

## 슈퍼 픽셀기반 무인항공 영상 영역분할 및 분류

# Super-Pixel-Based Segmentation and Classification for UAV Image

김인규 · 황승준 · 나종필 · 박승제 · 백중환\*

한국항공대학교 정보통신공학과

In-Kyu Kim · Seung-Jun Hwang · Jong-Pil Na · Seung-Je Park · Joong-Hwan Baek\*

Information & Telecommunication Engineering, Korea Aerospace University, Gyeonggi-do, 412-791, Korea

### [요 약]

최근 무인항공기는 군사용뿐만 아니라 민간용으로도 많이 사용되고 있다. 무인항공기는 미리 입력된 좌표에 따라 GPS 정보를 이용하여 자동비행한다. 그러나 제밍이나 외부 교란에 의해 GPS 신호를 수신할 수 없으면 자동비행이 불가능 해진다. 이러한 문제를 해결하기 위한 한 방법으로, 본 연구에서는 무인기에 탑재된 카메라로부터 촬영된 영상으로부터 실시간으로 특정 영역을 검출하고 인식하는 알고리즘을 제안한다. 실시간 분류와 기계 학습에 사용할 특징을 추출하기 위한 전처리 과정으로 군집화 알고리즘인 그래프 기반 분할 알고리즘을 사용하여 슈퍼 픽셀화 하였다. 다양한 컬러모델 및 혼합 컬러 모델을 비교 분석하여 가장 이상적인 혼합 모델을 선정하고, 분류 알고리즘으로는 적은 트레이닝 데이터로도 뛰어난 분류 성능을 낼 수 있는 서포트 벡터 머신을 사용하였다. 무인항공 영상으로부터 18개의 컬러와 텍스처 특징 벡터를 추출하고 학습 및 예측과정을 통해 하천, 비닐하우스, 논 등 3 종류의 영역을 실시간으로 분류하였다.

### [Abstract]

Recently UAV(unmanned aerial vehicle) is frequently used not only for military purpose but also for civil purpose. UAV automatically navigates following the coordinates input in advance using GPS information. However it is impossible when GPS cannot be received because of jamming or external interference. In order to solve this problem, we propose a real-time segmentation and classification algorithm for the specific regions from UAV image in this paper. We use the super-pixels algorithm using graph-based image segmentation as a pre-processing stage for the feature extraction. We choose the most ideal model by analyzing various color models and mixture color models. Also, we use support vector machine for classification, which is one of the machine learning algorithms and can use small quantity of training data. 18 color and texture feature vectors are extracted from the UAV image, then 3 classes of regions; river, vinyl house, rice field are classified in real-time through training and prediction processes.

**Key word** : Super pixel, Image segmentation, Image classification, UAV, Clustering.

<http://dx.doi.org/10.12673/jant.2014.18.2.151>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 10 April 2014; Revised 20 April 2014  
Accepted (Publication) 18 April 2014(30 April 2014)

\*Corresponding Author; Joong-Hwan Baek

Tel: +82-2-300-0125

E-mail: [jhbaek@kau.ac.kr](mailto:jhbaek@kau.ac.kr)

## I. 서론

무인항공기(unmanned aerial vehicle)란 조종사가 탑승을 하지 않고 지상 통제소에서 통신을 통해 비행체를 조정하거나 미션에 따라 자동 비행을 하는 항공기를 말한다. 그렇기 때문에 조종사의 생명에 위협이 되는 군사 목적으로 많이 활용되고 있다. 적외선 및 광학 카메라, 혹은 레이더를 갖추어서 적의 위치나 지형을 정찰하거나 혹은 특정 목표물을 공격하는 용도로 많이 사용되고 있다.

그리고 현재에는 군사목적뿐만 아니라 민간분야에서도 다양한 용도로 활용하고 있다. 특히 카메라로부터 영상을 획득하고 처리를 함으로써 산불을 감시하고 예방하는 것뿐만 아니라, 도로의 교통 상황을 감시하여 교통정보 방송에 활용되거나 송유관, 송전선, 도로 등 선형적으로 연결된 지형지물을 감시하여 문제점을 찾아내는 목적으로 많이 사용되고 있다[1]-[3]. 이처럼 비전기반 무인항공기 영상처리의 중요성이 커지고 이 분야에 대한 많은 기술이 개발 및 연구 되고 있다[4][5].

무인항공기가 선형적인 지형지물을 따라 자동비행을 통해 감시하기 위해서는 위성 항법 장치(GPS; global positioning system) 정보를 사용하여야 한다. 하지만 재밍(jamming)이나 외부 교란으로 위성신호를 수신할 수 없으면 GPS 정보를 사용할 수 없으므로, GPS 정보를 배제한 상황에서도 지속적인 감시 및 정찰을 위해서는 다른 종류의 센서를 사용하여야 한다.

본 연구에서는 비전기반 무인항공기가 비전 센서를 사용하여 GPS 신호를 수신할 수 없을 때를 대비하여 특정영역을 분할 및 분류하는 기법을 다루고 있다. 비전 기반 무인항공기에서 중요한 것은 실시간 영상처리가 가능해야 한다는 것이다. 하지만 영상 전체에 대해서 기계학습 알고리즘을 적용시켜서 분류하기에는 연산 량이 많아지므로 실시간으로 영상처리 하기가 힘들어진다. 그렇기 때문에 군집화(clustering)과정으로 그래프 기반 이미지 분할(graph-based image segmentation) 알고리즘을 사용하여 슈퍼 픽셀(super-pixels)을 만들었다. 슈퍼 픽셀을 만드는 과정에서는 컬러 모델을 기반으로 하였다. 일반적으로 가장 많이 사용되는 컬러 모델로는 RGB(red green blue) 컬러 모델, HSV(hue saturation value)컬러 모델 등이 있지만 더 세밀하게 군집화 하기 위해서 본 논문에서는 Lab 컬러 모델과 XYZ 컬러 모델을 혼합하여 5채널 Labxy 혼합 컬러 모델을 사용하였다.

분류 과정에서는 비행영상으로부터 트레이닝 데이터(training data)를 수집해야 하기 때문에 개수가 제한되어 있고, 이런 경우에는 기계학습 알고리즘 중 서포트 벡터 머신(support vector machine)이 좋은 결과를 보여준다. 그리고 2개 이상의 N-클래스를 분류를 해야 하기 때문에 본 논문에서는 서포트 벡터 머신을 사용하여 분류 하였다. 선형적인 지형지물인 하천과 특정 영역인 비닐하우스, 그리고 논으로 3-클래스로 분류 하였다. 트레이닝 및 분류에 사용한 특징 벡터로는 컬러와 텍스처로 나뉜다.

## II. 군집화(Clustering)

### 2-1 분할(Segmentation) 알고리즘

이미지 분할 알고리즘이란 영상의 특정 영역을 정확히 구별해 내는 알고리즘이다. 앞/뒤 프레임의 차이를 이용하여 움직임이 있는 오브젝트(object)를 분할하거나, 군집화 과정을 통해 분할하는 등 여러 가지 알고리즘이 존재하고 있다[6].

본 연구에서는 실시간 영상처리를 위한 전처리 과정으로 영상을 군집화하여 슈퍼 픽셀을 만드는 것에 더 초점을 두었다. 군집화 알고리즘에서 많이 쓰이고 있는 알고리즘은 K-평균 알고리즘이다. 하지만 K-평균 알고리즘은 처음에 K개의 군집을 결정해야 하는 단점이 있다. 무인항공기로부터 실시간으로 들어오는 영상이므로 항상 같은 개수의 군집을 정할 수가 없으므로 그래프 기반 이미지 분할 알고리즘을 사용하였다.

#### 1) 그래프 기반 이미지 분할

그래프 기반 이미지 분할 알고리즘은 영역(혹은 픽셀) 간 관계의 차이점을 결정하고 병합하는 알고리즘이다. 그래프 기반 이미지 분할 알고리즘은 최소비용 신장 트리 알고리즘을 기초로 한다. 신장 트리는 그래프의 모든 꼭지점을 포함하면서 순환되지 않는 방식으로 연결되어 있는 최소 서브 그래프를 의미한다. 최소비용 신장 트리는 신장 트리의 각 연결선에 가중치를 부여 했을 때, 가중치의 총합이 가장 적은 신장 트리를 의미하며 그래프 내에 있던 원래 연결선들만을 포함하고, 꼭지점의 개수가 n개이면 n-1의 연결선을 가지게 된다.

영역 R에 대한 내부 차이는 최소비용 신장 트리의 가중치가 가장 큰 연결선으로 식(1)과 같이 정의한다. 꼭지점이 연결된 한 개 이상의 연결선을 가지고 있는 서로 다른 영역의 차이는 두 영역의 가장 작은 가중치를 가진 연결선으로 식(2)로 정의된다. 그 후, 최소 내부 차이와 비교를 통해 영역들의 병합을 결정한다. 최소 내부 차이는 식(3),(4)을 이용하여 구한다. 여기서  $\tau(R)$ 은 병합 결정 문턱치 값으로써, k는 상수 파라미터(constant parameter)이다[7].

$$(R) = \max_{e \in MST(R)} w(e), \quad (e \in MST(R)) \quad (1)$$

$$Dif(R_1, R_2) = \min_{(e = (v_1, v_2) | v_1 \in R_1, v_2 \in R_2)} w(e) \quad (2)$$

$$MInt(R_1, R_2) = \min(\int (R_1) + \tau(R_1), \int (R_2) + \tau(R_2)) \quad (3)$$

$$\tau(R) = k / |size(R)| \quad (4)$$



(a) 원 영상 (b) 블러링 영상

그림 1. 원본영상 과 블러링 영상

Fig 1. Origin image and blurring image.

## 2) 슈퍼 픽셀 알고리즘

그래프 기반 분할 알고리즘을 통해 군집화를 하여 슈퍼 픽셀화 하였다. 군집과정에서 사용되는 가중치는 각 픽셀의 컬러를 사용하였다. 전처리 과정으로 영상축소와 블러링(blurring)을 하였다. 영상축소의 경우는 연산 량을 줄어든 채 함으로써 실시간 영상처리를 가능하게 한다[8]. 비행영상으로부터 들어오는 영상은 자연 현상물이 많기 때문에 픽셀간 거리 수치의 변화량이 크다. 변화량이 크면 같은 영역으로 군집하는 과정에서 문제가 발생할 수 있기 때문에 영상을 축소한 후 3X3 평균필터를 사용하여 블러링 하였다. 그 결과를 그림 1에서 볼 수 있다.

이후 슈퍼 픽셀화를 진행한다. 알고리즘은 표 1에 요약 하였다. Case 1에서는 두 픽셀 모두 영역에 포함되어있지 않기 때문에 크기가 1이 되고, 그렇기 때문에 문턱치 값은 k이다. 과정 6은 후처리 과정으로 너무 작은 군집된 영역들은 병합하는 작업을 하며, min(size)는 파라미터 값이다.

표 1. 슈퍼 픽셀화 알고리즘

Table 1. Super-pixels algorithm.

1. n개의 모든 픽셀의 8-방향 픽셀에 관한 거리 $d_1, \dots, d_m$ 을 구한다.
2. 픽셀간 거리를 가중치로 하여 낮은 순으로 두 개의 픽셀의 연결선 $e_1, \dots, e_m$ 로 정렬한다. 여기서 $e = (P_i, P_j)$ 이다.
3. 과정 4를 $q = 1, \dots, m$ 만큼 반복한다.
4. 군집된 영역을 R이라 한다면, 다음의 Case에 따라 실행한다.
Case1: $P_i \notin R_i^{q-1}, P_j \notin R_j^{q-1}$ 이고, $w(e_q) \leq k$ 이면 새로운 군집된 영역 $R_k^q = (P_i, P_j)$ 을 생성한다.
Case2: $P_i \in R_i^{q-1}, P_j \notin R_j^{q-1}$ 이고, $w(e_q) \leq \text{Min}(R_i^{q-1}, e_q)$ 이면 군집된 영역 $R_i^{q-1}$ 에 $P_j$ 를 추가한다.
Case3: $P_i \in R_i^{q-1}, P_j \in R_j^{q-1}$ 이고, $w(e_q) \leq \text{Min}(R_i^{q-1}, R_j^{q-1})$ 이면 병합한다.
5. 과정 4의 결과인 군집된 영역들의 집합 $S = (R_1, \dots, R_r)$ S = ()라 한다면 과정 6을 $q = 1, \dots, r$ 만큼 반복한다.
6. $R^q \leq \text{min}(size)$ 이면 $R^{q+1}$ 과 병합한다.

## 2.2 슈퍼 픽셀 기반 이미지 분할

RGB컬러 모델이나 HSV컬러 모델은 단일 모델이기 때문에 컬러 기반으로 슈퍼 픽셀화 했을 때, 하천의 끝부분이나 비닐하우스의 경계 부분이 완벽하게 되지 않아 Lab 컬러 모델과 XYZ 컬러 모델의 혼합으로 일반 3차원 컬러 모델이 아닌 5차원의 Labxy 컬러 모델을 사용하여 슈퍼 픽셀화 하였다.

Lab 컬러 모델의 L은 밝기인 명도(luminosity)를 말하며 a조합은 녹색에서 적색의 보색, b조합은 황색에서 청색의 보색을 뜻한다. XYZ 컬러 모델은 RGB 컬러 파장으로 표현할 수 없는 색이 존재하여서 표준 색 관찰자로 만든 컬러 모델이다. X, Y, Z는 각각 적색, 녹색, 파란색과 비슷한 색깔을 가진다. RGB 컬러 모델에서 XYZ 컬러 모델로 변환하는 식은 (5)와 같다. Lab 컬러 모델은 식(6)-(11)을 이용하여 XYZ 컬러 모델로부터 계산된다. Labxy 컬러 기반의 거리는 식 (12)-(14)로 구한다. 식(12)에서 i는 현재 픽셀을 의미하고 j는 i픽셀 주변에 있는 픽셀이며, Lab 컬러모델에 관련된 거리이다. 식(13)은 XY 컬러의 거리이며, 최종 거리는 식(14)와 같고, 결과화면은 그림 2와 같다.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.242671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (5)$$

$$\text{(if, } Y > 0.008856) L = 116 \times Y^{1/3} - 16 \quad (6)$$

$$\text{(if, } Y \leq 0.008856) L = 903.3 \times Y \quad (7)$$

$$a = 500(f(X) - f(Y)) + 128 \quad (8)$$

$$b = 200(f(X) - f(Z)) + 128 \quad (9)$$

$$\text{(if, } t > 0.008856) f(t) = t^{1/3} \quad (10)$$

$$\text{(if, } t \leq 0.008856) f(t) = 7.787t + 16/116 \quad (11)$$

$$d_{Lab} = \sqrt{(L_i - L_j)^2 + (a_i - a_j)^2 + (b_i - b_j)^2} \quad (12)$$

$$d_{xy} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2} \quad (13)$$

$$d_{Labxy} = d_{Lab} + d_{xy} \quad (14)$$

하천 영역이나 비닐 하우스 영역을 보면 RGB 컬러 모델과 HSV 컬러 모델보다 더 정확하게 군집화 되어 있는 것을 확인할 수 있다. 본 논문에서는 하천과 논, 그리고 비닐하우스를 분류하기 때문에 하천 영역과 비닐하우스 영역은 정확히 분류가 되어야 한다.



(a) 원영상 (b) k = 50 (c) k = 300

그림 2. 슈퍼 픽셀화 결과 영상

Fig 2. Result Images based on super-pixels.

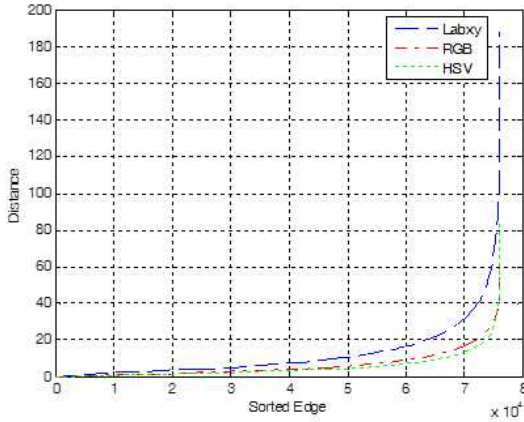


그림 3. 낮은 순으로 정렬된 에지  
Fig 3. Edges sorted in descending order.

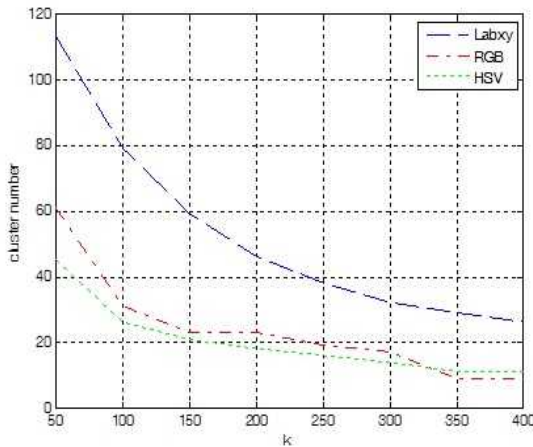


그림 4. 상수 k값에 따른 군집 영역 개수  
Fig 4. Cluster numbers according to constant k.

그림 3에서 말하는 에지(edge)란 한 픽셀과 그 주변에 있는 픽셀의 연결선선을 의미한다. Labxy 컬러의 경우는 두 개의 거리를 더함으로써 구해지기 때문에 RGB 컬러 모델과 HSV 컬러 모델보다 변화량이 더 큰 것을 확인할 수 있다. 변화량이 크다는 것은 군집화 과정에서 하천이나 비닐하우스 영역에 대해 더 완벽하게 분할된다는 것을 의미한다.

그림 4를 보면 상수 k값이 커짐에 따라 클러스터의 개수가 줄어드는 것을 확인할 수 있다. 그리고 같은 상수 k값을 가진다면 Labxy 컬러 모델이 더 많은 군집 영역 개수를 가지고 있는 것을 확인할 수 있다. 군집 영역 개수가 많은 것은 더 세밀하게 나누어졌음을 의미한다.

III. 분류(Classification)

3-1 서포트 벡터 머신(Support Vector Machine)

서포트 벡터 머신은 블라디미르 베프닉(vladimir vapnik)과 그의 AT&T Bell 연구소 팀이 개발한 식별 방법으로 최근 몇 년 동안에 이론적인 발전뿐만 아니라, 실제 구현되어 데이터 마이닝(data mining)분야는 물론 얼굴 인식과 같은 패턴인식 응용분야에도 널리 사용되고 있다.

서포트 벡터 머신은 교사 학습(supervised learning)을 통해 트레이닝 데이터들을 토대로 최적의 클래스를 나눌 수 있는 초평면(hyperplane)을 찾는다. 이 초평면에 가장 가까운 트레이닝 데이터가 서포트 벡터이다. 식(15)를 활용하여 초평면으로부터 한 점 간의 거리를 구한다.

$$d = \frac{w^T x + b}{\|w\|} \tag{15}$$

w는 가중치 벡터(weight vector)이고 b는 바이어스이다. 선형 분류 함수의 일차 직선을 데이터 점과 만날 때까지 확장한 폭을 선형 분류기의 마진(margin)이라고 정의하면, 마진의 값이 최대가 되는 마진의 중심을 이등분하는 초평면을 최적 분류 초평면(optimal separating hyperplane)이라고 정의한다. 최대 마진을 선택할 경우, 분류 경계 부근에서 발생하는 작은 오차가 최소화된다. 마진의 최대화를 구하는 과정이 바로 서포트 벡터 머신의 학습 과정이다.

3-2 특징 벡터(Feature Vector)

기계학습 알고리즘에서 특징 벡터의 결정은 매우 중요한 부분이다. 어떠한 특징 벡터를 선택하느냐에 따라 분류의 성능이 결정지어지기 때문이다. 본 연구에서는 9개의 컬러 특징과 9개의 텍스처 특징을 사용하여 1x18의 크기를 가지는 특징 벡터를 사용하였다.

1) 컬러 특징

컬러 특징은 군집 영역의 RGB 컬러 모델의 평균과 Lab 컬러 모델의 평균, 그리고 XYZ 컬러 모델의 평균을 사용하였다.



(a) 논 (b) 하천 (c) 비닐하우스

그림 5. 트레이닝 데이터 영역 샘플  
Fig 5. Training data for region sample.

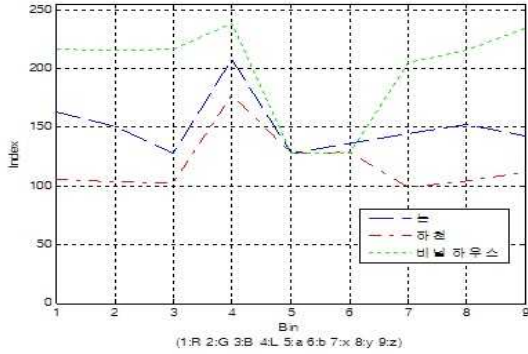


그림 6. 데이터 영역의 컬러 평균  
Fig 6. Color mean of data region.

그림 6은 그림 5의 영역들의 각각의 컬러 평균값을 의미한다. 슈퍼 픽셀 과정에서 Lab 컬러와 XYZ 컬러를 사용하여 군집화 하였기 때문에 Lab 컬러와 XYZ 컬러를 추가 하였다.

2) 텍스처 특징

텍스처란 영역이 가지고 있는 질감, 무늬의 모양이나 패턴 등을 말한다. 일정한 패턴의 무늬의 경우에는 2차원 푸리에 변환(fourier transform)을 하여 주파수 축으로 변경하여 패턴의 방향성 및 크기 등을 알 수가 있다. 하지만 본 연구에서 3개의 클래스 중 2개인 논, 하천의 경우에는 인위적으로 만든 것이 아닌 자연물이기 때문에 일정한 패턴이 존재할 확률이 적다. 그리고 인위적으로 만든 비닐하우스의 경우에는 패턴이 존재한다고 해도 비행의 방향에 따라서 패턴의 방향성이 달라지기 때문에 주파수 축으로 변환하는 방법은 사용할 수가 없다.

따라서 소벨 필터(sobel filter)를 사용하여 그래디언트 맵(gradient map)을 생성하고 각각의 픽셀의 방향성은 배제하고 식(16)을 이용하여 매그니튜드(magnitude)에 대해서만 히스토그램화 하였다.

$$Magnitude = \sqrt{(xsobel_{(x,y)})^2 + (ysobel_{(x,y)})^2} \tag{16}$$

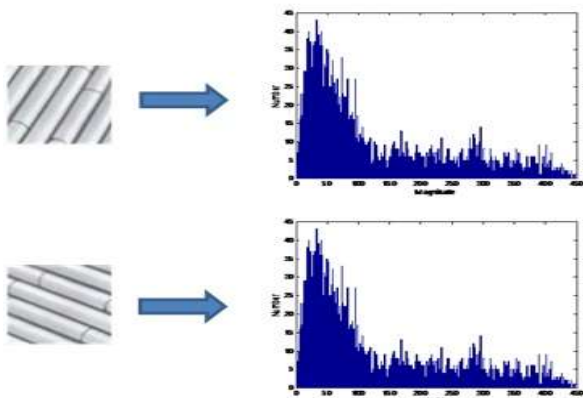
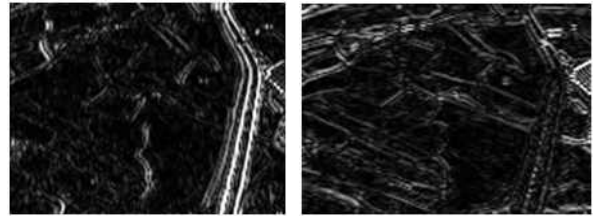
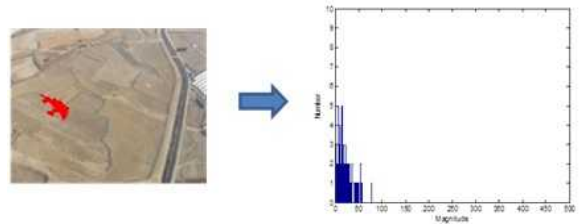


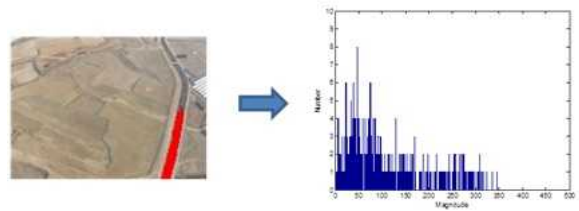
그림 7. 회전에 강인한 매그니튜드 히스토그램  
Fig 7. Rotation invariant magnitude histogram.



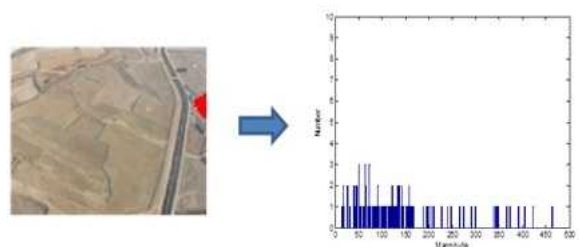
(a) x방향 미분  
(b) y방향 미분  
그림 8. 그래디언트 맵  
Fig8. Gradient map.



(a) 논의 매그니튜드 히스토그램



(b) 하천의 매그니튜드 히스토그램



(c) 비닐 하우스의 매그니튜드 히스토그램

그림 9. 3개 클래스의 매그니튜드 히스토그램  
Fig 9. Magnitude histogram of 3-classes.

그림 7은 65x65 비닐하우스 영상을 90°로 회전하여도 같은 히스토그램을 가지는 것을 보여주고 있다. 실제 비행 영상을 소벨 필터를 적용시키고 3개의 클래스의 매그니튜드 히스토그램은 그림 8 과 같다. 그림 9의 결과를 보면 두 가지의 문제점이 발생한다. 첫 번째 문제는 영역의 크기가 다르기 때문에 히스토그램이 일정하지 않다. 그렇기 때문에 정규화 과정이 필요하여 식(17)과 같이 진행하였다. H는 히스토그램을 나타내고 N은 빈(bin)의 개수이다.

$$\|H\|_{L1} = \sum_N abs(H_N) = 1 \tag{17}$$





**김 인 규 (In-Kyu Kim)**

2011년 2월 : 한국항공대학교 정보통신공학 (공학사)  
2013년 2월 ~ 현재: 한국항공대학교 정보통신공학 (공학석사)  
※관심분야 : 영상처리, 멀티미디어 3D VISION



**황 승 준 (Seung-Jun Hwang)**

2012년 2월 : 한국항공대학교 정보통신공학 (공학사)  
2014년 2월 : 한국항공대학교 정보통신공학 (공학석사)  
2014년 3월 ~ 현재 : 한국항공대학교 정보통신공학 박사과정  
※관심분야 : 영상처리, 패턴인식



**나 종 필 (Jong-Pil Na)**

2013년 8월 : 한국항공대학교 정보통신공학과 (공학사)  
2013년 9월 ~ 현재 : 한국항공대학교 정보통신공학 석사과정  
※관심분야 : 영상처리, 물체추적



**박 승 제 (Seung-Je Park)**

2014년 2월 : 한국항공대학교 정보통신공학과 (공학사)  
2014년 3월 ~ 현재 : 한국항공대학교 정보통신공학 석사과정  
※관심분야 : 영상처리



**백 중 환 (Joong-Hwan Baek)**

1981년 2월 : 한국항공대학교 항공통신공학 졸업(공학사)  
1987년 7월 : 오클라호마주립 대학원 전기 및 컴퓨터공학 (공학석사)  
1991년 7월 : 오클라호마주립대학원 전기 및 컴퓨터공학 (공학박사)  
1992년 ~ 현재 : 한국항공대학교 항공전자 및 정보 통신공학부 교수  
※관심분야 : 영상처리, 패턴인식, 영상압축, 멀티미디어