

Genelleştirilmiş Gauss Yoğunluk Modellemesi ile Müzik Aletlerinin Sınıflandırılması

Music Instrument Classification Using Generalized Gaussian Density Modeling

M. Erdal Özbeke, F. Acar Savacı

Elektrik Elektronik Mühendisliği Bölümü
İzmir Yüksek Teknoloji Enstitüsü, İzmir
{erdalozbek, acarsavaci}@iyte.edu.tr

Özetçe

Bu çalışmada farklı müzik aletlerinin yalıtılmış nota örneklerinin tek boyutlu dalgacık ayrıştırması uygulanarak elde edilen altbant katsayıları genelleştirilmiş Gauss yoğunluğu ile modellenmiştir. Sadece model parametrelerinin kullanılmasıyla, iki farklı yoğunluk arasındaki Kullback-Leibler uzaklığı hesaplanarak müzik aletlerinin sınıflandırması yapılmıştır. Dalgacık ayrıştırmasında kullanılan farklı temel dalgacık işlevlerinin etkileri incelenmiştir.

Abstract

In this work, subband coefficients obtained from one dimensional wavelet decomposition of isolated note samples of different instruments has been modeled using generalized Gaussian density. By using only model parameters, the classification of music instruments has been done by calculating the Kullback-Leibler divergence between two different densities. The effect of different mother wavelet functions used in wavelet decomposition has also been investigated.

1. Giriş

Çoklu ortam uygulamaları ve kullandıkları verilerdeki çok hızlı artış, verilerin etkin bir şekilde ayrıştırılması ve sınıflandırılması gereğini ortaya çıkarmıştır. Bu sınıflandırma ya da öbekleme işlemi müzik kayıtlarındaki bilginin kullanım amacına göre değişmektedir. Örneğin pop, caz, klasik gibi belirli bir müzik türünden olduğu, şarkıcının ya da bestecinin adı, çalan müzik aleti ve hatta müzik aletinin çaldığı notaların bulunması ve müzik yazısına aktarılması için çalışmalar bulunmaktadır [1, 2, 3, 4]. Bunlar ve benzeri çalışmalarda ses tanıma sistemlerinde kullanılan teknikler yardımıyla problemin çözülmesi için birçok araştırma yapılmış ve çok geniş bir öznitelik kümesi elde edilmiştir. MPEG-7 standardında da bulunan bu kümenin elemanlarına örnek olarak özilinti katsayıları, sıfır geçiş sayısı, izgesel akı, izgesel kitle merkezi, mel frekans kepsral katsayılarını verebiliriz [1, 3, 5].

Sadece zaman ya da frekans temelli yaklaşımların müzik işaretlerini tamamıyla ifade edememesi nedeniyle zaman-frekans gösterimlerine gerek duyulmuştur [6]. Zaman-frekans

ortamındaki enerji dağılımının öznitelik bulunması için kullanılması ile geleneksel yaklaşımlar dalgacık altbantlarının enerjilerini öznitelik olarak kullanmışlardır. Böylece frekans ortamındaki enerji dağılımının görüntü verilerindeki örüntüleri belirlemesinden yararlanarak dalgacık katsayılarının altbantlarının marjinal dağılımının modellenmesi yapılmıştır [7]. Bu şekilde istatistiksel yaklaşım kullanılarak dalgacık ayrıştırması sonucunda elde edilen altbantların katsayılarının dağılımının müzik türlerini ayrıştırmada kullanılabileceği gösterilmiştir [8].

Bu çalışmada, genelleştirilmiş Gauss yoğunluk dağılımı ile modellenen müzik aletlerinin sınıflandırılması gerçekleştirilmiştir. Yalıtılmış nota örneklerinin tek boyutlu dalgacık ayrıştırması ile elde edilen dalgacık katsayılarının dağılımının model parametreleri bulunarak, iki farklı yoğunluk arasındaki Kullback-Leibler uzaklığına dayalı bir sınıflandırma yapılmıştır. Ayrıca, farklı temel dalgacık işlevlerinin sınıflandırma başarımına etkileri incelenmiştir.

İzleyen bölümde genelleştirilmiş Gauss yoğunluk modellemesi ve Kullback-Leibler uzaklığı kullanılarak benzerlik ölçümü ele alınacaktır. Üçüncü bölümde yalıtılmış nota örnekleri kullanılarak yapılmış benzetim çalışmaları ve sonuçları sunulacaktır. Son bölümde ise sonuçlar özetlenerek değerlendirilecektir.

2. Genelleştirilmiş Gauss Yoğunluk Modellemesi

Dalgacıklar verilen bir işareti birçok altbanda ayrıştırmak için etkilidirler. Dalgacık ayrıştırması sonucunda elde edilen altbantların enerji dağılımının öznitelik olarak kullanılması, bu enerji dağılımının modellenmesini ortaya çıkarmıştır. Deneyler belirli bir altbanddaki dalgacık katsayılarının marjinal yoğunluğunun bulunması için olasılık yoğunluk işlevi kullanılabileceğini göstermiştir. Tek boyutlu dalgacık ayrıştırması ile elde edilen katsayıların dağılımı için genelleştirilmiş Gauss yoğunluk modeli kullanılması önerilmiştir [7].

Verilen bir x işareti için genelleştirilmiş Gauss yoğunluk modeli

$$p(x; \alpha, \beta) = \frac{\beta}{2\alpha\Gamma(1/\beta)} e^{-(|x|/\alpha)^\beta} \quad (1)$$

ile verilir. Burada $\Gamma(\cdot)$ Gamma işlevidir ve $z > 0$ olmak üzere

$$\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt \quad (2)$$

ile tanımlanır.

Genelleştirilmiş Gauss yoğunluk modelindeki α oran parametresi olarak adlandırılır ve olasılık yoğunluk işlevi tepesinin genişliğini gösterirken, β da biçim parametresi olarak tepenin azalma hızıyla ters orantılıdır. Genelleştirilmiş Gauss yoğunluk modeli için iki özel durum ise dağılımın Gauss ($\beta = 2$) ve Laplace ($\beta = 1$) dağılımı durumudur.

N seviyeli bir dalgacık ayrıştırması için $N + 1$ adet altbant katsayısı, ($D_1, D_2, D_3, \dots, D_N, Y_N$) bulunmaktadır. Burada D_i, Y_i sırasıyla i . seviye detay ve yaklaşım katsayılarını göstermektedir. Her bir dalgacık altbandının detay ve yaklaşım katsayılarının dağılımları genelleştirilmiş Gauss yoğunluk modelindeki iki parametre (α, β) ile tanımlanabilir [7].

2.1. Kullback-Leibler uzaklığı kullanarak genelleştirilmiş Gauss yoğunlukları arasındaki benzerlik ölçümü

Genelleştirilmiş Gauss yoğunluk modelindeki birbirine bağımlı iki parametre (α, β) en büyük olabilirlik kestirimi ile elde edilmektedir. Belirli bir β değeri için α ,

$$\alpha = \left(\frac{\beta}{L} \sum_{i=1}^L |x_i|^\beta \right)^{1/\beta} \quad (3)$$

olarak bulunurken, β Newton-Raphson döngülü yöntemi ile sayısal olarak

$$1 + \frac{\Psi(1/\beta)}{\beta} - \frac{\sum_{i=1}^L |x_i|^\beta \log |x_i|}{\sum_{i=1}^L |x_i|^\beta} + \frac{\log \left(\frac{\beta}{L} \sum_{i=1}^L |x_i|^\beta \right)}{\beta} = 0 \quad (4)$$

denkleminin çözümü olarak bulunmaktadır [7]. (Burada $\Psi(\cdot)$ digamma işlevidir ve $\Psi(z) = \Gamma'(z)/\Gamma(z)$ ile verilir). Daha sonra bu iki değer kullanılarak, iki farklı yoğunluk arasındaki benzerlik Kullback-Leibler uzaklığı kullanılarak hesaplanmaktadır.

Verilen iki olasılık dağılım işlevi $f(s)$ ve $g(s)$ arasındaki Kullback-Leibler uzaklığı

$$D(f||g) = \int_{-\infty}^{\infty} f(s) \log \left(\frac{f(s)}{g(s)} \right) ds \quad (5)$$

ile tanımlanır. Burada $D \geq 0$ 'dır ve eşitlik ancak ve ancak $f(s)$ ve $g(s)$ aynı dağılımlar ise geçerlidir. Bilindiği gibi Kullback-Leibler uzaklığı simetrik olmadığından bir uzaklık değil ancak iki dağılımın yakınlığını gösteren istatistiksel bir yöntemdir.

İki dalgacık altbandının benzerliğini, (1) denklemini (5)'de yerleştirerek ve bazı işlemler sonucunda

$$D(p(\cdot; \alpha_1, \beta_1) || p(\cdot; \alpha_2, \beta_2)) = \log \left(\frac{\beta_1 \alpha_2 \Gamma(1/\beta_2)}{\beta_2 \alpha_1 \Gamma(1/\beta_1)} \right) + \left(\frac{\alpha_1}{\alpha_2} \right)^{\beta_2} \frac{\Gamma((\beta_2 + 1)/\beta_1)}{\Gamma(1/\beta_1)} - \frac{1}{\beta_1} \quad (6)$$

olarak elde etmek mümkündür.

Farklı altbandlardaki dalgacık katsayılarının birbirinden bağımsız olduğu varsayımıyla iki farklı işaretin (I_1, I_2) benzerliği, j altband indisi olmak üzere

$$D(I_1, I_2) = \sum_{j=1}^{N+1} D(p(\cdot; \alpha_1^{(j)}, \beta_1^{(j)}) || p(\cdot; \alpha_2^{(j)}, \beta_2^{(j)})) \quad (7)$$

ile verilir [7]. Böylece Kullback-Leibler uzaklığı ile hesaplanan uzaklıkların hepsi bir benzerlik ölçümü altında toplanmış olmakta ve çıkarılan öznitelikler için normalizasyon uygulanması gerekmemektedir.

3. Yapılan Sınıflandırma Çalışmaları ve Sonuçları

Bu çalışmada Iowa Üniversitesi Elektronik Müzik Stüdyosu [9] nota örnekleri kullanılmıştır. Bu notalar tek kanallı (mono) yalıtılmış notalardan oluşmuş ve 44100Hz'de örneklenmiştir. Veritabanından kullanılan yaylı müzik aletlerine ait örnekler içinden sadece 'yayla çalınan' nota örnekleri alınmış, farklı bir sesin elde edildiği 'çekilerek çalınan' örnekler çalışmalara dahil edilmemiştir. Örneklerden sadece kuvvetli (fortissimo (ff)) olarak kayıt edilmiş olanları kullanılmıştır.

Yalıtılmış nota örneklerinin başlangıç ve bitiş anlarını belirlemek için basit bir işlev olarak ortalama enerji bilgisinden yararlanılmıştır. Bir nota işareti x 'in L uzunluğundaki örneklerinin

$$E_{ort} = \frac{1}{L} \sum_{i=1}^L x_i^2 \quad (8)$$

ile verilen ortalama enerjisinin %10'undan büyük olduğu ilk ve son değerler arasında asıl işaretin var olduğu kabul edilerek nota örneklerinin başlangıç ve sonlarında bulunan sessiz bölümler çıkarılmıştır.

Üç seviyeli ($N = 3$) dalgacık ayrıştırması uygulanan her bir örnek için genelleştirilmiş Gauss yoğunluk model parametreleri α ve β bulunarak, $\{(\alpha_1, \beta_1), (\alpha_2, \beta_2), (\alpha_3, \beta_3), (\alpha_4, \beta_4)\}$ değerlerinden oluşan 8×1 'lik bir öznitelik vektörü elde edilmiştir.

Her bir müzik aleti için bütün nota örneklerini arka arkaya ekleyerek bir müzik aleti işaret örneği oluşturulmuştur. Bu örneğin her bir altband için müzik aletini temsil eden α ve β parametreleri bulunarak, her bir nota örneği için Kullback-Leibler uzaklığının hesaplanacağı referans değer olarak saklanmıştır. Daha sonra, her bir altband için her bir nota örneğinin α ve β parametreleri bulunarak, dalgacık katsayılarının dağılımı arasındaki benzerlik denklem (6)'da verildiği gibi Kullback-Leibler uzaklığı hesaplanarak bulunmuştur. Bir nota örneği ile müzik aleti işaret örneği arasındaki benzerlik, denklem (7)'de verilen toplam yardımıyla bulunmuş ve nota örneği en az uzaklığa sahip olan müzik aleti sınıfına ait olarak sınıflandırılmıştır.

İlk olarak sekiz adet üflemlerli müzik aletinin sınıflandırılması gerçekleştirilmiş ve elde edilen doğru sınıflandırma başarımları Tablo 1'de verilmiştir. Her satırda koyu olarak belirtilmiş sonuçlar verilen bir müzik aletinin elde ettiği en iyi sınıflandırma başarımlarını yüzdesini göstermektedir. Sonuçlardan sadece Eb Klarinet, Bas Flüt ve Soprano Saksofon'un doğru olarak tanımlandığı durumlar olduğu

Tablo 1: Sekiz adet üflemeli müzik aletinin sınıflandırma başarımı

Başarım (%)	Bas Klarnet	<i>E♭</i> Klarnet	<i>E♭</i> Klarnet	Alto Flüt	Bas Flüt	Flüt	Alto Saksofon	Soprano Saksofon
Bas Klarnet	6.5	4.3	8.7	8.7	13.1	13.1	21.7	23.9
<i>E♭</i> Klarnet	0	17.4	28.3	19.6	2.1	2.1	10.9	19.6
<i>E♭</i> Klarnet	2.6	17.9	28.2	20.5	0	2.6	2.6	25.6
Alto Flüt	0	0	5.4	29.7	40.6	8.1	2.7	13.5
Bas Flüt	0	0	5.3	18.4	65.8	0	0	10.5
Flüt	0	7.8	29.9	23.4	15.6	9.0	1.3	13.0
Alto Saksofon	3.1	0	1.6	10.9	48.4	10.9	9.4	15.7
Soprano Saksofon	0	4.7	12.5	29.7	3.1	17.2	3.1	29.7

görülmektedir. Müzik aletlerinin yapısı ve kullandıkları frekans aralıklarındaki farklılık nedeniyle aynı müzik aleti olmasına rağmen farklı müzik aleti gibi sınıflandırılmamın yüksek olduğu görülmektedir.

Tablo 2'de ise sadece Klarnet, Flüt ve Saksofon müzik aletleri isimlendirilmesiyle yapılmış sınıflandırma sonuçları gösterilmektedir.

Tablo 2: Klarnet, Flüt ve Saksofon müzik aletleri sınıflandırma başarımı

Başarım (%)	Klarnet	Flüt	Saksofon
Klarnet	50.8	11.7	37.5
Flüt	28.4	37.9	33.7
Saksofon	29.7	26.6	43.7

Görüldüğü gibi farklı müzik aleti sınıfına ayırma oranı yüksek olsa da, müzik aletleri arası ayırım daha belirgindir ve en büyük başarımlar yüzdeleri doğru sınıflandırmada elde edilebilmiştir.

Bir diğer çalışma olarak hem yaylı hem de üflemeli müzik aletlerin birarada bulunduğu durum ele alınmıştır. Telli Kontrbas ve Viyolonsel gibi yaylı müzik aletlerin yanı sıra Fagot, Obua ve Tuba gibi üflemeli müzik aletlerinden oluşan bir küme için elde edilen sonuçlar Tablo 3'te verilmiştir. Bu tablodan da görüleceği gibi müzik aletlerinin farklı ailelere ait olması durumunda sınıflandırma daha başarılı olmaktadır.

Tablo 3: Yaylı ve üflemeli müzik aletleri sınıflandırma başarımı (db1)

Başarım (%)	Telli Kontrbas	Fagot	Viyolonsel	Obua	Tuba
Telli Kontrbas	80.6	11.2	5.1	2.1	1.0
Fagot	12.5	77.5	2.5	2.5	5.0
Viyolonsel	5.3	13.3	63.7	14.2	3.5
Obua	0	0	8.6	91.4	0
Tuba	5.4	0	13.5	0	81.1

Çalışmada şimdiye kadar verilen sonuçlar, müzik işaretinin tek boyutlu dalgacık ayrıştırmasında kullanılan temel dalgacık işlevinin Daubechies işlevi 'db1', olarak seçildiği durumları göstermektedir. Dalgacık ayrıştırmasında kullanılan temel dalgacık işlevinin sonuçlara etkisinin irdelenmesi amacıyla farklı temel dalgacık işlevleri ile benzetimler tekrarlanmıştır. Tablo

4'te Daubechies 'db2' temel dalgacık işlevi kullanılarak elde edilmiş sonuçlar verilmiştir. Telli Kontrbas müzik aletinde önemli bir başarımlar düşüşü gözlenmiştir.

Tablo 4: Yaylı ve üflemeli müzik aletleri sınıflandırma başarımı (db2)

Başarım (%)	Telli Kontrbas	Fagot	Viyolonsel	Obua	Tuba
Telli Kontrbas	48.0	34.7	8.2	7.1	2.0
Fagot	7.5	80.0	0	2.5	10.0
Viyolonsel	1.8	13.3	61.9	22.1	0.9
Obua	0	0	5.7	94.3	0
Tuba	2.7	5.4	0	0	91.9

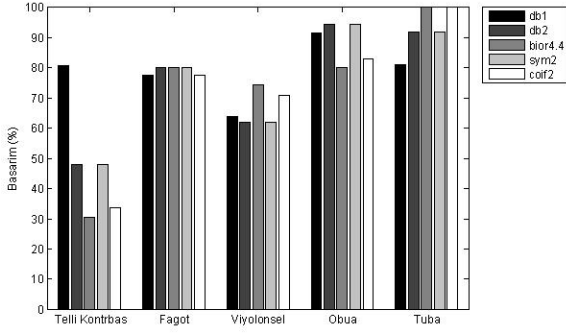
Farklı bir seçim olarak [8]'de verildiği gibi dikgen 'bior4.4' temel dalgacık işlevi kullanıldığında ise sonuçlar Tablo 5'te verildiği gibidir.

Tablo 5: Yaylı ve üflemeli müzik aletleri sınıflandırma başarımı (bior4.4)

Başarım (%)	Telli Kontrbas	Fagot	Viyolonsel	Obua	Tuba
Telli Kontrbas	30.6	49.0	6.1	13.3	1.0
Fagot	7.5	80.0	2.5	2.5	7.5
Viyolonsel	0	3.6	74.3	21.2	0.9
Obua	0	0	20	80.0	0
Tuba	0	0	0	0	100

Benzer şekilde symlet 'sym2' ve coiflet 'coif2' temel dalgacık işlevleri kullanılarak sınıflandırma da gerçekleştirilmiş ve karşılaştırma amacıyla bütün doğru sınıflandırma sonuçları Şekil 1'de verilmiştir.

Sonuçlardan gözlemlendiği gibi farklı temel dalgacık işlevlerinin sınıflandırmaya çok önemli bir etkisi görülmemektedir. Müzik aletinin doğru tanımlama başarımı, öncelikle müzik aletine göre değişmekte, temel dalgacık işlevlerinin değişimi bir ölçüde müzik aleti ve sınıflandırma yapılan kümedeki diğer müzik aletlerinin neler olduğuna bağlı kalmaktadır. Sonuçlardan Telli Kontrbas için 'db1' hariç %50'yi geçmeyen oranlar elde edildiği ancak Obua ve Tuba için %80'lerin üzerinde başarılı sınıflandırma yapılabildiği görülmektedir. Ancak tablolardan da görüldüğü üzere belirli bir temel dalgacık işlevi için yanlış müzik aleti olarak sınıflandırma yapılmış olabileceği gibi hiç hatasız başarımlar



Şekil 1: Farklı dalgacık işlevlerine göre başarımlar oranları

oranın da elde edilmesi mümkün olmuştur. Bunun nedenleri olarak temel dalgacık işlevlerinin yapılarının yansıma kullanılan müzik aletlerinin farklı frekans ya da nota aralıklarında çalınabilmelerinden ötürü farklı sayıda nota örneklerine sahip olmaları gösterilebilir.

4. Değerlendirme

Bu çalışmada yalıtılmış nota örneklerinin tek boyutlu dalgacık ayrıştırması ile elde edilen dalgacık katsayılarının dağılımı kullanılarak müzik aletlerinin sınıflandırılması yapılmıştır. Dalgacık katsayılarının dağılımlarının genelleştirilmiş Gauss yoğunluk dağılımı ile modellenmesi sonucunda elde edilen model parametreleri ile, farklı dağılımlar arasındaki Kullback-Leibler uzaklığı kullanılarak sınıflandırma gerçekleştirilmiş ve doğru sınıflandırma oranları elde edilmiştir. Ayrıca dalgacık ayrıştırmasında kullanılan farklı temel dalgacık işlevlerinin sınıflandırma sonuçlarına etkileri incelenmiştir.

Sekiz adet üflemeli müzik aleti sınıflandırılmasında başarımlar, farklı müzik aleti sınıflarına benzerlik nedeniyle düşük olmuştur. Müzik aletinin alt sınıflara ayrılmadan sadece adına göre sınıflandırıldığında başarımların yükseldiği görülmüştür. Benzer şekilde birbirinden farklı müzik aletleri kümelerinin kullanılmasının başarımlarını arttıracaklarını düşünerek yaylı ve üflemeli müzik aletlerinden oluşturulan bir küme için başarımlar oranları hesaplanmıştır. Elde edilen daha yüksek oranlar sonucunda, müzik aleti sınıflandırmasından önce daha genel sınıf etiketlerinin kullanılmasının yararlı olacağı düşünülmektedir.

Farklı temel dalgacık işlevlerinin başarımlar sonuçlarına doğrudan etkisi gözlemlenmiş, aynı müzik aleti için başarımlar oranlarının yakın olarak bulunması, işlevlerin müzik aleti ayırmakta önemli olduğunu ancak müzik aletinin diğerlerinden ayrılmasında tek başına yeterli olmadığını göstermektedir.

5. Kaynakça

- [1] Herrera-Boyer, P., Peeters, G., Dubnov, S., "Automatic classification of musical instrument sounds", *Journal of New Music Research*, 32(1):3–21, 2003.
- [2] Tzanetakis, G., Cook P., "Musical genre classification of audio signals", *IEEE Trans. on Speech and Audio Processing*, 10(5):293–302, 2002.

- [3] Kostek, B., "Musical instrument classification and duet analysis employing music information retrieval techniques", *Proc. of the IEEE*, 92(4):712–729, 2004.
- [4] Klapuri, A. P., "Automatic music transcription as we know it today", *Journal of New Music Research*, 33(3):269–282, 2004.
- [5] Essid, S., Richard, G., David, B., "Instrument recognition in polyphonic music based on automatic taxonomies", *IEEE Trans. on Audio, Speech, and Language Processing*, 14(1):68–80, 2006.
- [6] Pielemeier, W. J., Wakefield, G. H., Simoni, M. H., "Time-frequency analysis of musical signals", *Proc. of the IEEE*, 84(9):1216–1230, 1996.
- [7] Do, M. N., Vetterli, M., "Wavelet-based texture retrieval using generalized Gaussian density and Kullback-Leibler distance", *IEEE Trans. on Image Processing*, 11(2):146–158, 2002.
- [8] Tzagkarakis, C., Mouchtaris, A., Tsakalides, P., "Musical genre classification via generalized Gaussian and alpha-stable modeling", *IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, May 14-19 2006, Toulouse, France, 217–220.
- [9] The University of Iowa Electronic Music Studios. <http://theremin.music.uiowa.edu>