

天気図を対象とした類似図形検索について

谷口, 倫一郎
九州大学大学院総合理工学研究科情報システム学専攻

澄川, 道俊
日本アイ・ビー・エム株式会社 | 九州大学大学院総合理工学研究科情報システム学専攻

河口, 英二
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/10.15017/17634>

出版情報 : 九州大学大学院総合理工学報告. 7 (1), pp.115-123, 1985-08-01. 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類似図形検索について

谷口倫一郎*・澄川道俊**・河口英二*
(昭和60年3月30日 受理)

Similarity Retrieval of Weather Charts

Rin-ichiro TANIGUCHI*, Michitoshi SUMIKAWA** and Eiji KAWAGUCHI*

Major problems of PDBS (Pictorial DataBase System) lie not only in its storage structure but also in its retrieving mechanism to search out a required picture easily. Most of PDBS's which have been developed are based on the frameworks of DBMS (DataBase Management System) which deals with symbolical data. However, such frameworks cannot be applied to some kinds of associative retrieval, for example, retrieval of similar pictures to the input one.

In this paper, we present the retrieval of weather charts based on the semantic similarity of the pictures. We define the similarity between two weather charts as probability that they are similar. On the basis of the semantic structure, the similarity can be inferred from lower-leveled similarities of constituents of weather charts. In order to calculate the above probability inference, we adopt Dempster & Shafer's model of probability. This paper fully shows the methodology of semantic similarity retrieval of weather charts.

1. 緒 言

図形情報の検索を行なうシステムは通常、図形データベースシステムと呼ばれているが、近年研究が盛んになり、各種のものが開発されてきている。図形検索では、図形情報の集合から希望する画像情報を如何に効率良く取り出すかということが基本的な問題であるが、これには大きな二つの問題、つまり、大量の図形情報の効率的蓄積技術の問題とどのように図形を検索するかという検索方式の問題が含まれている。この二つの問題は必ずしも独立に議論できる訳ではなく、またどのような図形情報を検索するかによってもかなり解決の方法が異なってくる。

しかし、従来の図形データベースシステムの検索方式に関しては、既に確立した記号情報に関するデータベース理論の枠組みを利用するものがほとんどであった。つまり、図形情報の内容(主体や属性)を記号化してデータベース化し、それを従来のDBMS(DataBase Management System)で用いられる方法で蓄積・検索するという方式であった。従って、この方式では図形を検索するために、検索する図形情報の内容を記

号化して、更にそれを検索言語に置換するという過程を人間が行なう必要があった。しかし、図形情報は情報量、換言すれば図形のバリエーションが多く、その置換過程は図形検索を行なおうとする者にとってかなり負担を感じる所である。特に図形情報の場合、ある図形との比較によって検索を行ないたい場合が多いので、従来の方式では非効率といえる。また、単に図形に記述された事実に関する検索だけでなく、与えられた図形に何らかの観点から類似な図形を取り出すという類似検索も図形情報独特なものである。このような場合、比較する図形を直接システムに入力した後に補助的な検索コマンドを与えるという形式が効率的であり、そのための検索手法の確立が重要となってくる。

このような立場から、我々は天気図を対象として、図形情報を検索する場合に図形を query の一部として与え、類似等の条件で検索を行なう方式について研究を進めてきた。本論文では特に、天気図における類似の評価方法及びそれに基づいた類似検索について述べる。

2. 類似検索の問題点

図形の類似性を考えた場合、大きく分けて、図形を単に画素の配列として考え、数学的に類似度を求めて類似性を評価する構造的見地と、図形の表現してい

* 情報システム学専攻

** 情報システム学専攻修士課程 [現在(株)日本IBM]

る意味を考慮に入れて類似性を考える意味の見地という2つの異なった立場が考えられる。構造的類似性と意味的類似性を比較した場合、一般には前者の方が機械的なアルゴリズムとして実現することが容易であるのは明らかである。従って、意味的な相違が明確に

図形の構造的性質に反映される場合は、構造的類似性を評価の方が効率的であると考えられる。

ここで対象とする天気図に関しては、気象情報の観点から類似性を評価しなければならないが、天気図の場合は記号化された情報が多いため、必ずしも構造的類似性では十分でない場合が多い。例えば、日本列島の西方に高気圧、東方に発達した低気圧の存在する典型的な冬型の天気図 (Fig. 1) について、この西方の高気圧の記号 *H* と東方の低気圧の記号 *L* とを入れ替えたとしても、他の部分が全く等しければ、図形の構造上からは著しい差異はなく、むしろ入れ替えた部分だけが異なる極めて類似した図形ということができる。しかし、気象条件を表現する天気図として見た場合、気圧記号を入れ替えると、気象的な概念が変化するという以前に、等圧線・前線・気圧値等の関係が不正になってしまう。また、単に高気圧と低気圧の関係だけを見ても、西高東低の気圧配置の典型的な冬型の天気図から、日本海側でフェーン現象の起こり易い春先の気候を表わす天気図に変化することになり、類似の評価には意味的な評価を加えなければならないことは明らかである。

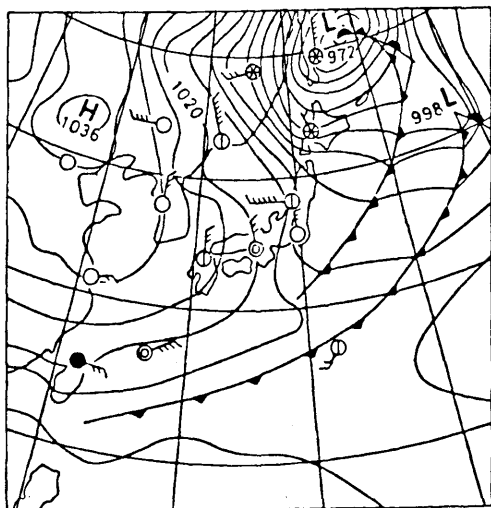


Fig. 1 Example of weather chart

Table 1 Semantic structure of weather chart

constituent	attribute	position	pressure	number	wind force	weather	wind direction
low pressure	L	○	○				
high pressure	H	○	○				
tropical depression	T.D	○	○				
typhoon	T	○	○	○			
warm front		○					
cold front		○					
stationary front		○					
occluded front		○					
isobar		○	○				
local weather		○			○	○	○

3. 意味的類似性の検討

3.1. 天気図の意味と類似性

天気図の意味的類似性を考えた場合、まず、天気図に表現される“意味”がどのような構造をしているかを明確にしなければならない。既に我々は、気象情報理解システム ISOBAR²⁾の開発に際して、天気図を **Table 1** に示すような事物と属性の関係として分析を行っている。従って、天気図の意味的類似性はこれらの事物と属性によって評価が可能であると考えられる。この場合、事物は意味を構成する上で基本となる主体的要素であるので、各事物の表現する情報の類似性、換言すれば各事物における属性値の一致の程度が、全体の一致の程度、即ち図形の意味的類似性に貢献するというのが自然な考え方である。問題は、各事物における属性値の一致の程度をどのような評価するかということと、そのプリミティブな一致の程度をどのように全体の類似性へ統合するかという二点にある。

そこで、前者については、異った二枚の天気図において対応する事物の類似の程度（類似度）を $[0, 1]$ の値で表現することとし、事物の各属性値の差に基づいて、差が小さいほど類似度が大きくなるような評価関数を各属性毎に定めた。各属性における評価関数については、後で詳細に述べるが、基本的にはこれらは各属性値の差が“類似している”及び“類似していない”のカテゴリーに属する確率として考える事ができる。一方、全体の類似度を評価する方法であるが、各事物の属性の類似の程度を確率として考えると、それらの確率を基に天気図全体として“類似している”及び“類似していない”のカテゴリーに属する確率を推定するという問題に置換することができる。

3.2. 類似度と確率推定

従来、確率の推定問題の多くは事前確率と条件付確率に基づく Bayes の定理の応用として考えられて来た。しかし、全体を構成する要素の数が増加すると、条件付確率によって要素間の関係を明確にしておくことは困難になってくる。また、天気図の類似性を考えた場合、二枚の天気図が類似性に関するアプリアリな確率を求めることは困難である。そこで、本研究では Dempster の組み合わせ理論³⁾に基づく確率の推定を用いた。

この Dempster の理論は、Shafer によって整理されたが⁴⁾、その理論の確率の公理は以下のように与えら

れる。

$$\begin{aligned} m(\phi) &= 0 \\ 0 &\leq m(A) \leq 1 \\ \sum_{A \subseteq \Theta} m(A) &= 1 \end{aligned}$$

ただし、 $m(A)$ を全体集合 Θ の部分集合 A に対する基本確率 (basic probability number) と呼ぶ。

この基本確率は「部分集合 A であるが、 A の真部分集合であると断定できないもの」の“信用度”を表わす測度と定義される。例えば、 $m(\Theta)$ は「どの真部分集合に属しているか分からない」ことに対する“信用度”換言すれば「 Θ の不明確さ」を意味することになるが、この不明確さの概念の導入によってその基本確率の持つ系への信頼性を表現することが可能になる。また、従来の確率論での集合 A に属している“確率” $P(A)$ に対応する信用度、即ち、 A の真部分集合に属している信用度を含めて、 A である信用度 $Bel(A)$ は

$$Bel(A) = \sum_{B \subseteq A} m(B)$$

として与えられることになる。

この基本確率を基に全体集合 Θ に対する 2 つの基本確率 m_1, m_2 を組み合わせた基本確率 m は Dempster の組み合わせ則 (Dempster's Rule of Combination, 以下では DRC と略) として以下のように与えられる。

$$m(A) = \frac{\sum_{A_i \cap B_j = A} m_1(A_i) m_2(B_j)}{1 - \sum_{A_i \cap B_j = \phi} m_1(A_i) m_2(B_j)}$$

この式の分母は全ての組み合わせのうち矛盾するものを除くことを意味し、分子は目的とする集合 A がどのような組み合わせで表現されるかということの意味している。つまり、目的とする集合の基本確率は、それを表現する組み合わせが、実際に矛盾しない組み合わせのうちどの程度を占めていることになるかということと求められる。以下ではこの組み合わせを \oplus で表現し、 $m = m_1 \oplus m_2$ と記す。

上述の確率モデルで類似度の評価を行なうということは、「系は『類似』 (A_1)、『非類似』 (A_2)、『不明』 (= 全体集合 Θ) に分類されるとした上で、各々に基本確

率 $m(\cdot)$ を割り当てる」ことと考えることができる。この場合、 $\Theta = A_1 \cup A_2, A_1 \cap A_2 = \phi$ とする。尚、これからの議論に於いて、2枚の天気図 I_1, I_2 の類似に関する基本確率であることを明確にする時には、 $m(A)$ を特に $m(A; I_1, I_2)$ と表現する。

3.3. 各事物の類似度

3.1., 3.2. で、類似性を評価するための基本的な考え方を述べたが、確率の組み合わせの要素となる各事物の属性の類似度については、厳密な議論が困難である。従ってここでは、以下に述べるような類似度を実験的に定めた。

① 各地の天気 m_{LW}

「各地の天気」は属性として、位置・天気・風向・風力を持っているが、位置は全ての天気図で同一であるので、類似度への影響は考慮しなくてもよい。そこで天気・風向・風力の各々について類似度に対応する基本確率 $m_{weather} \cdot m_{wind} \cdot m_{windF}$ を定義し、それをDRCによって組み合わせたものを「各地の天気」の類似度評価の基本確率 m_{LW} とした。

「天気」に関しては、属性値が非数値であるため、その類似性を数値に変換するための関数 $f_{weather}$ を **Table 2** のように定めた上でその基本確率を定義した。また、風向に関しても、属性値は非数値であるが、北を1、北北西を16として時計回りに、数値に変換した上で、基本確率を定義した。尚、風向が定義されていない場合、つまり風力0の場合は、風向も0

とする。

実際の定義は以下の通りである。

$$m_{weather}(A_1; I_1, I_2) = C_{weather} p_{weather}(I_1, I_2)$$

$$m_{weather}(A_2; I_1, I_2) = C_{weather} - m_{weather}(A_1; I_1, I_2)$$

$$m_{weather}(\Theta; I_1, I_2) = 1 - C_{weather}$$

$$p_{weather}(I_1, I_2) = \frac{1}{N_{pos}} \sum_{x \in pos} f_{weather}(weather(I_1, x), weather(I_2, x))$$

N_{pos} : 天気図中に出現する「各地の天気」の数 (= 15)

pos : 「各地の天気」の存在する位置の集合

$f_{weather}$: Table 2 参照

$weather(I, x)$: 天気図 I の位置 x の天気

$$m_{windD}(A_1; I_1, I_2) = C_{windD} p_{windD}(I_1, I_2)$$

$$m_{windD}(A_2; I_1, I_2) = C_{windD} - m_{windD}(A_1; I_1, I_2)$$

$$m_{windD}(\Theta; I_1, I_2) = 1 - C_{windD}$$

$$p_{windD}(I_1, I_2) = \frac{1}{N_{pos}} \sum_{x \in pos} f_{windD}(windD(I_1, x), windD(I_2, x))$$

$$f_{windD}(a, b) = \begin{cases} 1 & a=0 \wedge b=0 \\ 0 & a=0 \vee b=0 \\ \max\left(0, \frac{\alpha - |a-b|}{\alpha}\right) & |a-b| \leq 8 \\ \max\left(0, \frac{\alpha - (16 - |a-b|)}{\alpha}\right) & |a-b| > 8 \end{cases}$$

Table 2. Definition of $f_{weather}$

	A	B	C	D	E	F	G	H	I	J	K	L	M
A clear	1.0	0.8	0.5	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
B fare	0.8	1.0	0.6	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
C overcast	0.5	0.6	1.0	0.5	0.5	0.5	0.0	0.0	0.0	0.0	0.3	0.3	0.3
D rain	0.0	0.0	0.5	1.0	0.7	0.8	0.0	0.0	0.0	0.0	0.7	0.6	0.6
E snow	0.0	0.0	0.5	0.7	1.0	0.8	0.0	0.0	0.0	0.0	0.4	0.7	0.7
F sleet	0.0	0.0	0.5	0.8	0.8	1.0	0.0	0.0	0.0	0.0	0.4	0.7	0.7
G fog	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0	0.0
H smog	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0	0.0
I sandstorm	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0	0.0
J blizzard	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	1.0	0.0	0.0	0.0
K thundershower	0.0	0.0	0.3	0.7	0.4	0.4	0.0	0.0	0.0	0.0	1.0	0.5	0.5
L hailstone	0.0	0.0	0.3	0.6	0.7	0.7	0.0	0.0	0.0	0.0	0.5	1.0	0.8
M hail	0.0	0.0	0.3	0.6	0.7	0.7	0.0	0.0	0.0	0.0	0.5	0.8	1.0

$windD(I, x)$: 天気図 I の位置 x の風向

風力 $0 \rightarrow windD = 0$

$$m_{windF}(A_1; I_1, I_2) = C_{windF} p_{windF}(I_1, I_2)$$

$$m_{windF}(A_2; I_1, I_2) = C_{windF} - m_{windF}(A_1; I_1, I_2)$$

$$m_{windF}(\Theta; I_1, I_2) = 1 - C_{windF}$$

$$p_{windF}(I_1, I_2) = \frac{1}{N_{pos}}$$

$$\sum_{x \in pos} f_{windF}(windF(I_1, x), windF(I_2, x))$$

$$f_{windF}(a, b) = \max\left(0, \frac{\beta - |a - b|}{\beta}\right)$$

$windF(I, x)$: 天気図 I の位置 x の風力

α, β : 類似度の感度を調整する定数 ($\alpha = 4, \beta = 12$)

これらの式は、各地における天気・風向・風力の相違の平均と考えてよい。ここで $C_{weather}, C_{windD}, C_{windF}$ は系に対する「天気」「風向」「風力」の信頼度と考えることができる。

上記の基本確を基にした「各地の天気」の類似度に対応する基本確率 m_{LW} は、ここでは

$$m_{LW} = m_{weather} \oplus m_{windD} \oplus m_{windF}$$

としている。

② 寒冷前線 m_{CF}

寒冷前線の持つ属性は位置のみであり、属性値としては各寒冷前線に対して、その通過位置の列が与えられている。ここでは、各寒冷前線に対し通過位置の列を代表的通過位置1点と傾(0~7)に変換し、その値に基づいて類似度の基本確率 m_{CF} を定義した。基本的には、2つの天気図で対応する寒冷前線の位置が近いほど基本確率が大きくなるよう定めた。また、複数の寒冷前線が存在する場合は、各々に最も類似している ($m(A_1)$ が最大) 前線に対応させて基本確率を求め、それらを DRC で組み合わせさせている。

$$m_{CF}(; i, I_2) = \left(\bigoplus_{j \in CF(I_1)} \hat{m}_{CF}(; i, I_2) \right) \oplus \left(\bigoplus_{j \in CF(I_2)} \hat{m}_{CF}(; j, I_1) \right)$$

$$\bigoplus_{j=1}^n m_1 = m_1 \oplus \dots \oplus m_n$$

$CF(I)$: 天気図 I の寒冷前線の集合

$\hat{m}_{CF}(; i, I)$: $m'_{CF}(A_1; i, j)$ ($j \in CF(I)$) が最大となる時の $m'_{CF}(; i, j)$

$$m'_{CF}(; i, j) = m_{CF}^{pos}(; i, j) \oplus m_{CF}^{dir}(; i, j)$$

$m_{CF}^{pos}(; i, j)$: 寒冷前線 i, j の位置の類似度

$m_{CF}^{dir}(; i, j)$: 寒冷前線 i, j の傾の類似度

$$m_{CF}^{pos}(A_1; i, j) = C_{CF}^{pos} p_{pos}(position(i), position(j))$$

$$p_{pos}(a, b) = \max\left(0, \frac{\gamma - distance(a, b)}{\gamma}\right)$$

γ : 定数 (250 pixel length)

$$m_{CF}^{dir}(A_1; i, j)$$

$$= C_{CF}^{dir} p_{dir}(direction(i), direction(j))$$

$$p_{dir}(a, b) = \max\left(0, \frac{\alpha - |a - b|}{\alpha}\right)$$

$CF(I_1) = \phi \wedge CF(I_2) = \phi$ の場合は

$$m_{CF}(\Theta; I_1, I_2) = 1$$

いずれか一方のみ空集合の場合

$$m_{CF}(A_1; I_1, I_2) = 0,$$

$$m_{CF}(A_2; I_1, I_2) = C_{CF}^{pos}$$

③ 温暖前線 m_{WF}

④ 閉塞前線 m_{OF}

⑤ 停滞前線 m_{SF}

寒冷前線と同様。

⑥ 低気圧 m_{LP}

低気圧の持つ属性は気圧値及び位置であり、ここではそれらの属性値に基づいて類似度の基本確率 m_{LP} を定義した。基本的には、2つの天気図で対応する低気圧の位置が近いほど、気圧値が近いほど基本確率が大きくなるように定めた。また、複数の低気圧が存在する場合は、寒冷前線の組合と同様、各々に最も類似している ($m(A_1)$ が最大) の低気圧を対応させて基本確率を求め、それらを DRC で組み合わせさせている。

$$m_{LP}(; I_1, I_2) = \left(\bigoplus_{i \in LP(I_1)} \hat{m}_{LP}(; i, I_1) \right) \oplus \left(\bigoplus_{j \in LP(I_2)} \hat{m}_{LP}(; j, I_1) \right)$$

$LP(I)$: 天気図 I の低気圧の集合

$\hat{m}_{LP}(; i, I)$: $m'_{LP}(A_1; i, j)$ ($j \in LP(I)$) が最大となる

$$m'_{LP}(i, j) = m_{LP}^{pos}(i, j) \oplus m_{IS}^{pres}(i, j)$$

$m_{LP}^{pos}(i, j)$: 低気圧 i, j の位置の類似度

$m_{IS}^{pres}(i, j)$: 低気圧 i, j の気圧の類似度

$$m_{LP}^{pos}(A_1; i, j) = C_{LP}^{pos} p_{pos}(position(i), position(j))$$

$$m_{IS}^{pres}(A_1; i, j) = C_{IS}^{pres} p_{pres}(pressure(i), pressure(j))$$

$$p_{pres}(a, b) = \max\left(0, \frac{\delta - |a - b|}{\delta}\right)$$

δ : 定数 (30 mbar)

$LP(I_1) = \phi \wedge LP(I_2) = \phi$ の場合は

$$m_{LP}(\Theta; I_1, I_2) = 1$$

いずれか一方のみ空集合の場合

$$m_{LP}(A_1; I_1, I_2) = 0,$$

$$m_{LP}(A_2; I_1, I_2) = C_{LP}^{pos}$$

⑦ 高気圧 m_{HP}

⑧ 熱帯性低気圧 m_{TD}

⑨ 台風 m_{TP}

低気圧と同様.

⑩ 等圧線 m_{IS}

等圧線の持つ属性は気圧値及びその通過位置であるが、本質的な情報は、複数の等圧線によって表現される気圧分布という2次元の広がりを持つものである。従って、等圧線の類似性はこれらの情報を反映するものでなければならないが、2次元な情報は、本研究で分析した天気図の意味構造には直接表現されていない。そこでここでは、便宜的に天気図を1024 (= 32 × 32) の小領域に等分し、各小領域の等圧線の本数・平均傾・平均気圧を求めて、それらの値に基づいて類似性の評価を行なった。この3つの情報は各々、その小領域の気圧分布の勾配・勾配の向き・平均気圧を近似的に表現していると考えられる。

具体的には、まず本数・平均傾・平均気圧の情報に対して各小領域での相違の平均を求めて、それに基づいて各情報の類似度に対応する基本確率を求め、更にそれらをDRCで組み合わせている。

$$m_{IS} = m_{IS}^{no} \oplus m_{IS}^{pres} \oplus m_{IS}^{dir}$$

m_{IS}^{no} : 等圧線の本数に関する類似度

m_{IS}^{pres} : 等圧線の気圧に関する類似度

m_{IS}^{dir} : 等圧線の傾に関する類似度

$$m_{IS}^{no}(A_1; I_1, I_2) = C_{IS}^{no} p_{no}(I_1, I_2)$$

$$p_{no}(I_1, I_2) = \frac{1}{N_{ISO}}$$

$$\sum_{i \in ISO} \frac{\min(N(I_1, i), N(I_2, i))}{\max(N(I_1, i), N(I_2, i))}$$

$N(I, i)$: 天気図 I における最小領域 i の等圧線数

ISO : $N(I_1, i) > 0 \wedge N(I_2, i) > 0$ である小領域の集合

N_{ISO} : ISO の要素数

$$m_{IS}^{pres}(A_1; I_1, I_2) = C_{IS}^{pres} p_{IS}^{pres}(I_1, I_2)$$

$$p_{IS}^{pres}(I_1, I_2) = \frac{1}{N_{ISO}}$$

$$\sum_{i \in ISO} p_{pres}(pressure'(I_1, i), pressure'(I_2, i))$$

$$m_{IS}^{dir}(A_1; I_1, I_2) = C_{IS}^{dir} p_{IS}^{dir}(I_1, I_2)$$

$$p_{IS}^{dir}(I_1, I_2) = \frac{1}{N_{ISO}}$$

$$\sum_{i \in ISO} p_{dir}(direction'(I_1, i), direction'(I_2, i))$$

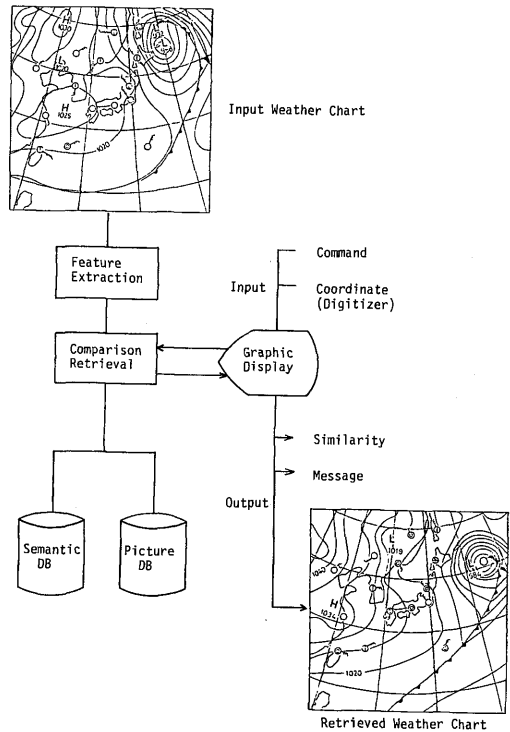


Fig. 2 Configuration of experimental system

$pressure'(I, i)$: 天気図 I の小領域 i で等圧線平均気圧

$direction'(I, i)$: 天気図 I の小領域 i で等圧線平均傾

上述の①～⑩の基本確率を基にした天気図の類似度に対応する基本確率 m_{SIM} は

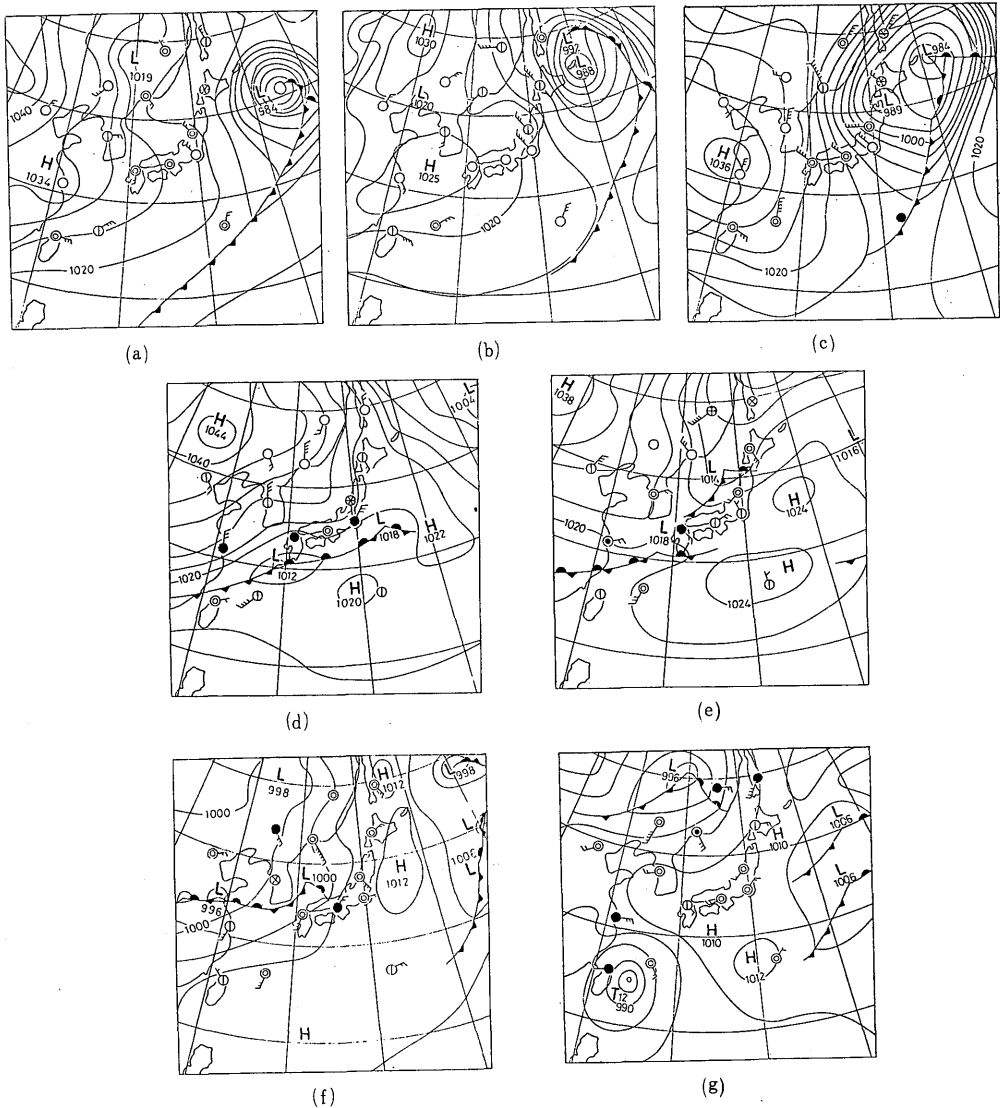
$$m_{SIM} = m_{LW} \oplus m_{CF} \oplus m_{WF} \oplus m_{OF} \oplus m_{SF} \oplus m_{LP} \oplus m_{HP} \oplus m_{TD} \oplus m_{TP} \oplus m_{IS}$$

であり、この $m_{SIM}(A_1; I_1, I_2)$ が大きい程、天気図 I_1, I_2 が類似していると評価する。

4. 実 験

4.1. 天気図検索の評価システム

天気図検索の実験を行なうために作成したシステムの構成を Fig. 2 に示す。Picture DB は、天気図表示のための原画の集合であるが、ここでは記憶容量の削減を図るため DF-画像表現⁵⁾を用いて符号化し、蓄積



している。一方、Semantic DBには、意味的類似度を評価するための Picture DB内にある天気図の事物とその属性値の表、及び等圧情報の類似度を評価するための分割小領域の等圧線本数・平均傾・平均気圧の集合が蓄積されている。また、ここでは述べていないが、天気図の構造的類似度を評価するための複雑度・素面スペクトラム⁵⁾も蓄積している。

特徴抽出部は Semantic DB に蓄えられた情報と比較すべき情報を入力天気図から抽出する部分であり、天気図認識システム⁶⁾等からなっている。現在のところ、認識速度等の問題があるため以下の実験では既に Semantic DB に蓄積された情報相互の類似度の計算を行ない、本論で述べた意味的類似性の評価の妥当性を検討した。

比較検索部は、類似度の計算や、天気図の事実の真偽を判定して検索を行なう内容検索⁷⁾の述語計算を行なっている。また、通常検索のコマンドはキーボードから入力するが、特定の地域をグラフィックディスプレイの画面上で直接指定できるようタブレットディスプレイも利用している。

4.2. 実験結果

実際の類似度の評価を行なう際の、基本確率の信頼度 $\{C_i\}$ は経験的に以下のように定めた。

$$\begin{aligned} C_{\text{weather}} &= 0.1, C_{\text{windD}} = 0.05, \\ C_{\text{windF}} &= 0.05, \\ C_{CF}^{\text{pos}} &= C_{WF}^{\text{pos}} = C_{DF}^{\text{pos}} = C_{SF}^{\text{pos}} = 0.2, \\ C_{CF}^{\text{dir}} &= C_{WF}^{\text{dir}} = C_{DF}^{\text{dir}} = C_{SF}^{\text{dir}} = 0.1, \\ C_{LF}^{\text{pos}} &= C_{HP}^{\text{pos}} = C_{DP}^{\text{pos}} = C_{SP}^{\text{pos}} = 0.2, \\ C_{LP}^{\text{pos}} &= C_{HP}^{\text{pos}} = C_{DP}^{\text{pos}} = C_{SP}^{\text{pos}} = 0.1, \\ C_{IS}^{\text{no}} &= 0.1, C_{IS}^{\text{res}} = 0.1, C_{IS}^{\text{dir}} = 0.1 \end{aligned}$$

これは、各事物の類似度がほぼ同一の信頼度を持つことを仮定しているが、気象状況は基本的には気圧分布に基づくので、等圧線の類似度の信頼度を多少高くしている。また、各地の天気は局所的情報であるので、信頼度を低くしている。

上述の値を用いて類似検索を行った例を Fig. 3 に示す。(a) に対し (b)(c), (d) に対し (e), (f) に対し (g) が各々類似天気図として検索されたものである。ここで類似というのは基本確率が $m(A_1) > m(A_2)$ となったこととしている。尚、実験に用いた天気図は約 30 枚である。

この結果からも判明するように、システムが類似と

判定した天気図は、人間が判断してもほぼ類似と考えられるものである。しかし、Fig. 3 (f), (g) に関しては、必ずしも類似とは判断できないが、これは現在の基本確率の導出法では低気圧、高気圧の数が多いと系の不明確さ $(m(\Theta))$ が減少するので、それらの類似度の高さが全体の類似度に大きく影響しているためと考えられる。

5. 結 言

本論文では、天気図を対象として図形の類似性に基づいた検索方法について述べた。ここでは、天気図の意味構造に基づいて、全体の類似性を、天気図を構成する事物情報の類似性から推定評価するという手法を用いた。その際、類似度を確率モデルとして捉え、Dempster の組合わせ則に従って類似度の評価を行なった。実験結果より、人間が見て類似していると判断できる天気図の対が、本手法でもほぼ類似しているものとして評価されることが判明した。

本手法で類似度を評価する場合の問題点は、各事物に関する類似度に対応する基本確率を求める際の信頼度 $\{C_i\}$ をどのように定めるかという点にある。この点に関しては、更にデータの件数を増やしてその妥当性の再評価を進めると共に、気象の専門家としての「知識」を導入する必要があり、今後の課題といえる。また、現在のシステムではこの $\{C_i\}$ 及び各事物の類似度計算のパラメータは固定にしているが、検索目的に応じてそれらの値を対話的に変更出来るようにすることも必要である。

謝 辞

日頃御指導賜わる、本研究科情報システム学専攻、田町常夫教授に感謝致します。また、天気図の意味構造に関する貴重な助言を頂いた本学医学部附属病院、横田将生講師にも謝意を表します。

参 考 文 献

- 1) Tamura, H. and Yokoya, N.: "Image database system: a survey", Pattern Recognition, Vol. 17, No. 1 (1984).
- 2) Kawaguchi, E., Yokota, M., Endo, T., Taniguchi, R. and Tamati, T.: "An Information Understanding System of Basic Weather Report", Trans. of IECE Japan. Vol. E64, No. 2 (1981).
- 3) Dempster, Arthur P.: "Upper and lower probabilities induced by a multivalued mapping", Annals of Mathematical

- Statistics, Vol. 38 (1967).
- 4) Shafer, G.: A mathematical theory of evidence, Princeton University Press, Princeton (1976).
 - 5) 遠藤, 河口: "DF-画像表現の性質と情報圧縮への応用", 電子通信学会論文誌. Vol. J62-D, No. 2 (1979).
 - 6) Taniguchi, R., Yokota, M., Kawaguchi, E. and Tamati, T.: "Knowledge-based Picture Understanding of Weather Chart", Pattern Recognition, Vol. 17, No. 1 (1984).
 - 7) Yokota, M., Taniguchi, R. and Kawaguchi, E.: "Language-Picture Question-Answering through Common Semantic Representation and Its Application to the World of Weather Report", in Natural Language Communication with Pictorial Information Systems, L. Bolc ed., Springer-Verlag (1984).