ニューロン) の膜電位を増加させる興奮性シナプス後電位 (Excitatory Post Synaptic Potential (EPSP)) の大きさは典型的には 1 mV 以下であり、神経細胞の静止膜電位と発火しきい値の差 (約 20 ミリボルト) に比べて極めて小さいと考えられてきた。しかし近年、大多数の EPSP がサブミリボルト程度の

ものであるが、まれに典型的な EPSP サイズの数十倍 (数ミリボルト以上) の強度をもつ巨大な EPSP が観測され、その強度分布が具体的にはロングテールな対数正規分布でよく近似できることが明らかにされている^{16, 17)}。Teramae らは、発火の生成の詳細な生物学的メカニズムは考慮せずに、入力的时间積分によって細胞間の膜電位が徐々に上昇し、ある閾値を超えると発火が生成される現象に着目して作られた最も単純なスパイクングニューロンモデルである leaky integrate-and-fire (LIF) モデルを用いて、ロングテールな分布に従うシナプス荷重により大脳皮質の発火活動の特徴が担われている可能性を示した¹⁸⁾。そして、このシナプス荷重の対数正規性によって、連想記憶における記憶の想起精度が高められることや¹⁹⁾、海馬における記憶の固定化 (memory consolidation) に重要な役割を担うバースト発火の伝搬を生起させることなどが報告されている²⁰⁾。更に、大多数の小さな EPSP を持つシナプス結合ではランダムネットワーク性を示すが、大きな EPSP を持つシナプス結合は短い平均経路長と高いクラスタ性を持ったスモールワールドネットワークを形成することが報告されている²¹⁾。しかし、このネットワークの二重構造の機能性や自発的発火活動への影響についての報告は少ない²²⁾。

このような中で本研究では、ネットワークの二重構造と自発的発火活動の複雑性との関連を明らかにするため、Teramae らが示した対数正規性を有するシナプス荷重を持ったスパイクングニューラルネットワーク¹⁸⁾ に対して、強い EPSP を持つシナプスによって形成されるネットワークに Watts-Strogatz モデルによるスモールワールド性を導入する。そしてマルチフラクタル解析の手法により、ニューロン集団の発火活動の時系列の複雑性を評価する。

2 モデルと評価指標

2.1 スパイクニューラルネットワーク

各ニューロンの膜電位 v のダイナミクスは LIF モデルによる (1) 式で与えられる.

$$\frac{dv}{dt} = -\frac{1}{\tau_m}(v - V_L) - g_E(v - V_E) - g_I(v - V_L) + I_{\text{ex}} \quad (1)$$

ここで、膜電位の時定数 τ_m は、20[ms](興奮性ニューロン) と 10[ms](抑制性ニューロン) に設定する. 漏れ電流, 興奮性シナプス後電流, 抑制性シナプス後電流は $V_L = -70$ [mV], $V_E = 0$ [mV], $V_I = -80$ [mV] に設定する. 興奮性/抑制性シナプスコンダクタンスである g_E/g_I [ms⁻¹] は (2) 式で表される. I_{ex} はポアソン過程によって生成される外部入力であり, $0 \leq t \leq 100$ [ms] において発火頻度 0.3 [Hz] で入力を与え, $t > 100$ [ms] では 0 [Hz] とする.

$$\frac{dg_X}{dt} = -\frac{g_X}{\tau_s} + \sum_j G_{X,j} \sum_{s_j} \delta(t - s_j - d_j), \quad X = E, I \quad (2)$$

ここで、 $\delta(t)$ はデルタ関数であり, G_j, d_j, s_j はそれぞれシナプス荷重, 遅れ時間, ニューロン j からの発火時刻を表す. 減衰定数 τ_s は 2[ms] に, d_j は $[d_0 - 1 : d_0 + 1]$ [ms] の一様乱数とし, 興奮性-興奮性のシナプス結合においては $d_0 = 2$, 他のシナプス結合においては, $d_0 = 1$ とする. 発火の閾値は $V_{\text{thr}} = -50$ [mV] で発火後は v を $V_r = -60$ [mV] にリセットする. 尚, 不応期は 1[ms] とする.

興奮性-興奮性の EPSP は, 対数正規による確率分布

$$p(x) = \frac{\exp[-(\log x - \mu)^2 / 2\sigma^2]}{\sqrt{2\pi\sigma x}} \quad (3)$$

で与えられる. 本研究では, $\sigma = 1.0, \mu - \sigma^2 = \log(0.2)$ に設定する. 尚, x が 15[mV] を超える場合は再度乱数を発生させる. 興奮性シナプスの結合確率は 0.1 とし, EPSP が 10[mV] 以上の興奮性ニューロン間の結合では, リンクの繋ぎかえ確率を β とした Watts-Strogatz モデルでネットワークを形成する.

また, $G_i = x/100$ により EPSP から G_i への変換を行う. 抑制性-興奮性及び抑制性-抑制性のシナプス荷重は $G_i = 0.018, 0.002, 0.0025$ で与え, 抑制性シナプスの結合確率は 0.5 とする. 興奮性-興奮性のシナプス伝達は EPSP の強度に依存して, $P_E = a/(a + EPSP)$ ($a = 0.1$ mV) で失われるとする.

2.2 Wavelet leaders によるマルチフラクタル解析

離散信号 $X(t)$ の離散 Wavelet 変換の係数は, (4) 式で与えられる.

$$d_X(j, k) = \int_R X(t) 2^{-j} \psi_0(2^{-j}t - k) dt \quad (4)$$

($j = 1, 2, \dots, k = 1, 2, \dots$)

ここで, ψ_0 はコンパクトサポートな mother wavelet 関数とする.

1次元の wavelet leaders は (5) 式で表される.

$$L_X(j, k) = \sup_{\lambda' \subset 3\lambda_{j,k}} |d_X(j, k)| \quad (5)$$

ここで, $\lambda = \lambda_{j,k} = [k2^j, (k+1)2^j)$ はスケール 2^j による時間間隔を, $3\lambda_{j,k} = \lambda_{j,k-1} \cup \lambda_{j,k} \cup \lambda_{j,k+1}$ は隣接時間を表す.

特異値スペクトルは, この wavelet leaders によって (6) 式で与えられる²³⁾.

$$D(k) = \inf_{q \neq 0} (1 + qh - \zeta_L(q)) \quad (6)$$

ここで, h と q はそれぞれ Hölder 指数と wavelet leaders のモーメントを表す. スケーリング指数 $\zeta_L(q)$ と構造関数 $S_L(q, j)$ は (7) 式と (8) 式で表される.

$$\zeta_L(q) = \liminf_{j \rightarrow 0} \left(\frac{\log_2 S_L(q, j)}{j} \right) \quad (7)$$

$$S_L(q, j) = \frac{1}{n_j} \sum_{k=1}^{n_j} |L_X(j, k)|^q \quad (8)$$

ここで, n_j はスケール 2^j の場合の X のサンプル数を表す. 本研究では, 上記のマルチフラクタル解析を $-5 \leq q \leq 5$ の区間で実施し, 解析には Wavelet Toolbox of MATLAB (<https://jp.mathworks.com/products/wavelet.html>) を用いる.

3 結果

Fig. 1 に, EPSP が 10[mV] 以上の興奮性シナプス結合で形成されるネットワークにおいて, $\beta = 0.2, 0.6, 1.0$ に設定した場合のラスタプロット ((a)) と興奮性ニューロン集団と抑制性ニューロンの発火率の時系列 ((b)), 発火率の時系列のパワースペクトル ((c)) を示す. これらの結果から全ての β の場合で不規則な低発火頻度 (興奮性ニューロン集団で 2.5 Hz 程度) の自発的発火活動が生じていることが確認できる.

次に, この興奮性ニューロン集団の発火率に対して (Fig. 1 (b)), マルチフラクタル解析を実施し, 特異値スペクトル $D(h)$ を算出した結果を Fig.2 に, $D(h)$ の 1次キュムラント (c_1) と 2次キュムラント (c_2) を Table. 1 に示す. $\beta = 0.2$ の場合では, $D(h)$ は $0.07 \lesssim h \lesssim 0.63$ の範囲に分布し, 分布のピークである大域的 Hölder 指数は $c_1 \approx 0.25$ を取る. 一方, $\beta = 0.6$ の場合では, $D(h)$ は $0.1 \lesssim h \lesssim 0.97$ の範囲に分布し, 大域的 Hölder 指数は $c_1 \approx 0.39$ となる. 更に, ネットワークのスモールワールド性が失われ, ネットワークがランダム化する $\beta = 1.0$ では, $D(h)$ は $0.07 \lesssim h \lesssim 0.45$ の範囲で分布し, 大域的 Hölder 指数は $c_1 \approx 0.24$ まで減少する. また, $D(h)$ の分布の広がり, 興奮性ニューロン集団の発火率のマルチフラクタル性を表しており, その程度は c_2 の絶対値に比例し, $\beta = 0.6$ の場合で最も高い値 $c_2 \approx -0.11$ を示す.

Table 1: Cumulants of singularity spectrum (first order cumulant: c_1 and second order cumulant: c_2).

	c_1	c_2
$\beta = 0.2$ case	0.25	-0.056
$\beta = 0.6$ case	0.39	-0.11
$\beta = 1.0$ case	0.24	-0.039

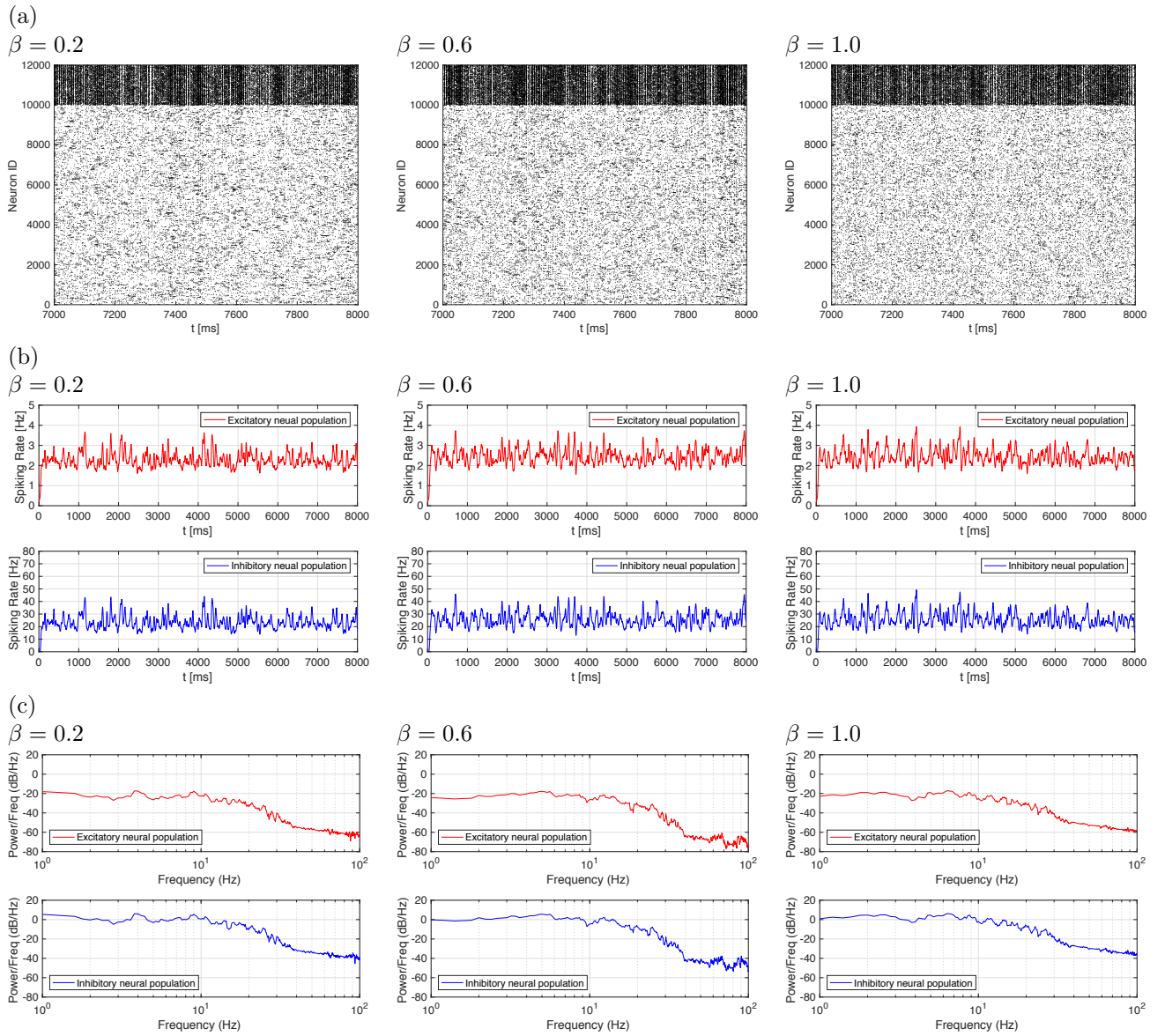


Fig. 1: Characteristic of spiking activity in spiking neural network with coexisting small-world network and random network. (a) Raster plot. (b) Time series of spiking rate. (c) Power spectrum of spiking rate.

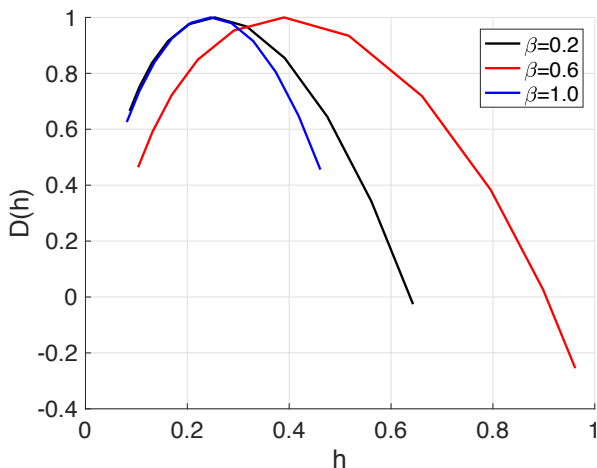


Fig. 2: Singularity spectrum $D(h)$ for time series of spiking rate in excitatory neural population as function of Hölder exponent h .

4 おわりに

本研究では、Teramae らが示した自発的発火活動を示すスパイクングニューラルネットワーク¹⁸⁾に対して強い EPSP を持ったシナプス結合で形成されるネットワークにスモールワールド性を導入した。そして、興奮性ニューロン集団の発火活動のマルチフラクタル解析を実施した。その結果、スモールワールド性を制御するリンクの繋ぎかえパラメータを変化させることで、発火活動の複雑性を表す大域的 Hölder 指数と複雑性の時間スケール依存性を反映したマルチフラクタル性が最大化することが明らかとなった。

今後の課題としては、このような時間スケール依存性を持った高い複雑性を示す自発的発火活動の機能性の評価が挙げられる。

参考文献

- 1) M. D. McDonnell and L. M. Ward, “The benefits of noise in neural systems: bridging theory and experiment,” *Nature Reviews Neuroscience*, vol. 12, no. 7, pp. 415–426, 2011.

- 2) A. R. McIntosh, N. Kovacevic, and R. J. Itier, "Increased brain signal variability accompanies lower behavioral variability in development," *PLoS Comput Biol*, vol. 4, no. 7, p. e1000106, 2008.
- 3) T. Takahashi, R. Y. Cho, T. Murata, T. Mizuno, M. Kikuchi, K. Mizukami, H. Kosaka, K. Takahashi, and Y. Wada, "Age-related variation in EEG complexity to photic stimulation: a multiscale entropy analysis," *Clinical Neurophysiology*, vol. 120, no. 3, pp. 476–483, 2009.
- 4) D. D. Garrett, N. Kovacevic, A. R. McIntosh, and C. L. Grady, "Blood oxygen level-dependent signal variability is more than just noise," *The Journal of Neuroscience*, vol. 30, no. 14, pp. 4914–4921, 2010.
- 5) —, "The importance of being variable," *The Journal of Neuroscience*, vol. 31, no. 12, pp. 4496–4503, 2011.
- 6) T. Takahashi, "Complexity of spontaneous brain activity in mental disorders," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 45, pp. 258–266, 2013.
- 7) A. C. Yang and S.-J. Tsai, "Is mental illness complex? from behavior to brain," *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, vol. 45, pp. 253–257, 2013.
- 8) J.-n. Teramae and T. Fukai, "Computational implications of lognormally distributed synaptic weights," *Proceedings of the IEEE*, vol. 102, no. 4, pp. 500–512, 2014.
- 9) H. Kada, J.-N. Teramae, and I. T. Tokuda, "Effective suppression of pathological synchronization in cortical networks by highly heterogeneous distribution of inhibitory connections," *Frontiers in computational neuroscience*, vol. 10, 2016.
- 10) N. M. Timme, S. Ito, M. Myroshnychenko, S. Nigam, M. Shimono, F.-C. Yeh, P. Hottowy, A. M. Litke, and J. M. Beggs, "High-degree neurons feed cortical computations," *PLoS computational biology*, vol. 12, no. 5, p. e1004858, 2016.
- 11) A. S. Ecker, P. Berens, G. A. Keliris, M. Bethge, N. K. Logothetis, and A. S. Tolias, "Decorrelated neuronal firing in cortical microcircuits," *science*, vol. 327, no. 5965, pp. 584–587, 2010.
- 12) A. Renart, J. De La Rocha, P. Bartho, L. Hollender, N. Parga, A. Reyes, and K. D. Harris, "The asynchronous state in cortical circuits," *science*, vol. 327, no. 5965, pp. 587–590, 2010.
- 13) W. R. Softky and C. Koch, "The highly irregular firing of cortical cells is inconsistent with temporal integration of random epsps," *Journal of Neuroscience*, vol. 13, no. 1, pp. 334–350, 1993.
- 14) Y. Shu, A. Hasenstaub, and D. A. McCormick, "Turning on and off recurrent balanced cortical activity," *Nature*, vol. 423, no. 6937, pp. 288–293, 2003.
- 15) M. Abeles and Y. Prut, "Spatio-temporal firing patterns in the frontal cortex of behaving monkeys," *Journal of Physiology-Paris*, vol. 90, no. 3, pp. 249–250, 1996.
- 16) S. Song, P. J. Sjöström, M. Reigl, S. Nelson, and D. B. Chklovskii, "Highly nonrandom features of synaptic connectivity in local cortical circuits," *PLoS biology*, vol. 3, no. 3, p. e68, 2005.
- 17) S. Lefort, C. Tómm, J.-C. F. Sarria, and C. C. Petersen, "The excitatory neuronal network of the c2 barrel column in mouse primary somatosensory cortex," *Neuron*, vol. 61, no. 2, pp. 301–316, 2009.
- 18) J.-n. Teramae, Y. Tsubo, and T. Fukai, "Optimal spike-based communication in excitable networks with strong-sparse and weak-dense links," *Scientific reports*, vol. 2, 2012.
- 19) N. Hiratani, J.-N. Teramae, and T. Fukai, "Associative memory model with long-tail-distributed hebbian synaptic connections," *Frontiers in computational neuroscience*, vol. 6, p. 102, 2013.
- 20) Y. Omura, M. M. Carvalho, K. Inokuchi, and T. Fukai, "A lognormal recurrent network model for burst generation during hippocampal sharp waves," *Journal of Neuroscience*, vol. 35, no. 43, pp. 14585–14601, 2015.
- 21) 寺前順之介, "脳の情報通信メカニズムと「揺らぎ」を生かした脳型情報処理," *電子情報通信学会誌*, vol. 99, no. 10, pp. 999–1004, 2016.
- 22) V. V. Klinshov, J.-n. Teramae, V. I. Nekorkin, and T. Fukai, "Dense neuron clustering explains connectivity statistics in cortical microcircuits," *PloS one*, vol. 9, no. 4, p. e94292, 2014.
- 23) S. Jaffard, B. Lashermes, and P. Abry, "Wavelet leaders in multifractal analysis," *Wavelet analysis and applications*, pp. 201–246, 2007.

四元数ニューラルネットを用いた9軸モーションセンサによる 前腕姿勢の推定

○ハフィズディン フィルダウス 小川毅彦 (拓殖大学)

Estimation of Forearm Posture by Nine-Axis Motion Sensor Using Quaternion Neural Network

* H. Firdaus and T. Ogawa (Takushoku University)

Abstract— Human motion and posture are quantitatively analyzed and evaluated using various sensors. Geometric algebra including quaternions can be used to describe the geometry of three dimensional space. Moreover, we can learn and estimate nonlinear relation between quaternion input and output by neural networks expanded to quaternion. In this study, we applied a quaternion neural network to estimate forearm posture from data with the nine-axis motion sensor.

Key Words: Quaternion neural networks, forearm posture, nine-axis motion sensor

1 はじめに

人間の動作を定量的に評価するために、各種センサを用いた動作解析が行われる¹⁾。動作を表現する物理量には、位置、速度、加速度や角度、角速度、角加速度等があり、3軸センサにより3次元データとして計測できる。そこで、動作を評価したい部位に多軸センサを取り付けて多次元データを計測し、そのデータを解析して動作の評価や認識が行われる。複数の物理量を同時に計測し、多次元データを得るための多軸モーションセンサが開発されている²⁾。本研究では、加速度・角速度・方位をそれぞれ3軸で計測できる9軸モーションセンサによる動作評価を考える。

空間の幾何学の記述に、四元数をはじめとする幾何学的代数を用いることができる³⁾。近年、幾何学的代数が見直され、ベクトルの代わりに四元数を空間表現に用いる例が増えてきている。特にコンピュータグラフィックスなど3次元空間内でのモデルの姿勢表現でしばしば四元数が用いられる⁴⁾。四元数を用いることで、ベクトルを用いる際に生じるいくつかの問題が解決できる。本研究では、9軸センサで得られる位置・加速度・角速度の計測データと人間の動作・姿勢を四元数で表現し、これらの間の関係を学習・推定することを考える。

近年、四元数に拡張されたニューラルネットが提案され、理論と応用の両面で検討が進められている⁵⁾⁶⁾。四元数ニューラルネットは、四元数入出力間の非線形関係を学習・推定することが可能である。四元数ニューラルネットのモデルとしては、層状モデルをはじめさまざまなものが提案されている。本研究では、層状モデルを対象として、加速度・角速度・方位の3次元データと、人間の3次元姿勢の関係を教師あり学習・推定を考える。

本研究では、9軸モーションセンサによる人間の前腕姿勢の推定を考える。モーションセンサによって前腕姿勢を推定することで、コンピュータとのインタフェースや、ロボットの遠隔制御などへの応用が考えられる。モーションセンサからのデータと姿勢との関係を、ニューラルネットによって学習し、さらに与えられたデータから姿勢を推定する。本研究では、モーション

センサの多次元データからの前腕姿勢の推定に、四元数ニューラルネットを適用することを考え、シミュレーションによって効果を示す。

2 9軸モーションセンサ

動作を表現する物理量として、位置、速度、加速度または角度、角速度、角加速度等がある。動作解析にはこれらを計測できるセンサや方法が必要である。速度・加速度については、それぞれ時間微分・積分の関係にあるので、いずれかが判れば残りは計算で求めることができる。角速度・角加速度についても同様である。さらに位置または角度が判れば、動作の初期位置や基準位置を設定することができる²⁾。

本研究では、ZMP社のIMU-Z 9軸モーションセンサを用いて人間の動作の学習・推定を行うことを考える。対象とする物理量は、加速度と角速度、方位である。それぞれ3軸センサによる3次元データを計測するため、合計9軸センサで9次元のデータを得ることができる。IMU-Zモーションセンサの外観と、USBによるPCへの接続の様子をFig.1に示す。

IMU-ZモーションセンサとPCの接続は、CAN-USBによる有線接続とBluetoothによる無線接続が可能である。また、複数のセンサモジュールを接続し、人体の各部の運動を計測することが可能である。本研究ではモーションセンサによって計測されたデータと姿勢の間の関係のニューラルネットによる学習の基本的な動作を確認することを目的とするため、CAN-USBによる有線接続かつ、1個のモジュールのみを用いることとする。CAN-USBによる有線接続の方法をFig.2に示す。

以下では、IMU-Zモーションセンサで測定できる物理量について述べる²⁾。

2.1 加速度

IMU-Zには加速度センサとして3軸加速度センサが使われている。モジュール内部には1方向の加速度を検出できる素子が6個配置され、 x, y, z の3軸において正負の加速度を検出する。測定範囲は $\pm 2g$ または $\pm 4g, \pm 8g$ で分解能は12bitである。加速度はセンサ内部でA/D変換され、デジタル信号として送信される。

2.2 角速度

角速度センサとして2軸角速度センサが用いられる。モジュール内部には、 x, y 軸周りの角速度を検出できる素子を持つ2軸センサと、 x, z 軸周りの角速度を検出できる2軸センサが配置され、それらの組み合わせで直交3軸の正負の角速度を検出する。測定範囲は $\pm 500\text{deg/s}$ または $\pm 2000\text{deg/s}$ で、分解能は12bitである。角速度についてもセンサ内部でA/D変換され、デジタル信号として送信される。

2.3 方位

方位計測には3軸地磁気センサを用いる。モジュール内部にある地磁気センサで、直交する x, y, z の3軸方向の磁気強度を検出し、それぞれの軸ベクトルの向きと大きさを合成して地磁気の方向を決定する。さらに決定した地磁気の方向に基づいて3次元の方位を求める。磁気の測定範囲は $\pm 0.7\sim 4.5\text{gauss}$ で、分解能は12bitである。方位についてもセンサ内部でA/D変換され、デジタル信号として送信される。



Fig. 1: ZMP IMU-Z nine-axis motion sensor, (a) sensor module, and (b) with CANUSB adapter.

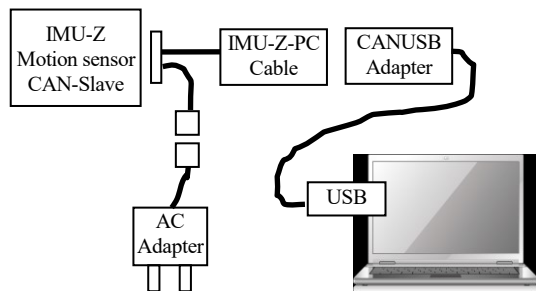


Fig. 2: Connection between IMU-Z and PC via CAN-USB.

3 前腕姿勢の推定

手話や身振り手振りなどの情報表現や、運搬や操作など対象物の制御など、人間の前腕動作の解析は重要な問題である。前腕動作の解析にはさまざまなセンサの使用が考えられるが、制御意図を推定するには筋電位、離れた地点から動作結果を認識するためには画像や超音波等、動作主体から見た動作結果を判別するためにはモーションセンサの使用が有効と考えられる。

著者らはこれまで、筋電位を用いた前腕動作の推定に四元数ニューラルネットを適用する検討を行ってきた⁷⁾。筋電位は筋肉を制御する神経パルスの重畳であり、筋電位を計測・解析することによって、筋肉の制御意図を推定することができる。しかし、筋電位と変

位の関係は非線形かつ負荷にも左右されるため、情報表現の判別への使用では誤差が問題となる。そこで本研究では、情報表現を直接計測できるモーションセンサを使用した前腕動作解析を行う。

本研究では、身振り手振りなどの情報表現をとらえてその動作を判別・認識することを目的に、モーションセンサによる前腕動作の推定を行う。手首付近に取り付けたモーションセンサによって、前腕の加速度・角速度・方位を計測し、それらをニューラルネットで学習・推定することを試みる。

人間の前腕は関節の回転により3次元空間でさまざまな姿勢をとることができる。本研究では、前腕の3次元空間内での姿勢とモーションセンサによるセンサデータとの関係を学習・推定することを考える。具体的には、前腕の屈曲・伸展（ピッチ）、回外・回内（ロール）、外旋・内旋（ヨー）回転を対象とし、それらに伴う加速度、角速度、方位のセンサから得られたデータとの関係を学習・推定する。本研究で試す前腕の回転動作をFig.3に示す⁸⁾。

3.1 センサデータの処理

前腕の屈曲・伸展に関する動作による9軸モーションセンサの波形をFig.4に示す。これらの物理量は、順に加速度、角速度、方位で、それぞれ3軸のデータである。これらの波形から、ニューラルネットの学習・テスト用のデータを作成する。本研究では、角速度をトリガとしてこの時点のセンサ値を、それぞれの姿勢を表すデータとみなし、学習・テスト用データとする。なお、9軸のセンサデータはそれぞれのデータの最大値と最小値によって、 $[-1.0, 1.0]$ に正規化して学習・推定用データとする。今回は、それぞれの姿勢に対し8個、合計48個のデータを作成した。

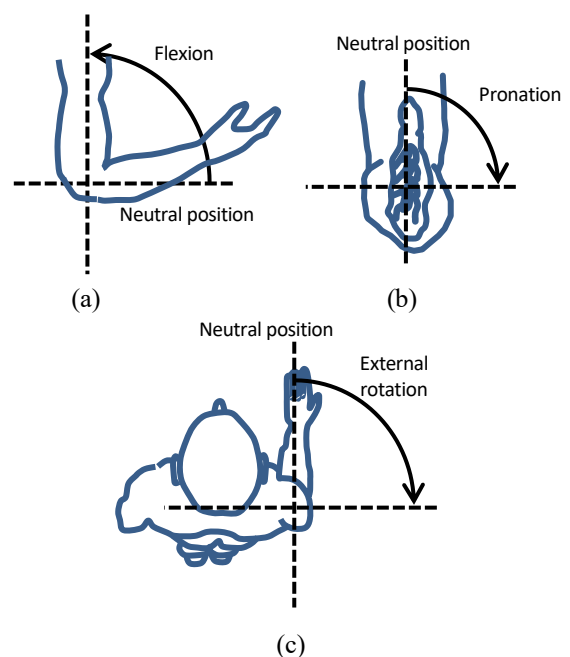


Fig. 3: Three forearm motions; (a) flexion, (b) pronation, and (c) external rotation

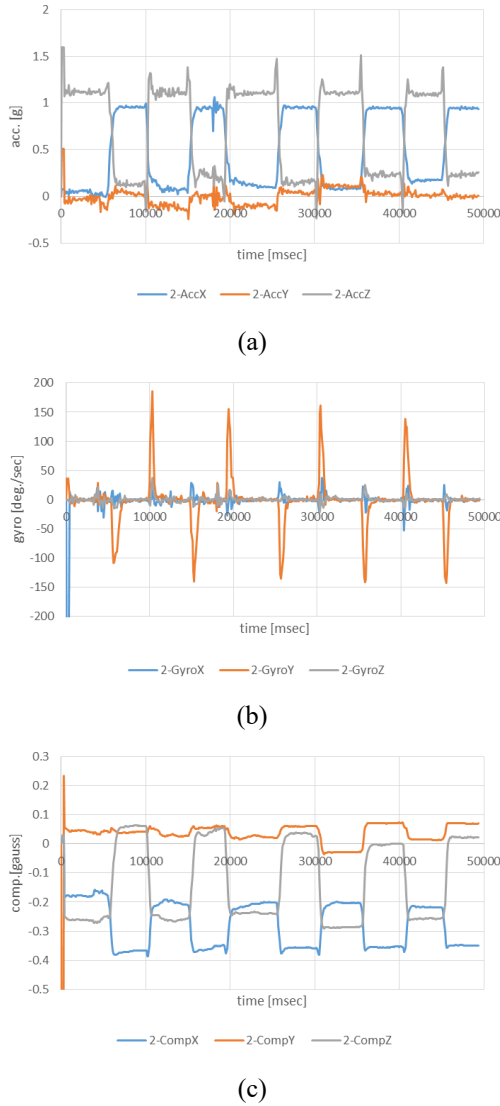


Fig. 4: Sensor data of nine-axis motion sensor by forearm motion on flexion and extension.

4 四元数ニューラルネット

本研究では、前腕動作に関わる物理量である加速度・角速度・方位と前腕姿勢の間の関係を四元数ニューラルネットによって学習・推定することを目的とする。前腕の姿勢を3つの関節の回転によって表現することを考え、9軸センサデータと関節の回転を四元数で表現する。これらの間の関係を四元数ニューラルネットで表現する。

本研究では入力、荷重および出力を四元数に拡張した多層型四元数ニューラルネットモデルを用いる。層数は入力層・中間層・出力層の3層型とする。中間層と出力層には、四元数入力の各部にシグモイド関数を適用する各部独立型の四元数ニューロンを用いる。ニューロンの伝達関数は、入力 $\mathbf{s} = s_0 + is_1 + js_2 + ks_3$ に対して

$$f(\mathbf{s}) = f(s_0) + if(s_1) + jf(s_2) + kf(s_3), \quad (1)$$

と表せる。ただし $f(x) = \tanh(x)$ である。学習時は四元

数に拡張された誤差逆伝搬法により、四元数入出力から四元数荷重の更新を行う。誤差関数は

$$E = \frac{1}{2} \sum_r (\mathbf{d}_r - \mathbf{y}_r) (\overline{\mathbf{d}_r - \mathbf{y}_r}) \quad (2)$$

とする。 \mathbf{d}_r と \mathbf{y}_r は r 番目素子に与えられた出力とネットワーク出力をそれぞれ表す。さらに中間・出力層間荷重の修正量 $\Delta \mathbf{w}_{qr}$ と、入力・中間層間荷重の修正量 $\Delta \mathbf{w}_{pq}$ はそれぞれ、

$$\begin{aligned} \Delta \mathbf{w}_{qr} &= ((\mathbf{d}_r - \mathbf{y}_r) \otimes (\mathbf{1} - \mathbf{y}_r) \otimes (\mathbf{1} + \mathbf{y}_r)) \cdot \overline{\mathbf{v}_q} \\ \Delta \mathbf{w}_{pq} &= \left((\mathbf{1} - \mathbf{v}_q) \otimes (\mathbf{1} + \mathbf{v}_q) \otimes \sum_r \delta_r \cdot \overline{\mathbf{w}_{qr}} \right) \cdot \mathbf{x}_p, \end{aligned} \quad (3)$$

となる。ただし、

$$\delta_r = (\mathbf{d}_r - \mathbf{y}_r) \otimes (\mathbf{1} - \mathbf{y}_r) \otimes (\mathbf{1} + \mathbf{y}_r) \quad (4)$$

とする。ここでは、“ \otimes ”の記号によって四元数の要素ごとの乗算を表す。また、 \mathbf{v}_q および \mathbf{x}_p は、それぞれ q 番目中間素子からの出力および、 p 番目入力を表す。さらに、 $\mathbf{1}$ は各要素が1の四元数とする。これらの荷重修正量に基づいて、

$$\mathbf{w}^{new} = \mathbf{w}^{old} - \varepsilon_t \Delta \mathbf{w} \quad (5)$$

に基づき入力を更新する。ここで ε_t は学習係数、 \mathbf{w} は \mathbf{w}_{qr} または \mathbf{w}_{pq} を意味する。この手順を繰り返すことで四元数入出力間の関係を学習する⁷⁾。

5 シミュレーション

四元数ニューラルネットによるモーションセンサデータからの前腕姿勢推定の動作を示すために、上記の学習・推定のデータを用いて計算機シミュレーションを行った。ネットワークのパラメータはTable 1の通りである。入力は3つの物理量の各3軸データを四元数の各部に対応させ、加速度 $\mathbf{a} = ia_1 + ja_2 + ka_3$ 、角速度 $\boldsymbol{\omega} = i\omega_1 + j\omega_2 + k\omega_3$ 、方位 $\mathbf{d} = id_1 + jd_2 + kd_3$ とし、出力は前腕の回転による姿勢 $\mathbf{r} = ir_1 + jr_2 + kr_3$ とした。

前腕姿勢のデータは、それぞれの動作に対し8個あるので、そのうち7個を学習データ、残りの1個をテストデータとした。データのばらつきを考慮し、テストデータを変えた8通りの学習データ・テストデータの組で計算機シミュレーションを行った。学習の際の誤差の推移をFig. 5に示す。8回の試行ともに、学習回数100000回で誤差は十分に減少しており、学習が正しく行われたことを確認した。

学習した四元数ニューラルネットに対し、テストデータを与えて前腕の回転姿勢 $\mathbf{r} = ir_1 + jr_2 + kr_3$ を推定した。それぞれの回転姿勢に対する推定結果の四元数の各要素 (r_1, r_2, r_3) をFig. 6に示す。y軸周りの回転すなわち回内動作を表す r_2 の推定値のばらつきがやや大きいものの、ほぼそれぞれの回転を表す推定値が得られている。推定結果の四元数を3次元グラフ上にプロットしたものをFig. 7に示す。図より、それぞれの姿勢が分類されていることがわかる。

以上、四元数ニューラルネットによる9軸モーションセンサからの前腕姿勢推定の基本動作を確認した。

Table 1 Network parameters

Parameters	Values
Number of input neurons	3
Number of hidden neurons	5
Number of output neurons	1
Learning rate ε	0.001
Momentum rate α	0.9
Number of training epoch	100000

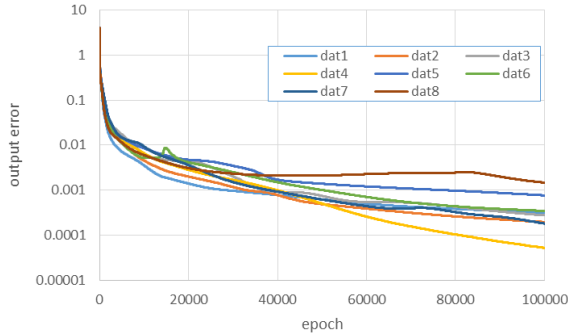
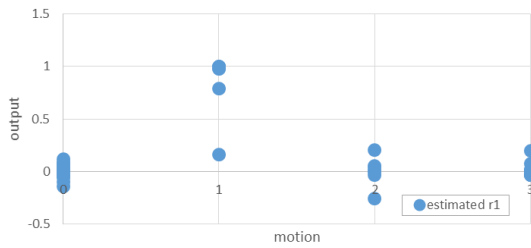
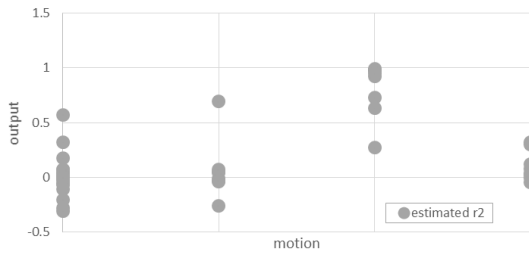


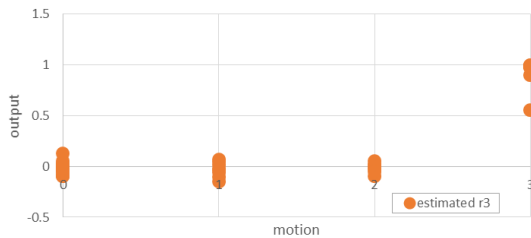
Fig.5 Mean squared output error in learning.



(a)



(b)



(c)

Fig.6 Each element of the output quaternion for each motion, (a) r_1 , (b) r_2 , and (c) r_3 ; motion = 0, 1, 2, 3 represent flexion, pronation, and external rotation, respectively.

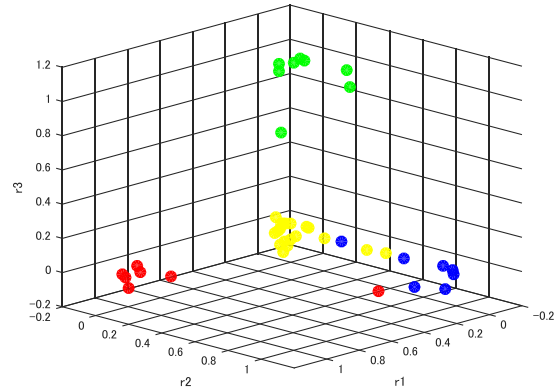


Fig.7 Simulation result of the estimated quaternions as the 3-D plot; red, blue, green, and yellow markers mean the flexion, pronation, external rotation, and neural positions.

6 おわりに

本研究では、9軸モーションセンサからの前腕姿勢推定の問題を対象に、四元数ニューラルネットの導入を検討した。ニューラルネットの構成は3層型モデルとした。手首部に取り付けたモーションセンサで計測される加速度・角速度・方位の各3軸データと、前腕の屈曲・伸展、回外・回内、外旋・内旋の運動の関係を四元数ニューラルネットで学習・推定することを試みた。本研究では測定実験で得たモーションセンサによるデータで計算機シミュレーションを行い、前腕動作の学習・推定を示した。今後は姿勢や回転の量的関係についての検討、スポーツ動作の解析やロボット制御への応用のための検討などを考えている。

謝辞

本研究の一部は科学研究費補助金（基盤 C，課題番号 26330284）の援助を受けた。

参考文献

- 1) 牧川, 吉田, 南部, 塩澤, 岡田: ヒト心身状態の計測技術, コロナ社 (2010)
- 2) 齊藤浩一: e-nuvo IMU-Z で学ぶ 9 軸モーションセンサ, ZMP パブリッシング (2010)
- 3) 金谷健一: 幾何学と代数系 Geometric Algebra -ハミルトン, グラスマン, クリフォード, 森北出版 (2014)
- 4) F. Dunn, I. Parberry, 松田晃一訳: 実例で学ぶゲーム 3D 数学, オライリージャパン (2008)
- 5) T. Nitta: An extension of the back-propagation algorithm to quaternions, Proc. of Int'l Conf. on Neural Information Processing, 1, 247/250 (1996)
- 6) N. Matsui, T. Isokawa, H. Kusamichi, F. Peper and H. Nishimura: Quaternion neural network with geometrical operators, Journal of Intelligent and Fuzzy Systems, 15, 3-4, 149/164 (2004)
- 7) 小川, フィルダウス: 四元数ニューラルネットを用いた筋電位による前腕姿勢の推定, 第 10 回コンピュータショナル・インテリジェンス研究会 (2016)
- 8) 牧川, 吉田: 運動のバイオメカニクス, コロナ社 (2008)

並列 RBF ネットワークに基づく紙の汚れ計測

○名取直毅 長谷部光威 古賀敏之 齊藤廣大 (株式会社東芝)

Paper Soil Measurement based on Parallel RBF Networks

* N. Natori, M. Hasebe, T. Koga and H. Saito (Toshiba Corporation)

Abstract— This document describes a new method for paper soil measurement. The method is based on radial basis functions (RBFs). A measurement system including a number of dedicated RBF networks connecting in parallel is proposed. The networks are modeled on the Weber-Fechner law in psychophysics in order to match human sense of fitness. This makes a contribution to enhancement in the plausibility.

Key Words: RBF networks, Paper soil measurement, Weber-Fechner law, Psychophysics

1 はじめに

センシングと信号処理に基づく外観検査の自動化技術は、様々な産業分野に応用されている。外観検査の主な目的は、対象物に付着した異物や対象物の傷、欠陥を検出することである。例えば、紙を媒体とする出版物の外観検査においては、紙の破れや穴、印刷ムラやインク滲みなどが具体的な検出項目となる。一般にこうした外観検査を自動化する場合、対象物から何らかの特徴量を抽出し、抽出された特徴量があらかじめ設定された正常範囲内か否かを確認し、もし範囲外であれば異物や傷、欠陥が「有る」とみなす。外観検査においては、対象物の物理特性を反映して抽出される何らかの特徴量（例えば輝度やコントラストなど）が仮に連続的な値をとるとしても、そこから異物や傷、欠陥の「有無」を判断することが重要である。

これに対して、同じ紙であっても流通した紙幣を対象物とする場合は、汚れの「度合い」が問題となる。なお、ここで、汚れとは、上述の異物や傷、欠陥といった製造工程上の直接的かつ外的な要因によるものではなく、紙の変色や印刷の退色など、主に通常使用や経年劣化に起因する状態変化を意味するものとする。例えば、海外各国の中央銀行は、市中銀行を通して流通紙幣を定期的に回収し、汚れ度合いに応じてそれらを再流通させるか廃棄処分するかを判断している。また、その判断基準は、その国や地域の政策、その時々での紙幣の発行量や流通量などによっても異なり一定ではない。

こうした中で、古くは熟練者（人間）が汚れ度合いを目視で判断していたが、紙幣流通量の増加に伴い、現在では多くの国が紙幣処理機による自動化、省力化を推進している。従来の紙幣処理機に搭載されている汚れ度合い計測モジュールは、外観検査の自動化技術を応用したもので、紙幣から抽出した物理特徴量（特定領域の輝度やコントラストなど）を汚れ度合い計測量とみなし、エンドユーザ自身が状況に応じて閾値パラメータを調整した上で、再流通可能か否かを判断する。しかしながら、こうした従来の汚れ度合い計測モジュールは、以下の2つの理由によってエンドユーザの要望を十分に満たしていない。1つは、汚れ度合いという概念が感覚的なものであるため、物理量の領域で閾値パラメータを調整しても、それにより生じる判断結果への影響が熟練者の感覚とずれること。もう1つは、紙幣には製造上のばらつきがあり、例えば単純に輝度が高いほど汚れ度合いが低いとは言えず、いわ

ゆる多峰性があるため、判断結果の順序が熟練者の感覚と食い違うことである。（熟練者はその多峰性を理解し視覚情報を高度に補正した上で汚れ度合いを判断していると考えられる。）

こうしたことから、筆者らは、まず、精神物理学（psychophysics）におけるヴェーバー-フェヒナーの法則¹⁾（Weber-Fechner law）を応用し、人間が「紙幣の輝度」という網膜刺激を受けてから「汚れ度合い」という脳内感覚を得るまでのプロセスを単純な関数でモデル化した。次に、特徴空間内で紙幣が多峰性をもって分布することを RBF ネットワークで自己組織的に表現した。これらの手法に基づき、熟練者の感覚との整合性を高め、かつ、紙幣製造ばらつきの影響を低減する新たな汚れ度合い計測技術を開発し、当社紙幣処理機向けにアルゴリズムを実装した。

本論文では、まず、提案手法について説明し、次に、輝度変化を模擬した加工データを用いて本手法によって自己組織化される紙幣の輝度と汚れ度合いの関係を示し、最後に、熟練者によって選別された再流通可能な紙幣（fit notes）および廃棄処分すべき紙幣（unfit notes）を当社紙幣処理機で処理して従来手法との比較を行う。

2 提案手法

2.1 ヴェーバー-フェヒナーの法則とその応用

精神物理学においては、古くからヴェーバー-フェヒナーの法則が知られている。これによれば、物理界の刺激 S に対する精神界の感覚 R は、Fig. 1に示すような対数関数で表される。

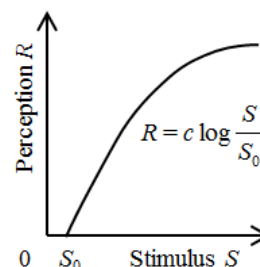


Fig. 1: Weber-Fechner law

ただし、 S_0 は刺激閾値と呼ばれ、感覚を生じさせる刺激の最小値である。

提案手法では、熟練者の感覚との整合性を向上させ

るべく、紙幣の輝度と汚れ度合いをそれぞれ刺激 S と感覚 R とみなし、この法則にしたがって紙幣の輝度を汚れ度合いへ変換する。ただし、筆者らがこれまでに行ってきた熟練者へのヒアリングや心理学的実験などから、汚れ度合いという感覚には「綺麗」と「汚い」という二極性があると考えられている。仮に、刺激 S と感覚 R に符号を与え、正領域を綺麗、負領域を汚いと定義すると、Fig. 2のような特性となる。

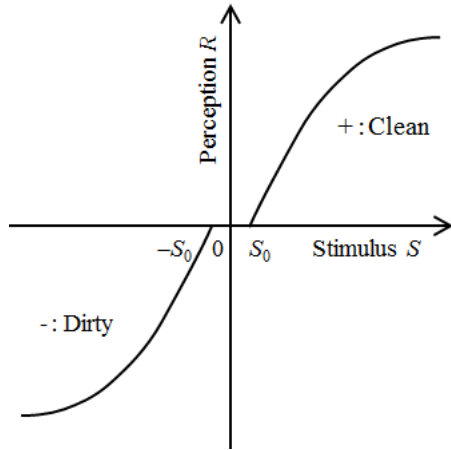


Fig. 2: Bipolar soil measurement

ここで、刺激閾値 S_0 の値が十分に小さいと仮定すると、特性全体としてはS字の曲線に近づくことが分かる。そこで、提案手法では、刺激 S に対応する紙幣の輝度 x と感覚 R に対応する適合度 y の関係をFig. 3に示すようなガウス関数にしたがう曲線で置き換える。

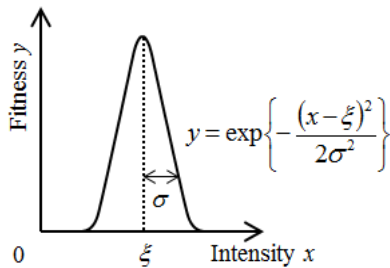


Fig. 3: Gaussian function

ただし、 ξ は中心パラメータ、 σ は幅パラメータである。なお、汚れ度合い (soil degree) と適合度 (fitness) は、どちらもこの分野の用語であるが、汚れ度合いを y' とし、ともに $[0, 1]$ の値をとり得るとすると、これらは、

$$y' = 1 - y$$

の関係にあり本質的には同じとみなしてよい。ここで、 ξ は、適合度 y が最大値となる x であり、具体的には、製造直後の汚れのない紙幣がもたらす輝度である。前述の通り、汚れには紙の変色と印刷の退色があるが、一般に、変色は輝度が下がる方向、退色は輝度が上がる方向にあるため、Fig. 3に示すように ξ を中心とする両側の特性が必要となる。

2.2 並列 RBF ネットワークによる自己組織化

これまでの、紙幣の輝度 x と適合度 y とが1対1であるかのように述べてきた。事実、従来の汚れ度合い計測の中には、主に演算コスト上の理由から、入力される

膨大な情報量のデータから特定の領域を切り出し、その領域に含まれる平均輝度 (スカラー量) を紙幣の輝度 x として処理するものも少なくない。しかし、特定の領域に限定せず、かつ、紙幣全面に一律ではなく偏在するような汚れも捉えたいという紙幣処理機市場の要求は以前より存在し、例えば複数の領域を設定するといった従来延長上の改良手法も見られるが、近年、演算プロセッサやメモリなどのハードウェア性能が飛躍的に向上していることと相まって、汚れ度合い計測を大幅に高性能化、高機能化する機運が高まっている。

そこで、本論文では、紙幣全面を格子状に覆う多数の局所領域を定義し、それぞれの局所領域の中で赤・緑・青の輝度から局所適合度を求め、それらの局所適合度を重み付き加算することで全面の適合度を得る手法を提案する。前述の通り、紙幣製造ばらつきの影響により、赤・緑・青の輝度からなる特徴空間内で局所適合度には多峰性が生じる。これに対応すべく、提案手法では複数のガウス関数型球状基底関数 (Radial Basis Function; RBF) ユニットからなるRBFネットワークをそれぞれの局所領域に配置し、さらに、それらのRBFネットワークを並列結合する。本論文では、これを「並列RBFネットワーク」と呼ぶ。それぞれのRBFネットワークは独立したネットワークパラメータをもつが、それらは、並列RBFネットワーク全体の出力誤差を最小化する方向へと自己組織化される。

紙幣全面の適合度 Y を以下のように表す。

$$Y = \sum_i w_i y_i$$

ただし、 y_i は i 番目の局所領域から求まる局所適合度、 w_i は y_i の重みパラメータである。本手法では、局所適合度 y_i を i 番目の局所領域に配置されたRBFネットワークの出力とする。すなわち、 i 番目のRBFネットワークの j 番目のRBFユニットの出力を o_{ij} 、 o_{ij} の重みパラメータを m_{ij} とすると、局所適合度 y_i は、

$$y_i = \sum_j m_{ij} o_{ij}$$

で求まる。前述の通り、本手法のRBFはガウス関数型であるから、 o_{ij} は、

$$o_{ij} = \exp \left\{ - \sum_k \frac{(x_{ik} - \xi_{ijk})^2}{2\sigma_{ijk}^2} \right\}$$

のように表される。ただし、 x はその局所領域から得られる輝度、 ξ は中心パラメータ、 σ は幅パラメータ、 k はデータの次元を表す。

この並列RBFネットワークに含まれる各ネットワークパラメータは、以下の出力誤差 ε を最小化するように、学習則によって逐次的に自動調整される。

$$\varepsilon(t) = \frac{1}{2} \{A - Y(t)\}^2$$

なお、上式において、 A はいわゆる教師信号であり、具体的には、熟練者が判断した汚れ度合いに基づく値である。本論文では、基本的な学習則である最急降下法を用いて、時刻 $t+1$ における出力 $Y(t+1)$ が時刻 t における出力 $Y(t)$ よりも教師信号 A に近づくよう、出力誤差 ε の偏微分 (勾配) 方向へと各ネットワークパラメータを更新する。すなわち、

$$w_i(t+1) = w_i(t) - \alpha \frac{\partial \varepsilon(t)}{\partial w_i}$$

$$m_{ij}(t+1) = m_{ij}(t) - \alpha \frac{\partial \varepsilon(t)}{\partial m_{ij}}$$

$$\xi_{ijk}(t+1) = \xi_{ijk}(t) - \alpha \frac{\partial \varepsilon(t)}{\partial \xi_{ijk}}$$

$$\sigma_{ijk}(t+1) = \sigma_{ijk}(t) - \alpha \frac{\partial \varepsilon(t)}{\partial \sigma_{ijk}}$$

ただし、 α は学習率（一定）である。Fig. 4は、ある局所領域に配置されたRBFネットワーク内のRBFユニットが自己組織化された様子を可視化したものである。

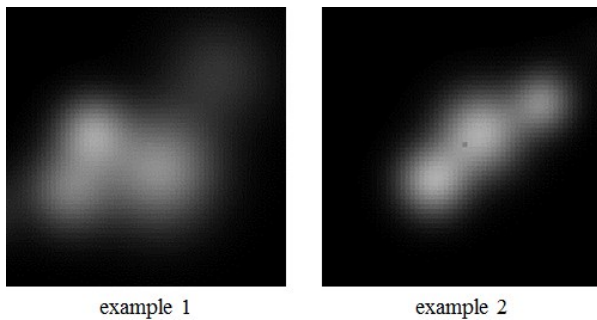


Fig. 4: Visualization of RBF units in feature space

Fig. 4からも分かるように、局所適合度は、いくつかのRBFユニット出力の重み付き加算で表される。

3 加工データによる適合度特性の確認

自己組織化された並列 RBF ネットワークの適合度特性を確認するために、輝度変化（すなわち汚れ）を模擬した加工データ 10 件でシミュレーションを実施した。オリジナルの輝度を 100%とすると、これを一律に 50%から 105%の倍率で変化させたものを合計 240 件（オリジナル 10 件を含む）作成した。学習においては「学習率」や「学習回数」などのハイパーパラメータが存在するが、本論文ではこれを実験的に求めることとした。

まず、Fig. 5 から Fig. 8 は、学習率をそれぞれ 0.5, 0.1, 0.01, 0.001 として学習したときの適合度特性を示したものである（学習回数は 4000 回）。

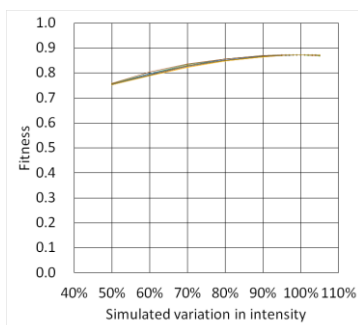


Fig. 5: Fitness curve ($\alpha=0.5$, iter=4000)

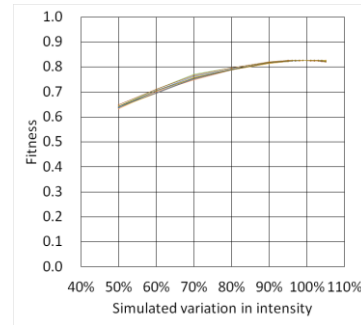


Fig. 6: Fitness curve ($\alpha=0.1$, iter=4000)

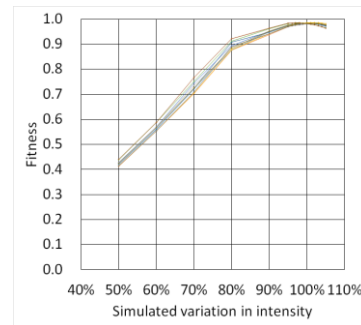


Fig. 7: Fitness curve ($\alpha=0.01$, iter=4000)

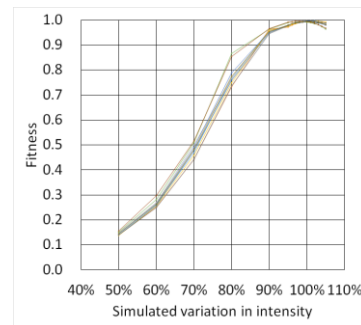


Fig. 8: Fitness curve ($\alpha=0.001$, iter=4000)

これらの図から、所望の S 字曲線に最も近いのは学習率が 0.001 のときであることが分かる。よって、今回の紙幣処理機への適用にあたっては、学習率 0.001 を採用した。

次に、学習回数による適合度特性の違いを比較する。Fig. 8 の 4000 回に対し、Fig. 9 および Fig. 10 は、学習回数をそれぞれ 2000 回、1000 回に減らしたときの適合度特性を示したものである（学習率は 0.001）。

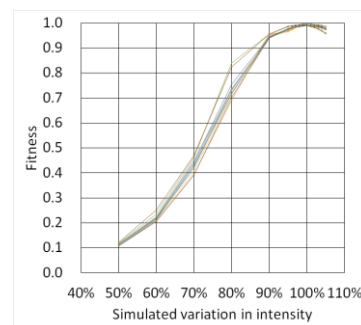


Fig. 9: Fitness curve ($\alpha=0.001$, iter=2000)

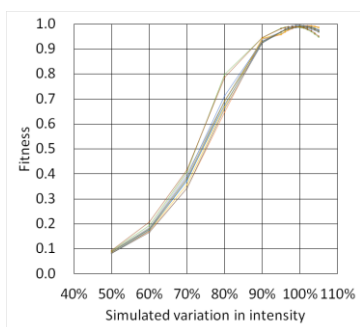


Fig. 10: Fitness curve ($\alpha=0.001$, iter=1000)

これらの図には違いがほとんど見られないが、Fig. 8の曲線が最も滑らかであるため、今回の紙幣処理機への適用にあたっては、学習回数 4000 回を採用した。

4 紙幣処理機での従来手法との比較

提案手法に基づくアルゴリズムを当社紙幣処理機へ実装した。紙幣処理機には従来手法（特定領域の平均輝度）に基づくアルゴリズムも実装されており、紙幣処理機のログデータを収集することで、両アルゴリズムの比較を行える。今回、熟練者の一定基準によって選別された再流通可能の紙幣（fit notes）40 枚および廃棄処分すべき紙幣（unfit notes）60 枚を処理した。

Fig. 11 および Fig. 12 は、それぞれ従来手法および提案手法による汚れ度合い計測のログデータからヒストグラムを作成したものである。

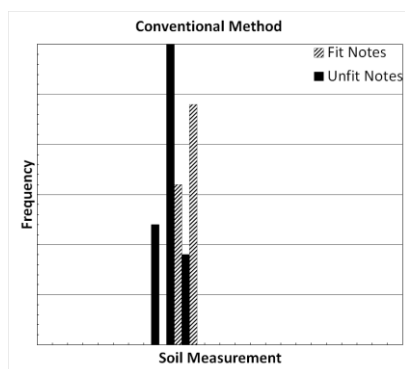


Fig. 11: Soil measurement on actual machine (Conventional method)

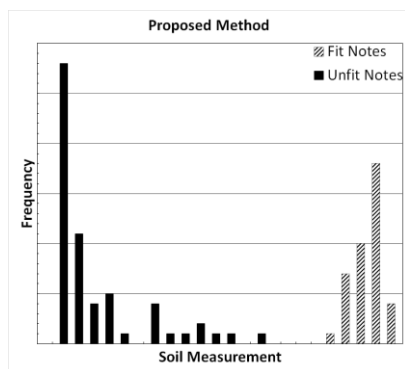


Fig. 12: Soil measurement on actual machine (Proposed method)

これらの図から、従来手法では fit notes か unfit notes

かの判断が難しいような場合でも、提案手法では熟練者と同様に両者を明確に判断し、熟練者の感覚との整合性が向上していることが分かる。

5 おわりに

精神物理学におけるヴェーバー-フェヒナーの法則を応用した物理量から感覚量への変換と、並列結合した大規模な RBF ネットワークの自己組織化に基づき、新たな汚れ度合い計測技術を開発した。これまでに、当社紙幣処理機である海外向け整理機や海外向け鑑査機の一部に適用され、今後も適用範囲を拡大予定である。

本研究は、当社 AI 活用による最先端要素技術への取り組みの一環として位置付けられており、今後も GAN 手法などを応用したデータ自動生成や、ハイパーパラメータの自動最適化などを検討していく。

参考文献

- 1) 池田：心理測定法, p.135/p.147, 放送大学教育振興会, (1993)

不精密ルールを考慮したラフ集合による知識獲得及び ファジィ判別分析

○関宏理 (大阪大学) 山下亮 (伊藤忠商事) 乾口雅弘 (大阪大学)

Knowledge Acquisition by Rough Set Considering Imprecise Rules and Fuzzy Discriminant Analysis

*H. Seki (Osaka University), R. Yamashita (ITOCHU Corporation) and M. Inuiguchi (Osaka University)

Abstract— This paper firstly gives new interpretation for the fuzzy inference model with imprecise rules. Second, it proposes how to obtain the minimal decision rules of the imprecise rules by using rough set theory. Finally, the proposed method is applied to medical diagnosis, and shown to superior to the conventional model from point of view of accuracy and number of rules in this paper.

Key Words: Approximate reasoning, Fuzzy inference model, Rough set, Imprecise rules

1 はじめに

Mamdani がファジィ推論の概念をスチームエンジン実験装置の制御へ適用して以来, 様々な分野でファジィ推論の研究と応用が行われてきた¹⁾.

ファジィ推論のルールはファジィIf-Then ルールで記述されるため, 直感的にも理解しやすくエキスパートの知識を表現できることからエキスパートシステムとして様々な分野へ応用されてきた. しかしながら, すべての入力項目はファジィIf-Then ルールの前件部にセットされるため, 入力数が増加すると規則数も膨大になってしまう.

一方, 決定に最小限必要な属性や条件の長さが極小な決定ルール (以後, 極小決定ルール) を決定表から抽出する方法として, ラフ集合に基づく方法がある^{2, 3)}. 矛盾を含んだデータを扱う場合, 極小決定ルールを使用することにより矛盾を取り除くことが可能である. また, ラフ集合に基づく手法では不精密ルールを獲得する方法⁴⁾が提案されており, ファジィ推論も同様に不精密ファジィルールを用いた推論モデル⁵⁾が提案されており, 共に判別分析において良好な結果が得られている. しかしながら, 不精密ルールを用いたファジィ推論も従来のファジィ推論モデルと同様に入力数が増加すると規則数も増加してしまう.

したがって本研究では, ラフ集合に基づく手法を用いて精密および不精密な極小決定ルールを獲得し, ファジィ推論モデルを提案する. また, 理論的性質として等価性を用いることにより不精密ルールの解釈として, 後件部を求める重みはファジィ集合の面積とみなすことが可能であることから, 後件部ファジィ集合をもつ不精密ルールと解釈できることを示す. また, 提案手法を医療診断へ応用することにより, その有効性を示す.

2 準備

2.1 ファジィ推論モデル

本章では代表的なファジィ推論モデルである Min-Max-重心モデル (Mamdani の推論モデル)¹⁾ と代数積-加算-重心モデル^{7, 8, 9, 10)}, 簡略化ファジィ推論モデル¹¹⁾ について述べる.

一般的なファジィ推論モデルの規則は以下のような IF-THEN ルールで構成される.

$$\text{Rule } R_i = \begin{cases} x_1 \text{ is } A_i^1 \text{ and } x_2 \text{ is } A_i^2 \text{ and } \cdots \text{ and } x_n \text{ is } A_i^n \\ \rightarrow y \text{ is } B_i \end{cases} \quad (1)$$

ここで, x_1, x_2, \dots, x_n は前件部の入力変数, $A_i^1, A_i^2, \dots, A_i^n$ はファジィ集合, B_i は後件部ファジィ集合を表す. また $i = 1, 2, \dots, M$ であり, M は規則の総数を表す. なお, 後件部のファジィ集合 B_i が関数に置き換えられた場合を T-S 推論モデル⁶⁾, 定数に置き換えられた場合を簡略化ファジィ推論モデル¹³⁾ と呼ぶ.

2.1.1 Min-Max-重心モデル

本節ではまず, Mamdani の推論モデルとして知られる Min-Max-重心モデル¹⁾ について説明する. Min-Max-重心モデルの規則は式 (1) のように与えられる.

入力 " $x_1^0, x_2^0, \dots, x_n^0$ " とファジィルール " $A_i^1, A_i^2, \dots, A_i^n \rightarrow B_i$ " から導かれた各々の推論結果 B'_i は以下のように与えられる (図 1 参照).

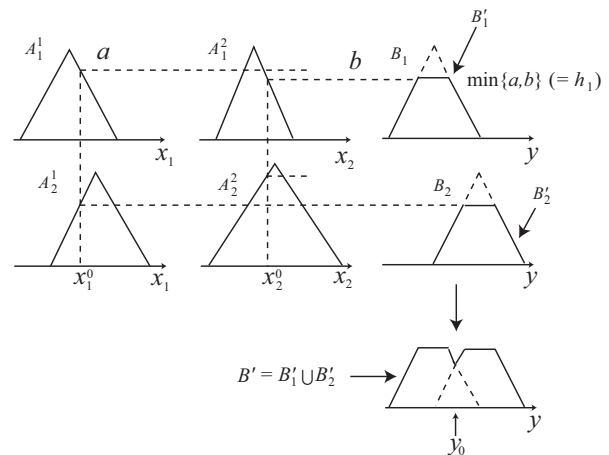


Fig. 1: Min-Max-重心モデル

前件部 “ $A_i^1, A_i^2, \dots, A_i^n$ ”へ入力 “ $x_1^0, x_2^0, \dots, x_n^0$ ”が与えられたとき、その適合度 h_i は

$$h_i = \min\{A_i^1(x_1^0), A_i^2(x_2^0), \dots, A_i^n(x_n^0)\} \quad (2)$$

のように与えられる。したがって、その規則の推論結果 B'_i は

$$B'_i(y) = \min\{h_i, B_i(y)\} \quad (3)$$

のように導くことができる。

式 (1) の最終推論結果 B' は B'_1, B'_2, \dots, B'_M を演算 \max で結合することにより得ることができる。すなわち、

$$B'(y) = \max\{B'_1(y), B'_2(y), \dots, B'_M(y)\} \quad (4)$$

となる。

推論結果 B' に対する代表値 y_0 は以下のように B' の重心を求めることにより得られる。

$$y_0 = \frac{\int y \cdot B'(y) dy}{\int B'(y) dy} \quad (5)$$

2.1.2 代数積-加算-重心モデル

次に本節では、代数積-加算-重心モデル^{7, 8, 9, 10)}について述べる (図2参照)。代数積-加算-重心モデルの規則は Min-Max-重心モデルと同様に、式 (1) で与えられる。

入力 “ $x_1^0, x_2^0, \dots, x_n^0$ ”とファジィルール “ $A_i^1, A_i^2, \dots, A_i^n \rightarrow B_i$ ”から得られた各推論結果 B'_i は以下のように求められる。

前件部 “ $A_i^1, A_i^2, \dots, A_i^n$ ”へ入力 “ $x_1^0, x_2^0, \dots, x_n^0$ ”が与えられたとき、その適合度 h_i は以下のように求められる。

$$h_i = A_i^1(x_1^0) \cdot A_i^2(x_2^0) \cdots A_i^n(x_n^0) \quad (6)$$

ここで “ \cdot ”は代数積を意味する。したがって、その推論結果 B'_i は以下のように導かれる。

$$\begin{aligned} B'_i(y) &= A_i^1(x_1^0) \cdots A_i^n(x_n^0) \cdot B_i(y) \\ &= h_i \cdot B_i(y) \end{aligned} \quad (7)$$

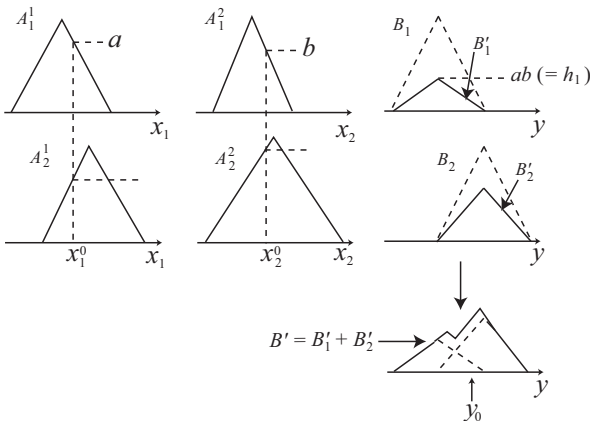


Fig. 2: 代数積-加算-重心モデル

式 (1) の最終推論結果 B' は B'_1, B'_2, \dots, B'_M を代数和 (+) で統合することにより得ることができる。すなわち、

$$B'(y) = B'_1(y) + B'_2(y) + \cdots + B'_M(y) \quad (8)$$

となる。

推論結果 B' の代表値 y_0 は重心法により以下で求めることができる。

$$y_0 = \frac{\int y \cdot B'(y) dy}{\int B'(y) dy} \quad (9)$$

2.2 簡略化ファジィ推論モデル

ファジィルールの後件部の実数値を前件部適合度で荷重平均して求める推論法を簡略化ファジィ推論法¹¹⁾と呼ぶ。前件部は複数の条件の連言で与えられ、前件部適合度は前件部ファジィ集合の帰属度の各条件の適合度の代数積 (もしくは \min) で与えられる。簡略化ファジィ推論モデル¹¹⁾では、 j 番目のルールは

$$x_1 \text{ is } A_j^1, x_2 \text{ is } A_j^2, \dots, x_n \text{ is } A_j^n \rightarrow y \text{ is } y_j \quad (10)$$

という形で与えられる。ここで、 A_j^i は j 番目のルールの i 番目の前件部ファジィ集合、 y_j は j 番目のルールの後件部実数値を表す。ルール数を m とするとき、入力 $\mathbf{x}^0 = (x_1^0, \dots, x_n^0)$ に対する推論値 y_0 は次式で算出される。

$$y_0(\mathbf{x}^0) = \frac{\sum_{j=1}^m h_j y_j}{\sum_{j=1}^m h_j}, \quad h_j = \prod_{i=1}^n A_j^i(x_i^0) \quad (11)$$

本研究では、ラフ集合により求められたルールに見合ったファジィルールを考える。ラフ集合による条件部極少化により、前件部ファジィ集合の数が n 以下に削減される。

2.3 不精密ルールを用いた簡略化ファジィ推論モデル

通常、ルールの後件部はクラス D_i からなり、それにより D_i への帰属が推定される。このようなルールを精密ルールと呼ぶ。これに対し、後件部がクラスの合併集合 $D_i \cup D_j \cup D_k$ である場合は、そのルールから D_i か D_j か D_k に帰属することしか推定できない。このようなルールを不精密ルールと呼ぶ⁴⁾。合併するクラス数 p を定め、全組合せの合併クラスに対する不精密ルールを抽出すれば、推定精度が向上することが知られている⁴⁾。この不精密ルールの概念を簡略化ファジィ推論モデルに導入するモデルが提案されている⁵⁾。合併クラスが p 個の場合の不精密ファジィルールは次式のようなになる。

$$\begin{aligned} x_1 \text{ is } A_j^1, x_2 \text{ is } A_j^2, \dots, x_n \text{ is } A_j^n \\ \rightarrow y \text{ is } y_j^1 \text{ or } y_j^2 \text{ or } \cdots \text{ or } y_j^p \end{aligned} \quad (12)$$

ここで、 $y_j^i (i = 1, 2, \dots, p)$ を加味して最終推論結果 y_j を算出することになるが、本研究では、重み $w_j^i, i =$

$1, \dots, p$ を与え, $y_j = (\sum_{i=1}^p w_j^i y_j^i) / (\sum_{i=1}^p w_j^i)$ と定める. 複数のルールによる入力 \mathbf{x}^0 に対する推論値は次式で求められる.

$$y_0(\mathbf{x}^0) = \frac{\sum_{j=1}^m h_j \left(\frac{\sum_{i=1}^p w_j^i y_j^i}{\sum_{j=1}^p w_j^i} \right)}{\sum_{j=1}^n h_j} \quad (13)$$

2.4 ラフ集合^{2, 3)}

2.4.1 近似

議論の対象となる全体集合を U , U 上の同値関係を R とする. U の R による同値類 $[x]_R$ は, 同値関係 R のもとで互いに区別できない対象を表わす. 部分集合 $X \subseteq U$ に対して, R による X の下近似 $R_*(X)$ 及び上近似 $R^*(X)$ が以下のように定義される.

$$R_*(X) = \{x \in U \mid [x]_R \subseteq X\} \quad (14)$$

$$R^*(X) = \{x \in U \mid [x]_R \cap X \neq \emptyset\} \quad (15)$$

下近似と上近似の対 $(R_*(X), R^*(X))$ を X のラフ集合と呼ぶ. X のラフ集合は, 集合の包含関係の意味で, X の上下方向からの近似を表わす. $R_*(X)$ は X への帰属が疑わしくない対象の集合と解釈され, $R^*(X)$ は X に帰属する可能性のある対象の集合と解釈される.

2.5 ラフ集合と決定表

決定表は $(U, C \cup D, V, \rho)$ の4項組で表される. ここで, U は対象の集合, C は条件属性集合, D は決定属性集合である. V は属性値の集合, ρ は $U \times C$ or D から V の中への写像を表す. 表1に決定表の例を示す. この表は6人の患者 (Patient) の集合を頭痛 (Headache: 以後 H), 筋肉痛 (Musclepain: 以後 M), 体温 (Temperature: 以後 T) から流感 (Flu: 以後 F) か否かを類別したデータである. $U = \{x_1, x_2, \dots, x_6\}$, $C = \{H, M, T\}$, $D = \{F\}$, $V = \{\text{yes, no, normal, high, very high}\}$ である. 対象の集合が p 個の決定クラス D_k ($k = 1, 2, \dots, p$) に分割されたとすると, 条件属性の部分集合 $A \subseteq C$ に基づき, 各決定クラスの下近似, 上近似は次のように定義される.

$$A_*(D_k) = \{x \in U \mid [x]_A \subseteq D_k\} \quad (16)$$

$$A^*(D_k) = \{x \in U \mid [x]_A \cap D_k \neq \emptyset\} \quad (17)$$

ただし, $[x]_A$ は A による対象 x の同値類を表わす.

表1において, 流感である患者の集合 $D_1 = \{x_1, x_2, x_3, x_4\}$ と流感でない患者の集合 $D_2 = \{x_5, x_6\}$ とする. また, $A = \{H, M, T\}$ とすると

Table 1: 決定表の例

Patient	Headache	Musclepain	Temperature	Flu
x_1	no	yes	high	yes
x_2	yes	no	high	yes
x_3	yes	yes	very high	yes
x_4	no	yes	very high	yes
x_5	no	yes	normal	no
x_6	yes	no	high	no

$A_*(D_1) = \{x_1, x_3, x_4\}$, $A^*(D_1) = \{x_1, x_2, x_3, x_4, x_6\}$, $A_*(D_2) = \{x_5\}$, $A^*(D_2) = \{x_2, x_5, x_6\}$ が得られる.

決定表に関する指標として次式に示す近似精度 $\alpha_A(D_k)$ と近似の質 $\gamma_A(D_k)$ が提案されている¹²⁾.

$$\alpha_A(D_k) = \frac{|A_*(D_k)|}{|A^*(D_k)|} \quad (18)$$

$$\gamma_A(D_k) = \frac{|A_*(D_k)|}{|D_k|} \quad (19)$$

$|X|$ は集合 X の要素数である. $\alpha_A(D_k)$ は属性集合 A により決定クラス D_k がどの程度近似できるかを示し, $\gamma_A(D_k)$ は属性集合 A の情報により決定クラス D_k がどの程度の要素が明確に判定できるかを示している.

表1において, $D_1 = \{x_1, x_2, x_3, x_4\}$, $D_2 = \{x_5, x_6\}$, $A = \{H, M, T\}$ とすると, $\alpha_A(D_1) = 3/5 = 0.6$, $\gamma_A(D_1) = 3/4 = 0.75$, $\alpha_A(D_2) = 1/3 \cong 0.33$, $\gamma_A(D_2) = 1/2 = 0.5$ となる.

2.5.1 決定表と決定ルール

決定表は, 条件属性の値に対する決定属性の値を示す決定ルールを与えている. 例えば, 表1の第2行は

$$[H = \text{yes}] \wedge [M = \text{no}] \wedge [T = \text{high}] \Rightarrow [F = \text{yes}]$$

なるルール, すなわち, 「頭痛があり, 筋肉痛がなく, 体温が高ければ, 流感である」という決定ルールを示している.

決定ルール $\Delta \Rightarrow \Gamma$ の質を評価するために, 次のような指標が考えられている.

$$\text{Cer}(\Gamma \mid \Delta) = \frac{|\Gamma \wedge \Delta|}{|\Delta|} \quad (20)$$

$$\text{Cov}(\Gamma \mid \Delta) = \frac{|\Gamma \wedge \Delta|}{|\Gamma|} \quad (21)$$

$$\text{Supp}(\Gamma \mid \Delta) = \frac{|\Gamma \wedge \Delta|}{|U|} \quad (22)$$

$\text{Cer}(\Gamma \mid \Delta)$ は確信度と呼ばれ, そのルールがどの程度正しいかを示している. $\text{Cov}(\Gamma \mid \Delta)$ は被覆度と呼ばれ, そのルールにより結論 Γ を説明できる対象の割合を示している. $\text{Supp}(\Gamma \mid \Delta)$ は支持度と呼ばれ, そのルールを満たす対象が全体の何割にあたるかを示している.

表1において, $\Delta = [H = \text{no}] \wedge [M = \text{yes}]$, $\Gamma = [F = \text{yes}]$ とすると,

$$\text{Cer}(\Gamma \mid \Delta) = \frac{|x_1, x_4|}{|x_1, x_4, x_5|} = \frac{2}{3} \cong 0.67$$

$$\text{Cov}(\Gamma \mid \Delta) = \frac{|x_1, x_4|}{|x_1, x_2, x_3, x_4|} = \frac{2}{4} = 0.5$$

$$\text{Supp}(\Gamma \mid \Delta) = \frac{|x_1, x_4|}{|x_1, x_2, x_3, x_4, x_5, x_6|} = \frac{2}{6} \cong 0.33$$

となる.

2.5.2 ラフ集合による知識獲得

ここでは, 決定表に内在する極小決定ルールをすべて導出する方法である決定行列について述べる. 決定表 $(U, C \cup D, V, \rho)$ が与えられたとき, 決定属性集合 $B \subseteq D$ の属性値に基づき, 対象の集合が p 個の決定ク

ラス $D_k (k = 1, 2, \dots, p)$ に分割されたとする. このとき, 決定クラス D_k に応じて決定行列の (i, j) 成分は式 (23) のように定義される.

$$M_{ij}^k = \{(a, \rho(x_i, a)) \mid \rho(x_i, a) \neq \rho(x_j, a)\} \\ i \in K_k^+ = \{i \mid x_i \in C_*(D_k)\}, j \in K_k^- = \{j \mid x_j \notin D_k\} \quad (23)$$

ここで, $\mathcal{L}(M_{ij}^k)$ を (i, j) 成分の要素の論理和とすると, 式 (24) で示す論理式の主加法標準形を求めれば, 各連言項が $x \in D_k$ を導く条件の長さが極小な決定ルールの特徴部となる.

$$\bigvee_{i \in K_k^+} \bigwedge_{j \in K_k^-} \mathcal{L}(M_{ij}^k) \quad (24)$$

このような操作をすべての決定クラス D_k について行えば, 決定表内に内在する極小決定ルールのすべてが求められる.

表 1 において, $D_1 = \{x_1, x_2, x_3, x_4\}$ とした場合, 決定クラス D_1 の決定行列は以下ようになる.

	x_5	x_6
x_1	{(T,high)}	{(H,no),(M,yes)}
x_3	{(H,yes),(T,very high)}	{(M,yes),(T,very high)}
x_4	{(T,very high)}	{(H,no),(M,yes),(T,very high)}

式 (24) より

$$\bigvee_{i \in K_k^+} \bigwedge_{j \in K_k^-} \mathcal{L}(M_{ij}^k) \\ = [\text{T} = \text{high}] \wedge ([\text{H} = \text{no}] \vee [\text{M} = \text{yes}]) \\ \vee ([\text{H} = \text{yes}] \vee [\text{T} = \text{very high}]) \wedge ([\text{M} = \text{yes}] \vee \\ [\text{T} = \text{very high}]) \\ \vee [\text{T} = \text{very high}] \\ = ([\text{H} = \text{no}] \wedge [\text{T} = \text{high}]) \vee ([\text{M} = \text{yes}] \wedge [\text{T} = \text{high}]) \\ \vee ([\text{H} = \text{yes}] \wedge [\text{M} = \text{yes}]) \vee [\text{T} = \text{very high}]$$

これにより, 次に示す r_1, r_2, r_3, r_4 の 4 つの決定ルールが得られる.

$$r_1 : [\text{H} = \text{no}] \wedge [\text{T} = \text{high}] \Rightarrow [\text{F} = \text{yes}] \\ r_2 : [\text{M} = \text{yes}] \wedge [\text{T} = \text{high}] \Rightarrow [\text{F} = \text{yes}] \\ r_3 : [\text{H} = \text{yes}] \wedge [\text{M} = \text{yes}] \Rightarrow [\text{F} = \text{yes}] \\ r_4 : [\text{T} = \text{very high}] \Rightarrow [\text{F} = \text{yes}]$$

3 不精密ファジィルールの解釈および学習法

本章では不精密ルールを用いたファジィ推論の一解釈方法を述べる. 次に, ラフ集合に基づいて得られた簡潔ルールを簡略化ファジィ推論に適用して従来法と同様に学習するニューロファジィ手法を示す.

3.1 等価性による不精密ルールの解釈

文献¹⁴⁾では, ファジィ推論の各規則の統合に加算演算を用いた場合, その推論結果は後件部ファジィ集合の面積と重心から計算することができるが示されている. この後件部ファジィ集合の面積は後件部実数値への重みとみなすこともできる. すなわち, 式 (13) の重み w_j^i をファジィ集合の面積, 後件部実数値 y_j^i を

後件部ファジィ集合の重心とみなすことにより, 後件部を容易にファジィ集合へ置き換えることが可能である. したがって, 不精密ルールの後件部は p 個のファジィ集合の重なりとみなすことができ, 規則としても理解しやすいことがわかる.

3.2 学習アルゴリズム

ラフ集合に基づく手法でルールの前件部を簡略化し, 得られた簡潔ルールをファジィルールに変換して初期化する. この際, 簡潔ルールの条件部の分布に応じて入力空間をファジィ分割し前件部ファジィ集合を共通化した. 得られたファジィルールを基に $p = 1, p = 2$ ともに考慮して, 学習を行う.

ファジィルールのメンバシップ関数や後件部実数値をニューラルネットワークと同様に学習するニューロファジィ手法が提案されている¹¹⁾. 本研究では不精密ルールを用いた簡略化ファジィ推論モデルにニューロファジィ手法を適用し, 前件部ファジィ集合のパラメータ (中心 a_j^i , 幅 b_j^i) と後件部実数値 y_j^i , および重み w_j^i を学習する. 評価関数を

$$E = \frac{1}{2} \sum_{k=1}^l (y_0(\mathbf{x}_k^0) - t_k)^2 \quad (25)$$

とし, この最小化により学習する. ただし, $\{(\mathbf{x}_k^0, t_k) \mid k = 1, \dots, l\}$ は学習用データである. オンライン最急降下法により次の各パラメータの更新式が得られる.

$$y_j^{i \text{ new}} = y_j^{i \text{ old}} - \alpha (y_0(\mathbf{x}_k^0) - t_k) \frac{h_j}{\sum_{j=1}^m h_j} \quad (26)$$

$$a_j^{i \text{ new}} = a_j^{i \text{ old}} - \beta (y_0(\mathbf{x}_k^0) - t_k) \frac{h_j}{\sum_{j=1}^m h_j} \\ \cdot \frac{y_j^i \sum_{j=1}^m h_j - \sum_{j=1}^m h_j y_j^i}{(\sum_{j=1}^m h_j)^2} \\ \cdot \frac{h_j}{b_j^i (1 - \frac{x_k^i - a_j^i}{b_j^i})} \quad (27)$$

$$b_j^{i \text{ new}} = b_j^{i \text{ old}} - \gamma (y_0(\mathbf{x}_k^0) - t_k) \frac{h_j}{\sum_{j=1}^m h_j} \\ \cdot \frac{y_j^i \sum_{j=1}^m h_j - \sum_{j=1}^m h_j y_j^i}{(\sum_{j=1}^m h_j)^2} \\ \cdot \frac{h_j (x_k^i - a_j^i)}{(b_j^i)^2 (1 - \frac{x_k^i - a_j^i}{b_j^i})} \quad (28)$$

$$w_j^{i \text{ new}} = w_j^{i \text{ old}} - \delta (y_0(\mathbf{x}_k^0) - t_k) \frac{h_j}{\sum_{j=1}^m h_j} \\ \cdot \frac{y_j^1 (w_j^1 + w_j^2) - (w_j^1 y_j^1 + w_j^2 y_j^2)}{(w_j^1 + w_j^2)^2} \quad (29)$$

ここで, $\alpha, \beta, \gamma, \delta$ は学習係数で, 試行錯誤により決定する.

4 医療診断への応用

本章では提案手法の性能を実際に評価するため, 医療診断データに適用し, 従来モデル (簡略化ファジィ推論モデル) との比較, 検討を行う. 医療データは, 医師が実際に診断した 145 の糖尿病データを用いる¹⁵⁾.

糖尿病データは5入力1出力となっており、各入力項目は相対体重、空腹時血糖値、ブドウ糖値、インシュリン値、SSPG 値の5つで、範囲をすべて [0,1] に正規化する。糖尿病データは1群（臨床的糖尿）、2群（化学的糖尿）、3群（正常）に分類されており、推論に使う理想出力値は、それぞれ、1群を1、2群を0.5、3群を0としている。

なお、出力は3群に分類し、推論出力値0.75以上を1（臨床的糖尿）、0.25以上0.75未満を0.5（化学的糖尿）、0.25未満を0（正常）とする。

145のデータの73を教師データ、72を評価データとしてランダムに振り分け、1,000回学習を行い、従来法と比較する。それぞれ精密ルールと $p=2$ とする不精密ルールによるモデル化を考える。提案法では、出力値を専門家の知識に従い3群に分けこれをファジィルールに変換して初期化した。この際、簡潔ルールの条件部の分布に応じて入力空間をファジィ分割し前件部ファジィ集合を共通化した。これによりパラメータ数を削減した。従来法では、各入力空間を三つにファジィ分割し、その全組合せに対するファジィルールを生成した。後件部実数値は精密ルールで $y_j = 0$ 、不精密ルールで $y_j^1 = 0, y_j^2 = 1$ と定めた。

学習回数が1000回のときの10パターンのデータにおける提案モデルと従来モデルである簡略化ファジィ推論モデルの正答率を表2に示す。

Table 2: 正答率

データ	総ルールを用いた場合（従来法）		ラフ集合による簡潔ルールの場合（提案法）	
	$p=1$	$p=2$	$p=1$	$p=2$
1	50.0000	69.4444	93.0556	93.0556
2	51.3889	48.6111	97.2222	97.2222
3	54.1667	54.1667	93.0556	93.0556
4	58.3333	65.2778	94.4444	94.4444
5	56.9444	55.5556	95.8333	94.4444
6	62.5000	66.6667	97.2222	97.2222
7	44.4444	56.9444	97.2222	97.2222
8	58.3333	61.1111	97.2222	97.2222
9	52.7778	52.7778	94.4444	94.4444
10	51.3889	54.1667	66.6667	83.3333
平均	54.0278	58.4722	92.6389	94.1667

従来モデルと提案モデルで、提案モデルの方が正答率が大きく向上した。これは、ラフ集合を用いて簡潔ルールを抽出した後にファジィ化しているため、システムに冗長なものを取り除いて学習を行いシステム構築を行ったためだと考える。従来モデルでは総ルールを学習していたため、規則の競合が生じてしまいシステムとしての精度を低くしてしまっていたが、ラフ集合による知識獲得を行ったことでこの問題が解消されたためだと考えられる。また、 $p=1$ と比べると $p=2$ のほうが正答率が向上していることがわかる。これは不精密ルールによるパラメータ数の増加により、精度が向上しているからである。

次に、従来モデルのパラメータ数とルール長およびルール数を表3に、提案モデルの平均パラメータ数とルール長およびルール数を表4に示す。

Table 3: 従来モデルのルール長とルール数、パラメータ数

ルール長	ルール数	パラメータ数	
		$p=1$	$p=2$
5	243	2673	3159

Table 4: 提案法のルール長とルール数、パラメータ数の平均

ルール長	ルール数	パラメータ数	
		$p=1$	$p=2$
1.46	3.5	13.9	20.9

以上の結果を踏まえると、ルール数やパラメータ数が少なくなり、ラフ集合を用いることで非常に簡潔なルールを抽出できていることがわかり、これにより大幅に計算時間を短縮することができた。 $p=2$ とすることでパラメータ数が増加しており、パラメータ数が増加したことにより精度が高まることわかる。

5 おわりに

本研究ではまず等価性を用いることにより後件部の重みはファジィ集合の面積と解釈できることを述べた。このことから不精密ルールを用いたとしてもルールとして直感的に理解しやすい。

また、ラフ集合による知識獲得を行うことにより、矛盾を含む医療データからも精度の良いルールが獲得でき、不精密ルールも獲得できることが示された。さらに提案手法を医療診断へ応用することによりファジィ判別分析を行い、提案手法では極小決定ルールを用いていることから従来モデルよりも規則数が少なく、ルール長が短いにもかかわらず従来モデルよりもはるかに良好な結果を得ることができた。

これらのことから、提案手法はビッグデータや矛盾を有したデータでも応用が可能であることが期待される。

今回使用した糖尿病データでは、後件部が3クラスであったため合併クラス数 $p=2$ しか考慮しなかったが、他のデータを扱う際には、後件部のクラス数によっては $p=2$ 以外も考慮することが考えられる。今後の課題としては適切な p の値、メンバシップ関数の分割数やパラメータの初期値を設定することなどが考えられる。

謝辞 本研究の一部は、日本学術振興会科学研究費補助金若手研究(B) (課題番号: 15K16065) の補助による。

参考文献

- 1) E. H. Mamdani: Application of fuzzy algorithms for control of simple dynamic plant, Proc. IEE, **121** 1585/1588 (1974).
- 2) Z.Pawlak: Rough sets, International Journal of Information Computer Science, **11-5**, 341/356 (1982).
- 3) 水口, 水本: 被覆度を考慮したラフ集合による知識獲得及びファジィ判別分析, バイオメディカル・ファジィ・システム学会大会講演論文集, 35/38 (2005)
- 4) 浜川: 不精密ルールの有用性に関する研究, 大阪大学大学院基礎工学研究科 2015 年度修士論文 (2016)
- 5) 山下, 関, 乾口: 不精密ルールを用いたファジィ推論による医療診断システムの構築第 61 回システム制御情報学会研究発表講演会論文集, 223-3, 2pages (2017)

- 6) T. Takagi and M. Sugeno: Fuzzy identification of systems and its applications to modeling and control, *IEEE Trans. Syst., Man, Cybern.*, **15-1**, 116/132 (1985)
- 7) M. Mizumoto: Fuzzy controls under various fuzzy reasoning methods, *Information Sciences*, **45**, 129/151 (1988).
- 8) M. Mizumoto: Fuzzy controls under product-sum-gravity method and new fuzzy control methods, *Fuzzy Control Systems* (ed. A. Kandel and G. Langholz), CRC Press, 275/294 (1993).
- 9) B. Kosko: *Neural Networks and Fuzzy Systems*, Prentice Hall (1992).
- 10) B. -G. Hu, G. K. I. Mann, and R. G. Gosine: A systematic study of fuzzy PID controllers—function-based evaluation approach, *IEEE Transactions on Fuzzy Systems*, **9-5**, 699/712 (2001).
- 11) 市橋, 渡辺: 簡略ファジィ推論を用いたファジィモデルによる学習型制御; *日本ファジィ学会誌*, **2-3**, 429/437 (1990)
- 12) 森, 田中, 井上: ラフ集合と感性, *海文堂*, 163/181 (2004).
- 13) H. Ichihashi: Iterative fuzzy modeling and a hierarchical network, *Proc. 4th IFSA Congress of Engineering*, 49/52 (1991).
- 14) H. Seki and M. Mizumoto: On the equivalence of fuzzy inference methods—part 1: basic concept and definition, *IEEE Trans. Fuzzy Systems*, **19-6**, 1097/1106 (2011).
- 15) D.F. Andrews and A.M. Herzberg: *Data: A Collection of Problems from Many Fields for the Students and Research Worker*; Springer (1985)

重み付き投票法による離散化と Artificial Bee Colony を用いた 設計最適化法

横谷元 ○ 畠中利治 (大阪大学)

Design Optimization based on Artificial Bee Colony and Weighted Voting Method

G. Yokoya and *T. Hatanaka (Osaka University)

Abstract— In this paper, we discuss a design optimization scheme by using Artificial Bee Colony and a kind of discretizing method based on probabilistic voting for an alternative candidate. Artificial Bee Colony is used for a global continuous search that corresponds to a physical perspective of target design problem and a discretizing part is a choice of parts where we face to an actual design phase. In addition, it works as a local search in the optimization procedure. By applying a benchmark problem, an applicability of the proposed approach is shown in numerically.

Key Words: Design Optimization, Black-Box Function Optimization, Artificial Bee Colony, Local Search

1 はじめに

マーケットのグローバル化や環境意識の変化は消費者のニーズの多様化をもたらしている。特に、自動車に代表されるモノづくりでは、多様なニーズに対応するため、従来の大量生産から多品種少量生産へのシフトが進んでおり、多品種の設計を効率化することを目的としたさまざまな設計支援技術の必要性が高まっている。特に、設計、試作とテストからなる工業製品の設計サイクルを支援する計算機技術は重要であり、HPCI戦略プログラムにおいても分野4として次世代ものづくり¹⁾が取り上げられ、研究と技術開発が進められてきた。この次世代ものづくりの一分野として、単純に試作と試験に対応する部分をコンピュータシミュレーションに置き換えていくことだけではなく、設計における意思決定支援の技術として、多目的設計探査²⁾の研究がなされている。多目的設計探査の代表的なアプローチは、多目的最適化を主に進化計算によって実行し、多様な解候補を得るとともに、その解候補からの知識獲得からなる設計支援の枠組みを提供するものである。

この立場から、小平らは複数車種の同時設計として、車体を構成する車種あたり74部品の板厚を3車種にわたって同時に設計する問題を設定し、車両の基本特性となる剛性、低周波特性、衝突安全性と設計上の要請に起因する変数間の大小関係とからなる車種あたり18個、計54個の制約条件付きの3車種の合計重量の最小化と車種間の共通部品数の最大化の2目的最適化問題を定式化している。また、この問題に対して、実際に車両構造の評価を有限要素法を用いて行い、制約条件を満たすパレート解を多目的進化計算によって得ることによる設計支援技術の開発が行われている^{3,4)}。さらに、進化計算は多目的設計探査における有力な解探索ツールであると考えられているが、従来のベンチマークには、このような実設計問題に見られる特性が備わっていないという問題点⁵⁾があり、設計支援に資する手法を確立する目的で、応答局面法によって3車種の車両設計問題をモデル化したベンチマークが公開されている⁶⁾。

本研究では、このベンチマーク問題を題材とし、設計支援のための多数制約条件下の最適化問題において、制約を充足しつつ、有望な解候補を得るための手法を検討

している。この問題は、車両を構成する部材の板厚に代表されるように、設計変数は本質的には連続変量であり、目的関数や制約条件の評価においても物理量の評価がなされる一方で、設計問題としては部品選択の形で離散的に解候補が存在するクラスの問題であり、我々は、連続変数の最適化と部品の選択としての離散化を組み合わせたアルゴリズムを検討している。

提案法では、具体的には、連続最適化に加藤らが提案した収束性能を強化した算術交叉人工蜂コロニー (Arithmetic Crossover Artificial Bee Colony, 以下、AC-ABC) 法を制約付き関数最適化問題向けに改良した方法を用い、さらにAC-ABCが与える有望解から実際の部材の候補を絞り込む探索すなわち、多数の候補の組み合わせから制約を満たしつつより3車種の合計重量が小さくなる点を選択する探索に、候補の擁立と投票からなる進化的な投票法を用いている。以下では、2章でベンチマーク問題の概要を述べ、3章で、提案法の概要とAC-ABCと投票法のそれぞれについて述べる。また、4章で実行条件と計算結果の一例を紹介し、最後の今後の課題についてまとめる。

2 ベンチマーク問題

本稿で取り上げるベンチマーク問題⁶⁾は、実設計問題に資する最適化法を開発する目的で提供されているものである。具体的には、車両を構成する部材の板厚に対して、複数の車種に対する車体の剛性や衝突安全性を有限要素法を用いたシミュレーションを実際に行う代わりに、応答局面法を用いてモデル化した目的関数と多数の制約条件により構成され、1車種あたり18個の制約(3車種合計で54個の制約)と車両重量の最小化および、3車種の共通部品数の最小化の2目的最適化問題として定式化されている。なお、この3車種は、車格が異なる3つの車種に対応しており、設計変数は車種ごとに74個ある。各々の設計変数には上下限と選択可能な板厚の離散値が指定されている。以下、本文での説明では、車種のインデックスを $i = 1, 2, 3$ で与え、本文および図表での表記は順にSUV, CDW, C5Hとする。また、設計変数、制約条件、目的関数をそれぞれ次の記号で表記する。車種ごとの設計変数は $D = 74$ 次元のベクトルであり、 $\mathbf{x}_i = (x_{i,1}, x_{i,2}, \dots, x_{i,D})^T$ と書く。 T

はベクトルの転地を表す。制約条件は、車種ごとに関数 $g_{i,k}(\mathbf{x}_i)$, $\{i = 1, 2, 3, k = 1, 2, \dots, P\}$, $P = 18$ に対する不等式条件として与えられる。具体的な制約条件は文献⁶⁾に示されているが、車種ごとに14個の制約条件は、車体の衝突安全性や剛性に関わるものであり、4つは設計上の要請による変数の大小関係である。

ベンチマーク問題は、3車種の総重量の単目的問題と、車両重量の最小化をできるだけ多くの共通部品を用いて実現する2目的最適化問題に対応しているが、本研究では、主たる目的を車両重量の最小化ととらえ、共通部品数が考慮できる探索法を与えることを目的とした。

すなわち、目的関数は、 $mass_i(\mathbf{x}_i)$ を i 番目の車種の車重とすると、 $\sum_{i=1}^3 mass_i(\mathbf{x}_i)$ で与えられる。なお、 \mathbf{x} は、 $x_{i,j}$, $\{i = 1, 2, 3, j = 1, 2, \dots, 74\}$ を要素とする3車種にわたる説明変数の222次元ベクトルである。

ここで、 $x_{i,j}$ は部材の板厚に対応する変数であり、本質的には連続値をとる変数である。単純な意味での最適化を図るのであれば、連続値として扱うことも可能であるが、いわゆる設計問題を考えるときや、部品の共通化を意図する場合には、選択可能な部材に限定した離散変数とみなすことが自然であり、本研究で提案する方法では、連続値による探索によって絞り込んだ近傍での離散化を、改めて最適化問題ととらえ、部品候補を提示しうる枠組みを構成する。なお、このベンチマーク問題では、変数値の候補は一定の規則で離散化されており、基本的な板厚の離散幅は、1.0未満の場合は0.05刻み幅で、1.0以上の場合は0.2刻み幅で与えられている。例えば、設計空間の下限と上限がそれぞれ0.8と1.4の場合、候補となる板厚は、0.8, 0.85, 0.9, 0.95, 1.0, 1.2, 1.4となる。部品の共通化とは、同じ次元の要素に対して、3つの車種で同じ板厚の部材を用いることに相当する。この離散化は、ある設計空間上の1点 \mathbf{x} をその近傍に存在する多数の候補点のいずれかへの置き換えに相当する。この操作は、ある車種のある要素 $x_{i,j}$ に対し、通常はその値を挟む2つもしくはそれ以上の個数の候補点のいずれかと置き換える操作となるが、仮に上下2点を候補としても、車種あたりで少なくとも 2^{74} 通りの組み合わせが存在し、全組合せを調べることは現実的ではなく、有効な探索法が要求されることになる。

3 提案法の概要

多数の制約条件が存在する関数最適化問題であることから、次のように車種ごとの制約違反の個数 $p_i(\mathbf{x}_i)$ をペナルティとした適合度関数に対する進化計算による関数最適化を行うことを考える。

$$f(\mathbf{x}) = \sum_{i=1}^3 mass_i(\mathbf{x}_i) + \beta \sum_{i=1}^3 p_i(\mathbf{x}_i) \quad (1)$$

この問題に対して、提案法は、一定の評価回数を AC-ABC による探索と、得られた実数値の優良解の候補から、その近傍の離散化された解候補への置き換えのための探索の組合せへ配分する方法を適用するものである。

制約条件を充足する解を得るために、制約違反数および違反量の取り扱いは重要であるが、本研究ではその議論は行わず、車両重量と制約違反数のバランスをとるパラメータ β による重み付き和の形とし、 β を車両重量

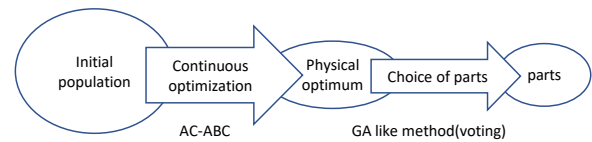


Fig. 1: 提案法の概要

の減少に応じて、段階的に減少させ、探索過程においては制約を満たさない個体の情報も利用可能になるようにした。

3.1 算術交叉 ABC による探索

Artificial Bee Colony (ABC)⁷⁾ は、Karaboga らが提案した群知能の一種である。探索点は、蜜蜂の3つの振る舞いに対応した3つのフェーズ；employed bee, onlooker bee, scout bee の行動を行うことで解探索を行う。それぞれ、次のように説明されている。

Employed bee：働き蜂のフェーズ：近傍個体との情報共有により次の探索点を決定して、新しい解候補が優れている場合には移動し、そうでない場合は元の位置にとどまる。

Onlooker bee：見物蜂のフェーズ：ルーレット選択によって1個体を選択し、Employed bee と同様の探索を行う。

Scout bee：偵察蜂のフェーズ：あらかじめ定められたステップ数で解更新がない（移動しない）個体は、Scout bee フェーズに移り、探索領域内で初期化される。

ABC は、Employed bee の操作がランダムに選択された1つの次元に対するものであることから、集団の多様性が維持されやすく、多峰性関数の最適化に優れているとされるが、一方で高次元の問題における収束の速さが犠牲になっている。本研究では、現実的设计問題を取り扱うことが目標であり、目的関数の評価回数には制限があることを前提としている。そこで、収束性能の改良をはかった加藤らによる交叉を導入した AC-ABC を採用することにした⁸⁾。AC-ABC では、Onlooker bee に対して、ルーレット選択で選ばれた個体と交叉率を満たす次元に対しては、2つの個体の内分点に要素をとることで、探索範囲を拡げている。なお、本研究では、ペナルティ違反量を組み込んだ目的関数の性質を考慮し、探索範囲を外分点へも拡張するようにパラメータをとることにしている。

また、評価値の高い個体による探索に計算資源を多く投入するための Elite bee フェーズとして、算術交叉による探索を行うことにした。Elite bee フェーズは以下のようなフェーズと位置付けることができる。

Elite bee：集団内で高い評価値をもつ個体のみによる探索のフェーズ。エリート保存の観点から確定的に選択され、次世代に情報が引き継がれることになる。Onlooker bee に代えて用いる。さらに、Scout Bee のフェーズでは、Elite 近傍での個体生成も一定の確率で起こるように設定した。

Fig.2 に本研究で用いた AC-ABC の手順を示す。以下では、標準的な AC-ABC からの変更点を中心に、手順を具体的に説明する。

初期個体生成：初期個体は、探索範囲の上限と下限の間で次元ごとに独立な一様乱数を用いて生成する。設計変数間の大小関係を規定する制約条件については、これらの条件を満たすように生成することとした。

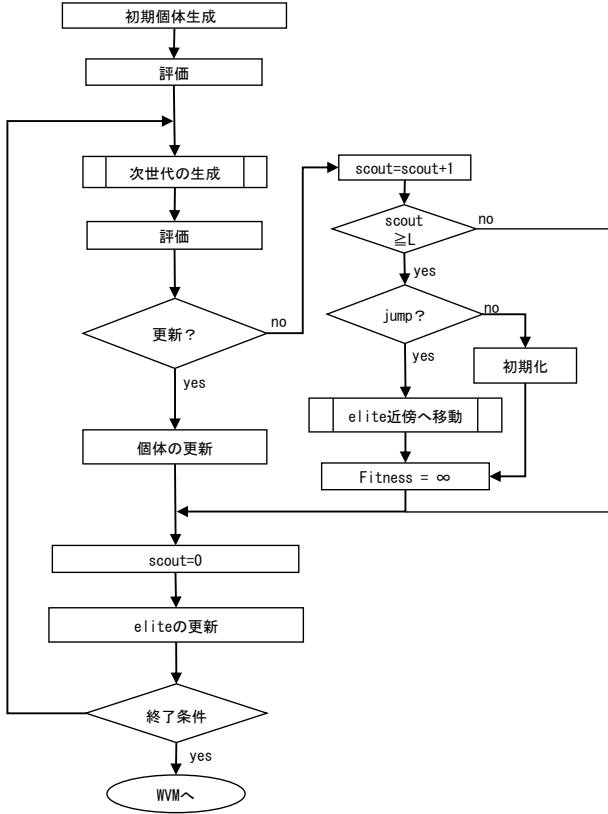


Fig. 2: AC-ABC の概要

次世代の生成：以下の式に従って、 i 番目の個体に対して、ランダムに選んだ個体 j との算術交叉により個体を生成する。

$$\mathbf{x}_i^{t+1} = \lambda \mathbf{x}_i + (1 - \lambda) \mathbf{x}_j \quad (2)$$

ここで、 λ は、区間 $[-0.2, 1.2]$ の一様乱数で与えることとした。なお、Elite Bee は現世代における $f(\mathbf{x})$ が上位の個体であり、以下の操作により次世代の個体を近傍から生成する。

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i + r\delta(\text{upper}_{i,j} - \text{lower}_{i,j}) \quad (3)$$

ここで、 r は、範囲 $[-1, 1]$ の一様乱数であり、 δ は探索範囲の大きさを定めるパラメータである。

個体の更新：目的関数が改善した個体は更新される。改善しない個体は、改善しなかった回数を保存するカウンター ($scout$) に 1 を加算する。一定の回数 (この回数を L とする) 続けて更新されない ($scout > L$) とき、Scout bee フェーズへ移行する。

Scout Bee の更新：通常の初期個体を生成するほか、確率的に Elite bee 個体の近傍で初期化されるようにする。ここでは、以下のように集団内の最良個体の近傍で初期化するようにした。

$$\mathbf{x}_i^{t+1} = \mathbf{x}_i^{best} + r\delta(\text{upper}_{i,j} - \text{lower}_{i,j}) \quad (4)$$

なお、すべての操作で、生成された個体の位置が上限下限を超える場合は、上限値下限値にセットすることとした。また、世代ループあたりの目的関数の評価回数をそろえるため、Scout Bee の評価は Skip し、次世代での個体の更新によらず、位置と評価値が保持されるようにしている。

3.2 重み付き投票法

AC-ABC の結果、連続変数として有望な解候補が与えられる。しかし、実際の決定変数は離散値から選択されるため、得られた解候補 \mathbf{x} をその近傍の離散値に置き換えることが必要になる。もっとも単純な手段は、四捨五入や切り捨てといった演算により離散値に置換することであるが、置換先で制約違反が発生する可能性が高く、制約違反を回避するように置換することは容易ではない。

本研究では、この離散化に対して、次元ごとの置換先の候補の擁立と投票からなる探索法を導入する。ここでは、まず、与えられた離散値の候補から次元ごとに解候補を作成する。各離散値の候補は自身が採択される採択率をもつ。次元ごとに設定され、 p_{ij} を次元 i における j 番目の候補とすると、採択率は次のように書ける。

$$\begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1d} \\ p_{21} & p_{22} & \cdots & p_{2d} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nd} \end{pmatrix} \quad (5)$$

ただし、 $\sum_{i=1}^n p_{ij} = 1, j = 1, 2, \dots, d$ である。

初期の採択率の設定は、たとえば、

$$p_{sj} = (n - s) / \sum_{s=1}^n s \quad (6)$$

によって行う。ここで、 n は生成される離散値の解候補の個数、 s は代表点からの距離の昇順のインデックスであり、 $1, \dots, n$ のいずれかの値をとる。この場合、連続値の解候補に近いほど初期の採択率は高く設定される。次に、各次元の採択確率に従って個体を生成し、目的関数を計算する。評価が高い個体のみが投票に参加でき、投票では評価値の良い個体の各次元の離散値の候補の採択率を

$$p_{s_a,j} = p_{s_a,j} + s_a \gamma \quad (7)$$

$$p_{s_b,j} = p_{s_b,j} - s_b \gamma (n - s_b) / \sum_{s=1, \neq s_b}^n s \quad (8)$$

に従って変更させる。ここで、 s_a, s_b は、それぞれ評価値が上位と下位の解候補のインデックスを表す。

この繰り返しによって、採択率が変化し、目的関数を改善しない (制約違反を増加させない) 解候補が得やすくなると期待できる。選択確率の上限 R_{max} を設けることで、局所解への収束を防ぐことも可能である。

この手続きは、連続値の代表点に対する近傍探索とみなすことができ、提案法の基本的な戦略は、連続最適化による探索と代表点近傍での離散値による探索の局所探索との組み合わせであるといえる。数値例では、探索の前半に ABC による大域的探索、後半に重み付き投票法による局所探索を実行しているが、繰り返し実行することも可能である。

4 数値例

ベンチマークに対して、Table 1 と 2 に示す条件で数値実験を行った。

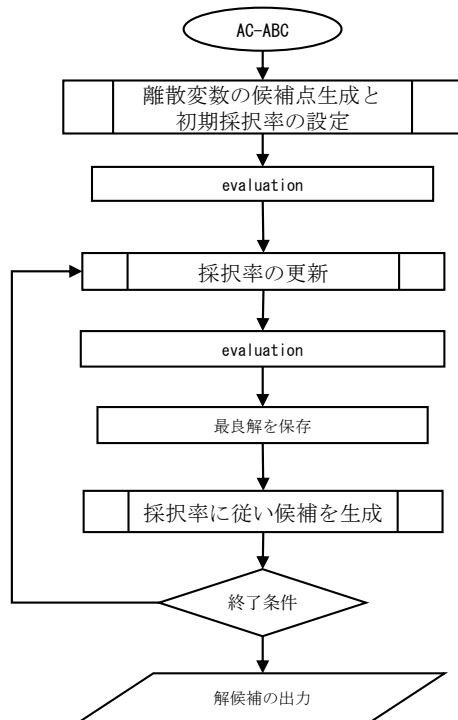


Fig. 3: WVM の概要

Table 1: パラメータ設定 (AC-ABC)

個体数	終了世代	elite 数	L	δ	Jump 確率
300	100	30	5	0.05	0.4

Table 2: パラメータ設定 (投票法)

個体数	終了世代	投票個体	γ	N	R_{max}
60	166	3	0.03	2	0.8

3 車種それぞれの最良個体の重量の世代変化と制約違反量をそれぞれ Fig.3 と Fig.4 に示す. 横軸は連続した評価回数で, 縦軸は各車両の重量および, 制約違反数に比例するペナルティである. 評価回数 30,000 回で AC-ABC から WVM に切り替わる. AC-ABC によって解候補が更新され, 30,000 回に達したのち, 離散化され車両重量の改悪が発生するが, その後, 繰り返しにより一定の改善がなされる. 制約に関しては, CDW で離散化に伴う制約違反が生じているが, その解消がなされていることもわかる.

5 おわりに

本研究では, 多数の設計変数かつ多数の制約条件のもとでの設計最適化のための ABC と解候補の離散化からなる探索法を示した. これは, 想定される問題では, 設計変数が本質的に連続変数であるが選択できる値は離散値であることに対応している. 離散化では, 解候補の近傍に存在する多数の離散値の候補からより良い候補を得るため目的で重み付き投票法を与えた. 評価回数の制限がある場合, ABC による大域的探索と重み付き投票法への計算資源の配分や, 実行時のパラメータ設定は試行錯誤が必要であり, その設定は今後の課題である. また, 他の連続最適化法との組み合わせも可能であり, それらの調査も今後の課題である.

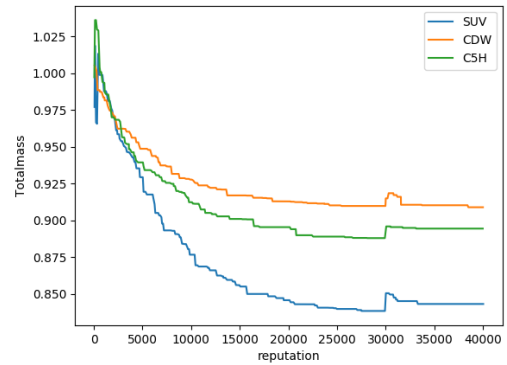


Fig. 4: 3 車種の重量の世代推移

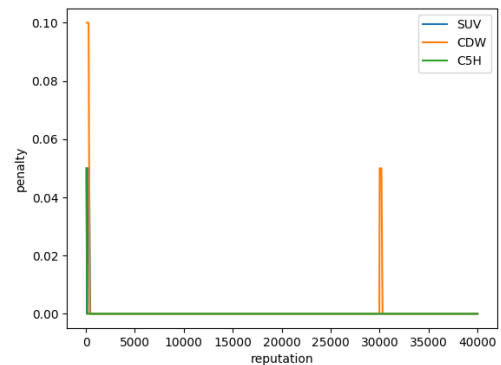


Fig. 5: 制約違反量の変化

謝辞

本研究の一部は, JSPS 科研費 JP15K00338 の助成を受けた. ここに謝意を表す.

参考文献

- 1) 加藤千幸, 京と次世代ものづくり, 第 25 回計算力学講演会講演論文集, pp. F-49-52, 2012.
- 2) 大山聖, 立川智章, 渡辺毅, スーパーコンピュータ「京」を用いた大規模多目的設計探索, 最適化シンポジウム講演論文集, pp. 1-5, 2014.
- 3) 大山聖, 小平剛央, 他, スーパーコンピュータ「京」を用いた複数車種の車両構造同時設計最適化, 第 28 回計算力学講演会講演論文集 (CD-ROM), 2015.
- 4) 小平剛央, 大山聖, 他, 大規模並列計算を用いた複数車体構造の同時設計最適化, 第 29 回計算力学講演会講演論文集 (CD-ROM), 2016.
- 5) 田邊遼司, 大山聖, 制約付き多目的最適化ベンチマーク問題の問題点, 進化計算シンポジウム 2016 講演論文集, pp. 340-347, 2016.
- 6) 小平剛央, 釘持寛正, 大山聖, 立川智章, 応答曲面法を用いた複数車種の同時最適化ベンチマーク問題の提案, 進化計算学会論文誌, Vol. 8, No. 1, pp.11-21, 2017.
- 7) D. Karaboga and B. Basturk, "A powerful and efficient algorithm for numerical function optimization: Artificial bee colony (ABC) algorithm," Journal of Global Optimization, Vol. 39, pp. 459-471, 2007.
- 8) 加藤達郎, 前田陽一郎, 高橋泰岳, 算術交叉を用いた改良型 Artificial Bee Colony アルゴリズム, 第 28 回ファジィシステムシンポジウム講演論文集, pp. 430-435, 2012.

罰を受ける状況に関する情報の抽象化と強化学習効率化への利用

坂下悠太¹ ○村田純一 (九州大学)

Abstraction of Punished Situations for Efficient Reinforcement Learning

Y. Sakashita¹ and *J. Murata (Kyushu University)

Abstract— A method is proposed that first extracts abstract expressions of situations where reinforcement learning agents receive punishments and then utilizes them for efficient learning. The method focuses on ‘bad’ situations, and it obtains a general description of situations that can be used to predict ‘bad’ situations and thus to avoid those situations, resulting in more efficient learning. A navigation task of an autonomous robot is used as an example problem, and its simulations demonstrate validity of the method.

Key Words: Transfer learning, Constraints, Generalization, Autonomous navigation

1 はじめに

転移学習 (transfer learning)¹⁾ は, あるタスクに関する学習を行う時に, 対象タスクとは別のタスクで既に学習した成果を利用するものである. 最近, 深層学習 (deep learning) への関心が大きく高まるにつれて, 対象タスクに関するデータ不足を補う手段としての転移学習が注目されてきている. 転移学習は深層学習だけでなく, 学習を効率的に行うためにいろいろな学習手法において活用されている. 強化学習においても, 学習で獲得された価値関数²⁾ や方策³⁾ を他のタスクの学習に活用する転移学習⁴⁾ の研究がなされている. これらの研究では, 主として学習によって得られた「良いこと」の転移を行っている. 一方, 「悪いこと」の利用も有効であると考えられる. Takano⁵⁾ らは2つの環境の類似度を各環境で罰を受ける行動の一致率によって定め, 転移学習において転移すべき内容の選択に用いている. 著者らは, さらに積極的に, 罰を受ける状況に関する情報を抽象化し, それを転移させて活用する方法を考案した⁶⁾. 本論文では, 学習中に得た成果を他のタスクに転移させて利用するだけでなく, 学習中のタスクにおいてもリアルタイムに学習の効率化に活用することを考える. この方法の有効性を, カメラ画像に基づいてナビゲーションを行う移動ロボットのシミュレーションによって確認する.

2 考え方

2.1 「悪いこと」に関する仮説

この論文で取り扱う「悪いこと」の活用は以下の仮説に基づいている.

仮説0 「悪いこと」の方が「良いこと」よりも頻繁に体験される. したがって,

仮説1 「悪いこと」の方が「良いこと」よりも一般性が高く利用価値がある.

仮説2 「悪いこと」の方が「良いこと」よりも早く知ることができる.

これらの仮説を証明することはできないが, 次のような例から, ある種のタスクについては妥当であると推測することができる. 自律移動ロボットが, 机や椅子

などが置かれている部屋の中から, 部屋の出入口を通って外へ出るナビゲーションタスクを考えてみよう. ロボットがランダムに移動すると, 机, 椅子や壁に衝突する「悪いこと」の方が, 出入口を通って外へ出る「良いこと」よりも頻繁に起こる (仮説0). 頻繁に起こりうる衝突に関する知識は利用価値がある (仮説1). また, 出入口へ移動する際の「良い」移動方向は, 出入口の位置などに依存し, 部屋が異なれば変わってしまう. 一方, 机や椅子に衝突することは「悪い」ということは, 部屋が異なっても共通であって一般性が高い (仮説1). さらに, 頻繁に体験されるということは, そのことに関する学習が速く進行することを意味する (仮説2).

2.2 強化学習での利用

強化学習では, 「良いこと」とは高い (正の) 報酬を得ることである. 一方, 「悪いこと」の端的な例は罰 (負の報酬) を受けることである. タスク達成時にのみ正の報酬が得られる問題が多いこと, 一方で, 罰はタスク遂行途中でも受けることが多いことを考えると, 上記の仮説は強化学習が扱う問題の多くで成り立つと考えられる.

強化学習は「良い行動」を行う良い方策を学習によって獲得するものであり, 罰を受けるような「悪い行動」を学習中に多く試行するのは時間の浪費につながる. そこで, あるタスクにおいて, 罰を受ける状態と行動の対を抽象化し, それを, 罰を受ける行動の予測とその回避に利用すれば, 対象タスクでの学習を効率的に行い, 学習速度を向上させることができる. さらに, 仮説2が主張しているように, 罰を受ける事態は, タスクが達成されて正の報酬を得る以前に発生することから, 罰に関する抽象化された情報は強化学習が完了する前に獲得することができ, 学習完了後に他のタスクに転移して利用するだけでなく, 当該タスクの強化学習実行中にも利用可能である.

なお, 行動価値の値が低い行動も「悪い」行動である. しかし, 強化学習は, 価値関数を推定する際に, 低い価値の値を精度よく推定することを重視しないため, 推定された価値関数に基づいて価値の低さを判定してもその信頼性は保証されない.

一般に, 抽象化⁷⁾とは情報の本質部分以外の一部を欠落させることである. 強化学習で獲得する方策は, 与えられた状態 s に応じた適切な行動 a を規定する写像

¹現在三菱電機勤務. Currently with Mitsubishi Electric.

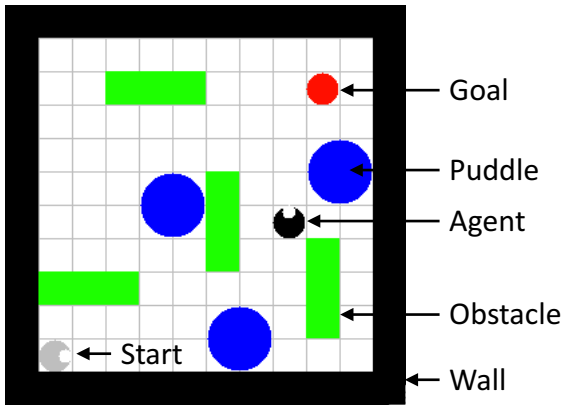


Fig. 1: An example environment.

$\pi: s \mapsto a$ である。方策中の特定のルール $s_0 \rightarrow a_0$ において状態の表現 s_0 を抽象化すると、このルールが当てはまる状態が増える、つまり方策の一般化につながり、他のタスクへの転移が容易になる。さらに、抽象化によって、方策が表しているルールを人間が解釈できるようになる可能性もある。

2.3 提案内容の実現と検証に用いた例題タスク

「悪いこと」が起こる状況をどのように抽象化すれば良いかは、対象とするタスク、特に学習エージェントが知覚する状態に依存する。そこで、例題タスクを用いて提案内容の実現と検証を行う。

使用した例題タスクは以下の通りである。Fig. 1 に示すような壁（黒）に囲まれた環境に、障害物（緑）や水たまり（青）がある。学習エージェントは Start から Goal（赤）まで移動する最適経路を学習する。障害物や水たまりにエージェントが進むことはできない。環境はマス目に区切られており、エージェントは1マスを占有する。エージェントは前方をカメラで撮影し、その画像（Fig. 2）を環境の状態の観測値として利用する。1画像は 32×16 ピクセルの RGB 値で表される。RGB 値は $[0, 1]$ に正規化し、1画像を $32 \times 16 \times 3$ 個の成分を持つベクトル s で表現する。エージェントは1タイムステップに、1マス前進、その場で90度左旋回、180度旋回、90度右旋回の4つのうちのいずれかの行動を行う。行動はベクトル a で表される。これは、前進、左旋回、180度旋回、右旋回に対応した4個の成分を持つベクトルであり、実行した行動に対応する成分を1、その他の成分を0として表現する。エージェントが、壁、障害物あるいは水たまりが存在するマス目に進もうとした場合は進むことができず、罰（負の報酬）を受けた上でしばらく停止する。ゴールに到達したときには正の報酬を受ける。

3 罰を受ける状況の抽象化とその利用

3.1 例題における罰を受ける状況の抽象化

この論文で扱う例題においては、罰を受けた複数の状態-行動対（以下では「体験」と呼ぶ）の間の差異に関する情報を欠落させる抽象化を行う。これによって、どのような状態の時にどのような行動を行えば罰を受けるのかについての、一般性のある知識を獲得する。

例題では状態はカメラ画像の各ピクセルの RGB 値を成分とするベクトルで表現される。状態 s と行動 a を並べたベクトルを「体験」 $x = [s^T a^T]^T$ と定義する。

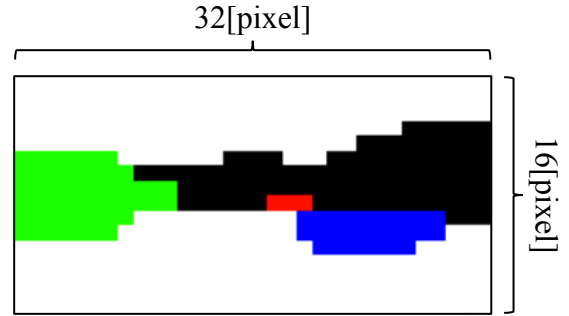


Fig. 2: An example image that the agent captures.

また、ある体験をした際に罰を受けたかどうかを y で表す。 $y = 1$ は罰を受けたことを、 $y = 0$ は受けなかったことを表す。抽象化は、体験 x を $y = 1$ の罰を受けるクラスと $y = 0$ の罰を受けないクラスに分類する分類器を構成し、 $y = 1$ のクラスに対応する複数の体験 x の間の差異を欠落させることによって行う。このための分類器として Learning Vector Quantization (LVQ) を用いる。LVQ は、クラス内の複数のサンプルベクトルを、ユークリッド距離が近い重みベクトルで代表する。したがって、同じ重みベクトルによって代表される複数のベクトル間の差異が欠落することになる。また、代表となる重みベクトルを見れば、抽象化された情報を人間が解釈することができる。ここで、同じクラス内のベクトルがすべて単一の重みベクトルで代表されるとは限らない。適切な本数の重みベクトルで代表する必要がある。

以上を踏まえて、以下の手順によって LVQ を用いた抽象化を行う⁶⁾。

準備

- 強化学習の過程で収集した体験と罰の有無の組のデータを $D = \{(x_i, y_i)\}_{i=1}^N = D_p \cup D_u$ とする。ただし、 D_p は罰を受けた（すなわち $y = 1$ である）データの集合、 D_u は罰を受けなかった（ $y = 0$ である）データの集合であり、それぞれ N_p 個、 N_u 個のデータを含むものとする。当然 $N_p + N_u = N$ である。
- 罰を受けたクラスに対応する LVQ 重みベクトルの個数を n_p 、罰を受けなかったクラスに対応する重みベクトルの個数を n_u で表すことにし、いずれも初期値を1とする。

手順

1. D 中の各クラスに属する体験 x のデータから、それぞれ n_p 個、 n_u 個をランダムに選んで、LVQ の重みベクトルとする。
2. データ D を $(n_p + n_u)K$ 回繰り返し用いて LVQ の重みベクトルの学習を行う。 K は定数である。 D 内のデータの提示順序はランダムに変更する。
3. 学習済の LVQ を用いて D 内の経験データ x のクラス分類を行い、誤り率 r_p 、 r_u を次式を用いて計算する。

$$r_p = (D_p \text{ 中で罰無しと誤分類されたデータ数}) / N_p$$

$$r_u = (D_u \text{ 中で罰ありと誤分類されたデータ数}) / N_u$$

4. $r_p < \theta$ かつ $r_u < \theta$ であれば終了. そうでなければ, 誤り率が θ 以上であったクラスの重みベクトル数 (n_p または n_u , あるいは両方) を 1 増加させて, 手順 1. に戻る. ここで θ は適切な定数の誤差率許容上限である.

以上の操作により, 罰の有無を誤り率 θ 未満で正しく分類することができる, 適切な個数の重みベクトルを持つ LVQ を得ることができる. LVQ を用いているため, ここで行われている抽象化は, ある重みベクトルに注目したとき, ユークリッド距離の意味でこの重みベクトルに近い複数の異なる体験値ベクトル \mathbf{x} の間の差異を無視する抽象化である. あるいは, 体験 \mathbf{x} の空間をボロノイ分割する抽象化であるとも言える. また, 得られた重みベクトルは, 相互の差異が無視できる複数の体験値ベクトルの代表であり, そのうちの状態部分 \mathbf{s} を画像に復元することによって, どのような状況で罰を受けるのかを人間が視覚的に解釈することもできる.

なお, LVQ の学習については適切な文献⁸⁾を参照されたい.

3.2 例題における罰を受ける状況抽象化の利用

強化学習において罰を受ける行動は良くない行動であり, その行動を試行するのは学習を行う上で無駄である. そこで, 罰を受けるかどうかの予測を上記で得た LVQ を用いて行う. 強化学習の実施中, ある状態 \mathbf{s} において ε -greedy や softmax などの適切な行動選択法によって行動 \mathbf{a} を選択したとする. このときの $[\mathbf{s}^T \mathbf{a}^T]^T$ を LVQ に入力し, 罰を受けるかどうかを予測する. 罰を受けると予測された場合は, 確率 $1 - \rho$ でこの行動 \mathbf{a} を破棄し, 再度行動を選びなおす.

LVQ による罰の有無の予測は完璧ではないため, 上記では, 罰ありと予測された行動も確率 ρ で許容する. これは, 通常とは逆の意味の exploration である.

4 例題を用いた検証

提案手法が強化学習を効率化することができることを, 2.3 節に示した例題を用いて検証する. 検証は, 転移学習に利用した場合と, 同一の学習タスクでリアルタイムに利用した場合の 2 ケースについて行う. 両者に共通する報酬設定を Table 1 に示す. 表中の loss of time は, 壁など進行不可能な場所に進もうとした時に, 停止し次に移動可能となるまでのタイムステップ数である. これは衝突によるダメージの回復に要する時間などに相当する. 強化学習には Q 学習を用いる. 行動選択には ε -greedy を用い, ε の値はエピソード数に応じて直線的に減少する. 1 エピソードはゴール到達時または 400 タイムステップ経過時に終了する. また, 両ケースで共通して用いる Q 学習および LVQ 学習のパラメータ設定を Table 2 に示す. LVQ の誤差率上限値 θ を 0.05 に設定したことに対応して, LVQ が罰を受けると予測した行動を許容する確率 ρ も同じ 0.05 に設定している.

4.1 転移学習への利用

Figure 1 に示した環境でのナビゲーションタスクについて Q 学習を 400 エピソード行い, 得られたデータを用いて LVQ の学習を行った. その後, 得られた LVQ を Fig. 3 に示す別の環境での Q 学習に利用する転移学

Table 1: Rewards and loss of time.

Situation	Reward	Loss of time
Arriving at the goal	1.0	0
Stepping in a puddle	-0.2	20
Bumping into a wall	-0.05	5
Bumping into an obstacle	-0.05	5

Table 2: Common parameters in the experiments.

Learning	Parameter	Value
Q learning	Discount factor	0.9
	Learning rate	0.5
LVQ learning	Learning rate	0.1
	K	1000
	θ	0.05
	ρ	0.05

習を行った. ターゲットとする Fig. 3 の環境では, 障害物や水たまりの配置が Fig. 1 の環境と異なっている. また, 比較のために, Fig. 3 の環境において LVQ を利用しない通常の Q 学習も行った.

Figure 4 に, Fig. 3 の環境での Q 学習中に, Start から Goal に到達するまでのタイムステップ数が, エピソードの進行に伴ってどのように変化するかを図示する. 転移あり (図中 RL using LVQ と表示), 転移なし (図中 Normal RL と表示) のいずれの場合についても 100 回の実験を行い, 図にはその平均値を示している. Goal までの到達に要するタイムステップ数が速く減少する方が学習は速い. この図から, LVQ によって獲得した罰を受ける状況に関する情報を用いることにより, 別環境での強化学習の速度が向上することが確認できる. なお, このタスクでは, エージェントが障害物などに衝突しそうな場合, loss of time として指定された長さの時間, エージェントは停止する. このため, 罰を受ける行動の「悪さ」が強調されている. しかしながら, この loss of time をゼロとしても, Fig. 5 に示すように, 転移学習は学習効率化に効果がある.

4.2 学習中の同時利用

2.1 節の仮説に述べたように, 「悪いこと」(ここで扱っている例題では罰を受けること) は, 「良いこと」(ゴールに到達すること) よりも, 早く頻繁に経験される. そのため, 「悪いこと」についての抽象化を行う LVQ の学習を強化学習と同時に実施することができ, しかも, その学習は強化学習よりも速く進行する. そこで, LVQ の学習と Q 学習と同時に実施し, かつ, LVQ によって獲得される罰を受ける状況の情報をリアルタイムに Q 学習に活用することができる. これを, Fig. 1 の環境について行った結果を, Fig. 6 と Fig. 7 に示す. Fig. 6 は loss of time が存在する場合, Fig. 7 はそれが存在しない場合の結果である.

図から, Q 学習と同時に LVQ の学習を行ってそれをリアルタイムに Q 学習に利用すると, Q 学習の速度を向上させることができることが確認できる. ただし, loss of time を考慮しない場合の結果を示した Fig. 7 では, 20 エピソード付近で一時的に, LVQ を利用した場合の方が学習が遅くなっている. これは, Q 学習の初期段階では, LVQ による罰を受ける状況の学習を実施するのに十分なデータが得られておらず, LVQ が正しくない結果を与えたためであると考えられる.

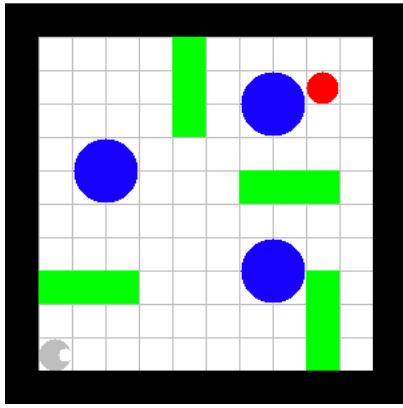


Fig. 3: The target environment of transfer learning.

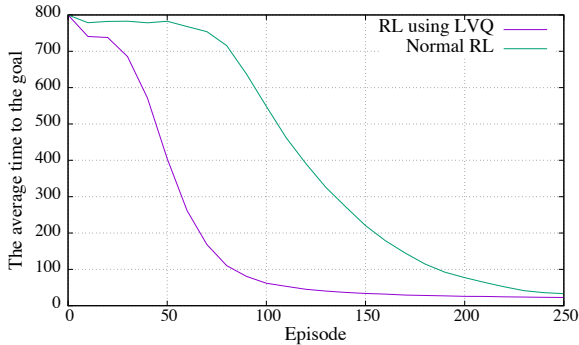


Fig. 4: Time to the goal vs episode, averaged over 100 runs, with transfer learning (violet curve) and without it (green curve). Loss of time is present.

4.3 LVQによる抽象化の結果

学習が終了したLVQの重みベクトルは、罰を受ける状況を抽象化した代表値を表す。これらをFig. 8に示す。図中、画像は重みベクトルの状態 s の部分を表し、画像の下のカッコ内の数値は行動 a の部分を表す。上部5枚の画像は罰を受ける状況の重みベクトル、下の1枚は罰を受けない状況の重みベクトルである。この環境では、エージェントが遭遇する可能性がある体験(状態と行動の組)は300通りある。LVQはこの300通りを6通りの代表的な状態に抽象化して表現している。

罰を受ける状況の代表画像には、壁(黒)、障害物(緑)、水たまり(青)の色が含まれている。また、罰を受けるのは行動を表すベクトル a の第1成分が1に近い、すなわち、前進行動を取るときであることが示

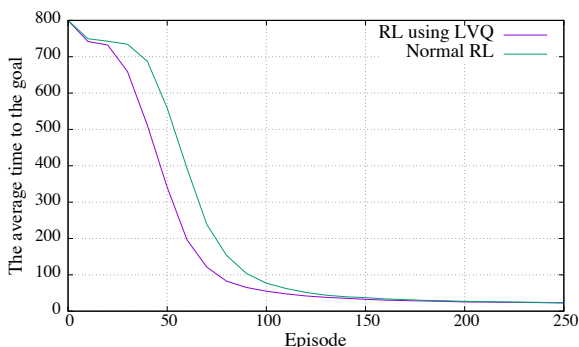


Fig. 5: Time to the goal vs episode, averaged over 100 runs, with transfer learning (violet curve) and without it (green curve). Loss of time is not present.

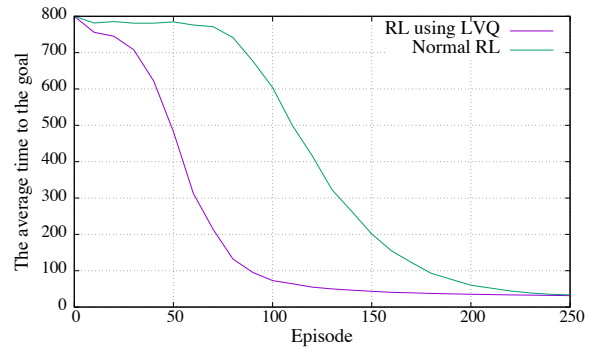


Fig. 6: Time to the goal vs episode, averaged over 100 runs, with real time use of LVQ (violet curve) and without it (green curve). Loss of time is present.

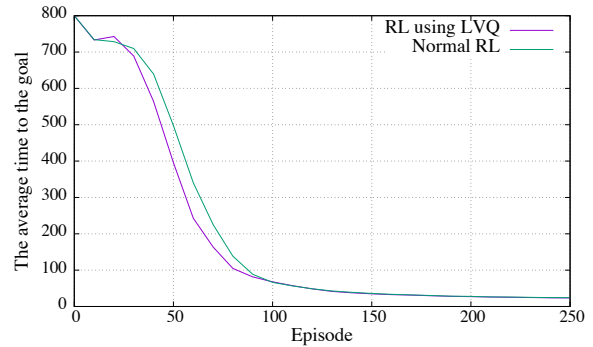


Fig. 7: Time to the goal vs episode, averaged over 100 runs, with real time use of LVQ (violet curve) and without it (green curve). Loss of time is not present.

されている。これに対し、罰を受けない状況の代表画像は、多様な画像の平均値に近い、形も色も特徴のないものになっており、行動ベクトルも4種類の行動に対応する各成分の値の間に特に差異はない。

LVQの学習を行ったFig. 1の環境で、罰を受けない状況であるにも関わらず誤って罰を受けるとLVQが誤判定する状況の個数は8であった。同じLVQが、Fig. 3に示す新しい未学習の環境で同様の誤判定をする状況は11個であった。このことから、LVQは両環境にほぼ共通する一般的な知識を獲得していると判断できる。

Figure 8に示した抽象化の結果は、前の節で述べたように強化学習の高速化に有効である。しかし、人間が解釈するには抽象化の程度がやや高すぎる。そこで、LVQの学習の際の許容誤り率 θ を0.05から0.02に厳しくして学習を行った。より低い誤り率を満たすために、原画像から欠落させる情報は少なくしなければならず、抽象化の程度は低くなる。このようにして得たLVQの重みベクトルをFig. 9に示す。抽象化の程度が低くなったため、300通りの体験ベクトルは18個の重みベクトルで代表されている。これらの抽象度の低い状態ベクトル画像と行動ベクトルの値から、壁、障害物、水たまりが前方にある場合に前進すると罰を受けることを容易に読み取ることができる。

5 おわりに

強化学習の枠組みの中で、「悪いこと」である罰を受ける状況に関する情報を強化学習実行中に取得してそれを抽象化し、これを強化学習の効率化に利用する方法を提案した。特に、あるタスクで抽象化を行ってそ

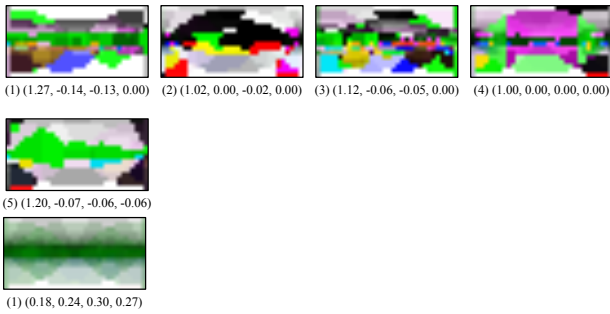


Fig. 8: Images represented by LVQ weight vectors ($\theta = 0.05$). The images represent the state (\mathbf{s}) part while the four numbers in the parentheses represent the action (\mathbf{a}) part. The top five corresponds to punished situations while the bottom to unpunished situations.

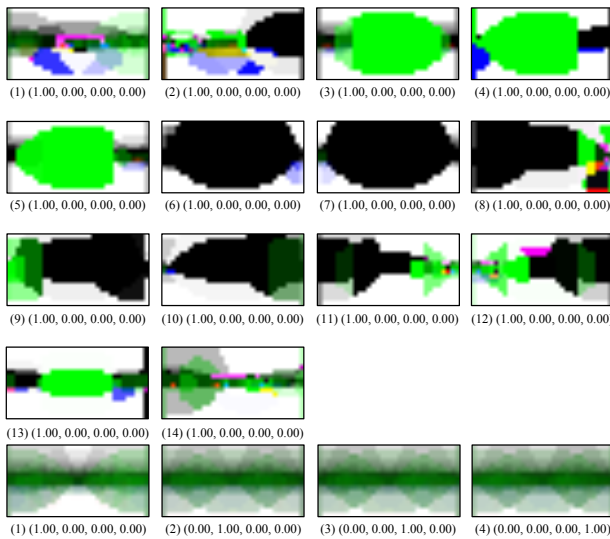


Fig. 9: Images represented by LVQ weight vectors ($\theta = 0.02$). The images represent the state (\mathbf{s}) part while the four numbers in the parentheses represent the action (\mathbf{a}) part. The top fourteen corresponds to punished situations while the bottom four to unpunished situations.

れを別のタスクで利用する転移学習としての活用だけでなく、強化学習と抽象化を同時進行で行い、抽象化の成果をリアルタイムで活用することも提案した。障害物が存在する環境でゴールを目指す自律移動ロボットのナビゲーションシミュレーションを例題タスクとし、Learning Vector Quantization を抽象化の実現手段として用いて、提案内容の実現と検証を行った。得られた結果から、転移学習とリアルタイム活用のいずれにおいても、罰を受ける状況に関する抽象化された情報は、強化学習の速度向上に有効であること、さらに抽象化された結果は人間による解釈も可能であること、を確認した。

本論文では「悪いこと」に着目した。「悪いこと」はやってはいけないことであり、有害な行動を避けるための制約を課す根拠となる。制約は自由を束縛する手枷足かせという不都合なものであるとは限らず、高い自由度をもつ候補の中から適切なものを見出す際のガイドにもなる。ニューラルネットワークの学習や回帰分析に用いられる誤差関数の正則化項はその一例であ

る。また、最近大きな注目を集めている深層ニューラルネットワークは極めて高い自由度を持つが、その自由度を畳み込み層などによって制約していることが、深層ニューラルネットワークがうまく機能している理由であると考えられる。しかしながら、最終的な学習成果を決定するのはデータである。制約はそれをガイドするために用いられる。そのため、「これをすべき」という自由度を大きく制約するタイプの制約よりも、「少なくともこれは避けるべき」という自由度を残すタイプの制約の方が望ましい。ここに「悪いこと」に着目することの利点が存在する。

上記の正則化項や畳み込み層は人間が考案して組入れている。すなわち利用価値の高い制約の源は人間である。これに対し、本研究では強化学習を高速化する利用価値の高い制約を学習によって見出すことを行った。例題として用いたタスクでは、見出した制約は人間には自明に近いものであって、それ自身は大きく特筆すべきものではないかもしれない。また、「エージェントは壁を通り抜けられない」などの、環境とエージェントの相互関係のうちエージェントに害になるものを表すモデルであると見ることもできる。しかしながら、ここで行ったような、自由度の適切な低減をガイドする有用な制約を自動発見することは、コンピュータショナル・インテリジェンスの一つの方向を示していると考えられる。

参考文献

- 1) S.J.Pan and Q.Yang: A Survey on Transfer Learning, IEEE Trans. Knowledge and Data Engineering, **22**, 1345/1359 (2010).
- 2) M.Bowling and M.Veloso: Reusing learned policies between similar problems. Proc. AI-98 Workshop on New Trends in Robotics (1998).
- 3) F.Fernández and M.Veloso: Probabilistic policy reuse in a reinforcement learning agent. Proc. 5th International Joint Conf. on Autonomous Agents and Multi-agent Systems, 720/727 (2006).
- 4) M.E.Taylor and P.Stone: Transfer learning for reinforcement learning domains: A survey. The Journal of Machine Learning Research, **10**, 1633/1685 (2009).
- 5) T.Takano, et al.: Transfer learning based on forbidden rule set in actor-critic method, International Journal of Innovative Computing, Information and Control, **7**, 5, 2907/2917 (2011).
- 6) 坂下, 村田: 罰を受ける状態-行動対の Learning Vector Quantization を用いた抽象化による強化学習の効率化, 計測自動制御学会システム・情報部門学術講演会 SSI2016, 159/162 (2016).
- 7) R.A.Brooks. Intelligence without representation. Artificial Intelligence, **47**, 1-3, 139/159 (1991).
- 8) T. コホネン 著, 徳高平蔵, 岸田悟, 藤村喜久郎 訳: 自己組織化マップ, シュプリンガー・フェアラーク東京 (1996).

報酬分配を用いた Deep Q-Network の実現

中矢 裕太 ○ 長名 優子 (東京工科大学)

Deep Q-Network using Reward Distribution

Y. Nakaya and *Y. Osana (Tokyo University of Technology)

Abstract— In this paper, we propose a Deep Q-Network using reward distribution. Deep Q-Network is based on the convolutional neural network which is a representative method of Deep Learning and the Q Learning which is a representative method of reinforcement learning. In the Deep Q-Network, when the game screen (observation) is given as an input to the convolutional neural network, the action value in Q Learning for each action is output. This method can realize learning that acquires a score equal to or higher than that of a human in plural games. The Q Learning learns using the greatest value in the next action, so a positive reward is propagated. However, since negative rewards can not be of greatest value, they are not propagated in learning. Therefore, by distributing negative rewards in the same way as Profit Sharing, the proposed method learn to not take wrong actions. Computer experiments were carried out, and it was confirmed that the proposed method can learn with almost the same speed and accuracy as the conventional Deep Q-Network. Moreover, by introducing reward distribution, we confirmed that learning can be performed so as not to acquire negative reward in the proposed method.

Key Words: Deep Q-Network, Reward Distribution, Profit Sharing

1 はじめに

近年、画像認識や音声認識の分野で従来手法よりも優れた性能を示すとして Deep Learning¹⁾ が注目されている。Deep Learning は多数の層を持つ階層型ニューラルネットワークであり、畳み込みニューラルネットワーク²⁾ や Deep Belief Network (DBN)³⁾ などがその代表的なモデルである。階層型ニューラルネットワークに関する研究はパーセプトロン⁴⁾ に始まり、誤差逆伝播法⁵⁾ の提案によって注目され、1980年代の半ばから1990年代の前半にかけて研究が非常に盛んに行われた。誤差逆伝播法では、学習データに対する出力と教師信号との誤差を求め、誤差関数の重みに対する勾配に基づいて学習を行う。しかしながら、誤差逆伝播法による学習では、3層程度のネットワークではうまくいくものの、それ以上の多層のネットワークになると勾配が急速に小さくなったり、発散してしまったりする勾配消失問題と呼ばれる現象が発生し、学習が困難であった。そのような中で、2006年に Deep Belief Network³⁾ が提案され、再び、多層構造を持つ階層型ニューラルネットワークが注目されるようになった。Deep Belief Network では、多層のネットワークを制約ボルツマンマシン⁶⁾ と呼ばれる単層のネットワークに分解し、入力層に近い層から順番に学習していくことで多層のネットワークの学習を実現している。また、画像認識の分野や音声認識などの分野でも従来手法よりも大幅に性能が改善するとして、Deep Learning の有効性が広く知られるようになった⁷⁾。従来の画像認識や音声認識では、有効な特徴量を事前に抽出し、それに対して分類を行っていたが、Deep Learning ではデータから有効な特徴量を自動的に抽出することができる。

また、環境との相互作用により適切な政策を獲得するための学習方法として、強化学習に関するさまざまな研究が行われている⁸⁾。強化学習では報酬を適切に設定しておくことで、未知の環境においても試行錯誤を繰り返すことにより学習を進めることができるという特徴がある。強化学習の学習方法は、Profit Sharing⁹⁾

に代表される経験強化型の学習と、Q Learning¹⁰⁾ に代表される環境同定型の学習とに大きく分けられる¹¹⁾。経験強化型の学習は、エージェントが報酬獲得までに得た経験を報酬に基づいて分析し、学習を繰り返していくことから経験強化型と呼ばれる。この手法は学習の速度を重視した手法であり、比較的速く確率的政策を学習させることができるが、決定的政策を学習させるのは難しい。環境同定型の学習法は、獲得した報酬を用いて環境におけるエージェントの状態を評価し、エージェントの行動ごとに報酬獲得により近づけるような良い状態に遷移していくことを目指し、環境を把握していくような学習を行うことから環境同定型と呼ばれる。この手法は決定的政策を学習させることができるが、環境を知るためにより多く試行を繰り返す必要があり、問題によっては決定的政策を得ることが困難な場合がある。

2013年に Volodymyr Mnih らが提案した Deep-Q-Network¹²⁾ と呼ばれる手法は、強化学習の一種の Q Learning を Deep Learning の一手法である畳み込みニューラルネットワークで実現したものである。Deep-Q-Network は様々なゲームにおいて人間と同程度もしくはそれ以上の記録を出し、有効性が確認されている。

本研究では、Profit Sharing における報酬分配を取り入れた Deep-Q-Network を提案する。Q Learning では次の行動の最大の価値をもとに学習を行うので、正の報酬については時間とともに伝播されていく。しかし、負の報酬は最大の価値にはなり得ないので、学習において伝播されない。そこで、提案手法では、負の報酬についてのみ Profit Sharing で行われている報酬分配を行い、間違った行動をとらないことをより学習できるようにしている。

2 畳み込みニューラルネットワーク

畳み込みニューラルネットワーク²⁾ は、畳み込み演算を行う畳み込み層とプーリングを行うプーリング層を含む階層型のニューラルネットワークである。畳み込み層において画像に含まれる特徴を自動的に抽出す

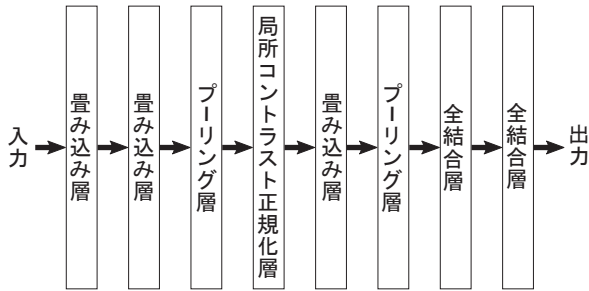


Fig. 1: 畳み込みニューラルネットワークの構造

るように学習が行われる。

2.1 構造

畳み込みニューラルネットワークは、図1に示すような多層構造のネットワークであり、畳み込み層とプーリング層の組み合わせを複数重ねた構造を持つ。複数の畳み込み層の後にプーリング層がくることもある。

畳み込み層は、重みとして表現されるフィルタの濃淡パターンに類似した局所的な特徴を抽出することができる。プーリング層は、畳み込み層で抽出された特徴の位置に対する感度を低下させる働きがある。畳み込み層とプーリング層の組み合わせが繰り返される中で、プーリング層の後ろに局所コントラスト正規化層が挿入されることもある。局所コントラスト正規化層では、前の層の出力のコントラストの正規化を行う。最後に、全結合層を通して、最終的な出力が出力される。

2.2 学習

畳み込みニューラルネットワークでは、確率的勾配降下法を用いて入力に対する出力と教師信号の誤差が小さくなるように重みを学習していく。畳み込みニューラルネットワークでは、畳み込み層のフィルタに相当する重みは、同じフィルタに対応する重みと同じ値になる必要がある。これを重み共有といい、畳み込みニューラルネットワークでは重み共有を考慮した上で誤差が小さくなるように確率的勾配降下法により学習を行うことになる。

確率的勾配降下法は学習データの一部を使ってパラメータの更新を行う方法であり、用いる学習データを毎回変えてパラメータの更新を行うことで、違う目的関数を最小化するように更新が行われるために局所解にトラップされるリスクを低減することができる。

確率的勾配降下法において、重みの更新は以下のように行われる。

$$\mathbf{w} \leftarrow \mathbf{w} - \eta \frac{\partial E_s(\mathbf{w})}{\partial \mathbf{w}} \quad (1)$$

ここで、 \mathbf{w} は学習すべきパラメータ (重みやしきい値) をまとめたベクトル (以下、単に重みと呼ぶ)、 η は学習係数である。また、 $E_s(\mathbf{w})$ はパラメータ \mathbf{w} を用いて求めたサンプル s に対する出力に関する誤差関数であり、回帰問題を扱う場合には誤差関数として

$$E_s(\mathbf{w}) = \frac{1}{2} \sum_{p \in C^s} \sum_j \left(x_j^{(L)} - d_j^{(p)} \right)^2 \quad (2)$$

のような二乗誤差を用いる。ここで、 C^s はサンプル s に含まれるデータの集合、 $x_j^{(L)}$ は第 L 層 (出力層) の

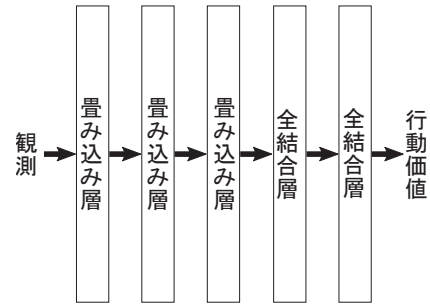


Fig. 2: Deep Q-Network の構造

ニューロン j の出力、 $d_j^{(p)}$ はデータ p の教師信号の j 番目の成分である。なお、これ以降の説明では、誤差 $E_s(\mathbf{w})$ のことを単に E と記述する。

3 Deep Q-Network

ここでは、提案モデルのもととなる Deep Q-Network¹²⁾ について説明する。

3.1 概要

Deep Q-Network は、Q Learning¹⁰⁾ における行動価値の学習を行う畳み込みニューラルネットワーク²⁾ である。Deep Q-Network では、ゲームのプレイ画面を観測として畳み込みニューラルネットワークに入力し、その状態におけるすべての行動価値を出力するように学習を行う。Q Learning はすべての状態と行動の組を学習することで最適な価値関数を得ることができるが、コンピュータゲームのように状態と行動の組み合わせが高次元になると、収束するまでにかかる時間が膨大になってしまう。それに対し、畳み込みニューラルネットワークは高次元のデータを扱うことに長けている。

Deep Q-Network は Atari2600 の複数のゲームに適用されており、それぞれのゲームに対してネットワーク構造や学習アルゴリズムを調整することはせずに、従来手法に勝る結果が得られることが示されている。また、多くのゲームにおいて、人間と同程度もしくはそれ以上の記録を出し、一部のゲームでは熟練した人間にも勝る成績を収め、有効性が確認されている。

3.2 構造

Deep Q-Network の構造を図2に示す。図2を見ても分かるように、Deep Q-Network は畳み込みニューラルネットワークに基づくモデルであり、3層の畳み込み層と2層の全結合層から構成されている。ゲームのプレイ画面を観測として畳み込みニューラルネットワークに入力し、その観測に対応する各行動に対する行動価値が出力される。第1~4層では出力関数として正規化線形関数を用いている。出力層である全結合層のニューロン数は扱う問題においてとり得る行動の数と同じになる。

なお、Deep Q-Network で学習する問題は、各観測とその観測における各行動の行動価値との関係を学習する回帰問題としてとらえることができるため、出力層である第5層の出力関数は恒等写像となる。

3.3 学習

Deep Q-Network では、観測であるゲームの画面の情報を入力とし、その観測におけるそれぞれの行動価値を出力とすることで、回帰問題として学習をおこな

う。出力となる行動価値は Q Learning における行動価値を用いるため、学習の際に用いられる誤差関数は

$$E = \frac{1}{2} \left(r_\tau + \gamma \max_{a' \in C^A(o_{\tau+1})} q(o_{\tau+1}, a') - q(o_\tau, a_\tau) \right)^2 \quad (3)$$

で与えられる。ここで、 r_τ は時刻 τ における報酬、 $C^A(o_{\tau+1})$ は観測 $o_{\tau+1}$ においてエージェントのとり得る行動の集合、 γ は割引率、 $q(o_\tau, a_\tau)$ は観測 o_τ において行動 a_τ をとることの価値を表す。

Deep Q-Network にゲーム画面 o_τ が入力されると、それに対し、観測 o_τ におけるすべての行動の価値が出力層から出力される。その行動価値に基づき、 ϵ -greedy 法で行動を決定する。 ϵ -greedy 法では、 ϵ ($0 \leq \epsilon \leq 1$) の確率ですべての行動からランダムに行動を選択し、 $1 - \epsilon$ の確率でルール の価値が最も高い行動を選択する。観測 o_τ において行動 a を選択する確率 $P(o_\tau, a)$ は

$$P(o_\tau, a) = \begin{cases} (1 - \epsilon) + \frac{\epsilon}{|C^A|} & \left(a = \operatorname{argmax}_{a' \in C^A} q(o_\tau, a') \text{ のとき} \right) \\ \frac{\epsilon}{|C^A|} & \text{(それ以外)} \end{cases} \quad (4)$$

で与えられる。ここで、 $|C^A|$ はエージェントのとり得る行動の種類の数であり、Deep Q-Network の出力層のニューロン数と同じ数になる。

選択された行動 a_τ が実行され、次の状態 $o_{\tau+1}$ に遷移する。また、行動 a_τ をとったことにより、スコアやゲームの状態などをもとに報酬 r_τ が与えられる。

Q Learning の行動価値を畳み込みニューラルネットワークを用いて近似するだけでは学習が不安定であり発散してしまうので、Deep Q-Network では Experience Replay、Fixed Target Q-Network、Reward Clipping と呼ばれる学習における工夫がなされている。

3.3.1 Experience Replay

強化学習によって与えられるデータは、時系列的に連続したものとなっている。そのため、得られたデータを学習にそのまま利用すると、連続して与えられるデータ間に相関が出てしまう。そこで、観測、行動、報酬、次の観測の組み合わせ $o_\tau, a_\tau, r_\tau, o_{\tau+1}$ を経験としてメモリに蓄積しておき、それをランダムにサンプリングしてミニバッチを作成し、確率的勾配降下法による学習に利用するという形をとっている。これを Experience Replay という。メモリには、最新の 1000 万個の観測、行動、報酬、次の観測の組み合わせを蓄えておく。

式 (3) において

$$r_\tau + \gamma \max_{a' \in C^A(o_{\tau+1})} q(o_{\tau+1}, a') \quad (5)$$

の部分が教師あり学習における教師信号に相当するものとなる。Experience Replay では、この教師信号に相当する部分の値を蓄積しておいたメモリから取り出した値を利用して計算することになる。

3.3.2 Fixed Target Q-Network

3.3.1 でも述べたように、式 (5) で表される部分が教師信号に相当するが、この部分の値を求めるには学習中のネットワークの重みを利用することになる。そのため、更新された重みを用いて計算された行動価値によって選択された行動やそれによって得られた報酬などを利用して式 (5) によって教師信号を求めるとなると、同じ入力 (観測) に対する教師信号が毎回変わってしまい、学習が非常に不安定になってしまう。そこで、Deep Q-Network では、3.3.1 でも述べたようにまずメモリからサンプリングしてミニバッチを作成し、それを用いて学習を行う間は教師信号の作成に用いるネットワークの重みは固定しておき、同じ入力 (観測) に対する教師信号が変化しないようにしている。ミニバッチによる一定回数の学習が終わったら、学習後のネットワークを教師信号の作成に用いるネットワークにコピーして上書きし、次からの学習にはそれを用いる。このように教師信号 (Target) の作成に用いるネットワークの重みをミニバッチによる一定回数の学習が終わるまで固定しておくことを Fixed Target Q-Network という。

3.3.3 Reward Clipping

報酬として与える値は問題によって異なるが、報酬の値を負なら -1 、正なら 1 、それ以外は 0 に固定することを Reward Clipping という。Reward Clipping を行うことにより、報酬に重み付けすることはできなくなるが、学習を速く進めることができるようになる。

さらに、学習の安定性を向上させるために重みの更新時に用いられる誤差関数の微分に対してもクリッピングを行う。式 (3) で与えられる二乗誤差において、Deep Q-Network から出力される行動価値 $q(o, a)$ に対する教師信号に相当する部分を $d(o, a)$ と表すと、誤差関数は

$$E = \frac{1}{2} (d(o, a) - q(o, a))^2 \quad (6)$$

と表すことができる。誤差関数の微分は

$$\frac{\partial E}{\partial \mathbf{w}} = \frac{\partial E}{\partial q(o, a)} \frac{\partial q(o, a)}{\partial \mathbf{w}} = (q(o, a) - d(o, a)) \frac{\partial q(o, a)}{\partial \mathbf{w}} \quad (7)$$

となる。ここで、誤差関数の行動価値 (出力) $q(o, a)$ に関する微分に相当する部分を

$$e \equiv q(o, a) - d(o, a) \quad (8)$$

と定義し、 e を

$$e \leftarrow \begin{cases} 1 & (1 \leq e) \\ e & (-1 < e \leq 1) \\ -1 & (e \leq -1) \end{cases} \quad (9)$$

のように $[-1, 1]$ の範囲にクリッピングする。このとき、式 (7) の誤差関数の微分は

$$\frac{\partial E}{\partial \mathbf{w}} = \begin{cases} \frac{\partial q(o, a)}{\partial \mathbf{w}} & (1 \leq e) \\ e \frac{\partial q(o, a)}{\partial \mathbf{w}} & (-1 < e \leq 1) \\ -\frac{\partial q(o, a)}{\partial \mathbf{w}} & (e \leq -1) \end{cases} \quad (10)$$

となる。これを積分することにより

$$E = \begin{cases} \frac{1}{2}e^2 & (-1 < e \leq 1) \\ |e| & (\text{それ以外}) \end{cases} \quad (11)$$

が得られる。つまり、誤差関数の $q(o, a)$ に関する微分 e を $[-1, 1]$ の範囲にクリッピングすることは、 e の値により、2種類の誤差関数を用いていることに相当する。 $|e|$ の値が大きいつま、り教師信号と出力との差が大きいつま、り重みの更新量が大きくなりすぎないように制限することで学習の安定性を向上させている。

3.4 学習の流れ

Deep Q-Network の学習は以下のようなアルゴリズムで行われる。

```

1: Replay Memory  $D$  の初期化
2: Network  $N$  を初期化
3: Target Network  $N_{\text{target}}$  を初期化  $N_{\text{target}} = N$ 
4: 行動確率  $\varepsilon$  を初期化  $\varepsilon = \varepsilon_{\text{ini}}$ 
5: for episode = 1 ...  $E$  do
6:    $\tau = 1$ 
7:    $ep\_end = false$ 
8:   while not  $ep\_end$  do
9:     現在の画面  $x_\tau$  から状態  $s_\tau$  を作成
10:     $o_\tau$  に基づき行動選択  $a_\tau =$ 
    { ランダムな行動 (確率  $\varepsilon$ )
    { または  $\text{argmax}_a q(s, a; N)$  (確率  $1 - \varepsilon$ )
11:     $a_\tau$  を実行し、報酬  $r_\tau$  と次の画面  $x_{\tau+1}$  と終了判定  $ep\_end$  を観測
12:     $x_{\tau+1}$  から状態  $s_{\tau+1}$  を作成
13:     $D$  に遷移 ( $s_\tau, a_\tau, r_\tau, s_{\tau+1}, ep\_end$ ) を追加
14:    if  $D$  にサンプルが  $M$  以上たまっていると then
15:       $D$  の中からランダムにサンプルを抜き出し、サンプル数  $M$  のミニバッチを作成する
16:      ミニバッチ  $j$  について順伝播を実行
17:      誤差  $E$  を求め、確率的勾配降下法を実行
18:    end if
19:     $T_{\text{update}}$  回ごとに Target Network の更新  $N_{\text{target}} = N$ 
20:     $\tau += 1$ 
21:     $\varepsilon_{\text{min}} < \varepsilon$  なら  $\varepsilon_r$  だけ  $\varepsilon$  を減少
22:  end while
23: end for

```

4 報酬分配を用いた Deep Q-Network

ここでは、提案する報酬分配を用いた Deep Q-Network について説明する。

提案モデルは、Deep Q-Network に基づいたモデルであり、行動価値を学習する際に Q Learning における価値の更新に加え、負の報酬が与えられた場合には Profit Sharing で行われる報酬分配の方法を利用して負の報酬を過去にさかのぼって分配する。

Q Learning では、次の状態において価値が最大となる行動における価値を利用して学習を行うため、正の報酬については時間とともに伝播されていく。しかし、負の報酬につながるような行動の価値が最大にはなり

得ないので、学習において伝播されず、そのような行動の価値を下げるような方向には学習は進まない。

そこで、負の報酬についてのみ報酬分配を行い、負の報酬の獲得につながるような行動をとらないことをより学習しやすくする。報酬分配は以下のように行う。

$$r_\tau = \frac{1}{(|C^{s_\tau}| + 1)^{\tau_r - \tau}} \quad (12)$$

ここで、 r_τ は時刻 τ における報酬、 $|C^{s_\tau}|$ は時刻 τ の状態においてとることができる行動数、 τ_r は負の報酬を獲得した時刻を表す。Profit Sharing では報酬の分配はエピソードが終了した後に行うが、提案モデルでは、報酬の入手時に時刻をさかのぼりながら分配を行う。

報酬分配を用いた Deep Q-Network の学習は、3.3 で述べた Deep Q-Network の学習アルゴリズムの 11 行目以降を変更した以下のようなアルゴリズムで行われる。

```

1: Replay Memory  $D$  の初期化
2: Network  $N$  を初期化
3: Target Network  $N_{\text{target}}$  を初期化  $N_{\text{target}} = N$ 
4: 行動確率  $\varepsilon$  を初期化  $\varepsilon = \varepsilon_{\text{ini}}$ 
5: for episode = 1 ...  $E$  do
6:    $\tau = 1$ 
7:    $ep\_end = false$ 
8:   while not  $ep\_end$  do
9:     現在の画面  $x_\tau$  から状態  $s_\tau$  を作成
10:     $s_\tau$  に基づき行動選択  $a_\tau =$ 
    { ランダムな行動 (確率  $\varepsilon$ )
    { または  $\text{argmax}_a q(s, a; N)$  (確率  $1 - \varepsilon$ )
11:     $a_\tau$  を実行し、報酬  $r_\tau$  と次の画面  $x_{\tau+1}$  と終了判定  $ep\_end$  を観測
12:    if  $r_\tau < 0$  のとき then
13:       $r_{\text{tmp}} = r_\tau$ 
14:      for time = 1 ...  $T$  do
15:         $r_{\text{tmp}} = r_{\text{tmp}} \times \frac{1}{(|C^{s_{\tau-time}}| + 1)}$ 
16:         $r_{\tau-time} = r_{\text{tmp}}$  に  $D$  を更新
17:      end for
18:    end if
19:     $x_{\tau+1}$  から状態  $s_{\tau+1}$  を作成
20:     $D$  に遷移 ( $s_\tau, a_\tau, r_\tau, s_{\tau+1}, ep\_end$ ) を追加
21:    if  $D$  にサンプルが  $M$  以上たまっていると then
22:       $D$  の中からランダムにサンプルを抜き出し、サンプル数  $M$  のミニバッチを作成する
23:      ミニバッチ  $j$  について順伝播を実行
24:      誤差  $E$  を求め、確率的勾配降下法を実行
25:    end if
26:     $T_{\text{update}}$  回ごとに Target Network の更新  $N_{\text{target}} = N$ 
27:     $\tau += 1$ 
28:     $\varepsilon_{\text{min}} < \varepsilon$  なら  $\varepsilon_r$  だけ  $\varepsilon$  を減少
29:  end while
30: end for

```

5 計算機実験

提案手法の動作を確認し、有効性を示すために第 6 章で述べた Atari2600 の 3 つのゲームを題材として計算機実験を行った。以下にその結果を示す。

Table 1: 畳み込みニューラルネットワークに関する実験条件

層名	フィルタサイズ	ストライド	出力サイズ	出力関数
入力	-	-	84 × 84 × 4	-
畳み込み層 1	8 × 8	4	20 × 20 × 32	ReLU
畳み込み層 2	4 × 4	2	9 × 9 × 64	ReLU
畳み込み層 3	3 × 3	1	7 × 7 × 64	ReLU
全結合層 1	-	-	512	ReLU
全結合層 2	-	-	行動種類数	恒等関数

Table 2: 学習に関する他の実験条件

学習エピソード数	E	$10^3 \sim 5 \times 10^3$
ϵ の初期値	ϵ_{ini}	1
ϵ の減少量	ϵ_r	$1/10^6$
ϵ の最低値	ϵ_{min}	0.1
評価エピソードにおける ϵ	ϵ'	0.05
Replay Memory のサイズ	D_{max}	10^5
ミニバッチサイズ	M	32
割引率	γ	0.99
Target Network の更新間隔	T_{update}	10^4

5.1 実験条件

表 1 に提案手法と比較対象の Deep Q-Network で用いる畳み込みニューラルネットワークに関する条件を示す。畳み込みニューラルネットワークへは 4 フレーム分のゲームのプレイ画面を入力する。Atari2600 のゲームの画面は 210×160 ピクセルの RGB 画像であるが、これをそのまま利用すると計算量が非常に大きくなってしまふ。そのため、RGB 画像をグレースケール化し、 110×84 ピクセルに縮小した後、中心の 84×84 ピクセルを切り出したものを 4 フレーム分まとめて 1 つの画像としたものを入力としている。また、出力層の出力は行動価値であるので、出力層のニューロン数は学習するゲームにおいてとることのできる行動の種類の数と同じになる。

また、学習に関する他の条件を表 2 に示す。行動選択は ϵ -greedy 法で行う。 ϵ の値は学習開始時には 1 に設定するため、完全にランダムに行動が選択される。 ϵ の値は 1 回行動するごと (1 ステップごと) に $1/10^6$ ずつ 0.1 になるまで減少させていく。また、10 エピソードごとに 1 エピソードを評価に用いるが、評価エピソードにおける ϵ の値 (ϵ') は 0.05 としている。ゲームによって 1 エピソードあたりのステップ数が異なるため、各ゲームにおいて ϵ の値が最小値になり、学習がある程度終了したと思われるエピソード数を選んだため、学習エピソードはゲームごとに別の値になっている。また、Replay Memory サイズは元の Deep Q-Network の実験においては 10^6 であったが、本実験では、使用する計算機の関係上 10^5 とした。サンプル数がこの値を上回ったときは一番初めのサンプルから順に上書きすることで、最新の 10^5 個のサンプルを学習に使用することができる。

5.1.1 スコアの推移

ここでは、提案する報酬分配を用いた Deep Q-Network と従来の Deep Q-Network において学習を行い、スコアの推移の比較を行った。

(1) Pong

図 3 に Deep Q-Network において Pong を題材として学習を行ったときのスコアの推移を示す。Pong では、

自分のスコアから相手のスコアを引いたものが最終的なスコアとなっており、0 を超えたら勝利したことになる。スコアは最低が -21 、最高が 21 となっている。学習エピソード数は 1000、評価エピソード数は 112 である。15 エピソード目くらいまではスコアが -21 に近いことが多いが、そこから 70 エピソード目くらいにかけて徐々にスコアが伸びていき、最初の 10 エピソードのスコアの平均が -20.9 であったのに対し、最後の 10 エピソードのスコアの平均は 14.7 となった。学習エピソード全体にかかったステップ数は 2408647 ステップ、評価エピソード全体にかかったステップ数は 363930 ステップであった。

図 4 に報酬分配を用いた Deep Q-Network において Pong を題材として学習を行ったときのスコアの推移を示す。15 エピソード目くらいまではスコアが -21 に近いことが多いが、そこから 65 エピソード目くらいにかけて徐々にスコアが伸びていき、最初の 10 エピソードのスコアの平均が -20.8 であったのに対し、最後の 10 エピソードのスコアの平均は 13.5 となった。学習エピソード全体にかかったステップ数は 3016489 ステップ、評価エピソード全体にかかったステップ数は 477353 ステップであった。

図 5 に Deep Q-Network と報酬分配を用いた Deep Q-Network における学習時のスコアの推移の比較を行った結果を示す。図 5 を見るとスコアの推移に大きな違いは見られないことが分かる。

(2) Breakout

図 6 に Deep Q-Network において Breakout を題材として学習を行ったときのスコアの推移を示す。学習エピソード数は 5000、評価エピソード数は 556 である。100 エピソード目くらいまではスコアが 0 に近いことが多いが、そこから 300 エピソード目くらいにかけて徐々にスコアが伸びていき、最初の 10 エピソードのスコアの平均が 2.7 であったのに対し、最後の 10 エピソードのスコアの平均は 17.6 となった。学習エピソード全体にかかったステップ数は 1964570 ステップ、評価エピソード全体にかかったステップ数は 332824 ステップであった。

図 7 に報酬分配を用いた Deep Q-Network において Breakout を題材として学習を行ったときのスコアの推移を示す。120 エピソード目くらいまではスコアが 0 に近いことが多いが、そこから 200 エピソード目くらいにかけて徐々にスコアが伸びていき、最初の 10 エピソードのスコアの平均が 1.6 であったのに対し、最後の 10 エピソードのスコアの平均は 14.8 となった。学習エピソード全体にかかったステップ数は 2221370 ステップ、評価エピソード全体にかかったステップ数は 180375 ステップであった。

図 8 に Deep Q-Network と報酬分配を用いた Deep

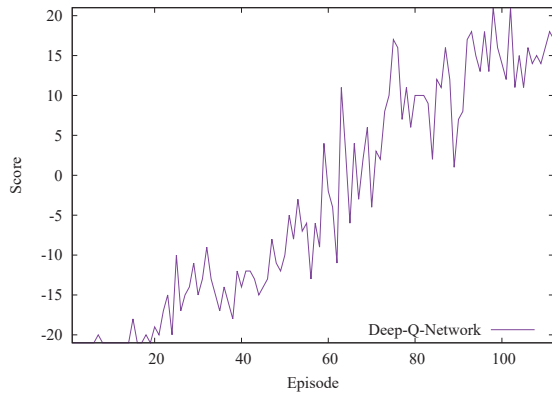


Fig. 3: Deep Q-Network におけるスコア推移 (Pong)

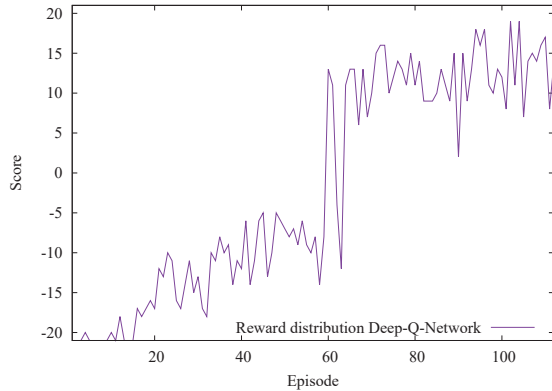


Fig. 4: 報酬分配を用いた Deep Q-Network におけるスコア推移 (Pong)

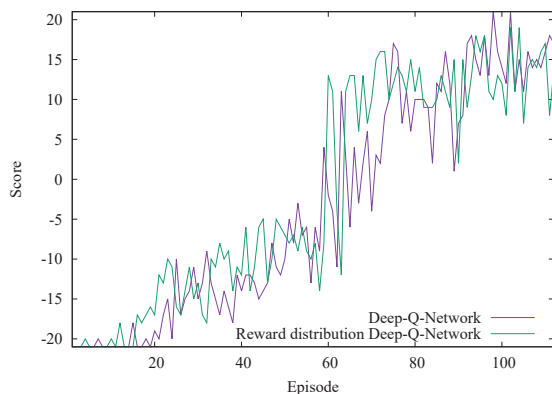


Fig. 5: スコアの推移の比較 (Pong)

Q-Network における学習時のスコアの推移の比較を行った結果を示す。また、図 8 ではスコアの上下の変動が激しく、比較がしにくいため、同じデータを 20 エピソードごとの平均として表示したのが図 9 である。報酬分配を用いた Deep Q-Network の方が少し早い段階で収束しているが、最終的なスコアにほとんど差は見られないことが分かる。

(3) Asterix

図 10 に Deep Q-Network において Asterix を題材として学習を行ったときのスコアの推移を示す。学習エピソード数は 5000、評価エピソード数は 556 である。最初の方はスコアが 100 から 500 くらいの範囲で、最後の方はスコアが 200 から 600 くらいの範囲で上下している。最初の 10 エピソードのスコアの平均が 240 であったのに対し、最後の 10 エピソードのスコアの平均

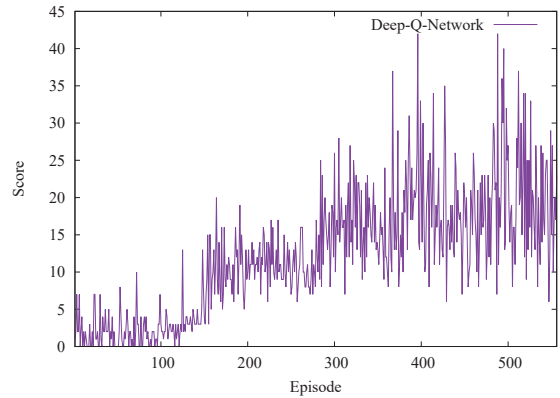


Fig. 6: Deep Q-Network におけるスコア推移 (Breakout)

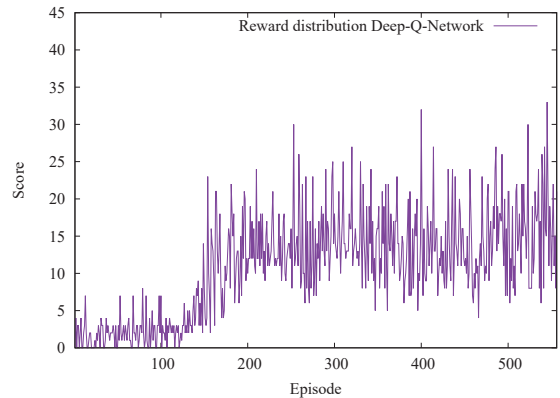


Fig. 7: 報酬分配を用いた Deep Q-Network におけるスコア推移 (Breakout)

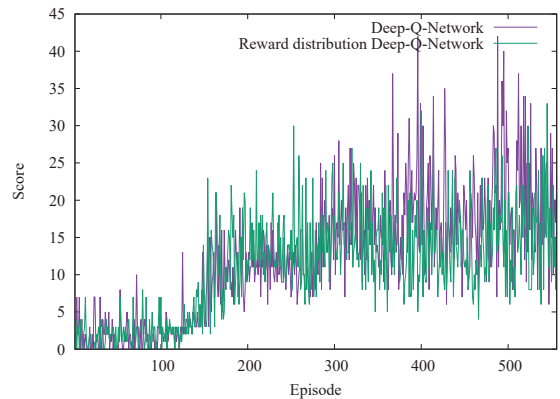


Fig. 8: スコアの推移の比較 (Breakout)

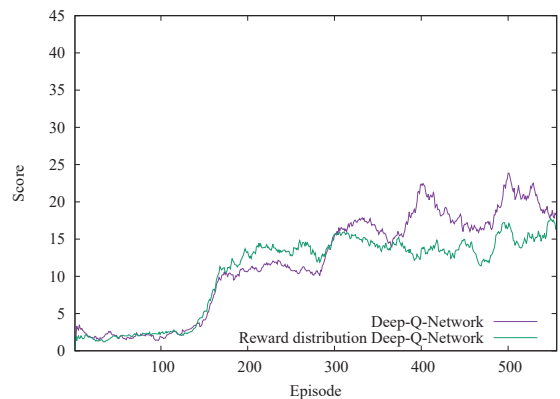


Fig. 9: スコアの推移 (20 エピソードごとの平均) の比較 (Breakout)

は 385 となった。学習エピソード全体にかかったステップ数は 1467501 ステップ、評価エピソード全体にかかったステップ数は 191082 ステップであった。

図 11 に報酬分配を用いた Deep Q-Network において Asterix を題材として学習を行ったときのスコアの推移を示す。最初の方はスコアが 100 から 500 くらいの範囲で、最後の方はスコアが 300 から 700 くらいの範囲で上下している。最初の 10 エピソードのスコアの平均が 305 であったのに対し、最後の 10 エピソードのスコアの平均は 480 となった。学習エピソード全体にかかったステップ数は 1554778 ステップ、評価エピソード全体にかかったステップ数は 180375 ステップであった。

図 12 に Deep Q-Network と報酬分配を用いた Deep Q-Network における学習時のスコアの推移の比較を行った結果を示す。また、図 12 ではスコアの上下の変動が激しく、比較がしにくいため、同じデータを 20 エピソードごとの平均として表示したのが図 13 である。Asterix は、Deep Q-Network が苦手としている問題である。最終的なスコアは提案手法である報酬分配を用いた Deep Q-Network の方が高くなっているが、いずれの手法でもエピソードごとのスコアの変動が激しく、学習結果に有意な差があるとまでは言いきれない。

5.2 負の報酬の獲得量の推移

5.1.1 で述べたようにスコアの推移に関しては提案手法と従来の Deep Q-Network との間に明確な差が見られなかった。提案手法では、負の報酬の分配を行っているため、適切に学習が行われていれば学習が進むにつれて負の報酬が獲得されることが少なくなると考えられる。

(1) Pong

図 14 は、Pong の学習において学習エピソードを 10 万ステップごとに区切り、その区間において獲得した負の報酬の回数をグラフにしたものである。図 14 より、報酬分配を行わない従来の Deep Q-Network に比べ、提案手法では負の報酬の獲得回数が 85%ほどになっていることが分かる。

(2) Breakout

図 15 は、Breakout の学習において学習エピソードを 10 万ステップごとに区切り、その区間において獲得した負の報酬の回数をグラフにしたものである。図 15 より、報酬分配を行わない従来の Deep Q-Network に比べ、提案手法では負の報酬の獲得回数が 90%ほどになっていることが分かる。

(3) Asterix

図 16 は、Asterix の学習において学習エピソードを 10 万ステップごとに区切り、その区間において獲得した負の報酬の回数をグラフにしたものである。図 16 より、報酬分配を行わない従来の Deep Q-Network に比べ、提案手法では負の報酬の獲得回数が 85%ほどになっていることが分かる。

以上の結果よりいずれの場合にも提案手法では Deep Q-Network に比べて学習速度が上がったり、精度が上がっているわけではないが、より負の報酬を獲得しないように学習できていることが分かった。

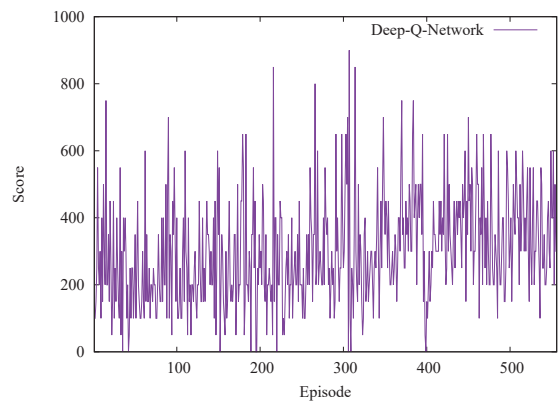


Fig. 10: Deep Q-Network におけるスコア推移 (Asterix)

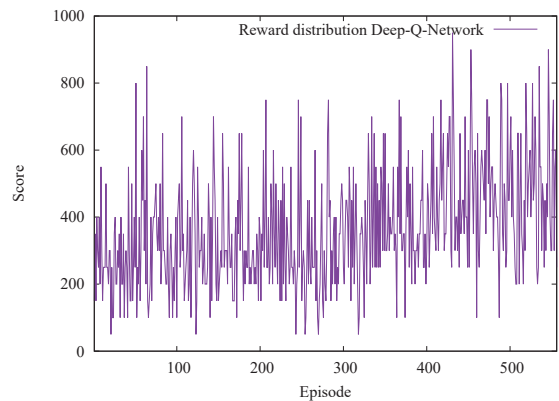


Fig. 11: 報酬分配を用いた Deep Q-Network におけるスコア推移 (Asterix)

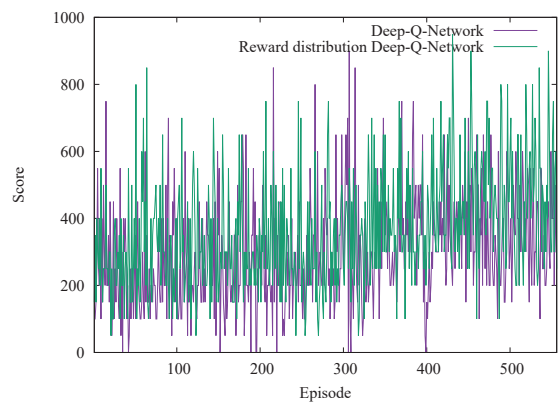


Fig. 12: スコアの推移の比較 (Asterix)

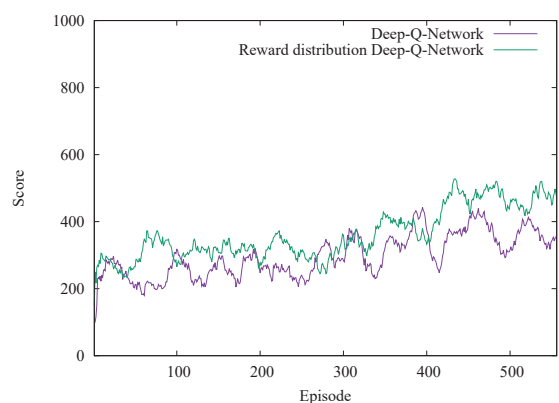


Fig. 13: スコアの推移 (20 エピソードごとの平均) の比較 (Asterix)

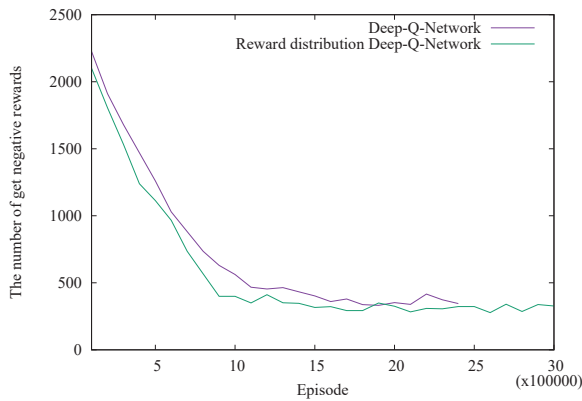


Fig. 14: 学習エピソードにおける負の報酬獲得回数の推移 (Pong)

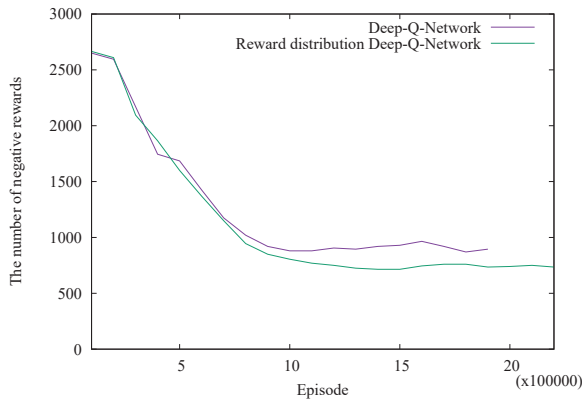


Fig. 15: 学習エピソードにおける負の報酬獲得回数の推移 (Breakout)

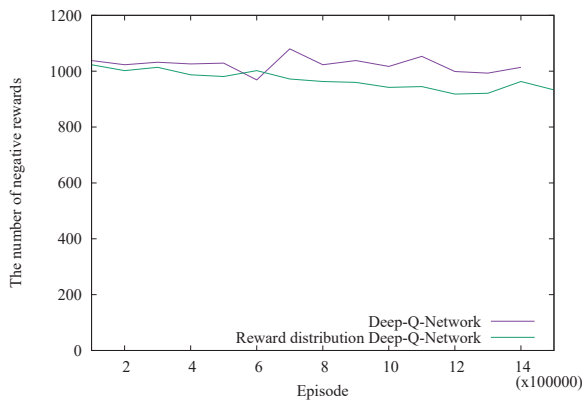


Fig. 16: 学習エピソードにおける負の報酬獲得回数の推移 (Asterix)

6 おわりに

本論文では、報酬分配を用いた Deep Q-Network を実現した。報酬分配を用いた Deep Q-Network は、従来の Deep Q-Network に基づいたモデルであり、行動価値を学習する際に Q Learning における価値の更新に加え、負の報酬が与えられた場合には Profit Sharing で行われる報酬分配の方法を利用して負の報酬を過去にさかのぼって分配する。Q Learning では、次の状態において価値が最大となる行動における価値を利用して学習を行うため、正の報酬については時間とともに伝播されていく。しかし、負の報酬につながるような行動の価値が最大にはなり得ないので、学習において伝播されず、そのような行動の価値を下げるような方向

には学習は進まない。提案手法の基本的な流れは従来の Deep Q-Network と同じであるが、Q Learning において負の報酬は伝播されないという部分に着目し、負の報酬を獲得したときのみ、Profit Sharing における報酬分配を行い、負の報酬をとるような行動をとらないように学習を行う。

Atari2600 の 3 つのゲームにおいて計算機実験を行い、提案手法と従来の Deep Q-Network の比較を行った。いずれのゲームにおいても、Deep Q-Network に比べて学習が速くなることも精度が上がることもないことが分かった。しかし、負の報酬の獲得回数は減っており、より負の報酬を獲得しないように学習できていることが確認できた。

参考文献

- 1) A. Krizhevsky, I. Sutskever and G. E. Hinton : ImageNet classification with deep convolutional neural networks. In Advances in Neural Information Processing System, pp.1097–1105, 2012.
- 2) Y. LeCun, L. Bottou, Y. Bengio and P. Haffner : Gradient-based learning applied to document recognition. Proceedings of the IEEE, Vol.86, No.11, pp.2278–2324, 1998.
- 3) G. E. Hinton, S. Osindero and Y. Teh : “A fast learning algorithm for deep belief nets,” Neural Computation, Vol. 18, pp.1527–1544, 2006.
- 4) F. Rosenblatt : “The Perceptron: a probabilistic model for information storage and organization in the brain,” Psychological Review, Vol. 65, No. 6, pp.386–408, 1958.
- 5) D. E. Rumelhart, J. L. McClelland and the PDP Research Group : Parallel Distributed Processing, Explorations in the Microstructure of Cognition Vol. 1: Foundations, The MIT Press, 1986.
- 6) G. E. Hinton : “A practical guide to training restricted boltzmann machines,” in Neural Network: Tricks of the Trade, Springer Berlin Heidelberg, pp.599–619, 2012.
- 7) A. Krizhevsky, I. Sutskever and G. E. Hinton : “ImageNet classification with deep convolutional neural networks,” Proceedings in Neural Information Processing Systems, pp.1097–1105, 2012.
- 8) R. S. Sutton and A. G. Barto : Reinforcement Learning : An Introduction, The MIT Press, 1998.
- 9) J. J. Grefenstette : “Credit assignment in rule discovery systems based on genetic algorithms,” Machine Learning, Vol.3, pp.225–245, 1988.
- 10) C. J. C. H. Watkins, and P. Dayan : “Technical Note: Q-Learning”, Machine Learning, Vol.8, pp.55–68, 1992.
- 11) 宮崎和光, 山村雅幸, 小林重信 : “ k -確実探索法 : 強化学習における環境同定のための行動選択戦略,” 人工知能学会誌, Vol.10, No.3, pp.454–463, 1995.
- 12) V. Mnih *et al.* : “Human-level control through deep reinforcement learning,” Nature, No.518, pp.529–533, 2015.
- 13) I. H. Witten : “An adaptive optimal controller for discrete-time Markov environments,” Information and Control, Vol.34, pp.286–295, 1977.
- 14) 宮崎和光, 山村雅幸, 小林重信 : “強化学習における報酬割当ての理論的考察,” 人工知能学会誌, Vol.9, No.4, pp.580–587, 1994.

構造適応型 Deep Belief Network 学習法による 検診データの学習速度の向上

○鎌田 真 (広島市立大学大学院情報科学研究科情報科学専攻)
市村 匠 (県立広島大学経営情報学部経営情報学科)

Improvement of Learning Speed For Comprehensive Medical Examination Data by Adaptive Structure Learning Method of Deep Belief Network

*S. Kamada (Dept. of Intelligent Systems, Graduate School of Information Sciences,
Hiroshima City University)

T. Ichimura (Faculty of Management and Information Systems,
Prefectural University of Hiroshima)

Abstract— Deep Belief Network (DBN) has a deep architecture that can represent multiple features of input patterns hierarchically with the pre-trained Restricted Boltzmann Machines (RBM). We have developed the adaptive structural learning method of DBN, that can discover an optimal number of hidden neurons in RBM and layers in DBN. The adaptive DBN obtains the highest classification accuracy among the other traditional methods for some image benchmark data sets. Recently, the learning method for the combination of image data and some text or numerical data called multi modal data has been proposed, but the preprocessing for their data such as the filter composition is required. In this paper, the classification accuracy and its computational time for the multi-modal data including not only image data but also numeric data and text data is verified on the adaptive DBN.

Key Words: Multi-Modal Data; Data Arrangement Method; Deep Belief Network; Restricted Boltzmann Machine, Adaptive Structure Learning; Comprehensive Medical Examination Data

1 はじめに

近年, Deep Learning(深層学習)¹⁾が人工知能の理論的な研究だけでなく, 産業界を中心に様々な場面でも利用され始めている。ILSVRC と呼ばれる画像認識の世界大会では, GoogleNet²⁾ や ResNet³⁾ で知られるように, 人間の認識能力を大きく上回る Deep Learning モデルが発表されている。このような背景の中, Deep Learning は画像認識だけでなく, 数値, テキスト, バイナリデータが混在したマルチモーダルデータに対しても適用され始めている。文献⁴⁾では, 時系列の数値データを画像に変換して分類を行う手法が提案されている。また富士通は, マルチモーダルデータに対する Deep Learning フレームワークとして Zinrai⁵⁾を開発している。ここでは, Deep Learning で学習を行う前に, カオス理論や Topological data analysis(TDA)を用いてマルチモーダルデータの前処理や特徴抽出を行っている。このような事前の処理は, 学習の精度を高めるために必要な処理だが, 一方で, 近年の産業界における IoT のデータ収集の観点では, 結果を出力するまでにかかる学習時間も重要な要素となっている。

Deep Learning の手法の1つとして, 事前学習を行った複数の Restricted Boltzmann Machine(RBM)⁶⁾を積み重ねることで, 複雑な入力パターンを表現できる Deep Belief Network(DBN)⁷⁾がある。我々は, 入力データ空間に応じて最適と考えられる RBM の隠れニューロン数および隠れ数を学習中に自動で求める構造適応型学習法を開発している⁸⁾。構造適応型学習法を画像ベンチマークデータセット CIFAR-10 及び CIFAR-100⁹⁾に適用した結果, 従来手法よりも高い分類精度が得られることが分かっている^{10), 11)}。

本論文では, 提案手法を画像データだけでなく数値やテキストが含まれたマルチモーダルデータに適用する。ここでは, 構造適応型 RBM の学習中における入力パターンに応じて, 最終的な学習時間を削減できるようにマルチモーダルデータの並びを入れ替える手法を提案する。提案手法をマルチモーダルデータとして Android アプリケーション “ひろしま観光マップ” により収集された主観的データ¹²⁾と広島県環境保険協会から提供された医療検診データ¹³⁾に適用した結果を報告する。

2 構造適応型 Deep Belief Network

2.1 Restricted Boltzmann Machine

RBM⁶⁾は, Fig. 1 のように, 可視層と隠れ層の2層から構成されるネットワーク構造を持ち, 確率分布に基づいて入力データに含まれる特徴を隠れニューロン上で学習する。RBM の学習では, 与えられた入力データに対して, 式 (1) のエネルギー関数を最小にするパラメタ $\theta = \{\mathbf{b}, \mathbf{c}, \mathbf{W}\}$ が最尤推定により求められる。

$$E(\mathbf{v}, \mathbf{h}) = - \sum_i b_i v_i - \sum_j c_j h_j - \sum_i \sum_j v_i W_{ij} h_j, \quad (1)$$

$$p(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} \exp(-E(\mathbf{v}, \mathbf{h})), \quad (2)$$

$$Z = \sum_{\mathbf{v}} \sum_{\mathbf{h}} \exp(-E(\mathbf{v}, \mathbf{h})), \quad (3)$$

ここで, v_i は i 番目の可視ニューロン, h_j は j 番目の隠れニューロンである。 b_i は v_i に対するパラメタ, c_j は h_j に対するパラメタ, W_{ij} は v_i と h_j 間の重みである。 $\mathbf{v} \in \{0, 1\}^I$ と $\mathbf{h} \in \{0, 1\}^J$ はそれぞれ可視層と隠れ層のベクトルであり, I と J は可視ニューロン数, 隠れニューロン数である。式 (2) は入力 \mathbf{v} と \mathbf{h} の確率分布

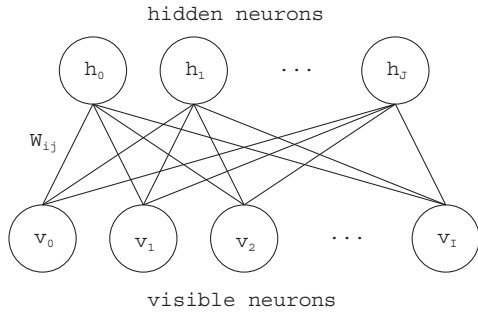


Fig. 1: Structure of RBM

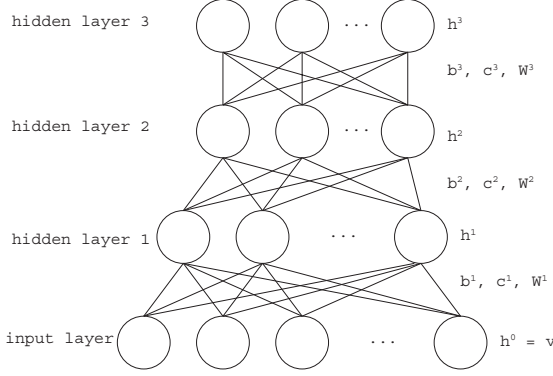


Fig. 2: Hierarchical network structure of DBN

である。式 (3) の Z は分配関数 (partition function) であり、 v と h の全ての 2 値のパターンの組み合わせに対するエネルギーを示す。RBM の学習では Contrastive Divergence (CD) 法¹⁴⁾ が使われ、これは可視層と隠れ層間の前向きと後ろ向き計算を繰り返すことで、モデルの出力を推定するサンプリング手法である。

2.2 Deep Belief Network

DBN は、事前学習を行った RBM を複数積み重ねることで、入力データの特徴を階層的に学習する Deep Learning 手法である⁷⁾。Fig. 2 は、3 つの RBM を持つ DBN のネットワーク構造を示している。入力データを $h^0 = v$ とすると、 $l (> 0)$ 層の RBM の j 番目の隠れニューロンは式 (4) により出力が求められる。

$$p(h_j^l = 1 | h^{l-1}) = \text{sigm}(c_j^l + \sum_i W_{ij}^l h_i^{l-1}), \quad (4)$$

ここで、 c^l と W^l は、それぞれ l 層の RBM の隠れニューロンのパラメータと重みである。 $\text{sigm}()$ はシグモイド関数である。DBN 自体は RBM を積み重ねたモデルであるため教師なし学習モデルであるが、最上位層に出力層を加え、ソフトマックス法等により計算された出力値と教師信号との誤差を最小化することで、教師あり学習に適用できる。ソフトマックス法では、与えられた入力データがカテゴリ k に属する確率 y_k が、式 (5) により求められる。

$$y_k = \frac{\exp(z_k)}{\sum_j^M \exp(z_j)}, \quad (5)$$

ここで、 z_j は、出力層における j 番目のニューロン出力活性化値である。 M は出力層のニューロン数である。

2.3 ニューロン生成/消滅アルゴリズム⁸⁾

我々は、階層型ニューラルネットワークにおける文献¹⁵⁾ の手法に基づき、RBM の学習中において最適な隠

れニューロン数を自動で求めるニューロン生成/消滅アルゴリズムによる構造適応型学習手法を提案した⁸⁾。ニューロン生成は、学習中におけるパラメータの変分に基づいて行われる。一般的に学習データに対して十分な数の隠れニューロンがあれば、学習が進むにつれて重みベクトルの変分は小さな値に収束する。一方、収束しない場合は、学習データに対するネットワークの表現能力が不足しており、データが持つ特徴 (確率分布) を隠れニューロン上で表現できていないことが考えられる。このような場合、関連する位置に隠れニューロンを挿入することで、多くの次元でデータを表現できると考えられ、学習が収束する。我々の調査では、RBM の 3 つのパラメータ $\theta = \{b, c, W\}$ のうち、CD 法による学習の収束に関係しているパラメータは入力信号に関係しない c と W であると仮定し¹⁶⁾、ニューロン生成の条件を式 (6) に示すように定義した。

$$(\alpha_c \cdot dc_j) \cdot (\alpha_W \cdot dW_j) > \theta_G, \quad (6)$$

ここで、 dc_j 、 dW_j はそれぞれ j 番目の隠れニューロンのパラメータの変分、重みの変分を示し、 α_c 、 α_W は各パラメータの変分のスケールを調整するために使われるパラメータであり、 θ_G は閾値で予め与えられる。学習中に式 (6) を満たす隠れニューロンが存在する場合、Fig. 3(a) のように、その隠れニューロンの近傍で属性 (パラメータや重みの値) を継承するように新しいニューロンが挿入される。

一方、入力データに対して十分な数の隠れニューロンが生成された後、出力に寄与していない、すなわち冗長な値を出力する隠れニューロンが存在する場合がある。構造適応型 RBM では、一定の学習の後、式 (7) を満たす隠れニューロンが存在すれば、ニューロン消滅アルゴリズムを適用し、その隠れニューロンを Fig. 3(b) のように消去する。式 (7) では、すべての入力データに対する隠れニューロンの出力平均値を観察し、一定の閾値以下のニューロンを冗長なニューロンとし、削除する。

$$\frac{1}{N} \sum_{n=1}^N p(h_j = 1 | v_n) < \theta_A, \quad (7)$$

$$p(h_j = 1 | v_n) = \sigma(b_j + \sum_i W_{ij} v_i), \quad (8)$$

ここで、 v_n は N 個の入力データセット中のある入力データである。 $p(h_j = 1 | v_n)$ は RBM の可視層に入力データ v_n が与えられたときの隠れニューロン $h_j \in \{0, 1\}$ の発火確率を示している。ここで、 $\sigma()$ はシグモイド関数で $[0, 1]$ を出力する関数である。 θ_A は閾値で予め与えられる。

2.4 層の生成条件¹⁰⁾

構造適応型 RBM におけるニューロン生成アルゴリズムでは、学習中の WD (パラメータ c と W の変分) を観察したが、構造適応型 DBN¹⁰⁾ では、個々の RBM ではなく、ネットワーク全体の WD とエネルギーを式 (9) と式 (10) で観察した。

$$\sum_{l=1}^k (\alpha_{WD} \cdot WD^l) > \theta_{L1}, \quad (9)$$

$$\sum_{l=1}^k (\alpha_E \cdot E^l) > \theta_{L2}, \quad (10)$$

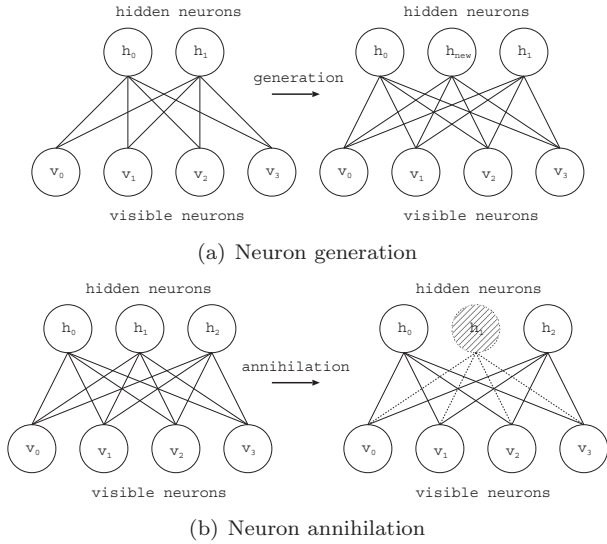


Fig. 3: Adaptive Learning method of RBM

ここで、 WD^l は l 層の RBM における c^l と W^l の変分で、 E^l はエネルギー関数を示す。 α_{WD} と α_E は層ごとの値を調整するためのパラメタであり、 θ_{L1} 、 θ_{L2} は適切な閾値である。 k 層の RBM の学習中に式 (9) 及び式 (10) を同時に満たす場合、入力データを表現するためにネットワーク全体の表現能力が不足していると考え、 k 層の学習が終了した時点で、 $k+1$ 層の RBM を生成し、学習を行う。

3 マルチモーダルデータ学習法

2.4 節で述べた構造適応型 RBM 及び構造適応型 DBN は、3 つのパラメタ $\theta = \{b, c, W\}$ のうち、 c, W の変分を観察し、与えられた入力データを表現するために最適と考えられる構造を求めていた。一方、パラメタ b は、入力データに対するパラメタであり、その変分はデータの特徴やデータ配列の並び等、入力信号の与え方によって変動する。本節では、数値や画像等の複数種類のデータが混在するマルチモーダルデータを、パラメタ b の変分と可視層と隠れ間の入出力のパターンに応じて、入力信号を入れ替えることで、分類能力を劣化させることなく、学習速度を短縮する手法を提案する。

3.1 データの扱い方法

本節では、マルチモーダルデータ中の画像データと数値等が含まれる CSV データの構造について述べる。Fig. 4 は、画像と CSV のデータ構造を示している。画像は 2 次元であるが、RBM における可視ニューロンは 1 次元のベクトルであるため、Fig. 4 のように、画像の左上から右下に向けて、各行ごとのピクセル値を Image Line として抽出している。Image Line には、複数の画像ブロック $IBlock_k$ が含まれているとする。すなわち、各画像データは、 $IBlockList = \{IBlock_1, \dots, IBlock_k, \dots, IBlock_K\}$ と表現される。 K は画像ブロックの数である。各画像の大きさは $N \times N$ ピクセルとする。

CSV の各データは、 L 個の項目が含まれる長さ M のベクトルとする。また、ある項目 l に関するデータを $CBlock_l$ とする。すなわち、CSV データは、 $CBlockList =$

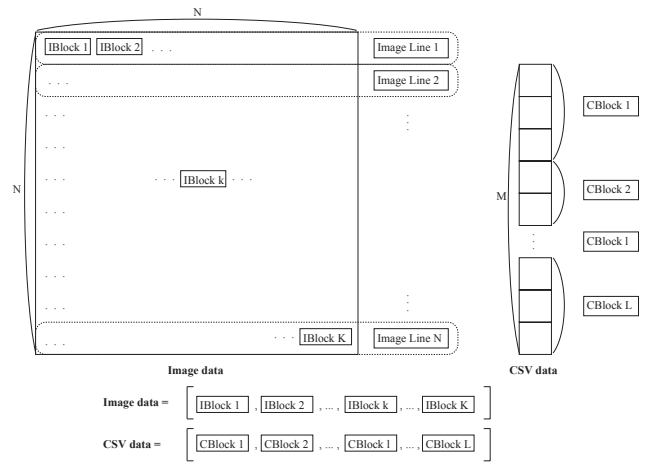


Fig. 4: Data Structure of Image and CSV

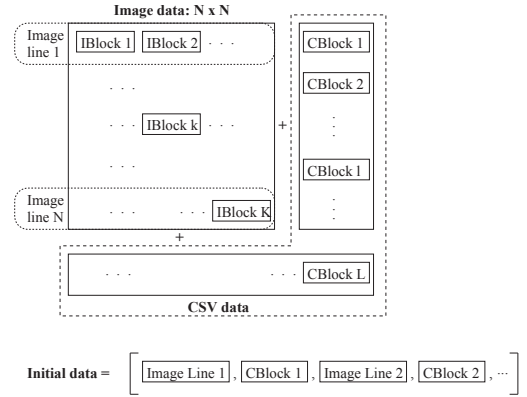


Fig. 5: Initial Arrangement of Image and CSV

$\{CBlock_1, \dots, CBlock_l, \dots, CBlock_L\}$ と表現する。 L は CSV ブロック数である。ここで項目とは、例えば検診データでは各血液検査や問診の結果等に該当する。各項目によってデータの長さは異なる。例えば、数値で表現される血液検査で、連続値を正常値・異常値(下)・異常値(上)の 2 値の one-hot ベクトルで表すと、3 ビットで表現する。

Fig. 5 は、学習前における画像ブロックと CSV ブロックの並びの初期値を示している。ここでは、元の画像の形状に対する自然な拡張として、画像の各行である Image Line と CSV の各ブロックである CBlock を交互に結合している。

3.2 アルゴリズム

Fig. 5 に示すように、学習を始める段階では、各入力データは画像の各行と CSV の各項目が交互に結合されたベクトルである。これらのデータの並びは単純に与えたものであり、それらが最適である保証はない。このため、学習状況に対しデータの並びをソートし、最適化する学習法を構築する必要がある。

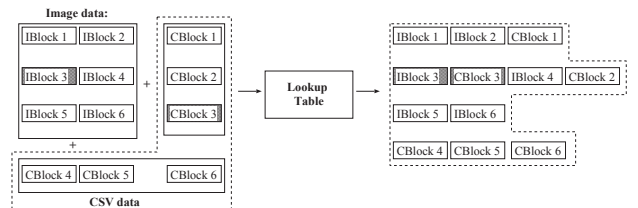


Fig. 6: Mapping with Look up Table

Algorithm 1 Multi Modal Data Learning procedure

- 1: 隠れ層 h の中で, 出力値 h_j が 1 で, かつ h_j の振動 (WD) が予め定められた閾値以下である隠れニューロンを $h^S \in h$ とする.
- 2: **for all** $h_j \in h^S$ **do**
- 3: h_j から可視層 v の計算を行う.
- 4: 画像ブロック, CSV ブロックの各領域の中で, $v_i = 1$ となる領域を求める. この領域が一定の値を超える画像ブロック, CSV ブロックの集合を, 入れ替え候補のブロックとしてそれぞれ $I\text{BlockCand} \in I\text{BlockList}$, $C\text{BlockCand} \in C\text{BlockList}$ とする.
- 5: **for all** $I\text{Block}_k \in I\text{BlockCand}$ **do**
- 6: $I\text{Block}_k$ の近傍ブロックを求める. ここで, 近傍とは, 1次元の可視ニューロンに対して, ニューロン間の順序的な位置が一定の値よりも小さいことを指す.
- 7: $C\text{BlockCand}$ の中で, b の振動 (WD) が最も高い CSV ブロックを, 入れ替えを行う CSV ブロックとして, $C\text{Block}_l$ とする.
- 8: **if** $C\text{Block}_l$ が $I\text{Block}_k$ の近傍に含まれていない場合 **then**
- 9: $C\text{Block}_l$ の位置を, $I\text{Block}_k$ の隣の位置に変更する.
- 10: $C\text{Block}_l$ を処理済みのブロックとし, $C\text{BlockCand}$ から除外する.
- 11: 現在の入力データの並びで $LookupTable$ を更新する.
- 12: **end if**
- 13: **end for**
- 14: **end for**

Algorithm 2 Inference procedure

Require: V : 入力データ, $Model$: 学習後のネットワーク.

$LookupTable$: 元の入力の位置と変換後の位置の対応を記録するテーブル.

- 1: $LookupTable$ に記録されている変換情報に基づいて, 元の入力データ V の順序を入れ替え, これを V' とする.
- 2: V' を学習済みのネットワーク $Model$ に与え, 推論を行う.

学習中における WD の状況に応じて, 出力値が定まった隠れニューロンから可視層への計算を行い, 発火する可視ニューロンのパターンを求めることが基本となる. ここで, ある画像ブロックと CSV ブロックの位置に対応する可視ニューロンがあれば, 両者に関連があると見なし, CSV ブロックの位置を画像ブロックの隣になるように, 可視ニューロンの位置を変更する. 例えば, $I\text{Block}_k$ と $C\text{Block}_l$ を入れ替える場合, 元の入力データの並びを

$\{\dots, I\text{Block}_k, I\text{Block}_{k+1}, \dots, C\text{Block}_l, C\text{Block}_{l+1}, \dots\}$ とすると, 入れ替え後の並びを

$\{\dots, I\text{Block}_k, C\text{Block}_l, I\text{Block}_{k+1}, \dots, C\text{Block}_{l+1}, \dots\}$ とする. **Algorithm 1** は, 提案手法のアルゴリズムを示している. なお, 入力データの入れ替え後, 元のデータ位置から入れ替え後の位置の対応を記録するた



(a) Register a new sightseeing spot (b) Google map with new locations

Fig. 7: Android smartphone application

Table 1: Health check items

Category	Name	Data type	Range
Basic test	Patient ID	Integer	
	Age	Integer	(10 - 134)
	Sex	Code	[Male, Female]
	Date	Integer	
	Height	Float	(117 - 196.7)
	Weight	Float	(27.6 - 175)
	BMI	Float	(11.9 - 57.3)
	Abdomen	Float	(53 - 157)
	Eye sight(right)	Float	(0 - 9.915)
	Eye sight(left)	Float	(0 - 9.915)
	Hearing(right, 1000)	Code	[Normal, Abnormal]
	Hearing(right, 4000)	Code	[Normal, Abnormal]
	Hearing(left, 1000)	Code	[Normal, Abnormal]
	Hearing(left, 4000)	Code	[Normal, Abnormal]
Blood pressure	Blood pressure(Max)	Integer	(70 - 257)
	Blood pressure(Min)	Integer	(26 - 148)
Urine	Protein	Code	[(-), (±), (1+), (2+), (3+)]
	Occult blood	Code	[(-), (±), (1+), (2+), (3+)]
	Urobilinogen	Code	[(-), (±), (1+), (2+), (3+)]
Blood analysis	WBC	Integer	(1200 - 26000)
	RBC	Integer	(234 - 672)
	Hb	Float	(5.5 - 22.3)
	Ht	Float	(20.6 - 65.2)
	PLT	Float	(2.7 - 112.6)
Liver function	GOT	Integer	(5 - 1134)
	GPT	Integer	(4 - 1909)
	Gamma GTP	Integer	(4 - 2329)
	ALP	Integer	(39 - 1758)
	LDH	Integer	(77 - 620)
	ChE	Integer	(103 - 621)
	ZTT	Float	(1 - 43.8)
	Total Bilirubin	Float	(0.1 - 4.8)
	TP	Float	(5.6 - 9.3)
	Alb	Float	(3.3 - 5.4)
Lipid	A/G	Float	(0.6 - 2.8)
	TC	Integer	(94 - 418)
Lipid	LDL	Integer	(4 - 357)
	HDL	Integer	(17 - 205)
	TG	Integer	(17 - 2628)
Diabetes	Sugar urine	Code	[(-), (±), (1+), (2+), (3+)]
	Blood sugar	Integer	(41 - 441)
Diabetes	HbA1c	Float	(4.7 - 12.7)
	Uric acid	Integer	[1, 3, 4, 6, 7]
Uric acid	Creatinine	Float	(0.28 - 13.68)
Kidney function	BUN	Integer	(5 - 59)
	eGFR	Float	(3.9 - 224.7)
	CRP	Code	[(-), (±), (1+), (2+), (3+)]
Infection	Hbs antigen	Code	[(-), (+)]
	Hbs antibody	Code	[(-), (+)]
	Hbc antibody	Code	[(-), (+)]
	Hcv antibody	Code	[(-), (±), (1+), (2+)]
	Pepsinogen	Code	[(-), (+)]
Additional test	Pylori	Code	[(-), (+)]
	Amylase	Float	(27 - 1335)
	ASO	Float	(10 - 393)
	CEA	Float	(0.2 - 10)
	CA15-3	Float	(4.600-19.400)
	TTT	Float	(0.3 - 10.6)
	Fecal occult blood	Code	[(-), (+)]
Other	Health questionnaire	Code	[1, 2]

めの表として, “Look up Table” が作成される. 学習後のネットワークを用いて推論を行う場合は, Fig. 6 および **Algorithm 2** に示すように, “Look up Table” を用いてデータの変換が行われる.

Table 2: Medical Image

種類	部位	撮影方法	がん検診
胸部 X 線	胸部	X 線	
胸部 CT	胸部	CT	○
胃部 X 線	胃部	X 線	○
マンモグラフィ	胸部	マンモグラフィ (X 線)	○

4 実験

4.1 データセット

本論文では、複数種類のデータが含まれているマルチモーダルデータとして、観光に関する主観的データ¹²⁾と、約 11 万人の検診データ¹³⁾の 2 種類のデータに対し、提案手法を適用し、分類精度と学習速度を評価した。

観光データは、我々が開発した MPPS アプリケーション (Mobilephone Based Participatory Sensing System) である「ひろしま観光マップ」¹⁸⁾を用いて収集されたもので、Fig. 7 に示すように、観光地を訪れた旅行者の主観的情報として位置情報、5 段階評価値、コメント文、写真を収集することができる。教師信号を与えるために、階層成長型 SOM (Growing Hierarchical Self-Organizing Map)^{19), 20)}を用いて、収集された 974 個のデータを 7 個のクラスに分類した。

検診データは、広島県環境保健協会¹³⁾から提供された定期健康診断データである。2012 年から 2015 年の 4 年間で、118,165 人の記録であり、表 1 に示す健康診断項目と表 2 に示す 4 種類の医療画像から構成される。

提案手法の学習パラメタとして、勾配の学習手法を Stochastic Gradient Descent (SGD)、学習係数を 0.01、バッチサイズは 100、学習の最大反復回数を 200 回、初期隠れニューロン数を 300 とした。構造適応型 RBM のパラメタとして、 $\theta_G = 0.050$, $\theta_A = 0.100$ とし、構造適応型 DBN のパラメタとして、 $\theta_{L1} = 0.1$, $\theta_{L2} = 0.1$ とした。これらのパラメタの値は、予備実験の結果による最良の組み合わせである。

4.2 実験結果

表 3, 表 4 は、それぞれひろしま観光マップ、検診データに対する実験結果である。各表は、各層におけるテストデータに対する正答率および学習にかかった計算時間 (CPU Time: sec) を示している。正答率は、10-fold Cross Validation 結果であり、10 回の試行における平均値、標準偏差、最大値、最小値を求めた。計算時間については、2 つの計算機を用いて比較を行い、Tesla の GPU が搭載された GPU 計算機 (Time1) (CPU: Intel(R) 24 Core Xeon E5-2670 v3 2.3GHz, GPU: Tesla K80 4992 24GB \times 3, Memory: 64GB, and OS: Cent OS 6.7 64 bit) と GTX 1080 が搭載された PC (Time2) (CPU: Intel(R) Core(TM) i5-4460 @ 3.20GHz, GPU: GTX 1080 8GB, Memory: 16GB, and OS: Fedora 23 64 bit) を用いて計測した。

実験の結果から、提案手法は、通常の DBN や構造適応型 DBN に比べて、学習精度を維持したまま、合計学習時間を削減できていた。最終層における学習精度は、構造適応型 DBN と提案手法で大きな違いはなかった。通常の DBN と提案手法による計算時間を比較した結果、ひろしま観光マップでは 25.9%、検診デー

タでは約 29.4%削減することができた。削減した計算時間 (分) はそれぞれ 26.9 分, 569.8 分であり (Time2), 14.3 分, 162.2 分であった (Time1) であった。

5 おわりに

様々な Deep Learning の学習手法が提案されており、画像認識の精度については、大変高い能力をもっていることが注目されている。また、画像だけでなく、数値やテキストデータ等の複数種類のデータが混在したマルチモーダルデータへの適用が期待されている。我々は、RBM や DBN に着目し、入力データに対し最適と考えられるネットワーク構造を自動で求める構造適応型 RBM、構造適応型 DBN を開発した。これらの手法は、RBM の 2 つの係数に対する変分をもとに構造変化を履行していたが、本論文では入力信号の与える順序を学習状態に応じて変更することで、 \mathbf{b} に関連した変化を小さくすることができたと考えている。その結果、提案手法により精度を維持したまま約 30% 計算時間を短縮に成功した。今後は、時系列データ等、他の種類のマルチモーダルデータに対しても提案手法を適用し、精度の検証を行う。

謝辞

本研究開発は総務省 SCOPE (受付番号: 162308002) の委託、人工知能財団による助成、および JSPS 科研費 (課題番号: JP17J11178) の助成を受けたものです。

参考文献

- 1) Y. Bengio: *Learning Deep Architectures for AI*, Foundations and Trends in Machine Learning archive, vol.2, no.1, 1/127 (2009)
- 2) C. Szegedy, W. Liu, et al.: *Going Deeper with Convolutions*, 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1/9 (2015)
- 3) K. He, X. Zhang, S. Ren, J. Sun: *Deep residual learning for image recognition*, arXiv preprint arXiv:1512.03385 (2015)
- 4) Z. Wang, T. Oates: *Imaging time-series to improve classification and imputation*, Proc. of the 24th International Conference on Artificial Intelligence (IJCAI15), 3939/3945 (2015)
- 5) 富士通, 時系列 Deep Learning: <http://journal.jp.fujitsu.com/2016/03/10/01/>, (2017 年 12 月 2 日閲覧)
- 6) G. E. Hinton: *A Practical Guide to Training Restricted Boltzmann Machines*, Neural Networks, Tricks of the Trade, Lecture Notes in Computer Science (LNCS, vol.7700), 599/619 (2012)
- 7) G. E. Hinton, S. Osindero and Y. Teh: *A fast learning algorithm for deep belief nets*, Neural Computation, vol.18, no.7, 1527/1554 (2006)
- 8) S. Kamada and T. Ichimura: *An Adaptive Learning Method of Restricted Boltzmann Machine by Neuron Generation and Annihilation Algorithm*, Proc. of 2016 IEEE International Conference on Systems, Man, and Cybernetics (IEEE SMC 2016), 1273/1278 (2016)
- 9) A. Krizhevsky: *Learning Multiple Layers of Features from Tiny Images*, Master of thesis, University of Toronto (2009)
- 10) 鎌田真, 市村匠: **忘却機能を持った Deep Belief Network による構造適応型学習**, 計測自動制御学会第 9 回コンピュータ・インテリジェンス研究会講演論文集, 92/97 (2016)
- 11) S. Kamada and T. Ichimura: *An Adaptive Learning Method of Deep Belief Network by Layer Generation Algorithm*, Proc. of IEEE TENCON2016, 2971/2974 (2016)

Table 3: Classification Accuracy (KankouMap)

Model	Layer	Ave.	Std.	Max.	Min.	Sort	Iterations	Time1	Time2
Traditional DBN	1	0.821	0.011	0.829	0.819	0	489	4.5	9.1
	2	0.846	0.008	0.861	0.867	0	411	3.3	6.7
	3	0.862	0.009	0.863	0.865	0	430	3.8	6.3
	4	0.862	0.012	0.869	0.861	0	379	3.8	6.8
	5	0.859	0.010	0.864	0.863	0	354	3.9	7.0
Total							19.3	35.9	
Adaptive DBN	1	0.915	0.010	0.921	0.910	0	429	3.9	7.3
	2	0.934	0.008	0.940	0.929	0	356	3.1	5.9
	3	0.954	0.012	0.961	0.949	0	329	3.2	6.1
	4	0.978	0.009	0.9800	0.972	0	338	3.4	6.5
	5	0.982	0.010	0.985	0.978	0	331	3.9	6.9
Total							17.5	32.7	
Multi-Modal Learning	1	0.920	0.010	0.915	0.928	363	370	3.8	7.2
	2	0.945	0.010	0.950	0.941	28	299	2.5	4.6
	3	0.953	0.008	0.959	0.948	4	278	2.6	4.9
	4	0.971	0.009	0.974	0.967	0	256	2.7	5.0
	5	0.982	0.010	0.984	0.979	0	260	2.7	65.2
Total							14.3	26.9	

Table 4: Classification Accuracy (Medical Data)

Model	Layer	Ave.	Std.	Max.	Min.	sort	Iterations	Time1	Time2
Traditional DBN	1	0.833	0.011	0.851	0.818	0	500	44.1	151.8
	2	0.862	0.008	0.873	0.850	0	500	37.2	128.3
	3	0.864	0.009	0.877	0.850	0	430	34.7	122.1
	4	0.892	0.010	0.904	0.875	0	456	37.8	130.3
	5	0.907	0.006	0.916	0.893	0	421	36.8	128.9
	6	0.911	0.010	0.930	0.895	0	433	39.1	131.8
Total							229.7	793.0	
Adaptive DBN	1	0.835	0.007	0.847	0.825	0	500	43.1	150.4
	2	0.861	0.008	0.870	0.842	0	440	29.3	99.2
	3	0.864	0.010	0.874	0.845	0	402	31.1	110.0
	4	0.896	0.007	0.907	0.883	0	411	33.5	111.9
	5	0.912	0.012	0.941	0.901	0	422	35.9	120.3
	6	0.944	0.008	0.953	0.932	0	406	35.2	124.0
Total							208.1	715.69	
Multi-Modal Learning	1	0.854	0.012	0.870	0.832	502	413	36.6	130.4
	2	0.879	0.011	0.898	0.861	127	367	25.1	89.2
	3	0.878	0.006	0.888	0.871	23	305	23.7	88.0
	4	0.927	0.009	0.941	0.915	0	299	24.5	86.9
	5	0.942	0.008	0.953	0.925	0	307	25.7	86.3
	6	0.942	0.008	0.953	0.931	0	295	26.6	89.0
Total							162.2	569.8	

- 12) ITProducts, ひろしま観光マップ: <https://play.google.com/store/apps/details?id=jp.itproducts.KankouMap>, (2017年12月2日閲覧)
- 13) 広島県環境保健協会: <http://www.kanhokyo.or.jp/>, (2017年12月2日閲覧)
- 14) G.E.Hinton: *Training products of experts by minimizing contrastive divergence*, Neural Computation, vol.14, no.8, 1771/1800 (2002)
- 15) T.Ichimura and K.Yoshida Eds.: *Knowledge-Based Intelligent Systems for Health Care*, Advanced Knowledge International (ISBN 0-9751004-4-0) (2004)
- 16) 鎌田真, 市村匠, 原章: ニューロン生成/消滅アルゴリズムによる構造適応型 *Restricted Boltzmann Machine*, 計測自動制御学会第8回コンピュータショナル・インテリジェンス研究会講演論文集, 90/96 (2015)
- 17) S.Kamada and T.Ichimura: *A Structural Learning Method of Restricted Boltzmann Machine by Neuron Generation and Annihilation Algorithm*, Neural Information Processing, vol.9950 of the series Lecture Notes in Computer Science, 372/380 (2016)
- 18) N.D.Lane, E.Miluzzo, L.Hong, D.Peebles, T.Choudhury and A.T.Campbell: *A survey of mobile phone sensing*, IEEE Communications Magazine, Vol.48, No.9, 140/150 (2010)
- 19) A.Rauber, D.Merkl, M.Dittenbach: *The growing hierarchical self-organizing map: exploratory analysis of high-dimensional data*, IEEE Transactions on Neural Networks, Vol.13, No.6, 1331/1341 (2012)
- 20) T.Ichimura, S.Kamada and K.Kato: *Knowledge Discovery of Tourist Subjective Data in Smartphone Based Participatory Sensing System by Interactive Growing Hierarchical SOM and C4.5*, Intl. J. Knowledge and Web Intelligence, Vol.3, No.2, 110/129 (2012)

深層生成モデルとエンコーダを用いた自動修正法 および住宅設計への適用

○植田考哉 瀬尾昌孝 西川郁子 (立命館大学)

Application of Generative Adversarial Networks with Encoder Model to Building Construction of Ordinary Houses

* T. Ueda, M. Seo and I. Nishikawa (Ritsumeikan University)

Abstract— Generative Adversarial Networks (GAN) are composed of two networks of discriminator and generator, which learns a distribution of a given data set by adversarial learning. We introduce an encoder to infer a latent space of the obtained generator. Encoder is trained not only as an inverse mapping of the generator, but also as an auto-corrector of an invalid data to be reconstructed as a valid data through the generator with a minimum modification. Proposed method is first applied to image data of a handwritten figure, to be modified to another figure. It is also applied to voxel data of a construction design, to be corrected to attain a required quake resistance.

1. はじめに

深層学習による生成モデルが多数提案されているが、最も注目されているものの一つが Goodfellow ら¹⁾によって提案された Generative Adversarial Nets (以下 GAN) である。ここでは、隠れ変数で表現される潜在変数空間上の写像としてデータを生成する生成器

(以下 Generator) と、その Generator から出力された生成データか或いは実観測データかを識別する識別器 (以下 Discriminator) の 2 つの深層ネットワークが相互に学習する。つまり Generator は、Discriminator がデータの統計性から観測データと識別しがたいデータ生成を学習することで、観測データの分布を獲得する。対象物 (例えば、「イス」) に対する十分多くの観測データを用いることで、学習後の Generator は観測データに含まれていない、しかしヒトが見ても「イス」と思えるデータを生成するようになる。

本報告では、さらに符号化器 (以下エンコーダ) を組合せることで、学習済みの Generator を用いて、その潜在変数空間の特定領域から、一定の条件を満たすデータのみを生成することを試みる。ここでエンコーダは、データから潜在変数空間への写像だが、設定された条件を満たすデータに対しては、Generator が同じデータを生成するよう自己符号化器となり、他方、条件を満たさないデータに対しては、条件を満たすデータのうち復元誤差が最小となるデータを Generator が生成するよう、いわばデータの自動修正機能の実現を目指す。

計算機実験として、まず、認識や分類課題に広く用いられてきた手書き数字データセット MNIST で提案法を検証する。2 種の数字を対象に、手書き数字の自動修正を行う。次に、3 次元建築物を対象に、強度制約を満たすよう部材配置の自動修正を行う。

以下では、まず 2 章で GAN を紹介し、続く 3 章で、提案するエンコーダを付き GAN によるデータの自動修正法を説明する。提案法を用いた数値実験として、4 章では MNIST の手書き数字データ、5 章では住宅部材配置データをそれぞれ対象とし、6 章でまとめる。

2. 深層生成モデル

2.1. Generative Adversarial Nets

Goodfellow ら¹⁾の GAN では、Generator(式中では写像 G で表わす)および Discriminator(同じく写像 D で表わす)の 2 つの深層ネットワークが、同時に、下式(1)の min-max 問題を逐次的に求解することで学習を行う。

$$\min_G \max_D V(G,D) = E_{\mathbf{x} \sim p_{data}(\mathbf{x})} \log D(\mathbf{x}) + E_{\mathbf{z} \sim p(\mathbf{z})} \log(1 - D(G(\mathbf{z}))) \quad (1)$$

ここで $G(\mathbf{z})$ は、潜在空間 \mathbf{z} 上のある確率分布 $p(\mathbf{z})$ のもとで \mathbf{z} からデータを生成する写像である。また $D(\mathbf{x})$ は、入力データ \mathbf{x} が与えられたときに、それが観測データ (の分布 $p_{data}(\mathbf{x})$ から与えられた) か、或いは Generator が生成したデータ ($p(\mathbf{z})$ から生成された Generator による分布 $p_g(\mathbf{x})$ から与えられた) か、を識別する写像である。両者とも深層ネットワークで与える。式(1)で与えられる $V(G,D)$ に対して、Discriminator はこれを最大化することによって、観測データの分布と Generator の分布との JS ダイバージェンスを求めて識別を実現し、Generator はこれを最小化することで観測データの分布を獲得するように、それぞれが学習する。

2.2. Wasserstein GAN

前節の GAN とは異なる分布間距離を用いる学習法として、Arjovsky ら²⁾は Wasserstein GAN (以下 WGAN) を提案した。WGAN では、下式(2)で与えられる Earth Mover's Distance (あるいは Wasserstein 距離) と呼ばれる分布間距離を用いる。その定義は、式(2)のように上界値で与えられ、ここで f_w はリプシッツ定数 1 のリプシッツ連続な任意の写像である。学習過程でモード崩壊と呼ばれる状態に陥ることがある前節の GAN と比較して、安定した学習を行うことができる。

$$W(p_{data}, p_g) = \sup_{\|f_w\|_L \leq 1} E_{\mathbf{x} \sim p_{data}(\mathbf{x})} [f_w(\mathbf{x})] - E_{\mathbf{x} \sim p_g(\mathbf{x})} [f_w(\mathbf{x})] \quad (2)$$

式(2)の表式は、Discriminator と Generator を用いて、次のように最大値の表式に変形できる。このとき f_w は Discriminator そのものであり、Discriminator はパラメータ w に関する最大化問題を解くことで学習する。

$$\max_D E_{\mathbf{x} \sim p_{data}(\mathbf{x})} [D(\mathbf{x})] - E_{\mathbf{z} \sim p(\mathbf{z})} [D(G(\mathbf{z}))] \quad (3)$$

Discriminator の学習は、上界値で与えられる Wasserstein 距離を求めるために式(3)の最大化に対応するが、2つの分布間距離をより正確に求めることがデータの識別に相当する。同時に、Generator の学習は、その距離の最小化に対応するが、それによって観測データの分布 $p_{data}(\mathbf{x})$ を内部に獲得する。

なお、WGAN では定数 1 のリップシッツ連続を満たす f_w が必要となる。Gulrajani ら³⁾は、定数 1 のリップシッツ連続を満たす Discriminator を獲得するのに代えて、Discriminator の学習に用いる目的関数に、以下のペナルティ項を加えて学習している (以下 WGAN-GP³⁾。

$$\begin{aligned} & E_{\mathbf{x} \sim p_{data}(\mathbf{x})} [D(\mathbf{x})] - E_{\mathbf{z} \sim p(\mathbf{z})} [D(G(\mathbf{z}))] \\ & - \lambda E_{\hat{\mathbf{x}} \sim p(\hat{\mathbf{x}})} [(\|\nabla_{\hat{\mathbf{x}}} D(\hat{\mathbf{x}})\|_2 - 1)^2] \quad (4) \\ & \text{ただし } \hat{\mathbf{x}} = \varepsilon \mathbf{x} + (1 - \varepsilon) \tilde{\mathbf{x}} \\ & \mathbf{x} \sim p_{data}, \tilde{\mathbf{x}} \sim p_g, \varepsilon \sim U[0,1] \\ & \lambda = 10 \end{aligned}$$

3. 提案モデル

3.1. Variational Auto-encoder

前章の GAN とは異なる深層生成モデルに、Kingma ら⁴⁾によって提案された Variational Auto-encoder (以下 VAE)がある。これは自己符号化器の名のとおり、エンコーダおよびデコーダを用いて、データを一旦低次元の潜在変数空間での表現を経て再構築し、その再構築誤差最小化を学習する。ここでデコーダは、単純な確率分布で与えられる低次元の潜在空間からデータを生成する写像であり、逆にエンコーダは、与えられたデータに対応する潜在空間中の点を推論する写像として学習される。一般に VAE では、GAN と比較して安定した学習ができる。VAE の安定性と GAN の精練性を両立させる試みとして、VAE/GAN⁵⁾や α -GAN⁶⁾がある。両者はどちらも、エンコーダおよび Generator の学習において、再構築誤差項を置く。学習した Generator では、潜在空間に、観測データの特徴量と分布を反映した多様体が獲得されると考えられる⁷⁾。

3.2. GAN with Encoder

本稿では、観測データを用いて学習した GAN の Generator に対して、その潜在空間を推論するエンコーダを導入する。つまり Generator をデコーダとして、再構築誤差を最小化するようにエンコーダを学習する。その際、設定した条件を満たさないデータ \mathbf{x}_{NG} に対しては、条件を満たすデータ \mathbf{x}_{OK} のうち再構築誤差が最小

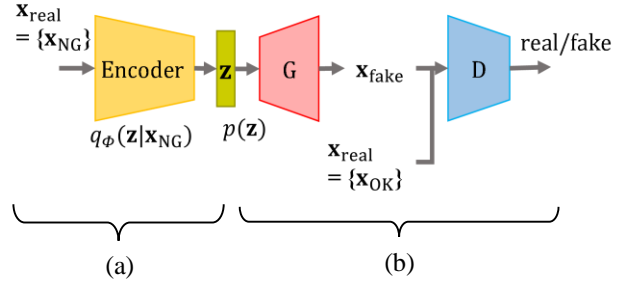


Fig. 1: Proposed model of GAN with encoder

のものを Generator の潜在空間から探して生成させるよう、自動修正機能も担わせる。

学習に用いるモデルを Fig.1 に示す。図中の \mathbf{x}_{real} は観測データ、 \mathbf{x}_{OK} は Generator に生成させたい条件を満たしたデータを示す。Generator の学習に用いる観測データ \mathbf{x}_{real} は、条件を満たしているもののみとする。

まず、Generator および Discriminator のみを GAN で学習させる (Fig.1 b)。図では、確率分布 $p(\mathbf{z}) = \mathcal{N}(\mathbf{z}|0, \mathbf{I})$ に従って Generator で生成されたデータを \mathbf{x}_{fake} と示している。学習後に生成される \mathbf{x}_{fake} には、一般には条件を満たすもの \mathbf{x}_{OK} のみでなく満たさないもの \mathbf{x}_{NG} もありうると考えられる。

次に、エンコーダのみを学習させる。エンコーダへの入力データにも、条件を満たすもの \mathbf{x}_{OK} も満たさないもの \mathbf{x}_{NG} もあるとする。図では満たさないデータ \mathbf{x}_{NG} として示されている。エンコーダの学習は、Generator を経て再構築された誤差の最小化とし、下式(5)で表される L2 ノルムで評価する。これにより、エンコーダに入力されるデータと Generator から出力されるデータが最も近い潜在変数の推論器としてエンコーダを学習する。

$$E_{\mathbf{z} \sim q_{\phi}(\mathbf{z}|\mathbf{x}_{NG})} [\|\mathbf{G}(\mathbf{z}) - \mathbf{x}_{NG}\|_2] \quad (5)$$

潜在空間 \mathbf{z} 上の確率分布 $p(\mathbf{z})$ は正規分布としたが、エンコーダもまた入力データ \mathbf{x} に応じて、潜在空間 \mathbf{z} 上の正規分布 $q_{\phi}(\mathbf{z}|\mathbf{x}) = \mathcal{N}(\mathbf{z}|\boldsymbol{\mu}_{\phi}, \boldsymbol{\sigma}_{\phi}^2 \mathbf{I})$ のパラメータ $\boldsymbol{\mu}_{\phi}$ 、 $\boldsymbol{\sigma}_{\phi}^2$ を出力するものとした。 \mathbf{z} は以下の Reparametrization trick⁴⁾によって確率的に与える。

$$\begin{aligned} \mathbf{z} &= \boldsymbol{\mu}_{\phi}(\mathbf{x}) + \varepsilon \odot \boldsymbol{\sigma}_{\phi}(\mathbf{x}) \\ \text{ただし } \varepsilon &\sim \mathcal{N}(0, \mathbf{I}), \odot \text{ は各要素の積} \quad (6) \end{aligned}$$

ここで、式(5)の再構築誤差最小化のみで評価すると、Generator の潜在空間全域から探索することになる。そこで、エンコーダ出力の分布を、Generator 潜在空間の分布 $p(\mathbf{z})$ に近付けることを考える。すなわち、(5)式に両分布の KL ダイバージェンスをペナルティ項として加え、再構築誤差とともに最小化することでエンコーダを学習する (以下「L2+KL 損失モデル」とする)。

$$\begin{aligned} & E_{\mathbf{z} \sim q_{\phi}(\mathbf{z}|\mathbf{x}_{NG})} [\|\mathbf{G}(\mathbf{z}) - \mathbf{x}_{NG}\|_2] \\ & + \lambda D_{KL}(q(\mathbf{z}|\mathbf{x}_{NG}) \| p(\mathbf{z})) \quad (7) \end{aligned}$$

KL 項の係数 λ はハイパーパラメータであり、探索範囲を制御する。

しかしながら、KL 項の存在は、正確な推論モデルの獲得を困難にする場合がある。ここでの Generator は、 \mathbf{x}_{OK} を生成するよう学習しており (GAN の学習に用いる観測データは全て \mathbf{x}_{OK})、潜在空間上の $p(\mathbf{z})$ は \mathbf{x}_{OK} データの分布を反映している。そのため、 \mathbf{x}_{NG} データに対して L2 ノルムを最小とるように学習したエンコーダ出力の分布は、 $p(\mathbf{z})$ の分布とは一般に異なる。そこで、KL 項に替えて、以下の最小化によるエンコーダの学習を考える。

$$E_{\mathbf{z} \sim p(\mathbf{z})} [\|E(G(\mathbf{z})) - \mathbf{z}\|_2] \quad (8)$$

E はエンコーダ

つまり、Generator によって生成されたデータには、全て対応する \mathbf{z} が存在するため、それらに対してはエンコーダは正にその \mathbf{z} を推論するよう学習する。このときエンコーダは、Generator の逆関数として機能する。

以上より本稿では、(5)式と(8)式を交互に用いたエンコーダの学習を提案する(以下「L2, G^{-1} 損失モデル」とする)。次の4章では、MNIST 手書き文字データを用いて、手書き文字の自動修正を行う。その際、「L2+KLモデル」「 G^{-1} のみモデル」、「L2, G^{-1} モデル」の3つの異なる損失関数でエンコーダを学習し、特に要修正データに対する結果を比較する。

4. 手書き文字データ MNIST に対する実験

本章では、手書き文字データである MNIST を対象とする。従来から MNIST は分類や認識課題の学習データとして広く用いられ、深層生成モデルでも低次元の潜在変数空間に比較的分かりやすい特徴量を見出すことができる。'0'から'9'までの10種類の手書き数字データが予め10のラベル付きで用意されている。そこで以下では、仮想的な条件として、ある一つのラベル(数字'9'とする)のデータを \mathbf{X}_{OK} 、それとは異なる一つのラベル(数字'7'とする)のデータを \mathbf{X}_{NG} とおき、前節で提案した GAN with Encoder を用いて、各手書き文字'7'からそれに近い'9'に「修正」する実験を行う。

以下に、使用したデータおよびネットワークと学習における各設定を示す。

- WGAN の学習に用いる \mathbf{X}_{OK} データ : MNIST のラベル'9'の 5000 データ
- エンコーダの学習に用いる \mathbf{X}_{NG} データ : MNIST のラベル'7'の 5000 データ
- 手書き文字画像データ : サイズ 28×28 のグレースケール
- Generator の潜在変数空間の次元数 : 16
- エンコーダの学習におけるバッチサイズ : 100
- \mathbf{X}_{NG} の全 5000 データを一巡する 50 バッチを入力してエンコーダのパラメータ更新する繰り返しを 1epoch とし、全 400epoch で学習
- 「L2, G^{-1} 損失モデル」によるエンコーダの学習では、 G^{-1} 損失関数(式(8))による更新 1 回に対して、L2 損失関数(式(5))による更新を 50 回行う。

Table 1 に Generator, Discriminator およびエンコーダとして用いた各ネットワークの構造をまとめて示す。

D
Input data $\mathbf{x} \in R^{28 \times 28 \times 1}$
Convolution, filter size=4×4, stride=2, 64, lReLU
Convolution, filter size=4×4, stride=2, 128, lReLU
output: full connection→1
G
$\mathbf{z} \in R^{16} \sim \mathcal{N}(0, I)$
dense→4×4×128
Deconvolution, filter size=4×4, stride=2, 128, ReLU
Deconvolution, filter size=4×4, stride=2, 64, ReLU
Deconvolution, filter size=4×4, stride=2, 6, Sigmoid
E
Input data $\mathbf{x} \in R^{28 \times 28 \times 1}$
Convolution, filter size=4×4, stride=2, 64, ReLU
Convolution, filter size=4×4, stride=2, 128, ReLU
output: $\boldsymbol{\mu} \rightarrow 16, \boldsymbol{\sigma}^2 \rightarrow 16$

Table 1: Network architectures of Generator (G), Discriminator (D), and Encoder (E)

学習の手順を示す。まず Generator および Discriminator の学習には、観測データとして \mathbf{X}_{OK} データ'9'のみを用い、WGAN-GP の学習を行う。その結果 Generator は、 \mathbf{X}_{OK} データの分布を内部に獲得し、生成すると期待される。その後、学習済みの Generator のパラメータを固定し、次は \mathbf{X}_{NG} データ'7'のみを用いてエンコーダを学習させる。この学習の目標は、入力された \mathbf{X}_{NG} データに (L2 ノルムの意味で) 最も近い \mathbf{X}_{OK} データに対応する潜在表現を見出すことである。そこで、学習の過程として、エンコーダへの入力データと Generator からの出力データにおける再構築誤差の変化を見る。ここでは、式(7)と(8)に示した損失関数として、「L2+KLモデル」で $\lambda = 0, 1$, および 10^{-5} の3つの値と、「 G^{-1} のみモデル」、式(5)と式(8)を組み合わせた「L2, G^{-1} モデル」の5つの損失関数による学習を比較した。

Fig.2 にそれらの損失関数でエンコーダを学習したときの再構築誤差の変化を示す。「L2+KLモデル」($\lambda = 0, \lambda = 10^{-5}$)、「L2, G^{-1} モデル」で学習したエンコー



Fig. 2: Reconstruction error during the learning

だが、再構築誤差を最小にする潜在変数を探索できている。特に、「L2+KL モデル」で $\lambda = 0$ （つまり「L2のみ」）の時に400epoch後に最小の再構築誤差を与えた(誤差0.01616)。

これらの学習の結果、エンコーダに入力されたデータが「修正」されてGeneratorから出力されたデータをFig.3に例示する。図(a)に示したラベル'7'の100データをエンコーダに入力した結果、図(b)に示す各データが出力された。図(a), (b)の10行10列の各位置のデータは、対応する入出力対である。対を見比べると、誤差を小さく保ったまま、'9'に分類されるデータが出力されていることが見られる。なお、ここでのエンコーダの損失関数は「L2のみ」である。



(a) Input data



(b) Corrected data

Fig. 3: Example of correction of '7' to '9'

ここでの実験は、比較的小さい2次元画像データを用いて、異なるデータクラスへの「修正」を行った。しかし、出力されたデータが所望のクラスか否かの基準は必ずしも明確ではない(学習済みの分類器を用いることは勿論できるが)。そこで次の5章では、より複雑な3次元立体データを用い、さらに形状ではなく力学的な条件を充足しているか否かの2クラスに対して、同様の実験を行う。

5. 建築設計における実験

本章では、3次元構造をもつデータを対象として同様の自動修正を試みる。ここでの課題は、以下のような条件を満たす3次元データの生成と自動修正である。

5.1. 建造物の部材配置

建造物として、例えば、2階や3階建ての一般住宅を建てる際、要求として部屋割りや扉、窓、階段や吹抜けなどの建物外形が与えられたとき、耐震基準を満たすよう、柱や梁などの建築部材を配置する必要がある。一般に、部材を増やすほど強度が増すが、同時に部材コストも増える。そのため、できるだけ少ない部材で強度制約を満たすような配置を求めることになるが、部材が規格化された工業化住宅の場合には、どんな厚みや太さを持つどの種の部材をいくつどの位置に配置するか、を決める組合せ最適化ともみなせる。他方、強度条件は、各部材に各方向から長期・短期の力を加えた際に生じる歪みの上限値などである。一般の住宅でも数百以上にのぼる力学条件として、構造計算シミュレーションによって制約の充非が求められる。

ここでは、ある3階建て一般住宅を対象に部材配置を考える。まず、部材配置状態を15cm立方のボクセルデータで表わす。各ボクセルは、その領域を占有する部材割合を値として持つ。Fig.4に梁などの水平部材と柱などの鉛直部材に分けてボクセルデータで表わした配置例を示す。カラーバーに示されるボクセル値0-1を表している。

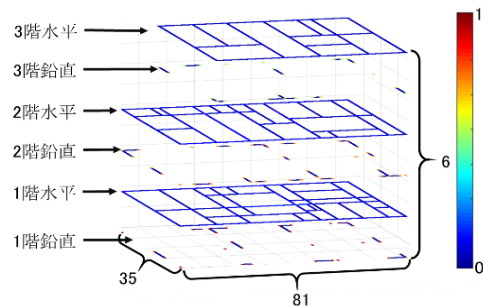


Fig. 4: Example of member placement of ordinary three-story house

強度は、複数部材の相互位置関係に依存し、ある部材の変更や位置ずれにより何処の強度がどう変わるかは単純ではなく、建物全体に対する構造計算によってのみ正確に求められる。つまり、ボクセルデータは、部材配置をほぼ一意に表現するが(厳密には一意ではない)、その立体形状と、強度充足という力学条件の対応は自明ではない。それに対して、以下では、Generatorには制約を充足する配置を生成させること、さらに、エンコーダには非充足な配置を与えて充足配置に自動修正させることを試みる。

5.2. 学習に用いる部材配置データ

- GANの学習に用いる X_{OK} データ：全強度制約を満たす部材配置15000データ
- エンコーダの学習に用いる X_{NG} データ：最低1つの強度制約を満たさなかった部材配置15000データ。(部材量や部材配置は X_{OK} データに近い、構造計算でも定量的に良質な X_{NG} データ)
- 部材配置データ：サイズが幅35×奥行81×高6のボクセルで部材占有率を[0,1]で表現

5.3. GAN による部材配置の生成

まず, Generator と Discriminator からなる GAN によって \mathbf{X}_{OK} データの生成を学習する. ここでは, 3 種の GAN を用いて学習を行い, 比較する. 比較方法として, 生成される部材配置データの制約充足を評価する. ただしここでは, 簡便法として構造計算は行わず, \mathbf{X}_{OK} と \mathbf{X}_{NG} を判別する深層ネットワークを用いて評価した.

GAN および WGAN-GP を比較する. ここではいずれも DCGAN⁷⁾を用いた. 加えて, 2 段階の学習を行うことで洗練されたデータを生成するために Zhang ら⁸⁾が提案した Stack GAN も用いて比較した. この概要を Fig.5 に示す. まず 1 段階目の生成として, 図中 G_{s1} で示される Generator が, 互いに依存関係が強い部材群の配置を学習する. その学習が終了した後, 2 段階目の生成として, 図中 G_{s2} で示される Generator が, 残りの部材を含む全部材の配置を学習する. 以下ではこれを Stack WGAN-GP とする. そこの各学習で用いる損失関数は以下で与える(ただし, ペナルティに関する項は省いた).

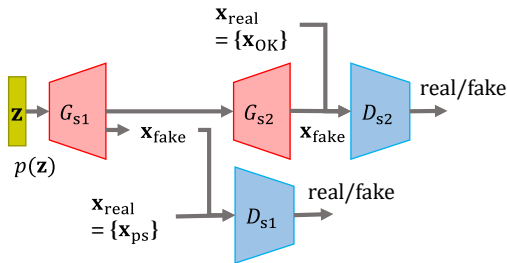


Fig. 5: Proposed model of Stack WGAN-GP: First G_{s1} and D_{s1} pair are trained to generate a partial design \mathbf{x}_{ps} , then, G_{s2} and D_{s2} pair are trained to generate a whole design \mathbf{x}_{OK} while G_{s1} is fixed.

$$\max_{D_{s1}} E_{\mathbf{x}_{ps} \sim p_{ps}(\mathbf{x})} [D_{s1}(\mathbf{x}_{ps})] - E_{\mathbf{z} \sim p(\mathbf{z})} [D_{s1}(G_{s1}(\mathbf{z}))] \quad (9)$$

$$\max_{D_{s2}} E_{\mathbf{x}_{ok} \sim p_{data}(\mathbf{x}_{ok})} [D_{s2}(\mathbf{x}_{ok})] - E_{\mathbf{x}_{s1} \sim G_{s1}(\mathbf{z}), \mathbf{z} \sim p(\mathbf{z})} [D_{s2}(G_{s2}(\mathbf{x}_{s1}))] \quad (10)$$

ここで, \mathbf{x}_{ps} は 1 段階目で扱う部材のみを配置したデータであり, \mathbf{x}_{OK} は全部材が配置されたデータである. G_{s1} は式(9)を最小化するように, G_{s2} は式(10)を最小化するように学習する. まず G_{s1} と D_{s1} を学習し, その後 G_{s1} を固定して, G_{s2} と D_{s2} を学習する. 学習後のデータ生成は, 潜在変数 $\mathbf{z} \sim p(\mathbf{z})$ から G_{s1} , G_{s2} の順に写像して, 最終出力を得る.

上記の 3 つをそれぞれ DCGAN, WGAN-GP, Stack WGAN-GP とし, 各ネットワークの構造を Table 2 に示す. Discriminator および Generator の最適化手法にはいずれも Adam⁹⁾を適用し, ネットワークの各層には Batch Normalization¹⁰⁾(BN)を用いた. Stack WGAN における G_{s2} には Residual Block¹¹⁾をネットワークの中間に持つ. なお, 潜在変数空間の次元数は 100 とした.

これら 3 つに対して, 評価用ネットワークを用い, 生成データのうち \mathbf{X}_{OK} データと判定された割合を比較

D, D_{s1}, D_{s2}
Input data $\mathbf{x} \in R^{35 \times 81 \times 6}$
Convolution, filter size=5×5, stride=2, 64. BN, lReLU
Convolution, filter size=5×5, stride=2, 128. BN, lReLU
Convolution, filter size=4×4, stride=2, 256. BN, lReLU
Convolution, filter size=4×4, stride=2, 512. BN, lReLU
output: full connection→1
G, G_{s1}
$\mathbf{z} \in R^{100} \sim \mathcal{N}(0, I)$
dense→2×5×512
Deconvolution, filter size=4×4, stride=2, 256. BN, ReLU
Deconvolution, filter size=4×4, stride=2, 128. BN, ReLU
Deconvolution, filter size=5×5, stride=2, 64. BN, ReLU
Deconvolution, filter size=5×5, stride=2, 6. Sigmoid
G_{s2}
Input data $\mathbf{x}_{ps} \in R^{35 \times 81 \times 6}$
Convolution, filter size=5×5, stride=2, 64. BN, ReLU
Convolution, filter size=5×5, stride=2, 128. BN, ReLU
Residual Block
Convolution, filter size=5×5, stride=2, 128. BN, ReLU
Convolution, filter size=5×5, stride=2, 128.
BN, ReLU
Deconvolution, filter size=5×5, stride=2, 64. BN, ReLU
Deconvolution, filter size=5×5, stride=2, 6. Sigmoid

Table 2: Network architectures of the proposed DCGAN, WGAN-GP and Stack WGAN-GP: Adam optimizer (learning rate $\alpha = 0.0002$, $\beta_1 = 0.5$ and $\beta_2 = 0.9$) is used in all cases.

する. 評価用ネットワークは VGG-16¹²⁾を用い, 前述の \mathbf{X}_{OK} と \mathbf{X}_{NG} 各 15000 データで 2 クラス分類を学習した. テストデータ各 1000 に対する分類精度は 91% であった.

学習過程での判定結果を Fig.6 に示す. 学習では, バッチサイズを 200 とした. 全 15000 データを一巡し Generator を 75 回更新するのを 1epoch とし, 1epoch ごとに 100 データずつ生成し, 評価用ネットワークにより判定した. 100 データのうち \mathbf{X}_{OK} と判定された割合を示している. さらに図には, 20epoch ごとの平均を示した. なお, 3 つのモデルの学習回数はそれぞれ DCGAN : 500epoch, WGAN-GP : 1000epoch, Stack WGAN-GP : 1 段階目および 2 段階目で各 1000epoch とした.

図(a)に示す DCGAN では, 生成されたデータの多くが形状的にも不良で, 学習が困難であった. 図(b)の WGAN-GP では最大 84%, 図(c)の Stack WGAN-GP では最大 92% の生成データが制約充足と判定された. そこで次節では, Stack WGAN-GP で学習済み Generator を用いて, エンコードによる \mathbf{X}_{NG} データの自動修正を学習する. そこの制約充足判定には, 厳密に構造計算を行った.

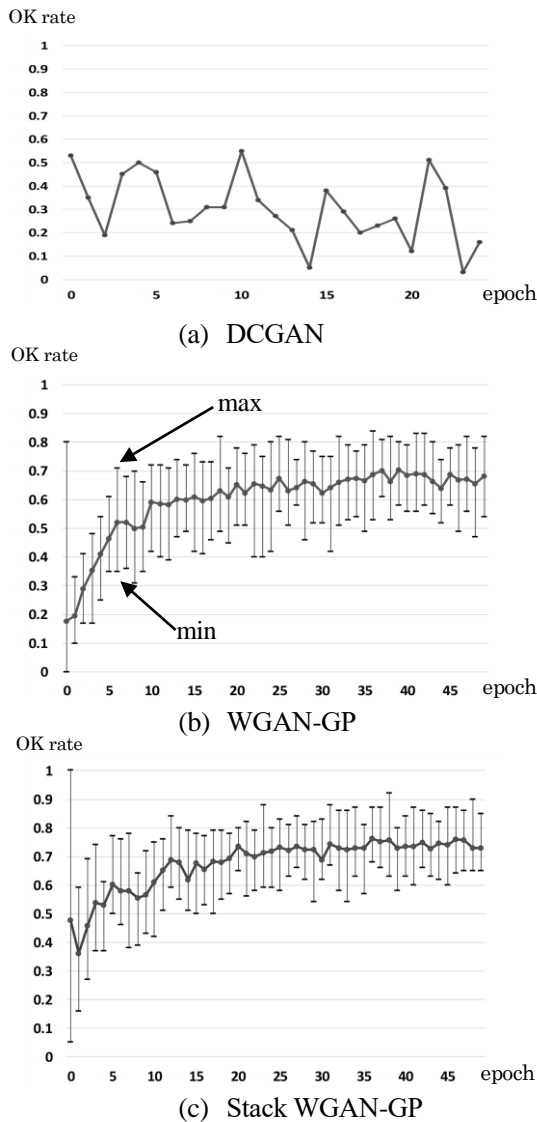


Fig. 6: Generation rate of 'valid' design evaluated by a network during the learning of each GAN.

5.4. データエンコーダによる配置修正

Table 3 にエンコーダのネットワーク構造を示す。学習時は式(5)および式(8)を交互に最小化しパラメータを更新する。ただし $G(\mathbf{z})$ を $G_{s2}(G_{s1}(\mathbf{z}))$ とし、式(5)で 50 回更新した後式(8)で 1 回更新した。

評価として、エンコーダの学習に用いていない制約違反データ 100 例を用いて、再構築データの制約充非を構造計算により求めた。その結果、制約充足となっ

E
Input data $\mathbf{x} \in R^{35 \times 81 \times 6}$
Convolution, filter size=5×5, stride=2, 64. BN, ReLU
Convolution, filter size=5×5, stride=2, 128. BN, ReLU
Convolution, filter size=5×5, stride=2, 256. BN, ReLU
Convolution, filter size=5×5, stride=2, 512. BN, ReLU
output: $\mu \rightarrow 100, \sigma^2 \rightarrow 100$

Table 3: Network architecture of encoder

たデータ 38, 制約違反のままのデータ 53, 構造計算が不能であったデータ 9 となった。制約充足となった場合には、部材位置は変えず太さのみを変えるなどの部材配置の微小修正で実現していた。制約違反であったデータの多くは、部材の一部が欠損しており、計算不能となったものも同様であった。この理由と対応については、検討を進めている。

6. おわりに

深層生成器の表現能力を活かしつつ、潜在変数空間を読み解くエンコーダや、分類器などを組み合わせることで、所望の条件を満たすデータへの自動修正を試みた。個別の適用事例における課題とともに、学習モデルの検討もさらに進めたい。

参考文献

- 1) Ian J. Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, Yoshua Bengio, Generative Adversarial Nets, in NIPS 2014.
- 2) Martin Arjovsky, Soumith Chintala, and Leon Bottou, Wasserstein GAN, arXiv:1701.07875v2, 2017.
- 3) Ishaan Gulrajani, Faruk Ahmed, Martin Arjovsky, Vincent Dumoulin, Aaron Courville, Improved Training of Wasserstein GANs, arXiv preprint arXiv:1704.00028, 2017.
- 4) Diederik P. Kingma, Max Welling, Auto-Encoding Variational Bayes, in NIPS 2014.
- 5) Anders Boesen Lindbo Larsen, Søren Kaae Sønderby, Hugo Larochelle, Ole Winther, Autoencoding beyond pixels using a learned similarity metric, arXiv:1512.09300v2, 2016.
- 6) Mihaela Rosca, Balaji Lakshminarayanan, David Warde-Farley, Shakir Mohamed, Variational Approaches for Auto-Encoding Generative Adversarial Networks, arXiv:1706.04987v1, 2017.
- 7) Alec Radford & Luke Metz, Soumith Chintala, Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks, in ICLR 2016.
- 8) Han Zhang, Tao Xu, Hongsheng Li, Shaoting Zhang, Xiaogang Wang, XiaoLei Huang, StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks, Dimitris Metaxas. arXiv:1612.03242. 2016.
- 9) Diederik P. Kingma, Jimmy Lei Ba, Adam: A method for stochastic optimization, in ICLR 2015.
- 10) S.Ioffe, C.Szegedy, Batch Normalization: Accelerating deep network training by reducing, ArXiv:1502.03167, 2015.
- 11) Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun, Deep Residual Learning for Image Recognition, arXiv:1512.03385. 2015.
- 12) Karen Simonyan, Andrew Zisserman, Very deep convolutional networks for large-scale image recognition, in ICLR 2015.

複雑な環境下における可変型測域センサアレイの知的制御

○北井 瑛佳 (首都大) 戸田 雄一郎 (首都大) 武居 直行 (首都大)
和田 一義 (首都大) 久保田 直行 (首都大)

Intelligent Control for Variable Sokuiki Sensor Array in the Complicated Environment

* S. Kitai, Y. Toda, N. Takesue, K. Wada, N. Kubota (Tokyo Metropolitan University)

Abstract— Recently, in order to reduce secondary disasters at the disaster site, the expectation of rescue robots has been increasing. In the disaster site, many kinds of environments are mixed. Therefore, in order to make robots perform the environmental sensing in the complicated environment, it is necessary to change the measurement area and density according to the environment information. Therefore, we developed a Variable Sokuiki Sensor Array as a 3D-distance measurement system for changing measurement area and density. In this paper, we propose an intelligent control method for the sensor array based on Multi-Objective Behavior Coordination, and show results of proposed method in a simulation experiment.

Key Words: Robot Vision, Laser Range Finder, 3D Point Cloud processing

1 緒言

近年、災害現場での二次災害を減少させるために、頑健に行動できる歩行ロボットの開発が望まれている。災害現場でロボットが人に変わって作業を行うためには、遠隔操作やロボットが自律的に行動する必要がある。未知の環境下でロボットが自律的に行動するためには、3次元地図を構築し、その中から障害物や段差、動的物体などの移動に必要な環境情報を抽出する環境センシングを行う必要がある。このような3次元地図構築技術や環境認識技術において、近年、LIDARなどの3次元測距センサを用いた研究が盛んに行われている[1,2]。災害現場では、ロボットは、屋外環境では家屋の倒壊や地割れ、地面の陥没などにより道路上に平地領域と瓦礫領域が混在している中を行動する必要がある。また、屋内環境では階段や梯子などを昇降する必要があり、さらに、どちらの環境においても他のロボットや人などの動的物体を回避する必要がある。このように、災害現場では様々な環境が混在した非常に複雑な環境の中を行動しなければならない。このような複雑な環境下での環境センシングでは、ロボットの置かれている環境や、移動方法などによって計測したい領域や範囲、必要な計測密度などが異なる。そのため、一度の計測のみで環境センシングに十分な地図を構築するのは困難であり、地図を常に更新する必要がある。しかしながら、従来の単一のLIDAR等を使用した計測手法では地図全体を更新するため、計測に時間がかかることや、データ量が膨大になり、SLAMなどの計算コストが増大してしまうといった問題が存在する。そこで、我々は計測領域や計測密度を変化させることのできる3次元距離計測システムとして、複数のLaser Range Finder (LRF) を使用した可変型測域センサアレイを開発し[3]、本センサを用いた環境認識技術について提案を行ってきた[4,5,6,7]。そこで、本稿では取得した環境情報から、複数の基本行動の重みを動的に更新することで滑らかな動作を実現する手法である多目的行動調停[8]の考え方をを用いて、ロボットの置かれている環境や移動方法に適した計測を行うために、可変型測域センサアレイの計測範囲、計測密度の制御を行うための方法論を提案する。

2 可変型測域センサアレイ

2.1 可変型測域センサアレイの概要

本研究において開発を行った可変型測域センサアレイを Fig.1 に示す。4つのLRFと2つのサーボモータから構成されている。LRFは上部の2つにUST-10LX(北陽電機株式会社)、側面の2つにUST-20LX(同社)を搭載しており、LRFを回転させるアクチュエータとして、FHA8C(オリエンタルモーター株式会社)を2つ搭載している。LRFおよびアクチュエータの仕様を Table 1, Table 2 に示す。

本センサアレイはモータの回転軸とLRFの光軸を一致させることによって最低限の座標変換により3次元復元が可能なセンサアレイとなっている。また、上部と側面に設置されたLRFの取り付け角度を変更していることにより、上部のLRFでは周囲全体を満遍なく計測することが可能で、側部のLRFではロボットの前方方向をより密に計測することが可能になっている。本センサアレイはサーボモータの回転速度を制御することで計測密度を、回転範囲を制御することで計測範囲を変更することが可能なセンサアレイシステムとなっており、本稿では認識した環境に合わせ、サーボモータの回転範囲を知的に制御することを目的とする。

2.2 ODEによるセンサアレイとロボットのモデル

本稿では、オープンソースの物理演算エンジンであるOpen Dynamics Engine (ODE)を用いて、Fig.2のような4脚ロボットとセンサアレイのモデルを作成し、シミュレーション環境上において制御実験を行う。また、シミュレーション環境内のセンサアレイで計測された距離データは実際のLRFと同じ形式でデータを保存することが可能になっており、Fig.3のように3次元地図に復元することが可能となっている。

Table 1: Specification of LRF^[9]

型 式	UST-10LX	UST-20LX
検出距離及び検出体	検出保証値：0.06m~10m (白ケント紙) ：0.06m~4m (反射率10%時) 最大検出距離：30m (出力限界値)	検出保証値：0.06m~20m (白ケント紙) ：0.06m~8m (反射率10%時) 最大検出距離：60m (出力限界値)
測距精度	±40mm ^{s1}	
走査角度	270度	
走査時間	25msec (モータ回転数 2400rpm)	
角度分解能	0.25度	
インターフェース	Ethernet 100BASE-TX	

※1.精度は出荷時の基準環境下による基準反射板(白ケント紙)による精度です。

Table 2: Specification of FHA-8C^[10]

減速比	50
瞬間最大トルク[N・m/A]	3.3
最大回転速度[r/min]	120
トルク定数	1.3
一方向位置決め精度[sec]	120
検出分解能	400000
エンコーダ形式	アブソリュート

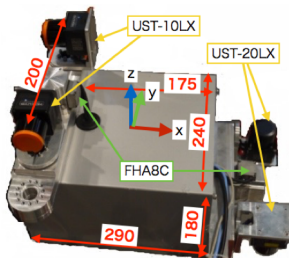


Fig. 1: Variable Sokuiki Sensor Array.

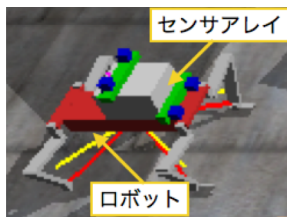


Fig. 2: Robot and Variable Sokuiki Sensor Array Model of Simulation Environment

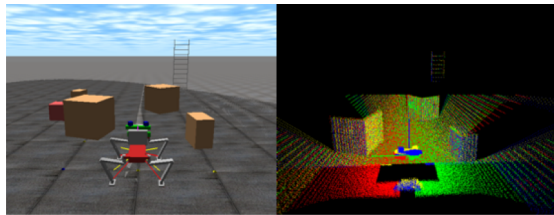


Fig. 3: 3D-map of Simulation Environment.

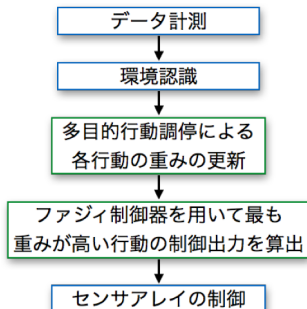


Fig. 4: 3D-map of Simulation Environment.

3 センサアレイの知的制御

本章では多目的行動調停の考え方をを用いて、知覚したロボット周囲の環境に適したサーボモータの回転速度 ω 及び回転範囲 $\theta_{min}, \theta_{max}$ を出力することで計測範囲を変化させ、可変型測域センサアレイを知的に制御する手法を提案する。全体のアルゴリズムのフローチャートを Fig.4 に示す。環境認識に関する手法としては特徴点抽出[4],梯子検出[5],動的物体検出[6]を提案してきた。本稿における多目的行動調停の基本行動は次の3つの計測手法 1.平地歩行計測, 2.動的物体追従計測, 3.梯子計測と定義し、それぞれの制御出力はファジィ制御を用いて算出する。また、従来手法では取得した環境入力情報に応じて多目的行動調停により、各行動の重み付けを行い統合的な出力を算出していた。しかしながら、本稿の場合は、例えば、平地歩行計測と動的物体追従計測の重みが大きくなった際に、その2つの物体の位置が大きく離れていた場合を考えると、統合的な出力では2つの物体の中間的な位置を計測するような出力が算出されてしまうことで、どちらの物体も計測できなくなってしまうことが考えられる。そこで、本稿では更新された重みの中で一番高い重みになった行動がロボットにとって現在の環境で一番必要な計測であると考え、ファジィ制御により算出されたその計測手法の出力をそのまま最終的な制御出力として選択し、センサアレイの制御を行う。

3.1 ファジィ制御による各行動の出力決定

本稿における基本行動は簡略化ファジィ推論を用いたファジィ制御[6,7]によって算出する。各行動における前件部の入力情報は以下の通りとする。

1.平地歩行計測：

平地を歩行する際に使用し、障害物に対する局所的計測と自己位置推定、動的物体検出を行うために、センサ側面に搭載されている LRF は固定し、センサ上部の2つの LRF をロボットから一番近い障害物に対し局所的に計測を行うように制御する。前件部の入力情報としてはロボットに一番近い障害物までの距離、方向、大きさとする。

2.動的物体追従計測：

動的物体を検出した際に、動的物体の挙動を確認するために、センサ上部に搭載された LRF を、動的物体を追従するように計測を行えるように制御を行う。前件部の入力情報としては動的物体の位置、方向、大きさ、移動速度、移動方向とする。

3.梯子計測：

梯子を登るために、梯子の詳細なデータを取得するための計測を行うために、全ての LRF を梯子周囲の範囲に対し局所的に計測を行うように制御する。前件部の入力情報としては梯子の位置、横幅とする。

3.2 多目的行動調停

ここでは、環境認識によって得られた環境知覚情報をもとに、多目的行動調停を用いて各行動の重みを算出する手法について述べる。行動 k の時の制御出力を y_{kj} 、重みを wgt_k 、行動の総数を r ($=3$) とすると、重み wgt_k は知覚した時系列センサ情報に従い、(1)式により更新される。ここで、 α は忘却係数であり、センサ情報が初期値と同値の場合に累積の重み更新幅が 0 になることを防ぐ目的がある。

$$wgt_k \leftarrow \frac{\alpha \cdot wgt_k + \Delta wgt_k}{\sum_{k=1}^r (\alpha \cdot wgt_k + \Delta wgt_k)} \quad (1)$$

ここで、 Δwgt_k は重み更新則であり、次のように設定される。

$$\begin{bmatrix} \Delta wgt_1 \\ \Delta wgt_2 \\ \Delta wgt_3 \end{bmatrix} = \begin{bmatrix} dw_{1,1} & \cdots & dw_{1,5} \\ \vdots & \ddots & \vdots \\ dw_{3,1} & \cdots & dw_{3,5} \end{bmatrix} \begin{bmatrix} si_1 \\ \vdots \\ si_5 \end{bmatrix} \quad (2)$$

si_m は環境知覚情報を表し、入力された情報に応じて 0~1 の値をとる。本稿では検出された障害物や動的物体、梯子の位置情報とロボットの自己位置の情報をもとに以下の 5 つを定義する。 si_1 は障害物に対する危険度を表しており、ロボットとロボットに最も近い障害物との距離 l_o が $L_o = 1.2[m]$ 以内になった際に急激に上昇するように設定する。 γ_o, η_o は危険度の上昇の勾配を決定する任意の定数であり、本稿では $\gamma_o = 1.0, \eta_o = 2.0$ とした。 si_2 は si_1 と同様に動的物体に対する危険度を表しており、 l_m はロボットと動的物体との距離、 L_m, γ_m, η_m はそれぞれ任意の定数であり、本稿では $L_m = 1.5[m], \gamma_m = 1.0, \eta_m = 2.0$ とした。 si_3 も動的物体に対する危険度であり、一回前の計測により取得された l_m を l'_m とし、 $l'_m > l_m$ のとき、すなわち動的物体がロボットに近づいてきている場合は 1 を、離れていく場合は 0 を返す。 si_4 はロボットと梯子までの距離 l_l が $L_l = 1.0[m]$ 以内になった際に、急激に梯子計測重みが大きくなるように設定する。

$$si_1 = \begin{cases} \exp(\gamma_o \cdot (L_o - l_o) / L_o) / \eta_o & (\text{if } L_o > l_o) \\ 1.0 & (\text{otherwise}) \end{cases} \quad (3)$$

$$si_2 = \begin{cases} \exp(\gamma_m \cdot (L_m - l_m) / L_m) / \eta_m & (\text{if } L_m > l_m) \\ 1.0 & (\text{otherwise}) \end{cases} \quad (4)$$

$$si_3 = \begin{cases} 1.0 & (\text{if } l'_m > l_m) \\ 0.0 & (\text{otherwise}) \end{cases} \quad (5)$$

$$si_4 = \begin{cases} 1.0 & (\text{if } L_l > l_l) \\ 0.0 & (\text{otherwise}) \end{cases} \quad (6)$$

$$si_5 = 1 - si_2 \quad (7)$$

また、 $dw_{k,m}$ は行動 k が si_m によって受ける影響を定める定数パラメータであり、以下のように設定した。

$$\begin{aligned} dw_1 &= \{0.6, 0.0, 0.0, 0.0, 0.4\} \\ dw_2 &= \{0.0, 0.7, 0.2, 0.0, 0.0\} \\ dw_3 &= \{0.0, 0.0, 0.0, 0.7, 0.2\} \end{aligned} \quad (8)$$

これらを用いて各行動の重みを算出し、重みが最大となった行動のファジィ制御器によって算出された出力をもとにセンサアレイの制御を行う。

4 シミュレーション実験

本実験では、提案手法により環境に適した計測が行われていることを確認するため、シミュレーション環境上に置いて、センサアレイの制御実験を行った。前述のシミュレーション環境上に Fig. 5 のように障害物、動的物体、梯子を設置した実験環境を構築した。本実験では、ロボットから見て x 方向に 7[m]、 y 方向に -1[m] の位置に設置した梯子を目標位置としてロボットを歩行させ、ロボット周囲の環境に応じて最適な計測を行えることを確認する。また、本稿ではロボットの自己位置及び障害物、動的物体、梯子の位置は既知の情報として扱う。実験におけるロボットの歩行の様子を Fig. 6 に、その時の多目的行動調停により算出された各行動の重みの推移を Fig. 7 に示す。実験結果より、スタート地点では周囲に障害物しかないため、平地歩行計測の重みが最大になっているが、1.4[m] 手前の位置からは、障害物の後ろにある動的物体が見えたことにより、動的物体計測の重みが増していることがわかる。その後、1.6[m] の位置で動的物体が通り過ぎるのを待機している間は動的物体が通り過ぎていくにつれて動的物体計測の重みが減少し、平地歩行計測の重みが再び上昇している。1.7[m] 付近で再び動的物体計測の重みが増しているが、これはロボットが移動を開始したことにより動的物体との距離が近づき、 si_3 が影響したことによるものだと考えられる。また、6.0[m] 付近からは梯子との距離が 1.0[m] 以内になったことで、梯子計測の重みが増し、最大になっていることがわかる。これらの結果から、本提案手法を用いて、ロボット周囲の環境に適した計測が行えていることがわかる。

5 まとめ

本稿では、災害現場などの複雑な環境下において、ロボットの歩行や作業に必要な情報を取得する環境センシングを行うために、ロボット周囲の環境や、ロボットの行動計画に合わせ、可変型測域センサアレイの計測範囲を知的に制御する手法について提案した。また、シミュレーション環境上で障害物や動的物体、梯子などを混在させた環境を構築し、ロボットを移動させセンサアレイの制御を行いながら環境計測を行ったことで本手法の有用性を示した。本研究では、本稿では多目的行動調停における各種パラメータや、どのような入力情報を与えると、どの行動が選択されるかを決定する環境知覚情報 si を経験的に定めた。しかし、

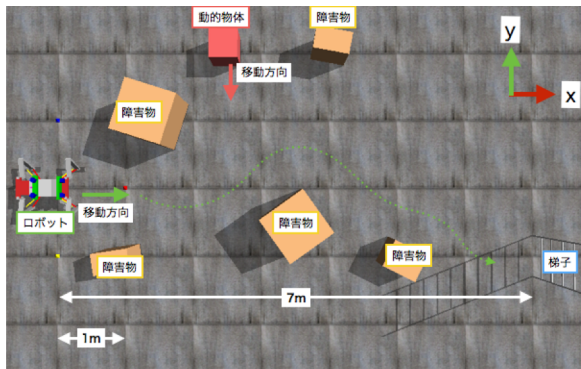


Fig. 5: Experiment of Simulation Environment.

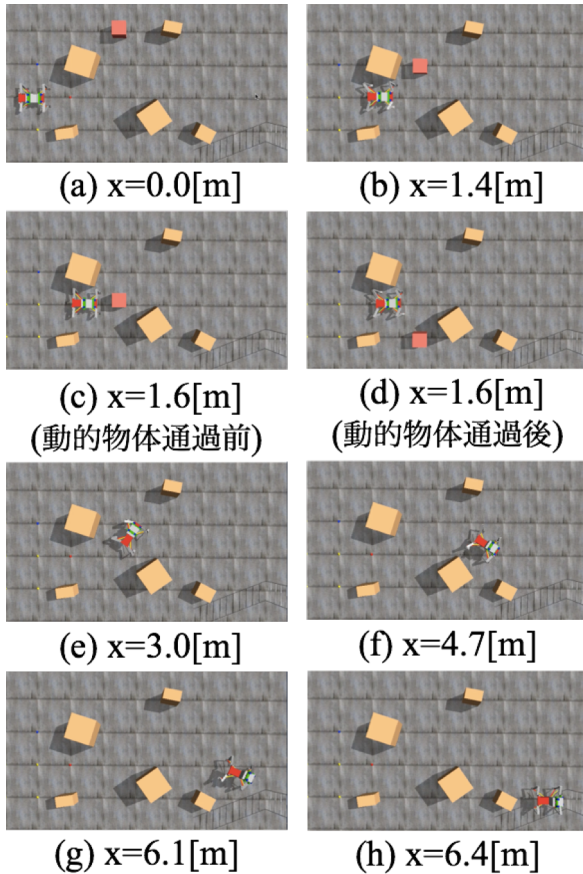


Fig. 6: State of robot locomotion in experiment

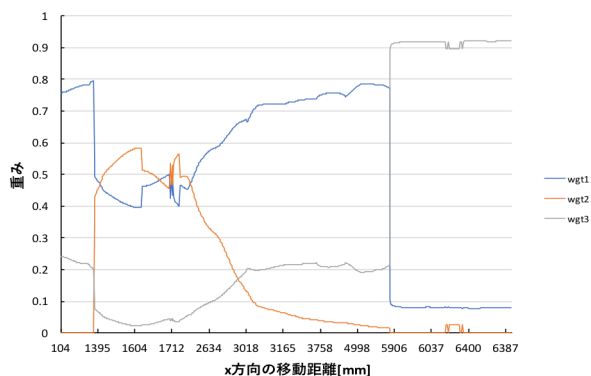


Fig. 7: Experimental result of transition of weights

実際の災害現場などではより様々な環境が考えられ、さらにロボットの作業も多岐に渡ると考えられる。従って、パラメータ設定や環境知覚情報を人が設計するのは困難になると考えられる。

そこで、強化学習などの教師なし学習を用いて、未知の環境下でも最適な制御出力を学習できるようにしていきたい。

6 謝辞

本研究は、総合科学技術・イノベーション会議により制度設計された革新的研究開発推進プログラム (ImPACT)により、科学技術振興機構を通して委託されたものです。

参考文献

- 1) 石川繁樹, 浅香俊一: 移動障害物を含むような動的変化を伴う走行環境における自律移動ロボットの走行誘導方式, 日本ロボット学会誌, 11.6, 856/867,(1993).
- 2) Kita Shohei, Jae Hoon Lee, and Shingo Okamoto: 3D Map Building Using Mobile Robot with Scan Device., International Journal of Mechanical Engineering and Robotics Research(IJMERR), Vol. 5, No. 3, 191/195 (2016)
- 3) 武居直行, 戸田雄一郎, 北井瑛佳, 和田一義, 久保田直行: 可変型測域センサアレイの機構設計, 第27回インテリジェント・システム・シンポジウム (FAN2017 in OKAYAMA), 1B2-2, 61/62 (2017)
- 4) 戸田雄一郎, 北井瑛佳, 武居直行, 和田一義, 久保田直行: 物体把持のための3次元点群の位相構造の学習と特徴抽出, 第27回インテリジェント・システム・シンポジウム (FAN2017 in OKAYAMA), 1B2-1, 57/60 (2017)
- 5) 戸田雄一郎, 北井瑛佳, 武居直行, 和田一義, 久保田直行: 測域センサアレイを用いた DYNAMIC TIME WARPING に基づく3次元点群からの特徴点抽出, 第22回ロボティクスシンポジウム, 2B5, 105/110 (2017)
- 6) 北井瑛佳, 戸田雄一郎, 武居直行, 和田一義, 久保田直行: 可変型測域センサアレイによる動的物体の追従制御, 日本ロボット学会学術講演会, RSJ2017,AC3C1-05, (2017)
- 7) 北井瑛佳, 戸田雄一郎, 武居直行, 和田一義, 久保田直行: 環境センシングのための可変型測域センサアレイの知的制御, ロボティクス・メカトロニクス講演会 Robomec2017 予稿集 1P1-Q06 (2017)
- 8) 能島裕介, 小島史男, 久保田直行: 多目的行動調停に基づく移動ロボットの行動獲得, 日本機械学会論文集(C編), 68巻671号, No.01-1066, 141/147 (2002)
- 9) 北陽電機株式会社, <http://www.hokuyo-aut.co.jp> (2016/12/15).
- 10) (株)ハーモニックドライブシステムズ, <https://www.hds.co.jp> (2016/12/15).

MA を用いた絵画的迷路の自動生成

○今宮明則 小野典彦 永田裕一 (徳島大学)

MA for Automatic Picturesque Maze Generation

* A. Imamiya, N. Ono, Y. Nagata (Tokushima University)

Abstract— Picturesque maze is a category of maze, whose solution path shows up as a picture. When you reach from the entrance to the exit without mistake, you can draw a picture. In this paper, we discuss a method to automatically generate a solution path close to a given binary image. Although a simulated annealing algorithm has already been proposed to this problem, this paper proposes a memetic algorithm. To apply memetic algorithm to this problem, we develop a crossover operator suitable to this problem by modifying edge assembly crossover for the TSP.

Key Words: Picturesque Maze, Memetic Algorithm

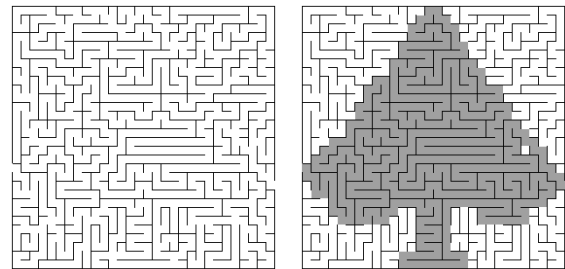
1 はじめに

迷路は老若男女が楽しめる古典的パズルの一つであり、これまでに様々な種類のものが多く作成されてきた。絵画的迷路は迷路の一種で、入口から出口まで正解路部分のみを通り、その経路を塗りつぶすことで何らかの絵や記号が完成するというものである(Fig.1)。絵画的迷路はパズルを解く楽しさだけでなく、解いた後に絵が完成するという芸術的な楽しさも持ち合わせており、特に子供を中心として人気がある。

絵画的迷路を扱う本にある迷路は、生成するアルゴリズムを公開していないか、自動的に迷路を生成したものではないとされており、学術的な研究も少ない。2009年に岡本らが絵画的迷路の構成法について提案してから、いくつかの論文が発表されているが¹⁾³⁾⁴⁾、新しい分野としての発展の余地は多く残されている。

絵画的迷路の生成は正解路が絵にならなければならないという制約のため、満足のいく迷路を作成することが難しい。また良い迷路かどうかの評価は、絵の完成度、経路の複雑さ、解いた時の達成感などがかわってくる。このような条件の問題はコンピュータで自動生成するのは難しく、人の手作業によって作成されるものが多い。この問題をコンピュータによって解決することができれば、絵画的迷路作成にかかる時間の大幅な削減を見込むことができる。

先行研究としてSA(Simulated Annealing)を用いた手法¹⁾が考案されており、絵と迷路の多少の食い違いを許容することで迷路の生成を実現している。絵画的迷路の自動生成に対してGA(Genetic Algorithm)を適応するのは適切な交叉方法を考案するのが難しく、開発が難しいと考えられてきた。一方GAは経路最適化問題やゲーム情報学の分野で非常に目覚ましい成果をもたらしており、人間が手掛けるよりも良い性能のものを



Problem

Answer

Fig. 1: Example of a picturesque maze.

作り出す可能性を秘めている。絵画的迷路にもGAを適応できれば従来手法よりも良い迷路を生成することが期待でき、新しい切り口を開くことになる。本論文ではGAと局所探索を組み合わせたMA(Memetic Algorithm)を用いることで既存手法より食い違いの少ない正解路を生成することを目的としている。

2 絵画的迷路構成問題

迷路には非常に多くの種類のものが存在し、次元数や外観などいくつかの観点で分類される。本論文では二次元格子状の全域木型迷路を対象としており、以下にその特徴をまとめる。

- 横 C_{wd} マス、縦 C_{hg} マスの長方形格子状の迷路である。
- 各マスは上下左右最大4マスと接する。その間に壁がある場合は直接移動出来ない。壁がない場合は直接移動出来る。
- 外周部マスは迷路の外部に進めないよう壁で隔てられている。ただし、入口と出口が1つずつ存在し、その部分だけ外部とつながっている。
- 正解路は重複しないマスの配列 $P = \{p_1, \dots, p_n\}$ で定義される。 p_1 は入口マス、 p_n は出口マスであり、すべての p_i と p_{i+1} ($i = 1, 2, \dots, n-1$)は壁のない隣り合


```

procedure SA()
1 :初期解 $P_{start}$ の生成
2 :  $T_{start}, T_{decay}, T_{end}$ の設定
3 :  $T_{now} := T_{start}; P := P_{start};$ 
4 : REPEAT
5 : 近傍解 $P'$ の作成
6 : IF  $(F(P') - F(P)) < 0$  THEN
7 :  $x := x'$ ;
8 : ELSE
9 :  $\exp(- (F(P') - F(P)) / T_{now})$ の確率で
 $P := P'$ ;
10: ENDIF
11:  $T_{now} := T_{now} \times T_{decay};$ 
12: UNTIL  $T_{now} < T_{end}$ 
13: RETURN with 最良解

```

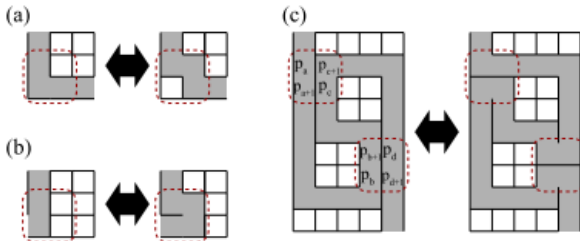


Fig. 4: Method of generating neighbor solution of SA.¹⁾
池田, 橋本(2012) 『確率的最適化を用いた絵画的迷路生成』
p1629 図5から引用

ランダムに生成したものを使用する。SAの特徴といえるLine9は一定確率で解を改悪する場合でも解を更新する。また使用する近傍オペレータは3種類あり、それらをFig.4に示す。(a)は正解路をずらす機能、(b)は正解路を増減させる機能、(c)は正解路の経路順を入れ替える機能を持っている。(c)の近傍オペレータを使用しても評価値が変わることはないが、解探索の自由度が増すため結果として性能も向上する。

このSAはある程度の時間で食い違いピクセルを許容した迷路を生成することができ、絵の大きさが変化しても対応できる。しかしながら複雑な絵で迷路を生成しようとする局所解に陥ってしまう恐れがあり、絵と大きく食い違う迷路が生成されることがある。詳しくは「確率的最適化を用いた絵画的迷路生成」¹⁾を参考にされたい。

絵の一致する量だけ比べれば3.1の手法を用いて迷路を生成したほうが良いように見えるが、手法の性質

```

Procedure MA( $N_{pop}, N_{ch}$ )
1 :初期集団 $\{x_1, x_2, \dots, x_{N_{pop}}\}$ の生成
2 : REPEAT
3 :  $i \in \{1, \dots, N_{pop}\}$ をランダムに集団に割り振る;
4 : FOR  $i := 1$  TO  $N_{pop}$  DO
5 :  $p_A := x_i; p_B := x_{i+1}; (x_{N_{pop}+1} = x)$ 
6 :  $\{c_1, c_2, \dots, c_{N_{ch}}\} := EAX(p_A, p_B);$ 
7 :  $x_i := \text{Select\_Best}(c_1, c_2, \dots, c_{N_{ch}});$ 
8 :  $x_i := SA(x_i);$ 
9 : END FOR
10: UNTIL 終了条件が満たされる;
11: RETURN with 集団中の最良個体;

```

により複雑な経路は生成しづらく、単調な経路になりやすい。経路の複雑さや面白さを求めるならば、経路に多様性を持たせ、様々な経路の迷路を生成できた方が良いため、3.2の手法の方が向いている。

本論文では3.2の手法のように食い違いピクセルを許容して、経路の変更を容易にし、多くの経路を生成することで迷路の質の向上を図っている。

4 MAを用いた正解路構成法

本研究で提案するMAでは交叉としてEAX(Edge Assembly Crossover)²⁾を用いる。EAXはTSP(巡回セールスマン問題)を解くために作られた交叉であるが、絵画的迷路構成問題に適用できるように改良する。またEAXは中間個体を生成する前の手順を変えることで、より効果的に良い評価値の経路を見つけられる場合がある。本章ではEAXの概要と中間個体の生成方法について述べる。次に正解路以外の部分の生成方法についても簡単に示す。

4.1 手順

MAを用いるためには初期集団が必要となる。初期集団は3.2で示したSAを用いていくつかの解を生成し構成する。またMAはGAと局所探索を組み合わせる手法であるが、本論文ではMAの局所探索部分にも3.2で示したSAを使用する。ただし、MAでは何度も局所探索を呼び出すため、計算時間が長くなるように、局所探索と近くなるような温度設定でSAを用いる。使用するMAのアルゴリズムをProcedure MAに示す。

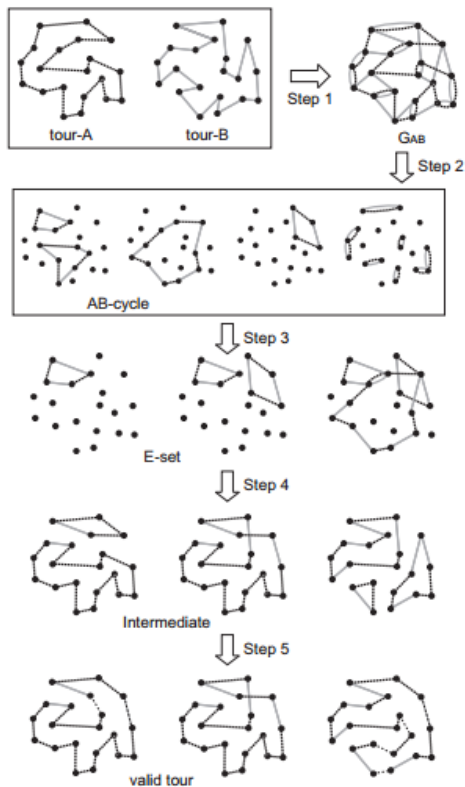


Fig. 5: EAX basic steps.²⁾

永田(2017)『局所的な交叉EAXを用いたGAの高速化とTSPへの適応』p544 図1から引用

procedure MAにおいて、 N_{pop} と N_{ch} はそれぞれ集団サイズ、各親から生成される子個体の数を示している。Line1では先に述べたようにSAによって初期集団が生成される。Line6ではEAXを用いて各両親 p_A , p_B から N_{ch} 個の子個体が生成される。Line7では N_{ch} 個の生成子個体の中から最も評価値の良い個体を選択し、集団中の x_i と置き換える。ただし生成した子個体のすべてが p_A よりも悪かった場合、置き換えは行われず、 x_i がそのまま残るものとする。Line8では局所探索としてSAを用いて近傍解を生成する。ただしLine7で x_i が置き換えられなかった場合はこの処理を行わない。

4.2 絵画的迷路構成問題にEAXを適応

EAXはTSP用の交叉として開発され、各都市間の距離を短縮し、コストを削減するために使われる。しかし迷路生成の場合、短い経路の迷路を作ることに意味はなく、絵に近い迷路を作成する必要がある。そこでEAXの評価値を経路の長さではなく、食い違いピクセルの数で処理させることで、経路の大域的な変更、改善を見込むことができる。

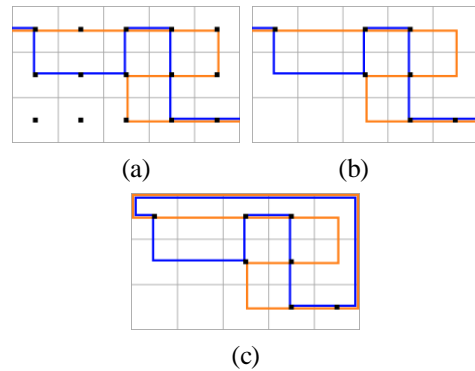


Fig. 6: Example of route substitution.

4.2.1 EAXの概要

TSPで使用されるEAXはFig.5のような手順で行われる。tour-Aとtour-Bは親となる2つの順回路を表し、以下の5つのステップで子個体を生成する。

Step1: tour-Aとtour-Bの経路を全て組み合わせたグラフ G_{AB} を作成する。

Step2: G_{AB} をtour-Aの枝、tour-Bの枝、tour-Aの枝…と交互に辿っていき出来上がった閉路をAB-cycleとして記憶する。

Step3: 記憶したAB-cycleの中から使用する閉路を何らかの手順で取り出しE-setとする。

Step4: E-setに含まれるtour-Bの枝とE-setに含まれないtour-Aの枝を組み合わせて中間個体を生成する。

Step5: 中間個体が2つ以上の閉路であった場合1つの閉路になるまで近い閉路の枝を付け替える。出来上がった1つの閉路が生成される子個体になる。

EAXがTSPで非常に良い結果を出している理由の一つに、Step5で閉路を統合している際、局所探索のように枝を付け替えていることが挙げられる。つまり交叉しながら局所探索ができるため、MAのように別途局所探索を組み合わなくても、MAと同等の結果を得られるのである。

4.2.2 TSPと絵画的迷路構成問題の違い

TSPの解はすべての点(都市)を通る閉路で表される。EAXは交叉前の経路がこれらの条件(i)すべての点を通る、(ii)閉路である)を満たすことを前提として作られている。迷路では点が存在しないが、1マスの中心を点とみなすことでTSPの点(都市)のように扱うことができる。しかし絵画的迷路の解は個体ごとに通る点異なるため、(i)を満たさない。また入口と出口が一致しないため閉路が生成されることはなく、(ii)を満たさない。このためEAXをそのまま絵画的迷路の生成に用

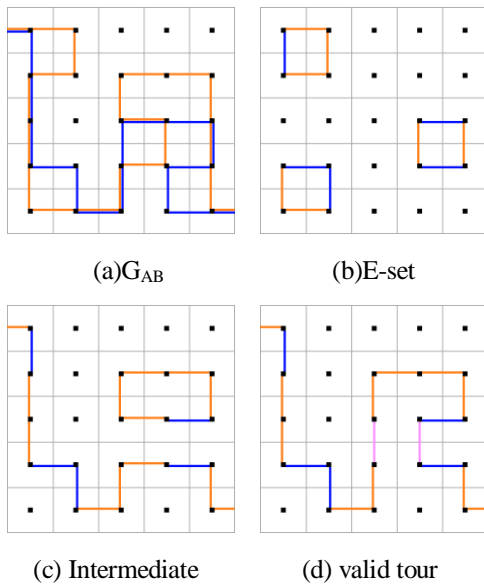


Fig. 7: Example of crossover in the maze.

いることはできない。そこでEAXを適用できるようにするために、絵画的迷路の経路を条件(i)(ii)を満たすような経路に置き換える操作を導入する。

- ・条件(i)を満たす方法

Fig.6(a)のような G_{AB} があったとする。tour-Aとtour-Bはそれぞれ通る点異なるため、条件(i)を満たしていない。そこで両親の経路が共に通る点以外を全て削除し、片方の親の連続する(削除されない2点間の)経路を1つの枝とみなす(Fig.6(b))。この処理によりtour-A、tour-Bの経路はどちらも同じ点を通るようになるため、条件(i)を満たすことができる。

- ・条件(ii)を満たす方法

Fig.6(b)の経路はtour-A、tour-Bのどちらも経路が途切れているため、閉路ではない。そこで入口と出口を繋ぐ経路を付け足す(Fig.6(c))。これにより条件(ii)を満たすようになる。

Fig.6(c)の経路は条件(i)(ii)を満たすため、TSPと同様に扱うことができる。したがってEAXも適用できるようになる。

また先に述べたようにTSPでは距離が評価関数であったが、絵画的迷路構成問題では食い違いピクセル数に変更する。すなわち、Fig.5のStep5では隣接する閉路の枝を付け替えるため評価値が変化しなくなる。迷路でEAXを適応した一例をFig.7に示す。

迷路生成でEAXを使う利点はFig.7の(c)(d)にあり、中間個体の枝を付け替えることによって、両親が通らない経路を含む子個体を生成できることにあり。これにより局所探索の近傍オペレータが使用できる経路の

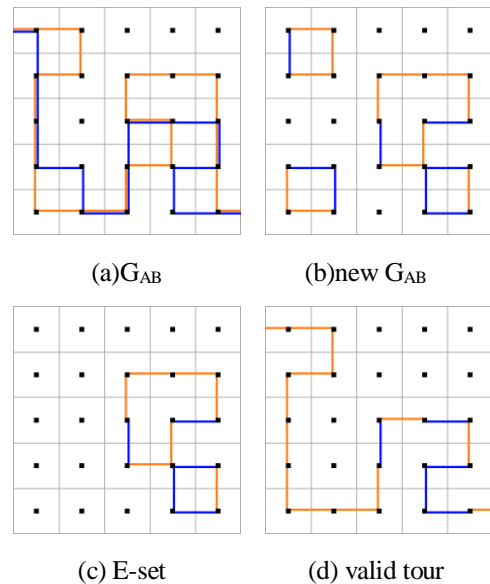


Fig. 8: Example of route exclusion.

数が増え、探索範囲を広げることができる。また付け替える枝をある程度操作できるので、ワープするマスが存在しないという制約条件を常に満たしながら交叉することができる。

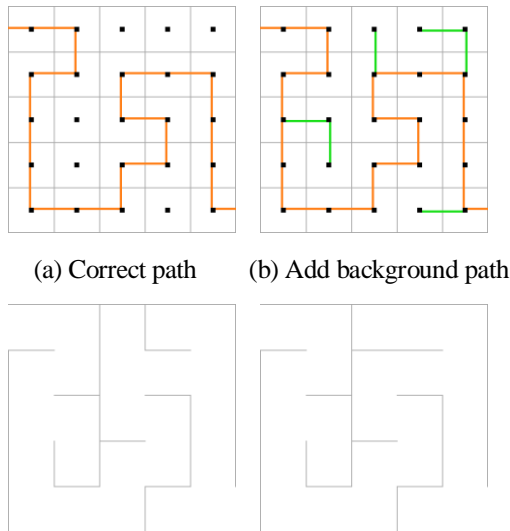
4.3 EAXの工夫

これまでに示した方法で正解路部分は作成することができるが、4.2.1で述べたStep1とStep3を迷路の性質を考慮して改良することで、より食い違いの少ない迷路を生成できる可能性がある。

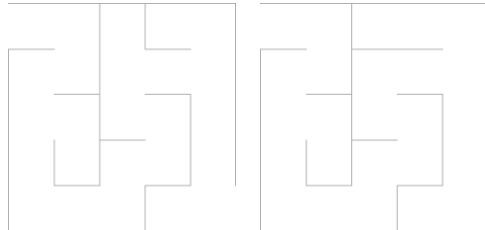
<Step1で共通経路を除外>

4.2.1で述べたように通常のEAXはStep1で両親の経路を全て使用して G_{AB} を作成する。しかしながら迷路生成でEAXを実行した場合評価値が変わる可能性があるのは、どちらか一方の親しか通らない部分をE-setにしたときだけである。また迷路生成ではStep5を行う際自由に枝を付け替えることができず、隣接する閉路の枝しか付け替えられないため、複雑なE-setが生成されると一本の経路にできない場合がある。そこでFig.8のように G_{AB} を作成した後 G_{AB} の中から両親とも通る経路をあらかじめ除外し(b)のような G_{AB} を作成する。その後の操作は同じように行うことで、どちらか一方の親しか通らない経路を含むE-setを使用することができる。

G_{AB} の経路を除外することで生成される子個体の多様性が低くなるというデメリットも存在する。メリットとどちらが大きいのかの実験が必要である。



(a) Correct path (b) Add background path



(c) Maze to be created (d) Integrated maze

Fig. 9: Generation method of background path.

<Step3でE-set数を変化>

Fig.5のE-setの部分で示されているようにE-setの数は変えることができる。E-setの数が少ないと局所的な交叉となり出来上がる子個体はtour-Aに近いものになる。またE-setの数が多くと大域的な交叉となり子個体はtour-Bに近いものになる。すべてのAB-cycleをE-setにすると出来上がる子個体はtour-Bに一致してしまうため使用する数は多くできないが、どの程度の数がいいのか実験する必要がある。これはTSP用のEAXでも実験されており、評価値が変化しなくなりだした頃にE-set数を増やすと良い解が見つかりやすくとされている。本論文の実験では初期解の時点である程度良い評価値の経路を使用しているため、E-set数を固定して比較を行う。

4.4 背景路部分の生成

背景路部分の生成は既存の手法¹³⁾で述べられているものを用いる。Fig.9にその例を示す。(a)のような正解路が生成されていた時、正解路を木、不正解部分のマスをもノードとして見る。これらを繋ぐ全域木を生成したものが(b)である。(b)で加えた背景路部分は複数考えられるが、任意のものを使用してよい。これを元にして迷路を生成したものが(c)である。(c)のままだも迷路としての条件は満たしているが、迷路の規模が大きくなると行き止まりの数が増えるため、難易度が下がりパズルとしての面白さに欠ける。そこで隣接する行き止まりを統合することで数を減らし難易度を上げていく。

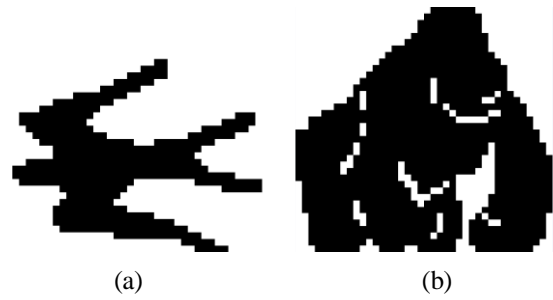


Fig. 10: Image to use.

減衰率	平均	最良
0.99995	6.90	4
0.9995	15.80	8

Table. 1: Evaluation value of (a) using SA.

減衰率	平均	最良
0.99995	36.20	28
0.9995	60.55	42

Table. 2: Evaluation value of (b) using SA.

不正解部分は正解路部分が完成した後から付け加えるので、正解路と不正解部分は独立して考えることができる。

5 実験と考察

本章ではSAのみを用いた実験、MAを用いた実験で、それぞれどのように評価値が推移したかを示す。SAのみの実験は本論文の本題ではないが、ベースライン性能としてSAのみでどの程度の評価値の迷路が生成されるかを確認しておくことは非常に重要だと考えられる。

5.1 SAのみを用いた実験

基本的なパラメータ設定や使用する画像は3.2で述べた手法のものをそのまま利用する。以下にそれらの具体的な内容を示す。

- 開始温度 $T_{start} = 10.0$
- 減衰率 $T_{decay} = 0.9995, 0.99995$
- 最終温度 $T_{end} = 0.1$
- 評価値 $f = \sum(\text{食い違いピクセル} \times \text{重み})$
- 使用する画像 Fig.10((a)40×30ピクセル, (b)40×38ピクセル)
- 重み $w = 1.0, 2.0, 100.0$ (黒ピクセルの重みを1, 黒ピクセルに接する白ピクセルの重みを2.0, 接しない白ピクセルの重みを100.0)

この設定で減衰率を変更しどのような評価値、迷路が生成されるかを確かめる。Table.1, Table.2は20回実行した減衰率ごとの平均評価値、最良評価値を示したものである。これらから減衰率を高くするほど良い評価値の迷路が作られやすいことが分かる。これよりも減衰率を上げれば平均評価値は良くなるかもしれないが、最良評価値に関してはこの辺りが頭打ちであると考えられる。また、減衰率を0.9995にした場合の迷路一つ当たりの生成時間が一秒程度であるのに対して、減衰率を0.99995にすると迷路一つ当たりの生成時間が十数秒かかってしまう。従って多くの迷路を生成したい場合は減衰率を下げる必要がある。

これらの表はあくまで一例であり、全く同じ条件でも多少結果が食い違うことがある。これは確率的に経路が変更されるため当然のことであるが、大抵このあたりの結果になる。

5.2 MAを用いた実験

初期解生成は SA を用いて行う。減衰率は 5.1 で結果の良かった 0.99995 を使用する。また結果の比較をしやすくするため使用する初期集団は試行回数ごとに同じものを使用する。交叉に関しては集団サイズを大きくするほど初期解を多く生成しなければならず、一世代あたりの処理時間も大きくなるので、集団サイズをあまり大きくできない。

以上を踏まえて各パラメータは以下のように設定した。

初期解

- ・ 5.1 と同様 (減衰率は 0.99995 とする)

交叉

- ・ 集団サイズ 10
- ・ 世代数 500
- ・ E-set を構成する AB-cycle の数 1~3
- ・ 子個体生成数 10

Table.3, Table.4 は MA をそれぞれ 5 回ずつ行い、500 世代目の集団内の平均評価値の平均、集団内の最良評価値の平均、最良評価値を示している。また E-set 数 1, 2, 3 の時はの共通経路の除外を行っておらず、経路除外の時は E-set 数を 1 で実験した。

Table.1~4 より、SA のみよりも MA を使用した方が良い評価値の迷路を多く作っていることが分かる。

Table.3 は Fig.10(a) を用いて生成した迷路の評価値だが、設定を変えてもあまり大きな変化は見られなかつ

	平均	最良平均	最良
E-set 数 1	3.96	2.8	2
E-set 数 2	3.88	2.4	2
E-set 数 3	4.08	2.4	2
経路除外	3.28	2.8	2

Table. 3: Evaluation value of (a) using MA.

	平均	最良平均	最良
E-set 数 1	27.10	21.2	20
E-set 数 2	27.24	21.4	18
E-set 数 3	26.92	20.4	18
経路除外	23.14	17.4	15

Table. 4: Evaluation value of (b) using MA.

た。これは世代数を 500 に設定したことに原因があると考えられる。Table.3 で最良評価値を見つけた際の世代数は、経路除外をした時が 50~100 世代、その他の時は 100~150 世代が多かった。したがって Fig.10(a) の画像のような迷路にしやすい図形の場合、世代数を減らしてもよい解の経路を見つけられると考えられる。

Table.4 は Fig.10(b) を用いて生成した迷路の評価値である。Table.4 より E-set 数を変化させても大きな変化は見られなかった。ただし経路除外をした場合の評価値は他に比べてよい経路が多く生成された。これは交叉の際両親の良い部分を使用する可能性が高く、効率よく解を探索できたからだと考えられる。また Fig.10(a) の画像を用いても経路除外をした方がよい解を早く見つけられているので、高速化する場合はこの手法を取り入れた方がよいのだろう。しかし Fig.10(b) の画像を使用した場合、最良評価値の経路を見つけた際の世代数は特に特徴がなかったため、500 世代ではよい解を見つけきれないのかもしれない。

6 おわりに

本論文は絵画的迷路構成問題において MA を使用し、有用である可能性を示した。これは他のパズルや教育などの分野にも応用できると考えられる。しかし複雑な絵(白ピクセルが黒ピクセルの中に複数存在するなど)になった場合、単純な絵に比べると評価値は良くなりづらい。これらを解決できるアルゴリズムを提案できれば、評価値を食い違いピクセル以外に設定する余地も生まれ、面白い迷路を作れるようになるだろう。

また E-set の数を世代数ごとに変化させる、E-set の決め方を意図的なものにするなど、EAX のパラメータは

変更できる点が多くある。これらなどを変化させたときにどのような結果になるか今後実験していきたい。

参考文献

- 1) 池田心, 橋本隼一: 確率的最適化を用いた絵画的迷路生成, 情報処理学会論文誌, Vol.53, No.6, pp.1625-1634 (2012).
- 2) 永田裕一: 局所的な交叉EAXを用いたGAの高速化とTSPへの適応, 人工知能学会論文誌, Vol.22, No.5H, pp.542-552 (2007).
- 3) 岡本吉央, 上原隆平: 絵画的迷路の作り方, 数理解析研究所講究録, Vol.1649, pp.58-65 (2009)
- 4) 中井亮平, 岡本吉央: 絵画的迷路作成アルゴリズムの改善, 数理解析研究所講究録, Vol.1691, pp.162-166 (2010)
- 5) 池田心: 絵画的迷路のある拡張, 組み合わせゲーム・パズルミニプロジェクト第5回ミニ研究集会(2010)
- 6) RPGドット, 追加パック16(動物セット)
<http://www.geocities.co.jp/Milano-Cat/3319/muz/002.html>

車両配送問題における追加注文を考慮した配送計画システムの提案

○織田雄大 小野典彦 永田裕一 (徳島大学)

Proposal of delivery planning system considering additional order in vehicle delivery problem

* Y.oda and N.ono and Y.Nagata (University of Tokushima)

Abstract— We propose a delivery planning system considering additional orders using metaheuristics in the vehicle delivery problem. In order to create a delivery planning system, dynamic scheduling that takes into account the situation of additional orders is required. In order to cope with additional orders, the proposed system first creates a schedule, in which a specified number of vehicles have enough room to deal with the additional orders, the additional orders are then assigned to these vehicles by a rescheduling procedure.

Key Words: Meta-heuristic, Vehicle Routing Problem

1 はじめに

人々の生活において、宅配便などの輸送サービスはなくてはならない存在である。特に近年、ネット通販などの普及により、配送業務の需要はより高くなっている。通販などによる注文増加と、ドライバー不足などの問題が続く状況の中で、効率的な配送計画を立案する配送支援システムの構築が望まれている。

最適化問題の一つである配送経路問題 (VRP: Vehicle Routing Problem) では、特定の拠点を荷物の集積地(デポ)として、与えられた制約を守りながら、総移動コストならびに、配送に必要な車両の台数が適切なものを求める問題である。VRPは様々な業種において配送形態をモデル化できる問題クラスである。そのため適応できる問題は多岐にわたるが、実問題への応用を考えるときには、様々なモデルや細かな条件を考慮しなければならず、最適化をしようとする問題に対してはあらかじめモデル・制約条件を扱う問題を定義しておく必要がある。VRPにおいては、各顧客の座標や、荷物の需要量、配達サービスにおける時間指定などのような時間枠、顧客先での作業時間、配送に使える車両の台数や、最大積載量が挙げられる。

筆者らは、時間枠付き車両配送問題 (VRPTW) をベースに、配送中に追加配送が入ってもスケジュールを変えずに対応できるような最適化を目指す。VRPTWでは、VRPの制約に時間枠制約や積載量制約を加えた問題である。この問題に対して、筆者らは配送業務を行っている企業からデータをお借りして、実際に起こりうる問題のモデル化を行い、

制約条件を加える。また、顧客間での移動時間や移動距離は、地理情報システムを利用することで現実の経路に則した道路情報でのモデリングを行う。

本研究では、モデル化した問題を最適化するシステムを構築するために、VRPで使用されている典型的な近傍を用いた局所探索およびタブー探索を構築する。計算機実験により、構築した最適化システムの妥当性を検証する。

2 問題の定義

筆者らが取り組んでいる問題の基本となる、WRPTWの定義を行う。

2.1 WRPTWの定義

$G=(V, E)$ を $n+1$ の頂点からなる完全有向グラフとする。ここで、 $V=\{0, 1, \dots, n\}$ はノード集合、 E は枝集合であり、各枝には距離が与えられている。ノード0はデポ、それ以外のノード $1, \dots, n$ を顧客の集合、各顧客 v に対し、荷物量 q_v ($q_0=0$)、サービスに要する時間 s_v ($s_0=0$)、配送時間枠 $[e_v, l_v]$ が既知であるとする。VRPTWはデポから出発した複数台の配送車両がすべての顧客に対して荷物を配達してからデポに戻ってくるような配送計画において、最適な配送計画を発見する問題である。また、各車両には共通の容量 Q があり、配送車両はこの容量 Q を超えて荷物を積載することができないという容量制約が存在する。また、各配送車両は時間 e_0 にデポを出発した後、各車両は各顧客に荷物を配達していくが、顧客ごとに配送可能な時間枠が存在して

いるので、この条件を満たしながら、時間 t_0 までにデポに帰ってくる必要があるという時間枠制約も存在する。このとき、配送車両が配送可能な時間枠よりも早く顧客についてしまった場合は、待ち時間が生じるものとする。

2.2 実問題モデリング

VRPTWは様々な業務形態に対応した問題クラスであるため、適応できる実問題は多い。実問題においては状況ごとに細かい条件を考慮しなければならず、起こりうる制約条件をあらかじめ定式化しておく必要がある。ここでは実際の企業の顧客データをもとに必要とされる制約を設定する。以下が指定した制約である。

(1) 複数人作業制約

大がかりな作業を行う場合など、1顧客に対して複数人の作業員がサービスを行う制約である。複数人で作業が必要な場合は、複数台の車両が合流するものとする。また、1台の車両には1人の作業員しか乗っていないものとする。この時に必要作業員数を満たさない場合、条件を満たすまで先に到着した方は待機するものとする。また、複数人が1台に乗車し作業を行う場合、デポと配送先を往復する作業日程になることが多いため、別枠で考えることができる。このため、そのような場合は、今回のスケジュールに含めないものとする。

(2) 追加配送制約

追加配送に確実に対応できるように、あらかじめ指定した車両については確実に指定時間までにはデポに戻ってくることにする。

(3) 休憩制約

各作業員はある時間枠の中で昼休憩をとるものとする。このとき、作業員はどこで休憩してもよいものとするが、顧客先で作業している最中には休憩を取らないものとする。

2.3 地理情報システム

VRPTWでは、移動コストを地点間の直線距離と定義されている。しかし、実際に道路を走行する場合

は必ずしも目的地に対して道路が直線に敷かれているとは限らないため、実問題を解く際には道路に則した移動コスト情報を取得する必要がある。

そこで、デジタル道路地図を提供しているゼンリン社のNAVI-APIを利用することで経路情報を取得する。また、経路探索時に使用する経路の設定を行うことができる。今回は、有料道路を使用しない最短時間での移動経路を探索する設定を用いる。下図のFig. 1では徳島大学から徳島大学病院への移動時間および、移動距離の二つの移動コストを取得している。



Fig.1経路情報取得例(徳島大学から徳島大学病院までの経路)

配送先の顧客リストがあれば顧客間の移動コストをまとめたマトリックスを作成することができ、配送スケジュール作成時に利用することができる。下図のFig.2では移動時間並びに移動距離のマトリックスを確保している。マトリックスを利用することで、移動コストの保存と経路の変更が可能であるという利点がある。移動コストを保存することで経路探索時にコストを計算する必要がなく、経路探索に繰り返し利用可能である。経路の変更については、APIでは見つけられないが、良いルートがあった場合、経路の書き換えが可能である。

	目標地点		
	徳島大学	徳島市役所	鳴門市役所
出発点 徳島大学	0	4	29
徳島市役所	5	0	34
鳴門市役所	28	30	0

	目標地点		
	徳島大学	徳島市役所	鳴門市役所
出発点 徳島大学	0	1674	13295
徳島市役所	2235	0	14748
鳴門市役所	12656	13902	0

Fig2. マトリックス作成例

上: 移動時間(分) 下: 移動距離(m)

Fig. 2の例は, 徳島大学, 徳島市役所, 鳴門市役所の3地点における移動時間と移動距離のマトリクスである. 例えば, 徳島大学を出発し, 鳴門市役所に向かった場合, 移動時間は29分で移動距離は13295mの移動距離が必要となる. 出発点と目的地店が同一の場合, 移動コストは時間, 距離ともに0で表される. また, 往復路で移動コストに違いがある. これは, 実際の道路には一方通行などの制限があるため, こうした事情を考慮した場合, 行きと帰りでは通過する経路が異なることがあるためである.

3 解法的设计

VRPTWはNP困難に属する組み合わせ最適化問題であることが知られている. 本研究では, 配送計画システムの実問題へ対応させることを目指しているため, ある程度の高速性が求められる. したがって, タブー探索を中心としたメタ・ヒューリスティックを用いた近似解を構築する.

3.1 基本フレームワーク

本研究の提案手法では3つのフェーズからなる近似解法を用いる.

(1) 第1フェーズ

ここでは初期解を生成する. 初期解ではまず, 1顧客に対して1台の配送車両が配送するような, 顧客と車両が1対1対応した配送計画を作成する. (Fig. 4の左図)

(2) 第二フェーズ

生成した初期解をもとに経路数の最小化を行うことで, あらかじめ定めた目標のルート数まで削減することを目指す.

(3) 第三フェーズ

総移動コストの最小化を進める. 本研究では移動コストの最小化のほかに追加注文に対応するための余剰を設ける必要もあるため, 十分な余剰を確保することも併せて行う. コスト最小化を行う評価関数については3.3節で説明する. (Fig. 4の右図)

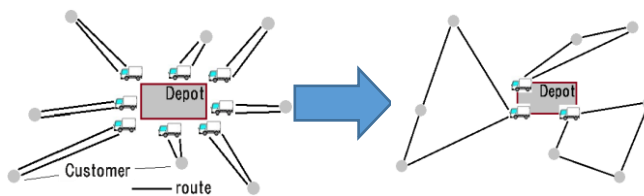


Fig. 3 各フェーズにおける経路図例

3.2 探索手法

本研究ではタブー探索を用いた近似解法を構築する. 近傍の定義として挿入近傍, 交換近傍, 2opt近傍の3つを用いる.

(1) 挿入近傍

挿入は, 顧客と顧客(あるいはデポ)の間に新たに顧客を入れる操作である. 挿入近傍ではルートを削減することを目的としているため, 1顧客にしか配送しない車両のみを対象としている. Fig. 4に挿入近傍の例を示す.

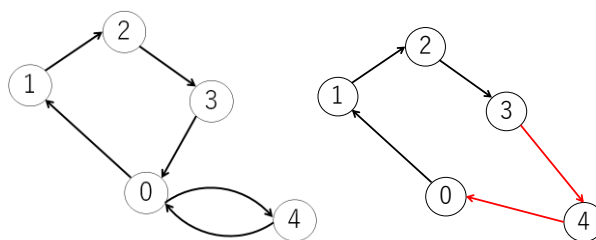


Fig. 4 挿入近傍の例

左図は挿入が行われる前の状態. 右図は挿入が行われた後の状態.

(2) 交換近傍

交換は選択された2顧客を交換する操作である. Fig. 5に交換近傍の例を示す.

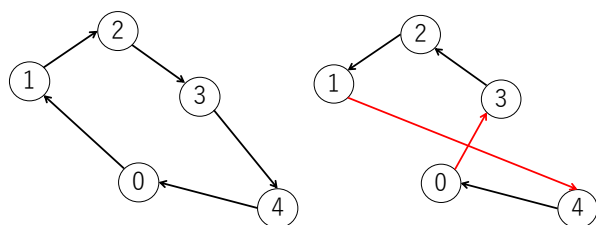


Fig. 5 交換近傍の例

左図は交換が行われる前の状態. 右図は交換が行われた後の状態.

(3) 2opt近傍

2optでは2つのルート間で辺を切った後に繋ぎなおす近傍操作を行う。本研究では2opt近傍を3つのパターンに分けて運用している。

(3.1) 2opt-1近傍

Fig. 6に2opt-1の例を示す。顧客2から3と顧客6から7の間の辺の入れ替えを行っている。

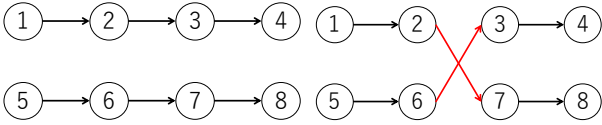


Fig. 6 2opt-1の例

左図は2opt-1が行われる前の状態。右図は2opt-1が行われた後の状態。

(3.2) 2opt-2近傍

Fig. 7に2opt-2の例を示す。2opt-1と同様に顧客2から3と顧客6から7の間の辺の入れ替えを行っているが、繋ぎ変え方が異なる。

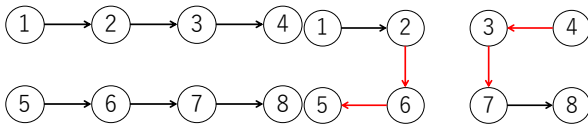


Fig. 7 2opt-2の例

左図は2opt-2が行われる前の状態。右図は2opt-2が行われた後の状態。

(3.3) ルート削減近傍

ルート削減近傍は2opt-2をもとにしたルートを削減するための近傍操作である。Fig. 8にルート削減近傍の例を示す。顧客2からデポと、顧客4からデポの間の辺を繋ぎ変えている。デポからデポに向かうルートは削除される。

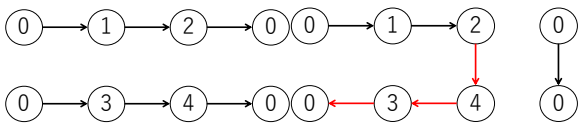


Fig. 8 ルート削減近傍の例

左図はルート削減近傍が行われる前の状態。右図はルート削減近傍が行われた後の状態。

3.2 タブー探索

タブー探索法は、現在の解 x から、近傍 $N(x)$ に含まれるタブーリスト T には含まれない解の中で再了解 x' に移動するという探索手法である。そのため、改悪解への遷移も行われる。また、これまでに見つかった最良解は x^* として記録する。タブーリスト T は、初期状態は空であるが、解が移動するたびに遷移パターンを記録する(遷移パターンは後述)。ただし、最良解 x^* より良い解が見つかった場合に限っては禁止解への遷移を許可する最良移動戦略を用いる。解の移動が行われるたびにタブーリストを更新し続けていくと、いずれ近傍内の解すべてが禁止解になってしまうので、解の移動を一定数行った後に、記録した移動パターンを削除していくこととする。このときタブーに記録しておく期間をタブー期間とする。また、終了条件を満たす際に、それまでの探索で得られている最良の解を出力する。

本研究ではタブーとして記録するものは顧客である。挿入近傍では挿入した顧客と、挿入された顧客の2顧客がタブーリストに記録される。交換近傍では、交換された2顧客がタブーリストに記録される。2-optでは経路の分断に使用された2顧客をタブーとする。

3.3 評価関数

配送計画システムを作成する際には、人件費や労働時間を含めた配送コストを削減するとともに利用者の要望を満たすスケジュールを探索する必要がある。本研究では追加注文が入った場合、それに対して円滑に対応できるようなスケジューリングを目指しているので、その条件を満たす評価関数を導入する必要がある。

本実験で用いた評価関数は総作業量を削減するための関数 $D(x)$ および、スケジュールに余裕を持たせるためのペナルティー項 $P(x)$ を用いる。このペナルティー項 $P(x)$ では、あらかじめ指定した車両が指定時間を超過するたびにペナルティーが増加するため、指定した車両のスケジュールに時間の余裕をつくりやすくする目的がある。これらのことを考慮して、本研究では以下の評価関数を用いる。

評価関数: $E(x) = D(x) + \alpha P(x)$

$D(x)$: 総作業時間(デポを出発してから帰ってくるまでの各車両の時間の合計)

α : 重み $P(x)$: ペナルティー項

ペナルティー項: $P(x) = \sum_{i=1} (T_i - T')$

T_i : 車両 i がデポに戻ってくる時間

T' : あらかじめ指定した時間

4 実験結果

本研究ではVRPTWに休憩制約や、複数人作業制約、追加注文に対応するための評価関数など、実問題に対応するための条件を盛り込んだ問題を取り扱う。

設定した問題では初期の顧客数85に対して、15顧客の追加配送が入る。各顧客での作業時間は顧客ごとに異なり、作業時間が5分で済む顧客や作業時間が60分かかる顧客も存在する。サービスを開始できる時間は顧客ごとに異なるが、設定した問題では8-12時の時間帯と、12-18時の時間帯、8-18時の時間帯の3つがあり、これらの時間帯に沿って荷物を配送しなくてはならない。荷物量についても顧客ごとに異なり、荷物量1の顧客から荷物量75までの需要がある。1つの顧客には1台の配送車両で作業することになるが、複数人作業制約に設定している顧客については2台の車両で作業を行う。追加注文の荷物を配送する車両については、デポに帰還して荷物を積み込む必要があるが、この時にデポで荷物を車両に積むための作業時間は一律20分と定める。

デポは徳島大学常三島キャンパスと設定し、どの車両も始業時間を8時、最遅終業時間は19時とする。また、すべての車両は8時には配送を開始することとする。車両の搭載できる荷物は200を限界と定め、それ以上の荷物は搭載できないものとする。

昼休憩制約については、11時30分から13時30分までの間に1時間の休憩をとるものとする。

4.1 実験

最適化の目的として、スケジュール全体の最適化をする際に評価関数1, 評価関数2の2つ用いて実験を行う。

評価関数1は総作業時間 $D(x)$ の最小化のみを行う。評

価関数2は総作業時間 $D(x)$ に、ペナルティー項 $P(x)$ と重みを加えて最小化する。実験では、ペナルティー項を加えた場合とそうでない場合の比較を行う。

追加注文に使う車両はあらかじめ5台と定め、内1台は必ず午前中に配送を終えるものとする。また、余剰時間は、追加注文に使用する車両がデポに戻ってきたときの時間と、終業時間の差の合計である。Table1にそれぞれの実験結果を表示する。

4.2 結果

作業時間は、顧客についてから作業を行う時間である。移動時間は車両が目的地到達に必要な時間である。休憩時間は、各作業者が休憩を行う時間。休憩時間は一律60分と設定した。待ち時間は、各配送車両がスケジュール内でサービス開始するまでにサービス開始予定時刻よりも前に到着した場合のサービス開始予定時刻まで車両が待機した時間の合計である。総就業時間は、各車両の移動、待ち時間、作業時間、休憩時間を合計したものである。余剰時間は終業時間と、追加配送用に確保した車両がデポに戻ってきた時間との差の合計である。

最適化の結果、追加配送を行う前のスケジュールでは、総就業時間のみを対象とした最適化の方が良い結果になっている。Fig. 9に実際の配送スケジュールを示す。また、Fig. 12に配送経路を示す。

一方で追加配送用に用意した5台の車両の余剰時間についてはペナルティー項を導入したものが多く時間を確保できた。Fig. 10に実際の配送スケジュールを示す。また、Fig. 13に配送経路を示す。

追加配送を実際に行う場合、ペナルティー項を導入したものは追加配送に対応できているが、総就業時間最適化の場合には追加配送に対応することができないことが確認できた。Fig. 11に実際の配送スケジュールを示す。また、Fig. 14に配送経路を示す。この結果から、頻繁に追加注文が入る配送業務において、スケジュールに余裕を持たせることでスケジュールを変更することなくそれらに円滑に対応できる可能性を示せた。今後は、より円滑に追加注文に対応できるようにするために、追加配送を行う指定時間付近にはデポの近くに車両がいるようにする評価関数を実装し、評価を行う予定である。

顧客数 85	総就業時間	車両数	待機時間	余剰時間
総就業時間最適化のみ	5101	10	33	849
ペナルティー項込み	5217	10	39	1072

顧客数 85+追加 15	総就業時間	車両数	待機時間
総就業時間最適化のみ			
ペナルティー項込み	5647	10	39

Table1:実験結果

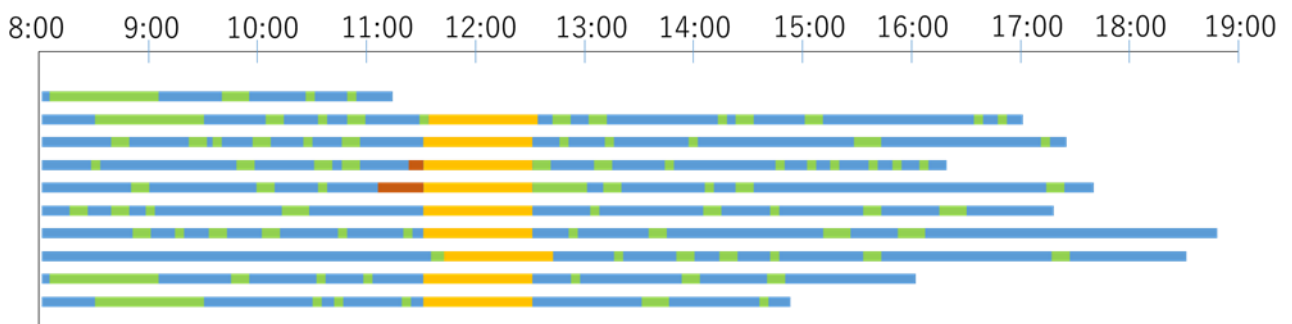


Fig. 9:総就業時間最適化のみの結果

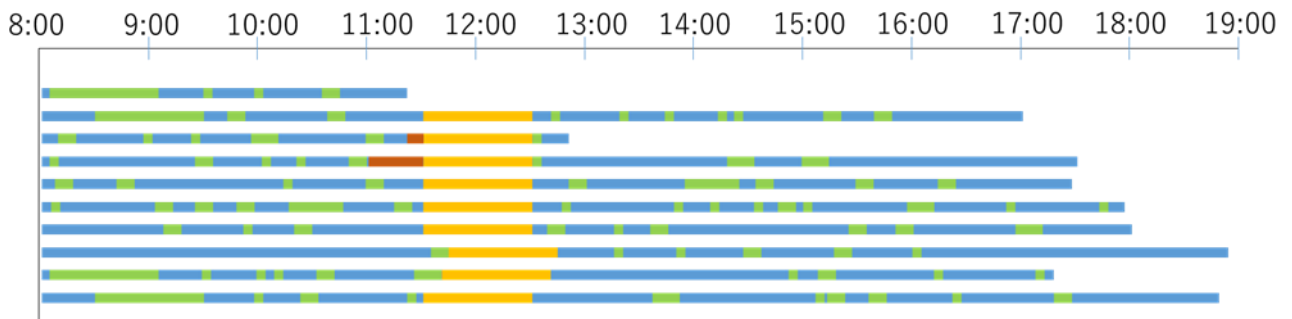


Fig. 10:評価関数にペナルティー項を加えた結果

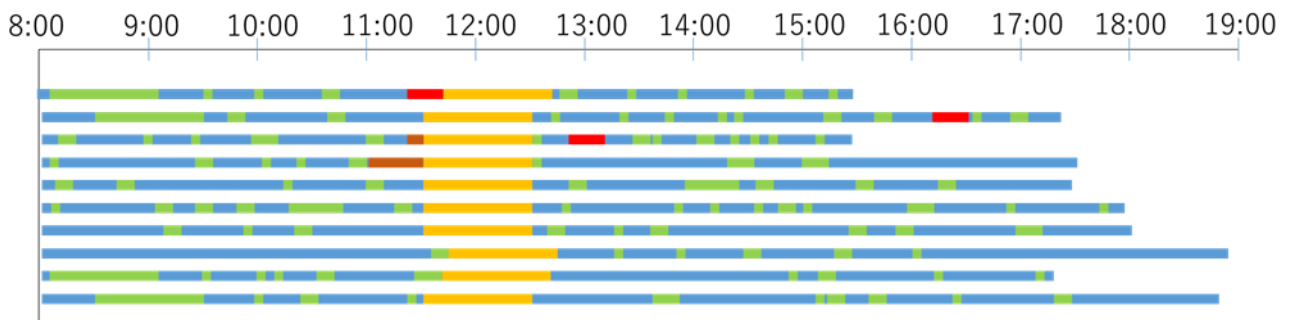


Fig. 11:上記の評価関数にペナルティー項を加えたものに追加配送を行った結果

緑色:顧客での作業時間 青色:移動時間 茶色:待機時間 赤色:デポでの作業時間 黄色:休憩時間

以下には地図上に経路を表示する. 車両ごとに色分けしており, 同じ色のルートが1台の配送ルートになっている.



Fig. 12: 総作業時間のみを考慮した配送経路図 (85 顧客)



Fig. 13: ペナルティーを考慮した配送経路図 (85 顧客)

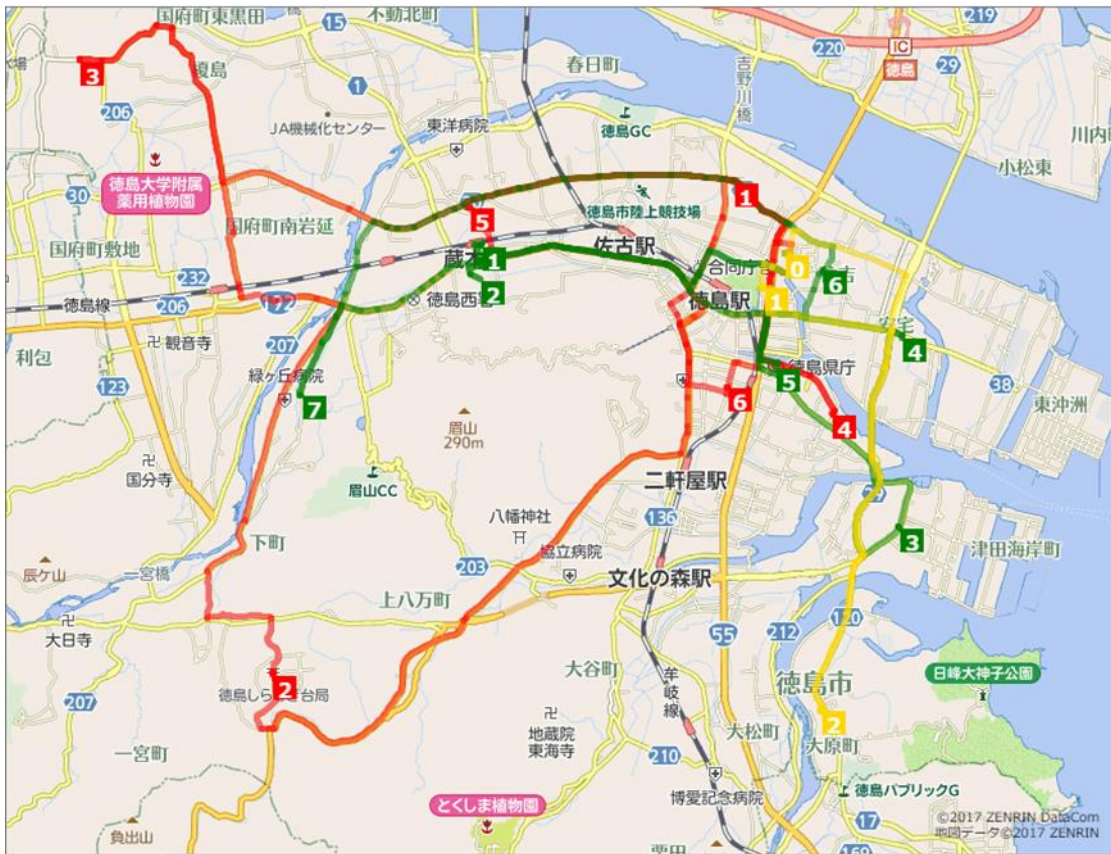


Fig. 14: 追加配送を行った配送経路図(15 顧客)

参考文献

- 1) 岡野 祐之: 配送経路最適化の適応: 銀行における配送例として, 日本オペレーションズ・リサーチ学会, 11 (1999)
- 2) 片井 修, 玄 光男, 大野 勝久, 石渕 久生, 川上 浩司, 辻村 奏寛, 半田 久志, 林 林, 岡本 東: 進化技術ハンドブック, 近代科学社, 1/31 (2010)
- 3) 沖 展彰: 実問題制約付き車両配送問題の対する配送計画システムの提案, 徳島大学大学院, 2/2 (2017)
- 4) 永田 裕一, 小林 重信: Memetic Algorithm による Vehicle Routing Problem with Windows の効率的近似解法

遺伝子ネットワークにおける周期発現パターン遷移列の アイソクロンの計算法

○森 禎弘 黒江康明 (京都工芸繊維大学)

A computational method for isochrons of cyclic expression pattern sequences in gene regulatory networks

*Y. Mori and Y. Kuroe (Kyoto Institute of Technology)

Abstract— There exist periodic phenomena in organisms such as circadian rhythm and those are considered to be generated by gene regulatory networks. Some periodic phenomena in organisms are affected by some stimuli and timings of the oscillations change. For example, circadian clocks in organisms are considered to be affected by light stimuli and phases of the clocks shift, that is timings of the oscillations change. In order to understand such mechanisms, phase shifts caused by perturbations are investigated. Isochron and asymptotic phase are often used for analysis of phase shift of periodic phenomena. However, it is almost impossible to obtain asymptotic phases and isochrons analytically. Therefore, they are computed approximately by using numerical methods. In this paper, we show a computational method of isochrons in gene regulatory networks. In the method, we compute isochrons by using maps obtained analytically. Therefore we can obtain isochrons without using numerical method solving nonlinear differential equations. In this paper, we evaluate the method by performing a numerical example.

Key Words: gene regulatory network, isocline, phase

1 はじめに

生体には概日リズムなど様々な周期現象が存在し、それらは遺伝子ネットワークによって生成されていると考えられている。これらの周期現象は、周りの環境からの影響や相互作用などによる影響を受ける。このような摂動を受けると状態は周期軌道上からずれたところに移動するが、時間がたつと元の周期解軌道に戻る。このとき、一般にその位相は変化する。すなわち、摂動を受けなかったときは異なるタイミングで振動するようになる。例えば、外国に飛行機で移動して時差ぼけ状態になったとき、概日リズムを実現する体内時計は環境の光刺激を受けてその位相が調整され、周りの環境の昼夜のサイクルに合わせたリズムを刻むようになると考えられている。

このような現象を理解するために、周期現象において摂動により位相がどのように変化するかに興味を持たれ解析されている^{1, 2)}。このとき、摂動によってどの位置に移ったかという情報から元の周期軌道に戻ったときの位相の変化量が直ちに分かるならば有用である。そこで、以下に説明するアイソクロンや位相を用いて摂動による位相の変化について解析が行われている。安定な周期解軌道の吸引領域内の異なる複数の点から始まった解軌道が周期軌道に収束したとき、それらが同じタイミングで振動するようになる場合がある。そのような点の集合をアイソクロンという^{3, 4)}。同じアイソクロン上の点は同じタイミングで振動する周期解軌道に収束することからもともと同一位相をもつと考え、これを基にして吸引領域内の位相を定義する。このように定義すると、摂動によって周期軌道から外れたときの状態の位置と位相の変化量を対応付けることができる。

吸引領域内の位相やアイソクロンを求めることができれば、摂動を受けた直後の状態量から位相の変化が分かり、摂動に対する位相の変化を解析するとき

用である。しかし、安定な周期解軌道をもつネットワークのダイナミクスは非線形であるため、一般にアイソクロンや位相を解析的に求めることは困難である。そのため、通常はそれらを数値的に求めることになる。本稿では、文献5)で提案している遺伝子ネットワークにおけるアイソクロンの計算法を示す。この計算法では、遺伝子ネットワークのモデルとして区分線形ネットワークモデルを用いる。そして、遺伝子ネットワークが周期発現パターン遷移列をもつとし、その周期動作についてアイソクロンを求める。非線形モデルであるが、その非線形微分方程式を数値的に解くことなくアイソクロンを近似的に求めることができる簡単な計算法で、モデルのダイナミクスから解析的に求めたポアンカレ写像とポアンカレ断面から始まる解軌道が再びポアンカレ断面と交差するまでの時間間隔を表す写像を用いた繰り返し計算により、ポアンカレ断面上の十分な密度の点集合の点の位相を近似的に求める。そして、それらの点の位相をもとにして、モデルのダイナミクスから解析的に求めた写像を用いてアイソクロンを求める。提案法は、このような方法であるため、遺伝子ネットワークモデルの非線形微分方程式を数値的に解くことなく簡単にアイソクロンを求めることができる。本稿では、数値例により提案法の評価を行う。文献6)では、提案法による周期解軌道の近傍におけるアイソクロンの計算例を示している。この数値例の周期解軌道は大域的に安定であるため、より広い領域にアイソクロンが存在する。そこで、より広い領域でアイソクロンを求めた数値例を示し、提案法を評価する。

2 アイソクロンと位相

ここでは、アイソクロンと位相の定義を示す。システムが周期 T の周期解軌道をもつとする。この周期軌道上の位相は、周期軌道上のある点を基準として定義される。すなわち、周期軌道上のある点 γ_0 の位相を 0 とし、周期軌道上の点 x の位相 θ は、 γ_0 からその点に

到達するまでにかかる時間を Δt として

$$\theta = \omega \Delta t$$

で定義される。ここで、 $\omega = \frac{2\pi}{T}$ である。また、定義からわかるように位相 θ は 0 から 2π の値をとる。

次に、アイソクロンとそれに基づく吸引領域内の点に対する位相を定義する。安定な周期解軌道 Γ の吸引領域内の点を初期値として解軌道を考えて、周期解軌道 Γ に収束した時に同じタイミングでその解軌道が振動するようになる初期点の集合が存在する。そのような点の集合をアイソクロンという³⁾。アイソクロン \mathcal{I} の任意の点 $x_0 \in \mathcal{I}$ と周期軌道の位相 θ の点 γ_θ を初期値とする解軌道をそれぞれ $x(t, x_0), \gamma(t, \gamma_\theta)$ とし、

$$\lim_{t \rightarrow \infty} \|x(t, x_0) - \gamma(t, \gamma_\theta)\| = 0$$

となったとする。ここで、 $\|x\|$ は $x \in R^n$ の適当なノルムである。 $x(t, x_0)$ は位相 θ の点から始まった周期解軌道 $\gamma(t, \gamma_\theta)$ と同じタイミングで振動する周期解軌道に収束するので、そのアイソクロン上の点から始まる解軌道は $\gamma(t, \gamma_\theta)$ と同じ位相をもともともっていると考えられる。そこで、そのアイソクロンの位相を θ とし、そのアイソクロンに属する点の位相は θ であると定義する。すなわち、アイソクロンと吸引領域内の点の位相を以下のように定義する。

【定義 (アイソクロン)】 x_0 から始まる解軌道を $x(t, x_0)$ 、周期軌道上の位相 θ の点 γ_θ から始まる解軌道を $\gamma(t, \gamma_\theta)$ とする。このとき、位相 θ のアイソクロン \mathcal{I}_θ は、

$$\mathcal{I}_\theta = \left\{ x_0 \in \mathcal{B} \mid \lim_{t \rightarrow \infty} \|x(t, x_0) - \gamma(t, \gamma_\theta)\| = 0 \right\}$$

である。

【定義 (吸引領域内の点の位相)】 吸引領域内の点 x_0 から始まる解軌道を $x(t, x_0)$ とする。また、同じ時刻に周期軌道上の位相 θ から始まる解軌道を $\gamma(t, \gamma_\theta)$ とする。

$$\lim_{t \rightarrow \infty} \|x(t, x_0) - \gamma(t, \gamma_\theta)\| = 0$$

となるならば、 x_0 の位相は θ である。

このようにアイソクロンや位相を定義すると、摂動により周期解軌道から外れた解軌道が周期軌道に収束したとき、どのようなタイミングで振動するかが摂動を受けた後の解軌道の場所から直ちに分かるようになり、摂動による位相の変化の解析を行うときに有用である。

3 問題設定

文献5)では遺伝子ネットワークを対象として、そのアイソクロンの計算法を提案している。本稿では、この提案法の評価を行う。本稿で対象とする遺伝子ネットワークのモデルは、次式の区分線形微分方程式モデルである⁷⁾。

$$\begin{aligned} \dot{x}_i(t) &= -d_i x_i(t) + f_i(w_{i1}, w_{i2}, \dots, w_{im_i}, \\ &\quad y_1(t), y_2(t), \dots, y_n(t)), \\ x_i(0) &= x_{(0,i)} \end{aligned} \quad (1)$$

$$y_i(t) = h(x_i(t)), \quad i = 1, 2, \dots, n \quad (2)$$

ここで、 h はしきい値関数で、次式で定義される。

$$h(x_i) = \begin{cases} 0 & \text{if } x_i < 0 \\ 1 & \text{if } x_i \geq 0 \end{cases} \quad (3)$$

$x_i(t)$ は i 番目の遺伝子の生成物の正規化された濃度を表す。 $x_{(0,i)}$ はその初期状態である。 n は遺伝子の数である。 $y_i(t)$ は遺伝子 i の発現レベルを表し、その値が 1 ならば発現している、0 ならば発現していないことを表す。 f_i は遺伝子間の相互作用を表す関数である。 w_{ij} は f_i のパラメータで、 m_i は f_i のパラメータの個数を表す。 d_i は i 番目の遺伝子の生成物の分解速度に関する正実数パラメータである。

以下ではこのモデルをベクトル形式で次式のように表す。

$$\dot{x}(t) = -Dx(t) + f(w, y(t)), \quad x(0) = x_0 \quad (4)$$

$$y(t) = H(x(t)) \quad (5)$$

ここで、 $x(t) = [x_1(t) \ x_2(t) \ \dots \ x_n(t)]^T$ であり、 $x_0 = [x_{(0,1)} \ x_{(0,2)} \ \dots \ x_{(0,n)}]^T$ である。 y は $y(t) = [y_1(t) \ y_2(t) \ \dots \ y_n(t)]^T$ であり、各遺伝子の発現レベルを表すベクトルである。ここではこの y を遺伝子ネットワークの発現パターンと呼ぶ。 f は $f = [f_1 \ f_2 \ \dots \ f_n]^T$ である。 w は $w^T = [w_1^T \ w_2^T \ \dots \ w_n^T]$ で、 $w_i = [w_{i1} \ w_{i2} \ \dots \ w_{im_i}]^T$ である。 D は対角行列で、 $D = \text{diag}\{d_1, d_2, \dots, d_n\}$ である。 $H(x)$ は、 $H(x) = [h(x_1) \ h(x_2) \ \dots \ h(x_n)]^T$ である。

本稿では、式 (4)、(5) の遺伝子ネットワークが周期 T の安定な周期解軌道 Γ をもつとし、周期解軌道 Γ のときに現れる周期発現パターン遷移列を、 p を遷移回数として

$$y^{*(0)} \rightarrow y^{*(1)} \rightarrow \dots \rightarrow y^{*(p)}, \quad y^{*(0)} = y^{*(p)} \quad (6)$$

とする。ただし、 $y^{*(r)}$ と $y^{*(r+1)}$ 、 $r = 0, 1, \dots, p-1$ で異なる要素は一つのみとする。本稿の問題は、この周期解軌道 Γ をもつ遺伝子ネットワークのアイソクロンを求めることである。

以下に本稿で用いる記号を定義する。式 (4)、(5) の遺伝子ネットワークが発現パターン y をとる x の n 次元実数空間の領域を Ω_y と定義する。すなわち、

$$\Omega_y = \{x \mid y = H(x)\}$$

と定義する。さらに、発現パターン y に対して $e(y)$ を

$$e(y) = D^{-1}f(w, y)$$

と定義する。また、安定な周期解軌道 Γ の吸引領域を \mathcal{B} で表す。 $y^{*(r)}$ と $y^{*(r+1)}$ で異なる要素の添字を i_r と表記する。すなわち、 $y_{i_r}^{*(r)} \neq y_{i_r}^{*(r+1)}$ 、 $y_i^{*(r)} = y_i^{*(r+1)}$ 、 $i \neq i_r$ である。また、 $\Omega_{y^{*(r-1)}}$ と $\Omega_{y^{*(r)}}$ の境界を S_r 、 $r = 1, 2, \dots, p$ と定義する。ただし、周期発現パターンの場合には $y^{*(0)} = y^{*(p)}$ であるので、 $\Omega_{y^{*(p-1)}}$ と $\Omega_{y^{*(0)}}$ の境界であることを強調するときには、 $\Omega_{y^{*(p-1)}}$ と $\Omega_{y^{*(p)}} (= \Omega_{y^{*(0)}})$ の境界 S_p を S_0 と表記する。

4 アイソクロンの計算法

一般には、モデルの非線形微分方程式を解いてアイソクロンを求めるが、本稿で示すアイソクロンの計算法では、モデルの非線形微分方程式を解くことなく吸引領域内のポアンカレ断面上の点の位相を求め、それらの情報をもとにアイソクロンを求める。本節では、まず位相の計算法を示し、次に、ポアンカレ断面上の位相の計算結果を利用したアイソクロンの計算法を示す。このとき、周期解軌道 Γ の S_0 との交点 γ_0 を基準として周期軌道上の位相を定義する。また、以下では

$$e(y^{*(r)}) \in \Omega_{y^{*(r+1)}}, r = 0, 1, \dots, p-1 \quad (7)$$

が成り立っているとす。

4.1 吸引領域内の位相の計算法

安定な周期解軌道の吸引領域内の点 x_0 の位相は、位相の定義に従って次のような考え方に基いて近似的に求めることができる。 x_0 と γ_0 から同時に始まった解軌道の時間発展をそれぞれ考える。 x_0 から始まる解軌道 $x(t, x_0)$ が S_0 上で周期解軌道 Γ に十分近づいたときの時刻を求める。この時刻に $x(t, x_0)$ は周期解軌道上で位相 0 の場所に存在することになる。また、その時刻における γ_0 から始まった解軌道 $\gamma(t, \gamma_0)$ が存在する場所の位相は、その時刻と周期解軌道の周期より求めることができる。 $x(t, x_0)$ は周期解軌道上における位相 0 のところに存在することから、 $\gamma(t, \gamma_0)$ が存在する場所の位相は周期解軌道上における x_0 が存在する場所と γ_0 が存在する場所の位相差に対応する。よって、その位相差から直ちに x_0 の位相の近似値を得ることができる。

このような考え方で、式 (4), (5) の微分方程式を数値的に解くことなく位相を求める方法を示すために、 S_0 をポアンカレ断面としたときのポアンカレ写像と解軌道がポアンカレ断面と交差してから次に交差するまでの時間間隔を表す写像を示す。ポアンカレ写像は、 x_0 から始まる解軌道が S_0 と交差したときに周期解軌道 Γ に十分に近づいているかどうかを調べるために利用される。時間間隔の写像は、解軌道が周期解軌道に十分近づいたときの時刻を求めるときに利用される。

先にポアンカレ写像を示す。そのために、 $\Omega_{y^{*(r)}}, r = 0, 1, \dots, p-1$ 内の点 x に対してつぎの写像を導入する。

$$\begin{aligned} \psi_i^{(r+1)}(x) &= e_i(y^{*(r)}) - (e_i(y^{*(r)}) - x_i) \\ &\quad \times \left(\frac{e_{i_r}(y^{*(r)})}{e_{i_r}(y^{*(r)}) - x_{i_r}} \right)^{\frac{d_i}{d_{i_r}}}, \\ & \quad i = 1, 2, \dots, n, \quad r = 0, 1, \dots, p-1 \end{aligned}$$

この写像は、 $\Omega_{y^{*(r)}}$ 内の点 x から始まる解軌道が境界 S_{r+1} と交差する点 $\psi^{(r+1)}(x), r = 0, 1, \dots, p-1$ を与える⁹⁾。

さて、式 (7) の条件が成り立つとしているので、 $\Omega_{y^{*(0)}}$ 内の点 x_0 から出発する解軌道は、式 (6) の周期発現パターン遷移列に対応して境界 $S_1, S_2, \dots, S_{p-1}, S_p$ と順に交差する⁸⁾。このとき、 S_r と交差する点は、 $\psi^{(k)}, k = 1, 2, \dots, r$ を用いて x_0 に対して

$$\Psi^{(r)}(x_0) = \psi^{(r)} \circ \psi^{(r-1)} \circ \dots \circ \psi^{(1)}(x_0) \quad (8)$$

で定義される写像により与えられる。ここで、 $g \circ f(x) = g(f(x))$ である。 $\Psi^{(p)}(x_0)$ は S_0 上の x_0 から始まる解軌道が S_0 と交差する点を与える。よって、 S_0 をポアンカレ断面としたときのポアンカレ写像は、 $\Psi^{(p)}(x)$ となる⁹⁾。以下では、表記の簡単のため、このポアンカレ写像を $\Psi(x)$ と表す。

次に、解軌道がポアンカレ断面と交差してから次にポアンカレ断面と交差するまでにかかる時間を表す写像を求めるため、解軌道が領域 $\Omega_{y^{*(r)}}$ 内を通過するのにかかる時間を表す写像を示す。 S_0 上の点 x_0 から始まった解軌道が S_r と交差した時刻を $t = 0$ とする。このときの S_r との交点は $\Psi^{(r)}(x_0)$ である。時刻 $t = 0$ に $\Psi^{(r)}(x_0)$ で S_r と交差した解軌道は領域 $\Omega_{y^{*(r)}}$ 内を通過して $e(y^{*(r)}) \in \Omega_{y^{*(r+1)}}$ へ向かい、 S_{r+1} と $\Psi^{(r+1)}(x_0)$ で交差し、遺伝子ネットワークの発現パターンは $y^{*(r)}$ から $y^{*(r+1)}$ へと変化する⁸⁾。このときの、領域 $\Omega_{y^{*(r)}}$ 内の解軌道は、 $\Psi^{(r)}(x_0)$ を用いて

$$\begin{aligned} x_i(t) &= \Psi_i^{(r)}(x_0) \exp(-d_i t) \\ &\quad + e_i(y^{*(r)}) (1 - \exp(-d_i t)), \\ & \quad i = 1, 2, \dots, n \end{aligned} \quad (9)$$

と表される。この軌道が S_{r+1} と交差するときに $x_{i_r} = 0$ となることから、 S_r 上の点 $\Psi^{(r)}(x_0)$ から始まる解軌道が境界 S_{r+1} と交差するまでにかかる時間 $T^{(r+1)}(\Psi^{(r)}(x_0))$ は

$$\begin{aligned} T^{(r+1)}(\Psi^{(r)}(x_0)) \\ = -\frac{1}{d_{i_r}} \left\{ \log \frac{e_{i_r}(y^{*(r)})}{e_{i_r}(y^{*(r)}) - \Psi_{i_r}^{(r)}(x_0)} \right\} \end{aligned} \quad (10)$$

で得られる。以上のことは、任意の $r \in \{0, 1, \dots, p-1\}$ に対して成り立つので、解軌道がポアンカレ断面 S_0 と x_0 で交差してから、次にポアンカレ断面と交差するまでの時間間隔 $T_p(x_0)$ は

$$T_p(x_0) = \sum_{r=1}^p T^{(r)}(\Psi^{(r-1)}(x_0)) \quad (11)$$

となる¹⁰⁾。ただし、 $\Psi^{(0)}(x) = x$ である。

x_0 が S_0 上ではない場合、 x_0 から始まった解軌道がはじめて S_0 と交差するまでの経過時間とそのときの交点を求める必要がある。それらはつぎのようにして求めることができる。 x_0 が $\Omega_{y^{*(v)}}$ 内の点であるとする。二つの写像 $\Psi^{(r)}$ と $T^{(r)}$ を求めたときの議論は、 $\Omega_{y^{*(r)}}$ 内の任意の点に対して成り立つので、 $\Omega_{y^{*(v)}}$ 内の点 x_0 から始まる解軌道が S_0 と交差する点 \bar{x} と交差するまでにかかる時間 \bar{T} は $\psi^{(r)}$ と $T^{(r)}$ を用いて次式で表すことができる。

$$\bar{x} = \psi^{(p)} \circ \psi^{(p-1)} \circ \dots \circ \psi^{(v+1)}(x_0) \quad (12)$$

$$\bar{T} = \sum_{r=v+1}^p T^{(r)}(\bar{\Psi}^{(r-1)}(x_0)) \quad (13)$$

$$\begin{aligned} \bar{\Psi}^{(k)}(x_0) &= \psi^{(k)} \circ \psi^{(k-1)} \circ \dots \circ \psi^{(v+1)}(x_0), \\ & \quad k = v+1, v+2, \dots, p-1 \end{aligned} \quad (14)$$

ただし、 $\bar{\Psi}^{(v)}(x) = x$ である。

以上で示した写像を使って吸引領域内の点 x_0 の位相を求めるための考え方を説明する。 $x_0 \in \Omega_{y^{*(v)}}$ なる x_0 から始まる解軌道 $x(t, x_0)$ が S_0 と初めて交差する時刻 \bar{T} と交点の座標 \bar{x} は、式 (13) と式 (12) よりそれぞれ求められる。 S_0 と交差してからつぎに S_0 と交差する時刻と交点の座標は、式 (11) を用いて $\bar{T} + T_p(\bar{x})$ 、ポアンカレ写像を用いて $\Psi(\bar{x})$ とそれぞれ求めることができる。さらに式 (11) とポアンカレ写像を繰り返し用いて、 S_0 上で周期軌道 Γ に十分近づいたときの時刻 t^k がわかる。また、周期軌道 Γ 上の位相 0 の点 γ_0 から $x(t, x_0)$ と同時に始まった解軌道 $\gamma(t, \gamma_0)$ が時刻 t^k に周期軌道上のどの位相の場所に存在するかは、周期が分かっているので簡単に求めることができ、その場所の位相が $x(t, x_0)$ と $\gamma(t, \gamma_0)$ の位相差であることから、 x_0 の位相はその位相差から簡単に求めることができる。

このような考え方から次のアルゴリズムが得られる。 ϵ_T と ϵ_x は適切に設定する。

【位相を求めるアルゴリズム】

step 1 吸引領域内の点 $x_0 \in \Omega_{y^{*(v)}}$ から始まる解軌道が S_0 平面と交差したときの時刻 $t^{(0)}$ と交点の座標 $x^{(0)}$ を式 (12), (13) を用いて求める。 $k = 0$ とする。

step 2 $x^{(k)}$ を通過した解軌道が、次に S_0 と交差する時刻 $t^{(k+1)}$ と交点 $x^{(k+1)}$ を式 (11) と式 (8) を用いて

$$\begin{aligned} t^{(k+1)} &= t^{(k)} + T_p(x^{(k)}), \\ x^{(k+1)} &= \Psi(x^{(k)}) \end{aligned}$$

と求める。また、 $\bar{T}^{(k+1)} = T_p(x^{(k)})$ とする。

step 3 $k = k + 1$ とする。 $|T - \bar{T}^{(k)}| > \epsilon_T$ あるいは $\|\gamma_0 - x^{(k)}\| > \epsilon_x$ ならば step 2 へ、そうでなければ step 4 へ。

step 4 $t^{(k)}$ を T で割ったときの余り \hat{T} を求め、 $\Delta t = T - \hat{T}$ とする。

step 5 x_0 の位相の近似値 θ は $\theta = \omega \Delta t$ である。

解析的に求めたポアンカレ写像などの繰り返し計算により位相を求めており、簡単な計算で位相が得られる。

4.2 アイソクロンの計算

ここではアイソクロンを求める方法を説明する。提案法では、ポアンカレ断面上の位相を求め、ポアンカレ断面上の点から指定した位相差の場所にある点を解析的に求めることでアイソクロンを求める。この方法の説明のために、指定した位相差の点を解析的に求める方法を示す。

位相 θ_A の点 x_A から、 x_A を通る解軌道上の位相 θ_B の点 x_B を求めるとし、 $x_A \in \Omega_{y^{*(v)}}$ であるとする。同じ解軌道上の二つの点の位相差は、片方の点からもう一方の点へ解軌道が到達するまでにかかる時間に対応する。 $\theta_B > \theta_A$ の場合、 x_A と x_B の位相差から、その時間は

$$\Delta t = \frac{\theta_B - \theta_A}{\omega} \quad (15)$$

と求められる。 $\theta_B \leq \theta_A$ の場合は、

$$\Delta t = \frac{2\pi - (\theta_A - \theta_B)}{\omega} \quad (16)$$

と求められる。 x_A から Δt だけ進んだ場所が求めるべき x_B である。領域 $\Omega_{y^{*(v)}}$ 内の x_A から始まる解軌道の式は、式 (9) において $\Psi^{(r)}(x_0)$ を x_A に、 $e(y^{*(r)})$ を $e(y^{*(v)})$ にそれぞれ置き換えることで

$$\begin{aligned} x_i(t) &= x_A \exp(-d_i t) \\ &\quad + e_i(y^{*(v)})(1 - \exp(-d_i t)), \\ i &= 1, 2, \dots, n \end{aligned} \quad (17)$$

と得られる。その軌道が S_{v+k} と交差する時刻 $t^{(k)}$ と交差する点の座標は、

$$t^{(k)} = \sum_{j=v+1}^{v+k} T^{(j)}(\bar{\Psi}^{(j-1)}(x_A)) \quad (18)$$

と $\bar{\Psi}^{(k+1)}(x_A)$ で与えられる。ただし、 $j \geq p$ なら $j = j - p$ として計算する。このとき、 $t^{(k)} < \Delta t \leq t^{(k+1)}$ となる k を求めると、 x_A から Δt だけ進んだ解軌道は吸引領域内の領域 $\Omega_{y^{*(r)}}$ に存在することがわかる。ただし、 $r = v + k$ とし、 $r \geq p$ なら $r = r - p$ とする。そこで、次式で得られる $\hat{T}^{(k+1)}$ を求め、 $\hat{T}^{(k+1)} < 0$ となる最小の k を求める。

$$\begin{aligned} \hat{T}^{(k+1)} &= \Delta t - t^{(k+1)} \\ &= \Delta t - (t^{(k)} + T^{(v+k+1)}(\bar{\Psi}^{(v+k)}(x_A))) \\ &= \hat{T}^{(k)} - T^{(v+k+1)}(\bar{\Psi}^{(v+k)}(x_A)) \end{aligned} \quad (19)$$

S_{v+k} と交差するときの交点は $\bar{\Psi}^{(v+k)}(x_A)$ 、その交点から x_B に到達するまでの経過時間は $\hat{T}^{(k)}$ として求められているので、 x_B の座標を

$$\begin{aligned} x_{(B,i)} &= \bar{\Psi}^{(v+k)}(x_A) \exp(-d_i \hat{T}^{(k)}) \\ &\quad + e_i(y^{*(v+k)})(1 - \exp(-d_i \hat{T}^{(k)})), \\ i &= 1, 2, \dots, n \end{aligned} \quad (20)$$

と求めることができる。以上の考え方より x_B を求めるアルゴリズムはつぎのようになる。

【指定した位相差の座標を求めるアルゴリズム】

step 1 $\theta_B > \theta_A$ のときは式 (15)、そうでないときは式 (16) により Δt を求める。

step 2 $\hat{T}^{(0)} = \Delta t$, $x^{(0)} = x_A$, $k = 0$ とする。

step 3 $\hat{T}^{(k+1)} = \hat{T}^{(k)} - T^{(r+1)}(x^{(k)})$ と $x^{(k+1)} = \psi^{(r+1)}(x^{(k)})$ を求める。

step 4 $\hat{T}^{(k+1)} < 0$ ならば step 6 へ。

step 5 $r = r + 1$ とする。 $r = p$ なら $r = 0$ とする。

step 6 $k = k + 1$ として step 2 へ。

step 7 $x_0 = x^{(k)}$, $\hat{T} = \hat{t}^{(k)}$ として x_B を求める.

$$\begin{aligned} x_{(B,i)} = & x_{(0,i)} \exp(-d_i \hat{T}) \\ & + e_i (y^{*(r)}) (1 - \exp(-d_i \hat{T})), \\ & i = 1, 2, \dots, n \end{aligned} \quad (21)$$

また、解軌道上で位相 θ_B の場所を探すときに、時刻を遡って探すことも同様の方法で可能である。この計算においても先の【位相を求めるアルゴリズム】で利用したポアンカレ写像などの繰り返し計算のみで簡単に所望の座標を求めることができる。

ポアンカレ断面上の十分な密度の点集合に対して位相を求め、このアルゴリズムによりそれらの点を通る解軌道上の同位相の点の集合を求めることでアイソクロンを求めることができる。位相 θ_B のアイソクロンを求めるときは、次のようにすればよい。

【アイソクロンを求めるアルゴリズム】

step 1 ポアンカレ断面上の十分な密度の点に対して、【位相を求めるアルゴリズム】によりそれぞれの位相を求める。

step 2 位相の値を求めたポアンカレ断面上の点に対して、【指定した位相差の場所を求めるアルゴリズム】によりその点を通る軌道上を時間について順方向に位相 θ_B の場所を求める。

step 3 位相の値を求めたポアンカレ断面上の点に対して、その点を通る軌道上を時間について逆方向に位相 θ_B の場所を求める。

これにより、ポアンカレ断面上の位相を求めた点に対応して位相 θ_B のアイソクロンに属する点の集合が求められる。このアルゴリズムで必要な計算は、先に述べたように解析的に求められたポアンカレ写像などの計算の繰り返しであるため、簡単にアイソクロンを求めることができる。

5 数値例

ここでは、本稿で示した計算法によりアイソクロンを求めた結果を示す。相互作用関数は、

$$\begin{aligned} f_i(a^{(i)}, y) = & a_0^{(i)} + \sum_{j=1}^n a_j^{(i)} y_j + \sum_{j=1}^{n-1} \sum_{k=j+1}^n a_{jk}^{(i)} y_j y_k \\ & + \dots + a_{12\dots n}^{(i)} y_1 \dots y_n, \quad i = 1, 2, \dots, n \end{aligned}$$

がよく用いられる。ここでもこれを用いることにする。

遺伝子の数は2個で、次の遺伝子ネットワークのアイソクロンを求めた結果を示す。

$$\begin{aligned} \dot{x}_1(t) = & -x_1(t) \\ & + (3 - y_1(t) - 7y_2(t) + 3y_1(t)y_2(t)) \end{aligned} \quad (22)$$

$$\begin{aligned} \dot{x}_2(t) = & -3x_2(t) \\ & + (-4 + 5y_1(t) + 2y_2(t)) \end{aligned} \quad (23)$$

この遺伝子ネットワークは、Fig. 1 に示す安定な周期解軌道をもつ。周期解軌道の周期は1.0314で、このときの発現パターンは、

$$[1 \ 1]^T \rightarrow [0 \ 1]^T \rightarrow [0 \ 0]^T \rightarrow [1 \ 0]^T \rightarrow [1 \ 1]^T \quad (24)$$

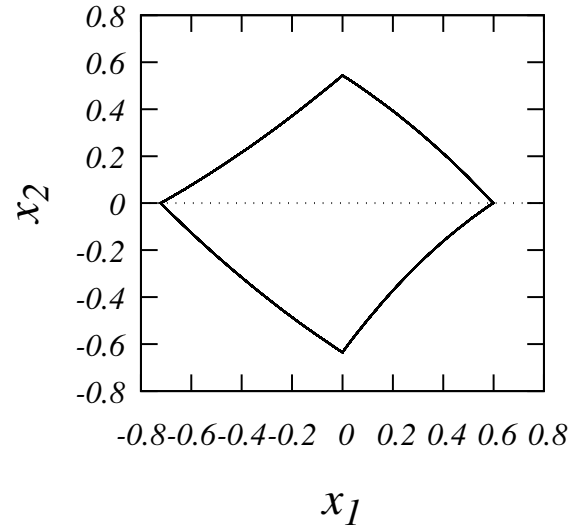


Fig. 1: Periodic orbit

Table 1: Phases of points on Poincaré map

x_1	0.25	0.5	0.75	1.0
位相	2.682	0.560	5.564	4.635
x_1	1.25	1.5	1.75	2.0
位相	3.887	3.266	2.719	2.229
x_1	2.25	2.5	2.75	3.0
位相	1.796	1.394	1.017	0.672
x_1	3.25	3.5	3.75	4.0
位相	0.349	0.0425	6.035	5.759
x_1	4.25	4.5	4.75	5.0
位相	5.495	5.244	5.01	4.779

と変化する。

まず、本稿で示した方法により、ポアンカレ断面 S_0 上の位相を求める。ポアンカレ断面 S_0 は、 $S_0 = \{x | x_1 \geq 0, x_2 = 0\}$ であり、周期解軌道は S_0 と $[0.5982 \ 0.0]^T$ で交差する。ポアンカレ断面上で、 x_1 の値を0から5まで0.001刻みで変化させ、5000点の位相を求めた。その結果の一部をTable 1に示す。これらのポアンカレ断面上の点を通る解軌道上において位相が $0.0, 0.2\pi, 0.4\pi, 0.6\pi, 0.8\pi, 1.0\pi, 1.2\pi, 1.4\pi, 1.6\pi, 1.8\pi$ の点を求めることで、それぞれの位相に対応するアイソクロンを求めた。その結果をFig. 2に示す。Fig. 2において、 $[0.5982 \ 0.0]^T$ を通る赤い曲線が位相0のアイソクロンである。この位相0のアイソクロンから原点に向かって $0.2\pi, 0.4\pi, 0.6\pi, 0.8\pi, 1.0\pi, 1.2\pi, 1.4\pi, 1.6\pi, 1.8\pi$ の順にその位相のアイソクロンが得られている。

この数値例の周期解軌道は大域的に安定である。そこで、より広い状態空間内のアイソクロンの計算結果をFig. 3に示す。周期解軌道の近傍だけでなくより広い領域でアイソクロンを求めることができている。

6 おわりに

生体には概日リズムなど様々な安定な周期現象が存在し、それらは遺伝子ネットワークによって生成されていると考えられている。また、それらは環境などからの影響により位相が調整されることも多い。その仕組みを調べるために、摂動により位相がどのように変化するかに興味を持たれ調べられている。このとき、安定な周期軌道の吸引領域内の位相やアイソクロンを求

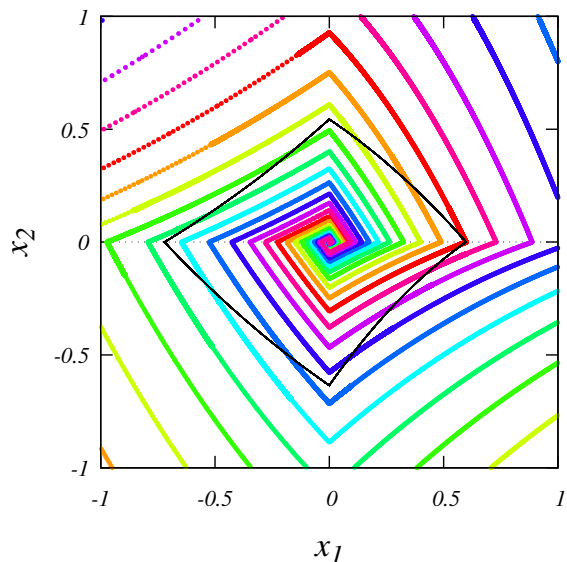


Fig. 2: Isochrons of phases of $0.0, 0.2\pi, 0.4\pi, 0.6\pi, 0.8\pi, 1.0\pi, 1.2\pi, 1.4\pi, 1.6\pi, 1.8\pi$ in $[-1, 1] \times [-1, 1]$

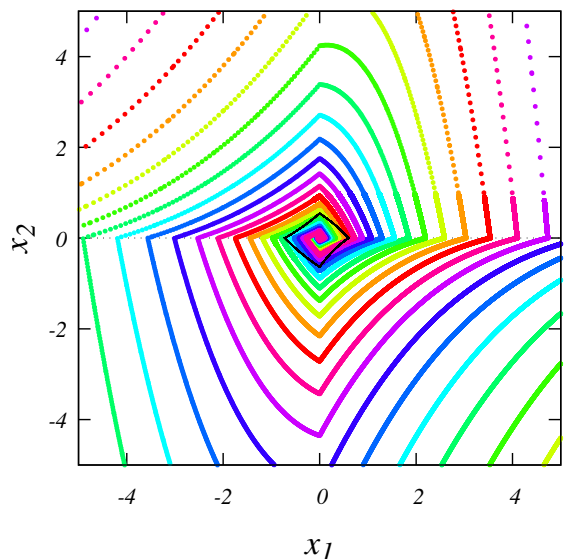


Fig. 3: Isochrons of phases of $0.0, 0.2\pi, 0.4\pi, 0.6\pi, 0.8\pi, 1.0\pi, 1.2\pi, 1.4\pi, 1.6\pi, 1.8\pi$ in $[-5, 5] \times [-5, 5]$

めることができれば、解析のときに有用である。しかし、位相やアイソクロンを解析的に求めることは困難で、一般に数値的に求めることになる。

本稿では区分線形ネットワークモデルを対象にして文献5)で筆者らが提案している遺伝子ネットワークのアイソクロンの計算法を示した。一般には遺伝子ネットワークのモデルの非線形微分方程式を数値的に解いて位相を求めるが、提案法では非線形微分方程式を解くことなく、区分線形ネットワークモデルであることを利用し、解析的に求めた写像を用いた計算によりアイソクロンを簡単に求めることができる。本稿では、数値例によりこの方法の評価を行い、周期解軌道の近傍だけでなくより広い領域でアイソクロンを求めることができることを示した。

参考文献

- 1) 蔵本 由紀:リズム現象の世界, 東京大学出版会 (2005)
- 2) 郡 宏, 森田善久:生物リズムと力学系, 共立出版 (2011)
- 3) A. T. Winfree: Patterns of phase compromise in bi-

ological cycles, Journal of Mathematical Biology, **1**, 73/95 (1974)

- 4) K. Josic, E. T. Shea-Brown, and J. Moehlis: Isochron, Scholarpedia, **1**-8, 1361 (2006)
- 5) 森 禎弘, 黒江康明: 遺伝子ネットワークにおけるアイソクロンの計算法, 第23回インテリジェント・システム・インポジウム FAN2013 講演論文集, 63/68 (2013)
- 6) 森 禎弘, 黒江康明: 遺伝子ネットワークにおけるアイソクロンの簡単な計算法, 第4回コンピューショナル・インテリジェンス研究会講演論文集, 102/107 (2013)
- 7) L. Glass: Classification of biological networks by their qualitative dynamics, Journal of Theoretical Biology, **54**-1, 85/107 (1975)
- 8) 森 禎弘, 黒江 康明, 森 武宏: 発現パターンに基づく遺伝子ネットワークの設計法, 計測自動制御学会論文集, **44**-11, 936/945 (2008)
- 9) Y. Mori and Y. Kuroe: Synthesis method of gene regulatory networks having cyclic expression pattern sequences –realization by introducing Poincarè map–, Proc. of SICE Annual Conference 2011, 544/549 (2011)
- 10) Y. Mori and Y. Kuroe: Synthesis method of gene regulatory networks having desired periodic expression sequences, Proc. of 2012 International Conference on Systems, Man, and Cybernetics, 1159/1164 (2012)

量子ダイナミクスによる群ロボットの知的タスク処理

○伊丹哲郎 (RIDC) 松井伸之 (兵庫県立大学)

概要 群ロボットの運動が「量子ダイナミクス」により制御される。課されるタスクは、閉じた平面上の物体の搬送(トラッキング)である。各ロボット, また搬送される物体も, 厳密にニュートン力学で記述される。各ロボットに特別なポテンシャルを作用することで, 量子効果が与えられる。結果として, タスク処理に必要なロボット台数の情報を系統的に設計することができる。

キーワード: 巨視的な物体の搬送システム, 群ロボット, 量子力学的波動のニュートン力学による表現

1 はじめに

著者らは, スォーム・ロボティクス¹⁾の視点から物体をロボット群により搬送する研究を行ってきた。ロボットを衝突フリーに設計するのではなく, むしろ積極的に衝突させる, あるいは衝突を利用する, というのが提案システムの特徴である。この目的のため各ロボットは可能な限り安価に仕上げられている。すなわち外界センサを搭載せず互いのコミュニケーションも必要としない。つまりわれわれのロボットは五感をまったく持たないのである。また念のために追記しておくが, 物体もロボットも「巨視的」である。すなわち彼らは $\sim 1[\text{m}]$, $\sim 1[\text{kg}]$ のオーダーの大きさを持ち, $\sim 1[\text{s}]$ オーダーの時間で仕事をする。

われわれの巨視的物体の搬送系^{2, 3, 4)}では, ボルツマン分布をポテンシャル場で変形する。これにより, 搬送される物体の運動を制御する。いっぽうブラウンモータ⁵⁾では, 温度勾配を操作して液体分子の熱揺らぎを制御することにより駆動部の適切な運動を図る。すなわち確率分布関数の温度場による変形が入力因子である。これらの巨視的物体の搬送系, またその着想の元であるブラウンモータ, いずれにおいても, 物体の運動や駆動部が分布関数という連続的に領域全体にわたる「場」あるいは「波」で制御される, と理解することができる。であるならば, 直接に「波」で物体を動かすことができるか? その波として量子物理にあらわれる波動関数 Ψ を取る。すると実はその波とは, 物体, しかし搬送される物体より小さなロボット, で表現されるのである。そしてロボット台数は連続的な波動関数をどの程度で近似するか, を評価基準として決めることができる。すなわち本研究は, 巨視的物体を「波」で搬送する試みである。この波はなんら「媒体」を要するものではなく, 実はロボット群で表現され, ふうの機械として設計される。

はじめに2で, 波を表現するための量子ダイナミクスを示す。ここでは全系のシュレディンガー方程式から出発し, その波動関数にわたる期待値をとる。これにより物体 B のニュートンの運動方程式とロボット群のシュレディンガー方程式を得る。これらに基づき, 波の作用による物体の駆動が状態方程式の形で具体的に与えられる。これは, 物体重心が量子揺らぎを持たない条件下でなされる。ここで, ポテンシャル外場の「形」が物体運動の制御に対する操作入力であること, が明示される。3は, この量子の波による物体搬送の数値シミュレーション結果を示す。この波がなんら媒体は必要とせず, 実はロボットというメカニカルな機構で生成されることを, 4で明らかにする。最後に5で議論を

まとめ, 課題を示す。付録Aでは, 確定的なニュートン力学が量子力学の波動を表現することを説明する。

2 量子ダイナミクス

2次元の矩形領域に物体ひとつとひとつあるいは複数多数台のロボットを入れる。質量 M の被搬送物体 B の座標を \vec{X} , 物質波である質量 m のロボット R の座標を \vec{x} とする。全体系の波動関数 $\Psi(\vec{x}, \vec{X}; t)$ のシュレディンガー方程式は

$$\begin{aligned} iH_R \frac{\partial \Psi}{\partial t} &= \hat{H} \Psi \\ &= -\frac{H_R^2}{2M} \frac{\partial^2 \Psi}{\partial \vec{X}^2} - \frac{H_R^2}{2m} \frac{\partial^2 \Psi}{\partial \vec{x}^2} \\ &\quad + (V_{\text{cnt}}(\vec{x}; \mathbf{u}) + V_B(|\vec{x} - \vec{X}|) + V_{B0}(\vec{X})) \Psi \quad (1) \end{aligned}$$

である。ここでプランク定数 $(\hbar/(2\pi))$ に相当する定数を H_R と表した。右辺の第1, 2項は, 物体とロボットそれぞれの運動エネルギーである。第3項は制御入力 \mathbf{u} で決まる外力を表す。第4, 5項はそれぞれ物体とロボットの衝突, 壁反力のエネルギーである。

物体 B を古典力学で, 一方 R のロボット群を量子力学的な物質波で表せれば, ロボット群による物体搬送を波と物体の相互作用の形で表現できる。このために以下で全体系のシュレディンガー方程式から出発し, エーレンフェストの定理⁶⁾に基づいて波動関数にわたる期待値を取ることで B に対するニュートンの運動方程式を計算する。一方で B の座標のみにわたる期待値を取ると, これはロボット R の群の物質波の方程式を与える。

物理量 $\hat{\omega}$ の期待値 $\langle \omega \rangle \equiv \int d^2 \vec{X} \int d^2 \vec{x} \Psi^* \hat{\omega} \Psi$ と B の運動量の演算子表現 $\hat{P} = -iH_R \frac{\partial}{\partial \vec{X}}$ を使うと, 物体の運動方程式として

$$\begin{aligned} M \frac{d^2 \langle \vec{X} \rangle}{dt^2} &= \frac{d}{dt} \int d^2 \vec{X} \int d^2 \vec{x} \Psi^* \left(-iH_R \frac{\partial}{\partial \vec{X}} \right) \Psi \\ &= \left\langle -\frac{\partial}{\partial \vec{X}} \left(V_B(|\vec{x} - \vec{X}|) + V_{B0}(\vec{X}) \right) \right\rangle \quad (2) \end{aligned}$$

を得る。ここでは, 部分積分において領域を囲む壁表面での $\Psi = \frac{\partial \Psi}{\partial \vec{x}} = 0$ の境界条件を使う。また(2)では $\frac{\partial V_{\text{cnt}}}{\partial \vec{x}} = 0$ に従う。なお本稿では, 波が物体を実際に駆動するかどうか, に焦点をあてる。このため, モーメントによる物体の回転は, 単に単純化という理由で,

計算対象としない．また同じ理由でロボットは質点として扱い，その大きさはロボットに対する壁反力のパラメータとしてのみ考慮される．

全系のシュレディンガー方程式 (1) を物体座標 \vec{X} にわたり積分する．物体の運動エネルギー項 $-\frac{H_R^2}{2M} \frac{\partial^2 \Psi}{\partial \vec{X}^2}$ は発散であるから $\int d^2 \vec{X}$ の下でゼロになる．ゆえにロボット側のみの波動関数を

$$\psi(\vec{x}; t) \equiv \int d^2 \vec{X} \Psi(\vec{X}, \vec{x}; t) \quad (3)$$

と定義すれば，それが従う波動方程式は

$$iH_R \frac{\partial \psi}{\partial t} = -\frac{H_R^2}{2m} \frac{\partial^2 \psi}{\partial \vec{x}^2} + \int d^2 \vec{X} (V_B + V_{B0}) \Psi + V_{cnt} \psi \quad (4)$$

となる．

物体の量子揺らぎを無視し物体側の規格化定数を N ($\int_{\Omega} d^2 \vec{X} N \delta(\vec{X} - \vec{X}(t)) = 1$) として，全系の波動関数を，

$$\Psi(\vec{x}, \vec{X}; t) = N \delta(\vec{X} - \vec{X}(t)) \times \psi(\vec{x}; t) \quad (5)$$

の変数分離形に仮定する．ここで $\vec{X}(t) \equiv \langle \vec{X} \rangle$ ．この仮定の下で (2) は，

$$\begin{aligned} \frac{d}{dt} \vec{P}(t) &= -\frac{\partial}{\partial \vec{X}(t)} V_{B0}(\vec{X}(t)) \\ &- \frac{\partial}{\partial \vec{X}(t)} \int d^2 \vec{x} \psi^*(\vec{x}; t) V_B(|\vec{x} - \vec{X}(t)|) \psi(\vec{x}; t) \end{aligned} \quad (6)$$

である．ここで $d^2 \vec{x} \psi^* \psi$ は，ロボットが平面領域 $d^2 \vec{x}$ に存在する確率である．すなわち (6) 右辺第 2 項は，力 $-\frac{\partial}{\partial \vec{X}} V_B(|\vec{x} - \vec{X}|)$ に確率の重みをつけた加算平均である．次に変数分離の仮定 (5) の考慮と (4) の積分 $\int d^2 \vec{X}$ の実行は次の波動方程式を導く．

$$iH_R \frac{\partial \psi}{\partial t} = -\frac{H_R^2}{2m} \frac{\partial^2 \psi}{\partial \vec{x}^2} + \left((V_{cnt}(\vec{x}; \mathbf{u}) + V_B(|\vec{x} - \vec{X}(t)|) + V_{B0}(\vec{X}(t))) \right) \psi, \quad (7)$$

ここで注意すべきは，物体とロボットの相互作用 V_B が右辺に入るのは当然のこととして，(ロボット位置には依存しないはずの) 物体 $\vec{X}(t)$ の受ける壁反力 V_{B0} も考慮されるべきことである．

以上の (6), (7) が，物質波としてのロボット群による物体運動の制御を決める状態方程式となる．すなわち (7) によりロボットの波動関数 ψ が制御入力 \mathbf{u} で制御され，その結果を受けて (6) で $\vec{X}(t)$ が制御される．

3 波による物体駆動

計算パラメータは表 1 に示すとおりである．SI 単位を取る．辺長 2 の正方形領域の中で，中央に置いた一つの物体 (径 $2 \times R_B = 0.5$ ，質量 $M = 37.5$) を，その周囲のどこか \vec{x}_R に置いたと想定するロボット (径 $a_R = 0.1$ ，質量 $m = 0.3^8$) の衝突により搬送する．「 \vec{x}_R に置い

た想定」という意味は，波によって物体を動かすのであるが，その波の初期値 (初期分布) が位置 \vec{x}_R に中心を持つガウス波束である，という意味である．ガウス波束の広がりにはロボット径 a_R とする．ロボットの波動

Table 1: Parameters

meaning	parameter	value
geometry	$[-S_1, S_1]$	$S_1 = 1[\text{m}]$
	$\times [-S_2, S_2]$	$S_2 = 1[\text{m}]$
number of robots	N_0	4[-]
radius	robot: $\frac{a_R}{2}$	0.05[m]
	object: R_B	0.25[m]
mass	robot: m	0.3[kg]
	object: M	37.5[kg]
coefficient of restitution	e	1[-]
collision potential ⁷⁾	σ_S	4[J]
	σ_v	4[J]
	n_S	12[-]
	n_v	12[-]
wall reaction	c_R	4[J]
	c_B	4[J]
	n_{cR}	12[-]
	n_{cB}	12[-]
friction coefficients	robot: γ	0[kg/s]
	object: δ	0[kg/s]
quantum fluctuation	H_R	0.1[J · s]

方程式 (7) は有限差分法で計算し，とくに時間更新は $O(dt^2)$ まで正しい方法⁹⁾ を使う．

3.1 ポテンシャル外場が作用しないとき: $V_{cnt} = 0$

外場 V_{cnt} が作用しないときでも，波は初期分布から拡散する．拡散した波は物体に到達し，これを動かすことができる．波の拡散と物体がそれに伴い動く様子を，ロボットが 1 台のときに，図 1 に示す．その駆動

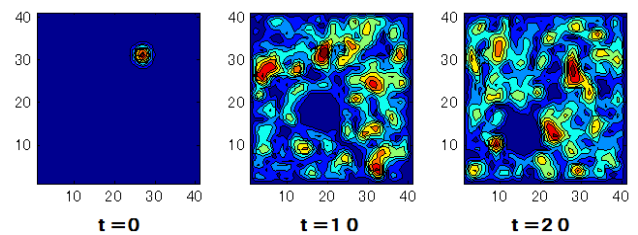


Fig. 1: 外場がない ($V_{cnt} = 0$) ときの波と物体の動き．ロボット $N_R = 1$ 台．

力は波の揺らぎを表すパラメータ H_R で制御できるが，本稿では表 1 のとおり $H_R = 0.1$ としている．ロボット台数が増えると，このような波による駆動の結果として，物体はどんな運動をするだろうか？ その様子を図 2 で示す．このうちの青実線で示される $N_R = 1$ 台 ($\vec{x}_R = [0.5, 0.3]$) では，波の裏側への回り込みが遅く物体は主として左下方向に駆動され続ける．いっぽう $\vec{x}_R = [-0.25, 0.1]$ を追加した $N_R = 2$ (緑破線)，さらに 3 台目を $\vec{x}_R = [-0.5, -0.6]$ に置いた場合 (赤点線) あるいは 4 台目を $\vec{x}_R = [0.7, -0.5]$ とした場合 (水色一点鎖

線) では, X_1 は波の回り込みの効果で右方向への駆動が見られる. しかし X_2 は系統的な上方向への運動は発生していない. それでは, 搬送の要求値 $\vec{X}^{req}(t)$ が

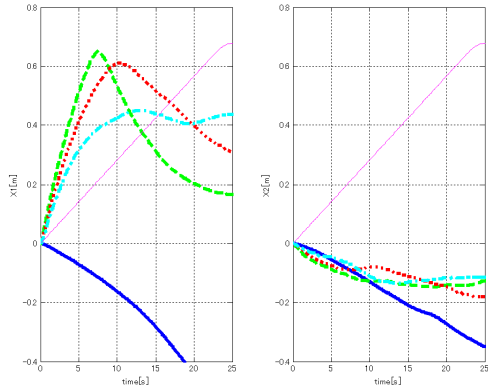


Fig. 2: 外場 $V_{cnt} = 0$ での物体の搬送.

図 2 のマゼンタ色の細実線であって, これにトラッキングする必要があるとすれば, どのように波の運動を制御すればよいか? それを次に考える.

3.2 外場 V_{cnt} のフィードバックによる生成

波を積極的に制御するためには, 外場 V_{cnt} を操作する必要がある. ここではもっとも単純と言う意味で次の線形の形を取り,

$$V_{cnt}(\vec{x}) = a_1 x_1 + a_2 x_2 \quad (8)$$

パラメータ a_1 と a_2 を操作入力 u とする. この a_1, a_2 については, これら先行値 (フィードフォワード) を \vec{X}^{req} から決め, それで不足する量をフィードバックで補うというのが本来の設計法である. しかし現状ではその先行値計算の方法が未だ無い. そこですべてをフィードバックでまかなう. すなわち

$$\vec{a} \equiv \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = C(\vec{X}^{req} - \vec{X}) \quad (9)$$

である. 一般には C は非対角行列であり, また比例成分のみならず微分や積分成分も含めて表現しているとする. ただし本稿では対角項のみの比例成分に限定し, a_1 と a_2 で比例定数も等しいとする.

$$a_i = P \times (X^{req}_i - X_i), \quad i = 1, 2. \quad (10)$$

ロボット台数によって設定すべき P 値がどう変化するか, を見る. 図 3 は, $N_R = 1$ 台で比例制御の定数が $P = 20$ のときの, 波の伝播とそれによって動かされる物体のプロファイルを示す. これは時間トレンド

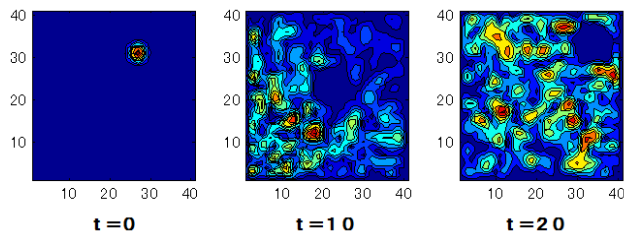


Fig. 3: フィードバックで決まる外場 (V_{cnt}) の下での波と物体の動き. ロボット $N_R = 1$ 台, 比例定数 $P = 20$.

で見ると図 4 の青実線で示される. 緑破線は $N_R = 2$,

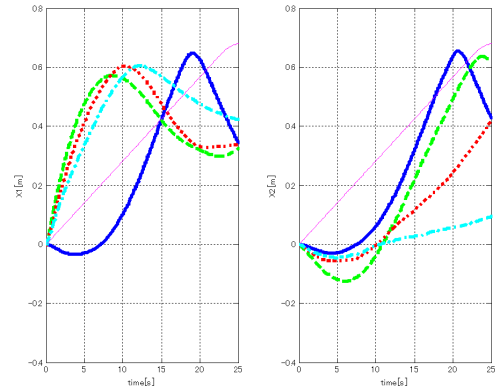


Fig. 4: フィードバックによる外場 V_{cnt} 下の物体の搬送.

$P = 10$, 赤点線は $N_R = 3$, $P = 5$ さらに水色一点鎖線は $N_R = 4$, $P = 2$ の計算結果である. 必ずしも完全なトラッキングは得られておらず, また $N_R = 4$ (水色一点鎖線) の X_2 のようにトラッキングの遅れが目立つものもある. しかし定性的には, N_R が大きいほど比例定数 P が小さくてもよい, と言える. なおここで注意すべきは比例定数 P が正值であればマゼンタ色細実線の \vec{X}^{req} に追従できることである. 波はある程度の時間遅れを伴って物体に到達するから, このような目標値に対しては, 最初のうちは X_1 も X_2 も目標値より小さい. 従って (10) で $P > 0$ なら a_1, a_2 は正になる. ということはポテンシャルの形を考えれば分かる通り, 通常の質点力学であればこれは $F_1 = -a_1, F_2 = -a_2$ という「左下向き」の力 \vec{F} を与えるはずである. しかし逆に正值の a_i により波が「左下向き」に駆動される結果として, 物体裏側への回り込みが発生する. このため物体を「右上方向に」, すなわち \vec{X}^{req} に追従する方向に駆動できるのである. この質点力学の予想とは逆方向に物体が駆動される現象は, 著者らの従来研究でも見られたものである^{2, 3, 4}. すなわちこれは連続体の「波の回り込み」という共通要因を表現している.

4 メカニカルな機構による物体駆動

複素波動関数 ψ に従う波動の場を何らかの媒体にわたり実際に生成することは困難である. しかし, 質点の確定的な力学による量子力学の表現^{10, 11, 12}) によれば, ロボットに揺らぎ力が作用することが, ψ の存在と同じ効果をもたらす. すなわちロボットのシュレディンガー方程式 (7) は, 実関数 C, S^q を使って $\psi = e^C e^{i\frac{S^q}{\hbar R}}$ と局座標表示することで次の運動方程式と等価になる.

$$m \frac{d^2 \vec{x}}{dt^2} = -\frac{\partial}{\partial \vec{x}} \left((V_{cnt}(\vec{x}; u) + V_B(|\vec{x} - \vec{X}(t)|) + V^q[\psi(\vec{x}; t)]) \right) \quad (11)$$

$$V^q = -\frac{\hbar R^2}{2m} \left(\vec{\nabla}^2 C + (\vec{\nabla} C)^2 \right), \quad (12)$$

ここで右辺第 3 項が量子揺らぎ力である. また物体の壁反力は \vec{x} に依存しないのでその偏微分はゼロとした. また揺らぎポテンシャルの計算は¹³) に従う.

本来は波動関数 ψ が与える確率密度に応じたロボット台数を平面上の各点に配置し, その全台数 N_R がど

の程度なら波による駆動と同様の結果になるか、を見るべきである。しかし ψ はガウス波束で与えているため、この事情を再現するには N_R として非現実的な台数値をとることになる。そこでここではロボット台数 N_R に応じてそれぞれの波動関数 (波束) を設定することとする。

ロボットには図 5 に示すように量子ダイナミクスによる力が加算される。なお本図は例を示すために弱いフィードバックによる V_{cnt} をかけている。外場がない ($V_{cnt} = 0$) なら量子ダイナミクスが作用しない限りロ

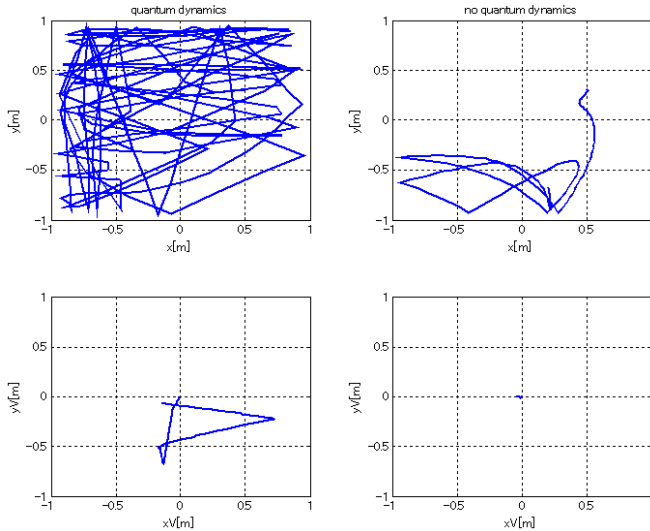


Fig. 5: ロボット $N_R = 1$ 台の閉平面内の動き。外場 V_{cnt} は弱いフィードバック ($P = 0.1$) で決まる。

ボットはまったく動かない。これでは量子ダイナミクスの「大きさ」を見ることができないため、 $P = 0.1$ として外場を弱く引加している。これにより、 V_{cnt} だけでなく量子ダイナミクスを受けるロボットは、領域全体を動き回りこの運動が物体を動かす。すなわち外場がないときでも、ロボットは量子ダイナミクスにより物体を動かす。そのときの物体の搬送トレンドを図 6 に、左からロボット台数を $N_R = 1, 2, 3, 4$ として示す。図中の赤点線はそれぞれ波による駆動の結果である。またマゼンタ細実線は搬送の要求値を示す。このときロボットに作用する力はどの程度であるか？これを図 7 に、 $N_R = 1$ 台の場合のロボットの動きとともに示す。図中の青実線が壁反力 (縦軸ゼロ点は $+10[N]$ で表示)、緑破線は物体との衝突力 (縦軸は $-10[N]$ 表示) を与える。量子ダイナミクスによる力は赤点線である。では外場をフィードバックで与えるときはどうなるか？これを図 8 に示す。やはり左から $N_R = 1, 2, 3, 4$ 台のロボットを設定するものとしている。このときのロボットに作用する力を、 $N_R = 1$ 台のケースで示すと図 9 のとおりである。ここではマゼンタ実線 (縦軸原点を $-20[N]$ 表示) で示されるような外場 V_{cnt} による力 $-\frac{\partial V_{cnt}}{\partial \vec{x}}$ も存在する。

5 まとめと議論

本稿では、相互通信も外界センサも搭載しないロボット群による巨視的物体の搬送のひとつの拡張形を示した。すなわち物体は波で搬送され、その波は媒体を必要とせずロボットというメカニカル機構で表現される。またその台数は十分に量子力学的確率を再現すべき、という原則で決めることができる。この点で著者らのポ

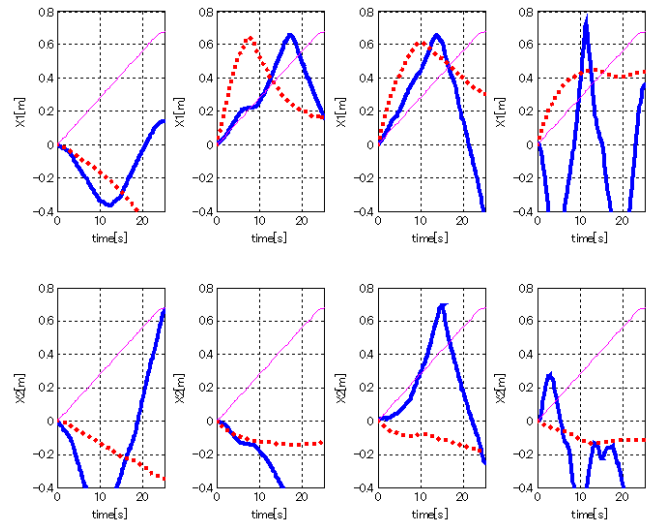


Fig. 6: 量子ダイナミクスに従うロボット (左から $N_R = 1, 2, 3, 4$) による物体の搬送。外場は作用しない; $V_{cnt} = 0$ 。

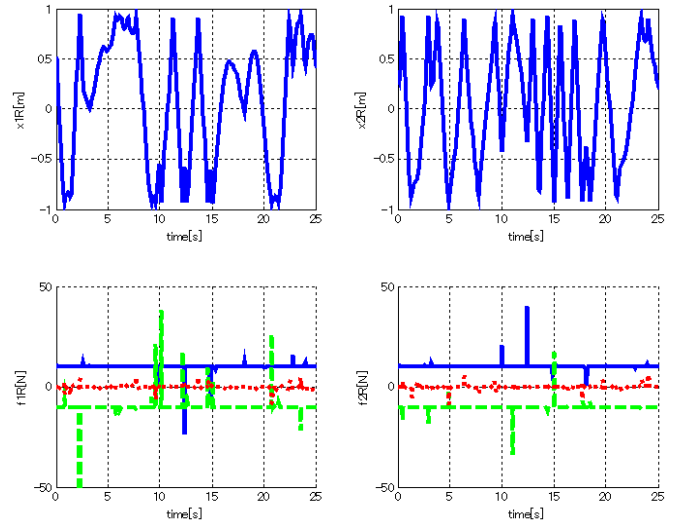


Fig. 7: ロボット $N_R = 1$ 台に作用する力の時間トレンド。外場は作用しない; $V_{cnt} = 0$ 。

ルツマン分布に基づく従来研究^{2, 3, 4)}に比べ、より系統的な台数設定が可能となる。波および対応するロボットによる物体搬送のシミュレーション計算も示した。ただしロボットには V^q による力を感じるための、位置 \vec{x} のセンサ搭載が必要となる。また実験検証が必要なのはいうまでもなく、単体ロボットとしてボール型⁸⁾あるいはディスク型¹⁴⁾を設計・製作中である。本研究の一部は科研費 (基盤研究 (C)16K00337) の助成を受けたものである。

参考文献

- 1) R.Pfeifer and J.Bongard, *How the Body Shapes the Way We Think: A New View of Intelligence*, Bradford(2006).
- 2) T.Itami, Macroscopic Group Robots inspired By “Brownian Motion,” in *Handbook of Research on Design, Control, and Modeling of Swarm Robotics*, 2016.
- 3) 伊丹, 相互通信と外界センサを持たないロボット少数台が示す “知的” 挙動, 第 8 回 CI 研究会 (広島), 2015.
- 4) 伊丹, 連続体描像に基づくロボット群の集団運動の解析, *Trans.SICE*,43-3,2012.
- 5) P.Hänggi and F.Marcheson, Artificial Brownian Motors: controlling transport on the nanoscale. *Review*

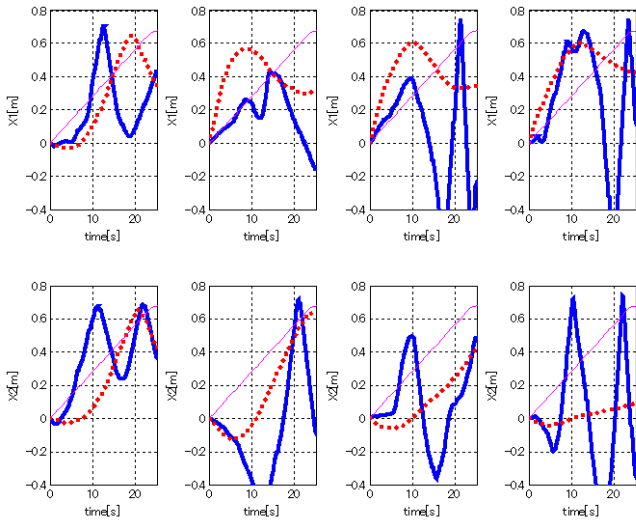


Fig. 8: 量子ダイナミクスに従うロボット (左から $N_R = 1, 2, 3, 4$) による物体の搬送. 外場 V_{cnt} はフィードバックで決まる.

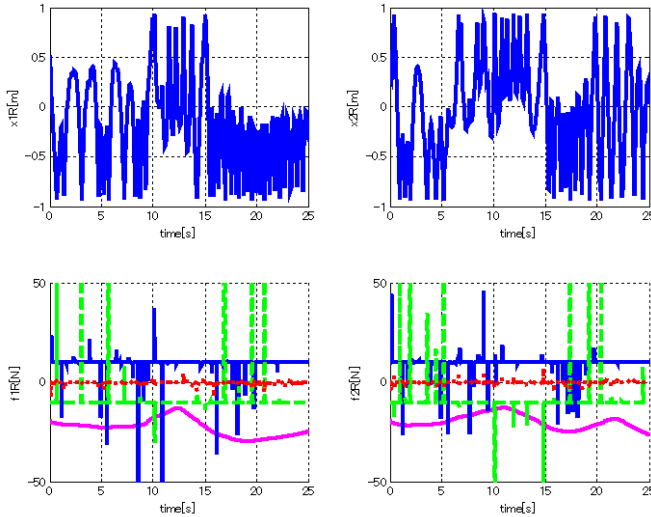


Fig. 9: ロボット $N_R = 1$ 台に作用する力の時間トレンド. 外場 V_{cnt} はフィードバックで決まる.

of *Modern Physics*, **81**, 383-442(2009).

- 6) L.I.Schiff, *Quantum mechanics*, McGraw-Hill(1968).
- 7) グールド, トポチク, 計算物理学入門, ピアソンエデュケーション 2000.
- 8) <http://www.elekit.co.jp/product/4d522d39383032.tama>
- 9) 飯高, 量子ダイナミクス入門 (パリティ物理学コース), 丸善, 1994.
- 10) D.Bohm, "A suggested interpretation of the quantum theory in terms of 'hidden' variables," *Phys.Rev.* **85**, 166-193(1952).
- 11) P.R.Holland, *The quantum theory of motion* (Cambridge niv. Press, 1995)
- 12) M.Gondran and A.Gondran, "Numerical simulation of the double slite interference with ultracold atoms," *Am.J.Phys.* **73**(6), 507-515(2005)
- 13) P.A.Machado, *Computational Approach to Bohm's Quantum Mechanics*, Dr.Thesis, McMaster niversity, 2007.
- 14) 井出, ボールで動くロボットの製作 (サイエンス・シリーズ), パワー社, 2001.

- 15) 伊丹, 松井, 乾, 全, 量子力学的手法によるシステムと制御 (計測・制御テクノロジーシリーズ第6巻), コロナ, 2017.

A ニュートン力学による波動の表現

特定の形を持つ揺らぎポテンシャルの作用下においては, N 質点のニュートン力学が量子力学の波動を表現する. ここではこの等価性の説明を与え, また古典的物体があつて境界条件が変化する系の取り扱いを確認する.

対象系は n 次元空間中の N 個の質点群である. それぞれの質点の質量は m とし, 外場ポテンシャル V が作用する. また質点は互いに衝突するが, これは単に後述の式を簡単にするため 2 体間のみとし, ポテンシャル V_v で表現する. この質点群に「揺らぎポテンシャル」として, パラメータ H_R で大きさが制御され実関数 $C(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ の微分 ($\vec{\nabla}_i \equiv \frac{\partial}{\partial \vec{x}_i}$) から作られる

$$V^q(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \equiv -\frac{H_R^2}{2m} \sum_i \left(\vec{\nabla}_i^2 C + (\vec{\nabla}_i C)^2 \right) \quad (13)$$

を作用する. ここで $C = \ln R(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ となる正値関数 R を使うと (13) は次のようにも表現できる.

$$V^q = -\frac{H_R^2}{2m} \frac{1}{R} \sum_i \vec{\nabla}_i^2 R \quad (14)$$

個々の質点 $i (i = 1, 2, \dots, N)$ の運動方程式は

$$m \frac{d^2 \vec{x}_i}{dt^2} = -\frac{\partial}{\partial \vec{x}_i} \left(V(\vec{x}_i) + \sum_{j \neq i} V_v(\vec{x}_j, \vec{x}_i) + V^q(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \right) \quad (15)$$

であるが, これをまとめて値関数 $S(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)$ によるハミルトン-ヤコビ方程式で表現すると次となる.

$$-\frac{\partial S}{\partial t} = \frac{1}{2m} \sum_i \left(\frac{\partial S}{\partial \vec{x}_i} \right)^2 + \sum_i V(\vec{x}_i) + \frac{1}{2} \sum_i \sum_{j \neq i} V_v(\vec{x}_i, \vec{x}_j) + V^q(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \quad (16)$$

すなわち質点の速度は以下のフィードバック則の形¹⁵⁾で与えられる.

$$\vec{v}_i \equiv \frac{d\vec{x}_i}{dt} = \frac{1}{m} \frac{\partial S(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N)}{\partial \vec{x}_i} \quad (17)$$

さてこの (16) に R を乗じると, (14) に注意して

$$-R \frac{\partial S}{\partial t} = \frac{R}{2m} \sum_i \left(\frac{\partial S}{\partial \vec{x}_i} \right)^2 + \left(\sum_i V(\vec{x}_i) + \frac{1}{2} \sum_i \sum_{j \neq i} V_v(\vec{x}_i, \vec{x}_j) \right) R - \frac{H_R^2}{2m} \sum_i \vec{\nabla}_i^2 R \quad (18)$$

そこでこの関数 R の時間発展が次のようなものであったと考える。

$$\frac{dR^2}{dt} = -\sum_i \vec{\nabla}_i \cdot \vec{v}_i R^2 \quad (19)$$

あるいは C を使えば

$$\frac{dC}{dt} = -\frac{1}{2} \sum_i \vec{\nabla}_i \cdot \vec{v}_i \quad (20)$$

ここで

$$\frac{d}{dt} = \frac{\partial}{\partial t} + \sum_i \frac{d\vec{x}_i}{dt} \cdot \frac{\partial}{\partial \vec{x}_i} = \frac{\partial}{\partial t} + \sum_i \vec{v}_i \cdot \frac{\partial}{\partial \vec{x}_i} \quad (21)$$

また (17) に注意して, (19) は次に導く。

$$\frac{\partial R}{\partial t} = -\frac{1}{m} \sum_i \frac{\partial R}{\partial \vec{x}_i} \cdot \frac{\partial S}{\partial \vec{x}_i} - \frac{R}{2m} \sum_i \vec{\nabla}_i^2 S \quad (22)$$

そして絶対値が R で位相が $\frac{S}{\hbar R}$ である複素数の波

$$\psi(\vec{x}_1, \vec{x}_2, \dots, \vec{x}_N) \equiv \text{Re} e^{i\frac{S}{\hbar R}} \quad (23)$$

を定義すれば, その時間空間変化を支配する複素方程式

$$i\hbar R \frac{\partial}{\partial t} \psi = -\frac{\hbar R^2}{2m} \sum_i \vec{\nabla}_i^2 \psi + \left(V(\vec{x}_i) + \frac{1}{2} \sum_{j \neq i} V_v(\vec{x}_j, \vec{x}_i) \right) \psi \quad (24)$$

が, その両辺を $e^{i\frac{S}{\hbar R}}$ で除した上で, (16) と (22) を, その実部と虚部としてまとめて表現していることは簡単な計算から分かる。すなわち (19) あるいは (20) と (15) は複素数の波動を表現するのである。

次に, 古典的物体があつて境界条件が変化する系の取り扱いを確認する。対象は, 物体が波あるいは波を表現するロボットと相互作用する系である。簡単のためロボットを 1 台として定式化する。時点 t での物体の位置を $\vec{X}(t)$ として, 2 つの基礎式のうち (15) は

$$m \frac{d\vec{v}}{dt} = -\frac{\partial}{\partial \vec{x}} \left(V(\vec{x}, \vec{X}(t)) + V^q(\vec{x}, \vec{X}(t)) \right) \quad (25)$$

また (20) は

$$\frac{dC(\vec{x}, \vec{X}(t))}{dt} = -\frac{1}{2} \vec{\nabla} \cdot \vec{v} \quad (26)$$

である。いずれの式においても t のあらわな依存性は表記しておらず物体の位置が t に依存することのみを強調している。次に物体側であるが, ふたつの可能性がある。ひとつは波の作用により運動が決まるとするもので, それは

$$M \frac{d\vec{V}}{dt} = \left\langle -\frac{\partial V(\vec{x}, \vec{X}(t))}{\partial \vec{X}(t)} \right\rangle_{wave} \quad (27)$$

である。ここで $\int d^2 \vec{x} R^2 = 1$ という規格化条件の下で

$$\langle \dots \rangle_{wave} \equiv \int d^2 \vec{x} R^2 (\dots) \quad (28)$$

と定義される。またもうひとつはロボットの衝突によるものでそれは

$$M \frac{d\vec{V}}{dt} = -\frac{\partial V(\vec{x}, \vec{X}(t))}{\partial \vec{X}(t)} \quad (29)$$

のように $\langle \dots \rangle_{wave}$ をはずしたものである。

肘付き制御によるマニピュレータの作業精度と エネルギー効率の改善

○李 想 見浪 護 (岡山大学)

Merits Concerning Energy-consumption and Trajectory-tracking Accuracy Derived from Elbow-bracing Manipulator

*X.Li M.Minami (Okayama University)

Abstract— Considering that humans perform handwriting task with reduced powers by contacting elbow or wrist on a table, it is reasonable to deem that manipulators can save energy and simultaneously accomplish writing tasks precisely like humans by bracing intermediate links such as elbow or wrist. First this paper discusses equation of motion of robot under bracing constraint condition, based on the robot's dynamics with constraint condition including motor dynamics. Then a control method to utilize the constraint dynamics is proposed to control simultaneously bracing force and hand's trajectory in work space. Even though the model used for the simulation analyses is simple four-links manipulator, the simple structure can help understand even more clearly the effects got by bracing part of the manipulator. This paper demonstrates the merits of the strategy to utilize bracing by comparing the contacting motions with non-contacting motions, including; (1) the energy consumption can be reduced; (2) the hand trajectory tracking becomes accurate; (3) there is an optimum contacting point that minimize the energy consumption on condition that trajectory-tracking task be given to the hand.

Key Words: Bracing manipulator, Energy-consumption, Numerical simulation

1 緒言

運動学に基づいた冗長マニピュレータは幅広く研究され、その効果は Chirikjian と Burdick¹⁾ により紹介されている。しかし、高い冗長性を持つマニピュレータはリンク数が増えるにつれ重量が増大する。そのため超冗長マニピュレータは冗長性を利用して形状を多様に変更できる反面、ハンドの可搬重量は制限されていた。多くの研究では障害物回避^{2,3)} や形状の最適化などの冗長性を利用する方法⁴⁾ が議論されているが、现阶段では超冗長マニピュレータは実用的レベルには達していない。

Fig.1 に示すように、人間は字を書く動作において手首や肘を机につくことで、より少ない力で正確に字を書けることを経験的に知っている。冗長マニピュレータでも肘を環境に対して拘束することで省エネルギーでしかも高精度の作業が実現できる可能性がある。そこで、次に環境に拘束されて運動するマニピュレータについて考える。

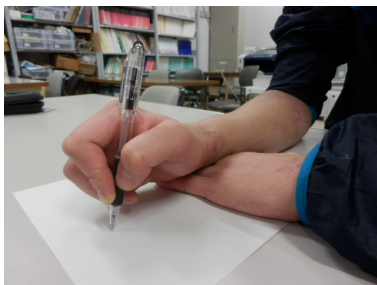


Fig. 1: Human's writing motion

Whitcomb⁵⁾ は、ロボットの拘束状態の運動と制御についての確かな分類を行った。すなわち (a): 変形しないロボットと変形する環境を前提にしたモデルベース制御^{6,7)}, (b): 位置/速度制御をベースにした変形しないロボットと変形する環境を前提にした制御⁸⁾ という分

類である。また Oussama Khatib は (b) の範疇で、接触動作を制御するために複数点接触の運動学モデル^{9,10)} を提案した。最後に (c): 変形しないロボットと変形しない環境を前提にした制御法¹¹⁾ という分類もある。山根と中村は、この枠組みの中でヒューマノイドの歩行¹²⁾ とダイナミクスフィルターの概念を提案している¹³⁾。

以上はロボットの肘付き運動を議論するための、超冗長マニピュレータと拘束状態の運動についての今までの研究の流れである。これまでに環境に接触している冗長マニピュレータの有効性と精度を議論した研究例は存在する。West と Asada¹⁴⁾ は拘束されたマニピュレータの位置/力同時制御コントローラを設計するための一般的な運動学的接触モデルを提案した。

本報では、(c) の変形しないロボットと変形しない環境を前提にしたロボット^{11,12,13)} の制御方法について考える。この前提条件のもとでは、拘束条件と運動方程式から (1) 式に示す代数方程式が得られる。

$$A f_n = a - B \tau \quad (1)$$

f_n は抗力、 A と a 、 B は次章で定義されるベクトルと行列、 τ は入力トルクベクトルである。式 (1) はロボットの手先が拘束されているときの入力トルクと抗力の代数関係を表している。上式は、二足歩行の分野で Hemami¹⁵⁾ が導出し、ロボティクスの分野では Peng¹⁶⁾ により最初に適用された。Peng は τ を入力とし、 f_n を出力と考えることで、式 (1) を力センサーとして用いている。本論文では逆に目標抗力 f_{nd} を達成するための入力トルク τ を計算するために用いる。

著者らは、環境と接触するロボット作業について、抗力を利用することで重力の影響を抑え、より少ない消費エネルギーで高精度の作業を実行できる冗長マニピュレータを提案してきた¹⁷⁾。この結果に基づいて本報では講演発表会^{18,19,20)} で発表してきた肘付き拘束状態のロボットの制御について抗力、ハンド位置および肘位置を同時に制御する制御方法を提案し、ハンド軌道

追従特性と消費エネルギーの面から肘付き運動の有効性について議論する。さらにエネルギー最少規範に基づく肘付き位置とハンド負荷の影響に関する考察を行い、最適肘付き位置と目標ハンド軌道および手先負荷との関係を議論する。

2 拘束運動のモデリング

2.1 肘付き拘束運動

本節では複数点拘束を表現するモデリングを行う。\$n\$ リンクマニピュレータの中間リンクが \$p\$ 個の拘束

$$\begin{aligned} C(\mathbf{r}(\mathbf{q})) &= [C_1(\mathbf{r}_1(\mathbf{q})), C_2(\mathbf{r}_2(\mathbf{q})), \dots, C_p(\mathbf{r}_p(\mathbf{q}))]^T \\ &= \mathbf{0} \end{aligned} \quad (2)$$

を受けているときの運動方程式について考える。ここで \$\mathbf{q}\$ は関節角ベクトルであり、\$\mathbf{r}_i\$ は拘束を受けている第 \$i\$ リンクの位置である。\$\mathbf{r}_i\$ と \$\mathbf{q}\$ の関係及び \$\dot{\mathbf{r}}_i\$ と \$\dot{\mathbf{q}}\$ の関係を以下に求める。

$$\mathbf{r}_i = \mathbf{r}_i(\mathbf{q}) \quad (3)$$

$$\dot{\mathbf{r}}_i = \mathbf{J}_i(\mathbf{q})\dot{\mathbf{q}}, \quad \mathbf{J}_i(\mathbf{q}) = [\tilde{\mathbf{J}}_i(\mathbf{q}), \mathbf{0}] \quad (4)$$

式(4)では、\$\mathbf{J}_i\$ は \$m \times n\$ の行列、\$\tilde{\mathbf{J}}_i\$ は \$m \times i\$ の行列、ゼロ部分行列 \$\mathbf{0}\$ は、\$m \times (n - i)\$ である。

本報では、拘束条件が複数存在する場合を考え、抗力の作用方向を表す係数ベクトルと摩擦力の係数ベクトルはそれぞれ、

$$\left(\frac{\partial C_i}{\partial \mathbf{q}^T} \right)^T / \left\| \frac{\partial C_i}{\partial \mathbf{r}^T} \right\| = \mathbf{j}_{ci}^T \quad (5)$$

$$\left(\frac{\partial \mathbf{r}_i}{\partial \mathbf{q}^T} \right)^T \frac{\dot{\mathbf{r}}_i}{\|\dot{\mathbf{r}}_i\|} = \mathbf{j}_{ti}^T \quad (6)$$

の \$n \times 1\$ のベクトルとなる。さらに

$$\mathbf{J}_c^T = [\mathbf{j}_{c1}^T, \mathbf{j}_{c2}^T, \dots, \mathbf{j}_{cp}^T] \quad (7)$$

$$\mathbf{J}_t^T = [\mathbf{j}_{t1}^T, \mathbf{j}_{t2}^T, \dots, \mathbf{j}_{tp}^T] \quad (8)$$

$$\mathbf{f}_n = [f_{n1}, f_{n2}, \dots, f_{np}]^T \quad (9)$$

$$\mathbf{f}_t = [f_{t1}, f_{t2}, \dots, f_{tp}]^T \quad (10)$$

と定義する。\$\mathbf{J}_c^T\$、\$\mathbf{J}_t^T\$ は \$n \times p\$ 行列であり、\$\mathbf{f}_n\$、\$\mathbf{f}_t\$ は \$p \times 1\$ のベクトルである。これらを考慮すると、\$p\$ 個の点でリンクが拘束されているマニピュレータの運動方程式は、

$$\begin{aligned} M(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) + D\dot{\mathbf{q}} \\ &= \boldsymbol{\tau} + \sum_{i=1}^p (\mathbf{j}_{ci}^T \mathbf{f}_{ni}) - \sum_{i=1}^p (\mathbf{j}_{ti}^T \mathbf{f}_{ti}) \\ &= \boldsymbol{\tau} + \mathbf{J}_c^T \mathbf{f}_n - \mathbf{J}_t^T \mathbf{f}_t \end{aligned} \quad (11)$$

となる。また、式(2)を時間 \$t\$ で2回微分し、\$\ddot{\mathbf{q}}\$ の拘束条件を求めると、

$$\dot{\mathbf{q}}^T \left[\frac{\partial}{\partial \mathbf{q}} \left(\frac{\partial C}{\partial \mathbf{q}^T} \right) \right] \dot{\mathbf{q}} + \left(\frac{\partial C}{\partial \mathbf{q}^T} \right) \ddot{\mathbf{q}} = \mathbf{0} \quad (12)$$

が得られる。マニピュレータが常に拘束面に拘束されるためには、式(11)の解 \$\mathbf{q}(t)\$ が時間 \$t\$ に無関係に式(2)

を満たさなければならない。式(2)の時間微分によって得られた式(12)を満たす \$\ddot{\mathbf{q}}\$ と式(11)の \$\ddot{\mathbf{q}}\$ が同じ値をとるとき、式(11)の \$\mathbf{q}(t)\$ は式(2)を満たすことになる。

ここで抗力 \$\mathbf{f}_n\$ と摩擦力 \$\mathbf{f}_t\$ について、動摩擦係数 \$K_i\$ を用いて

$$\begin{aligned} \mathbf{f}_t &= \mathbf{K} \mathbf{f}_n, \quad \mathbf{K} = \text{diag}[K_1, K_2, \dots, K_p] \quad (13) \\ 0 &< K_i < 1, (i = 1, 2, \dots, p) \end{aligned}$$

と表される²¹⁾から、式(11)は以下ようになる。

$$\begin{aligned} M(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) + D\dot{\mathbf{q}} \\ &= \boldsymbol{\tau} + (\mathbf{J}_c^T - \mathbf{J}_t^T \mathbf{K}) \mathbf{f}_n \end{aligned} \quad (14)$$

2.2 肘付き時の抗力の導出

本節では \$\mathbf{f}_n\$ の導出方法について述べる。まず、式(11)、(12)から \$\ddot{\mathbf{q}}\$ を消去し、\$(\partial C / \partial \mathbf{q}^T) M^{-1} (\partial C / \partial \mathbf{q}^T)^T = M_c\$ と置くと、

$$\begin{aligned} M_c \mathbf{f}_n &= \left\| \frac{\partial C}{\partial \mathbf{r}^T} \right\| \left(\frac{\partial C}{\partial \mathbf{q}^T} \right) M^{-1} (\mathbf{J}_t^T \mathbf{K} \mathbf{f}_n + D\dot{\mathbf{q}} + \mathbf{h} \\ &\quad + \mathbf{g} - \boldsymbol{\tau}) - \left\| \frac{\partial C}{\partial \mathbf{r}^T} \right\| \dot{\mathbf{q}}^T \left[\frac{\partial}{\partial \mathbf{q}} \left(\frac{\partial C}{\partial \mathbf{q}^T} \right) \right] \dot{\mathbf{q}} \end{aligned} \quad (15)$$

が得られる。さらに

$$\mathbf{B} = \left\| \frac{\partial C}{\partial \mathbf{r}^T} \right\| \left(\frac{\partial C}{\partial \mathbf{q}^T} \right) M^{-1} \quad (16)$$

$$\mathbf{a} = \mathbf{B} \{ D\dot{\mathbf{q}} + \mathbf{h} + \mathbf{g} \} - \left\| \frac{\partial C}{\partial \mathbf{r}^T} \right\| \dot{\mathbf{q}}^T \left[\frac{\partial}{\partial \mathbf{q}} \left(\frac{\partial C}{\partial \mathbf{q}^T} \right) \right] \dot{\mathbf{q}} \quad (17)$$

と置くと、式(15)は、

$$M_c \mathbf{f}_n = \mathbf{B} \mathbf{J}_t^T \mathbf{K} \mathbf{f}_n - \mathbf{B} \boldsymbol{\tau} + \mathbf{a} \quad (18)$$

となる。さらに

$$\mathbf{A} = M_c - \mathbf{B} \mathbf{J}_t^T \mathbf{K} \quad (19)$$

と置くことにより、式(1)が得られる。抗力 \$\mathbf{f}_n\$ と入力トルク \$\boldsymbol{\tau}\$ の関係は代数方程式で表されることがわかる。\$\mathbf{f}_n\$ は \$p\$ 次元ベクトルであり、\$\boldsymbol{\tau}\$ は \$n\$ 次元ベクトルで \$n > p\$ であるから、\$\mathbf{f}_n\$ を実現する \$\boldsymbol{\tau}\$ には拘束冗長性がある。

2.3 ロボット・モータの連立方程式

モータの電流をベクトル \$\mathbf{I}\$ で表し、式(14)にモータのダイナミクスを加えると次式となる。

$$\begin{aligned} (M(\mathbf{q}) + \mathbf{J}_m)\ddot{\mathbf{q}} + \mathbf{h}(\mathbf{q}, \dot{\mathbf{q}}) + \mathbf{g}(\mathbf{q}) + (D + D_m)\dot{\mathbf{q}} \\ &= \mathbf{K}_m \mathbf{I} + (\mathbf{J}_c^T - \mathbf{J}_t^T \mathbf{K}) \mathbf{f}_n \end{aligned} \quad (20)$$

ここで \$\mathbf{J}_m\$ はモータのロータの慣性モーメントを表す対角行列、\$\mathbf{D}_m\$ はモータの粘性抵抗係数行列、\$\mathbf{K}_m\$ はモータのトルク定数行列である。式(2)と式(11)の関係と同様に、式(12)を満たす \$\ddot{\mathbf{q}}\$ と式(11)の \$\ddot{\mathbf{q}}\$ が同じ値をとるとき、式(11)の \$\mathbf{q}(t)\$ は式(2)を満たすことになる。さらに次式のモータの印可電圧と電流との関係

$$L \frac{d\mathbf{I}}{dt} = \mathbf{v} - \mathbf{R}\mathbf{I} - \mathbf{K}_m \dot{\mathbf{q}} \quad (21)$$

を、式(12)および式(11)と連立させると、モータのダイナミクスを含み、 p 点で肘を拘束された n リンクマニピュレータの運動方程式は、次のように表される。

$$\begin{bmatrix} \mathbf{M} + \mathbf{J}_m & -(\mathbf{J}_c^T - \mathbf{J}_t^T \mathbf{K}) & \mathbf{0} \\ \frac{\partial \mathbf{C}}{\partial \mathbf{q}^T} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{L} \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}} \\ \mathbf{f}_n \\ d\mathbf{I}/dt \end{bmatrix} = \begin{bmatrix} \mathbf{K}_m \mathbf{i} - \mathbf{h} - \mathbf{g} - (\mathbf{D} + \mathbf{D}_m) \dot{\mathbf{q}} \\ -\dot{\mathbf{q}}^T \left[\frac{\partial}{\partial \mathbf{q}} \left(\frac{\partial \mathbf{C}}{\partial \mathbf{q}^T} \right) \right] \dot{\mathbf{q}} \\ \mathbf{v} - \mathbf{R}\mathbf{I} - \mathbf{K}_m \dot{\mathbf{q}} \end{bmatrix} \quad (22)$$

3 拘束冗長性を利用する肘付き位置/力制御

モータのダイナミクスを考慮しない場合、目標抗力 \mathbf{f}_{nd} を実現する解 $\boldsymbol{\tau}$ は式(1)より、

$$\boldsymbol{\tau} = \mathbf{B}^+(\mathbf{a} - \mathbf{A}\mathbf{f}_{nd}) + (\mathbf{I} - \mathbf{B}^+\mathbf{B})\mathbf{l} \quad (23)$$

となる。ただし \mathbf{B}^+ は \mathbf{B} の疑似逆行列である。 $\text{rank}(\mathbf{I} - \mathbf{B}^+\mathbf{B}) = n - p$ であり、 $\mathbf{I} - \mathbf{B}^+\mathbf{B}$ は無次元の行列であるから、 \mathbf{l} はトルクの次元を持つ。 \mathbf{l} を新しい入力とし、 \mathbf{B} の零空間 $\mathbf{I} - \mathbf{B}^+\mathbf{B}$ を通して手先目標軌道 \mathbf{r}_d を追従するためや肘付き位置の制御に使うことができる。疑似逆行列の性質によって、 \mathbf{l} に任意の入力を加えても \mathbf{f}_{nd} の実現には影響を与えない。よって \mathbf{l} をハンドの軌道追従などのタスクに用いることができ、このタスクと \mathbf{f}_{nd} を実現するタスクは非干渉化できる。

ここでは \mathbf{l} の決定方法について述べる。本報のシミュレーションでは、4リンクマニピュレータの4自由度をそれぞれ肘の抗力の制御に1自由度、肘の位置制御に1自由度、手先の位置制御に2自由度の合計4自由度全てを利用してマニピュレータの制御を行う。

$$\mathbf{l} = \tilde{\mathbf{J}}_{2y}^T [\mathbf{K}_{p2y}(y_{d2} - y_2) + \mathbf{K}_{d2y}(\dot{y}_{d2} - \dot{y}_2)] + \mathbf{J}_4^T [\mathbf{K}_{p4}(\mathbf{r}_{d4} - \mathbf{r}_4) + \mathbf{K}_{d4}(\dot{\mathbf{r}}_{d4} - \dot{\mathbf{r}}_4)] \quad (24)$$

ここで、 $\tilde{\mathbf{J}}_{2y}^T$ は式(4)で定義されている $\tilde{\mathbf{J}}_2^T$ を構成する列ベクトル $\tilde{\mathbf{J}}_2^T = [\tilde{\mathbf{j}}_{2y}^T, \tilde{\mathbf{j}}_{2z}^T]$ の中の第1列である。また、 \mathbf{K}_{p2y} 、 \mathbf{K}_{d2y} はそれぞれFig.2に示す第2リンクの y 軸方向の位置と速度の制御ゲインであり、 \mathbf{J}_4 は第4リンクのヤコビ行列であり、 \mathbf{K}_{p4} 、 \mathbf{K}_{d4} はそれぞれ第4リンクの位置と速度の制御ゲイン行列である。

式(23)は直接トルクを発生させることができるダイレクトモータで駆動されるロボットでは実現できるが、通常のモータの入力は電圧である。本報では、式(23)のコントローラの代わりに、次式を用いる。

$$\mathbf{v} = \mathbf{K}_v [\mathbf{B}^+(\mathbf{a} - \mathbf{A}\mathbf{f}_{nd}) + (\mathbf{I} - \mathbf{B}^+\mathbf{B})\mathbf{l}] \quad (25)$$

4 肘付きハンド軌道追従シミュレーション

この章では4リンクマニピュレータによるシミュレーションを行う。シミュレーションで用いるモデルをFig.2に示す。4リンクマニピュレータの物理パラメータは以下のように設定した。リンクの質量 $m_i = 1.0$ [kg]、長さ $l_i = 0.5$ [m]、関節の粘性摩擦係数 $D_i = 2.9$ [N·m·s/rad]、トルク定数 $K_i = 0.2$ [N·m/A]、抵抗 $R_i = 0.6$ [Ω]、インダクタンス $L_i = 0.1$ [H]、モータ

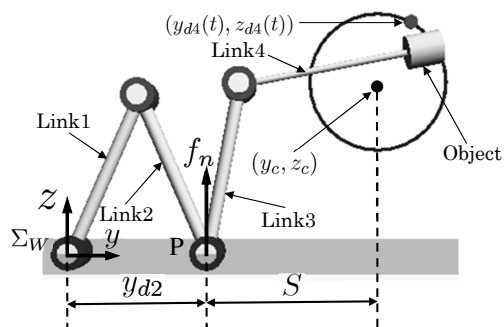


Fig. 2: Simulation model

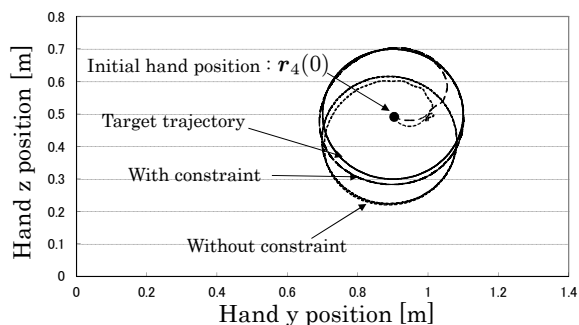


Fig. 3: Position time profile of hand

の慣性モーメント $I_{mi} = 1.64 \times 10^{-4}$ [kg·m²]、減速比 $k_i = 3.0$ 、減速機の粘性摩擦係数 $d_{mi} = 0.1$ [N·m·s/rad] ($i = 1, 2, 3, 4$)である。

また手先の目標軌道を次式で与える。

$$y_d(t) = 0.2 \cos \frac{2\pi}{10} t + y_c \quad (26)$$

$$z_d(t) = 0.2 \sin \frac{2\pi}{10} t + z_c \quad (27)$$

4.1 肘付きの有効性

この節では肘付きの有効性について示す。評価する指標として消費エネルギーと手先制御の精度を用いる。また、消費エネルギーの導出方法を以下に示す。時刻 $0 \sim T$ [s]間のマニピュレータの i リンクでのモータ消費エネルギーは次式で与えられる。

$$E_i(T) = \int_0^T v_i(t) I_i(t) dt \quad (28)$$

$$E_{sum}(T) = \sum_{i=1}^4 E_i(T) \quad (29)$$

ここでは、マニピュレータの第2関節を肘とみなし、肘を拘束する場合としない場合の2種類のシミュレーションを行った。Fig.2に示す目標軌道の中心位置を $(y_c, z_c) = (0.9, 0.5)$ [m]と設定し、初期ハンド位置を $(0.9, 0.5)$ 、初期肘付き位置を $(0.4, 0)$ と設定した。初期時刻 $t = 0$ から $t = 30$ [s]までの肘を付く場合と付かない場合それぞれの手先軌道と消費エネルギーをFig.3, Fig.4に示す。Fig.3より、肘を拘束すれば手先の精度は良くなるのが分かる、また、Fig.4より消費エネルギーは1/8~1/10程度削減できているのが分かる。つまり、肘付き運動が有効であることが分かる。

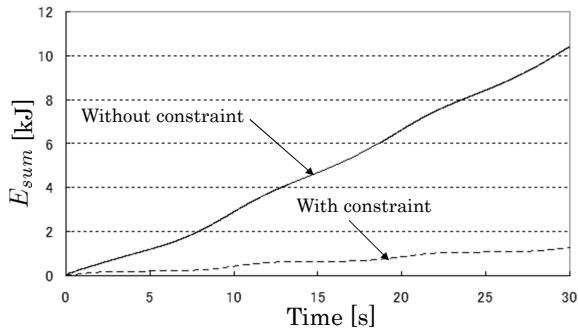
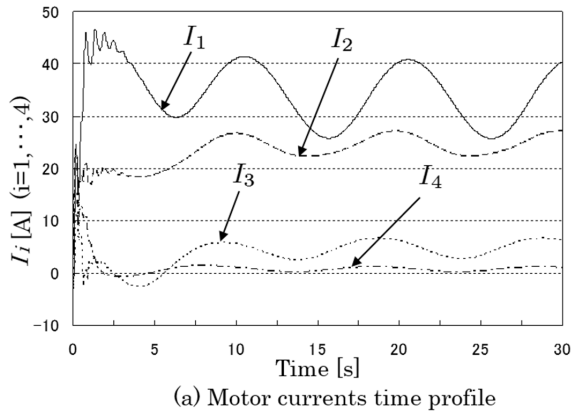
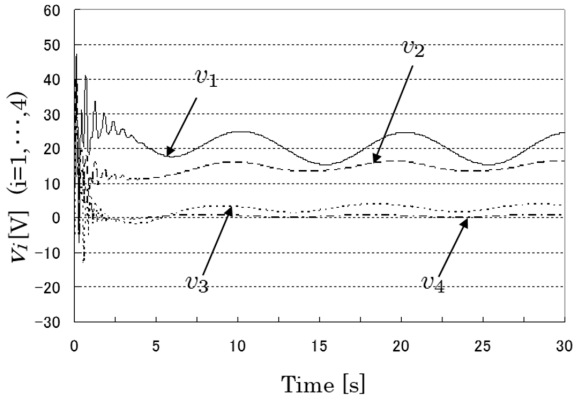


Fig. 4: Comparison of energy consumption defined by Eq.(29)



(a) Motor currents time profile

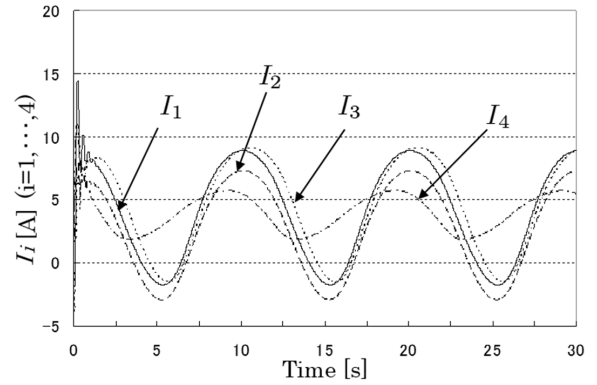


(b) Input voltage time profile

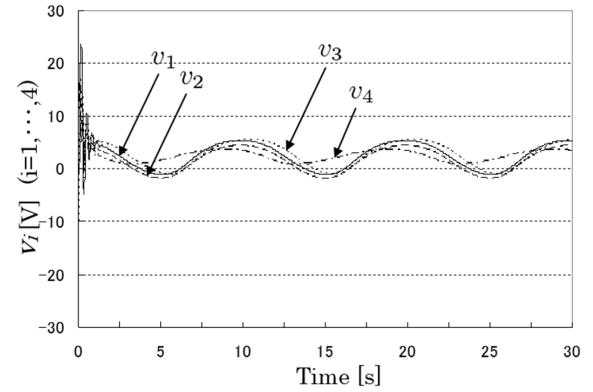
Fig. 5: Input voltages and currents of motors without bracing

肘を付かない場合と肘を付く場合の消費エネルギーの差について検討するため、Fig.5に肘を付かない場合の $t = 0$ から $t = 30$ [s] の (a) モータ電流 (b) モータ印可電圧を示す。式 (26), 式 (27) よりハンドの軌道追従の周期は 10[s] であるので Fig.5 には周期 10[s] の変動と $t = 0$ 直後の過渡応答が見られる。

この肘付きなしの運動と比較するために Fig.6 に肘を付く場合を示す。Fig.5(a) と Fig.6(a) の第一関節のモータ電流 I_1 を比較すると約 1/10 ほど小さいことが分かる。また Fig.5(b) と Fig.6(b) の第一関節のモータ印可電圧 v_1 を比較すると約 1/7 ほど小さくなっている。他の関節のモータの比較も同様であり、その結果 Fig.4 の消費エネルギーの差が生じたものと考えられる。



(a) Motor currents time profile



(b) Input voltage time profile

Fig. 6: Input voltages and currents of motors with 2nd link bracing

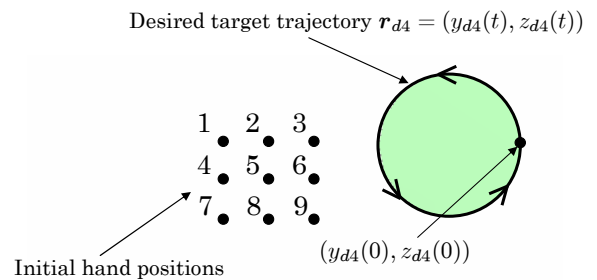


Fig. 7: Initial hand positions

4.2 ハンド軌道追従特性と初期値依存性

Fig.7 にハンドの目標軌道とその初期位置を示す。ハンドの初期位置として異なる位置の 1~9 の 9 点を設定した。まず肘付き拘束を利用する場合について考察する。肘付き抗力の目標値はリンク質量総計を 4 [kg] と設定したことを考慮し、 $f_{n2d} = 30$ [N] と与えた。式 (24) の y_{2d} は $y_{2d} = 0.4$ [m] とし、ゲインは $\mathbf{K}_{p4} = \text{diag}[100, 100]$, [N/m], $\mathbf{K}_{d4} = \text{diag}[33, 33]$, [Ns/m], $K_{p2y} = 150$, [N/m], $K_{d2y} = 75$, [Ns/m] と設定した。また、初期ハンド位置は Fig.7 に示す 1~9 の位置であり、初期肘付き位置は全ての場合について (0.4, 0) と固定した。1~9 の初期位置からスタートしたハンドの軌道追従結果を Fig.8(a) に示す。時間の経過とともにすべての軌道は一つの軌道に収束している。次にハンドの軌道追従ゲインを $\mathbf{K}_{p4} = \text{diag}[200, 200]$, $\mathbf{K}_{d4} = \text{diag}[66, 66]$ とした場

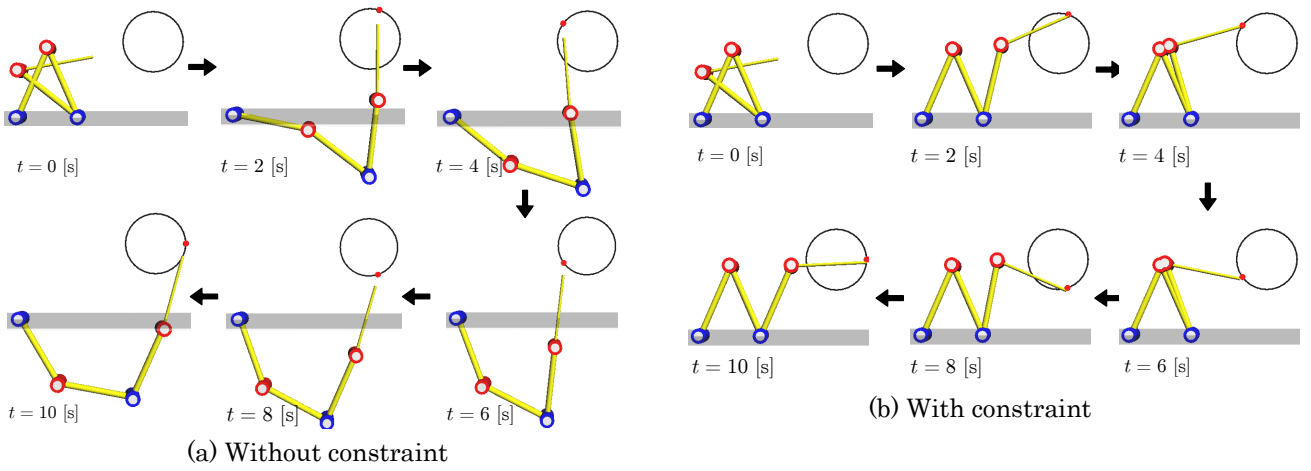
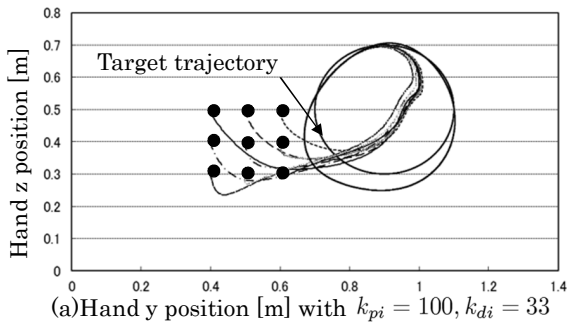
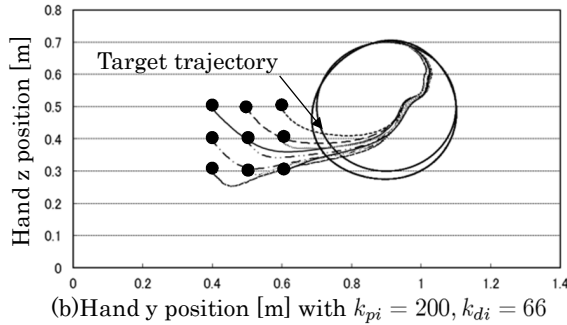


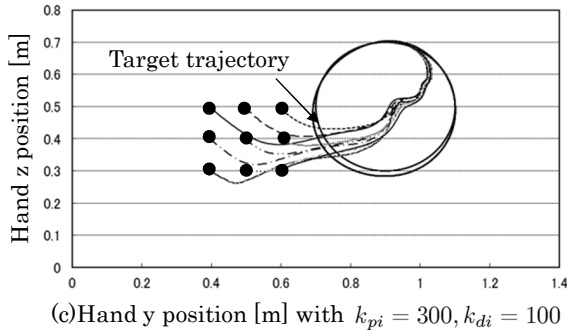
Fig. 10: Transition diagram of the manipulator's configuration



(a) Hand y position [m] with $k_{pi} = 100, k_{di} = 33$



(b) Hand y position [m] with $k_{pi} = 200, k_{di} = 66$

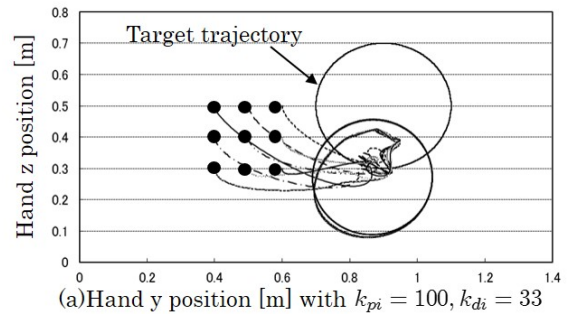


(c) Hand y position [m] with $k_{pi} = 300, k_{di} = 100$

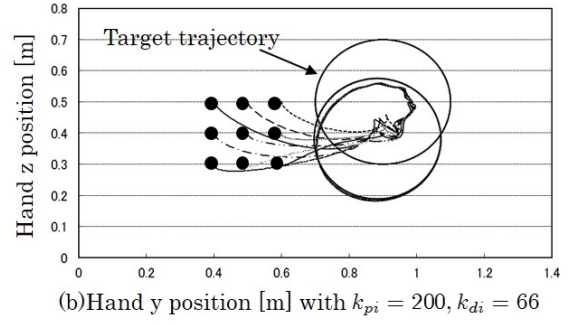
Fig. 8: Position time profiles of hand with constraint

合の結果を Fig.8(b) に、 $\mathbf{K}_{p4} = \text{diag}[300, 300]$, $\mathbf{K}_{d4} = \text{diag}[100, 100]$ とした場合の結果を Fig.8(c) に示す。

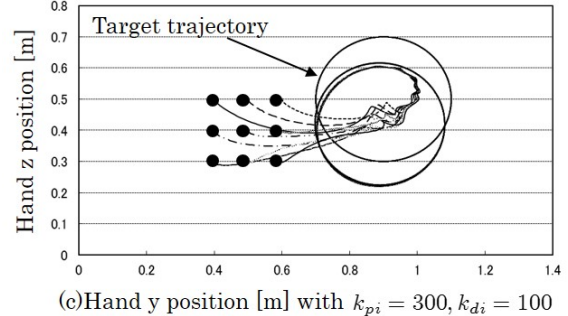
以上の肘付き拘束条件を与えた場合の結果と比較するため、肘を拘束しない条件の軌道追従結果を Fig.9 に示す。この場合の制御則は式 (25) の肘付き拘束時の制御則と同等の以下の式で表される。拘束条件がない場合、 $\partial C / \partial \mathbf{q}^T$ は零ベクトルと考えられるから、 $\mathbf{B} = \mathbf{0}$



(a) Hand y position [m] with $k_{pi} = 100, k_{di} = 33$



(b) Hand y position [m] with $k_{pi} = 200, k_{di} = 66$



(c) Hand y position [m] with $k_{pi} = 300, k_{di} = 100$

Fig. 9: Position time profiles of hand without constraint

であり、式 (23) より、 $\tau = \mathbf{l}$ となる。また、式 (25) より、肘付き位置の制御部分は存在しないので、 \mathbf{l} は右辺第 2 項のみとなる。この \mathbf{l} を式 (25) に代入すると、次式の制御則が得られる。

$$\mathbf{v} = \mathbf{K}_v [\mathbf{J}_4^T (\mathbf{K}_{p4} (\mathbf{r}_{d4} - \mathbf{r}_4) + \mathbf{K}_{d4} (\dot{\mathbf{r}}_{d4} - \dot{\mathbf{r}}_4))] \quad (30)$$

拘束条件を与えないこと以外のすべての条件は、Fig.8 の条件と等しい。図より明らかに軌道追従誤差が大き

くっており、肘付き拘束の利点は明らかである。

Fig.10はFig.7に示す初期ハンド位置1から目標軌道追従運動を始めた場合における、(a)肘付きなし、(b)肘付きありのロボットの形状の推移を示している。Fig.10(a)では、ロボットの形状が大きく変化している。これは、もともと式(25)、式(25)で与えられる制御方法はロボットの形状に関する制御を行っていないため、肘付きのない場合の制御式(30)も形状を制御するように意図されていないためである。これに対して、(b)では全体の形状制御を行っていないにも関わらず肘付きにより大きな形状の変化は見られない。

ここで、Fig.8(c)の場合について考える。シミュレーションにおいて目標軌道と手先軌道の誤差 e を以下のように定義する。

$$e(t) = \sqrt{(y_{d4} - y_4)^2 + (z_{d4} - z_4)^2} \quad (31)$$

また、初期位置1～9の中で i 点から始まるハンド軌道 (y_{4i}, z_{4i}) と j 点から始まる軌道 (y_{4j}, z_{4j}) の誤差 $e_{ij}(t)$ を以下のように定義する。

$$e_{ij}(t) = \sqrt{(y_{4i} - y_{4j})^2 + (z_{4i} - z_{4j})^2} \quad (i, j = 1, 2, \dots, 9, i \neq j) \quad (32)$$

ただし、 i, j はFig.7における1～9を表す。

Fig.11を見ると、1～9の初期位置からスタートした式(32)与えられるハンド軌道の軌道追従誤差 $e_{ij}(t)$ は、ほぼ重なっており、すべて3秒以降0.019[m]以下と減少している。次に初期位置1と2の軌道の差 $e_{12}(t)$ をFig.12に、その拡大図をFig.13に示す。3秒以降軌道の差は、 5.8×10^{-5} [m]以下に収まり続けている。表示はしていないが他のすべての軌道間の差も同様であり、ハンド軌道追従実験の初期位置による違いは3秒以降は無視できる程度に減少していると判断できる。以上より $t = 3$ [s]以降は初期値に依存した影響は無視できる程度に減少した。

4.3 肘付き位置の最適化

この節では肘付き位置の最適化について述べる。前節のシミュレーションより、 $t = 3$ [s]以降、初期位置の影響は無いと考えてよいことを確認できたので、 $t = 3$ [s]以降の消費エネルギー $E_{sum}^*(T)$ を評価の指標として肘付き位置の最適化を行う。

$$E_i^*(T) = \int_3^T v_i(t) I_i(t) dt \quad (33)$$

$$E_{sum}^*(T) = \sum_{i=1}^4 E_i^*(T) \quad (34)$$

ハンド負荷質量 M をパラメータとして0.0, 0.2, ..., 1.2[kg]と変化させる。また目標軌道の中心位置を $(y_c, z_c) = (0.8, 0.5), (0.9, 0.5), (1.0, 0.5)$ と与え、その3通りをA, B, Cと名前を付けて表すものとする。また目標肘付き位置 y_{d2} 及び初期肘付き位置 $y_2(0)$ をFig.14の w_1 から w_{17} に示す。さらに初期ハンド位置を目標軌道の中心として与えたため、A, B, Cそれぞれの目標軌道に対し、初期ハンド位置をA(0.8, 0.5), B(0.9, 0.5), C(1.0, 0.5)と設定した。

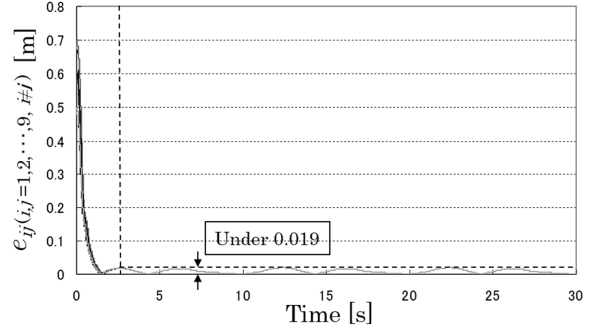


Fig. 11: Error of the hand

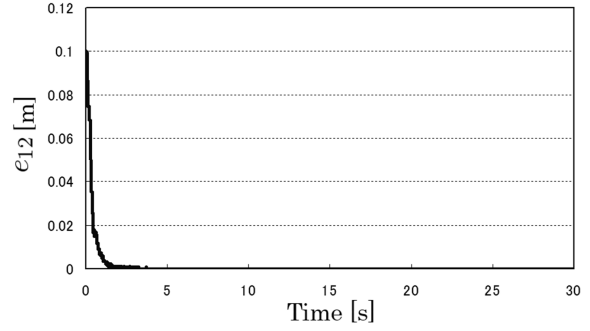


Fig. 12: Error profile between the trajectory started from initial point1 in Fig.7 and the trajectory started from point2

Fig.15はハンド負荷がゼロの条件であり、その縦軸は式(34)で与える消費エネルギーである。またFig.15の横軸である肘付き位置とはFig.2における作業座標系 Σ_W の原点から肘付き位置P地点までの距離を表す。Fig.15より、ハンド負荷 $M = 0$ の場合、目標軌道の中心位置 (y_c, z_c) がA: (0.8, 0.5), B: (0.9, 0.5), C: (1.0, 0.5)と移動するにつれて、最適肘付き位置は、0.4, 0.5, 0.6[m]であることが分かる。図中のA, B, Cは上記の目標軌道の位置を表すA, B, Cに対応している。また図よりA, B, Cと目標軌道中心位置がFig.2に示すy軸正方向に変化するにつれ、曲線の最低値を与える最適肘付き位置もy軸正方向に移っている。

次に、手先に付けた対象物の質量 M を0.2～1.2[kg]まで0.2[kg]ずつ変化させて、それぞれの場合に対してFig.14のように肘付き位置を変化させてシミュレーションを行った。肘付き位置の変化と目標軌道中心位置の変化に対する消費エネルギーのグラフをFigs. 16～21に示す。

ここでFig.2に示す目標軌道中心位置と肘付き位置の間の距離を“ S ”とし、 $S = y_c - y_{d2}$ で与えるものとする。目標軌道中心位置がAの場合の各ハンド負荷に対応する最適な肘付き距離 S は、Fig.15の場合、 $S = 0.8 - 0.4 = 0.4$ [m]となる。Fig.16の場合は、ほぼ $y_{d2} = 0.45$ で最低であるから $S = 0.8 - 0.45 = 0.35$ となる。同様にして、Figs. 15～21より S は、0.40[m]($M = 0.0$)[kg], 0.35($M = 0.2$), 0.30($M = 0.4$), 0.25($M = 0.6$), 0.20($M = 0.8$), 0.15($M = 1.0$), 0.15($M = 1.2$)となり、これを図示したものをFig.22に示す。

Fig.22より、ハンド負荷 M が増加するにつれ、目標軌道に近い位置で肘を付く方が消費エネルギーが少なくなることを示している。

Figs. 15～21のそれぞれの消費エネルギーはU字型

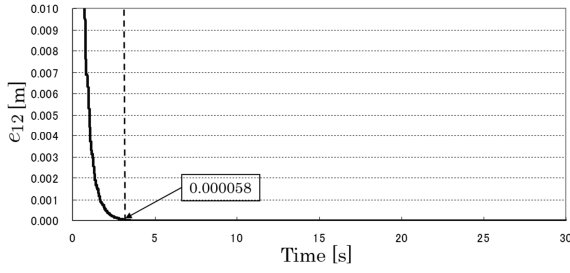


Fig. 13: Expanded error profile in Fig.12

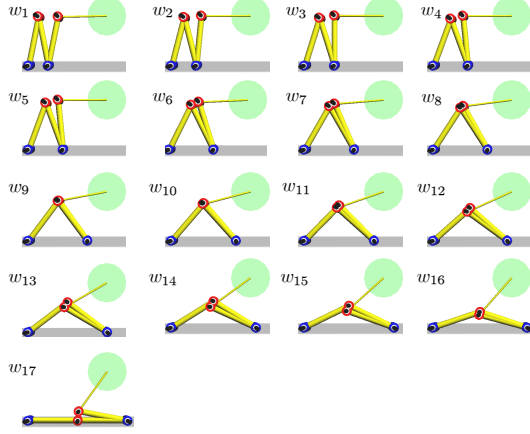


Fig. 14: Desired elbow-bracing position in simulation:

- $w_1 : y_{d2} = y_2(0) = 0.20$, $w_2 : y_{d2} = y_2(0) = 0.25$
 $w_3 : y_{d2} = y_2(0) = 0.30$, $w_4 : y_{d2} = y_2(0) = 0.35$
 $w_5 : y_{d2} = y_6(0) = 0.40$, $w_6 : y_{d2} = y_2(0) = 0.45$
 $w_7 : y_{d2} = y_9(0) = 0.50$, $w_8 : y_{d2} = y_2(0) = 0.55$
 $w_9 : y_{d2} = y_2(0) = 0.60$, $w_{10} : y_{d2} = y_2(0) = 0.65$
 $w_{11} : y_{d2} = y_2(0) = 0.70$, $w_{12} : y_{d2} = y_2(0) = 0.75$
 $w_{13} : y_{d2} = y_2(0) = 0.80$, $w_{14} : y_{d2} = y_2(0) = 0.85$
 $w_{15} : y_{d2} = y_2(0) = 0.90$, $w_{16} : y_{d2} = y_2(0) = 0.95$
 $w_{17} : y_{d2} = y_2(0) = 1.00$, where y_{d2} is given by Eq.(24)

をしており、肘付き位置が目標軌道に近すぎる場合、エネルギー消費量は増加する。また、Figs. 15~21 よりハンド負荷が大きくなるにつれて消費エネルギーが最低となる最適肘付き位置が判別しにくくなるのが分かる。

さらに、Fig.15 よりハンド負荷 M が零の場合の最小のエネルギーは、A、B、C に関わらず 0.8 kJ であり、 M が増加するにつれ、Fig.16 より、最小のエネルギーは 1.1 kJ ($M=0.2$)、Fig.17 より 1.6 ($M=0.4$) である。同様にして 2.0 ($M=0.6$)、2.5 ($M=0.8$)、3.1 ($M=1.0$)、3.8 ($M=1.2$) となる。以上をまとめて Fig.23 に示す。図より手先負荷の増加とともに最低消費エネルギーが増加することがわかる。

また、Figs. 15~21 において、肘付き位置 y_{d2} を $w_1 : 0.2$ から $w_{17} : 1.0$ まで変化させたとき、最大のエネルギーと最小のエネルギーの差が M が増えるにつれて増加することが分かる。肘付き位置に依存した消費エネルギーの差は、ハンド負荷が大きいくほど顕著であり、重量物を把持しているほど、肘付きを最適化する効果が大きいことが分かる。

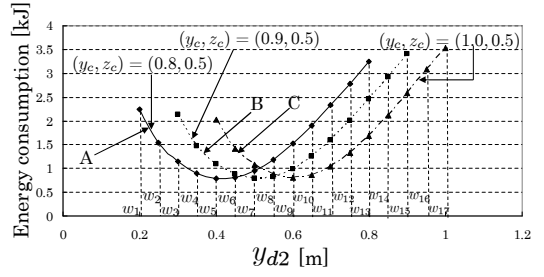


Fig. 15: Evaluation of energy consumption ($M=0.0$)

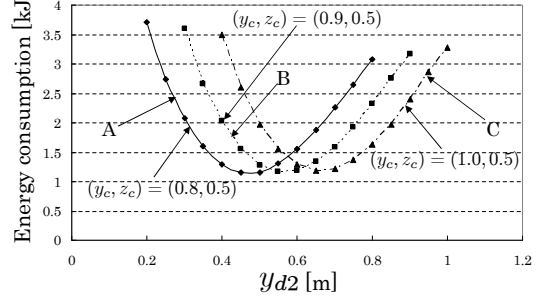


Fig. 16: Evaluation of energy consumption ($M=0.2$)

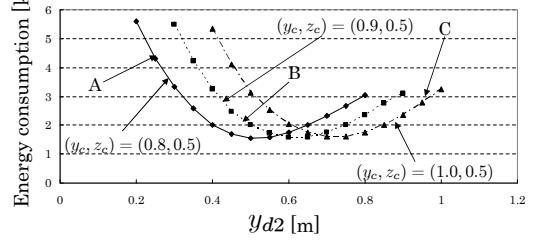


Fig. 17: Evaluation of energy consumption ($M=0.4$)

5 おわりに

本論文では、ハンド軌道追従精度とエネルギー消費の点でロボットの肘付き運動のメリットを、肘付きを行わないロボットの運動と肘付きを行うロボットの運動を比較することで議論した。その結果肘付きによってロボットの軌道追従精度は向上し、かつエネルギー消費は減少することを明らかにした。さらにエネルギー消費を最小化する肘付き位置は、目標ハンド軌道の位置と手先負荷重量に依存して変化することを示した。今後は、本報のシミュレーションで得られた結果から、実時間での肘付き位置最適化制御を行う予定である。

参考文献

- 1) G. S. Chirikjian and J. W. Burdick: A Hyper-Redundant Manipulator; *IEEE Robotics and Automation Magazine*, pp.22-29 (1994)
- 2) K. Glass, R. Colbaugh, D. Lim and H. Seraji: Real-time collision avoidance for redundant manipulators; *IEEE Trans. on Robotics and Automation*, Vol.11, pp.448-457 (1995)
- 3) H. Seraji and B. Bon: Real-Time Collision Avoidance for Position-Controlled Manipulators; *IEEE Trans. on Robotics and Automation*, Vol.15, No.4, pp.670-677 (1999)
- 4) S. Hirose and R. Chu: Development of a light weight torque limiting M-Drive actuator for hyper-redundant manipulator Float Arm; *Robotics and Automation, 1999. Proc. of IEEE International Conference*, Vol.4, pp.2831-2836 (1999)
- 5) J. Roy and L. L. Whitcomb: Adaptive Force Control of Position/Velocity Controlled Robots: Theory and Experiment; *IEEE Transactions on Robotics and Automation*, Vol.18, No.2, pp.121-137 (2002)

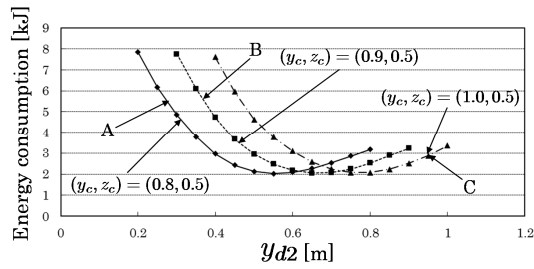


Fig. 18: Evaluation of energy consumption($M=0.6$)

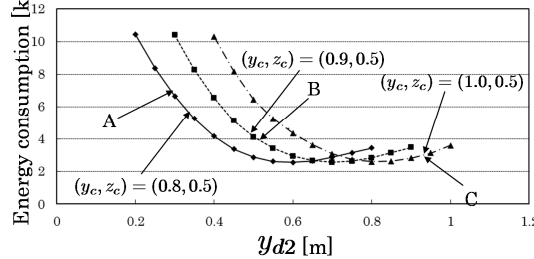


Fig. 19: Evaluation of energy consumption($M=0.8$)

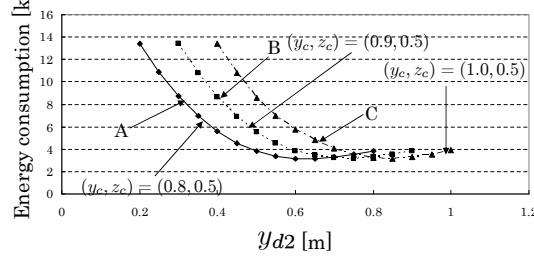


Fig. 20: Evaluation of energy consumption($M=1.0$)

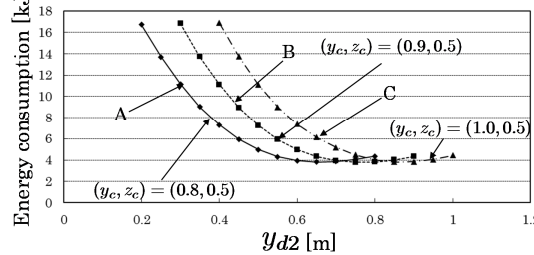


Fig. 21: Evaluation of energy consumption($M=1.2$)

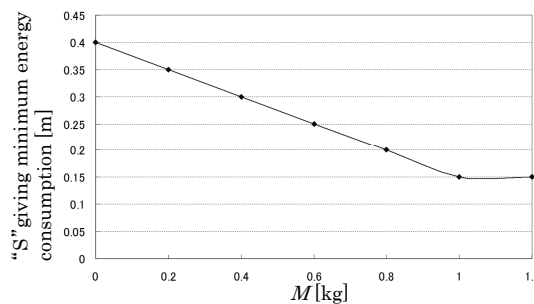


Fig. 22: Best bracing distance “S” to give minimum energy consumption when the center of the target trajectory being set at A(0.8,0.5)[m]

- 6) B. Siciliano and L. Villani: A passivity-based approach to force regulation and motion control of robot manipulators; *Automatica*, Vol.32, No.3, pp.443-447 (1996)
- 7) L. Villani, C. C. de Wit and B. Brogliato: An exponentially stable adaptive control for force and position tracking of robot manipulators; *IEEE Trans. Automat. Contr.*, Vol44, pp.778-

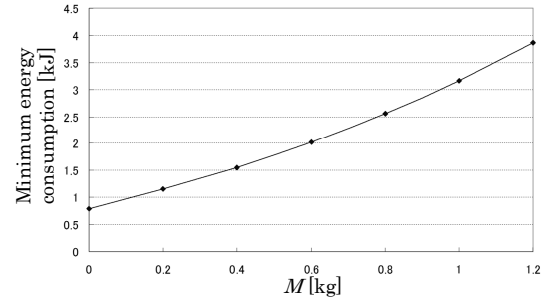


Fig. 23: Minimum energy consumption given by best bracing distance “S” by changing hand load “M” from 0 to 1.2 [kg]

- 802 (1999)
- 8) J. D. Schutter and H. V. Brussel: Compliant robot motion 2. A control approach based on external control loops; *Int. J. Robot. Res.*, Vol.7, No.4, pp.18-33 (1988)
- 9) J. Park and O. Khatib: Multi-Link Multi-Contact Force Control for Manipulators; *Proc. of 2005 IEEE Int. Conf. on Robotics and Automation*, pp.3624-3629 (2005)
- 10) A. Petrovskaya, J. Park and O. Khatib: Probabilistic Estimation of Whole Body Contacts for Multi-Contact Robot Control; *Proc. of IEEE International Conference on Robotics and Automation*, pp.568-573 (2007)
- 11) T. Yoshikawa: Dynamic Hybrid Position/Force control of Robot Manipulators—Description of Hand Constraints and Calculation of Joint Driving Force; *IEEE J. on Robotics and Automation*, Vol.RA-3, No.5, pp.386-392 (1987)
- 12) K. Yamane and Y. Nakamura: O(N) Forward Dynamics Computation of Open Kinematic Chains Based on the Principle of Virtual Work; *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp.2824-2831 (2001)
- 13) K. Yamane and Y. Nakamura: Dynamics Filter - Concept and Implementation of On-Line Motion Generator for Human Figures; *IEEE Transactions on Robotics and Automation*, vol.19, no.3, pp.421-432 (2003)
- 14) H. West and H. Asada: A Method for the Design of Hybrid Position/Force Controllers for Manipulators Constrained by Contact with the Environment; *Proc. of IEEE Int. Conf. on Robotics and Automation*, pp.251-260 (1985)
- 15) H. Hemami and B. F. Wyman: Modeling and Control of Constrained Dynamic Systems with Application to Biped Locomotion in the Frontal Plane; *IEEE Trans. on Automatic Control*, Vol.AC-24, No.4, pp.526-535 (1979)
- 16) Z. X. Peng and N. Adachi: Position and Force Control of Manipulators without Using Force Sensors (in Japanese); *Trans. of the Japan Society of Mechanical Engineers(C)*, Vol.57, pp.1625-1630 (1991)
- 17) W. Gu, H. Kataoka, F. Yu, T. Maeba, M. Minami and A. Yanou: Control of Hyper-Redundancy Mobile Manipulator with Multi-Elbows braced for High Accuracy/Low-Energy Consumption; 第 21 回インテリジェント・システム・シンポジウム, 1A1-4 (2011)
- 18) 糸島 道之, 前場 友秀, 見浪 護, 矢納 陽: 位置/力空間における肘付き冗長性を利用したマニピュレータの制御; システム・情報部門学術講演会, 第 21 回インテリジェント・システム・シンポジウム, 2B2-3 (2011)
- 19) 糸島 道之, 尾崎 壮紘, 前場 友秀, 見浪 護, 矢納 陽: 肘付拘束冗長性を利用したハンド軌道追従と拘束運動の非干渉制御; 第 2 回コンピューター・シヨナル・インテリジェンス研究会, pp.57-64 (2012)
- 20) 近藤 大介, 糸島 道之, 見浪 護, 矢納 陽: 拘束状態を利用するマニピュレータの肘付き位置の最適化; 第 57 回システム制御情報学会研究発表講演会, 347-1 (2013)
- 21) 河村 末久, 矢野 章成, 樋口 誠宏, 杉田 忠彰: 研削加工と砥粒加工, 加工学基礎 2, 共立出版 (1984)

PHD フィルタを用いたバス乗客の行動把握

○*南 哲志 **Pham Minh Tuan *橘 完太 (*工学院大学, **ダナン工科大学)

Bus Passengers Activity Recognition Using PHD Filter

○* T. Minami, **M. Pham and *K. Tachibana

(*Kogakuin University, **Danang University of Science and Technology)

Abstract— In Japan, as the declining birthrate and the aging of society progresses, future needs for route buses as public transportation facilities are rising. Then improvement for bus routes and schedules is required. However, existing system can collect only data of passengers paid fee, so there is no method for collecting accurate data of passenger such as commutation ticket users and free ticket users. Therefore, in this paper, we consider a method that can recognize bus passenger activity using PHD filter through images obtained inside of a bus.

Key Words: PHD filter, object tracking

1 はじめに

日本の路線バス利用者は 1946 年の年間延べ約 100 億人から年々減少の傾向を辿っているが、現在でも年間延べ 40 億人程度の利用があり¹⁾、特に交通網の発達していない地域では重要な交通手段となっている。一方で全国 244 社のうち 73%、大都市部を除くと 170 社のうち 89%が赤字の収支となっており²⁾、乗客の需要に合わせた路線の統廃合やダイヤの再編によって収益性を上げる必要に迫られている。しかし現状では乗客の需要は荒いメッシュの住民統計や年に数度の計数調査のデータによってのみ定量化されているに留まっている。実際には現金や、見せるのみで精算を行う定期券や老人向けの公共サービスであるシルバーパスを用いた利用者も多いと考えられるが、その数を安価かつ継続的に把握する手法は確立されていない。またカウンティングだけでなく、どの停留所からどの停留所までバスを利用したか、といった停留所単位での需要も必要となる。

そこで本研究では、車載カメラで記録したバス車内の映像を用いて、乗客の計数・行動を把握する。そこでこれまで不明確であった乗客の乗降を定量的に記録し、より厳密な需要の把握の一助となる手法を検討する。バスの乗客を把握する問題には、例えば、(1)入り口と出口が固定されている、(2)オクルージョンが多発する、(3)車内の位置によって乗客の行動が変化する、(4)奥行を考慮すべき位置にカメラが設置されている、などの特徴がある。また、乗客のプライバシーに関する問題もある。そういった特徴や問題に対し、本研究では近年注目されている PHD (Probability Hypothesis Density) フィルタ³⁻⁸⁾を用いて対処する。しかし、いずれの先行研究も 3 次元空間上の奥行を考慮する必要のない応用先である。本稿では、バスの車内を模した状況下で撮影した動画を用いて、バス車内をセルに区切り乗客の状態遷移をモデル化し、3 次元空間上の奥行も考慮した手法を提案する。

2 提案手法

近年、物体追跡の分野ではPHDフィルタという手法が注目されている。PHDフィルタとは、複数の状態を推定できるよう、再帰的に事後密度を計算する手法である。時刻毎に追跡対象の数や観測数が変化する場合のモデル化に適している³⁾。PHDフィルタは以下の式

で与えられる。

$$\Omega_{k|k-1} \sim Q_k(\Omega_k | \Omega_{k-1}) \quad (1)$$

$$z_k \sim R_k(z_k | \Omega_k) \quad (2)$$

ここで、 k は時刻を表す。 $\Omega_{k|k-1}$ は予測された状態の集合の分布であり、 z_k は観測された画像である。 Q_k は一時刻前の状態の集合から得られた状態の集合の分布であり、 R_k は状態の集合と観測された分布を比較した結果の分布である。Fig. 1. にPHDフィルタの流れ図を示す。

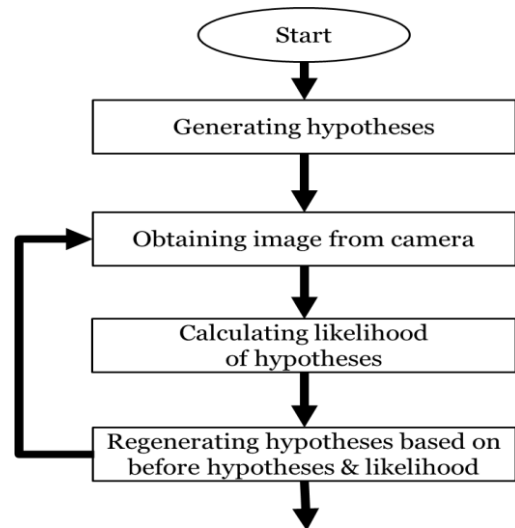


Fig. 1. Flowchart of PHD filter

本問題では、全ての乗客は出入口から乗降車し、座席の位置もほとんど固定であるため、バス車内の出入口や座席の位置はおおよそFig. 2. のように決まる。

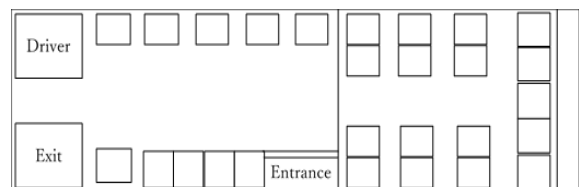


Fig. 2. Layout of bus

また、乗客の行動範囲もかなり制限できるので、問題の単純化のために提案手法ではバス車内をFig. 3. のように複数のセルに分割する。

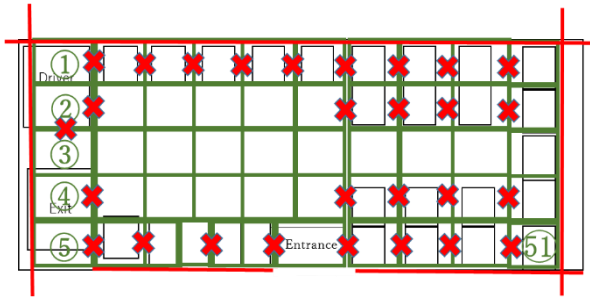


Fig. 3. Layout partitioned into cells

それぞれのセルの人数は制限され、既に占有されているセルに移動する確率を下げ、座席セルにいる場合は着席状態と見なし乗客の高さを下げる、などの工夫が考えられる。それぞれのセルは番号を持ち、赤い線とバツ印は乗客の移動が不可能であることを示す。また、カメラの位置はおおよそ③番から⑤番の付近と想定される。

また、本問題では乗客らは歩行や着席など、単純な行動しか取らない。そこで時刻 k での状態集合 Ω_k を
$$\Omega_k = \{(cell_id, X, Y, Z, h, w, b)\}_{i=1}^{J_k}$$
 (3)

と定義する。 Ω_k それぞれは仮説と言い換えることもできる。 J_k はその仮説での乗客の数、 $cell_id$ はその乗客のいるセル番号、 X, Y, Z は乗客の座標、 h, w は乗客の高さ、幅を表す。 b は乗客の行動であり、 walking, standing, sitting, getting off, sliding から成る。 sliding は2人以上用の座席で、座席位置を移動する状態を指す。乗客の座標・幅・高さに関しては、分散を設定し、平均0の正規分布でぶらす。乗客の行動 b に関して、状態遷移モデルは

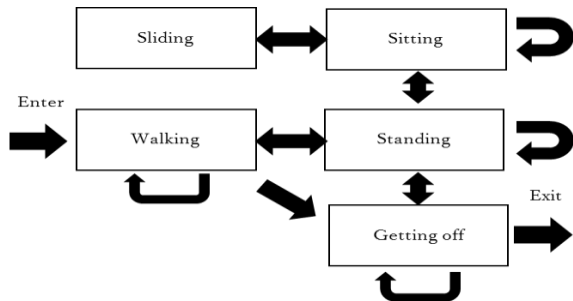


Fig. 4. Behavior of passenger transition diagram

のように考えられる。

また、状態遷移確率は乗客の位置によって変わる。例えば座席のセルにいれば sitting への遷移確率が上がり、出口付近のセルにいれば getting off への遷移確率が上がる。例として出口付近の遷移確率を Table 1. に示す。その他の場所での状態遷移に関しては、本稿の末尾に付記した。

Table 1. Behavior transition probability of the cell near the exit

current \ next	Walking	Standing	Getting off	Sitting	Sliding
Walking	0.25	0.25	0.5	0	0
Standing	0.25	0.25	0.5	0	0
Getting off	0.15	0.15	0.7	0	0
Sitting	0	0	0	0	0
Sliding	0	0	0	0	0

状態遷移モデルが定義でき仮説を立てられるようになったので、次に、立てた仮説が観測された画像と比較してどれほど尤もらしいかを定義するため、観測モデルを考える。仮説によって画像上に矩形状のモデルが構築できるので、必要なはその矩形が如何に乗客を捉えているかの指標である。尤度 *likelihood* を

$$likelihood = \frac{|A \cap B|}{|A \cup B|} \quad (4)$$

と定義し用いる。ここで A は矩形の面積の総和、 B は画像全体の特徴量である。振動に関しては Enhanced Correlation Coefficient (ECC)、輝度の変化に関しては輝度の変化に頑健な表色系を用いることで対応できると考えられる⁹⁾。

3 実験

3.1 実験概要

実験はバスの車内を模した場で行った。 Fig. 5. にバスの車内を模した実験環境を示す。

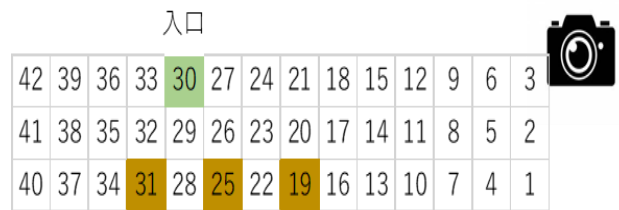


Fig. 5. Layout of experiment

それぞれの数値はセル番号を示し、19, 25, 31 番のセルは座席、30 番のセルは入口を表す。カメラの位置は3番のセル付近である。各セルは 50×50 cm である。

先行研究にてこの問題で背景差分法を用いる有用性が示された⁹⁾ので本実験でも背景差分を用いて特徴量とする。また、比較対象として奥行を考慮せず画面上の座標のみを状態として遷移させるモデルを構築した。比較対象の状態集合 Ω_k は以下の式で与えられる。

$$\Omega_k = \{(u_l, v_t, u_r, v_b)\}_{i=1}^{J_k} \quad (5)$$

ここで、 (u_l, v_t) は画面上の矩形の左上の座標、 (u_r, v_b) は画面上の矩形の右下の座標を表す。

状態遷移モデルのパラメータとして、提案手法では、 X, Z, h の分散を5、 h の分散を2としてそれぞれ平均0の正規分布から実現値を生成し、状態を遷移させる。比較手法では、 u_l, v_t, u_r, v_b は分散を10として同様に実現値を生成し、状態を遷移させる。また、どちらの手法も仮説数は1000、フレームレートは毎秒10フレーム、背景差分を行う際の2値化の閾値は200とした。

実験には30fpsで12秒間、3人が入口から各座席に着席していく 800×600 px の動画を使用した。 Fig. 6. は実験動画の1フレームを例として取り出してきたものである。



Fig. 6. One frame of experiment movie

3.2 結果



Fig. 7. One frame of comparative method

Fig. 7. は、比較手法のあるフレームを取り出した画像である。1 仮説を描写しており、それぞれの矩形は仮説を表す。乗客 1 を 3 人分、乗客 3 を 2 人分と捉えるなど人数は不正確で、乗客 2 が追えていないことが分かる。



Fig. 8. One frame of proposal method

Fig. 8. は提案手法のあるフレームを取り出した画像である。乗客数は正確で、3 乗客とも追えている。

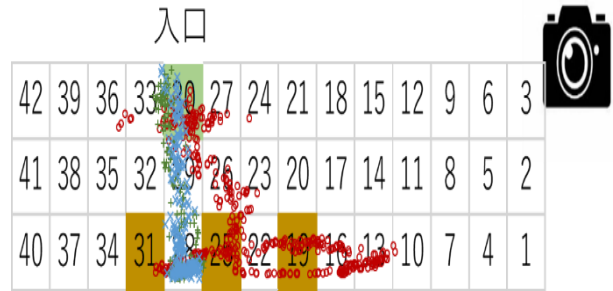


Fig. 9. Trajectory of passengers

Fig. 9. は提案手法の仮説を二次元上に描写したものである。赤が乗客 1、青が乗客 2、緑が乗客 3 を表す。人数は正確に追えているが、奥行方向には完全に追いついていないことが分かる。

4 考察

4.1 提案手法の有意性

比較手法で乗客の把握が不正確だったのは大きく分けて 2 つ理由が考えられる。理由の一つは、ピクセル値だけを状態とすると奥行方向の情報を持たないため、乗客がカメラの手前に来れば分散を大きくしないと対応しきれないことである。使用した動画にて手動で人間を切り出してみると、同じ乗客でも 30 番のセル（入口付近）では $30 \times 280\text{px}$ 程度だったのが、19 番のセル（一番手前の座席）で立っている状態だと $170 \times 550\text{px}$ 程度と幅は 5 倍、高さは 2 倍ほどになっている。Table 5. で示した X, Y のぶらし方、すなわち分散を大きくすれば対応できるが、分散を大きくすればするほど乗客を同一視しやすくなってしまいう問題があり、好ましくない。もう一つの理由は、入口の制約条件を設定できないことである。例えば「入口付近に乗客がいるときに乗客は追加しない」とすると、入口付近のピクセルを入口であるなどして対応することが考えられるが、実際に入口でない場所に乗客がいたとしても追加しないパターンがあるため不適切である。「新しく乗客を追加してから数フレームは追加しない」という設定も考えられるが、ある乗客が乗車してから次の乗客が乗車するまでの最低フレーム数は不定であるため、不適切である。提案手法では奥行情報を持ち出入口の情報も管理しているので、これらの問題は解決される。また、いずれにせよ人間の個人的な情報は用いていないので、プライバシーの問題に対しても対応が取れていると考えられる。

4.2 提案手法の改善

しかし Fig. 9. などの結果を見ると、まだまだ乗客の把握は不十分であるなどの問題がある。そこで手法を改善するために、いくつかの方法が考えられる。まずそのセルが占有されているかどうかのリストを作成し、占有されていればその行動を制限することである。また乗客の思考モデルとして、「空いている座席があればそちらに向かう可能性が高まる」なども導入できる

だろう。また最終的な状態を確認すると、本来全員が sitting になっているはずがそうはなっていないので、「一度 sitting に遷移したら h が大きく変化するまで状態は遷移しない」などとしても良いかもしれない。それから、オクルージョンが起きたときのフラグを設定し、隠れ状態では x, h の分散を低減することで隠れ状態時の把握精度を向上させることが可能だろう。尤度や特徴量に関しても、今回は誰もいない状態との差分を特徴量として画像中の特徴量をどれだけ捉えているか、といった単純なものを採用したので、人画像のデータベースを用いた学習手法の適用も有効であると考えられる。また、バス車内で実際に人間がどれくらい移動しているかを実測し、より厳密な分散や、実際に人間がどういった状態の遷移を行っているかを実測し、より厳密な遷移確率を適用することで、モデルの精度向上が見込める。更に今回は対象としなかったが、バスが走行中は乗客の移動が減るため、スキップするフレーム数を増やすことも考えられる。

5 おわりに

本稿では、バス車内にあるカメラで撮影された画像に PHD フィルタを適用し、乗客の行動を把握する手法を示した。実験としてバス車内を模した状況下で撮影された動画を用い、提案手法の有効性を示した。また、実験結果から浮かび上がった問題点に対する解決策と手法に対する改善策・今後の展望について述べた。

参考文献

- 1) 国土交通省, 自動車輸送統計月報, 第 54 巻第 6 号, pp. 7. (2017)
- 2) 公益社団法人日本バス協会, バス事業の現状と取り組みについて, 社会資本整備審議会資料, pp.3. (2016)
- 3) B. Vo and W. Ma, "The Gaussian mixture probability hypothesis density filter," IEEE Trans. Signal Process., vol. 54, no. 11, pp. 4091/4104, (2006)
- 4) Si, Weijian, Liwei Wang, Zhiyu Qu. "Multi-target tracking using an improved gaussian mixture CPHD filter." Ed. Stefano Mariani. Sensors (Basel, Switzerland) 16.11, (2016)
- 5) Fu, Z., Feng, P., Naqvi, S. M., & Chambers, J. A. "Robust particle PHD filter with sparse representation for multi-target tracking." Digital Signal Processing (DSP), (2016)
- 6) Zhang, Yongquan, Hongbing Ji, and Qi Hu. "A box-particle implementation of standard PHD filter for extended target tracking." Information Fusion 34, pp. 55/69, (2017)
- 7) Wu, W., Jiang, J., Liu, W., Feng, X., Gao, L., Qin, X. "Augmented state GM-PHD filter with registration errors for multi-target tracking by Doppler radars." Signal Processing 120, pp. 117/128, (2016)
- 8) Xue, W., Yan, L. H., Qian, T., & Lei, P. "A multiple target measurement retrieval algorithm based on KN-neighborhood membership degree P-PHD filtering." Communication Software and Networks (ICCSN), 2016 8th IEEE International Conference on. IEEE, pp 615/619 (2016).
- 9) Tetsushi Minami, Kanta Tachibana, "A Study on Observation Model of PHD Filter for Estimating Behavior of Bus Passenger", The 16th International Symposium on Advanced Technology (2017)

付録

Table 3. Behavior transition probability of cell near the entrance

current\next	Walking	Standing	Getting off	Sitting	Sliding
Walking	0.5	0.5	0	0	0
Standing	0.6	0.4	0	0	0
Getting off	0	0	0	0	0
Sitting	0	0	0	0	0
Sliding	0	0	0	0	0

Table 4. Behavior transition probability of the aisle cell

current\next	Walking	Standing	Getting off	Sitting	Sliding
Walking	0.5	0.5	0	0	0
Standing	0.5	0.5	0	0	0
Getting off	0	0	0	0	0
Sitting	0	0	0	0	0
Sliding	0	0	0	0	0

Table 5. Behavior transition probability of the double seat cell

current\next	Walking	Standing	Getting off	Sitting	Sliding
Walking	0	0.5	0	0.5	0
Standing	0	0.5	0	0.5	0
Getting off	0	0	0	0	0
Sitting	0.15	0.15	0	0.55	0.15
Sliding	0	0.1	0	0.7	0.2

Table 6. Behavior transition probability of the rearmost seat cell

current\next	Walking	Standing	Getting off	Sitting	Sliding
Walking	0	0.5	0	0.5	0
Standing	0	0.5	0	0.5	0
Getting off	0	0	0	0	0
Sitting	0.15	0.15	0	0.55	0.15
Sliding	0	0.1	0	0.7	0.2