

Ses Dosyalarının ENF Tabanlı Adli Analizine Kısa Zamanlı Fourier Dönüşümü Parametre Seçimlerinin Etkisi

The Effect of Short-Time Fourier Transform Parameters Choice on ENF-based Forensic Analysis of Audio

¹Ali Berk YALINKILIÇ , ²Saffet VATANSEVER 

^{1,2}Bursa Teknik Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi, Mekatronik Mühendisliği Bölümü, Bursa, TÜRKİYE

¹aliberk425@gmail.com, ²saffet.vatansever@btu.edu.tr (Sorumlu Yazar)

Araştırma Makalesi/Research Article

ARTICLE INFO

Article history

Received : 1 February 2023
Accepted : 11 April 2023

Keywords:

ENF,
Electric Network
Frequency,
Audio and Video Forensics,
Time-of-Recording,
Time-Stamping.

ABSTRACT

Due to the rapidly developing computer technology, digital audio, image, and video can be modified with various techniques. These modifications may be on the media content or the metadata. In this context, forensic analysis of digital media is of great importance. Electrical network frequency (ENF) based forensic analysis is a significant tool that can be used for checking file integrity and detecting time-of-recording. One of the most preferred methods for ENF signal estimation is the Short-Time Fourier Transform (STFT) based approach. The choice of STFT window size and STFT hop size can directly affect the accuracy of ENF signal to be estimated from media, and hence the performance of the ENF-related forensic applications. This work investigates how the STFT parameters choice affects the performance in time-of-recording verification. A comparative analysis is made for various STFT window sizes and STFT hope sizes by experimenting with different audio lengths.

© 2023 Bandırma Onyedi Eylül University, Faculty of Engineering and Natural Science. Published by Dergi Park. All rights reserved.

MAKALE BİLGİSİ

Makale Tarihleri

Gönderim : 1 Şubat 2023
Kabul : 11 Nisan 2023

Anahtar Kelimeler:

ENF,
Elektrik Şebeke Frekansını,
Ses ve Videoların Adli
Analizi,
Kayıt Zamanı,
Zaman Damgası.

ÖZET

Hızla gelişmekte olan bilgisayar teknolojisi sayesinde çeşitli tekniklerle dijital ses, görüntü ve video dosyaları üzerinde modifikasyonlar yapılabilmektedir. Bu modifikasyonlar doğrudan dosya içeriğinde olabileceği gibi bazen de meta data üzerinde olabilmektedir. Bu bağlamda dijital medyaların adli analizi büyük önem arz etmektedir. Elektrik şebeke frekansını (ENF) tabanlı adli analiz yaklaşımı dosya bütünlük kontrolünde ve dosyaların kayıt zamanı tespitinde kullanılabilen önemli bir araçtır. ENF sinyali kestiriminde en çok tercih edilen yöntemlerden biri, kısa zamanlı Fourier dönüşümü (Short-Time Fourier Transform - STFT) temelli yaklaşımdır. STFT yönteminde, pencere boyutu ve kaydırma miktarı parametrelerinin seçimi büyük öneme sahip olup, kestirimi yapılan ENF sinyali doğruluğunu, dolayısıyla da ENF tabanlı adli analiz uygulamalarının performansını doğrudan etkileyebilmektedir. Bu çalışmada, STFT parametreleri seçiminin, ENF tabanlı dosya kayıt zamanı doğrulamada performansa ne derece etki ettiği araştırılmıştır. Farklı uzunluktaki ses dosyaları, çeşitli STFT pencere boyutu ve STFT kaydırma miktarlarına göre ayrı ayrı test edilerek karşılaştırmalı bir analiz yapılmıştır.

© 2023 Bandırma Onyedi Eylül Üniversitesi, Mühendislik ve Doğa Bilimleri Fakültesi. Dergi Park tarafından yayınlanmaktadır. Tüm Hakları Saklıdır.

1. GİRİŞ

Günümüz bilgi çağında dijital kayıtların çok kolay bir şekilde manipüle edilebilir ya da yapay olarak oluşturulabilir olması; ses, görüntü ve video dosyalarının adli analizini önemli bir araştırma konusu haline getirmiştir. Bu bağlamda, elektrik şebeke frekansı (ENF - Electric Network Frequency) tabanlı adli analiz son yıllardaki en popüler ve en önemli araştırma konuları arasında göze çarpmaktadır. ENF, elektrik şebeke geriliminin frekansı olup, harcanan ve üretilen güçteki farklılığa bağlı olarak bir nominal değer (Türkiye’de 50 Hz) etrafında (genelde ± 0.1 Hz) sürekli dalgalanmalar yapar [1]. Enterkonnekte bir elektrik şebekesi üzerindeki tüm jeneratörlerin senkron çalışması sonucu şebeke üzerindeki tüm noktalarda aynı ENF değişimleri gözlenir [1].

ENF, elektrik şebeke gerilimi kaynaklı elektromanyetik alan ya da akustik gürültünün (akustik şebeke gürültüsü) var olduğu ortamlarda yapılan ses kayıtlarına istemsiz olarak girişim yapar [2-7]. Ayrıca, ENF, elektrik şebekesinden beslenen bir ışık kaynağı ile aydınlatılan ortamlarda yapılan video kayıtlarına gömülür [8-19]. Zaman veya frekans uzayı temelli çeşitli tekniklerle bu kayıtlardan ENF’in zamana bağlı değişimleri (ENF sinyali) yakalanabilir [2, 6, 20-23].

ENF sinyali; dosya kayıt zamanı doğrulama, içerik doğrulama, videolarda ses ve görüntü senkronizasyonu, elektrik şebekesi tanılama gibi çeşitli adli analiz işlemlerinde kullanılabilir. Örneğin; adli bir vakada suçun işlendiği tarihte, olay yerinden başka bir yerde (ev, işyeri, çocuk parkı, cadde, sokak vs.) sanığa ait güvenlik kamerası ile çekilmiş bir video tespit edilip bu videonun kayıt zamanı zamanı doğrulanabilirse ilgili tarihte şüpheli kişinin başka yerde olduğu kanıtlanabilir. Bir başka örnek uygulama olarak; bilinmeyen bir sebeple gizli kalmış geçmişe ait video dosyalarının gün yüzüne çıkması durumunda dosyaların orijinalliği, bütünlüğü ve çekildiği tarih ve saat aydınlığa kavuşturulabilir.

Elektrik üretimindeki arz ve talep dengesizliği, dolayısıyla da ENF’teki değişimler zaman içinde kısa süreliğine de olsa benzerlikler gösterse de zaman uzadıkça bu değişimler eşsiz hale gelir. Bu sebeple, bir ses ya da video dosyasının ENF tabanlı adli analiz uygulamalarının performansı kayıt dosyasının uzunluğu ile, dolayısıyla da dosyadan kestirimi sağlanan ENF zaman serisinin (ENF sinyalinin) uzunluğu ile doğru orantılıdır.

Medya dosyalarından (ses ya da video) ENF sinyali kestiriminde sıklıkla tercih edilen en önemli yöntemlerden biri kısa zamanlı Fourier dönüşümü (Short-Time Fourier Transform - STFT) tabanlı yaklaşımdır [2, 3, 22-25]. STFT yöntemi ile ENF sinyali kestiriminde, STFT pencere boyutu ve kaydırma miktarı parametrelerinin seçimi, medyadan elde edilen ENF sinyalinin doğruluğuna, dolayısıyla da ENF tabanlı adli analiz uygulamalarının performansına doğrudan etki edebilmektedir.

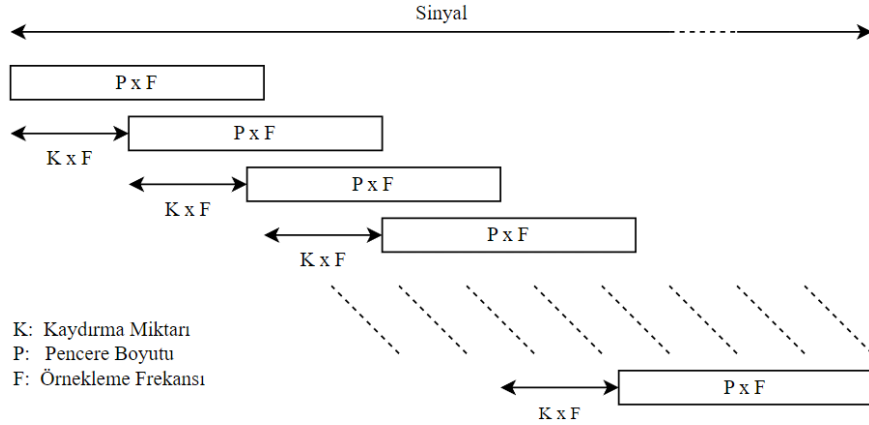
Bu çalışmada, farklı uzunluktaki ses dosyaları için STFT parametre seçimlerine göre ENF tabanlı kayıt zamanı doğrulamada performansın nasıl etkilendiği üzerine kapsamlı bir araştırma yapılmıştır. Bu amaçla, ENF-WHU veri tabanındaki ([22, 23]) 130 adet ses dosyasının ilk 90 sn., ilk 120 sn., ilk 180 sn., ve ilk 240 sn.’lik klipleri 8 sn. ile 64 sn. arasında değişen STFT pencere boyutları ve 1 sn. ile 32 sn. arasında değişen STFT kaydırma miktarları için ayrı ayrı test edilmiş ve performans analizi yapılmıştır. Daha iyi performans elde edebilmek için, hangi koşullarda hangi parametre seçimlerinin yapılmasının daha uygun olabileceği araştırılmıştır.

2. KURAMSAL TEMELLER

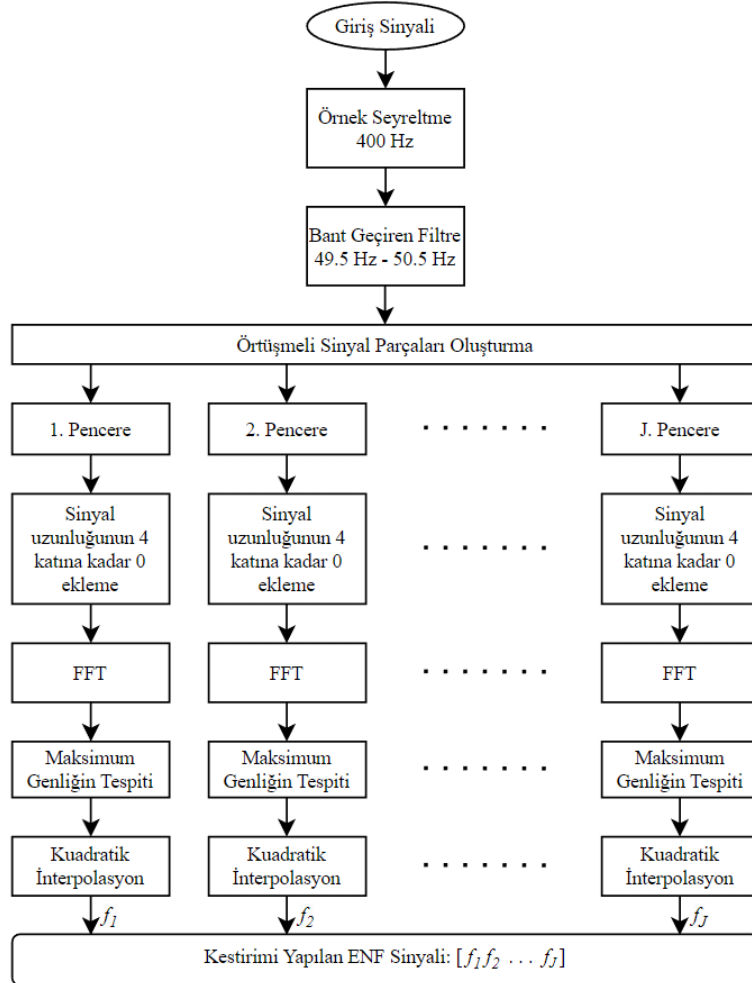
Sayısal medya dosyalarının (ses ya da video) adli analizi kapsamında kullanılmak üzere, ENF sinyali kestiriminde en çok tercih edilen yöntemlerden biri kısa zamanlı Fourier dönüşümü (Short-Time Fourier Transform - STFT) yaklaşımıdır [2, 3, 22-25]. Kısa zamanlı Fourier dönüşümü, bir sinyalin frekans ve faz bileşenlerinin zamana bağlı değişimlerini hesaplamada kullanılabilen önemli bir yöntemdir. STFT, birbirini izleyen iki temel aşamadan oluşmaktadır. İlk aşamada, STFT uygulanacak ana sinyal belirli oranda örtüşmeli olarak eşit uzunluktaki sinyal parçalarına ayrılır. Bu işlem, Şekil 1’de resmedilmiştir. Burada pencere boyutu (P), her bir sinyal parçasının saniye cinsinden uzunluğunu belirler. Sinyalin her 1 saniyelik kısmındaki örnek sayısı (örnekleme frekansı) F ile ifade edilirse, her bir sinyal parçası $F \times K$ örnekten oluşmaktadır. Kaydırma miktarı (K), saniye cinsinden kaydırma boyutunu belirler. Buna göre ardışık iki sinyal parçası arasındaki örtüşme miktarı $(P-K) \times F$ örnek kadardır. STFT’nin ikinci aşamasında, hızlı Fourier dönüşümü (Fast Fourier Transform - FFT) algoritması yardımı ile, ilk aşamada elde edilen her bir sinyal parçasının ayrık Fourier dönüşümü hesaplanır. Ardından, her bir dönüşüm için istenen frekans bandındaki frekans ve/veya faz bilgisi elde edilir.

STFT tabanlı ENF kestirim algoritması ([3]), Şekil 2’de blok diyagramı halinde sunulmuştur. Bu algoritmada ilk olarak gereksiz işlem yükünden kurtulmak için sinyalin örnekleme frekansı nominal ENF frekansının en az 2 katı olacak şekilde azaltılır (Nyquist kriteri [26]). ENF harmoniklerinin de kullanılabilmesi düşünülerek örnekleme frekansını 400 Hz’e indirgemek iyi bir seçim olacaktır. Ardından, aşağı örneklenmiş sinyal merkez frekansı nominal frekans olan bir bant geçiren filtreden geçirilir. ENF’in nominal değerden ± 0.1 Hz kadar sapmalar yapabileceği ([1]) düşünüldüğünde, 49.5 Hz ile 50.5 Hz aralığındaki bir frekans geçirme bandı, bu filtre için iyi bir seçim olacaktır. Filtreleme işlemi sonucu ilgilenilmeyen frekanslardan yüksek oranda arındırılan sinyal, J adet örtüşmeli sinyal parçasına bölünür (Şekil 1’de olduğu gibi). Burada J , seçilen STFT pencere boyutu ve kaydırma miktarına göre sinyalden elde edilebilen sinyal parçalarının sayısını ifade eder. Her bir sinyal parçasından 1 ENF örneği elde edileceği düşünüldüğünde, J aynı zamanda seçilen STFT parametrelerine göre sinyalden elde edilebilecek ENF örneği sayısını ifade eder. Daha sonra, her bir sinyal parçasının sonuna sıfırlar eklenip (zero padding) FFT’si alınır. FFT işleminden önce sıfır ekleme işlemi, sinyal parçasının frekans spektrumuna ait frekans

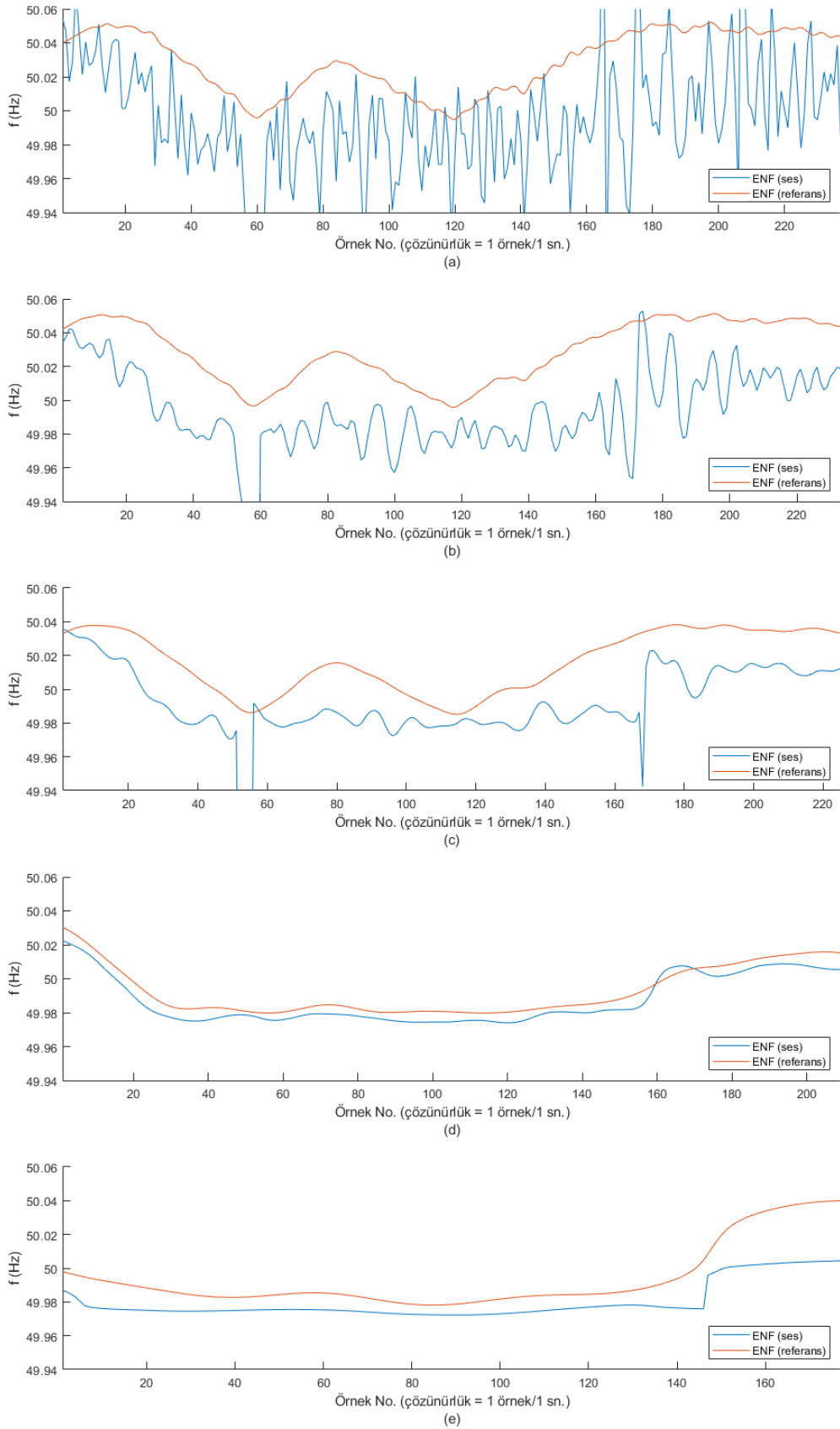
noktaları arasındaki mesafeyi azaltarak daha küçük frekans değişimlerini yakalayabilmeye olanak sağlar. Ancak, gereğinden fazla sıfır eklemek hafıza kullanımını önemli ölçüde artırır. Dahası, hafıza yetersizliğinde diskin hafızasının kullanımına, sonrasında da işlemlerin durma seviyesine gelecek kadar yavaşlamasına sebep olur. Bu sebeple, eklenen sıfırlarla birlikte sinyal boyutu 4 katını geçmeyecek şekilde ekleme yapmak uygun olacaktır. Sıfır ekleme işlemi ile yeterince küçültülemeyen frekans aralıkları kuadratik (ikinci derece) interpolasyon ile istenen seviyelere getirilir [3]. Bir sinyal parçası için uygulanan FFT ve kuadratik interpolasyon sonucu, frekans spektrumunda maksimum genliğe karşılık gelen frekans değeri o sinyal parçası için hesaplanmış ENF örneğine karşılık gelir. Sırasıyla her bir parça için ENF örneklerinin elde edilmesi sonucu ENF zaman serisi, yani ENF sinyali hesaplanmış olur.



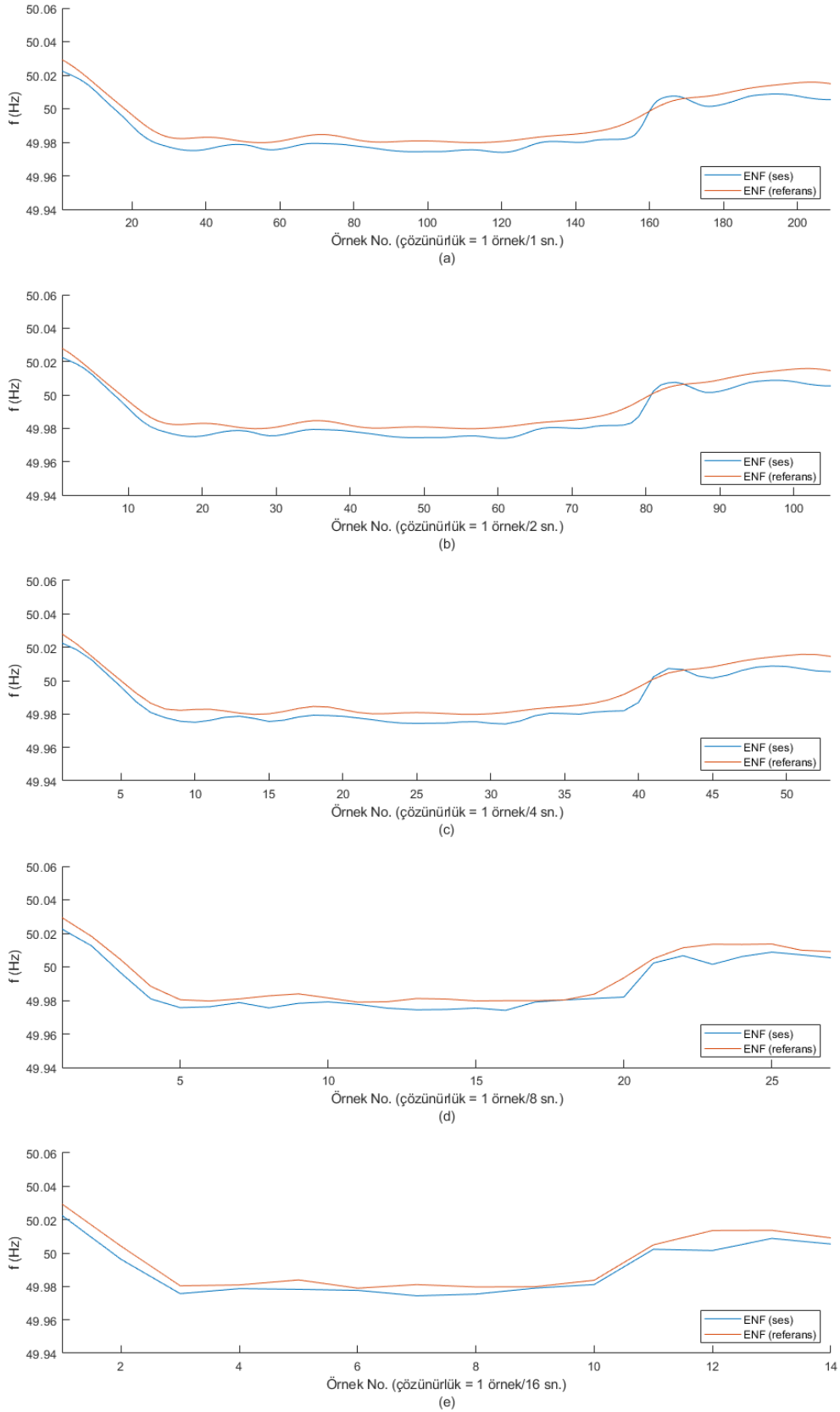
Şekil 1. Kısa zamanlı Fourier dönüşümünde örtüşmeli sinyal parçaları oluşturma işlemi [3].



Şekil 2. Kısa Zamanlı Fourier Dönüşümü Tabanlı ENF Kestirim Algoritması [3].



Şekil 3. 1 sn. uzunluklu STFT kaydırma miktarı için farklı STFT pencere boyutu ((a) 4 sn. (b) 8 sn. (c) 16 sn. (d) 32 sn. (e) 64 sn.) seçimlerine göre örnek bir ses dosyasından kestirimi sağlanan ENF sinyalleri ile aynı parametre seçimlerine göre elde edilmiş referans ENF sinyalleri karşılaştırması. (a), (b), ve (c) için ses dosyasının kayıt zamanı doğrulanamamış olup hatalı eşleşme elde edilmiştir. (d) ve (e) için ses dosyasının kayıt zamanı doğrulanabilmiştir.



Şekil 4. 32 sn. uzunluklu STFT pencere boyutu için farklı STFT kaydırma miktarı ((a) 1 sn. (b) 2 sn. (c) 4 sn. (d) 8 sn. (e) 16 sn.) seçimlerine göre örnek bir ses dosyasından kestirimi sağlanan ENF sinyalleri ile aynı parametre seçimlerine göre elde edilmiş referans ENF sinyalleri karşılaştırması. (a), (b), ve (c) için ses dosyasının kayıt zamanı doğrulanabilmiştir. (d) ve (e) için ses dosyasının kayıt zamanı doğrulanamamış olup hatalı eşleşme elde edilmiştir.

3. ENF HESABINDA STFT PARAMETRELERİ SEÇİMİNİN ÖNEMİ VE ETKİSİ

Ses ya da video dosyalarının ENF tabanlı adli analiz uygulamalarında elde edilen performans, kestirimi sağlanan ENF sinyalinin doğruluğu ve çözünürlüğü ile doğrudan ilişkilidir. Bu sebeple, incelenen medya dosyasından yüksek doğrulukta ve uygun çözünürlükte ENF hesabı kritik önem arz etmektedir. Medya dosyalarından STFT yaklaşımı ile ENF sinyali kestiriminde, STFT pencere boyutu ve kaydırma miktarı parametrelerinin seçimi hesaplanan ENF sinyalinin çözünürlüğünü belirleyen ve doğruluğuna etki eden iki anahtar etmendir.

STFT tabanlı ENF sinyali kestiriminde seçilen pencere boyutu ve kaydırma miktarı, sinyalden elde edilebilecek örtüşen sinyal parçası sayısını (Şekil 1), dolayısıyla da medyadan elde edilebilecek ENF zaman serisinin uzunluğunu (kestirilen ENF sinyalindeki örnek sayısını) tayin eder. STFT kaydırma miktarı tek başına, ENF çözünürlüğünü, yani kaç saniye aralıklarla ENF örneklerinin hesaplanacağını belirler. STFT pencere boyutu, örtüşmeli sinyal parçalarının uzunluğunu (Şekil 1), dolayısıyla da her bir ENF örneği kestiriminde (Şekil 2) kullanılan veri sayısını belirler. Buna göre, pencere boyutunun çok küçük ya da çok büyük olması, ENF hesabını olumsuz yönde etkileyebilmektedir. Çok küçük bir pencere boyutu, FFT işleminde yetersiz veri kullanımına bağlı olarak düşük güçlü ve gürültülü frekans bileşenleri elde etmeye, dolayısıyla da kestirimi sağlanan ENF sinyalinde bozulmalara neden olabilmektedir. Çok büyük bir pencere boyutu kullanımı ise yüksek güçlü frekans bileşenleri elde etmeyi sağlasa da ENF değişimindeki detayları kaybetmeye neden olabilir.

Özet olarak, STFT parametre seçimlerine göre; ENF çözünürlüğü, ENF örneği hesabında kullanılan verinin boyutu ve elde edilen ENF örneği sayısı şekillenmekte ve bunlara bağlı olarak medyadan elde edilen ENF sinyalinin doğruluğu ve hassasiyeti değişebilmektedir. Bu bağlamda, ses ve video dosyalarının kayıt zamanının ENF sinyali kullanılarak doğrulanmasında STFT parametrelerinin doğru seçimi, kestirilen ENF sinyali ve referans ENF sinyali arasındaki ilintiyi doğru noktada yakalamaya yardımcı olmaktadır.

Şekil 3'te, 1 sn. uzunluklu sabit STFT kaydırma miktarı için sırasıyla 4 sn., 8 sn., 16 sn., 32 sn. ve 64 sn.'lik STFT pencere boyutu seçimlerine göre 240 sn. uzunluktaki örnek bir ses dosyasından ayrı ayrı kestirimi sağlanan ENF sinyalleri ile aynı parametre seçimlerine göre hesaplanmış referans ENF sinyalleri görülmektedir. Buna göre, 32 sn. ve 64 sn. uzunluklu STFT pencereleri kullanımı ile ses dosyasından yüksek doğrulukta ENF kestirimi sağlanabildiği halde, 4 sn., 8 sn. ve 16 sn.'lik STFT pencereleri kullanılarak kestirimi sağlanan ENF sinyalleri oldukça gürültülü olarak elde edilmiştir. Hesaplanan ENF sinyallerinin referans sinyale benzerliğini ölçebilmek için açık ifadesi aşağıda verilen normalize çapraz korelasyon işlemi kullanılmıştır.

$$c(k) = \frac{\sum_{n=1}^N [F_r(n) - \mu_r] [F_s(n-k) - \mu_s]}{\sqrt{\sum_{n=1}^N [F_r(n) - \mu_r]^2 \sum_{n=1}^N [F_s(n) - \mu_s]^2}} \quad (1)$$

Burada, F_r ve F_s sırasıyla referans ENF ve ses dosyalarında kestirimi sağlanan ENF sinyallerini, μ_r ve μ_s ise ortalama değerleri ifade eder. Normalize çapraz korelasyon sonucu, ses ENF sinyali ile referans ENF sinyali (ses kaydının yapıldığı zaman periyodunu içeren güne ait) arasındaki en yüksek benzerlik (normalize çapraz korelasyon işlemi ile) 32 sn. ve 64 sn. uzunluklu STFT pencereleri için doğru noktada (lag point) tespit edilerek ses dosyasının kayıt zamanı doğrulanabilmiştir. 4 sn., 8 sn. ve 16 sn. uzunluklu STFT pencereleri ile hesaplanan maksimum korelasyon katsayısı ise yanlış noktada (lag point) bulunmuş olup ses dosyasının kayıt zamanı doğrulanamamıştır (Kayıt zamanı doğrulama işlemi Bölüm 4'te daha detaylı anlatılmıştır). Referans ENF sinyallerinin, küçük boyutlu STFT pencereleri için dahi yüksek doğrulukta hesaplanabilmiş olması, zayıflatılmış şebeke sinyalinden doğrudan elde edilmelerinin doğal bir sonucudur (ENF frekans bandına girişim yapan başka sinyal yoktur).

Şekil 4'te, 32 sn. uzunluklu sabit STFT pencere boyutu için sırasıyla 1 sn., 2 sn., 4 sn., 8 sn. ve 16 sn.'lik STFT kaydırma miktarı seçimlerine göre 240 sn. uzunluktaki örnek bir ses dosyasından ayrı ayrı kestirimi sağlanan ENF sinyalleri ile aynı parametre seçimlerine göre hesaplanmış referans ENF sinyalleri görülmektedir. 1 sn., 2 sn. ve 4 sn. uzunluklu kaydırma miktarları kullanılarak ses dosyasından kestirimi sağlanan ENF sinyalleri ile ses dosyasının kayıt zamanı doğrulanabilmiştir. Buna karşın, 8 sn. ve 16 sn. uzunluklu kaydırma miktarları kullanımı ile ses dosyasının kayıt zamanı doğrulanamamış olup hatalı eşleşme elde edilmiştir. Çok yüksek kaydırma miktarları için doğru eşleşme bulunamamış olması, yetersiz sayıda ENF örneği elde edilmesinden kaynaklanmaktadır. 8 sn.'lik kaydırma miktarı için 8 sn.'de 1 ENF örneği, 16 sn.'lik kaydırma için ise 16 sn.'de 1 ENF örneği hesaplanmış olup frekans değişimindeki detaylar tam olarak yakalanamamıştır.

4. BULGULAR

Bu bölümde, STFT pencere boyutu ve kaydırma miktarı seçiminin ENF temelli medya kayıt zamanı doğrulama performansını nasıl etkilediği, ENF-WHU veri tabanına ([22, 23]) ait uzunlukları 4.5 dk. ile 16 dk. arasında değişkenlik gösteren 130 adet ses dosyası üzerinde yapılan testlerle incelenmiştir. Ses dosyalarının tamamı akustik şebeke gürültüsü ya da elektromanyetik alan içeren ortamlarda kaydedilmiş olup, ENF içerdiği bilinmektedir. ENF-WHU veri tabanı ayrıca, bu ses kayıtlarının yapıldığı günlere ait zayıflatılmış şebeke sinyali kayıtlarını (referans (ground-truth) ENF sinyali kestirimi için) da içermektedir.

ENF-WHU veri tabanındaki her bir ses dosyasının sırasıyla ilk 240 sn., ilk 180 sn., ilk 120 sn. ve ilk 90 sn.'lik kliplerinden kısa zamanlı Fourier dönüşümü (STFT) yardımıyla farklı pencere boyutları (8 sn., 16 sn., 32 sn., 64 sn.) ve seçilen pencere boyutunun en fazla yarısı olacak şekilde farklı kaydırma miktarlarına göre ayrı ayrı

kestirimi yapılan ENF sinyalleri, aynı yöntemle zayıflatılmış şebeke sinyali kaydından (ses dosyalarına ait kayıt zaman periyodunu kapsayan) elde edilen referans ENF sinyalleri (24 saatlik) ile normalize çapraz korelasyon işlemine sokulmuş ve maksimum korelasyon katsayısının yeri analiz edilmiştir. Maksimum korelasyon katsayısının bulunduğu nokta ses kaydının başlatılma anına tekabül ediyorsa, yani referans ENF sinyalinin başlangıç anına maksimum korelasyon katsayısının bulunduğu zaman noktası (lag point) eklendiğinde (saniye cinsinden) ses kaydının başlangıç anına karşılık geliyorsa (30 sn.'lik tolerans dahilinde) kayıt zamanının doğru bir şekilde tespit edildiği sonucuna ulaşılır.

Tablo 1. 130 adet ses dosyası üzerinde yapılan ENF tabanlı kayıt zamanı doğrulama analizinde, **64 sn.** uzunluklu STFT pencere boyutu için farklı STFT kaydırma miktarı ve farklı klip uzunluğu seçimine göre elde edilen doğru eşleşme sayıları.

Klip Uzunluğu	STFT Kaydırma Miktarına Göre Doğru Eşleşme Sayısı (130 dosyada)					
	1 sn.	2 sn.	4 sn.	8 sn.	16 sn.	32 sn.
Tamamı	117	118	115	115	112	100
268 sn.	107	109	108	103	68	46
240 sn.	99	99	99	91	60	27
180 sn.	54	56	55	39	26	5
120 sn.	7	7	7	4	4	1
90 sn.	2	1	0	1	0	-

Tablo 2. 130 adet ses dosyası üzerinde yapılan ENF tabanlı kayıt zamanı doğrulama analizinde, **32 sn.** uzunluklu STFT pencere boyutu için farklı STFT kaydırma miktarı ve farklı klip uzunluğu seçimine göre elde edilen doğru eşleşme sayıları.

Klip Uzunluğu	STFT Kaydırma Miktarına Göre Doğru Eşleşme Sayısı (130 dosyada)				
	1 sn.	2 sn.	4 sn.	8 sn.	16 sn.
Tamamı	115	115	113	113	107
268 sn.	97	99	99	100	78
240 sn.	99	102	102	99	73
180 sn.	97	99	98	86	51
120 sn.	71	71	64	40	12
90 sn.	43	43	32	18	1

Tablo 3. 130 adet ses dosyası üzerinde yapılan ENF tabanlı kayıt zamanı doğrulama analizinde, **16 sn.** uzunluklu STFT pencere boyutu için farklı STFT kaydırma miktarı ve farklı klip uzunluğu seçimine göre elde edilen doğru eşleşme sayıları.

Klip Uzunluğu	STFT Kaydırma Miktarına Göre Doğru Eşleşme Sayısı (130 dosyada)			
	1 sn.	2 sn.	4 sn.	8 sn.
Tamamı	105	102	98	93
268 sn.	80	83	85	81
240 sn.	81	84	84	80
180 sn.	80	82	82	71
120 sn.	69	69	67	44
90 sn.	54	51	42	21

Tablo 4. 130 adet ses dosyası üzerinde yapılan ENF tabanlı kayıt zamanı doğrulama analizinde, **8 sn.** uzunluklu STFT pencere boyutu için farklı STFT kaydırma miktarı ve farklı klip uzunluğu seçimine göre elde edilen doğru eşleşme sayıları.

Klip Uzunluğu	STFT Kaydırma Miktarına Göre Doğru Eşleşme Sayısı (130 dosyada)		
	1 sn.	2 sn.	4 sn.
Tamamı	91	93	92
268 sn.	68	67	67
240 sn.	67	67	63
180 sn.	61	68	66
120 sn.	59	60	51
90 sn.	44	44	36

Tablo 1, Tablo 2, Tablo 3, ve Tablo 4'te; sırasıyla 64 sn., 32 sn., 16 sn. ve 8 sn. uzunluklu pencere boyutları için farklı klip uzunluklarına ve farklı kaydırma miktarlarına göre elde edilen doğru eşleşme sayıları görülmektedir. Buna göre, belirli bir STFT pencere boyutu için kaydırma miktarının 1 sn. ya da 2 sn. seçimi klip uzunluğu fark etmeksizin performansta önemli bir değişime sebep olmamıştır. 4 sn.'lik bir kaydırma miktarı seçildiğinde ise, pencere boyutu ve klip uzunluğuna bağlı olarak performansta kısmi farklılıklar görülmüştür. Bu bulgu, ENF'in çok ani değişimler göstermemesinin doğal bir sonucu olarak düşünülebilir. Kaydırma miktarının 4 sn. 'den daha büyük seçilmesi, klip uzunluğu fark etmeksizin doğru eşleşme sayısında genel bir azalmaya, dolayısıyla da performansta bir düşüşe sebep olmuştur.

Uzun klipler (240 sn.'den büyük) için, STFT pencere boyutu büyüdükçe doğru eşleşme sayısında genel bir artış görülmüştür. Diğer bir deyişle, bu uzunluktaki klipler için pencere boyutunu yüksek seçmek performansı kayda

değer oranda arttırmıştır. Ancak kısa klipler (180 sn. ve 180 sn.'den küçük) için STFT pencere boyutu büyüdükçe performansın düştüğü, yani doğru eşleşme sayısında bir azalma gözlenmiştir. Bu bulgu, kısa kliplerde büyük pencere boyutunu seçiminin hesaplanan, dolayısıyla da benzerlik testinde değerlendirmeye katılan ENF örneği sayısının önemli ölçüde azalmasının doğal bir sonucu olarak düşünülebilir.

240 sn.'den büyük ses klipleri için en iyi performansın 64 sn.'lik pencere boyutu için elde edildiği görülmüştür. 240 sn. ile 120 sn. aralığındaki ses klipleri için en yüksek performans 32 sn. uzunluklu pencere boyutu kullanıldığında elde edilmiştir. 120 sn.'lik ses klipleri için 16 sn. uzunluklu pencere boyutu kullanılarak elde edilen performans, 32 sn. uzunluklu pencere boyutu kullanılarak elde edilen performansa çok yakındır. 90 sn.'lik ses klipleri için en iyi performans 16 sn. uzunluklu pencere boyutu kullanılarak elde edilmiştir. Sinyal boyutu düştükçe performanstaki genel düşüşün sebebi, ENF sinyalinin kısa aralıklarla benzer örüntüler göstermesinden kaynaklanır. Sinyal boyutu büyüdükçe, yani kestirimi sağlanan ENF zaman serisinin uzunluğu arttıkça, bu benzerlikler azalmakta ve gittikçe daha emsalsiz hale gelmektedir.

5. TARTIŞMA

Guang Hua ve arkadaşları tarafından ENF-WHU veri tabanındaki videoların tamamı üzerinde tüm klip uzunlukları kullanılarak 16 sn.'lik STFT penceresi ve 1 sn.'lik STFT kaydırma miktarı seçimi ile yapılan karşılaştırmalı deneylerde ([23]); [27]'deki yöntem ile 105 eşleşme, [20]'deki yöntem ile 116 eşleşme, [23]'teki yöntem ile ise 129 eşleşme bulunmuştur. Karşılaştırılan bu yöntemlerin her biri, çeşitli filtreleme işlemleri ve veri iyileştirmeleri uygulanarak ve çoklu sayıda harmonik kullanarak elde edilen sonuçlardır. Bu çalışmada elde edilen sonuçlar ise, tek bir harmonik üzerinden dosyalara hiçbir ön işleme uygulamadan yalnız STFT ile elde edilen sonuçlardır. Bu durumda bile, 64 sn. pencere boyutu ve 2 sn. kaydırma miktarı için 118 eşleşme (Tablo 1) ile, [20] ve [27]'deki yöntemlerin 16 sn. pencere ve 1 sn. kaydırma kullanılarak uygulanmış durumuna göre daha iyi performans elde edilmiştir. [20] ve [27] için daha doğru parametre seçimleri yapılmış olsa sonuçlarda kayda değer ölçüde iyileşmeler olması beklenmektedir. Dahası, bu veri tabanı üzerindeki her bir ses dosyasının kendi uzunluğuna özgü farklı STFT parametre seçimlerinin yapılması ile [23]'e yakın ya da daha iyi sonuçlar elde edilmesi ihtimal dâhilindedir.

6. SONUÇ

Bu çalışmada, medyadan (ses ya da video dosyası) elektrik şebeke frekansı (ENF) sinyali kestiriminde tercih edilen en önemli yöntemlerden biri olan kısa zamanlı Fourier dönüşümü (STFT) yaklaşımında; STFT pencere boyutu ve kaydırma miktarı parametreleri seçiminin, hesaplanan ENF sinyali doğruluğuna ve dolayısıyla da ENF tabanlı adli analiz uygulamaları performansına ne derece etki edebildiği incelenmiştir. Bu amaçla ENF-WHU veri tabanındaki ([22, 23]) 130 adet ses dosyasının sırasıyla ilk 240 sn., ilk 180 sn., ilk 120 sn. ve ilk 90 sn.'lik kliplerinden farklı STFT pencere boyutu (8 sn., 16 sn., 32 sn. ve 64 sn.'lik) ve kaydırma miktarlarına (pencere boyutunun en fazla yarısı olacak şekilde 1 sn., 2 sn., 4 sn., 8 sn., 16 sn. ve 32 sn.'lik) göre ayrı ayrı hesaplanan ENF sinyalleri ile kayıt zamanı doğrulama testleri yapılmış ve her birinin performansına etkisi analiz edilmiştir. 240 sn.'den daha büyük ses klipleri için en yüksek performans 64 sn. uzunluklu pencere boyutu kullanıldığında elde edilmiştir. 120 sn. ile 240 sn. aralığındaki ses klipleri için en iyi performansın 32 sn.'lik pencere boyutu için elde edildiği görülmüştür. 90 sn.'lik ses klipleri için ise en iyi performansın 16 sn. uzunluklu pencere boyutu kullanılarak elde edildiği görülmüştür. Belirli bir pencere boyutu için kaydırma miktarını arttırmak performansın genel olarak düşmesine sebep olmakla birlikte en iyi sonuçlar 2 sn.'lik kaydırma miktarı için elde edilmiştir. 180 sn.'den büyük klipler için kaydırma miktarının 4 sn. seçimi de benzer sonuçlar vermiştir. Ancak kaydırma miktarının 4 sn.'den büyük seçilmesi klip uzunluğundan bağımsız olarak performansın kayda değer oranda düşürdüğü görülmüştür.

Bu çalışmada, benzer konuda çalışmalar yapacak araştırmacılara kayıt süresi uzunluğuna bağlı olarak STFT parametre seçimlerine dikkat edilmesi yönünde ışık tutması amaçlanmış ve literatürde sıklıkla kullanılan 2'nin katları olan pencere boyutları ve kaydırma miktarları analiz edilmiştir. Gelecek çalışmalarda, 8 sn. ile 64 sn. arasında test edilmeyen diğer pencere boyutları analiz edilerek performansı optimize eden parametrelerin tespit edilmesi hedeflenmiştir.

Yazar Katkıları

Ali Berk YALINKILIÇ - Deneyleri yapmış ve makaleyi yazmıştır.

Saffet VATANSEVER - Araştırmanın ortaya çıkmasını sağlamış; deneylerin yapılmasında, makale yazımında, düzeltmelerinde ve iyileştirmelerinde yardımcı olmuştur.

Çıkar Çatışması

Makale yazarları aralarında herhangi bir çıkar çatışması olmadığını beyan ederler.

Teşekkür

Bu çalışma Bursa Teknik Üniversitesi Bilimsel Araştırma Projeleri birimi tarafından 211N022 numaralı proje kapsamında desteklenmiştir.

KAYNAKÇA

- [1] M.H. Bollen and I.Y. Gu “Signal Processing of Power Quality Disturbances,” John Wiley and Sons, 2006.
- [2] C. Grigoras “Digital audio recording analysis—the electric network frequency criterion”, *International Journal Speech Language Law*, vol. 12, no. 1, pp. 63–76, 2005.
- [3] A. Cooper, “The electric network frequency (ENF) as an aid to authenticating forensic digital audio recordings—an automated approach”, *Audio Engineering Society Conference*, pp. 1–10, 2008.
- [4] E.B. Brixen “Techniques for the authentication of digital audio recordings”, *Audio Engineering Society Conference*, vol. 122, pp. 1–8, 2007.
- [5] J. Chai, F. Liu, Z. Yuan, R. Conners, and Y. Liu “Source of ENF in battery-powered digital recordings”, *Audio Engineering Society Conference*, vol. 135, pp. 1–7, 2013.
- [6] N. Fechner and M. Kirchner “The humming hum: Background noise as a carrier of ENF artifacts in mobile device audio recordings”, *8th International Conference IT Secure Incident Management IT Forensics*, pp. 3–13, 2014.
- [7] S. Vatansever and A.E. Dirik “Forensic analysis of digital audio recordings based on acoustic mains hum”, *24th Signal Processing Communication Application Conference*, pp. 1285–1288, 2016.
- [8] R. Garg, A.L. Varna, and M. Wu “Seeing ENF: Natural time stamp for digital video via optical sensing and signal processing”, *19th ACM International Conference Multimedia*, pp. 23–32, 2011.
- [9] R. Garg, A.L. Varna, A. Hajj-Ahmad, and M. Wu “Seeing ENF: Powersignature-based timestamp for digital multimedia via optical sensing and signal processing”, *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 9, pp. 1417–1432, 2013.
- [10] H. Su, A. Hajj-Ahmad, R. Garg, and M. Wu “Exploiting rolling shutter for ENF signal extraction from video”, *IEEE International Conference Image Processing*, pp. 5367–5371, 2014.
- [11] M. Wu, A. Hajj-Ahmad, and H. Su “Techniques to extract ENF signals from video image sequences exploiting the rolling shutter mechanism; and a new video synchronization approach by matching the ENF signals extracted from soundtracks and image sequences”, *U.S. Patent US9916857B2*, 2015.
- [12] S. Vatansever and A.E. Dirik “Videoların ENF tabanlı adli kanıt analizine ışık kaynağı etkisi”, *Journal of Uludag University Faculty of Engineering*, vol. 22, pp. 53–64, 2017.
- [13] S. Vatansever, A.E. Dirik, and N. Memon “Detecting the presence of enf signal in digital videos: A superpixel-based approach”, *IEEE Signal Processing Letters*, vol. 24, no. 10, pp. 1463–1467, 2017.
- [14] S. Vatansever, A.E. Dirik, and N. Memon “Analysis of rolling shutter effect on enf based video forensics”, *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 9, pp. 2262–2275, 2019.
- [15] S. Vatansever “Modern techniques in forensic analysis of multimedia signals”, Ph.D. Dissertation, Department of Electronics Engineering, Bursa Uludağ University, Bursa, 2019.
- [16] J. Choi and C.-W. Wong “ENF signal extraction for rolling-shutter videos using periodic zero-padding”, *IEEE International Conference on Acoustics, Speech, and Signal Processing*, pp. 2667–2671, 2019.
- [17] S. Fernández-Mendiña and F. Pérez-González “Temporal localization of non-static digital videos using the electrical network frequency”, *IEEE Signal Processing Letters*, vol. 27, pp. 745–749, 2020.
- [18] P. Ferrara, G. Draper-Gil, I. Sanchez, H. Junklewitz, and L. Beslay “Modelling gop structure effects on ENF-based video forensics”, *International Conference on Digital Forensics and Cyber Crime*, pp. 121–138, 2020.
- [19] D. Nagothu, Y. Chen, A. Aved, and E. Blasch “Authenticating video feeds using electric network frequency estimation at the edge”, *EAI Endorsed Transactions on Security and Safety*, vol. 7, no. 24, pp. 168648, 2021.
- [20] D. Bykhovsky and A. Cohen “Electrical network frequency (ENF) maximum-likelihood estimation via a multitone harmonic model”, *IEEE Transactions on Information Forensics and Security*, vol. 8, no. 5, pp. 744–753, 2013.
- [21] S. Vatansever, A.E. Dirik, and N. Memon “ENF based robust media time-stamping”, *IEEE Signal Processing Letters*, vol. 29, pp. 1963–1967, 2022.
- [22] G. Hua and H. Zhang “ENF signal enhancement in audio recordings”, *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 1868–1878, 2020.
- [23] G. Hua, H. Liao, H. Zhang, D. Ye, and J. Ma “Robust ENF estimation based on harmonic enhancement and maximum weight clique”, *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 3874–3887, 2021.
- [24] C. Grigoras, A. Cooper, and M. Michalek “Best practice guidelines for ENF analysis in forensic authentication of digital evidence”, *Forensic Speech Audio Analysis Working Groups*, vol. 1, no. 1, pp. 1–10, 2009.
- [25] A. Hajj-Ahmad, R. Garg, and M. Wu “Instantaneous frequency estimation and localization for ENF signals”, *Asia-Pacific Signal and Information Processing Association Annual Summit and Conference*, pp. 1–10, 2012.
- [26] J.G. Proakis and D.G. Manolakis “Digital Signal Processing: Principles, Algorithms, and Applications”, 4th Edition, Pearson, 2007.
- [27] A. Hajj-Ahmad, R. Garg and M. Wu “Spectrum Combining for ENF Signal Estimation”, *IEEE Signal Processing Letters*, vol. 20, no. 9, pp. 885–888, 2013.