

# Belirlenimci Benzerlik Yöntemlerine Dayalı Model Tabanlı Ses Dizisi Hizalaması

## Model Based Audio Sequence Alignment Based on Deterministic Similarity Methods

Doğaç Başaran\*, Ali Taylan Cemgil†, Emin Anarım\*

Elektrik ve Elektronik Mühendisliği Bölümü\*

Bilgisayar Mühendisliği Bölümü†

Boğaziçi Üniversitesi

İstanbul Türkiye

Email:dogac.basaran,taylan.cemgil,anarim@boun.edu.tr

**Özetçe** —Bu çalışmada, çoklu ses dizilerinin hizalanması problemi için ilinti ve Hamming uzaklığını temel alan iki yeni üretim modeli önerilmektedir. Burada, ikişerli hizalama durumları ele alınmış olsa da, yöntem kolayca çoklu ses dizilerinin de hizalanması için genellenabilir. Gerçek veriler üzerinde yapılan benzetim çalışmaları, iki dizinin örtüşüğü ve örtüşmediği, çok gürültülü durumlarda doğru parametre seçimiyle yöntemimizin başarılı olduğunu göstermektedir.

**Anahtar Kelimeler**—*Ses hizalama, ses eşleştirme, ses parmak izi, Bayesçi modelleme*

**Abstract**—In this work, we proposed two new generative models for multiple audio sequence alignment based on correlation and Hamming distance. Here, we focused on pairwise alignment however the framework extends directly to multiple alignment cases. The simulation results on real data sets suggest that our method is very robust and efficient for overlapping or non-overlapping cases under very noisy conditions with proper choices of model parameters.

**Keywords**—*Audio alignment, audio matching, audio fingerprinting, Bayesian modelling*

## I. GİRİŞ

Ses hizalama veya ses parmak izi problemi literatürde, bir ses dizisinin büyük bir veri tabanına eşleştirildiği bir tanılama problemi olarak tanımlanır. Bu konuda yüksek eşleştirme başarımına sahip, gürültüye dayanıklı birçok ses parmak izi yöntemi bulunmaktadır [3],[4],[5],[6]. [1] ve [2]'de, ses hizalama sorunu değişik bir açıdan ele alınmıştır. Burada, birçok eşzamansız alıcı olduğu ve bunların birbirlerinden bağımsız şekilde kayıt aldıkları varsayılmıştır. Amaç bu ses kayıtlarını genel bir zaman çizgisinde birbirlerine göre hizalamaktır. Buradaki temel zorluklar şöyle sıralanabilir;

- Ses dizileri örtüşmeyebilirler.
- Hiçbir ses dizisi tüm zaman çizgisini kaplamak zorunda değildir.

- Genel ses hizalama fikrindeki gibi temiz bir veri tabanı bulunmamaktadır.

Bu açıdan bakıldığında hizalama, bağıl konumu bilinmeyen video kliplerini eşzamanlı hale getirme [7] veya gürültülü kayıtlardan arka plandaki bir ses dizisini yeniden oluşturmak gibi birçok başka probleme de uygulanabilir. Olası bir uygulama da, bir konserin kaydını seyircilerin birbirlerinden bağımsız olarak aldıkları kayıt parçalarından oluşturmaktır. Benzer yaklaşımlar genetik alanında DNA sarmallarını küçük dizilerden oluşturma probleminde [8] ve kısmi olarak örtüşen resimlerden panoramik bir resim oluşturma probleminde de karşımıza çıkmaktadır.

Hizalama sorunundaki iki önemli kriter hesaplama hızı ve gürbüzlüktür. Bu iki kriter açısından da hizalama genellikle ham veride değil, öznitelik uzayında uygulanır. Eğer öznitelik uzayı gürültüye ve çevresel etkenlere karşı bağışık olursa, hizalama başarımının daha yüksek olacağı açıktır. Kısa süreli pencerelerdeki sinyal enerjisi [3], yerel renk doygunluğu (Kroma) enerji dağılımı [4], pozitif spectrum farkı [5], [9] ve spektrumun frekans ve zaman boyutlarında birinci farkının kendisi ve eşiklenmiş hali [6] hizalama sorununda en çok kullanılan özniteliklerdir.

Model yaklaşımında kaynakların başlangıç noktalarına göre hizalarına rastgele değişkenler atanmakta ve daha sonra bu hiza değişkenlerinin eldeki veriye koşullu sonsal olasılık dağılımı hizalama için bir ceza fonksiyonu olarak kullanılmaktadır. Daha önce çoklu ses dizilerini hizalama sorunu için, şablon bir üretim modeli tanımlanmış ve döngülü şekilde çalışan, ikişerli biçimde hizalamanın kullanıldığı ardışık bir algoritma önerilmişti [1]. [2]'de ise yaklaşık Bayes çıkarımı yöntemlerinden Gibbs örnekleyicisi ile sonsal dağılımı örneklemeye fikri sunulmuş ve benzetimli tavlama yaklaşımının burada kullanımı sergilenmişti. Bu çalışmada, aynı problem kurgusunu ikişerli hizalama durumları için, belirlenimci benzerlik yöntemlerine dayalı yeni üretim modelleriyle ele alıyoruz. Bundaki amacımız, [1]'da önerilen ardışık algoritmadaki ikişerli hizalama kısmının başarımını arttırmak ve bir altprogram oluşturmaktır. Önerilen yöntem, çoklu ses dizilerini hizalamak için genellenabilir [1],[2].

Hızalama problemi için benzerlik ölçütlerini direk olarak kullanmak da mümkündür ancak ses dizilerinin örtüşmediği veya birtakım verinin eksik olduğu durumlarda bu yöntemlerin nasıl kullanılacağı açık değildir. Bu ölçütlerin ceza fonksiyonu olarak kuvvetinden yararlanıp, aynı zamanda örtüşmeyen durumların üstesinden gelmek için, [1]'de tanımlanan şablon üretim modeli izlenerek, bu ölçütleri temel alan yeni üretim modelleri oluşturulmuştur. Yeni modellerin sonsal dağılımları benzerlik ölçütlerine benzemekte, ve aynı zamanda örtüşmeme durumu için de bir puan belirleyebilmektedirler. Bu çalışmada ilinti ve Hamming uzaklığı ölçütleri üzerinde durulmuştur.

Raporun geri kalanı şu şekilde düzenlenmiştir: Bölüm 2'de belirlenimci benzerlik ölçütlerinden ilinti ve Hamming uzaklığı yöntemlerine değinilmiş ve Bölüm 3'te önerilen belirlenimci benzerlik ölçütlerine dayalı modeller ayrıntılı bir şekilde tanımlanmıştır. Sonuçlar Bölüm 4'te verilmektedir.

## II. BELİRLENİMCİ BENZERLİK ÖLÇÜTLERİYLE HIZALAMA

İlinti, en bilinen belirlenimci benzerlik ölçütlerinden biridir. Ham veri üzerinde uygulanabildiği gibi öznelik uzayında da uygulamak mümkündür. Sezgisel olarak, doğru hizalanan ses dizilerinin örtüşen kısımlarının birbirine benzemesi gerekir. Dolayısıyla her olası hizalama için benzerlik ölçütü hesaplandığında, doğru hizalamada ölçütün en yüksek değeri vermesi beklenir. İkişerli ses dizisi hizalama problemlerinde, amaç birbirlerine göre hizalamak olduğundan, kaynaklardan bir tanesini sabit tutup, diğerini onun üzerinden kaydırmak suretiyle tüm olası hizalamalar taranabilir [1].

İkili öznelik dizileri için, örtüşen kısımların bitsel olarak karşılaştırılması bir benzerlik ölçütü olarak kullanılabilir. Buna Hamming uzaklığı adı verilir. Bu durumun bir örneği Şekil 1'de gösterilmiştir. Örnekte  $\tau = 1 \dots T$  genel zaman çerçeve dizinini,  $r_k$  ise  $k$ 'inci kaynağın genel zamana göre başlangıç noktasını ifade eden değişkenlerdir.  $x_{k,n,f}$ ,  $k$ 'inci kaynağın  $f$ 'inci altbandındaki  $n$ 'inci katsayıyı ifade etmektedir. Bu örnekte 2 tane kaynak ( $x_1, x_2$ ) ve her kaynağın 2 altbandı bulunmaktadır. Her iki kaynak da 3 elemanlıdır. Kaynaklardan  $x_1$  genel zaman çerçeve dizininde 2. sıradan,  $x_2$  ise 3. sıradan başlamaktadır.

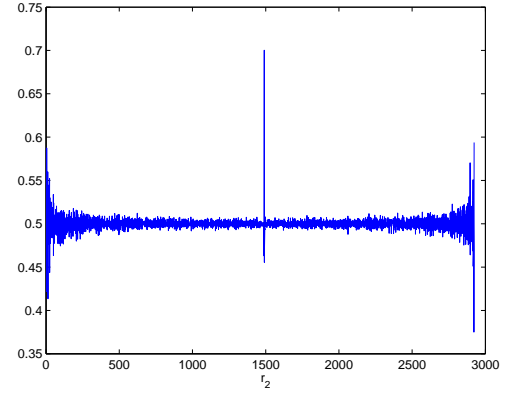
Bu benzerlik ölçütündeki ana fikir şudur, doğru hizalanmış olan kaynakların örtüşen ve genel zaman dizininde aynı noktaya denk gelen kısımlarında, aynı altbanddaki katsayıların çoğunlukla aynı olması beklenir. Yani Şekil 1'deki örnekte  $(x_{1,1,1}, x_{2,0,1}), (x_{1,1,2}, x_{2,0,2}), (x_{1,2,1}, x_{2,1,1})$  ve  $(x_{1,2,2}, x_{2,1,2})$  katsayıları birbirlerine benzemelidir. Örtüşen kısımdaki bit sayısı hizalamaya göre değişeceğinden standartlaştırma adına aynı olan bit sayısının toplam örtüşen bit sayısına oranı gürbüz bir benzerlik ölçütü olarak kullanılabilir. Bu örnekteki durumda örtüşen 4 tane bitten 3 tanesi birbirine eşittir, o yüzden oran 3/4 olarak hesaplanır.

Bu ölçüt 0 ile 1 arasında bir sayı vermektedir ve eğer değeri 1'e yakınsa bu benzerliğin kuvvetli olduğu anlamına gelir. Şekil 2'de bunun bir örneği verilmiştir. Bu örnekte, [6]'de anlatılan yöntemle özneliklerine ayrılmış iki ses dizisinin her olası hizalama için hesaplanmış benzerlik değerleri çizilmiştir. Görüldüğü gibi genel olarak 0.5 civarında olan değer doğru hizalamaya ulaşıldığında 0.7 civarında bir doruk oluşturmuştur. Bu ölçüt ile ilgili başka önemli gözlemse, ses kaynaklarının

Şekil 1. İki ses dizisinin ikili özneliklerinin benzerliği

$\tau =$	1	2	3	4	5
$r_1 = 2$		$x_{1,0,1} = 1$	$x_{1,1,1} = 1$	$x_{1,2,1} = 0$	
		$x_{1,0,2} = 0$	$x_{1,1,2} = 1$	$x_{1,2,2} = 1$	
$r_2 = 3$			$x_{2,0,1} = 1$	$x_{2,1,1} = 0$	$x_{2,2,1} = 1$
			$x_{2,0,2} = 1$	$x_{2,1,2} = 0$	$x_{2,2,2} = 1$

Şekil 2. İkili öznelikler kullanılarak hizalama



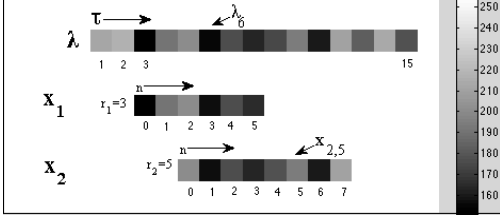
daha çok örtüştükları durumlarda daha gürbüz bir sonuç vermesidir. Örtüşen kısım daha büyük olduğunda sezgisel olarak aynı olan bitlerin oranının daha gerçekçi bir sonuç vereceğini düşünmek yanlış olmaz.

## III. BELİRLENİMCİ BENZERLİK ÖLÇÜTLERİNE DAYALI MODEL BAZLI HIZALAMA

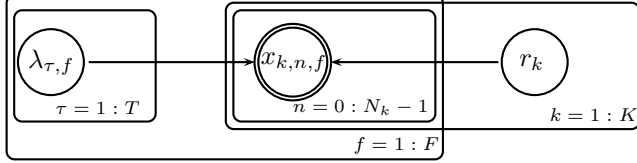
Bu bölümde, [1] ve [2]'de verilen modeli özetleyip, ilinti ve bit bazında kıyaslama yöntemlerinin yeni üretim modelleriyle bütünleştirilmesini açıklayacağız. Şekil 3'de, model basit bir örnekle gösterilmektedir. Burada öznelikler pozitif katsayılar olup, renkler katsayının değerine göre çizilmiştir. Modelin ana fikri şudur: *Doğru şekilde hizalanmış öznelik dizileri, alta yatan, ortak ancak gözlemlenemeyen bir başka öznelik dizisinin gürültülü gerçekleşmeleri veya fonksiyonlardır* [1], [2]. Saklı öznelik dizisi,  $\lambda_\tau$  ile ifade edilmektedir.  $\tau$ ,  $r_k$  ve  $x_{k,n,f}$ , Bölüm II'de anlatıldığı şekilde kullanılmıştır. Şekil 3'de gösterimin kolaylığı açısından kaynaklarda frekans dizini gösterilmemiştir.  $N_k$  ise  $k$ 'inci kaynağın uzunluğunu göstermektedir. Bu örnekte dizilerin uzunlukları  $N_1 = 6, N_2 = 8$ , başlangıç noktaları ise  $r_1 = 3, r_2 = 5$  olarak alınmıştır. Diziler birçok noktada örtüşmektedir örneğin  $x_{1,2}$  ve  $x_{2,0}$ ,  $\tau = 5$  genel zamanında hizalanmışlardır. Görüldüğü gibi  $x_{1,2}$  ve  $x_{2,0}$  birbirlerine çok yakın değerlere sahiptirler zira her ikisi de aynı saklı kaynak  $\lambda_5$ 'in gürültülü gözlemlenmiş halleridir.

Bu fikrin uygulandığı şablon üretim modeli [1]'de şu şekilde verilmiştir;

Şekil 3. Model Açıklaması.  $\lambda$  gizli dizi,  $x_1$  ve  $x_2$  gözlemlenmiş diziler



Şekil 4. Grafik Modeli



$$\lambda_{\tau,f} \sim p(\lambda_{\tau,f})$$

$$r_k \sim p(r_k) = \prod_{\tau=1}^{T-N_k+1} \pi_{k,\tau}^{[r_k=\tau]}$$

$$x_{k,n,f} \sim p(x_{k,n,f}|r_k, \lambda_{1:T,1:F}) = \prod_{\tau=1}^T p(x_{k,n,f}|r_k, \lambda_{\tau,f})^{[n=\tau-r_k]}$$

Burada, saklı öznitelik dizisindeki katsayılar,  $\lambda_{\tau,f}$ , önsel bağımsız olarak kabul edilmiştir ve hizalama değişkeni  $r_k$ , bir-biçimli olarak varsayılmıştır.  $[\cdot]$  ifadesi içindeki ifade doğru olduğunda 1, yanlış olduğunda 0 vermektedir. Gözlem modelindeki  $[n = \tau - r_k]$  ifadesi, eğer  $x_{k,n,f}$  katsayısı genel zaman  $\tau$ 'ya hizalanmışsa, saklı dizinin yalnızca  $\lambda_{\tau,f}$  katsayısına bağımlı olduğunu göstermektedir. Dolayısıyla gözlenen katsayılar aynı hizada olmadıkları sürece farklı saklı katsayıya koşullu olacaklardır. Grafik modeli Şekil 4'de verilmiştir.

Burada amaç, gözlemlenen dizilerin en olası hizalanmasını ( $r_{1:2}^*$ ) bulmaktır. Bu, birleşik koşullu sonsal dağılım,  $p(r_{1:2}|x_{1:2,0:N_k-1})$ 'ın birincil modunu bulmaya denk gelir. Hizalama konusunda bir önsel bilginin olmadığını varsayarsak, olabilirlik, sonsal ve birleşik dağılımlar orantılı hale gelir. Dolayısıyla,  $\Phi(r_{1:2}) = p(x_{1:2,0:N_k-1}, r_{1:2})$  bir ceza fonksiyonu olarak kullanılabilir. Önsel ve olabilirlik dağılımları Gamma ve Ters Gamma, Gauss ve Gauss, Bernoulli ve Bernoulli gibi eşlenik dağılımlardan seçilirse,  $\Phi(r_{1:2})$  analitik olarak, birleşik sonsal dağılım ( $p(x_{1:2,0:N_k-1}, r_{1:2}, \lambda_{1:T,1:F})$ ) saklı katsayılar  $\lambda_{\tau,f}$  üzerinden toplanarak türetilebilir [1],[2]. O halde en iyi hizalama  $\Phi(r_{1:2})$ 'nın logaritmasını en çok büyüten hizalama olarak bulunur yani,

$$r_{1:2}^* = \arg \max_{r_{1:2}} \mathcal{L}(r_{1:2}) = \arg \max_{r_{1:2}} \log \Phi(r_{1:2})$$

Bu aynı zamanda Bayes model seçme sorunu olarak da düşünülebilir [10]. Bu şekilde yaklaşıldığında en iyi hizalamayı bulmak, gözlemlenen veriyi en iyi şekilde açıklamak için değişik  $r_{1:2}$  yapılandırmalarını kıyaslamaya denk gelir.

Bölüm II'de bahsedildiği gibi, ikişerli hizalama probleminde amaç dizileri birbirlerine göre hizalamak olduğundan bir kaynağı sabitleyip diğeri üzerinden ceza fonksiyonu hesaplanabilir. Teknik olarak kaynakların örtüşmedikleri tüm durumlar

ceza fonksiyonunda aynı puanı vereceğinden [1],  $r_1 = N_2 + 1$  ve  $T = N_1 + 2N_2 - 1$  seçilerek ceza fonksiyonu sadece ikinci kaynağın hizalama değişkeni üzerinden hesaplanabilir. Sadece  $r_2 = 0$  değeri, kaynakların örtüşmediği durumu ifade etmekte, kalan tüm  $r_2$  değerleri diğer tüm olası hizalamaları içermektedir. O halde ceza fonksiyonu yalnızca ikinci kaynağın hizalama değişkenine bağlıdır ( $\mathcal{L}(r_2)$ ).

#### A. İLİNTİYE DAYALI ÜRETİM MODELİ (İM)

İlintiye dayalı bir üretim modeli oluşturmak kasıt,  $\mathcal{L}(r_2)$  ceza fonksiyonundaki terimlerden birinin ilinti veya ilinti benzeri olmasıdır. Bu önsel ve olabilirlik dağılımlarının seçimleriyle mümkündür. İlintiye dayalı bir model için önsel ve olabilirlik dağılımlarının Gauss seçilmesi gerekir. Saklı öznitelik dizi katsayılarının gözlemlenmiş kaynak dağılımlarının saklı ortalamaları oldukları varsayılmıştır. Uygun olması açısından saklı katsayı değişkeni için  $\lambda_{\tau,f}$  yerine  $\mu_{\tau,f}$  kullanılmıştır. Şu halde önsel ve olabilirlik dağılımları şunlardır,

$$\mu_{\tau,f} \sim \mathcal{N}(\mu_{\tau,f}; 0, \nu)$$

$$x_{k,n,f}|r_k, \mu_{\tau,f} \sim \prod_{\tau=1}^T \mathcal{N}(x_{k,n,f}; \mu_{\tau,f}, \alpha)^{[n=\tau-r_k]}$$

$\mathcal{N}$  Gauss dağılımını ifade etmektedir.  $\nu$  ve  $\alpha$  sırasıyla önsel ve olabilirlik dağılımlarının varyans değerleridir.  $\mathcal{L}(r_2)$  fonksiyonu şu şekilde hesaplanmıştır,

$$\mathcal{L}(r_2) = A(x_1, x_2, \nu, \alpha) + \frac{1}{\alpha \frac{\alpha}{\nu} + 2\alpha} \sum_{f=1}^F \sum_{\tau=1}^T x_{1,\tau-N_2-1,f} x_{2,\tau-r_2,f}$$

Denklemdaki ikinci terim ilinti terimini, birinci terim  $A(x_1, x_2, \nu, \alpha)$  da kalan tüm terimleri ifade etmektedir.

Bu model reel katsayılarından oluşan öznitelik dizileri için uygundur.  $\nu$  ve  $\alpha$  değerlerini uygun şekilde seçerek elde edilen ceza fonksiyonundaki önemi değiştirilebilir. Örneğin,  $\alpha$  değerini yüksek seçmek verinin etkisini azaltacaktır ki bu anlamlıdır zira eğer gözlemlerin varyansını yüksekse veri çok güvenilir değildir ve hizalamaya etkisi az olmalıdır.  $\nu$  değerinin etkisi tek başına  $\alpha$  kadar açık değil ancak  $\frac{\alpha}{\nu}$  oranına bakılacak olursa,  $\nu$  değerini  $\alpha$ 'ya göre düşük seçmek yine verinin etkisini azaltan yönde bir etki yapar.

#### B. HAMMING UZAKLIĞINA DAYALI ÜRETİM MODELİ (HM)

Aynı Bölüm III-A'de açıklandığı gibi önsel ve olabilirlik dağılımlarını uygun seçerek, ceza fonksiyonunda Hamming uzaklığını içeren bir terim elde edilmiştir. Bu kez önsel ve olabilirlik dağılımları Bernoulli ve koşullu Bernoulli dağılımları olarak aşağıdaki gibi seçilmiştir,

$$\lambda_{\tau,f} \sim \mathcal{BE}(\lambda_{\tau,f}; \alpha, \lambda)$$

$$x_{k,n,f}|r_k, \lambda_{\tau,f} \sim \prod_{\tau=1}^T \mathcal{P}(x_{k,n,f}|r_k, \lambda_{1:T,f})^{[n=\tau-r_k]}$$

$\mathcal{BE}$  Bernoulli dağılımını,  $\mathcal{P}(x_{k,n,f}|r_k, \lambda_{1:T,f})$  da koşullu Bernoulli dağılımını ifade etmektedir ve şu şekilde tanımlıdır,

$$\mathcal{P}(x_{k,n,f}|r_k, \lambda_{1:T,f}) = (w_{i,j})^{\sum_{i=0}^1 \sum_{j=0}^1 [x_{k,n,f}=i][\lambda_{\tau,f}=j]}$$

Burada  $w_{i,j}$ ,  $\lambda_{\tau,f} = j$  iken  $x_{k,n,f} = i$  olma olasılığını vermektedir. Bu çalışmada eğer  $i \neq j$  ise  $w_{i,j} = w$ , eğer  $i = j$  ise  $w_{i,j} = 1 - w$  olduğu varsayılmıştır. Aynı zamanda

önsel dağılımın parametresi  $\alpha_\lambda = 0.5$  olarak seçilmiş, böylece saklı dizinin katsayılarında eşit miktarda 1 ve 0 bulunduğu varsayılmıştır.  $\mathcal{L}(r_2)$  fonksiyonu şu şekilde hesaplanmıştır,

$$\mathcal{L}(r_2) = \sum_{\tau=1}^T \sum_{f=1}^F \log \left( 0.5(1-w)^{([x_{1,\tau-N_2-1},f=0]+[x_{2,\tau-r_2},f=0])} \right. \\ \left. w^{([x_{1,\tau-N_2-1},f=1]+[x_{2,\tau-r_2},f=1])} \right. \\ \left. + 0.5w^{([x_{1,\tau-N_2-1},f=0]+[x_{2,\tau-r_2},f=0])} \right. \\ \left. (1-w)^{([x_{1,\tau-N_2-1},f=1]+[x_{2,\tau-r_2},f=1])} \right)$$

Bu fonksiyona göre, eğer örtüşen aynı alt banttaki 2 gözlem katsayısı birbirine eşitse toplama katkısı  $0.5(1-w)^2 + 0.5w^2$  şeklinde, eğer eşit değilse  $(1-w)w$  şeklinde olur. Dolayısıyla Bölüm II'de açıklandığı gibi aynı olan ve farklı olan bitlerin sayısı hesaplanarak ceza fonksiyonu hesaplanabilir.

#### IV. SONUÇLAR

Deneyler için kullanılan kaynaklar, değişik gürlüğü tip ve seviyelerinde, uzunlukları birbirlerinden farklı, değişik ortamlarda kaydedilmiş örneklerden oluşmaktadır. Doğru hizalamalar her deneyde bilinmektedir. 20 tane örtüşen, 20 tane örtüşmeyen kaynak üzerinden yürütülen deneylerin sonuçları sırasıyla Tablo I ve Tablo II'de listelenmiştir. Karşılaştırma amacıyla örtüşen kaynaklar için ilinti ve Hamming uzaklığı yöntemleri de ceza fonksiyonu olarak kullanılmış, İM ve HM sonuçlarıyla birlikte Tablo I'de gösterilmiştir. Sonuçlar doğru hizalama, yanlış hizalama ve örtüşmeme olarak 3 sınıfta incelenmiştir. Tablo II'de ise örtüşmeyen kaynaklar için İM ve HM sonuçları örtüşmeme ve yanlış hiza olarak gösterilmiştir.

İlintiye dayalı model reel öznitelikler için, Hamming uzaklığına dayalı model ise ikili öznitelikler için uygundur. Modelleme yönteminin esnekliği nedeniyle bu kriterleri sağlayan herhangi öznitelik çıkarımı uygulanabilir. Bu çalışmada, İkili öznitelik çıkarma işlemi için [6]'de anlatılan yöntem izlenmiştir. Bu yöntem kısaca spektrumun bantlara ayrılıp hem frekans hem zaman yönünde birincil farkın alınması ve ardından eşikleme vasıtasıyla bitlerin elde edilmesidir. Bu işlemin sonucunda oluşan 1'ler ve 0'ların sayısı aşağı yukarı eşit çıkmaktadır, dolayısıyla önsel dağılımın parametresini  $\alpha_\lambda = 0.5$  seçmek yanlış olmaz. İM içinse yine [6]'deki yöntem, yalnızca son eşikleme kısmı çıkarılarak uygulanmış, dolayısıyla reel öznitelikler elde edilmiştir.

Modelleme yaklaşımının başarımını etkileyen en önemli etmenlerden biri parametre seçimidir. Yapılan deneyler sonucunda bir parametre kümesinin bütün kaynaklar için doğru çalışması mümkün olmamıştır. Fakat kaynaklar gürlüğü miktarına göre iki sınıfa ayrılıp her iki sınıf için ayrı parametre setleri seçildiğinde hem örtüşen hem örtüşmeyen örneklerde başarı oranı çok yüksek çıkmıştır. İM'de  $\nu$  ve  $\alpha$  varyans parametreleri,  $\frac{\alpha}{\nu}$  oranı yüksek olacak şekilde seçildiğinde başarımlar daha yüksek olmuştur. Bu da esasında bu modelin elde edilen veriyi çok iyi açıklayamadığı anlamına gelir. Öte yandan HM'de,  $w$  değeri 1'e yakın olduğunda, aynı olan bitlerin ceza fonksiyonuna olan katkısı farklı olan bitlerin katkısından çok daha yüksek olur. Tam tersi durumda  $w$ , 0.5'den biraz büyük olduğundaysa aynı olan bitlerin katkısı yine daha büyük olduğu halde, iki değer birbirine çok daha yakındır. Buradan az gürlüklü olan durumlarda  $w$  değerinin yüksek seçilmesi, yüksek gürlüklü durumlarda ise 0.5'e daha yakın seçilmesi

Tablo I. ÖRTÜŞEN KAYNAKLAR İÇİN HIZALAMA BAŞARIMI

	D. Hiza	Y. Hiza	Örtüşmeme
İlinti	17	3	—
Bit bazlı kıyaslama	17	3	—
İM	17	0	3
HM	19	0	1

Tablo II. ÖRTÜŞMEYEN KAYNAKLAR İÇİN HIZALAMA BAŞARIMI

	Örtüşmeme	Y. Hiza
İM	16	4
HM	18	2

gerektiği çıkabilir zira  $w$  küçükken bitlerin aynı olması ile farklı olması aşağı yukarı aynı etkiyi yapmaktadır, bu da gürlüklü veriye daha az önem göstermek anlamına gelir.

Sonuçlarda görüleceği gibi modelleme yöntemi, hizalama sorunu için oldukça yüksek bir başarıma sahiptir. Eldeki verilerden özellikle ciddi miktarda rüzgar gürlüğü bulunan birkaç örnekte ulaşılan hizalama sonuçları, kullanılan özniteliklerin ve yöntemin gürlüğü açısından dikkat çekicidir. Örtüşmeyen örneklerden birinde, bir oyuncunun repliği farklı prova zamanlarında kaydedilmiştir. Ritmik yapısı ve tonlamasıyla birbirlerine çok benzemeleri sebebiyle yanlış hizalama olarak sonuç elde edilmiştir. Fakat bu örnek için de doğru parametre seçimi ile örtüşmeme sonucu elde edilebilmektedir. [1] ve [2]'deki modellerle kıyaslandıklarında, özellikle bahsi geçen zor senaryolarda HM'nin hem daha başarılı hem de daha hızlı olduğu gözlemlenmiştir.

#### V. TEŞEKKÜR

Bu çalışma 2007K120610 numaralı TAM Projesi kapsamında Türkiye Devlet Planlama Teşkilatı tarafından desteklenmiştir. Ali Taylan Cemgil, bu çalışmada, TÜBİTAK tarafından 110E292 Bayesci Tensör ayrıştırma (BAYTEN) projesi kapsamında desteklenmektedir.

#### KAYNAKÇA

- [1] D. Basaran, A. T. Cemgil and E. Anarım, "Model Based Multiple Audio Sequence Alignment," in Proc. IEEE Workshop on Applications of Signal Processing to Audio and Acoustics WASPAA '11, pp. 13-16, 2011.
- [2] D. Basaran, A. T. Cemgil and E. Anarım, "Model Tabanlı Ses Dizisi Hizalaması", in Proceedings of the 19'th IEEE Signal Processing and Communications Applications Conference (SIU'11), 2011
- [3] Wang, A.L, "An Industrial-Strength Audio Search Algorithm", InProc. ISMIR, Baltimore, USA, 2003.
- [4] M. Muller, F. Kurth and M. Clausen, "Audio Matching via Chroma-based statistical features", In Proc. Int. Conf. on Music Info. Retr. ISMIR-05, pages 288-295, London, 2005.
- [5] S. Dixon and G. Widmer, "Match: A music alignment tool chest", in Proc. ISMIR, London, GB, 2005
- [6] Haitsma J., Kalker T., "A Highly Robust Audio Fingerprinting System". in Proc. ISMIR Paris, France, 2002
- [7] Junsong Yuan; Qi Tian; Ranganath, S.; , "Fast and robust search method for short video clips from large video collection," Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on , vol.3, no., pp. 866- 869 Vol.3, 23-26 Aug. 2004
- [8] Weber J. L., Myers E. W., "Human Whole-Genome Shotgun Sequencing", Genome Res. 1997 7: 401-409
- [9] J.P. Bello, L. Daudet, S. Abdallah, C. Duxbury, M. Davies, and M. Sandler, "A tutorial on onset detection in musical signals", IEEE Transactions on Speech and Audio Processing, vol. 13, no. 5, pp. 1035-1047, 2005.
- [10] C. M. Bishop, *Pattern Recognition and Machine Learning*, Springer, 2006