

선형판별법에 의한 GMS 영상의 객관적 운형분류

서애숙 · 김금란
기상연구소 위성기상부
(1990년 2월 10일 받음; 1990년 2월 20일 수리)

Objective Cloud Type Classification of Meteorological Satellite Data Using Linear Discriminant Analysis

Suh Ae-Sook and Kim Kum-Lan
Meteorological Research Institute, Satellite Meteorology Division
(Received February 10, 1990; Accepted February 20, 1990)

Abstract

This is the study about the meteorological satellite cloud image classification by objective methods. For objective cloud classification, linear discriminant analysis was tried. In the linear discriminant analysis 27 cloud characteristic parameters were retrieved from GMS infrared image data. And, linear cloud classification model was developed from major parameters and cloud type coefficients. The model was applied to GMS IR image for weather forecasting operation and cloud image was classified into 5 types such as Sc, Cu, CiT, CiM and Cb. The classification results were reasonably compared with real image.

1. 서 론

광범위한 지역의 연속적이고 실시간적으로 제공되는 기상위성 영상자료는 종관규모현상에서부터 중·소규모 대류현상에 이르기까지 그 현상을 이해하고 예보하는 데 매우 중요한 자료로 이용되고 있다. 그러나 기상위성 영상자료를 보다 신속 정확하게 예보에 활용하기 위해서

는 무엇보다 영상의 객관적인 운형분류 및 분석이 이루어져야 할 것이다.

현재는 분석자가 적외·가시영상 자료를 종합하여 주관적인 해석방법으로 운형을 분류한 다음 구름분석도(nephanalysis)를 작성하고 있다. 그러나 최근 새로 설치된 기상위성 수신·분석 시스템에는 영상자료의 디지털화, 영상자료 계산처리 능력 증대, 응용소프트웨어 등에 힘입어 운형분류에 대한 정량적인 연구가 가능하게 되었다. 이와 같은 운형분류 방법은 그 동안 많은 학자들에 의해 연구·개발되어 왔다(Parikh, 1981; Inoue, 1987; Aso et al., 1988). 운형분류 변수로서는 대부분 적외영상과 가시영상 자료로부터 추출된 휘도온도, 평균, 편차, 반사도 등을 기초로 하였고 이 변수들로부터 상·중·하층운, 혼합형, 평균, 편차, 반사도 등을 기초로 하였고 이 변수들로부터 상·중·하층운, 혼합형, 적운 등의 운형을 분류하였다(Lilijas, 1981).

본 연구에서의 운형분류는 기상위성 적외영상의 스펙트럴 특징(spectral feature), 종관적 특징(synoptic feature) 및 텍스쳐 특징(textural feature)으로부터 27개 변수를 추출하고 이 변수와 계수값에 1차 선형판별 분석법을 적용하여 운형을 Sc(Stratocumulus), Cu(Cumulus), CiT(Cirrus only), CiM(Cirrus with low level clouds) 및 Cb(Cumulonimbus)로 분류하였다.

이와 같은 연구를 수행함으로써 기상위성 영상의 운형분류 업무가 종전의 주관해석에서 객관해석이 되고 구름분석 업무도 정확도가 향상될 뿐만 아니라 기상예보 향상에도 기여하게 될 것이다.

2. 자료 및 방법

본 연구에 사용된 자료는 GMS-4 위성이 관측한 0530 GMT 10 January 1990 우리 나라 부근 IR 영상이며 분류를 위해 사용된 방법은 선형판별 분석방법이다. 이 방법을 간단히 요약하면 Fig. 1과 같다. 먼저 GMS 적외영상의 0.5×0.5 (lat. \times lon.) 격자점 내에 포함된 TBB 자료로부터 27가지의 운형분류 변수를 스펙트랄, 종관, 텍스쳐 특징별로 추출한다. 전형적인 운형별 training sample data를 Supervisor analysis(Schowegerdt, 1983)방법으로 결정한 다음, 운형별 P개의 변량 X_p 와 계수 a 에 의한 1차 선형판별 함수식을 도출한다. 이 함수식을 실제 영상에 응용하여 운형을 분류한다. 분류된 운형을 디스플레이 및 하드카피를 하는데, 이 방법에 대한 자세한 내용은 뒤에 기술하였다.

가. 운형분류 변수 추출

1) 스펙트랄 특징(Spectral Feature)

스펙트랄 특징변수는 관측영역 내의 적외영상 자료에서 위경도 0.5×0.5 간격 격자점 내에 포함된 픽셀(pixel)의 TBB 히스토그램(Fig. 2)으로부터 구할 수 있다.

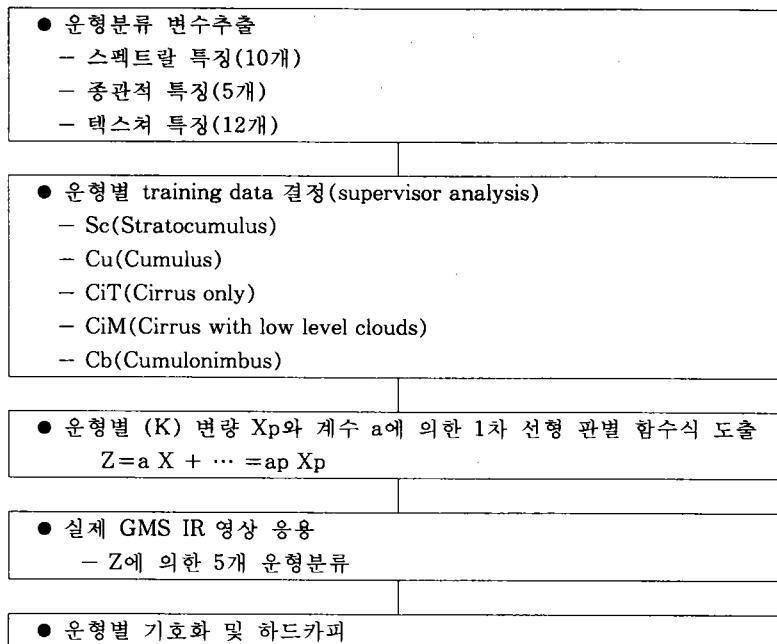


Fig. 1. Flow chart of objective cloud classification using linear discriminant analysis.

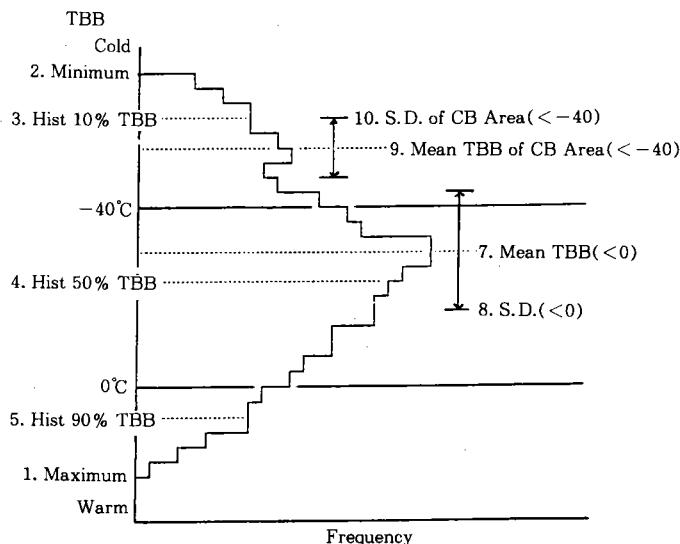


Fig. 2. TBB histogram within 0.5×0.5 (lat. \times lon.)

Fig. 2는 위경도 0.5×0.5 격자 간격 내에 포함된 적외영상 픽셀(pixel) 값에 대한 TBB 히스토그램으로서 가로축은 빈도수, 세로축은 TBB 온도를 나타낸다. 여기서 1, 2변수는 각각 TBB maximum(최대 TBB 온도값), TBB minimum(최소 TBB 온도값)이고, 3, 4, 5 변수는 각각 TBB 히스토그램에서 10%, 50%, 90% 누적빈도수를 나타내는 TBB 온도값이다. 또 6, 7, 8 변수는 각각 0°C 이하값을 갖는 픽셀수, 평균값 및 표준편차이고 9, 10 변수는 각각 -40°C 이하(Cb 영역) 값을 나타내는 픽셀의 TBB 평균값 및 표준편차이다. 이를 정리하면 Table 1과 같다.

Table 1. Cloud classification parameter extracted from spectral feature
in 0.5×0.5 grid TBB histogram

Parameter No.	Description
1	TBB Maximum
2	TBB Minimum
3	Histogram 10% TBB
4	Histogram 50% TBB
5	Histogram 90% TBB
6	Number of Cloud Pixel($<0^{\circ}\text{C}$)
7	Mean TBB($<0^{\circ}\text{C}$)
8	Standard Deviation($<0^{\circ}\text{C}$)
9	Mean TBB of CB Area($<-40^{\circ}\text{C}$)
10	Standard Deviation of CB Area($<-40^{\circ}\text{C}$)

스펙트랄 특징에서는 위와 같이 10개의 특징적인 운형분류 변수를 추출하였다.

2) 종관특징(Synoptic Feature)

종관특징 변수는 위성영상에 나타난 전체적인 구름영상의 남북규모, 확장범위 및 형태 등 종관적인 특징을 나타내는 변수로서 스펙트랄 특징과 유사한 방법으로 격자점 내의 TBB 온도로부터 구할 수 있으며 추출된 변수는 Table 2와 같다.

Table 2. Cloud classification parameter extracted from synoptic feature
in 0.5×0.5 grid TBB histogram

Parameter No.	Description
11	Mean of Mean TBB($<0^{\circ}\text{C}$) within 1.5×1.5 (lat. \times lon.)
12	Mean of Mean TBB($<0^{\circ}\text{C}$) within 2.5×2.5 (lat. \times lon.)
13	Mean of Mean TBB($<0^{\circ}\text{C}$) within 3.5×3.5 (lat. \times lon.)
14	N-S Size(Numerical $\times 0.5 = N\text{-}S$ Degree)(Hist. 90%, TBB $<0^{\circ}\text{C}$)
15	Normalized Distance from Southern Edge(Hist. 90%, TBB $<0^{\circ}\text{C}$)

11, 12, 13 변수는 위경도 0.5 격자점 내의 0°C 이하인 TBB의 평균값(7번째 스펙트랄 변수)에 대한 각각의 위경도 $1.5(3 \times 3)$, $2.5(5 \times 5)$, $3.5(7 \times 7)$ 격자점 내의 평균값이다. 14 변수는 구름의 남북크기, 15변수는 구름의 남단 가장자리에서부터 북단 가장자리까지의 거리로서 표준화된 값이다. Fig. 3은 Synoptic feature를 나타낸 것이다.

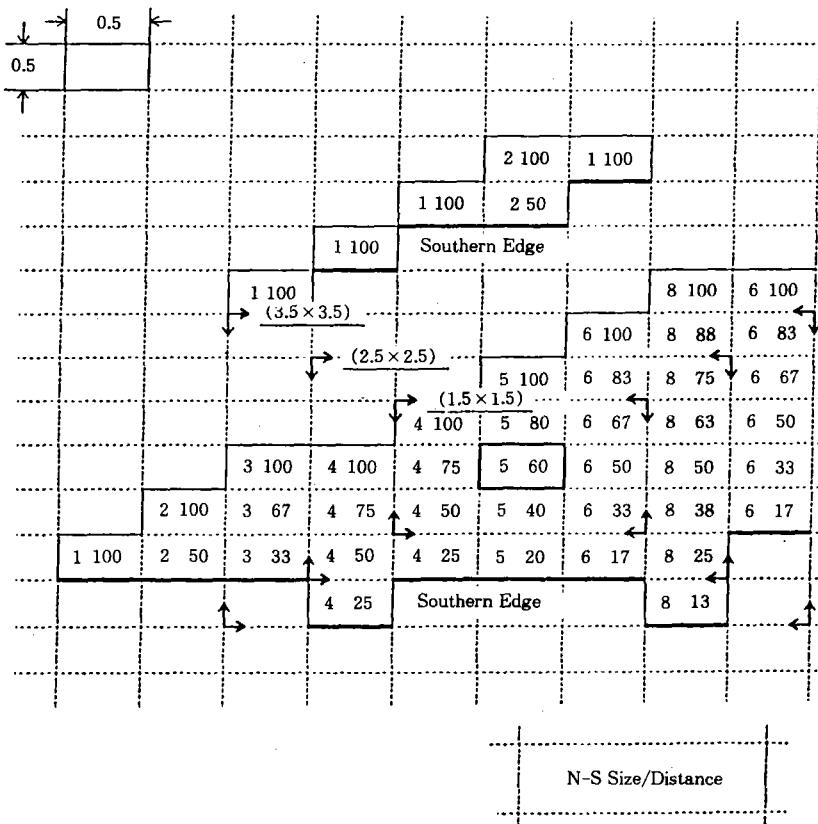


Fig. 3. Synoptic feature within 1.5×1.5 , 2.5×2.5 , 3.5×3.5 (lat. \times lon.).

3) 텍스쳐 특징(Texture Feature)

텍스쳐 특징은 관측영역 내의 표면상태에 관한 정보로서 구름의 밀도, 형태, 방향성 등에 관한 변수이다. 이는 0.5×0.5 격자점 내의 픽셀 간의 거리($\rho=1, 2, 4$)와 방향($\theta=\text{horizontal, vertical, right diagonal, left diagonal}$)에 대한 TBB 온도차로부터 구한다.

ρ, θ 에 대해 추출된 텍스쳐 특징변수는 Table 3 및 Fig. 4와 같다.

텍스쳐 특징에서는 두 픽셀 간의 거리와 방향에 따라 Fig. 4와 같이 12개의 텍스쳐 특징변수를 추출하였다.

Table 3. Cloud classification parameter extracted from texture feature in 0.5 grid TBB histogram

Parameter No.	Description
16	$\rho=1 \theta=0$ (두 픽셀간의 거리가 1, 수직방향)
17	$\rho=2 \theta=0$ (두 픽셀간의 거리가 2, 수직방향)
18	$\rho=4 \theta=0$ (두 픽셀간의 거리가 4, 수직방향)
19	$\rho=1 \theta=45$ (두 픽셀간의 거리가 1, 수평방향)
20	$\rho=2 \theta=45$ (두 픽셀간의 거리가 2, 수평방향)
21	$\rho=4 \theta=45$ (두 픽셀간의 거리가 4, 수평방향)
22	$\rho=1 \theta=90$ (두 픽셀간의 거리가 1, 수평방향)
23	$\rho=2 \theta=90$ (두 픽셀간의 거리가 2, 수평방향)
24	$\rho=4 \theta=90$ (두 픽셀간의 거리가 4, 수평방향)
25	$\rho=1 \theta=-45$ (두 픽셀간의 거리가 1, 수평방향)
26	$\rho=2 \theta=-45$ (두 픽셀간의 거리가 2, 수평방향)
27	$\rho=4 \theta=-45$ (두 픽셀간의 거리가 4, 수평방향)

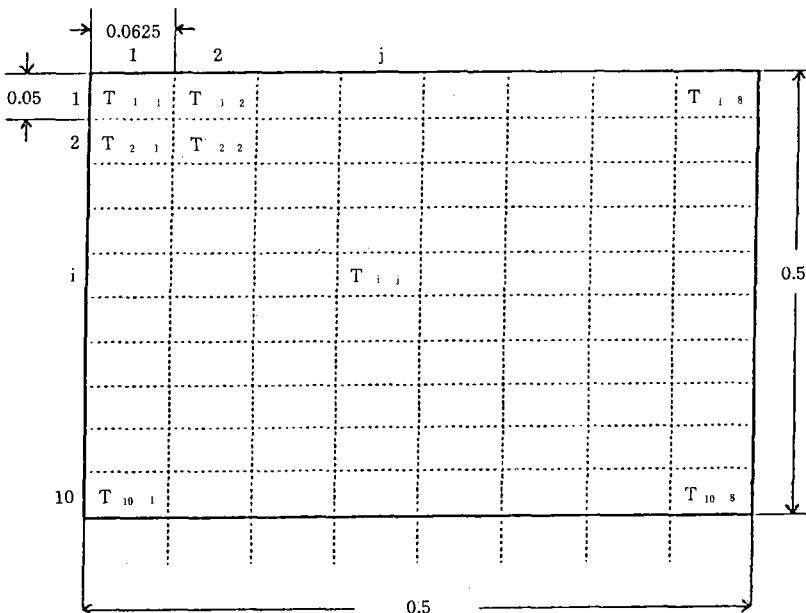


Fig. 4. Texture feature within 0.5×0.5 (lat. \times lon.).

나. 선형판별 분석법(Linear Discriminant Analysis)

1) 선형판별 함수 유도

P개의 변수 x_1, \dots, x_p 이 2개 group에서 관측된 값이 Table 4와 같다고 하자.

Table 4. Observational value of parameter(P) in 2 groups

(a) 1 group						(b) 2 group					
변량 개체	X ₁₁	X ₂₁	X _{p1}	Z ₁	변량 개체	X ₁₁	X ₂₁	X _{p1}	Z ₁
	1	X ₁₁ ⁽¹⁾	X _{p1} ⁽¹⁾	Z ₁ ⁽¹⁾		2	X ₁₁ ⁽²⁾	X _{p1} ⁽²⁾	Z ₁ ⁽²⁾
2	⋮	⋮	⋮	⋮	⋮	2	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n ₁	X _{1n₁} ⁽¹⁾	X _{pn₁} ⁽¹⁾	Z _{n₁} ⁽¹⁾		n ₂	X _{1n₂} ⁽²⁾	X _{pn₂} ⁽²⁾	Z _{n₂} ⁽²⁾	

이 2개의 group 중 어느 group에 속하는지를 알 수 없는 1개의 관측치 $x_i = (x_{i1}, \dots, x_{ip})$ 가 있을 때 관측된 P 변량값을 근거로 해서 이 값이 1 group에 속하는지 2 group에 속하는지를 판별해 주는 rule를 만든다. 따라서 선형판별 함수는 P개의 변량에 대해

예의 한 선형합성(1차식에 의한 합성변량) Z 를 만들고 이 값에 의해 판별한다. 계수 $a_1, \dots, +a_p$ 는 관측된 2개의 group자료를 더욱더 잘 판별할 수 있도록 해 준다. 여기서 계수 $a_1, \dots, +a_p$ 값이 구해지면 식 (1)을 이용해서 Table 1의 각 자료에 대한 선형결합 Z 값(판별득점이라고도 부름-discriminant score)을 구하면

이다.

여기서 k : group a : coefficient

p: parameter x: observational data

에 의해 속하는 group이 판별된다. 즉 Z값 중 가장 최대값을 나타내는 Z값의 group으로 판별 분석된다.

이는 Fig. 5에서 2 group의 평균 간의 표준화 거리를 최대로 하는 것에 상당한다.

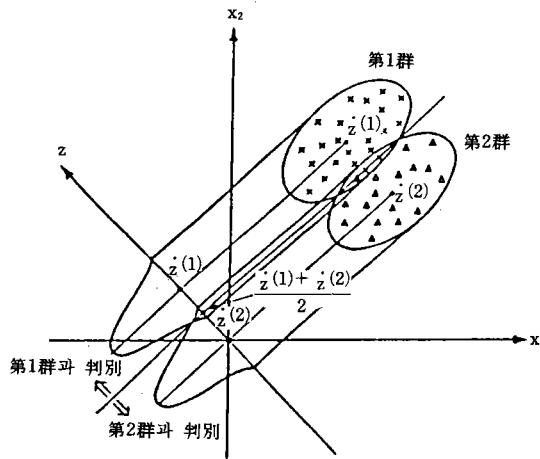


Fig. 5. 2 groups classification by linear discriminant function.

2) 운형별 training data 결정 및 운형분류

선형 판별분석법을 응용하기 전에 먼저 운형을 분류하고자 하는 영상에서 전형적인 운형 패턴을 가진 training sample data를 결정해야 한다.

결정방법은 training sample의 사용여부에 따라 감독(supervisor)과 무감독(unsupervisor) 방법으로 나누는데, 본 연구에서는 감독방법을 택하였다. 이는 각 운형의 특징을 대표할 수 있는 training data를 먼저 구하고 이 training data가 나타내는 특징이 그 운형의 특징을 설명한다는 가정하에 이 특징을 가지고 영상을 분류하는 방법이다. 따라서 얼마나 정확한 training data를 사용했느냐에 따라 결과가 달라지므로 정확한 training data 결정은 매우 중요하다. 이렇게 하여 운형별 전형적인 training data를 구한 다음 운형별 각각의 변수에 대한 계수값을 계산하고 식 (2)의 Z값에 의해 운형을 분류한다.

3. 결과 및 토의

본 연구를 위해 선정된 사례는 GMS 기상위성 영상으로서 0530 GMT 10 January 1990 관측된 IR 영상(Fig. 6)이다. 관측영역은 위도 27-47.5°N, 경도 110-140°E이다.

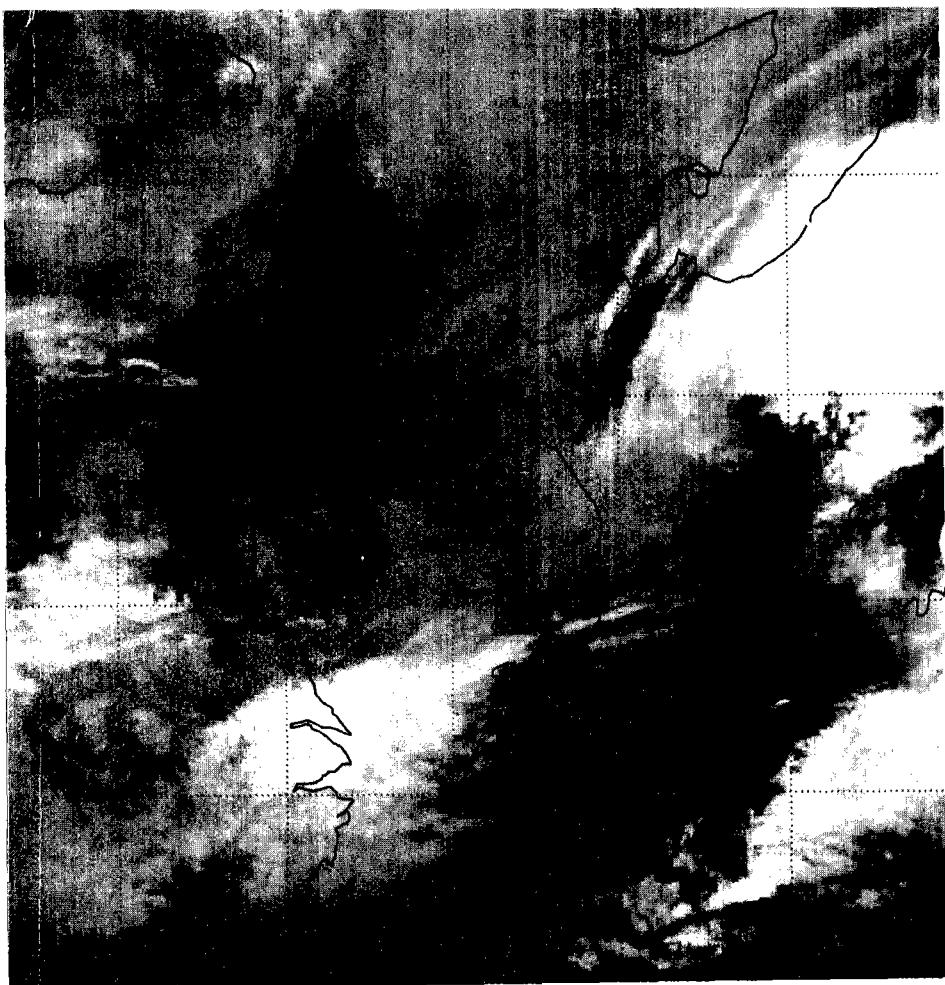


Fig. 6. GMS IR image on 0530 GMT 10 January 1990.

가. 운행분류 변수

운형분류를 위해 추출된 변수는 GMS IR 영상자료로부터 스펙트랄 특징 10개, 종관적 특징 5개, 텍스쳐 특징 12개 총 27개이다. 그 예로써 Fig. 7에 나타내었다.

Fig. 7. Maximum TBB parameter within 0.5×0.5 (lat. \times lon.) in spectral feature.

Fig. 7은 0.5×0.5 (lat. \times lon.) 범위 내 TBB 히스토그램에서 maximum TBB 값이다. 전 영역의 TBB 값을 한 장의 용지에 132컬럼 이상의 숫자를 수록할 수 없으므로 1개의 값을 2자리 수로 표기하기 위해 값의 단위를 조정하였다. +TBB는 -TBB, -TBB는 +TBB를 의미하며 *는 10°C 이상의 값으로 parameter 계산에는 사용되지 않는다.

나. 운형별 training sample data 결정

운형별 training sample data는 supervisor analysis 방법으로 결정하였으며 운형은 5종으로

서 Sc(Stratocumulus), Cu(Cumulus), CiT(Cirrus only), CiM(Cirrus with low level clouds) 및 Cb(Cumulonimbus)이다. 5종 운형에 대한 전형적인 training sample data 및 범위는 Fig. 8과 같다.

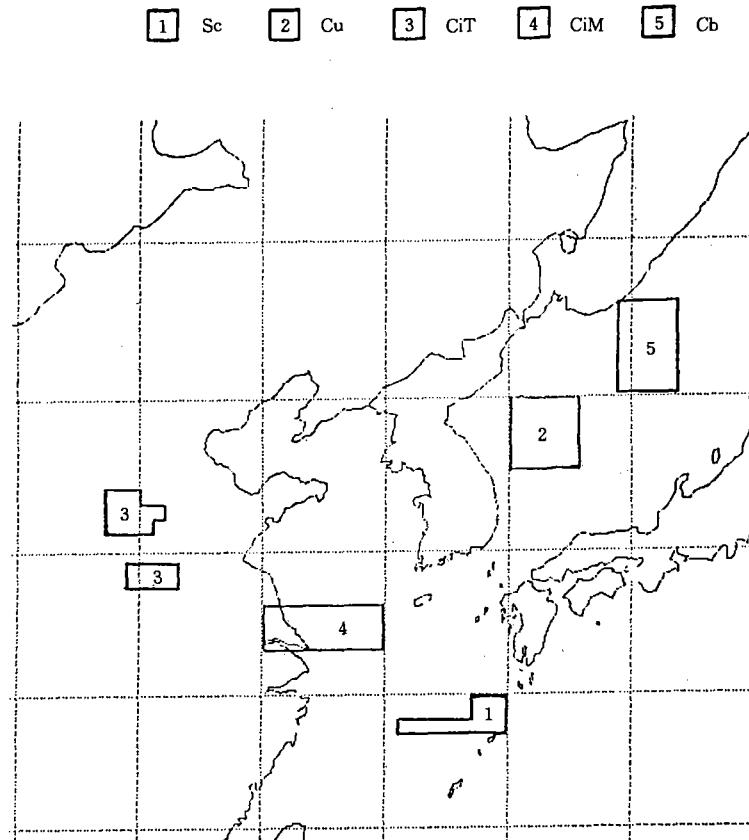


Fig. 8. Training sample area of typical Sc, Cu, CiT, CiM, Cb by supervisor analysis.

Fig. 8은 0530 GMT 10 January 1990 GMS IR 영상에서 Sc, Cu, CiT, CiM, Cb를 대표할 수 있는 자료 영역이다.

다. 선형판별 분석법에 대한 운형분류 모델

1차 선형 판별 함수식을 도출하기에 앞서 Wilks통계방법으로 각 운형별 총 27개 변수인자 중 운 특징분류에 가장 기여도가 큰 변수로 스펙트랄 특징 중 Max TBB, Min TBB, H10,

H50, H90로 5개 변수, 텍스쳐 특징 중 $\rho=1, \theta=0, \rho=2, \theta=0, \rho=4, \theta=0$ 으로 3개 변수 총 8개 변수를 추출하였다. 운형별 주요변수에 대한 평균, 분산, 표준편차 및 covariance matrix는 Table 5와 같다.

Table 5. Mean, variance, standard deviation and covariance matrix of Sc about main 8 parameters

	Var	Mean	Variance	S.D.
1	MAX	12.7056	7.9970	2.8279
2	MIN	5.3222	1.4089	1.1870
3	H10	6.3278	0.9845	0.9922
4	H50	8.1833	3.7038	1.9245
5	H90	10.9667	6.5588	2.5610
6	1, 0	0.9500	0.2085	0.4567
7	2, 0	1.4389	0.5013	0.7081
8	4, 0	1.7500	0.6485	0.8053

Covariance Matrix								
	1 Max	2 Min	3 H10	4 H50	5 H90	6 1, 0	7 2, 0	8 4, 0
1	Max	7.9970						
2	Min	1.4816	1.4089					
3	H10	1.7881	0.8523	0.9845				
4	H50	4.3489	1.0675	1.5564	3.7038			
5	H90	6.8520	1.0225	1.5451	4.1688	6.5588		
6	1, 0	0.9915	0.1176	0.1897	0.6285	0.9865	0.2085	
7	2, 0	1.5827	0.2014	0.3283	1.0025	1.5525	0.3191	0.5013
8	4, 0	2.0621	0.1882	0.3750	1.1909	1.9835	0.3315	0.5232
								.6485

운형별 주요변수에 대한 계수값은 Table 6과 같다.

Table 6. Coefficient (a) of Sc, Cu, CiT, CiM and Cb on main parameter by first linear discriminant analysis

parameter	1 Sc	2 Cu	3 CiT	4 CiM	5 Cb
Min	6.3054	-1.8301	-3.3095	-3.1729	-3.9543
H90	0.0736	0.6413	1.4512	0.7873	0.9088
1, 0	4.5377	1.1051	-0.1812	6.9999	2.0145
4, 0	-0.6299	-1.5378	-2.6883	-3.9434	-3.6639

Table 6으로부터 운형별 주요변수에 대한 계수로부터 Z(discriminant score)를 계산할 수 있다. 따라서 1개의 관측치에 대한 운형분류는 5종 운형에 대해 Z값을 계산한 다음 이 중 maximum Z값을 나타내는 운형으로 분류가 된다.

라. 실제응용 결과

0530 GMT 10 January 1990 GMS IR 영상자료에 앞의 Table 6의 운형별 주요 변수에 대한 계수값을 이용하여 각각의 관측치에 대한 운형분류를 시도하였다. 그 결과는 Fig. 9와 같다.

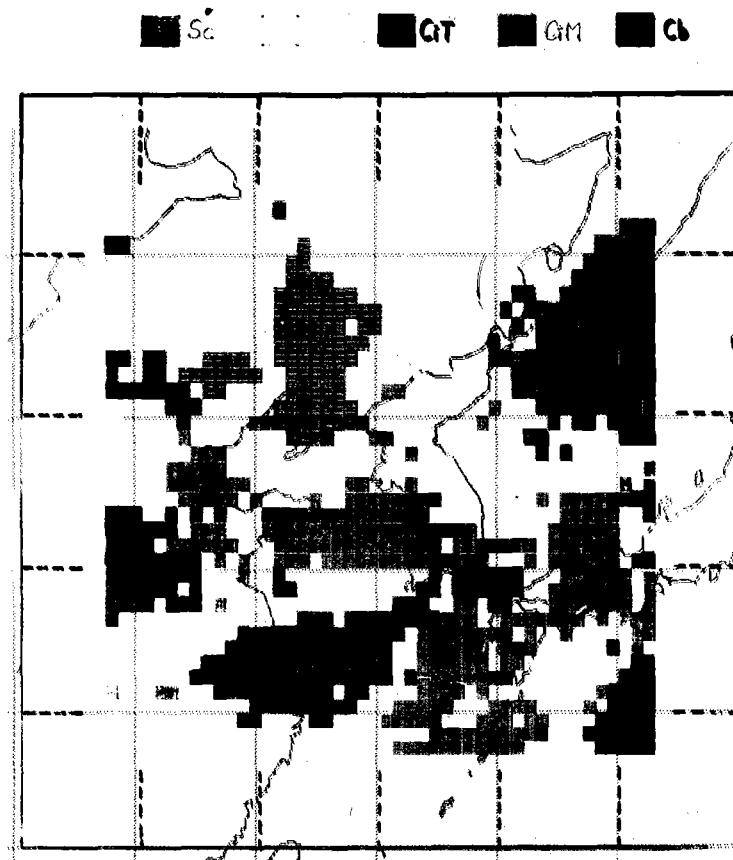


Fig. 9. Cloud type classification of GMS IR image by linear discriminant analysis.
Orange, yellow, violet, brown and black color represent Sc, Cu, CiT, CiM
and Cb, respectively.

보다 쉽게 운형구분이 용이하도록 운형별로 칼라색을 지정하여 출력하였다. 여기서 주황색은 Sc, 노랑색은 Cu, 보라색은 CiT, 고동색은 CiM, 검정색은 Cb를 의미한다. 실제영상 Fig. 6과 비교해 보면 결과가 잘 일치함을 알 수 있다.

4. 결 론

GMS 기상위성 적외영상 자료의 객관적인 운형분류를 위해 1차선형 판별 분석기법을 이용하여 5종 운형(Sc, Cu, CiT, CiM, Cb) 분류를 시도해 보았다. 운형분류 변수로는 스펙트랄 특징 10개, 종관 특징 5개 및 텍스쳐 특징 12개로 총 27개의 변수를 추출하였다. 이 중 분류에 가장 주요한 변수는 TBB Max, TBB Min, H10, H50, H90, $\rho=1$, $\theta=0$, $\rho=4$, $\theta=0$ 총 8개였다.

5종 운형별 주요변수에 대한 계수산출로 5종 운형판별 모델을 설계하였으며, 이를 0530 GMT 10 January 1990 GMS IR 영상에 응용하여 Sc, Cu, CiT, CiM, Cb운형을 분류하였다. 이 결과는 실제영상과 비교해 볼 때 잘 일치함을 알 수 있다.

본 연구를 수행함으로써 분석자마다 주관적인 해석방법의 구름분석도 작성업무를 객관화할 수 있으며 위성자료를 수신과 동시에 실시간적으로 신속하게 제공할 수 있으므로 기상예보 향상에 크게 기여하게 될 것이다.

참고문헌

- Aso, T., K. Oosawa, and Y. Takeuchi, 1988: Cloud classification in satellite imagery. A latitudinal and seasonal variation of cloud characteristic parameters, *Meteorological Satellite Center Technical Note*, **16**, 9-20.
- Inoue, T., 1987: A cloud type classification with NOAA 7 split-window as measurements, *J. Geophy. Res.*, **92**, 3991-4000.
- Liljas, E., 1981: Analysis of clouds and precipitation through automated classification of AVHRR data, *SMHI Reports RMK*, **32**, 1-25.
- Parikh, J., 1977: A comparative study of cloud classification techniques, *Remote Sensing of Environment*, **6**, 67-81.
- Schowengerdt R. A., 1983: *Techniques for Image Processing and Classification in Remote Sensing*, Academic Press, pp. 249.