

天気図を対象とした図形認識実験

谷口, 倫一郎
九州大学大学院総合理工学研究科情報工学専攻

遠藤, 勉
九州大学大学院総合理工学研究科情報システム学専攻

田町, 常夫
九州大学大学院総合理工学研究科情報システム学専攻

<https://doi.org/10.15017/17492>

出版情報 : 九州大学大学院総合理工学報告. 1 (1), pp.73-80, 1979-10-09. 九州大学大学院総合理工学
研究科
バージョン :
権利関係 :



天気図を対象とした図形認識実験

谷口倫一郎*・遠藤 勉**・田町常夫**

(昭和54年7月26日 受理)

Some Experiments on Machine Recognition of Weather Chart

Rin-ichiro TANIGUCHI, Tsutomu ENDO
and Tuneo TAMATI

It is one of the most important aspects of artificial intelligence researches to construct a system which has an ability to understand pictures. Generally a picture contains various kinds of information, so it is necessary to select more meaningful information among them. One of good ways is to extract information representable in natural language. The system which we present as an example aims to describe weather charts in natural language. This system consists of three major processes: 1. Extraction of component patterns, 2. Recognition of component patterns, 3. Syntax analysis (Recognition of relations between component patterns).

Cognitive process in human beings is considered to use a priori knowledge about the objects. So the system needs to have such knowledge represented in some manner.

First we present how a priori knowledge about weather charts aids to recognize them and how it is organized in computer program. Second the fundamental processings such as noise elimination, thinning, template matching and curve following, are presented. Third the effectiveness of these processings are shown through some experimental results. Finally the problems for further research are pointed out.

1. 緒 言

人工知能研究における大きな課題の一つに図形の持つ意味を機械に理解させるといった問題がある。我々は従来より図形を計算機に入力してその意味内容を自然語で記述させるという研究を行ってきたが¹⁾、現在その応用として、天気図の機械認識の研究を進めている²⁾。一般に図形は非常に多くの情報量を持っている。このような図形を自然語で記述する場合、意味のある情報を取り出すことが重要である。本研究では対象を天気図としているが、図形を自然語で記述するにはどのような過程が必要であるかを研究することが主な目的である。

天気図は気象情報の一表現法である。他の表現法として自然語による気象通報文がある。これらは同一のものを表現しているので、天気図を自然語で記述する

には、気象通報文を構成するに必要な情報を天気図から抽出すればよいことになる。またその記述の正当性を、自然語出力と入力天気図を比較して判定できる。このように天気図は、言語と図形の意味の対応構造を研究するという点で適当な対象といえる。本稿では、現在開発中の天気図認識システムの概要及び実験について述べ、ここで用いた処理の有効性を示す。また今後の問題点についても述べる。

2. 知識の導入

「天気図を理解する」とは大略次の三過程が行われることであると考えられる。

1) 要素図形の抽出

天気図を構成する基本的要素を抽出する。

2) 局所的認識

要素図形を認識し、その持つ情報を獲得する。

3) 大局的認識

要素間の関係を把握し、天気図を記述するのに必要

* 情報工学専攻修士課程

** 情報システム学専攻

な情報を獲得する。

人間が物を見て理解する場合、見ている対象に関する知識に基づき、目に入る世界をある程度予測していると考えられる。従って、天気図を理解するシステムにも何らかの形でその知識をもたせる必要がある。本節では、天気図を機械が認識するのに必要とする知識について、及びそれらの知識をアルゴリズムの上でどのように実現させるかについて述べる。

2.1 天気図の構成

本システムの対象としている天気図の一例を Fig. 1 に示すが、これは日本気象協会発行の天気図集成に準拠したものである。このシステムで考える要素図形 (Component) は 11 種類で、Table 1 にそれらの持つ意味的情報を示す。気象情報の中で大きな意味を持つ前線に関しては、天気図上では一つの要素図形としてではなく、前線記号 (Front symbol) の組み合わせとして表現される (Table 2 参照)。

2.2 天気図認識に用いるアプリオリな知識

天気図に関する知識としてあらかじめシステムに与えておくようにしたものは、以下の (a) — (i) である。

(a) 天気記号 (Weather symbol) は同一の大きさの円で囲まれ (以後この円を外周円と呼ぶ)、その内側のみが互いに異なり、この部分だけに天気の情報を持つ。またその種類は 14 種で、拡大、縮小、回転等の変形は受けていない。

(b) 天気記号の出現位置は 15 地点に限られる。天気図によっては、その地点に天気記号が存在しないこともあるが、限定された地点以外に出現することはない。ただし、限定されるといっても画像データ上では入力によって数画素の位置ずれが生ずることがある。

(c) 天気記号に連結する要素は、風記号 (Wind symbol) のみである。天気記号に連結するものがない場合には、「風力 0」を示す。

(d) 風記号は、天気記号の外周円から外側に延びる直線 (風向軸と呼ぶ) の向きで風向を示し、風向軸から分岐する枝 (風力枝と呼ぶ) の数で風力を示す。ただし風力 0 は (c) で述べたとおりである。

(e) 前線記号は、天気記号を除くと唯一の面の部分を持った要素である。残りの図形はすべて線画である。

(f) 地形線 (Coast line)、緯経線 (Latitude,

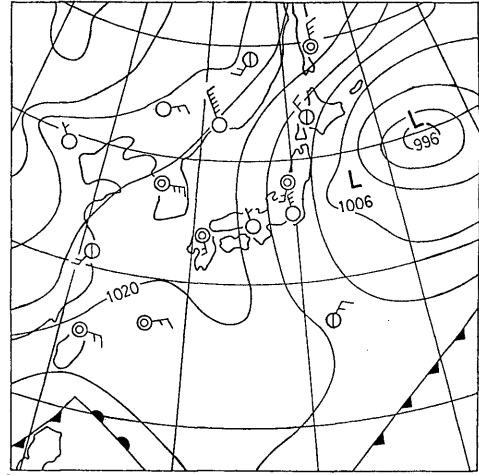


Fig. 1 Example of weather chart

Table 1 Picture components in weather chart

Component	Pictorial pattern	Information from the component
Weather symbol	☉ ●	Location, variety of weather
Wind symbol	☼ ☼☼ ☼☼☼	
High pressure	H	Location, central pressure
Low pressure	L	
Tropical depression	T, D	
Typhoon	T	Location, variety of shape, direction
Front symbol	—▲▲	
Isobar	—	Locus, pressure value
Coast line	—	No information about weather
Latitude, longitude	—	
Numeral	01...9	Pressure value

Table 2 Classification of front lines

Front line	Pictorial pattern	Relation between front symbols
Warm front	—▲▲	▲ is not adjacent to other symbols.
Cold front	—▲▲	▲ is not adjacent to other symbols.
Stationary front	—▲▲	▲ and ▲ are adjacent and have opposite directions.
Occluded front	—▲▲	▲ and ▲ are adjacent and have same directions.

longitude) の形、位置、大きさはほぼ一定である。ただし、他の要素によって一部が消去されていることがある。

(g) 数字 (Numeral) の出現位置は等圧線 (Isobar)、文字 (Low pressure 等) の近傍で、その数値はそれぞれの気圧値を示す。

(h) 文字の種類は 4 種。(H, L, T, D)

(i) 文字、数字の大きさは一定である。

認識システムでこれらの知識を利用する際は、それぞれ次のようにする。

(a), (b) → 〇 大きさの決まった円を探索し、天気記号の位置を確定する。

○その際の探索範囲は、期待される出現位置の近傍に限定する。

○天気記号の認識にはテンプレートマッチングを用いる。

(c) →○天気記号の外側境界を追跡して、天気記号及び風記号を抽出する。

(d) →○風向・風力は後述の細線化後、線分追跡を行い認識する。風力は「風力1」の場合を除き風記号の端点の数に一致する(5.3参照)。

(e) →○天気図の中で面部分を探索し、前線記号の位置を確定する。

○前線記号は総て、二種類のパターン(「三角型」, 「半丸型」)のいずれかに、回転の変形を加えたものである。回転を考慮してテンプレートマッチングを行い認識する。(前線としての認識はTable 2に示した前線記号間の関係を認識する必要がある。別に行う。)

(f) →○面要素をすべて消去し、以後の処理を簡単化するため細線化を行う。

○辞書を参照しながら線分追跡を行い、地形線、緯経線を除去する。辞書は、地形線、緯経線を適当な間隔でサンプリングしたもので、これらを目標として線分追跡を行う。

以上の(a)–(f)については、既にそれぞれの処理プログラムが現システムで稼動している。(g)–(i)部分については現在開発中である(3.参照)。

3. 天気図認識システムの構成

本システムにおける天気図認識プログラムは、概略Fig. 2の流れ図に示す通りである。前処理としての雑音除去操作後、2.2に従って認識を行う。Fig. 3の中で、破線はその処理の結果得られた情報が他の処理に使われることを示す。各処理過程については4.及び5.で詳述する。なお(10)の構造解析(Syntax analysis)には次の処理が含まれる。

- 1) Table 2に示した前線記号間の関係を認識し、前線を識別する。
- 2) 複数桁の数字を一つの気圧値として認識する。
- 3) 2)の結果をその近傍に存在する等圧線あるいは低気圧等の持つ気圧値と関連づける。
- 4) 他の要素によって消されている等圧線を補う。
- 5) 気圧値が付記されていない等圧線の気圧値を推定する。

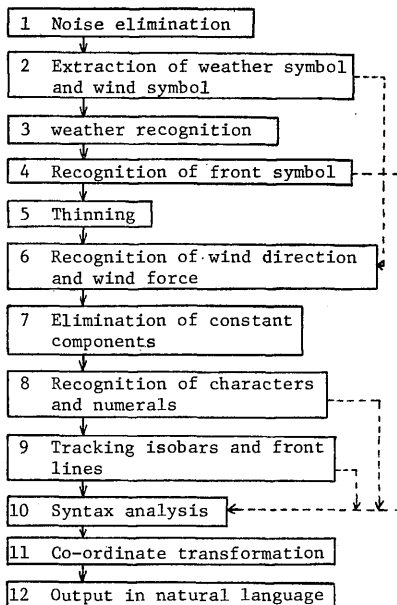


Fig. 2 Flow chart of weather chart recognition system

6) 要素図形間の関係を記述する。(「低気圧の中心から南西に寒冷前線が延びている。」等)

4. 基本アルゴリズム

本節では、このシステムで用いられる基本的図形処理手法である1. 雑音除去, 2. 細線化, 3. テンプレートマッチング, 4. 線分追跡について述べる。

4.1 雑音除去

PDS から入力された画像データを二値化する際には、量子化雑音が発生しやすい。この雑音は黒画素の孤立点、突起点あるいは白画素の孤立点、突起点となる。一般にこのような雑音を含む入力パターンに対する完全な図形処理アルゴリズムは極めて難しく、現実的アルゴリズムでは特に細線化処理に悪影響を及ぼす。例えば Fig. 3 に示すように意味のない分岐点や閉路を生ずる。そこで3×3画素の簡単なマスクを用いて黒白両画素の孤立点と突起点を除去するようにした。



Fig. 3 Effect of noise

4.2 細線化

線画を PDS により入力すると、実際の画像データ上では数画素の線幅を持つ。そこでこのような線画の処理を簡単化するために細線化を行う。細線化のアルゴリズムとしては Hilditch⁹⁾ のものを用いた。

4.3 テンプレートマッチング

入力される図形パターンが高度に規格化されている場合、つまり拡大、縮小、回転の変形が加わらず、どの入力パターンも標準パターン（テンプレート）とほとんど差異がない場合、標準パターンの中から最もよく入力パターンに一致するものを選び、それを入力パターンであると識別する方法がテンプレートマッチングである。「一致」の程度（類似度）を評価する方法はいくつかあるが、ここではテンプレートの黒白分布と入力パターンの黒白分布の一致した画素の割合を用いている。一般に入力パターンの位置がはっきりと決まっていない場合があるので、縦横数画素の範囲ですべての点で類似度を求め、その範囲で最大のものをそのテンプレートとの類似度とする。すなわち、

$$\alpha(x, y) = 1 - \frac{\sum_{i=1}^N |p_i(x, y) - q_i|}{N}$$

$$\alpha_m = \text{Max}_{\substack{u-s \leq x \leq u+s \\ v-s \leq y \leq v+s}} \alpha(x, y)$$

- q_i テンプレートの i 番目の画素値 (0 又は 1)
- $p_i(x, y)$ パターンの基準座標を画面上の (x, y) においた時に、テンプレートの i 番目画素に対応する入力パターンの画素値
- N テンプレートの画素数
- (u, v) 期待される入力パターンの位置
- s ずれの補正範囲

で、 α_m が求めるテンプレートとの類似度である。 α_m を与える (x, y) が補正されたパターンの位置である。なお、 $|p_i(x, y) - q_i|$ は「排他的論理和」に相当する評価値である。特にテンプレートと入力パターンが完全一致の場合は $\alpha_m = 1$ となる。実際の処理では α_m がある閾値を越えた場合、テンプレートと一致したと判定する。複数のテンプレートから一つを決定する場合は類似度を最大にするものとする。

4.4 線分追跡

ここで述べる線分追跡は細線化された線図形を対象としたものである。したがって通常の基本的操作は 8

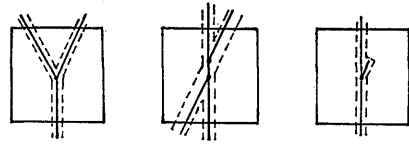


Fig. 4 Examples of branching points

近傍の画素を走査して次に進む点を決定し、その点を次の走査の中心とするという操作の繰り返しである。線分を追跡する過程で特別な処理を必要とするのは、端点に遭遇した場合、及び分岐点に遭遇した場合である。まず端点に遭遇した場合はそこで線分追跡を中止する。新たな追跡はシステムにより指示される。分岐点に遭遇した場合は分岐点の種類を判別する必要がある。これは、細線化処理の影響により交差点が二つの分岐点に変形したり、あるいは短い枝（雑音枝）が出現するためである (Fig. 4 参照)。種類を判定するには、分岐点に遭遇した時に走査範囲を拡大して分岐の数を調べて、分岐点、交差点、雑音枝のいずれであるかを決定する。以後の処理は、雑音枝の場合は雑音枝を無視して追跡を続行する。分岐点、交差点の場合は追跡を続行する分岐先を決定して追跡を行う。追跡先の決定に際しては、交差点の場合なめらかに連なる方向を選択することが望ましいので、交差点の向かい合う分岐先を優先する。また追跡しない分岐先は、システムに保持しておき、次の追跡時の情報として用いる。

5. 実験

本節では現在稼働中の天気図認識システムにおける 1. 天気記号の抽出と認識、2. 前線記号の認識、3. 風向・風力の認識、および 4. 地形線、緯経線の除去について、そのアルゴリズムと具体的な処理状況を示す。

5.1 天気記号の抽出と認識

(1) 天気記号の位置の確定

2.2 で述べたように、天気記号は同一大の円で囲まれている。しかも、多少の位置ずれは生ずるがほぼ出現位置は決まっている。天気記号の位置を確定するために 4.3 で述べたテンプレートマッチングを用いて類似度を計算し、 α_m を与える (x, y) を天気記号の位置とする。このときの位置ずれの補正範囲は縦横とも ± 10 画素 ($s=10$) とする。ただし α_m がある閾値を越えない場合は、この近傍に天気記号はないものと判

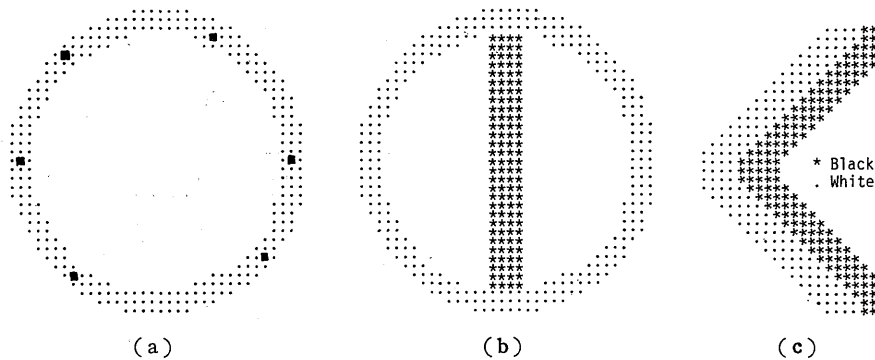


Fig. 5 Examples of templates

断する。

α_m を求めるには $21 \times 21 = 441$ 点で α を求める必要があるが、実際に閾値を越える α は少なく無駄な計算をしていることになる。そこで計算量を減らすために二段テンプレートマッチングを採用した。すなわち、まず特徴的な少数画素のパターンを調べ、 α が閾値を越える可能性があるか推定する。可能性のあるものについてのみ α を計算し、それらの中で最大を α_m と決定することにした。Fig. 5(a) にここで用いたテンプレートを示す。外周円のテンプレートのうちまず 6 点の画素を調べ、4 画素以上黒の場合のみ α を求めるという簡単な方法で効率的に α_m を求めることができた。

(2) 天気記号と風記号の抽出

天気記号の外周円には風記号のみが連結するので(2.2(c)), (1) の後に外周円の外側境界を追跡し、その内領域を取り出すことによって天気記号及び風記号が抽出できる。

(3) 天気記号の認識

天気記号は規格化されたパターンであるので、テンプレートマッチングの手法を用いて識別する。2.2(a) で述べたように天気記号の外周円には天気に関する情報はなく、内側のパターンの形状だけが天気を表しているの、内側部分のみに関するテンプレートマッチングにより記号を同定することにした。Fig. 5(b) は「晴」を示すテンプレートであり、そのうち * の部分だけが実際の記号の同定のために用いられる。このような内部パターンだけのマッチングを評価することにより、外周円を含めた場合に比べよりの確な同定ができた。天気記号は認識後消去する。また風

記号については残った部分と併せて細線化を行う。

5.2 前線記号の認識

(1) 探索範囲の限定

2.2(e) に述べたように前線記号は天気記号以外では唯一の面部分を持った要素図形である。5.1 の処理で天気記号は総て抽出消去されているのでこの段階で画面上に面部分が残っていればそれは前線記号である。前線記号は天気記号と異なり出現位置は不定である。したがってその探索は画面の全範囲にわたり計算量も非常に多くなる。そこで精密な探索操作の前に全画面を前線の存在が可能な領域(候補領域)とそうでない領域に分割しておくことが望ましい。画面から前線記号を含む小画面を切り出せば、その部分は他の部分に比べ黒画素の割合が大である。そこで全画面をいくつかの小領域に分割し、各領域での黒画素の割合を求め、それがある閾値を越える領域を候補領域とすることが考えられる。しかし、割合の高くなる領域として他に等圧線密度の高い領域(台風の近傍等)がある。このような領域との混同を避けるため我々のシステムでは、 4×4 画素の黒小領域の割合を求めることにした。実際には全画面を 128×128 画素からなる 64 個の小領域に分割している。Fig. 10(a) の場合の分布を Fig. 6 に示す。

(2) 前線記号の認識

まず面部分を抽出し、前線記号の位置を決定する。面部分の探索は 8×8 の黒画素のみからなる正方形が 1 つ以上接続して存在する部分を抽出することによって行う。ここでも二段テンプレートマッチングを適用して計算時間の短縮を図っている。

次に前線記号の同定を行う。前線記号は「半円形」

0	1	3	1	15	4	2	3
2	0	1	4	11	5	2	5
5	5	1	5	18	6	14	5
0	3	3	4	15	8	2	2
2	1	4	6	0	0	2	2
7	8	0	0	0	0	5	22
3	33	0	0	1	0	16	2
27	1	23	2	2	1	13	2

Fig. 6 Histogram of 4×4 black regions

と「三角型」の二種類であるが、回転という変形が加わっている。ここでは 11.25° ごとに回転したテンプレート 32 個を用意して、回転に対処したマッチングを行っている。一般に前線記号の向きは近傍にある前線記号と同一もしくは逆であることが多い (Table 2 参照)。従って、あらかじめ近傍に存在する面部分の位置関係で前線記号の向きを推定し、マッチングを行うテンプレートの向きを制限することができる。類似度がある閾値を越えなかった場合はすべての向きのテンプレートとマッチングをとる。

前線記号を識別するには、その境界の形状が判明すれば十分であるので、マッチングは境界部分だけで行っている。Fig. 5(c) にそのためのテンプレートの一例を示す。前線記号の認識終了後面部分を消去する。

5.3 風向、風力の認識

5.1 の処理で抽出され、さらに細線化された風記号について天気記号の外周円に連結した風向軸の方向を求めれば風向が決まる。ただし図法の関係で地図上の方位 (東西南北の方向) は画面上の x-y 座標軸の方向と異なっているので、この段階では x-y 座標上での方向を求めるまでとする。実際の出力に用いられる「東西南北」は Fig. 2 (11) の座標変換処理後に決定される。

風力は風向軸から分岐する枝の数で示される。したがって基本的には枝の端点数を求めればよい。この場合問題となるのは端点数が共に 2 である「風力 1」と「風力 2」の判別である (Fig. 7 参照)。この区別のためには Fig. 7 に示す a, b を用い、適当な閾値 T_1, T_2 により

$$b/a < T_1 \rightarrow \text{風力 1}$$

$$b/a > T_2 \rightarrow \text{風力 2}$$

と決定することにした。端点数を求めるには細線化さ

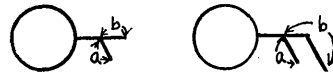


Fig. 7 Wind force 1 and 2

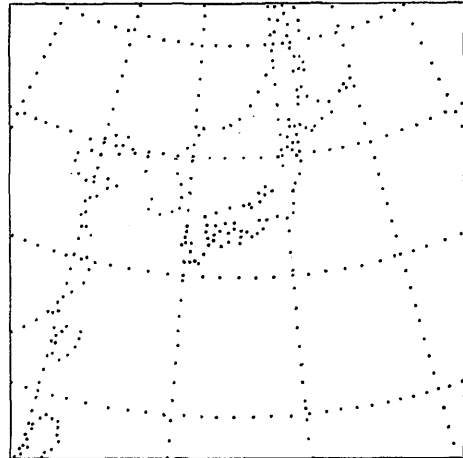


Fig. 8 Representative points

れた風記号の各画素でその 8 近傍の画素値 (0, 又は 1) の和が 1 となる点の数を求めればよい。但し、外周円に接した点は除く。

5.4 地形線、緯経線の除去

我々が天気図中の地形線、緯経線を他の線画から識別できるのは、明らかに我々は事前に地形線、緯経線の形を知っているからである。機械に対しても同様な識別能力を期待するならば、何らかの方法で事前に知識を与えざるを得ない。ここでは一つの試みとして、地形線、緯経線を 400 個程度の代表点で表現し、そのデータを機械の知識とした (Fig. 8 参照)。細線化された天気図から地形線、緯経線を除去する基本的な方法は最も近い二つの代表点を天気図の線分上に割当てその二点を結ぶ路を追跡除去するという操作の繰り返しである。ここでこの二点のうち追跡を開始する点を出発点、追跡の目標となる点を目標点と呼ぶことにする。これらは代表点の辞書の中に記述されている。

この手法を用いる上で特別な処理を必要とするのは、代表点が天気図上に発見できなかった場合、及び追跡途中で端点に遭遇した場合である。前者の場合、出発点が発見できなかった時は、それに最も近い点を仮の出発点として処理を行い、また目標点が発見できなかった場合は二点間の路の除去は放棄する。後者の

Weather information
 ((Tokyo KAISEI (3 62)) (Osaka KAISEI (1 60)) (Fukuoka KUMORI (2 30)) (Niigata KUMORI (2 35))
 (Sapporo HARE (2 57)) (Naha KUMORI (2 14)) (Chichijima HARE (2 7)) (Taipei KUMORI (2 19))
 (Shanghai HARE (2 50)) (Peking KAISEI (1 59)) (Seoul KUMORI (3 17)) (Choshun KAISEI (2 13))
 (Khabarovsk HARE (2 40)) (Vladivostok KAISEI (6 59)))
 Front information
 (((T (72 102) 27) (T (144 152) 27) (C (228 159) 4) (C (281 108) 4)
 ((T (790 118) 29) (T (842 191) 28) (T (897 261) 28) (T (949 327) 28)))

Fig. 9 Information extracted from weather chart (Fig. 1)

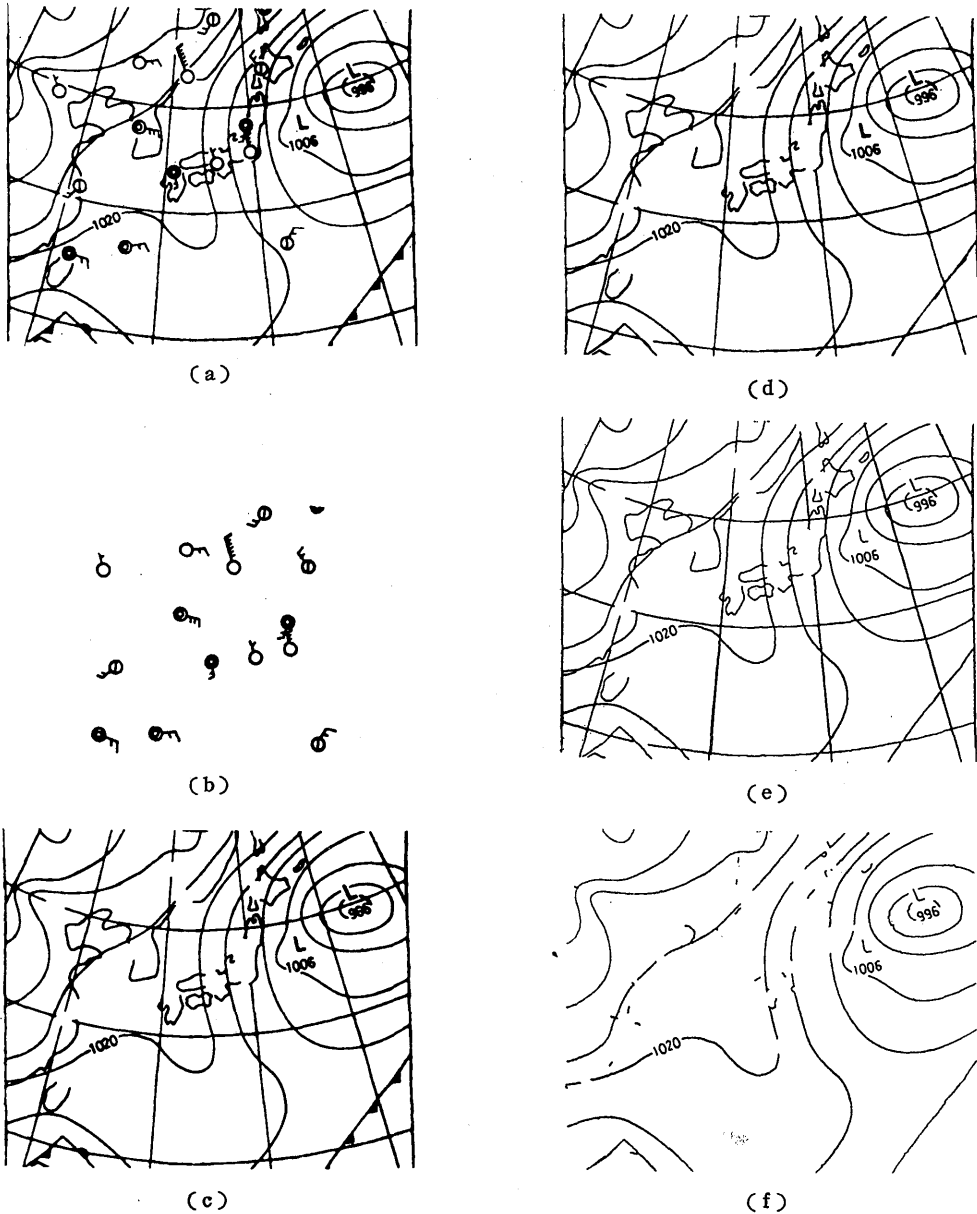


Fig. 10 Experimental result

場合は出発点から端点までを除去する。

この手法を用いることにより、入力図形の線幅の変化や位置ずれに対してもかなり対応できると考えられる。

5.5 実験結果

今までに述べた処理により天気図から読み取られた情報は、Fig. 9 に示す通りとなる。また処理の経過を Fig. 10 に示す。Fig. 9 における天気情報 (Weather information) は、(地名 天気 (風力 風向)) から成っている (風向は $x-y$ 座標系における傾きを 0—63 で表している)。また、前線情報は (前線記号 {T: 三角型, C: 半丸型} 位置 (x 座標, y 座標) 向き {0—31}) から成っている。Fig. 10 は (a) が雑音除去後、(b) が抽出された天気記号、風記号、(c) が残りの部分、(d) が前線記号の認識消去後、(e) が細線化後、(f) が地形線、緯経線の除去後の画面をそれぞれ示している。(f) においていくつかの小片が残っているが、これは代表点が発見できなかつたか、あるいは端点に遭遇して追跡を中断したことによるものである。また等圧線が一部消去されているが、これは等圧線と地形線が重なって一本となり、それを地形線と判断して消去してしまったためである。

以上簡単に実験結果について述べたが、天気図を自然語で記述するには要素図形からどのような意味を取り出す必要があるのかを検討し、それらに適応した処理アルゴリズムをそれぞれ用いることによって、要素図形の認識が的確に行われることが示された。これによって、要素図形間の関係の認識が適切に行われれば、計算機を用いて天気図の自然語による記述が可能であるという見通しが得られた。

6. 結 言

本稿では、天気図から意味情報を取り出すための一つのアプローチを提案し、それに基づいて実験を行った。Fig. 2 に示すシステムの地形線、緯経線の除去までは一応満足のできる結果が得られている。天気図のように扱うデータ量が多い場合は、いかに無駄な処理を省くかが問題となろう。

今後改良すべき点としては、地形線、緯経線の除去の能力を向上させることがある。また残る過程についても、処理方式の検討、実験を行っている。特に構造解析における等圧線の補間、気圧値の推定等には解決すべき問題が含まれている。

謝辞 日頃有意義な御助言を頂く本学情報認識研究室の皆様には感謝の意を表します。

参 考 文 献

- 1) 岡田, 田町: “図形の意味解釈とその自然語記述”, 信学論(D), J59-D, (昭51-05).
- 2) 谷口, 遠藤, 田町: “天気図を対象とした図形認識システム”, 昭54信学会全国大会, No. 1354.
- 3) Kawaguchi, E., Yokota, M., Endo, T. and Tamati, T.: “An Understanding System of Natural Language and Pictorial Pattern”, 6th IJCAI (1979).
- 4) Hilditch, J.: “Linear skeletons from square cupboards”, Machine Intelligence 4 (1969).
- 5) Rosenfeld, A., Kak, A. C.: “Digital Picture Processing”, Academic Press (1976).