

ニューラルネットワークによるナンバープレートの位置検出

藤吉 弘亘[†] 梅崎 太造[†] 今村 友彦[†] 金出 武雄^{††}

Area Extraction of the License Plate Using Artificial Neural Network

Hironobu FUJIYOSHI[†], Taizo UMEZAKI[†], Tomohiko IMAMURA[†],
and Takeo KANADE^{††}

あらまし 本論文では、ナンバープレートの重心位置を出力するように学習させたニューラルネットワークを使用して、プレート領域を抽出する方法を提案し、その有効性を示す。学習パターンの提示位置は、乱数で与えられるため、毎回異なるパターンがニューラルネットワークに学習され、汎化能力の高いニューラルネットが形成される。学習パターン中に含まれるナンバープレートの最適学習面積、バンパーとヘッドライト部分の抑制学習による効果、および学習パターンの拡大縮小と濃度値変換による効果について検討する。地下駐車場で撮影された595台の車に対して評価実験を行った結果、それぞれ98.5%、98.7%および100%の検出率を得た。

キーワード ナンバープレート、ニューラルネットワーク、バックプロパゲーション

1. まえがき

自動車は、機動性に優れた乗物として広く利用されており、今日の生活になくてはならないものの一つである。しかしながら、その入退出に関するセキュリティシステムは、いまだ十分に完備されておらず、犯罪に利用される場合も多い。そのため、不審者の出入りを監視する必要性の高い学校、病院、研究機関、および行政機関等の公共的施設においては、車の入退出管理が早急に要求されている。また、駐車場や高速道路等での車の入退出管理を考えた場合、料金所における人員の削減、チケットの紛失、支払い作業にかかる時間等の問題、およびスピード違反や盗難車の発見にも貢献できる。

車を識別するための要素としては、車種、車色、ナンバープレート（以下プレートと略す）等が挙げられるが、この中で個々の車を唯一同定できるのはプレートのみであり、所有者等の情報を調べることも可能である。

プレートの文字情報を自動認識するには、まずプレートがどの位置にあるかを高精度に検出することが

重要であり、その後の文字領域切出しと切出された文字の認識性能に大きく影響する。

既に、プレートの位置検出方法としては、ハフ変換を用いてプレート輪郭（垂直線と水平線）の直線性を求める方法や、垂直エッジを追跡する方法等が報告されている[1]~[7]。しかし、これらの方法では、プレートと車両が同系列色であったり、暗い環境下で採取された映像の場合認識できないときがある。

本論文では、プレートの輪郭線のみ注目するのではなく、プレート全体の特徴（プレートらしき）を利用する方法について述べる。まず2.では位置検出システムの構成を示し、3.でプレートの位置検出をニューラルネットワークを用いて行う場合のネットワークモデル学習法、および検出方法について述べる。次に4.で位置検出実験に用いるプレートの画像データベースについて述べ、5.では今回提案する位置検出システムの有効性を示す。

2. 位置検出システムの構成

プレートの位置検出システムは図1のようにビデオカメラ白黒256階調のA-D変換器および位置検出部により構成される。

3. プレート位置の検出方法

3.1 ネットワークモデル

プレートの重心位置検出に使用するニューラルネッ

[†] 中部大学工学部電子工学科, 春日市
Faculty of Engineering, Chubu Univ., 1200 Matsumoto-cho
Kasugai-shi, 487 Japan

^{††} カーネギーメロン大学ロボティクス研究所, 米国
Robotics Institute, Carnegie Mellon University, 5000 Forbes Ave.,
Pittsburgh, PA 15213, USA.

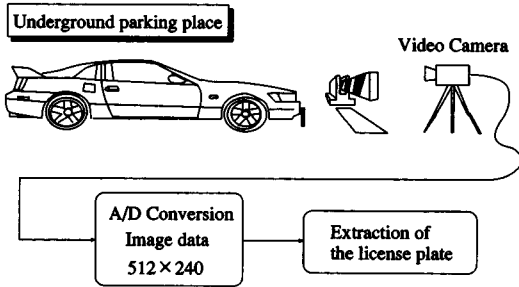


図1 システム構成
Fig.1 System.

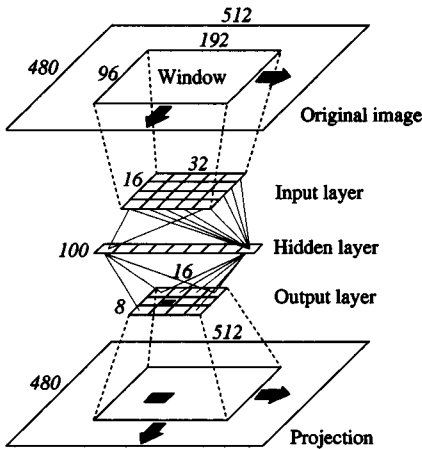


図2 ネットワークモデル
Fig.2 Neural network model.

ネットワークは3層の階層型で学習にはバックプロパゲーション法を用いる [8]~[10]. ユニットのしきい値関数には次式で示されるシグモイド関数を使用する.

$$f(x) = \frac{1.0}{1.0 + \exp(-x/U_0)} \quad (1)$$

ネットワークモデルを図2に示す. 各層の構成ユニット数は以下のとおりである.

- Input layer : 512 (32×16)
- Hidden layer : 100
- Output layer : 128 (16×8)

3.2 学習パターンと教師信号

まず, 学習パターンの作成方法について述べる. 入力画像から切り出した 192×96 の小領域に, 3×3 サイズの sobel 変換でエッジ強調を施し, 32×16 に圧縮したものを学習パターンとする (図3参照). データを平滑化することで“プレートらしさ”の情報を生成している. なお, 切出しポイント (小領域の左上座標) の決定は乱数で行う.

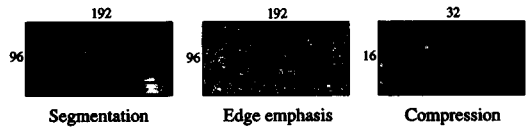


図3 入力パターン例
Fig.3 Example of input pattern.

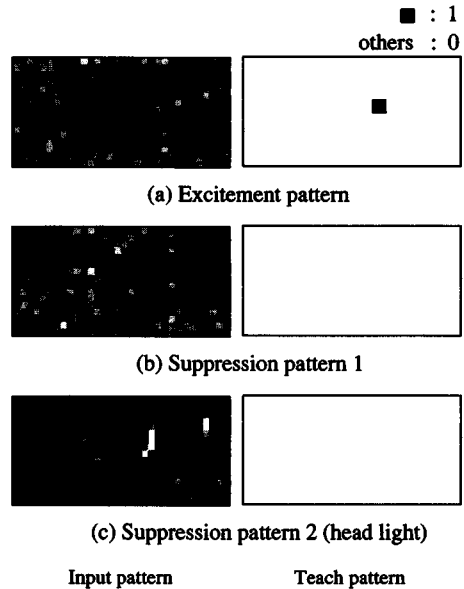


図4 学習パターン例
Fig.4 Example of learning pattern.

学習パターンには図4のように興奮パターンと抑制パターンを用いる. ここで興奮パターンは, 切り出した領域の中にプレート面積が α%以上含まれているものとする (図4(a)参照). また抑制パターン1は, プレート面積の 25 ~ α%未満の領域が含まれているもので, 抑制パターン2はプレート形状に類似したヘッドライト部分である (図4(b), (c)参照). 興奮パターンに対する教師信号はあらかじめ視察によって求めたプレートの重心座標に対応するユニットを1, その他を0とする. また, 抑制パターンに対しては, すべてのユニットを0とする.

学習パターンの提示方法として, ニューラルネットワークに興奮パターンのみを提示する方法と, 抑制パターンも混在して提示する方法の2種類について検討する. なお, 学習の終了回数は, 30000 とする.

3.3 プレート面積の計算方法

学習パターン中に含まれているプレートの面積の度

合によって、ニューラルネットワークの学習しやすさや性能が変化する。図5のように W を切出し窓の横幅、 H をその高さとする、切出し窓の面積 S_w は次のようになる。

$$S_w = HW \tag{2}$$

次に、プレート面積 S_p は、 (x_1, y_1) をプレートの左上、 (x_2, y_2) をプレートの右下の座標とすると次式で計算できる。

$$S_p = (x_2 - x_1 + 1)(y_2 - y_1 + 1) \tag{3}$$

切出し画像中に含まれるプレート面積 S_i は次式で表される。

$$S_i = S_w \cap S_p \tag{4}$$

具体的には (R_{xi}, R_{yi}) を切出し窓の左上の座標とすると、次式で計算できる。切出し窓とプレートの位置関係によって (a)~(i) の場合に分けられる。

$$S_i = \begin{cases} (a) (x_2 - R_{xi} + 1)(y_2 - R_{yi} + 1) \\ (b) (R_{xi} + W - x_1 + 1)(y_2 - R_{yi} + 1) \\ (c) (x_2 - R_{xi} + 1)(R_{yi} + H - y_1 + 1) \\ (d) (R_{xi} + W - x_1 + 1)(R_{yi} + H - y_1 + 1) \\ (e) (x_2 - R_{xi} + 1)(y_2 - y_1 + 1) \\ (f) (R_{xi} + W - x_1 + 1)(y_2 - y_1 + 1) \\ (g) (x_2 - x_1 + 1)(y_2 - R_{yi} + 1) \\ (h) (x_2 - x_1 + 1)(R_{yi} + H - y_1 + 1) \\ (i) (x_2 - x_1 + 1)(y_2 - y_1 + 1) \end{cases} \tag{5}$$

故に、面積 S_i の S_p に対する割合 α [%] は次式のようにになる。

$$\alpha = \frac{S_i}{S_p} \times 100[\%] \tag{6}$$

3.4 学習誤差曲線

ニューラルネットワークの学習を行う際の誤差について述べる。ここで、ニューラルネットワークに学習させる切出しパターンは、プレート面積の60%以上を含む。このとき、 n 回学習したときの学習誤差を次式のように定義する。

$$E_n = \frac{1}{N} \sum_{i=1}^N (O_i - T_i)^2 \tag{7}$$

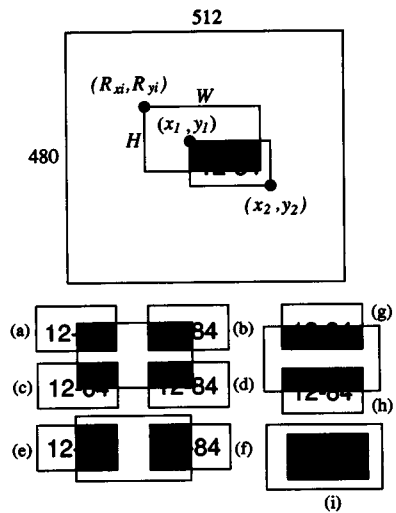


図5 面積の計算方法
Fig.5 Calculation method of area.

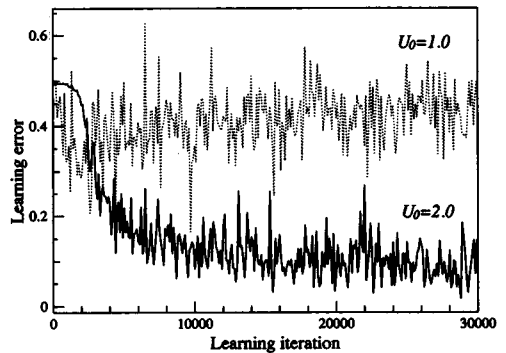


図6 学習誤差曲線
Fig.6 Learning error curves.

- N : 学習に用いる入力画像の数
(ここでは $N = 20$)
- O_i : 学習パターン i におけるニューラルネットワークの出力信号
- T_i : 学習パターン i に対する教師信号

学習誤差曲線を図6に示す。シグモイド関数の傾き U_0 を1.0とすると、学習誤差は発散するが、2.0だと誤差は振動しながら減少する。これは、学習パターンの切出し位置が、乱数によって決定されるからである。毎回異なる学習パターンが提示されるため、若干神経の応答を鈍くする方がよいことがわかる。

3.5 重心位置の学習効果

ニューラルネットワークが、プレートの重心位置を

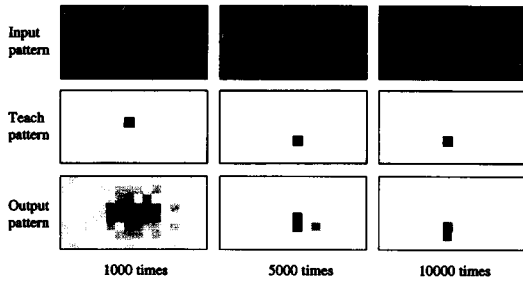


図7 重心位置の学習とそのときの検出能力
Fig.7 Learning of center position and detection ability.

学習する過程について述べる。図7は、入力信号に対して、重心位置を学習させるときの教師信号と各学習回数における出力を観察したものである。ここで、シグモイド関数の傾きは2.0である。1000回学習したときは、出力信号は一点に定まらない。しかし、5000回、10000回と学習が進行するにつれて、出力信号の精度が向上していくことがわかる。

3.6 ガウス窓による検出精度の向上

ニューラルネットワークの出力値を512×480のサイズ上に投影する場合、投影ポイントは離散的分布をなす。そのため、重心付近でなくても、投影加算値が大きくなることはしばしば起こる。そこで、出力値に次式に示すような2次元のガウス窓 $W(x, y)$ を掛けて投影加算することにより、近似的に連続的な分布を作り、検出精度の向上を目指す。

$$W(x, y) = \exp\left(-\frac{x^2 + y^2}{2\sigma^2}\right) \quad (8)$$

$$\left(-\frac{l}{2} < x < \frac{l}{2}, -\frac{l}{2} < y < \frac{l}{2}\right)$$

σ^2 : 2次元ガウス窓の分散

l : ガウス窓の大きさ

3.7 検出方法

プレート重心の位置検出は、以下の手順で行う。

(1) 3.2の学習パターンの作成と同じ方法で、切り出した圧縮パターンをニューラルネットワークに入力して出力値を計算する。

(2) 16×8の出力値を192×96の大きさに投影拡大し、ガウス窓を掛けて加算する。

(3) 図8のように、切出し窓を任意のシフト幅で x, y 両方向に走査する。ここで用いるシフト幅は、8, 16, 32, 64[pixel]である。

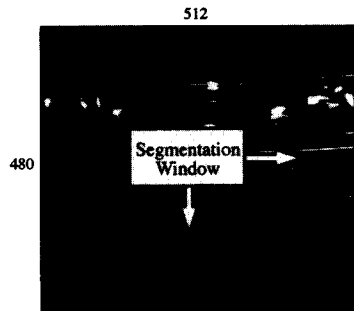


図8 切出し窓の走査
Fig.8 Scanning method of the segmentation window.

(4) (1)~(3) を繰り返す、入力画像を全探索する。

(5) 投影加算した中で最も出力値が大きい位置をプレートの重心とする。

3.8 評価基準

あらかじめ視察により求めた重心位置を (x_0, y_0) ニューラルネットワークにより自動検出された位置を (x', y') とすると、そのユークリッド距離 d は、次式のようになる。

$$d = \sqrt{(x_0 - x')^2 + (y_0 - y')^2} \quad (9)$$

ここで $d \leq 12$ [pixel] のとき検出成功とする。12以内とはニューラルネットワークの出力層において入力画像の1/12の大きさに圧縮されているためである。

4. 画像データ

評価実験に使用する画像は、実環境を想定して地下駐車場で一時停止した自動車を正面から CCD カメラにより撮影したものである。ビデオ映像を512×240画素、白黒256階調に量子化してあるが、実際には画像の縦横比が1:1になるように、横2ラインごとと同じデータで埋め、512×480のサイズにしてある。図9に入力画像例を示す。595台分の画像をデータベースとする。なお、データベースの中には、緑色地のプレートが全体の22%程度含まれている。

自動車の停止位置により撮影距離が変動しているため、プレートの大きさは一定ではない。595台のデータに対するプレート横幅の分布を図10に示す。平均値122画素から±20画素以上差のあるものが13台含まれる。一般に多量の学習データを用いれば、汎化能力の高いニューラルネットワークを構築できるが、可能

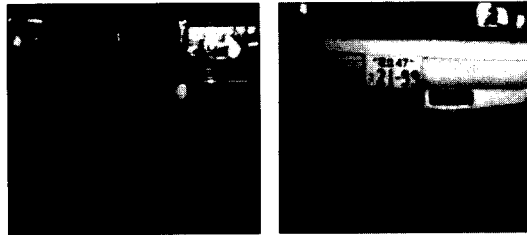
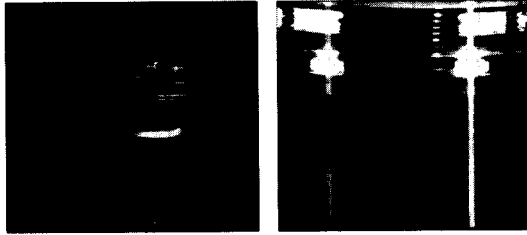
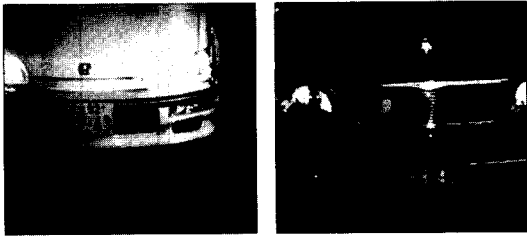


図9 入力画像例
Fig. 9 Examples of input image.

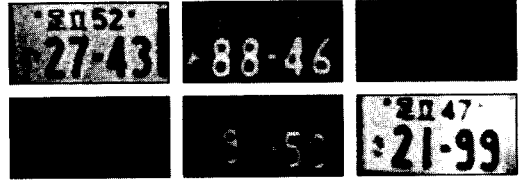


図11 検出成功例
Fig. 11 Examples of success.

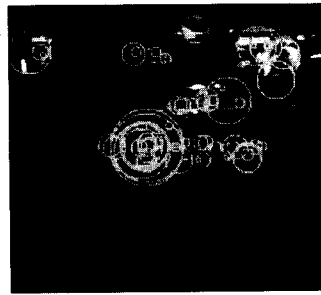


図12 投影加算例
Fig. 12 Example of projection.

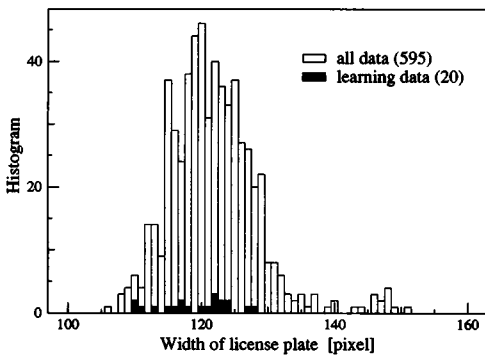


図10 プレート幅の分布
Fig. 10 Width distribution of license plate.

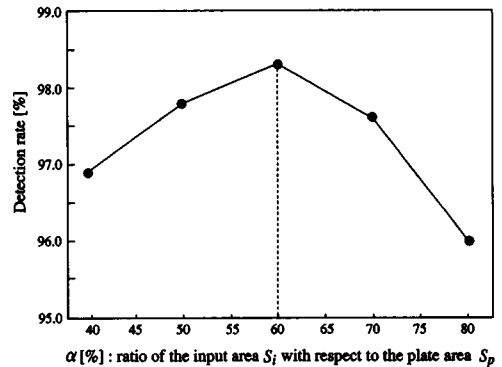


図13 最適学習面積
Fig. 13 Optimum learning area.

な限り少量のデータで学習できることが望ましい。本論文で用いるニューラルネットワークの学習用データとしては、595台の中から無作為に取り出した20台分のデータを用いる。また、昼間の屋外で撮影した6台分を5.4で述べる拡大縮小による学習パターン増加の実験に用いる。

5. 評価実験

5.1 最適学習面積の検討

学習パターンに含まれているプレート面積が小さい場合、正確な重心位置をニューラルネットワークに教えこむのは困難であり、検出誤りが増加する一原因となる。そこで、プレート面積が少ない学習パターンは興奮学習に含めないようにする。切出し窓のシフト幅は xy 両方向ともに32[pixel]とする。図11に、プレート面積が60%以上含まれている切出しパターンを学習したときの検出成功例を示す(式(9)参照)。図12

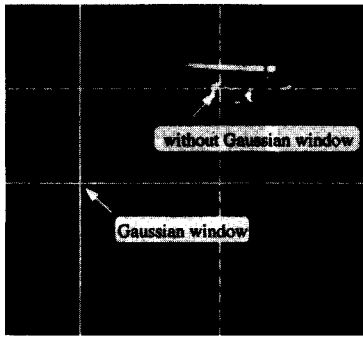


図 14 ガウス窓の効果
Fig.14 Effect of Gaussian window.

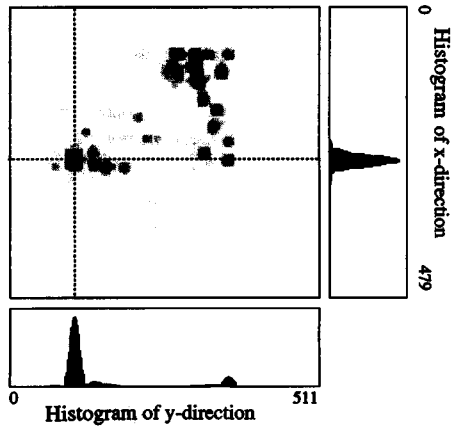


図 16 ガウス窓掛け時の投影加算値
Fig.16 Example of projection using Gaussian window.

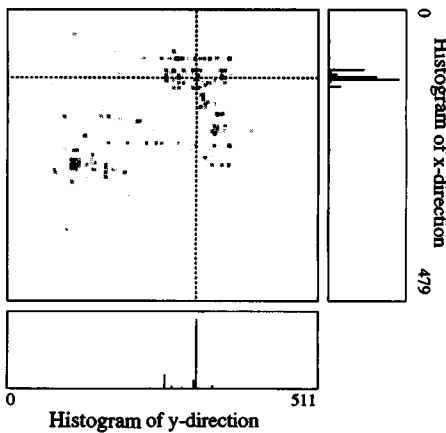


図 15 ガウス窓なしの投影加算値
Fig.15 Example of projection without Gaussian window.

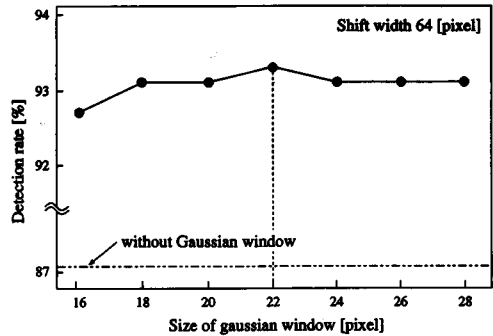


図 17 ガウス窓の大きさと検出率
Fig.17 Detection rate with respect to the size of Gaussian window.

は投影加算例である。円の大きさは、ニューラルネットワークの反応の強さである。プレートの重心付近にニューラルネットワークの反応が集中しているが、プレート周辺部やヘッドライト部分にも検出誤りが見られる。

位置検出結果を図 13 に示す。プレート面積が 60%以上含まれるときの検出率が最もよく、97.5 (580/595)%である。以後、この値 ($\alpha = 60$) を最適学習面積とする。

5.2 ガウス窓の効果

正方形なガウス窓を出力値に掛けて投影加算を行う場合の検出率に与える影響について検討する。ここで、窓の大きさを 16~28 画素まで変化させて実験する。より明白な効果を出すため切出し窓のシフト幅は、 xy 両方向ともに 64[pixel] とする。図 14 にガウス窓の効果例を示す。

最大出力値がヘッドライトからプレート重心位置に移動している。このときの投影加算値を図 15、図 16 に検出結果を図 17 に示す。ガウス窓なしの場合と比較すると検出率が 87.1 (518/595)% から 93.3 (555/595)% に向上した。ガウス窓の大きさは 22 画素のときが最もよく、以後の実験では、この値を用いる。

5.3 抑制学習

プレートの最適学習面積を求める実験で、プレート周辺部分とヘッドライト部分に検出誤りが多いことが判明している。そこで、これら領域の学習時、出力を抑制するように学習させる。抑制学習後の検出結果を図 18 に示す。シフト幅 32[pixel] での検出率は、98.7 (587/595)% で 1 台分よくなっている程度で、顕著な差は特に見られない。しかし、64[pixel] のシフト幅では、プレート周辺部分のみを抑制したときの検出率

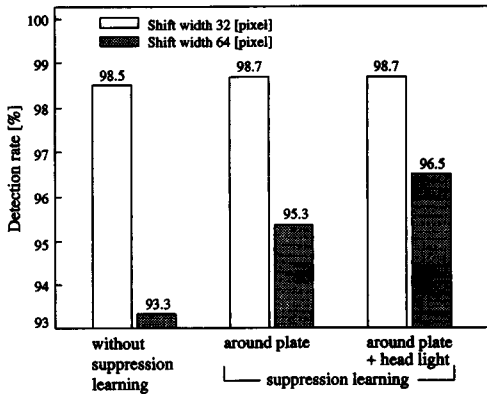


図 18 抑制学習による効果
Fig. 18 Effect of suppression learning.

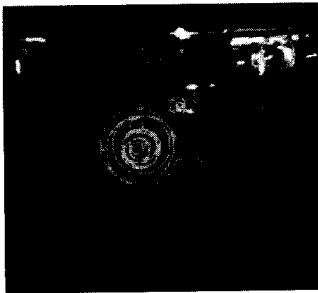


図 19 抑制学習後の投影加算例
Fig. 19 Example of projection after suppression learning.

は、93.3 (555/595)%から 95.3 (567/595)%に向上した。更に、ヘッドライト部分も抑制することにより、96.5 (574/595)%の検出率が得られた。

図 19 に抑制学習後の投影加算例を示す。ニューラルネットワークのプレート周辺部分やヘッドライト部分の出力誤りが図 12 と比較して減少していることがわかる。

5.4 拡大縮小による学習パターンの増加

学習に用いた 20 台の画像に拡大縮小変換を行い、学習パターン数を疑似的に増加したときの重心検出結果を図 20 に示す。シフト幅 32[pixel] のとき、99.3 (591/595) %に検出率が向上したが、プレート部分の階調が著しく低いときと、プレート周辺形状が複雑なときに検出誤りがある。

これらの欠点を補うため、595 台以外のデータベースからプレート周辺形状が複雑な 3 台と、プレート部分の階調を低くした 3 台を新たに学習データとして加えた。シフト幅 64[pixel] のとき 98.3 (585/595)、

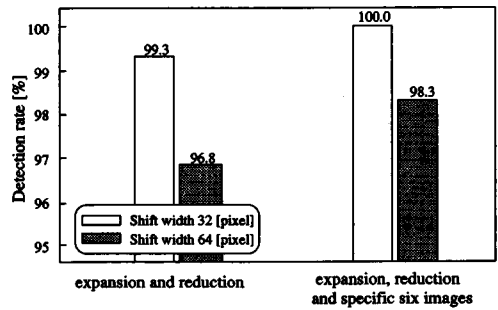


図 20 学習パターンの増加による効果
Fig. 20 Effect of increasing learning samples.

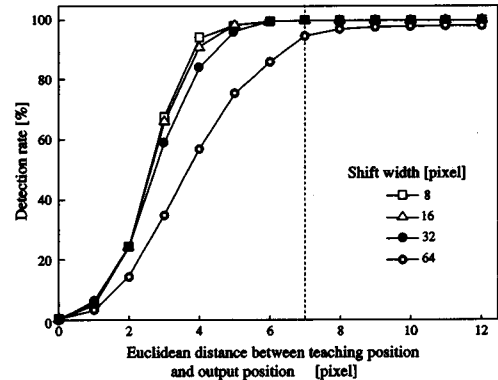


図 21 シフト幅の影響
Fig. 21 Effect of shift width.

32[pixel] 以下で、100.0 (595/595) %の検出率を得た(図 21 参照)。

なお、1 台当りのプレート抽出に要する時間は、シフト幅 32[pixel] のとき Alpha AXP/300 で約 1 秒シフト幅 64[pixel] のとき 0.3 秒である。

6. む す び

本論文では、ニューラルネットワークを使用したプレートの位置検出法を提案し、その有効性を明らかにした。学習パターンの提示位置は、乱数により決定されるため、同じデータがネットワークに学習されることは基本的にない。

まずはじめに、切出し窓に含まれるプレートの最適学習面積と出力値の投影加算時に掛けるガウス窓の効果について調べた。学習パターンにプレート面積の 60%以上が含まれるものを用いた場合、595 台の車に対してシフト幅 32[pixel] のとき 97.5%の検出率を得た。また、ガウス窓を掛けることにより、離散分布と

なるニューラルネットワークの出力値を連続分布に近い状態へ置き換えることで更に検出率は向上した。

次に、抑制学習の効果について検討した。プレート周辺部分やヘッドライト部分における出力値を抑制することにより、ニューラルネットワークの出力誤りが減少し、98.7%の検出率を得ることができた。更に、入力画像に拡大縮小変換を施して、学習パターン数を増加することで、99.3%の検出率を得ている。疑似的に学習パターン数を増加することでニューラルネットワークの汎化能力が向上することを示した。

最後に、595台以外のデータ6台をこれまでの学習パターンに付加して使うことにより、100.0%の検出率を得ることができた。

本手法を用いることにより、従来困難とされてきたプレートと車両が同系列色の場合でも、高精度な位置検出が行えることを明らかにした。

謝辞 本論文で使用した画像データを提供して下さったオムロン(株)ファジィ推進センター大八木雅之センター長、開発課江島秀二氏に感謝致します。

文 献

- [1] 高橋裕子, 前田英作, 塩 昭夫, 石井健一郎, “駐車場管理自動化のための画像認識技術,” NTT R&D, vol.41, no.4, pp.493-500, 1992.
- [2] 飯田泰久, 中山博之, 宮本一正, 藤田一郎, 浦田秀夫, “高速画像処理装置の開発とその応用,” 三菱重工技報, vol.27, no.1, pp.76-80, 1990.
- [3] 出口正俊, 加藤健司, 三谷雅古, 日根野谷俊男, “旅行時間計測用ナンバープレート読取装置の開発,” 住友電気・第139号, pp.8-13, 1991.
- [4] 加藤博光, 南部 元, 福井英登, 依田文夫, 青木仁志, “ナンバープレート認識技術,” 三菱電機技報, vol.62, no.2, pp.8-12, 1988.
- [5] 崔 亨振, 安居院猛, 中嶋正之, “アダプティブパラメータ平面領域制限ハフ変換を用いたナンバープレート領域抽出法,” 信学論 (D-II), vol.J72-D-II, no.4, pp.597-604, April 1989.
- [6] 尾上博和, 塩野 充, “漢字部分を含めたナンバープレート全体の文字認識実験,” 信学技報, PRU92-47, 1992.
- [7] 三島忠明, 金崎守男, 高藤政雄, 鈴木優人, 小林芳樹, 太田秀夫, 柴田敏郎, “画像処理応用車番認識装置の検討,” 信学技報, PRU86-94, 1986.
- [8] 今井勝次, 郷原一寿, 内川嘉樹, “3層BPモデルを用いた横書き文字列の認識,” 信学技報, PRU91-3, 1991.
- [9] 西村治彦, 小山直樹, “バックプロパゲーション法における学習性能の入力パターン分解能依存性,” 情報処理学会論文誌, vol.35, no.11, pp.2331-2337, 1994.
- [10] D.E. Rumelhart, J.L. McClelland, and the PDP Research Group, “Parallel Distributed Processing,” 1, The MIT Press, 1986.

(平成8年10月21日受付)



藤吉 弘亘 (学生員)

平4中部大・工・電子卒, 平6同大学院修士課程了。平9, 同大学院博士後期課程了。現在, 同大学研究生。画像処理, パターン認識の研究に従事。



梅崎 太造 (正員)

昭57豊橋技科大・工・情報卒, 昭62名大大学院博士課程了。同年4月名大助手。平2中部大講師。平4より同大助教授。平5~6カーネギーメロン大学客員研究員。工博。音声・画像情報処理, 聴覚障害児教育の研究に従事。情報処理学会, 日本音響学

会各会員。



今村 友彦 (正員)

平6中部大・工・電子工卒。現在, 日立信情報通信。在学中, 画像認識の研究に従事。



金出 武雄 (正員)

昭43京大・工・電子卒, 昭48同大学院博士課程了。同年同大情報工学科助手。昭51同助教授。昭55米国カーネギーメロン大学計算科学科高等研究員。昭60同教授。現在, 同 U.A. and Helen Whitaker 教授。同大ロボティクス研究所所長。工博。計算機視覚, ロボットの腕, 自律走行車, VLSI センサに関する研究に従事。IEEE Fellow, AAAI Fellow。