

ニューラルネットワークによる背景想起を利用した 物体抽出システムの開発

藤本 健司* 神田 嵩臣**

Development of Object Extraction Using Recalling Background by Neural Network

Kenji FUJIMOTO* Takaomi KANDA**

ABSTRACT

This paper describes how to develop the object extraction system based on human skill. In the existing studies, researchers have proposed that the object extraction should be approached by background subtraction method or optical flow method. However, these method suffers from a number of disadvantages. In this study, we proposed a new object extraction system using recalling background images by Neural Network. The experimental results showed that the system we proposed were very effective for moving images. We concluded that our system are effective for moving images.

Keywords: neural network, background subtraction method, object extraction

1. はじめに

近年、デジタルカメラやデジタルビデオカメラなどの電子光学機器の発達により、様々な分野で物体認識や移動物体認識技術への需要が高まっている。しかし、画像認識の研究が始まった1960年代から約50年たった現在においても、物体認識においては、未だに人間の顔の正面画像を除いては、実用的な精度で認識が可能な対象はほとんどなく、動画像から、移動物体を抽出・認識するという人間には非常に容易な作業に関しては、未だ実用精度で使用できるものがほとんどないという状況である。

しかし、ここ近年、コンピュータの演算速度や並列処理能力の向上、記憶媒体の記憶容量の増大化などにより、複雑な画像処理が容易に行えるようになってきており、新たな画像認識技術への期待が急速に高まってきている。このような流れの中で、今まで有効的な認識が行えなかった3次元空間(実空間)上を移動する物体を動画像中から抽出・認識する技術について様々な手法が提案され研究されている。

その中でも、主に次の2つの手法が動画像からの移動物体抽出に利用されている。

- (1) 背景差分などの差分処理を用いる方法^{(1),(2)}
- (2) 物体の移動方向を示すオプティカルフローを用いる方法⁽³⁾

しかし、(1)の背景差分法では、照明の変動や背景物体の変動といった背景の変化への対応が難しく、様々な雑音が前景として抽出されてしまう。また、ほとんどの場合でカメラは静止していなければならず、使用には様々な制約が伴う。また、(2)のオプティカルフローを利用する場合には膨大な計算が必要なためリアルタイム処理には向いていないという問題点がある。

そこで、本研究では上記の2つの方法のうち、背景差分の持つ問題を克服した新しい物体抽出・認識手法の開発を目的とする。現在、提案する手法では、まず、差分を行う背景をニューラルネットワーク(Neural Network:以下 NN)⁽⁴⁾が有する画像の記憶・想起ができる性質を利用し、物体を含んだ画像が入力されたときに学習した背景画像を想起させ作成する。この想起した背景画像を使用することにより、雑音や物体が入った時に起こる背景の変動について柔軟な対応ができると考えられる。

* 電子工学科 准教授

** 専攻科 電気電子工学専攻

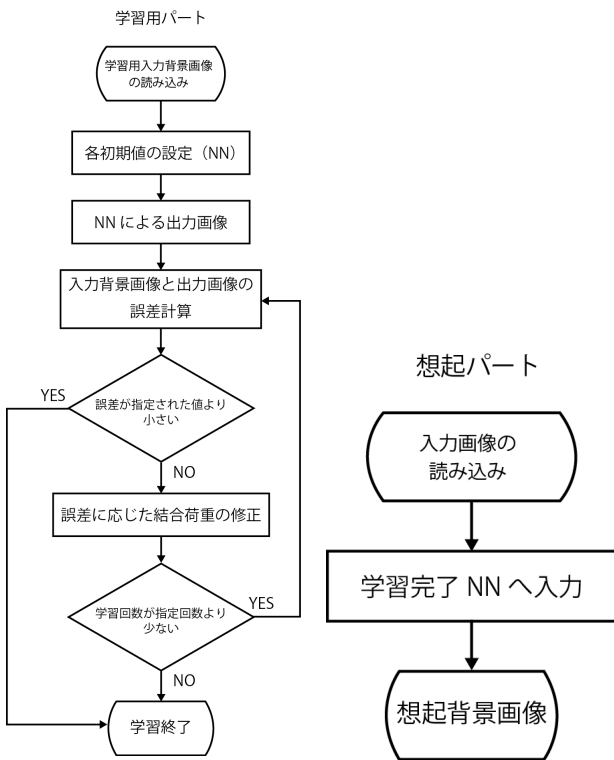
2. 提案するシステム

2.1 システム概要 今回用いる移動物体抽出用システムは図1のようにNNを用いた想起背景用パートと移動物体抽出を行う画像処理パートの2つのパートから構成されている。



図1 提案する移動物体抽出システム

2.1.1 背景作成用 NN ここでは、想起背景作成用 NN の処理の流れと画像処理パートの処理の流れの説明を行う。



(a) 学習用パート (b) 想起用パート
図2 想起背景作成用 NN の処理の流れ

図2に示すのが想起背景作成用 NN の処理の流れである。ここで使用する NN は、図2(a)と図2(b)のような2つの流れから成り立っている。まず、学習用パートでは、学習用の背景画像を複数用意し、NNに学習を行わせる。今回使用している NN では学習用アルゴリズムとして一般的によく使用されている誤差逆伝搬法を用いている。

図2(a)のフローチャートにもあるように、学習終了条件が2つ設定してあり、そのどちらかを満たすと学習が完了する。この学習用パートでは、ケースによって

膨大な計算量を必要とするため処理に時間がかかってしまうが、一度学習が完了すれば、完了時の全結合荷重を保存しておくことで、いつでも瞬時にその状態を再現することができるため想起背景画像を作成する際には学習による計算時間は問題にならない。

図2(b)の想起パートでは、学習完了後の NN に任意の入力画像を入れることで学習時に記憶した背景を想起させる作業を行っている。もちろん、ここで使用する入力画像については学習用の背景画像とは別の画像である。

2.1.2 画像処理パート 次に画像処理パートについて説明する。図3に画像処理パートの流れを示す。また、図中には、想起背景画像作成用 NN からの流れが分かりやすいように、点線部分で囲まれた部分に想起パートの部分からの処理の流れを入れている。

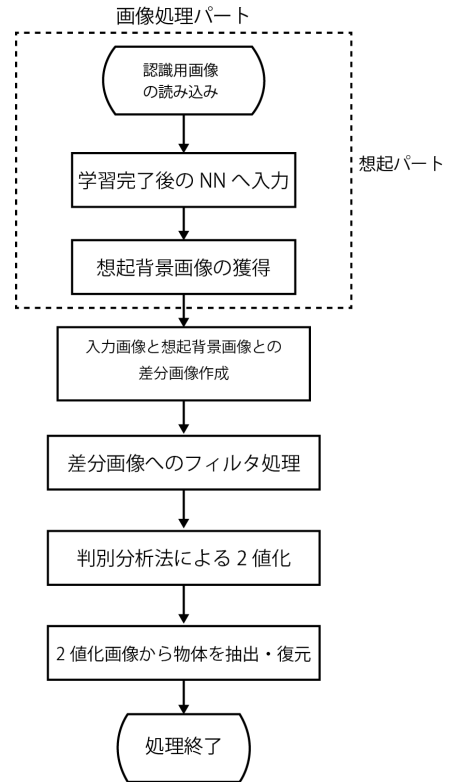


図3 画像処理パートの処理の流れ

まず、学習の完了した NN にカメラなどから取り込んだ画像を入力することで想起背景画像を得る。次に、その背景画像とカメラから取り込んだ画像の差分画像を取る。その次に、得られた差分画像にフィルタ処理を行うが、ここで用いるフィルタ処理には、我々が考案した各色メディアン法とグレーメディアン法を用いた。これらの手法については次節以降で説明を行う。最後に判別分析法により2値化を行い、物体の抽出・復元を行う流れとなっている。

2.1.3 各色メディアン法 ここでは各色メディアン (Each Color Median :以下 ECM) 法について説明する。

ECM 法の概略を図 4 に示す。

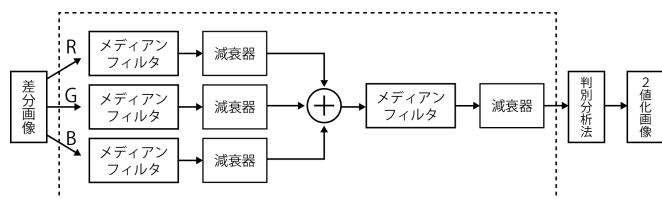


図 4 ECM 法(図中点線部分)

まず、差分画像を RGB のそれぞれ単色の画像に分ける。そして、各画像に対して 3×3 ピクセルのメディアンフィルタを適用する。その後、減衰器により、画像全体の輝度値の平均値を各画素から引く作業を行う。最後に、それぞれの減衰器から出てきた出力を足し合わせ、再び、メディアンフィルタと減衰器を通すことでグレースケール化を行う。

2.1.4 グレーメディアン法 ここでは、グレーメディアン (Gray Scale Median : 以下 GSM) 法について説明する。GSM 法の概略を図 5 に示す。

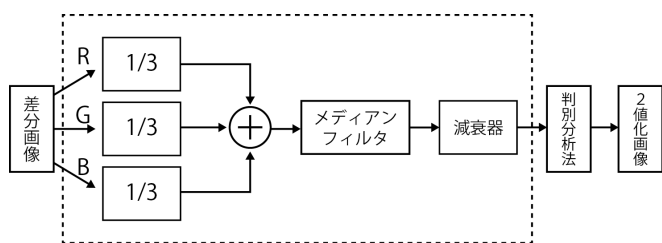


図 5 GSM 法 (図中点線部分)

GSM 法では、RGB 毎の各画像について作業を行うことは同じであるが、メディアンフィルタをかけて平均化するのではなく、単純に各画像の輝度を $1/3$ して足し合わせることで平均化を行っている。その後、メディアンフィルタと減衰器を通すことでグレースケール化を行っている。

2.2 背景差分法と提案する手法との違い 最初に、背景差分法について簡単に説明する。図 6 はあらかじめ設定しておいた基準画像と入力画像の画像間差分をとる最も単純な背景差分法の例である。

背景画像 (基準画像)

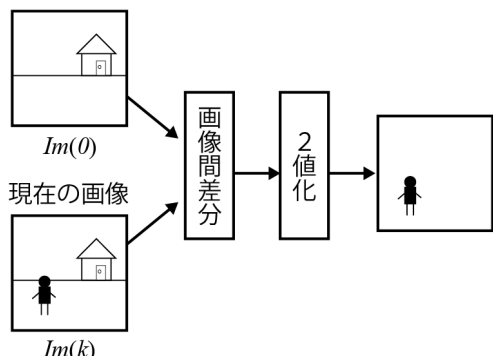


図 6 背景画像との画像間差分による変化検出

ここでは、まず変化のない (移動物体などが無い) 背景画像を基準画像 $Im(0)$ とし、現時点での入力画像を $Im(k)$ としている。そうすると、基準画像と入力画像の画像間差分は、以下の式で与えられる。

$$\text{基準画像との画像間差分} = TH(|Im(k) - Im(0)|) \quad (1)$$

この式より、動きの変化のあった部分の画素が黒として検出される。ここで、 $|Im|$ は画像の各画素値の絶対値計算、 $TH(Im)$ は濃淡画像の 2 値化を意味する。このような単純な方法は、照明変動などの外乱が小さい場合については有効である。

しかし、照明変動などを考慮する場合には、図 7 に示すような連続画像の画像間差分による方法などを使用しなくてはならない。

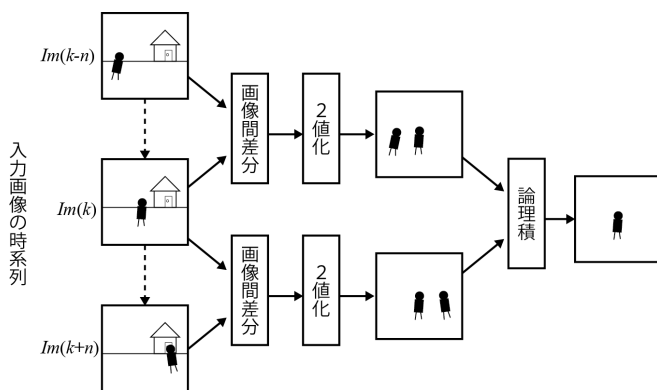


図 7 連続画像の画像間差分による変化検出

一方、提案する手法では、基準画像を用意しなくても、今まで説明してきた通り、NN により入力された画像から基準画像にあたる想起背景画像が提供される。また、この想起背景画像は、NN の低雑音におけるロバスト性を引き継いでいると考えられるため、多少の外乱が入った場合においても、その入力画像に適した背景画像を想起できると期待できる。そのため、図 7 で表したような連続画像の場合においても特別な作業を行う必要がない。

ただし、物体などが画像に入ってきた場合、大きな外乱となるため、想起背景画像に雑音に乗る可能性がある。この部分の問題点を処理する方法として 2.1.3 や 2.1.4 で説明した ECM 法や GSM 法を用いている。

以上が従来の背景差分法と提案する手法との違いとなっている。

3. 実験方法

今回は、通常背景差分法で行われている連続画像間の物体検出と背景差分法が苦手とされている 3 つのケースについて実験を行った。以下にそれぞれのケースを記載する。なお、NN に使用した学習用画像は、

ケース毎に 60[s]~120[s]間撮影した動画から取得した 300 枚の画像を用いた。但し、最後のケースについては 3000 枚の画像を学習画像に用いた。

3.1 連続した画像に関する物体検出 2.2 の図 7 で説明したような連続した画像から、移動する物体の検出を行った。この際、カメラは固定とし、ECM 法や GSM 法を使用しない場合と使用した場合についての実験を行った。図 8 に学習画像の一例を示す。



(a)1 フレーム目 (b)100 フレーム目 (c)200 フレーム目 (d)300 フレーム目
図 8 連続した画像に関する学習画像例

3.2 照度が増加した場合の物体検出 背景差分法が苦手とする照度を変化させた場合の物体検出を行った。カーテンの開閉により照度を変化させ 3.1 と同様に ECM 法や GSM 法を使用しない場合と使用した場合についての実験を行った。図 9 に学習画像の一例を示す。



(a)1 フレーム目 (b)100 フレーム目 (c)200 フレーム目 (d)300 フレーム目
図 9 照度が増加した場合の学習画像例

3.3 手振れするカメラを用いた場合の物体検出 背景差分法では、通常カメラを固定して使用するが、その際カメラが何らかの影響によりぶれたりすると画像間の差分画像が大きく変わってしまう。このケースでは、あえて手振れを起こして検出しにくい状況を作成し、物体検出の実験を行った。図 10 に学習画像の一例を示す。



(a)1 フレーム目 (b)90 フレーム目 (c)180 フレーム目 (d)270 フレーム目
図 10 手振れしたカメラを用いた場合の学習画像例

3.4 カメラをスイングさせた場合の物体検出 3.3 でも述べたとおり、背景差分法のほとんどはカメラを固定して使用する。そこで、カメラをスイング移動させながら、物体検出の実験を行った。図 11 に学習画像の一例を示す。

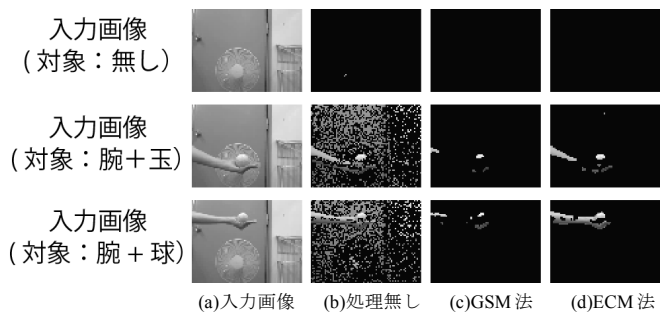


(a)1 フレーム目 (b)200 フレーム目 (c)400 フレーム目 (d)600 フレーム目
図 11 カメラをスイングさせた場合の学習画像例

4. 実験結果

4.1 連続した画像に関する物体検出 連続した画像に関する物体検出の結果を図 12 に示す。単純に想起背景画像と差分を取った後、グレースケール化したものと GSM 法、ECM 法を使用した場合の結果を並べてある。一番左から、入力画像、処理なしの結果、GSM 法の結果、ECM 法の結果に並んでいる。図 12 を見る限り、単純にグレースケール化した場合には、完全に物体の抽出はできておらず、たくさんの雑音領域が乗っていることが分かる。それに対して ECM 法を用いた場合には、物体の抽出部分の一部が少し欠けているケースもあるが物体のほとんどの部分を抽出できている。

また、本ケースにおいては、GSM 法で抽出できた領域は少なく、GSM 法には向いていないということが分かった。



(a)入力画像 (b)処理無し (c)GSM 法 (d)ECM 法
図 12 連続画像を入力した場合の物体抽出結果

4.2 照度が増加した場合の物体検出 照度が増加した場合の物体検出の結果を図 13 に示す。結果の順番は 4.1 と同様である。なお、入力画像 1 と 2 については、学習画像と照度を変えていない。入力画像 3 と 4 については、学習画像と照度が変わっており、入力画像 3 は照度が低く、入力画像 4 は照度が高くなっている。

処理を行わない場合には、照度が増加しない場合は、問題なく物体を検出できているが、照度が増加した場合には、物体以外の雑音を多く含んだ結果となった。ECM 法を使用した場合には、照度が増加しない場合、照度が低くなる場合ともに物体の大部分を検出できているが、入力画像 4 の場合には、画像中の照度の高い部分が差分画像に現れている。しかし、GSM 法を使用した場合には、入力画像 4 のケースでは若干雑音のようなものが残っているが他の手法に比べて良い物体抽

出を行えている。

このことから、GSM法は、何も処理しない場合や、ECM法に比べて急激な変化や外乱に強いと考えられる。

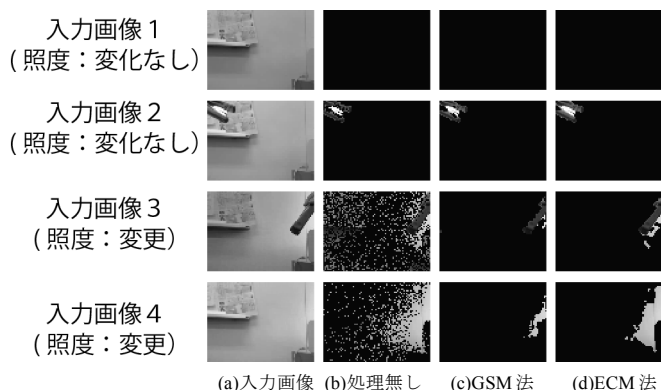


図 13 連続画像を入力した場合の物体抽出結果

4.3 手振れするカメラを用いた場合の物体検出 カメラを手振れさせた場合の物体検出の結果を図 14 に示す。入力画像 1 は対象物体は何も入っていない。入力画像 2~4 には対象物体として、実験者の手の一部とホッチキスが入っている。

処理無しの場合には、広範囲に雑音のようなものに乗っており、移動物体以外の背景の一部についても検出している。また、一部に白い塊のようなものが出ていますが、これは、背景のホワイトボードが照明を反射している部分に反応していると考えられる。

GSM法とECM法を使用した場合には、やはり処理無しの場合の結果と同じような場所に雑音のような白い塊が見られた。ただ、処理無しの場合に比べると広範囲の雑音は抑えられており、GSM法、ECM法共に移動物体の検出が行えていることが分かる。

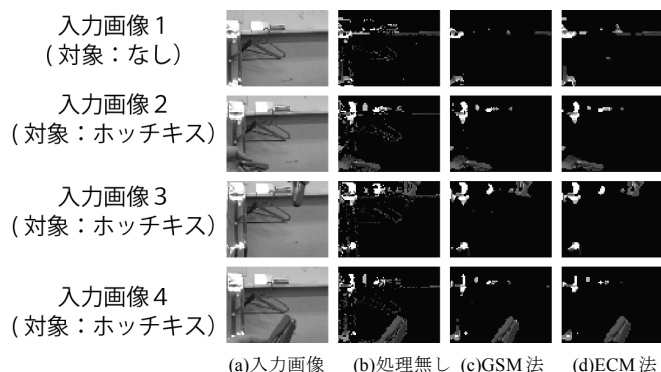


図 14 手振れしたカメラを用いた場合の物体抽出結果

4.4 カメラをスイングさせた場合の物体検出 カメラをスイングさせた場合の物体検出結果を図 15 に示す。

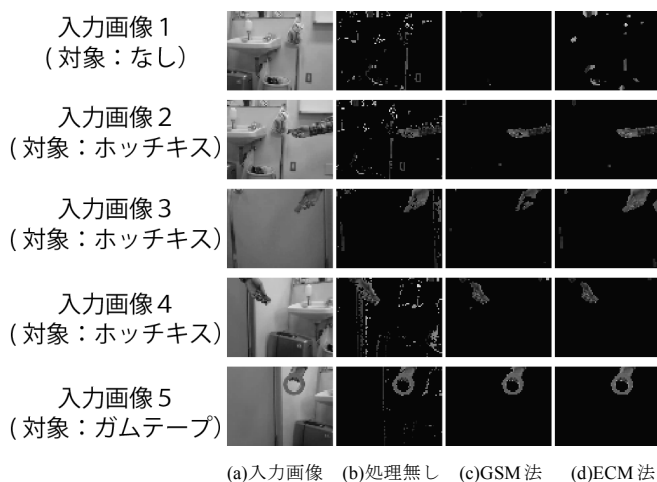


図 15 カメラをスイングさせた場合の物体抽出結果

入力画像 1 には、対象物体が入っていないものを使用し、入力画像 2~4 に関しては対象物体として実験者の手の一部とホッチキスを使用した。入力画像 5 では、対象物体として、実験者の手の一部とガムテープを用いた。

処理無しの場合には、どの入力画像に対しても物体の検出は行えているが、やはり広範囲にわたって必要のない部分が検出されてしまう結果となった。

GSM法を用いた場合には、全ての入力画像に対して、処理無しの場合に見られた雑音のほとんどが消えており、物体の抽出を行うことができています。しかし、物体の一部が欠けるなど、全ての部分が抽出されている訳ではない。入力画像 3 を入れたときの結果はその傾向が顕著に表れている。

ECM法を用いた場合には、やはり処理無しの場合に比べて雑音は抑えられているが、一部照度の変化の大きい部分については、誤検出されてしまっている(入力画像 1, 入力画像 3)。しかし、GSM法のように抽出した物体が大きく欠けたりすることはなく、物体の形状を良く検出できていることが分かった。

5. まとめ

今回 4 つのケースについて提案するシステムを用いて実験を行った。その結果を表 1 にまとめる。表 1 を見ても分かるように、想起背景画像を用いて得られた差分画像をそのまま利用する方法だと、どのケースにおいても対象物体は抽出できているが、物体以外にも抽出している部分が多く、そのままでは物体抽出として使用するには問題がある。GSM法を使用する場合には、カメラを固定した連続画像では、良い結果を得られなかったものの、その他のケースでは処理無しのように多く見られた誤抽出の部分のほとんどを消去した抽出画像を得ることができた。ECM法を用いた場合においても、照度などの外乱が大きくなりすぎると一部誤抽出が生じてしまうが、外乱がさほど大きくないケ

ースにおいてはかなり正確に物体の抽出を行うことができた。

表1 各種手法を用いたときの検出結果

手法 ケース	想起背景を そのまま利用	GSM 法を使用	ECM 法を使用
連続画像 (カメラ固定)	抽出可能 (雑音が非常に多い)	ほとんど抽出できず	抽出可能 (一部画像に欠け)
照度の変動	抽出可能 (雑音が非常に多い)	抽出可能 (一部雑音あり)	抽出可能 (一部雑音あり)
手振れ	抽出可能 (雑音が非常に多い)	抽出可能 (一部雑音あり)	抽出可能 (一部雑音あり)
スイング	抽出可能 (雑音が非常に多い)	抽出可能 (一部画像に欠け)	抽出可能 (一部雑音あり)

以上のことから、今回の実験結果より、背景差分法で苦手とするようなケースにおいて、提案する手法は、それぞれのケースにおいて特別な処理を必要とせず、有効な結果を得ることができたと言える。今回の実験結果では定量的に評価できなかったため、今後は定量的に評価する手法などを取り入れ、より本システムの有効性を確認していく予定である。

参考文献

- (1)篠原利章 他：“背景差分法における移動物体の検出精度の向上とカメラの揺れによる誤検出の低減手法の検討”，第2回画像センシングシンポジウム講演論文集,pp.249-254,1996
- (2)佐久間喜郎，伊東潔，増田功：“フレーム間差分を用いた侵入物体検出法”，テレビジョン学会技術報告,IPCV '90-27, AIPS '90-50, pp.1-6, 1990.9
- (3)岡田隆三，白井良明，三浦純，久野義徳：“オペティカルフローと距離情報に基づく動物体追跡”，信学論D-II Vol.J80-D-II No.6, pp.1530-1538,1997.6
- (4)吉富康成，“ニューラルネットワーク”，朝倉書店,2002