

기상위성 영상의 자동분할 및 분류에 관한 연구

○ 김 은 진 * 김 평 육 * 김 정 선 *

* 한국항공대학 대학원 전자공학과

Automatic segmentation and classification of meteorological satellite images

Kim , Eun Jin Kim , Pyung Luck Kim , Jung Sun
Dept.of Electronics Eng. Hankuk Aviation College

ABSTRACT : This paper describes an automated procedure to determine cloud-types in infrared meteorological satellite images. This algorithm consists of two parts ; (1) the segmentation procedure and (2) the classification procedure. Statistical pattern recognition techniques were combined with the segmentation techniques. A texture was used to determine whether or not a cluster represented a distinct cloud object or should be merged with other clusters to form a cloud-type object.

I. 개 요

기상위성 데이터로부터 구름의 자동분류는 Automatic Cloud Mapping, 바람장율 결정하기 위한 구름의 추적, 지구 대기권의 방사 연구와 같은 분야에서 구름 유형의 식별에 대한 필요로부터 시작되었고, 구름의 분포에 대한 관측은 기상위성의 기본적인 항목으로서 초기 TIROS 위성 시대로부터 현재에 이르기까지 가장 많은 관측이 실시되어오고있다. 구름의 분포로부터 직접 태풍이나 폭풍의 위치, 진행상황 및 발달과정 등의 많은 정보를 얻을 수 있다. 이러한 분석을 위한 방법으로는 개개 화소의 그레이 레벨에 의하여 주어진 정보만을 이용하는 간단한 스테쉬홀드 기법으로부터 패턴인식 기법을 이용한 알고리즘과 각 화소에 대한 완전한 복사전도 방정식의 해에 이르기까지 그 범위가 다양하다.

디지털 위성 데이터로부터 구름의 정보를 추출하기 위한 패턴인식 기법은 일반적으로 다음과 같은 단계를 포함한다.(1) 잘 정의된 카테고리, 구름의 정보를 유한한 이산 수로 유도.(2) 구름의 특징을 계산하기 위하여 적당한 스펙트럴 영역의 결정. (3) 분류를 위한 영역 또는 윈도우로 위성영상을 분할하는 방법의 결정. (4) 특징벡터의 정의. (5) 판별 모직의 유형을 정의. (6) 판별함수의 형태 또는 판별 알고리즘의 형태를 정의. 본고에서는 GMS-3 영상의 윈도우 내에서 목표물 또는 패턴의 검출을 위한 분할과 분류 방법에 관하여 논하였다. 이를 위하여 먼저 윈도우를 그레이 레벨의 집락(Cluster)으로 자동분할하고, 분류과정에서는 각 집락의 Class를 식별한다. 이러한 식별은 그레이 레벨 집락의 통계량 또는 집락간의 통계량의 비교에 기초를 둔 것이다. 영상 분할은 스테쉬홀드 기법과[1] 영역확장 기법[2]을 포함하고있으며 영상 데이터의 분할을 위

한 스테쉬올드의 선택은 Edge 정보[3] 또는 세션화된 Edge Map과 영역 경계간의 일치성으로 결정할 수 있다.

II. 분할 공정

본 알고리즘은 그레이 레벨의 발생빈도에 따라 순서화된 집합을 집락화시키기 위하여 Fisher의 알고리즘을 적용하였다[4]. 본 알고리즘과 다른 일반적인 알고리즘과의 차이는 국부적 최적화 공정에 의한 분할보다는 전체적으로 최적한 분할이 동작 프로그램에 의하여 얻어진다는 점이다. Fisher 알고리즘에서는 가능한 모든 분할에 대해서 어려함수가 계산되므로 초기시작값이 필요없지만 원하는 집락의 최대수 K_{max} 는 주어하여야 한다. M 개의 목표물을 K 개의 집락으로 분할하는 방법의 수는 다음과 같다.

$$(1/K!) \sum_{i=1}^K (-1)^{K-i} \binom{K}{i} \frac{M!}{i!} \quad (1)$$

식(1)은 대략 $M^K/K!$ 로 근사화될 수 있다. M 개의 목표물 $(1, 2, \dots, M)$ 을 다음과 같이 주어진 K 개의 집락으로 분할하는 것을 $P(M, K)$ 로 표시하자.

$$(I_1, I_1+1, \dots, I_2-1), (I_2, I_2+1, \dots, I_3-1), \dots, (I_K, I_K+1, \dots, M) \quad (2)$$

여기서 $I_i=1$ 이고, $I_i = I_{i-1} + 1$ 이다. $1 < I_i < M$ 에서 $D(I, J)$ 를 집락의 직경이라 하자.

$$D(I, J) = \sum_{L=1}^J |X(L) - \bar{X}| \quad (3)$$

$$\bar{X} = \frac{\sum_{L=1}^J X(L)}{J - I + 1} \text{ 이고,}$$

M 개의 목표물을 K 개의 집락으로 분할함에 있어서 어려함수는 다음과 같다.

$$e[P(M, K)] = \sum_{J=1}^K D(I_J, I_{J+1}-1) \quad (4)$$

본 알고리즘은 다음과 같은 4가지 기본단계로 구성되었다.

(제 1 단계)

$1 \leq I < J \leq M$ 의 모든 가능한 I, J 에서 집락 $(I, I+1, \dots, J)$ 의 직경 $D(I, J)$ 를 계산한다.

(제 2 단계)

$2 \leq I \leq M$ 일 때, 두 집락으로 I 개의 목표물의 모든 분할에서 $2 \leq J \leq I$ 의 범위에 걸친 어려를 구한다. 이 때 I 개의 목표물을 최적한 두 개의 집락으로 분할할 때의 어려는 다음과 같다.

$$e[P(I, 2)] = \min [D(I, J-1) + D(J, I)]$$

최적 분할 $P(I, 2)$ 에서 두 번째 집락의 하측 경계는 $B(I, 2)$ 로서 $M \times K$ 크기의 행렬에 저장된다.

(제 3 단계)

$3 \leq I \leq K$ 인 각 I 인 다음 식에 의하여 최적분할 어려를 계산한다.

$$e[P(I, J)] = \min [e[P(I-1, I-1)] + D(I, I-1)], I \leq J \leq I$$

(제 4 단계)

최적 분할 $P(M, K)$ 는 첫 번째로 구한 J 에 의하여 $1 \leq I \leq J, 1 \leq I \leq M$ 에서 $e[P(M, K)]$ 의 표로부터 구한다. 따라서 $e[P(I, J)] = e[P(I-1, I-1)] + D(I, M)$ 이 되고 최종 집락은 $(J, J+1, \dots, M)$ 이 된다. 이 때 J^* 를 구하면, 다음과 같다.

$$e[P(J-1, K)] = e[P(J-1, K-1)] + D(J, J-1)$$

목표물을 K 개의 집락으로 분할하는 $P(M, K)$ 에서의 통계적 모델은 관측 $X(1), X(2), \dots, X(I_2-1)$ 이 정상밀도함수 $f(x; \theta_1) = \exp\{-1/2(x - \theta_1)^2 / \sqrt{2}\pi}\theta_1$ 로부터 독립이고, 기타 $X(I_k), X(I_k+1), \dots, X(M)$ 도 정상밀도함수 $f(x; \theta_k) = \exp\{-1/2(x - \theta_k)^2 / \sqrt{2}\pi}\theta_k$ 로부터 독립이다.

만일 관측 X 가 정상밀도함수를 갖고, 최적분할 $\hat{P}(M, K)$ 에서 $K+1$ 개의 집락이 $P(M, K)$ 의 K 개 집락중의 어느 하나로 나눔으로써 구해진다면, MSQ (Mean-square ratio)는 다음과 같게된다.

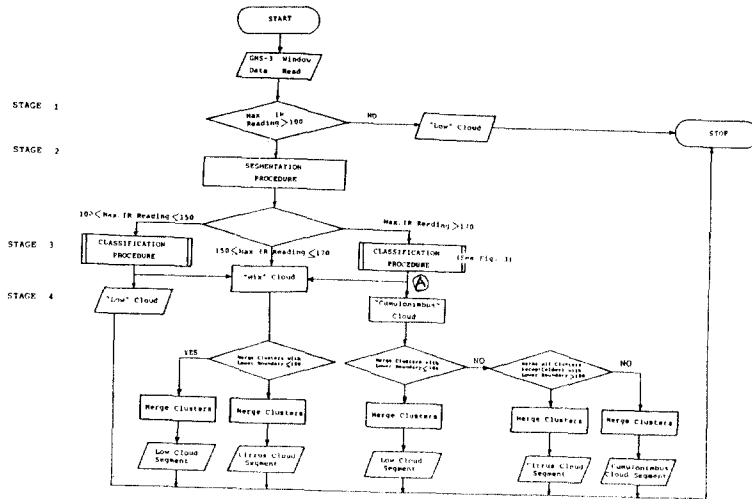
$$MSQ = (M-K-1) \frac{e[\hat{P}(M, K)]}{e[P(M, K+1)]} - 1$$

III. 분류 공정

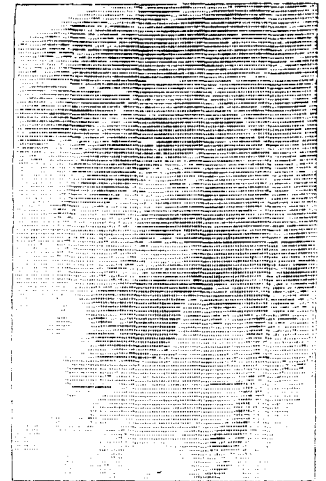
본 연구에서 수행한 GMS-3 구름 윈도우의 분할 및 분류 효용도를 그림 1.에 나타내었다. 이 알고리즘은 다음과 같은 기본 단계로 구성되었다.

(제 1 단계)

최대 그레이 레벨값의 스테쉬올드 100 (7°C 등가



(그림 1) GMS-3 영상 분석을 위한 흐름도



(그림 2) 한반도 영역(128x128) 영상

적외선 온도)이하의 데이터의 모든 표본은 하층운(사실상은 안개,육지,해면)으로 분류한다. 이때 $k=9$ 의 값에서 다른 모든 표본들에 적용된다.

(제 2 단계)

각 스택휘도값들을 나머지 표본들내에서 가능한 구름 유형의 패턴을 결정하는데 적용한다.

(제 3 단계)

본 단계는 분류공정으로서 적외선 데이터의 값들이 100과 150 사이인 표본들은 적란운을 함유하고 있는가를 판별하기 위하여 테스트되고, 150 - 170 범위의 데이터는 혼운으로서 분류된다. 즉, 이 표본들은 하층운과 권운 양자를 함유하고있는 것으로 판별한다. 또, 기상학적으로 강우량을 크게 좌우하는 적란운은 데이터값이 170 이상으로 판별한다. 분류공정은 각 윈도우를 다음 세 부류 중의 하나로 분류한다.

가. 하층운 :- 적운, 층운과/또는 층적운만을 함유하고있는 윈도우.

나. 혼운 :- 권운과 하층운을 함유하고있는 윈도우

다. 적란운 :- 적란운과 하층운을 함유하고있는 윈도우.

부류 판별은 동일한 크기의 집락으로서 크기가 다른 두 그레이 레벨 범위에 속해있는 점들에 대하여 계산된 특징값과 온도가 가장 낮은 집락에 속해있는 점들

에 대하여 계산된 텍스처 특징값과의 비교에 의하여 이루어진다. 이 값의 범위는 그림 3.에 정의되었다.

윈도우내에서 온도가 가장 낮은 구름부분은 구름의 유형을 결정하는데 쓰여진 텍스처 특징은 단위 영역당의 Edge Strength E 이다[6]. E는 주어진 영역내의 모든 데이터 점에 걸친 Robert's Gradient의 평균값으로 정의된다. 분류공정에서 2x2 배열내의 한 점 A에서 Robert's Gradient는 대략적으로 다음과 같다.

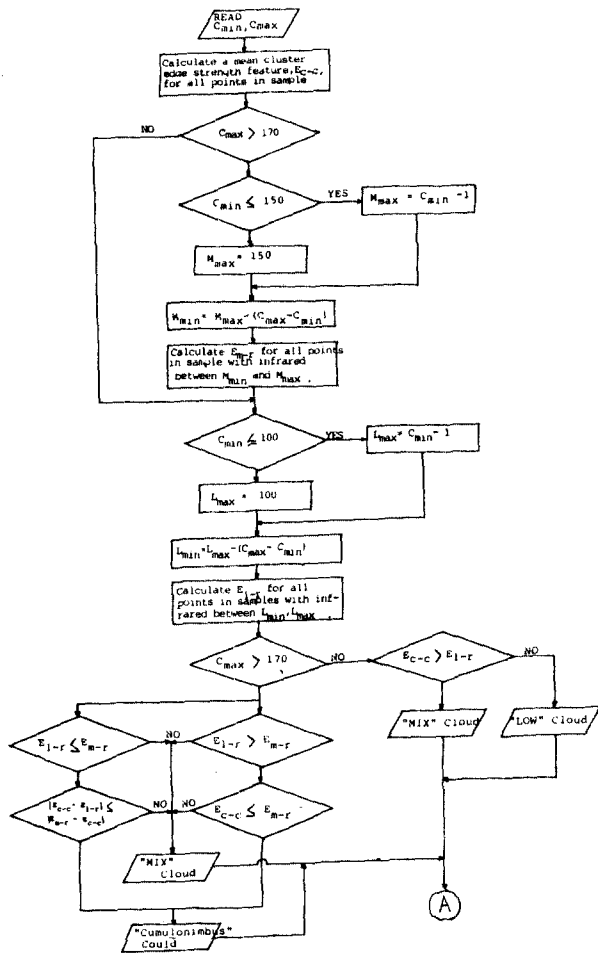
$$\begin{array}{cc}
 & 2 \times 2 \text{ 어레이} \\
 \begin{array}{|c|c|} \hline A & B \\ \hline C & D \\ \hline \end{array} \\
 \max(|A-D|, |B-C|)
 \end{array}$$

(제 4 단계)

집락은 하층운 부분과 권운 부분을 판별하기 위한 온도 임계값을 이용하여 병합된다. 분류공정에서 적란운으로 분류된 표본은 온도가 가장 낮은 집락으로 구성되고, 하층운 부분은 100 이하의 값을 갖는 모든 집락을 병합시킴으로써 얻는다. 또, 권운 부분은 적란운 표본에서 온도가 가장 낮은 집락을 제외한 100 이상의 값을 집락을 병합시킴으로써 얻을 수 있다.

IV. 실험 데이터

실험에 쓰여진 데이터는 1985년 10월 5일 00Z.



(그림 3) 분류 알고리즘

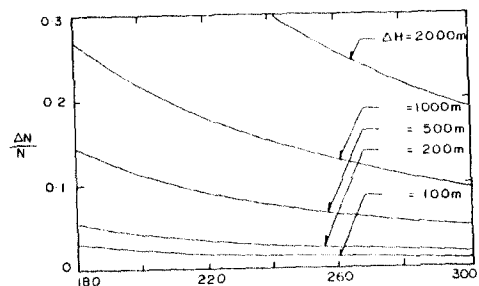
1986년 6월 13일 0903Z, 0950Z의 GMS-3 적외선 디지털 영상이다. 이 영상은 1434x1078의 분해능을 가지고 있으며 그레이 레벨은 0 - 255 까지이다. 한편, GMS-3는 고도 36000 Km. 순시 각속도는 0.14 mr이고 동작파장은 10.5 - 12.5 μm 의 열적외선이다.

가. 처리 영역

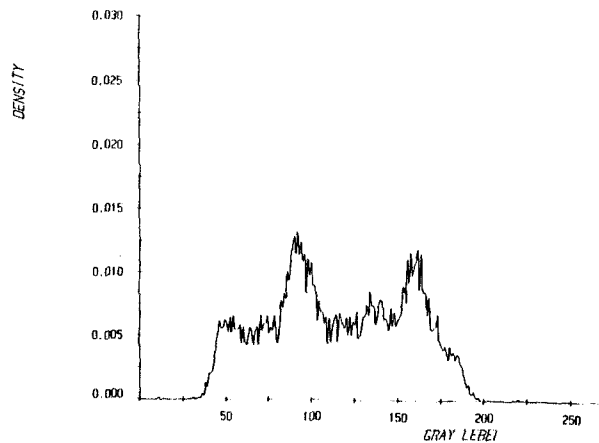
원래의 GMS-3 영상으로부터 128x128 단위로 데이터를 추출하여 다시 중복되지 않은 3x3 윈도우로 나누어 처리하였다.

나. 온도 추정 및 구름의 정의

기상위성 영상으로부터의 온도 추정은 적외선역 10.5 - 12.5 μm 에 의해 표면으로부터의 방사량을 측정하여 반사율과 대기의 흡수에 의한 오차를 보정하여 행한다. 구름의 온도가 결정되면 기상학적으로 산출되고있는 고도와 기온과의 상관으로부터 그 구름의 고도를 추정할 수 있다. 구름은 7°C 이하의 성분으로 정의되는데 이는 해면 온도 5°C 낮은 온도이다. 온도가 높은 지표에서는 방사량 N이 크기 때문에 같은 온도 분해능에 대하여 온도의 정도가 양호하게 되고 반면, 구름의 높이를 결정하는 온도 측정의 경우는 구름 자체의 온도가 낮아 N이 상대적으로 작으므로 온도의 분해능이 떨어진다. 본 연구에서 온도값의 산출은 다음 식에 준하였다(7). 그림 4.는 방사 측정 오차 대 온도와의 관계를 고도차 100, 200, 500, 1000, 2000M에서 계산한 것이다.



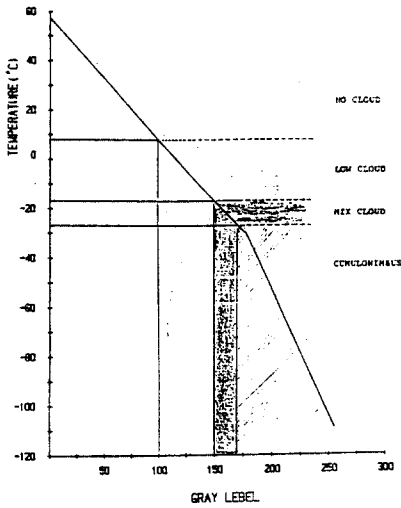
(그림 4) 고도차 2000, 1000, 500, 200, 100 M 에서 방사량 차와 온도와의 관계



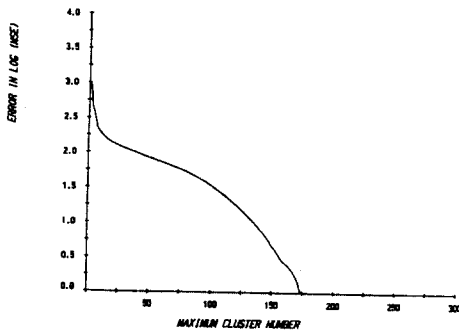
(그림 5) 그림 2. 영상의 히스토그램

$$T = 330 K - (C / 2), \quad C \leq 176 \text{ 일 때}$$

$$T = 418 K - C, \quad C \geq 176 \text{ 일 때} \quad (6)$$



(그림 6) 그레이 레벨과 온도와의 관계 및 구름의 정의



(그림 7) 그림 2의 영상에 적용한 Fisher 알고리즘에서 최대 집합수와 여러 V. 결론

(1) 하층운 형

하층운만이 일어나는 경우는 주로 해안에 접한 지역에서 잘 발생하였으며 적외선 구름 사진에서는 안개 혹은 층운 등 하층운의 식별이 용이하지 않으므로 이들을 모두 한 부류로 분류하였으나 이러한 문제는 가시역 사진에서는 쉽게 해결할 수 있다. 그렇지만 적외선 영상만을 이용하고자 하는 경우에는, 일반적으로 하층운은 온도가 높으며 온도에서 국지적 변화가 작다. 따라서 하층운 표본에서는 인접 데이터와의 온도상 변화가 작기 때문에 CONTRAST 나 ENTROPY, ASM 과 같은 텍스처 특징을

병행시키면 우수한 결과가 예상된다.

(2) 적란운 형

적란운은 발달 정도에 따라 온정고도가 다르고 모두 적란운의 온정 부분에서 흠어진 권운의 규모에 따라 다르게 나타난다. 일반적으로 적란운이 발달할수록 온정고도가 낮아지며 모두적란운의 온정부분에서 흠어진 권운이 길게 뻗는다. 적란운은 다른 유형의 구름보다 강우량을 크게 좌우하므로 이 유형의 추습이 기상학적으로 중요한 관심사가 되고 있다. 본 연구에서는 온도 특징과 텍스처 특징 Edge Strength E 하나만을 이용하였으나 엔트로피 특징을 부가하여 판별 로직을 구성시킨다면 온운, 권운, 또는 적란운 중의 하나로부터 하층운을 분리시키는 데 유용할 것이고, 적란운으로부터 온운을 분리하기 위한 유용한 특징 파라메터가 될 것이다.

참 고 문 헌

- [1] A.Rosenfeld and A.Kak, Digital Picture Processing: New York, Academic, 1976
- [2] S.W.Zucker, "Region growing: Childhood and adolescence," Computer Graphics Image Processing, vol. 5, pp. 382-399, 1976
- [3] J.Herkert, B.Romondi, Goddard, and W.Callcott, " An overview of GOES data flow and processing facilities," NOAA tech.memo NESS 64, pp. 2-20, Mar. 1975.
- [4] John A. Hartigan, Clustering Algorithms, John Wiley and sons, 1975.
- [5] R.O.Duda, P.E.Hart, Pattern Classification and Scene Analysis: New York, Wiley 1973.
- [6] Bonita G.Lee, R.T.Chin, and D.W.Martin, " Automated rain-rate classification of satellite images using statistical pattern recognition," IEEE Trans. on Geol. and Remote. Sensing, pp. 315-323, vol. 23, May 1985.
- [7] J.A.Parikh, "A comparative study of cloud classification techniques," Remote Sensing Environ., pp. 67-81, vol. 6, 1977.
- [8] C.Simmer, E.Raschke, and E.Ruprecht, "Cloud properties from two-dimensional histograms," IEEE 6th Int. Joint Confer. Pattern Recognition, pp. 529-532, 1982.