

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

**GERÇEK ZAMANLI VIDEO GÖRÜNTÜLERİNDEN İNSAN HAREKETLERİNİN  
TANINMASI VE YORUMLANMASI**

**ADEM GÜÇLÜ**

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

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

**İSTANBUL, 2012**

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

**GERÇEK ZAMANLI VİDEO GÖRÜNTÜLERİNDEN İNSAN HAREKETLERİNİN**  
**TANINMASI VE YORUMLANMASI**

Adem GÜÇLÜ 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 KARSLIĞİL  
Yıldız Teknik Üniversitesi

**Jüri Üyeleri**

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

Prof. Dr. Nizamettin AYDIN  
Yıldız Teknik Üniversitesi

Yrd. Doç. Dr. Hakan ERDOĞAN  
Sabancı Üniversitesi

\_\_\_\_\_

\_\_\_\_\_

\_\_\_\_\_

## ÖNSÖZ

---

Günümüzde bilginin hızına yetişmek, bilgiye erişmekten daha zor hale gelmiştir. Üstün bir hızla artan video kayıtları beraberinde otomatik algılama ve anlamlandırma gereksinimini de getirmiştir.

Artan bilgi birikimi ve anlık düzeye inen kamera kayıtları ile birlikte videodan anlam çıkarımının otomatikleşmesi zorunlu bir hal almıştır. Video anlamlandırma çalışmalarının uygulama alanındaki artan genişlik ile beraber, bilginin kaynağını oluşturan en önemli varlığın insan olduğunun düşünülmesi, insan hareketlerinin tanınmasını günümüzde etkin bir gereksinim haline getirmiştir. Belirtilen ihtiyaçlardan esinlenerek bu çalışma ortaya konulmuş, hareket tanıma esasına dayalı video anlamlandırabilen basit ve yenilikçi bir yapı tasarlanmıştır.

Çalışmam boyunca bana yol gösteren tez danışmanım Yrd. Doç. Dr. M. Elif Karşılığil'e, her koşulda yanımda olan oda arkadaşım Erkan Uslu'ya ve çalışma arkadaşlarıma tüm destekleri ve yardımları için teşekkür ederim.

Hayatım boyunca beni her zaman destekleyen ve yanımda olan aileme teşekkürlerimi bir borç bilirim.

Haziran, 2012

Adem GÜÇLÜ

## İÇİNDEKİLER

	Sayfa
SİMGE LİSTESİ.....	vi
KISALTMA LİSTESİ.....	vii
ŞEKİL LİSTESİ.....	viii
ÇİZELGE LİSTESİ .....	ix
ÖZET .....	x
ABSTRACT.....	xii
<b>BÖLÜM 1</b>	
GİRİŞ.....	1
1.1 Literatür Özeti .....	3
1.2 Tezin Amacı .....	6
1.3 Hipotez .....	6
<b>BÖLÜM 2</b>	
HAREKET TANIMANIN ESASLARI .....	8
2.1 Değişken Koşullarda Hareket Tanıyan Bir Sistem .....	9
2.2 Video Görüntülerinin Alınması .....	10
2.3 İlgi Alanı Çıkarımı.....	10
2.4 Optik Akış Güç Şekilleri .....	12
2.5 Sıralanmış Yönlü Gradyanlar .....	16
2.5.1 Yönlü Gradyanlar Histogramı .....	17
2.5.2 Yönlü Gradyanlar Histogramlarının Sıralanması.....	20
2.6 Hareket Sınıfı Belirleme .....	21
2.6.1 k-En Yakın Komşuluk Sınıflandırıcısının Kullanımı .....	22
2.6.2 Sıralanmış Yönlü Gradyanların k-En Yakın Komşuluk Sınıflandırıcısı ile Sınıflandırılması.....	23
<b>BÖLÜM 3</b>	

HAREKET ÖRÜNTÜLERİNİN TANINABİLİRLİĞİ .....	25
3.1 Basit Hareketlerin Tanınması .....	25
3.1.1 Weizmann Hareket Tanıma Veri Seti.....	26
3.1.2 Sistem Testi - Weizmann Hareket Tanıma Veri Seti Üzerinde Bütün Video Girdisinin Deneysel Sonuçları .....	27
3.1.3 Sistem Testi - Weizmann Hareket Tanıma Veri Seti Üzerinde Kısaltılmış Video Girdilerinin Deneysel Sonuçları .....	28
3.1.4 Weizmann Hareket Tanıma Veri Seti Üzerindeki Sistem Testi Sonuçlarının İncelenmesi .....	30
3.2 Görüntü Açısının Tanıyıcı Sistem Üzerine Etkisi.....	32
3.3 İnsan Hareketlerini Tanımak İçin Ne Kadar Süre Gerekli?.....	33
3.4 Değişken Koşulların Altında Sistemin İncelenmesi .....	35
3.4.1 Hareketin Tam Olarak Görünmediği Durumların Değerlendirilmesi. 35	
3.4.2 Düzgün Yapılmayan Hareketler Testi .....	37
<b>BÖLÜM 4</b>	
SONUÇ VE ÖNERİLER .....	40
KAYNAKLAR .....	42
EK-A.....	45
“20. IEEE SİNYAL İŞLEME VE İLETİŞİM UYGULAMALARI” KURULTAYINDA NİSAN 2012 TARİHİNDE SUNULAN “ARDIŞIK GRADYANLAR HİSTOGRAMI İLE İNSAN HAREKETLERİNİN TANINMASI” İSİMLİ BİLDİRİNİN ÖZET SAYFASI .....	45
ÖZGEÇMİŞ .....	47

## SİMGE LİSTESİ

---

$I_x$	X yönündeki deęişim
$u$	X yönündeki hareket bilgisi
$I_y$	Y yönündeki deęişim
$v$	Y yönündeki hareket bilgisi
$I_t$	Çerçevesel arası (zaman düzlemindeki) deęişim
$E$	Hareket enerjisi

## KISALTMA LİSTESİ

---

SVM	Destek Vektör Makineleri
HMM	Saklı Markov Modelleri
HOG	Yönlü Gradyanlar Histogramı
SURF	Hızlandırılmış Gürbüz Öznitelikler
PCA	Temel Bileşen Analizi
HOF	Yönlü Akış Histogramı
SIFT	Ölçekten Bağımsız Öznitelik Dönüşümü
k-NN	k-En Yakın Komşu
STS	Uzam-Zaman Betimleyicileri Yaklaşımı
BoW	Görsel Kelime Çantası
SBoW	Uzam Kısıtlı Görsel Kelime Çantası

## ŞEKİL LİSTESİ

	Sayfa
Şekil 2. 1 Gerçek zamanlı video görüntülerinden insan hareketlerinin tanınması ve yorumlanması sistem blok diyagramı .....	9
Şekil 2. 2 Weizmann hareket tanıma veri seti.....	10
Şekil 2. 3 İlgili alan çıkarımı – blob analizi sonuçları.....	12
Şekil 2. 4 Optik akış güç şekli örnekleri (normalize edilmiş) .....	14
Şekil 2. 5 Optik akış güç şekillerinin elde edilmesi (yalnız üst vücut bölgesi hareketli) .....	15
Şekil 2. 6 Optik akış güç şekillerinin elde edilmesi (tüm vücut hareketli) .....	15
Şekil 2. 7 İlgili alan üzerinde optik akış güç şekillerinin incelenmesi .....	16
Şekil 2. 8 Blok ve hücre bölgelerinin gösterimi .....	17
Şekil 2. 9 Gradyan yönleri.....	19
Şekil 2. 10 İlgili alan üzerinde 3x3 pencere ile yönlü gradyanlar histogramı hesapları yapılan bölgelerin gösterimi .....	19
Şekil 2. 11 Ardışık sıralanmış yönlü gradyanlar histogramının görselleştirilmesi .....	21
Şekil 3. 1 Hareket tanıma sıklıkla kullanılan veri setlerinden biri: Weizmann hareket tanıma veri seti .....	26
Şekil 3. 2 Bütün video girdisi hatalı örneği - seksek hareketinin koşma olarak tanınması.....	30
Şekil 3. 3 Kısaltılmış video girdisi hatalı örnekler - zıplama hareketinin seksek olarak tanınması.....	31
Şekil 3. 4 Kısaltılmış video girdisi hatalı örnekler – yerinde (dik) zıplama hareketinin elleri açarak zıplama (jumping jack) olarak tanınması.....	31
Şekil 3. 5 Kısaltılmış video girdisi hatalı örnekler – yan yürüme hareketinin yürüme hareketi olarak tanınması .....	31
Şekil 3. 6 Kısaltılmış video girdisi hatalı örnekler – eğilme hareketinin tek el sallama hareketi olarak tanınması .....	32
Şekil 3. 7 Hareket dereceleri testi girdisi.....	32
Şekil 3. 8 Hareketin kameraya göre yapılış derecesinin hareket tanımaya etkisi.....	33
Şekil 3. 9 Hareket tanıma için ardışık çerçeve sayısına göre başarı değişimi.....	34
Şekil 3. 10 Hareketin görünümlüğünü engelleyen objeler testi girdileri.....	36
Şekil 3. 11 Düzgün yapılmayan hareketler testi girdileri.....	38



## ÇİZELGE LİSTESİ

---

	Sayfa
Çizelge 2. 1	İkili geçiş algoritması ..... 11
Çizelge 2. 2	Yönlü gradyanların sıralanması ..... 20
Çizelge 2. 3	Zamanda kayan pencere yardımı ile sınıflandırma ..... 23
Çizelge 3. 1	Weizmann hareket tanıma veri seti üzerinde bütün video girdisiyle hareket tanıma hata matrisi..... 27
Çizelge 3. 2	Weizmann hareket tanıma veri seti üzerinde günümüz modellerinin başarı sonuçları ..... 28
Çizelge 3. 3	Weizmann hareket tanıma veri seti üzerinde kısaltılmış video girdisiyle hareket tanıma hata matrisi..... 29
Çizelge 3. 4	Hareketin görünümlüğünü engelleyen obje testi sonuçları..... 37
Çizelge 3. 5	Düzgün yapılmayan hareketler testi sonuçları..... 38

**GERÇEK ZAMANLI VİDEO GÖRÜNTÜLERİNDEN İNSAN HAREKETLERİNİN  
TANINMASI VE YORUMLANMASI**

Adem GÜÇLÜ

Bilgisayar Mühendisliği Anabilim Dalı

Yüksek Lisans Tezi

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

İnsan hareketlerinin tanınması ve yorumlanması video anlamlandırma çalışmalarının önemli bir parçasını oluşturmaktadır. Bu tez çalışmasında, video görüntülerini değerlendirerek yapılan hareketin ne olduğuna karar veren özgün bir sistem geliştirilmiştir. Böylece, otomatik olarak hareket algılayabilen gürbüz bir sistem oluşturulmuş ve ortam anlamlandırmaya ön adım olabilecek bir yapı ortaya konulmuştur.

Önerilen sistemde hareket zamansal bir örüntü olarak tanımlanmış ve bu durumu ifade etmek için anlık durağan görüntü bilgisi ile hareket bilgisinin birleşiminden faydalanılmıştır. Bu işlemin gerçekleşmesinde, öncelikle ön plan çıkarımı ile bilgi içermeyen bölge elenmiştir. Daha sonra ön plan üzerinde insan konumu ve koordinatları blob analizi ile çıkarılmıştır. Optik akış hesabı ile bu bölge üzerindeki hareket bilgisi elde edilmiş ve beraberinde bu bilgi hareketin gücünü ifade eden bir hale dönüştürülmüştür. Elde edilen optik akış güç şekli yönlü gradyanlar histogramı yardımıyla betimleyici bir hale dönüştürülmüştür. Betimleyiciler zaman düzleminde ardışık dizilerek, hareket bilgisinin ve durgun görüntü bilgisinin yanı sıra zaman kavramını da içeren son betimleyiciler elde edilmiştir. Bu betimleyicilerin güçlülüğü ve güvenilirliği esas alınarak zaman düzleminde bir pencere çerçevesinde kayan basit bir sınıflandırıcı ile tanımlayıcı oluşturulmuştur.

Bu tez çalışmasında, insan hareketlerinin hareket, durgun görüntü ve zaman bilgisinin birleştirilmesi ile etkili biçimde ifade edilebileceği ortaya konulmuştur. Sonuç olarak ise değişken boyutta, değişen koşullar altında başarıyla çalışabilecek, otonom hareket tanıyabilen ve bu sayede video anlamlandırabilen bir yapı oluşturulmuştur.

**Anahtar Kelimeler:** İnsan hareketlerinin algılanması, hareket tanıma, ardışık gradyanlar histogramı

**REAL TIME HUMAN ACTION RECOGNITION AND UNDERSTANDING**

Adem GÜÇLÜ

Department of Computer Engineering

MSc. Thesis

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

Human action recognition and interpretation constitutes an important part of the video understanding. In this study, a novel action recognition system is developed which can detect and understand human actions. Thus, a robust system that can understand actions automatically is produced, so that, a structure is put forward as a first step to a fully developed video sensing framework.

The proposed system defines the action as a temporal pattern which is described by still image information combined with movement information. To realize that, firstly, foreground is extracted and the region containing no beneficial data is eliminated. Then, the location and coordinates of area of interest are obtained using blob analysis. Optical flow is used to determine the movement in the interest region and the output is used to construct the movement power shape. Descriptors are constituted by histogram of oriented gradient analysis of optical flow power shapes. Descriptors lined up in the time plane to form identifiers that contain the concept of motion, still image and time domain. Based on the power and reliability of these descriptors, a simple sliding window classifier is formed in the time domain.

In this study, it is seen that human actions can be represented efficiently by using descriptors of motion, still image and time. As a result, an autonomous action recognition system is produced which can detect human actions in a way that is steady

to work successfully under changing conditions, hence, a structure is put forward for video sensing.

**Key words:** Human action recognition, action detection, sequential gradient histograms

### GİRİŞ

İnsan hareketlerinin algılanması ve yorumlanması video anlamlandırma çalışmalarının önemli bir parçasını oluşturmaktadır. İnsan hareketlerinin tanınması, içerdiği çok yönlülük ile birlikte, bilgisayarla görü çalışmalarının gözde alanlarından biri haline gelmiştir. Bu durumun altında yatan en önemli sebep; video arşivleme, görsel gözetim, insan-bilgisayar etkileşimi, jest-mimik tanıma, video düzenleme ve video içerik arama gibi çeşitli uygulama alanlarındaki problemlerin henüz keşfedilmemiş yüksek potansiyele sahip çözümlerinin olduğunun öngörülmesidir.

Hareket algılama ifadesi dahilinde bulunan, aslen büyük değişkenlikler barındıran fakat ortak bir sınıf içerisinde toplanması gereken veri kümesinin varlığının ortaya konulması ile birlikte, insan hareketlerinin algılanması çalışma ilgisi uyandıran zorlu bir alan olarak tanımlanmıştır.

Günümüzde gittikçe artan video kayıtları ise beraberinde akıllı çözümler gereksinimini getirmiştir. İnsan hareketlerini algılayabilen bir sistemin, bilgi çıkarımı konusunda öncü ve daha karmaşık algılayıcılara yardımcı bir yapı oluşturabileceği öngörülmüştür.

İnsanın farkında bile olmadan doğal olarak gerçekleştirdiği hareketlerin algılanmasını ve tanımasını yapay zeka çalışmaları ile gerçeklemek hayal edilenden zordur. Hareket kendi içinde bir bütünlüğe ve periyodikliğe sahip olsa bile hareketler farklı insanlar tarafından farklı şekilde gerçekleştirilebilmektedir. Işık, kıyafet, ortam hareketleri gibi unsurlar da hareketin tanımlanabilmesine pozitif veya çoğunlukla negatif yönde etki edebilmektedir. Bu nedenlerle aynı sınıfa ait hareketler çok yüksek farklılıklar içerebilmektedir.

Hareket tanıma çevrimiçi veya çevrimdışı olarak tanımlanan çeşitli uygulamalarda kullanılabilir. Örneğin; daha önceden kayıt edilmiş olan videolarda futbol maçlarında top kapma anlarının bulunması, haberlerde el sıkışma hareketlerinin bulunması veya dans görüntüsünde istenilen bir hareketin bulunması gibi video aramada kolaylık sağlayan çevrimdışı uygulamalara konu olabilmektedir. Çevrimiçi çalışmalara örnek olarak ise otomatik gözetleme verilebilir. Örneğin; metro gibi toplu taşıma araçlarının izlenmesi, suçlu veya şüpheli tespiti veya bir diğer örnek olarak etraftaki nesnelere-denetleyicileri kontrol etme, basit yönetimleri gerçekleştirme gibi akıllı ev uygulamaları üzerindeki çalışmalar da gösterilebilir. Örneklerden de anlaşılacağı üzere hareket tanıma konusu çoktan günlük hayatımız içerisinde bulunan ve gelecek vaat eden bir çalışma alanıdır.

Hareket herhangi bir cismin durağan bir noktaya göre değişimidir. Hareket tanıma olarak nitelendirilen işlem sıralı bir biçimde, akış halinde devam eden görüntüde yapılan hareketleri değerlendirerek etiketlemektir.

Hareket kavramı derinlemesine incelenecek olursa, belli bir hiyerarşiye sahip olarak kabul edilebilir. Moeslund, çalışmasında hareketi primitif eylem, eylem ve aktivite olarak üç gruba ayırmıştır [1]. Bir uzuv hareketiyle ifade edilebilen eylemlere primitif eylem denilmiştir. Primitif eylemlerden oluşan, muhtemelen, döngüsel olarak ifade edilebilen hareketlere ise eylem denmiştir. Aktivitenin ise birkaç eylemi bir arada içeren anlamlı ve yorumlanabilir bir olay bütünü olduğunu ifade etmiştir. Örneğin; “sol bacak ileri” bir primitif eylem, “koşma” bir eylemdir. “Engelli koşu” ise zıplama ve koşma hareketlerini içeren bir aktivitedir. Primitif eylem, eylem, aktivite tanımlamaları sırasıyla düşük seviye, orta seviye ve yüksek seviye görü çalışmaları olarak da literatürde ifade edilmiştir.

İnsan hareketlerinin tanınmasında genel olarak videodan görüntü özelliklerinin çıkarılması ve buna karşılık gelen hareket sınıfı ile etiketlenmesi esası yaygındır. Böylece hareket tanıma iki ana adıma ayrılabilir: Özellik çıkarımı ve Sınıflandırma. Sınıflandırma başarı yüzdesine elbette etki etmesine rağmen, özellik çıkarımı adımı hem bu yüzdeyi ana şekilde etkileyen adım hem de sistemin bütününe nasıl çalıştığını betimleyen modeli oluşturan esas bölümdür. Böylece özellik çıkarımının sistemin ana

adımı olarak algılanması doğrudur. Bu adımda yaşanan zorluklar yahut kullanılan yöntemlerin karakteristikleri doğrudan başarıya etki etmektedir.

Sınıflar içi ve sınıflar arası varyasyonlar hareket tanıma konusunda başarıya etki eden en büyük etmenlerden biridir. Örneğin; yürüme hareketi insandan insana adım büyüklüğü, yürüme hızı, stil ve tip olarak büyük farklılıklar göstermektedir. Ortam ve çekim koşulları da hareket tanımlamaya doğrudan etki etmektedir. Kamera hareketi veya istenmeyen hareketlerin bulunduğu dinamik ortamlardaki çekimler hareket tanımayı zorlaştırmaktadır. İnsanların belirli kısımlarının veya komple insanın görüntü içinde kaybolması; kamera açısının veya ilgi alanı açısının değişmesi gibi durumlar hareket algılamayı zorlaştıran etmenlerdendir. Diğer bütün etmenler mükemmel hale getirilse dahi hareket tanımlamanın tanımının içine giren zamansal değişkenler boyutu olaya etki etmektedir. Zamanlama ve hareketin başlangıç-yapılış-bitiş gibi ortak olarak ifade edilmesi gereken durumlar insandan insana pek çok farklılık göstermektedir ve bu durum hareket tanımadaki genelleme yapma durumunu zorlaştırmaktadır.

### 1.1 Literatür Özeti

Bilgisayarla görme, bilgisayar bilimlerinin ilgi çeken çalışma alanlarından birini oluşturmuştur [1], [2]. Hareket tanıma ise başta basit tanımlamalar, sınıflandırmalar, gruplandırmalar yapılan bilgisayarla görme alanında günümüzde gittikçe popülerliğini artıran önemli bir kısmı oluşturmaktadır. İnsanların bireysel farklılıkları nedeniyle karşılaşılan ifade, duruş, kıyafet ve hareket yapılış şekli farklılıkları; ayrıca, kamera hareketleri, ışıklandırma varyasyonları ve perspektifin etkisi gibi dış etkenler hareket algılama konusunda büyük değişiklikler içeren, fakat ortak sınıf içerisinde toplanması gereken veri bilgisinin varlığını ortaya koymaktadır.

Hareket tanıma üzerine farklı yapılar kullanılarak çeşitli yaklaşımlar ileri sürülmüştür. Bu yaklaşımları esas farkı oluşturan özellik çıkarımı adımlarıyla kategorize etmek mümkündür. Özellik çıkarımı adımıyla kullanılacak özellikler ideal olarak kişinin görünümünden, arka plandan, bakış açısından ve eylem gerçekleştirme şeklinden bağımsız, genellenebilir olmalıdır. Hareket zaman kavramını da içerdiğinden özellikler, genel olarak, zaman kavramını da içeren 3 boyuttan veya görüntüden çıkartılan özelliklerin ardı sıra dizilmesiyle elde edilen zamansal kavramlarla ifade edilmektedir.



Özellik tanımlamaları iki ana grupta toplanmıştır: bütünsel tanımlama ve yerel tanımlama [2].

Bütünsel tanımlayıcılar video girdisinin anlık görüntüsünden bilgi çıkarımını bir bütün olarak ele almaktadırlar. Bu görüntü üzerinde öncelikle insan bulunur, daha sonra ilgi bölgesinden komple bir tanımlayıcı çıkarılır. Bilgiyi bir bütün halinde tuttıkları için bütünsel tanımlayıcılar sınıflandırma açısından güçlüdürler. Fakat doğru lokalizasyon gereksinimi duyduğu için gürültüye daha çok duyarlıdırlar. İlgi alanı çoğunlukla arka plan çıkarımı veya takip ile elde edilir. Bilgi elde etme yahut ifade etme şekli ise genel olarak silüetlerden, kenar özelliklerinden veya optik-akış özelliklerinden faydalanarak oluşturulmaktadır. Zaman kavramı bilgisi ise çoğunlukla görüntünün arka arkaya dizilmesiyle eklenir.

Silüet çıkarımı, literatürde sıklıkla karşılaşılabilen hareket ifade edim biçimlerinden biridir. Bu gösterimin ilk kullanımlarından biri, zamansal şablonlar çıkarılarak şekil değişiklikleri geçmişi ile hareketin ifade edilmesidir [3]. Belirtilen çalışmada şekil değişiklikleri incelenerek silüetlerin görsellerinin birleştirilmesi ile “Hareket Geçmişi Görüntüsü” elde edilerek insan hareketleri anlamlandırılmaya çalışılmıştır. Wang, ise çalışmasında silüetlere R transformu uygulamıştır [4]. Souvenir ve Babbs, R transformuna üçüncü boyut olarak zamanı katmış ve hareket tanıma işlemini gerçeklemiştir [5]. Weinland, çalışmasında silüetleri öklid uzaklıkları ile eşleyerek sınıflandırma yapmıştır [6]. Başka bir çalışmada ise silüetleri Chamfer uzaklığı ile sınıflandırmıştır [7]. Silüetlere bağlı olarak uzay-zaman şekilleri gösterimi de ortaya konulmuştur [8]. Gorelick ve arkadaşları, bu paralelde, silüeti t zaman aralığında takip edip uzay-zaman izlenimleri oluşturmuş ve bu görüntünün Poisson denklemiyle özellik çıkarımı yapmış ve hareket tanımlamıştır [9]. Lu ve Little, HOG (Histogram of oriented gradients – Yönlü gradyanlar histogramı) üzerine PCA (Principal component analysis – Ana bileşen analizi) uygulayarak bir hareket tanıma modeli tasarlamıştır [10].

Özellik çıkarımı adımıyla kullanılan bir diğer tanımlayıcılar ise yerel tanımlayıcılardır. Yerel tanımlayıcılar birbirinden bağımsız ufak parçacıklar halinde ifade edilirler. Yerel tanımlayıcılarda önce uzay-zaman ilgi noktaları elde edilir ve bu noktalar etrafında yerel özellikler çıkarılır. Daha sonra bu parçacıklar birleştirilerek son temsilci ortaya

konular. “Kelime Çantası” (bag of words) da denilen bu modelleme; ufak parçacıklar halinde bulunan özelliklerin birleştirilmesi ve temsilcilerinin seçilmesi esasına dayanmaktadır. Yerel tanımlayıcılar gürültü veya kısmi kayıplar gibi durumlarda bütünsel tanımlayıcılara göre daha başarılıdır. Fakat çoğunlukla yüksek miktarda bir önişleme gereksinimi duymakta, kamera hareketi gibi bir durumdan ise daha çok etkilenmektedirler. Uzay-zaman yaklaşımlarının yerel 3 boyutlu parçalar halinde kullanılmasının başarılı sonuçlar ortaya koyduğu çeşitli çalışmalarda gözlemlenmiştir [11], [12], [13], [14], [15].

Laptev, çalışmasında yerel HOG ve HOF (Histogram of oriented flow – Yönlü akış histogramı) tanımlayıcılarını kullanmıştır [16]. HOG’ un 3 boyutlu nasıl kullanılacağı Kläser [17] tarafından ortaya atılmıştır. Wang ise bir diğer çalışmasında gradyanların ve akış bilgisinin birleştirilmesinin ayrı ayrı kullanılmasına göre en iyi sonucu verdiğini bulmuştur [18].

Laptev ve Lindeberg, Harris köşe tanımlayıcısını 3 boyutlu hale getirdiler ve bunlarla parçalar çıkarımı yapıp hareket tanımladılar [19]. Willems çalışmasında SURF (Speeded-up robust features - Hızlandırılmış gürbüz öznitelikler) noktalarını 3 boyutlu hale getirmiştir ve böylece hareket tanımada kullanmıştır [20].

Özellik çıkarımı adımında elde edilen gösterimler ve betimleyiciler SVM (Support vector machines - Destek vektör makineleri ) [16], [21], HMM (Hidden markov models - Saklı markov modelleri) [7], [22] gibi çeşitli makine öğrenmesi metotları kullanılarak modellenmeye çalışılmıştır.

Genel olarak hareket tanıma literatürüne bakıldığında zaman, videodan bilgi çıkarımı konusu çok farklı yöntemler içerse de temelde bir hareket bilgisi ifadesinden ve bunun yanı sıra durgun görüntünün bilgisinin de aktarılmasından oluştuğu görülmektedir. Hareketli bilgi için optik akış kullanımı yaygındır [16], [23], [24]. Durgun görüntünün ayırt edici özellik olarak kullanılmasında ise HOG gibi kenar özellikleri, SURF gibi boyuttan bağımsız özellik çıkarıcılar veya çeşitli 2 boyutlu özellik çıkarıcıların 3 boyutlu hallerinin kullanımı yaygındır [10], [25]. Bu iki bilginin, durgun görüntü bilgisi ve hareket bilgisinin, beraber kullanımının ayrı kullanımına göre daha başarılı olduğu ortaya konulmuştur [18].

## 1.2 Tezin Amacı

Hareket tanıma konusu gelecek vaat eden ve diğer konulara da öncülük edebilecek, yeni çalışma ve uygulama alanları çıkarabilecek bir konudur. Bilgisayarla görü günümüzde gelişen video, kayıt ve görsellik teknolojisi ile birlikte popülerliğini, uygulama alanını ve şeklini gittikçe değiştirip artırmaktadır.

Hareket tanıma kavramı günümüzde insan hareketleri üzerine yoğunlaşmış durumdadır. Bu durumun elbette en büyük etkenlerinden biri bilgi ve uygulamanın insan takibi, insan algılaması üzerinde yoğunlaşmasıdır. Bu çalışmada da bu konu üzerinde durulmuş, hareket tanımlamaya derinlemesine bir inceleme yapılmış ve anlamlandırma çalışmalarına önyak olabilecek bir sistem tasarlanmıştır. Araştırmacıların insanların doğal olarak yapabildiklerini yapay zeka ile çalışan sistemlere yaptırabilme güdüsünün de etkisiyle son yıllarda hareket tanıma konusu çok düşük seviyelerden çok yüksek tanıma başarısına sahip modellere dönüşmüştür. Böylece yeni uygulama alanlarında da başarıyla uygulanabilecek yapılar ortaya konmuştur. Bu tez çalışmasında önerilen sistemde bahsedilen başarıyı sağlamak üzere basit bir yapı oluşturulmaya çalışılmıştır.

## 1.3 Hipotez

Bu çalışmada, insan hareketi tanımlama için, hareket bilgisi ifadesinin durgun görüntü bilgisi ile birleştirilmesiyle, her iki bilginin zaman düzleminde ortaklaşa kullanılabileceği bir sistem tasarlanmış ve gerçekleştirilmiştir. Sistemin tasarımında, optik akış üzerine uygulanan yönlü gradyanlar histogramı temel alınmış ve 3. boyut olarak düşünebileceğimiz zaman kavramı da dahil edilip boyuttan bağımsız hareket tanımlayabilen özgün bir yapı geliştirilmiştir. Optik akış yöntemi, klasik olarak kullanımı olan hareketin yönünü bulma amacındansa aslında bir filtre olarak düşünülebilecek şekilde uygulanmıştır. Optik akış ile ilgi alanında bulunan hareket halindeki parçaların hareket kuvvetleri belirlenmiştir. Böylece, optik akış bir filtre görevi ile ilgi alanı içerisinde bilgi ifade etmeyen kenarların ortadan kaybolmasını sağlamış ve bilgi ifade eden kenarların ise korunmasını sağlamıştır. Sonuç olarak ise hareketli bilgi ile görüntünün kenar özellikleri birleştirilebilmesine olanak sağlayan bir düzen oluşturulmuştur. Kenar özelliklerinin korunması ile de özellik çıkarımı adımında kenar

geçiş özelliklerini inceleyen yönlü gradyanlar histogramının efektif bir biçimde çalışmasına ortam hazırlanmıştır. Sonrasında ise zamanda kayan bir pencere yardımıyla çıkarılan özelliklere zaman kavramı aktarılmıştır. Böylece çok yüksek başarıya sahip özgün bir sistemin ortaya çıkması sağlanmıştır.

### HAREKET TANIMANIN ESASLARI

Hareketlerin zamanda değişen nitelikler olduğu ve bu süreçleri ifade edebilmenin kendi bütünü içinde ortak ve diğer sınıflarca kolayca ayırım gösterilebilecek nitelendiricilerle ortaya konulması gerektiğinden yola çıkarak, hareket zamansal bir örüntü olarak tanımlanmıştır. Değişken boyutta, değişen koşullar altında başarıyla çalışabilecek bir yapının esasını da bu örüntülerin başarıyla tespit edilebilmesinin oluşturacağı aşikardır.

Bu tez çalışmasında, verilen esas gözetilerek, “Hareket en iyi biçimde nasıl tanınabilir?” sorusunun cevabı aranmıştır. Öncelikle literatürdeki bilgiler üzerinde incelemeler yapılmış ve aslında hareketin ifadesinde hem durgun halde olduğu düşünülen anlık görüntünün, hem de hareket bilgisinin birleştirilmesi gerektiği sonucuna varılmıştır. Ayrıca hareketin aslında zaman düzlemi ile birlikte 3 boyutlu bir bilgi olduğu kavranmıştır. Çıkarılan sonuçlar neticesinde hem durgun anlık görüntünün, hem hareket bilgisinin hem de peşi sıra devam eden zaman kavramını ortaklaşa ifade edebilecek bir yapı ortaya konulmasının gerekliliği ön plana çıkmıştır.

Hareket tanıyabilecek bir yapının öncelikle görüntüdeki ilgi alanını kestirmesi gerekmektedir. İşlemler daha sonra bu ilgi alanından özellik çıkarımı ile devam edip doğru etiketi bulma ile sonlanacaktır. Bu tez çalışmasında ağırlık özellik çıkarımı adımına verilmiş olup, iyi bir özellik çıkarımı sayesinde çok basit sınıflandırıcılar kullanılarak dahi başarılı olunabileceği öngörülmüştür.

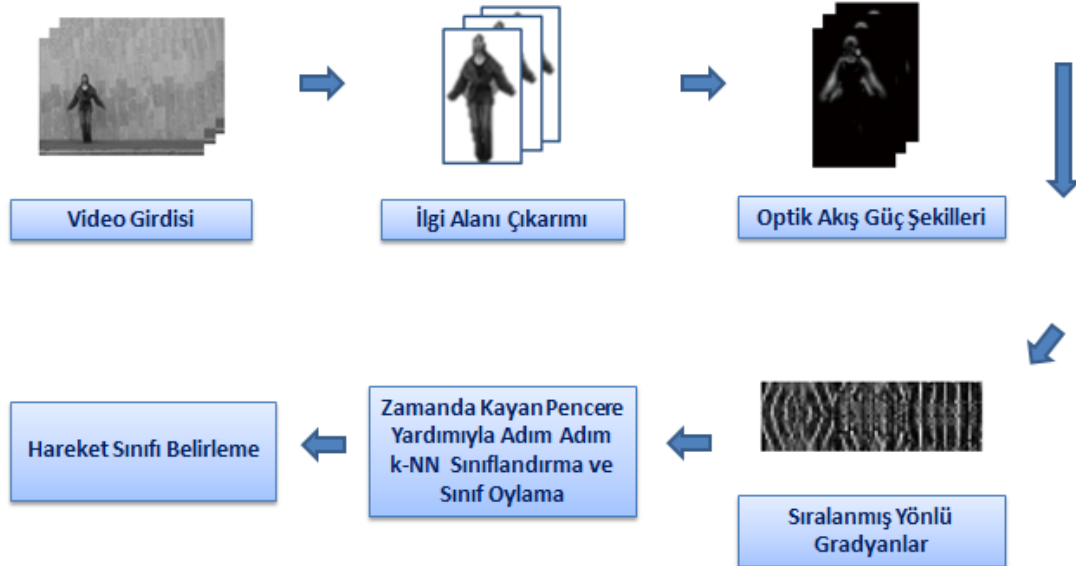
Hareketin aslında bütün farklılıklara rağmen tanınabilir bir yapıda olduğunu ortaya koyabilecek bir sistem, çeşitli koşullar altında başarıyla çalışabilmelidir. Bu durumu gerçekleyebilmek için sistemin tasarımında boyuttan, ortamdaki, kişilerden bağımsız

nitelendiriciler gerekmektedir. Bu nitelendiricilerin iyi seçilmesiyle beraber sınıflandırma aşamasında başarı çok daha iyi bir düzeye çıkacaktır.

## 2.1 Değişken Koşullarda Hareket Tanıyan Bir Sistem

Önerilen sistem, hareketlerin zamanda değişen nitelikler olduğu esastan yola çıkarak, hareketin zamansal örüntüsünü tanımlamayı hedeflemiştir. Sistem oluşumunda kullanılan algoritmaların doğru seçilmesiyle beraber ışık, renk, boyut gibi değişkenlerden minimum etkilenen bir yapıda olması sağlanmıştır. Böylece değişken koşullar altında başarıyla çalışabilen bir sistemin ortaya konulduğundan bahsedilir.

Sistemi genel yapısı şöyledir. Video girdisinden hareketsiz arka planın çıkartılması ile ilgili alanı olan ön plan görüntüsü elde edilir. Ön planda oluşan hareket optik akış yöntemi ile belirlenir. Optik akışta hareketin olduğu noktaların hareket güçlerine göre yüksekten düşüğe sırasıyla parlaktan koyuya doğru değerlendirilmesiyle optik akış güç şekli olarak da adlandırılan hareket bilgisini oluşturan şekil elde edilir. Bu şekillerin yönlü gradyanları hesaplanır ve arka arkaya dizilir. Oluşturulan bu yapıda, betimleyici dizinin içinde pencere yardımıyla gezilir ve sınıflandırma yapılır. Sistemin blok diyagramı Şekil 2.1’de verilmiş olup adımlar sırasıyla incelenecektir.



Şekil 2. 1 Gerçek zamanlı video görüntülerinden insan hareketlerinin tanınması ve yorumlanması sistem blok diyagramı

## 2.2 Video Görüntülerinin Alınması

Giriş bilgisi olarak kullanılan video görüntüleri sistemin amacına uygun olarak, insan hareketi içeren görüntülere aittir. Sistemde video girdisi olarak hareket tanıma alanında sıklıkla kullanılan bir veri seti olan Weizmann hareket tanıma veri seti (Şekil 2.2) kullanılmıştır [9]. Arka planı da içinde barındıran bu görüntülerden önce ilgi alanı çıkarılmalıdır.



Şekil 2. 2 Weizmann hareket tanıma veri seti

## 2.3 İlgi Alanı Çıkarımı

İlgi alanı video girdisinin insanı içeren kısmı olarak belirlenmiştir. Böylece, hareketin arka plandan bağımsız ve sadece insan üzerine odaklanması sağlanmıştır.

Renkli halde olan görüntüler bu adımda gri seviyeye dönüştürülmüştür, dönüştürme için renk değerleri olan kırmızı(R), yeşil(G) ve mavi(B) standart dönüşüm matrisi ile çarpılarak gri seviye değeri elde edilmiştir.

$$gri\ seviyesi\ deęeri = [0.299\ 0.587\ 0.114] * [R\ G\ B]^T \quad (2.1)$$

(2.1)' in uygulanması ile elde edilen gri seviye değerlerden oluşan görüntü bir sonraki adım için kullanıma hazır hale getirilmiştir.

İlgi alanı çıkarımı için ön planın arka plandan ayrılmasının video girdisi olan Weizmann hareket tanıma veri seti için yeterli olacağı görülmüş ve arka plan çıkarımı ile ön plan elde edilmiştir. Elde edilen ön plan üzerine blob analizi yapılmasıyla beraber insanı içeren bölgenin koordinatları belirlenmiştir. Bu analiz, bağlantılı olan kısımların

birleştirilip bir bölge olarak kabul edilmesi esasına dayanmaktadır (Şekil 2.3). İkili geçiş algoritması ise komşuluk bağları bulunması için kullanılmıştır. Bu algortmada iki geçiş ile bağlı bölgenin bulunmasını sağlar. Birinci geçişte her bir eleman için, arka plan değilse komşulara sırayla gözetilir, komşu yok ise tek olarak işaretlenir, komşu var ise en küçük etiketli komşunun etiketi ile şu anki eleman işaretlenir. İkinci geçişte ise her bir eleman için, arka plan değilse en küçük eşit etiketli eleman ile işaretlenir ve böylece bağlantılı bölge analizi yapılmış olur. Analiz sonucu ilgi alanının koordinatları elde edilir. Elde edilen koordinatlar daha sonraki adımlarda esas alınacak bölgeyi oluşturacaktır.

Çizelge 2. 1 İkili geçiş algoritması

---

**İkili Geçiş Algoritması**

---

```
1 : A = mxn boyutlu görüntü matrisi
2 : bağlantılar = []
3 : etiketler = mxn boyutlu A matrisinin arka plan değeri ile ilk değerlendirilmiş matris
4 :
5 : Birinci geçiş
6 :
7 : for i=1 to m
8 :   for j=1 to n
9 :     if A[i][j] is not arka plan
10:
11:       komsular = A[i][j] ile aynı değere sahip bağlı elemanlar
12:
13:       if komsular is boş
14:         bağlantılar[sonrakiEtiket] = sonrakiEtiket'in değeri
15:         etiketler[i][j] = sonrakiEtiket
16:         sonrakiEtiket += 1
17:
18:       else
19:         L = komsuların etiketleri
20:         etiketler[i][j] = min(L)
21:       for etiket in L
22:         bağlantılar[etiket] = union(etiketler[etiket], L)
23:
24 : İkinci geçiş
```

---

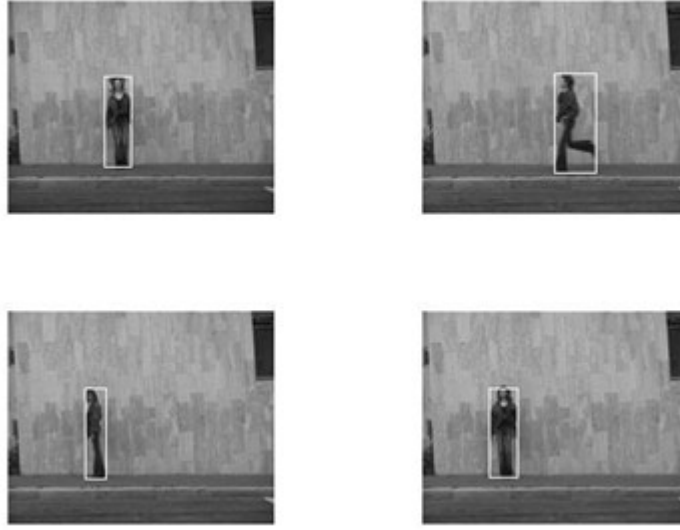


Çizelge 2. 1 (devam)

---

```
25:
26: for i=1 to m
27:   for j=1 to n
28:     if A[i][j] is not arka plan
29:       etiketler[i][j] = min(bağlantılar[etiketler[i][j]])
30:
31: return etiketler
```

---



Şekil 2. 3 İlgili alanı çıkarımı – blob analizi sonuçları

#### 2.4 Optik Akış Güç Şekilleri

Görüntüdeki hareketin algılanmasını sağlayan optik akış, geliştirilen sistemin de temel adımını oluşturmaktadır. Art arda gelen görüntülerden Horn-Schunck optik akış hesaplama yöntemi kullanılarak hareket bilgisi elde edilmiştir [26].

Horn-Schunck yönteminde piksellerin zaman bağlı parlaklığının değişmeyeceği kabul edilir. Görüntünün bir  $t$  zamanında  $(x, y)$  noktasının parlaklığı  $I(x, y, t)$  ise, nokta hareket etse dahi parlaklığında değişim olmamasından yola çıkılarak (2.2) elde edilir.

$$\frac{dI}{dt} = 0 \quad (2.2)$$

Zincir kuralı ile (2.3) oluşturulur,

$$\frac{\partial I}{\partial x} \frac{dx}{dt} + \frac{\partial I}{\partial y} \frac{dy}{dt} + \frac{\partial I}{\partial t} = 0 \quad (2.3)$$

$u$ ,  $x$  yönündeki;  $v$ ,  $y$  yönündeki değişimi göstermek üzere (2.4)

$$u = \frac{dx}{dt}, v = \frac{dy}{dt} \quad (2.4)$$

(2.3) ve (2.4)'ten iki bilinmeyenli (2.5) ortaya konulabilir.

$$I_x u + I_y v + I_t = 0 \quad (2.5)$$

(2.5) daha iyi bir şekilde (2.6) olarak düzenlenebilir.

$$I_{(x, y)} \cdot (u, v) = -I_t \quad (2.5)$$

Bu çıkarımlar ise tek başına optik akış özelliklerini çıkarmaya yetmemektedir. Komşu piksellerin birbirine benzer akış hareketi yapacağından yola çıkılarak  $x$  ve  $y$  bileşenlerinin Laplace değerlerinin karelerinin toplamı(2.6) duruma akıcılık kısıtını dahil edip denklemler sistemini çözmeyi sağlamaktadır.

$$\nabla^2 u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}, \nabla^2 v = \frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} \quad (2.6)$$

Böylece denklemler sistemi bir minimizasyon problemi haline dönüşmüştür. Görüntünün parlaklığın değişmezliği ve akıcılık kısıtındaki hataların minimize edilmesi optik akış değerlerini elde etmeyi sağlayacaktır. Parlaklık değişmezliğinin pratikte hiç hatasız olması sıfır değerini alması beklenmemektedir, fakat parlaklık değişim hatası ortaya çıkmaktadır.  $\varepsilon_b$  parlaklık değişimdeki hata ve  $\varepsilon_c$  akıcılık kısıtı hesabındaki hata ve  $\alpha^2$  denklik ağırlık faktörü olmak üzere çözüm (2.7) olarak ifade edilebilir.

$$\varepsilon^2 = \iint (\alpha^2 \varepsilon_c^2 + \varepsilon_b^2) dx dy \quad (2.7)$$

2.7'nin düzenlenmesi ile minimize edilmesi gereken optik akış çözüm denklemi 2.8 elde edilir.

$$\iint [(I_x u + I_y v + I_t)^2 + \alpha^2 (\|\nabla u\|^2 + \|\nabla v\|^2)] d_x d_y \quad (2.8)$$

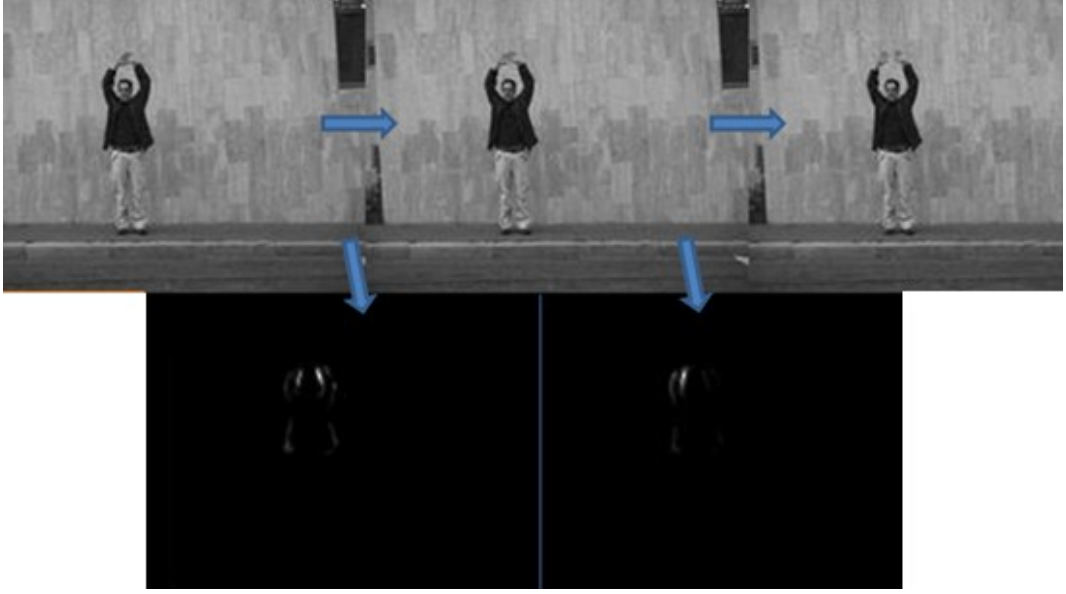
Horn-Schunck yönteminde (2.8)'in minimize edilmesi sonucu  $x$  ve  $y$  yönündeki hareket değişimi bilgisi miktarıyla birlikte elde edilir. Bu çalışmada, (2.9) kullanılarak,  $x$  yönündeki hareket değişimi  $u$  ve  $y$  yöndeki hareket değişimi  $v$  kullanılarak piksellerin optik akış güçleri( $E$ ) hesaplanmıştır. Bu güç değerleri hareketli kenar özelliklerini koruyarak hareketsiz kenarları ortadan kaldırmıştır ve böylece hareketi daha iyi ifade edebilecek bir yapı oluşturmuştur.

$$E = \sqrt{u^2 + v^2} \quad (2.9)$$

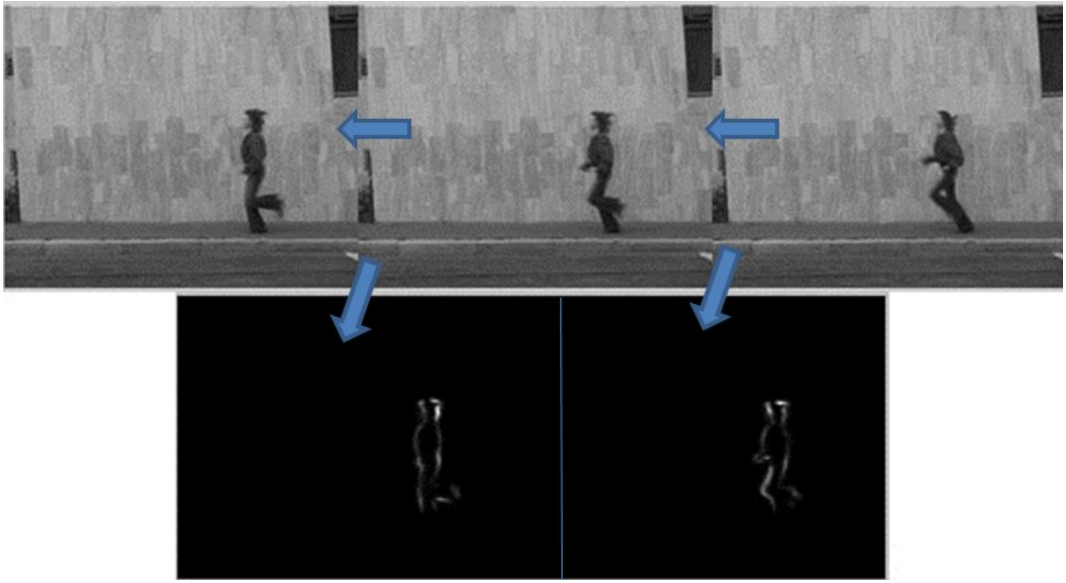


Şekil 2. 4 Optik akış güç şekli örnekleri (normalize edilmiş)

(2.9)'un uygulanmasının ardından çıkan güç değerlerinde hareketli bölgeler yüksek değerler alırken, hareketsiz bölgeler ise sıfır veya sıfıra yakın değerler alır. Bu değerler normalize edildiğinde ise Şekil 2.4'de görülebileceği gibi optik akış güç şekillerini oluşturur. Bu görüntüde hareketli bölgeler beyaza yakın, hareketsiz bölgeler ise siyah değer almaktadır. Bu şeklin avantajlarından biri hareketin olduğu kenarları bozmadan koruyabilmesidir. Hareket içermeyen noktalar ise optik akış hesaplanırken zaten ortadan kaybolduğu için bir sonraki adıma girdi olarak kullanılacak olan bu şekil istenilen kenar özelliklerin çıkarılmasında başarılı olmaktadır. Örneğin, Şekil 2.5'te yalnız üst vücut bölgesinin hareketli olduğu bir örnek gösterilmekte; Şekil 2.6'da ise tüm vücudun hareketli olduğu bir örnek gösterilmektedir. Hareketli kenarların kendini koruyup hareketsiz kenarların kaybolduğu rahatça gözlemlenebilmektedir.



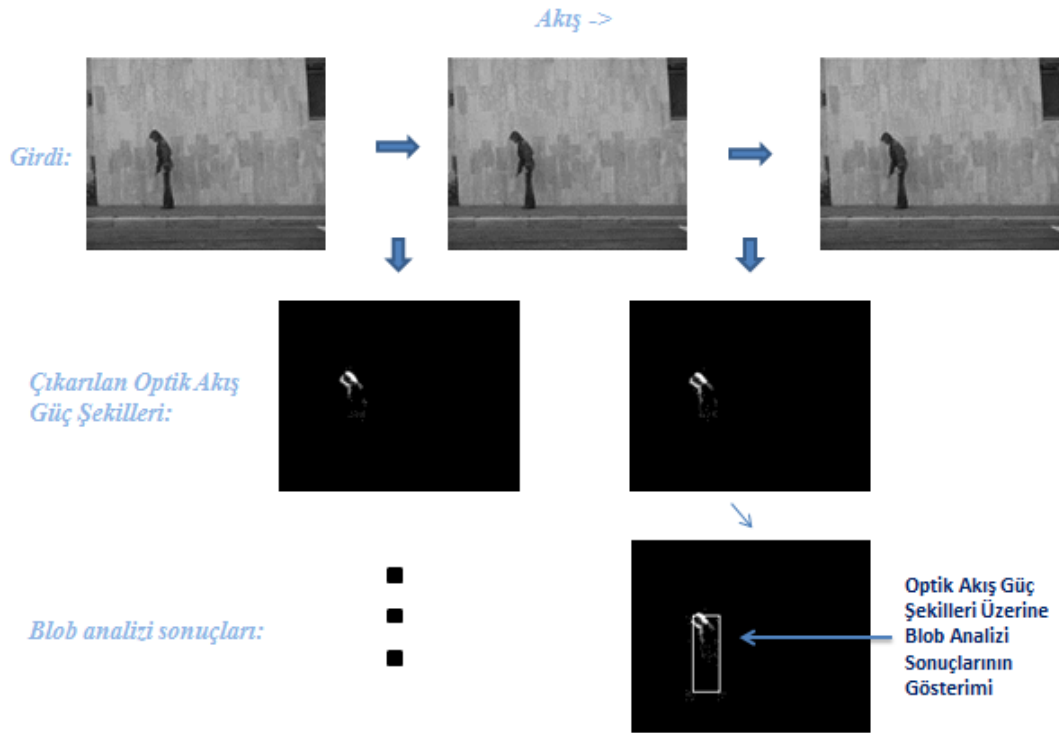
Şekil 2. 5 Optik akış güç şekillerinin elde edilmesi (yalnız üst vücut bölgesi hareketli)



Şekil 2. 6 Optik akış güç şekillerinin elde edilmesi (tüm vücut hareketli)

Optik akış güç şekilleri video girdisinin hareket bazında ifade edilmesini sağlamaktadır. Önceki adımda hesaplanan ilgi alanının bu adımda kullanılmasıyla beraber gri seviyeden oluşan ve “Ne kadar hareket etti?” sorusunun cevabını da barındıran insan içeren anlamlı bölge bu adımda hesaba hazır hale getirilmektedir. Normalde ilgi alanında insan görüntüsü üzerinde kenar özellikleri incelendiğinde bütün bilgiler içerilmekte ve bu ifade hareket eden bir varlık için tanımlamada yeterli olmamaktadır.

Bu adımda bulunan hareket ifadesi kenar özelliklerini koruyan Horn-Schunck optik akış yöntemiyle birlikte anlam ifade eden kenarlar korunmuştur. Hareket içermeyen dolayısıyla tanımlamada bilgi ifade etmeyecek olan kenarların çıkarılması sağlanmıştır. Böylece bir sonraki adımda kenar özelliklerine bağlı olan özellik çıkarımında, optik akış güç şekilleri, hem gerekli bilginin daha iyi ifade edilmesini sağlayan, hem de gereksiz bilgilerin yok edilmesini sağlayan, çok iyi bir ön işlem adımı olduğu öngörülmektedir.



Şekil 2. 7 İlgili alanı üzerinde optik akış güç şekillerinin incelenmesi

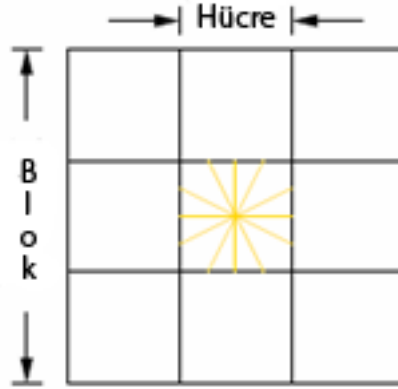
## 2.5 Sıralanmış Yönlü Gradyanlar

Optik akış güç şekilleri elde edildikten sonra elde edilen hareket bilgisini de içeren ilgili alanı üzerinde anlık durgun görüntü de denilebilecek görüntünün kenar özellikleri çıkarımı yapılmıştır. Durgun görüntünün tanımlayıcısının çıkarımı için insan görüntülerinde üstün başarı ortaya koyan HOG yöntemi kullanılmıştır [27]. HOG esasen insan tanıma amacıyla kullanılmış bir yöntemdir. Dolayısıyla insan hareketi tanımayaya birebir elverişli değildir. Bu çalışmada, yapılan bir uyarlamayla insan hareketi tanımlanabilecek bir yapıya kavuşturulmuştur. Sıralanmış yönlü gradyanlar hesabından önce HOG kullanılarak nasıl betimleyici çıkarıldığı incelenecektir.

### 2.5.1 Yönlü Gradyanlar Histogramı

Yönlü gradyanlar histogramı nesne tanıma için kullanılan bir tanımlayıcı çıkarma yöntemidir. Belirtilen metotta görüntüdeki yönlü gradyanlar yerel olarak hesaplanır ve yönleri sayılır. Bu açıdan bakıldığında SIFT [28] (Scale-invariant feature transform - Ölçekten bağımsız öznelik dönüşümü), SURF [29] gibi yöntemlere benzese de üst üste çakışan ızgara kullanımı ve yerel parlaklık normalizasyonu adımlarıyla beraber hem farklılığını ortaya koymakta, hem de yerel anlamda daha iyi bir betimleyici sunmaktadır.

HOG'un arkasındaki temel düşünce nesnelerin yerel kenar yönlerinin dağılımları ile ifade edilebilir olmasıdır. HOG'un uygulanmasında görüntü hücre denilen küçük bağlantılı bölgelere ayrılır (Şekil 2.8). Her hücre için bu hücrelerin içindeki piksellerin gradyan hesabı yapılır. Daha sonra bölgelerin histogramları birleştirilerek son tanımlayıcı oluşturulur. Bu son tanımlayıcı, geliştirilmiş başarımlar için blok denilen birleştirilmiş hücreler (Şekil 2.8) bazında normalize edilir. Bu normalizasyon adımı aydınlatma veya gölgeleme değişikliklerinin daha iyi ifadesi ile sonuçlanır.



Şekil 2. 8 Blok ve hücre bölgelerinin gösterimi

HOG tanımlayıcısının diğer tanımlayıcı yöntemler üzerinde birkaç temel avantajı vardır. HOG tanımlayıcısı yerelleştirilmiş hücreleri üzerinde çalıştığından, nesne yönelimi dışında, geometrik ve fotometrik dönüşümler değişmezliği konusunda benzer özellik çıkarımı adımlarına göre çok daha başarılıdır. Çünkü çoğu bahsedilen değişim sadece büyük bölgelerde görünür. Küçük bölgeler üzerinde hesaplanan HOG belirteçleri bu konu üzerinde daha başarılı bir yaklaşım ileri sürmektedir.

Yönlü gradyanlar hesabını yaparken ilk adım olarak gradyanların elde edilmesi gerekmektedir. Bunun için giriş görüntüsünün yatay ve dikey yönlerde kenarlarını belirlemeye yarayan türev maskeleri, sırasıyla (2.10) ve (2.11) uygulanır.

$$[-1,0,1] \quad (2.10)$$

$$[-1,0,1]^T \quad (2.11)$$

Dalal ve Triggs çalışmalarında daha karmaşık maskeleri de kullanmışlardır fakat en basit ama en efektif olan türev maskesi en iyi sonucu vermiştir [27]. Bu çalışmada da bu temelden yola çıkılarak kenar belirleme maskeleri kullanılmıştır.

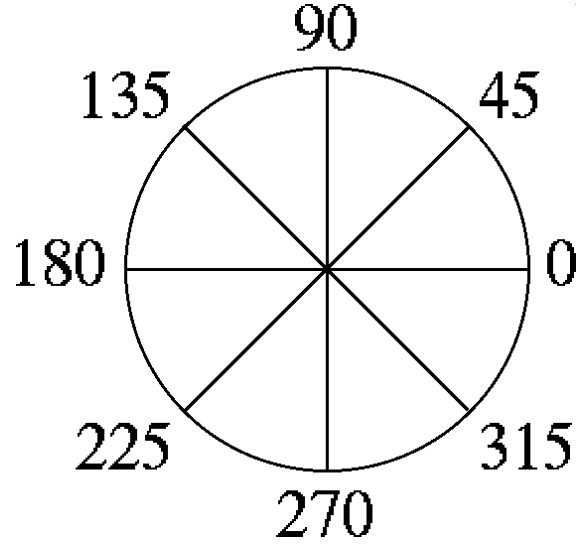
Maskelerin uygulanmasının ardından her bir piksel için elde edilmiş olan gradyanlardan büyüklük ve yön analizi yapılabilir.  $s_x$  x yönündeki,  $s_y$  y yönündeki gradyan değerini ifade etmek üzere (2.12) kullanılarak gradyan büyüklüğü elde edilebilir.

$$s = \sqrt{s_x^2 + s_y^2} \quad (2.12)$$

Gradyan yönü ise (2.13) kullanılarak elde edilir.

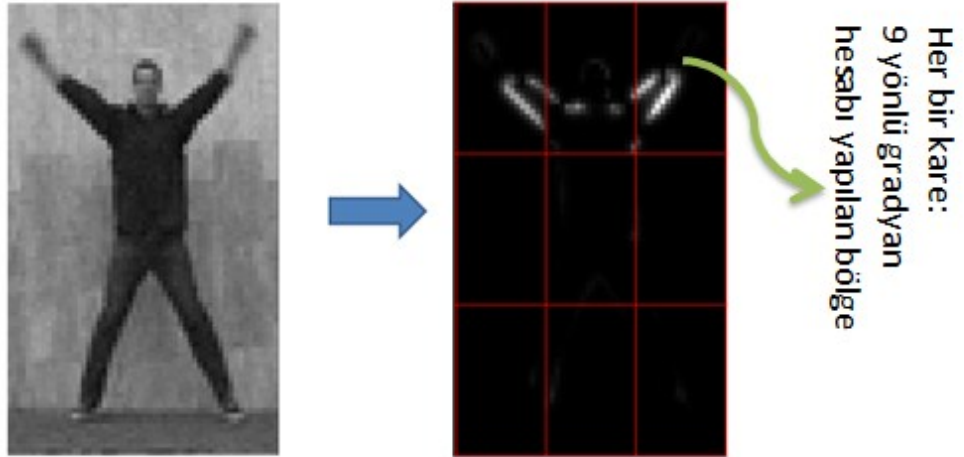
$$\theta = \tan^{-1}\left(\frac{s_x}{s_y}\right) \quad (2.13)$$

Bir sonraki adım ise hücre histogramlarının oluşturulmasıdır. Hücre denilen küçük piksel topluluklarının her biri gradyan değerlerine bağlı olarak histograma oy vermektedir. Hücreler dikdörtgensel şekilde olup, gradyan yönleri işaretli yani 0-360 derece arası (Şekil 2.9), yahut işaretli 0-180 derece arası hesaplanabilmektedir. Bu hesap aralıkları eşit bölmelere ayrılıp istenilen kadar histogram kanalından bakılıp hücre histogramları oluşturulabilmektedir. Yine Dalal ve Triggs çalışmalarında yaptıkları testlerde işaretli yani 0-180 derece arası 9 kanallı histogram almanın en iyi sonucu verdiğini ortaya koymuşlardır.



Şekil 2. 9 Gradyan yönleri

Hücre histogramları hesaplandıktan sonra bir sonraki adımda gradyan değerlerinin yerel değerlendirilmesi yapılmaktadır. Bu adım hücrelerin yerel olarak birleştirilmesi ve böylece blokların oluşturulmasını içermektedir. Bu adım ile birlikte aydınlanma ve karşıtlık değerlerine bağlılık oranı azaltılmaktadır. Dalal ve Triggs çalışmalarında 9 kanallı histogram ile birlikte 3x3'lük hücre bloklarının oluşturulmasının en iyi sonuçları verdiğini ortaya koymuşlardır [27]. Bu tez çalışmasında da Dalal ve Triggs'in insan tanıma için kullandığı ve en iyi sonuçları ürettiği gözlemlenen değerler esas alınmıştır.



Şekil 2. 10 İlgili alanı üzerinde 3x3 pencere ile yönlü gradyanlar histogramı hesabı yapılan bölgelerin gösterimi



İlgi alanının 3x3' lük parçalara ayrılması ayrıntılı olarak incelendiğinde ilgi alanı dikeyde yukarı, orta, aşağı kısımlara; yatayda sol, orta, sağ olarak kısımlara bölünmekte ve bu gösterim kafa, gövde, bacak ve kol kısımlarını daha iyi ifade etmeyi sağlamaktadır. Böylece hareketin ifadesi kolaylaşmakta, netleşmekte ve bu durum başarıya da pozitif olarak yansımaktadır. Bir optik akış görüntüsünün 3x3 lük bölümlenmesiyle 9 yönlü gradyan hesabı yapılmış ve bir şeklin en iyi gösterim şekli olan 3x3x9'luk gösterime sahip olmasını sağlamıştır [27]. Bu gösterim daha sonra 81'lik vektörler haline getirilip tek bir anlık imaj için tanımlayıcı oluşturulmuştur.

### 2.5.2 Yönlü Gradyanlar Histogramlarının Sıralanması

Yönlü gradyanlar histogtamı esasen obje tanıma için kullanılmıştır, dolayısıyla durgun görüntü üzerinde çalışmaya elverişlidir. Hareket tanıma olayında ise zaman kavramı devreye girmektedir. Bu kavramı yansıtabilmek ve daha iyi tanıma sonuçları elde edebilmek üzere; bir önceki adımda ortaya konulan ve hareket kavramını olaya dahil eden yönlü gradyanlar histogramı ayrıca bir zaman düzlemine ihtiyaç duymaktadır. Optik akış sadece iki çerçeve arasındaki geçişi incelemektedir. Optik akış değerleri üzerinden hesaplanan ve yönlü gradyan analizi yapılan değerler sadece iki çerçeve arasındaki bağıntıyı verecektir. Oysaki hareket denilen kavram birden çok çerçeveyi içermektedir. Bu bakış açısından yola çıkılarak optik akış güç şekilleri üzerinde yapılan hesap n adet çerçeve süresince devam ettirilmiş ve bu bilgi zaman kavramını da içeren bir hal almıştır.

Çizelge 2. 2 Yönlü gradyanların sıralanması

---

Yönlü Gradyanların Sıralanması
1 : $B = t \times 81(3 \times 3 \times 9)$ 'lik HOG betimleyici vektörü
2 : $t =$ çerçeve sayısı
3 : $n =$ zamandaki pencere büyüklüğü
4 : $S = [ ] [ ] [ ]$ Sıralanmış yönlü gradyanlar, ilk değeri olmayan, $(t-n+1) * n * 81$ boyutunda
5 :
6 : <b>for</b> $i=1$ <b>to</b> $t$
7 : <b>for</b> $j=1$ <b>to</b> $n$
8 : <b>if</b> $t >= n$

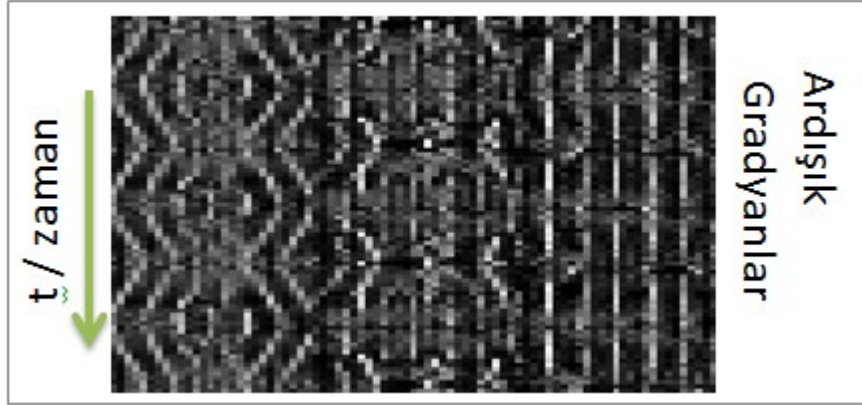
---

Çizelge 2. 2 (devam)

---

```
9:     for k=1 to 81
10:     S[t-n+1][n][k]=B[t][k]
11:
12: return S
```

---



Şekil 2. 11 Ardışık sıralanmış yönlü gradyanlar histogramının görselleştirilmesi

İlgi alanındaki bir optik akış güç şeklinin [27]'de belirtildiği üzere en yüksek başarıyı veren 3x3x9'luk HOG betimleyicilerinin yan yana dizilip 81'lik vektörler halindeki gösteriminin de zaman düzleminde art arda gelen  $n$  adet çerçevelerin dizilmesiyle hareket kavramından sonra zaman kavramını da barındıran sıralanmış yönlü gradyanlar elde edilmiştir. Sınıflandırmada bu aşamada elde edilen veri kullanılacak olup hem hareket, hem zaman, hem de kenar bilgisini en iyi biçimde ifade ettiği öngörüldüğünden hareket algısını tanımlamak için en başarılı veri olduğu düşünülmüştür.

## 2.6 Hareket Sınıfı Belirleme

$n$  zaman boyutundaki pencere büyüklüğü dolayısıyla art arda sıralanan HOG betimleyicisi adeti olmak üzere, sıralanmış yönlü gradyanlar ile elde edilen  $81 \times n$  boyutundaki zaman ve hareket kavramlarını içinde barındıran tanımlayıcıların hareketi çok başarılı bir şekilde ifade ettiği öngörülmektedir. Bu esastan yola çıkılarak gerçek zamanda çalışabilecek bir sistemin hız önceliği düşünülerek bu adımda çok basit bir sınıflandırıcı olan k-NN( k-Nearest Neighbour - k-en yakın komşuluk) kullanılmıştır.

### 2.6.1 k-En Yakın Komşuluk Sınıflandırıcısının Kullanımı

k-NN sınıflandırıcısı, örüntü tanıma konusunda sıklıkla kullanılan basit bir sınıflandırıcıdır. Verilen bir örnek için uzayda en yakın eğitim örneklerini bularak nesnelere sınıflandırır. Nesnenin sınıflandırılması için komşularının oy çokluğuna bakar ve böylece karar verir.

k-NN sınıflandırıcısının bir eğitim aşaması bulunmamaktadır. Eğitim verisi arama için kullanılacak veriyi oluşturur. Her sınıflandırma için, sınıflandırılacak örneğin benzerleri bu veri içerisinde aranır. Etiketlenmiş eğitim verisi üzerindeki her bir eğitim örneği ile gelen sınıflandırılacak örnek arasında uzaklık hesabı yapılır.  $p$  sınıflandırılacak örneğin her bir özelliği  $q$  da eğitim uzayındaki bir örneğin her bir özelliğini olmak üzere ve  $m$  vektör boyutunu belirtmek üzere, bu tez çalışmasında, Öklid uzaklığı(2.14) ve Manhattan uzaklığı (2.15) ile uzaklık hesabı yapılmıştır.

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_m - q_m)^2} = \sqrt{\sum_{i=1}^m (p_i - q_i)^2} \quad (2.14)$$

$$\sum_{i=1}^m |p_i - q_i| \quad (2.15)$$

Öklid uzaklığı k-NN için standart, en çok kullanılan uzaklık tanımlayıcısıdır. Manhattan uzaklığı ise sadeliği, daha basit oluşu dolayısıyla daha hızlı çalışmasından ötürü bu tez çalışmasında gerçek zaman hız koşulları düşünülerek esas olarak tercih edilmiştir.

Her bir eğitim betiğiyle karşılaştırılıp elde edilen uzaklık farklılıklarından sınıflandırılacak girdiye en yakın k tanesinin belirlenmesi ile en yakın komşuların belirlenmesi adımı tamamlanmaktadır. En yakın k adet tanımlayıcı betiğin incelenmesiyle, k genellikle 1,3,5,7... gibi tek değerlerle denenir. Bu en yakın k adet tanımlayıcıların sınıfları belirlenir. Sınıflandırılacak örneğin sınıf çokluğu bulunan gruba atanmasıyla sınıflandırma işlemi tamamlanır.

## 2.6.2 Sıralanmış Yönlü Gradyanların k-En Yakın Komşuluk Sınıflandırıcısı ile Sınıflandırılması

$81 \times n$  boyutundaki betimleyiciler k-en yakın komşuluk sınıflandırıcısı ile kendilerine en yakın betimleyicilerin eğitim verisinden elde edilmesiyle ve kendilerinin sınıflara ait olduğunun etiketlenmesiyle hareket tanımanın ilk adımı atılmış olmaktadır.

Hareket bir süreklilik içerdiğinden yalnız tek bir sınıflandırmanın yeterli olmayacağı ortadadır. Buradan yola çıkılarak k-en yakın komşuluk sınıflandırıcı zamanda kayan bir pencere yardımıyla sırasıyla girdiden elde edilen  $81 \times n$ 'lik vektörleri teker teker etiketlemekte kullanılmıştır. Böylece sınıflandırma işlemi artık bir oy verme işlemine dönüşmüştür. Stabil oylar sonucu hareket sınıfını belirlenmiş. Kısa zaman aralığında farklı hareket gruplarıyla etiketlenen betimleyiciler sistemin uyarısı niteliğinde olmuştur, çünkü milisaniyeler ile ifade edilecek çok kısa süre zarfında hareketin değişmeyeceği ortadadır. Böylece ortaya birbirinin doğrulayıcısı olabilecek bir yapı kendiliğinden ortaya çıkmıştır.

Zamanda kayan pencere ile tanıma yapma çeşitli avantajları ve dezavantajları da bulunmaktadır. Kayan pencere aslında aynı betimleyicilerin bir kısmını içeren bir sınıflandırma metodunu ortaya koymaktadır. Bu durum da kendi kendini doğrulayıcı yahut eğer bir düzen içermiyorsa kesinliğinden emin olunamayan durumu ortaya koymaya elverişli bir haldedir. Kayan bir pencere ile sınıflandırma yapmak ise bir diğer açıdan bakıldığında sürekli bir sınıflandırma yüküne sebebiyet vermekte ve bu da gerçek zamanlı olacak bir sistemi yavaşlatabilir olarak düşünülmektedir. Önerilen sistem bu açıdan çok hızlı bir sınıflandırıcı metodu olan k-en yakın komşuluk ile tasarlanmış ve bu sınıflandırıcıyı daha da hızlandırabilmek için Manhattan uzaklığı ifadesi uzaklık hesabında kullanılmıştır.

Çizelge 2. 3 Zamanda kayan pencere yardımı ile sınıflandırma

---

### Zamanda Kayan Pencere Yardımı ile Sınıflandırma

---

- 1 :  $S = (t-n+1) * n * 81$  boyutunda sıralanmış yönlü gradyanlar
  - 2 :  $t =$  girdi çerçeve uzunluğu
  - 3 :  $n =$  zamandaki pencere büyüklüğü
  - 4 :  $R = []$  sınıflandırma sonuçları, ilk değeri olmayan,  $(t-n+1)$  boyutunda
-

Çizelge 2. 3 (devam)

---

```
5 :  
6 : for i=1 to t-n+1  
7 :   etiket = S(t-n+1)[(1:n)(1:81)] 'in k-NN sonucu  
8 :   R[i] = etiket  
9 :  
10: return mode(R)
```

---

Hareket sonuç olarak oylar bütünlüğünün olduğu sınıf olarak belirlenmiştir. Düzenli ve istikrarlı oylamamanın görüldüğü durumlar kendilerini içten içe doğruladığından çok güvenilir bir hareket tanıma sonucunu ortaya koymaktadırlar. Fakat aynı kısa video içerisinde farklı hareket sınıfları ile etiketlenme içeren durumlar tanıma güvenilirliği az bir vakayı ifade etmektedir. Bu durum videonun uzun yahut kısa oluşunun harekete nasıl bir etkisinin olduğu sorusunu ortaya çıkarmıştır. Sistemin gürbüzlüğünün ispatı açısından ve bu durumun daha iyi analizini sağlayabilmek için bütün bir hareket ile etiketlenen videoların daha sonra küçük parçalara da bölünüp tanınması da bu çalışmada değerlendirilmiştir.

### HAREKET ÖRÜNTÜLERİNİN TANINABİLİRLİĞİ

İnsan hareketlerinin ortak bir örüntüsü olduğu ve bu örüntülerin tanınabilirliği esasından yola çıkılarak zaman, hareket ve ayrıca durgun haldeki bilginin derinlemesine incelenmesi ile video görüntülerinin insan hareketleri açısından yorumlanması sağlanabilir. Böylece hareket tanıyabilen bir sistemin varlığından bahsedilebilir.

Hareket tanıyabilen bir sistemin varlığının ortaya konulması gerçekçi bir hareket tanımlayıcı sistemin ilk adımı olarak da düşünülebilir. Bu sistemin gerçek şartlar altında çalışabilir olduğunu ispatlamak için ise ayrıntılı bir karşılaşılan durum – verilen tepki analizi yapılması gerekmektedir. Bu da ancak farklı senaryolara sistemin nasıl veya hangi derecede başarıyla cevaplar verdiği incelenerek anlaşılabilir.

#### 3.1 Basit Hareketlerin Tanınması

İnsan hareketi tanıyabilen bir sistemin ilk test unsuru elbette belirlenmiş koşullar altında, ideal şartlarda yapılan belirli adette hareketin tanınmasıdır. Bu durum için literatürde hareket tanıma gerçekleyen araştırmacıların sıklıkla kullandığı bir veri seti olan Weizmann veri seti kullanılmıştır [9]. Weizmann hareket tanıma veri setinin ortak kullanılan bir veri seti olması bu tez çalışmasında kullanılmasının esas nedenidir. Çünkü sistemin gerçekten genellenebilir bir yapıda olmasını ispatlayabilmek için ortak kullanılan bir veri setinde test edilmesi çok önemlidir. Özele indirgenmemiş genel bir veri seti herhangi bir sistemin test aşamasının da ana mekanizmasını oluşturmada önemlidir. Weizmann veri seti de bu açıdan bakıldığında ideal bir veri setidir.

### 3.1.1 Weizmann Hareket Tanıma Veri Seti

Weizmann hareket tanıma veri seti yaygın olarak kullanılan veri setlerinden birisidir. Durağan bir alanda çekilen insan hareketleri içeren görüntülerden oluşmaktadır. Sabit bir arka plan içerdiğinden ilk test için ideal bir ortam oluşturmaktadır. Ortam hareketlerinden bağımsız, ilgi alanı çıkarımı kolay bir veri seti basit insan hareketlerinin tanımlanması ve örüntülerin elde edimi için uygun bir ortam oluşturmaktadır.

Weizmann veri setinde 9 farklı insanın yapmış olduğu yürüme, koşma, zıplama, eğilme, yan yürüme, tek el sallama, iki el sallama, yerinde zıplama, elleri açarak sıçramak(jumping jack) ve seksek hareketinden oluşan 10 adet hareket bulunmaktadır.



Şekil 3. 1 Hareket tanımada sıklıkla kullanılan veri setlerinden biri: Weizmann hareket tanıma veri seti

Farklı insanlar tarafından yapılan farklı hareketler sabit bir arka plan ile birlikte hareketlerin insanlar tarafından yapılma farklılığının en iyi ölçülebileceği ortamı sağlamaktadır. Böylece farklı insanlar tarafından yapılan aynı sınıfta birleştirilmesi gereken hareketler acaba gerçekten birleştirilebilir mi sorusuna da cevap bulunabilecektir. Dış etkilerin minimum olması ise ilk aşamada sistemin gerçekten farklı insanların yaptığı aynı hareketleri aynı sınıfa toplayabileceğinin testi için uygun bir düzendir.

### 3.1.2 Sistem Testi - Weizmann Hareket Tanıma Veri Seti Üzerinde Bütün Video Girdisinin Deneysel Sonuçları

Sistemin ilk testi için video girdisi olarak Weizmann hareket tanıma veri seti kullanılmıştır. Weizmann hareket tanıma veri setinin içerdiği 9 insan ve 10 adet videodan oluşan toplam 90 video girdisi tek insan dışarıda bırak (leave-one-person-out) çapraz geçerlilik sınaması ile sisteme girdi olarak uygulanmıştır. Bu uygulama şekli diğer çalışmaların da kullandığı standart model olduğundan[30], sistemin diğer çalışmalara başarıları karşılaştırabilmek amacıyla kullanılan sınama yöntemi olmuştur. Böylece her seferinde tek insanın yapmış olduğu 10 adet hareket test olarak seçilmiş ve diğer insanların yapmış olduğu hareketler eğitim seti olarak kullanılmıştır. Toplamda 10 hareket 9 kere çaprazlanarak 90 adet test videosu oluşturulmuştur. Değişken zaman aralıklarına sahip bu videolarda videonun bir bütün olarak oylanması yapılmıştır ve hareketlerin tanınabilirliği ölçülmüştür.

Test videolarının bir bütün olarak sisteme verilmesi ile hareket nedir sorusunun cevabının tahmini üzerine hareket belirleme testi yapılmıştır. Daha sonra ayrıntılı olarak denenecek olan zaman boyutundaki n değeri, hareket uzunluğu en iyi sonucu verdiği empirik olarak gözlemlenen 10 alınmış, basitlik ve hız açısından Manhattan uzaklığı kullanılarak en yakın komşuluk bulunmuştur. Sistem bu sınıflandırmada Çizelge 3.1' de görülebileceği gibi yalnızca 3 hatalı sınıflandırma yapmış ve başarı oranının %96.67 olduğu ortaya konulmuştur. Bu sonuçlar ile sunulan sistem hareket tanıma konusunda günümüzde en başarılı olan çok daha karmaşık sistemlere eşit bir başarı ortaya koymuştur (Çizelge 3.2) [30].

Çizelge 3. 1 Weizmann hareket tanıma veri seti üzerinde bütün video girdisiyle hareket tanıma hata matrisi

Gerçek / Tahmin	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<b>(1) Eğilme</b>	9	0	0	0	0	0	0	0	0	0
<b>(2) Zıplama Eller</b>	0	9	0	0	0	0	0	0	0	0
<b>(3) Zıplama</b>	0	0	9	0	0	0	0	0	0	0
<b>(4) Dik Zıplama</b>	0	0	0	9	0	0	0	0	0	0



Çizelge 3. 1 (devam)

(5) Koşma	0	0	0	0	8	0	0	1	0	0
(6) Yan Yürüme	0	0	0	0	0	9	0	0	0	0
(7) Seksek	0	0	1	0	1	0	7	0	0	0
(8) Yürüme	0	0	0	0	0	0	0	9	0	0
(9) Tek El	0	0	0	0	0	0	0	0	9	0
(10) İki El	0	0	0	0	0	0	0	0	0	9

Çizelge 3. 2 Weizmann hareket tanıma veri seti üzerinde günümüz modellerinin başarı sonuçları

Yöntem	Başarı
STS(Spatio-temporal descriptors approach – Uzam-zaman betimleyicileri yaklaşımı)	<b>%96.67</b>
BoW (Bags-of-visual-Words – Görsel kelime çantası) hard voting(kesin-belirli oylama)	%86.11
BoW soft voting(belirsizlik hesabı ile oylama)	%90.00
SBoW(Spatially restricted BoW – Uzam kısıtlı görsel kelime çantası)	%85.00
Önerilen sistem	<b>%96.67</b>

### 3.1.3 Sistem Testi - Weizmann Hareket Tanıma Veri Seti Üzerinde Kısaltılmış Video Girdilerinin Deneysel Sonuçları

Sistemin gerçek yaşama uygulanabilirliğinin görülebilmesi açısından, ikinci bir test olarak, Weizmann hareket tanıma veri setindeki videolar n adetten oluşan çekim

parçalarına bölünmüştür. n=10 alınarak 4623 adet kısa video parçaları oluşturulmuştur. Böylece yalnızca 10 çerçeve ile hareket tanımlanabilir mi sorusunun cevabı aranmıştır. Sistemin bu koşullar altında da Çizelge 3.3’de görüldüğü üzere çok başarılı sonuçlar verdiği ortaya konulmuştur.

Çizelge 3. 3 Weizmann hareket tanıma veri seti üzerinde kısaltılmış video girdisiyle hareket tanıma hata matrisi

Gerçek / Tahmin	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
<b>(1) Eğilme</b>	519	0	0	0	0	0	0	0	30	0
<b>(2) Zıplama Eller</b>	5	626	0	8	0	0	0	0	0	0
<b>(3) Zıplama</b>	0	0	305	0	1	7	55	0	0	0
<b>(4) Dik Zıplama</b>	0	2	0	446	0	0	0	0	0	0
<b>(5) Koşma</b>	0	0	0	0	250	1	5	28	0	0
<b>(6) Yan Yürüme</b>	0	0	0	0	0	342	0	7	0	0
<b>(7) Seksek</b>	0	0	42	0	102	0	189	11	0	0
<b>(8) Yürüme</b>	0	0	0	0	0	0	0	543	0	0
<b>(9) Tek El</b>	5	0	0	0	0	0	0	0	556	2
<b>(10) İki El</b>	0	0	0	0	0	0	0	0	13	521

Sınıflandırma sonuçları incelediğinde 10'luk çerçeveleri olarak kısaltılmış Weizmann hareket tanıma veri seti videolarında %92.95'lik bir başarı yüzdesi olduğu görülmektedir. Basit ve efektif metotlar kullanan, gerçek zamanda çalışabilir bir sistem için bu başarı oranının oldukça başarılı olduğu düşünülmektedir.

10' luk çerçevelere bölünen videolarda değerlendirilen çerçeve sayısının az olmasının zorlukları ortaya çıkmaktadır. İnsan gözünün de algılamada zorlanacağı durumlar ortaya çıkmakta ve sistem başarısızlıklarının aslında bu sebeplerden kaynaklandığı anlaşılmaktadır.

### 3.1.4 Weizmann Hareket Tanıma Veri Seti Üzerindeki Sistem Testi Sonuçlarının İncelenmesi

Sistemin Weizmann hareket tanıma veri seti ile testinden çeşitli çıkarımlar incelenebilir hale gelmiştir. Weizmann videolarının bir bütün olarak testi olan 90 adet videonun tek insan dışarıda bırak çapraz geçerlilik sınaması incelendiğinde (Çizelge 3.1) sadece 3 hatanın olduğu gözlemlenmektedir. Bu hatalara yakından bakıldığında:

- koşma hareketi yürüme,
- seksek hareketi koşma,
- seksek hareketi ellerle beraber zıplama,

olarak tanındığı görülmüştür.



Şekil 3.2 Bütün video girdisi hatalı örneği - seksek hareketinin koşma olarak tanınması

Bütün video girdisi ile ölçülen ilk sistem tepkisi sonucunda elde edilen 3 hatanın aslında gözle dahi izah edilebilir bir halde olmadığı anlaşılmaktadır. Yapılan hareket insandan insana o kadar değişiklik göstermektedir ki Şekil 3.2'nin etiketsiz verildiği düşünülürse hangi hareketin hangi parçası olduğunu söyleyebilmek insan tarafından dahi gerçekleştirilebilse bile gerçekten çok zordur. Bu da aslında göstermektedir ki sistem ilk testten çok yüksek başarı oranına sahip, hatalı çıkarımları mantıklı izah edilebilir sonuçlar ortaya koymuştur.

Video girdisinin bir bütün olarak değerlendirilmesi elbette gerçek zamanlı, gerçekçi bir sistemin geçerli bir not alabilmesi için yeterli değildir. Bu sebeple 10 çerçeve uzunluklu bölümlere ayrılmış kısaltılmış videolar üzerinde yapılan test sistemin çalışmasının esasları hakkında daha net bir bilgi verecektir. %92.95'lik bir başarı elde edilen bu adımda değerlendirilen video parçasının kısalığının hareketin anlamlandırılabilmesine etkisi de gözler önüne serilmiştir. Beklendiği üzere videonun tamamının değerlendirilmesine göre daha düşük başarı elde edilen bu testte hareketin videonun

hangi kesilme noktasına geldiğinin belirsizliği ile beraber hareketlerin birbirine benzerliği, hangi hareketin ne olduğunun anlaşılmasının zorluğu da artmıştır. Fakat gerçek zamanda çalışabilecek bir sistemin bu testte de başarılı olması gerekmektedir.



Şekil 3. 3 Kısaltılmış video girdisi hatalı örnekler - zıplama hareketinin seksek olarak tanınması

Sistem hataları incelendiğinde aslında hata olarak görülen durumlar, gerçekten hareket algılamanın uç kısımlarını oluşturmakta, göz ile ayırt edilmesi gerçek dışı olan kısa durumları belirtmektedir.



Şekil 3. 4 Kısaltılmış video girdisi hatalı örnekler – yerinde (dik) zıplama hareketinin elleri açarak zıplama (jumping jack) olarak tanınması



Şekil 3. 5 Kısaltılmış video girdisi hatalı örnekler – yan yürüme hareketinin yürüme hareketi olarak tanınması

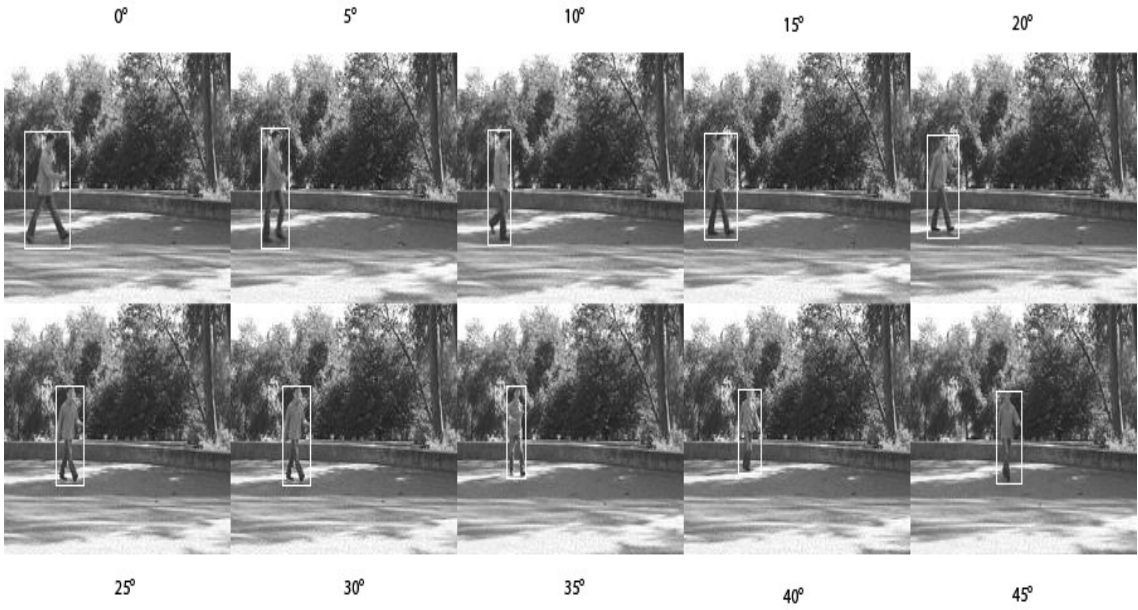
Bakış açısının önemi de başarıya çok büyük bir etki etmektedir. Şekil 3.6'da bu durum açıkça gözlenebilmektedir. Karşıdan çekilen bir görüntüde eğilme ve tek el sallama hareketinin ilgi alanına göreceli konumu birbirine benzeyebilmekte ve hareketli kenarların kenar özellikleri dahilinde işlem yapan sistem bu durumu yanlış yorumlayabilmekte ve bu da dolayısıyla başarıya negatif etki etmektedir.



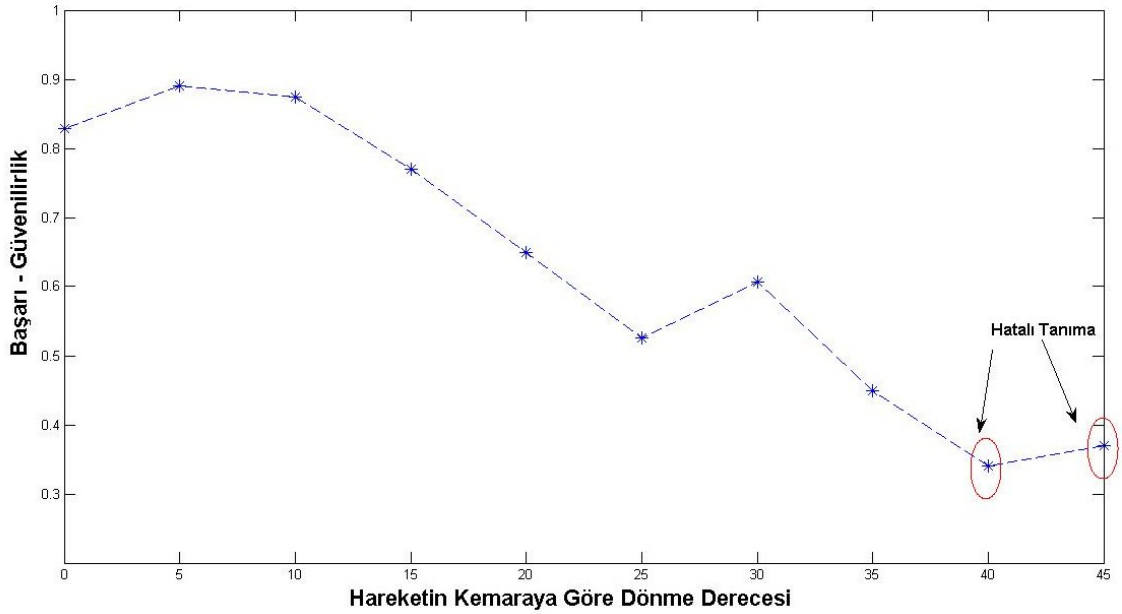
Şekil 3. 6 Kısaltılmış video girdisi hatalı örnekler – eğilme hareketinin tek el sallama hareketi olarak tanınması

### 3.2 Görüntü Açısının Tanıyıcı Sistem Üzerine Etkisi

Bilgisayarla görü çalışmalarının tek bir kayıt cihazı ile alınan görüntülerden yapılması açılım kavramının çok önemli bir dış etken olmasına neden olmaktadır. Weizmann hareket tanıma veri seti de tek yönden çekilmiş video görüntülerinden oluşmaktadır. Bu durum bir taraftan bakıldığında mükemmel hareket tanıyabilen bir sistemin bir başka açıya geçildiğinde hareketi tanıyamaz olmasına sebep olmaktadır. Bu düşünceden yola çıkılarak sistem değişik açılardan alınan görüntülerle test edilmiş ve değişen görüntü alma derecelerine tepkisi incelenmiştir.



Şekil 3. 7 Hareket dereceleri testi girdisi



Şekil 3. 8 Hareketin kameraya göre yapılış derecesinin hareket tanıma etkisi

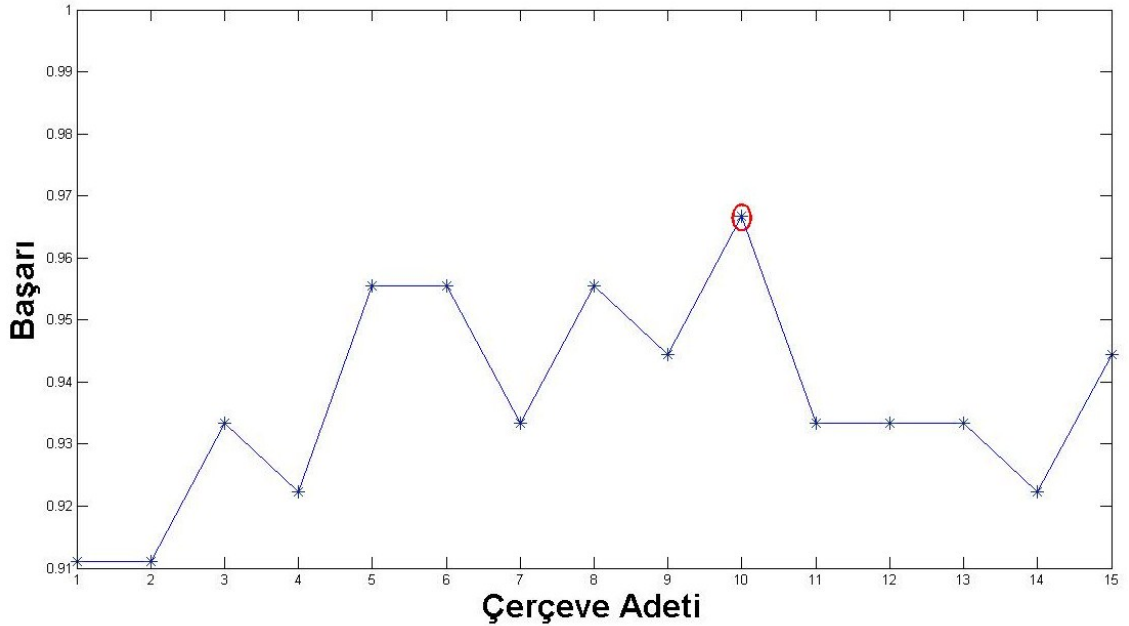
Hareketin kameraya göre yapılış derecesinin göreceli olarak değişmesi tahmin edildiği üzere hareketin tanınmasına negatif etki etmektedir. Weizmann hareket tanıma veri seti güvenilirlik testi olarak uygulanması tavsiye edilen, 0 dereceden başlayıp 45 derece dönme açısı uygulanana kadar 5'er derecelik aralıklarla yürüme hareketi yapılan videolar bu adımda test videosu olarak kullanılmıştır (Şekil 3.7). Böylece 10 farklı yürüme hareketi oluşturulmuştur. Küçük derecelerde hareket tanıma büyük bir güvenilirlik sağlamasına karşın artan derece değerleri ile beraber öncelikle bu güven düşmüş sonra ise hareket tanınamayacak hale gelmiştir (Şekil 3.8). Bu durumda elbette eğitimi veri setinin de önemi büyüktür. Daha önce hep kameraya 0 derece açı ile yapılan hareketlerle eğitilmiş bir veri setinin, açı farkı arttıkça tanıma başarısının azalması zaten beklenen bir durumdur. Bu durumun düzeltilebilmesi ise eğitim setinin değişik açılardan çekilmiş videolarla desteklenmesi ile çok kolay bir şekilde mümkündür.

### 3.3 İnsan Hareketlerini Tanımak İçin Ne Kadar Süre Gereklidir?

İnsan hareketlerinin zaman boyutunda da bilgi barındıran, sürekli bir bilgi olduğu açıktır. Bu tez çalışmasında hareket bilgisi, durgun görüntünün bilgisi ve yanında zamanda ilerleyen bilgilerin sıra dizilimi ile beraber görüntüden tanımlanıp algılanabilecek bütün bilgilerin -bu üç bilginin- beraberce kullanıldığı bir sistem

tasarlanmıştır. Ama bir diğer açıdan bakacak olursak zaman kavramı aslında göreceli bir kavramdır. Hareket zaman düzleminde değişken bir yapıya sahiptir. Bu bakış açısı hareket kavramının algılanmasının değişken zaman dilimlerinde değişken bir hal alacağını göstermektedir. Bu da aslında çıkarım olarak insan hareketlerini tanımak için ne kadar bir süre veya bir diğer dolaylı ifadeyle kaç çerçeve gerekmektedir sorusunu beraberinde getirmektedir.

Hareket algılamanın nasıl bir süreç olduğu aslında başlı başına büyük bir araştırma sürecini oluşturmaktadır. Ne kadar zaman sorusu bu kavramın göreceliğinin genellenebilirliğine giden evresinde cevap verilmesi gereken esas sorulardan biridir. Bu tez çalışmasında da çeşitli çerçeve adetlerinde hareketin algılanabilir olup olmadığı test edilmiş ve zaman kavramının başarıya etkisi incelenmiştir. Böylece hareketin algılanması ve tanınması için gereken zaman kavramı ne olmalıdır yorumunun yapılması da sağlanmıştır. Bahsedilen zaman kavramının algılanabilirliğe etkisini gösterebilmek için sistem birçok çerçeve boyutunda hareketi tanıma üzerine test edilmiş ve Şekil 3.8’de sonuçlar sunulmuştur.



Şekil 3. 9 Hareket tanıma için ardışık çerçeve sayısına göre başarı değişimi

Değişen çerçeve adetine göre sistemin başarısı incelendiğinde 10 çerçevede en iyi sonucu verdiği görülmüştür. Weizmann hareket tanıma veri setinin saniyede 25

çerçeve içeren görüntüler olduğu göz önüne alındığında 0,4 saniye gibi bir sürede sistem en yüksek başarıyı sağlamaktadır. Sonuçlar daha yakından incelendiğinde ise aslında hareket tanımının insan pozisyonuna ne kadar bağlı olduğu gözlemlenmektedir. Çünkü daha ilk çerçeve bile hareketin ne olacağı hakkında çok yüksek bir bilgi içermektedir. Sistem detaylı incelendiğinde ise 5 çerçevenin yeterli bir başarı düzeyinde çok daha hızlı bir sistem için kullanıma ideal olduğu söylenebilir. Bu da aslında sistemin 0,2 saniyede hareketi algıladığını ortaya koymaktadır. Elbette bu sonuçlar belirli hareketler içerisinde yapılan bir hareketin belirlenmesinde ifade edilmektedir. Aksi düşünülüğünde insan hareketi tanımlanamayacak kadar çoktur ve ifade edilmesi bile ayrı bir zorluktur.

### **3.4 Değişken Koşulların Altında Sistemin İncelenmesi**

Gürbüzlük günümüz sistemlerinin temel esaslarından biridir. Sistemin gerçekçiliğinin ortaya konulabilmesi için değişik durumlar altında sistemin verdiği tepkilerin gerçekçi olup olmadığı ve sistemin zorlu koşullar altında başarıya ulaşıp ulaşılmadığı gözlemlenmelidir. Bu esastan yola çıkılarak sistem çeşitli durumlar içeren senaryolarla denenmiş ve incelenmiştir.

#### **3.4.1 Hareketin Tam Olarak Görünmediği Durumların Değerlendirilmesi**

Sistemin değişken koşul incelemeleri testinde ilk olarak hareketin görünmesini engelleyen objelerle sistem test edilmiştir. Şekil 3.10 test videolarının genel bir görüntüsünü vermektedir. Gerçekçi hayat örnekleri çerçevesinde çanta taşıma, evrak çantası taşıma, köpek ile birlikte yürüme, ayak kısmını engelleyen kutular arkasında yürüme, hareketin belli bir müddet görülmesini engelleyen direk arkasında yürüme, hareketin görünürlüğünü engelleyen etek bağlı yürüme (Şekil 3.10), gibi çeşitli kısa videolar sistem testinde kullanılmıştır.





a) Çanta taşıyan hareketli



b) Evrak çantası taşıyan hareketli



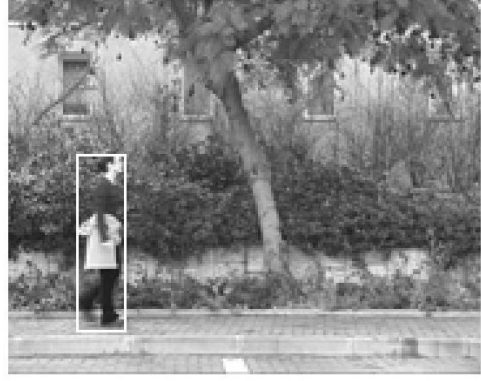
c) Köpek ile yürüyen hareketli



d) Ayak kısmının görünmesini engelleyen kutular



e) Hareketin belli bir müddet bütünüyle engellenmesi



f) Hareketin görünmesini engelleyici etek bağlı yürüme

Şekil 3. 10 Hareketin görünürlüğünü engelleyen objeler testi girdileri

Hareketin belli bir kısmının görünürlüğünü etkileyen objeler sistemin hareket tanınabilirliği güvenilirliğini belli bir oranda düşürse sistem hareketleri tanıyabilmiştir. Sistem oy çoğunluğunu sağlayamadığı durumlarda kararsız kalmıştır, bu durum çizelgelerde de kararsız olarak ifade edilmiştir. Çizelge 3.4'te tanıma sonuçları görülebilmektedir.

Çizelge 3. 4 Hareketin görünürlüğünü engelleyen obje testi sonuçları

Test Girdisi	Tutarlılık	Tanıma Sonucu
Çanta taşıyan hareketli	%59.18	Yürüme
Evrak çantası taşıyan hareketli	%81.25	Yürüme
Köpek ile yürüyen hareketli	%100	Yürüme
Ayak kısmının görünmesini engelleyen kutular	%83.67	Yürüme
Hareketin belli bir müddet bütünüyle engellenmesi	%45.1	Kararsız
Hareketi engelleyici etek bağlı yürüme	%43.9	Kararsız

### 3.4.2 Düzgün Yapılmayan Hareketler Testi

Düzgün yapılmayan yürüme hareketlerini esas alan bu bölümde çeşitli anormal yürüme şekilleri ele alınmıştır. 'Yürüme hareketine ne kadar benzeyen harekete yürüme denir?' sorusundan yola çıkılarak yapılan bu testlerde aslen yürüme hareketinden farklı denilebilecek hareketler içeren videolar değerlendirilmiştir. Şekil 3.11 test videolarının genel bir görüntüsünü vermektedir. Dizler yukarıya çekilerek yürüme, topallayarak yürüme, uyurgezer bir şekilde yürüme (Şekil 3.11) gibi kısa videolarla sistem test edilmiş ve sistemin hareket ayırt ediciliği ölçülmüştür. Sistem daha önce sadece Weizmann veri setinde bulunan 10 adet hareket ile eğitildiğinden, uyurgezer yahut

topallayarak yahut dizler yukarı çekilerek yürüme girdilerine yürüme mi diyeceği yoksa başka bir hareket olarak mı algılanacağı bu durum içerisinde gözlemlenebilecektir.



a) Dizler yukarıda yürüme



b) Topallayarak yürüme



c) Uyurgezer yürüme

Şekil 3. 11 Düzgün yapılmayan hareketler testi girdileri

Çizelge 3. 5 Düzgün yapılmayan hareketler testi sonuçları

Test Girdisi	Tutarlılık	Tanım Sonucu
Dizler yukarıda yürüme	%40.86	Kararsız
Topallayarak yürüme	%62.5	Yürüme
Uyurgezer yürüme	%38.18	Kararsız

Sistemin çeşitli düzgün yahut sıradan diyeceğimiz hareketler ile yapılan test sonuçları Çizelge 3.5'te görülmektedir. Bu çizelgeden de anlaşılacağı üzere sistem belli bir oranda bu hareketlere yürüme diyebilmektedir. Fakat çoğu zaman bu durum hareketin gerçekten ne olduğuna karar verme konusunda yeterli tutarlılığı ve dolayısıyla güveni oluşturmamaktadır.

### SONUÇ VE ÖNERİLER

Bu tez çalışmasında, şekil kenar bilgisinin hareket bilgisi ile birleşimi ile hareket algılama konusunda yüksek başarılar elde edilebileceği ortaya konulmuştur. Hareket bilgisinin veya sabit görüntünün bilgisinin tek başına kullanılmasındansa, bu bilgilerin birleştirilmesi daha başarılı sonuçlara ulaşmayı sağlamıştır.

Sistem, hareket eden parçaların kenar özelliklerini kullanarak boyuttan bağımsız hareket algılayabilen bir yapı oluşturmuş ve en yakın komşuluk gibi basit bir sınıflandırma yöntemiyle dahi çok başarılı sonuçlar ortaya koyabileceğini göstermiştir.

İnsanı içeren anlamlı bölgeye optik akış yöntemi uygulanması ile hareketi ifade eden bir tanımlayıcı haline dönüştürülmüş ve bu tanımlayıcının kenar özellikleri zaman düzleminde incelenmiştir. Bunu yaparken optik akış yöntemi klasik kullanımının dışında bir filtre olarak da düşünülebilenecek bir şekilde kullanılmış ve böylece hareket algılamayı sağlayan özgün bir sistem tasarlanmıştır.

Sistemin basitlik ve etkinlik olarak gerçek zamanda çalışmaya uygun olduğu gözlemlenmiştir. Bu durum, gerçek zamanda hareket algılama tabanlı bir karar mekanizması oluşturmayı sağlayabilecek bir yapı ortaya koyabilmenin ön adımı olarak düşünülebilir. Böyle bir sistemin gerçekleşmesiyle ise çok çeşitli kullanım alanlarında başarıyla uygulanabilir bir yapı ortaya konulması öngörülmektedir.

Önerilen hareket tanıma sistemi sayesinde insan-bilgisayar etkileşimi üst düzeye çıkarılabilecek ve böylece yeni nesil uygulamalar ortaya konulabilecektir. Aynı zamanda görsel gözetim alanında ilerleme sağlayacak bu yapı, bu alandaki zorlukları ortadan kaldıracaktır. Gürbüz bir hareket tanıma sistemi video erişim uygulamalarında çok

büyük kolaylıklar sağlayabilecek ve günümüz dünyasında önlenemez bir hızda artan bu alandan bilgi ve anlam çıkarımını mümkün hale getirebilecektir.

Hareket tanıma yalnız kendi içerisinde bir uygulamaymış gibi düşünülse de günümüz koşulları içerisinde kamera ve kayıt cihazlarının artması ile birlikte artık 'an' içerisine girmiş bir konumdadır. Anlık kayıtların analizi ve anlık kayıtların bilgiye dönüşümü için esas adımlardan biri olan hareket tanıma bu tez çalışmasıyla birlikte basit bir yapıda ortaya konulmuş, gerçekleşmiş ve başarımı test edilmiştir. Böylece gelecek çalışmalara pozitif bir yol açılmış ve sistemin uygunluğu ve yapılabilirliği ortaya konulmuştur.

## KAYNAKLAR

---

- [1] Moeslund, T.B., Hilton, A. ve V. Kruger, (2006). "A survey of advances in vision-based human motion capture and analysis", *Computer Vision and Image Understanding*, 104(2-3):90-126.
- [2] Poppe, R., (2010). "A survey on vision-based human action Recognition", *Image and Vision Computing*, 28(6):976-990.
- [3] Bobick A.F. ve Davis J.W., (2001). "The recognition of human movement using temporal templates", *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 23(3):257-267.
- [4] Wang, Y., Huang, K. ve Tan, T., (2007). "Human activity recognition based on R transform", *Computer Vision and Pattern Recognition (CVPR'07)*, June 2007, Minneapolis.
- [5] Souvenir, R. ve Babbs, J., (2008). "Learning the viewpoint manifold for action Recognition", *Computer Vision and Pattern Recognition (CVPR'08)*, June 2008, Anchorage.
- [6] Weinland, D., Ronfard, R. ve Boyer, E., (2006). "Free viewpoint action recognition using motion history volumes", *Computer Vision and Image Understanding*, 104(2-3):249-257.
- [7] Weinland, D. ve Boyer, E., (2008). "Action recognition using exemplar-based embedding", *Computer Vision and Pattern Recognition 2008*, 23-28 June 2008, Anchorage.
- [8] Yilmaz, A. ve Shah, M., (2005). "Actions sketch: A novel action representation", *Computer Vision and Pattern Recognition, IEEE Computer Society Conference on*, 20-26 June 2005, San Diego.
- [9] Gorelick, L., Blank, M., Shechtman, E., Irani, M. ve Basri, R., (2007). "Actions as Space-Time shapes", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(12):2247-2253.
- [10] Lu, W-L., Little J.J., (2006). "Simultaneous tracking and action recognition using the PCA-HOG descriptor", *Canadian Conference on Computer and Robot Vision (CRV'06)*, June 2006, Quebec City.
- [11] Laptev, I., (2005). "On Space-Time interest points", *International Journal of Computer Vision*, 64(2):107-123.

- [12] Shechtman, E. ve Irani, M., (2005). "Space-time behavior based correlation", Computer Vision and Pattern Recognition, IEEE Computer Society Conference on, 20-26 June 2005, San Diego.
- [13] Dollar, P., Rabaud, V., Cottrell, G. ve Belongie, S., (2005) "Behavior recognition via sparse spatio-temporal features", Visual Surveillance and Performance Evaluation of Tracking and Surveillance, Oct. 2005, Beijing.
- [14] Niebles, J., Wang, H. ve Fei-Fei, L., (2008). "Unsupervised learning of human action categories using Spatial-Temporal words", International Journal of Computer Vision, 79(3):299-318.
- [15] Gilbert, A., Illingworth, J. ve Bowden, R., (2008). "Scale invariant action recognition using compound features mined from dense spatio-temporal corners", ECCV '08: European Conference on Computer Vision, 2008, Berlin, 222-233.
- [16] Laptev, I., Marszałek, M., Schmid, C. ve Rozenfeld, B., (2008). "Learning realistic human actions from movies", Computer Vision and Pattern Recognition 2008, June 2008, Anchorage.
- [17] Kläser, A., Marszałek, M. ve Schmid, C., (2008). "A spatio-temporal descriptor based on 3d-gradients", British Machine Vision Conference (BMVC'08), September 2008, Leeds, United Kingdom.
- [18] Wang, H., Ullah M.M., Kläser A., Laptev, I. ve Schmid, C., (2009). "Evaluation of local spatio-temporal features for action Recognition", British Machine Vision Conference (BMVC'09), 2009, London, United Kingdom.
- [19] Laptev, I. ve Lindeberg, T., (2003). "Space-time interest points", International Conference on Computer Vision (ICCV'03), October 2003, Nice, France.
- [20] Willems, G., Tuytelaars, T. ve Van Gool, L.J., (2008). "An efficient dense and scaleinvariant spatio-temporal interest point detector", European Conference on Computer Vision (ECCV'08), October 2008, Marseille, France.
- [21] Schuldt, C., Laptev, I. ve Caputo, B., (2004). "Recognizing human actions: a local SVM approach", ICPR 2004, 23-26 Aug. 2004, Cambridge.
- [22] Jia, K. ve Yeung, D-Y., (2008). "Human action recognition using Local Spatio-Temporal Discriminant Embedding", Computer Vision and Pattern Recognition 2008, 23-28 June 2008, Anchorage.
- [23] Efros, A.A., Berg, A.C., Mori, G. ve Malik, J., (2003). "Recognizing action at a distance", ICCV 2003: 9th International Conference on Computer Vision, April 2003, Los Alamitos, CA, USA.
- [24] Danafar S. ve Gheissari N., (2007). "Action recognition for surveillance applications using optic flow and SVM", Asian conference on Computer vision (ACCV'07), 2007, Tokyo, Japan.
- [25] Thureau, C. ve Hlaváč, V., (2008). "Pose primitive based human action recognition in videos or still images", Conference on Computer Vision and Pattern Recognition (CVPR'08), June 2008, Anchorage, AK.



- [26] Horn, B.K.P. ve Schunck, B.G., (1981). "Determining optical flow", *Artificial Intelligence*, 17:185-203.
- [27] Dalal, N. ve Triggs, B., (2005). "Histograms of Oriented Gradients for Human Detection," *IEEE Conf. Computer Vision and Pattern Recognition*, June 2005.
- [28] Lowe D.G., (1999). "Object Recognition from Local Scale-Invariant Features", *Proceedings of the International Conference on Computer Vision*, September 1999.
- [29] Bay H., Ess A., Tuytelaars T. ve Van Gool L., (2008). "SURF: Speeded Up Robust Features", *Computer Vision and Image Understanding (CVIU)*, June 2008.
- [30] de Campos, T., Barnard, M., Mikolajczyk, K., Kittler, J., Yan, F., Christmas, W. ve Windridge, D., (2011). "An evaluation of bags-of-words and spatio-temporal shapes for action Recognition", *Applications of Computer Vision (WACV)*, 2011 IEEE Workshop, 5-7 Jan. 2011.

**“20. IEEE SİNYAL İŞLEME VE İLETİŞİM UYGULAMALARI” KURULTAYINDA  
NİSAN 2012 TARİHİNDE SUNULAN “ARDIŞIK GRADYANLAR HİSTOGRAMI  
İLE İNSAN HAREKETLERİNİN TANINMASI” İSİMLİ BİLDİRİNİN ÖZET  
SAYFASI**

# ARDIŞIK GRADYANLAR HİSTOGRAMI İLE İNSAN HAREKETLERİNİN TANINMASI

## HUMAN ACTION RECOGNITION WITH SEQUENTIAL GRADIENT HISTOGRAMS

Adem Güçlü, M. Elif Karşılıgil

Bilgisayar Mühendisliği Bölümü  
Yıldız Teknik Üniversitesi

{adem, elif}@ce.yildiz.edu.tr

### ÖZETÇE

*İnsan hareketlerinin algılanması ve yorumlanması video anlamlandırma çalışmalarının önemli bir parçasını oluşturmaktadır. Bu çalışmada, optik akıştan elde edilen hareket şiddeti bilgisinin oluşturduğu kenar özellikleri ardışık gradyanlar histogramı ile ifade edilerek yapılan hareketin ne olduğuna karar veren özgün bir sistem geliştirilmiştir. Sunulan sistem hareket tanıma alanında günümüzde en başarılı olan çok daha kompleks sistemlere eşit bir başarı ortaya koymuştur. Sistem alanda yaygın olarak kullanılan Weizmann veri setinde test edilmiş ve karşılaştırmalara yer verilmiştir.*

### ABSTRACT

Human action recognition and interpretation constitutes an important part of the video understanding. In this work, a novel action recognition system is developed that uses edge features obtained from optical flow power shapes which is represented as sequential gradient histograms. The presented system can achieve equal results to the complicated top action recognition systems of nowadays. The system is tested with the Weizmann dataset which is widely used in the field, and comparisons are given.

### 1. GİRİŞ

Görsel algılama ve insan hareketlerinin tanınması son yıllarda bilgisayarla görü dalının popüler çalışma alanlarından biri haline gelmiştir [1], [2]. Bu durumun altında yatan en önemli sebep; video arşivleme, görsel gözetim, insan-bilgisayar etkileşimi, jest tanıma, video düzenleme ve arama gibi uygulama alanlarındaki problemlerin henüz keşfedilmemiş yüksek potansiyele sahip çözümlerinin olmasıdır. İnsanların bireysel farklılıkları nedeniyle karşılaşılan ifade, duruş, kıyafet ve hareket yapılış şekli farklılıkları; ayrıca, kamera hareketleri, ışıklandırma varyasyonları ve perspektifin etkisi gibi dış etkenler hareket algılama konusunda büyük değişiklikler içeren, fakat ortak sınıf içerisinde toplanması gereken veri bilgisinin varlığını ortaya koymaktadır.

Hareket tanıma üzerine farklı yapılar kullanılarak çeşitli yaklaşımlar ileri sürülmüştür. Bu yaklaşımları esas farkı

oluşturan özellik çıkarımı adımlarıyla kategorize etmek mümkündür. Silüet çıkarımı literatürde sıklıkla karşılaşılabilen hareket ifade edim biçimlerinden biridir. Bu gösterimin ilk kullanımlarından biri, zamansal şablonlar çıkarılarak şekil değişiklikleri geçmişi [3] ile hareketin ifade edilmesidir. Bu gösterim daha sonraki çalışmalarda 3 boyutlu hale getirilmiştir [4]. Silüetlere bağlı olarak uzay-zaman şekilleri gösterimi [5], [6] ortaya konulmuştur. Son zamanlarda ise uzay-zaman yaklaşımlarının yerel 3 boyutlu parçalar halinde kullanılmasının başarılı sonuçlar ortaya koyduğu gözlemlenmiştir [7], [8], [9], [10], [11]. Bu gösterimler ile Destek Vektör Makineleri(SVM) [12], [13], Saklı Markov Modelleri(HMM) [14], [15] gibi çeşitli makine öğrenmesi metodları kullanılarak modellenmeye çalışılmıştır.

Videodan bilgi çıkarımı konusu çok farklı yöntemler içerirse de temelde bir hareket bilgisi ifadesinden ve bunun yanı sıra durgun görüntünün bilgisinin de aktarılmasından oluşmaktadır. Hareketli bilgi için optik akış kullanımı yaygındır [16], [17]. Durgun görüntünün ayırt edici özellik olarak kullanılmasında ise yönlü gradyanlar histogramı gibi kenar özellikleri çıkarıcıların kullanımı yaygındır [18], [19]. Bu iki bilginin beraber kullanımının ayrı kullanımına göre daha başarılı olduğu ortaya konulmuştur [20].

Bu çalışmada her iki bilginin ortaklaşa kullanılabilmesi bir sistem tasarlanmış ve gerçekleştirilmiştir. Optik akış yöntemi, klasik olarak kullanımı olan hareketin yönünü bulma amacındansa; hareket eden parçaların hareket kuvvetinin belirleyicisi olarak kullanılmıştır. Böylece hareketli bilgi ile görüntünün kenar özellikleri birleştirilebilmesine olanak sağlayan bir sistem geliştirilmiştir. Sistemde gerekli kenar özellikleri korunmuş, ilgi alanı dahilinde bilgi ifade etmeyen hareketsiz kısımlar optik akış sayesinde ortadan kaldırılmıştır. Sistemin tasarımında, optik akış üzerine uygulanan yönlü gradyanlar histogramı temel alınmış ve 3. boyut olarak düşünülebileceğimiz zaman kavramı da dahil edilip boyuttan bağımsız hareket tanımlayabilen özgün bir yapı ortaya çıkarılmıştır.

Bildirinin 2. kısmında önerilen sistem detaylarıyla anlatılmakta, 3. kısımda deneysel sonuçlar verilmekte ve 4. bölümde de sonuçlar değerlendirilmektedir.

## ÖZGEÇMİŞ

---

### KİŞİSEL BİLGİLER

**Adı Soyadı** : Adem GÜÇLÜ  
**Doğum Tarihi ve Yeri** : 11.10.1988, Pazaryolu  
**Yabancı Dili** : İngilizce  
**E-posta** : adem@ce.yildiz.edu.tr

### ÖĞRENİM DURUMU

Derece	Alan	Okul/Üniversite	Mezuniyet Yılı
Lisans	Bilgisayar Müh.	Fatih Üniversitesi	2010
Lise	Fen	Kurtuluş Lisesi	2006

### İŞ TECRÜBESİ

Yıl	Firma/Kurum	Görevi
2010 - 2012	Yıldız Teknik Üniversitesi	Araştırma Görevlisi

## **YAYINLARI**

### **Makale**

- 1.(2011) Eker O.F., Camci F., Guclu A., Yilboga H., Sevкли M., Baskan, S., A  
“Simple State-Based Prognostic Model for Railway Turnout  
Systems”, IEEE Transactions on Industrial Electronics, Vol. 58, Issue  
5, pp. 1718-1726, (2011)

### **Bildiri**

- 1.(2012) Aydemir M.S., Ergul U., Guclu A., Karşligil M.E., “Video  
Summarization Using Simple Action Patterns”, International  
Conference on Pattern Recognition (ICPR), 2012.
- 2.(2012) Güçlü A. ve Karşligil M.E., “Ardışık Gradyanlar Histogramı ile İnsan  
Hareketlerinin Tanınması”, 20. IEEE Sinyal İşleme ve İletişim  
Uygulamaları Kurultayı (SIU), 2012.
- 3.(2010) Guclu A., Yilboga H., Eker O. F., Camci F., Jennions I., “Prognostics  
with Autoregressive Moving Average for Railway Turnouts”, Annual  
Conference of Prognostics and Health Management Society, 2010.
- 4.(2010) Guclu A., Yilboga H., Eker O. F., Camci F., “Classification of Uncertain  
Data Streams Using Modified K-Nearest Neighbor Method: A Case  
Study on Railway Turnouts”, International Symposium on  
INnovations in Intelligent SysTems and Applications, 2010.
- 5.(2010) Yilboga H., Eker O.F., Guclu A., Camci F., “Failure Prediction on  
Railway Turnouts Using Time Delay Neural Networks”, IEEE  
International Conference on Computational Intelligence for  
Measurement Systems and Application, 2010.