

**KOCAELİ ÜNİVERSİTESİ \* FEN BİLİMLERİ ENSTİTÜSÜ**

**VIDEO İÇİN NESNE BÖLÜTLEMESİ**

**135931**

**YÜKSEK LİSANS TEZİ**

**Elektronik ve Haberleşme Müh. Oğuzhan URHAN**

**Anabilim Dalı: Elektronik ve Haberleşme**

**Danışman: Doç.Dr. Sarp ERTÜRK**

**TC. YÜKSEKÖĞRETİM KURULU  
DOKÜMANİZYON MERKEZİ**

**MAYIS 2003**

**KOCAELİ ÜNİVERSİTESİ \* FEN BİLİMLERİ ENSTİTÜSÜ**

**VIDEO İÇİN NESNE BÖLÜTLEMESİ**

**YÜKSEK LİSANS TEZİ**

**Elektronik ve Haberleşme Müh. Oğuzhan URHAN**

**Tezin Enstitüye Verildiği Tarih: 20 Mayıs 2003**

**Tezin Savunulduğu Tarih: 25 Haziran 2003**

**Y.C. YÜKSEKÖĞRETİM KURULU  
DOKÜMANTASYON MEMBERS**

**Tez Danışmanı**

**Doç.Dr. Sarp ERTÜRK**



**Üye**

**Doç. Dr. Işıl CELASUN**



**Üye**

**Yrd. Doç. Dr. Mehmet YAKUT**



**MAYIS 2003**

**KOCAELİ ÜNİVERSİTESİ \* FEN BİLİMLERİ ENSTİTÜSÜ**

**VIDEO İÇİN NESNE BÖLÜTLEMESİ**

**YÜKSEK LİSANS TEZİ**

**Elektronik ve Haberleşme Müh. Oğuzhan URHAN**

**Anabilim Dalı: Elektronik ve Haberleşme**

**Danışman: Doç.Dr. Sarp ERTÜRK**

135951

**MAYIS 2003**

# VIDEO İÇİN NESNE BÖLÜTLEMESİ

Oğuzhan URHAN

**Anahtar Kelimeler:** Video Bölütlemesi, Video Nesne Bölütlemesi, Video Nesne Uzayı Bölütlemesi, Hareketli Nesne Bölütlemesi, Blok Bazlı Bölütleme, Arka Plan Kaydı, MPEG-4.

**Özet:** Sayısal teknolojiler hayatın her alanında olduğu gibi resim, video v.b. gibi görüntü bilgisi içeren materyallerin işlenmesinde de uzun zamandır kullanılmaktadır. Başlangıçta analog olarak edip, üzerinde değişiklik yapıp yayınlanan video, artık sayısal olarak kaydedilip, sayısal kodlama teknikleri ile sıkıştırılıp, uydular üzerinden yayınlanmaktadır.

Analog video işaretinin sayısal olarak iletilmesi için yüksek bant genişliği gerekmektedir. Bu yüksek bant genişliği sorununun çözümü için bir çok video sıkıştırma tekniği geliştirilmiştir. Hareketli resim uzmanları grubu (MPEG – Moving Picture Expert Group) tarafından, videonun sayısal ortamlarda saklanması ve yayını için MPEG-1, MPEG-2 standartları geliştirilmiştir. Şu aşamada MPEG-2 sayısal video yayını (DVB – Digital Video Broadcasting) için yaygın olarak kullanılmaktadır.

Gelişen İnternet ve 3. Kuşak (3G – 3-Generation) mobil telefonlar teknolojileri üzerinden görüntü iletimi için daha yüksek sıkıştırmaya gereksinim duyulmaktadır. Ayrıca video içeriği ile etkileşim gereksinimi de artmaktadır. Bu amaçla MPEG tarafından nesne bazlı MPEG-4 video kodlama standardı öne sürülmüştür. MPEG-4, İnternet üzerinden video iletiminden, sayısal video yayına kadar birçok farklı ortam için kullanılabilir bir standart olarak kabul görmüştür.

Bu çalışmada MPEG-4 video kodlama standardının en temel ögesi olan video nesnelerinin (VOs) elde edilmesi için yeni bir bölütleme yöntemi önerilmiş ve gerçekleştirilmiştir. Önerilen yöntem blok bazlı arka plan kaydına dayanan zamansal bir bölütleme yöntemidir. Özellikle video konferans ve video telefon tipindeki kameranın sabit olduğu görüntüler için uygun olan bölütleme yöntemi, çok yavaş hareket içeren veya kısa süreli duraklayan nesnelere içeren görüntü dizinlerindeki başarılı bölütleme sonuçları ile literatürdeki diğer video bölütleme yöntemlerinden farklılık göstermektedir. Önerilen yöntemde eklenecek hareket dengeleme, gölge etkisinden kurtulma, bazı parametrelerin otomatik olarak bulunması gibi özellikler ile genel amaçlı bir bölütleme yöntemi ortaya çıkarılabilir.

## OBJECT SEGMENTATION FOR VIDEO

Oğuzhan URHAN

**Keywords:** Video Segmentation, Video Object Segmentation, Video Object Plane Segmentation, VOP, Moving Object Segmentation, Block Based Segmentation, Background Registration, MPEG-4.

**Abstract :** As in many areas of everyday life, digital technologies have been utilized for the processing of visual material such as pictures and video. Video information used to be recorded, processed and transmitted in analog form in traditional video broadcasting. Nowadays it is being recorded in digital form, compressed using digital coding techniques and transmitted over satellites.

For digital transmission of analog video information, high bandwidth is needed. In order to cope with high bandwidth requirements, many data compression techniques have been proposed. MPEG-1 and MPEG-2 standards have been developed by the Moving Picture Expert Group (MPEG) to store and broadcast digital video. Generally MPEG-2 is used for digital video broadcasting (DVB) currently.

Emerging 3G mobile and Internet technologies require efficient compression for the transmission of video information. Furthermore, interaction with video content is nowadays becoming more and more popular. The object based video coding standard MPEG-4 was developed for this purpose by MPEG. MPEG-4 has been proposed for video streaming over internet and digital video broadcasting.

In this work, a new video object segmentation method for obtaining video objects, which are the basic elements of MPEG-4, is developed and realized. The proposed method uses block based background registration, and is classified as a temporal segmentation method. It is particularly suited for videophone and videoconference applications where the camera is fixed. The method successfully segments sequences which contain slow movements and temporal object poses. With the addition of several features such as motion compensation, shadow elimination and adaptive parameter selection, the proposed technique can be developed into a general purpose video segmentation algorithm.

## ÖNSÖZ ve TEŞEKKÜR

Hızla gelişen ve maliyeti düşen tüm devre ve bilgisayar teknolojileri ile görüntü bilgisi hayatın hemen her alanında kullanılmaya başlanmıştır. Video ve resim bilgilerinin işlenmesi tıbbi cihazlardan, askeri donanımlara kadar çok geniş bir alanda uygulamalara sahiptir. Bu aşamada bilgisayar ve elektronik teknolojisinin iç içe geçtiği görüntü işleme bilim dalının önemini her geçen gün artmaktadır. Günümüzün vazgeçilmezlerinden olan bilişim teknolojisinin en büyük gereksinimlerinden biri olan görüntünün, kaydedilmesi, depolanması, ve iletilmesinde sayısal teknoloji kullanımı hızla artmaktadır. Sayısal teknoloji, beraberinde, sayısal biçime dönüştürülmüş görüntünün saklanması ve iletimi için sıkıştırılması gereksinimi ortaya çıkarmıştır.

Bu tez çalışmasında, sunduğu sıkıştırma verimliliği ve içerik temelli etkileşim olanakları ile bir çok uygulama için kabul görmüş MPEG-4 sayısal video kodlama standardının en temel ögesi olan video nesnelere elde edilmesinde kullanılabilecek yeni bir video nesne bölütlemesi yöntemi sunulmuştur. Gerçekleştirilen çalışmanın bu konuda çalışmakta olanlara ve araştırma yapmak isteyenlere katkısı olmasını dilerim.

Bana bu konuda tez çalışması yapmama olanak sağlayan, tezimin her aşamasında yönlendirmeleri ve güler yüzüyle sürekli destek olan değerli hocam Doç.Dr. Sarp ERTÜRK'e, yaşamım boyunca bana katkıda bulunan aileme ve hocalarıma, bu çalışmaya ayırdığım vakti anlayışla karşılayıp beni destekleyen müstakbel eşim Özge'ye teşekkür ederim.

Mayıs 2003, KOCAELİ

Oğuzhan URHAN

## İÇİNDEKİLER

ÖZET .....	ii
ABSTRACT.....	iii
ÖNSÖZ ve TEŞEKKÜR.....	iv
İÇİNDEKİLER .....	v
SİMGELER DİZİNİ ve KISALTMALAR.....	vii
ŞEKİLLER DİZİNİ.....	viii
BÖLÜM 1 : GİRİŞ.....	1
BÖLÜM 2 : SAYISAL VIDEOYA GİRİŞ.....	5
2.1. MPEG-1 Video Sıkıştırma Formatı .....	5
2.1.1. MPEG-1 veri yapısı .....	6
2.1.2. MPEG-1 I-resimi kodlama.....	8
2.1.3. MPEG-1 ara çerçeve (interframe) kodlama .....	9
2.1.4. MPEG-1 çerçeve kod çözümü .....	10
2.2. MPEG-2 Video Sıkıştırma Formatı .....	10
2.3. MPEG-4 Video Sıkıştırma Formatı .....	11
2.3.1. MPEG-4'ün özellikleri ve sunduğu işlevsellikler .....	11
2.3.2. MPEG-4'ün yapısı .....	13
2.3.3. Kenar kodlama .....	17
2.3.4. Hareket kestirim ve dengeleme araçları.....	18
2.3.5. Doku kodlama araçları .....	20
2.3.6. Hata düzeltme .....	21
2.3.7. Sprayt kodlama .....	22
2.3.8. Ölçeklenebilirlik.....	23

BÖLÜM 3 : VIDEO BÖLÜTLEME YÖNTEMLERİ .....	24
3.1. Giriş.....	24
3.2. Video Nesne Bölütlemesi Tekniklerine Genel Bakış .....	25
3.3. Hareket Bazlı 2-Boyutlu Video Nesnesi Bölütlemesi Teknikleri.....	27
3.3.1. Optik akış .....	28
3.3.2. Değişim algılama .....	28
3.4. Hareket Bazlı 3-Boyutlu Video Nesnesi Bölütlemesi Teknikleri.....	29
3.4.1. Hareketten yapı (structure from motion) .....	29
3.4.2. Parametrik yöntemler .....	30
3.5. Uzam-Zamansal Video Nesnesi Bölütlemesi Teknikleri .....	30
3.5.1. Uzamsal bölütleme yöntemleri .....	31
3.5.2. Zamansal bölütleme yöntemleri .....	34
BÖLÜM 4 : ÖNERİLEN VIDEO BÖLÜTLEME YÖNTEMİ .....	37
4.1. Giriş.....	37
4.2. Blok Bazlı Değişim Algılama .....	38
4.3. Nesne Kenarlarının Oluşturulması.....	43
4.4. Bölütleme Sonuçları.....	45
SONUÇLAR ve ÖNERİLER.....	50
KAYNAKLAR .....	52
KİŞİSEL YAYINLAR VE ESERLER .....	55
ÖZGEÇMİŞ .....	56



## SİMGELER DİZİNİ ve KISALTMALAR

B-VOP	Bidirectional VOP
CIF	Common Interface Format
DCT	Discrete Cosine Transform
DPCM	Differential Pulse Code Modulation
DVB	Digital Video Broadcasting
DVD	Digital Versatile Disc
EM	Expectation and Maximization
GOV	Group of Video Object Planes
ISO	International Organization for Standardization
ITU	International Telecommunication Union
I-VOP	Intra VOP
JPEG	Joint Picture Expert Group
MB	Macroblok
MBC	Model Based Coding
Mbps	Mega bit per second
MPEG	Moving Picture Expert Group (Hareketli Resim Uzmanları Grubu)
P-VOP	Predicted VOP
QCIF	Quad Common Interface Format
SAD	Sum of Absolute Difference
SIF	Standard Input Format
VCD	Video Compact Disc
VHS	Video Home System (Video Cassette Format ; Patented by JVC)
VLC	Variable Length Coding
VO	Video Object
VOL	Video Object Layer
VOP	Video Object Plane (Video Nesne Uzayı)
VS	Video Sequence

## ŞEKİLLER DİZİNİ

Şekil 2.1. MPEG-1 için 4:2:0 renk alt örnekleme formatı. ....	6
Şekil 2.2. MPEG-1 sıradüzensel veri yapısı. ....	6
Şekil 2.3. Ara resimler arası geçişler. ....	7
Şekil 2.4. I-Resim kodlayıcı. ....	8
Şekil 2.5. MPEG-4'ün sunduğu işlevsellikler. ....	12
Şekil 2.6. Örnek bir MPEG-4 bit dizisinin mantıksal yapısı. ....	15
Şekil 2.7. Video nesnelerinin kodlanması ve kodunun çözülmesi. ....	16
Şekil 2.8. Video nesne uzayının kod çözümü. ....	17
Şekil 2.9. MPEG-4'ün video nesne uzaylarını kodlarken desteklediği kipler. ....	18
Şekil 2.10. Video nesne uzayı doku kodlaması. ....	21
Şekil 3.1. "Akiyo #28" in bölütlenme öncesi ve sonrası durumu. ....	24
Şekil 3.2. Hareketli nesne bölütlemesi yöntemlerinin sınıflandırılması. ....	26
Şekil 4.1. Önerilen video nesnesi bölütleme yöntemin blok şeması. ....	38
Şekil 4.2. Piksel bazlı çerçeve farkı. ....	39
Şekil 4.3. Blok bazlı çerçeve farkı. ....	40
Şekil 4.4. Blok farkı bazlı değişim algılama. ....	41
Şekil 4.5. Blok bazlı morfoloji öncesi bölütleme sonuçları. ....	42
Şekil 4.6. Blok bazlı morfoloji. ....	43
Şekil 4.7. Kenar bloklarda piksel bazlı değişim algılama. ....	44
Şekil 4.8. Nesne kenarlarında polinomsal eğri uydurma. ....	44
Şekil 4.9. "Akiyo" görüntü dizini için bölütleme sonuçları. ....	46
Şekil 4.11. "Missam" dizini bölütleme sonuçları. ....	48
Şekil 4.12. "Grandma" dizini bölütleme sonuçları. ....	49

## BÖLÜM 1 : GİRİŞ

Her geçen gün gelişen sayısal teknoloji ile, halen bir çok alanda kullanılmakta olan analog bilgi yerine sayısal bilgi kullanımı giderek artmaktadır. Bu doğrultuda görüntü sistemlerinin yapısı da analog teknolojilerden sayısal teknolojilere doğru kaymaktadır. 10 yıl öncesine kadar video bilgisi kaydının, depolanmasının ve iletiminin neredeyse tamamı analog olarak yapılmıydı. Örneğin geleneksel televizyon yayınında görüntü, analog olarak video kasetlere kaydedilip, analog genlik kiplenimi (amplitude modulation) kullanılarak havadan iletilmektedir. Günümüzde sayısal kayıtlar yapılmakla birlikte kullanılan televizyon alıcıları hala sadece analog yayın alabildikleri için yayınının iletiminin çoğu analog olarak yapılmaktadır.

Analog video sinyali zaman değişkenine bağlı tek boyutlu elektriksel bir işaret olarak ele alınabilir. Yapı itibariyle iki boyutlu olan görüntünün taranması sonucu tek boyutlu işaret elde edilebilmektedir. En çok kullanılan tarama yöntemleri; sürekli tarama (progressive scanning) ve geçmeli taramadır (interlaced scanning). Avrupa'da kullanılan geleneksel televizyon yayında analog PAL (Phase Alternation Lines) video işareti tek bir görüntü veya çerçeve (frame) için 625 satır içermektedir ve her bir çerçeve iç içe geçmiş iki alandan (field) oluşmaktadır. Geçmeli tarama kullanılan bu tip video işareti; öncelikle tek satırlardan oluşan görüntü ekrana basılmakta, sonra çift satırlar ekrana basılmaktadır. Belirli şartlar altında, insan beyni, gözün retinasına odaklanmış hareketli bir görüntü ile çok hızlı bir durağan çerçeve dizisi arasındaki farkı anlayamamaktadır. Psiko-görsel araştırmalar insan gözünün yüksek hızda (saniyede 50 kereden fazla) ardı ardına gelen çerçeveler arasındaki geçişin farkına varamadığını göstermiştir.

Analog bir PAL çerçevesi, 400.000 den fazla resim bileşeni (picture element-pel, pixel) içermektedir. Yine PAL sisteminde çerçeve tazeleme hızınının 25 olduğu göz önüne alınırsa bir saniyelik bir görüntü için 11 milyondan fazla piksel iletilmesi

gerekmektedir. Her bir pikselin üç temel renk bileşeni (kırmızı, yeşil, mavi – Red,Green,Blue) ile 8'er bit temsil edildiği durumda her bir piksel için 24 bit gerekmektedir. Bu durumda gerçek zamanlı görüntü iletimi için 253Mbps'lik bir iletim hızına ihtiyaç duyulmaktadır. Sayısal görüntüde bilgi fazlalığından oluşan bu yüksek bant genişliği gereksinimi verilerin sıkıştırılmasını ön plana çıkarmaktadır.

Sayısal video halen kullanılmakta olan bir çok uygulamada analog videonun yerini almaktadır. Buna önemli bir örnek sayısal televizyon yayınlarının başlaması ile birlikte hızlı bir şekilde yayılmasıdır. Diğer bir önemli örnek ise analog video kasetlerin yerini VCD, DVD'lerin almasıdır. MPEG-1 ve MPEG-2 bu ortamların (media) kabulünde anahtar teknoloji olmuştur.

1988 yılında Uluslararası Haberleşme Birimine (ITU) bağlı Uluslararası Standartlar Örgütü (ISO) video, film, grafik ve metinsel materyallerin sıkıştırılmış olarak sunulması için uluslararası olarak tanınan bir dünya standardı üzerinde bir görüş birliği oluşturmak amacıyla, Hareketli Resimler Uzmanları Grubu'nu (MPEG) kurmuştur. Bu komitenin amacı, karmaşık işlevlerin çoğunu alıcıdan ziyade vericiye yükleyen daha basit, ekonomik ve esnek bir standart geliştirmektir (Long 1995). 1990 yılında, sayısal televizyon yayını göz önüne alınmadan, çoklu-ortam kaynaklarının sıkıştırılmış sayısal sunumunu ele almak amacıyla 1.5Mbps sabit bit hızı ve sürekli tarama kullanan MPEG-1 standardı (Benzetim Modeli 1- Simulation Model 1) ortaya atıldı ve 1992 yılında resmi olarak kabul edildi. MPEG-1 formatı kullanılarak 1.2Mbps ile sıkıştırılan ve açılan CIF (Common Interface Format) formatındaki görüntünün kalitesinin, VHS formatında kaydedilen analog videonun kalitesi ile aynı (veya VHS'den daha iyi) olduğu görülmüştür (Tekalp 1995). Ancak sabit bit oranı kullanan bu standardın kalitesi bazı uygulamalar için yeterli olmadığından daha kaliteli sıkıştırma yöntemleri için çalışmalar yürütülmüş ve bunun neticesinde 1995 yılında 2-20 Mbps değişken bit hızı ve geçmeli tarama kullanan MPEG-2 standardı oluşturulmuştur. MPEG-2, sayısal video yayını (DVB – Digital Video Broadcasting) için genel olarak kabul görmüştür ve yaygın olarak kullanılmaktadır.

Gün geçtikçe gelişen haberleşme, tüketici elektroniği, bilgisayar endüstrisi gibi teknolojiler artık çoklu ortam öğelerini de kullanır hale gelmiştir. Geniş bir bakış

açısı ile çoklu ortam; video bilgisini de içeren farklı kaynaklardan elde edilen bilgi ile etkileşim için bir iskelet olarak tanımlanabilir. Bir çoklu ortam standardının birçok uygulama için destek sağlaması beklenir ve bu uygulamalar birbirlerinden çok farklı özel gereksinimlere ihtiyaç duyabilir. Görsel bilgi içeren uygulamalar aşağıda verilen özelliklere göre birkaç grupta toplanabilir (Ebrahimi and Horne 2000).

- i) Veri tipi (imge, stereo imge, video)
- ii) Kaynak tipi (doğal imgeler, bilgisayarda oluşturulmuş imgeler, yazılar, grafikler, tıbbi imgeler)
- iii) Haberleşme tipi (tek noktadan tek noktaya iletişimde çok noktadan çok noktaya iletişime kadar geniş bir alanda)
- iv) İstenilen işlevsellik tipi (nesne değiştirme (manipulation), çevrimiçi düzenleme, sürekli iletim, hatalara karşı dayanıklılık)

MPEG-2 standardının analog televizyon yayından daha kaliteli görüntü sunmanın ve bu yayını sayısal olarak yapmanın dışında hemen hemen bir farkı bulunmaktadır. MPEG-1 ve MPEG-2, analog televizyon yayınında olduğu gibi çoklu ortam uygulamalarının gerektirdiği esneklikleri sağlayamamaktadır. Bundan dolayı MPEG (Moving Picture Experts Group) geniş erimli çoklu ortam uygulamaları için ortak bir platform oluşturmak amacıyla 1993 yılında başladığı çalışmalar sonucunda 1999 yılında MPEG-4 standardının ilk sürümünü öne sürmüştür.

MPEG-4 önceki video sıkıştırma formatlarından daha verimli olmasının yanı sıra sunduğu içerik temelli işlevsellikler ile sayısal televizyon yayınından, mobil çoklu ortam uygulamalarına, televizyon prodüksiyonlarından, oyunlara ve Internet üzerinden video iletimine kadar bir çok farklı alandaki uygulamalarda kullanılabilir.

MPEG-4'ün ana yapısını oluşturan içerik temelli işlevsellikler video nesnesi kavramının ortaya çıkması ile oluşturulmuştur. Bir video nesnesi bazı kriterler altında birlikte gruplanabilecek video bölgeleri olarak tanımlanabilir. Örneğin sabit bir arka plan üzerinde konuşma yapan bir sunucu bir video nesnesi olarak ele alınabilir. Bu video nesnesine ait biçim, hareket ve doku bilgisi ayrı bir video nesne

katmanında (video object layer – VOL) kodlanır ve kodu çözülür. Bu yapı, istenilen tek bir video nesnesine erişilmesine ve içerik temelli sunuma olanak sağlaması açısından önemlidir. MPEG-4'ün video kodlama yapısı ile ilgili ayrıntılı bilgi 2. bölümde verilmiştir.

Yukarıda anlatılan nedenlerden dolayı video nesne uzayı bölütlemesi (video object plane segmentation, VOP segmentation) veya hareketli nesne bölütlemesi olarak bilinen video nesnelерinin elde edilmesi içerik temelli sunum ve işlevselliklerin sağlanması için temel şarttır. Bu amaçla otomatik olarak ve kullanıcı desteği ile video nesnelерinin elde edilmesini sağlayan bir çok video nesne bölütlemesi tekniği geliştirilmiştir. Bu teknikler ile ilgili ayrıntılı bilgi 3. bölümde verilmiştir.

Bu çalışmada video nesnelерinin bölütlenmesi için özellikle video konferans ve video telefon tipi sabit arka plana sahip uygulamalarda kullanılabilecek bir bölütleme yöntemi önerilmiştir. Blok bazlı arka plan kaydına dayanan ve değişim algılama kullanılan yöntem uzam-zamansal (spatio-temporal) bir bölütleme yöntemi olarak ele alınabilir. Önerilen yöntemde klasik değişim algılama temelli yöntemlerin aksine sadece bir önceki çerçeve ile piksel bazlı farkın eşiklenmesi kullanılmamış, daha eski çerçeveler de kullanılarak yavaş hareket içeren görüntü dizinlerinin bölütlenmesi sağlanmıştır. Çerçeve farkı alınırken gürültü etkisini azaltabilmek için bu işlem blok bazlı yapılmıştır. Daha sonra blok bazlı ön plan ve arka plan bölgelerini elde edebilmek için blok bazlı morfolojik kapatma (morphologic closing) işlemi uygulanmıştır. Bir kez arka plan olarak işaretlenen bloklar için daha eski çerçevelere bakılmayıp, arka plan imgesindeki blok ile o anki imgedeki ilgili blok karşılaştırılarak bloğun durumuna karar verilmektedir. Ön plan olan bloklar için ise önceden belirlenmiş sayıda çerçeve geri gidilerek bloğun durumu belirlenir. Elde edilen ön plan bölgesi bloklar halinde olduğundan nesne kenarlarına daha fazla yaklaşabilmek için, kenar bloklarda piksel bazlı eşikleme kullanılmış ve son olarak daha yumuşak nesne kenarlarını elde etmek için nesne kenarlarında polinomsal eğri uydurması kullanılmıştır. Önerilen bölütleme yöntemi ilgili ayrıntılı bilgi, bölütleme sonuçları ve öneriler 4.bölümde verilmiştir

## BÖLÜM 2 : SAYISAL VIDEOYA GİRİŞ

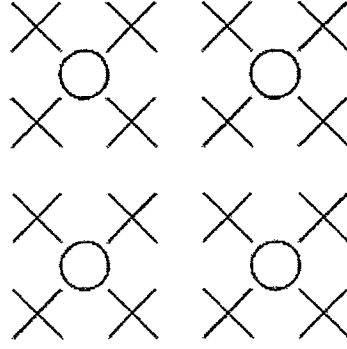
### 2.1. MPEG-1 Video Sıkıştırma Formatı

Bir ISO standardı olan MPEG-1, kodlanmış görüntü dizinlerinin sunumu ve bir kod çözme yönteminin yapısını standartlaştırmış ve CIF video formatının ve bu video ile ilgili sesin 1.5Mbps ile sayısal kayıt ortamlarında saklanmasına olanak sağlamıştır (Tekalp 1995). Önerilen yapı, hareket dengelemesi (motion-compensation), ayrık kosinüs dönüşümü, niceme ve değişken uzunluklu kodlama gibi işlemleri içermektedir. Bazı hangi algoritmalarla yapılacağı özel olarak belirtilmemiş, böylece tasarımcılara kodlayıcı tasarımında büyük esneklikler sağlamıştır.

MPEG-1 sadece sürekli tarama (noninterlaced or progressive scanning) yöntemini desteklemektedir. Hedeflenen 1.5Mbps'lik bit oranına ulaşmak için genellikle giriş video işareti MPEG-1 standart giriş formatına (Standart Input Format – SIF) dönüştürülür. Sıkıştırma bakımından RGB formatından daha verimli olması nedeni ile YUV renk uzayında temsil edilen görüntü işaretleri üzerinde işlem yapılır. MPEG-1 SIF formatında 4:2:0 renk alt örnekleme formatı kullanılır. Bu alt renk örnekleme formatı Şekil 2.1.'de gösterilmiştir.

Işıklılık kanalı 352 x 240 pikseldir ve genellikle saniyede 30 çerçeve gösterilir. Işıklılık ve renk bileşenlerinin her bir pikseli 8 bit ile temsil edilmektedir. Resim büyüklüğü, tazeleme oranı gibi parametreler söz diziminde belirlenebilmektedir ancak donanımsal uygulamalara yardımcı olması için satır sayısı en fazla 576 ile, bir satırdaki piksel sayısı en fazla 720 ile ve tazeleme oranı saniyede en fazla 30 çerçeve ile sınırlandırılmıştır (Tekalp 1995).

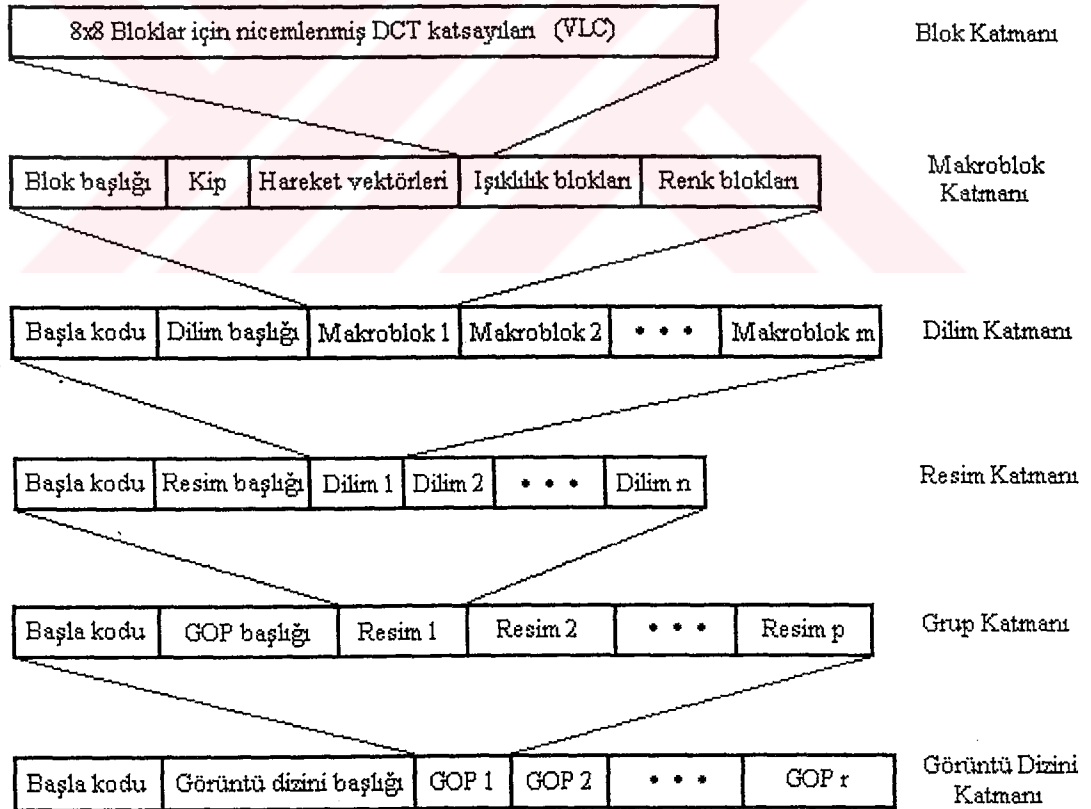




Şekil 2.1. MPEG-1 için 4:2:0 renk alt örnekleme formatı.

### 2.1.1. MPEG-1 veri yapısı

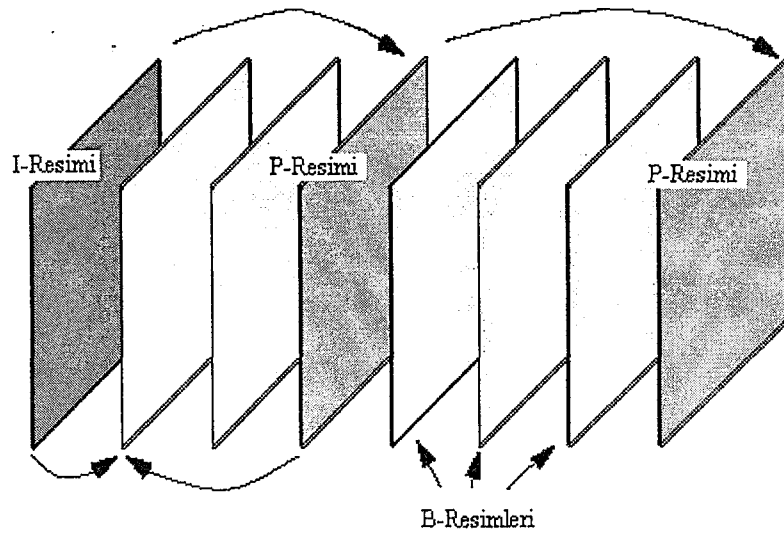
MPEG-1, kod çözücünün verileri belirli bir şekilde yorumlamasına olanak sağlayan altı katmanlı sıra düzensel (hierarchical - hiyerarşik) bir veri yapısı kullanır. Bu sıra düzensel veri yapısı Şekil 2.2.'de gösterilmiştir.



Şekil 2.2. MPEG-1 sıradüzensel veri yapısı.



- i) Görüntü dizinleri (sequences) birkaç grup resimden oluşmaktadır.
- ii) Resimler grubu (Group of Pictures – GOP) resimlerden meydana gelmektedir.
- iii) Resimler dilimlerden (slice) meydana gelir ve birinci kodlama ünitesidir. Bir resim bir adet ışıklılık (luminance) ve 2 adet renk (Cr, Cb) olmak üzere üç dikdörtgensel matristen oluşmaktadır. Sıkıştırma kipine (mode) bağlı olarak 4 tip resim bulunmaktadır. I-resimler (Intra-Pictures), JPEG benzeri bir yöntemle DCT olarak kodlanırlar. Bu resimler rasgele erişim noktaları olarak görüntü dizisine yardımcı olmaktadır. P ve B olmak üzere iki tip ara çerçeve (interframe) kodlanmış resimler vardır. Bu resimlerde hareket dengelenmiş kestirim hataları DCT olarak kodlanır. P-resimlerinde sadece ileri tahmin kullanılır ve kodlama bir önceki I veya P resmine göre yapılır. B-resimlerinin tahmini ise diğer I ve P-resimlerine bağlı olarak ileri, geri veya iki yönlü olabilir. D-resimleri sadece her bloğa ilişkin DC bileşenini içerir ve imge dizisi içinde çok düşük bit oranlarında tarama yapmak için kullanılır. Bir grup resimdeki (GOP) I, P ve B resimlerinin sayısı kullanılacak uygulamanın erişim zamanı, bit oranı gereksinimi gibi özelliklerine bağlıdır. Resimler arası geçişler Şekil 2.3’de verilmiştir. Oluşacak hataları engellemek için art arda gelen en fazla 12 resimden 1 tanesi mutlaka I-resmi olmalıdır.

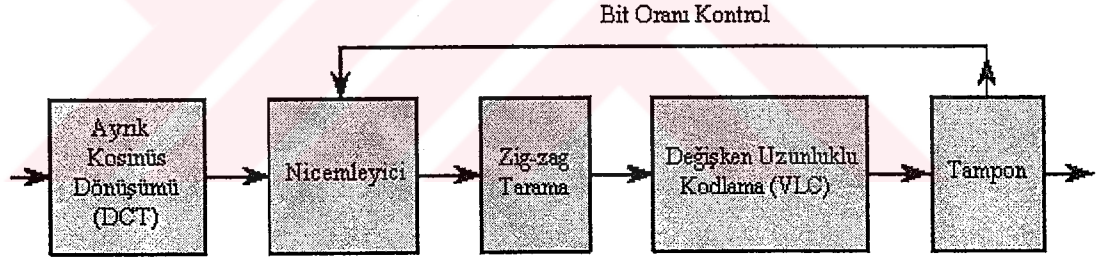


Şekil 2.3. Ara resimler arası geçişler.

- iv) Dilimler satır satır tarama sonucu meydana gelen makro blok (MB) dizinlerinden oluşmuştur. Bu dilim yapısı temel olarak hata kurtarmaya olanak sağlamaktadır.
- v) Makro bloklar genellikle 8x8 boyutunda 4 adet resim bloğundan oluşmaktadır. Işıklılık için 16x16 örnek, renk bilgileri için ise 8x8 örnek içerir.
- vi) Bloklar 8 x 8 piksel boyutundadır ve MPEG-1 veri yapısının en küçük DCT birimidir.

### 2.1.2. MPEG-1 I-resimi kodlama

MPEG-1'de I-resimlerinin kodlaması öncelikle bloklara ayrılan resmin ayrık kosinüs dönüşümüne (DCT) tabii tutulması ile başlar. Elde edilen ayrık kosinüs dönüşümü katsayıları önem derecelerine göre nicemleyici ile ağırlıklandırılır. Son olarak zig-zag tarama sonucu elde edilen veriler değişken uzunluklu kodlama kullanılarak kodlanır. Şekil 2.4.'de I-resim kodlama yapısı verilmiştir.



Şekil 2.4. I-Resim kodlayıcı.

- Ayrık Kosinüs Dönüşümü : Yapılan araştırmalar sonucunda bir bloğun enerjisini birkaç katsayıda yoğunlaştırma eğiliminde olan ayrık kosinüs dönüşümünün iyi sonuç verdiği ortaya çıkmıştır. Bu dönüşümün bir diğer özelliğe de ters ayrık kosinüs dönüşümünün sayısal uygulamasının kolay olmasıdır. 8x8 piksellik bloklar ayrık kosinüs dönüşümü ile frekans uzayında ifade edilmektedir. En düşük frekanslı bileşen DC katsayıdır ve bloğun ortalama parlaklığını gösterir (Ely 1995). Düşük frekans bileşenleri resmin genel hatlarını, yüksek frekans bileşenleri ise resim ile ilgili ayrıntıları gösterir.

- ii) Nicemleyici : Blok, frekans uzayına dönüştürülüp, dönüşüm katsayıları elde edildikten sonra, gösterdiği uzamsal frekans dağılımına göre farklı nicemleme matrisleri kullanılarak nicemlenir. Burada amaç gönderilecek bit sayısını olabildiğince azaltmaktır. İnsan görsel sistemi yüksek frekans bileşenlerine daha az duyarlı olduğu için bu bileşenler nicemlenirken yüksek hata oluşmasına izin verilir.
- iii) Zig-zag Tarama ve Değişken Uzunluklu Kodlama : Nicemlenmiş DC katsayılar arasındaki artıklık DPCM ile azaltılır ve elde edilen işaret değişken uzunluklu kodlama (VLC) ile kodlanır. Nicemlenmiş AC katsayılar zig-zag tarama kullanılarak [run, level] şeklinde çiftli forma dönüştürülür ve daha sonra değişken uzunluklu kodlama ile kodlanır. Burada “level”, sıfır olmayan katsayı değeri, “run” ise o anki katsayıdan önceki sıfır olan katsayıların sayısıdır. Kodlamada Huffman benzeri bir kod tablosu kullanılmaktadır.

### 2.1.3. MPEG-1 ara çerçeve (interframe) kodlama

Art arda gelen çerçevelerde genellikle değişimin az olacağı gerçeğinden yararlanmak için MPEG-1’de zamansal kestirim kipleri kullanılmıştır. Zamansal kestirim kiplerinden elde edilen kestirim hataları ayrık kosinüs dönüşümü ile kodlanır. MPEG-1’de ileri kestirim ve iki yönlü kestirim olmak üzere iki kestirim kipi bulunmaktadır.

- i) P-Resimleri : P-resimleri, bir önceki I veya P-resimini referans alarak, hareket dengelenmiş ileri tahmin kodlamasına olanak sağlar. Bu aşamada seçilen ara çerçeve kodlama türüne göre, tahmin hatalarının ayrık kosinüs dönüşümü ile kodlanıp kodlanmayacağına, hareket düzeltme yapılıp yapılmayacağına, uyarlamalı nicemleme (adaptive quantization) kullanılıp kullanılmayacağına, sadece hareket vektörlerinin iletilip iletilmeyeceğine veya ilgili bloğun önceki çerçeveye göre yer değiştirip değiştirmemesine göre tekrar gönderilip gönderilmeyeceğine karar verilir (Tekalp 1995).

- ii) B-Resimleri : B-resimleri, yüksek sıkıştırmaya izin verir ancak daha fazla

çerçevenin depolanmasına gerek duyar ve bir miktar gecikmeye yol açar. Tahmin olarak sadece önceki çerçeve, sadece sonraki çerçeve, veya her ikisini birden kullanılabilir. Kodlanan blok ile referans blok arasındaki hareketi elde etmek için blok uyumlaması (block matching) kullanılır. MPEG-1’de sabit bir blok uyumlaması yöntemi kullanılması şart koşulmamış, yöntem seçimi tasarımcıya bırakılmıştır (Ely 1995).

#### **2.1.4. MPEG-1 çerçeve kod çözümü**

MPEG-1 olarak kodlanmış video işaretinin tekrar elde edilmesi kodlayıcı tarafından yapılan işlemlerin tersinin yapılması ile sağlanır. Öncelikle gelen bit dizisi tekilleyerek DCT katsayıları ve hareket vektörleri, makro blok tipi gibi yan bilgileri elde edilir. Ek olarak B-resimlerinin kodlanmasında iki referans çerçeve kullanıldığından, kod çözücüde de iki çerçeve depolamaya gerek duyulmaktadır (Tekalp 1995).

#### **2.2. MPEG-2 Video Sıkıştırma Formatı**

MPEG-1 kullanılarak 1.2 Mbps ile sıkıştırılmış videonun kalitesinin birçok uygulama için yeterli olmadığı görülmüş ve yapılan araştırmalarda 4-6 Mbps ile sıkıştırılmış videonun (CCIR 601) çok iyi kalitede sonuç verdiği gözlenmiştir. Bunun üzerine MPEG-1’in uyumlu bir uzantısı olan ve farklı uygulamalar için birçok bit oranında (2 - 20Mbps) ve çözünürlükte kodlamaya olanak sağlayan MPEG-2 standardı geliştirilmiştir. MPEG-2’nin en önemli özellikleri

- Geçmeli taramaya, yüksek çözünürlüğe ve değişik renk alt örnekleme formatlarına izin verme
- Ölçeklenebilir bit dizilerinin kullanımına olanak sağlama
- Geliştirilmiş nicemleme ve kodlama seçenekleri sağlama

olarak sıralanabilir (Tekalp 1995).

Bir makro blok 8x8 boyutunda 4 ışıklılık bloğu ve bu blokla ilgili renk bloklarını içerir. MPEG-2, MPEG-1’de kullanılan 4:2:0, sadece yatay doğrultuda renk alt örnekleme yapan 4:2:2 ve alt renk örnekleme kullanmadan her ışıklılık bileşeni için renk bileşeni de kullanan 4:4:4 renk alt örnekleme formatlarını destekler.

### **2.3. MPEG-4 Video Sıkıştırma Formatı**

MPEG-1 ve MPEG-2 standartları bir çok uygulama için yeterli görsel kalite sağlamasına rağmen çoklu ortam uygulamalarının gereksinimlerini sağlayacak yeterli esnekliğe sahip değildir. Bundan dolayı Hareketli Resim Uzmanları Grubu (MPEG) geniş bir alana yayılmış olan çoklu ortam uygulamaları için bir standart geliştirmek üzere 1993 yılından itibaren çalışmaya başlamıştır.

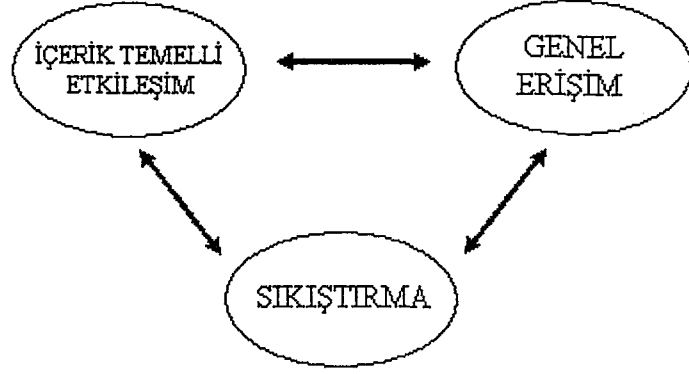
MPEG-1 ve MPEG-2’de sayısal video, analog video ile aynı işlevselliklere sahiptir, yani bilgi analog yerine sayısal formda sadece daha iyi kalitede sunulmaktadır ve içerik sunumunda herhangi bir farklılık yoktur. Ancak, eğer içerik de sayısal olarak sunulabilirse, kullanıcıya belirli bir içeriği izleme, erişme, değiştirme gibi birçok işlevsellik kolayca sağlanabilir. MPEG-4 bu gibi işlevsellikleri sağlamak için gerekli olan anahtar teknolojileri destekler.

Bu bölümde MPEG-4’ün iletimi, ses özellikleri v.b. gibi konulara girilmeyip sadece video kodlama aşaması incelenmiş ve Ebrahimi et al (2000) tarafından önerilen yapı izlemiştir.

#### **2.3.1. MPEG-4’ün özellikleri ve sunduğu işlevsellikler**

MPEG-4 video kodlama standardı, birkaç çeşit fonksiyonelliği destekleyerek çeşitli uygulamalara imkan sağlayan bir grup araçtan oluşur. MPEG-4’ün en önemli özellikleri Şekil 2.5.’de görüldüğü gibi üç gruba ayrılabilir.

**TC YÜKSEKÖĞRETİM KURULU  
EĞİTİM ARAŞTIRMA MERKEZİ**



**Şekil 2.5.** MPEG-4'ün sunduğu işlevsellikler.

- i) Sıkıştırma Verimliliği : Sıkıştırma verimliliği, sayısal televizyon yayına olanak sağlayan ve DVD teknolojisinin temeli olan MPEG-1 ve MPEG-2 de olduğu gibi en önemli özelliklerden biridir. Arttırılmış kodlama verimliliği ve birçok veri dizisinin aynı anda kodlanması MPEG-4 standardını kullanan uygulamaların kabul edilebilirliğini arttıracaktır.
- ii) İçerik Temelli Etkileşim : Video nesnelerinin video çerçeveleri yerine video nesnesi olarak kodlanması ve sunulması içerik temelli uygulamalara olanak sağlar ve bu MPEG-4'ün en önemli yeniliğidir. Nesnelerin etkin sunumunu esas alan nesne düzeltme (manipulation), bit dizisi düzeltme (editing) ve nesne temelli ölçekleme yeni içerik etkileşimlerine imkan verir.
- iii) Genel Erişim : MPEG-4'ün gürültüye eğilimli ortamlardaki sağlamlığı, mobil ağlar ve kablolu bağlantıları da içeren geniş bir ortamdan kodlanmış içeriğe erişime olanak sağlar. Buna ek olarak nesne temelli zamansal ve uzamsal ölçekli kodlama, kullanıcıya, kullanılabilir bant genişliği, bilgisayar kapasitesi ve güç harcama gibi sınırlı kaynakların nerede kullanılacağına karar vermesine olanak sağlar.

Yukarıda bahsedilen işlevselliklerin bazılarını desteklemek için MPEG-4, gelişigüzel (arbitrarily) biçimlendirilmiş video nesnelerini sunma yeteneğine sahip olmalıdır. Böylece her bir nesne değişik parametreler ve değişik kalitelerde kodlanabilir. MPEG-4'de video nesnelerinin biçimleri ikili veya gri-seviyeli (alpha) uzayda sunulabilir, doku ise nesnenin şeklinden ayrı olarak kodlanır. Düşük bit oranı gerektiren uygulamalarda MPEG-1 ve MPEG-2'de olduğu gibi çerçeve tabanlı



kodlama kullanılabilir. Hatalara karşı dayanıklılığı artırmada hızlı yeniden eş zamanlama (resynchronization) ve etkin şekilde hata kurtarmaya olanak sağlamak için bit dizisi seviyesinde özel koşullar dikkate alınır.

MPEG-4 görsel kodlama standardı aşağıda verilen üç değişik bit oranı için en iyi şekilde ayarlanmıştır.

- 64 kbit/s altında
- 64 kbit/s – 384 kbit/s
- 384 kbit/s – 4 Mbit/s

Yüksek kaliteli uygulamalar için, yüksek bit oranları önerilmekle birlikte, bu uygulamalar için düşük bit oranlarında kullanılan araçlar grubu ve bit dizisi yapısı kullanılmaktadır.

MPEG-4 hem sürekli tarama hem de geçmeli tarama kullanımına olanak sağlar. Desteklenen renk alt örnekleme formatı 4:2:0 dır. Bu alt örnekleme formatında Cr ve Cb örnekleri hem yatay hem düşey doğrultuda ışıklılık bileşenin yarısı kadardır. (MPEG sıkıştırma formatlarında sıkıştırma verimliliği daha yüksek olduğu için RGB renk uzayı yerine YUV renk uzayı kullanılmıştır. Buradaki U, Cb'yi , V ise Cr'yi temsil etmektedir).

### **2.3.2. MPEG-4'ün yapısı**

MPEG-4'ün öne sürdüğü temel görüş nesne-tabanlı sunumun temelini sağlayan sesli-görsel nesnelere. Böyle bir sunum etkileşimli uygulamalar için oldukça uygundur ve video ortamının içeriğine doğrudan erişim sağlar.

Bu kısımda, temel olarak doğal video nesnelere hakkında bilgi verilecektir. Bir video nesnesi ölçeklenebilir kodlamaya olanak sağlamak için bir veya birkaç katmandan oluşabilir. Ölçeklenebilir yapı, videonun tek başına temel katmandan veya buna eklenen birkaç geliştirme (enhancement) katmanıyla, katmanlı bir şekilde

oluşturulmasına olanak sağlar. Bu yapı, uygulamalara birçok değişik bant genişliği ve/veya hesapsal karmaşıklık isteklerine göre tek bir MPEG-4 video bit dizisi oluşturma olanağı sağlar. Durağan bir imgenin iki veya üç boyutlu nesnelere dönüştürülmesi yüksek seviyede ölçeklenebilirliğin istendiği özel bir durumdur. Bu işlevselliğe imkan vermek için MPEG-4 durağan dokunun dalgacık dönüşümü kullanılarak kodlanmasına imkan veren özel bir kipi destekler.

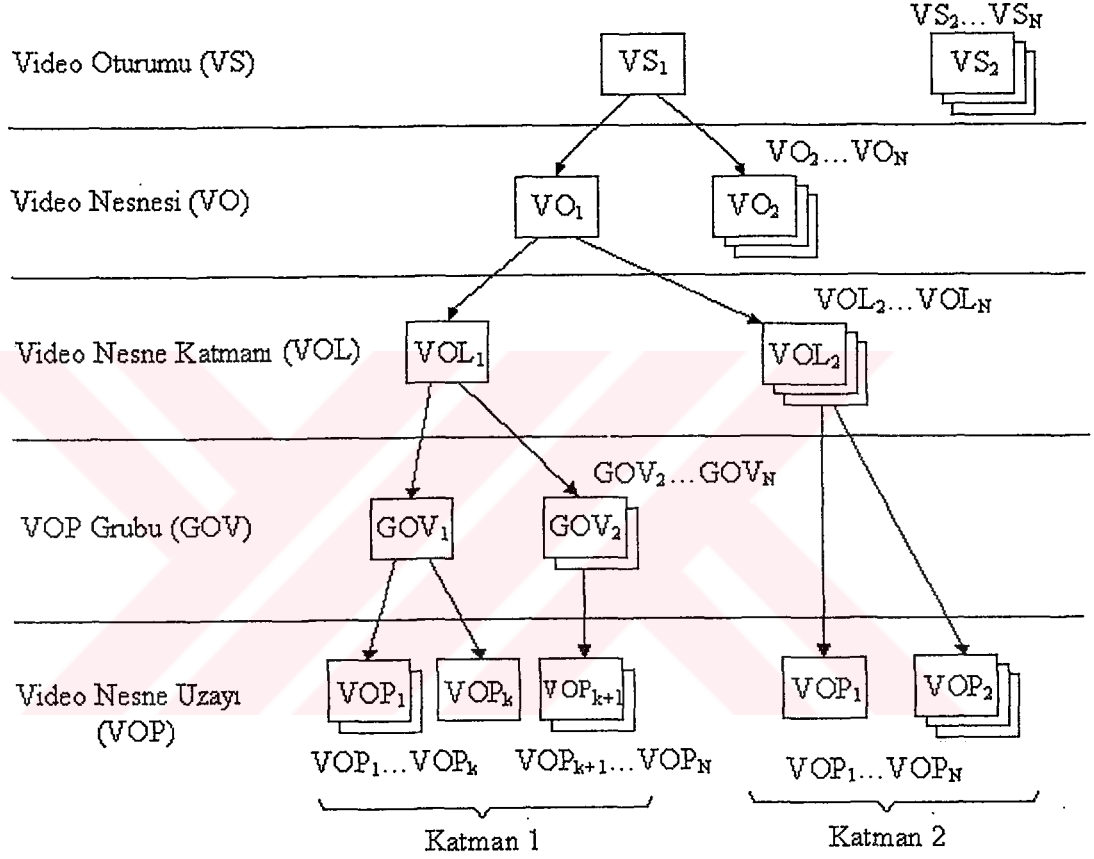
Bir MPEG-4 görsel ortamı bir veya daha fazla video nesnesinden oluşabilir. Her bir video nesnesi şekil, hareket ve doku gibi uzamsal ve zamansal bilgiler ile nitelendirilir. Video nesnelere için model bazlı kodlama (model based coding, MBC) imkanı da tanınmaktadır. Bazı uygulamalarda video nesnesi oluşturmanın zorluğu veya bunun getirdiği yük nedeni ile video nesnelere istenmeyebilir. MPEG-4 bu tip uygulamalarda MPEG-1 ve MPEG-2'de kullanılan dikdörtgen şeklindeki çerçevenin kodlanmasına olanak sağlar.

Görsel bilgi içeren bir MPEG-4 bit dizisi, görsel ortamın Şekil 2.6.'da gösterildiği gibi sıra düzensel bir yapıda ifade edilmesini sağlar. Bu yapının her bir seviyesine bit dizisi bazında başlama kodu olarak adlandırılan özel kodlar ile erişilebilir. Bu sıra düzensel yapıdaki her bir seviye aşağıda açıklanmıştır.

- Görsel Nesne Görüntü Dizini (Visual Object Sequence - VS) : 2-Boyutlu veya 3-boyutlu gerçek veya yapay nesnelere ve onların geliştirilmiş katmanlarını içeren bütün bir MPEG-4 manzarasıdır.
- ii) Video Nesnesi (Video Object – VO) : Bir video nesnesi ortamdaki belirli bir 2-boyutlu nesneye karşılık gelir. En basit biçimde bu dikdörtgensel bir çerçeve veya bir nesneye veya ortamın arka planına karşılık gelen gelişigüzel şekillendirilmiş bir nesne olabilir.
- iii) Video Nesne Katmanı (Video Object Layer – VOL) : Her bir video nesnesi uygulamaya bağlı olarak, ölçeklenmiş (çok katmanlı) veya ölçeklenmemiş (tek katmanlı) şekilde video nesne katmanında sunulmak üzere kodlanabilir. Video nesne katmanı ölçeklenebilir kodlamayı desteklemeye imkan verir. Bir video nesnesi



zamansal ve uzamsal ölçekleme kullanılarak kaba hatlardan iyi çözünürlüğe kadar kodlanabilir. Böylece kullanılabilir bant genişliği, bilgisayar gücü ve kullanıcının tercihi gibi parametrelere göre istenilen çözünürlük kod çözücünde sağlanabilir. MPEG-4'ün, işlevselliklerini taşıyan veya sadece kısa başlık bilgisi taşıyan iki tip video nesne katmanı vardır. Sadece kısa başlık bilgileri taşıyan yapı H.263 protokolü ile uyumludur.



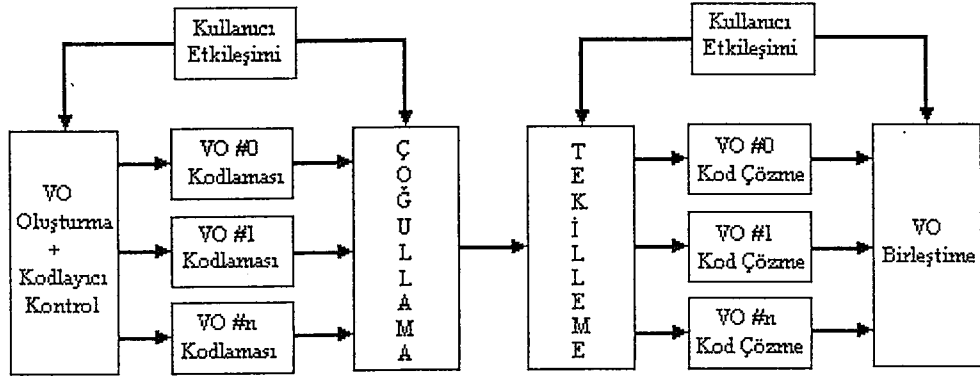
Şekil 2.6. Örnek bir MPEG-4 bit dizisinin mantıksal yapısı.

- iv) Video Nesnesi Uzayı Grubu (Group of Video Object Planes - GOV) : Video nesne uzayı grupları, video nesne uzaylarını gruplandırır. Video nesne uzayı grupları vasıtasıyla, bit dizisinde, video nesne uzaylarının birbirinden bağımsız olarak kodlandığı noktalar bulunabilir ve bu özellik bit dizisi içinde rasgele erişim noktalarına ulaşmaya olanak sağlayabilir. Video nesne uzayı grupları isteğe bağlıdır.

- v) Video Nesne Uzayı (Video Object Plane - VOP) : Bir video nesne uzayı, bir video nesnesinin zamanda örneklemesidir. Video nesne uzayları birbirlerinden bağımsız olarak veya birbirlerine bağımlı şekilde hareket dengeleme (compensation) kullanılarak kodlanabilirler. Geleneksel bir video çerçevesi dikdörtgensel köşeleri olan bir video nesne ile gösterilebilir. Bir video nesne uzayı birkaç farklı şekilde kullanılabilir. En genel kullanımda video nesne uzayı, bir video nesnesinin herhangi bir zaman anında kodlanmış video bilgisini içerir. Bu durumda video nesne uzayı, içinde hareket parametreleri, biçim bilgisi ve doku verisini de bulundurur ve bunlar makro bloklar kullanılarak kodlanır.

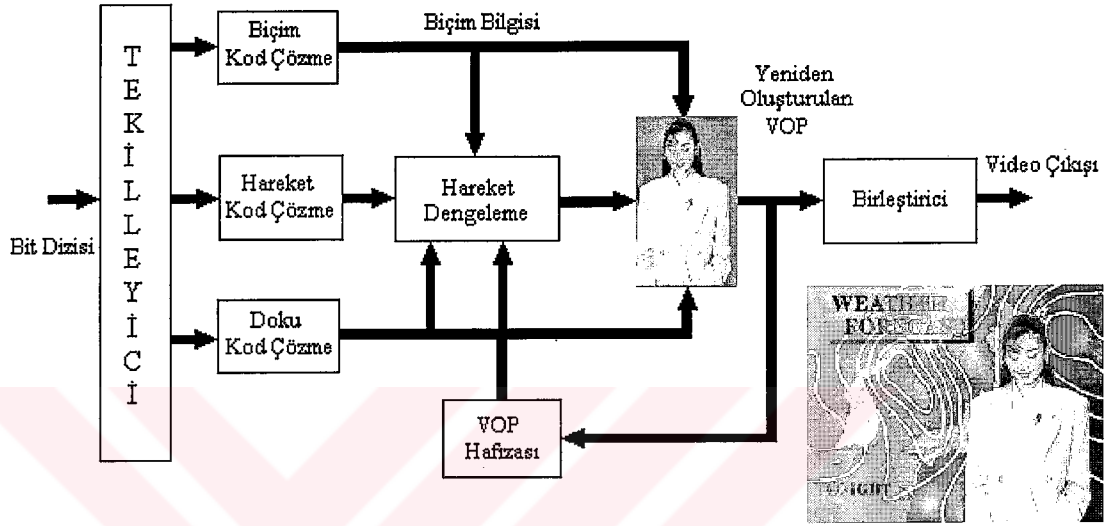
Video nesne uzayları aynı zamanda bir sprayt (sprite) kodlamak için de kullanılabilir. Sprayt genellikle gösterilen videodan daha büyük olan ve zaman boyunca sürekli olan bir video nesnesidir. Bir spraytı, parlaklığını değiştirerek veya uzamsal bozulmayı göz önüne alıp eğerek biraz değiştirmenin yolları vardır. Spraytlar, arka plan gibi büyük, durağan veya durağan olmayan alanı sunmak için kullanılır ve makro bloklar kullanılarak kodlanır.

Bir makro blok bir kısım ışıklılık bileşeni ve uzamsal olarak alt örneklenmiş renk bileşenlerinden oluşur. MPEG-4 kodlama standardı her makro blok için sadece 4:2:0 renk alt örnekleme formatını destekler. Her bir blok 8x8 piksel boyutundadır ve ayrık kosinüs dönüşümü kullanılarak kodlanır. Bir makro blok biçim bilgisi, hareket bilgisi ve doku bilgisi taşır.



Şekil 2.7. Video nesnelerinin kodlanması ve kodunun çözülmesi.

Şekil 2.7.'de MPEG-4'ün video nesnelere açısından genel kodlayıcı ve kod çözücü blok şeması verilmiştir. Her bir video nesnesi ayrı olarak kodlanır. Etkinlik ve geriye uyumluluk sebepleri ile video nesnelere, kendi video nesne uzayları yolu ile önceki MPEG standartlarına oldukça benzeyen bir melez kodlama şeması ile kodlanır. Şekil 2.8.'de bir video nesne uzayının kodunun çözülmesi gösterilmiştir.



Şekil 2.8. Video nesne uzayının kod çözümü.

### 2.3.3. Kenar kodlama

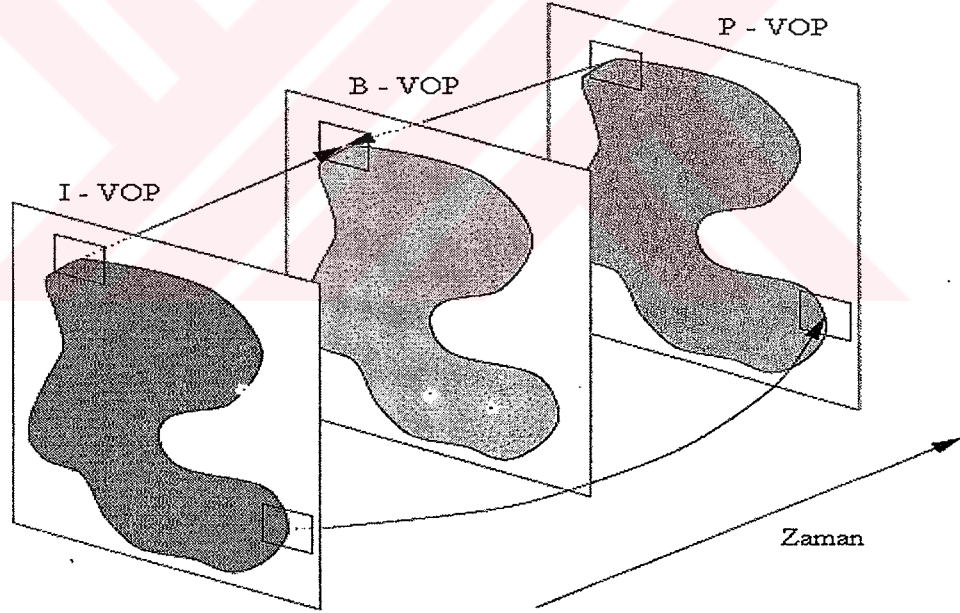
Biçim kodlama teknikleri bu tezin kapsamında olmadığından bu bölümde sadece MPEG-4 tarafından, gelişigüzel biçimlendirilmiş video nesnelere biçim bilgisinin kodlanmasında önerilen araçlar kısaca ele alınacaktır.

İlgili VOP için varolan biçim bilgisinin yanı sıra, biçim kodlama sistemi, biçim bilgisinin daha da fazla sıkıştırılması için aynı zamanda hareket kestirimi de kullanmaktadır. MPEG-4 video kodlama standardında video nesnesinin kendine özgü niteliğine bağlı olarak iki çeşit biçim bilgisi bulunmaktadır. Bunlar ikili ve gri-seviyeli biçim bilgileridir. İkili biçim bilgisindeki 1'ler ilgili bölümün (piksellerin) belirli bir zamanda video nesnesine ait olduğunu gösterir. İkili biçim bilgisi çoğunlukla video nesnesini sınırlayan dikdörtgensel çerçevenin boyutunda bir matris olarak gösterilir. Bu matrisin her elemanı pikselin video nesnesinin içinde veya dışında olmasına bağlı olarak mümkün olan iki değerden birini alır. Gri-seviyeli

biçim, ikili biçimin saydam (transparent) sunumuna olanak sağlayan ve örtüşme (aliasing) etkisini azaltan geliştirilmiş bir şeklidir. Bu bilgiler alfa uzayı (alpha plane) denilen imge dizinlerinde tutulur. (Sikora 1997)

#### 2.3.4. Hareket kestirim ve dengeleme araçları

Hareket kestirimi ve dengelemesi çoğunlukla çerçeveler arasındaki zamansal artıklıklardan yararlanarak video dizinlerini sıkıştırmak için kullanılır. MPEG-4’de hareket dengeleme için kullanılan yaklaşım diğer video kodlama standartlarına benzer. Temel fark diğer standartlarda kullanılan blok temelli tekniklerin, MPEG-4’de kullanılan video nesne uzayı yapılarına uyarlanmasıdır. MPEG-4 video nesne uzaylarının kodlanmasında üç kipi destekler. Bu kipler Şekil 2.9.’da gösterilmiştir.



Şekil 2.9. MPEG-4’ün video nesne uzaylarını kodlarken desteklediği kipler.

- Bir video nesne uzayı diğer video nesne uzaylarından bağımsız olarak kodlanabilir. Bu durumda kodlanacak video nesne uzayı Intra video nesne uzayı (I-VOP) olarak adlandırılır.

- Bir video nesne uzayı hareket dengeleme kullanılarak bir önceki kodu çözülmüş video nesne uzayından kestirilebilir. Bu durumda video nesne uzayları kestirilmiş video nesne uzayı (Predicted VOP – P-VOP) olarak adlandırılır
- Bir video nesne uzayı geçmiş ve gelecek video nesne uzaylarından kestirilebilir. Böyle video nesne uzayları iki yönlü aradeğerlenmiş video nesne uzayları (Bidirectional Interpolated VOP – B-VOP) olarak adlandırılır. Bu tip sadece I-VOP veya P-VOP lar kullanılarak da aradeğerlenebilir.

P-VOP ve B-VOP'ların kodlanması için hareket kestiriminin (motion estimation – ME) kullanılması gerekmektedir. Hareket kestirimi, makro bloklar için sadece video nesne uzayını kapsayan dikdörtgensel çerçeve içerisinde yapılır. Eğer bir makro blok tamamı ile video nesne uzayının içerisinde ise 16x16 piksel boyutundaki makro bloklarda, blok uyumlaması kullanılarak yapıldığı gibi, gelişmiş tahmin kipinde 8x8 piksel boyutundaki bloklar için de yapılır. Hareket vektörleri yarım-piksel doğrulukla hesaplanır.

Sadece bir kısım video nesne uzayına ait makro bloklar için hareket vektörleri, geliştirilmiş blok uyumlaması tekniği (modified block (polygon) matching technique) kullanılarak elde edilir. Burada uyumlama farklılığı, ilgili makro blokta sadece video nesne uzayına ait olan pikseller için mutlak farkların toplamı (sum of absolute difference – SAD) kullanılarak hesaplanır. Referans bloğun video nesne uzayının sınırında olması durumunda tekrarlamalı bir dolgu (padding) tekniği video nesnesi dışındaki piksellerin değerini belirler. Mutlak farkların toplamı daha sonra dolgulanmış pikseller kullanılarak da hesaplanır. Bu işlem, referans video nesne uzayının tahmini için aday pikseller ararken daha fazla olasılığa imkan sağlayarak etkinliği artırır.

P ve B tipi video nesne uzayları için hareket vektörleri aşağıda anlatılan şekilde kodlanır. Hareket vektörleri öncelikle farksal şekilde, önceden gönderilmiş blokların hareket vektörlerinin en fazla üçü kullanılarak kodlanır. Kesin sayı izin verilen hareket vektörü erimine bağlıdır. Azami erim MPEG-2'ye benzer bir şekilde

kodlayıcı ile seçilir ve kod çözücüye gönderilir. Daha sonra hareket vektörlerini kodlamak için değişken uzunluklu kodlama kullanılır.

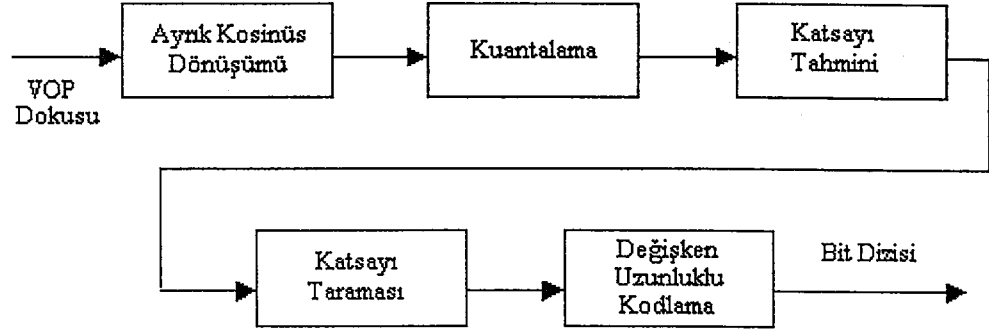
MPEG-4 aynı zamanda H.263'de de kullanılan üst üste binen (overlapped) hareket dengelemeyi de destekler. Bu genellikle daha düşük bit oranlarında daha iyi tahmin kalitesi verir. Burada makro bloğun her bir bloğu için komşu bloklar dikkate alınarak, o anki bloğun ve komşu dört bloğun hareket vektörü hesaba katılır. Her bir vektör bir piksel değeri kestirimi sağlar. Tahmin edilmiş gerçek değer, bütün bu kestirimlerin ağırlıklı ortalamasıdır.

### 2.3.5. Doku kodlama araçları

Bir video nesne uzayının doku bilgisi ışıklılık (Y) ve iki renk bileşeninden (Cb,Cr) oluşur. Intra video nesne uzayından (I-VOP) doku bilgisi doğrudan ışıklılık ve renk bileşenlerinden gelir. Hareket dengelenmiş video nesne uzaylarında doku bilgisi hareket dengelemeden sonra kalan artık hatayı gösterir. Doku bilgisinin kodlanmasında standart 8x8 blok tabanlı ayrık kosinüs dönüşümü kullanılır. Gelişigüzel biçimlendirilmiş video nesne uzayını kodlamak için 8x8 boyutundaki ızgara (grid) video nesne uzayının üzerine bindirilir. Bu ızgara kullanılarak video nesne uzayı içerisindeki 8x8 piksel boyutunda bloklar herhangi bir değişiklik yapılmadan kodlanır.

Video nesne uzayının her iki tarafında da bulunan bloklar sınır bloklar olarak adlandırılır ve onlara iç bloklardan farklı davranılır. Ayrık kosinüs dönüşümü uygulanan bloklar nicemlenir ve entropi değerini daha da azaltmak için ilgili blok için komşu bloklar kullanılarak katsayı tahmininde bulunulur. Bunu kodlanmış katsayılarla olan ortalama koşma mesafesini (run length) azaltmak için kullanılan katsayıların taranması takip eder. Daha sonra katsayılar değişken uzunluklu kodlama ile kodlanır. Bu süreç Şekil 2.10'da gösterilmiştir.





Şekil 2.10. Video nesne uzayı doku kodlaması.

### 2.3.6. Hata düzeltme

Bu işlevsellik mobil haberleşme gibi hatalara eğilimli ortamlarda evrensel erişim için önemlidir. MPEG-4 oluşacak hatalara karşı farklı seviyelerde sağlamlık ve karmaşıklık içeren birkaç mekanizma bulundurulur. Bu mekanizmalar ;

- i) Yeniden eş zamanlama (resynchronization) : Bir bit dizisinde hata düzeltme için en sık kullanılan yoldur. Benzeri olmayan işaretlerin bit dizisine eklenmesi mantığına dayanır ve bir hata durumunda kod çözücü bir dahaki işarete kadar olan bitleri atlar ve bir önceki noktadan itibaren verileri baştan alır (reset).
- ii) Veri bölme (data partitioning) : Bu yöntem hareket bilgisinin kodlanması için kullanılan bitlerle, doku bilgisini ayırır. Bir hata oluşması durumunda, örneğin hata sadece doku bilgisinde oluşmuşsa kodu çözülmüş hareket bilgisi kullanılarak etkin bir hata örtme uygulanabilir.
- iii) Başlık uzatma kodu (header extension code) : İkili kodlar doğru video kod çözümünün çok önemli olduğu durumlarda gerekenden fazla başlık bilgisi kullanılmasına olanak sağlar.
- iv) Tersine çevrilebilir değişken uzunluklu kodlar (reversible VLCs) : Bu değişken uzunluklu kodlar, kodu çözülen veride hatanın oluşma etkisini daha fazla azaltmaya olanak sağlar. Tersine çevrilebilir değişken uzunluklu kodlar ileri doğru çözülebileceği gibi geriye doğru da çözülebilen kodlardır. Bir hata oluşması ve bir

sonraki yeniden eş zamanlama işaretine kadar olan bitlerin atlanması durumunda, hatanın etkisini sınırlamak için bozulmuş bit dizisinden geriye doğru kod çözümü hala mümkündür.

### 2.3.7. Sprayt kodlama

Bir sprayt, bir video parçası boyunca ortamda bulunan bir video nesnesinin bölgelerinden oluşur. Buna açık bir örnek kamera hareketi boyunca arka plan olan piksellerin oluşturduğu arka plan mozaiği olarak bilinen “arka plan sprayt” dır. Bu aslında iletimin başında bir kez gönderilecek bir durağan imgedir. Spraytlar bazı durumlarda yüksek sıkıştırma verimliliği sağladığı için MPEG-4’e eklenmiştir. Zamanın herhangi bir anında, arka plan video nesne uzayı sprayt kesip kısaltılarak veya eğilerek uygun şekilde elde edilir. Sprayt tabanlı kodlama suni nesnelere için çok uygundur fakat aynı zamanda doğal ortamlardaki deforme olmadan hareket eden (rigid motion) nesnelere için de kullanılabilir. Video nesne uzayı gösterimine benzer şekilde bir sprayt için doku bilgisi bir ışıklılık ve iki renk bileşeninden oluşur. Bu üç bileşen ayrı ayrı işlenir fakat uygun ölçeklendirmeden sonra renk bileşenlerini işlerken kullanılan yöntemle ışıklılık bileşenini işlerken kullanılan yöntem aynıdır. Bir spraytın biçim ve doku bilgisi sadece kendi içinde kodlanan (Intra) video nesne uzaylarında olduğu gibi kodlanır.

Durağan spraytlar kodlama süreci başlamadan önce orjinal video nesne uzayları kullanılarak oluşturulur. Kod çözücü her durağan spraytı video parçasının tümünden önce alır. Sabit spraytlar, nicemlenmiş spraytın ilgili parametreler ile eğilmesi sonucu video nesne uzayının kolayca oluşturulabilmesine olanak tanınacak şekilde kodlanmaktadır.

Düşük esneklik içeren uygulamaları desteklemek için sprayt iletiminde birkaç olasılık dikkate alınır. Esnekliği sağlamanın bir yolu başlangıçta spraytın sadece bir kısmını iletmektir. İletilen kısım ilk video nesne uzayının yeniden oluşturulmasında yeterli olacaktır. Spraytın kalan kısmı gereklikçe veya bant genişliği izin verdikçe parçalar halinde iletilir. Bütün spraytı iletirken kullanılan diğer bir yöntem; spraytı



düşük kalitede iletmeye başlayan ve artık imgelerin iletilmesiyle kalitesi artan sürekli biçimdir. Pratikte bu yöntemlerin birleşimi kullanılabilir.

### **2.3.8. Ölçeklenebilirlik**

MPEG-4'de uzamsal ve zamansal ölçeklenebilirlik çoklu video nesne katmanları (Video Object Layer VOL) kullanılarak uygulanır. İki video nesne katmanı olması durumunda bunlar temel katman ve geliştirme katmanıdır (enhancement-layer). Uzamsal ölçeklenebilirlik bakımından geliştirme katmanı, temel katman tarafından sağlanan uzamsal çözünürlüğü artırır. Benzer şekilde zamansal ölçeklenebilirlik durumunda, istenilen çerçeve hızı temel katmanın sunduğu çerçeve hızı tarafından karşılanmıyorsa geliştirme katmanının kodu çözülebilir. Bundan dolayı zamansal ölçeklenebilirlik görüntü dizinindeki hareketin yumuşaklığını artırır. Nesne ölçekleme, video nesne tabanlı yapı tarafından doğrudan desteklenir.

## BÖLÜM 3 : VIDEO BÖLÜTLEME YÖNTEMLERİ

### 3.1. Giriş

MPEG-4'ün işlevsellikleri açıklanırken içeriğin (ortamdaki fiziksel nesnelerin) ayrı şekilde kodlanıp kodunun çözülmesinin içerik temelli sunumun temeli olduğu Bölüm 2'de vurgulanmıştı. İçerik temelli fonksiyonelliklere imkan verebilmek için MPEG-4 video nesne uzayı görüşünü öne sürmüştür. Giriş video dizininin her bir çerçevesi belirli sayıda gelişigüzel biçimlendirilmiş imge bölgesine ayrılır (Sikora 1997). MPEG-4 tarafından önerilen video nesne bölütlemesine bir örnek Şekil 3.1.'de gösterilmiştir.



**Şekil 3.1.** “Akiyo #28” in bölütlenme öncesi ve sonrası durumu.

Aynı video nesnesine ait video nesne uzaylarının biçim, hareket ve doku bilgisi ayrı bir video nesne katmanında kodlanır, iletilir ve kodu çözülür. Buna ek olarak bütün görüntü dizininin tekrar elde edilmesi için her bir video nesne katmanını tanımlayan bilgilere ihtiyaç vardır. Kodlanan video nesne uzayları, alıcıda her bir video nesne uzayı ve çözülmüş alfa kanalı bilgisi kullanılarak yeniden oluşturulur. Eğer arka plan içeriği “Akiyo” video dizininde olduğu gibi sabitse arka plan video nesne uzayı için sadece bir çerçeve kodlamak yeterlidir. Bundan dolayı ön plan ve arka plan video nesne uzayları, alıcıda farklı tekrarlanma oranlarına sahip olabilirler (Sikora 1997). Bu da yüksek sıkıştırmaya olanak sağlamaktadır.

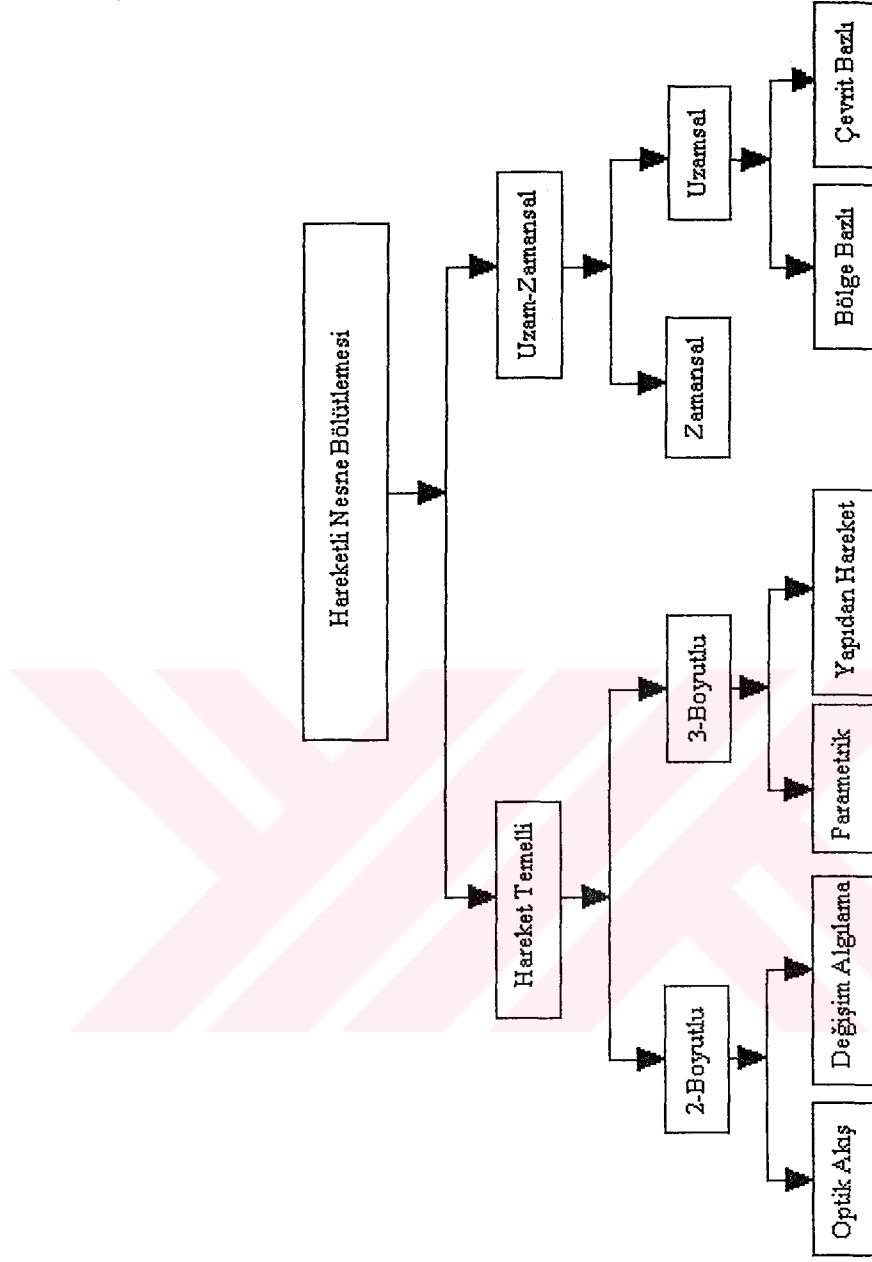
Yukarıda anlatılan nedenlerden dolayı bir görüntü dizini içerisinde hareket eden nesnelerin bulunması MPEG-4 video kodlama standardı için oldukça önemlidir. Bu amacı gerçekleştirmek için birçok video nesne uzayı bölütlemesi (hareketli nesne bölütlemesi) tekniği ileri sürülmüştür. Bu bölümde başlıca video nesne bölütlemesi teknikleri avantaj ve dezavantajları ile incelenecektir.

### **3.2. Video Nesne Bölütlemesi Tekniklerine Genel Bakış**

Literatürde video nesne bölütlemesi tekniklerini sınıflandırmak için birçok farklı yol önerilmiş fakat birçoğu bütün yöntemleri kapsamada yetersiz kalmıştır. Zang and Lu (2001) tarafından genel bir bakış açısı ile bu yöntemler sınıflanmıştır. Bu çalışmada Zang and Lu (2001) tarafından önerilen sınıflama yapısı ele alınarak video nesne bölütlemesi teknikleri sınıflandırılacaktır. Kullanılacak sınıflama yapısı Şekil 3.2.'de gösterilmiştir.

Hareketli nesne bölütlemesi yöntemleri hareket-bazlı teknikler ve uzam-zamansal teknikler olarak ikiye ayrılır. Hareket bazlı teknikler bölütlemede kullanılan hareket modelinin boyutuna göre 2-boyutlu ve 3-boyutlu teknikler olarak iki alt gruba ayrılır. 3-boyutlu yöntemler, çoğunlukla katı (rigid) nesne hareketinin üstesinden gelen hareketten yapı yöntemi (structure from motion) ve parçalı katı hareket ve 2-boyutlu ortamların üstesinden gelen parametrik yöntemler olarak ikiye ayrılır.

Hareket temelli bölütleme yöntemleri üç temel unsura ihtiyaç duyar. İlk unsur pikseller, köşeler, çizgiler, bloklar veya bölgelerden meydana gelebilen veri öğeleridir. İkinci unsur parametre veya hareket kestirimi gerektiren 2-boyutlu optik veya 3-boyutlu hareket parametrelerinden meydana gelebilen akış, hareket modeli veya hareket sunumudur. Üçüncü unsur azami son değer (maximum a posteriori - MAP), Hough dönüşümü, beklenti ve enbüyükleme (EM) den meydana gelebilen bölütleme ölçütüdür. Bundan dolayı tipik bir hareket bölütleme algoritması bu üç unsura bağlı olarak üç aşamadan oluşur. Bununla birlikte gürültü sorunu ve ortamın hareket karmaşıklığı nedeni ile bölütleme/kümeleme yöntemleri hareketin gösterilmesi aşamasında kullanılan hareket kestiriminden daha karmaşıktır ve bölütleme genellikle özyineli bir süreçtir.



**Şekil 3.2.** Hareketli nesne bölütlemesi yöntemlerinin sınıflandırılması.

Hareket bazlı bölütleme algoritmaları hareket sunumu ve kümeleme ölçütleri göz önüne alınarak sınıflanabilir. Hareket sunumu, hareket bölütlemeye önemli bir rol oynadığı için, hareket bölütleme teknikleri hareket kestirim algoritması tasarımına odaklanmıştır. Bundan dolayı hareket bölütlemesi en iyi olarak, kabul edilen hareket gösterimi tarafından tanımlanmakta ve ayırt edilebilmektedir. Hareket gösterimi ile tanımlanan her alt grupta, yöntemler kümeleme ölçütüne göre ayrılmıştır.

Sadece hareket bilgisini kullanan geleneksel hareket bazlı bölütleme yöntemleri genellikle katı veya kesikli katı hareket içeren ortamların üstesinden gelebilir. Hareket ve uzamsal bilgiyi birlikte kullanan uzam-zamansal bölütleme yöntemleri, imge bölütlemeye oluşan aşırı bölütleme sorununu aşmayı amaçlar ve gürültü duyarlılığının ve hareket bazlı yöntemlerdeki yanlışlıkların üstesinden gelebilir. Uzam-zamansal bölütlemeye, hareket temelli bölütlemeye kullanılan benzer hareket kestirim teknikleri kullandığından ve zamansal bölütleme genellikle baştan başa bölütleme sonuçlarına yol gösterdiğinden hareket temelli bölütleme olarak sınıflanacağı düşünülebilir. Ancak, bu grup bölütleme algoritmaları zamansal bölütleme sonuçlarını düzeltmekte ve iyileştirmekte uzamsal bilgi kullanıldığından hareket bazlı bölütlemeye farklılık gösterir. Bu yolla sadece hareket bazlı yöntemlerdeki bazı sorunların üstesinden gelinmez, bu yöntemler aynı zamanda katı olmayan (non-rigid) hareketlere de uygulanabilir.

Hareket temelli 2-boyutlu yöntemler düşük işlem yüküne ve kolay uygulanmasına karşın aşırı bölütlemeye neden olabilir ve güvenli olmayabilir. Hareket temelli 3-boyutlu yöntemler ise sabit arka plan, katı varsayımlar gibi şartlara bağlı olarak genel hatlarıyla güvenli sonuç vermekle birlikte genel ortamlara uygulanamamasından ve işlem yükünün gerçek zamanlı uygulamalar için fazla olmasından dolayı video nesne bölütlemesinde pek kullanılmamaktadır. Bu nedenle, literatürde son zamanlarda uzam-zamansal teknikleri kullanan bölütleme yöntemleri ile ilgili çalışmaların sayısı hızla artmaktadır. Yukarıda anlatılan nedenlerden ve MPEG'in video nesne uzayı bölütlemesi için uzam-zamansal bir yöntem önermesinden dolayı bu bölümde hareket temelli tekniklerden kısa bir şekilde bahsedilecek, uzam-zamansal teknikler daha detaylı olarak anlatılacaktır.

### **3.3. Hareket Bazlı 2-Boyutlu Video Nesnesi Bölütlemesi Teknikleri**

2-boyutlu video nesnesi bölütleme teknikleri optik akış kesikliği ve değişim algılama temelli bölütleme yöntemleri olarak ikiye ayrılır.

### 3.3.1. Optik akış

Bu grup yöntemler imge piksellerinin yer değişimine veya optik akışına göre bölütleme yaparlar. Bir pikselin yer değişimi veya optik akışı o pikselin bir çerçeveden bir sonraki çerçevedeki yeni yerine hareketi ile gösterilen bir hareket vektörüdür. Birçok yaklaşım piksel seviyesinde hareket vektörü kestirimi için optik akış kullanır ve bölütleme yapmak için tutarlı hareketin olduğu bölgedeki pikselleri kümeler. Optik akıştaki kesiklikleri kullanan bazı bölütleme yöntemleri önerilmiş fakat bu yöntemler tatmin edici sonuçlar vermemiştir (Zang and Lu 2001).

Optik akış yaklaşımının iki büyük dezavantajı vardır. İlki optik akış yöntemlerinin büyük hareketlerle başa çıkamaması ikincisi ise tutarlı hareket içeren bölgelerin birden fazla nesne içermeye ihtimali ve nesne çıkartımı için tekrar bölütlemeye ihtiyaç duymasıdır (Deng and Manjunath 1998). Ayrıca optik akış gürültüye aşırı şekilde duyarlıdır ve doğruluğu açıklık (aperture) ve kapatma (occlusion) sorunlarından dolayı sınırlıdır (Meier and Ngan 1998). Yukarıda anlatılan eksikliklerinden dolayı optik akış hareketli nesne bölütlemesinde doğrudan kullanılmayıp, bazı yöntemlerle birlikte kullanılmaktadır. Örneğin Meier and Ngan (1998) tarafından önerilen hareketli nesne bölütlemesi yönteminde optik akış, nesnelerin hareketinin algılanması için kullanılmıştır.

### 3.3.2. Değişim algılama

Optik akış genellikle çerçevedeki her piksel için hesaplanır ancak bir görüntü dizisinde art arda gelen iki çerçevedeki piksellerin çoğu ya hareketsizdir yada genel hareketi (global motion) takip eder. Bu nedenle sadece değişim gösteren bölgelere odaklanmak daha etkin ve işlem yükü açısından daha kazançlı olacaktır. Bu algoritmalar arda ardına gelen çerçevedeki zamanla değişen ve değişmeyen bölgeleri ayırt etmek için değişim algılamayla başlar. Herhangi bir pikselin hangi bölgeye ait olduğu çerçeve farkı (frame difference) kullanılarak belirlenir. Ortamdaki ışık değişikliği, kamera gürültüsü gibi etkenlerden dolayı bu çerçeve farkı bir eşik değer ile karşılaştırılır. Belirlenen eşik değerini aşan pikseller değişim algılama maskesinde

(change detection mask) işaretlenir. Değişim algılama ile ilgili ayrıntılı bilgi uzam-zamansal bölütleme yöntemleri anlatılırken verilecektir.

### **3.4. Hareket Bazlı 3-Boyutlu Video Nesnesi Bölütlemesi Teknikleri**

Yukarıda 3.3 başlığı altında incelenen 2-boyutlu yaklaşımlar 2-boyutlu imge uzayındaki açık hareketleri analize eder. Bu analizler sadece çerçevelerde varolan bilgilerle, hareketli nesnenin uzaydaki gerçek hareketi ve yapısı dikkate alınmadan yapılır. 2-boyutlu hareket modelleri basit fakat gerçekliği daha azdır. 3-boyutlu yaklaşımda nesnenin uzaydaki hareketi de göz önüne alınır. 3-boyutlu bölütleme yöntemleri hareketten yapı (structure from motion) ve parametrik (parametric) yaklaşım olarak ikiye ayrılır. Hareketten yapı yaklaşımı genellikle önemli derinlik bilgisi içeren 3-boyutlu ortamlarla başla çıkarken, parametrik yaklaşım derinlik bilgisini kullanmaz. Bu iki yaklaşım arasındaki diğer temel fark, hareketten yapı yaklaşımı katı hareketlere izin verirken, parametrik yaklaşım kesikli hareket katılığına izin verir.

#### **3.4.1. Hareketten yapı (structure from motion)**

Hareketten yapı yaklaşımı, 2-boyutlu imge uzayından 3-boyutlu geometriyi elde etmeye dayanır. Bu yaklaşımda insanın 3-boyutlu dünyada bulunmasına rağmen onun sadece 2-boyutlu izdüşümünü algılaması olgusundan esinlenilmiştir. Bu yaklaşım robot yön güdümü, nesne izleme, animasyon, aktif görü, 3-boyutlu kodlama ve mozaikleme gibi uygulamalarda kullanılmaktadır.

2-boyutlu hareketten 3-boyutlu yapıyı elde etmek, kestirilecek ortama göre bir boyut daha az bilindiğinden zor bir problemdir. Bu yapıyı açık ve kesin biçimde kullanmak için farklı basitleştirme varsayımları yapılmıştır. Önemli varsayımlardan birinde ortamdaki nesne katı hareketlidir veya buna denk olarak sadece kameranın hareketine izin verilmiştir. Diğer bir önemli konu, noktaların belirlenmesi ve iki çerçeve arasındaki önemli noktalar arasında bağlantı kurulmasıdır. Hareketten yapı teknikleri kamera modeli, parametre kestirimdeki doğrusallık ve kamera ve ortam



yapısının sınırlanmasına göre farklılık gösterir. Birçok yapıdan hareket tekniği, nesneye ait çoklu imgelerin izdüşümünde değişmeyen ve kamera hareketinin kolayca çıkartılabileceği geometrik ve cebirsel özellikleri kullanma temeline oturtulmuştur. Örneğin, gerekli matris, temel matris, ve çarpanlara ayırma (factorization) yöntemlerinin tamamı izdüşüm veya paralel projeksiyon değişmezlerini kullanarak iki veya üç görünümünden kamera parametrelerini elde etmektedir (Zang and Lu 2001).

Torr (1995) tarafından temel matris ile oluşturulmuş bir epipolar çizgi kullanan olasılıksal (stochastic) bir 3-boyutlu bölütleme yapısı geliştirilmiştir. MacLean (1996) giriş verisi olarak doğrusal sabitler kullanan bir beklenti ve enbüyükleme algoritması öne sürmüştür (Zang and Lu 2001).

### **3.4.2. Parametrik yöntemler**

Parametrik yöntemler yapıdan hareket yöntemindeki katılık varsayımının yerine kesikli katılık anlayışını temel alır. Parametrik modeller, ortamda fiziksel yapı varsayımı ile uzaydaki nesne hareketinin kullanılması ile oluşturulur. Nesnenin üç boyutlu hareketi genellikle 3-boyutlu ilgin hareket modeli ile modellenir ve bir rotasyon matrisi ve dönüşüm vektörü ile tanımlanabilir. 6 parametrelilik model paralel izdüşüm altındaki düzlemsel yüzeyler için, 8 parametrelilik model perspektif izdüşüm altındaki düzlemsel yüzeyler için, 12 parametrelilik model ise paralel izdüşüm altındaki parabolik yüzeyler için kullanılmaktadır.

### **3.5. Uzam-Zamansal Video Nesnesi Bölütlemesi Teknikleri**

MPEG grubu hareketli nesne bölütlemesi için morfolojik uzam-zamansal bölütleme ve yüksek dereceden istatistik bilgi kullanan hareket algılama (higher-order-statistic motion detection) yöntemlerini birleştirilmiştir (Perez et al 2001). Bu nedenle video nesne bölütlenmesi için uzam-zamansal teknikler üzerindeki çalışmalar giderek artmıştır. Bu bölümde uzam-zamansal teknikler detaylı olarak ele alınacaktır. Bölütlemede kullanılan zamansal ve uzamsal teknikler daha iyi bölütleme için genellikle tek başlarına değil birlikte kullanılırlar. Bu nedenle sınıflama yapılırken ana ölçüt olarak uzamsal bilgiyi kullanıp daha iyi bölütleme için zamansal bilgidan



yararlanan yöntemler uzamsal yöntemler kısmında, ana ölçüt olarak zamansal bilgiyi kullanıp daha iyi bölütleme için uzamsal bilgiden de yararlanan yöntemler zamansal yöntemler kısmında açıklanacaktır.

### 3.5.1. Uzamsal bölütleme yöntemleri

Ana ölçüt olarak uzamsal bilgi kullanıp daha iyi bölütleme için zamansal bilgiden yararlanan yöntemler bölge ve çevrit bazlı yöntemler olarak ele alınabilir. Bölge bazlı yöntemler genellikle birkaç aşamadan oluşan morfolojik yapılar kullanarak temelde imge bölütlemesi yaparlar. Daha sonra elde ettikleri bölgeleri hareket kestirimi veya izleme yöntemleri kullanarak takip ederek bölütleme sonuçlarını elde ederler. Çevrit bazlı yöntemler aktif çevrit kullanan yöntemler ve kenar algılama kullanıp, elde ettiği kenarları takip eden yöntemler olarak sınıflanabilir.

Zhong and Cheng (1999) başlangıçta o anki çerçevenin kenar ve renk bilgilerini ve bir önceki çerçeveyi dikkate alarak hesapladığı hareket bilgisini kullanarak imgeyi bölgelere ayırmaya çalışan bir yarı otomatik video nesne bölütlemesi tekniği önermişlerdir. Önerilen yöntemin bölütleme kısmının temelini CIE (L,u,v) renk uzayına dönüştürülmüş renk bilgisi oluşturmaktadır. CIE uzayına dönüştürülen renk bilgisi bölge birleştirmede kullanılmadan önce uyarlanır nicemleme (adaptive quantization) ve ortanca süzgeci (median filter) kullanılarak gürültüden kurtulmak için yumuşatılır. Kenar bilgileri Canny kenar algılama algoritması ile, hareket bilgileri ise sıra düzensel blok uyumlaması algoritması ile elde edilir. Elde edilen bu bilgiler kendi önerdikleri yöntem ile işlenerek bölge bölütlemesi tamamlanır. Daha sonra bölütleme sonucu elde edilen türdeş bölgeler ve kullanıcı tarafından belirlenen nesne sınırları dikkate alınarak bölgeler bir araya toplanır. Daha yüksek bölütleme doğruluğu için diğer yaklaşımların aksine hem nesne hem de arka plan için ilgin nesne parametreleri takip edilir. Bu yöntem birden fazla nesneyi takip edememektedir.

Zhong and Cheng (1999) tarafından önerilen yöntemin değişik bir uygulaması Zhou et al (2001) tarafından önerilmiştir. Bu çalışmada, diğerinin aksine kullanıcı yardımına gerek duyulmamaktadır. Önerilen yöntemde öncelikle üç aşamadan oluşan

morfolojik bölütleme kullanılmış daha sonra bölge birleştirme sonucu elde edilen nesne bölgeleri ilgin hareket parametreleri kullanılarak takip edilmiştir. Bölütlemenin ilk aşamasında bölütlemeyi kolaylaştırmak için morfolojik süzgeçler kullanılarak imge basitleştirilir. Daha sonra türdeş bölgelerin içleri birbirinden farklı olarak işaretlenir. Son olarak bu etiketlenilmiş işaretler bölge büyütme için watershed algoritmasına giriş olarak verilir. Bölge birleştirme aşamasında hareket bilgisinden yararlanmak için kendi önerdikleri optik akış kestirimi ve baskın hareket kestirimi ölçütü kullanmışlardır. Elde edilen nesnenin ilgin hareket parametreleri ile takibi ile hareketli nesne bölütlemesi tamamlanmıştır.

Zhou et al (2001) tarafından önerilen yapının benzeri daha önce Wang (1998) tarafından da kullanılmıştır. Bu iki yöntem arasındaki temel farklardan biri watershed algoritmasının fazla bölütleme yapmasını engellemek için kullandıkları süzgeçlerdir. İzleme aşamasında Wang (1998) hareket izdüşümü, işaret çıkartımı, geliştirilmiş watershed ve bölge bileştirme kullanılmıştır. Wang (1998) tarafından önerilen yöntem sadece tek bir video nesnesi için bölütleme yaparken Zhou et al (2001) tarafından önerilen yöntem birden fazla nesne için bölütleme yapabilir.

Tsaig and Averbuch (2002) tarafından video nesne bölütlemesi problemi sabit bir arka plan üzerindeki hareketli nesnelerin algılanması olarak düşünülmüştür. Bu nedenle öne sürdükleri yöntemin ilk aşamasını bütünsel hareketin dengelenmesi (global motion compensation) oluşturmaktadır. Bütünsel hareket, 8-parametrelilik izdüşümsel hareket modeli ile modellenmiş ve gradyan tabanlı bir teknikle kestirilmiştir. O anki çerçevenin ilk uzamsal bölütlemesi gradyan yaklaşım, watershed bölütlemesi ve uzam-zamansal birleştirme ile yapılmıştır. Sınıflama aşaması, ilk bölütleme sonucu elde edilen bölgeleri arka plan ve ön plan adayları olarak işaretleyen istatistiksel önem testi ile başlar. Her ön plan adayının hareketi sıradüzensel bir bölge uyumlaması ile kestirilir. Örtme sorununda kaynaklanacak yanlış hareketlerden sakınmak için kestirilen vektörler döngülü (iterative) bir düzeltme yapısından geçirilir ve gerekirse düzeltilir. Kestirilen hareket vektörleri daha sonra uzamsal bilgi ve bir önceki çerçeveden alınan bilgi ile birlikte Markov rasgele alan modeline (Markov random field model) dahil edilir. Markov rasgele alan modelinin eniyilemesi (optimization) ile sınıflama aşaması bitirilir. Son olarak

bölütleme sürecinin zamansal tutarlılığını sağlamak için dinamik bir hafıza kullanılmıştır. Her bir bölge görüntü dizinindeki bir sonraki çerçevede takip edilir ve buna göre bu hafıza güncellenir. Önerilen yöntemin avantajları hareket dengelemesi kullanması, birden fazla video nesnesini takip edebilmesi, yavaş hareket içeren görüntü dizinlerinde de iyi sonuç vermesi olarak sıralanabilir.

İmge bölütlemesi için kullanılan çevrit bazlı yöntemler son zamanlarda video nesne bölütlemesi için de kullanılmaya başlanmıştır. Aktif çevrit modelleri, dahili güç zorlamaları ile çizgi ve kenar gibi özelliklere çekilmelerine yol açan harici görüntü güçleri tarafından yönlendirilen enerjisi enküçüklenmiş eğrilerdir (spline) (Park et al 2001).

Park et al (2001) video bölütlemesi için yönlü bir aktif çevrit yapısı önermiştir. Bu yapı çevrit enerjisini en aza indirirken imgenin hem gradyan gücünü hem de gradyanın yönünü dikkate alır. Bu yapı kullanılarak ilk bölütleme yapıldıktan sonra, bir sonraki aşamada ilk bölütleme sonucunda elde edilen parametreler kullanılmaktadır. Ancak önerilen yapının verdiği bölütleme sonuçları tatmin edici değildir. Bunun en temel sebepleri aktif çevrit bazlı yöntemlerin karmaşık arka plan içeren görüntülerde iyi sonuç vermemesi ve gürültü nedeni ile bozulabilecek yön parametrelerinin bir sonraki bölütleme sonuçlarını kötü yönde etkilemesidir.

Precioso and Barlaud (2002) video bölütlemesi için uzamsal düzenli b-eğrisi aktif çevrit (regular spatial b-spline active contour) kullanan bir yöntem önermiştir. Nesne kenarlarının tahmininde kullanılan B-eğrisi, bir veya daha fazla eğri parçasından oluşan sürekli global parametresi olan parametrik bir eğridir. Önerilen yöntem video bölütlemesinde iyi sonuç veren bölgesel bazlı aktif çevritlerin yüksek işlem yükünü hafifletmek için çok çözünürlüklü düzenli örnekleme ile uzunluk cezasını birlikte kullanır.

Zaletelj and Tasic (2002) tarafından önerilen yöntem, görüntü dizinin ilk çerçevesinde kullanıcı tarafından belirlenen ana hatların eğrilik fonksiyonunun analizini yapar ve en iyi B-eğrisi biçim modelini elde eder. Daha sonra çevritin yer değiştirmesini takip etmek için 2-boyutlu ilgin hareket modelini kullanır. Önerilen

yöntem büyük nesne kapatmaları (major object occlusions) olmadan önce en az 50 çerçeve boyunca nesnenin kenarlarını takip edebilir. Ancak büyük nesne hareketlerini takip edebilmek için zamansal kestirim kullanılması gerekmektedir.

Sun et al (2003) tarafından önerilen VSnake yapısında aktif çevritin enerjisi tanımlanırken tek çevrit enerjisi yerine iki çevrit arasındaki enerji farkı kullanılmıştır. İmge çerçevesinin kenar enerjisi oluşturulurken çok-çözünürlüklü dalgacık açılımı (multi-resolution wavelet decomposition) uygulanmıştır. Nesnenin bir çerçeveden diğerine şekil değiştirmesi ile başa çıkabilmek için çevrit gevşetme (contour relaxation) kullanılmış ve gevşetme boyunca Viterbi algoritması kullanılarak çevrit güncellenmiştir. Bu yöntem temel aktif çevrit yapısından daha güvenilir olmakla birlikte diğer çevrit bazlı yöntemlerde olduğu gibi nesnedeki büyük biçim değiştirmesine karşı iyi sonuç vermemektedir.

Kim and Hwang (2002) tarafından çevrit veya bölge bazlı diğer yöntemlerden farklı olarak hareketli kenarların elde edilmesine dayanan bir bölütleme yöntemi önerilmiştir. Yöntem arka arkaya gelen iki çerçevenin farkından elde edilen iki katlı kenar haritasının oluşturulması ile başlar. Önceki çerçeveye ait kenar noktalarının çıkarılmasından sonra elde edilen hareketli kenarlar video nesne uzayının elde edilmesinde kullanılır. Ortamda birden fazla nesne olması durumunda bağlantılı bileşen analizi (connected component analysis) ve video nesnesinin arka arkaya gelen çerçevelerdeki yer değişiminin yumuşatılması kullanılmıştır. Yöntem göreceli olarak düşük işlem yükü ve etkin bölütleme sonuçları vermekle birlikte yavaş hareket içeren veya kısa süreli duran nesnelere için iyi sonuç vermeyebilir.

### **3.5.2. Zamansal bölütleme yöntemleri**

Ana ölçüt olarak zamansal bilgi kullanıp daha iyi bölütleme için zamansal bilgidenden yararlanan yöntemler temel olarak değişim algılama yöntemleridir.

Mech and Wollborn (1998) tarafından kamera hareketlerini ve ortam değişimlerini algılayabilen gürültüye karşı dayanıklı değişim algılama temelli bir bölütleme yöntemi önerilmiştir. Önerilen yöntem öncelikle, olabilecek belirgin (apparent)

kamera hareketlerini sekiz parametrelili bir hareket modeli kullanarak kestirir ve dengeler. Yöntem, ortam deęişikliği algıladıęında (scene cut detection) gerekirse algoritmayı baştan başlatılır. Önerilen yöntem bir bölgesel eşikleme teknięi kullanarak temelde deęişim algılama üzerine kurulmuştur. Bundan dolayı bölgesel komşuları dikkate alan bir gevşetme teknięi önerilmiştir. Daha tutarlı bölütleme sonuçları için deęişim algılama maskesinde bir hafıza yapısı kullanılmıştır. Deęişim algılama aşamasından sonra kestirilen yer deęiştirme vektörleri, deęişen bölgeleri hareketli nesne bölgesi ve açığa çıkan arka plan bölgesi olarak ayırmak için kullanılmıştır. Son olarak kestirilen nesne biçiminin doğruluęunu arttırmak için elde edilen nesnenin kenarları o anki imgenin kenar bilgilerine uydurulmuştur. Farklı tipteki görüntülerin üstesinden gelebilmek için kullanılan parametreler otomatik olarak algoritma tarafından belirlenmektedir. Önerilen yöntem en önemli eksiklięi kullandıęı hafıza yapısı nedeni ile kenarları çok iyi şekilde yapışamamasıdır.

Neri et al (1998) tarafından önerilen yöntem deęişim algılama mantığına dayanmaktadır. Muhtemel ön plan bölgeleri, çerçeveler arası farka yüksek dereceli istatistiksel (higher order statistics) önem testi uygulanarak algılanır. Bu aşamada sonuçlar tatmin edici olmadığından hareket bilgisi kullanılarak düzeltilmiştir. Ancak herhangi bir şekilde uzamsal kenar bilgisi kullanılmadıęı için elde edilen sonuçlar nesne kenarlarına yaklaşmamıştır.

Kim et al (1999) tarafından bütünsel hareket kestirimi için altı parametrelili ilgin hareket modeli kullanan ve elde ettięi bu parametrelere göre hareket dengelemeyi devreye sokup çıkararak zamansal bir bölütleme yöntemi önerilmiştir. Bu yöntem zamansal bölütleme ile nesnenin yerini belirler ve uzamsal bölütleme kullanarak nesne sınırlarının doğru bir şekilde elde edilmesini sağlar. Zamansal bölütleme, deęişinti (variance) karşılaştırmaya dayanan istatistiksel varsayım testini de içeren ışıklılık farkını temel almaktadır. Uzamsal bölütleme ise bir bölge büyütme algoritması olan watershed algılaması ile yapılmaktadır. Bunun için genelde olduęu gibi bu yöntemde de morfolojik süzgeçlerle imge basitleştirilir. Gradyan yaklaşımı ve imge piksellerini en uygun bölgeye atayan watershed algoritmalarının uygulanmasından sonra bölge birleştirme kullanılarak uzamsal bölütleme tamamlanır. Son olarak uzamsal bölütleme aşırı bölütlemeye yol açacağından



zamansal sonuçlar ve uzamsal sonuçlar birlikte değerlendirilerek hareketli nesnelere arka plandan ayrılır. Bu yöntem, yavaş hareket eden ve kısa süreli duran hareketli nesnelere bölütlenmesinde sorun yaratmaktadır.

Long et al (2001) tarafından önerilen yöntemde öncelikle basitleştirilmiş doğrusal bir model kullanarak bütünsel hareket kestirimi ve dengelemesi yapılmaktadır. Model parametrelerini kestirmek için sıra düzensel blok uyumlama ve en küçük kareler yaklaşımı kullanılmaktadır. Daha sonra art arda gelen çerçeveler arasındaki değişim, türevsel bir yapı kullanılarak elde edilir. Bundan sonra sistem önceden belirlenen ölçütlere göre nesneyi elde etmeye çalışır. Uzamsal bölütlenme aşamasında ise morfolojik süzgeçleme algoritmalarının yerine daha basit olan sıra düzensel uyarlamalı eşikleme (hierarchical adaptive thresholding – HAT) ve bölge birleştirme kullanılmıştır. Bu yöntem nesne kenarlarına Neri et al 1998 tarafından önerilen yöntemden daha fazla yaklaşmasına rağmen yavaş hareket eden nesne bölgelerinde etkisiz kalmaktadır.

Chein et al (2002) tarafından piksel bazlı arka plan kaydı temel alan bir bölütlenme algoritması önerilmiştir. Bu yöntem temel olarak arka arkaya gelen birkaç çerçevenin değişimini kullanır. Eğer bir piksel belirlenen bir süre boyunca hareket vermiyorsa arka plan imgesine aktarılır ve daha sonra her çerçeve bu arka plan imgesi ile karşılaştırılır. Arka plan imgesinden büyük farklılık gösteren pikseller ön plan olarak işaretlenir diğer pikseller ise arka plan olarak kabul edilir. Son olarak gürültülü bölgeleri yok etmek ve daha yumuşak nesne biçimi elde etmek için işleme sonrası aşamada morfolojik süzgeçler kullanılmıştır. Bir çok değişim algılama temelli yöntemin sorunu olan gölge etkisini aşmak için nesne biçimini koruyup, gölgeli bölgelerin üstesinden gelebilen morfolojik bir gradyan süzgeci kullanılmıştır. Yöntemin gerçek zamanlı olarak çalışması için günümüz mikro işlemcilerinin 32 veya 64 bitlik veri yollarına sahip olduğu gerçeği göz önüne alınarak 8 bitlik imge ışıklılık değerleri paralel olarak işlenmiştir. Yöntemi test ederken kullanılan görüntüler ya gürültüsüz ve yavaş hareket içermekte ya da gürültülü ve hızlı hareket içermektedir. Bu tip görüntülerde yöntem çalışmakla birlikte, hem yavaş hareket içerip hem de gürültülü görüntü dizinlerinde istenilen sonuçlar alınamayacaktır.

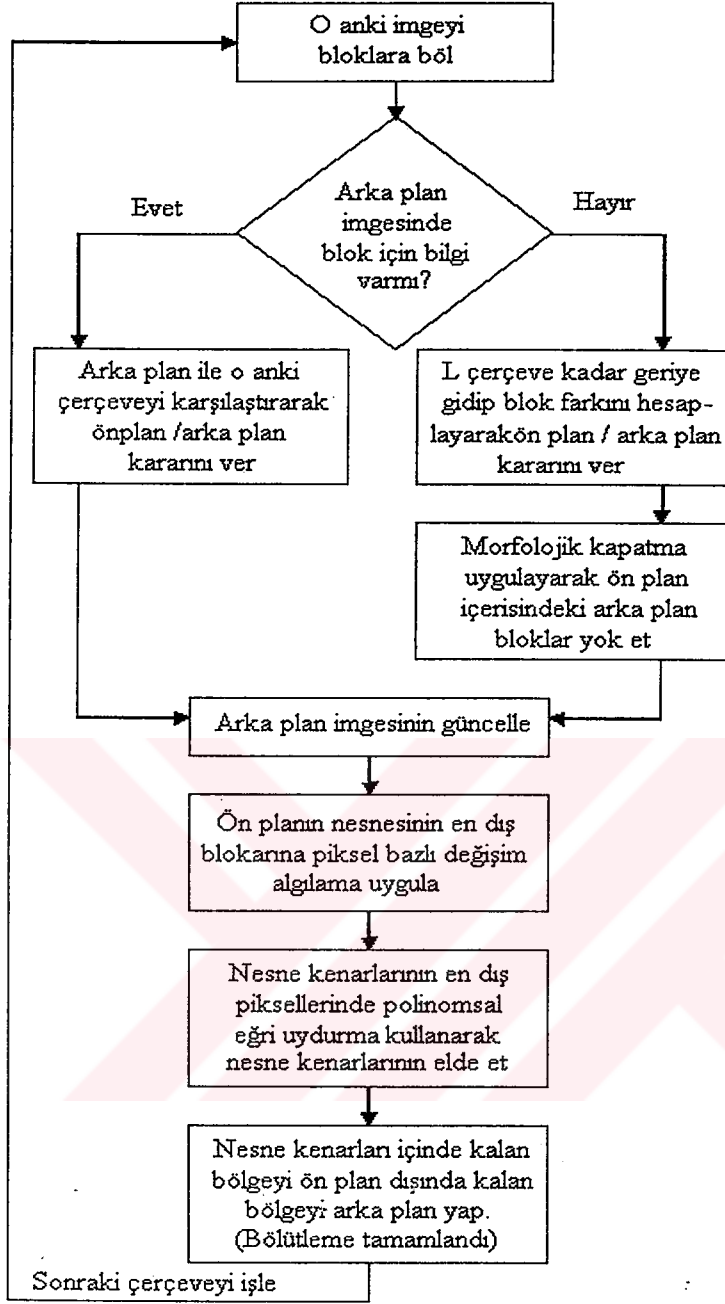
## BÖLÜM 4 : ÖNERİLEN VIDEO BÖLÜTLEME YÖNTEMİ

### 4.1. Giriş

Bu çalışmada MPEG-4 video kodlama standardının temel ögesi olan video nesnelerinin bölütlemesi için yeni bir yöntem ileri sürülmüştür. Önerilen bölütleme yöntemi 3. bölümde gösterilen sınıflama yapısı dahilinde uzam-zamansal bir yöntem olarak ele alınabilir. Daha kesin bir sınıflama yapılmak istenirse; temel kriter olarak değişim algılama kullanıldığından ve daha iyi bölütleme için uzamsal bilgiden de yararlandığı için zamansal yöntem olarak sınıflamak daha doğru olacaktır.

Önerilen yöntem blok bazlı arka plan kaydına dayanmaktadır ve özellikle yavaş hareket eden nesnelere bulunduran video konferans ve video telefon gibi uygulamalarda video nesnelerinin elde edilmesi için uygundur. Sadece art arda gelen çerçevelerin eşiklenmesini temel alan klasik değişim algılama (change detection) yaklaşımı yavaş hareket eden veya kısa süreli duraklayan nesnelere bölütlenmesinde yetersiz kalmaktadır. Bu çalışmada daha eski çerçeveleri de göz önüne alarak yavaş hareket ve kısa süreli nesne duraklamaları durumlarında da başarılı bölütleme yapılması amaçlanmıştır. Çerçeve farkı piksel yerine blok bazlı olarak alınmış ve bu sayede gürültü etkisi azaltılmaya çalışılmıştır. Blok bazlı çerçeve farkında hareketsiz olan bloklar arka plan resmine kaydedilmektedir. Sonraki aşamalarda arka plan bilgisi olan bloklar için sadece bir önceki çerçeveye bakılırken, arka plan bilgisi olmayan bloklar için önceden belirlenen sayıda çerçeve geri gidilip ilgili bloğun durumuna karar verilir. Blok bazlı morfolojik işlemler uygulanarak blok bazlı nesne maskesi elde edilir. Nesne kenarlarına yaklaşmak için nesne maskesinin en dış bloklarında piksel bazlı eşikleme kullanılarak bu bloklardaki piksellerin durumu belirlenir. Son olarak nesne kenarlarını yumuşatmak için düşük dereceli polinomsal eğri uydurma kullanılmıştır. Önerilen yöntemin blok şeması Şekil 4.1'de verilmiştir.

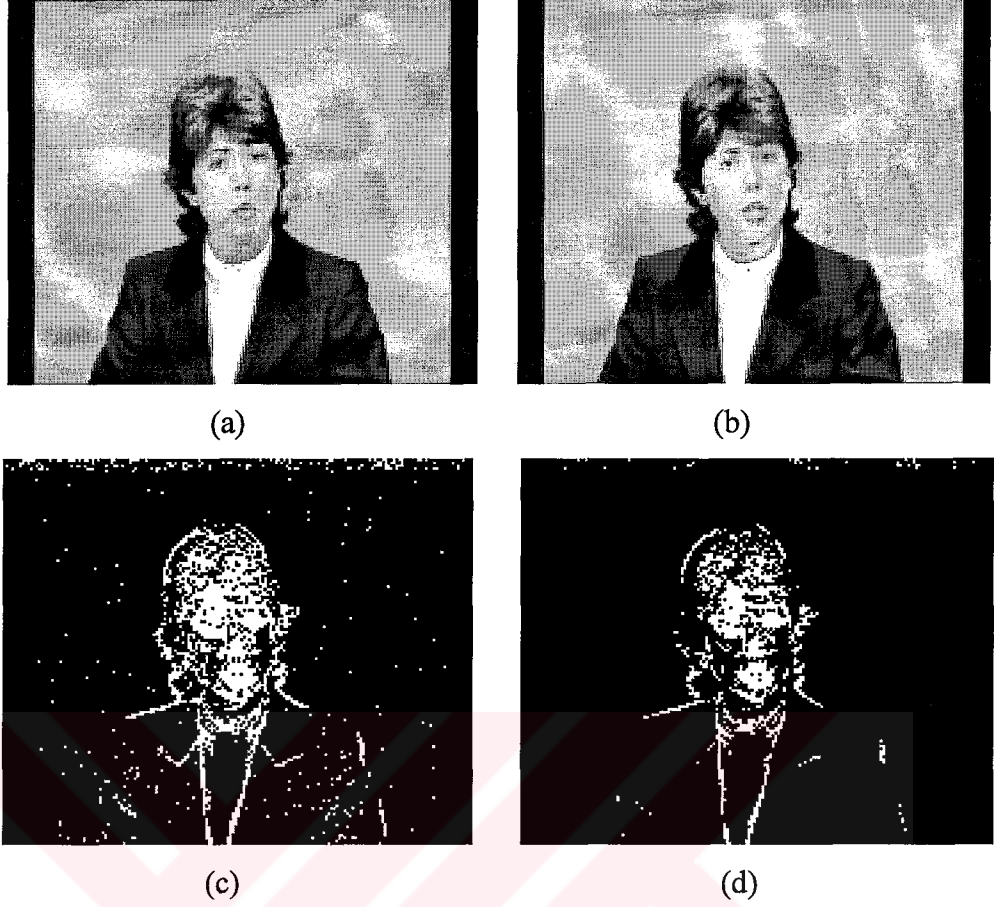




Şekil 4.1. Önerilen video nesnesi bölütleme yöntemin blok şeması.

#### 4.2. Blok Bazlı Değişim Algılama

Değişim algılama kullanan bölütleme yöntemleri imge dizisindeki değişimleri elde etmek için genellikle o anki çerçeve ile bir önceki çerçevenin piksel bazlı farkının eşiklenmesini kullanır. Şekil 4.1’de “Claire” görüntü dizisinin arka arkaya gelen iki çerçevesinin piksel bazlı farkının eşiklenmesi görülmektedir.



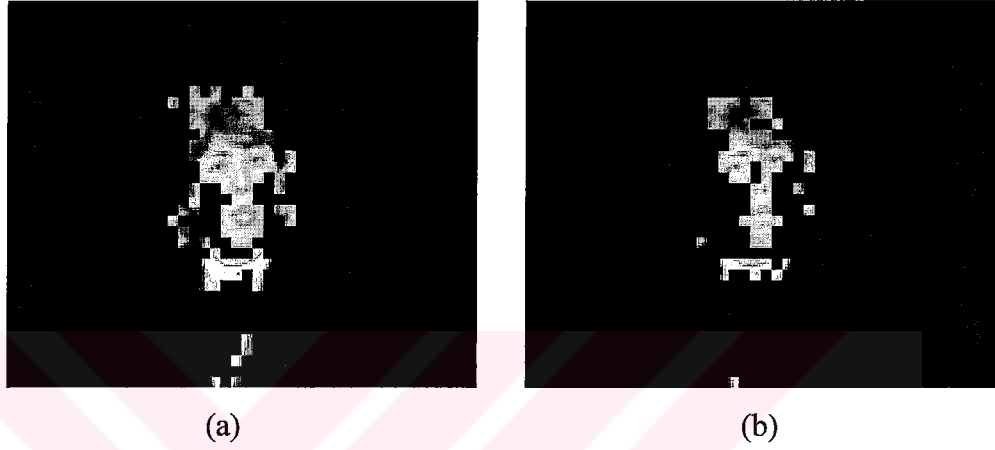
**Şekil 4.2.** Piksel bazlı çerçeve farkı. (a) “Claire” çerçeve #120. (b) “Claire” çerçeve #119. (c) Fark imgesi (eşik =2). (d) Fark imgesi (eşik=3).

Nesne hareketi göreceli olarak yavaş olduğu için düşük eşik değeri kullanılması nesne hareketinin daha doğru elde edilmesini sağlamaktadır ancak beraberinde arka planda algılama gürültüsü ortaya çıkmaktadır (Şekil 4.2(c)). Daha yüksek eşik değeri algılama gürültüsünün üstesinden gelebilir ancak bu durumda elde edilen nesnenin doğruluğu azalacaktır (Şekil 4.2(d)) .

Algılama gürültüsü sorununu aşmak için Chein at al (2002) tarafından piksel tabanlı morfolojik işlemler kullanılmıştır. Bu işlemler algılama gürültüsü üstesinden gelmekte ancak değişim algılama işleminin kendisinden çok daha fazla zaman harcamaktadır.

Bu çalışmada, karşılaşılan algılama gürültüsünü aşmak için piksel bazlı değişim algılama yerine blok farkı tabanlı değişim algılama yapısı önerilmiş ve kullanılmıştır.

Öncelikle imgeler üstü üste binmeyen (non-overlapping) bloklara ayrılmıştır. Blok boyutları 4x4 piksel olacak şekilde seçilmiştir. Blok farkı, her blok için tek bir fark ölçütü bulmak amacıyla blok içerisindeki bütün piksellerin karşılaştırılacak bloktaki ilgili piksellerle farkının toplamının ortalaması ile elde edilir. Ortalamanın alınması piksel bazlı gürültü etkisinin bloklarda eritilmesini sağlar. Şekil 4.3.'de "Claire" görüntü dizininin blok bazlı çerçeve farkı sonuçlarını göstermektedir.

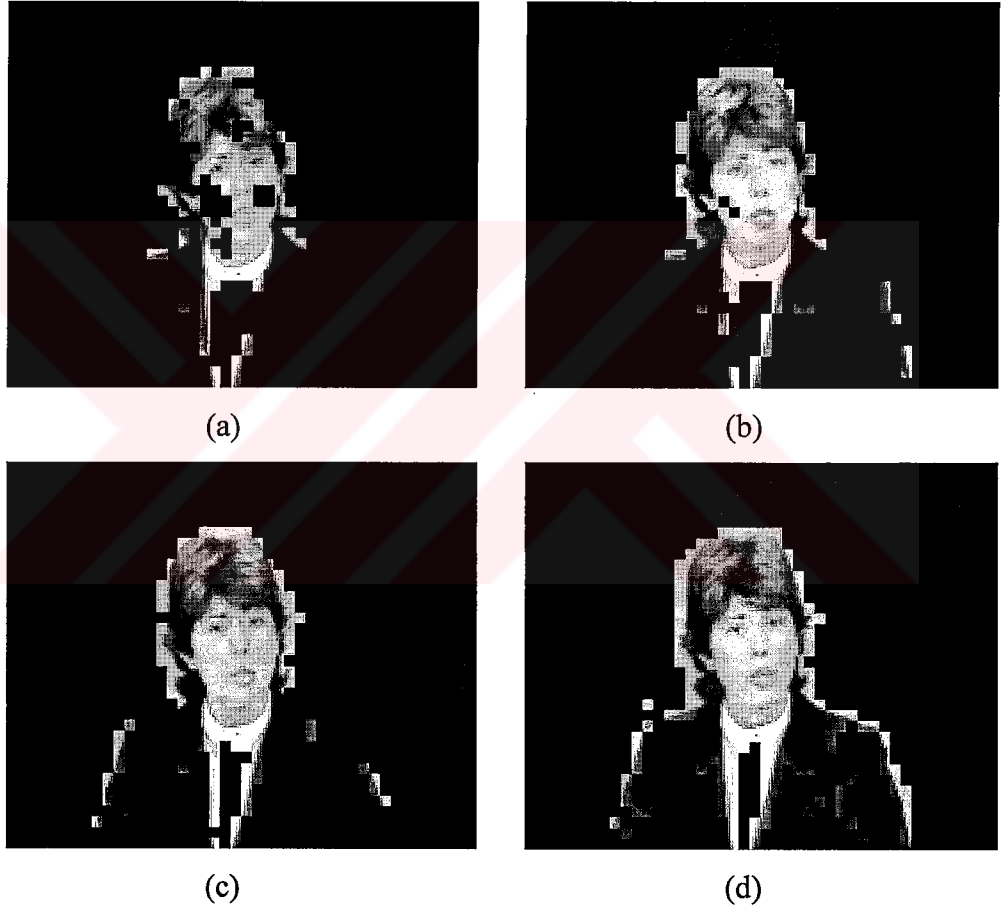


**Şekil 4.3.** Blok bazlı çerçeve farkı. (a) Blok bazlı çerçeve farkı #120,#119 (eşik=2). (b) Blok bazlı çerçeve farkı #120, #119 (eşik=3).

Nesne yavaş hareket ediyor veya geçici olarak duruyorsa o anki ve bir önceki çerçevenin farkı doğru bölütleme için yetersiz olacaktır. Yavaş hareketler göreceli olarak düşük eşik değeri kullanılarak algılanabilir fakat bu durumda algılama gürültüsü ortaya çıkacaktır. Nesne hareketindeki duraklamalar ise arka arkaya gelen çerçevelerin farkı ile algılanamaz. Baş-omuz (head-and-shoulder) tipi görüntü dizinlerinde genellikle nesnenin baş kısmı yeterli hareket gösterirken, omuz kısmı hareketsiz olduğu için arka plan olarak algılanır. Bu etki Şekil 4.2.'de açık şekilde görülmektedir. Göreceli olarak yüksek eşik kullanımı ise gürültüleri engellerken omuz bilgisinin kaybına neden olur. Örneğin, Meier and Ngan (1999) tarafından önerilen yöntem ile "Grandma" görüntü dizininde nesnenin sadece baş kısmı bölütlenebilmiştir.

Klasik değişim algılama kullanan bölütleme yöntemleri sadece bir önceki çerçeve farkının eşiklenmesini kullandığından yavaş hareket veya kısa süreli duraklamalar içeren görüntü dizinlerinde başarılı sonuç vermemektedir. Bu çalışmada, daha eski

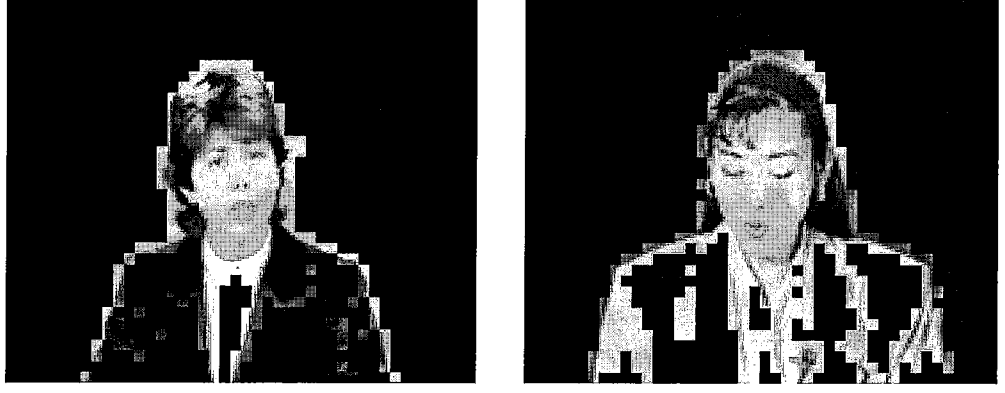
çerçevelerde kullanılarak bu tip görüntü dizinlerinin bölütlenmesi sağlanmıştır. Hesapsal yük göz önüne alındığında değişim algılama için geçmiş birkaç çerçevenin kullanılması kabul edilebilir, sadece gerekli hafıza kullanımı artmaktadır. Blok bazlı değişim algılama için “I” o anki çerçeveyi göstermek üzere I-5, I-10, I-15, I-25, ... dikkate alınmaktadır. Şekil 4.4.’de birkaç geçmiş çerçeve de kullanılarak elde edilen blok farkı bazlı değişim algılama sonuçları görülmektedir. Bu aşamada “Claire” görüntü dizini için 25 çerçeve kadar geriye gidilerek nesne hatları genel olarak elde edilmiştir.



**Şekil 4.4.** Blok farkı bazlı değişim algılama. (a) Blok bazlı fark imgesi (Claire #120, #115). (b) Blok bazlı fark imgesi (Claire #120, #110). (c) Blok bazlı fark imgesi (Claire #120, #105). (d) Blok bazlı fark imgesi (Claire #120, #095) (Eşik=4).

Eğer geri gidilen çerçevelerin hiç birinden ilgili blok için ön plan gelmiyorsa blok arka plan maskesinde arka plan olarak işaretlenip, arka plan imgesine kaydedilir. Bundan sonraki aşamalarda arka plan maskesinde arka plan olarak işaretlenen bloklar için sadece bir önceki çerçeveye bakılarak o bloğun durumuna karar verilir.

Eğer geri gidilen çerçevelerden en az birisinden ilgili blok için ön plan bilgisi geliyorsa blok ön plan olarak alınır. Bu şekilde yapılan blok bazlı arka plan kaydı ile elde edilen bölütleme sonuçları Şekil 4.5.'de verilmiştir.



(a)

(b)

**Şekil 4.5.** Blok bazlı morfoloji öncesi bölütleme sonuçları (a) Claire #120. (b) Akiyo #120.

Görüldüğü gibi piksel tabanlı fark yerine blok tabanlı fark kullanıldığı zaman kamera sensöründen kaynaklanan elde etme gürültülerinin üstesinden gelinmiştir ancak aynı dokuya sahip olan veya yavaş hareket eden bazı bölgeler blok farkı için koyulan eşğin altında kaldığından hala arka plan olarak algılanmaktadır. Bu sorunu aşmak için arka plan imgesi oluşturulmadan önce elde edilen blok bazlı imgeye, piksel bazlı morfolojik işlemlere göre işlem yükü daha az olan blok bazlı morfolojik kapatma (closing) uygulanmıştır.

Öncelikle sadece birlerden oluşan morfolojik yapı ile ön plan bölgeleri yayılmış (dilation) böylece iç bölgelerdeki arka plan olan bloklar ön plan yapılmıştır. Bu işlem sırasında ön plan bölgeleri de yayılmaktadır. Bu yayılmaları geri toparlamak için bu kez aşındırma (erosion) işlemi ile fazla yayılan önplan bölgeleri eski haline getirilmiştir. Şekil 4.6.'da yayma ve aşındırma işlemleri sonucu elde edilen bölütleme sonuçları görülmektedir.



(a)



(b)

**Şekil 4.6.** Blok bazlı morfoloji. (a) Şekil 4.5.(a)'da verilen Claire #120'nin blok bazlı yayma (dilation) işlemi sonrası durumu. (b) Şekil 4.6.(a)'daki imgenin blok bazlı aşındırma (erosion) sonrası durumu.

Bu şekilde güvenli olarak elde edilen arka plan blokları bir arka plan imgesine saklanır. Eğer ilgili blok için arka plan bilgisi daha önceden varsa, o anki imge ile arka plan bilgisi güncellenir. Var olan arka plan bilgilerini izlemek için bir arka plan maskesi kullanılır. Bir blok için bir kez arka plan bilgisi gelmiş ise o blok için daha eski çerçevelere bakmaya gerek yoktur. Bunun yerine sadece arka plan imgesinde var olan arka plan bilgisi ile o anki çerçevenin blok bazlı farkının eşiklenmesi kullanılarak bloğun o çerçevedeki durumuna karar verilir.

### 4.3. Nesne Kenarlarının Oluşturulması

Morfolojik kapatma işleminde sonra elde edilen bölütlenmiş imgeler bloklar şeklindedir. Nesne kenarlarına doğru şekilde yaklaşabilmek için ön plan nesnesinin en dış bloklarında piksel bazlı değişim algılama kullanılmıştır. Yavaş hareketleri ve kısa süreli duraklamaları göz önüne almak için yine daha eski çerçeveler de kullanılmıştır. Eğer bütün çerçevelerden ilgili pikselin fark değerleri belirlenen eşik değerinin altında kaldığı bilgisi gelirse piksel arka plan, aksi takdirde ön plan olarak değerlendirilir. Şekil 4.7. 'de ön plan nesnesinin en dış bloklarının piksel bazlı değişim algılama kullanıldığında elde edilen bölütleme sonuçları gösterilmiştir.





(a)



(b)

**Şekil 4.7.** Kenar bloklarda piksel bazlı değişim algılama. (a) Claire #120. (b) Akiyo #120.

Kenar bloklarda piksel bazlı değişim algılama kullanıldığında “Claire” ve “Akiyo” görüntü dizinlerinde açık şekilde görülen yanlış sınıflandırmalarla karşılaşmıştır. Düşük dereceli polinomsal eğri uydurma kullanılarak nesne kenarları yumuşak ve daha doğru bir şekilde elde edilmeye çalışılmıştır. Ön plan nesnesinin kenarlarındaki piksellerden yeni konumu belirlenecek olan pikselin konumu,  $L$  uzunluğundaki eğrinin ortasına gelecek şekilde yerleştirilir ve bu  $L$  uzunluğundaki eğriye düşük dereceli polinomsal eğri uydurma işlemi uygulanarak ilgili pikselin yeni konumu bulunur. Bu işlem nesne kenarlarındaki bütün pikseller için gerçekleştirilir. Belirlenen nesne kenarlarının içinde kalan pikseller ön plan olarak, dışında kalan pikseller ise arka plan olarak değerlendirilir. Bu işlem sonrasında elde edilen bölütleme sonuçları Şekil 4.8.’de verilmiştir.



(a)



(b)

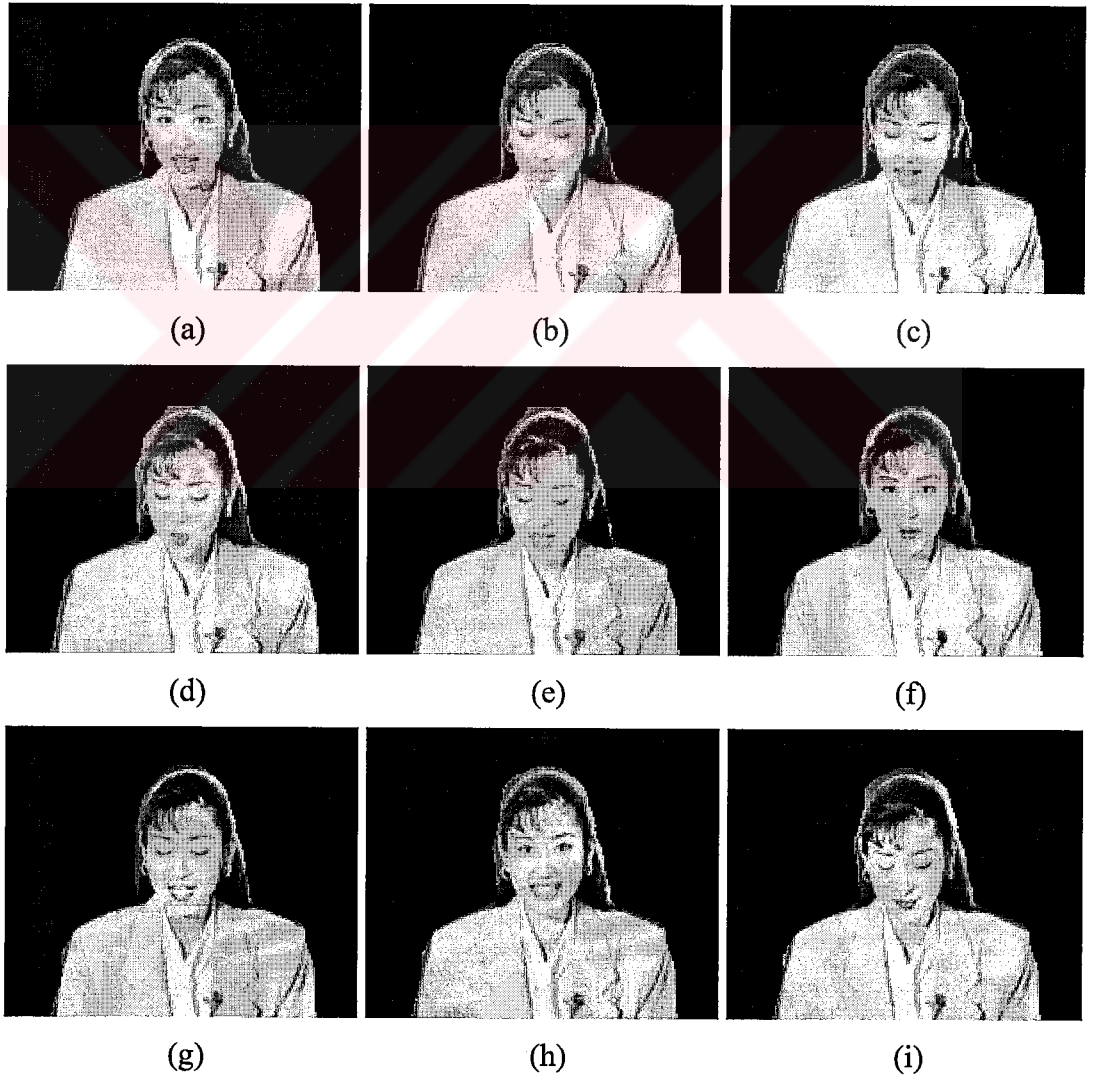
**Şekil 4.8.** Nesne kenarlarında polinomsal eğri uydurma. (a) Claire #120. (b) Akiyo #120.



#### 4.4. Bölütleme Sonuçları

Önerilen video bölütleme yöntemi bilinen bazı video konferans tipi görüntü dizinlerine uygulanmıştır (QCIF tipindeki görünü dizinleri, 176x144 piksel). Elde edilen sonuçlar aşağıda verilmiştir.

“Akiyo” : Bu görüntü dizininde elde etme gürültüsü olmadığından değişim algılama kullanılarak bölütleme yapmak gayet kolaydır. Bölütleme için bazı ara çerçeveler de kullanılarak sadece 25 çerçeve geri gidilmiştir. Bu görüntü dizini için elde edilen bölütleme sonuçları Şekil 4.9.’da verilmiştir.



Şekil 4.9. “Akiyo” görüntü dizini için bölütleme sonuçları. (a) Akiyo #030. (b) #060. (c) #090. (d) #120. (e) #150. (f) #180. (g) #210. (h) #240. (i) #270.

“Claire” : Bu görüntü dizisinde aşırı derecede elde etme gürültüsü bulunmaktadır. Bölütleme için sadece 25 çerçeve geriye gidilmesine rağmen sağlıklı bölütleme sonuçlarının elde edilebilmesi ancak arka plan imgesinin olduğu 75. çerçeveden sonra mümkün olmaktadır. Bölütleme sonuçları Şekil 4.10.’da verilmiştir.



Şekil 4.10. “Claire” dizini bölütleme sonuçları. (a) Claire #075. (b) #100. (c) #125. (d) #150. (e) #175. (f) #200. (g) #225. (h) #250. (i) #275. (j) #300. (k) #325. (l) #350.

“Missam” : Bu görüntü dizininde de “Claire”de olduğu gibi fazla miktarda elde etme gürültüsü bulunmaktadır. Ayrıca bölütleme için sadece ışıklılık bileşeni kullanıldığından ve kahverengi olan arka plan ile, görüntüdeki kişinin saç rengi ve kıyafetinin kol bölgesinin ışıklılık değerlerinin birbirine çok yakın olması bölütlemeyi zorlaştırmıştır. Bu görüntü dizininde bölütleme için 75 çerçeve kadar geriye gidilmiştir. “Missam” için bölütleme sonuçları Şekil 4.11.’de verilmiştir.



(a)

(b)

(c)



(d)

(e)

(f)



(g)

(h)

(i)

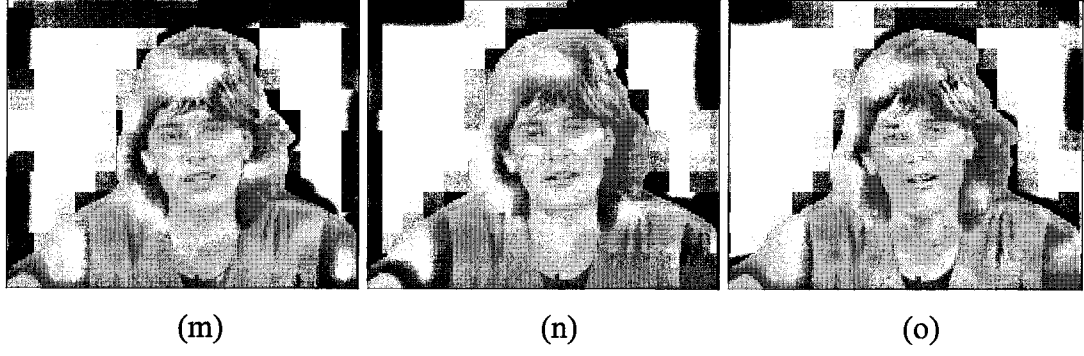


(j)

(k)

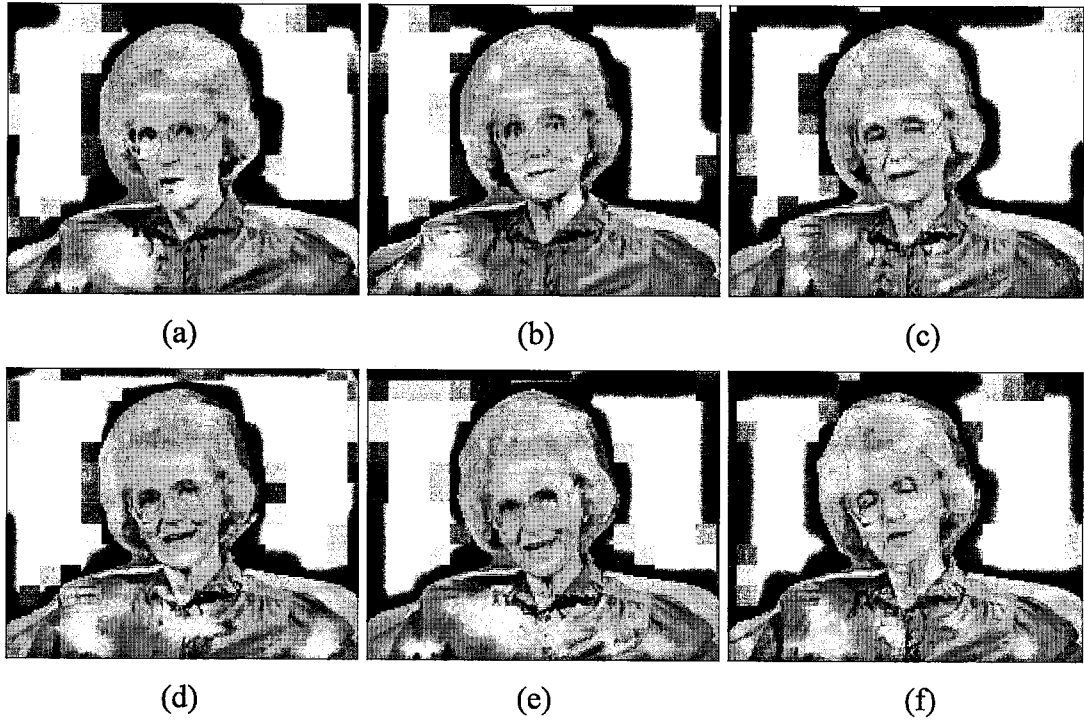
(l)

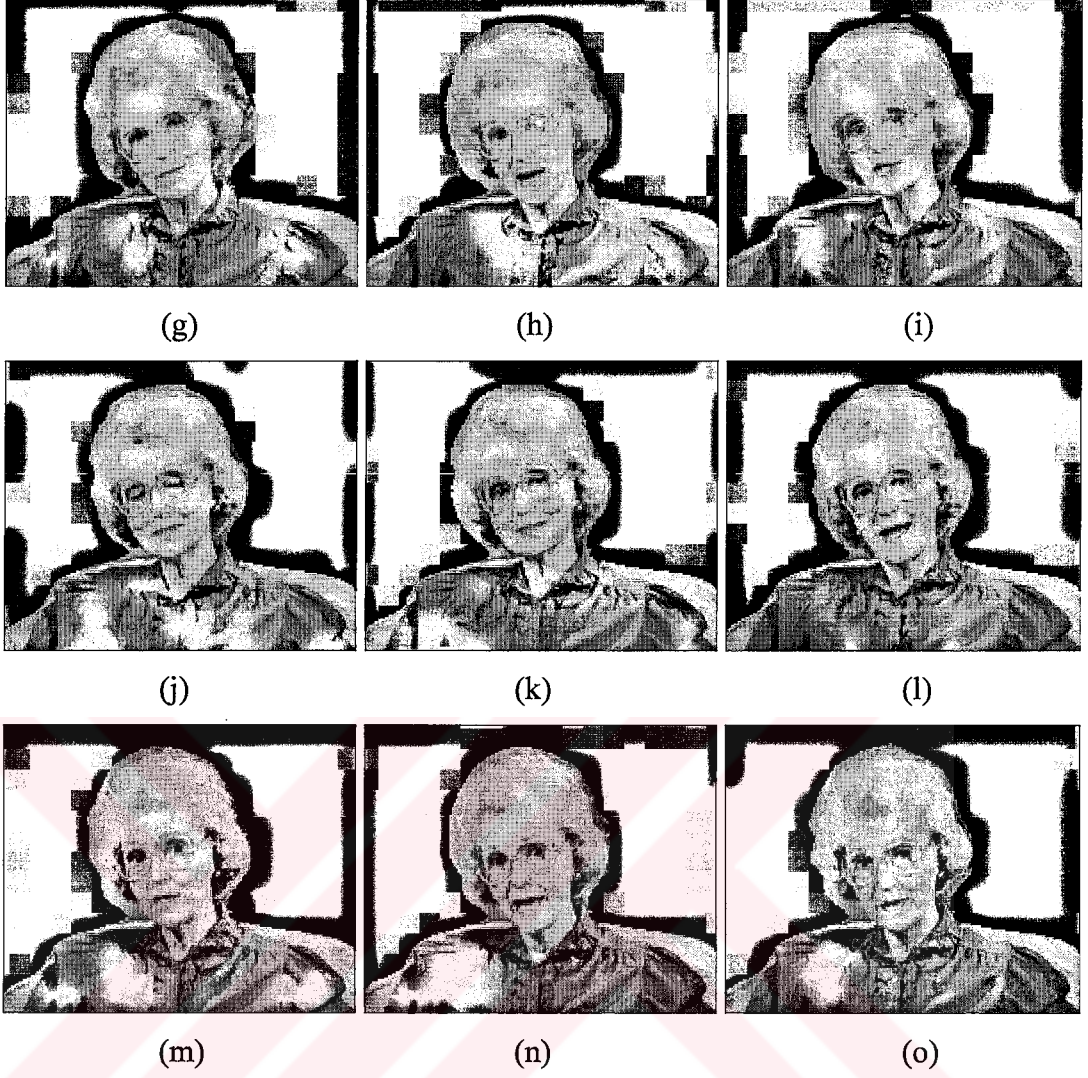




**Şekil 4.11.** "Missam" dizini bölütleme sonuçları. (a) Missam #080. (b) #085. (c) #090. (d) #095. (e) #100. (f) #105. (g) #110. (h) #115. (i) #120. (j) #125. (k) #130. (l) #135. (m) #135. (n) #140. (o) #145.

"Grandma" : Bu görüntü dizini hem elde etme gürültüsü, hem yavaş hareket hem de kısa süreli duraklamalar içermektedir. Bu nedenle bölütlenmesi bir hayli zor bir görüntü dizinidir. Literatürdeki birçok bölütleme yöntemi bu görüntü dizinini önerdikleri yöntemde iyi sonuç vermeyeceği için kullanmamıştır. Meier and Ngan (1999) tarafından önerilen bölütleme yöntemi ise bu görüntü dizininde bölütleme sonucunda sadece baş kısmı elde edebilmektedir. Bu görüntü dizini için çok yavaş hareket ettiği için 100 çerçeve kadar geriye gidilerek bölütleme yapılmıştır. Bölütleme sonuçları Şekil 4.12.'de gösterilmiştir.





**Şekil 4.12.** "Grandma" dizini bölütleme sonuçları. (a) Grandma #120. (b) #150. (c) #200. (d) #250. (e) #300. (f) #350. (g) #400. (h) #450. (i) #500. (j) #550. (k) #600. (l) #650. (m) #700. (n) #730. (o) #760.

## SONUÇLAR ve ÖNERİLER

Bu çalışmada MPEG-4 video kodlama standardının en temel ögesi olan video nesnelere elde edilmesinde kullanılacak zamansal bir bölütleme yöntemi önerilmiştir. Önerilen yöntem temel olarak değişim algılama kullanılmaktadır. Klasik değişim algılama temelli yöntemler art arda gelen çerçevelerin piksel tabanlı farkının eşiklenmesine dayanmaktadır. Ancak değişim algılama, elde etme gürültüsüne karşı aşırı derecede duyarlıdır ve yavaş hareket eden veya kısa süreli duraklayan nesnelere elde edilmesinde etkisiz kalmaktadır.

Klasik değişim algılama yönteminin bu eksikliklerinin üstesinden gelebilmek için çerçeve farkını piksel tabanlı değil blok tabanlı olarak ele alan yeni bir değişim algılama ölçütü önerilmiştir ve kullanılmıştır. Bu sayede elde etme gürültüsü blok bazlı farkın içerisinde eritilebilmektedir. Bununla birlikte yavaş hareket eden veya kısa süreli duraklayan nesnelere algılanmasında sadece bir önceki çerçeve değil geçmiş bazı çerçeveler de göz önüne alınmıştır. Böylece yavaş hareket eden veya kısa süreli duraklayan nesnelere bölütlenmesi sağlanmıştır. Örneğin, literatürde önerilen video nesnesi bölütleme yöntemlerinin test edilmesinde, çok yavaş hareket eden ve kısa süreli duraklayan "Grandma" görüntü dizini genellikle iyi bölütleme sonuçları elde edilemediğinden kullanılmamıştır. Meier and Ngan (1999) önerdikleri yöntemle bu görüntü dizininin sadece baş kısmı bölütleyebilmiştir.

Önerilen yöntemin bir diğer avantajı ise yüksek hesapsal yük getiren işlemler kullanılmadığından gerçek zamanlı çalıştırmaya uygun olmasıdır. Yavaş hareketlerin algılanması için geçmiş bazı çerçevelerin de kullanılması bir miktar işlem yükü getirmekle birlikte, diğer bölütleme yöntemlerinin eksikliği olan yavaş hareketlerin algılanması avantajı sağlanmıştır. Arka plan olan bölgeler için sadece bir çerçeve geriye gidilmesi bu işlem yükünü de bir miktar azaltmıştır.

Önerilen yöntem sabit bir arka plana gereksinim duymaktadır. Video telefon ve video konferans uygulamalarında kamera sabittir ve bu çalışmada bu uygulamalar

üzerinde durulmuştur. Yöntemin kameranın hareketli olduğu görüntü dizinlerinde bölütleme yapabilmesi için hareket dengelemesi kullanılması gerekmektedir. Bazı araştırmacılar kamera hareketi nedeni ile oluşacak hareketin dengelenmesi için bütünsel hareket kestirimi ve dengelemesi kullanmıştır (Mech and Wollborn 1998, Kim et al 1999, Tsaig and Averbuch 2002).

Önerilen yöntemde kullanılan eşik değerler her görüntü dizini için denenerek bulunmuştur. Arka plan kaydı yapılırken geçmiş bazı çerçevelerin dikkate alındığı durumlarda bloklar için kullanılan ortalama eşik değeri 1 ile 1.5 arasındadır. Arka plan bilgisi mevcut blokların o anki çerçeve için durumu belirlenirken kullanılan ortalama eşik değeri 1.5 ile 2 arasındadır. Piksel tabanlı ayırım için ise eşik değeri olarak 2 kullanılmıştır. İleri çalışmalarda bu eşik değerleri için uyarlamalı değeri tespiti üzerinde çalışılarak eşik değerleri otomatik olarak elde edilebilir.

Önerilen yöntem diğer değişim algılama temelli yöntemler gibi arka planda oluşacak hareketli gölge durumunda bu bölgeleri ön plan olarak değerlendirmektedir. Bunun için ileriki çalışmalarda fazla hesapsal yük getirmeyen işlemler kullanılarak gölge tespiti yapılarak, önerilen yöntemin gölge etkisi içeren görüntü dizinlerinde düzgün çalışması sağlanabilir.



## KAYNAKLAR

1. LONG, M., 1995. Understanding MPEG-2 Digital Video Compression. MLESAT Web Site , <http://www.mlesat.com/article7.html>
2. TEKALP, A.M., 1995. Digital Video Processing. Prentice Hall Press
3. EBRAHIMI, T. and HORNE, C., 2000. MPEG-4 Natural Video Coding – An overview. Signal Processing : Image Communication, Elsevier Press, Vol. 15, 365-385.
4. ELY, S.R., 1995. MPEG Video – A Simple Introduction. EBU (European Broadcasting Union) Technical Review, 12-23.
5. SIKORA, T., 1997. The MPEG-4 Video Standard Verification Model. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 7, No: 1, 19-31.
6. ZHANG, D. and LU, G., 2001. Segmentation of Moving Object in Image Sequence : A Review. Circuits, Systems and Signal Processing (Special Issue on Multimedia Communication Services), 20(2) : 143-183.
7. DENG, Y. and MANJUNATH, B.S, 1998. NeTra-V : Toward an Object-Based Video Representation. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8, No: 5, 616-626.
8. MEIER, T. and NGAN, K.N., 1998. Automation Segmentation of Moving Object for Video Object Plane Generation. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8, No: 5, 525-538.
9. TORR, P.H.S, 1995. Motion Segmentation and Outlier Detection. Ph.D Thesis, University of Oxford.
10. MACLEAN, W.J., 1996. Recovery of Egomotion and Segmentation of Independent Object Motion Using The EM-Algorithm. Ph.D Thesis, University of Toronto.
11. PEREZ, D.G., GU, C., SUN, M.T., 2001. Semantic Video Extraction Using Four-Band Watershed and Partition Lattice Operators. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No: 5, 603-618.
12. ZHONG, D. and CHANG, S.F., 1999. An Integrated Approach for Content-Based Video Object Segmentation and Retrieval. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No: 8, 1259-1268.

13. ZHOU, J.Y., Ong, E.P., Ko, C.C., 2000. Video Object Segmentation and Tracking for Content-Based Video Coding. International Conference on Multimedia and Expo. (ICME), Vol. 3, 1555-1558.
14. WANG, D., 1998. Unsupervised Video Segmentation Based on Watersheds and Temporal Tracking. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 8, No: 5, 539-646.
15. TSAIG, Y. and AVERBUCH, A., 2002. Automatic Segmentation of Moving Objects in Video Sequences : A Region Labeling Approach. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No: 7, 597-612.
16. PARK, H., SCHOEPFLIN T., and KIM, Y., 2001. Active Contour Model with Gradient Directional Information : Directional Snake. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 11, No: 2, 252-256.
17. PRECIOSO, F. and BARLAUD, M. , 2002. Regular Spatial B-Spline Active Contour for Fast Video Segmentation. International Conference on Image Processing (ICIP), Vol. 2, 761-764.
18. ZALETELJ, J. and TASIC, J.F, 2001. Video Object Segmentation Based on Edge Tracking. International Conference on Image Processing (ICIP), Vol. 2, 813-816.
19. SUN, S., HAYNOR, D.R. and KIM, Y., 2003. Semiautomatic Video Object Segmentation Using Vsnakes. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 13, No: 1, 75-82.
20. KIM, C. and HWANG, J.N., 2002. Fast and Automatic Video Object Segmentation and Tracking for Content-Based Applications. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No: 2, 122-129.
21. MECH, R. and WOLLBORN, M., 1998. A Noise Robust Method for 2D Shape Estimation of Moving Objects in Video Sequences Considering a Moving Camera. Signal Processing, Vol. 66, 213-217.
22. NERI, A., COLONNESE, S., RUSSO, G. and TALONE, P., 1998. Automatic Moving Object and Background Separation. Signal Processing, Vol. 66, 219-232.
23. KIM, M., CHOI, J.G., KIM, D., LEE, H., LEE, M.H., AHN, C. and HO, Y.S, 1999. A VOP Generation Tool : Automatic Segmentation of Moving Objects in Image Sequences Based on Spatio-Temporal Information. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No: 8, 1216-1226.

24. LONG, F., FENG, D., PENG, H. and SIU, W.C., 2001. Extracting Semantic Video Objects, IEEE Computer Graphics and Applications. Vol. 21, Issue 1, 48-55
25. CHEIN, S.Y., MA, S.Y. and CHEN L.G., 2002. Efficient Moving Object Segmentation Algorithm Using Background Registration Technique. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 12, No: 7, 577-586.
26. MEIER, T. and NGAN, K.N., 1999. Video Segmentation for Content-Based Coding. IEEE Transactions on Circuits and Systems for Video Technology, Vol. 9, No: 8, 1190-1203.



## KİŞİSEL YAYINLAR VE ESERLER

1. URHAN, O., ERTÜRK, S., 2003. Arka Plan Kaydına Dayanan Blok Bazlı Hızlı Video Nesne Bölütlemesi. 11. Sinyal İşleme ve İletişim Uygulamaları Kurultayı (SİU2003) Bildirileri Kitabı, 18-20 Haziran 2003, İstanbul, TÜRKİYE, 335-338.



## ÖZGEÇMİŞ

1980 yılında Erzurum'da doğdu. İlk öğrenimini Erzurum, orta ve lise öğrenimini Kocaeli'de tamamladı. 1997 yılında girdiği Kocaeli Üniversitesi Mühendislik Fakültesi Elektronik ve Haberleşme Mühendisliği Bölümü'nden 2001 yılında Elektronik ve Haberleşme Mühendisi olarak mezun oldu. 2001 yılında Kocaeli Üniversitesi Fen Bilimleri Enstitüsü Elektronik ve Haberleşme Ana Bilim Dalı'nda Yüksek Lisans öğrenimine başladı. 2003 yılında mezun olma durumundadır.

2001 yılında beri Kocaeli Üniversitesi Mühendislik Fakültesi Elektronik ve Haberleşme Bölümü Kontrol ve Kumanda Sistemleri Ana Bilim Dalı'nda Araştırma Görevlisi olarak görev yapmaktadır.