

Türk Müziği Enstrümanlarının Sınıflandırılması

Classification of Turkish Musical Instruments

M. Erdal Özbek, F. Acar Savacı

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

Özetçe

Bu çalışmada, Batı müziği enstrümanlarının tanınması amacıyla kullanılan özneliklerden mel frekans kepsral kat-sayıları ile Türk müziği enstrümanları sınıflandırılmıştır. Destek vektör makineleri ile yapılan çoklu sınıf sınıflandırma sonucunda farklı çekirdek parametreleri için karışım matrisleri bulunarak, başarımlar, duyarlılık ve belirleme oranları hesaplanmıştır. Ud örnekleri için %97'ye, tüm örnekler için ise ortalama olarak %90'a varan oranlarda yüksek başarımlar elde edilmiştir.

Abstract

In this work, Turkish musical instruments are classified with the features of mel frequency cepstral coefficients used for recognition of Western musical instruments. The performance, sensitivity and specificity ratios are calculated by finding the confusion matrices for different kernel parameters following multi-class classification performed by support vector machines. High performance reaching 97% for Ud samples and 90% in average for all samples is obtained.

1. Giriş

Müzik kayıtlarını otomatik olarak notaya dökmek için birçok yöntem önerilmiş ve kullanılmaktadır [1]. Bu yöntemlerin bir bölümü müzik işaretlerinden enstrümanların tanınması ve sınıflandırılması üzerinedir. Bu çalışmalarda temel olarak iki yaklaşım ele alınmaktadır. Yaklaşımlardan biri seslerin insanlar tarafından algılanmasına dayanan algılamacı yaklaşım; diğeri ise enstrümanın tınısı, türü, şekli, ya da üyesi bulunduğu enstrüman ailesi gibi enstrümanları gruplara ayırma temeline dayanan sınıflamacı yaklaşımdır. Çözüm önerileri genellikle sınıflandırılacak örneklerden özneliklerin seçilmesi, hesaplanması ve ardından bir öğrenme algoritması kullanılarak farklı sınıfları ayrıştırılması ve tanınmasına dayanmaktadır [2].

Müzik işaretlerinden bilgi edinilmesini sağlayan öznelikler için, konuşma işleme çalışmalarında sıkça kullanılan özneliklerden yararlanılmıştır. Bunlar arasında işaretin zaman ve frekans değişimlerine dayanan sıfır geçiş sayısı, özilinti işlevi, izgesel akı, izgesel kütle merkezi, mel frekans kepsral katsayıları (MFKK) gibi birçok öznelik bulunmaktadır. İnsan işitme sisteminin perde algılaması doğrusal olmadığından, 1000 mel olarak adlandırılan birimi 1000 Hz'e eşitleyerek doğrusal olmayan frekans değişimi modellenmiştir.

Bu modele göre bir işaretin mel değeri, f frekans değeri olmak üzere

$$mel(f) = 2595 \log\left(1 + \frac{f}{700}\right), \quad (1)$$

ile verilir. MFKK, işaretin güç izgesinin logaritmasının mel frekanslarında merkezlere sahip süzgeç bankasından geçirilip, ayrık kosinüs dönüşümünün alınması ile elde edilir [3]. MFKK, izgesel dağılımın şekli hakkında diğeri bir çok izgesel özneliklerden elde edilebilecek bilgiyi öz olarak içerdiğinden, konuşma işleme çalışmalarında olduğu gibi müzik işaret işleme çalışmalarında, ayrıca elde edilen başarımlardan dolayı özellikle de müzik enstrümanı sınıflandırma çalışmalarında yaygın olarak kullanılmaktadır [4, 5, 6, 7].

Sınıflandırılacak müzik kayıt örneklerinden çıkartılan bu öznelikler daha sonra birçok öğrenme algoritmaları yardımıyla sınıflara ayrıştırılmaktadır. Bunlar arasında destek vektör makineleri [8], sağladıkları yüksek başarımlar nedeniyle tercih edilmektedir [4, 7].

Ancak literatürde müzik enstrümanlarının tanınması ve sınıflandırılması üzerine yapılan çalışmaların çoğu Batı müziğini ele almaktadır. Uluslararası yayınlarda Türk müziği konusunda yapılan çalışmalar sayıca azdır. Yapılmış çalışmalarda ise daha çok makam, dizi, aralık, temel frekans ya da perde analizi yapılmakta, doğrudan enstrümanların sınıflandırılması ele alınmamaktadır [9, 10, 11].

Bu çalışmada MFKK özneliği ve destek vektör makineleri kullanılarak Türk müziği enstrümanlarının sınıflandırılması gerçekleştirilmiştir. Bu amaçla, izleyen bölümde destek vektör makineleri özetlenecektir. Yapılan çalışmalar ve sonuçlarını değerlendirme bölümü izleyecektir.

2. Destek Vektör Makineleri

Destek vektör makinelerinin temelleri istatistiksel öğrenme kuramına dayanmaktadır [8]. Bu kurama göre belirli bir öğrenme işlemi için sınırlı sayıda veri verildiğinde en iyi genelleştirme, sınıflandırma işlevinin kapasitesinin eğitim kümesinin boyutu ile uyumlu olduğu durumda elde edilir [12]. İlk olarak iki sınıfa ayırıcı olarak verilen destek vektör makinelerinde, n boyutlu eğitim örnekleri \mathbf{x}_i ve sınıf etiketi y_i ($y_i \in \{-1, 1\}$) olarak verildiğinde algoritma $\mathbf{w} \cdot \mathbf{x} + b = 0$ ile verilen en iyi ayırıcı yüzeyi bulmaya çalışır. Bu durumda

$$y_i (\mathbf{w} \cdot \mathbf{x}_i + b) - 1 \geq 0 \quad \forall i, \quad (2)$$

şeklinde yazılan optimizasyon problemi, en yakın eğitim örneği ile ayırıcı yüzey arasındaki uzaklık ($2/\|w\|$) en büyük olacak şekilde kısıtlandırılır. Ayırıcı yüzeye en yakın eğitim örnekleri, tüm eğitim örneklerinin küçük bir alt kümesini oluşturur ve destek vektörleri olarak adlandırılır [13, 14]. Sınıflandırma ise optimizasyon probleminin çözümünde kullanılan Lagrange kat-sayılarını (α_i) da içeren

$$f(\mathbf{x}) = \text{sgn} \left(\sum_{i=1}^n \alpha_i y_i \mathbf{x}_i \cdot \mathbf{x} + b \right) \quad (3)$$

işlevi kullanılarak gerçekleştirilir. Görüldüğü gibi çözüm, verilerin iç çarpımı cinsinden yazılabilmektedir.

Verilerin doğrusal olarak ayrıştırılmadığı durumda, veri uzayından verilerin doğrusal olarak sınıflandırılabilmesi daha büyük (genellikle sonsuz) boyutlu özel bir Hilbert uzayı olan doğuran çekirdekli Hilbert uzayına dönüşümü sağlayan Denklemin (4)'deki gibi bir çekirdek işlevi κ kullanılmaktadır.

$$\kappa(x_i, x_j) = \langle \phi(x_i) \cdot \phi(x_j) \rangle. \quad (4)$$

İki veri noktası x_i, x_j 'nin iç çarpımları çekirdek hilesi olarak bilinen bir değişim sayesinde ϕ işlevinin açık ifadesine gerek kalmadan hesaplanmaktadır. Böylece bir uzayda iç çarpım olup olmadığını belirleyen Mercer şartını sağladığı sürece en iyi ayırıcı yüzeyi oluşturmak için herhangi bir işlev kullanılabilir [15, 16]. En sık kullanılan çekirdek işlevlerinden birisi,

$$\kappa(x_i, x_j) = \exp \left(-\frac{\|x - z\|^2}{2\sigma^2} \right) \quad (5)$$

ile verilen Gauss işlevidir. Ayrıca belirli amaçlar ve uygulamalar için bir çok çekirdek işlevi bulunmaktadır [17].

Destek vektör makineleri esas olarak iki sınıf sınıflandırılması için tasarlanmıştır. Ancak sınıf sayısı daha çok olduğunda destek vektör makineleri bire-bir ve bire-tüm yöntemlerini kullanarak çok sınıflı sınıflandırma problemini çözebilmektedir. Her iki yöntem de çok sınıflandırma problemini iki sınıf sınıflandırma probleminin toplamı olarak ele almaktadır. Verilen bir k sınıf sınıflandırma problemi için, bire-tüm yöntemi bir sınıf ile kalan $k - 1$ sınıf arasında yüzeyi oluştururken, bire-bir yönteminde her olası sınıf çifti arasındaki $k(k - 1)/2$ sayısındaki yüzeyler oluşturulmaktadır. Her iki yöntemde de belirlenen bir seçme kuralı ile karar verilmektedir [18]. Her ne kadar yöntemin seçimi probleme dayalı olsa da, bire-tüm yöntemi çoğunlukla daha kabul edilebilir sonuçlar üretmektedir [14].

3. Çalışmalar ve Sonuçlar

Bu çalışmada Kanun, Keman, Kemeçe, Klarnet, Ney, Tambur ve Ud olmak üzere 7 adet Türk müziği enstrümanı ile çalınan ses örnekleri kullanılmıştır. Bu örnekler farklı sanatçıların farklı makamlardaki taksim eserlerini içeren toplam 293 adet kayıttan 5'er saniyelik bölümler ayrılarak elde edilmiştir. Bu kayıtların enstrümanlara göre dağılımı Tablo 1'de verilmektedir.

Çalışmanın ilk aşamalarında MIR [19], Auditory [20] ve MPEG-7 [21] yazılım paketleri kullanılarak harmonik izge merkezi, sapması, yayılımı ve değişimi, atak zamanı ve logaritması, izgesel ve zamansal kitle merkezi, parlaklık, entropi,

Tablo 1: Çalışmalarda kullanılan kayıt ve örnek sayıları.

| Enstrüman adı | Kayıt sayısı | Örnek sayısı |
|---------------|--------------|--------------|
| Kanun | 21 | 503 |
| Keman | 32 | 1190 |
| Kemeçe | 24 | 514 |
| Klarnet | 30 | 1412 |
| Ney | 44 | 971 |
| Tambur | 92 | 5958 |
| Ud | 50 | 1996 |
| Toplam | 293 | 12544 |

düzlük, harmonik olmama, düzensizlik, yamukluk, yayılma, sızır geçiş sayısı ve MFKK incelenmiştir. Az sayıda örnek üzerinden yapılan bu ilk inceleme sonucunda, sınıflandırma aşaması için literatürde de başarısı öne çıkmış olan MFKK özneliği seçilerek çalışmalar sürdürülmüş ve toplam sayısı 12544 olan her bir örnek için 13-boyutlu MFKK vektörü elde edilmiştir.

Elde edilen bu MFKK vektörleri düzgelenecek Spider [22] yazılım paketinin kullanıldığı destek vektör makinelerine öznelik olarak verilmiştir. Daha önceki çalışmalarımıza [23] dayanarak bu çalışmada çoklu sınıf sınıflandırma için bire-tüm yöntemi seçilmiştir. Elimizdeki özneliklerin rasgele olarak ayrılan yarısı destek vektör makinelerini eğitmek, kalan yarısı da test etmek amacıyla kullanılmıştır. Ancak aynı kayıttan hem eğitim hem de test için örnek alınmaması sağlanmıştır. Çekirdek işlevi olarak, yine önceki çalışmalarımıza dayanarak Gauss işlevi seçilmiş, parametre değeri σ ise, 0.1'den 1'e kadar 0.1'er adımla değiştirilmiştir. Sonuçlar karışım matrisleri olarak elde edilmiş, buradan duyarlılık ve belirleme değerleri hesaplanmıştır. Sonuçlar elde edilirken istatistiksel doğrulama amacıyla çapraz-onaylama yöntemi kullanılmış, çapraz-onaylama set sayısı olarak 10 alınmıştır. Set içindeki sınıfa ait eleman sayısı, [24]'te önerildiği gibi eşit oranlanacak şekilde seçilmiştir.

Sonuçların elde edildiği her bir parametre değeri için hesaplanan karışım matrisleri içinden, benzer değerlere örnek oluşturması açısından $\sigma = 1$ için bulunan değerler ve yüzdelik oranları, kullanılan eğitim örnekleri için Tablo 2'de, test örnekleri için ise Tablo 3'te verilmiştir. Her iki tablodan da gözleneceği gibi, Kanun'un başarımı diğer enstrümanlara göre düşük kalmaktadır, daha çok Klarnet ve Tambur ile karıştırılmaktadır. Bunun nedeni olarak Kanun'un çok sesliliği ve Türk müziğinin piyanosu olarak adlandırılması gösterilebilir. Ayrıca, Tambur'un daha çok Ney ile, Ney'in de hemen hemen diğer tüm enstrümanlarla karıştırıldığı görülmektedir.

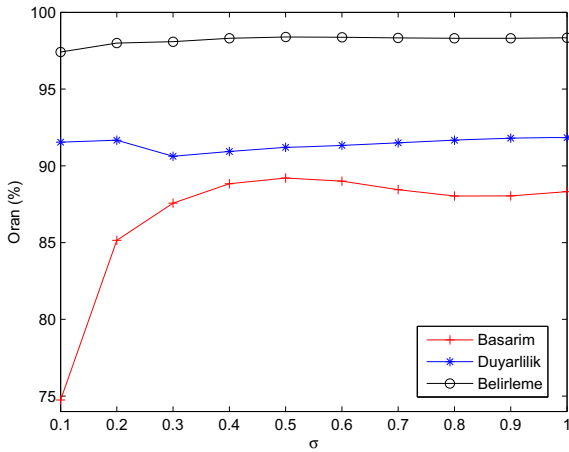
Test örneklerinden her bir parametre değeri için elde edilen ortalama başarımlar, duyarlılık ve belirleme oranları ise Şekil 1'de verilmiştir. Görüldüğü gibi en kötü durumda bile %75 civarında bir ortalama başarımlar elde edilmiş, çoğu σ değeri için %90'a varan başarımlar gözlenmiştir. Ortalama karışım matrisleri kullanılarak bulunan duyarlılık ve belirleme değerlerinin de %90'ın üzerinde olması, sonuçların istatistiksel olarak da doğruluğunu göstermektedir.

Tablo 2: Eğitim örnekleri için karışım matrisi ($\sigma = 1$)

| Sayı (Oran) | Kanun | Keman | Kemençe | Klarnet | Ney | Tambur | Ud |
|-------------|------------------|------------------|------------------|------------------|------------------|-------------------|-------------------|
| Kanun | 192 (%82) | 0 | 0 | 41 | 0 | 0 | 1 |
| Keman | 1 | 529 (%91) | 0 | 26 | 18 | 6 | 1 |
| Kemençe | 0 | 0 | 238 (%99) | 0 | 0 | 2 | 0 |
| Klarnet | 0 | 2 | 0 | 692 (%99) | 2 | 0 | 0 |
| Ney | 4 | 9 | 7 | 19 | 406 (%86) | 26 | 0 |
| Tambur | 1 | 19 | 2 | 7 | 58 | 2875 (%97) | 0 |
| Ud | 0 | 0 | 1 | 0 | 0 | 1 | 982 (%100) |

Tablo 3: Test örnekleri için karışım matrisi ($\sigma = 1$)

| Sayı (Oran) | Kanun | Keman | Kemençe | Klarnet | Ney | Tambur | Ud |
|-------------|------------------|------------------|------------------|------------------|------------------|-------------------|------------------|
| Kanun | 100 (%37) | 9 | 6 | 75 | 4 | 71 | 4 |
| Keman | 1 | 464 (%76) | 0 | 27 | 56 | 47 | 14 |
| Kemençe | 0 | 1 | 257 (%94) | 1 | 5 | 9 | 1 |
| Klarnet | 7 | 5 | 1 | 622 (%87) | 29 | 18 | 34 |
| Ney | 6 | 15 | 16 | 44 | 342 (%68) | 71 | 6 |
| Tambur | 25 | 71 | 39 | 25 | 159 | 2563 (%86) | 114 |
| Ud | 0 | 19 | 1 | 3 | 0 | 5 | 984 (%97) |



Şekil 1: Değişen parametre değerine göre ortalama başarımların, duyarlılıkların ve belirlenmelerin.

4. Değerlendirme ve Sonuç

Bu çalışmada Batı müziği işaretlerinin incelenmesinde ve enstrümanlarının sınıflandırılmasında başarısı bilinen mel frekans kepsral katsayıları kullanılarak Türk müziği enstrümanlarının sınıflandırılması gerçekleştirilmiştir. Destek vektör makineleri ile yapılan sınıflandırma sonucunda elde edilen başarımın yüksekliği, mel frekans kepsral katsayılarının Türk müziği kayıtlarının otomatik olarak notaya dökülmesi aşamasında öznelik olarak kullanılabilirliğini göstermiştir. Bu sonuçlara dayanarak, verilen bir kayıttan öncelikle enstrümanın belirlenip, daha sonra enstrüman bazında notaya dökme işleminin gerçekleştirilmesi için yapılan çalışmalara devam edilecektir.

5. Teşekkür

Bu çalışma Dr. Barış Bozkurt yürütücülüğünde TÜBİTAK-EEEAG tarafından desteklenen 107E024 numaralı "Klasik Türk müziği kayıtlarının otomatik olarak notaya dökülmesi ve otomatik makam tanıma" projesinin bir bölümü olarak ve proje çerçevesinde toplanılan kayıtlardan bir kısmı kullanılarak gerçekleştirilmiştir.

6. Kaynakça

- [1] A. Klapuri and M. Davy, Eds., *Signal Processing Methods for Music Transcription*. Springer, 2006.
- [2] P. Herrera-Boyer, G. Peeters, and S. Dubnov, "Automatic classification of musical instrument sounds," *Journal of New Music Research*, vol. 32, no. 1, pp. 3–21, September 2003.
- [3] J. R. Deller, J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*. New Jersey: Prentice Hall, 1987.
- [4] J. Marques and P. J. Moreno, "A study of musical instrument classification using Gaussian mixture models and support vector machines," Compaq Corporation, Cambridge Research Laboratory, Technical Report Series CRL 99/4, June 1999.
- [5] J. C. Brown, "Computer identification of musical instruments using pattern recognition with cepstral coefficients as features," *Journal of Acoustical Society of America*, vol. 105, no. 1, pp. 1933–1945, 1999.
- [6] A. Eronen and A. Klapuri, "Musical instrument recognition using cepstral coefficients and temporal features," in *Proc. of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'00)*, Istanbul, Turkey, 2000, pp. 753–756.

- [7] S. Essid, G. Richard, and B. David, "Instrument recognition in polyphonic music based on automatic taxonomies," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 14, no. 1, pp. 68–80, January 2006.
- [8] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer-Verlag, 1995.
- [9] C. Akkoç, "Non-deterministic scales used in traditional Turkish music," *Journal of New Music Research*, vol. 31, no. 4, pp. 285–293, December 2002.
- [10] O. Yarman, "A comparative evaluation of pitch notations in Turkish makam music," *Journal of Interdisciplinary Music Studies*, vol. 1, no. 2, pp. 43–61, 2007.
- [11] B. Bozkurt, "An automatic pitch analysis method for Turkish maqam music," *Journal of New Music Research*, vol. 37, no. 1, pp. 1–13, March 2008.
- [12] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, pp. 121–167, 1998.
- [13] N. Cristianini and J. Shawe-Taylor, *An Introduction to Support Vector Machines*. Cambridge University Press, 2000.
- [14] B. Schölkopf and A. J. Smola, *Learning with Kernels*. MIT Press, 2002.
- [15] C. Cortes and V. Vapnik, "Support vector networks," *Machine Learning*, vol. 20, pp. 273–297, 1995.
- [16] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Prentice-Hall, 1999.
- [17] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [18] J. Weston and C. Watkins, "Multi-class support vector machines," Department of Computer Science, Royal Holloway, University of London, Tech. Rep. CSD-TR-98-04, May 1998.
- [19] "MIR toolbox," <http://www.jyu.fi/music/coe/materials/mirtoolbox>.
- [20] "Auditory toolbox," <http://cobweb.ecn.purdue.edu/malcolm/interval/1998-010/>.
- [21] "MPEG-7 toolbox," <http://mpeg7.doc.gold.ac.uk/index.html>.
- [22] "Spider: Object-orientated machine learning library," <http://www.kyb.tuebingen.mpg.de/bs/people/spider>.
- [23] M. E. Özbek, N. Özkurt, and F. A. Savacı, "Musical instrument classification using wavelet ridges and support vector machines," submitted to IET Signal Processing.
- [24] R. Kohavi, "A study of cross-validation and bootstrap for accuracy estimation and model selection," in *International Joint Conference on Artificial Intelligence (IJCAI'95)*, vol. 2, Montreal, Québec, Canada, August 1995, pp. 1137–1143.