باسمه تعالی

سیستم‌های چندرسانه‌ای (۳۴۲-۴۰)
دانشکده مهندسی کامپیوتر
ترم بهار ۱۳۸۵
دکتر حمیدرضا ربیعی

تکلیف شماره ۵: تکنولوژی ویدیو، از آنالوگ تا دیجیتال

۱- مقدمه

در این آزمایش، مبانی تکنولوژی ویدیو را یاد می گیرید. در ابتدا اصول گرفتن و نمایش ویدیو را با توضیح عملکرد تلویزیون آنالوگ فرا می گیریم. همان طور که خواهیم دید، ویدیوهای آنالوگ به فرم اسکن Raster گرفته، ذخیره و انتقال داده می شوند. در ابتدا سیستم های تلویزیون تک رنگ را توضیح می دهیم و سپس به سراغ سیستم های تلویزیون رنگی می رویم. به شما دیدی درباره چگونگی انتخاب پارامترهای اساسی مشخصی همچون نرخ قاب و نرخ خط می دهیم.

محتوای طیفی سیگنال تلویزیون را بررسی می کنیم و چگونگی مالتی پلکس شدن مولفه های مختلف سیگنال برای تشکیل یک سیگنال مرکب را شرح می دهیم. در انتها ویدیوی CCIR601 و نوع دیجیتال شده سیگنال تلویزیون رنگی آنالوگ را توضیح می دهیم. ملاحظاتی را که برای انتخاب پارامترهای دیجیتالیزاسیون گوناگون داشته‌ایم ارائه می کنیم. همچنین چندین فرم دیگر ویدیوی دیجیتال شامل تلویزیون کیفیت بالا (HDTV) را نیز شرح خواهیم داد. همچنین استانداردهای فشرده سازی برای کاربردهای مختلف و فرمت های ویدیوی متناظرشان را شرح خواهیم داد.

۲- تکنولوژی تلویزیون تک رنگ

در سیستم تلویزیون آنالوگ حاضر، صفحه ویدیوی متحرک پیوسته در زمان رقاب های گسسته گرفته و نمایش می شود ولی نرخ قاب به اندازه‌ای سریع است که چشم آن را مانند حرکت پیوسته احساس می کند، نرخ قاب لازم به درخشندگی نمایشگر وابسته است. برای دید در شرایط نور اتاق معمولاً ٦٠ قاب بر ثانیه کافی است. هر قاب با مجموعه ای از خطوط نسبتاً افقی نشان داده می شود. بنابراین سیگنال ویدیو که ماهیتاً ۳ بعدی است (زمانی، عمودی و افقی) می تواند در یک سیگنال یک بعدی که با نام اسکن Raster شناخته می شود نشان داده شود. بنابراین اسکنRaster تسلسلی از خطوط پشت سر هم قاب ها می باشد. این روشی است که سیگنال ویدیو گرفته، انتقال داده و نمایش داده می شود.

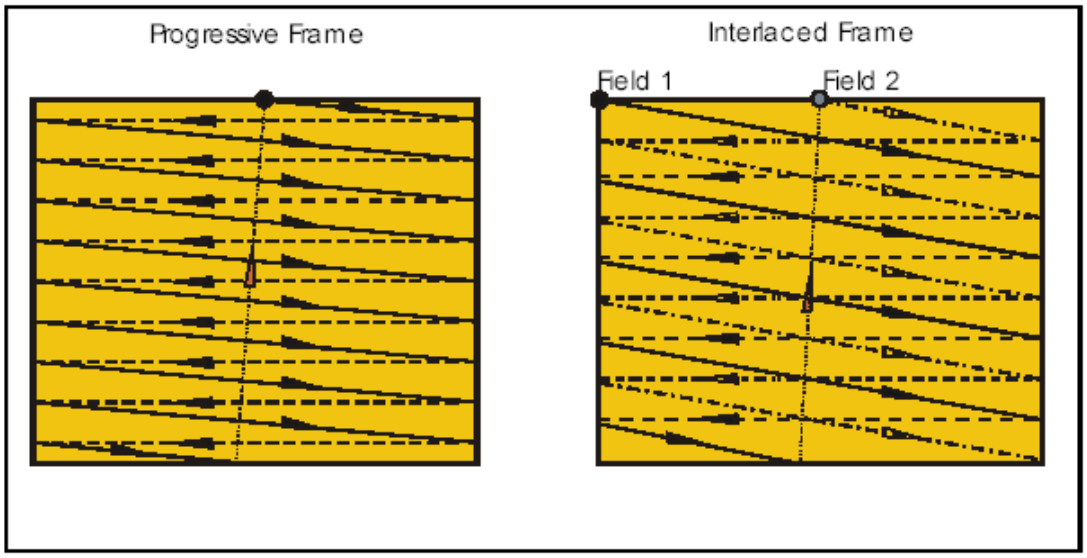
گرفتن تصویر با یک دوربین ویدیویی (یا نمایش با یک CRT) با اسکن کردن یا خواندن اطلاعات روشنایی مربوط به بالای تصویر آغاز می شود. سیگنال جاروب افقی نقطه اسکن را افقی به سمت راست تصویر جا به جا می کند. در همان زمان سیگنال جاروی عمودی نقطه اسکن را در نرخی بسیار آهسته تر از جاروب افقی به سمت پایین می برد. خط اسکن نتیجه شده از نظر افقی به سمت راست و از نظر عمودی به طور جزیی به سمت پایین زاویه دارد. وقتی اسکن به سمت راست تصویر رسید، جاروب افقی سریعاً به سمت چپ بر می گردد (زمان برگشت افقی). وقتی نقطه اسکن کننده به لبه پایین راست می رسد سیگنالی جاروب افقی سریعاً به بالای تصویر بر می‌گردد تا بتواند دوباره اسکن را شروع کند. بازگشت عمودی عبارت مورد استفاده برای این کار است و بازه بازگشت عمودی زمان لازم برای آن است. در خلال دو زمان بازگشت عمودی و افقی اطلاعات تصویر وجود ندارد.

۲-۲- خطوط بینابینی

قاب هایی که ۳۰ بار در ثانیه نمایش داده می شوند در صورت وجود اشیا با سرعت زیاد در تصویر چشمک ایجاد می کنند. همان طور که قبلاً گفتیم، چشم انسان می تواند چشمک زدن تصویر را در هر چیزی که با کمتر از ٦٠ قاب در ثانیه نمایش داده شود احساس کند. با تقسیم هر قاب به دو میدان که اولی از خطوط اسکن فرد و دومی از خطوط اسکن زوج تشکیل می شود انتقال دهنده می تواند هر میدان یا هر خط تصویر را در هر $\frac{1}{60}$ ثانیه بفرستد.

این کار از نظر بینایی مثل نمایش در هر $\frac{1}{60}$ ثانیه می ماند در حالی که کل تصویر هر $\frac{1}{30}$ ثانیه نشان داده شود. خطوط میانی چشمک زدن تصویر را حداقل می کند. مشکل چشمک زدن در تصاویر با خطوط افقی نازک باقی می ماند. چون خطوط فقط هر $\frac{1}{30}$ ثانیه نشان داده می شوند چشم، چشمک زدن را تشخیص می دهد. بیشتر تصاویر تلویزیونی عاری از خطوط نازک افقی هستند.

اگرچه خطوط بینابینی نرخ نمونه برداری زمانی ویدیو را در هر میدان افزایش می دهد، رزلوش عمودی را به نصف یعنی ٢٦٢/٥ خط بر میدان کاهش می دهد. بنابراین خطوط بینابینی راه حلی است که بین رزلوشن عمودی و زمانی مصالحه برقرار می کند. وقتی تلویزیون آنالوگ توسعه یافت، این راه حل برای تکنولوژی آن روز مناسب بود. ولی این کار، آرتیفکت هایی مانند لبه های عمودی دندانه دندانه هنگام حرکت افقی یک خط عمودی و خمیدگی خط را ایجاد می کند.

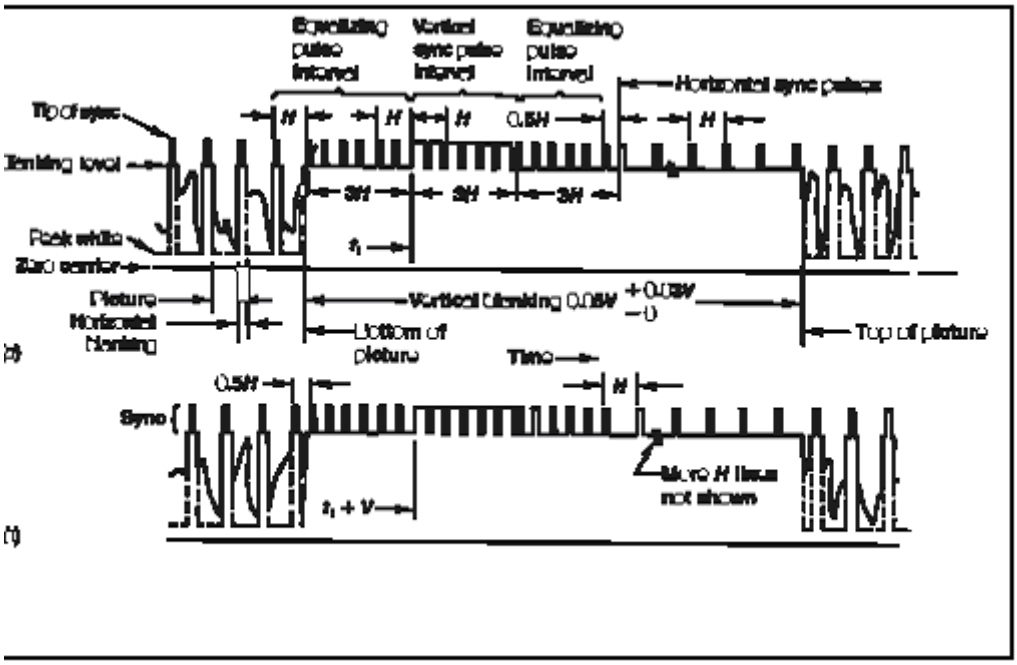


شکل ۱: قابهای Progressive و Interlaced

۲-۳- پارامترهای سیستم NTSC

در یک سیستم NTSC، ۳۰ قاب بر ثانیه استفاده می شود. هر قاب از ٥٢٥ خط اسکن تشکیل شده است، بازه خطی۶۳.۵ میلی ثانیه است.

بازگشت افقی ۱۰ms وقت می گیرد. بنابراین زمان واقعی اسکن هر خط ۵۳/۵ms است. بازگشت عمودی میدان های مجاور ۱۳۳۳ میکروثانیه زمان می گیرد که معادل زمان ۲۱ اسکن خط در هر میدان است. بنابراین، تعداد خطوط فعال۴۸۳ تا در هر فریم است. در خلال زمان بازگشت عمودی اطلاعات روشنایی نوشته نمی شود بنابراین فقط ٤ خط اطلاعات واقعی در هر قاب وجود دارد. شکل ۲ شکل موج های نمونه متناظر با ویدیویی NTSC را نشان می دهد.



شکل ۲ : شکل موج های نمونه متناظر با ویدیویی NTSC

۲-۴- پهنای باند سیگنال

پهنای باند اسکن Raster از روی نرخ خط قاب تخمین است. اول از همه، حداکثر فرکانس عمودی موقعی روی می دهد که خطوط سیاه و سفید در یک قاب پشت سر هم تکرار شوند. حداکثر فرکانس قابل نمایش در روی سیستم معمولاً کمتر از این حد تئوری است. فاکتور تضعیف Kell که باK نشان داده می شود به توابع دهانه نمایشگر و دوربین بستگی دارد. دوربین های تلویزیون معمولی K برابر ۷ دارند. حداکثر فرکانس عمودی با در نظر گرفتن فاکتور Kell به این صورت حساب می شود:

$$f_{v,\max} = K f_{s,y}/2$$

با فرض اینکه حداکثر فرکانس افقی شبیه عمودی باشد داریم:

$$f_{\max} = f_{h,\max}/T'_1 = \mathrm{IAR}\, K f_{s,y}/2\, T'_1 \,\mathrm{Hz}.$$

که هر خط در هر53.5 میلی ثانیه اسکن می شود و حداکثر فرکانس سیگنال ۱ بعدی raster برابر 4.2 MHZ است.

۲-۵- استفاده از سیگنال مرکب

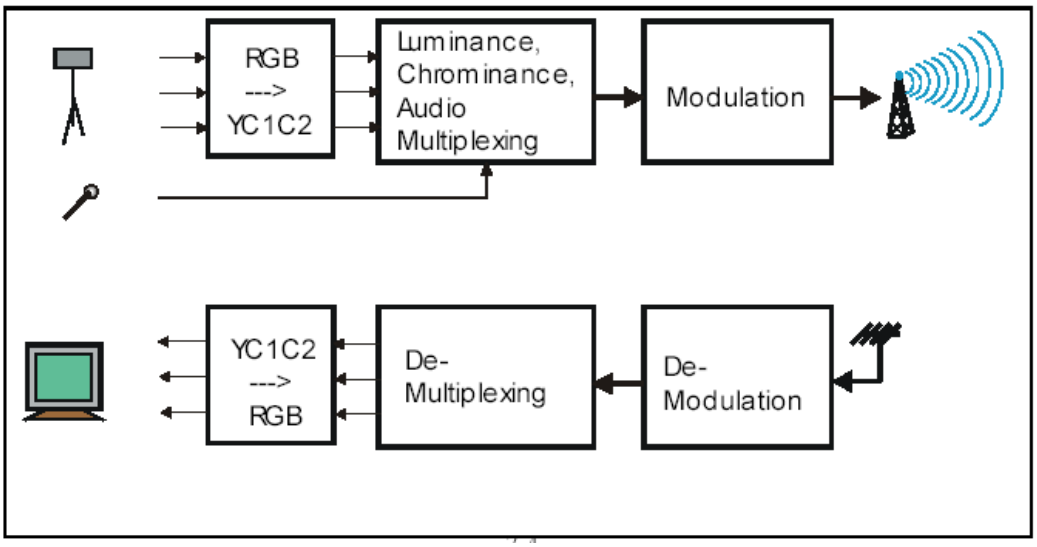
ویدیو هر $\frac{1}{30}$ ثانیه از انتهای منبع انتشار خوانده شده و در صفحه دریافت نوشته می شود.

سیگنال های جاروب عمودی و افقی و یک سیگنال که انتهای دریافت، شروع نوشتن اطلاعات روشنایی و پایان نوشتن اطلاعات در خلال زمان های بازگشت را می دهد همراه سیگنال روشنایی و به صورت مرکب ارسال می شود. سیگنال مرکب به وسیله دریافت کننده تمام اطلاعات موردنیاز برای همزمانی با منتشر کننده را در بر دارد. سیگنال مرکب، اطلاعات صدا را نیز با ویدیو همراه می‌کند. سیگنال صدا با فرکانس ٤/٥MHz در سیگنال ویدیو که پهنای باند ٤/٢MHz معدوله شده و جمع می شود.

در حالی که اسکن کردن و خواندن تصویر در نقطه انتشار روی می دهد، تلویزیون یا نمایشگر شما دقیقاً با سیگنال انتشار همزمان است. ایستگاه دریافت کننده سیگنال مرکب را کد گشایی می‌کند و اطلاعات روشنایی را به صفحه تلویزیون شما منتقل می کند و اطلاعات همزمانی عمودی و افقی را برای هماهنگی به کار می گیرد. سیگنال صدا به وسیله یک فیلتر میانگذر از سیگنال مرکب جدا می شود.

۳- از تک رنگ به تلویزیون رنگی

توضیحات بالا مربوط به تلویزیون سیاه و سفید بود. یکی از مهمترین محدودیت های تلویزیون رنگی منطبق بودن آن با سیستم های قبلی سیاه و سفید است. اولاً پهنای باند سیگنال ویدیویی رنگی باید داخل باند اختصاص داده شده به سیگنال تک رنگ جای بگیرد. ثانیاً، تمام سیگنال‌های رنگی باید طوری در یک سیگنال مرکب جای بگیرند که گیرنده سیاه و سفید بتواند مولفه روشنایی را استخراج کند. طراحی سیستم های تلویزیون رنگی به طوری که محدودیت های گفته شده را ارضا کند یکی از بزرگترین اختراعات قرن بیستم بوده است. شکل ۳، مراحل پردازش اصلی سیگنال تلویزیون رنگی شامل تولید، انتقال و دریافت را نشان می دهد.



شکل ۳: مراحل پردازش اصلی سیگنال تلویزیون رنگی شامل تولید، انتقال و دریافت

قسمت رنگ به این صورت به پردازش اضافه می شود که مولفه های سبز، قرمز و آبی یک سیگنال روشنایی (Y) و دو سیگنال رنگ (Q,I) تولید می کنند. سیگنال های روشنایی، رنگ، صدا و همزمانی مالتی پلکس می‌شوند تا یک سیگنال رنگی مرکب تولید کنند. وسیله دریافت کننده این

مولفه ها را دی مالتی پلکس می کند. سپس سیگنال Y، I، Q دوباره برای نمایش به R، G، B تبدیل می شود. تبدیل رنگ و ترکیب سیگنال ها را با جزییات بیشتری در قسمت بعد توضیح می‌دهیم.

سه سیستم تلویزیون رنگی مختلف در سراسر دنیا وجود دارد: سیستم NTSC که در آمریکای شمالی و بعضی قسمت های آسیا شامل ژاپن و تایوان استفاده می شود. سیستم PAL که در اروپای غربی و چین استفاده می شود و سیستم SECAM که در اروپای شرقی، فرانسه و خاورمیانه استفاده می شود. این سیستم ها را بر حسب رزلوشن زمانی و مکانی، مولفه رنگ و مکانیسم مالتی پلکس کردن مقایسه خواهیم کرد. اطلاعات بیشتر در مورد سیستم های تلویزیون رنگی در [۱] موجود است.

۳-۱- رزلوشن مکانی و زمانی

تمام سیستم های تلویزیون رنگی، مکانیسم اسکن بینابینی ۱: ۲ را برای گرفتن و پخش ویدیو استفاده می کنند. سیستم NTSC نرخ قاب ۵۹/۹٤HZ را استفاده می کند و ۵۲۵ خط بر قاب دارد. نرخ میدان به طور جزیی نسبت به ٦۰ هرتز استفاده شده در سیستم تک رنگ متفاوت است به طوری که یک فرکانس حامل رنگ برای مالتی پلکس کردن مولفه های رنگ و روشنایی ممکن می‌باشد.

سیستم های PAL و SECAM هر دو از نرخ قاب ۵۰ هرتز و ٦۲۵ خط بر قاب استفاده می‌کنند. این نرخ قاب ها طوری انتخاب شده اند که با فرکانس برق شهر کشورهای استفاده کننده تداخل نداشته باشد، آنها همچنین با فرکانس حداقل چشمک زدن سیستم

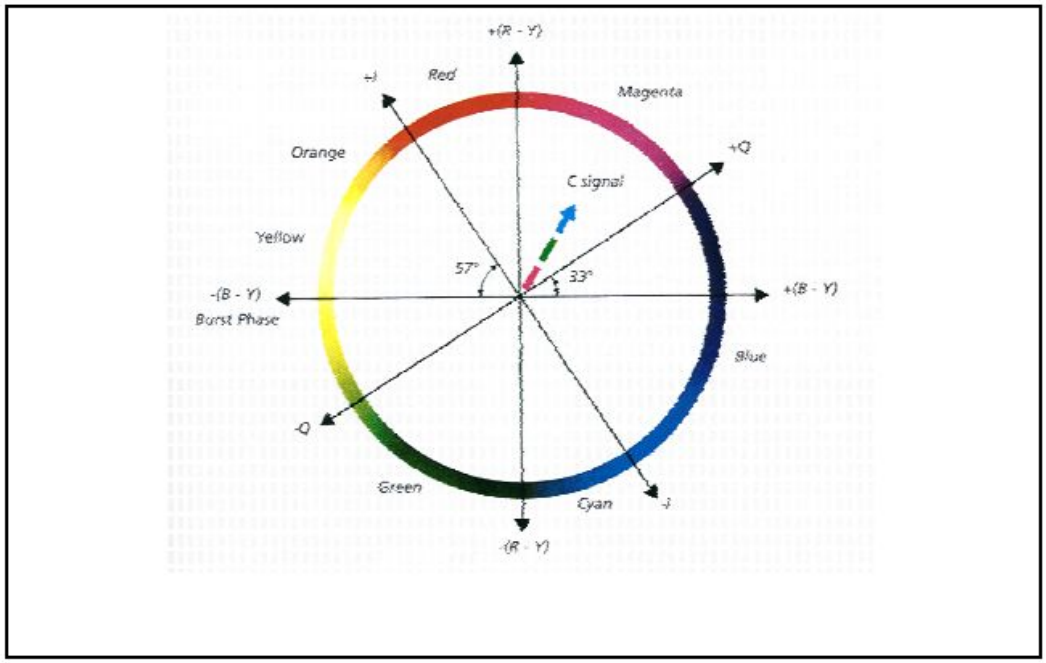
بینایی انسان تطابق دارند. تمام سیستم ها نسبت اندازه تصویر ٤:٣ دارند. پارامترهای NTSC، PAL و SECAM در جدول ١ خلاصه شده اند.

جدول ١: پارامترهای سیستم های NTSC، PAL و SECAM

Parameters	NTSC	PAL	SECAM
Field Rate (Hz)	59.95 (60)	50	50
Line Number/Frame	525	625	625
Line Rate (Line/s)	15,750	15,625	15,625
Color Coordinate	YIQ	YUV	YUV
Luminance Bandwidth (MHz)	4.2	5.0/5.5	6.0
Chrominance Bandwidth (MHz)	1.5(I)/0.5(Q)	1.3(U,V)	1.0 (U,V)
Color Subcarrier (MHz)	3.58	4.43	4.43
Audio Subcarrier	4.5	5.5/6.0	6.5
Total Bandwidth (MHz)	6.0	7.0/8.0	8.0

٣-٢- مختصات رنگ

سیستم های مختصات رنگ استفاده شده در این سه سیستم متفاوتند. برای ثبت و نمایش، هر سه سیستم از RGB استفاده می کنند ولی تعریف های کمی متفاوتی از طیف رنگ هر مولفه دارند. برای انتقال سیگنال ویدیو، به منظور کاهش پهنای باند موردنیاز و سازگار بودن با سیستم های تلویزیون سیاه و سفید مختصات رنگ/ روشنایی استفاده شده است مختصات YIQ از مختصات XYZ منتج می شود. در واقع، بر اساس رابطه بین مولفه های RGB و YIQ، مقدار Y از روی مقدار RGB که مولفهٔ روشنایی را تشکیل می دهد قابل محاسبه است. دو مولفه رنگ I و Q با چرخاندن ٣٣ درجه محورهای $B - Y$ و $R - Y$ به دست می آید. (شکل ٤) I با رنگ های محدودهٔ نارنجی – فیروزه ای و Q با محدوده رنگ های سبز – ارغوانی متناظر است. چون چشم انسان به تغییرات محدوده سبز – ارغوانی کمتر از محدوده زرد – فیروزه ای حساس است، مولفه Q می تواند در پهنای باند کمتری نسبت به مولفه I ارسال شود.



شکل ۴: طیفهای رنگ

مقادیر YIQ از طریق روابط زیر در سیستم NTSC از روی RGB به دست می آیند:

$$Y = 0.299\ R + 0.587\ G + 0.114\ B,$$
$$I = 0.596\ R - 0.274\ G - 0.322\ B,$$
$$Q = 0.211\ R - 0.523\ G + 0.311\ B.$$

مقادیر RGB با روابط زیر به YIQ مرتبطند:

$$R = 1.0\ Y + 0.956\ I + 0.621\ Q,$$
$$G = 1.0\ Y - 0.272\ I - 0.649\ Q,$$
$$B = 1.0\ Y - 1.106\ I + 1.703\ Q.$$

که R، G، B مقادیر نرمالیزه تصحیح گاما شده اند به طوری که (۱ و ۱ و ۱)= (R,G,B) با رنگ سفید مرجع در سیستم NTSC متناظر است.

یک خاصیت خوب دیگر مختصات YIQ این است که $\tan^{-1}(Q/I)$ غلظت و $\sqrt{I^2+Q^2}/Y$ اشباع را نشان می‌دهد. در یک ویدیوی مرکب NTSC، مولفه های I و Q در یک سیگنال مالتی پلکس می شوند به طوری که فاز سیگنال مدوله شده $\tan^{-1}(Q/I)$، است و اندازه آن $\sqrt{I^2+Q^2}/Y$ است. به خاطر اینکه تأثیر خطاها بر روی اندازه بیشتر از فاز است اطلاعات غلظت بهتر از اطلاعات اشباع منتقل می شود. این موضوع مطلوب است زیرا چشم انسان به غلظت بیشتر از اشباع حساس است. نام های I و Q از آن رو انتخاب شده اند که I همفاز با فرکانس مدولاسیون رنگ و Q عمود بر آن است. چگونگی مالتی پلکس کردن در بخش ٤-٣ گفته خواهد شد.

۳-۳- پهنای باند سیگنال

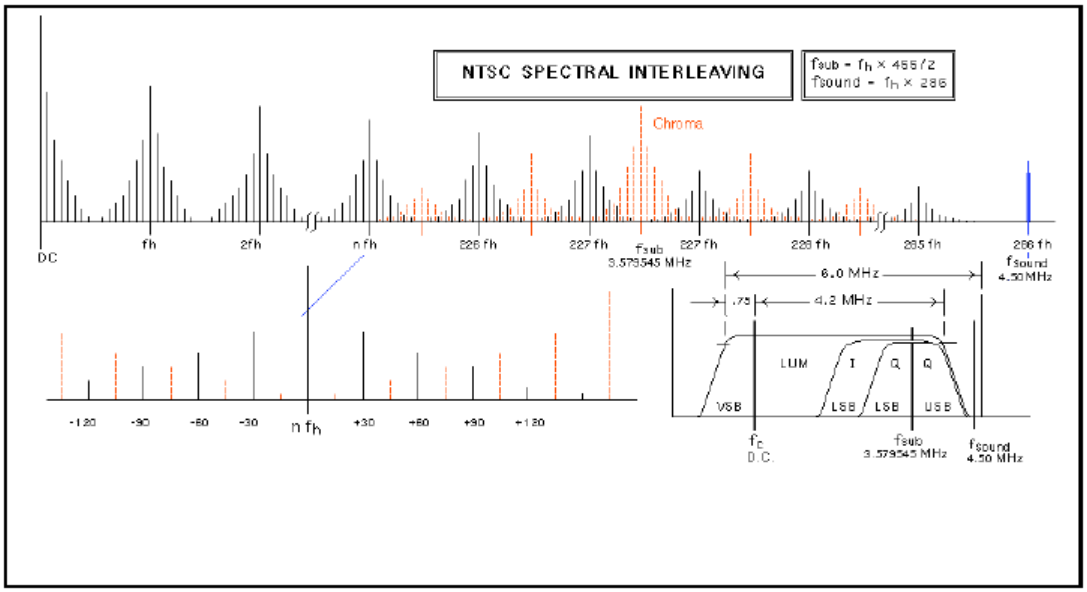
در بخش ٤-٢ پهنای باند ویدیویی NTSC تک رنگ را تحلیل کردیم. دیدیم که حداکثر فرکانس ٤/٢ مگاهرتز است. در یک سیستم تلویزیون رنگی، این پهنای باند تنها برای سیگنال روشنایی مصرف می شود. اگرچه سیگنال رنگ هم به طور بالقوه می تواند چنین پهنای باندی داشته باشد، معمولاً تغییرت سیگنال رنگ شدیداً از مولفه روشنایی کمتر است. همچنین سیستم بینایی انسان آستانه بسیار پایین تری برای مشاهده تغییرات رنگ دارد. عمدتاً، دو سیگنال رنگ پهنای باند بسیار باریکتری دارند. همان طور که قبلاً گفتیم چشم

به محدوده نارنجی – فیروزه ای بیشتر از سبز – ارغوانی حساس است بنابراین پهنای باند مولفه I برابر ۱/۵MHz و مولفه Q برابر ۰/۵MHz انتخاب می شود.

۳-۴- مالتی پلکس کردن روشنایی، رنگ و صدا

برای سازگار کردن سیگنال تلویزیون رنگی با سیستم تلویزیون سیاه و سفید، هر سه سیستم تلویزیون رنگی، فرمت ویدیویی مرکبی را استفاده می کنند که هر سه مولفه رنگ همراه با مولفه صدا در یک سیگنال مالتی پلس شده باشد. در اینجا مختصراً مکانیسم مورد استفاده NTSC را شرح می دهیم. در ابتدا دو مولفه رنگ با استفاده از مدولاسیون اندازه عمودی (QAM) در یک سیگنال ترکیب می شوند. فرکانس زیر حامل طوری انتخاب می شود که مضرب فردی از نصف نرخ خط باشد. این کار برای ارضاء معیارهای زیر انجام می گیرد: الف) باید به اندازه ای فرکانس بالا باشد که مولفه روشنایی انرژی بسیار کمی داشته باشد. ب) باید بین دو هارمونیک نرخ خط باشد که مولفه روشنایی قوی است. ج) باید به اندازه کافی از زیرحامل سیگنال صدا که در ٤/٥MHz است دور باشد. شکل ٥ ماکسیمم هارمونیک های سیگنال های روشنایی و رنگ در هم رونده را نشان می دهد.

در انتها، سیگنال صدا با استفاده از حامل ٤/٥MHz مدولاسیون فرکانسی (FM) می شود و به سیگنال ویدیویی مرکب اضافه می شود تا سیگنال مالتی پلکس شده نهایی را تولید کند. از آنجاییکه مولفه I پهنای باند ۱/۵MHz دارد سیگنال رنگ مدوله شده حداکثر فرکانسی تا ۵/۰۸MHz خواهد داشت. به منظور اجتناب از تداخل با سیگنال صوت، سیگنال صوت در نیمه بالایی باند فرکانسی به ۰/۵MHz محدود می شود. توجه کنید که نیمه پایینی باند فرکانسی سیگنال I به داخل بخش بالایی فرکانس سیگنال Y می رود. به همین خاطر، بعضی اوقات سیگنال I از هر دو طرف به ۰/۵MHz محدود می شود. در انتها کل سیگنال مرکب با پهنای باند حدود ٤/٧٥MHz به یک فرکانس حامل تصویر با استفاده مدولاسیون VSB، مدوله می شود به طوری که باند پایینی تنها ۱/۲۵MHz پایین تر از حامل تصویر می رود و کل سیگنال پهنای باندی حدود ٦MHz اشغال می کند. این فرایند شبیه سیستم تلویزیون تک رنگ است. حامل تصویر به کانال انتشار بستگی دارد. شکل ٥ گوشه پایین راست ترکیب طیفی سیگنال NTSC را نشان می دهد. پهنای باند سیگنال در سیستم های تلویزیون رنگی در جدول ۱ خلاصه شده است.



شکل ٥: ماکسیمم هارمونیک های سیگنال های روشنایی و رنگ در هم رونده

در گیرنده تلویزیون، سیگنال مرکب ابتدا به باند پایه و مدولاسیون می شود و سپس سه مولفه ویدیو و یک مولفه صدا دی مالتی پلکس می‌شوند. برای مجزا کردن سیگنال ویدیو و صدا می توان از یک فیلتر پایین گذر استفاده کرد، برای جدا کردن سیگنال رنگ از روشنایی به طور ایده آل از یک فیلتر شانه ای می توان استفاده کرد. بیشتر تکنولوژی های جدید TV یک فیلتر دیجیتال شانه ای را با

فرکانس های صفر در هارمونیک های متناظر با مولفه رنگ پیاده سازی می کنند، ولی تکنولوژی های قدیمی TV از مدار RC ساده برای انجام فیلترینگ پایین گذر با فرکانس قطع ۳MHz استفاده می کنند که مولفه I باعث آرتیفکت های تداخل رنگ و تداخل روشنایی می شود. تداخل رنگ به رنگ های خطای ایجاد شده با سیگنال روشنایی فرکانس بالا که به فرکانس زیر حامل رنگ نزدیک است گفته می شود. تداخل رنگ به الگوهای لبه فرکانس بالای اشتباه ایجاد شده بر اثر اطلاعات رنگ مدوله شده گفته می شود. برای نمایش اثرات فیلترهای مختلف، [۲] را ببینید. بعد از استخراج سیگنال رنگ، روش دمدولاسیون رنگ متناظر برای جدا کردن دو مولفه رنگ استفاده می شود. نهایتاً، سه مولفه رنگ برای نمایش به مختصات RGB برده می شوند.

۴- تکنولوژی ویدیویی دیجیتال

۴-۱- فرمت های ویدیویی دیجیتال CCIR601

در کوششی برای استاندارد کردن فرمت های دیجیتال استفاده شده در سیگنال های ویدیویی آنالوگ، CCIR، بخش رادیوی جامعه مخابرات بین المللی، ITU-R، توصیه CCIR601 را ایجاد کرد.

۴-۱-۱- رزلوشن مکانی سیگنال CCIR601

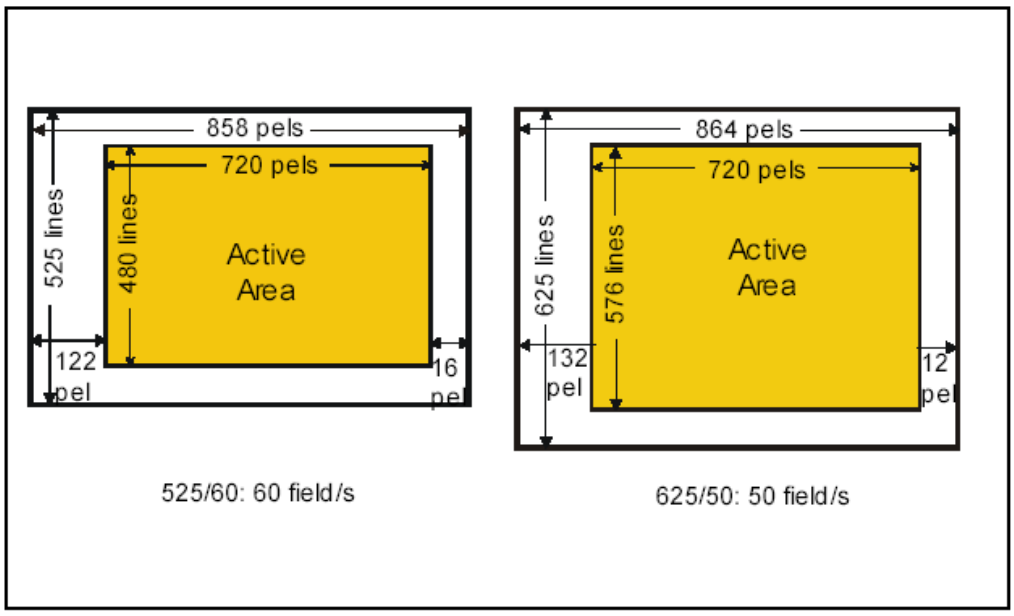
برای تبدیل یک اسکن raster به ویدیوی دیجیتال، تنها نیاز به نمونه برداری از شکل موج یک بعدی می باشد. نرخ نمونه برداری معادل برابر است با:

$$f_s = f_{s,x}\ f_{s,y}\ f_{s,t} = f_{s,x}\ f_l \text{ samples/second,}$$

در استاندارد CCIR601 نرخ نمونه برداری برای ارضا ۲ شرط انتخاب می شود: الف- رزلوشن نمونه برداری افقی باید با رزلوشن نونه برداری عمودی حتی الامکان برابر باشد و ب) همان نرخ نمونه برداری باید برای تمام سیستم ها به کار رود. عددی که به هر دوی آنها نزدیک است و شرط دوم را نیز ارضا می کند برابر است با:

$$f_s = 858\, f_l\,(\text{NTSC}) = 864\, f_l\,(\text{PAL/SECAM}) = 13.5 \text{ MHz}$$

تعداد پیکسل ها در هر خط۵۵۸ برای NTSC و ۵۶۴ برای PAL/SECAM است. این دو فرمت به نام سیگنال های 625/50 و CCIR601 525/60 شناخته می شوند و در شکل ٦ نمایش داده شده اند. تعداد خطوط فعال به ترتیب ۴۸۰ و ۵۷۶ در ۵۲۵ و۵۶۵ است. ولی تعداد پیکسل/خط فعال در هر دو برابر ۷۲۰ پیکسل است. بقیه نمونه های به دست آمده در خلال بازگشت های افقی و عمودی است که در ناحیه غیر فعال می افتد.



شکل ۶: فرمت های ویدئوی CCIR601

۴-۱-۲- مختصات رنگ

در کنار رزلوشن تصویر، توصیه CCIR601 یک مختصات رنگ دیجیتال به نام ycbcr تعریف می کند. برای مرور خوبی بر دلایل تاریخی طراحی این سیستم رنگ خوانندگان به [۳] ارجاع داده می شوند. اینجا ما فقط ماتریس انتقال برای پیدا کردن مقادیر انرژی RGB را می گوییم. فرض می‌شود که مقادیر RGB در محدوده (۲۵۵-۰) هستند. رابطه بین ycbcr و RGB به این صورت است:

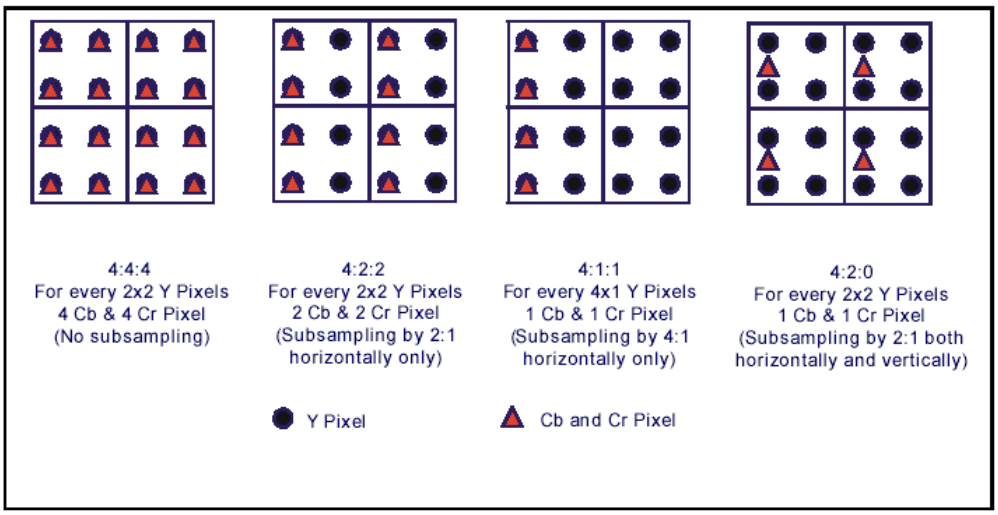
$$
\begin{aligned}
Y_d &= 0.257\,R_d + 0.504\,G_d + 0.098\,B_d + 16,\\
C_b &= -0.148\,R_d - 0.291\,G_d + 0.439\,B_d + 128,\\
C_r &= 0.439\,R_d - 0.368\,G_d - 0.071\,B_d + 128,
\end{aligned}
$$

$$
\begin{aligned}
R_d &= 1.164\,Y_d' + 0.0\,C_b' + 1.596\,C_r',\\
G_d &= 1.164\,Y_d' - 0.392\,C_b' - 0.813\,C_r',\\
B_d &= 1.164\,Y_d' + 2.017\,C_b' + 0.0\,C_r',
\end{aligned}
$$

همانطور که در سیستم های BAL/SECAM یا NTSC دیدیم. در مختصات ycbcr، y روشنایی را نشان می دهد و بین (۲۳۵-۱۶) مقیاس بندی شده است و cb و cr ویرایش های نرمالیزه شده اختلاف رنگ ها هستند. مقیاس بندی و شیفت برای قرار گرفتن در محدوده (۲۴۰-۱۶) انجام شده است.

۴-۱-۳- زیر نمونه برداری رنگ

نرخ نمونه برداری مکانی معرفی شده در قبل به مولفه روشنایی اشاره می کرد. برای مولفه های رنگ معمولاً نصف نرخ نمونه برداری استفاده می شود. این کار منجر به نصف شدن تعداد پیکسل‌ها در هر خط می شود ولی همان تعداد خطوط در هر قاب باقی می مانند. این فرمت ها با نام۴:۲:۲ که اشاره به دو نمونه cb و cr برای هر ٤ نمونه y دارد. برای کاهش بیشتر داده های موردنیاز CCIR601 فرمت ۴:۱:۱ ٤ را ارائه داده است که در آن مولفه رنگ در هر خط با فاکتور ٤ زیر نمونه برداری می شود. برای هر ٤ نمونه y یک نمونه cr و یک نمونه cb وجود دارد. این روش نمونه برداری، رزلوشن بسیار غیر متقارنی در جهتهای عمودی و افقی می دهد. بنابراین یک فرمت نمونه برداری دیگرابداع شد که نرخ مولفه های cr و cb را در هر دو جهت عمودی و افقی به نصف کاهش می دهد. در این فرمت نیز به ازاء هر ٤ نمونه y ی نمونه cr و یک نمونه cb وجود دارد. ولی برای اجتناب از اشتباه شدن با فرمت قبلی با نام ٤:۲:۰ شناخته می شوند. برای کاربردهای با رزلوشن بالا، فرمت ٤:٤:٤ تعریف شده است که مولفه های رنگ را درست در همان رزلوشن مولفه روشنایی نمونه برداری می کند. موقعیت های نسبی نمونه های رنگ و روشنایی برای فرمت‌های مختلف در شکل ۷ نشان داده شده است.



شکل ۷: موقعیت های نسبی نمونه های رنگ و روشنایی برای فرمت‌های مختلف

۴-۱-۴- نرخ های داده خام سیگنال های CCIR601 و استانداردهای فشرده سازی MPEG2

اطلاعات خام سینال CCIR601 به عامل زیر نمونه برداری رنگ وابسته است. فرمت ٤:٢:٢ از بقیه رایجتر است که ۲ مولفه رنگ به ازای دو نمونه y وجود دارند که همه با هم ۸ بیت را تشکیل می دهند.

بنابراین نرخ بیت معادل برای هر نمونه y ۱۶ بیت و نرخ داده خام ۲۱۶Mbs است. داده خام متناظر با ناحیه فعال 166Mbs است. با فرمت ٤:٢:٠ ۲ مولفه رنگ برای هر ٤ مولفه y وجود دارد و نرخ بیت معادل برای هر نمونه y، ۱۲ بیت است. بنابراین نرخ داده خام 162Mbs است که 124Mbs در ناحیه فعال است. در فرمت ٤:٤:٤ نرخ بیت معادل برای هر نمونه y، ۲٤ بیت است و نرخ داده خام324Mbs است که 249Mbs در ناحیه فعال است. رزلوشن و نرخ داده سیگنال های CCIR601 در جدول ۲ خلاصه شده‌اند.

توجه کنید که ما مختصات RGB را برای نمایش رزلوشن ویدیو استفاده کرده ایم. هر مولفه رنگ نیاز به رزلوشن یکسان با مولفه Y دارد و هر کدام نیاز به ۸ بیت دارد که نرخ داده خام 324Mbs را نتیجه می دهد. می بینیم که با انتخاب مختصات رنگ مناسب که اجازه زیر نمونه برداری شدن به مولفه های مشخصی را می دهد، می توانیم از قبل نرخ داده را بدون قربانی کردن کیفیت دید کاهش دهیم.

فرمت های CCIR601 در کاربردهای ویدیویی کیفیت بالا با ۴:۴:۴ و۴:۲:۲ عمدتاً برای تولید و ویرایش استفاده شده اند. درحالیکه۴:۲:۰ برای پخش ویدیو مثلاً برای روی DVD و VOD از فرمت ٤:٢:٠ استفاده می شود. استاندارد فشرده سازی MPEG2 عمدتاً برای فشرده سازی سیگنال های CCIR601 ٤:٢:٠ به کار رفته اند اگرچه می توانند برای رزلوشن های بالاتر یا پایین تر هم به کار بروند. یک سیگنال ٤:٢:٠ نمونه با نرخ داده فعال ۱۲٦mbps می تواند به ۸mbps–٤ فشرده شود.

۴-۲- فرمت های دیگر ویدیویی دیجیتال و کاربردهای آن

علاوه بر فرمت CCIR601، فرمت های ویدیوی دیجیتال متعدد دیگری نیز تعریف شده اند. اهم آنها عبارتند از CIF که توسط ITV-T تعریف شده و نصب رزلوشن ٤:٢:٠ CCIR601 در هر دو بعد عمودی و افقی را دارد و برای کاربردهای کنفرانس ویدیویی درست شده است. QCIF که ربع CT-F هست برای کاربردهای تلفن تصویری تعریف شده است. هیچ کدام آنها از خطوط بینابینی استفاده نمی کنند، استاندارد کدینگ TTV-T H.261 برای فشرده سازی ویدیو در فرمت‌های p×64kbps با P=1,2,…,30 برای انتقال بر روی خطوط ISPN که تنها اجازه نرخ های انتقال در ضرایب 64 kbps را می دهند درست شده به طور کلی، سیگنال CIF با یک نرخ داده خام ۳۷/۳mbps می تواند به ۱۲۸ تا ۳۸٤kbps با کیفیت قاب قبول فشرده شود در حالی که سیگنال QCIF با نرخ داده خام ۳mbps. ۹ می تواند به ۱۲۸kbps–٦٥ فشرده شود. استاندارد اخیر H.263 می تواند به کیفیتی بهتر از H.261 در همان نرخ بیت برسد. به طور مثال، فشرده سازی تصویر QCIF با ۲۰kbps در عین کیفیت شبیه یا بهتر از H.261 در ٦٤kbps ممکن است. این موضوع تلفتصویری بر روی یک خط مودم ۲۸/۸kbps را ممکن می کند.

به طور موازی با کوشش های ISO-MPEG (CCIR) TTV هم تعدادی فرمت ویدیویی دیجیتال تعریف کرد. فرمت SIF $\frac{1}{4}$ اندازه نحیه فعال سیگنال CCR601 را دارد و تقریباً مثل CIF است. این فرمت برای کاربردهای ویدیویی با کیفیت متوسط مانند بازی های ویدیویی و فیمل های cv استفاده شده است. مثل CCIR601 ۲ فرمت SIF موجود است. SIF-25 با نرخ قاب ۳۰ هرتز و ۲٤۰ خط و SIF-625 ب نرخ قاب ۲۵ و ۲۸۸ خط همچنین مجموعه متناظری از فرمت SIF-I وجود دار که ۲:۱ از خطوط بینابینی استفاده می کند. الگوریتم MPEG-1 می تواند یک ویدیویی SIF نمونه با نرخ داده خام ۳۰mbps را تا ۱/۱ mbps با کیفیت برابر با رزلوشن دید شده بر روی VHS VCR فشرده کند. که از کیفیت انتشار تلویزیون کمتر است. نرخ ۱/۱mbps امکان تماشا بر روی CDROM ها را ه دستیابی به نرخ ۱/٥mbps را دارند فراهم می کند. توزیع فیلم های MPEG1 بر روی CDها ورود ویدیوی دیجیتال به بازار مصرف را در اوایل دهه ۱۹۹۰ اعلام کرد. تکنولوژی MPEGZ سنگ بنای نسل بعدی سیستم های تلویزیون است که کاملاً دیجیتال بوده و از تکنولوژی فشرده سازی و انتقال دیجیتال استفاده می کنند. جدول ۲، جزئیات فرمت های

ویدیویی بحث شده دربالا را و کنار کاربردهای اصلی آنها، روش های فشرده سازی و نرخ بیت های فشرده شده نشان می دهد. توضیحات بیشتر درباره استانداردهای فشرده سازی در آزمایش بعدی می آید.

فرمت CCIR601 فرمت تصویر استاندارد برای تلویزیون دیجیتال (DTV) خواهد بود. برای افزایش بیشتر کیفیت ویدیو فرمت های HDTV متعددی توسط موسسه SMPTE استاندارد شده‌اند

جدول ۲: استانداردهای فرمت ویدئو

Video Format	Y Size	Color Sampling	Frame Rate (Hz)	Raw Data Rate (Mbps)
HDTV Over air. cable, satellite, MPEG2 video, 20-45 Mbps				
SMPTT296M	1280x720	4:2:0	24P/30P/60P	265/332/664
SMPTE295M	1920x1080	4:2:0	24P/30P/60I	597/746/746
Video production, MPEG2, 15-50 Mbps				
CCIR601	720x480/576	4:4:4	60I/50I	249
CCIR601	720x480/576	4:2:2	60I/50I	166
High quality video distribution (DVD, SDTV), MPEG2, 4-10 Mbps				
CCIR601	720x480/576	4:2:0	60I/50I	124
Intermediate quality video distribution (VCD, WWW), MPEG1, 1.5 Mbps				
SIF	352x240/288	4:2:0	30P/25P	30
Video conferencing over ISDN/Internet, H.261/H.263, 128-384 Kbps				
CIF	352x288	4:2:0	30P	37
Video telephony over wired/wireless modem, H.263, 20-64 Kbps				
QCIF	176x144	4:2:0	30P	9.1

که در جدول ۲ لیست شده اند. یک تمایز عمده HDTV نسبت ۱۶:۹ آن در برابر ۴:۳ در SDTV است. رزلوشن تصویر در هر دو بعد عمودی و افقی دوتا سه برابر شده است. همین طور اسکن متوالی برای کاهش آرتیفکت های خطوط بینابینی استفاده شده است. یک نمونه کیفیت بالا در استاندارد فشرده سازی ویدیویی MPEG2 برای فشرده سازی ویدیویی HDTV عرضه شده است. معمولاً این کار می تواند نرخ داده را به ۲۰mbps کاهش دهد در عین اینکه کیفیت را بالا نگه می‌دارد.این نرخ بیت از آن رو انتخاب شده است که رشته بیت ترکیب شده با صدا در هنگام انتقال با تکنیک های مدولاسیون دیجیتال می تواند در کانال ۶MHz جا بگیرد که پهنای باند اختصاصی برای انتشار HDTV در آمریکا است.

۵- آزمایش ها

۱- یک تصویر رنگی RGB را بگیرید و به سه مولفه تجزیه کنید و به تصویر ycbcr تبدیل کنید. سه مولفه cr، cb و y را جداگانه به عنوان سه تصویر ذخیره کنید. هر تصویر را جداگانه به عنوان یک تصویر سیاه سفید با استفاده از imshow نشان دهید. همبستگی بین رنگ هر پیکسل در تصویر RGB اصلی و مقادیر cr، cb و Yاش را پیدا کنید. یک برنامه matlab برای این کار بنویسید. برنامه شما باید تصویر اصلی RGB را در فرمت BMP بخواند و برای هر پیکسل مقادیر R، G و B را به مقادیر y، cb و cr تبدیل کند و آنها را در آرایه های جداگانه ذخیره کند. در آخر هر آرایه را در یک فایل تصویر جداگانه ذخیره کند. از "help bmpread" و "help imshow" برای پیدا کردن پارامترهایی که باید در این توابع به کار ببرید استفاده کنید.

۲- اختیاری: گفتیم که مولفه y (روشنایی) جزییات بیشتری از مولفه های رنگ (cb یا cr) دارد. بنابراین می توانیم آن را زیر نمونه برداری کنیم و با فاکتور ۲ در هر دو جهت عمودی و افقی بدون از دست دادن جزییات). برای نشان دادن این امر، ابتدا سعی کنید که هر سه مولفه تولید شده آزمایش ۱ (y، cb و cr) را با فاکتور ۲ در هر جهت عمودی و افقی زیر نمونه برداری کنید. برای ساده سازی، پیش فیلتر انجام ندهید. سپس در هر دو جهت عمودی و افقی بیش نمونه برداری کنید و از درونیابی خطی استفاده کنید.

تصاویر اصلی y، cb و cr را با تصویر پس از زیر و بیش نمونه برداری مقایسه کنید، آیا در مولفه y اعوجاج های بیشتری نسبت به مولفه های cr و cb احساس می کنید؟ می توانید این آزمایش را با ابزار تغییر نرخ matlab یا photoshop انجام دهید.

۳- یک تصویر با خطوط بینابینی را به یک دنباله پشت سر هم تبدیل کنید. یک قاب بینابینی شده تشکیل شده از دو میدان را بخوانید، عکس عمل بینابینی را برای تولید دو قاب پشت سر هم انجام دهید.

این کار را به این صورت انجام دهید که در داخل هر میدان مثلاً یک میدان زوج، که فقط حاوی خطوط زوج است، یک خط فرد را با میانگیری از خطوط فرد بالا و پایین به دست آورید. قاب بینابینی شده اصلی و دوقاب عکس بینابینی شده را نمایش دهید. آیا اثرات دندانه ای دندانه ای (لبه‌های عمودی مانند ردیف دندان) بر روی قاب بینابینی شده اصلی را می بینید؟ چرا؟ قاب های عکس بینابینی شده چطور؟ آیا آنها از نظر دید رضایتبخش تر هستند؟ مشاهداتتان را بنویسید (برای خواندن یک تصویر سیاه سفید می توانید از 'fopen' و 'fread' استفاده کنید).

٤- یک ویدیوی دیجیتال را با استفاده از تجهیزات دیجیتالیزاسیون ویدیو در آزمایشگاه استفاده کنید (دوربین یا VCR و برد ثبت ویدیو). توصیه می شود که نوار VHS یا S-VHS حامل کلیپ‌های ویدیویی که می خواهید دیجیتال کنید را بیاورید.

۶- مراجع

[1]. B. Grob and C. Hemdon, Basic Television and Video Systems, 6th ed., Glencoe/McGraw Hills, 1999.

[2]. J. F. Blinn, "NTSC: Nice Technology, Super Color", IEEE Computer Graphics and Applications Magazine, pages 17-23, Mar. 1993.

[3]. A. N. Netravali and B. G. Haskell, Digital Pictures: Representation, Compression and Standards, Plenum Press, 2nd Edition, 1995.

[4]. D. H. Pritchard, "US Color Television Fundamentals", IEEE Trans. Consum. Electron., CE-23:467-478, 1977.