

VIDEO MÜHENDİSLİĞİ / BÖLÜM 01

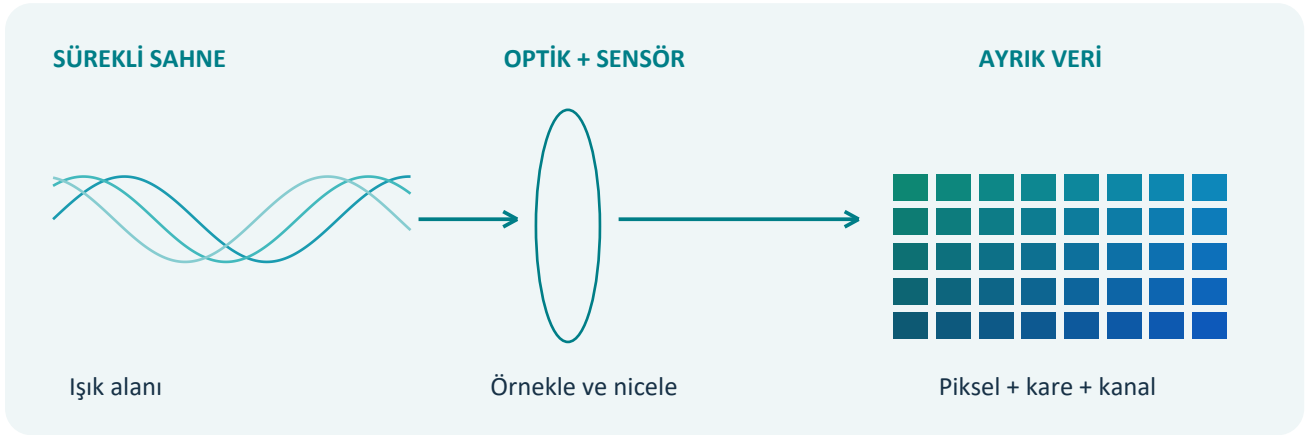
Video Nedir?

Işıktan Piksele

Işıktan FPGA'e • 20 bölümlük serinin ilk yazısı

Telefonunuzla bir video çektiğinizde aslında ne kaydediyorsunuz? Bir manzarayı mı, hareketi mi, yoksa arka arkaya dizilmiş fotoğrafları mı? Bugünkü yazımızda bu soruya mühendislik tarafından bakacağız. Merak etmeyin; önce fikri kuracak, matematiği gerektiği yerde kullanacağız.

Bir kamera, dünyayı olduğu gibi kutusuna koyamaz. Işığın içinden belirli bir bakış açısını seçer; bu görüntüyü konum, zaman ve renk bakımından sınırlı sayıda ölçüme dönüştürür. Yani elimizdeki ışık, oldukça zengin bir dünyayı sonlu sayıda bite sığdırmaktır.



Şekil 1. Işıktan sayılara: optik görüntü, sensörde örneklenir; ölçümler sayısal değerlere dönüştürülür. Özgün şema.

Peki, bu dönüşümde neyi seçiyoruz?

Objektif bakış açısını belirler. Sensördeki ölçüm noktaları uzamsal ayrıntıyı sınırlar. Kare hızı, zaman içindeki değişimi ne sıklıkla kaydettiğimizi söyler. Renk kanalları ve bit derinliği de her ölçümün nasıl temsil edileceğini belirler.

Başlangıç fikrimiz: Video, görüntünün uzayda ve zamanda örneklenmiş, değerleri nicelenmiş temsidir. Bir başka deyişle; nereden, ne zaman ve hangi hassasiyetle ölçüm yaptığımızın sonucudur.

Bu yazıda renk, hareket algısı, örnekleme ve veri hızını aynı zincirin parçaları olarak ele alacağız. Temel elektronik bilgisi yeterli; formüller ana fikrin önüne geçmeyecek.

01 / IŞIĞI SAYILARA ÇEVİRMEK

Dünya sürekli, hafızamız sınırlı

Bir sahnedeki ışığı yalnızca “şu pikselin parlaklığı” diye anlatamayız. Işık nereden geliyor, hangi yöne gidiyor, dalga boyu ne ve zaman içinde nasıl değişiyor? Bunların hepsi önemlidir.

Biraz matematik isteyenlere: $L(x, y, z, \theta, \phi, \lambda, t)$

Bu gösterim, üç konum, iki yön, dalga boyu ve zaman olmak üzere yedi değişkenli plenoptik fonksiyonu anlatır. Geleneksel 2B video, bunun seçilmiş bir görünümünü kaydeder.

Dünyadaki bilgi	Geleneksel videodaki karşılığı
Bakış yönü ve derinlik	Objektifin seçtiği perspektifte 2B görüntü
Dalga boyu dağılımı	Genellikle üç renk kanalıyla temsil
Görüntü düzlemindeki konum	Satır ve sütunlardan oluşan örnekleme ızgarası
Zaman içindeki değişim	Belirli aralıklarla alınan kareler

Örnekleme ile niceleme aynı şey mi?

Örnekleme, ölçümün konumunu ve zamanını seçmektir. **Niceleme** ise ölçülen değeri sonlu sayıdaki seviyeden biriyle temsil etmektir. Kareli bir kâğıda manzara çizdiğinizizi düşünün: karelerin sıklığı ayrıntıyı, kalem kutunuzdaki ton sayısı ise geçişlerin hassasiyetini sınırlar.

Sonuçta veriyi $V[m, n, k, c]$ gibi düşünebiliriz: satır, sütun, kare ve kanal. Buradaki dört indis, dört uzamsal boyut anlamına gelmez. Görüntü x ve y boyunca, zaman da t boyunca değişir; renk kanalı bunlara eklenen bir bileşendir.

Küçük bir ayrıntı daha: Pek çok renkli sensör, her noktada üç rengi birden ölçmez. Renk filtre dizisinden gelen ölçümler işlenerek tam renkli görüntü oluşturulur. Bu işi ISP bölümünde ayrıntılı konuşacağız.

02 / RENK

Neden üç renk kanalı işimize yarıyor?

Güneş ışığının spektrumu ile ekrandaki beyazın spektrumu aynı olmak zorunda değil. Buna rağmen ikisi de bize beyaz görünebilir. İlk bakışta garip geliyor, değil mi? Işın merkezinde gözümüzün rengi nasıl algıladığı var.

Normal trikromatik gündüz görmesinde S, M ve L koni sınıfları farklı dalga boyu aralıklarına duyarlıdır. Bu duyarlılıklar geniştir ve örtüşür. Yani gözümüzde üç dar bantlı “RGB ölçüm cihazı” bulunduğunu düşünmeyelim. Renk görme farklılıkları bu genel modelin dışındadır. [1]

RGB ve LMS aynı koordinat sistemi değildir.



Şekil 2. Ekranın RGB primerleri üç koni sınıfını birlikte uyarır. Şema bir duyarlılık grafiği değildir; renkler yalnızca anlatım amaçlıdır.

Farklı spektrum, benzer renk algısı

Belirli bir gözlem ve uyum koşulunda farklı spektrumlar aynı koni yanıtlarını üretebilir. Buna **metamerizm** diyoruz. Ekranın bütün ışık spektrumunu yeniden üretmek yerine görsel sistemimizde uygun yanıtı oluşturması, üç primerle renk üretiminin temelini açıklar.

Mühendislik karşılığı: Üç uygun primerle çok sayıda renk üretilebilir; ancak belirli bir ekranın renk gamı sınırlıdır. Doğadaki her rengi, her parlaklıkta birebir üretmesi beklenmez.

Kamera ile göz neden anlaşılamayabilir?

Kameranın filtreleri ve spektral duyarlılığı, insan konileriyle tam örtüşmez. Özellikle farklı LED aydınlatmalar altında, gözle doğal görünen bir yüz kamerada farklı çıkabilir. Bu yüzden renk meselesi yalnızca “R, G ve B değerlerini ayarladık” seviyesinde bitmez.

Serinin renk bilimi bölümünde primerleri, beyaz noktasını ve transfer fonksiyonlarını ayrı ayrı ele alacağız. Şimdilik üç kanalın işe yaramasının nedenini bilmemiz yeterli.

03 / HAREKET

Kareler neden hareket gibi görünür?

Çocukken defterin köşesine küçük çizimler yapıp sayfaları hızlıca çevirdiniz mi? Her sayfa sabittir, ama çizdiğiniz top hareket eder gibi görünür. Videoda da ardışık görüntüler, görsel sistemimizin hareketi işlemesine uygun bir dizi oluşturur.

Bunu yalnızca “görüntü retinada bir süre kalıyor” diye açıklamak eksik kalır. Hareket algısı, görüntünün konum ve zaman içindeki değişimini işleyen sinirsel mekanizmaları da içerir. **Titreme görmek** ile **hareketi akıcı görmek** farklı sorulardır. [2]

24 farklı kare / saniye



Aynı kareyi 3 kez sunmak



Tekrar gösterim, yeni hareket örneği eklemes. Şema zaman ölçekli değildir.

Şekil 3. Kare tekrarları ışığın sunum sıklığını artırabilir; kayıta bulunmayan yeni hareket konumları üretmez. Kavramsal gösterim. [2]

24 fps, her koşulda aynı deneyim demek mi?

Hayır. Kamera hareketi, nesnenin hızı, pozlama süresi, ekranın çalışma biçimi ve izleme koşulları sonucu değiştirir. Bu nedenle “şu kare hızının üstü kesin akıcıdır” şeklinde herkes için geçerli tek bir eşik vermek doğru olmaz.

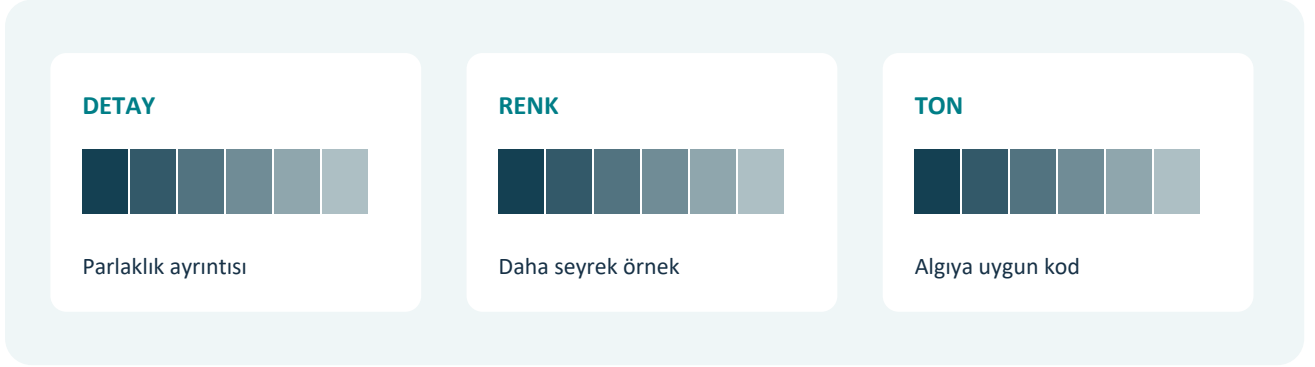
Klasik film projektörlerinde bir kare birden fazla ışık darbesiyle gösterilebilir. Örneğin 24 kare/s kaydı, kare başına üç ışık darbesiyle 72 darbe/s sunmak mümkündür. Yeni hareket bilgisi yine 24 kare/s düzeyindedir. Modern ekranların ışık üretme biçimi ise projektörle aynı olmak zorunda değildir.

Tasarımda akılda tutalım: Kare hızı, yenileme hızı ve ışığın modülasyon sıklığı birbirinin yerine kullanılmamalı. Aynı sayıyı yazıyor olmaları, aynı şeyi anlattıkları anlamına gelmez.

04 / ALGI

Her ayrıntıyı aynı hassasiyetle görmüyoruz

Bir yüzün hatlarındaki bozulmayı hemen fark ederken hareketli çimenlerin içindeki küçük bir hatayı gözden kaçırabiliriz. Görüntü kodlamasında hangi bilginin korunacağına karar verirken bu algısal farklar önemlidir.



Şekil 4. Üç tasarım kararı: parlaklık ayrıntısını korumak, renk örneklerini azaltmak ve tonları uygun bir kodlamayla temsil etmek. Kavramsal şema.

Parlaklık ayrıntısı ve renk ayrıntısı

Birçok izleme koşulunda ince parlaklık ayrıntılarına, ince renk ayrıntılarından daha duyarlıyız. Kroma alt örnekleme bu farktan yararlanır. 4:2:0 biçiminde, 4:4:4'e göre her bir kroma bileşeninin örnek sayısı dörtte bire iner. Bu, **kroma örneklerinin %75 azalmasıdır**; tüm görüntü verisinin %75 azalması değildir.

Aynı bit derinliğinde, alfa ve depolama boşlukları hariç toplam örnek yükü yarıya iner. Dosya boyutu aynı oranda küçülmek zorunda değildir. Küçük renkli yazılarda ve keskin renk sınırlarında kayıp belirginleşebilir. [6]

Tonlar ve kontrast

Kontrast duyarlılığı ayrıntının büyüklüğüne, parlaklığa ve izleme koşullarına bağlıdır. Çok ince ayrıntı her durumda görünmez değildir; izleme mesafesi de hesabın içindedir. Bu yüzden sıkıştırma "yüksek frekansları atalım" gibi tek başına çalışan bir kural yoktur.

Doğrusal olmayan kodlama, seviyeleri algıyla daha uyumlu dağıtmaya yardımcı olur. Luma (Y'), doğrusal olmayan RGB bileşenlerinden türetilen sinyaldir; fiziksel ışıklılık ve algılanan parlaklıkla aynı değildir. Gamma, PQ ve HLG de aynı fonksiyon değildir. [7, 8]

Maskeme: hatanın bulunduğu yer önemli

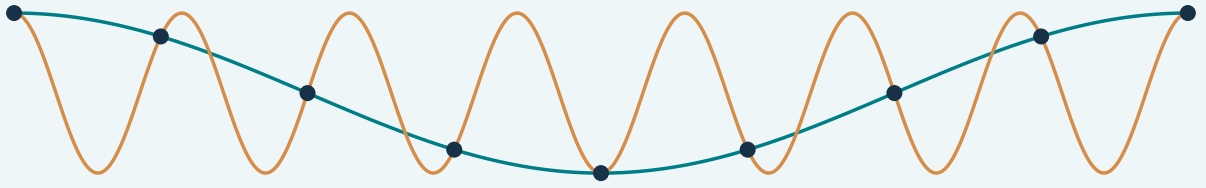
Bir desen, başka bir desenin fark edilmesini zorlaştırabilir; buna görsel maskeme denir. Etki, desenlerin frekansı ve kontrastı gibi özelliklere bağlıdır. Bu yüzden aynı hata farklı görüntü bölgelerinde farklı ölçüde fark edilebilir. [9]

05 / ÖRNEKLEME

Çizgili gömlek ile dönen tekerleğin ortak noktası

Bir çekimde ince çizgili gömleğin üzerinde dalgalar oluştuğunu görmüşsünüzdür. Ya da tekerleğin geriye dönüyormuş gibi görüldüğünü. İki örnekte de örnekleme, sahnedeki değişimi yeterince ayırt edemez. Ortaya **aliasing**, yani farklı sinyallerin aynı örneklerle temsil edilmesi çıkar. [3]

AYNI ÖRNEKLER, İKİ FARKLI SİNYAL



Noktalar örneklerdir. Örnekleme yetersizse hızlı değişim yavaş görünebilir.

Şekil 5. Turkuaz düşük frekanslı ve turuncu yüksek frekanslı eğri aynı siyah örnek noktalarından geçer. Bu örneklerden hangisinin asıl sinyal olduğunu tek başına çıkaramayız.

Temel koşul: Bant sınırlı bir sinyali ideal olarak yeniden kurabilmek için örnekleme frekansı, en yüksek sinyal frekansının iki katından büyük seçilir: $f_s > 2f_{\max}$. Gerçek filtrelerin geçiş bölgesi için de pay bırakılır. [3]

Videoda uzamsal örnekleme sensör üzerindeki piksel aralıklarıyla, zamansal örnekleme kare hızıyla ilişkilidir. Piksel saati bu örneklerin okunma veya taşınma hızını anlatır; tek başına yatay uzamsal örnekleme sıklığı değildir.

Peki, çözüm ne?

Örnekleme öncesinde ayırt edilemeyecek kadar hızlı değişimleri azaltırız. Optik alçak geçiren filtre, optik ve sensörün toplam cevabı; yeniden boyutlandırmada ise uygun sayısal filtreler bu işte rol oynar. OLPF kullanılmaması aliasing riskinin bittiği anlamına gelmez.

Pozlama süresi, ışığı zaman boyunca bütünleştirerek hareket bulanıklığı üretir ve bazı zamansal aliasing etkilerini azaltır. 24 fps'de 180° kapak açısı yaklaşık 1/48 saniyedir. Ama bu seçim, aliasing'i her sahnede önleyen ideal bir filtre değildir.

Interlace notu: Tek ve çift satır alanları farklı anlara aittir. Bu yüzden onları aynı anda çekilmiş tek bir kare gibi işlemek hareketli kenarlarda taraklanmaya yol açabilir. İkinci bölümde bu zaman farkına döneceğiz.

06 / SİNYAL ZİNCİRİ

Kameradan göze kadar uzun bir yol var

Video yalnızca kamerada oluşup ekrana gitmez. Arada renk dönüşümü, sıkıştırma, paketlenme ve aktarım gibi adımlar bulunabilir. Aşağıdaki zincir tipik bir örnek; her sistemin bütün blokları ayrı ayrı içermesi gerekmez.



Şekil 6. Tipik uçtan uca video zinciri. Sıkıştırılmamış SDI gibi bir yolda codec ve konteyner blokları atlanabilir. CFA: renk filtre dizisi; ADC: analog-sayısal dönüştürücü.

Bloklar çalışıyor; görüntü neden yanlış?

Her blok, aldığı verinin nasıl yorumlanacağını bilmek zorunda. 1920×1080 bilgisi tek başına yeterli değil. Renk matrisi, transfer karakteristiği, sinyal aralığı veya kroma konumu farklı yorumlandığında görüntü değişebilir.

Uyuşmazlık	Görülebilecek belirti
Full / limited aralığı	Yükselmiş siyahlar veya kırılan tonlar
BT.601 / BT.709 matrisi	Özellikle ten ve doygun renklerde kayma
SDR / HDR yorumu	Yanlış parlaklık ve ton görünümü
Ters interlace alan sırası	Harekette ileri-geri sıçrama hissi

07 / FORMAT

“1080p video” demek neden yetmiyor?

Size bir dosya verildi ve “1080p, sorunsuz çalışır” denildi. Güzel; ama saniyede kaç kare? Hangi renk kodlaması? Kaç bit? İki cihazın aynı dili konuşması için bu ayrıntıları netleştirmemiz gerekiyor.

Parametre	Örnek / açıklama	Seride
Aktif çözünürlük	1920 × 1080	2
Tarama tipi	Progressive veya interlaced	2
Kare / alan hızı	59,94 kare/s veya alan/s açık yazılmalı	5
Piksel / görüntü en-boy oranı	PAR 1:1 / DAR 16:9	2
Renk primerleri	BT.709	3
Transfer karakteristiği	Örn. BT.709; PQ / HLG ayrı tercihler	3
Matris katsayıları	BT.709 veya RGB için uygun tanım	4
Kroma formatı	4:4:4 / 4:2:2 / 4:2:0	4
Kroma örnek konumu	Örn. yatay eş konumlu	4
Nicemleme aralığı	Full veya limited	4
Bit derinliği	Bileşen başına 10 bit	4
Zamanlama / taşıma hızı	Arayüze ve video moduna bağlı	2, 6

Pratik yaklaşım: Hata ararken önce iki uçta bu tanımların eşleştiğini kontrol edin. Kabloyu değiştirmek bazen işe yarar; yanlış renk yorumunu ise tek başına düzeltemez.

Bu bilgiler dosyada, arayüzde veya sistem yapılandırmasında bulunabilir. FFmpeg, renk primerleri, transfer karakteristiği ve sinyal aralığını ayrı tanımlar; RFC 4175 ise sıkıştırılmamış IP video için örnekleme, boyut ve renk parametrelerini tarif eder. Hepsini tek bir “1080p” etiketiyle anlatamayız. [10, 11]

BT.1886 ekran tarafındaki bir EOTF tanımıdır; kaynak sinyalin transfer etiketiyle otomatik olarak eş anlamlı kabul edilmemelidir. Bu ayrımı renk bilimi bölümünde açacağız.

08 / VERİ HIZI

Bir saatlik video ne kadar yer kaplar?

Gelelim işin hafıza ve bant genişliği tarafına. 1920×1080 , 60 kare/s, bileşen başına 10 bit ve 4:2:2 örnekleme kullandığımızı düşünelim. Sıkıştırma ve ek taşıma yüklerini şimdilik dışarıda bırakıyoruz.

Adım adım hesap:

$1920 \times 1080 \times 60 = 124.416.000$ piksel/s

4:2:2 ve 10 bit: ortalama 20 bit/piksel

$124.416.000 \times 20 = \mathbf{2,48832 \text{ Gbit/s}}$

Bir saat: $2,48832 \times 10^9 \times 3600 / 8 = \mathbf{1,119744 \text{ TB}}$

BİR SAATLİK KAYIT • ONDALIK BİRİMLER

Ham aktif görüntü

1,12 TB

2,488 Gbit/s

H.265 örneği

4,5 GB

10 Mbit/s varsayımı

Şekil 7. Ham hesap boşluksuz paketlenmiş aktif örnek yüküdür; disk dosyası boyutu değildir. H.265 için 10 Mbit/s ortalama video hızı varsayılmıştır; ses ve ek yük hariçtir.

1080p60 için 3G-SDI nominal hat hızı 2,97 Gbit/s'tir. Tüm hat bitlerinin bir saatlik karşılığı yaklaşık 1,34 TB'dir. Aktif örnek yükü, hat hızı ve dosya boyutu farklıdır. Örneğin Y210, her 10 bitlik örneği 16 bitlik alanda saklar; dosya bu nedenle daha büyük olabilir. [13, 14]

10 Mbit/s ortalama video hızıyla bir saat yaklaşık 4,5 GB tutar; ses ve ek yük hariçtir. Aktif ham yüke göre oran yaklaşık 249'dur. Bu, eşit kaliteyle sağlanacak bir sıkıştırma vaadi değildir. TB ve GB burada ondalık birimlerdir.

Bu bölümden neyi cebimize koyalım?

Video; ışığı seçme, örnekleme, sayısallaştırma ve yeniden gösterme işidir. Renk, hareket, bant genişliği ve algı aynı tasarımın parçalarıdır. Birindeki karar diğerini etkiler. Sonraki yazıda piksel, çözünürlük ve tarama geometrisini açacağız. Piksel gerçekten her zaman kare midir? Oradan devam edelim.

KAYNAKLAR

Kaynaklar ve ileri okuma

[1] Stockman: Cone fundamentals and CIE standards

[Kaynağı aç](#)

[2] Hoffman, Karasev, Banks: Temporal presentation protocols

[Kaynağı aç](#)

[3] Stanford CS248A: Sampling and Anti-aliasing

[Kaynağı aç](#)

[4] Adelson ve Bergen: The Plenoptic Function

[Kaynağı aç](#)

[5] Basler: PGI ve demosaicing

[Kaynağı aç](#)

[6] Microsoft: 8-bit YUV formats

[Kaynağı aç](#)

[7] Poynton: Frequently Asked Questions About Color

[Kaynağı aç](#)

[8] ITU-R BT.2100

[Kaynağı aç](#)

[9] Heeger: Spatial Frequency Channels

[Kaynağı aç](#)

[10] FFmpeg: pixel format definitions

[Kaynağı aç](#)

[11] IETF RFC 4175: Uncompressed Video

[Kaynağı aç](#)

[12] ITU-R BT.1886: Reference display EOTF

[Kaynağı aç](#)

[13] Altera: 3-Gbps SDI Video

[Kaynağı aç](#)

[14] Microsoft: 10-bit and 16-bit YUV formats

[Kaynağı aç](#)

[15] Stanford: Sampling theory, aliasing

[Kaynağı aç](#)