

주성분분석을 이용한 적조생물 영상 인식

박선*, 이경효*, 이성로*, 임양섭**, 신준우**, 권장우**, 박석천***

*목포대학교

**정보통신산업진흥원

***경원대학교

e-mail:sunpark@mokpo.ac.kr

Red Tide Image Recognition using PCA

Sun Park*, Gyeonghyo Lee*, Seong Ro Lee*, Yangseop Lim**, Jun Woo

Sin**, Jangu Kwuan**, Seokcheon Park***

*Mokpo National University

**NIPA

***Kyungwon University

요 약

적조의 피해가 증가함에 따라서 적조에 대한 많은 연구가 이루어지고 있으나 자동으로 적조류를 분류하는 연구는 미흡한 실정이다. 적조류는 영상 객체를 일치 할 수 있는 기준 중심 특징이 없기 때문에 인식이 어렵다. 이 때문에 기존에 연구들은 단순히 몇 종류의 적조류 만을 분류에 이용하고 있다. 본 논문은 주성분분석과 영상 객체의 원형율을 이용한 새로운 적조류 인식 방법을 제안한다.

1. 서론

적조(red tide)란 플랑크톤의 대량 번식으로 바다물의 색깔이 변하는 자연 현상을 말하는 것이었으나, 최근에는 적조 피해가 늘어나면서 적조를 유해조류의 대번식이란 의미로 사용하고 있다. 국내에서는 연안 산업화의 진행으로 적조현상이 증가하였으며, 90년대 이래로 양식장 등 수산업에 큰 피해를 주고 있다. 적조의 수산 피해를 최소화하기 위하여서 국립수산과학원에서는 선박, 육상, 항공 감시를 통하여 적조예보를 발령하고 있다[1].

국내의 대부분 적조 감시는 사람의 손에 의존하고 있기 때문에 많은 시간과 예산이 소모되고 있다. 이 때문에 자동화된 효율적인 적조정보시스템이나 분류 방법 등이 필요하다. 적조류에 대한 연구를 진행하기 위해서는 적조류의 종류와 적조생물 밀도(개체/ml)를 정확히 파악해야 한다. 그러나 국내에서의 적조류 분류 및 개체 수 파악에는 대부분 수작업에 의존하고 있다. 적조류의 개체 수를 파악에는 센서를 이용하여 자동화 방법[2]에 대하여서 어느 정도 진행되고 있으나 수집된 적조생물 영상자료로부터 적조류를 분류하는 연구는 아직 미흡한 편이다. 이는 국내의 연안 및 근해역에서는 60종 이상의 적조류가 발생하고 있으며, 이들의 형태 또한 수만곡이나 다양한 모양을 보이고 있어서 자동화된 분류방법을 이용하는 데에 제한 사항으로 작용하고 있다. 특히 적조류 영상 간에 일치시킬 수 있는 기준점이 없어서 적조류 영상 인식을 더욱 어렵게 하고 있다. 예를 들면, 사람의 얼굴인식은 어느 정도 동일한 형태의 특징인 눈, 코, 입, 귀 등 영상 간에 일치시킬 수 있는 기준 특징들을 가지고 있다. 그러나 적조류는 너무 다양한 형태를 보이기 때문에 이러한 기준 특징들을 파악하기 힘들다.

본 논문에서는 국내 해역에서 발생하는 적조류를 자동으로 분류하기 위한 방법에 대하여서 연구한다. 적조류 자동 분류를 위하

여서 본 논문의 제안한 방법은 주성분분석(PCA; principal component analysis)방법을 이용하여서 63종의 적조류를 학습하고, 적조류 영상 간에 일치시키기 위하여 영상 객체의 원형율을 이용하여 기준 특징을 추출하였으며, 영상 객체의 엔트로피를 이용하여 정확한 적조류 영상을 인식한다.

2. 관련연구

2.1 적조 영상 인식

Jiang의 저자들은 베이지안 분류자와 SVM을 이용한 적조류 분류방법을 제안하였다. 이들의 방법은 베이지안 분류자를 이용하여서 오염되거나 알려지지 않은 적조류를 제거하고, SVM을 이용하여서 적조류를 분류한다[3]. 그러나 알려지지 않은 적조류나 오염된 객체를 인식하는 데에는 좋은 성능을 보이지 않는다. 또한, 적조 분류의 학습에는 8종의 적조류만 사용하였다. Jiang의 저자들은 적조류 인식을 위하여 2단계 분류자를 사용한 방법을 제안하였다. 첫 번째 분류자인 SVDD(support vector data description)는 오염된 객체나 알려지지 않은 적조류를 제거하는 데에 사용하며, 두 번째 분류자인 SVM(support vector machine)은 적조류를 미리 분류된 분류 계층으로 분류한다[4]. 이들의 방법은 우수종인 8종의 적조류만 학습에 이용하였다. Lili의 저자들은 적조류 분류를 위하여 SVM과 FCM(fuzzy c-means)군집방법을 이용하였다. 이들 방법은 먼저 SVM을 이용해서 검정 표본에 분류 표시를 하고, 확률 방법으로 잘못 분류된 자료를 FCM을 이용하여서 다시 분류 표시한다[5]. 이들 방법 역시 5종의 적조류를 학습하는 데에 이용하였다. 이들 방법[3, 4, 5]들은 소량의 학습된 적조류 이외의 적조류를 인식하는 데에 는 어려움을 가지고 있다.

2.2 주성분 분석

주성분 분석(PCA; principal component analysis)은 상관관계가 높은 자료를 선형 변환하여 고차원의 정보를 유지하면서 저차원으로 차원을 축소하거나 중복을 제거하는 통계적 방법이다. 주성분 분석은 주로 신호처리, 특징 추출, 자료의 차원 감소, 영상 인식 등 많은 분야에서 사용된다. 특히 영상 인식 분야에서는 영상 자료의 특징 추출에 매우 효율적인 방법의 하나이다[6].

주성분 분석을 이용하여서 영상으로부터 고유벡터를 얻는 방법은 다음과 같다. 하나의 영상은 $N^2 \times 1$ 벡터로 표현되며, 학습 영상은 M 개의 영상 X_1, X_2, \dots, X_M 에 대한 평균으로 표현된다. M 개의 영상에 대한 평균은 식(1)과 같다.

$$\bar{X} = \frac{1}{M} \sum_{i=1}^M X_i \quad (1)$$

각 영상 X_i 는 식(1)에서 계산된 영상의 평균과의 차인 편차 $\Phi = X_i - \bar{X}$ 를 구함으로써 일반화한다. 영상의 편차로 구성된 행렬 A 에 A^T 를 곱하여서 식(2)과 같이 공분산행렬을 계산한다.

$$C = \frac{1}{M} \sum_{n=1}^M (X_n - \bar{X})(X_n - \bar{X})^T \quad (2)$$

공분산 행렬은 고유값과 고유벡터를 구하는 데에 이용한다. 즉, 영상의 공분산 행렬은 주성분 분석에 의해 다음 식(3)과 같이 나타낼 수 있다. 그리고 이로부터 고유벡터 e_i 와 고유상수 λ_i 를 구할 수 있다.

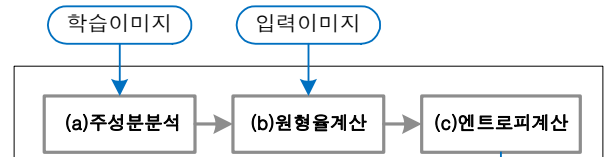
$$C = UAU^T \quad (3)$$

여기서 U 는 공분산 행렬의 고유벡터 e_i 를 고유값에 따라서 정렬한 $U = [e_1, e_2, \dots, e_M]$ 이고, U^T 는 U 의 전치행렬이다. A 는 고유상수 λ_i 를 i 번째 열의 원소이고 모든 비 대각 원소가 0인 대각 행렬이다. 이렇게 구해진 고유공간에 평균을 뺀 영상을 입력하여 투영시키면 식(4)과 고유벡터 영상 인식에 사용되는 가중치 벡터 W 를 구할 수 있다[6].

$$W = U^T \Phi \quad (4)$$

3. 제안방법

본 논문에서 제안한 적조류 인식 과정은 그림1과 같이 주성분 분석, 원형율계산, 엔트로피계산 단계로 구성된다. 주성분분석단계에서는 국내 해역에서 발생하는 적조류 영상에 주성분분석을 적용하여서 학습한다. 원형율계산 단계에서는 입력영상과 학습영상의 원형율을 계산하고, 이 원형율에 가장 근접한 영상을 선택하여서 후보 인식 영상 집단을 구성한다. 엔트로피계산단계에서는 구성된 후보 인식 영상 집단의 각각의 영상의 엔트로피를 계산하여서 가장 엔트로피가 유사한 영상을 인식영상으로 선택한다.



(그림 1) 제안된 적조류 인식 블록도

3.1 주성분분석에 의한 적조류 영상 학습

이 장에서는 주성분분석을 이용하여 국내 해역에서 발견되는 63종의 적조류를 학습한다. 학습에 사용되는 63종의 학습 영상은 국립수산물품질관리원 적조정보 홈페이지[1]에서 가져온 자료이다.

그림1(a)과 같이 주성분분석에 의한 적조류 영상 학습방법은 다음과 같다. 첫 번째 단계는 학습영상의 벡터집합을 구성한다. 이때 각 영상의 크기를 115×95 픽셀로 고정하였으며, 대부분의 영상은 중앙에 중심이 일치하도록 하였다. 두 번째 단계에서는 배경과 빛에 의하여 발생하는 에러를 줄이기 위해 평균과 분산을 기준으로 모든 영상을 정규화 한다. 세 번째 단계에서는 식(1)을 이용해서 평균 영상 벡터를 계산하고, 계산된 평균 영상 벡터와 식(2)을 이용하여 학습 영상 벡터집합의 공분산 행렬을 계산한다. 네 번째 단계에서는 식(3)을 이용하여 주성분을 계산한다. 마지막 단계에서는 식(4)과 입력 영상을 이용하여 적조류를 인식할 수 있다.

3.2 원형율계산에 의한 인식 후보 적조류 선택

주성분분석만을 이용하여 적조류 영상을 인식할 경우에는 적조류 영상에 대한 중심 기준점이 없기 때문에 영상에 객체가 위치하는 장소나 각도에 따라서 전혀 다른 영상을 인식하는 문제점을 가지고 있다. 이러한 문제점을 해결하기 위해서 본 논문에서는 그림1(b)과 같이 적조류 객체의 원형율을 계산하여 입력 영상에 가장 근접한 원형율을 선택하는 방법을 제안하였다. 즉 영상 객체가 얼마나 원형모양에 가까운가를 판단하여 유사한 영상으로 분류한다. 본 논문에서는 원형모양에 가까울수록 원형율이 1의 값을 가지며, 수평에 가까울수록 0의 값을 가진다. 다음 식(5)은 영상 객체의 원형율을 계산하는 식이다[7].

$$R(I) = \frac{4 \times \pi \times area}{perimeter^2} \quad (5)$$

여기서 I 는 115×95 행렬로 구성된 그레이 영상이고, π 는 원주율로 3.1416, $area$ 는 원형율을 구하고자하는 객체의 면적, $perimeter$ 는 객체의 주변길이다.

3.3 엔트로피 계산에 의한 적조류 인식

입력 영상에 대한 원형율을 계산하더라도 서로 비슷한 원형율 값을 가지는 적조류 영상이 있다. 이러한 인식 후보 적조류로부터 입력 영상에 일치하는 적조류를 인식하기 위하여 본 논문에서는 그림1(c)과 같이 각 후보 적조류 영상의 엔트로피를 계산하여 입력 영상의 엔트로피에 가장 가까운 적조류 영상을 인식한다. 각 영상에 대한 엔트로피 계산은 식(6)과 같다. 본 논문에서는 엔트로피계산 이전에 영상 객체를 흑백영상으로 변환하여 잡음을 제거하였다.

$$Entropy(I_{bw}) = -\sum_{i=1}^m p_i \log_2(p_i) \quad (6)$$

여기서 I_{bw} 는 115×95 행렬로 구성된 흑백 영상이고, p 는 흑백영상 I_{bw} 의 히스토그램으로 나타 낼 때에 히스토그램의 수이다[7, 8].

4. 결론

유해조류인 적조의 피해가 증가함에 따라서 국내에서는 적조에 대한 많은 연구가 이루어지고 있다. 그러나 현재까지는 자동으로 적조류를 인식하는 연구는 미흡한 실정이다. 본 논문은 이러한 문제를 해결하기 위하여서 주성분분석과 영상 객체의 원형율을 이용한 새로운 적조류 인식 방법을 제안한다. 제안된 방법은 주성분분석을 이용하여 학습영상을 고유벡터로 만들고, 각 학습영상과 입력영상의 원형율을 계산하여 입력 영상에 가장 근접한 영상을 선택하여 후보 영상을 선택한다. 선택된 후보 영상의 엔트로피를 계산하여 엔트로피가 가장 일치하는 영상을 최종 영상으로 선택한다. 제안방법은 엔트로피와 원형율을 이용하여 적조류 영상의 중심기준점이 없이도 영상을 인식할 수 있다.

Acknowledgment

“본 연구는 지식경제부 및 정보통신산업진흥원의 대학 IT연구센터 지원사업의 연구결과로 수행되었음” (NIPA-2011-C1090-1121-0007), 이 논문은 2010년도 정부(교육과학기술부)의 재원으로 한국연구재단의 대학중점연구소 지원사업으로 수행된 연구임(2010-0028295)

참고 문헌

- [1] 국립수산과학원 적조정보 홈페이지
<http://portal.nfrdi.re.kr/redtide/index.jsp>
- [2] M. Turk, A. Pentland, "Eigenfaces for Recognition", *Journal of cognitive Neuro-science*, 3(1), pp.71-86, 1991.
- [3] T. Jiang, C. Wang, B. Wang, J. Xie, N. Jiao, T. Luo, "Real-time Red Tide Alge Classification Using Bayes Classifier and SVM", *In proceeding of International Conference on Bioinformatics and Biomedical Engineering (ICBBE)*, pp.2888-2891, May 2008.
- [4] T. Jiang, C. Wang, B. Wang, J. Xie, N. Jia, T. Luo, "Real-time Red Tide Algae Recognition using SVM and SVDD", *In proceeding of International Conference on Intelligent Computing and Intelligent System (ICIS)*, pp.602-606, Oct. 2010.
- [5] X. Lili, J. Tao, X. Jiozhen, Z. Shaoping, "Red Tide Algae Classification Using SVM-SNP and Semi-supervised FCM", *In proceeding of International Conference on Education Technology and Computer (ICETC)*, pp.389-392, Jun. 2010.
- [6] 한화용, "패턴인식 개론 MATLAB 실습을 통한 입체적 학습", 한빛미디어, pp.241-277, 2009.

- [7] R. C. Gonzalez, R. E. Woods, S. L. Eddins, "Digital Image Processing Using MATLAB", New Jersey, *Prentice Hall*, 2003.
- [8] J. Han, M. Kamber, "Data Mining Concepts and Techniques", *Morgan Kaufmann*, 2006.