

İlintropi ile Müzik İşaretlerinin Temel Frekanslarının İzlenmesi

Fundamental Frequency Tracking of Musical Signals with Correntropy

M. Erdal Özbek, F. Acar Savacı

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

Özetçe

Bu çalışmada, notaya dökme amacıyla müzik işaretlerinin temel frekans bileşenleri izlenmiştir. Bu amaçla ilintropi işlevinin özilinti işlevi gibi işaretlerin temel frekansını bulmak amacıyla kullanılabilirdiği gösterilmiştir. Yöntemin başarımı farklı nota ve melodi örnekleri için YIN algoritması ile karşılaştırılarak değerlendirilmiştir. İlintropi işlevinin YIN algoritması kadar başarılı olduğu gösterilmiştir.

Abstract

In this work, the fundamental frequencies of the musical signals are tracked for transcription. For this purpose, it is shown that the correntropy function can be used like the autocorrelation function for finding the fundamental frequencies of signals. The successfulness of the method is evaluated by comparing with the YIN algorithm for different note and melody samples. The correntropy function is shown to be successful as the YIN algorithm.

1. Giriş

Müzik işaretlerinin ayrıştırılması ve sınıflandırılması üzerine yapılan çalışmalar son yıllarda gittikçe artmaktadır. Temel işaret işleme tekniklerinin ve örüntü tanıma algoritmalarının birarada kullanıldığı bu çalışmaların önemli bir bölümü, seslerin notaya dökülmesi amacıyla yönelmiştir [1, 2]. Bilindiği gibi müzik, seslerin gösterilmesi amacıyla kullanılan özel işaretler olan notalar ile ifade edilir. Yatay ekseninde zaman, dikey ekseninde ise nota işaretlerinin yerleştirildiği dizelerle gösterilir. Müzik seslerinin notalarla işaretlenmesi ya da notaya dökülme işlemi hangi müzik enstrümanının, ne kadar süre ile, hangi notayı çaldığının belirlenmesidir. Bu yüzden belirli bir enstrüman için nota ya da perde frekanslarının bulunması önem taşımaktadır.

Nota frekansını bulmak amacıyla, verilen bir nota sesi işaretinin tekrarlanma hızı ya da periyodu kullanılabilir. Periyod, periyodik ya da yaklaşık periyodik işaretler için tanımlanabilir ve bu periyodun tersi de işaretin temel titreşim frekansını (F0) verir. En bilinen ve basit F0 belirleme yöntemi zaman ortamında hesaplanan özilinti işlevidir. Örnek sayısı N olan bir $x(n)$ işaretinin özilinti işlevi,

$$R(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} x(n)x(n+\tau) \quad (1)$$

ile tanımlanır. Burada τ gecikme parametresi olarak bilinmektedir. Özilinti işlevindeki sıfır gecikme değeri işaretin enerjisini verir. Bu gecikmeden sonraki ilk tepedeki gecikme değeri işaretin periyodunu ve dolayısıyla F0 değerini belirler. Özilinti işlevinin hesaplanmasının kolay ve işaret üzerinden doğrudan gerçekleştirilebilmesinden dolayı, hem konuşma hem de müzik işaretlerinde sıklıkla kullanılmaktadır [3, 4]. Konuşma ve müzik işaretlerinin durağan olmaması nedeniyle kısa-zaman özilinti işlevinin bir pencere ile birlikte kullanılması önerilmekte, bu şekilde doğuşkan (harmonik) frekansların belirlenmesi de engellenerek oktav hatalarının önüne geçilmiş olmaktadır [2, 5]. Bu ve benzeri sorunların giderilmesi için önerilen YIN algoritması literatürde yaygın olarak kullanılmaktadır [3]. Basitliği ve ayarlanması gereken az değişkene sahip olmasının sağladığı kullanım kolaylığı ile, hem konuşma hem de müzik işaretlerinin temel frekans incelenmesinde referans olarak başvurulmaktadır.

Özilinti işlevinin hesaplanmasında ortaya çıkan ama F0'ı bulurken problem oluşturabilecek bir özelliği de tepelerinin, özellikle F0'ı gösteren tepenin geniş olmasıdır. Bu sorunu çözmek için daraltılmış özilinti işlevi önerilmiştir [6]. Ancak daha dar bir tepe genişliği için bulunması gereken terimlerin artması hesaplama karmaşıklığını arttırmaktadır. Önerilen bir diğer yöntem de toplam özilinti işlevidir ve ilk olarak seslerin algılanması için sunulmuştur [7]. Bu yöntemde işaret bir süzgeç bankasından geçirilmekte ve bütün alt-band ya da kanallardan gelen özilintiler toplanmaktadır. Daha sonra toplamın oluşturduğu tepedeki gecikme değeri kullanılarak F0 belirlenmektedir. Ancak bu yöntemin de her bir alt-bandda yapılan özilinti işlevi hesaplamasından dolayı hesaplama karmaşıklığı yüksektir.

Önceki çalışmamızda ilintropi [8] olarak adlandırdığımız genelleştirilmiş ilinti işlevi önerilmiştir [9]. İlintropi işlevine benzer özelliklere sahip bu işlev, bilgi kuramıyla öğrenme ve çekirdek yöntemlerine dayanmaktadır [9, 10, 11, 12]. Çekirdek işlevleri kullanılarak ikinci dereceden istatistikler içeren özilinti işlevi yerine yüksek dereceden istatistiklere sahip olan ilintropi kullanılabilir. Bu özelliğe dayanarak, ilintropi işlevinin kullanıldığı konuşma işaret örnekleri için perde belirleme algoritması da önerilmiştir [13, 14].

Bu çalışmada ilintropi işlevi kullanılarak F0 değerlerinin belirlendiği gösterilmiş ve başarımı YIN algoritması ile karşılaştırılmıştır. Bu amaçla izleyen bölümde çekirdek yöntemleri ve ilintropi işlevi özetlenecektir. Sonraki

bölümde yapılan çalışmalar anlatılacak ve elde edilen sonuçlar gösterilecektir. Son bölümde başarımlar değerlendirilecektir.

2. Çekirdek Yöntemleri ve İlintropi

Çekirdek yöntemleri, iç çarpımları kullanan çekirdek işlevi aracılığı ile giriş uzayındaki veriyi genellikle daha yüksek boyutlu bir öznelik uzayına taşımakta kullanılırlar. Bu uzay, özel bir Hilbert uzayı olan doğuran çekirdekli Hilbert uzayıdır. İki veri noktası x_i, x_j 'nin iç çarpımları

$$\kappa(x_i, x_j) = \langle \phi(x_i) \cdot \phi(x_j) \rangle. \quad (2)$$

ile verilen çekirdek işlevi kullanılarak hesaplanabilmektedir. En yaygın olarak kullanılan çekirdek işlevi

$$\kappa(x_i, x_j) = \frac{1}{\sqrt{2\pi\sigma}} \exp \left\{ -\frac{(x_i - x_j)^2}{2\sigma^2} \right\} \quad (3)$$

ile verilen Gauss çekirdek işlevidir. Bakışlımlı bir işlev olan Gauss çekirdek işlevinde σ , çekirdek parametresi olarak bilinmekte ve öznelik uzayında verinin dağılımını kontrol etmektedir.

İlintropi işlevi, verilen bir rastgele süreç $x(t), t \in T$ için T indis kümesi olmak üzere $T \times T$ 'den R^+ 'a tanımlanan,

$$V(t_1, t_2) = E[\kappa(x(t_1), x(t_2))] \quad (4)$$

ile ifade edilen bir işlevdir [9]. Burada $E[\cdot]$ istatistiksel beklentiyi göstermektedir ve görüldüğü üzere özilinti işlevinin bir çekirdek işlevi kullanılarak bulunması ile tanımlanmaktadır. Verilen N örneklili bir $x(n)$ işaretini için ilintropi işlevi, Denklem (1)'deki gibi yazılabilir [9]:

$$V(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} \kappa(x(n), x(n + \tau)). \quad (5)$$

Bu durumda ilintropi işlevi, öznelik uzayı için çekirdek işlevi aracılığı ile hesaplanan ilinti işlevi gibi düşünülebilir. Bilindiği gibi ilinti işlevi verinin sadece ikinci dereceden istatistiklerini yakalamaktadır. Ancak [9, 10]'da gösterildiği gibi ilintropi işlevinin işaret işleme ve makina öğrenme problemlerine uygulanmasını doğuran bir çok özelliği bulunmaktadır. Bizim durumumuz için, Denklem (4)'deki ifadede, Denklem (3)'te verilen çekirdek işlevini kullanıp daha sonra Taylor seri açılımını uygularsak

$$V(t_1, t_2) = \frac{1}{\sqrt{2\pi\sigma}} \sum_{k=0}^{\infty} \frac{(-1)^k}{(2\sigma^2)^k k!} E[(x(t_1) - x(t_2))^{2k}] \quad (6)$$

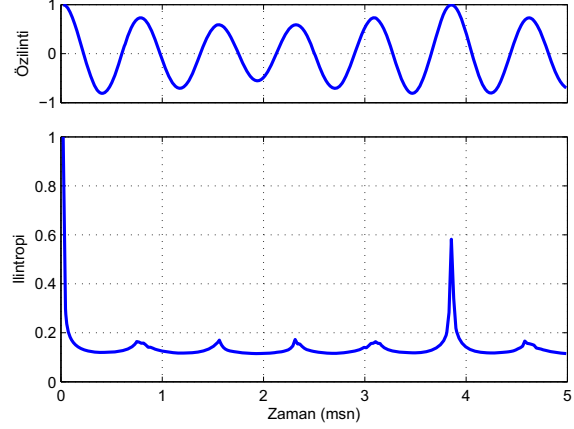
ile verilen ilintropi ifadesinde rastgele süreç $(x(t_1) - x(t_2))$ 'nin çift dereceden momentlerini elde ederiz. Yani, ilintropi rastgele süreçlerin yüksek derecede istatistiklerini içermektedir. Bu da ilinti işlevine karşı üstünlüğünü oluşturmaktadır.

3. Çalışmalar ve Sonuçlar

Çalışmalarımızda nota ve melodi örnekleri kullanılarak ilintropi işlevinin F0 bulma başarısı incelenmiştir. Nota örnekleri, Iowa Üniversitesi müzik enstrümanlarının [15] önceki çalışmalarımızda ayrıştırdığımız [16] nota örneklerinden

seçilmiştir. Melodi örnekleri ise kişisel kütüphanemizde bulunan klasik müzik CD'lerinden alınmıştır.

İlintropi işlevinin ilinti işlevine olan üstünlüğünü göstermek amacıyla öncelikle tek bir notanın ilintropisi bulunmuş ve özilinti işlevi ile karşılaştırılmıştır. Bunun için örnek olarak seçilen Obua C4 notasının düzgelenmiş özilinti ve ilintropi işlevleri Şekil 1'de gösterilmiştir.



Şekil 1: Obua C4 notasının özilinti ve ilintropi işlevleri.

Görüldüğü gibi, ilintropi işlevi ilinti işlevine göre daha dar bir tepeye sahiptir. Ayrıca ilintropi işlevinde, F0 bulunması amacıyla belirlenmesi gereken en büyük genlikli tepe, nispeten daha küçük genlikli diğer tepelerin bastırılmış olması nedeniyle daha belirgindir. Bu şekilde F0 daha hassas da ölçülebilir.

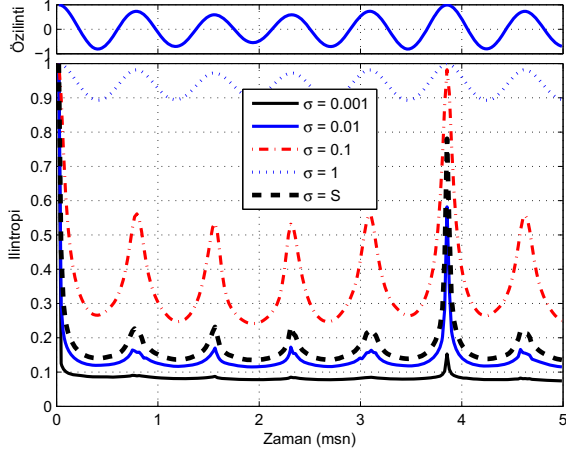
Bilindiği gibi ilintropi işlevinin hesaplanması çekirdek yöntemlerine dayanmaktadır. Bu yöntemlerde kullanılan çekirdek işlevinin ve parametresinin önemi bilinmekteyse de, nasıl seçileceği belirsizliğini korumaktadır. Ancak, ilintropi işlevi hesaplamasında kullanılan Gauss çekirdek işlevinin parametresinin uygun bir seçimi için [13]'de önerildiği gibi Silverman kuralı uygulanabilir. Bu kurala göre, N uzunluğundaki bir işaretin parametresi

$$\sigma_S = 0.9AN^{-1/5} \quad (7)$$

ile hesaplanabilmektedir. Bir dağılımda ya da sıralanmış bir veri dizisinde, verinin birinci çeyrek değer olarak bilinen ilk %25'i ile üçüncü çeyrek değer olarak bilinen %75'i arasında kalan aralığa çeyrek değerler genişliği denilmektedir. Denklem (7)'de A ile gösterilen değer, işaretin standart sapması ile çeyrek değerler genişliğinin 1.34 ile oranlanmasıyla bulunan değerden küçük olanıdır.

Şekil 1'de verilen Obua C4 nota örneğinin farklı σ değerleri için hesaplanan ilintropi işlevleri Şekil 2'de gösterilmiştir. Şekilde S ile belirtilen işlev, çekirdek parametresi değerinin Silverman kuralı ile hesaplanması sonucu elde edilen ilintropi işlevini göstermektedir.

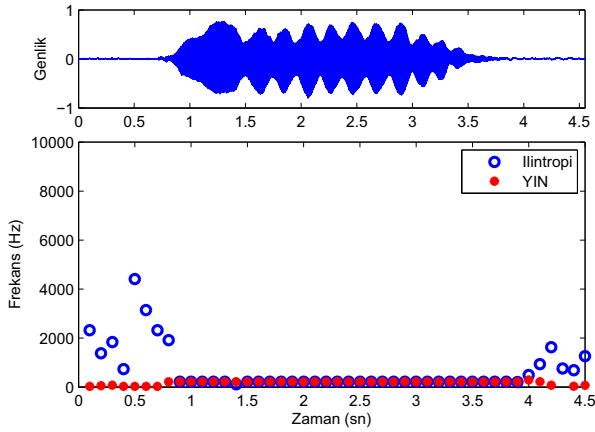
Şekil 2'den de görüldüğü gibi çekirdek parametresinin seçimi F0 bulunması için önem taşımakta, büyük σ değerleri için ilintropi işlevi özilinti işlevine yakınsamaktadır. Bu nedenle bundan sonraki örnekler için ilintropi işlevindeki çekirdek



Şekil 2: Obua C4 notasının farklı σ değerleri için ilintropi işlevleri.

parametre değeri, Denklem (7) kullanılarak Silverman kuralı ile belirlenmiştir.

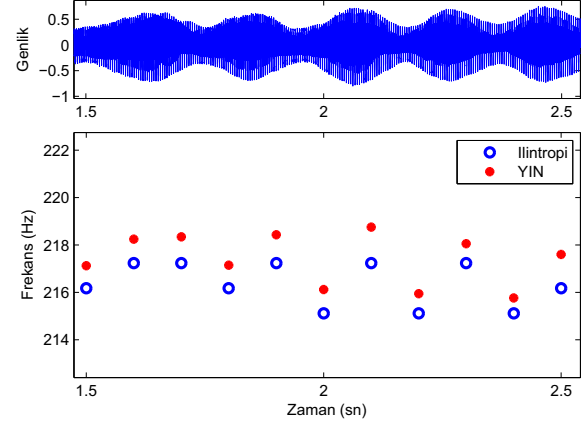
İlintropi işlevi ile bulunan F0 değerlerinin doğruluğunu denetlemek için, ilintropi işlevi ile YIN algoritması kullanılarak elde edilen F0 değerleri karşılaştırılmıştır. Bu amaçla kullanılan örnekler 0.1 saniye olarak belirlenen bir pencere boyutunda ele alınmış, her bir pencere için işaretin hem ilintropi işlevinden belirlenen hem de YIN algoritması ile bulunan F0 değerleri gösterilmiştir. Örneklerden ilki olarak seçilen Alto Flüt enstrumanının A3 nota örneği ve yapılan karşılaştırma Şekil 3'te verilmiştir.



Şekil 3: Alto Flüt A3 notası için ilintropi ve YIN algoritması karşılaştırılması.

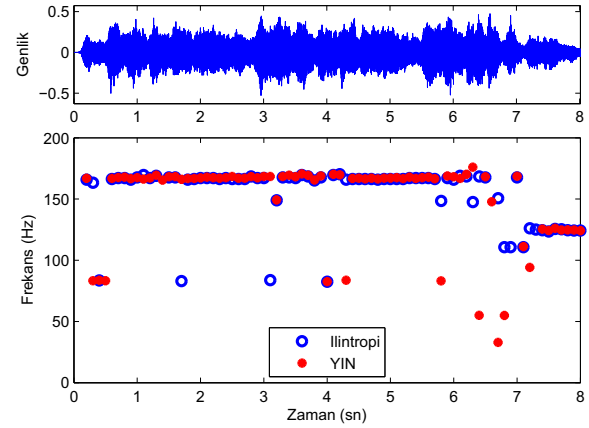
Ayrıklaştırılmış nota örneğinin başlangıcında ve sonunda bulunan bölgelerde farklılıklar olmasına rağmen, işaretin kararlı bulunduğu bölgede YIN algoritması ile A3 notasının frekans karşılığı olan 220 Hz çevresindeki aynı nota frekansları bulunmuştur. Küçük frekans farklılıklarının daha iyi

görülebilmesi amacıyla Şekil 4'te, Şekil 3'ün yakınlaştırılmış hali sunulmuştur. Frekans değerlerindeki farklılıkların birkaç Hz seviyesinden büyük olmaması ilintropi işlevinin başarısını ortaya koymaktadır.



Şekil 4: Alto Flüt A3 notası için ilintropi ve YIN algoritması karşılaştırılması, Şekil 3'ün yakınlaştırılmış hali.

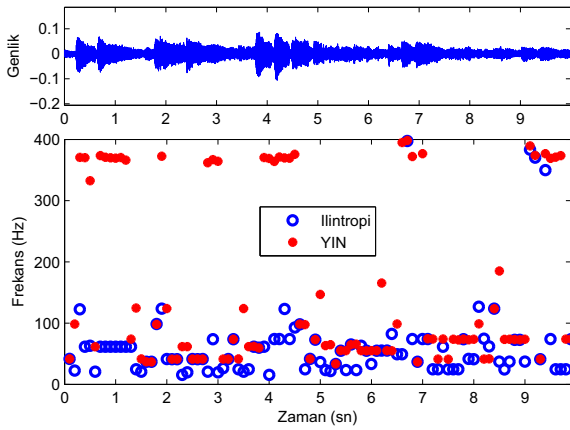
İlintropi işlevinin YIN algoritması ile karşılaştırıldığı bir diğer müzik işareti olarak, Vivaldi'nin Le Quattro Stagioni (Dört Mevsim) adlı eserinden keman enstrumanının bir örneği Şekil 5'te verilmiştir.



Şekil 5: Keman örneği için ilintropi ve YIN algoritması karşılaştırılması.

Görüldüğü gibi, tek bir nota işaretinin frekansı gibi belirli ve sabit bir F0 değerine sahip olmayan, zamanla farklı notaların bulunduğu bir müzik işaretinde de ilintropi işlevi YIN algoritması ile yakın değerler bulmaktadır. Bir başka örnek olarak da Rodrigo'nun Concierto de Aranjuez (Aranjuez konçertosu) eserinden bir gitar örneği kullanılmış ve karşılaştırma sonucu Şekil 6'da verilmiştir.

Benzer şekilde bu örnekte de ilintropi işlevinin YIN



Şekil 6: Gitar örneği için ilintropi ve YIN algoritması karşılaştırılması.

algoritması gibi F0 değerlerini bulmaktadır. Hızlı değişen F0 değerlerine sahip örnekler için pencere boyutunun seçiminin önemi ortaya çıkmaktadır. Elde edilen farklı sonuçların bazılarının frekans değerlerinin tam katlarında olduğu gözönüne alınarak oktav hatalarından oluştuğu, ayrıca gitarın çoksesliliğinin de bu duruma etkisi olduğu görülmektedir.

4. Değerlendirme

Bu çalışmada ilintropi kullanılarak müzik işaretlerinin notaya dökülmesi amacıyla F0 temel titreşim frekansları bulunmuş ve başarımlarını YIN algoritması ile karşılaştırılmıştır. Çekirdek parametresinin seçiminin önemli olduğu ilintropi işlevi kullanılarak özilinti işlevine göre daha dar tepe genişliği bulunarak, özilinti işlevine göre üstünlüğü gösterilmiştir. Her bir pencere için hesaplanan F0 değerinin zamandaki değişimi YIN algoritması ile yapılan karşılaştırmalar ile sunulmuş, aynı nota işaretinin frekansları elde edilerek yöntemin başarısı gösterilmiştir. Oluşan farklılıkların ve oktav hatalarının pencere boyutunun seçiminden kaynaklandığı, zamanda ve frekansta hassasiyet için uygun pencerelerin kullanılması gerekliliği görülmektedir. Ayrıca, verilen örneklerde tek bir frekans izlemesi gerçekleştirilmiş, bir pencerede birden çok frekans dikkate alınmamıştır. Müzik işaretlerinin notaya dökülmesi amacıyla birden çok frekansın izlenebildiği, bu amaçla da uygun pencere belirlenmesi üzerine çalışmalara devam edilecektir.

5. Kaynakça

- [1] 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.
- [2] A. Klapuri and M. Davy, Eds., *Signal Processing Methods for Music Transcription*. Springer, 2006.
- [3] A. de Cheveigné and H. Kawahara, "YIN, A fundamental frequency estimator for speech and music," *Journal of*

Acoustical Society of America, vol. 111, no. 4, pp. 1917–1930, April 2002.

- [4] G. Monti and M. Sandler, "Monophonic transcription with autocorrelation," in *Proc. of the COST G-6 Conference on Digital Audio Effects (DAFX-00)*, Verona, Italy, December 2000, pp. 257–260.
- [5] J. R. Deller, J. G. Proakis, and J. H. L. Hansen, *Discrete-Time Processing of Speech Signals*. New Jersey: Prentice Hall, 1987.
- [6] J. C. Brown and M. S. Puckette, "Calculation of a "narrowed" autocorrelation function," *Journal of Acoustical Society of America*, vol. 85, no. 4, pp. 1595–1601, April 1989.
- [7] R. Meddis and M. J. Hewitt, "Virtual pitch and phase sensitivity of a computer model of the auditory periphery. I: Pitch identification," *Journal of Acoustical Society of America*, vol. 89, no. 6, pp. 2866–2882, June 1991.
- [8] M. E. Özbek and F. A. Savacı, "İlintropi kullanarak müzik aletlerinin ayrıştırılması (Separation of musical instruments using correntropy)," in *IEEE 16th Signal Processing and Communications Applications Conference (SIU'08)*, Didim, Turkey, April 2008.
- [9] I. Santamaría, P. P. Pokharel, and J. C. Principe, "Generalized correlation function: Definition, properties, and application to blind equalization," *IEEE Trans. on Signal Processing*, vol. 54, no. 6, pp. 2187–2197, June 2006.
- [10] W. Liu, P. P. Pokharel, and J. C. Principe, "Correntropy: Properties and applications in non-Gaussian signal processing," *IEEE Trans. on Signal Processing*, vol. 55, no. 11, pp. 5286–5298, November 2007.
- [11] B. Schölkopf and A. J. Smola, *Learning with Kernels*. MIT Press, 2002.
- [12] J. Shawe-Taylor and N. Cristianini, *Kernel Methods for Pattern Analysis*. Cambridge University Press, 2004.
- [13] J.-W. Xu and J. C. Principe, "A novel pitch determination algorithm based on generalized correlation function," in *Proc. of the 17th IEEE Signal Processing Society Workshop on Machine Learning for Signal Processing (MLSP'07)*, Thessaloniki, Greece, August 2007, pp. 270–275.
- [14] —, "A pitch detector based on a generalized correlation function," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 16, no. 8, pp. 1420–1432, November 2008.
- [15] The University of Iowa Electronic Music Studios, <http://theremin.music.uiowa.edu>.
- [16] M. E. Özbek, C. Delpha, and P. Duhamel, "Musical note and instrument classification with likelihood-frequency-time analysis and support vector machines," in *15th European Signal Processing Conference (EUSIPCO)*, Poznań, Poland, September 2007, pp. 941–945.