

KONUŞMA İŞARETLERİNİN KONUŞMACIDAN BAĞIMSIZ GENELLEŞTİRİLMİŞ ÖNCEDEN TANIMLI TEMEL SES VE ZARF VEKTÖR BANKALARI İLE OPTİMUM MODELLENMESİ

Ümit GÜZ

Hakan GÜRKAN

Sıddık B. YARMAN

İşık Üniversitesi, Mühendislik Fakültesi, Elektronik Mühendisliği Bölümü
80670, Maslak, İstanbul

e-posta: guz@isikun.edu.tr hakan@isikun.edu.tr yarman@isikun.edu.tr

Anahtar sözcükler: Konuşma İşaretlerinin Modellenmesi, Veri Kodlama, Veri Sıkıştırma

ABSTRACT

*In this paper a new modeling and representation method of speech signals is introduced. The crucial feature of the proposed method is its independence from speaker. The reconstruction procedure is based on the generation of the so called Generalized Predefined Signature and Envelope Banks (GPSEB). In this method, due to its non-stationary behaviour, speech signals blocked into frames and each frame is assumed quasi-stationary. Any frame of speech signal $X_i(t)$ is modeled by the form of $X_i(t)=C_i w_i(t) v_i(t)$. Each frame is represented in terms of index numbers which corresponds to best representative Generalized Signature and Envelope Vectors and appropriate frame scaling coefficient in the GPSEB. It has been shown that the proposed method provides fast reconstruction algorithm and substantial compression ratio according to hearing quality.**

1. GİRİŞ

Konuşma işaretlerinin işlenmesi, iletilmesi ve depolanması modern iletişim sistemlerinde oldukça geniş uygulama alanlarına sahiptir. Özellikle, konuşma işaretlerinin modellenmesi ya da yeniden oluşturulması için gerekli olan bilgi miktarının azaltılması, sayısal konuşma işaretlerinin iletilmesini ve depolanmasını sağlayan sistemlerin kapasitesini belirgin bir biçimde arttırmaktadır. Konuşma işaretlerinin iletimi ya da yeniden oluşturulmaları sırasında konuşma işaretinin kendisinden ve konuşmacıdan bağımsız, bilginin önemli özelliklerini kaybetmeden ve yüksek sıkıştırma oranları ile oluşturulması oldukça önem taşımaktadır. Son yıllarda işaret işleme alanında, işaretlerin gösterimi, öznel bulma, sınıflandırma, boyut indirgeme, kodlama, sıkıştırma ve yeniden oluşturma tekniklerine ilişkin yeni yöntemler geliştirilmiştir. İşaretlerin modellenmelerine yönelik daha önce geliştirilen

yöntemler değerlendirildiğinde bunların iki sınıfta toplanmış oldukları görülür.

- Zaman Alanı Yöntemleri (Darbe Kod Modülasyonu, Entropi Kodlaması, Öngörüye Dayalı Yöntemler, Vektör Nicemleme).
- Dönüşüm Alanı Yöntemleri (Temel Bileşenler Dönüşümü-TBD, Karhunen Loeve Dönüşümü-KLD, Ayrık Fourier Dönüşümü, Ayrık Kosinüs Dönüşümü).

Bu yöntemlerden alt-uzay ya da dönüşüm tabanlı boyut indirgeme ve sıkıştırma yöntemleri son yıllarda geniş uygulama alanı bulan yöntemlerdendir.

Daha önce sunulan çalışmalarımızda [1-6] Konuşma ve EKG işaretlerinin sırası ile, sinusoidal yapıda seçilen Tanım Fonksiyonları (TF) ve enerjinin büyük bir kısmının içerildiği Temel Tanım Frekanslarının belirlenmesi ve TF'lerin dikleştirilerek Temel Tanım Fonksiyonlarının elde edilmesinden sonra yeniden oluşturulmaları gerçekleştirilmiştir. Daha sonra yapılan çalışmalarda [7-9], herhangi bir kaynağa ilişkin ses, konuşma ve EKG işaretleri çift tabanlı bir referans çizelgesi ile modellenmişlerdir. Referans çizelgesi, KLD sonucu elde edilen özvektörlerin benzer olanlarının indirgenmesi ile oluşturulan ve ses vektörleri olarak adlandırılan vektörler ile özgün işaretin kullanılması ile elde edilen iyileştirme vektörlerinden oluşturulmuştur. Özgün işaret, referans çizelgede herbir çerçevenin temsil edildiği ses ve iyileştirme vektörlerinin indislerinin karşı tarafta da bulunduğu varsayılan aynı referans çizelgesine gönderilmesi ile yeniden oluşturulmaktadır. Algoritmanın hızlandırılması, yeniden oluşturma hatasının ve kullanılan vektör sayılarının indirgenmesi amacıyla geliştirilen yeni yaklaşımlarımızda [10-12] ses, konuşma, EKG ve EEG işaretleri, Önceden Tanımlı Fonksiyon Bankaları (ÖTFB) olarak adlandırılan bankalar (Temel Tanım ve Zarf) yardımı ile modellenmektedir. Geliştirilen yöntemin en belirgin özelliği, herhangi bir konuşma işaretinin, konuşmacıdan ve söz konusu konuşma işaretinden bağımsız olarak aynı dildeki veya başka bir dildeki konuşma işaretinden üretilebilmesidir. Konuşma işaretleri durağan yapıda olmadıklarından yarı durağan kabul edilebilecekleri uygun ve eşit uzunluktaki çerçevelere ayrılarak modellenmektedirler. Herbir

* Bu çalışmanın bir bölümü, Prof. Dr. Sıddık B. YARMAN yönetiminde, Ümit GÜZ'ün doktora çalışması olarak tamamlanmıştır [11].

çerçeve $X_i(t)$ olmak üzere, ses ya da konuşma işaretleri $X_i(t) \approx C_i w_i(t) v_i(t)$ biçiminde modellenmektedir. Getirilen Zarf Bankası kavramı ile yeniden oluşturma işlemi en az hata ile gerçekleştirilmiş ve yeniden oluşturma sırasında yüksek sıkıştırma oranları elde edilmiştir.

Bu çalışmada, daha önce yapılan yaklaşımlarda [10-12] Temel Tanım ve Zarf vektörlerinin indirgeme işlemleri ile elde edilmesinde kullanılan uzaklık ölçütü yerine ilinti katsayısı kullanılmaktadır. Bu durumda indirgeme miktarı ve çok daha az indis kullanılması nedeni ile yeniden oluşturma hızı ve sıkıştırma oranı oldukça artmaktadır.

2. YENİ MODELLEME YÖNTEMİ

$X(n)$ ayrık ortamda kaydedilmiş gerçel, ortalama değeri sıfır olan, belirli bir aralıkta ve uzunluğu N olan bir konuşma işaretini temsil etmektedir. Bu işaret çerçeve çerçeve analiz edildiğinde, $X_i(n)$ herhangi bir i . çerçevesini temsil etmek üzere yöntemin temelini oluşturan aşağıdaki temel yaklaşım yapılabilir ve buna ilişkin alt tanımlamalar verilebilir.

Temel Tanım

(a) Herhangi bir konuşma işaretine ilişkin X_i ile temsil edilen herhangi bir i . çerçeve,

$$X_i \cong C_i W_i V_i \quad (1a)$$

yaklaşıklık bağıntısı ile gösterilebilir. Bu bağıntıda C_i gerçel bir sabit, $V_i^T = [v_1, v_2, v_3, \dots, v_{L_F}]$ ise bir satır vektördür. $C_i V_i$ vektörü, En Küçük Kareler (EKK) anlamında X_i çerçevesinin hemen hemen en yüksek enerjisini taşır. Başka bir deyişle $C_i V_i$ vektörü X_i çerçevesine hatanın karelerinin toplamını en aza indirecek biçimde tek bir terim ile en iyi yaklaşan vektördür $X_i \approx C_i V_i$.

$L_F \times L_F$ köşegen matrisi (1a) denklemindeki EKK hatasını en aza indirecek biçimde geliştirmeyi sağlayan bir zarf terimi olarak davranmaktadır.

$$W_i = \text{diag}[w_{i1} \quad w_{i2} \quad w_{i3} \quad \dots \quad w_{iL_F}]$$

bağıntısında L_F herhangi bir çerçevedeki toplam eleman sayısını göstermektedir.

(b) Sürekli zaman bölgesinde, (1a) denklemini aşağıdaki biçimi alır.

$$X_i(t) = C_i w_i(t) v_i(t) \quad (1b)$$

Temel tanıma bağlı olarak aşağıdaki tanımlamalar yapılabilir.

Tanım 1a : V_i vektörü, bir C_i katsayısı ile birlikte X_i vektörünün hemen hemen en yüksek enerjisini taşıdığından Temel Ses Vektörü olarak adlandırılır.

Tanım 1b : (1a) tanımına benzer bir yaklaşımla, $v_i(t)$ sürekli zaman bölgesi fonksiyonu Temel Ses Fonksiyonu olarak tanımlanır.

Tanım 2a : Köşegen W_i matrisi, $C_i V_i$ vektörünün zarfını özgün işaretin X_i çerçevesine uydurduğundan dolayı Zarf Matrisi olarak adlandırılır.

Tanım 2b : (2a) tanımına benzer bir yaklaşımla, $w_i(t)$ sürekli zaman bölgesi fonksiyonu Zarf Fonksiyonu olarak adlandırılır.

Tanım 3 : C_i gerçel sabiti her bir çerçeve için elde edilen Ölçekleme Katsayısı olarak tanımlanmıştır. Temel tanımın doğruluğunu kanıtlamak için aşağıdaki işlemler gerçekleştirilmektedir.

Ayrık zamanlı bir konuşma işareti $x(n)$, aşağıdaki biçimde yazılabilir.

$$x(n) = \sum_{i=1}^N x_i \delta_i(n-i) \quad (2a)$$

Bu denklemde, $\delta_i(n)$ birim örneği, x_i , N uzunluğundaki $x(n)$ dizisinin genliğini temsil etmektedir. $x(n)$ aynı zamanda vektör/matris gösteriminde aşağıdaki biçimde ifade edilebilir.

$$X^T = [x(1) \quad x(2) \quad \dots \quad x(N)] = [x_1 \quad x_2 \quad \dots \quad x_N] \quad (2b)$$

X Ana Çerçeve Vektörü olarak adlandırılmakta olup 16, 32 gibi eşit uzunluktaki çerçevelere bölünür. M_F ile gösterilen Çerçeve Matrisi, çerçeve vektörleri cinsinden,

$$M_F = [X_1 \quad X_2 \quad \dots \quad X_{N_F}] \quad (3)$$

$$X_i^T = [x_{(i-1)L_F+1} \quad x_{(i-1)L_F+2} \quad \dots \quad x_{iL_F}], \quad i=1,2,\dots,N_F \quad (4)$$

biçiminde yazılabilir. Burada N_F , X vektöründeki çerçevelerin sayısını temsil etmektedir ($N_F = N/L_F$). Her bir X_i , birimlik vektörler tarafından biçimlendirilmiş bir vektör uzayına açılabilir $\{V_{ki}, k=1,2,3,\dots,L_F\}$.

$$X_i = \sum_{k=1}^{L_F} c_k V_{ki} \quad (5)$$

$$c_k = (X_i)^T V_{ki} \quad (6)$$

Buradaki V_{ki} , hata vektörünün beklenen değerini en aza indirecek biçimde belirlenmektedir.

$$\varepsilon = X_i - \sum_{k=1}^{L_F} c_k V_{ki} \quad (7)$$

Bu işlem " V_{ki} 'nin EKK anlamında belirlenmesi" olarak adlandırılmaktadır. V_{ki} vektörleri, X_i dizilerinin R_i özilinti matrislerinin özvektörleri cinsinden hesaplanmaktadır. R_{i^r} 'ler, yarı kesin artı, gerçel, simetrik, Toeplitz matrislerdir. Yukarıdaki EKK yaklaşımı aşağıdaki özdeğer problemi ile açıklanabilmektedir.

$$R_i V_{ki} = \lambda_{ki} V_{ki}, \quad k=1,2,\dots,L_F \quad (8)$$

Bu durumda, λ_{ki} ve V_{ki} sırası ile özdeğer ve özvektörleridir. R_i 'lerin özdeğerleri gerçel ve pozitifdir. Ayrıca, V_{ki} özvektörlerinin tümü birimdir. Her bir X_i için elde edilen R_i matrislerinin özdeğerleri, karşı geldiği özvektörleri ile birlikte büyükten küçüğe doğru sıraya dizildiğinde ($\lambda_{1i} \geq \lambda_{2i} \geq \lambda_{3i} \geq \dots \geq \lambda_{L_F i}$) herhangi bir i . çerçevenin toplam enerjisi,

$$X_i^T X_i = \sum_{k=1}^{L_F} x_{ki}^2 = \sum_{k=1}^{L_F} c_{ki}^2 = \sum_{k=1}^{L_F} \lambda_{ki} \quad (9)$$

biçiminde yazılabilmektedir.

Her bir X_i için özdeğerler sıralamasında, en büyük değerli özdeğere karşı gelen özvektör, enerjisi en yüksek özvektör olup, X_i 'yi en iyi biçimde temsil eder. En yüksek enerjili özvektörler işaretteki en büyük değişim yönünü gösterirler ve temel bileşenler olarak

adlandırılırlar. Bu durumda, (5) bağıntısı en yüksek enerjili ilk p tane temel bileşenin alınması ile yuvarlatılabilir.

$$X_i \cong \sum_{k=1}^p c_{ki} V_{ki} \quad (10)$$

bağıntısında $p=1$ olması durumunda V_{li} özvektörleri özel olarak Ses Vektörleri olarak adlandırılırlar. EKK anlamında en yüksek enerjili taşıyan ses vektörleri her bir çerçeveyi özgün işarete ilişkin çerçevelere en az hata ile yaklaşıtıracaktır.

$$X_i = c_{li} V_{li} \quad (11)$$

Bu durumda, L_F uzunluğundaki çerçevelerin hemen hemen tüm enerjisi (5) bağıntısındaki ilk terime aktarılacak, diğer terimler enerji anlamında çok az bir katkıya sahip olduklarından gözardı edileceklerdir. (11) nolu eşitlikteki V_{li} özvektörleri özgün işarete ilişkin yararlı bilginin hemen hemen tümünü içermektedirler. Her bir X_i için tanımlanan Köşegen Zarf Matrisi W_i 'nin katılması ile (11) numaralı eşitlikte elde edilen yaklaşıtırlmış bağıntı aşağıdaki biçime dönüştürülmüş olmaktadır.

$$X_i = C_i A_i V_{li} \quad (12)$$

W_i matrisinin köşegen elemanları (a_{ir}), V_{li} elemanları (v_{lir}) ve X_i elemanları (x_{ir}) cinsinden

$$(a_{ir} = x_{ir} / (C_i v_{lir}), \quad r = 1, 2, \dots, L_F) \quad (13)$$

biçiminde elde edilir.

Çok sayıda konuşma işareti üzerinde yapılan incelemelerde, her bir X_i için elde edilmiş zarf vektörlerinin $a_i(n) = (a_{ir})$ çerçeve indisine göre $n=1, 2, 3, \dots, L_F$ ve temel ses vektörlerinin $v_i(n) = (v_{lir})$ çerçeve indisine göre $n=1, 2, 3, \dots, L_F$ dalga biçimlerinin birçoğunun birbirlerini yineleyen benzerlikler gösterdiği saptanmıştır. Konuşma işaretlerinin her bir çerçevesi için elde edilmiş ses ve zarf vektörleri içerisinde, benzer dalga biçimlerine sahip olan vektörler bir ilinti katsayısı (ρ_{YZ}) bağıntısı yardımı ile indirgenmişlerdir.

$$Y = [y_1 \quad y_2 \quad \dots \quad y_L], \quad Z = [z_1 \quad z_2 \quad \dots \quad z_L] \quad \text{ve}$$

$$\rho_{YZ} = \frac{\sum_{i=1}^L (y_i \cdot z_i) - \left(\sum_{i=1}^L y_i \right) \left(\sum_{i=1}^L z_i \right) / L}{\sqrt{\left(\sum_{i=1}^L y_i^2 - \left(\sum_{i=1}^L y_i \right)^2 / L \right) \cdot \left(\sum_{i=1}^L z_i^2 - \left(\sum_{i=1}^L z_i \right)^2 / L \right)}} \quad (14)$$

İndirgenmiş ses ve zarf vektörleri sırası ile "Temel Ses" ve "Zarf" bankaları adı verilen bankalarda toplandırılmışlardır.

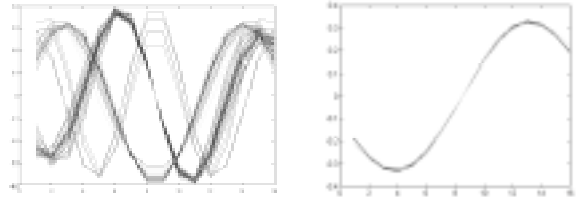
Sonuç olarak, herhangi bir i . çerçeve X_i , EKK anlamındaki yaklaşım hatasını en aza indiren bir temel ses vektörünün bir c_i katsayısı ile birlikte temel ses bankasından çekilmesi ile oluşturulmaktadır $X_i \cong c_i V_{li}$. İndirgenmiş temel ses vektörleri temel ses bankası adı altında toplandırılmıştır $\{V_{n_v}(n), n_v = 1, 2, \dots, N_V\}$. N_V sayısı "Temel Ses Bankası" içerisinde yeralan toplam "Temel Ses Vektörü" sayısını vermektedir. Benzer biçimde, indirgenmiş zarf vektörleri veya

köşegenleştirilmiş zarf matrisleri zarf bankası adı altında toplandırılmıştır $\{W_{n_w}(n), n_w = 1, 2, \dots, N_W\}$.

N_W sayısı "Zarf Bankası" içerisindeki toplam "Zarf Vektörü" sayısını vermektedir. Temel ses ve zarf vektörleri sürekli zaman bölgesi fonksiyonları olarak sırası ile $\{v_{n_v}(t), n_v = 1, 2, \dots, N_V\}$ ve $\{w_{n_w}(t), n_w = 1, 2, \dots, N_W\}$ olarak ifade edilebilirler.

Sonuç olarak, herhangi bir konuşma işaretine ilişkin bir X_i , temel ses $v_i(t)$ ve zarf $w_i(t)$ fonksiyonları ile, bir C_i katsayısının çarpımı biçiminde yeniden oluşturulmaktadır $X_i(t) \cong C_i w_i(t) v_i(t)$.

Yeni modelleme yöntemi iki temel algoritma tarafından gerçekleştirilmektedir.



Şekil 1. Benzer Özvektörler

Algoritma 1: Genelleştirilmiş Temel Ses ve Zarf Bankalarının Oluşturulması

Girişler: X, L_F

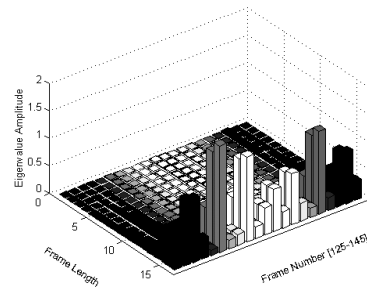
Adım 1: Toplam çerçeve sayısının hesaplanması $N_F = N / L_F$.

Adım 2: X 'in çerçeve vektörlerine ayrılması (X_i).

Adım 3: Her bir X_i için, özilinti matrisinin (R_i) hesaplanması.

Adım 4: Her bir R_i için, R_i 'nin özdeğerlerinin (λ_{ki}) ve karşılık gelen özvektörlerinin (V_{ki}) hesaplanması.

Adım 5a: λ_{ki} 'lerden en büyük özdeğere (λ_{ri}), $\lambda_{ri} = \max\{\lambda_{1i}, \lambda_{2i}, \lambda_{3i}, \dots, \lambda_{L_F i}\}$ karşı gelen özvektörün (V_{ri}) belirlenmesi. V_{ri} 'nin çerçeve indisinin V_{li} olarak değiştirilmesi.



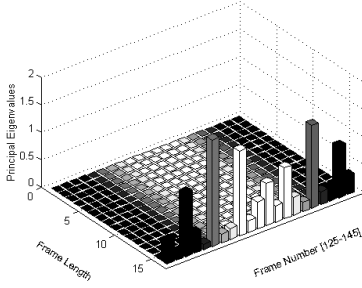
Şekil 2. [125-145] Çerçevelerinin Özdeğerleri

Adım 5b: C_i 'nin EKK anlamında $X_i = C_i V_{li}$ yaklaşımını elde edecek biçimde bulunması.

Adım 6: Adım 5'in tüm X_i 'ler için $i=1, 2, \dots, N_F$ gerçekleşmesi. Bu adımın sonunda, her bir X_i için en yüksek enerjili özvektörler elde edilmiş olmaktadır.

Adım 7: Adım 6'da elde edilmiş tüm özvektörlerin hızlı ve verimli bir algoritma ile karşılaştırılması ve benzer özvektörlerin elenmesi. İndirgenmiş sayıdaki

V_{li} özvektörlerinden Temel Ses Bankasının $\{V_{n_v}(n), n_v = 1, 2, \dots, N_V\}$ oluşturulması.



Şekil 3. [125-145] Çerçevelerinin Temel Özdeğerleri

Adım 8: Herbir X_i veya Adım 5b' de hesaplanan herbir $C_i V_{li}$ için Köşegenleştirilmiş Zarf Matrisinin köşegen elemanlarının $a_{ir} = x_{ir} / (C_i V_{li})$, $r=1, 2, \dots, L_F$ denklemi yardımı ile hesaplanması.

Adım 9: Adım 7 ile benzer biçimde, elde edilen tüm zarf vektörlerinin hızlı ve verimli bir algoritma ile karşılaştırılması ve benzer zarf vektörlerinin elenmesi. İndirgenmiş sayıdaki özgün vektörlerden Zarf Bankalarının $\{W_{n_w}(n), n_w = 1, 2, \dots, N_W\}$

oluşturulması. Burada N_W , Zarf Bankasındaki özgün zarf vektörlerinin toplam sayısını göstermektedir.

Algoritma2:Konuşma İşaretlerinin Genelleştirilmiş Önceden Tanımlı Temel Ses ve Zarf Bankaları Yardımı ile Yeniden Oluşturulması

Girişler: Modellenek konuşma işareti ($X(n), n=1, 2, \dots, N$), Algoritma 1'in kullanılması ile oluşturulmuş Önceden Tanımlı Bankalar, L_F : Alt çerçeve uzunluğu (Herbir çerçevede eşit sayıda bulunan örnek sayısı).

Adım 1: Modellenek konuşma işaretinin ($X(n)$) Algoritma 1 de olduğu gibi çerçeve vektörlerine ayrılması.

Adım 2: (a) Herbir çerçeve için, özgün ve yaklaşılmış çerçeve vektörleri arasındaki uzaklığı veya hatayı ($\delta = \|X_i - C_k V_k\|^2$, $k=1, 2, \dots, N_V$) en aza indiren uygun bir temel ses vektörünün V_k temel ses bankasından çekilmesi.

(b) V_K vektörüne karşı gelen K indis sayısının belirlenmesi ($X_i \approx C_K V_K$).

Adım 3: Uygun zarf vektörlerinin (köşegenleştirilmiş matris) W_R , $\delta_R = \min\{\|X_i - C_K W_r V_K\|^2, r=1, 2, \dots, N_W\}$ hata tanımını en aza indirecek biçimde önceden tanımlı zarf bankasından çekilmesi.

(b) W_R vektörüne karşı gelen R indis sayısının belirlenmesi. Bu adım sonlandırıldığında, özgün çerçeveyi en iyi biçimde temsil edecek temel ses vektörü V_K ve zarf vektörü W_R uygun bir seçim ile belirlenmiş olacaktır. Başka bir deyişle X_i çerçeve vektörü en iyi biçimde W_R ve V_K vektörlerinin kullanılması ile oluşturulabilecektir ($X_i \approx W_R V_K$).

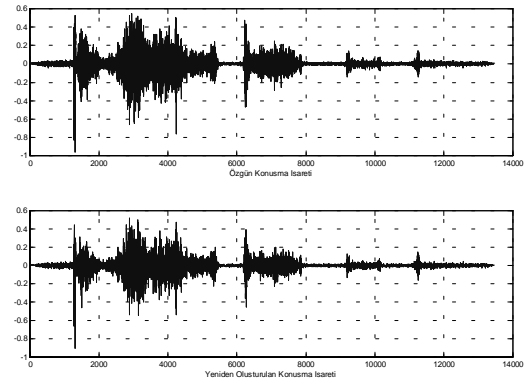
Adım 4: Elde edilen sabitlenmiş W_R ve V_K vektörlerini kullanarak, bir bütünsel en küçük (global minimum) hata değeri bulacak biçimde yeni çerçeve

katsayılarının (C_i) hesaplanması. Bu adımda, çerçeve vektörü yukarıda hesaplanan değerler yardımı ile, $X_i \approx C_i W_R V_K$ biçiminde elde edilir.

Adım 5: Yukarıdaki adımlar, yeniden oluşturulacak işaretin ($X(n)$) herbir çerçevesinin oluşturulması için yinelenir. Herbir X_i , kabul edilebilir bir hata ($\delta_{Global} = \|X_i - C_i W_R V_K\|^2$) sınırı içerisinde, C_i ve sırası ile ses ve zarf vektörü indisleri K ve R ile temsil edilmiş olur.

3. UYGULAMA VE SONUÇLAR

Algoritmaların başarımını test etmek için 8 bit, 8KHz'de örneklenmiş ses işaretleri (müzik, farklı dillerde konuşma işaretleri) kullanılmıştır. Aşağıda verilen örnek uygulamada, çerçeve sayısı ile karşılaştırıldığında Genelleştirilmiş Önceden Tanımlı Bankaların oldukça az sayıda vektör ile oluşturulabildiği görülmektedir. Herhangi bir ses işareti konuşma işaretinin kendisinden ve konuşmacıdan bağımsız olarak bu bankalar yardımı ile yeniden oluşturulabilmektedir. İndirgeme için karşılaştırma yapılırken ilti katsayısının 0.99 değerinden büyük olması durumu dikkate alınmıştır.



Şekil 4. (a) Özgün Konuşma İşareti, (b) Yeniden Oluşturulan Konuşma İşareti

Tablo 1. Yeniden Oluşturma Parametreleri

Gen. Banka için Konuşma İşareti)	İngilizce+ Fransızca+İskandinav
Kişi	Bayan
Gen. Banka Çerçeve Sayısı	2500
Temel Ses Vektörü Sayısı	171
Zarf Vektörü Sayısı	1946
Yeniden Oluşturulan İşaret	İngilizce
Kişi	Erkek
Özgün İşaret Çerçeve Sayısı	841
Çerçeve Uzunlukları	16
İndirgeme için ρ_{YZ}	0.99

4. SONUÇLAR

Konuşma işaretlerinin iletilmesinde, işaretin tamamının gönderilmesi yerine, herbir çerçeveyi temsil eden ve Genelleştirilmiş Önceden Tanımlı Temel Ses ve Zarf Vektörleri Bankasında yer alan Temel ses ve zarf vektörlerinin indislerin gönderilmesi konuşma işaretinin yeniden oluşturulması için yeterli

olmaktadır. Getirilen ilinti katsayısı kavramı ile indirgeme oranı artırılmış yeniden oluşturma işlemi daha yüksek sıkıştırma oranları ve daha az hata ile yapılmıştır. Temel Ses ve zarf vektörlerinin sayısı oldukça azalacağından bunları temsil etmek için gerekli olan bit sayısı da azalacaktır. Bu durumda, sıkıştırma oranı doğrudan bankanın boyutu ile orantılı olacaktır. Bankanın boyutları sıkıştırma oranı ve duyma kalitesini etkileyecektir. Yöntem, duyma kalitesine bağlı olarak önemli ölçüde sıkıştırma oranları vermektedir. Yöntemin en belirgin özelliği, yeniden oluşturma işleminin konuşmacıdan bağımsız olarak yapılabilmesidir. Başka bir deyiş ile, herhangi bir konuşma işareti yine herhangi bir konuşmacı için üretilmiş bankalardaki vektörlerden oluşturulabilmektedir. Yöntem, konuşmacıdan ve dilden bağımsız olmakla birlikte, konuşmacıya ya da bir dile ilişkin bankaların elde edilebileceği biçime dönüştürülebileceğinden konuşmacı tanılama veya bir dile ilişkin özelliklerin sınıflandırılması gibi amaçlar için de kullanılabilir.

KAYNAKLAR

- [1] Karaş M., Yarman B. S., "A New Approach for Representing Discrete Signal Waveforms via Private Signature Base Sequences", 12th European Conference on Circuit Theory and Design (ECCTD'95), Istanbul, Türkiye, pp. 875-878, August 27-31, 1995.
- [2] Akdeniz R., Yarman B. S., "Reconstruction of Speech Signals via Reduced One and Two Sound Syllables", 14th National Conference on Information Technology (BİLİŞİM'97), Istanbul, Türkiye, sayfa 118-122, 3-6 Eylül 1997.
- [3] Karaş A. M., Yarman B. S., "A New Method for the Compression of ECG Signals: The YARKAR Method", ICSPAT'97, San Francisco, USA, September 14-17, 1997.
- [4] Akdeniz R., Yarman B. S., "Speech Coding by Signature Base Sequences", 6. Sinyal İşleme ve Uygulamaları Kurultayı (SİU'98), Ankara, Türkiye, sayfa 178-183, 28-30 Mayıs 1998.
- [5] Akdeniz R., Yarman B. S., "Turkish Speech Coding by Signature Base Sequences", International Conference on Signal Processing Applications & Technology (ICSPAT), Toronto, Kanada, pp. 1291-1294, September 13-16, 1998.
- [6] Yarman B. S., Akdeniz R., "Generation of Optimum Signature Base Sequences for Speech Signals", First IEEE Balkan Conference on Signal Processing, Communication, Circuits and Systems, Istanbul, Turkey, Conference Digest – CD, pp. 1-4, June 1-3, 2000.
- [7] Güz Ü., Gürkan H., Yarman B. S., "Çift Tabanlı Referans Tablolarıyla Ses İşaretlerinin Modellenmesinde Özgün Bir Yaklaşım", SİU'2001 IEEE Sinyal İşleme ve Uygulamaları Kurultayı, Gazi Magusa, K.K.T.C., Cilt 2, sayfa 693-698, 25-27 Nisan 2001.
- [8] Gürkan H., Güz Ü., Yarman B. S., "EKG (ElectroCardioGram) İşaretlerinin Özgün Temel Tanım Fonksiyonları ile Modellenmesi", SİU'2001 IEEE Sinyal İşleme ve Uygulamaları Kurultayı, Gazi Magusa, K.K.T.C., Cilt 2, sayfa 587-592, 25-27 Nisan 2001.
- [9] Akdeniz R., Güz Ü., Gürkan H., Yarman B. S., "Farklı Ses Kaynaklarından Üretilen Temel Tanım Dizileri ile Konuşma İşaretlerinin Modellenmesi", SİU'2001 IEEE Sinyal İşleme ve Uygulamaları Kurultayı, Gazi Magusa, K.K.T.C., Cilt 1, sayfa 128-133, 25-27 Nisan 2001.
- [10] Güz Ü., Yarman B. S., Gürkan H., "A New Method to Represent Speech Signals via Predefined Functional Bases", Proceedings of ECCTD'01 European Conference on Circuit Theory and Design, Espoo, Finlandiya, Vol. II, pp. 5-8, August 28-31, 2001.
- [11] Güz Ü., "Türkçe Konuşma İçin Optimum Temel Tanım Fonksiyonlarının Belirlenmesinde Yeni Bir Yaklaşım", Doktora Tezi (Tez Danışmanı: Prof. Dr. B. Sıddık Yarman), İstanbul Üniversitesi, Fen Bilimleri Enstitüsü Elektronik Müh. Anabilim Dalı, Şubat 2002.
- [12] Güz Ü., Gürkan H., Yarman B. S., "Ses İşaretlerinin Modellenmesi için Optimum Temel Tanım Fonksiyonlarının Belirlenmesinde Yeni Bir Yaklaşım", SİU'2002 Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Denizli, Türkiye, Cilt 2, sayfa 691-696, 12-14 Haziran 2002.
- [13] Gürkan H., Güz Ü., Yarman B. S., "EEG (ElectroEncephaloGram) İşaretlerinin Optimum Temel Tanım Fonksiyonları ile Modellenmesinde Yeni Bir Yaklaşım", SİU'2002 Sinyal İşleme ve İletişim Uygulamaları Kurultayı, Denizli, Türkiye, Cilt 2, sayfa 1204-1209, 12-14 Haziran 2002.
- [14] Newman A. J., "Model Reduction via the Karhunen Loeve Expansion Part I: An Exposition" ISR Technical Research Report T.R. 96-32, 2 April 1996.
- [15] Jolliffe I. T., "Principal Component Analysis", Springer-Verlag New York Inc., 1986.
- [16] Strang G., "Linear Algebra and its Applications", Academic Press, New York, 1980.
- [17] Zaim İ., A Novel Method to Represent ECG Signals via Predefined Personalized Signature and Envelope Functions, Lisans Tezi (Tez Danışmanı: Prof. Dr. B. Sıddık YARMAN), Işık Üniversitesi Müh. Fak., Haziran 2002.