

**T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**GERÇEK ZAMANLI GÜVENLİK KAMERASI VİDEOLARININ İNSAN
TABANLI ÖZETLENMESİ**

M. SAİD AYDEMİR

**YÜKSEK LİSANS TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI
BİLGİSAYAR MÜHENDİSLİĞİ PROGRAMI**

**DANIŞMAN
YRD. DOÇ. DR. M. ELİF KARSLIGİL**

İSTANBUL, 2012

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**GERÇEK ZAMANLI GÜVENLİK KAMERASI VİDEOLARININ İNSAN
TABANLI ÖZETLENMESİ**

M.Said AYDEMİR tarafından hazırlanan tez çalışması 10.07.2012 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda **YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Tez Danışmanı

Yrd. Doç. Dr. M.Elif KARSLIGİL
Yıldız Teknik Üniversitesi

Jüri Üyeleri

Yrd. Doç. Dr. M.Elif KARSLIGİL
Yıldız Teknik Üniversitesi

Prof. Dr. Nizamettin AYDIN
Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. Hakan ERDOĞAN
Sabancı Üniversitesi

ÖNSÖZ

Günümüzde birçok özelliđi bir arada bulunduran, taşınabilir ve kullanım kolaylığı ile dikkat çeken cihazların artışı ile birlikte inşaların teknolojiye olan ilgisi de artmıştır. Bu teknolojik gelişmeler sayesinde özellikle video verisinin saklanması ve paylaşılmasında ciddi bir artış söz konusu olmuştur.

Bu çalışma, güvenlik kamerası videolarının insan tabanlı özetlenmesi üzerinedir. Video verisinin özetlenmesi, görüntü akışı esnasında insan hareket özelliklerinin belirlenmesi ve hareket değişimlerinin belirlenmesiyle, çıkış video verisinin zaman boyutunda küçültülmesidir.

Tez çalışmam boyunca beni destekleyen, yol gösteren ve tez öğrencisi olmaktan onur duyduğum değerli tez danışmanım Yrd. Doç. Dr. Mine Elif Karslıgil'e teşekkürü bir borç bilirim. Bugüne kadar yetişmemizde emeđi olan YTÜ Bilgisayar Mühendisliğindeki bütün hocalarıma ve desteklerini her zaman hissettiğim çalışma arkadaşlarıma da teşekkürlerimi sunarım.

Bu tezi hayatım boyunca beni her zaman destekleyen ve yanımda olan aileme ve üzerimde emeđi olan herkese ithaf ediyorum.

Temmuz, 2012

M. Said AYDEMİR

İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ	vi
KISALTMA LİSTESİ	vii
ŞEKİL LİSTESİ.....	viii
ÇİZELGE LİSTESİ	x
ÖZET	xi
ABSTRACT	xiii
BÖLÜM 1	
GİRİŞ	1
1.1 Literatür Özeti	3
1.2 Tezin Amacı	6
1.3 Hipotez	7
BÖLÜM 2	
SİSTEM TASARIMI.....	8
2.1 Ön İşleme	9
2.1.1 Arka Plan Çıkarımı.....	9
2.1.2 Blob Analizi	12
2.2 Özellik Çıkarımı.....	13
2.2.1 Optik Akış.....	13
2.2.1.1 Horn ve Schunck Optik Akış Modeli	15
2.2.2 Eğimlerin Histogramı(Histogram of Gradients)	19
2.3 Özelliklerin Değerlendirilmesi.....	21
2.3.1 İnsan Tespiti.....	22
2.3.1.1 Destek Vektör Makineleri	23
2.3.2 İnsan Takibi	24
2.3.2.1 Kalman Filtresi.....	25
2.3.3 Kalıp Eşleme Yöntemleri	28

2.3.4	Hareket Kalıbının Bulunması ve Kalıp Tekrarının Aranması	30
BÖLÜM 3		
UYGULAMA		32
3.1	Weizmann İnsan Hareketi Veriseti	32
3.2	Tek Kişi Yandan Dik Videolarla Oluşturulmuş Veriseti.....	34
3.3	Tek Kişi Tepeden Çekilen Videolarla Oluşturulmuş Veriseti	40
3.4	Birden Fazla Kişinin Olduğu Videolarla Oluşturulmuş Veriseti	45
BÖLÜM 4		
SONUÇ VE ÖNERİLER		49
KAYNAKLAR.....		52
ÖZGEÇMİŞ.....		55

SİMGE LİSTESİ

$P(j)$	Dağılım ağırlık fonksiyonu
$P(x j)$	Aitlik ihtimali
μ	Sınıf ortalaması
σ	Sınıf varyansı
X_t	Sınıfa ait Gauss Dağılım Modeli
$\omega_{i,t}$	Gauss Dağılım ağırlığı
$\Sigma_{i,t}$	Gauss Dağılımı kovaryans matrisi
Π	Gauss olasılıksal yoğunluk fonksiyonu
M_{10}	Birinci dereceden moment
M_{20}	İkinci dereceden moment
$I(x, y)$	Gri seviye resim
$p(x, y)$	Pikselin parlaklığı
(x, y, t)	Pikselin zamana bağlı parlaklığı
I_x	x noktasındaki şiddetin türevi
∇I	Uzaysal şiddet gradyenini
$\vec{v}$	Pikselin t anındaki optik akışı
$\vec{v}_n$	Normal genliğe sahip hız vektörü
$\hat{n}$	Normal hız doğrultusu
θ	Piksel yönelim açısı
G	Pikselin Kenarı
S_y	Yatay Sobel Filtresi
S_d	Dikey Sobel Filtresi
ρ	Korelasyon katsayısı
$g(c,r)$	Pikselin gri değeri
(c,r)	Aranan pikselin koordinatları

KISALTMA LİSTESİ

HOG	Yönlü Gradyan Histogramı(Histogram of Gradients)
DVM	Destek Vektör Makinesi
GMM	Gaussian Mixture Model(Gauss Karışım Model)
SMM	Saklı Markov Modeli
EM	Expectation Maximization
OF	Optik Akış(Optic Flow)
INRIA	Institut National De Recherche En Informatique Et En Automatique
MIT	Massachusetts Institute of Technology
TAN	Tanjant

ŞEKİL LİSTESİ

	Sayfa
Şekil 1. 1 Video akımı [1]	1
Şekil 1. 2 Video akışının yapısı	2
Şekil 1. 3 Video akışı ve çıkarılmış özeti	3
Şekil 1. 4 Video özetleme sistemlerinin genel yapısı	3
Şekil 1. 5 Hareket değişimi tespit yaklaşımı [16]	6
Şekil 1. 6 Test videosu sol üst köşede görülmektedir. Sahne değişim sonuçları sağ üst, insan tarama sonuçları da sol alt köşede gözlenmektedir [17].....	6
Şekil 2. 1 İnsan Tabanlı Video Özetleme Sisteminin Genel Yapısı	9
Şekil 2. 2 Örnek bir çerçeve ve çerçeveye ait ön plan maskesi	11
Şekil 2. 3 Blob analizi sonucu elde edilen dikdörtgenin çerçeve üzerinde gösterimi .	13
Şekil 2. 4 Direğin hareket alanı ve optik akışı [24]	14
Şekil 2. 5 (x,y,t) konumundan, (x+dx, y+dy, t+dt) konumuna hareket eden bir noktanın renk bilgisi değişmezliği.....	15
Şekil 2. 6 (a) Belirsizlik Problemi, (b) Hareket kısıtlama denklemi normal gelen hız ..	17
Şekil 2. 7 İki boyutlu hareket kısıtlama denkleminin grafiksel gösterimi.....	18
Şekil 2. 8 İki çerçeve arasındaki optik akış.....	19
Şekil 2. 9 HOG algoritmasının genel adımları	19
Şekil 2. 10 Dokuz lokal bölgeye ayrılmış insan görüntüsü	21
Şekil 2. 11 Weizmann Veriseti'ne ait hog özellikleri	21
Şekil 2. 12 Pozitif ve negatif örnekler.....	22
Şekil 2. 13 Yeni oluşturulan pozitif örnekler	23
Şekil 2. 14 (a) İki sınıf için hiper düzlemler, (b) Optimum hiper düzlem ve destek vektörleri	23
Şekil 2. 15 Kernel fonksiyonu ile verinin daha yüksek bir boyutta ifade edilmesi.....	24
Şekil 2. 16 Ayrık Kalman Filtresi döngüsü	25
Şekil 2. 17 Alan tabanlı eşleme örneği	29
Şekil 2. 18 Gri seviye bir görüntü ve Sobel Operatörü ile bulunan kenarları.....	30
Şekil 3. 1 Weizmann Veriseti'nden yararlanılarak oluşturulmuş videoda arka arkaya gelen iki görüntü.....	33
Şekil 3. 2 Bulunamayan harekete ait birbiri arasında 6 görüntü fark olan iki görüntü	34
Şekil 3. 3 Yürüme ve yürümeden koşmaya geçiş başlangıç görüntüleri	34
Şekil 3. 4 Koşma ve koşmadan yürümeye geçiş başlangıcı	35
Şekil 3. 5 Yürüme-koşma ve koşma-yürüme geçişleri.....	35
Şekil 3. 6 Elinde nesne ile yürüyen kişinin yürüme şeklini değiştirme başlangıçları ...	36

Şekil 3. 7	Şüpheli kutu bırakan kişinin görüntüsü	37
Şekil 3. 8	Şüpheli kutu bırakma anı	37
Şekil 3. 9	Şüpheli kutunun alınmasında hareket değişimleri	38
Şekil 3. 10	Bir videoya ait korelasyon değerlerine ait sonuçlar	39
Şekil 3. 11	Veri setinin oluşturulduğu kamera görüş açısı	40
Şekil 3. 12	Yürüyüş görüntüsü açısı değişiminin tespit edilmesi.....	41
Şekil 3. 13	Bir kişiye ait yürüme kalıpları başlangıç görüntüleri.....	41
Şekil 3. 14	Bir kişiye ait koşmaya geçiş başlangıcı ve koşma kalıpları başlangıç görüntüleri.....	42
Şekil 3. 15	Özellik çıkarımı düzgün yapılamayan görüntü örnekleri	43
Şekil 3. 16	Hareket başlangıçları tespit edilemeyen görüntüler	44
Şekil 3. 17	Farklı hareket olarak algılanan hareket başlangıç görüntüleri	45
Şekil 3. 18	Çoklu insan veriseti hareket değişim tespitleri.....	46
Şekil 3. 19	Çoklu insan veriseti dış ortam görüntüsü	46
Şekil 3. 20	Çoklu insan veriseti dış ortam hareket değişimi örneği	47
Şekil 3. 21	Çoklu insan veriseti çakışma örneği.....	48

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 2. 1 Sistemin sözde kodu	31
Çizelge 3. 1 Weizmann Veriseti ile yapılan test sonuçları	33
Çizelge 3. 2 Sistemin dik kamera görüntüsündeki başarısı	38
Çizelge 3. 3 Sistemin verisetindeki başarısı	43
Çizelge 3. 4 Sistemin verisetindeki başarısı	47

GERÇEK ZAMANLI GÜVENLİK KAMERASI VİDEOLARININ İNSAN TABANLI ÖZETLENMESİ

M. Said AYDEMİR

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

Tez Danışmanı: Yrd. Doç. Dr. M. Elif Karslıgil

Bu tez çalışmasında güvenlik kameralarından alınan video görüntülerinin insan tabanlı özetlemesini yapan bir sistem tasarlanmış ve gerçekleştirilmiştir. Önemli bir olay gerçekleştiğinde, bazen kayıtların saatler süren bölümlerinin incelenmesi gerekmektedir. Aynı zamanda yapılan kayıtların tamamının saklanması mümkün olmadığından, belirli aralıklarla eski kayıtlar silinmekte, bu durumda da gerektiğinde istenilen kayıtlara ulaşmak her zaman mümkün olmamaktadır. Bu nedenlerden dolayı bu görüntülerin özetlenerek kısaltılması ve sadece anlamlı parçalarının kullanılması bir gereklilik haline gelmiştir.

Çalışma gerçekleştirilirken ön işlem adımları olarak Gauss Karışım Modeli ile ön plan çıkarımı yapılmıştır. Daha sonra çıkarılan ön plan Blob Analizi ile çerçeve içine alınmıştır. Çerçeve içine alınan nesnenin hareket bilgisini elde etmek için optik akış yöntemi kullanılmıştır. Hareket bilgisi çıkarılan nesneye yönlü gradyanlar yöntemi uygulanarak özellik çıkarımı yapılmıştır.

Her nesne için çıkarılan yönlü gradyan özellikleri kullanılarak, çevrimdışı olarak oluşturulan eğitim seti yardımıyla Destek Vektör Makinesi yöntemi kullanılarak nesnenin insan olup olmadığına karar verilmiştir. Eğer nesne insan ise Kalman Filtresi ile insan takibi gerçekleştirilir. Yönlü gradyan bilgilerini taşıyan histogramlara zaman bilgisi eklenerek, video boyunca nesneye ait hareket bilgisi sürekli hale getirilmiştir. Sürekli hale getirilen gradyan bilgisi üzerine Şablon Eşleme yöntemi uygulanarak hareket kalıpları ve bu hareket kalıplarının tekrar başlangıçları bulunmuştur. Bulunan

hareket kalıplarının geiş noktaları da hareket geişleri olarak işaretilenmiştir. Hareket geişleri özet video için videonun önemli anlarıdır.

Çalışma 4 ayrı veri setinde test edilmiştir. Sistemin özetleme başarısı, periyodik hareketleri tespit etme başarısıdır. İlk olarak test edilen veri seti genel kullanıma açık olan Weizmann İnsan Hareketi Veri Seti'dir. Bu veri setinden 6 kişiye ait 9 farklı hareket için özetleme işlemi gerçekleştirilmiştir. Bu örneklerde %89,47 özetleme başarısı tespit edilmiştir. İkinci kullanılan veri setinde güvenlik kamerası kayıtlarında olabilecek çeşitli durumlar için senaryolar 3 farklı kişi tarafından gerçekleştirilmiştir. Veri seti üzerinde gerçekleştirilen test çalışmaları sonucunda %93,44 başarı oranı elde edilmiştir. İkinci veri setiyle benzer senaryoların farklı kamera açısıyla gerçekleştirildiği üçüncü bir veri seti üzerinde yapılan denemeler sonucunda %89,29 başarı oranı yakalanmıştır. Birden fazla kişinin bulunduğu videolarda ise %75,76 gibi bir başarı gözlemlenmiştir.

Anahtar Kelimeler: İnsan tabanlı video özetleme, güvenlik kamerası, gradyanların histogramı, Gauss karışım modeli, optik akış

HUMAN BASED REAL TIME SURVEILLANCE VIDEO SUMMARIZATION

M. Said AYDEMİR

Department of Computer Engineering

MSc. Thesis

Advisor: Asst. Prof. Dr. M.Elif KARSLIGİL

In this thesis, a human based video summarization system, which receives video frames from surveillance cameras, was designed and implemented. When an important event happens, sometimes surveillance video records should be examined for hours. At the same time, old entries are deleted at regular intervals because storing of all records is not feasible. In this case it is not always possible to reach the needed records. For these reasons, shortening by summarizing and using only the significant parts of these records have become a necessity.

In this work, as a pre-processing step Gaussian Mixture Model is used for foreground extraction. Then the foreground was taken into the frame by using Blob Analysis. Optical flow method is used to obtain information on the object motion within frames. Histogram of Gradients method was applied to motion information of the object for feature extraction.

Histogram of gradient features extracted for each object by using Support Vector Machine method and offline generated training set are used to decide the object whether human or not. If the object is human, human tracking is performed with the Kalman Filter. During the video motion information of object has been made continuous by adding time information to histograms with gradient information. Template matching method is applied on the continuous gradient information to find beginnings of movement patterns and beginnings of repetition of the movement patterns. Transition points of the movement patterns are also marked as movement

transitions. Movement transitions are the important video moments for the summary video.

This study was tested on 4 different datasets. Summarization success rate of the system is the success of detecting periodic movements. First, Weizmann Human Action Database, which is available for public use, is used to test. In this dataset, summary is achieved for 6 people with 9 different actions. %89,47 summarization success rate was found in these examples. In the second dataset various situation scenarios that may occur in surveillance video records were performed by 3 different people. As a result of testing on this dataset %93,44 success rate was obtained. %89,29 success rate was achieved as a result of test with the third dataset, which contains scenarios similar to the second dataset with a different camera angle. %75,76 success rate is observed with multiple human videos.

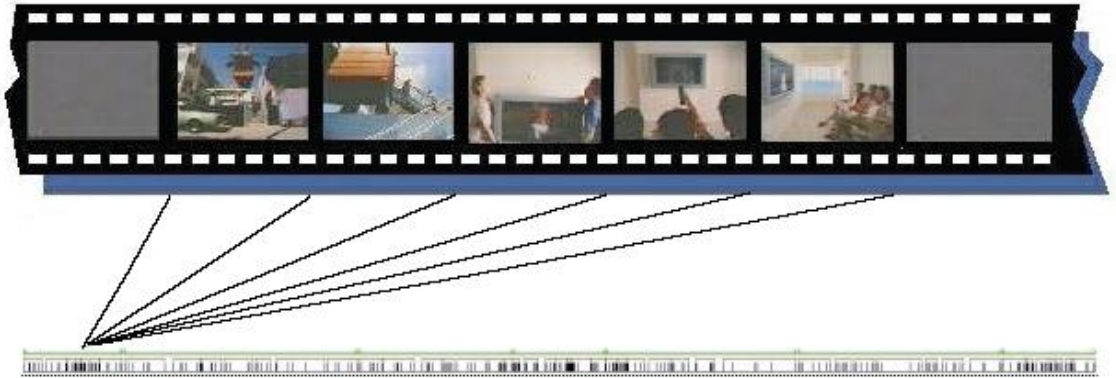
Key words: Human based video summarization, surveillance camera, histogram of gradients, Gaussian mixture Model, optical flow

BÖLÜM 1

GİRİŞ

Günümüzde güvenlik kameralarının kullanımı çok yaygınlaşmıştır. Önemli bir olay gerçekleştiğinde, bazen kayıtların saatler süren bölümlerinin incelenmesi gerekmektedir. Fakat yapılan kayıtların tamamının saklanması mümkün olmadığından, belirli aralıklarla eski kayıtlar silinmekte, bu durumda gerektiğinde istenilen kayıtlara ulaşmak her zaman mümkün olmamaktadır. Bu nedenlerden dolayı bu görüntülerin özetlenerek kısaltılması ve sadece anlamlı parçalarının kullanılması bir gereklilik haline gelmiştir.

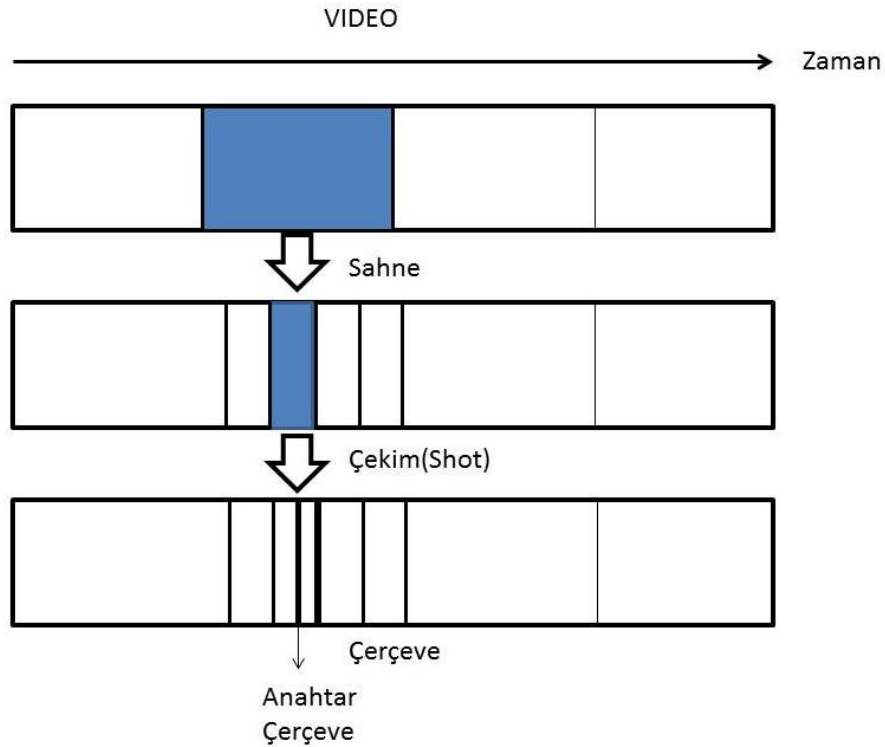
Video, arka arkaya resimlerden meydana gelen bir veri bütünüdür. Şekil 1.1’de görülebileceği gibi zaman içinde değişen resimler videoyu oluşturmaktadır. Bu şekildeki görüntüler, belirli zaman aralıklarıyla alınan görüntülerdir.



Şekil 1. 1 Video akımı [1]

Bir videoyu oluşturan herbir resime çerçeve denilir. Tek bir kameradan kaydedilmiş, sıralı çerçeveler bütünü oluşturulan yapılara çekim(shot) denilmektedir. Bir çekimin

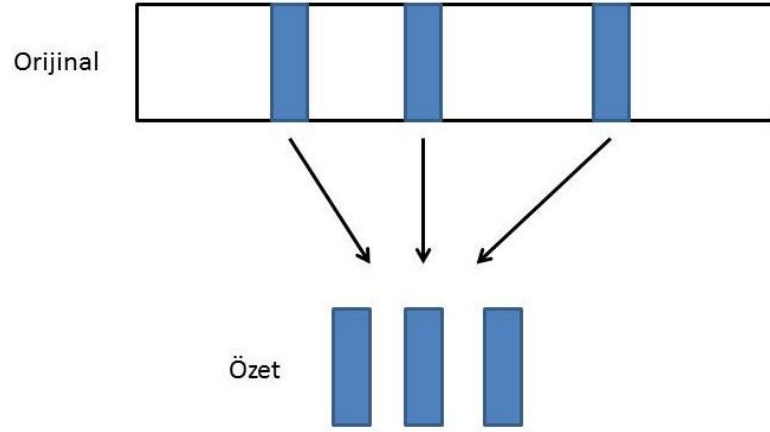
özelliklerini en iyi gösteren çerçeveye anahtar çerçeve(key frame) denilmektedir. Çekimlerin bir araya gelmesiyle oluşan, anlamsal bütünlüğe sahip video parçalarına sahne denilir. Çekimler yapısal sınırlar, sahneler de mantıksal sınırlarla belirlenmektedir [2]. Şekil 2.1' de bir video akışının yapısı görülmektedir.



Şekil 1. 2 Video akışının yapısı

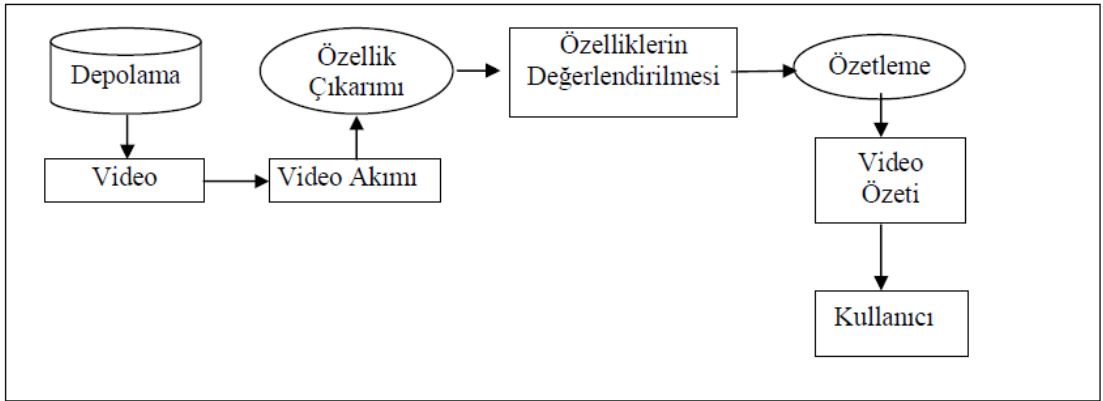
Özetleme; bir yazı, konu söz veya filmin içeriğini daha az sözle anlatmak, özünü vermek, kısaltmaktır.(Türk Dil Kurumu). Video özetlemede amaç, orijinal video hakkında mümkün olduğunca çok bilgi içeren sahnelerden oluşan bir resim kümesi elde etmektir [3]. Video gözden geçirmede(skimming) ise amaç, ses ve görüntü bilgisini beraber kullanarak içerik bilgisi elde etmektir. İçerik bilgisiyle videoyu en iyi temsil eden sahneler elde edilir [4].

Şekil 1.3'te, orijinal video akımında siyah ile gösterilen kısımlar sahne değişiminin olduğu resimleri, beyaz ile gösterilen ise aynı sahnenin devam ettiği yani değişimin olmadığı resimleri göstermektedir. Şekil 1.3'te de görüldüğü gibi değişen sahneler özeti oluşturmaktadır. Bu sahneler ise, ardışık görüntülerdeki hareket, renk, doku ya da şekil gibi özelliklerin değişiminin incelenmesi ile belirlenmektedir [5].



Şekil 1. 3 Video akışı ve çıkarılmış özet

Şekil 1.4’te video özetleme uygulamalarının genel adımları görülmektedir. Video görüntülerinden özellik çıkarımı ve değerlendirilmesi en önemli adımları oluşturmaktadır.



Şekil 1. 4 Video özetleme sistemlerinin genel yapısı

Bu çalışmada da genel olarak güvenlik kameraları düşünülerek fakat farklı alanlarda da kullanılabilecek insan tabanlı bir video özetleme sistemi gerçekleştirilmiştir. Güvenlik kameralarının günümüzdeki sayısı ve kayıt edilen video büyüklükleri düşünüldüğünde bu çalışmanın faydalı olacağı düşünülmüştür.

1.1 Literatür Özeti

Video görüntülerinin özetinin çıkarılmasında, çoğu zaman anahtar çerçeve(key frames) çıkarımı ile bir özetleme gerçekleştirilir [6]. Son yıllarda gerçekleştirirken çalışmalarda genel olarak hareket odaklı anahtar çerçeve çıkarımı [7], bilgi teorisi tabanlı çekim çıkarımı [8], renk dağılım karakteristiği histogramı tabanlı özetleme [9], hareket eden

obje takibine ve yörünge çıkarımına dayalı özetleme [10], Saklı Markov Modeli'nin tamamlayıcı özelliklerle kullanılmasıyla çekim çıkarımı [11] gibi bazı çekim ve hareket odaklı çalışmalar yapılmıştır. Yapılan çalışmalarda genel olarak sahne seçimi ve çerçevenin önem derecesini belirleme temel karar noktalarıdır [12] [13] [14]. Geçici şablonlar [15] ve silüet tünelleri [16] yöntemleri çekim tabanlı özetleme yapısından kurtularak insan tabanlı özetlemeyi amaçlayan çalışmalardır. Silüet tünelleri yöntemi gerçek manada insan odaklı özetlemeyi hedeflese de kullandığı videolardan dolayı tam manasıyla bir özetleme işlemi gerçekleştirmemektedir. Bu çalışmada kullanılan videolar ayrı ayrı kaydedilmiş hareket videolarının birleştirilmesiyle oluşturulmuştur.

Hareket odaklı anahtar çerçeve çıkarımı daha çok arabaları hedef alarak trafik güvenlik kameraları için tasarlanmış bir yöntemdir. Bu yöntemde öncelikle ortalama sabit bir hareket hareket akışı hesaplanmaktadır. Daha sonra bu hareket akışına bağlı olarak her hareketli nesnenin hareket bilgisi elde edilir. Nesnelerin hareket bilgisi ortalama hareket bilgisine göre değerlendirilerek anahtar çerçeve çıkarımı gerçekleştirilir.

Bilgi teorisi tabanlı çekim çıkarımı, çerçeveler arasındaki karşılıklı bilgi(mutual information) ve ortak entropi(joint entropy) hesabına dayanır. Bu yöntem hem ani geçişleri hem de yavaş çekim geçişlerinde başarılı sonuçlar vermiştir. Bilgi teorisi ölçütü çerçeveler arasındaki bilgiyi, çerçeveleri birbirinden çıkararak elde edilen bilgiden daha iyi temsil ettiği için daha iyi sonuçlar vermiştir. Dolayısıyla anahtar çerçeve çıkarımı için de karşılıklı bilgi yöntemini kullanılmaktadır.

Renk dağılım karakteristiği histogramı tabanlı özetleme, bir resimdeki aynı renkteki piksellerin birbirlerine komşu olup olmadıklarını değerlendirerek resmin uzaysal karakteristiğini çıkarır. Bu özellik sayesinde, piksellerin komşuluklarına göre histogram oluşturulur. Bu histogramlar video içeriğindeki değişimlere, renk dağılımlarına göre oluşturulan normal histogramlardan daha duyarlıdır. Bu da bu yöntemin normal histogram dağılımı yönteminden daha verimli ve başarılı özetleme yapmasına olanak sağlamaktadır.

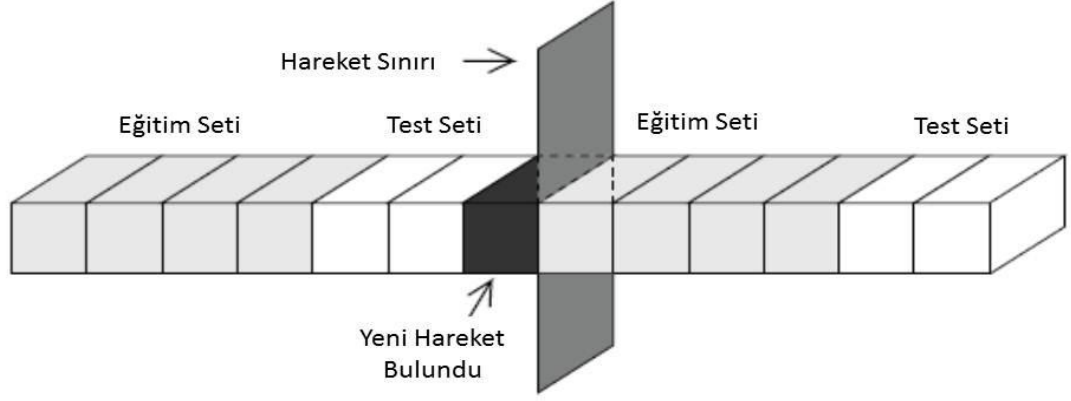
Hareket eden obje takibine ve yörünge çıkarımına dayalı özetleme yönteminde, öncelikle hareketli nesneler tespiti temel alınarak video parçalara ayrılır. Daha sonra

her hareketli nesne için yörünge çıkarımı gerçekleştirilir. Anahtar çerçeveler, hareketli nesne yörüngesinin o video parçasını temsil edip etmemesine göre seçilir.

Makine öğrenmesi tabanlı Saklı Markov Modeli'nin tamamlayıcı özelliklerle kullanılması yönteminde, çekim çıkarımı yapılırken hem şekil hem de renk bilgisinden yararlanılmıştır. İki özelliğin beraber kullanılması yöntemin avantajını iki katına çıkarmıştır. Yöntem uygulanırken öncelikle farklı çekim geçişleri için geçici karakteristik çıkarımı yapılmıştır. Bu sayede eğitilmiş Saklı Markov Modeli kullanılarak çekim geçişlerinin otomatik olarak bulunması amaçlanmıştır. Eğitilmiş SMM kullanmak eşik değeri bulma gibi bir problemi de ortadan kaldırmıştır. Daha sonra birbirini tamamlayıcı iki özellik olan istatistiksel köşe değişim oranı ve Hue Koyuluk Değeri renk histogramı farkı kullanılmıştır. Köşe özelliği kullanılarak şekil, renk özelliği kullanılarak görünüş özetlemesi gerçekleştirilir.

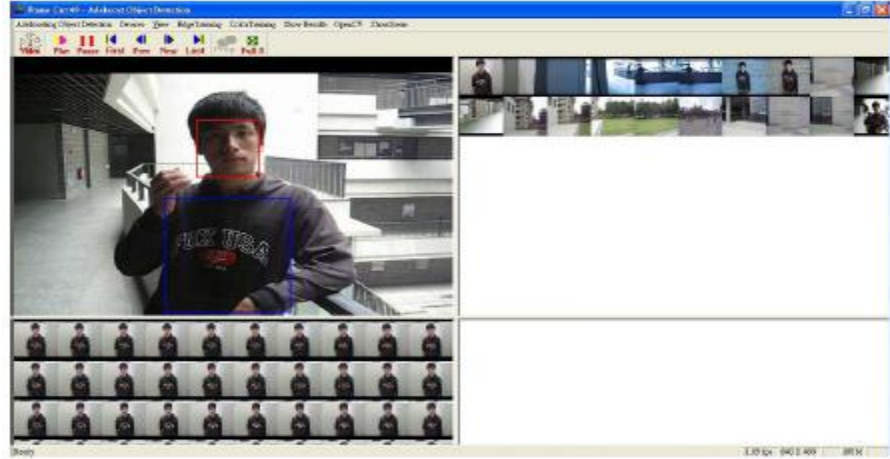
Geçici şablonlar yönteminde resim akışındaki her bir nokta hareket bilgisi içeren statik bir vektör olarak ifade edilir. Bu vektörün iki temel elemanı bulunur. İlki, iki terimli bir değer olan o noktada hareket olup olmadığı bilgisidir. İkinci ise hareket var ise bu hareketin ne kadar yakın bir zamanda olduğu bilgisidir. Yöntemde belirlenen bazı hareketlerin statik vektörleri çıkarılarak eğitim seti oluşturulur. Test amaçlı kullanılan videonun da statik vektörleri çıkarılarak eğitim setindeki statik vektör setlerinden hangisiyle eşleştiği kontrol edilir.

Silüet tünelleri yönteminde ilk olarak yapılan her çerçevede bulunan insanın silüetini çıkarmaktır. Silüet çıkarımı, ön plan çıkarımı yapılarak gerçekleştirilir. Her çerçeve için çıkarılan silüetlerden, yine her çerçeve için 10 adet özellik vektörü oluşturulmaktadır. Bunlardan 2' si zaman boyutunda, geriye kalanlar da silüetin şekil değerlerinin tutulmasında kullanılmaktadır. Şekil 1. 5' te görüleceği gibi her hareket değişiminden sonra belli sayıdaki çerçeve eğitim için kullanılmaktadır. Yani eğitim için kullanılan bu bölgede hareketin değişmediği, yeni bir harekete geçiş olmadığı, kabul edilmektedir. Bu çalışmada hareket değişimleri tek hareketi kapsayan farklı videoların birleştirilmesiyle oluşturuldukları için hareket değişimi esnasında olabilecek tekrar etmeyen anlamsız hareketler göz ardı edilmiştir.



Şekil 1. 5 Hareket değişimi tespit yaklaşımı [16]

Bir başka insan tabanlı özetleme yönteminde insanın öncelikle yüzü tespit edilerek daha sonraki çerçevelerde bulunan yüzün aynı kişi olup olmadığına bakılmaktadır [17]. İlk önce yüz bulma yapıldıktan sonra vücudun geri kalan bölümlerinden de renk bilgisi kullanılarak özellik çıkarımı yapılmaktadır. Bu çalışmada tam anlamıyla insan hareketine yönelik bir özetleme çalışması değildir. Ayrıca kameraya bakan kişinin yüzünün kamerayı karşıdan gördüğü varsayılmaktadır. İnsan bulunmayan görüntülerde de sahne değişiminden yararlanarak özetleme yapılmıştır. Şekil 1. 6'da çalışmaya ait örnek çıktı görülmektedir.



Şekil 1. 6 Test videosu sol üst köşede görülmektedir. Sahne değişim sonuçları sağ üst, insan tarama sonuçları da sol alt köşede gözlenmektedir [17].

1.2 Tezin Amacı

Son yıllarda artan bir hızla gelişen multimedya teknolojilerden dolayı kaydedilen videoları düzgün bir şekilde değerlendirip arşivlemek hem zaman hem de kayıt yeri

açısından pek mümkün değildir. Bu nedenden dolayı uzun video kayıtlarını incelemek ve arşivlemek yerine anlamlı video parçalarıyla işlem yapmak gereği ortaya çıkmıştır. Bunun sonucunda da birçok video özetleme sistemi geliştirilmiştir.

Genel olarak kaydedilen video sayısındaki artış güvenlik kamerası kayıtlarına da yansımıştır. Bu yansımanın sonucu olarak da artık kayıtları incelemek ve arşivlemek işgücü ve zaman açısından kabul edilemez noktalara gelmiştir. Bu tez çalışmasında, güvenlik kamerası veya insan odaklı çalışan kamera görüntülerinden anlamlı parçaları çıkararak videoyu özetleyen bir sistem tasarlanmıştır ve gerçekleştirilmiştir. Temel olarak güvenlik kameralarına odaklanılsa da, insan tabanlı bir sistem olduğu için insanın olduğu her çeşit videodaki kullanılabilecek bir çalışma yapılması amaçlanmıştır.

1.3 Hipotez

Video özetleme için, her video dizisinde tekrar eden hareketler ve görüntülerin sadece tek bir defa görüntülenmesi yeterli bir bilgidir. Birden daha fazla bulunan bir görüntünün oluşturduğu gereksiz bilginin elenmesi gerekir. Buradan yola çıkarak her bir görüntüde tekrar eden hareketler bütünü elenebilirse, saf anlamlı bilgi elde edebilir.

Bir video çerçeve akımı olduğundan dolayı öncelikle her bir çerçevenin özellik çıkarımı, o çerçeveyi ifade eden anlamlı bilgi çıkarımı yapılabilirse, tüm video için düzgün bir bilgi çıkarımı yapılmış olunur. Her çerçeve için oluşturulan bilgilere zaman bilgisini de eklenirse, bir video akımında tekrar eden elenmesi gereken bilgi akımını bulunabilir. Eğer incelenen videoda bir hareket yani ardışık çerçeve bütünleri bir kere bulunuyorsa, anlamlı bilgi olduğu için elenme işlemine dâhil olmaz.

Bir videoda tekrar eden hareketlerin bulunabilmesi için, öncelikle o hareket bütünün, kalıbının bulunması gerekir. Daha sonra bu kalıbın videonun geri kalanında aranmasıyla tekrar eden hareketlerin elenme işlemi yapılabilir. Bu noktada karşı karşıya kalınabilecek iki ana problem ortaya çıkar. Birincisi kalıbın başlangıç noktası belli değildir. İkincisi de kalıbın büyüklüğü belli değildir. Bu problemler, öncelikle başlangıcı sabit kabul edip, kalıp büyüklüğünü aramakla çözülebilir. Eğer belli bir kalıp büyüklüğüne ulaşıldığı halde hala bir kalıp bulunamadıysa, kalıp başlangıcı kaydırılarak amaçlanan özetleme işlemi gerçekleştirilebilir.

BÖLÜM 2

SİSTEM TASARIMI

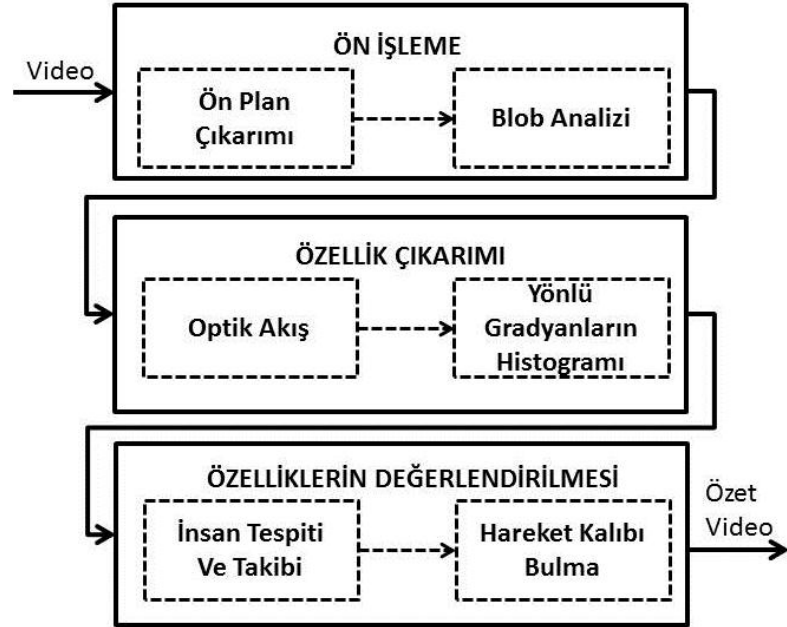
Bu tez çalışmasında insan görüntüsü içeren videolardan, periyodik hareketlerin belirlenmesiyle video özetleme yapan bir sistem tasarlanmış ve gerçekleştirilmiştir. Geliştirilen sistem ön işleme, özellik çıkarımı ve özelliklerin değerlendirilmesi bölümlerinden oluşmaktadır. Şekil 2. 1' de sistemin yapısı görülmektedir.

Şekilde görüldüğü gibi ön işleme adımı iki alt adımdan oluşur. Öncelikle arka plan çıkarılır. Arka plan elde edildikten sonra, blob analizi yapılarak ön plandaki nesne çerçeve içine alınır. Bu işlem, işlenecek kısmın küçültülerek işlem yükünden kurtulmayı ve özellik çıkarımında gereksiz bilgi elde edilmemesini sağlar.

Ön işleme adımlarından sonra gerçekleştirilen özellik çıkarımı bölümü de iki adımdan oluşmaktadır. Çerçeve içine alınan yani nesnenin bulunduğu bölgenin optik akışı hesaplanarak, hareket bilgisine sahip bölgeler ön plana çıkarılıp ağırlığı artırılır. Daha sonra Yönlü Gradyanların Histogramı yöntemi kullanılarak her nesnenin gradyan bilgileri çıkarılarak, her nesne gradyanların histogramı ile ifade edilir. Elde edilen histogramlar bütününden bir nesne elde edilir.

Özellik çıkarımı sonucundan elde edilen her nesnenin yönlü gradyanlar histogramı kullanılarak çerçeve içindeki nesnenin insan olup olmadığına karar verilir. İnsan belirleme işlemi için Destek Vektör Makineleri(Support Vector Machines-SVM) yöntemi kullanılır. Çerçeve içindeki nesne insan ise hareket ve konum bilgileri kullanılarak Kalman Filtresi ile nesne takibi gerçekleştirilir. Nesnelerin histogramları bir araya getirilerek oluşturulan nesne şablon eşleme yöntemiyle analiz değerlendirilerek

hareket kalıpları bulunur. Bulunan hareket kalıplarının değişimi değerlendirilerek özet video oluşturulur.



Şekil 2. 1 İnsan Tabanlı Video Özetleme Sisteminin Genel Yapısı

2.1 Ön İşleme

Ön işleme aşamasında amaç, görüntünün arka planının belirlenmesi sayesinde ön plandaki hareketli nesnelerin elde edilmesi ve insan olan nesnelerin belirlenmesidir. Ön işleme adımı 2 alt adımdan oluşur.

2.1.1 Arka Plan Çıkarımı

Hareketli nesnelerin belirlenebilmesi için arka plan ile ön planın ayrılması gereklidir. Uygulamada ön plan çıkarımı için kullanılan yöntem Gauss Karışım Modeli(Gaussian Mixture Model-GMM)'dir [18].GKM ışık değişiminden etkilenmemesi, arka planın sabit olmadığı ortamlarda, özellikle açık alanda istikrarlı sonuçlar vermesinden ötürü tercih edilmiştir. Çünkü bu çalışmada asıl hedeflenen güvenlik kameraları olduğundan dolayı iç ortamda ve açık alanda aynı istikrarda çalışması amaçlanmıştır.

Gauss Karışım Modeli istatistiğe dayalı bir eğiticişiz öğrenme algoritmasıdır. Bu model, aynı sınıfa ait verilerin normal dağılıma uygun bir şekilde sınıf merkezi etrafında

dağıldığını öngörür. Algoritma, önceden belirlenen sınıf sayısı için, örneklerin en az varyansla toplandığı sınıf merkezlerini ve bu merkezlere ait örneklerin standart sapmalarını bulur.

Modelin matematiksel ifadesi şu şekildedir:

$P(j)$ j. sınıfın dağılımdaki ağırlığı, başka bir deyişle bu sınıftaki elemanların olasılığını gösterirken K sınıf için Eşitlik 2. 1'deki bağıntı elde edilir.

$$\sum_{j=1}^K P(j) = 1 \quad (2.1)$$

X elemanının j. sınıfa ait olma olasılığına $P(x|j)$ denilirse, bu olasılığın Gauss dağılımı Eşitlik 2. 2'deki gibidir.

$$P(x|j) = \frac{1}{(2\pi)\sigma_j^2} e^{-\frac{(x-\mu)^2}{2\sigma_j^2}} \quad (2.2)$$

Bu eşitlikte, x sınıflandırılacak veri, j aidiyetine bakılan sınıf, μ söz konusu sınıfın ortalaması ve σ ise söz konusu sınıfın varyansını göstermektedir.

Bu durumda sınıflandırılmak istenilen bir x parçacığının j sınıfına hangi oranda ait olabileceği (2.3)'teki gibi bulunabilir.

$$P(j|x) = \frac{P(x|j)P(j)}{\sum_j P(x|j')P(j')} \quad (2.3)$$

O halde bütün x parçacılarının söz konusu sınıflara ait olma ihtimali bütün x'ler için (2.3)'teki olasılıkların toplamıdır. Bu olasılığı en yüksek yapacak sınıf ortalamalarının ve varyanslarının ne olacağı bulunduğu sınıflama işlemi yapılmış olunur. Bütün bu parametreleri bulma işlemi literatürde Beklenti Maksimizasyonu (Expectation Maximization-EM) [19] olarak bilinen yöntemle hesaplanır. Başlangıçta her sınıfın ortalama değeri, standart sapmaları ve ağırlıkları rastgele belirlenir. Her iterasyonda bu değerler çevrimiçi olarak güncellenir. Bu yüzden GMM işlem yükü ağır olan bir algoritmadır.

Bu çalışmada her parçacık bir piksele karşılık gelmektedir. Sınıf sayısı ön plan ve arka plan için olmak üzere ikidir. Yöntemde bir pikselin zaman serisine bağlı değerlerine piksel süreci adı verilmiştir. Herhangi bir t anında, I_m bir resim dizisiyse, belirli bir $\{x_0, y_0\}$ pikseline ait geçmiş (2.4) ile ifade edilir.

$$\{X_1, \dots, X_t\} = \{I_m\{x_0, y_0, i\}: 1 \leq i \leq t\} \quad (2.4)$$

Her piksel için $\{X_1, \dots, X_t\}$ geçmiş bilgilerinden yararlanılarak her piksel için K adet Gauss Dağılımı modellenir. Bir piksel için olasılık değeri (2.5)'de görüldüğü gibi hesaplanır.

$$P(X_t) = \sum_{i=1}^K \omega_{i,t} * \eta(X_t, \mu_{i,t}, \Sigma_{i,t}) \quad (2.5)$$

(2.5)'de K dağılım sayısını, $\omega_{i,t}$ t zamanındaki i 'inci Gauss Dağılımı'nın ağırlığını, $\mu_{i,t}$ t zamanındaki i 'inci Gauss Dağılımı'nın ortalamasını, $\Sigma_{i,t}$ t zamanındaki i 'inci Gauss Dağılımı'nın kovaryans matrisini, η ise Gauss olasılıksal yoğunluk fonksiyonunu ifade eder.

Bu dağılımlardan empirik olarak belirlenen ilk p adedi arka plan, p ile K arasında kalan adedi de ön plan olarak kabul edilir. Yeni gelen piksel dağılım hangi sınıfa daha yakınsa, o sınıfa ait olarak kabul edilir.

Şekil 2. 2'de veri setinde bulunan bir videodan alınan bir çerçeve ve o çerçeveye ait Gauss Karışım Modeli kullanılarak çıkarılan ön plan görülmektedir.



Şekil 2. 2 Örnek bir çerçeve ve çerçeveye ait ön plan maskesi

2.1.2 Blob Analizi

Blob analizi [20] yöntemi bir resim üzerinde birbirine bağlı bölgelerin bir bütün olarak konumunun ve büyüklüğünün tespit edilmesini sağlar. Bu çalışmada blob analizi yöntemi kullanılarak ön plan olarak çıkarılan nesnenin çerçeve içine alınması amaçlanmıştır. Analiz uygulandıktan sonra ön planın bir dikdörtgen çerçeve içine alınması istendiği için gerekli parametreler dikdörtgen çerçevenin merkezi, kısa ve uzun kenar uzunlukları olacaktır [21].

Blob analizi, aynı mantıksal seviyedeki bitişik piksel bölgelerinin bulunmasıdır. Bu analiz nesne içerisindeki blobların büyüklükleri ve merkez noktası bilgilerinin elde edilmesi için kullanılır. (X_c, Y_c) merkez noktası koordinatları olmak üzere Eşitlik 2. 6' daki gibi hesaplanır.

$$X_c = \frac{M_{10}}{M_{00}} \quad Y_c = \frac{M_{01}}{M_{00}} \quad (2.6)$$

M_{00} , M_{10} ve M_{01} Eşitlik 2. 7' de görüldüğü gibi ikili seviye yani siyah beyaz $I(x, y)$ resminden hesaplanan birinci dereceden momenti göstermektedir.

$$M_{00} = \sum_x \sum_y I(x, y) \quad M_{10} = \sum_x \sum_y x \cdot I(x, y) \quad M_{01} = \sum_x \sum_y y \cdot I(x, y) \quad (2.7)$$

Yine benzer şekilde ikinci dereceden moment de (2.8)'deki gibi hesaplanır.

$$M_{20} = \sum_x \sum_y x^2 \cdot I(x, y) \quad M_{02} = \sum_x \sum_y y^2 \cdot I(x, y) \quad (2.8)$$

İkinci dereceden momentin hesaplanmasıyla a uzun kenar ve b kısa kenar olmak üzere dikdörtgenin kenar uzunlukları (2.9)'daki gibi elde edilir.

$$a = \sqrt{6(p + r + \sqrt{q^2 + (p - r)^2})} \quad (2.9)$$
$$b = \sqrt{6(p + r - \sqrt{q^2 + (p - r)^2})}$$

2.9'da kullanılan p, q ve r değerleri (2.10)'daki gibi hesaplanır.

$$p = \frac{M_{20}}{M_{00}} - x_c^2 \quad q = 2\left(\frac{M_{20}}{M_{00}} - x_c y_c\right) \quad r = \frac{M_{02}}{M_{00}} - y_c^2 \quad (2.10)$$

Blob analizi ile elde edilen dikdörtgen çerçevenin gri seviye ve bölütleme sonrası ikili seviyedeki görüntüsü Şekil 2. 3’de görülmektedir.



Şekil 2. 3 Blob analizi sonucu elde edilen dikdörtgenin çerçeve üzerinde gösterimi

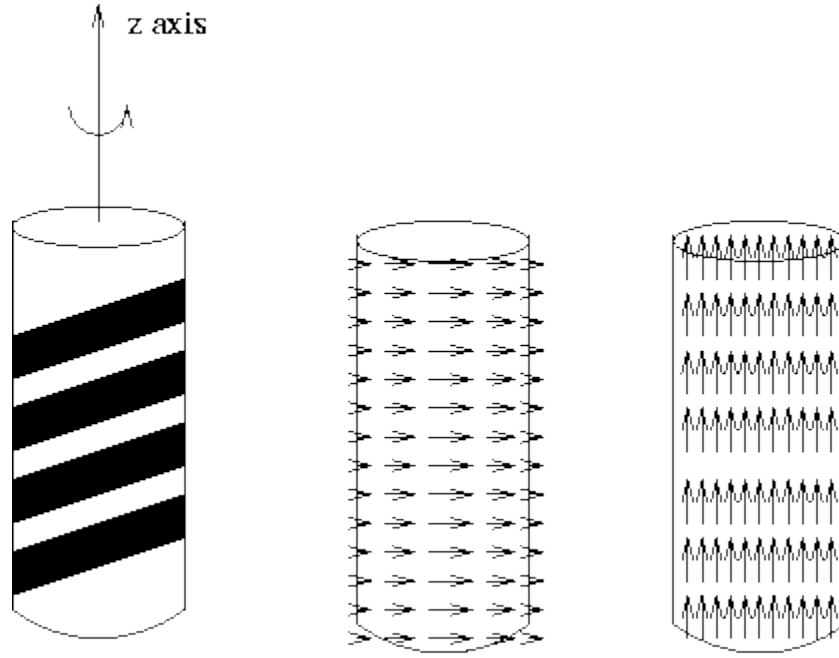
2.2 Özellik Çıkarımı

Özellik çıkarım aşamasında nesnenin yönlü gradyan özellikleri elde edilir. Çünkü yönlü gradyan özellikleri her nesne ve nesnelere ait hareketler için ayırt edici bir özelliktir. Bu nesnelerin bu özelliklerinin çıkarılmasıyla hem nesnenin insan olup olmadığına karar verilebilecek hem de nesneye ait hareketlerin ayırt edici özellikleri elde edilecektir.

Özellik çıkarım adımıyla öncelikle optik akış [22] kullanılır. Optik akış sayesinde hareketli parçaların ağırlıkları daha fazla olacağından dolayı sistemin hareketli kısımlar üzerine odaklanmasına olanak sağlanır. Hareket odaklı yapılan bir çalışmada hareket bilgisinin ön plana çıkarılması gereklidir. Optik akış aslında bir sonraki adımda kullanılacak olan Yönlü Gradyanların Histogramı(Histogram of Gradients-HOG) [23] yönteminin bir ön adımı olarak kullanılmaktadır.

2.2.1 Optik Akış

Optik akış, 3 boyutlu sahnede gözlemci hareket ettiğinde, görüntüde meydana gelen gözle görülür hareketin yön ve hızını tanımlar. Genellikle optik akış hareket alanına karşılık gelir fakat bu denklik her durum için geçerli değildir. Örneğin Şekil 2. 4’teki dönen direğin hareket alanı ve optik akışı birbirinden farklıdır.



Şekil 2. 4 Direğin hareket alanı ve optik akışı [24]

Optik akışta ardışık çerçevelerde bir pikselin renginin değişmediği ve bir pikselin yakın komşuları ile birlikte hareket ettiği kabul edilir.

Optik akış hesaplanması için birçok yöntem ileri sürülmüştür. Optik akış aşağıda bahsedilen üç gruptan birine bağlı olarak sınıflandırılabilir:

1. Fark teknikleri: Uzay-zamansal şiddet türevlerinden görüntü hızı hesaplanır.
2. Frekans temelli teknikler: hız ayarlı filtre çıkışındaki enerji bilgisi kullanılır.
3. Eşleştirme teknikleri: az sayıda görüntüden (genellikle iki veya üç görüntü dizisinden) değişik görüntü özelliklerini eşleştirerek görüntü hareketi hesaplanır.

Bu üç grup yaklaşım arasında geliştirilen uygulamaya göre performans farklılıkları görülse de birçok bakımdan birbirlerine denk olarak gösterilirler. Genel yapısı bakımından bu teknikler üç ana aşamada gerçekleştirilir:

1. Görüntü üzerinde ilgilenilen işaret yapısını elde edebilmek ve işaret/gürültü oranını artırmak için alçak geçiren veya bant geçiren filtre kullanılarak ön filtreleme veya yumuşatma.
2. Hızın normal bileşenlerini hesaplamak için temel ölçümlerin hesaplanması.

3. İki boyutlu akış alanını üretebilmek için ön filtreleme ve temel hesaplamaların entegrasyonunun yapılması.

Bu çalışmada Horn ve Schunck yöntemi kullanılmıştır [25].

2.2.1.1 Horn ve Schunck Optik Akış Modeli

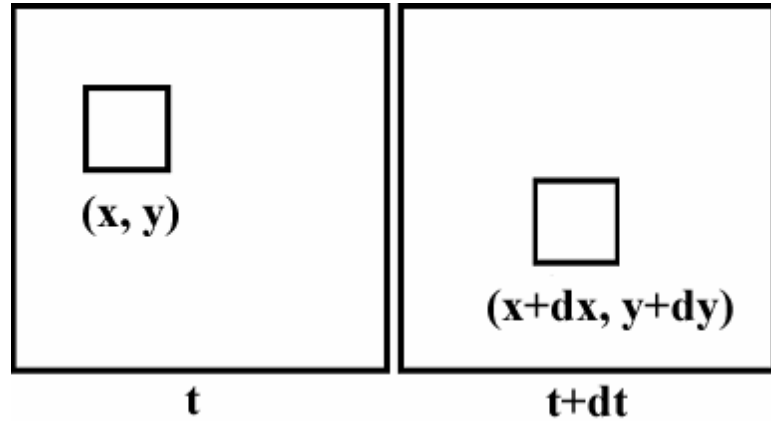
Nesne üzerinde bulunan (x, y) pikselinin parlaklığının $I = (x, y, t)$ olduğu ve piksellerin zamana bağlı parlaklıklarının sabit olduğu varsayımlarında bulunulur Eşitlik 2. 11 yazılabilir.

$$\frac{dI}{dt} = 0 \quad (2.11)$$

Bu eşitlikten yola çıkarak nesnenin parlaklığı daha önceki ve daha sonraki çerçevelerde de sabit kalacaktır.

$$I(x, y, t) = I(x + dx, y + dy, t + dt) \quad (2.12)$$

Nesnenin parlaklık bilgisi yani renk özellikleri zamana bağlı olarak değişiyorsa hareket tespiti imkânsızlaşır ve pikselin bir sonraki nesnede yeri hesaplanamaz. Bundan dolayı, nesnenin t anındaki renk özelliklerinin Şekil 2. 5’de görüldüğü gibi $t + dt$ anında da aynı olduğu kabul edilir.



Şekil 2. 5 (x, y, t) konumundan, $(x+dx, y+dy, t+dt)$ konumuna hareket eden bir noktanın renk bilgisi değişmezliği

Parlaklığın zamansal düzlemde değişmediği varsayımı ile $p(x, y)$ pikselinin (dx, dy, dt) değişimi için bir noktadaki şiddeti Taylor Açılımı (2.13) ile gösterilebilir.

$$I(x + dx, y + dy, t + dt) = I(x, y, t) + dx \frac{\partial I}{\partial x} + dy \frac{\partial I}{\partial y} + dt \frac{\partial I}{\partial t} \quad (2.13)$$

Şiddet değişmezliği varsayımınca $I(x, y, t) = I(x + dx, y + dy, t + dt)$ eşitliği sağlandığı için Eşitlik 2. 14 elde edilir.

$$dx \frac{\partial I}{\partial x} + dy \frac{\partial I}{\partial y} + dt \frac{\partial I}{\partial t} = 0 \quad (2.14)$$

Şiddet değişmezliği varsayımına ek olarak zamansal değişimin en az seviyede olduğu, $dt \rightarrow 0$, varsayımı yapıldığı için zamana göre türev alınarak eşitlik tekrar düzenlenir(2.15-2.17).

$$\frac{dx}{dt} \frac{\partial I}{\partial x} + \frac{dy}{dt} \frac{\partial I}{\partial y} + \frac{dt}{dt} \frac{\partial I}{\partial t} = 0 \quad (2.15)$$

$$\frac{\partial I(x, y)}{\partial t} = - \left(\frac{dx}{dt} \frac{dy}{dt} \right) \left(\frac{\frac{\partial I}{\partial x}}{\frac{\partial I}{\partial y}} \right) \quad (2.16)$$

$$I_t = -(v_x \ v_y) \begin{bmatrix} I_x \\ I_y \end{bmatrix} \quad (2.17)$$

Şekil 2. 17'de $v_x = \frac{dx}{dt}$ ve $v_y = \frac{dy}{dt}$ optik akışın x ve y bileşenlerini I_x, I_y ve I_t , (x,y,t) noktasındaki şiddetin türevleridir. Eşitlik 2. 17 (2.18)'deki gibi tekrar düzenlenirse;

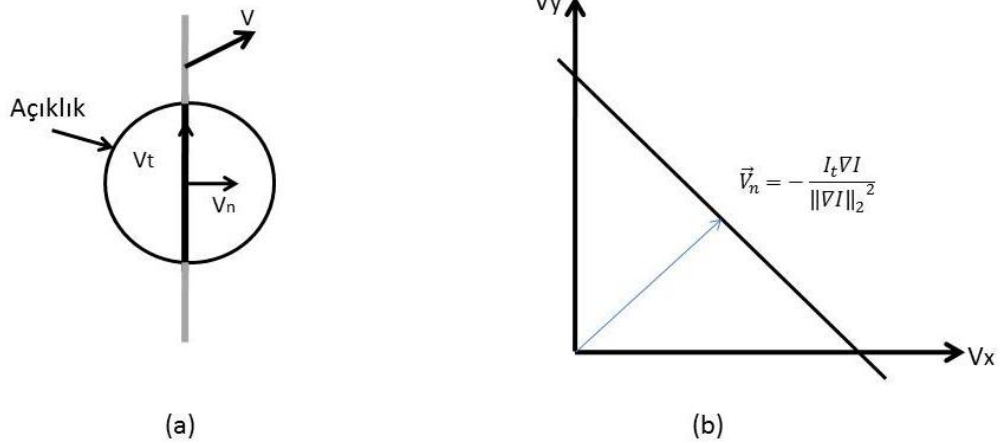
$$(I_x, I_y) \cdot (v_x, v_y) = -I_t \quad (2.18)$$

başka bir ifadeyle

$$\nabla I \cdot \vec{v} = -I_t \quad (2.19)$$

(2.19) eşitliği şeklinde ifade edilebilir. Eşitlik 2. 19'daki $\nabla I = (I_x, I_y)$, uzaysal şiddet gradyenini ve $\vec{v} = (v_x, v_y)$, (x, y) pikselinin t anındaki optik akışıdır. Eşitlik 2. 19 iki boyutlu hareket kısıtlama denklemi olarak adlandırılmaktadır.

İki boyutlu hareket kısıtlama denklemi iki bilinmeyenli bir denklemdir. Bu denklemin çözümü sonucunda Şekil 2. 6(b)' deki gibi bir doğru denklemi elde edilir. Doğrultusu bu doğru üzerinde belirli bir (x,y) üzerinde olan birden fazla hız bileşeni elde edilebilir. Bu nedenle sadece hareket kısıtlama denkleminde elde edilen doğru, hızın tam olarak yerini bulmaya yeterli değildir. Bu problem belirsizlik problemi olarak adlandırılır. Şekil 2. 6(a)'da yukarı ve sağa ilerleyen bir $\vec{V}$ vektörü ve dairesel belirsizlik resmedilmiştir. Bu doğrunun gerçek hızı $\vec{v} = (v_x, v_y)$ belirsizlik probleminden dolayı elde edilemez ama hareket kısıtlama doğrusu üzerinde en küçük genliğe sahip normal hız, $\vec{v}_n$ hesaplanabilir. Bu durumda, yerel şiddet yapısına normal gelen hızın bileşeni örneğin kenar bileşeni elde edilebilir. Sonuç olarak, iki bilinmeyenli bir denklemi çözebilmek için aynı bilinmeyenlere sahip ikinci bir denklem gerektiğinden dolayı, hareket kısıtlama denkleminde ek olarak aynı bilinmeyenlere sahip başka denklem elde etmeye çalışılmıştır.

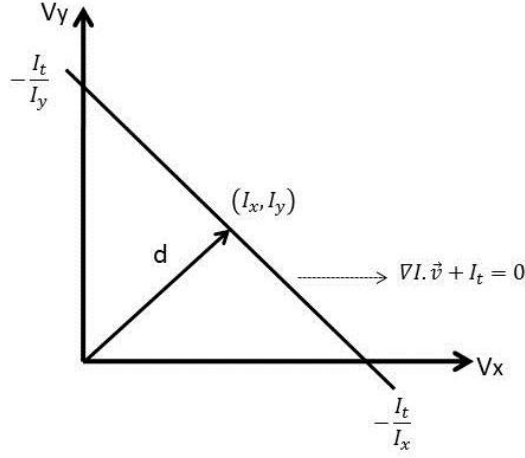


Şekil 2. 6 (a) Belirsizlik Problemi, (b) Hareket kısıtlama denkleminde normal gelen hız. Normalin genliği ve doğrultusu ($\vec{v}_n = v_n \hat{n}$), şiddet türevleri I_x, I_y ve I_t ile hesaplanabilir.

$$v_n = -\frac{I_t}{\|\nabla I\|_2} \quad \text{ve} \quad \hat{n} = \frac{(I_x, I_y)}{\|\nabla I\|_2} \quad (2.20)$$

Eşitlik 2. 20'deki v_n ve $\hat{n}$ sırasıyla, normal hız genliği ve normal hız doğrultusudur. Normal hız vektörü, Eşitlik 2. 21'de gösterildiği gibi bu ikisinin noktasal çarpımıdır. Burada $\vec{v}_n$, birim normal hız ve $\nabla I = (I_x, I_y)$ uzaysal şiddet gradyanıdır.

$$\vec{V}_n = -\frac{I_t(I_x, I_y)}{\|\nabla I\|_2^2} \quad (2.21)$$



Şekil 2. 7 İki boyutlu hareket kısıtlama denkleminin grafiksel gösterimi

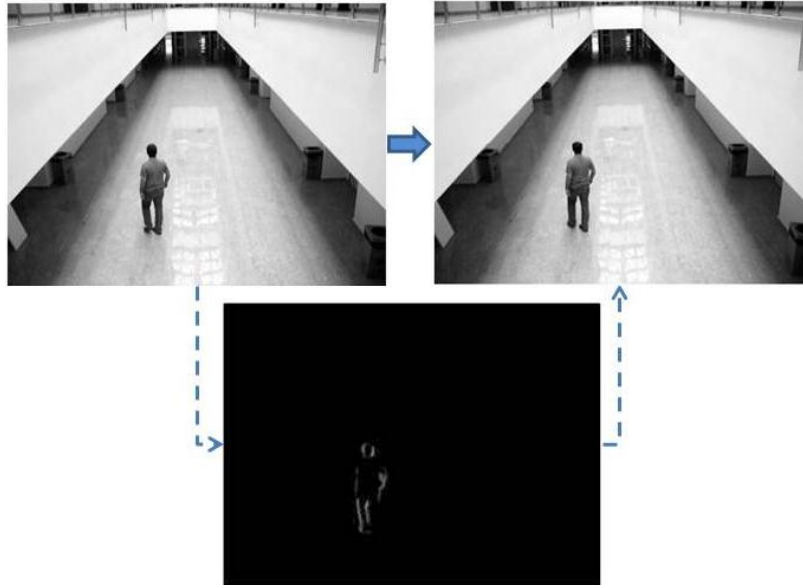
İki boyutlu hareket kısıtlama denkleminde elde edilen doğru Şekil 2. 7’de grafiksel olarak gösterilmiştir. $\nabla I \cdot \vec{v} + I_t = 0$ Eşitliğinden elde edilen x ve y yönündeki hız eşitlikleri (2.20) ve (2.21)’de hesaplanarak, x ve y düzelemlerini kesim noktaları belirlenmiştir.

$$v_y = -\frac{I_x}{I_y}v_x - \frac{I_t}{I_y} \quad (2.22)$$

$$v_x = -\frac{I_y}{I_x}v_y - \frac{I_t}{I_x} \quad (2.23)$$

Buradan $d = \frac{-I_t}{\sqrt{I_x^2 + I_y^2}}$, denklemin üzerindeki belirli bir şiddet gradyanı doğrultusundaki

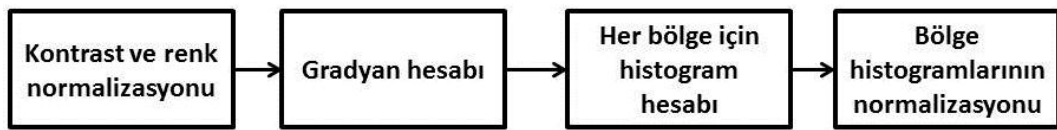
optik akış bileşeni hesaplanır. Şekil 2. 8’ de görüntü içinde insan bulunan bir video çerçevesi ve bu çerçevedeki insana ait optik akış sonucu görülmektedir.



Şekil 2. 8 İki çerçeve arasındaki optik akış

2.2.2 Eğimlerin Histogramı(Histogram of Gradients)

Görüntüdeki piksellerin yönelim(θ) ve büyüklük değerlerinin karakteristiği olarak da adlandırılabilir. Yönlü Gradyanların Histogramı(Histogram of Gradients-HOG) algoritmasının kullanımı Dalal ve Shashua [26] tarafından ortaya atılmıştır. HOG yöntemindeki temel amaç görüntüyü lokal histogramlar olarak ifade etmektir. Bu histogram grupları, görüntünün yerel bir bölgesindeki gradyanların yönelimlerinde, gradyanların büyüklüklerinin toplandığı histogramlardır. Yöntemin genel adımları Şekil 2. 9'daki gibidir.



Şekil 2. 9 HOG algoritmasının genel adımları

Bir görüntünün HOG değerlerinin çıkarılması için gerekli olan hesaplamalar aşağıdaki gibidir. İlk olarak S_y yatay ve S_d dikey Sobel filtreleri uygulanarak görüntünün, I_x ve I_y olmak üzere x ve y yönündeki kenarları belirlenir(2.15 ve 2.16). Sobel filtresi uygulanmış I_x ve I_y görüntüleri kullanılarak, her pikselin gradyanını(G) ve her kenarının yön açılarını(θ) hesaplamak için (2.17) ve (2.18) kullanılır.

$$I_x = I * S_y \quad (2.15)$$

$$I_y = I * S_d \quad (2.16)$$

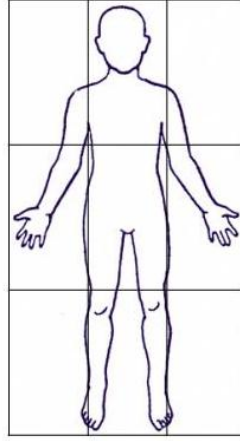
$$|G| = \sqrt{I_x^2 + I_y^2} \quad (2.17)$$

$$\theta = \tan^{-1} \frac{I_x}{I_y} \quad (2.18)$$

HOG uygulanırken görüntü çeşitli sayıda lokal parçalara ayrılabilir. Bu çalışmada, görüntüde dikdörtgen içine alınan nesne 3 satır 3 sütun olmak üzere 9 lokal bölgeye ayrılarak kullanılmıştır. Deneysel sonuçlar özellikle insan tabanlı çalışmalarda HOG'un 9 bölgeye ayrılmasının daha başarılı sonuçlar verdiğini göstermektedir.

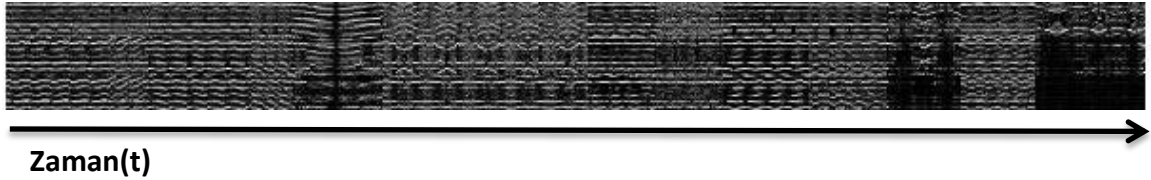
HOG algoritması uygulanırken oluşturulan histogramda piksellerin yönelim açılarının gruplanması ile daha başarılı sonuçlar elde edildiği gözlemlenmiştir. Bu gruplama işlemi 0-360 aralığında yer alan açısal değerlerin istenilen bir aralığa çekilmesi ile mümkün olmaktadır. Örneğin normalde her açı değeri için histogram oluşturulması gerekirken, 0-45 aralığındaki değerler tek bir histogram ile ifade edilmektedir. Bu çalışmada değerler 9 aralığa toplanarak işlem yapılmıştır.

HOG yönteminin birçok avantajı vardır ve en önemlisi de boyut bağımsız olmasıdır. Yönlü gradyanlar kullanıldığı için nesnenin boyutunun histogram değerlerine etkisi yoktur. Bu özelliği sayesinde görüntüde yaklaşan veya uzaklaşan objenin özellikleri boyutuna bağlı olarak değişmez. Kenar özelliklerini kullandığı için nesnelerin hareketlerindeki karakteristik özellikleri başarılı şekilde ifade eder. Özellikle insan üzerinde uygulandığında, görüntüyü bölgelere ayırıp parçalı bir şekilde uygulanması sayesinde insanın hareket eden farklı bölgelerinin ayrı ayrı değerlendirilip sonrasında birleştirilmesiyle daha fazla bilgi taşıyan özellikler elde edilmiş olunur. Şekil 2. 10 çerçeve içindeki bir insanın 9 lokal bölgeye ayrılmış görüntüsüdür.



Şekil 2. 10 Dokuz lokal bölgeye ayrılmış insan görüntüsü

Yukarıda görüldüğü gibi insanın hareket eden parçalarının farklı bölgelere ayrılmasıyla hareketlerin karakteristiğinin çıkarılmasını sağlayacaktır. Bu çalışmada olduğu gibi yukarıdaki şekilde de 9 bölge için 9’ar adet gradyan histogram değer oluşturulur. Dolayısıyla her nesne 81 adet histogram değer şeklinde ifade edilir. Şekil 2. 11’da Weizmann İnsan Hareketi Veri Seti’nde [27] Lena isimli örneğe ait videosuna ait histogram değerleri görülmektedir. Şekilde sütundaki her vektör bir çerçevede bulunan nesneye ait histogramları ifade etmektedir.



Şekil 2. 11 Weizmann Veriseti’ne ait hog özellikleri

2.3 Özelliklerin Değerlendirilmesi

Bu çalışmada, histogram olarak ifade edilen nesneye ait hareket kalıbını bulma işlemini gerçekleştirmeden önce nesnenin insan olup olmadığı değerlendirilir. İnsan tespiti gerçekleştirilmek için insan ve insan olmayan, pozitif ve negatif, örneklerden oluşan çevrimdışı bir eğitim seti oluşturularak sınıflandırıcı eğitilmiştir. Eğitim işlemi gerçekleştirilirken örneklere ait HOG özellikleri çıkarılmıştır. Sınıflandırıcı olarak Destek Vektör Makinesi(Support Vector Machine-SVM) kullanılmıştır. Video çerçevesinden elde edilen nesnenin insan olup olmadığını tespit etmek için eğitim setindeki gibi test

nesnesinin de HOG özellikleri kullanılarak hangi sınıfa ait olduğu tespit edilmiştir. İnsan olduğu tespit edilen nesnelerin takibi için Kalman Filtresi kullanılmıştır.

İnsan olduğu tespit edilen nesneler için HOG özellikleri ardışık olarak sıralandığında yani zaman bilgisi eklendiğinde buradan belirli aralıklarla tekrarlayan kalıpların bulunması gereklidir. Çıkarılan özellikler arka arkaya sıralandığında belirli desenlerin tekrar ettiği gözle görülmektedir. Fakat bu desenlerin başlangıç noktaları ve kalıp uzunlukları bilinmemektedir. Tekrar eden kalıp bulma işlemi için bu çalışmada kullanılan yöntem şablon eşleme(template matching) [28] yöntemidir.

2.3.1 İnsan Tespiti

İnsan tespiti gerçekleştirilirken eğitim setinde INRIA veriseti, MIT yaya veriseti [29] ve Caltech veriseti [30] kullanılmıştır. Inria veriseti 64x128 boyutunda 2416 pozitif, 1218 negatif örnek bulunmaktadır. MIT veriseti yine 64x128 boyutunda 924pozitif örnek içerirken, Caltech veriseti çeşitli kategorilerden oluşan 329 negatif örnek bulundurmaktadır. Toplam 3340 pozitif, 1547 negatif örnekle eğitim gerçekleştirilmiştir. Şekil 2. 12’da pozitif ve negatif örneklere ait görüntüler bulunmaktadır.



Şekil 2. 12 Pozitif ve negatif örnekler

Sınıflandırıcı eğitilirken 3340 pozitif örneğin ortadan ikiye bölünmesiyle vücudun sadece üst kısmının ve sadece alt kısmının görüldüğü durumlar için de toplam 6680 pozitif örnek verisetine eklenmiştir. Sınıflandırıcı toplam 10020 pozitif ve 1547 negatif örnekle eğitilmiştir. Pozitif örneklerin ikiye bölünmesiyle oluşturulan pozitif örnekler Şekil 2. 13’de görülmektedir.

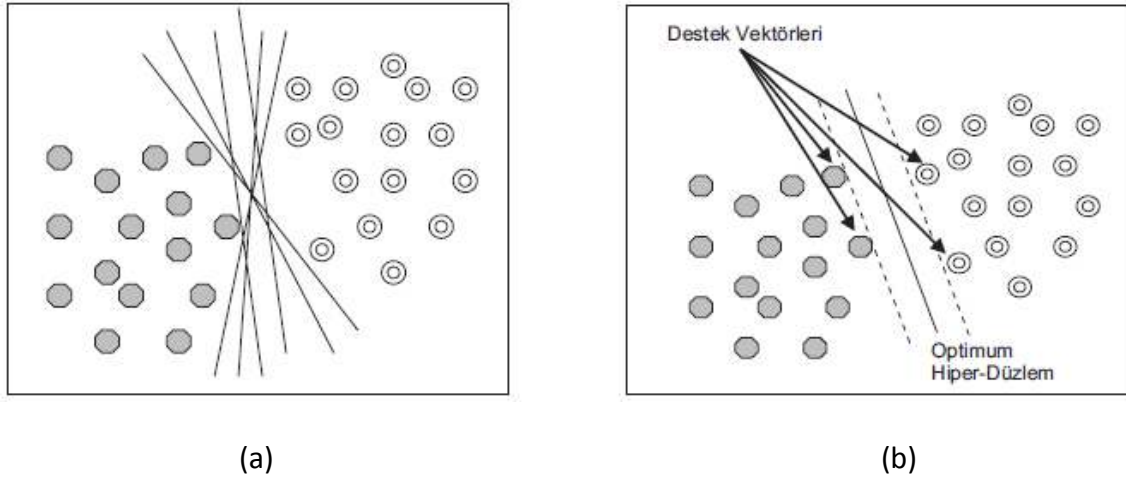


Şekil 2. 13 Yeni oluşturulan pozitif örnekler

Şekildeki pozitif örneklerle eğitim kişinin vücudunun tamamının herhangi bir nedenden dolayı görünmediği durumlarda test edilen nesnenin insan olarak sınıflandırılmasını sağlamıştır.

2.3.1.1 Destek Vektör Makineleri

Destek Vektör Makineleri(DVM)'nin çalışma prensibi iki sınıfı birbirinden ayırabilen en uygun karar fonksiyonunun tahmin edilmesi, başka bir ifadeyle iki sınıfı birbirinden en uygun şekilde ayırabilen hiper düzlemin tanımlanması esasına dayanmaktadır. Hiper düzlem iki sınıfı birbirinden ayıran düzleme verilen isimdir. Şekil 2. 14(a) hiper düzlem örnekleri ve Şekil 2. 14(b)'de optimum hiper-düzlem ve destek vektörleri görülmektedir.

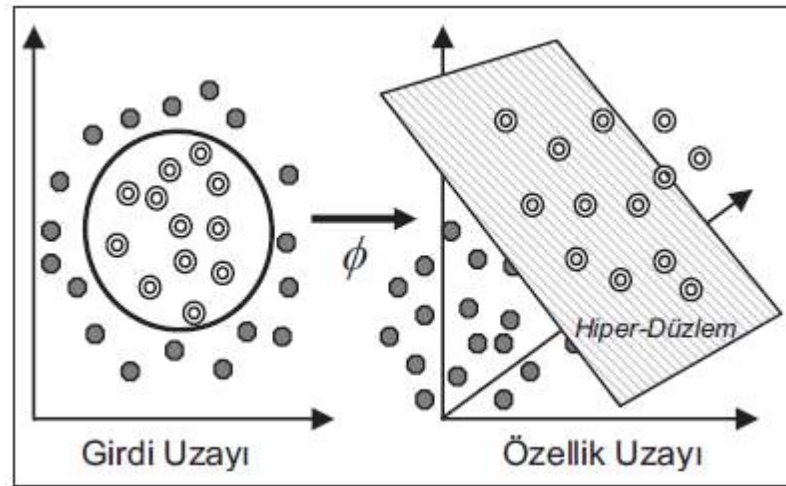


Şekil 2. 14 (a) İki sınıf için hiper düzlemler, (b) Optimum hiper düzlem ve destek vektörleri

Şekildeki optimum hiper düzlemin belirlenebilmesi için bu düzleme paralel ve sınırları oluşturacak iki hiper düzlem belirlenmesi gerekir. Bu hiper düzlemleri oluşturan noktalar destek vektörleri olarak adlandırılırlar.

Doğrusal olarak ayrılamayan veriler, özellik uzayı olarak tanımlanan yüksek boyutlu bir uzayda görüntülenir. Böylece verilerin doğrusal olarak ayrımı yapılabilmekte ve sınıflar

arasındaki hiper düzlem belirlenebilmektedir. Destek Vektör Makinelerinde kernel fonksiyonu olarak tanımlanan uzay değıştirme fonksiyonları yardımıyla doğrusal olmayan dönüşümler yapılabilmekte ve bu şekilde verilerin yüksek boyutta doğrusal olarak ayırımına imkân sağlanmaktadır. DVM ile gerçekleştirilecek bir sınıflandırma işlemi için kullanılacak kernel fonksiyonu ve bu fonksiyona ait optimum parametrelerin belirlenmesi gereklidir. Şekil 2. 15’de girdi uzayında doğrusal olarak ayıramayan verinin özellik uzayında hiper düzlemle ayrılabildeği görölmektedir.



Şekil 2. 15 Kernel fonksiyonu ile verinin daha yüksek bir boyutta ifade edilmesi

Bu çalışmada farklı kernel fonksiyonu performanslarının değeriendirilmesi sonucunda polinomsal kernel fonksiyonu, dönüşüm fonksiyonu olarak seçilmiştir. Yine bu kernelda parametre olarak kullanılması gereken polinom derecesi başarı sonuçları değeriendirilerek 3 olarak belirlenmiştir.

2.3.2 İnsan Takibi

Nesne takibinin amacı sıralı çerçeveler arasında nesneler arası ilişkinin kurulup yörünge, hız, yön gibi zamansal özelliklerin çıkarılmasıdır. Tespit edilen nesnenin çerçeve çerçeve takip edilmesi önemli ve zor bir işlemdir. Akıllı gözetleme sistemleri için önemli bir kısım olan nesne takibi olmadan takip edilen nesnenin birbiriyle bağlantılı özelliklerinin çıkarılması ve üst seviye davranış analizinin yapılması mümkün değildir. Bu çalışmada üzerinde yoğunlaşılın hareketli nesne insandır. Takip mekanizmasına geçilmeden önce hareketli önplan çıkarımı yapılarak nesnenin insan

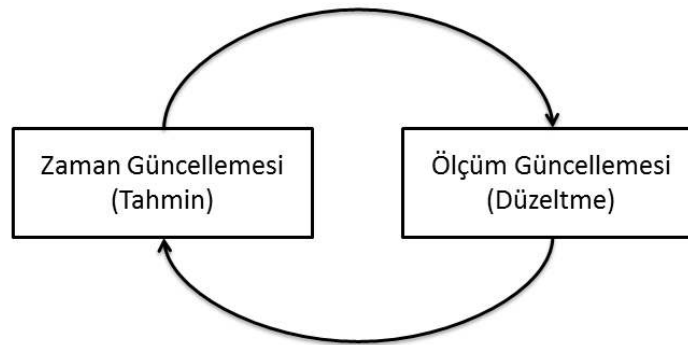
olup olmadığına karar verilmekte ve insan olduğuna karar verilen hareketli nesnenin kalman filtresi ile takibi yapılmaktadır.

Kalman filtresi mühendislikte ve gömülü sistemlerde çokca kullanılan, işlenen işaret içersindeki gürültüyü azaltırken gerekli bilgiyi koruyabilen bir metottur. Çok çeşitli işlemlerin değişkenlerinin tahmininde kullanılan ve lineer sistemler üzerinde çalışabilen kalman filtresi uygulamada da düzgün tahmin sonuçları verdiği için seçilmiştir.

2.3.2.1 Kalman Filtresi

R.E.Kalman [31] 1960 yılında özyinelemeli ayırık veri lineer filtreleme probleminin çözümüne yönelik makale yayınlamıştır. O günden bugüne kalman filtresi özellikle otonomluk ve destekli gezinim başta olmak üzere çok geniş alanlarda ve uygulamalarda kullanılmıştır. Nesne Takibi için de kullanılabilen bu yönteminin istatistiksel ve hesaplama dayalı kökenini ve uygulamasını Welch ve Bishop [32] yaptıkları çalışmayla detaylı bir şekilde anlatmışlardır.

Kalman filtresi özyinelemeli bir tahmin aracıdır. Bunun anlamı önceki adımda tahmin edilen durum ile o anki ölçümün bir sonraki durumun tahminin yapılabilmesi için yeterli olmasıdır. Bu yöntem gürültü ve sistem dinamikleri hakkındaki bilgiyi gürültülü ölçümden gelen belirsizliği gidermede kullanır.



Şekil 2. 16 Ayırık Kalman Filtresi döngüsü

Kalman filtresi bir işlemi geri besleme kontrolünü kullanarak tahmin eder. Belli bir zamanda işlem durumunu tahmin eder, sonrasında ise ölçümden geri beslemeyi edinir. Kalman filtresi denklemleri zaman güncellemesi ve ölçüm güncellemesi olarak ikiye ayrılır. Zaman güncellemesi denklemleri mevcut durumu ilerletmekle ve bir sonraki

adımda kullanılmak üzere önsellik tahminini elde etmek için hata kovaryans tahminini güncellemekle sorumludurlar. Ölçüm güncellemeleri ise geri bildirimden sorumludurlar. Zaman güncelleme denklemleri aynı zamanda tahmin edici olarak düşünülebiyecekken ölçüm güncelleme denklemleri düzeltici denklemler olarak düşünülebilir. Şekil 2.16'da görüleceği üzere sonuç tahmin algoritması bu iki eşitlik grubunun birlikteliği ile çalışır.

$$\hat{x}_t^- = A\hat{x}_{t-1} + B u_{t-1} \quad (2.19)$$

$$P_t^- = AP_{t-1}A^T + Q \quad (2.20)$$

Eşitlik (2.19) ve (2.20)'de görülen eşitlikler zaman güncellemesi denklemleridir. İlk denklemden geçen $\hat{x}_k^-$ t zamandaki tahmin edilen önsel durumu ifade eder. A durum geçiş matrisi iken B ise u_k kontrol giriş vektörüne uygulanan kontrol giriş modelidir. Eşitlik (2.20)'de geçen P_t^- t zamanındaki önsel hata kovaryans matrisi, Q ise işlem gürültü kovaryans matrisidir.

$$K_t = P_t^- H^T (H P_t^- H^T + R)^{-1} \quad (2.21)$$

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-) \quad (2.22)$$

$$P_t = (I - K_t H) P_t^- \quad (2.23)$$

(2.21), (2.22) ve (2.23)'teki eşitlikler ölçüm güncelleme denklemleridir. Bu eşitliklerle tahmin edilen durum bilgileri gerçek ölçümler kullanılarak düzeltilir. Eşitlik (2.21)'de geçen K_t kalman kazancıdır. Kalman kazancı ağırlıklı düzenleyici olarak görev yapar. İşlem gürültüsünün kovaryansı olarak ifade edilen R büyüdükçe Kalman kazancı küçülecektir. H ifadesi gözlem uzayından gerçek durum uzayına geçiş yapmayı sağlayan gözlem modelidir. Eşitlik (2.22)'de ifade edilen z_t ise t anındaki ölçüm bilgisini gösterir.

Hareketli bir nesnenin fiziksel olarak t anındaki konum bilgisinin elde edilebilmesi için eşitlik (2.24)'te ifade edildiği gibi t-1 anındaki konum, hız ve ivme bilgilerine ihtiyaç duyulur.

$$x_t = x_{t-1} V_{t-1} t + \frac{1}{2} a_{t-1} t^2 \quad (2.24)$$

Bu tez çalışmasında hareketli nesne takibi yapılırken blob analizi sonucu elde edilen nesnenin orta noktasının konum ve hız bilgisinden faydalanılmış ve ivme 0 kabul edilerek eşitlik (2.25) ile gösterilen fiziksel sistem dinamiği kullanılmıştır.

$$x_t = x_{t-1} + V_{t-1} dt \quad (2.25)$$

Bu ifade matris formunda yazılacak olursa eşitlik (2.26) elde edilir.

$$\begin{bmatrix} x_{x,t} \\ x_{y,t} \\ V_{x,t} \\ V_{y,t} \end{bmatrix} = \begin{bmatrix} 1 & 0 & dt & 0 \\ 0 & 1 & 0 & dt \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{x,t-1} \\ x_{y,t-1} \\ V_{x,t-1} \\ V_{y,t-1} \end{bmatrix} \quad (2.26)$$

Görüldüğü gibi (2.19)'da ifade edilen kalman eşitliği elde edilmiş olur. A matrisi ve $\hat{x}_t^-$ 'in karşılıkları (2.27)'deki gibidir. Dolayısıyla kontrol giriş matrisi olan B'nin değeri 0 olacaktır.

$$\hat{x}_t^- = A \hat{x}_{t-1} \Rightarrow \hat{x}_t^- = \begin{bmatrix} x_{x,t} \\ x_{y,t} \\ V_{x,t} \\ V_{y,t} \end{bmatrix}, A = \begin{bmatrix} 1 & 0 & t & 0 \\ 0 & 1 & 0 & t \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \hat{x}_t^- = \begin{bmatrix} x_{x,t-1} \\ x_{y,t-1} \\ V_{x,t-1} \\ V_{y,t-1} \end{bmatrix} \quad (2.27)$$

(2.27)'de görülen $x_{x,t}$ ve $x_{y,t}$ ifadeleri takip edilen noktanın x ve y koordinatlarının t anındaki tahminini gösterirken $V_{x,t}$ ve $V_{y,t}$ ifadeleri de aynı noktaların x ve y koordinat yönündeki hız tahminlerini göstermektedir.

Filtrenin gerçekleşmesinde ölçüm kovaryans matrisi olan R'nin hesaplanması nesnenin sabit konumda olduğu durumda sistemin bir süre çevrim dışı olarak çalıştırılması sonucu elde edilir. Tahmin edilen işlemin aynı zamanda gözlemlenmesi imkânının olmamasından dolayı işlem gürültü kovaryans matrisi Q'nun belirlenmesi oldukça zordur. Bu yüzden ölçüm değerlerinin güvenilir olduğu varsayımı yapılmıştır. Takip edilen nesnenin başlangıç konumu ve hızının tam olarak bilindiği durumlarda P işlem gürültüsü kovaryans matrisine 0 değeri verilebilir fakat hız ve konum bilgisi tam olarak bilinmiyorsa P matrisinin diyagonalini oluşturan sayılara büyük değerler atanarak kalman filtresinin çalışması sağlanabilir.

Gerçekleştirilen sistemde optik akış kullanıldığından dolayı nesnenin sabit durumda olması önplan çıkarımı yapılmasına imkân vermeyecektir. Bu yüzden R kovaryans matrisinin başlangıç değerlerine 1 ile 0 arasında değişen küçük değerler atanmıştır. Aynı zamanda hareketli nesnenin başlangıçtaki konum ve hız bilgisi bilinmediğinden dolayı P matrisinin diyagonalinde bulunan sayılara 100 değeri atanarak filtrenin düzgün sonuçlar üretmesi sağlanmıştır.

2.3.3 Kalıp Eşleme Yöntemleri

Bu çalışmada, bir resimdeki görüntü parçasının diğer resimdeki yerinin bulunması için Şablon Eşleme [33] yöntemi kullanılmıştır. Bu yöntemin kullanılmasının nedeni gradyan histogramları çıkarılan çerçevelerin, histogramlarının zaman bilgisi eklenerek ardışık sıralanmasıyla insana ait hareket kalıplarının tekrar eden kalıplar şeklinde bir görüntü oluşturduğunun gözlemlenmesidir. Şablon Eşleme metotları aşağıdaki gibi sıralanabilir.

- Alana dayalı eşleme(en küçük kareler yöntemi ve çapraz korelasyon)
- Detaylara dayalı eşleme(nokta detayla, kenar bulma, bölge çıkarımı)
- Görüntü ilişkilerine dayalı eşleme(piksellerin komşuluk ilişkileri)

Alana dayalı görüntü eşleme yönteminin amacı, bir görüntü üzerindeki bir parçanın diğer görüntüdeki karşılığının gri değerleri yardımıyla bulunmasıdır.

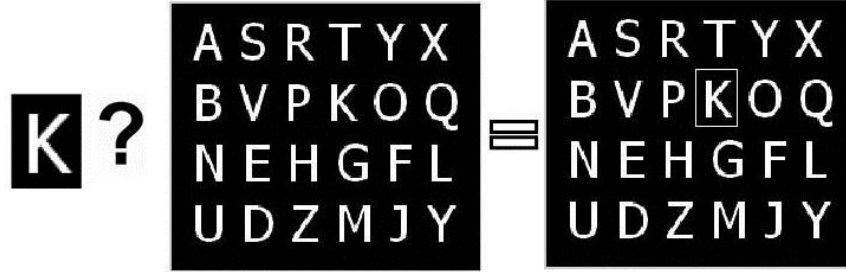
İki görüntü arasındaki gri değer varyasyonları korelasyon veya en küçük kareler yöntemiyle değerlendirilerek benzerlik tespit edilir. Alana dayalı eşleme yönteminde çapraz ilişki ve en küçük kareler eşlemesi kullanılır.

Detaya dayalı görüntü eşleme tekniğinde görüntülerin sembolik tanımlamaları kullanılır. Detaya dayalı görüntü eşleme iki aşamadan oluşur. Birincisi, tüm görüntülerde ilgili detayların ve özelliklerinin çıkarılması, ikincisi ise benzer detayların tüm görüntülerde tanımlanmasıdır. Detay bazlı eşleştirmede gri değerleri görüntüden nesnelerin detaylarını çıkarmak için kullanılır. Her bir detay nokta kenarlar ve alanlardan oluşur [34].

$$\rho = \frac{\sum_{i,j} [g_1(c_1, r_1) - \mu_1] [g_2(c_2, r_2) - \mu_2]}{\sqrt{\sum_{i,j} [g_1(c_1, r_1) - \mu_1]^2 \sum_{i,j} [g_2(c_2, r_2) - \mu_2]^2}} \quad (2.28)$$

$$\mu_1 = \frac{1}{n} \sum_{i,j} g_1(c_1, r_1) \quad , \mu_2 = \frac{1}{n} \sum_{i,j} g_2(c_2, r_2) \quad (2.29)$$

(2.28)'da gösterilen ρ korelasyon katsayısını, $g(c,r)$ pikselinin gri değerini, c_1, r_1 aranan görüntünün piksel koordinatları, c_2, r_2 içinde arama yapılan görüntünün piksel koordinatlarının penceredeki toplam piksel sayısını ve i,j de korelasyon penceresindeki piksel indeksini gösterir.



Şekil 2. 17 Alan tabanlı eşleme örneği

Nokta detaylar yönteminin ana fikri ayrı ayrı görüntüdeki noktaları yüksek olasılıklı bölgelerde tanımlamaktır. Fakat bu alanlar eşleştirme için kullanışlı olmalıdır. Detay noktalarını eşleştirmenin kolay yolu bölge bazlı eşleştirmedir. Detay noktalarını sınıflandırılmanın en kolay yolu yeterli bilgiye sahip görüntü parçalarını bulmakla olur. Bu sayede korelasyon başarılı olmak için en yüksek olasılığa sahip olur. Şekil 2.17'de sadece "K" harfi bulunan resim en solda ve içinde "K" dahil 24 harf bulunan içinde arama yapılan resim ortada aralarında toplam sembolü ile gösterilmektedir. Eşitlik ifadesinin sağ tarafında da birinci resmin ikinci resim içinde bulunduğu bölge çerçeve içinde gösterilmektedir.

Kenar bulma tabanlı yöntemin uygulanması için hali hazırda bir çok kenar bulma yöntemi mevcuttur ve bunların her biri ayrı ayrı karakteristik özelliğe sahiptir. Operatörlerin seçimi kişisel öncelik veya özel deneyimlere göre farklılık gösterir. Kenar bulma, kenar piksellerini ve bir bölgedeki grup halindeki pikseller olan "kenar

konturları”nı belirlemeyi içerir. Bu iki adım arasında önemli farklılıklar vardır. Örneğin beyaz pikselleri bilinen kenar piksellerine sıralamak, siyahları da diğer bütün piksellere yerleştirmek, tüm kenarların rahatça görülmesini sağlar ama kenarlar arasındaki komşuluk ilişkileri kesin olarak belirlenemez. Fakat kenar konturları da yalnızca kapalı alan olarak ifade edilir ve bu da kenar eşleştirme metotlarının çoğuna uygun değildir. Bu metotta sıfır köşeli konturlar öncelikle piksel piksel eşleştirilir. Daha sonra büyük alanlar içeren sınıflarda eşleştirme yapılır. Şekil 2. 18’de Sobel operatörü ile yapılmış bir kenar bulma işlemi gösterilmiştir [34]. Bu tez çalışmasında alan tabanlı çapraz korelasyon yöntemi kullanılmıştır.



Şekil 2. 18 Gri seviye bir görüntü ve Sobel Operatörü ile bulunan kenarları

2.3.4 Hareket Kalıbının Bulunması ve Kalıp Tekrarının Aranması

Hareket kalıbının bulunması için iki temel problem vardır. Bir video içinde ne hareket kalıplarının başlangıç çerçeveleri, ne de kalıpların uzunlukları bilinmektedir. Bu problemin çözümü için öncelikle başlangıç noktası sabit kabul edilerek hareket kalıbı büyüklüğü artırılır. Eğer hareket kalıbı büyüklüğü belirli bir değere geldiği halde kalıp bulunamadıysa, kalıp başlangıç noktası bir kaydırılarak ve kalıp büyüklüğü başlangıç değerine alınarak kalıp eşleme yöntemiyle arama yapılmaya devam edilir. Aşağıdaki algoritmada hareket değişimin başlangıç noktalarının ve hareket kalıplarının nasıl bulunduğu dair algoritmanın sözde kodu görülmektedir. Korelasyon değerleri görüntü eşleme yönteminin kullanılmasıyla elde edilen korelasyon değerleridir.

Çizelge 2. 1 Sistemin sözde kodu

Kalıp Çıkarımı ve Kalıp Arama	
1 :	İklendirilmiş başlangıç noktası
2:	İklendirilmiş kalıp büyüklüğü
3:	WHILE (Video sona ermediği sürece)
4 :	IF (Hareket kalıbı bulunmadı)
5 :	Kalıp büyüklüğünü artır
6 :	IF ((Kalıbın korelasyon değeri > Otomatik korelasyon değeri) & (Toplam HOG farkı < Eşik değeri))
7 :	Hareket Kalıbı Bulundu, Hareket Değişimi Bulundu
8 :	Hareket uzunluğu = Kalıp büyüklüğü
9:	END IF
10:	IF (Kalıp büyüklüğü > 2 saniyedeki toplam çerçeve sayısı)
11:	İklendirilmiş kalıp büyüklüğü, Hareket başlangıç noktasını bir kaydır
12:	END IF
13:	ELSE
14:	Olası hareket kalıbı tekrarı başlangıcı = Bulunan hareket kalıbı bitişi + 1
15:	WHILE (Video sona ermediği sürece)
16:	IF ((Olası hareket kalıbı korelasyon değeri > Bir sonraki olası hareket kalıbı başlangıcı) & (Olası hareket kalıbı korelasyon değeri > Bir önceki olası hareket kalıbı başlangıcı) & (Olası hareket tekrarı korelasyon değeri > Ortalama hareket kalıbı korelasyon değeri))
17:	Hareket kalıbı tekrarı bulundu
18:	Hareket kalıbı tekrarı arama başlangıç noktasını, bulunan kalıp tekrarının bitişine kaydır
19:	ELSE
20:	Hareket kalıbı tekrarı arama başlangıç noktasını bir kaydır
21:	END IF , İklendirilmiş kalıp büyüklüğü
22:	END WHILE
23:	END IF
24:	END WHILE
25:	Özet video başlangıç noktaları = Bulunan hareket değişimi noktaları

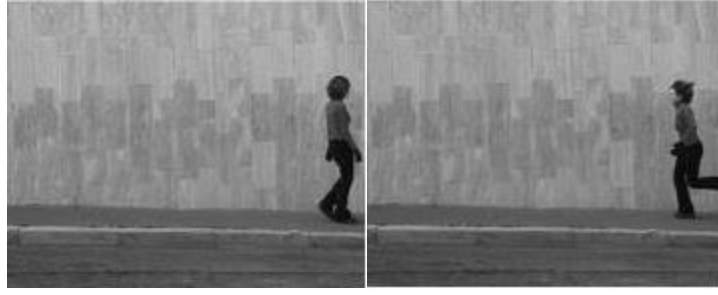
BÖLÜM 3

UYGULAMA

Geliştirilen uygulama fark verisetleri üzerinde güvenlik kamerası görüntülerin oluşabilecek durumlar baz alınarak test edilmiştir. Kullanılan verisetleri genel olarak 4 ana başlık altında toplanabilir. Bunlar sırasıyla genel kullanıma açık olan Weizmann İnsan Hareket Veriseti, tek kişilik 90 derecelik veriseti, tek kişilik farklı açılardan oluşan veriseti ve çoklu insan veriseti olarak sıralanabilir.

3.1 Weizmann İnsan Hareketi Veriseti

Geliştirilen uygulamaya daha önce yine insan tabanlı özetlemeyi hedef alan çalışma baz alınarak Weizmann İnsan Hareketi Veriseti üzerinde test edilmiştir. Bu veriseti genel kullanıma açıktır. Veri setinde ayrı ayrı videolar halinde 9 kişiye ait 10 farklı hareket vardır. Yapılan test işlemleri sonucunda alınan sonuçlar Çizelge 3. 1' de görülmektedir. Elde edilen başarılı sonuçlara rağmen veriseti incelendiğinde, özetleme işlemi verisetinde parça parça bulunan hareket videolarının birleştirilmesiyle elde edilmiştir. Yani temelde videoda bir bütünlük yoktur. Hareket geçişi bulunması için yapay bir verisetidir ve bu videoda bulunan hareket geçişlerinin bulunması gerçek görüntülerdeki geçişlere göre daha kolaydır. Bunun nedeni de hareketler arasında keskin geçişlerin olmasıdır. Şekil 3. 1' de arka arkaya gelen iki çerçeve arasındaki geçiş görülmektedir.



Şekil 3. 1 Weizmann Veriseti’nden yararlanılarak oluşturulmuş videoda arka arkaya gelen iki görüntü

Çizelge 3. 2 Weizmann Veriseti ile yapılan test sonuçları

Video No.	Çerçeve Sayısı	Bulunan Kalıp / Kalıp Sayısı	Bulunan Hareket / Hareket Sayısı	Özet Video Çerçeve Sayısı
1	637	29 / 33	8 / 9	168
2	496	27 / 31	8 / 9	171
3	546	22 / 24	8 / 9	174
4	733	25 / 36	8 / 9	178
5	533	23 / 26	8 / 9	163
6	746	33 / 36	11 / 12	236
Toplam:	3691	159 / 186 (%85,48)	51 / 57 (%89,47)	1090

Weizmann Veriseti’ndeki 9 kişiden test verisi olarak kullanılan kişiler Çizelge 3. 1’deki sırasıyla Daria, Denis, Eli, İra, Shahar ve Lena’dır. Ayrıca tüm kişilerde eğilme hareketi tekrarlı olmadığı için kullanılmamıştır. Uygulamanın ilk test aşaması olduğu için tekrar eden hareketler bulunan veri kullanılmıştır. Her kişi için bulunamayan sadece 1 adet hareket vardır ve bunun neden olduğu çeşitli sayıda bulunamayan hareket tekrarı vardır. Şekil 3. 2’de bulunamayan harekete ait iki adet görüntü görülmektedir. Şekilde de görüleceği gibi hareket eden kişi kutu içine alındığında yani genel hareket bilgisi yok edildiğinde, fazla bir özellik değişimi görülmemektedir. Özellikle HOG yönteminin

kullandığı kenar bilgisinde çok az değişim olduğu için sistem bu hareketi bulmakta başarısız olmaktadır.



Şekil 3. 2 Bulunamayan harekete ait birbiri arasında 6 görüntü fark olan iki görüntü

3.2 Tek Kişi Yandan Dik Videolarla Oluşturulmuş Veriseti

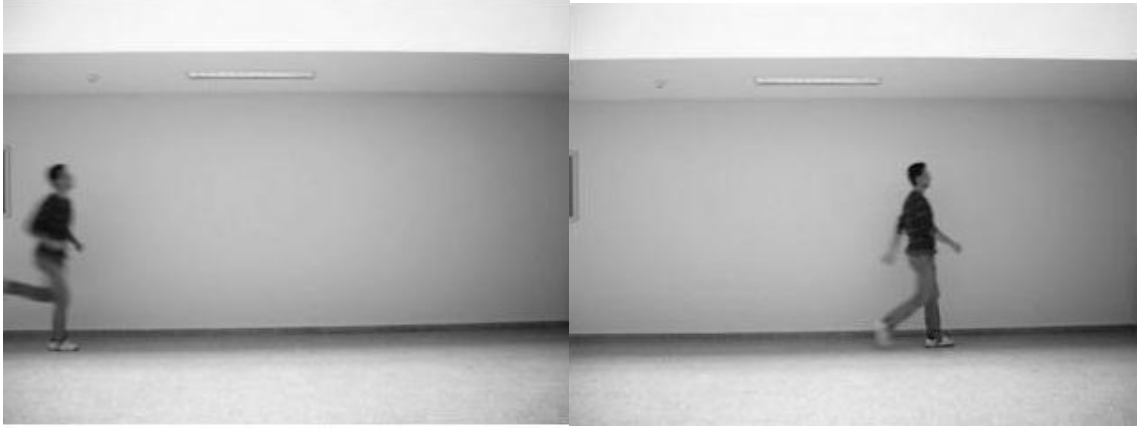
Weizmann verisetinden sonra kendi oluşturduğumuz verisetini üzerinde sistemin çalışması kontrol edilmiştir. Bu veriseti kişiyi yan taraftan 90 derece ile görecektir. Bu veriseti kişiyi yan taraftan 90 derece ile görecektir ve kişinin hareket ettiği seviyeden yapılan çekimler oluşturulmuştur. Oluşturulan verisetinde 3 farklı kişiyle çeşitli durumlar oluşturulmuştur. Bunlar yürümeden koşmaya geçiş, koşmadan yürümeye geçiş, yere bir obje bırakılıp başka biri tarafından tekrar alınması, yere bir objenin düşürülmesi, elinde bir obje ile yürüme gibi bir güvenlik kamerasının görüntülerinde oluşabilecek durumlar oluşturulmuştur. Şekil 3. 3'te bir kişiye ait yürüme başlangıcı ve yürümeden koşmaya geçiş başlangıcı görülmektedir.



Şekil 3. 3 Yürüme ve yürümeden koşmaya geçiş başlangıç görüntüleri

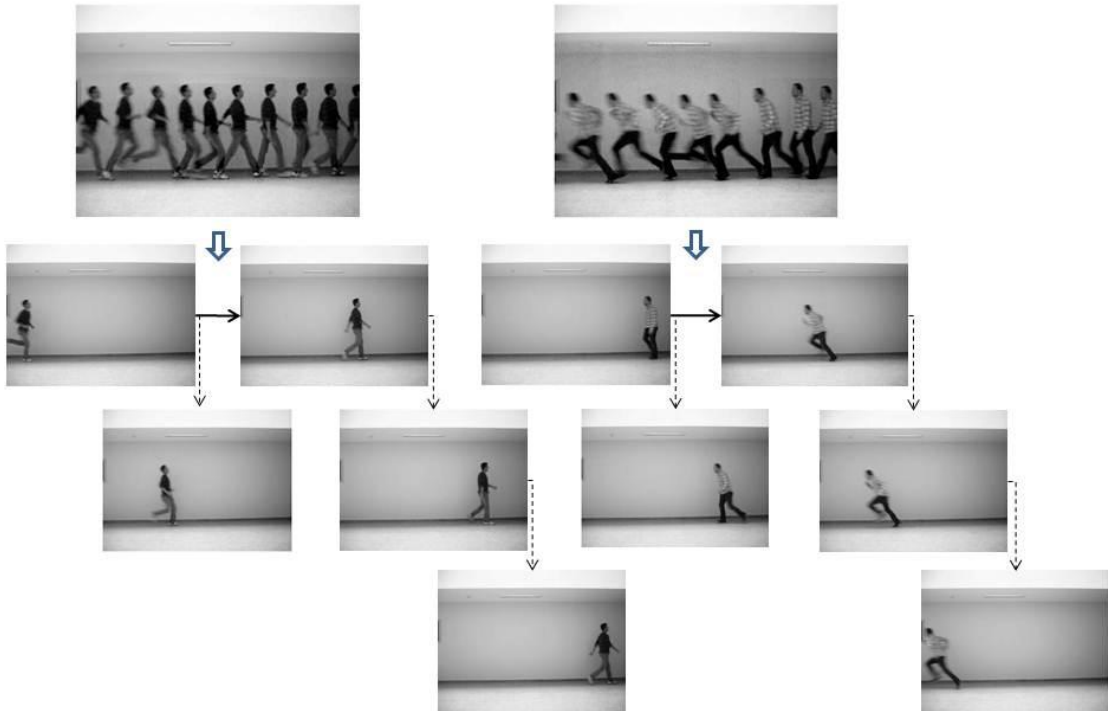
Yukarıdaki videoda sistem gayet başarılı çalışarak tek bir görüntü dizisi olan yürüme ve yürümeden koşmaya geçişi doğru şekilde tespit etmiştir. Böyle bir olayın tespit

edilmesi, olası bir gasp ve hırsızlık anını yakalanabileceğini göstermektedir. Yürümeden koşmaya geçişte olduğu gibi koşmadan yürümeye geçiş durumunda da sistem başarılı sonuçlar vermektedir. Böyle bir geçiş Şekil 3. 4'te görülmektedir.



Şekil 3. 4 Koşma ve koşmadan yürümeye geçiş başlangıcı

Şekil 3. 5 yürüme koşma hareket geçişlerini adım adım, Şekil 3. 3 ve Şekil 3. 4'te görülen hareket geçişlerinin daha ayrıntılı gösterimidir.



Şekil 3. 5 Yürüme-koşma ve koşma-yürüme geçişleri

Yukarıda en üstte bulunan iki adet görüntü kişilere ait video akışının üst üste birleştirilmesiyle oluşturulmuştur. Üstten ikinci sıradaki 4 görüntü hareket geçişlerini

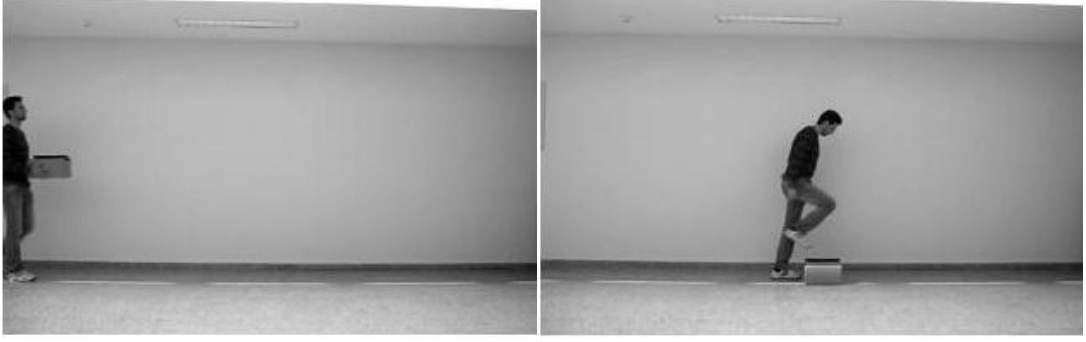
göstermektedir. Bu görüntülerden kesik çizgili ok ile gösterilenler de aynı hareketin tekrar başlangıç noktasıdır. Son iki görüntü satırı tekrar görüntülerinin tekrar başlangıçlarını göstermektedir. Genel olarak incelersek toplam 4 hareket geçişi, bu geçişlerin de toplamda 10 adet tekrarı vardır.

Yürüme ve koşma geçişler bir güvenlik kamerası görüntülerinde oluşabilecek en genel görüntülerdir. Fakat bu görüntülerde kişiler ellerinde farklı nesneler taşıyabilirler ve bu nesneler kişinin özellik çıkarımı yaptığımız en çok hareket eden vücut parçaları olan kol ve bacaklarının bir kısmını kapatabilir ve bu da sistem başarısını değiştirebilir. Şekil 3. 6' da böyle bir duruma ait ekran görüntüsü görülmektedir. Görüntüde elinde büyük bir nesne taşıyan bir kişi görülmektedir. Bu kişinin 2 farklı yürüme şekli görülmektedir. İlkinde daha büyük adımlarla yürümesine rağmen ikinci görüntüden sonraki görüntülerde adımlarını küçülterek yürümüştür. Bu değişim sistem tarafından başarılı bir şekilde tespit edilmiştir.



Şekil 3. 6 Elinde nesne ile yürüyen kişinin yürüme şeklini değiştirme başlangıçları

Tek kişinin içinde olduğu olası durumlarda başarılı olmasıyla sistemin farklı durumlarda nasıl davrandığı test edilmiştir. Bu durumlardan güvenlik kameraları için en fazla karşılaşılabilecek olanı şüpheli paket bırakılması ve bu paketin daha sonra başka bir kişi tarafından alınarak götürülmesidir. Bu duruma ait görüntüler Şekil 3. 7, Şekil 3. 8 ve Şekil 3. 9'de görülmektedir.



Şekil 3. 7 Şüpheli kutu bırakan kişinin görüntüsü

Şekil 3. 7’de elinde kutu ile gelen bir kişinin kutuyu getirirken ve bırakıp giderken tespit sistem tarafından tespit edilen 2 görüntüsü görülmektedir. Burada sistem kutunun bırakılma anını kaçırmıştır. Bunun nedeni de bırakma anında tekrar eden bir hareket bulunmamasıdır. Fakat bulunan hareketlerin nerede bittiği bilindiği için eğer arada bulunmayan tek bir hareket varsa onun başlangıç noktası bulunabilir. Şekil 3. 8’de sistemin hareketin bittiği yerden sonraki yeni hareketin başlangıcını işaretlediği yer görülmektedir.



Şekil 3. 8 Şüpheli kutu bırakma anı

Kutuyu bırakan kişiden sonra kutuyu alıp giden kişiye ait yürüyerek gelme, kutuyu alma ve kutuyla beraber yürüme görüntüleri Şekil 3. 9’da görülmektedir. Kutu bırakma anının bulunamayıp kutunun alınma anının ek bir kontrole gerek kalmadan bulunmasının nedeni kutu alınırken yavaş hareket edilmesidir. Bu nedenle sistem kutu alma hareketini tekrar eden bir hareket olarak işaretlemiştir.



Şekil 3. 9 Şüpheli kutunun alınmasında hareket değişimleri

Sistemin bu verisetindeki farklı durumlar için çalışmasının başarı oranı Çizelge 3. 2’de görülmektedir.

Çizelge 3. 3 Sistemin dik kamera görüntüsündeki başarıları

Video No.	Çerçeve Sayısı	Bulunan Kalıp / Kalıp Sayısı	Bulunan Hareket / Hareket Sayısı	Özet Video Çerçeve Sayısı
1	719	14 / 14	6 / 6	76
2	510	7 / 10	5 / 6	94
3	360	7 / 7	2 / 2	40
4	210	4 / 4	2 / 2	47
5	600	9 / 9	4 / 4	66
6	779	12 / 13	5 / 6	91
7	1348	13 / 13	7 / 7	122
8	4526	53 / 55	26 / 28	360
Toplam:	9052	119 / 125 (%95,2)	57 / 61 (%93,44)	896

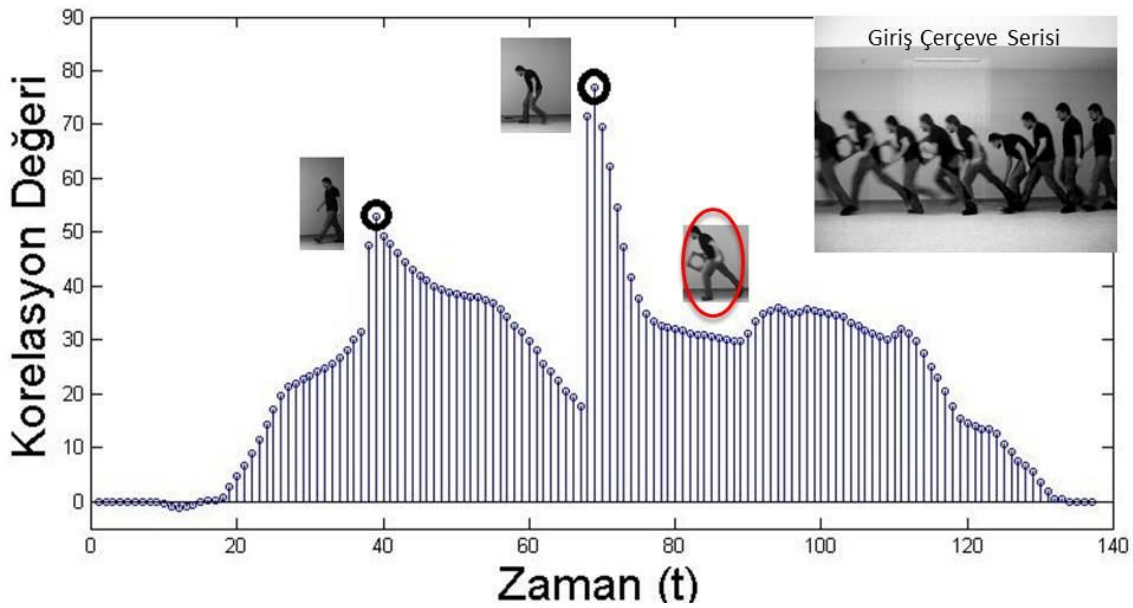
Çizelgede gösterilen videoların içerikleri;

- Koşma yürüme geçişleri
- Yere nesne bırakıp, bir süre sonra alıp koşarak uzaklaşma
- Yere nesne bırakıp hemen ardında alıp yürüme
- Elinde büyük bir nesne ile farklı şekillerde yürüme

- Yere nesne bırakıp, bir süre sonra alıp yürüyerek uzaklaşma
- Yere bir nesne bırakıp nesnenin üzerinde oturup bekleme
- Yürüyerek gelip, bir noktada sabit bekleme
- Yukarıdaki tüm hareketlerin farklı zaman aralıklarında farklı kombinasyonlar ile yapılması

gibi çeşitli olası durum senaryolarına sahiptir. Çizelgedeki çerçeve sayısı her videonun kaç çerçeveden oluştuğunu göstermektedir. Çizelgedeki dördüncü sütun videoda sistemin tespit ettiği hareket değişimlerinin, işaretlenmiş hareket değişimlerine oranını göstermektedir. Üçüncü sütun, dördüncü sütunun devamı niteliğinde olup tekrarlayan hareket kalıplarının toplamda kaç kere tespit edildiğinin, videoda tespit edilmesi gereken kalıp sayısına oranlanmasıyla oluşturulmuştur. Çizelgenin sağındaki sütun ise sistemin oluşturduğu özet videonun çerçeve sayısını göstermektedir. Özet video oluştururken, hareket değişimleriyle beraber bulduğu hareket kalıbı uzunluğundan yararlanmaktadır.

Sistemin bir koşma hareketi olmasına rağmen bulamadığı bazı örnekler de vardır. Şekil 3. 10'da başarılı ve başarısız olunan tekrarlara ait korelasyon değerleri bulunmaktadır.



Şekil 3. 10 Bir videoya ait korelasyon değerlerine ait sonuçlar

Yukarıdaki şekilde yürüyerek gelme ve yerdeki cismi alma anındaki başarılı olarak tespit edilen ilk iki görüntünün korelasyon değerleri siyah çember içine alınmıştır. Bunların dışında bulunamayan bir hareket başlangıcı olan yerden cismi aldıktan sonra koşma hareketi sistem tarafından tespit edilememiştir. Tespit edilen hareket geçişlerinin korelasyon değerlerine bakıldığında kendinden önceki ve sonraki korelasyon değerinden bir hayli büyük olduğu görülmektedir. Bulunamayan hareket geçişine bakıldığında net olarak yerel bir maksimum değer görülememektedir. Giriş çerçeve serisi incelendiğinde yerden cisim alındıktan sonra her ne kadar yapılan hareket bir koşma hareketi olsa da bacakların ve vücudun üst kısmının duruş şekli bşr hareket tekrarı yakalamanın mümkün olmayacağını göstermektedir. Farklı şekillerde atılan adımlar sistemin hareket tekrarını yakalamasını engellemiştir.

3.3 Tek Kişi Tepeden Çekilen Videolarla Oluşturulmuş Veriseti

Sistem kişiyle aynı seyiveden ve yaklaşık olarak 90 dereceden alınan görüntülerde gayet başarılı sonuçlar vermiştir. Bu sonuçlar başarıyı değerlendirmek için yeterli değildir çünkü bir güvenlik kamerası sistemindeki güvenlik kamerası kişileri daha yukarıdan görecektir şekilde, kişinin ulaşamayacağı şekilde yerleştirilir. Bu nedenden dolayı 5 kişiye ait görüntülerden oluşan bir veriseti oluşturulmuştur. Bu veri setinde de daha önce oluşturulan veri setinde oluşturulan hareket geçişleri gerçekleştirilmiştir. Ayrıca bu verisetinde kişiler kameradan uzaklaşıp yaklaşmaktadır yani kişinin boyutu büyüyüp küçülmektedir. HOG'un ölçekten bağımsız olarak çalışmasıyla kişinin kameraya yaklaşip uzaklaşması sistemin başarısını etkilememiştir. Şekil 3. 11'de kameranın kişiyi hangi açıdan gördüğünün bir örneği görülmektedir.



Şekil 3. 11 Veri setinin oluşturulduğu kamera görüş açısı

Sistem yukarıdaki şekildeki gibi bir kamera açısıyla test edildiğinde kişi yönünü değiştirdiğinde bunu farklı bir hareket olarak tespit etmektedir. Bunun nedeni de gradyanların histogramı yönteminde kişi yön değiştirdiğinde çıkarılan kenar özellikleri değiştiği için farklı hareket algısı oluşmaktadır. Şekil 3. 12’de Çerçevenin alt kısmından görüntüye giren bir kişinin önce sol ileriye doğru çapraz, daha sonra ileriye doğru düz ve son olarak da sağ ileriye doğru çapraz yürüyüşüne ait tespit edilen geçişler görülmektedir.



Şekil 3. 12 Yürüyüş görüntüsü açısı değişiminin tespit edilmesi

Sistem daha önceki veri setinde de test edildiği gibi kişinin yürümeden koşmaya geçmesi gibi durumlarda da kişinin hem yukarıdan hem de açılı bir şekilde görüntüsünün olmasına rağmen başarılı sonuçlar vermiştir. Bir kişiye ait örnek yürüme kalıp başlangıçları Şekil 3. 13’te, yürümeden koşmaya geçiş ve koşma hareketinin tekrarlarının başlangıç görüntüleri Şekil 3. 14’te görülmektedir.



Şekil 3. 13 Bir kişiye ait yürüme kalıpları başlangıç görüntüleri



Şekil 3. 14 Bir kişiye ait koşmaya geçiş başlangıcı ve koşma kalıpları başlangıç görüntüleri

Yukarıdaki görüntülerde görüldüğü üzere yürümenin son kalıp başlangıcı ile koşma hareketine başlangıç görüntüsü arasında bir boşluk görülmektedir. Bu görüntüler hareketlerin başlangıç görüntüleri olmasından dolayı arada kalıp uzunluğu kadar bir fark olması normal bir durumdur. Fakat burada kişi koşmadan yürümeye geçişte ne yürüme hareketine dâhil ne de koşma hareketine dâhil olabilecek kısa süreli bir hareket geçişi gerçekleştirmiştir. Hareket geçişlerinde oluşan kalıp uzunluğundan daha uzun olan boşlukların nedeni de bu kısa süreli geçiş hareketlerinin varlığıdır. Sistemin yukarıdan çekilen görüntülerle oluşturulan verisetindeki başarı oranları Çizelge 3. 3'te görülmektedir.

Çizelge 3. 4 Sistemin verisetindeki başarısı

Video No.	Çerçeve Sayısı	Bulunan Kalıp / Kalıp Sayısı	Bulunan Hareket / Hareket Sayısı	Özet Video Çerçeve Sayısı
1	450	9 / 9	4 / 4	107
2	1065	15 / 16	6 / 6	145
3	660	12 / 13	5 / 5	124
4	1317	7 / 16	3 / 5	489
5	457	13 / 13	3 / 3	67
6	1260	13 / 17	4 / 5	149
Toplam:	5092	69 / 84 (%82,14)	25 / 28 (%89,29)	896

Sistem genel manada başarılı olsa da başarısız olduğu görüntüler de vardır. Bunları inceleyecek olursak öncelikle dört numaralı videoya bakacak olursak, burada iki adet problem göze çarpmaktadır. Bunlardan ilki kişinin kamera görüntüsünden çok fazla uzaklaştığı için görüş alanı sınırlarını zorladığı için kişi her çerçevede kutu içine alınamamıştır. Dolayısıyla düzgün bir özellik çıkarımı gerçekleştirilememiştir. Bulunamayan harekete ait iki görüntü Şekil 3. 15’de görülmektedir.



Şekil 3. 15 Özellik çıkarımı düzgün yapılamayan görüntü örnekleri

Şekil 3. 15’de de görüldüğü gibi kişi hem kameraya uzaklığından hem de koyu tonlardaki elbiselerinden dolayı arka plan ile neredeyse aynı tonlardadır. Dolayısıyla bu

örneklerde arka arkaya gelen tüm çerçevelerde istenilen şekilde kişi kutu içine alınamamakta, bu nedenle de özellik çıkarımında arka arkaya gelen çerçevelerin özelliklerinden tekrar arayak hareket geçişi tespiti yapılamamaktadır.

Sistemin tespit edemediği görüntüler arasında sıfır derece ile kameradan koşarak uzaklaşan veya yaklaşan kişilere ait görüntülerde istikrarlı bir şekilde özetleme yapılamamaktadır. Kişi eğer yürüyor ise her zaman olmakmak kaydıyla sistem özetleme işlemini yapabilmektedir. Sistemin istikrarlı olarak çalışabileceği minimum açı değeri 15 derecedir. Yani kişi minimum 15 derece ile kameraya yaklaşmalı veya uzaklaşmalıdır. Şekil 3. 16'da başarısız olunan kameradan koşarak uzaklaşan bir kişiye ait görüntüler görülmektedir.

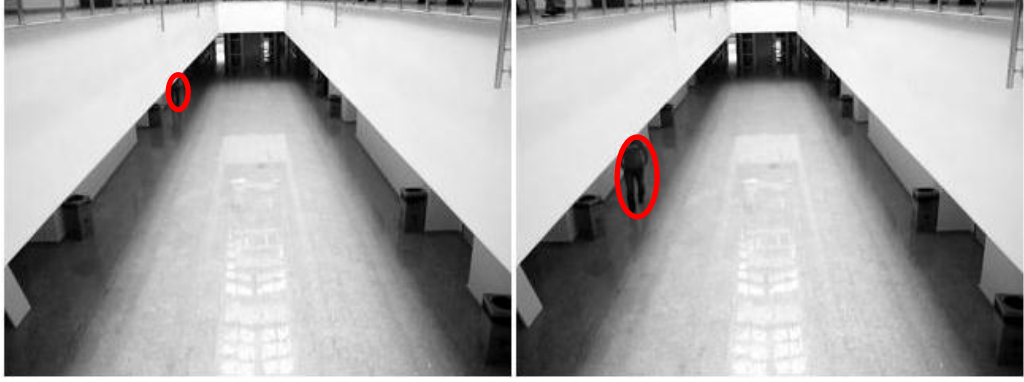


Şekil 3. 16 Hareket başlangıçları tespit edilemeyen görüntüler

Sistemin yukarıdaki örnekte doğru çalışmamasının sebebi gradyanların histogramı çıkarılırken kullanılan kenar özelliklerinde yeterince değişim olmamasıdır. Ayrıca hareket hızlı bir şekilde gerçekleşmesi de diğer bir etkindir fakat yavaş gerçekleşmesi durumunda da daha önce bahsedildiği gibi istikrarlı bir şekilde çalışmamaktadır. Bu problemin çözümü aynı alanı göreceğ biçimde, birbirine dik yani 90 derece olacak şekilde yerleştirilen iki kameradan alınan görüntülerin işlenip, iki kamera görüntüsünün işlenmesiyle alınan sonuçların harmanlanmasıyla daha başarılı sonuçlar elde edilecektir.

Sistem vücudun belli bir bölümü görüntüde olduğunda da başarılı olarak çalışmaktadır. Fakat görüntüde bir kişinin önce sadece bacakları daha sonra vücudunun tamamı görülüyorsa, bu yeni bir hareket olarak algılanmaktadır. Çünkü ilk durumda kişinin sadece bacaklarından özellik çıkarımı yapılırken, ikinci durumda tüm vücudunun özellik

çıkartımı yapılır ve bu iki durumda oluşturulan histogram vektörleri temsil ettikleri bölgeler bakımında farklı yapılardadır. Şekil 3. 17’ de vücudunun önce bir kısmı, sonra tamamı görünen bir kişiye ait hareket değişimi olarak tespit edilen görüntüler bulunmaktadır.

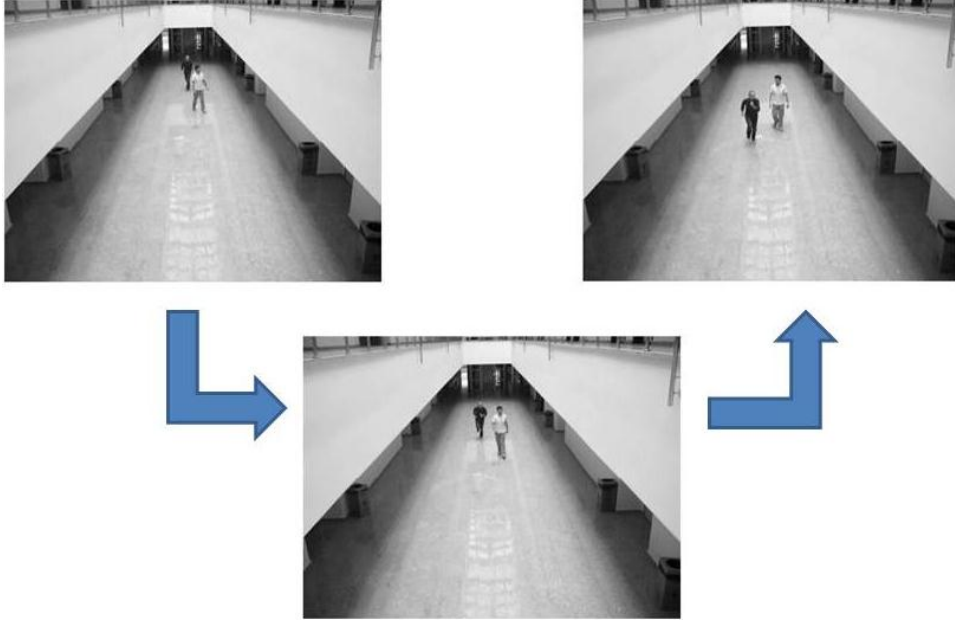


Şekil 3. 17 Farklı hareket olarak algılanan hareket başlangıç görüntüleri

Yukarıdaki görüntülerde bir kişiye ait yürüme hareketi görülmektedir. Bu görüntülerde kişinin vücudunun ilk kısmı sol taraftaki görüntüde ve devam eden görüntülerde görülmemektedir. Fakat daha sonraki görüntülerde kişinin vücudunun neredeyse tamamı görülmektedir. Kişi yürüme hareketine devam ettiği halde sistem farklı bir harekete geçiş olarak algılamıştır. Bunun nedeni de kişi ilk önce kişinin sadece vücudunun belli bir bölümünden özellik çıkarımı yapılmasına rağmen, ilerleyen görüntülerde tüm vücutundan özellik çıkarımı yapılmıştır. Bu özelliklerin değerlendirilmesiyle de devam eden bir hareket olan yürüme farklı bir hareket olarak tespit edilmiştir.

3.4 Birden Fazla Kişinin Olduğu Videolarla Oluşturulmuş Veriseti

Güvenlik kameralarından elde edilecek videolarda tek kişilik görüntülerden daha çok birden fazla kişinin görüntüye dâhil olacaktır. Bu nedenle birden fazla kişinin bulunduğu videolar üzerinde sistem test edilmiştir. Oluşturulan çoklu insan veriseti iki farklı ortamda farklı kamera açılarıyla çekilen videolardan oluşmaktadır. Test işlemlerindeki ilk 3 video kapalı ortamda, diğer 3 video da açık alanda çekilen görüntülerden oluşmaktadır. Test edilen kapalı ortam videosunda sistem tarafından tespit edilen hareket değişimlerine ait örnek çerçeveler Şekil 3. 18’de görülmektedir.

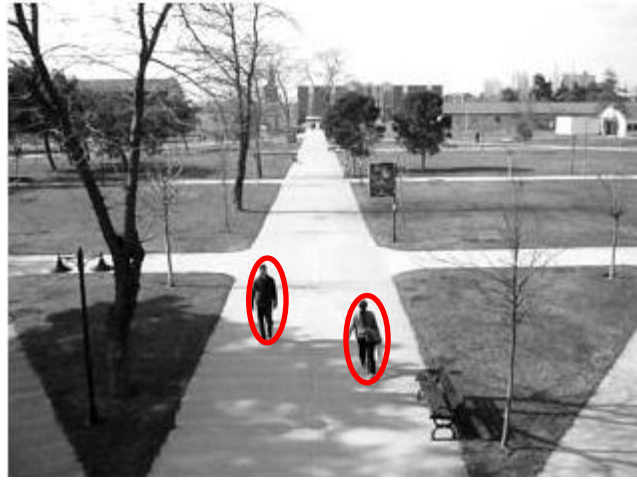


Şekil 3. 18 Çoklu insan veriseti hareket değişim tespitleri

Tek kişinin bulunduğu videolardaki gibi bazı test senaryoları çoklu insan ortamları için de gerçekleştirilmiştir. Şekil 3. 18’de yürümeden koşmaya geçişe ait hareket videosuna örnek çerçeveler görülmektedir. Şekildeki ilk çerçeve görüntüdeki beyazlı kişinin hareketine göre özetlenirken, daha sonra gelen 2 çerçeve siyahlı kişinin koşu hareketine bağlı hareket değişimi olarak tespit edilmiştir.



Şekil 3. 19 Çoklu insan veriseti dış ortam görüntüsü



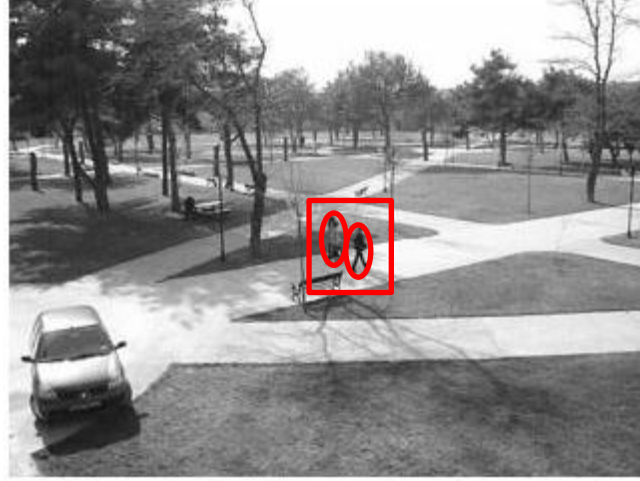
Şekil 3. 20 Çoklu insan veriseti dış ortam hareket değişimi örneği

Dış ortamda kaydedilen video görüntülerine ait örnek geçiş çerçeveleri Şekil 3. 19' da ve Şekil 3.20' de görülmektedir. Açık alan videoları belirle senaryolara göre oluşturulmamıştır. Görüntülerdeki kişilerin yapılan kayıtlardan haberi olmadığı için belirli hareket geçişleri yoktur. Sadece kişilerin yürüme esnasında yürümelerindeki değişimler tespit edilmiştir. Dış ortam görüntüleri 2 farklı kamera açısı ile kaydedilmiştir. Çoklu insan verisetine ait başarı oranları Çizelge 3. 4' te görülmektedir.

Çizelge 3. 5 Sistemin verisetindeki başarısı

Video No.	Çerçeve Sayısı	Bulunan Kalıp / Kalıp Sayısı	Bulunan Hareket / Hareket Sayısı	Özet Video Çerçeve Sayısı
1	1019	17 / 21	6 / 7	280
2	1799	59 / 73	16 / 21	396
3	2101	46 / 71	11 / 18	597
4	994	18 / 29	3 / 5	126
5	894	46 / 46	7 / 7	117
6	1170	38 / 44	7 / 8	157
Toplam:	7977	224 / 284 (%78,87)	50 / 66 (%75,76)	1673

Sistemin başarısının çoklu insan videolarında tek kişili videolara göre düşmesinin nedeni insan takibini düzgün bir şekilde yapılamamasıdır. Yani görüntüdeki insanların sistem tarafından zaman zaman tek bir insanmış gibi algılanmasıdır. İnsanların hareketleri sırasında çakışma nedeniyle takip işleminin yeteri kadar doğru yapılamamasından dolayı başarı düşmüştür. Şekil 3. 21’de bu yanlışlığa neden olan örnek bir durum görülmektedir.



Şekil 3. 21 Çoklu insan veriseti çakışma örneği

Sistem daha başarılı bir insan takip sistemi ile entegre bir şekilde çalıştırıldığında çoklu insan verisetleri için tekli insan verisetlerinde olduğu kadar başarılı sonuçlar vermesi muhtemeldir.

SONUÇ VE ÖNERİLER

Video özetleme uygulamalarında karar verilmesi gereken en önemli etken özetlemenin temel alacağı özelliktir. Günümüzde video kaydı yapabilecek multimedya araç sayısının artması ve bu kayıtlarda temel öğenin genel olarak insan olması, videolardan bilgi çıkarımında temel etmenin insan olması yönünde bir yönelim olmasına sebep olmuştur. Bu nedenle video özetleme veya videoda önemli an çıkarımı yapılırken insan üzerinden özellik çıkarımı yapılmamıştır. Çünkü bir kayıtta garip bir durum varsa bu durum kendi kendine oluşamaz, mutlaka olayın içinde bir insan etkisi vardır.

Bilgi çıkarımı yapılmadan önce görüntüde nereden bilgi çıkarımı yapılacağı seçilmelidir. İnsan tabanlı bir çıkarım yapılacaksa, öncelikle insanın bulunduğu görüntü bölgesi işaretlenmelidir ve görüntünün tamamından ayrılarak işaretlenen kısım üzerinde işlem yapılmalıdır. İşaretlenen bölge üzerinde işlem yapmak, işlem yükünü azaltacağı gibi özellik çıkarılırken gereksiz bilgi çıkarımını engelleyerek sadece anlamlı bilgilerin çıkarımını sağlar.

İşlem yapılacak bölge çıkarıldıktan sonra bu bölgeden özellik çıkarımı yapılmalıdır. Çıkarılacak özelliği işaretlenen bölgedeki diğer parçalara oranla daha belirgin hale getirmek, özelliğin baskın olmasında önemli bir rol oynayacaktır. Daha sonra çıkarılan özelliklerin işlenip, analiz edilmesiyle bir görüntü üzerinden bilgi çıkarılarak o görüntünün işleme işlemi tamamlanacaktır.

Bu tez çalışması kapsamında video özetleme alanı için insan tabanlı bir özetleme ve anormal durum tespit edici sistem geliştirilmiştir. Bunu gerçekleştirme üzere önce ön plan çıkarımı yapılmıştır ve çıkarılan ön planda insanın bulunduğu bölge kutu içine alınarak işaretlenmiştir. Daha sonra bu işaretli bölgenin hareket özellikleri çıkarılmıştır

ki, bu da hareket eden kısımların belirleyici özellik olmasını sağlamıştır. Hareket özellikleri çıkarılan bölgenin gradyan özellikleri kullanılarak her çerçeve için özellik çıkarımı gerçekleştirilmiştir. Çıkarılan özellik sayesinde her çerçeve bir özellik vektörü şeklinde ifade edilebilir hale getirilmiştir. Özellik vektörleri arka arkaya sıralanarak zaman bilgisi eklenmiştir. Zaman bilgisi eklenen çerçeve özellikleri, kalıp benzerliği yöntemiyle analiz edilerek tekrar eden hareketler ve bu hareketlerin geçişleri elde edilmiştir. Tekrar etmeyen hareketler veya ani değişimler de olağandışı durum olarak işaretlenmiştir.

Sistemin başarısı insan tabanlı özet çıkarımı yapan başka bir çalışmada gerçekleştirildiği gibi Weizmann İnsan Hareketi Veriseti kullanılarak sistem testleri gerçekleştirilmiştir. Bu verisetindeki 6 kişiye ait 9 farklı hareket üzerinde yapılan testlerde tek bir hareket dışında diğer bütün hareket geçişleri ve hareket kalıbı tekrarları başarılı bir şekilde tespit edilmiştir. Bulunamayan harekette yeteri kadar gradyan değişimi olmadığı için sistem başarısız olmuştur. Sistem hareket geçişlerini %89.47 gibi bir başarı oranıyla bulmuştur. Bu veriseti insan hareketi tabanlı bir özetleme için özel üretilmemiş olup, insan hareketi tanıma için oluşturulmuş bir verisetidir. Bu nedenle giriş videosu oluşturulurken ayrı ayrı kaydedilmiş hareketler arka arkaya birleştirilerek, bütün bir giriş videosu oluşturulmuştur. Bu da hareket geçişleri arasında keskin farklılıklar oluşturmuştur. Bu nedenle sistemin test edilmesi için yeni bir veriseti oluşturulmuştur.

Oluşturulan ilk verisetinde görüntüler kişinin hareketini 90 derecelik bir açı ile görecektir. Bu şekilde kayıt yapan bir kamera ile kaydedilmiştir. Veriseti 3 kişiyle güvenlik kamerası görüntülerinde oluşabilecek çeşitli senaryoların gerçekleşmesiyle oluşturulmuştur. Sistemin bu verisetinde hareket geçişi bulma başarısı %93.44 olarak tespit edilmiştir. Sistemin geçiş tespitinde başarısız olduğu hareketlerdeki tekrarlar net bir şekilde belli olmadığından sistem hareketi algılamayı başaramamıştır.

Sistemin başarısı hem genel bir verisetinde hem de oluşturulan verisetinde test edildikten sonra gerçek bir güvenlik kamerası açısına sahip görüntü kayıtlarıyla 5 kişiye ait görüntülerden oluşan yeni bir veriseti üzerinde test edilmiştir. Bu verisetinde yapılan denemelerde kişinin kameraya yaklaşması yani boyutunun değişmesi sistemin başarısını etkilememektedir. Bu veriseti üzerinde yapılan test çalışmalarında da %89.29

gibi bir başarı oranı yakalanmıştır. Kişinin vücudunun tamamının görünmediği görüntülerde de sistem başarılı bir şekilde çalışmaktadır. Kameraya direk yaklaşan veya uzaklaşan kişiler için stabil sonuçlar alınamamakta, bazen başarılı bazen başarısız olunmaktadır. Başarılı olunan örneklerde kişinin hareketleri yavaştır. Hızlı bir şekilde uzaklaşan veya yaklaşan örneklerde başarısız olunmuştur. Bunun çözümü aynı bölgeyi birbirine dik olacak şekilde gören senkron iki kamerayla kayıt yapmaktır. Özetlemenin de bu iki kamera kaydından alınan sonuçların birleşimiyle yapılması sistemin başarısını artıracaktır.

Yapılan test sonuçlarından anlaşılaçağı üzere insan tabanlı özetleme yöntemi yeni bir video özetleme yöntemi olarak kullanılabilir. Bunun yanında olağandışı durumları tespit için de bir yaklaşım oluşturabilir. İnsan odaklı bir yaklaşımın diğer özetleme yaklaşımlarından hem güvenlik sistemleri açısından hem video etiketleme açısından daha faydalı bir yöntem olduğı açık bir şekilde göz önündedir.

İnsan tabanlı video özetleme yöntemini daha da geliştirmek üzere yapılabilecek gelecek çalışmalar, yöntemin sadece insanlar üstünde değil diğer canlılar üzerinde, askeri alanların güvenliğı ve spor karşılaşması özetleyecek şekilde uyarlamaktır. Sistemin başarısını artırmak için birden fazla kamera görüntüsünün işlenmesiyle yapılacak özetleme işlemi sistemin başarısını artıracacağı gibi daha güvenilir sonuçlar alınmasını da sağlayacaktır. Ayrıca ön plan çıkarımında yapılacak iyileştirmelerle, işlenmek üzere işaretlenen bölgenin daha az hata ile gerçekleştirilmesi, insan tabanlı özetleme yöntemini geliştirmek üzere gerçekleştirilebilecek bir çalışmadır.

KAYNAKLAR

- [1] Dimitrova, N., (2002). Applications of video-content analysis and retrieval, 9:42- 55.
- [2] Xiong, Y., Radhakrishnan, Z., Divakaran, R., Huang, A. ve Rui, T.S., (2004). A Unified Framework for Video Summarization, Browsing and Retrieval.
- [3] Merialdo, B., Huet, B. ve Yahiaoui, I., (2003). Automatic Video Summarization.
- [4] Zhang, L., Cao, Y., Ding, G. ve Wang, Y., (2008). "A Computable Visual Attention Model for Video Skimming", 10th IEEE International Symposium on Multimedia, 667-672.
- [5] Gensel, P., Martin, J. ve Mulhem, H., (2003). Adaptive Video Summarization, Singapore.
- [6] Ulges, A., Schulze, C., Breuel, T.M. ve Borth, D., (2008). Keyframe Extraction for Video Tagging & Summarization, Proc. Informatiktage, 45-48.
- [7] Wu, Y., Yu, S., Chen, T. ve Li, C., (2009). "Motion-focusing key frame extraction and video summarization for lane surveillance system", ICIP, 16th IEEE International Conference.
- [8] Cernekova, Z., Pitas, I. ve Nikou, C., (2006). "Information theory-based shot cut/fade detection and video summarization", Circuits and Systems for Video Technology, IEEE Transactions, 16:82-91.
- [9] Chang, J., Hu, R.M. ve Zhang, H., (2009). "Summarization Extraction Method for Surveillance Video Based on Color Spatial Distribution Characteristic", in Computational Intelligence and Software Engineering.
- [10] Zhong, J., Yuting, S., Rongrong, Q. ve Jintao, M., (2010). "Surveillance video summarization based on moving object detection and trajectory extraction", International Conference Signal Processing Systems (ICSPS) , 2:250-253.

- [11] Zhang, W., Lin, J., Chen, X., Huang, Q. ve Liu, Y., (2006). "Video Shot Detection Using Hidden Markov Models with Complementary Features", *Innovative Computing, Information and Control*, 3:593-596.
- [12] Ma, Y., Zhang, H. ve Ngo, C., (2005). "Video summarization and scene detection by graph modeling", *Circuits and Systems for Video Technology, IEEE Transactions*, 15:296-305.
- [13] Zhang, X. ve Zhou, J., (2005). "Video Shot Boundary Detection Using Independent Component Analysis", *Acoustics Speech and Signal Processing(ICASSP)*, 2:541- 544.
- [14] Kasturi, R., Strayer, S.H. ve Gargi, U., (2000). "Performance characterization of video-shot-change detection methods", *Circuits and Systems for Video Technology, IEEE Transactions*, 10:1-13.
- [15] Davis, J.W. ve Bobick, A.F., (2001). "The recognition of human movement using temporal templates", *Pattern Analysis and Machine Intelligence, IEEE Transactions*, 23:257-267.
- [16] Guo, K., Ishwar, P. ve Konrad, J., (2010). "Action change detection in video by covariance matching of silhouette tunnels", *Acoustics Speech and Signal Processing (ICASSP)*, 1110-1113.
- [17] Chen, D.Y., Chu, K.C., Liu, Y.C. ve Chen, Y.S., (2010). "Human subject-based video browsing and summarization", *Machine Learning and Cybernetics (ICMLC)*, 6:2796-2801.
- [18] Grimson, C. ve Stauffer, W.E.L., (1999). "Adaptive Background Mixture Models for Real-Time Tracking", *Computer Vision and Pattern Recognition*, 2:2.
- [19] Mladenovic, V.M., Porrat, D. ve Lutovac, M.D., (2011). "The direct execution of the expectation-maximization algorithm using symbolic processing", *Telecommunication in Modern Satellite Cable and Broadcasting Services (TELSIKS)*, 1:265-268.
- [20] Lindeberg, T., (1993). "Detecting Salient Blob-Like Image Structures and Their Scales with a Scale-Space Primal Sketch: A Method for Focus-of-Attention", *International Journal of Computer Vision* 11, 3:283–318.
- [21] Lu, S., Tsechpenakis, G., Metaxas, D.N., Jensen, M.L. ve Kruse, J., (2005). "Blob Analysis of the Head and Hands: A Method for Deception Detection", in *Proceedings of the 38th Annual Hawaii International Conference*.

- [22] Barron, J.L., Fleet, D.J., Beauchemin, S.S. ve Burkitt, T.A., (1992). "Performance of optical flow techniques", *Computer Vision and Pattern Recognition*, 236-242.
- [23] Dalal, N. ve Triggs, B., (2005). "Histograms of oriented gradients for human detection", *Computer Vision and Pattern Recognition*, 1:886-893.
- [24] Zanker, J.M. ve Fisher, N., (2001). "The directional tuning of the barber-pole illusion", *Perception*, 30:1321 – 1336.
- [25] Al Kanawathi, J., Mokri, S.S., Ibrahim, N., Hussain, A. ve Mustafa, M.M., (2009). "Motion detection using Horn Schunck algorithm and implementation", *Electrical Engineering and Informatics International Conference(ICEEI)*, 1:83-87.
- [26] Gdalyahu, Y., Hayon, G. ve Shashua, A., (2004). "Pedestrian detection for driving assistance systems: Single-frame classification and system level performance", in *In Proceedings of IEEE Intelligent Vehicles Symposium*.
- [27] Blank, M., Shechtman, E., Irani, M., Basri, R. ve Gorelick, L., (2007). "Actions as space-time shapes", *IEEE Transactions Pattern Analysis and Machine Intelligence*, 29:2247–2253.
- [28] Wang, C.Y., (1985). *Edge Detection Using Template Matching*.
- [29] Papageorgiou, C. ve Poggio, T., (2000). "A trainable system for object detection", *International Journal of Computer Vision(IJCV)*, 38:15-33.
- [30] Fergus, R., Perona, P. ve Fei, L., (2004). "Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories", *Computer Vision and Pattern Recognition(CVPR)*.
- [31] Kalman, R.E., (1960). "A New Approach to Linear Filtering and Prediction Problems", *Transaction of the ASME-Journal of Basic Engineering*, 35-45.
- [32] Bishop, G. ve Welch, G., (2006) "An Introduction to the Kalman Filter", *University of North Carolina, Technical Report TR95-041*.
- [33] Chen, S.L., Chung, L.S., Chen, Y.C. ve Chang, S.W., (2004). "Automatic License Plate Recognition", *IEEE Transactions On Intelligent Transportation Systems*, 5:42-53.
- [34] Wintz, R.C. ve Gonzalez, P., (1987). *Digital Image Processing. 2nd. Ed.. USA: Addison-Wesley Publishing Company*, 503.

ÖZGEÇMİŞ

KİŞİSEL BİLGİLER

Adı Soyadı :M.Said AYDEMİR
Doğum Tarihi ve Yeri :30 Nisan 1987, Denizli
Yabancı Dili :İngilizce, Almanca
E-posta :said@ce.yildiz.edu.tr

ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Bilgisayar Mühendisliği	Yıldız Teknik Üniversitesi	2010
Lise	Fen Bilimleri	Denizli Anadolu Lisesi	2005

İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2011-2012	Yıldız Teknik Üniversitesi, Elektrik-Elektronik Fakültesi, Bilgisayar Mühendisliği Bölümü, Donanım Anabilim Dalı	Araştırma Görevlisi

2010	IAS Bilgi İşlem Danışmanlık San. Ve Tic. A.Ş.	AR-GE Mühendisi
2009	Pamukkale Üniversitesi Bilgi İşlem Daire Başkanlığı	Yazılım Stajyeri
2008	Garanti Bilişim Teknolojisi ve Ticaret T.A.Ş.	Yazılım Stajyeri
2007	KoçSistem Bilgi ve İletişim Hizmetleri A.Ş.	Donanım Stajyeri

YAYINLARI

Bildiri

- 1.(2012) Aydemir M.S., Ergul U., Guclu A., Karsligil M.E., “Video Summarization Using Simple Action Patterns”, International Conference on Pattern Recognition (ICPR), 2012.