

Hiperspektral Görüntülerde Gürültülü Bantların Bilişim Kuramsal Tespiti

Information-Theoretic Noisy Band Detection in Hyperspectral Imagery

Mustafa Ergül, Fatih Nar, Emre Akyılmaz, Nigar Şen
Uzay ve Savunma Teknolojileri (SDT)
Ankara, Türkiye
{mergul, fnar, eakyilmaz, nsen}@sdt.com.tr

Müjdat Çetin
Sabancı Üniversitesi
İstanbul, Türkiye
mccetin@sabanciuniv.edu

Özetçe—Hiperspektral görüntüler farklı dalga boylarında ait oldukları materyaller ile ilgili spektral bilgiler taşıyan ardışık yüzlerce banttan oluşmaktadır. Bununla birlikte, atmosferik etkiler ve sensör yapısı gibi pratik faktörlerden dolayı bazı spektral bantlar yüksek seviyede gürültü içermektedir. Bu bildiride, gürültülü bantları otomatik ve etkin olarak tespit eden karşılıklı bilgi miktarı kavramına dayalı bir algoritma önerilmiştir. Önerilen yaklaşımın etkinliği ve doğruluğu AVIRIS ve TELOPS sensörleri ile toplanan hiperspektral verilerle doğrulanmıştır. Test sonuçları önerilen yaklaşımın diğer algoritmalara göre daha üstün performans sergilediğini göstermektedir.

Anahtar Kelimeler — gürültü azaltma; gürültülü bant tespiti; karşılıklı bilgi miktarı; bilişim kuramı; hiperspektral görüntü işleme.

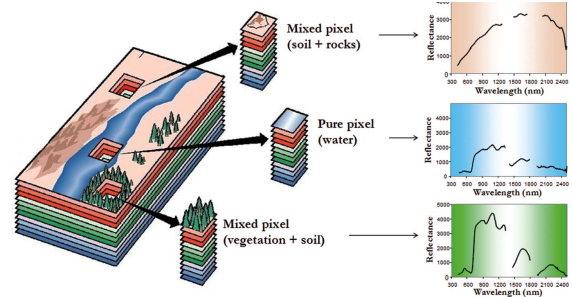
Abstract—Hyperspectral imagery consists of hundreds of successive bands that carry spectral information about the underlying materials at various wavelengths. However, due to practical factors such as atmospheric effects and sensor characteristics, some spectral bands contain high amounts of noise. In this paper, an effective information-theoretic algorithm based on mutual information that automatically detects such noisy bands is proposed. The effectiveness and accuracy of the proposed approach is validated on hyperspectral images collected by the AVIRIS and TELOPS sensors. Experimental results show that the proposed method outperforms the other algorithms in the literature.

Keywords — hyperspectral denoising; noisy band detection; mutual information; information theory; hyperspectral image processing.

I. GİRİŞ

Son yıllarda meydana gelen teknolojik gelişmelerle birlikte hiperspektral görüntüleme sistemlerinin tarım, mineraloji, jeoloji, çevrebilim, savunma sanayii vb. alanlara ait çeşitli uzaktan algılama uygulamalarında

kullanımı hızla yaygınlaşmaktadır. Geleneksel görüntüleme tekniklerinin aksine, hiperspektral veriler yüzlerce ardışık dar bant aralıklarından toplanmış spektral bilgilerden oluşmaktadır (Şekil 1). Bu sayede, farklı materyaller için farklı spektral imzalar içeren çok bantlı verilerden materyal tanıma ve hedef tespiti gibi uygulamaları daha başarılı bir şekilde gerçekleştirme kapasitesine sahip güçlü bir araç ortaya çıkmaktadır.



Şekil 1: Hiperspektral görüntüleme veri küpü yapısı ve spektral imza örneği [J].

Bununla birlikte, atmosferik etkiler ve sensör yapısı gibi faktörlerden dolayı hiperspektral (HS) verilerde hem uzamsal hem de spektral boyutta farklı seviyelerde gürültü meydana gelmektedir. Bu gürültülerin varlığı, eldeki verinin görselliğini bozmasının yanı sıra görüntü üzerinde gerçekleştirilecek sınıflandırma, spektral ayrıştırma, hedef tespiti gibi işlemlerin performansını da düşürmektedir. Bu yüzden, hiperspektral verilerdeki gürültüleri azaltmak ve ardından gelebilecek olası işlemlerden önce veri kalitesini yükseltmek hayati önem taşımaktadır.

HS verilerde var olan en kritik gürültü; yansıtıcı spektral bant bölgelerinde atmosferdeki oksijen ve su buharı parçacıklarının varlığından, yayıcı spektral bant bölgesinde ise sensör yapısından kaynaklanmaktadır. Düşük sinyal-gürültü oranına (SGO) sahip ve materyalle ilgili herhangi bir bilgi taşımayan bu spektral bantlar, olası görüntü analizi işlemlerine geçmeden önce atılmalıdır.

Literatürdeki çalışmalarda genellikle bu bantlar araştırmacı tarafından manuel olarak atılmaktadır. Ama bu işlem çok sayıda veri içeren durumlarda hem çok yorucu olmakta hem de tam otomatik bir sistem geliştirilmesini imkânsız hale getirmektedir. Son yıllarda bu sorunun üstesinden gelmek için düşük SGO değerine sahip gürültülü bantları otomatik olarak tespit eden ve veriden çıkaran metotlar önerilmeye başlanmıştır. [2,3]'te spektral bantlar arasındaki ilinti katsayısı kullanılarak gürültülü bantlar belirlenmeye çalışılmıştır. Bu algoritmalarındaki temel yaklaşım ilinti katsayısının ardışık temiz bantlar arasında yüksek, ardışık gürültülü bantlar arasında ise düşük değerlere sahip olacağı perspektifine dayanmaktadır. Nguyen ve arkadaşları [3], ilinti katsayısını uzamsal bilgiyi dikkate almadan global olarak hesaplarken; Bhushan ve diğerleri [2] ise bantlar arasındaki ilintiyi uzamsal bilgileri de göz önünde bulundurarak bloklar arasında hesaplamaktadır. Bir diğer çalışmada [4], dördün ağaç veri yapısı ve oransal kırılma boyutu kavramı benimsenerek verinin uzamsal boyutta yerel özelliklerini hesaba katan bir yöntem önerilmiştir.

Şu ana kadar literatürde önerilen yöntemler, iki rastgele değişkenin doğrusal ilişkisini ölçen ilinti katsayısına dayalı olduğu için gürültülü bantların belirlenmesinde yetersiz kalmaktadır. Bu çalışmada ise, bahsedilen problemin üstesinden gelmek için bantlar arasındaki doğrusal olmayan ilişkileri de hesaba katan karşılıklı bilgi miktarı (KBM) kavramına dayalı bir algoritma önerilmiştir. Bilişim kuramı alanında yaygın olarak kullanılan bu kavram hiperspektral görüntülerde boyut azaltma amacıyla bant seçme işleminde de kullanılmaktadır [5,6]. Bu yöntemlerde, yer gerçekliği gibi bir referans verisiyle yüksek seviyede istatistiksel benzerlik gösteren spektral bantlar “*anlamli bantlar*” olarak seçilir.

Bildirinin bundan sonrası şu şekilde organize edilmiştir. Bölüm 2’de KBM kavramı ve bu kavrama dayalı gürültülü bantların tespiti için önerilen algoritma anlatılmıştır. Deneyler sırasında kullanılan veri kümelerinin özelliklerinden Bölüm 3’te bahsedilmiştir. Bölüm 4, önerilen algoritmanın performans analizini içeren deney sonuçlarını sunmaktadır. Son olarak, bildiri sonuç ve gelecek çalışmalar ile sonlandırılmıştır.

II. ÖNERİLEN ALGORİTMA

A. Karşılıklı Bilgi Miktarı

Karşılıklı bilgi miktarı (KBM), iki rastgele değişken arasındaki istatistiksel bağımlılığı ölçmek için kullanılan bilişim kuramı alanındaki temel kavramlardan biridir. Birleşik olasılık dağılımı $p_{AB}(a, b)$, marjinal olasılık dağılımı $p_A(a)$ ve $p_B(b)$ olarak verilen iki rastgele değişken A ve B için KBM değeri aşağıdaki gibi hesaplanır:

$$I(A, B) = \sum_{a,b} p_{AB}(a, b) \log \frac{p_{AB}(a, b)}{p_A(a) \cdot p_B(b)} \quad (1)$$

Ayrıca, Shannon entropisi tanımı kullanılarak Denklem (1) entropi cinsinden aşağıdaki gibi de ifade edilebilir.

$$\begin{aligned} I(A, B) &= H(A) + H(B) - H(A, B) \\ &= H(A) - H(A|B) = H(B) - H(B|A) \end{aligned} \quad (2)$$

Bu ibarede, $H(A)$ ve $H(B)$ marjinal entropi değerlerini, $H(A, B)$ birleşik entropi değerini, $H(A|B)$ ve $H(B|A)$ ise koşullu entropi değerlerini belirtmektedir.

Sezgisel olarak ise, KBM A ve B rastgele değişkenlerinin paylaştığı bilgi miktarını ölçmektedir. Yani, bir değişken gözlemlendiğinde veya bilindiğinde diğer değişkenin belirsizliğinin ne kadar azaldığını ölçmektedir. İlinti katsayısından farklı olarak, KBM iki rastgele değişkenin ilişkisinin doğrusal olup olmadığı ile ilgilenmemektedir. Diğer bir fark ise; ilinti katsayısı değişkeni olasılık dağılımına ihtiyaç duyulmadan ikinci dereceden momentler kullanılarak hesaplanırken, KBM değişkeninin olasılıksal dağılım bilgisini de hesaba katarak bir değer üretmesidir.

B. Gürültülü Bantların Belirlenmesi

HS veriler dar spektral genişliğe sahip çok sayıda banttan oluştuğundan, ardışık bantlar çoğu zaman benzer bilgiler taşımaktadır. Bununla birlikte, bu olgusal gerçeklik düşük SGO’ya sahip gürültülü bantlarda sağlanamamaktadır. Gürültülü bantların tespiti işlemi bu gözleme dayanarak, ardışık spektral bantlar arasındaki benzerliği ölçen KBM metriği vasıtasıyla gerçekleştirilir.

KBM metriğinin hesaplanması sırasında, ihtiyaç duyulan marjinal ve birleşik olasılık dağılımlarının kaba birer kestirimi olarak, her banttaki piksel değerlerinden oluşturulan marjinal ve birleşik histogramlar kullanılmaktadır. Her bir spektral bandın ardışık spektral bantlara benzerliği aşağıdaki gibi hesaplanmaktadır:

$$d(i) = \max(I(i, i-1), I(i, i+1)) \quad (3)$$

i , spektral bant numarasını; I , karşılıklı bilgi değerini; $d(i)$ ise ilgili bandın komşu bantlarına benzerlik seviyesini temsil etmektedir. Yukarıda hesaplanan benzerlik değeri ne kadar düşük ise ilgili spektral bandın gürültülü bant olarak belirlenme olasılığı o oranda artmaktadır. Bu yüzden, herhangi bir spektral bant bir önceki veya sonraki spektral banda benzerlik gösteriyorsa o bant gürültüsüz olarak işaretlenir.

Ayrıca, bazı HS verilerde kötü veya eksik piksellerden ötürü çok miktarda tuz-biber gürültüsü meydana gelmektedir. Bu gürültü tipi, ilgili pikselin bütün bantlarında meydana geldiği için normalde yüksek seviyede benzerlik gösteren spektral bantların KBM değerleri ve dolayısıyla benzerlik değeri düşebilmektedir. Bu yüzden, bu çalışmada benzerlik hesaplamasına geçmeden önce uzamsal boyuttaki tuz-biber gürültü

pikselleri her bir bant için medyan filtre ile temizlenmektedir.

III. HİPERSPEKTRAL TEST VERİSİ KÜMESİ

Önerilen algoritmanın performans analizini gerçekleştirmek için AVIRIS ve TELOPS sensörleri ile toplanmış 4 farklı veri kümesi kullanılmıştır. AVIRIS sensörü tarafından toplanan ve literatürde yoğun bir şekilde kullanılan “Indian Pines” ve “Salinas” adlı veriler, yansıtılık spektral bant bölgesinde dalga boyu 0.4 - 2.5 μm aralığında 224 spektral banttan oluşmaktadır. Indian Pines verisi düşük uzamsal çözünürlüğe sahip olup 145x145 piksel, yüksek çözünürlüklü (~ 3.7 m/pixel) Salinas verisi ise 512x217 piksel içermektedir. Yayıncılık spektral bant bölgesinde bilgi içeren TELOPS verileri ise dalga sayısı yaklaşık olarak 800 – 1350 cm^{-1} aralığında 165 ve 660 adet spektral bant içeren iki veri kümesinden oluşmaktadır. Yerden toplanan bu veriler yüksek çözünürlüklü olup 320x256 piksel içermektedir. İçerik olarak, 660 bantlı veri çeşitli gömülü hedeflerin yerleştirildiği toprak ve taşları; 165 bant içeren veri ise bina ve arabaların yer aldığı bir şehir merkezini kapsamaktadır.

Indian Pines ve Salinas veri kümeleri, atmosferik etkilerden ve sensör yapısından kaynaklı sırasıyla 32 ve 21 gürültülü bant içermektedir. TELOPS verilerinin gürültülü bant sayısı ise 18/165 ve 90/660 olarak belirlenmiştir.

IV. DENEYSEL SONUÇLAR

Önerilen yaklaşımın geçerliliğini doğrulamak ve literatürdeki diğer yaklaşımlara göre etkinliğini kanıtlamak için çeşitli testler gerçekleştirilmiştir. Yer gerçekliği belirlenmesi sırasında, su soğurma spektral bölgesindeki ve yüksek gürültü seviyesine sahip spektral bantlar gürültülü bant sınıfı; diğer bantlar ise bilgi verici veya temiz bant sınıfı olarak işaretlenmiştir. Performans değerlendirmesi sırasında her algoritma sonucu için yer gerçekliğine göre kesinlik-geri getirme eğrisi (KGE) üretilmiştir ve kıyaslama bu eğriler üzerinden yapılmıştır. Her bir banda ait benzerlik değeri üzerinden eşikleme işlemi gerçekleştirilerek eğri üzerindeki farklı çalışma noktaları elde edilmiştir.

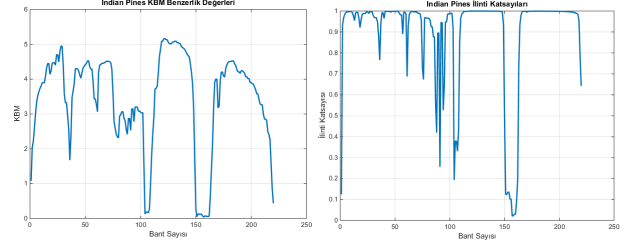
Önerilen yaklaşımın diğer yöntemlere göre üstünlüğünü doğrulamak için literatürde yaygın olarak kullanılan ilinti katsayısı ve sinyal-gürültü oranı (SGO) karşılaştırma için kullanılmıştır. İlinti katsayısına dayalı benzerlik değeri şu şekilde hesaplanmıştır:

$$\rho(i, i + 1) = \frac{E[(A_i - \mu_i)(A_{i+1} - \mu_{i+1})]}{\sigma_i \cdot \sigma_{i+1}} \quad (4)$$

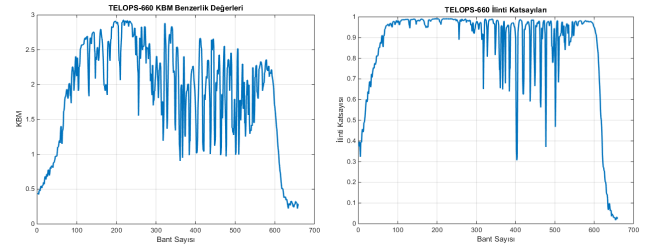
Denklem 4’te $\rho(i, i + 1)$, i . ve $(i + 1)$. bantlar arasındaki ilinti katsayısı; A_i , verinin i . bandını; μ_i , i . bandın ortalama değerini; σ_i , i . bandın standart sapmasını temsil etmektedir. İlgili banda ait benzerlik değeri ise KBM hesaplamasında olduğu gibi hesaplanmıştır.

$$d_\rho(i) = \max(\rho(i, i - 1), \rho(i, i + 1)) \quad (5)$$

Indian Pines ve TELOPS-660 veri kümesi kullanılarak elde edilen KBM benzerlik değeri ve ilinti katsayısı benzerlik değeri grafikleri Şekil 2 ve Şekil 3’te gösterilmektedir.



Şekil 2: Indian Pines verisine ait KBM (soldaki) ve İlinti katsayısı benzerlik değeri (sağdaki).



Şekil 3: TELOPS-660 verisine ait KBM (soldaki) ve İlinti katsayısı benzerlik değeri (sağdaki).

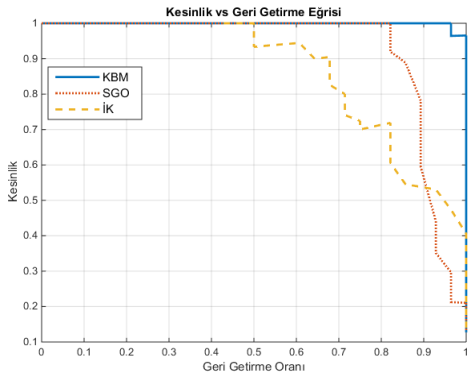
İlgili çizgeler dikkatlice incelendiğinde, ilinti katsayısına dayalı benzerlik değerlerinin KBM katsayısına dayalı yöntemden farklı olarak temiz veya bilgi taşıyan bantlarda da (Indian Pines: 80-95, TELOPS-600: 300:500) düşük değerlere sahip olduğu görülmektedir. Bundan dolayı, ilinti katsayısına dayalı gürültülü bantların tespiti sonucunda çok sayıda yanlış alarm meydana gelmektedir. İlinti katsayısı yerine KBM değerlerinin kullanılması sonucunda ise algoritmanın kesinlik oranının arttığı gerçekleştirilen deneyler sonucunda da ortaya çıkmaktadır.

Testler sırasında kullanılan diğer bir karşılaştırma yöntemi, bir sinyaldeki gürültü seviyesini ifade eden sinyal-gürültü oranı metriğine dayanmaktadır. Bu yaklaşımda, düşük SGO değerli bantlar gürültülü bant olarak işaretlenmektedir. SGO değerlerinin kestirilmesi sırasında ise referans görüntüsü olarak her bir banda ayrı ayrı uygulanan Wiener süzgeç sonucu kullanılmaktadır.

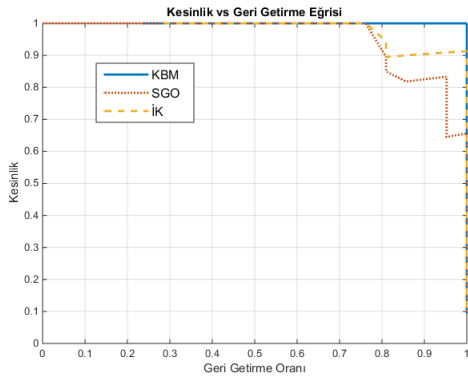
$$SGO(i) = \frac{w(A_i)^2}{(A_i - w(A_i))^2} \quad (6)$$

$w(\cdot)$, Wiener filtreyi; A_i , HS verinin i . bandını belirtmektedir. Filtre çıktısının enerjisi sinyal enerjisi, orijinal veri ile filtre çıktısının farkının enerjisi de gürültü enerjisi olarak değerlendirilmiştir.

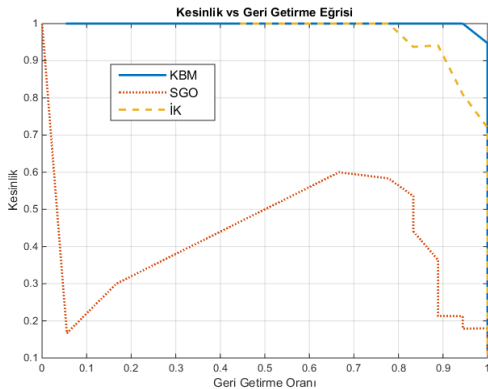
Önerilen yaklaşımın ve diğer iki algoritma sonucunun kesinlik-geri getirme eğrisi her bir veri kümesi için Şekil 4 - 7'de sunulmuştur. Bu grafiklerde sunulan test sonuçlarından da görüldüğü üzere önerilen yöntem hem SGO hem de İK tabanlı yaklaşımlardan istikrarlı bir şekilde daha üstün performans göstermiştir. Bunun yanında, ilinti katsayısına dayalı yöntem Salinas ve TELOPS-165 veri kümesinde SGO yaklaşımına göre daha başarılı performans göstermesine rağmen diğer iki veri kümesinde aynı başarıyı gösterememiştir. Bu sonuçlar, önerilen KBM tabanlı algoritmanın aksine İK ve SGO tabanlı yaklaşımların gürbüz ve istikrarlı sonuç üretmediğini göstermiştir.



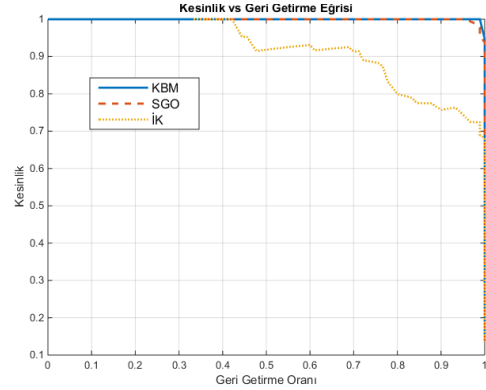
Şekil 4: Indian Pines HS verisine ait algoritma KGE çizgeleri.



Şekil 5: Salinas HS verisine ait algoritma sonucu KGE çizgeleri.



Şekil 6: TELOPS-165 HS verisine ait algoritma KGE çizgeleri.



Şekil 7: TELOPS-660 HS verisine ait algoritma KGE çizgeleri.

Ayrıca, önerilen yöntemin sensör farkı (AVIRIS, TELOPS), spektral bant bölgesi (VNIR-SWIR, LWIR), spektral bant sayısı vb. faktörlerden bağımsız olarak benzer performansı yakaladığı deneyler sonucunda anlaşılmaktadır. Önerilen yaklaşım hız açısından analiz edildiğinde ise karşılıklı bilgi miktarı değerinin hesaplama maliyetinin ilinti katsayısı hesaplama maliyetine yakın olduğu görülecektir. n piksel ve m bant sayısı olmak üzere ikisi de $\mathcal{O}(nm)$ hesaplama karmaşıklığına sahiptir.

V. VARGILAR

Bu bildiriye, hiperspektral görüntülerde atmosferik etkiler ve sensör yapısı gibi faktörlerden kaynaklanan gürültülü bantların tespiti için etkin bir algoritma önerilmiştir. Önerilen yaklaşım bilişim kuramı alanında yaygın olarak kullanılan karşılıklı bilgi miktarı kavramına dayanmaktadır. Farklı sensör, spektral bant bölgesi ve bant sayısında toplanmış verilerle gerçekleştirilen deneyler sonucunda önerilen yaklaşımın diğer yöntemlere göre daha gürbüz ve başarılı sonuçlar ürettiği gözlemlenmiştir. Gelecek çalışmalarda, bilişim kuramı kavramlarının hiperspektral görüntülerde gürültü azaltma probleminin çözümüne katkısı araştırılacaktır.

KAYNAKÇA

- 1 Ciznicki M., Kurowski K, and Plaza A. "Graphics processing unit implementation of JPEG200 for hyperspectral image compression." *Journal of Applied Remote Sensing*, 6(1):1-14, 2012.
- 2 Bhushan D, Sowmya V, Manikandan M.S., "An effective pre-processing algorithm for detecting noisy spectral bands in hyperspectral imagery." *IEEE Sym. on Ocean Elec.* p.34-39, 2011.
- 3 Nguyen V.H, Hakil K. "Noisy Band Removal Using Band Correlation in Hyperspectral Images." *Korean Journal of Remote Sensing*, 25(2):263-270, 2009.
- 4 Chang A.J, Yong-II K. "Noise Band Extraction of Hyperion Image using Quad tree Structure and Fractal Characteristic." *Korean Journal of Remote Sensing*, 26(5):489-495, 2010.
- 5 Wang B, Wang X, Chen Z. "Spatially Entropy based Mutual Information in Hyperspectral Band Selection for Supervised Classification." *Int. J. of Numerical Analysis and Modeling*, 9(2):181-192, 2012.
- 6 Guo B, Gunn S, Damper B, Nelson J. "Adaptive band selection for hyperspectral image fusion using mutual information." *IEEE Int. Conf. on Information Fusion*, p.1-8., 2005