

İSTANBUL TEKNİK ÜNİVERSİTESİ ★ FEN BİLİMLERİ ENSTİTÜSÜ

**FARKLI DALGACIK TABANLARINDA DIRAC
VIDEO KODLAYICI PERFORMANS ANALİZİ**

YÜKSEK LİSANS TEZİ

Deniz ÖZENLİ

Anabilim Dalı : Elektronik ve Haberleşme Mühendisliği

Programı : Telekomünikasyon Mühendisliği

AĞUSTOS 2011

**FARKLI DALGACIK TABANLARINDA DIRAC
VIDEO KODLAYICI PERFORMANS ANALİZİ**

YÜKSEK LİSANS TEZİ

Deniz ÖZENLİ

(504091357)

Tezin Enstitüye Verildiği Tarih : 26 Ağustos 2011

Tezin Savunulduğu Tarih : 08 Eylül 2011

Tez Danışmanı : Prof. Dr. Melih PAZARCI (İTÜ)
Diğer Jüri Üyeleri : Prof. Dr. Bilge GÜNSEL (İTÜ)
Prof. Dr. Mühittin GÖKMEN (İTÜ)

AĞUSTOS 2011

ÖNSÖZ

Tez çalışmalarım sırasında büyük destek aldığım danışman hocam Prof. Dr. Melih PAZARCI başta olmak üzere tüm yardımları geçen herkese teşekkürlerimi bir borç bilirim. Bu tezi sevgili aileme ithaf ediyorum.

Ağustos 2011

Deniz ÖZENLİ
(Mühendis)

İÇİNDEKİLER

Sayfa

ÖNSÖZ.....	v
İÇİNDEKİLER.....	vii
KISALTMALAR	ix
ÇİZELGE LİSTESİ.....	xi
ŞEKİL LİSTESİ.....	xiii
ÖZET.....	xvii
SUMMARY	xix
1. GİRİŞ	1
1.1 Tezin Amacı	2
1.2 Literatür Özeti	2
1.3 Ele Alınan Problem	7
1.4 Çalışmada Yapılanlar	7
2. ISO VE JVT EKİPLERİNİN SON DÖNEM VIDEO ÇÖZÜCÜ	
ALGORİTMALARI	9
2.1 MPEG-2	9
2.1.1 MPEG-2'de temel resim türleri.....	10
2.1.1.1 Ayrık kosinüs dönüşümü (DCT).....	12
2.1.1.2 Hareket kestirimi ve kompanzasyonu	14
2.1.1.3 Geçmeli ve geçmesiz video dizisi modelleri.....	15
2.1.1.4 YUV video dizileri.....	16
2.1.2 MPEG-2'de temel işlem blokları	17
2.1.2.1 Giriş-çıkış bit oranı kontrolü.....	19
2.1.2.2 Profiller	20
2.1.3 MPEG-2'de temel işlem seviyeleri	23
2.2 H.264 ve Genel Özellikleri.....	23
2.2.1 H.264'te temel profil yapıları.....	24
2.2.2 H.264'de temel dönüşüm teknikleri.....	26
2.2.3 Ölçekleme ve kuantalama.....	28
2.2.4 Entropi kodlama	29
2.2.5 Kodlama verimliliği arttıran temel araçlar	30
2.2.5.1 Bloklaşma giderici filtreleme	30
2.2.5.2 İçsel kestirim ve makro blok kenarları eklentileme işlemi	31
2.2.5.3 Makro blok segmentasyonu	33
2.2.6 Sonuçlar	34
3. DIRAC ALGORİTMASI	37
3.1 Dirac Kodlayıcı Yapısı.....	37
3.2 Dirac Dalgacık Sistematiği ve Kaldırma (Lifting) Yöntemi	40
3.2.1 Biortogonal dalgacıklar ve 'Bior 5/3' yapısı	42
3.3 Kuantalama.....	46
3.4 Dirac'da Çerçeve İçi ve Çerçeveler Arası Kestirim	46
3.5 Serpiştirilmiş Üstel Golomb Kodlaması ve İkileştirme Mekanizması.....	47

3.6 Aritmetik Kodlama	49
3.7 Dirac Kod Çözücü Yapısı	50
4. DENEYSEL SONUÇLAR.....	51
4.1 Giriş	51
4.2 QCIF ve CIF Videolarda 1, 1/2, 1/4 ve 1/8 Pel Hareket Kestirimleri İçin Farklı Dalgacık Tabanlarında Dirac Performans Analizi	53
4.3 Bior 5/3 Dalgacığın Dirac Performansına Etkisi ve Farklı Dalgacık Tabanlarında Dirac Performansı Genel Analizi	58
4.4 Bior 5/3 Tabanında Dirac Performansının H.264 ile Farklı Açılardan Kıyaslanması	62
5. SONUÇLAR ve ÖNERİLER	67
EKLER.....	71

KISALTMALAR

AVC	: Advanced Video Coding
API	: Application Programming Interface
BBC	: British Broadcasting Corporation
CABAC	: Context-Based Adaptive Binary Arithmetic Coding
CAVLC	: Context-Based Adaptive Variable Length Coding
CBR	: Constant Bit Rate
CIF	: Common Intermediate Format (352x288)
4CIF	: Four Common Intermediate Format (704x576)
DCT	: Discrete Cosine Transform
EFS	: Encoded File Size
GOP	: Group of Picture(s)
HDTV	: High Definition Television (1920x1080)
HF	: High Frequency
HH	: High - High
HL	: High – Low
IDCT	: Inverse Discrete Cosine Transform
ITU	: International Telecommunication Union
IQ	: Inverse Quantization
JVT	: Joint Video Team
LH	: Low - High
LL	: Low – Low
MCP	: Motion Compensated Prediction
MPEG	: Moving Picture Experts Group
MSE	: Mean Square Error
MV	: Motion Vector
OBMC	: Overlapped Block-Based Motion Compensation
Q	: Quantization
QCIF	: Quarter Common Intermediate Format (176x144)
QF	: Quality Factor
PSNR	: Peak Signal to Noise Ratio
RDO	: Rate Distortion Optimization
SAD	: Sum of The Absolute Difference
SDTV	: Standard Definition Television (720x576)
SMPTE	: Society of Motion Picture and TV Engineers
SSIM	: Structural Similarity Index
VBR	: Variable Bit Rate
VCEG	: Video Coding Experts Group
VLC	: Variable Length Coding
VLD	: Variable Length Decoding
VLSI	: Very Large Scale Integration
WMV	: Windows Media Video
YUV	: Luminance and Chrominance Color Components

ÇİZELGE LİSTESİ

Sayfa

Çizelge 2.1 : MPEG-2 temel işlem seviyeleri [8,9].....	23
Çizelge 3.1 : Dirac'da kullanılan temel dalgacık fonksiyonları [3].	40
Çizelge 3.2 : Serpiştirilmiş üstel Golomb kodlaması [3,36].....	48
Çizelge 4.1 : Analizlerde kullanılan video dizilerinin temel bilgileri.	52
Çizelge 4.2 : Daubechies 9/7 dalgacık yapısından doğan yeni durumlar için kaldırma yöntemi katsayıları [38].	54
Çizelge 4.3 : 'Foreman.qcif' için hareket kestirimi mekanizmasının farklı dalgacık tabanlarında sıkıştırmaya etkisi.	55
Çizelge 4.4 : 'Miss-america.qcif' için hareket kestirimi mekanizmasının farklı dalgacık tabanlarında sıkıştırmaya etkisi.	56
Çizelge 4.5 : 'Bus.cif' için hareket kestirimi mekanizmasının farklı dalgacık tabanlarında sıkıştırmaya etkisi.	57
Çizelge 4.6 : 'Carphone.qcif' için Bior. 5/3 dalgacığının da katılarak tüm dalgacık tabanlarında Dirac performans analizi.	58
Çizelge 4.7 : 'Bus.cif' için Bior. 5/3 dalgacığının da katılarak tüm dalgacık tabanlarında Dirac performans analizi.	59
Çizelge 4.8 : 'Harbour.4cif' için Bior. 5/3 dalgacığının da katılarak tüm dalgacık tabanlarında Dirac performans analizi.	60

ŞEKİL LİSTESİ

Sayfa

Şekil 1.1	: Dalgacık dönüşümünün kaldırma yöntemiyle gerçekleşmesi [5].	6
Şekil 2.1	: MPEG-2 genel kodlayıcı yapısı [6].	9
Şekil 2.2	: MPEG-2 video dizisi hiyerarşisi [6].	10
Şekil 2.3	: MPEG-2 Temel resim türleri-1 [7].	10
Şekil 2.4	: MPEG-2 Temel resim türleri-2 [7].	11
Şekil 2.5	: Kod çözücüye iletilmek için bekleyen henüz sıralanmamış resim dizisi [6].	11
Şekil 2.6	: Kod çözücüye iletilecek olan ve sıralanmış resim dizisi [6].	11
Şekil 2.7	: Ayrık kosinüs dönüşümünde giriş verisinin '2n' ile periyodik olması [8,10].	13
Şekil 2.8	: İki boyutta DCT taban vektörleri [6].	13
Şekil 2.9	: Genel hareket kestirimi mekanizması [6].	14
Şekil 2.10	: MPEG-2 hareket kestirimi örneği [6].	15
Şekil 2.11	: Geçmeli ve geçmesiz video dizisi modelleri (Üstteki dizi geçmesiz model, alttaki dizi ise geçmeli modeldir.) [6].	15
Şekil 2.12	: YUV video örnekleme biçimleri (Soldan sağa sırasıyla 4:2:0 4:2:2 4:4:4 ve beyaz-Y açık gri-U koyu gri-V) [6].	16
Şekil 2.13	: MPEG-2 kodlayıcı bölümü [8,9].	17
Şekil 2.14	: MPEG-2 kod çözücü bölümü [6].	17
Şekil 2.15	: MPEG-2 video bit dizisi akışı içerik modeli [6,11,12].	18
Şekil 2.16	: MPEG-2'de kullanılan temel kuantalama matrisleri [6].	19
Şekil 2.17	: SNR profil MPEG-2 video kodlayıcı [8,9].	21
Şekil 2.18	: Uzaysal profil MPEG-2 video kod çözücü [8,9].	22
Şekil 2.19	: H.264'de temel profil yapıları [1,13].	25
Şekil 2.20	: H.264 genel kodlayıcı yapısı [1].	26
Şekil 2.21	: H.264'de çerçeve içi kestirim modunda dönüşüm organizasyonu [1].	27
Şekil 2.22	: Kroma katsayılarının çerçeve içi kestirim modunda Hadamard dönüşümüne hazırlanması [1].	28
Şekil 2.23	: Genel anlamda H.264 kodlayıcısı içyapısı [1,13].	28
Şekil 2.24	: Örnek cif bir video dizisi için CABAC ve CAVLC performans karşılaştırması ('Mobile.cif' için yapılmıştır.) [1].	29
Şekil 2.25	: H.264'de temel rezüdi tarama biçimleri.	30
Şekil 2.26	: H.264'de bloklaşma etkisini giderici filtreleme işlemi [1].	31
Şekil 2.27	: Filtreleme işleminin video kalitesine etkisi [12].	31
Şekil 2.28	: Sınır eklenmesi (boundry padding) [6].	32
Şekil 2.29	: Çerçeve içi kestirim [6].	32
Şekil 2.30	: 4x4'lük makro bloklar için çerçeve içi kestirim modları [1].	33
Şekil 2.31	: Farklı hareket vektörleri kombinezonları [6].	33
Şekil 2.32	: H.264'de örnek bir çerçeve için makro blok segmentasyonu [6].	34

Şekil 3.1	: Genel Dirac kodlayıcı mimarisi [1].	37
Şekil 3.2	: Hiyerarşik hareket kestirimi [17].	38
Şekil 3.3	: Dirac'da temel kestirim tarama yapıları [1].	38
Şekil 3.4	: Hareket kestiriminde örtüşen blokların gösterimi [1].	39
Şekil 3.5	: Dalgacık analizi eş değer modeli [18,19].	41
Şekil 3.6	: Kaldırma (lifting) metodu [5,18].	41
Şekil 3.7	: Dalgacık alt bant sistematığı [20,21].	42
Şekil 3.8	: 'Bior 5/3' süzgeç sistematığı [29].	44
Şekil 3.9	: 'Bior 5/3' süzgeç yapısı (sürekli zaman gösterimi).	45
Şekil 3.10	: 'Bior 5/3' süzgeç yapısı (ayrık zaman gösterimi-orijine taşınmış hali).	45
Şekil 3.11	: Ölü bölge kuantalaması [1].	46
Şekil 3.12	: Dirac'da RG (resimler grubu)'yi oluşturma [1].	46
Şekil 3.13	: Dirac entropi kodlama blok gösterim [1,35].	47
Şekil 3.14	: Aritmetik kodlama mekanizması ('Squeeze kelimesi üzerinde örneklendirilmiştir. ') [15].	49
Şekil 3.15	: Dirac kod çözücü sistematığı [1,32].	50
Şekil 4.1	: Analizlerde kullanılan QCIF-YUV videolar (Soldaki 'miss- america.qcif' ortadaki 'carphone.qcif' ve sağdaki 'foreman.qcif') [41].	51
Şekil 4.2	: Analizlerde kullanılan CIF-YUV ve 4CIF-YUV videolar (Soldaki 'bus.cif' ortadaki 'stephan.cif' ve 'sağdaki 'harbour.4cif') [41].	52
Şekil 4.3	: Yeni dalgacık modelleri olarak sunulan Durum 1, 2, 3 ve 4 için analiz yapısı [38,42].	54
Şekil 4.4	: Yeni süzgeç yapısı (Bior. 5/3) ile Dirac-Le Gall süzgecin karşılaştırılması (Carphone. QCIF).	60
Şekil 4.5	: Yeni süzgeç yapısı (Bior. 5/3) ile Le Gall süzgecin karşılaştırılması (Bus. CIF).	61
Şekil 4.6	: Yeni süzgeç yapısı (Bior. 5/3) ile Le Gall süzgecin karşılaştırılması (Harbour. 4CIF).	61
Şekil 4.7	: Dirac (bior 5/3) için performans kıyaslaması (kodlama zamanı).	62
Şekil 4.8	: Dirac (bior 5/3) için performans kıyaslaması (kodlama zamanı).	63
Şekil 4.9	: Dirac (bior 5/3) için performans kıyaslaması (kod çözme zamanı).	63
Şekil 4.10	: Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Miss-america.QCIF (PSNR-Y).	64
Şekil 4.11	: Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Stephan.CIF (PSNR-Y).	64
Şekil 4.12	: Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Miss-america.QCIF (SSIM).	65
Şekil 4.13	: Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Stephan.CIF (SSIM).	65
Şekil A.1.1	: Örnek bir YUV video çerçevesi (En üstteki tüm YUV çerçevesi, ikinci sıra soldaki 'Y', sağdaki 'U' ve en alttaki de 'V' bileşenleridir.) [45].	72
Şekil A.1.2	: Elektromanyetik spektruma göre insan gözünün çalışma evreleri ve farklı dalga boylarına verdiği tepki-1 [46].	74
Şekil A.1.3	: Elektromanyetik spektruma göre insan gözünün çalışma evreleri ve farklı dalga boylarına verdiği tepki-2 [46].	74
Şekil A.1.4	: Dış ortam parlaklık seviyesine göre insan gözünün verdiği tepkinin incelenmesi [46].	75
Şekil A.1.5	: YUV uzayı farklı örneklenme biçimleri (En üst sol: 4:2:0, sağ: 4:2:2, en alt: 4:4:4 örnekleme. 4:2:0 ve 4:4:4 'de kesikli çerçeveye alınmış kısımlar 4 pikseli göstermektedir.) [45].	76

Şekil A.2.1 : Dirac SSIM mekanizması [1,47].	77
Şekil A.2.2 : PSNR ile SSIM farkının irdelenmesi [1].	78

FARKLI DALGACIK TABANLARINDA DIRAC VIDEO KODLAYICI PERFORMANS ANALİZİ

ÖZET

Video sıkıştırmanın belki daha önceden hiç taşımadığı kadar çok misyon ve önem taşıdığı bir çağda yaşamaktayız. İnternet ve yüksek kaliteli medya teknolojilerinin hızlı gelişimi ile birlikte video kodlama tekniklerindeki iyileştirmelere duyulan ihtiyaç giderek artmaktadır. Dirac video çözücü bu anlamda bu konudaki araştırmacılara yeni gelebilir fakat özellikle açık kaynak kodlarını kullanıcıya sunması ve lisans ödemeleri gerektirmemesi gibi nitelikleri bakımından dikkat çekicidir. H264/MPEG 4 PART-10 AVC video çözücü sayısal platformda artan ihtiyaçları karşılamak için geliştirilmiş ve üst seviye niteliklerle donanmış son dönemdeki en iyi video kodlayıcı olarak düşünülse de, Dirac, özellikle düşük bit oranlarında görüntü kalitesindeki performansının etkileyiciliği, diğer kodlayıcılara kıyasla daha az hesaplama yükü gereksinimi, ‘basitliği koru’ gibi özel ve kullanıcı dostu bir çalışma felsefesiyle işe başlaması ve belki de her şeyden önemlisi herhangi bir lisans ödemesi ödemesi gerektirmemesi açısından kullanıcıya yepyeni olanaklar getirmektedir.

Bu tez, farklı dalgacık tabanlarında QCIF, CIF ve 4CIF giriş test video dizileri için hem sabit hem de değişken bit oranı ölçümlerini dikkate alarak Dirac video kodlayıcının performansını analiz etmek üzere hazırlanmıştır. Analizler PSNR, SSIM ve sıkıştırma oranı kriterlerine göre değerlendirilmiştir. Ayrıca bu tezde, kodlayıcı ve kod çözücü performansları, çözücü kalite faktörü sabit tutulup çerçeve indeksi üzerinden incelenerek normal şartlarda çözücünün incelenen video üzerinde ortaya çıkardığı sıkıştırma etkisi göz önüne alınarak en iyi şekilde irdelenmeye çalışılmıştır.

DIRAC VIDEO ENCODER PERFORMANCE ANALYSIS IN BASES OF DIFFERENT WAVELETS

SUMMARY

We are living in a century that video compression has a more important mission than it has ever had. With the rapid growth of internet and high-definition media technology, the need for improvement in video encoding techniques has been increasing constantly. The Dirac video codec may be new for researchers in this area in this sense but in the view of its having open source codes and requiring no licence payments, it is worthy of attention. H264/MPEG 4 Part-10 AVC video codec is currently known as the best video codec that has been developed in order to cover the increasing demands of the digital platform and is also equipped with advanced level skills, however Dirac has brought new means. Especially its performance of video quality at low bit rates, reduced computational burden relative to other coders, its special and user friendly 'keep it simple' philosophy, and its not requiring 'licence fees' are the most important aspects.

This thesis was prepared to analyze the performance of Dirac video codec for QCIF CIF and 4CIF input sequences in basis of different wavelets by considiring CBR and VBR measurements. The analysis were evaluated in terms of PSNR, SSIM and compression ratio criteria. Furthermore, in this thesis, the performances of encoder and decoder were observed at constant quality and the resulting compression under normal conditions was considered over frame index.

1. GİRİŞ

Sayısal haberleşme açısından baktığımızda, video sıkıştırma, önemi küçümsenemez ve çokluortam (multimedia) kullanımını son derece verimli hale getiren birleştirici bir rol üstlenmiş konumdadır. Bu anlamda pek çok farklı sıkıştırma teknikleri, kısıtlı depolama olanakları ve sınırlı haberleşme bandı sebebiyle sıkça ihtiyaç duyulan ve günlük hayatımızda farkında olarak ya da olmadan kullandığımız haberleşme birimleri olarak kendilerine sağlam bir yer edinmiş durumdadırlar. Bu tezde ele aldığımız Dirac video kodlayıcı da bu anlamda, bahsettiğimiz ölçütler dikkate alınarak incelendiğinde etkileyici bir performansa sahip olması açısından kullanıcılara yeni olanaklar getirmektedir [1].

BBC tarafından geliştirilmiş, açık kaynak kodlarını kullanıcılarla paylaşan ve son dönemde adından sıkça söz ettiren bir video çözücü olarak karşımıza çıkan Dirac, ismini ünlü İngiliz bilim adamı Paul Dirac'dan almaktadır. Kodlayıcının temel amacı, web üzerindeki videolardan yüksek çözünürlüklü içeriklere kadar bütün bu video dizilerinde olabildiğince yüksek kalitede sıkıştırılmış yeni içerikler üretebilmektir. Üretilen bu yeni formattaki içeriklerin, H264 ve WMV gibi var olan diğer video formatlarının üzerinde bir kalitede olması amaçlanmaktadır. Dirac, bu anlamda ilk gerçek sınavını da 2008'deki Beijing Olimpiyatları'nda BBC tarafından yayınlarda kullanılmak üzere vermiştir ve bu konudaki uzmanların beğenisini kazanmıştır. kalitesi ve genel anlamda başarı oranı bakımından belirtilen bu formatlardan daha ileri seviye görülmektedir. Dirac üreticilerinin, yüksek çözünürlüklü videolar için MPEG-2'den iki kat daha az bit oranı ile üst seviye video içeriği elde edilebileceği iddiaları da bunu destekleyici niteliktedir [2]. Bu anlamda aslında Dirac video kodlayıcının temel işleyişini iki bölüme ayırabiliriz: **1)** Bit dizini için bir sıkıştırma spesifikasyonu tanımlanması ve bunun kod çözücüye yollanması ve kod çözücünün ele alınışı. **2)** Belirtilen sıkıştırma operasyonu ve tersi işleminin gerçekleştirilmesi için bir yazılım dizini tanımlanması.

Kodlama işlemi her ne kadar karmaşık bir işlem olarak tanımlansa da diğer video kodlayıcılara kıyasla Dirac için kodlama ve kod çözme zamanının genel anlamda oldukça kısa olduğu söylenilebilir. Ayrıca sonuçlar analiz edilebilirlik açısından yazılımsal pek çok parametre ile desteklemiş olduğundan işler daha da kolaylaşmış durumdadır [2].

Bu tezde, Dirac video kodlayıcının performansı, QCIF, CIF ve 4CIF giriş video dizileri için farklı dalgacık tabanlarında analiz edildi ve sonuçlar karşılaştırıldı. Ayrıca Dirac video çözücünün dalgacık tabanlı bir kodlayıcı içermesi itibariyle kullanılan dalgacık tabanlarının ekstra bir öneme sahip olduğunun bilinciyle transformasyon sırasında yararlanılan dalgacık fonksiyonlarının ‘lifting’ işlemine etkisi ve kalite açısından çözücüye getirdikleri incelenmeye çalışıldı.

1.1 Tezin Amacı

Bu tez, Dirac video kodlayıcı performansını farklı dalgacık tabanlarında analiz etmeyi ve hazırda varolan ve Dirac'ın kendi veritabanındaki dalgacık dağarcığında yer alan setlere yeni bir dalgacık daha ekleyerek performansı yukarı çekmeyi amaçlamaktadır. Ayrıca farklı hareket vektörü karaklılıklarındaki analizleri de sunarak uzaysal domendeki kodlanacak olan veriden sağlanacak olan kazancın nasıl değiştiğine yönelik önemli çıkarımlar yapılmasına da olanak sağlamaktadır. Buradaki analiz en azından çözücüdeki çeşitlemeyi arttırarak kullanıcıya daha yeni olanaklar temin edebilecek niteliktedir.

1.2 Literatür Özeti

Genel olarak video çözücüler açısından baktığımızda yazılımsal ya da donanımsal pek çok yöntemle iç içe, bir dizin araştırma grubuyla karşı karşıya kalmaktayız. Ancak bunların çoğunluğunun yazılım ağırlıklı olduğu çıkarımını yapmak hiç de zor olmayacaktır zira, video gibi bilgisayar üzerinden akan bir veri yığınıyla ilgilenildiği ve pek çok farklı algoritmik işlemlerle bu diziler üzerinde değişiklikler yapmanın mümkün olduğu göz önüne alındığında söz konusu durumun odak noktasını çıkarabilmekteyiz. Ele alınan problem temelde bir yazılım problemidir. Dolayısıyla video işlemeyi ve çözücü sistematüğini diğer haberleşme konularından ayıran en ince noktalardan biri de budur diyebiliriz.

Genel anlamda video çözücüler sahasına bakarsak, 1990' larda MPEG-2' nin yaptığı atılımdan sonra 1999'da yaklaşık %20'lik bir kod kazancı iyileştirmesiyle 'MPEG-4 Part 2' ismiyle yeniden karşımıza çıkan video sıkıştırma algoritmalarının bu kolu temelde ayrık kosinüs dönüşümüne dayalı olmakta birlikte hareket kompanzasyonu ve kestrimindeki efektif yaklaşımı ile dikkat çekici olarak görülebilirdi. 2000'li yıllarda 'H.264/MPEG-4 Part 10' çok daha verimli bir organizasyonla karşımıza çıkarken özellikle düşük bit oranlarında bile kayda değer seviyede iyi derecede PSNR oranları içerebilmesi dikkatlerin birden bu algoritmanın üzerinde toplanmasına sebep olmuştu [2]. Bu arada bahsedilen PSNR denilen hata ölçüm parametresinin görsel kaliteyi ne derece yansıttığı da tartışma konusu olmuştu. Pek çok otorite artık PSNR'ın tahtının sallandığını kabullenmekte ancak bunun yerine hangi yeni parametre yada parametreleri kullanmak gerektiği konusunda görüş birliğine varamamaktaydılar [3]. Bu sebepten ötürü, bu dönemden itibaren ortaya çıkan yeni video sıkıştırma algoritmalarında hata ölçüm parametresi olarak, çözücülerin içlerinde özel olarak tanımlanmış pek çok yeni sistematığe rastlamak mümkündü. Ve son olarak, 2004'te kullanıcıyla buluşan Dirac algoritmasının ise zaten PSNR olarak MPEG-2 ve 4'ten genel anlamda üstün olduğu önemli ölçüde gösterilmişti. Ayrıca 2005'te, hantal diyebileceğimiz aritmetik kodlayıcı kısmının da önerilen yeni bir sistemle hızlandırılması oldukça basit bir sistemle video sıkıştırmada neler yapılabileceğini gösterir nitelikteydi [4]. Tüm bunların dışında herhangi bir lisans ödemesi gerektirmemesi ve hata metriği olarak da SSIM ve kendine özel metrik değerlendirmesine sahip oluşu bu yapıyı benzersiz kılıyordu.

Dirac çözücüsü hakkında ortaya çıkan gelişmeleri birçok noktada tasnif etmek mümkündür. İyileştirmeler çözücünün ilgilenilen bloklarına bağlı olup kimi çalışmalar video çözücüsünün sadece belirli bir bloğuna dairken kimileri ise tüm çözücü modelini içerebilen çalışmalardır [1,2]. İşin en başında Dirac'ın ortaya çıkışından başlamak gerekirse 2004 yılına gitmek gerekir. Yeni yüzyılla birlikte internet uygulamalarının video basamağının gerek geliştirilen yeni algoritmalarla gerekse bu yöndeki kullanıcı tarafındaki talebin gözle görülür şekilde artmasıyla, o ana kadar mevcut olan sıkıştırma metodlarının yetersizliği bu konuda çalışan uzmanların dikkatini çekmiş ve yepyeni standartlar türemiştir. Fakat bu standartların büyük bir çoğunluğu lisans ödemeleri ve uygulanan teknolojinin kullanıcıya açıklığı noktasında zorluklar içeriyordu. Özellikle lisans ödemelerinin fazlalığı bireysel

kullanıcıların bu algoritmalarla tanışmasını sıkıntılı bir hale getiriyordu. 2004' te Tim Borer ve ekibinin BBC yayıncılık şirketiyle yaptıkları ortak bir tasarım olarak karşımıza çıkan Dirac tüm bu sıkıntıları ortadan kaldırıcı nitelikteydi. Ortaya çıkan bu yeni yapıda dalgacık transformunun etkin bir şekilde kullanılması ve genel örgü olarak diğer algoritmalarla kıyaslandığında oldukça basitlik içermesi önemli avantajlar olarak görünmekteydi. En kritik artısı ise, herhangi bir lisans ödemesi gerektirmemesiydi. Kullanıcı istediği gibi kodları indirip kullanabilmekte böylece mevcut yazılımı geliştirmeye yönelik -daha önceden pek denenmemiş bir şekilde- kullanıcıdan da gelen desteklere açık olduğu belirtilmiş oluyordu [1,2,3].

2005 yılında Dirac'ın sahip olduğu aritmetik kodlayıcının verimliliğini arttırmak üzere H.264 yapısına benzer daha farklı bir yapı önerilmiş ancak kodlama zamanı ve PSNR performansı açısından halen istenilen seviyeye gelinememişti. Diğer yandan Dirac'ın değişik açılardan performansını analiz etmeye yönelik pek çok çalışmada da önemli veriler elde edilmişti. Bunların kimileri bu kodlayıcıyı diğer temel ve üst seviye çözücülerle karşılaştırmaya yönelik iken bazıları da Dirac performansını çözücünün kendi içinde değerlendiriyor, verimliliği arttırabilmek için çeşitli öneriler sunuyorlardı. Birtakım makalelerde, Dirac performansının benzerlik indeks metriğine (SSIM) göre değerlendirilmesinin daha olumlu olacağı önerilirken bazı bilimsel yayınlarda da Dirac'ın kendi iç hata metrik parametresinin kalite kıyaslaması için daha uygun bir kıstas olacağı vurgulanmıştır. Bu noktada, her ne kadar Dirac, performans ayarlaması bakımından kendine özgü bir kalite metriği tanımlasa da, diğer çözücülerle Dirac'ın kıyası yapılmak istendiği zaman PSNR, SSIM ve bu gibi diğer yaygın metrikleri kullanmanın şu an için daha mantıklı olacağı açıktır. Çünkü halen pek çok otorite için bu parametreler çözücülerini kıyaslamada önemli mihenk taşları hükmündedirler. Yine genel başarımlar karşılaştırması açısından bakılırsa, bu konuda Dirac'daki her bir dalgacığın bilgisayar mikroişlemcisini ne kadar yorduğu hakkında da önemli çalışmalar ortaya konmuştur. Bu tür çalışmalar işlem kompleksliğini sergilemek açısından önem arz edip kıyas yapmaktan daha çok iç performansı arttırmaya yönelik çıkarımlar sunmaktadır [2,4].

Dirac video çözücüsünü, hareket kompanzasyonu ve eşleştirmesi yönünden inceleyen çalışmalar da olmuştur. Bu tür çalışmalarda Dirac'ın kullandığı 1/8' e kadar çıkabilen kararlılıktaki hareket vektörlü hareket eşleştirmesi teknikleri incelenmiş, hibrid ve hiyerarşik düzendeki kestirim organizasyonuna sahip Dirac iç blokları

açıklanmıştır. Bunun yanında örtüşen bloklarla hareket kestrimi yapıldığı için özellikle çerçevelerin içlerindeki geçiş noktalarında ve objelerin sınırlarında ('boundry' ve 'edge') daha iyi sonuçlar elde edildiğini ve Dirac'ın bu noktada diğer algoritmalarından belirgin şekilde öne çıktığını vurgulayan makaleler de mevcuttur.

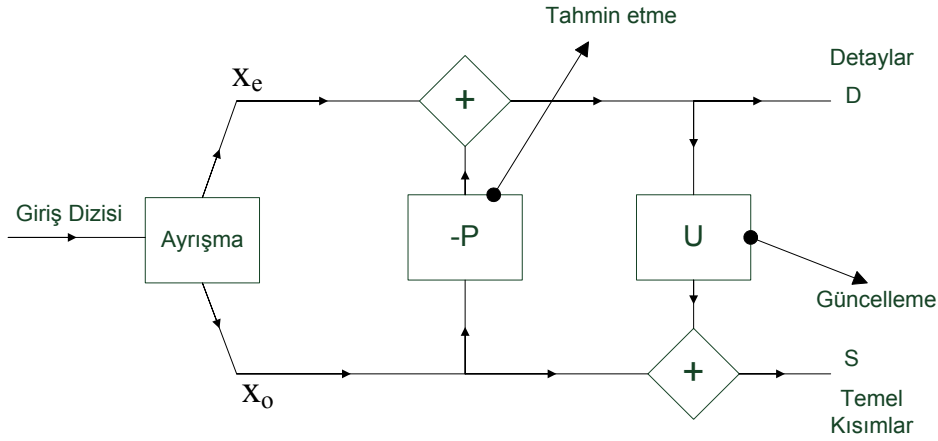
Dirac'ın kullandığı ikilileştirme (binarization) sistemi de araştırmalara konu olmuştur. Zira, Dirac burada serpiştirilmiş Golomb kodu kullanmaktadır. Söz konusu araştırmalarda da geçtiği üzere bu yöntem klasik Golomb koduna göre avantajlıdır çünkü bu yöntem kod çözücü tarafında işlem yükünü azaltarak kod çözme işlemini kısaltmaktadır.

Yine önemli bir husus olarak Dirac Pro serisinden de bahsedilmesi gerekir. Bu serinin en önemli yanı katı uygulama (hardware) olarak ilk defa video çözücü setlerinin gerçekleştirilmesidir. Bu, Dirac için oldukça önem taşımaktadır. 2008 'deki Beijing olimpiyatlarında ilk ciddi sınavını başarıyla veren Dirac video çözücü için kullanıcıyla buluşma adına bu çalışma farklı bir anlam taşımaktadır. Dirac Pro serisinde VLC ve aritmetik kodlama içiçe uygulanmakta ve kodlama kaybı bu yolla azaltılmaya çalışılmaktadır. Dirac Pro 270, bu yöndeki en son atılım olarak dikkat çekerken bu kodlayıcıyı DP400 serisi olarak kit şeklinde satın alabilmek BBC'nin önderliğinde mümkün olabilmektedir [2].

Esasen burada inceleyeceğimiz konu dalgacıklar tabanında olduğu için, Dirac uygulamalarını bu ekseninde incelemek ayrı bir önem taşımaktadır. Literatürde, Dirac'ın dalgacık kullandığı ve bu yönden diğer belli başlı video sıkıştırma algoritmalarından farklı özellikler taşıdığı pek çok kez vurgulansa da bu temelde, Dirac için çok da fazla inceleme yapıldığı söylenemez. Dirac yapısında kullanılan dalgacık filtrelerinin birbirlerine göre kıyaslamasını yapan ve işlem yükü açısından bizlere fikir veren makaleleri, dalgacıklar üzerine yazılmış genel çalışmalar içerisinde de görmek mümkündür. Zaten BBC'nin Dirac için seçtiği dalgacık dönüşüm sistematığının de esasen bu çalışmaların bir ürünü olduğunu söylemek yerinde olacaktır.

Dirac, dalgacık kaldırma (lifting) yöntemini kullandığından dolayı literatürdeki bazı çalışmalarda buna da yer verilmiştir. JPEG 2000'i de kapsayarak anlatılan Dirac kaldırma dönüşümü sistematığı kullanılan dalgacık katsayıları da birlikte verilerek bazı çalışmalarda incelenmiştir. Burada özellikle kaldırma prosedürünün klasik dalgacık dönüşümüne göre üstünlüklerinden bahsedilmiş ve Dirac video çözücünün

tam sayı ya da reel sayı (integer to integer or floating point) tabanında dalgacık dönüşümü niteliklerine sahip olduğunun üzerinde durulmuştur. Tam sayı domenindeki dönüşümün bilgisayar işlemcisini daha az yorduğu sonucuna dayalı olarak uygulamalar geliştirilmiştir [2].



Şekil 1.1 : Dalgacık dönüşümünün kaldırma yöntemiyle gerçekleştirilmesi [5].

Tüm bu söylenenlerden sonra genel olarak belirtmelidir ki çeşitli sebeplerden dolayı ki; bu sebeplerin başında H.264 ve MPEG gibi üst seviye algoritmaların mali durumlarından ötürü Dirac'dan çok daha iyi tanıtılması ve pazarlanması gibi olgular gelmektedir, üzerinde durulan Dirac video sıkıştırma algoritması hakkında çok sayıda ve ileri seviyede akademik çalışmalar yapıldığı pek de söylenemez. Bu noktada çıkarılacak temel sonuç, Dirac gibi son dönemin oldukça popüler dönüşüm metodu olan dalgacık transformunu kullanan, karmaşıklığı az hibrid hareket eşleştirme metoduyla zengin bir uzaysal kodlama kazancına sahip olan ve lisans ödemeleri gibi sıkıntılardan kullanıcıyı kurtaran bir çözünün geliştirilmeye oldukça açık olduğudur. Bu tezdeki temel amaç da her ne sebepten olursa olsun bir şekilde arka planda kalmış olan Dirac algoritmasını yeniden nazarlara sunmaktır ve çıkarılan bu sonuçlardan hareketle video sıkıştırmada yeni olanaklarla kullanıcıyı buluşturmak.

1.3 Ele Alınan Problem

Literatürde yeralan çalışmaların pek çok alanda Dirac adına yetersiz kaldıkları söylenebilir zira, belirtildiği üzere oldukça gelişmeye açık bir sistem üzerine kurulu bu yapıyı değişik açılardan derinlemesine ele almak mümkündür. Buradaki Dirac analizi dalgacık temelinde kapsamlı bir şekilde incelenecek çalışmalar bu yönde gerçekleştirilecektir. Ele alınan problem temelde başında söylenildiği bir yazılım problemidir ama bu teknik bir sorundur ve yazılımın bu noktada bir amaç değil bir araç olduğunu unutmamak gerekir. Dolayısıyla matematiksel çıkarımların yazılıma uyarlanması olan video işleminin kritik bir kolu olan video çözücüleri ve Dirac'ı bu noktayı da göz önünde bulundurarak değerlendirmek yararlı olacaktır.

Dirac için, dalgacık tabanlarında performans analizi geçmiş çalışmalarda pek fazla yapılmamış olup bu noktada iyice irdelene yapılması mümkündür ve gereklidir. Ayrıca başka bir nokta bu tabanlara farklı kararlılıktaki hareket vektörlerinin etkisini incelemek ve burada kullanıcıyı daha verimli bir kodlamaya nasıl götürmek mümkün olabilir sorusunu cevaplamaktır.

Tezde, literatürde yer almamış bir şekilde, değişik kararlılıktaki hareket vektörleri analizi ile birleşik olarak, farklı dalgacık tabanlarında Dirac performans analizi gerçekleştirilmiş, ayrıca 'bior' gibi farklı özellikteki bir dalgacığı Dirac'da kullanarak mevcut dalgacık veritabanı genişletilmeye çalışılmış, değişik çalışmalar için çok daha farklı dalgacık çeşitleri de önerilmiştir. Bu noktadaki temel hedef ise, mevcut Dirac çözücüsünü genel performans olarak bir adım daha ileriye taşıyabilmektir.

1.4 Çalışmada Yapılanlar

Tez çalışmasının Dirac performans analizinde izlenen yöntem aşağıdaki şekilde özetlenmiştir:

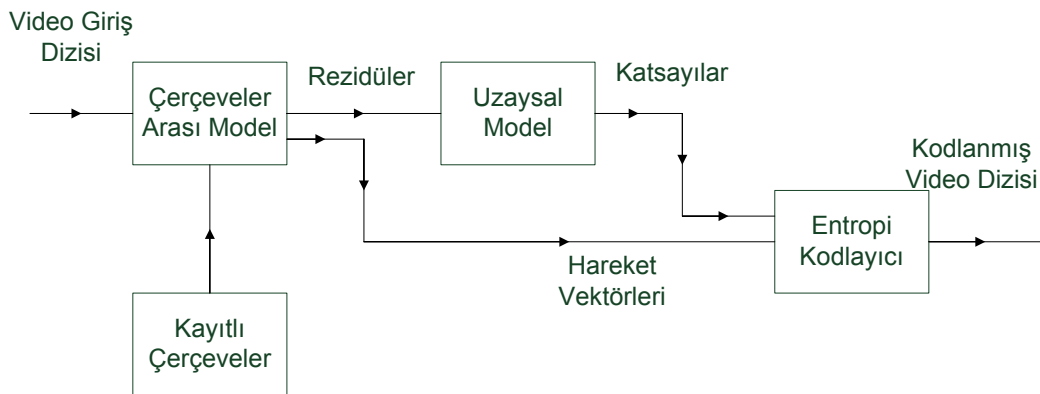
Bölüm 2'de Dirac dışındaki video çözücülere genel anlamda bir bakış yapılmıştır. Bu bölümde, diğer video sıkıştırma algoritmalarının önemli niteliklerine değinilmiş ve geçmiş çalışmalardan da yararlanılarak Dirac ile ilgili başarımların analizi daha da geniş bir profile incelenmesine olanak sağlanmıştır. Bölüm 3'de Dirac algoritması geniş bir şekilde ele alınmış ve bütün yanlarıyla irdelenmiştir. Bölüm 4'de ise bu tezdeki odak nokta olan dalgacık analizinin detaylarına bakılmış bir sonraki bölümde gerçekleştirilecek olan analizlere gerekli dayanaklar oluşturulmuştur. Burada Dirac'da

kullanılmış olan temel dalgacık filtrelerine de detaylı olarak yer verilmiş ve kaldırma yöntemiyle yapılan Dirac dönüşüm sistemiği irdelenmiştir. Bölüm 5'de, yapılan tüm analizler sunulmuştur. Bu kısımda, Dirac hem farklı dalgacık tabanlarında QCIF, CIF ve 4CIF dizileri kullanılarak incelenmiş hem de analizler daha önce de belirtildiği gibi çeşitli kararlılıktaki hareket vektörlerini kullanarak geliştirilmiştir. En son bölümde ise yorumlar ve ileriye dönük yapılabilecek daha farklı çalışmalar önerilmiştir. Ekler kısmında ise YUV video yapısı ve SSIM performans metriği sunulmuştur.

2. ISO VE JVT EKİPLERİNİN SON DÖNEM VIDEO ÇÖZÜCÜ ALGORİTMALARI

2.1 MPEG-2

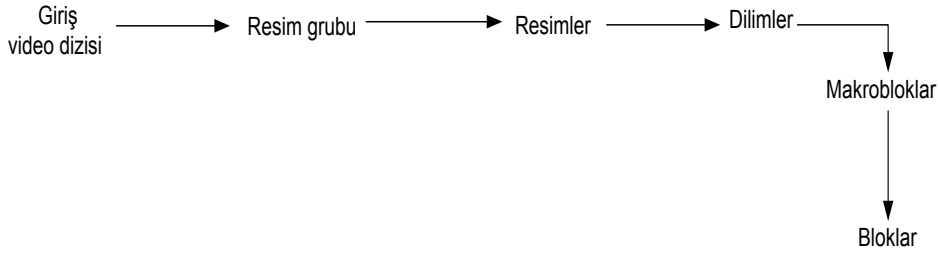
MPEG-1 ortaya çıkarıldığında pek çok açıdan yenilikleri içinde barındıran bir nitelikteydi. Zaten profesyonel olarak ortaya atılan ilk video çözücülerden olması itibariyle bu anlamda daha sonraki gelişmeler adına öncü olabilecek seviyedeydi. Video çözücüler dünyasındaki bu ilk kıpırdanmalar MPEG-1 ile yaşanmıştı ancak böyle olmasına rağmen MPEG-1'in artan ve gelişen haberleşme teknolojileri isteklerine cevaplar özellikte olmadığı açıktı [6]. Bu beklenen bir şeydi zira, bahsedildiği gibi bu çözücü, daha sonra geliştirilebilecek çok daha hızlı ve güvenilir yapıdaki kodlayıcılara bir eşik belirlemek, ilerlemelere ortam hazırlamak amacını da taşıyordu. İşte MPEG-2 bu bakış açısında, MPEG-1 bünyesinde ve onun üzerine inşa edilmiş yeni bir video çözücü olarak görülebilir [7]. Bu bölümde MPEG-2 sistematığı özetle ele alınacak, getirdiği yeniliklerin yanında dezavantajlarından da bahsedilerek Dirac odağındaki bu tez ve ileri çalışmalar için çıkarılan önemli sonuçlar önerilerle birlikte sunulacaktır.



Şekil 2.1 : MPEG-2 genel kodlayıcı yapısı [6].

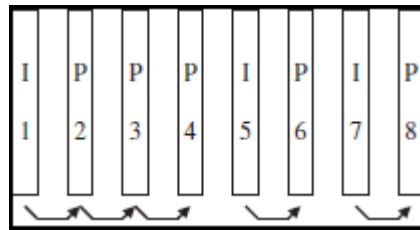
2.1.1 MPEG-2'de temel resim türleri

Öncelikle bahsetmek gerekir ki burada bahsedilecek temel video çözücü kavramlarından bazıları daha sonra anlatılacak H.264 ve Dirac gibi kodlayıcıları da temel teşkil ettiğinden sonra yeniden tekrarlanmayacak sadece genel anlamda ele alınacaktır.



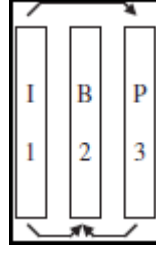
Şekil 2.2 : MPEG-2 video dizisi hiyerarşisi [6].

MPEG-2 de diğer temel video çözücüler gibi video dizisinde akan resimlere belli isimler atar ve bunu çerçevenin öz niteliklerine göre yapar. Bu anlamda belirtilecek ilk resim türü 'I' tipi çerçevedir. Bu içsel çerçeveyi tanımlamaktadır ve karakteristiği gereği en çok bilgiyi içeren çerçeve türüdür zira ondan sonra o dizide akan diğer tüm resimler bu türden kestirilmeye çalışılır. Bilgi olarak bu öz niteliği sağlamayan çerçeveler 'I' tipi olarak atanmaz ve kestirimde kullanılmaz. ('P' tipi seçilenler dışında.)



Şekil 2.3 : MPEG-2 Temel resim türleri-1 [7].

'P' tipi çerçeveler için kestirilmiş resim türleri tanımı yapılabilir ancak bunlar birazdan bahsedilecek olan çift yönlü kestirilmiş türden farklıdır. Zira, bunlar tek yönde kestirilmiş ve kendileri de bir başka resmin kestiriminde rol almışlardır. Çift yönde kestirilen 'B' tipi resimler ise hem 'P', hem de 'I' tipi resimlerden yararlanılarak tahmin edilmişlerdir fakat kendileri herhangi bir kestirim işleminde kullanılmazlar. Bu temel yapı hem MPEG-4 genel yapısıyla hem de Dirac ile son derece benzerlik taşımaktadır. Şekil 2.4'de görülen 'B' tipi çerçevelerin önemli bir özelliği de hem ileri hem geri yönde kestirime olanak vermesidir ki bu özellik bir tek bu tip resim çerçevelerinde bulunmaktadır [6,7,8].



Şekil 2.4 : MPEG-2 Temel resim türleri-2 [7].

Temel olarak resim türleri böyle iken elbette bahsetmek gerekir ki bir de bu çerçevelerin bir de akış sırası vardır. Bu akış sırası son derece önemlidir zira eğer kodlayıcı ya da kod çözücü de herhangi bir sebepten ötürü bu sıra bozulacak olursa video sıkıştırma işleminde ihmal edilemeyecek sorunlar ortaya çıkabilecektir. Burada kod çözücü, belirlenmiş resim dizilerini kod çözücünün sistematığına göre ayarlanmış bir sıra içerisinde gönderir ki bu kod çözücüde bu sıraya göre çözülür. Bu sıra MPEG ailesinin tüm fertlerinde hemen hemen aynıyken Dirac'da resim isimlerinin değişik olmasından dolayı küçük bir farklılıkla yine aynıdır.

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
I	B	B	P	B	B	P	B	B	P	B	B	P	B	B	I

Şekil 2.5 : Kod çözücüye iletmek için bekleyen henüz sıralanmamış resim dizisi [6].

I	P	B	B	P	B	B	P	B	B	P	B	B	I	B	B
1	4	2	3	7	5	6	10	8	9	13	11	12	16	14	15

Şekil 2.6 : Kod çözücüye iletilecek olan ve sıralanmış resim dizisi [6].

Şekil 2.5 ve Şekil 2.6 kod çözücüye iletilmeden önce resim dizilerinin nasıl sıralandığını anlamak için önemli ölçüde fikir verir niteliktedir. Burada bilinmesi gereken ve aslında MPEG-2'nin hızını yavaşlatan önemli dezavantajlardan biri de bu sıralamaya riayet edilmesinin bir zorunluluk olmasıdır. Bu zorunluluk kodlayıcının uygun çerçevelerin ardı sıra gelişini beklemesi demektir. Bu başlı başına bir gecikmeye sebep olduğu gibi kod çözücüde de yine bu sıranın takip edilmesi yükümlülüğü ayrı bir zaman gecikmesine yol açar. Özellikle 'B' tipi resimlerin kod çözücüde çözülebilmesi için öncelikle kod çözücüye 'I' ve 'P' tipi resimlerin buraya ulaşmış olması gerekir. Bu yüzden bu tür çerçevelerin çözümünde önemli gecikmeler

yaşanır ki ileri video çözücülerde (MPEG-4 Part-10 vb.) bu duraksamaların önü alınmaya çalışılmıştır.

Tüm bu bakış açıları altında MPEG-2 kodlayıcısı ve kod çözücüsünün biraz daha detaylarına girmeden temel bazı araçlara değinmekte yarar vardır. Gelecek birkaç bölümde bu temel araçlardan bahsedilerek MPEG-2 algoritmasının detayları özetlenmeye çalışılacaktır.

2.1.1.1 Ayırık kosinüs dönüşümü (DCT)

MPEG-2 video sıkıştırma algoritmasının ve aslında daha sonraki kısımda bahsedilecek olan H.264 algoritmasının en kritik noktalarından biri ayırık kosinüs dönüşümü (discrete cosine transform) 'dür [7,8,9].

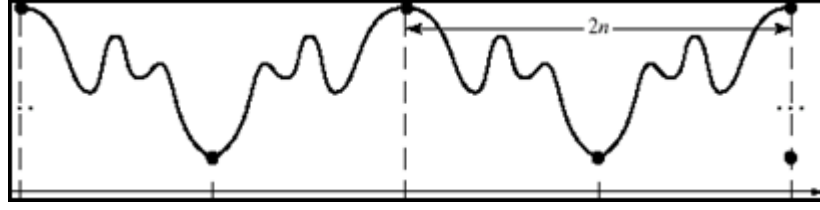
$$F_{vu} = \frac{1}{4} C_v C_u \sum_{y=0}^{N-1} \sum_{x=0}^{N-1} S_{yx} \cos\left(\frac{v\pi(2y+1)}{2N}\right) \cos\left(\frac{u\pi(2x+1)}{2N}\right) \quad (2.1)$$

Denklem 2.1'de [6,10] temel ayırık zamanlı dönüşüm eşitliği görünmektedir. Burada S_{yx} $N \times N$ 'lik bir giriş terimidir ve 'xy' düzleminden 'uv' düzlemine aktarma işlevi söz konusudur. Bu yolla DC ve AC terimler ayrıştırılır ve kodlayıcıya yollanır. C_v ve C_u :

$$C_u = \begin{cases} \frac{1}{\sqrt{2}} & u = 0 \\ 1 & \text{diğer hallerde} \end{cases} \quad (2.2)$$

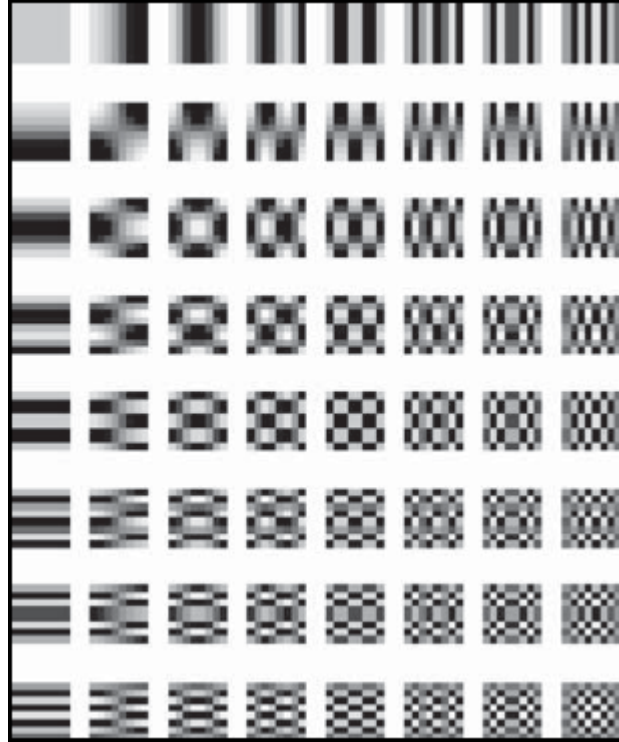
$$C_v = \begin{cases} \frac{1}{\sqrt{2}} & v = 0 \\ 1 & \text{diğer hallerde} \end{cases} \quad (2.3)$$

Ayrık kosinüs dönüşümü orijinal verinin enerjisinin daha az katsayıda toparlanıp sıkıştırılabilirliği açısından üst seviye bir operasyondur. Burada bahsedilmeyecek olan ancak enerji sıkıştırmasında yine bir o kadar üstün olan Karhunen-Loeve dönüşümünün getirdiği kompleksliği indirgeyen bu dönüşüm aslında ayırık fourier dönüşümünden başka bir şey değildir. Sadece DFT'den farkı, veriyi 'n' nokta ile değil verinin kendisi ve ters çevrilmiş olmak üzere toplam '2n' ile periyodik saymasıdır [10]. Şekil 2.7 bunun bir örneğidir. Burada 'n' orijinal verinin nokta sayısıdır. İki boyutlu bir çerçevede bunu DCT uygulanan her veri için x ve y doğrultularında düşünmek mümkündür.



Şekil 2.7 : Ayrık kosinüs dönüşümünde giriş verisinin ' $2n$ ' ile periyodik olması [8,10].

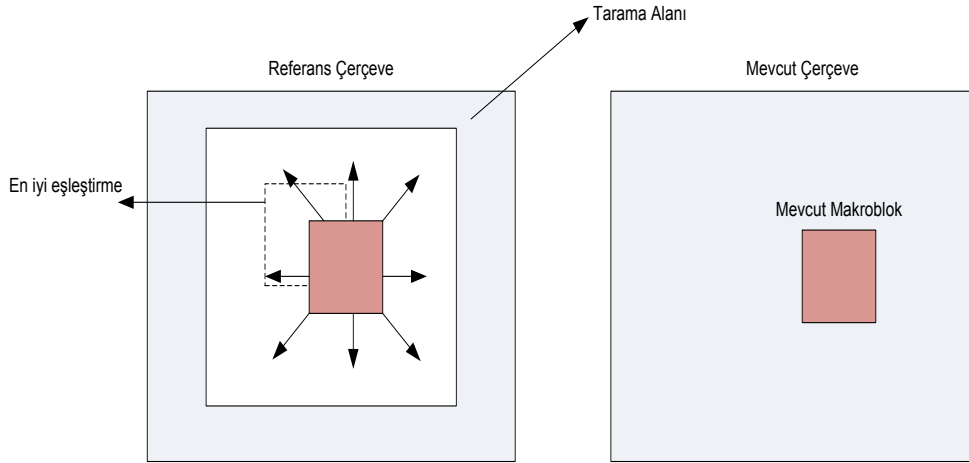
Önemli bir başka bahsedilmesi gereken şey de bu dönüşüm uygulanırken 8×8 blokların kullanılmasıdır. Bu gereksinim kodlama ve kod çözme aşamasında bloklaşma etkisine yol açar. MPEG-1 ve 2'de ve JPEG görüntü kodlayıcıda bu bloklaşma etkisinin önemli bir kalite düşürücü sebep olarak karşımıza çıktığı söylenebilir. Ancak MPEG-4 ailesinde ve özellikle H.264'te bu tür etkiler yumuşatıcı bir takım filtrelerle giderilmeye çalışılmıştır [11]. En son olarak bahsedilecek olan ve asıl konuyu teşkil eden Dirac'da ise bu tür etkiler minimize edilmiştir zira Dirac'da dalgacık dönüşümü gibi çok daha farklı ve yeni bir sistematik kullanımı söz konusudur.



Şekil 2.8 : İki boyutta DCT taban vektörleri [6].

2.1.1.2 Hareket kestirimi ve kompanzasyonu

MPEG-2 sisteminin blok kodlama yapısının çekirdeği verildiğine göre ileri analiz açısından bir az detaya inmekte yarar olacaktır. Öncelikle temel yapı DCT üzerine kuruludur. Ancak dönüşüme girip kuantalanan sadece asıl görüntü bileşenleri değildir. Bu noktada çok farklı ama bir o kadar da derin bir sekme dikkat çeker ki bu adım hareket kestirimi ve eşleştirmesidir. Eğer kodlanan bilgi sadece video çerçevelerinden oluşsaydı kod kazancımız inanılmaz kötü olurdu. Bunun yerine gelen video dizileri arasındaki geçişler esnasında sadece değişen kısmın kodlanması ve sabit olan yani katı (still) formdaki bölgelerin yollanmaması söz konusudur. Bu adımı gerçeklemek için Şekil 2.3 [3,6] incelenmelidir. Söz konusu adımda çerçeveler arasındaki hareket vektörleri belirlenir ve belirlenen hareket vektörlerine göre eşleştirme yapılır ve bu adım duruma göre birden çok kez yapılarak çok daha iyi kestirimler elde etmek mümkündür. Belirtilen hareket kestirimi mekanizması kullanıcının isteği ve kullanma amacına göre farklılık gösterebilir (4x4, 8x8 vb.) .



Şekil 2.9 : Genel hareket kestirimi mekanizması [6].

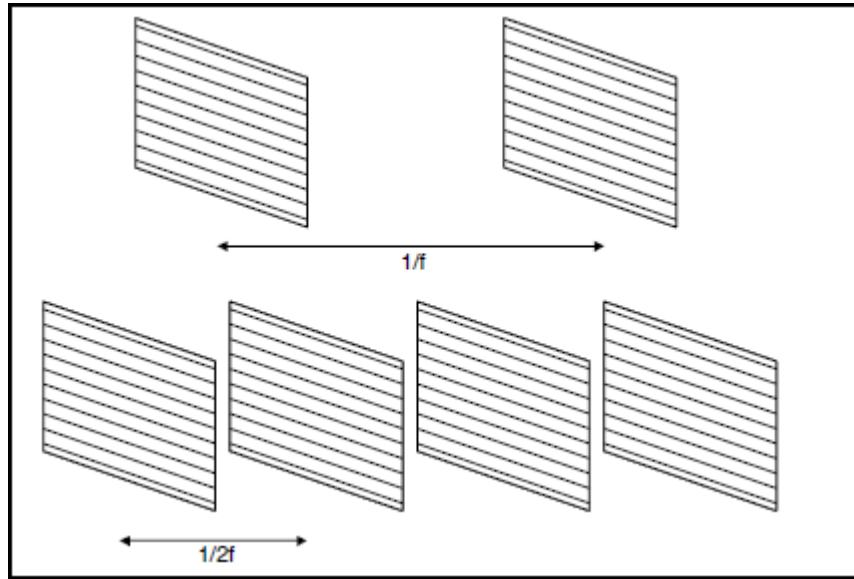
Bu noktaya kadar anlatılanlar bölümün başında da arz edildiği gibi diğer video kodlayıcılardan çok da farklılık içermeyen kısımlardır. Ancak MPEG-2 mekanizmasının anlaşılabilmesi için öncelikle kullanılan bloklar için temeliyle ele alınmalıdır ki konunun özü ve daha sonra vurgulanacak olan artı ve eksi noktalar anlaşılabilir. MPEG-2'deki hareket kestirimi mekanizmasına anlaşılabilirlik teşkil etmesi açısından Şekil 2.10 [6] güzel bir örnektir. Burada 16x16'lık bloklarla bir hareket kestirimi söz konusudur. Ele alınan blokların küçülmesi halinde işlemsel yükün artacağı açıktır. İşte bu aşamada kestirilen hareket vektörleri hem hareket eşleştirmesi işlemini gerçekleştirmek hem de kodlayıcıya gönderilmek üzere kullanılırlar.



Şekil 2.10 : MPEG-2 hareket kestirimi örneği [6].

2.1.1.3 Geçmeli ve geçmesiz video dizisi modelleri

Video kodlayıcının genel olarak detaylarına girmeden önce bahsedilecek önemli unsurlardan biri de geçmesiz (progressive) ve iç içe geçmeli (interlaced) video dizilerinin varlığıdır.



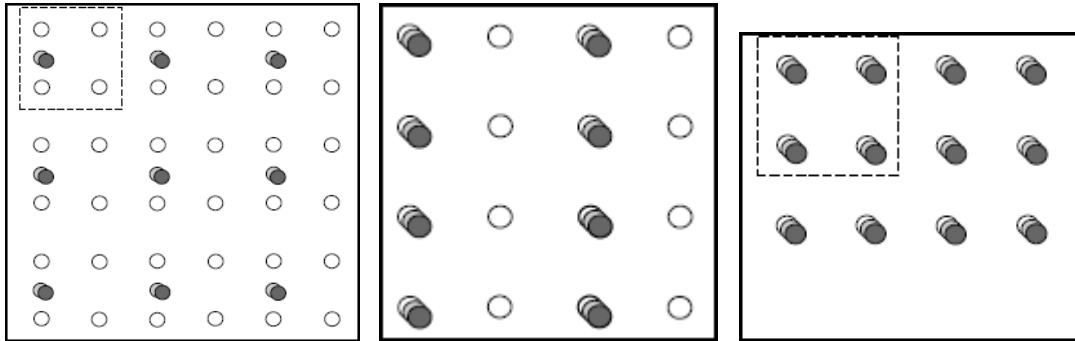
Şekil 2.11 : Geçmeli ve geçmesiz video dizisi modelleri (Üstteki dizi geçmesiz model, alttaki dizi ise geçmeli modeldir.) [6].

Analog televizyon teknolojisinin gelişimiyle sayısal teknolojiye doğru hızla ilerlenirken video dizilerinin görüntülenmesinde halen önemli bir sorun bulunmaktaydı. Çerçeve hızlarının yeteri kadar yüksek olmamasından kaynaklanan parlama etkisi (flicker effect) gerek görsel gerekse teknik anlamda işte bu sorunun

temelini teşkil ediyordu [6]. Bu yıllarda ortaya atılan geçmeli dizilerin kullanılması fikri aslında yeni bir çığır açma demek değildi ama gerekli bant genişliğini arttırmadan çerçeve hızını Şekil 2.11'de görüldüğü üzere arttırmak ve bahsedilen parlama etkisini büyük ölçüde sınırlamak demekti. Bu teknoloji yapısı gereği katot ışın tüplü TV'ler için tasarlanmasıyla birlikte, ilave bir takım işlemlerle günümüz LCD ya da plazmalarına da uygulanabilmektedir. Aslında aşamalı akan dizilere uyumlu tasarlanan ve parlama etkisinin büyük ölçüde aşılmış olduğu son teknoloji görüntüleme araçlarında bile bu yöntemin uygulanması mümkündür ve avantajlar sağlayabilmektedir.

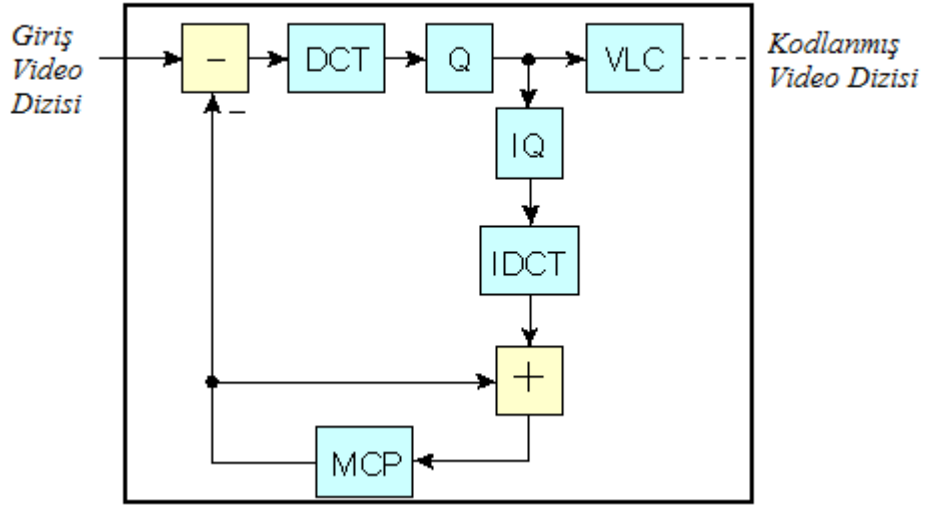
2.1.1.4 YUV video dizileri

Son olarak YUV serilerinden kısaca bahsetmek konunun akışı itibariyle önem arz etmektedir. Şekil 2.12'de video kodlanması açısından çerçevelerin örneklenme tarzları görülmektedir. Özellikle 4:2:0 yöntemi YUV videoları açısından en çok kullanılan tarzdır. Zira Y parlaklık (luma) bileşeni gözün daha hassas olduğu kısım olduğundan bu parçadan daha fazla örnek alıp diğer renk farklarının bilgisini içeren bileşenleri daha az örneklemek yayın sırasında bant genişliğini verimli kullanmak açısından önemlidir. Bu konu ek bölümde ele alınacağından burada bu kadarla yetinilecektir.



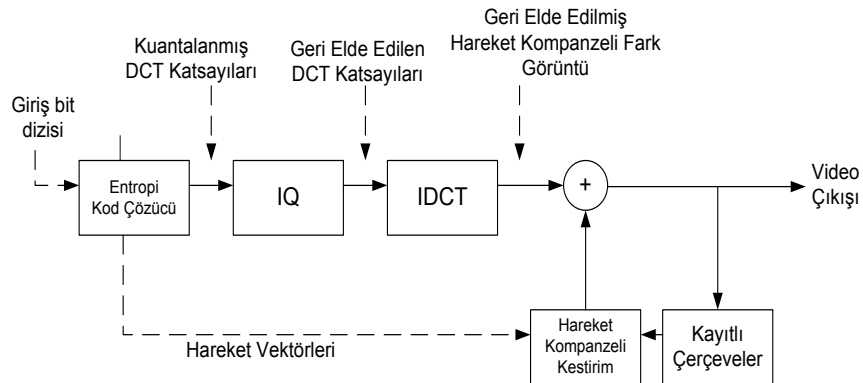
Şekil 2.12 : YUV video örneklemme biçimleri (Soldan sağa sırasıyla 4:2:0 4:2:2 4:4:4 ve beyaz-Y açık gri-U koyu gri-V) [6].

2.1.2 MPEG-2'de temel işlem blokları



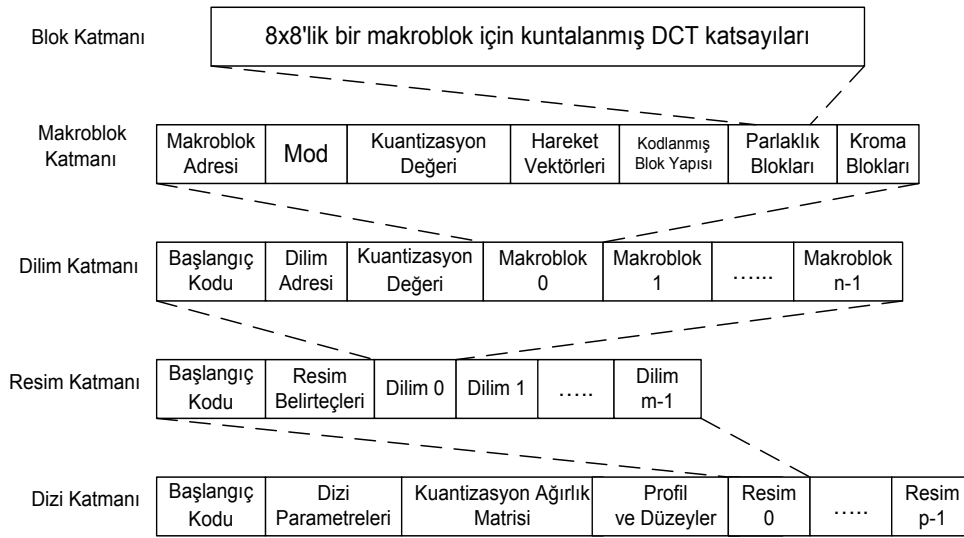
Şekil 2.13 : MPEG-2 kodlayıcı bölümü [8,9].

Şekil 2.13'de giriş video dizisi üzerinde normal dönüşümsel kodlamanın yanı sıra önceki kısımlarda da bahsedildiği üzere hareket tahmini ve kompanzasyonu gibi uzaysal domendeki işlemlerle kodlama kazancının arttırılmaya çalışıldığı görülmektedir [10,11]. Hareket kestirimi için hem çerçeve içi hem de çerçeveler arası yöntemler kullanılmakla birlikte çerçeveler arası kestirim resimler grubu içerisinde daha yaygındır. Böylece kodlanan referans çerçeveler yanında fark görüntüler de kodlanarak daha sonra gelen çerçevelerin tümünü karşıya yollamak yerine bu farkların gönderilmesi yoluyla önemli kod kazançları elde etmek mümkün olmaktadır.



Şekil 2.14 : MPEG-2 kod çözücü bölümü [6].

Kod çözücü kısmında ise kodlanan referans görüntülerine ek olarak fark kısımlar da çözülür ve bu referans dizilere eklenerek girişteki orijinal diziye minimum hatayla ulaşılmaya çalışılır. Burada diğer bahsedilecek bir husus da hareket vektörlerinin kod çözücüye yollanmasıdır. Hareket kompanzasyonu olarak bahsedilen kodlama kazancı sağlayan aşama hareket vektörlerinin belirlenmesi ve bunların elde edildikleri koordinat bilgileriyle kod çözücüye yollanması sayesinde gerçekleşir. Koordinat bilgisi de yollanılmalıdır zira kod çözücü hangi hareket vektörünün görüntünün hangi bölgesinden elde edildiğini bilemez ve bu önemlidir.



Şekil 2.15 : MPEG-2 video bit dizisi akışı içerik modeli [6,11,12].

Önemli bir başka taraf da kestirim mod bilgisinin karşı tarafa iletilmesidir. Kod çözücünün doğru çözümü yapabilmesi için, hareket vektörleri, bunların koordinat bilgileri ve aradaki fark görüntünün çözülmesinin dışında ileri yönde mi, iki yönlü olarak mı yoksa çerçeve içi bir kestirim mi yapılmış olduğu sorusu kod çözücüde cevaplanmalıdır. Tüm bu yardımcı bilgiler ışığında kod çözücüde en doğru video çıkışının elde edilmesine çalışılır.

Şekil 2.15'e bakılacak olursa, MPEG-2'de detaylı bir bit dizisi organizasyonu olduğunu söylemek mümkün olacaktır. Bu şekildeki en alttaki bit dizisinin hangi bilgileri hangi kısımlarında taşıdığı dikkat çekicidir. Bu organizasyon MPEG-1'de de vardır ama bu kadar organize bir şekilde tasarlandığı söylenemez. Burada makro blok tabakasında 4 parlaklık ve 2 kroma bileşeni bilgisi bulunur. 4:2:0 video örnekleme yapısı için ki bu yapı genel olarak en çok tercih edilen bir sistemattir; her biri 'n' adet makro blok içermek üzere toplam 'm' adet dilim tanımlanır [12,13].

Önemli bir başka parametre de kuantalama değerinin ya da matrisinin tanımlanmasıdır. Şekil 2.16'da görülen kuantalama matrisleri bilgileri Şekil 2.15'de görülen kodlanmış bit dizisinde yer alır ve burada soldaki matris çerçeveler arası kestirimde kullanılırken sağdaki ise içsel kestirim için kullanılır.

8	16	19	22	26	27	29	34
16	16	22	24	27	29	34	37
19	22	26	27	29	34	34	38
22	22	26	27	29	34	37	40
22	26	27	29	32	35	40	48
26	27	29	32	35	40	48	58
26	27	29	34	38	46	56	69
27	29	35	38	46	56	69	83

16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16
16	16	16	16	16	16	16	16

Şekil 2.16 : MPEG-2'de kullanılan temel kuantalama matrisleri [6].

2.1.2.1 Giriş-çıkış bit oranı kontrolü

Bu konu hassas bir önem arz etmekle birlikte video çözücünün çalışma mekanizmasına bağlı olarak çok kritik değişiklikler gösterir. Bu noktada temel bazı tanımları yaparak MPEG-2'nin bit oranı kontrolüne eğilmekte yarar olacaktır.

Öncelikle, değişken bit oranında çalışmak çoğu kez tercih edilse de giriş veri dizisine ve yapılan kestirim ve alt bir takım işlemlerin karmaşıklığına bağlı olarak çıkış bit oranı yukarılara tırmanabilir ki bu da çok arzulanan bir durum değildir. Bu yüzden bazı durumlarda sabit bit oranlarında çalışmak arzulandır ki bu tarz çalışmalar çalışılan video çözücünün kalitesini ortaya koyma anlamında önemlidir. Çünkü belirtilen bit oranında meydana gelen sapmanın büyüklüğü nispetinde istenilen verim elde edilememiş olur [10,12]. İşte bu belirtilen sabit bit oranlarının kod çözücü tarafından sağlanabilmesi işlemini genel olarak kod çözücülerde tampon bölge denilen kısım yapar. Buraya değişken bit uzunluklarında gelen kodlanmış bilgiler kod çözücüye sabit bit uzunluklarında gönderilir. Aynı tampon kısımdan bir de kod çözücüde vardır ki o da gelen bu sabit uzunluktaki verileri değişken uzunluklarda bit dizilerine çevirir ve çıkış elde edilir. Burada dikkat edilecek önemli bir husus, kodlayıcı ve kod çözücüdeki tampon bölgelerin uzunluklarının aynı olması gerektiğidir [10,12,13].

MPEG-2 için bu yapıya bakarsak geliştirilen bu tampon bölgeler maksimize edilmiş durumdadır. Ancak burayı maksimum uzunlukta tutmak, sıkıştırma işlemini bir o kadar da geciktirecektir. Bu gecikme ise tampon boyutu ile kanal bit oranı bölümünden yaklaşık olarak çıkarılabilir. Örneğin MPEG-2'nin 6 Mbit/s bit hızında çalıştırıldığı varsayılırsa ve böyle bir sistemin 1.8 Mbit'lik bir tampon bölüme sahip olduğu düşünülürse buradaki gecikme (kodlayıcı ve kod çözücünden doğan toplam gecikme kastediliyor.) yaklaşık 300 ms olacaktır [13,14]. Tampon bölümün taşıyabileceği bit kapasitesi düşürülürse gecikme de azalır ancak kod çözme işleminin sonucunda karşımıza çıkan sıkıştırılıp açılmış yeni video dizisi orijinale nazaran kalite olarak oldukça kötüleşecektir ki bu da istenmeyen bir durumdur. İşte giriş-çıkış bit oranı kontrolü bahsedilen bu bilgilerin ışığında tam bir ödünleşme problemidir. İyi video çözücüler bu noktada kaliteli sıkıştırma işlemi gerçekleştirirken bunu da hızlı yapabilen sistemlerdir ve günümüzde bu optimizasyon problemi oldukça iyi derecelerde çözüme kavuşmuş durumdadır denilebilir.

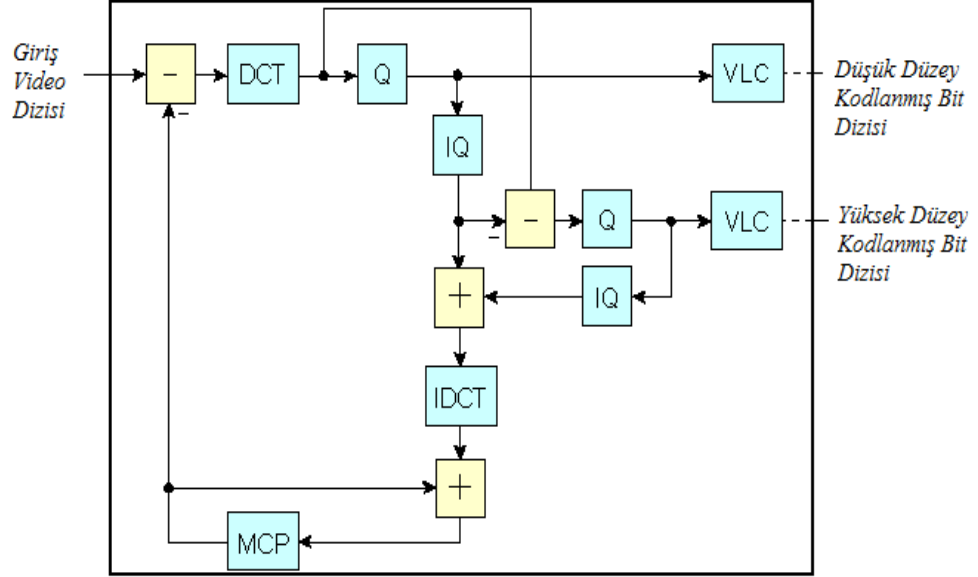
2.1.2.2 Profiller

MPEG-2'de resim boyutu ve bit oranı gibi özel kontrolleri farklı ölçülerde gerçekleyen ve kullanıcıya birbirinden farklı olanaklar sunabilen iki temel profilden bahsetmek mümkündür. Bu bölümde MPEG-2'nin barındırdığı profiller ve profillerdeki farklı seviyeler dikkate alınacak sonuçlar kısmıyla da MPEG-2 bölümüne nihayet verilecektir.

- **Ölçeklenemez profiller:** Temel olarak iki tipi bulunduğu söylenebilir. Öncelikle basit profil denilen kısmı ele alınırsa, bu profilin isminden de anlaşılacağı üzere temel seviyedeki çalışmalar için tasarlanmış olduğu görülür. Küçük tampon bölge genişliğiyle ve 'B' tipi çerçeveleri kullanmamasıyla dikkat çeken bu profil, düşük işlem gecikmesi sağlaması yönünden avantaj sağlar. Video konferans sistemleri gibi yerlerde kullanabilen bu yapının yaklaşık olarak 100 ms'lik sistem gecikmesi sağlamasıyla hız yönünden iyi noktadadır denilebilir [12,14].

Bir diğer bilinen ölçeklenemez profil tipi temel profil denilen kısımdır ki bu yapı 'B' tipi çerçeveleri destekler ve daha çok tercih edilen bir profildir. Ancak tampon bilge genişliğinin daha fazla olmasından dolayı basit profile göre yaklaşık 120 ms'lik ilave bir gecikme getirir [12,14].

- **Ölçeklenebilir profiller:** Bu tip profiller içinde temel anlamda üç tarz yapıdan bahsetmek mümkündür.
İlk olarak SNR profilinden bahsetmek gerekirse, bu yapı DCT katsayılarına daha verimli şekilde sıkıştırma işleminde yer vermek üzere bir takım güdümlerle bloklarına sahiptir.



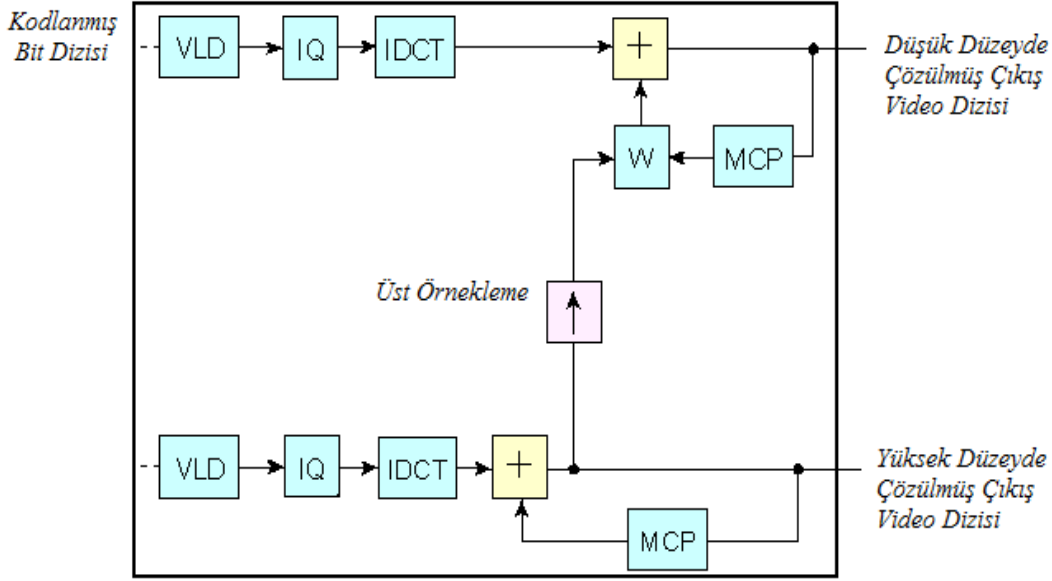
Şekil 2.17 : SNR profil MPEG-2 video kodlayıcı [8,9].

Aslında SNR profili ölçeklenemez tarzdaki profillere benzer tarzda çalışmaktadır. Burada önemli bir farklılık, hareket kompanzasyonu sırasında ortaya çıkan fark görüntünün tekrar kuantalanıp fark görüntünün oluşmasında hatayı azaltıcı şekilde sürecin içine sokulmasıdır. Bu profil karasal televizyon yayını sistemlerinde başarıyla kullanılmıştır [8,9].

Bir diğer profil uzaysal profildir. Bu yapı farklı çözünürlüklerde yapılan kestirimleri karar mekanizması için kullandığından ötürü bu adı almıştır ve ölçeklenebilir olması da aslında bu özelliğinden ötürüdür. Burada eğer uzaysal ölçeklenebilirlik tanımlanacak olursa, bu özelliğin düşük seviyede çözülmüş video dizilerinin yüksek seviyedeki çerçevelerin kestirimlerinden elde edilmesinden kaynaklandığı söylenebilir. Bu noktada dikkat edilecek husus, eğer yüksek seviyede çözülmüş video dizilerinin çözünürlüğü düşük seviyedekinden yüksekse, bu seviyede çözülen dizilerin de yüksek seviyeye uyum sağlaması açısından enterpole edilip örnek sayısı artırılarak çözünürlüğü yükseltilir [12,14]. Bu açıdan konuyu anlayabilmek için Şekil 2.18 incelenmelidir.

Burada 'W' uyarlamalı ağırlıklandırma fonksiyonu 'VLD' ise değişken uzunluklu kod çözücüü sembolize etmektedir. Bu tür sistemler yüksek içerikli TV yayınlarında başarıyla uygulanmaktadır.

Bu anlamda bahsedilecek son profil ise yüksek seviye profildir ve bu kısım daha kompleks çalışmalar için tercih edilir. 4:2:0 örneklemesinin yanında 4:2:2 tarzı çalışmalara da olanak vermektedir.



Şekil 2.18 : Uzaysal profil MPEG-2 video kod çözücü [8,9].

Üstte anlatılan bu profiller MPEG-1 tabanında geliştirilmekle birlikte düşük bit oranlarında sağladığı önemli başarımı sayesinde pek çok imkânları kullanıcılara kazandırmışlardır. MPEG-1 özellikle aşamalı video dizilerine uyumluyken MPEG-2'de iç içe geçirilmiş diziler de kullanılmış ve bu noktada ilerleme kaydedilmiştir [6,14]. Hız noktasında da belirtilen profil yelpazesıyla kullanıcılara çeşitli tercih imkanları sunulurken farklı alanlarda oluşabilecek farklı ihtiyaçlara uygun yeni nitelikler video çözücüye kazandırılmıştır. MPEG-2'nin yine de pek çok eksikleri bulunduğunu tüm bunlarla beraber gözden kaçırmamak gerekir ki bunların başında her şeye rağmen sıkıştırma bit oranları ve hızlarının tatmin edici seviyede olmamasıydı. Gelişen teknolojiyle birlikte yeni ihtiyaçlara cevap verecek daha özgün ve hızlı yapılara ihtiyaç duyulduğu önemli bir gerçektir ve bu yapının temelinde ortaya çıkacak olan H.264 bu noktada MPEG-2'nin pek çok eksikliğini giderici nitelikteydi [13]. Bir sonraki bölümde ele alınacak bu video çözücü, özellikle düşük bit oranlarında bile sağladığı yüksek performansla dikkatleri üzerine çekmiştir ve halen video çözücüler dünyasının pek çok açıdan zirvesini işgal etmiş durumdadır.

2.1.3 MPEG-2'de temel işlem seviyeleri

Son olarak bahsedilmesi gereken önemli bir husus da MPEG-2'deki temel işlem seviyeleridir bu her seviyenin temel özellikleri Çizelge 2.1'de gösterilmiştir. Her seviyenin temel video sıkıştırma parametreleri adına kendi içinde belirli sınırlamaları vardır ve her kullanıcının sunulan bu farklı imkânlardan kendi tercihi ve çalışmaları doğrultusunda farklı şekilde istifadesi mümkündür. Temel yayıncılık anlamında MPEG-2'de standart içerikli TV için temel profil kullanılması en azından gerekirken yüksek içerikli yayınlar için ise yüksek-1440 seviyesi tercih edilir.

Çizelge 2.1 : MPEG-2 temel işlem seviyeleri [8,9].

Seviyeler	Maks. Çerçeve Genişliği (Piksel seviyesinde)	Maks. Çerçeve Yüksekliği (Piksel seviyesinde)	Maks. Çerçeve Hızı (Hz)	Maks. Bit Oranı (Mbits/s)	Tampon Bölge Genişliği (Bits)
Düşük	352	288	30	4	475136
Temel	720	576	30	15	1835008
Yüksek-1440	1440	1152	60	60	7340032
Yüksek	1920	1152	60	80	9781248

2.2 H.264 ve Genel Özellikleri

Gelişmiş bir ürün olan bu standart VCEG ve MPEG video takımlarının bileşkesi olan JVT ekibinin ortaya çıkardığı yeni bir video sıkıştırma algoritmasıdır [1,13]. 2009'da son şeklini alan bu ürünün kullanıcıyla tanışması 2003'ün başlarında olmuştur. Bu yapı üst seviye hata başarımı performansı ve düşük bit oranlarında dahi görsel olarak tatmin edici sonuçlar vermesi açısından MPEG-2, MPEG-4 Part 2 ve H.263'den çok daha ileri seviyede olarak görülebilir. Ayrıca tasarım karmaşıklığı açısından bu sayılan video çözücülerden fazla bir yükü kullanıcıya yüklemeyiz. Geniş video uygulamaları sahasının pek çok noktasında kullanılmaya elverişli olan bu yapı sayısal sinema ve HDTV teknolojisinin vazgeçilmezleri arasında kendine yer bulmuş durumdadır [13]. VLSI teknolojisi için de elverişli olan bu yapı kodlama verimliliği ile gerçekleştirilen işlem yükü arasında önemli bir denge sistematığı kurar. Dönüşümsel kodlamanın kullanıldığı hibrid bir çözücü olan bu yapı giriş video dizinin çerçeveleri arasındaki istatistiksel ilişkiyi azaltmak için de entropi kodlayıcıdan yararlanır.

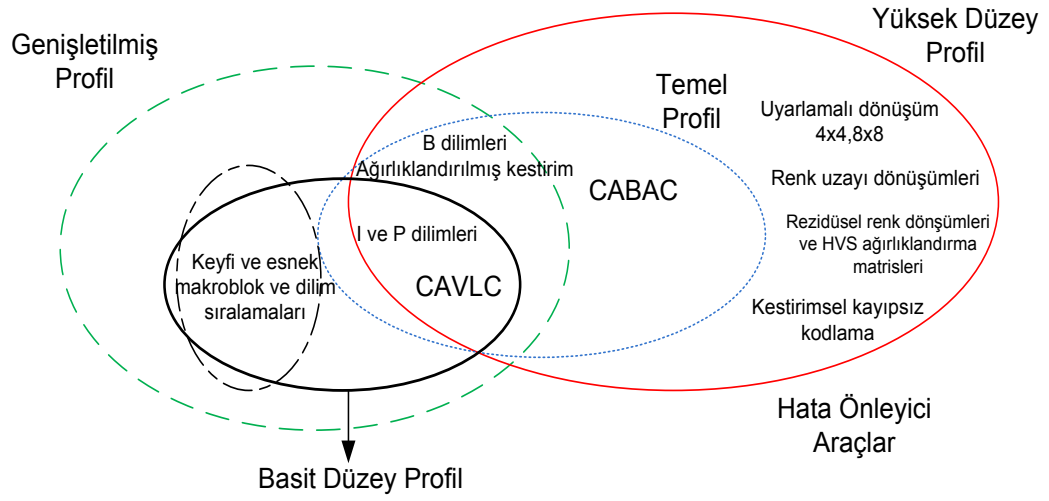
4x4 ve 8x8 tamsayı dönüşümlerin kullanıldığı H.264, aynı zamanda değişken boyutlarda makro blok dizilerinden ve çoklu referanslı kestirim gibi MPEG-2 tabanlı bir takım niteliklerden bloklaşma giderici süzgeçler ve ileri seviye entropi kodlayıcı gibi üst seviye bir takım araçlarla maksimum seviyede faydalanır [1,13]. Sonuç olarak tüm bu özellikleriyle H.264 önemli ölçüde diğer video çözücülerden öne çıkmakta ve uzmanların vazgeçilmezleri arasında kendisine sağlam bir yer edinmiş durumdadır. Ancak tüm bunlara rağmen H.264'ün de önemli dezavantajları vardır ki bunlardan giriş bölümünde kısmen bahsedilmiştir. Şu kadarını söylemek gerekir ki H.264 DCT'ye benzer bir dönüşüm anlayışına dayanan ve her şeye rağmen üst seviye bir kompleksliği bünyesinde barındıran bir kodlayıcıdır. Ayrıca günümüz teknolojisinin ve haber çağının getirdiklerinin dikkate alınmasıyla bu tarz video çözücülerin kullanıcılara ulaşmasının kolaylığından bahsetmek gerekirken ne yazık ki bu tarz geliştirilen video çözücü yapılarının bireysel çalışmalardan ziyade kurumsal bazda kitlelere hitap etmekte olduğu görülmektedir. Zira istenen lisans ödemeleri genele hitap etmekten uzaktır. Bu da dikkate alındığında yeni, basit ve bireysel kullanıcıya da en az kurumlar kadar hitap edebilen ve başarımının yanında hızı da yüksek bir video çözücüye daha ihtiyaç olduğu açıktır. İşte bu noktada ilerleyen kısımlarda önerilen Dirac'ın önemi daha da iyi anlaşılmaktadır.

2.2.1 H.264'te temel profil yapıları

Aslında bir standartlar ailesi olan H.264'te çeşitli spesifik çalışmalara yönelik oluşturulmuş temel profil yapıları bu bölümde ele alınacaktır.

- **Başlangıç seviye profil:** Bu yapı 'B' tipi çerçeveleri desteklemez ve CAVLC desteği üzerine kuruludur. Temel H.264 profilleri arasında en az kodlama verimliliği garantileyen bu profil basit seviyedeki uygulamalar için geliştirilmiştir [13].
- **Temel profil:** Bu profil daha kapsamlı ve genele hitap eden bir tür olarak tanımlanabilir. CAVLC'yi desteklediği gibi CABAC ile çalışmaya da olanak tanır. 'B' tipi çerçeveleri destekler. Üst seviye performans sağlar [6,13].
- **Genişletilmiş profil:** Bilinen klasik çerçeve yapılarının yanı sıra SP ve SI olmak üzere iki yeni dilim türünü de destekler. Başlangıç seviye profilin tüm özelliklerini bünyesinde barındırmakla birlikte artı olarak iç içe girmiş video dizilerine de destek verir. Kompleks yapısıyla dikkat çeken bu profil üstün başarımlar sağlar [13].
- **Yüksek seviye profil:** Daha çok HDTV yayınları için tercih edilen bu profilin başlıca dört türü bulunmaktadır [6,13].

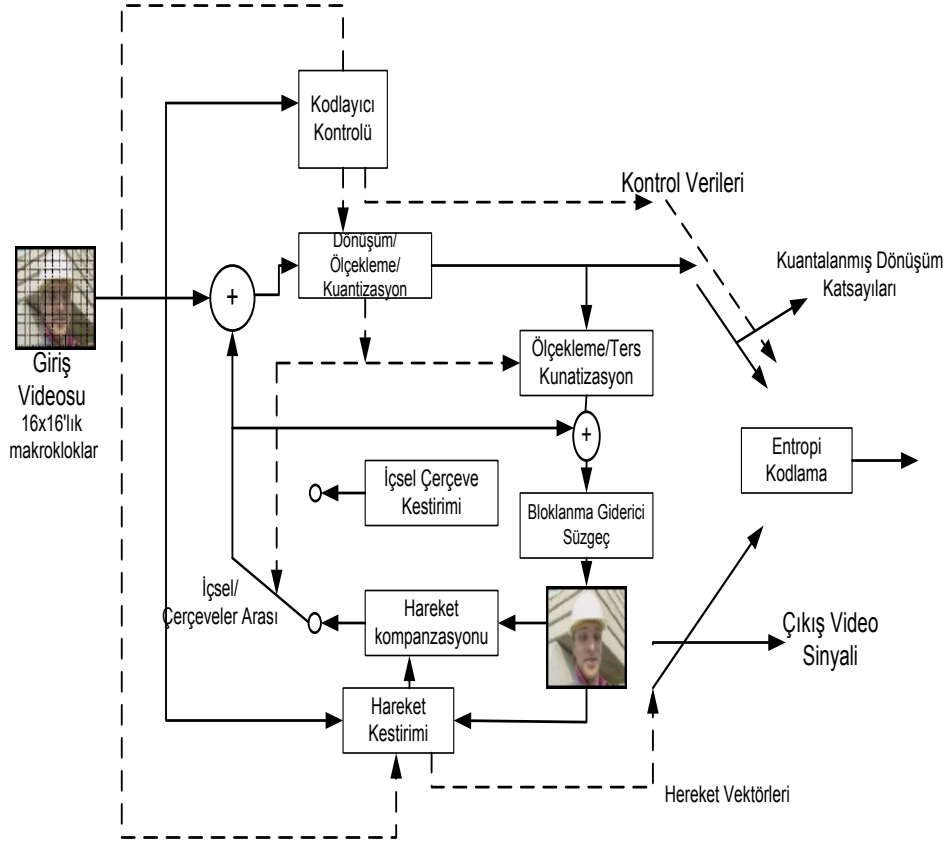
- Standart üst seviye profil:** Normal 8 bitlik video dizileri ve 4:2:0 örnekleme türüne göre dizayn edilmiştir.
- Üst seviye profil 10:** 4:2:0 örnekleme için 10 bits/sample seviyesine kadar çıkabilir.
- Üst seviye 4:2:2 profil:** 4:2:0 örnekleme için 10 bits/sample seviyesine kadar çıkabilir.
- Üst seviye 4:4:4 profil:** 4:4:4 örnekleme için 12 bits/sample seviyesine kadar çıkabilir. Ayrıca tamsayı seviyesinde renk uzayı dönüşümü de yapılabilir.



Şekil 2.19 : H.264'de temel profil yapıları [1,13].

Şekil 2.19'da H.264'deki tüm profil yapılarının ana hatları görülmektedir. Burada dikkat çeken bir diğer husus da CAVLC ve CABAC gibi ileri kodlama tekniklerinin kullanılmasıdır. CAVLC, CABAC'a göre karmaşıklık açısından daha kullanışlı olsa da genel anlamda başarımda CABAC'dan geridedir. Zaten CABAC daha çok üst seviye H.264 profillerinde karşımıza çıkan bir özelliktir.

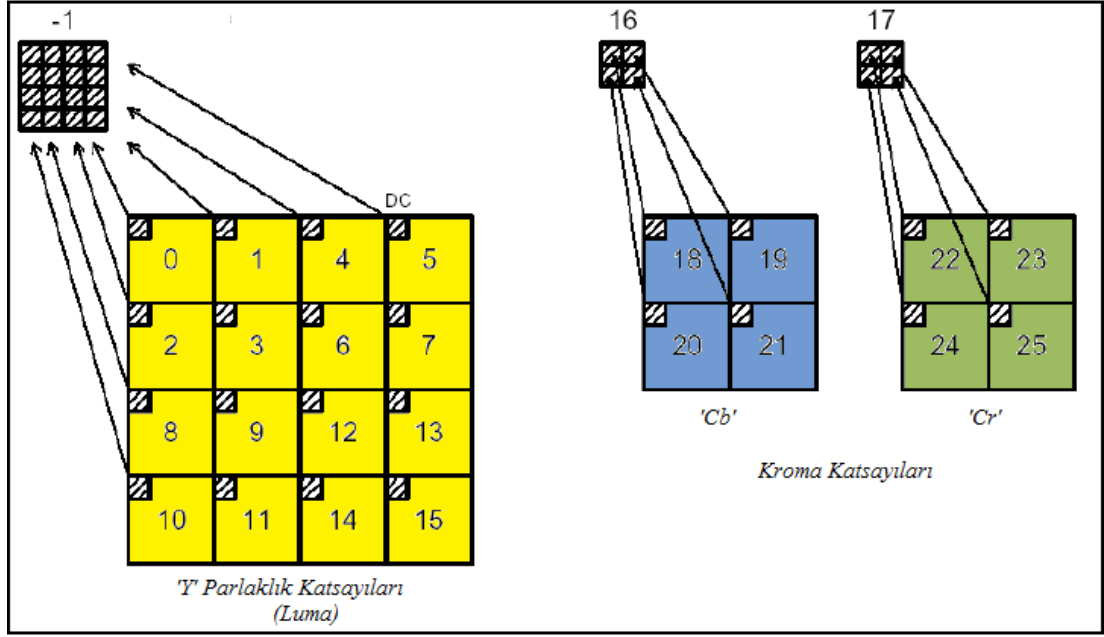
Şekil 2.20'de genel H.264 kodlayıcı yapısı görülmektedir. Burada gerçektende üst seviyede organize edilmiş bir kodlayıcı yapısı gözlemlenmektedir. Genel anlamda MPEG-2 üzerine kurulu bir yapı olduğundan MPEG ailesinin sahip olduğu hareket kestirimi ve kompanzasyonu, DCT anlayışına yakın bir dönüşüm ve kodlama ve zig-zag tarama teknikleri gibi belli başlı nitelikler burada da yerlerini almışlardır. Ancak birazdan ele alınacağı üzere H.264, dönüşüm teknikleri ve bloklaşma etkisini giderici farklı yapılarıyla önemli değişiklikler içermektedir. Ayrıca makro blok boyutlarının çeşitliliği ve çerçeveler arasında sıra değişimi gibi yeni nicelikler H.264 tarafından kullanıcıya sunulmuştur.



Şekil 2.20 : H.264 genel kodlayıcı yapısı [1].

2.2.2 H.264'de temel dönüşüm teknikleri

H.264 genel olarak 4x4 ya da 8x8 makro blok boyutlarında tamsayı dönüşüm odağında işlem temellidir. Daha çok 4x4 makro blok boyutları kullanılır. Bunun dışında H.264'de bir de Hadamard dönüşümü kullanılır ki bu dönüşüm sadece 4x4 tamsayı dönüşüm ile 16x16 çerçeve içi kestirim modu tercih edildiği zaman devreye girer [13,15]. Bu durumda 4x4 boyutlarında tamsayı dönüşümüne uğrayan her bir makro bloğun DC katsayıları birleştirilir ve parlaklık bileşenleri için 4x4'lük, kroma bileşenleri için de 2x2'lik oluşan DC setleri bir de Walsh-Hadamard dönüşümüne uğrarlar ve böylece katsayılar arasındaki uzaysal ilişki minimize edilmeye çalışılır [1,10,13].



Şekil 2.21 : H.264'de çerçeve içi kestirim modunda dönüşüm organizasyonu [1].

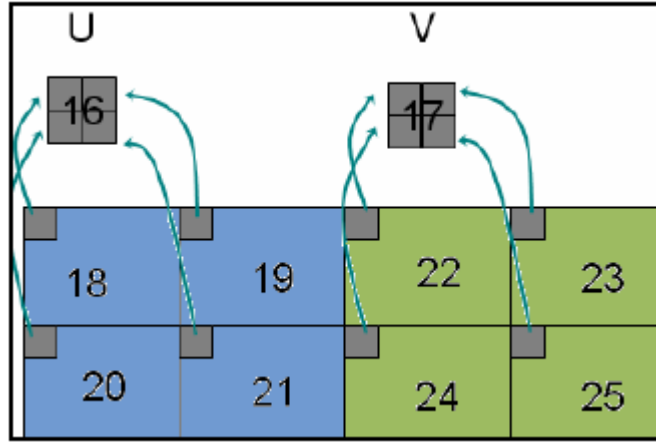
$$Y = AXA^T = \begin{bmatrix} a & a & a & a \\ b & c & -c & -b \\ a & -a & a & a \\ c & -b & b & -c \end{bmatrix} \begin{bmatrix} x_{00} & x_{01} & x_{02} & x_{03} \\ x_{10} & x_{11} & x_{12} & x_{13} \\ x_{20} & x_{21} & x_{22} & x_{23} \\ x_{30} & x_{31} & x_{32} & x_{33} \end{bmatrix} \begin{bmatrix} a & b & a & c \\ a & c & -a & -b \\ a & -c & -a & b \\ a & -b & a & c \end{bmatrix} \quad (2.4)$$

$$a = \frac{1}{2}, \quad b = \sqrt{\frac{1}{2}} \cos\left(\frac{\pi}{8}\right), \quad c = \sqrt{\frac{1}{2}} \cos\left(\frac{3\pi}{8}\right).$$

Şekil 2.21 ve denklem (2.4)'de H.264'de çerçeve içi kestirim modunda dönüşüm organizasyonunun temelleri gözlenebilmektedir. Burada bu dönüşüm daha farklı şekilde faktörize edilebilir ama en genel hali bu şekildedir ve bu dönüşüm H.264 için özel tanımlanmış, DCT'ye benzeyen fakat farklı bir dönüşüm türüdür. Bu noktada daha fazla ayrıntı için [1] ve [16] incelenmelidir. Bu noktada bahsedilmesi gereken bir diğer dönüşüm olan Hadamard dönüşümü için kullanılan genel yapı, denklem (2.5)'de verilmiştir.

$$Y_D = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} \begin{bmatrix} x_{D00} & x_{D01} & x_{D02} & x_{D03} \\ x_{D10} & x_{D11} & x_{D12} & x_{D13} \\ x_{D20} & x_{D21} & x_{D22} & x_{D23} \\ x_{D30} & x_{D31} & x_{D32} & x_{D33} \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix} // 2 \quad (2.5)$$

Burada '/' işlemi matristeki her bir elemanın 2 ile bölümünün sonucunun en yakın tam sayıya yuvarlandığını gösterir. Bu yapı 4:2:0 örnekleme için bu şeklidir ancak 4:2:2 ve 4:4:4 yapılarına da uyarlanabilir ki o takdirde makro blok boyutlarında önemli değişiklikler söz konusu olacaktır.

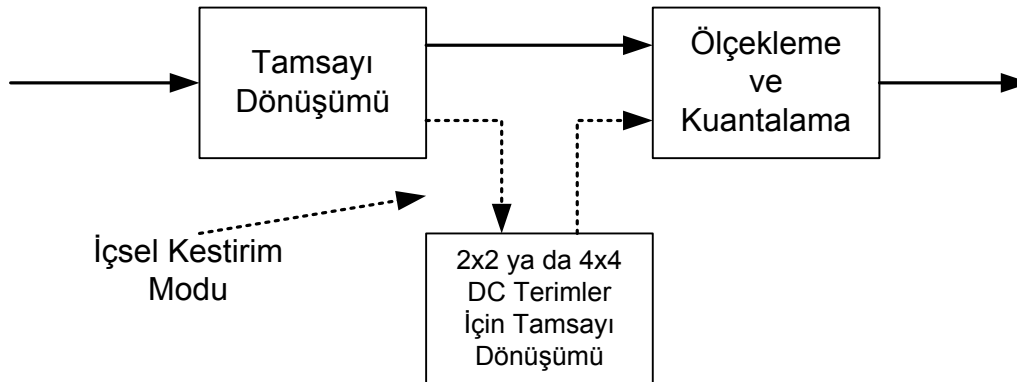


Şekil 2.22 : Kroma katsayılarının çerçeve içi kestirim modunda Hadamard dönüşümüne hazırlanması [1].

2.2.3 Ölçekleme ve kuantalama

Bu nokta bir takım detay prosedürleri sayılmazsa entropi kodlayıcının girişine gelmeden önce giriş dizisinin son geçirdiği işlem olarak tanımlanabilir. Burada ölçekleme faktörü ile kuantalama değeri arasında belirli bir formülizasyon tanımlanabilir. Aslında bu formül kuantalanmış giriş değerinin entropi kodlayıcı girişine nasıl yansıdığını göstermek için önem arz etmektedir. Bu noktada denklem (2.6)'ya bakılabilir. Burada 'I' girişi 'O' ise çıkışı sembolize etmektedir. 'SF' ölçekleme faktörünü, 'Q' ise kuantalama adımını temsil etmektedir.

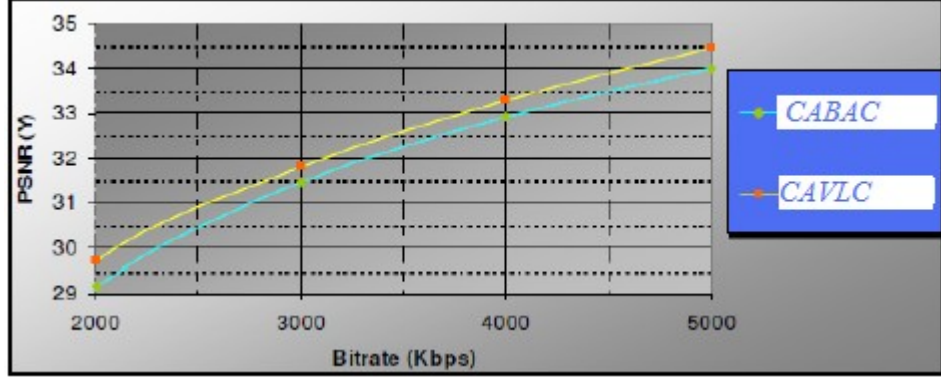
$$O_{ij} = I_{ij} \text{round} \left(\frac{SF_{ij}}{Q_{step}} \right) \quad (2.6)$$



Şekil 2.23 : Genel anlamda H.264 kodlayıcısı içyapısı [1,13].

2.2.4 Entropi kodlama

H.264 temel olarak iki farklı entropi kodlama metodu içermektedir. Bunlardan biri bir ikili aritmetik kodlama metodu olan CABAC diğeri de işlemsel olarak çok daha fazla kolaylık getiren CAVLC'dir.

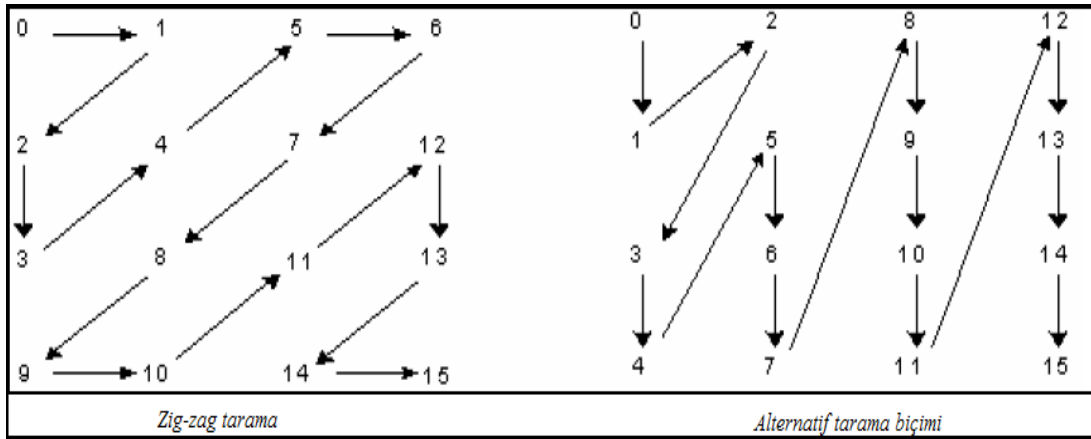


Şekil 2.24 : Örnek cif bir video dizisi için CABAC ve CAVLC performans karşılaştırması ('Mobile.cif' için yapılmıştır.) [1].

İlk olarak CAVLC'den bahsedecek olursak bu yapı değişken uzunluklu kodlamanın bir alt türüdür. Girişine gelen sıfır ve ± 1 katsayılarının toplam sayılarını kodlar. Diğer katsayılar seviyelerine göre kodlanır.

CABAC ise aritmetik kodlama yapısını kodlanır ve CAVLC'ye göre daha karmaşık bir yapıdadır. Girişine gelen katsayıların olasılık giriş dağılımı bu kodlayıcı için gerektiğinden genel de ve H.264'de içerik modelleme araçları ile birlikte kullanılır. H.264'de hareket vektörleri ve kuantalanmış katsayıların sağlıklı ve en kaliteli seviyede kodlanması bu yapı ile sağlanır.

Son olarak bahsedilmesi gereken şeylerden biri de kuantalanmış katsayıların kuantalanması sırasında izlenen metotlardır ki burada hemen zig-zag tarama dikkat çekecektir. Bu yapı MPEG ve JPEG ailesinin değişmez bir parçası haline gelmiştir. Ancak H.264'da alternatif tarama metotları da geliştirilmiştir ki bunlardan biri Şekil 2.25'de verilmektedir. Burada şundan da bahsedilmelidir ki bu katsayılar daha sonra ikilileştirme işlemine sokulmak üzere eksponansiyel-Golomb kodlama bloğuna gönderilirler ki bu konu Dirac bölümünde detaylarıyla ele alınacağından şimdilik bilgilendirme amacıyla bu kadarıyla yetinilecektir.



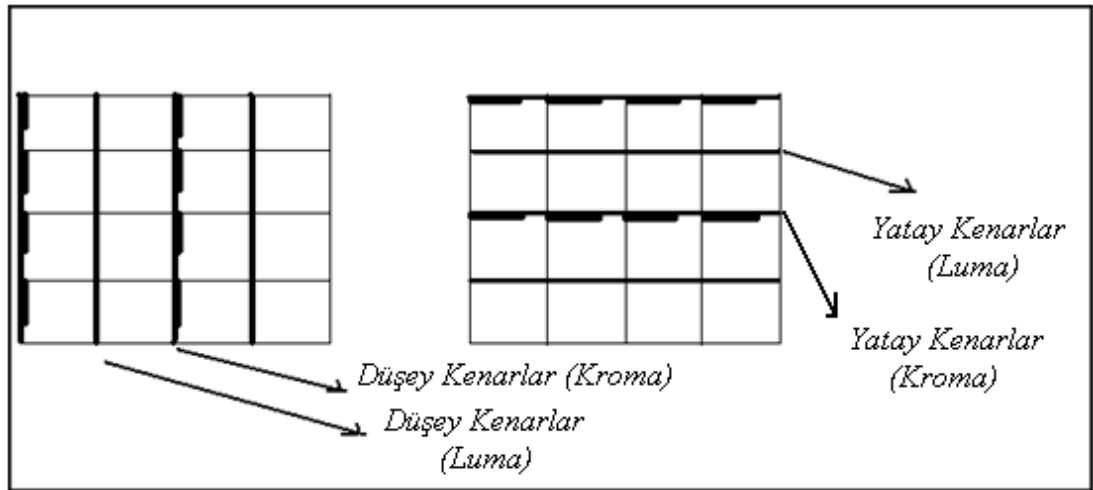
Şekil 2.25 : H.264'de temel rezüdi tarama biçimleri.

2.2.5 Kodlama verimliliği arttıran temel araçlar

Bu bölümde H.264'de kodlama verimliliği sağlayan ve video çözücünün çalışma yelpazesini genişleten önemli bazı araçlardan ve bunların temel niteliklerinden bahsedilecektir.

2.2.5.1 Bloklaşma giderici filtreleme

Bu yapı oldukça büyük önem arz etmektedir. Özellikle yapılan DCT'ye benzer işlem ve makro bloklara ayırma prosedürü video dizilerinde bloklaşma etkisi meydana getirdiğinden ötürü bu tür bir filtreleme işlemine gerek duyulmaktadır. Zira MPEG-2'de böyle bir yapı H.264'deki kadar etkin olmadığından ötürü sıkıştırılmış video dizilerinde bloklaşma etkileri belirginlik kazanmaktaydı ve bu video kalitesini oldukça olumsuz etkilemekteydi. Bu eksiklik H.264'de bu filtreleme sistemi ile önemli ölçüde giderilmiştir [1,13]. Ayrıca video dizisi içinde çerçeveler arası kestirim gibi modlar mevcut olduğundan kestirim yoluyla diğer çerçevelere de yayılacak olan, çerçevelerin birisinde oluşan fazlaca bloklaşma etkisinin ve böylece büyümesi muhtemel olan bir hata yığının önü bu yapıyla alınmıştır. Bu noktada Şekil 2.26 incelenmelidir. Burada önce 16x16'lık makro blok kümesinin düşey kenarları daha sonra da yatay kenarları filtrelendir. Bu filtreleme işlemi sadece resmin genel olarak kenarlarına uygulanmaz [1]. Zaten burada amaç 4x4'lük makro bloklar arasındaki geçişleri yumuşatmaktır. Sınır noktalarında ise eklentileme metoduyla zaten bloklaşma etkisi az olduğundan buralarda düşey ya da yatay olarak filtreleme işlemi daha seyrek tutulmaktadır [1].



Şekil 2.26 : H.264'de bloklaşma etkisini giderici filtreleme işlemi [1].

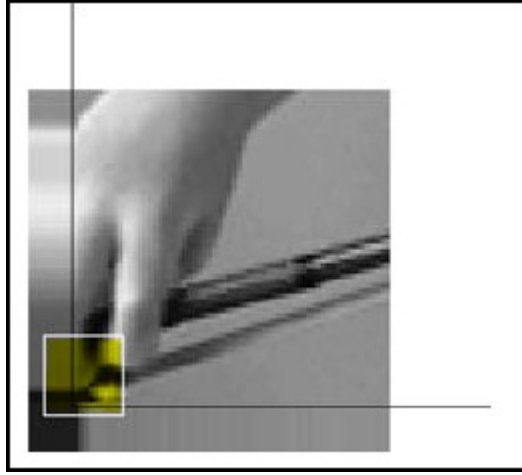
Yine Şekil 2.27 dikkatle incelenirse kalite faktörü yüksek tutulmasına rağmen filtreleme işlemi kullanılmadığında video kalitesinin ne kadar değiştiği hakkında daha sağlıklı bir yorum yapmak mümkün olacaktır.



Şekil 2.27 : Filtreleme işleminin video kalitesine etkisi [12].

2.2.5.2 İçsel kestirim ve makro blok kenarları eklentileme işlemi

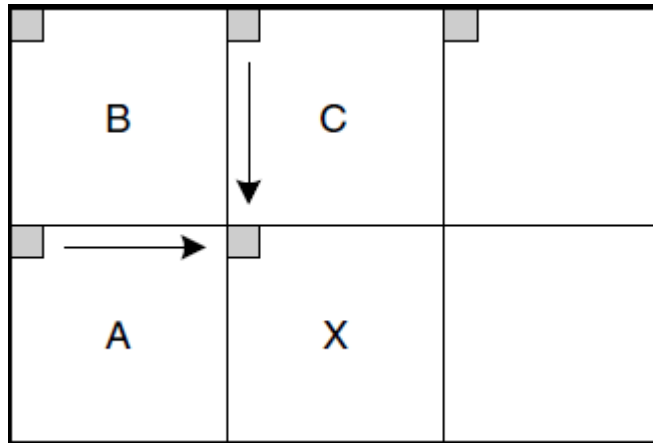
Bir diğer önemli husus hareket eşleştirilmesi sırasında makro blokların tam olarak 16x16'lık ya da 8x8'lik olabilmesi için çerçevelere eklentiler yapılır [6,13]. Bu eklemeler otomatik olarak çözücü tarafından kullanıcıdan bağımsız olarak gerçekleştirilecektir. Burada temel amaç, en iyi hareket eşleştirmesinin tam sınır noktalarında olması durumunda eşleştirmenin ve kestirimin en iyi şekilde yapılabilmesidir.



Şekil 2.28 : Sınır eklenmesi (boundry padding) [6].

Şekil 2.28'de referans obje uzayına makro blokların eşleştirilmesi sırasında yapılan sınır eklenmesi görülmektedir. Bu özellik bilhassa en iyi eşleştirmenin hesaplanması açısından önemlidir.

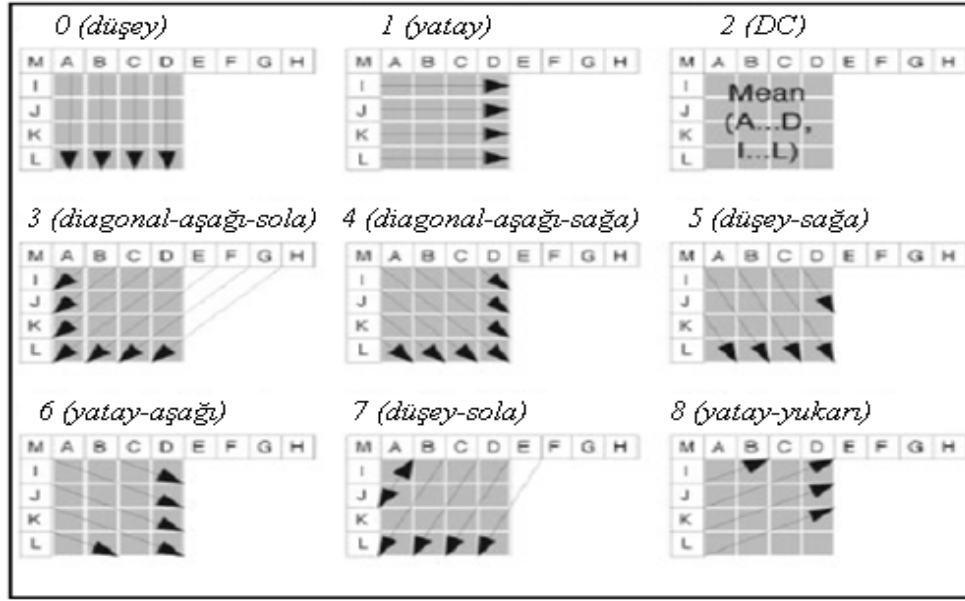
Önemli bir başka husus da çerçeve içi kodlamanın ve hareket kestiriminin varlığıdır. Çerçevelerin pek çoğunda, içsel olarak bir ilişkisinin varlığı söz konusudur. MPEG-4/Part 10 'da da bu özellikten faydalanılır. Ayırık kosinüs dönüşümü sırasında elde edilen DC ve AC katsayıları her bir makro blokta tanımlı olmak üzere başka komşu makro bloklardan kestirilebilirler [6,13]. DC katsayılar yine DC'ler cinsinden AC'ler ise AC katsayılar cinsinden tahmin edilebilir ki bu işlem ekstra bir yük getirmesiyle birlikte kodlama yükünü düşüreceğinden genel olarak avantaj getirmektedir.



Şekil 2.29 : Çerçeve içi kestirim [6].

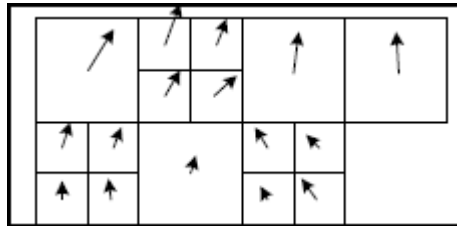
Şekil 2.29'da B ve C makro bloklarının yardımıyla X bloğundaki DC katsayısının kestirimi görülmektedir. Burada aslında küçük bir maliyet hesabı yapılmaktadır ve bu bağlamda önemli bir diğer nokta da 4x4'lük makro bloklar için her bir makro blok içinde kestirim yönünü tayin etmek için yapılan mod seçimidir. Bu seçim,

kullanıcı tarafından da yapılabilir ki üst seviye çalışmalar da bu tercih edilebilir ancak daha çok kodlayıcının kendisi tarafından bit oranı-distorsiyon optimizasyonu sırasında belirlenir. Şekil 2.30'de 4x4'lük makro bloklar için olası içsel kestirim modları sunulmuştur. 4x4'lük bloklar için toplamda 9, 16x16'lık bloklar için ise toplamda 4 adet çerçeve içi kestirim modu bulunmaktadır [13]. 16x16'lık makro bloklar kalite hassasiyetinin çok yüksek olmadığı yerlerde daha az karmaşıklık içerdiğinden dolayı tercih edilebilir. Burada bu modların çeşitliliği çalışma esnekliği açısından H.264 için önemli bir artıdır.



Şekil 2.30 : 4x4'lük makro bloklar için çerçeve içi kestirim modları [1].

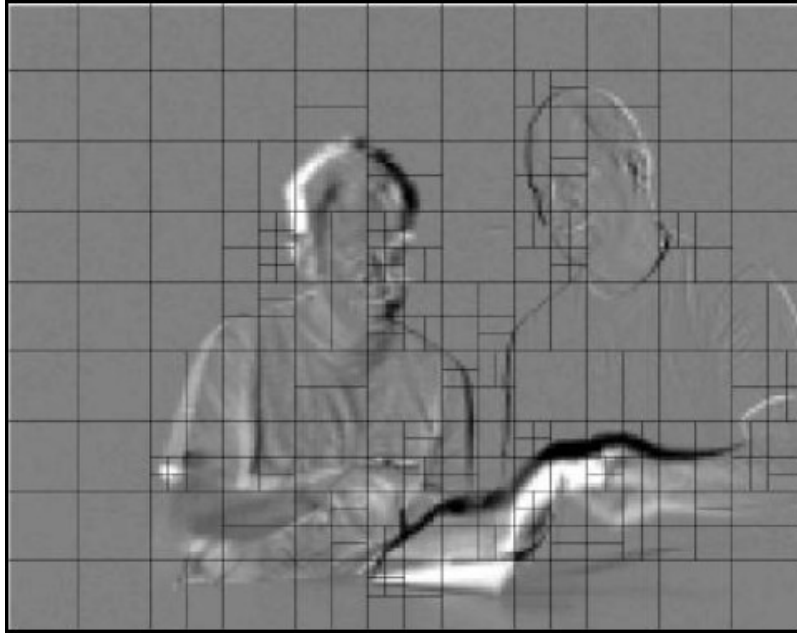
2.2.5.3 Makro blok segmentasyonu



Şekil 2.31 : Farklı hareket vektörleri kombinasyonları [6].

Şekil 2.31'de H.264'ün hareket vektörleri oluşturma mekanizmasından küçük bir kesit sunulmuştur. H.264'ün mekanizması 16x16'lık blokların dışında şekilde de görüldüğü üzere daha küçük blok yapılarına izin verebilmektedir. Özellikle objelerin sınır bölgelerinde ve hareketliliğin fazla olduğu ya da olabileceği bölgelerde bu mod çok büyük bir avantaj sağlamaktadır. Kod çözücüye bu durumda bir hareket vektörü

yerine birden çok hareket vektörü bilgisinin ulaştırılması gerektiği de unutulmamalıdır [1,6]. Şekil 2.32'de örnek bir segmentasyon görülmektedir. Bu işlem hareket vektörlerinin çıkarılması sırasında kestirimin doğruluğunun arttırılmasına yönelik önemli bir işlemdir. Özellikle video çerçeveleri içindeki resimlerin kenar noktaları başta olmak üzere yer değiştirebilme ihtimali yüksek olan bölgeler daha fazla makro blok segmentasyon işlemine uğrarlar ve bu noktalarda daha detaylı bir analiz yapılır.



Şekil 2.32 : H.264'de örnek bir çerçeve için makro blok segmentasyonu [6].

2.2.6 Sonuçlar

Genel olarak bakıldığında H.264 ile birlikte video çözücüler dünyasında pek çok yeniliğin de birlikte geldiğini söylemek abartı sayılmayacaktır. Video işleme ve sıkıştırma sahasının pek çok alanında kendine yer edinen bu kodlayıcının belli başlı bazı özelliklerine yeniden vurgu yapılarak bu bölüm nihaytlendirilecektir.

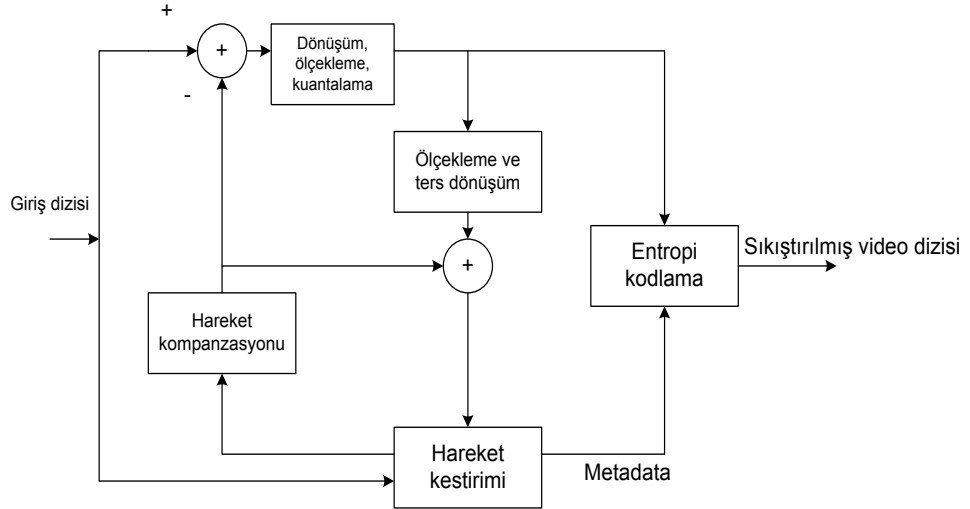
H.264'de çoklu referanslar kullanarak çerçeveler arası kestirim yapmak mümkündür ki burada referans çerçevenin kestirilecek çerçeveye olan uzaklığına bakılarak ağırlıklandırma yapılır ve bu MPEG-2'ye nispeten önemli bir artıdır. Yine uzaysal domende geniş bir mod seçimi olanağıyla çerçeve içi kestirim olanağını kullanıcıya sunan H.264 hata önleyici çeşitli küçük araçlarıyla da dikkat çekicidir. İyi seviyede bloklaşma giderici süzgeç yapısı içermesi, çerçeveye göre uyarlamalı makro blok segmentasyonu ve verimli aritmetik kodlayıcı bölümü ile bunlara örnekler bulmak

mümkündür. Ayrıca farklı resim grupları arasında çerçeve değişimlerinin sağlanabilmesi, dilim sıralamalarının kullanıcı tarafından belirlenebilmesi ve Hadamard dönüşümü gibi yüksek oranlarda sıkıştırma sağlayabilen ek dönüşüm olanaklarının kullanıcıya sunulması 4:4:4, 4:2:2 ve 4:2:0 gibi farklı örnekleme uzaylarına uyumluluğu kullanıcıları cezbeden önemli diğer niteliklerdir. Tüm bu bakış açısı altında geniş bir bit oranı yelpazesinde H.264 diğer kodlayıcılara nispeten önemli şekilde üstün sayılabilir ki [12]'de bunu gösteren birkaç örnek bulunmaktadır. Bu simülasyonlar da henüz Dirac kullanılmamıştır. En genel başarımları kıyaslanması Dirac anlatıldıktan sonra yeniden ele alınacaktır. Sonuç itibarıyla, yapılan bu son video çözümlerin genel hatlarına bakış çatısı altında, Dirac daha iyi anlaşılabilir ve Dirac'ın da bu video dünyasında kısa zamanda aldığı mesafeye binaen önemli daha üst başarımları yakalayabileceği çıkarımını yapmak pek de zor olmayacaktır.

3. DIRAC ALGORİTMASI

3.1 Dirac Kodlayıcı Yapısı

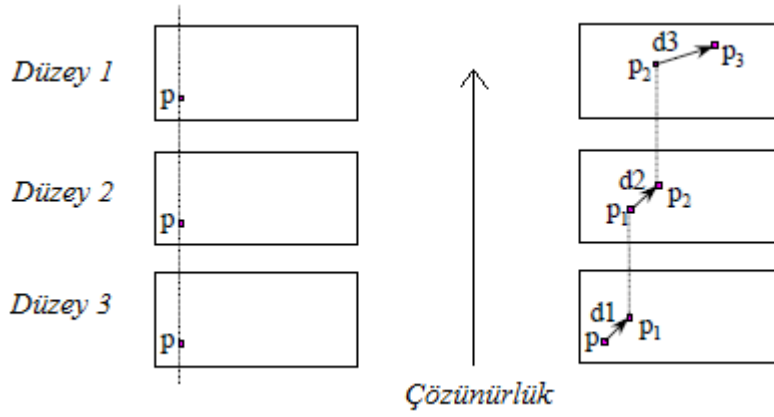
Dirac video çözücü her şeyden önce basitliği ile öne çıkmaktadır. Hareket kompanzasyonu ve eşleştirmesi için getirdiği basitlik ve en önemlisi dalgacık tabanlı kodlama bu çözücüde ilk göze çarpan ayırt edici özellikler olarak belirtilebilir. Pek çok üst seviye algoritmada 8x8'lik ya da 16x16'lık bloklar hareket kestirimi için kullanılırken Dirac 'da bunun için 4x4 'lük boyutlara kadar inilebilmektedir ki bu hareketin değişimini kestirmek açısından son derece önemlidir. Yine vektör analizi olarak 1/8 piksellik boyutlarda blok tabanlı (örtüşen bloklar kullanarak) analiz yapılarak ağırlıklandırılmış hareket vektörü çıkartılır ve bu hatanın en aza indirilmesi açısından avantaj taşımaktadır. İlerleyen bölümlerde belirtilen tüm bu sekmeler ele alınmaya çalışılacaktır. Şekil 3.1 'de [1] genel Dirac mimarisi görünmektedir. MPEG-2 ve özellikle H.264 mimarisiyle karşılaştırıldığında çok daha sade olduğu anlaşılabacaktır.



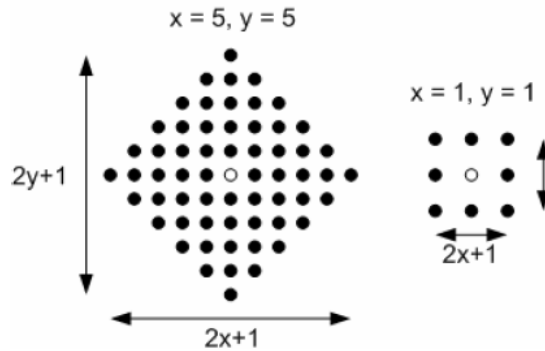
Şekil 3.1 : Genel Dirac kodlayıcı mimarisi [1].

Hiyerarşik hareket kestiriminin kullanılması diğer video çözücülerde pek kullanılmayan bir nitelik olarak Dirac sistematiğini basitleştirmektedir. Bu mimaride çoklu çözünürlük anlayışı öne çıkar. Şekil 3.2'deki yapıda görüldüğü üzere seviye 1'den itibaren referans ve o anki çerçevelerin eş zamanlı olarak çözünürlükleri düşey

ve yatay olarak ikiye kat düşürülür. Sistemin basitleştirilmesi burada sadece tam taramalı kestirim yerine elmas biçimli taramalı ve 9 noktalı küçük yapıda tam taramalı kestirim kullanılmasından ileri gelmektedir [1]. Böylece diğer çözücülerdeki gibi geniş boyuttaki tam taramalı yapılar kullanmak yerine elmas biçimli model ve sadece 9 noktadan ibaret olan tarama alanları belirlenir ve hareket vektörleri bu yolla belirlenir [1]. Burada ayrıca her seviyede elde edilen hareket vektörü bilgisi bir üst seviyedeki kestirime iletilerek daha detaylı bir hareket kestirimi de gerçekleştirilmiş olur. Bunlara ek olarak bir de çözünürlük düşürülmesi sırasında çerçevelerin her birine alçak geçiren filtreleme işlemi uygulanır. Burada [1]'de belirtilen filtre biçimleri kullanılmaktadır. Bu noktadaki temel amaç da çerçeveler üzerinde oluşmuş olan gürültüsel bileşenleri temizlemektir. Sonuçta Şekil 3.2'de belirtilen 'p' noktasının toplam yer değiştirmesi ' $d1+d2+d3$ ' olarak belirlenir. Burada üç seviye var olduğu varsayılmıştır. Bu gösterim içindir. Dirac'da normal olarak CIF videolar için 4, HD videolar için ise 6 olarak bu seviye sayısı belirlenir ve kestirim gerçekleştirilir.

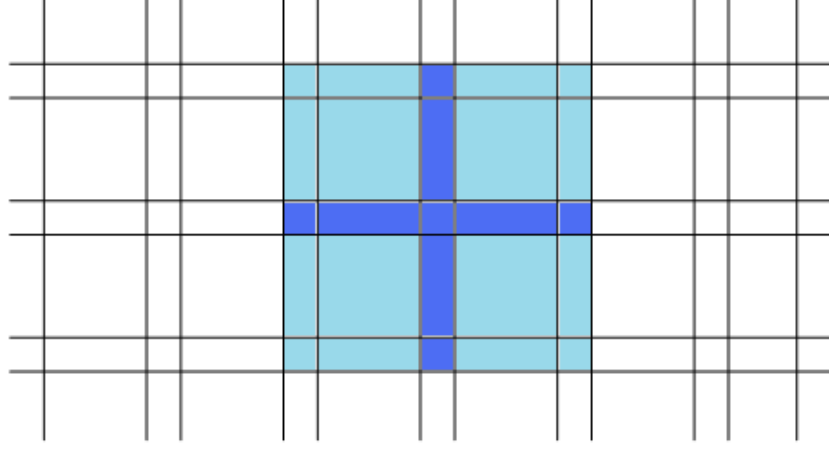


Şekil 3.2 : Hiyerarşik hareket kestirimi [17].



Şekil 3.3 : Dirac'da temel kestirim tarama yapıları [1].

Dirac'da kullanılan bir diğer farklı hareket kestirimi yaklaşımı örtüşen bloklarla bu işin yapılmasıdır. Yukarıda anlatılan hiyerarşik model gerçekleştirirken aynı zamanda kullanılan makro bloklar alt makro bloklara ayrılırlar. Burada makro blok boyutları kullanılan giriş video dizisinin boyutlarına bağlıdır.



Şekil 3.4 : Hareket kestiriminde örtüşen blokların gösterimi [1].

Şekil 3.4'de boyalı kısım tüm bir makro bloğu gösterirken açık mavi alanlar alt makro blokları göstermektedir. Koyu mavi kısımlar ise örtüşen kısımları sembolize etmektedir. Aslında bu metodun bir önemli avantajı örtüşen kısımların dört kez kestirime sokularak en iyi hareket vektörünün saptanmasında kolaylık sağlanmasıdır. Sonuçta da tüm alt makro bloklar için saptanan hareket vektörlerinin ağırlıklandırılmasıyla ortaya çıkan net hareket vektörü tüm bloğun hareket vektörü olarak kompanzasyona yollanır. Bu yöntem Dirac'da etkin olarak kullanılan ve kodlama verimini arttıracak önemli araçlardan biridir.

Bu kısımda son olarak Dirac'ın bir diğer sunduğu kolaylığın da yazılımsal araçların kullanıcılara ücretsiz sunulması olduğu belirtilmelidir. Dirac kodlayıcı ve kod çözücü yapısı C++ ve C dillerinde geliştirilmiş olup kolay bir biçimde API dosyaları tanımlanmak mümkündür [3]. Kullanıcılar diledikleri gibi bu ana kodlar üzerinde değişiklikler yapıp kendi özel kodlayıcı ve kod çözücü yapılarını gerçekleyebilirler. Bu yolla ortaya çıkan gelişmelerle Dirac'ı çok daha iyi yerlere getirebilmek mümkündür.

3.2 Dirac Dalgacık Sistematığı ve Kaldırma (Lifting) Yöntemi

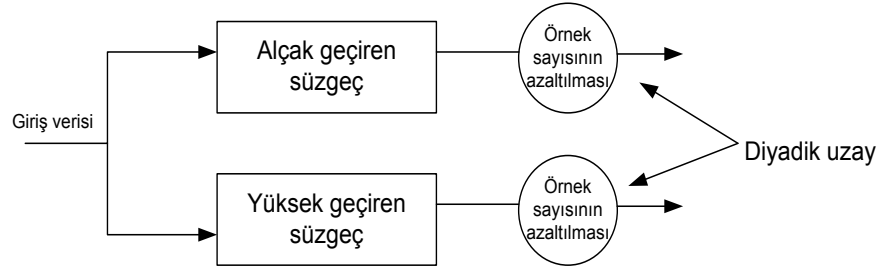
Dirac, giriş video dizisini kodlarken dalgacık metodunu kullanır. Bir önceki bölümde ele alındığı üzere dalgacık metodu giriş verisinin frekans ve zaman bilgisinin en uygun bir şekilde birlikte elde edimini sağlar. Ayrıca DCT 'deki bloklaşma etkisi de kaybolduğundan bunun için ekstra, bloklaşma etkisini giderici süzgeçlere de gerek kalmayacaktır [18,19].

Dalgacık mimarisi pek çok açıdan esneklik getirmektedir. Farklı taban fonksiyonlarının kullanımına izin veren bir dönüşüm olması kodlayıcının farklı uygulamalar için verim çeşitliliği sağlamasına olanak verir ki bu video çözücüler için çok önemli bir getiridir [20,21,22] .

Dirac video çözücünün kullandığı dalgacık fonksiyonlarını sıralarsak Çizelge 3.1 'i incelememiz gerekecektir. Tabii ki belirtilen dalgacıkların Şekil 3.5'deki dalgacık analizi eş değer modelindeki gibi eş değer bir model tabanında incelenmesi gerekmektedir [5].

Çizelge 3.1 : Dirac'da kullanılan temel dalgacık fonksiyonları [3].

Dalgacık Türü	Alçak geçiren süzgeç uzunluğu	Yüksek geçiren süzgeç uzunluğu	Öteleme
Deslauriers-Dubuc (9,7)	9	7	1
Deslauriers-Dubuc (9,7)	13	7	1
Le Gall 5/3	5	3	1
Daubechies (9,7)	9	7	1
Haar 0	2	2	0
Haar 1	2	2	1

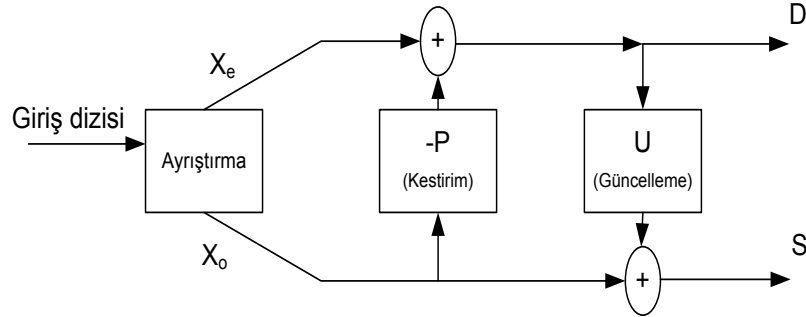


Şekil 3.5 : Dalgacık analizi eş değer modeli [18,19].

Dalgacık fonksiyonları en çok tercih edilen diyadik uzayda çalıştırılırlar [22]. Fakat Dirac bu yapıyı yalnız başına kullanmaz zira işlemsel olarak daha fazla avantaj sağlayan başka bir yöntem daha tercih edilir ki o da kaldırma yöntemi denilen bir tür katsayı kuantalamasıdır. Aslında yapılan şey yine dalgacık dönüşümüdür ancak bu sefer klasik konvolüsyon yerine kestirim ve güncelleme adıyla iki yeni aracı sahne almış olarak görürüz [23,24,25].

$$S = x_0 + Güncelleme(D) \quad (3.1)$$

$$D = x_e + Kestirim(U) \quad (3.2)$$

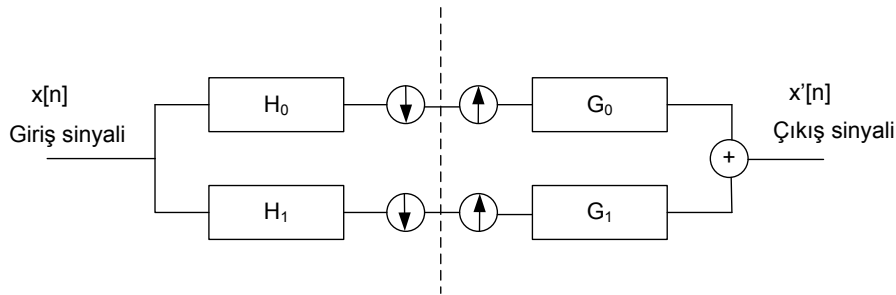


Şekil 3.6 : Kaldırma (lifting) metodu [5,18].

Eşitlik (3.1) ve (3.2) [18,19] ve Şekil 3.6, Dirac 'da işleyen kaldırma metodunu anlamak için bir ön adım teşkil ederler. Bu metodun sonucunda aynen klasik yöntemdeki gibi detay (detail) ve yaklaşım (approximation) kısımlar elde edilecektir [18,19] .

3.2.1 Biortogonal dalgacıklar ve 'Bior 5/3' yapısı

Biortogonal dalgacıklar en çok tercih edilen dalgacık türleridir. Zira, ancak bu koşulu yani biortogonalite özelliğini sağlayan dalgacık türleri sorunsuz bir şekilde, ayrıştırılan giriş sinyalinin tekrardan geri elde edilmesini sağlayabilirler (Perfect Reconstruction). Bu aile dışında tercih edilen dalgacık türleri daha çok özel uygulamalar için geliştirilmiş farklı niteliklere sahip formlardır. Bu kısımda biortogonal dalgacık sistematiği özetle ele alınacak ve Dirac için bu tezde önerilen 'bior' dalgacık ailesinden olan 'Bior 5/3' dalgacık türünden bahsedilecektir.



Şekil 3.7 : Dalgacık alt bant sistematiği [20,21].

Şekil 3.7 dalgacık kodlamasının temel sistematiği görülmektedir. Burada ' H_0 ' ve ' H_1 ' sırasıyla analiz alçak geçiren ve yüksek geçiren filtreleri iken ' G_0 ' ve ' G_1 ' ise sırasıyla sentez alçak geçiren ve yüksek geçiren filtreleridir. Sorunsuz bir analiz ve sentez işleminin sonucunda $x[n] = x'[n]$ koşulunun sağlanması beklenir [26,27]. Sonuç olarak eğer yapılan tüm filtreleme ve 2 ile sırasıyla alt ve üst örnekleme işlemlerinin yapılmasıyla 'Z' domeninde aşağıdaki eşitlik elde edilecektir:

$$X'(Z) = \frac{1}{2} [G_0(Z)H_0(Z) + G_1(Z)H_1(Z)]X(Z) + \frac{1}{2} [G_0(Z)H_0(-Z) + G_1(Z)H_1(-Z)]X(-Z) \quad (3.3)$$

Denklem (3.3)'e göre giriş sinyali ile çıkış sinyali arasında eşitliğin sağlanması için aşağıdaki koşullar sağlanmalıdır:

$$G_0(Z)H_0(-Z) + G_1(Z)H_1(-Z) = 0 \quad (3.4)$$

$$G_0(Z)H_0(Z) + G_1(Z)H_1(Z) = 2 \quad (3.5)$$

Denklem (3.4) ve (3.5) giriş sinyalinin sorunsuz elde edilebilmesi için gerekli olan koşullardır [25]. Bu koşulları sağlayan filtre takımı ise biortogonal filtre yapısını tanımlar. Biortogonal yapı biraz daha detaylı irdelenmek istenirse denklem (3.6)'dan çıkacak eşitlikler öne çıkacaktır [26,28].

$$\begin{bmatrix} H_0(-Z) & H_0(-Z) \\ H_0(-Z) & H_1(-Z) \end{bmatrix} \begin{bmatrix} G_0(Z) \\ G_1(Z) \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad (3.6)$$

Buradan ‘ $G_0(Z)$ ’ ve ‘ $G_1(Z)$ ’ bulunmaya çalışılırsa; Z ve $-Z$ için ifadeler düzenlenerek:

$$G_0(Z)H_0(Z) + G_0(-Z)H_1(-Z) = 2 \quad (3.7)$$

$$G_1(Z)H_1(Z) + G_1(-Z)H_1(-Z) = 2 \quad (3.8)$$

$$G_1(Z)H_0(Z) + G_1(-Z)H_0(-Z) = 0 \quad (3.9)$$

$$G_0(Z)H_1(Z) + G_0(-Z)H_1(-Z) = 0 \quad (3.10)$$

eşitlikleri ortaya çıkmaktadır. Bu eşitlikler biortogonal dalgacık yapısını tanımlayan temel eşitliklerdir. Eğer bu denklemler zaman domeninde yazılırsa [26]’den de yararlanılarak,

$$\sum_k g_0[k]h_0[2n-k] = \delta[n] \quad (3.11)$$

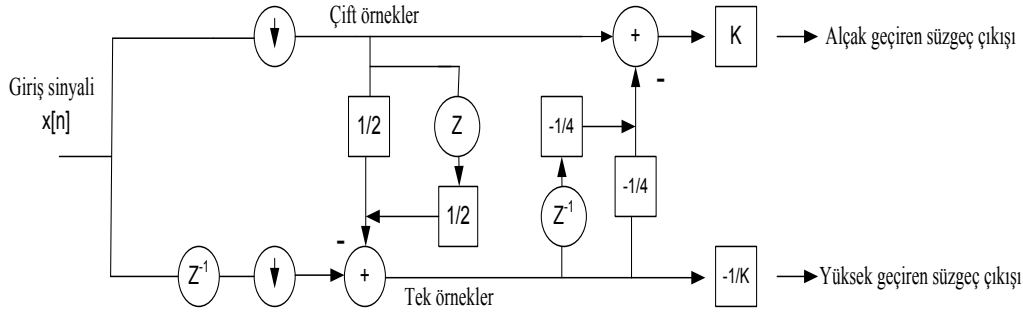
$$\sum_k g_1[k]h_1[2n-k] = \delta[n] \quad (3.12)$$

$$\sum_k g_1[k]h_0[2n-k] = 0 \quad (3.13)$$

$$\sum_k g_0[k]h_1[2n-k] = 0 \quad (3.14)$$

eşitlikleri bulunacaktır [26]. Burada yukarıdaki denklemler incelenirse, analiz edilen herhangi bir 'X' giriş sinyali için biortogonal bir yapının aslında iki farklı ortogonal yapıdan meydana geldiği görülecektir. Bu yapı belirtilen bu özellikleri ile birlikte pek çok esneklik sağlamaktadır. Zira analiz edilen sinyalin sorunsuz geri elde edilmesini direkt olarak sağlamakla birlikte 'ayna filtre' yapılarında olduğu gibi ilave sınırlamalar ya da koşullar gerektirmez [26,29]. Bu noktada 'bior' ailesi detaylandırılacak olunursa bu kombinezonda uygulanan filtrelerin simetrik olması da uygulama açısından büyük kolaylık getirir. Ayrıca belirtilmelidir ki 'bior' dalgacık yapısını diğer klasik biortogonal dalgacıklardan ayıran yanı da işte bu özelliği yani analiz ve sentez filtrelerinin simetrik yapı da olmalarıdır [26,28].

'Bior 5/3' dalgacık yapısı en küçük biortogonal yapı olarak da bilinir. Burada analiz alçak geçiren süzgeç uzunluğu 5, yüksek geçiren süzgeç uzunluğu da 3 olarak alınır.



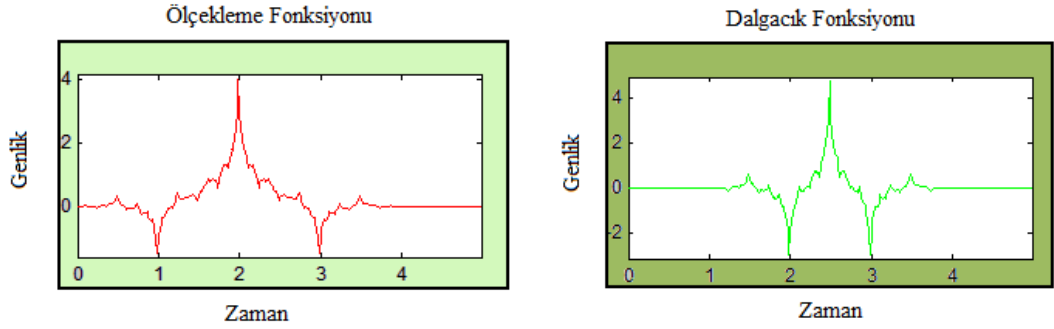
Şekil 3.8 : 'Bior 5/3' süzgeç sistematığı [29].

Şekil 3.8'da görüleceği üzere oldukça basit bir şekilde uyarlanabilen bir yapıda olduğundan dolayı 'Bior 5/3' süzgeci pek çok video ve görüntü işleme algoritmasında da tercih edilen ve ileride de çokça tercih edilebilecek yapıdadır. Bu yapıyı burada söz konusu olan Dirac algoritmasında kullanabilmek için ise 5/3 süzgeç yapısını algoritmaya tanımlamak gerekmektedir [33]. Ayrıca ölçekleme katsayıları da göz önüne alınmalıdır zira kaldırma metodunda kullanılan analiz süzgeçleri bu katsayılarla bir bütün teşkil ederler. Bu yapıyı Dirac'a tanıtmak için kullanılacak temel organizasyon şu şekilde tanımlanabilecektir [26,29,30]:

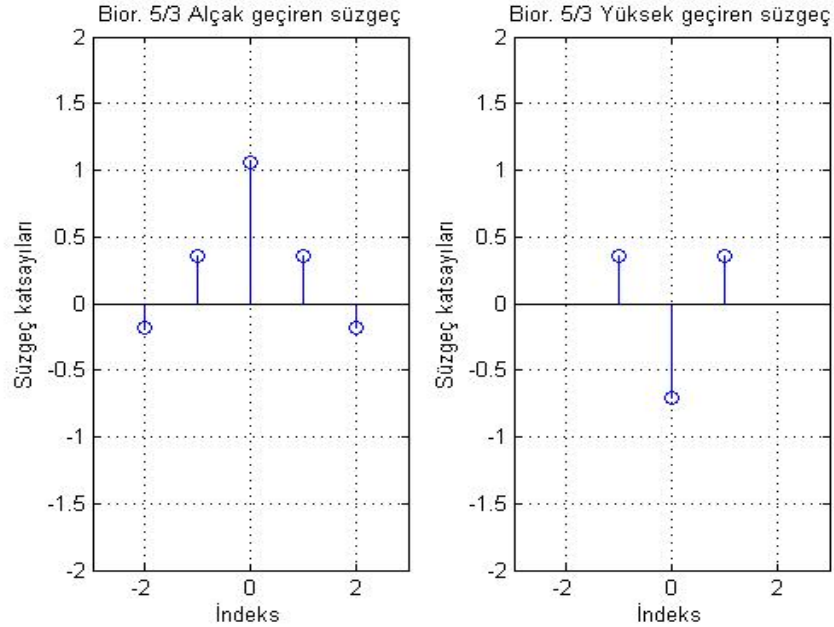
- **Böl:** Öncelikle dönüşümün gerçekleştirilmesi için giriş verisi çift ve tek örnekler olmak üzere iki kısma bölünür [31].
- **Kestir:** Bir önceki kısımda iki kısma ayrılan giriş sinyalinin çift örnekleri kullanılarak sinyalin tek örneklenmiş kısmı kestirilir.
- **Güncelle:** Çift örneklenmiş kısımlar bir önceki adımda kestirilmiş tek örneklenmiş kısım kullanılarak güncellenir [31,32].
- **Kontrol:** Bu adım eğer kullanılan algoritma, 5/3 gibi kısa adımlı süzgeçleri kullanmayıp daha uzun süzgeçleri tercih ediyorsa devreye girer ve kestirim ile güncelleme işlemlerinin devamını sağlarlar.
- **Ölçekle:** Son adım olan bu aşamada ise tüm güncelleme ve kestirim adımlarından sonra elde edilen katsayılar ölçeklenir ve çıkışlar alınır [34].

Bu belirtilen adımlar tek bir kaldırma adımı için geçerli olup Dirac algoritması 4 adıma kadar kaldırma yöntemiyle çıkış verebilmektedir ki bu da üst kısımda bahsedilen bu tüm aşamaların 4 kez tekrarlanması demektir. Bu her ne kadar zaman alıcı gibi görünse de öncelikle bu yapının yani kaldırma yönteminin klasik dalgacık dönüşümüne göre yarı yarıya kadar işlem yükünü azaltabildiği belirtilmelidir [35]. Ayrıca kaldırma yönteminin 4 aşamaya kadar götürülmesi sıkıştırmanın verimliliği

açısından da önem taşımaktadır. Bunlara ek olarak Dirac bahsedilen bu kaldırma yöntemini tamsayı yaklaşımı içerisinde yapar. Bu da ek bir avantaj getirmektedir zira süzgeç katsayılarının reel sayı olarak uzun basamaklar teşkil etmeleri işlem yükünü arttıran bir etmen olduğundan böylece bu olumsuz etki de bertaraf edilmiş olmaktadır. Sonuç olarak tüm bu bakış açıları ve avantajlar göz önüne alınarak 'Bior 5/3' yapısı Dirac algoritmasına tanıtılmış ve sonuçlar gözlemlenmiş ve diğer dalgacık süzgeçleri ile karşılaştırmalar gerçekleştirilmiştir. Bu kısım deneysel sonuçlar bölümünde ele alınacaktır. Uygulanan 5/3 süzgeç yapısını daha iyi irdelemek için Şekil 3.9 ve Şekil 3.10 incelenebilir.



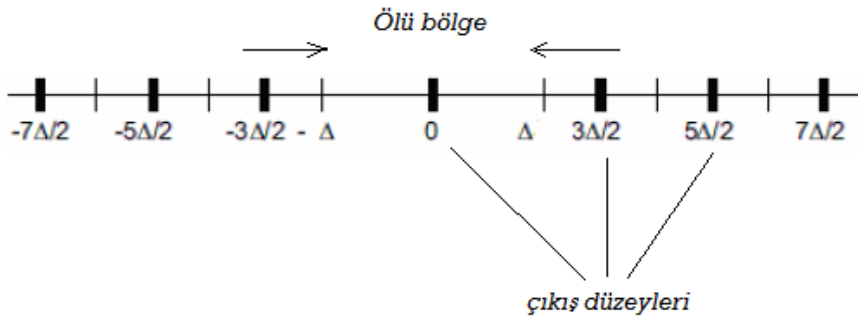
Şekil 3.9 : 'Bior 5/3' süzgeç yapısı (sürekli zaman gösterimi).



Şekil 3.10 : 'Bior 5/3' süzgeç yapısı (ayrık zaman gösterimi-origine taşınmış hali).

3.3 Kuantalama

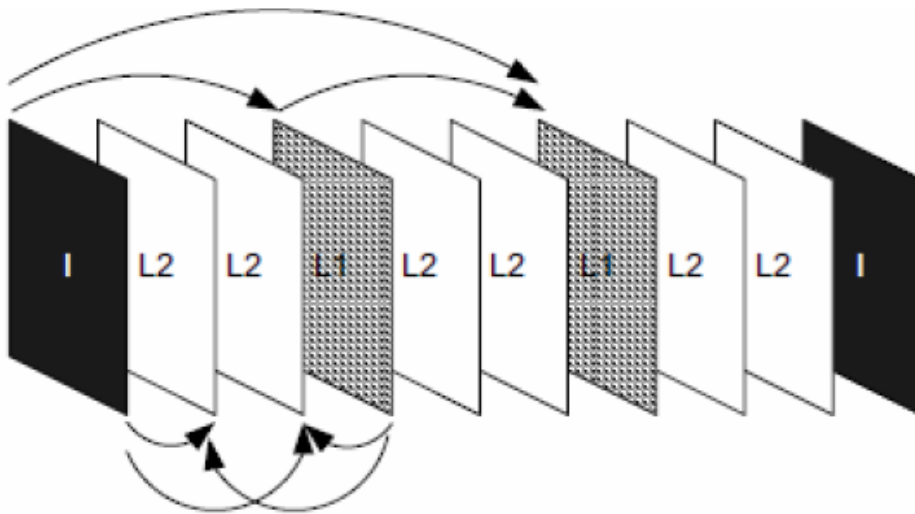
Dirac ölü bölge kuantalamasını kullanır. Bu doğrusal olmayan bir kuantalama çeşididir. Dolayısıyla skalar kuantalama gibi çok sık kullanılan doğrusal tarzdaki kuantalayıcılardan farklıdır. Burada amaç belirlenen aralığa göre düşük değerli katsayıların elenmesidir. Şekil 3.11 'de [1] bu kuantalayıcının çalışma mekanizması özetlenmiştir. Belirtilmelidir ki bu parametre direkt olarak kalite faktörü denilen 'QF' katsayısından etkilenir ve kodlama başlangıcında belirtilen bu katsayının değerine göre yapılacak kuantalama hassasiyeti belirlenir.



Şekil 3.11 : Ölü bölge kuantalaması [1].

3.4 Dirac'da Çerçeve İçi ve Çerçeveler Arası Kestirim

Dirac 'da da tıpkı MPEG serisinde olduğu gibi gerek çerçeve içi olarak gerek de çerçeveler arası kestirim yapmak mümkündür [1,2]. Oluşturulan RG 'da 'L1', 'L2' ve 'I' çerçeveleri olmak üzere toplam üç tip çerçeve bulunur. Bu yapı Şekil 3.12 'deki haliyle diğer video çözücülerle de benzerlik taşımaktadır.



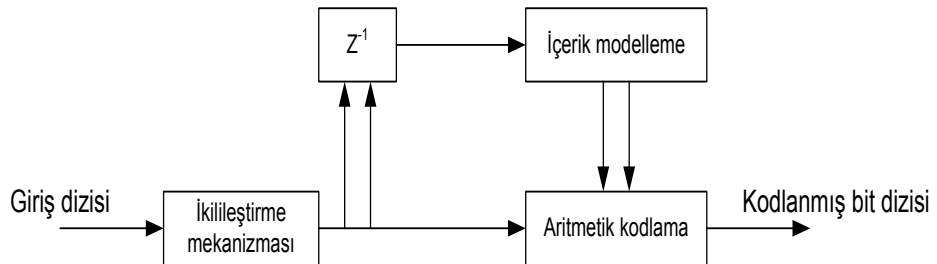
Şekil 3.12 : Dirac'da RG (resimler grubu)'yi oluşturma [1].

Dirac'da kullanılan en ilginç mekanizmalardan birisi hareket kestirimi sırasında dikkat çeker. Burada piramitsel bir hiyerarşi içerisinde düşük çözünürlükten yüksek çözünürlüklü resme kadar sekmeler oluşturulur ve her bir adım için bir hareket vektörü tespit edilir [1]. Sonuç hareket vektörü hesaplanan bu vektörlerinin ağırlaklandırılmışı olarak görülebilir. Bu kısım daha önceki bölümlerde bahsedildiğinden sadece burada çerçeve özelliklerine değinilecektir.

Öncelikle 'I' tipi çerçeveler ele alınacak olunursa bunlara içsel çerçeve tipi de denir ve MPEG-2'deki yapıyla aynıdır [35]. Bu tip çerçeveler diğer çerçeveleri kestirmede kullanılırlar ve ayrıca çerçeve içi kestirim mekanizmasında da rol oynarlar. 'L1' tipi çerçeveler MPEG-2'deki 'P' tipi çerçeveleri andırmaktadır. Bu çerçeveler kendisinden sonra ya da önce gelen 'L2' tipi çerçevelerin kestiriminde görev alırlar. 'L2' tipi çerçeveler ise herhangi bir kestirim görevinde bulunmazlar ve bunlar MPEG-2'deki 'B' tipi yapıyla bire bir benzerlik gösterirler [35]. Bu çerçeve tiplerinin tayin edilişi ve sıralanışı da MPEG-2 ile benzerlik gösterir ve hatta MPEG-2'den daha çok esneklik gösterebilir [35]. Şekil 3.12'de Dirac'daki çerçeve dizilimine bu anlamda daha yakından bakılabilir. Çerçevelere uygulanacak dalgacık filtrelerinin seçimi ise kullanıcıya bağlı olmakla birlikte genelde çerçeve içi ve çerçeveler arası dönüşümde aynı süzgeçler kullanılması sonuçların gözlemi ve değerlendirilmesi noktasında tutarlılık açısından uygun görülür. Ancak farklı kombinezonlar da denenebilir.

3.5 Serpiştirilmiş Üstel Golomb Kodlaması ve İkileştirme Mekanizması

Dirac'da ikilileştirme işlemi çok değerli katsayıları bit dizinine çevirmek için kullanılır. Oluşan bit dizisi Şekil 3.13'de de görüleceği üzere aritmetik kodlayıcıya gönderilir ve kodlanır. Burada amaçlanan içeriksel olarak, çoklu değerde çalışıldığında bazı değerlerin çok az oluşabileceğinden dolayı aritmetik kodlayıcıdaki kodlama veriminin düşmesi söz konusu olduğundan bunun yerine ikili bit dizisinde çalışarak kazancı yukarıya çekmektir. Zira bu noktada içerik modelleyici '0' ve '1'lerin oluşma olasılıklarına göre işleyecektir ve bu olasılıklar ilk duruma göre çok daha yüksektir. Bu aşamada Dirac serpiştirilmiş üstel Golomb kodlamasını kullanır.



Şekil 3.13 : Dirac entropi kodlama blok gösterim [1,35].

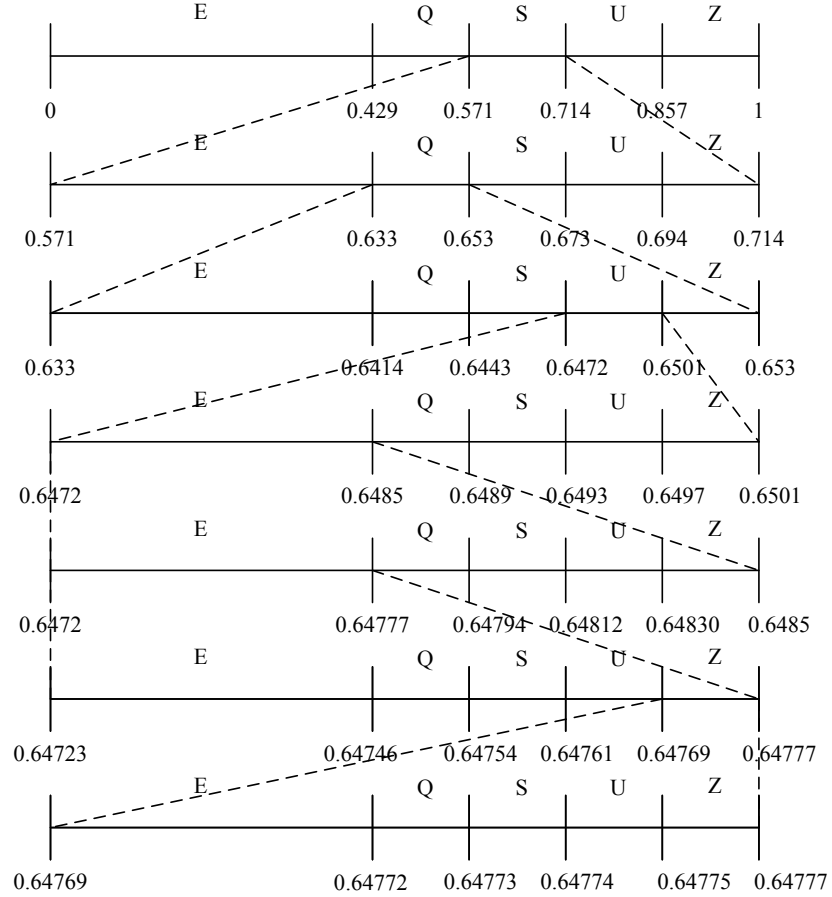
Çizelge 3.2 : Serpiştirilmiş üstel Golomb kodlaması [3,36].

Üstel Golomb bit dizisi	Kuantalanmış piksel değeri	Serpiştirilmiş üstel Golomb bit dizisi
0	0	1
100	1	001
101	2	011
11000	3	00001
11001	4	00011
11010	5	01001
11011	6	01011

Çizelge 3.2’de görüleceği üzere dalgacık dönüşümüne uğramış katsayılar serpiştirilmiş üstel Golomb koduna çevrilirler. Elbette ki bu sıradan bir ikilleştirme işlemi değildir. Burada bütün değerler ‘0’ ile başlatılıp ‘1’ ile bitecek şekilde kodlanır ve ‘ $\text{floor}[\log_2(\text{değer}+1)]$ ’ işlemiyle elde edilen sonuç, kodlanacak olan değerın ‘ID’ serisinin kaç haneli olacağını belirler [36]. Ancak burada işaretli değerler için sistemin ele alındığı hatırlatılmadır. Eğer değerler işaretli değerler olurlarsa (çoğunlukla bu durum geçerlidir.) o takdirde son bit işaretçi konumunda olur ve negatif değerler için ‘1’ ve pozitif değerler için ‘0’ olarak sadece son bitte değişiklik yapılır. Genel olarak temel mantığın aynı olduğu söylenebilir. Buradan sonra kodlanmış olan asıl değer bitleri ‘ID’ bitleri arasına dağıtılırlar [36]. Sonuçta kodlanmış bu dizi hem aritmetik kodlayıcıya hem de çözülmek üzere içerik modelleyiciye yollanır. İçerik modelleyicide hem bu diziler çözülür hem de elde edilen dalgacık dönüşüm katsayılarına göre ‘baba-oğul’ hiyerarşisi altında aritmetik kodlayıcı için gerekli olan giriş simge olasılıkları üretilir [31,35]. Bu noktada belirtilmelidir ki aslında Dirac’da içerik modelleme kısmı, bahsedilen bu bilgiler ışığında hem ‘baba-oğul’ yaklaşımını kullanması hem de serpiştirilmiş Golomb modelinden yararlanması açısından kodlama kazancını önemli ölçüde arttıran bir mekanizma sayılabilir. Fakat H.264 ile bu noktadan kıyaslanırsa bu tarzın kesinlikle bir ‘CABAC’ yapısı kadar kapsamlı, verimli fakat bir o kadar da karmaşık olamayacağı aşikârdır. Son olarak serpiştirilmiş Golomb kodunun kullanılmasının H.264’e göre bir avantaj olduğu da söylenmelidir. Çünkü bu yapıda asıl değer bitleri dağıtıldığı için kod çözücünde bu değer bitlerinin çözülmesi daha hızlı olabilecektir. H.264 ise klasik üstel metodu kullandığından bu noktada biraz gecikme söz konusudur.

3.6 Aritmetik Kodlama

Aritmetik kodlayıcı genel olarak tercih edilen Huffman entropi kodlayıcısından özellikle uzun veri yapıları için kodlama kazancı açısından çok daha avantajlıdır ve bu yüzden de Dirac’da tercih edilir.



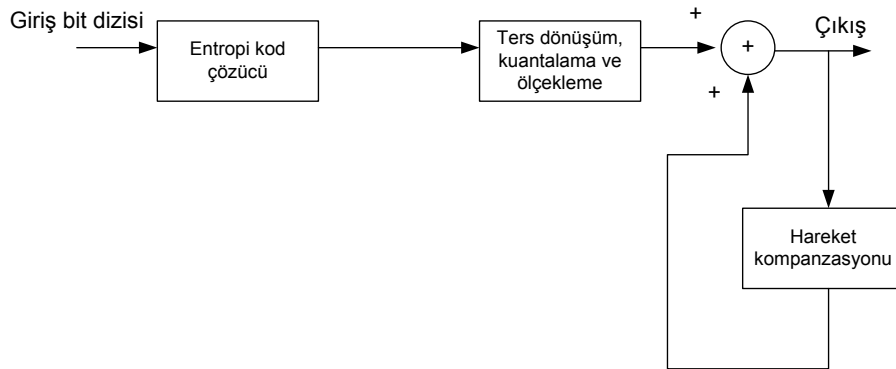
Şekil 3.14 : Aritmetik kodlama mekanizması (‘Squeeze kelimesi üzerinde örneklendirilmiştir.’) [15].

Şekil 3.14’de ‘Squeeze’ kelimesi için örnek bir aritmetik kodlama gösterilmiştir. Şekildeki kodlamada olasılık eksenini şu şekilde belirlenmiştir: $E=1/7$, $Z=S=U=Q=1/7$. Bu örnekten yola çıkıldığında görülecektir ki eğer bu kelime ‘Huffman’ ile kodlansaydı 15 bit gönderilmesi gerekirken burada şekilden de görüleceği üzere 0,64749’un en yakın ikili karşılığı olarak 0,101001011101 bit dizisi yollanır. Burada baştaki sıfır biti yollanmazken yaklaşık 3 bitlik bir kazanç söz konusu olmaktadır ki başında da belirtildiği gibi uzun yapılarda bu kazancın daha da artması beklenilmektedir [15]. Burada sadece mekanizmanın nasıl işlediği gösterilmek istenmiştir.

Son olarak eklenilmek istenirse, bu yapı H.264 ya da MPEG'deki uyarlamalı kodlayıcı organizasyonlarından çok daha basit ve oldukça verimli bir yapıdır denilebilir. Ayrıca Dirac'daki aritmetik kodlayıcı mekanizması 2005 yılında [4]'de ele alındığı üzere bir adım daha ileri götürülmüş ve klasik yapının doğurduğu zamansal gecikme minimize edilmiştir.

3.7 Dirac Kod Çözücü Yapısı

Dirac yukarıda bahsedilenler ışığında basit olan sistematüğını kod çözücü kısmında da korumaktadır. Şekil 3.15'da görüleceğı üzere kodlama işleminde kod çözücüye yollanan kodlanmış veri bitleri kod çözücüde kolayca çözölür. Burada kod çözücüye yollanan bit dizini içinde hem giriş video dizisine ait çerçevelerin önemli bir kısmının bilgileri hem de hareket kestirimi yoluyla elde edilmiş fark görüntüler mevcuttur. Ayrıca kod çözücüde hareket kompanzasyonunun doğru yapılabilmesi için kodlayıcıda kestirim sırasında elde edilip kod çözücüye yollanan hareket vektörlerinin ve koordinatlarının bilgileri de kod çözücüde çözölür. Böylece çözölün fark görüntü bileşenleri ile hareket vektörlerinden yararlanılarak elde edilen hareket kompanzasyonu yapılmış görüntüler birleştirilir ve yüksek bir kalite performansı ile çıkış elde edilir

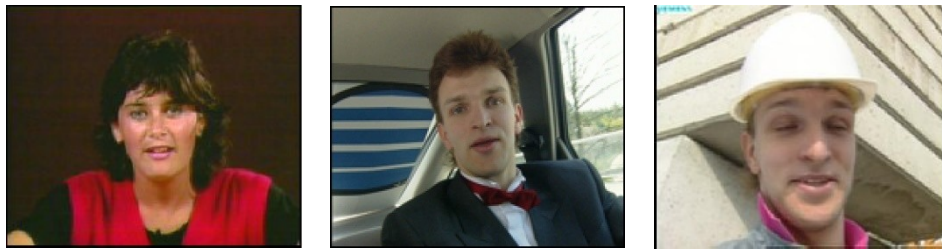


Şekil 3.15 : Dirac kod çözücü sistematüğü [1,32].

4. DENEYSEL SONUÇLAR

4.1 Giriş

Yapılan analizlerde öncelikle QCIF, CIF ve 4CIF video dizileri için tüm dalgacık tabanlarında Dirac performans analizi PSNR ve SSIM parametreleri tabanlı olarak gözlemlenmiş ve ilk olarak [38]'daki matematiksel çıkarımın Dirac'a uyarlanması olan farklı dalgacıklar da hesaba katılarak farklı hareket vektörleri durumları için tüm sonuçlar elde edilmiştir. Burada ve tüm analizlerde Dirac 1.0.2 (2009 sürümü) kullanılmış ve gerekli kaynak kodlar C++ tabanında Visual Studio 2008 derleyicisi kullanarak derlenmiştir. Ayrıca H.264 için de yine [39]'den yararlanarak C++ tabanında yazılmış JM (sürüm 18.0) kod dizini kullanılmıştır. Bahsi geçen Dirac 1.0.2 ve JM kod dizini 'Release mod' da çalıştırılmıştır. Ölçüm sonuçlarının önemli bir kısmında [40]'den sağlanan analiz programı kullanılmıştır. Analizlerde kullanılan PC'nin temel özellikleri de şöyledir: Intel(R) Pentium(R) Dual CPU T 2390 @ 1,86 GHz ve 2,38 MB RAM bellek ve 32 Bit işletim sistemi. Dirac ve bu bölümün sonraki kısımlarında performans kıyaslaması için kullanılacak olan H.264 için analizlerde QP 'default' değerler alınmıştır. (Dirac için 5, H.264 için 28) Burada amaçlanan farklı dalgacık tabanlarındaki Dirac performansını önceki kısımlarda anlatılan bilgiler ışığında ilk olarak kendi içinde değerlendirmek daha sonra da H.264 ile kıyaslayarak durum değerlendirmesi yapmaktır. Yine bahsedilmelidir ki bu noktada bir de Dirac için yeni bir yapı olarak Bior. 5/3 dalgacığı önerilmiş olduğundan yapılan analizler bu dalgacığın Dirac performansına yaptığı katkılar ışığında ayrıca değerlendirilecektir.



Şekil 4.1 : Analizlerde kullanılan QCIF-YUV videolar (*Soldaki 'miss-america.qcif' ortadaki 'carphone.qcif' ve sağdaki 'foreman.qcif'*) [41].



Şekil 4.2 : Analizlerde kullanılan CIF-YUV ve 4CIF-YUV videolar (*Soldaki 'bus.cif' ortadaki stephan.cif' ve 'sağdaki 'harbour.4cif'*) [41].

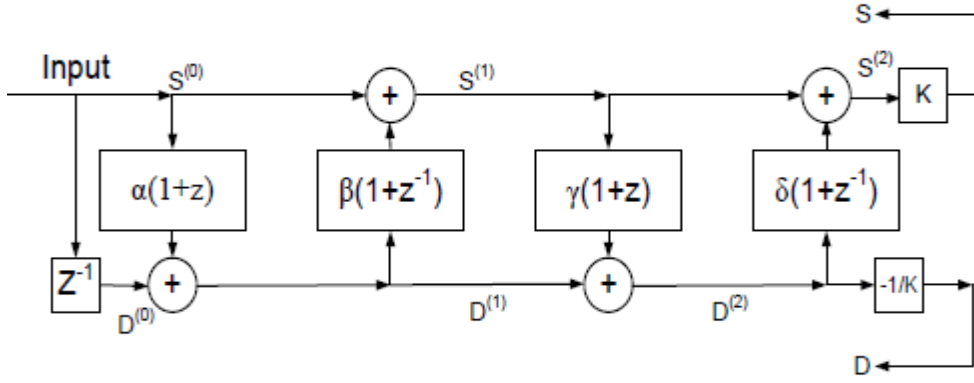
Çizelge 4.1 : Analizlerde kullanılan video dizilerinin temel bilgileri.

Temel bilgiler	Miss- america	Carphone	Stephan	Foreman	Bus	Harbour
Süre (s)	6	9	14	12	7	15
Resim frekansı (Hz)	25	25	6.4285	25	21.42	20
Genişlik (piksel)	176	176	352	176	352	720
Uzunluk (piksel)	144	144	288	144	288	576
Bit hızı (KB/s)	930	567,3	977.55	928,1	3182,1	119,7
Boyut (KB)	5569	14182	13365	11138	22275	178176
Top. çerçeve sayısı	150	382	90	300	150	300

4.2 QCIF ve CIF Videolarda 1, 1/2, 1/4 ve 1/8 Pel Hareket Kestirimleri İçin Farklı Dalgacık Tabanlarında Dirac Performans Analizi

Aşağıdaki çizelgelerde 'Foreman.qcif', 'Miss-america.qcif' ve 'Bus.cif' video dizileri hakkında farklı hareket kestirimleri için yine farklı dalgacık tabanlarında PSNR, SSIM ve kodlanmış dosya boyutu (EFS) baz alınarak bir takım analizler yapılmıştır. QCIF ve CIF video dizilerinin tercih edilmesinde, hızlı sonuç alabilme etkili olmuştur. Burada önemli bir nokta öncelikle belirtilmelidir ki Dirac, gelen tüm giriş video dizilerini 1, 1/2, 1/4 ve 1/8 hareket kestirimleri için sıkıştırma işlemini gerçekler ve SSIM olarak en iyi sonuç veren hareket vektörü analizi temel alınarak sıkıştırma işlemi neticelendirilir [35]. Dolayısıyla hemen söylenilebilir ki aşağıda yapılan analizler, en iyi PSNR ve SSIM oranlarının görülmesi açısından oldukça önemlidir. Çünkü böylece kullanılan herhangi bir dalgacık için en iyi olabilecek hareket kestirimi analizi bulunabileceği gibi aynı zamanda çalışılacak herhangi bir giriş videosu için bu tarz bir analizle PSNR, SSIM, dosya boyutu ve kodlama zamanı dikkate alınarak performansı istenilen yönde en iyiye taşıma mümkün olabilecektir.

Aşağıdaki çizelgelerde normal Dirac dalgacık modellerinin yanı sıra [37,38]'daki matematiksel analizden yararlanılmıştır. Çizelgelerin hepsinde ortak nokta olarak diyebileceğimiz şey 1/4 pel analizinde en iyi kalitede sonuçların yakalanabildiğidir. Üstte de belirtildiği gibi Dirac bu analizi herhangi bir giriş video dizisi için bütün durumlar için yaptığından ötürü 1/4' de en iyi sonucun yakalandığını görmek kullanıcıyı daha başlangıçta doğru bir hareket kestirimi analizinin seçimini yapabilmesi için yönlendirecektir. Böyle bir analizle kodlama zaman kazancı buradaki bütün test videoları için 4-5 saniye dolaylarındadır ki bu Dirac gibi hızlı bir kodlayıcı için önemli bir süredir. Aslında 1/4'de sonuçların en iyi çıkacağı en başından belliydi denebilir. Zira pek çok görüntü işleme algoritmasında da 1/4 pel boyutlarında hareket kestirimi analizi en iyi sonuç verdiğinden zaten tercih edilir. Ancak bu bir kural değildir zira bir takım videolarda çözünürlük, hareket sıklığı ve yönleri gibi etmenlerden dolayı farklı pel boyutlarında analizlerde bizi en iyi sonuca (PSNR ya da SSIM olarak) ulaştırabilir. Nitekim Çizelge 4.3, Çizelge 4.4 ve Çizelge 4.5'de bunun bazı örneklerine rastlayabiliriz. Özellikle Çizelge 4.3'ye dikkat edilmelidir.



Şekil 4.3 : Yeni dalgacık modelleri olarak sunulan Durum 1, 2, 3 ve 4 için analiz yapısı [38,42].

Burada Durum 1, 2, 3 ve 4 [38]'de önerilen yapılardır ve burada ilk olarak bu çalışmada Dirac'da uygulanmıştır. Bu durumlar Daubechies 9/7 orijinal yapısından [38,42,43]'de yapılan matematiksel yaklaşımlarla çıkarılan yeni yapılardır denilebilir. Bu yapıların üstteki çizelgelerde hem QCIF hem de CIF videolarda önemli ölçüde performansı iyileştirdiği görülmektedir. Performansın iyiye gittiğini kodlama zamanı açısından da gözlemlemek mümkündür. Burada normal Daubechies kaldırma yöntemi katsayılarında $\alpha=-1,58613434236$, $\beta=-0,05298011857$, $\gamma=0,88291107553$ ve $\delta=0,44350685204$ olarak alınmaktadır. Bu katsayıların uzunluğu tam sayı yaklaşımlı dalgacık dönüşümü ile kısmen giderilse de yapılan dalgacık dönüşümü işleminde yine bu katsayıların uzunluğundan dolayı işlem süresi uzamaktadır. Bu bir çerçeve için küçük bir süre olsa bile çerçeve sayısı fazla olan video dizilerinde önemli kodlama süresi gecikmelerine yol açabilmektedir. Önerilen yeni durumlar bu açıdan da işlem yükünü azaltıcı yöndedir. Çizelge 4.2 bu anlamda gözlemlenmelidir. Çizelgedeki her bir durum Daubechies 9/7 temelli farklı bir kaldırma yöntemini yansıtmaktadır [38].

Çizelge 4.2 : Daubechies 9/7 dalgacık yapısından doğan yeni durumlar için kaldırma yöntemi katsayıları [38].

Dalgacıklar	α	β	γ	δ
Durum 1	-5/4	-1/9	9/16	16/27
Durum 2	-4/3	-9/100	25/39	1079/2000
Durum 3	-3/2	-1/16	4/5	15/32
Durum 4	-8/5	-25/484	121/135	9369/21296

Çizelge 4.3 : 'Foreman.qcif' için hareket kestirimi mekanizmasının farklı dalgacık tabanlarında sıkıştırmaya etkisi.

Dalgacıklar	Performans parametreleri	Hareket kestirimi analizi (pel)			
		1	1/2	1/4	1/8
Le Gall 5/3	PSNR (dB)	32	32,184	32,33	32,326
	SSIM	0,916	0,919	0,921	0,920
	EFS (kB)	199	198	206	215
Daubechies 9/7	PSNR (dB)	33	33,27	33,51	33,354
	SSIM	0,922	0,924	0,926	0,925
	EFS (kB)	217	213	207	229
Deslauriers-Dubuc 13/7	PSNR (dB)	32,194	32,46	32,612	32,464
	SSIM	0,916	0,919	0,920	0,918
	EFS (kB)	202	200	207	216
Deslauriers-Dubuc 9/7	PSNR (dB)	31,94	32,26	32,34	32,24
	SSIM	0,913	0,917	0,918	0,916
	EFS (kB)	198	198	205	214
Haar 0	PSNR (dB)	32,4	32,6	33	32,92
	SSIM	0,932	0,919	0,924	0,922
	EFS (kB)	225	223	228	237
Haar 1	PSNR (dB)	32,72	32,957	32,4	33,31
	SSIM	0,919	0,921	0,925	0,924
	EFS (kB)	227	224	229	239
Durum 1	PSNR (dB)	32,89	33	33,33	33,17
	SSIM	0,92	0,921	0,923	0,922
	EFS (kB)	211	207	214	223
Durum 2	PSNR (dB)	32,95	32,2	33,47	33,25
	SSIM	0,921	0,923	0,926	0,923
	EFS (kB)	213	209	216	225
Durum 3	PSNR (dB)	33	33,3	33,47	33,37
	SSIM	0,923	0,925	0,927	0,925
	EFS (kB)	217	213	218	228
Durum 4	PSNR (dB)	32,98	33,3	33,5	33,3
	SSIM	0,922	0,924	0,926	0,924
	EFS (kB)	218	214	220	230

Çizelge 4.4 : 'Miss-america.qcif' için hareket kestirimi mekanizmasının farklı dalgacık tabanlarında sıkıştırılmaya etkisi.

Dalgacıklar	Performans parametreleri	Hareket kestirimi analizi (pel)			
		1	1/2	1/4	1/8
Le Gall 5/3	PSNR (dB)	36,8	36,75	37,11	36,77
	SSIM	0,958	0,958	0,958	0,956
	EFS (kB)	58	49	48	48
Daubechies 9/7	PSNR (dB)	38,57	38,49	38,65	38,42
	SSIM	0,954	0,952	0,952	0,948
	EFS (kB)	64	54	52	52
Deslauriers-Dubuc 13/7	PSNR (dB)	37	36,83	37	36,97
	SSIM	0,958	0,956	0,956	0,955
	EFS (kB)	59	50	49	49
Deslauriers-Dubuc 9/7	PSNR (dB)	36,52	36,32	36,61	36,5
	SSIM	0,956	0,954	0,955	0,953
	EFS (kB)	58	50	48	48
Haar0	PSNR (dB)	37	36,8	37,19	37
	SSIM	0,959	0,956	0,958	0,956
	EFS (kB)	67	57	55	55
Haar1	PSNR (dB)	38,46	38,12	38,58	38,48
	SSIM	0,961	0,959	0,96	0,958
	EFS (kB)	68	58	56	56
Durum 1	PSNR (dB)	38,48	38,41	38,63	38,45
	SSIM	0,958	0,957	0,958	0,956
	EFS (kB)	62	52	50	50
Durum 2	PSNR (dB)	38,67	38,59	38,7	38,579
	SSIM	0,959	0,958	0,958	0,956
	EFS (kB)	62	53	51	51
Durum 3	PSNR (dB)	38,75	38,75	39	38,78
	SSIM	0,961	0,960	0,961	0,958
	EFS (kB)	63	53	51	51
Durum 4	PSNR (dB)	38,44	38,36	38,45	38,35
	SSIM	0,952	0,950	0,950	0,947
	EFS (kB)	64	54	52	52

Çizelge 4.5 : 'Bus.cif' için hareket kestirimi mekanizmasının farklı dalgacık tabanlarında sıkıştırmaya etkisi.

Dalgacıklar	Performans parametreleri	Hareket kestirimi analizi (pel)			
		1	1/2	1/4	1/8
Le Gall 5/3	PSNR (dB)	28,87	29,86	30,31	30,02
	SSIM	0,889	0,903	0,909	0,907
	EFS (kB)	537	513	515	519
Daubechies 9/7	PSNR (dB)	29,83	31	31,47	31,41
	SSIM	0,90	0,914	0,92	0,918
	EFS (kB)	613	576	572	595
Deslauriers-Dubuc 13/7	PSNR (dB)	28,94	30,15	30,46	30,31
	SSIM	0,88	0,902	0,909	0,906
	EFS (kB)	545	520	521	544
Deslauriers-Dubuc 9/7	PSNR (dB)	28,72	29,92	30,21	30,24
	SSIM	0,88	0,902	0,906	0,905
	EFS (kB)	538	515	516	536
Haar0	PSNR (dB)	29,65	30,57	31,04	30,94
	SSIM	0,89	0,91	0,916	0,916
	EFS (kB)	616	582	579	601
Haar1	PSNR (dB)	29,78	30,73	31,49	31,41
	SSIM	0,9	0,91	0,919	0,918
	EFS (kB)	617	582	579	602
Durum 1	PSNR (dB)	29,42	30,55	30,98	30,88
	SSIM	0,893	0,906	0,912	0,91
	EFS (kB)	553	520	521	542
Durum 2	PSNR (dB)	29,5	30,68	31,19	31
	SSIM	0,89	0,91	0,915	0,912
	EFS (kB)	570	537	534	556
Durum 3	PSNR (dB)	29,78	30,75	31,32	31,33
	SSIM	0,9	0,91	0,917	0,917
	EFS (kB)	597	562	560	584
Durum 4	PSNR (dB)	29,78	30,91	31,52	31,45
	SSIM	0,901	0,914	0,92	0,919
	EFS (kB)	614	576	572	597

4.3 Bior 5/3 Dalgacığın Dirac Performansına Etkisi ve Farklı Dalgacık

Tabanlarında Dirac Performansı Genel Analizi

Bior 5/3 yapısı bölüm 3'de detaylarıyla ele alınmıştı. Orada da gösterilmek istenmişti ki Bior 5/3 dalgacık yapısı Dirac'ın kendi iç dalgacık süzgeçlerinden daha farklı ve basit bir yapıdadır ve analizi kolaydır. İşte bu yapı Dirac dalgacık veri tabanına yazılımsal olarak eklenmiş ve PSNR, SSIM ve bir sonraki kısımda ele alınacak olan kodlama ve kod çözme zamanı açısından elde edilen deneysel bulgular bu kısımda sunulmuştur.

Çizelge 4.6 : 'Carphone.qcif' için Bior. 5/3 dalgacığının da katılarak tüm dalgacık tabanlarında Dirac performans analizi.

Dalgacık Türü	PSNR-Y (dB)	SSIM	EFS (kB)
Le Gall 5/3	32,66	0,93269	265
Daub. 9/7	33,76	0,93539	287
DD 9/7	32,63	0,93018	265
DD 13/7	32,9	0,93246	268
Bior. 5/3	34,12	0,94613	323
Haar 0	33,29	0,93622	295
Haar 1	33,718	0,93773	297

Çizelge 4.6, Çizelge 4.7 ve Çizelge 4.8'de PSNR-Y, SSIM ve EFS açısından sonuçlar görülmektedir. Eklenen Bior 5/3 yapısı QCIF'de genel olarak 1-1,5 dB civarında iyileşme sağlarken SSIM olarak da 0,01-0,035 arasında artı getirmektedir. Bu, zaten 0 ile 1 arasında seyreden ve Dirac'ın temel aldığı metrik olması açısından büyük önem arz eden SSIM performansına katkı sağlama noktasında önemlidir. Buradaki en önemli dezavantaj denilebilecek kısım bu dalgacığın EFS'yi arttırmasıdır. Ebettteki

bu istenmeyen bir şeydir, ancak Dirac'ın sıkıştırma performansı, QF'in 5 olarak alındığı dikkate alınırsa 1/20'sine hatta daha büyük boyutlardaki giriş dizileri için 1/50'sine kadar dosya boyutunu indirebilecek bir güçtedir. Dolayısıyla çok kritik dosya depolama problemlerinin olmadığı yerlerde bu yapı en iyi performansı elde edebilmek için rahatlıkla tercih edilebilir durumdadır. Ayrıca Dirac'ın 'default' olarak belirtilmiş DD 9/7 dalgacığına göre 4CIF boyutlarında sağladığı yaklaşık 15 saniyelik kodlama ve yine bir o kadar da kod çözme zaman kazancı zamanlamanın hayati olduğu noktalarda oldukça önemlidir.

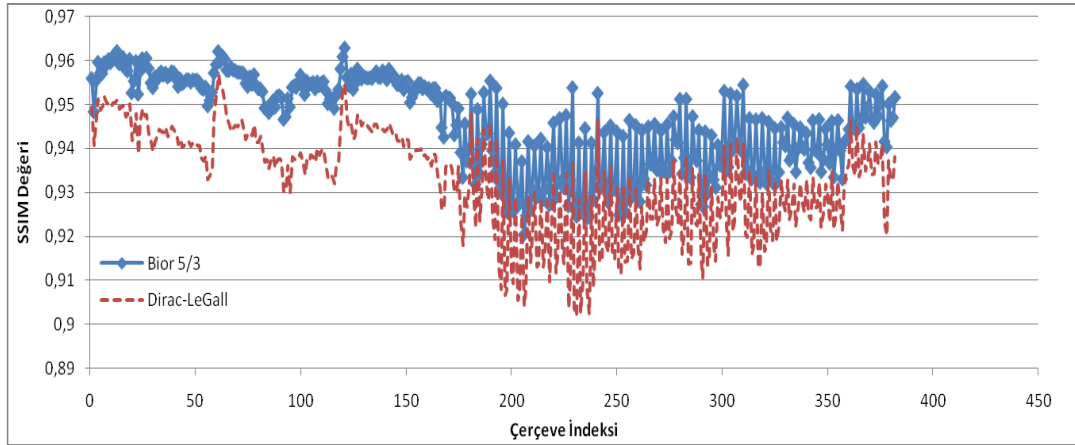
Çizelge 4.7 : 'Bus.cif' için Bior. 5/3 dalgacığının da katılarak tüm dalgacık tabanlarında Dirac performans analizi.

Dalgacık Türü	PSNR-Y (dB)	SSIM	EFS (kB)
Le Gall 5/3	30,709	0,91227	521
Daub. 9/7	31,64	0,92097	572
DD 9/7	30,59	0,90933	517
DD 13/7	30,814	0,91165	521
Bior. 5/3	31,79	0,9286	664
Haar 0	31,266	0,9164	579
Haar 1	31,49	0,91730	579

Son olarak çizelgelerde Haar dalgacığının performansı da etkileyici gözükabilir ancak bu dalgacık daha çok küçük boyutlarda video dizileri için tercih edilmelidir zira burada esas alınan PSNR-Y metriği kullanıcı için aldatıcı olabilir. Zaten genel olarak 0.9 ile 1 arasında bulunan SSIM değeri (iyi bir video kodlayıcı için kodlama sonucunda) eğer dikkate alınacak olursa Çizelge 4.8'de de görülebileceği üzere en iyi sonuca göre önemli bir görsel kayıptan söz edebiliriz ki bu kaybın artan video boyutu ile birlikte büyümesi beklenebilir.

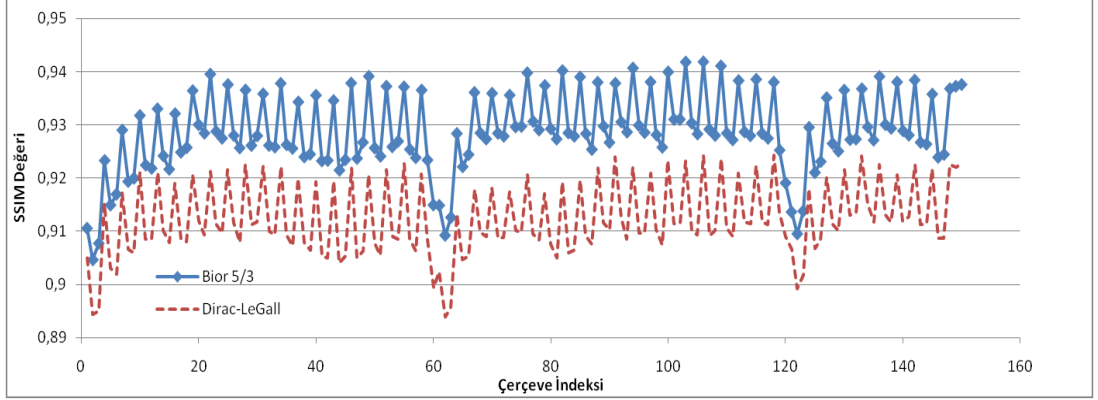
Çizelge 4.8 : 'Harbour.4cif' için Bior. 5/3 dalgacığının da katılarak tüm dalgacık tabanlarında Dirac performans analizi.

Dalgacık Türü	PSNR-Y (dB)	SSIM	EFS (MB)
Le Gall 5/3	31,07	0,912	3,19
Daub. 9/7	32,21	0,925	3,69
DD 9/7	31,46	0,918	3,13
DD 13/7	31,62	0,919	3,20
Bior. 5/3	31,94	0,924	4,48
Haar 0	30,83	0,9	3,82
Haar 1	31	0,9	3,81

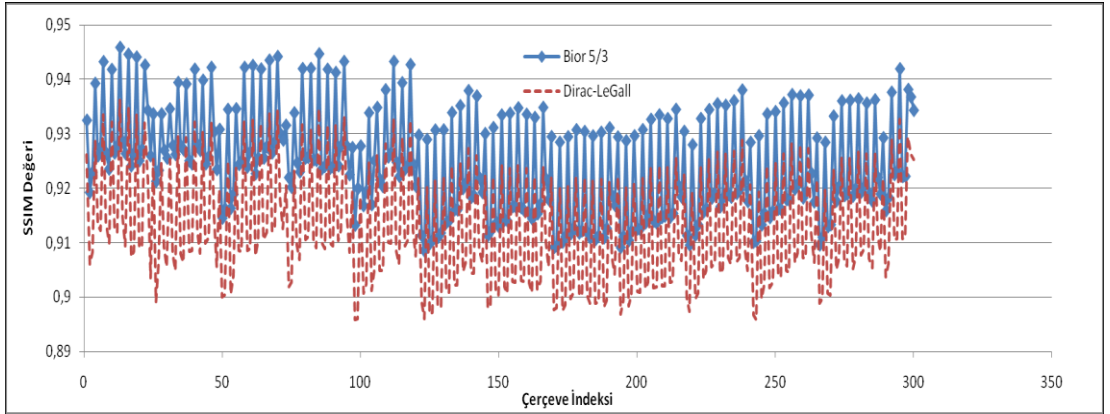


Şekil 4.4 : Yeni süzgeç yapısı (Bior. 5/3) ile Dirac-Le Gall süzgecin karşılaştırılması (Carphone. QCIF).

Şekil 4.4, Şekil 4.5 ve Şekil 4.6 'den çıkarılan sonuçlara göre girilen yeni 5/3 süzgecin VBR'de daha iyi performans sağladığı açıktır. Bu yeni yapının PSNR metriğine dayalı analizlere kıyasla SSIM analizlerinde çok daha iyi sonuç görüldüğünü söylemek gerekir ancak daha sonra da bahsedileceği üzere bu yeni süzgecin en büyük dezavantajı kodlanmış dosyanın boyutunun büyümesine sebep olmasıdır.



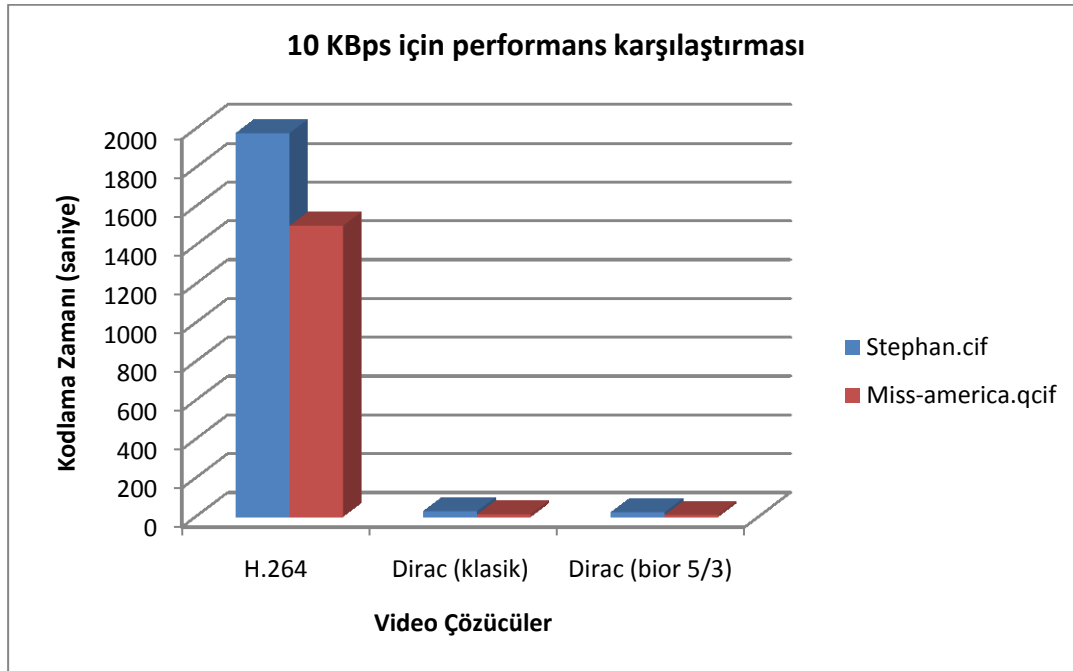
Şekil 4.5 : Yeni süzgeç yapısı (Bior. 5/3) ile Le Gall süzgecin karşılaştırılması (Bus. CIF).



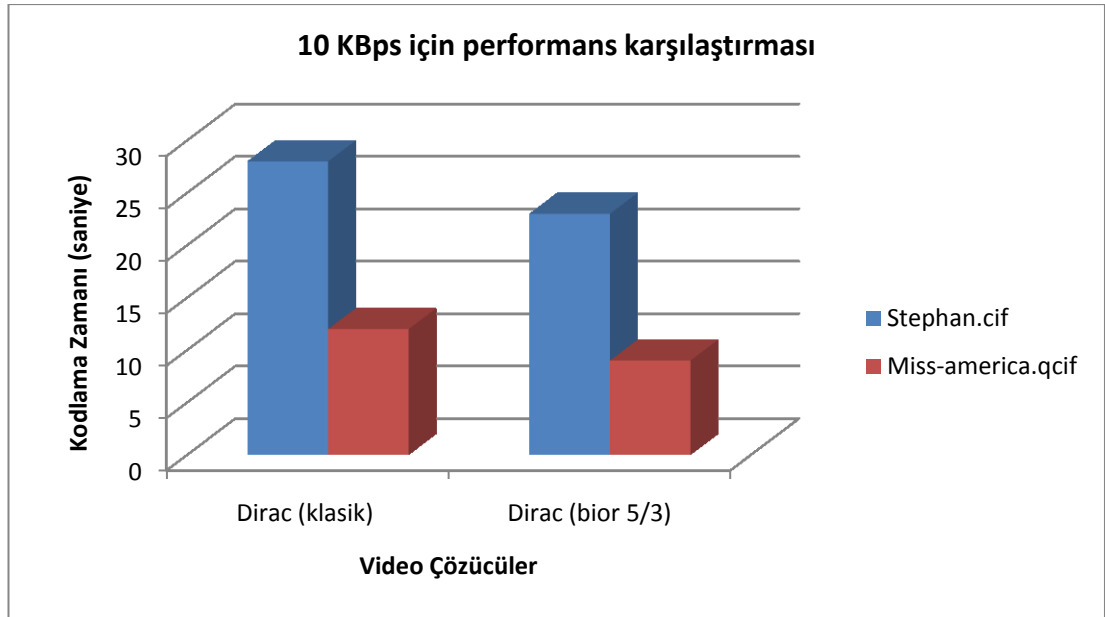
Şekil 4.6 : Yeni süzgeç yapısı (Bior. 5/3) ile Le Gall süzgecin karşılaştırılması (Harbour. 4CIF).

4.4 Bior 5/3 Tabanında Dirac Performansının H.264 ile Farklı Açılardan Kıyaslanması

Bu kısımda Dirac video kodlayıcının özellikle düşük bit oranlarındaki performansını irdelemek ve günümüzde en önde gelen video çözücülerden biri olan H.264 ile bir kıyaslama yapmak için bir takım gözlemler sunulmuştur. Öncelikle kodlama ve kod çözme zamanı açısından bir kıyas yapılırsa Dirac'ın H.264'den oldukça üstün olduğu görülecektir ki buradaki çalışmalara benzer anlamda çalışmalar [1,2]'de de yapılmıştır. Ancak burada H.264'ün test sürümü kullanıldığı hatırlatılmalıdır. Eğer optimize edilmiş sürümler tercih edilirse aşağıdaki grafiklerde görünen zaman farkının yine Dirac lehinde olmakla birlikte önemli ölçüde azaltılabilmesi söz konusudur. Bu noktada 10KB/s kodlanmış veri hızlarında zamanlama açısından bir kıyas yapılmıştır. Burada Şekil 4.7 ve Şekil 4.8'de ortaya çıkan zaman farkı diğer bit oranlarında da yaklaşık olarak korunmaktadır. Dolayısıyla bir örnek teşkil etmesi ve düşük bit oranlarında Dirac'ın performansını yansıtmaları açısından bu grafikler burada sunulmuştur. Ayrıca burada esas önemli nokta bu tezde sunulan yeni yapının (bior 5/3) Dirac'a getirdiği kodlama ve kod çözme zamanı açısından artırlardır. Bu kısımda QCIF ve CIF videolardan yararlanılmıştır zira performans kıyaslaması noktasında bu videolar hem Dirac için hem de genel anlamda kullanıcılar için yönlendirici olabilecek niteliktedir. [44]

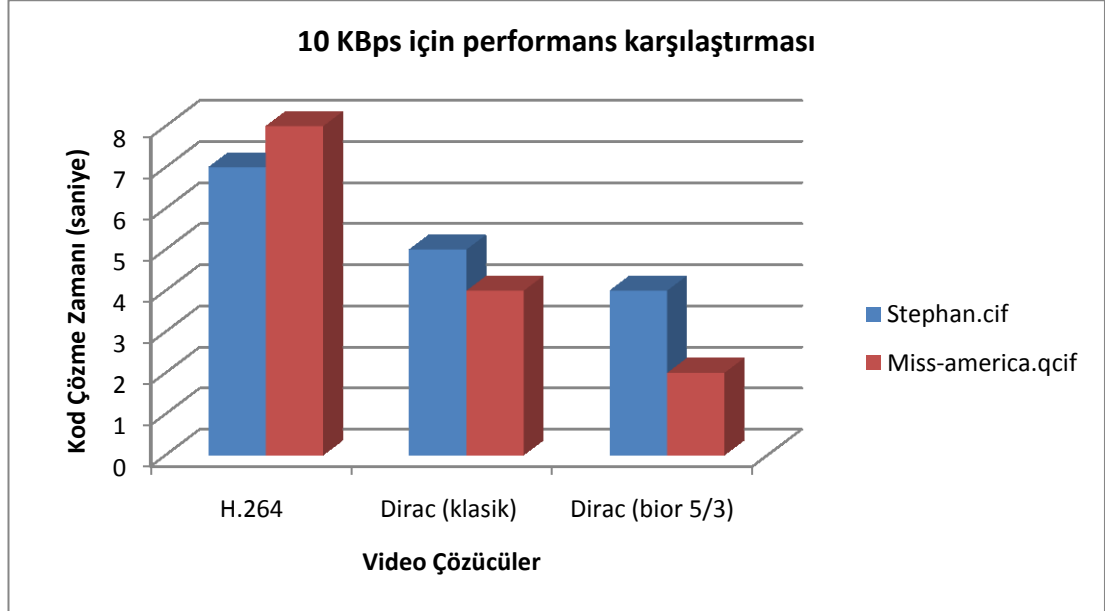


Şekil 4.7 : Dirac (bior 5/3) için performans kıyaslaması (kodlama zamanı).

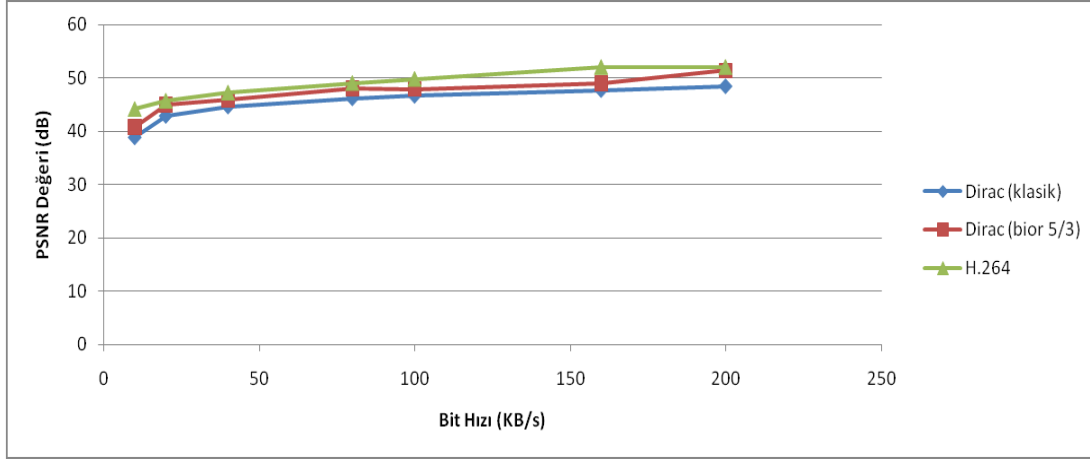


Şekil 4.8 : Dirac (bior 5/3) için performans kıyaslaması (kodlama zamanı).

Burada kod çözme açısından çok önemli bir fark yoktur ancak özellikle Şekil 4.7'deki kodlama zamanındaki H.264 ile Dirac ve yeni Dirac (bior 5/3) arasındaki fark aşikârdır ve pek çok uygulama için kritik bir parametre olarak görülebilir.

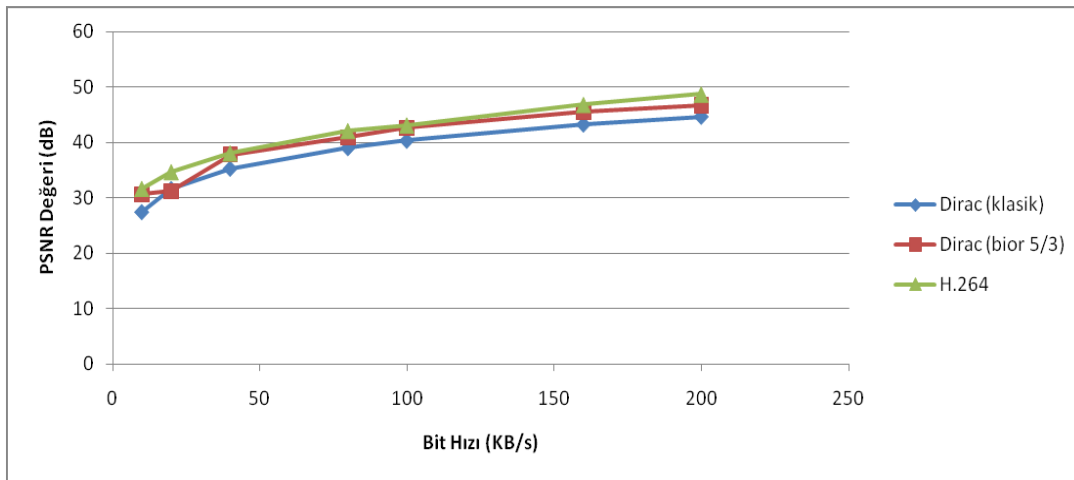


Şekil 4.9 : Dirac (bior 5/3) için performans kıyaslaması (kod çözme zamanı).

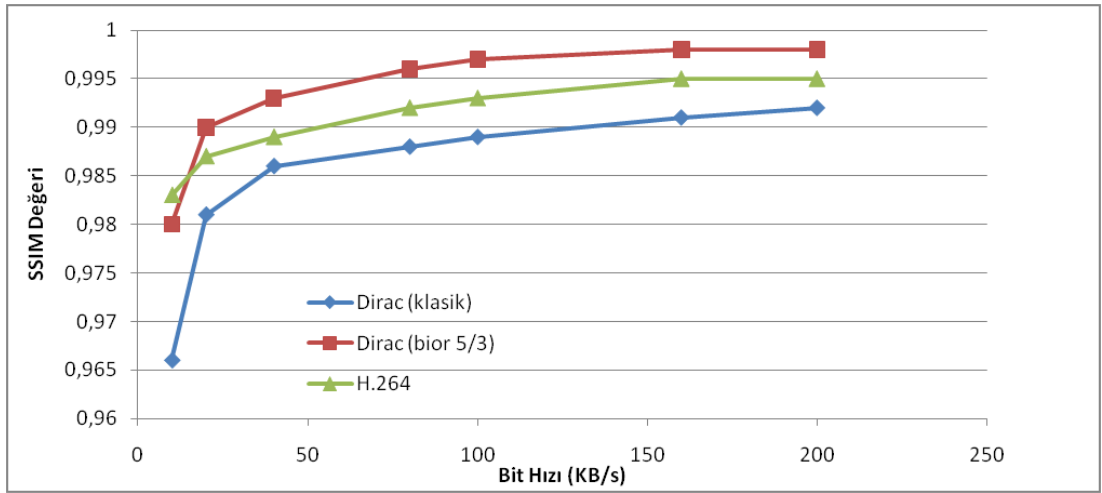


Şekil 4.10 : Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Miss-america.QCIF (PSNR-Y).

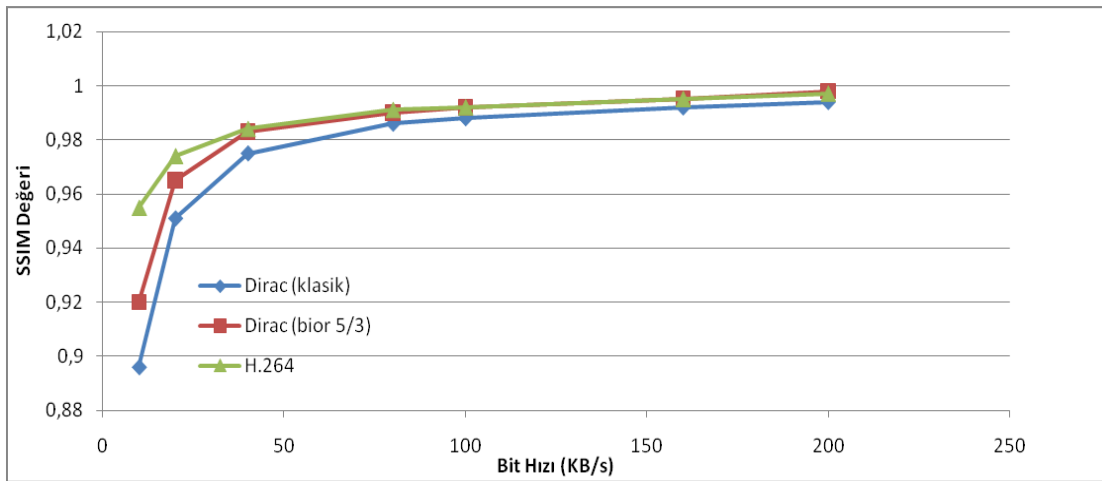
Şüphesiz PSNR-Y ve SSIM açısından yeni bior 5/3 dalgacık anlayışının Dirac performansına yaptığı katkılar Şekil 4.10, Şekil 4.11, Şekil 4.12 ve Şekil 4.13 den de görüleceği üzere sabit bit oranlarında da azımsanamayacak niteliktedir. Normalde Dirac H.264 performansından özellikle SSIM'de belirgin şekilde uzak iken yeni önerilen yapı ile H.264'ün yakalanması hatta genel performans açısından geçilmesi söz konusudur. Bu noktada Şekil 4.12 ve Şekil 4.13 irdelenebilir. Ayrıca tezin çeşitli yerlerinde bahsedildiği üzere Şekil 4.10, Şekil 4.11, Şekil 4.12 ve Şekil 4.13'de de gözlemlenebileceği gibi Dirac düşük bit oranlarında [1,2] çeşitli analizler sonucunda zaten MPEG-2'den üstün bulunmuştur. Burada da H.264 ile arasında özellikle yeni yapıyla birlikte neredeyse fark kapanmıştır. Burada özellikle bu anlamda [1] ve [2]'den çok daha iyi sonuçlar gözlemlenmiştir.



Şekil 4.11 : Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Stephan.CIF (PSNR-Y).



Şekil 4.12 : Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Miss-america.QCIF (SSIM).



Şekil 4.13 : Sabit bit akış oranlarında yeni Dirac yapısı performans kıyaslaması- Stephan.CIF (SSIM).

5. SONUÇLAR ve ÖNERİLER

Dirac performansını iyileştirmek üzere zaten Dirac 'da var olan dalgacık setlerine yenilerinin eklenmesi amaçlanmıştır ve PSNR ile SSIM temelinde sonuçlar gözlemlenmiştir. Analizi çeşitlemek amacıyla farklı QCIF, CIF ve 4CIF video dizileri tercih edilmiştir. Burada genel olarak Dirac performansı farklı dalgacık tabanlarında irdelenmiş olup Bior. 5/3 yapısı da önerilerek Dirac için yeni bir dalgacık olarak sunulmuştur. Önerilen bu dalgacık Dirac 'default' dalgacığı olarak seçilerek Dirac video kodlayıcı da genel olarak performans iyileşmesi gözlemlenebilmiştir. Tezin başından itibaren öncelikle genel video çözücü çalışma prensibini anlamak için diğer video çözücüler hakkında da kullanıcıya geniş bir bilgi yelpazesi içeren ana hatlar ele alınmış daha sonra da Dirac çözücüsü ve temel farkları irdelenmiştir. Bior. 5/3 yapısının sunulmasıyla da analizler hakkında kullanıcıyı daha kolay yorum yapabilecek konuma getirebilmek amaçlanmıştır.

Sonuç olarak EFS dışında dezavantajı yok denilebilecek Dirac için yeni bir dalgacık yapısıyla Dirac dalgacık veri tabanı daha da güçlü kılınmıştır. Buradan sonra yapılabilecek şeylerin başında elbette Dirac'ı video çözücü dünyasında zirveye taşıyabilmek olmalıdır. Açık teknoloji denilen tarzda bir model olarak kullanıcılarla BBC tarafından paylaşılan bu video çözücünün en çok tercih edilen bir noktaya gelmesi şüphesiz ki kolay bir süreç değildir. İnternet teknolojilerinin bu denli ilerlediği ve kıran kırana bir rekabetin söz konusu olduğu bir alanda sadece bu tezde üstünde durulan dalgacık mekanizmasının değil belki daha bu mekanizma gibi pek çok argüman ve çalışma sekmesinin ileri seviyelere taşınması gerekmektedir. Ayrıca Dirac'ın da MPEG ailesi gibi ileri seviye bir ses çözücüsüne de ihtiyacı olduğu açıktır. Bu konuda [1]'de ve [2]'da pek çok öneriler sunulsa da ve yakın zamanda Dirac için bir ses kodlayıcısının da gerçekleşmiş olduğu söz konusu olsa da bu konuda da daha çok yol alınması gerektiği açıktır. Tüm bu bahsedilen geniş alanda Dirac için daha pek çok öneri ve çalışmalar oluşturulup önemli çalışmalar ortaya çıkarılabilir. Dirac bu konuda tüm albenisiyle video dünyası gezginlerinin ilgi ve gayretlerini beklemektedir.

KAYNAKLAR

- [1] **ArunaRavi**, August 2009: "Performance Analysis and Comparison of Dirac Video Codec Scheme", Master of Science in Electrical Engineering, The University of Texas at Arlington.
- [2] **T. Borer and T. Davies**, July 2005: "Dirac video compression using open technology", BBC EBU Technical Review.
- [3] **Url-1** <<http://dirac.sourceforge.net>>, alındığı tarih 05.03.2011.
- [4] **Hendrik Eeckhaut, Benjamin Schrauwen, Mark Christiaens, and Jan Van Campenhout**, August 2005: Speeding up Dirac's entropy coder. In Proceedings of the 5th WSEAS International Conference on Multimedia, Internet and Video Technologies, pp.120-125, Corfu, Greece.
- [5] **Ingrid Daubechies and Wim Sweldens**, 1997: "Factoring Wavelet Transforms Into Lifting Steps", *Applied and Computational Mathematics Transaction..*
- [6] **John Arnold, Michael Frater, Mark Pickering**, Digital Television Technology and Standards, *The University of New South Wales, ADFA Canberra, ACT, Australia, Wiley, 2007.*
- [7] **Dr. Tim Ferguson**, Digital video coding and compression Lecture Notes, School of Computer Science and Software Engineering, Monash University, AUSTRALIA.
- [8] **ISO/IEC 11172** : 'Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mbit/s'.
- [9] **ISO/IEC 13818** : 'Generic coding of moving pictures and associated audio (MPEG-2)'.
- [10] **Rao, K.R. and YIP, P.**, 'Discrete cosine transform: algorithms, advantages, applications' (Academic Press, 1990).
- [11] **Wells, N.D.**, 'Component codec standard for high-quality digital television', *Electronics & Communication Engineering Journal*, August 1992, Vol 4, (4), pp. 195-202.
- [12] **Url-2** <<http://wwwam.hhi.de/mpeg-video/>>, alındığı tarih 05.07.2011.
- [13] **Iain E. G. Richardson**, 2003: H.264 and MPEG-4 Video Compression, Video Coding for Next-generation Multimedia, Wiley.
- [14] **Jain, A.K.**, 'Fundamentals of digital image processing' (Prentice Hall, 1989).
- [15] **Url-3** <<http://AlgorithmDetails.htm>>, alındığı tarih 03.08.2011.

- [16] **Masahiro Iwahashi and Masahiro Ohnishi**, '*Analytical Evaluation of Integer DCT*' International Symposium on Communications and Information Technologies (ISCIT 2004) Sapporo, Japan, October 26 - 29, 2004.
- [17] **Url-4** <<http://DIRAC/motiondatabase/Hierarchicalestimation.mht>>, alındığı tarih 08.08.2011.
- [18] **Howard L. Resnikoff, Raymond O. Wells**, "*Wavelet Analysis: The Scalable Structure of Information*", Springer, 1998.
- [19] **Ambikairajah E.**, Lecture Notes on Wavelet Transforms, University of New South Wales, 2000.
- [20] **M. Vetterli and J. Kovacevic**, Wavelets and Subband Coding, (Prentice Hall: 1995).
- [21] **G. Strang and T. Nguyen**, *Wavelets and Filter Banks*, (Wellesley-Cambridge Press: 1996).
- [22] **Catherine Bénéteau, Caroline Haddad, David Ruch and Patrick J. Van Fleet**, University of South Florida, Lecture Notes, 2009.
- [23] **CCIR Recommendation**, 'Encoding parameters of digital television for studios', 601-1 XVIth Plenary Assembly Dubrovnik 1986, Vol. **XI**, Part 1, pp. 319-328.
- [24] **Url-5** <<http://www.bilsen.com/aic/>>, alındığı tarih 10.08.2011.
- [25] **Url-6** <[http:// Biorthogonal and legall filters\ biorfilters\FilterSynthesis.htm](http://Biorthogonal%20and%20legall%20filters/biorfilters/FilterSynthesis.htm)>, alındığı tarih 01.08.2011.
- [26] **P. P. Vaidyanathan**, "*Multirate systems and filter banks*," pp. 457-538 1992.
- [27] **JPEG 2000 Image Coding System**, JPEG 2000 Final Committee Draft Version 1.0, 16th Mar., 2000.
- [28] **Sue Twelves, Mike Wu, and Andrew White**, JPEG2000 Wavelet Transform Using StarCore Application Note.
- [29] **H. S. Malvar, D. H. Staelin**, "The LOT: transform coding without blocking effects," IEEE Transactions on Acoustics, Speech, and Signal Processing, Vol. **37**, No. 4, pp. 553-559, April 1989.
- [31] **P. Ogunbona, I. Kharitonenko, and P. John**, JPEG2000 – The New Wave in Image Compression, (CCTV Magazine, pp. 36-38, Jan. 2000).
- [32] **J. M. Shapiro**, "Embedded image coding using zerotrees of wavelet coefficients," IEEE Transactions on SignalProcessing," Vol. **41**, No. 12, pp. 3445-3462, Dec. 1993.
- [33] **Url-7** <[DIRAC/Binarization and context modeling/wlt_coeff_coding.xht](http://DIRAC/Binarization%20and%20context%20modeling/wlt_coeff_coding.xht)>, alındığı tarih 10.05.2011.
- [34] **Patrick J. Van Fleet**, Center for Applied Mathematics, University of St. Thomas St. Paul, MN USA, 2007.
- [35] **Taubman, Marcellin**, Sections 4.2, 6.1-6.4, 7.

- [36] **André Filgueiras de Araujo, Yuzo Iano**, On the Performance of Promising Dirac Video Codec, Departamento de Comunicações, Faculdade de Engenharia Elétrica e de Computação, Universidade Estadual de Campinas, Campinas, Brazil.
- [37] **Jie Liang**, Lecture Notes, Engineering Science, Simon Fraser University, 2005.
- [38] **Tilo Strutz**, “Lifting Parameterisation of the 9/7 Wavelet Filterbank and its Application in Lossless Image Compression”, Proceedings of the 8th WSEAS International Conference on signal processing, robotics and automation.
- [39] **Url-8** <<http://iphome.hhi.de/suehring/tml/>>, alındığı tarih 08.08.2011.
- [40] **Url-9** <http://compression.ru/video/quality_measure/video_measurement_tool_en.html>, alındığı tarih 05.05.2011.
- [41] **Url-10** <<http://trace.eas.asu.edu/yuv>>, alındığı tarih 01.07.2011.
- [42] **Barua, S., Kotteri, K.A. Bell, A.E.Carletta, J.E**, 2004: “Optimal Quantized Lifting Coefficients for the 9/7 Wavelet”, Acoustics, Speech, and Signal Processing *Proceedings (ICASSP '04). IEEE International Conference*, Vol. 5, V - 193-6.
- [43] **Zhong Guangjun, Cheng Lizhi and Chen Huowang**, 2001: “A Simple 9/7 Tap Wavelet Filter Based on Lifting Scheme”, *IEEE International Image Processing Conference*, Vol. 2, pp. 249-252.
- [44] **K.V.Suchethan Swaroop and K.R.Rao**, IEEE Video Processing, Performance Analysis and Comparison of JM 15.1 and Intel IPP H.264 Encoder and Decoder.
- [45] **Url-10** <<http://en.wikipedia.org/wiki/YUV>>, alındığı tarih 25.07.2011.
- [46] **Url-11** <http://www.telescope-optics.net/eye_spectral_response.htm>, alındığı tarih 24.07.2011.
- [47] **Z. Wang, A. C. Bovik, H. R. Sheikh and E. P. Simoncelli**, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, Vol. 13, No. 4, pp. 600-612, Apr. 2004.

EKLER

EK A.1 : YUV Uzayı ve İnsan Gözüne Uyumluluğu

EK A.2 : SSIM ve Dirac

EK A.1

YUV Uzayı ve İnsan Gözüne Uyumluluğu

YUV renk uzayı, RGB renk uzayına alternatif olarak sunulmuş, yeni olmamakla birlikte pek çok açıdan gizemini halen koruyan nitelikte farklı bir platformdur. Burada 'Y' parlaklık, 'U' ve 'V' ise renk farklılıklarını temsil eden bileşenlerdir. İlk olarak siyah-beyaz görüntüleme altyapısı üzerine renkli TV ya da görüntüleme mekanizmasını nasıl kolaylıkla yerleştirebiliriz sorusunun cevabı olarak sunulan bu yapı 'Y' bileşeninin ağırlıklı olarak bilgi taşıdığı 'U' ve 'V' kısımlarının ise daha az önemsendiği bir organizasyon olarak dikkat çekmektedir. Bu özelliği sayesinde RGB uzayından farklı olarak, değişik örnekleme olanaklarını kullanıcıya sunarak kullanılan bant genişliğinden tasarrufa gitme fikri yeni bir organizasyondur ve sıkıştırma algoritmalarında da bu tür kombinezonlar dikkate alınır. Şekil A.1.1'de özellikle 'Y' bileşenine dikkat edilmelidir. Bu bileşen parlaklığın bir ölçüsü olarak biçimlendirilmiştir ve gözün en hassas tepki gösterdiği kısımdır. Denklem (A.1.1), (A.1.2), (A.1.3) ve denklem takımı (A.1.4), (A.1.5) ve (A.1.6)'ya bakıldığında YUV renk uzayı ile RGB renk uzayı arasında var olan sistematik dönüşüm görülmektedir [45,46]. Burada $k_b = 0.114$ ve $k_r = 0.299$ olarak alınmaktadır [45]. Bu katsayılar uluslararası telekomünikasyon birliği (ITU) tarafından önerilen değerlerdir.



Şekil A.1.1 : Örnek bir YUV video çerçevesi (En üstteki tüm YUV çerçevesi, ikinci satır soldaki 'Y', sağdaki 'U' ve en alttaki de 'V' bileşenleridir.) [45].

$$Y = k_r R + (1 - k_b - k_r)G + k_b B \quad (\text{A.1.1})$$

$$U = 0.5 (B - Y) / (1 - k_b) \quad (\text{A.1.2})$$

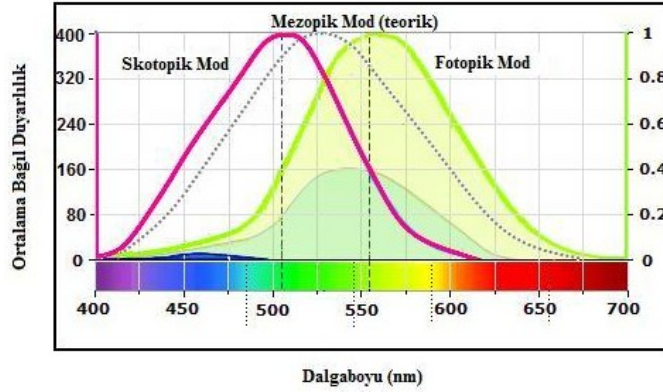
$$V = 0.5 (R - Y) / (1 - k_r). \quad (\text{A.1.3})$$

$$R = Y + U(1 - k_r) / 0.5 \quad (\text{A.1.4})$$

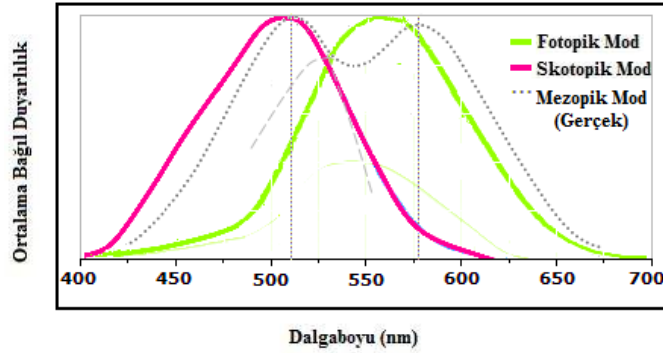
$$G = Y - 2k_b U(1 - k_b) / (1 - k_b - k_r) - U - 2k_r V(1 - k_r) / (1 - k_b - k_r) \quad (\text{A.1.5})$$

$$B = Y + U(1 - k_r) / 0.5. \quad (\text{A.1.6})$$

Bu noktada dikkat edilecek en önemli nokta, belirtilen bu katsayılar dikkate alındığında gözün en çok hassasiyet gösterdiği parlaklık değeri olan 'Y' nin belirlenmesi için 'G' bileşeninin yani 'Yeşil' rengin en belirgin rolü üstlendiği görülmektedir. Zira gözün elektromanyetik spektrum açısından en çok duyarlılık gösterdiği renk yeşildir. 'de insan gözünün farklı dalga boylarındaki çalışma mekanizması genel olarak gözlemlenebilir. Burada üç farklı çalışma evresi dikkat çekmektedir ve parabolik eğriler gözün tepki fonksiyonlarını yansıtmaktadır. İlk olarak 'Skotopik Evre' ye değinecek olursak bu aşamada, çubuk hücreler retina tabakasında egemen rol üstlenirler ve bu hücreler parlaklık ya da dış ortamın aydınlıklılık oranına karşı koni hücrelerinden çok daha az hassastırlar [46]. İkinci evre olan 'Mezopik Evre' ise bu iki ayrı hücre gruplarının eşit oranda görev üstlendiği ve etkidiği evre olarak tanımlanabilir [46]. Son evre olan 'Fotopik Evre' Skotopik evrenin tam tersi bir mekanizmada çalışıldığı durumdur [46]. Şekil A.1.3 'deki grafikte dikkat çeken nokta burada Mezopik evrenin gerçek insan gözünden alınan verilerle oluşturulan tepki fonksiyonunun verililişidir. Zira üstteki grafik teorisel Mezopik evreyi yansıtmaktadır ve idealde olması gerekeni sunar. Bu iki grafiği de yorumlarsak neden eşitlik (A.1.1), (A.1.2) ve (A.1.3) 'de 'G' bileşenine daha çok ağırlık verildiği anlaşılabacaktır çünkü gözün elektromanyetik açıdan en hassas olduğu renk bileşeni yeşildir [46].

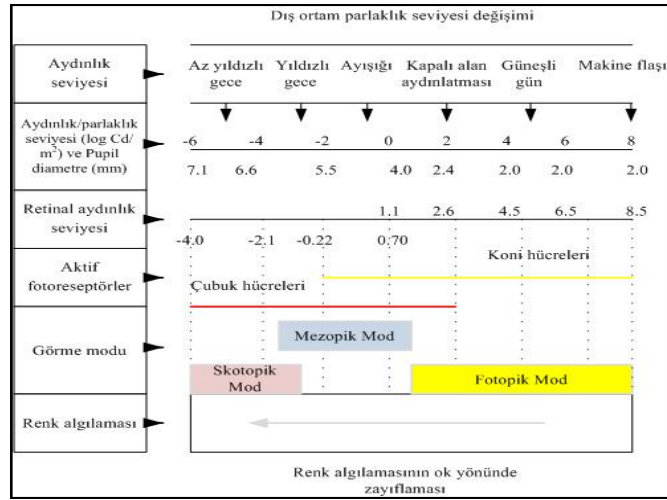


Şekil A.1.2 : Elektromanyetik spektruma göre insan gözünün çalışma evreleri ve farklı dalga boylarına verdiği tepki-1 [46].



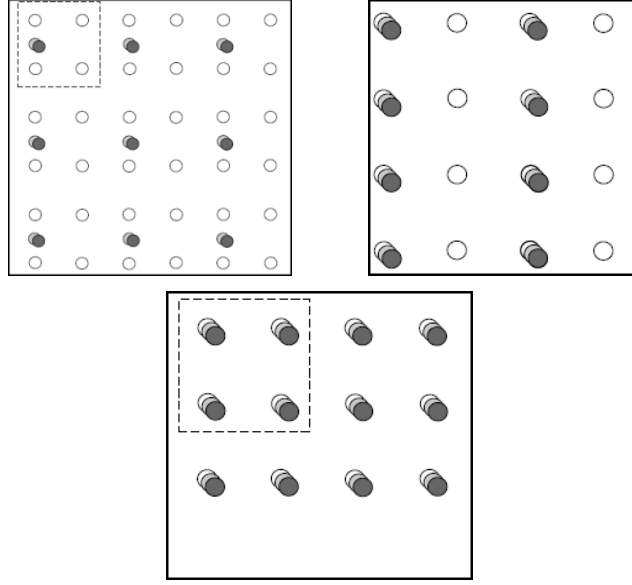
Şekil A.1.3 : Elektromanyetik spektruma göre insan gözünün çalışma evreleri ve farklı dalga boylarına verdiği tepki-2 [46].

Şekil A.1.4'de insan gözünün biraz daha detaylı analizine gidilmiştir. Bu noktadan daha ileriye gidilmeyecektir zira el aldığımız konunun temelleri burada her şeyi ile anlatılamayacak kadar derinlik arz etmektedir. Burada özellikle dikkat edilecek nokta koni hücrelerinin ve çubuk hücrelerinin retinadaki etkinliğinin değişiminin gözlemlenmesi ve retina tabakasının elektromanyetik etkiye verdiği tepkidir. Şekil A.1.2 ve Şekil A.1.3 de dikkate alındığında parlaklık olarak tanımlanan bileşenin en önemli etmeni olan yeşil renk değerinin değişimi ve dolayısıyla spektrumda ilerleme söz konusudur. Bu da gözün görme modunu değiştirecek temel mekanizmayı işletmektedir.



Şekil A.1.4 : Dış ortam parlaklık seviyesine göre insan gözünün verdiği tepkinin incelenmesi [46].

Yukarıda belirtilenler dikkate alındığında YUV uzayının bir önemli artısı daha öne çıkmaktadır. Bu uzay kullanılarak video işleme mekanizması yürütüldüğünde RGB uzayına göre ciddi anlamda sistem bant genişliğinden tasarruf etmek mümkündür. Çünkü insan gözünün bahsedildiği üzere parlaklık değerine karşı daha duyarlı olması diğer 'U' ve 'V' bileşenlerinin daha az örneklenebilmesine olanak verir ki bu da normalde ve en çok kullanılan 24 bitlik RGB uzayından 16 ya da 12 bitlik banda geçiş imkanı demektir [46]. Özellikle 4:2:0 uzayı denilen ve 'U' ve 'V' bileşenlerinin 'Y' ye göre dörtte bir oranında örneklendiği kombinezon bant genişliğinden yarı yarıya tasarruf edilmesini sağlar ki bu çok kritik bir öneme haizdir [6]. Şekil A.1.5 dikkate alınırsa YUV uzayından yararlanılarak video işlemede kullanılan temel örnekleme kombinezonları görülecektir. Burada 4:2:0 en çok tercih edilen ve Dirac'da da kullanılan tarzıdır ve bant genişliğini yarı yarıya kullanıcıya kazandırır. Ancak bu avantajı daha belirgin kılacak olan şey görüntü metriğinin de ona göre seçilmesidir, zira PSNR gibi klasik metrikler insan gözünün YUV uzayına bağlı niteliklerini ayırt edemediğinden görsel kalite kötüleşir ve işte bu noktada da bir sonraki bölümde bahsedileceği Dirac buna önlem olarak SSIM'i seçmiş ve kaliteden ödün vermemiştir.



Şekil A.1.5 : YUV uzayı farklı örneklenme biçimleri (En üst sol: 4:2:0, sağ: 4:2:2, en alt: 4:4:4 örnekleme. 4:2:0 ve 4:4:4 'de kesikli çerçeveye alınmış kısımlar 4 pikseli göstermektedir.) [45].

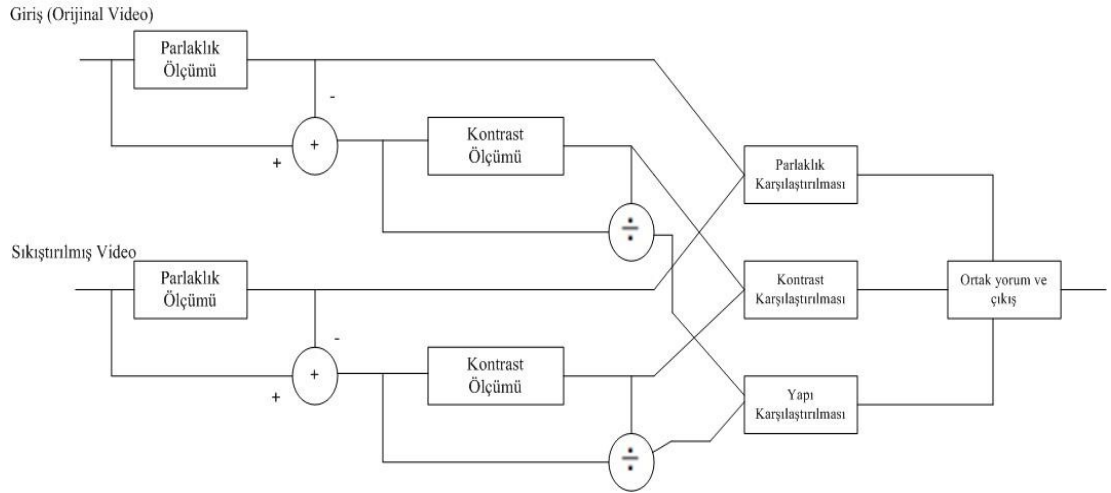
Bu bölümde YUV uzayına geçilmesinin ne kadar önemli olduğu ve getirdiği avantajların özetle de olsa taşıdığı amacı arz edilmeye çalışılmıştır. Burada sunulanlar bir sonraki bölümde belirli özellikleri üstünde durulacak olan Dirac video kodlayıcının SSIM mantığına dayalı yapısını anlamada önem taşımaktadır. Bahsedilen YUV uzayının elektromanyetik açıdan göze uyumluluk avantajını daha da hissedilir kılan SSIM anlayışı Dirac'ın da getirdiklerini kavramaya yardım edecektir.

EK A.2

SSIM ve Dirac

SSIM mekanizması Şekil A.2.1'da görülmektedir. Mekanizmanın çalışması özetlenirse, öncelikle ortalama piksel değeri tespit edilerek orijinal ve sıkıştırılmış video dizilerinden ayrı ayrı çıkarılır. Kalan sinyallerinin standart saçılmaları belirlenir ve parlaklık ölçümleri alınmış sinyaller bu saçılma değerine bölünerek normalize edilir [47]. En son yapı video sinyalinin iskeletini oluşturduğundan bu kısım yapı analizinin yapıldığı bölmeye gönderilir. Bu ölçüm sisteminde en önemli nokta parlaklık, zıtlık (kontrast) ve en sondaki temel yapı değerlendirmesinin birbirinden bağımsız değerlendirilmesiyle bu değerlerdeki olası değişimlerin birbirlerini etkilemesini önlemektir. Zira PSNR gibi klasik değerlendirmelerde bütün etkileri ayrıştırılmadan hep birlikte göz önüne alınır ve özellikle nazara verilen 'Y' değerinin etkisi tam olarak değerlendirilememekle birlikte zıtlık ve yapı analizleri de gözün olağan tepkisine uygun sonuçlandırılmaz.

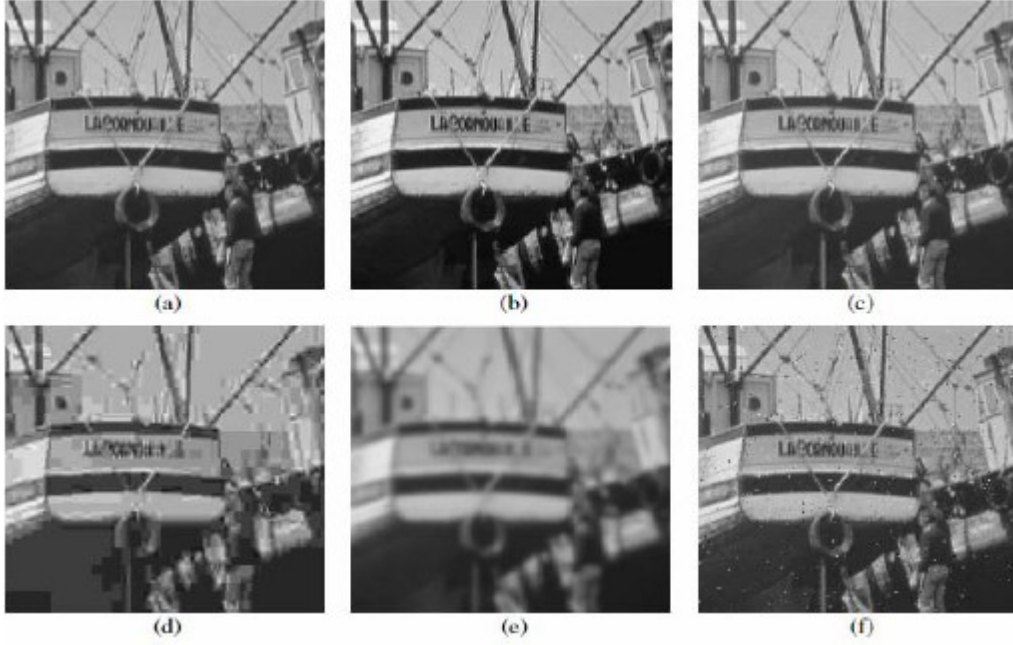
$$SSIM = [(2\mu_x\mu_y+c_1)(2\sigma_{xy}+c_2)] / [(\mu_x^2+\mu_y^2+c_1)(\sigma_x^2+\sigma_y^2+c_2)] \quad (\text{A.2.1})$$



Şekil A.2.1 : Dirac SSIM mekanizması [1,47].

Eşitlik (A.2.7)'de görüldüğü üzere SSIM ortalama ve varyans değerler üzerine çalışır ve bu haliyle bu yapı sadece luma (parlaklık) birleşenlerine uygulanır ki gözün elektromanyetik açıdan en hassas olduğu parça üzerinden çalışan bu metrik Dirac'ı da diğer video kodlayıcılardan basitliğine rağmen öne çıkarır. SSIM değeri sırasıyla minimum ve maksimum olmak üzere '-1' ve '1' arasında tüm değerleri alabilir. '0' değeri ilişkisizliği '-1' ters yönde yapısal ilişkiyi, '1' ise tam ilişkiyi gösterir [1,47].

Benzetim sonuçları bu bağlamda incelenmelidir. Ayrıca (3)'de ' σ_{xy} ' çapraz kovaryansı, ' c_1 ' ve ' c_2 ' ise sırasıyla ' k_1L ' ve ' k_2L ' olarak tanımlanırlar. Burada ' k_1 ' ve ' k_2 ' uluslar arası standartlarca sırasıyla 0.01 ve 0.03 olarak kabul edilirler. ' L ' ise dinamik piksel aralığı olarak kabul edilir ve örneğin 8 bitlik bir görüntüde 2^8 olmak üzere 255 olarak alınır [1]. Şekil A.2.2 SSIM'in PSNR ile farkını irdelemek için bu noktada incelenebilir.



Şekil A.2.2 : PSNR ile SSIM farkının irdelenmesi [1].

Burada (a) orijinal görüntüdür ('boat.pgm'). (b), (c), (d), (e) ve (f)'deki görüntüler ise çeşitli bozucu etkilerle distorsyona uğratılmışlardır. Bu noktada dikkat çeken nokta belirtilen tüm bu bozulmuş imajlarda MSE değerinin 210'a eşit olmasıdır. Ancak dikkat çekicidir ki göz burada belirgin bir şekilde farklı etkilerle bozulmuş görüntülerin birbirlerinden farklarına duyarlı bulunmaktadır ki bu da MSE'nin tüm imajlarda eşit çıkmasını görsel olarak anlamsız kılmaktadır zira örnek olarak (c) insan gözü açısından yapısal olarak orijinal imaja en yakın durumdadır. Aynı bozulmuş imajların ölçülmüş SSIM değerlerine bakılırsa b) SSIM = 0,9168, c) SSIM = 0,99, d) SSIM = 0,6949, e) SSIM = 0,7052 ve f) SSIM = 0,7748 bulunur. Bu sonuçlar da SSIM'in görsel olarak insan gözüne uyumluluğunu açık bir şekilde ortaya koyar.

ÖZGEÇMİŞ



Ad Soyad: Deniz Özenli

Doğum Yeri ve Tarihi: 22 Haziran 1988

Adres: Kartaltepe Mah., Koşuyolu Sk., Sungur Apt. Bakırköy/İSTANBUL

Lisans Üniversitesi: İstanbul Üniversitesi, Mühendislik Fakültesi, Elektrik-Elektronik Mühendisliği