

自己組織化写像の拡張によるモーションキャプチャシステムの頑健化

叶内 円†, 島田 敬士‡, 有田 大作*†, 谷口 倫一郎‡

九州大学大学院システム情報科学府†
九州大学大学院システム情報科学研究院‡
〒 819-0395 福岡県福岡市西区元岡 744
財団法人九州先端科学技術研究所*

〒 814-0001 福岡県福岡市早良区百道浜 2 丁目 1-22

E-mail: {kanouchi,atsushi,rin}@limu.is.kyushu-u.ac.jp, arita@isit.or.jp

あらまし 本稿では、実時間ビジョンベースモーションキャプチャシステムの頑健化のために、高速化した追加学習型の自己組織化写像を利用した事例ベース手法を提案する。追加学習型の自己組織化写像は、オンライン処理の中で逐次的に各部位の特徴点情報を関連付けて学習することができる。そのため、特徴点抽出に失敗した時には、抽出に成功した情報のみを利用して、抽出に失敗した特徴点を補完することが可能である。しかし、事例を追加的に学習する場合、学習データ数が増加するにつれて計算時間も増加するという問題がある。提案手法では、逐次学習型の自己組織化写像の学習法を高速化することで、実時間で安定にモーションデータを学習することが可能である。

キーワード 自己組織化マップ, ビジョンベースモーションキャプチャシステム, 学習の高速化

Improvement in Robustness of Real-Time Vision-based Real-time Human Posture Analysis by the enhancement of Self-Organizing Map

Madoka Kanouchi†, Atsushi Shimada‡, Daisaku Arita*† and Rin-ichiro Taniguchi‡

Department of Intelligent Systems, Kyushu University†‡

744, Motoooka, Nishi-ku, Fukuoka, 819-0395, Japan

Institute of Systems, Information Technologies and Nanotechnologies*

2-1-22, Momochihama, Sawara-ku, Fukuoka, 814-0001, Japan

E-mail: {kanouchi,atsushi,rin}@limu.is.kyushu-u.ac.jp, arita@isit.or.jp

Abstract We propose a approach to improve the accuracy of estimating feature points of human body on a vision-based motion capture system using the accelerated incremental learning type of the Self-Organizing Map (SOM). The incremental learning type of the SOM has the ability to correlate the feature points of each body part in online sequential learning. Therefore, we can complement some feature points which are failed to extract using the successful information. However, the typical incremental learning type of the SOM has a problem that its training time increases as the number of training data increases. Using our proposed method which is speeding up the learning method of the incremental learning type of SOM, it can stably learn motion data on real-time.

Key words Self-Organizing Map, Vision-Based Motion Capture System, Speeding Up of Learning

1. はじめに

Kohonen らにより提案された自己組織化マップ (Self-Organizing Map : SOM)[1, 2] は人工ニューラルネットワークのひとつである。SOM は教師なし競合近傍学習により、高次元データをデータ間

の類似性を保持した低次元 (2 次元がよく用いられる) のマップに非線形写像し、それを記憶しておくことができる。また、SOM には連想記憶という機構がある。これは、学習時に入力信号に対する出力信号を組み合わせとして学習しておくことで、認識時には入力信号あるいは出力信号のいずれかし

か入力されない場合でも、他方の信号を連想することができるというものである。SOMを学習の方式で分類すると、一括学習型と追加学習型の2種類に大別される。一般にSOMの学習は学習データがすべて揃ってから学習を開始するいわゆる一括学習型が主流であり、データの収集と学習を同時に行う逐次的な学習には対応が難しい。追加学習型は、入力データを逐次的に学習することが可能であるため、オンラインクラスタリングやインタラクティブ性を必要とするロボット制御などへ応用可能である。本稿では、追加学習型SOMを利用したビジョンベースモーションキャプチャシステムにおける特徴点抽出の頑健化を実時間でを行う手法を提案する。

本研究で想定しているビジョンベースモーションキャプチャシステムは、人体に特別なマークを付ける必要がなく、カメラで撮影した画像のみから人物の姿勢を推定するものである。伊達らによって発表されているビジョンベースモーションキャプチャシステム[3]では、複数台のカメラで被計測者を撮影し、各カメラの画像から人物の頭部と手足の位置を抽出、計測する。そして全カメラの情報を統合することで各身体部位の3次元位置（以下、3次元特徴点）を推定している。頭部や手足の3次元特徴点は、他の部位を推定する時に利用されるため、精度よく推定されている必要がある。しかし被計測者の姿勢によっては頭部や手足がカメラに写らないことがあるため、一部の部位の3次元特徴点の推定を誤ってしまうことがある。本研究では、このように抽出に失敗した特徴点のことを特に「欠落特徴点」と呼ぶ。

ビジョンベースモーションキャプチャシステムの高精度化を行うために、人体の取りうる姿勢を事例として事前に学習させ、利用時には事例と推定姿勢とのマッチングにより推定精度を上げる研究は数多く取り組まれている[4, 5, 6]。先行研究では、事前に人体の取りうる姿勢を学習するが、どのような姿勢を学習しておくべきかというのは使用する場面において異なるものであり、またできるだけ多様な姿勢に対応するため大量の事例が必要になる。例えば大量の事例が準備できたとしても、実際の利用時に新しい姿勢の事例が必要になった場合には再学習の必要がある。本研究においても、SOMに人体の取りうる姿勢を学習させるという点は共通であるが、島田らが提案したように追加学習型のSOMを利用して学習もオンラインで行う[7]という点で異なっている。学習をオンラインで行うと、新しい姿勢をその都度追加学習できるため、このような問題が生じない。またSOMを利用していることから、近傍学習により事例の数が少数で

あっても類似事例を複数ニューロンに記憶させることが可能であるため学習効率が向上する。

しかし、追加学習にはニューロンの枯渇、データの忘却、位相破壊などの解決すべき問題があり、これまでに様々な取り組みがなされているが、完全な解決法は得られていない。山田らはシナプス硬直条件とシナプス半硬直条件を利用したSOMの追加学習法[8]を提案しており、これはデータの忘却問題を解決するものの、ニューロンの枯渇と位相破壊については解決していない。ニューロンを適応的に追加することによりニューロンの枯渇問題を解決する手法として、Growing Self-Organizing Map (GSOM)[9]、Growing Hierarchical Self-Organizing Map (GHSOM)[10]、Growing Neural Gas (GNG)[11]、Evolving Self-Organizing Map (ESOM)[12]、Self-Organizing Incremental Neural Network (SOINN)[13]などが提案されている。これらは位相破壊は起こらないが、データの忘却問題については言及していない。これら3つの課題を解決している追加学習型SOMとして、島田らにより密度可変型自己組織化マップ (Variable Density Self-Organizing Map:VDSOM) [14]が提案されている。

VDSOMにより一般的な追加学習時の問題は解決されるが、インタラクティブなシステムへ応用するには、これらの問題に加えて、即応性の問題を解決しなければならない。一般に、SOMの学習はニューロンの数に比例して学習時間も増加する傾向にある。追加学習型のSOMは、学習過程でニューロンを適応的に追加するため、学習が進むにつれてニューロン数が増加し、それに伴い学習時間も増加することになる。これは前述の全ての追加学習法にあてはまる。従って、追加学習型のSOMを実時間でインタラクティブ可能なレベルで利用できるように、その学習法を高速化する必要がある。

そこで本研究では、高速化の実現のために、局所更新法、局所探索法、不要ニューロン消去法をVDSOMに導入する。局所探索法では、学習時に入力に最も類似した結合係数ベクトルを持つ最大活性化ニューロンを探索する際に、その探索範囲を局所的にすることによって高速化を行う。この時、時間的に連続する入力はその前後で類似した値を持つという仮定と、SOMのマップ上では類似したデータが近傍に存在するという特徴を利用している。局所更新法では、ニューロンの持つ結合係数ベクトルを入力に近づけるように更新する際に、最大活性化ニューロンから位置的に離れたニューロンの結合係数ベクトルを更新せずに、更新範囲を限定することで、高速化を行う。不要ニューロン消去法は、あまり参照されていなかったり、他のニューロンに比べて入力との類似度が低いニューロンを消去

し、ニューロン増加に伴う処理時間を抑える手法である。以上を VDSOM に組み込んだ手法を高速密度可変型自己組織化マップ (Accelerated Variable Density Self-Organizing Map:AVDSOM) と呼ぶ。

オンライン学習には高速な追加学習の必要があるため、高速な追加学習型 SOM である AVDSOM が利用できる。全ての部位の姿勢推定が成功した場合、その正しい姿勢情報を AVDSOM に追加学習させる。一方で、一部の部位の姿勢推定を失敗した場合、すなわち、欠落特徴点が存在する場合は、AVDSOM の連想記憶能力を利用して欠落特徴点を追加学習した事例に基づき補完する。本研究では、このように学習時の情報を連想記憶により呼び起こす手続きを「想起」と呼び、学習と想起を必要に応じて切り替えながら姿勢推定を行うことで、モーションキャプチャシステムの高精度化を行う。SOM の学習過程と想起過程の切り替え方については、モーションキャプチャシステムで抽出された特徴点には「信頼度」が付加されているところを想定し、この信頼度の高さに応じて AVDSOM の振る舞いを決定する。以下では VDSOM, AVDSOM, 学習と想起の切り替え方について順に述べる。最後に、モーションキャプチャシステムに AVDSOM を組み込んで、実時間化と頑健化の実験結果について述べる。

2. 密度可変型自己組織化マップ

SOM で追加学習をする場合には、次の5つの問題が生じることが知られている。

問題 1 ニューロンが枯渇する

問題 2 過去に学習したデータが忘却される

問題 3 ニューロン間の位相が破壊される

問題 4 学習時間が追加学習過程で増加する

問題 5 不要ニューロンが存在する

問題 1 は、事前に設定したニューロン数をはるかに上回る数の入力データを学習する際に、あらかじめ設定したニューロン数では足りなくなるという問題である。問題 2 は、SOM では後に入力されるデータほどマップに大きな影響を与えるので、過去に入力されたデータほど忘却してしまう可能性が高くなるという問題である。問題 3 は、学習するデータが多すぎてマップが全ての入力を記憶できなくなると、マップの位相近傍性が保てなくなるとい問題である。問題 4 は、学習時間がマップ上のニューロンの数に比例して増加していく問題である。入力に応じてマップにニューロンを追加していく学習手法では、多くのデータを学習すれ

ばするほどニューロンが多くマップに追加される。その結果、学習時にニューロンの探索や更新に多くの時間がかかるようになる。本稿では実時間での学習を目指すので、この問題を解決することが重要になる。問題 5 は、マップ上のニューロン数が動的に変化する学習手法において、ニューロンの追加と競合近傍学習を繰り返しから、あまり参照されない不要なニューロンがマップ上に存在する問題である。

VDSOM では、シナプス硬直条件 [15] という結合係数ベクトル更新時の拘束条件と、ニューロンの適応的追加により問題 1-3 を解決する。シナプス硬直条件とは、学習データを十分学習したニューロン (以下、シナプス硬直ニューロンと呼ぶ) に適用され、それ以降結合係数ベクトルの値を更新しないようするものである。またその周辺のニューロンにはシナプス半硬直条件という、更新を行いにくくする条件が適用される。さらに、現在の学習データに対して選択された最大活性化ニューロンがシナプス硬直ニューロンである時、そのシナプス硬直ニューロンの 8 近傍に新しくニューロンを追加する。追加されたニューロンの結合係数ベクトルには、その近傍のシナプス硬直ニューロンに類似した初期値が与えられる。以上より、VDSOM はシナプス硬直条件により問題 2 を、シナプス半硬直条件とニューロンの適応的追加により問題 1 と 3 を解決していることになる。

3. 高速密度可変型自己組織化マップ

AVDSOM は更に問題 4, 5 を解決するように VDSOM を改良した手法である。問題 4 を解決するために、ニューロン増加に伴う処理時間の増加をできるだけ一定にするための局所更新法と局所探索法を導入する。更に、問題 5 を解決するために、不要ニューロン消去法も導入する。この方法には、ニューロンの増加自体を抑える作用もある。

3.1 局所探索法

VDSOM の学習アルゴリズムには、入力に最も類似する結合係数ベクトルを持つニューロンである最大活性化ニューロンを探索するステップがあるが、ニューロン数が増えマップが大きくなると、探索時間がかかりすぎてしまう。これを解決するために、以下の特徴を利用して探索範囲を限定し、処理時間を削減する。

連続性のあるデータを時系列に沿って入力すると、入力ベクトルの値も連続的に変化すると仮定する。一方、VDSOM のマップでは近傍ニューロン間の類似性が保たれている。この二つの特徴から、直前の入力データに対する最大活性化したニューロンの近くに、現在の入力データに対する最大活

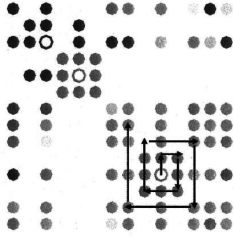


図 1: 局所探索法における探索の様子

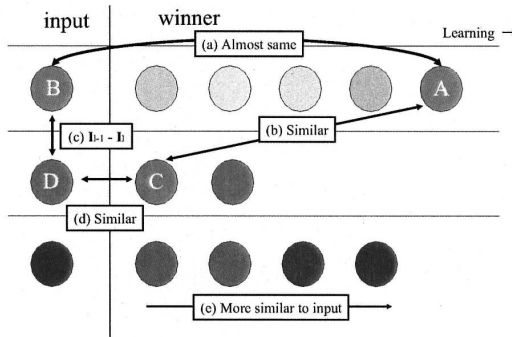


図 2: d_c の決定方法

性化ニューロンがある確率が高い。また、VDSOM は同じデータをシナプス硬直条件を満たすまで何度も学習する。この時は、同じ入力データに対する最大活性化ニューロンを探索するので、同様に直前の学習に対する最大活性化したニューロンの近くに、現在の学習に対する最大活性化ニューロンがある確率が高い。この特徴から、最大活性化ニューロンを探索する時に、直前の入力に対する最大活性化ニューロンの周りから走査し、入力ベクトルとの類似度が高いニューロンが見つければそれを最大活性化ニューロンとして扱っても問題ないと言える。

探索の具体的な方法として、まず直前の入力に対する最大活性化ニューロンの周りから渦巻状(図 1)に $N \times N$ の範囲を探索する。そのなかで最も入力ベクトルと類似するニューロンを選択し、入力ベクトルとの距離が式 (1) に示す d_c より小さければそれを最大活性化ニューロン c とする。

$$d_c = \frac{\|I_{l-1} - I_l\|}{2^{(l+1)}} \quad (1)$$

ここで、 l は同じ学習に対する離散時間、 l は入力データの番号である。 I_l は l 番目に入力されたデータである。図 2 に示すように、前入力データの学習の最後の最大活性化ニューロン A の結合係数ベクトルは、ほとんど前入力データの入力ベクトル B

と同じ値 (a) である。それに対して、入力が切り替わった時最初に選択される最大活性化ニューロン C は、前最大活性化ニューロン A とその近傍ニューロンの中から現入力 D に最も類似するニューロンである (b)。よって、現入力に切り替わった時の入力データ D とその最大活性化ニューロン C の差 (d) は、前入力と現入力の差 (c) に前後することがわかる。一方、ひとつの入力データに対してシナプス硬直条件が適用されるまで学習をする時には、学習の最初の最大活性化ニューロンより、最後の方が入力データとの差はより小さくなる (e)。以上から、 d_c を式 (1) のように決定した。

d_c より入力ベクトルとの距離が大きければ引き続き走査し、 d_c より距離が小さいニューロンを見つけた時点でそれを最大活性化ニューロンとする。

この方法を用いると、マップが大きくなりニューロンが増えても、一定の範囲を探索すれば良いので処理時間の削減になる。この手法を局所探索法と呼び、VDSOM に取り入れる。

3.2 局所更新法

VDSOM は学習過程において、最大活性化ニューロンとその近傍のニューロンを入力データに近づけるように、各ニューロンの結合係数ベクトルを更新する。この時、マップ上で最大活性化ニューロンに近いニューロンほど大きく更新され、最大活性化ニューロンから十分遠いニューロンはほとんど更新されない。本研究では、このようなほとんど更新されないニューロンに対しては更新自体を行わないことにより処理時間を削減する。

VDSOM においてニューロンの更新には、近傍関数の式 (2) により、最大活性化ニューロンからの距離によって更新の重み付けがなされる。

$$h_{c,u}(t) = \exp\left(-\frac{\|x_c - x_u\|^2}{2\sigma^2(t)}\right) \quad (2)$$

x_u はニューロン u のマップ上での位置である。ニューロン c は最大活性化ニューロンを表す。 $\sigma(t)$ は標準偏差の値である。この式 (2) により、最大活性化ニューロンからの距離が十分遠いニューロンに対しては、重みがゼロに近い値となる。事実上、この近傍関数において、マップ上で最大活性化ニューロンからの距離が、標準偏差の値である $\sigma(t)$ の 5 倍以上にあるニューロンに対する重み付けはほとんど 0 である。よって、これ以上遠いニューロンを十分遠いニューロンとみなし、最大活性化ニューロンからの距離が $\sigma(t)$ の 5 倍以内にあるニューロンだけを更新する。

3.3 ニューロンの消去法

VDSOM ではマップに適応的にニューロンを追加し、競合近傍学習によってシナプス硬直ニューロ

間に中間データを多数生成する。この中間データはデータのぶれを吸収し想起能力を向上させるという利点があるが、中にはあまり参照されないニューロンも生成される。このような不要ニューロンは有用性がない上に学習時間を増加させるだけなので、消去を行いより無駄のないマップとして成長させる手法を提案する。

不要になったニューロンとは、入力データからの参照がないため、マップ上でシナプス硬直ニューロンから遠のいてしまったニューロンのことである。各ニューロンに「活性値」をもたせ、シナプス硬直ニューロンに位置的に近いほど活性値を大きく設定することにより、不要ニューロンを探索する。その重みは、シナプス硬直ニューロンの活性値を最大にして、式(3)によってそれぞれのニューロンで硬直ニューロンの活性値の加重平均をとる。

$$v_p = \sum_{q=1}^Q \exp\left(-\frac{\|x_q - x_p\|^2}{2\sigma^2}\right) \quad (3)$$

Q は硬直ニューロンの数、 p は注目ニューロン、 q は硬直ニューロンを示す。活性値の低いニューロンに対して、その淘汰率 r を設定し、該当ニューロンの $r\%$ を消去する。このように活性値の低いすべてのニューロンを消去しないのは、シナプス硬直ニューロンの周辺だけにニューロンが偏ったマップを生成しないようにするためである。ニューロンの消去を行うタイミングは、マップ全体のニューロン数が、シナプス硬直ニューロンの λ 倍を超えたときとする。

3.4 学習アルゴリズム

AVDSOM の学習アルゴリズムを以下に示す。

Step 1 マップ上の全てのニューロン u の結合係数ベクトル \mathbf{W}_u にランダム値を設定する。

Step 2-A 直前の入力に対する最大活性化ニューロンを中心に、渦巻状に $N \times N$ 個のニューロンを走査し、その時点で、式4を満たすニューロンを c' とする。

$$\|\mathbf{I} - \mathbf{W}_{c'}\| = \min_u (\|\mathbf{I} - \mathbf{W}_u\|) \quad (4)$$

Step 2-B c' が式(5)を満たすなら c' を最大活性化ニューロン c とし、**Step 3** に飛ぶ。

$$\|\mathbf{I} - \mathbf{c}'\| < d_c \quad (5)$$

ここで、 d_c は式(1)に示した最大活性化ニューロンを決める閾値である。

Step 2-C 式(5)を満たさなければ引き続き走査をし、(5)を満たした時点でそのニューロンを最大活性化ニューロン c とする。

Step 3 ニューロン c にシナプス硬直条件が適用されている場合、ニューロン c の8近傍に新たにニューロンを追加する。そうでない場合は、**Step 5** に飛ぶ。

Step 4 新たに追加されたニューロン n の結合係数ベクトルの初期値を次のように決める。ニューロン n の最も近くにあるシナプス硬直ニューロン f がニューロン c である場合は式(6)を用い、そうでない場合は式(7)を用いる。

$$\mathbf{W}_n = h_{c,n} \mathbf{W}_c \quad (6)$$

$$\mathbf{W}_n = \frac{h_{c,n} \mathbf{W}_c + h_{f,n} \mathbf{W}_f}{h_{c,n} + h_{f,n}} \quad (7)$$

$h_{c,u}(t)$ は式(2)に示した近傍関数である。

Step 5 式(8)によりニューロン u ($\|x_u - x_c\| < 5\sigma(t)$) の結合係数ベクトルを更新する。ただし、シナプス硬直ニューロンの結合係数ベクトルは更新しない。

$$\mathbf{W}_u(t+1) = \mathbf{W}_u(t) + H(d)\alpha(t)h_{c,u}(\mathbf{I}(t) - \mathbf{W}_u(t)) \quad (8)$$

$\alpha(t)$, $\sigma(t)$ は次式に示す単調減少関数である。

$$\alpha(t) = \frac{-\alpha_0(t-T)}{T} \quad (9)$$

$$\sigma(t) = \sigma_i \left(\frac{\sigma_f}{\sigma_i}\right)^{t/T} \quad (10)$$

ここで、 T は学習回数の上限、 α_0 は学習率の初期値、 σ_i , σ_f はそれぞれ分散 $\sigma(t)$ の初期値と最終値を表す。最後に、 $H(d)$ はシナプス半硬直条件のための関数で、次式で表される。

$$H(d) = \frac{1 - \exp(-d \cdot k)}{1 + \exp(-d \cdot k)} \quad (11)$$

ここで、 d はニューロン u と、このニューロン u の最も近くにあるシナプス硬直ニューロンとのマップ上の距離を表す。また、 k は $H(d)$ の傾きを決定する係数である。

Step 6 シナプス硬直条件 $U < d_f$ を満たすまで、同じ学習データに対して **Step 2** ~ **Step 5** の学習を繰り返す。

Step 7-A シナプス硬直条件を満たすニューロンが現れ、全ニューロン数がシナプス硬直ニューロン数の λ 倍以上であったら、ニューロンの消去を行う。そうでなければ **Step 8** に飛ぶ。

Step 7-B 全シナプス硬直ニューロンの活性値 $v_q (0 \leq v \leq 1)$ を最大値 1 に設定し、その他のニューロンの活性値を式 (3) によって計算する。

Step 7-C 式 (12) を満たすニューロンの $r\%$ を消去する。

$$v_q < TH_{delete} \quad (12)$$

$TH_{delete} (0 \leq TH_{delete} \leq 1)$ は消去に関する閾値である。

Step 8 Step 2 ~ Step 7-C の学習を繰り返す。

4. 欠落特徴点の想起法

AVDSOM とモーションキャプチャの連携の方法は、島田らが提案した方法 [7] に準ずる。モーションキャプチャシステムが出力した 3 次元特徴点がすべて正しく求められている場合には、その情報を AVDSOM で追加学習する。一方、全ての 3 次元特徴点が正しく求められていない場合、正しく推定された特徴点情報のみを利用して、欠落特徴点を AVDSOM で想起する。3 次元特徴点が正しく求められているかどうかの判断には、3 次元特徴点座標に対して「信頼度」を設ける。我々のシステムでは、信頼度によって、AVDSOM が学習をするか想起をするかを切り替えることを想定しているが、3 次元特徴点座標に信頼度を付加する方法については、取り扱うモーションキャプチャシステムに依存するため本研究では取り扱わないこととし、全ての特徴点に適切に信頼度が付加されていることとする¹。

5. 実験と考察

5.1 実験の設定

実験では、カーネギーメロン大学の Graphics Lab が公開しているモーションキャプチャデータベース (<http://mocap.cs.cmu.edu/>) を利用した。本実験で利用したモーションデータは、体にマーカをつけて計測された精度の高いものである。身体姿勢を 26 点の特徴点で示し、各特徴点の回転情報が記録されている。今回は各ノードの 3 次元座標を計算し、その座標値を学習、想起用のデータとして利用する。ここの実験ではビジョンベースの手法で獲得した特徴点情報ではなく、マーカベースで獲得された精度の高い特徴点情報を利用することで、提案手法の有効性を正確に検証するためである。

今回使用した 7513 フレームから成るモーションデータにおいて、最初の 2700 フレームで行った動

¹我々のビジョンベースモーションキャプチャシステム [3, 16] では、人体モデル特徴点とモーションキャプチャにより推定された特徴点の整合性や特徴点の位置関係に基づいた信頼度付加、あるいはフレーム間での特徴点追跡結果に基づく信頼度付加を考えている。

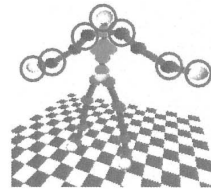


図 3: 実験に利用した特徴点

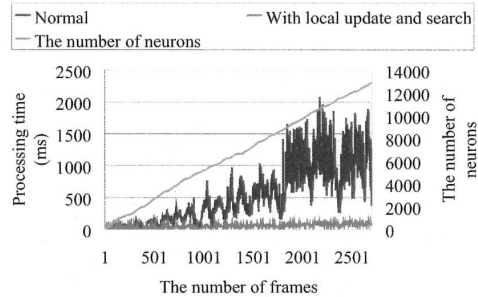


図 4: 処理時間の比較

作を残りのフレームで 2 回繰り返している。このモーションデータはほとんど上半身しか動かず、下半身との相関性が薄いと考えられたので、頭、右肩、右肘、右手、左肩、左肘、左手の 7 点だけを用いて実験を行った (図 3)。従って、AVDSOM で学習する入力ベクトルの次元数は 21 となる。実験で利用した計算機は、Pentium4 3.6GHz の CPU と 2GB のメモリを搭載している。

AVDSOM の初期のマップサイズは 10×10 (100 個のニューロン) とし、各パラメータは $\alpha_0 = 0.3, \sigma_i = 10.0, \sigma_f = 0.3, k = 0.6, T = 1000$ とした。1 フレーム目から順々に AVDSOM に学習データを入力し、1 フレーム分のデータの学習が完了後次のフレームのデータを入力した。従って、同じデータが AVDSOM で学習されるのは 1 回限りである。

5.2 高速化の検証

VDSOM に局所探索法と局所更新法を組み込んだ結果、どれだけ高速化が行えたか検証する。学習用の 2700 フレームのモーションデータを学習し、その処理時間を比較する。局所探索の範囲は $N = 11$ として実験を行った。

図 4 にニューロン数と学習の処理時間の推移を示す。通常の VDSOM はニューロン数の増加に伴って処理時間が大きく増加している。その一方、局所更新法と局所探索法を組み込んだ場合、ニューロン数が増えてもほぼ一定の時間で更新が行えてい

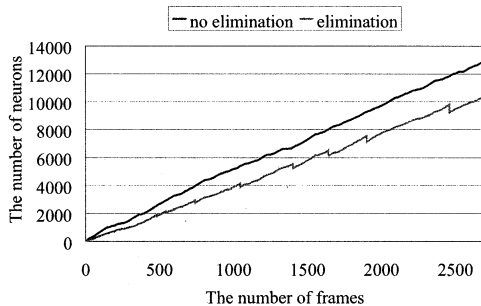
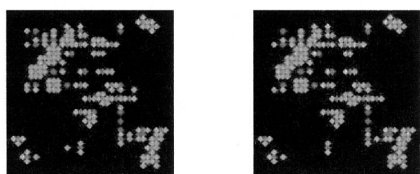


図 5: ニューロン数の推移



(a) 消去前

(b) 消去後

図 6: 消去の前後でのマップの様子

ることがわかる。このときの平均処理時間は $53ms$ である。以上より、これらの高速化手法は効果があったことが証明された。

5.3 不要ニューロン消去の検証

AVDSOM のニューロン消去能力について検証を行った。高速化の実験と同様に、学習用の 2700 フレームのモーションデータを学習した。 $\lambda = 4.0$, $r = 20$, $TH_{delete} = 0.117$ として実験を行った。

ニューロン数の推移を図 5 に示す。最終的な数は、消去ありが 12953 個、消去なしが 10415 個である。グラフ中で急激にニューロンが減っているところで消去が起きており、全ニューロン数が学習したデータ数の約 $\lambda(4)$ 倍の値を保っていることが分かる。消去が起きた時のマップの様子を図 6 に示す。赤色が硬直ニューロンで、緑色が普通のニューロンを表している。硬直ニューロンから離れたことにより活性値が TH_{delete} より小さくなったニューロンのうちの一部が削除されている。この図において、硬直ニューロン数は 48 個で、消去前の全ニューロン数は 198 個、消去後は 189 個であった。

5.4 想起精度の検証

AVDSOM を用いた結果、想起結果にどのように影響が出ているかを調べるために、7513 フレームのうち残りの 4713 フレームに対して故意に右肘の信頼度を 0 として想起を行い、真値との誤差を検証

表 1: 右肘の想起位置の誤差の平均値と標準偏差：ニューロン消去に関する比較 (単位: mm)

	VDSOM	AVDSOM	TM
標準偏差	21.0873	21.2050	21.8089
平均	22.9324	23.0154	24.1019

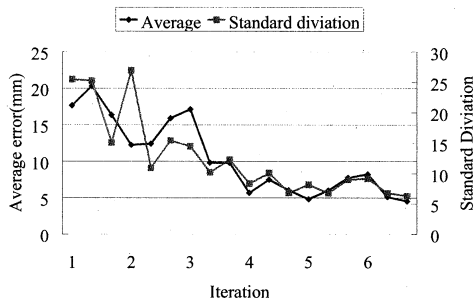


図 7: 右肘の想起位置の誤差の平均値と標準偏差の推移：学習と想起を切り替えた場合

した。また比較のために、2700 フレームの学習データをテンプレートとしたテンプレートマッチング (TM) も行った。表 1 にその結果を示す。VDSOM を用いた結果、TM より良い結果が出ていることがわかる。これは VDSOM は近傍学習により中間データを作成できるので、学習後のニューロンが 2700 フレーム分の以上の姿勢を生成することができるためである。その VDSOM に比べて、ニューロンを約 2000 個消去している AVDSOM の想起精度は VDSOM とほとんど変わらないという結果を得た。このことから、本当に不要なニューロンが消去できているといえる。

5.5 想起と学習を切り替える場合の想起能力の検証

前節で AVDSOM の想起能力の有用性が示されたので、想起と学習を DOC に応じて切り替えるような、より実際の応用に近い実験を行う。7513 フレームのデータを 2 回繰り返す中で、25% の確率で学習を行うように設定する。学習をしない時は、それまでの AVDSOM の学習結果を利用し、右肘の信頼度を 0 として想起を行う。7513 フレームのうち約 900 フレームごとにデータが変わるので、その度に想起誤差と標準偏差を記録した。

結果を図 7 に示す。学習が進むにつれて誤差の平均値と標準偏差が共に減少しているのがわかる。よって、追加学習を行えば行うほどマップがより多くのデータを学習し、想起能力は徐々に向上する

ことが明らかになった。このことから、モーションキャプチャにおいて想起と学習を切り替えながら追加学習していくことは、その後の人体姿勢推定精度を向上させることが確認できた。

6. おわりに

本研究では、実時間ビジョンベースモーションキャプチャシステムの頑健化のために追加学習モデルの高速化の必要性から、VDSOMの改良版である、AVDSOMを提案した。AVDSOMはVDSOMが解決できる追加学習の一般的な問題点を解決できる上に、局所更新法と局所探索法により学習時間を短縮し、不要なニューロンの消去を行いより無駄の無いマップが作成できることが示された。モーションキャプチャへの応用に関して、人体姿勢推定結果に応じて想起と学習を切り替えることにより徐々に想起能力を向上させる方法を提案し、その有用性が証明された。

今後の課題として、人体姿勢推定結果のある時点の姿勢として学習するだけでなく、時系列情報も含めて人体動作を学習、想起する仕組みを考案することが挙げられる。

参考文献

- [1] T. Kohonen. "Self-Organization and Associative Memory". 1989.
- [2] T. Kohonen. "Self-Organizing Maps". 1995.
- [3] N. Date, H. Yoshimoto, D. Arita, and R. Taniguchi. "Real-time Human Motion Sensing based on Vision-based Inverse Kinematics for Interactive Applications". *International Conference on Pattern Recognition*, Vol. 3, pp. 318–321, 2004.
- [4] 田中秀典, 中澤篤志, 竹村治雄. "ポリュームデータの細線化とグラフマッチングを用いた事例ベース人体姿勢推定". 画像の認識・理解シンポジウム (MIRU2007), pp. 1414–1419, 2007.
- [5] 佐川裕一, 下坂正倫, 森武俊, 佐藤知正. "3次元ボクセルを利用した高速オンライン人物姿勢推定". 画像の認識・理解シンポジウム (MIRU2007), pp. 157–163, 2007.
- [6] 谷本貴頌, 星野聖. "自己増殖型SOMを用いた画像データベースからの手指形状の実時間推定". 情報処理学会論文誌 CVIM 13, Vol. 47, No. SIG 5, pp. 100–107, 2006.
- [7] 島田敬士, 叶内円, 有田大作, 谷口倫一郎. "追加学習型自己組織化写像を利用した実時間人体姿勢計測の頑健化". 信学技報パターン認識・メディア理解 (PRMU2007-129), pp. 13–18, 2007.
- [8] 山田剛生, 服部元信, 森澤正之, 伊藤洋. "自己組織化特徴マップを用いた追加学習可能な連想メモリ". 電子情報通信学会 信学技法 NC98-90, pp. 93–100, 1999.
- [9] D. Alahakoon, S. K. Halgamuge, and B. Sirinivasan. "Dynamic Self Organizing Maps With Controlled Growth for Knowledge Discovery". *IEEE Transactions on Neural Networks*, Vol. 11, No. 3, pp. 601–614, 2000.
- [10] Michael Dittenbach, Dieter Merkl, and Andreas Rauber. "The Growing Hierarchical Self-organizing Map". *Proceedings of the IEEE-INNS-ENNS International Joint Conference on Neural Networks*, Vol. 6, pp. 15–19, 2000.
- [11] B Fritzsche. "A Growing Neural Gas Network Learns TopologiesAdvances in Neural Information Processing Systems". *Neural Networks*, Vol. 7, pp. 625–632, 1995.
- [12] Da Deng and Nikola Kasabov. "ESOM: An Algorithm to Evolve Self-organizing Maps from On-line Data Streams". *Proceedings of IJCNN 2000*, Vol. VI, pp. 3–8, 2000.
- [13] F. Shen and O. Hasegawa. "An Incremental Network for On-line Unsupervised Classification and Topology Learning". *Neural Networks*, Vol. 19, pp. 90–106, 2006.
- [14] 島田敬士, 谷口倫一郎. "密度可変型自己組織化マップによる追加学習の実現法". 日本神経回路学会誌, Vol. 14, No. 2, pp. 71–78, 2007.
- [15] 近藤修平, 二見亮弘, 星宮望. "特徴マップを用いたパターン認識における追加学習能力の向上". 信学技報, NC95-161, pp. 55–66, 1996.
- [16] T. Saiki, A. Shimada, D. Arita, and R. Taniguchi. "A Vision-based Real-time Motion Capture System using Fast Model Fitting". *Proceedings of 14th Korea-Japan Joint Workshop on Frontiers of Computer Vision*, pp. 185–190, 2008.