

ナンバープレートの位置検出

今村 友彦*, 藤吉 弘亘*, 梅崎 太造**, 金出 武雄***

Area Extraction of the License Plate

Tomohiko IMAMURA*, Hironobu FUJIYOSHI*, Taizo UMEZAKI** and Takeo KANADE***

(Received May 31, 1995)

We propose a system which extracts the plate area by using artificial neural networks(N.Ns) which detect the center position of the license plate. A coordinates position of the learning pattern is given as the random numbers. Therefore different patterns are learned by the N.Ns every time and N.Ns with higher generalization are built up. We investigated the optimum learning area of the plate included in the learning pattern, effects of specific learning in the bumper and the headlight, and effects of size and density conversions. From experimental results, we obtained higher detection rates of 98.5 %, 98.7 % and 100.0 % respectively (shift width 32 pixels).

Keywords : license plate, neural network, back propagation

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、及び行政機関等の公共施設においては、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、及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り出しと切り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。

本稿では、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使って行う

*電気工学専攻大学院生 (Graduate Student, Electrical Engineering)

**電子工学科 (Department of Electronic Engineering)

***カーネギーメロン大学ロボティクス研究所 (Robotics Institute, Carnegie Mellon University)

場合のネットワークモデル, 学習法, 及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, 及び位置検出部により構成される。

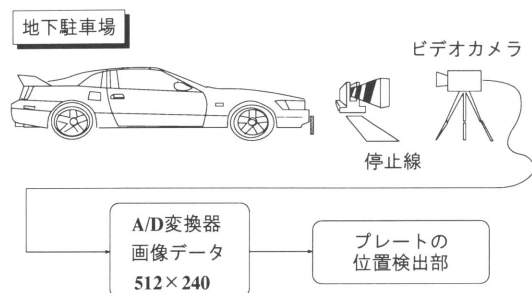


図1 システム構成

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]}。ユニットの閾値関数には, シグモイド関数 $f(x) = 1.0 / (1.0 + \exp(-x/u_0))$ を使用する。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, 以下の通りである。

- 入力層 : 512 (32×16)
- 中間層 : 100
- 出力層 : 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)。なお, 切り出しポイント(小領域の左上座標)の決定は乱数で行う。

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 $n\%$ 以上含まれているものである(図 4(a)参照)。また, 抑制パターン 1 は, プレート面積の $25 \sim n\%$ 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

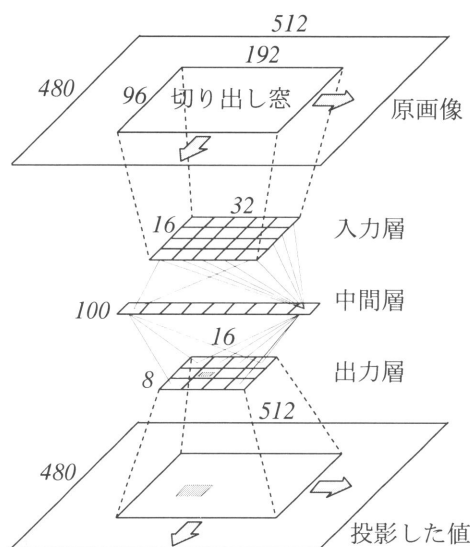


図2 ネットワークモデル

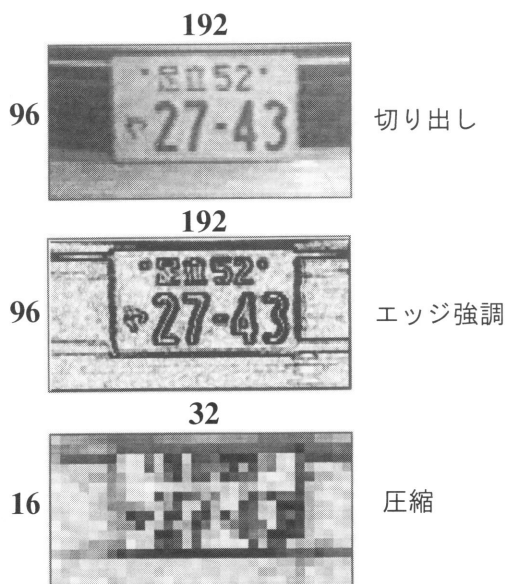


図3 入力パターン例

号は, 予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, 全てのユニットを 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。

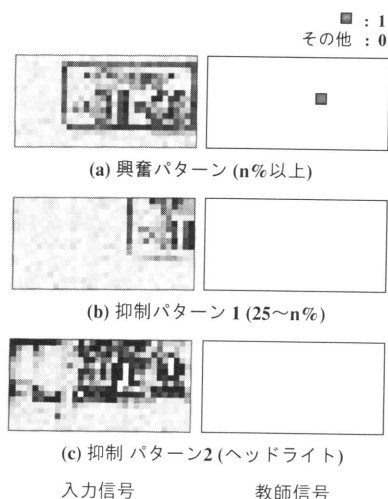


図4 学習パターン例

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

$$S_w = HW \quad (1)$$

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。

$$S_p = (x_2 - x_1 + 1)(y_2 - y_1 + 1) \quad (2)$$

切り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。

$$S_i = S_w \cap S_p \quad (3)$$

具体的には、 (R_{xi}, R_{yi}) を切り出し窓の左上の座標とすると、次式で計算できる。切り出し窓とプレートの位置関係によって、①～⑨の場合に分けられる。

$S_i =$

$$\begin{cases} \text{① } (x_2 - R_{xi} + 1)(y_2 - R_{yi} + 1) \\ \text{② } (R_{xi} + W - x_1 + 1)(y_2 - R_{yi} + 1) \\ \text{③ } (x_2 - R_{xi} + 1)(R_{yi} + H - y_1 + 1) \\ \text{④ } (R_{xi} + W - x_1 + 1)(R_{yi} + H - y_1 + 1) \\ \text{⑤ } (x_2 - R_{xi} + 1)(y_2 - y_1 + 1) \\ \text{⑥ } (R_{xi} + W - x_1 + 1)(y_2 - y_1 + 1) \\ \text{⑦ } (x_2 - x_1 + 1)(y_2 - R_{yi} + 1) \\ \text{⑧ } (x_2 - x_1 + 1)(R_{yi} + H - y_1 + 1) \\ \text{⑨ } (x_2 - x_1 + 1)(y_2 - y_1 + 1) \end{cases} \quad (4)$$

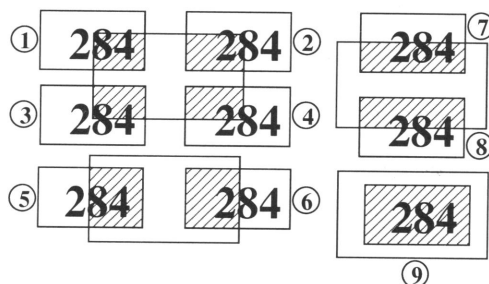
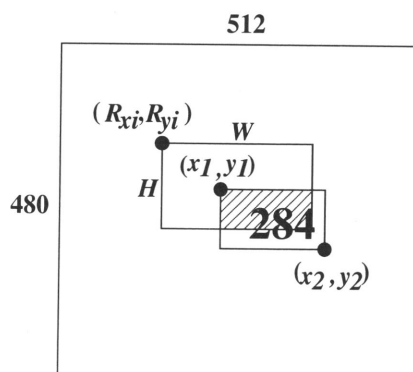


図5 面積の計算方法

故に、面積の割合 $n[\%]$ は次式のようにになる。

$$n = \frac{S_i}{S_p} \times 100 [\%] \quad (5)$$

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り出しパターンは、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。

$$E_n = \frac{1}{N} \sum_{i=1}^N (O_i - T_i)^2 \quad (6)$$

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を1.0にすると、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、学習パターンの切り出し位置が、乱数によって決定されるからである。毎回異なる

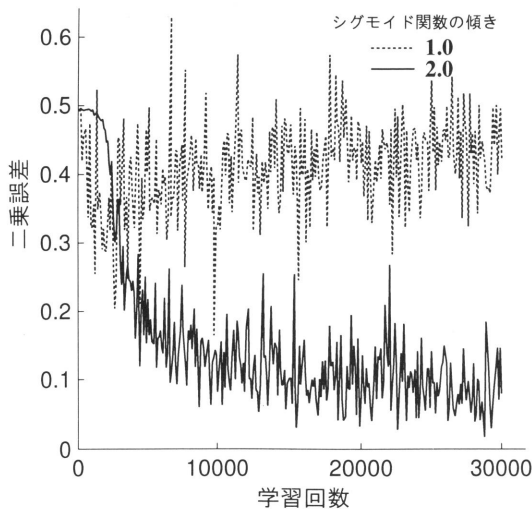


図6 学習誤差曲線

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良いことがわかる。

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、出力信号の精度が上がっていることがわかる。

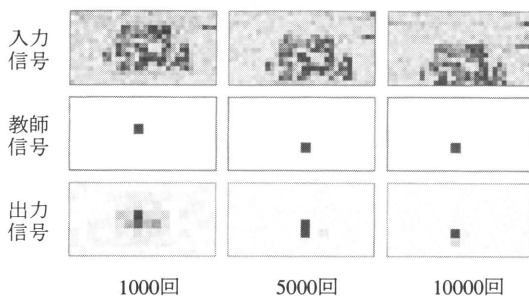


図7 重心位置の学習とそのときの検出能力

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。

$$W(x, y) = \exp\left(-\frac{x^2 + y^2}{2}\right) \quad (7)$$

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

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、切り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。
- (4) (1)～(3)を繰り返して、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。

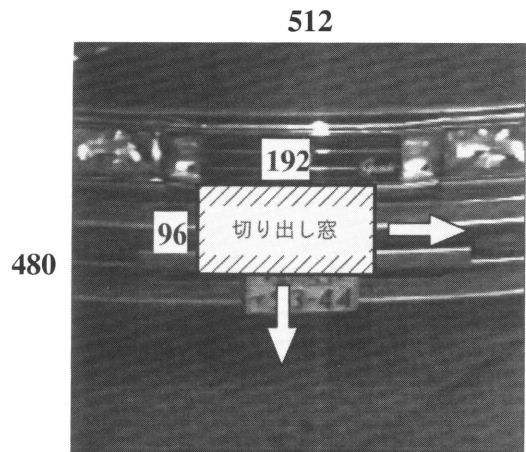


図8 入力方法

3.8 評価基準

予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。

$$d = \sqrt{(x_0 - x')^2 + (y_0 - y')^2} \quad (8)$$

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 $1:1$ に近くなるように、横 2 ライン毎に同じデータで埋め、 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 (図 9 (a) 参照)。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、可能な限り少量のデータで学習で

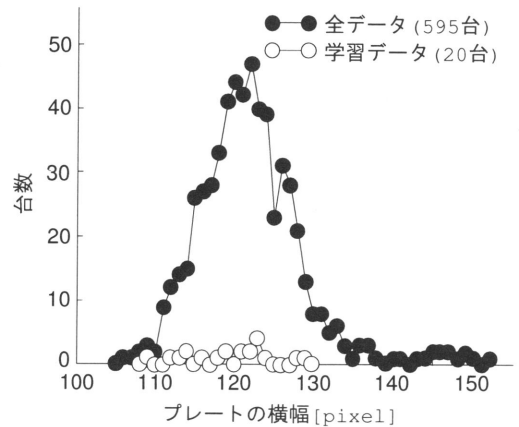


図10 プレートの横幅

きることが望ましい。本稿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原因の一つになっている。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。切り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り出しパターンを学習したときの投影加算例を示す。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



(a) 緑色地



(b) 白色地

図 9 入力画像例

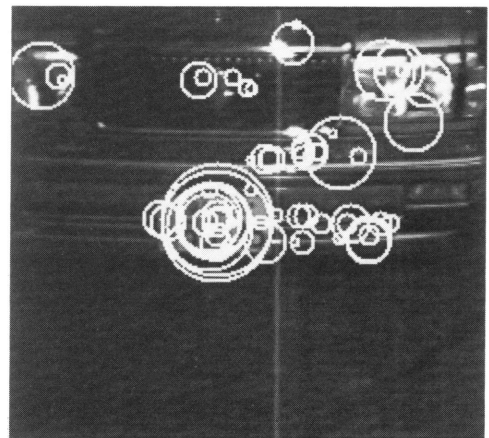


図11 投影加算例

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12に示す。プレート面積が60%以上含まれるとき、検出率は、97.5(580/595)%であった。以後、この値を最適学習面積とする。

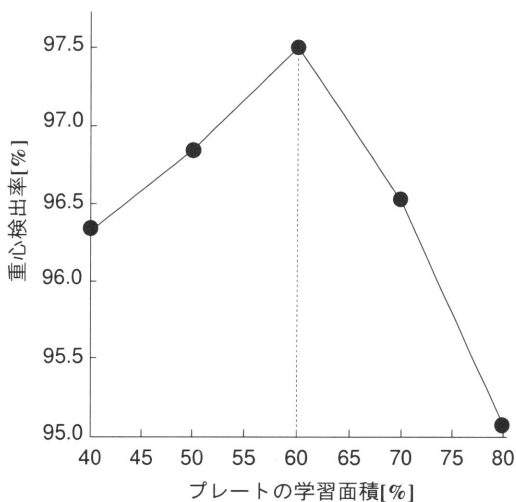


図12 最適学習面積

5.2 ガウス窓の効果

正方形ガウス窓を出力値に掛けて投影加算を行う場合、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、切り出し窓のシフト幅は、xy両方向ともに64[pixel]とする。図13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14、図15に、検出結果を図16に示す。ガウス窓なしの場合

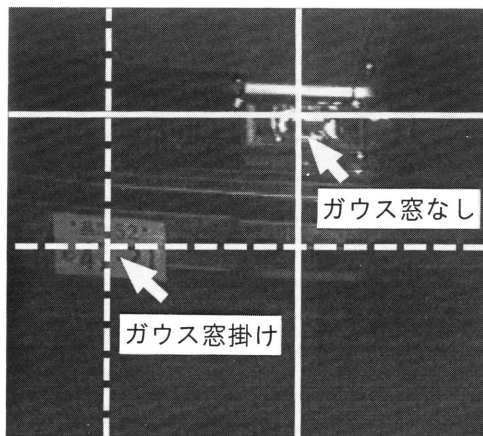


図13 ガウス窓の効果

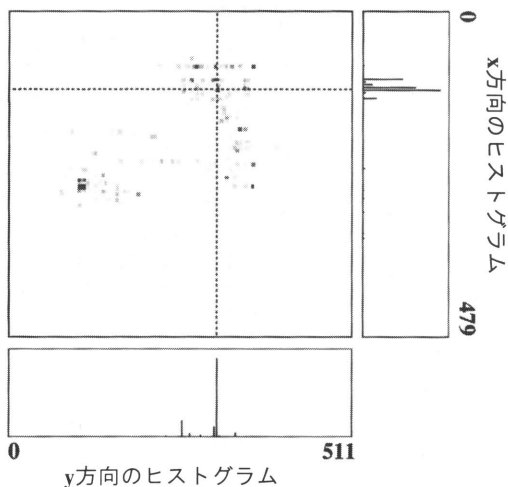


図14 ガウス窓なしの投影加算値

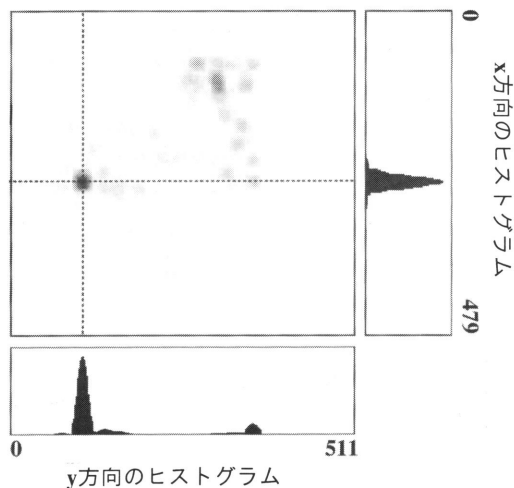


図15 ガウス窓掛け時の投影加算値

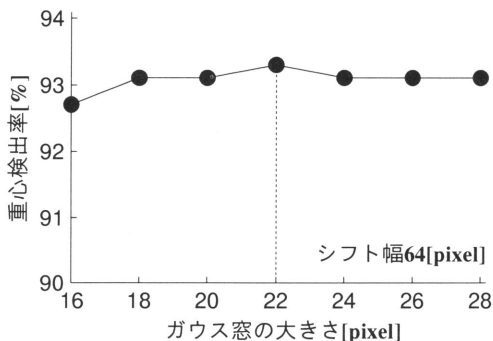


図16 ガウス窓の大きさと検出率

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画素のときが最も良い。以後の実験では、すべてこの値を用いる。

5.3 抑制学習

プレートの最適学習面積を求める実験において、プレート周辺部分とヘッドライト部分の検出誤りが多かった。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台分良くなっているが、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は、93.3(555/595)%から

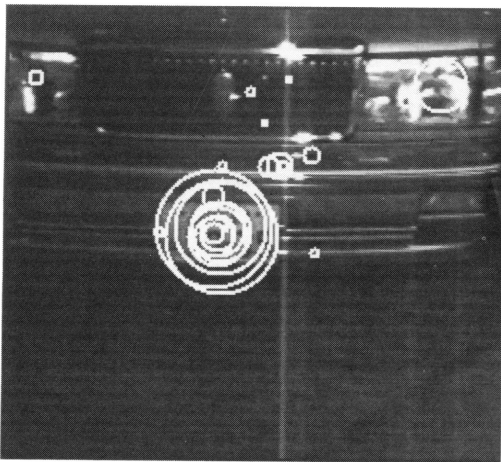


図17 抑制学習後の投影加算例

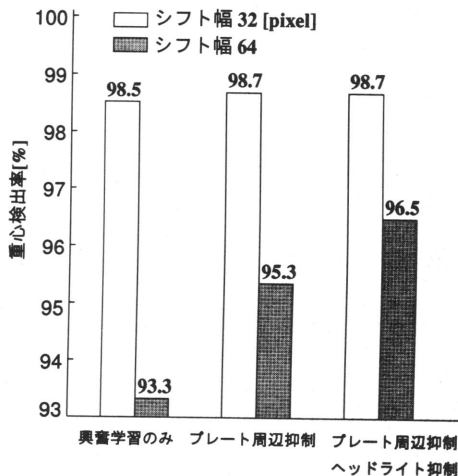


図18 抑制学習による効果

95.3(567/595)%に向上した。さらに、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17に抑制学習後の投影加算例を示す。

図11と比較すると、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、学習パターン数を増加したときの重心検出結果を図19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、プレート周辺形状が複雑なときに検出誤りがあ

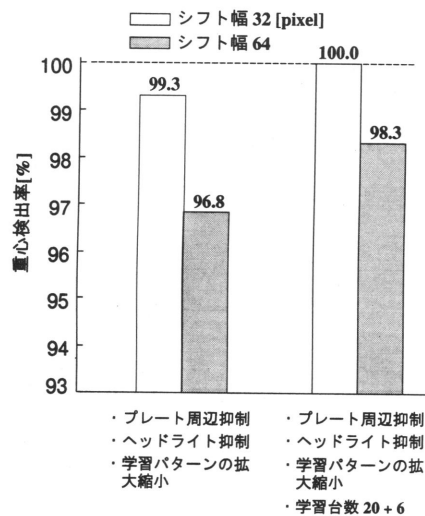


図19 学習パターンの増加による効果

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)%, 32[pixel]以下で、100.0(595/595)%の検出率を得た(図20参照)。

なお、1台当たりのプレート抽出に要する時間は、シフト幅32[pixel]のときHP9000/712で11.9秒、シフト幅64[pixel]のとき3.5秒であった。

6. むすび

本稿では、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、乱数によって決定されるた

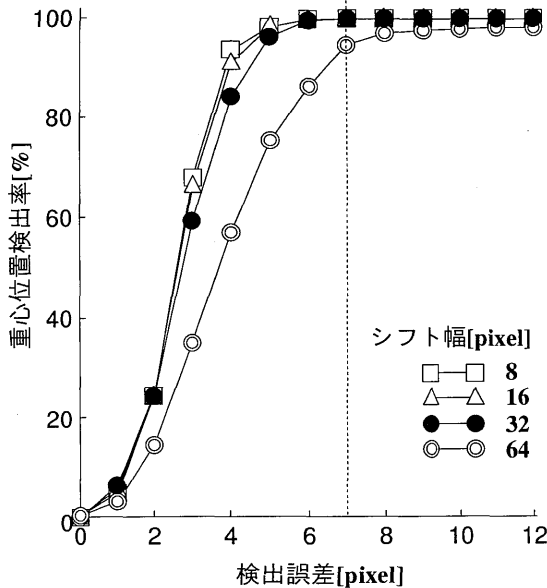


図20 シフト幅の影響

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無い。

まずはじめに、切り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、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、さらに1%向上した。一方、シフト幅64では87.1%から93.3%に上昇し、顕著な差が得られた。

次に、抑制学習についての検討を行った。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って、ニューラルネットワークの出力誤りが減少し、98.7%(シフト幅32)と96.5%(シフト幅64)の検出率を得ることができた。さらに、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%(シフト幅32)となった。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を上げることができた。

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%の検出率を得た。

本手法を用いることにより、従来、難しかったプレー

トと車両が同系列色の場合でも、高精度な位置検出が行われるようになった。

謝 辞

本稿で使用した画像データを提供して下さいったオムロン(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。

参 考 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), J72-D-II, 4, pp. 597-604(1989-04).
- [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, pp. 15-22(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, pp. 31-36(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, pp.17-22(1991).
- [9] 西村治彦, 小山宣樹: "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] Rumelhart D. E., McClelland J. L. and the PDP Research Group: "Parallel Distributed Processing", 1, The MIT Press(1986).