



معرفی درس
موضوعات اصلی درس:

- ← معرفی چند رسانه ای
- ← موارد استفاده
- ← مرور مفاهیم موجود در این حوزه
- ← معرفی بلوک های اصلی چند رسانه ای
- ← نحوه تولید فایل های چند رسانه ای
- ← تکنولوژی ها و ابزارها و نرم افزارهای این حوزه

اهداف درس

- ایجاد مهارت‌های لازم در دانشجویان برای استفاده و تولید عناصر و محصولات چندرسانه‌ای، شناخت تکنولوژی‌ها و استانداردها و آگاهی از فاکتورهای موثر در کارایی سیستم‌های چندرسانه‌ای
- در پایان ترم دانشجو باید قادر باشد حداقل با یک نرم‌افزار در حوزه چندرسانه‌ای کار کند. ضمن اینکه باید با مفاهیم موجود در این حوزه و انواع فایل‌ها آشنایی داشته باشد.

ضرورت یادگیری سیستم های چند رسانه ای

- استفاده از چندرسانه‌ای در انتقال آسان تر مفاهیم
- آموزش
- تبلیغات
- شبیه‌سازی
- بازی‌ها
- صفحات وب



سرفصل مباحث



- آشنایی با مالتی مدیا
- کاربرد چندرسانه‌ای
- بلوک‌های اصلی چندرسانه‌ای:
 - متن
 - تصویر، گرافیک و رنگ
 - صوت
 - انیمیشن
 - ویدئو
- فشرده‌سازی
- سخت‌افزارها و نرم‌افزارها
- چندرسانه‌ای و اینترنت
- تولید چندرسانه‌ای
- ابزارهای تولید

منابع

- ➔ Vaughan, T. (2008) "Multimedia: Making It Work", 7th Edition, Osborne-McGraw Hill
- ➔ McGraw-Hill, G. (2010) "Introduction to Multimedia", McGraw-Hill
- ➔ Chapman, N. & Chapman, J. (2004) "Digital Multimedia", John Wiley and Sons

چند رسانه ای

چند رسانه ای (Multimedia) چیست؟

ترکیب متن، تصاویر ثابت و متحرک، و صوت با استفاده از تکنولوژی کامپیوتر

کلمات

نوشته، گفتگو، آواز

تصاویر

ترسیم، نمودار، طرح، نقاشی، عکس، دکمه، انیمیشن یا ویدئو

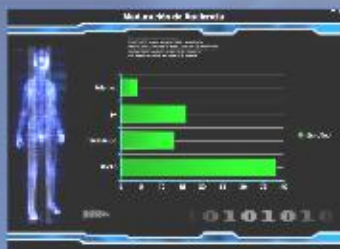
صوت (audio یا sound)

گویندگی (voice-over)، جلوه های صوتی، موسیقی



محصولات چند رسانه ای

- ارائه
- محتوای آموزشی
- بازی
- شبیه سازی
- صفحه وب







مباحث این جلسه

- تعریف چند رسانه ای و سیستم چند رسانه ای
- مزایا
- کاربردها
- مشخصات
- اجزا
- چالش ها



مقدمه

انسان معمولاً

۱۵٪ چیزی را که می شنوند

۲۵٪ چیزی را که می بینند

بیشتر از ۶۰٪ چیزی را که با آن تعامل دارند

به خاطر می سپارد



تعریف

- ترکیب متن، تصاویر ثابت و متحرک، و صوت با استفاده از تکنولوژی کامپیوتر
- توانایی ارائه اطلاعات از طریق صوت و ویدئو و انیمیشن همراه با رسانه های سنتی (نظیر متن، رسوم گرافیکی و تصاویر)
- یکپارچه سازی (توسط کامپیوتر) متن، گرافیک، ترسیم، تصاویر ثابت و متحرک، انیمیشن، صوت، و هر رسانه دیگری که در آن هر نوعی از اطلاعات می تواند به شکل دیجیتال-نمایش، ذخیره، منتقل و پردازش شود

مزایا

- انتقال داده در چندین فرمت به شکل همزمان
- امکان دسترسی به اطلاعات در بهترین حالت متناسب با سبک شناختی (cognitive style)
- ترجیح هر فرد برای نحوه دریافت اطلاعات
- افزایش بهره وری



تاریخچه

- روزنامه!
- استفاده از متن، گرافیک و تصویر
- قرن ۱۹ رادیو
- قرن ۲۰ تلویزیون
- دهه ۷۰ تولد اینترنت



سیستم چند رسانه ای

سیستمی که قادر به

پردازش

ذخیره سازی

تولید

دستکاری

تبدیل

داده ها و برنامه های چند رسانه ای است



مشخصات سیستم چند رسانه ای

● کنترل شده توسط کامپیوتر

● یکپارچه

● دیجیتال

● واسط کاربر تعاملی

● قدرت پردازشی بالا

● فایل سیستم قادر به کار با چند رسانه ای

● فرمت هایی که چند رسانه ای را پشتیبانی می کنند

● ورودی و خروجی کارآمد و سریع

● سیستم عامل

● پشتیبانی حافظه

● پشتیبانی شبکه

● ابزارهای نرم افزاری



اجزا

- وسایل ثبت و ضبط
- وسایل ذخیره سازی
- شبکه های ارتباطی
- سیستم های کامپیوتری
- وسایل نمایش



چالش ها

- نیاز به رندر کردن چند مدیا به شکل همزمان
- وابستگی زمانی داده ها به یکدیگر
- دیجیتال کردن داده ها
- حجم بالای اطلاعات چند رسانه ای





کاربردها

- WWW
- نرم افزارهای آموزشی چندرسانه ای
- ویدئو کنفرانس
- ویدئو بر حسب تقاضا (VoD)
- تلویزیون های تعاملی
- گروه افزار
- بازی
- واقعیت مجازی
- سیستم های تولید و ویرایش ویدئوی دیجیتال
- سیستم های پایگاه داده چندرسانه ای



مباحث این جلسه

- چند رسانه ای در اینترنت
- چند رسانه ای در تجارت
- چند رسانه ای در آموزش
- چند رسانه ای در خانه
- واقعیت مجازی
- بلوک های سازنده چند رسانه ای
- متن



چند رسانه ای در اینترنت

● توسعه اینترنت

تولد، وب، کاوشگرها

on-demand media

Internet radio, Internet video and satellite radio

● کاربرد:

تبلیغات

تجارت

آموزش

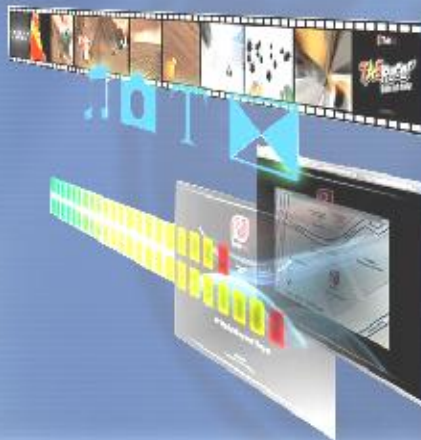
به اشتراک گذاری دانش

بازی ها و تفریحات دیگر

● نیازهای چند رسانه ای در اینترنت

سرعت

پهنای باند



چند رسانه ای در تجارت

● تجارت الکترونیک

Business-to-Business

Business-to-Consumer

تبلیغات

● کاربرد چند رسانه ای در محل کار

آموزش کارکنان

کار در منزل

تلکنفرانس

شبکه شرکت ها



چند رسانه ای در آموزش

- ◀ یادگیری از راه دور
- ◀ کلاس های مجازی
- ◀ یادگیری مبتنی بر کامپیوتر

• مزایای استفاده از چند رسانه ای در آموزش

- انتقال مطلب
- تعامل آموزنده با سیستم
- آزمایش و تست بر اساس شبیه سازی



چند رسانه ای در خانه

- اخبار
- مجلات اینترنتی
- کتاب های آنلاین
- توزیع موسیقی
- بازی ها
- تعاملات اجتماعی
- پزشکی از راه دور



واقعیت مجازی

- محیط های شبیه سازی شده بر اساس کامپیوتر که حضور فیزیکی در محیط مذکور را شبیه سازی می کنند
- یک محیط چند رسانه ای و سه بعدی
- ارتباط با سیستم از طریق وسایل ورودی نظیر ماوس، کیبورد، میکروفون، حسگرها
- کاربرد: آموزش، تست، تفریح و ...



واقعیت مجازی

توسعه از طریق:

نرم افزارهای CAD
سخت افزارهای تسریع گرافیک
head mounted displays
stereoscopic displays
haptic systems

مفاهیم واقعیت مجازی:

شبیه سازی، تعامل، تصنیع، Immersion، حضور از راه دور، ارتباطات شبکه



بلوک های سازنده چند رسانه ای

متن

نمونه هایی از متن

- صفحات وب
- زیر نویس ویدئو
- آموزش مبتنی بر کامپیوتر
- ارائه (presentation)
- چت
- اطلاعات ارائه شده در بازی ها



فرمت متن

← مفهوم: نحوه نمایش متن

← اجزا

فونت

رنگ و اندازه

تنظیم

فاصله

فرمت های پیشرفته: سایه، بالانگاشت (X^2)، زیروند (X^2)،
... ، watermark

تکنولوژی های فونت

← Bitmap

استفاده از مجموعه ای از نقاط یا پیکسل ها برای نمایش متن

کنترل بیشتر روی ظاهر متن

رندر شدن سریع و ساده

← مقیاس پذیر

کاراکترها به شکل فرمول های ریاضی ذخیره می شوند

نمایش فونت، به شکل outline و رنگ آن به شکل Bitmap

تغییر اندازه بدون از بین رفتن کیفیت

Postscript, TrueType, OpenType

| | | |
|--------------|-------|--------------|
| <i>a b c</i> | x y z | 1 2 3 |
| a b c | x y z | 1 2 3 |

مواردی در استفاده از فونت

- عدم استفاده از فونت های متفاوت در یک پروژه!
- استفاده از فونت هایی که در سیستم های مختلف وجود داشته باشند
- استفاده از فونت های Bitmap
- استفاده محدود از فونت های زیبا برای جلوه های ویژه و تاکید
- استفاده از گزینه های **Bold**، **Italic** و **Underline**
- استفاده از فونت، گزینه های سبک نگارش (**style**)، اندازه و رنگ
- انتخاب رنگ زمینه و فونت
- پرهیز از اشتباهات گرامری و املايي



سیستم های چند رسانه ای

متن و رنگ

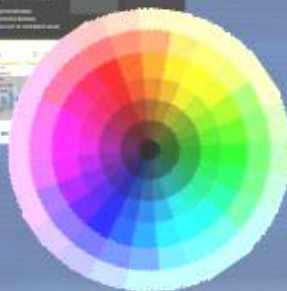
مباحث این جلسه

ذخیره سازی متن

فرمت فایل های متنی

متن در چند رسانه ای

رنگ



ذخیره سازی متن

← ذخیره سازی به شکل کاراکتر

هر کاراکتر، یک بایت
ذخیره سازی فرمت به شکل کاراکتر یا باینری

- محدودیت فضا برای ذخیره سازی مطرح نیست
- محدودیتی در ترتیب و زمان ذخیره سازی نیست، مگر محدودیت های معنایی نوع خاصی از فایل

فایل متنی

← فایل متنی، text file یا flatfile

← فرمت های فایل های متنی



.TXT
.RTF
.LOG
.DOCX



.TXT

- سند استاندارد متنی که شامل متن بدون فرمت است
- ذخیره سازی ساده تر
- فضای کمتر
- عدم بروز مشکلاتی چون:
Endianness
Padding bytes
- تصحیح ساده تر خطاهای فایل
- بی نظمی یا انتروپی پایین

| Windows | Linux |
|-----------------------------|---------|
| Microsoft Notepad | Leafpad |
| Microsoft WordPad | gedit |
| Microsoft Word 2010 | KWrite |
| Corel WordPerfect Office X5 | |
| Helios TextPad | |
| Notepad++ | |
| Notepad2 | |



.RTF

- Rich Text Format File
- پشتیبانی از Rich text
- متنی که دارای فرمت، فونت، رنگ و ... است
- قابلیت ویرایش توسط اکثر برنامه های پردازش متن
- ذخیره سازی:

This is some **bold** text.

```
{\rtf1\ansi{\fonttbl\font0\swiss Helvetica;}\pard  
This is some {\b bold} text.\par }
```

LOG



.LOG

- ثبت وقایع همراه با تاریخ وقوع آن ها
- ایجاد توسط سیستم عامل یا برنامه های نظارتی

Common Log Format (CLF)
Extended Common Log Format
W3C Standard

|  Windows |  Mac OS |
|--|---|
| Microsoft Notepad | Apple Console |
| | Apple TextEdit |

DOCX



.DOCX

- شامل متن، تصویر، فرمت متن، سبک، و سایر تنظیمات سند
- ذخیره سازی با استفاده از فرمت **Open XML** به شکل مجموعه ای از فایل ها و فولدرهای جداگانه در یک بسته فشرده **zip**.

فایل XML
docProps
Word
_rels

|  Windows |  MacOS |
|---|---|
| Microsoft Word 2010 | Microsoft Word 2011 |
| Microsoft Word Viewer with Office 2007 Compatibility Pack | Apple Pages |
| Microsoft Works 9 | Planamesa NeoOffice |
| Corel WordPerfect Office X5 | Panergy docXConverter |
| Panergy docXConverter | |

سایر فرمت ها

- .DOC .
- .EML .
- .LIT .
- .LST .
- .MSG .
- .ODT .
- .PAGES .
- .SIG .
- .TEX .
- .WPS .
- .WPD .

متن در چند رسانه ای

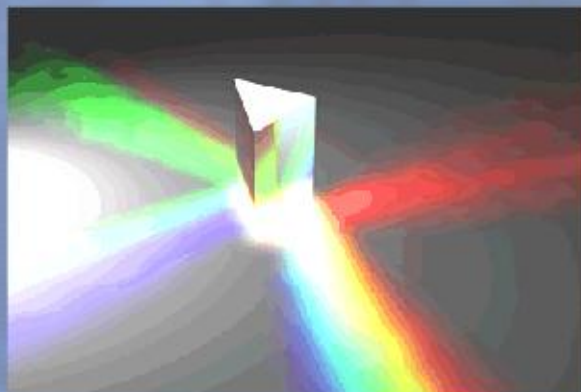
- کلمات معمول ترین روش برقراری ارتباط هستند
- متن روشی انعطاف پذیر برای ارائه اطلاعات است
- نباید متن جایگزین سایر اجزا شود، بلکه باید به شکل مکمل استفاده شود

بلوک های سازنده چند رسانه ای

رنگ

نور و طیف

- نور یک موج الکترومغناطیسی است
- رنگ نور با محتوای طول موج آن مشخص می شود
- نور مرئی، موج الکترومغناطیسی در محدوده ۴۰۰ تا ۷۰۰ نانومتر است



دید انسان

- چشم مانند یک دوربین کار می کند و شبکیه چشم نظیر لنز روی تصویر تمرکز می کند (به شکل وارونه)
- شبکیه شامل آرایه ای از میله ها و سه نوع مخروط است
- میله ها زمانی که میزان نور کم است، وارد عمل می شوند و یک تصویر در سایه های خاکستری تولید می کنند
- برای میزان نور بالاتر هر مخروط یک سیگنال تولید میکند. بدلیل متفاوت بودن رنگدانه هایشان، سه نوع مخروط به نورهای قرمز (R)، سبز (G) و آبی (B) حساس هستند

رنگ

- نور منعکس شده
- مشخصه هایی که به درک رنگ کمک می کنند:
- درخشندگی (Brightness): میزان نور دریافتی توسط چشم انسان از اشیا
- ته رنگ (Hue): رنگ غالبی که بیننده دریافت می کند
- پررنگی (Colorfulness): میزان کمتر یا بیشتر به چشم آمدن دو ناحیه با ته رنگ یکسان
- روشنایی (Lightness): میزان احساس درخشندگی یک ناحیه توسط چشم انسان نسبت به رنگ سفید مرجع
- رنگینی (Chroma): میزان پررنگی یک ناحیه نسبت به درخشندگی رنگ سفید مرجع
- اشباع (Saturation): میزان پررنگی یک ناحیه نسبت به درخشندگی آن ناحیه



دلیل تعریف فضاهای رنگ

- از لحاظ کارایی و پیچیدگی به کارگیری، کار کردن با مشخصه ی طول طیف رنگ ها، کاری است بسیار مشکل
- نیاز به پیدا کردن فضایی چند بعدی و تعریف مجدد رنگ ها در فضایی با ابعاد به مراتب کوچکتر و استفاده از حداکثر ابعاد قابل درک (یعنی ۳ بعد)
- تفاوت فضاهای رنگ با یکدیگر ناشی از تعاریف متفاوتی است که برای هر کدام از ابعاد و مرکز این ابعاد متعامد، در هر کدام از این فضاها ارائه شده است
- **فضای رنگ خطی**: تغییرات در یکی از ابعاد متناظر با همان میزان تغییر در مفهومی است که با آن بعد متناظر است

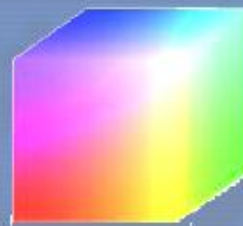
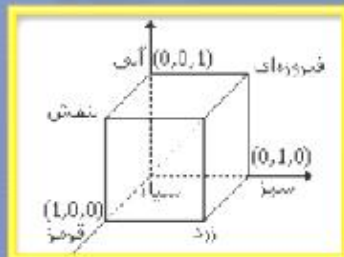
روش های مختلف نمایش فضای رنگ

- ➡ استفاده از ابعاد (محورک) پایه (T_1, T_2, T_3)
- ➡ استفاده از رنگینگی و درخشندگی (t_1, t_2, γ) :

$$t_1 = \frac{T_1}{T_1 + T_2 + T_3} \quad , \quad t_2 = \frac{T_2}{T_1 + T_2 + T_3}$$

فضای رنگ RGB

- یکی از اولین استانداردها
- قرمز (با طول طیف رنگ ۷۰۰ نانومتر)
- سبز (با طول طیف رنگ ۵۴۶/۱ نانومتر)
- آبی (با طول طیف رنگ ۴۳۵/۸ نانومتر)



فضای رنگ XYZ

- یک مجموعه رنگ پایه ی جدید برای برطرف کردن مشکلات فضای رنگ RGB
- برای مشاهده گرهای مختلف نظیر انسان، سنسور و ... تعریف واحدی برای دریافت اطلاعات رنگی موجود نیست
- میزان درخشندگی رنگ در این فضا برخلاف فضای قبل فقط به یک پارامتر (Y) بستگی دارد

تبدیل فضاهای رنگی RGB و XYZ به یکدیگر:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.412453 & 0.357580 & 0.180423 \\ 0.212671 & 0.715160 & 0.072169 \\ 0.019334 & 0.119193 & 0.950227 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \quad \begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 3.240479 & -1.537150 & -0.498535 \\ -0.969255 & 1.875992 & 0.041556 \\ 0.055640 & -0.204043 & 1.057311 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

فضای رنگ xyY

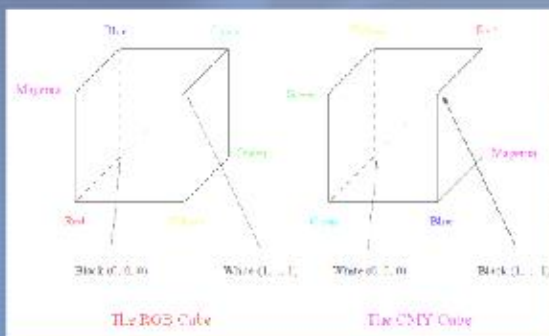
- از فضای XYZ مشتق شده است و نمایش کروما-لومای آن است
- یک فضای دوبعدی مستقل از شدت نور ارائه می دهد

Y: روشنایی

xy: رنگینگی

- دو مقدار رنگ که فقط در شدت نور متفاوت هستند، در این فضا دارای مقادیر یکسانی می باشند

فضای رنگ CMY



فیروزه ای (Cyan)
بنفش (Magenta)
زرد (Yellow)

$$\begin{bmatrix} C \\ M \\ Y \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} - \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

تبدیل RGB به CMY

فضای رنگ CMYK

جزء K یعنی رنگ سیاه به فضای CMY اضافه شده است

تبدیل از فضای CMYK به CMY:

$$C = \min(1, C(1-K)+K)$$

$$M = \min(1, M(1-K)+K)$$

$$Y = \min(1, Y(1-K)+K)$$

تبدیل از فضای CMYK به CMY:

$$K = \min(C, M, Y)$$

$$C = (C-K)/(1-K)$$

$$M = (M-K)/(1-K)$$

$$Y = (Y-K)/(1-K)$$

فضای رنگ YUV - YIQ

فضاهای رنگ انتقال تلویزیونی

YUV و YIQ: استاندارد تلویزیون های آنالوگ

YCbCr: استاندارد تلویزیون دیجیتال

اطلاعات رنگی به دو جزء روشنایی و رنگینگی تبدیل می شوند

● تطابق با تلویزیون های سیاه و سفید و کاربردهای فشرده سازی



فضای رنگ YUV - YIQ

- YUV استاندارد سیستم‌های NTSC
- YIQ استاندارد سیستم‌های PAL
- از روی مدل Gamm Corrected RGB (R'G'B') بدست آمده‌اند
- حساسیت سیستم بینایی انسان به تغییرات روشنایی بیشتر از حساسیت آن به ته رنگ و اشباع است
- بیشتر پهنای باند به پارامتر روشنایی (Y) اختصاص می‌یابد
- پارامتر روشنایی (Y) از دو پارامتر مشخص کننده رنگ جدا است

فضای رنگ HSL

- HSL (Hue, Saturation, Lightness)
- شبیه‌ترین فضا به سیستم بینایی انسان است
- استفاده از S و H به شیوه‌ای که انسان‌ها رنگ را تشخیص می‌دهند، بسیار نزدیک است
- L از بقیه‌ی مشخصه‌های رنگ جدا شده است
- Hue: توصیف‌کننده‌ی رنگ خالص
- Saturation: یک رنگ خالص تا چه حد توسط نور یا رنگ سفید رقیق شده است
- Lightness: میزان شدت رنگ احساس شده توسط بیننده



فضاهای رنگ مفهوما یکنواخت

- تفاوت مفهومی بین دو رنگ، برابر تفاضل آن ها در فضای رنگ XYZ نیست
- ✓ فضای رنگ یکنواخت: فاصله فضایی بین دو رنگ، با قضاوت بیننده انسانی از شباهت دو رنگ، متناظر است

• $L^*a^*b^*$ و $L^*u^*v^*$

- L^* مقدار لومینانس یک رنگ در یک مقیاس نور یکنواخت است
- u^* و v^* نسخه های تبدیل ریاضی یافته ی رنگینگی هستند
- a^* و b^* بیانگر تغییرات رنگی یکنواخت بر روی محورهای زرد-سبز و زرد-آبی هستند

| کاربردها | مهمترین ویژگیها | فضای رنگ |
|---|--|--------------------------|
| <ul style="list-style-type: none"> • نمایشگرهای CRT • دوربین های تلویزیونی • گرافیک کامپیوتری • ... | <ul style="list-style-type: none"> • سیستم رنگ جمع پذیر، شامل رنگهای پایه قرمز، سبز، آبی • بر پایه ی نظریه ی سه رنگی • فقط مخصوص سیستمهای نمایشی • پیاده سازی و کاربرد آسان • غیرخطی و نیمه شهودی • وابسته به وسیله | RGB |
| <ul style="list-style-type: none"> • چاپ رنگی • عکس رنگی • نقاشی • صرفه جویی در هزینه به دلیل استفاده از رنگ سیاه | <ul style="list-style-type: none"> • رنگ های مکمل: فیروزه ای، بنفش، زرد، سیاه • بکارگیری آسان • تبدیل به آن از فضای RGB مشکل است (مگر با بعضی ساده‌انگاری‌ها) • وابسته به دستگاه • چاپگرها و وسایل نظیر آنها اغلب دارای رنگ سیاه جداگانه هستند • غیرخطی و غیرشهودی | CMY(K) |
| <ul style="list-style-type: none"> • پردازش تصویر • جداسازی لومینانس می تواند در پردازش تصویر و سایر کاربردها، بسیار مفید واقع شود. | <ul style="list-style-type: none"> • اغلب این فضاها به یک تبدیل خطی از فضای RGB بدست می آیند • غیرخطی ولی بسیار شهودی • وابسته به دستگاه | HSL HSI HSV HCI |

موسسه آموزش عالی غیر انتفاعی رایانه ای تهران

جزوه درس : سیستم های چند رسانه ای

چکیده : پنجم

| فضای رنگ | مهمترین ویژگیها | کاربردها |
|--|---|--|
| <ul style="list-style-type: none"> • YUV • YIQ • YCbCr • YCC | <ul style="list-style-type: none"> • YUV و YIQ فضاهای آنالوگ برای PAL و NYSC هستند. • YCbCr برای سیستم های دیجیتال بکار می رود. • روشی برای از رنگینگی (نور را از رنگ) جدا می کنند • وابسته به دستگاه • غیرشهودی، مگر اینکه متخصص تلویزیون باشید | <ul style="list-style-type: none"> • انتقال تلویزیونی • انتقال تصویر رنگی • پردازش تصویر • فشرده سازی تصویر • ذخیره سازی تصویر |
| <ul style="list-style-type: none"> • XYZ | <ul style="list-style-type: none"> • ریشه و پایه ی علوم رنگ سنجی • X, Y و Z برای تمامی رنگ ها دارای مفادیر مثبت می باشند. • Y جزء لومینانس رنگ می باشد • غیر خطی • بصورت بکخواحت قابل مشاهده نیستند • X=Y=Z ستاقل R=G=B با انرژی نوری یکسان می باشد | <ul style="list-style-type: none"> • به عنوان فضای رنگ استاندارد، اغلب در تعریف فضاهای رنگی دیگر به کار می رود • معاوضه ی داد هپای تصویر رنگی بین دستگاه های مختلف. • کالیبراسیون برای دستگاه های مختلف |
| <ul style="list-style-type: none"> • L*U*V* • L*a*b | <ul style="list-style-type: none"> • بر پایه ی سیستم بینایی انسان • تقریبا خطی با درک بصری • مستقل از دستگاه • کاملا غیرشهودی | <ul style="list-style-type: none"> • کالیبراسیون برای دستگاه های مختلف نظیر چاپگرها، دوربین های دیجیتال و ... • معاوضه ی داد هپای تصویر رنگی بین دستگاه های مختلف • مدیریت رنگ • پردازش تصویر |



صدا Sound

صدا یک پدیده موجی نظیر نور است اما میکروسکوپی است و شامل مولکول های هواست که تحت عمل برخی از دستگاه های فیزیکی فشرده شده و گسترش یافته است از آنجا که صدا یک موج فشار است، مقادیر پیوسته دارد صوت دارای خواص و رفتار موج عادی است:



● انعکاس : برخورد و بازگشت

● شکست : تغییر زاویه هنگام ورود به رسانه ای با تراکم متفاوت

● انکسار : خم شدن اطراف مانع

دیجیتال سازی صدا



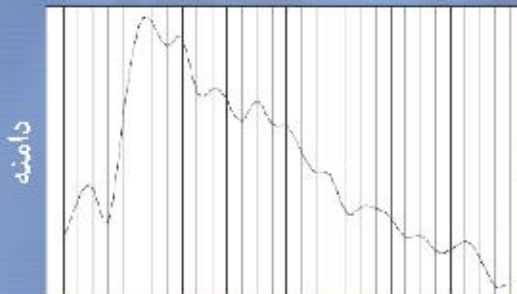
دیجیتال سازی یعنی تبدیل به یک رشته از اعداد مقادیر دامنه این موج به یک متغیر یک بعدی (زمان) وابسته اند برای دیجیتال کردن نیاز به نمونه برداری سیگنال در هر دو بعد است: زمان و دامنه نمونه برداری به معنی اندازه گیری کمیت مطلوب، معمولاً در فواصل زمانی مساوی (فرکانس نمونه برداری)

دیجیتال سازی صدا

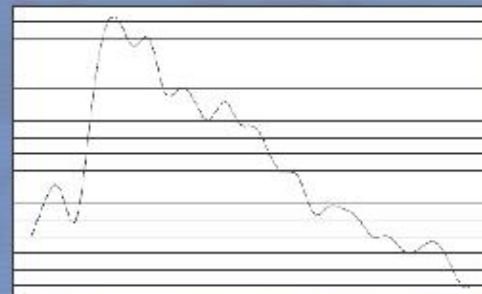
سرعت نمونه برداری برای صوت:

۸ کیلو هرتز (۸۰۰۰ نمونه در ثانیه) تا ۴۸ کیلو هرتز → قضیه نایکوئیست

نمونه برداری در بعد دامنه موج، کوانتیزاسیون نام دارد (شکل ب)



زمان
(الف)



زمان
(ب)

نرخ سیگنال به نویز SNR

نرخ قدرت سیگنال صحیح به نویز

Signal to Noise Ratio

معیاری برای کیفیت سیگنال

SNR معمولاً با واحد دسی بل اندازه گیری می شود

مقدار SNR، در واحد دسی بل، به شکل لگاریتم پایه ۱۰ از مربع ولتاژ تعریف شده است:

$$SNR = 10 \log_{10} \frac{V_{signal}^2}{V_{noise}^2} = 20 \log_{10} \frac{V_{signal}}{V_{noise}}$$

مقادیر SNR برای چند مثال

| | | | |
|-----|------------------------|----|----------------------|
| ۸۰ | راديو با صدای بلند | ۰ | آستانه شنوایی |
| ۹۰ | صدای قطار (از ایستگاه) | ۱۰ | خش خش برگ |
| ۱۰۰ | میخ پرچ کن | ۲۰ | اتاق خیلی آرام |
| ۱۲۰ | آستانه ناراحتی | ۴۰ | اتاق با صداهای متوسط |
| ۱۴۰ | آستانه درد | ۶۰ | گفتگو |
| ۱۶۰ | آسیب به گوش | ۷۰ | خیابان شلوغ |

نرخ سیگنال به نویز کوانتیزاسیون (SQNR)

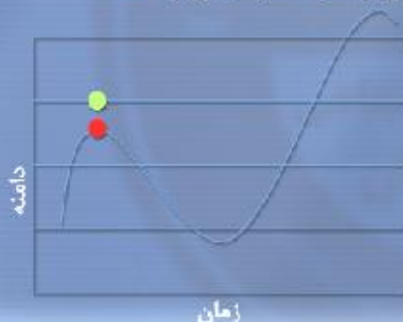
Signal to Quantization Noise Ratio

کیفیت کوانتیزاسیون توسط SQNR مشخص می شود

نویز کوانتیزاسیون: تفاوت بین مقدار واقعی سیگنال آنالوگ، در زمان خاصی از نمونه برداری،

و نزدیک ترین مقدار فاصله کوانتیزاسیون حداکثر این خطا می تواند به اندازه نیمی از فاصله

نمونه برداری باشد



کوانتیزاسیون خطی و غیر خطی

خطی: نمونه ها به شکل مقادیری که به طور یکنواخت کوانتیزه شده اند، ذخیره می شود
غیر خطی: سطوح با فواصل بهتری تنظیم می شوند که انسان به بهترین نحو آن را می شنود
ابتدا یک سیگنال آنالوگ از فضای اولیه S به فضای نظری ۲ تبدیل می شود
سپس مقادیر بدست آمده، به طور یکنواخت کوانتیزه می شوند μ -law یا u-law



فیلترینگ صوت

قبل از نمونه برداری و تبدیل آنالوگ به دیجیتال، سیگنال های صوتی برای حذف فرکانس های ناخواسته فیلتر می شوند
سخنرانی: ۵۰ هرتز تا ۱۰ کیلو هرتز
موسیقی: ۲۰ هرتز تا ۲۰ کیلو هرتز
در بخش تبدیل دیجیتال به آنالوگ، ممکن است فرکانس های بالا در خروجی دوباره تولید شوند
سیگنال ورودی با یک سری از توابع پلکانی که شامل تمام فرکانسهای ممکن هستند، جایگزین می شود

صدا های مصنوعی



Synthetic Sounds

۱. مدولاسیون فرکانس

۲. ساخت جدول موج

- تولید صوت از سیگنال های دیجیتال
- نمونه های دیجیتالی واقعی صدا از سازهای واقعی در حافظه بر روی کارت صدا ذخیره می شوند
- توسط نرم افزار ترکیب و ویرایش می شوند

واسط دیجیتال وسایل موسیقی



MIDI : Musical Instrument Digital Interface

- زبان برنامه نویسی است که رویدادهایی را صوت را تولید می کنند، کد می کند
- استاندارد دی که توسط صنعت موسیقی الکترونیکی برای کنترل دستگاه هایی که صوت تولید می کنند، به تصویب رسیده است
- کامپیوترها باید رابط MIDI داشته باشند
- کارت صوتی همچنین باید مبدل های دیجیتال به آنالوگ و برعکس را داشته باشد



سیستم های چند رسانه ای

صوت

مباحث این جلسه

- مفاهیم MIDI
- تبدیل MIDI به WAV
- کوانتیزاسیون و انتقال
- مدولاسیون پالس کد
- انواع فرمت های فایل های صوتی

AAIC / AIF / MPA / MID / WAV / WMA / MP3

مفاهیم MIDI

کانال های MIDI برای جداسازی پیام ها استفاده می شوند :



- کانال : چهار بیت آخر پیام
- ۱۶ کانال که از ۰ تا ۱۵ شماره گذاری شده اند، وجود دارد
- معمولا هر کانال با ابزار خاصی مرتبط است
- می توان ابزار را در خلال پیام تغییر داد

مفاهیم MIDI



طنین : اصطلاحی در MIDI برای تعیین ساز
آوا : طنین و زیر و بمی که همزمان تولید می شود
ابزار : (یا کارت صدایی) که چند-طنینی است قادر است بسیاری از صداها را همزمان پخش کند

تبدیل MIDI به WAV



برنامه های مختلفی برای تبدیل بین فرمت های MIDI و WAV وجود دارد. این برنامه ها شامل فایل های lookup بزرگ هستند که تلاش می کنند خروجی از پیش تعریف شده WAV را برای پیام های MIDI، جایگزین کنند.

کوانتیزاسیون و انتقال

کوانتیزاسیون و انتقال: کدگذاری داده

Compression - Expanding :Companding

روش μ -law برای Companding سیگنال های صوتی، معمولاً با یک الگوریتم که افزودگی زمانی موجود در سیگنال های صوتی را استخراج می کند، ترکیب می شود.

تولید خروجی کوانتیزه شده برای صوت، مدولاسیون پالس کد (Pulse Code Modulation - PCM) نامیده می شود:

- DPCM ●
- DM ●
- ADPCM ●

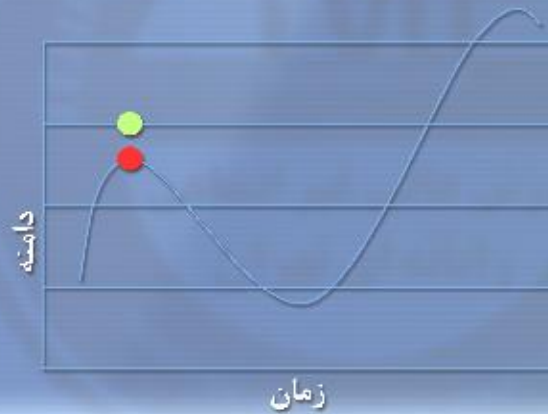


مدولاسیون پالس کد

کوانتیزاسیون: انتخاب نقاط انفصال در دامنه و نگاشت هر مقدار در یک فاصله به یکی از سطوح خروجی

مجموعه مرزهای فاصله ها: مرزهای تصمیم گیری

مقادیر: سطوح بازسازی



انواع فرمت های فایل های صوتی



.AIF

Audio Interchange File Format

کیفیت صوتی سی دی غیر فشرده

مورد استفاده برای ذخیره سی دی های صوتی استاندارد

توسعه توسط شرکت کامپیوتری اپل، بر اساس فرمت IFF. هنرهای الکترونیک

نرخ نمونه برداری: ۴۴.۱ کیلوهرتز، ۱۶ بیت و دارای دو کانال صدای استریو



.WAV

WAVE Audio File

فایل صوت دیجیتال استاندارد برای ذخیره سازی داده های به شکل موج

ضبط صوت با نرخ نمونه برداری و نرخ بیت مختلف

نرخ نمونه برداری: ۴۴.۱ کیلوهرتز، ۱۶ بیتی، استریو

مشابه AIF

مبتنی بر فرمت تبادل منابع

اغلب در رایانه های مبتنی بر ویندوز

DTS - WAV File

فایل های ایجاد شده توسط برنامه های صوتی DTS

حاوی داده های صوتی برای کانال های متعدد

قابل پخش توسط برنامه هایی که کدگذاری آن را پشتیبانی می کنند



.MP3

MPEG -1 Audio Layer -3

فرمت صوتی فشرده در "لایه ۳"

فرمت معمول برای ذخیره فایل های موسیقی

کیفیت صدای استریو، ۱۶ بیتی در فایلی با تقریباً یک دهم ساینز WAV، یا AIF.

کیفیت فایل MP3 تا حد زیادی به نرخ بیت استفاده شده برای فشرده سازی بستگی دارد

۱۲۸، ۱۶۰، ۱۹۲ و ۲۵۶ کیلوبیت بر ثانیه

توسط اکثر پخش کننده های موسیقی قابل حمل مانند iPod Apple و Microsoft Zune پشتیبانی می شود



.WMA

Windows Media Audio File

فایل صوتی فشرده سازی شده با Windows Media compression

فرمت اختصاصی طراحی شده توسط مایکروسافت، شبیه به فرمت MP3

قابل تبدیل به فرمت های استاندارد دیگر

اغلب برای پخش موسیقی از وب استفاده می شود

قابل پخش در پخش کننده قابل حمل Microsoft's Zune و Disney Mix Stick MP3 Player

.MID

Musical Instrument Digital Interface

استفاده توسط برنامه های ایجاد و ویرایش فایل های موسیقی و همچنین دستگاه های سخت افزاری MIDI

حاوی داده های موسیقی

چه نت هایی نواخته شده

زمان نواخته شدن هر نت

مدت نواخته شدن هر نت

آوا (بلندی) هر نوت

رمز کردن داده های MIDI به فرمت استاندارد سازگار بین انواع مختلف برنامه های نرم افزاری و ابزارهای سخت افزاری

شامل داده های کنترل کننده برای دستگاه

حاوی مشخصه های داده موسیقی

.MPA

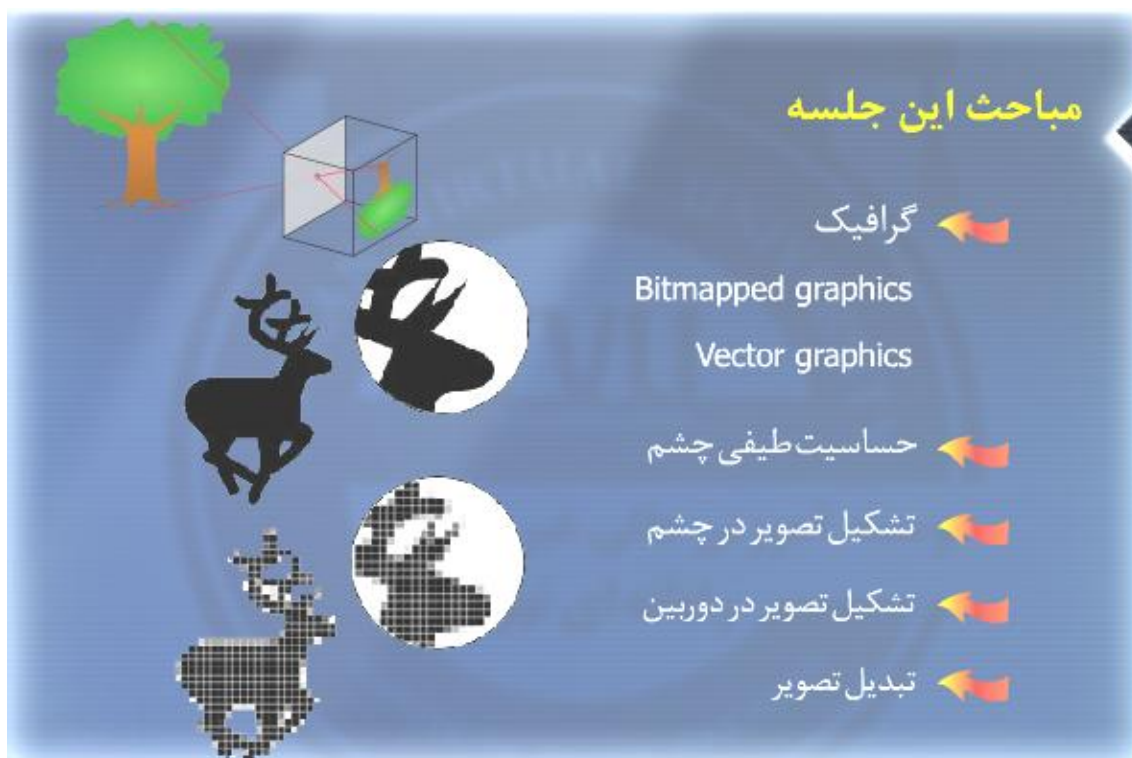
MPEG-2 Audio File

فایل صوتی فشرده با فشرده سازی MPEG لایه دوم

کاهش حجم فایل با حداقل اثر در کیفیت کلی صدا

بر اساس یک استاندارد کدگذاری صوتی 1 - MPEG





گرافیک

نمودار، نقاشی، عکس، دکمه ها و ...

گرافیک می تواند سیاه و سفید، Grayscale، رنگی، و ثابت یا متحرک باشد

نمایش رویه ها، واضح کردن داده، بیان مفاهیم، اضافه کردن نمود تصویری

کیفیت گرافیک

وضوح (resolution) : تعداد پیکسل ها در هر اینچ

عمق رنگ: تعداد رنگ های متمایز (از ۲ تا ۲۴ بیت)

تصاویر ایجاد شده یا ویرایش شده روی کامپیوتر:



Bitmapped graphics

Vector graphics

Bitmapped graphics



Raster Graphics

شامل شبکه هایی از پیکسل های رنگی است

می تواند تصویری تشکیل شده از ترکیبات نواحی مجزا (continuous - tone) باشد

ویرایشگرهای آن برنامه های نقاشی (paint) نام دارند

تغییر سایز تصویر باعث از بین رفتن وضوح می شود

Vector graphics

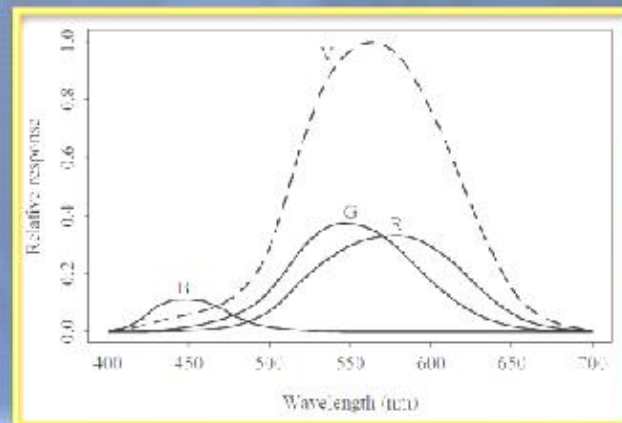
از فرمول های ریاضی برای تعریف خطوط، منحنی ها و سایر مشخصه ها استفاده می شود
فایل های کوچکتری هستند



ویرایشگرهای آن برنامه های طراحی (draw) نامیده می شوند
با تغییر سایز، وضوح آن ها از بین نمی رود
به شکل bitmap نمایش داده می شوند

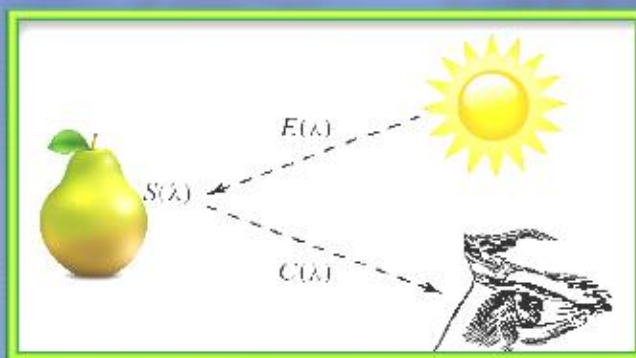
حساسیت طیفی چشم

- چشم به وسط طیف مرئی نور بیشتر حساس است
- حساسیت گیرنده های ما تابعی از طول موج است
- تابع منحنی نورانی-کارایی: $V(\lambda)$



تشکیل تصویر در چشم

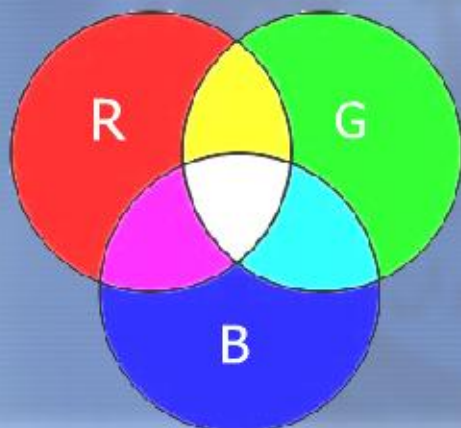
سطوح مقدار متفاوتی از نور را در طول موج های مختلف منعکس می کنند.
سطوح تیره انرژی کمتری را نسبت به سطوح روشن منعکس می کنند.
نور از منبع نور، به سطحی برخورد می کند، منعکس می شود و توسط توابع چشم فیلتر می شود.



$$C(\lambda) = E(\lambda) S(\lambda)$$

تشکیل تصویر در دوربین

یک دوربین استودیو دارای سه سیگنال است که در موقعیت هر پیکسل تولید می شود.
سیگنال های آنالوگ به دیجیتال تبدیل می شوند، به عدد صحیح تبدیل شده و ذخیره می شود.
اگر دقت استفاده شده ۸ بیت باشد، حداکثر مقدار هر کدام از رنگ ها (R و G و B) ۲۵۵ است.
چشم نور ساطع شده از صفحه نمایش را می بیند.



تبدیل تصویر

به دو بخش شکسته می شود :

● بخش هندسی

کاهش مقیاس بلوک های $2n \times 2n$ به $n \times n$

● بخش جرمی (massic)

تبدیل مقادیر همه پیکسل ها به سطح خاکستری یکسان

تغییر مقیاس دادن contrast

تغییر روشنایی

تغییر در الگوی بیت ها ، انعکاس و یا چرخش



مباحث این جلسه



- انواع فایل های Bitmap ←
 - .JPG
 - .BMP
 - .GIF
 - .PNG
 - .TIF
- انواع فایل های Vector ←
 - .SVG
 - .AL
 - .EPS
- فایل های تصویری سه بعدی ←

انواع فایل های Bitmapped



| | |
|--|-----------------|
| JPEG Image File | .jpg ● |
| Bitmap Image File | .bmp ● |
| Graphical Interchange Format File | .gif ● |
| Portable Network Graphic | .png ● |
| Tagged Image File | .tif ● |
| Adobe Photoshop Document | .psd ● |
| Photo Album | .abm ● |
| DICOM Image File | .dcm ● |
| JPEG Image File | .jpeg ● |
| Stereo JPEG Image | .jps ● |
| JPEG 2000 Image File | .jpx ● |
| Nikon Electronic Format RAW Image File | .nef ● |
| Picture Clipping File | .pictclipping ● |
| Tagged Image File Format | .tiff ● |
| Thumbnail Image File | .thm ● |

JPG



توسعه دهنده: Joint Photographic Experts Group ●

فرمت تصویر فشرده سازی شده ●

کاهش حجم

کاهش کیفیت

پشتیبانی تا رنگ های ۲۴ بیت ●

استفاده: ●

برای ذخیره سازی تصاویر دیجیتال

انتشار تصویر در وب



.BMP

- تصویر Raster فشرده نشده
- شبکه ای مستطیلی از پیکسل ها
- محتویات فایل

Header : شناسه، اندازه فایل، پهنا، ارتفاع، گزینه های رنگ، نقطه شروع داده Bitmap

پیکسل های bitmap

هر پیکسل می تواند عمق متفاوتی از رنگ را داشته باشد



.GIF

Graphical Interchange Format File

- تا ۲۵۶ رنگ
- محدوده رنگ ها: از پیش تعریف شده یا وفق یافته با رنگ های داخل تصویر
- فشرده سازی بدون از دست دادن کیفیت
- استفاده شده در وب
- قابلیت ذخیره سازی به شکل تصویر متحرک
- قابلیت داشتن پیکسل های شفاف

.PNG



- Portable Network Graphic
- شامل رنگ های Bitmap
- فشرده سازی بدون از دست دادن کیفیت
- استفاده برای ذخیره سازی تصاویر وب
- افزایش پشتیبانی از رنگ نسبت به GIF.
- ۸ بیت برای شفافیت
- نمی تواند متحرک باشد

.TIF



- Tagged Image File
- ذخیره سازی تصاویر با تعداد زیادی رنگ
- پشتیبانی از لایه ها و صفحات متعدد
- قابلیت ذخیره سازی بدون فشرده سازی یا استفاده از فشرده سازی JPG

انواع فایل های Vector



| | |
|--------------------------------------|---------|
| Scalable Vector Graphics File | .svg ● |
| Adobe Illustrator File | .ai ● |
| Encapsulated PostScript File | .eps ● |
| Drawing File | .drw ● |
| PostScript File | .ps ● |
| ConceptDraw PRO Document | .cdd ● |
| ConceptDraw MINDMAP Document | .cdmm ● |
| CorelDRAW Image File | .cdr ● |
| Computer Graphics Metafile | .cgm ● |
| Enhanced Windows Metafile | .emf ● |
| Windows Compressed Enhanced Metafile | .emz ● |
| Flash XML Graphics File | .fxg ● |
| Visio Drawing File | .vsd ● |
| Xara Xtreme Drawing | .xar ● |

.SVG



- Scalable Vector Graphics File
- توسعه دهنده: Adobe Systems
- فرمت استاندارد برای نمایش گرافیک Vector روی وب
- توصیف تصاویر با استفاده از xml



.AI



- Adobe Illustrator File
- توسعه دهنده: Adobe Systems
- تشکیل شده از خطوطی که نقاط را به هم متصل می کنند
- استفاده برای لوگوها
- برخی نرم افزارها برای باز کردن این فایل ها آن ها را به Bitmap تبدیل می کنند



.EPS



- Encapsulated PostScript File
- توسعه دهنده: Adobe Systems
- فایل PS، شامل تصاویر vector و bitmap و متن
- شامل یک تصویر پیش نمایش تعبیه شده
- پشتیبانی شده در اکثر سیستم عامل ها

فایل های تصویری سه بعدی

- مدل ها ، انیمیشن ها و نقشه های سه بعدی
- تشکیل شده از چندضلعی ها و رئوسی که شکل مدل را ایجاد می کنند
- شامل اطلاعات بافت ها ، روشنایی و سایه
- انواع فرمت ها :



3D Object File .obj

Maya Project File .ma

IGES File .iges



ویدئو در چند رسانه ای



- نمایش ترکیبی از تصویر، صوت و متن (فریم ها) با نرخ فریمی ثابت
- مزایا:
- کمک به انتقال واقعیت
- تاثیرگذاری بیشتر
- موارد استفاده:
- شبیه سازی ویدئویی بروسه ها و تعاملات
- بازی ها که شبیه سازی دنیای واقعی هستند
- وب سایت ها
- ویدئو کنفرانس
- تبلیغات

مفاهیم

- کیفیت ویدئو
- تفکیک پذیری و وضوح فریم های ویدئو
- نرخ فریم
- streaming video : انتقال ویدئو روی شبکه به شکلی که دریافت کننده قبل از دریافت کل فایل شروع به پخش می کند
- نرم افزار: پخش، ویرایش و توزیع ویدئو
- سخت افزار
- Camcorders : ضبط ویدئو به شکل دیجیتال یا آنالوگ
- Webcams : درج ویدئوی زنده و streaming در صفحه وب
- Capture cards : تبدیل ویدئوی آنالوگ به دیجیتال
- Video tuners : ایجاد امکان دریافت و ضبط سیگنال های تلویزیونی برای کامپیوتر



انواع سیگنال های ویدئو

Component video

- از سه سیگنال ویدئویی جداگانه برای رنگ های قرمز، سبز و آبی استفاده می کند
- برای هر طرح تفکیک رنگ، بدلیل عدم وجود crosstalk بین سه کانال، بهترین نمایش رنگ ها را دارد
- به پهنای باند بالا و همزمانی مناسب بین سه مولفه نیاز دارد

انواع سیگنال های ویدئو

Composite Video

- دارای یک سیگنال که ترکیبی از رنگ (chrominance) و نور (luminance) است
- رنگ ترکیبی از دو مولفه رنگی I و Q، یا U و V است
- دو سیگنال می توانند جدا شده و ارسال شوند و درگیرنده دوباره ترکیب شوند
- در اتصال تلویزیون و VCR فقط یک کابل استفاده می شود و سیگنال ها ترکیب می شوند
- بدلیل ترکیب سیگنال ها، تداخل غیرقابل اجتناب است

انواع سیگنال های ویدئو

S-Video

● video-Super یا Separated video

- از دو سیگنال، یکی برای روشنایی و دیگری برای ترکیبی از رنگ ها استفاده می کند
- دلیل قرار دادن روشنایی در سیگنالی جداگانه، این است که اطلاعات سیاه و سفید برای درک بینایی مهم است

انواع سیگنال های ویدئو

● یک سیگنال آنالوگ بک تصویر متغیر در زمان را نمونه برداری می کند

● progressive scanning: فریم ها در فواصل زمانی ثابت به شکل خط به خط

ترسیم می شوند

● interlaced scanning: ابتدا خطوط با شماره فرد و سپس خطوط زوج ترسیم

می شوند که نتیجه آن یک فریم است

انواع سیگنال های ویدئو

● horizontal retrace : پرش از Q به R

● vertical retrace : پرش از T به U



NTSC

● National Television System Committee

● استاندارد تلویزیونی که بیشتر در آمریکا و ژاپن مورد استفاده است

● نسبت تصویر ۴:۳


● ۵۲۵ خط پویا در هر فریم

● ۳۰ فریم در ثانیه

● استفاده از interlaced scanning







Phase Alternating Line

استاندارد تلویزیونی که بیشتر در غرب اروپا، چین، هند و بسیاری از دیگر نقاط جهان استفاده می شود

نسبت تصویر ۴:۳، ۶۲۵ خط اسکن در هر فریم، ۲۵ فریم در ثانیه، interlaced scanning

استفاده از مدل رنگ YUV



Systeme Electronique Couleur Avec Mémoire

۶۲۵ خط اسکن در هر فریم، ۲۵ فریم در ثانیه، نسبت تصویر ۴:۳، interlaced scanning

تفاوت با PAL در کدگذاری رنگ ها:

سیگنال های U و V در SECAM مجزا هستند

در هر خط اسکن، فقط یکی از U یا V ارسال می شود

مقایسه

| TV System | Frame Rate (fps) | # of Scan Lines | Total Channel Width (MHz) | Bandwidth Allocation (MHz) | | |
|-----------|------------------|-----------------|---------------------------|----------------------------|--------|--------|
| | | | | Y | I or U | Q or V |
| NTSC | 29.97 | 525 | 6.0 | 4.2 | 1.6 | 0.6 |
| PAL | 25 | 625 | 8.0 | 5.5 | 1.8 | 1.8 |
| SECAM | 25 | 625 | 8.0 | 6.0 | 2.0 | 2.0 |

ویدئوی دیجیتال

مزایای ویدئوی دیجیتال

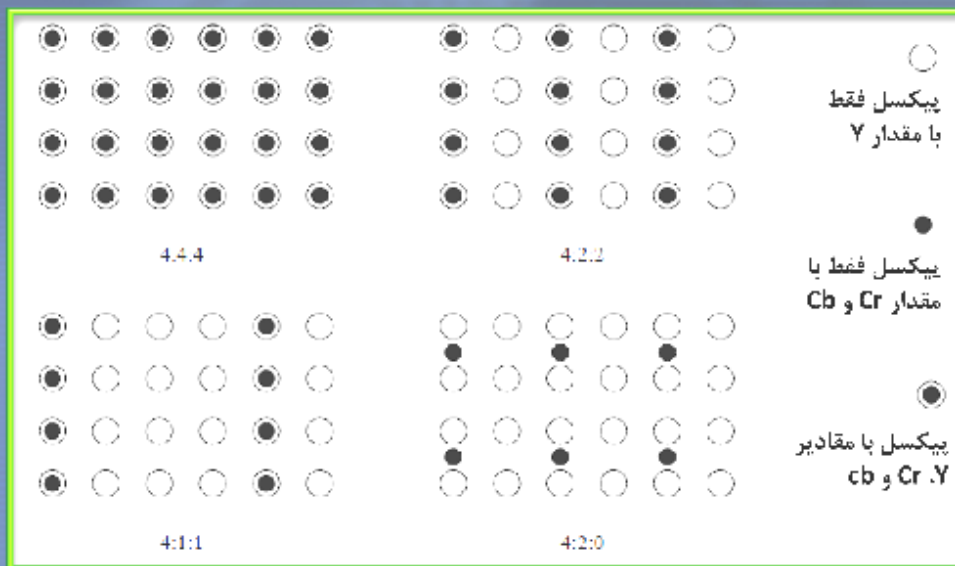


- ذخیره سازی فایل ها روی دستگاه های دیجیتال و حافظه
- پردازش فایل ها
- استفاده در برنامه های کاربردی
- دسترسی مستقیم (در مقایسه با پردازش خطی ویدئو)
- عدم کاهش کیفیت در صورت کپی شدن یا ضبط مجدد
- کدگذاری ساده تر و تحمل نویز بالاتر

ویدئوی دیجیتال

- از فضای رنگ YCbCr استفاده می کند
- برای کاهش حجم و پهنای باند مورد نیاز نمونه برداری رنگ صورت می گیرد
- ۴ نوع نمونه برداری وجود دارد؛ براساس اینکه چه تعداد پیکسل، به ازای هر ۴ پیکسل اولیه، ارسال می شود

نمونه برداری رنگ



استانداردهای ویدئوی دیجیتال

| | CCIR 601 525/60 NTSC | CCIR 601 625/50 PAL/SECAM | CIF | QCIF |
|------------------------|----------------------------|---------------------------------|-----------|-----------|
| Luminance resolution | 720 × 480 | 720 × 576 | 352 × 288 | 176 × 144 |
| Chrominance resolution | 360 × 480 | 360 × 576 | 176 × 144 | 88 × 72 |
| Color Subsampling | 4:2:2 | 4:2:2 | 4:2:0 | 4:2:0 |
| Aspect Ratio | 4:3 | 4:3 | 4:3 | 4:3 |
| Fields/sec | 60 | 50 | 30 | 30 |
| Interlaced | Yes | Yes | No | No |

فرمت فایل های ویدئویی

| فرمت فایل های ویدئویی | | | | |
|--|--------------------------|----------------------------|----------------------------|------------------------|
| MPEG | AVI | WMV | FLV | |
| MPEG-1 یا MPEG-2 | استفاده از فشرده سازی کم | استفاده از فشرده سازی زیاد | استفاده از فشرده سازی زیاد | فشرده سازی |
| پایین | بالا | پایین | خیلی پایین | حجم |
| بالا | بالا | نامناسب | مناسب | کیفیت تصویر |
| کم | زیاد | متوسط | کم | سازگاری با نرم افزارها |
| Home DVD - Player MP4- برای اشتراک روی وب | -ویدئوی اولیه در ویرایش | -ایمیل -اشتراک روی وب | -پخش روی وب | موارد استفاده |



فشرده سازی

فرآیند کدگذاری که به طور موثری تعداد بیت‌های لازم برای نمایش اطلاعات خاصی را کاهش می‌دهد.

```
graph LR; A[کد گذاری (Compression)] --> B[ذخیره‌سازی / انتقال روی شبکه / ...]; B --> C[کد گشایی (Decompression)];
```

انواع فشرده‌سازی:

- Lossless
- Lossy

فشرده سازی

نرخ فشرده‌سازی = $\frac{\text{تعداد بیت‌ها قبل از فشرده‌سازی}}{\text{تعداد بیت‌ها بعد از فشرده‌سازی}}$

$$\eta = H(S) = \sum_{i=1}^n p_i \log_2 \frac{1}{p_i}$$

S : اطلاعات منبع
 p_i : احتمال وجود S_i در S

- مجموع وزن دار $\log_2 1/p_i$
- متوسط حجم اطلاعات در هر نماد در S
- حد پایین متوسط تعداد بیت‌های لازم برای هر نماد

Run-Length Coding (RLC)

- شکل ساده‌ای از فشرده‌سازی اطلاعات
- از حافظه متبوع استفاده می‌کند.
- اگر منبع اطلاعات طوری باشد که نمادها به تشکیل گروه‌های پیوسته تمایل داشته باشند، می‌توان این نمادها و طول گروه هر یک را برای کد کردن استفاده نمود.

Run-Length Coding (RLC)

مثال: فرض کنید تصویری سیاه و سفید داریم که به ترتیب زیر نمایش داده شده است

```
WWWWWWWWWWBWWWWWWWWWWWWWWWBBBWWWWWWWW  
WWWWWWWWWWWWWWWWWWWWWWWWWBWWWWWWWWWW
```

9W1B12W3B24W1B8W

کاربردها:

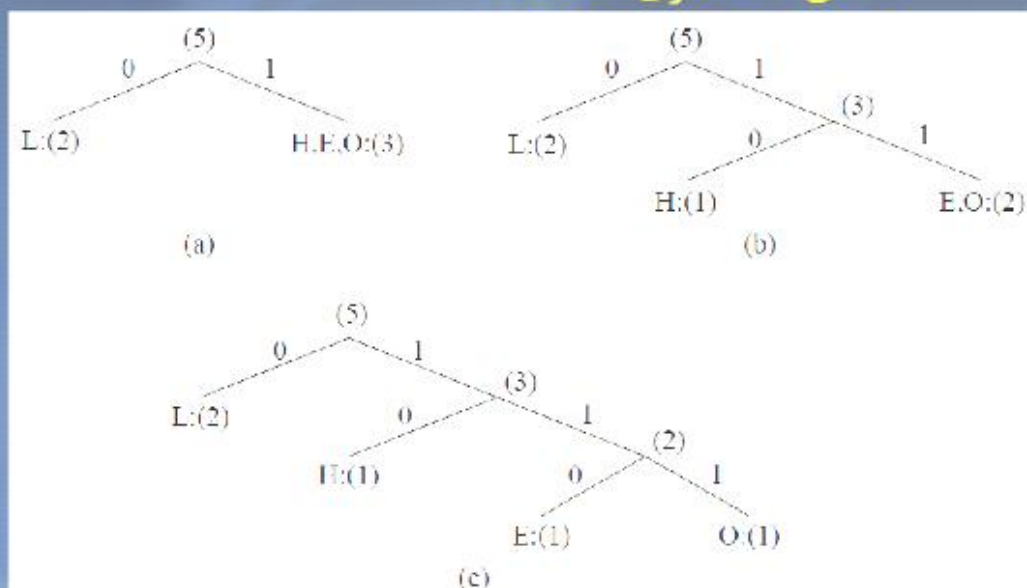
- بخشی از الگوریتم JPEG
- در دستگاه‌های فکس

Variable-Length Coding (VLC)

- نمادها را بر اساس شمار فراوانی رخدادشان مرتب کنید.
- نمادها را به صورت بازگشتی به دو قسمت تقسیم کنید، هر کدام با تعداد شماره های تقریباً برابر، تا جایی که هر قسمت فقط یک نماد داشته باشد.
- مثال: کدگذاری HELLO

| نماد | H | E | L | O |
|-------|---|---|---|---|
| تعداد | ۱ | ۱ | ۲ | ۱ |

مثال کدگذاری HELLO



مثال کدگذاری HELLO

| نماد | تعداد تکرار | $\log_2 1/p_i$ | کد | تعداد بیت‌های استفاده شده |
|------|-------------|----------------|-----|---------------------------|
| L | 2 | 1.32 | 0 | 2 |
| H | 1 | 2.32 | 10 | 2 |
| E | 1 | 2.32 | 110 | 3 |
| O | 1 | 2.32 | 111 | 3 |

تعداد کل بیت‌ها : 10

کدگذاری هافمن

- خصوصیت منحصر به فرد پیشوندی: هیچ یک از کدها پیشوند کد دیگری نیست بشر
 - بهینگی: حداقل افزونگی کد
- کد دو نماد با کمترین فرکانس، فقط در بیت آخر تفاوت دارد.
- نمادهایی که فرکانس بیشتری دارند کد هافمن کوتاه‌تری دارند.
- طول کد برای S کمتر از $n + 1$ است.



Dictionary-based Coding

Lempel-Ziv-Welch ●

- از codeword های با طول ثابت برای نمایش رشته‌های با طول متغیر استفاده می‌کند.
- یک دیکشنری به شکل پویا ساخته می‌شود.
- مدخل‌های طولانی‌تر را در دیکشنری وارد می‌کند و سپس یک کد برای عنصر تولید می‌شود.

الگوریتم LZ

```
BEGIN
  s = next input character;
  while not EOF
    { c = next input character;
      if s + c exists in the dictionary
        s = s + c;
      else
        { output the code for s;
          add string s + c to the dictionary with a new code;
          s = c;
        }
    }
  output the code for s;
END
```

مثال ABABBABCABABBA

| S | C | خروجی | کد | رشته |
|-----|-----|-------|----|------|
| A | B | 1 | 4 | AB |
| B | A | 2 | 5 | BA |
| A | B | | | |
| AB | B | 1 | 6 | ABB |
| B | A | | | |
| BA | B | 5 | 7 | BAB |
| B | C | 2 | 8 | BC |
| C | A | 3 | 9 | CA |
| A | B | | | |
| AB | A | 4 | 10 | ABA |
| A | B | | | |
| AB | B | | | |
| ABB | A | 6 | 11 | ABBA |
| A | EOF | 1 | | |

| کد | رشته |
|----|------|
| 1 | A |
| 2 | B |
| 3 | C |

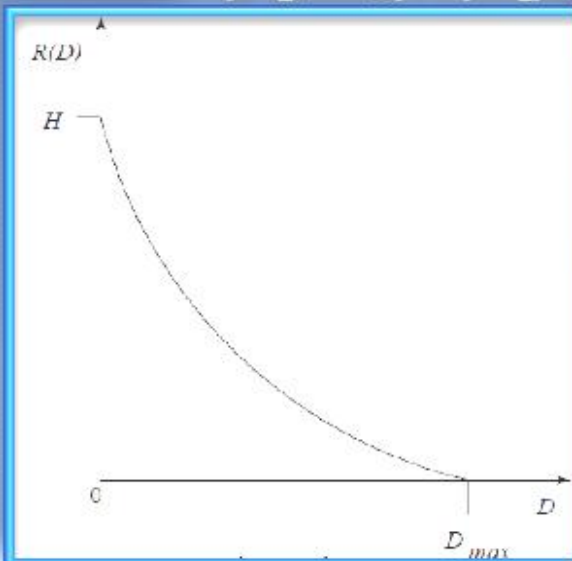
خروجی نهایی: ۱۲۴۵۲۳۴۶۱

کدگذاری حسابی

یک رشته از کاراکترها با استفاده از تعداد ثابتی از بیت‌ها برای هر کاراکتر نمایش داده می‌شود. کاراکترهای با فرکانس تکرار بیشتر، با تعداد بیت‌های کمتری ذخیره می‌شوند. هر پیام به شکل (a, b) نمایش داده می‌شود. وقتی پیام طولانی‌تری می‌شود، طول این فواصل کوتاه‌تر و تعداد بیت‌های مورد نیاز برای نمایش آن بیشتر می‌شود.

فشرده سازی با اتلاف

الگوریتم‌های فشرده سازی بدون اتلاف، غالباً نرخ فشرده‌سازی مناسبی ندارند.



نمودار تابع Rate-Distortion

کوآنتیزاسیون

نگاشت مقادیر خروجی متمایز به مجموعه‌ای کوچکتر و گسسته سه شکل متفاوت:

- یکنواخت (Uniform)
- غیر یکنواخت (Nonuniform)
- برداری (Vector)

کوآنتیزاسیون عددی یکنواخت

دامنه مقادیر ورودی را به فواصل مکانی یکسان تقسیم می کند.
خروجی یا مقدار بازسازی (Reconstruction Value) برای هر بازه، نقطه میانی آن بازه
در نظر گرفته می شود.

طول هر بازه اندازه گام نامیده شده و با علامت Δ نشان داده می شود.

انواع:

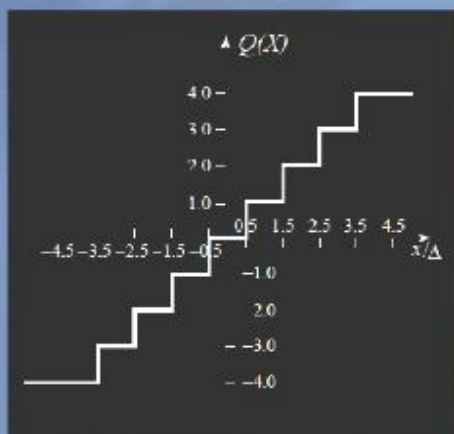
Midrise: دارای سطوح خروجی به تعداد زوج

Midtread: سطوح خروجی به تعداد فرد (شامل صفر)

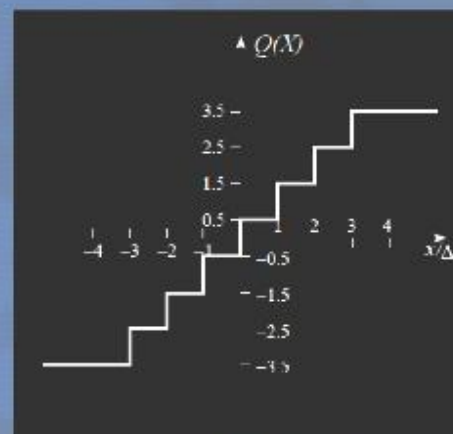
کوآنتیزاسیون عددی یکنواخت

$$Q_{midrise}(x) = \lceil x \rceil - 0.5 \quad \Delta = 1 \text{ برای}$$

$$Q_{midtread}(x) = \lfloor x + 0.5 \rfloor$$

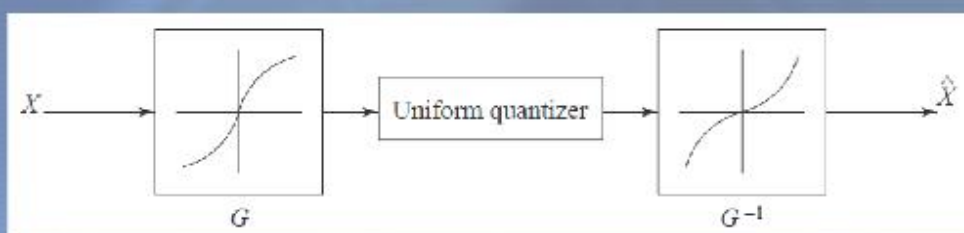


Midtread (ب)



Midrise (الف)

کوآنتیزاسیون غیر یکنواخت



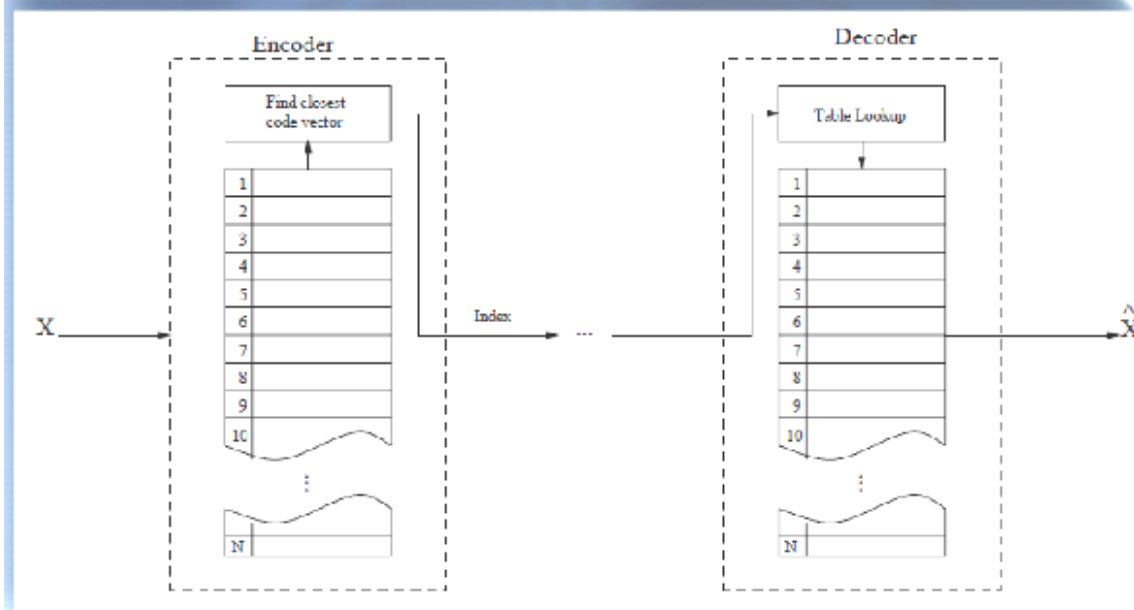
یک کوآنتیزاسیون غیر یکنواخت شامل تابع فشرده ساز G ، یک کوآنتایزر یکنواخت و یک تابع بسط دهنده G^{-1} است.

مثال: μ -law و A-law

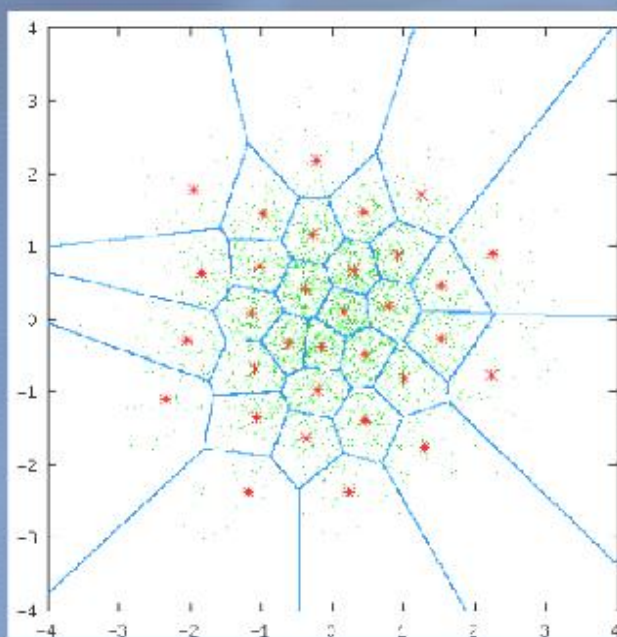
کوآنتیزاسیون برداری

اگر فشرده سازی بر روی گروهی از نمونه ها عمل کند، نتیجه بهتری نسبت به کاربرد روی تک نمادها دارد. الحاق تعدادی از نمونه های متوالی، بردار نامیده می شود. به جای قرار دادن مقادیر بازسازی به عنوان نماینده هر نمونه ورودی، بردارهای n قطعه ای (نمونه ای) در کوآنتایزر برداری مورد استفاده قرار می گیرند. مجموعه این بردارهای رمز Code Book را می سازند.

کوآنتیزاسیون برداری



کوآنتیزاسیون برداری





کدگذاری انتقالی

Transform Coding

منطق اصلی:

اگر γ نتیجه یک تغییر خطی (T) بردار X ورودی، به گونه ای باشد که قطعات γ خیلی کمتر به هم وابسته باشند، γ می تواند مؤثرتر از X کد شود
اگر بیشتر اطلاعات موجود در منبع به صورت دقیق توسط اجزای اولیه یک بردار انتقال توصیف شوند، بقیه اجزا از اهمیت کمتری برخوردار هستند و می توانند با ضریب بزرگ تری کوانتایز شوند.

مثال: Discrete Cosine Transform (DCT)

تبدیل کسینوسی گسسته (DCT)

تصویر را به قسمت ها (یا طیف های زیر باندها) که دارای اهمیت متفاوت (با توجه به کیفیت بصری عکس) تقسیم می کند.

تبدیل سیگنال یا تصویر از حوزه فضایی به حوزه فرکانس



تبدیل کسینوسی گسسته (DCT)

یک تبدیل خطی است.

به طور کلی یک تبدیل (یا تابع) T خطی است، اگر:

$$T(\alpha p + \beta q) = \alpha T(p) + \beta T(q)$$

وظیفه DCT: تجزیه کردن سیگنال اصلی به اجزا DC و AC

$$F(u, v) = \frac{C(u)C(v)}{4} \sum_{i=0}^7 \sum_{j=0}^7 \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} f(i, j)$$

تبدیل کسینوسی گسسته (DCT)

وظیفه IDCT: بازسازی سیگنال

$$\tilde{f}(i, j) = \sum_{u=0}^7 \sum_{v=0}^7 \frac{C(u)C(v)}{4} \cos \frac{(2i+1)u\pi}{16} \cos \frac{(2j+1)v\pi}{16} F(u, v)$$

| | | | | | | | |
|---|----|----|----|----|----|----|----|
| 8 | 16 | 24 | 32 | 40 | 48 | 56 | 64 |
|---|----|----|----|----|----|----|----|

DCT

| | | | | | | | |
|-----|-----|---|----|---|----|---|-----|
| 100 | -52 | 0 | -5 | 0 | -2 | 0 | 0.4 |
|-----|-----|---|----|---|----|---|-----|

Truncate

| | | | | | | | |
|-----|-----|---|----|--------------|---------------|--------------|----------------|
| 100 | -52 | 0 | -5 | 0 | -2 | 0 | 0.4 |
|-----|-----|---|----|--------------|---------------|--------------|----------------|

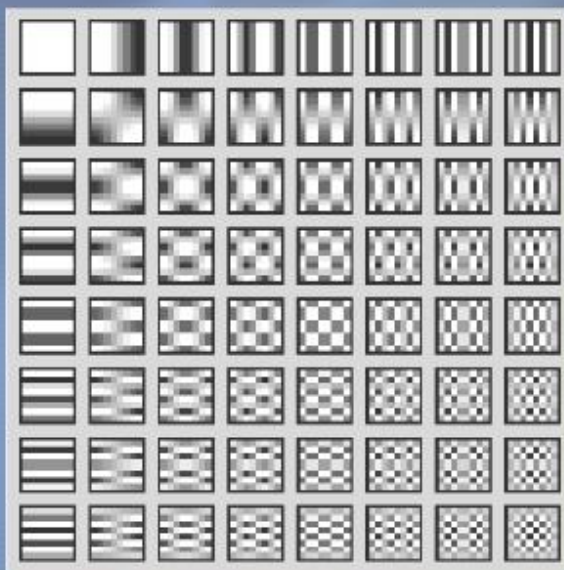
IDCT

| | | | | | | | |
|---|----|----|----|----|----|----|----|
| 8 | 15 | 24 | 32 | 40 | 48 | 57 | 63 |
|---|----|----|----|----|----|----|----|

عملیات اصلی DCT

تصویر ورودی N در M است.
 $f(i,j)$ مقدار پیکسل در سطر i و ستون j است.
 $F(u,v)$ ضریب DCT در سطر k_1 و ستون k_2 در ماتریس DCT است.
برای بسیاری از عکس ها ، بیشتر انرژی سیگنال در فرکانس های پایین نهفته است.
ورودی یک آرایه 8 در 8 از اعداد صحیح است.
آرایه خروجی ضرایب DCT شامل اعداد صحیح است (از -1024 تا 1023).

توابع پایه DCT



نمایش گرافیکی توابع پایه یک DCT دو بعدی 8×8

تبدیل Wavelet

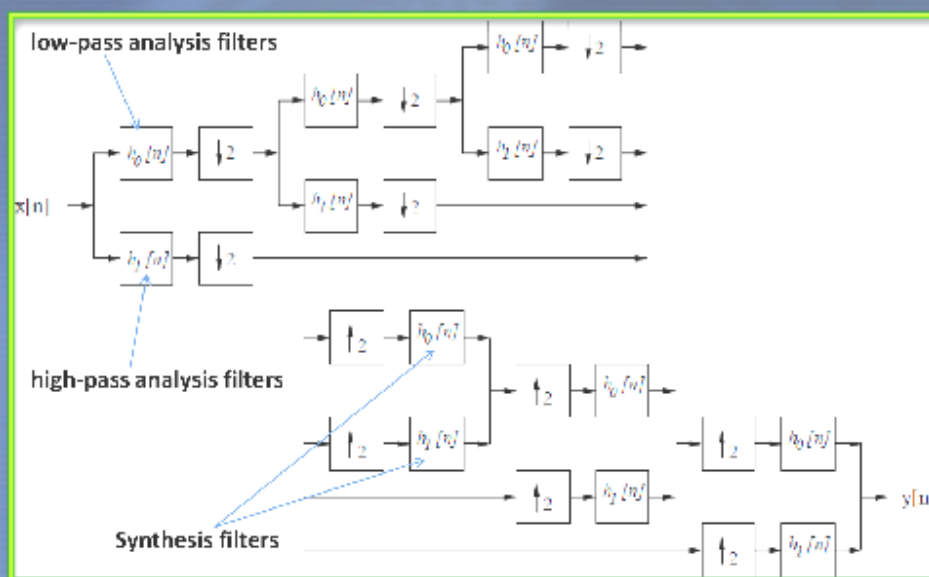
هدف از Wavelet تجزیه سیگنال ورودی به قطعاتی است که کار کردن با آنها ساده تر باشد و قابلیت بازسازی تقریبی سیگنال اصلی از روی قطعات را داشته باشد.

دو نوع تبدیل Wavelet :

تبدیل Wavelet پیوسته (CWT)

تبدیل Wavelet گسسته (DWT)

تبدیل Wavelet



تبدیل Wavelet دو بعدی

برای یک تصویر ورودی N در N :

هر ردیف از تصویر را با $h_0[n]$ و $h_1[n]$ تبدیل کنید. ستون های فرد را از آرایه ها حاصل حذف کنید و آرایه ها را به هم وصل کنید تا یک ردیف ایجاد شود.

بعد از اینکه تمام ردیف ها تبدیل شدند، هر ستون حاصل را با $h_0[n]$ و $h_1[n]$ تبدیل نمایید. این بار ردیف های فرد را حذف کنید و آرایه های باقیمانده را به هم وصل کنید.

تصویر نهایی اکنون حاوی 4 تا subband است: HL ، LH ، LL و HH .

تبدیل Wavelet دو بعدی

| | |
|----|----|
| LL | HL |
| LH | HH |

تبدیل یک سطحی

| | | |
|-----|-----|-----|
| LL2 | HL2 | HL1 |
| LH2 | HH2 | |
| LH1 | | HH1 |

تبدیل دو سطحی



JPEG

- یک استاندارد بین المللی
- استفاده از کدگذاری تبدیل گسسته کسینوسی (DCT)
- یک تصویر، تابعی از z (یا به طور قراردادی x و y) در فضای مختصات می باشد.
- استفاده از DCT دوبعدی به عنوان یک مرحله در JPEG برای بدست آوردن پاسخ فرکانسی که $F(u, v)$ در دامنه فرکانس فضایی مختصات می باشد.

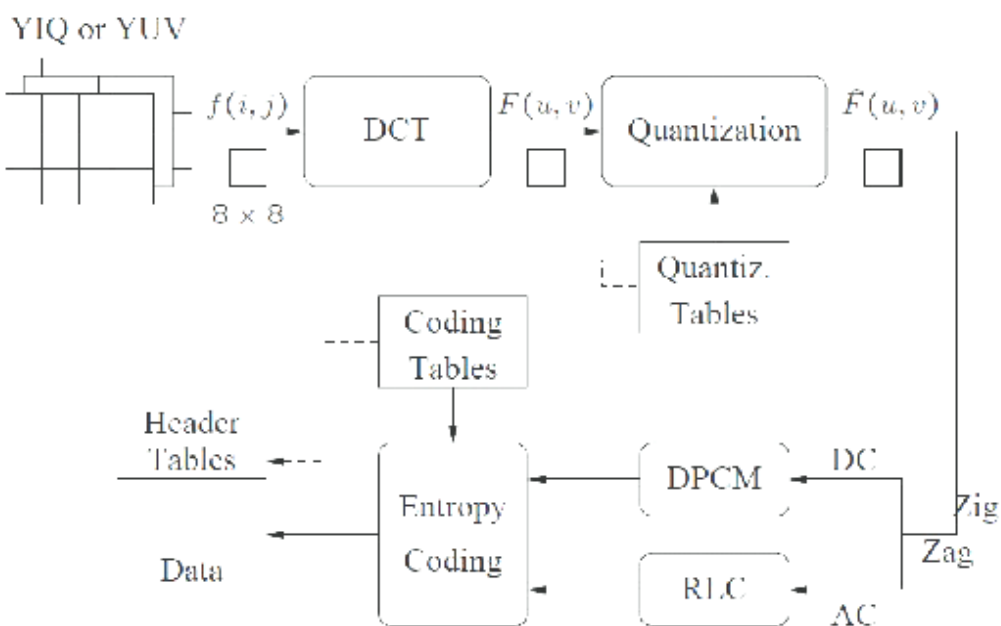
نکاتی در مورد JPEG

- محتوای مفید تصویر به طور نسبتاً آهسته‌ای در طول تصویر تغییر می کند
- انسان اتلاف در اجزای فرکانس های بالای فضا را کمتر از اتلاف در اجزای فرکانسی پایین متوجه می شود.
- هوش بینایی (دقت در تشخیص) برای شدت نور یا سطح خاکستری (سیاه و سفید) بسیار بیشتر از رنگ است.

مراحل اصلی در فشرده سازی تصویر JPEG

- تبدیل RGB به YIQ یا YUV
- اجرای تبدیل DCT بر روی بلوک های تصویر
- کوآنتیزه شدن
- پیمایش به صورت زیگزاگ
- کدگذاری آنتروپی

مراحل اصلی در فشرده سازی تصویر



DCT بلوکهای تصویر

- هر تصویر به بلوک های ۸ در ۸ تقسیم می شود .
- DCT دو بعدی روی هر بلوک تصویر $f(i, j)$ اعمال می شود و خروجی، ضرایبی از DCT به صورت $F(u, v)$ برای هر بلوک خواهد بود .
- اثر جدا کردن هر بلوک از محتوای نزدیک و مجاورش گریزناپذیر است .
(Blocking Artifact)

کوآنتیزاسیون

$$\hat{F}(u, v) = \text{round} \left(\frac{F(u, v)}{Q(u, v)} \right)$$

- دلیل اصلی برای اتلاف در فشرده سازی
- عناصر $Q(u, v)$ در پایین ترین گوشه راست مقادیر بزرگتری دارند . این عمل باعث اتلاف بیشتر در فرکانس های فضایی بالاتر خواهد شد

RLC و DPCM روی ضرایب DC و AC

RLC مقادیر $F^*(u,v)$ را به گروه های

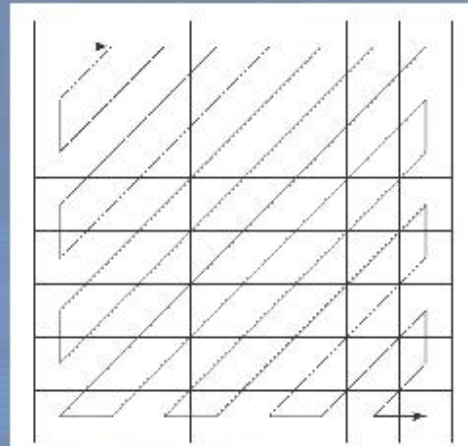
{next non-zero value#zeros-to-skip,}

تبدیل می کند. برای بدست آوردن ترتیب طولانی تر از

صفرها: اسکن زیگزاگ برای تبدیل ماتریس ۸ در ۸ مربوط

به $F^*(u,v)$ ، به بردار با طول ۶۴ استفاده می شود.

ضرایب DC با استفاده از DPCM کدگذاری می شوند



مدهای استفاده شده در JPEG

● حالت ترتیبی (Sequential Mode)

کدگذاری در یک اسکن بالا به پایین یا چپ به راست

● حالت تصاعدی (Progressive Mode)

نسخه های با کیفیت پایین تصویر را به طور سریع انتقال می دهد و بعد بقیه اطلاعات

را برای افزایش کیفیت استفاده می کند.

● حالت سلسله مراتبی (Hierarchical Mode)

تصاویر در مراحل متعددی انتقال می یابند و کیفیت را بطور تصاعدی بالا می برند.

● حالت بدون اتلاف (Lossless Mode)



مباحث این جلسه

استانداردهای فشرده سازی تصویر

JPEG2000 ●

کدگذار و رمزگشا

JPEG-LS ●

JBIG و JBIG-2 ●

فشرده سازی ویدئو

JPEG2000

- فشرده سازی بدون اتلاف و با اتلاف در یک بیت استریم
- فشرده سازی با نرخ بیت پایین
- قابلیت کار با تصاویر بزرگ
- یک معماری مشترک بازسازی
- انتقال در محیط های دارای نویز
- ارسال به صورت تصاعدی
- کدگذاری نواحی مورد نظر (ROI)
- تصویرسازی توسط کامپیوتر
- اسناد ترکیبی و مکانیزم های متاداده ای
- ۲۵۶ کانال رنگی



JPEG-LS

- فشرده سازی بدون اتلاف یا تقریباً بی اتلاف تصاویر با رنگ پیوسته
- هدف: فشرده سازی بهتر تصاویر پزشکی
- استفاده از الگوریتم LOCO-I
- Low COmplexity LOSSless Compression for Images
- مزیت اصلی: پیچیدگی کمتر

JBIG و JBIG-2

- Joint Bi-level Image Experts Group
- استانداردهای فشرده سازی تصویر دو سطحی
- هدف اصلی: ایجاد امکان مدیریت اسناد در فرم الکترونیکی
- کدگذاری تصاویر اسکن شده از متن های پرینت شده یا دست نویس و متن تولید شده توسط کامپیوتر
- قابلیت کدگذاری و دیکد کردن تصاعدی
- کدگذاری مبتنی بر مدل

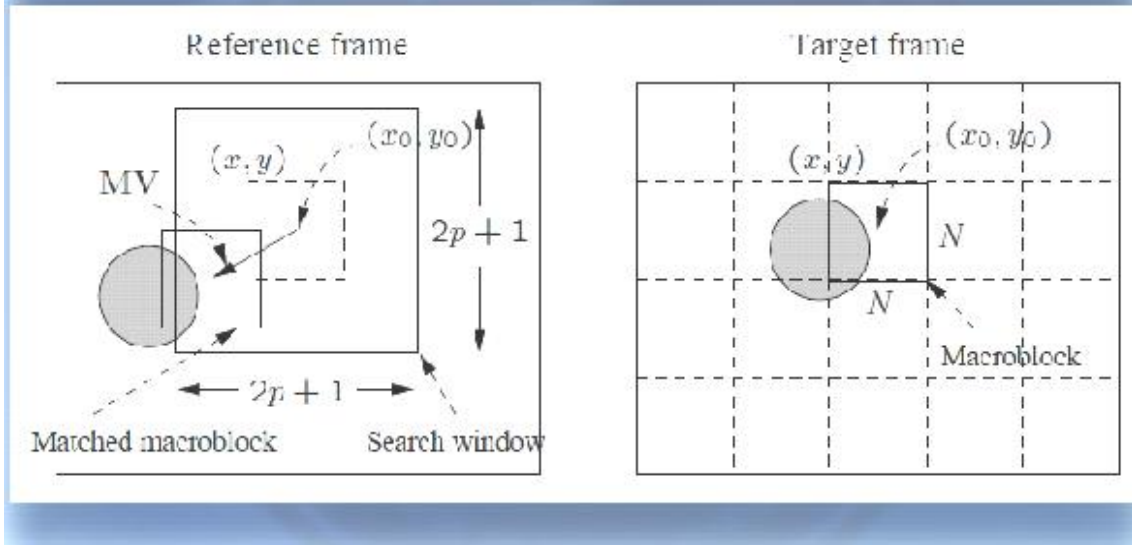
فشرده سازی ویدئو

- یک راه حل واضح برای فشرده سازی ویدئویی، رمزگذاری مبتنی بر فریم های قبلی است.
- افزونگی زمانی (Temporal Redundancy)
- فشرده سازی ویدئویی توسط جبران حرکت
برآورد حرکت (جستجوی بردار حرکت)
پیش بینی مبتنی بر جبران حرکت
به دست آوردن خطای پیش بینی (تفاوت)

فشرده سازی ویدئو

- یک راه حل واضح برای فشرده سازی ویدئویی، رمزگذاری مبتنی بر فریم های قبلی است.
- افزونگی زمانی (Temporal Redundancy)
- فشرده سازی ویدئویی توسط جبران حرکت
برآورد حرکت (جستجوی بردار حرکت)
پیش بینی مبتنی بر جبران حرکت
به دست آوردن خطای پیش بینی (تفاوت)

جبران حرکت





مباحث این جلسه

H.261 ←

H.263 ←

MPEG ←

MPEG-1

MPEG-2

H.261

- فشرده سازی مبتنی بر MC
- کاربرد : video conferencing , videophone و سایر خدمات سمعی بصری
- نرخ بیت : $p \times 64 \text{ kbps}$ (p از ۱ تا ۳۰ است).
- تاخیر کدگذار ویدئویی : کمتر از ۱۵۰ میلی ثانیه

ترتیب فریم

دو نوع فریم تصویری تعریف شده است:

- درون فریمی (I-frames)
- میان فریمی (p-frames) - پیش بینی رویه جلو

I P P P I P P P I

جبران حرکت

- تخمین حرکت: به هر ماکرو بلاک در P frame-Target، یکی از ماکرو بلاک های قبلی (I یا P frame)، که بهترین تطبیق را با آن دارد، اختصاص داده می شود.
- پیش بینی: بر اساس فریم قبلی (forward prediction) صورت می گیرد.
- خطای پیش بینی: تفاوت بین ماکرو بلاک و ماکرو بلاکی که با آن تطبیق شده است، به DCT و کدگذاری های بعدی آن ارسال می شود.

H.263

- اصلاح شده برای ویدئو کنفرانس و سایر خدمات سمعی-بصری انتقال داده شده در شبکه های تلفنی انتقال عمومی (PSTN)
- نرخ بیت: کمتر از 64 KBPS
- استفاده از کدگذاری پیش بینی برای ذخیره سازی بردار حرکتی در frame-p جهت کاهش از افزونگی زمانی
- استفاده از دقت نیم پیکسل
- درون یابی دو جهتی

MPEG

MPEG-1 ●

سازگاری با قالب های:

352 x 240 for NTSC video at 30 fps
352 x 288 for PAL video at 25 fps

نمونه برداری رنگی ۴:۲:۰

```
graph TD; VCD[VCD] --> MPEG1([MPEG-1]); DVD[DVD] --> PS([PS]); DVB[DVB] --> TS([TS]); DVB_S2[DVB-S2] --> MPEG4([MPEG-4]); PS --> MPEG2([MPEG-2]); TS --> MPEG2
```

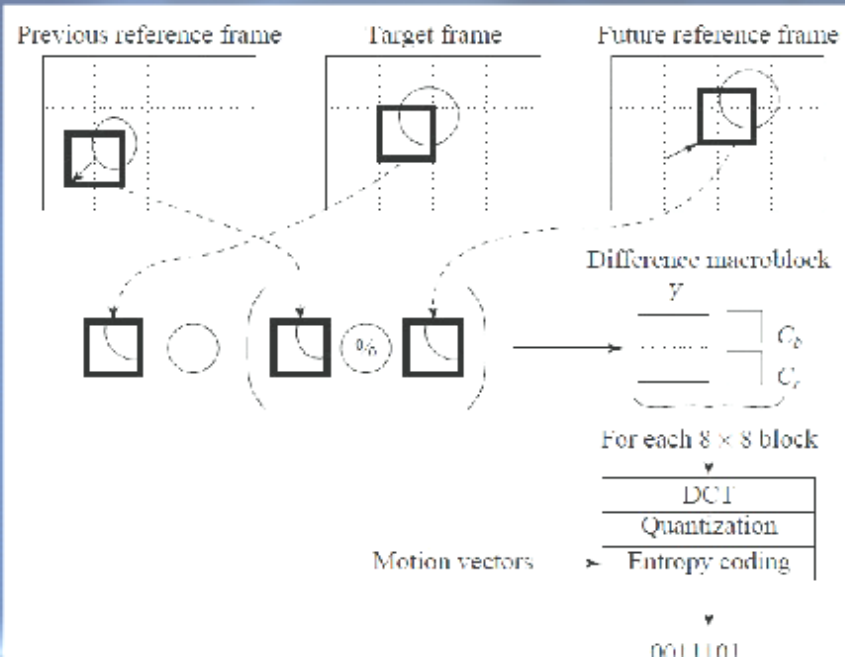
جبران حرکت در MPEG-1

Previous frame Target frame Next frame

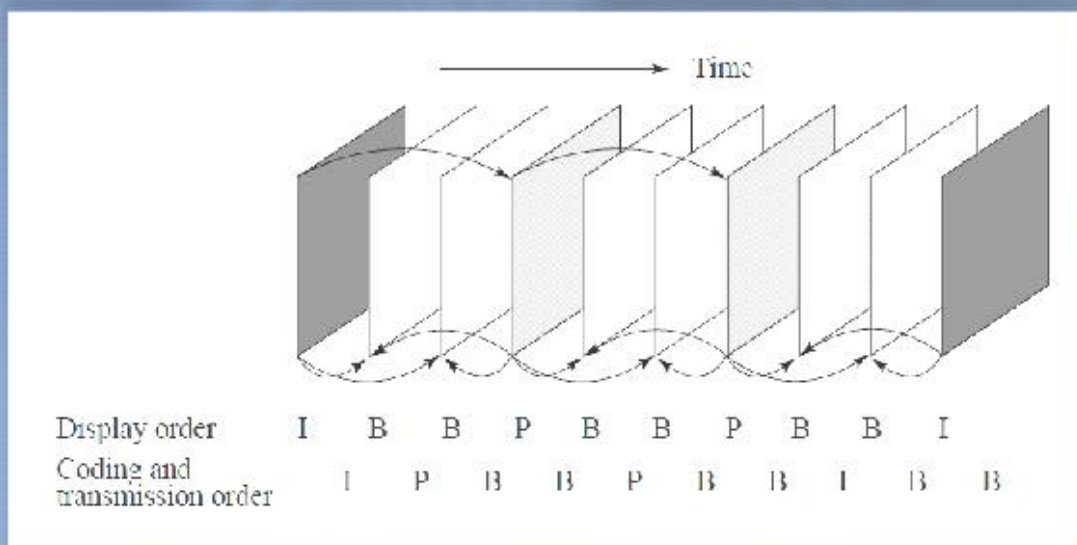
جستجوی دو طرفه

● تعریف B - frame

نمایش رمزگذاری B-frame



ترتیب فریم ها



MPEG-2

برای ویدئوهای با کیفیت بالا با جریان بیتی بیش از 4Mbps
۷ پروفایل برای موقعیت ها و برنامه های مختلف:

Simple, Main, SNR scalable, Spatially scalable, High, 4:2:2, Multiview

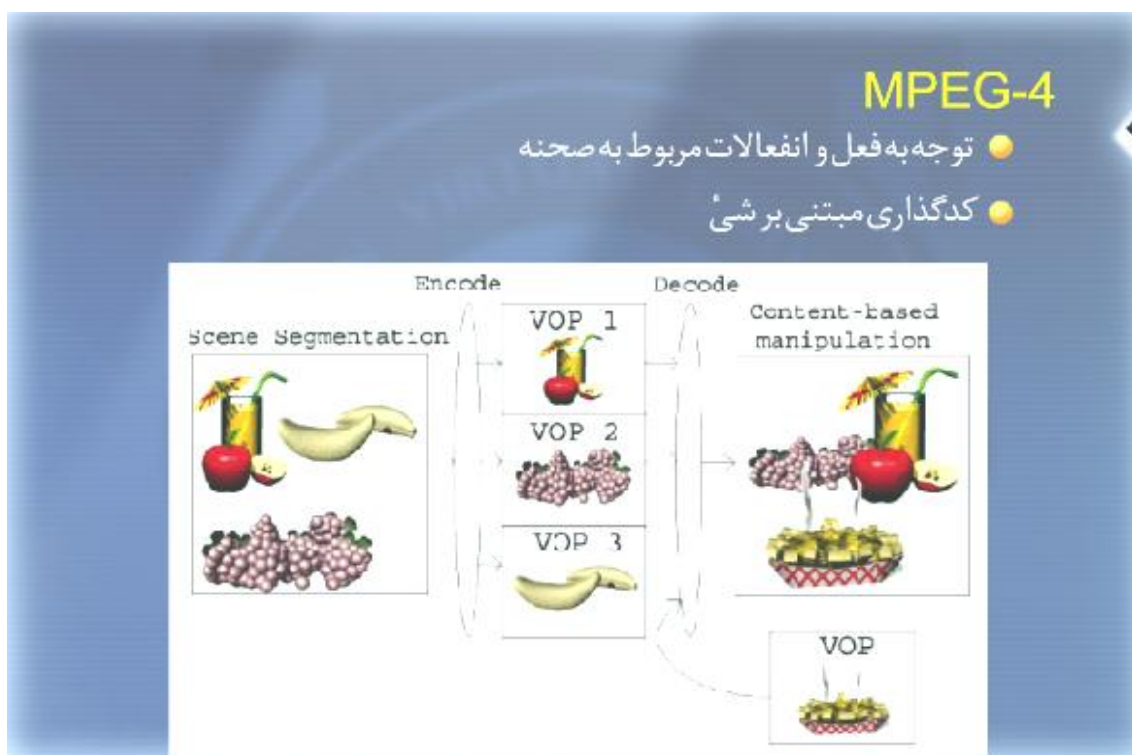
کاربرد: تلویزیون، HDTV
پشتیبانی از ویدئوی Interlaced:
top-field
bottom-field

پشتیبانی از ویدئوی Interlaced

(a) Frame-picture vs. Field-pictures

(b) Field Prediction for Field pictures





ساختار سلسله مراتبی

Video-object Sequence (VS)

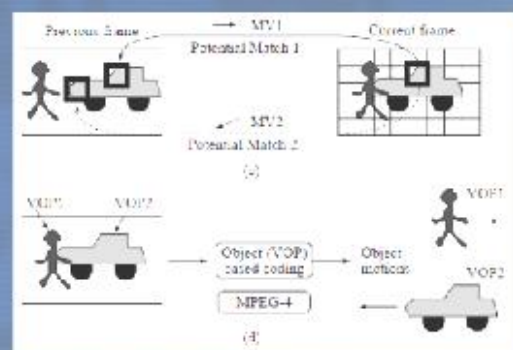
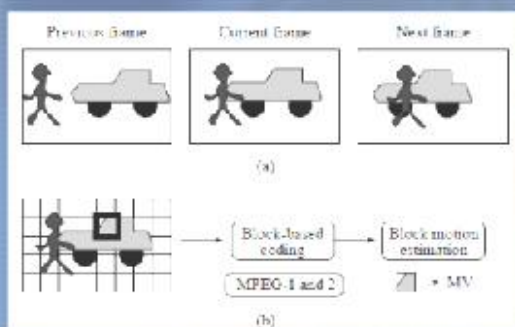
Video Object (VO)

Video Object Layer (VOL)

Video Object Plane (VOP)

Group of VOPs (GOV)

مقایسه کدگذاری مبتنی بر شیء و مبتنی بر فریم



جبران حرکت

● تصحیح حرکت در کدگذاری مبتنی بر شیء:

I-VOP

P-VOP

B-VOP

● مشکل: VOPها ممکن است دارای اشکال دلخواه باشند

● مراحل:

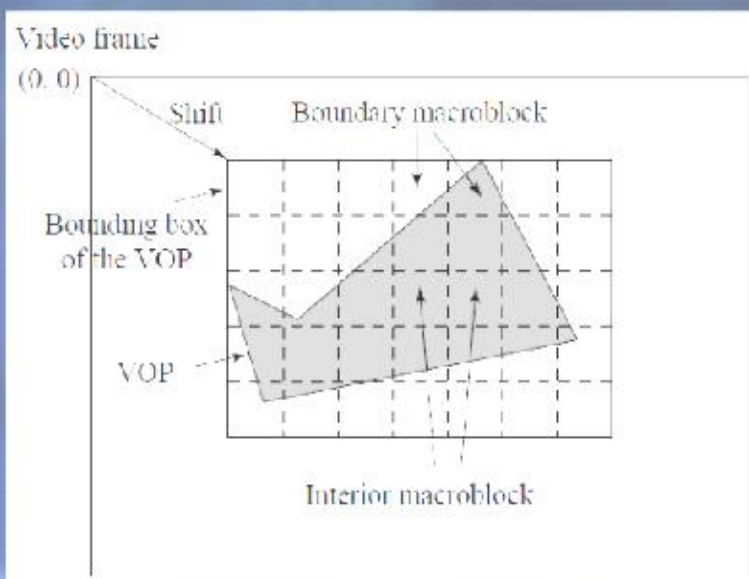
تخمین حرکت

پیش بینی مبتنی بر بردار حرکت

کدگذاری اشکال حرکت

● هر VOP به تعداد زیادی ماکرو بلاک تقسیم می شود

جبران حرکت



MPEG-7

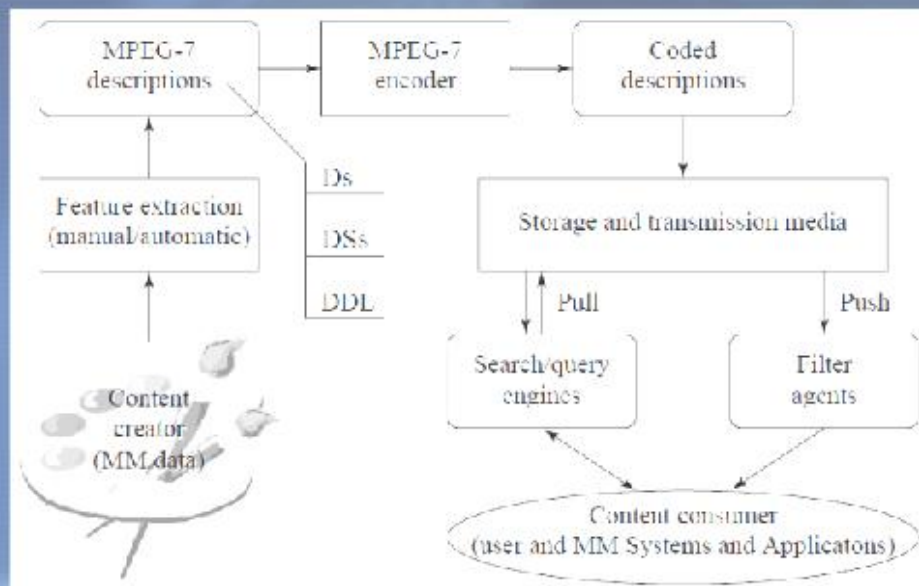
● پاسخ گویی به نیاز بازیابی صوتی-تصویری، مبتنی بر محتوا

● Multimedia Content Description Interface

● داده MPEG-7 شامل:

تصاویر ثابت، گرافیک، مدل های سه بعدی، صوت، گفتگو، اطلاعات ترکیب

کاربردهای MPEG-7



توصیف محتوای چند رسانه ای

● توصیفگرها ((Descriptors (D)

انتخاب بر اساس مقایسه عملکرد، کارایی و اندازه
سطح پایین توصیفگرهای بصری بر اساس خصوصیات بصری شامل رنگ، بافت،
شکل، حرکت، مکان و غیره

● شمای توصیف (Description Scheme)

● زبان تعریف توصیف (Description Definition Language)

زبانی مبتنی بر XML است که مجموعه ساختارهایی برای کار با انواع رسانه دارد

توصیف محتوای چند رسانه ای

● شمای توصیف (Description Scheme)

عناصر اصلی: نوع داده‌ها و ساختار ریاضی، ساختن، ابزارهای طرح
مدیریت محتوا: تشریح رسانه‌ها، ایجاد و تشریح تولید، تشریح کاربرد محتوا
تشریح محتوا: تشریح ساختاری
توصیف ادراکی
هدایت و دستیابی: خلاصه سازی، جزء بندی کردن و تجزیه، تغییر و دگرگونی محتوا
سازماندهی محتوا: جمع آوری، مدل سازی، تغییر و دگرگونی محتوا
تعامل با کاربر: تقدم کاربر



مباحث این جلسه

فشرده سازی صوت

PCM در فشرده سازی

کدگذاری تفاضلی صوت

Lossless Predictive Coding ●

Differential PCM ●

فشرده سازی صوت

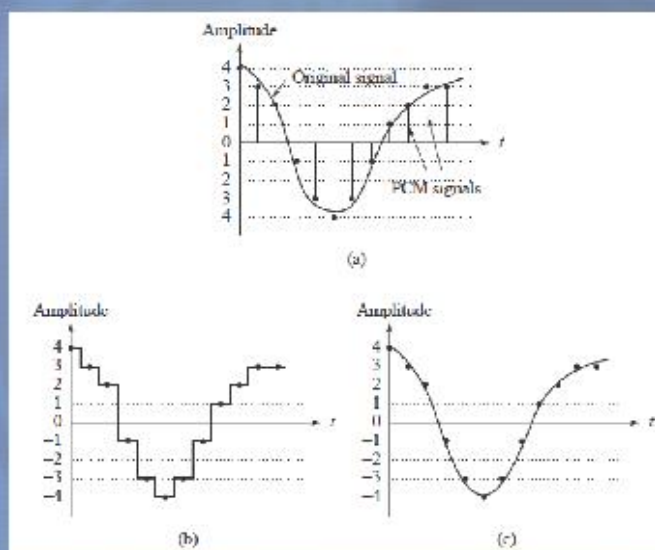
- برای فشرده سازی داده می توان یک جریان بیتی را که تعداد بیت های کمتری دارد، به متداول ترین مقادیر در سیگنال اختصاص داد
- برای سیگنال صوتی ما ابتدا PCM را برای دیجیتال سازی به کار می بریم.
- ADPCM ,DPCM ,Lossless Predictive Coding

مدولاسیون پالس کد

هر طرح فشرده سازی دارای سه مرحله است:

- داده های ورودی به نمایش جدیدی که ساده تر یا برای فشرده سازی کارآمدتر است، تبدیل می شوند.
- ممکن است بخشی از اطلاعات از دست برود. کوانتیزاسیون پراتلافترین قدم است.
- رمزگذاری که شامل انتساب یک **codeword** و تشکیل یک جریان داده باینری به هر سطح خروجی و یا هر نماد است

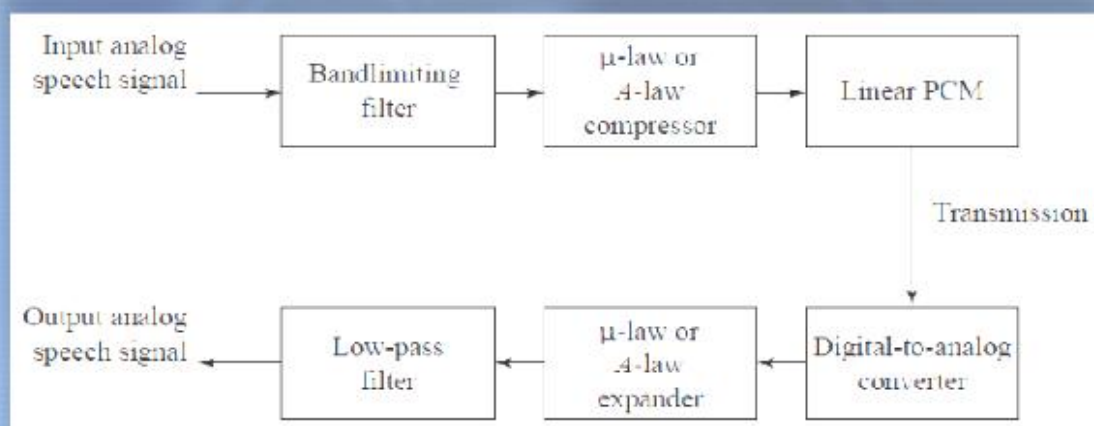
مدولاسیون پالس کد



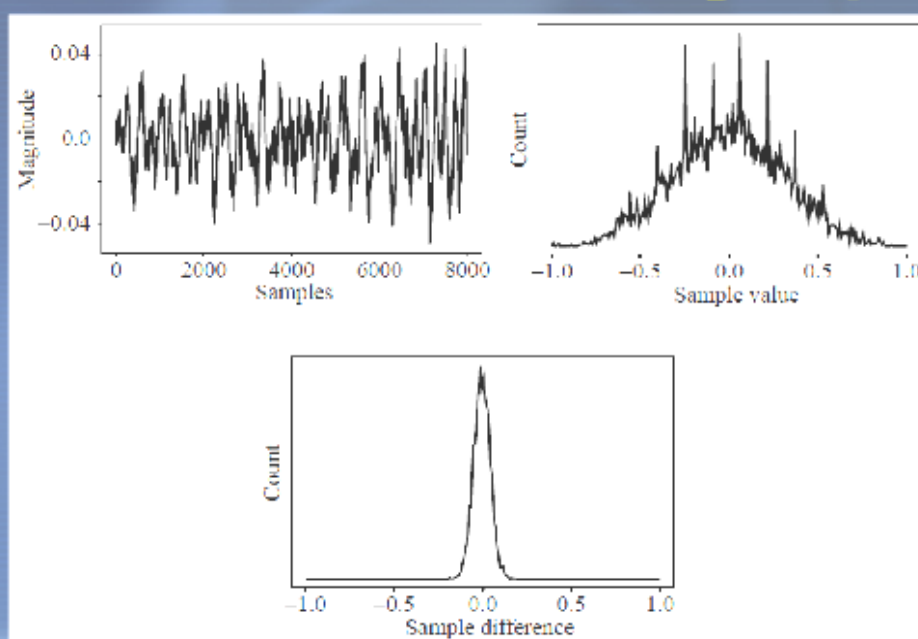
مدولاسیون پالس کد در فشرده سازی صوت

- نرخ نمونه برداری ۲۰ کیلوهرتز
- استفاده از کوانتایزر یکتواخت برای گسسته سازی
- تعداد بیت های مورد نیاز برای هر نمونه: ۸ یا ۱۲ بیت
- برای سیگنال تلفنی فرض می شود ماکزیمم فرکانس ۴ کیلوهرتز است. بنابراین نرخ نمونه برداری ۸ در نظر گرفته می شود.

انتقال سیگنال های تلفنی PCM



کدگذاری تفاضلی صدا



کدگذاری تفاضلی صدا

$$\hat{f}_n = f_{n-1}$$

$$e_n = f_n - \hat{f}_n$$

$$\hat{f}_n = \sum_{k=1}^{2 \text{ to } 4} a_{n-k} f_{n-k}$$

Lossless Predictive Coding ●

پیش بینی نمونه بعدی بر اساس نمونه های قبل

ارسال تفاضل مقدار اصلی و پیش بینی شده

Differential PCM (DPCM) ●

مشابه LPC

استفاده از کوانتایزر



سیستم های چند رسانه ای

فشرده سازی صوت



ADPCM

- Adaptive DPCM
- وفق دادن کدگذار با ورودی
- کوانتایزر
- Forward adaptive quantization
- Backward adaptive quantization

$$\hat{f}_n = \sum_{i=1}^M a_i \tilde{f}_{n-i}$$

پیش بینی کننده:

- اساس استانداردهای فشرده سازی گفتار ITU نظیر G.721, G.723, G.726 و G.727
- ورودی پیش فرض: نمونه های ۱۶ بیتی PCM کد شده با μ -law

G.726 ADPCM

- کوانتایزر یکنواخت
- اندازه های متفاوت codeword ها - نرخ ۱۶، ۲۴، ۳۲ و ۴۰ کیلوبیت در ثانیه
- در نرخ نمونه برداری ۸ کیلوبیت بر ثانیه
- تعریف a که یک ثابت افزایشنده است:
- استفاده از Backward Adaptive Quantizer

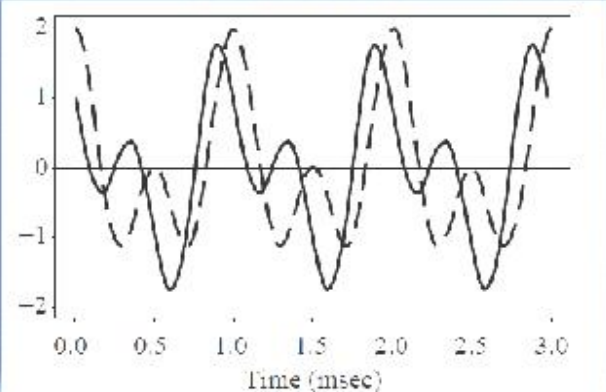
$$e_n = s_n - \hat{s}_n$$
$$g_n = e_n / a$$

Vocoder

- Voice Coder
- مورد استفاده: زمانی که سیگنال های آنالوگ مثل سیگنال های مودم استفاده نمی شوند
- مدلسازی گفتار بطوریکه ویژگی های آن در کمترین بیت های ممکن ضبط شوند
- استفاده از یک مدل شکل موج گفتار در زمان (LPC)
- سیگنال را به اجزای فرکانسی شکسته و آنها را مدلسازی می کند.

عدم حساسیت فاز

- فاز (phase): یک انتقال در آرگومان زمان در یک تابع زمان
- حساسیت فاز
- بازسازی دقیق شکل موج گفتار در واقع غیر ضروری به نظر می رسد



Vocoder

- Channel Vocoder** ●
استفاده از یک بانک فیلتر برای جدا کردن فرکانس های متفاوت
تحلیل سیگنال برای تعیین گام (pitch) گفتار
- Formant Vocoder** ●
Formant: اجزای سازنده فرکانسی که در یک نمونه گفتار وجود دارند.
رمزگذاری مهمترین فرکانس ها
- LPC Vocoder** ●
ویژگی های برجسته گفتار مستقیماً از شکل موج خطی استخراج می شود

CELP

Code Excited Linear Prediction

از یک توصیف پیچیده تر برای مشخصه های صوت استفاده می کند :

یک مجموعه کامل (Codebook) از بردارهای مشخصه گفتار، با گفتار واقعی نگاشت می شود

نرخ بیت بالا (۴.۸۰۰ تا ۹.۶۰۰ بیت در ثانیه)

صوت شبیه تریه صدای واقعی و پیوسته است

کدگذارها: LTP و STP

MPEG

- یک بانک فیلتر بر روی ورودی بکار می برد تا آن را به مولفه های فرکانساش تجزیه کند.
- بطور همزمان از یک مدل برای تخصیص قطعه بیت استفاده می شود.
- تخصیص رشته بیت به قطعات با استفاده از کوانتیزاسیون صورت می گیرد

لایه های MPEG

- هدف: یک tradeoff خوب بین کیفیت و نرخ بیت
- لایه ۱: می تواند بسته به نرخ بیت در دسترس کیفیت کاملاً خوبی داشته باشد.
- لایه ۲: پیچیدگی بیشتری دارد و برای استفاده در بخش صوتی دیجیتال پیشنهاد شد.
- لایه ۳: بیشترین پیچیدگی را دارد و در اصل برای ارسال صوتی در خطوط ISDN پیشنهاد شد.
- بیشترین افزایش پیچیدگی در رمزگذار هست، نه در رمزگشا.



مقدمه ای بر شبکه های کامپیوتری

| OSI | TCP / IP | |
|--------------|---------------------------------|--|
| Application | Application | FTP, Telnet, SMTP/MIME HTTP, SNMP, etc. |
| Presentation | | |
| Session | | |
| Transport | Transport | TCP (connection-oriented) UDP (connectionless) |
| Network | Internet | IPv4, IPv6, RSVP |
| Data link | Network access (LLC and MAC) | X.25, Ethernet, Token ring, FDDI, PPP/SLIP, etc |
| Physical | Physical | 10/100Base-T, 1000Base-T, Fibre Channel, etc. |

لایه انتقال

TCP

- اتصال گرا (Connection-oriented)
- فقط برای شبکه های بسته ای (Packet switching)
- متکی به لایه IP برای تحویل پیام به مقصد
- ارسال پیام به شکل بسته های داده، تشخیص خطا، ارسال مجدد، تسهیم، ترتیب بسته ها
- سربار بالای ارسال مجدد

UDP

- بدون اتصال
- تسهیم و تشخیص خطا
- بدون تایید
- سریع اما غیر قابل اعتماد

لایه شبکه

● سرویس ها :

◀ آدرس دهی بسته

آدرس دهی جهانی دستگاه ها در تمام شبکه های به هم پیوسته

ارسال بسته در درون شبکه های محلی یا ارسال با استفاده از hub یا ارسال نقطه به

نقطه با استفاده از switch

ارسال بسته در درون شبکه های گسترده، گذرگاه ها و یا روترها که از جداول مسیریابی استفاده می کنند .

◀ قطعه قطعه شدن بسته

● بدون اتصال

تکنیک های تسهیم

FDM (Frequency Division Multiplexing)

● کانال های متفاوت فرکانسی برای ارسال در نظر گرفته می شوند .

● مدولاسیون سیگنال آنالوگ در فرستنده صورت می گیرد .

● گیرنده از فیلترهای میانگذر برای دریافت سیگنال استفاده می کند .

تکنیک های مدولاسیون :

, Amplitude Modulation (AM)

, Frequency Modulation (FM)

و Phase Modulation (PM)

. Quadrature Amplitude Modulation (QAM)

تکنیک های تسهیم

WDM(Wavelength Division Multiplexing)

نسخه ای از FDM برای انتقال در فیبرهای نوری

کانال های متفاوت نوری

پهنای باند و ظرفیت بسیار بالا

انواع: DWDM , WWDM

تکنیک های تسهیم

(Time Division Multiplexing) TDM

انتقال مستقیم اطلاعات

تبدیل آنالوگ به دیجیتال (PCM)

انتقال در بعد زمان

دو نوع همزمان و غیرهمزمان

انواع: Synchronous TDM , Asynchronous TDM



سیستم های چند رسانه ای

شبکه های چند رسانه ای

مباحث این جلسه

خصوصیات داده های چند رسانه ای

کیفیت ارسال داده های چند رسانه ای

مدل های کاری چند رسانه ای

پهنای باند مورد نیاز برای کاربردها

دامنه تحمل تاخیر

کیفیت سرویس دریافتی

کیفیت عملکرد برای پروتکل IP

انتقال اولویت بندی شده

خصوصیات داده های چند رسانه ای

- حجم و تراکم بالا (Voluminous): به نرخ بیت بالایی احتیاج دارند.
- بلادرنگ و تعاملی (Real-time and interactive): به همزمانی بالا میان صدا و تصویر و کمترین تاخیر برای lip-sync احتیاج دارند.
- Sometimes bursty: نرخ داده به شدت نوسان دارد.

کیفیت ارسال داده های چند رسانه ای

کیفیت سرویس (QoS) به عوامل گوناگونی بستگی دارد:

- نرخ داده (Data Rate)
- تاخیر (Latency)
- اتلاف بسته یا رخ دادن خطا (Packet Loss or Error)
- حرکت نامنظم اتفاقی (Jitter)
- انحراف همزمانی (Sync Skew)

مدل های کاری چند رسانه ای

- **پلاذرنگ (Real-time)**: ترافیک دو طرفه، تاخیر کمتر و حرکت نامنظم اتفافی پایین.
- **داده های اولویت دار (Priority Data)**:
ترافیک دو طرفه، تاخیر کمتر و از بین رفتن داده ها بصورت اندک، با ارسال های اولویت دار.
- **نقره ای (Silver)**: تاخیر و حرکت نامنظم اتفافی متوسط، ترتیب و همگام سازی بالا، ترافیک یک طرفه.
- **Best Effort**: به هم زمانی نیازی نیست.
- **برنزی (Bronze)**: ضمانتی برای ارسال وجود ندارد.

پهنای باند مورد نیاز برای کاربردها

| Application Speed | Requirement |
|------------------------------|-------------------|
| Telephone | 16 kbps |
| Audio-conferencing | 32 kbps |
| CD-quality audio | 128-192 kbps |
| Digital music (QoS) | 64-640 kbps |
| H. 261 | 64 kbps - 2 Mbps |
| H. 263 | < 64 kbps |
| DVI video | 1.2 - 1.5 Mbps |
| MPEG-1 video | 1.2 - 1.5 Mbps |
| MPEG-2 video | 4 - 60 Mbps |
| HDTV (compressed) | > 20 Mbps |
| HDTV (uncompressed) | > 1 Gbps |
| MPEG-4 video-on-demand (QoS) | 250 - 750 kbps |
| Videoconferencing (QoS) | 384 kbps - 2 Mbps |

دامنه تحمل تاخیر

دامنه تحمل تاخیر و حرکت نامنظم اتفاقی در کاربردهای صدا و تصویر دیجیتال

| Application | Avg Latency Tolerance (msec) | Avg Jitter Tolerance (msec) |
|------------------------------|------------------------------|-----------------------------|
| Low-end videoconf. (64 kbps) | 300 | 130 |
| Compressed voice (16 kbps) | 30 | 130 |
| MPEG NTSC video (1.5 Mbps) | 5 | 7 |
| MPEG audio (256 kbps) | 7 | 9 |
| HDTV video (20 Mbps) | 0.8 | 1 |

کیفیت سرویس دریافتی

هرچند کیفیت عملکرد به طور معمول با پارامترهای تکنیکی اندازه گیری می شود، با این حال نتیجه مجموعه ای از کارایی هایی است که درجه رضایت کاربر از عملکرد را تعیین می کند.

● نظم از تاخیر مهم تر است.

● صحت زمانی از صدا و کیفیت تصویر مهم تر است.

کیفیت عملکرد برای پروتکل IP

- IP یک تکنولوژی ارتباطی Best Effort است و رسیدن به کیفیت عملکرد بر روی آن با متدهای مسیریابی موجود بسیار مشکل است.
- پهنای باند زیاد، کیفیت عملکرد را بهبود می بخشد اما در هر قسمت از یک شبکه مرکب قابل دسترسی نیست.
- DiffServ: دسته بندی بسته ها در شبکه
- ToS در IPv4
- کلاس ترافیک برای IPv6
- MPLS (Multiple Protocol Label Switching)
- تسهیل پیوستن IP به تکنولوژی های لایه OSI

انتقال اولویت بندی شده

- برای کم کردن میزان مشاهده از بین رفتگی (اتلاف زیاد بسته ها یا میزان خطا) در ازدحام شبکه ای به کار می رود.
- اولویت بندی برای انواع رسانه ها
- اولویت بندی برای صداهای فشرده سازی نشده
- اولویت بندی برای تصاویر
- اولویت بندی برای ویدئوهای فشرده شده



مباحث این جلسه

- چند رسانه ای روی شبکه های ATM
- انتقال MPEG-2
- انتقال MPEG-4
- رسانه بر حسب تقاضا (Media-on-Demand(MOD))

چند رسانه ای روی شبکه های ATM

● شبکه ATM

- به پالس های ساعت وابستگی ندارد.
- اتصال گرا است.
- اطلاعات به صورت بسته ارسال می گردد.
- سرعت بالایی دارد.
- بسته ها را در یک مسیر و به ترتیب ارسال می کند.
- کیفیت بالا است.

● نرخ داده های ویدئویی در ATM

- نرخ داده ثابت (CBR)
- نرخ داده متغیر (VBR)
- rt-VBR (نرخ داده متغیر همزمان) ✓
- nrt-VBR (نرخ داده متغیر غیر همزمان) ✓
- نرخ داده در دسترس (ABR)
- نرخ داده غیر مشخص (UBR)

AAL: لایه انطباق در ATM

- تبدیل فرمتهای مختلف داده های کاربر به جریان های داده ATM و بالعکس

- انواع مختلف پروتکل های AAL

AAL1: پشتیبانی همزمان، نرخ داده ثابت

AAL2: طراحی شده برای نرخ داده های متغیر صدا و تصویر فشرده شده

AAL3/4: نرخ داده های متغیر، پشتیبانی از ارسال اتصال گرا و بدون اتصال

AAL5: ارسال داده های چند رسانه ای، پشتیبانی از تمام گروه های داده ها

و سرویس های تصویری

ATM در MPEG-2

- MPEG-2 از AAL5 استفاده می کند.
- در هنگام ساخت یک کانال اتصال مجازی مقادیر QoS زیر باید مشخص باشند:

حداکثر تأخیر انتقال سلول

حداکثر تأخیر

میزان از بین رفتن سلول

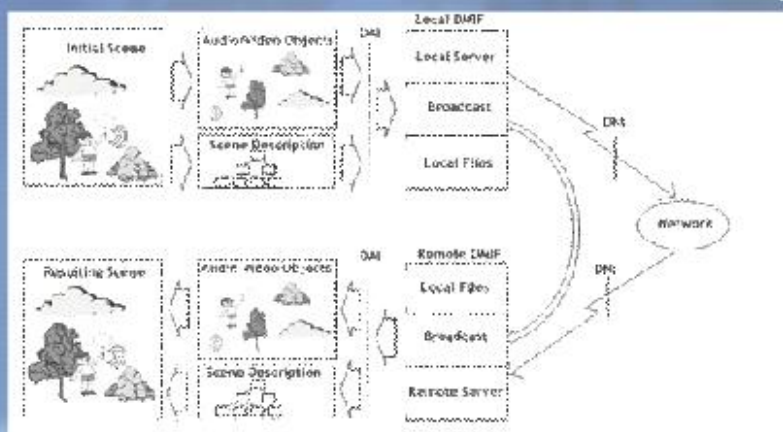
میزان خطای سلول

میزان انسداد سلولهای دارای خطای شدید

انتقال MPEG-4

● DMIF

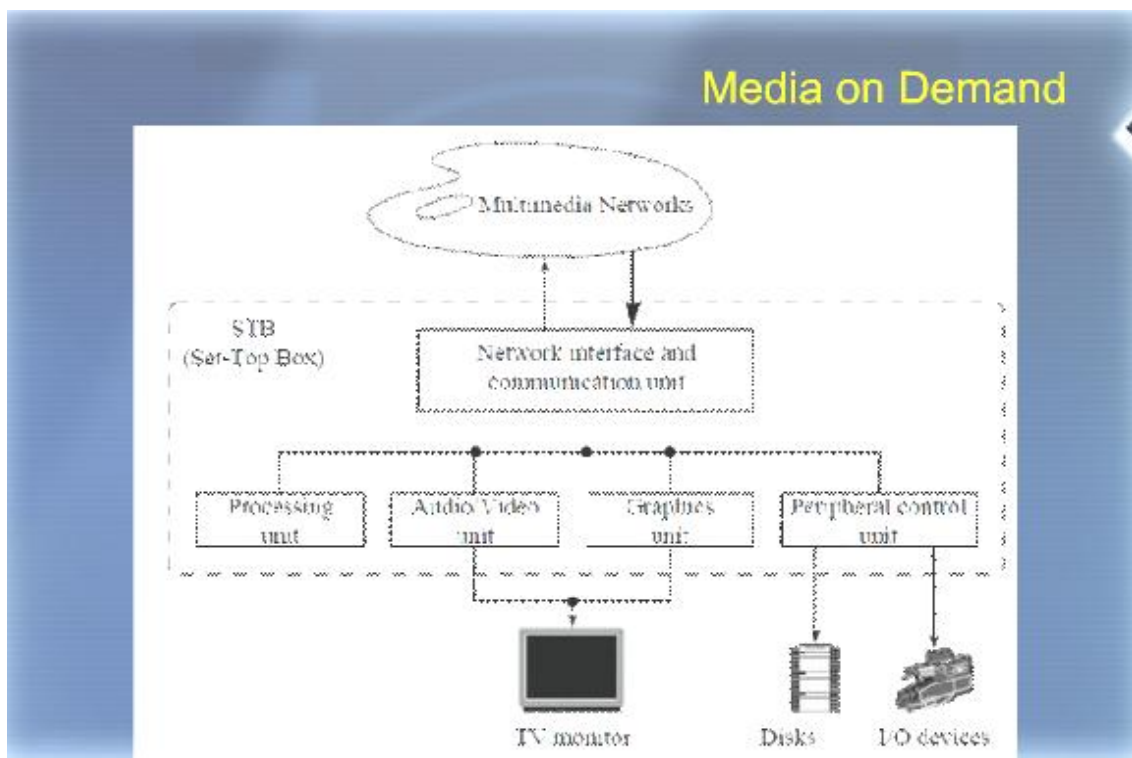
کنترل دسترسی شبکه
پخش رسانه (کابلی و ماهواره ای)
رسانه محلی بر روی دیسکها



Media on Demand

● خدمات (ITV) Interactive TV :

- تلویزیون
- ویدئو بر حسب تقاضا
- خدمات اطلاعاتی
- سرگرمی های تعاملی
- تجارت الکترونیک
- دسترسی به کتابخانه های دیجیتالی و کالاهای آموزشی







The screenshot shows a window titled "littlequiz.a4p" with a "Level 1" indicator. It features a flowchart starting with a "welcome screen" box, leading to a diamond-shaped decision box labeled "U". From the diamond, three paths lead to boxes labeled "q1", "q2", and "q3", which then converge into a "results" box. Below the flowchart is a code block with the following content:

```
-- load an MPEG file
extFileName of MediaPlayer "theMpegPath" =
  "c:\windows\media\hcm33.mpg";
-- play
extPlayCount of MediaPlayer "theMpegPath" = 1;
-- put the MediaPlayer in frames mode (not time mode)
extDisplayMode of MediaPlayer "theMpegPath" = ;
-- if want to start and end at specific frames
extSelectionStart of MediaPlayer "theMpegPath" = 103;
extSelectionEnd of MediaPlayer "theMpegPath" = 1997;
-- start playback
get extPlay() of MediaPlayer "theMpegPath";
```

تالیف چند رسانه ای

Multimedia Authoring ●

تولید محصول چند رسانه ای ●

راهکارهای تالیف چند رسانه ای: ●

زبان های اسکریپتی

نمایش به صورت صفحات اسلاید

روش سلسه مراتبی

نمایش فلوجارتی / نمادی



The screenshot shows a window titled "Quest Frame Edit - Check Calc". It has a menu bar with "Frame", "Edit", "View", "Object", and "Help". Below the menu is a toolbar with icons for "Graphics", "Animation", "Audio/Video", "Controls", and "Interact". The main area contains a script with the following text:

```
// This example starts up the Windows Calculator.
// If the user minimizes the calculator and then
// tries to start it up again, the Calculator is
// brought to the top instead of starting up a new instance of it.
// To use this in your title, copy the following three lines into
// your frame as well as the watch for. Then modify the watch
// for to watch for whatever button or event you have designated
// will launch the calculator. You do not need to use the
// small program text at the title design level for this to work
|
WORD wStatus; //Copy these 3 lines of
HWND hWnd; //code into your frame
char szMsg(80);

Graphic the (calc bin) "openCalc"
Panel

//Copy this watch for into your frame and modify it to watch
//for whatever event you would like to launch the calculator
Watch for: "OpenCalc" LButtonClicked then...
```

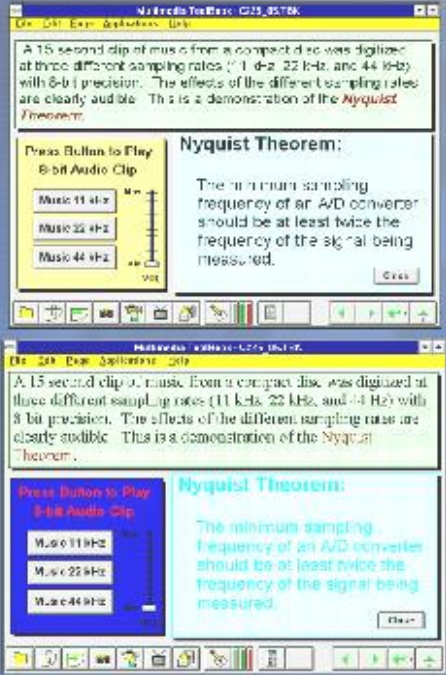
تالیف چند رسانه ای

نمایش در قالب فریم

نوع کارت / متن

Cast / Score / Scripting





ارائه چند رسانه ای

سبک های گرافیکی

- قوانین رنگ ها
- قلم ها
- استفاده از تفاضل رنگ ها
- سیستم رنگی (R,G,B)
- رنگ مناسب پس زمینه

ارائه چند رسانه ای

گذارهای ویدئویی

نشان دادن تغییر صحنه

انواع گذارها:

- برش (Cut)
- Wipe
- ادغام (Dissolve)
- Cross Dissolve
- Dither Dissolve

$$D = (1 - \alpha(t)) \cdot A + \alpha(t) \cdot B$$

ارائه چند رسانه ای

Video L

Video R

Viewport

Max X

```
for t in 0..t_max
  for x in 0..x_max
    if ( x / x_max < t / t_max )
      R = R_L ( x + x_max * [1 - t / t_max], t )
    else
      R = R_R ( x - x_max * t / t_max, t )
```

ارائه چند رسانه ای

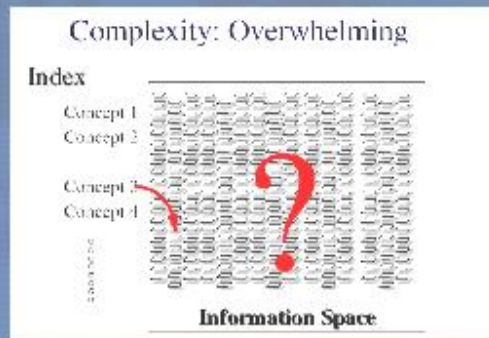
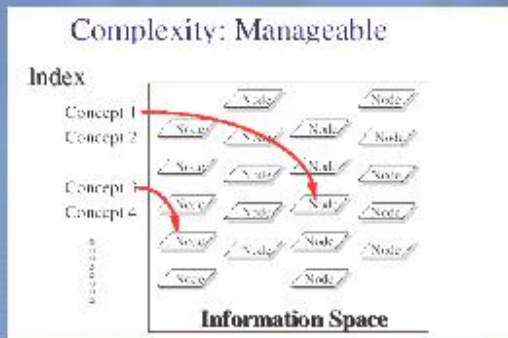
سکوهای کامپیوتری
فرمت و تفکیک پذیری ویدئو
نیازهای سخت افزاری
روش های ارائه



ابرتصویر (Hyper Image)

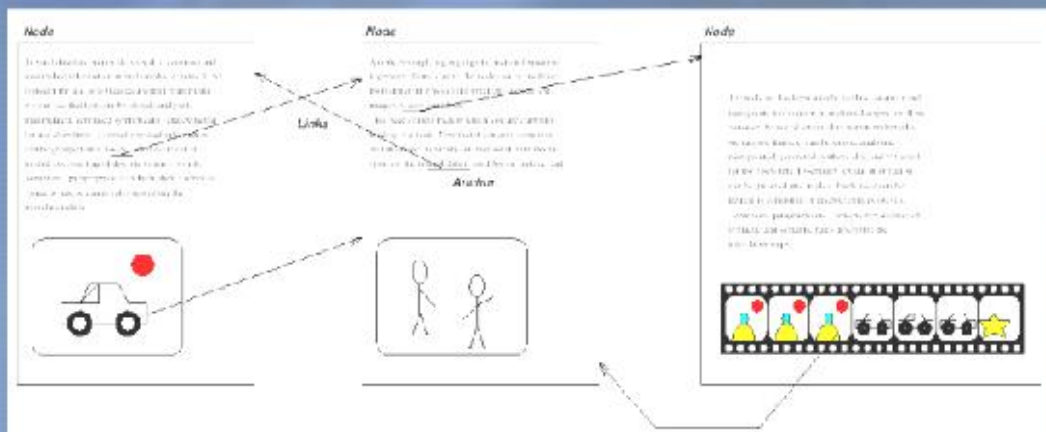
● خطی سازی (Linearization)

● خارجی سازی (Externalization)



ابرتصویر (Hyper Image)

ابرمتن ← ابرتصویر



زبان مدل سازی واقعیت مجازی

Virtual Reality Modeling Language

یک زبان مفسر است.

هدف آن امکان ایجاد و قرار دادن اشیاء رنگی در محیط سه بعدی است.

VRML 1.0: امکان ایجاد بسیاری از اشکال ساده سه بعدی را می دهد.

VRML 2.0:

امکان تعامل با محیط را فراهم می سازد.

انیمیشن و صداگذاری را میسر می کند.

اشیاء جدید دیگری جهت تولید آسان تر محیط مجازی اضافه شده است.

پشتیبانی از JavaScript و Java

اجزای VRML

Shape Node ●

Material Node ●

Texture Node ●

Image Texture

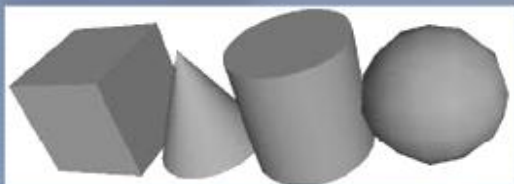
Movie texture

Pixel texture

Background Node ●

Panorama Node ●

Time Sensor ●



دنیای VRML

● مکان و جهت دید (Orientation)

● محدوده دید دوربین (Field-of-View)

● سه نوع منبع نور:

نور مستقیم (Directional Light)

نور نقطه ای در تمام جهات (Point Light)

نور نقطه ای (Spot Light)

● متحرک سازی و تعامل



چند اصطلاح

HDTV

HDMI

3D TV

Multimedia DataBase

Augmented Reality

Lingo