

N. N. を用いた生産情報読み取りシステムの開発

YU QIYUE[†] 三谷 大志[†] 服部 公央亮[†] 田口 亮[†] 梅崎 太造[†] 保黒 政大^{††} 本野 明大[‡]
 名古屋工業大学大学院産業戦略工学専攻[†] 中部大学工学部電子情報工学科^{††} 株式会社明和 e テック[‡]

1 はじめに

商品や工業製品などの生産情報を取得するために、QR コードの自動検出・識別が盛んに研究されている。一方で、欠損や撮影環境などにより、QR コードを読み取れず、生産情報が取得できない場合も多い。そこで、本研究ではニューラルネットワークを利用して QR コードの自動検出を行い、さらにコードが読み取れない場合は QR コード付近に存在する文字を認識することで、生産情報のロバストな自動読み取りを可能にするシステムを提案する。

2 概要

本研究では、ニューラルネットワークを利用して QR コードの自動検出・切り出しを行い、認識システムに入力する。QR コードが読み取れない場合は、QR コードの上に存在する文字領域を切り出し、文字認識ニューラルネットワークにより、生産情報を取得する。切り出す際に、傾き補正を施して画像中帳票の傾きに対応する。処理の流れを図 1 に示す。本研究で使用した帳票の撮影画像例を図 2 に示す。

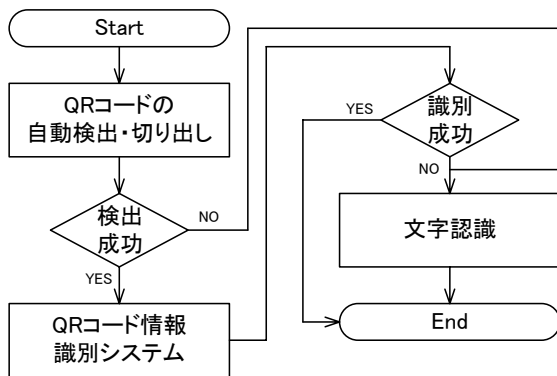


図1 処理の流れ

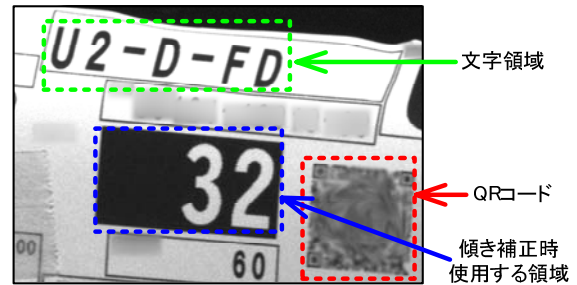


図2 帳票の撮影画像

3 QR コードの自動検出

生産情報を識別するためには、帳票上に存在する QR コードを検出して、撮影画像から切り出すことが必要である。本研究ではニューラルネットワークを利用した QR コード自動検出を行う。

ニューラルネットワークには入力層と出力層 30×30 ユニット、中間層 256 ユニットの三層型^[1]を使用する。学習する際には、ガウシアン平滑化した撮影画像から QR コード画像を切り出し、拡大、縮小、回転のランダム化を行う。その後 30×30 [pixel] に圧縮し、入力層に入力する。教師データはガウス分布で変化する濃淡画像とした。ニューラルネットワークの構成を図 3 に示す。

シグモイド係数 6.0, 9.0, 12.0 において、最大 1 万セットの学習を行い、学習回数毎に検出実験をした。QR コードの誤検出率を図 4 に示す。

図 4 より、シグモイド係数 9.0 で 1 万セット学習時に最も低い誤検出率 0.1 [%] が得られた。

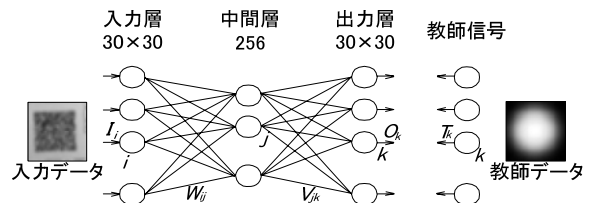


図3 ニューラルネットワークの構成

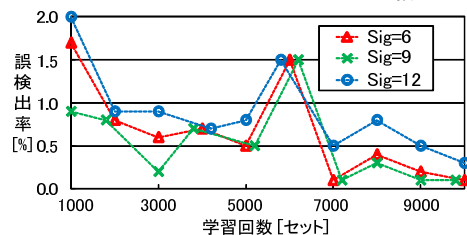


図4 誤検出率

The Development of An Automatic System for Reading Product Information

Qiyue Yu[†], Hiroshi Mitani[†], Koosuke Hattori[†], Ryo Taguchi[†], Taizo Umezaki[†], Masahiro Hoguro^{††}, Akihiro Motono[‡]

[†]Nagoya Institute of Technology

^{††}Chubu University

[‡]MEIWA e-TEC co.,Ltd

4 傾き補正

ニューラルネットワークで撮影画像中の QR コードを検出して、QR コード自動認識システムで識別を行う。しかし、QR コードの欠損や撮影環境の明るさなどにより、QR コードが読み取れない場合も多い。そこで、QR コード認識失敗時には文字で書かれた生産情報を読み取る。

撮影画像より、文字領域において文字の傾きが存在することが確認できた。文字の誤認識率を減らすために、画像の傾き補正が必要である。本研究では、まず輪郭追跡により QR コード左側に存在する黒い四角形の領域の頂点座標を求めた。次に、求めた頂点座標を利用して射影変換行列を求め、全体画像を射影変換することで傾き補正をした。傾き補正処理の流れを図 5 に示す。

頂点の座標を求める際に、まず、ラベリングした画像を輪郭追跡して、図 6 に示す輪郭の方向コードを求めた。次に、方向コードを用いて四角形の 4 辺を取得した上で交点の座標を四角の頂点座標として取得した。求めた 4 辺と頂点を図 7 に示す。

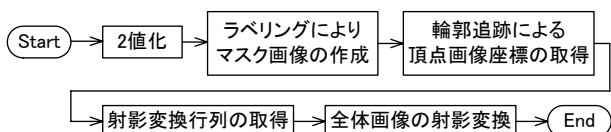


図 5 傾き補正処理の流れ

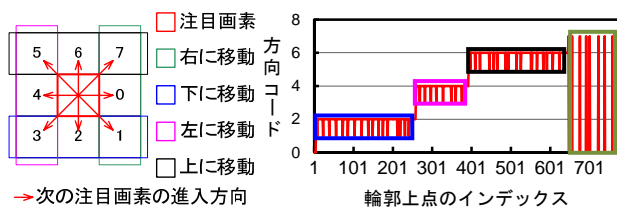
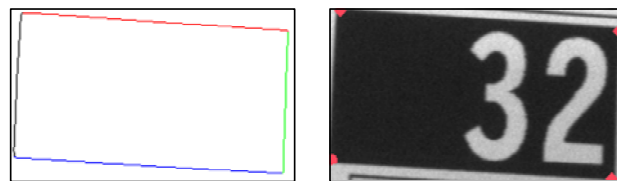


図 6 輪郭追跡により取得した方向コード



(a) 求めた 4 辺 (b) 求めた頂点

図 7 頂点座標の取得

5 文字認識

全体画像の傾きを補正した後、文字領域を QR コード位置から固定値で切り出した。切り出した文字領域画像を図 8 に示す。

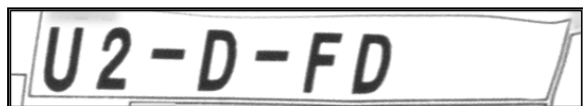


図 8 切り出した文字領域画像

文字以外の領域が文字認識に悪影響を与えないように、本研究では、式(1)と式(2)で表わす投影分散ヒストグラムを利用することで、文字をセグメンテーションした。

$$S_x(y) = \sum_{j=0}^{H-1} I(j,y) \quad V_x(y) = \frac{1}{H} \sum_{j=0}^{H-1} \left(I(j,y) - \frac{S_x(j)}{W} \right)^2 \quad (1)$$

$$S_y(x) = \sum_{i=0}^{W-1} I(x,i) \quad V_y(x) = \frac{1}{W} \sum_{i=0}^{W-1} \left(I(x,i) - \frac{S_y(i)}{H} \right)^2 \quad (2)$$

ここで、 $S_y(x)$ と $S_x(y)$ はそれぞれ縦 y 、横 x の位置における 1 ラインの輝度の合計であり、 $V_y(x)$ と $V_x(y)$ はその分散値である。ただし、 W 、 H は画像の幅と高さ、 $I(x,y)$ は座標 (x,y) の輝度とする。文字セグメンテーションの結果の一例を図 9 に示す。

文字認識には三層型のニューラルネットワークを使用した。文字セグメンテーションにより切り出した文字「-」, 数字 1-9, 英字大文字 A-Z, ノイズの計 37 種類に対して、学習(2 万 7 千セット)と認識^[2]をした。ニューラルネットワークの中間層を 128 ユニット、入力層を 45×60 ユニット、出力層を 37 ユニットとし、シグモイド係数を 5.0, 8.0, 11.0 とした。文字認識の平均正解率を図 10 に示す。シグモイド係数 11.0 で 2 万 7 千セット学習時に最も高い平均正解率 99.1 [%] が得られた。



図 9 文字セグメンテーションの結果

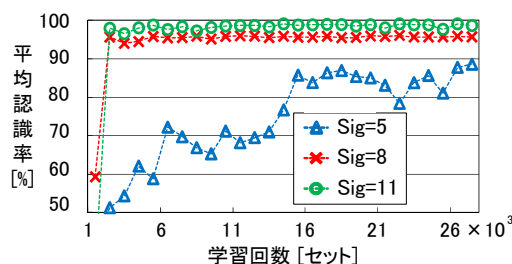


図 10 文字認識の平均正解率

6 まとめ

本研究ではニューラルネットワークで画像中の QR コードを検出し、QR コード認識失敗時に文字認識を行うことで生産情報のロバストな読み取りシステムを開発した。QR コードの検出率 99.9 [%], 文字認識率 99.1 [%] が得られた。

参考文献

- [1] 殿村正延, 中山謙二: 多層パーセプトロンにおける内部情報最適化アルゴリズムの提案, 信学論, 2001
- [2] 奥田忠義, 道木慎二, 石田宗秋: $\Delta \Sigma$ 変調に基づくパルスニューラルネットワークによる BP 学習, 信学論, 2005.