

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**İNSAN HAREKETLERİNİN TAKİBİNDE KARŞILAŞILAN PROBLEMLERİN
ÇÖZÜMÜNE YENİ YAKLAŞIMLAR**

DOKTORA TEZİ

Yük. Müh. Muhammed Fatih TALU
(05131202)

Anabilim Dalı: Elektrik - Elektronik Mühendisliği

Tez Danışmanları: Prof. Dr. Mehmet CEBECİ
Doç. Dr. İbrahim TÜRKOĞLU

Tezin Enstitüye Verildiği Tarih: 06 Temmuz 2010

TEMMUZ - 2010

T.C.
FIRAT ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

**İNSAN HAREKETLERİNİN TAKİBİNDE KARŞILAŞILAN PROBLEMLERİN
ÇÖZÜMÜNE YENİ YAKLAŞIMLAR**

DOKTORA TEZİ

Yük. Müh. Muhammed Fatih TALU
(05131202)

Tezin Enstitüye Verildiği Tarih: 06 Temmuz 2010

Tezin Savunulduğu Tarih: 26 Temmuz 2010

Tez 2. Danışmanı: Doç. Dr. İbrahim TÜRKOĞLU (Fırat Üniv.)

Diğer Jüri Üyeleri: Prof. Dr. Bekir KARLIK (Mevlana Üniv.)

Doç. Dr. Ali KARCI (İnönü Üniv.)

Yrd. Doç. Dr. Ahmet ÇINAR (Fırat Üniv.)

Yrd. Doç. Dr. Mustafa TÜRK (Fırat Üniv.).....

TEMMUZ – 2010

ÖNSÖZ

Bu tezin hazırlanmasında, deneyimi, bilgisi ve önerileriyle araştırma ve geliştirmeyi yönlendiren tez danışmanlarım sayın Prof. Dr. Mehmet CEBECİ ve sayın Doç. Dr. İbrahim TÜRKOĞLU'na, bilgi ve tecrübesinden yararlandığım sayın Doç. Dr. Deniz ERDOĞMUŞ'a (ABD, Northeastern Üniversitesi, Elektrik ve Bilgisayar Mühendisliği), tezin anlaşılır sade bir üslup ile yazılabilmesi için kıymetli vaktini ayırıp tezi gözden geçiren sayın Dr. Resul DAŞ'a ve sayın Arş. Gör. Ramazan POLAT'a, zor çalışma sürecinde sabır ve özveri ile hep yanımda olan eşim Meryem TALU'ya, tez uygulamalarının gerçekleşmesinde finansal katkı sağlayan Fırat Üniversitesi Bilimsel Araştırmalar Projeleri Koordinasyon Birimine (FÜBAP) teşekkürlerimi sunarım.

Muhammed Fatih TALU

Elazığ - 2010

İÇİNDEKİLER

	<u>Sayfa No</u>
1. GİRİŞ	1
1.1. AMAÇ.....	2
1.2. TEZİN ORGANİZASYONU VE LİTERATÜRE KATKISI	3
2. NESNE TAKİP YÖNTEMLERİ.....	5
2.1. NESNE SUNUMU	6
2.2. NESNE ÖZELLİKLERİ VE SEÇİMİ	7
2.3. NESNE YAKALAMA	8
2.4. NESNE TAKİP YÖNTEMLERİ	8
2.4.1. Nokta Tabanlı Nesne Takibi	9
2.4.2. Çekirdek Tabanlı Nesne Takibi	17
2.4.3. Siluet Tabanlı Nesne Takibi.....	22
3. KALABALIK ORTAMLARDA ÇOKLU İNSAN TAKİP YÖNTEMİNİN GELİŞTİRİLMESİ.....	25
3.1. ÖN PLAN NESNELERİNİN TESPİTİ	26
3.2. SEVİYE KÜMELERİ YARDIMIYLA NESNE SUNUMU	28
3.3. ORTALAMA KAYMA ALGORİTMASI.....	29
3.4. VERİ BAĞI PROBLEMİ VE KİŞİ ETİKETLEME	31
3.5. KAPATMANIN YAKALANMASI.....	32
3.6. UYGULAMA SONUÇLARI VE DEĞERLENDİRİLMESİ	34
4. TEN YAKALAMA TABANLI İSTATİSTİKSEL EYLEM ALGILAMA YÖNTEMİNİN GELİŞTİRİLMESİ.....	39
4.1. ÖN PLAN NESNELERİNİN TESPİTİ	39
4.2. TEN YAKALAMA YÖNTEMİNİN KULLANILMASI	39
4.3. NESNELERİN İLİŞKİLENDİRİLMESİ	40
4.4. EYLEMİN ALGILANMASI.....	42
4.5. UYGULAMA SONUÇLARI VE DEĞERLENDİRİLMESİ	42
5. TEK KAMERA YARDIMIYLA 3-BOYUTLU İNSAN POZ TAKİP YÖNTEMİNİN GELİŞTİRİLMESİ.....	46
5.1. İNSAN POZ TAKİBİ PROBLEMİ	46

5.1.1. Tek Kameralı Poz Takip Yöntemleri	47
5.1.2. Çok Kameralı Poz Takip Yöntemleri.....	50
5.2. GELİŞTİRİLEN İNSAN POZ TAKİP YÖNTEMİ	51
5.3. MOCAP VERİTABANI.....	53
5.4. MODELLER.....	53
5.4.1. Vücut İskelet Modeli.....	54
5.4.2. Vücut Görünüm Modeli	55
5.4.3. Kamera Modeli.....	56
5.5. ÖN İŞLEME ADIMI	57
5.5.1. Ön Plan Nesnelerinin Tespiti	58
5.5.2. Ön Plan Nesneleri İçerisinden İnsan Silüetinin Çıkarımı	58
5.5.3. Başlangıç Parametrelerinin Ayarlanması	59
5.6. İSTATİSTİKSEL POZ TAKİBİ	63
5.6.1. Bayesian Takip Modeli	63
5.6.2. Durum Vektörünün Belirlenmesi.....	64
5.6.3. Öncesel Dağılımın Hesaplanması	64
5.6.4. Maksimum Olabilirlik Değerinin Hesaplanması	67
5.7. UYGULAMA SONUÇLARI VE DEĞERLENDİRİLMESİ	68
6. SONUÇ VE DEĞERLENDİRME	74
6.1. SONUÇLARIN İRDELENMESİ.....	74
6.2. ÖNERİLER	76
KAYNAKLAR.....	77
ÖZGEÇMİŞ	91

ŞEKİLLER LİSTESİ

Sayfa No

Şekil 1-1. Konularına ve seviyelerine göre bilgisayarlı görme çalışma alanının sınıflandırılması.	1
Şekil 2-1. Sık kullanılan şekilsel sunum türleri.....	6
Şekil 2-2. Sık kullanılan görsel sunum türleri.	7
Şekil 2-3. Genel nesne özellikleri.....	7
Şekil 2-4. Nesne takip yöntemlerinin sınıflandırılması.	9
Şekil 2-5. Parçacık süzgecindeki seçme adımının çalışma şekli.	14
Şekil 2-6. Farklı bant genişliğine sahip gaussian çekirdek fonksiyon grafikleri.....	19
Şekil 2-7. Bir işaretçinin görünüm profilinin elde edilmesi.	23
Şekil 2-8. Basit bir nesnenin farklı çözünürlük seviyelerindeki sunumu.	24
Şekil 2-9. Ardışık imgelerde işaretçi arama işlemi.	24
Şekil 3-1. Geliştirilen çoklu insan takip yönteminin akış diyagramı.	25
Şekil 3-2. Yakalanan hareketli insan nesneleri için seviye kümesi gösterimi.....	29
Şekil 3-3. Ortalama kayma vektörünün hareketi.....	30
Şekil 3-4. Ortalama kayma yöntemi kullanılarak kısmi kapatma probleminin giderilmesi.	32
Şekil 3-5. Geliştirilen çoklu insan takip yönteminin kısmi kapatma olayı anındaki davranışı.	34
Şekil 3-6. Kalabalık ortamda çoklu insan takip yönteminin örnek video kaydı üzerindeki sonuçları	36
Şekil 4-1. Ten yakalama tabanlı istatistiksel eylem algılama yönteminin akış diyagramı..	39
Şekil 4-2. Ten yakalama tabanlı istatistiksel eylem algılama yönteminin farklı kişiler üzerindeki uygulama sonuçları.	43
Şekil 4-3. Geliştirilen insan eylem algılama yöntemi kullanılarak ilacın ağıza götürülme eyleminin tespit edilmesi.	44
Şekil 5-1. İnsan hareket takibi yöntemlerinin sınıflandırılması	47
Şekil 5-2. Geliştirilen poz takip yönteminin akış diyagramı.....	52
Şekil 5-3. Özel elbise giydirilen örnek bir figüran resmi.....	53
Şekil 5-4. Vücut iskelet modeli ve modelin ağaç yapısının sunumu.....	54

Şekil 5-5. Görünüm modelleri.....	55
Şekil 5-6. Perspektif izdüşüm tekniği kullanılarak 3-boyutlu yapay insan modelinin 2-boyutlu imge düzlemine izdüşümünün alınması.	56
Şekil 5-7. Ön işleme adımının akış diyagramı.	58
Şekil 5-8. “ <i>asf</i> ” uzantılı örnek bir dosya içeriğinin görünümü.....	60
Şekil 5-9. İnsan vücut iskelet modelinin gösterimi.	61
Şekil 5-10. İki eklem noktası arasındaki euler açılarının gösterimi.	61
Şekil 5-11. Yapay insan modeline ait noktaların vücut iskeleti üzerine giydirilmesi.	62
Şekil 5-12. Gerçekleştirilen poz tahmin yönteminin “koşma” videosu üzerindeki sonuçları.	70
Şekil 5-13. Gerçekleştirilen poz tahmin yönteminin “zıplama” videosu üzerindeki sonuçları.	71
Şekil 5-14. Gerçekleştirilen poz tahmin yönteminin “yürüme” videosu üzerindeki sonuçları.	72

TABLÖLAR LİSTESİ

	<u>Sayfa No</u>
Tablo 2-1. Çekirdek fonksiyonlarının formülasyonları.	18
Tablo 3-1. Çoklu insan takip yönteminin farklı video kayıtları üzerindeki işlem süresi sonuçları.	37
Tablo 4-1. Ten yakalama tabanlı istatistiksel eylem algılama yönteminin takip sonuçları.	44
Tablo 5-1. Vücut iskelet modelindeki düğümlerin isimlendirilmesi.	55
Tablo 5-2. Gerçekleştirilen poz tahmin yönteminin vücut bölümleri üzerindeki ortalama piksel hatası.	73
Tablo 5-3. Gerçekleştirilen poz tahmin yönteminin farklı video kayıtları üzerindeki performansı.....	73

KISALTMALAR

PCA	: Principal Component Analysis (Temel Bileşen Analizi)
ICA	: Independent Component Analysis (Bağımsız Bileşen Analizi)
MS	: Mean Shift (Ortalama Kayma)
KF	: Kalman Filter (Kalman Süzgeci)
PF	: Particle Filter (Parçacık Süzgeci)
NP	: Nondeterministic Polynomial (Belirleyici Olmayan Polinom)
GMM	: Gaussian Mixture Model (Gaussian Karışık Model)
RGB	: Red Green Blue (Kırmızı Yeşil Mavi)
HSV	: Hue Saturation Value (Renk Doygunluk Değer)
HSL	: Hue Saturation Lightness (Renk Doygunluk Hafiflik)
LAB	: Luminescence Red/Green Blue/Yellow (Hafiflik Kırmızı/Yeşil Mavi/Yeşil)
LUV	: Luminescence Saturation Hue Angle (Hafiflik Doygunluk Renk Açısı)
MTT	: Multiple Target Tracking (Çoklu Hedef Takibi)
MHT	: Multiple Hypothesis Tracking (Çoklu Hipotez Takibi)
JPDA	: Joint Probability Data Assosiation (Ortak Olasılıklı Veri Bağı)
VISOR	: Video Surveillance Online Repository (Çevrimiçi Video İzleme Deposu)
AMISE	: Asymptotic Mean Integrated Squared Error (Asimtotik Ortalama Karesel Hata)
TPR	: True Positive Rate (Doğru Pozitif Oran)
FPR	: False Positive Rate (Yanlış Pozitif Oran)
TP	: True Positive (Doğru Pozitif)
FP	: False Positive (Yanlış Pozitif)
TN	: True Negative (Doğru Negatif)
FN	: False Negative (Yanlış Negatif)

SEMBOLLER LİSTESİ

t	: Zaman
M	: Ağırlık Merkezi
Σ	: Kovaryans
X^t	: Durum Vektörü
Z^t	: Ölçüm Vektörü
$f(.)$	: Sistem Modeli
$h(.)$	: Ölçüm Modeli
W^t	: İşlem Modelindeki Beyaz Gürültü
N^t	: Ölçüm Modelindeki Beyaz Gürültü
$p(X^t Z^{1,...,t})$	: Sonrasal Olasılık Yoğunluk Fonksiyonunu
$p(Z^t X^t)$	: Maksimum Olabilirlik Fonksiyonunu
$\bar{\Sigma}^t$	: Kovaryans Tahmini
D	: Durum Dönüşüm Matrisi
Q^t	: Beyaz Gürültünün Kovaryansı
M	: Ölçme Matrisi
K	: Kalman Kazancı
$s_t^{(n)}$	: Partikül
$\pi_t^{(n)}$	: Partikül Ağırlığı
$c^{(n)}$	: Kümülatif Ağırlık
z_t	: Ölçümler
$v_{i,l}$	: Yenilik
β_i^l	: Sonrasal Yoğunluk
$\tau_{i,l}$	: Hızlandırıcı Değişken
$K(.)$	: Çekirdek Fonksiyonu

u	: Çekirdek Fonksiyonu Giriş Parametresi
Δx	: Hareket Vektörü
Γ	: Seviye Sunum Kapalı Eğrisi
Φ	: Seviye Sunum Fonksiyonu
C	: Normalizasyon Terimi

ÖZET

Bilgisayarların veri işleme kapasitelerinin gelişmesi, yüksek kaliteye sahip ucuz kameraların üretilmesi ve görsel izleme, denetleme ve yönetme ihtiyacının ortaya çıkmasına bağlı olarak günlük yaşamdaki birçok alanda görsel takip sistemlerinin kullanılma gereksinimi, son yıllarda araştırmacıların nesne takibi konusuna olan ilgisini arttırmıştır.

Bu tez çalışmasında, birbirinden farklı üç gerçek dünya problemine çözüm sağlayan üç yeni görsel takip yöntemi sunulmaktadır. İlk takip yönteminde, kalabalık ortamlarda gezinti yapmakta olan insanların takibi amaçlanmaktadır. Gerçekleştirilen çoklu insan takip yöntemi ile sabit bir video kameranin görüş açısı içerisinde hareket halinde olan her kişiye, eşsiz bir kimlik numarası atanmaktadır. Ortamın değişen ışık miktarına ve dinamik arka plan görünümüne hızlı bir şekilde adapte olabilen çoklu insan takip yöntemi, kişi görünümlerinin kısmi kapatmaya maruz kalması gibi istenmeyen durumlarda takip işlemini yüksek bir başarımla oranı ile sürdürebilmektedir.

İkinci takip yöntemi, ten yakalama tabanlı istatistiksel insan eylem algılama yöntemidir. Gerçekleştirilen bu yöntem ile Alzheimer hastalarının tedavilerine destek sağlanmıştır. Önerilen takip yöntemi kullanılarak tedavi süresince belirli aralıklarda ilaç kullanması gerekli olan Alzheimer hastalarının vaktinde ilacı kullanıp kullanmadığının tespiti yapılabilmektedir. İlacın hasta tarafından alınıp alınmadığı bilgisi, hastanın el ve yüz uzuvlarının görsel takibi sonucunda elde edilir. Önerilen takip yöntemi yaklaşık %92 oranında doğru karar verebilmektedir.

Üçüncü takip yönteminde ise tek kamera yardımıyla elde edilen 2-boyutlu video imgeleri kullanılarak 3-boyutlu vücut poz tahmini yapılmaktadır. Gerçekleştirilen poz tahmin yöntemi, 3-boyutlu yapay bir insan modelinin 2-boyutlu imge üzerine izdüşümünün hesaplanarak vücut eklem açılarının tahmin edilmesine dayanır. Güncel eklem açılarını tahmin edebilmek için her bir vücut bölümüne ait renk ve şekil bilgileri bir Bayesian çatısı altında birleştirilir. Tahmini eklem açıları kullanılarak 3-boyutlu yapay insan modelinin benzetimi gerçekleştirilir. Elde edilen 3-boyutlu yapay insan modeli yardımıyla “yürüme”, “koşma” ve “zıplama” hareketleri başarılı bir şekilde takip edilmektedir.

Anahtar Kelimeler: Nesne Takibi, Çoklu İnsan Takibi, İnsan Eylem Algılama, İnsan Poz Tahmini, Poz Takibi.

SUMMARY

New Approaches To Solve Encountered Problems in Tracking of Human Movements

The development of data processing capacity of computers, the production of advanced but inexpensive cameras and the requirement of use of visual tracking systems in many areas of daily life depending on the emergence of the needs such as visual tracking, monitoring and managing have increased interest of researchers in the object tracking issue in recent years.

In this thesis, three novel visual tracking methods, which can provide solutions to three different real world problems, are presented. In the first tracking method, the tracking people wandering in crowded scenes is intended. The developed multiple human tracking method assigns a unique identity number to each person in view of fixed video camera angle. Being able to adapt to varying the light amount and dynamic background information of environment, the proposed multiple human tracking method could maintain the tracking process for a high success rate in unintended conditions such as exposure to partial occlusion of human view.

The second tracking method is the statistical human action recognition method based on skin detection. This realized method provided support to treatment of Alzheimer's patients. By means of the proposed tracking method, it is determined that Alzheimer's patients, who must use drug at regular intervals, use the given drug in time or not. The information whether the patients received the drugs or not, is obtained after visual tracking hands and faces of them. The proposed tracking method gives the right decision at approximately 92% success ratio.

In the third tracking method, 3-D body pose estimation of a moving person is performed by using 2-D image sequences obtained with the help of a single camera. Pose estimation method performed is based on estimates of body joint angles by taking the projection of 3-D model of an artificial person on 2-D image. To estimate current body joint angles, the color and shape information of each body segment are combined under a Bayesian framework. A simulation of 3-D artificial human body obtained by using body joint angles is performed. The movements of “walking”, “running” and “jumping” are tracked successfully with the help of the 3-D artificial human model obtained.

Key Words: Object Tracking, Multi Human Tracking, Human Activity Recognition, Human Pose Estimation, Pose Tracking.

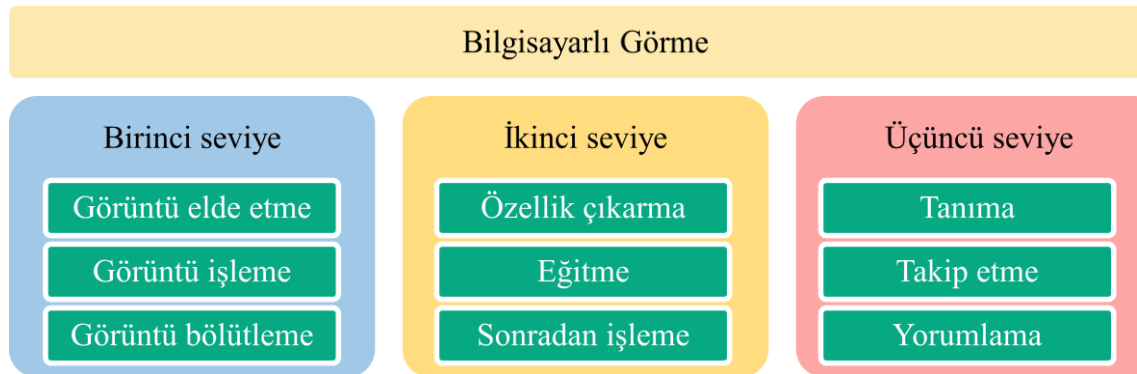
1. GİRİŞ

Bilgisayar görmesi konusu, 3-boyutlu görsel dünya bilgisinin elektronik ortamlarda anlamlı hale dönüştürülme sürecini kapsar. Şüphesiz ki; bu dönüştürme sürecinin doğru, etkin ve hızlı işleyebilmesi, kusursuz görme algısına ve anlamlandırma yetisine sahip insanın, göz ve beyin organları arasında gerçekleşen ve henüz tam anlaşılamamış biyolojik ve psikolojik faaliyetlerin dikkatli bir şekilde incelenmesine bağlıdır [1].

İnsan görüntü algılama sistemi, ortamda bulunan nesnelerin geometrik şekillerini, renksel desenlerini, hareketlerini ve birbirleri ile olan ilişkilerini anlamlı hale dönüştürülebilmesi için sırayla ortama ait görüntü bilgisinin elde edilmesine, bu bilginin işlenmesine ve son olarak mevcut bilginin algılanmasına ve yorumlanmasına ihtiyaç duymaktadır [2]. Bahsedilen tüm biyolojik ve psikolojik süreçlerin yapay sistemler tarafından modellenlenerek elektronik ortamlarda işlerliği ile ilgilenen bilgisayarlı görme çalışma alanı, konularına ve seviyelerine göre sınıflandırılarak Şekil 1-1’de gösterilmiştir.

Tez çalışmasının temel araştırma konusu olan nesne takibi, video imgelerinin işlenerek hareket halindeki nesnelerin algılanması ve takip edilmesi problemleri ile ilgilenmektedir. Daha teknik bir ifadeyle, nesne takibi, ardışık video imgelerini kullanarak imge içerisindeki hareketli nesne veya nesnelerin pozisyon bilgilerinin elde edilmesi ile ilgilenmektedir [3]. Nesne takip edicinin başarısı, bilgisayarlı görme konusunun ilk iki seviyesinin başarısı ile doğrudan ilgilidir. Bu nedenle, bir nesne takip edici tasarlanacağı zaman, bilgisayarlı görmenin tüm seviyeleri bir bütün halinde titiz bir şekilde düşünülmelidir.

Nesne takibi yöntemleri, tıbbi tedavi alanlarından endüstriyel sistemlere, askeri



Şekil 1-1. Konularına ve seviyelerine göre bilgisayarlı görme çalışma alanının sınıflandırılması.

alanlardan eğitime kadar çok geniş bir yelpazede kullanılmaktadır. Daha spesifik örnekler verilecek olursa, 4-boyutlu insan tomografi imgeleri kullanılarak akciğer tümörünün imhası [4, 5], trafik akışının düzenlenmesi için taşıt takibi [6], web sayfalarının işlevselliğini arttırmak için gözdeki iris hareketinin takibi [7] veya kalabalık bir ortamda sahibini takip edebilen robot tasarımı [8] gibi birbirinden farklı alanlarda nesne takip yöntemlerinin etkili bir şekilde kullanılmaya başlaması gösterilebilir. Geniş kullanım alanına sahip olmasından dolayı nesne takibi, son yıllarda araştırmacıların dikkatini çeken önemli bir konu haline gelmiştir.

Nesne takibi yöntemleri iki temel görevi içermektedir [9]: 1) Takip edilecek nesnenin güncel imge içerisindeki konumunun tespit edilmesi (Object Lifting); 2) İmgeler boyunca nesneler arasındaki veri bağı ilişkisinin kurulması (Data Association). Birbirinden farklı bu iki görevi aynı anda gerçekleştiren nesne takip yöntemleri, güncel gözlem bilgisini ve geçmiş imge bilgilerini tek bir çatı altında kullanarak takip edilen nesnenin bir sonraki durum tahminini gerçekleştirebilmeyi amaçlamaktadır. Bu görevleri birbirinden ayrı ele alan takip yöntemlerinde ise, öncelikle her bir imgede, bölütleme ve nesne yakalama algoritmaları kullanılır ve bağımsız nesneler belirlenir. Daha sonra, yakalanan bu nesnelerin ardışık imgeler boyunca takip edilebilmesi için nesneler arasındaki veri bağı ilişkilerinin kurulması gerekir.

1.1. Amaç

Bu tezin temel amacı, yeni görsel takip yöntemleri geliştirerek gerçek dünya problemlerine etkili ve kullanılabilir çözümler sunabilmektir. Bu temel amaç doğrultusunda aşağıda belirtilen özel ve somut amaçlara ulaşılması hedeflenmektedir;

1. Tek ve sabit bir kamera kullanarak kalabalık insan gruplarının olduğu ortamlardan elde edilen video imgelerinin işlenmesi sonucunda bağımsız hareket etmekte olan her bir insanın konum bilgisinin tespit ve takip edilmesi.
2. Ten bilgisi yardımıyla hareket halindeki bir insanın el ve yüz uzuvlarını takip ederek, önceden belirlenen bir eylemin takip edilen kişi tarafından icra edilip edilmediğinin tespit edilmesi.
3. 2-boyutlu video imgelerinin işlenmesiyle hareket halinde olan bir insanın 3-boyutlu vücut poz tahmininin yapılması ve yapılan tahmin sonuçlarının kullanılarak sayısal ortamda yapay bir insan modelinin 3-boyutlu benzetiminin gerçekleştirilmesi.

1.2. Tezin Organizasyonu ve Literatüre Katkısı

Tez çalışmasının birinci bölümünde, teze genel bir bakış açısı kazandırılmaya yönelik temel bilgiler verilmiştir. Diğer bölümlerin organizasyonu ile birlikte tezde yapılan katkılar aşağıda sunulmaktadır.

Bölüm 2’de, öncelikle nesne takibi konusuyla ilişkili olan nesne sunumu, özellik çıkarma ve nesne yakalama konuları anlatılarak nesne takibi ile olan bağları irdelenmiştir. Daha sonra nesne takip probleminin genel bir tanımı yapılarak, literatürde yer alan nesne takip yöntemleri yapılarına göre sınıflandırılmış ve her bir sınıf alt bölümleri ile birlikte detaylı olarak incelenerek sunulmuştur. Özellikle, tez çalışmasının temel çalışma alanı olarak belirtilen istatistiksel nesne takibi yöntemi ayrıntılı bir şekilde incelenmiş ve uygulama alanları üzerinde daha geniş ve kapsamlı bir şekilde durulmuştur.

Bölüm 2’de, tez çalışmasının literatürdeki yerine açıklık getirilerek dayandığı temel konular verilip, tez çalışmasında önerilen ve geliştirilen yöntemlerin anlaşılabilmesine yönelik alt yapı bilgileri kapsamlı bir şekilde sunulmuştur.

Bölüm 3’de, kalabalık ortamlarda insan takibini gerçekleştirebilen yeni bir çoklu insan takip yöntemi sunulmuştur. Güncel nesne takip yöntemlerinin avantajlarını tek bir çatı altında kullanabilen bu yöntem, kişi görünüm engellemesi olarak bilinen “kısmi kapatma” probleminin çözümü için yeni bir metod kullanmaktadır. Geliştirilen çoklu insan takip yöntemi, dinamik arka plan değişimine ve ortamın ışık değişimine karşı kısa sürede adapte olabilmektedir. Önerilen çoklu insan takip yöntemi farklı video kayıtları üzerinde test edilmiştir. Elde edilen takip sonuçları, önerilen yöntemin gerçek zamanlı uygulamalarda kullanılabileceğini göstermiştir.

Bölüm 4’de, ten yakalama tabanlı yeni bir istatistiksel insan eylem algılama yöntemi geliştirilmiştir. Geliştirilen bu yöntem, Alzheimer hastalarının tedavilerine katkı sağlayabilmek amacı ile bu hastalara ait problemin çözümünde kullanılmıştır. Tedavi etkinliğinin artırılmasını hedefleyen bu yardımcı sistem, hasta hareketlerini anlamlandırabilmek için el ve yüz uzuvlarının takibini gerçekleştirilerek önceden belirlenen bir eylemin kişi tarafından icra edilip edilmediğini tespit edebilmektedir. Farklı kişilerden elde edilen video kayıtları üzerinde test aşaması gerçekleştirilen yöntemin önceden belirlenen bir eylemi yaklaşık %92 oranında başarılı algılayabildiği görülmüştür.

Bölüm 5’de, yeni bir insan poz tahmin yöntemi önerilmiştir. Önerilen poz tahmin yöntemi, tek kamera yardımıyla elde edilen ve farklı insan hareketlerini içeren hazır bir

video veri tabanındaki bilgileri kullanmaktadır. Geliştirilen poz tahmin yöntemi, 2-boyutlu video imgelerini kullanarak 3-boyutlu insan pozunu tahmin edebilmektedir. Yapılan poz tahmin işlemi, önceden tasarlanmış 3-boyutlu yapay bir insan modelinin imge üzerine 2-boyutlu perspektif izdüşümü ile imgedeki insan silüetinin üst üste örtüştürülmesi prensibine dayanmaktadır. Geliştirilen yöntemin test aşamasında “yürüme”, “koşma” ve “zıplama” hareketlerini içeren hazır video kayıtları kullanılmıştır. Her bir 3-boyutlu vücut bölümü için elde edilen takip sonuçları matematiksel olarak verilmiştir.

Bölüm 6'de, tezin sonuçları irdelenmiş ve orijinal katkılar vurgulanmıştır. Ayrıca ileriye dönük uygulama alanları ve öneriler tartışılmıştır.

Bu tez çalışması, Fırat Üniversitesi Bilimsel Araştırma Projeleri Koordinasyon Birimi tarafından *FÜBAP-1912* numaralı proje ile desteklenmiştir.

2. NESNE TAKİP YÖNTEMLERİ

Nesne takip problemi, video imgeleri içerisinde hareket etmekte olan bir nesnenin gerçekleştirdiği hareket güzergâhını tahmin etme zorluğudur. Daha teknik bir ifadeyle, ardışık video imgelerinin analizi sonucunda hareketli nesnelerin algılanarak her bir nesneye eşsiz bir etiket (örneğin numara) verme işlemidir [11]. Bu problem, deterministik olmayan polinom (NP) tipinde bir problemdir ve problemin çözümünü güç kılan faktörler aşağıda listelenmiştir [12];

- 3-boyutlu gerçek nesne görüntülerinin 2-boyutlu imge alanında sunulması ile meydana gelen bilgi kaybı
- İmgelerdeki görüntü bilgisinin gürültü içermesi
- Video çekimi esnasında ani ışık değişimlerin yaşanması
- Arka plan görünüm bilgisinin değişken olması
- Nesne hareketlerinin karmaşık oluşu
- Nesne şekillerinin karmaşık oluşu
- Bazı nesnelerin ayırt edilebilir fiziksel bir yapıya veya renge sahip olmayışı
- Nesne görünüşlerinin engellenmesi (occlusion)
- Nesne görünüşlerinin bozulması (clutter)
- Gerçek zamanlı uygulamaların gereksinimleri

Bahsedilen bu zorlukların üstesinden gelebilmek ve tatmin edici bir takip işlemi gerçekleştirebilmek için literatürde doyurucu sayıda yaklaşım önerilmiştir. Farklı yapısal özelliklere sahip bu yaklaşımların doğru anlaşılabilmesi ve geliştirilebilmesi için bir sınıflandırma yapılmıştır. Önerilen her bir nesne takip yöntemine aşağıdaki dört soru yöneltilmiştir;

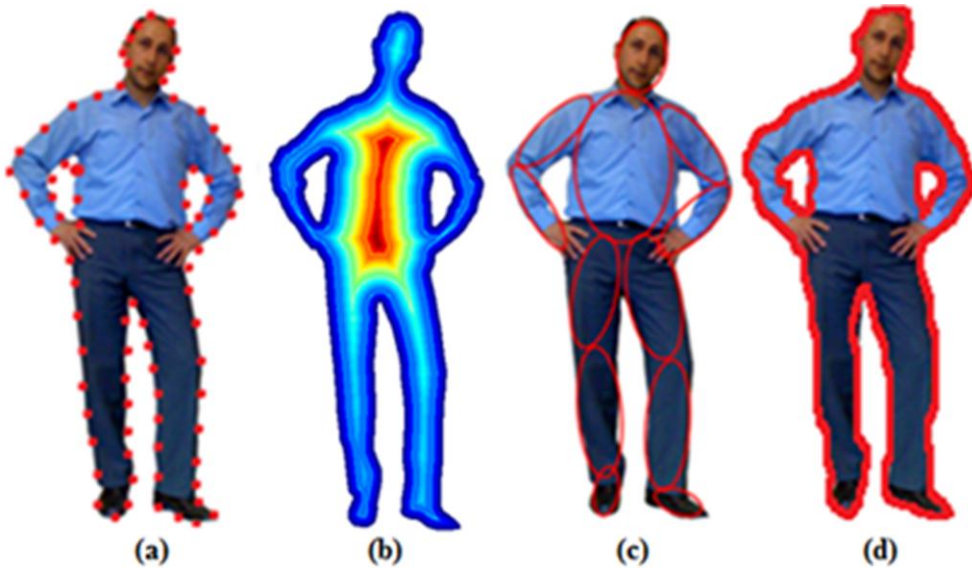
- 1) Takip edilen nesne için hangi nesne sunumu seçilmiştir?
- 2) Takip edilen nesnenin ayırt edici özellikleri nelerdir?
- 3) Takip edilen nesnenin ilk pozisyon tespiti nasıl gerçekleştirilir?
- 4) Başlangıç pozisyonu bilinen bir nesnenin imgeler boyunca takibi nasıl gerçekleştirilir?

Nesne takip edicilere yöneltilen bu dört sorunun ilk üçü, nesne sunumu, özellik seçimi ve nesne yakalama konuları ile ilgili olmasına karşın sonuncu soru doğrudan nesne takibi konusu ile alakalıdır. Ancak, bir nesne takip edicinin tasarlanabilmesi için sırasıyla

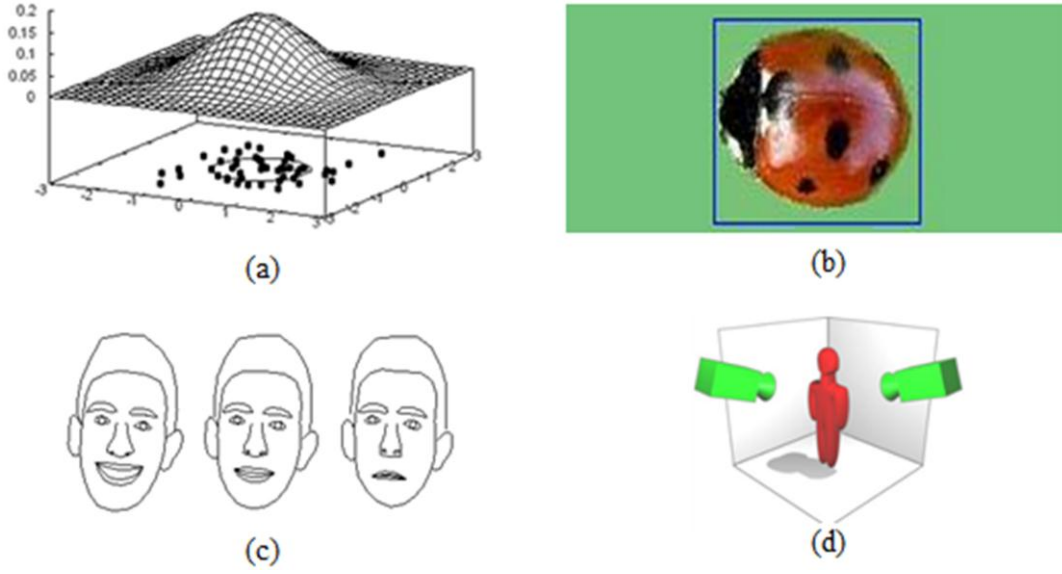
nesne sunumu, özellik seçimi, nesne yakalama ve nesne takibi aşamalarının titizlikle belirlenmiş olması gerekir [3]. Çünkü bu aşamaların başarısı nesne takip edicinin başarısını doğrudan etkilemektedir.

2.1. Nesne Sunumu

Nesne sunumu ile nesne takibi arasında güçlü bir ilişki bulunmaktadır. Nesne sunumu, takip edilecek nesnenin şekil ve görünüm bilgilerinin nasıl ifade edileceğini belirler. Örneğin, yerdeki böcekleri, gökyüzündeki kuşları, tarlada çalışan insanları veya akciğer üzerindeki tümörü, bağımsız bir nesne olarak ifade edebilmek için bu nesnelere ait şekil ve görünüm bilgilerine ihtiyaç duyulmaktadır. Nesnelere ait şekil ve görünüm bilgileri değişik türlerde ifade edilmektedir [3]. Literatürde sık kullanılan nesnenin şekilsel sunum türleri şunlardır: 1) Noktasal sunum [13, 14]; 2) Geometrik sunum [15]; 3) Siluet veya dış hat sunumu [16-18]; 4) Yapay modeli sunum [19]; 5) İskelet sunumu [20]; 6) Seviye kümesi sunumu [109, 110]. Şekil 2-1’de sık kullanılan şekilsel sunum türleri gösterilmektedir. Literatürde sık kullanılan nesnenin görsel sunum türleri şunlardır: 1) Olasılık yoğunluğu [21-26]; 2) Şablonlar [27]; 3) Aktif şekil/görünüm modelleri [28-30]; 4) Çoklu görüş modelleri. Şekil 2-2’de literatürde nesne takibiki alanında sık kullanılan görsel sunum türlerinden birkaçtanesi gösterilmiştir.



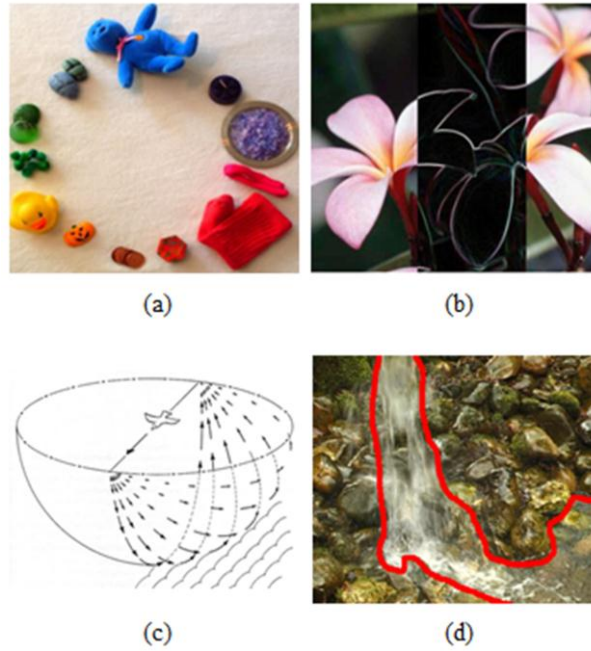
Şekil 2-1. Sık kullanılan şekilsel sunum türleri (a) İşaretçi sunumu (b) Seviye kümeleri sunumu (c) Yapay model sunumu (d) Dış hat sunumu.



Şekil 2-2. Sık kullanılan görsel sunum türleri **(a)** Olasılık yoğunluğu **(b)** Şablonlar **(c)** Aktif şekil/görünüm modelleri **(d)** Çoklu görüş modelleri.

2.2. Nesne Özellikleri ve Seçimi

Doğru nesne özelliklerinin seçimi nesne takibi algoritmalarının başarısı üzerinde önemli bir rol oynar. İmge içerisinde nesnenin ayırt edici özelliklerinin doğru belirlenmesi, nesne takip işleminin başarısını artırır. Şekil 2-3'te gösterilen genel nesne özellikleri şunlardır [31]: 1) Renk özelliği [32-34]; 2) Kenar özelliği [35, 36, 169]; 3) Optik akış özelliği [37, 38]; 4) Doku özelliği [39, 40]. Bahsedilen nesne özellikleri birçok nesne takip



Şekil 2-3. Genel nesne özellikleri **(a)** Renk özelliği **(b)** Kenar özelliği **(c)** Optik akış özelliği **(d)** Doku özelliği.

uygulamasında statik olarak el ile belirlenmektedir [15, 16, 19, 20, 26, 27]. Oysa, nesne özelliklerini otomatik belirleyebilen yöntemler literatürde yer almaktadır. Bu yöntemler için Temel Bileşen Analizi (PCA) [41] ve Adaboost [42] örnek olarak verilebilir. Nesne özelliklerinin otomatik belirlenmesi, nesne takip edicinin kullanımını kolaylaştırmasının yanında sağlamlılığında artmasına destek sağlamaktadır.

2.3. Nesne Yakalama

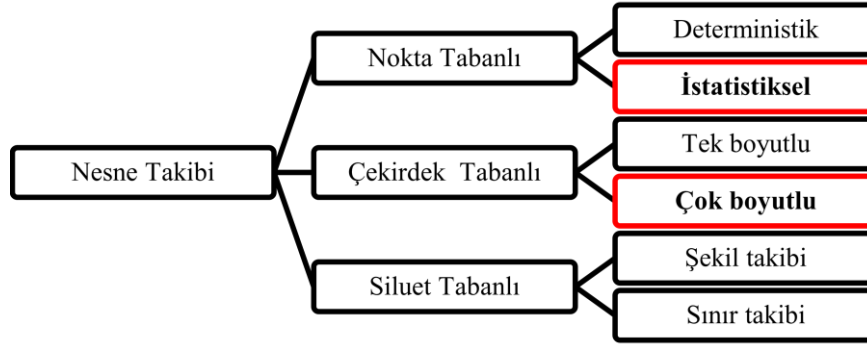
Genellikle nesne takip uygulamalarında gerçekleştirilen ilk adım, güncel imge içerisinde takip edilecek nesnenin yakalanması işlemidir [43]. Bu aşamanın başarılı olup olmaması takip işleminin başarısını doğrudan etkilemektedir.

Nesne yakalama konusu imge bölütleme başlığı altında incelenmektedir. İmgedeki aynı karakteristiğe sahip alanların çıkartılması, arka plandan ayrılması ve belirgin bir hale getirilmesi işlemine imge bölütleme işlemi denir [44-46]. Bölütleme işlemine, bir cadde imgesindeki arabaların tespiti, pilotun havadayken yerdeki nesneleri ayırt edebilmesi veya iç içe nesnelerin ayrılması örnek olarak verilebilir. Cadde imgesi için düşünülürse, yapılan ilk işlem imgenin bölütlenmesi, ikinci işlem ise bölütlenen imge parçalarının potansiyel araba boyutlarına karşılık gelip gelmediğinin kontrol edilmesidir [47-49]. Nesne yakalama ve imge bölütleme konusu hakkında detaylı bilgi [51, 44-46] nolu çalışmalarda bulunmaktadır.

2.4. Nesne Takip Yöntemleri

Nesne takip yöntemleri, genel özellikleri bakımından nokta tabanlı, çekirdek tabanlı ve siluet tabanlı olmak üzere üç başlık altında gruplandırılmıştır [3]. Şekil 2-4'te yer alan bu sınıflandırmadaki kırmızı çerçeve içerisine alınan yöntemler, tez çalışmasında gerçekleştirilen uygulamalarda kullanılmıştır.

Nesne takip yöntemleri literatürde çok farklı uygulama alanlarında kullanılmaktadır. Örneğin; solunum esnasında hareket etmekte olan akciğer tümörünün hareket güzergahının tespiti [4, 5, 52, 53]; şüpheli aktivitelerin veya istenmeyen hareketlerin tespit edilebilmesi için sabit bir ekran görüntüsünün izlenmesi [54]; çoklu ortam veritabanında bulunan videolarına hızlı bir şekilde ulaşılabilmesi [55]; insan vücut uzuvlarının takip edilerek yapılan hareket hakkında bilginin üretilmesi [19, 20]; trafik akış



Şekil 2-4. Nesne takip yöntemlerinin sınıflandırılması [3].

yoğunluğunun belirlenerek sürücünün bilgilendirilmesi veya seyir halindeki araçların takibi [6, 56]; aracın şoförsüz seyahat edebilmesi veya yol çizgilerini takip edebilmesi [57] gibi birçok farklı alanda nesne takip yöntemlerinin uygulamalarını görmek mümkündür.

2.4.1. Nokta Tabanlı Nesne Takibi

Nokta tabanlı takip yöntemleri iki adımdan oluşmaktadır [13]; 1) Her bir nesnenin tek bir nokta ile ifade edilmesi (nesne yakalama); 2) Belirlenen noktalar ile önceki imgede tespit edilen noktalar arasındaki nokta benzerlik değerlerinin hesaplanması (veri bağı problemi). Nesnelerin nokta ile ifade edilebilmesi için başlangıç pozisyon bilgisine ihtiyaç duyulmaktadır. Bunun için harici bir nesne yakalama yöntemi kullanılmakta veya kullanıcı tarafından el ile belirlenmektedir. Nokta benzerliğinin kurulabilmesi işlemi, özellikle nesne görünümünün kaybolması, yanlış nesnelerin yakalanması, nesnenin imgeye ilk giriş/çıkışı gibi durumlarda karmaşık bir hale dönüşebilir. Nokta benzerlik probleminin çözümü için önerilmiş yöntemler genel olarak iki kategori altında incelenmektedir: 1) Belirleyici yöntemler; 2) İstatistiksel yöntemler. Belirleyici yöntemler, benzerlik problemini sınırlamak için hareket bilgilerini kullanır. Diğer taraftan istatistiksel yöntemler, nesne ölçümlerini kullanarak nokta benzerliğinin kurulmasında belirsizlikleri (gürültü değerlerini) göz önüne alır.

2.4.1.1. Belirleyici Yöntemler

Belirleyici yöntemler, $t - 1$ anında imgede bulunan herhangi bir nesneyi t anında bulunan tek bir nesneye bağlama maliyetini (benzerlik maliyeti) tanımlamaktadır [60]. Bu

işlemi yaparken nesne hareketleri üzerinde gerekli sınırlamalar yapılmaktadır. Örneğin; nesnelerin birbirlerine olan uzaklıkları, maksimum veya minimum hızları, katılık veya yumuşaklık özellikleri gibi bilgiler nesne hareketlerini doğrudan etkilemektedir. ile ilgili olmaktadır. yakınlık, minimum hız, Benzerlik maliyetinin minimum seviyeye indirgenmesi, optimizasyon problemi olarak formüle edilmektedir. Literatürde sık kullanılan optimizasyon algoritmalarına Hungarian [58] ve greedy algoritması [59] örnek olarak verilebilir.

2.4.1.2. İstatistiksel Yöntemler

İstatistiksel yöntemler, nesne takip probleminin çözümünde kullanılabildiği gibi zamanla değişim gösteren herhangi bir sistemin durum tahminini gerçekleştirmek için de kullanılabilmektedir. Örneğin, nesne sınırlarının izlenmesi, aktivitelerin algılanması, hareket yapılarının belirlenmesi gibi sistemlerde istatistiksel yöntemler sıklıkla kullanılmaktadır [61].

Gerçek ortamlardan elde edilen video kayıtları gürültü bilgisini içermekle birlikte takip edilen nesne birtakım istenmeyen etkilere maruz kalabilir. Tüm bu olumsuzluklara rağmen, istatistiksel yöntemler, nesnenin durum tahminini gerçekleştirebilmek için güncel ölçüm değerlerini ve model belirsizliklerini (gürültü) kullanabilmektedir. Nesnenin güncel imgedeki pozisyon, hız veya ivme verileri ölçüm bilgisi olarak kullanılmaktadır.

Nesne takip problemlerinin çözümünde istatistiksel yöntemlerin nasıl kullanıldığını doğru anlayabilmek için problemi net bir şekilde tanımlamak gerekmektedir. Ekranda hareket eden bir nesne düşünelim. Takip edilecek nesnenin durum bilgisi $X^t: t = 1, 2, \dots$ şeklinde bir dizi olarak tanımlanmaktadır. Zaman boyunca durum bilgisi üzerindeki değişim aşağıdaki dinamik denklem ile ifade edilmektedir;

$$X^t = f(X^{t-1}) + W^t \quad (2.1)$$

Denklem (2.1.)’deki $f(.)$, doğrusal olmayan sistem modelidir ve X durum bilgisinin $t - 1$ anından t anına geçerken nasıl bir dönüşüme maruz kalacağını tanımlar. $W^t: t = 1, 2, \dots$ beyaz gürültüdür. Ölçüm verisi ile durum bilgisi arasındaki ilişki ise aşağıdaki gibi ifade edilir;

$$Z^t = h(X^t) + N^t \quad (2.2)$$

Denklem (2.2)'deki N^t beyaz gürültüdür ve W^t den bağımsızdır. İstatistiksel yöntemlere dayalı nesne takip edicilerin temel amacı, t anına kadarki tüm ölçüm değerlerini göz önüne alarak X^t durum değişkenini tahmin etmektir. Yani, sonrasal olasılık yoğunluk fonksiyonu $p(X^t|Z^{1,...,t})$ değerini elde etmek demektir. Teorik olarak en uygun çözüm, problemi iki adımda çözebilen tekrarlamalı Bayesian süzgeci [59] yöntemini kullanmaktır. Bayes süzgeci, tahmin ve düzeltme olarak bilinen iki adım içermektedir. Tahmin adımı, dinamik bir eşitlik kullanmakta ve güncel durumun $t - 1$ anındaki öncesel olasılık yoğunluk fonksiyonu $p(X^t|Z^{1,...,t-1})$ değerini hesaplamaktadır. Düzeltme adımı ise, güncel ölçümün maksimum olabilirlik fonksiyonunu, yani $p(Z^t|X^t)$ 'i kullanmaktadır. Buna göre, sonrasal olasılık yoğunluk fonksiyonu aşağıdaki denklem kullanılarak elde edilmektedir;

$$p(X^t|Z^{1,...,t}) = \frac{p(X^t|Z^{1,...,t-1}) p(Y^t|X^t)}{p(Y^t|Z^{1,...,t-1})} \quad (2.3)$$

Denklem (2.3)'teki $p(Y^t|Z^{1,...,t-1})$ değişkeni normalizasyon katsayısını göstermektedir. İmge içerisinde sadece tek bir nesnenin olması durumunda, nesnenin durum bilgisi tahmin ve düzeltme adımı kullanılarak rahatlıkla elde edilmektedir. Diğer taraftan, imge içerisinde birden fazla nesnenin olması durumunda, elde edilen ölçümler ile ilgili nesneler arasında gerekli bağlantıların kurulma ihtiyacı ortaya çıkar.

Tek nesneli durum tahmini

İmge içerisindeki tek bir nesnenin takibi gerçekleştirilirken aşağıda belirtilen durumların mevcut olması halinde en uygun çözüm Kalman süzgeci yöntemi [62, 63] tarafından sağlanmaktadır.

- f^t ve h^t fonksiyonları doğrusal olmalı
- Nesnenin başlangıç durumu X^1 Gaussian dağılımına sahip olmalı
- İşlem gürültüsü Gaussian dağılımına sahip olmalı

Yukarıda bahsedilen durumlardan herhangi birinin mevcut olmaması halinde en uygun çözüm Parçacık süzgeci yöntemi [64] tarafından sağlanmaktadır.

Kalman süzgeci yöntemi: Kalman süzgeci yöntemi, doğrusal sistemlerin durum tahminlerini gerçekleştirmek için kullanılmaktadır. Burada dikkat edilmesi gereken nokta; sistem gürültüsünün Gaussian dağılımına sahip olması gerekliliğidir. Kalman süzgeci yöntemi, Bayesian süzgeci yöntemine benzer olarak tahmin ve düzeltme adımlarını içermektedir. Tahmin adımında, değişkenlerin yeni durum tahminini gerçekleştirmek için durum modelini kullanmaktadır. Şöyle ki;

$$\bar{X}^t = D\bar{X}^{t-1} + W \quad (2.4)$$

$$\bar{\Sigma}^t = D\Sigma^{t-1}D^T + Q^t \quad (2.5)$$

Denklem (2.4) ve (2.5)'de kullanılan $\bar{X}^t$ ve $\bar{\Sigma}^t$ değişkenleri, t anındaki durum ve kovaryans tahminlerini vermektedir. Her iki denklemde kullanılan D değişkeni, t ile $t - 1$ anındaki durum değişkenleri arasındaki ilişkiyi tanımlayan durum dönüşüm matrisidir. Q değişkeni, W gürültüsünün kovaryansıdır.

Düzeltilme adımında, güncel gözlem değeri (Z^t) nesne durumunu güncellemek için kullanılmaktadır. Buna göre Kalman süzgeci yönteminin düzeltme adımı aşağıdaki eşitliklerle tanımlanmaktadır;

$$K^t = \bar{\Sigma}^t M^t [M^t \bar{\Sigma}^t M^t + R^t]^{-1} \quad (2.6)$$

$$X^t = \bar{X}^t + K^t \underbrace{[Z^t - M\bar{X}^t]}_v \quad (2.7)$$

$$\Sigma^t = \bar{\Sigma}^t - K^t M \bar{\Sigma}^t \quad (2.8)$$

Denklem (2.7)'deki v değişkeni, Kalman yeniliği olarak adlandırılmaktadır. M değişkeni ölçme matrisini göstermektedir. K , durum modelinin yayılımı için kullanılan kalman kazancıdır. Dikkat edilmesi gereken nokta şudur ki; güncellenen X^t durumu yine bir Gaussian dağılımını göstermektedir. Eğer f^t ve h^t fonksiyonlarının doğrusal olmadığı bir durum söz konusu olursa, o zaman bu fonksiyonlar Taylor seri açılımları kullanılarak doğrusallaştırılabilir. Bahsedilen bu süzgeç tekniği literatürde Genişletilmiş Kalman Süzgeci (EKF) [64] olarak bilinir. Birkaç değişik versiyonu bulunan Kalman süzgeç yöntemi, veri bağı problemi çözüldükten sonra çoklu nesnelerin takip edilebilmesinde kullanılabilmektedir [65].

Parçacık süzgeci yöntemi: Son yıllarda özellikle bilgisayarlı görme alanında çok popüler olan parçacık süzgeci yöntemi, nesne takip problemlerine çözüm olarak sıklıkla kullanılmaktadır. Kalman süzgecinin bir dezavantajı, durum değişkenlerinin Gaussian dağılımına sahip olma zorunluluğunun bulunmasıdır. Böylece, Kalman süzgeci Gaussian dağılımına sahip olmayan sistemler için zayıf bir tahmin edicidir. Bu dezavantaj, parçacık süzgeci yöntemiyle aşılabilmektedir [66].

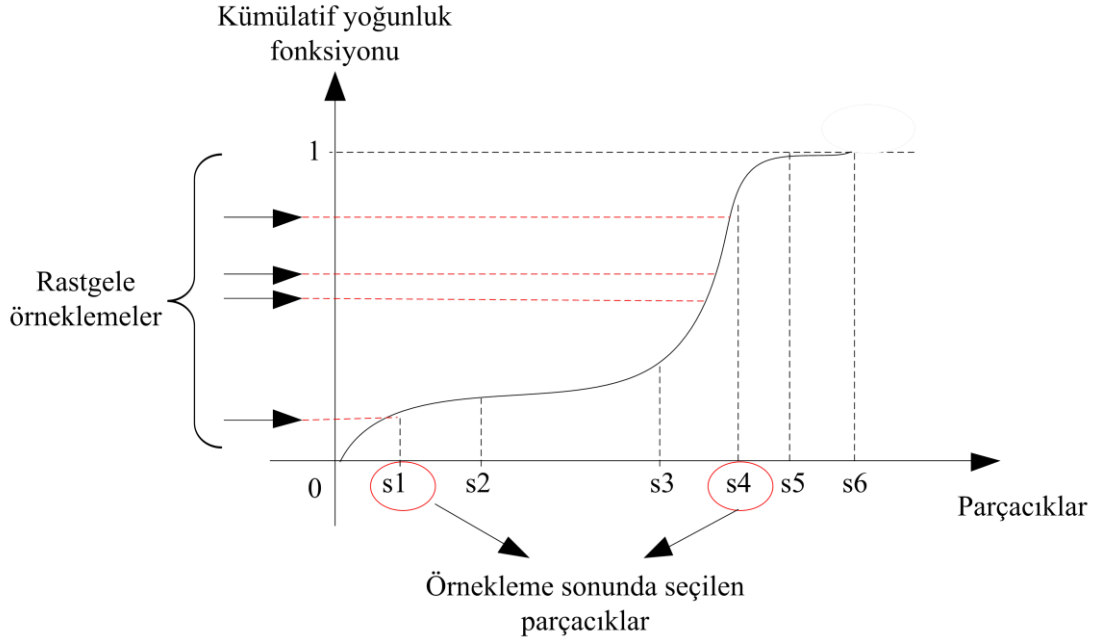
Parçacık süzgecinde, t anındaki şartsal durum yoğunluğu π ağırlığına sahip N tane parçacık içeren örnek küme ile ifade edilir ve bu örnek küme $\{s_t^{(n)}: n = 1, \dots, N\}$ şeklinde gösterilebilir. Bu kümedeki her bir parçacığın ağırlığı $\pi_t^{(n)}$ (örnekleme olasılığı) ile ifade edilir. Ağırlıklar parçacığın önemi veya onun gözlenme frekansı olarak tanımlanabilir. Her bir $(s_t^{(n)}, \pi_t^{(n)})$ 'nin hesaplama maliyetinin azaltılması için birikmiş (kümülatif) bir ağırlık değeri $c^{(n)}$ değişkeni ile aynı zamanda hafızaya yüklenir. Buna göre, son parçacığın birikmiş ağırlık değerleri ($c^{(N)} = 1$) olmaktadır. t anındaki yeni örnekler $t - 1$ anındaki hafızada mevcut olan $S_{t-1} = \{(s_{t-1}^{(n)}, \pi_{t-1}^{(n)}, c_{t-1}^{(n)}): n = 1, \dots, N\}$ örnekleme şeması kullanılarak elde edilir. En genel örnekleme şeması aşağıda bahsedildiği gibi *önem örnekleme* şemasıdır. Bu şema, *seçme*, *tahmin* ve *düzeltilme* olmak üzere üç temel adımın tekrarlı bir şekilde gerçekleşmesi prensibine dayanır. Bu adımlar şu şekilde gerçekleşir:

(1) *Seçme adımı:* S_{t-1} örnekleme şemasından rastgele N tane $\hat{s}_t^{(n)}$ örneği seçilir. Seçme işlemi yapılırken $r \in [0, 1]$ olmak kaydıyla rastgele bir değişken üretilir ve $c_{t-1}^{(j)} > r$ ve $\hat{s}_t^{(n)} = s_{t-1}^{(j)}$ şartlarını sağlayan en küçük j indeksine sahip örnekler seçilir. Bu adımda gerçekleşen seçme operasyonu Şekil 2-5'de gösterilmiştir.

(2) *Tahmin adımı:* Seçilen her bir $\hat{s}_t^{(n)}$ örneği için $s_t^{(n)} = f(\hat{s}_t^{(n)}, W_t^{(n)})$ fonksiyonu kullanılarak yeni bir örnek üretilir. Burada $W_t^{(n)}$ sıfır ortalamaya sahip bir Gaussian gürültüsüdür ve f negatif olmayan bir fonksiyondur ($f(s) = s$).

(3) *Düzeltilme adımı:* $s_t^{(n)}$ örneklerine ait $\pi_t^{(n)}$ ağırlıkları z_t ölçümleri kullanılarak hesaplanır. Bunun için $\pi_t^{(n)} = p(z_t | x_t = s_t^{(n)})$ eşitliği kullanılır. Burada $p(\cdot)$, Gaussian dağılımı ile modellenebilir.

Elde edilen yeni S_t örnekleri, $\varepsilon_t = \sum_{n=1}^N \pi_t^{(n)} f(s_t^{(n)}, W)$ eşitliğinde kullanılmakta ve yeni nesne pozisyonları hesaplanmaktadır. Parçacıkların başlangıç ağırlık değerleri iki



Şekil 2-5. Parçacık süzgecindeki seçme adımının çalışma şekli.

şekilde belirlenmektedir: 1) Örnekleme dizisini kullanan sistemlerin eğitimiyle; 2) İlk ölçüm değerleri ($s_0^{(n)} \sim X_0$) kullanılarak. Başlangıç değerleri için ilk ölçümler kullanıldığında, her bir örneğin ağırlık değeri $\pi_0^{(n)} = \frac{1}{N}$ eşit olarak dağıtılmaktadır. Ayrıca en iyi parçacık örneklerinin korunması için düşük ağırlıklı olanların elenmesi gerekir. Bunun için ek bir örnekleme algoritmasına ihtiyaç duyulur. Burada dikkat edilmesi gereken nokta, sonrasal yoğunluğun Gaussian olmak zorunda olmayışıdır. Parçacık süzgeci ile ilgili daha kapsamlı bilgi [69-70] numaralı çalışmalarda mevcuttur.

Çok nesneli durum tahmini

Kalman ve parçacık süzgeci yöntemleri, bir anda tek bir ölçüm değerinin olduğunu farz etmektedir. Yani bu süzgeç yöntemleri, imge içerisinde aynı anda sadece tek bir nesnenin durum tahminini gerçekleştirebilmektedir. Bu nedenle, çoklu nesnelerin takibi işlemi için veri bağı ve durum tahmini problemlerinin çözümünü bir arada gerçekleştirebilen yöntemlere ihtiyaç duyulmaktadır [67, 68].

Çoklu nesne takibi yönteminde, ölçüm değerleri kullanılarak hedeflerin güncel durum vektörleri tahmin edilmektedir. Eğer ölçüm değeri üreten nesne sayısı sabit ise ve biliniyorsa, çoklu nesne takip problemi, tekli nesne takip problemine indirgenebilir. Böylece, birbirinden bağımsız hedeflerin durum tahminleri için standart bir süzgeç

algoritması (Kalman veya Genişletilmiş Kalman) kullanılabilir. Ancak, takip edilecek nesne sayısının bilinmesi, pratikte nadiren görülen bir durumdur. Örneğin, sabit bir kamera yardımıyla otoyoldaki trafik akışının veya okul bahçesindeki öğrencilerin izlendiği düşünüldüğünde, kameranın görüş alanı içerisine sürekli nesne giriş çıkışı olabilir. Bununla birlikte, takip edilen hedef nesnelerin üst üste gelmesi veya aniden kaybolması gibi durumlar görülebilir. Bu nedenle, takibi gerçekleştirecek hedef nesne sayısının ve ölçüm bilgisinin bilinmesi pratikte az rastlanan bir durumdur. Hedef sayısının belli olmaması veya zaman içerisinde bu sayının değişime uğraması, "veri bağı" veya "veri ilişkisi" probleminin ortaya çıkmasına neden olur. Böylece, takip probleminin karmaşıklığı da artmış olur [71].

Veri bağı probleminin en basit çözümü, en yakın komşu yaklaşımıdır. Ancak takip edilen nesnelerin birbirlerine yakın olduğu durumlarda, en yakın komşu yaklaşımı uygun sonuç vermemektedir. Veri bağı problemi, nesne takibi işleminin başarısını doğrudan etkilemektedir. Literatürde veri bağı problemi için sunulan çözüm önerileri iki farklı sınıf altında toplanmaktadır: 1) Tüm Komşulu Veri Bağı (örneğin, JPDA); 2) Eşsiz Komşulu Veri Bağı (örneğin, MHT [37, 73-75, 125, 170]). Bu yöntemlerin çalışma prensipleri arasındaki temel fark şu şekilde özetlenebilir: MHT’de her bir ölçüm değeri, önceden oluşturulan izlerden sadece biri ile ilişkilendirilir. JPDA’da ise, tüm ölçüm değerleri, izlerin tamamını güncellemek için kullanılır. MHT ve JPDA yöntemleri hakkında daha kapsamlı bilgi için [72, 76] numaralı çalışmalar incelenebilir.

Ortak Olasılıklı Veri Bağı: Bir nesnenin video imgeleri içerisinde gerçekleştirmiş olduğu hareket güzergâhını “iz” olarak adlandıralım ve N nesne için N tane iz’in olduğunu varsayalım. Bunun yanında t anında m tane ölçüm değerinin olduğunu ve bu ölçümlerin var olan iz’lere atanmak istenildiğini düşünelim. İz sayısının zaman ilerledikçe değişmeyeceği düşünülerek iz’ler ile ölçüm atamalarının bir kümesini η ile ifade edelim. $v_{i,l}$, l . iz ile i . ölçüm arasındaki yenilik değeridir ve bu durum aşağıdaki gibi ifade edilir;

$$v^l = \sum_{i=1}^{m_k} \beta_i^l v_{i,l} \quad (2.9)$$

Denklem (2.9)’daki, l . iz’e sahip nesnenin i . ölçüme ait sonrasal yoğunluk değeri β_i^l ile ifade edilir ve aşağıdaki gibi hesaplanır.

$$\beta_i^l = \sum_{\eta} P(\eta_l(k)|Z^t) \tau_{i,l}(\eta) \quad (2.10)$$

Denklem (2.10)'daki, $\tau_{i,l}$, $i = 1, \dots, m_k$ ve $l = 1, \dots, N$ indeklerine bağlı hızlandırıcı değişkendir. Eğer $z_i(k)$ ölçümü l izi ile bağlantılı ise 1 diğer durumlarda 0 sonucunu üretir.

Veri bağı problemini çözebilmek için önerilen bu yöntemin bir uygulaması [87]'de görülmektedir. Bu yöntemin en büyük dezavantajı, izlenecek nesne sayısının sabit olma zorunluluğudur. Ayrıca bu yöntem, veri bağı probleminin çözümü için sadece ardışık iki imge bilgisine ihtiyaç duyar.

Çoklu Hipotez Takibi: Veri bağı probleminin çözümünde ardışık iki imgenin kullanılması durumu, istenmeyen nesne ilişkilerinin elde edilme olasılığını artırır. Bunun yerine, nesneler arasındaki veri bağı ile ilgili bir karar verilirken güncel birkaç imgenin birlikte incelenmesi, daha doğru bir kararın verilmesinde etkili olmaktadır. Çoklu hipotez takibi algoritması [76], her bir nesne için birkaç muhtemel ilişkinin sürdürülebilmesi anlayışına dayanmaktadır. Nesnenin nihai izi, zamanla elde edilen ilişkiler içerisindeki en büyük olasılık değerine sahip olan izlerin bir kümesidir. Algoritma, görüş alanına yeni giren nesneler için yeni izler oluşturabilir veya görüş alanındaki nesne izlerini yok edebilir. Aynı zamanda kısmi veya tam üst üste gelme problemine de çözüm sağlayabilir. Daha açık bir ifadeyle, nesne ölçümlerinden bazıları kaybolursa bile, nesneye ait izler sürdürülebilir.

Çoklu hipotez takibi algoritması tekrarlanan bir algoritmadır ve her bir adımda güncel hipotezlerden oluşan bir iz kümesi bulunur. Her bir hipotez, ayrık izlerin bir toplamıdır ve her biri için sonraki imgede bir tahmin yapılır. Daha sonra tahminler güncel ölçümler ile karşılaştırılır ve her bir hipotez için bir ilişki kümesi kurulur. Her bir yeni hipotez, yeni ölçümlere dayanan bir iz kümesini ifade eder. Dikkat edilmesi gereken nokta; herhangi bir güncel ölçüm değeri, görüş alanına yeni giren bir nesneye, daha önceden takip edilen bir nesneye veya sahte bir nesneye ait olabilir. Ancak her nesneye bir ölçüm atanmayabilir. Çünkü takip edilen nesne görüş alanının dışına çıkabilir veya bir nesneyle ilgili bir ölçüm elde edilemeyebilir. Bu durumda nesne kaybolabilir ya da gürültüden dolayı algılanamaz.

Çoklu hipotez takibi algoritması muhtemel tüm nesne bağlantılarını kapsamlı bir şekilde ifade etmektedir. Bu nedenle algoritmanın hafıza ve zaman tüketimi aşırı derecede fazladır. Bu dezavantajı ortadan kaldırmak için [77] numaralı çalışmada istatistiksel bir

yöntem önerilmektedir. Bu yöntem ile nesneler arasındaki her bir bağlantı, bağımsız bir rastgele değişken olarak düşünülmektedir. Böylece nesne ilişkileri düşük olan iz'ler göz ardı edilmektedir. Bunun sonucunda, hafıza daha verimli kullanılmış olacaktır. Ayrıca, çoklu hipotez takibi ile görüntüsü engellenen veya deforme olan nesnelerin takibi gerçekleştirilmektedir. Özellikle, kalabalık ortamlarda insan takibinin yapılabilmesi konusu, son yıllarda ilgi çeken bir çalışma alanı olmuştur. Bu konuda yapılan çalışmalarda [78-86], görüş alanına giren her bir insana eşsiz bir numara verilir ve takip işlemi başlar. Sonraki ardışık imgelerde, kişiler ilk atanan bu numaralar ile ifade edilmektedir. Önerilen takip yöntemlerinin en önemli özelliği, insan görünümünün bir kısmının kaybolması (bir başka insan veya nesne tarafından engellenmesi) durumunda bile takip etme işleminin devam edebilmesidir.

2.4.2. Çekirdek Tabanlı Nesne Takibi

Çekirdek tabanlı nesne takip yöntemlerinde, ilk olarak takip edilecek nesne basit bir geometrik şekil içerisine alınır ve görünüm bilgisinin olasılık yoğunluk dağılımı elde edilir. Sonra elde edilen yoğunluk dağılımı ardışık video imgeleri boyunca takip edilir. Olasılık yoğunluk dağılımı, takip edilecek nesnenin her pikselinin diğer pikseller üzerindeki etkisini ifade eder [21, 22].

Çekirdek tabanlı nesne izleyiciler, parametrik olmayan tahmin ediciler gurubu içerisinde yer almaktadır. Parametrik olmayan sistemlerde sabit bir fonksiyon yapısı söz konusu değildir ve bir tahmin gerçekleşeceği zaman dağılıma ait tüm veriler göz önünde bulundurulur. Ancak, parametrik sistemlerde sabit bir fonksiyon yapısı ve parametre değerleri bulunmaktadır [88].

Çekirdek tabanlı nesne takip yönteminin iyi anlaşılabilmesi için öncelikle histogram sunumunu ve böyle bir sunumun sahip olduğu dezavantajların bilinmesi gerekmektedir. Bir nesnenin histogram bilgisi elde edilirken renklerin hangi aralıkta kaç eşit parçaya bölünmesi gerektiği bilinmeli ve bu parçaların başlangıç ve bitiş noktalarının hangi değerler olacağı belirlenmelidir. Bu zorunluluklar, histogram sunumunun cazibesini azaltır ve bu sınırlamaların olmadığı çekirdek yoğunluk fonksiyonlarının kullanılmasına neden olur [89]. Histogramdaki bu bağımlılıkları kaldırabilmek amacıyla, çekirdek tabanlı nesne takip ediciler, nesne üzerindeki her bir pikseli merkez kabul ederek piksellere ait olasılık yoğunluk değerlerini elde eder. Burada kullanılan çekirdek fonksiyonunun

yumuşak veya sertliği, yapılan tahmininin yumuşak veya sert olmasına neden olur. Bu sayede histogram sunumu kullanıldığı zaman elde edilen dezavantajlar ortadan kalkmış olur [89].

Literatürde sıkça kullanılan çekirdek fonksiyonlarının listesi Tablo 2-1’de verilmiştir. Tablo 2-1’de kullanılan $K(\cdot)$ çekirdek fonksiyonunun u giriş parametresi aşağıdaki gibi hesaplanır;

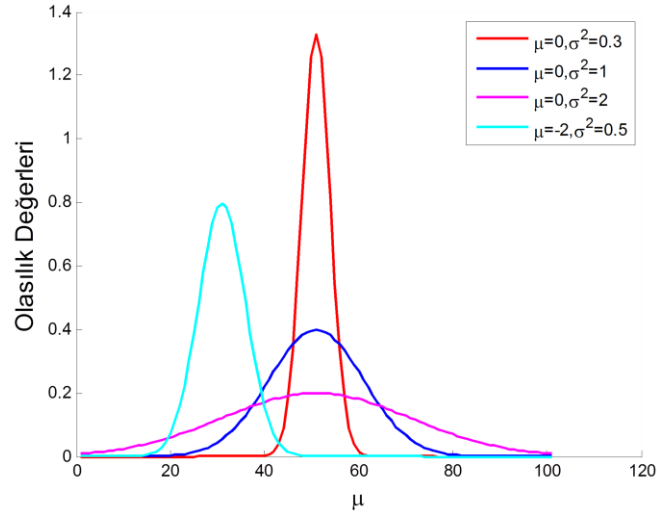
$$u = \frac{x-\mu}{\sigma} \quad (2.11)$$

Denklem (2.11)’deki μ ve σ simgeleri ile $K(\cdot)$ çekirdek fonksiyonunun ortalama (merkez) ve bant genişliği (kovaryans) değişkenleri ifade edilir. En sık kullanılan çekirdek fonksiyonu Gaussian çekirdek fonksiyonudur. Ortalama değeri sıfır olan ve değişik bant genişliklerine sahip Gaussian çekirdek fonksiyonlarının grafikleri Şekil 2-6’da gösterilmiştir.

Çekirdek tabanlı yaklaşımın en önemli dezavantajı, hesaplama maliyetinin yüksek olmasıdır. Bu durum bu yöntemin gerçek zamanlı uygulamalarda kullanımını zorlaştırmaktadır. Bu dezavantajın ortadan kaldırılabilmesine yönelik çalışmalar yapılmıştır [101]. Yapılan çalışmada, korelasyon tabanlı bir nesne takip yöntemi sunulmuştur. Bu yöntemde, takip edilecek nesnenin görünümü olasılık yoğunluk değerleri ile yapılmamaktadır. Bunun yerine ten yakalama algoritmaları kullanılarak insan yüzünün tespiti sağlanmakta ve imgeler boyunca korelasyon tekniği kullanılarak takip işlemi gerçekleştirilmektedir.

Tablo 2-1. Çekirdek fonksiyonlarının formülasyonları [90].

Çekirdek Fonksiyon Adı	Çekirdek Fonksiyonu ($K(u)$)
Uniform	$\frac{1}{2}I(u < 1)$
Triangle	$(1 - u)I(u < 1)$
Epanechnikov	$\frac{3}{4}(1 - u^2)I(u < 1)$
Quartic	$\frac{15}{16}(1 - u^2)^2I(u < 1)$
Triweight	$\frac{35}{32}(1 - u^2)^3I(u < 1)$
Gaussian	$\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}u^2\right)$
Cosinus	$\frac{\pi}{4}\cos\left(\frac{\pi}{2}u\right)I(u < 1)$



Şekil 2-6. Farklı bant genişliğine sahip gaussian çekirdek fonksiyon grafikleri [24].

Çekirdek tabanlı nesne takibi yöntemleri tek boyutlu ve çok boyutlu olmak üzere iki alt sınıf altında incelenir. Takip edilecek nesnenin görünüm bilgisi aşağıda gösterildiği gibidir;

$$X_i = \begin{pmatrix} X_{i1} \\ \vdots \\ X_{id} \end{pmatrix}, \quad i = 1, \dots, n \quad (2.12)$$

Buna göre, $d = 1$ olması durumunda takip edilecek nesnenin görünüm bilgisi tek boyutlu, $d > 1$ olması durumunda ise çok boyutludur.

2.4.2.1. Tek boyutlu çekirdek yoğunluk tahmini

Tek boyutlu X_i veri kümesinin herhangi bir x veri elemanının X dizisi üzerindeki katkısı (olasılık yoğunluk değeri veya yoğunluk tahmini) aşağıdaki şekilde hesaplanmaktadır;

$$f(x) = \frac{1}{n} \sum_{i=1}^n K\left(\frac{x-X_i}{h}\right) \quad (2.13)$$

Denklem (2.13)'teki $K(\cdot)$ çekirdek fonksiyonu, x merkez noktasına ve h bant genişliğine sahiptir. Bu denklemde $\int K(t)dt = 1$ eşitliğinin sağlanma zorunluluğu vardır.

Comaniciu [93] yapmış olduğu çalışmasında, çekirdek tabanlı nesne takibi alanında birçok çalışmaya temel teşkil edecek bir nesne takip yöntemi önermiştir. *Ortalama kayma* (Bölüm 3.3'te detaylıca bahsedilmiştir) olarak adlandırılan bu yöntemde, ilk olarak takip edilen nesnenin uzaysal görünümü olasılık yoğunluk değerleri şeklinde ifade edilmektedir. Daha sonra bu yoğunluk dağılımının zirve noktası (nesnenin merkezi) imgeler boyunca takip edilmeye çalışılmaktadır. Nesneye ait yoğunluk dağılımındaki tepe noktasının imgeler boyunca sürekli hareket halinde olduğu düşünülmektedir.

[99]'da önerilen yöntemde, model ve hedef dağılımlara ait yumuşatılmış çekirdek yoğunluk tahminleri arasında simetrik bir benzerlik fonksiyonu kullanılarak nesne takibi işlemi gerçekleştirilmiştir. Model ve hedef görünümüne ait olasılık yoğunluk fonksiyonlarının merkez noktaları üzerinde yapılan bu yumuşatma işlemi, hesaplama maliyetini azaltmakla birlikte gürültü duyarlılığının azalmasını da sağlamıştır.

[100]'de önerilen nesne takip yönteminde, model ve hedef dağılımlar üzerinde var olan arka plan bilgisi çıkarılır ve elde edilen sonuçlar yeni bir benzerlik fonksiyonuna girdi olarak verilir. Nesne takibi için tekrarlı bir şekilde ortalama kayma algoritması kullanılmaktadır.

2.4.2.2. Çok boyutlu çekirdek yoğunluk tahmini

Çok boyutlu bir veri kümesinin olasılık yoğunluk dağılımı şu şekilde hesaplanır;

$$\begin{aligned}\hat{f}_h(x) &= \frac{1}{n} \sum_{i=1}^n \frac{1}{h^d} K\left(\frac{x-X_i}{h}\right) \\ &= \frac{1}{n} \sum_{i=1}^n \frac{1}{h^d} K\left(\frac{x_1-X_{i1}}{h}, \dots, \frac{x_d-X_{id}}{h}\right)\end{aligned}\quad (2.14)$$

Denklem (2.14)'de kullanılan bant genişlikleri $h = (h_1, \dots, h_d)^T$ şeklinde açık yazılabilir. Buna göre Denklem (2.14) tekrar düzenlenerek aşağıdaki gibi yazılabilir;

$$\hat{f}_h(x) = \frac{1}{n} \sum_{i=1}^n \frac{1}{h_1 \dots h_d} K\left(\frac{x_1-X_{i1}}{h_1}, \dots, \frac{x_d-X_{id}}{h_d}\right)\quad (2.15)$$

[95]'de gerçekleştirilen nesne takip uygulamasında, ortalama kayma yöntemi birden fazla nesnenin aynı anda takip edilmesinde kullanılmıştır. Ortalama kayma yöntemi, nesne büyüklüğünün imge içerisinde sabit boyutta olduğunu varsayar. Oysa takip edilen

nesnenin kameraya yaklaşp uzaklaşması, imge içerisinde kapladığı alanın sürekli değişmesine yol açar. Bu durum, nesne takip yöntemlerinin birçoğunun takip işlemini sürdürememesine neden olur. [96-98]'de, ortalama kayma algoritmasının bu dezavantajına çözüm olabilecek yöntemler sunulmuştur. [96] ve [97]'de, ortalama kayma algoritması simetrik olmayan çekirdek fonksiyonları ile birlikte kullanılmıştır. Asimetrik çekirdek fonksiyonunun kullanımı, nesneye ait olmayan arka plan piksellerinin nesne modeli dışında kalmasına neden olur. Böylece [93]'deki takip yönteminin performansı daha da artmış olur. [98]'de, ortalama kayma ve parçacık süzgeçleme yöntemlerinin birlikte kullanıldığı yeni bir nesne takibi yöntemi önerilmiştir. Bu yöntem sayesinde nesnenin imge içerisinde büyüyüp küçülmesi sonucunda ortaya çıkan problem çözülebilmektedir. Bahsedilen yöntem iki aşamadan oluşmaktadır. İlk olarak imgeler arasındaki nesne değişimi Δx ortalama kayma yöntemi kullanılarak hesaplanır. Bütün halinde bir sonraki imgeye taşınan ve kenarları tam üst üste oturmamış nesnelerin örtüşmesini sağlayabilmek için nesne kenarları üzerine yerleştirilen işaretçiler parçacık süzgeci yöntemi kullanılarak en uygun pozisyonlara yerleştirilir. Böylece nesnenin görünümünün imge içerisinde büyüyüp küçülme problemi ortadan kalkmış olur.

2.4.2.3. *Bant genişliği seçimi*

Bir çekirdek fonksiyonunun tahmin kalitesi, $K(.)$ fonksiyonunun seçiminden daha çok h bant genişliğinin belirlenmesine bağlıdır [91]. Uygun bant genişliğinin seçimi nesne takip edicinin takip kararlılığı açısından çok önemlidir. Bant genişliği değeri, nesne yüzeyindeki renk değişim yapısına ait nicel bir bilgiyi ifade etmektedir. Bu değerin küçük seçilmesi, nesne yüzeyindeki renk yumuşaklığının azalmasına neden olur. Büyük seçilmesi yumuşaklığı artırır [92].

Uygun bant genişliği değerini elde etmek için sık kullanılan genel bir yöntem Denklem (2.16 ve 2.17)'de verilmiştir. Bu eşitlikler yardımıyla dağılımın ortalama karesel hatası ($AMISE$) değeri minimum seviyeye indirgenir. Bu değer, takip edilecek nesneye ait görünüm bilgisinden elde edilir. Bant genişliğinin belirlenmesinde $AMISE$ değerinin kullanılması, nesne üzerindeki tüm karakteristik özelliklerin korunması anlamına gelir.

$$h_{optimal} = argmin AMISE \quad (2.16)$$

$$AMISE(h) = \int_{-\infty}^{\infty} E \left(\hat{f}_h(y) - f_h(y) \right)^2 dy \quad (2.17)$$

Dağılımın $AMISE$ değerinin hesaplanmasında kullanılan f fonksiyonu bilinmeyen yoğunluğu ifade eder. $\hat{f}$, f fonksiyonunun n . örneğine dayalı tahmin bilgisidir. E simgesi ile beklenen değer ifade edilir.

2.4.3. Siluet Tabanlı Nesne Takibi

Kare veya elips gibi basit geometrik şekiller ile ifade edilemeyen, el, baş, omuz gibi karmaşık geometrik şekilli nesnelerin takibi için siluet tabanlı yöntemler, uygun bir nesne sunumu sağlamaktadır. Siluet tabanlı nesne takip edicilerin amacı, önceki imgeler kullanılarak üretilen nesne modelini güncel imge içerisinde bulabilmektir [28, 29]. Nesne modeli olarak nesnenin renk histogramı, nesne kenarı veya sınırı seçilebilir. Siluet tabanlı nesne takip edicileri *şekil karşılaştırmacılar* ve *sınır takip ediciler* olarak iki alt kategori altında incelemek mümkündür. Şekil karşılaştırmalı yaklaşım, güncel imge içerisinde nesne siluetini aramaktadır. Sınır izlemeli yaklaşım ise bazı enerji fonksiyonlarının minimum seviyeye indirgenmesiyle nesnenin güncel imgedeki yeni yerini belirlemeye çalışır. Bunun için başlangıç aşamasında belirlenen nesne sınırları kullanılmaktadır. Bu yöntemlerin her ikisi de geçici bir alana, nesne bölütleme işlemini uygulayan yöntemler şeklinde düşünülebilir. Yine her iki yöntemde, güncel nesne konumlarının belirlenmesi için öncesel bilgiye ihtiyaç duymaktadır [30, 102, 103].

2.4.3.1. Şekil karşılaştırmacılar

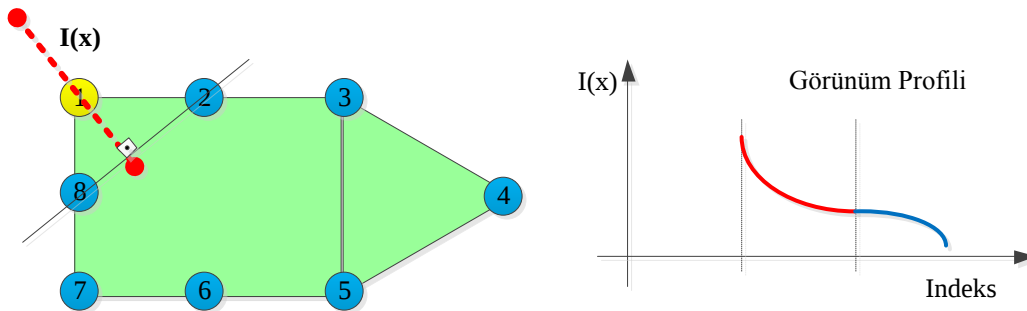
Şekil karşılaştırmacı yöntemlerde nesne siluetinin güncel imgeden bir sonraki imgeye aktarıldığı varsayılır. Güncel imgedeki nesne silueti ve onunla ilişkili nesne modeli bir sonraki imgede aranır. Arama işlemi, bir önceki imgede elde edilen nesne silueti ve modeli ile güncel imgede yakalanan nesneler arasındaki benzerlik değerleri hesaplanarak gerçekleştirilir. Bu nedenle, silueti net olmayan (örneğin bulut gibi) nesneler doğru ve sağlam takip edilemez. Genellikle nesne modeli için kenar haritası kullanılır. İmgeler arasında geçiş yaparken siluet ve kenar üzerindeki değişim göz önüne alınarak nesne modelinin sürekli güncellenmesi gerekir.

2.4.3.2. Sınır takibi

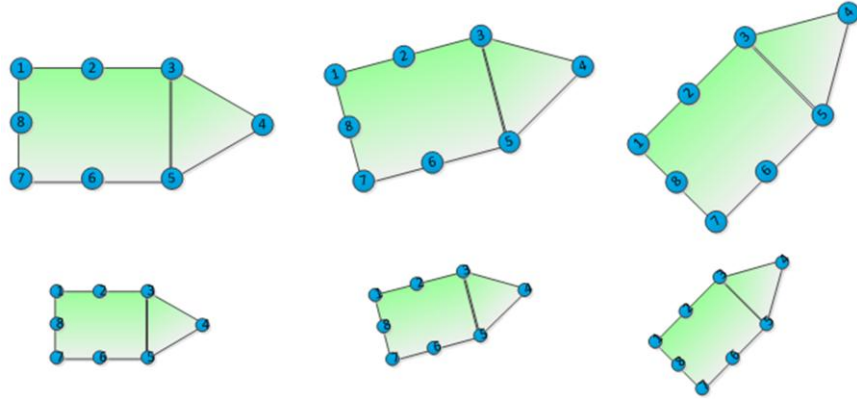
Sınır takibi yöntemleri, güncel imgedeki sınır çizgilerini başlangıç kabul ederek bir sonraki imgede yeni nesne sınırlarını aramaktadır. Sınır takibi algoritmalarının uygun sonuç verebilmesi için güncel imgedeki nesne sınırları ile bir sonraki imgede bulunan nesne sınırlarının kesişimi sıfır olmamalıdır.

Sınır değerlendirme algoritmaları nesne sınırları üzerine işaretçiler koyarak her bir işaretçinin görünüm profilini belirlemeye çalışır [104]. Şekil 2-7’de nesne üzerindeki bir işaretçiye ait görünüm profilinin nasıl elde edildiği gösterilmiştir. İşaretçi görünüm profilinin hesaplanması için işaretçiye komşu olan ve ön ve arkasında bulunan diğer iki işaretçi kullanılmaktadır. Bu iki işaretçiden geçen doğruya dik olan doğru bulunur. Son olarak da bulunan bu dik doğru üzerinde belirli bir uzunluktaki görünüm bilgisi kodlanır.

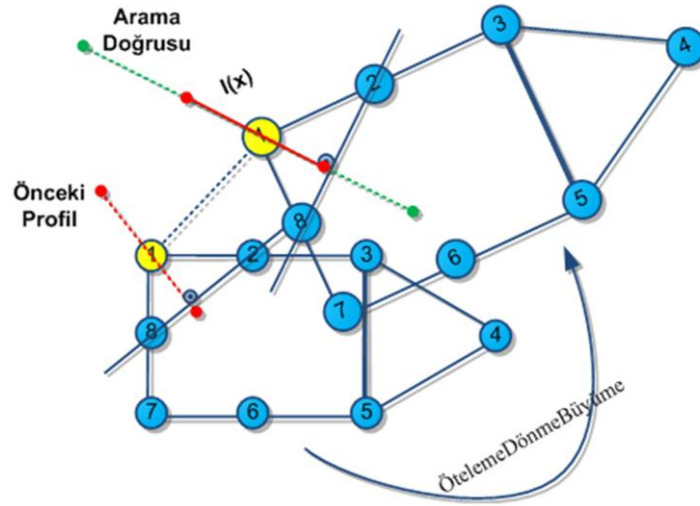
Sınır işaretçilerinin tek bir görünüm profili ile ifade edilmesi bu noktaların takip edilebilme ihtimalini azaltır. Bu nedenle her bir işaretçi için birden fazla görünüm profil bilgisine ihtiyaç duyulur. [103]’de önerilen nesne takip yöntemi, işaretçiler için birden fazla görünüm profili ihtiyacını, nesnenin farklı çözünürlükteki sunumlarını kullanarak sağlamıştır. Önerilen yöntem sayesinde, aynı nesneye ait birden fazla imge elde etmek yerine tek bir imgenin farklı çözünürlükteki görüntüleri elde edilmiş ve her bir işaretçi için farklı çözünürlük seviyelerindeki görünüm profilleri kullanılmıştır. Şekil 2-8’de basit bir nesnenin farklı çözünürlükteki sunumları gösterilmiştir. Son olarak, her bir işaretçinin görünüm profili bir sonraki imge içerisinde aranır. Yapılan bu işlem Şekil 2-9’da gösterilmektedir. Önceki imgede bulunan işaretçi profili bir sonraki imge üzerinde büyük bir arama doğrusu üzerinde gezdirilir. Maksimum benzerliğin sağlandığı nokta yeni işaretçiye ait olan noktadır. Bu nokta bir sonraki imge içerisinde bulunması istenilen noktadır. Nesne hareketinin algılanabilmesi için tüm işaretçilerin sonraki imgedeki yerleri tespit edilmektedir.



Şekil 2-7. Bir işaretçinin görünüm profilinin elde edilmesi.



Şekil 2-8. Basit bir nesnenin farklı çözünürlük seviyelerindeki sunumu.



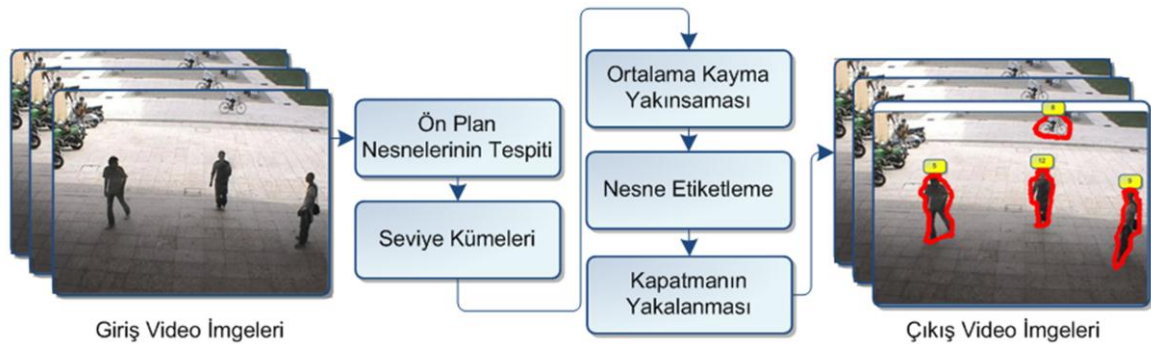
Şekil 2-9. Ardışık imgelerde işaretçi arama işlemi.

3. KALABALIK ORTAMLARDA ÇOKLU İNSAN TAKİP YÖNTEMİNİN GELİŞTİRİLMESİ

Bu bölümde, kalabalık ortamlarda çoklu insan takibini gerçekleştirebilen melez esaslı yeni bir takip yöntemi geliştirilmiştir. Önerilen takip yönteminin amacı, sabit bir arka plan görünümüne sahip video imgelerinde hareket halinde olan kişilerin her birine eşsiz bir etiket değeri vermek ve kişileri kamera görüş alanı içerisinde kaldıkları süre boyunca aynı etiket numarası ile takip edebilmektir.

Geliştirilen çoklu insan takip yönteminin akış diyagramı Şekil 3-1’de verilmiştir. Buna göre, geliştirilen çoklu insan takip yöntemi ardışıl beş işlemin tekrarlı kullanımından oluşmaktadır: 1) Ön plan nesnelerinin tespiti; 2) Ön plan nesnelerinin seviye kümeleri yöntemi kullanılarak sunumu; 3) Ortalama kayma yöntemi kullanılarak nesnelerin takip edilmesi; 4) Takip edilen her bir nesnenin etiketlenmesi; 5) Kapatma probleminin çözülmesi.

Önerilen çoklu insan takip yöntemi, sağlam takip performansı sağlayan ve Bölüm 3.3’te detaylı bahsedilen ortalama kayma yönteminin geliştirilmiş bir biçiminin kullanılmasına dayanmaktadır. Geleneksel ortalama kayma yönteminde, video boyunca takip edilecek nesnenin boyut ve eğim bilgisi sabit olarak kabul edilmektedir [93, 100]. Bu nedenle, klasik ortalama kayma yönteminde nesne sunumu için dikdörtgen veya elips gibi simetrik bir geometrik şekil kullanılmaktadır. Oysa hareketli bir nesnenin takibi esnasında nesnenin boyut ve eğim bilgisi değişebilir. Bundan dolayı takip edilecek nesnenin geometrik şekli simetrik olmayabilir. Bu tür durumlarda klasik ortalama kayma yöntemi, takip edilen nesneyi kaybeder ve istenmeyen takip sonuçlarının elde edilmesine yol açar. Buna karşılık, geliştirilen çoklu insan takip yöntemi içerisinde kullanılan ortalama kayma



Şekil 3-1. Geliştirilen çoklu insan takip yönteminin akış diyagramı.

yönteminde, nesneyi ifade etmek için basit geometrik şekiller kullanmak yerine takip boyunca hareketli nesnenin siluet bilgisi kullanılmaktadır. Bu bilgi ortalama kayma yönteminden önce ön plan nesnelerinin tespiti işleminin gerçekleştirilmesi ile elde edilir. Böylece, nesnenin boyut ve eğim bilgisinin değişmesine rağmen, takip işlemi başarılı bir şekilde devam edebilmektedir.

Klasik ortalama kayma yönteminde, nesnenin renk sunumu için olasılık yoğunluk fonksiyonu kullanılmaktadır. Olasılık yoğunluk fonksiyonu, yüksek hesaplama maliyetine sahip güçlü bir nesne sunumu sağlamaktadır. Bu durum nesne takip edicinin sağlamlılığını arttırmasına karşın bu takip edicinin gerçek zamanlı takip sistemlerinde tercih edilme olasılığını azaltır. Buna karşın, geliştirilen çoklu insan takip yöntemi, kişilere ait görünüm bilgisini olasılık yoğunluk fonksiyonu ile değil seviye kümeleri yöntemi ile ifade etmektedir. Bu sayede takip edicinin daha hızlı çalışmasına karşılık sağlamlılığında önemli ölçüde bir değişikliğin olmadığı görülmektedir.

Klasik nesne takip edicilerin [93, 100] en büyük dezavantajlarından birisi kapatma olarak bilinen görünüm engellenmesi problemidir. Geleneksel nesne takip ediciler bu tür durumlara karşı tepkisizdir. Geliştirilen çoklu insan takip yönteminde, kısmi görünüm engellenmesi problemine farklı bir çözüm yöntemi sunulmaktadır. Bu yöntem kullanılarak kalabalık bir ortamda birbiri görünümünü kısmi engelleyen insanlar olmasına rağmen, takip işlemi başarılı bir şekilde sürdürülebilmektedir.

Bölümün bundan sonraki kısmında önerilen çoklu insan takip yönteminin her bir aşaması detaylı bir şekilde ifade edilmiştir.

3.1. Ön Plan Nesnelerinin Tespiti

Bu bölümde, imge içerisinde hareket etmekte olan ön plan nesnelerinin dinamik olarak değişebilen arka plan görünümünden nasıl ayırt edildiği hakkında bilgiler verilmektedir. Buna göre, renkli video imgesinin her pikseli, dinamik arka plan çıkarım algoritmaları yardımıyla siyah/beyaz sınıflarından birine atanmaktadır. Elde edilen sonuç imgesinde, siyah pikseller hareketli ön plan nesnelerini gösterirken beyaz pikseller hareketsiz arka plan nesnelerini göstermektedir.

Ön plan nesnelerinin tespit yöntemleri tekrarlamalı ve tekrarlamasız olmak üzere iki kategori altında incelenmektedir [105]. Tekrarlamasız teknikler, ön ve arka plan nesnelerini ayırt etmek için kayan pencere yaklaşımını [106] kullanmaktadır. Bu yaklaşım

ile boyutu belirlenmiş bir pencere, video imgeleri üzerinde gezdirilmektedir. Bu yaklaşımın dezavantajı, hafızaya yüklenmiş sabit sayıdaki video imgesi üzerinde çalışabilmesidir. Tekrarlamalı tekniklerde, geçmiş imgeler hafızaya yüklenmez ve tek bir arka plan modeli belirlenerek zamanla güncellenir. Bu sayede tekrarlamalı tekniklerin hafıza kullanımı çok daha iyidir. Literatürde farklı süzgeçleme teknikleri (Median, Kalman, Gaussian karışımı) kullanılarak ön plan nesnelerinin tespiti yapılmaktadır. Kalman süzgeci yöntemi, arka plan modelini tek bir Gaussian fonksiyonu ile ifade ederek bu dağılımın değişimini imgeler boyunca takip eder. Gaussian karışımı yönteminde ise imgedeki her bir pikselin renk dağılımı için ayrı bir Gaussian modeli kullanılmakta ve tüm modellerin eş zamanlı çalışması sağlanmaktadır.

Klasik ön plan çıkarım yöntemlerinde piksel olasılığının hesaplanabilmesi için t anında RGB renk uzayındaki x^t pikselinin güncel renk değeri aşağıdaki gibi kontrol edilmektedir;

$$p(x^t|BG) > \epsilon \quad (3.1)$$

Denklem (3.1)'deki ϵ değişkeni, güncel eşik değerini göstermektedir ve çok küçük bir sabit değere sahiptir. Bu değerinin sabit olması, sınıflandırıcının birçok durumda istenilen sonucu üretememesine neden olur. Çünkü ortamın arka plan görünümü zamanla değişebilir (örneğin, ışık yoğunluğunun gün boyu değişebilmesi). Bu nedenle dinamik arka plan çıkarım tekniklerine ihtiyaç duyulmaktadır.

Geliştirilen çoklu insan takip yönteminde, piksel seviyesinde Gaussian karışım modelini kullanan dinamik arka plan çıkarım yöntemi [50, 108] kullanılmıştır. Bahsedilen yöntemde, her bir piksel için olasılıksal bir model inşa edilmekte ve güncel imgenin piksel değerleri bu modeller kullanılarak sınıflandırılmaktadır. [50] numaralı çalışmada geliştirilen arka plan çıkarım tekniği imgedeki her bir piksel için B bileşenli Gaussian karışım modelini kullanarak sınıflandırma işlemini gerçekleştirmektedir. Bu karışım modeli aşağıdaki gibi ifade edilmektedir;

$$p(x|X_T, BG) \sim \sum_{m=1}^B \pi_m \eta(x, \mu_m, \sigma_m^2) \quad (3.2)$$

$$\eta(x, \mu, \sigma^2) = \eta\left([x_r, x_g, x_b], [\bar{\mu}_r, \bar{\mu}_g, \bar{\mu}_b], [a_r, a_g, a_b]^2\right)$$

$$\propto \exp \left[-\left(\frac{x_r - \mu_r}{\sigma_r} \right)^2 - \left(\frac{x_g - \mu_g}{\sigma_g} \right)^2 - \left(\frac{x_b - \mu_b}{\sigma_b} \right)^2 \right] \quad (3.3)$$

Denklem (3.2)'de kullanılan μ_m ve σ_m değişkenleri sırasıyla GMM'nin ortalama ve varyans değerlerini ifade etmektedir. Denklem (3.2)'deki π değişkeni, GMM'de kullanılan bileşenlerin ağırlık değerlerini göstermektedir. GMM algoritmasında her piksel için gerekli olan bileşen sayısı otomatik hesaplanmakta ve π, μ ve σ değişkenleri uygun bir tekrarlı fonksiyon kullanılarak güncellenmektedir. Güncelleme işlemi, imge içerisindeki hareketsiz pikseller üzerinde gerçekleştirilmektedir. Bu sayede istenmeyen arka plan görünümü ön plan nesnelerinden ayrılmaktadır. Bu yöntemin yanında daha karmaşık arka plan çıkarım modelleri de mevcuttur [107].

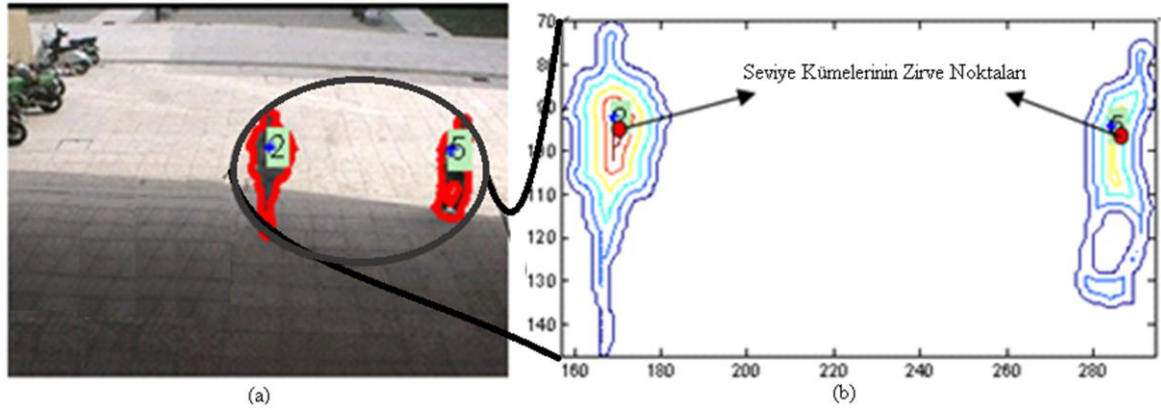
3.2. Seviye Kümeleri Yardımıyla Nesne Sunumu

Seviye küme yöntemi, basit olmakla birlikte şekil ve modellerin takip edilebilmesinde çok güçlü sayısal bir sunum sağlayabilmektedir [109, 110]. İki boyutlu bir düzlemde seviye kümesi sunumu kapalı bir Γ eğrisi tarafından gerçekleştirilmektedir. İki boyutlu φ fonksiyonunun sıfır seviyeli kümesi aşağıdaki gibi ifade edilmektedir;

$$\Gamma = \{(x, y) | \varphi(x, y) = 0\} \quad (3.4)$$

Denklem (3.4)'teki φ fonksiyonu, seviye kümesi fonksiyonu olarak bilinir ve Γ eğrisi ile sınırlanan nesne alanının iç kısmında kalan pikselleri, nesne kenarına uzaklığı oranında pozitif bir değer ile ifade eder. Nesne dışındaki tüm pikseller sıfır alınarak göz ardı edilir. Seviye kümesi sunumu, ortalama kayma algoritmasında kullanılan olasılık yoğunluk sunumu kadar güçlü bir destek sağlayamaz. Çünkü bir nesnenin seviye kümesi kullanılarak sunumunun gerçekleştirilebilmesi için sıfır seviyesine sahip kapalı Γ eğrisinin bilinmesi gerekir. Nesnenin dış hat çizgisi (kenarı) ile birlikte nesne dışında kalan kısmın sıfır olması ve nesne içindeki piksellerin dış hat çizgisine olan uzaklığına göre pozitif değer alabilmesi için seviye küme fonksiyonu aşağıdaki gibi ifade edilmektedir;

$$\vartheta(x) = C\varphi(x)I(x) \quad (3.5)$$



Şekil 3-2. Yakalanan hareketli insan nesneleri için seviye kümesi gösterimi **(a)** Arka plan çıkarım algoritması sonucu **(b)** Seviye küme sunumu.

Denklem (3.5)'teki C katsayısı normalizasyon terimidir ve $\frac{1}{\sum \varphi(x)I(x)}$ formülü kullanılarak hesaplanmaktadır. Seviye kümelerinin ortalama kayma algoritması içerisinde kullanılması ile yoğunluk tahmini aşağıdaki gibi hesaplanmaktadır;

$$f(x) = \frac{1}{n} \sum_{i=1}^n \vartheta(x_i) \quad (3.6)$$

Şekil 3-2'de, seviye kümesi yöntemi kullanılarak örnek bir video imgesi içerisinde hareket etmekte olan iki kişiye ait nesne sunumu gösterilmektedir.

3.3. Ortalama Kayma Algoritması

Ortalama kayma algoritması parametrik olmayan bir sınıflandırma tekniğidir. Bu algoritma, nesneye ait olasılık dağılımının en yoğun olduğu yeri bulmayı sağlamaktadır. Bu işlem yapılırken, sınıf sayısını belirten öncesel bir bilgiye ihtiyaç duyulmaz ve nesnenin şekli üzerinde herhangi bir sınırlama gerçekleşmez. Ortalama kayma algoritmasının bahsedilen bu özellikleri nesne takip uygulamalarında onu sürekli cazip kılmıştır. Bu algoritma ilk olarak Fukunaga ve Hostetler [111] tarafından önerilmiş ve daha sonra Comanicu [93] tarafından geliştirilmiştir.

Ortalama kayma algoritmasının çalışma prensibi şu şekildedir: İmgedeki hareketli bir nesnenin t ve $t + 1$ anlarındaki başlangıç ve bitiş pozisyonlarının $\hat{x}$ ve $\hat{x}'$ olarak ifade edildiği düşünülmektedir. Buna göre iki imge arasındaki nesne hareketi $\hat{x}' = \hat{x} + \Delta x$ eşitliği ile ifade edilmektedir. Bu eşitlikte Δx , nesnenin imgeler arasındaki

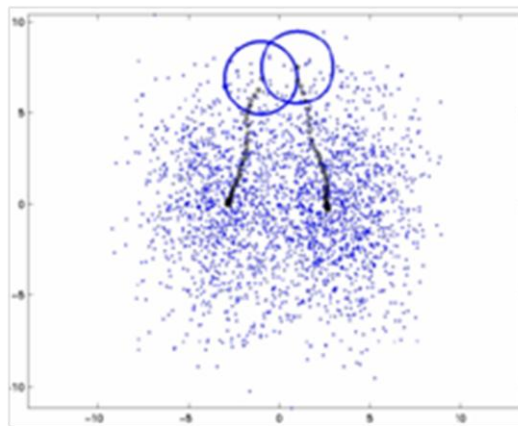
kayma vektörü olup nesne merkezinin gerçekleştirmiş olduğu hareketin mesafe ve yön bilgisini içermektedir. Buna göre Δx kayma vektörü aşağıdaki gibi hesaplanmaktadır;

$$\Delta x = \frac{\sum_i K(x_i - \hat{x}) w(x_i) (x_i - \hat{x})}{\sum_i K(x_i - \hat{x}) w(x_i)} \quad (3.7)$$

Denklem (3.7)'deki $K(\cdot)$, nesne alanını ifade eden h bant genişliğine sahip çekirdek fonksiyonudur ve radyal simetrik bir yapıya sahiptir. w ağırlık katsayısı ise Denklem (3.8)'de verilen Bhattacharya benzerlik formülü [94] kullanılarak elde edilmektedir. Model ve aday nesnelerin renk dağılım fonksiyonlarının p ve q olduğu düşünüldüğünde, x pikselindeki Bhattacharya ağırlık katsayısı aşağıdaki denklemde belirtildiği gibi hesaplanmaktadır;

$$w(x) = \sqrt{q(l(x))/p(l(x))} \quad (3.8)$$

Denklem (3.8)'deki w ağırlık değeri kullanılarak hesaplanan Δx kayma vektörü, nesnenin olasılık yoğunluk dağılımı üzerinde tepe noktaya doğru hareket etmekte ve her harekette vektör boyutu küçülmektedir. Hareket vektörü yoğunluk dağılımının tepe noktasına vardığında sifıra yakın bir değer almaktadır. Hareket vektörünün sifıra yakın olması, nesnenin merkez noktasına ulaşıldığı anlamını taşımaktadır. Şekil 3-3'te iki farklı nesneye ait olasılık yoğunluk dağılımları gösterilmekte ve dağılımlar üzerinde rastgele alınan iki noktanın ortalama kayma yöntemini kullanarak dağılımların tepe noktalarına nasıl ulaştığı gösterilmektedir. Tepe noktasına yaklaştıkça hareket vektörünün küçüldüğüne dikkat edilmelidir.



Şekil 3-3. Ortalama kayma vektörünün hareketi

3.4. Veri Bağı Problemi ve Kişi Etiketleme

Veri bağı problemi, ardışıl iki imgede yakalanan nesnelerin aynı etiket numarası ile birbirlerine bağlanabilmeleri olarak ifade edilmektedir. Birden fazla insanın hareket halinde olduğu gerçek ortamlardan elde edilen video imgeleri birbirinden ayrık olduğu için veri bağı problemi meydana gelmektedir. Önerilen çoklu insan takip yönteminde, imgeler arasındaki veri bağı probleminin çözülebilmesi için yeni bir yöntem kullanılmaktadır. Bu yöntem sayesinde imge içerisinde tespit edilen her bir insana hangi etiket numarasının verileceği belirlenmektedir.

Geliştirilen çoklu insan takip yöntemi, veri bağı probleminin çözülebilmesi için şu şekilde bir yol izlemektedir: t anında imgede yer alan nesnelerin durum vektörü θ olsun. Durum vektörü, imgedeki nesne sayısını, nesnelerin imge içerisindeki pozisyon bilgisini ve nesne pozunu parametre bilgilerini içerir. Buna göre θ durum vektörü şu şekilde ifade edilir;

$$\theta^{(t)} = \{\text{Obj}_1, \text{Obj}_2, \dots, \text{Obj}_n\} = \{(k_1, m_1), \dots, (k_n, m_n)\} \in \Theta_n. \quad (3.9)$$

Bu denklemde k_i , i . nesneye atanan eşsiz etiket değeridir ve bu nesnenin tüm parametreleri m_i ile gösterilir. Θ ile çözüm uzayı kastedilir. t anındaki nesnelerin etiketlenebilmesi için $t - 1$ anındaki tanımlanan nesne bilgileri ve t anındaki m_i parametreleri birlikte kullanılır. Buna göre, $t - 1$ ve t imgelerindeki nesneler arasında üç farklı nesne ilişkisinin olduğu kabul edilir: 1) İlişkili nesne; 2) Yeni nesne, 3) Ölü nesne. Varsayılan nesne ilişkileri aşağıdaki gibi ifade edilebilir;

$$\text{Obj}_i \propto \begin{cases} \text{İlişkili nesne} & k_i^{(t)} = k_i^{(t-1)} \\ \text{Yeni nesne} & m_i^{(t-1)} = \emptyset \\ \text{Ölü nesne} & m_i^{(t)} = \emptyset \end{cases} \quad (3.10)$$

Denklem (3.10)'daki $k_i^{(t)} = k_i^{(t-1)}$ ifadesi, $k_i^{(t)}$ nesnesinin takip edilen bir nesne olduğunu, $m_i^{(t-1)} = \emptyset$ ifadesi $k_i^{(t-1)}$ nesnesinin yeni bir nesne olduğunu, $m_i^{(t)} = \emptyset$ ifadesi ise $k_i^{(t-1)}$ nesnesinin ölü bir nesne olduğunu göstermektedir.

3.5. Kapatmanın Yakalanması

Kapatma olarak tabir edilen nesne görünümünün engellenmesi problemi, takip edilen bir nesnenin görünüm bilgisinin kaybedilmesi anlamına gelmektedir. Bu problemin meydana gelmesi, nesne takip edicilerin takip kararını verebilmede zorlandığı andır.

Takip edilen bir nesnenin görünümü üç farklı şekilde engellenebilir [3]: 1) Nesnenin kendisi tarafından engellenmesi; 2) Bir başka nesne tarafından engellenmesi; 3) Arka plan görünümü tarafından engellenmesi. Şekil 3-4(a)'da, nesnelerin birbirlerinin görünümünü engellemesi durumu gösterilmektedir. Şekilde görüldüğü gibi, arka plan çıkarım algoritması hareket eden iki insan nesnesini tek bir hareketli bölge olarak tespit etmiştir. Bir başka ifadeyle, önceki imgede etiketlenen iki insan nesnesi güncel imgede tek bir hareketli nesne olarak algılanmıştır. Dolayısıyla, kameraya daha yakın olan insan görünümü daha uzakta olan insan görünümünü engellemektedir.

Geliştirilen çoklu insan takip yöntemi, kısmi kapatma probleminin gerçekleştiği



(a)



(b)

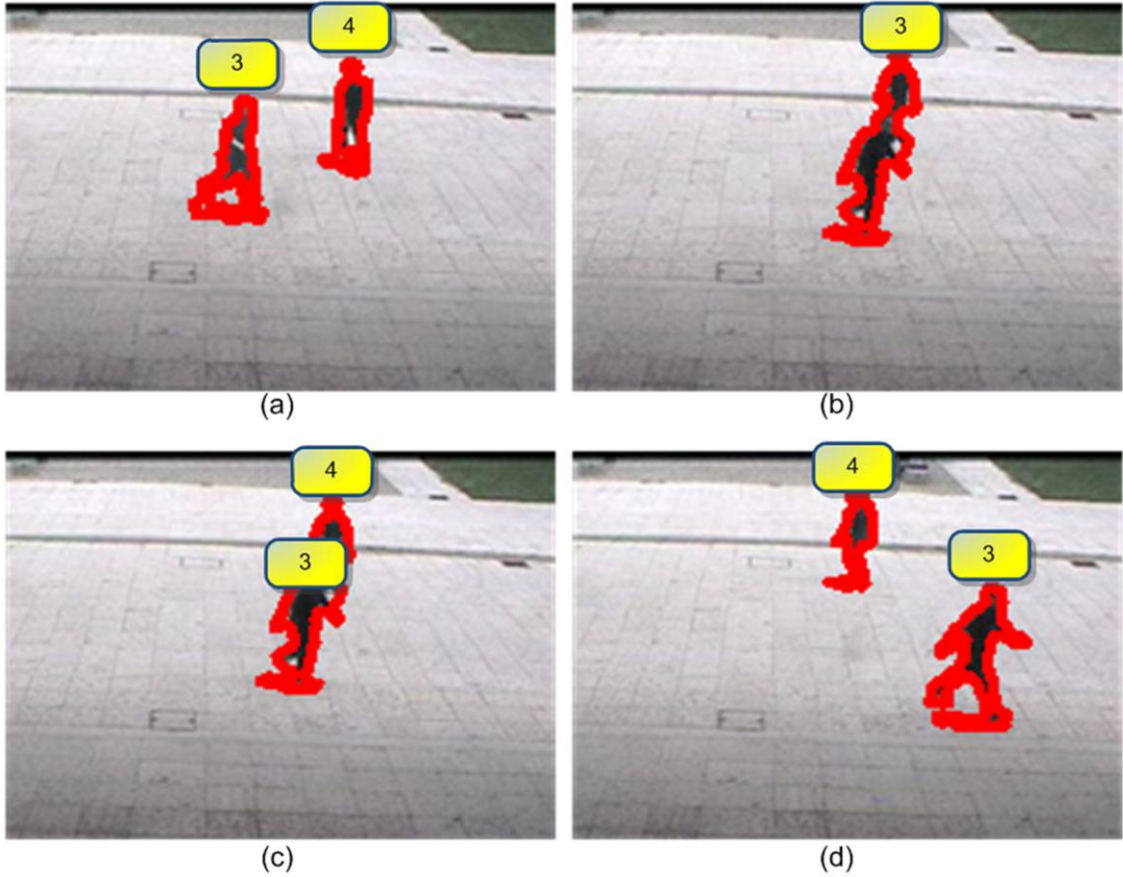
Şekil 3-4. Ortalama kayma yöntemi kullanılarak kısmi kapatma probleminin giderilmesi **(a)** Hareket halindeki iki insanın birbirlerinin görünümünü engellemesi, **(b)** Güncel imgedeki zirve noktalarının bulunması için önceki imgede elde edilen başlangıç noktalarının ortalama kayma algoritmasında kullanımı ve hareket vektörlerinin tespiti.

imgelerde kapatmadan önceki takip işlemlerini doğru bir şekilde devam ettirebilmek için yeni bir yöntem kullanmaktadır. Bu yöntem, seviye kümesi sonucu üzerinde ortalama kayma algoritmasının belirli bir yarıçap kullanılarak çalışmasına dayanmaktadır. Buna göre, kapatma olayından etkilenen her kişi için Denklem (3.11)'de gösterilen r yarıçap değeri hesaplanarak seviye kümesi sonucu ile birlikte ortalama kayma algoritmasına gönderilmektedir.

$$r_i = \frac{h_i + w_i}{(\text{toplam kişi sayısı})^\varepsilon}, \quad i = 1 \dots N \quad (3.11)$$

Denklem (3.11)'deki h_i ve w_i değişkenleri, kapatmanın meydana geldiği i . insan silüetinin yükseklik ve genişlik uzunluklarıdır. *toplam kişi sayısı* değişkeni, kapatma olayından etkilenen toplam kişi sayısını ifade etmektedir. ε değişkeni ise optimum yarıçap değerini elde edebilmek için kullanılan katsayıdır ve optimizasyon teknikleri kullanılarak elde edilebilir. Bu tez çalışmasında gerçekleştirilen uygulamada $\varepsilon = 2$ olarak belirlenmiştir. Böylece, takibi gerçekleştirilen her bir kişi için bir yarıçap değeri elde edilmektedir. Daha sonra bu yarıçap bilgisi seviye küme sonucu ile birlikte ortalama kayma algoritmasına gönderilerek kişilerin güncel imgedeki pozisyon bilgileri başarılı bir şekilde hesaplanmaktadır. Şekil 3-4(b)'de, kısmi kapatma durumunda geliştirilen çoklu insan takip yönteminin davranışı gösterilmiştir. Ancak önerilen bu yöntem sadece kısmi kapatmanın gerçekleştiği durumlarda çalışmaktadır. Geliştirilen insan takip yöntemi, tam kapatma durumunda istenilen sonucu vermeyi garanti etmemektedir.

Şekil 3-5'te, kısmi kapatma durumunun olduğu örnek bir video imgesi gösterilerek bu durumda geliştirilen insan takip yönteminin davranışı gösterilmiştir. Şekil 3-5(a)'da, yakalanan iki insan silüetine 3 ve 4 numaralı etiketler verilmiştir. Şekil 3-5(b)'de, 3 numaralı insan silüeti ile 4 numaralı insan silüetinin kısmi kapatma anı gösterilmektedir. Bu anda 4 numaralı insana ait etiket değeri kaybolmuştur. Şekil 3-5(c)'de, kısmi kapatmanın meydana geldiği bu imge üzerinde geliştirilen çoklu insan takip yönteminin çalıştırılması ile elde edilen sonuç gösterilmiştir. Görüldüğü gibi Şekil 3-5(b)'de etiket numarası kaybolan insanın tekrar orijinal etiket numarası elde edilmiştir. Böylece, kalabalık bir ortamda hareket halinde olan insanların birbirlerinin görünümünü kısmi engellemeleri durumunda, önerilen çoklu insan takip yönteminin takip işlemini kararlı bir şekilde devam ettirebildiği görülmektedir.



Şekil 3-5. Geliştirilen çoklu insan takip yönteminin kısmi kapatma olayı anındaki davranışı **(a)** Görünüm engellemesinden önceki durum **(b)** Engellemenin olduğu ve 4 numaralı insanın kaybolduğu durum **(c)** Geliştirilen algoritmanın görünüm engellemesi olduğu andaki sonucu **(d)** Engellemeden sonraki durum.

3.6. Uygulama Sonuçları ve Değerlendirilmesi

Çoklu insan takibi konusunda önerilen her hangi bir yöntemin aşağıda belirtilen ihtiyaçları tam anlamıyla cevaplaması gerekir [138];

- Kamera görüş alanına giren veya görüş alanından çıkan insanları otomatik olarak algılayabilmesi ve takip etmesi gerekir.
- Takip edilen kişinin kameraya yaklaşması veya uzaklaşmasıyla değişen görünümü algılayabilmesi gerekir.
- Gerçek zamanlı uygulamaların gereksinimlerine cevap verebilmesi gerekir.
- Ortamın ışık değişimine ve gürültüye duyarlı olması gerekir.
- Tek kamera kullanılmalıdır.
- Kısmi ve tam kapatma durumlarında anlamlı cevaplar verebilmesi gerekir.

Kalabalık ortamlarda çoklu insan takibi için önerilen ve gerçekleştirilen yöntem, yukarıda belirtilen gereksinimleri eksiksiz karşılayabilmektedir. Ancak son maddede belirtilen kapatma olaylarından sadece kısmi kapatma olayına anlamlı cevap verebilmektedir. Tam kapatma durumunda, takip edilen kişiye ait görüntü bilgisi tamamen kaybolduğu için takip işlemi sonlandırılmaktadır.

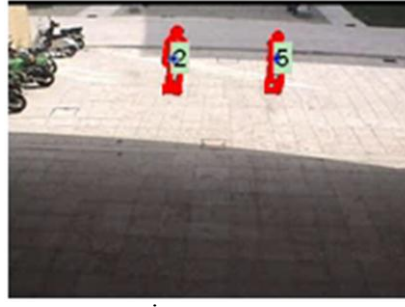
Bu bölümde, tez kapsamında gerçekleştirilen kalabalık ortamlarda insan takibi yöntemine ilişkin elde edilen sonuçlar verilmiştir. Gerçek video kayıtlarının işlenebilmesi, bağımsız hareket eden her bireyin tespiti, etiketlenmesi ve takibinin yapılabilmesi için MATLAB (Lisans No:585775) yazılım ortamında yukarıdaki bölümlerde açıklanan süreç uygun bir program geliştirilmiştir. Geliştirilen program yardımıyla VISOR [112] çevirim içi video izleme deposunda verilen ve araştırmacıların erişimlerine açık olan gerçek video kayıtları analiz edilmiştir.

Gerçekleştirilen ilk uygulamada, “Camera3exp3_1.mpeg” olarak isimlendirilen 384x288 boyutlarına sahip 4387 imge içeren video kaydı kullanılmıştır. Sabit bir kamera ile elde edilen bu video kaydının örnekleme frekansı 29 fps dir. Video kaydının toplam süresi 2.52 dakikadır. Uygulama için bu video kaydının seçilme nedeni, kısmi kapatma olaylarını içeriyor olmasıdır. Bu sayede geliştirilen çoklu insan takip yönteminin kısmi kapatma durumlarında nasıl bir davranış sergileyeceği daha açık görülebilmektedir. Ayrıca, bu video kaydında ani ışık değişimlerin olduğu da görülmektedir. İstenmeyen tüm bu harici etkenler, geliştirilen takip yönteminin performans değerlendirilmesinde önemli rol oynamaktadır. Şekil 3-6’da, çoklu insan takip yönteminin bu video kaydının ilk 1 dakikası üzerindeki takip sonuçları gösterilmiştir. Elde edilen sonuçlar göstermektedir ki; geliştirilen çoklu insan takip yöntemi, kamera görüş açısı içerisine girip çıkan kişileri otomatik olarak algılayabilmekte ve ilk defa gördüğü bir kişiye yeni ve eşsiz bir etiket numarası vererek imgeler boyunca bu kişiyi başarılı bir şekilde takip edebilmektedir.

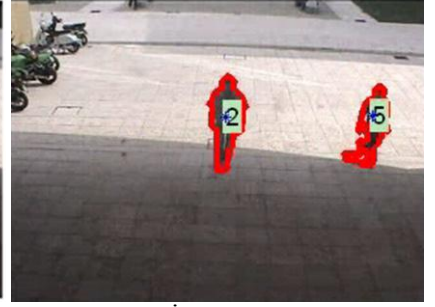
Şekil 3-6’da verilen 1585 indeksli imgede kısmi kapatma olayı meydana gelmektedir. Bu imgede, 19 ve 15 numaralarıyla etiketlenmiş kişilerin görünüşleri üzerinde kısmi kapatma olayı meydana gelmesine rağmen geliştirilen çoklu takip yöntemi, her iki kişiyi hatasız bir şekilde takip edilebilmektedir. Ayrıca bu video kaydında, kamera görüş açısı içerisine toplam 22 defa insan giriş çıkışı yapılmıştır. Gerçekleştirilen çoklu insan takip yöntemi, imgeye yapılan tüm giriş çıkışları eksiksiz bir şekilde tespit ederek kişileri etiketlemiş ve etiketlenen her bir bağımsız kişinin imge içerisindeki gezintisini başarılı bir şekilde takip edebilmiştir.



İmge 1002



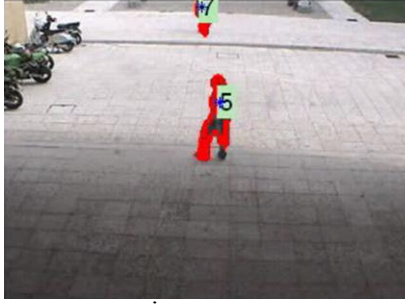
İmge 1019



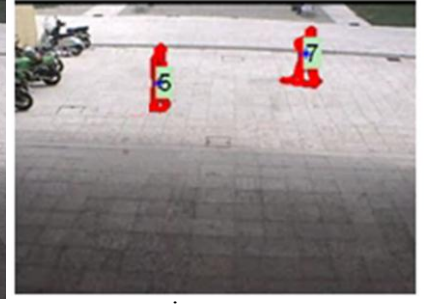
İmge 1051



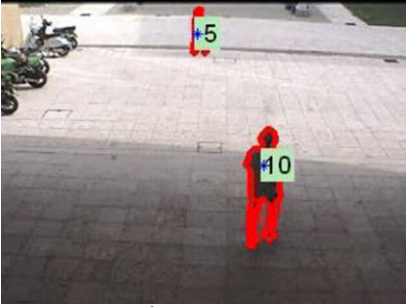
İmge 1089



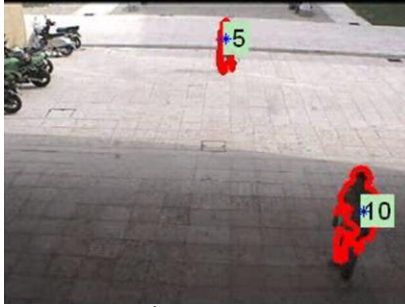
İmge 1120



İmge 1244



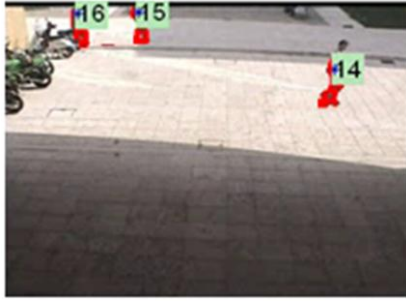
İmge 1282



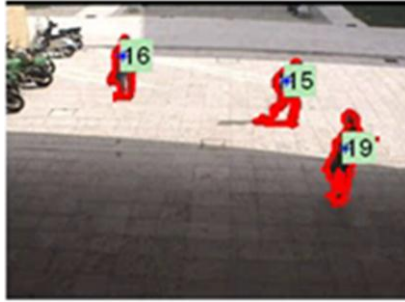
İmge 1299



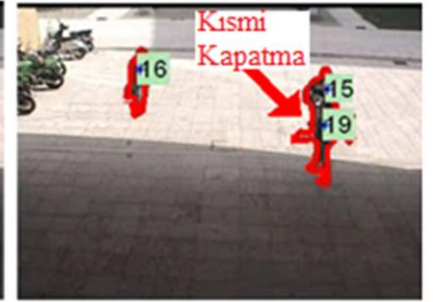
İmge 1380



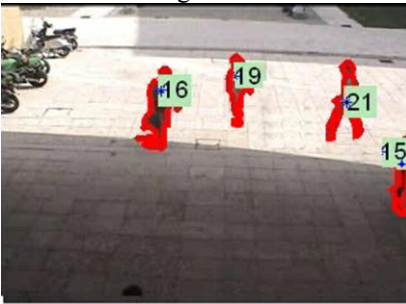
İmge 1510



İmge 1570



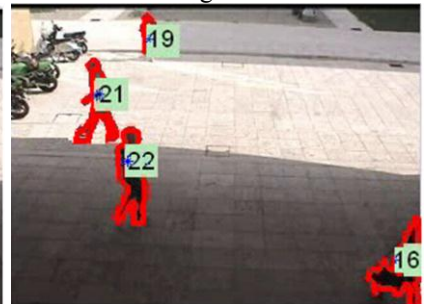
İmge 1585



İmge 1605



İmge 1645



İmge 1660

Şekil 3-6. Kalabalık ortamda çoklu insan takip yönteminin örnek video kaydı üzerindeki sonuçları

Önerilen çoklu insan takip yönteminin ikinci uygulaması, “Visor_video2_5_mpeg” olarak isimlendirilen ve 320x240 boyutlarına sahip 1304 imge içeren video kaydı üzerinde gerçekleştirmiştir. 10 fps örnekleme oranına sahip bu video dosyası da, sabit bir kamera kullanılarak elde edilmiştir. Test için seçilen ikinci video kaydına ait görüntü boyutu ve örnekleme oranı değerlerinin birinci video kaydından daha düşük olması, bu video kaydının daha kalitesiz imgeler içerdiği anlamına gelmektedir. Bu durum, insan takip yönteminin başarısını doğrudan olumsuz etkilemektedir. Önerilen çoklu insan takip yönteminin bu video kaydı üzerindeki sonuçları Şekil 3-7’de gösterilmiştir. Tam kapatma olayı Şekil 3-7’deki 462 indeks numarasına sahip imgede meydana gelmektedir. Bu imgede, iki insan görünümü üst üste çakışmıştır. Önerilen takip algoritması, bu durumda iki insanı tek bir insan olarak algılamış ve istenilmeyen bir sonuç üretmiştir. Bahsedildiği gibi önerilen yöntem sadece kısmi kapatma problemine çözüm sunabilir. Daha açık bir ifadeyle, kişilerin birbirlerinin görünümünü tamamen değil kısmen engellemesi durumunda istenilen çözüm sağlanabilir. Aynı videonun 488 indeks numaralı imgesinde ise dinamik arka plan çıkarım algoritmasının hareket halindeki insanı algılayamamasından dolayı “yanlış yakalama” olayı gerçekleşmiştir. Sonuç olarak, incelenen bu video kaydının görüş alanı içerisine toplam 24 insan giriş çıkışı gerçekleşmiş ve 21 kişinin takibi başarılı bir şekilde yapılabilmesine karşılık 3 kişinin takibi başarısızlıkla sonuçlanmıştır.

Geliştirilen çoklu insan takip yönteminin uygulamaları, 1GB RAM ve 3.2GHz işlemci hızına sahip bir masa üstü bilgisayar kullanılarak gerçekleştirilmiştir. Ayrıca yazılım geliştirme aracı olarak teknik hesaplamalar ve matematiksel problemlerin çözümü ve analizi için tasarlanmış olan MATLAB kullanılmıştır. Önerilen çoklu insan takip yönteminin çalışma hızı imgeden imgeye değişmekle birlikte ortamala değerler Tablo 3-1’de verilmiştir. Her iki uygulamada imgeler ardışıl olarak değil 10’ar atlamalı bir şekilde işlenmiştir. Elde edilen takip sonuçları önerilen yöntemin gerçek zamanlı uygulamalarda kullanılabilirliği ümidini arttırmıştır.

Geliştirilen çoklu insan takip yönteminin avantaj ve dezavantajları aşağıdaki gibi kısaca özetlenebilir;

Tablo 3-1. Çoklu insan takip yönteminin farklı video kayıtları üzerindeki işlem süresi sonuçları.

Video Adı	İmge Sayısı	Çözünürlük	Toplam işlem süresi (dk)
Camera3exp3_1.mpeg	4387	384x288	3.65
Visor_video2_5_mpeg	1304	320x240	0.72

Avantajları:

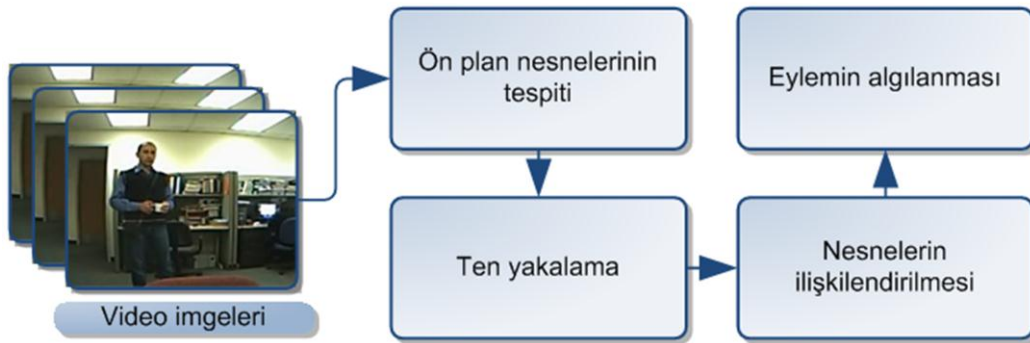
- 1) Gerçekleştirilen çoklu insan takip yöntemi, kısmi kapatma olayının gerçekleştiği durumlarda takip işlemini başarılı bir şekilde sürdürebilmektedir.
- 2) Kişilerin imge içerisinde ilk görüldüğü anda yeni indeks atama işlemi, imge içerisinde gezinti yaptığı süre içerisinde atanmış indeksi takip etme işlemi ve imgeden çıktığında ise indeks silme işlemi otomatik olarak gerçekleşmektedir.
- 3) İmge içerisine insan giriş çıkışına izin verilerek takip edilecek kişi sayısı sınırlandırılmamıştır.
- 4) Ortamın değişen ışık miktarına ve dinamik arka plan bilgisine hızlı bir şekilde adapte olabilmektedir.

Dezavantajları:

- 1) Tam kapatma olayı gerçekleştiğinde takip işlemi başarısızlıkla sonuçlanmaktadır.
- 2) Ortamdaki ışık miktarının hızlı bir şekilde değişime uğraması anında yanlış kişi algılamaları meydana gelmektedir.
- 3) Nadiren olsa kişilerin imgeye giriş ve çıkış anlarında kişi boyutlarının küçük olması nedeni ile yanlış algılamalar gerçekleşebilmektedir.

4. TEN YAKALAMA TABANLI İSTATİSTİKSEL EYLEM ALGILAMA YÖNTEMİNİN GELİŞTİRİLMESİ

Bu bölümde, belirli vücut uzuvlarının eşzamanlı takibi ile önceden tanımlanmış bir eylemi tespit edebilen ten yakalama tabanlı olarak geliştirilmiş yeni bir istatistiksel eylem algılama yöntemi sunulmaktadır. Geliştirilen yöntemin amacı, hareketli bir kişinin el ve yüz uzuvlarını takip ederek kişinin önceden tanımlanan bir eylemi gerçekleştirip gerçekleştirmediğini tespit etmektir. Geliştirilen eylem algılama yönteminin akış diyagramı Şekil 4-1’de gösterilmektedir. Buna göre, geliştirilen eylem algılama yöntemi ardışıl dört işlemin tekrarlı kullanımından oluşmaktadır: 1) Ön plan nesnelerinin tespiti; 2) Ten yakalama; 3) Nesnelerin ilişkilendirilmesi (veri bağı problemi); 4) Eylem algılama. Bölümün bundan sonraki kısmında önerilen eylem algılama yönteminin herbir aşaması detaylı bir şekilde ifade edilmiştir.



Şekil 4-1. Ten yakalama tabanlı istatistiksel eylem algılama yönteminin akış diyagramı.

4.1. Ön Plan Nesnelerinin Tespiti

Geliştirilen eylem algılama yöntemi, imge içerisindeki hareketli insan silüet bilgilerini elde edebilmek için Bölüm 3.1’de bahsedilen ön plan nesnelerinin tespiti yöntemini kullanmaktadır.

4.2. Ten Yakalama Yönteminin Kullanılması

Bu bölümde, ten rengi bilgisini kullanılarak yakalanan insan silüeti üzerindeki yüz ve el uzuvlarının diğer vücut uzuvlarından nasıl ayırt edilebildiği anlatılmaktadır. Ten

renginin algılanabilmesi için farklı renk uzaylarını kullanabilen ten yakalama yöntemleri mevcuttur [33, 113, 114]. Örneğin, Bayesian sınıflandırıcı, Gaussian karışımı modeli ve eliptik sınır modeli, ten yakalama uygulamalarında sıkça kullanılmaktadır. Bu yöntemlerin hesaplama maliyetlerinin yüksek olması gerçek zamanlı sistemlerde tercih edilmemelerine neden olmuştur.

Renk uzayı seçimi, ten yakalama işleminin en kritik adımlarından biridir. Literatürde RGB, HSV, HSI ve HSL gibi farklı renk uzayları mevcuttur. Fakat hangi renk uzayının ten yakalama işleminde daha etkili sonuçlar sağladığı konusunda henüz net bir bulgu elde edilememiştir [34].

Geliştirilen eylem analiz yönteminde, [115] nolu çalışmada gerçekleştirilen ten yakalama yöntemi kullanılmıştır. Bu yöntemin tercih edilmesinin nedeni, imge içerisindeki bir pikselin ten rengi olup olmadığını tespit edebilmek için basit, hızlı ve sağlam bir teknik kullanıyor olmasıdır. [115] nolu çalışmaya göre; (R, G, B) değerlerine sahip bir pikselin ten rengine sahip olup olmadığını anlaşılabilmesi için aşağıdaki sabit eşikleme tekniği kullanılmaktadır;

$$\begin{aligned} &(R > 95) \& (G > 40) \& (B > 40) \& \\ &(Max(R, G, B) - Min(R, G, B) > 15) \& \\ &(|R - G| > 15) \& (R > G) \& (R > B) \end{aligned} \tag{4.1}$$

Bu eşikleme sonucu yakalanan insan silüetinin ten rengine sahip olmayan kısımları yok edilerek el ve yüz uzuvlarının ortaya çıkması sağlanmaktadır.

4.3. Nesnelerin İlişkilendirilmesi

Ten rengine sahip vücut bölümlerinin tespitinden sonra bu bölümlerin el veya yüz nesnesi olma olasılıkları hesaplanmaktadır. Bunun işlem için el ve yüz nesnelerinin geçmiş imgelerde görülme sıklıkları kullanılmaktadır. Geçmiş ve güncel imge arasındaki veri bağı ilişkisinin kurulabilmesi için en yakın komşu yöntemi kullanılmaktadır. Böylelikle el veya yüz nesnelerinin bir önceki imgede var olup olmadığı, eğer varsa hangi görülme olasılığına sahip olduğu öğrenilmektedir. Ten rengine sahip vücut uzuvlarının görülme olasılığı β güncelleme katsayısı kullanılarak yenilenmektedir. Güncelleme işlemi aşağıdaki gibi gerçekleştirilir;

$$FP = \beta * FP + (1 - \beta) \quad (4.2)$$

$$HP = \beta * HP + (1 - \beta) \quad (4.3)$$

Denklem (4.2) ve (4.3)'de kullanılan FP ve HP değişkenleri, yakalanan vücut uzvunun el ve yüz olma olasılığını ifade etmektedir. Bu değerler belirli bir eşik değeri üzerine çıktığı zaman yakalanan nesne el veya yüz olarak etiketlenir. El ve yüz nesneleri arasındaki ayırım, nesnelerin geometrik yapısı kullanılarak gerçekleştirilmektedir. Bu nesnelerin iki seviyeli görüntüleri incelendiği zaman görülür ki, el nesnesi, üzerinde bir delik olmaksızın katı bir yapıya sahipken, yüz nesnesi üzerinde sağ ve sol gözlerden kaynaklanan iki delik bulunmaktadır. Bu ayırt edici fark kullanılarak el ve yüz nesneleri birbirinden ayırt edilmektedir. Ancak yine de bahsedilen bu ayırt edici farkın video kaydındaki tüm imgelerde ortaya çıkarılabilmesi mümkün değildir. Bazı imgelerde el nesnesi delikli olabileceği gibi yüz nesnesi de, gözlerin kapalı olması durumunda delikli olmayan bir yapıya sahip olmaktadır. Bu tür durumlarda el veya yüz nesnesinin kaybedilmemesi için istatistiksel bir yöntem kullanılmaktadır. Daha açık bir ifadeyle, el veya yüz nesnenin güncel imgede aniden kaybolması durumunda geliştirilen eylem algılama yöntemi, nesnenin silinmesi yerine olasılık değerini azaltmaktadır. Böylece, bazı imgelerde el veya yüz nesnesinin yakalanamamasının sisteme olan negatif etkisi ortadan kaldırılmaktadır.

İmgeler arasındaki nesne ilişkilerinin kurulması işleminde karşılaşılan diğer problem, el veya yüz nesnelerinin belirli bir süre hareketsiz kalması halinde arka plan nesnesi olarak algılanmasıdır. Bu problemin çözümü için tespit edilen el ve yüz nesnelerinin arkaplan güncellenmesine tabi tutulmaması gerekmektedir. Bunun için ön plan nesnelerinin tespiti algoritması ile ten yakalama algoritmasının sonuçları aşağıdaki gibi kullanılmaktadır;

$$\text{Kesişim}(t) = \text{Arka_plan}(t - 1) \cap \text{Ten_imgesi}(t) \quad (4.4)$$

$$\text{Arka_plan}(t) = \alpha * \text{Arka_plan}(t - 1) + (1 - \alpha) * \text{Ten_imgesi}(t) \quad (4.5)$$

$$\text{Arka_plan}(t) = (\text{Arka_plan}(t) * (1 - \text{Ten_imgesi}(t))) + \text{Kesişim}(t) \quad (4.6)$$

Denklemlerde kullanılan α değişkeni, arka plan yenilenme katsayısını göstermektedir. Arka plan yenilenme işlemi, imgenin tamamı üzerinde gerçekleşmektedir. Sadece el ve yüz nesnelerinin olduğu imge bölgeleri arka plan yenilenmesi işlemine tabi tutulmamaktadır.

Böylece, el veya yüz nesnelerinin belirli bir süre hareketsiz kalması durumunda arka plan nesnesi olarak algılanmayacağı garanti edilir.

4.4. Eylemin Algılanması

Bu bölümde, önceden belirlenen bir eylemin video kaydı içerisinde gerçekleştirilip gerçekleştirilmediğine karar verilmektedir. Bu işlemi gerçekleştirebilmek için tespit edilmek istenen eyleme ait özelliklerin sisteme önceden yüklü olması gerekmektedir. Gerçekleştirilen eylem algılama yöntemi, sadece el ve yüz nesnelerinin hareketlerini kullanmaktadır. El ve yüz nesnesinin tespitinden sonra bu nesnelerin imgeler boyunca yapmış oldukları konum değişimleri izlenerek yapılan eylem hakkında karar verilmektedir. El ve yüz nesneleri arasındaki ilişkileri ifade etmek için nesneler arasındaki uzaklık değerleri göz önüne alınmaktadır.

Gerçekleştirilen eylem algılama yöntemi uygulamasında, Alzheimer hastalarının ilaç alma eylemini icra edip etmediği incelenmektedir. Daha açık bir ifadeyle, takibi gerçekleştirilen kişinin, oda içerisindeki konumu belli olan bir ilaç kutusundaki ilacı alıp ağzına götürmesi eylemi tespit edilmek istenmiştir. Bu nedenle, ilaç kutusu açıldıktan sonra el ve yüz nesnelerinin yakalanması, takip edilmesi ve bu nesnelerin birleşip birleşmediğinin tespiti yapılmaktadır. Takip edilen nesnelerin konum değişimleri izlenerek bahsedilen ilacı alma eyleminin yapılıp yapılmadığı hakkında karar verilmektedir. Bunun için, yüz nesnesinin genellikle sabit kalması ve el nesnesi ile yüz nesnesi arasındaki mesafenin azalarak sıfır seviyesine ulaşması gerekmektedir.

4.5. Uygulama Sonuçları ve Değerlendirilmesi

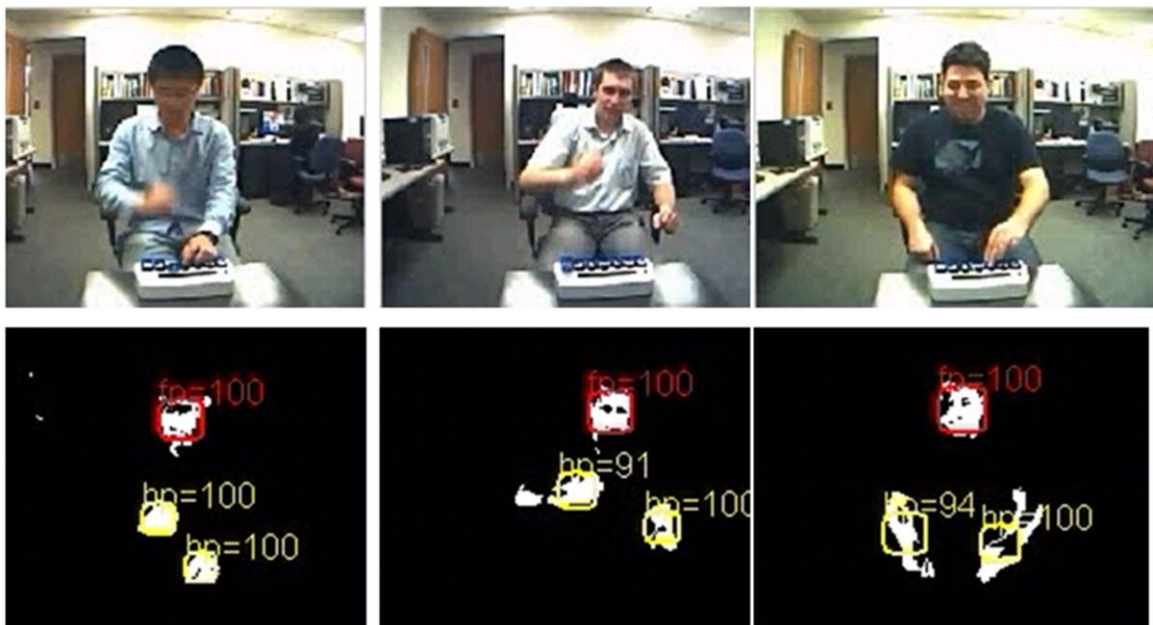
Ten yakalamaya dayalı istatistiksel eylem algılama yöntemi, Alzheimer hastalarının tedavilerine destek verebilmek amacıyla hasta hareketlerinin algılanmasına yönelik bir uygulamada kullanılmıştır. Alzheimer hastalarının tedavilerinde, hasta hareketlerinin algılanması ile tedavi etkinliğinin artırılması hedeflenmiştir.

Genellikle 65 yaş üzeri kişilerde görülen Alzheimer hastalığı, unutkanlık hastalığı olarak bilinmektedir. Hastanın tam olarak iyileşmesini sağlayıcı bir tedavi olmamakla birlikte hastalığın ilerlemesini yavaşlatabilen ve belirtilerini azaltabilen ilaçlar geliştirilmiştir [116]. Gerçekleştirilen uygulamada, tedavi süresince belirli aralıklarda ilaç

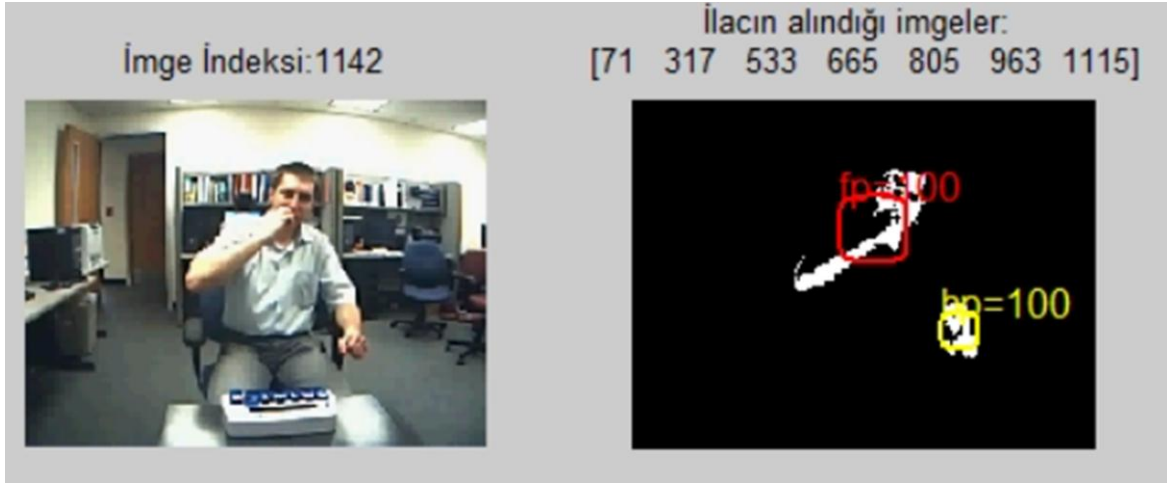
kullanması gereken Alzheimer hastalarının bu eylemi gerçekleştirip gerçekleştirmediğinin tespiti yapılmaktadır. Hastanın düzenli ilaç kullanımına yardımcı olmayı amaçlayan sistem, ilaç alma eylemi hakkında hastayı bilgilendirmektedir.

Alzheimer hastasının ilaç kutusundan ilacı alıp ağzına götürdüğünü tespit edebilmek için odanın belirli bir noktasına özel tasarlanmış ilaç kutusu yerleştirilmiştir. Sekiz adet küçük kutucuğa sahip bu ilaç kutusu, üzerindeki kutu kapaklarından hangisinin açıldığına dair bilgiyi bir sinyal üreterek bildirebilme özelliğine sahiptir. Böylece, hasta tarafından hangi ilacın kullanıldığı anlaşılmaktadır. Ayrıca, ilaç kutusunun bu özelliği, eylem algılama yönteminin tüm video kaydı üzerinde çalışması yerine sadece herhangi bir ilaç kapağı açıldıktan sonra çalışmasını sağlamıştır.

Gerçekleştirilen uygulamada, Alzheimer hastası olarak düşünülen 5 farklı kişinin özel tasarlanmış ilaç kutusundaki ilaçları alma eylemleri sabit bir kamera vasıtasıyla kaydedilmiştir. Her bir video kaydında ortalama 7 ilaç alma eylemi gerçekleştirilmiştir. Önerilen eylem algılama yöntemi kullanılarak geliştirilen uygulamada, ilaç kutusu açıldıktan sonra el ve yüz nesnelerinin yakalanması, etiketlenmesi, imgeler boyunca takip edilmesi ve takip sonucu ilaç alma eylemi hakkında karar verebilmesi sağlanmıştır. Uygulama çalışmasında, arka plan güncelleme değeri $\alpha = (1 - 1/150)$ olarak belirlenirken, el ve yüz nesnesinin olasılık değeri $\beta = (1 - 1/50)$ olarak belirlenmiştir. Bu değerler kullanılarak yakalanan ve olasılık değerleri güncellenen el ve yüz nesnelerine ait sonuçlar Şekil 4-2’de gösterilmektedir. El ve yüz nesneleri üzerindeki değerler, güncel



Şekil 4-2. Ten yakalama tabanlı istatistiksel eylem algılama yönteminin farklı kişiler üzerindeki uygulama sonuçları.



Şekil 4-3. Geliştirilen insan eylem algılama yöntemi kullanılarak ilacın ağıza götürülme eyleminin tespit edilmesi.

imgedeki el veya yüz olma olasılığını ifade etmektedir. Yüz nesnesi kırmızı renkli daire ile gösterilirken, sağ ve sol el nesneleri sarı renkli daireler içerisine alınmıştır. Önerilen yöntemde, el ve yüz nesne sınıflarının bir araya geldiği imgeler, ilacın alındığı imgeler olarak değerlendirilir ve bu image indeksleri sırayla kaydedilir. Şekil 4-3’de, el ve yüz nesnelerinin birleşme anı gösterilmektedir. Ayrıca bu andan önce ilacın alındığı diğer image indeksleri de kaydedilerek imgenin üst kısmında listelenmiştir.

Ten yakalamaya dayalı istatistiksel eylem algılama yöntemi, 5 farklı kişiden elde edilen ve toplam 35 ilaç alma eylemini içeren video kayıtları üzerinde test edilmiştir. Test sonucundan elde edilen takip sonuçları Tablo 4-1’de verilmiştir. Test sonuçlarına göre, önerilen eylem algılama yöntemi, 5 farklı kişinin gerçekleştirdiği toplam 35 ilaç alma eylemini, %92 oranında bir başarımla algılamıştır. El ve yüz nesnelerinin birleşme ve ayrılma olaylarının gerçekleştiği toplam 3 imgede yanlış algılama problemi yaşanmıştır. Bu problemi, el nesnesinin yüz nesnesi üzerinde titrek bir davranış sergilemesinden kaynaklanmaktadır. Bu durumda, önerilen eylem algılama yöntemi gerçekleştirilen tek bir ilaç alma eylemini ardı ardına gerçekleştirilen iki eylem olarak algılamaktadır. Tablo 4-1’deki test sonuçları Denklem (4.7) ve (4.8) içerisinde kullanılarak, geliştirilen eylem algılama yönteminin *TPR* (True Positive Rate) ve *FPR* (False Positive Rate) değerlerinin

Tablo 4-1. Ten yakalama tabanlı istatistiksel eylem algılama yönteminin takip sonuçları.

		Gerçek durum		toplam
		pozitif	negatif	
Test sonucu	pozitif	35 (TP)	3 (FP)	38(P)
	negatif	0 (FN)	0 (TN)	0 (N)
	toplam	35 (P)	3 (N)	

1 olduğu görülmektedir. Elde edilen bu sonuçlar, geliştirilen eylem algılama yönteminin sağlamlılığını ortaya koymaktadır.

$$TPR = TP/P = TP/(TP + FN) \quad (4.7)$$

$$FPR = FP/N = FP/(FP + TN) \quad (4.8)$$

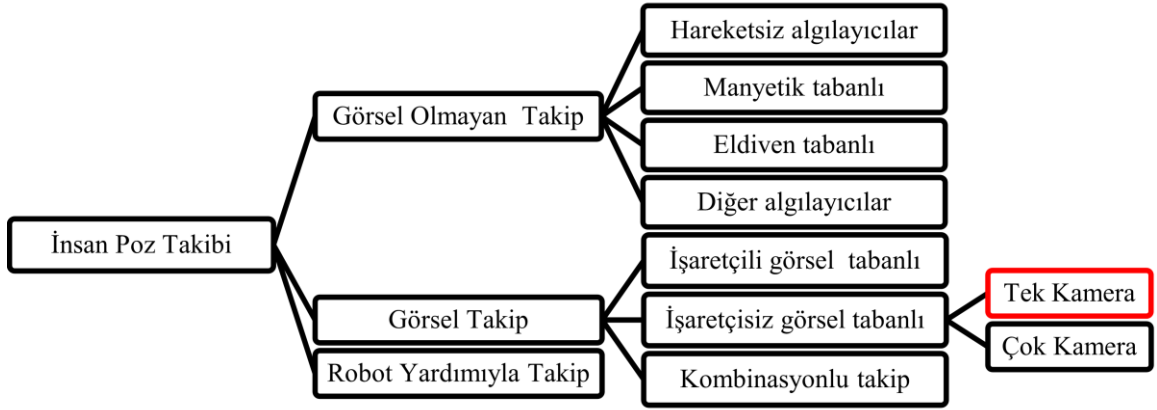
5. TEK KAMERA YARDIMIYLA 3-BOYUTLU İNSAN POZ TAKİP YÖNTEMİNİN GELİŞTİRİLMESİ

5.1. İnsan Poz Takibi Problemi

İnsan poz takibi, insan eylemlerinin tanınmasını sağladığı gibi insan-insan ve insan-çevre arasındaki etkileşim hakkın da bilgi vermektedir. Poz takibi konusu, insan hareketlerinin anlaşılmasına ihtiyaç duyan birçok uygulama için büyük bir öneme sahiptir. Bu konu hakkında önerilen çözüm yöntemleri temel olarak üç farklı başlık altında gruplanmaktadır [118]: 1) Görsel olmayan takip yöntemleri; 2) Görsel takip yöntemleri; 3) Robot yardımıyla takip. Yapılan bu sınıflandırma Şekil 5-1’de gösterilmiştir. Görsel olmayan takip teknikleri hareketsiz algılayıcılar [119, 120], manyetik algılayıcılar [121], eldiven kullanımlı takip [43] ve diğer algılayıcılar [122, 123] olarak sıralanabilir. Hareketsiz algılayıcılara örnek olarak hız ve yön tespiti yapabilen hızölçer ve jiroskop algılayıcıları verilebilir. Bu algılayıcılar pozisyon ve açı tespitini yapamazlar. Manyetik algılayıcılar, küçük boyutlara ve yüksek örnekleme oranına sahiptir. Bu tür algılayıcılarda, kapatma olarak bilinen görünüm engellenme problemine rastlanmaz [124]. Fakat algılayıcının yavaş çalışması ve titreşen bir sinyal üretmesi bu tür algılayıcıların sık tercih edilmemesini neden olur. Eldiven tabanlı algılayıcılar, parmak hareketlerini elektrik sinyaliye dönüştürerek el ile yapılan hareketlerin algılanmasını sağlar. [43] nolu çalışmada bu amaçla üretilen bir eldiven tabanlı hareket algılama sistemi bulunmaktadır. Bununla birlikte, robot yardımıyla elektromekanik veya elektromanyetik izleme stratejileri kullanarak insan hareketleri izlenebilmektedir. Ancak bahsedilen tüm bu görsel olmayan takip yöntemlerinin hem maliyeti yüksek hem de izlenecek kişinin üzerinde ekipmanlar taşıma zorunluluğundan dolayı uygulanması zordur.

Görsel takip yöntemleri, işaretçili, işaretçisiz ve bu iki yöntemin birleşimi olmak üzere üç alt başlıkta incelenmektedir. İşaretçi kullanan yöntemlerde genellikle insan vücudundaki eklem bölgeleri üzerine görsel vericiler yerleştirilir ve bu vericilerin konumları bir algılayıcı cihaz kullanılarak tespit edilir. [126] nolu çalışmada, işaretçi tabanlı insan hareketlerini algılayabilen örnek bir sistem gösterilmektedir.

İnsan vücut pozunun izlenebilmesi önünde üç temel zorluk bulunmaktadır: 1) Zemine ait görüntü bilgisinin ani ve sürekli değişim gösterebilmesi; 2) İnsan vücudunun



Şekil 5-1. İnsan hareket takibi yöntemlerinin sınıflandırılması [118].

birçok farklı poz sergileyebilmesi; 3) Vücudu oluşturan bölümlerin birbiri görünümünü engelleyebilmesi [118]. Bu konuda önerilecek yeni bir yöntemin bahsedilen bu zorlukları aşabilmesi ve insan poz takibini tatmin edici bir seviyede gerçekleştirebilmesi gerekmektedir.

Bu tez çalışmasında, uygulama maliyetinin düşük olması ve geniş bir kullanım alanına sahip olması nedeniyle işaretçi kullanmayan görsel takip yöntemleri ele alınmaktadır. İşaretçi kullanmayan görsel takip yöntemleri iki grup altında incelenmektedir: 1) Tek kamera kullanan teknikler [117, 150-152]; 2) Birden çok kamera kullanan teknikler [19, 153-154]. Birden çok kamera kullanımı problemin çözümünü basitleştirmiştir. Fakat birçok gerçek dünya uygulamalarında maliyetin düşük olması ve kullanım alanının geniş olması nedeniyle tek kamera kullanımı tercih edilmektedir.

5.1.1. Tek Kameralı Poz Takip Yöntemleri

Bu alanda önerilen yöntemler, herhangi bir algılayıcı kullanılmaksızın sadece tek bir kamera yardımıyla insan vücut pozunun takibini hedeflemektedir. Tek kameralı poz takip yöntemlerine örnek olarak parçacık süzgeci yöntemi [117, 127-129], genetik algoritma yöntemi [144], şablon karşılaştırma yöntemi [145], eksen tabanlı gösterim yöntemi [146], optik akış yöntemi [147] ve inanç yayılımı yöntemi [148] verilmektedir.

Poz takip probleminin çözümünde parçacık sayısı yüksek olan klasik parçacık süzgeci yönteminin kullanılması, hesaplama maliyetinin yüksek olmasına neden olur. Bu nedenle [133]'de, klasik parçacık süzgecinin bu dezavantajını ortadan kaldırmak için iki seviyeli hiyerarşik bir yöntem önerilmektedir. Yapılan bu çalışmada, kullanılan parçacık

sayısı azaltılarak vücut hareketlerinin algılanabilmesi sağlanmaktadır. Benzer bir çalışma [140]'da görülmektedir. Bu çalışmada, parçacık sayısı düşük bir süzgeç kullanan model tabanlı bir sistem ile insan hareketlerinin algılaması sağlanmıştır. Önerilen yöntem, iki seviyeli bir yapı içermektedir. İlk seviyede yapılandırma alanı büyük parçalar haline bölünür ve bazı tahminler gerçekleştirilir. İkinci aşamada, bu tahminlerin etrafında daha olası ve dar bir yapılandırma kullanılarak yeni parçacıklar üretilerek insan hareketleri bu parçacıkların takibi ile elde edilmektedir. [130]'da, parçacık süzgeci yöntemi içerisinde tekrarlamalı bir olabilirlik fonksiyonu kullanılarak vücut hareketleri tespit edilmiştir. Ayrıca bu çalışmada geniş açılı kamera kullanılmış ve takip edilecek vücut bölümü odanın ortasına ve tepesine kurulan tümsek bir kamera ile takip edilmiştir.

[117]'de, insan vücut hareketlerinin tespiti için üç aşamalı bir yöntem önerilmektedir: 1) Zemin görüntüsü elendikten sonra insan olduğu tespit edilen görüntü bölümleri elips içerisine alınır ve önyüz nesnesi olarak kabaca takip işlemi başlatılır; 2) Yüz, omuz ve ayaklar tespit edilir ve bir önceki maddenin sonucu ile örtüştürülür; 3) Veri Sürücülü Markov Zincirli Monte Carlo yöntemi kullanılarak 3-boyutlu vücut pozisyonu belirlenir. İnsan hareketlerinin yüksek bir performansla izlenebilmesine rağmen sistemin hesaplama maliyetinin yüksek olması, önerilen yöntemin gerçek zamanlı uygulamalar için elverişsiz olmasına neden olmuştur.

Poz takibi problemi bazı çalışmalarda optimizasyon problemi olarak ele alınmıştır. Örneğin, [128]'de, yapay bir insan modeli ile gözlenen insan verisi arasındaki benzerliğin tespiti için bir eşleştirme fonksiyonunun optimizasyonu gerçekleştirilir. Önerilen yöntemin dezavantajı, ilk video imgesinde insan vücut pozisyonlarının elle belirlenmek zorunda olmasıdır. Benzer bir çalışma [152]'de görülmektedir. Bu çalışmada, model ve gözlem sonuçları arasındaki karşılaştırmayı gerçekleştiren fonksiyonlar optimize edilmiştir.

[131] ve [146], insan hareketlerinin görsel çözümlemesi üzerine hazırlanan iki tez çalışmasıdır. [131]'de, klasik Bayesian izleyiciler detaylı bir şekilde ele alınırken, [146]'da, parça tabanlı koordinat sistemi kullanılarak elde edilen özellik yörüngelerine dayalı bir eylem tanıma yöntemi anlatılmaktadır. Parça tabanlı koordinat sistemi, eksen tabanlı bir gösterime dayalı olarak oluşturulmuştur. Hareket çözümlemesinde kullanılan özellikler vücut silüeti kullanılarak elde edilmiştir.

[143]'de hazırlanan tez çalışması ile Bayes yöntemleri kullanılarak çoklu nesnelerin takibi konusu detaylı bir şekilde incelenmiş ve literatürde bulunan istatistiksel nesne takip algoritmaları kullanılarak insan hareketleri algılanmıştır. [144]'de, ilk olarak yüksek

boyutlu imgeler Temel Bileşen Analizi (PCA) kullanılarak düşük bir boyuta indirgenir. Böylelikle hareket bilgisini algılamak için daha düşük boyutlu bir veri uzayı elde edilir. Sonraki adımda ise düşük boyutlu çözüm uzayı üzerinde genetik algoritma yöntemi kullanılarak hareket analizi gerçekleştirilmiş ve insan hareket karakteristiği elde edilmeye çalışılmıştır.

3-boyutlu insan poz takibi için istatistiksel bir yaklaşım, [148]'de önerilmiştir. Bu yaklaşım üç adımdan oluşmaktadır: 1) İnsan silueti üzerindeki yüz, omuz ve baş gibi vücut uzuvlarının bağımsız pozisyon tahmini; 2) Tahmini uzuv koordinatlarının inanç yayılım algoritması [158, 159] kullanılarak pozisyon bilgisinin netleştirilmesi; 3) Parçacık süzgeci yöntemi ile vücut uzuvlarının takibi. Bu çalışmanın dezavantajı, insan vücudunun sadece üst kısmını takip ediyor olmasıdır.

[150]'de, sadece birkaç sınırlı hareket incelenmiş ve bu hareketlerin algılanabilmesi için paralel algoritma tekniği ile yazılmış bir parçacık süzgeç algoritması tanıtılmıştır. Gerçek zamanlı uygulamalar için ümit verici sonuçlar elde edilmiştir. [151]'de, basit bir yapay insan modeli kullanılarak zıplama hareketinin algılanmasını sağlayan bir yöntem önerilmiştir. Modelin izdüşümünü kullanmak yerine eklem noktalarının ters kinematik dönüşümü alınarak poz tahmini gerçekleştirilmiştir.

Tek kameralı insan poz takip sistemleri bazen işaretçi tabanlı görsel algılayıcılar ile birlikte kullanılmıştır. Örneğin, [127], [129] ve [136] makalelerinde, insan hareketlerinin kayıtlı olduğu hazır bir veritabanı kullanılarak vücut poz tahmini yapılmıştır. Bu veritabanında, işaretçi tabanlı görsel algılayıcılar kullanılarak elde edilen belirli insan hareketlerine ait veriler saklanmaktadır. Bu çalışmalarda, veritabanında kayıtlı olan hareketler sistem tarafından eğitilir ve daha sonra eğitilmiş sistem, parçacık süzgecinin öncesel bilgisi olarak kullanılır. Yeni imgedeki hareketin tahmini, bu öncesel bilgi ile birlikte güncel durum tahmini sonuçlarının birleşimini sağlayan bir süzgeç ile yapılır.

[145]'de, vücuttaki belirli eklem bölgeleri üzerine konulan işaretçiler tek bir kamera ile izlenerek vücut pozunun takibi yapılmaktadır. İlk olarak eklem bölgelerindeki açılar, insan iskeletinin ters kinematiği kullanılarak tahmin edilmektedir. Daha sonra 3-boyutlu insan vücut modelinin 2-boyutlu perspektif izdüşümü hesaplanmaktadır. Son olarak elde edilen izdüşümü verisi ile görüntü verisi arasında bir şablon eşleştirme (histogram eşleştirme) tekniği kullanılarak vücut pozunu takip edilmektedir.

[156]'da, dikdörtgen şekiller ile modellenen insan silueti, histogram tabanlı bir takip yöntemi kullanılarak takip edilir. Önerilen yöntem, vücut kinematiğini göz ardı

ederek karmaşık dinamik denklemler ile ilgilenmemektedir. Bu sayede gerçek zamanlı uygulamalarda kullanılabileceği iddia edilmektedir. [160] makalesinde sporsal aktivitelerin anlaşılabilmesi için vücut poz takibi üzerine bir çalışma yapılmıştır. [161] çalışması, insan poz takibi konusu üzerinde hazırlanmış bir yüksek lisans tezidir. Bu çalışmada literatürde bulunan poz takip yöntemleri detaylı bir şekilde incelenmiştir. [162] ve [163] çalışmalarında içerik tabanlı poz takip yöntemleri önerilmektedir.

5.1.2. Çok Kameralı Poz Takip Yöntemleri

Poz takip sistemlerinde birden fazla kamera kullanımı maliyetin artmasına yol açmakta ve kameralar arasında entegre çalışabilme problemini ortaya çıkarmaktadır. Buna karşılık daha doğru bir poz takibi yapılabilmesine olanak sağlamaktadır. [19]'da birden fazla kameranın kullanıldığı örnek bir poz takip sistemi gerçekleştirilmiştir. Önerilen yöntemde, 3-boyutlu elipsler ile modellenen yapay bir vücut modeli ile video imgelerinden elde edilen 2 boyutlu insan silueti arasındaki maksimum benzerliğin tespiti yapılarak insan poz takibi gerçekleştirilmiştir. Maksimum benzerliğin hesaplanabilmesi için maksimum olabilirlik teorisini kullanılmıştır. Bu yöntemde, vücut hareketinin takibi için yakalanan insan silueti üzerinde geçici noktalar belirlenmiş ve bu noktalardan geçen dik vektörler kullanılmıştır. Önerilen poz takip yöntemi, sistemin başlangıç parametrelerine karşı duyarlılık göstermektedir.

Çok kameralı poz takip sistemlerinde klasik parçacık süzgeci algoritmalarının aksine optimizasyon yöntemleri de kullanılmıştır. Örneğin [102]'de önerilen poz takip yönteminde, ilk olarak sistem, önceden belirlenen poz hareketleri ile eğitilir. Daha sonra deterministik bir optimizasyon yöntemi ve olasılıksal bir örnekleme yönteminin birleşiminden oluşan melez bir yöntem ile poz takibi gerçekleştirilir. Sonuç olarak sağlamlık ve etkinlik kıstasları için doyurucu sonuçlar elde edilmiştir.

[108]'de insan bedeninin üst kısmında gerçekleşen hareketleri algılayan bir tahmin yöntemi önerilmektedir. Üç adımdan oluşan poz tahmin yöntemi, ilk olarak olası poz tahminlerini gerçekleştirmektedir. Daha sonra maksimum olabilirlik fonksiyonu ile en uygun insan pozunu seçilmektedir. Son olarak, seçilen en uygun poz tahmininin yapay bir model üzerinde benzetimi gerçekleştirilmektedir.

[105]'de önerilen çok kameralı insan poz takip sisteminde, 3-boyutlu vücut verisinin kullanılmış olması eğitim aşaması gereksinimini ortadan kaldırmıştır. Bu

çalışmada önerilen yöntem, gerçek zamanlı çalışma hızına sahip olup imge içerisindeki insan pozunun tahmini için 30 fps harcamaktadır.

5.2. Geliştirilen İnsan Poz Takip Yöntemi

Bu bölümde, insan poz takibi problemine yeni bir çözüm yöntemi sunulmaktadır. Önerilen çözüm yöntemi tek kamera vasıtasıyla gerçek dünya ortamından elde edilen video kayıtları üzerinde çalışmaktadır. Elde edilen 2-boyutlu imgeler dizisindeki her bir imgede vücut pozunun tahmini yapılır. Yapılan poz takibi, MATLAB ortamında tasarlanan 3-boyutlu yapay insan modeli üzerinde uygulanarak görsel model benzetimi gerçekleştirilmektedir.

Önerilen poz tahmin yönteminde dört aşamalı bir çözüm sunulmaktadır. İlk adımda, çevrim dışı olmak üzere video kaydının ilk imgesindeki insan silueti elde edilir. Daha sonra siluet üzerindeki eklem bölgeleri belirlenir ve son olarak eklem noktaları arasındaki Euler açı değerleri hesaplanmaktadır. İkinci aşamada, yapay insan modelini oluşturan vücut bölümlerine ait 3-boyutlu nokta koordinatları ileri kinematik denklemler kullanılarak hesaplanır. Üçüncü aşamada, poz tahmininin gerçekleşeceği bir sonraki imge işleme alındıktan sonra hareket halindeki insan silueti çıkarılır ve bu siluet üzerinde hesaplanan 3-boyutlu yapay insan modeli noktalarının 2-boyutlu perspektif izdüşümü hesaplanır. Son aşamada ise, izdüşümü alınan her bir vücut noktasının insan silueti ile hangi oranda örtüştüğünün değerlendirilmesi yapılmaktadır. Bu değerlendirilmenin yapılabilmesi için maksimum olabilirlik yöntemlerinden yararlanılmaktadır. Bu yöntemler yardımıyla, siluet ile izdüşüm bilgilerinin maksimum benzerliğini sağlayan eklem açıları bulunmaktadır. Güncel imgede insan pozunu belirleyen bu eklem açıları kullanılarak MATLAB ortamında 3-boyutlu vücut pozunun benzetimi gerçekleştirilir. Şekil 5-2’de, geliştirilen poz takip yönteminin akış diyagramı gösterilmektedir. Bölümün bundan sonraki kısmında geliştirilen poz takip yönteminin aşamaları detaylı bir şekilde ifade edilmektedir.

5.3. MOCAP Veritabanı

Geliştirilen poz takip yöntemi, Carnegie Mellon Üniversitesi tarafından 2006 yılında hazırlanmış olan MOCAP (Motion Capture) [164] adlı veritabanında kayıtlı bulunan belirli video kayıtlarını (yürüme, koşma, sıçrama videoları) ve bu kayıtlara ait işaretçi verilerini kullanmaktadır. MOCAP veritabanı, $3 \times 8m^2$ 'lik bir dikdörtgen alan içerisinde yapılan insan hareketlerine ait görüntü ve işaretçi verilerini içermektedir. Bu dikdörtgen alan 12 adet Vicon Infrared MX-40 kamera ile izlenmiştir ve her bir kamera 120 Hz hızına ve 4 MP çözünürlüğe sahiptir. Yürüme, koşma ve sıçrama gibi değişik hareketler yapan figüranların eklem koordinatlarını tespit edebilmek için özel bir elbise kullanılır. Şekil 5-3'de, toplam 41 adet işaretçiye sahip bu özel elbiseyi giymiş bir figüran resmi gösterilmektedir. Eklem bölgelerinin 3-boyutlu konum bilgilerini doğru bir şekilde elde edebilmek için işaretçiler vücut bölümleri üzerine özenle yerleştirilmiştir. Her bir video kaydı için elde edilen 3-boyutlu işaretçi konum bilgisi “c3d” uzantısına sahip bir dosyaya kaydedilir. Bununla birlikte, hareketi yapan figüranın vücut iskelet bilgisi ve vücut hareket bilgisi “asf” ve “amc” uzantılara sahip iki farklı veri dosyasına kaydedilmektedir.

5.4. Modeller

Geliştirilen poz takip yöntemi, belirli başlangıç model bilgisine ihtiyaç duymaktadır. Bu bilgiler, pozu takip edilecek kişinin vücut iskelet modeli, vücut



Şekil 5-3. Özel elbise giydirilen örnek bir figüran resmi

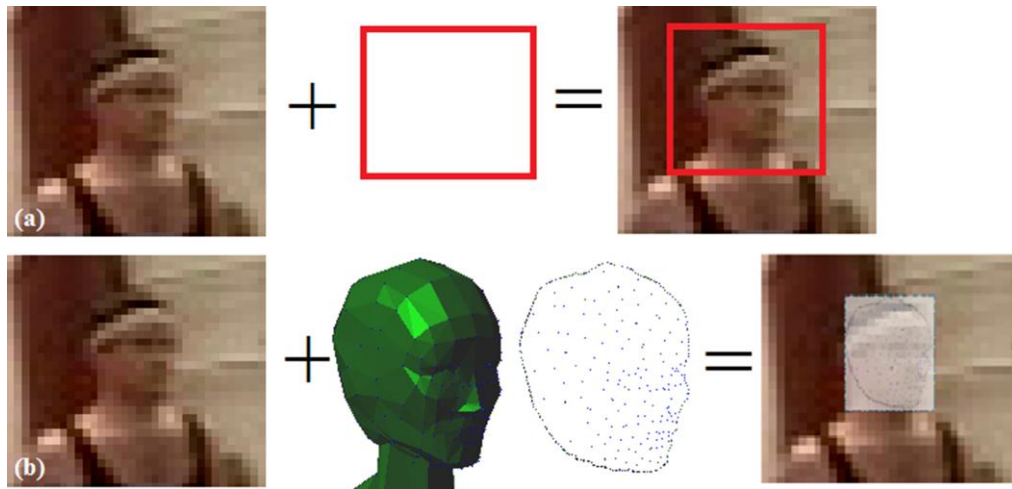
Tablo 5-1. Vücut iskelet modelindeki düğümlerin isimlendirilmesi.

J1	J2	J3	J4	J5	J6	J7	J8	J9	J10
Göbek	Sol kalça	Sol diz	Sol ayak bileği	Sol ayak	Sağ kalça	Sağ diz	Sağ ayak bileği	Sağ ayak	Köprücük
J11	J12	J13	J14	J15	J16	J17	J18	J19	J20
BoyunBaş		Sol omuz	Sol kol	Sol el bileği	Sol el	Sağ omuz	Sağ kol	Sağ el bileği	Sağ el

ağacın ilk seviyesindeki kök düğüm olan J1 düğümüdür ve hemen sonraki adımda ikinci seviyedeki J6, J10 ve J2 düğümlerinin koordinatları hesaplanmaktadır. Tablo 5-1’de, insan vücut iskelet modelindeki her bir düğüm isimlendirilmiştir.

5.4.2. Vücut Görünüm Modeli

Bu bölümde, vücut modelindeki 19 aygıtın (her bir vücut bölmesinin) renksel görünüm modeli elde edilmektedir. Genellikle vücut bölümlerinin görünüm modelleri, Şekil 5-5(a)’da görüldüğü üzere dikdörtgen veya elips gibi simetrik geometrik şekiller vasıtasıyla elde edilmektedir. Takip edilecek vücut bölümü Şekil 5-5(a)’da görüldüğü gibi geometrik şekil içerisine alınarak şekil içerisinde kalan piksellerin renk bilgisi elde edilmektedir. Böyle bir görünüm modelinin seçimi, sistemin takip hızının artmasına neden olmaktadır. Fakat simetrik şekillerin kullanımı, seçilen geometrik şeklin içerisine takip edilen nesneye ait olmayan arka plan bilgisinin girmesine neden olmaktadır. Böylece takip sağlamlığı ve güvenliği azalmaktadır. Bahsedilen dezavantajın üstesinden gelebilmek için, geliştirilen poz takip sistemi, Şekil 5-5(b)’de gösterildiği gibi farklı bir vücut görünüm modelini kullanmaktadır. Buna göre, 3-boyutlu yapay insan modelinin her bir vücut



Şekil 5-5. Görünüm modelleri (a) Klasik model (b) Önerilen model.

bölümüne ait noktaların izdüşümü hesaplanır ve imge üzerine yerleştirilir. Bu sayede sadece takip edilecek vücut bölümüne ait renk değerleri kodlanır, arka plan renk değerleri görünüm modeli içerisine giremez. Bu görünüm modelinin kullanımı, takip edilen nesnenin sunumunu iyileştirerek sistemin takip performansının artmasını sağlar.

5.4.3. Kamera Modeli

2-boyutlu imge görüntüleri sabit bir noktaya yerleştirilmiş tek bir video kamera kullanılarak elde edilir ve poz tahmini için imgelerdeki insan silueti çıkarılır. Geliştirilen poz tahmin sisteminin de, poz tahminini gerçekleştirebilmesi için 3-boyutlu yapay vücut modelinin aynı kamera koordinatlarından perspektif izdüşümünün [165] hesaplanmasına ihtiyacı vardır. Perspektif izdüşüm tekniği ile yapay vücut modeline farklı açı ve pozisyonlardan bakabilmek mümkündür. Şekil 5-6'da, 3-boyutlu yapay bir insan modelinin 2-boyutlu imge düzlemi üzerine izdüşümü gösterilmiştir. Buna göre, perspektif izdüşüm için aşağıda belirtilen değişkenlere ihtiyaç vardır;

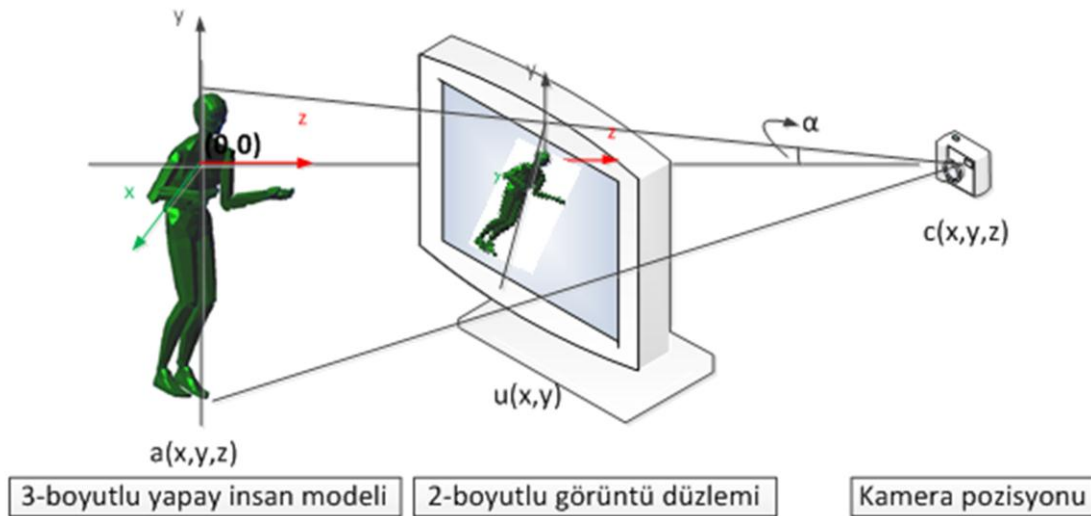
$a_{x,y,z}$: İzdüşümü alınacak 3-boyutlu nokta koordinatı.

$c_{x,y,z}$: Kameranın 3-boyutlu pozisyonu.

$\theta_{x,y,z}$: Kameranın 3-boyutlu açısı.

$e_{x,y,z}$: İzleyen kişinin, ekran yüzeyine göre olan pozisyonu.

$u_{x,y}$: $a_{x,y,z}$ noktasının 2-boyutlu izdüşümü.



Şekil 5-6. Perspektif izdüşüm tekniği kullanılarak 3-boyutlu yapay insan modelinin 2-boyutlu imge düzlemine izdüşümünün alınması.

İlk olarak, $c_{x,y,z}$ tarafından tanımlanan koordinat sistemi içerisindeki $a_{x,y,z}$ noktası bir başka $d_{x,y,z}$ noktasına dönüştürülür;

$$[\theta_x, \theta_y, \theta_z] = [-\theta_x, -\theta_y, -\theta_z], \quad (5.1)$$

$$\begin{bmatrix} d_x \\ d_y \\ d_z \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos(\theta_x) & \sin(\theta_x) \\ 0 & -\sin(\theta_x) & \cos(\theta_x) \end{bmatrix} * \begin{bmatrix} \cos(\theta_y) & 0 & -\sin(\theta_y) \\ 0 & 1 & 0 \\ \sin(\theta_y) & 0 & \cos(\theta_y) \end{bmatrix} * \begin{bmatrix} \cos(\theta_z) & \sin(\theta_z) & 0 \\ -\sin(\theta_z) & \cos(\theta_z) & 0 \\ 0 & 0 & 1 \end{bmatrix} * \left(\begin{bmatrix} a_x \\ a_y \\ a_z \end{bmatrix} - \begin{bmatrix} c_x \\ c_y \\ c_z \end{bmatrix} \right) \quad (5.2)$$

Daha sonra Denklem (5.3)'teki homojen koordinatlar kullanan matris formu yardımıyla elde edilen $d_{x,y,z}$ noktasının 2-boyutlu imge düzlemi üzerine izdüşümü alınır.

$$\begin{bmatrix} f_x \\ f_y \\ f_z \\ f_w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & -e_x \\ 0 & 1 & 0 & -e_y \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/e_z & 0 \end{bmatrix} * \begin{bmatrix} d_x \\ d_y \\ d_z \\ 0 \end{bmatrix}. \quad (5.3)$$

Son olarak, $a_{x,y,z}$ noktasının koordinatları görüntü üzerindeki izdüşüm ile ilişkilendirilir;

$$\begin{aligned} u_x &= f_x/f_w \\ u_y &= f_y/f_w \end{aligned} \quad (5.4)$$

İzleyicinin ekran yüzeyine olan uzaklığı e_z ile gösterilir ve bu değer direkt olarak bakış açısıyla ilgilidir: $\theta_z = 2 \cdot \tan^{-1}(1/e_z)$.

5.5. Ön İşleme Adımı

Önerilen yöntem ile insan poz tahmininin yapılabilmesi için video kaydının ilk imgesi üzerinde sistemin başlangıç parametrelerinin belirlenmesine ihtiyaç duyulmaktadır. Bu aşama ön işlem adımı olarak adlandırılır ve bu aşamada 3-boyutlu yapay insan modeli ile sonraki video imgelerinden çıkarılacak insan silueti arasındaki ilişkinin kurulabilmesi için başlangıç parametreleri ayarlanır. Şekil 5-7'de, ön işlem aşaması akış diyagramı şeklinde gösterilir. Temel olarak, ön işlem adımı alt üç adımdan oluşur: 1) İmge

içerisindeki hareketli nesnelerin tespiti; 2) Morfolojik işlemler; 3) İnsan silueti ile yapay modelin örtüştürülmesi. Bu bölümün devamında alt bölümler hakkında detaylı bilgi verilmiştir.

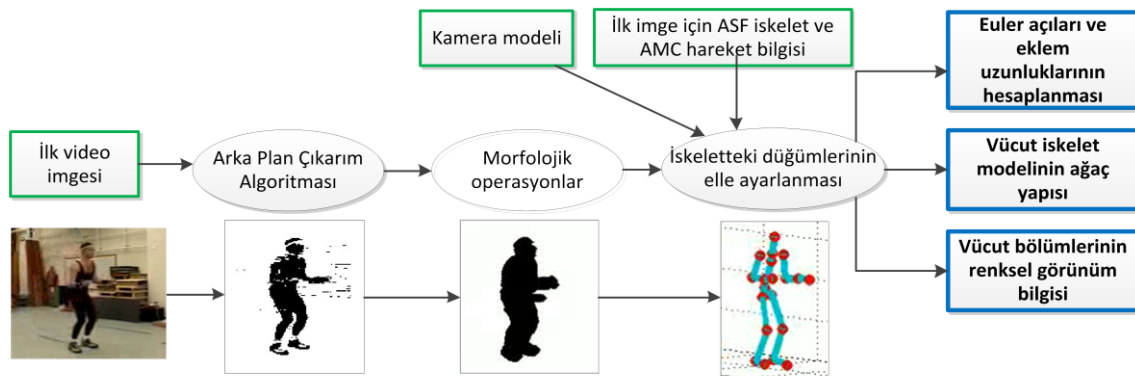
5.5.1. Ön Plan Nesnelerinin Tespiti

İnsan poz tahmini uygulamasında imgelerdeki ön plan nesnelerinin tespiti için kullanılan yöntem Bölüm 3.1’de bahsedilen yöntem ile tamamen aynıdır. Detaylı bilgi için Bölüm 3.1’e bakınız.

5.5.2. Ön Plan Nesneleri İçerisinden İnsan Siluetinin Çıkarımı

Dinamik arka plan çıkarım algoritması hareketli ön plan nesnelerini tespit etmekle birlikte, arka planda hareketsiz olan ve renk değeri ani değişimler gösteren birçok pikseli de yakalar. Bunun nedeni, gerçek dünya ortamından elde edilen video imgeleri üzerinde birçok gürültü bilgisinin mevcut oluşu, ortamdaki ışık miktarının ani değişebilmesi ve ön plan nesnelerinin hızlı hareket etmesi olabilir. Bu aşamada hareketli ön plan nesneleri içerisinden insan siluetinin çıkarılması gerekir. İnsan olmayan ve ön plan nesnesi olarak yakalanmış olan nesneleri temizlemek sırayla iki adımda gerçekleştirilir: 1) Morfolojik operasyonlar; 2) Süzgeçleme.

Elde edilen iki seviyeli imge üzerinde temel iki morfolojik işlem uygulanır: i) Açma [166]; ii) kapama [167]. Açma operasyonu ile aşınmış nesne pikselleri temizleneceği için nesnenin piksel sayısında azalma meydana gelir. Böylece nesne etrafında ve nesneye ait olmayan pikseller elenmiş olur. Kapatma operasyonunda ise nesne, belirlenen yapısal



Şekil 5-7. Ön işleme adımının akış diyagramı (sistem giriş ve çıkışları sırayla yeşil ve mavi renklerle gösterilmektedir).

element kullanılarak genişletilir. Böylece çok yakın bir alan içerisinde ayırık olan nesne bölümlerinin birleşmesiyle nesnenin tek parça haline gelmesi sağlanır.

Morfolojik operasyonlardan sonra insan silüetinin dışında hala istenmeyen yapay nesneler olabilir. Bunların elenmesi için her bir ön plan nesnesi hakkında aşağıdaki özellikler elde edilir ve bu özellikler değerlendirilerek insan silüeti dışındaki diğer tüm ön plan nesneleri imgeden çıkarılır. Her bir ön plan nesnesi için şu özellikler elde edilir;

Alan: Ön plan nesnesindeki piksel sayısı bu nesnenin alanı olarak değerlendirilir.

Katılık: Dışbükey örtü içerisindeki piksel sayısının nesne üzerindeki piksel sayısına oranı olarak hesaplanır.

İlginçlik: Nesne görünümünün çizgiye veya daireye daha çok benzediğini belirten bir değer.

Bir birinden bağımsız elde edilen bu özelliklerin aşağıdaki sınırlar içerisinde olup olmadığı kontrol edilir. Ekranda sadece tek bir hareketli insanın olduğu kabul edildiği için, aşağıdaki sınırlar içerisinde olan ve en büyük alan değerine sahip nesne insan silüeti olarak belirlenir.

$$1 > \text{Katılık} > 0.5,$$

$$1 > \text{İlginçlik} > 0.8 \quad (5.5)$$

5.5.3. Başlangıç Parametrelerinin Ayarlanması

Gerçekleştirilen insan poz takip sistemi, ilk adımda başlangıç parametrelerinin belirlenmesine ihtiyaç duyar. Gerçekleştirilen sistemin üç başlangıç parametresi vardır: 1) Eklem açıları ve uzunlukları; 2) Vücut iskelet modelinin ağaç yapısı; 3) Vücut bölümlerinin renksel görünüm bilgisi. Sistemin başlangıç parametrelerinin belirlenmesi için video kaydının ilk imgesi kullanılır. Sonraki imgelerde bu aşamaya gerek duyulmadan poz takibi işlemi gerçekleşir.

Başlangıç parametrelerinden ilk ikisinin elde edilebilmesi için, işlenen video kaydına ait asf uzantılı dosya okunur. Asf uzantılı örnek bir dosya Şekil 5-8’de gösterilmiştir. Metin formatına sahip bu dosya, figüran üzerindeki tüm işaretçi bilgilerini içerir. Gerçekleştirilen poz takip yazılımı bu metin dosyasını okuyarak işaretçilere ait koordinat bilgilerini, eklem uzunluklarını ve vücut iskelet modelinin ağaç yapısını elde eder. Şekil 5-9’da, asf dosyası okunduktan sonra elde edilen bilgiler kullanılarak çizilen

vücut iskelet görünümü gösterilmektedir. İşaretçi koordinatları belirlendikten sonra, bitişik iki işaretçi arasındaki Euler açılarının belirlenmesi gerekir. Euler açıları Şekil 5-10'da gösterilmiştir. Bu açıların hesaplanabilmesi için işaretçi koordinat bilgileri aşağıdaki gibi kullanılır;

$$x = x_1 - x_2 \quad y = y_1 - y_2 \quad z = z_1 - z_2 \quad (5.6)$$

$$r = \sqrt{x^2 + y^2 + z^2} \quad (5.7)$$

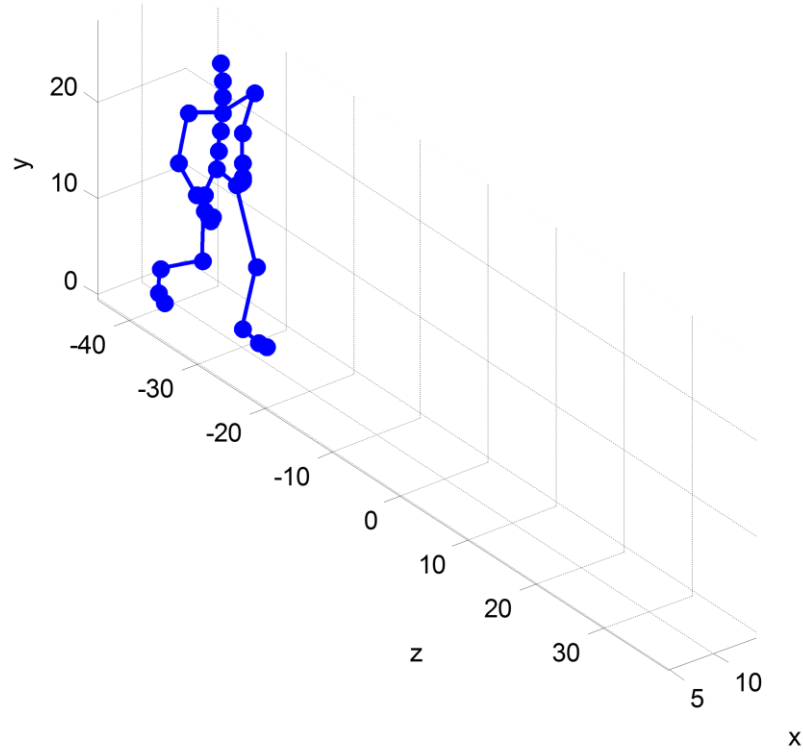
$$\emptyset = \arccos\left(\frac{z}{r}\right) \quad (5.8)$$

$$\theta = \text{atan2}(y, x) \quad (5.9)$$

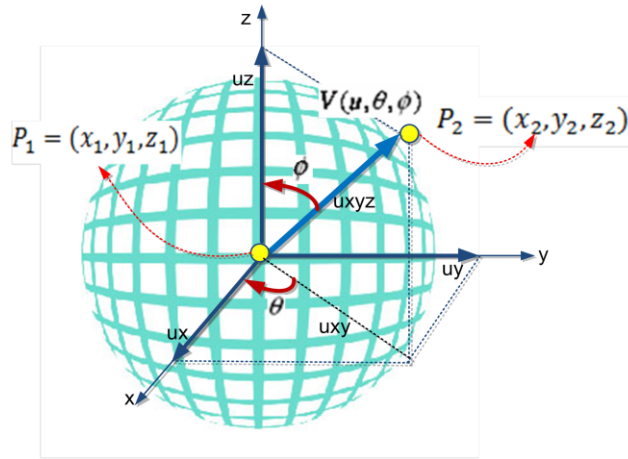
Son olarak, üçüncü başlangıç parametresi olan vücut bölümlerinin renksel görünüm bilgilerinin elde edilmesi gerekir. Bu bilginin elde edilebilmesi için üç aşamalı bir işlem gerçekleştirilir. Birinci aşamada, vücut iskelet modeli üzerine Oxford Üniversitesi Robotik araştırma gurubu tarafından hazırlanan 3-boyutlu yapay insan modeli [168] giydirilir. İkinci olarak, bu yapay modelin imge üzerine perspektif izdüşümü alınır. Son olarak, izdüşüm sonucu imge üzerine düşen noktalar, vücut bölümlerinin görünüm bilgilerini elde

```
# AST/ASF file generated using VICON BodyLanguage
# -----
:version 1.10
:name VICON
:units
  mass 1.0
  length 0.45
  angle deg
:documentation
  .ast/.asf automatically generated from VICON data using
  VICON BodyBuilder and BodyLanguage model FoxedUp or BRILLIANT.MOD
:root
  order TX TY TZ RX RY RZ
  axis XYZ
  position 0 0 0
  orientation 0 0 0
:bonedata
  begin
    id 1
    name lhipjoint
    direction 0.566809 -0.746272 0.349008
    length 2.40479
    axis 0 0 0 XYZ
  end
  begin
    id 2
    name lfemur
```

Şekil 5-8. “asf” uzantılı örnek bir dosya içeriğinin görünümü



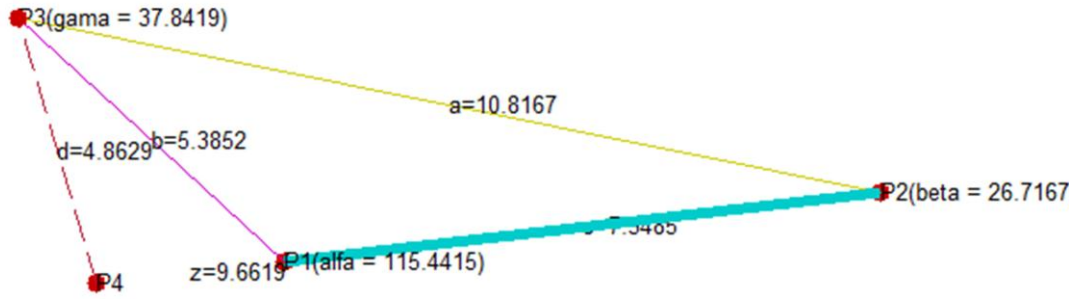
Şekil 5-9. İnsan vücut iskelet modelinin gösterimi.



Şekil 5-10. İki eklem noktası arasındaki euler açılarının gösterimi.

etmek için kullanılır. Böylece her uzvun renk değerleri kodlanmış olur.

Vücut iskelet modelindeki her hangi bir hareketin 3-boyutlu yapay vücut modeline taşınabilmesi için iskelet modelindeki eklem noktalarının yapay modeldeki vücut noktalarına bağlanması gerekir. Şekil 5-11’de iskelet model ile yapay model arasındaki bu bağlama işlemi gösterilmiştir. Buna göre, Şekil 5-11’deki P_1 ve P_2 eklem çubuğunun birlikte yaptığı hareketin P_3 noktasına taşınabilmesi için P_3 noktasının $P_1 - P_2$



Şekil 5-11. Yapay insan modeline ait noktaların vücut iskeleti üzerine giydirilmesi.

doğrusunu dik kesen $P4$ noktasına ait koordinatların hesaplanması gerekir. $P4$ noktasının koordinatları aşağıdaki denklemler kullanılarak hesaplanmaktadır;

$$\beta = \arccos\left(\frac{a^2 + c^2 - b^2}{2 \cdot a \cdot c}\right) \quad (5.10)$$

$$d = a \cdot \sin(\beta) \quad (5.11)$$

$$z = \sqrt{a^2 - d^2} \quad (5.12)$$

$$P4 = P2 - \frac{z}{c}(P2 - P1) \quad (5.13)$$

$P4$ noktası ve bu noktanın $P3$ 'e olan d uzaklığı hesaplandıktan sonra $P1 - P2$ doğrusunda meydana gelecek herhangi bir hareket sonrasında $P3$ noktasının yeni koordinatları aşağıdaki denklemler ile hesaplanabilir. Böylelikle iskelet modelde gerçekleşecek bir hareket aynı şekilde 3-boyutlu yapay insan üzerinde uygulanabilmektedir.

$$\vec{v}_1 = P4 - P3 \quad (5.14)$$

$$\vec{v}_2 = P4 - P2 \quad (5.15)$$

$$\vec{v}_3 = \vec{v}_1 \times \vec{v}_2 \quad (5.16)$$

$$\vec{v}_3 = \frac{v_3}{\sqrt{v_3^2}} \quad (5.17)$$

$$P5 = P4 + \vec{v}_3 \quad (5.18)$$

3-boyutlu yapay insan modelinin hareket edebilmesi sağlandıktan sonra 2-boyutlu imge üzerine izdüşümünün hesaplanması gerekir. Bunun için Bölüm 5.4.3'te bahsedilen perspektif izdüşüm yöntemi kullanılır. İzdüşüm sonrası elde edilen 2-boyutlu vücut noktalarının imge üzerindeki renk değerleri kullanılarak her bir vücut bölümünün renksel görünüm modeli elde edilir.

5.6. İstatistiksel Poz Takibi

Bu aşamada, güncel imgedeki 2-boyutlu insan silueti ile maksimum örtüşmeyi sağlayacak 3-boyutlu yapay insan modeline ait Euler açılarının istatistiksel bir model kullanılarak tahmini yapılmaktadır. Tahmin işlemi Bayesian çatısı altında gerçekleşir ve genel vücut pozunun tahmininin yapılabilmesi için vücut bölümlerinin bağımsız bir şekilde takibine ihtiyaç duyulur. Vücut bölümlerinin takip edilebilmesi için ön işleme adımından elde edilen eklem açıları, eklem uzunlukları, iskelet modelinin ağaç yapısı ve vücut bölümlerinin renksel görünüm bilgisine ihtiyaç duyulur.

Geliştirilen poz takip yöntemi, vücut bölümlerinin güncel imgedeki şekil ve görünüm bilgileri ile geçmiş imgelerdeki bilgilerini tek bir Bayesian çatısı altında değerlendirerek vücudun güncel poz tahminini yapabilmeyi sağlar.

5.6.1. Bayesian Takip Modeli

Poz tahmin problemindeki amacımız, t anına kadar elde edilen $I_1, \dots, I_t$ gözlem değerlerini (kısaca $I_{1..t}$ olarak gösterilebilir) kullanarak t anındaki s_t durum vektörünü hesaplamaktır. Bu amaca yönelik en uygun yanıtı Bayesian tahmin teoremi vermektedir. Buna göre güncel imgedeki poz tahmini problemi Bayes teoremi kullanılarak şu şekilde ifade edilebilir;

$$\begin{aligned} s_{t*} &= \arg \max_{(s_{t*} \in \Theta)} \{P(s_t | I_{1..t})\} \\ &= \arg \max_{(s_{t*} \in \Theta)} \left\{ \overbrace{P(I_t | s_t)}^{\text{olabilirlik}} \overbrace{P(s_t | I_{1..t-1})}^{\text{öncesel dağılım}} \right\} \end{aligned} \quad (5.19)$$

Θ sembolü ile çözüm uzayı ifade edilir ve 19 vücut bölümü için çözüm uzayı $\Theta = \bigcup_{n=1}^{19} \Theta_n$ olarak yazılabilir. Buna göre, t anındaki imgede insan poz tahmininin yapılabilmesi için sırayla ardışık üç işlemin gerçekleştirilmesi gerekmektedir: 1) s_t durum vektörü belirlenmesi; 2) Belirlenen durum vektörünün $P(s_t | I_{1..t-1})$ öncesel dağılımın hesaplanması; 3) Belirlenen durum vektörünün güncel imge üzerindeki $P(I_t | s_t)$ olabilirlik fonksiyonunun hesaplanması.

5.6.2. Durum Vektörünün Belirlenmesi

t anında I_t imgesindeki insana ait durum vektörü s_t olarak ifade edilir ve yapılan uygulamada insan vücudunun 19 bölmeye ayrıldığı düşünülürse s_t durum vektörü şu şekilde yazılabilir;

$$s_t = \{(k_1, X_1), \dots, (k_{19}, X_{19})\} \quad (5.20)$$

k_j , j . vücut bölümüne verilen eşsiz etiket değeridir. X_j ise j . vücut bölümünün durum vektörünü gösterir ve aşağıdaki gibi ifade edilir;

$$X_j = \{g_j, m_j, I_j, e_j\} \quad (5.21)$$

Her bir vücut bölümünün g_j , geometrik şekil özelliklerini; m_j , hareket özelliklerini; I_j , görünüm özelliklerini ve e_j , eklem açılarını ifade etmektedir. g_j geometrik şekil özellikleri $g_j = (o_j, x_j, y_j, h_j, f_j, i_j)$ şeklinde ifade edilir. Her bir vücut bölümünün, o_j , dönme açısını; x_j , x eksenindeki değerini; y_j , y eksenindeki değerini; h_j , uzunluğunu; f_j , genişliğini ve i_j , eğimini ifade etmektedir. m_j hareket özellikleri $m_j = (v_j, a_j)$ olarak ifade edilir. v_j , hızını; a_j , eğimini ifade etmektedir. I_j , vücut bölümüne ait piksellerin renk histogram bilgisini ifade etmektedir. Son olarak her bir eklem Euler açıları $e_j = (\theta_j, \phi_j)$ şeklinde ifade edilir. Buna göre t anındaki durum vektörü aşağıdaki gibi ifade edilmektedir;

$$s_t = \left\{ (k_1, o_1, x_1, y_1, h_1, f_1, i_1, v_1, a_1, I_1, \theta_1, \phi_1), \dots, (k_{19}, o_{19}, x_{19}, y_{19}, h_{19}, f_{19}, i_{19}, v_{19}, a_{19}, I_{19}, \theta_{19}, \phi_{19}) \right\} \quad (5.22)$$

5.6.3. Öncesel Dağılımın Hesaplanması

$P(s_t | I_{1,\dots,t-1})$ ile ifade edilen öncesel dağılım fonksiyonu iki alt fonksiyona ayrıştırılabilir;

$$P(s_t | I_{1,\dots,t-1}) \propto P(s_t) P(s_t | I_{1,\dots,t-1}) \quad (5.23)$$

Zamana bağlı olmayan $P(s_t)$ aşağıdaki gibi ifade edilir;

$$P(s_t) = \prod_{j=1}^{19} p(|D_j|)p(X_j) \quad (5.24)$$

$p(|D_j|)$ ile vücut bölümünün imge üzerine olan izdüşümü ifade edilir ve $|D_j|$, izdüşüm sonrası elde edilen görüntü alanıdır. Buna göre $p(|D_j|)$ şu şekilde modellenir;

$$p(|D_j|) = \overbrace{\exp(-\lambda_1 |D_j|)}^{\text{büyük nesne eleyici}} \overbrace{[1 - \exp(-\lambda_2 |D_j|)]}^{\text{küçük nesne eleyici}} \quad (5.25)$$

Denklemdaki birinci ve ikinci kısımlar, çok büyük ve çok küçük alana sahip vücut bölümlerini elemek için kullanılır. Böylece, vücut bölümlerinin üst üste örtüşmesinden meydana gelen büyük alan değerleri ile gürültüden meydana gelen küçük alan değerleri poz tahmin sistemini kararsız bir sonuca sürükleyemez.

$p(X_j)$, vücut bölümlerine ait parametrelerin öncesel bilgisini ifade eder ve her bir parametre bağımsız değerlendirilir. Buna göre öncesel bilgi aşağıdaki gibi ifade edilebilir;

$$p(X_j) = p(x_j, y_j)p(h_j)p(f_j)p(i_j)p(v_j)p(a_j)p(l_j)p(o_j)p(\theta_j)p(\phi_j) \quad (5.26)$$

$P(x_j, y_j)$, imgedeki insan silüetini çevreleyen dikdörtgen üzerinde tekbiçimli (uniform) bir dağılım göstermektedir.

$P(h_j)$, $[h_{min}, h_{mak}]$ değerleri ile sınırlanmış bir $\mathcal{N}(\mu_h, \sigma_h^2)$ Gaussian dağılımı ile modellenir. μ_h , silüete göre ayarlanmış yapay modelin eklem uzunluklarıdır. σ_h ise $\mu_h/10$ olarak alınır. $P(f_j)$, $[f_{min}, f_{mak}]$ değerleri ile sınırlanmış bir $\mathcal{N}(\mu_f, \sigma_f^2)$ Gaussian dağılımı ile modellenir. μ_f , yapay modeli oluşturan eklemleri dik kesen çizgilerin uzunluklarıdır. σ_f ise $\mu_f/5$ olarak alınır.

Vücut bölümlerinin hız ve ivme modellerinin oluşturulabilmesi için Mocap veritabanında bulunan örnek bir video kaydına ait hareket bilgileri kullanılır. Her bir eklem merkezinin zamana göre pozisyon verisi hesaplandıktan sonra bu verinin ilk türevi ile hız, ikinci türevi ile ivme bilgisi elde edilir. Böylece her bir ekleme ait $P(v_j)$ hız ve $P(a_j)$ ivme bilgisi, $[v_{min}, v_{mak}]$ ve $[a_{min}, a_{mak}]$ değerleri ile sınırlanmış doğrusal bir model ile ifade edilir.

$P(I_j)$ ile vücut bölümlerinin renksel görünüm modelleri ifade edilir. Her bir modelin renk histogramı, yapay modelin imge üzerine izdüşümü sonucu elde edilir. Kırmızı, yeşil ve mavi renk değerlerini içeren her bir imgenin renk histogramı 512 parça ile ifade edilir. Histogram tabanlı nesne fonksiyonlarını optimize edebilmek için literatürde sık kullanılan “ortalama kayma” algoritması [93, 96] kullanılır.

$P(i_j)$, $P(o_j)$, $P(\theta_j)$ ve $P(\emptyset_j)$ vücut sınırları göz önüne alınarak aşağıda belirtildiği gibi tekbiçimli bir dağılım ile modellenir;

$$p(\theta_j) = \begin{cases} \frac{1}{us_{\theta_j} - as_{\theta_j}} & as_{\theta_j} \leq \theta_j \leq us_{\theta_j} \\ 0 & \theta_j < as_{\theta_j} \text{ veya } us_{\theta_j} < \theta_j \end{cases} \quad (5.27)$$

Denklem (5.27)’de kullanılan us ve as değişkenleri, eklem açısının üst ve alt sınırını göstermektedir. Belirtilen sınır değerleri sabittir. Kişiden kişiye farklılık göstermez.

Denklem (5.9)’daki eşitliğin ikinci terimi $P(s_t | I_{1,...,t-1})$ şu şekilde hesaplanır. s_{t-1} ile geçmiş gözlemlerden gerekli bilgilerin kodlandığı göz önüne alınırsa, $P(s_t | s_{t-1})$ aşağıdaki gibi ifade edilir;

$$P(s_t | s_{t-1}) = \prod_{i=1}^{19} P(X_i^t | X_i^{t-1}) \quad (5.28)$$

Her bir vücut bölümü için tanımlanan geçici bilgi şu şekilde ifade edilir;

$$P(X_i^t | X_i^{t-1}) = \begin{cases} P(X_i^t | X_i^{t-1}), & k_i^t = k_i^{t-1} \\ P(X_i^t), & X_i^{t-1} = \emptyset \\ P(X_i^{t-1}), & X_i^t = \emptyset \end{cases} \quad (5.29)$$

Denklem (5.29)’daki $P(X_i^t)$ ve $P(X_i^{t-1})$ değişkenleri, geçmiş ve güncel gözlemleri ifade edilir. Eğer, t anında X_i^t ve X_i^{t-1} bilgisi mevcutsa, vücut bölümlerine ait özelliklerin (pozisyon, hız, yükseklik, genişlik ...) modellenebilmesi için sabit hız ve sabit Gaussian gürültüsü kullanılır. İki zaman dilimi arasındaki geçici tahminin yapılabilmesi için kullanılabilecek en iyi yöntem Kalman süzgeci yöntemidir. $X_i^t = \emptyset$ eşitliği ile k_i^{t-1} . vücut bölümü, vücudun bir başka bölümü tarafından engellendiği kastedilir. $X_i^{t-1} = \emptyset$ ile k_i^t vücut bölümünün yeniden görünmeye başladığı ifade edilir.

5.6.4. Maksimum Olabilirlik Değerinin Hesaplanması

Bu bölümde, aday durum kümesi s_t , $p(I_t|s_t)$ ile ifade edilen maksimum olabilirlik fonksiyonu tarafından elde edilir. Önerilen maksimum olabilirlik fonksiyonu iki temel bileşen içerir: 1) Vücut bölümlerinin renksel görünümü; 2) Şekilsel bilgisi. Daha formal bir ifadeyle, $p(I_t|s_t)$ fonksiyonu şu şekilde yazılabilir;

$$P(I_t|s_t) = \prod_{j=1}^{19} p(I_t|X_j^t) \\ \propto \prod_{k=1}^{19} p(I_t^{\text{renk}}|X_j^t) p(I_t^{\text{sekil}}|X_j^t) \quad (5.30)$$

Denklem (5.30)'daki $p(I_t^{\text{renk}}|X_j^t)$ ve $p(I_t^{\text{sekil}}|X_j^t)$ değişkenleri, j . vücut bölümünün renk ve şekil bilgisi için maksimum olabilirlik değerlerini göstermektedir. Renksel maksimum olabilirlik değeri için 3-boyutlu yapay insan modelinin siluet üzerine düşen izdüşüm pikselleri kullanılır. Buna göre renk bilgisi için maksimum olabilirlik değeri şu şekilde ifade edilebilir;

$$p(I_t^{\text{renk}}|\theta_j^t \& \phi_j^t) = \frac{1}{19} \sum_{i=1}^{19} K(p_i^j, \hat{p}_i^j, h^j) \quad (5.31)$$

Denklem (5.31)'deki p^j ve $\hat{p}^j$, j . vücut bölümünün model ve aday renk dağılımlarını gösterir. h bant genişliğine sahip $K(\cdot)$, radyal simetrik çekirdek fonksiyonu aşağıdaki gibi ifade edilir;

$$K(p, \hat{p}, h) = \frac{1}{(2\pi)^{d/2} h^d} e^{-\|p - \hat{p}\|^2 / 2h^2} \quad (5.32)$$

Her bir vücut bölümünün şekilsel olabilirlik değeri aşağıdaki gibi tanımlanır;

$$p(I_t^{\text{sekil}}|\theta_j^t \& \phi_j^t) = \frac{1}{19} \sum_{i=1}^{19} (m_i^j \cap B) \quad (5.33)$$

Denklem (5.33)'teki m değişkeni, imge üzerine izdüşümü alınan noktaların koordinatlarını ifade etmektedir. B ise Bölüm 3.1'de detaylı bahsedilen dinamik arka plan çıkarım

algoritmasının iki seviyeli sonucudur. İzdüşüm noktaları ile arka plan çıkarım sonucunun kesişimini ifade eden Denklem (5.33) ile izdüşümü alınan vücut noktalarının kaç tanesinin imge üzerindeki siluet içerisine düştüğü tespit edilmektedir.

5.7. Uygulama Sonuçları ve Değerlendirilmesi

Önerilen insan poz takip yöntemi, Mocap veritabanındaki “yürüme”, “zıplama” ve “koşma” hareketlerini içeren 3 video kaydı üzerinde test edilmiştir. Elde edilen test sonuçları şekil ve tablolar halinde listelenerek önerilen yöntemin başarımı değerlendirilmiştir. Uygulamada kullanılan video kayıtları 240x352 piksel² boyutlarında ve 4 MP çözünürlüğe sahiptir. Bu video kayıtları saniyede 30 imge yakalayabilen Vicon Kızılötesi MX-40 kameralar ile elde edilmiştir. Ayrıca her video kaydında farklı vücut parametrelerine sahip bir figüran bulunmakta ve bu figüranlar yaklaşık 3x8 metrelik bir dikdörtgen alan içerisinde farklı hareketler gerçekleştirmektedirler.

Gerçekleştirilen ilk uygulamada önerilen insan poz takip yöntemi, Mocap veritabanında 01_01 kodu ile indekslenmiş 744 imge içeren “koşma” adlı video kaydı üzerinde test edilmiştir. Elde edilen poz takip sonucu Şekil 5-12’de gösterilmiştir. Şekil 5-12’deki ilk ve ikinci satırlarda orijinal video imgeleri ve bu imgeler üzerinde dinamik arka plan çıkarım algoritmasının çalışma sonucu gösterilmektedir. Üçüncü satırda, önerilen poz takip algoritmasının 3-boyutlu yapay model sonucu gösterilmektedir. Gösterilen takip sonucundan da anlaşılabacağı üzere, en büyük takip hatası eller ve ayaklarda meydana gelmektedir. Çünkü iskelet modelinin ağaç yapısı hatırlandığı zaman görülür ki, el ve ayak uzuvları ağaç modelinin en alt seviyesinde bulunmaktadır. Alt seviyedeki düğümler üst seviyedeki düğümlerde meydana gelen takip hatasından doğrudan etkilenir. Dolayısıyla, bu düğümlerin takibi sonucunda elde edilen hata değeri üst seviyede meydana gelen takip hatalarını kapsamaktadır.

Şekil 5-12(c) ve 5-13(c)’ye dikkatli bakılırsa en büyük takip hatasının sol el üzerinde gerçekleştiği görülür. Bunun iki nedeni vardır. Birincisi, eller, vücuttaki en hızlı hareketin meydana geldiği uzuvlardır. Ellerin hızlı hareket etmesi takip doğruluğunun azalmasını sağlamıştır. İkincisi, elde edilen video kayıtlarında kamera pozisyonu vücudun sağ tarafında olduğu için hareket esnasında sol elin görünümü arasına kaybolmakta ve vücut gövdesi arkasına gizlenmektedir. Bu nedenle, önerilen takip algoritmasının takip hatası en fazla bu uzuvlarda meydana gelmektedir. Şekil 5-12’nin 4. ve 5. satırlarında 3-

boyutlu yapay modelin perspektif izdüşüm sonucu ve insan iskelet modeli gösterilmektedir. Gerçekleştirilen tüm uygulamalarda perspektif izdüşüm hesaplanırken, kamera ve izleyicinin pozisyonları sabit bir nokta olarak alınmıştır.

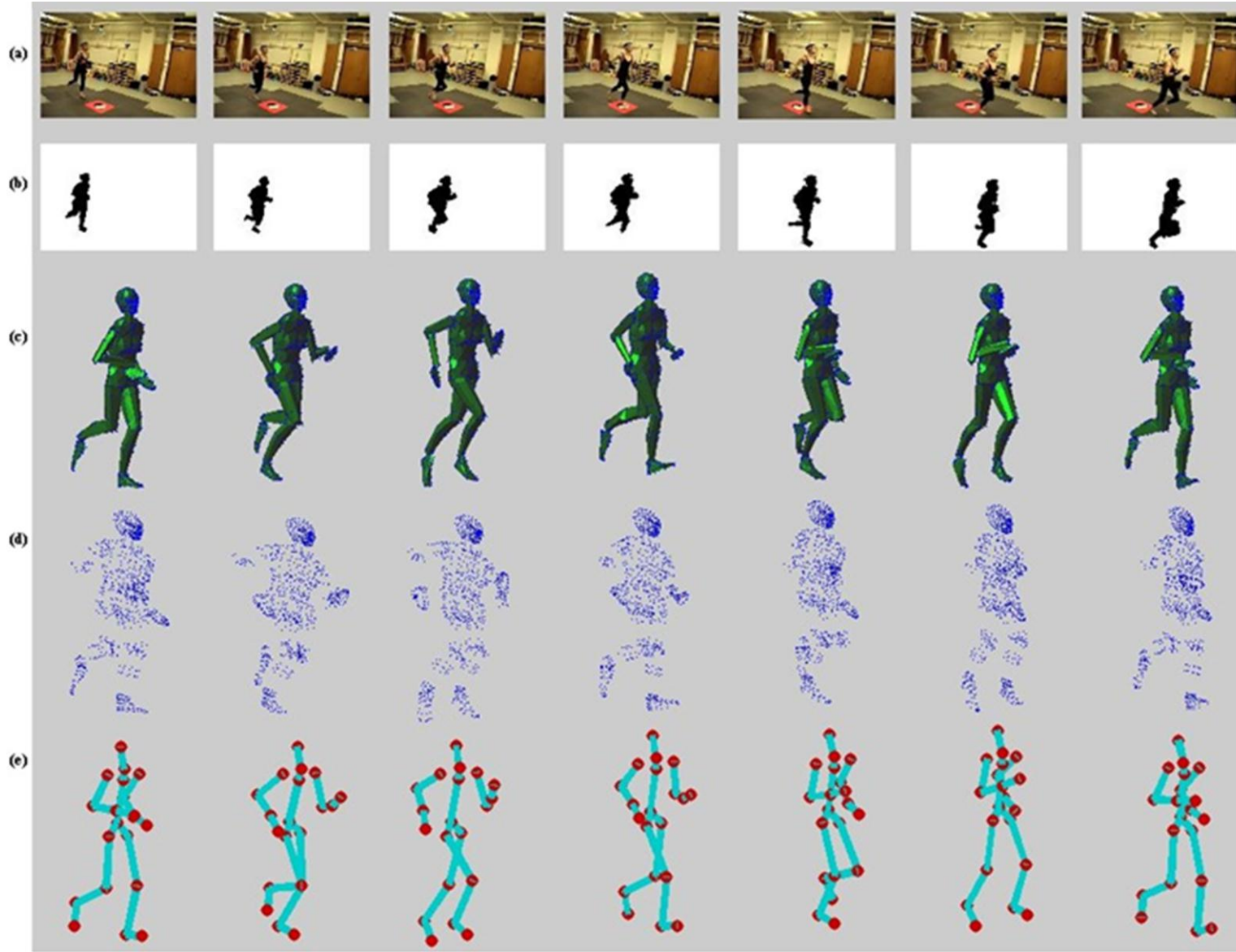
Önerilen poz takip yönteminin ikinci uygulaması “zıplama” adlı video kaydı üzerinde test edilmiştir. Bu video kaydı 240 imge içermektedir ve Mocap veritabanında 08_10 kodu ile indekslenmektedir. Önerilen poz takip yönteminin bu video kaydını üzerindeki takip sonuçları Şekil 5-13’de gösterildiği gibidir.

Önerilen poz takip yönteminin son uygulaması “yürüme” adlı video kaydı üzerinde gerçekleştirilmiştir. Bu video kaydı toplam 130 imge içermektedir ve Mocap veritabanında 02_03 kodu ile indekslenmektedir. Önerilen poz takip sisteminin bu video kaydını üzerindeki takip sonuçları Şekil 5-14’de gösterildiği gibidir.

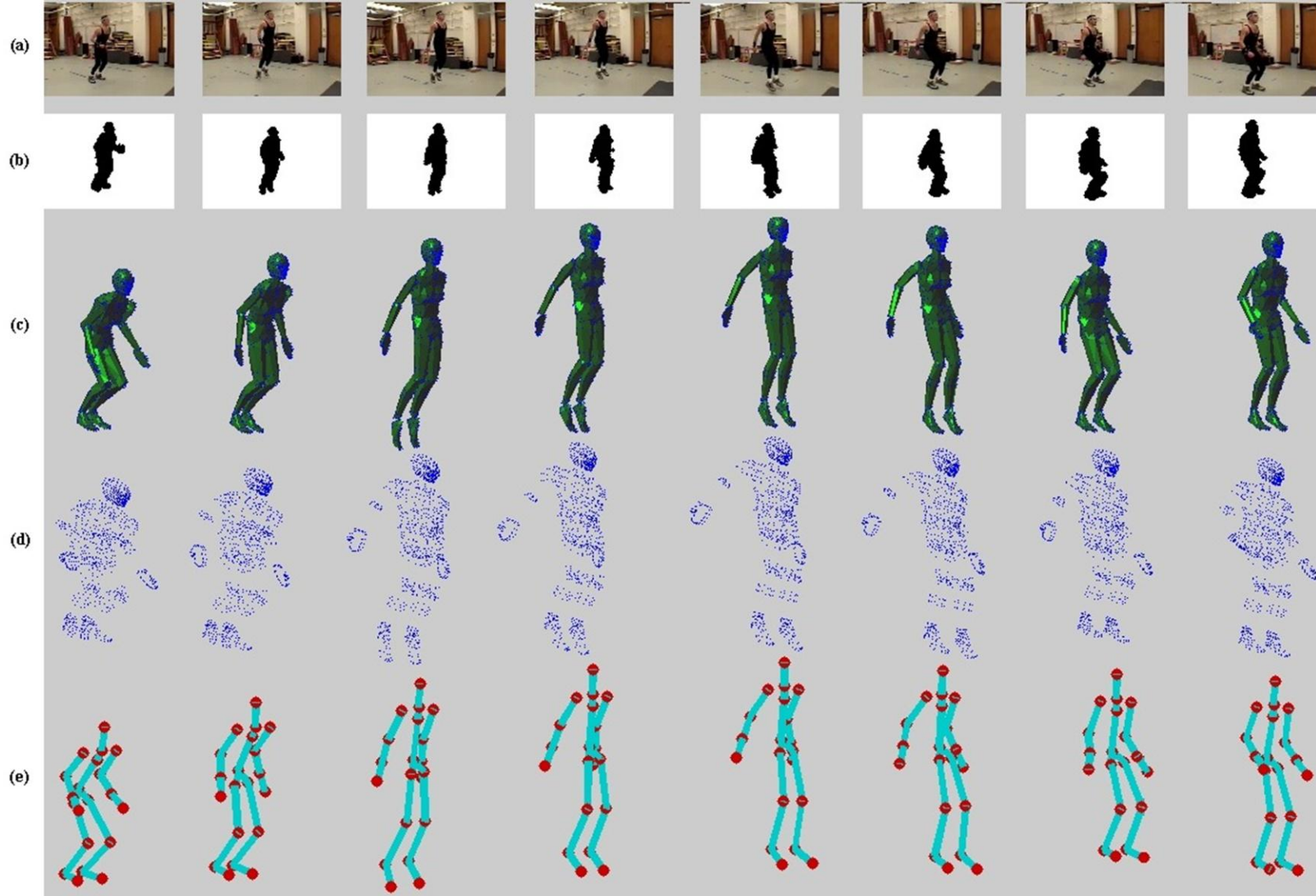
“Koşma”, “zıplama” ve “yürüme” video kayıtları için önerilen poz takip yönteminin takip doğruluğunu belirlemek ve hangi vücut bölümlerinin hangi oranlarda takip edebildiğini görebilmek amacıyla, önerilen takip yöntemine ait sonuçlar ile orijinal hareket bilgileri karşılaştırılmaktadır. Orijinal hareket bilgileri, Mocap veritabanında kayıtlı bulunan “amc” uzantılı dosyalarda bulunmaktadır. Bu dosyalar hareket boyunca figuran üzerindeki işaretçilerin 3-boyutlu koordinat eksenindeki değişim bilgilerini içermektedir. Önerilen poz tahmin yöntemi ile orijinal hareket bilgileri arasındaki hata değeri aşağıdaki gibi hesaplanmaktadır;

$$\mathbf{err} = \frac{\sum_{i=1}^{FN} \sqrt{(x_i - x'_i)^2 + (y_i - y'_i)^2}}{FN} \quad (5.34)$$

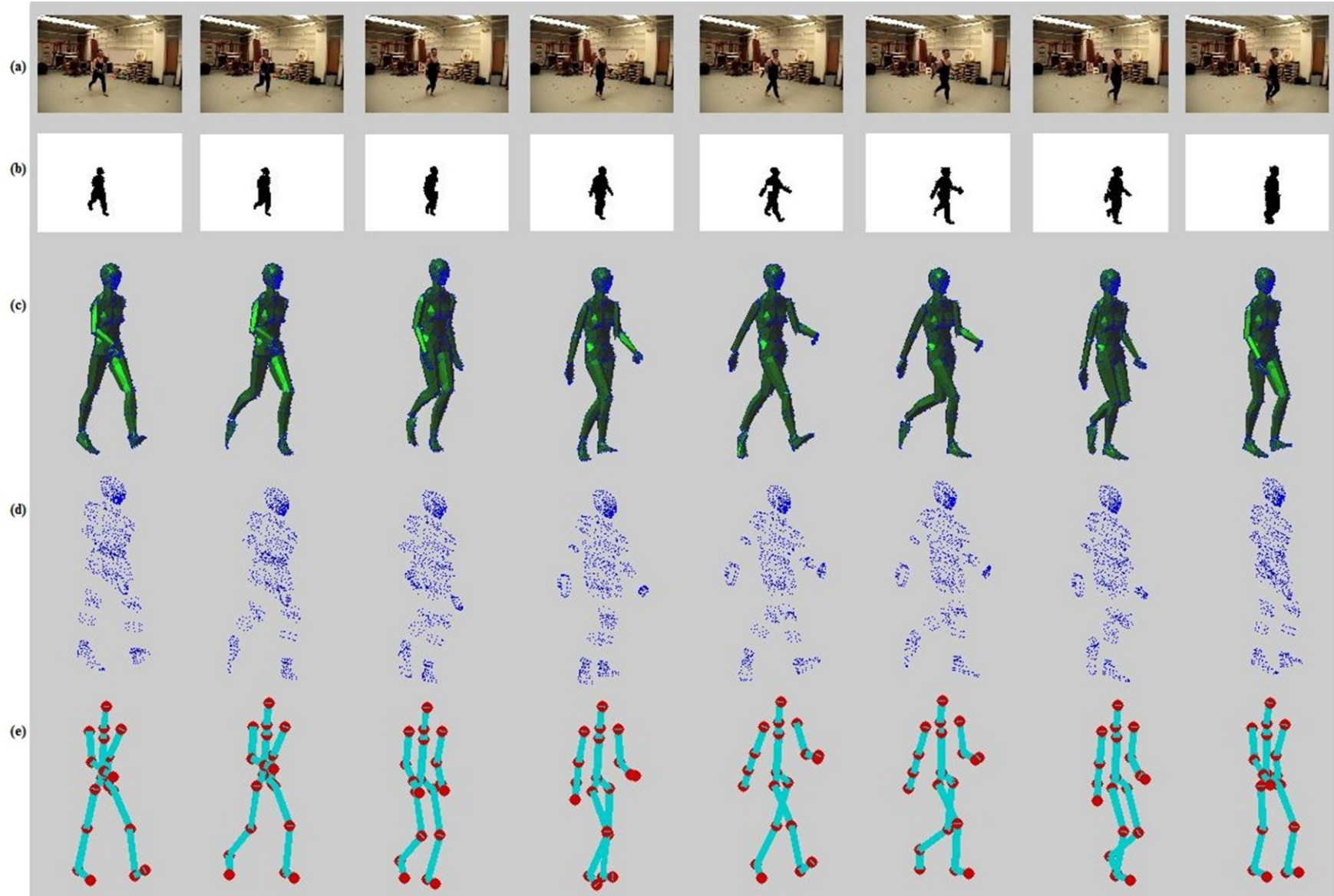
Denklem (5.34)’deki FN değişkeni, video kaydındaki toplam imge sayısını; x_i ve y_i değişkenleri, Mocap veritabanındaki orijinal işaretçi koordinat bilgilerini; x'_i ve y'_i değişkenleri önerilen poz takip sistemi sonucu elde edilen işaretçi koordinat bilgilerini ve err değişkeni ise ortalama hata değerini göstermektedir. Bu değerlendirme ölçütü kullanılarak “yürüme”, “koşma” ve “zıplama” videoları için önerilen poz takip sistemi tarafından elde edilen vücut uzuvları üzerindeki ortalama piksel hataları Tablo 4’de gösterildiği gibidir. Tablo 5’de ise önerilen poz tahmin yönteminin farklı video kayıtları üzerindeki performans değerleri gösterilmektedir. Bu tabloda, dinamik arka plan çıkarımı için harcanan zaman ile poz takip işlemi için harcanan zaman ayrı listelenmiştir.



Şekil 5-12. Gerçekleştirilen poz tahmin yönteminin “koşma” videosu üzerindeki sonuçları **(a)** Orjinal imge **(b)** İnsan silueti **(c)** 3-boyutlu yapay insan modeli **(d)** Yapay modelinin imge üzerindeki izdüşüm sonucu **(e)** İnsan iskelet modeli.



Şekil 5-13. Gerçekleştirilen poz tahmin yönteminin “zıplama” videosu üzerindeki sonuçları **(a)** Orjinal imge **(b)** İnsan silueti **(c)** 3-boyutlu yapay insan modeli **(d)** Yapay modelinin imge üzerindeki izdüşüm sonucu **(e)** İnsan iskelet modeli.



Şekil 5-14. Gerçekleştirilen poz tahmin yönteminin “yürüme” videosu üzerindeki sonuçları **(a)** Orjinal imge **(b)** İnsan silueti **(c)** 3-boyutlu yapay insan modeli **(d)** Yapay modelinin imge üzerindeki izdüşüm sonucu **(e)** İnsan iskelet modeli.

Tablo 5-2. Gerçekleştirilen poz tahmin yönteminin vücut bölümleri üzerindeki ortalama piksel hatası.

Eklem	Ortalama Piksel Hatası		
	Koşma	Zıplama	Yürüme
Sol Dirsek	26.45	24.77	19.38
Sol Bilek	31.30	28.54	24.95
Sol Diz	23.97	17.35	16.45
Sol Ayak	24.33	19.89	17.82
Sağ Dirsek	15.45	14.75	10.66
Sağ Bilek	22.34	19.23	15.48
Sağ Diz	12.79	8.20	6.35
Sağ Ayak	14.55	11.79	9.22

Tablo 5-3. Gerçekleştirilen poz tahmin yönteminin farklı video kayıtları üzerindeki performansı

Video	Algoritma	İşlem Zamanı	İmge Sayısı	Başarım
Yürüme	Arka plan çıkarım	160 ms	130	%93
	3-boyutlu poz tahmini	120 ms		
Koşma	Arka plan çıkarım	580 ms	744	%82
	3-boyutlu poz tahmini	450 ms		
Zıplama	Arka plan çıkarım	300 ms	240	%88
	3-boyutlu poz tahmini	250 ms		

Geliştirilen 3-boyutlu insan poz takibi yazılımı, veri görselleme ve analizi ile nümerik hesaplama amaçlı yüksek seviyeli bir programlama dili olan MATLAB 2008B platformu kullanılarak hazırlanmıştır. Hazırlanan yazılım çift çekirdekli 2.66 GHz Pentium IV işlemci ve 2GB RAM belleğe sahip bir masaüstü bilgisayarda çalıştırılmıştır. PC’de kullanılan işletim sistemi Windows7’dir.

6. SONUÇ VE DEĞERLENDİRME

Görüntü tabanlı çözüm sistemlerinin hayatımızdaki birçok alanda yaygın bir şekilde kullanılmasına neden olan birkaç önemli etken şu şekilde sıralanabilir;

- 1) Yüksek kaliteye sahip görüntü bilgisinin düşük maliyetlerde elde edilebilmesi.
- 2) Hızlı veri işleme ve geniş veri depolama aygıtlarının üretilmesi.
- 3) Tek bir görüntü yardımıyla birçok bilginin bir anda ifade edilebilmesi.
- 4) Görüntü ile ifade edilen bilginin insan tarafından hızlıca algılanabilmesi.

Görüntü tabanlı sistemler içerisinde önemli bir yer işgal eden nesne takip sistemleri, nesnelere ait hareket olgusunun algılanması ve bu olgunun ardışık imgeler boyunca takip edilmesi ile ilgilenmektedir [55].

Bu tez çalışmasında, üç farklı gerçek dünya problemine çözüm sağlayabilen üç yeni görsel takip yöntemi geliştirilmiş ve uygulamaları gerçekleştirilmiştir. Geliştirilen ilk takip yöntemi, kalabalık bir ortamda gezinti yapmakta olan insanların imgeler boyunca takip edilmesini ve bağımsız her kişinin farklı etiket değerleri ile ifade edilmesini amaçlamaktadır. Gerçekleştirilen ikinci takip yöntemi yardımıyla hareket halindeki tek bir insanın el ve yüz takibi sonucu belirli bir eylemi gerçekleştirip gerçekleştirmediğinin tespiti yapılmaktadır. Üçüncü takip yönteminde ise tek kamera yardımı ile elde edilen 2-boyutlu imgelerdeki hareketli insan verisi kullanılarak 3-boyutlu insan poz tahmini gerçekleştirilmektedir.

6.1. Sonuçların İrdelenmesi

- 1) Geliştirilen ilk görsel takip yöntemi, [78, 79, 81, 83, 135, 137, 138, 149, 151] çalışmalarının aksine, kalabalık ortamlarda gezinti yapmakta olan birden fazla insanın eş zamanlı takip işlemini gerçekleştirirken insan vücut sunumu için basit bir geometrik şekil (örneğin elipsoid) kullanmak yerine takip edilen her bir insanın vücut siluet bilgisini kullanmaktadır. Takip edilen kişilerin görsel ve şekilsel sunumlarındaki bu iyileştirme, önerilen çoklu insan takip yönteminin sağlamlılığını arttırmakla birlikte takip edilen insan görünümünün büyüüp küçülmesinden kaynaklanan problemlerin [93, 100] ortadan kalkmasını sağlamıştır. Ayrıca geliştirilen çoklu insan takip yöntemi, görünüm

engellenmesi olarak bilinen “kapatma” probleminin çözümü için [78, 79] çalışmalarından farklı olarak yeni bir method kullanmıştır. Önerilen bu yeni method yardımıyla “kısmi kapatma” olayının gerçekleştiği imgelerde takip işlemleri herhangi bir aksamaya uğramadan başarılı bir şekilde sürdürülmektedir. Kapatma probleminin çözümü için önerilen bu yaklaşımın dezavantajı “tam kapatma” olayına istenilen yanıtı veremiyor olmasıdır. Bununla birlikte, önerilen takip yöntemi, ortamın değişen ışık miktarına ve dinamik arka plan bilgisine hızlı bir şekilde adapte olabilmektedir.

- 2) Gerçekleştirilen ikinci takip yöntemi, klasik el ve yüz takip yöntemlerinden [101, 171, 172] farklı olarak hareket halindeki tek bir insanın el ve yüz uzuvlarının takibi sonucunda kişi tarafından belirli bir eylemin gerçekleştirip gerçekleştirmediğinin tespitini yapılabilmektedir. Gerçekleştirilen takip yönteminin uygulaması, Alzheimer hastalarının tedavilerine destek verebilmek amacıyla Alzheimer hastalarına ait bir problemin çözümünde kullanılmıştır. 5 farklı figüranın yer aldığı bu uygulamanın sonuçları, önerilen eylem takip yönteminin istenilen eylemi başarılı algılama oranının %92 olduğunu göstermiştir.
- 3) Gerçekleştirilen üçüncü takip yöntemi yardımıyla 2-boyutlu video imgeleri kullanılarak 3-boyutlu insan poz tahmini benzetimi yapılmaktadır. Bu yöntemde [117, 128, 130, 131, 140, 146, 162, 163] çalışmalarının aksine 2-boyutlu insan silüetine ait hem görsel hemde şekilsel bilgi tek bir Bayesian çatısı altında incelenir ve önceden tasarlanmış 3-boyutlu yapay insan modelinin eklem bölgelerindeki açı değerleri tahmin edilmektedir. İmge içerisindeki 2-boyutlu insan silüeti ile 3-boyutlu yapay insan modeli arasındaki ilişki perspektif izdüşüm yöntemi kullanılarak giderilmektedir. Önerilen yöntemin test aşamasında, MOCAP veritabanında kayıtlı bulunan “yürüme”, “koşma” ve “zıplama” eylemlerine ait video kayıtları kullanılmıştır. Farklı figürlara ait eylemler üzerinde gerçekleştirilen test sonuçları, önerilen poz tahmin yönteminin hareket vektörü düz olan eylemlerde genel vücut pozunun doğru bir şekilde tahmin edilebildiği göstermiştir. Elde edilen matematiksel poz tahmin sonuçlarına göre; el ve ayak gibi küçük boyutlara sahip ve hızlı hareket edebilen vücut uzuvlarında yüksek oranda poz tahmin hataları meydana gelmektedir.

Fakat küçük boyutlu vücut uzuvlarında meydana gelen bu hatalar, genel vücut pozunun tahminini etkileyecek boyutlara ulaşmamaktadır.

6.2. Öneriler

- 1) Tez çalışmasında önerilen ilk yöntem olan kalabalık ortamlarda insan yakalama ve etiketlendirme yöntemi içerisinde kullanılan kısmi kapatma algoritması yerine daha gelişmiş ve tam kapatma durumunu tespit edebilen yeni bir yöntem geliştirilebilir. Ayrıca, önerilen çoklu insan takip yönteminin ani ışık değişimlerine karşı olan duyarlılığı geliştirilerek yanlış insan algılama problemi ortadan kaldırılabilir.
- 2) Geliştirilen çoklu insan takibi yöntemi içerisinde kapatma probleminin çözümü için kapatma olayına maruz kalan hareketli her bir kişinin yarıçap hesabı (Denklem 3.11) yapılırken kullanılan β katsayısının değeri optimizasyon teknikleri yardımıyla hesaplanarak takip performansının artması sağlanabilir.
- 3) Tez çalışmasında önerilen çoklu insan poz tahmin yöntemi ile istatistiksel ten yakalama tabanlı insan hareket tespit yöntemi birlikte kullanılarak daha karmaşık vücut hareketlerine ait poz tahminleri gerçekleştirilebilir.
- 4) Tez çalışmasında önerilen insan poz tahmin yöntemi, herhangi bir canlının veya nesnenin poz tahmini için kullanılabilir. Bunun için gerçek dünya ortamında hareket etmekte olan canlının vücut eklem bilgilerinin 3-boyutlu koordinat bilgilerine ihtiyaç duyulmaktadır. Harici bir sistem kullanılarak bu bilgiler elde edildiği takdirde her hangi bir canlıya ait poz tahmini gerçekleştirilebilir.
- 5) Doğada birçok canlının şekil ve yapısında bulunan *altın oran* değeri, insan poz tahmini yöntemi içerisinde de kullanılabilir. Bu sayede, poz tahmini gerçekleştirilecek kişinin vücut bölümlerine ait uzunluk bilgileri altın oran kullanılarak elde edilebilir.

KAYNAKLAR

- [1] http://en.wikipedia.org/wiki/Human_visual_system_model.2010.
- [2] http://en.wikipedia.org/wiki/Computer_vision. 2010.
- [3] Yilmaz, A., Javed, O., and Shah, M., 2006. "Object Tracking: A Survey", ACM Comput. Surv., 38(4):13.
- [4] Harada T, Shirato H, Ogura S, Oizumi S, Yamazaki K, Shimizu S, Onimaru R, Miyasaka K, Nishimura M and Dosaka-Akita H., 2002. "Real-time tumor-tracking radiation therapy for lung carcinoma by the aid of insertion of a gold marker using bronchofiberscopy", Cancer; 95(8):1720-7.
- [5] Shirato H., 2003. "Feasibility of insertion/implantation of 2.0-mm-diameter gold internal fiducial markers for precise setup and real-time tumor tracking in radiotherapy". Int. J. Radiat. Oncol. Biol. Phys. 56 240–7.
- [6] Bas, E., Salman, S., and Tekalp, A.M., 2007. "Automatic Vehicle Counting from Video for Traffic Flow Analysis", Proceedings of the 2007 IEEE Intelligent Vehicles Symposium, Istanbul, Turkey.
- [7] Nielsen, J. and Pernice, K. 2010. "Eyetracking Web Usability",New Riders, California USA, ISBN-10: 0-321-49836-4.
- [8] Talu, M.F., Soyguder, S., and Aydogmus, O., 2010. "An implementation of a novel vision-based robotic tracking system", Sensor Review, Vol:30, Issue:3, pp. 225-232.
- [9] David A. F., Arikan, O., Ikemoto, L., O'Brien, J., and Ramanan, D., 2005. "Computational Studies of Human Motion:Part 1, Tracking and Motion Synthesis", Computer Graphics and Vision, 1(2/3):77–254.
- [10] D.J. Sheskin, 2003. "Handbook of Parametric and Nonparametric Statistical Procedures", Chapman & Hall/CRC.
- [11] http://en.wikipedia.org/wiki/Video_tracking. 2010.
- [12] Bettencourt, K., Somers, D. C., 2007. "Effects of task difficulty on multiple object tracking performance", Journal of Vison, 7(9):898, 898a.

- [13] Veenman, C., Reinders, M., and Backer, E., 2001. "Resolving motion correspondence for densely moving points", IEEE Trans. Patt. Analy. Mach. Intell. 23(1):54–72.
- [14] Shafique, K. and Shah, M., 2003. A non-iterative greedy algorithm for multi-frame point correspondence. IEEE International Conference on Computer Vision (ICCV). 110–115.
- [15] Comaniciu, D. Ramesh, V. Meer, P., 2003. "Kernel-based object tracking", IEEE Transactions on Pattern Analysis and Machine Intelligence, 25(5): 564- 577.
- [16] Yilmaz, A., Li, X. and Shah, M., 2004. "Contour-Based Object Tracking with Occlusion Handling in Video Acquired Using Mobile Cameras", IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 26, No. 11.
- [17] Talu, M.F., Türkoğlu, İ., 2009. "A Novel Object Recognition Method Based on Improved Edge Tracing", Application of Information and Communication Technologies (AICT), pages: 1-5, Azerbaijan.
- [18] Talu, M.F., Tatar, Y., 2003. "A new edge tracing method developed for object recognition", Turkish Symposium on Artificial Intelligence and Neural Network (TAINN), pages: 434:438, Çanakkale.
- [19] Horaud, R., Niskanen, M., Dewaele, G., Boyer, E., 2009. "Human Motion Tracking by Registering an Articulated Surface to 3D Points and Normals", Transactions On Pattern Analysis And Machine Intelligence, Vol. 31, No. 1.
- [20] Ali, A. and Aggarwal., J., 2001. "Segmentation and recognition of continuous human activity", IEEE Detection and Recognition of Events in Video. 28-35.
- [21] Ushakov, N.G., 2001, "Density of a probability distribution", in Hazewinkel, Michiel, Encyclopaedia of Mathematics, Kluwer Academic Publishers, ISBN 978-1556080104.
- [22] Evans, M.; Hastings, N.; and Peacock, B. 2000. "Probability Density Function and Probability Function." §2.4 in Statistical Distributions, 3rd ed. New York: Wiley, pp. 9-11.
- [23] [http://en.wikipedia.org/wiki/Kernel_\(statistics\)](http://en.wikipedia.org/wiki/Kernel_(statistics)). 2010.
- [24] http://en.wikipedia.org/wiki/Normal_distribution. 2010.
- [25] Flannery, W. H., Teukolsky, B. P., and Vetterling, W. T., 1992. "Numerical Recipes in FORTRAN", The Art of Scientific Computing, Cambridge University Press, pp. 554-556.

- [26] Bohyung Han; Comaniciu, D.; Ying Zhu; Davis, L.S.; 2008, “Sequential Kernel Density Approximation and Its Application to Real-Time Visual Tracking”, IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(7): 1186 – 1197.
- [27] R. T. Collins, 2003. “Mean-shift Blob Tracking through Scale Space”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol.2. pp. 234-240.
- [28] Cootes, T. F., Edwards, G. J., and Taylor, C. J., 1998. “Active Appearance Model”, Proc European Conference on Computer Vision, Vol. 2, pp. 484-498, Springer.
- [29] Cootes, T. F., Taylor, C. J., Cooper, D. H., and Graham, J., 1995. “Active Shape Models-Their training and application”, Computer vision and image understanding, Vol. 61. No:1, pp.38-59.
- [30] Lee, S.-W., Kang, J., Shin, J., Paik, J., 2007. “Hierarchical active shape model with motion prediction for real-time tracking of non-rigid objects”, IET Comput. Vis., 2007, 1, (1), pp. 17–24.
- [31] W. L. Martinez, A. R. Martinez, 2002, “Computational Statistics Handbook with MATLAB”, Chapman & Hall/CRC. Chapter 8.
- [32] Fieguth, P. and Terzofoulos, D. 1997. Color-based tracking of heads and other mobile objects at video framerates. In IEEE Conference on Computer Vision and Pattern Recognition. 21-27.
- [33] Kim, D. H. and Kim, M., 2006. “Skin Color Analysis in HSV Color Space and Rendering with Fine Scale Skin Structure”, Springer-Verlag Berlin Heidelberg, CGI, LNCS 4035, pp. 254–264.
- [34] Vezhnevets, V., Sazonov, V., Andreeva, A., 2003. “A Survey on Pixel-Based Skin Color Detection Techniques”, in Proc. Graphicon-2003, pages: 85—92.
- [35] Talu, M.F., Tatar, Y., 2005. "Model Tabanlı Yapay Sinir Ağı Kullanılarak Geliştirilmiş Bir Kenar Çıkarma Metodu", TBV Bilgisayar Bilimleri ve Mühendisliği Dergisi, Cilt/vol:1, Sayı/Number:01, Sayfa 77-84.
- [36] Talu, M. F., Tatar, Y., 2004. "Model tabanlı yapay sinir ağları kullanılarak görüntü kenarının yakalanması", Fırat Üniversitesi, Yüksek lisans tezi.
- [37] Kelson R. T. Aires, Andre M. Santana, Adelardo A. D. Medeiros, 2008. “Optical Flow Using Color Information”, ACM New York, NY, USA. ISBN 978-1-59593-753-7.

- [38] David J. Fleet and Yair Weiss, 2006. "Optical Flow Estimation". in Paragios et al.. Handbook of Mathematical Models in Computer Vision. Springer. ISBN 0387263713.
- [39] Amiaz, T., Fazekas, S., Chetverikov, D., and Kiryati, N., 2007. "Detecting regions of dynamic texture". In: Int. Conf. on Scale Space and Variational Methods in Computer Vision (SSVM 2007), LNCS 4485, pp. 848-859.
- [40] Fazekas, S., and Chetverikov, D., 2007. "Analysis and performance evaluation of optical flow features for dynamic texture recognition". Signal Processing: Image Communication, 22:680-691.
- [41] Jolliffe I.T., 2002. "Principal Component Analysis", Series: Springer Series in Statistics, 2nd ed., Springer, NY, XXIX, 487 p. ISBN 978-0-387-95442-4.
- [42] Tieu, K., and Viola, P., 2004. "Boosting image retrieval". Int. J. Comput. Vision 56, 1, 17–36.
- [43] <http://www.vrealities.com/cyber.html>. 2010.
- [44] Russ, J.C., 2002. "The Image Processing Handbook", CRC Press LLC North Carolina State University.
- [45] Jahne, B., 1997. "Digital Image Processing Concept, Algorithms, and Scientific Applications", Springer, 4th Edition.
- [46] Ritter, G. X., Wilson, J. N., 2001. "Handbook of Computer Vision Algorithms in Image Algebra", Boca Raton London New York Washington, D.C.
- [47] Gonzalez, R., and Wintz, P., 1987. Digital Image Processing. Reading, MA: Addison Wesley, second ed.
- [48] Todd, R.R, 2005. "Digital Image Sequence Processing, Compression, and Analysis", CRC Press, London, ISBN 0-8493-1526-3.
- [49] Ioannis, P., 1993. Digital Image Processing Algorithms, Prentice hall international series in signal processing.
- [50] Zivkovic, Z., Heijden, F. V., 2005. "Efficient adaptive density estimation per image pixel for the task of background subtraction", Pattern recognition letters, Volume 27(7), Pages: 773-780.
- [51] Chen, L., Chen, J., Liao, H., 2000. "Object Segmentation for Video Coding", 15th International Conference on Pattern Recognition (ICPR'00), vol. 3, pp.3387.
- [52] Berbeco, R. I., Jiang, S. B., Sharp, G. C., Chen, G. T., Mostafavi, H., and Shirato H. 2004. Integrated radiotherapy imaging system (IRIS): design considerations of

tumour tracking with linac gantry-mounted diagnostic x-ray systems with flat-panel detectors. *Phys. Med. Biol.* 49 243–55.

- [53] Berbeco, R. I., Mostafavi, H., Sharp, G. C. and Jiang, S. B. 2005. Towards fluoroscopic respiratory gating for lung tumours without radiopaque markers *Phys. Med. Biol.* 50 4481–90.
- [54] Haritaoglu, I., Harwood, D., and Davis, L. 2000. W4: Real-time surveillance of people and their activities. *IEEE Trans. Patt. Anal. Mach. Intell.* 22, 8, pp.809–830.
- [55] Chen, S., and Shyu, M., 2000. “Object Tracking and Multimedia Augmented Transition Network for Video Indexing and Modeling”, 12th IEEE International Conference on Tools with Artificial Intelligence (ICTAI 2000), pages: 250 – 257.
- [56] Ocakli, M., 2007. “A video tracker system for traffic monitoring and analysis”, Middle East Technical University, Master Dissertation.
- [57] Alarcon, J.; Salvador, R.; Moreno, F.; Cobos, P.; Lopez, I., 2006. “A new Real-Time Hardware Architecture for Road Line Tracking Using a Particle Filter”, *IEEE Industrial Electronics*, Page(s): 736 – 741.
- [58] Kuhn, H., 1955. “The hungarian method for solving the assignment problem”. *Naval Research Logistics Quart.* 2, 83–97..
- [59] Shafique, K. and Shah, M., 2003. “A non-iterative greedy algorithm for multi-frame point correspondence”.In *IEEE International Conference on Computer Vision*, Vol:27, No:
- [60] Guopu Zhu; Shuqun Zhang; Xijun Chen; Changhong Wang, 2007, “Efficient Illumination Insensitive Object Tracking by Normalized Gradient Matching”, *IEEE Signal Processing Letters*, 14(12): 944 – 947.
- [61] Serry, D., Koller-Meier, S., and Gool., L. V. 2004. Probabilistic object tracking using multiple features. In *IEEE International Conference of Pattern Recognition (ICPR)*. 184-187.
- [62] Schmidt, S. F., 1981. “The Kalman filter - Its recognition and development for aerospace applications“,*Journal of Guidance Control and Dynamics*, 0731-5090 vol.4 no.1 (4-7).
- [63] Julier, S.J, and Uhlmann, J. K, 1997. “A New Extension of the Kalman Filter to Nonlinear Systems”, *SPIE AeroSense Symposium*, April 21–24, Orlando.
- [64] Tanizaki, H., 1987. Non-gaussian state-space modeling of nonstationary time series. *J. Amer. Statist. Assoc.*82, 1032–1063.

- [65] Leven, W.F., Lanterman, A.D., 2009, "Unscented Kalman Filters for Multiple Target Tracking With Symmetric Measurement Equations", Proceedings of SPIE, the International Society for Optical Engineering, 5810:56-67.
- [66] Xue J, Zheng N, Geng J, Zhong X., 2008, "Tracking multiple visual targets via particle-based belief propagation", IEEE Trans Syst Man Cybern B Cybern. 38(1):196-209.
- [67] Smal, K. Draegestein, N. Galjart, W. Niessen, E. Meijering, 2008, "Particle Filtering for Multiple Object Tracking in Dynamic Fluorescence Microscopy Images: Application to Microtubule Growth Analysis", IEEE Transactions on Medical Imaging, 27(6):789-804.
- [68] Chang, W.Y., Chen, C.S., Jian, Y.D., 2008. "Visual Tracking in High-Dimensional State Space by Appearance-Guided Particle Filtering", 17(7):1154-67.
- [69] Maskell, S. Gordon, N., 2001, "A tutorial on particle filters for on-line nonlinear/non-Gaussian Bayesian tracking", 2(2):1-15.
- [70] Doucet, A.; Johansen, A.M., 2008, "A tutorial on particle filtering and smoothing: fifteen years later". Technical report, Department of Statistics, University of British Columbia.
- [71] Oh, S., Russell, S., and Sastry, S. S., 2009, Markov chain Monte Carlo data association for multi-target tracking. IEEE Transactions on Automatic Control, 54(3), 481–497.
- [72] Bar-shalom, Y. and Foreman, T., 1988. "Tracking and Data Association". Academic Press Professional, ISBN:0-120-79760-7.
- [73] Bar-Shalom, Y., Li, X.-R., and Kirubarajan, T., 1995. "Multitarget-Multisensor Tracking: Principles and Techniques". YBS.
- [74] Blackman, S., Popoli, R., 1999. "Design and Analysis of Modern Tracking Systems". Artech House Radar Library.
- [75] Stone, L. D., Barlow, C. A., Corwin, T. L., 1999. "Bayesian Multiple Target Tracking". Artech House, Boston, London.
- [76] Reid, D. B., 1979. "An algorithm for tracking multiple targets". IEEE Trans. Autom. Control 24, 6, 843–854.
- [77] Streit, R. L. and Luginbuhl, T. E., 1994. "Maximum likelihood method for probabilistic multi-hypothesis tracking", In Proceedings of the International Society for Optical Engineering (SPIE.) vol. 2235. 394–405.

- [78] Weiming Hu Xue Zhou Min Hu Maybank, S., 2009. "Occlusion Reasoning for Tracking Multiple People", IEEE Transactions on Circuits and Systems for Video Technology, 19(1):114-121.
- [79] Saad M. Khan, Mubarak Shah, 2009, "Tracking Multiple Occluding People by Localizing on Multiple Scene Planes", IEEE Transactions on Pattern Analysis and Machine Intelligence, 31(3):505-519.
- [80] Salari, V. and Sethi, I. K., 1990, "Feature point correspondence in the presence of occlusion", IEEE Trans.Patt. Analy. Mach. Intell. 12(1):87–91.
- [81] Zhao, T., Nevatia, R., Wu, B. 2008. Segmentation and Tracking of Multiple Humans in Crowded Environments. IEEE Trans. on pattern analysis and machine intelligence. Vol:30, No:7.
- [82] Fleuret, F.; Berclaz, J.; Lengagne, R.; Fua, P., 2008, "Multicamera People Tracking with a Probabilistic Occupancy Map", IEEE Transactions on Pattern Analysis and Machine Intelligence, 30(2): 267 – 282.
- [83] Yan Lu; Payandeh, S., 2008, "Cooperative hybrid multi-camera tracking for people surveillance", Canadian Conference on Electrical and Computer Engineering, p:001365 – 001368.
- [84] Munder, S.; Schnorr, C.; Gavrila, D.M., 2008, "Pedestrian Detection and Tracking Using a Mixture of View-Based Shape–Texture Models", IEEE Transactions on Intelligent Transportation Systems, 9(2): 333 - 343.
- [85] Sankaranarayanan, A.C.; Veeraraghavan, A.; Chellappa, R., 2008, "Object Detection, Tracking and Recognition for Multiple Smart Cameras", Proceedings of the IEEE, 96(10): 1606 – 1624.
- [86] Nascimento, J.C.; Marques, J.S., 2007, "Robust Shape Tracking With Multiple Models in Ultrasound Images", IEEE Transactions on Image Processing, 17(3):392-406.
- [87] Chang, Y. L., and Aggarwal, J.K, 1991. "3d structure reconstruction from an ego motion sequence using statistical estimation and detection theory", In Workshop on Visual Motion. 268–273.
- [88] Chunhua Shen; Brooks, M.J.; van den Hengel, A., 2007, "Fast Global Kernel Density Mode Seeking: Applications to Localization and Tracking", IEEE Transactions on Image Processing, 16(5): 1457 – 1469.
- [89] <http://school.maths.uwa.edu.au/~duongt/seminars/intro2kde>. 2010.

- [90] http://homepages.inf.ed.ac.uk/rbf/CVonline/LOCAL_COPIES/AV0405/MISHRA/kde.html. 2010.
- [91] V. C. Raykar, R. Duraiswami, 2006, “Fast optimal bandwidth selection for kernel density estimation”, Proceedings of the sixth SIAM International Conference on Data Mining, pp. 524-528.
- [92] Zhang, Xibin, King, Maxwell L. and Hyndman, Rob J, 2004, “Bandwidth Selection for Multivariate Kernel Density Estimation Using MCMC”, Monash Econometrics and Business Statistics Working Papers, No 9/04.
- [93] Comaniciu, D., and Meer, P. 2002. Mean shift: A robust approach toward feature space analysis. IEEE Trans. Patt. Analy. Mach. Intell. 24, 5, 603-619.
- [94] Bhattacharyya, A. (1943). "On a measure of divergence between two statistical populations defined by their probability distributions". Bulletin of the Calcutta Mathematical Society 35: 99–109.
- [95] Zhimin Fan; Ying Wu; Ming Yang, 2007, “Multiple Collaborative Kernel Tracking”, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2: 502 – 509.
- [96] Huiyu Zhou, Yuan Yuan and Chunmei Shi, 2008, “Object tracking using SIFT features and mean shift”, Computer Vision and Image Understanding, 113(3):345-352.
- [97] Yilmaz, A., 2007, “Object Tracking by Asymmetric Kernel Mean Shift with Automatic Scale and Orientation Selection”, IEEE Conference on Computer Vision and Pattern Recognition, pp.1-6.
- [98] Talu, M.F., Erdogmus, D., 2009, "A Hybrid Object Tracking Method for Scaled and Oriented Objects", INISTA, Trabzon. Pp.319-322.
- [99] Yang, C., Duraiswami R., and Davis, L., 2005. “Efficient Mean-Shift Tracking via a New Similarity Measure”, Computer Vision and Pattern Recognition. Pages: 176—183.
- [100] Talu, M.F., Türkoğlu, İ., ve Cebeci, M., 2010, “Real-Time Kernel Based Object Tracking using Mean Shift”, IEEE 18. Sinyal İşleme ve İletişim Kurultayı, Diyarbakır.
- [101] Talu, M.F., Erdogmus, D., "Real Time Face and Hand Tracking With Correlation", Innovation in Intelligent Systems and Applications (INISTA), pages:374:377, Trabzon, Haziran 2009.

- [102] Bertalmio, M., Sapiro, G., and Randall, G., 2000. "Morphing active contours". IEEE Trans. Patt. Analy. Mach.Intell. 22, 7, 733–737.
- [103] Cordea, M. D., Petriu, E. M., Petriu, D. C., 2008, "Three-Dimensional Head Tracking and Facial Expression Recovery Using an Anthropometric Muscle-Based Active Appearance Model", IEEE Transactions on Instrumentation and Measurement, 57(8): 1578 – 1588.
- [104] Jehan-Besson, S., and Barlaud, M., 2001. "Video Object Segmentation Using Eulerian Region-Based Active Contours," Proc. IEEE Int'l Conf. Computer Vision.
- [105] Cheung, S-C.S., Kamath. C., 2004. "Robust Techniques for Background Subtraction in Urban Traffic Video", In Proceedings of the SPIE, Volume 5308, pages: 881–892.
- [106] Smith, M., 2006. "Background Subtraction for Urban Traffic Monitoring using Webcams", Master Thesis, University of Amsterdam.
- [107] Elgammal, A., Duraiswami, R., Harwood, D. and Davis, L., 2002. "Background and Foreground Modeling Using Non-Parametric Kernel Density Estimation for Visual Surveillance," Proc. IEEE, vol. 90, no. 7, pages: 1151-1163.
- [108] Zivkovic, Z., Krose, B., 2004. "An EM-like Algorithm for Color Histogram-based Object Tracking," Proceedings of IEEE Conference on Computer Vision and Pattern Recognition, vol 1, pp. 798-803.
- [109] Sethian, J. A., 1999. "Level Set Methods and Fast Marching Methods: Evolving Interfaces in Computational Geometry", Fluid Mechanics, Computer Vision, and Materials Science. Cambridge University Press. ISBN 0-521-64557-3.
- [110] Osher S.J. and Fedkiw., R., 2002. "Level Set Methods and Dynamic Implicit Surfaces", Springer-Verlag, ISBN 0-387-95482-1, page: 95.
- [111] Fukunaga, K. and Hostetler, L., 1975. "The estimation of the gradient of a density function, with applications in pattern recognition". IEEE Transactions on Information Theory, 21(1), 32–40.
- [112] <http://www.openvisor.org/>.2010.
- [113] Lee, J. Y., and Yoo, S. I. 2002. An elliptical boundary model for skin color detection. In Proc. of the 2002 International Conference on Imaging Science Systems and Technology.
- [114] Gomez, G., and Morales, E. 2002. Automatic feature construction and a simple rule induction algorithm for skin detection. In Proc. of the ICML Workshop on Machine Learning in Computer Vision, 31–38.

- [115] Peer, P., Kovac, J., and Solina, F., 2003. "Human skin colour clustering for face detection". EUROCON, International Conference on Computer as a Tool.
- [116] <http://www.alzheimernedir.com/uzmanGorus.asp>. 2010.
- [117] Lee, M.W., Nevatia, R., 2009. "Human Pose Tracking in Monocular Sequence Using Multilevel Structured Models". IEEE Transactions On Pattern Analysis And Machine Intelligence, Vol. 31, No. 1.
- [118] Zhou, H., Hu, H., 2008. "Human motion tracking for rehabilitation—A survey". Biomedical Signal Processing and Control 3, pp. 1–18.
- [119] H. Zhou, H. Hu, 2005. "Inertial motion tracking of human arm movements in stroke rehabilitation", Proceedings of IEEE International Conference on Mechatronics and Automation, Canada.
- [120] H. Zhou, H. Hu, 2005. "Kinematic model aided inertial motion tracking of human upper limb", Proceedings of International Conference Info. Acqu, Hong Kong.
- [121] O. Suess, S. Suess, S. Mularski, B. Kuhn, T. Picht, S. Hammersen, R. Stendel, M. Brock, T. Kombos, 2006., "Study on the clinical application of pulsed dc magnetic technology for tracking of intraoperative head motion during frameless stereotaxy", Head Face Medicine, vol:2:10.
- [122] Roche, A., X. Pennec, G. Malandain, N. Ayache, 2001. "Rigid registration of 3D ultrasound with MR images: a new approach combining intensity and gradient information", IEEE Trans. Med. Imag. 20 (10), 1038–1049.
- [123] X. Pennec, P. Cachier, N. Ayache, 2001. "Tracking brain deformations in time sequences of 3D US images", Proceedings of International Conference Information Processing in Medical Imaging.
- [124] J. Lenz, 1990. "A review of magnetic sensors", Proc. IEEE 78, 973–989.
- [125] Bettencourt, K., Somers, D. C., 2007. "Effects of task difficulty on multiple object tracking performance", Journal of Vision, 7(9):898, 898a.
- [126] <http://www.charndyn.com/>.2010.
- [127] Rius, I., Varona, J., González, J., Villanueva, J., 2006. "Action-specific motion prior for efficient bayesian 3D human body tracking". Proceedings of the 18th International Conference on Pattern Recognition, ICPR. Pages: 2907-2921.
- [128] Shen, S., Deng, H., and Liu, Y., Model based human motion tracking using probability evolutionary algorithm. Lecture notes in computer science, ISSN 0302-9743, 2008.

- [129] Caillette, F., Galata, A., and Howard, T., Real-time 3-D human body tracking using learnt models of behaviour. In British Machine Vision Conf, 2005.
- [130] McKenna, S.J., Nait-Charif, H., “Tracking human motion using auxiliary particle filters and iterated likelihood weighting”, Image and Vision Computing 25, pp. 852–862, 2007.
- [131] Erdem, S., Human Motion Analysis Via Axis Based Representations. Phd Thesis, The Graduate School Of Natural And Applied Sciences Of Middle East Technical University, 2007.
- [132] David, M., A real-world system for human motion detection and tracking. California Institute of Technology, 2003.
- [133] M. Vondrak, L. Sigal, and O. Jenkins. Physical simulation for probabilistic motion tracking. To appear in Computer Vision and Pattern Recognition, CVPR, page in press, Anchorage, AK, USA, Jun 2008.
- [134] Yuanqiang Dong; DeSouza, G.N.; 2009, “A new hierarchical particle filtering for markerless human motion capture”, Computational Intelligence for Visual Intelligence, IEEE Workshop on, Page(s):14 – 21.
- [135] Wei-Lwun Lu, Kenji Okuma, James J. Little, 2009, “Tracking and recognizing actions of multiple hockey players using the boosted particle filter”, Image and Vision Computing, Volume 27, Issues 1-2, Pages 189-205.
- [136] Rius, I., Varona, J., González, J., Roca, X., 2009. “Action-specific motion prior for efficient Bayesian 3D human body tracking”, Pattern Recognition, Volume 42, Issue 11, Pages 2907-2921.
- [137] Rafael Muñoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas, A. Carmona-Poyato, 2009. “People detection and tracking with multiple stereo cameras using particle filters”, Journal of Visual Communication and Image Representation, Volume 20, Issue 5, July 2009, Pages 339-350.
- [138] Yonggang Zhang; Ning Li; Yuxin Zhao; 2009, “3-D Human Tracking Based on Particle Filter”, Computational Sciences and Optimization, International Joint Conference on, Volume 2, 24-26 April 2009 Page(s):273 – 277.
- [139] Zhaowen Wang, Xiaokang Yang, Yi Xu, Songyu Yu, 2009, “CamShift guided particle filter for visual tracking”, Pattern Recognition Letters, Volume 30, Issue 4, Pages 407-413.

- [140] Yuanqiang Dong, Guilherme N. DeSouza, 2009. “A New Hierarchical Particle Filtering for Markerless Human Motion Capture”, *Computational Intelligence for Visual Intelligence*, Page(s): 14-21.
- [141] Tobias J., Esther K., and Luc V., *Monocular Tracking with a Mixture of View-Dependent Learned Models*. Springer- Verlag Berlin Heidelberg, AMDO, pp. 494–503, 2006.
- [142] Tobias J., Esther K., and Luc V., *Multi-activity Tracking in LLE Body Pose Space*. Springer- Verlag Berlin Heidelberg, *Human Motion 2007*, LNCS 4814, pp. 42–57, 2007.
- [143] Smith, K., *Bayesian Methods For Visual Multi-Object Tracking With Applications To Human Activity Recognition*. Phd Thesis, The University Of Illinois, Urbana-Champaign, USA, 2007.
- [144] Zhao, X., Liu, Y., *Generative tracking of 3D human motion by hierarchical annealed genetic algorithm*. *Pattern Recognition* 41, pp. 2470 – 2483, 2008.
- [145] Zou, B., Chen, S., Shi, C., Marie, U., *Automatic reconstruction of 3D human motion pose from uncalibrated monocular video sequences based on markerless human motion tracking*. *Pattern Recognition* 42, pp.1559 – 1571, 2009.
- [146] Kale, N.A., *Human Posture Recognition From Detected Body Parts*. Master Of Science Thesis, Eskisehir Osmangazi University, 2007.
- [147] Nakata, T., *Recognizing Human Activities in Video By Multi-Resolutional Optical Flows*. *Intelligent Robots and Systems*. IEEE, 2006.
- [148] Olivier Bernier, Pascal Cheung-Mon-Chan, Arnaud Bouguet , 2009, “Fast nonparametric belief propagation for real-time stereo articulated body tracking”, *Computer Vision and Image Understanding*, Volume 113, Issue 1, January 2009, Pages 29-47.
- [149] David, A. F., Okan A., Leslie I., James B. and Deva R., *Computational Studies of Human Motion:Part 1, Tracking and Motion Synthesis*. *Computer Graphics and Vision*, Vol. 1, No 2/3, pp. 77–254, 2006.
- [150] Jenkins, O. C., Gonz’alez, G., Loper, M., 2006. “Dynamical Motion Vocabularies for Kinematic Tracking and Activity Recognition”, *Computer Vision and Pattern Recognition Workshop (IEEE)*, pages: 147:152.

- [151] Sminchisescu, C. and Triggs, B., 2003. "Kinematic Jump Processes For Monocular 3D Human Tracking", IEEE International Conference on Computer Vision and Pattern Recognition, pages: 69-76.
- [152] Shen, S., Deng, H., and Liu, Y., 2008. "Probability Evolutionary Algorithm Based Human Motion Tracking Using Voxel Data", The 2008 IEEE World Congress on Evolutionary Computation (CEC).
- [153] Ni, B., Kassim, A.A., and Winkler, S., "A Hybrid Framework for 3-D Human Motion Tracking". Transactions On Circuits And Systems For Video Technology, Vol. 18, No. 8, IEEE, 2008.
- [154] Séverine Dubuisson, Jonathan Fabrizio, 2009, "Optimal recursive clustering of likelihood functions for multiple object tracking", Pattern Recognition Letters, Volume 30, Issue 6, 15 April 2009, Pages 606-614.
- [155] Jian Yao, Jean-Marc Odobez , 2009, "Multi-Person Bayesian Tracking with Multiple Cameras", Multi-Camera Networks, 2009, Pages 363-388.
- [156] Ikizler, N., Duygulu, P., 2009. "Histogram of oriented rectangles: A new pose descriptor for human action recognition", Image and Vision Computing. Pages: 1515–1526.
- [157] Singh, V.K. and Nevatia, R., 2009. "Monocular human pose tracking using multi frame part Dynamics", Motion and Video Computing, page(s): 1-8.
- [158] Y. Weiss, 2000. "Correctness of local probability propagation in graphical models with loops", Neural Computation 12 (1), pages:1–41.
- [159] J.S. Yedidia, W.T. Freeman, Y. Weiss, 2005. "Constructing free-energy approximations and generalized belief propagation algorithms, IEEE Transactions on Information Theory 51 (7) pages: 2282–2312.
- [160] Lin, H., Chen, T., Chen, C., Hsieh, C. and Lie, W., 2009. "Human Pose Estimation From Monocular Image Captures", IEEE international conference on Multimedia and Expo, Pages: 994-997.
- [161] Kale, N.A, 2007. "İnsan Vücut Parçalarından Pozun Tanınması". Yüksek Lisans Tezi, Eskisehir Osmangazi Üniversitesi.
- [162] Sedai, S., Bennamoun, M., and Huynh, D., 2009. "Context-based Appearance Descriptor for 3D Human Pose estimation from Monocular Images", Digital Image Computing: Techniques and Applications, 978-0-7695-3866-2.

- [163] Ouyang, Y., Zhang, S., 2009. “Human Pose Tracking algorithm based on Skeleton-Texture model”, ETP International Conference on Future Computer and Communication. 978-0-7695-3676-7.
- [164] <http://mocap.cs.cmu.edu/>. 2010.
- [165] http://en.wikipedia.org/wiki/3D_projection#Perspective_projection. 2010.
- [166] [http://en.wikipedia.org/wiki/Closing_\(morphology\)](http://en.wikipedia.org/wiki/Closing_(morphology)). 2010.
- [167] [http://en.wikipedia.org/wiki/Opening_\(morphology\)](http://en.wikipedia.org/wiki/Opening_(morphology)). 2010.
- [168] http://www.robots.ox.ac.uk/~wmayol/3D/nancy_matlab.html. 2010.
- [169] Canny, J., A Computational Approach To Edge Detection, IEEE Trans. Pattern Analysis and Machine Intelligence, 8:679–714, 1986.
- [170] Cox, I. and Hingorani, S., 1996. “An efficient implementation of reid’s multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking”, IEEE Trans. Patt. Analy. Mach. Intell. 18, 2, 138–150.
- [171] G. Zhao, L. Chen, G. Chen, 2005. “A simple 3D face tracking method based on depth information”, Machine Learning and Cybernetics, page(s): 5022 - 5027 Vol. 8.
- [172] Kim, M., Kumar, S., Pavlovic, V., Rowley, H.; 2008. “Face tracking and recognition with visual constraints in real-world videos”, Computer Vision and Pattern Recognition, ISSN: 1063-6919, On page(s): 1 – 8.

ÖZGEÇMİŞ

Muhammed Fatih TALU

Fırat Üniversitesi Enformatik Bölümü 23119, Elazığ

Tel: +90-424-237 00 00 # 3161

E.posta: fatihthalu@firat.edu.tr, fatihthalu@gmail.com

- 1979** Elazığ-Palu’da doğdu.
- 1993 – 1996** Elazığ Mehmet Akif Ersoy lisesinden mezun oldu.
- 1997 – 2001** Fırat Üniversitesi Mühendislik Fakültesi, Bilgisayar Mühendisliği bölümünden mezun oldu.
- 2001 – 2003** Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği Anabilim Dalında “*Model Tabanlı Yapay Sinir Ağları Kullanılarak Görüntü Kenarlarının Yakalanması*” konulu tez çalışması ile Yüksek Lisans derecesini aldı.
- 2001 – ...** Fırat Üniversitesi Rektörlüğü, Enformatik Bölümü’nde öğretim görevlisi olarak görev yapmaktadır. Aynı zamanda, Bilgi İşlem Daire Başkanlığı bünyesindeki ağ altyapısı bilişim hizmetleri ile Cisco Ağ Akademisi eğitmenliği görevlerini yürütmektedir.
- 2005 – ...** Fırat Üniversitesi, Fen Bilimleri Enstitüsü, Elektrik – Elektronik Mühendisliği Anabilim Dalında doktora tez çalışması devam etmektedir.
- 2008 – 2009** Tübitak’ın 2224 nolu desteğinden yararlanarak doktora tez çalışmasının bir kısmını Amerikada bulunan Oregon Bilim ve Sağlık Üniversitesinde (Oregon Health and Science University) gerçekleştirdi.