

天気図中の線図形群識別のためのアルゴリズム

後藤恒久 Le Phuoc Minh 吉田雄二 福村晃夫
(名古屋大学 工学部)

§1. ま え が き

我々は、主に線図形により構成される画像データを解析して種々の情報を自動的に抽出し、構造化して効率的に格納するシステムを構成する研究の一環として、天気図データを対象とした画像処理システムの構成を行なっている。^[1]

一般に、天気図中には、等圧線、地形線、経線・緯線、前線などの線図形群、および各種記号（天気、前線、風力など）が含まれている。本講では、はじめに我々が開発している天気図自動処理システムにおける線図形群の処理過程について述べる。次に、線図形群の中で経線・緯線、および等圧線をとりあげ、これらを識別、抽出する手順を実験結果と共に示す。

§2. 天気図自動処理の概要

2.1. 天気図データ

ここで対象とするデータは、朝日新聞朝刊に掲載されている天気図の中央部およそ $3.6\text{cm} \times 3.6\text{cm}$ の部分を F.S.S. を用いて 500×500 で標本化したものである（図1）。各成分の処理に際しては、これを文献[2]の方法で2値化したものを扱う。標本化された天気図中に存在する図形は、図中における位置、および形状により表1のように分類される。以下では、図中の位置が一定なクラス1, 2を統合して不変成分、図中の位置が変化するクラス3, 4を統合して可変成分と呼ぶ。

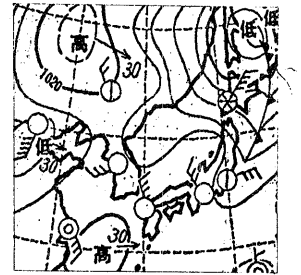


図1. 天気図データの例

2.2. 天気図処理の流れ

2.1で述べた各クラスに属するパターン of の処理は図2の流れにそって行なわれる。これらの処理のうちで、気圧記号^[3]、天気記号^[4]、前線記号^[5]についてはすでにその処理アルゴリズムを開発済である。また、風力記号、数字などについては§4で述べる等圧線の抽出の過程で雑音として除去される。これらを個別に分類・識別するアルゴリズムは現在開発中である。

表1. 天気図中のパターン

形状 \ 位置	不変	可変
不変	(クラス1) 経線・緯線 地形線	(クラス3) 気圧記号 数字
可変	(クラス2) 天気記号 風力記号	(クラス4) 等圧線 前線

§3. 経線・緯線の抽出

経線・緯線の情報、天気図データの標本化の際に生ずる位置ずれや回転の検出、表示範囲の異なる天気図の自動識別などに用いることができる。標本化された経線・緯線は、長さ10~20画素、線幅2~4画素の小規模塊状図形がほぼ直線

的に分布したものとなる。これらは多くの場合孤立しているが、地形線、等圧線などと交差することがある。その場合には上述の形状についての特徴が変化する。

ここで述べる方法は、孤立している場合および経線と緯線とが交差している場合に適用可能である。この方法では、はじめに経線、緯線を構成すると考えられる塊状図形（これを候補図形と呼ぶ）が抽出される。次に、これらの図形を線分を用いて表現し、それらを経線または緯線として連結する。最後に、この結果に基づいて直線近似および円弧近似を行なうことにより、経線および緯線の全体の形状がそれぞれ抽出される。以下各手順について詳しく述べる。

3.1. 候補図形の抽出

《候補図形を細線化して抽出するアルゴリズム》

入力：2値化された天気図データW

出力：細線化された候補図形の集合C

手順：

step1. W中の連結成分（塊状図形）のうちで、その面積Sが次式を満たすものをすべて抽出する。

$$S_{MIN} \leq S \leq S_{MAX} \dots \textcircled{1}$$

step2. 抽出された各図形を細線化⁶⁾し、長さ2以下のヒゲを取り除く。

step3. 線図形を構成するセグメントの長さの総和Lが、 $L_{MIN} \leq L \leq L_{MAX} \dots \textcircled{2}$ を満たすものをすべて求め、その集合をCとする。

3.2. 候補図形の線分表現

《候補図形を線分表現するアルゴリズム》

入力：細線化された候補図形の集合C

出力：線分表現された候補図形の集合S

手順：各候補図形に対して次の手順を適用する。なお以下では、各図形の端点、分岐点をそれぞれ T_i, B_j のように表わす。

step1. 与えられた候補図形の端点を T_1, \dots, T_n とする。端点 T_i, T_j 間の距離を $l_{ij} (1 \leq i < j \leq n)$ 、 l_{ij} のうち最大のものを l_{ij}^* とする。

step2. (a) $L/l_{ij}^* < p_1$ のとき、すなわち図形全体がほぼ直線的であるとき、
i) $L \leq p_2$ ならば、 T_i, T_j を両端とする線分でその図形を線分表現する。

ii) $L > p_2$ ならば、その図形を候補から除外する。

ここで、 p_2 は孤立している経線・緯線の長さの性質に基づくパラメータである。

(b) $L/l_{ij}^* \geq p_1$ のとき、step3を行なう。

step3. 図形中の分岐点の数に従って以下の手順を適用する。いずれの場合にも、指定される条件が満たされない時は図形を候補から除外する。

(a) 分岐点をもたない場合——2つの端点 T_1, T_2 間の距離を l_{12} とし、図

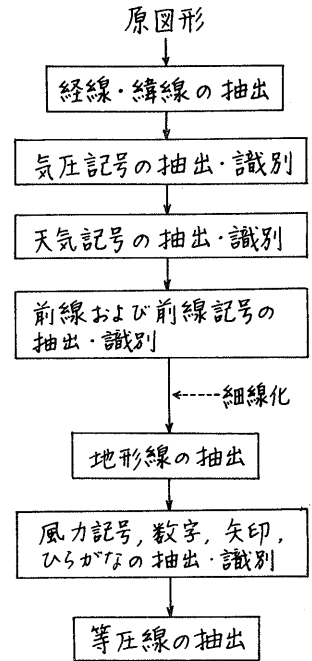
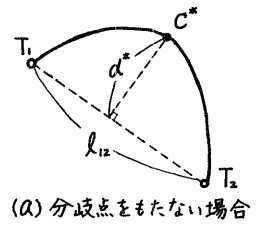
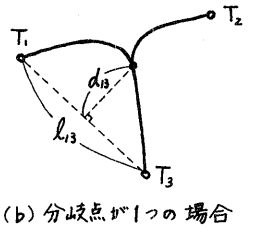


図2. 天気図処理の流れ

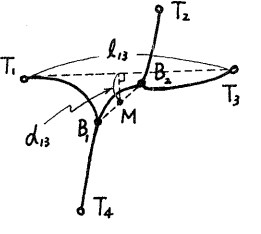
形中で線分 $\overline{T_1 T_2}$ との距離が最大となる点を C^* , そのときの距離を d^* で表わす (図3-a)。このとき, $d^*/l_{12} > p_3$ ----- ③ が満たされるならば, T_1 と C^* , および T_2 と C^* を両端とする2つの線分でその図形を表現する。



(b) 分岐点が1つの場合——分岐点を B_1 , 端点を T_1, T_2, T_3 とする。2つの端点 T_i, T_j 間の距離を l_{ij} , 分岐点 B_1 と線分 $\overline{T_i T_j}$ との距離を d_{ij} とする (図3-b)。 d_{ij}/l_{ij} の最小値 $d_{ij}^*/l_{ij}^* = r_{ij}^*$ が, $r_{ij}^* < p_4$ ----- ④ を満たすならば, T_i^* と T_j^* , およびもう一つの端点と B_1 を両端とする2つの線分でその図形を表現する。



(c) 分岐点が2つの場合——距離が4以下の2つの分岐点を B_1, B_2 とし, 両者の中点を M とする。また4つの端点を T_1, \dots, T_4 とする。 l_{ij} を (b) の場合と同様に, また d_{ij} を M と線分 $\overline{T_i T_j}$ との距離と定義する (図3-c)。このとき,



$d_{13}/l_{13} < p_4$, かつ $d_{24}/l_{24} < p_4$ ----- ⑤ が満たされるならば, T_1 と T_3 , および T_2 と T_4 を両端とする2つの線分でその図形を表現する。

3.3. 経線・緯線の結合

ここでは, 線分表現された候補図形 (数字や短い等圧線も含まれる) の中から, 経線, 緯線を構成するものを選び出して結合する。

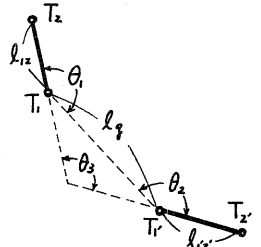
(1) 候補図形の分類 3.1 および 3.2 で抽出された線分をそれが水平方向となす角 θ により次の3つに分類する。

- ① 経線 ($1.2 \text{ rad} < \theta \leq 1.9 \text{ rad}$)
- ② 緯線 ($0 \text{ rad} \leq \theta \leq 0.4 \text{ rad}, 2.7 \text{ rad} < \theta \leq \pi \text{ rad}$)
- ③ その他 (上記以外)

ここで, θ の範囲は経験的に与えた。

(2) 結合可能性の判定 図4のように, 2つの線分 $\overline{T_1 T_2}, \overline{T_1' T_2'}$ があるとき,

図3. 複数個の線分で表現する場合の特徴量



(条件1) $l_g \leq c \cdot l_{12}$, または $l_g \leq c \cdot l_{12}'$ ----- ⑥

(条件2) $\pi - d\theta \leq \theta_1 \leq \pi + d\theta$
 かつ $\pi - d\theta \leq \theta_2 \leq \pi + d\theta$
 かつ $\pi - d\theta \leq \theta_3 \leq \pi + d\theta$ ----- ⑦

図4. 結合可能性の判定に用いる特徴量

が同時に満たされるならば, 線分 $\overline{T_1 T_2}$ と $\overline{T_1' T_2'}$ とは結合可能であるとし, 次に述べる操作を行なう。

(3) 線分の置き換え 結合可能な2つの線分 $\overline{T_1 T_2}, \overline{T_1' T_2'}$ が与えられて図5に示すように線分 $\overline{T_2 T_2'}$ と T_1, T_1' との距離をそれぞれ L_1, L_1' とする。このとき,

$\max(L_1, L_1') \leq d$ ----- ⑧

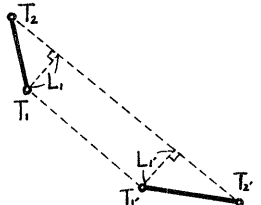


図5. 線分の置き換え

が満たされるならば、2つの線分を1つの線分 T_1T_2 で置き換える。⑧式が満たされないときは3つの線分 $\overline{T_1T_2}$, $\overline{T_1T_3}$, $\overline{T_2T_3}$ で表わす。従って、結合された線分は1本または3本の線分により表わされることになる。後者の場合、(2)の結合可能性の判定において、条件1では両端点(T_1 , T_2)間の直線距離を、また条件2ではそれぞれの端点側にある線分の角度を用いる。

(4) 結合アルゴリズム (1)で①と②に分類された線分の集合の各々について以下を行なう。

step1. $c = 1.0$ とする。

step2. 同一線分上にない2つの端点の組合せに対して、 l_j の小さいものから順に以下を行なう。

(a) (2)の条件により結合可能性の判定を行なう。結合可能と判定された場合は(b)へ。

(b) (3)で述べた方法により線分の置き換えを行なう。

step3. $c = 2.0$ として step2 を繰り返す。

step4. ②に分類された線分の集合については終了。①に分類された線分の集合については、 $c = 3.0$, $c = 4.0$ として step2 を繰り返す。

3.4. 経線・緯線の近似

3.3で述べた方法により結合が行なわれた線分のうち、①に分類されたものは直線で、②に分類されたものは円弧で近似する。近似の方法は最小二乗誤差近似とし、標本点としては(3)の方法によって置き換えられた各線分の端点を用いる。

3.5. 実験結果

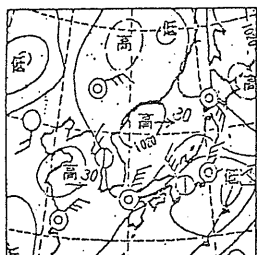
以上に述べた方法を実際の天気図データに適用した結果を図6に示す。なお、各パラメータの値は次のように設定した。

$$S_{MIN} = 15, S_{MAX} = 150, L_{MIN} = 7, L_{MAX} = 40$$

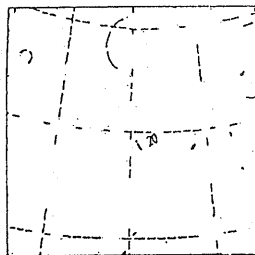
$$P_1 = 1.3, P_2 = 25, P_3 = 0.4, P_4 = 0.2, d = 5$$

$$d\varphi = \pi/8, \quad d\theta = \begin{cases} \pi/4 & (0 < l_j \leq \bar{l}/2 \text{ のとき}) \\ \pi/6 & (\bar{l}/2 < l_j \leq \bar{l} \text{ のとき}) \\ \pi/8 & (\bar{l} < l_j \text{ のとき}) \end{cases}$$

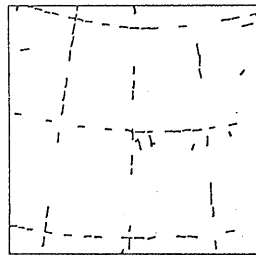
ここに \bar{l} は線分の平均長で、 $\bar{l} = 14$ である。計算時間は、3.1, 3.2の候補図形の抽出とその線分表現の処理において7.96秒^{*}, 3.3の経線・緯線の結合の処理において25.9秒^{**}, 3.4の近似の処理において0.15秒^{**}であった。



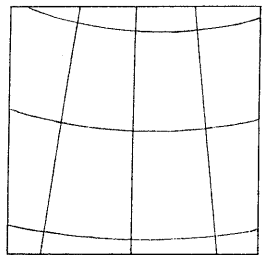
(a) 原図形



(b) 3.1のstep1の処理結果



(c) 候補図形の線分表現



(d) 経線・緯線の近似

図6. 経線・緯線の抽出実験の結果の一例

* FACOM M-200 使用

** FACOM 230-38 使用 (演算速度は M-200 の約 1/40)

§4. 等圧線の抽出

ここでは、文献[3],[4]の方法によって気圧記号と天気記号を除去した天気図データ(2値)を細線化⁽⁶⁾した図形から、地形線、数字、風力記号などの雑音を除去して等圧線を抽出する方法について述べる。

線図形は、あらかじめ端点または分岐点を結ぶ分岐のない、セグメントと呼ばれる線図形に分解され、セグメントごとに種々の情報が集められる。以下の処理はこの情報にもとづいて行なわれる。情報としては次の項目があげられる。

(a) 始点と終点の座標

(b) 8方向指数系列とその長さ

(c) 不変成分テンプレート中に含まれている部分の長さ

(d) 始点および終点から出る枝の数、およびその枝へのポイント

ここで不変成分テンプレートとは、28枚の天気図を重ねて得られた図形を閾値処理⁽⁶⁾によって2値化して得られる図形である(図7)。

得られたセグメント集合から等圧線以外のセグメントを除去することにより等圧線のセグメントを抽出する。この処理の流れを図8に示す。

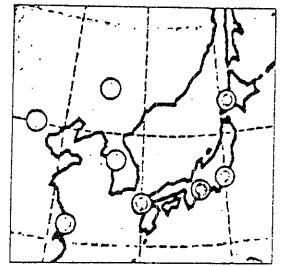


図7. 不変成分テンプレート

4.1. ノイズの除去

セグメントの集合Sの中には、細線化によって生じたヒゲや、孤立した短いセグメントとして現われる背景部分の雑音が含まれる。

ここでは、長さ10以下のヒゲ、および長さ10以下の孤立したセグメントを雑音として除去する。さらに、その結果分岐点の不変成分テンプレートに含まれる割合のほぼ等しい2本のセグメントの連結点になった場合は、2つのセグメントを統合して1本のセグメントとする。こうして得られたセグメントの集合を S_1 とする。

4.2. 不変成分の除去

次には、地形線に代表される不変成分と考えられる線図形を除去する。

図9は、2本の等圧線と1本の地形線が交差している部分を細線化⁽⁶⁾して得られる図形の状況を示している。不変成分テンプレートを重ねた時に、不変成分領域に含まれている割合の大きいセグメントを線図形としての不変成分とすれば、図中の S_T, S_r, S_L, S_C, S_e がこれに該当する。しかし、これらを一律に除去すると等圧線の連続性が保存さ

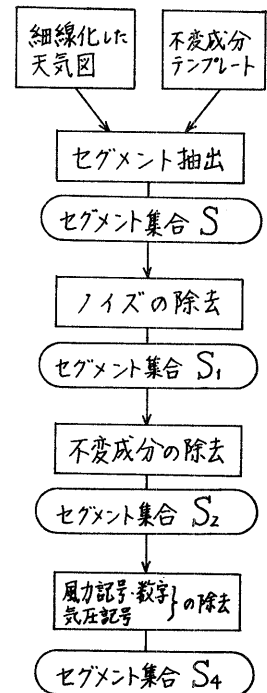


図8. 等圧線を抽出する処理の流れ

れない。すなわち、 S_L および S_R は除去しない方が望ましい。

以上のような考え方に従って不変成分の除去を行なう。具体的なアルゴリズムを与える前に、図9の S_L , S_R のようなセグメントを接続セグメントと呼び、これを次のように定義する。

〔定義1〕 不変成分除去前のセグメントの集合を S_1 、不変成分に含まれる割合の小さい（以下では経験的に0.6と定めた）セグメントの集合を S_2 とする。セグメント s ($s \in S_1$) が次のすべての条件を満たすならば s を接続セグメントと呼ぶ。

- (a) $s \in S_1 - S_2$
- (b) 長さ ≤ 100
- (c) s はその両端点で S_2 の要素の1つと連結している。それらを s_1, s_2 とする。
- (d) $(s \text{ の長さ}) \leq (s_1 \text{ の長さ}) + (s_2 \text{ の長さ})$
- (e) $(s \text{ の長さ}) > 20$ の場合は、 s が1つの直線 l で近似される。さらに、 s の端点の近傍で s_1, s_2 を直線 l_1, l_2 で近似したときに、それらの直線の間次式が成り立つ。

$$\begin{cases} (l \text{ と } l_1 \text{ の角度}) \geq \frac{5}{8}\pi \\ (l \text{ と } l_2 \text{ の角度}) \geq \frac{5}{8}\pi \end{cases}$$

なお、曲線図形を直線近似する方法については文献[7]の方法を用いた。不変成分を除去するアルゴリズムは次のようである。

《不変成分除去アルゴリズム》

入力：4.1の方法によりノイズを除去したセグメント集合 S_1

出力：不変成分除去後のセグメントの集合 S_2

手順：

- step1. $S_2 = \phi$
- step2. $s \in S_1$ として、 $(s \text{ が不変成分テンプレートに含まれる割合}) < 0.6$ を満たすセグメント s を S_1 から取り除き S_2 の要素とする。
- step3. S_1 と S_2 から接続セグメントの集合 S_3 を求める。
- step4. S_3 が空になるまで以下を繰り返す。
 - (a) 最も短い接続セグメント s' を S_3 から取り除き S_2 の要素とする。
 - (b) (a) の操作により接続セグメントの条件を満たさなくなったセグメントを S_3 から除外する。
- step5. S_2 を出力する。

4.3. 風力記号, 数字, 気圧記号の除去

不変成分を除去した後に残る成分は、等圧線の他に風力記号, 数字, 矢印, ひらがな, 前線, および文献[3]の方法で除去できなかった気圧記号である。これらのうちで、数字および気圧記号はそれ自身の占める面積が小さく、かつ各部分を直線近似⁽⁶⁾すると多くの直線成分に分解されるという性質を有する。また、風力記号はそれを構成するセグメントの幾つかが平行であるという性質を有する。これらの成分は以上の性質を利用して除去される。

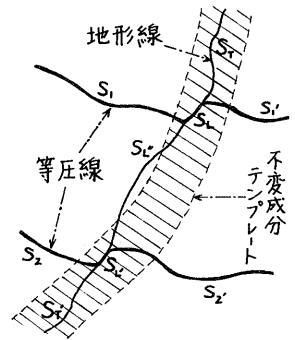


図9. 不変成分除去の例

これらを除くアルゴリズムを与える前に、必要な定義を与える。

[定義2] 1つの連結線図形を構成しているセグメントの組を“segment group”と呼ぶ。

[定義3] segment group G を包含する最小長方形の大きさが $(x \times y)$ のとき、 $(x \times y)$ を G のサイズと呼ぶ。

[定義4] segment group G のすべてのセグメント s_i について、 s_i 上に端点をもつ線分によって s_i を直線近似する操作をセグメント分解と呼び、得られた直線成分のそれぞれを直線セグメントと呼ぶ。

《風力記号、数字、気圧記号を除去するアルゴリズム》

入力：不変成分を除去したセグメント集合 S_2

出力：風力記号、数字、気圧記号を除去したセグメント集合 S_4

手順：

step1. 次の条件を満たす segment group G を S_2 から求める。なお、ここでは G に含まれるすべてのセグメント s について l を s の長さ、 d を s の端点間の距離とする。

(a) s が孤立しているセグメントの場合 $30 \leq l \leq 120$

(b) s の両端点が他のセグメントと連結している場合 $d \leq 40$

(c) s の一方の端点が分岐点の場合 $d \leq 50$

この条件を満たす segment group がなければ終了（そのときのセグメント集合 S_2 が出力 S_4 となる）。あれば step2 へ。

step2. G をセグメント分解する。

(a) G のサイズが (34×34) 以下のとき、

i) 直線セグメントの数が2以下であり、かつそれらの間の角度が $\pi/2$ 以下ならば G を除去する。

ii) 直線セグメントの数が2以上ならば G を除去する。

(b) G のサイズが (34×34) 以上のとき、 G の中に平行な2本の直線セグメントがあれば G を除去する。

step3. step1 へ。

ここで、2本の直線セグメントが平行とは、それらのなす角度が 20° 以内のときのことをいう。

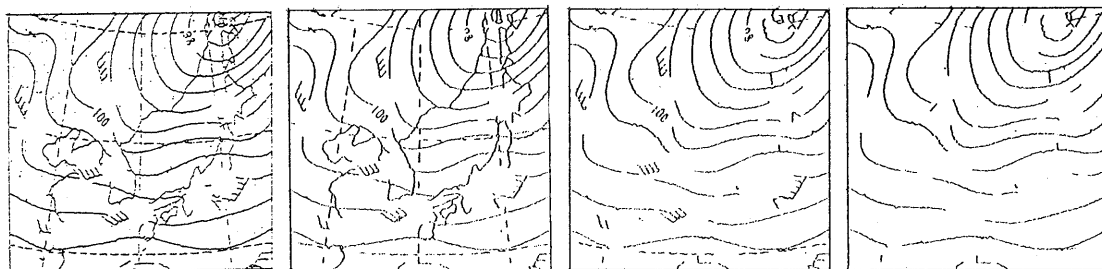
4.4. 実験結果

以上に述べた一連の手続を実際の天気図データに適用した結果を図10に示す。不変成分の除去については満足の結果が得られたが、風力記号、数字や気圧記号の除去では、それ自身が等圧線などに連結しているために除去されないで残る場合があることが知られる。また、なめらかな数字 '1' は等圧線のセグメントと区別できないために除去されなかった。各処理での計算時間、出力されたセグメント数、およびセグメントの長さの総和の27枚の天気図データに

表2. 実験結果の各処理毎の平均値

	出力されたセグメント数	セグメントの長さの合計	計算時間*
セグメント抽出	939	9647	15分 34秒
ノイズ除去	291	8200	41
不変成分除去	161	5797	20
風力記号、数字などの除去	110	4680	1 21

* FACOM 230-38 使用



(a) セグメント抽出

(b) ノイズの除去

(c) 不変成分の除去

(d) 風力記号・数字
気圧記号の除去

図10. 等圧線の抽出実験の結果の一例

対する平均値を表2に示す。(ただし、セグメント抽出には細線化処理も含まれる。)

§5. あとがき

天気図の経線・緯線、および等圧線の抽出については一応満足のゆく処理手順が得られた。経線・緯線については、結合、近似された結果ならびに全体的な構造に着目して等圧線など他の線図形と交わる部分を抽出すること、および3.3の(1)の手順で用いた日の範囲を自動的に設定することなどが今後の課題となる。一方、等圧線の抽出については、等圧線などと連結した風力記号や数字の除去、および切断された等圧線を接続して等圧線図として復元することが今後の課題である。

謝辞 日頃熱心に御指導頂く本学の本多波雄教授、並びに本多・福村研究室の諸氏に感謝します。

参考文献

- (1) 吉田, 福村: 気象図自動処理システムの構成, コンピュータ・ビジョン研資1-1 (昭54.5)
- (2) 大津: 濃度分布からの閾値決定法, 信学情報部門全大, 145 (昭52.8)
- (3) 林, 吉田, 福村: 高速テンプレート・マッチング法と気象図への応用, 信学技報, PRL78-90 (昭54.3)
- (4) 後藤, 吉田, 福村: 天気図中の天気記号の認識, 信学情報部門全大, 122 (昭54.10)
- (5) 伊藤, 吉田, 福村: 天気図中から前線を抽出する一方法, 昭55情報全大発表予定
- (6) 横井, 鳥脇, 福村: 標本化された2値図形のトポロジカルな性質について, 信学会論文誌 Vol.56-D No.11 (昭48.11)
- (7) 岡崎, 吉田, 福村: 曲線図形の自動記述と再生およびその等高線への応用, 信学会パターン認識と学習研資 (昭54.10)