

**SINIF İÇİ ve SINIFLAR ARASI DAĞILIMLARDAN
ELDE EDİLEN ÖLÇÜTLERİN BİRLEŞTİRİLEREK
FONEM TANIMADA KULLANILMASI**

Mehmet KOÇ

YÜKSEK LİSANS TEZİ

Elektrik-Elektronik Mühendisliği Anabilim Dalı

Ağustos 2006

**COMBINING CRITERIA OBTAINED FROM
WITHIN AND BETWEEN CLASS SCATTERS
FOR PHONEME RECOGNITION**

Mehmet KOÇ

MASTER OF SCIENCE THESIS

Department of Electrical-Electronics Engineering

August 2006

SINIF İÇİ ve SINIFLAR ARASI DAĞILIMLARDAN ELDE EDİLEN ÖLÇÜTLERİN
BİRLEŞTİRİLEREK FONEM TANIMADA KULLANILMASI

Mehmet KOÇ

Osmangazi Üniversitesi
Fen Bilimleri Enstitüsü
Lisansüstü Yönetmeliği Uyarınca
Elektrik-Elektronik Mühendisliği Anabilim Dalı
Telekomünikasyon-Sinyal İşleme Bilim Dalında
YÜKSEK LİSANS TEZİ
Olarak Hazırlanmıştır

Danışman: Yrd.Doç.Dr. Rifat EDİZKAN

Ağustos 2006

Mehmet KOÇ' in YÜKSEK LİSANS tezi olarak hazırladığı “Sınıf İçi ve Sınıflar Arası Dağılımlardan Elde Edilen Ölçütlerin Birleştirilerek Fonem Tanımada Kullanılması” başlıklı bu çalışma, jürimizce lisansüstü yönetmeliğinin ilgili maddeleri uyarınca değerlendirilerek kabul edilmiştir.

Üye : Yrd.Doç.Dr. Rifat EDİZKAN

Üye : Prof.Dr. Atalay BARKANA

Üye : Doç.Dr. M.Bilginer GÜLMEZOĞLU

Fen Bilimleri Enstitüsü Yönetim Kurulu'nun tarih ve sayılı kararıyla onaylanmıştır.

Prof. Dr. Abdurrahman KARAMANCIOĞLU

Enstitü Müdürü

ÖZET

Sınıflandırıcı birleřtirmedeki amaç örüntü tanımada en iyi sınıflandırma başarımını elde etmektir. Sınıflandırıcıların üstün özelliklerini birleřtirerek daha iyi sınıflandırma yapılabilir. Sınıflandırıcı topluluğunun başarımı, sınıflandırıcıların tek başlarına kullanılmasıyla elde edilen başarımdan daha iyi olacaktır. Örüntü sınıflamada sınıf içi ve sınıflar arası dağılımları kullanan ölçüt sınıflandırma başarımını arttıracaktır. Bu çalışmada, sırasıyla sınıf içi ve sınıflar arası dağılımlardan ortak vektör yaklaşımı(OVY) ve ana bileşen analizi(PCA) kullanarak elde edilen sınıflandırıcılar yapay sinirağı ile birleřtirilmiřtir. İki-sınıf problemi için hedeflenen sınıflandırıcının başarımı, iki boyutlu yapay veriler ve TIMIT veri tabanındaki seslilerin sınıflandırılmasından elde edilmiřtir. Sınıflandırıcının başarımı ayrıca sınıf içi ve sınıflar arası dağılımları kullanan doğrusal ayırtaç analizi(LDA)'nin başarımları ile karşılařtırılmıřtır.

Anahtar Kelimeler: Sınıflandırıcı birleřtirme, sesli tanıma, altuzay yöntemleri, ortak vektör yaklaşımı, doğrusal ayırtaç analizi

SUMMARY

The aim of classifier combination is to achieve the best classification performance in pattern recognition. Better classification can be achieved by combining the superior characteristics of the classifiers. The performance of ensemble classifier will be better than the performance of each individual classifiers. In pattern classification, the metric that considers within-class and between-class scatters will improve the classification performance. In this study, the classifiers that are derived from within-class and between-class scatters using the common vector approach(CVA) and the principal component analysis(PCA) respectively are combined in a neural network architecture. The performance of proposed classifier for two-class problem is obtained from the classification of two dimensional artificial data and the vowels in TIMIT database. The performance of the classifier is also compared with the performance of linear discriminant analysis(LDA) that uses within-class and between-class scatters.

Key Words: Combining classifier, vowel recognition, subspace methods, common vector approach, linear discriminant analysis.

TEŐEKKÜR

Yüksek lisansın gerek ders, gerekse tez aşamasında, bana danışmanlık ederek, beni yönlendiren ve her türlü olanağı sağlayan ve büyük sabır gösteren danışmanım Yrd.Doç.Dr. Rifat EDİZKAN'a; danışmanlığını esirgemeyen, yol gösteren, Prof.Dr. Atalay BARKANA'ya teşekkür ederim.

Ayrıca yüksek lisans boyunca benden maddi, manevi desteklerini esirgemeyen, sabır gösteren aileme de teşekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
ÖZET	IV
SUMMARY	V
TEŞEKKÜR.....	VI
İÇİNDEKİLER	VII
ŞEKİLLER DİZİNİ	VIII
1. GİRİŞ	1
2. ALTUZAY TEKNİKLERİ	4
2.1 Ortak Vektör Yaklaşımı (OVY).....	4
2.1.1 Yetersiz Veri Durumu ($n \geq m$)	4
2.1.2 Yeterli Veri Durumu ($n < m$)	10
2.2 Ana Bileşen Analizi (PCA).....	13
2.3 Doğrusal Ayırtaç Analizi(LDA)	16
3. SINIFLANDIRICI BİRLEŞTİRME	20
3.1 Sınıflandırıcı Birleştirme Yöntemleri	20
3.2 Sınıflandırıcı Birleştirme Mimarıleri	23
3.3 Sınıf-İçi ve Sınıflar Arası Dağılımlardan Elde Edilen Ölçütlerin Birleştirilmesi	24
3.3.1 Enküçültme Problemi	24
3.3.2 Enbüyütme Problemi	25
3.3.3 Sınıf içi ve Sınıflar arası Ölçütlerin Birleştirilmesi	26
3.3.4 Sınıflandırıcı Birleştirmenin İki-Sınıf Problemine Uygulanması	30
4. DENEYSEL ÇALIŞMALAR	34
4.1 TIMIT Veritabanı	34
4.2 Sesli Sınıflama	35
4.2.1 Özdeğer Seçimleri	37
4.2.2 Ölçekleme Katsayısı	38
4.2.3 Sınıflama Başarımları	39
5. SONUÇLAR VE ÖNERİLER	45
6. KAYNAKLAR DİZİNİ	47

ŞEKİLLER DİZİNİ

Sayfa

Şekil 2.1 Aynı sınıfa ait iki boyutlu iki vektör ve bu vektörlerin ortak vektörlerinin gösterilişi.....	7
Şekil 2.2 C_1 ve C_2 sınıfları için ana bileşenler.....	19
Şekil 2.3 C_1 ve C_2 sınıfların ait vektörlerin ana bileşen üzerine izdüşümü.....	19
Şekil 3.1 Paralel sınıflandırıcı birleştirme.....	23
Şekil 3.2 Seri sınıflandırıcı birleştirme.....	23
Şekil 3.3 OVY in iki sınıf probleminde yapay sinir ağı formunda gösterilmesi.....	25
Şekil 3.4 PCA sınıflandırıcısının yapay sinir ağı şeklinde gösterilmesi.....	26
Şekil 3.5 OVY ve PCA sınıflandırıcılarının yapay sinir ağı ile birleştirilmesi.....	27
Şekil 3.6 $w_{1_{c_1}} - w_{1_{c_2}}$ düzleminde oluşturulan ızgara.....	29
Şekil 3.7 Normlar ve ağırlıklar arasındaki ilişki.....	30
Şekil 3.8 C_1 ve C_2 sınıfları.....	31
Şekil 3.9 Test kümesinin OVY ile sınıflandırılması.....	32
Şekil 3.10 Test kümesinin PCA'den elde edilen ölçüt ile sınıflandırılması.....	32
Şekil 3.11 Test kümesinin LDA ile sınıflandırılması.....	33
Şekil 3.12 OVY ve PCA'den elde edilen ölçütlerin birleştirilmesi ile test kümesinin sınıflandırılması.....	33
Şekil 4.1 'ey' seslisinin ortak değişinti matrisinin özdeğerleri.....	38
Şekil 4.2 'ey' seslisi için $F_{1_{c_1}}$ ve $\frac{1}{F_{2_{c_1}}}$ normlarının histogramları.....	39
Şekil 4.3 $w_{1_{c_1}}$ ve $w_{1_{c_2}}$ ağırlıklarının ızgara üzerinde aldıkları değerlere göre 'ey' ve 'ae' seslilerinin birleştirilen sınıflandırıcı ile tanınma yüzdeleri.....	44
Şekil 4.4 Şekil 4.3'deki grafiğin bir kesitinin büyütülmüş hali.....	44

1. GİRİŞ

Sınıflandırıcı birleřtirmedeki amaç daha iyi bir sınıflandırma başarımı elde etmektir. Sınıflandırıcıların sınıflandırmada kullandıkları üstün özellikleri birleřtirilerek, sınıflandırıcıların tek başına yaptığı sınıflandırmadan daha iyi bir sınıflandırma yapılabilir.

Örüntü sınıflamada sadece sınıf içi dağılımları veya sadece sınıflar arası dağılımları kullanmak, bir çok durum için istenilen başarımları vermemektedir. Bunun için sınıf içi ve sınıflara arası dağılımları birlikte kullanan ölçütler sınıflamada kullanılmaktadır. Bu ölçütlerden en çok bilineni Fisher'in doğrusal ayırtaç analiz ölçütüdür(Duda, et al., 2001; Bishop, 1996; Katz, et al., 2002). Bu ölçütlerin yanında çoklu sınıflandırıcılar birleřtirilmesi de çoklu sınıf problemlerine uygulanmaktadır. Çoklu sınıflandırıcılar ile karmaşık örüntü tanıma problemlerinde istenilen başarımlar elde edilebilmektedir. Çoklu sınıflandırıcılar, yüz tanıma(Lu, et al., 2003), el yazısı karakter tanıma(Cao, et al., 1995), kişisel kimlik doğrulama(Kittler, et al., 1998), ses tanıma (Felföldi, 2003) ve kişi tanıma(Radova and Psutka, 1997) gibi çeşitli örüntü tanıma problemlerine uygulanmıştır.

Birleřtirilecek sınıflandırıcıların çıkış türlerine göre birleřtirme yöntemine karar verilir. Birleřtirilen sınıflandırıcılardan sadece sınıf etiketleri çıkış olarak alınıyorsa çoğunluk oyu veya etiket sıralaması kullanılır. Sınıflandırıcı çıkışları sonsal olasılık gibi sürekli ise ortalama veya doğrusal bileşim kullanılabilir. Eğer sınıflandırıcı çıkışları bulanık üyelik değerleri veya güven değeri ise bulanık kurallar, güven fonksiyonları veya Dempster-Shafer teknikleri kullanılır.

Birleřtirilecek sınıflandırıcıların seçiminde dikkat edilmesi gereken konulardan biri çeşitlilik(Kuncheva, 2004). İki sınıflandırıcı eğer hataları farklı veri kümelerinde yapıyorsa bu sınıflandırıcılar çeşitlidir denilebilir(Dietterich, 2000). Birleřtirilecek

sınıflandırıcıların hatalarının farklı kümelerde olması sınıflandırma başarımını arttıracak etkenlerdendir.

Çoklu sınıflandırıcılara hibrit yöntemler, karar birleştirme, çoklu uzmanlar, sınıflandırıcı topluluğu, algılayıcı kaynaşımı da denir. Çoklu sınıflandırıcılar birkaç şekilde gruplanabilir. Bunlardan ilki sınıflandırıcıların birleştirilmesine göre olanıdır. Sınıflandırıcı birleştirme paralel, seri ve hiyerarşik olmak üzere üç ana sınıfta toplanabilir. Çoklu sınıflandırıcılar ayrıca kaynaşımın giriş ve çıkışı arasındaki haritalamaya görede doğrusal veya doğrusal olmayan haritalama olarak da sınıflandırılabilir. Doğrusal birleşim en basit yöntemlerdendir. Bu yöntemlerde birleştirilecek her bir sınıflandırıcı belirli bir ağırlıkla çarpılır. Ağırlıklı ortalama(Heskes, 1997), bulanık integral(Gader et al, 1996) doğrusal birleşim yöntemlerdendir. Bunun yanında çoğunluk oyu (Ji and Ma, 1997) ve enbüyük, doğrusal olmayan çoklu sınıflandırıcı yöntemleridir. Birleştirme yöntemleri gösterim şekilleri olarak da iki gruba ayrılabilir. Bunlardan ilkinde farklı sınıflandırıcılar giriş vektörünün aynı gösterimini kullanır. Diğerinde ise farklı sınıflandırıcılar giriş vektörünün farklı gösterimlerini kullanırlar. Bu farklı gösterimler, farklı algılayıcılardan veya veri kümesinden farklı parametreler üretilerek elde edilir.

Bu çalışmada sınıf içi ve sınıflar arası dağılımları için ölçütler OVY ve PCA'den elde edilmiştir. OVY yönteminde bir sınıfa ait vektörlerindeki farklılıklar atıldığında geriye o sınıfın değişmez özellikleri içeren vektör kalır. Bu vektöre ortak vektör denir (Gülmezoğlu et al., 1999; Gülmezoğlu, et al., 2001). OVY'de bir sınıfa ait vektörlerin oluşturduğu uzay, fark ve farksızlık olmak üzere iki alt uzaya ayrılır. OVY'de yeterli($n < m$) ve yetersiz($n \geq m$) veri durumu için gerekli matematiksel türetmeler yapılmıştır. Burada m , sınıfa ait vektör sayısını; n ise vektör boyutunu göstermektedir. OVY ile sınıflandırma, bilinmeyen vektörün farksızlık uzayına izdüşümü ile ortak vektör arasındaki Öklit uzaklığına bakılarak yapılır. Bilinmeyen vektör, ortak vektörüne uzaklığı en küçük olan sınıfa atanır. PCA yöntemi aslında boyut indirgeme yöntemidir(Kirby and Sirovich, 1990; Wang and Paliwal, 2002). Fakat sınıflandırma içinde kullanılabilir (Günel, et al., 2005). PCA yönteminde çok boyutlu veriler en az bilgi kaybıyla daha düşük boyutlara indirgenmektedir. PCA, izdüşümü alınan verilerin saçılımını enbüyük

yapan, boyut indirgeyen doğrusal bir izdüşüm belirler. Böylece en az bilgi kaybıyla boyut indirgenmiş olur. Ancak boyut indirgeme genelde sınıflandırma için önemli olan bazı bilgilerin kaybolmasına neden olur.

Tezin 2. bölümünde OVY, PCA ve LDA hakkında bilgiler verilmiştir. Üçüncü bölümde sınıflandırıcı birleştirme yöntemleri anlatılmıştır. Ayrıca bu bölümde, OVY'den ve PCA'den üretilen iki sınıflandırıcının birleştirilmesi ve mimarisi hakkında bilgiler verilmiştir. Dördüncü bölümde, 3. bölümde önerilen sınıflandırıcının ve LDA yönteminin iki boyutlu veriler ve TIMIT veritabanına uygulanmasıyla elde edilen sonuçlar karşılaştırılmıştır. Beşinci bölümde, elde edilen sonuçlar hakkında bazı yorumlar ve gelecek için öneriler bulunmaktadır.

2. ALTUZAY TEKNİKLERİ

Bu bölümde sınıf içi ve sınıflar arası dağılımlardan elde edilen altuzaylar hakkında bilgi verilecektir.

2.1 Ortak Vektör Yaklaşımı (OVY)

Bir sınıftaki verilerin o sınıfa ait olan özellikleri ve verilerin birbirinden farklılık gösteren özellikleri vardır. Ortak vektör, bir sınıfa ait özellik vektörlerindeki farklılıklar çıkarıldıktan sonra geriye kalan ve o sınıf için değişmeyen özelliklerin oluşturduğu vektöre denir(Gülmezoğlu, et al., 2001). Ortak vektör yaklaşımı ses, kişi tanıma ve yüz tanıma gibi örüntü tanıma uygulamalarında kullanılan bir altuzay yöntemidir (Gülmezoğlu et al., 1999; Gülmezoğlu, et al., 2001; Gülmezoğlu, et al, 2004; Edizkan, et al, 2005; Çevikalp, et al., 2006). OVY'nin temelinde bir sınıfa ait özellik vektörlerinin ortak özellikleri gösteren tek vektörün varlığı yatar. Örneğin; ses sinyali söyleyen kişiden ve çevresel faktörlerden dolayı farklılıklar gösterebilir. Ortak vektör yaklaşımında, bir kelime sınıfındaki söyleyişlerde çevresel ve kişisel faktörlerden dolayı oluşan farklılıklar elenir ve bu kelime sınıfının ortak özelliklerini taşıyan ve sınıf için değişmez olan bir vektör elde edilir. Ortak vektörde, yetersiz($n \geq m$) ve yeterli($n < m$) veri durumu olmak üzere iki durum söz konusudur. Burada m sınıftaki veri sayısı n ise verilerin boyutunu göstermektedir.

2.1.1 Yetersiz Veri Durumu ($n \geq m$)

R^n n -boyutlu bir vektör uzayı olsun. Bir sınıfa ait öznitelik vektörleri

$$\vec{a}_i = \vec{a}_{i,dif} + \vec{x} \quad i=1,2, \dots,m \quad (2.1)$$

biçiminde yazılabilir. (2.1) denkleminde, m tane vektör denklemi ve $m+1$ tane bilinmeyen vektör vardır. Bu durumda \vec{x} için sonsuz çözüm vardır..Ortak vektörü elde etmek için aşağıdaki eniyileme ölçütünü ele alalım.

$$F_1 = \left\| \vec{a}_{1,dif} \right\|^2 + \cdots + \left\| \vec{a}_{m,dif} \right\|^2 \quad (2.2)$$

$$F_1 = \left\| \vec{a}_1 - \vec{x} \right\|^2 + \cdots + \left\| \vec{a}_m - \vec{x} \right\|^2 \quad (2.3)$$

F_1 'in \vec{x} 'e göre türevi alınırsa

$$\begin{aligned} \frac{\partial F_1}{\partial \vec{x}} &= 2(\vec{a}_1 - \vec{x}) + \cdots + 2(\vec{a}_m - \vec{x}) \\ &= 2 \left[(\vec{a}_1 + \vec{a}_2 + \cdots + \vec{a}_m) - m\vec{x} \right] \\ &= 2m \left[\frac{\vec{a}_1 + \vec{a}_2 + \cdots + \vec{a}_m}{m} - \vec{x} \right] \\ &= 2m \left[\vec{a}_{ave} - \vec{x} \right] \\ \vec{x} &= \vec{a}_{ave} \end{aligned} \quad (2.4)$$

elde edilir. Bu en iyileme ölçütüne göre ortak vektör, sınıfın ortalamasına eşit çıkar.

Diğer bir en iyileme ölçütü ile de bir ortak vektör bulunabilir. Bu ölçüt için fark altuzayı tanımlanmalıdır. Bir sınıf ait öznitelik vektörlerinin \vec{a}_i , ($i=1,2,\dots,m$), seçilen bir referans vektörüne göre farkları $\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_{m-1}\}$ fark kümesini oluşturur.

$$\begin{aligned} \vec{b}_1 &= \vec{a}_2 - \vec{a}_1 \\ \vec{b}_2 &= \vec{a}_3 - \vec{a}_1 \\ &\vdots \\ \vec{b}_{m-1} &= \vec{a}_m - \vec{a}_1 \end{aligned} \quad (2.5)$$

$\mathbf{A} = \{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_{m-1}\}$ tarafından gerilen altuzayı \mathbf{B} ile gösterelim. Bu durumda

$$\mathbf{B} = span\mathbf{A} = span\{\vec{b}_1, \vec{b}_2, \dots, \vec{b}_{m-1}\} \quad (2.6)$$

olur ve \mathbf{B} 'ye $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_m$ vektörlerinin fark uzayı adı verilir. Bu taban vektörleri kullanarak Gram-Schmidt ortagonalleştirme metoduyla ortanormal $\{\vec{z}_1, \vec{z}_2, \dots, \vec{z}_{m-1}\}$ vektör kümesi elde edilir.

$$\begin{aligned}
\vec{d}_1 &= \vec{b}_1 & \vec{z}_1 &= \frac{\vec{d}_1}{\|\vec{d}_1\|} = \frac{\vec{b}_1}{\|\vec{b}_1\|} \\
\vec{d}_2 &= \vec{b}_2 - \langle \vec{b}_2, \vec{z}_1 \rangle \vec{z}_1 & \vec{z}_2 &= \frac{\vec{d}_2}{\|\vec{d}_2\|} \\
&\vdots & &\vdots \\
\vec{d}_{m-1} &= \vec{b}_{m-1} - \sum_{i=1}^{m-2} \langle \vec{b}_{m-1}, \vec{z}_i \rangle \vec{z}_i & \vec{z}_{m-1} &= \frac{\vec{d}_{m-1}}{\|\vec{d}_{m-1}\|}
\end{aligned}$$

Bu durumda herhangi bir \vec{a}_i vektörünün bu ortonormal taban vektörleri üzerine izdüşümü aşağıdaki ifadeden elde edilir:

$$\bar{a}_i = \sum_{k=1}^{m-1} \langle \vec{a}_i, \vec{z}_k \rangle \vec{z}_k \quad i = 1, 2, \dots, m \quad (2.7)$$

Ortak vektör herhangi bir \vec{a}_i öznelik vektörü ve bu vektörün \mathbf{B} fark alt uzayına izdüşümünün farkı olarak tanımlanır.

$$\tilde{a}_{com} = \vec{a}_i - \bar{a}_i \quad i = 1, 2, \dots, m \quad (2.8)$$

Aşağıdaki eniyileme ölçütünü ele alalım.

$$F_2 = \|\vec{a}_{1,dif} - \bar{a}_1\|^2 + \|\vec{a}_{2,dif} - \bar{a}_2\|^2 + \dots + \|\vec{a}_{m,dif} - \bar{a}_m\|^2 \quad (2.9)$$

Burada $\vec{a}_{i,dif}$ ($i=1,2,\dots,m$)'ler (2.1)'deki fark vektörleridir. F_2 ölçütünün enküçük yapılabilmesi için $\vec{a}_{i,dif}$ fark vektörlerinin \bar{a}_m 'lere yakın olmaları gerekir. F_2 'nin (2.1) 'e bağlı olarak enküçük olabilmesi için ortak vektör

$$x_{opt} = \tilde{a}_{com} \quad (2.10)$$

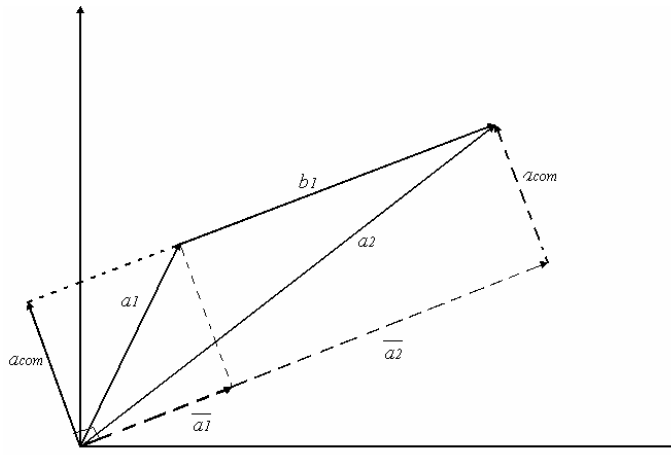
olarak seçilir. (2.1) ve (2.9) kullanılarak

$$F_2 = \left(\|\vec{a}_1 - \vec{x} - \bar{a}_1\|^2 + \dots + \|\vec{a}_m - \vec{x} - \bar{a}_m\|^2 \right)_{x=\tilde{a}_{com}} \quad (2.11)$$

$$= \left\| \vec{a}_1 - \tilde{a}_{com} - \bar{a}_1 \right\|^2 + \dots + \left\| \vec{a}_m - \tilde{a}_{com} - \bar{a}_m \right\|^2 \quad (2.12)$$

yazılabilir. Buradan (2.8) 'yi kullanarak $F_2 = 0 + 0 + \dots + 0 = 0$ elde edilir.

$\tilde{a}_{com} = \tilde{a}_i = \vec{a}_i - \bar{a}_i$ ($i = 1, 2, \dots, m$) vektörü i 'nin seçiminden bağımsızdır. Diğer bir deyişle F_2 'den elde edilen ortak vektör bir sınıf için tektir ve seçilen referans vektöründen bağımsızdır (Gülmezoğlu, et al., 1999).



Şekil 2.1 Aynı sınıfa ait iki boyutlu iki vektör ve bu vektörlerin ortak vektörlerinin gösterilişi

Örnek: $\left\{ \vec{a}_1 = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}, \vec{a}_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \vec{a}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \right\}$ aynı sınıfa ait vektörlerin oluşturduğu bir küme

olsun.

$$\vec{b}_1 = \vec{a}_2 - \vec{a}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} \quad \vec{b}_2 = \vec{a}_3 - \vec{a}_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix}$$

$$\vec{d}_1 = \vec{b}_1$$

$$\vec{z}_1 = \frac{\vec{d}_1}{\|\vec{d}_1\|} = \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix}$$

$$\vec{d}_2 = \vec{b}_2 - \langle \vec{b}_2, \vec{z}_1 \rangle \vec{z}_1 = \begin{bmatrix} -1 \\ -2 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

$$\vec{z}_2 = \frac{\vec{d}_2}{\|\vec{d}_2\|} = \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}$$

$$\vec{a}_{com} = \vec{a}_1 - \langle \vec{a}_1, \vec{z}_1 \rangle \vec{z}_1 - \langle \vec{a}_1, \vec{z}_2 \rangle \vec{z}_2 \rightarrow$$

$$\vec{a}_{com} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} - (-2) \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \left(-\frac{1}{\sqrt{2}} \right) \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \end{bmatrix}$$

veya

$$\vec{a}_{com} = \vec{a}_2 - \langle \vec{a}_2, \vec{z}_1 \rangle \vec{z}_1 - \langle \vec{a}_2, \vec{z}_2 \rangle \vec{z}_2 \rightarrow$$

$$\vec{a}_{com} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} - (-1) \begin{bmatrix} 0 \\ -1 \\ 0 \end{bmatrix} - \left(-\frac{1}{\sqrt{2}} \right) \begin{bmatrix} -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/2 \\ 0 \\ 1/2 \end{bmatrix}$$

Yetersiz veri durumunda bir sınıfa ait ortak vektör, ortak değışinti matrisinin sıfır özdeğerlerine karşılık gelen özvektörlerin doğrusal bileşimleri doğrultusundadır (Gülmezoğlu, et al., 2001). Bir sınıfa ait ortak değışinti matrisi

$$\Phi_{n \times n} = \sum_{i=1}^m (\vec{a}_i - \vec{a}_{ave}) (\vec{a}_i - \vec{a}_{ave})^T \quad (2.13)$$

şeklinde tanımlanır. Ortak değışinti matrisi vektörlerin, ortalama vektörüne göre değışintilerini karakterize etmektedir. $m \leq n$ durumunda Φ ortak değışinti matrisinin sıfırdan farklı özdeğerlerine karşılık gelen özvektörleri \mathbf{B} fark altuzayı için ortanormal taban oluşturur. Bu durumda fark altuzayının ortagonal tümleyeni olan \mathbf{B}^\perp , sıfıra eşit olan özdeğerlere karşılık gelen özvektörler tarafından gerilir. Ortak vektör fark uzayındaki her vektöre dik olduğundan, \mathbf{B}^\perp farksızlık uzayında yer almalıdır. O halde ortak vektör, Φ 'nin sıfıra eşit özdeğerlere karşılık gelen özvektörlerin doğrusal bir bileşimidir. Bu durumda ortak vektör bir vektörün sıfırdan farklı özdeğerlere karşılık gelen özvektörlerin

üzerine izdüşümlerinin vektörün kendisinden çıkarılmasıyla da elde edilebilir (Gülmezoğlu, et al., 2001).

Örnek 2: Vektör sınıfı olarak Örnek 1'deki sınıfı ele alalım.

Sınıfın ortalama vektörü $\vec{a}_{ave} = \frac{1}{3}(\vec{a}_1 + \vec{a}_2 + \vec{a}_3) = \begin{bmatrix} 2/3 \\ 1 \\ 1/3 \end{bmatrix}$ olur. Ortak deęişinti matrisi ise

$$\begin{aligned} \Phi &= \sum_{i=1}^3 (\vec{a}_i - \vec{a}_{ave})(\vec{a}_i - \vec{a}_{ave})^T \\ &= \begin{pmatrix} 1/3 \\ 1 \\ -1/3 \end{pmatrix} \begin{pmatrix} 1/3 & 1 & -1/3 \end{pmatrix} + \begin{pmatrix} 1/3 \\ 0 \\ -1/3 \end{pmatrix} \begin{pmatrix} 1/3 & 1 & -1/3 \end{pmatrix} + \begin{pmatrix} -2/3 \\ -1 \\ 2/3 \end{pmatrix} \begin{pmatrix} -2/3 & -1 & 2/3 \end{pmatrix} \\ &= \begin{bmatrix} 2/3 & 1 & -2/3 \\ 1 & 2 & -1 \\ -2/3 & -1 & 2/3 \end{bmatrix} \end{aligned}$$

elde edilir. Buradan Φ 'nin özdeęerleri

$$\lambda_1 = 3.1196$$

$$\lambda_2 = 0.2137$$

$$\lambda_3 = 0$$

bulunur. Bu özdeęerlere karşılık gelen özvektörler ise sırasıyla

$$\vec{u}_1 = [-0.4389 \quad -0.7840 \quad 0.4389]^T$$

$$\vec{u}_2 = [0.5544 \quad -0.6207 \quad -0.5544]^T$$

$$\vec{u}_3 = [-0.7071 \quad 0 \quad -0.7071]^T$$

elde edilir. \vec{a}_1 vektörünün \vec{u}_1 ve \vec{u}_2 üzerine izdüşümleri

$$\langle \vec{a}_1, \vec{u}_1 \rangle \vec{u}_1 = [0.8809 \quad 1.5735 \quad -0.8809]^T$$

$$\langle \vec{a}_1, \vec{u}_2 \rangle \vec{u}_2 = [-0.3809 \quad 0.4265 \quad 0.3809]^T$$

olur. Bunların toplamı ise

$$\langle \vec{a}_1, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{a}_1, \vec{u}_2 \rangle \vec{u}_2 = [0.5 \quad 2 \quad -0.5]^T$$

dır. Bu durumda ortak vektör

$$\vec{a}_{com} = \vec{a}_1 - \langle \vec{a}_1, \vec{u}_1 \rangle \vec{u}_1 + \langle \vec{a}_1, \vec{u}_2 \rangle \vec{u}_2 = [0.5 \quad 0 \quad 0.5]^T$$

elde edilir. Bu da Örnek 1'de elde edilen ortak vektörle aynıdır.

$$\begin{aligned} \vec{a}_{com} &= \langle \vec{a}_1, \vec{u}_3 \rangle \vec{u}_3 = \langle \vec{a}_2, \vec{u}_3 \rangle \vec{u}_3 = \langle \vec{a}_3, \vec{u}_3 \rangle \vec{u}_3 \\ &= [0.5 \quad 0 \quad 0.5]^T \end{aligned}$$

\mathbf{B} fark altuzayının boyutu $(m-1)$ olduğundan \mathbf{B}^\perp nin boyutuda $(n-m+1)$ olur. Bu durumda $u_1 > u_2 > \dots > u_n$ olmak üzere \mathbf{B} ye izdüşüm matrisi

$$P_{cov} = \sum_{i=1}^{m-1} \vec{u}_i \vec{u}_i^T \quad (2.14)$$

ve \mathbf{B}^\perp nin izdüşüm matrisi

$$P_{cov}^\perp = \sum_{i=m}^n \vec{u}_i \vec{u}_i^T \quad (2.15)$$

olur. Burada \vec{u}_i ler ortak değişinti matrisinin özvektörleridir.

2.1.2 Yeterli Veri Durumu ($n < m$)

Bir sınıfa ait vektörlerin sayısı vektörlerin boyutundan büyük ise yeterli veri durumu ortaya çıkar. Yeterli veri durumunda fark uzayını geren özvektörler bütün uzayı gerer ve ortak vektör $\vec{0}$ olur. Bu durumda da bir \vec{a}_{com} tahmini yapılabilir. Bunun için ortak değişinti matrisinin n tane olan özdeğerini küçükten büyüğe doğru sıralayalım. Bu özdeğerlerden en küçük m tanesine karşılık gelen özvektörler \mathbf{B}^\perp için bir altuzay tanımlasınlar. \mathbf{B}^\perp ve \mathbf{B} birbirlerine dik olduklarından geriye kalan özdeğerlere karşılık gelen özvektörlerde fark uzayı için bir taban oluştururlar. Sınıfa ait herhangi bir vektörün \mathbf{B} altuzayına izdüşümü fark vektörü

$$\vec{a}_{i,dif} = \sum_{i=k+1}^n (a_i^T u_i) u_i \quad (2.16)$$

verir. Buradan sınıf vektörlerini

$$\vec{a}_i = \vec{a}_{i,dif} + \vec{a}_{com} + \varepsilon_i \quad i = 1, 2, \dots, m \quad (2.17)$$

şeklinde yazabiliriz. Fark vektörleri ortak değişinti matrisinin en büyük $n-k$ tane özdeğerine karşılık gelen özvektörleri ile elde edildiğine göre geriye kalan k tane özdeğere karşılık gelen özvektörlerle de ortak vektör elde edilebilir. Bu durumda ortak vektör

$$\vec{a}_{com} = \sum_{i=1}^k \left(\vec{a}_{com}^T u_i \right) u_i \quad (2.18)$$

olarak yazılabilir. Eğer hataların karelerinin en küçük olması istenirse

$$F_3 = \frac{1}{2} \sum_{j=1}^m \varepsilon_j^T \varepsilon_j = \frac{1}{2} \sum_{j=1}^m (\vec{a}_i - \vec{a}_{i,dif} - \vec{a}_{com})^T (\vec{a}_i - \vec{a}_{i,dif} - \vec{a}_{com}) \quad (2.19)$$

şeklinde bir eniyileme ölçütü tanımlanabilir. F_3 'ün \vec{a}_{com} ye göre türevi alınırsa, F_3 ortak vektöre göre en küçük yapılmış olur. \vec{a}_{com} fark vektörünün u_i ($i=1,2,\dots,k$) yönünde bileşeni olmadığı için eniyileme işlemi

$$\vec{a}_{com}^T u_i = \vec{a}_{com,i} = \vec{a}_{ave}^T u_i \quad i = 1, 2, \dots, k \quad (2.20)$$

sonucunu verir. Burada \vec{a}_{ave} sınıfın ortalama vektörünü gösterir. O halde ortak vektör

$$\vec{a}_{com} = \sum_{i=1}^k \left(\vec{a}_{ave}^T u_i \right) u_i \quad (2.21)$$

şeklinde veya

$$P_{cov}^\perp = \sum_{i=1}^k \vec{u}_i \vec{u}_i^T \quad (2.22)$$

olmak üzere

$$\vec{a}_{com} = P_{cov}^{\perp} \vec{a}_{ave} \quad (2.23)$$

şeklinde hesaplanır.

Örnek 3: $\left\{ \vec{a}_1 = \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \vec{a}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \vec{a}_3 = \begin{bmatrix} 3 \\ 2 \end{bmatrix} \right\}$ aynı sınıfa ait vektör kümesi olsun.

$$\vec{a}_{ave} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad \Phi = \begin{bmatrix} 6 & 0 \\ 0 & 2 \end{bmatrix}$$

olur. Buradan özdeğer ve özvektörler,

$$\begin{aligned} \lambda_1 &= 6 & \vec{u}_1 &= \begin{bmatrix} 1 \\ 0 \end{bmatrix} \\ \lambda_2 &= 2 & \vec{u}_2 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned}$$

elde edilir. \mathbf{B}^{\perp} yi gelecek özvektör olarak $\vec{u}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ seçeceğiz. Çünkü bu vektör küçük

özdeğere karşılık geliyor.

$$\begin{aligned} \vec{a}_1 &= \vec{a}_{1,dif} + \vec{a}_{com} + \vec{\varepsilon}_1 \\ \vec{a}_2 &= \vec{a}_{2,dif} + \vec{a}_{com} + \vec{\varepsilon}_2 \\ \vec{a}_3 &= \vec{a}_{3,dif} + \vec{a}_{com} + \vec{\varepsilon}_3 \end{aligned}$$

$$\begin{aligned} \vec{a}_{com} &= \langle \vec{a}_{ave}, \vec{u}_2 \rangle \vec{u}_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \\ \vec{a}_{1,dif} &= \langle \vec{a}_1, \vec{u}_1 \rangle \vec{u}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \vec{a}_{2,dif} &= \langle \vec{a}_2, \vec{u}_1 \rangle \vec{u}_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \\ \vec{a}_{3,dif} &= \langle \vec{a}_3, \vec{u}_1 \rangle \vec{u}_1 = \begin{bmatrix} 3 \\ 0 \end{bmatrix} \end{aligned} \Rightarrow \begin{aligned} \vec{\varepsilon}_1 &= \vec{a}_1 - \vec{a}_{1,dif} - \vec{a}_{com} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \vec{\varepsilon}_2 &= \vec{a}_2 - \vec{a}_{2,dif} - \vec{a}_{com} = \begin{bmatrix} 0 \\ -1 \end{bmatrix} \\ \vec{\varepsilon}_3 &= \vec{a}_3 - \vec{a}_{3,dif} - \vec{a}_{com} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{aligned}$$

Karar Kuralı: Yetersiz veri durumu ($n \geq m$) için bir sınıfa ait herhangi bir vektörün fark uzayına izdüşümünün vektörün kendisinden çıkarılması ile elde edilen vektör daima ortak vektörü verir. Yeterli veri durumu ($n < m$) için ise bir sınıfa ait vektörün farksızlık uzayına izdüşümü, o sınıfın ortak vektöre yakın olacaktır. Yeterli veri durumu için ortak vektör, sınıf ortalamasının farksızlık altuzayına izdüşümünden elde edilmektedir.

Herhangi bir \vec{x} vektörü için \vec{x}_{rem} kalan vektörü

$$\vec{x}_{rem} = \vec{x} - \sum_{j=1}^{m-1} (\vec{x}^T \vec{u}_j) \vec{u}_j \quad (2.24)$$

formülüyle bulunur. Burada \vec{u}_j ($j=1,2,\dots,m-1$), fark uzayının ortonormal taban vektörleridir. \vec{x}_{rem} , ayrıca \vec{x} 'nin farksızlık uzayına izdüşümüyle de hesaplanabilir.

Herhangi bilinmeyen \vec{x} vektörünün $C^{(i)}$ ($i=1,2,\dots,K$) sınıfına sınıflanmasında ortak vektör ile kalan vektör \vec{x}_{rem} arasındaki Öklid mesafesine bakılır. Ortak vektör yaklaşımında

$$C^* = \arg \min_i \left[\left\| \vec{x}_{rem}^{(i)} - \vec{a}_{com}^{(i)} \right\|^2 \right] \quad i = 1, 2, \dots, K \quad (2.25)$$

karar kuralı kullanılır. Buradaki $\vec{x}_{rem}^{(i)}$, \vec{x} vektörünün i inci sınıftan elde edilen kalan vektörü, $\vec{a}_{com}^{(i)}$ ise bu sınıfa ait ortak vektördür.

2.2 Ana Bileşen Analizi (PCA)

Ana bileşen analizi, sinyal işleme uygulamalarında kullanılan (Zhang, et al, 2001; Duin, et al, 2000), çok boyutlu verileri, az boyutlu verilere enaz bilgi kaybıyla indirgeyen bir dönüşüm tekniğidir (Kramer, 1991). Bu yöntem aynı zamanda Karhunen-Loeve dönüşümü de denir (Bishop, 1996). PCA aynı zamanda sınıflandırma içinde kullanılmaktadır (Oja, 1983; Günel, et al., 2005).

Bu dönüşüm yapılırken ortak değışinti matrisinin özvektörleri özdeğerlerine göre büyükten küçüğe doğru sıralanır. İlk ana bileşen özdeğeri en büyük olan özvektör

yönündedir ve değişiminin en büyük olduğu doğrultuyu gösterir. İkinci ana bileşen özdeğeri en büyük ikinci özvektördür ve bir sonraki en büyük değişiminin doğrultusunu gösterir ve ilk ana bileşenden bağımsızdır. PCA ile sınıflandırma yapmak genelde iyi sonuçlar vermeyebilir. Çünkü PCA ile boyut indirilmesi yapılırken sınıflandırmada önemli olabilecek bazı bilgileri eleyebilir (Gülmezoğlu, et al., 2001).

PCA'de amaç d -boyutlu uzaydaki $\vec{x}_i = (x_1, x_2, \dots, x_d)$ ($i = 1, 2, \dots, d$) vektör kümesini $m < d$ olmak üzere m -boyutlu bir uzaydaki $\vec{z}_i = (z_1, z_2, \dots, z_m)$ vektörlerine dönüştürmektir.

\vec{x} vektörü hiç bir veri kaybına uğramadan

$$\vec{x} = \sum_{i=1}^d z_i \vec{u}_i \quad (2.26)$$

şeklinde gösterilebilir. Burada \vec{u}_i ler ortanormal vektörlerdir.

$$z_i = \vec{u}_i^T \vec{x} \quad \text{ve} \quad \vec{u}_i^T \vec{u}_j = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \quad (2.27)$$

\vec{x} vektörünü

$$\vec{x} = \sum_{i=1}^m z_i \vec{u}_i + \sum_{i=m+1}^d z_i \vec{u}_i \quad (2.28)$$

şeklinde yazabiliriz. Boyut indirildikten sonra

$$\vec{x} = \tilde{\vec{x}} = \sum_{i=1}^m z_i \vec{u}_i + \sum_{i=m+1}^d b_i \vec{u}_i \quad (2.29)$$

elde edilir. Burada b_i 'ler sabit sayılardır ve her vektör için aynıdır. Bu bir boyut indirilmesidir. Çünkü orjinal \vec{x} vektörü d boyutlu ve her boyutunda değişkenlik vardır.

Ama \vec{x} vektöründen elde edilen yeni \vec{z} vektörünün b_i ler sabit olduğu için $m(m < d)$ boyutunda değişkenlik vardır. Veri kümesindeki N tane vektörü \vec{x}^n , $n = 1, 2, \dots, N$ olmak üzere ele alalım. Amacımız \vec{x} vektörünü en iyi temsil edecek şekilde (2.29) daki \vec{u}_i ve b_i 'leri seçmektir.

\vec{x}^n vektöründeki boyut indirgemesinden oluşan hata

$$\vec{x}^n - \tilde{x}^n = \sum_{i=m+1}^d (z_i - b_i) \vec{u}_i \quad (2.30)$$

ile gösterilir. Bu hatayı en küçük yapmak bilgi kaybını da en küçük yapmak anlamına gelir. (2.27) daki ortonormallik bağıntısını kullanılırsa, en küçük yapılması gereken hata fonksiyonu

$$E_m = \frac{1}{2} \sum_{n=1}^N \left\| \vec{x}^n - \tilde{x}^n \right\|^2 = \frac{1}{2} \sum_{n=1}^N \sum_{i=m+1}^d (z_i^n - b_i)^2 \quad (2.31)$$

şeklini alır. E_m 'nin b_i 'ye göre türevi sıfıra eşitlenirse;

$$b_i = \frac{1}{N} \sum_{i=1}^N z_i^n = \vec{u}_i^T \bar{X} \quad (2.32)$$

elde edilir. Burada $\bar{x} = \frac{1}{N} \sum_{n=1}^N \vec{x}^n$ sınıfın ortalama vektörünü göstermektedir. b_i ,

(2.31)'de yerine koyulursa

$$E_m = \frac{1}{2} \sum_{m+1}^d \vec{u}_i^T \Phi \vec{u}_i \quad (2.33)$$

elde edilir. Burada Φ , $\left\{ \vec{x}^n \right\}$ vektör sınıfının değişinti matrisidir ve

$$\Phi = \sum_{n=1}^N \left(\vec{x}^n - \bar{x} \right) \left(\vec{x}^n - \bar{x} \right)^T \quad (2.34)$$

şeklinde tanımlanır. Bundan sonra geriye uygun \vec{u}_i vektörlerini seçerek E_m 'yi enküçük yapılması kalır. E_m , taban vektörleri

$$\Phi \vec{u}_j = \lambda_j \vec{u}_j \text{ ve } \lambda_j \geq 0 \quad (2.35)$$

eşitliğini sağladığında enküçük yapılıır.

(2.35)'i (2.33)'de yerine koyulursa ve (2.27)'deki ortanormallik bağıntısı kullanılırsa hata fonksiyonunun enküçükteki değeri

$$E_M = \frac{1}{2} \sum_{i=m+1}^d \lambda_i \quad (2.36)$$

olarak elde edilir. Enküçük hata $d-m$ tane en küçük özdeğeri ve onlara karşılık gelen özvektörleri eleyerek elde edilir.

2.3 Doğrusal Ayırtaç Analizi(LDA)

Doğrusal ayırtaç analizi, örüntü tanımada kullanılan bir sınıflandırma yöntemidir (Swets and Weng, 1996). Doğrusal ayırtaç analizindeki amaç sınıflar arası dağılımı enbüyük yaparak farklı grupların örneklerini birbirinden ayırırken, sınıf içi değişintiyi en küçük yapmaktır. Böylece boyut indirgemesi yaparken aynı zamanda da sınıflandırmada kullanılacak bilgiler korunacaktır. LDA'de sınıflandırma ölçütünü oluşturmak için sınıf içi ve sınıflar arası saçılım kullanılır. Sınıf içi saçılım, sınıfların herbirinin beklenen ortak değişintisidir. Sınıf içi saçılım matrisi

$$S_w = \sum_j p_j \text{cov}_j \quad (2.37)$$

ile hesaplanır. Burada j sınıf indeksini, cov_j sınıf ortak deęişinti matrisini, ve p_j ise sınıfın önsel olasılıęını göstermektedir.

Sınıflar arası saçılım ise

$$S_b = \sum_j (\mu_j - \mu_g)(\mu_j - \mu_g)^T \quad (2.38)$$

formülüyle hesaplanır. Burada μ_j , j inci sınıfın ortalamasını; ve μ_g , bütün verilerin ortalamasını

$$\mu_g = \sum_j p_j \mu_j \quad (2.39)$$

göstermektedir. S_b üyeleri sınıfların ortalama vektörleri olan veri kümesinin ortak deęişinti matrisi olarak düşünülebilir.

Eniyileme ölçütü J_r ,

$$J_r = (S_w)^{-1} S_b \quad (2.40)$$

formülüyle hesaplanır. J_r nin sıfırdan farklı özdeęerlerine karşılık gelen özvektörler dönüşüm matrisimizi oluşturur. Çünkü sıfır özdeęere karşılık gelen özvektörler farksızlık atluzayını gererler ve bu doğrusal bağımlılıęı gösterir. Bu özvektörlerin sınıflandırmada bir katkısı olmaz. Herhangi bir L -sınıf problemi için daima $L-1$ tane sıfırdan farklı özdeęer vardır. W dönüşüm matrisi olmak üzere dönüşüm

$$\vec{y} = W\vec{x} \quad (2.41)$$

ile elde edilir. LDA'da aşıęıdaki karar kuralı kullanılarak sınıflama yapılır:

$$C^* = \arg \min_i \left\{ \left\| W \left(\vec{x} - \vec{x}_{ort}^{(i)} \right) \right\| \right\} \quad i = 1, 2, \dots, L \quad (2.42)$$

Burada \vec{x} bilinmeyen vektör, $\vec{x}_{ort}^{(i)}$ i inci sınıfın ortalaması, L ise sınıf sayısıdır. Uzaklık ölçütü olarak ise Öklid kullanılır (Balakrishnama, et al., 1999).

Örnek: $C_1 = \left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \end{bmatrix}, \begin{bmatrix} 2 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix} \right\}$, $C_2 = \left\{ \begin{bmatrix} -1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 1 \\ -3 \end{bmatrix} \right\}$ iki sınıf olsun.

Sınıfların ortalamaları: $\mu_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$ ve $\mu_2 = \begin{bmatrix} 0 \\ -2 \end{bmatrix}$

C_1 'in ortak değişinti matrisi: $S_1 = \sum_{i=1}^4 (\vec{x}_i - \vec{\mu}_1)(\vec{x}_i - \vec{\mu}_1)^T = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$, $\vec{x}_i \in C_1$

C_2 'nin ortak değişinti matrisi: $S_2 = \sum_{i=1}^4 (\vec{x}_i - \vec{\mu}_2)(\vec{x}_i - \vec{\mu}_2)^T = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$, $\vec{x}_i \in C_2$

$$S_w = 0.5S_1 + 0.5S_2 = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \quad \Rightarrow \quad S_w^{-1} = \frac{1}{16} \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$$

$$S_b = (\mu_1 - \mu_g)(\mu_1 - \mu_g)^T + (\mu_2 - \mu_g)(\mu_2 - \mu_g)^T \\ = \begin{bmatrix} 1/2 & 2 \\ 2 & 8 \end{bmatrix}$$

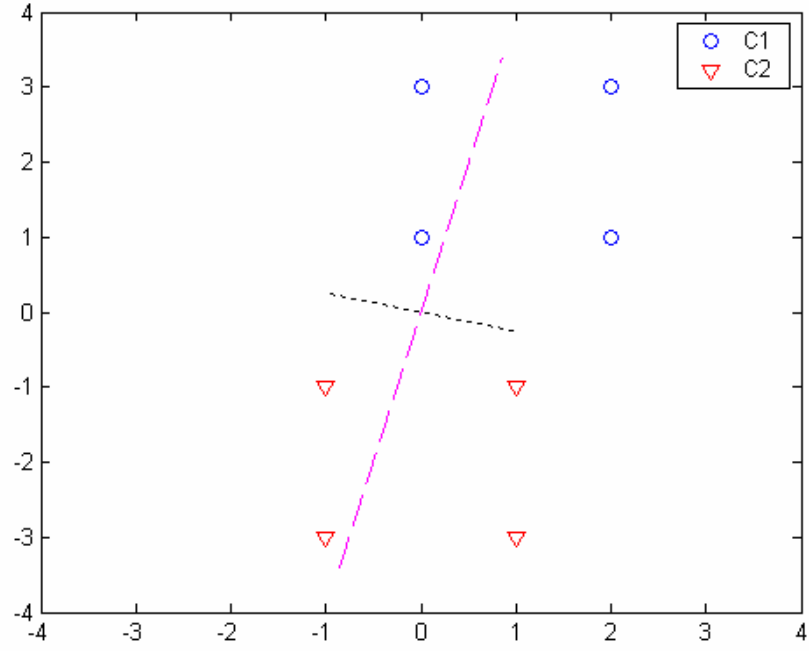
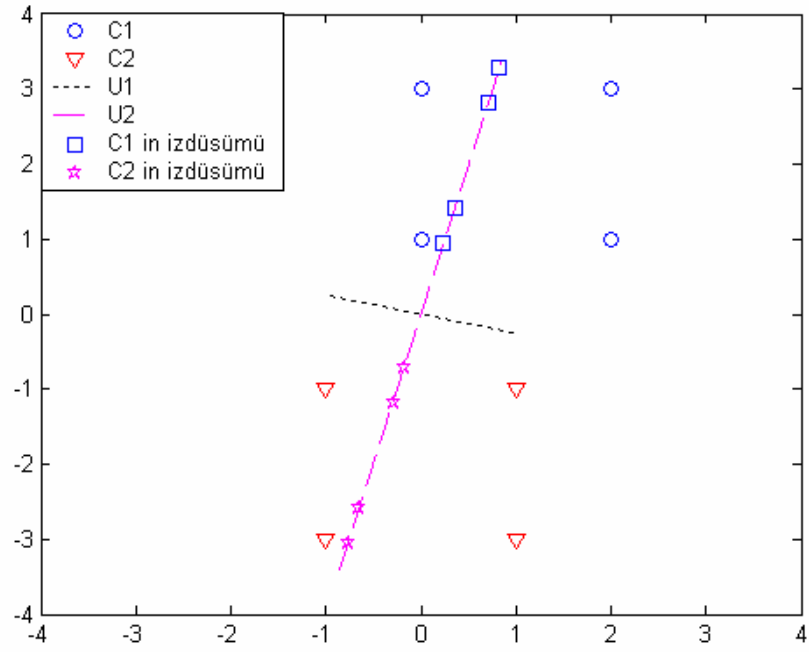
$$C_r = S_w^{-1}S_b = \frac{1}{16} \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \begin{bmatrix} 1/2 & 2 \\ 2 & 8 \end{bmatrix} = \begin{bmatrix} 1/8 & 1/2 \\ 1/2 & 2 \end{bmatrix}$$

$$\lambda_1 = 0 \quad \Rightarrow \quad u_1 = \begin{bmatrix} 0.9701 \\ -0.2425 \end{bmatrix}, \quad u_2 = \begin{bmatrix} 0.2425 \\ 0.9701 \end{bmatrix}$$

λ_2 değerine karşılık gelen özdeğeri alırsak $w = [0.2425 \quad 0.9701]$ olur. Buradan

C_1	C_2
0.9701	-1.2127
2.9104	-3.1530
1.4552	-0.7276
3.3955	-2.6679

olarak elde edilir.

Şekil 2.2 C_1 ve C_2 sınıfları için ana bileşenlerŞekil 2.3 C_1 ve C_2 sınıfların ait vektörlerin ana bileşen üzerine izdüşümü

3. SINIFLANDIRICI BİRLEŞTİRME

Çoklu sınıf problemi ve gürültülü girişlerin olduğu durumlarda sınıflandırmada istenilen başarımlar elde edilememektedir. Ama çoklu sınıflandırıcı sistemleri bu problemler için iyi çözümler üretebilir. Bu sistemlerde farklı sınıflandırıcıların üstün özellikleri birleştirilerek daha iyi sınıflandırma yapılabilir.

Birleştirilecek sınıflandırıcıların seçilmesinde dikkat edilmesi gereken önemli konulardan biri, birleştirilen sınıflandırıcıların birbirlerini tamamlamasıdır. Sınıflandırıcılardan birisi hatasız sınıflandırma yapıyor ise sınıflandırıcı birleştirmeye gerek yoktur. Ama sınıflandırıcılar hata yapıyorsa bu hataların farklı kümelerde olmasına dikkat edilmelidir (Kuncheva, 2004).

Dietterich (2000) sınıflandırıcı birleştirmenin, tek sınıflandırıcıdan niçin daha iyi çalışabileceği hakkında üç ana sebep sunmuştur:

- İstatistiksel: Bir veri kümesi üzerinde iyi çalışan sınıflandırıcı kümemiz olsun. Sınıflandırıcılardan herhangi birini seçmek yerine bunların ortalamasını almak yanlış sınıflandırıcıyı seçme riskini azaltır.
- Hesapsal: Bazı eğitim algoritmaları (tepe-inme, rastgele arama) farklı yerel eniyi değere yakınsayabilirler. Fakat bunların farklı şekillerdeki toplamları küresel en iyiye daha iyi yakınsayabilir.
- Temsili: Sınıflandırıcıların tanımlı olduğu uzaydaki sınıflandırıcılardan hiçbiri en iyi sınıflandırıcı olmayabilir. Ama bu sınıflandırıcıların birleşimi ile en iyi sınıflandırıcı elde edilebilir.

3.1 Sınıflandırıcı Birleştirme Yöntemleri

Çoklu sınıflandırıcılar için çeşitli yöntemler geliştirilmiştir (Xu, et al., 1992). Bunlardan en iyi bilinen yöntem çoğunluk oyu (Ji and Ma, 1997) yöntemidir. Diğer

oylama yöntemleri ise enküçük, enbüyük, ortanca, ortalama (Taniguchi and Tresp, 1997) ve çarpım (Tax, et al.2000) yöntemleridir. Diğer yöntemler; ağırlıklı ortalama(Heskes, 1997), Borda sayısı, Bayes birleşimi, bulanık integral, bulanık bağlayıcılar, bulanık şablonlar, Dempster Shafer teorisi ve olasılıksal şablonlardır.

Ortalama: Bu yöntemde tekil sınıflandırıcıların çıkışlarının ortalaması alınır. Sınıflandırmada ise ortalamanın en yüksek olduğu sınıfa yapılır.

$$Q(x) = \arg \max_{j=1}^N \left(\frac{1}{K} \sum_{i=1}^K y_{ij}(x) \right) \quad (3.1)$$

Burada N sınıf sayısı, $y_{ij}(x)$ ise x girişi için j . sınıfın i . sınıflandırıcıdan çıkış değeridir.

Ağırlıklı Ortalama: Ağırlıklı ortalama, ortalama yöntemiyle benzerdir, farklı olarak sınıflandırıcıların çıkışları ağırlıklarla çarpılır. Yani

$$Q(x) = \arg \max_{j=1}^N \left(\frac{1}{K} \sum_{i=1}^K w_i y_{ij}(x) \right) \quad (3.2)$$

Çoğunluk Oyu: Burada doğru sınıf farklı sınıflandırıcılar tarafından en çok seçilen sınıftır. Eğer bütün sınıflandırıcılar farklı sınıfları seçer ise veya eşitlik olursa doğru sınıf olarak çıkış değeri en yüksek olan seçilir.

Enbüyük: Doğru sınıf olarak en büyük çıkış değeri veren seçilir.

$$Q(x) = \arg \max_{i=1}^K y_i(x) \quad (3.3)$$

Burada K sınıflandırıcı sayısı, $y_i(x)$ ise x vektörünün i . sınıflandırıcıdaki çıkış değeridir.

Nash: Bu yöntemde herbir sınıflandırıcı giriş vektörü için 0 ile 1 arasında değer atar. Bu giriş vektörü için bütün sınıflandırıcılardan elde edilen değerler çarpılır. Vektör çarpım değeri en yüksek olan sınıfa atanır.

$$Q(x) = \arg \max_{j=1}^N \prod_{i=1}^K y_{ij} \quad (3.4)$$

Borda Sayısı: Herhangi bir j sınıfı için, Borda sayısı, herbir sınıflandırıcı tarafından sıralamada j inci sınıfın ardına yerleştirilen sınıf sayısının toplamıdır. Eğer $B_i(j)$ i sınıflandırıcı tarafından j 'den sonraya yerleştirilen sınıf sayısı ise j sınıfı için Borda sayısı

$$B(j) = \sum_{i=1}^K B_i(j) \quad (3.5)$$

ile gösterilir. Örüntü en büyük Borda sayısına sahip sınıfa atanır.

Bayes Birleşimi: c^i , i . sınıfa ait hata matrisi olsun. Bu hata matrisinin elemanları c_{jk}^i 'ler, j . sınıfta oldukları halde k . sınıfta tespit edilen veri sayısı olsun. Bu durumda bir i sınıflandırıcısının j sınıfına ait bir x vektörünü k sınıfına atanmasının koşullu olasılığı

$$P(x \in q_j \mid \lambda_i(x) = j_i) = \frac{c_{jk}^i}{\sum_{j=1}^N c_{jk}^i} \quad (3.6)$$

Buradan, eğer sınıflandırıcıların birbirinden bağımsız olduğunu varsayarsak, x vektörünün j sınıfına ait güven değeri yaklaşık olarak

$$Belief(j) = \frac{\prod_{i=1}^K P(x \in q_j \mid \lambda_i(x) = j_i)}{\sum_{j=1}^N \prod_{i=1}^K P(x \in q_j \mid \lambda_i(x) = j_i)} \quad (3.7)$$

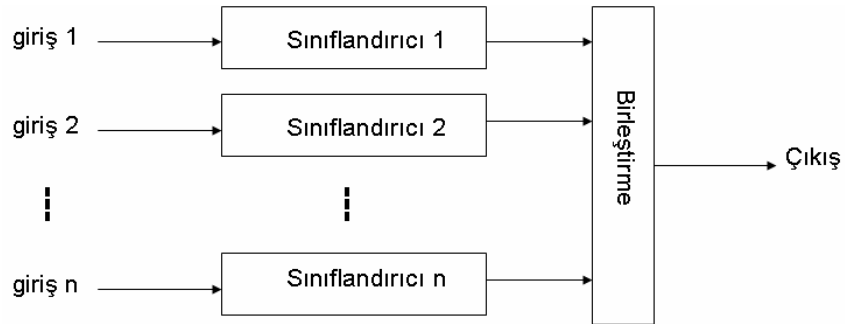
olarak hesaplanabilir. x vektörü güven değeri en büyük olan sınıfa atanır.

3.2 Sınıflandırıcı Birleştirme Mimarileri

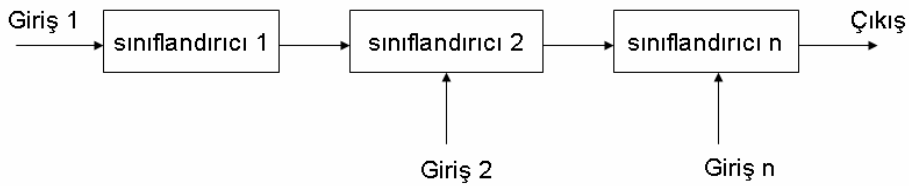
Sınıflandırıcı birleştirme mimarileri sınıflandırıcıların birleştirilmelerine göre üç ana grupta toplanabilir:

Paralel: Sınıflandırıcıların herbirinin çıkışları birbirinden bağımsız olarak alınır ve birleştirici tarafından birleştirilir.

Seri: Sınıflandırıcılar ardışıl olarak işleme sokulur. Verilen bir örüntü için olası sınıflar yeni bir sınıflandırıcı işleme alındıkça azalır.



Şekil 3.1 Paralel sınıflandırıcı birleştirme



Şekil 3.2 Seri sınıflandırıcı birleştirme

Hiyerarşik: Sınıflandırıcılar karar ağacı sınıflandırıcısına benzer bir yapıda birleştirilirler. Bu mimarinin avantajı farklı tipteki özniteliklerin ayırt edici özelliklerini başarılı bir şekilde kullanabilmesidir (Felföldi,2003).

3.3 Sınıf-İçi ve Sınıflar Arası Dağılımlardan Elde Edilen Ölçütlerin

Birleştirilmesi

Sınıf içi ve sınıflar arası dağılımları göz önüne alan sınıflandırmada, örüntünün kendi sınıfının ortalamasına yakın ve diğer sınıfların ortalamasından uzak yapacak ölçütler kullanılır. Sınıf içi ve sınıflar arası dağılımlardan elde edilen ölçütlerin birleştirilmesinde iki-sınıf problemi göz önüne alınacaktır. Bu bölümde; $C_1 = \{\vec{a}_1, \vec{a}_2, \dots, \vec{a}_{N_1}\}$ ve $C_2 = \{\vec{a}_{N_1+1}, \vec{a}_{N_1+2}, \dots, \vec{a}_{N_1+N_2}\}$, iki sınıfa ait öznitelikleri; N_1 ve N_2 sırasıyla C_1 ve C_2 sınıflarının öznitelik sayısını gösterecektir.

3.3.1 Enküçültme Problemi

Sınıf içi dağılıma göre belirlenen ölçütün enküçük olması, bir sınıfın elemanlarının o sınıfın ortalamasına yakın olması ile sağlanır. Bu durumda iki-sınıf problemi için aşağıdaki ölçütlerin enküçük olması gerekir.

$$F_{1_{c_1}} = \sum_{c_1} \left\| P_{1_{c_1}}^\perp (\vec{a}_i - \vec{a}_{ave_{c_1}}) \right\|^2 \quad (3.8)$$

$$F_{1_{c_2}} = \sum_{c_2} \left\| P_{1_{c_2}}^\perp (\vec{a}_i - \vec{a}_{ave_{c_2}}) \right\|^2 \quad (3.9)$$

Burada $P_{1_{c_1}}^\perp$ ve $P_{1_{c_2}}^\perp$ matrisleri C_1 ve C_2 sınıflarının farksızlık uzaylarına izdüşüm matrisleri ve $\vec{a}_{ave_{c_1}}$ ve $\vec{a}_{ave_{c_2}}$ vektörleri bu sınıfların ortalamalarıdır.

Bilindiği gibi OVY, ortak değişinti matrisinin sıfırdan farklı özdeğerlere karşılık gelen bütün özvektörlerin yönündeki öznitelikleri eler.

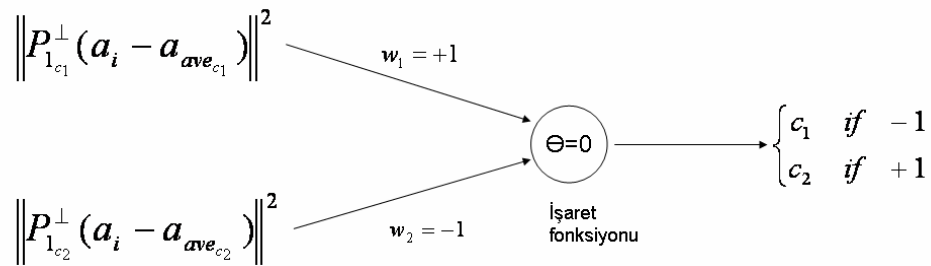
$$\mathfrak{R}^n = span \left\{ \underbrace{\vec{u}_1, \vec{u}_2, \dots, \vec{u}_{m_1}}_{\substack{\text{en küçük özdeğerde} \\ \text{karşıışgelen} \\ \text{özvektörler}}}, \underbrace{\vec{u}_{m_1+1}, \dots, \vec{u}_n}_{\substack{\text{en büyük özdeğerde} \\ \text{karşıışgelen} \\ \text{özvektörler}}} \right\} \quad (3.10)$$

$$P_{1_{c_1}}^\perp = \sum_{i=1}^{m_1} \vec{u}_i \vec{u}_i^T \quad (3.11)$$

$P_{1_{c_1}}^\perp$, ortak deęişinti matrisinin farksızlık uzayını geren, en küçük özdeęerlerine karşılık gelen özvektörlerinden oluşturulur. Buradan OVY karar ölçütü \vec{a}_x sınıflandırılacak vektör olmak üzere

$$c_* = \arg \min_k \left\| P_{1_{c_i}}^\perp (\vec{a}_x - \vec{a}_{ave_{c_i}}) \right\|^2 \quad (3.12)$$

şeklinde yazılabilir. İki-sınıf problemi için bu ölçüt tek algılayıcılı yapay sinir ağı şeklinde gösterilebilir.



Şekil 3.3 OVY in iki sınıf probleminde yapay sinir ağı formunda gösterilmesi

3.3.2 Enbüyütme Problemi

Sınıflar arası dağılımdan elde edilen ölçütün en büyük olması bir sınıfın elemanlarının dięer sınıfın ortalamasına uzak olması ile sağlanabilir. Bu durumda iki-sınıf problemi için aşağıdaki ölçütlerin enbüyük olması beklenir.

$$F_{2_{c_1}} = \sum_{c_1} \left\| P_{2_{c_1}} (\vec{a}_i - \vec{a}_{ave_{c_2}}) \right\|^2 \quad (3.13)$$

$$F_{2_{c_2}} = \sum_{c_1} \left\| P_{2_{c_2}} (\vec{a}_i - \vec{a}_{ave_{c_1}}) \right\|^2 \quad (3.14)$$

Burada $P_{2_{c_1}}$ ve $P_{2_{c_2}}$ matrisleri sınıflar arası dağılımların fark uzaylarının izdüşüm matrisleridir ve PCA'den elde edilmişlerdir. Tanıma için sıfırdan farklı ana bileşenler kullanılır.

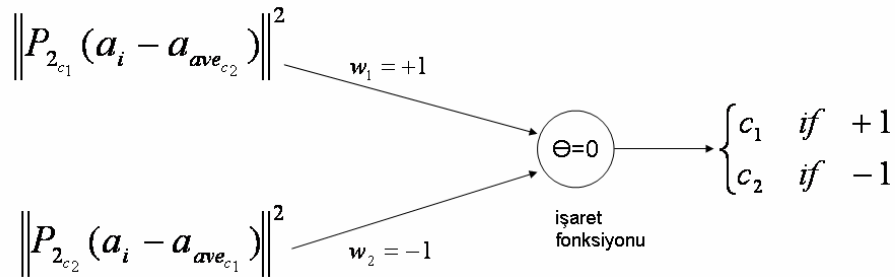
$$\mathcal{R}^n = \text{span} \left\{ \underbrace{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_{m_2}}_{\substack{\text{en küçük özdeğerlere} \\ \text{karşılaşılabilen} \\ \text{özvektörler}}}, \underbrace{\vec{v}_{m_2+1}, \dots, \vec{v}_n}_{\substack{\text{en büyük özdeğerlere} \\ \text{karşılaşılabilen} \\ \text{özvektörler}}} \right\} \quad (3.15)$$

$$P_{2_{c_1}} = \sum_{i=m_2+1}^n \vec{v}_i \vec{v}_i^T \quad (3.16)$$

PCA, ortak değışinti matrisin en küçük özdeğerlere (sıfır özdeğerler dahil) karşılık gelen özvektörler yönündeki öznitelikleri eler. PCA yönteminde karar ölçütü

$$c_* = \arg \max_k \left\| P_{2_{c_i}} (\vec{a}_x - \vec{a}_{ave_{c_j}}) \right\|^2 \quad i, j, k = 1, 2 \quad i \neq j \quad (3.17)$$

şeklinde yazılabilir. Burada \vec{a}_x sınıflandırılacak vektörü göstermektedir. Bu ölçüt de enküçültme ölçütü gibi yapay sinir ağı şeklinde gösterilebilir.

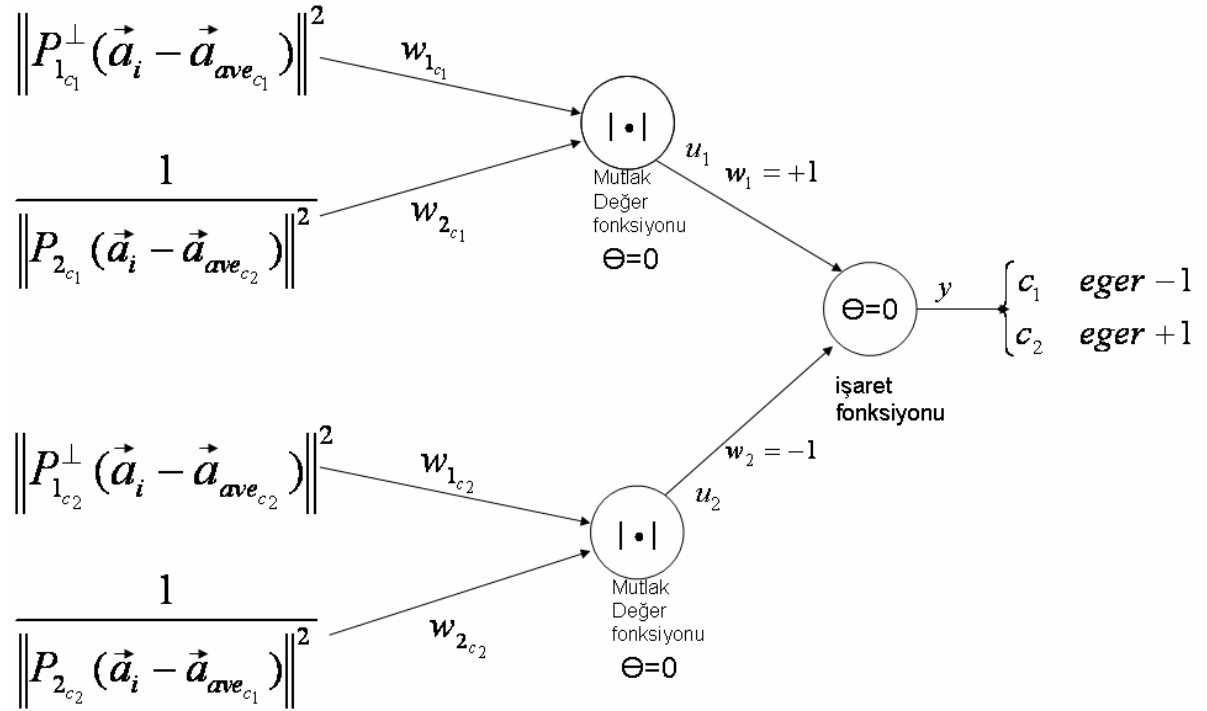


Şekil 3.4 PCA sınıflandırıcısının yapay sinir ağı şeklinde gösterilmesi

3.3.3 Sınıf içi ve Sınıflar arası Ölçütlerin Birleştirilmesi

Sınıflandırıcı birleştirmedeki amacımız enküçültme ve enbüyütme ölçütlerini yapay sinir ağı yardımıyla birleştirmektir. Bunun yapılabilmesi için bu iki ölçütün enküçültme veya enbüyütme şeklinde ifade edilmesi gerekmektedir. Enbüyütme

ölçütünde eğer her normun -1 inci kuvveti alınırsa, ölçüt enküçültme ölçütüne dönüşecektir.



Şekil 3.5 OVY ve PCA sınıflandırıcılarının yapay sinir ağı ile birleştirilmesi

Şekil 3.5'de birleştirilen ölçütün iki-sınıf problemi için sinir ağı şeklinde gösterimi verilmiştir. Buradaki \vec{a}_i vektörünün birinci ve ikinci sınıfın OVY ve PCA'den elde edilen altuzaylara izdüşümlerinden elde edilen normları, sinir ağının girişlerini oluşturmaktadır. Bu girişlerden CVA'den elde edilenler $w_{1_{c_1}}$ ve $w_{1_{c_2}}$ ağırlıklarıyla, PCA'den elde edilenler ise $w_{2_{c_1}}$ ve $w_{2_{c_2}}$ ile çarpılır. Birleştirilen ölçüt enküçültme ölçütü olduğundan sinir ağının ilk katmanından elde edilen u_1 ve u_2 çıkışları dan en küçük olana göre \vec{a}_i vektörünün sınıflandırılması yapılır. y çıkışı sıfırdan küçük ise \vec{a}_i birinci sınıfa, sıfırdan büyük ise ikinci sınıfa atanır.

Amacımız sinir ağının hata miktarını en küçük yapmaktır. Hataları en küçük yapmak için $w_{1_{c_1}} - w_{1_{c_2}}$ uzayında

$$\left| w_{1_{c_1}} \right| \leq 1, \left| w_{2_{c_1}} \right| \leq 1 \quad (3.18)$$

ve

$$w_{1_{c_1}}^2 + w_{2_{c_1}}^2 = 1, w_{1_{c_2}}^2 + w_{2_{c_2}}^2 = 1 \quad (3.19)$$

olmak üzere kare bir bölge ele alalım. Bu kısıtlamalar $w_{1_{c_1}} - w_{2_{c_1}} - w_{1_{c_2}} - w_{2_{c_2}}$ uzayında boyut indirilmesi sağlar. Kare bölgeyi $w_{1_{c_1}} - w_{1_{c_2}}$ düzleminin birinci bölgesinde alalım ve burada Şekil 3.6'daki gibi bir ızgara oluşturalım. $w_{2_{c_1}}$ ve $w_{2_{c_2}}$ ağırlıkları kısıt ilişkilerinden hesaplanabilir. $w_{2_{c_1}}$ ve $w_{2_{c_2}}$ elde edilirken (3.18) ve (3.19)'deki denklemlerden faydalanılır. Bu denklemlerden $w_{2_{c_1}}$ ve $w_{2_{c_2}}$ sıfırdan büyük ve sıfırdan küçük olmak üzere ikişer değeri çıkar. Şekil 3.7'de bir sınıfın OVY ve PCA'den elde edilen normlar ve bu normların çarpıldığı ağırlıklar arasındaki ilişki gösterilmektedir. Şekil 3.7'den

$$F_1 w_1 + \frac{1}{F_2} w_2 = 0 \quad (3.20)$$

elde edilir. (3.20)'dan w_1 ($w_{1_{c_1}}$ ve $w_{1_{c_2}}$) sıfırdan büyük seçildiğinden w_2 ($w_{1_{c_1}}$ ve $w_{1_{c_2}}$) sıfırdan küçük olmak zorundadır. Şekil 3.5'de u_1 ve u_2 çıkışlarının sıfırdan küçük olmaları metriğin çalışmasını olumsuz olarak etkileyeceğinden sinir ağının ilk katmanına mutlak değer fonksiyonu eklenmiştir.

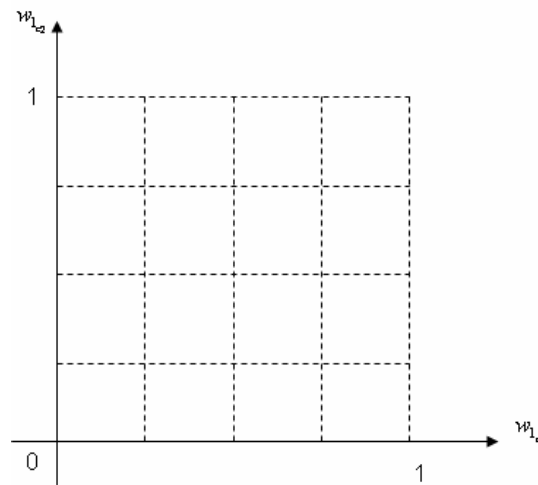
Buradan yapay sinir ağını çıkışındaki hata miktarı hesaplanabilir. Şekil 3.5'e bakıldığında C_1 sınıfının bir elemanı C_2 'ye sınıflandırılırsa yapay sinir ağı +1 çıkışı verecektir. Bu durumda C_1 sınıfının elemanlarının yanlış sınıflandırılmasından dolayı oluşacak toplam hata

$$\text{Hata}_{C_1} = \frac{1}{2} \sum_{c_1} \left[\text{sgn} \left\{ \left\| P_{1_{c_1}}^\perp (a_i - a_{ave_{c_1}}) \right\|^2 w_{1_{c_1}} + \frac{1}{\left\| P_{2_{c_1}} (a_i - a_{ave_{c_2}}) \right\|^2} w_{2_{c_1}} \right\} - \left\| P_{1_{c_2}}^\perp (a_i - a_{ave_{c_2}}) \right\|^2 w_{1_{c_2}} - \frac{1}{\left\| P_{2_{c_2}} (a_i - a_{ave_{c_1}}) \right\|^2} w_{2_{c_2}} \right\} + 1 \right] \quad (3.21)$$

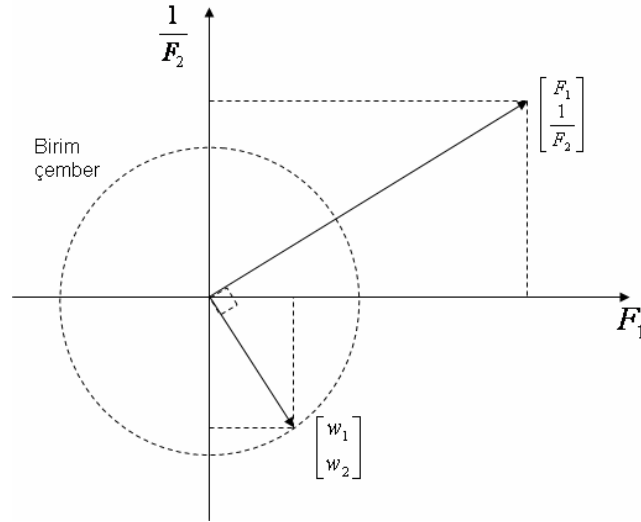
formülüyle hesaplanabilir. Benzer şekilde C_2 sınıfının bir elemanı C_1 sınıfında sınıflandırılırsa Şekil 3.5'deki sinir ağı -1 çıkışı verecektir. Böylece C_2 sınıfının elemanlarının yanlış sınıflandırılmasından dolayı oluşacak toplam hata

$$\text{Hata}_{C_2} = -\frac{1}{2} \sum_{c_2} \left[\text{sgn} \left\{ \left\| P_{1_{c_1}}^\perp (a_i - a_{ave_{c_1}}) \right\|^2 w_{1_{c_1}} + \frac{1}{\left\| P_{2_{c_1}} (a_i - a_{ave_{c_2}}) \right\|^2} w_{2_{c_1}} \right\} - \left\| P_{1_{c_2}}^\perp (a_i - a_{ave_{c_2}}) \right\|^2 w_{1_{c_2}} - \frac{1}{\left\| P_{2_{c_2}} (a_i - a_{ave_{c_1}}) \right\|^2} w_{2_{c_2}} \right\} - 1 \right] \quad (3.22)$$

formülüyle hesaplanabilir.



Şekil 3.6 $w_{1_{c_1}} - w_{1_{c_2}}$ düzleminde oluşturulan ızgara



Şekil 3.7 Normlar ve ağırlıklar arasındaki ilişki

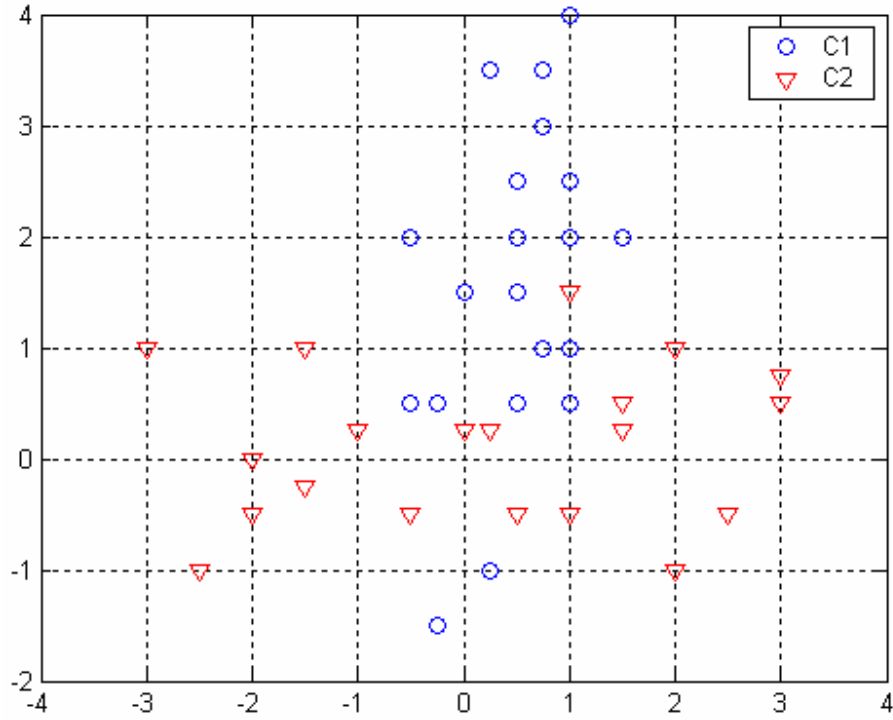
3.3.4 Sınıflandırıcı Birleştirmenin İki-Sınıf Problemine Uygulanması

Sınıflandırıcı birleştirmedeki başarımlar iki boyutlu yapay veriler üzerinde denenmiştir. Bu örnekte iki boyutlu iki sınıflı yapay veri kümesi üzerinde sınıflandırıcıların başarımları karşılaştırılacaktır. Kullanılan sınıflar C_1 ve C_2 olmak üzere her birinde 20 eleman bulunmaktadır ve Şekil 3.8’de C_1 ve C_2 gösterilmektedir.

Düzlem üzerinde x - ve y - eksenleri boyunca 0.1 aralıklarla alınan noktalar test kümesi olmak üzere OVY ile sınıflandırma sonuçları Şekil 3.9’de verilmektedir. Şekil 3.9’e bakıldığında iki sınıfı ayıran sınır çizgileri kolayca görülebilir. Benzer şekilde aynı test kümesini kullanarak PCA’den elde edilen ölçüt ve LDA ile oluşturulan grafiklerde sınır çizgileri sırasıyla Şekil 3.10 ve Şekil 3.11’de rahatlıkla görülebilmektedir. Şekil 3.12’de ise OVY ve PCA’den elde edilen ölçütlerin birleştirilmesi ile elde edilen yeni sınıflandırıcıdan elde edilen test grafiği verilmiştir. Burada Şekil 3.6’daki ızgara üzerinde gezilerek ve (3.18) bağıntıları da kullanılarak $w_{1_{c_1}} = 0.914$ $w_{1_{c_2}} = 0.999$

$w_{2_{c_2}} = -0.04471$ $w_{2_{c_1}} = -0.40571$ olarak hesaplanmıştır. Tablo 3.1’de OVY, PCA, LDA ve birleştirilen sınıflandırıcının sınıflandırma başarımları verilmiştir. Tablo 3.1’den

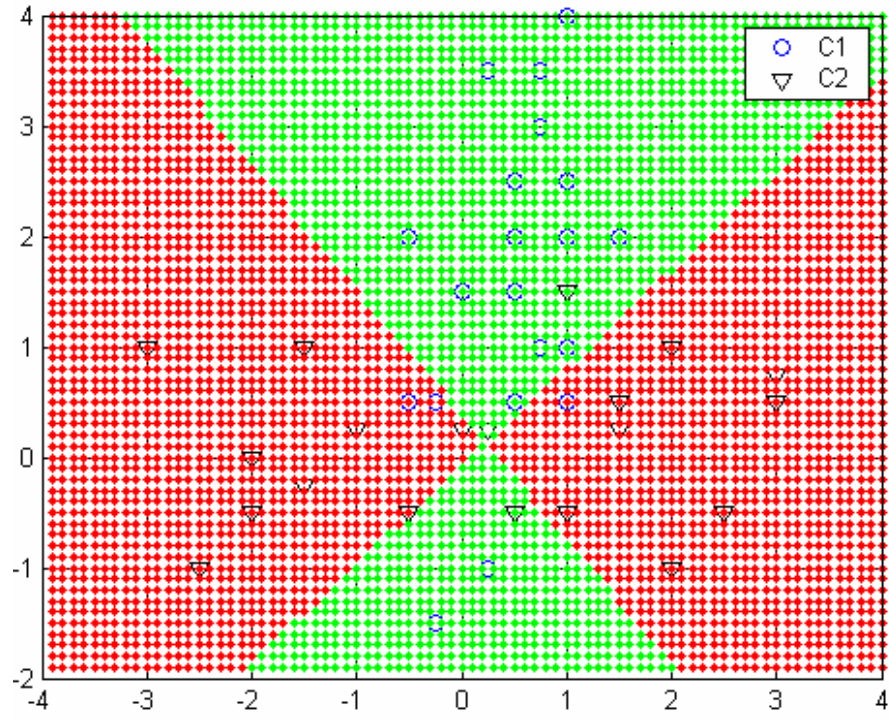
de görüldüğü gibi yapay veri üzerinde, birleştirilen sınıflandırıcı ile diğerler sınıflandırıcılara göre daha iyi sınıflandırma başarımı elde edilmiştir.



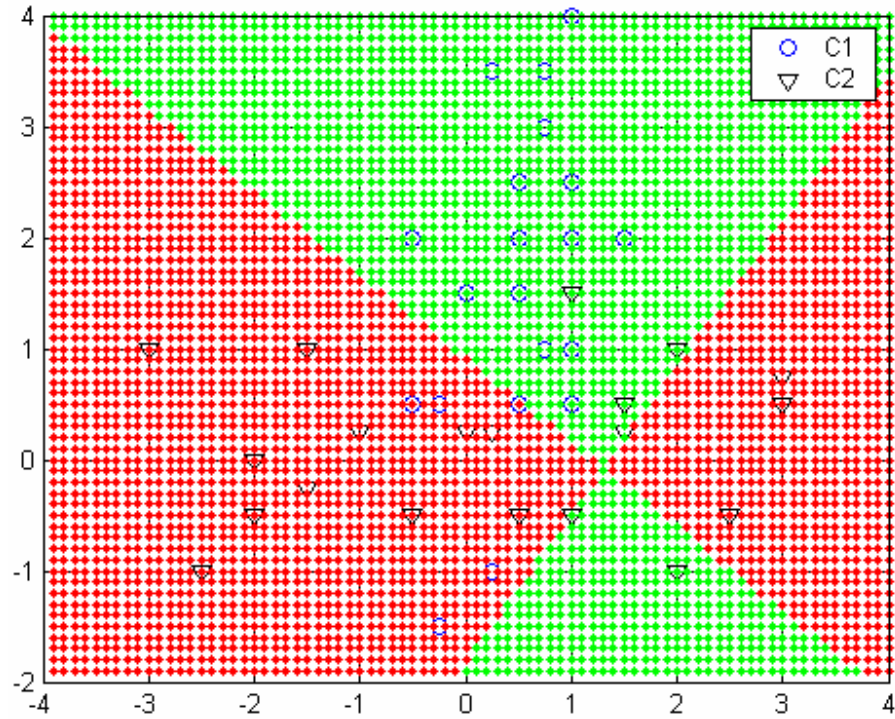
Şekil 3.8 C₁ ve C₂ sınıfları

	Tanıma Oranları (%)
PCA	72.5
LDA	75
OVY	85
OVY+PCA	90

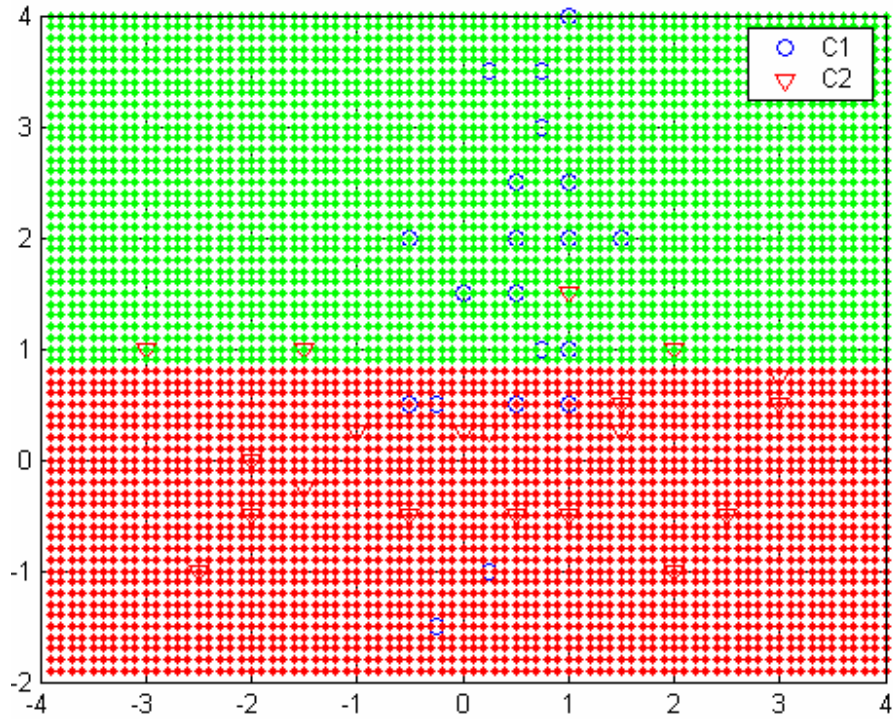
Tablo 3.1 PCA, LDA, OVY ve birleştirilmiş sınıflandırıcının tanıma başarımları



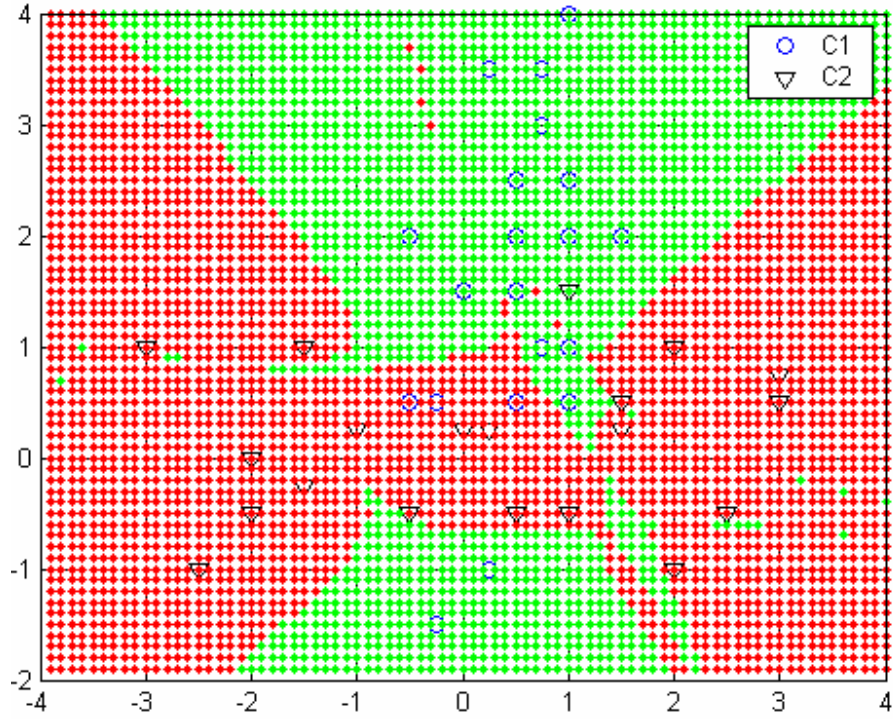
Şekil 3.9 Test kümesinin OLV ile sınıflandırılması



Şekil 3.10 Test kümesinin PCA'den elde edilen ölçüt ile sınıflandırılması



Şekil 3.11 Test kümesinin LDA ile sınıflandırılması



Şekil 3.12 OVY ve PCA'den elde edilen ölçütlerin birleştirilmesi ile test kümesinin sınıflandırılması

4. DENEYSEL ÇALIŞMALAR

Bu bölümde altuzay teknikleri ile sınıflandırıcı birleştirmenin sesli sınıflamadaki başarımları verilmektedir. Deneysel çalışmalarda TIMIT veritabanındaki sesliler kullanılmıştır.

4.1 TIMIT Veritabanı

TIMIT veritabanı otomatik ses tanıma sistemlerinin geliştirilmesinde kullanılan standart bir veri tabanıdır. TIMIT toplamda 6300 cümle içerir. ABD nin 8 ana leçe bölgesinden 630 konuşmacı tarafından 10'ar cümle söylenir. Tablo 4.1'de konuşmacı sayılarının 8 bölgeye göre dağılımı cinsiyet sayılarında belirtilerek verilmiştir. Yüzdeler oranlar parantez içinde verilmiştir. Bir konuşmacının lehçe bölgesi onun çocukluğunda yaşadığı coğrafi bölgedir. Coğrafi bölgeler ABD'de bilinen lehçe bölgeleridir. Fakat bunlar içinden Western bölgesi (dr7)'nin lehçe sınırları tam olarak bilinmemektedir ve 8. bölgenin konuşmacılarının çoğu çocukluklarında başka yerlere taşınmışlardır.

TIMIT'deki yazı materyali Stanford Araştırma Enstitüsü (SRI) tarafından hazırlanmış 2 "sibboleth" cümle, Massachusetts Teknoloji Enstitüsü (MIT) tarafından hazırlanmış 450 tane fonetik olarak kompakt cümle ve Texas Instruments (TI) tarafından hazırlanmış fonetik olarak çeşitli 1890 cümle içermektedir. Lehçe cümleleri (sa cümleleri) konuşmacıların lehçesel farklılıklarını ortaya çıkarmak için kullanılmıştır. Bu cümleler bütün konuşmacılar tarafından söylenmiştir. Tablo 4.2, TIMIT'deki cümlelerin türlerine göre dağılımlarını ve hangi tür cümleyi kaç kişinin gösterdiğini özetlemektedir.

Bölge (dr)	Lehçe Bölgeleri	#Erkek	#Bayan	Toplam
1	New England	31 (63%)	18 (27%)	49 (8%)
2	Northern	71 (70%)	31 (30%)	102 (16%)
3	North Midland	79 (67%)	23 (23%)	102 (16%)
4	South Midland	69 (69%)	31 (31%)	100 (16%)
5	Southern	62 (63%)	36 (37%)	98 (16%)
6	New York City	30 (65%)	16 (35%)	46 (7%)
7	Western	74 (74%)	26 (26%)	100 (16%)
8	Army Brat(taşınanlar)	22 (67%)	11 (33%)	33 (5%)
Toplam		438 (70%)	192 (30%)	630 (100%)

Tablo 4.1 Konuşmacıların lehçe dağılımları

Cümle Türü	#Cümle	#Konuşmacı	Toplam	#Cümle/Konuşmacı
diyalekt(SA)	2	630	1260	2
Kompakt (SX)	450	7	3150	5
Çeşitli (SI)	1890	1	1890	3
Toplam	2342	6300		10

Tablo 4.2 TIMIT ses materyali

4.2 Sesli Sınıflama

Yapılan deneylerde TIMIT veri tabanındaki 16 sesli kullanıldı. Deney kümesinde eğitim ve test kümelerindeki 'sa' cümlelerini kullanılmamıştır. Çünkü bu cümleler eğitim ve test kümelerinde geçtiklerinden test kümesinde tanıma oranlarında yapay bir artışa neden olmaktadır (Lee and Hon, 1989; Hataoka and Waibel, 1990). Deneylerde kullanılan sesliler:

'iy'	beet	bcl b IY tcl t
'ih'	bit	bcl b IH tcl t
'eh'	bet	bcl b EH tcl t
'ey'	baıt	bcl b EY tcl t
'ae'	bat	bcl b AE tcl t
'aa'	bott	bcl b AA tcl t
'aw'	bout	bcl b AW tcl t
'ay'	bite	bcl b AY tcl t
'ah'	but	bcl b AH tcl t
'ao'	bought	bcl b AO tcl t
'oy'	boy	bcl b OY
'ow'	boat	bcl b OW tcl t
'uh'	book	bcl b UH kcl k
'uw'	boot	bcl b UW tcl t
'ux'	toot	tcl t UX tcl t
'er'	bird	bcl b ER dcl d

Tablo 4.3 Deneylerde kullanılan sesliler

Deneylerde kullanılan özniteliklerden uzunluğu 10 ms ve ardışık iki pencerenin örtüşmesi 2.5 ms olacak şekilde 19 adet pencere çıkarıldı. Her bir pencereden 12 MFCC ve çerçeve enerjisinin logaritması olmak üzere toplam 13 parametre üretildi.

Bu durumda 145 ms'den daha uzun olanlarda ağırlık merkezinin sağından 9, solundan 9 olmak üzere 19 çerçeve alındı. 145 ms den kısa olanlarda ise seslinün uzunluğu 145 ms'ye denk gelecek şekilde sıfır eklendi. Ağırlık merkezi bulunurken

$$\bar{x} = \frac{\sum_{j=1}^n x_j m_j}{\sum_{j=1}^n m_j} \quad (4.1)$$

formülü kullanıldı (Adams, 2003). Burada x_j , ses sinyalinin j inci örneği, m_j bu örneğin momentidir. Çerçeve enerjileri, e_i i inci çerçevenin enerjisinin logaritması olmak üzere,

$$(e_i)_N = e_i - \max\{e_i\}, \quad i=1,2,\dots,m \quad (4.2)$$

formülüyle normalize edilmektedir. Öznitelik vektörü oluşturulurken her bir pencereden üretilen parametreler (4.3)'deki gibi ucuca eklenmiştir. Böylece her bir sesli için 247 boyutlu vektör üretilmiştir.

$$\vec{x} = \begin{bmatrix} \vec{x}_1^T : \vec{x}_2^T : \dots : \vec{x}_i^T \end{bmatrix} \quad i = 1, 2, \dots, 19 \quad (4.3)$$

Burada \vec{x}_i her bir pencereden üretilen öznitelik vektördür.

4.2.1 Özdeğer Seçimleri

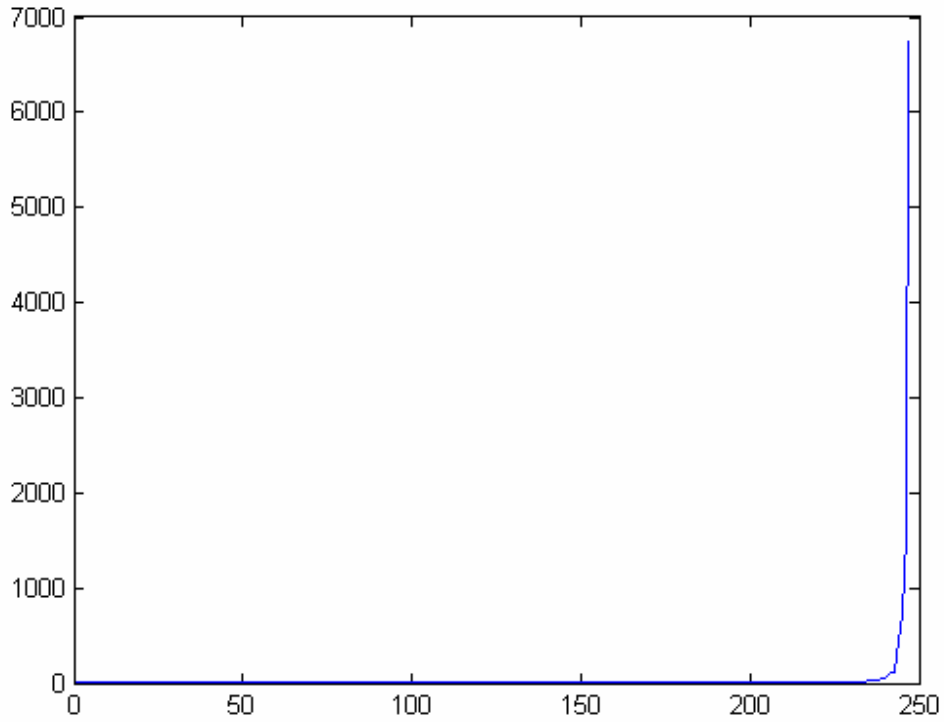
Deneylerde OVY'nin farksızlık altuzayına izdüşüm matrislerini elde ederken veri sayısı boyut sayısından fazla olduğundan yeterli veri durumu ortaya çıkmaktadır (Gülmezoğlu, et al., 2001). Yeterli veri durumunda farksızlık altuzayını gelecek özdeğer sayısı iki şekilde belirlenebilir. Bunlardan ilkinde özdeğer sayısı dirsek noktasından elde edilir. Bu yöntemde ortak değişinti matrisinin özdeğerleri küçükten büyüğe doğru sıralanır. Enerji değişiminin artmaya başladığı yere göre yaklaşık olarak özdeğer sayısı seçilir. Diğer yöntem ise enerji ölçütüdür. Bu ölçütte yine ortak değişinti matrisinin özdeğerleri en küçük olandan itibaren belirli bir enerji oranını elde edecek şekilde seçilir (Oja, 1983).

$$\frac{\sum_{i=1}^k \lambda_i}{\sum_{i=1}^n \lambda_i} \leq \kappa \quad (4.4)$$

Burada λ_i ortak değişinti matrisinin özdeğerlerini, k seçilen, n toplam özdeğer sayısını, κ ise $0 < \kappa \leq 1$ olacak şekilde gerçel bir sayıyı göstermektedir.

Şekil 4.1'de 'ey' seslisinin ortak değişinti matrisinin özdeğerlerinin küçükten büyüğe doğru sıralanmış olarak çizimi verilmiştir. OVY'nin farksızlık altuzayına ait izdüşüm matrisini elde etmek için 247 özvektörden özdeğerleri en küçük 235 tanesini seçilmiştir. Bu seçimi yaparken dirsek noktası yöntemi kullanılmıştır. Sınıflar arası

dağılımdan elde edilen ölçüde kullanılan izdüşüm matrisini elde ederken de 247 vektörden en büyük 12 özdeğere karşılık gelen özvektörler kullanılmaktadır. Bu çalışmada sınıf içi ve sınıflar arası ölçütlerden elde edilen özdeğer sayısı her sınıf için aynı alınmıştır.

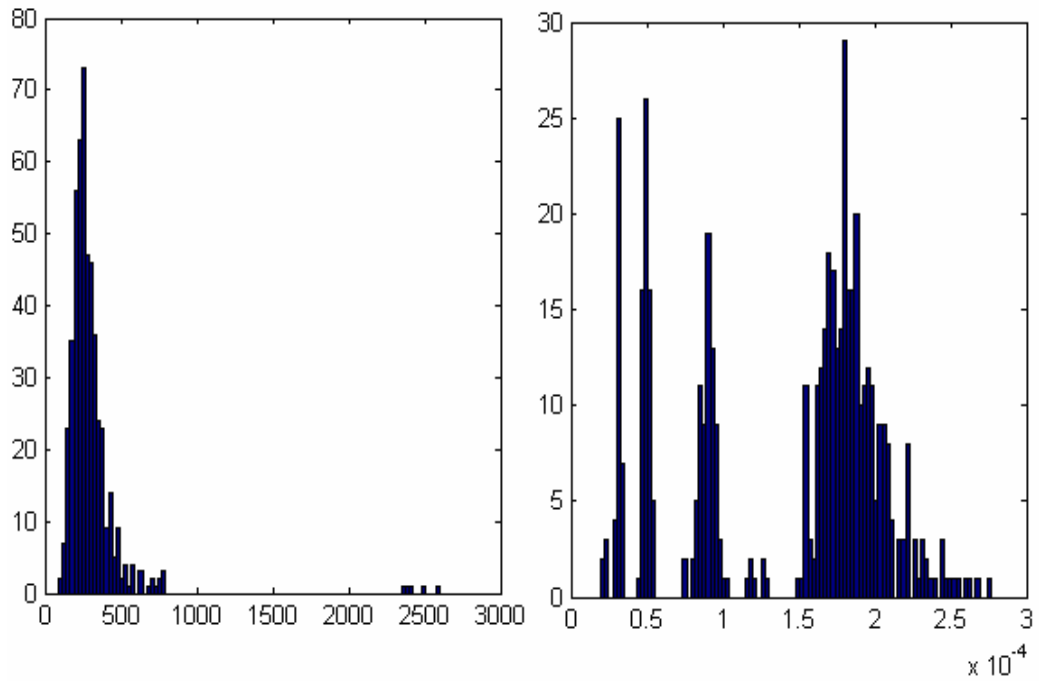


Şekil 4.1 'ey' seslisinin ortak değişinti matrisinin özdeğerleri

4.2.2 Ölçekleme Katsayısı

Ölçekleme katsayısı, farklı aralıklardaki sayıları aynı aralığa getirmek için kullanılan sayıdır. 'ey' seslisinin $F_{1_{c1}}$ ve $\frac{1}{F_{2_{c1}}}$ normlarının histogramları Şekil 4.2'de verilmiştir. Şekillerden de görüldüğü gibi bu iki normun büyüklükleri arasında büyük fark var. Bu durumda $w_{1_{c1}}, w_{2_{c1}}, w_{1_{c2}}$ ve $w_{2_{c2}}$ ağırlıklarını hesaplanırken kullanılan ızgara yöntemi kullanışlı bir yol olmayacaktır. Çünkü ızgara üzerindeki nokta aralıkları her azaltıldığında hesaplama maliyeti artacaktır. $\frac{1}{F_{2_{c1}}}$ normunun birleştirilmiş ölçüde etkili

olabilmesi için belirli bir ölçekleme katsayısıyla çarpılmıştır. İncelemeler sonucu farklı ölçekleme katsayısı değerlerinin TIMIT veritabanı için farklı başarımlar gösterdiği görülmüştür. Farklı ölçekleme katsayılarıyla farklı tanıma oranları elde edilmesine rağmen bu oranlar birbirine yakın çıkmaktadır. Yapılan deneylerde ölçekleme katsayısı olarak 1×10^5 kullanılmıştır.



Şekil 4.2 'ey' seslisi için F_{1c1} ve $\frac{1}{F_{2c1}}$ normlarının histogramları

4.2.3 Sınıflama Başarımları

Sınıf içi ve sınıflar arası dağılımlardan elde edilen ölçütlerin birleştirilmesini kullanan sınıflandırıcıyı çoklu sınıf probleminde test etmeden önce ilk olarak II. bölümdeki altuzayların sesli tanımadaki başarımları elde edilmiştir. Tablo 4.4'de CVA, PCA ve LDA'nin sınıflandırma başarımları verilmektedir. Bu sonuçlar, ileride karşılaştırma yapmada kullanılacaktır. Tablo 4.4'e bakıldığında en iyi sınıflandırmanın LDA ile elde edildiği görülmektedir. PCA'nin ise sınıflandırma başarımına bakıldığında, aslında PCA'nin sınıflandırmada tek başına iyi bir yöntem olmadığı görülmektedir. CVA'nin ise LDA'ye yakın sonuç verdiği görülmektedir.

Sesli	Eğitim			Test		
	OVY	PCA	LDA	OVY	PCA	LDA
'iy'	46.2	10.6	73.2	34.8	9.4	72.4
'ih'	55.0	61.8	50.6	38.0	57.8	41.6
'eh'	28.2	14.0	43.0	18.6	10.8	43.8
'ey'	72.2	15.4	71.6	67.0	13.4	69.8
'ae'	48.0	2.2	58.0	40.2	0.8	52.0
'aa'	52.2	3.0	55.6	36.0	2.2	47.6
'aw'	83.6	82.0	67.2	75.0	89.5	65.5
'ay'	76.4	70.4	68.6	65.4	64.2	63.2
'ah'	23.8	4.8	45.4	15.4	3.6	39.0
'ao'	41.6	5.8	47.2	43.0	5.8	53.6
'oy'	87.3	27.7	76.3	74.4	22.4	68.8
'ow'	61.6	7.6	64.4	49.6	8.4	54.6
'uh'	49.5	51.1	41.7	45.0	51.0	42.0
'uw'	64.6	4.8	71.4	56.0	2.7	64.0
'ux'	52.6	2.2	56.8	41.8	1.5	55.8
'er'	77.4	16.6	76.4	60.4	12.2	69.0
Ortalama	57.5	23.7	60.0	47.5	22.2	55.6

Tablo 4.4 OVY, PCA, LDA yöntemlerinin sesli tanımadaki başarımları

İki sınıf problemi deneysel çalışmada iki şekilde ele alındı. Önce sınıflar ikiye ayrılarak alınarak sınıflama yapıldı.. Daha sonra deneylerde iki sınıf problemi bir sınıf ve geriye kalan diğer sınıflar göz önüne alınarak uygulandı. Ağırlıkları hesaplamak için önce Şekil 3.6'daki ızgara üzerindeki her bir $w_{1_{c_1}}$ ve $w_{1_{c_2}}$ değeri için (3.19) kullanılarak $w_{2_{c_1}}$ ve $w_{2_{c_2}}$ değerleri elde edilmiştir. Daha sonra bu ağırlıklardan eğitim kümesinde en iyi tanıma başarımını veren ağırlıklar deneylerde kullanılmıştır.

Birleştirilen sınıflandırıcı ile LDA'yı karşılaştırabilmek için enküçültme ve enbüyütme probleminden elde edilen ölçütler kullanılarak LDA1 sınıflandırma kriteri tanımlanmıştır. Bu kriterde sınıflama aşağıdaki şekilde yapılmaktadır.

$$C^* = \arg \min_i \frac{\left\| \left(P_1^{(i)} \right)^\perp \left(\vec{a}_x - \vec{a}_{ave}^{(i)} \right) \right\|}{\left\| P_2^{(i)} \left(\vec{a}_x - \vec{a}_{ave,rem}^{(i)} \right) \right\|} \quad (4.5)$$

Burada $a_{ave}^{-(i)}$ ve $a_{ave,rem}^{-(i)}$ sırasıyla i inci sınıfın ve i inci sınıf dışındaki sınıfların ortalamasını göstermektedir..

‘ey’ ile ‘ae’ seslileri için sınıflandırma sonuçları Tablo 4.5’de verilmektedir. Burada birleştirilen ölçüde kullanılan ağırlıklar $w_{1_{c_1}} = 0.99$, $w_{2_{c_1}} = -0.1411$, $w_{1_{c_2}} = 0.93$ ve $w_{2_{c_2}} = -0.3676$ olarak elde edilmiştir. Tablo 4.6 , ‘ih’ ile ‘eh’ seslileri için elde edilen başarımları göstermektedir. Bu başarımlar $w_{1_{c_1}} = 0.99$, $w_{2_{c_1}} = -0.1411$, $w_{1_{c_2}} = 0.93$ ve $w_{2_{c_2}} = -0.3676$ ağırlık değerleri kullanılarak elde edilmiştir.

Yöntem	Eğitim(%)	Test(%)
OVY	90.8	89.3
PCA	56.6	54
LDA	97.1	91.1
LDA1	86.8	82.4
OVY+PCA	91.6	90

Tablo 4.5 'ey' ve 'ae' seslilerinin sınıflandırma sonuçları

Yöntem	Eğitim(%)	Test(%)
OVY	77.2	68.8
PCA	61.5	58
LDA	87.8	78.4
LDA1	62.9	59.2
OVY+PCA	79.8	69

Tablo 4.6 'ih' ve 'eh' seslilerinin sınıflandırma sonuçları

İki-sınıf problemi bir sınıf ve diğer sınıflar şeklinde göz önüne alındığında OVY, PCA, LDA ve birleştirilen sınıflandırıcı için Tablo 4.7’deki başarımlar elde edilmiştir. Bu deneylerde kullanılan ağırlıklar Tablo 4.8’de verilmiştir. İkinci sınıfta, ilk sınıfın 15 katı veri olduğu için tanınma yüzdesi hesaplanırken birinci ve ikinci sınıfın tanıma yüzdelerinin ortalaması alındı. Eğitim kümesinde birleştirilen sınıflandırıcı ile genel itibariyle OVY, PCA ve LDA’dan daha iyi tanıma başarımları elde edilmiştir. Fakat test

kümesinde çoğunlukla LDA daha iyi tanıma başarımı vermiştir. Test kümesinde birleştirilen ölçüt ile sadece iki testte en iyi tanıma başarımı elde edilmiştir.

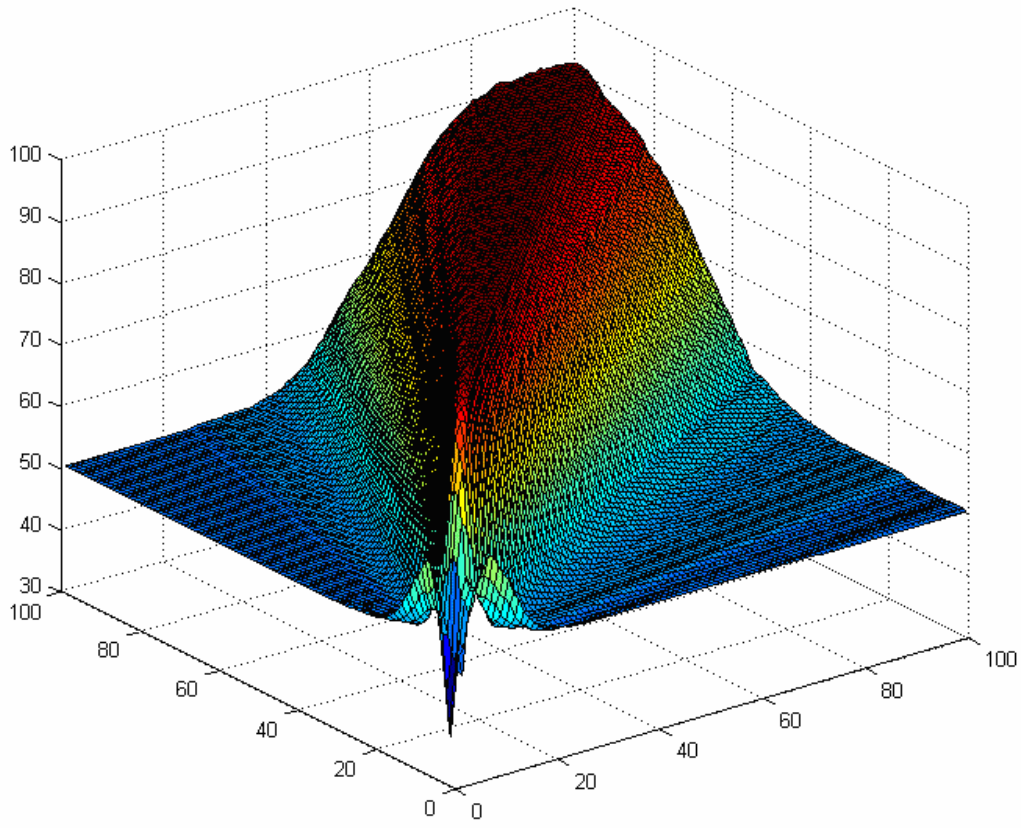
Şekil 4.3'de $w_{1_{c_1}}$ ve $w_{1_{c_2}}$ ağırlıklarının ızgara üzerinde aldıkları farklı değerlere göre, 'ey' ve 'ae' seslilerinin tanıma yüzdelerinin grafiği verilmiştir. Grafikte tanıma yüzdesinin yüksek değer aldığı noktalar, genel olarak $w_{1_{c_1}}$ ve $w_{1_{c_2}}$ 'nin birbirine eşit olduğu yerlere yakın noktalardır. Bu da OVY'den elde edilen sınıflandırıcının, PCA'den elde edilen sınıflandırıcıdan daha baskın olduğunu gösterir. Şekil 4.4'de, Şekil 4.3'deki grafiğin bir kesitinin büyütülmüş olarak gösterimi vardır. Şekil 4.4'deki yüzey grafiğinde girinti, çıkıntı çok fazla olduğundan, bayır inişi yöntemi ile en iyi tanıma yüzdesini veren ağırlıkları bulmak mümkün olmayabilir.

Sesli	Eğitim					Test				
	OVY	PCA	LDA	LDA1	OVY+PCA	OVY	PCA	LDA	LDA1	OVY+PCA
'iy'	87.3	64.7	91.2	74.5	88.2	85.3	66.7	90.0	75.6	87.3
'ih'	78.8	71.7	80.9	72.6	79.1	73.0	72.3	74.8	73.4	72.2
'eh'	79.2	64.0	77.4	67.1	79.8	74.7	67.0	76.6	68.4	74.9
'ey'	86.6	59.4	88.1	69.4	87.2	83.5	61.2	83.6	70.9	83.6
'ae'	84.3	61.9	85.7	64.5	84.3	81.3	61.3	81.5	64.1	81.3
'aa'	83.4	58.9	85.2	72.9	84.2	81.9	58.6	78.6	71.3	80.1
'aw'	87.0	69.2	85.1	72.4	88.5	83.7	73.4	84.8	75.9	83.0
'ay'	87.2	68.7	86.6	72.3	87.7	84.3	69.9	85.6	73.1	82.1
'ah'	77.7	65.1	81.7	68.9	82.6	70.2	63.1	76.2	66.5	76.8
'ao'	84.7	58.8	84.6	77.6	84.8	83.8	58.4	84.8	77.0	83.2
'oy'	90.4	76.0	88.3	77.9	92.1	85.8	75.0	85.0	75.8	86.4
'ow'	84.1	58.8	84.2	65.9	85.0	82.0	59.3	79.3	66.2	81.7
'uh'	80.2	71.5	78.8	72.3	80.4	77.0	70.1	73.8	71.2	76.1
'uw'	90.0	59.5	87.0	80.0	90.1	86.5	56.2	84.2	77.7	86.3
'ux'	85.3	59.4	87.3	70.5	85.7	81.6	62.8	85.4	72.4	81.3
'er'	89.8	56.7	90.0	85.3	90.9	87.2	58.7	86.2	84.3	86.0
Ortalama	84.8	64.0	85.1	72.8	85.7	81.4	64.6	81.9	72.7	81.4

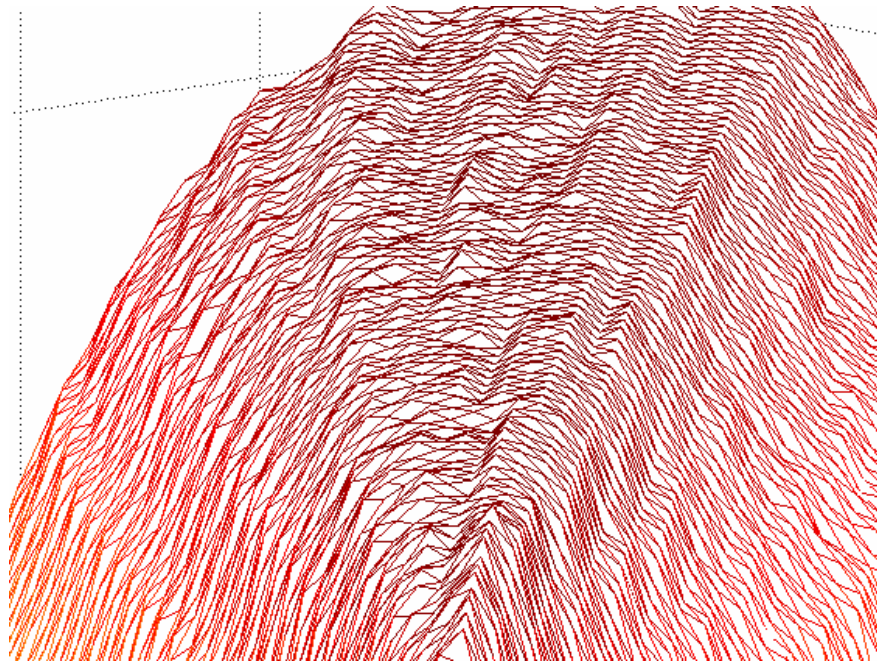
Tablo 4.7 Deneylerde OVY, PCA, LDA, LDA1 ve birleştirilen ölçüt ile elde edilen tanıma başarımları

Sesli	Ağırlıklar			
	<i>Sınıf-1 (C1)</i>		<i>Sınıf-2 (C2)</i>	
	w_1	w_2	w_1	w_2
'iy'	0,18	-0,984	0,2	-0,98
'ih'	0,99	-0,141	0,98	-0,199
'eh'	0,88	-0,475	0,92	-0,392
'ey'	0,97	-0,243	0,92	-0,392
'ae'	1	0	1	0
'aa'	0,92	-0,392	0,88	-0,475
'aw'	0,97	-0,243	0,88	-0,475
'ay'	0,98	-0,199	0,89	-0,456
'ah'	0,87	-0,493	0,99	-0,141
'ao'	0,97	-0,243	1	0
'oy'	0,99	-0,141	0,94	-0,341
'ow'	0,99	-0,141	0,94	-0,341
'uh'	0,99	-0,141	0,99	-0,141
'uw'	0,4	-0,917	0,4	-0,917
'ux'	0,25	-0,968	0,25	-0,968
'er'	0,52	-0,854	0,49	-0,872

Tablo 4.8 Deneylerde kullanılan ağırlıklar



Şekil 4.3 W_{1c_1} ve W_{1c_2} ağırlıklarının ızgara üzerinde aldıkları değerlere göre 'ey' ve 'ae' seslilerinin birleştirilen sınıflandırıcı ile tanınma yüzdeleri



Şekil 4.4 Şekil 4.3'deki grafiğin bir kesitinin büyütülmüş hali

5. SONUÇLAR ve ÖNERİLER

Bu çalışmada, sınıf içi ve sınıflar arası dağılımlardan elde edilen ölçütlerin sınır ağları ile nasıl birleştirileceği gösterilmiştir. Buradaki amaç her iki ölçütü kullanarak çoklu sınıflamadaki başarıyı arttırmaktır. İyi bir sınıflandırıcı hem sınıf içi ve hem de sınıflar arası dağılımları göz önüne almalıdır. Bu çalışmada önerilen ölçüt birleştirme yönteminde her iki dağılım göz önüne alınmaktadır. Önerilen ölçüt birleştirme yönteminin başarıyı, sınıf içi ve sınıflar arası dağılımları gözönüne alan LDA ile karşılaştırılmıştır.

Önerilen sınıflandırıcı birleştirme yönteminin başarıyı iki boyutlu veriler kullanılarak iki-sınıf probleminde test edildi. Elde edilen sonuçlara göre, iki boyutlu veriler için sınıflandırıcı birleştirme yöntemi OVY, PCA ve LDA'den daha iyi başarımlar vermektedir. Bu sonuç önerilen sınıflandırıcı birleştirme şeklinin gerçek sınıflandırma problemleri içinde sınıflandırma başarımlarını arttıracak şekilde yorumlanmıştır.

Deneyel çalışmalarda OVY, PCA ve LDA yöntemlerinin tek başlarına TIMIT veri tabanındaki seslileri sınıflandırma başarımları elde edildi. Elde edilen sonuçlara göre LDA'nin daha iyi başarımlar verdiği görülmüştür. PCA'in ise sınıflandırmada başarısız olduğu görüldü. Son deneylerde ise OVY, PCA, LDA, LDA1 ve önerilen sınıflandırıcı iki tür iki-sınıf probleminde sesli tanımadaki başarımlarına göre karşılaştırıldı.

Bu çalışmada, birleştirilen iki sınıflandırıcıdan elde edilen normları aynı aralığa getirmek için ölçekleme katsayısı kullanıldı. Deneylerde kullanılan eğitim kümelerinde genelde birleştirilen ölçüt ile daha iyi tanıma başarımları elde edilmiştir. Ancak test kümelerinde çoğunlukla LDA ile daha iyi sınıflandırma başarımları elde edilmiştir. Birleştirilen sınıflandırıcı ile hem eğitim hem de test kümelerinde LDA1'e göre daha iyi sınıflandırma başarımları elde edilmiştir. Birleştirilen sınıflandırıcı sadece iki testte daha iyi sonuç vermiştir. Test kümelerinde çoğunlukla CVA birleştirilen ölçütten daha iyi sonuçlar vermiştir. Ayrıca CVA ve önerilen sınıflandırıcının başarımları birbirine çok

yakın çıkmıştır. Buradan PCA'nin sınıflandırma başarımını istenilen oranda arttırmadığı söylenebilir.

Sonuçlardan da anlaşılacağı gibi sınıflandırıcı birleştirme yöntemi LDA'den daha iyi başarımlar vermemektedir. Bunun sebeplerini araştırırken öncelikle birleştirilen iki sınıflandırıcının birbirini ne kadar iyi tamamladıklarına bakılmalıdır. Bunun yapılması için sınıflandırıcıların çeşitliliği ölçülmelidir. Sınıflandırıcıların çeşitliliğinin ölçülmesi için çeşitli yöntemler verilmiştir (Kuncheva, 2004; Ranawana and Palade, 2006). Oluşturulan mimari üzerinde çeşitli değişiklikler yapılabilir. Sınıflandırıcı birleştirmede $w_{1_{c_1}}, w_{2_{c_1}}, w_{1_{c_2}}$ ve $w_{2_{c_2}}$ ağırlıkları hesaplanırken ızgara yöntemini kullanıldı. Bu yöntem yerine bayır-inme, bayır-çıkma yöntemleri veya klasik yapay sinir ağı algoritmalarıyla $w_{1_{c_1}}, w_{2_{c_1}}, w_{1_{c_2}}$ ve $w_{2_{c_2}}$ için sınıflandırmadaki başarımı arttıran değerler elde edilebilir.

OVY ve PCA'den elde edilen sınıflandırıcıların farklı şekillerde birleştirilmesiyle daha iyi sonuçlar elde edilebilir. Bu çalışmada çıkıştaki ağırlıklar $\{-1,+1\}$ olarak alınmıştır. Bu ağırlıklar değişken olarak ele almak ve bu ağırlıkların eniyi değerlerini çeşitli ölçütlere göre elde etmek sınıflandırma başarımını arttırabilir.

6. KAYNAKLAR DİZİNİ

- Adams, R.A., Calculus a Complete Course, Pearson Education Inc. pp. 432
- Balakrishnama, S., Ganapathiraju, A., and Picone, J., Linear Discriminant Analysis for Signal Processing Problems, Southeastcon Proceedings. IEEE, 1999
- Bishop, C.M., Neural networks for Pattern Recognition, Aston University, Birmingham, UK, 1996, pp 310-313
- Cao, J., Ahmadi, M., Shridhar, M., Recognition of Handwritten Numerals with Multiple Feature and Multi-stage Classifier, Pattern Recognition, Vol. 28, No. 2, pp. 153-160, 1995
- Çevikalp, H., Neamtu, M., and M. Wilkes, Discriminative common vector method with kernels, IEEE Trans. on Neural Networks
- Dietterich T. G., Ensemble Methods in Machine Learning, Lecture Notes in Computer Science, Cagliari, Italy, 2000, Springer, pp. 1-15
- Duda, R.O., Hart, P.E., Stork, D.G., Pattern Classification, Second Edition, John Wiley and Sons Inc., 2001, pp. 117-121
- Duin, R.P.W., Loog, M., and Umbach, R.H., Multi-class Linear Feature Extraction by Nonlinear PCA, ICPR15, Proc. 15th Int. Conference on Pattern Recognition, 2000
- Edizkan, R., Gülmezoğlu, M.B., Ergin, S., Barkana, A., Improvements on Common Vector Approach for Multi Class Problems, 13th European Signal Processing Conference, Antalya, 2005
- Felföldi, L., Kocsor, A., Tóth, L.: Classifier Combination in Speech Recognition, Periodica Polytechnica, Electrical Engineering, Vol. 47, No. 1-2, pp. 125-140, 2003
- Gader, P., Mohamed, M., and Keller, J., Fusion of handwritten word classifiers. Pattern Recognition Letters, vol. 17, no. 6, pp. 577–584, 1996.
- Gülmezoğlu, M.B., DZhafarov, V., Barkana, A., The common vector approach and its relation to principal component analysis, IEEE Trans., Speech and Audio Processing, vol9. no.6, 2001.

- Gülmezoglu, M.B., Cevikalp, H., Barkana, A., A new point of view to common vector approach, Signal Processing and Communications Applications Conference, 2004. Proceedings of the IEEE 12th, 2004, pp. 732 - 734
- Gülmezoğlu, M. B., Dzhafarov, V., Keskin, M., and Barkana, A., A Novel Approach To Isolated Word Recognition, IEEE Trans. Speech and Audio Processing, vol. 7, pp. 620-628, 1999
- Günel, S., Ergin, S., Gerek, Ö.N., Spam e-mail recognition by subspace analysis, Proc. INISTA 2005, Istanbul, 2005.
- Hataoka, N., Waibel, A., Speaker-Independent Phoneme Recognition on TIMIT Database Using Integrated Time-Delay Neural Networks (TDNNs), Proceedings of the International Joint Conference on Neural Networks (IJCNN), Vol. I, pp. 57-62, 1990
- Heskes, T., Balancing between bagging and bumping. Neural Information Processing Systems, vol. 9, pp. 466–472. MIT Press, Cambridge, 1997
- Ji, C. and Ma, S. Combinations of weak classifiers. IEEE Transactions on Neural Networks, vol. 8, no. 1, pp. 32–42, 1997
- Katz, M., Meier, H.-G., Dolfing, H., Klakow, D., Robustness of Linear Discriminant Analysis in Automatic Speech Recognition, International Conference on Pattern Recognition, vol. 3, pp. 371-374
- Kirby, M., Sirovich, L., Application of the Karhunen-Loeve Procedure and the Characterization of Human Faces. IEEE Transactions on Pattern Analysis and Machine Intelligence, pp. 103-108, 1990
- Kittler, J., Hatef, M., Ruin, R.P.W., Matas, J., On Combining Classifiers, IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 20, no. 3, 1998
- Kramer, M. A. , Nonlinear Principal Component Analysis Using Autoassociative Neural Networks, Neural Computation 9 (7), 1493--1516
- Kuncheva, L.I., Combining Pattern Classifiers: Methods and Algorithms, John Wiley and Sons, Inc, 2004, pp 295
- Lee, K.F., Hon, H.W., Speaker-Independent Phone Recognition Using Hidden Markov Models, IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. 37, No. 11, Nov. 1989

- Lu, X., Wang, Y., and Jain, A.K., Combining Classifiers for Face Recognition, Proceedings of International Conference on Multimedia and Expo, Vol. 3, pp. 13-16, 2003
- Radova, V., Psutka, J., An Approach to Speaker Identification Using Multiple Classifiers. Proc. of the IEEE Int. Conf. on Acoustics, Speech and Signal Processing, 1997, pp. 1135–1138
- Ranawana, R., Palade, V., Multi Classifier Systems – A Review and Roadmap for Developers, Journal of Hybrid Intelligent Systems, 2006
- Swets, D.L., Weng, J.J., Using Discriminant Eigenfeatures for Image Retrieval, IEEE Trans. Pattern Analysis and Machine Intelligence, vol.18, no. 8, pp. 831-836, 1996
- Taniguchi, M. and Tresp, V., Averaging regularized estimators. Neural Computation, vol. 9, pp. 1163–1178, 1997
- Tax, D., Breukelen, M. V., Duin, R., and Kittler, J., Combining multiple classifiers by averaging or by multiplying?, Pattern Recognition, vol. 33, pp. 1475–1485, 2000
- Wang, X., Paliwal, K.K., A modified minimum classification error (MCE) training algorithm for dimensionality reduction, Journal of VLSI Signal Processing Systems, Vol. 32, pp. 19-28, 2002
- Xu, L., Krzyzak, A., and Suen, C., Methods of combining multiple classifiers and their applications to handwriting recognition. IEEE Transactions on Systems, Man, and Cybernetics, vol. 22, no. 3, pp. 418–435, 1992
- Zhang, B., Fu, M., Yan, H., A Nonlinear Neural Network Model of Mixture of Local Principal Component Analysis: Application to Handwritten Digits Recognition, Pattern Recognition