

**YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**SAKLI MARKOV MODEL KULLANILARAK
GÖRÜNTÜDEN GERÇEK ZAMANLI TÜRK İŞARET DİLİ
TANIMA SİSTEMİ**

Elektronik ve Haberleşme Mühendisi Hakan HABERDAR

**FBE Bilgisayar Mühendisliği Anabilim Dalında
Hazırlanan**

YÜKSEK LİSANS TEZİ

Tez Danışmanı : Yrd. Doç. Dr. Songül ALBAYRAK

İSTANBUL, 2005

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	iv
KISALTMA LİSTESİ.....	v
ŞEKİL LİSTESİ.....	vi
ÇİZELGE LİSTESİ	vii
ÖNSÖZ	viii
ÖZET	ix
ABSTRACT	x
1. GİRİŞ	1
1.1 İşaret Dili Tanıma Sistemleri	2
1.2 Tez Çalışmasında Tasarlanan Sistem	4
2. TÜRK İŞARET DİLİ.....	7
2.1 Türk İşaret Dili Sözlüğü.....	8
2.2 Türk İşaret Dili Kılavuzu	8
2.3 Tez Çalışmasında Kullanılan Kelimeler	10
3. MARKOV MODELLERİNİN TEMEL ÖZELLİKLERİ	14
3.1 Ayırık Markov Prosesleri	14
3.2 Saklı Markov Model	16
3.3 Saklı Markov Modellerinin Üç Temel Problemi.....	18
3.3.1 Tanıma Problemi	18
3.3.2 Durum Dizisinin Bulunması.....	21
3.3.3 Model Parametrelerinin Öğrenilmesi	22
3.4 Model Seçimi	24
3.5 Sürekli Gözlemler.....	24
3.5.1 Durum Geçiş Olasılıkları	25
3.5.2 Sürekli Saklı Markov Modelde Üç Temel Problem	26
3.6 Saklı Markov Modelin Video Görüntüsünden Hareket Tanımaya Uygulanması ..	28
4. ÖZELLİK SEÇİMİ	31
4.1 Ortamın Aydınlatılması	31
4.2 Görüntü Yakalama Aygıtı	32
4.3 Görüntüde Ellerin Bulunması.....	32
4.3.1 Renkli Görüntünün Temel Özellikleri	33
4.3.2 Renk Uzayları.....	33
4.3.2.1 RGB Renk Uzayı	33

4.3.2.2	YIQ ve YUV Renk Uzayları	35
4.3.2.3	HSI, HSV ve YCbCr Renk Uzayları	36
4.3.3	Ten Rengi Tanıma	37
4.3.4	El ve Yüz Olabilecek Şekillerin Değerlendirilmesi	40
4.4	Hareket Takibi	42
4.5	Anahtar Çerçeve Seçimi	42
4.6	Global Özelliklerin Çıkarılması	44
4.7	Yerel Özelliklerin Belirlenmesi	46
5.	TÜRK İŞARET DİLİ TANIMA SİSTEMİ	49
5.1	Sistem Yapısı ve Temel Tanıma Prensipleri	49
5.1.1	İşaretin Global Özellik Grubunun Belirlenmesi	49
5.1.2	İşaretin Yerel Özelliklere Göre Seçilmesi	50
5.2	Kelimeleri Ayrı Ayrı Tanıma	52
5.3	Cümle Tanıma	52
6.	DENEYSEL TANIMA SONUÇLARI	54
6.1	Kelimelerin Global Özellikler Kullanılarak Tanınması	54
6.2	Kelimelerin Yerel Özelliklerin Kullanılarak Tanınması	58
7.	SONUÇLAR VE ÖNERİLER	62
	KAYNAKLAR	65
	EKLER	67
	Ek 1a Temel Olasılık Kavramları	68
	Ek 1b Koşullu Olasılık	68
	Ek 1c Olasılık Yoğunluk Fonksiyonu	68
	Ek 1d Beklenen Değer	68
	Ek 1e Lineer Bağımlılık	69
	Ek 1f Normal (Gauss) Dağılım	69
	Ek 1g Çoklu Rastlantı Değişkeni	70
	Ek 1h Parametre Kestirimi	70
	ÖZGEÇMİŞ	72

SİMGE LİSTESİ

Δt	İşaretler arası durma eşik süresi
A	Durum geçiş olasılıkları matrisi
B	Gözlem olasılıkları matrisi
H	Herhangi bir global özellik grubundaki üye sayısı
K	Eğitim verisindeki gözlem dizisinin sayısı
k	Seçilen en yakın k komşu
M	Ayrık gözlem dizisinin uzunluğu
N	Durum sayısı
O	Gözlem dizisi
O^k	Eğitim setindeki k. gözlem dizisi
o_t	Sistemin t anında ürettiği gözlem
Q	Durum dizisi
q_t	Sistemin t anında bulunduğu durum
s	Eğitim verisinde her grup üyesinden olan örnek sayısı
T	Gözlem dizisinin uzunluğu
V	Ayrık gözlem vektörünün alabileceği değerler kümesi
$\alpha_t(i)$	İleri-yön değişkeni
$\beta_t(i)$	Geri-yön değişkeni
$\gamma_t(i)$	t anında i durumunda olma olasılığı
λ	Saklı Markov model parametre grubu
Π	İlk durum olasılıkları vektörü

KISALTMA LİSTESİ

CCD	Charge Coupled Device
CCL	Connected Component Labeling
HMMs	Hidden Markov Models
HSI	Hue Saturation Intensity
HSV	Hue Saturation Value
kNN	k Nearest Neighbor
NTSC	National Television Standards Committee
PAL	Phase Alternating Line
RGB	Red Green Blue
TİD	Türk İşaret Dili
TSL	Tint Saturation Luma
TÜBA	Türkiye Bilimler Akademisi
YCbCr	Luminance, Mavi Renklilik, Kırmızı Renklilik
YIQ	Luminance Inphase Quadrature

ŞEKİL LİSTESİ

Şekil 1.1 Hareket ifadeli kelime [1].....	1
Şekil 1.2 Parmak hecelemeli kelime.....	1
Şekil 1.3 İşaret dili tanıma sistemleri.....	3
Şekil 1.4 Görüntüde sağ el, sol el ve yüzün ayırt edilmesi.....	5
Şekil 2.1 TİD’de kelime ekleri vücut hareketleri yada parmakla yön belirtilerek verilir.....	7
Şekil 2.2 J ve Ğ harflerinin TİD karşılıkları (MEB, 1995).....	9
Şekil 2.3 Kılavuzda kullanılan hareket okların anlamları.....	9
Şekil 2.4 Hareket sırasında ellerin izlediği yörüngeler ve global özellik grupları.....	11
Şekil 3.1 3 durumdan oluşan bir Markov model.....	15
Şekil 3.2 Sistem bir durumda iken bu durumun tetiklediği gözlemi ortaya çıkarır.....	17
Şekil 3.3 İleri-yön değişkeni gözlem dizisini iki parçaya ayırır.....	19
Şekil 3.4 İleri-yön ve geri-yön değişkenleri.....	20
Şekil 3.5 Geri-yön algoritmasında yineleme adımı.....	21
Şekil 3.6 $\xi_i(i, j)$ ’nin hesaplanması.....	23
Şekil 3.7 Dört durumlu soldan-sağa saklı Markov model topolojisi.....	24
Şekil 3.8 Gözlem üretmeyen başlangıç ve son durumları eklenen model.....	25
Şekil 3.9 Baum-Welch algoritmasına başlamadan önce başlangıç değeri bulma adımları.....	27
Şekil 3.10 Baum-Welch algoritmasının adımları.....	28
Şekil 3.11 Görüntü çerçevesinden gözlem dizisi oluşturulması.....	29
Şekil 3.12 Saklı Markov model ile hareket tanıma.....	29
Şekil 4.1 Ortamın aydınlatılması.....	31
Şekil 4.2 İçinde çok çeşitli renkler içeren, Ebru Sanatı ile yapılmış örnek resim [3].....	33
Şekil 4.3 (a) Kırmızı bileşen (b) Yeşil bileşen (c) Mavi bileşen.....	34
Şekil 4.4 Verilen örnek resmin Y (gri seviyeli) bileşeni.....	35
Şekil 4.5 (a) Orijinal resim (b) H bileşeni (b) S bileşeni (b) I bileşeni (Russ, 1999).....	36
Şekil 4.6 (a) Orijinal resim (b) Aday ten rengi piksellerin bulunduğu resim.....	38
Şekil 4.7 (a) Orijinal Resim (b) Ten olmayan bölgelerin seçilmesi.....	40
Şekil 4.8 CCL algoritmasında p pikselinin a,b,c,d ile gösterilen 4 komşuna bakılır.....	40
Şekil 4.9 CCL algoritması ile ten rengi olan piksellerden ellerin ve yüzün bulunması.....	41
Şekil 4.10 Ellerin merkezlerinin bulunması.....	42
Şekil 4.11 Anahtar çerçevelerin seçilmesi.....	43
Şekil 4.12 İşaret sırasında ellerin konumu.....	43
Şekil 4.13 Koordinat sistemi.....	44
Şekil 4.14 Aynı işaretin farklı kamera açılarıyla tekrarlanması.....	44
Şekil 4.15 Orijini yüzün merkez noktası olan koordinat sistemi.....	45
Şekil 4.16 Elin aldığı biçim : (a) acımak, (b) Aksaray, (c) korkmak, (d) subay.....	46
Şekil 4.17 Ellerin anahtar çerçevelerdeki durumu.....	47
Şekil 4.18 c1 global özellik grubuna ait iki elli işaretlerde hareketin sonunda ellerin biçimi: (a) ağır, (b) bırakmak, (c) dinlenmek, (d) kabul etmek, (e) olmak, (f) yer.....	47
Şekil 4.19 d5 global özellik grubuna ait işaretlerde hareketin başında ellerin biçimi: (a) fırıncı, (b) futbolcu, (c) oyuncakçı.....	48
Şekil 5.1 Global özelliklerin belirlenmesi.....	49
Şekil 5.2 Tüm modellerin O gözlemine üretme olasılıkları.....	50
Şekil 5.3 Karşılaştırma öncesi yapılan ön işlemler.....	50
Şekil 6.1 Yakın global ve yerel özelliklere sahip işaretler.....	61
Şekil 6.2 Yüz ifadesinin ve dudakların işaretlere kattığı anlam.....	61

ÇİZELGE LİSTESİ

Çizelge 2.1 Sistemin tanıdığı kelimeler ve global özellik grupları	12
Çizelge 2.2 Global özellik gruplarındaki kelime sayıları	13
Çizelge 6.1 Eğitim verisinin özellikleri	54
Çizelge 6.2 Tek elleri hareketlere ait global özellik grubu belirleme sonuçları.....	55
Çizelge 6.3 Hatalı sınıflandırılan gruplar (tek elleri hareketler)	55
Çizelge 6.4 İki elleri hareketlere ait global özellik grubu belirleme sonuçları	56
Çizelge 6.5 Hatalı sınıflandırılan gruplar (iki elleri hareketler)	57
Çizelge 6.6 Tüm işaretlerin beraber değerlendirilmesi.....	58
Çizelge 6.7 Yerel özelliklerin kullanılması.....	59
Çizelge 6.8 Yerel özelliklere ait test sonuçları.....	60

ÖNSÖZ

Danışmanım Sayın Yrd. Doç. Dr. Songül ALBAYRAK'a tez çalışmam boyunca beni yönlendirdiği, bilgisini, değerli görüşlerini ve deneyimlerini bana aktardığı için çok teşekkür ederim.

Yıldız Teknik Üniversitesi ailesine katıldığım ilk günden beri, her aşamada, her zaman desteğini esirgemeyen, Sayın Yrd. Doç. Dr. Banu DİRİ'ye teşekkür ederim.

Hayatımın ve çalışmalarımın her aşamasında bana destek olan arkadaşım Rubabe ŞİNASİ'ye; oda arkadaşlarım Arş. Gör. M. Fatih AMASYALI ve Arş. Gör. Nilgün ERDEM başta olmak üzere YTÜ Bilgisayar Mühendisliği bölümündeki bütün arkadaşlarıma teşekkür ederim.

Hayatım boyunca desteklerini hiçbir zaman eksik etmeyen, benim için her türlü fedakârlıkta bulunan, beni yüreklendiren ve bana güvenen sevgili anneme, sevgili babama, sevgili ablama ve sevgili kardeşime çok teşekkür ederim.

ÖZET

İnsan makine etkileşimi konusundaki çalışmalar son yıllarda geniş ilgi ve uygulama alanları bulmuştur. Bu uygulama alanlarından bir tanesi insanların makineler vasıtasıyla birbirleriyle iletişimini kolaylaştıran sistemlerdir. İnsan hareketlerini doğru şekilde anlayabilmek için öncelikle insanın çevresiyle iletişime geçerken kullandığı öğeler dikkatle incelenmelidir. İnsan hareketinin önemli bir parçasını oluşturan el hareketlerini kavrayabilmek, insan makine etkileşimini kolaylaştırmak için önemli bir adımdır.

İşaret dili işitme engellilerin birbirleriyle iletişim kurmasını sağlayan, el hareketleri ve yüz ifadesi gibi öğelerden oluşan görsel bir dildir. İşarete anlam katmada daha etkin rol oynayan el hareketlerinin meydana getirdiği özellikler elin izlediği yol (global özellikler) ve elin aldığı biçim (yerel özellikler) olmak üzere ikiye ayrılır. İşaret dilini oluşturan el hareketlerini tanıyabilen bir sistem tasarlayabilmek, işitme engellilerin diğer kişilerle kolayca iletişimini sağlayacağı için son 15 yılda bu konudaki araştırmalar artmıştır. Tasarlanan tanıma sistemleri temelde elektronik eldiven tabanlı sistemler ve görüntü tabanlı sistemler olmak üzere ikiye ayrılmaktadır.

Günümüze kadar çeşitli ülkelerin işaret dilleri için tanıma sistemleri tasarlanmasına rağmen Türk İşaret Dili (TİD) tanıma konusunda doyurucu bir çalışma yoktur. Bu tez çalışmasında, TİD hareketlerini tanıyabilen bir sistem tanıtılacaktır. Geliştirilen sistem iki aşamalı olup, işaret edilen hareket önce global özelliklerine göre saklı Markov model kullanılarak gruplandırılmakta ve sonra belirlenen gruptaki üyelerden k en yakın komşu algoritması kullanılarak en uygun olanı seçilmektedir. TİD hareketlerine ait bir görüntü arşivi olmadığı için, hareketler sistemi tasarlayan kişi tarafından öğrenilip gerçekleştirilmiş ve böylece kişiye bağımlı bir sistem tasarlanabilmiştir.

TİD'ne ait 172 hareket kullanılarak sistemin başarısı test edilmiş ve geliştirilen sistem ile kabul edilebilir oranlarda tanıma yapmanın mümkün olduğu görülmüştür.

Anahtar kelimeler: Türk İşaret Dili Tanıma, El Hareketi Tanıma, Saklı Markov Model, Görüntüden Hareket Analizi, İnsan Makine Etkileşimi.

ABSTRACT

Researchers study on designing intelligent interfaces to provide more comfortable interaction between human and machine. The interfaces generally analyze body motion and decide what to do according to motion. Gestures, which are generated by hands and arms, are special case of body motion. If a system that recognizes sign language (SL) gestures can be designed, the deaf-dumb people will have the chance of communicating with the others.

SLs are characterized by manual and non-manual features. The non-manual features include facial expressions, and the manual features can be divided into 2 main categories: global features (hand movement) and local features (hand shape). Studies about SL recognition have begun to be seen in the literature for 15 years, and to date many systems have been designed. These systems can be categorized into two main groups: data glove based systems and vision based systems.

In the literature, there is not a comprehensive study on recognition of Turkish Sign Language (TSL). The system proposed in the thesis recognizes TSL gestures in two steps. First, the global feature group of performed gesture is determined by using Hidden Markov Models, and then the local features are used with k nearest neighbor algorithm to select the corresponding word from candidates that are previously derived from the global features. Because of having no TSL gesture database, a signer dependent system can be designed.

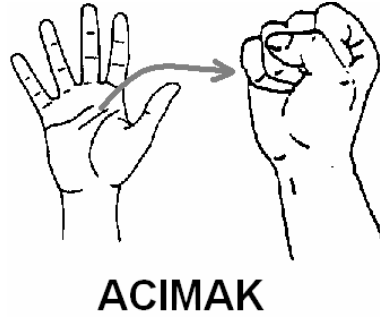
172 signs are used to show that the system can achieve satisfactory results in TSL recognition.

Keywords: Turkish Sign Language Recognition, Hand Gesture Recognition, Hidden Markov Model, Motion Analysis from Video, Human Machine Interaction.

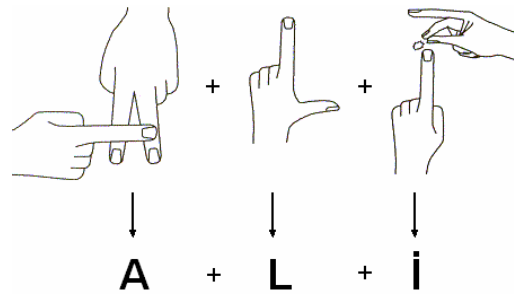
1. GİRİŞ

Son yıllarda, araştırmacılar insan ve bilgisayar arasındaki etkileşimi daha kolay ve daha kullanışlı hale getirebilmek için akıllı arabirimler tasarlamaya çalışmaktadır. Bu arabirimler genellikle vücut hareketlerini analiz etmekte ve ne yapılacağına harekete göre karar vermektedir. El ve kol kullanılarak yapılan hareketler vücut hareketlerinin özel bir parçasıdır ve bu tarz hareketleri tanıyabilen sistemler tasarlamak insan bilgisayar etkileşimi açısından çok önemli bir adımdır.

İşaret dili, işitme engellilerin kendi aralarında iletişim kurmasını sağlayan görsel bir dildir. Elin şekli, elin hareketi, vücut hareketi ve yüz ifadesi işaret dilini meydana getiren öğelerdir. İşaret dilindeki her hareket konuşma dilindeki bir kelimeye yada alfabedeki bir harfe denk gelir. Sık kullanılan kelimeler için hareketler mevcutken (Şekil 1.1), nadir kullanılan kelimeler ile özel isimler için harf harf parmak heceleme yapılır (Şekil 1.2).



Şekil 1.1 Hareket ifadeli kelime [1]



Şekil 1.2 Parmak hecelemeli kelime

İşaret dili konuşma dilinden etkilense de kendine has bir dilbilgisi yapısı vardır [1]. Her ülkenin işaret dili kendine özgüdür, örneğin İngiliz İşaret Dili ile Amerikan İşaret Dili birbirinden farklıdır (Hai ve Sutherland, 2004). Ülkemizdeki işitme engellilerin kendi aralarında iletişim kurmak için kullandıkları Türk İşaret Dili (TİD) sık kullanılan kelimeler ile alfabedeki 29 harfe karşı gelen hareketlerden oluşmaktadır. Şayet, işaret diline ait hareketleri

tanıyabilen bir sistem tasarlanabilirse, işitme engellilerin diğer insanlarla da kolayca iletişim kurmasına olanak sağlanacaktır.

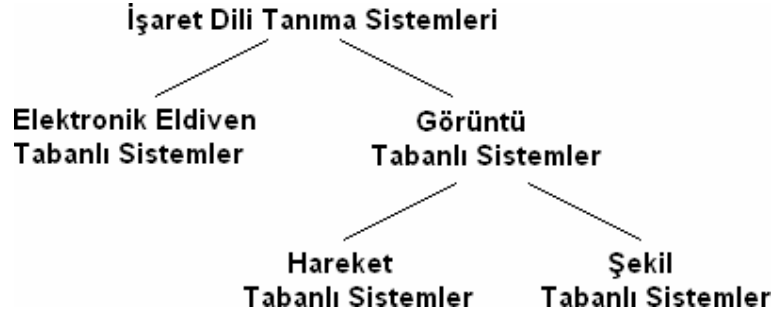
İşaret dilindeki hareketlerin karakteristiğini ele ait özellikler ile yüz ifadesi ve vücut duruşu belirler (Grobel ve Assan, 1997). Elin işaret diline kattığı özellikler global ve yerel (local) özellikler olmak üzere ikiye ayrılır. Hareket sırasında ellerin izlediği yörünge hareketin global özelliklerini meydana getirirken, elin aldığı biçim de hareketin yerel özelliklerini oluşturur. Ticari anlamda kullanılabilir bir işaret dili tanıma sistemi tasarlanabilmesi için hem elin global özelliklerinin hem de elin şekli ve yönü gibi yerel özelliklerinin beraber değerlendirilmesi gerekmektedir (Imagawa vd., 2000).

Bu çalışmada standart bir bilgisayar ile CCD (Charge Coupled Device) kameradan oluşan ve çeşitli Türk İşaret Dili hareketlerini tanıyabilen bir sistem tasarlanmıştır. Bu sistemde, işaretleri gerçekleştiren kişinin elektronik yada renkli eldivenler kullanmasına gerek yoktur. Ten renginin özelliklerinden (Hsu vd., 2002) faydalanarak ellerin ve yüzün yeri bulunmaktadır. Görüntüyü oluşturan çerçevelerden çıkarılan özellikler ile saklı Markov modelleri (HMMs, Hidden Markov Models) ve k en yakın komşu (kNN, k Nearest Neighbor) algoritması kullanılarak gerçek zamanlı çalışan bir tanıma sistemi tasarlanmıştır.

TİD hareketlerini barındıran bir veritabanı olmadığından ve şimdiye kadar bu konuda doyurucu bir çalışma yapılmadığından ötürü, hareketler sistemi tasarlayan kişi tarafından öğrenilmiş ve gerçekleştirilmiştir. Bu nedenle kullanıcıya bağımlı bir sistem geliştirilebilmiştir.

1.1 İşaret Dili Tanıma Sistemleri

İşaret dili tanıma konusundaki çalışmalar, makine öğrenmesi sistemlerinde son 15 yılda görülmeye başlanmış ve günümüze kadar birçok sistem tasarlanmıştır. Bu sistemler temelde elektronik eldiven tabanlı ve görüntü tabanlı sistemler olmak üzere ikiye ayrılmaktadır. Önceleri, tasarlanan sistemlerin çoğu alıcılara dolu elektronik eldivenler barındıran pahalı sistemlerdir (Takahashi ve Kishino, 1991; Liang ve Ouhyoung, 1996). İşareti gerçekleştiren kişinin elektronik eldiveni giymesi gerektiği için kullanımı kolay olmayan, kişinin hareketlerini kısıtlayan özelliklere sahiptirler. Elektronik eldiven tabanlı sistemlerin dezavantajlarından dolayı, son 10 yılda görüntü tabanlı sistemler üzerine çalışmalar artmıştır. Görüntü tabanlı sistemler kendi içinde, hareket tabanlı sistemler (Starner, 1995) ve şekil tabanlı sistemler (Imagawa vd., 2000) olmak üzere temelde iki kategoriye ayrılmaktadır (Şekil 1.3).



Şekil 1.3 İşaret dili tanıma sistemleri

Hareket tabanlı sistemler elin işaret sırasındaki hareketi, şekil tabanlı sistemler ise elin herhangi bir andaki duruşu ile ilgilenir. Ancak çok sayıda hareket tanıyabilen güvenilir ve kullanılabilir sistemlerin tasarlanabilmesi için her iki yaklaşımın beraber uygulanması gerekmektedir.

Görüntüden işaret dili tanıma konusunda yapılan ilk çalışmalar arasında en dikkat çeken, Starner'ın 1995 yılında gerçekleştirdiği Amerikan İşaret Dili tanıma sistemidir. Bir kamera ve standart bir bilgisayarın kullanıldığı sistemde tanıma işlemi saklı Markov model ile sağlanır. Starner, elin duruşunun ayrıntılı şekilde incelenmesinin gerekli olmadığını düşünmüş ve bunun yerine elin şeklinin, yönünün ve yörünge bilgisinin kaba ifadesini kullanmıştır. Elleri takip etmenin kolay olabilmesi için işareti gerçekleştiren kişi sağ eline sarı, sol eline kavuniçi renkte sıradan eldivenler giymiştir. Kullanıcıya bağımlı olan sistemde izleme, analiz etme ve tanıma işlemlerinin tamamı gerçek zamanlı olarak yapılmıştır. Zamir, fiil, sıfat ve isimden oluşan 40 kelimeden türetilen 494 cümle, 395 tanesi eğitim, 99 tanesi test amaçlarıyla kullanılmıştır. Yapılan testler sonunda bazı kelimeler için sadece hareket bilgisinin yeterli olamadığı ve parmakların duruşunun da önem kazandığı belirlenmiştir. Kelimelerin türü ve cümlelerin kaç kelimeden oluştuğu bilinerek yapılan testlerde %99 başarı, cümle yapısının ve sayısının bilinmediği testlerde ise %91 oranında başarı elde edilmiştir. Sistem eldiven kullanılmadan da test edilmiş, ancak gerçek zaman performansı yakalanamamış ve başarı oranı %75 dolaylarına düşmüştür. Starner'ın bu çalışması sonraki yıllarda yapılan işaret dili tanıma sistemlerine ışık tutmuştur.

1996 yılında Yoshino vd. tarafından tasarlanan Japonca İşaret Dili tanıma sistemi, üzerinde çeşitli renklerde bantlar bulunan bir eldiven kullanılarak gerçekleştirilmiştir. Sistem hem parmak heceli kelimeleri hem de hareket ifadeli kelimeleri tanıyabilmektedir. Doğrudan elin duruşunu incelemek yerine, elin şekline göre gözüken bantların renk kombinasyonu ve yoğunluğu kullanılmıştır. Bu nedenle eldivenin üzerine renkli bantlar yerleştirilirken, elin

farklı duruşlarının aynı renk kombinasyonunu oluşturmamasına dikkat edilmiştir.

Başlarda yapılan çalışmaların çoğunda kısıtlı sayıda hareketi tanıyabilen sistemler tasarlanabilmiştir. Doğal olarak, ticari anlamda kullanılabilecek bir sistemin çok daha fazla sayıda hareketi tanınması gerekmektedir.

1997 yılında Grobel ve Assan'ın CCD kamera kullanarak tasarladıkları kişiye bağlı sistem Hollanda İşaret Dili'ne ait 262 hareketi %94 başarı tanımıştır. Grobel ve Assan (1997), Yoshino vd.'nin (1996) kullandığı eldivene benzer eldivenler kullanmıştır, ancak tek bir eldiven kullanmak yerine, her iki el içinde eldiven kullanılmıştır. İşaretleri yapan kişi sağ eline, parmakları, avuç içini ve elin arkasını belli edecek şekilde üzerinde 7 farklı renk bulunan eldiven ve sol eline üzerinde 8 farklı renk bulunan eldiven giymiştir. Ellerin hareket sırasında izledikleri rota, elin şekli, yönü gibi özelliklerden faydalanan Grobel ve Assan saklı Markov modelleri kullanarak tanıma yapmışlardır.

Imagawa vd.'nin (2000) tasarladıkları sistem, ten renginden faydalananak elleri bulmakta ve iki aşamalı bir tanıma işlemi önermektedir. Öncelikle tanınmaya çalışılan hareketin global özelliklerine göre mümkün olan aday kelimeler seçilmekte ve sonrasında seçilen bu işaretlerden yerel özellikler kullanılarak işaret edilen kelime belirlenmektedir. Sistemin başarısını ölçmek için postanedeki bir konuşmada geçebilecek Japonca İşaret Dili'ne ait 160 kelime seçilmiştir. Tasarlanan sistemin kamusal bir alanda ve belli bir amaca yönelik kullanılacağı düşünülerek, bir konu ile ilgili kelimeleri tanıyan sistemler tasarlamak doğru gözükmemektedir. Yapılan testler sonunda yakın global özelliğe sahip kelimeleri yerel özelliklerinden faydalananak tanıma başarısı %90 civarında gerçekleşmiştir.

Bugüne kadar tasarlanan sistemlere bakıldığında, Amerikan (Starner, 1995), Japon (Yoshino vd., 1996), Hollanda (Grobel ve Assan, 1997), Çin (Chen vd., 2003), Arap (Mohandes vd., 2004), Kore (Chang ve Pengwu, 2004), İrlanda (Hai ve Sutherland, 2004) İşaret Dilleri hakkında çalışmalar olmasına rağmen Türk İşaret Dili tanıma konusunda doyurucu bir çalışma yoktur.

1.2 Tez Çalışmasında Tasarlanan Sistem

Bu çalışmanın amacı, bir adet CCD kameradan alınan görüntünün standart bir bilgisayar yardımıyla işlenmesi sonucu Türk İşaret Dili'ne ait çeşitli hareketleri tanımdır. Öncelikle ellerin ve yüzün yeri Hsu vd.'nin (2002) önerdiği renkli resimlerde ten rengi tanıma yöntemi ile bulunmaktadır. Daha sonra görüntüdeki yüz, sağ ve sol el, bağlantılı bileşen etiketleme

(CCL, Connected Component Labeling) yöntemi (Gonzalez ve Woods, 1993) kullanılarak ayırt edilmektedir (Şekil 1.4).



Şekil 1.4 Görüntüde sağ el, sol el ve yüzün ayırt edilmesi

Arka plandan elleri ve yüzü ayırt etmek için ten rengi özelliklerinden faydalanılması nedeniyle, sistem bazı kısıtlamaları beraberinde getirmektedir.

- Eller birbirine temas etmemelidir.
- Eller yüze temas etmemelidir.
- Sağ el, sol elin her zaman sağında olmalıdır.
- Hareket boyunca eller ve yüz kameranın görüş alanı içinde olmalıdır.

Koç Üniversitesi ve Türkiye Bilimler Akademisinin (TÜBA) ortak çalışması sonucu oluşturulan Türk İşaret Dili Sözlüğü'ndeki [1] işaretler incelenmiş ve yukarıda belirtilen kısıtlamalara uygun olarak sistemin tanıyabileceği 172 hareket seçilmiştir. Gerçek zamanlı çalışabilen bir sistem tasarlandığı için, hem sistem hızını yükseltecek hem de tanıma oranını etkilemeyecek şekilde çerçeve oranı ve görüntü çözünürlüğü seçilmiştir.

İşareti gerçekleştiren kişinin yüzü bulunmasına rağmen, bu çalışmanın kapsamında hareket tanıma işleminde doğrudan kullanılmamıştır. Türk İşaret Dili'nde yer alan hareketlerin uzunlukları çok çeşitlidir. Bu nedenle hareketi içeren görüntüyü oluşturan tüm çerçeveleri değerlendirmek yerine, içlerinden sadece dört tanesini seçilmiş ve bunlara anahtar çerçeve adı verilmiştir. Tanıma işlemi iki aşamalı olarak yapılmakta, önce ellerin konum bilgileri saklı Markov model ile değerlendirilip hareketin ait olduğu global özellik grubu belirlenmekte ve sonra da bu gruptaki işaretlerin yerel özelliklerine bakılarak harekete karşı gelen kelime seçilmektedir.

2. TÜRK İŞARET DİLİ

Türk İşaret Dili tarihinin Osmanlı dönemine kadar uzanmasına karşın, TİD hakkında Milli Eğitim Bakanlığı'nın 1995'te yayınladığı görsel kılavuz dışında henüz bir yazılı materyal, arşiv yada sözlük yoktur [1]. Türk İşaret Dili'nin tarihçesiyle ilgili bilgiler, işaret dili görsel bir dil olduğu ve dolayısıyla kağıda geçirilmesi zor olduğu için oldukça yetersiz kalmaktadır. Türk tarihinde işaret dilinin varlığı ve eğitimde kullanımıyla ilgili arşivler Osmanlıca olduğu için bu konuda yoğun bir arşiv çalışması gerekmektedir. Şu ana kadar edinilen bütün bilgiler en azından Osmanlı işaret dilinin batıda kullanılan işaret dilleriyle bir ilişkisi olmadan geliştiği ve bu açıdan oldukça özgün bir işaret dili olduğudur [1].

TİD'deki toplam işaret sayısı bilinmemekle birlikte yerli ve yabancı araştırmacılar tarafından incelemeler devam etmektedir. Türk İşaret Dili sık kullanılan kelimelere ve alfabedeki 29 harfe karşı gelen hareketlerden oluşur. Kelime ekleri için özel bir yapı bulunmamakla birlikte, vücut yönü yada parmak işaretleri kullanılarak işaret edilen kelimeye ek anlamlar katılabilmektedir (Şekil 2.1).



Şekil 2.1 TİD'de kelime ekleri vücut hareketleri yada parmakla yön belirtilerek verilir

TİD hakkında yazılı kaynak eksikliği dolayısıyla literatürde Türk İşaret Dili tanıma sistemleri konusunda doyurucu bir çalışma yoktur. Araştırmacılar için önerilebilecek kaynaklar Koç Üniversitesi ve TÜBA'nın ortaklaşa oluşturduğu Türk işaret Dili Sözlüğü [1] ile Milli Eğitim Bakanlığının 1995 yılında yayınladığı Türk İşaret Dili Kılavuzu'dur.

2.1 Türk İşaret Dili Sözlüğü

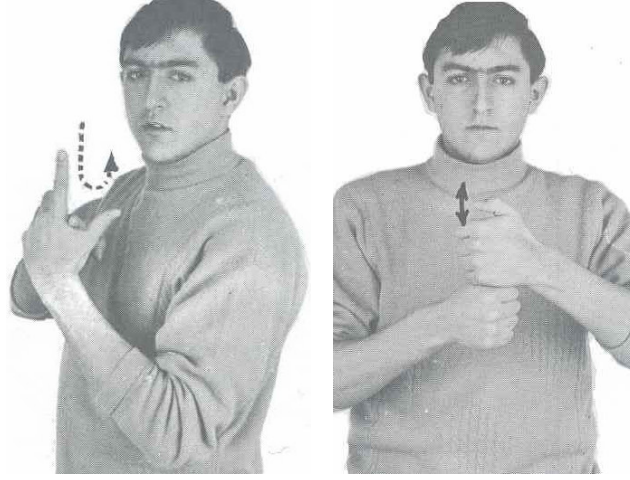
Bu tez çalışması sürecince, Koç Üniversitesi ve Türkiye Bilimler Akademisi'nin birilikte yürüttüğü üç yıllık araştırma sürecinin ürünü olan TİD Sözlüğü'nden [1] faydalanılmıştır. Tez çalışmasının temelini oluşturan hareketlerin tamamı bu sözlükten seçildiği için, burada sözlüğün nasıl meydana geldiğinden kısaca bahsedilecektir.

Sözlük oluşturulurken ilk aşamada, İstanbul'da işaret dili kullanan 20 kişinin günlük sohbetleri Beyoğlu ve Aksaray İşitme Engelliler Dernekleri'nde kameraya alınmıştır. Toplam süresi yaklaşık 10 saati bulan bu veriler, dilbilimsel ve psikolojik yöntemlerle toplanan diğer verilerle birleştirilip, hem TİD hem de Türkçe bilen iki ayrı tercüman tarafından Türkçeye çevrilmiştir. Tercüme edilen işaret grupları daha sonra kelimelere bölünüp benzerlikleri karşılaştırılmıştır [1]. Son olarak bu sohbetlerde kullanılan işaretler, erken yaşta işaret dili öğrenmiş 3 ayrı işitme engelli ile sözlük için tekrar kameraya çekilmiş ve yine onlar tarafından onaylanmıştır [1].

İnternet ortamında bulunan kapsamlı ilk kaynak olan Türk İşaret Dili Sözlüğü'nde sözlük ve kavramsal kategoriler altında toplam yaklaşık 750 işaret bulunmaktadır.

2.2 Türk İşaret Dili Kılavuzu

Türk İşaret Dili hakkında yayınlanan tek Türkçe yazılı kaynak Milli Eğitim Bakanlığı Özel Eğitim Rehberlik ve Danışma Hizmetleri Genel Müdürlüğü'nün 1995 yılında çıkardığı görsel kılavuzdur. Kılavuz işitme engelli bireylerin günlük ihtiyaçlarını karşılayabilecek yaklaşık 2000 adet kelimeden oluşmaktadır. Kılavuzun kolay kullanılabilmesi için kelimeler konu başlıklarına göre gruplandırılarak, bölümler halinde sunulmuştur. Kelimelere ve alfabedeki harflere karşı gelen ifadeler, resimler ve hareketi belirten işaret okları yardımıyla anlatılmaktadır (Şekil 2.2).



Şekil 2.2 J ve Ğ harflerinin TİD karşılıkları (MEB, 1995)

Resimlerin üzerindeki hareket okları uluslararası GESTUNA normlarına uygun olarak verilmiş (Şekil 2.3) ve her hareket okunun ve işaretinin anlamı açıklanmıştır.

	Devamlı düz hareketler		Kesik kesik devamlı hareketler
	Aynı yöne düz hareketin tekrarı		İki yöne düz hareketin tekrarı
	Aynı yöne paralel hareketler		Ters yöne paralel hareketler
	İki parmağın harekete açık başlayıp kapalı bitirmesi		Küçük kıpırdatma hareketleri
	Bütün parmakların harekete kapalı başlayıp açık bitirmesi		Dairesel hareketler
	Bütün parmakların harekete açık başlayıp kapalı bitirmesi		Kavisli hareketler
	İki veya daha çok parmakla kalın bir hat çizimi		Parmaklarla yapılan dalgalanma hareketleri
	Hareketin ani bir duruşla bitirilmesi		Parmaklarla yapılan spiral hareketler
	Parmağın vücuda tam teması		Beş parmak ile herhangi bir parmağın sürtünmesi veya fiske atma hareketi

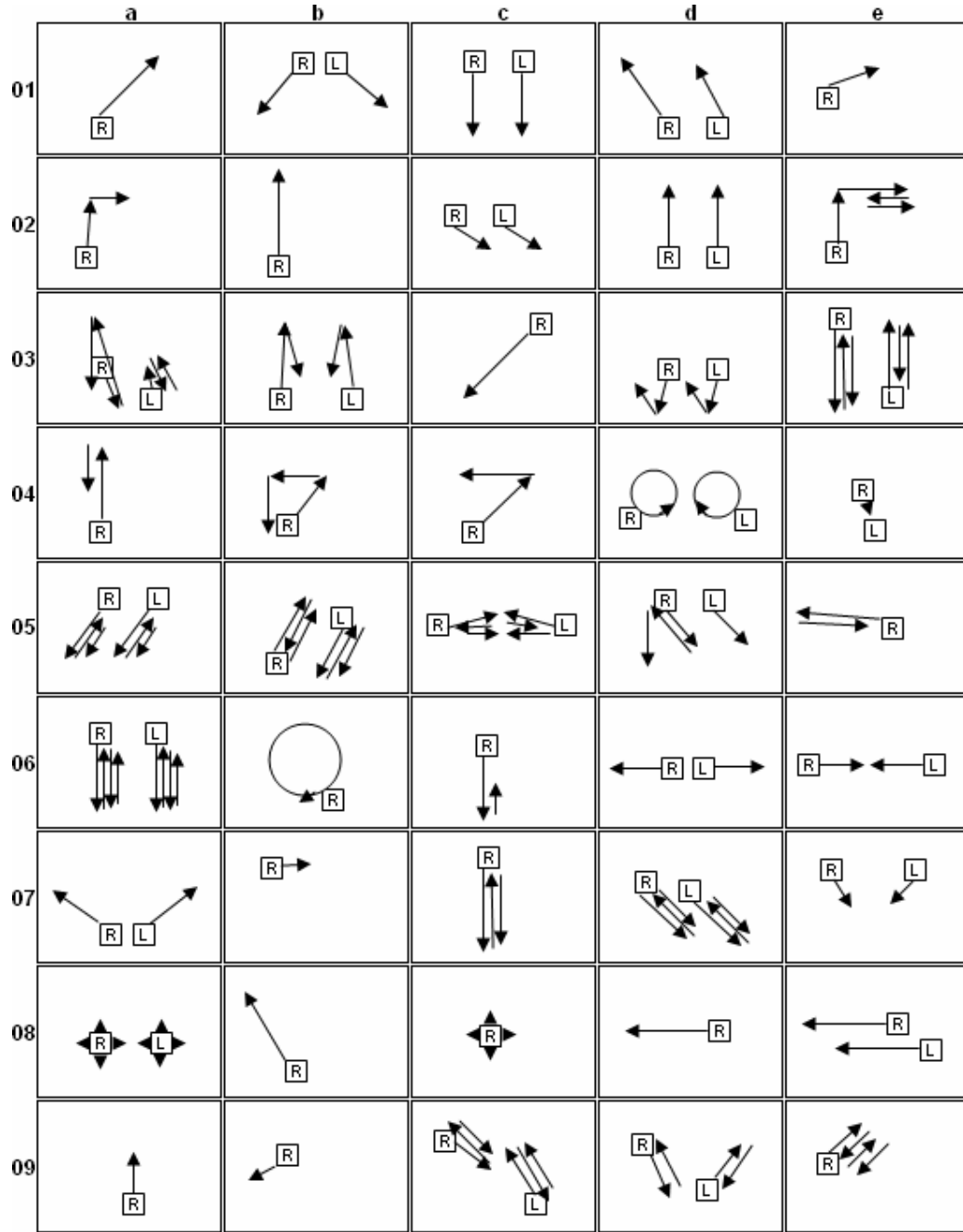
Şekil 2.3 Kılavuzda kullanılan hareket oklarının anlamları

2.3 Tez Çalışmasında Kullanılan Kelimeler

Tez çalışmasında tanıma için kullanılan işaretler Bölüm 1.2’de anlatılan sistem kısıtlamaları göz önüne alınarak seçilmiştir. Seçilen 172 tane işaret, hareket sırasında ellerin izlediği yörüngeye göre 45 farklı alt gruba (a1,..., a9, b1,..., b9, c1,...,c9, d1,..., d9, e1,..., e9) ayrılmıştır. Bu gruplar hareketlerin global özelliklerine göre belirlendiği için global özellik grupları olarak adlandırılmıştır. Şekil 2.4’te R sağ ele, L sol ele karşı gelmek üzere ellerin izlediği yörüngeler verilmiştir.

Buradan elde edilen veriler ışığı altında sadece global özelliklere göre sınıflandırma yapan bir sistem tasarlamamanın yetersiz olduğu anlaşılmaktadır. Çünkü 172 hareketin ancak 45 farklı alt grup meydana getirmesi, sistemin sadece global özellikleri kullanarak tanıyabileceği işaret sayısını 45 ile sınırlı kalacağı anlamına gelmektedir.

Sistemin tanıyabildiği işaretler ile işaretlerin ait olduğu global özellik grupları Çizelge 2.1’de verilmiştir. Çizelge 2.1’deki gruplar incelendiğinde birçok kelimenin aynı global özellik grubuna ait olduğu görülmektedir. Örneğin Şekil 2.4’te a8 olarak verilen global özellik grubundan 26 adet kelime vardır. Bu durum, sistemin daha fazla sayıda kelime tanıyabilmesini sağlamak için sınıflandırma yaparken global özelliklerin yanında yerel özelliklerin de kullanılması gerekliliğini göstermektedir.



Şekil 2.4 Hareket sırasında ellerin izlediği yörüngeler ve global özellik grupları

Çizelge 2.1 Sistemin tanıdığı kelimeler ve global özellik grupları

Kelime	Grup	Kelime	Grup	Kelime	Grup	Kelime	Grup
acımak	a1	değil	d2	kazak	a8	seçmek	e9
açık	b1	deli	b2	kısa	e6	seramik	a8
ağır	c1	devam	a5	kış	a8	serbest	b1
Aksaray	a1	dinlenmek	c1	kızartmak	c8	sert	a8
alışveriş	c9	disko	b5	kızmak	b6	sevinçli	a8
almak	e1	doğmak	d3	kiraz	b2	sevinmek	a8
anlamak	b2	durmak	b1	kirli	a4	sevmek	c3
anne	a2	eğitmek	a5	kitap	e3	sıfır	c8
araba	e3	eğlence	c5	korkmak	a1	sıkılmak	a9
arka	b2	Eskişehir	a8	kova	b2	sınav	b1
atmak	c2	fabrika	e3	kovmak	c8	soğuk	a8
aynı	a8	fazla	c8	kötü	c8	sonra	b8
ayran	b2	fena	c8	kurbağa	c7	subay	a1
az	a9	Fenerbahçe	c8	küçük	e6	şaka	e4
bankacı	a3	fırıncı	d5	limon	b2	şampanya	e3
başka	e2	futbolcu	d5	lunapark	e3	şişman	a8
başkan	b2	gibi	a8	marangoz	d7	talebe	c8
bebek	b2	gidiş-geliş	e5	masa	d6	tam	d8
bebek arabası	b3	giyinmek	a6	merak	c8	taşınmak	e8
belediye otobüsü	c3	göndermek	d1	mimar	d9	tava	c8
bezelye	a9	hafif	a8	minibüs	c8	taze	a8
bırakmak	c1	hata	b6	motosiklet	d6	tembel	a8
bilgisayar	a8	hava	a8	muhasabe	e7	top	a8
bira	b2	hazır	d2	ne zaman	b2	toprak	c8
biraz	c8	hevesli	d4	nefes	c8	tren	d7
bisiklet	e3	hizmetli	c5	nerede	a6	tuvalet	c8
borç	b2	ikiz	c6	neyse	a9	uyumak	c2
boş	b2	incelemek	e3	normal	b1	uzun	d6
bozuk	b1	ip	d6	oda	a8	üşümek	a8
börek	a8	istemek	c8	olmak	c1	üzgün	a9
bugün	a9	iş	d4	oyun	a8	üzülmek	c8
büyümek	a4	işitmek	b7	oyuncakçı	d5	var	c8
canlı	b4	iyi	c8	öğretmen	c8	vermek	b9
cansız	c4	kabahat	e4	önemli	c8	vicdansız	b4
cin	b1	kabul etmek	c1	pahalı	c8	virgül	b9
cömert	c2	kalabalık	a8	para	c8	viski	b2
çadır	b1	kalkmak	d2	pilav	b2	yakın	c8
çağırarak	a4	kapalı	e6	pişirmek	e3	yastık	a8
çalışmak	d4	karpuz	a7	radio	a8	yer	c1
çay	e4	kavanoz	e4	rahat	a8	yumuşak	a8
çingirak	a9	kavga	e3	reklam	e3	zaman	b2
çocuk	b2	kavun	d6	saat	a9	zayıf	b2
çok	a9	kayıp	d6	sandal	d4	zor	b2

Şekil 2.4'te verilen 45 global özellik grubunun 25 tanesi iki elle 20 tanesi ise tek elle gerçekleştirilen işaretlere aittir ve her grupta farklı sayılarda işaretler mevcuttur. Çizelge 2.2'de gruplardaki kelime sayıları verilmiştir.

Çizelge 2.2 Global özellik gruplarındaki kelime sayıları

Global Özellik Grubu	a1	a2	a3	a4	a5	a6	a7	a8	a9
Gruba Ait Kelime Sayısı	4	1	1	3	2	2	1	26	9
Global Özellik Grubu	b1	b2	b3	b4	b5	b6	b7	b8	b9
Gruba Ait Kelime Sayısı	8	19	1	2	1	2	1	1	2
Global Özellik Grubu	c1	c2	c3	c4	c5	c6	c7	c8	c9
Gruba Ait Kelime Sayısı	6	3	2	1	2	1	1	24	1
Global Özellik Grubu	d1	d2	d3	d4	d5	d6	d7	d8	d9
Gruba Ait Kelime Sayısı	1	3	1	4	3	6	2	1	1
Global Özellik Grubu	e1	e2	e3	e4	e5	e6	e7	e8	e9
Gruba Ait Kelime Sayısı	1	1	10	4	1	3	1	1	1

Özetlemek gerekirse, tez çalışmasında tasarlanan Türk İşaret Dili Tanıma Sistemi 172 hareketi tanıyabilmektedir ve bu hareketlerin 78 tanesi tek elle, 94 tanesi de iki elle gerçekleştirilen işaretlerdir.

3. MARKOV MODELLERİNİN TEMEL ÖZELLİKLERİ

Hangi sınıfa ait olduğu bulunmaya çalışılan örneği oluşturan veri dizisinin elemanlarının, birbirinden bağımsız olduğunu varsaymak sınıflandırma işlemini kolaylaştırır. Böylece her bir elemana ait maksimum olabilirlik (Maximum likelihood) derecesini hesaplayıp bunları çarparak, o örneğin hangi sınıfa ait olduğu bulunabilir. Ancak, bu varsayım veri elemanlarının dizilimlerinin birbirine bağımlı olduğu durumlarda kullanılamaz (Alpaydın, 2004). Örneğin, İngilizcede t harfinden sonra h gelme olasılığı çok yüksekken, x gelme olasılığı çok düşüktür. Benzer şekilde, bir DNA dizisini oluşturan elemanların dizi içindeki yerleri rasgele değildir.

Bu bölümde, bu tür özelliklere sahip örneklerin sınıflandırılmasında başarılı sonuçlar veren saklı Markov modellerinin temel özellikleri anlatılacaktır.

3.1 Ayırık Markov Prosesleri

Her zaman, N farklı $(S_1, S_2, \dots, S_N)$ durumdan birinde olan bir sistem düşünüldüğünde, sistemin t anındaki durumu q_t ile sembolize edilmek üzere, $q_t = S_i$ ifadesi, sistemin t anında S_i durumunda olduğunu gösterir. (3.1)'de verilen sabit aralıklı ayırık zaman diliminde sistemin bir durumdan diğer bir duruma geçmesi olasılığı önceki durumların ne olduğuna bağlıdır:

$$P(q_t = S_j \mid q_{t-1} = S_i, q_{t-2} = S_k, \Lambda) \quad (3.1)$$

Özel bir durum olan birinci derece Markov modelinde ise, $t+1$ anındaki durumun sadece t anındaki duruma bağlı olduğu varsayılır (3.2).

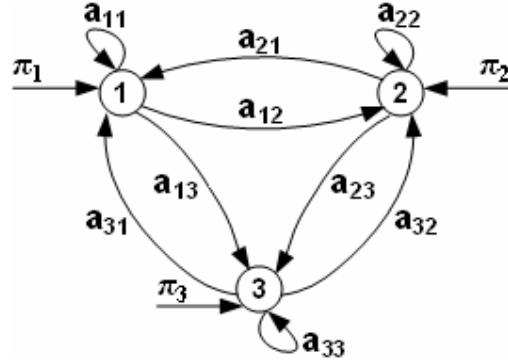
$$P(q_{t+1} = S_j \mid q_t = S_i, q_{t-1} = S_k, \Lambda) = P(q_{t+1} = S_j \mid q_t = S_i) \quad (3.2)$$

Modeli basitleştirmek için yapılan diğer bir varsayım da, (3.3)'de ifade edilen durum geçiş olasılıklarının zamandan bağımsız olduğudur.

$$a_{ij} \equiv P(q_{t+1} = S_j \mid q_t = S_i), \quad a_{ij} \geq 0 \text{ ve } \sum_{j=1}^N a_{ij} = 1 \quad (3.3)$$

Yukarıda belirtilen ikinci varsayıma dayanarak S_i durumundan S_j durumuna geçiş olasılığının her zaman aynı olduğu söylenebilir. Yani sistemin hangi durumlarda olduğunu gösteren gözlem dizisinde, bu geçişin dizinin hangi sırasında gerçekleştiği ile a_{ij} arasında bir

bağlantı yoktur. Durum geçiş olasılıkları $N \times N$ boyutundaki $\mathbf{A}=[a_{ij}]$ matrisinde verilir. (3.3)'deki koşuldan da anlaşılacağı üzere $\mathbf{A}$ matrisinin her satırının toplamı bire eşittir. Bu model Şekil 3.1'de gösterilen bir olasılıksal makine (Stochastic automaton) olarak düşünülebilir.



Şekil 3.1 3 durumdan oluşan bir Markov model

Burada tek özel hal sistemin ilk anda hangi durumda olacağıdır, bu da (3.4)'de verilen π_i ile ifade edilir.

$$\pi_i \equiv P(q_1 = S_i), \quad \sum_{i=1}^N \pi_i = 1 \quad (3.4)$$

Bu olasılıklar, N elemanlı ve elemanlarının toplamı 1'e eşit olan $\Pi=[\pi_i]$ ilk durum olasılıkları vektörü ile verilir. Gözlemlenebilir Markov modelde, durumlar, yani sistemin hangi durumda olduğu gözlenebilir. Her t anında q_t 'nin ne olduğu, bir başka ifade ile sistemin hangi durumdan hangi duruma geçtiği gözlem dizisi olarak verilir. Gözlem dizisini, sistemin olabileceği durumlar oluşturur. Burada belirtilmesi gereken bir nokta da şudur ki sistemin bulunduğu her durum gözlemlenebilir fiziksel bir olayı (örn:Yazı-Tura atma) ifade eder. Durum dizisine denk düşen gözlem dizisi $O=o_1o_2 \wedge o_T = Q=q_1q_2 \wedge q_T$ ile verilirse, modelde bu gözlemin gerçekleşme olasılığı,

$$\begin{aligned} P(O=Q | A, \Pi) &= P(q_1)P(q_2 | q_1)P(q_3 | q_2) \wedge P(q_T | q_{T-1}) \\ P(O=Q | A, \Pi) &= P(q_1) \prod_{t=2}^T P(q_t | q_{t-1}) \\ P(O=Q | A, \Pi) &= \pi_{q_1} a_{q_1 q_2} \wedge a_{q_{T-1} q_T} \end{aligned} \quad (3.5)$$

olarak hesaplanır. π_{q_1} ilk durumun q_1 olma olasılığını, $a_{q_1 q_2}$ sistemin q_1 durumundan q_2 durumuna geçme olasılığını verir, geri kalan bu şekilde devam eder. Tüm gözlem dizisinin

meydana gelme olasılığını bulabilmek için bu olasılıkların çarpılması yeterlidir. Sistem parametreleri Π vektörü ile A matrisi verildiğinde, rasgele gözlem dizileri (=durum dizileri) oluşturabilir. Tabii olarak, burada önemli olan, sistemin ürettiği gözlem dizileri verildiğinde sistem parametrelerini hesaplayabilmektedir.

Modeli meydana getiren parametrelerin (A ve Π) hesaplanabilmesi için, T uzunluğundaki K tane gözlem dizisinin meydana getirdiği eğitim verisi kullanılabilir. Bu durumda, ilk durum olasılık vektörü

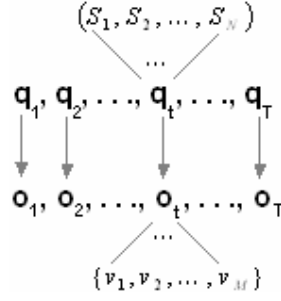
$$\pi = \frac{\sum_{k=1}^K 1(q_1^k = S_i)}{K} \quad (3.6)$$

ile hesaplanır. (3.6)'dan da anlaşıldığı üzere, sistemin S_i durumu ile başlayan gözlem dizisi üretme olasılığının, S_i durumu ile başlayan gözlem dizisi sayısının tüm gözlem dizisi sayısına oranı ile hesaplanabilir. Durum geçiş olasılıkları matrisi elemanı a_{ij} ise, verilen K gözlem dizisi içinde S_i 'den S_j 'ye geçişlerin sayısının S_i 'den tüm durumlara geçiş sayısına oranı ile hesaplanabilir (3.7).

$$a_{ij} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k-1} 1(q_t^k = S_i \text{ ve } q_{t+1}^k = S_j)}{\sum_{k=1}^K \sum_{t=1}^{T_k-1} 1(q_t^k = S_i)} \quad (3.7)$$

3.2 Saklı Markov Model

Gözlemlenebilir Markov prosesinden farklı olarak, saklı Markov modelde sistemin herhangi bir t anında hangi durumda olduğu bilinmez, ancak sistem bir durumda iken bu durumun tetiklediği gözlemi ortaya çıkarır (Şekil 3.2).



Şekil 3.2 Sistem bir durumda iken bu durumun tetiklediği gözlemi ortaya çıkarır

Gözlenebilir Markov modelde sistem S_i durumunda iken, eldeki gözlem sadece sistemin i .durumda olduğudur. Saklı Markov modelde ise, sistem bir durumda iken, M tane farklı gözlemden $\{v_1, v_2, K, v_M\}$ bir tanesi, olasılık fonksiyonuna bağlı olarak elde edilir. Bu olasılık $b_j(m)$ ile gösterilir ve sistemin S_i durumunda iken v_m gözlemini üretme olasılığını verir. Sistemin olabileceği durumlar (S_1, S_2, K, S_N) ile elde edilebilecek gözlemler $\{v_1, v_2, K, v_M\}$ arasındaki bağıntıyı $N \times M$ boyutundaki $\mathbf{B} = [b_j(m)]$ gözlem matrisi verir (3.8).

$$b_j(m) \equiv P(O_t = v_m \mid q_t = S_j), \quad b_j(m) \geq 0 \text{ ve } \sum_{m=1}^M b_j(m) = 1 \quad (3.8)$$

Gözlenebilir Markov model hatırlanırsa, gözlem dizisi, sistemin her t anı için bulunduğu durumdur. Saklı Markov modelde ise gözlem dizisi $O = o_1 o_2 \Lambda o_T$ 'yi, gözlem olasılık matrisindeki değerlere göre her t anı için sistemin ürettiği gözlemler meydana getirir ve durum dizisi Q belirgin değildir. Zaten saklı Markov modelde saklı denilmesinin nedeni de sistemin bulunduğu durumların bilinmemesidir. Bununla birlikte, elde edilen gözlemler sayesinde sistemin hangi durumda olabileceği tahmin edilebilir. Burada bir O gözlem dizisini, çok farklı sayıda Q durum dizisinin meydana getirebileceği aşıkardır, tabii önemli olan bu durum dizilerinden olabilirlik olasılığı en yüksek olanı belirlemektir.

Bir saklı Markov model aşağıdaki elemanlardan oluşur:

- 1) N , modeldeki durum sayısı: $S = \{S_1, S_2, \dots, S_N\}$
- 2) M , elde edilebilecek farklı gözlem sayısı: $V = \{v_1, v_2, \dots, v_M\}$
- 3) A , durum geçiş olasılıkları matrisi: $\mathbf{A} = [a_{ij}]$, $a_{ij} \equiv P(q_{t+1} = S_j | q_t = S_i)$
- 4) B , gözlem olasılıkları matrisi: $\mathbf{B} = [b_j(m)]$, $b_j(m) \equiv P(O_t = v_m | q_t = S_j)$
- 5) Π , ilk durum olasılıkları vektörü: $\mathbf{\Pi} = [\pi_i]$, $\pi_i \equiv P(q_1 = S_i)$

N ve M belirlendikten sonra, sistem parametreleri $\lambda = (\mathbf{A}, \mathbf{B}, \mathbf{\Pi})$ saklı Markov modeli oluşturur. λ verildiğinde, herhangi uzunlukta, herhangi sayıda gözlem dizisi üretilebilir. Doğal olarak burada amaçlanan, eğitim seti verildiğinde parametrelerin (λ) tahmin edilmesidir.

3.3 Saklı Markov Modellerinin Üç Temel Problemi

Bir λ modelinin ürettiği çok sayıda gözlem dizisi verildiğinde çözümlenmesi gereken 3 problem vardır:

- λ verildiğinde, bu modelin gözlem dizisi $O = o_1 o_2 \dots o_T$ 'yi üretme olasılığı, $P(O | \lambda)$.
- λ ve O verildiğinde, bu gözlem dizisini oluşturma ihtimali en yüksek olan durum dizisinin $Q = q_1 q_2 \dots q_T$ bulunması olasılığı, $P(Q | O, \lambda)$.
- Gözlem dizilerinden oluşan eğitim seti $X = \{O^k\}_{k=1}^K$ verildiğinde, bu eğitim setini en yüksek oranda gerçekleyen model parametreleri λ 'nın hesaplanması, $P(X | \lambda)$.

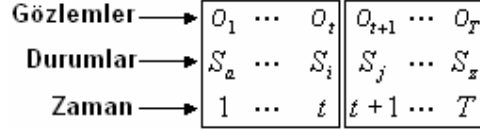
Bu problemler sırasıyla aşağıda incelenecektir.

3.3.1 Tanıma Problemi

λ ile sembolize edilen bir saklı Markov modelin ürettiği gözlem dizisi $O = o_1 o_2 \dots o_T$ verildiğinde, bu gözlem dizisini gerçekleyen N^T tane farklı durum dizisi $Q = q_1 q_2 \dots q_T$ olabilir. Bu nedenle O 'nun elde edilme olasılığı

$$P(O | \lambda) = \sum_{OLASILAR} P(O, Q | \lambda) \quad (3.9)$$

ile hesaplanır. Ancak burada olası durum dizisi sayısı çok fazla olduğu için işlem karmaşıklığı açısından uygun bir çözüm değildir. Bu nedenle $P(O|\lambda)$ hesaplaması için ileri-yön (Forward) ve geri-yön (Backward) algoritması kullanılır. (3.10)'da belirtilen ileri-yön değişkeni $\alpha_t(i)$, t anına kadar $(o_1 \wedge o_t)$ kısmi gözlem dizisinin elde edilme ve t anında durumun S_i olması olasılığıdır. Böylece gözlem dizisi $1K t$ ve $t+1K T$ olmak üzere ikiye ayrılır (Şekil 3.3).



Şekil 3.3 İleri-yön değişkeni gözlem dizisini iki parçaya ayırır

$$\alpha_t(i) \equiv P(o_1 \wedge o_t, q_t = S_i | \lambda) \quad (3.10)$$

İleri-yön algoritması yinelemeli olarak çalışır ve işlem karmaşıklığı $\phi(N^2T)$ 'dir.

İleri-yön algoritma adımları aşağıda verilmektedir:

- İlk değer atama:

$$\begin{aligned} \alpha_1(i) &\equiv P(o_1, q_1 = S_i | \lambda) \\ \alpha_1(i) &= P(o_1 | q_1 = S_i, \lambda) P(q_1 = S_i) \\ \alpha_1(i) &= \pi_i b_i(o_1), \quad t=1 \text{ ve } i=1, K, N \end{aligned} \quad (3.11)$$

- Yineleme:

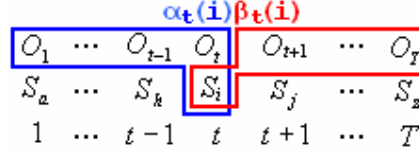
$$\begin{aligned} \alpha_{t+1}(j) &\equiv P(o_1 \wedge o_{t+1}, q_{t+1} = S_j | \lambda) \\ &\quad \wedge \\ &\quad \wedge \\ \alpha_{t+1}(j) &= \left[\sum_{i=1}^N \alpha_t(i) a_{ij} \right] b_j(o_{t+1}), \quad j=1, K, N \end{aligned} \quad (3.12)$$

- Son adım:

$$P(O | \lambda) = \sum_{i=1}^N \alpha_T(i) \quad (3.13)$$

(3.12)'de verilen ifadede, $\alpha_t(i)$ ilk t gözlemi ve sistemin S_i durumunda olmasını, a_{ij} S_i durumundan S_j durumuna geçişi, $b_j(o_{t+1})$ ise $t+1$ anında o_{t+1} gözleminin elde edilmesi olasılığını ifade eder. Her t anı için tüm durumların ileri-yön değerleri hesaplandıktan sonra

(3.13)'de verilen ifade ile gözlemin model tarafından üretilme olasılığı hesaplanır. İleri-yön değişkenine benzer şekilde, geri-yön değişkeni $\beta_t(i)$ 'yi tanımlanarak (Şekil 3.4) aynı olasılık hesaplanabilir.



Şekil 3.4 İleri-yön ve geri-yön değişkenleri

(3.14)'de verilen $\beta_t(i)$, t anında S_i durumunda olma ve $(o_{t+1} \wedge o_T)$ kısmi gözlemini elde etme olasılığını verir.

$$\beta_t(i) \equiv P(o_{t+1} \wedge o_T | q_t = S_i, \lambda) \quad (3.14)$$

Geri-yön algoritmasının adımları aşağıda verilmiştir:

- İlk değer atama:

$$\beta_{T-1}(i) = \sum_{j=1}^N a_{ij} b_j(o_T), \quad i = 1, K, N \quad (3.15)$$

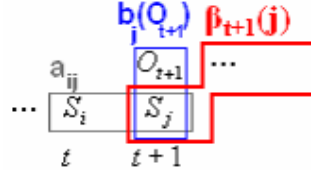
- Yineleme:

$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j), \quad i = 1, K, N \quad (3.16)$$

- Son adım:

$$P(O | \lambda) = \sum_{i=1}^N \beta_1(i) \pi_i b_i(o_1) \quad (3.17)$$

(3.16)'da verilen ifade de a_{ij} , S_i durumunda iken sonraki N olası S_j durumuna geçmeyi ve $b_j(o_{t+1})$, S_j durumunda iken o_{t+1} gözlemini elde etme olasılığını belirtir. $\beta_{t+1}(j)$ ise $t+1$ anından sonraki tüm gözlemlere karşı gelir ve bu bağıntıyı Şekil 3.5'te verildiği gibi ifade edilebilir.



Şekil 3.5 Geri-yön algoritmasında yineleme adımı

3.3.2 Durum Dizisinin Bulunması

Saklı Markov modelin ikinci problemi modelin ürettiği gözlem dizisi $O = o_1 o_2 \wedge o_T$ verildiğinde, bu gözlem dizisini oluşturma ihtimali en yüksek olan durum dizisi $Q = q_1 q_2 \wedge q_T$ 'yi bulma işleminin çözümlenmesidir. (3.18)'de verilen $\gamma_t(i)$, t anında sistemin S_i durumunda olma olasılığını verir.

$$\gamma_t(i) \equiv P(q_t = S_i | O, \lambda)$$

$$\gamma_t(i) = \frac{\alpha_t(i)\beta_t(i)}{\sum_{j=1}^N \alpha_t(j)\beta_t(j)} \quad (3.18)$$

Burada $\alpha_t(i)\beta_t(i)$ çarpımı tüm gözlem dizisini 2 parçaya bölüp tamamını taramaktadır ve bu çarpım normalize edilerek her t anında $\sum_{i=1}^N \gamma_t(i) = 1$ olması sağlanır. Elde edilen olasılıklardan her t anı için olabilirlik olasılığı en yüksek olanı (3.19) alınarak, durum dizisi oluşturulabilir.

$$q_t^* = \arg \max_i \gamma_t(i) \quad (3.19)$$

Ancak sadece bu olasılığa bakmak, $a_{ij} = 0$ olduğu durumda bile t anında sistemin S_i 'den S_j 'ye geçtiği sonucunu doğurabilir. Bu nedenle, hesaplamalara durum geçiş olasılıklarını katan Viterbi algoritması kullanılır (Alpaydın, 2004). (3.20)'de verilen $\delta_t(i)$ değişkeni, t anına kadar olabilirlik derecesi en yüksek olan durum dizisini ifade eder ve t anında S_i durumuna erişilir.

$$\delta_t(i) \equiv \max_{q_1 q_2 \wedge q_{t-1}} p(q_1 q_2 \wedge q_{t-1}, q_t = S_i, o_1 \wedge o_t | \lambda) \quad (3.20)$$

$t = 1$ anından başlanarak, yinelemeli olarak her t anı için en optimum yol bulunur, Viterbi algoritmasının adımları aşağıda verilmiştir.

- İlk değer verme:

$$\delta_1(i) = \pi_i b_i(o_1), \quad i = 1, K, N$$

$$\Psi_1(i) = 0$$

- Yineleme:

$$\delta_t(j) = \max_i \delta_{t-1}(i) a_{ij} \cdot b_j(o_t) \quad i = 1, K, N$$

$$\Psi_t(i) = \arg \max_i \delta_{t-1}(i) a_{ij}$$

- T anı için:

$$path^* = \max_i \delta_T(i), \quad i = 1, K, N$$

$$q_T^* = \arg \max_i \delta_T(i)$$

- Geriye doğru giderek $t = T - 1, K, 1$

$$q_t^* = \Psi_{t+1}(q_{t+1}^*)$$

Viterbi algoritması bölüm 3.3.1’de verilen ileri-yön algoritmasına çok benzer, ancak Viterbi algoritmasında t anında j . durum için hesap yapılırken bu durumun öncesi olabilecek tüm i . durumlarını hesaba katmak yerine sadece en büyük olasılığa sahip olan durum katılır.

3.3.3 Model Parametrelerinin Öğrenilmesi

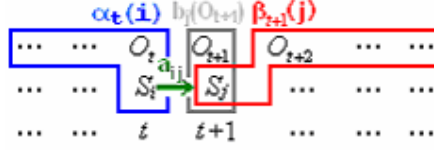
Model parametrelerini $\lambda = (\mathbf{A}, \mathbf{B}, \mathbf{\Pi})$ hesaplayabilmek için verilen eğitim verisinden $\mathbf{X} = \{O^k\}_{k=1}^K$ faydalanılır. Bu aşamada (3.21)’de açık ifadesi verilen, t anında sistemin S_i ve $t + 1$ anında S_j durumunda olmasının olasılığı $\xi_t(i, j)$ tanımlanır (Şekil 3.6).

$$\xi_t(i, j) \equiv P(q_t = S_i, q_{t+1} = S_j | O, \lambda)$$

$$\Lambda$$

$$\Lambda$$

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{\sum_{k=1}^N \sum_{l=1}^N \alpha_t(k) a_{kl} b_l(o_{t+1}) \beta_{t+1}(l)} \quad (3.21)$$

Şekil 3.6 $\xi_t(i, j)$ 'nin hesaplanması

Bir beklenti yükseltme (Expectation maximization) işlemi olan Baum-Welch algoritması kullanılarak model parametreleri hesaplanır.

Baum-Welch algoritması adımları aşağıda verilmiştir:

- Mevcut $\lambda = (\mathbf{A}, \mathbf{B}, \mathbf{\Pi})$ parametreleri kullanılarak $\xi_t(i, j)$ ve $\gamma_t(i)$ değerleri hesaplanır.
- Bir önceki adımda belirlenen $\xi_t(i, j)$ ve $\gamma_t(i)$ değerleri ile yeni parametre takımı $\lambda_{YENI} = (\mathbf{A}, \mathbf{B}, \mathbf{\Pi})$ elde edilir.
- Durma şartı:

$P(O | \lambda_{YENI}) > P(O | \lambda)$ olduğu sürece ilk adıma dönülür.

Eğitim seti olarak $X = \{O^k\}_{k=1}^K$ verildiğinde, model parametrelerinin hesaplanmasında (3.22), (3.23) ve (3.24) kullanılır.

$$a_{ij} = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \xi_t^k(i, j)}{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \gamma_t^k(i)} \quad (3.22)$$

$$b_j(m) = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \gamma_t^k(j) 1(o_t^k = v_m)}{\sum_{k=1}^K \sum_{t=1}^{T_k-1} \gamma_t^k(j)} \quad (3.23)$$

$$\pi_i = \frac{\sum_{k=1}^K \gamma_1^k(i)}{K} \quad (3.24)$$

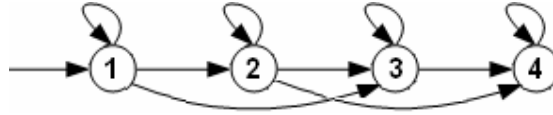
Eğitim setinde K tane eleman olduğu için durma kontrolü yaparken $P(X | \lambda)$ olasılığı (3.25)'de verildiği şekilde hesaplanır.

$$P(X | \lambda) = \prod_{k=1}^K P(O^k | \lambda) \quad (3.25)$$

3.4 Model Seçimi

Saklı Markov model topolojisinin karmaşıklığı eldeki verinin özelliklerine göre seçilmelidir (Alpaydın, 2004). Şekil 3.1’de verilen tüm durumlar arasında karşılıklı geçişlerin olduğu ergodik modelde **A** matrisinin tüm elemanları sıfırdan farklı olacağından hesaplamalar için gereken işlem karmaşıklığı yüksektir. Bu nedenle, uygulamaya göre model seçilirken geçişlerin sınırlı olduğu topolojiler tercih edilir. Böylece ileri-geri yön algoritmasının karmaşıklığı düşürülmüş olur.

Örneğin ses tanıma uygulamalarında, zaman ilerledikçe durum indis numarasının artmasına yada aynı kalmasına olanak sağlayan geçişlere izin veren soldan-sağa saklı Markov model (Şekil3.7) kullanılır.



Şekil 3.7 Dört durumlu soldan-sağa saklı Markov model topolojisi

Bu kısıtlama, özellikleri zaman boyunca değişen ses, görüntü gibi gözlem dizilerinin modellenmesini kolaylaştırmaktadır (Alpaydın, 2004). Soldan-sağa modelde, $j < i$ olduğunda $a_{ij} = 0$ şartını sağlayacak şekilde, durumların kendileri dışında sadece indis numarası daha büyük olan durumlara geçmesine izin verilir ve model adını bu özelliğinden alır.

Modelin karmaşıklığını belirleyen bir diğer nokta da durum sayısı N 'dir. Durum dizisi saklı olduğu için durum sayısı da bilinemez ama modeli eğitmeye başlamadan önce verinin karakteristik özelliğine göre N deneme yanılma yapılarak seçilmelidir.

3.5 Sürekli Gözlemler

Şu ana kadar bahsedilen durumlarda saklı Markov modelin ürettiği gözlemlerin kesikli olduğu, yani üreyen gözlemlerin $\{v_1, v_2, K, v_M\}$ gibi bir kümenin elemanı olduğu varsayımı yapılmıştır. Ancak hareket tanıma uygulamasında seçilecek özelliklerin sürekli değerler alması kuvvetle muhtemeldir. Bu sorunu aşmak için bir vektör kuantalama yöntemi kullanarak değerlerin sürekli halden kesikli hale dönüştürülmesi sağlanır (Alpaydın, 2004) yada saklı Markov modelin parametreleri sürekli değer alabilen gözlemlerle çalışabilir duruma getirilir (Young vd., 2002). Tabii bu durumda saklı Markov modelin temel üç problemine cevap olan ileri-geri yön, Viterbi ve Baum-Welch algoritmalarının da yeniden gözden geçirilmesi gerekir.

3.5.1 Durum Geçiş Olasılıkları

Doğal ve toplumsal olaylarda karşılaşılan rasgele değişkenlerden birçoğunun normal dağılıma uyması istatistiğin merkez limit teoremi ile açıklanabilir (Cerit ve Yüksel, 1998). Bu bağlamda, modelin ürettiği gözlemlerin olasılık yoğunluk fonksiyonlarının (3.26)'da verilen normal dağılıma uygun olduğu varsayılmıştır.

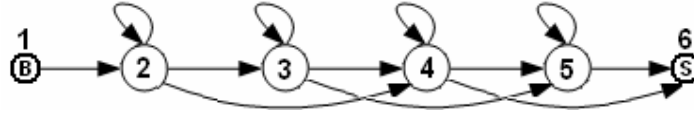
$$N(o; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^n |\Sigma|}} e^{-\frac{1}{2}(o-\mu)^T \Sigma^{-1} (o-\mu)} \quad (3.26)$$

Olabilecek gözlem sayısı (yada gözlemin alabileceği değer) sonsuz sayıda olduğu için, **B** gözlem olasılıkları matrisi yerine (3.27)'de verilen gözlem olasılıkları dağılım fonksiyonu kullanılır.

$$b_j(o_t) = N(o_t; \mu_j, \Sigma_j) \quad (3.27)$$

$b_j(o_t)$, t anında S_j durumunun o_t gözlemini üretme olasılığını verir. S_j durumuna ait olan ortalama vektörü μ_j ile kovaryans matrisi Σ_j parametreleri eğitim sürecinde hesaplanır.

(3.4)'te verilen ilk durum olasılığı vektörü **II**, modele gözlem üretmeyen başlangıç ve son durumları eklenerek (Şekil 3.8), **A** durum geçiş olasılıkları matrisin ilk satırını oluşturabilir.



Şekil 3.8 Gözlem üretmeyen başlangıç ve son durumları eklenen model

Şekil 3.8'den de anlaşılacağı üzere yeni durum geçiş olasılıkları son satırı sıfırlardan oluşur. (3.21)'de verilen ifade, (3.22)'de yerine konulup gözlem olasılık yoğunluk fonksiyonu eklenerek (3.28)'deki forma dönüştürülür ve eğitim seti olarak $X = \{O^k\}_{k=1}^K$ kullanılarak yeni geçiş olasılıkları matrisinin elemanlarını kestirilir (Young vd., 2002).

$$\hat{a}_{ij} = \frac{\sum_{k=1}^K \frac{1}{P_k} \sum_{t=1}^{T_k-1} \alpha_t^k(i) a_{ij} b(o_{t+1}^k) \beta_{t+1}^k(j)}{\sum_{k=1}^K \frac{1}{P_k} \sum_{t=1}^{T_k-1} \alpha_t^k(i) \beta_t^k(i)}, \quad 1 < i < N \text{ ve } 1 < j < N \quad (3.28)$$

Burada P_k , O^k gözleminin tamamının model tarafından üretilme olasılığı $P(O^k | \lambda)$ 'ya karşı gelir. Gözlem üretmeyen başlangıç durumundan, gözlem üreten düğümlere geçiş olasılıklarını

hesaplamak için (3.29) kullanılır. Burada elde edilen değerler ilk durum olasılıkları vektörüne denk gelmektedir.

$$\hat{a}_{1j} = \frac{1}{K} \sum_{k=1}^K \frac{1}{P_k} \alpha_1^k(j) \beta_1^k(j), \quad 1 < j < N \quad (3.29)$$

Gözlem üretmeyen son duruma geçiş olasılıkları ise (3.30) yardımıyla elde edilir.

$$\hat{a}_{iN} = \frac{\sum_{k=1}^K \frac{1}{P_k} \alpha_T^k(i) \beta_T^k(i)}{\sum_{k=1}^K \frac{1}{P_k} \sum_{t=1}^{T_k} \alpha_t^k(i) \beta_t^k(i)}, \quad 1 < i < N \quad (3.30)$$

3.5.2 Sürekli Saklı Markov Modelde Üç Temel Problem

Bölüm 3.3'te ayrıntılı olarak incelenen üç temel problemin çözümü olan ileri-geri yön, Viterbi ve Baum-Welch algoritmalarının da sürekli gözlem vektörleri ile uyumlu hale getirilmesi gerekir. Bu noktadaki esas sorun (3.27)'de verilen modeli oluşturan her duruma ait kovaryans matrisi ile ortalama vektörünün (Σ_j, μ_j) nasıl hesaplanacağıdır (Young vd., 2002). Model tek bir durumdan oluşmuş olsa, hesaplamalar çok kolaylaşacak basitçe (3.31) ve (3.32)'deki ifadeler kullanılabilecekti.

$$\hat{\mu}_j = \frac{1}{T} \sum_{t=1}^T o_t, \quad 1 < j < N \quad (3.31)$$

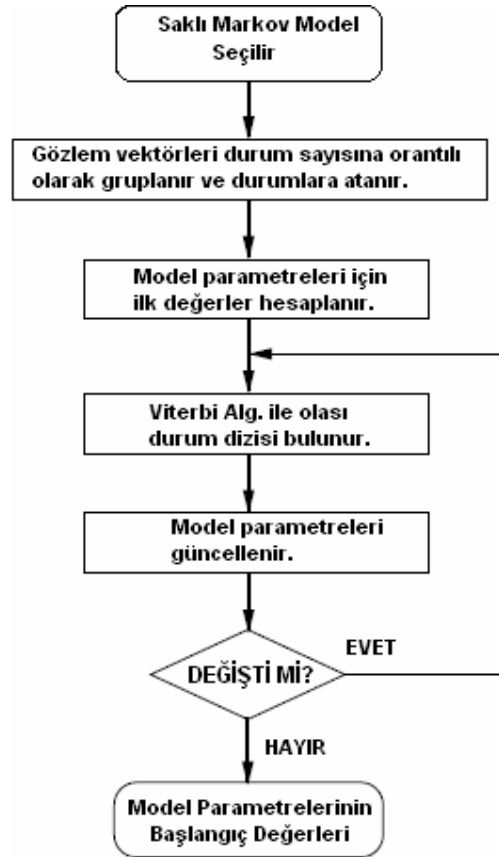
$$\hat{\Sigma}_j = \frac{1}{T} \sum_{t=1}^T (o_t - \mu_j)(o_t - \mu_j)', \quad 1 < j < N \quad (3.32)$$

Oysa ki pratikte modelin birden fazla durum içereceği aşıkardır ve gözlem vektörlerinin durumlara direk atanması, durum dizisinin saklı olmasından ötürü mümkün değildir. Fakat kovaryans matrisine ve ortalama vektörüne en azından başlangıç değerlerinin atanabilmesi için bazı varsayımlar yapılır. Gözlem vektörleri modelin durum sayısına göre gruplara ayrılıp, (3.31) ve (3.32) kullanılarak Σ_j, μ_j için başlangıç değerleri elde edilir (Young vd., 2002). Daha sonra başlangıç değerleri ile Viterbi algoritması kullanılarak olasılığı en yüksek durum dizisi belirlenip Σ_j, μ_j 'in değerleri güncellenir. Bu işleme kestirilen değerler değişmeyene kadar devam edilir. Böylece Baum-Welch algoritmasında gerekli olan başlangıç değerleri elde edilmiş olur (Şekil 3.9). İleri-yön ve geri-yön algoritmalarında gözlem olasılıkları matrisinin elemanları yerine, doğrudan gözlem olasılık yoğunluk fonksiyonu kullanılarak çözüm

sağlanır. Bu değişiklikler ışığı altında Baum-Welch algoritması kullanılarak modelin tüm parametrelerinin kestirimi yapılır. Durum geçiş olasılıklarının hesaplanmasında kullanılan (3.28), (3.29) ve (3.30)'un yanı sıra Σ_j, μ_j parametrelerinin kestirimi için (3.33) ve (3.34) kullanılır.

$$\hat{\mu}_j = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_t^k(j) o_t^k}{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_t^k(j)} \quad (3.33)$$

$$\hat{\Sigma}_j = \frac{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_t^k(j) (o_t^k - \hat{\mu}_j)(o_t^k - \hat{\mu}_j)'}{\sum_{k=1}^K \sum_{t=1}^{T_k} \gamma_t^k(j)} \quad (3.34)$$

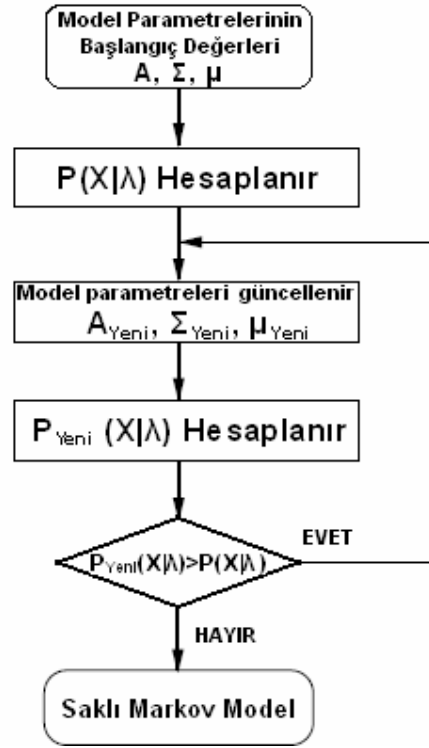


Şekil 3.9 Baum-Welch algoritmasına başlamadan önce başlangıç değeri bulma adımları

(3.33) ve (3.34)'de verilen ifadelerde, (3.31) ve (3.32)'den farklı olarak işlemlere (3.18)'de tanımlanan $\gamma_t(i)$ katılmış, böylece gözlem vektörünün direk bir duruma atanması yerine o gözlem vektörünün o durum tarafından üretilme olasılığı değerlendirilmiş olur.

Eğitim seti olarak $X = \{O^k\}_{k=1}^K$ verildiğinde, Şekil 3.10'da özetlenen Baum-Welch algoritmasının adımları gerçekleştirilerek model parametreleri kestirilmiş olur. Durma şartı

$$P(X|\lambda) = \prod_{k=1}^K P(O^k|\lambda) \text{ olasılığının artarak belli bir noktaya gelmesidir.}$$

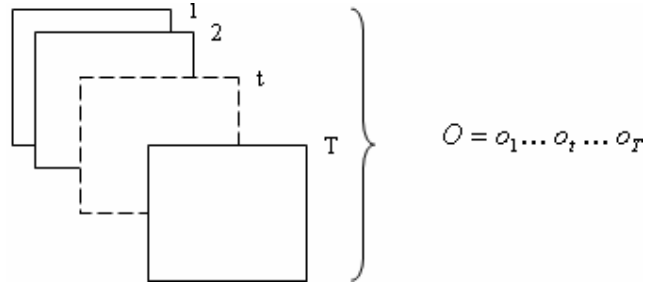


Şekil 3.10 Baum-Welch algoritmasının adımları

3.6 Saklı Markov Modelin Video Görüntüsünden Hareket Tanımayaya Uygulanması

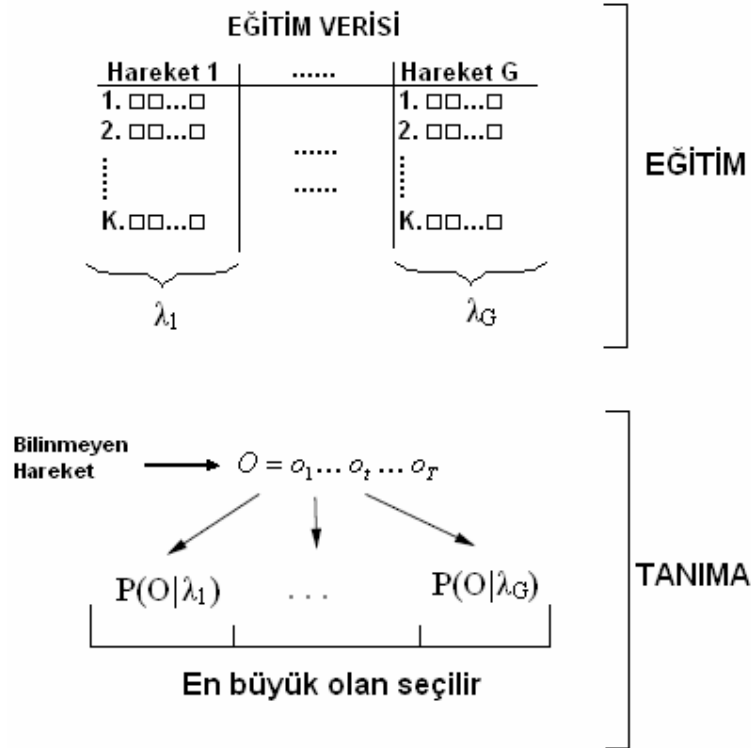
Saklı Markov model ses tanımayaya başarı ile uygulanmaktadır. Hareket tanımayaya ilk uygulaması ise Yamato vd. (1992) tarafından 25x25 çözünürlüğünde tenis vuruşlarını içeren 6 farklı görüntüyü tanıma çalışmalarıdır. Verinin bu kadar düşük miktarda olmasına rağmen, elde edilen tanıma sonuçlarının hayli iyi olması saklı Markov modellerin hareket tanıma için ne denli uygun olduğunu göstermeye yetmiştir.

Hareketi sınıflandırmak için saklı Markov model kullanıldığında her sınıfa karşı gelen bir model bulunmalıdır. Örneğin, 45 farklı hareketi sınıflandırmak için her görüntü için parametreleri farklı 45 model belirlenmelidir. Herhangi bir kelimeye karşı gelen bir işaret dili hareketi, gözlem dizisi $O = o_1 K o_t K o_T$ ile ifade edilebilir. Bu durumda, hareket arka arkaya gelen görüntü çerçevelerinin oluşturduğu bir yapı olarak düşünülürse, bir harekete karşı gelen O gözlem dizisinin her elemanı da bir çerçeveye karşı gelir (Şekil 3.11).



Şekil 3.11 Görüntü çerçevesinden gözlem dizisi oluşturulması

Dizideki o_t elemanı, t anındaki görüntü çerçevesinden çıkartılmış özellikleri içeren gözlem vektörüdür. Tabii burada şu belirtilmelidir ki, uygulamaya göre özelliklerin sayısı ve aldığı değerler değişebilir. Eğer gözlem vektörünün elemanları $V = \{v_1, v_2, K, v_M\}$ gibi bir kümeden sınırlı değerler alıyorsa, gözlem vektörü ayrıktır denir. Şayet görüntü çerçevesinden elde edilen özellikler sürekli değerli ise mutlaka bir kuantalama metodu ile bu değerler ayrık hale getirilmelidir (Alpaydın, 2004) yada Bölüm 3.5’de incelenen sürekli saklı Markov model kullanılmalıdır. Bu çalışmada sürekli saklı Markov model kullanılarak hareket tanıma çalışması yapılmıştır. Öncelikle eğitim verisi kullanılarak, G ayrı saklı Markov model eğitilip model parametreleri belirlenmiştir (Şekil 3.12).



Şekil 3.12 Saklı Markov model ile hareket tanıma

O gözlem dizisinin hangi kelimeye karşı geldiğini anlamak için, sistemin tanıdığı G tane harekete karşı gelen G farklı saklı Markov modelin bu gözlemi üretme olasılıkları $P(O | \lambda_i)$, $i = 1, K, G$, ileri-yön yada geri-yön algoritmaları kullanılarak hesaplanır. En büyük olasılığı sağlayan modelin bu gözlemi ürettiği kabul edilir ve O gözlemine modele karşı gelen kelimedir denilir.

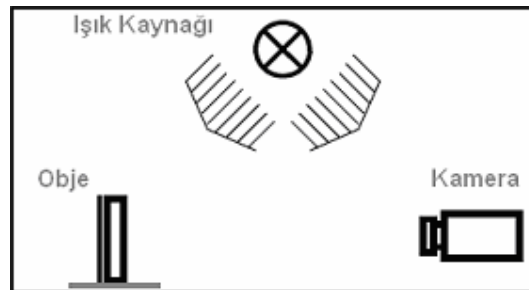
4. ÖZELLİK SEÇİMİ

Basit bir hareket tanıma sistemi ışık kaynağı, görüntü yakalama aygıtı ve verinin işleneceği bilgisayardan oluşur. Sistemi tasarlamamanın önemli bir adımı, sistemin çalışacağı ortam dikkate alınarak uygun özellik seçimidir. Elde edilen verinin sağlıklı değerlendirilebilmesi için öncelikle verinin sağlıklı bir ortamda alınması gerekir. Farklı ışık koşullarında aynı ton ve doygunluk seviyesine sahip renklerin farklı değerlerde algılanması nedeniyle görüntünün sayısal ortama aktarımı sırasında ortam şartlarının sisteme etkisi büyüktür. Benzer şekilde kullanılan görüntü yakalama aygıtının özellikleri de sistemin işleyişini değiştirebilecek öneme sahiptir.

4.1 Ortamın Aydınlatılması

Görüntü işleme uygulamasında önemli noktalardan bir tanesi sistemin çalışacağı ortama uygun bir ışık kaynağı seçmektir. Böylece görüntüyü iyileştirme aşaması için harcanacak zaman minimuma indirilmiş olur. Olası ışık kaynaklarından birkaçını şöyle sıralanabilir: gün ışığı; çıplak ampul; flüoresan lamba ve halojen lamba.

Genelde gün ışığını görüntü işleme sistemlerinde aydınlatma kaynağı olarak düşünmek doğru olmaz çünkü ışığın rengi ve parlaklığı gün içersinde değişmektedir. Çıplak ampuller ucuz bir seçenek olsa da zaman zaman kamera ile istenmeyen etkileşimler meydana geldiği için görüntünün kalitesi düşmektedir. Bir diğer noktada doğrudan akım ile çalışmaları nedeniyle çok ısınmaları ve buna bağlı olarak düzgün dağılımlı aydınlatma sağlayamamalarıdır. Flüoresan lambaların homojen aydınlatma alanı çok geniştir ve çok ısınmazlar. Tek olası dezavantajı kaplama maddesinden kaynaklanan bazı spektral sınırlamalardır ki bazen uygulamaya bağlı olarak bu istenilen bir özellik olabilir. Uygun özelliklerinden ötürü flüoresan lambalar birçok görüntü işleme sisteminde kullanılmaktadır (Erhardt-Ferron, 2000). Bu nedenlerden ötürü, tez çalışmasında tasarlanan sistemde ortamın aydınlatılmasında 2 tane 20 W Philips® Energy Saver flüoresan lamba kullanılmıştır (Şekil 4.1).



Şekil 4.1 Ortamın aydınlatılması

Tabii burada önemli bir başka noktada kameranın, ışık kaynağının ve objenin konumudur. Işın başında veriyi en doğru şekilde alabilmek için, Şekil 4.1’de gösterildiği üzere kameranın ve ışık kaynağının objenin aynı tarafında olduğu aydınlatma yapısı kullanılmıştır (Erhardt-Ferron, 2000).

4.2 Görüntü Yakalama Aygıtı

Birçok görüntü işleme sisteminde CCD (Charge Coupled Device) alıcılar büyük öneme sahiptir. CCD kameralarda mekanik parça yoktur, gelen ışık ışığa duyarlı yarı iletken elemanlar içeren CCD alıcının üzerine düşer. Piksel olarak düşünülebilecek bu elemanlar matris yada satır şeklinde sıralıdırlar. Birçok sistemde tercih edilmesinin nedenlerinden bir tanesi düşük ışıktaki bile kaliteli görüntü yakalamaya imkan vermesidir (Erhardt-Ferron, 2000). Veri kalitesini artırabilmek için görüntü yakalama aygıtı olarak Philips® PCVC 840K model kamera kullanılmıştır. 1/3” VGA CCD alıcısı olan kamera saniyede 60 çerçeve oranına sahiptir ve 1 lüksten daha düşük ışık olan ortamda bile kaliteli görüntü iletimi yapabilmektedir [2].

Dengeli bir sistemin tasarlanabilmesi için kullanılan kameranın parlaklık, doygunluk, beyaz dengesi vb. gibi çerçeve iletimi sırasında ortam ve görüntü şartlarına bağlı olarak otomatik değişebilen parametrelerin sabit bir değerde kalması sağlanmalıdır (Sangwine ve Horne, 1998).

4.3 Görüntüde Ellerin Bulunması

Görüntü tabanlı işaret dili tanıma sistemlerinin çoğunda, kullanıcının çeşitli türlerde renkli eldiven giymesi istenmektedir. Bu yaklaşım gerçek zamanlı tasarım ihtiyaçlarını karşılamak ve tanınacak hareket sayısını arttırmak için kullanılmaktadır. Ancak, hastane, banka vb. gibi alanlarda çalışan hareket tanıma sisteminin kullanıcının giydiği eldivene bağlı olması çok uygun gözükmemektedir. Bu nedenle, tasarlanan sistemde insanın ten rengine ait özellikler kullanılarak ellerin takibi gerçekleştirilmiştir. Ayrıca, diğer araştırmalarda genellikle kullanılan 320x240 çözünürlük yerine, 160x120 çözünürlükteki görüntü ile çalışılmış, böylece ten rengi belirlemek için gereken zaman minimuma indirmek amaçlanmıştır.

Tanırma sisteminde Hsu vd.’nin (2002) önerdikleri ten rengi tanıma yöntemi kullanılmıştır. Yöntemi incelemeye başlamadan önce renk uzaylarını incelemek faydalı olacaktır.

4.3.1 Renkli Görüntünün Temel Özellikleri

İnsanların renkleri algılaması gözün mevcut ışığa tepki vermesi ve beynin bu tepkiyi sinirsel işlemlerden geçirmesi ile gerçekleşir. Renkler, insan gözünde video sistemlerinde kullanılan üç temel renge (RGB, Kırmızı Yeşil Mavi) kabaca karşı gelen uzun, orta ve kısa dalga boylarının lineere yakın oranlarda meydana getirdiği kombinasyonlar olarak algılanır. Bir rengi oluşturan temel özellikler renklilik (Chrominance) ve aydınlık (Luminance) bileşenleridir. Renklilik renk tonu ve doygunluğunu verirken, aydınlık bileşeni de parlaklık bilgisini verir (Russ, 1999).

4.3.2 Renk Uzayları

Renkli görüntüyü ifade etmek için, saklama ortamları ve uygulamaların gereksinimlerine uygun olacak şekilde çeşitli renk uzayları kullanılmaktadır. Kullanılacak uzayın seçimi yapılacak işleme göre değişmektedir. Bir rengi bir uzayda temsil etmek için genelde şu öğeler kullanılır: temel renk bileşenleri; parlaklık; renk tonu; doygunluk.

4.3.2.1 RGB Renk Uzayı

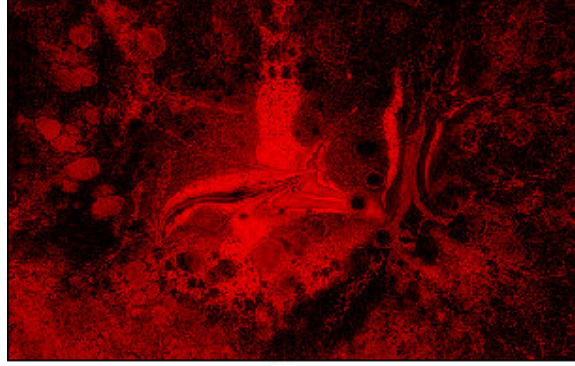
Genellikle renklerin ekranda gösterilmesi için kırmızı, yeşil ve mavi bileşenlerden oluşan RGB uzayı kullanılır. Üç temel rengin 0 ile 255 arasında değerler alan parlaklık oranlarının toplamalı karışımıyla elde edilir. Kırmızı ile yeşilin toplamından sarı, yeşil ile mavinin toplamından turkuvaz, kırmızı ile mavinin toplamından da siyahımsı renkleri elde edilir. Üç temel rengin aynı parlaklık seviyelerinin toplanmasıyla gri tonları elde edilirken, üç bileşenin saf hallerinin toplamı ile beyaz renk elde edilmiş olur. Şekil 4.2’de verilen örnek resmin kırmızı, yeşil ve mavi renk bileşenleri Şekil 4.3’de verilmiştir.



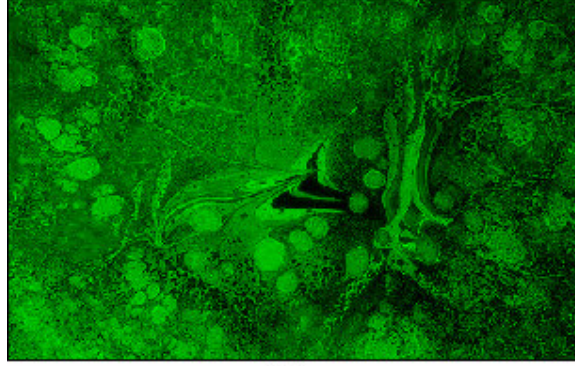
Şekil 4.2 İçinde çok çeşitli renkler içeren, Ebru Sanatı ile yapılmış örnek resim [3]

Aşağıda verilen temel renk bileşenleri ile orijinal resim karşılaştırıldığında renk oluşumunda bileşenlerin katkısı kolayca anlaşılmaktadır. Örneğin, orijinal resmin merkezi civarında

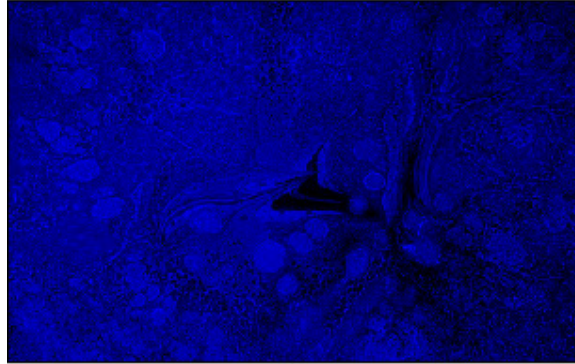
bulunan kırmızı yoğunluklu bölgenin, yeşil (Şekil 4.3(b)) ve mavi (Şekil 4.3(c)) renk bileşenlerine ait resimlerde koyu tonlu renklerle ifade edildiği görülmektedir. Bunun nedeni, o bölgedeki renge, yeşil ve mavi bileşenin katkısının sıfıra yakın olması, dolayısıyla siyaha yakın koyu tonlarda renkler oluşmasıdır.



(a)



(b)



(c)

Şekil 4.3 (a) Kırmızı bileşen (b) Yeşil bileşen (c) Mavi bileşen

RGB uzayında her bir piksel 3 bayt ile gösterildiğinden, görüntüyü hafızada bu formatta saklamak çok yer kaplar. Bundan dolayı RGB uzayı görüntü saklama yada iletimi açısından elverişli değildir. Görüntü işleme açısından da pek uygun olmayan bu uzayda yapılan işlemler uzun sürmektedir. Ayrıca yapılan bazı görüntü iyileştirme işlemlerinde renk atlamaları meydana gelmektedir (Gonzalez ve Woods., 1993).

4.3.2.2 YIQ ve YUV Renk Uzayları

RGB uzayında saklanan görüntünün çok yer kaplaması nedeniyle, televizyon yayıncılığında kullanılan bant genişliğini en aza indirebilmek için, donanıma bağlı olarak YUV yada YIQ (Luma, Inphase, Quadrature) uzayları kullanılır. Y rengin parlaklık bilgisini taşır ve renkli görüntünün gri seviyeli halini (Şekil 4.4) ifade eder, UV ya da IQ bileşenleri de renk bilgisini taşır (Russ, 1999).



Şekil 4.4 Verilen örnek resmin Y (gri seviyeli) bileşeni

Avrupa'da yaygın olarak kullanılan PAL (Phase Alternating Line) sistemlerde YUV uzayı, Amerika'da kullanılan NTSC (National Television Standards Committee) sistemlerde ise YIQ uzayı kullanılmaktadır. Bu uzaylar ile RGB uzayı arasındaki ilişki (4.1) ve (4.2)'de verilmektedir.

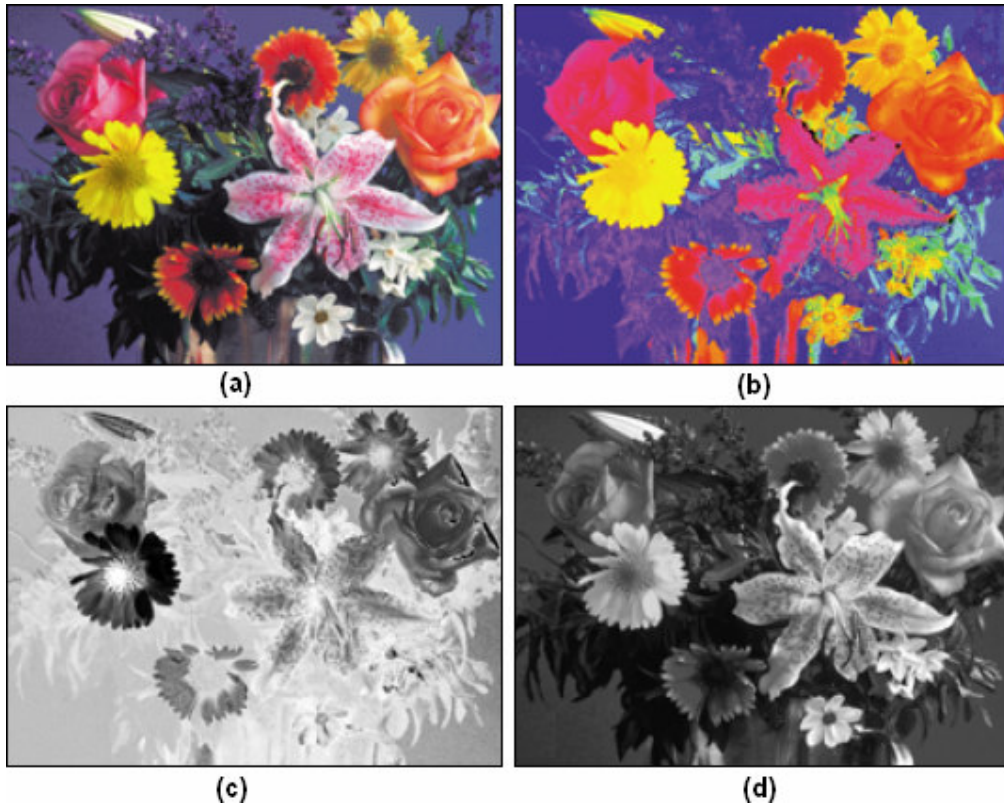
$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.1 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.1)$$

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.523 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad (4.2)$$

Böylece görüntünün parlaklık özelliği ile renklilik bileşenlerini birbirinden ayrılmış olur. YUV uzayının yeni teknolojilerde kullanılmasının nedenlerinden bir tanesi insanın görme duyusunun, uzaysal bileşenlerden renkliliğe parlaklıktan daha az duyarlı olmasıdır. Bu nedenle YUV uzayında kroma bileşenleri parlaklık bileşenine göre daha az örneklenir (Russ, 1999). Y bileşeni görüntünün siyah-beyaz televizyonlardaki bileşenini taşımaktadır.

4.3.2.3 HSI, HSV ve YCbCr Renk Uzayları

Görüntü işleme uygulamaları açısından HSI (Hue Saturation Intensity), HSV (Hue Saturation Value) ve YCbCr (Luminance, Mavi renklilik, Kırmızı renklilik) uzayları daha uygun gözükmetedir. Çünkü bu uzayların renk bileşenlerini ayırım biçimi, insan görme sisteminin algılama şekli ile örtüşmektedir. HSI uzayı bileşenleri, H (Hue) temel renk tonunu, S (Saturation) ise rengin miktarını ifade eder. Örneğin pembe ve kırmızı renkler için H değeri aynı iken farklı oranlarda kırmızı renk içermeleri nedeniyle S değerleri farklıdır. I (Intensity) değeri ise rengin parlaklığını verir (Russ, 1999). Şekil 4.5(a)'da verilen görüntünün H, S ve I bileşenleri sırasıyla verilmiştir.



Şekil 4.5 (a) Orijinal resim (b) H bileşeni (b) S bileşeni (b) I bileşeni (Russ, 1999)

RGB uzayından HSI uzayına dönüşümler (4.3)'de verilmiştir.

$$\left. \begin{aligned}
 H &= \cos^{-1} \left\{ \frac{\frac{1}{2}[(R-G) + (R-B)]}{[(R-G)^2 + (R-B)(G-B)]^{1/2}} \right\} \\
 S &= 1 - \frac{3}{(R+G+B)} [\min(R, G, B)] \\
 I &= \frac{1}{3}(R+G+B)
 \end{aligned} \right\} R, G, B \in [0,1] \quad (4.3)$$

(4.1), (4.2) ile (4.3) karşılaştırıldığında RGB uzayından YUV ve YIQ uzayına dönüşümün ve bunu donanımsal olarak gerçekleştirilmesinin daha kolay olduğu görülmektedir. Bu açıdan bakıldığında YUV ve YIQ uzaylarının neden televizyon yayıncılığında kullanıldığı kolayla anlaşılmaktadır.

Görüntü işleme açısından uygun olan bir diğer renk uzayı da HSV uzayıdır. HSV renk uzayı insanın algılamasına uygun olarak açısıl bir koordinat sistemine sahiptir. Kırmızı, mavi ve yeşilin doğrusal olmayan bir şekilde yeniden organize edilmiş şeklidir. HSI uzayına benzer şekilde, H (renk tonu) temel rengin dalga boyunun ölçüsüdür, S (doygunluk) ise rengin saflık ölçüsüdür ve V bileşeni de rengin parlaklığını verir.

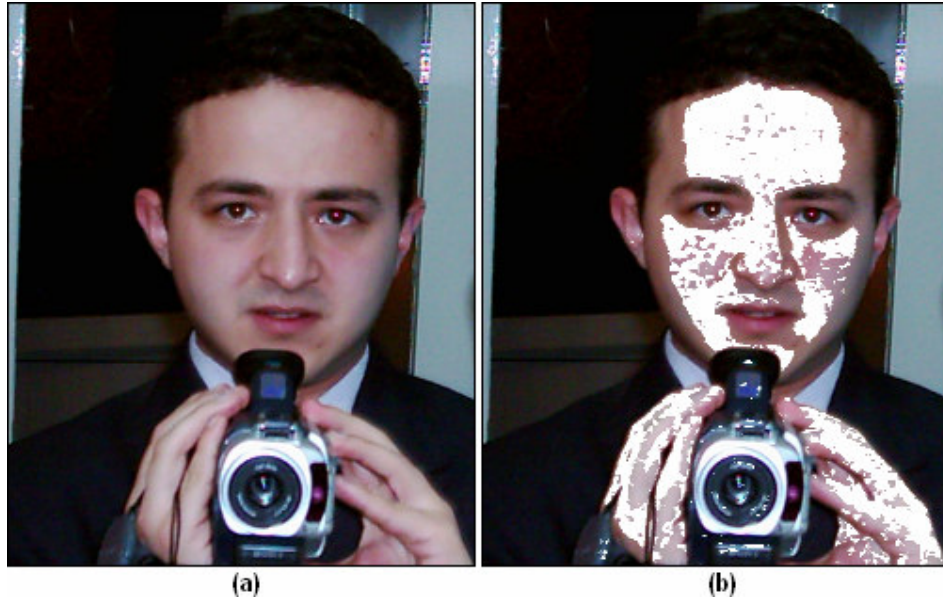
YCbCr uzayı televizyon sinyali kodlama uzayları YUV ve YIQ'den bağımsız olmakla birlikte, Y bileşeni bu uzaylarda olduğu gibi rengin parlaklık bilgisini taşır. Renklilik bileşenleri ise renk ve doygunluk bilgisini verir. YCbCr uzayı video görüntü kodlamasında ve görüntü sıkıştırma uygulamalarında kullanılmaktadır (Sangwine ve Horne, 1998). Renk bileşenlerini ayırım biçimi bakımından görüntü işleme çalışmalarına oldukça uygundur.

4.3.3 Ten Rengi Tanıma

Yüz tanıma sistemlerinde, en büyük sorun görüntüde yüzün bulunmasıdır. Genelde yüzün konumu konusunda bazı ön kabuller yapılarak arama yapılır, ancak bu ön kabuller çoğu zaman gerçek hayat uygulamaları için doğru sonuç vermez. Hsu vd. 2002'de geliştirdikleri yüz tanıma algoritması, renkli resimlerde yüzü bulurken ilk adım olarak görüntüde ten rengi olabilecek pikselleri başarı ile ayırt edebilmektedir. Renkli görüntüde ten rengini bulabilmek için yapılması gereken ön önemli adım Bölüm 4.3.2'de bahsedilen renk uzaylarından doğru olanını seçebilmektir.

Ten rengi tanıma konusunda farklı renk uzaylarında birçok çalışma yapılmış ve en iyi

sonuçların TSL (Tint Saturation Luma) uzayında elde edildiği belirlenmiştir (Terrillon vd., 2000). Ancak RGB uzayından TSL uzayına çift taraflı dönüşüm bağıntıları ağır matematiksel ifadeler içerdiğinden, gerçek zamanlı uygulamalar için pek uygun değildir. Bu nedenle, Hsu vd. (2002) rengin renklilik ve parlaklık bileşenlerini ayırma şekli TSL uzayına benzeyen YCbCr uzayını kullanılarak ten rengi tanıma uygulaması yapmışlardır. Yapılan çalışmaların çoğunda ten rengi olan pikseller için renklilik bileşenlerinin parlaklık bileşeninden bağımsız olduğu varsayımı yapılmıştır, ancak ten rengi tonları lineer olmayan bir şekilde parlaklığa bağlıdır. Hsu vd. (2002) ise yüz veritabanından alınan ten rengi örneklerini YCbCr uzayında belirlemiş ve bazı matematiksel fonksiyonlar yardımıyla dönüşümler yaparak bu bağımlılığın ortadan kaldırılmasını sağlamışlardır. Böylece YCbCr uzayında belirli sınırlar içinde kalan pikseller ten rengi olarak değerlendirilmiştir (Şekil 4.6).



Şekil 4.6 (a) Orijinal resim (b) Aday ten rengi piksellerin bulunduğu resim

YCbCr renk uzayında, renklilik bileşenleri (Cb ve Cr) parlaklık (Y) bileşeninin fonksiyonu olarak $C_b(Y)$ ve $C_r(Y)$ şeklinde ifade edilirse, $C_b'(Y)$ ve $C_r'(Y)$ ise dönüştürülmüş renklilik bileşenleridir. Gerekli dönüşümler aşağıda verilmiştir.

$$C'_i(Y) = \begin{cases} (C_i(Y) - \bar{C}_i(Y)) \cdot \frac{W_{C_i}}{W_{C_i}(Y)} + \bar{C}_i(K_h), & Y < K_l \text{ yada } K_h < Y \\ C_i(Y), & Y \in [K_l, K_h] \end{cases} \quad (4.4)$$

$$W_{C_i}(Y) = \begin{cases} WL_{C_i} + \frac{(Y - Y_{\min}) \cdot (W_{C_i} - WL_{C_i})}{K_l - Y_{\min}}, & Y < K_l \\ WH_{C_i} + \frac{(Y_{maks} - Y) \cdot (W_{C_i} - WH_{C_i})}{Y_{maks} - K_h}, & K_h < Y \end{cases} \quad (4.5)$$

$$\overline{C_b}(Y) = \begin{cases} 108 + \frac{(K_l - Y) \cdot (118 - 108)}{K_l - Y_{\min}}, & Y < K_l \\ 108 + \frac{(Y - K_h) \cdot (118 - 108)}{Y_{maks} - K_h}, & K_h \leq Y \end{cases} \quad (4.6)$$

$$\overline{C_r}(Y) = \begin{cases} 154 - \frac{(K_l - Y) \cdot (154 - 144)}{K_l - Y_{\min}}, & Y < K_l \\ 154 - \frac{(Y - K_h) \cdot (154 - 132)}{Y_{maks} - K_h}, & K_h \leq Y \end{cases} \quad (4.7)$$

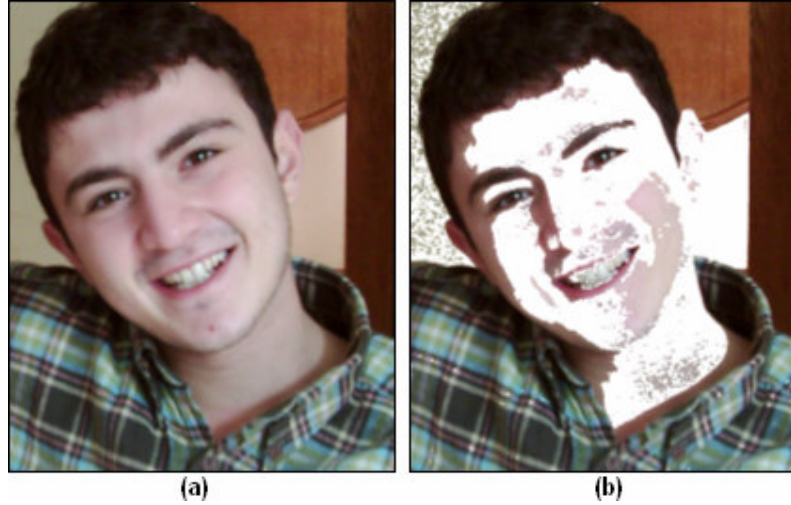
(4.4) ve (4.5)'de C_i ifadesi C_b ya da C_r 'ye karşı gelir. (4.4), (4.5), (4.6) ve (4.7)'de verilen sabitlerin değerleri: $W_{C_b} = 46,97$, $WL_{C_b} = 23$, $WH_{C_b} = 14$, $W_{C_r} = 38,76$, $WL_{C_r} = 20$, $WH_{C_r} = 10$, $K_l = 125$ ve $K_h = 188$ 'dir (Hsu vd., 2002). Bu sabitler veritabanındaki örnek ten rengi verilerinden elde edilen değerlerdir. $Y_{\min}$ ve Y_{maks} sabitleri YCbCr uzayında sırasıyla 16 ve 236 değerlerine eşittir ve ten rengi olabilecek pikseller için eliptik model $C'_b C'_r$ uzayına dönüştürülmüştür (4.7) ve (4.8).

$$\frac{(x - ec_x)^2}{a^2} + \frac{(y - ec_y)^2}{b^2} = 1 \quad (4.8)$$

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} C'_b - c_x \\ C'_r - c_y \end{bmatrix} \quad (4.9)$$

(4.8) ve (4.9)'da verilen sabitlerin değerleri: $c_x = 109,38$, $c_y = 152,02$, $\theta = 2,53$, $ec_x = 1,60$, $ec_y = 2,41$, $a = 25,39$ ve $b = 14,03$ 'dir. (4.8) ve (4.9) ifadelerini sağlayan x ve y değerleri ten rengi bölgelerinin $C'_b C'_r$ uzayında oluşturduğu kümeye denk gelmektedir ve eliptik alan içinde kalan pikseller ten rengi bölgeleridir (Hsu vd., 2002).

Yapılan testlerde görülmüştür ki, görüntüdeki bazı piksellerin ten rengine çok benzemesi nedeniyle hatayla ten olarak belirlenmesi mümkün olmaktadır (Şekil 4.7).



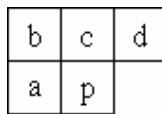
Şekil 4.7 (a) Orijinal Resim (b) Ten olmayan bölgelerin seçilmesi

Bu nedenle, aday ten rengi pikseller belirlendikten sonra bazı ek kontroller yapılarak bu bölgelerin gerçekten el ya da yüz olup olamayacağı kontrol edilmelidir.

4.3.4 El ve Yüz Olabilecek Şekillerin Değerlendirilmesi

Aday ten rengi pikseller belirlendikten sonra oluşturdukları bölgeler bağlantılı bileşen etiketleme (CCL, Connected Component Labeling) yöntemiyle gruplandırılır. CCL algoritması görüntüde ten olabilecek piksellerin komşuluklarına bakarak kaç farklı grup olduğunu belirler (Gonzalez ve Woods, 1993). Aynı grupta bulunan tüm piksellere aynı etiket değeri atanır. CCL algoritmasının adımları aşağıda verilmiştir:

- Ten rengi olan pikseller görüntüde sol üst köşeden başlayarak kontrol edilir.
- Pikselin sol komşusunun grup etiketine bakılır.
- Eğer komşu etiketlenmiş ise komşu etiketinin değeri mevcut piksele verilir.
- Eğer komşu etiketlenmiş ama değeri mevcut piksel ile aynı değilse, bu iki etiketin aynı gruba ait olduğu bilgisi not edilir.
- Bu işlemler Şekil 4.8’de gösterilen 4 komşuya sırasıyla uygulanır.
- Tüm pikseller gözden geçirildikten sonra aynı grup olduğu belirtilen etiketler düzenlenir.



Şekil 4.8 CCL algoritmasında p pikselinin a,b,c,d ile gösterilen 4 komşuna bakılır

Burada en çok üyeye sahip ilk üç grup yüz ve eller olarak kabul edilir, diğer grupları oluşturan üyelerin gürültü olduğu kabul edilip değerlendirme dışı bırakılır. En fazla üyeye sahip olan C_1 grubu yüz olarak kabul edilir, ikinci ve üçüncü en büyük gruplar C_2 , C_3 eller olarak kabul edilir. Görüntü sol taraftan taranarak bulunan ilk el grubunun sol el olduğu varsayımı yapılır. Buna ek olarak, her üç grubun boyutlarının gerçekten 160x120 çözünürlüğündeki bir görüntüde el ve yüze ait olup olamayacağı kontrol edilir. $F_{\min} - F_{\max}$ ve $H_{\min} - H_{\max}$ çiftleri 160x120 çözünürlükteki bir görüntüde sırasıyla ellerin ve yüzün olabileceği minimum ve maksimum değerleri göstermektedir. Bu eşik değerleri eğitim verisi olarak kullanılan görüntülerden elde edilmiştir. S_1 , S_2 ve S_3 sırasıyla görüntüde yüz ve eller olarak belirlenen grupların boyutları olmak üzere:

- Eğer $F_{\min} < S_1 < F_{\max}$ ise C_1 grubu yüzdür.
- Eğer $H_{\min} < S_2 < H_{\max}$ ve $H_{\min} < S_3 < H_{\max}$ ise C_2 ve C_3 grupları eldir.

Burada yapılan kabullerin doğru olabilmesi için aşağıdaki koşullar varsayılmıştır:

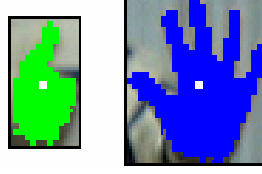
- i. Sol el her zaman sağ elin solundadır.
- ii. Eller hiçbir zaman yüzün üst sınırını geçemez.

CCL algoritması ile görüntüde eller ve yüz Şekil 4.9'da gösterildiği gibi ayırt edilebilmektedir.



Şekil 4.9 CCL algoritması ile ten rengi olan piksellerden ellerin ve yüzün bulunması

Eller ve yüz belirlendikten sonra, elleri ve yüzü oluşturan piksellerin minimum ve maksimum koordinatları kullanılarak elde edilen el ve yüz çerçevesinin merkezleri (Şekil 4.10) bir sonraki görüntüde ellerin ve yüzün yerlerini hızlıca bulabilmek için kullanılır.



Şekil 4.10 Ellerin merkezlerinin bulunması

4.4 Hareket Takibi

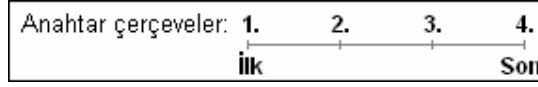
Tüm piksellerin taranarak uygun aday ten rengi piksellerin elde edilmesi ve sonrasında yapılan işlemler gerçek zamanlı çalışan sistemlerden beklenen performansı düşürebilecek durumdadır. Bu nedenle eller ve yüz belirlendikten sonra, aynı işlemleri tekrar tekrar yapmak yerine ellerin ve yüzün merkez koordinatlarını bir sonraki çerçevede kullanarak, yerlerini daha hızlı bulmak mümkün olmaktadır. Merkez koordinat olan piksellerin 8 komşusuna yinelemeli olarak bakılarak ellerin ve yüzün yeri bulunmaktadır (Starner, 1995). Zaman zaman ellerin hızlı hareketi nedeniyle bu yöntem sonuç vermediğinde, tüm pikselleri teker teker kontrol edilerek arama yapılmış ve eller bulunduktan sonra tekrar merkeze bağlı arama yapılmıştır. Bu yöntem döngüsü eller kaybedildiği sürece birbiri ardına devam etmektedir.

Burada hareket takibi için yapılabilecek bir diğer çalışmada çerçeve farkını alarak belli bir eşik değerinin üzerinde olan bölgelerde hareket olduğunu varsayıp bu bölgelerde eli aramaktır (Zhao vd., 2004). Tabii bu işlem sistemi daha fazla yoran bir çözümdür.

4.5 Anahtar Çerçeve Seçimi

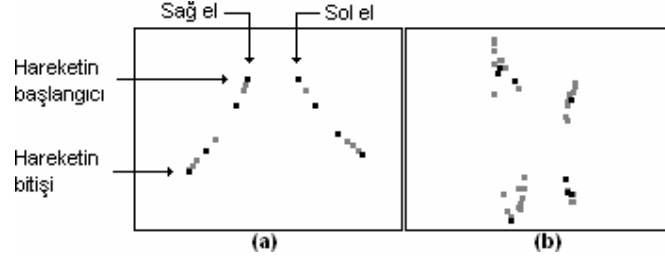
İşaret dilini meydana getiren kelimeler zamanda ve uzayda değişir. Şöyle ki, bir kişi aynı işareti iki kere tekrarladığında her iki hareket arasında farklılıklar ortaya çıkabilmektedir. Atlamalı geçişlere sahip olan bir saklı Markov model bu farklılıkların üstesinden gelerek tanıma yapabilir (Starner, 1995). Bununla birlikte, tüm işaretler için değişmeyen sayıda duruma sahip aynı saklı Markov modeli kullanmak uygun olmaz çünkü sistemin tanıdığı işaretler arasında uzun ve kısa süreli işaretler vardır.

Bu problemi aşmak için hareketi meydana getiren tüm çerçeveleri dikkate almak yerine anahtar çerçeve adı verilen 4 çerçeve kullanılmıştır. Bir hareketin başlangıç ve bitiş anları belirlendikten sonra hareket süresi üç eşit parçaya bölünmüş ve Şekil 4.11’de gösterildiği gibi anahtar çerçeveler seçilmiştir.



Şekil 4.11 Anahtar çerçevelerin seçilmesi

İki farklı işarete ait hareketler sırasında ellerin izlediği konum bilgisi Şekil 4.12’de verilmiştir. Burada gri kareler hareket sırasında tüm çerçevelerde ellerin yerlerini gösterirken, siyah kareler ise anahtar çerçevelerde ellerin konumlarını göstermektedir.



Şekil 4.12 İşaret sırasında ellerin konumu

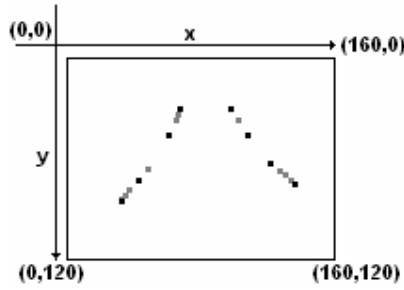
Şekil 2.4’te verilen global özellik grupları incelendiğinde, sistemin tanıyabildiği hareketlerin bir kısmının tek elle bir kısmının da iki elle gerçekleştirilen hareketler olduğu görülmektedir. İki elle gerçekleştirilen hareketlerden Şekil 4.12 (a) ve (b)’de verilenler dikkatlice incelendiğinde her iki işarete ait farklı özellikler dikkat çekmektedir. Şöyle ki, Şekil 4.12 (a)’da verilen işarete her iki elde simetrik hareketler yapıldığı ancak Şekil 4.12 (b)’de sağ ve sol elin hareketlerinin simetrik olmadığı görülmektedir. Bu asimetrinin nedeni iki elle gerçekleştirilen bazı işaretlerde ellerin farklı hareketi yapması ve hareketin farklı sürelerde tamamlanmasıdır. Bazı durumlarda bir elin hareketi tamamlamışken diğer elin hareketi devam etmektedir. Şekil 2.4’e bakıldığında, örneğin a3 ve d5 gruplarında her iki elin farklı hareketleri yaptığı kolayca anlaşılmaktadır. Bu farklılığa bağlı olarak, her iki el için farklı iki anahtar çerçeve seçici kullanılmıştır. İki elle gerçekleştirilen hareketlerde, sağ ve sol elin hareketi kendi içinde farklı üç parçaya bölünmüş ve dolayısıyla her iki el için farklı anahtar çerçeveler seçilmiştir.

Hareketlerin başlangıç ve bitiş noktaları belirlenirken, işareti gerçekleştiren kişinin işaretler arasında kısa süreli beklemeler yaptığı varsayılmıştır. Böylece ellerin konumları belirlenen bir eşik değerinden fazla yer değiştirirse işaretin başladığı ve konumları sabitlendiğinde de işaretin bittiği sonucu çıkarılabilir. Ayrıca bazı durumlarda elin merkezinin değişmemesine rağmen parmakların hareket ettirilmesiyle işaret yapılmaktadır. Böyle bir durumda ellerin merkez noktalarının sabit kalması hareketin olmadığı anlamına gelmemektedir. Bu nedenle, ek bir kontrol yapıp Şekil 4.10’da gösterilen el dış çerçevelerinin boyutlarının değişip

değişmediğine dikkat edilmektedir.

4.6 Global Özelliklerin Çıkarılması

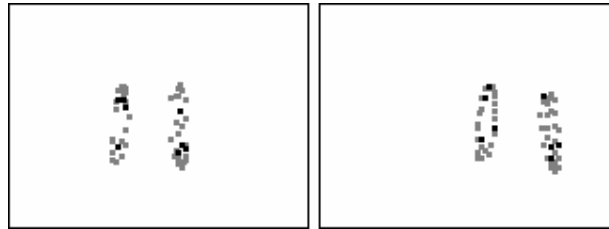
Önceki bölümlerde, işaret dilini meydana getiren hareketlerin ayırt edici özelliklerinin global ve yerel özellikler olmak üzere ikiye ayrıldığından bahsedilmiştir. Tez çalışmasında tasarlanan sistem, bu özellikleri ayrı ayrı değerlendirdiği için iki aşamalı bir tanıma yapmaktadır. İlk aşama olarak, hareket sırasında ellerin izlediği yörünge bilgisi yani işaretin global özellikleri değerlendirilmiştir. Seçilen 4 anahtar çerçevede ellerin yatay ve dikey koordinatları (Şekil 4.13) kullanılarak (4.10)'da verilen özellik matrisi, saklı Markov modellere girdi olarak verilmiştir.



Şekil 4.13 Koordinat sistemi

$$h = \begin{bmatrix} x_R^1 & y_R^1 & x_L^1 & y_L^1 \\ x_R^2 & y_R^2 & x_L^2 & y_L^2 \\ x_R^3 & y_R^3 & x_L^3 & y_L^3 \\ x_R^4 & y_R^4 & x_L^4 & y_L^4 \end{bmatrix} \quad (4.10)$$

Ancak yapılan testlerde (x, y) koordinat sisteminin orijininin çerçevenin sol üst köşesi olarak belirlenmenin (Şekil 4.13) bazı sorunlar ortaya çıkardığı belirlenmiş ve sağlıklı tanıma sonuçları alınabilmesi için işareti yapan kişinin her defasında kameraya aynı açıda bakması gerektiği görülmüştür. Şekil 4.14'te gösterildiği gibi, aynı hareket farklı zamanlarda tekrarlandığında, kamera açısına bağlı olarak ellere ait çok farklı konum bilgileri elde edilebilmektedir.



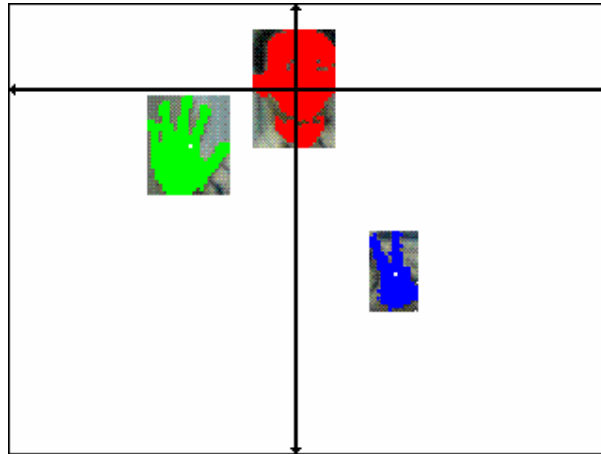
Şekil 4.14 Aynı işaretin farklı kamera açılarıyla tekrarlanması

Bu sorunu aşmak için ellerin anahtar çerçevelerdeki konum bilgilerini doğrudan kullanmak yerine çerçeveler arasındaki konum değiştirme bilgisini kullanma yoluna gidilmiştir. Bu durumda (4.10)'da verilen özellik matrisi, k anahtar çerçevenin indisi ve $\Delta x^k = x^k - x^{k-1}$ olmak üzere (4.11)'deki hali almıştır.

$$h = \begin{bmatrix} \Delta x_R^2 & \Delta y_R^2 & \Delta x_L^2 & \Delta y_L^2 \\ \Delta x_R^3 & \Delta y_R^3 & \Delta x_L^3 & \Delta y_L^3 \\ \Delta x_R^4 & \Delta y_R^4 & \Delta x_L^4 & \Delta y_L^4 \end{bmatrix} \quad (4.11)$$

Ancak bu durumda da ayırt edici bir özellik olan, ellerin harekete başlarkenki konumları bilgisi (Şekil 2.4) kullanılamamış olur. Sadece konum değiştirmelerini kullanmak farklı konumlarda harekete başlayan ama eşit miktarda yer değiştirme yapılan hareketleri ayırt etmeyi güçleştirir. Sonuç olarak farklı bir koordinat sistemi kullanıp, ellerin hareketin başında hangi konumda olduğu bilgisinden faydalanılmasına karar verilmiştir.

Bölüm 1.2'de belirtildiği üzere sistem çerçeve içerisinde tanıma için kullanılsa bile yüz bölgesini bulabilmektedir. İşaretlerin gerçekleştirilmesi sırasında ellerin yüze göre olan bağıl konumları işareti gerçekleştiren kişinin kameraya göre nerede durduğundan bağımsızdır. İşaretçi, görüntü çerçevesinin soluna ya da sağına yakın dursa bile yüzü referans almak Şekil 4.14 ile gösterilen problemin üstesinden gelmektedir. Bu nedenle, yüzün merkez noktasının, kullanılacak yeni koordinat sisteminin orijini olmasına karar verilmiştir (Şekil 4.15).



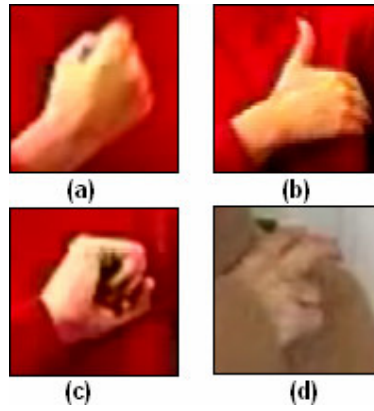
Şekil 4.15 Orijini yüzün merkez noktası olan koordinat sistemi

Harekete başlamadan önce ellerin konum bilgisinin de özellik matrisine eklenmesi sonucu (4.12)'de verilen global özellik matrisi elde edilmiş olur.

$$h = \begin{bmatrix} x_R^1 & y_R^1 & x_L^1 & y_L^1 \\ \Delta x_R^2 & \Delta y_R^2 & \Delta x_L^2 & \Delta y_L^2 \\ \Delta x_R^3 & \Delta y_R^3 & \Delta x_L^3 & \Delta y_L^3 \\ \Delta x_R^4 & \Delta y_R^4 & \Delta x_L^4 & \Delta y_L^4 \end{bmatrix} \quad (4.12)$$

4.7 Yerel Özelliklerin Belirlenmesi

Sistemin tanıyabildiği hareketler incelendiğinde sadece global özelliklere göre sınıflandırma yapmanın yetersiz olduğu açıkça görülmektedir. Aynı global özellik grubuna sahip olup yerel özellikleri, yani hareket sırasında elin biçiminin farklı birçok işaret bulunmaktadır. Örneğin a1 global özellik grubuna (Şekil 2.4) ait 4 tane işaret bulunmaktadır ve bu işaretler gerçekleştirilirken elin aldığı biçimler Şekil 4.16’da verilmiştir.

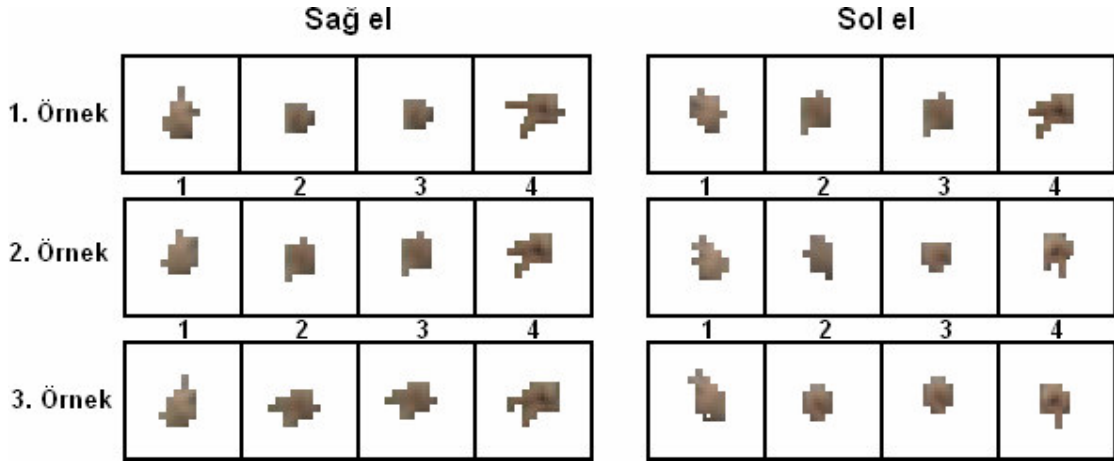


Şekil 4.16 Elin aldığı biçim : (a) acımak, (b) Aksaray, (c) korkmak, (d) subay

Şekil 4.16’da verilenler incelendiğinde, çok sayıda kelime tanıyabilen bir sistem tasarlayabilmek için yerel özelliklerin de hesaba katılması gerektiği anlaşılmaktadır.

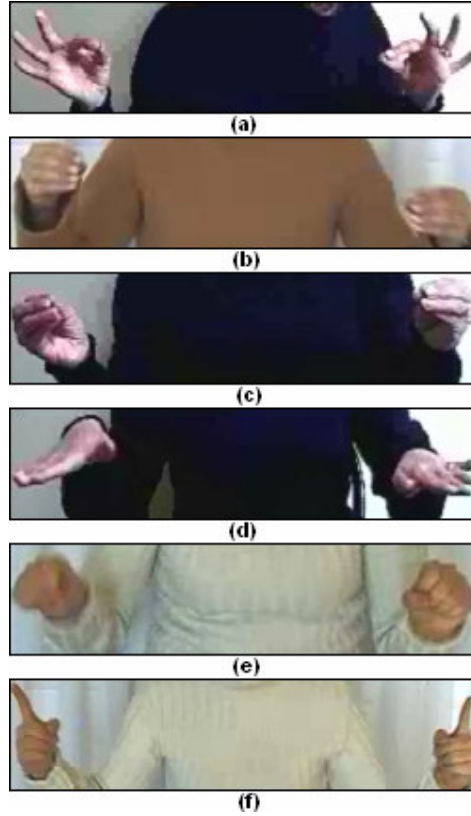
Yapılan denemelerde, aynı hareketin birden fazla tekrarlanması durumunda, elin aldığı biçimin her zaman tamamen aynı kalmamakla birlikte çok da fazla değişmediği gözlenmektedir. Şekil 4.17’de aynı hareketin 3 defa tekrarlanması sırasında, anahtar çerçevelerde ellerin biçimleri gösterilmektedir.

Buna ek olarak, hareketler üzerinde yapılan denemelerde farklı işarete ait hareketlerin 2. ve 3. anahtar çerçevelerinde elin aldığı biçimin zaman zaman aynı olduğu gözlemlenmiştir. Bu nedenle hem gerçek zamanlı çalışma şartlarının sağlanması hem de daha belirgin özelliklerin kullanılabilmesi için sadece hareketin başladığı ve bittiği anda (1. ve 4. anahtar çerçeve) elin biçiminin değerlendirilmesinin yeterli olduğu sonucuna varılmıştır.



Şekil 4.17 Ellerin anahtar çerçevelerdeki durumu

Tek elle gerçekleştirilen hareketlerde doğrudan hareket eden el değerlendirilmiş, diğer elin aldığı biçim önemsizlenmiştir. İki elleri hareketlerde de çok farklı bir durum söz konusu değildir. Şekil 4.18’de c1 global özellik grubuna (Şekil 2.4) ait 6 kelimenin işareti sırasında ellerin hareket sonunda aldığı biçim gösterilmektedir.

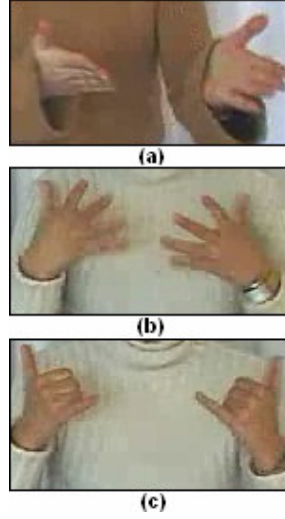


Şekil 4.18 c1 global özellik grubuna ait iki elleri işaretlerde hareketin sonunda ellerin biçimi:
(a) ağır, (b) bırakmak, (c) dinlenmek, (d) kabul etmek, (e) olmak, (f) yer

Yukarıda verilen şekiller incelendiğinde yerel özellikleri belirlerken sadece bir elin aldığı

biçimin değerlendirilmesinin yeterli olacağı gözükmemektedir. Ancak hangi elin kullanılacağı bir diğer sorundur. Şekil 4.18’de verilen ve her iki elinde benzer hareketler yaptığı simetrik hareketlerde hangi elin kullanılacağı fark etmemektedir. Ancak, her iki elin farklı sürelerde ve biçimlerde hareket gerçekleştirdiği simetrik olmayan hareketler de ise durum farklıdır.

Şekil 2.4’te verilen d5 global özellik grubundaki hareketler incelendiğinde, bu gruptaki 3 kelimeye ait hareketin sonunda elin aldığı biçimin aynı olmasına rağmen hareketin başında ellerin aldığı biçimin farklı olduğu görülmüştür (Şekil 4.19).



Şekil 4.19 d5 global özellik grubuna ait işaretlerde hareketin başında ellerin biçimi: (a) fırıncı, (b) futbolcu, (c) oyuncakçı

Bu gözlemler ışığı altında, yerel özellik incelemesi yapılırken işarete ait sadece 1. ve 4. anahtar çevrelerin kullanılmasının ve sadece sağ elin belirtilen çerçevelerdeki biçiminin değerlendirilmesinin yeterli olduğu sonucuna varılmıştır.

5. TÜRK İŞARET DİLİ TANIMA SİSTEMİ

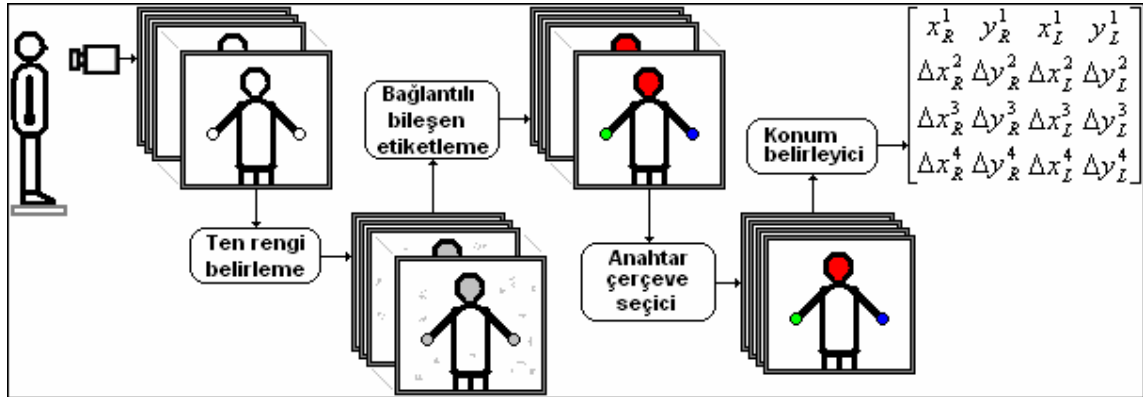
Bu tez çalışmasında, görsel ve algoritmik bölümleri Java programlama dili kullanılarak oluşturulan Türk İşaret Dili Tanıma Sistemi tasarlanmıştır.

5.1 Sistem Yapısı ve Temel Tanıma Prensibi

Türk İşaret Dili Tanıma Sistemi, gerçekleştirilen hareketi tek bir CCD kamera yardımıyla almakta ve seçilen özellikleri kullanarak iki aşamalı bir tanıma işleminden sonra işaretin ifade ettiği kelimeyi belirleyebilmektedir.

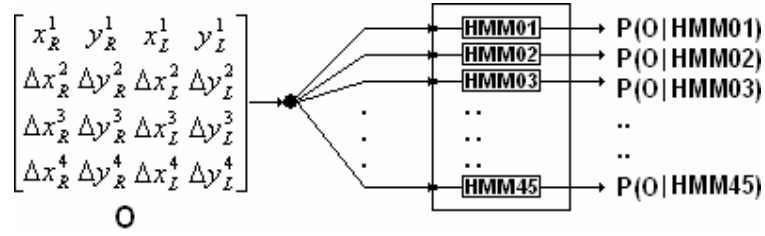
5.1.1 İşaretin Global Özellik Grubunun Belirlenmesi

İlk aşamada, gerçekleştirilen hareketin sadece global özellikleri değerlendirilerek hangi global özellik grubuna ait olduğu belirlenir. CCD kameradan alan görüntü birbirini takip eden çerçevelerden oluşmaktadır. Çerçevelerde ten olabilecek pikseller belirlenir ve bağlantılı bileşen etiketleme metodu kullanılarak eller ve yüz bulunur. Daha sonra anahtar çerçeveler seçilerek, ellere ait konum bilgilerini içeren global özellikler çıkarılır (Şekil 5.1).



Şekil 5.1 Global özelliklerin belirlenmesi

Konum bilgilerinin oluşturduğu O gözlem dizisi 45 farklı saklı Markov modele verilerek, bu gözlemin hangi model tarafından üretilmiş olabileceğine bakılır (Şekil 5.2). Tüm modeller için, bu gözlemin o model tarafından üretilme olasılıkları sırasıyla $P(O | HMM_i)$, $i = 1, K, 45$ hesaplanır.

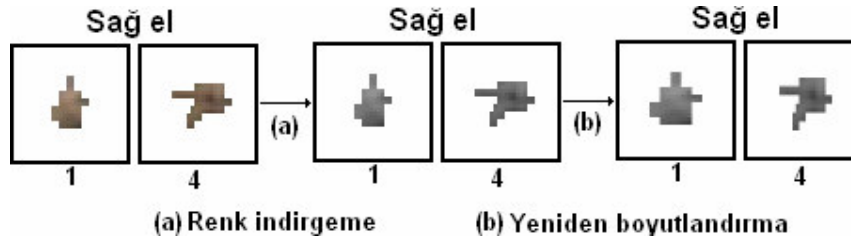


Şekil 5.2 Tüm modellerin O gözlemini üretme olasılıkları

En yüksek olabilirlik olasılığına $P_{maks}(O | HMMj)$ sahip olan modele karşı gelen j global özellik grubu seçilir. Gerçekleştirilen hareketin global özellik grubu belirlendikten sonra, bu gruptaki hangi kelime olduğuna karar verebilmek için Bölüm 4.8’de anlatıldığı üzere hareketin başında ve sonunda (1. ve 4. anahtar çerçeve) sağ elin biçimine dikkat edilir.

5.1.2 İşaretin Yerel Özelliklere Göre Seçilmesi

Aynı hareketi gerçekleştirirken istemeden elin biçiminde değişiklikler olabilmektedir. Aslında aynı hareket artarda tekrarlandığında elin biçimi her zaman biraz farklı olmaktadır. Bu nedenle, her hareket altışar defa tekrarlanmış bu altı örnek yerel özelliklere göre kelime seçme aşamasında kullanılmıştır. Örneğin, gerçekleştirilen hareketin global özellikleri değerlendirilerek a8 grubuna ait olduğunun belirlendiği varsayılırsa, bu gruba ait olan 26 işaretten hangisinin olduğunu belirlemek için yerel özellikler kullanılmaktadır. Sağ elin biçimi ile veritabanında bu 26 işarete ait olan 156 (26x6) örnek ile karşılaştırılır. Sağlıklı sonuç alabilmek için karşılaştırma yapmadan önce bazı ön işlemler uygulanır (Şekil 5.3).



Şekil 5.3 Karşılaştırma öncesi yapılan ön işlemler

1. ve 4. anahtar çerçevede sağ el bulunarak, elin dışındaki bölgeler görüntüden çıkarılır. 24 bit derinliğindeki renkli görüntü, (5.1)’de verilen rengin parlaklık bileşeni (Y) kullanılarak 256 gri seviyeye indirgenir.

$$Y = 0.299R + 0.587G + 0.114B \quad (5.1)$$

Kameraya olan uzaklık nedeniyle zaman zaman elin boyutu değiştiği ve farklı işaretlerde elin boyutu farklı olduğu için, bir boyut denkleştirmesi yapılarak elin dış çerçevesi 28x28

ebatlarında olacak şekilde yeniden boyutlandırılır. Ön işlemler sonunda, elde edilen 1. ve 4. anahtar çerçevedeki görüntü eğitim setindeki örnekler ile piksel piksel karşılaştırılır.

İlk adımda saklı Markov modelden elde edilen sonuçlara göre seçilen global özellik grubunda H adet farklı hareket olduğu ve eğitim verisinde bu H hareketten sadece 1 örnek olduğu varsayılırsa, $p_1^o(i, j)$ ve $p_4^o(i, j)$ ön işlemler sonucu elde edilip tanınmaya çalışılan harekete ait 1. ve 4. çerçevedeki sağ el görüntüsü olmak üzere, $p_1^o(i, j)$ ve $p_4^o(i, j)$ gruptaki H işarete ait örneklerle (5.2)'de gösterildiği gibi karşılaştırılır .

$$\left. \begin{aligned} \Delta c_1^x &= \sum_{i=1}^{28} \sum_{j=1}^{28} |p_1^x(i, j) - p_1^o(i, j)| \\ \Delta c_4^x &= \sum_{i=1}^{28} \sum_{j=1}^{28} |p_4^x(i, j) - p_4^o(i, j)| \\ \Delta c^x &= \Delta c_1^x + \Delta c_4^x \end{aligned} \right\} x = 1, K, H \quad (5.2)$$

Karşılaştırma işlemi 256 gri seviyeli pikseller arasında basit fark alma ile sağlanır. 1. ve 4. anahtar çerçeve için hesaplanan fark değerleri toplanır ve Δc^x , $x = 1, K, H$ elde edilir. Böylece tanınmaya çalışılan p^o hareketinin mevcut global özellik grubundaki H üyeye ne kadar benzediği bulunmuş olur.

Burada belirtmek gerekir ki, sağlıklı bir tanıma yapabilmek için eğitim verisinde global özellik gruplarındaki işaretlerden 1'den fazla örnek saklamak gereklidir. Bu durumda, H üyeden hangisinin en çok p^o 'a benzediğine, yani hangi işarete karşı geldiğine ise k en yakın komşu algoritması (Duda vd., 2001) kullanılarak karar verilir.

s, ilk aşamada belirlenen global özellik grubundaki her hareketten eğitim verisinde kaç tane örnek olduğunu göstermek üzere, k en yakın komşu algoritmasının adımları aşağıda verilmiştir.

- p^o ile gruptaki $H \cdot s$ adet örnek arasındaki fark (5.2)'de verilenlere göre hesaplanır.
- Δc^x , $x = 1, K, H \cdot s$ farkları küçükten büyüğe doğru sıralanır.
- k adet en küçük fark seçilir.
- Seçilen k adet en yakın komşuda, global özellik grubundaki işaretlerden hangisinden en çok örneği varsa, p^o o harekettir denilir.

Sistem tasarlanırken yapılan denemeler sonucunda s değeri 6 olarak ve k değeri de 5 olarak belirlenmiştir.

5.2 Kelimeleri Ayrı Ayrı Tanıma

Türk İşaret Dili Tanıma Sistemi'nin çalışma şeklini temelde iki gruba ayırabilir. Bunlardan ilki kelimeleri ayrı olarak tanımadır. Bu uygulamada işareti gerçekleştiren kişi işaret dili hareketlerini ayrı ayrı yapmaktadır (birbiri ardına değil) ve bu nedenle pratikte kullanılabilecek bir çalışma şekli değildir. Ancak sistemin izole olarak hareketleri tanıma başarısını görebilmek ve tabii ki cümle tanıyabilen bir sistem geliştirebilmek için mutlaka dikkatlice incelenmesi gereken bir adımdır. Sistemin bu şekilde çalışırken sonuca gitme adımları aşağıda maddeler halinde verilmiştir:

- Anahtar çerçeveler seçilir.
- Ellerin ve yüzün yerleri bulunur.
- 4 anahtar çerçeveden global özellikler çıkarılır.
- Global özellikleri içeren gözlem matrisi saklı Markov modellere verilir ve hareketin ait olduğu global özellik grubu belirlenir.
- Seçilen global özellik grubunda eğer birden fazla üye varsa, yerel özellikler ile k en yakın komşu algoritması kullanılarak gerçekleştirilen hareketin hangi işarete denk geldiği belirlenir.

5.3 Cümle Tanıma

Gerçek anlamda kullanılabilen bir sistemin birbiri ardına yapılan işaretleri tanınması gerekmektedir. Bunu sağlayabilmek için işareti gerçekleştiren kişinin aşağıda belirtilen varsayımlara uygun olarak hareket etmesi gerekmektedir:

- İşaretler arasında eller durur.
- Bir işaret bittikten sonra, yeni işarete geçmek için eller yeni konum alır.

Ellerin durgun kabul edilebilmesi için, birbirini takip eden çerçevelerde ellerin belirlenen Δt süresince sabit kalması gerekmektedir. Sabit kalma durumu, ellerin merkezlerinin aynı kalması ya da ellere ait dış çerçevenin boyutunun değişmemesi olarak tanımlanmaktadır.

Sistem cümle içindeki bir işareti tanıırken izlediği adımlar şunlardır:

- Eller Δt süresince sabit kalıyorsa: İşaret başlar.
- Eller Δt süresince sabit kalıyorsa: İşaret sona erer.
- İşaret için anahtar çerçeveler seçilir ve Bölüm 5.2’de anlatılan tanıma işlemleri yapılır.
- Eller sonraki işareti gerçekleştirmek için yeni konumuna ilerler.

Bu 4 adım birbirini takip eden işaretler için tekrarlanarak devam eder.

6. DENEYSEL TANIMA SONUÇLARI

Türk İşaret Dili Tanıma Sistemi hareket değerlendirme işlemini iki sınıflandırmayı arka arkaya gerçekleştirerek yaptığı için, sistemin test edilmesi de iki aşamalı olarak gerçekleştirilmiştir. İlk olarak, sistemin tanıyabildiği 172 hareket global özelliklerine göre sınıflandırılmış ve test sonuçları verilmiştir. İkinci aşamada, global özellik grupları içindeki hareketler kendi içlerinde yerel özelliklere göre sınıflandırılmış ve test sonuçları verilmiştir.

6.1 Kelimelerin Global Özellikler Kullanılarak Tanınması

Sistemin, işaret dili hareketlerini sadece global özellikleri kullanarak sınıflandırma performansını belirlemek için 172 işarete ait hareketler bir kişi tarafından onar defa tekrarlanmıştır. Böylece, elde edilen 1720 örneğin 1032 (6x172) tanesi 45 farklı saklı Markov modeli eğitmek için, geriye kalan 688 (4x172) tanesi sistemi test etmek için kullanılmıştır. Çizelge 6.1’de eğitim verisine ait ayrıntılı bilgiler verilmektedir.

Çizelge 6.1 Eğitim verisinin özellikleri

İşaret türü	Tek eli	İki eli
Global özellik grubu sayısı	20	25
Toplam işaret sayısı	78	94
Toplam örnek sayısı	780	940
Eğitim için kullanılan örnek sayısı	468	564
Test için kullanılan örnek sayısı	312	376

Sistemin tek ve iki elle gerçekleştirilen hareketleri sınıflandırma yeteneği öncelikle ayrı ayrı incelenmiştir. Tek eli hareketlerin oluşturduğu 20 global özellik grubuna denk gelen 20 saklı Markov model 78 farklı işarete ait 468 eğitim örneği ile eğitilip, 312 test örneği ile de test edilmiştir. Eğitim ve test aşamalarında (6.1)’de verilen global özellikler kullanılmıştır.

$$h = \begin{bmatrix} x_R^1 & y_R^1 \\ \Delta x_R^2 & \Delta y_R^2 \\ \Delta x_R^3 & \Delta y_R^3 \\ \Delta x_R^4 & \Delta y_R^4 \end{bmatrix} \quad (6.1)$$

Burada başarılı sınıflandırmadan kasıt, işaretin ait olduğu global özellik grubunu doğru belirleyebilmektir. Yapılan testlerde elde edilen sonuçlar Çizelge 6.2’de verilmiştir.

Çizelge 6.2 Tek eli hareketlere ait global özellik grubu belirleme sonuçları

Global Özellik Grubu	Gruptaki Kelime Sayısı	Test Örneği Sayısı	Doğru Sınıflandırılan Örnek Sayısı	Başarı Oranı
a1	4	16	12	75,00%
a2	1	4	3	75,00%
a4	3	12	12	100,00%
a9	9	36	34	94,44%
b2	19	76	71	93,42%
b4	2	8	8	100,00%
b6	2	8	8	100,00%
b7	1	4	4	100,00%
b8	1	4	4	100,00%
b9	2	8	8	100,00%
c3	2	8	8	100,00%
c4	1	4	4	100,00%
c6	1	4	4	100,00%
c7	1	4	3	75,00%
c8	24	96	93	96,88%
d8	1	4	4	100,00%
e1	1	4	3	75,00%
e2	1	4	4	100,00%
e5	1	4	4	100,00%
e9	1	4	4	100,00%
Toplam:		312	295	94,55%

312 test örneğinin 295 tanesinin global özellik grubu doğru sınıflandırılmış ve %94,55 başarı elde edilmiştir. Sistemin verdiği hatalı sonuçlar değerlendirilmiş ve yanlış sınıflandırılan gruplar Çizelge 6.3'te verilmiştir.

Çizelge 6.3 Hatalı sınıflandırılan gruplar (tek eli hareketler)

Gerçek grup	a1	a1	a1	a1	a2	a9	a9	b2	b2	b2	b2	b2	c7	c8	c8	c8	e1
Hatalı sonuç	a2	e1	e1	e1	a1	b2	b2	a9	a9	a9	b8	a4	c6	a9	b7	b9	a1

Hatalı sınıflandırılan gruplar ayrıntılı olarak incelendiğinde, hataların büyük bir bölümünün birbirine yakın global özellik gruplarından kaynaklandığı görülmektedir. Örneğin, a1 ile e1 grupları karşılıklı olarak en çok hatalı sınıflandırılan gruplardır ve bu iki grubun global özellikleri Şekil 2.4'te incelendiğinde elin harekete yaklaşık benzer konumda başladığı ve devamında yakın hareketler yaptığı görülmektedir. Sistemin hatalı değerlendirmeler

yapmasının bir diğer nedeni de işareti gerçekleştiren kişinin bazı hareketleri hatalı gerçekleştirmesinden kaynaklanmaktadır. Bu denemeler sonucunda Türk İşaret Dili Tanıma Sistemi'nin yaklaşık %95 başarı oranı ile tek elli hareketleri kendi içinde gruplandırabildiği görülmüştür.

Sistemin tanıyabildiği işaretlerin büyük bir kısmını iki elle gerçekleştirilen hareketler oluşturmaktadır. İki elli hareketlerin oluşturduğu 25 global özellik grubuna denk gelen 25 saklı Markov model 94 farklı işarete ait 564 eğitim örneği ile eğitilip, 376 test örneği ile de test edilmiştir. Her iki elde işaret sırasında konum değiştirdiği için sistemin eğitim ve test aşamasında (4.12)'de verilen global özellik matrisi kullanılmıştır. Yapılan testlerde elde edilen sonuçlar Çizelge 6.4'de verilmiştir.

Çizelge 6.4 İki elli hareketlere ait global özellik grubu belirleme sonuçları

Global Özellik Grubu	Gruptaki Kelime Sayısı	Test Örneği Sayısı	Doğru Sınıflandırılan Örnek Sayısı	Başarı Oranı
a3	1	4	3	75,00%
a5	2	8	6	75,00%
a6	2	8	7	87,50%
a7	1	4	3	75,00%
a8	26	104	103	99,04%
b1	8	32	27	84,38%
b3	1	4	4	100,00%
b5	1	4	4	100,00%
c1	6	24	20	83,33%
c2	3	12	10	83,33%
c5	2	8	8	100,00%
c9	1	4	3	75,00%
d1	1	4	3	75,00%
d2	3	12	10	83,33%
d3	1	4	4	100,00%
d4	4	16	16	100,00%
d5	3	12	12	100,00%
d6	6	24	23	95,83%
d7	2	8	6	75,00%
d9	1	4	4	100,00%
e3	10	40	36	90,00%
e4	4	16	16	100,00%
e6	3	12	12	100,00%
e7	1	4	3	75,00%
e8	1	4	4	100,00%
Toplam:		376	347	92,29%

376 test örneğinin 347 tanesinin global özellik grubu doğru sınıflandırılmış ve %92,29 başarı elde edilmiştir. Sistemin verdiği hatalı sonuçlar değerlendirilmiş ve yanlış sınıflandırılan gruplar Çizelge 6.5’te verilmiştir.

Çizelge 6.5 Hatalı sınıflandırılan gruplar (iki elli hareketler)

Gerçek Grup	a3	a5	a5	a6	a7	a8	b1	b1	b1	b1	b1	c1	c1	c1	c1
Hatalı Sonuç	d7	a6	d7	a5	d1	e4	c1	c1	c2	e7	e7	b1	b1	c2	e7
Gerçek Grup	c2	c2	c9	d1	d2	d2	d6	d7	d7	e3	e3	e3	e3	e7	-
Hatalı Sonuç	b1	b1	e3	a7	d1	a7	a8	a6	a6	c9	c9	c9	c9	b1	-

Tek elli hareketlere benzer şekilde, hataların büyük bir kısmının birbirine yakın global özellik gruplarından kaynaklandığı görülmüştür. Örneğin, e3 ile c9 grupları en çok karıştırılan gruplar olmuştur. Bunun dışında, hataların bir kısmı işareti gerçekleştiren kişinin yaptığı yanlış hareketlerden kaynaklanabilmekte bir diğer kısmı da çerçevelerde ellerin kaybedilmesine bağlı olarak, çok farklı anahtar çerçeve seçilmesi nedeniyle meydana gelebilmektedir. İki elli hareketler ile ilgili denemeler sonucunda Türk İşaret Dili Tanıma Sistemi’nin yaklaşık %92 oran ile bu hareketleri kendi içinde başarıyla gruplandırabildiği görülmüştür.

Her iki tür işaret türü incelendiğinde sınıflandırma başarısının iki elli hareketlerde biraz daha düşük olması beklenen bir sonuçtur. Çünkü iki elli hareketler tek elli hareketlere göre daha karmaşık bir yapıya sahiptir.

Gerçek zamanlı çalışan bir sistem de tek ve iki elli işaretleri ayrı ayrı değerlendirmek için yapılan işaretin ne tür olduğunu belirleyebilen bir ön işlem birimi kullanılabilir. Ön işlem birimi biraz zaman kaybına neden olsa da, farklı iki global özellik matrisinin kullanılmasına olanak sağlayacaktır. Bunun yanı sıra, sistem tasarlanırken tek ve iki elli işaretleri beraber değerlendirilmesinin sistem performansına etkisi incelenmiş ve bu nedenle tek elli işaretler için sadece sağ el kullanıldığından (6.2)’de verilen global özellik matrisi kullanılmıştır.

$$h = \begin{bmatrix} x_R^1 & y_R^1 & x_L^1 & y_L^1 \\ \Delta x_R^2 & \Delta y_R^2 & 0 & 0 \\ \Delta x_R^3 & \Delta y_R^3 & 0 & 0 \\ \Delta x_R^4 & \Delta y_R^4 & 0 & 0 \end{bmatrix} \quad (6.2)$$

Tek elli işaretlerde sol elin hareketi olmadığı için konum değiştirmeler sıfır olarak alınmıştır. Bu aşamada tüm hareketlerin oluşturduğu 45 global özellik grubuna denk gelen 45 saklı Markov model 172 farklı işarete ait 1032 eğitim örneği ile eğitilip, 688 test örneği ile de test edilmiş ve elde edilen sonuçlar Çizelge 6.6’da verilmiştir.

Çizelge 6.6 Tüm işaretlerin beraber değerlendirilmesi

İşaret Sayısı	172
Test örneği sayısı	688
Doğru Sınıflandıran Örnek Sayısı	642
Başarı Oranı	93,31%

Hatalı sınıflandırılan 46 test örneği ayrıntılı olarak incelendiğinde, Çizelge 6.3 ve Çizelge 6.5’te belirtilen sonuçların biri bir aynısı olduğu görülmüştür. Bu sonuçtan, sistemin tek ve iki elli hareketleri beraber sınıflandırmasının herhangi bir ek hataya neden olmadığı ve %93,31 başarı oranı ile global özellik gruplarının belirlendiği görülmüş ve işaretin türünü belirleyen bir ön birimin kullanılmasına gerek olmadığına karar verilmiştir.

6.2 Kelimelerin Yerel Özelliklerin Kullanılarak Tanınması

Önceden global özelliklerine göre doğru sınıflandırıldığı varsayılan ortak global özelliğe ait hareketler kendi içlerinde yerel özelliklerine göre sınıflandırılmıştır. Aynı global özelliğe sahip birden fazla kelimenin olduğu gruplarda, yerel özellikler kullanılarak işaret edilen kelime tanınmaya çalışılmıştır.

Birden fazla üyeye sahip 24 global özellik grubundaki 151 işaretin her biri işareti gerçekleştiren kişi tarafından onar kere tekrarlanmış ve 1510 örneğin 906 tanesi eğitim amacı ile 604 tanesi de test işlemlerinde kullanılmıştır. Eğitim ve test verisine ait ayrıntılı bilgiler Çizelge 6.7’de verilmiştir.

Çizelge 6.7 Yerel özelliklerin kullanılması

Global Özellik Grubu	İşaret Türü	Gruptaki Kelime Sayısı	Toplam Örnek Sayısı	Eğitim Örneği Sayısı	Test Örneği Sayısı
a1	Tek elli	4	40	24	16
a4	Tek elli	3	30	18	12
a5	İki elli	2	20	12	8
a6	İki elli	2	20	12	8
a8	İki elli	26	260	156	104
a9	Tek elli	9	90	54	36
b1	İki elli	8	80	48	32
b2	Tek elli	19	190	114	76
b4	Tek elli	2	20	12	8
b6	Tek elli	2	20	12	8
b9	Tek elli	2	20	12	8
c1	İki elli	6	60	36	24
c2	İki elli	3	30	18	12
c3	Tek elli	2	20	12	8
c5	İki elli	2	20	12	8
c8	Tek elli	24	240	144	96
d2	İki elli	3	30	18	12
d4	İki elli	4	40	24	16
d5	İki elli	3	30	18	12
d6	İki elli	6	60	36	24
d7	İki elli	2	20	12	8
e3	İki elli	10	100	60	40
e4	İki elli	4	40	24	16
e6	İki elli	3	30	18	12

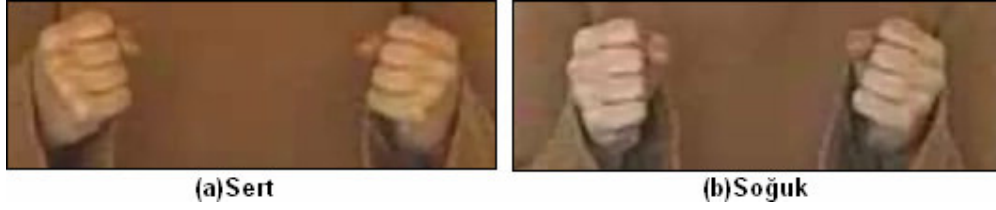
Türk İşaret Dili Tanıma Sistemi'nin yerel özelliklere göre değerlendirme yapması tanıma işleminin ikinci aşaması olduğu için, bir önceki aşamada işaretlerin global özelliklerine göre doğru gruplandırıldığı kabul edilip, sistemin yerel özellikler üzerindeki başarısı incelenmiştir. Bu nedenle hareketler bulundukları global özellik grubu içindeki üyeler ile değerlendirilmiştir. Sağ elin 1. ve 4. anahtar çerçevedeki durumuna göre hangi hareketin yapıldığı anlaşılmaya çalışılmış ve k en yakın komşu algoritmasından faydalanılmıştır.

Aynı global özellik grubunda yer alan işaretler için yapılan testlerde elde edilen sonuçlar Çizelge 6.8'de verilmiştir.

Çizelge 6.8 Yerel özelliklere ait test sonuçları

Global Özellik Grubu	Gruptaki Kelime Sayısı	Test Örneği Sayısı	Doğru Sınıflandırılan Örnek Sayısı	Başarı Oranı
a1	4	16	15	93,75%
a4	3	12	12	100,00%
a5	2	8	8	100,00%
a6	2	8	8	100,00%
a8	26	104	92	88,46%
a9	9	36	32	88,89%
b1	8	32	29	90,63%
b2	19	76	65	85,53%
b4	2	8	8	100,00%
b6	2	8	8	100,00%
b9	2	8	8	100,00%
c1	6	24	22	91,67%
c2	3	12	11	91,67%
c3	2	8	8	100,00%
c5	2	8	8	100,00%
c8	24	96	85	88,54%
d2	3	12	12	100,00%
d4	4	16	16	100,00%
d5	3	12	12	100,00%
d6	6	24	21	87,50%
d7	2	8	8	100,00%
e3	10	40	37	92,50%
e4	4	16	15	93,75%
e6	3	12	12	100,00%
Toplam:		604	552	91,39%

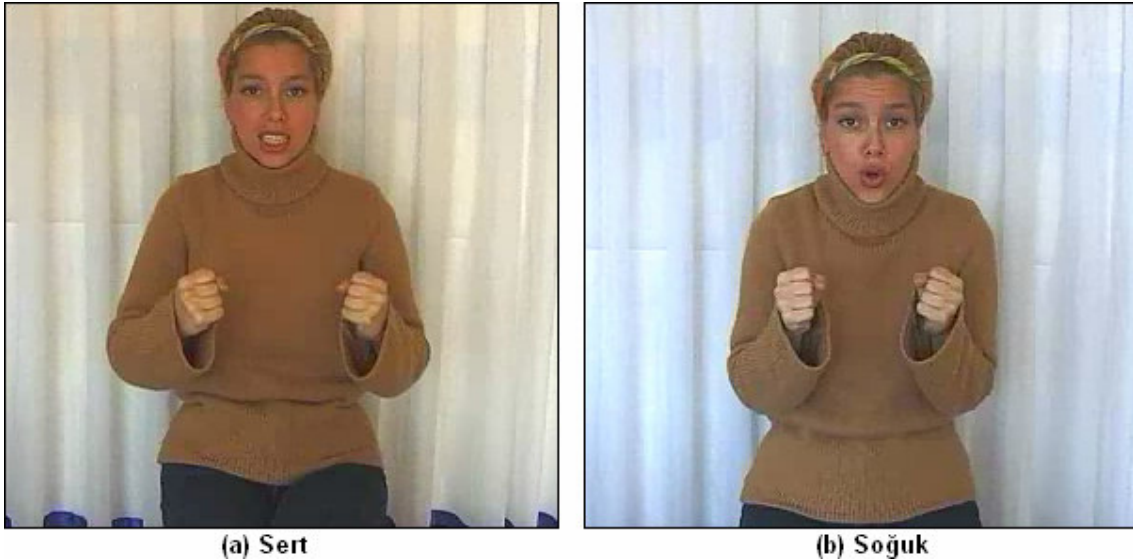
Elde edilen sonuçlar incelenmiş ve hataların aynı grup içindeki benzer el duruşları nedeniyle kaynaklandığı saptanmıştır. Örneğin 26 işareti içeren a8 global özellik grubuna ait 104 test örneğinin 12 tanesi hatalı sınıflandırılmıştır. Bu gruptaki işaretlerin yerel özellikleri dikkatlice incelenmiş ve yerel özellikleri birbirine çok yakın olan “sert” ve “soğuk” kelimelerine ait işaretler Şekil 6.1’de gösterilmiştir.



Şekil 6.1 Yakın global ve yerel özelliklere sahip işaretler

Şekil 6.1 incelendiğinde belirlenen özellikler ile tez çalışmasında tasarlanan Türk İşaret Dili Tanıma Sistemi'nin bu iki işareti ayırt edemeyeceği kolayca görülmektedir. Burada akla gelen ilk soru iki işaretin bu kadar yakın olmasının nasıl mümkün olduğudur.

Ancak, unutulmamalıdır ki işaret dili hareketlerine anlam veren öğeler elin şekli, elin hareketi, vücut hareketi ve yüz ifadesidir. Bu tez çalışması kapsamında sadece elin hareketi ile şekli incelendiği ve yüzdeki ifadeyle ilgilenilmediği için bu iki işaret ayırt edilememiştir. Şekil 6.2'ye bakılacak olursa yüz ifadesinin işaretlere kattığı anlam çok daha rahat anlaşılacaktır.



Şekil 6.2 Yüz ifadesinin ve dudakların işaretlere kattığı anlam

Bu tür işaretlerin ayırt edilebilmesi için sistemin yüz ifadelerini de değerlendirir hale getirilmesi gerekmektedir, ancak tez çalışması kapsamında bu uygulama yapılmayacaktır.

7. SONUÇLAR VE ÖNERİLER

Türk İşaret Dili'ne ait herhangi bir hareket veritabanı olmadığı için, sistemin tanıdığı işaretler sistemi tasarlayan kişi tarafından gerçekleştirilmiştir. Bu nedenle, işaretleri öğrenme ve tekrarlama aşaması oldukça uzun zaman almaktadır. İşitme engelli olmayan bireyler için gerçekleştirilen işaretler zihinde yeterince yer etmediği için zaman zaman işaretlerin hatalı yapılabildiği gözlemlenmiştir. Buna bağlı olarak farklı zamanlarda, aynı özellikleri test ederken farklı sonuçlar alınabilmektedir. Ayrıca, işaretleri sadece bir kişinin gerçekleştirmesinin sonucu olarak tasarlanan sistem kişiye bağımlı kalmaktadır. Oysaki ticari anlamda kullanılan bir sistemin mutlaka kişiden bağımsız olması gereklidir. Bu nedenle Türk İşaret Dili uzmanları ya da eğitmenleri ile beraber ortak çalışma yapılmalı ve böylece kişiden bağımsız tanıma yapabilen sistemlerin tasarlanabilmesi sağlanmalıdır.

Bazı ülkelerin işaret dillerine ait hareket veritabanları mevcuttur. Bir araştırma projesi kapsamında, çok sayıda Türk İşaret Dili hareketi içeren ve aynı işaretin farklı kişilerce tekrarlanması sonucu oluşturulacak bir veritabanı hem araştırmacıların işlemlerini kolaylaştırılacak hem de Türk İşaret Dili tanıma sistemlerin tasarlanabilmesinin yolunu açacaktır. Her ülkenin kendine has bir işaret dili olduğu için Türk İşaret Dili Tanıma Sistemleri konusundaki gelişmelerin kendi öz çalışmamıza ve kaynaklarımıza bağlı olduğu unutulmamalıdır.

Görüntü tabanlı işaret dili tanıma sistemlerinin birçoğunda kullanıcının çeşitli renklerde plastik yada yün eldivenler giydiği görülmektedir. Ancak bu yaklaşım her ne kadar hastane, banka, vb. gibi alanlarda kullanılan bir sistem için çok pratik bir yol gibi gözükmesine de, kullanılmadıkları zaman da sistemin tanıyabildiği hareketlere büyük sınırlamalar getirmektedir. Bu nedenle, sistem tasarlanırken eldiven kullanılıp kullanılmayacağına iyi karar verilmelidir. Eğer tasarlanan sistemde, ellere herhangi bir eldiven giyilmemesi isteniyorsa, sistemin sahip olacağı kısıtlamaları aşmak için el ve kolların eklem bağlantıları bilgilerinden faydalanılabilir. Tabii bu eklemeler sistemin yavaşlaması sonucunu ortaya çıkarabilir. Bilgisayar sistemlerinin, büyük bir ivme ile gelişen donanım teknolojilerine bağlı olarak, her geçen gün hızının artması sadece görüntü işleyerek herhangi bir eldiven kullanılmadan gerçek zamanlı çalışabilen sistemlerin tasarlanabileceği konusunda umut vermektedir.

Tez çalışmasında tasarlanan sistemde tek bir kamera kullanıldığı için gerçekleştirilen hareketlerin derinlik bilgisi kaybolmakta ve bazı hareketlerin karakteristik özellikleri gözden

kaçabilmektedir. Buna ek olarak, yapılan çeşitli denemelerde, işareti gerçekleştiren kişinin karşısında duran bir kameranın zaman zaman elin duruşuna ait bazı bilgileri elde edemediği görülmüştür. Bu nedenle, uygun açılırla sisteme yerleştirilen iki adet kamera kullanılarak hem derinlik bilgisi elde edilebilir hem de elin başka açılardan duruşunun incelenmesi sağlanabilir. Sistemin gerçek zamanlı çalışmasını sağlayabilmek için uzun görüntü iyileştirme çalışmalarından uzak durulmalıdır ve bunu yapabilmek için de sistem de mutlaka CCD kameralar kullanılmalı ve ortamın aydınlatma şartları dikkatle sağlanmalıdır.

Türk İşaret Dili Tanıma Sistemi iki safhalı bir tanıma gerçekleştirdiği için ilk aşamada yapılan bir hata, doğru sınıflandırmayı tamamen imkânsız bir hale getirmektedir. Yapılan testlerde ya işareti gerçekleştiren kişiye bağlı olarak yada birbirine yakın global özellik grupları nedeniyle hatalar olabildiği görülmektedir. Birbirine yakın grupları birleştirmek bir çözüm olarak düşünülse de, aynı grupta fazla sayıda üye olması ikinci aşamanın çok zaman almasına neden olabilmektedir. Bu nedenle, global özellik grubu belirleme aşamasında sadece saklı Markov modelin ürettiği en yüksek olasılık olan grubu seçmek yerine, belirli hata oranlarına bağlı olarak en yüksek olasılığa sahip birkaç tanesini belirlenebilir. Belirlenen gruplar ile yerel özelliklerin uygun şekilde kullanılmasıyla global özellik grubu seçilebilir.

Yapılan testlerde aynı global özellik grubu içinde yakın yerel özelliklere sahip işaretlerin sağlıklı olarak tanınmadığı görülmüştür. Tabii bu noktada şu hatırlatılmalıdır ki, bazı işaretlerin anlamına yüz ifadesinin, dudakların ve vücut duruşunun büyük etkisi olmaktadır. Tez çalışması kapsamında tasarlanan Türk İşaret Dili Tanıma Sistemi'nin mevcut çalışma yöntemi ile aynı global özellik grubundaki yakın yerel özelliklere sahip işaretleri büyük doğrulukla ayırt edebilmesi mümkün değildir. Bu nedenle, sistemin tanıdığı kelime sayısı arttıkça yüz ifadesinin ve dudakların durumu da incelenmelidir. Ancak, yüz ifadesini ayrıntılı olarak değerlendirme bugünün donanım koşulları ile gerçek zamanlı çalışan bir sistem tasarlamayı güçleştirmektedir.

Birbiri ardına gerçekleştirilen işaretler arasında belirgin duruşların gerekli olması, Türk İşaret Dili Sistemi'nin en dezavantajlı yönü olarak düşünülebilir. Tez çalışması süresince bu kısıtlamanın aşılammamasının en önemli nedeni, şu ana kadar TİD tanıma sistemleri konusunda bir çalışma yapılmamış olması ve özellikle TİD hakkında yeterince kaynak bulunmamasıdır. İşaretler arası durma kısıtlamasını aşabilmek için, öncelikle TİD işaretleri hakkında derin incelemeler ve araştırmalar yapmak gereklidir. Bunun sağlanması için en uygun yol bir araştırma projesi kapsamında çalışmalar yapmaktır. Türk İşaret Dili hareketlerinin özellikleri ayrıntılı bir şekilde belirlendikten sonra, uygun saklı Markov modelleri seçilerek, işaretler

arasında durmaların olmadığı sistemler tasarlamak mümkündür.

Sistemin bazı kısıtlamalara sahip olması ve tanıyabildiği kelime sayısının 172 ile sınırlı olması ticari anlamda kullanılabilir bir sistem tasarlamak için incelenen özellik sayısının arttırılması gerekliliğini ortaya çıkarmaktadır. Bununla birlikte, tez çalışmasında tasarlanan Türk İşaret Dili Tanıma Sistemi'nin TİD konusunda bir ilk olması ve iki aşamalı olarak elde edilen tanıma sonuçlarının yaklaşık %93 ve %91 oranlarında olması nedeniyle ileride yapılacak çalışmalara yön vereceği umut edilmektedir.

KAYNAKLAR

- Alpaydın, E., (2004), Introduction to Machine Learning, The MIT Press, Cambridge.
- Bayazıt, M. ve Oğuz, B., (1994), Mühendisler İçin İstatistik, Birsen Yayınevi, İstanbul.
- Cerit, C. ve Yüksel, M., (1998), Olasılık Problemleri, İstanbul Teknik Üniversitesi Fen-Edebiyat Fakültesi, İstanbul.
- Chang, C. ve Pengwu, C., (2004), “Gesture Recognition Approach for Sign Language Using Curvature Scale Space and Hidden Markov Model”, IEEE International Conference on Multimedia and Expo, ICME 2004, 1187–1190.
- Chena, Y., Gao, W., Fang, G., Yang, C. ve Wang, Z., (2003), “CSLDS: Chinese Sign Language Dialog System”, IEEE International Workshop on Analysis and Modeling of Faces and Gestures, 17 Oct. 2003, 236 – 237.
- Duda, O.D., Hart, P.E. ve Stork D.G. (2001), Pattern Classification 2nd Edition, Wiley-Interscience Publication, USA.
- Erhardt-Ferron, A., (2000), Theory and Applications of Digital Image Processing, University of Applied Sciences, Offenburg.
- Gonzalez, R.C. ve Woods, R.E., (1993), Digital Image Processing, Addison-Wesley, Massachusetts.
- Grobel, K. ve Assan, M., (1997), “Isolated Sign Language Recognition using Hidden Markov Models”, International Conference on Computational Cybernetics and Simulation, 12-15 Oct. 1997, 1:162–167.
- Hai, W. ve Sutherland, A., (2004), “Irish Sign Language Recognition Using PCA, Multi-scale Theory, and Discrete Hidden Markov Models”, Proceedings of The Irish Machine Vision Conference, 2004, 66–73.
- Hsu, R., Abdel-Mottaleb, M. ve Jain, A.K., (2002), “Face Detection in Color Images”, Pattern Analysis and Machine Intelligence, 24(5): 696–706.
- Imagawa, K., Matsuo, H., Taniguchi, R., Arita, D., Lu, S. ve Ilgi, S., (2000), “Recognition of Local Features from Camera-based Sign Language Recognition System”, 15th International Conference on Pattern Recognition, 3-7 Sept. 2000, 4:849–853.
- Liang, H. ve Ouhyoung, M., (1996), “A Sign Language Recognition System Using Hidden Markov Model and Context Sensitive Search”, ACM VRST.
- Mohandes, M., A-Buraiky, S., Halawani, T. ve Al-Baiyat, S., (2004), “Automation of the Arabic Sign Language Recognition”, Int. Conf. on Information and Communication Technologies: From Theory to Applications, 19-23 Apr. 2004, 479–480.
- Russ, J.C., (1999), The Image Processing Handbook 3rd Edition, CRC Press, Florida.
- Sangwine, S. J. ve Horne, R. E. N., (1998) , The Colour Image Processing Handbook, Chapman&Hall, London.
- Starner, T., (1995), Visual Recognition of American Sign Language Using Hidden Markov Models, Master’s Thesis, MIT, Media Laboratory.
- Takahashi, T. ve Kishino, F., (1991), “Hand Gesture Coding Based on Experiments Using a Hand Gesture Interface Device”, SIGCHI Bulletin, 23(2):67–73.

Terrillon, J.C., Shirazi, M. N., Fukamachi, H. ve Akamatsu, S., (2000), “Comparative Performance of Different Skin Chrominance Models and Chrominance Spaces for the Automatic Detection of Human Faces in Color Images,” Proc. IEEE International Conference on Face and Gesture Recognition, 54-61.

Yamato, J., Ohya, J. ve Ishii, K., (1992), “Recognizing Human Action in Time-sequential Images Using Hidden Markov Models”, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 1992, 379–385.

Yetişkinler İçin İşaret Dili Kılavuzu, T.C. Milli Eğitim Bakanlığı Özel Eğitim ve Rehberlik ve Danışma Hizmetleri Genel Müdürlüğü, 1995, Ankara.

Yoshino, K., Kawashima, T. ve Aoki, Y., (1996), “Recognition of Japanese Sign Language From Image Sequence Using Color Combination”, International Conference on Image Processing, 1996, 3:511 – 514.

Young, S., Evermann, G., Hain, T., Kershaw, D., Moore, G., Odell, J., Olason, D., Povey D., Valtchev, V. ve Woodland, P., (2002), The HTK Book, Cambridge University Engineering Department.

Zhao, Y., Zhang, Z. ve Gao. Z., (2004), “A Simple and Workable Moving Objects Segmentation Method”, ELMAR, June 2004, Croatia.

INTERNET KAYNAKLARI

[1] <http://turkisaretdili.ku.edu.tr>

[2] <http://www.philips.com/pcestuff>

[3] <http://www.turk.ch>

EKLER

Ek 1a	Temel Olasılık Kavramları
Ek 1b	Koşullu Olasılık
Ek 1c	Olasılık Yoğunluk Fonksiyonu
Ek 1d	Beklenen Değer
Ek 1e	Lineer Bağımlılık
Ek 1f	Normal (Gauss) Dağılım
Ek 1g	Çoklu Rastlantı Değişkeni
Ek 1h	Parametre Kestirimi

Ek 1a Temel Olasılık Kavramları

Bir denemenin muhtemel sonuçları olan eşit şanslı basit olaylara karşı gelen S örnek uzayı N adet noktadan oluşuyorsa ve bunlardan n tanesi A olayına karşı geliyorsa A olayının P(A) ile gerçekleşme olasılığı (Ek 1.1) ile tanımlanır.

$$P(A) = \frac{n}{N}, \quad 0 \leq P(A) \leq 1 \quad (\text{Ek 1.1})$$

Ek 1b Koşullu Olasılık

A ve B herhangi iki olay ve P(B)>0 olsun. Önce B olayının olması koşulu ile A olayının gerçekleşmesinin koşullu olasılığı P(A|B) ile ifade edilir ve (Ek 1.2) ile tanımlanır.

$$P(A | B) = \frac{P(AB)}{P(B)} \quad (\text{Ek 1.2})$$

A ve B olaylarının simetrik konumda olmalarını dikkate alarak P(A)>0 olur (Cerit, 1998) ve bu durum (Ek 1.3)'de verilmiştir.

$$P(A, B) = P(B)P(A | B) = P(A)P(B | A) \quad (\text{Ek 1.3})$$

Ek 1c Olasılık Yoğunluk Fonksiyonu

X rastlantı değişkeni sürekli olsun ve bir [a,b] kapalı aralığı içinde değerler alsın. Bu aralıkta X rastlantı değişkeninin alabilmesi olası sonsuz sayıda değer vardır. Bunlardan özel olarak bir tanesini, örneğin bir $c \in [a, b]$ değerini alması, olası olmasına karşın, olasılığı sıfır olacaktır. Bunun yanı sıra X rastlantı değişkeninin [a, b]'nin herhangi bir alt aralığında yer alması söz konusu olabilir. $a \leq c \leq d \leq b$ olsun, $x \in [c, d]$ olmasının olasılığını veren (Ek 1.4)'teki f(x)

$$P(c \leq X \leq d) = \int_c^d f(x)dx \quad (\text{Ek 1.4})$$

fonksiyonuna, X rastlantı değişkeninin olasılık yoğunluk fonksiyonu denilir. Her f(x) olasılık yoğunluk fonksiyonu (Ek 1.5)'de verilen iki koşulu gerçekler (Cerit, 1998).

$$(i) f(x) \geq 0$$

$$(ii) \int_a^b f(x)dx = 1 \quad (\text{Ek 1.5})$$

Ek 1d Beklenen Değer

Kesikli X rastlantı değişkeninin alması olası değerler $x_1, x_2, \dots, x_N$ olsun, X'in olasılık fonksiyonu $P(X = x_i) = f(x_i)$, $i = 1, 2, 3, \dots, N$ ile formüle edilsin. X rastlantı değişkeninin beklenen değeri (Ek 1.6) ile tanımlanır.

$$E[X] = \sum_{i=1}^N x_i \cdot f(x_i) \quad (\text{Ek 1.6})$$

X rastlantı değişkeni N değeri eşit olasılıklarla alıyorsa (Ek 1.7)'de belirtilen durum olur.

$$E[X] = \frac{1}{N} \sum_{i=1}^N x_i \quad (\text{Ek 1.7})$$

X rastlantı değişkeni sürekli ve olasılık yoğunluk fonksiyonu $f(x)$ olsun. X'in beklenen değeri (Ek 1.8) ile tanımlanır.

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x) dx \quad (\text{Ek 1.8})$$

Ek 1e Lineer Bağımlılık

$f(x)$ olasılık fonksiyonuna sahip kesikli X rastlantı değişkeninin varyansı σ^2 ya da $\text{Var}(X)$ notasyonu ile gösterilir ve (Ek 1.9) ile tanımlanır.

$$\sigma^2 = E[(X - \mu)^2] = \sum_{i=1}^N (x_i - \mu)^2 f(x_i) \quad (\text{Ek 1.9})$$

Burada μ , X rastlantı değişkeninin alması olası değerlerinin ortalamasıdır ve (Ek 1.10) ile tanımlanır.

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i \quad (\text{Ek 1.10})$$

Varyans, X'in aldığı değerlerin beklenen değer ($\mu \equiv E[X]$) civarında değişimin ölçütüdür. Kovaryans ise iki rastlantı değişkeni arasındaki lineer ilişkiyi ifade eder. Eğer X'in var olması Y'nin var olmasını artırıyorsa, kovaryans pozitif ve X'in var olması Y'nin meydana gelmesini azaltıcı yönde ise kovaryans negatiftir. Eğer aralarında bağımlılık yoksa kovaryans sıfırdır. Öte yandan kovaryansın sıfır olması X ile Y'nin bağımsız olması anlamına gelmez (Alpaydın, 2004). X ile Y rastlantı değişkenleri arasındaki ilişkinin derecesini ölçen ve yönünü gösteren boyutsuz büyüklüğe korelasyon katsayısı denir. Korelasyon katsayısı (Ek 1.11) ile tanımlanır ve $[-1, +1]$ arasında değerler alır.

$$\rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \quad (\text{Ek 1.11})$$

X ve Y gibi iki değişken arasındaki korelasyon katsayısı, bu değişkenler arasındaki lineerlik derecesinin bir ölçütüdür. +1 ve -1'e yakın $\rho(X, Y)$ değerleri yüksek bir lineerlik derecesini, 0'a yakın değerleri ise lineerliğin azaldığını gösterir. Değişkenlerden birinin değeri artarken diğerinin değeri de artıyorsa, iki değişken arasında pozitif korelasyon vardır. Değişkenlerden birinin değeri artarken diğer değişkenin değerinde azalma varsa değişkenler arasında negatif korelasyon vardır. Değişkenlerden birinin değeri değişirken diğerinin değerinde değişme yoksa aralarında korelasyon yoktur denilir (Cerit, 1998). Burada yine hatırlatmak gerekir ki, iki değişken bağımsız olduğunda korelasyonun sıfır olmasına rağmen, korelasyonun sıfır olması iki değişkenin bağımsız olduğunu göstermez. Değişkenler lineer olmayan biçimde bağımlı olabilirler (Alpaydın, 2004). Eğer değişken sayısı ikiden fazla ise değişkenlerden birisi bağımlı, diğerleri bağımsız değişken kabul edilir. Böyle durumlarda ikili olarak değişkenler arasındaki kovaryansa bakılabilir.

Ek 1f Normal (Gauss) Dağılımı

Sürekli X rastlantı değişkeninin olasılık yoğunluk fonksiyonu (Ek 1.12) ile tanımlanır.

$$N(x; \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad -\infty < x < \infty \quad (\text{Ek 1.12})$$

(Ek 1.12) ile verilen fonksiyona normal olasılık yoğunluk fonksiyonu ve X rastlantı değişkenine de normal dağılımlıdır denir (Cerit, 1998). Burada varyansın pozitif kareköküne standart sapma denir ve σ ile gösterilir. Doğal ve toplumsal olaylarda karşılaşılan rasgele değişkenlerden birçoğunun normal dağılıma uyması istatistiğin merkez limit teoremi ile açıklanabilir. Bu teoreme göre bir değişkenin normal dağılması için şu koşulların gerçekleşmesi yeterlidir:

- i. Bu değişkeni çok sayıda etken, etkileri birbiriyle toplanacak şekilde etkilemelidir.
- ii. Bu etkiler birbirinden bağımsız olmalıdır.

Şunu belirtmek gerekir ki, normal dağılmış bir rasgele değişkenin $(\mu - 3\sigma, \mu + 3\sigma)$ aralığının dışında bir değer alması olasılığı ihmal edilebilecek kadar küçüktür (Bayazıt ve Oğuz, 1994).

Ek 1g Çoklu Rastlantı Değişkeni

Birçok uygulamada X örneğine karşı gelen veri matrisi (Ek 1.13)'teki gibi gösterilir.

$$X = \begin{bmatrix} X_1^1 & X_2^1 & \Lambda & X_d^1 \\ X_1^2 & X_2^2 & \Lambda & X_d^2 \\ \text{M} & & & \\ X_1^N & X_2^N & & X_d^N \end{bmatrix} \quad (\text{Ek 1.13})$$

Burada d , seçilen özelliklere, N ise bağımsız gözlem sayısına karşı gelir. Özellikleri oluşturan ölçümler farklı boyutlarda olabilir. Bir tanesi nümerik bir değerken diğeri ayrık bir değer olabilir (Alpaydın, 2004).

Ek 1h Parametre Kestirimi

(Ek 1.14)'de verilen μ , ortalama vektörü olarak adlandırılır. Her elemanı, X matrisinin bir sütununun ortalamasıdır.

$$E[x] = \mu = [\mu_1, \mu_2, \dots, \mu_d]^T \quad (\text{Ek 1.14})$$

X_i 'nin varyansı σ_i^2 ile, X_i ve X_j değişkenlerinin kovaryansını (Ek 1.15) ile ifade edelim.

$$\sigma_{ij} \equiv \text{Cov}(X_i, X_j) = E[(X_i - \mu_i)(X_j - \mu_j)] \quad (\text{Ek 1.15})$$

d değişkeni düşündüğümüzde, d tane varyans ve $d(d-1)/2$ kovaryans tanımlanabilir ve bu değerler genelde Σ ile gösterilen $d \times d$ boyutlu kovaryans matrisinde gösterilir:

$$\Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \Lambda & \sigma_{1d} \\ \sigma_{21} & \sigma_2^2 & \Lambda & \sigma_{2d} \\ \text{M} & & & \\ \sigma_{d1} & \sigma_{d2} & & \sigma_d^2 \end{bmatrix} \quad (\text{Ek 1.16})$$

Kovaryans matrisinde köşegen terimleri varyansları, diğer elemanlar ise kovaryansları verir.

Ortalama vektörünü, μ 'nün kestirilen değerini m ile gösterirsek, m_i elemanları (Ek 1.17) ile verilir.

$$m_i = \frac{\sum_{t=1}^N x_i^t}{N}, i = 1, K, d \quad (\text{Ek 1.17})$$

Kovaryans matrisi Σ 'nin kestirilen değerini S ile gösterirsek (Ek 1.18) ve (Ek 1.19)'da verilen ifadeler ile tanımlanır.

$$s_i^2 = \frac{\sum_{t=1}^N (x_i^t - m_i)^2}{N} \quad (\text{Ek 1.18})$$

$$s_{ij} = \frac{\sum_{t=1}^N (x_i^t - m_i)(x_j^t - m_j)}{N} \quad (\text{Ek 1.19})$$

X örneğinin korelasyon katsayıları (Ek 1.20) ile verilen elemanlardan oluşan R korelasyon matrisi ile tanımlanır (Alpaydın, 2004).

$$r_{ij} = \frac{s_{ij}}{s_i s_j} \quad (\text{Ek 1.20})$$

ÖZGEÇMİŞ

Doğum tarihi 17.05.1979

Doğum yeri İstanbul

Lise 1993-1997 Haydarpaşa Lisesi

Lisans 1997-2002 İstanbul Teknik Üniversitesi Elk-Elektronik Fak.
Elektronik ve Haberleşme Mühendisliği Bölümü

Yüksek Lisans 2002-2005 Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği

Çalıştığı kurum

2002-2005 Yıldız Teknik Üniversitesi, Elk.-Elektronik Fak.
Bilgisayar Mühendisliği Bölümü, Bilgisayar
Bilimleri Anabilim Dalı, Araştırma Görevlisi