

# KENDİ KENDİNİ DÜZENLEYEN HARİTALAR YÖNTEMİYLE TÜRKÇE SESLİ HARFLERİN SINIFLANDIRILMASI VE TANINMASI

*Emrah YÜRÜKLÜ\**  
*Osman H. KOÇAL\*\**

**Özet:** Lokal dinamik modelleme teknikleri kullanılarak zaman serilerini modellemek özellikle son yıllarda oldukça başarılı sonuçlar vermektedir. Kohonen'in 1990 yılında sunduğu 'Kendi Kendini Düzenleyen Haritalar' yöntemi ile lokal dinamik modelleme tekniğine farklı bir bakış açısı kazandırılmıştır. Bu yöntem ile, zaman serilerinden türetilen lokal dinamik modeller, sinyalin tüm dinamikleri oldukça başarılı ve kolay bir yöntemle gösterebilmektedir. Zamanla bu teknik pek çok alanda kendine uygulama alanı bulmuş, gerek Kohonen, gerekse diğer uzmanlar tarafından pek çok farklı versiyonu türetilmiştir. Yapılan çalışmada SOM yöntemi kısaca açıklanmış ve bu yöntem yardımıyla Türkçe sesli harfler için sınıflandırma ve tanıma uygulaması yapılmış ve sonuçları tartışılmıştır.

**Anahtar Kelimeler:** Kendi Kendini Düzenleyen Haritalar, Lokal Dinamik Modelleme, Ses Tanıma.

## The Classification and Recognition of Turkish Vowels with Self-Organizing Maps

**Abstract:** The easiness of putting the model into practice, and making signal or system dynamics and structure observable has made dynamic modeling for time series very popular for the last years. By the years, new versions and approaches of dynamic modeling have been developed and applied to different kind of signals and systems. Kohonen's suggestion was a new approach to local dynamic modeling which he had offered in 1990. The new technique's name was 'Self Organizing Maps'. The innovation of the new approach was its needlessness of the memory for saving the history of time series. It was because the whole model is updated with the new sample of time series. New versions of this technique are introduced in a lot of different kinds of applications by the years. In this work, 'Self Organizing Maps' technique is applied to Turkish vowels and worked on the advantages of the technique.

**Key Words:** Self-Organising Maps, Local Dynamic Modelling, Speech Recognition.

## 1. GİRİŞ

Sistem tanımlama ve zaman serisi tahminleri, en eski mühendislik problemlerinden biridir. Genel olarak tahmin yöntemi yaklaşımlarında, sistemin zaman serisini üreten parametrelerin arayışı üzerine yoğunlaşmıştır. Geleneksel yaklaşım istatistik ve lineer modeller üzerine kurulmuştur. Genel olarak tahminler yapılırken zaman serisinin lineer bir sistem tarafından üretilen sinyal ve Beyaz Gauss Gürültüsünün (WGN) toplamından oluştuğu varsayılmıştır. Böylelikle lineer modellemenin hedefi, zaman serisinin güç spektrumu ile uyuşacak modelin parametrelerin kestirimi ile sınırlanmıştır.

1990'lı yılların sonlarına doğru içinde dinamik modelleme adı verilen yeni bir perspektif ortaya çıkmıştır. Bu yeni yaklaşımda, zaman serileri, deterministik, otonomus (girişsiz) ve dinamik sistemlerin çıkışları olarak değerlendirilmiştir. Zaman serilerinin karmaşası, lineer durumdaki gibi dış kaynaklı rasgele uyarımların değil, dinamik sistemlerin

\* TOFAŞ Türk Otomobil Fabrikası A.Ş., ArGe, Elk./Eln. Tasarım Yöneticiliği, Bursa.

\*\* Yalova Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, Yalova.  
İletişim Yazarı: E. Yürüklü (emrah.yuruklu@tofas.com.tr)

yüksek dereceden ve nonlinear doğası ile ilişkilendirilmiştir. Bu yaklaşımda model sistem, nonlinear ya da zamanla değişen parametrelere sahip lineer bir sistem olarak seçilmiştir.

*Kendi Kendini Düzenleyen Haritalar* metodu, ilk defa 1990 yılında Kohonen tarafından yayınlanan “The Self-Organizing Map” isimli makale ile tanınmıştır. Bu tarihten sonra da, başta Kohonen olmak üzere, birçok gelişim ve adaptasyonlar görmüş ve çeşitli uygulamalarda kullanılmıştır. Birçok sinirsel ağ uygulamalarında diğer metotlardan üstünlüğünü kanıtlamış ve üzerinde en çok çalışılan yöntemlerden biri olmuştur. Bu yeni yöntemin en büyük avantajlarından biride zaman geçmiş değerlerini tutmak için bir hafızaya ihtiyaç duymamasıdır. Çünkü sistem her yeni gelen örnek ile kendini tümünden yenilemektedir. Büyük veritabanlarını dolaşmak (Kohonen 1995, Wang 1996), makinelerin ve işlem süreçlerinin durum diyagramlarını çıkartmak ve böylece sinyalleri modelleme, sınıflandırma ve karşılaştırma yapmak için (Kohonen 1995, Principe ve ark. 1998, Wang 1996, Lapidot 2001), modellenen sistemin çıkışları hakkında geleceğe yönelik tahminler yapmak ve böylece gürültü bastırmak için (Principe ve ark. 1998, Wang 1996, Lapidot 2001) *Kendi Kendini Düzenleyen Haritalar* kullanılarak birçok çalışmalar yapılmış ve başarılarla imza atılmıştır. Günümüzde de *Kendi Kendini Düzenleyen Haritalar* metodu üzerinde halen çalışmalar devam etmekte yeni versiyonları ve farklı alanlar için uygulamaları türetilmektedir.

## 2. KENDİ KENDİNİ DÜZENLEYEN HARİTALAR

Kohonen, SOM (Kendi Kendini Düzenleyen Haritalar) algoritmasını ilk olarak rasgele boyuttaki bir giriş sinyalini daha düşük (bir ya da iki) boyutlu ayrık gösterimli topolojik komşular halinde temsil etmek amacıyla geliştirmiştir.  $\Phi: X \rightarrow A$  gösterimi,  $X$  giriş uzayını,  $A$  ayrık çıkış uzayına dönüştüren SOM haritasını ( $\Phi$ ) temsil etsin. SOM haritası  $\Phi$ , Kohonen’in kendi ifadesiyle “Bilginin yapı yoğunluğuna yaklaşmak üzere,  $X$  giriş sinyal uzayına uydurulmuş,  $A$ , noktaların elastik ağı”dır (Kohonen 1995).  $A$  uzayına ait her olasılık elemanı ( $i \in A$ ) bir referans vektör ( $w_i$ ) ile indekslenir. SOM algoritmasının yürütülmesinin sonunda bu olasılık elemanları  $x$  giriş vektörünü temsil eden kümelerle ait hücreler olacaktırlar.  $x$  giriş sinyaline SOM algoritmasının tepkisi hangi hücrenin kazandığına karar verilmesiyle başlar.  $i^\circ$  notasyonu kazanan hücreyi ve  $A$  ise  $A$  uzayındaki nokta sayısını belirtmek üzere, kazanan hücrenin belirlenmesi şu kritere göre yapılmaktadır:

$$i^\circ = \min \text{dist}(w_i - x) \quad i=1, \dots, A \quad (1)$$

Giriş sinyali olan  $x$  vektörü ile  $A$  uzayındaki tüm hücreler arasındaki mesafe, Öklid Mesafesi gibi bir uzaklık hesabıyla hesaplanır ve bu hesaplama sonucunda en yakın mesafedeki  $w_i$  hücresi yarışmayı kazanan hücre olur ( $i^\circ$ ) (Kohonen 1995). Bu bakış açısıyla,  $w_i$  referans vektörlerinin ağırlık vektörlerine dönüştürülmesi ile SOM bir vektör nicemleyici olarak ta görev yapabilmektedir.

SOM’un amacının yüksek yoğunluklu bilgi olması sebebiyle,  $A$  uzayındaki hücrelerin  $X$  uzayı ile uyuyacak şekilde yakınsaması gerekmektedir. En yakın mesafedeki uyum için ise uygun bir öğrenme kuralı gerekmektedir. SOM algoritmasının kullandığı eğitim ve  $w_i$  katsayılarını güncelleme algoritmasının en genel hali,  $n$  işlem anındaki örneğin indisi olmak üzere, eşitlik (2)’te görülmektedir.

$$w_i(n+1) = \begin{cases} w_i(n) + \eta(n)[x(n) - w_i(n)], & i = i^\circ \\ w_i(n), & \text{diğer} \end{cases} \quad (2)$$

Burada görülen öğrenme kuralı Kohonen tarafından önerilen Yarışmacı Öğrenme Kuralının en genel halidir.  $\eta(n)$  katsayısı öğrenme oranıdır ve sabit bir sayı veya gelen örneğin indisine göre değişen bir katsayı şeklinde de olabilmektedir. Kohonen öğrenme kuralı ile diğer yarışmacı öğrenme kurallarının arasındaki fark Kohonen öğrenmesinin ağırlık katsayılarını değiştirme şeklindedir. Kohonen öğrenme kuralında kazanan hücre ile birlikte diğer hücrelerinde bir kısmının veya hepsinin ağırlık katsayılarının güncellenmesi belli bir kurala bağlanabilmektedir.

### 3. SES SİNYALLERİ İÇİN SOM HARİTALARININ OLUŞTURULMASI

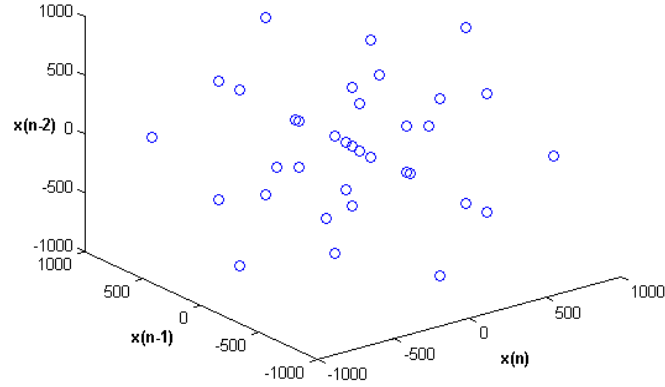
Bu güne kadar konuşmacı ve ses tanıma üzerine pek çok çalışma yapılmış ve günümüzde artık çok başarılı sonuçlar elde edilmektedir. Son yıllarda Türkçe sesler üzerine yapılan çalışmalarda belirgin bir artış görülmektedir. Bu çalışmada Türkçe Sesli harflerin SOM modellenmesi yapılacak daha sonra konuşmacı ve sesli harf sınıflandırmaları yapılarak ses ve konuşmacı tanımada SOM Haritalarının başarımı araştırılmıştır. Çalışmada, ses ve konuşmacı tanıma işlemlerinin genelde periyodikimsi bir yapı sergileyen sesli harfler üzerinde çalışılarak yapıldığı düşünülerek sesli harfler baz alınmıştır (Lapitod 2001).

Bir sinyalin SOM Haritasının çıkarılmasından önce SOM Algoritması için belirlenmesi gereken bazı parametreler bulunmaktadır. Bunlar öncelikle sinyalin kaç boyutlu uzaya gömüleceği (faz uzayının gömme boyutu), faz uzayının gecikme süresi ve ayrıca SOM Algoritmasının diğer başlangıç koşulları (haritadaki hücre sayısı) ve güncelleme (öğrenme katsayısı ve güncelleme kuralı) parametreleridir. Bu parametrelerin her biri modellenecek sinyale göre değişiklikler göstermektedir. Giriş sinyaline uygun parametrelerin belirlenebilmesi için çeşitli yöntemler bulunmaktadır. Bu çalışmada; gömme boyutunun belirlenebilmesi için *Hatalı Komşuluklar Yöntemi* (Abarbanel 1996), gecikme süresinin belirlenmesi için de *Ortak Enformasyon Hesabı* (Abarbanel 1996) kullanılmıştır.

Deney seti olarak beşi erkek ve üçü bayan olmak üzere toplam sekiz kişinin her birinin toplam sekiz sessiz harfi on kez konuşmasından elde edilmiş 640 adet ses kaydı kullanılmıştır. Ses kayıtları olabildiğince sessiz ortamlarda 11025 Hz frekansında 8bit mono olarak PCM formatında kaydedilmiştir. Her ses kaydının süresi 600 ms belirlenmiştir. Bu durumda her bir ses kaydı 6614 örnekten oluşmaktadır.

#### 3.1 Ses Dosyaları için SOM Parametrelerinin Seçimi

Hatalı Komşular Yöntemi Yardımıyla tüm deney seti üzerinde yapılan çalışmada gömme boyutunun tüm deney setinin %88'inde üç, %12'sinde de dört olarak hesaplandığı görülmüştür. Gözlem yapılabilmesi ve deney setinin büyük çoğunluğunun gömme boyutunun üç olarak hesaplanması sebebiyle tüm deney seti için gömme boyutu üç olarak alınmıştır. Gecikme süresinin seçiminde ise her kayıt için Ortak Enformasyon Yöntemiyle hesaplanmış ve SOM Haritalarının oluşturulmasında da bu gecikme süreleri kullanılmıştır.



**Şekil 1:**

*Ses verileri için uygulanacak SOM Algoritmasındaki ağırlık hücrelerinin başlangıç koşullarını teşkil eden üç boyutlu vektör uzayı*

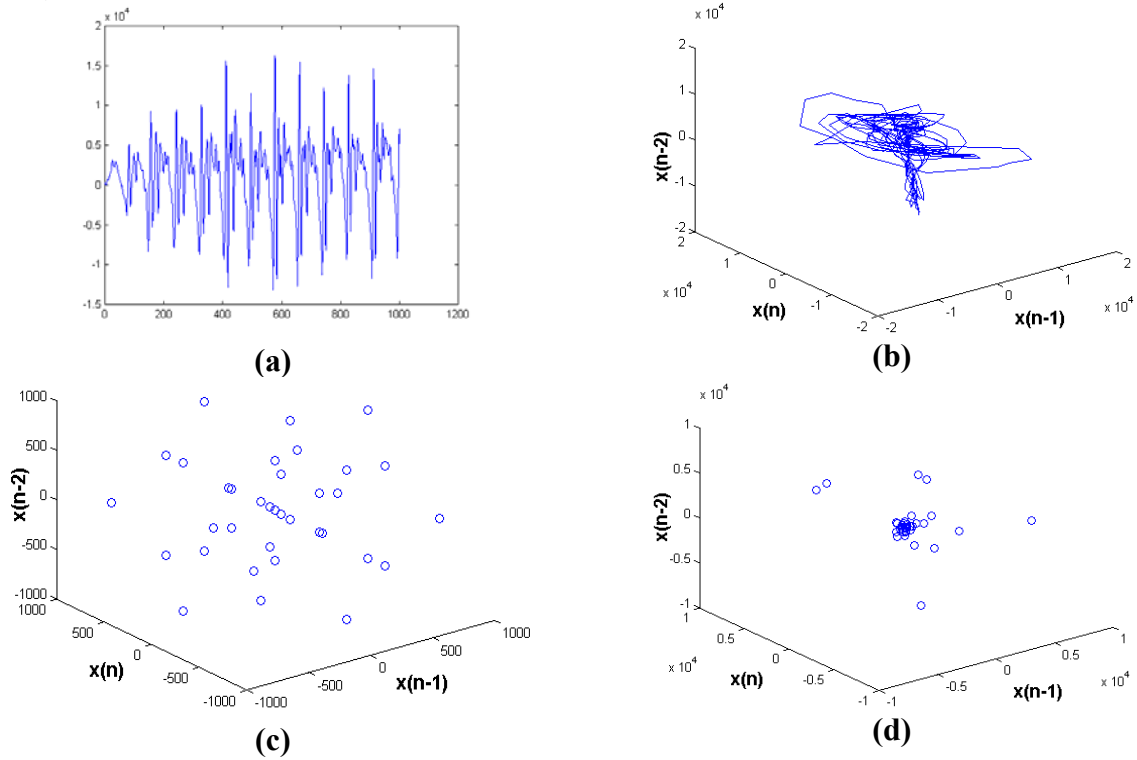
Başlangıç koşulları olarak, SOM Haritasında 35 adet hücre eşit uzaklıklara yerleştirilerek şekil.1’de gösterilen vektör uzayı kullanılmıştır. SOM Algoritmasının eşitlik (2)’de bahsedilen güncelleme kuralı olarak ise eşitlik (3)’te gösterilen güncelleme kuralı kullanılmıştır.

$$w_i(n+1) = \frac{mc_i(n)}{mc_i(n)+1} \cdot w_i(n) + \frac{x(n)}{mc_i(n)+1} \quad (3)$$

Burada  $x$  o andaki veriye ait örnek vektörü,  $w_i$  kazanan i.h ücreyi,  $mc_i$  ise  $i$ . hücrenin o ana kadarki toplam kazanma sayısını temsil etmektedir. Bu kural kullanılarak örnek sayısının artması ile birlikte, son gelen örneklerin, SOM Haritası üzerindeki etkisinin azalması hedeflenmiştir.

### 3.2 Örnek Bir Ses Sinyalinin SOM Haritasının Oluşturulması

Şekil.2 ile bir ses sinyalinin SOM Haritasının oluşturulmasının zaman serisinden SOM Haritasına kadar adımları gösterilmiştir. Şekil.2(a) modellenecek ses sinyalin zaman serisi şeklinde gösterimidir. Şekil.2(b) ise ses sinyalinin 3 gömme boyutu ve 1 gecikme süresi ile elde edilmiş faz uzayının gösterimidir. Şekil.2(c) üç boyutlu SOM Haritasının başlangıç durumudur. Şekil.2(d) ise SOM Haritasının kaynak ses sinyali ile eğitilmesi sonucu elde edilen SOM Haritasıdır. Burada SOM Haritasının (Şekil.2(d)) sinyalin faz uzayı ile benzerliğine dikkat edilmelidir. SOM algoritması aslında sinyalin faz uzayını nicemleyerek bütün faz uzayını belli sayıda vektör ile temsil edilmesini sağlamaktadır. Faz uzayının boyutu ve gecikme süresi doğru seçilirse faz uzayının sinyalin tüm dinamiklerini gösterebildiği, dolayısı ile içerdiği kabul edilir (Abarbanel 1996). Böylelikle tüm dinamikleri içeren bir faz uzayı ile modellenen bir SOM Haritasının sinyale ait dinamikleri içermeye kapasitesine sahip olduğunu söyleyebiliriz.



**Şekil 2:**

*Bir Ses Sinyalinin Zaman Serisi (a) ve Faz Uzayı Gösterimi (Gömme Boyutu=3, Gecikme Süresi=1)(b), ve SOM Haritasının Başlangıç (c) ve Eğitim Sonrası Durumu (d).*

#### 4. DENEYSEL SONUÇLAR

Deney setimizde bulunan 640 adet ses sinyalinin tümünün SOM Haritaları, Kısım 3.1’de belirtilen şartlar ile oluşturulmuştur. Böylelikle 35x3 matristen oluşan 640 adet SOM Haritası elde edilmiştir. Bu veritabanındaki, SOM Haritaları arasında kişiler ya da sesler bazında bir ilişkinin var olup olmadığının sorgulanması için ise KNN sınıflandırma yöntemi kullanılmıştır. Bu sınıflandırma yapılırken KNN Yöntemi için  $K$  komşuluk sayısı 5 olarak seçilmiştir.

Tablo 1 ile veritabanındaki SOM Haritaları arasında kişiler ve sesli harfler bazında yapılan sınıflandırma performansları yer almaktadır. Bu performans sonuçları ile aynı seslere ait ses kayıtlarından çıkartılan SOM haritalarının aralarında belli bir benzerlik olduğunu söyleyebiliriz. Fakat aynı kişilere ait SOM haritalarının aralarındaki benzerlik bu kadar belirgin değildir. Bu durum kişilere ait dinamiklerin SOM haritası oluşturulurken bir şekilde bastırılmasından kaynaklanmaktadır. Fakat SOM Haritaları sesli harfler ile ilgili dinamikleri gayet iyi korumakta ve bu korunum sayesinde sesli harfler arasında bir benzerlik görülmektedir.

**Tablo 1. Kişiler Bazında ve Harfler Bazında Yapılan Sınıflandırma Performansları**

Kişi	Bay1	Bay2	Bay3	Bay4	Bay5	Bayan1	Bayan2	Bayan3	Ortalama
<b>Doğru Karar Verme Oranı (%)</b>	80	67	55	75	83	85	54	64	70.4
<b>Sesli Harf</b>	‘a’	‘e’	‘ı’	‘i’	‘o’	‘ö’	‘u’	‘ü’	<b>Ortalama</b>
<b>Doğru Karar Verme Oranı (%)</b>	98	88	95	100	85	86	91	94	92.1

## 5. SONUÇ

Yapılan çalışmada, ilk defa Kohonen tarafından tanıtılmış Kendi Kendini Düzenleyen Haritalar (SOM) yöntemi kullanılarak Türkçe ses karakterlerinin özellik haritaları çıkartılmıştır. Daha sonra çıkartılan özellik haritaları kendi aralarında karşılaştırılmış ve benzer kişilerin veya benzer seslere ait ortak özellikler araştırılmıştır. Yapılan çalışmalar göstermiştir ki, SOM yöntemi ile elde edilen ses sinyallerine ait özellik haritalarında, farklı kişilere ait aynı sesli harflerin kayıtları arasında ortak bazı özellikler bulunmaktadır.

Çalışmanın önemi, konuşma tanıma işleminin, konuşma sinyalinin içerisinde rasgele diyebileceğimiz tavırlar gösteren sessiz harfler yerine daha çok yarı-periyodik yapı gösteren sesli harfler üzerinde çalışılarak yapıldığı düşünülürse daha iyi anlaşılacaktır. Dolayısıyla bu çalışmanın gelecek çalışması önce hece tanıma daha sonra da kelime tanıma olacaktır. Konuşma tanıma işlemini yüksek doğrulukla başarabilen bir sistem tasarlandığında ise, böyle bir sistemin kullanılabilmesi alanlar sınırsızdır. Bu sistem ile insansız telefon operatörleri, sesli komutlarla çalışan makineler, konuşulani dikte eden konuşma-yazı operatörleri gibi büyük kolaylıkların yolu açılacaktır. Günümüzde böyle bir sistem tasarlamak için pek çok çalışmalar yapılmaktadır. Fakat ne yazık ki bu çalışmaların büyük çoğunluğu İngilizce dilindedir. Yapılan bu çalışma ile Türkçe dili için böyle bir sistem tasarımının ön adımları atılmıştır. Önerilen yöntem Türkçe dili için gayet başarılı sonuçlar elde etmekte ve daha sonraki çalışmalar için umut vaat etmektedir.

## KAYNAKLAR

1. Kohonen, T. (1995). *Self-Organizing Maps*. New York: Springer-Verlag.
2. Abarbanel, H.D.I. (1996). *Analysis of Observed Chaotic Data*. New York: Springer.
3. Wang, L. (1996). *Local Dynamic Modeling with Self-Organizing Feature Map*, *PhD Dissertation*, University of Florida, Division of Electrical and Computer Engineering.
4. Lapidot, I. (2001). *Unsupervised Speaker Recognition*, *PhD Dissertation*, Ben-Gurion University of the Negev, Beer-Sheva.
5. Principe, J.C., Wang, L., Motter, M.A. (1998). *Local Dynamic Modeling with Self-Organizing Maps and Applications to Nonlinear System Identification and Control*, *Proceedings of the IEEE*, Vol. 86. No. 11.

Makale 19.03.2012 tarihinde alınmış ve 03.04.2012 tarihinde kabul edilmiştir.