

T.C.
GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ
MÜHENDİSLİK VE FEN BİLİMLERİ
ENSTİTÜSÜ

GÖZ-BAKIŞI TABANLI 3 BOYUTLU
BİLGİSAYARLA GÖRME
UYGULAMALARI

Ulaş VURAL
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI

GEBZE

2012

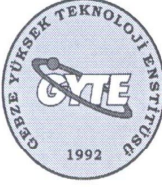
T. C.
GEBZE YÜKSEK TEKNOLOJİ ENSTİTÜSÜ
MÜHENDİSLİK VE FEN BİLİMLERİ
ENSTİTÜSÜ

GÖZ-BAKIŞI TABANLI 3 BOYUTLU
BİLGİSAYARLA GÖRME
UYGULAMALARI

Ulaş VURAL
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ
ANABİLİM DALI

TEZ DANIŞMANI
Doç. Dr. Yusuf Sinan AKGÜL

GEBZE
2012



DOKTORA TEZİ JÜRİ ONAY SAYFASI

G.Y.T.E. Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu'nun **13/07/2012** tarih ve **2012/25** sayılı kararıyla oluşturulan jüri tarafından **03/09/2012** tarihinde tez savunma sınavı yapılan **Ulaş VURAL**'ın tez çalışması **Bilgisayar Mühendisliği** Anabilim Dalında **DOKTORA** tezi olarak kabul edilmiştir.

JÜRİ

ÜYE

(TEZ DANIŞMANI) : Doç. Dr. Yusuf Sinan AKGÜL

ÜYE

: Doç. Dr. Abdullah BAL

ÜYE

: Yrd. Doç. Dr. Mehmet GÖKTÜRK

ÜYE

: Doç. Dr. Fatih Erdoğan SEVİLGİN

ÜYE

: Yrd. Doç. Dr. Tarkan AYDIN

ONAY

G.Y.T.E. Mühendislik ve Fen Bilimleri Enstitüsü Yönetim Kurulu'nun
...../...../20... tarih ve/..... sayılı kararı.

İMZA/MÜHÜR

ÖZET

TEZ BAŞLIĞI : GÖZ-BAKIŞI TABANLI 3 BOYUTLU BİLGİSAYARLA GÖRME UYGULAMALARI

YAZAR : ULAŞ VURAL

Günlük hayattaki problemlerin bilgisayarlar ve tamamen otomatik sistemler tarafından çözülmesi arzulanır. Tamamen otomatik sistemler insan hatalarından bağımsızdır ve bu sistemler oldukça verimli çalışırlar. Bununla birlikte mevcut teknolojik seviye insanların rahatlıkla çözebildikleri birçok problem için henüz yeterli değildir. Bu nedenle, uygulanabilir çözümlerin büyük bir bölümü önemli karar alma mekanizmalarında insanları kullanmaktadır.

Günümüz teknolojileri genellikle daha yüksek otomasyon seviyelerini amaçlarlar fakat insanın döngüdeki başarımını incelemezler. İnsan başarımını ölçen sistemler ise çoğu zaman benzetimlerle sınırlıdır. Bu tez çalışmasında, insan operatörlerin güvenlik gözetleme sistemlerinin vazgeçilemez bir parçası olduğu ve bu sistemlerinin döngüde-insan sistemi olarak düşünülmesi gerektiği gösterilmiştir.

Önerilen ilk çalışma operatörlerin nesnelere bakıp bakmadığını analiz etmektedir. Yöntem gözden kaçırılan kısımların doğrusal olmayan bir özetini operatörlere tekrar göstermektedir [Vural and Akgul, 2009]. Bu sayede operatörlere ikinci bir şans verilmiştir. Bu çalışma, yüksek çözünürlüklü güvenlik videolarında çalışacak şekilde genişletilmiştir [Vural and Akgul, 2011a]. Yapılan bir diğer çalışma videoları sentezlerken sadece göz-bakış konumlarını kullanmak yerine operatörlerin dikkat seviyelerini de hesaba katmaktadır. İnsanların dikkat seviyeleri sınırlıdır ve nesnel ve öznel değişkenlere bağımlıdır. Bu nedenle üretilecek videoların operatörün durumuyla uyumlu olması önemlidir. Deneyler, operatörlerin dikkat seviyelerine uygun üretilmiş videoların sistemin gürbüzlüğünü arttığını göstermiştir. Tezde duyurulan dördüncü çalışma hangi nesne özelliklerinin operatör için ilginç olduğunu analiz eder [Vural and Akgul, 2012b]. Yöntem ilginç nesneleri videoda daha erken konumlara çeker. Ayrıca, geliştirilen yöntemler x-ışınli bagaj görüntüleme sistemleri de denenmiştir.

SUMMARY

THESIS TITLE : EYE-GAZE BASED 3D COMPUTER VISION APPLICATIONS

AUTHOR : ULAS VURAL

It is desired that daily life problems are completely solved by computers and fully automated systems. Fully automated systems are free from the human errors and they can work efficiently, however the current technological level is not sufficient for solving many problems which can be easily handled by humans. Therefore, most of the feasible solutions use humans on decision critical parts.

Today's advanced technologies aim to increase the automation levels but they do not generally analyze the performances of operators. The systems that measure the human performances are generally limited to simulation environments. In this dissertation, we show that the human operators are indispensable parts of visual surveillance tasks and also the video surveillance system should be thought as an instance of a human-in-the-loop system.

The first proposed method analyzes whether the operator is looking at an object or not. The method then shows a non-linear synopsis video of overlooked parts again to the operator [Vural and Akgul, 2009]. By doing so, a second chance is given to the operator. This method is then extended to work with high-definition surveillance videos at real-time rates [Vural and Akgul, 2011a]. Another proposed method synthesizes surveillance videos by not only using the eye-gaze position but also taking account of operators' attention levels. Attention levels of human operators are limited and depend on subjective and objective variables. Thus, it is important that a video should fit in with the operator's state. The experiments show that the robustness of the surveillance systems are increased by showing synthetically adjusted videos [Vural and Akgul, 2011b]. The forth method proposed in this dissertation analyzes the object properties which are interesting for the human surveillance operator. The method places the interesting objects to the earlier times positions [Vural and Akgul, 2012b]. We also use eye-gaze based metrics on x-ray based baggage inspection.

TEŞEKKÜR

Lisansüstü çalışmaların her aşamasında desteğini gördüğüm, danışman hocam Doç. Dr. Yusuf Sinan AKGÜL'e, yönlendirmeleri, tavsiyeleri ve en önemlisi bana karşı olan sabrı ve güveni için,

Tez izleme komitemde bulunan Doç. Dr. Abdullah BAL'a ve Yrd. Doç. Dr. Mehmet GÖKTÜRK'e tezimin ilerleme sürecindeki önerileri ve katkıları için,

Doç. Dr. Fatih Erdoğan SEVİLGİN'e beni bilgisayarla görme alanına yönlendirmesi, yüksek lisans ve doktora eğitimimdeki büyük katkıları ve tez savunma jürimde yer aldığı için,

Yrd. Doç. Dr. Tarkan AYDIN'a tez jürimde yer aldığı için,

Tüm çalışma arkadaşlarıma dostlukları ve huzurlu bir çalışma ortamı sağladıkları için,

Eşim Özlem Vural'a, oğluma ve tüm aileme manevi destekleri ve sabırları için,

ve TÜBİTAK'a doktora eğitim boyunca yurtiçi doktora burs programıyla beni desteklediği için,

çok teşekkür ederim.

Bu tez çalışması Bilim, Sanayi ve Teknoloji Bakanlığı tarafından 00754-STZ-2010-2 kodlu "X Işınlı Paket Görüntüleme Cihazı Tasarımı" projesi altında desteklenmiştir.

Bu tez çalışması TÜBİTAK tarafından 110E033 numaralı "Operatör Başarım Ölçütü Ve Video Özet Yoğunluğu Kullanan Gözetleme Ve Güvenlik Sistemi" projesi altında desteklenmiştir.

İÇİNDEKİLER DİZİNİ

	<u>Sayfa</u>
ÖZET	İV
SUMMARY	V
TEŞEKKÜR	VI
İÇİNDEKİLER DİZİNİ	VII
SİMGELER VE KISALTMALAR DİZİNİ	İX
ŞEKİLLER DİZİNİ	X
ÇİZELGELER DİZİNİ	XI
1. GİRİŞ	1
1.1. Döngüde-İnsan ve Bilgisayarla Görme	1
1.1. Göz-Bakışı Takip Sistemleri	4
1.2. İçerik	6
2. GÖZ-BAKIŞI TABANLI GERÇEK ZAMANLI VİDEO ÖZETLEME	9
2.1. Doğrusal Olmayan Video Özetleme Yöntemi	13
2.2. Yöntem	16
2.2.1. Sıklık Tabanlı Arka Alan Çıkartma Yöntemi	17
2.2.2. Operatörlerinin Göz-Bakışı Pozisyonlarının Takip Edilmesi	19
2.3. Deneyler	22
2.4. Yorumlar	26
3. YÜKSEK ÇÖZÜNÜRLÜKLÜ VİDEOLARDA PARALEL ÖZETLEME	27
3.1. Yöntem	27
3.2. Deneyler	31
3.3. Yorumlar	36
4. OPERATÖR DİKKAT SEVİYESİ TABANLI VİDEO GÖZETLEME	37
4.1. Çoklu Nesne Takibinde İnsan Psikolojisi	41
4.2. Video Yoğunluğu Ayarlama Sistemi	43
4.2.1. Problem Tanımı	43
4.2.2. Enerji Fonksiyonu	45
4.2.3. Eniyileme	47

	<u>Sayfa</u>
4.3. Deneyler	48
4.4. Yorumlar	52
5. OPERATÖR İLGİSİ TABANLI GÜVENLİK VIDEO SENTEZİ	53
5.1. Yöntem	55
5.2. Deneyler	57
5.3. Yorumlar	60
6. X-IŞINLI BAGAJ İZLEME SİSTEMLERİ İÇİN GÖZ-BAKIŞI TAKİBİ	63
6.1 X-Işınlı Paket Görüntüleme Teknolojileri	64
6.2 Yöntem	66
6.3 Deneyler	68
6.4 Yorumlar	69
7. SONUÇLAR	71
KAYNAKLAR	74
ÖZGEÇMİŞ	89
EKLER	

SİMGELER VE KISALTMALAR DİZİNİ

3B	: Üç Boyutlu
AÇ	: Açık Çember
AK	: Açık Kare
BT	: Bilgisayarlı Tomografi
D	: Desensiz
DP	: Dinamik Programlama
İBE	: İnsan Bilgisayar Etkileşimi
KK	: Koyu Kare
KÇ	: Koyu Çember
KYM	: Kırmızı Yeşil Mavi
MİB	: Merkezi İşlem Birimi
ms	: milisaniye
PETS	: Performance Evaluation of Tracking and Surveillance
YİÖ	: Yerel İkili Örüntü

ŞEKİLLER DİZİNİ

<u>Sekil</u>	<u>Sayfa</u>
2.1. İlk deney videosu ve bu videodan elde edilen özet videolar.	12
2.2. Doğrusal-olmayan video özetleme yönteminin adımları.	14
2.3. Gerçek-zamanlı video özetleme için ikili tampon yaklaşımı.	17
2.4. Göz-bakışı tabanlı özetleme yönteminin temel şeması.	19
2.5. İlk video için enerji matrisleri ve minimum enerjili yollar.	21
2.6. Farklı hareket türleri için göz hareketlerinin incelenmesi.	23
2.7. İkinci deney videosu ve ilgili özetlerinden örnek çerçeveler.	25
3.1. Girdi video uzayı ve enerji matrisi üzerinde paralelleştirme.	29
3.2. DP alternatifi yığın tabanlı asgari enerjili yol bulma yöntemi.	30
3.3. İlk deney videosu için enerji matrisleri ve minimum enerjili yollar.	32
3.4. İkinci deney videosu ve ilgili özetlerinden örnek çerçeveler.	35
4.1. Video yoğunluk ayarlama sistemi.	39
4.2. Nesne sayısı, iş yükü ve dikkat seviyesi ilişkisi	42
4.3. Doğru dilimi modeli ve Douglas-Peucker algoritmasının ilk üç adımı.	44
4.4. Sentetik test arayüzünden örnek çerçeveler.	49
4.5. Önerilen sistemin PETS veri kümesindeki görsel sonuçları.	51
5.1. Operatör ilgisi tabanlı video sentezleme sisteminin genel yapısı.	55
5.2. Sentetik deneyde kullanılan nesne türleri ve YİÖ resimleri	59
5.3. Sentetik testlerdeki nesne sayılarının zamanla değişim grafikleri.	60
5.4. Farklı ilginç nesne tipleri için sentezlenmiş videolardan örnek çerçeveler.	61
6.1. Malzemelerin X-ışını kullanılarak sınıflandırılması, [Akay, 2012].	67
6.2. Örnek X-ışını resimleri ve sınıflandırma sonuçları, [Akay, 2012].	67
6.3. Tehdit unsuru içeren nesnelerden örnek görüntüler	69

ÇİZELGELER DİZİNİ

<u>Çizelge</u>	<u>Sayfa</u>
3.1. Yöntemlerin birinci test videosunda karşılaştırılması.	33
3.2. İlk video üzerinde yapılan iyileştirmelerin etkileri.	34
3.3. İkinci test videosu üzerinde yapılan iyileştirmelerin etkileri.	34
4.1. Psikolojik deneylerin sayısal sonuçları	50

1. GİRİŞ

1.1. Döngüde-İnsan ve Bilgisayarla Görme

Döngüde-insan sistemler bilgisayar ya da makine sistemleri ile insanların etkileşim içinde olduğu sistemlerdir [Rothrock and Narayanan, 2011]. Bu tür sistemler günümüzde oldukça yaygın şekilde kullanılmaktadır. Fabrikalardaki otomasyon sistemleri, güvenlik sistemleri, hava trafik yönetimi sistemleri gibi sistemler döngüde-insan sistemlerdir. Özellikle yukarıda sayılan alanlarda insan operatörlerin tamamen devre dışı bırakılamaması bu sistemlerde insanların verimli ve etkin kullanımını zorunlu kılmaktadır [Keval and Sasse, 2006].

Otomasyon sistemlerinin erken dönemlerinde yapılan çalışmalar genel olarak insanı tamamen dışlayacak çözümleri amaçlamışlardır [Lundmark, 2012]. Bu çalışmalar sonucunda birçok alanda bilgisayar sistemlerinin son kararı verecek yetkinliğe gelemeyeceği, yeterince hızlı ve doğru kararlar veremeyeceği görülmüştür. Bununla birlikte otomatikleşmiş kısımlarda oluşan hatalar oldukça azdır ve insan operatörler otomasyon döngüsünün zayıf halkalarıdır [Cranor, 2008].

Günümüzde otomasyon sistemleri, problemlerin tamamen otomatik çözümü bulmak yerine insan operatörleri verimli ve etkili şekilde kullanmayı amaçlamaktadır [Pretlove and Skourup, 2007]. Bu sistemler operatörler tarafından rahat anlaşılabilir ara sonuçları operatörlere sunarlar ve operatörün verdiği tepkiye bağlı olarak çalışmaya devam ederler. Daha gelişmiş sistemler ise operatörün durum değişikliklerini de izlerler, çünkü insanların duyu yeteneklerinin sınırlı ve öznel olması [Allen et al., 2004; Alvarez and Franconeri, 2007], iş yükü algılarındaki değişiklikler [Thiffault and Bergeron, 2003], sosyal ve duygusal durum değişiklikleri ve ortamdaki dış etmenler [Oron-Gilad et al., 2002] operatörlerin kararlarının doğruluklarını ve operatörün çalışma hızını etkileyebilen faktörlerdir [Hua and Yingzi, 2011]. Bu faktörler genel olarak öznel ve temel olarak modellenmesi bile oldukça zordur. Bu nedenle operatörlerin belli donanımlar yardımıyla gözetlenmesi ve onların öznel durumlarıyla ilgili bilgilerin toplanması ve işlenmesi gerekmektedir. Toplanan

operatör verilerine göre kullanılan arayüzler iyileştirilebilmekte [Komogortsev et al., 2011], yeni işe başlayan bir operatörün sıklıkla yapabileceği hatalar görülebilmekte [Law et al., 2004] ve operatörlerin yorgun olup olmadığı anlaşılabilmektedir [Fletcher and Zelinsky, 2009].

Güvenlik gözetleme sistemleri genel olarak insan operatörlerin vazgeçilemediği alanlardan birisidir. Bu sistemlerin tamamen otomatikleştirilmesi günümüzde mümkün değildir [Keval and Sasse, 2006]. Bilgisayarla görme tabanlı şüpheli hareket algılama sistemleri çok kısıtlı sayıda hareket için çalışabilmekte ve bu kısıtlı durumda bile ortalama bir insan operatörden çok daha başarısız sonuçlar vermektedir. Günümüzde yapılan çalışmaların birçoğu basit görüntü değişikliklerinde operatöre sesli ve görsel uyarılar veren arayüzlerin tasarımıyla sınırlıdır. Güvenlik gözetleme işi oldukça dikkat gerektiren bir iştir ve bu alanda çalışan operatörlerin çok kısa bir çalışma süresi sonunda dikkat seviyelerin düştüğü bilinmektedir [Allen et al., 2004]. Günümüzde kameraların ve kayıt ortamlarının fiyatlarının da düştüğü göz önüne alınırsa işgücü maliyeti en önemli işletim maliyeti olmaktadır [Dick and Brooks, 2003]. İnsan operatörlerin verimli ve etkili şekilde kullanılması için operatörlerin güvenlik gözetleme sistemi döngüsüne dâhil edilmesi gerekmektedir. Bu sayede operatörlerin öznel durumlarına uygun güvenlik videoları üretilebilecek ve sistemlerin daha gürbüz olması sağlanacaktır.

Güvenlik gözetleme görevi bir çeşit çoklu-nesne takibi problemidir [Takala and Pietikainen, 2007]. Bu problem birçok farklı açıdan ele alınmış ve altındaki psikolojik etmenler ve kısıtlar araştırılmıştır. İnsanların takip edebilecekleri nesnelerin hızlarının [Alvarez and Franconeri, 2007] ve sayısının sınırlı olduğu [Allen et al., 2004; Cavanagh and Alvarez, 2005; Franconeri et al., 2007], çoklu-nesne takibi yapan kişilerin bu nesnelerin belli bir civarına odaklandıkları [Fehd and Seiffert, 2008; Fehd and Seiffert, 2010], insanların ilgi duydukları nesnelere daha uzun süre sıklımadan [Trick et al., 2009] bakabildikleri saptanmıştır. Bu bilgiler psikoloji alanında bilinmesine rağmen güvenlik gözetleme sistemlerinde kullanılmamıştır.

Güvenlik operatörlerinin göz-bakışı konum bilgilerinin kullanılarak operatörlerin baktığı ve bakmadığı yerlerin analizinin yapılabileceği bu tez çalışması içinde gösterilmiştir [Vural and Akgul, 2009] ve çalışma yüksek çözünürlüklü güvenlik videolarıyla gerçek-zamanda çalışabilecek şekilde iletilemiştir [Vural and Akgul, 2011a]. Operatörlerin sadece bakma bilgisinin elde edilmesi yeterli değildir bunun için operatörlerin baktığı konumları gerçekten görüp görmediğinin incelenmesi gerekir. Yapılan bir diğer çalışma ile operatörün dikkat seviyesi analiz edilmiştir ve operatöre dikkat seviyesine uygun yoğunlukta güvenlik videosu sentezlenmiştir [Vural and Akgul, 2011b]. Bir diğer yapılan çalışma ise operatörlerin ilgi duyduğu hareket ve nesne türlerinin özniteliklerinin öğrenilmesi ve istenilen öznitelikteki hareketlerin daha erken gösterildiği videolar üretilmesidir [Vural and Akgul, 2012b]. Bu çalışma özellikle videoların yeniden izlenmesini oldukça hızlandıracaktır. Güvenlik operatörlerinin göz hareketlerine göre dikkat ve tecrübe durumunun ölçüldüğü bir diğer uygulama ise x-ışını tabanlı bagaj izleme sistemleri için kullanılmıştır. Yapılan çalışma ile güvenlik operatörlerinin doğru ve dikkatli bir şekilde sistemi kullanıp kullanmadıkları analiz edilmiştir.

Yukarıda sayılan çalışmaların yapılabilmesi için gerekli olan operatör durum bilgisinin tamamı göz-bakışı analizi ile elde edilmiştir. Bilgisayarla görme yöntemleri göz bebeğinin bulunması ve çeşitli ölçütlerin elde edilmesi için insan-bilgisayar etkileşimi için daha önceden de kullanılmaktaydı. Bir diğer deyişle bilgisayarla görme yeni bir kullanıcı arabirimi için bir ön işleme yöntemi olarak kullanılmaktaydı. Yapılan bu tez çalışmasıyla insandan elde edilen ölçütler gelişmiş bir bilgisayarla görme sistemine aktarılmaktadır.

Bir sonraki alt bölümde göz-bakışı takip sistemleri hakkında detaylı bilgi verilecektir. Tez çalışmasının içeriği ve yapılan çalışmaların kısa açıklamaları ise bir sonraki alt bölümdedir.

1.1. Göz-Bakışı Takip Sistemleri

İnsanların görme sistemleri, gerçek zamanda birçok zor görevi başarabilen karmaşık biyolojik sistemlerdir. Görme ile ilgili görevleri yerine getirmek için beyinin yarısı, tüm algılayıcıların yüzde yetmişi ve dört milyar nöron kullanılır [Coltekin, 2005]. Görme sistemi tarafından elde edilen algıların yorumlanması ile insanlar çok önemli bilgilere ulaşmaktadır [Ballard and Ozcandarlı, 1988]. Çok sayıda karmaşık görevi yerine getiren bir sistemin yüksek verimde çalışması bilgisayarla görme çalışmaları için ilham kaynağı olmuştur. Birçok problemin çözümü için gözün yapısından esinlenen metotlar geliştirilmiştir.

Göz-bakışının izlenmesi ve bu bilginin bilgisayarla görme problemlerinde kullanılması son dönemde oldukça yaygınlaşmıştır [Mohamed et al., 2008]. Göz-bakışı üzerine incelemeleri dört dönem içinde gruplandırmak mümkündür [Duchowski, 2002]:

1. (1879 – 1920 Yılları Arası) : Bu dönemde göz-bakışının basit hareketleri ve kavramları bulunmuştur. Bunların içinde göz-bakışındaki sıçramalarının gecikmesi, sıçramaların bastırılması ve algı alanı kavramları vardır.
2. (1930 – 1958 Yılları Arası) : Araştırmalar daha çok uygulamaya dönmüştür. Uygulamalı psikoloji deneyleri yapılmıştır ve göz-bakışı hareketlerinin davranışsal temelleri araştırılmıştır [Findlay and Gilchrist, 2001].
3. (1970 – 1998 Yılları Arası) : Doğru ve ucuz göz-bakışı takip ve kayıt cihazları geliştirilmiştir.
4. (2000 – Günümüze) : Göz-bakışı bilgisi bilgisayarla görme ve grafik uygulamalarında kullanılmaya başlamıştır.

Göz-bakışı bilgisinin takip edilmesi için geliştirilen cihazlar birkaç farklı şekilde sınıflandırılabilir. Morimoto ve Mimica [Morimoto and Mimica, 2005] göz-bakışı takip cihazlarını insan vücuduna takılanlar [Matsuyama, 2007; Topal et al., 2008] ve kamera tabanlı sistemler olarak ayırmaktadır. Kamera tabanlı sistemler insanları daha az rahatsız ederken gün ışığı gibi etkenlerden daha fazla

etkilenmektedir. Bir diğer sınıflandırma şemasında ise sistemler öncelikle etkileşimli ve tanısal olarak gruplandırılmaktadır [Duchowski, 2002]. Tanısal takip cihazları sadece göz-bakışının konumlarını kaydetmekte kullanılır ve bu tür sistemler kullanıcıyla etkileşime girmezler. Bu tür sistemler daha çok hastalıklarının tanımlanabilmesi için kullanılmaktadır. İnsanların göz-bakışı hareketleri izlenerek Parkinson hastalığı teşhis edilebilmektedir [Hillaire et al., 2008]. Etkileşimli sistemler tanısal sistemlere göre daha geniş kullanım alanlarına sahiptir ve kendi içlerinde tekrar alt gruplara ayrılırlar. Bu sistemler “seçici sistemler” ve “göz-bakışına bağlı sistemler” olarak adlandırılır. Seçici sistemler daha çok insan-bilgisayar etkileşimi alanında birer girdi cihazı tanımlarlar [Nonoka, 2003; Talmi and Liu, 1999]. Bu tür sistemler gözle kontrol edilebilen arabirimler sunarlar. Göz-bakışına bağlı sistemler ise kullanıcı ile etkileşim içinde kullanıcıdan aldıkları bilgilere göre işlem yapar ve bu işlem sonuçlarını kullanıcılara geri döndürür. Göz-bakışının bulunduğu noktaların daha detaylı şekilde gösterilmesi bu tür sistemlere örnek olarak verilebilir [Hillaire et al., 2008; Murphy and Duchowski, 2001; Murphy and Duchowski, 2007].

İnsanların hangi uyarıcılara ilgi gösterdikleri çok sayıda psikolojik araştırmaya konu olmuştur. İnsanların ilgi alanlarının ve yoğunluklarının ölçülmesi için yapılan bu araştırmalar başta ticari kaygılarla yapılsa da daha sonra bilgisayar arabirimlerinin tasarlanması için kullanılmıştır. İlginin yoğunlaştığı konu veya alanın bilinmesi insanların o anki düşüncelerinin anlaşılması için önemlidir. İnsanların baktıkları yerler ile ilgileri arasında bir ilişki olduğu düşünülmektedir [Harris and Jenkin, 2001]. Bu nedenle göz-bakışı hareketlerinin incelenmesi ile insanların düşüncelerinin ve ilgi duydukları şeylerin tahmin edilmesi mümkün olabilmektedir. Bu bilginin ilk başta ticari bir ihtiyacı karşılayabileceği düşünülmüştür. Reklam amacıyla oluşturulacak videolarda izleyicilerin ilgilerini arttırmak ve ilgiyi belli yerlere yönlendirmek için göz-bakışı konum bilgilerini tahmin edecek sistemler geliştirilmiştir [Barth et al., 2003; Böhme et al, 2005; Marat et al., 2009]. Bu çalışmaların bir benzer şekli videolar yerine resimler üzerinde çalışmaktadır. Sistem göz-bakışını takip ederek resimleri gruplandırabilmektedir [Oyekoya and Stentiford, 2004]. Bunun dışında ilginin belli bir alanda toplanmasını sağlayacak arabirimlerin

insanları istenilen görevleri yapmaya gizlice yönlendirebileceği görülmüştür [Vertegaal, 2002]. Bu tür arabirimleri insan bilgisayar etkileşimi için kullanarak operatörlerin verimliliğini arttırmak mümkün olmaktadır. Bu alanda başlıca iki türlü çalışma yapılmıştır. İlk olarak insanların bilgisayarlarla daha doğal yollardan haberleşebilmesi için girdi cihazları tasarlanmıştır. Bu tür cihazlar göz-bakışı konum bilgisine bakarak imleci hareket ettirebilmekte ve sanal klavye ile metin girişine izin vermektedirler. Bu tür sistemler fiziksel engelli insanlar için faydalı olduğu gibi, özel uygulamalar için verimliliği de arttırmaktadır. Göz-bakışı bilgisinin bilgisayarla kullanıldığı ikinci alan ise göz-bakışı konumuna göre nesnelerin çözünürlüğünü, netliğini, vb. değiştirmektir. Bu tür uygulamalar ile insanların ilgileri belli bir amaca hizmet edecek şekilde yönlendirilmektedir [Hillaire et al., 2008; Murphy and Duchowski, 2001; Murphy and Duchowski, 2007]. Göz-bakışı konum bilgisi ilginin önemli olduğu diğer alanlarda da kullanılmıştır. İnsan dikkatinin önemli olduğu alanlardan biri de sürücülüktür. Sürücülerin dikkatsizlikleri, uykulu veya dalgın oluşları önemli kazalara neden olabilmektedir. Sürücülerin dikkatlerini ölçmek için göz-bakışının konum bilgisini kullanan bilgisayarla görme metotları geliştirilmiştir [D'Orazio et al., 2007; Ji and Yang, 2002].

1.2. İçerik

Yapılan tez çalışması, insan-bilgisayar etkileşimi ve döngüde insan sistemleri temel almaktadır ve genel olarak bu sistemlerdeki insan faktörünün verimli ve etkin bir şekilde kullanılması için yöntemler sunmaktadır. Bölüm 1.1'de döngüde-insan sistemlerin genel bir tanımı ve bu alandaki mevcut eksiklikler detaylı şekilde verilmiştir. Tez çalışması içinde insan faktörünün izlenmesi için göz-bakışı tabanlı ölçütlerin kullanılması tercih edilmiştir. Göz-bakış yönü takip teknolojileri ve göz-bakışı takibi sırasında karşılaşılan problemler Bölüm 1.2'de anlatılmıştır.

Bölüm 1'de verilen arka alan bilgilerinden sonra tez çalışması boyunca geliştirilen yöntemler açıklanmıştır. Bölüm 2'de öncelikle güvenlik gözetleme problemi bir döngüde insan problemi olarak tanımlanmıştır [Vural and Akgul, 2009]. Mevcut güvenlik gözetleme sistemlerinin birçoğunun aksine geliştirilen sistem

güvenlik operatörlerinin çevrimiçi geri beslemelerini almaktadır. Bir başka deyişle, insan operatörler sistem döngüsüne dahil edilmiştir. İnsan operatörlerin döngüye alınmasıyla operatörlerin öznel durumlarından kaynaklanan gözden kaçırma gibi problemlere çözüm üretilmesinin önü açılmıştır. Bölüm 2’de açıklanan sistem operatörlerin göz bakışı konumlarını alarak onların ekranda hangi noktalara baktıklarını saptamaktadır. Sistem bakılmayan bölgelerdeki hareketlerin doğrusal olmayan bir özetini (Bkz. Bölüm 2.1.) elde etmektedir. Gerçek-zamanlı olarak elde edilen bu özet görüntü operatörlere tekrar gösterilebilmekte ve onlara gözden kaçırdıkları hareketleri izleme şansı sunulmaktadır. Dinamik programlama tabanlı yöntemin detayları Bölüm 2.2’de verilmiştir. Geliştirilen sistem gerçek güvenlik videoları üzerinde denenmiş ve etkinlikleri Bölüm 2.4’de gösterilmiştir. Geliştirilen yöntem operatörlerin başarımının ölçülmesi için de kullanılabilir.

Bölüm 2’de anlatılan yöntem yüksek çözünürlüklü güvenlik videolarında gerçek zamanlı olarak çalışmamaktadır. Güvenlik gözetleme pazarında yüksek çözünürlüklü kameraların hızla yaygınlaşması bu problemin çözümünü zorunlu kılmaktadır. Bölüm 3’de yapılan çalışma ile Bölüm 2’de anlatılan sistemin yüksek derecede paralel hesaplanabilir olduğu gösterilmiştir [Vural and Akgul, 2011a]. Bu bölümde, doğrusal olmayan video özetinin elde edilmesi için dinamik programlama yerine yığın tabanlı bir sistemin kullanılabileceği gösterilmiştir. Bölüm 3.1’de özetleme yöntemi ve özet video uzayının paralel şekilde nasıl inşa edildiği detaylı şekilde açıklanmıştır. Gerçek güvenlik videoları üzerinde yapılan deneylerle sistemin yüksek çözünürlüklü videolarda gerçek zaman gereksinimlerini karşılayabildiği gösterilmiştir.

Bölüm 2’de ve Bölüm 3’de anlatılan sistemlerin iki temel problemi mevcuttur. Bu problemlerden ilki, doğrusal olmayan video özetleme yöntemlerinin genel problemidir. Bu tür özetleme algoritmaları çok sıkı özetler üretebilmektedir fakat güvenlik gözetleme operatörlerin bu denli yoğun videoları etkin şekilde takip edebilmesi mümkün değildir. Güvenlik gözetleme işi, psikoloji alanında çokça çalışılmış olan çoklu nesne takibi probleminin bir örneğidir. Nesne sayısının yüksek olması operatörlerin iş yüklerini arttırmakta ve onların erken yorulmasına neden olabilir. Bölüm 2’de ve Bölüm 3’de anlatılan sistemlerin bir diğer sıkıntısı ise

operatörlerin psikolojik durumlarını dikkate almamasıdır. Bölüm 4’de geliştirilen sistem ile ilk kez video yoğunluğu ayarlama kavramı tanımlanmıştır [Vural and Akgul, 2011b]. Bu bölümde anlatılan sistem, güvenlik operatörlerinin çevrimiçi algı seviyelerini ölçmekte ve onların takip yeteneklerine uygun yoğunlukta videolar sentezlemektedir.

Bölüm 5’de yapılan çalışma, daha önceki bölümlerde anlatılan ve operatörleri döngüye alan güvenlik gözetleme sistemlerinin bir örneği olmakla birlikte ilk kez operatörün ilgi duyduğu nesnelerin özelliklerinin öğrenilmesi sağlanılmıştır [Vural and Akgul, 2012b]. Geliştirilen sistem, bu yönüyle makine öğrenmesi algoritmalarının daha önceki bölümlerde anlatılan yöntemlerle bütünleştirilebileceği göstermiştir. Beşinci bölümde anlatılan yöntem kayıtlı güvenlik videolarının hızlı bir şekilde incelenmesi için operatörlere destek olacaktır. Tam otomatik yöntemler sınırlı sayıda sınıf için çalışmaktadır ve sınıflandırma hataları kabul edilebilir sınırların oldukça üzerindedir. Bu bölümde duyurulan yöntem ile operatörlerin ilgi duydukları nesnelerin daha erken izlenmesi sağlanmıştır. Operatörün ilgi değişimine göre kendini uyarlayabilmesi de sistemin önemli bir avantajıdır.

Güvenlik gözetleme görevlerine bir diğer örnek de X-ışınlı bagaj tarama sistemlerinden elde edilen görüntülerin incelenmesidir. Bu sistemler bir video görüntüsü yerine zamanla değişen iki boyutlu X-ışını resimlerini içermektedir. Bu değişikliğe rağmen operatörlerin algı seviyelerin belirlenmesi sistemlerin etkili şekilde çalışması için önemlidir. Bölüm 6’da anlatılan sistem ile operatörlerin algı seviyelerinin ölçümü sağlanmış ve operatörlerin sanal tehdit nesnelerini yakalama başarımları ölçülmüştür.

Bölüm 7’de yapılan çalışmaların sonuçları yorumlanmıştır ve gelecekte yapılabilecek çalışmalar hakkında öngörülerde bulunulmuştur.

2. GÖZ-BAKIŞI TABANLI GERÇEK ZAMANLI VIDEO ÖZETLEME

Alışveriş merkezleri, metro istasyonları ve sokakların artan güvenlik ihtiyaçlarının karşılanabilmesi için çok sayıda kamera kullanılmaktadır [Koskela, 2000]. Güvenlik odalarında çalışan operatörler bu kameralardan gelen görüntüleri büyük bir monitör matrisi üzerinden izlerler. Operatörlerin izlemeleri gereken olaylar çok fazla sayıda olduğunda ilgi sadece bu olayların çok küçük bir alt kümesinde yoğunlaşır [Preece et al., 1994]. Bu durumda bazı önemli olaylar kaçırılabilir. Bu sorunun en basit çözümü güvenlik görevlilerinin sayısını artırmaktır. Bu çözüm güvenlik sisteminin kararlılığını artırırken, sistemin verimliliği düşer. İş gücü maliyeti, toplam yazılım ve depolama ortamı maliyetlerinden çok daha fazla bir hale gelir ve toplam maliyeti artıran en önemli kalemi oluşturur [Dick and Brooks, 2003].

Güvenlik sistemleri pazarının ucuz ve güvenilir ürünlere olan ihtiyacı giderek artmaktadır. Sektörde iş gücünün daha verimli olarak kullanılabilmesi için çok sayıda araştırma yapılmaktadır [Ahmad et al., 2007]. Bu araştırmaların bir kısmı güvenlik operatörlerinin sistemle etkileşimini iyileştirmek üzerinedir. İnsan bilgisayar etkileşimi alanında yapılan bu çalışmalarla operatörlerin sistemleri daha rahat yönetebilmeleri sağlanmaktadır. Bir operatör, fare ve klavye kullanmak yerine el ve parmak hareketleriyle kameraları yönetebilmektir [Iannizzotto et al., 2005]. Bunun dışında geliştirilen uygulamalar bilinen sınırlı sayıdaki saldırı ya da şüpheli hareketi tanımlayabilmektedir. Sistem bilinen hareketleri algıladığında operatörü sesli ve görsel olarak uyararak operatörün dikkatini ilgili görüntüye yönlendirmektedir [Steiger et al., 2005]. Gelişmiş ve tamamen otomatik güvenlik sistemleri insan takibi [Siebel and Maybank, 2004], nesne ve eylem tanıma [L'opez et al., 2006] gibi karmaşık mekanizmalara ihtiyaç duymaktadır [Hu et al., 2004]. Bu tür sistemler sadece belirlenmiş yerler ve konumlar için çalışabilmekte ve yüzde yüz başarı sunamamaktadır. Tüm geliştirilen bu sistemler güvenilirliği arttırsa da operatörler yüksek iş yükünde önemli olayları gözden kaçırabilmektedir.

Güvenlik sistemlerin olay anında kullanılmasının dışında güvenlik kayıtlarının izlenmesi de olay sonrası delil ve suçluların bulunması için kullanılmaktadır.

Depolanan görüntülerin boyutları güvenlik kameralarının sayıları, kameraların aktif görüntü aldığı süre ve çerçeve oranlarına bağlı olarak artmaktadır. Bu kadar büyük bir video veri tabanında istenen bilginin aranması oldukça fazla zaman ve işgücü gerektirmektedir [Hampapur et al., 2007]. Bu sorunu çözmek için otomatik arama yöntemleri geliştirilmiştir. Video görüntülerinin indekslenmesi ve bu görüntülerin bir veri tabanı içerisinde toplanması ile aramaların hızlandırılması için bir yöntem olarak kullanılmıştır [Dick and Brooks, 2003]. Aramaların otomatikleştirilmesi için hem içerik tabanlı hem de anlam-tabanlı video arama algoritmaları geliştirilmiştir. İçerik tabanlı arama algoritmaları nesnelerin şekline, rengine veya dokusuna bakarken anlam tabanlı arama yöntemleri videodaki eylemlerin hangi olaylara benzediğini çözmeye çalışmaktadır [Hu et al., 2004]. Anlam tabanlı yöntemler daha gelişmiş olmakla birlikte düşük-seviye resim özniteliklerinin gerçek anlamlarını bulmak için karmaşık mekanizmalara ihtiyaç duymaktadırlar [Snoek and Worring, 2005].

Operatörlerin sistemleri daha verimli kullanması için yapılan çalışmalar ve video arama algoritmaları ile birlikte etkileşimli güvenlik sistemleri için her alana dokunulmuş gibi gözükse de, bu iki alanın ortasında hala önemli bir boşluk bulunmaktadır. Sistemin verimliliğini ve güvenilirliğini arttırmaya çalışan ilk gruptaki yöntemler başarılı şekilde çalışsa da operatörlerin gözden kaçırma riskini sıfırlamamaktadır. Bir operatör, önemli bir olayı gözden kaçırdığında geri dönüşü olmamaktadır. Güvenlik kayıtlarının sonradan incelenmesi ise ancak olay sonrasında mümkün olmaktadır. Bu durumda güvenlik sistemi insanların zarar görmesini engelleyememekte ve sadece suçlunun bulunmasına yardımcı olabilmektedir. Yapılan çalışma ile bu boşluğu kapatacak bir sistem geliştirilmiştir. Geliştirilen sistem sadece operatörlerin göz-bakışı konum bilgisini kullanmaktadır. Sistem ne içerik ve anlam bilgisine bakmakta ne de videoları etiketlemektedir, sistemin yaptığı iş operatörlerin gözden kaçırdığı yerlerin bir özetini gerçek zamanda çıkartarak operatörlere ikinci bir şans vermektir. Sistemin bir diğer avantajı da geçmiş olayları ararken güvenlik operatörlerinin kaçırdıkları ya da dikkatle izledikleri yerlerin özetlerinin çıkartılabilecek olmasıdır. Bu yöntem diğer video arama yöntemleriyle birleştirilebilir ve video üzerindeki aramalar çok daha kısa süreler içerisinde tamamlanabilir.

Geliştirilen yöntem operatörün göz-bakışı konum bilgisine bakarak operatörün ilgi alanını bulmaktadır. Göz-bakışı tabanlı ilgi alanları iki boyutlu resim analizi için sıkça kullanılmıştır. Göz-bakışı bilgisi resimler üzerinde anlamsal bir bilgi olarak daha önceden kullanılmıştır [Jaimes et al., 2001; Jing and Lansun, 2008]. Bu yönü ile diğer içerik tabanlı resim analiz algoritmalarıyla bütünleştirilmiştir. Resimler sadece nesneleri içerirken, videolar hem nesneleri hem de eylemleri içermektedir, bu nedenle videolarda semantik kurallar bulmak oldukça zordur. Bu zorluk nedeniyle göz-bakışı bilgisinin videolarda kullanılması kısıtlı kalmıştır. Yapılan çalışma, bu sorunun etrafından dolaşılabilceğini göstermiştir. Tez çalışmamızın bu bölümünde herhangi bir anlamsal kural bulmak için uğraşmak yerine sadece insanların videoları nasıl izlediği ve hareketli nesneleri nasıl takip ettiğini [Jacob, 1991] gözlemlenmiştir. Psikolojik çalışmalar insanların aynı anda beş ila sekiz hareketli nesneyi takip edebildiğini ve bu sayıdan daha fazlasını ise gözden kaçırdığını göstermektedir [Franconeri et al., 2007; Pylyshyn and Storm, 1988; Sears and Pylyshyn, 2000]. Bu da sistemin yoğun olduğu zamanlarda operatörlerin bazı olayları gözden kaçırabileceğini göstermektedir. Göz-bakışı çok fazla sayıda hareketli nesne için sıçrama yapmak yerine hareketli nesnelerin ortasında bir yere odaklanmaktadır [Fehd and Seiffert, 2008]. Bu noktanın yakın çevresindeki eylemler görülmüş, uzaktakileri ise gözden kaçırılmış olarak kabul edilmiştir.

Video özeti çıkartılması için oldukça fazla sayıda yöntem mevcuttur [Komlodi and Marchionini, 1998; Truong and Venkatesh, 2007]. En çok kabul gören yöntemler, az etkinlik olan çerçevelerin atılması üzerine kuruludur [Kim and Hwang, 2000; Li et al., 2000]. Bu basit yöntemler videoyu belli sayıdaki anahtar çerçeveden daha kısa olacak şekilde özetleyemezler. Bu yöntemlerin bir diğer sıkıntısı ise belli bir eşik değere ihtiyaç duymalarıdır. Eşik değerlerin belirlenmesindeki zorluklar burada da oldukça önemli sıkıntılar yaratmaktadır. Yüksek eşik değerleri özet videonun boyunu arttırırken, düşük değerler önemli eylemlerin kaçırılmasına neden olabilirler.

Tüm çerçeveyi atarak özet çıkartan yöntemlerin bir diğer sorunu ise atılan çerçevede hem izlenen hem de izlenmeyen görüntülerin olmasıdır. Bir çerçevedeki görüntü operatör tarafından kısmen izlenmiş olabilir, bu durumda yöntemler ya bu tür çerçevelerin tamamını özete dâhil edecek ya da tamamını dışlayacaktır. İlk

durumda özet videonun boyutu büyürken, ikinci durumda ise istenilen eylemler kaybedilecektir. Bu nedenlerden ötürü; çerçeve dışlama tabanlı özetleme yöntemleri göz-bakışı tabanlı özetleme yapmak için uygun değildirler. Göz-bakışı tabanlı özetleme için daha elverişli olacak bir yöntem doğrusal olmayan video özetleme yöntemidir [Acha et al., 2006]. Doğrusal olmayan video özetleme yöntemi üç boyutlu video uzayında enerji minimizasyonu problemi olarak ele alınabilir. Bu yöntem ile piksellerdeki kronolojik bilginin değişmesine izin verilmektedir. Yöntemin kullanılmasıyla elde edilen özet videoda farklı zamanlarda olmuş olaylar aynı çerçeve içerisine toplanabilmektedir. Örneğin iki insanın farklı zamanlarda teker teker sahnede yürüdüğü bir videonun (Bkz. Şekil 2.1(a)) özetinde insanlar peş peşe yürüyormuş gibi gözükabilmektedir (Bkz. Şekil 2.1(b)).



(a) 366 çerçeve uzunluğundaki videodan seçilmiş örnek çerçeveler gözükmemektedir. Küçük mavi daire güvenlik operatörünün baktığı yeri göstermektedir.



(b) Göz-bebeği bilgisi kullanılmadan çıkartılan video özetinin uzunluğu 185 çerçevedir.



(c) İzlenen kısımların özeti 147 çerçeve uzunluğundadır.



(d) Operatörün gözden kaçırdığı kısımların özeti 130 çerçeve uzunluğundadır.

Şekil 2.1. İlk deney videosu ve bu videodan elde edilen özet videolar.

Yapılan çalışmanın en özgün tarafı doğrusal olmayan video özetleme yöntemi [Yildiz et al., 2008] ile göz-bakışı bilgisinin bir arada kullanılmasıdır. Temel doğrusal olmayan video özetleme algoritmasında özetlenecek videodan kaç tane çerçevenin atılacağı belirtilmek zorundayken bu çalışma ile böyle bir sınır verilmesi gerekliliği de ortadan kaldırılmıştır. Göz-bakışı tabanlı doğrusal olmayan video özetleme yöntemi güvenlik sistemleri için verimliliği ve güvenilirliği arttırıcı etkileri olacaktır. Geliştirilen sistem gerçek-zamana yakın hızlarda çalışmaktadır. Bunun anlamı operatör tarafından izlenmiş (Bkz. Şekil 2.1(c)) ve/veya gözden kaçırılmış (Bkz. Şekil 2.1(d)) kısımların özetleri güvenlik kamerasındaki eylem bittiğinde yeniden izlemeye hazır olmasıdır.

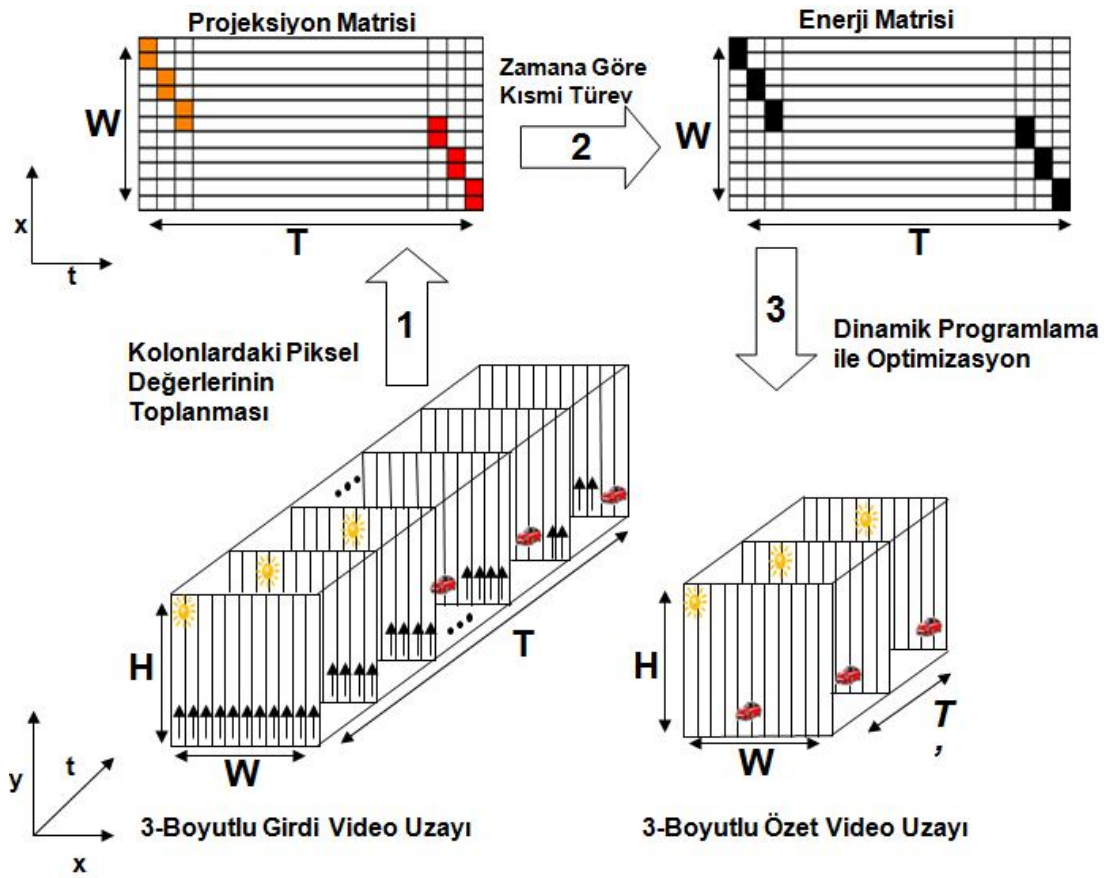
2.1. Doğrusal Olmayan Video Özetleme Yöntemi

Video özetleme yöntemleri, güvenlik sistemlerinin işletim maliyetlerini azaltırlar. Video içerisinde arama yapmak için gerekli olan insan gücünün azalmasının yanında depolama ortamlarına duyulan ihtiyaç da azalır. Geliştirilen sistemde video özetleme önceki sistemlere göre oldukça farklı şekilde kullanılmıştır. Göz-bakışı bilgisinin sistemle bütünleşmesiyle güvenlik operatörlerinin hangi eylemleri izlediği veya kaçırdığı bilinebilmektedir. Bu bilgi kullanılarak sadece izlenen ya da gözden kaçırılan eylemlerin özetlerinin çıkartılabilmesi mümkün olmuştur. Bu tür bir özet çıkartılması sadece doğrusal olmayan video özetleme yöntemi ile mümkün olabilmektedir.

Doğrusal olmayan video özetleme yöntemi oldukça verimli bir şekilde çalışabilmektedir. Yöntemin bu kadar hızlı çalışabilmesi temel bir gözleme dayanmaktadır. Yapılan gözlemler sonucunda x eksenleri yeryüzüne paralel şekilde yerleştirirken güvenlik kameralarındaki hareketlerin genelde yatay eksende olduğu gözlemlenmiştir. Bu gözlem ile problem üç-boyutlu yapısının iki-boyuta düşmesi sağlanmıştır. Boyutun düşürülmesi için üç-boyutlu video uzayı y eksenine üzerine yansıtılmıştır. Bu yansıtma sonrası problemin boyutu düşerken y eksenindeki hareket bilgisi kaybedilmiştir (Bkz. Şekil 2.2(Adım-1)). Yansıtma sonrası yatay hareket bilgisi iki boyutlu P yansıtım matrisinde tutulmaya başlanmıştır. $W \times H$ boyutlarındaki

çerçevelerden oluşan T uzunluğundaki bir video yansıtma sonrasında $W \times T$ uzunluğundaki P matrisinde tutulur. P matrisinin her bir elemanı V girdi videosundaki bir kolonu göstermektedir. Bu elemanların değerleri ilgili kolondaki piksel değerlerinin toplamı olarak bulunur. Toplama işlemi gri değerler üzerinden yapılır:

$$P(w, t) = \sum_{h=1}^H V(w, h, t), \forall w, t, c, s.t. w \in [1, W], t \in [1, T] \quad (2.1)$$



Şekil 2.2. Doğrusal-olmayan video özetleme yönteminin adımları.

Yansıtma sonrasında problemin boyutu düşmüş olsa da P matrisindeki değerler piksellerin renk değerlerinin toplamıdır ve bu şekilde kullanılamaz. Bu değerlerin optimizasyon içinde kullanılabilmesi için ek bir işleme gerek vardır. İkinci adımda P matrisi ile aynı boyutlarda bir E matrisi elde edilir (Bkz. Şekil 2.2(Adım-2)). Bu matrisin elemanları P matrisinin zamana göre kısmi türevlerinin alınmasıyla bulunur

(Bkz. Eşitlik 2.2). Hareket bilgisi, video kolonlarındaki renk değişimlerinden elde edilir.

$$E(w, t) = \left| \frac{\partial P(w, t)}{\partial t} \right|, \forall w, t, s.t. w \in [1, W], t \in [2, T] \quad (2.2)$$

Bu yöntem ile çerçeveler yerine kolonlar videodan çıkartılmaktadır. Hangi kolonun videodan çıkartılacağına karar vermek için dinamik programlama kullanılmaktadır. Dinamik programlama E enerji matrisi üzerinde çalışmaktadır. E matrisi üzerindeki yüksek enerji değerleri olası hareketleri, düşük enerji değerleri ise arka planı göstermektedir. E üzerinde dinamik programlamayla bulunan minimum enerjili yollar videodan çıkartılmaktadır. E matrisi üzerinde bulunan yolun her bir elemanı video uzayında bir kolona denk gelmektedir. Yol bulunurken kullanılan kısıtlar ile her bir x pozisyonundan sadece bir tek kolon seçilir. Bu yöntem ile elde edilen yol video uzayında farklı zamanlardan kolonlar içeren bir çerçeveye denk gelir. Bu çerçeve videodan atılarak videonun kısaltılması sağlanır. Video üzerinden bir çerçeve atıldığında E matrisindeki değerlerin küçük bir kısmı değişir. Değişen bölgeler yeniden hesaplandıktan sonra dinamik programlama ile yeni bir yol bulur. Sistem verilen yol sayısı kadar çalışır ve özet video elde edilir. Bu yöntem Avidan ve Shamir'in doğrusal olmayan resim boyutlandırma metodunun videolara uygulanması olarak düşünülebilir [Avidan and Shamir, 2007]. Resim boyutlandırma problemi için kullanılan enerji matrisi E_{resim} orijinal I resminin uzamsal türevi olarak hesaplanır. Bu yöntem ile kenarlar ve dokulu bölgeler korunacaktır.

$$E_{\text{resim}} = \sqrt{\left(\frac{\partial I}{\partial x} \right)^2 + \left(\frac{\partial I}{\partial y} \right)^2} \quad (2.3)$$

Resmin yatay olarak küçültülebilmesi için E_{resim} üzerinde dikey bir yolun bulunması gerekir. Bulunacak yol her bir satırdan bir piksel içermelidir. Bu kural her bir yol çıkartıldığında resmin her satırının aynı sayıda eleman içermesini zorunlu kılar. $W \times H$ boyutundaki bir resimde, dikey bir yol şu şekilde tanımlanabilir

$$S^d = \{(kolon(h), h)\}, s.t. \forall h, h \in [1, H], |kolon(h) - kolon(h-1)| \leq 1, \quad (2.4)$$

burada $kolon(h)$, h satırındaki kolonun konumunu vermektedir. Dikey yol S^d her bir satırdan bir tane olmak üzere h tane piksel içermektedir. İki komşu satırdaki elemanların kolon konumları arasındaki fark ise en fazla bir birim olabilir. Aynı şekilde yatay yol S^y aşağıdaki gibi tanımlanabilir:

$$S^y = \{(w, satir(w))\}, s.t. \forall w, w \in [1, W], |satir(w) - satir(w-1)| \leq 1. \quad (2.5)$$

Minimum enerjili yatay ve dikey yollar bulunarak resim yeni boyutlarına indirilir. Minimum enerjili yol dinamik programlama kullanılarak bulunur. Dinamik programlama ile her bir yolun toplam enerjisi öncelikle M tablosuna doldurulur ve minimum enerji değeri elde edilir. M tablosu tamamen doldurulduğunda minimum değerler tablonun en alt satırında yer alırlar. En alt satırdaki minimum değerden geri gelinerek yol elde edilir. M tablosundaki değerler aşağıdaki özyineleme kullanılarak hesaplanır.

$$M(w, h) = E_{resim}(w, h) + \min\{M(w-1, h-1), M(w, h-1), M(w+1, h-1)\} \quad (2.6)$$

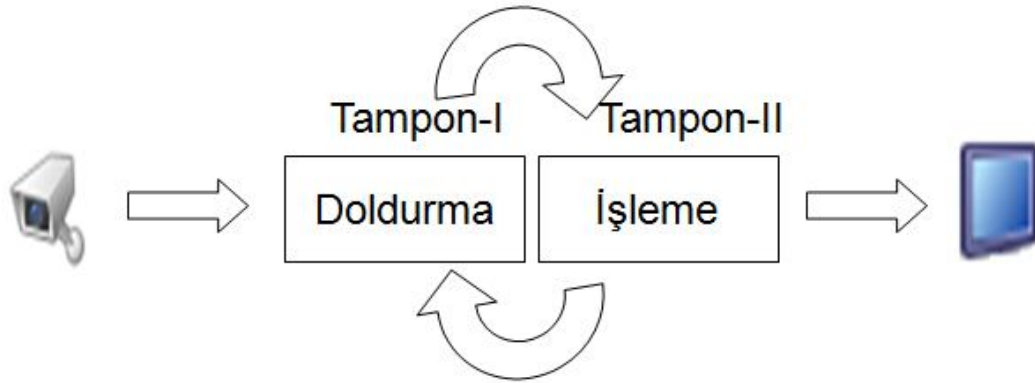
Bu yöntem iki boyutlu resimlere uygulanabildiği gibi üç boyutlu video uzaylarına da uygulanabilir. En kaba haliyle iki boyutlu minimum enerjili yolları aramak yerine üç boyutlu minimum enerjili yüzeyler aranabilir, bu durumda dinamik programlamanın üstel zaman gereksinimi olacaktır. Bunun yerine bölümün başında anlatıldığı şekilde bir yansıtma yapılması ve problemin iki boyutlu probleme indirilmesi mümkündür.

2.2. Yöntem

Geliştirilen yöntem Yıldız ve diğerlerinin yansıtma yöntemini kullanarak problemin boyutunu düşürmektedir [Yıldız et al., 2007]. Yeni yöntem daha sonra sıklık tabanlı arka alan analizi yaparak arka alanı çıkartmaktadır. Bu analiz ile sistem özet video boyutunu otomatik bir şekilde belirleyebilmektedir. B arka alan matrisindeki hareket bilgisi göz-bakışı konumuna göre filtrelenebilir. Bu

filtreleme sonucunda izlenen ya da izlenmeyen hareketler belirlenmektedir. Yöntem son aşamada bu veriler üzerinde dinamik programlamayı çalıştırmakta ve göz-bakışı konumuna bağlı doğrusal olmayan video özeti elde edilmektedir.

Sistemin gerçek-zamanlı çalışması için iki adet tampon bellek kullanılmıştır. Her bir tampon ayrı kanal programcıkları tarafından işlenmektedir. Bu programcıklardan biri kendi tamponunu doldurmakta ve her çerçeve geldiğinde P yansıma matrisinin ilgili satırını hesaplamaktadır. Yansıma matrisinin satırlarının hesaplanması tüm çerçevelere bağlı olmadığı için tek bir programcık bu iki işi bir arada yapabilmektedir. İlk programcık işini bitirdiğinde diğer tamponu doldurmaya başlar ve dolu tamponu diğer programcıya bırakır. İkinci programcık enerji matrisini hesaplar ve bu matris üzerindeki minimum enerjili yolları bulur.



Şekil 2.3. Gerçek-zamanlı video özetleme için ikili tampon yaklaşımı.

Geliştirdiğimiz yöntemin detaylı açıklamasına öncelikle sıklık tabanlı arkan alan çıkartma yöntemini açıklanarak devam edilecektir. Göz-bakışı bilgisinin sisteme dâhil edilmesi daha sonra anlatılacaktır.

2.2.1. Sıklık Tabanlı Arka Alan Çıkartma Yöntemi

Yıldız ve diğerlerinin [Yıldız et al., 2007] geliştirdiği video özetleme yöntemi oldukça hızlı çalışsa da yöntemin doğrudan kullanılmasının geliştirilen sistem için sıkıntılar doğurabileceği gözlenmiştir. Bu sıkıntılardan ilk ikisi E enerji matrisinin hesaplama yönteminden kaynaklanmaktadır. Enerji değerleri renk değerlerinin farkı

olarak hesaplandığından enerji değeri arka alan renginden farklı nesneler için yüksek çıkmaktadır. Bu değer nesnelerin renkleri arka alan renklerine yaklaştıkça azalmaktadır. Bu şekilde hesaplamanın bir diğer problemi ise ışık değişimlerinin hareket olarak algılanmasıdır. Bu problemlerin dışında sistem kaç tane çerçeveyi atması gerektiğini dışarıdan verilen bir parametre yardımıyla belirleyebilmektedir. Bu tür bir parametrenin güvenlik sistemindeki görüntüler için belirlenmesi mümkün değildir. Bunun için sistemin otomatik olarak atılacak çerçeve sayısını belirlemesi gerekmektedir. Son olarak, önceki yöntem göz-bakışı bilgisinin sistem içine dâhil edilmesi için gerekli mekanizmalara sahip değildir.

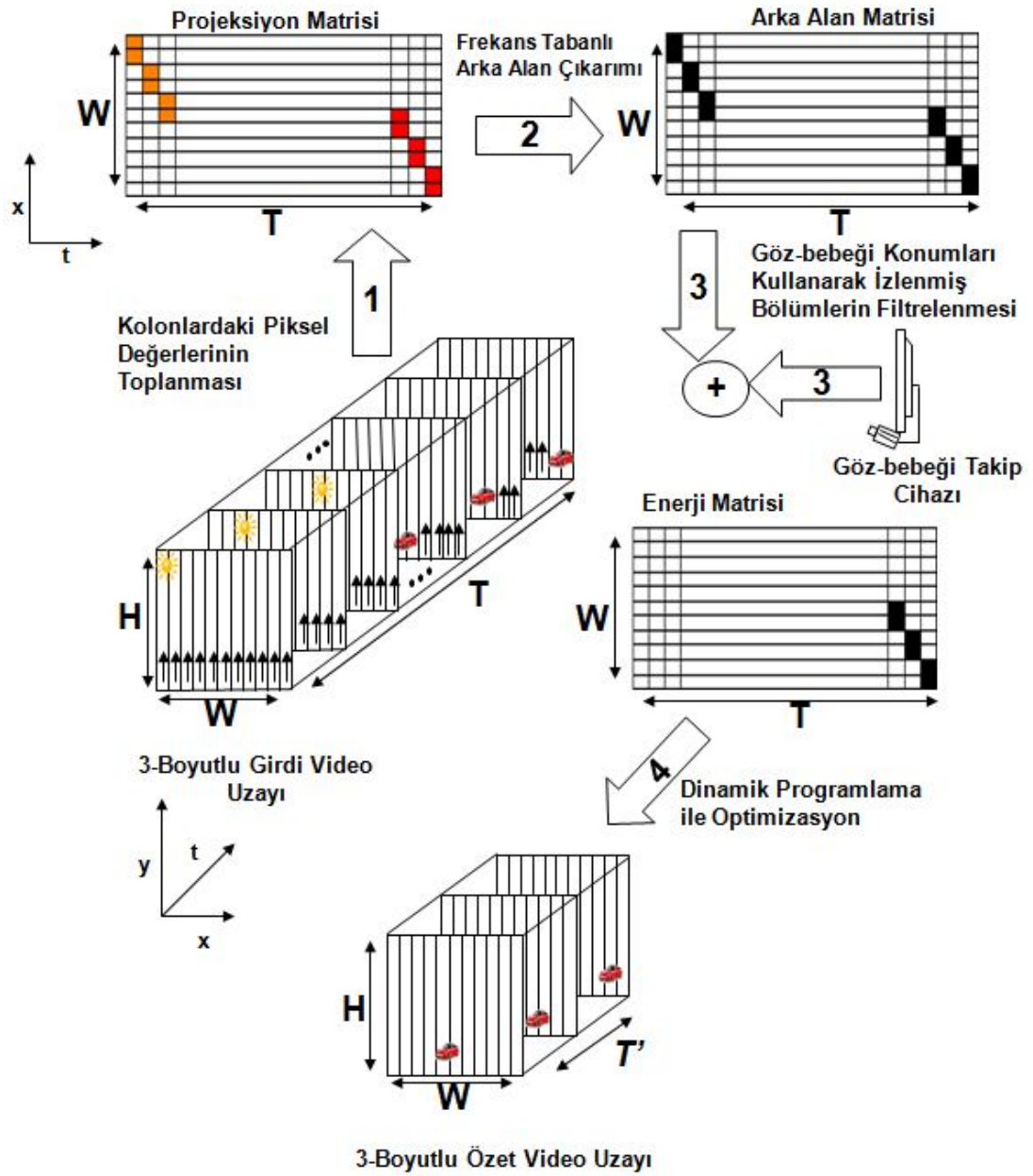
Yapılan çalışmalarda bu sorunlara çözüm üretebilecek yeni sıklık tabanlı arka alan çıkartma yöntemi geliştirilmiştir. Arka alan ve eylem bilgileri B arka alan matrisi içinde tutulmaktadır. Bunun için öncelikle P yansıma matrisindeki değerler bir S ölçekleme parametresi kullanılarak $[0, S]$ arasına ölçeklenmektedir. Sınırlı bir S değerinin kullanılmasıyla sistem büyük toplam değerlerinden kurtarılmış ve sıklık grafiği tabanlı sıklık analizi mümkün olmuştur.

Geliştirdiğimiz arka alan çıkartma metodu S boyutunda bir A sıklık grafiği dizisi kullanmaktadır. Yöntem P yansıma matrisi üzerindeki her bir satırdaki ölçeklenmiş değerleri sayarak A dizisini doldurmaktadır. $w \in [1, W]$ satırı için sıklık grafiği dizisi aşağıdaki şekilde hesaplanır.

$$A_w[P(w, t)] = A_w[P(w, t)] + 1 \forall t, 1 \leq t \leq T. \quad (2.7)$$

Eylem ve arka alanların ayrılması için Zhang ve Nayar'ın yöntemine [Zhang and Nayar, 2006] benzer bir yöntem kullanılmıştır. Hareketli nesnelerin sıklık grafiğindeki değerlerinin arka alandan küçük olması gerektiğini kabul edilmiştir ve belli bir eşik değerin altında kalan değerler “EYLEM” olarak, diğerleri ise “ARKA ALAN” olarak işaretlenmiştir.

$$B(w, k) = \begin{cases} \text{EYLEM}, & \text{eger } A_w(k) \leq \text{esik} \\ \text{ARKA ALAN}, & \text{diğer durumlarda.} \end{cases} \quad (2.8)$$



Şekil 2.4. Göz-bakışı tabanlı özetleme yönteminin temel şeması.

2.2.2. Operatörlerinin Göz-Bakışı Pozisyonlarının Takip Edilmesi

Geliştirilen sistem E enerji matrisini hesaplamak için hem B arka alan matrisine hem de operatörlerin göz-bakışı konumlarına ihtiyaç duymaktadır. Geliştirilen sistemde LC Technologies [LC Technologies, 1997] firmasının geliştirdiği göz-bakışı takip cihazını kullanılmıştır ama güvenlik operatörlerini

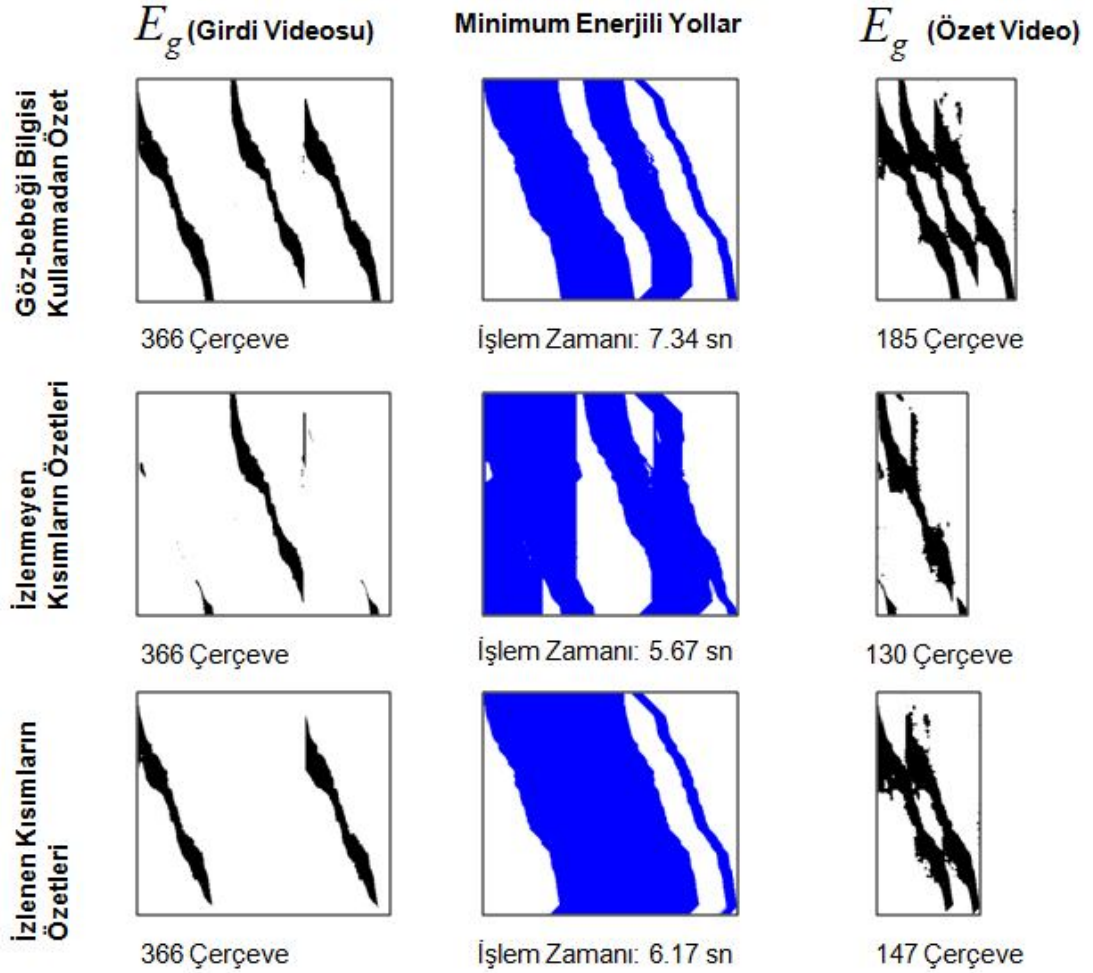
rahatsız etmeyecek herhangi bir takip yöntemi de sistemle çalışacaktır. Göz-bakışı konum takip cihazı, geliştirilen uygulamayla iletişim içinde çalışmakta ve operatörlerin göz-bakışlarının x ve y konumlarını sisteme göndermektedir. Öncelikle bir çerçevenin izlenip izlenmediğinin analizi yapılmaktadır. $L(t)$, t . zamandaki çerçevenin izlenme etiketi

$$L(t) = \begin{cases} \text{İZLENMİŞ}, & \text{eger } G_x(t) \in [0, W] \wedge G_y(t) \in [0, H], \\ \text{İZLENMEMİŞ}, & \text{diğer durumlarda,} \end{cases} \quad (2.9)$$

şeklinde bulunmaktadır. Burada $G_x(t)$ ve $G_y(t)$ sırasıyla t . çerçevedeki x ve y koordinat bilgisini vermektedir. Bu etiketleme ile göz kırpma sonucu oluşabilecek yanıltıcı etkilerden de sistemi kurtarmak mümkün olmaktadır. Eğer göz-bakışı sadece bir kaç çerçeve yok ise sistem göz-bakışı bilgisini önceki geçerli değerle doldurmaktadır. Bu süre geçildiğinde tüm çerçeve “İZLENMEMİŞ” olarak etiketlenmektedir.

“İZLENMİŞ” çerçeveler için daha kapsamlı bir analizin yapılması gereklidir. Bu analiz sonucunda operatörün çerçevenin hangi bölgesine yoğunlaştığı belirlenir. Çok gelişmiş algılama ve takip yetenekleri olsa da, insanlar çok fazla sayıda hareketli nesne geldiğinde sadece bunun küçük bir alt kümesi başarılı bir şekilde takip edilebilirler [Sears and Pylyshyn, 2000]. Bu tür durumların tespit edilmesi ile bir eylemin gözden kaçırılıp kaçırılmadığı anlaşılabilir. İnsanların takip edebileceği hareketli nesnelerin sayısı sınırlı olsa da insan beyni bu tür bir durumda takip edilebilecek nesne sayısını yükseltebilir. Bunun için, insanlar hareketli nesneleri teker teker izlemek yerine bu nesnelerin orta noktalarına odaklanırlar. Tek bir nesne için onun yakınında bir yere odaklanırken (Bkz. Şekil 2.6(a)), birden çok hareketli nesne olduğunda bu nesnelerin arasında bir yere odaklanırlar (Bkz. Şekil 2.6(b)). Bu yöntem görece başarılı olsa da insanların en fazla 5-8 nesneyi takip edebildiği anlaşılmıştır [Fehd and Seiffert, 2008]. İnsanların doğrudan nesnelerin üzerine odaklanmak yerine nesne/nesnelerin yakın bir çevresine odaklanması, insanların sadece odak noktalarını değil o noktanın çevresindeki belli bir alanı görebildiğini göstermektedir. Bu alanın uzağında eylemler varsa bunların gözden kaçırılmış

olabileceği kabul edilir. Geliştirilmiş olan sistemde, bu alanın dairesel bir alan olduğu kabul edilmiştir. Çemberin yarıçapı deneysel olarak ekran boyutlarının çeyreği olarak belirlenmiştir.



Şekil 2.5. İlk video için enerji matrisleri ve minimum enerjili yollar.

Geliştirilen yöntemde göz-bakışının gördüğü alandaki eylemlere ağırlık vermek için bir ω ağırlık fonksiyonu tanımlanmıştır. Ağırlık fonksiyonu pozitif değerlerden oluşuyorsa bu fonksiyon ω^+ olarak tanımlanmıştır. ω^+ fonksiyonu kullanıldığında görülen yerlerin özete eklenmesi sağlanır. Eğer ağırlık fonksiyonu negatif değerlerden oluşuyorsa ω^- olarak tanımlanır ve gözden kaçan eylemlerin video özetine eklenmesi sağlanır. Geliştirdiğimiz yöntemde videodaki kolonların bir kısmının atılması mümkün olmadığı için iki boyutlu bir ω fonksiyonu gereksizdir.

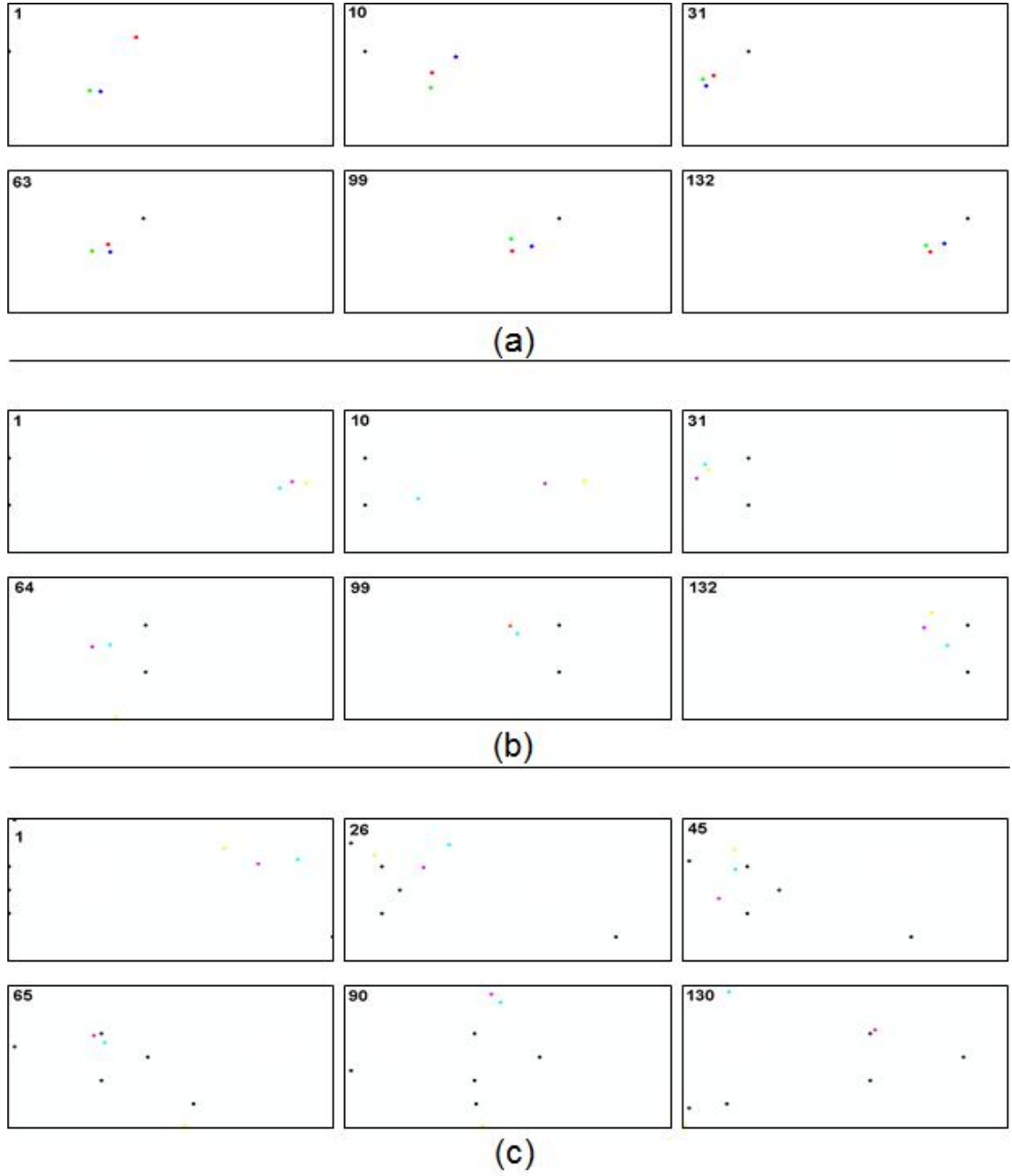
Bu ağırlık fonksiyonunu kolay bir şekilde uygulamak için tek boyutlu $2r+1$ boyutunda bir dizi kullanılmıştır.

Göz-bakışı bilgisini içeren enerji matrisi $E_{Bakış}$, ağırlık fonksiyonu ω ve arka alan matrisi B kullanılarak aşağıdaki şekilde elde edilebilir (Bkz. Eşitlik 2.10).

$$E_{bakış}(w, t) = \begin{cases} B(w, t)\omega[G_x(t) - w], & \text{eger } [G_x(t) - w] \leq r \\ B(w, t), & \text{diger durumlarda} \end{cases} \quad (2.10)$$

2.3. Deneyler

Deneyleri iki grup içinde açıklamak mümkündür. İlk grup deneylerde insanların hareketli nesneleri algılama ve takip etme yöntemleri ve mekanizmaları incelenmiştir. Bu deney göz bebeklerinin konumlarıyla hareketler arasındaki ilişkinin anlaşılması için gereklidir. Deney için altı adet sentetik video üretilmiş ve bu videolar bir test grubuna izlettirilmiştir. Her bir videoda farklı hareket karakteristiklerinde ve sayılarda hareketli nesneler bulunmaktadır. Yapılan sentetik deneyin sonuçları Şekil 2.6'da verilmiştir. Buradaki siyah noktalar hareketli nesneleri diğer renk noktalar ise üç farklı deneğin göz-bakışı konumlarını göstermektedir. İnsanlar sabit hızla hareket eden tek bir hareketli nesneyi takip ederken nesnenin üzerine odaklanmak yerine ona yakın bir bölgeye odaklanmaktadır (Bkz. Şekil 2.6(a)), yine aynı şekilde aynı yöne hareket eden iki noktanın izlenmesi sırasında da gözler bu iki nesne arasında sıçramalar yapmak yerine nesnelerin ortasında bir yere odaklanarak takip etmektedir (Bkz. Şekil 2.6(b)). İnsanlar nesnelerin sayıları, hareket yönleri ve konumları değiştiğinde zorlanmakta ve gözleri ani sıçramalar yapmaktadır. İlk iki durumda deneklerin göz-bakışı konumları hemen hemen aynı rotayı izlerken, çok sayıda ve değişik şekilde hareket eden nesnelerinden oluşan son test videosunda, her bir deneğin göz-bakış konumlarının oldukça farklı şekilde değiştiği görülmüştür (Bkz. Şekil 2.6(c)).



Şekil 2.6. Farklı hareket türleri için göz hareketlerinin incelenmesi.

Yapılan deneyler insanların ilk kez bir hareketli nesneyi algıladıklarında ve daha sonrasında o nesneyi nasıl takip ederken göz bebeklerinin konumlarının nasıl değiştiğini göstermiştir. Bir insan ilk kez bir hareketli nesne ile karşılaştığında öncelikle ilk iki saniye kadar göz-bakışı ani bir şekilde büyük sıçrama hareketleri yapmaktadır. Bu süreden sonra tüm deneklerin göz bebeklerin kararlı bir yörüngeye oturduğu ve hareketli nesneyi bu şekilde takip ettiği gözlenmiştir. İlk baştaki sıçrama

sırasındaki kararsızlık bazı eylemlerin kaçırılmasına neden olabilir. Bunun dışında deneyler Fend ve Seiffert'in tezlerini doğrulamıştır [Fehd and Seiffert, 2008]. İnsanlar hareketli nesnelerin üzerlerine odaklanmak yerine çevresine odaklanmayı tercih etmektedir. Bu durum, bir ilgi penceresi kullanmak gerektiğini göstermektedir.

İkinci grup deneylerde, sistem iki farklı güvenlik kamerası kaydı ile test edilmiştir. Göz-bakışı bilgisi kullanarak hem izlenen hem de gözden kaçırılan kısımların özetleri çıkartılmıştır. Bunun dışında göz-bakışı bilgisinin kullanmayan doğrusal olmayan video özeti de deneylerde elde edilmiştir. Her iki video için de üç adet özet elde edilmiştir ve sonuçların bazı çerçeveleri gösterilmiştir. Her iki video laboratuvar ortamında çekilmiştir ve operatör belli noktaları izlemeye zorlanmıştır. Tüm video kayıtları 320 x 240 boyutunda ve saniyede 15 çerçeve içerecek şekilde çekilmiştir. Deneyler için S ölçekleme parametresi 255 ve eşik değeri 5 olarak seçilmiştir.

İlk deney videosunda çerçevelerin tamamen gözlemlendiği ya da gözden kaçırıldığı durum incelenmiştir. Sahneye önce bir kişi girip yürümektedir. İlk kişi sahneden çıktığında ikinci bir kişi hemen hemen aynı rotayı izleyerek yürümektedir. İkinci kişi sahneyi terk ettiğinde birinci kişi tekrar sahneye girmekte ve aynı yolda yürümektedir. Operatörün birinci kişiyi izlemesi ve ikinci kişi sahneye girdiğinde izlemeyi kesmesi istenmiştir. Bu senaryonun örnek çerçeveleri Şekil 2-1'de gösterilmektedir. Bu örnek çerçevelerin dışında ilk senaryo için çalışma zamanları, enerji matrisleri ve minimum enerjili yollar Şekil 2-5'de gösterilmiştir. İlk video 24 saniye uzunluğundadır ve sistem gözden kaçırılan kısımların özetini 5,67 saniyede, izlenen kısımların özetini ise 6,17 saniyede hesaplamaktadır. Buradan görülebileceği gibi sistemin çalışma zamanı bulunan yol sayısı ile ters orantılıdır, yani dinamik programlama ile optimizasyon oldukça hızlıdır. Sistemi yavaşlatan kısım videonun yeninden oluşturulma kısmıdır ki özet videonun boyutu büyüdükçe çalışma süresi de artmaktadır.



(a) 359 çerçeveden oluşan bir girdi videosundan örnek çerçeveler gösterilmiştir. Küçük mavi daire operatörün göz-bebeğini göstermektedir.



(b) Göz-bebeği bilgisini kullanmayan lineer olmayan video özeti 209 çerçeve uzunluğundadır. İşlem süresi 7,78 saniyedir.



(c) Operatörün gözden kaçırdığı kısımların özetidir. 95 çerçeve uzunluğundadır. Özet video 4,39 saniyede üretilmiştir.



(d) Operatörün izlediği kısımların özeti 137 çerçevedir ve 5.47 saniyede üretilmiştir.

Şekil 2.7. İkinci deney videosu ve ilgili özetlerinden örnek çerçeveler.

Son deney senaryosunda ise bir çerçeve içindeki belli bazı hareketlerin kaçırılma durumu incelenmiştir. Bu senaryoda bir çanta çalınmakta fakat operatör başka bir taraftaki eylemle ilgilenmektedir. Sistem hırsızlık olayını tekrar operatöre göstermektedir. Bu işi 24 saniyelik video için 4,39 saniyede yapabilmektedir. Özet videoda bazı bozulmalar gözükmemektedir. Bunun nedeni sabit ilgi penceresi yarıçapıdır. Pencere bir hareketin sadece küçük bir kısmını içeriyorsa hareketin diğer kısımları atıldığında göze hoş gelmeyen bozulmalar oluşabilmektedir. Bunun için kullanılabileceğini düşündüğümüz bir yöntem hareket tabanlı dilimlemedir. Bu

yöntem ile pencerenin dokunduğu eylemlerin tamamı özete eklenebilir. Gerçek-zamanlı çalışma için sisteme böyle bir parça eklenmesi uygun görülmüştür.

2.4. Yorumlar

Yapılan çalışma ile yeni bir gerçek-zamanlı göz-bakışı bilgisini kullanan video özetleme sistemi ortaya konmuştur. Sistemin geliştirilmesi için doğrusal olmayan video özetleme algoritması göz-bakışı bilgisini kullanacak şekilde geliştirilmiştir. Video içinde izlenmiş ya da gözden kaçırılmış bölümler içinde arama yapma imkânı doğmuştur. Bu sistem ile güvenlik operatörlerine bir üst denetim mekanizması geliştirilmiştir. Elde edilen video özetleri sonucu operatörlerin performansları kıyaslanabilecek, çalışması gereken operatör sayısı ve operatörlerin çalışma süreleri belirlenebilecektir.

Sistem güvenlik operatörlerini bir göz takip cihazına bağladığı için operatörlere çalışma sırasında rahatsızlık verebilir. Bu durumda operatörlere temas etmeyen yöntemler göz-bakışının elde edilmesi için kullanılabilir. Bunun dışında sistemin bir diğer sıkıntısı kolonların yansıtılması sırasında bazı hareketlerin kaybedilmesidir. Bu tür kayıplar sadece dikey yönlü hareketler için geçerli olduğundan güvenlik sistemleri için etkisinin sınırlı olacağı düşünülmektedir. Yapılan deneylerde sadece hareketli nesnelerin ilgi uyandıracığı düşünülse de gerçekte göz-bakışının konumlandığı herhangi bir ilgi konusu için de önerilen sistem çalışacaktır.

3. YÜKSEK ÇÖZÜNÜRLÜKLÜ VİDEOLARDA PARALEL ÖZETLEME

Güvenlik gözetleme sistemlerinde kullanılan kamera sistemlerin birçoğu eski teknolojileri kullanmaktadır. Çözünürlük ve netlik gibi problemlerinden ötürü güvenlik kamera sistemlerinden elde edilen görüntüler suçlu teşhisinde kullanılamamakta ve çoğu durumda mahkemelerce delil olarak kabul görmemektedir [Keval and Sasse, 2006]. Bu durum neticesinde bazı devletlerin güvenlik kurumları güvenlik gözetleme sistemlerinin etkinliğini tartışmaya başlamışlardır [Keval and Sasse, 2010].

Önümüzdeki yıllarda düşük çözünürlüklü güvenlik kameralarının daha yüksek çözünürlüklü kameralarla değiştirilmesi beklenmektedir [Sasse, 2010]. Yeni nesil güvenlik kameraları daha yüksek kalitede videolar üretecektir. Bu videoların bilgisayarlarda işlenmesi ile daha doğru bir biçimde nesnelerin ve hareketlerin tanınması mümkün olacaktır. Yüksek çözünürlükteki video verisinin işlenmesi için ihtiyaç duyulacak zamansa çok daha yüksek olacaktır. Bu nedenle, günümüzde gerçek-zaman gereksinimi karşılayan bazı sistemler yeni nesil kameralar kullanıldığında güvenlik sektöründe çalışmayacaktır. Daha önce geliştirmiş olduğumuz sistem (Bkz. Bölüm 2) düşük çözünürlüklü güvenlik videolarında gecikmeli gerçek-zaman gereksinimlerini karşılarken yüksek çözünürlük videolarda gecikmenin aşırı arttığı görülmüştür. Bununla birlikte, sistemdeki bu gecikmeleri makul zamanlara çekebilecek iyileştirmelerin olduğu ve Bölüm 2’de açıklanan sistemde önemli miktarda zamana ihtiyaç duyan birçok alt parçanın yüksek derecede paralel hesaplanabilir olduğu anlaşılmıştır [Vural and Akgul, 2011a].

3.1. Yöntem

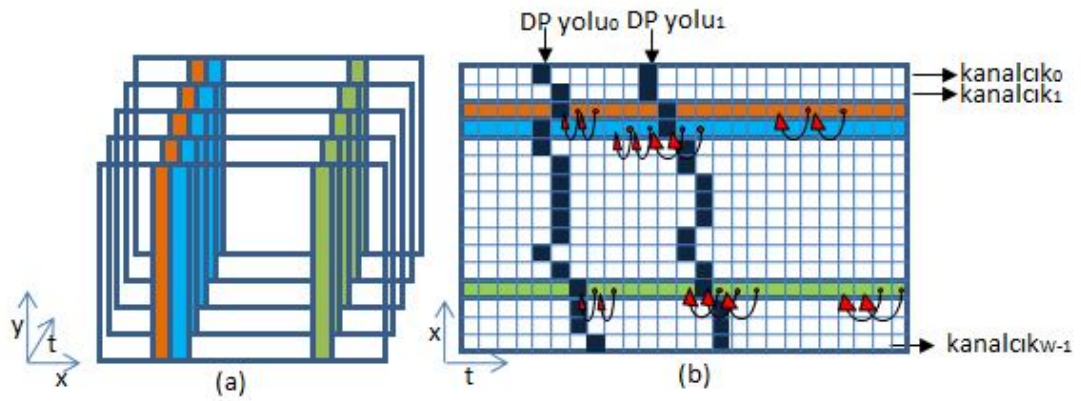
Sistem üzerinde yapılan analizlerde önceki sistemin en fazla zaman harcayan bölümünün videonun yeniden sentezlenmesi aşaması olduğu görülmüştür. Yüksek enerjili video kolonlarının özet videodaki yeni yerlerine kopyalanması sistemin en yavaş olduğu bölümdür. Yapılan deneylerde asgari enerjili yol sayısı arttıkça daha

hızlı özetleme yapıldığı görülmüştür. Bir diğer deyişle taşınacak kolon miktarı azaldıkça sistem hızlanmaktadır. Tüm sistemin hızlandırılması için sistemin en yavaş aşamasının öncelikli olarak hızlandırılması gerektiği düşünülmüştür. Videonun yeniden sentezlemesi kısmının hızlandırılması için iki temel iyileştirme yapılmıştır. Bu iyileştirmelerden ilki gereksiz kopyalamaların engellenmesidir. İkinci iyileştirme ise kopyalama işleminin paralel bir algoritma kullanılarak yapılmasıdır.

Temel sistem [Yildiz et al., 2008], bir minimum enerjili yolun mutlaka arka alan kolonlarından oluştuğundan emin olmakla birlikte, geri kalan kolonların hareket mi yoksa arka alan mı olduğunu bilmemektedir. Bunun altında yatan temel neden hızlı doğrusal-olmayan video özetleme yönteminin gerçek sayı değerlerinden oluşan bir enerji matrisi kullanmasıdır. Kullanılan bu gerçek sayı değerli enerji matrisi kolonların hareket etiketi hakkında kesin bir şey söyleyememektedir. Bu nedenle bu yöntem asgari enerjili yolların dışında kalan tüm video kolonlarını taşımak zorundadır. Daha önce geliştirilmiş [Vural and Akgul, 2009] ve Bölüm 2’de açıklanmış olan gerçek zamanlı video özetleme yönteminde kullanılan ikili enerji haritası, daha verimli bir şekilde kullanılarak sistem hızlandırılacaktır. İkili enerji haritası arka alan ve eylem içeren kolonları birbirlerinden ayırmaktadır. Bu ayırımdan yararlanılarak sadece hareket içeren kolonların özet videoya taşınması sağlanacaktır. Bunun için, özet video tamamen arka alanı içeren çerçevelerden oluşturulacaktır. İlkendirme çerçevesinin bulunması için ise sistemin bulduğu ilk sıfır enerjili yol kullanılacaktır. Bu yola denk düşen video kolonları eylem içermeyeceğinden sahnenin durağan halinin bir resmi elde edilecektir. Bu iyileştirme çok sayıda gereksiz kopyalamayı engelleyerek önemli ölçüde iyileşme sağlamıştır.

Yapılacak ikinci iyileştirme ise kopyalama aşamasının paralel hesaplanmasıdır. Doğrusal olmayan video özetleme yönteminde bir video kolonu özet videoda sadece aynı kolon indeksine sahip bir yere taşınabilir. Bir başka deyişle bir video kolonu sadece zaman eksenini boyunca kaydırılabilir. Bu kaydırma işlemi diğer kolonlardaki kaydırma işlemlerinden tamamen bağımsızdır. Yeni sistemde E enerji matrisinin her bir satırına bir kanalcık atanmıştır. Her bir kanalcık sadece kendi kolon indeksindeki video kolonlarının taşınması işine bakmaktadır. Bir girdi videosundaki tüm çerçeveler için aynı kolona denk düşen pikseller tek bir kanalcık tarafından

işlenecektir. Bu sistemde kolon sayısı kadar paralel yapılabilecek bağımsız iş olduğu anlamına gelmektedir (Bkz. Şekil 3.1(a)). Şekil 3'de her bir farklı renk, farklı bir kolon indeksini göstermektedir. Aynı renkle gösterilmiş kolonlar, videonun farklı çerçevelerindeki aynı kolon indeksine sahip kolonlardır. Verilen girdi videosunun E enerji matrisi Şekil 3.1(b)'de verilmiştir. Bu enerji matrisinde bulunan iki tane asgari enerjili yol örnek olarak gösterilmiştir. Her bir farklı renge bir kanalcık atanmıştır. Bu kanalcıklar ilgili kolon indeksindeki taşıma işlemlerini yapacaktır.

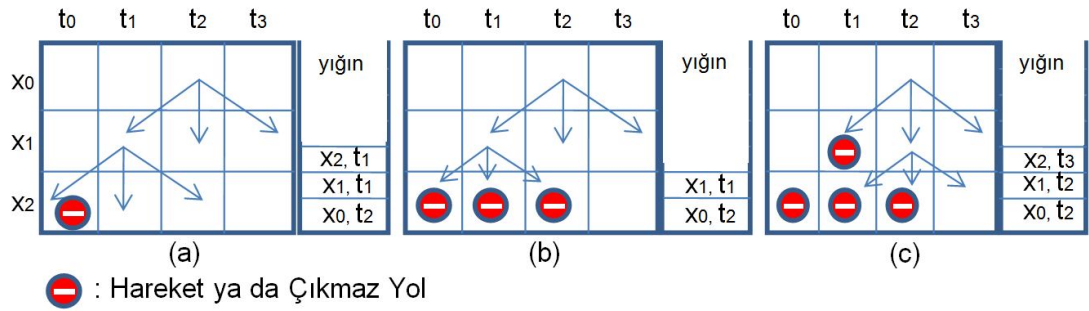


Şekil 3.1. Girdi video uzayı ve enerji matrisi üzerinde paralelleştirme.

Yıldız ve diğerleri [Yildiz et al., 2008] tarafından geliştirilen dinamik programlama tabanlı video özetleme algoritmasının çok zaman harcadığı bir başka kısım ise optimizasyon aşamasıdır. Dinamik programlama hızlı olmakla birlikte toplam çalışma zamanı açısından bakıldığında eski sistemlerin en çok zaman harcadığı ikinci kısımdır. Dinamik programlama ile asgari enerjili yolların bulunması sırasında tüm olası yolların enerjileri bulunmakta ve bu yollardan en düşük enerjili olanı seçilmektedir.

Bu en düşük enerjili yol seçildikten sonra işaretlenmekte ve enerji matrisinden çıkartılmaktadır. Sistem her bir yolu matristen çıkarttığında matristedeki video kolon komşulukları değişmekte ve ilgili kolonlar için enerji matrisinin yeniden hesaplanması gerekmektedir. Hızlı doğrusal olmayan video özetleme yöntemi gerçek sayılı enerji matrisinde [Yildiz et al., 2008], gerçek-zamanlı operatörler başarımlı ölçütü tabanlı video özetleme yöntemi [Vural and Akgul, 2009] ise ikili enerji matrisi üzerinde dinamik programlama kullanmıştır. Enerji haritasında ikili değerler tutmak

dinamik programlamadan daha verimli başka bir yöntemi mümkün kılmaktadır. Geliştirilen yöntemde dinamik programlama kullanılmadan ikili enerji matrisi üzerinde asgari enerjili yollar bulunmuştur. Bir yol sadece arka alana ait kolonlara denk düşen elemanlardan oluşabilir. Enerji matrisinde arka alan kolonuna denk gelen elemanların değeri 0, hareketli bölgelerin değeri ise 1'dir. Yığın veri yapısı kullanılarak enerji matrisinin en üst satırından en alt satırına kadar incek şekilde bir yol bulunur. Yolu oluşturmak için her bir satırdan bir eleman seçilerek yığına atılır. Yığına son satırdaki bir eleman eklendiğinde bir yol bulunmuş olur. Yığına sadece değeri sıfır olan ve daha önce başka bir yola eklenmemiş elemanlar eklenir (Bkz. Şekil 3.2(a)). Yığın veri yapısı kullanıldığından çıkmaz yolların bulunup geri dönülmesi ve başka bir yolun denenmesi oldukça kolay olmaktadır. Son satırdan önceki bir satırdaki elemandan alt satırda gidilebilecek başka bir yol kalmamışsa bu eleman “çıkamaz yol” olarak işaretlenir ve yığından çıkartılır (Bkz. Şekil 3.2(b)). Yığının en üstündeki elemandan yeni bir yol aranır (Bkz. Şekil 3.2(c)). Bu yöntem ile dinamik programlamadaki en iyi yol araması ve matris güncellemesi işlemlerinin zaman karmaşıklığından sistem kurtarılmıştır.



Şekil 3.2. DP alternatifi yığın tabanlı asgari enerjili yol bulma yöntemi.

Bu iyileştirmelerin dışında, P yansıtım ve E enerji matrislerinin hesaplanmalarında da paralellik söz konusudur. Her bir matris elemanı farklı kanalcıklar kullanılarak hesaplanabilir. Bununla birlikte bu alt parçalar mevcut seri sistemlerde de çok fazla zaman harcamadıklarından yapılacak iyileştirmelerin getirilerinin sınırlı olacağı düşünülmüştür. P yansıtım matrisinin hesaplanmasında bir boru hattı mekanizması kullanılarak her bir video çerçevesi elde edildiğinde o çerçeveye ait P matrisi kolonunun hesaplanması sağlanabilir.

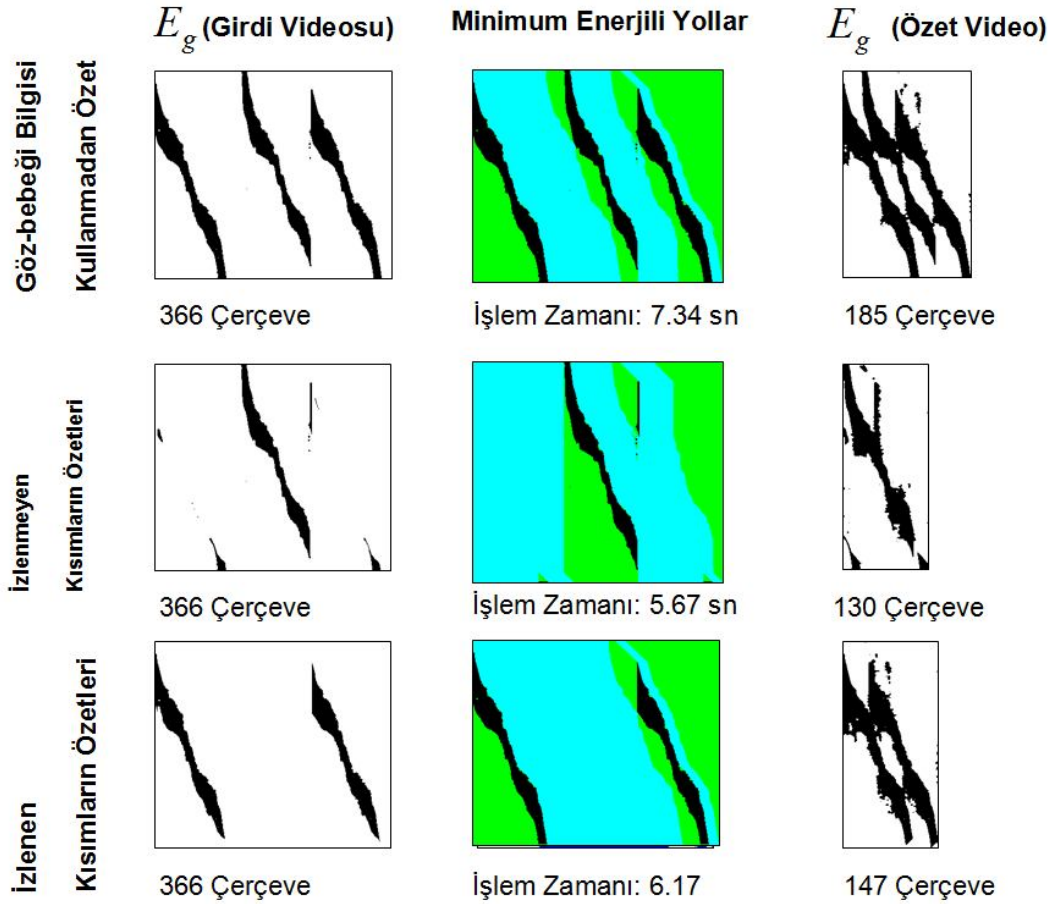
Gerçekleştirilen iyileştirmeler farklı boyutlardaki örnek videolarda ve farklı donanımlardaki bilgisayarlarda denenmiştir. Elde edilen sonuçlar, gerçekleştirilen sistemin yüksek çözünürlüklü videolarda gerçek zaman gereksinimlerini karşılayacağını göstermiştir.

3.2. Deneyler

Sistem iki farklı güvenlik kamerası kaydı ile test edilmiştir. Göz-bakışı bilgisi kullanarak hem izlenen hem de gözden kaçırılan kısımların özetleri çıkartılmıştır. Bunun dışında göz-bakışı bilgisinin kullanmayan doğrusal olmayan video özeti de deneylerde elde edilmiştir. Her iki video için de üç adet özet elde edilmiştir ve sonuçların bazı çerçeveleri gösterilmiştir. Her iki video laboratuvar ortamında çekilmiştir ve operatör belli noktaları izlemeye zorlanmıştır. Tüm video kayıtları 320 x 240, 640 x 480, 1024 x 768 ve 1920 x 1080 çözünürlüklerinde test edilmişlerdir. Deneyler için S ölçekleme parametresi 255 ve eşik değeri 5 olarak seçilmiştir. Deneyler farklı donanım yapılarına sahip iki kişisel bilgisayar üzerinde gerçekleştirilmiştir. Kullanılan ilk bilgisayarın üzerinde 1 GB bellek ve üstün-kanalcık teknolojisine sahip Intel Pentium 4 3.2 Ghz MİB, ikinci bilgisayarda ise 2 GB bellek ve Intel QuadCore 2.5 Ghz MİB bulunmaktadır.

İlk deney videosunda çerçevelerin tamamen gözlemlendiği ya da gözden kaçırıldığı durum incelenmiştir. Sahneye önce bir kişi girip yürümektedir. İlk kişi sahneden çıktığında ikinci bir kişi hemen hemen aynı rotayı izleyerek yürümektedir. İkinci kişi sahneyi terk ettiğinde birinci kişi tekrar sahneye girmekte ve aynı yolda yürümektedir. Operatörün birinci kişiyi izlemesi ve ikinci kişi sahneye girdiğinde izlemeyi kesmesi istenmiştir. Bu senaryonun örnek çerçeveleri Şekil 2.1’de gösterilmektedir. Bu örnek çerçevelerin dışında ilk senaryo için çalışma zamanları, enerji matrisleri ve minimum enerjili yollar Şekil 3.3’de gösterilmiştir. İlk video 24 saniye uzunluğundadır ve sistem gözden kaçırılan kısımların özetini 5,67 saniyede, izlenen kısımların özetini ise 6,17 saniyede hesaplamaktadır. Buradan görülebileceği gibi sistemin çalışma zamanı bulunan yol sayısı ile ters orantılıdır. Sistemi

yavaşlatan kısım videonun yeniden oluşturulma kısmıdır ki özet videonun boyutu büyüdükçe çalışma süresi de artmaktadır.



Şekil 3.3. İlk deney videosu için enerji matrisleri ve minimum enerjili yollar.

Siyah renkli bölgeler eylemlerin izdüşümünü vermektedirler. Mavi ile işaretlenen bölgeler minimum enerjili yolları göstermektedir. Yeşille işaretlenen bölgeler de maviyle işaretlenen bölgeler gibi arka alan kolonlarına denk düşmektedir. Yeşille işaretlenen bu arka alan kolonlarından geçecek yol bulunamamıştır. Geliştirilen sistem yeşil renkli alanlara denk düşen video kolonlarını kopyalanmaz. Şekilde girdi ve özet videolarının çerçeve sayıları ve tek kanalcıklı video özetleme yönteminin [Vural and Akgul, 2009] işleme süreleri de verilmiştir.

İlk video için tek kanalcıklı eski yöntemin [Vural and Akgul, 2009] sonuçları ile paralelleştirilmiş ve iyileştirilmiş yöntemin [Vural and Akgul, 2011a] sonuçları karşılaştırılmıştır (Bkz. Çizelge 3.1). Çizelgede dört farklı çözünürlük için iki farklı

bilgisayar donanımına göre çalışma zamanları verilmiştir. Verilen çalışma zamanları on çalıştırmanın verdiği ortalama sonuçlardır. Üstün-kanalcık destekli bilgisayar üzerinde 2,5, dört çekirdekli sistemde ise 5 kata kadar hızlanma olduğu görülmüştür. Dört çekirdekli sistemde ortalama çalışma süresi yaklaşık olarak dörtte bire düşmüştür. Ayrıca ilk bilgisayardaki sınırlı belleğin hızlanmayı azalttığı görülmüştür. Geliştirilen sistem düşük çözünürlüklü videolarda yüksek çözünürlüklü videolara oranla daha fazla hızlanmıştır.

Çizelge 3.1. Yöntemlerin birinci test videosunda karşılaştırılması.

	Intel P4 3.2Ghz – HT		Intel QuadCore 2.5Ghz	
	Tek Kanalcıklı	Çok Kanalcıklı	Tek Kanalcıklı	Çok Kanalcıklı
320x240	7,349 (sn)	2,532 (sn)	4,578 (sn)	0,859 (sn)
640x480	23,51 (sn)	10,17 (sn)	16,06 (sn)	2,844 (sn)
1024x768	104,3 (sn)	39,98 (sn)	48,30 (sn)	10,81 (sn)
1920x1080	182,1 (sn)	87,76 (sn)	126,4 (sn)	30,07 (sn)

Her bir iyileştirmenin sonuçlara etkisi ayrıca gösterilmiştir. Çizelge 3.2’de ve Çizelge 3.3’de tek kanalcıklı [Vural and Akgul, 2009] özetleme yöntemi ile yapılan iyileştirmelerin bu sisteme göre etkisi incelenmiştir. Çizelge 3.2’de ve Çizelge 3.3’de, A: Tek kanalcıklı [Vural and Akgul, 2009] sistemi, B: Dinamik programlama kullanmayan tek kanalcıklı sistemi, C: Dinamik programlama kullanmayan çok kanalcıklı sistemi ve D: Çok kanalcıklı, dinamik programlama kullanmayan ve gereksiz video kolonu taşıması yapmayan sistemi göstermektedir. Sonuçlar saniye cinsindendir. Yüzdelere göre iyileşmeyi göstermektedir. Yüksek çözünürlüklü videolarda dinamik programlamanın toplam zaman içinde aldığı zaman görece daha az olduğu görülmüştür. Bu nedenle dinamik programlamanın kaldırılması yüksek çözünürlüklü videolarda sadece %20 etki ederken düşük çözünürlüklü videolarda %50 iyileşme olmaktadır. Çok kanalcık kullanılması ise sistemleri üç ila beş kat hızlandırmaktadır. 1024x768 çözünürlükte

saniyede 25 çerçeve işleyebilen yöntem, 1920x1080 çözünürlük için saniyede 15 çerçevenin altında kalmış ve saniyede 11 çerçeve işleyebilmiştir. Bununla birlikte günümüzde bu çözünürlükte çalışan güvenlik videoları azami saniyede 10 çerçeve ile çalışmaktadır.

Çizelge 3.2. İlk video üzerinde yapılan iyileştirmelerin etkileri.

	A	B	C	D
Tam Özet (1920 x 1080)	126,4	101,0 (%125)	52,59 (%240)	30,07 (%420)
Tam Özet (1024 x 768)	48,30	39,94 (%121)	22,38 (%216)	10,81 (%447)
Tam Özet (320 x 240)	4,578	2,375 (%193)	1,094 (%418)	0,859 (%533)
İzlenen (320 x 240)	3,962	1,782 (%222)	1,031 (%384)	0,750 (%528)
Gözden Kaçan (320 x 240)	4,047	2,031 (%199)	1,203 (%336)	0,843 (%480)

Çizelge 3.3. İkinci test videosu üzerinde yapılan iyileştirmelerin etkileri.

	A	B	C	D
Tam Özet (320 x 240)	3,782	2,406 (%157)	1,297 (%292)	0,797 (%475)
İzlenen (320 x 240)	3,632	1,516 (%240)	0,969 (%375)	0,698 (%520)
Gözden Kaçan (320 x 240)	3,657	2,324 (%157)	1,188 (%308)	0,765 (%478)

Son deneyde senaryosunda ise bir çerçeve içindeki belli bazı hareketlerin kaçırılma durumu incelenmiştir (Bkz. Şekil 3.4). Tablolardaki, A: Tek kanalcıklı [Vural and Akgul, 2009] sistemi, B: Dinamik programlama kullanmayan tek kanalcıklı sistemi, C: Dinamik programlama kullanmayan çok kanalcıklı sistemi ve D: Çok kanalcıklı, dinamik programlama kullanmayan ve gereksiz video kolonu taşıması yapmayan sistemi göstermektedir. Bu senaryoda bir çanta çalınmakta fakat operatör başka bir taraftaki eylemle ilgilenmektedir. Sistem hırsızlık olayını tekrar

operatöre göstermektedir. Bu işi 24 saniyelik video için 4,39 saniyede yapabilmektedir. Özet videoda bazı bozulmalar gözükmemektedir. Bunun nedeni sabit bir ilgi penceresi yarıçapı kullanılmasıdır. Pencere bir hareketin sadece küçük bir kısmını içerdiğinde hareketin diğer kısımları atıldığından göze hoş gelmeyen bozulmalar oluşabilmektedir. Bunun için hareket tabanlı bir dilimleme yöntemi kullanılmıştır. Bu yöntem ile pencerenin dokunduğu eylemlerin tamamı özete eklenebilir.



(a) 359 çerçeveden oluşan bir girdi videosundan örnek çerçeveler gösterilmiştir. Küçük mavi daire operatörün göz-bakışını göstermektedir.



(b) Göz-bakışı bilgisini kullanmayan lineer olmayan video özeti 209 çerçevedir. İşlem süresi 7,78 saniyedir.



(c) Operatörün gözden kaçırdığı kısımların özetidir. 95 çerçeve uzunluğundadır. Özet video 4,39 saniyede üretilmiştir.



(d) Operatörün izlediği kısımların özeti 137 çerçevedir ve 5,47 saniyede üretilmiştir.

Şekil 3.4. İkinci deney videosu ve ilgili özetlerinden örnek çerçeveler.

3.3. Yorumlar

Bu bölümde yapılan çalışma ile Bölüm 2’de gerçekleştirmiş olduğumuz sistemin yüksek çözünürlüklü yeni nesil güvenlik kameralarıyla birlikte kullanılabileceği gösterilmiştir. Yığın tabanlı minimum enerjili yol bulma yönteminin dinamik programlama tabanlı önceki yöntemlere göre çok daha hızlı çalıştığı gösterilmiştir. İki boyutlu yansıtılmış uzay üzerinde tanımlanan video özetleme probleminin yüksek derecede paralel hesaplama özelliğine sahip olduğu gösterilmiş ve yapılan deneylerde dört kata kadar hızlanma sağlanmıştır. Geliştirilen sistemin, teknoloji marketlerindeki mevcut yüksek çözünürlüklü kameraların hızlarında çalışabileceği gösterilmiştir.

Bölüm 2’de ve bu bölümde gerçekleştirilen sistemler operatörlerin bakış alanlarındaki hareketlerin görüldüğü varsayımına dayanmaktadır. Bununla birlikte bakılan bir bölgede olay görülmemiş ya da operatör tarafından anlaşılamamış olabilir. Operatörlerin algı seviyeleri öznel ve zamanla değişkendir. Bölüm 4’de insan operatörlerin algı seviyelerinin analizi yapan ve operatörlerin anlık algı seviyelerine uygun güvenlik videosu sentezleyen sistem açıklanacaktır.

4. OPERATÖR DİKKAT SEVİYESİ TABANLI VIDEO GÖZETLEME

Dijital güvenlik sistemleri günlük hayatın önemli bir parçası haline gelmiştir [Koskela, 2000]. Her geçen gün düşen fiyatlar, dijital güvenlik sistemlerinin popülerliğinin artmasını sağlamıştır. Yaygın olarak kullanılan bu sistemler oldukça büyük miktarlarda güvenlik verisi üretmeye başlamışlardır. Bununla birlikte, bu verilerin tamamen otomatik olarak işlenmesi günümüzde mümkün değildir [Keval and Sasse, 2006]. Otomatik sistemlerin her türden hareketi algılayabilecek seviyede olmaması, insan operatörleri vazgeçilmez bir şekilde güvenlik gözetleme sistemlerinin kilit parçası haline getirmektedir [Dick and Brooks, 2003]. İnsan operatörlerin hata yapmaya açık oluşları [Wallace et al., 1997; Warm et al., 1996] ve sistemin ana maliyet kalemini oluşturmaları göz önüne alındığında operatörlerin en verimli ve kararlı şekilde çalıştırılmasının gerekli olduğu açıktır.

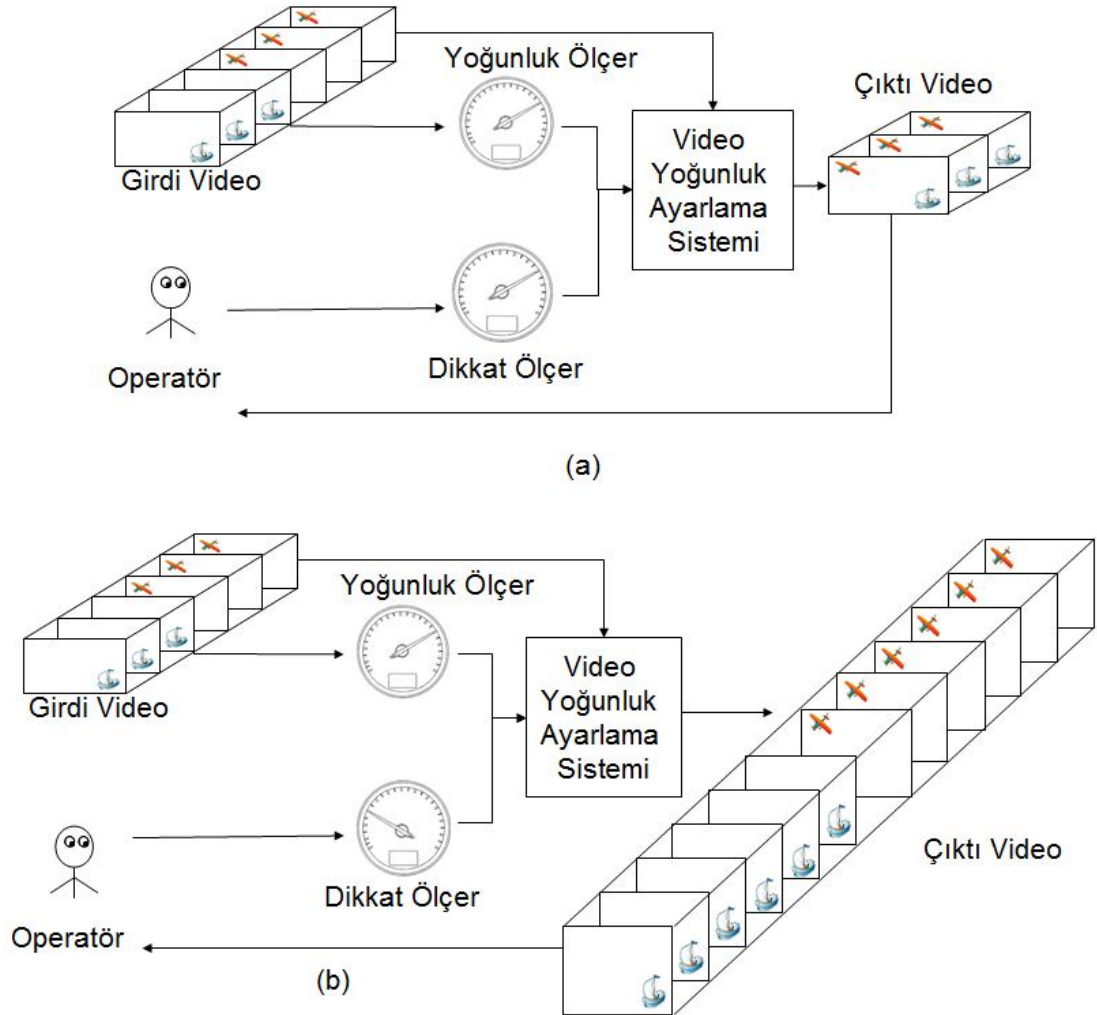
Operatörlerin daha kısa süre içerisinde daha çok görüntüyü izlemesi için video özetleme sistemleri geliştirilmiştir. Bu tür sistemler işgücü maliyetlerinin azaltılmasına yardımcı olmaktadır. Bununla birlikte kullanılan yöntemlerin bir kaç temel noktada sıkıntıları vardır. Doğrusal özet çıkartma yöntemleri [Li et al., 2000] az hareket içeren sahneleri atarak operatörlerin iş yükünü azaltırlar ama bu sistemler hareketli kısımları atmadan çok sıkı özetler çıkartmazlar [Kim and Hwang, 2000]. Bu nedenle olayların kaçırılması için üretilen özetler genellikle uzun tutulur. Son dönemde popülerleşen doğrusal olmayan video özetleme yöntemler daha sıkı özet görüntüler üretebilirler [Acha et al., 2006]. Bunu başarmak için farklı zamanlarda olmuş olan olayları aynı çerçeve içerisine taşırlar. Hareketlerin sadece kronolojik bilgilerini göz ardı eden bu sistemler hareketleri kaybetmezler ve hareketlerin yapılarını korurlar.

Hem doğrusal hem de doğrusal-olmayan video özetleme yöntemlerinin temel problemlerinden biri insan psikolojisini dikkate almamalarıdır. Mevcut tüm özetleme yöntemleri kısıtlı bir video hacmine mümkün olan en fazla sayıda hareketi yerleştirmeye çalışırlar. Bu tür bir yöntem hareketler arası çakışmaları hesaba katıyor olsa bile üretilen özetlerin çerçeve başına düşen hareket sayısı oldukça yüksektir.

İnsan operatörlerin ise çoklu-nesne izleme yetenekleri hem sınırlı hem de değişkendir. Çoklu hareket takip kapasitesi kişiden kişiye değiştiği gibi aynı kişinin takip kapasitesi de zaman içinde artıp azalabilmektedir. Takip edilebilen nesne sayısı öznel olarak değişebildiği gibi takip edilen nesnelerin özellikleri de takip edilebilirliği belirlemektedir. Genel olarak yavaş hareket eden nesneler hızlılara göre [Alvarez and Franconeri, 2007], az sayıda nesnenin takibi kalabalık sahneler göre [Allen et al., 2004; Cavanagh and Alvarez, 2005] ve aynı yöne doğru hareket eden nesnelerin zıt yönler doğru hareket eden nesnelere göre [Scholl et al., 2001] daha kolay takip edildiği bilinmektedir. Bilimsel kaynaklarda insanın sabit bir dikkat modelini kullanan yöntemler bulunmaktadır [Ma et al., 2002; Ren et al., 2007]. Bu yöntemlerin kullandıkları modeller çok geneldir, parametreleri sistemin tasarım aşamasında verilmiştir ve kişilere göre ayarlanmaları oldukça zordur. Durağan modellerin bir diğer sıkıntısı ise insanın dikkat seviyesi sürekli ve sabit bir seviyede düşünmesidir. Gerçekte ise adanmış dikkat seviyesinin sürekli yüksek bir seviyede tutulması sadece bir kaç dakika için mümkündür. Adanmış dikkat gerektiren işler çok yorucudur ve yorulan insanın dikkat seviyesi düşer. Güvenlik operatörlerinin dikkat modelleri operatörlerin öznel ilk durumlarını ve anlık iş yüklerini dikkate almalıdır [Thiffault and Bergeron, 2003]. Operatörün dikkat seviyesi çevre ortamlarından da etkilenebileceği için operatörün izlenmesi gerçek zamanlı dikkatin ölçümü için zorunludur [Oron-Gilad et al., 2002].

Çalışmanın bu bölümünde, operatörün dikkat seviyesine göre güvenlik videosunun yoğunluğunu ve takip edilebilirliğini ayarlayan bir sistem geliştirilmiştir (Bkz. Şekil 4.1). Bu çalışma iki temel yenilik getirmektedir. Bunlardan ilki güvenlik görevlilerinin anlık dikkat seviyelerini ölçülmesi diğeri ise gecikmeli gerçek zamanlı doğrusal olmayan video yoğunluk görüntülerinin oluşturulmasıdır. Sistem öncelikle operatörlerin göz bebeği büyüklüklerini, göz kapanma sıklıklarını ve sıçrama sıklıklarını göz önüne alarak operatörün anlık dikkat seviyesini belirler. Dikkat seviyesi ölçülmüş operatör için sistem onun takip edebilme kapasitesine uygun görüntü sentezlenir. Güvenlik operatörü kamera görüntüsünü izlemek yerine sürekli olarak kendi algısına uygun olarak sentezlenmiş görüntüleri izler. Sistem uygun görüntüyü sentezlemek için öncelikle arka alanda videodaki hareketleri takip ederek,

bulur. Bulunan bu görüntülerin hızları ve yönleri de değiştirilerek uygun yoğunluktaki görüntü elde edilir. Bu görüntüler belli bir gecikmeyle operatöre gösterilir. Bu gecikme dikkat seviyesi yüksek bir operatör için bir kaç saniyeden uzun olmayacaktır. Bununla birlikte dikkat seviyesi çok düşük bir operatör için dakikaları bulabilir ama algısı kapalı bir operatörün bir eylemi kaçırdıktan sonra yapacağı hiç bir şey kalmamaktadır. Sistemin sentezlediği videoların izlenmesiyle eylem kaçırma oranlarının düştüğü ve operatörlerin verimliliklerinin yükseldiği deneylerde gösterilmiştir.



Şekil 4.1. Video yoğunluk ayarlama sistemi.

Geliştirilen sistemin çeşitli avantajları ve getirdiği yenilikler vardır. İlk olarak, geliştirilen sistem güvenlik operatörünü dijital güvenlik gözetleme sistemi döngüsü

içine yerleştirmektedir. Operatörden alınan anlık geri besleme çevrimdışı dikkat modeli kullanan sistemlerin birçok sorununa çözüm olmaktadır. İnsan operatörlerin güvenlik gözetleme sistemlerinin kritik elemanı olduğu göz önüne alındığında, operatörün döngüye yerleştirilmesinin sistemin genel verimliliği ve kararlılığını önemli ölçüde arttıracığı açıktır. Gerçekleştirilen sistem, uyarlamalı otomasyon tabanlı sistemlerin bir örneğidir [Parasuraman, 2007]. İş yükünün ayarlanması ile operatörlerde yorgunluğun ve stresin azaltılması sağlanacaktır [Byrne and Parasuraman, 1996]. İkinci önemli yenilik ise problemin tanımıyla ilgilidir. Şimdiye kadar olan tüm yöntemler operatörlerin verimliliklerini arttırmak için onlara mümkün olduğunca yoğun özet görüntüleri göstermekteydi. Bu yöntemler operatörlerin verimliliklerini arttırırken sistemin kararlılığını düşürebilmekteydi. Video yoğunluk ayarlama yöntemi özetlemeden çok daha fazlasıdır. Dikkat seviyesi yüksek olan operatörler seviyelerinin yüksekliğiyle orantılı yoğunlukta özetler izlerken (Bkz. Şekil 4.1(a)), dikkat seviyesi düşük olan operatörler gerçek videodan daha uzun hale gelmiş bir çıktıyı izleyebilirler (Bkz. Şekil 4.1(b)). Operatörlerin iş yüklerini sürekli arttıran ve problemi bir sırt çantası problemi (İng. Knapsack problem) olarak gören mevcut yöntemlerin aksine geliştirilen sistem problemi tesis yerleşim problemi olarak tanımlamaktadır. Geliştirilen yöntemde her bir hareketli nesne operatörün dikkat seviyesi de hesaba katılarak video içerisinde en anlaşılabilir konuma yerleştirilmektedir. Tesis yerleştirme problemi NP-Hard bir problemidir. Güvenlik gözetleme sistemlerinde kullanılacak bir algoritmanın hızlı olarak çalışabilmesi önemlidir. Üçüncü yenilik olarak ise video yoğunluk ayarlama probleminin gerçek zamanlı olarak çözülmesini sağlayacak bir meta-sezgisel algoritmanın uygulanmasıdır.

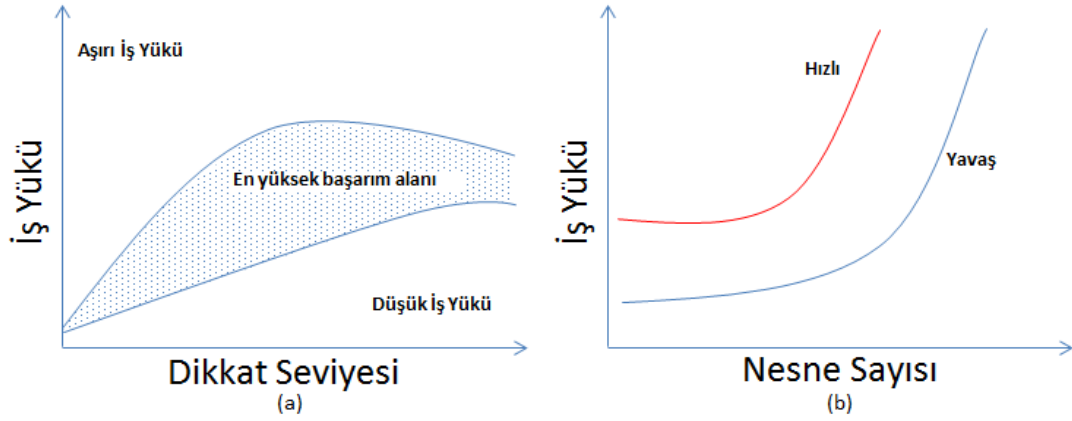
Gerçek ve sentetik videolar üzerinde sistem deneyleri yapılmıştır. Sentetik test ortamında yapılan çalışmalarda insanların dikkat seviyelerinin iş yüküyle nasıl değiştiği incelenmiştir. Bu deneyde farklı iş yüklerinde çalışan kişilerin başarıları değerlendirilmiştir. Geliştirilen sistem daha sonra gerçek videolar üzerinde farklı parametreler ve operatör dikkat seviyeleri için çalıştırılmış ve sistemin çıktıları gösterilmiştir.

4.1. Çoklu Nesne Takibinde İnsan Psikolojisi

Gözetleme görevlerinin birçoğu bağımsız şekilde hareket eden çok sayıda nesnenin gözle takibine dayanır. Bu nedenle, güvenlik operatörlerinin dikkat seviyelerini ölçebilmek için öncelikle insanların birden çok hareketi nasıl takip ettiğinin anlaşılması gereklidir. Psikoloji kaynakları bu konuda çok sayıda yayın içermektedir ve bu geniş bilimsel zemin sistemin temellerinin anlaşılması için oldukça faydalıdır [Fehd and Seiffert, 2010; Ma and Huang, 2009; Sears and Pylyshyn, 2000].

Psikolojik çalışmalar çoklu nesne takibini dikkat, takip stratejileri ve sınırları, yorgunluk ve iş yükü başlıkları altında incelemiştir. İnsanlar bu tür işlerde çok başarılı olmamaktadır çünkü insanın görsel ve algısal kapasitesi sınırlıdır. Bir çok çalışma göstermiştir ki insanlar belli bir sahnede hareket eden nesnelerin çok küçük bir alt kümesini dikkatli bir şekilde takip edebilmektedirler [Franconeri et al., 2007; Pylyshyn and Storm, 1988] (Bkz. Şekil 4.2(a)). Bazı çalışmalar insanın çoklu nesne takibi için bölünmüş dikkat kullandığını ve dört ila altı nesneyi eşzamanlı olarak takip edebildiğini belirtmektedir [Alvarez and Franconeri, 2007; Cavanagh and Alvarez, 2005]. Bununla birlikte insanların bir zamanda adanmış dikkat ile sadece tek bir nesneyi takip edebildiği bilinmektedir. İnsanların takip yeteneği kişinin öznel durumuna bağlı olduğu gibi nesnelerin hızlarına ve dağılımlarına da bağlıdır. Eğer nesneler sanal bir grup oluşturmuşçasına hareket ediyorsa grup içindeki nesnelerin tamamı tek bir nesne gibi takip edilebilirler [Scholl et al., 2001]. Bunun için insanlar sanal nesnenin merkezine odaklanarak takip kapasitelerini en verimli şekilde kullanırlar [Fehd and Seiffert, 2010]. Bunların dışında insanlar yavaş hareket eden nesneleri daha çok sayıda ve rahat takip edebilirler (Bkz. Şekil 4.2(b)). Yavaş nesnelerin takibinde göz yumuşak takip gerçekleştirir ve bu şekilde izlemenin doğruluğu sıçramalara oranla çok daha yüksektir [Alvarez and Franconeri, 2007]. Yumuşak izlemenin sıçramalarla izlemeye oranla daha az yorucu olması da uzun süreli çalışmalarda yorgunluğa bağlı performans düşüşünü azaltır [Horowitz et al., 2004]. Nesnelerin sahne içindeki konumları da takip edilebilirliği belirleyen bir

başka etmemidir. Birbirlerini kapatan ve sürekli olarak görüntü üzerinde çakışan hareketlerin takip edilmesi zordur [Franconeri et al., 2010].



Şekil 4.2. Nesne sayısı, iş yükü ve dikkat seviyesi ilişkisi.

Dikkat seviyesi ve süreleri öznel ve kişinin duygusal ve zihinsel durumuna [Gu et al., 2005], uzmanlığına [Allen et al., 2004] ve iş yükü algısına göre değişir [Warm et al., 1996]. Uzman güvenlik operatörleri diğer kişilere göre daha çok sayıda nesneyi takip edebilirler ve dikkatleri daha uzun süre yüksek seviyede seyreder. İş yükünün az olduğunu düşünen operatör işten sıkılabilir [Thiffault and Bergeron, 2003] ve çevredeki diğer uyarıcılarla ilgilenebilir. Bununla birlikte iş yükünün fazla olduğu durumlarda ise operatörler belli bir süre sonra yorulacak ve dikkat seviyeleri düşecektir [Wolfe et al., 2007]. Operatörlerin dikkat seviyelerini dış uyarıcılar kullanarak yükselten sistemler mevcuttur [Lavine et al., 2002; İnce ve Göktürk, 2009]. Bununla birlikte; operatörlerin dikkat seviyelerini yüksek tutmak için uzun süre zorlanmaları stres ve yorgunluk kaynağıdır ve belli bir süreden sonra dramatik dikkat kayıplarına neden olmaktadır.

Geliştirilen yöntem yukarıda değinilen psikoloji teorilerini temel almaktadır. Operatörün dikkat seviyesi göz-bebeği metrikleri kullanılarak hesaplanmaktadır ve sentezlenen görüntülerle operatörler her zaman en iyi performans alanında tutulmaya çalışılmaktadır (Bkz. Şekil 4.2(a)).

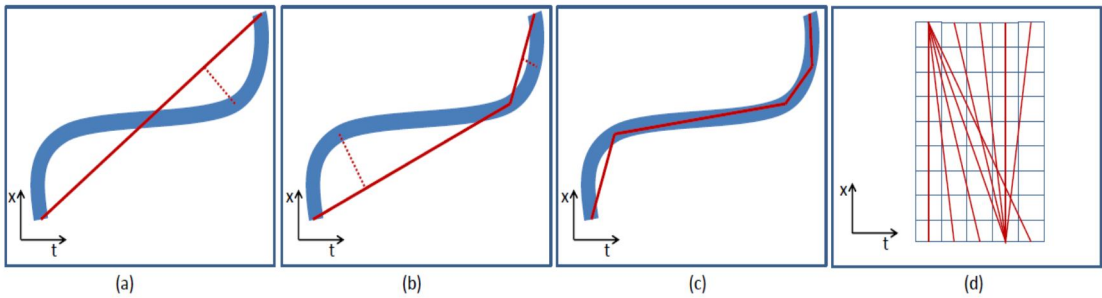
4.2. Video Yoğunluğu Ayarlama Sistemi

Video yoğunluk ayarlama sistemi için öncelikle hareketli nesnelerin videodan çıkartılması ve bu hareketlere ait üç boyutlu rotaların bulunması gerekmektedir. Geliştirilen sistem herhangi bir üç boyutlu video takip algoritmasıyla çalışabilecek durumdadır. Takip algoritmasının başarısı görsel bozulmaların en aza indirgenmesi için önemlidir. Her bir hareketin takibi ve ayrıştırılmasıyla hareketler bir kuyrukta beklemeye başlarlar. Enerji fonksiyonu minimum yapacak şekilde hareketler bu kuyruktan çekilerek sentezlenmiş video uzayına yerleştirilirler. Video sentezleme problemi NP-hard bir problem olan tesis yerleşim problemi [Liu, 2009] olarak tanımlandığı için arama uzayını azaltacak bir doğru modeli geliştirilmiş ve kullanılmıştır. Hareketlerin doğrular ile modellenmesi arama uzayını azalttığı gibi hareketlerin hızlarının ve yönlerinin değiştirilmesinde de hızlı hesaplamayı mümkün kılmaktadır. Doğru modeli ile atama yapılabilecek konumlar önceden bilindiği için tüm çarpışma durumları önceden hesaplanabilmekte ve böylece çalışma zamanına yük binmemektedir. Doğru modelinin kullanılmasının getireceği bir kısıt ise üç boyutlu hareket rotalarının y bilgisinin düşürülmesidir. Doğru modelinin daha detaylı incelemesi ve problemin tanımı bir sonraki alt bölümde verilecektir. Sistemin enerji fonksiyonu ve bu fonksiyonun yinelemeli yerel arama [Stutzle, 2006] ile çözülmesi ise takip eden alt bölümlerde anlatılacaktır.

4.2.1. Problem Tanımı

V_{Girdi} girdi video uzayı içindeki her bir a hareketini $V_{Çıktı}$ sentez videosu üzerindeki L aday konumlardan birine yerleştirilebilecek bir tesis olarak düşünülmektedir. Tüm hareketlerin bir kümesi olan A , girdi videosu V_{Girdi} üzerinde blob tabanlı bir hareket dilimleme ve izleme algoritması kullanılarak elde edilir. Video yoğunluk ayarlama sistemi A kümesi içindeki her bir a eylemine L konum kümesindeki bir l konum etiketi atar. Tüm A kümesi içindeki hareketlerin etiketleri bir C düzenlenişi oluşturur. Yoğunluk ayarlama sistemi, E_{Toplam} toplam enerjisini en az yapan C^* düzenlenişini bularak, çıktı videosu $V_{Çıktı}$ 'yı üretir.

Toplam enerji E_{Toplam} , veri enerjisi E_{Veri} , çarpışma enerjisi $E_{Çarpışma}$, ve çerçeve yoğunluk enerjisi $E_{Yoğunluk}$ değerlerinin doğrusal birleşimidir. E_{Veri} veri enerjisi diğer tesislerden bağımsız olarak bir tesisin l konum etiketi alma maliyetini verir. Her tesis kendi kazancını maksimize edecek bir konumu arar. Video yoğunluğu ayarlama probleminde her bir hareket bir tesisi göstermektedir ve her bir hareket kendi asıl değerlerini korurken en iyi izlenebilir konuma yerleşmeyi amaçlar. Tesis yerleşim probleminde her bir tesis ayrıca diğer tesislerin çıktılarına ihtiyaç duyar ve onlara yakın olarak ihtiyaç maliyetlerini azaltmayı amaçlar. Video yoğunluğu ayarlama probleminde hareketler arasında ayrı ayrı ilişki yoktur ve buradaki ilişki tesis yerleşim problemindeki tersi gibi düşünülebilir. Tesis yerleşim probleminde yakın olmak genelde iyiye, video yoğunluğu ayarlama hareketlerin birbirlerini kesmemeleri tercih edilmektedir. $E_{Çarpışma}$ enerji terimi farklı hareketlerin birbirini kesen l konumlarına yerleştirilmesini cezalandırmak için kullanılmaktadır.



Şekil 4.3. Doğru dilimi modeli ve Douglas-Peucker algoritmasının ilk üç adımı.

Tesis yerleşim problemi NP-hard bir problemdir bu nedenle arama uzayının daraltılması verimli bir algoritmanın geliştirilebilmesi için önemlidir. Hareketler tüm olası piksel konumları yerine daha sınırlı sayıdaki aday doğru dilimleri üzerine yerleştirilecektir. Doğru dilimleri modeli çoklu nesne takibi için geçerli psikolojik önermeleri hızlı bir şekilde mevcut hareketlere uygulamaya da izin verecektir. Doğru modelinin uygulanabilmesi için öncelikle hareketlerin iki boyutlu izdüşümleri hesaplanır. Hesaplanan izdüşümleri genellikle bir doğru vermezler. Bu izdüşümleri doğru parçacıklarıyla yakınsanabilir. Bunun için Douglas-Peucker doğru yakınsama algoritması [Douglas and Peucker, 1973] kullanılmıştır (Bkz. Şekil 4.3(a,b,c)). Her bir hareket üç adet doğru parçacığıyla ifade edilmiştir. Eğer doğru parçacıklarının

eğimlerinin işaretleri farklıysa burada hareketin sürekliliği içinde bir geri dönüş olduğu anlaşılır ve hareket ikiye bölünerek A kümesine işlenir. Bu işlem sonunda bir a hareketini ifade eden tüm doğru parçacıklarının eğimlerinin işaretlerinin aynı olması garanti edilir. Bir a hareketine ait doğru parçacığı eğiminin işareti farklı bir l konumuna atanırsa hareketin yönü değişmiş olur. Bu tür değişimlerle sanal hareket grupları yaratılabilir. Bunun dışında bir hareketin hızının değiştirilmesi de yine eğimin büyüklüğüyle oynanarak yapılabilir. Bir a hareketi kendi eğim değerinden daha küçük eğimli bir l konumuna atanarak yavaşlatılabilir ya da daha yüksek bir eğim değerine sahip l konumuna atanarak hızlandırılır. Hareketlerin çarpışma durumlarının incelenmesi oldukça fazla zaman gerektirir. Doğru parçacığı modeli kullanılması bu süreyi oldukça iyi bir şekilde azaltır. $V_{çıktı}$ video uzayı için tüm olası doğru konumları önceden bulunur ve bu parçacıkların birbirlerini kesip kesmediği bir matris içine yazılır (Bkz. Şekil 4.3(d)). Bu kesişim matrisin çalışma zamanında kullanılması oldukça hızlıdır.

4.2.2. Enerji Fonksiyonu

Verilen bir C düzenlenişinin, A hareket kümesinin, $V_{çıktı}$ çıktı videosunun ve DS dikkat seviyesinin toplam enerjisi;

$$E_{Toplam}(A, V_{çıktı}, C, DS) = \alpha_1 \sum_a E_{Veri}(A[a], C[a]) + \alpha_2 \sum_{a \neq b} E_{Çarpışma}(C[a], C[b]) + \alpha_3 \sum_a E_{Yoğunluk}(V_{çıktı}, C, DS), \quad (4.1)$$

olarak ifade edilir. Burada $a, b \in A$ ve $c[a] \in L$ 'dir.

Verilen bir l konum etiketi için a hareketinin veri enerjisi

$$E_{Veri}(a, l) = \beta_1 E_{Yön}(a, l) + \beta_2 E_{Hız}(a, l) + \beta_3 E_{Görünürlük}(l) + \beta_4 E_{Kronoloji}(a, l), \quad (4.2)$$

şeklinde hesaplanır.

Bir hareketin yönü değiştirilerek hareketlerin sanal bir nesne oluşturması sağlanabilir. Sanal nesne gruplarının izlenmesi daha kolaydır ama geri doğru yürüyen

insanlar gibi alışılmadık görüntülerin oluşmasına neden olabilir. Hareketin asıl yönünün korunması için bir maliyet yansıtılacaktır. Hareketin yönünde bir değişiklik olup olmadığı Denklem.3 kullanılarak hesaplanmaktadır. $a.doğru[0].eğim$, a hareketine uydurulmuş ilk doğru parçasığının eğimi olduğunda

$$E_{Yön} = \begin{cases} 1, & \text{eğer } (a.doğru[0].eğim * l.eğim < 0), \\ 0, & \text{diğer,} \end{cases} \quad (4.3)$$

yön enerjisidir. Hareketlerin yönleri değiştirilerek daha az çarpışmayla daha kısa videolar elde edilebilir.

Takip edilebilirliği belirleyen bir diğer önemli faktör ise hızdır. Geliştirilen sistem nesnelerin hızlarını ayarlayabilmektedir. Yüksek dikkat seviyesindeki bir operatör için nesnelerin hızları arttırılarak daha yoğun bir görüntü elde edilebilmektedir. Dikkat seviyesi düşük bir operatör için nesnenin hızı gerçek hızının altına indirilebilmektedir. Hız terimi

$$E_{Hız} = \sum_s a.doğru[s].oran \left| |a.doğru[s].son - a.doğru[s].baş| - |Zaman(a.doğru[s].son, l) - Zaman(a.doğru[s].baş, l)| \right|$$

olarak hesaplanır ve aşağıdaki şartlar sağlanmaktadır: $\sum_s a.doğru[s].oran = 1$ ve $Zaman(t, l) = (t - l.b) / l.m$.

Görünürlük terimi çok kısa bir zaman içinde çok hızlı harekete izin veren aday doğru konumlarını cezalandırır. Bu tür konumlar çıktı videonun kısa olmasını sağlarlar fakat hareketlerin takibini oldukça zorlaştırırlar. Görünürlük cezaları bir kez hesaplanmakta ve bir tüm optimizasyon süresince kullanılabilir. Hareketlerin kronolojik sıralarının değişmesine izin verilmektedir fakat bir hareketin çok uzun süre kuyrukta kalmaması için bu terim gereklidir. Ayrıca bu terim peş peşe gelen hareketlerin aynı sentez video olma ihtimalini de arttırmaktadır.

$$E_{görünürlük}(l) = \begin{cases} 2^{T_1 - (l.son - l.baş)} & \text{eğer } (l.son - l.baş) \leq T_1, \\ 2^{T_1} & \text{diğer.} \end{cases} \quad (4.4)$$

Hareketlerin kronolojik sıralarının değişmesine izin verilmektedir fakat bir hareketin çok uzun süre kuyrukta kalmaması için bu terim gereklidir. Ayrıca bu terim peş peşe gelen hareketlerin aynı sentez video olma ihtimalini de arttırmaktadır.

$$E_{kronoloji}(a, l) = |a.baş * (UZUNLUK(V_{çıkı}) / UZUNLUK(V_{girdi})) - l.baş| \quad (4.5)$$

Çarpışmalar ve kapanmalar çoklu hareket takibini zorlaştıran önemli etmenlerdir. Bu durumların önlenmesi ve görsel bozulmaların azaltılması için kesişen aday doğrulara atama yapılması istenilmez. Çarpışma cezalarının önceden hesaplanarak bir matris içinde tutulması optimizasyon sırasındaki zaman ihtiyacını önemli ölçüde azaltmaktadır. Eşitlik 4.6'daki ceza değerleri sistem tarafından bir kez hesaplanmaktadır.

$$E_{Çarpışma}(k, l) = \begin{cases} 0 & \text{eğer}((l.baş \leq k.baş) \text{ ve } (l.son \leq k.son)) \\ 0 & \text{eğer}((l.baş \leq k.son) \text{ ve } (l.son \leq k.son)) \\ 10 & \text{eğer}(|l.baş - k.baş| \leq T_2 \text{ ve } |l.son - k.son| \leq T_2) \\ 5 & \text{eğer}(|l.baş - k.baş| \leq T_2 \text{ veya } |l.son - k.son| \leq T_2) \\ 1 & \text{diğer,} \end{cases} \quad (4.6)$$

Her bir çerçevedeki hareket sayısı bir H sıklık grafiğinin üzerinde tutulmaktadır. Bu sıklık grafiğindeki değerlerin, operatörün takip edebileceği nesne sayısından uzaklığı bir enerji terimi olarak kullanılmaktadır.

$$E_{Yoğunluk}(a, l) = \sum_{f=0}^{UZUNLUK(V_{çıkı})} |H(f) - DS| \quad (4.7)$$

4.2.3. Eniyileme

Verilen bir dikkat seviyesi için video yoğunluğunun ayarlanması karmaşık bir optimizasyon problemidir. Minimum enerjiye sahip C^* düzenlenişinin verimli bir şekilde hesaplanması gereklidir.

$$C^* = \underset{C}{argmin} (E_{Toplam}(A, V_{çıkı}, C, DS))$$

Minimum enerjili düzenlenişi bulmak için yinelemeli yerel arama kullanılmıştır. Yinelemeli yerel arama oldukça basit yapısına rağmen güçlü bir meta sezgiseldir ve video yoğunluğunun ayarlanması için yeterince hızlıdır. Bu meta-sezgiselin üç önemli kısmı vardır:

1 - İlkendirme: İlk düzenlenişin oluşturulması için her bir a hareketine rasgele bir l etiketi atanabilir. Bununla birlikte geliştirilen sistemde hırslı bir algoritma kullanarak sadece E_{veri} enerjisini optimize eden bir düzenlenişle başlamak tercih edilmiştir.

2 - Yerel Arama: Bu aşamada her bir a hareketi tek tek aktif hareket olarak seçilmiş ve boşta bulunan tüm l konumları denenmiştir. Toplam enerjiyi düşüren bir konum bulunduğunda o konumun etiketi alınmıştır.

3 - Düzenleme: Bu aşamada C düzenleniş içindeki bir miktar hareket rasgele seçilmiş ve etiketleri rasgele değiştirilmiştir. Bu aşama, eniyilemenin yerel en iyilere takılmasını engellemeyi amaçlar.

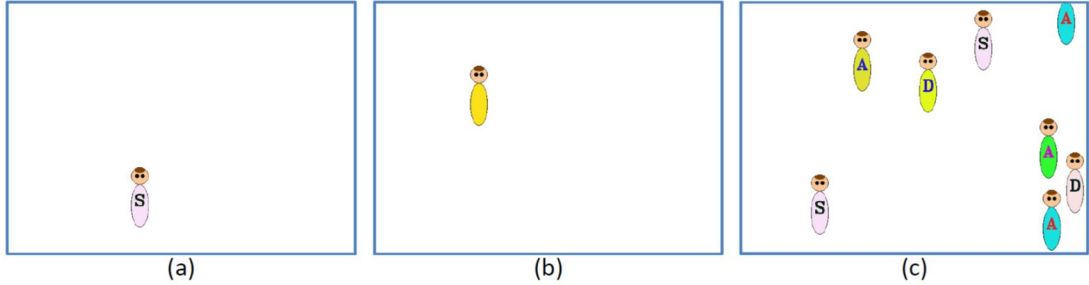
Geliştirilen sistemde yerel arama ve düzenleme adımları bir kaç kez tekrarlanmıştır. A hareket kümesindeki hareketler DS dikkat seviyesine uygun sayıda eleman içerecek şekilde rasgele gruplanmıştır.

4.3. Deneyler

İlk deney, çoklu nesne takibinin farklı iş yükleri altında ne kadar başarılı olduğunu ölçmektedir. Bu deney için sentetik bir test ortamı geliştirilmiştir.

Test yazılımı verilen iş yüküne uygun olarak insana benzeyen hareketli grafik nesneler üretmekte ve bunları kendi belirlediği hızlarda ekranda oynatmaktadır. Hareketli test nesnelerinin üzerinde 'A', 'S' ve 'D' etiketleri bulunmaktadır (Bkz. Şekil 4.4(a)). Bu etiketler rasgele zamanlarda kısa süre sönerek alarm vermektedir (Bkz. Şekil 4.4(b)). Operatörlerin bu süre zarfında doğru etikete basarak alarmı gördüğünü belirtmesi gerekmektedir. Operatör alarmı görüp etiketi hatırlamaması halinde ise boşluk tuşuna basmaktadır. Bu durumda operatörün alarmı fark ettiği ama nesneyi yeterince iyi takip edemediği, kısa süreli hafızasının yeterli olmadığı anlaşılmaktadır. Operatörün zamanında ve doğru olarak belirttiği alarmları "YAKALANDI", zamanında hiç bir tuşa basılmayanları "KAÇIRILDI" ve boşluk tuşuna basılanlarla

birlikte yakalanan alarmların toplamını ise "İZLENDİ" olarak puanlandırılmaktadır. Operatörlerin hedefi en çok sayıda alarmı "YAKALANDI" olarak puanlandırmaktır.



Şekil 4.4. Sentetik test arayüzünden örnek çerçeveler.

Operatörlere test ortamındaki görüntüler izletilirken, operatörlerin göz bebeği ölçütleri ve göz bakış konumları kafaya monte edilmiş bir göz-bakışı takip cihazıyla kaydedilmiştir. Deneylerde Arrington Research firmasına ait bir göz takip cihazı kullanılmıştır [Arrington, 1997]. Sentetik deneyler iki temel kısma ayrılmıştır. İlk kısımda iş yoğunlukları önceden ayarlanmış üç senaryo vardır. Bu senaryolar düşük iş yükü (Bkz. Şekil 4.4(a)), aşırı iş yükü (Bkz. Şekil 4.4(c)) ve karışık iş yükü altında deneklerin başarımlarını ölçmektedir. Her bir deney toplam 12000 çerçeveden oluşmaktadır. Zamana bağlı olarak başarımın nasıl değiştiğini ölçmek için bu 12000 çerçeve ilk 3000 çerçevenin 4 kez tekrarıyla üretilmiştir. Çizelge 4.1'de düşük iş yükü altındaki deneklerin açık bir şekilde diğer gruplara göre başarılı olduğu görülmektedir. Düşük iş yükü grubunun ortalama alarm yakalama oranı %85'dir ve bu karışık iş yükü grubunun en iyisinden %30 daha iyidir. Karışık iş yükü grubu altındaki deneklerin kaçırdığı alarmların %95'i aşırı yüklendikleri bölümlerdir. Bununla birlikte operatörleri sürekli düşük iş yoğunluğu altında tutmak sitem verimliliği için kötüdür. Karışık iş yükü altındakiler düşük iş yoğunluğundaki deneklere oranla 4 kat fazla hareket izlemişlerdir, aşırı iş yükü altındaki denekler ise düşük iş yükündekilere oranla aynı süre içinde 7 kat daha fazla hareketli nesneyi izlemişlerdir.

Yapılan bu deneyin göz ölçütleri ve görev puanı ile ilgili ilgi çekici sonuçları olmuştur. Göz kapama sıklığı fazla olan deneklerin diğerlerine oranla daha fazla alarmı kaçırdığı belirlenmiştir. Ayrıca deneklerin iş yükleri arttıkça sıçrama

hareketlerinin belirgin şekilde arttığı da gözlenmiştir. Bu nedenle sıçrama hareketinin sıklığı iş yükünün belirlenmesi için bir sonraki deneyde tercih edilmiştir.

Sentetik deneyin ikinci kısmında denekler, iş yüklerinin onların takip yeteneğine göre değiştiği uyarlamalı iş yükü altında çalışmışlardır. Bu denekler önce az sayıda nesneyi izlemeye başlamışlardır ve hareketli nesne sayısı her 30 saniyede bir arttırılmıştır. Sıçrama oranı bir önceki 30 saniyelik periyoda göre %10 artığında ya da göz kapama sıklığı %5 artığında bir nesne sistemden çıkartılmıştır. Ortalama sonuçlar Çizelge 4.1'de verilmiştir.

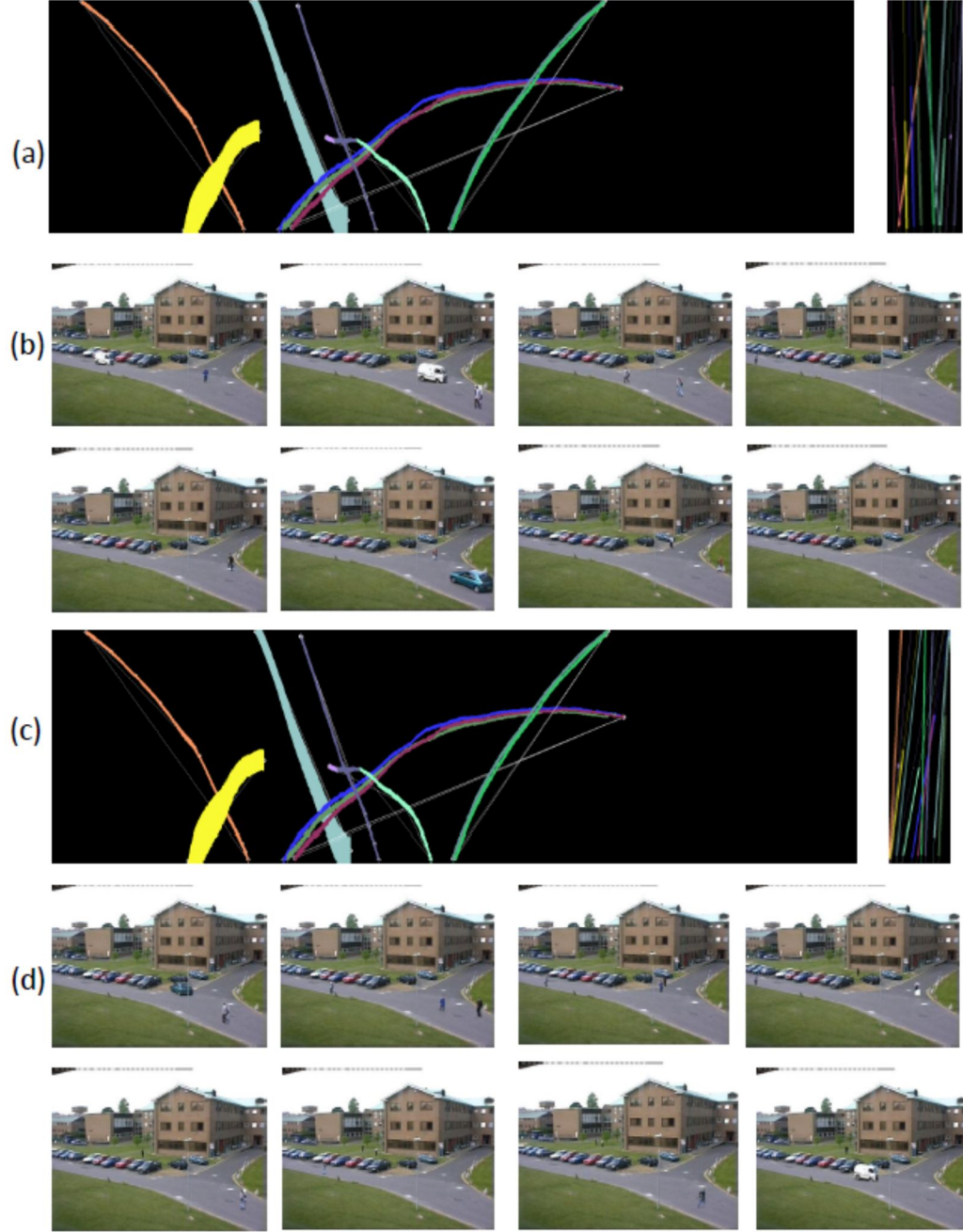
Çizelge 4.1. Psikolojik deneylerin sayısal sonuçları

	Düşük İş Yükü	Aşırı İş Yükü	Karışık İş Yükü	Uyarlamalı İş Yükü
Katılımcı Sayısı	3	3	3	5
Çerçeve Başına Ortalama Nesne	0,76	5,26	3,01	3,34
Ortalama Alarm Sayısı	40	101	74	87
Yakalama (%)	86,7	49,8	56,8	83,9
Yanlış Alarm (%)	4,17	5,28	11,2	4,75
Kaçırılan Alarm (%)	1,67	20,7	18,4	2,30
İzlenen + Yakalanan Alarm (%)	98,3	79,3	81,6	97,7
1000 Hareket Başına Kaçan Alarm	0,07	0,33	0,37	0,04

Uyarlamalı iş yükü altında çalışan deneklerin performansları diğer gruplarla karşılaştırıldığında oldukça iyi görünmektedir. Her bin hareketli nesne için alarm kaçırma oranı düşük iş yükü altındakilerin neredeyse yarısı kadardır, yüksek iş yükü altındakilerin ise onda birine yakındır. Bu başarısının yanında uyarlamalı iş yükü altında çalışan kişiler düşük iş yükü altındakilerin neredeyse 5 katı fazla hareketli nesneyi izlemişlerdir. Uyarlamalı iş yükü altındaki çalışanların "YAKALANDI" olarak etiketleme başarısı ise çok hafif bir şekilde düşük iş yükü altındakilerden kötüdür. Bunun nedeni neredeyse 5 kat fazla iş yükü altında kısa süreli hafızanın yeterli şekilde eniyilenememesidir.

Sistem ayrıca PETS [PETS, 2001] veri kümesindeki gerçek videolar üzerinde denenmiş ve farklı enerji parametreleri ve iş yükleri için çıktılar üretilmiştir. Çerçeve başına 2 ve çerçeve başına 3 hareketli nesne olacak şekilde optimize edilmiş görüntülere ait örnek çerçeveler Şekil 4.5'de gösterilmiştir. Hareketlere ait ilk ve eniyileme sonrası rotalarda şekilde verilmiştir. Geliştirilen eniyileme sezgiselimiz

teorik olarak $O(a.l)$ zaman karmaşıklığındadır ve standart bir Pentium-4 işlemcili bilgisayarda bir çerçevenin işlenmesi 5 ms sürmektedir.



Şekil 4.5. Önerilen sistemin PETS veri kümesindeki görsel sonuçları.

4.4. Yorumlar

Yapılan çalışma ile insan operatörler video yoğunluk ayarlama sistemi döngüsüne yerleştirilmiştir. Geliştirilen sistem operatörlerin algı seviyelerini ölçmekte ve video yoğunluğunu operatörün o anki kapasitesine göre belirleyebilmektedir. Bu sistemin kullanılmasıyla operatörler en yüksek başarılı takip bölgesinde tutulmakta, sistemin verimliliği ve kararlılığı arttırılmaktadır.

Geliştirilen sistem video yoğunluğu ayarlama problemini tesis yerleşim problemi olarak tanımlamaktadır. Bu yönüyle elde olan ayrıştırılmış hareketleri çıktı video uzayı içerisinde en iyi konumlara yerleştirmeyi amaçlamaktadır. Sistem, video yoğunluk ayarlama problemini meta-sezgisel kullanarak hızlı bir şekilde gerçekleştirmektedir.

Sistem, psikolojik çalışmalarca çoklu hareket takibi için gösterilen birçok etmeni modellemektedir. Hareketli nesnelerin hızları, yönleri, sayıları ve sahne üzerindeki dağılımları belirlenebilmektedir.

Geliştirilen sistem Bölüm 2’de ve Bölüm 3’de açıklanan göz-bakışı tabanlı sistemlerin iki temel problemine çözüm getirmiştir. Öncelikli olarak bu bölümlerde açıklanan yöntemlerin operatörlerin baktıkları yerlerdeki hareketleri gördükleri kabulüne ihtiyaç ortadan kaldırılmış ve daha doğru bir şekilde videodaki hareketlerin görülüp görülmediğinin analizi yapılabilir olmuştur. Bu bölümde açıklanan yöntem, doğrusal olmayan özetleme yöntemlerinin ürettiği ve genelde rahat bir şekilde operatörler tarafından takip edilemeyecek derece yoğun videoların üretilmesi probleminin de üstesinden gelmiştir.

Bölüm 5’de yapılacak çalışmalarla güvenlik operatörlerinin ilgi duydukları hareket ve nesnelere ait özniteliklerinin çıkartılması ve ilgi duyulan eylemlerin operatörlere öncelikli olarak gösterilmesi için geliştirmiş olduğumuz bir yöntem açıklanacaktır.

5. OPERATÖR İLGİSİ TABANLI GÜVENLİK VIDEO SENTEZİ

Dijital güvenlik gözetleme sistemleri ucuzlayan kurulum maliyetleriyle birlikte oldukça popülerleşmiştir [Sasse, 2010]. Birçok durumda müşteriler daha iyi bir kapsama alanı için daha çok güvenlik kamerası talep etmektedir. Kapsama alanının geniş tutulması arzulanan bir durum olsa da kamera sayısının artmasıyla görsel veri miktarı da artmaktadır. Kameralardan gelen bu verilerin şüpheli nesneler ya da hareketler için incelenmesi gerekmektedir. Bu verilerin otomatik olarak incelenebilmesi için çok sayıda yöntem geliştirilmiştir [Yu et al., 2011]. Tamamen otomatik sistemler oldukça hızlıdır [Antic and Ommer, 2011] fakat sınırlı sayıda nesne için çalışabilirler ve başarımları istenilen seviyelerin çok altındadır [Keval and Sasse, 2006]. İnsan operatörler kritik durumlarda gelişmiş güvenlik gözetleme sistemlerinden genelde daha iyi kararlar verirler [Vural and Akgul, 2011b]. İnsan operatörlerin istihdam edilmesi pahalıdır ve gözle yapılan inceleme zor ve zam alıcıdır. Bu nedenle, insan operatörlerin çevrim içinde kullanıldığı bir sistem tasarlanarak operatörlerin verimlerinin artırılması gerekmektedir.

Video özetleme, operatörlerin verimliliklerini arttırmak için kullanılan yaygın bir yaklaşımdır [Li et al., 2001]. Özetleme yöntemleri video arşivi içindeki tüm görüntüleri olduğundan daha kısa bir zamanda gösterirler. Anahtar çerçeve tabanlı özetleme yöntemleri az hareket içeren çerçeveleri atarak çalışırlar. Bu tür yöntemler biraz zorlandığında hareket içeren çerçeveleri kolayca kaybederler. Doğrusal olmayan video özetleme yöntemleri [Acha et al., 2006] ise çerçeve atmak yerine hareketlerin video zaman pozisyonlarını değiştirirler. Bu yöntemler çerçeveleri atmadıkları için hareketleri kaybetmezler. Bununla birlikte doğrusal olmayan yöntemlerle elde edilen özetler çok sayıda hareket içerirler ve izlenebilir değillerdir.

İnsanların gözle takip etme yetenekleri oldukça sınırlıdır [Alvarez and Franconeri, 2007] ve kalabalık sahnelerdeki önemli olayları kolaylıkla gözden kaçırabilirler. Güvenlik gözetleme sistemlerinin kararlılığını arttırabilmek için insan psikoloji hesaba katılmalıdır çünkü bir operatör tarafından takip edilebilecek nesne sayısı operatörün deneyimine [Allen et al., 2004], yorgunluğuna [Franconeri et al.,

2007] ve iş yükü algısına [Warm et al., 1996] göre değişir. Yakın zamanda, göz-bakışı temelli metrikler güvenlik operatörlerinin başarımını ölçmek için kullanılmaya başlanmıştır. Bölüm 2’de ve Bölüm 3’de, operatörlerin monitör üzerindeki baktıkları yere göre gördükleri ya da gözden kaçırdıkları hareketleri tespit eden sistemler açıklanmıştır. Bu sistemler [Vural and Akgul, 2009; Vural and Akgul, 2011b], operatörlere gözden kaçırdıkları kısımların doğrusal olmayan bir özetini göstererek onlara ikinci bir şans verir. Bölüm 4’de açıklanmış olan bir diğer göz-bakışı tabanlı özetleme yöntemiye [Vural and Akgul, 2012b], özet videonun çok fazla sayıda hareket içermesi problemi üzerine eğilmiştir. Bu yöntem özet videonun yoğunluğunu operatörün algı seviyesine göre ayarlar. Geliştirilmiş olan göz-bakışı tabanlı bu yöntemler operatörleri destekler ve sistemin kararlılığını artırır. Göz-bakışı tabanlı ölçütler insanların ilgilerini ölçmek için de kullanılırlar. Özellikle internet siteleri ve vitrin tasarımı için insanların nerelere baktığı önemlidir. Bununla birlikte göz-bakışı ile güvenlik operatörlerinin ilgisi çevrim-içinde-insan güvenlik gözetleme sistemlerinde şuana kadar kullanılmamıştır.

Çalışmanın bu bölümünde, operatörün ilgisine bağlı olarak güvenlik videosu sentezleyen bir yöntem duyurulmuştur. Güvenlik operatörleri genellikle belli karakterdeki nesnelere, kişilere ve olaylara diğerlerinden daha fazla ilgi gösterirler. İlgi duyulan nesneler mevcut göreve ve operatörün motivasyonuna göre değişir. Bunların dışında operatörün ilgisi monitörde görülen bir olay ya da nesne sonrası değişebilir. Bu nedenle çevrimiçi bir öğrenme algoritması sistemin vazgeçilmez bir parçası olacaktır. Bu bölümde, güvenlik videosu sentezleme problemi bir çevrimiçi öğrenme problemi olarak tanımlanmıştır ve geliştirilen öğrenme şemasıyla videodaki hareketler ilgi çekici ya da ilgi çekici olmayan şekilde iki sınıfa ayrılmıştır. Her hareketin resim ve hareket tabanlı öznitelikleri çıkartılmıştır ve bir hareket belli bir süre operatör tarafından takip edilmişse bu öznitelikler ilgi çekici olarak belirlenmiştir. Videonun izlenmeyen kısımlarında bulunan hareketlerin daha önceden dinamik olarak öğrenilmiş öznitelikler kullanılarak ilgi çekici olup olmadığına karar verilir. İlgi çekici özniteliklere sahip hareketlerin videodaki kronolojisi değiştirilerek daha erken gösterilmesi sağlanır. Bu sayede insanların videoda arama yapmaları kolaylaşır. Benzer bir yöntem [Pritch et al., 2009] videodaki hareketleri çok sayıda

ilişkili kapsam kutuları bir nesne tüpü olarak ifade edilir ve her a_i eylemi için bir $t_i \in T$ tüpü vardır. Hareketli nesnelerin ve onların dikdörtgen kapsama kutularının bulunması için bir çeşit Gauss karışımı tabanlı ön plan/arka plan ayırma algoritması kullanılmıştır.

Eylemler kümesi A bulunduktan sonra, her bir hareket için alt seviye resim ve hareket öznitelikleri X çıkartılır (Bkz. Şekil 5.1(b)). Nesnelerin dokusunu göstermek için yerel ikili örüntü sıklık grafikleri kullanılmıştır. Yerel ikili örüntü tabanlı öznitelikler her bir kapsama kutusu için elde edilir ve sekiz-sepetli bir sıklık grafiği oluşturulur. Kapsama kutusu içindeki piksellerin KYM (Kırmızı-Yeşil-Mavi) renk uzayındaki her bir renk kanalı için sekizer sepetli sıklık grafikleri çıkartılır. Bu resim tabanlı özniteliklerin dışında kapsama kutularının boyutları, kutuların kenarlarının oranları ve merkezleri de öznitelik olarak kullanılmaktadır. Hareketin yönü ve hızı hareket özniteliklerini oluşturmaktadır. Hareket öznitelikleri her kapsama kutusu için değil hareketin tamamı için sadece bir kez oluşturulur. x_i öznitelik vektörünün içinde toplan 48 resim ve hareket tabanlı öznitelik yer alır.

İkili bir sınıflandırıcı ile ilginç hareketlerin öğrenilmesi için gentleboost [Friedman et al., 2000] tabanlı bir çevrimiçi boosting algoritması geliştirilmiştir. Geliştirilen ikili boosting algoritması D boyutlu x öznitelik vektörlerinin bir kümesi olan $X \in \mathbb{R}^D$ ve bu kümenin her elemanını bir sınıf etiketine eşleyen $Y = \{-1, 1\}$ kümesini girdi olarak alır. Çevrimiçi ikili boosting algoritması sürekli olarak eğitilir ve eğitim sırasında $\{x_i, y_i\}, i = 1..N$ olan bir veri seti kullanılır. Bu veri setinde N adet eğitim örneği bulunmaktadır ve $x_i \in X, y_i \in Y$ şartları sağlamaktadır. Eğitim setinde kullanılan bir a_i hareketinin y_i sınıf etiketi hareket ilginç ise +1 değil ise -1 değerlerinden birini alır. Eğer bir a_i hareketi operatör tarafından belirlenmiş bir eşik değerinden daha uzun süre gözle takip edilmişse o hareket ilginç olarak etiketlenmektedir (Bkz. Şekil 5.1(f)). Sistem göz-bakışı takip cihazından gelen konum bilgileriyle hareketlere ait kapsama kutularının örtüşmesine bakarak nesnenin takip edilip edilmediğine karar vermektedir. Burada gözden kaçırılmaması gereken bir konu da sistemin operatörden gelen geri beslemeler ve yeni hareketlerle sürekli olarak eğitildiğidir, böyle bir eğitim sadece çevrimiçi bir makine öğrenme

algoritmasıyla sağlanabilir (Bkz. Şekil 5.1(e)). Eğitim sonuçları C sınıflandırıcısı girdi videoda tespit edilen hareketlerin öznitelik vektörlerine uygulanarak hareketlerin sınıfları tespit edilir (Bkz. Şekil 5.1(c)). Henüz operatör tarafından izlenmemiş tüm a_i hareketleri ve bu hareketlere ait $y_i = C(x_i)$ sınıf etiketleri bir Q kuyruğunda tutulurlar. Her a_i hareketi için seçim önceliğini belirleyen $P(a_i)$ seçim önceliği fonksiyonu aşağıdaki şekilde hesaplanır:

$$P(a_i) = C(x_i) + \lambda R(a_i), \quad (5.1)$$

Burada λ ağırlıklandırma katsayısıdır ve R fonksiyonu bir hareketin bekleme süresiyle orantılı olarak artan rastgele bir sayı üretmektedir. Böylece hem ilginç hareketlerin hem de kuyrukta uzun süre bekleyen hareketlerin seçilme şansı artmaktadır (Bkz. Şekil 5.1(d)). Psikolojik çalışmalar insan operatörlerin en fazla dört bağımsız hareket eden nesneyi takip edebildiğini göstermektedir [Alvarez and Franconeri, 2007]. Bu nedenle sentezlenen videolarda en fazla dört nesnenin görünmesine izin verilecektir.

5.2. Deneyler

Sistem hem yapay hem de gerçek videolar kullanılarak test edilmiştir. İlk olarak, 120 dakika uzunluğunda yapay bir deney videosu üretilmiştir. Sistem kontrollü bir şekilde bu video üzerinde denenmiştir. Yapay video beş farklı desende ve altı farklı renkte toplam otuz çeşit nesneyi içermektedir. Bu nesneler rastgele yönlerde ve rastgele hızlarda hareket etmektedirler. Video içerisinde toplam 1200 farklı hareket bulunmaktadır. Şekil 5.3(a)'daki grafik her bir nesnenin birikimli adedini göstermektedir ve şekilden de görülebileceği gibi hepsi düzenli bir şekilde artmaktadır.

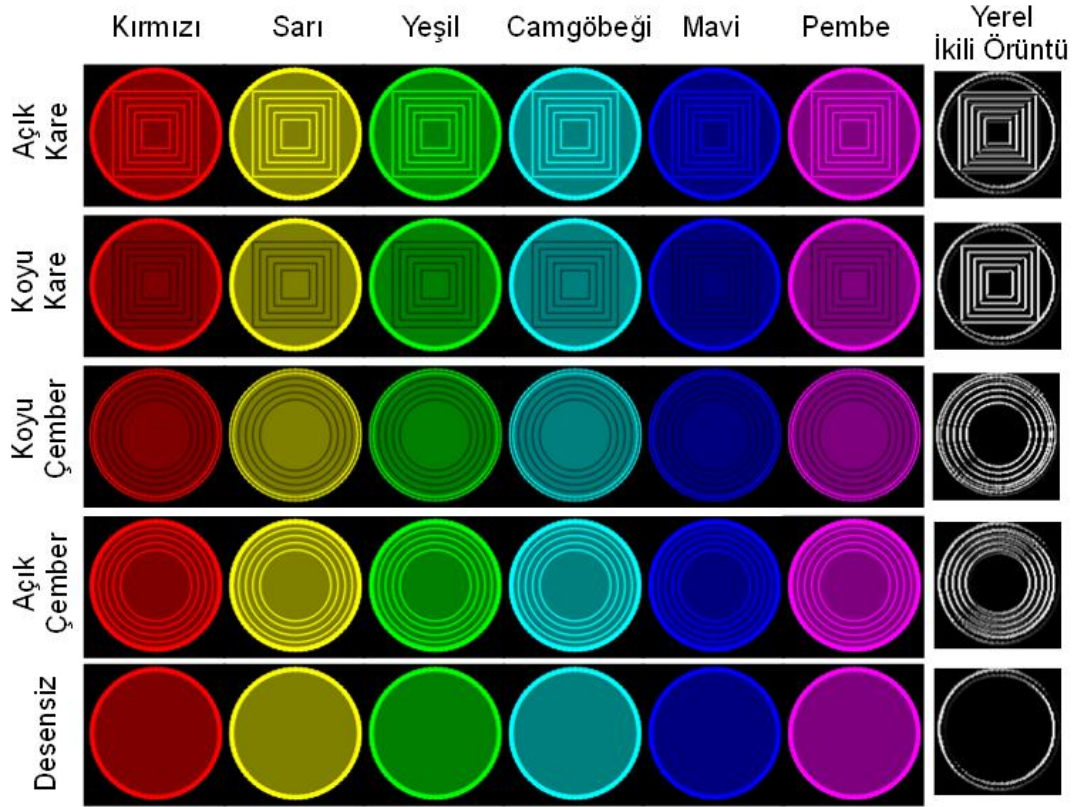
Bu yapay video, geliştirilen video özetleme sistemine gerçek bir güvenlik video girdisi gibi verilmiştir. Deneylerde Arrington Research [Arrington, 1997] firmasının ürettiği göz-bakışı takip cihazı operatörlerin göz-bakışı konumlarını elde etmek için kullanılmıştır. Geliştirilen sistem yapay video üzerinde beş farklı senaryo için denenmiştir. Tüm senaryolarda operatörlerden özel bir nesne tipini ve/veya

nesne tiplerini ilk on dakika boyunca takip etmesi istenmiştir, daha sonra bu nesne tipleri değiştirilerek operatörün yeni nesneleri on dakika boyunca izlemesi istenmiştir. Her bir test için nesnelerin özellikleri Çizelge 5.1'de verilmiştir.

Çizelge 5.1. Sentetik deneylerdeki ilgi duyulan nesne özelliklikleri.

	İlk 10 Dakika		İkinci 10 Dakika	
Deney	Renk	Desen	Renk	Desen
1	Kırmızı	Hepsi	Pembe	Koyu Çember
2	Kırmızı	Koyu Çember	Mavi	Koyu Çember
3	Mavi	Desen Yok	Hepsi	Açık Kare
4	Pembe	Açık Çember	Pembe	Desensiz, Koyu Kare
5	Sarı	Açık Kare	Sarı, Camgöbeği	Açık Çember

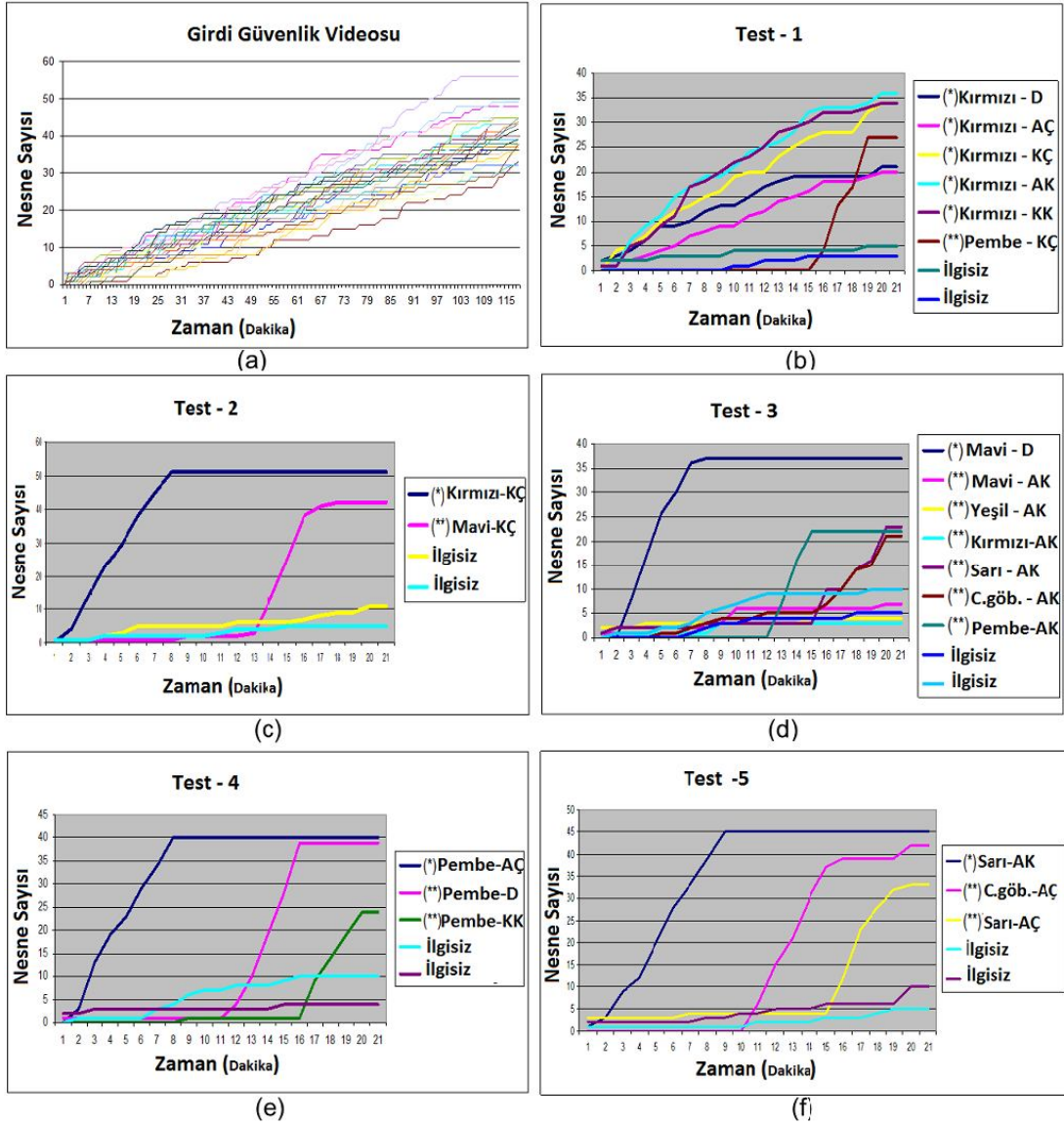
Her bir senaryoda operatörün izlediği nesnelerin birikimli toplamı kaydedilmiştir. Şekil 5.3(b)'de 1. senaryo için izlenen nesnelerin birikimli toplamının grafiği verilmiştir. İlk on dakika için ilgi çekici olan özellikler (*) ile ikinci on dakika için ilgi çekici olan özellikler (**) ile gösterilmiştir. İlk on dakika içinde tüm kırmızı renkli nesnelerin toplamı hızla artmaktadır çünkü operatörden kırmızı nesneleri takip etmesi istenilmiştir. Sistem operatör kırmızı renkli nesnelere baktıkça bu özneliği öğrenmiş ve bekleyen diğer kırmızı renkli nesneleri öne getirmiştir. İkinci on dakika boyunca ise operatörden pembe renkli ve koyu dairesel desenleri olan nesneleri takip etmesi istenilmiştir ve bu nesnenin toplam izlenilme sayıları ani şekilde artmıştır. İlginç olmayan tüm diğer nesne türlerine ait birikimli toplam nesne sayıları ise tüm deney süresince daha yavaş artmıştır. Şekillerin karmaşıklarının artmaması için sadece en hızlı artan iki ilgisiz nesne türü grafiklerde gösterilmiştir. Diğer senaryolar için de benzer sonuçlar elde edilmiştir (Bkz. Şekil 5.3(c, d, e, f)). İlk yirmi dakika içerisinde takip edilen nesnelerin hemen hemen tamamı operatöre gösterilmiştir. Bu durum grafiklerdeki artım miktarının doyumundan da anlaşılabilmektedir.



Şekil 5.2. Sentetik deneyde kullanılan nesne türleri ve YİÖ resimleri.

Üniversite yerleşkesinde çekilen gerçek video görüntüleri üzerinde de deneyler yapılmıştır. Orijinal girdi videosu 120 dakika uzunluğundadır. İlk deney için operatörden gri renkli arabaları takip etmesi istenmiştir. Sentezlenmiş videonun bazı örnek çerçeveleri Şekil 5.4'de gösterilmiştir. Orijinal girdi videosundaki ve sentezlenmiş videodaki çerçeve numaraları resimlerin sol üst köşesinde verilmiştir. Orijinal çerçeve numarası yüksek olan gri renkli arabalar sentezlenmiş videoda daha düşük çerçeve numaraları almışlardır yani bir başka deyişle daha erken gösterilmişlerdir. İkinci deneyde ise operatörlerin insanları takip etmeleri istenilmiştir. Her iki senaryonun erken çerçevelerinin aynı olduğuna dikkat edilmelidir. Bu durum, sistem henüz ilginç öznitelikleri öğrenemediği içindir. Bunun dışında hareket seçiminde erken olan olaylara belli bir öncelik verildiğinden (Bkz. Eşitlik 5.1), çok beklemiş ilgisiz nesnelerin görülme olasılığı artmaktadır. Bu durumun bir örneği Şekil 5.4(b)'de görülmektedir. Yapay ve gerçek deneyler operatörlerin ilgi değişiminde sistemin kendini dış bir müdahaleye gerek duymadan

uyarlayabildiğini göstermiştir. Gerçekleştirilen sistem uzun bir güvenlik videosunun oldukça hızlı ve rahat bir şekilde izlenebilmesine olanak vermektedir.



Şekil 5.3. Sentetik testlerdeki nesne sayılarının zamanla değişim grafikleri.

5.3. Yorumlar

Çalışmamızın bu bölümünde, çevrimiçi operatör ilgisini kullanan yeni bir güvenlik videosu sentezleme yöntemi geliştirilmiştir. Geliştirilen yöntem operatörün ilgilendiği nesneleri bulmak için onun göz-bakışı konum bilgilerini kullanmaktadır

ve ilgilenilen nesneleri videonun erken bölümlerine yerleştirmektedir. Yapılan deneyler sistemin ilgi değişimlerinde kendini hızlı bir şekilde uyarlayabildiğini göstermiştir. Sistem güvenlik videosu içindeki arama zamanını düşürmekte ve operatöre öncelikli olarak ilgi duyduğu nesneleri göstermektedir. Sistemin sınıflama başarımı oldukça yüksektir ve sistem nesne ve hareket modellerine ihtiyaç duymamaktadır. Bu özelliğiyle güvenlik videolarının çevrimiçi ve çevrimdışı izlenmesinde tüm nesne ve hareket türleri için kullanılabilir.



(a) Tüm gri arabalar operatör için ilgi çekicidir.



(b) İnsanlar ilginç, arabalar ise değildir.

Şekil 5.4. Farklı ilginç nesne tipleri için sentezlenmiş videolardan örnek çerçeveler.

Geliştirilen yöntem bilgisayarla görme, insan bilgisayar etkileşimi ve makine öğrenmesini bir arada kullanmaktadır. Operatör ilgisine duyarlı ikili sınıflandırma yenidir. Sistem sınıflandırma hatlarına karşı gürbüzdür çünkü sınıflandırma sonucu ilgisiz çıkan hareketler filtrelenmezler sadece videoda daha sonraki bir konuma

ötelenirler. Bu çalışmanın çok operatörlü sistemlere genişletilmesi ve her bir operatöre onun ilgi durumuna uygun işin dağıtması planlanmaktadır. Bunun dışında geliştirilen yöntemin operatörlerin gözetleme stratejilerinin çözümlenmesinde ve özne ilgilerinin belirlenmesinde de kullanılabileceği düşünülmektedir.

6. X-IŞINLI BAGAJ İZLEME SİSTEMLERİ İÇİN GÖZ-BAKIŞI TAKİBİ

Nesnelerin içeriklerini tahribatsız şekilde denetleme sistemleri günümüz dünyasında oldukça yoğun şekilde kullanılmaktadır. Bu tür sistemlerin kullandığı en yaygın, pratik, ucuz ve taşınabilir olan teknoloji X-ışını tabanlıdır. Tahribatsız içerik denetleme sistemlerinin kullanım alanları çok fazla olup bunlardan başta gelenleri tıbbi görüntüleme, endüstriyel muayene, askeri ve güvenlik amaçlı paket tarama sistemleridir.

X-ışınli paket tarama sistemlerindeki teknolojik gelişmelere rağmen tamamen otomatik tarama yapan sistemler mevcut değildir. Yüksek güvenlik gereksinimleri ve şüpheli nesne algılama teknolojisindeki yetersizliklerden dolayı güvenlik operatörleri X-ışınli paket tarama sistemlerin önemli bir parçasıdır. Yapılan gözlemler, operatörlerin geçmemesi gereken tehlikeli cisimlerden %20 kadarını fark etmediklerini göstermiştir [Strecker, 1998]. Bu durumun önüne geçmek için birçok yöntem denenmiştir. Bunlardan en başta geleni, operatörlere gösterilen görüntülerin iyileştirilmesi olmuştur [Kase, 2002; Ghylis et al., 2008]. Görüntü iyileştirmelerinin ne kadar etkili olduğu konusunda gerektiği kadar çalışma olmasa da, yapılan geliştirmelerin tehlikeli madde bulmada yardımcı olup olmadığını araştıran çalışmalar vardır [Schwaninger, 2004; Klock, 2005]. Güvenlik operatörlerinin hangi pozisyonları tercih etmesiyle ilgili çalışmalar, pozisyonla personel duruşu arasındaki ilişkiyi ölçmeyi amaçlamıştır [Drury et al., 2008].

Bilimsel kaynaklarda operatörlerin başarımları hakkında yapılan çalışmalara rastlansa da, bu çalışmalar genellikle çevrim dışı özellikte olup, operatörlerin anlık başarımları hakkında fazla bilgi üretmemektedir. Bu konuda yapılan araştırmalar sadece operatörlerin çalışma koşullarının önceden belirlenmesini ve iyileştirilmesini amaçlamaktadır ve çalışma anında operatörle ilgili herhangi bir bilgi edinme yöntemine başvurmamaktadır.

Güvenlik amaçlı X-ışını cihazları genel olarak operatör eğitim, bilgi ve dikkatine bağlı olarak verimlilik alınmaktadır. Sistem tamamen operatörün kontrol

ve sorumluluğundadır. Bu cihazları kullanan operatörlerin eğitim seviyelerinin yüksek olmayışı, iş yoğunluklarının fazlalığı gibi etmenler operatörlerin algı seviyelerinin düşmesine sebep olabilmektedir. Bu nedenle operatörlerin algı seviyelerinin denetlenmesi önemlidir. Ancak X ışın tarayıcı operatörlerinin anlık başarımlarını ölçen sistemlere bilimsel kaynaklarda rastlanmamıştır. Bu tür sistemlerde doğrudan olmasa da, kapalı devre kapalı sistemler üzerinde operatör davranışlarını anlık izleyen sistemler geliştirilmiştir. [Vural and Akgul, 2009; Vural and Akgul, 2011b].

Operatörler paket denetimini iki farklı görüntü sunan iki monitörü izleyerek yapmaktadırlar. Genel olarak efektif atom numarasından elde edilen renklendirilmiş görüntüyü gösteren monitörde şüpheli bir nesne tespit ettiklerinde, yoğunluk bilgisinden elde edilen görüntüyü gösteren ikinci monitördeki resimde şüpheli nesnenin karşılığını bularak kontrol ederler. Bu süreçte operatörün hatalı karşılaştırma yapıp yapmadığının tespiti amacıyla, operatörün göz bebeği izlenip nereye baktığı bulunup, her iki monitörde de aynı nesneye bakıp bakmadığının tespiti yapılacaktır.

Bu sistemlerde daha önce operatörleri de sürece dâhil eden, onların yüklerini ciddi manada hafifleten ve denetimlerini yapan bir sistem geliştirilmemiştir. Çalışmanın bu bölümünde, göz bebeği takibi ve algı – yorgunluk seviyesi ölçümü ile operatörlerin kontrolleri yapılacaktır ve iş yüklerinin ayarlanması sağlanacaktır. Ayrıca yapılan çalışmanın operatörler için ideal değişim zamanlarını belirlemek için de kullanılabileceği düşünülmektedir.

6.1 X-Işınlı Paket Görüntüleme Teknolojileri

X-ışınlı paket görüntüleme cihazları genel olarak X-ışınının yayılım özelliğini kullanarak, tahribatsız olarak, cihaza yerleştirilen paketlerin içindeki malzemelerin çeşitli algılayıcılar ve resim işleme algoritmaları kullanarak operatörün incelemesini sağlayan cihazlardır. Temel X-ışını teknolojisine dayanan bu cihazlar 1920'lerden beri kullanılmaktadır [Kramers, 1923]. Kullanılan temel teknoloji, X ışınlarının

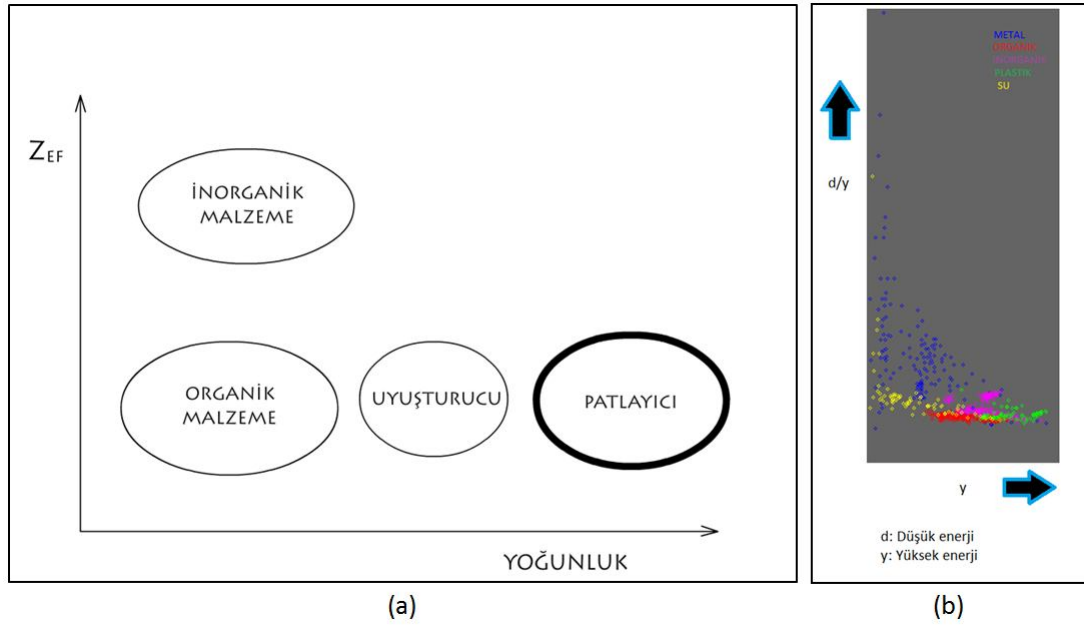
maddeler içerisindeki yayılımına dayansa da, geçen zaman içerisinde bu konuda oldukça büyük gelişmeler kaydedilmiştir. Bu amaçla kurulan ilk sistemler geleneksel enerji görüntüleme teknikleri kullanmaktadır [Peschmann and Harmann, 1993; Crawford and Ruth, 1999; Ying et al., 2006]. Bu sistemler bir kaynaktan çıkan X ışınlarının, x-ışının yayılım özelliği sayesinde, hedef paketten geçtikten sonra elde edilen ölçümlerinden yoğunluk ve kütle gibi özelliklerini çıkarabilmektedirler. Bu yöntemle elde edilen bilgilerin oldukça faydalı olmasına rağmen, yoğunluk ve kütle bilgilerinin patlayıcı gibi tehlikeli organik ve inorganik maddeleri ayırt etmeye yetmeyeceği bilinmektedir. Tek enerji ölçümüne dayalı sistemlerin problemlerine bir çözüm olarak, çift enerjili görüntüleme teknikleri geliştirilmiştir [Shi, 2000; Singh and Singh, 2003a]. Bu teknikler sayesinde, taranan nesnelerin yoğunluklarının yanı sıra nesnelerin efektif atom numaraları da elde edilmektedir. Efektif atom numaraları, nesnelerin gerçek atom numaralarından farklı olmasına ve taranan nesnelerin bileşiminden oluşmuş sanal bir nesnenin atom numarasına denk gelmesine rağmen, tehlikeli madde bulunmasında oldukça etkili olmuştur. Çift enerji sistemlerinden elde edilen görüntüleri değerlendirmek için çeşitli teknikler denenmiştir. Bunlardan bir tanesi, enerji seviyelerini sabit tutmak yerine değiştirmeyi denemiştir [Europ Scan, 1996]. Bir başka metotta ise değişik enerji seviyeleri değişik algılayıcılar aracılığıyla ölçülmüştür [Michette and Buckley, 1993]. Bir başka çalışmada, değişik cisimlerin kalınlıklarının yoğunluk algılamasını etkilemesinin önüne geçilmeye çalışılmıştır [Shi, 2000].

X ışını tarama sistemlerinin güvenilirliklerini daha da artırmak için başka çalışmalar da yapılmıştır. Bu çalışmalar genelde aynı nokta üzerinden birden fazla ışın geçirerek maddelerin özelliklerini daha iyi anlamak ve operatörler için daha kolay görüntülenebilir hale getirmek üzerinde yoğunlaşmıştır [Evans et al., 2001; Wang and Evans, 2003; Singh and Singh, 2003b]. Bu çalışmalar genellikle 3 boyutlu (3B) X ışını başlığı altında toplanabilir. 3B X ışını metotları genellikle tıbbi görüntüleme alanından esinlenerek geliştirilmiş olmasına rağmen taranacak paketlerin geçiş hızları ve taranan nesnelerin özellikleri gereği, bu tür metotlar üzerinde paket taraması için özel geliştirmeler yapılmıştır. 3B X ışın tarayıcılarının başında tıbbi görüntüleme alanından doğrudan alınmış bilgisayarlı tomografi vardır.

Bilgisayarlı tomografi (BT) [Schmidt, 1989], bir paketin yaklaşık 100 kadar değişik açıdan X ışın görüntüsünün alınmasını gerektirmektedir. Bu teknolojinin uygulamaları [Smith and Krall, 1996], paket tarama alanında kullanılmış olmasına rağmen, bu tür sistemler oldukça pahalıya mal olmakta ve 3B hacim elemanlarının bulunması fazla bilgisayar gücü gerektirmektedir. BT paket tarama sistemlerinin 1980’li yıllara göre daha ucuz olmasına rağmen, her türlü durum için kullanılması fiyat ve cihaz kurulum ihtiyaçları bakımından uygun değildir. Bu nedenle, BT gibi aynı nokta üzerinden onlarca ışın geçirmek yerine sadece iki ışın geçirerek bilgi elde etme prensibine dayanan stereoskopik görüntüleme yöntemleri geliştirilmiştir [Vural ve ark., 2012a; Akay ve ark., 2012; Hon and Evans, 2004]. X-ışını tabanlı tarayıcılar ve stereoskopik görüntüleme hakkında daha detaylı bilgi için Abdullah Akay’ın yüksek lisans tezi incelenebilir [Akay, 2012].

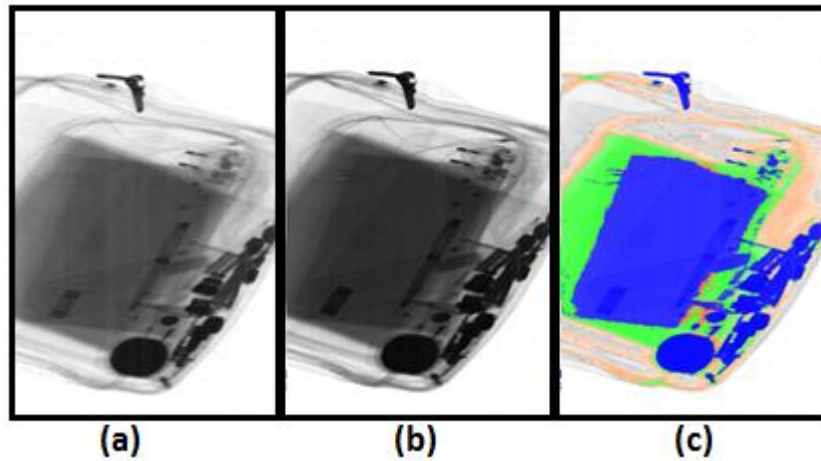
6.2 Yöntem

X-ışını tabanlı bagaj görüntüleme sistemleri için şüpheli nesnelerin bulunması gerekmektedir. Bu nesnelerin bulunması için iki farklı enerji seviyesinde X-ışını dalgası gönderilerek geçen malzemenin efektif atom numarası bulunur. Efektif atom numarası yüksek enerji seviyesinde geçen x-ışını miktarının düşük enerjide geçen x-ışını miktarına bölünmesiyle elde edilen bir orandır. Bulunan efektif atom numarası ve malzemenin yüksek enerji seviyesindeki X-ışını geçirgenliği kullanılarak malzemeler sınıflandırılabilir (Bkz. Şekil 6.1(a)). Bunun için öncelikle önemli miktarda gürültü içeren X-ışını geçirgenlik resimleri gürültüden arındırılır ve normalize edilirler. Geliştirilen sistemde bu resimlerdeki malzemeleri sınıflandırmak için destek vektör makinesi tabanlı bir makine öğrenmesi yöntemi kullanılmıştır ve malzemeler beş farklı sınıfa atanacak şekilde sınıflandırılmıştır (Bkz. Şekil 6.1(b)).



Şekil 6.1. Malzemelerin X-Işını kullanılarak sınıflandırılması, [Akay, 2012].

Sınıflandırma yapılabilmesi için sistem bir kez eğitilir ve daha sonra çalışma sırasında gelen resimler sadece sınıflandırma işlemine tabii tutulurlar. Sınıflandırma için hem yüksek enerji seviyesinde elde edilen resimler (Bkz. Şekil 6.2(a)) ve hem de düşük enerji seviyesinde elde edilen resimler (Bkz. Şekil 6.2(b)) kullanılırlar. Bu resimler sınıflandırma sonrası renklendirilirler (Bkz. Şekil 6.2(c)) ve operatörlerin daha kolay karar vermeleri için kullanılırlar.

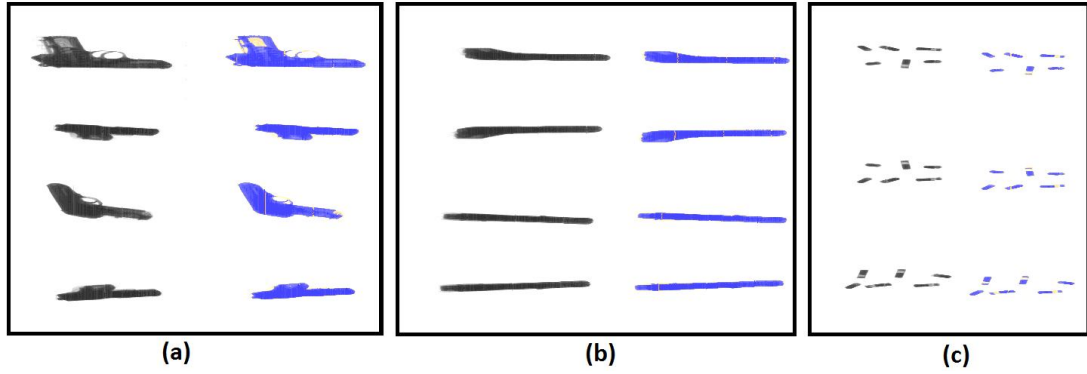


Şekil 6.2. Örnek X-ışını resimleri ve sınıflandırma sonuçları, [Akay, 2012].

Güvenlik operatörleri X-ışınlı paket tarayıcılarından gelen renkli görüntülerle birlikte yüksek/düşük enerji resimlerini bir arada kullanarak karar verirler. Bu nedenle genellikle iki farklı monitör kullanırlar. Bu monitörlerden birinde renklendirilmiş görüntü akarken diğerinde ise siyah beyaz enerji resmi bulur. Operatörün iki resme de bakıp bakmadığının anlaşılması operatörün işini doğru yapıp yapmadığının analizi için önemlidir. Yine aynı şekilde iki resme bakan bir operatörün resimlerde aynı nesnenin bulunduğu konumlara belli bir süre odaklanması beklenmektedir. Bu bilgilerin alınması için operatörlerin göz-bakış konumları izlenecektir. Bunu için Arrington Research [Arrington, 1997] firmasının geliştirmiş olduğu göz-bakışı takip cihazı kullanılacaktır. Göz-takip cihazından elde edilen göz-bebeği büyüklüğü ve göz kırpma sıklıkları kullanılarak güvenlik operatörün dikkat seviyesi de analiz edilmiş olacaktır.

6.3 Deneyler

Operatörlerin dikkat seviyelerinin ölçülmesi ve gözleme işini doğru yapıp yapmadığını ölçmek için sentetik bir test ortamı hazırlanmıştır. Sentetik ortamda güvenlik operatörlerine belli oranda tehdit unsuru içeren nesne barındıran çantalar gösterilmiştir ve operatörlerin bu nesneleri yakalayıp yakalayamadığına bakılmıştır. Operatörlerin yanlış alarm verdiği durumlar da ayrıca sayılmıştır. Deney için öncelikle 10 adet içinde tehdit unsuru içermeyen çanta görüntüsü elde edilmiştir. Tehdit unsuru içeren (Bkz. Şekil 6.3) ve içermeyen çok sayıda nesnenin enerji resimleri elde edilmiştir ve bu nesneler sentetik olarak çantaların içine rastgele yerleştirilmişlerdir. İki adet tabancaya (Bkz. Şekil 6.3(a)) ve mermilere ait (Bkz. Şekil 6.3(c)) görüntüler alınmıştır. Tabancalar x-ışını algılayıcılarına göre dik yerleştirilerek de görüntüleri alınmıştır (Bkz. Şekil 6.3(b)) ve bu görüntülerden deneyimsiz bir operatörün tehdit unsurunu bulmasının zor olacağı görülmüştür.



Şekil 6.3. Tehdit unsuru içeren nesnelerden örnek görüntüler.

Tehdit içeren nesneler daha düşük oranda çantaların içine eklenmiştir ve yapılan resim analiziyle eklenen nesnenin en az %20 oranında görülebilir olması sağlanmıştır. Deneye katılan güvenlik operatörlerine düşük enerji resimleri ve renkli resimler bir arada gösterilecek ve göz bakış konumları kaydedilecektir. Operatörlerden tehlikeli bir nesne gördüklerinde fare ile o nesneyi işaretlemeleri istenecektir. Operatörlerin göz-bebeği büyüklükleri, göz kırpma sıklıkları, iki monitöre birden bakma oranları ile şüpheli nesne bulmadaki başarıları arasındaki ilişki incelenecektir.

6.4 Yorumlar

X-ışını tabanlı bagaj tarama sistemleri yaygın olarak kullanılmaktadır. Bu sistemler için geliştirilen sistemler daha iyi görüntülerin üretilmesini amaçlamakta ve sistemin içinde çevrimiçi olarak bulunan operatörün anlık durumlarını kontrol etmemektedir. Burada geliştirilecek yöntem ile operatörlerin dikkat seviyelerinin ölçülmesi sağlanacaktır. Göz-bakışının farklı monitörler üzerinde odaklandığı konumlar kullanılarak operatörün doğru bir izleme yapıp yapmadığı anlaşılabilecektir. Geliştirilecek bu yöntem ile operatörlerin tecrübeleri ölçülebileceği gibi, yeni operatörlerin eğitimleri için kullanılabilecektir. Bu yöntemin eğitim aşamasında kullanılmasıyla genel olarak operatörleri yanıltan durumlar ortaya çıkacak ve operatörlerin doğru yerlere bakmaları sağlanabilecektir.

Geliştirilecek yöntem ile operatörlerin yorgunluk durumlarına bakılabilecektir ve sabit aralıklı görev değişimleri yerine dinamik görev değişim aralıklarının belirlenebileceği düşünülmektedir. Sistem doğru analiz edilmediği düşünülen görüntülerin bir başka operatör tarafından tekrar izlenmesine de olanak verecektir.

7. SONUÇLAR

Döngüde-insan sistemler oldukça yaygındır ve bu sistemlerin içindeki insan faktörünün yakın gelecekte döngü dışına alınması mümkün değildir. Bununla birlikte, bu tez çalışmasında sistem içindeki insan etkisinin değerlendirilmesinin sistemlerin kararlılığını ve verimliliğini önemli ölçüde arttıracak gösterilmiştir. Tez çalışmasında, güvenlik gözetleme sistemlerindeki operatörlerin doğru ve etkin bir şekilde kullanılabilmesi için dört farklı çözüm getirilmiştir.

İlk olarak güvenlik gözetleme sistemlerindeki personelin gözden kaçırdığı hareketlerin yakalanması sağlanmıştır [Vural and Akgul, 2009]. Operatörlere gözden kaçırdığı olayların kısa bir özetinin gösterilmesi güvenlik gözetleme sistemlerinin başarımlarını arttıracaktır. Yine aynı şekilde gözden kaçırılan olayların bir özeti başka bir operatör ya da yetkili tarafından izlenebilecektir. Bu durum güvenlik operatörleri için nesnel başarı ölçütleri tanımlanabilecektir. Operatörlerin gözden kaçırdıkları videoların uzunluklarını azaltacak şekilde iş ortamları ve çalışma saatleri düzenlenebilecektir. Göz takip cihazlarının ucuzlaması, gözlüksüz yeni takip yöntemleri ve güvenlik pazarının büyümesi bu tür sistemlerin yaygın bir şekilde kullanılmasının önünü açacaktır. Güvenlik videoların dinamik programlama kullanılarak özetlenmesi oldukça hızlıdır. Bununla birlikte göz bakışı tabanlı özetleme sistemleri üç boyutlu video uzaylarında çalışan özetleme algoritmalarıyla birlikte de kullanılabilir.

Güvenlik gözetleme personelinin bir üst mekanizma tarafından gözetlenmesi oldukça ilgi çekicidir. Yapılan çalışma popüler bilim dergisi NewScientist [Marks, 2009] ve ulusal ve yerel basın organları tarafından haber yapılmıştır. Bu çalışma ile insanların çevrimiçi olarak baktıkları yerler analiz edilmiştir ama çalışmanın temel olarak üç eksikliği vardır. İlk eksikliği yaygınlaşan yüksek çözünürlüklü videolarda gerçek-zamanlı olarak çalışamamasıdır. Bu problemin çözümü için algoritmanın paralelleştirilebileceği ve dinamik programlama yerine ikili enerji haritaları üzerinde yığın tabanlı yol bulma yönteminin kullanılabileceği gösterilmiştir [Vural and Akgul, 2011a]. Yapılan çalışma mevcut yüksek çözünürlüklü kameraların hızlarında gecikmeli gerçek-zamanlı olarak çalışabilmektedir. Önerilen sistemin ikinci eksikliği

ise sadece bakılan hareketleri tespit edebilmesi ama bakılan ve görülmeyen hareketleri bilememesidir. Bu yöntemlerin son eksikliği ise üretilen güvenlik videolarının çok yoğun olması ve operatörlerce izlenmesinin zor olmasıdır.

Videolardaki hareketlerin operatörler tarafından izlenebilir hale getirilmesi ve hareketlerinin özelliklerinin çevrimiçi olarak operatöre göre düzenlenmesi fikri yenidir [Vural and Akgul, 2011b]. Bunun için operatörlerin göz bebeği büyüklüğü ve göz kırpma sıklıkları kullanılarak operatörlerin dikkat seviyeleri ölçülmüştür. Operatörlerin dikkat seviyelerine uygun bir videonun oluşturulması için problem tesis yerleşim problemine dönüştürülmüştür. Oldukça zor olan bu problemin hızlı bir şekilde çözülebilmesi için hareketleri belirtecek doğru modelleri geliştirilmiştir. Geliştirilen bu yöntem ile operatörlerin optimal iş yüklerinde çalışması sağlanmıştır. Yapılan sentetik deneyler geliştirilen sistemin hem verimliliği hem de başarımı arttığını göstermiştir.

İnsan bilgisayar etkileşimi ve bilgisayarla görme algoritmalarının bir arada kullanıldığı algoritmalar oldukça azken son olarak buna makine öğrenmesi alanı da katılarak güvenlik operatörlerinin hangi özelliklere sahip hareketlere/nesnelere ilgi duydukları belirlenmiştir [Vural and Akgul, 2012b]. Buradaki ilgi operatörlerin göz-bakışı konumları kullanılarak saptanmış anlık ilgiyi göstermektedir ve sistem ilgi değişikliklerini uyum sağlayabilecek şekilde tasarlanmıştır. Bu çevrimiçi sistem kayıtlı güvenlik videolarındaki nesnelerin bulunması için de oldukça faydalı olacaktır.

Bu çalışmanın operatörlerin uzmanlık durumlarının, önyargılarının ve izleme stratejilerinin incelenmesi için de faydalı olabileceği düşünülmektedir. Bu sistem görsel işlerde çalışan kişilerin eğitimlerinde de kullanılabilecek, bir işi yeni öğrenen kişinin ilgisinin kaymasına sebep olabilecek etmenler bulunabilecektir.

Çalışmanın son kısmında sistem X-ışını tabanlı bagaj tarama sistemlerine uygulanmıştır. Bagaj tarama sistemlerinde çalışan güvenlik operatörlerin göz-bakışı ölçütleri alınarak performansları ölçülmüştür. Renkli ve siyah-beyaz X-ışını resimlerinin birlikte kullanılmasının başarımı nasıl etkilediği analiz edilmiştir.

Döngüde insan sistemlerde operatörlerin anlık başarımlarının ölçülmesinin daha birçok farklı alana uygulanabileceği düşünülmektedir. Fabrikalarda gözle yapılan kalite kontrol işleminde operatörler izlenmemektedir ve kişilerin ürünleri doğru biçimde değerlendirip değerlendirmedikleri bilinmemektedir. Bir başka alan ise insanların hangi tür özniteliklere daha çok ilgi duyduğu, ilgi duyulan özniteliklerle deneklerin kişilikleri arasındaki ilişkinin açığa çıkartılması gibi psikolojik çalışmalar olacaktır.

KAYNAKLAR

1. [Acha et al., 2006] Acha, A. R., Pritch, Y., and Peleg, S. (2006). Making a long video short: Dynamic video synopsis, IEEE Computer Vision and Pattern Recognition or CVPR, pp. I: 435–441.
2. [Ahmad et al., 2007] Ahmad, I., He, Z., Liao, M., Pereira, F., and Sun, M. (2007). Special issue on video surveillance, Circuits and Systems for Video Technology, IEEE Transactions on 17(9): 1271–1271.
3. [Akay ve ark., 2012] Akay, A., Öndin U., ve Akgul, Y. S., (2012). İki X-Işını Sensörü İle Malzeme Sınıflandırması, Akıllı Sistemlerde Yenilikler ve Uygulamaları Sempozyumu (ASYU), Trabzon.
4. [Akay, 2012] Akay, A., (2012). Çift sensörlü x-ışını tarayıcı tasarımı ve bilgisayarla görme incelemesi, Yüksek lisans tezi, Bilgisayar Müh. Bölümü, GYTE.
5. [Akgul et al., 2010] Akgul Y.S., Vural, U., and Yildiz, A.O., (2010). Video summary system, US Patent App. 12/728,172.
6. [Allen et al., 2004] Allen, R., McGeorge, P., Pearso, D., and Milne, A. B., (2004). Attention and expertise in multiple target tracking. Applied Cognitive Psychology, 18:337–347.
7. [Alvarez and Franconeri, 2007] Alvarez, G. and Franconeri, S., (2007). How many objects can you track?: Evidence for a resource-limited attentive tracking mechanism. Journal of Vision, 7(13).
8. [Antic and Ommer, 2011] Antic, B. and Ommer, B. (2011). Video parsing for abnormality detection, ICCV, pp. 2415–2422.
9. [Arrington, 1997] Arrington, K. (1997). Viewpoint eye tracker, Arrington Research.

10. [Atrey et al., 2011] Atrey, P. K., El Saddik, A. and Kankanhalli, M. S. (2011). Effective multimedia surveillance using a human-centric approach, *Multimedia Tools Appl.* 51: 697–721. URL: <http://dx.doi.org/10.1007/s11042-010-0649-1>
11. [Avidan and Shamir, 2007] Avidan, S., and Shamir, A. (2007). Seam carving for content-aware image resizing, *ACM Trans. Graph.* 26(3): 10.
12. [Buono and Simeone, 2010] Buono, P., and Simeone, A. L. (2010). Video abstraction and detection of anomalies by tracking movements, *AVI '10: Proceedings of the International Conference on Advanced Visual Interfaces*, ACM, New York, NY, USA, pp. 249–252.
13. [Ballard and Ozcandarli, 1988] Ballard, D.H., and Ozcandarli, A., (1988). Eye Fixation And Early Vision: Kinetic Depth, *Computer Vision., Second International Conference on* , vol., no., pp.524-531, 5-8.
14. [Barth et al., 2003] Barth, E.; Drewes, J.; Martinetz , T., (2003). Dynamic predictions of tracked gaze, *Signal Processing and Its Applications. Proceedings. Seventh International Symposium on* , vol.1, no., pp. 245-248 vol.1.
15. [Böhme et al, 2005] Böhme, M., Dorr, M., Krause, C., Martinetz, T., and Barth, E. (2005). Eye movement predictions on natural videos. *Neurocomputing*.
16. [Byrne and Parasuraman, 1996] Byrne, E. and Parasuraman, R., (1996). Psychophysiology and adaptive automation, *Biological Psychology* 42(3): 249–268.
17. [Castagnos and Pu, 2010] Castagnos, S. and Pu, P., (2010). Consumer decision patterns through eye gaze analysis, *Workshop on Eye Gaze in Intelligent Human Machine Interaction*.
18. [Cavanagh and Alvarez, 2005] Cavanagh, P. and Alvarez, G. (2005). Tracking multiple targets with multifocal attention, *Trends in Cognitive Sciences* 9(7): 349–354.

19. [Chen and Sen, 2008] Chen, B., and Sen, P. (2008). Video carving, In Short Papers Proceedings of Eurographics.
20. [Choudhary and Tiwari, 2008] Choudhary, V., and Tiwari, A. K. (2008). Surveillance video synopsis, ICVGIP '08: Proceedings of the 2008 Sixth Indian Conference on Computer Vision, Graphics & Image Processing, IEEE Computer Society, Washington, DC, USA, pp. 207–212.
21. [Coltekin, 2005] Coltekin, A., (2005). Stereo-foveation for anaglyph imaging, Proc. SPIE 5664, 48.
22. [Cranor, 2008] Cranor, L. F., (2008). A framework for reasoning about the human in the loop, Proceedings of the 1st Conference on Usability, Psychology, and Security, p.1-15, San Francisco, California.
23. [Crawford and Ruth, 1999] Crawford, C. R. and Ruth, C., (1999). Computed tomography scanning target detection using target surface normals, US Patent 5,901,198.
24. [Dick and Brooks, 2003] Dick, A. R., and Brooks, M. J., (2003). Issues in automated visual surveillance. In International Conference on Digital Image Computing: Techniques and Applications, pages 195-204.
25. [Douglas and Peucker, 1973] Douglas, D. and Peucker, T. (1973). Algorithms for the reduction of the number of points required to represent a digitized line or its caricature, Cartographica: The International Journal for Geographic Information and Geovisualization 10(2): 112–122.
26. [D'Orazio et al., 2007] D'Orazio, T., Leo, M., Guaragnella, C., and Distanti, A., (2007). A visual approach for driver inattention detection. Pattern Recognition, vol. 40, pp. 2341-2355.
27. [Drurya et al., 2008] Drurya, C. G. , Hsiaoa, Y. L., Josepha, C., Joshia, S., Lappa, J., and Pennathura, P. R., (2008). Posture and performance: sitting vs. standing for security screening, Ergonomics, Vol. 51, No. 3, pp. 290–307.

28. [Duchowski, 2002] Duchowski, A.T., (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments and Computing*, Vol. 34, No. 4 (2002) 455-470.
29. [Europ Scan, 1996] Europ Scan Inc. , Xcalibur, an innovative multi-energy X-ray explosive device detection systems, *Proceeding of the Second Explosives Detection Symposium and Aviation Security Conference*, pp. 230–235.
30. [Evans et al., 2001] Evans, P., Robinson, M., and Hon, H. W., (2001). Multiple view dual-energy X-ray imaging, *35th International Carnahan Conference on Security Technology*, pp. 103–107.
31. [Evans et al., 1998] Evans, J. P. O., Robinson, M., and Godber, S. X., (1998). Pseudo tomographic X-ray imaging for use in aviation security, *IEEE AES Systems Mag.* 13, pp. 25–30.
32. [Fehd and Seiffert, 2008] Fehd, H. M., and Seiffert, A. E. (2008). Eye movements during multiple object tracking: Where do participants look, *Cognition* 108(1): 201–209.
33. [Fehd and Seiffert, 2010] Fehd, H. and Seiffert, A. (2010). Looking at the center of the targets helps multiple object tracking, *Journal of vision* 10(4).
34. [Findlay and Gilchrist, 2001] Findlay, J. M., and Gilchrist, I. D., (2001). *Vision and Attention*, ch.5 pp.85-106, Springer.
35. [Fletcher and Zelinsky, 2009] Fletcher, L., and Zelinsky A., (2009). Driver Inattention Detection based on Eye Gaze Road Event Correlation. In: *The International Journal of Robotics Research*.
36. [Franconeri et al., 2007] Franconeri, S. L., Alvarez, G. A., and Enns, J. T. (2007). How many locations can be selected at once?, *J Exp Psychol Hum Percept Perform* 33(5): 1003–1012.

37. [Franconeri et al., 2010] Franconeri, S., Jonathan, S. and Scimeca, J. (2010). Tracking multiple objects is limited only by object spacing, not by speed, time, or capacity, *Psychological Science* 21(7): 920.

38. [Friedman et al., 2000] Friedman, J., Hastie, T. and Tibshirani, R. (2000). Additive logistic regression: a statistical view of boosting, *The annals of statistics* 28(2): 337–407.

39. [Gavin, 2004] Gavin, S.J.D., (2004). Behind the screens: Examining constructions of deviance and informal practices among cctv control room operators in the uk., *Surveillance and Society*, Vol. 2(2/3), pp. 376–395.

40. [Ghylin et al., 2008] Ghylin, K. M. , Schwaninger, A. , Drury, C. G. , Redford, J. , Lin, L., and Batta, R., (2008). Screening Enhancements: Why don't they enhance performance?, *PROCEEDINGS of the HUMAN FACTORS AND ERGONOMICS SOCIETY 52nd ANNUAL MEETING*.

41. [Green, 1999] Green, M. (1999). The appropriate and effective use of security technologies in us schools. a guide for schools and law enforcement, *Technical report*.

42. [Gryn et al., 2009] Gryn, J. M., Wildes, R. P., and Tsotsos, J. K. (2009). Detecting motion patterns via direction maps with application to surveillance, *Computer Vision and Image Understanding* 113(2): 291– 307.

43. [Gu et al., 2005] Gu, E., Stocker, C. and Badler, N. (2005). Do you see what eyes see? Implementing Inattentional Blindness, *Intelligent Virtual Agents*, Springer, pp. 178–190.

44. [Hampapur et al., 2007] Hampapur, A., Brown, L., Feris, R., Senior, A., Shu, C., Tian, Y., Zhai, Y., and Lu, M. (2007). Searching surveillance video, *AVSBS07*, pp. 75–80.

45. [Han et al., 2011] Han, B., Hamm, J. and Sim, J., (2011). Personalized video summarization with human in the loop, *2011 IEEE Workshop on Applications of Computer Vision*, pp. 51-57.

46. [Harris and Jenkin, 2001] Harris, R. L., and Jenkin, M., (2001). Vision and Attention, ch.1 pp.1-17, Springer.
47. [Hillaire et al., 2008] Hillaire, S., Lecuyer, A., Cozot, R., and Casiez, G., (2008). Using an Eye-Tracking System to Improve Camera Motions and Depth-of-Field Blur Effects in Virtual Environments, Virtual Reality Conference, 2008. VR '08. IEEE , vol., no., pp.47-50.
48. [Hon and Evans, 2004] Hon, H. W. and Evans, J. P. O., (2002). Multiple-view line-scan imaging, IEE Proc. Optoelectron. . Vol. 149. No. 2.
49. [Horowitz et al., 2004] Horowitz, T., Holcombe, A., Wolfe, J., Arsenio, H. and DiMase, J. (2004). Attentional pursuit is faster than attentional saccade, Journal of Vision 4(7).
50. [Hu et al., 2004] Hu, W., Tan, T., Wang, L., and Maybank, S. (2004). A survey on visual surveillance of object motion and behaviors, IEEE Trans. Syst., Man, Cybern pp. 334–352.
51. [Hua and Yingzi, 2011] Hua, C. and Yingzi, L., (2011). Modeling of operators' emotion and task performance in a virtual driving environment, International Journal of Human-Computer Studies, Volume 69, Issue 9, August 2011, Pages 571-586, ISSN 1071-5819, 10.1016/j.ijhcs.2011.05.003.
52. [Hutchinson et al., 1989] Hutchinson, T. E., White, K. P., Martin, W. N., Reichert, K. C., and Frey, L. A. (1989). Human-computer interaction using eye-gaze input, Systems, Man and Cybernetics, IEEE Transactions on 19(6): 1527–1534.
53. [Iannizzotto et al., 2005] Iannizzotto, G., Costanzo, C., Rosa, F. L., and Lanzafame, P. (2005). A multimodal perceptual user interface for video-surveillance environments, ICMI, pp. 45–52.
54. [İnce ve Göktürk, 2009] İnce, O., Göktürk, M., (2009). Akademik Bilişim 2009 konferansı dahilinde “Akademik Bilişim Konferansı Bildirileri” bildiri kitapçığındaki “Güvenlik Sistemi İzleyici Personelinin Görsel Tarama Davranışının Analizi”, 699-704 pp., Urfa Türkiye.

55. [Jacob, 1991] Jacob, R. (1991). The use of eye movements in human-computer interaction techniques: What you look at is what you get, *ACM Transactions on Information Systems* 9(3): 152–169.
56. [Jaimes et al., 2001] Jaimes, R., Pelz, J., Grabowski, T., Babcock, J., and fu Chang, S. (2001). Using human observers' eye movements in automatic image classifiers, *Proceedings of SPIE Human Vision and Electronic Imaging VI*.
57. [Ji and Yang, 2002] Ji, Q., and Yang, X., (2002). Real-time eye, gaze, and face pose tracking for monitoring driver vigilance, *Real-Time Imaging* 3 (8), pp. 1077–2014.
58. [Jing and Lansun, 2008] Jing, Z., and Lansun, S. (2008). A personalized image retrieval based on visual perception, *Journal of Electronics (China)*, Vol. 25.
59. [Kase, 2002] Kase, K, (2002). Effective Use of Color in X-ray Image Enhancement for Luggage Inspection, Masters Degree Thesis, The University of Tennessee, Knoxville.
60. [Keval and Sasse, 2006] Keval, H. and Sasse., M. A., (2006). Man or gorilla? Performance issues with cctv technology in security control rooms. 16th World Congress on Ergonomics Conference, International Ergonomics Association.
61. [Keval and Sasse, 2010] Keval, H. and Sasse, M.A., (2010). Not the usual suspects:A study of factors reducing the effectiveness of CCTV. In *The Security Journal*, 23, 2.
62. [Kim and Hwang, 2000] Kim, C., and Hwang, J.-N. (2000). An integrated scheme for object-based video abstraction, *MULTIMEDIA '00: Proceedings of the eighth ACM international conference on Multimedia*, ACM, New York, NY, USA, pp. 303–311.
63. [Klock, 2005] Klock, B. A., (2005). Test and evaluation report for X-ray detection of threats using different X-ray functions. *IEEE ICCST Proceedings*, 182-184.

64. [Komlodi and Marchionini, 1998] Komlodi, A., and Marchionini, G. (1998). Key frame preview techniques for video browsing, DL '98: Proceedings of the third ACM conference on Digital libraries, ACM, New York, NY, USA, pp. 118–125.
65. [Komogortsev et al., 2011] Komogortsev, O., Holland, C., and Camou, J., (2011). Adaptive eye-gaze-guided interfaces: design & performance evaluation. In Proceedings of the 2011 annual conference extended abstracts on Human factors in computing systems (CHI EA '11). ACM, New York, NY, USA, 1255-1260.
66. [Koskela, 2000] Koskela, H. (2000). The gaze without eyes: Video-surveillance and the nature of urban space, Progress in Human Geography 24(2): 243–265.
- 67.[Kramers, 1923] H. A. Kramers, On the theory of X-ray absorption and of the continuous X-ray spectrum, Philos. Mag. 46 (1923) 836–871.
68. [Lavine et al., 2002] Lavine, R., Sibert, J.L., Göktürk, M., Dickens, B., (2002). Eye-Tracking Measures and Human Performance in a Vigilance Task, Aviation Space and Environmental Medicine (ISI) , pp. 367-372.
69. [Law et al., 2004] Law, B., Atkins, M. S., Kirkpatrick, A. E., Lomax, A. J., and MacKenzie, C. L., (2004). Eye Gaze Patterns Differentiate Novice and Experts in a Laparoscopic Surgery Training Environment. In Proc. ETRA, ACM Press, 41–48.
70. [LC Technologies, 1997] LC Technologies (1997). The eyegaze communication system. URL: <http://www.eyegaze.com>
71. [Li et al., 2000] Li, F. C., Gupta, A., Sanocki, E., wei He, L., and Rui, Y. (2000). Browsing digital video, CHI '00: Proceedings of the SIGCHI conference on Human factors in computing systems, ACM, New York, NY, USA, pp. 169–176.
72. [Li et al., 2009] Li, T., Mei, T., Kweon, I.-S., and Hua, X.-S. (2009). Multi-video synopsis for video representation, Signal Process. 89(12): 2354–2366.

73. [Li et al., 2001] Li, Y., Li, Y., Zhang, T., Zhang, T., Tretter, D., and Tretter, D. (2001). An overview of video abstraction techniques, Technical report, Imaging Systems Laboratory, HP Laboratories, Palo Alto.
74. [Liu, 2009] Liu, B. (2009). Facility Location Problem, Theory and Practice of Uncertain Programming pp. 157–165.
75. [López et al., 2006] López, M. T., Fernández-Caballero, A., Fernández, M. A., Mira, J., and Delgado, A. E. (2006). Visual surveillance by dynamic visual attention method, Pattern Recogn. 39(11): 2194–2211.
76. [Lundmark, 2012] Lundmark, P., (2012). Higher Efficiency in Operations Can Be Achieved with More Focus on the Operator, Ergonomics - A Systems Approach, Dr. Isabel L. Nunes (Ed.), ISBN: 978-953-51-0601-2, InTech.
77. [Ma and Huang, 2009] Ma, W. and Huang, W. (2009). No capacity limit in attentional tracking: Evidence for probabilistic inference under a resource constraint, Journal of Vision 9(11).
78. [Ma et al., 2002] Ma, Y., Lu, L., Zhang, H. and Li, M. (2002). A user attention model for video summarization, Proceedings of the tenth ACM international conference onMultimedia, ACM, pp. 533–542.
79. [Marat et al., 2009] Marat, S., Phuoc, T., Granjon, L., Guyader, N., Pellerin, D., and Guerin-Dugue, A., (2009). Modelling spatio-temporal saliency to predict gaze direction for short videos. International Journal of Computer Vision (IJCV), 82(3):231–243.
80. [Marks, 2009] Marks, P., (2009). Eyeball spy turns the tables on Big Brother, New Scientist, 14 April, <http://www.newscientist.com/article/mg20227035.500-eyeball-spy-turns-the-tables-on-big-brother.html>.
81. [Matsuyama, 2007] Matsuyama T., (2007). Ubiquitous and wearable vision systems, Imaging beyond the pinhole camera, ch. 5, pp. 307, Springer, Netherlands.

82. [Michette and Buckley, 1993] Michette, A. G. and Buckley, C. J., (1993). X-ray Science and Technology, Institute of Physics Publishing, Bristol, pp. 1–44.
83. [Mohamed et al., 2008] Mohamed, A.O., Perreira Da Silva, M., Courboulay, V., (2008). A history of eye gaze tracking. Tech. Report: hal-00215967, ver. 1 - 24.
84. [Morimoto and Mimica, 2005] Morimoto, C. H., and Mimica, M. R. M., (2005). Eye gaze tracking techniques for interactive applications, *Computer Vision and Image Understanding*, vol. 98, no. 1, pp. 4–24.
85. [Murphy and Duchowski, 2001] Murphy, H., and Duchowski, A. T., (2001). Gaze-Contingent Level Of Detail Rendering. In *Eurographics*, Manchester, UK.
86. [Murphy and Duchowski, 2007] Murphy, H., and Duchowski, A. T., (2007). Hybrid image-_model-based gaze-contingent rendering, in *Proceedings of ACM Symposium on Applied Perception in Graphics and Visualization*, pp. 1–8.
87. [Nonoka, 2003] Nonoka, H., (2003). Communication Interface with Eye-Gaze and Head Gesture Using Successive DP Matching and Fuzzy Inference, *Journal of Intelligent Information Systems*, 21:2, 105–112.
88. [Norris and Armstrong, 1999] Norris, C., and Armstrong, G. (1999). *Cctv and the social structuring of surveillance, Surveillance of Public Space: CCTV, Street Lighting and Crime Prevention.*, Monsey: Criminal Justice Press.
89. [Oron-Gilad et al., 2002] Oron-Gilad, T., Ronen, A., Cassuto, Y., and Shinar, D., (2002). Alertness maintaining tasks while driving. In *Human Factors and Ergonomics Society Annual Meeting Proceedings*, volume 46, pages 1839–1843. Human Factors and Ergonomics Society.
90. [Oyekoya and Stentiford, 2004] Oyekoya, O.K., and Stentiford, F.W.M., (2004). Eye tracking as a new interface for image retrieval. *British Telecommunications Technology Journal* 22 (3).

91. [Parasuraman, 2007] Parasuraman, R. (2007). Adaptive automation for human-robot teaming in future command and control systems, Technical report, ARMY RESEARCH LAB ABERDEEN PROVING GROUND MD HUMAN RESEARCH AND ENGINEERING DIRECTORATE.

92. [Peschmann and Harmann, 1993] Peschmann, K. R. and Harmann, J., (1993). Automatic Concealed Object Detection System Having a Pre-Scan Stage, US Patent 5, 182, 764.

93. [Pets, 2001] PETS 2001 Benchmark Data (2001). Online.

94. [Preece et al., 1994] Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., and Carey, T. (1994). Human-Computer Interaction, Addison-Wesley Longman Ltd., Essex, UK, UK.

95. [Pretlove and Skourup, 2007] Pretlove, J. and Skourup, C., (2007). Human in the Loop, The human operator is a central figure in the design and operation of industrial automation systems. ABB Review, The corporate technical journal of the ABB Group. 1/2007, page 6-10.

96. [Pritch et al., 2009] Pritch, Y., Ratovitch, S., Hendel, A., and Peleg, S. (2009). Clustered synopsis of surveillance video, AVSS '09: Proceedings of the 2009 Sixth IEEE International Conference on Advanced Video and Signal Based Surveillance, IEEE Computer Society, Washington, DC, USA, pp. 195–200.

97. [Pylyshyn and Storm, 1988] Pylyshyn, Z. W., and Storm, R. W. (1988). Tracking multiple independent targets: evidence for a parallel tracking mechanism., *Spatial vision* 3(3): 179–197.

98. [Ren et al, 2007] Ren, R., Swamy, P., Jose, J. and Urban, J. (2007). Attention-based video summarisation in rushes collection, Proceedings of the international workshop on TRECVID video summarization, ACM, pp. 89–93.

99. [Rothrock and Narayanan, 2011] Rothrock, L. and Narayanan, S., (2011). Human-in-the-Loop Simulations: Methods and Practice, Springer-Verlag, New York.

100. [Sasse, 2010] Sasse, M. A. (2010). Not seeing the crime for the cameras?, *Commun. ACM* 53(2): 22–25.
101. [Schmidt, 1989] Schmidt, M., (1989). A review of X-ray computed axial tomography, *NDT-Australia* 26 (2) , pp. 42–47.
102. [Scholl et al., 2001] Scholl, B., Pylyshyn, Z. and Feldman, J. (2001). What is a visual object? Evidence from target merging in multiple object tracking, *Cognition* 80(1-2): 159–177.
103. [Schwaninger, 2004] Schwaninger, A., (2004) . Computer based training: A powerful tool to the enhancement of human factors. *Aviation Security International*, pp. 31-36.
104. [Sears and Pylyshyn, 2000] Sears, C. R., and Pylyshyn, Z. W. (2000). Multiple object tracking and attentional processing, *Canadian Journal of Experimental Psychology* 54(1): 1–14.
105. [Shi, 2000] Shi, X., (2000). Improving object classification in x-ray luggage inspection, Ph. D. dissertation, Virginia Polytechnic Institute and State University.
106. [Siebel and Maybank, 2004] Siebel, N. T., and Maybank, S. J. (2004). The advisor visual surveillance system, in M. Clabian, V. Smutny & G. Stanke (eds), *Proceedings of the ECCV 2004 workshop "Applications of Computer Vision" (ACV'04)*, Prague, Czech Republic, pp. 103–111.
107. [Singh and Singh, 2003a] Singh, S. and Singh, M., (2003). Explosives detection systems (EDS) for aviation security, *Signal Processing* 83, pp. 31–55.
108. [Singh and Singh, 2003b] Singh, S. and Singh, M., (2003). Review: Explosives detection systems (EDS) for aviation security, *Signal Processing*, Elsevier, Vol. 83.

109. [Slot et al., 2009] Slot, K., Truelsen, R., and Sporring, J. (2009). Content-aware video editing in the temporal domain, SCIA '09: Proceedings of the 16th Scandinavian Conference on Image Analysis, Springer-Verlag, Berlin, pp. 490–499.
110. [Smith and Krall, 1996] Smith, R. C. and Krall, P. R., (1996). Full volume dual-energy high speed computed tomography (CT) explosives detection system, Proceedings of the Second Explosives Detection Symposium and aviation Security Conference, pp. 260–263.
111. [Snoek and Worring, 2005] Snoek, C. G. M., and Worring, M. (2005). Multimodal video indexing: A review of the state-of-the-art, *Multimedia Tools Appl.* 25(1): 5–35.
112. [Steiger et al., 2005] Steiger, O., Cavallaro, A., and Ebrahimi, T. (2005). Real-Time Generation of Annotated Video for Surveillance, Proceedings of IEEE Workshop on Image Analysis for Multimedia Interactive Services, WIAMIS 2005, ISCAS, SPIE.
113. [Strecker, 1998] Strecker, H., (1998). Automatic detection of explosives in airline baggage using elastic X-ray scatter, *Medicamundi* 42 (2), pp. 30–33.
114. [Stutzle, 2006] Stutzle, T. (2006). Iterated local search for the quadratic assignment problem, *European Journal of Operational Research* 174(3): 1519–1539.
115. [Takala and Pietikainen, 2007] Takala, V., and Pietikainen, M., (2007). Multi-Object Tracking Using Color, Texture and Motion, *Computer Vision and Pattern Recognition, CVPR '07. IEEE Conference on* , vol., no., pp.1-7, 17-22.
116. [Talmi and Liu, 1999] Talmi, K., and Liu, J., (1999). Eye and gaze tracking for visually controlled interactive stereoscopic displays, *Signal Processing: Image communication Vol.14*, pp. 799-810.
117. [Thiffault and Bergeron, 2003] Thiffault, P. and Bergeron, J., (2003). Monotony of road environment and driver fatigue: a simulator study. *Accident Analysis & Prevention*, 35(3):381–391.

118. [Topal et al., 2008] Topal, C., Dogan, A., and Gerek, O. N., (2008). An Eye-Glasses-Like Wearable Eye Gaze Tracking System, 16. IEEE Sinyal İşleme ve İletişim Uygulamaları, Didim, Türkiye.

119. [Trick et al., 2009] Trick, L.M., Hollinsworth, H., and Brodeur, D., (2009). Multiple-object tracking across the lifespan: Do different factors contribute to diminished performance in different age groups? In Don Dedrick and Lana Trick (Eds). Computation, Cognition, and Pylyshyn. (pp. 79-99). MIT press.

120. [Truong and Venkatesh, 2007] Truong, B. T., and Venkatesh, S. (2007). Video abstraction: A systematic review and classification, ACM Trans. Multimedia Comput. Commun. Appl. 3(1): 3

121. [Vertegaal, 2002] Vertegaal, R., (2002). Designing attentive interfaces, in: Proc. of the Eye Tracking Research & Applications Symposium, New Orleans, LA.

122. [Vural and Akgul, 2009] Vural, U. and Akgul, Y. S., (2009). Eye-gaze based real-time surveillance video synopsis, Pattern Recognition Letters, Volume 30, Issue 12, Pages 1151-1159, ISSN 0167-8655, 10.1016/j.patrec.2009.03.002.

123. [Vural and Akgul, 2011a] Vural, U. and Akgul, Y. S., (2011). A parallel non-linear surveillance video synopsis system with operator eye-gaze input. In Video Surveillance. InTech.

124. [Vural and Akgul, 2011b] Vural, U. and Akgul, Y.S., (2011). Operator attention based video surveillance, Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on , vol., no., pp.1955-1962, 6-13 Nov. 2011, doi: 10.1109/ICCVW.2011.6130488.

125. [Vural ve ark., 2012a] Vural, U., Akay, A., and Akgul, Y.S., (2012). Stereo için güvenilirlik tabanlı çizge-kesme algoritması, 20. IEEE Sinyal İşleme ve İletişim Uygulamaları Konferansı, Fethiye, Türkiye.

126. [Vural and Akgul, 2012b] Vural, U. and Akgul, Y.S., (2012). Operator interest based surveillance video synthesis, 21st International Conference on Pattern Recognition, Tokyo, (Accepted.).

127. [Wallace et al., 1997] Wallace, E., Diffley, D., Baines, E. and Aldridge, J. (1997). Ergonomic design considerations for public area CCTV safety and security applications, Proceedings of the 13th Triennial Congress of the International Ergonomics Association, Tampere, Finland.
128. [Wang and Evans, 2003] Wang, T. W. and Evans, J. P. O., (2003). Stereoscopic dual energy X-ray imaging for target materials identification, IEEE Proceedings of Vision, Image and Signal Processing, Vol. 150, pp. 122 –130.
129. [Wang et al., 2009] Wang, X., Han, T. and Yan, S. (2009). An hog-lbp human detector with partial occlusion handling, ICCV, pp. 32–39.
130. [Warm et al., 1996] Warm, J., Dember, W. and Hancock, P. (1996). Vigilance and workload in automated systems, Automation and human performance: Theory and applications pp. 183–200.
131. [Wolfe et al., 2007] Wolfe, J., Place, S. and Horowitz, T. (2007). Multiple object juggling: Changing what is tracked during extended multiple object tracking, Psychonomic bulletin & review 14(2): 344–349.
132. [Yildiz et al., 2008] Yildiz, A., Ozgur, A., and Akgul, Y. (2008). Fast non-linear video synopsis, 23rd of the International Symposium on Computer and Information Sciences, Istanbul, Turkey .
133. [Ying et al., 2006] Ying, Z., Naidu, R., and Crawford, C. R., (2006). Dual energy computed tomography for explosive detection, Journal of X-Ray Science and Technology, pp. 235–256.
134. [Yu et al., 2011] Yu, C., Zheng, X., Zhao, Y., Liu, G. and Li, N. (2011). Review of intelligent video surveillance technology research, (EMEIT, Vol. 1, IEEE, pp. 230–233.
135. [Zhang and Nayar, 2006] Zhang, L., and Nayar, S. (2006). Projection defocus analysis for scene capture and image display, ACM Trans. Graph. 25(3): 907–915.

ÖZGEÇMİŞ

1980 yılında İstanbul’da doğmuştur. İlk ve orta öğretimini 1988–1996 yılları arasında sırasıyla Gebze Osmangazi İlkokulu ve Gebze Anadolu Lisesi’nde tamamlamıştır. Ordu Fen Lisesi’nde bir yıl okuduktan sonra Gebze Sarkuysan Lisesi’nden 1999 yılında mezun olmuş ve aynı yıl girdiği Kocaeli Üniversitesi Bilgisayar Mühendisliği Bölümü’nü 2003 yılında bitirmiştir. Yüksek lisans ve doktora eğitimlerine sırasıyla 2004 ve 2006 yıllarında Gebze Yüksek Teknoloji Enstitüsü bilgisayar mühendisliği bölümünde başlayan Ulaş Vural, 2005 yılından beri aynı kurumda araştırma görevlisi olarak çalışmaktadır.

EKLER

İNGİLİZCE –TÜRKÇE	
A	
Abstraction	Özet
Acceptance Change	Kabul Şansı, Seçilme Şansı
Adaptive Automation	Uyarlamalı Otomasyon
Attention	Dikkat
Axis	Eksen
B	
Background	Arka Alan
Bin	Sepet
Binary	İkili
Blink Rate	Göz Kırpma Sıklığı
Bounding Box	Kapsama Kutusu
Brightness	Parlaklık
Buffer	Tampon
C	
Computer Vision	Bilgisayarla Görme
Configuration	Düzenleniş
Content-Aware	İçeriği Duyarlı
Content-Based	İçerik Tabanlı
D	
Dead-End	Çıkmaz Sokak
Dynamic Programming	Dinamik Programlama
E	
Edge	Kenar
Eye Blinking	Göz Kırpma
Eye-Gaze	Göz Bakışı
Eye-Gaze Tracker	Göz Bakışı Takip Cihazı

F	
Facility Location Problem	Tesis Yerleşim Problemi
Field Of View	Görüş Alanı
Focusing	Odaklanma
Frame	Çerçeve
Frequency	Sıklık
H	
Heuristic	Sezgisel
Histogram	Sıklık Grafiği
Human Visual System	İnsan Görme Sistemi
Human-Computer Interaction	İnsan-Bilgisayar Etkileşimi
Human-In-The-Loop	Döngüde-İnsan
Hyper-thread	Üstün-kanalcık
I	
Intensity	Yoğunluk
Interactive	Etkileşimli
Iterated Local Search	Yinelemeli Yerel Arama
K	
Key Frame	Anahtar Çerçeve
Knapsack Problem	Sırt Çantası Problemi
L	
Labeling	Etiketleme
Line Segment	Doğru Parçası
Linear	Doğrusal
Literature	Kaynak, Bilimsel Kaynak
Local Binary Patterns	Yerel İkili Örüntü
Local Search	Yerel Arama
M	
Machine Learning	Makine Öğrenmesi
Manpower	İş Gücü
Monitoring	İzleme
Motion Features	Hareket Özellikleri
Multi-Object Tracking	Çoklu Nesne Takibi

N	
Node	Düğüm
O	
Offline	Çevrim Dışı
Online	Çevrimiçi
Operational Cost	İşletme Maliyeti
Optimization	Eniyileme
Overload	Aşırı Yükleme, Aşırı İş Yüğü
Overlooking	Gözden Kaçırma
P	
Partial Derivative	Kısmi Türev
Pattern	Örüntü
Pixel	Piksel
Processing Time	İşleme Zamanı
Projection	Yansıtma
Pupil	Göz Bebeği
Pupil Diameter	Göz Bebeği Çapı
R	
Real-Time	Gerçek-Zamanlı
Region Of Interest	İlgi Alanı
Reliable	Güvenilir
Resolution	Çözünürlük
Retrieval	Geri Bulma
S	
Saccade	Sıçrama
Scaling	Ölçekleme
Semantic	Anlamsal
Similarity	Benzerlik
Space-Time	Uzay-Zaman
Stable	Kararlı
Stack	Yığın
Surveillance	Güvenlik Gözetleme
Synopsis	Özet

T	
Texture	Doku
Thread	Kanalcık
Threshold	Eşik Değer
Tractability	İzlenebilirlik, Takip Edilebilirlik
Training Samples	Eğitim Örnekleri
U	
Underload	Düşük İş Yüğü
Unsupervised	Eğiticişiz
V	
Visual Artifact	Görsel Bozulma
Visual Field	Görme Alanı
W	
Workload	İş Yüğü

TÜRKÇE - İNGİLİZCE

A

Anahtar Çerçeve	Key Frame
Anlamsal	Semantic
Arka Alan	Background
Aşırı Yükleme, Aşırı İş Yüğü	Overload

B

Benzerlik	Similarity
Bilgisayarla Görme	Computer Vision
Bilimsel Kaynak	Literature

Ç

Çerçeve	Frame
Çevrim Dışı	Offline
Çevrimiçi	Online
Çıkmaz Sokak	Dead-End
Çoklu Nesne Takibi	Multi-Object Tracking
Çözünürlük	Resolution

D

Dikkat	Attention
Dinamik Programlama	Dynamic Programming
Doğru Parçası	Line Segment
Doğrusal	Linear
Doku	Texture
Döngüde-İnsan	Human-In-The-Loop
Düğüm	Node
Düşük İş Yüğü	Underload
Düzenleniş	Configuration

E	
Eğitici	Unsupervised
Eğitim Örnekleri	Training Samples
Eksen	Axis
Eniyileme	Optimization
Eşik Değer	Threshold
Etiketleme	Labeling
Etkileşimli	Interactive
G	
Gerçek-Zamanlı	Real-Time
Geri Bulma	Retrieval
Görme Alanı	Visual Field
Görsel Bozulma	Visual Artifact
Görüş Alanı	Field Of View
Göz Bakışı	Eye-Gaze
Göz Bakışı Takip Cihazı	Eye-Gaze Tracker
Göz Bebeği	Pupil
Göz Bebeği Çapı	Pupil Diameter
Göz Kırpma	Eye Blinking
Göz Kırpma Sıklığı	Blink Rate
Gözden Kaçırma	Overlooking
Güvenilir	Reliable
Güvenlik Gözetleme	Surveillance
H	
Hareket Özellikleri	Motion Features
İ	
İçeriği Duyarlı	Content-Aware
İkili	Binary
İlgi Alanı	Region Of Interest
İnsan Görme Sistemi	Human Visual System
İnsan-Bilgisayar Etkileşimi	Human-Computer Interaction
İş Gücü	Manpower
İş Yüğü	Workload
İşleme Zamanı	Processing Time
İşletme Maliyeti	Operational Cost
İzleme	Monitoring
İzlenebilirlik, Takip Edilebilirlik	Tractability

K	
Kabul Şansı	Acceptance Change
Kanalcık	Thread
Kapsama Kutusu	Bounding Box
Kararlı	Stable
Kaynak	Literature
Kenar	Edge
Kısmi Türev	Partial Derivative
M	
Makine Öğrenmesi	Machine Learning
O	
Odaklanma	Focusing
Ö	
Ölçekleme	Scaling
Örüntü	Pattern
Özet	Abstraction, Synopsis
P	
Parlaklık	Brightness
Piksel	Pixel
S	
Sepet	Bin
Sezgisel	Heuristics
Sıçrama	Saccade
Sıklık	Frequency
Sıklık Grafiği	Histogram
Sırt Çantası Problemi	Knapsack Problem
T	
Tampon	Buffer
Tesis Yerleşim Problemi	Facility Location Problem

U	
Uyarlamalı Otomasyon	Adaptive Automation
Uzay-Zaman	Space-Time
Ü	
Üstün-kanalcık	Hyper-thread
Y	
Yansıtma	Projection
Yazın	Literature
Yerel Arama	Local Search
Yerel İkili Örüntü	Local Binary Patterns
Yığın	Stack
Yinelemeli Yerel Arama	Iterated Local Search
Yoğunluk	Intensity