

فشرده سازی تصویر

به روش

MPEG

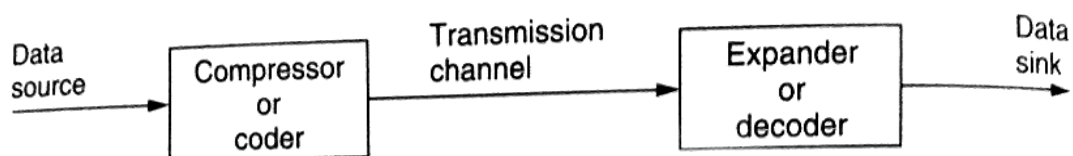
تعریف (MPEG)

MPEG در حقیقت خلاصه (Moving Picture Expert Group) گروه

متخصص تصاویر متحرک می باشد ، که توسط ISO برای فشرده سازی صوت و تصویر و انتقال آن به وجود آمد.

استاندارد MPEG بعنوان یک استاندارد فشرده سازی ، رمزنگاری را هم برای صوت و هم برای تصویر ارائه می کند . در واقع اصول کار برای هر دو یکسان است و در هر دو مورد اطلاعات اضافی که می توان بدون هیچ اثر محسوسی در کیفیت صدا یا تصویر حذف کرد را استخراج و حذف می کند. اما چگونگی این سلسله مراحل به دلیل تفاوت های ذاتی که در ساختار اطلاعات منبع برای صوت و تصویر وجود دارد، متفاوت می باشد.

خلاصه فشرده سازی به صورت شکل زیر می باشد :



شکل ۱-۲ بلوک دیاگرام فشرده سازی

Melec.ir

همان طور که در شکل مشخص است نرخ اطلاعات توسط کدر (compressor or coder) کاهش

می یابد و از طریق کانال منتقل می شود و توسط دکدر (expander or decoder) به نرخ اولیه باز

می گردد.

سیستم هایی که در آنها کدر (coder) پیچیده تر از دکدر (decoder) می باشد را سیستم نامتقارن

(Asymmetrical) می گوئیم . بر اساس این تعریف MPEG یک سیستم نامتقارن می باشد ، که در

آن Encoder نیاز دارد دارای قابلیت تطبیق یا به عبارتی دارای الگوریتم خاص باشد ؛ ولی دکدرها

کارهای مشخص شده ای را انجام می دهند . مزیت این روش در کاربرد هایی مثل پخش گسترده

تلویزیونی (Broadcasting) می باشد ؛ جایی که تعداد انکدرهای آن کم و قیمت آن ها بالا است ؛

ولی دکدرها (مانند رسیور های خانگی) دارای تعداد بالا و قیمت کم هستند .

در MPEG روشی که دکدر برای ترجمه بیت ها استفاده می کند دارای یک الگوی تعریف شده است.

مزیت این روش در این است که در طول زمان با گسترش الگوریتم های کد کردن، این دکدرها می

توانند با آن ها تطبیق شوند. MPEG محدود به فشردن سازی نمی شود و می تواند صدا و تصویر را

ترکیب کند و به صورت یک برنامه کامل ارسال کند .

3-1 انواع (MPEG) و کاربردهای مختلف

MPEG-1 : اولین نسخه MPEG است که به عنوان استاندارد رمزنگاری در سال 1992 رایج شد و بعد از آن MPEG2 که صدا و تصویری با کیفیت بالاتر ایجاد می کند و با MPEG-1 سازگار است طراحی شد.

هدف اصلی MPEG-1 رسیدن به تصویری با کیفیت متوسط و نرخ بیت ثابت 1/5Mb/S برای ذخیره تصویر و صوت روی CD-ROM بود. قسمت مربوط به تصویر از 1.15Mb/S استفاده می کند و باقیمانده 350Kb/s به وسیله صوت و داده های اضافی دیگری که مورد نیاز سیستم است استفاده می شود. به هر حال ، ویژگی های MPEG-1 خیلی قابل انعطاف است و امکان انتخاب پارامترهای مختلف را بسته به مصالحه بین پیچیدگی رمزنگار و سرعت فشرده سازی و کیفیت آن فراهم می کند. عمومی ترین کاربرد MPEG-1 در Video CD و نیز شامل فرمت معروف و پر کاربرد MP3 برای فشرده سازی صوت می باشد.

MPEG-2 : انتقال دهنده استانداردهای صوت و تصویر برای پخش تلویزیونی ؛ کاربرد آن در تلویزیون های دیجیتال (Over-The-Air) مثل DVB ، ATSC و ISDB ، تلویزیون های دیجیتال ماهواره ای ، تلویزیون های کابلی ، SVCD و با مقداری تغییر و تبدیل به عنوان VOB (Video Object) در حمل تصویر در DVD .

در مورد تفاوت های اساسی MPEG-1 نسبت به MPEG-2 می توان گفت که MPEG-1 در نرخ بیت حداکثر 1/856 Mb/S (برای کاربردهای مبتنی بر CD) استاندارد شده است و سیستم تصویر آن بصورت مرور تمام خط (Uninterlaced) می باشد. اما MPEG-2 سیستم تصویر مرور یک خط در میان (Interlaced) را با نرخ 4-9 Mb/S پشتیبانی می کند. هر دوی این استانداردها امکان ایجاد لایه های اضافی را که برای اضافه کردن انواع دیگر داده ها که در طول داده های تصویری لازم است فراهم می کنند. پهنای باند استفاده شده در MPEG-1 نسبت به MPEG-2 به مقدار زیادی کمتر

Melec.ir

بوده و کیفیت تصویر پایین است. میزان داده های پردازش شده بوسیله MPEG-2 چهار برابر MPEG-1 می باشد. در واقع MPEG-2 بیشتر در مواردی بکار می رود که پهنای باند نسبت به کیفیت از اهمیت بیشتری برخوردار باشد.

MPEG-3 : کاربرد اصلی در HDTV است. ولی چون MPEG-2 با کمی گسترش همان توانایی های MPEG-3 را پیدا می کرد جایگزین آن شد .

MPEG-4 : MPEG-1 را گسترش داده تا بتوانند از اجزا صوت و تصویر (Audio/Video Objects)، محتویات سه بعدی تصویر و انکد کردن با نرخ بیت کم را پشتیبانی کنند بعلاوه از ابزار کدینگ افزوده تر و پیچیدگی بیشتر برای دستیابی به فاکتور فشرده سازی بالاتر از MPEG-2 استفاده می کند؛ کاربرد در شبکه برای انتقال جریان رسانه ای (Media Stream) ، در تلفن های تصویری و پخش تلویزیونی می باشد که همه آن ها از مزایای فشرده سازی صوت و تصویر است .

- همچنین در MPEG-4 part 10 که معروف به H.264 می باشد از (Advance Video Coding) یا (AVC) استفاده می کند که در HD-DVD ها و دیسک های Blu-ray به کار می رود .

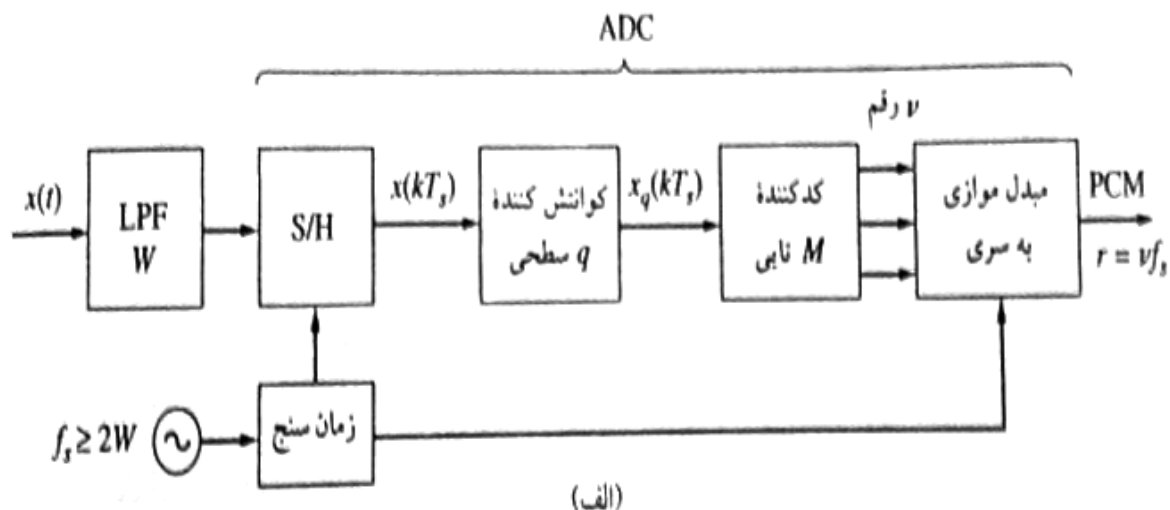
البته استانداردهای MPEG-7 و MPEG-21 نیز به وجود آمده اند که برای قالب های Multimedia طراحی شده اند .

1-4 دیجیتال سازی

مزایای موجود در سیگنال دیجیتال باعث می گردد تا در طراحی سیستم های پخش همگانی منابع صوت و تصویر از حالت آنالوگ به حالت دیجیتال یا دودویی حرکت شود .

برخی از مزیت های این کار عبارتند از:

- پایداری و ایمنی بیشتر در برابر نویز و عوامل خطا و رسیدن به پاسخ بهتر و (SNR) بالاتر.
 - قابلیت ذخیره سازی و بافر کردن به صورت بسیار انعطاف پذیر و در حجم بالاتر.
 - توانایی استفاده از الگوریتم های متنوع در پردازش دقیق سیگنال از جمله فشرده سازی.
 - بالاتر بردن ضریب امنیت و دسترسی به داده ها از طریق رمزنگاری.
 - قابلیت تکثیر متوالی از یک سیگنال دید **بلوک** **دیاگرام**انی در کیفیت سیگنال اصلی.
 - آسانتر شدن استفاده از جلوه های ویژه در سیگنال صدا و تصویر دیجیتال.
 - در سیستم های دیجیتال ایجاد تاخیر (delay) به آسانی قابل انجام است.
 - سیستم های دیجیتال مستقل از استانداردهای متنوع پخش (PAL,NTSC,SECAM..) می باشند.
- نحوه تبدیل آنالوگ به دیجیتال با استفاده از روش PCM می باشد ؛ که به صورت شکل زیر است:



شکل ۲-۲ (بلوک دیاگرام PCM)

که در آن سیگنال آنالوگ ابتدا وارد یک فیلتر پایین گذر شده و سپس با فرکانس f_s که بنابر اصل نایکوئیست $f_s \geq 2W$ نمونه برداری شده و پس از آن کوانتیزه می شود (دامنه آن به نزدیکترین عدد بر حسب گام کوانتتش تقریب زده می شود) و سیگنال دیجیتال حاصل می شود .

تنها عیب سیگنال دیجیتال شده حجم بسیار بالای اطلاعات می باشد که با فشرده سازی قابل حل است. حجم اطلاعات در سیگنال های آنالوگ با پهنای باند مشخص می شود. به حد پایین و بالای فرکانس سیگنال آنالوگ پهنای باند گویند. در سیستم های دیجیتال حجم اطلاعات حاصل ضرب تعداد نمونه برداری در تعداد سطوح دامنه های ممکن برای هر نمونه است. بنابراین یک سیگنال دیجیتال با دو پارامتر 1- تعداد بیت در هر نمونه، و 2- تعداد نمونه در هر ثانیه، مشخص شده و با این پارامترها حجم اطلاعات دیجیتال مشخص می شود.

1-2 فشرده سازی (Compression)

همانطور که قبلاً اشاره شد دیجیتال سازی سیگنال صوت و تصویر باعث افزایش بسیار زیاد حجم اطلاعات می شود. بنابراین برای ایجاد امکان پخش تلویزیونی تصاویر و یا برای ایجاد امکان ذخیره سازی و... به فشرده سازی سیگنال دیجیتال شده نیاز است.

فشرده سازی، کاهش نرخ بیت، کاهش دیتا و یا کد کردن منبع همگی اصطلاحاتی هستند که یک مفهوم را دارند. در اصل خود اطلاعات (یا تقریباً همان اطلاعات) با کمیت کمتری منتقل می شود. دلایلی که باعث شده فشرده سازی عمومیت پیدا کند، اجازه مینیاتورسازی را می دهد، که با استفاده از دیتا های ذخیره شده کمتر، همان زمان پخش، با سخت افزار کمتر حاصل شود.

با استفاده از فشرده سازی می توان با دیتا های ضبط شده کمتر زمان پخش واحد ذخیره شده را افزایش داد. بنابراین چگالی اطلاعات کمتر خواهد بود، این چگالی اطلاعات می تواند باعث کاهش ساخت تجهیزاتی شود که به پایداری بالا در برابر شرایط سخت و به نگهداری و رسیدگی کمتری نیاز دارند.

در سیستم های انتقال نیز باعث کاهش پهنای باند می شود که نتیجه آن کاهش هزینه ها می باشد اگر پهنای باند مورد نظر برای دیتای فشرده نشده کافی بود فشرده سازی اجازه می دهد که در همان پهنای باند سرعت انتقال بالاتری نسبت به زمان حقیقی آن داشته باشیم.

اگر پهنای باند داده شده قابل استفاده باشد با فشرده سازی کیفیت بالاتری در همان پهنای باند خواهیم داشت.

روشهای فشرده سازی از الگوریتم های فشرده سازی عمومی که برای هر نوع داده ای قابل اعمال است استفاده می کنند.

این الگوریتمها بر اساس قوانین کلی زیر می باشند:

اضافات مکانی ؛ که نظر به بستگی نقاط مجاور از لحاظ مکانی در تصویر می باشد . دقت قوه

بینایی ؛ که بر عدم حساسیت چشم به جزییات ریز تصاویر توجه می کند. برای تصاویر ثابت JPEG

استفاده می شود. و از **اضافات زمانی** ؛ که بر اساس بستگی نقاط متناظر از لحاظ زمانی بین

تصویرهای پی در پی در صورت داشتن تصاویر متحرک MPEG علاوه بر دو قانون اول بهره می

برند.

به همین طریق، روشهای فشرده سازی صدا از اصل رمزنگاری صوتی – روانی که بر اساس

ویژگیهای قوه شنوایی است، برای کاهش نرخ بیت و با حذف اطلاعات غیر قابل شنود بهره می برند.

2-2 انواع اصول فشرده سازی عمومی

(Run Length Coding) RLC: 1-2-2

وقتی که یک منبع اطلاعات عناصر مربوط به پیامی را تولید می کند که در آن یک جزء به طور متوالی تکرار می شود فرستادن کد عنصر و تعداد آنها، نسبت به تکرار کد به صرفه تر است. این عمل یک ضریب فشرده سازی متغیر ایجاد می کند که هر چه سری جزء تکرار شونده طولانی تر باشد ضریب فشرده سازی بزرگتر خواهد بود. این نوع رمزنگاری که در آن هیچ اطلاعاتی از بین نمی روند به عنوان روش بازگشت پذیر (Reversible) تعریف می شود. این روش معمولاً برای فشرده سازی فایل ها برای ذخیره سازی روی دیسک و یا انتقال بوسیله کامپیوتر (zip,...) و نیز در ماشین های fax به کار می رود.

(Variable Length Coding) VLC : 2-2-2

VLC به روش رمزنگاری با طول متغیر یا رمزنگاری آنتروپی می گویند. که در این روش کاهش نرخ بیت بر اساس این واقعیت استوار است که احتمال اتفاق یک عنصر تولید شده از منبع، توسط کدی که با n بیت در حالت هم احتمال بودن عناصر بوجود آمده است و 2^n احتمال ممکن را خواهد داشت، گاهی اوقات مشابه نمی باشد. بنابراین برای کاهش نرخ بیت لازم برای انتقال رشته های تولید شده به وسیله منبع، مناسب است که عناصر با تکرار بیشتر با طول کمتر از n بیت و عناصر با تکرار کمتر با تعداد بیت بیشتری، رمزنگاری شوند. که نتیجه آن، داشتن متوسط طول کمتر از طول ثابت n بیت می باشد.

اما چنانچه این کد گذاری بخواهد مستقیماً بر روی بیت ورودی انجام گیرد، باید احتمال پیشامد هر عنصر قابل تولید بوسیله منبع را از قبل دانست که البته در برخی از موارد از این موضوع شناخت

Melec.ir

داریم. به عنوان مثال در مورد حروف الفبا در یک زبان مفروض، استفاده از این روش برای فشرده سازی متن وجود دارد. این روش همچنین برای تصاویر ویدئویی که بوسیله DCT فشرده سازی شده اند، وقتی که انرژی روی تعداد کمی از عناصر متمرکز شده، قابل اجراست و هنگامی که تصاویر متحرک طوری اند که همه عناصر احتمال یکسان دارند قابل اجرا نمی باشد.

هدف رمزنگاری با طول متغیر (VLC) یا رمزنگاری آنتروپی، تا حد امکان نزدیک شدن به نرخ بیت متناظر با آنتروپی منبع است که متناظر با تعداد متوسط بیت برای هر عنصر می باشد .

معروفترین روش برای رمزنگاری با طول متغیر، الگوریتم هافمن است که در آن فرض می کنیم احتمال هر عنصر را می دانیم .

(Discrete Cosine Transform) DCT: 3-2-2

DCT (تبدیل کسینوسی گسسته) روش فشرده سازی بکار رفته برای تصاویر می باشد . تبدیل کسینوسی گسسته حالت ویژه ای از تبدیل فوریه اعمال شده بر روی سیگنال های گسسته (نمونه برداری شده) می باشد . در تبدیل فوریه یک سیگنال متناوب به سری هایی از توابع سینوسی و کسینوسی تبدیل می شود . سپس می توان سیگنال را بوسیله سری های ضرایب هر یک از این توابع، نمایش داد .

تحت شرایط مشخصی DCT سیگنال را تنها به یک سری از توابع کسینوسی هارمونیک هم فاز با سیگنال ، تبدیل می کند که این عمل بر خلاف فوریه تعداد ضرایب لازم برای توصیف سیگنال را به نصف کاهش می دهد .

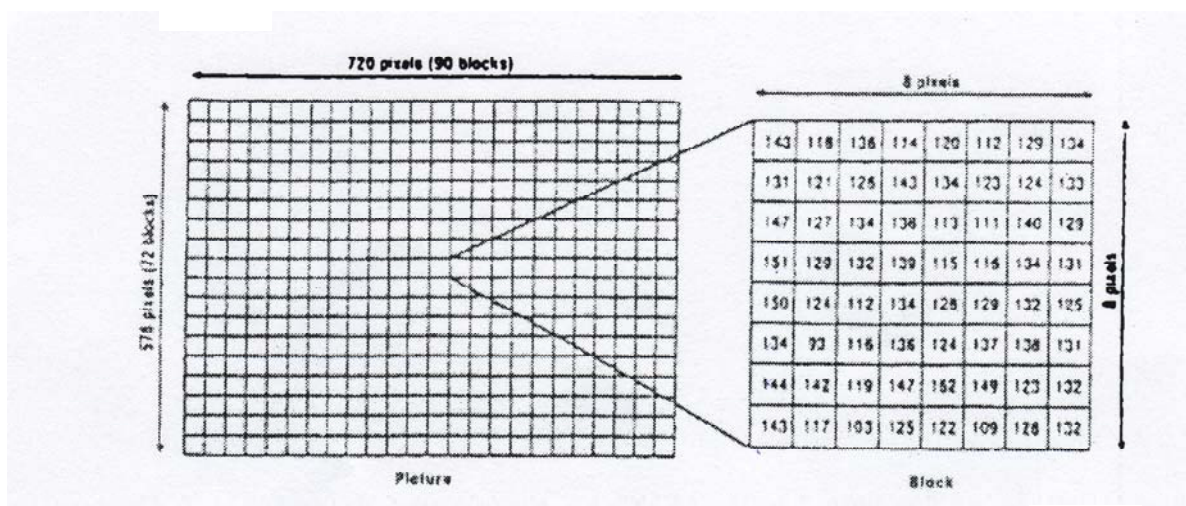
در مورد سیگنال تصویر ، سیگنال اصلی سیگنالی است که به صورت دوبعدی نمونه برداری شده است . بدین علت یک DCT دو بعدی در جهت های عمودی و افقی خواهیم داشت که مقادیر گسسته

Melec.ir

روشنایی یا رنگ از یک بلوک $n \times n$ را به بلوک ها یا ماتریس های ضرائب $n \times n$ تبدیل می کند که ضرائب موجود در این بلوک ها دامنه هر یک از توابع هارمونیک کسینوسی را نشان می دهند .

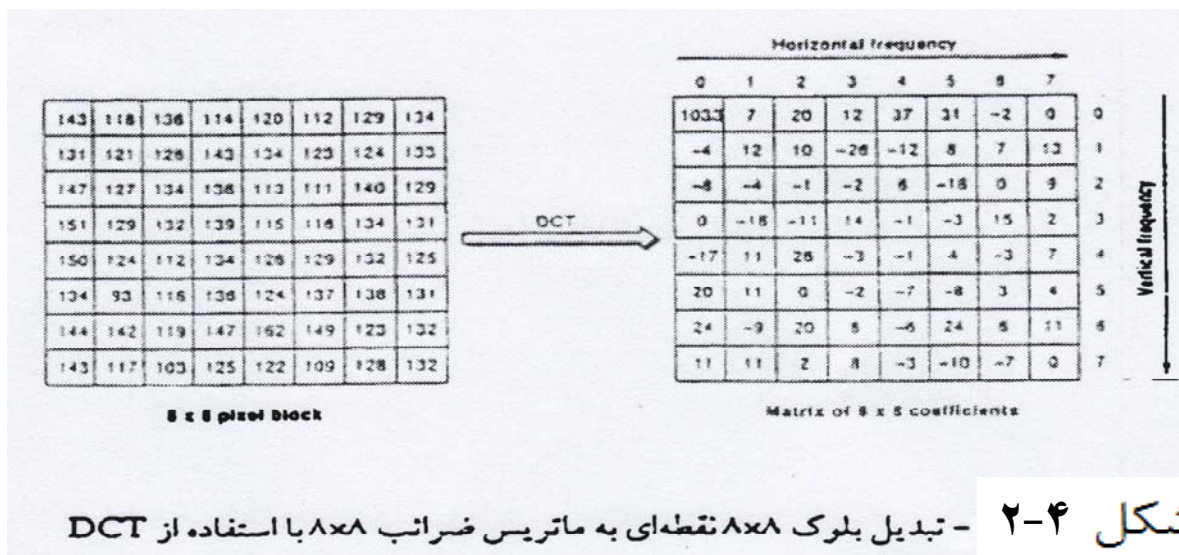
در بلوک تبدیل یافته ، ضرائب مربوط به فرکانس هایی هستند که در محور افقی از چپ به راست افزایش یافته و روی محور عمودی از بالا به پایین افزایش می یابند . اولین ضرائب در موقعیت (0 و 0) بیانگر فرکانس افقی و عمودی صفر است . بنابراین ضرائب dc نامیده می شود و پایین ترین ضرائب سمت راست بیانگر بالاترین مولفه فرکانسی در دو جهت می باشد .

به منظور کاهش پیچیدگی مدارها و زمان پردازش لازم ، اندازه بلوک انتخاب شده عموماً 8×8 می باشد و بصورتی که در شکل (1) دیده می شود :



شکل ۲-۳ تبدیل تصویر به بلوکهای 8×8 نقطه‌ای (مقادیر بیانگر میزان روشنایی نقاط هستند)

DCT آنرا به یک ماتریس ضرائب 8×8 تبدیل می کند (شکل 2) . با استفاده از این شکل ظاهر اصلی بلوک تصویر اصلی با میانگین گیری از ضرائب هریک از 64 مربع و جمع کردن نتایج ، بدست می آید .



بسته به جزییاتی که در بلوک اصلی وجود دارد ، ضرایب فرکانس بالا بزرگتر یا کوچکتر می شوند . اما عموماً دامنه ضرایب با فرکانس به طور نسبتاً سریع تغییر می کند که این امر بخاطر این است که در بیشتر تصاویر طبیعی فرکانس های مکانی بالا انرژی کمی دارند . بنابراین DCT ، اثر قابل توجهی در تمرکز انرژی روی تعداد نسبتاً کمی از ضرایب ، که در گوشه چپ بالا جای دارند ، دارد به علاوه این ضرایب از یکدیگر مستقل هستند .

تا این مرحله اطلاعاتی از بین نرفته است و عمل تبدیل DCT برگشت پذیر می باشد . به علت خصوصیات فیزیولوژیکی و روانی بینایی انسان (حساسیت کمتر به فرکانس های مکانی بالا) این امکان وجود دارد که بدون کاهش قابل مشاهده در کیفیت تصویر ، مقادیر زیر آستانه معین شده که با استفاده از یک تابع مشخص حد آن بیان می شود ، حذف شوند. مقادیر حذف شده با صفر جایگزین می شوند که بدین عمل آستانه گذاری گویند. ضرایب باقیمانده با دقتی که با افزایش فرکانس مکانی کاهش می یابد ، کوانتیزه می شوند که با این کار مجدداً میزان اطلاعات لازم برای رمزنگاری یک بلوک کاهش می یابد. این دو پردازش انجام شده برگشت ناپذیر هستند اما اثر کمی را روی کیفیت تصویر مشاهده شده خواهند داشت .

3-2 فشرده سازی تصویر (Video Compression)

1-3-2 فشرده سازی تصاویر ثابت (JPEG)

اولین تلاشهای علمی که هدفشان کاهش میزان اطلاعات لازم برای رمزنگاری تصاویر ثابت بود در سال 1980 شروع شد. هدف اولیه آنها کاهش قابل ملاحظه اندازه فایل های گرافیکی و عکس ها در ذخیره سازی و انتقال آنها بود. در سال 1990، سازمان جهانی استاندارد ISO، یک گروه کاری بنام JPEG تشکیل داد که وظیفه آن، تشریح دقیق یک استاندارد بین المللی برای تصاویر ثابت با قدرت تفکیک متفاوت در فرمت C_b ، C_r ، Y یا RGB بود. استاندارد بین المللی حاصل بنام JPEG معروف شد، و در سال 1993 تحت مرجع ISO/IEC10918 منتشر شد که به عنوان ابزاری برای فشرده سازی تصاویر ثابت به کار می رود. در اینجا گام های اصلی این استاندارد بررسی می شود چون روش JPEG تا حدود زیادی الهام بخش روش MPEG بوده است.

قابل توجه است که فشرده سازی JPEG بسته به کاربرد و ضریب فشرده سازی مورد نظر می تواند با اتلاف و بی اتلاف باشد. ما فقط روش فشرده سازی با اتلاف را بررسی خواهیم کرد زیرا رمزنگاری تصاویر درونی در MPEG از همان مراحل استفاده می کند. در ضمن اینکه در این روش بسته به محتویات تصویر، ضرایب فشرده سازی بیشتر از 10، بدون هیچ نقصان قابل توجهی در کیفیت تصویر بدست می آید.

فشرده سازی با اتلاف در شش مرحله قابل توضیح است:

تبدیل تصویر به بلوکها: تصویر عموماً در فرمت C_b C_r Y به بلوکهای اولیه $8*8$ نقطه ای تقسیم می شود.

Melec.ir

تبدیل کسینوسی گسسته: همانطور که قبلا گفته شد، DCT که به هر بلوک $Y C_b C_r$ اعمال می شود ، برای هر یک ، یک ماتریس $8*8$ با ضرایبی که فرکانس مکانی آنها با دور شدن از مبدا افزایش می یابد ، تشکیل می دهد که مولفه dc نشان دهنده متوسط رنگ یا روشنایی بلوک است . مقدار این ضرایب با دور شدن از مبدا سریعاً کاهش می یابد و مقادیر نهایی یک سری اعداد کوچک و حتی صفر هستند بنابراین اگر بلوک از رنگ و روشنایی یکنواخت تشکیل شده باشد فقط مقادیر dc صفر نخواهد بود و فقط این ضرایب فرستاده خواهد شد .

آستانه گذاری و کوانتیزاسیون: این مرحله ویژگی بینایی انسان را در نظر می گیرد . مخصوصاً این واقعیت که چشم نمی تواند جزئیات ریز یک سطح روشنایی مشخص را تشخیص دهد . این مرحله شامل صفر کردن مقادیر زیر آستانه از پیش مشخص شده و بعد از آن کوانتیزاسیون ضرایب باقیمانده می باشد که دقت آن برای افزایش فرکانس، کاهش می یابد .

مرور زیگزاگ: به جز ضریب DC که پردازشی جداگانه دارد ، 63 ضریب باقیمانده (AC) ، برای

تبدیل ماتریس به جریانی از داده ها که برای گام های بعدی رمزنگاری (RLC,VLC) مناسب

باشند، بطور زیگزاگ مرور می شوند.

RLC : به منظور ایجاد امکان استفاده بهینه از سری طولانی صفرهایی که بوسیله کوانتیزاسیون و

مرور زیگزاگ تولید می شوند ، تعداد وقوع صفرهای رمزنگاری شده و بدنبال آن اولین مقدار غیر

صفر آورده می شود. این عمل میزان اطلاعاتی که باید منتقل شوند را کاهش می دهد .

VLC : این مرحله بعنوان آخرین گام ، برای رمزنگاری ، از جدول تبدیل به این صورت که مقادیر

با تکرار زیاد با طول کوتاه و مقادیر باتکرار کمتر با طول بیشتر کد شوند، استفاده می شود.

2-3-2 فشرده سازی تصاویر متحرک (استانداردهای MPEG)

معرفی استاندارد و تاریخچه MPEG

در سال 1990، نیاز به ذخیره و بازیابی تصاویر متحرک و صدای متناظر در فرمت دیجیتال برای استفاده در کاربردهای چند رسانه ای در سطوح مختلف، ISO را به تشکیل یک گروه متخصص در همان زمینه مانند JPEG با اعضای از شاخه های مختلف (صنعت، ارتباطات راه دور، سازندگان تجهیزات خانگی و ...) وادار کرد که این گروه MPEG نامیده شد.

اولین نتیجه کار این گروه استاندارد بین المللی ISO/IEC11172 بود که به MPEG-1 معروف شد. هدف اصلی این استاندارد دسترسی به امکان ضبط برنامه ویدئویی زنده و صدای استریو روی CD-ROM یا CD-I (در آن زمان بصورت تک سرعت بود) که به نرخ بیت 1/5Mb/S منجر می شد. علاوه بر اضافات مکانی ذاتی که بوسیله JPEG برای تصاویر ثابت مورد استفاده قرار می گرفت، در رمزنگاری تصاویر متحرک از اضافات زمانی خیلی مهم بین تصاویر متوالی که یک رشته ویدئویی را بوجود می آورند، استفاده می شود.

با در نظر گرفتن این هدف که به نرخ فشرده سازی بسیار بالا (نرخ بیشتر از 100 نسبت به تصویر اصلی 4:2:2 CCIR) برای این کاربردها باید رسید، در ابتدای پردازش این مشکل را با قربانی کردن قدرت تفکیک حل کردند. فرمت استفاده شده برای تصاویری که رمزنگاری می شوند فرمت SIF است (این فرمت، تصاویر با مرور تمام خطوط 288*360 و یا 240*360@30Hz را بسته به استاندارد سیگنال ویدئوی اصلی تولید می کند) که تقریباً کیفیت مشابه با کیفیت ویدئوهای خانگی خواهد داشت.

الگوریتم های فشرده سازی صدا که برای کانال های صوتی همراه در این فرمت استفاده می شوند، به MUSICAM معروف هستند و در سیستم رادیوی دیجیتال اروپا (DAB) استفاده می شود.

- سیستم MPEG-1 دارای سه بخش متمایز است که در نوامبر 1992 منتشر شد.

- سیستم MPEG-1 که ساختار مالتی پلکسی آن را تعریف می کند .

- تصویر MPEG-1 که رمزنگاری تصویر MPEG-1 را تعریف می کند.

- صدای MPEG-1 که رمزنگاری صدای MPEG-1 را تعریف می کند.

به هر حال به دلیل اینکه MPEG-1 رمزنگاری تصویر با مرور یک خط در میان ویا تکامل تدریجی به سوی HDTV را در نظر نمی گیرد ، وضعیت تصویر MPEG-1 برای کاربردهای پخش مناسب نمی باشد . بنابراین گروه MPEG ، روی تعریف یک استاندارد قابل انعطاف بهینه شده برای پخش کار کردند. این استاندارد بین المللی به MPEG-2 معروف است .

MPEG-2 که مثل نسل های قبلی در سه قسمت مجزا مشخص می گردد ، در نوامبر 1994 منتشر

شد:

- سیستم MPEG-2 که جریان بیت MPEG-2 را تعریف می کند.

- تصویر MPEG-2 که رمزنگاری تصویر MPEG-2 را تعریف می کند.

- صدا MPEG-2 که رمزنگاری صدای MPEG-2 را تعریف می کند.

MPEG-2 علاوه بر دیگر کاربردها ، استاندارد رمزنگاری منبع مورد استفاده در سیستم پخش

تلویزیون دیجیتال اروپا (DVB) می باشد .

مروری بر اصول اولیه و محدودیت ها

از اصل نایکویست می دانیم در یک کانال ارتباطی با پهنای B ارسال با نرخ $r \leq 2B$ به صورت مستقل و بدون تداخل سمبل ها میسر می باشد .

در سرویس تلویزیونی دیجیتال باید به گونه ای باشد که ارسال تمامی اطلاعات پهنای باند مرسوم فعلی برای کانال تلویزیونی (6 تا 8Mhz) محدود گردد . در این صورت بر اساس نایکویست به نرخ 16Msymbol/s یعنی در ارسال حداکثر 16 M/s داریم ؛ در حالی که نرخ بیت ویدئوی استاندارد دیجیتال با کیفیت استودیو از 216 Mb/s تا 270 Mb/s رو به رو هستیم ، که اگر اطلاعات صوت و نیز کدینگ کانال به آن اضافه شود حجم آن باز هم زیاد شده که پهنای باند زیادی را لازم دارد . ولی با ابداع و توسعه روش های فشرده سازی مؤثر داده ها (نظیر MPEG) تلویزیون دیجیتال جنبه واقعی پیدا کرد .

سه روش فشرده سازی بر اساس سه ویژگی

1) کدینگ بر اساس تصاویر ساکن (still image) در اینجا از ویژگی تشابه مکانی بین پیکسل های سازنده تصویر در یک قاب ثابت استفاده می شود که حذف آن در احساس بیننده اشکال ایجاد نمی کند ، که مبنای فشرده سازی JPEG است .

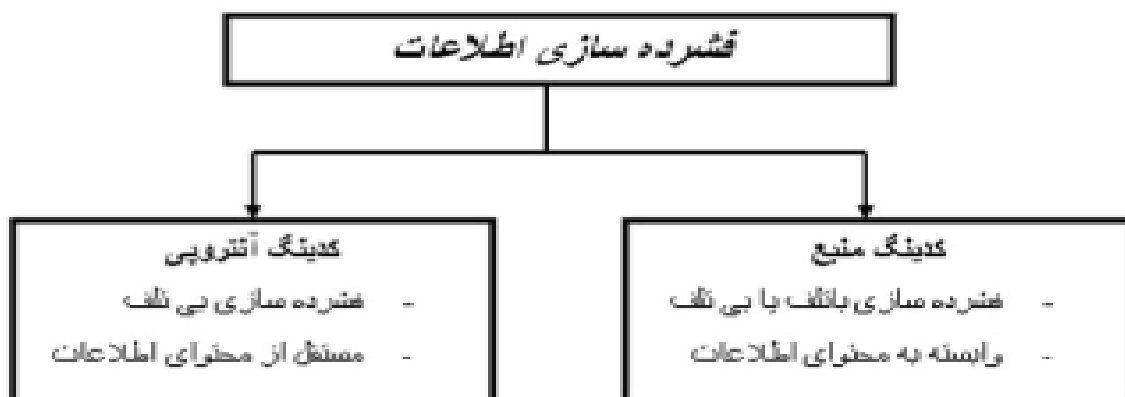
2) برای کدینگ تصاویر متحرک علاوه بر روش قبل از ویژگی تشابه عناصر مشترک بین قاب های متوالی تصویر در حوزه زمان نیز استفاده می گردد . می توان در ارسال اطلاعات تصاویر ، این داده های مشابه را حذف کرد ، که به آن کدینگ "بین قابی" گویند ، که مبنای فشرده سازی MPEG است .

3) برای هر دو نوع تصاویر ساکن و متحرک ، پس از فشرده سازی اولیه می توان از روش های آماری و ریاضی به منظور فشرده سازی بیشتر داده ها استفاده کرد که به آن کدینگ آنتروپی گویند ، که داده های اضافی و زائد یا افزونگی ها حذف می گردد .

فشرده سازی بی اتلاف و با اتلاف

در فشرده سازی بی تلف که کدینگ آنتروپی نام دارد داده های اصلی به روشی خاص کدگذاری می شوند ، و مجدداً در گیرنده به طور کامل بازسازی می شوند که البته در این حین یک سری از اطلاعات اضافی نیز به ناچار ذخیره و نگه داری می شوند ، که این نوع فشرده سازی برای اطلاعات مهم مناسب است ، که در نهایت این بهره فشرده سازی پایینی دارد .

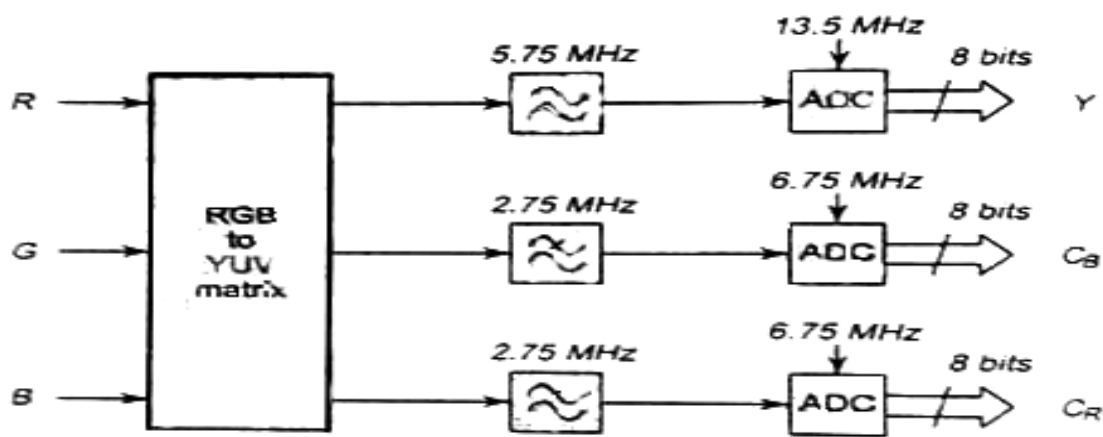
فشرده سازی با تلف یا کدینگ منبع اطلاعات ، که بر اساس سیستم های طبیعی ادراک انسانی می باشد و اقدام به حذف هر چه بیشتر اطلاعات اضافی می کند ، ممکن است مقداری از آنتروپی نیز از دست برود ؛ که این سیستم بر اساس محتوای اطلاعات فشرده سازی را انجام می دهد و دارای بهره فشرده سازی بالایی است .



شکل ۲-۵ (دسته بندی فشرده سازی)

سیگنال های پایه ای سازنده تصویر رنگی عبارتند از C_B ، C_R ، Y . که Y حامل اطلاعات سیگنال روشنایی تصاویر و C_B ، C_R سیگنال های تقاضا رنگ هستند . هر سه سیگنال بر مبنای ترکیب اطلاعات سه رنگ اصلی (RGB) ساخته می شوند.

هر عنصر از صفحه دارای سه جز می باشد که شامل Y مؤلفه روشنایی و C_B ، C_R می باشد که



شکل ۶-۲

$$Y = 8 \times 13.5 = 108$$

$$C_B = 8 \times 6.75 = 54$$

$$C_R = 8 \times 6.75 = 54$$

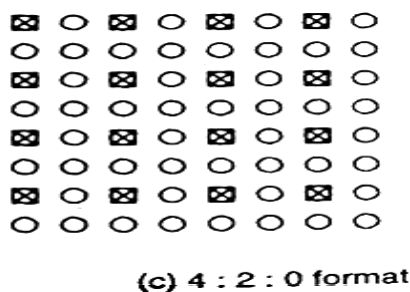
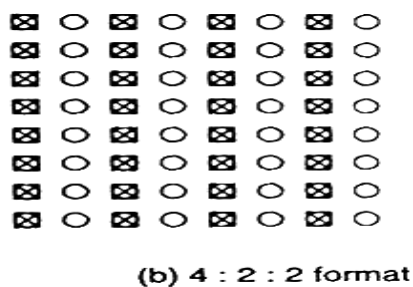
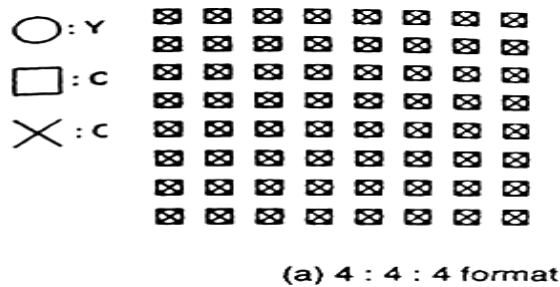
$$\text{Total} = 216 \text{ Mbit/s}$$

(ترکیب اطلاعات سه رنگ اصلی)

برای نمونه برداری قالب Y از فرکانس 13.5 Mhz استفاده می شود که محتوای قالب اول روشنایی و محتوای دو قالب دیگر اجزای رنگ می باشند و وابسته به قالب اول هستند .

در حالت 4:4:4 که هر پیکسل شامل این سه قالب می شود بیشترین حجم را پیدا می کند . برای انتقال نیز بالاترین نرخ بیت را دارد ، که به بیش از 405Mb/s می رسد . از آنجایی که حساسیت چشم نسبت به رنگ بسیار کمتر از روشنایی است، با کاهش نسبی اطلاعات مؤلفه رنگ هیچ گونه اختلال ظاهری در تصویر مشاهده نمی شود . و در آن فرکانس نمونه برداری هر سه عنصر برابر هم می باشد.

در حالت 4:2:2 تعداد پیکسل های رنگ به دست آمده در هر خط افقی نصف پیکسل های روشنایی است پس وضوح رنگ در بعد افقی نصف دقت روشنایی است و دلیل آن نمونه برداری سیگنال اختلاف رنگی با نصف فرکانس نمونه برداری روشنایی است (یعنی فرکانس 6.74Mhz) و نرخ بیت آن به 216 Mb/s می رسد. (شکل قبل)



در حالت 4:2:0 نمونه ها مانند قبل در بعد افقی و هم در بعد عمودی یک در میان شده است در این حالت یک وضعیت متقارن در تعداد و چیدمان نمونه های رنگ در دو بعد عمودی و افقی ایجاد می گردد، که از نظر بینایی چشم مناسب تر است و کیفیت تصویر نسبت به قبل تغییر محسوسی نخواهد کرد. از این قالب بندی برای انتقال و ارسال تصاویر و به عنوان قالب استاندارد تصاویر MPEG-2 استفاده می شود. در این حالت فرکانس نمونه برداری سیگنال روشنایی چهار برابر سیگنال اختلاف رنگی می باشد.

شکل ۷-۲ (حالت ها و فرمت های فشرده سازی)

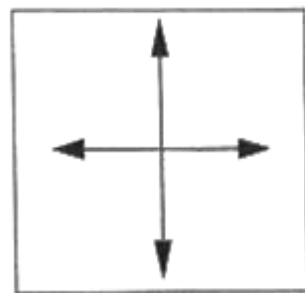
مروری بر مفاهیم پایه آنتروپی و افزونگی ها :

بر اساس نظریه اطلاعات هر بخشی از داده ها که از الگوی معین و تکراری پیروی کند ، طبق تعریف قابل پیش بینی است و جز داده های اضافی محسوب می شود . از سوی دیگر بخشی از داده ها از الگوی معین و خاصی پیروی نکنند و قابل پیش بینی نباشند به عنوان اطلاعات ضروری محسوب می شوند و آنتروپی نامیده می شوند و اختلاف بین نرخ بیت اطلاعات و کل نرخ بیت را افزونگی ها (redundancy) می باشد .

محتوای اطلاعات یا آنتروپی تابع این است که احتمال وقوع آن چقدر باشد ؛ مثلاً موج سینوسی بالاترین احتمال وقوع را دارد ، چون همه سیکل های آن مشابه هم هستند که بر اساس قضیه شانون سیگنالی که به طور کامل قابل پیش بینی باشد هیچ اطلاعاتی را حمل نمی کند . موج سینوسی چنین است و تک فرکانس می باشد و هیچ پهنای باندی ندارد .

فشرده سازی مکانی و زمانی

- اگر در فشرده سازی تصویر ، تصویر به طور جداگانه بدون اینکه وابستگی آن به تصویر های دیگر در نظر گرفته شود و محور زمان در آن دخیل نباشد به آن روش فشرده سازی "درون قابی" (intra-coded) یا کدینگ مکانی گویند .

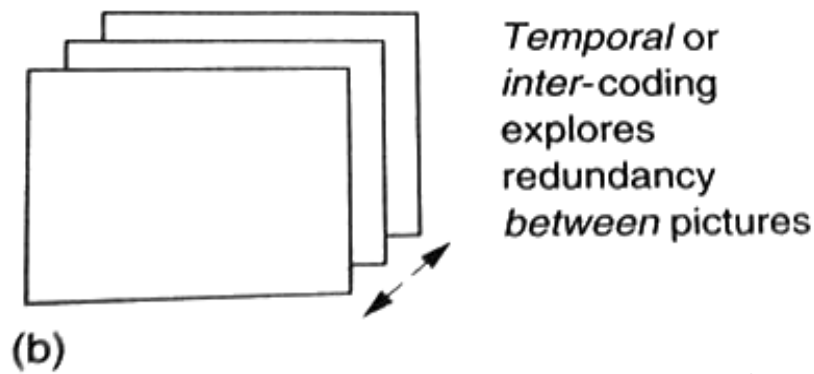


*Spatial or
intra-coding
explores
redundancy
within a picture*

(a)

شکل ۸-۲ (فشرده سازی درون قابی)

- فاکتور فشرده سازی بالاتر هنگامی به دست می آید که اضافات یک تصویر به تصویر بعدش را در نظر بگیریم که این خود شامل محور زمان می باشد . فشرده سازی “بین قابی” (inter-coded) یا کدینگ زمانی گویند.



شکل ۹-۲ (فشرده سازی بین قابی)

کدینگ بین قابی

اندیشه اصلی آن “ذخیره سازی تغییرات تصویری بین دو قاب متوالی” است که این تغییرات ، تغییراتی در شدت روشنایی ، رنگ و حرکت در عناصر تصویر می باشد ؛ مثلاً در مورد حرکت اگر به طریقی مقدار حرکت را پیش بینی کنیم می توانیم تفاوت با تصویر قبل را محاسبه کرده و آن را ارسال کنیم و میزان داده ها را کاهش دهیم . برای این ایده نیازمند برخی قاب های تصویر به عنوان قاب مرجع هستیم تا تغییرات تصویر را نسبت به آن ها بسنجیم به این ترتیب سه نوع قاب در MPEG داریم .

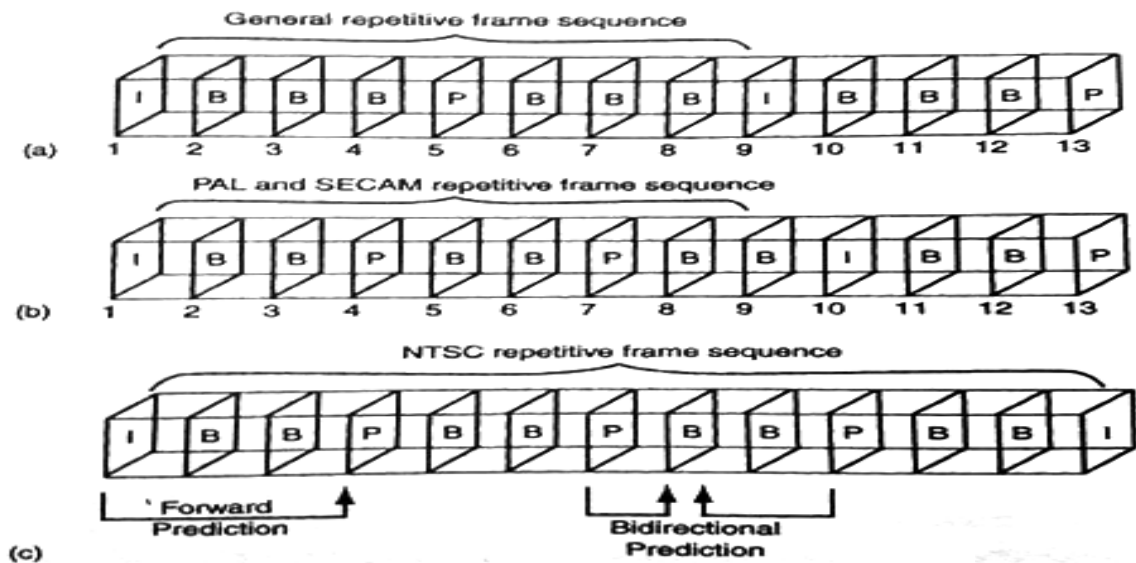
قاب نوع I (Intra-Coded Frame) : که همان قاب مرجع است این قاب کاملاً از الگوی

JPEG استفاده می کند و مستقل می باشد .

قاب های نوع P و B : کدینگ این قاب ها از روش کدینگ بر اساس پیش بینی استفاده می شود .

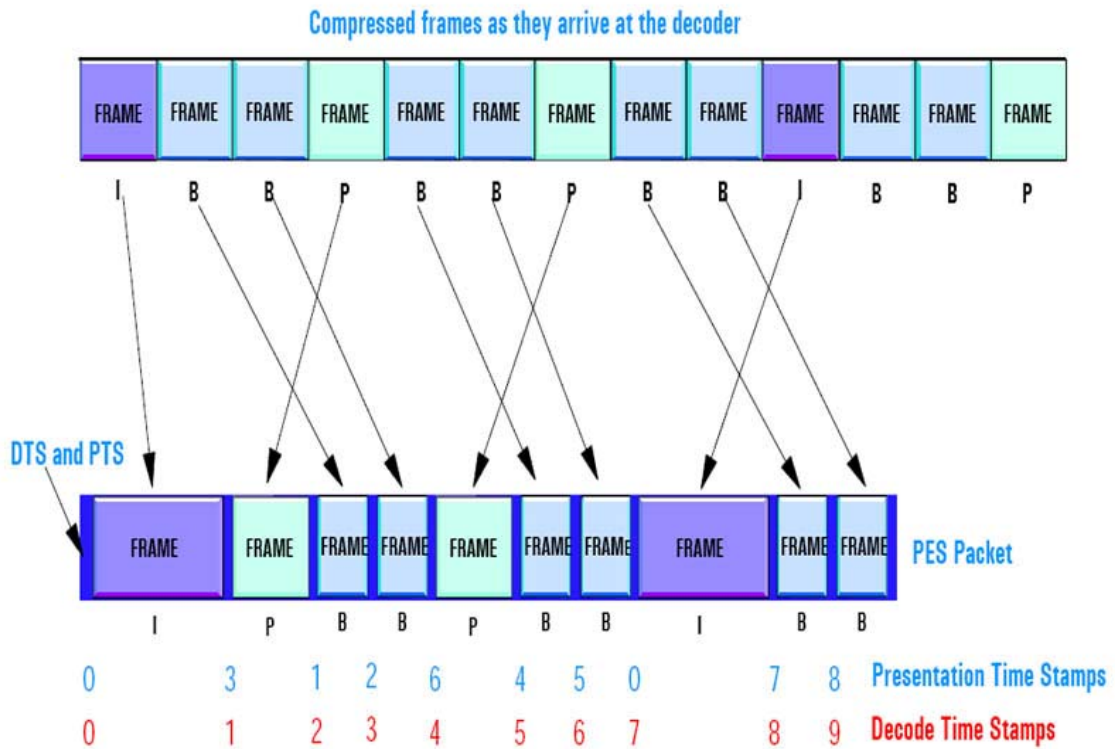
که در فریم P (Predicted-frame) از پیش بینی تصویر “پیش رو” که بر اساس تصویر مرجع

قبل می باشد انجام می پذیرد. و در فریم B (Bidirectional-frame) که این قاب با در نظر گرفتن قاب قبلی و بعدی خود کد می شود و به عنوان قاب دو جهتی شناخته می شود. که توالی این قاب ها گروه تصاویر (GOP) را تشکیل می دهند. که ترتیب های مختلف آن به شکل صفحه بعد است.



شکل ۱۰-۲ (روشهای کدینگ بین قابی)

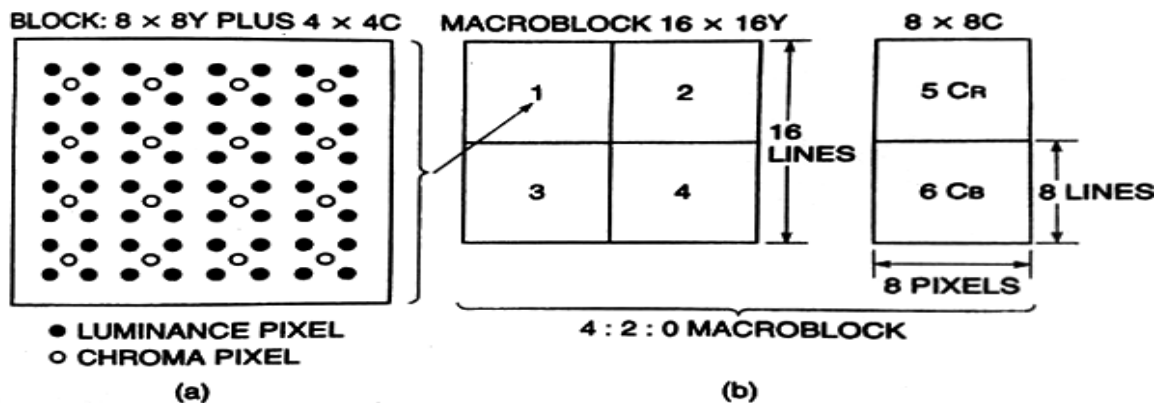
- نکته مهم اینکه توالی و ترتیب قاب ها هنگام کد شدن با ترتیب آن ها هنگام ارسال و دریافت یکسان نیست. مثلاً برای ارسال قاب P باید ابتدا قاب I موجود باشد؛ پس ابتدا قاب I ارسال می شود و سپس P ارسال می شود و همچنین چون قاب B نیاز به I و P دارد، پس هنگام ارسال قاب P قبل از B ارسال می شود؛ که به این کار رتبه بندی دوباره (Re-Ordering) می گویند. که زمان پخش و دکد کردن این فریم ها از روی ترتیب با DTS و PTS می باشد. مانند شکل زیر:



شکل ۱۱-۲ (فشرده سازی فریم ها هنگام رسیدن به دکدر)

ماکروبلوک در MPEG :

کوچکترین واحد تصویر که پردازش می شود را ماکروبلوک گویند که 16×16 پیکسل است. در الگوی $4 : 2 : 0$ هر ماکروبلوک شامل چهار بلوک 8×8 از پیکسل های روشنایی (Y) و دو بلوک 8×8 پیکسلی برای رنگ (سیگنال های Cr و Cb)



شکل ۱۲-۲ (ماکروبلوک تصویر ۱۶*۱۶ پیکسل)

بردار حرکت (Motion Vector)

تعیین بردار حرکت برای یافتن ماکروبلوک هایی می باشد که محتوای درونی آن در دو قاب تصویر متوالی تغییر نمی کند و فقط موقعیت آن ها در تصویر جابجا می شود و بردار حرکت بیانگر میزان جابجایی مکانی یک ماکروبلوک است .

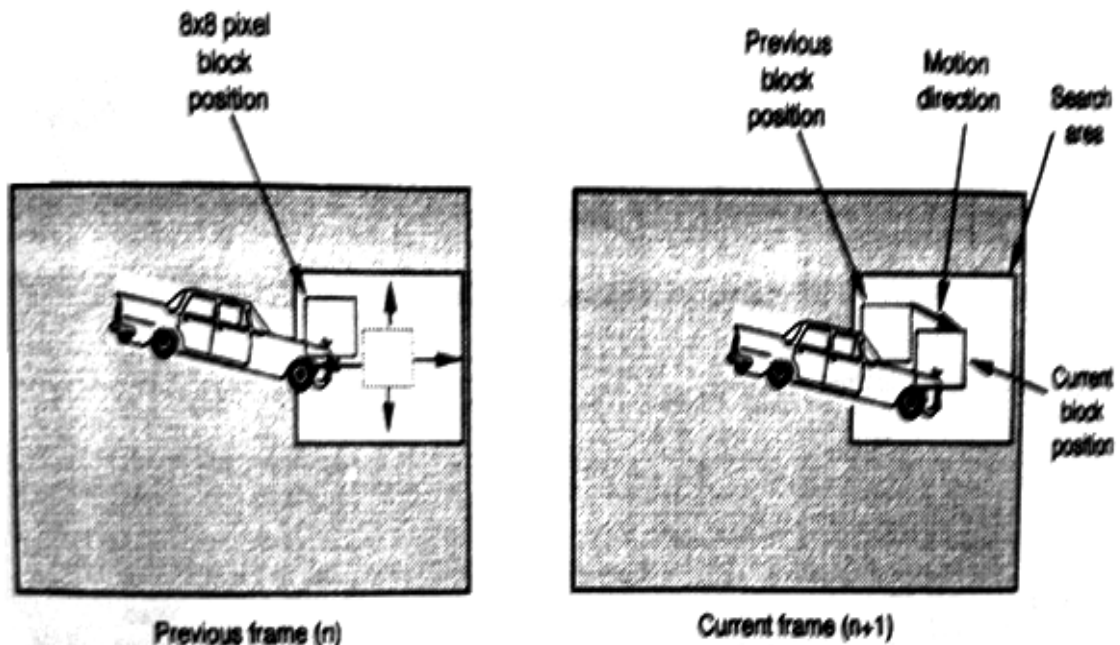
سه روش برای جستجوی ماکروبلوکی موجود عبارتند از:

تطبیق بلوکی ، تطبیق گرادیان یا شیب و همبستگی و شباهت فاز

- در روش تطبیق بلوکی :

برای پیدا کردن بردار حرکت ابتدا یک ماکروبلوک از تصویر فعلی با ماکروبلوک معادل در قاب قبل در تعدادی موقعیت مکانی متفاوت مقایسه می شود و اختلاف ها محاسبه می گردد و موقعیتی که کمترین اختلاف را داشته باشد به عنوان موقعیت درست انتخاب می شود و دیگر اطلاعات ماکروبلوک به دلیل تکراری شدن دیگر فرستاده نمی شود؛ حال اگر ماکروبلوک کاملاً منطبق پیدا

نشود و بین بلوک مرجع و بلوک پیدا شده مقداری خطا باشد ؛ علاوه بر بردار حرکت مقادیر خطا نیز ارسال می شود .



شکل ۱۳-۲ (تطبیق بلوک های متوالی)

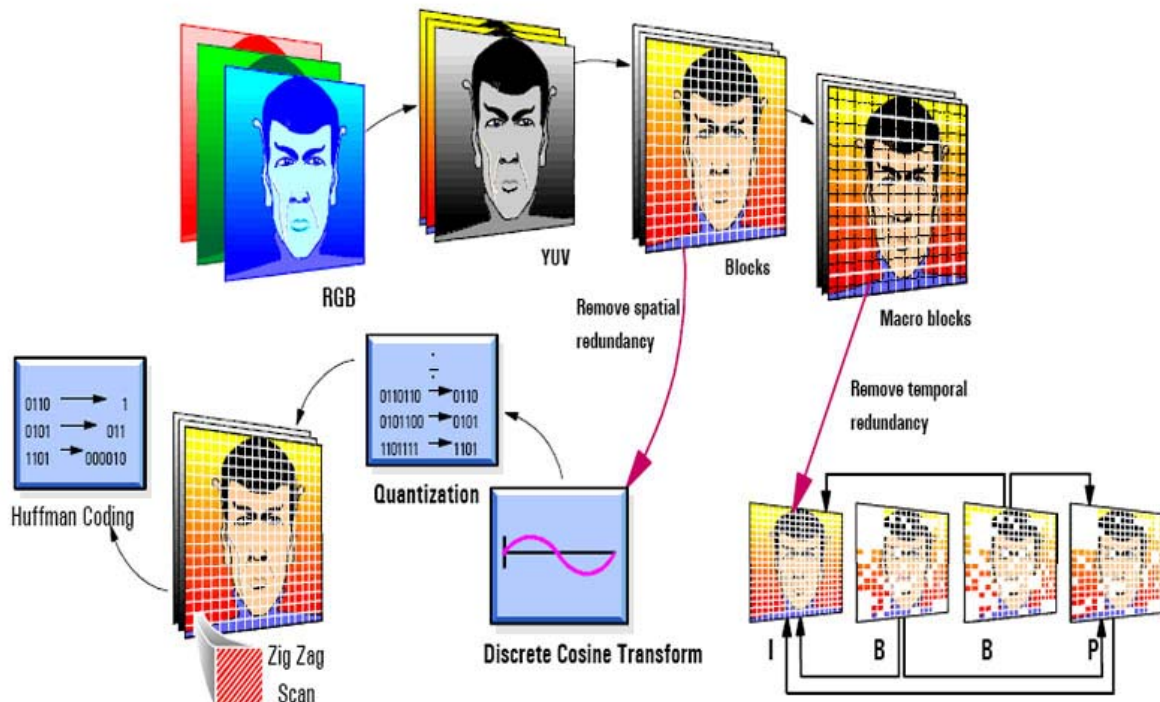
مراحل فشرده سازی MPEG

مراحل فشرده سازی به صورت شکل زیر است :

به این صورت است که ابتدا فریم های YUV از تصاویر RGB به دست می آیند ، در گام بعدی بلوک های 8x8 تصویر آماده می شوند که برای هر فریم Y، U، و V یک ماتریس 8x8 جداگانه داریم ، که کلاً برای یک تصویر PAL که شامل 720 سطر و 576 ستون 720* 576 تعداد کل بلوک های YUV به 19440 می رسد $[(720/8) \times (576/8) \times 3]$. در گام بعدی با استفاده از عملگر ریاضی DCT (تبدیل کسینوسی گسسته) عناصر ماتریس از حوزه مکان به حوزه فرکانس منتقل می شوند و یک ماتریس 8x8 جدید تشکیل می دهند ؛ به اولین عنصر این ماتریس

ضریب DC گویند ، در واقع میانگینی از دامنه تمامی پیکسل های درون بلوک است ، سایر عناصر ماتریسی DCT ضریب AC می باشند . در این ماتریس هر چه به سمت پایین و راست حرکت کنیم عناصر حاصل ، ضرایب فرکانس های بالاتر می باشند .

در گام بعدی که کوانتیزاسیون می باشد بر اساس ویژگی های روانشناسی بینایی انسان ، یک



شکل ۱۴-۲ (مراحل کوانتیزاسیون و فشرده سازی)

جدول کوانتیزاسیون (Q-Table) که یک ماتریس 8×8 می باشد تنظیم می شود و تک تک عناصر ماتریس DCT بر آن تقسیم می شود ، و مقدار آن ها گرد شده و ماتریس جدید به وجود می آید ؛ اصولاً آرایش اعداد در ماتریس کوانتیزه شده به گونه ای است ضرایب فرکانس های بالاتر صفر می باشد ، که مشکلی ایجاد نمی کند ، چون چشم انسان روی بخش های بسیار ظریف تصویر (شامل فرکانس های بالا) چندان حساس نیست .

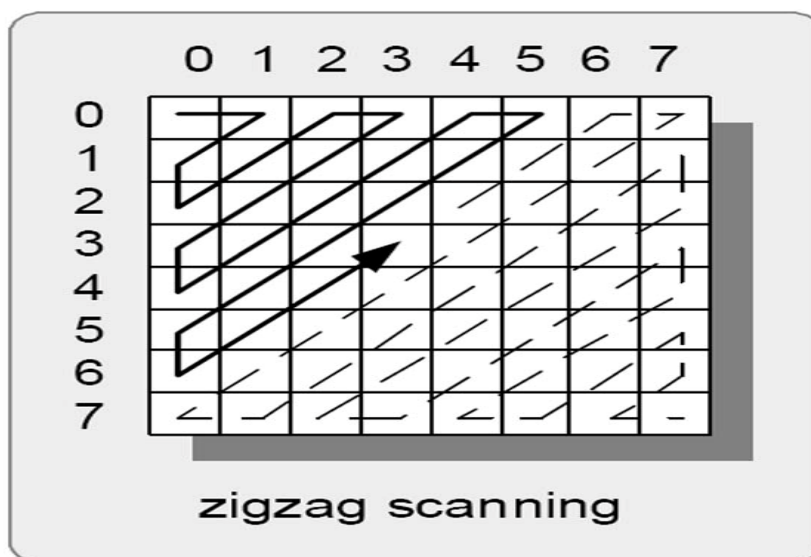
در گام بعدی که کدینگ بر مبنای طول تکرار می باشد ، ماتریس به وجود آمده از کوانتایزر 63 عنصر آن (ضرایب AC) به صورت زیگزاگ اسکن می شوند . این کار به این دلیل است تا بخت

Melec.ir

قرار گرفتن عناصر غیر صفر کنار هم بیشتر شود. که در این صورت ضرایب فرکانس بالا به صورت دنباله های طولانی غیر صفر پشت سر هم می شوند، که فرم RLC به این صورت است.

(اولین مقدار غیر صفر بعدی، تعداد صفر در رشته تکرار) $RLC =$

و آخرین عنصر آن EOB که معرف پایان بلوک می باشد است.



شکل ۱۵-۲ (اسکن تصویر به روش زیگزاگ)

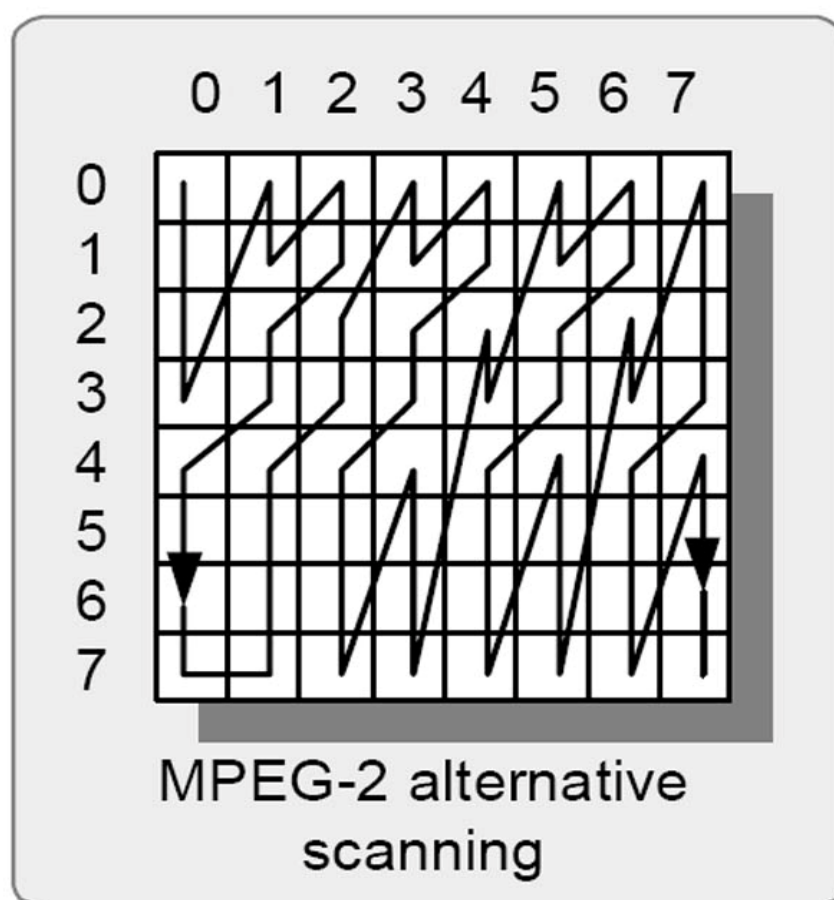
و برای ضرایب DC از تکنیک DPCM برای کاهش نرخ بیت ارسالی استفاده می شود.

در قدم آخر با استفاده از کدینگ آماری که همان کدینگ هافمن می باشد (بر مبنای احتمال وقوع سمپل ها می باشد) فشرده سازی می شود. که به این طریق I-frame ساخته می شود و از روی آن فریم های P و B تشکیل می شوند.

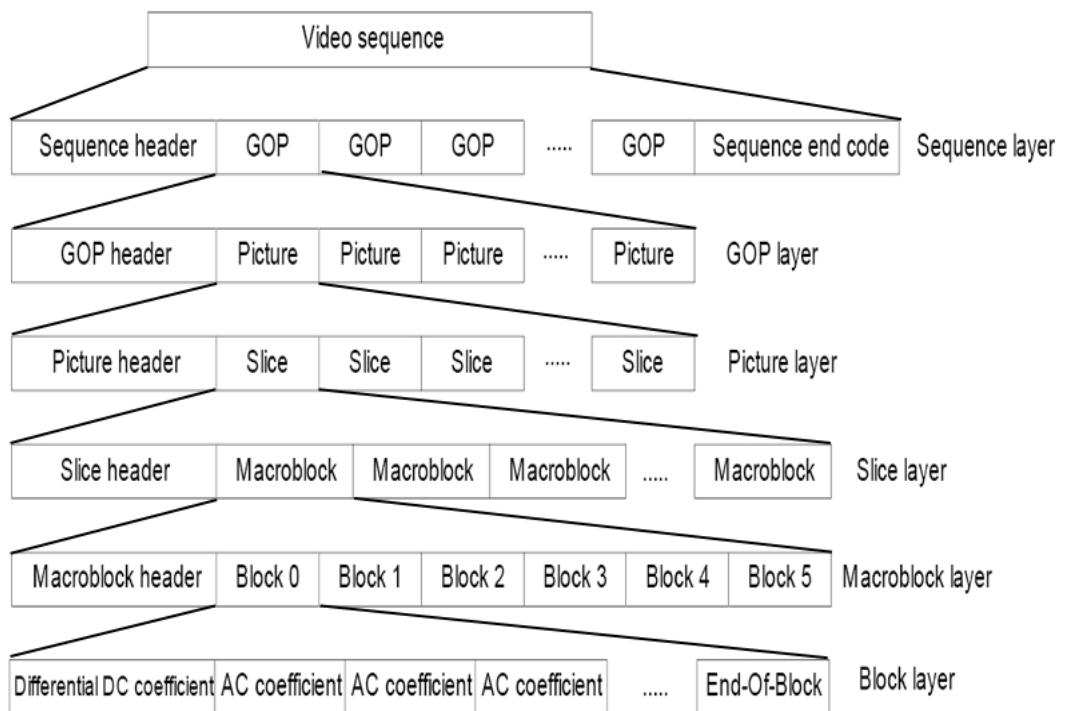
مرحله کوانتیزه کردن جزء فشرده سازی با تلف می باشد و کدینگ منبع است. ولی دو مرحله آخر فشرده سازی بی تلف بوده و کدینگ آنروپی می باشد. که در حالت با تلف به نسبت فشرده سازی 20 به 1 و در حالت بی تلف به نسبت 2 به 1 می رسیم.

Melec.ir

در MPEG-2 چون از تصاویر Interlace (در هم بافته) استفاده می‌کنیم و تشابه کمتری بین پیکسل‌ها در بعد عمودی وجود دارد، برای ضرایب DCT از اسکن متناوب (Alternate) استفاده می‌شود که نتیجه بهتری نسبت به اسکن زیگزاگ می‌دهند. (مانند شکل بعد)

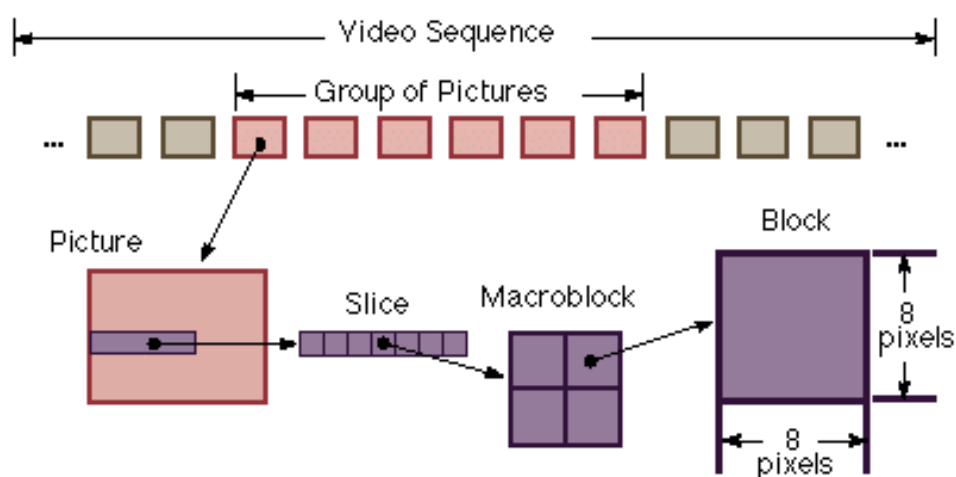


شکل ۱۶-۲ (اسکن تصویر به صورت متناوب)



شکل ۱۷-۲ (سلسله مراتب داده های تصویر ارسالی)

که هر کدام از اجزا به صورت زیر می باشند:



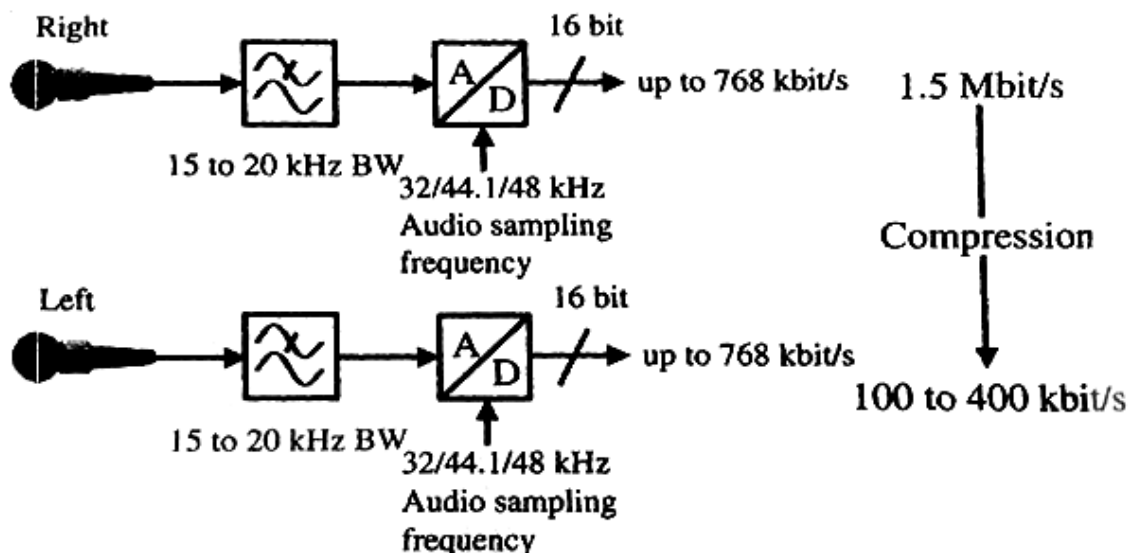
شکل ۱۸-۲ (اجزای تصویر ارسالی)

4-2 فشرده سازی صوت

سیگنال آنالوگ صدای تولید شده از یک سیگنال الکتریکی پیوسته با دامنه ولتاژ متغیر با زمان در بازه فرکانسی 20 Hz تا 20 KHz تشکیل شده است . تجهیزات استودیو صدا از نظر دامنه دو گروهند ، دامنه پایین مانند میکروفون و دامنه بالا مانند میکسر ها ، آمپلی فایر های صوتی و بلندگوها می باشند .

به دلیل گستره پویایی آکوستیکی بالا و دامنه وسیع شنوایی انسان ، فشرده سازی و کدینگ صدا باید با دقت بیشتر انجام گیرد ؛ به عنوان مثال اگر در نمونه برداری تصویر ، اختصاص دادن 8 بیت به هر نمونه تصویر يك پیکسل کافی به نظر می رسيد ، در صدا این مقدار اصلاً کافی نیست و در سیستم صوتی دیجیتال ، تعداد بیت لازم برای معرفی هر نمونه صوتی باید حداقل 12 بیت (برای کیفیت غیر حرفه ای) باشد . جالب اینکه از نظر مهندسی می توانیم گستره دامنه شنوایی انسان را مشابه یک سیستم دیجیتال با قدرت تفکیک حداقل 20 بیت فرض کنیم .

در صدای دیجیتال آنالوگ قبل از نمونه برداری توسط یک فیلتر پایین گذر از نظر باند فرکانسی محدود می شوند . نرخ نمونه برداری آن برای صوت با کیفیت پایین از فرکانس 32 KHz ، برای کیفیت در حد دیسک فشرده 44.1 KHz و برای کیفیت بالای صدای حرفه ای استودیویی 48KHz و حتی 96KHz استفاده می شود و حد تفکیک معمولاً حداقل 16 بیت است .



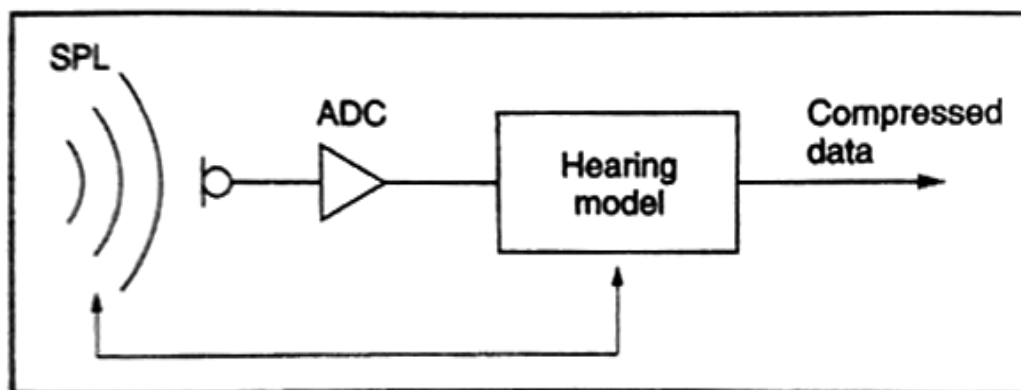
شکل ۱۹-۲ (نمونه برداری صوت در استودیو)

که برای کیفیت حرفه ای با نرخ نمونه برداری 48 بیت و حد تفکیک 16 بیت ، نرخ بیت تولید شده 768 Kbit/s می شود که این مقدار در استودیو 1.5 Mbit/s می باشد . هدف از فشردن سازی رساندن این مقادیر به 400 Kbit/s تا 100 Kbit/s است ؛ که به منظور بهینه کردن اشغال پهنای باند ارسال انجام می شود . که فشردن سازی بر اساس دو ایده اصلی بنا نهاده شده است : یکی وجود اضافات یا افزونگی ها و دیگری وجود اطلاعات نا محسوس یا غیر ضروری و قابل صرف نظر در صدا.

که یکی بر اساس فشردن سازی بر مبنای وجود تکرارها و شباهت ها و پیدا کردن الگوی خاص و تکرار شونده داده هاست ؛ که داده هایی که تابع الگوی معین و قابل پیش بینی هستند را اضافات محسوب می کنند . در این روش که روش "کاهش افزونگی های صدا" نام دارد فشردن سازی بی تلف محسوب می شود ، چون هیچ داده مفیدی از دست نمی رود . در این حالت داده ها ابتدا بدون توجه به ماهیت شان با توجه به روش های آماری خاص کد شده و فرستاده می شوند و در گیرنده عمل عکس انجام شده و کاملاً بازیابی می شوند ، که به یک نسبت فشردن سازی 2 به 1 می رسیم . در بخش دوم که نوعی فشردن سازی با تلاف محسوب می شود ، بر اساس خواص سیستم شنوایی

انسان است . که بر اساس آن بسیاری از اطلاعات که قابل تشخیص و درک برای انسان نیست حذف می گردد ، که دیگر قابل بازیابی نیست .

- هر دو بخش مجموعاً تا حداکثر 90% قادر به کاهش نرخ داده ها هستند .



شکل ۲۰-۲ (بلوک دیاگرام کاهش افزونگی صوت)

که SPL فشار ناشی از صوت بر حسب " نیوتون بر متر مربع " است .

یکی از روش های کدینگ صدا MUSICAM می باشد که سیگنال صوتی را به تعداد زیادی باند فرکانسی کوچک تقسیم می کند و عملیات پردازش روی باند های مجزا به نسبت های مختلف انجام می دهد . و روش دیگر ASPEC که در آن کدینگ بر مبنای توابع ریاضی می باشد ؛ در این روش با استفاده از DCT سیگنال صوتی از حوزه زمان به حوزه فرکانس رفته و مولفه های نامحسوس را حذف می کنیم. که در کدینگ صدای MPEG-1 از هر دو روش برای فشرده سازی استفاده شده که به عنوان استاندارد ISO/IEC 11172-3 پذیرفته شده است .

MPEG-1 از سه نوع طراحی با پیچیدگی های متفاوت و اصطلاحاً با سه لایه (لایه های I ، II و III) تعریف می گردد. که لایه اول نقش الگوریتم پایه را برای کدینگ صدا ایفا می کند و لایه های بعدی امکانات لایه اول را گسترش می دهند .

در حالت کامل تر از MPEG-1 کدینگ صدای MPEG-2 است که همه لایه های MPEG-1 را در بر می گیرد ؛ که علاوه بر آن لایه دوم به صورت صدای چند کاناله (Multi channel) توسعه یافته که با عنوان ISO/IEC13818-3 استاندارد گردیده و به منظور تأمین دو هدف بوده :

(1) توسعه صدا کیفیت بالا با 2 تا 5 کانال و در عین حال سازگار با روش های مرتبه پایین تر MPEG-1 .

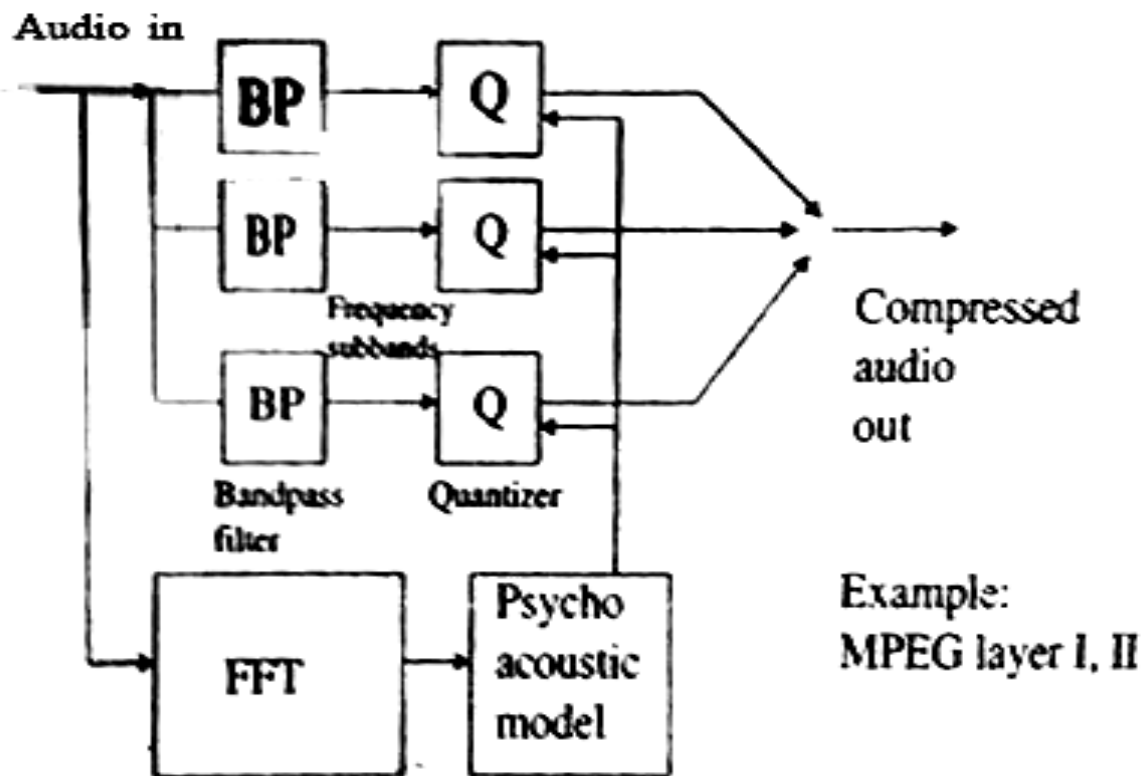
(2) کاهش نرخ بیت تولید شده در خروجی همراه با حفظ کیفیت مناسب صدا.

ارتباط تاریخی روش های کدینگ صدا در MPEG-1 و MPEG-2

در سیستم پخش تلویزیونی DVB از استاندارد ISO/IEC 13818-3 بهره می برد که توصیف کننده استفاده از MPEG در لایه های I و II می باشد . در آمریکا از سیستم ATSC و برای پخش و برای کدینگ از استاندارد Dolby AC3 استفاده می شود .

کدینگ صدای MPEG :

لایه های I و II: کدینگ این دو لایه شباهت زیادی به یکدیگر دارند . در این سیستم سیگنال صوتی آنالوگ ورودی پس از تبدیل شدن به حالت دیجیتال (PCM) از دو مسیر به صورت موازی حرکت می کند ؛ به صورت زیر :



شکل ۲-۲۱ (مثالی از مدل لایه MPEG I, II)

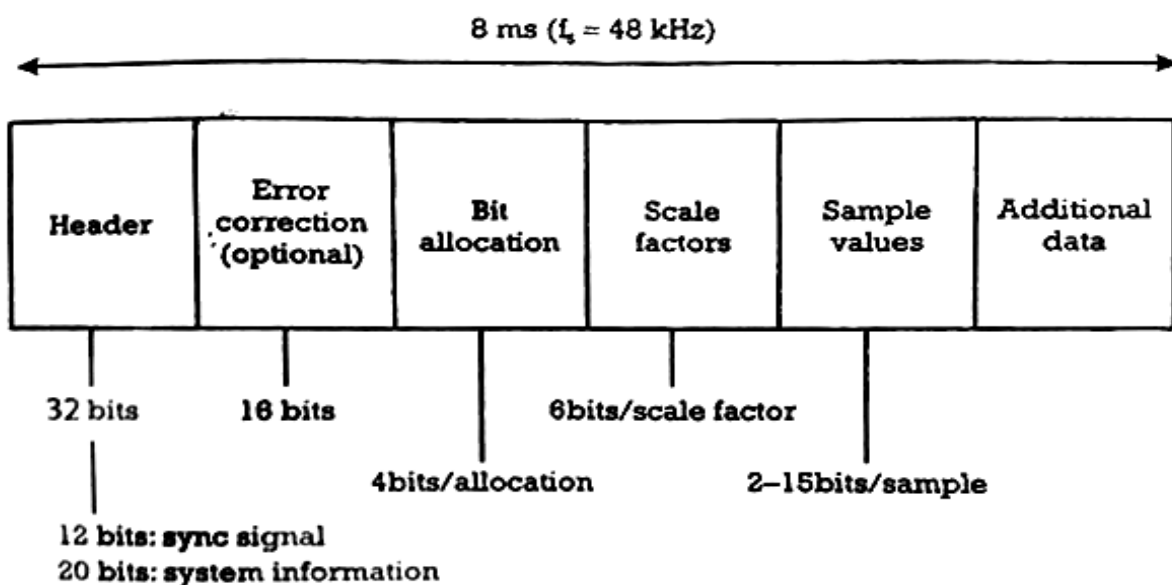
برای MPEG لایه 1 در مسیر اول طیف سیگنال ورودی توسط یک بانک فیلتر به 32 باند جزئی تقسیم می گردد که هر باند 750Hz پهنای باند دارد . که این تعداد فیلتر از لحاظ بهینه سازی در زمان تأخیر پردازش و نیز پیچیدگی سیستم مقداری مناسب است . به ازای هر فیلتر در این مسیر یک کوانتیزه کننده وجود دارد که می تواند هم اطلاعات یک باند جزئی را کاملاً حذف کند و هم در صورت لزوم عمل دوباره کوانتیزه کردن نمونه ها با دقت بیشتر را انجام دهد .

در مسیر دوم تبدیل فوریه برای 512 نقطه محاسبه می شود و مقادیر ماکزیمم در این نقاط به دست می آید ؛ و بر همین اساس الگوی پوشش صوتی به همراه نرخ بیت لازم در خروجی و در نتیجه تعداد مناسب سطوح کوانتیزه برای هر باند جزئی تعیین می شود . به این ترتیب هر چه میزان

پوشیده شدن دامنه ها بیشتر باشد دقت نمونه های ارسال شده بیشتر کاهش می یابد (چون هر چه میزان پوشش دامنه بیشتر باشد نیاز به دقت کمتری دارد.) که این کاهش دقت با کوانتیزه کردن دوباره نمونه های PCM و ساخت کد های جدید با طول کمتر می باشد، که اندازه این کاهش برای تمام واژه های جزئی گوناگون می توانند طول کد متفاوتی داشته باشند. که طبیعتاً مجموع بیت تولید شده در خروجی کمتر از تعداد بیت صدای دیجیتال ورودی است. که همان بهره کدینگ می باشد. که کاهش این نرخ بیت توسط کدینگ MPEG موجب کاهش پهنای باند فرکانس خروجی می شود.

طی پردازش صدا در MPEG ورودی صوتی PCM به تعدادی معین و ثابت دسته بندی می شود و به ازای هر باند جزئی 12 نمونه صوتی دیجیتال تشکیل یک واحد زمانی Time block را می دهند که از میان این 12 نمونه متوالی که در یک باند جزئی به دست می آید بزرگترین انتخاب شده و به عنوان عامل مقیاس (Scale Factor) تعریف می گردد که مقدار آن یک واژه 6 بیتی است. که این عامل مقیاس به ازای هر 12 نمونه وارد طبقه مدل سازی بر مبنای روانشناسی صدا (Psychoacoustic Model) می شود تا حداکثر سطوح کوانتیزه لازم برای هر گروه 12 تایی از نمونه ها محاسبه شود. در MPEG تعداد زیادی از نمونه ها صوتی PCM قالب بندی می شوند تا واحد پردازش بزرگتری را به وجود آورند که به آن قالب (Frame) گویند. در کدینگ لایه اول مجموعاً 384 نمونه زمانی PCM (با فرض 12 نمونه برای 32 باند جزئی موجود $32 \times 12 = 384$) که اگر با فرکانس 48KHz نمونه برداری شوند در 8ms زمان پردازش آن انجام می پذیرد.

که الگوی جریان بیت های MPEG-1 لایه 1 به صورت شکل زیر است :



شکل ۲۲-۲ (الگوی جریان بیت های MPEG لایه 1)

- در ابتدای این جریان Header قرار دارد که 12 بیت آن مربوط به سیگنال همزمانی و 20 بیت آن اطلاعات سیستم می باشد.

- پس از آن امکان تصحیح خطا و محافظت در گیرنده به صورت اختیاری با 16 بیت به عنوان داده های مخصوص تصحیح خطا در ساختار کد کننده وجود دارد که روش سیستم خطا یاب CRC می باشد.

- بعد از آن 32 گروه 4 بیتی از داده ها وجود دارد ، که بیانگر طول کد های هر باند جزئی می باشد . و به دکدر برای بازسازی نمونه های هر باند جزئی کمک می کند .

- بعد از آن مجموعه ای از 32 عامل مقیاس (scale factor) 6 بیتی قرار می گیرد ، که معرف ماکزیمم دامنه سنجیده شده در بین 12 نمونه که در یک باند جزئی است می باشد .

- بعد از این ها نوبت به اطلاعات اصلی می رسد ؛ که 32 مجموعه از داده ها می باشد که هر مجموعه متناظر با 12 نمونه صوتی است که از یک باند جزئی داریم که گفتیم این نمونه ها در هر باند جزئی مختلف طول های متفاوتی دارند (2bit تا 15bit). همچنین این جریان بیت می تواند حامل اطلاعات اضافی نیز باشد که در نهایت طول این جریان 8ms می شود .

در حالت کلی چون در فرکانس های بالاتر انرژی سیگنال کاهش می یابد برای کد کردن آن در این فرکانس تعداد سطوح کوانتیزه کمتر و در نتیجه تنوع بیت کمتر داریم .

جدول مقایسه لایه I و II و III

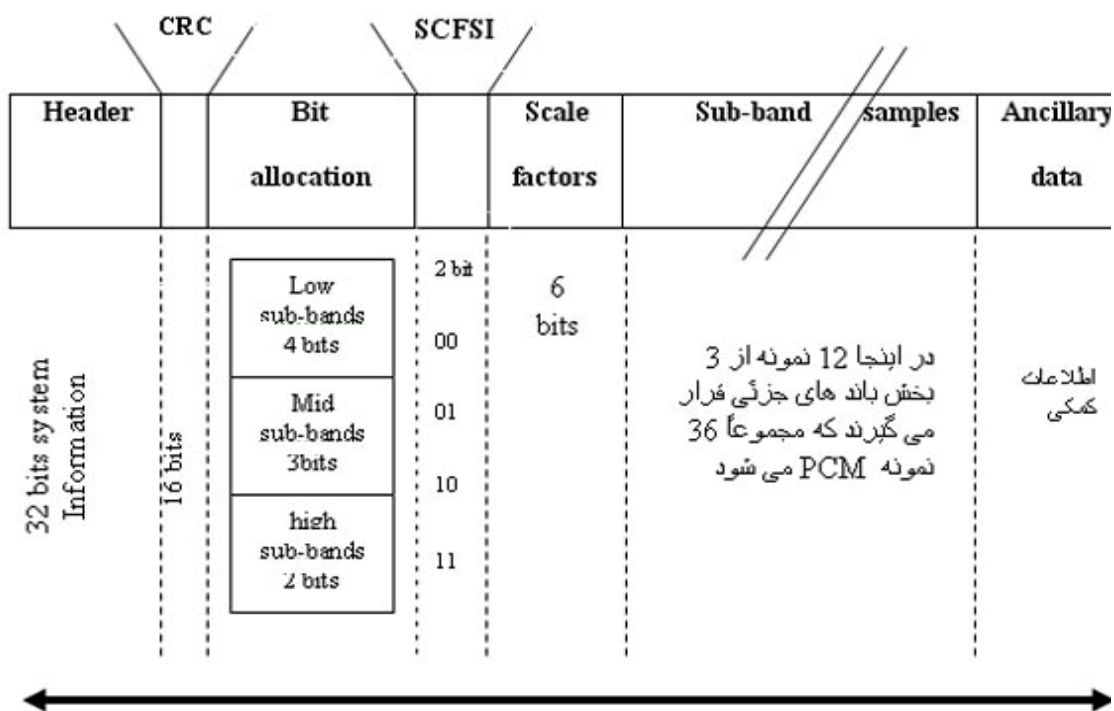
دوره	تعداد نمونه های صوتی	تعداد نمونه های هر باند جزئی	تعداد نقاط پردازش FFT (در حوزه فرکانس)	تعداد عامل مقیاس (scale factor)	پهنای هر باند جزئی	تعداد باند های جزئی	
زمانی حضور یک قالب	صوتی PCM (در قالب کامل)	هر باند جزئی	FFT (در حوزه فرکانس)	(scale factor)	باند جزئی	های جزئی	MPEG
8 ms	12 x 32 384	12	512	1	750 Hz	32	Layer I
24 ms	36 x 32 1152	36	1024	بیش از 1	750 Hz	32	Layer II
8,24 ms	1152	12-36	1024	-	750 Hz	32	Layer III

جدول ۲-۱ (مقایسه لایه های I, II, III)

یکی دیگر از اختلافات لایه اول و دوم در این است که در لایه دوم باند های جزئی به سه بخش پایین، متوسط و بالا تفکیک می شود . در فرکانس های پایین می توان 15 نوع متفاوت از واژه ها را برای نمونه ها ساخت ، با 4 بیت که نشان دهنده دقت بالاتر در کوانتیزه کردن می باشد .

در بخش متوسط 7 نوع متفاوت واژه که 3 بیت لازم دارد را می توان ساخت و در نهایت در فرکانس بالا 3 نوع متفاوت واژه را با 2 بیت می سازد .

الگوی زمانی لایه 2 به صورت زیر است که شباهت زیادی به الگوی لایه 1 دارد ، اما این لایه حامل اطلاعات جانبی بیشتری است که برای ارتقاء کیفیت صوتی و تطبیق با سیستم صدای فراگیر به کار برده می شود .



شکل ۲-۲) این فریم بر مبنای 1152 نمونه صوتی PCM که با فرکانس 48 KHz نمونه برداری

شده اند مرتب می شود . که زمان 24ms می باشد . (

Melec.ir

SCFSI (Scale Factor Select Information) : چون در لایه 2 قابلیت استفاده از سه

عامل مقیاس را داریم SCFSI به عنوان نشانگر تعداد این عامل های مقیاس استفاده می شود که دو بیت نیز می باشد (لایه 2 با سیستم لایه 1 نیز قابل تطبیق است).

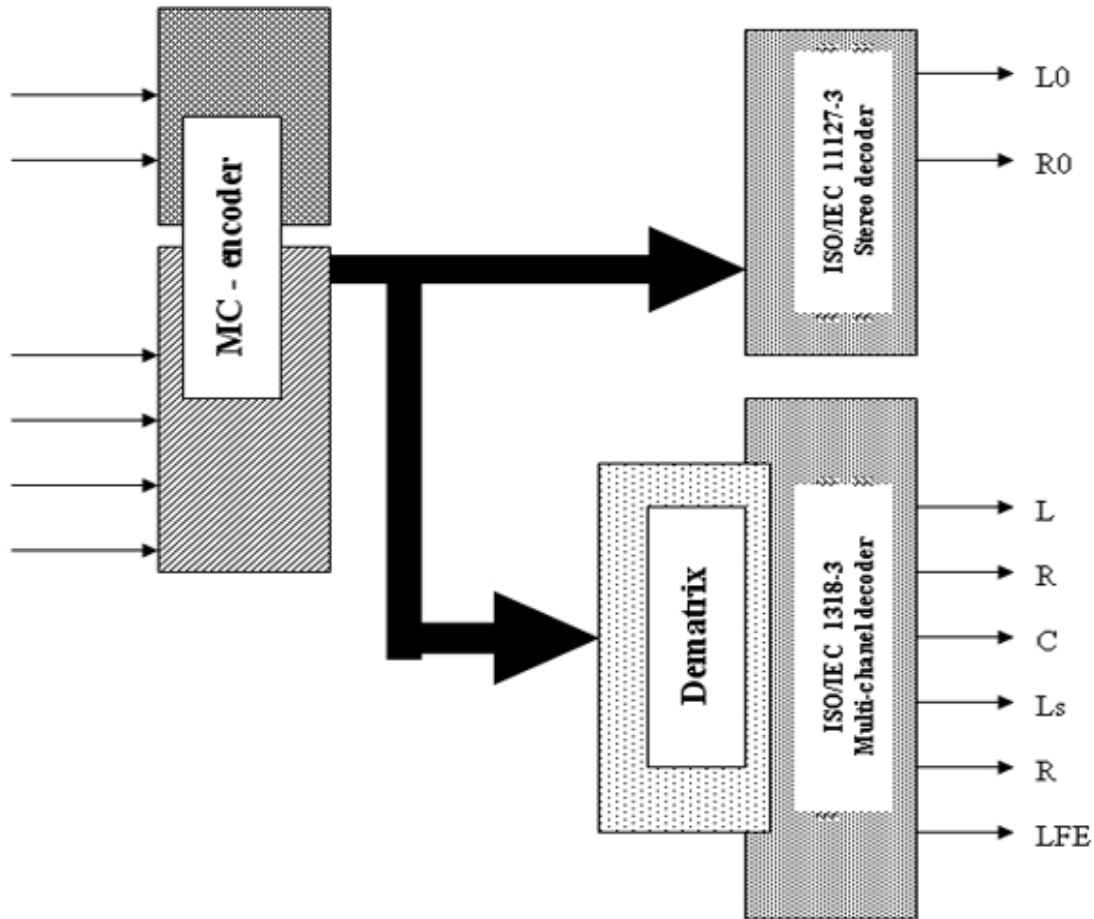
استاندارد پخش DVB فقط سیستم MPEG با لایه های I و II را پشتیبانی می کند . در استاندارد MPEG پهنای باند صدا 20 KHz می باشد و به ازای فرکانس نمونه برداری متفاوت نرخ بیت در خروجی می تواند 192Kbit/s در لایه 1 و تا 128Kbit/s در لایه 2 کاهش یابد .

لایه III :

این لایه پیچیده ترین لایه می باشد و دلیل اصلی آن به علت روبرو شدن با محدودیت شدید نرخ دیتا می باشد ، که در کاربرد های اینترنتی به عنوان MP3 معروف است . کدینگ آن بر مبنای ASPEC می باشد ، با اصلاحاتی خاص که به درجه شباهتی از لایه II برسد . کدر ASPEC در اصل مستقیماً از تابع DCT در نمونه های ورودی استفاده می کند ؛ در لایه 3 اصلاحاتی انجام می شود تا از تبدیل هیبرید استفاده شود و 32 باند موجود در لایه های 1 و 2 را ترکیب کند و اندازه بلوک 1152 نمونه را حفظ کند . در لایه 3 ، 32 باند جزئی به وجود آمده از بانک فیلتر توسط MDCT یا DCT اصلاح شده مجدداً پردازش می شود .

بر اساس ویژگی هایی که استاندارد DVB برای گیرنده صدا تعریف کرده ، روش های MPEG صدا در لایه های 1 و 2 هر دو باید توسط یک گیرنده IRD(Integrated Receiver Decoder) پشتیبانی شوند . که لایه 2 برای کد کردن صدا مناسب تر است . نرخ نمونه برداری 32 و 44.1 و 48 توسط گیرنده دیجیتال پشتیبانی می شود . گیرنده های IRD صداها را تک کاناله ، دو کاناله ، استریو الحاقی و استریو کامل را پشتیبانی می کند ، و قادرند حداقل یک زوج صدای استریو را از

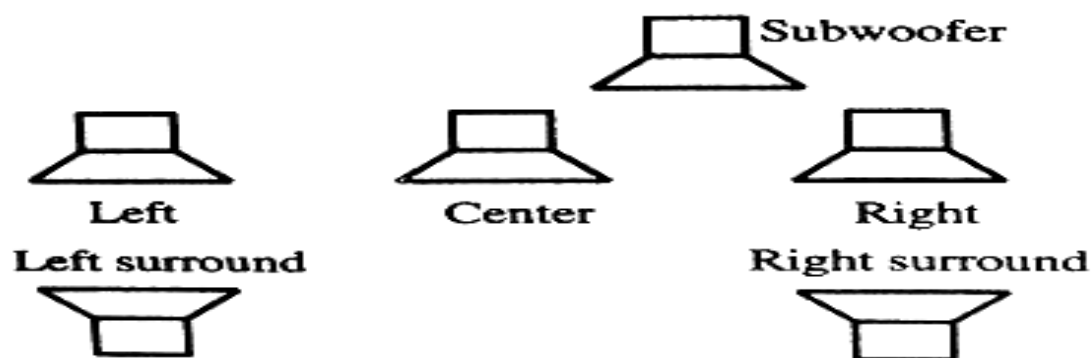
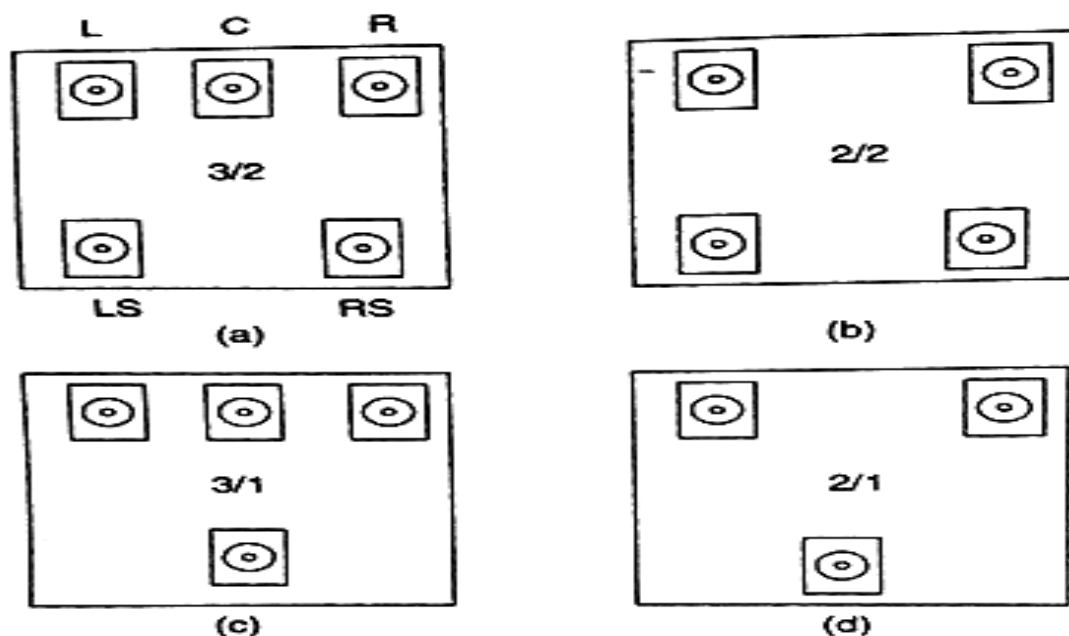
صدای چند کاناله MPEG-2 استخراج و دریافت کنند. که به این موضوع سازگاری با استاندارد پیشین گفته می شود مانند شکل زیر :



شکل ۲۳-۲ (استخراج و دریافت صدای چند گانه MPEG-2)

که نشان می دهد اگر 6 کانال کد شده در جریان MPEG-2 داشته باشیم هم می توانیم با یک مدار دی ماتریس و یک کد گشای حرفه ای هر 6 کانال را داشته باشیم و هم در روشی ساده تر با جمع و ترکیب این 6 سیگنال آن ها را به دو کانال ساده استریو تبدیل کنیم ، به گونه ای که تمام اطلاعات کانال های اولیه حفظ و به خروجی منتقل شود .

در شکل زیر الگوهای مختلف برای شرایط چند کاناله استاندارد MPEG و چیدمان بلند گوها و سیستم کامل (5.1) که شامل کانال اضافی LFE که در فرکانس پایین (15 تا 120Hz) می باشد که به آن گویند نشان داده شده است .



شکل ۲۴-۲ (الگوهای مختلف برای شرایط چندگانه استاندارد MPEG)

کدینگ به روش MPEG-2 AAC(Advanced Audio Coding)

به عنوان یک روش پیشرفته صدا می باشد . در ابتدا برای پخش رادیوی دیجیتال به کار می رفت ؛ بازدهی فشرده سازی را نسبت به روش های قبل افزایش داده مثلاً سیستم AAC با نرخ 96kbit/s کیفیت معادل کدینگ MPEG لایه 2 با نرخ بیت 192Kbit/s را می دهد ؛ یعنی 2 برابر فشردگی ولی کیفیت ثابت . سیستم MPEG2 AAC قابلیت پشتیبانی 48 کانال صوتی را دارد و نکته مهم اینکه این استاندارد سیستم های کدینگ پیش از خود (MPEG لایه 1 و 2) را پشتیبانی نمی کند و سازگار نیست .

AAC دارای سه نوع پروفایل است :

1 – پروفایل اصلی

2 – پروفایل با پیچیدگی پایین

3 - پروفایل با نرخ نمونه برداری مقیاس پذیر .

پروفایل اصلی پیچیده ترین وضعیت ساختاری سیستم را دارد ، که از تمام ابزار کدینگ استفاده می شود . پروفایل با پیچیدگی پایین به منظور کاهش حجم پردازش و میزان حافظه لازم ، برخی از ابزارهای اصلی را حذف می کند و توان ابزارهای دیگر را نیز محدود می کند . در پروفایل با نرخ نمونه برداری مقیاس پذیر ، سیگنال صدای ورودی به چهار طیف فرکانسی مساوی تقسیم می شود که هر کدام مولد جریان مستقلی می شوند .

مالتی پلکس کردن (Multiplexing)

مالتی پلکس کننده های دیجیتال (MUX) وظیفه ادغام و میان گذاری (interleaving) داده های گرفته شده از منابع مختلف را به عهده دارند . که این سیگنال شامل بیت ها یا بایت هایی از منابع مختلف است که لا به لای یکدیگر چیده شده اند ، تا در مجموع یک جریان واحد را به وجود آورند .

به طور کلی مالتی پلکس کننده باید وظایف زیر را انجام دهد :

1 (ساختن یک قالب اطلاعاتی با حداقل یک بیت از هر منبع ورودی به عنوان کوچکترین واحد

2 (تخصیص شکاف های زمانی ویژه به هر ورودی درون این قالب ها

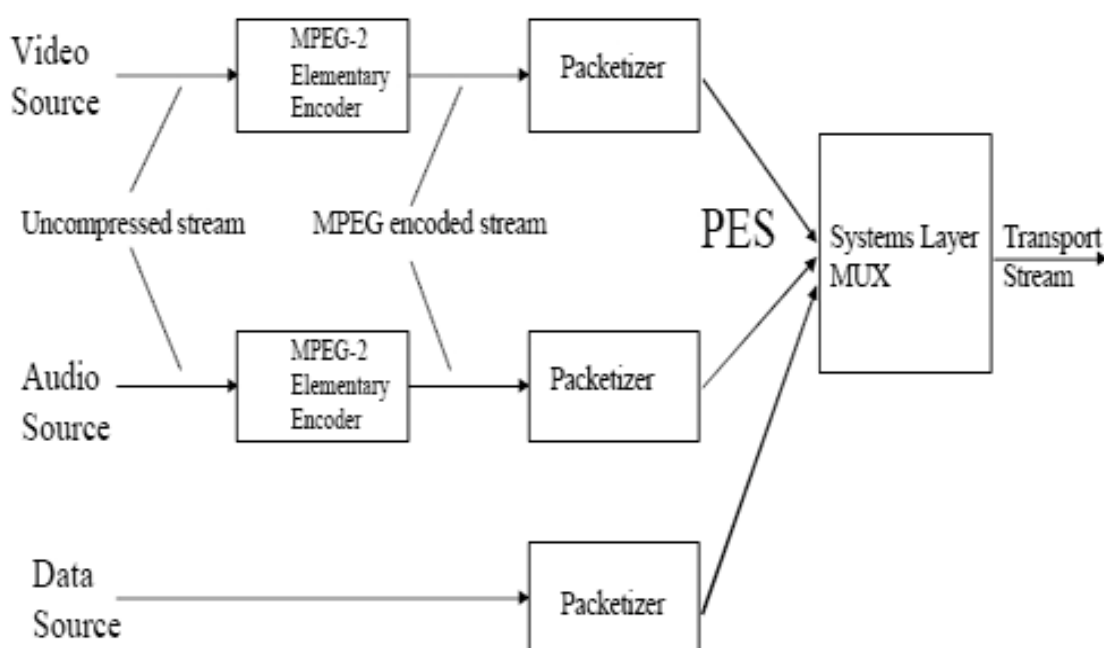
3 (اضافه کردن بیت های کنترل برای شناسایی و همزمان سازی قالب ها

4 (باید از چگونگی تغییرات نرخ اطلاعات در ورودی تعریف دقیق داشته باشد

در شیوه MPEG از روش سوئیچ بسته داده ها (Data Packet Switch) استفاده می شود؛ به صورتی که پیغام به بلوک هایی با طول کوچکتر تقسیم می شوند که بسته (Packet) نام دارند . هر بسته دارای یک (Header) سرآیند است . که شامل محتوای لازم برای مسیر یابی و بیت های اضافی برای کنترل است .

در MPEG-2 مالتی پلکس کردن داده ها با ایده بسته ای کردن اطلاعات مختلف و فرستادن آن ها در یک شاهره می باشد . همانطور که می دانیم اطلاعات مختلف از منابع صوت ، تصویر و داده های جانبی به صورت های مختلفی ممکن است کد شده باشند و مجموعاً این ها جریان های مقدماتی

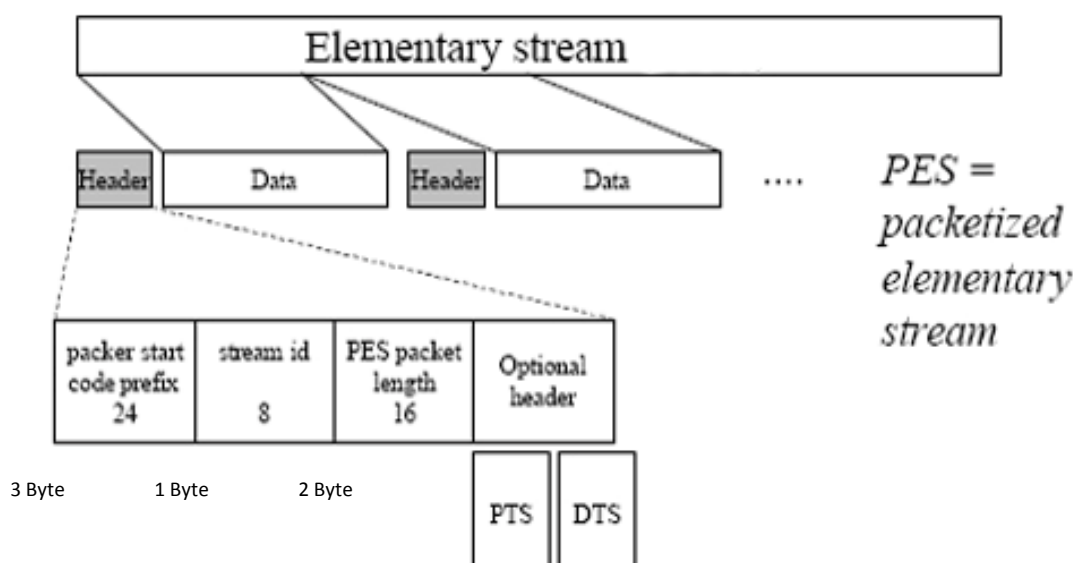
Elementary stream (ES) را تشکیل می دهند و بعد از آن این ها در بخشی از سیستم به نام بسته بندی کننده به صورت (PES) جریان های مقدماتی بسته بندی شده در می آیند . از آنجایی که ارسال و انتقال داده ها با طول متغیر و بلند بین سیستم های گوناگون پخش مناسب نیست . از این رو بسته های PES را به قطعات کوچکتر با طول ثابت تقسیم می کنند تا جریانی از بسته های مناسب برای انتقال به نام TS (transport stream) را بسازند . که روند آن به صورت زیر است :



شکل ۲۵-۲ (چگونگی بسته بندی به صورت PES برای انتقال)

تعریف PES (Packetized elementary stream)

بسته های اولیه صدا یا تصویر (PES) ساختاری با طول متغیر دارند ، و اطلاعات مفید آن ، داده هایی است که مرتباً از جریان های مقدماتی حاصل از منابع کد شده گرفته شده است . بسته های PES شامل header می باشند که معمولاً 6 بایت است .



شکل ۲۶-۲ (بسته بندی به صورت PES)

بعد از این 6 بایت ، بخشی به نام سرآیند اختیاری (Optional Header) ارسال می شود ، استفاده از آن جنبه اختیاری دارد و بستگی به مقتضیات ارسال دارد . که این بخش خود شامل تقسیمات کوچکتری است که مهمترین آن ها عبارتند از :

PTS و DTS (Decoding Time stamp) و (Presentation Time Stamp) که هر دو از 33 بیت تشکیل یافته اند .

PTS : وظیفه آن آگاه کردن دکدر، از زمان درست ورود داده های تصویر و صداست .

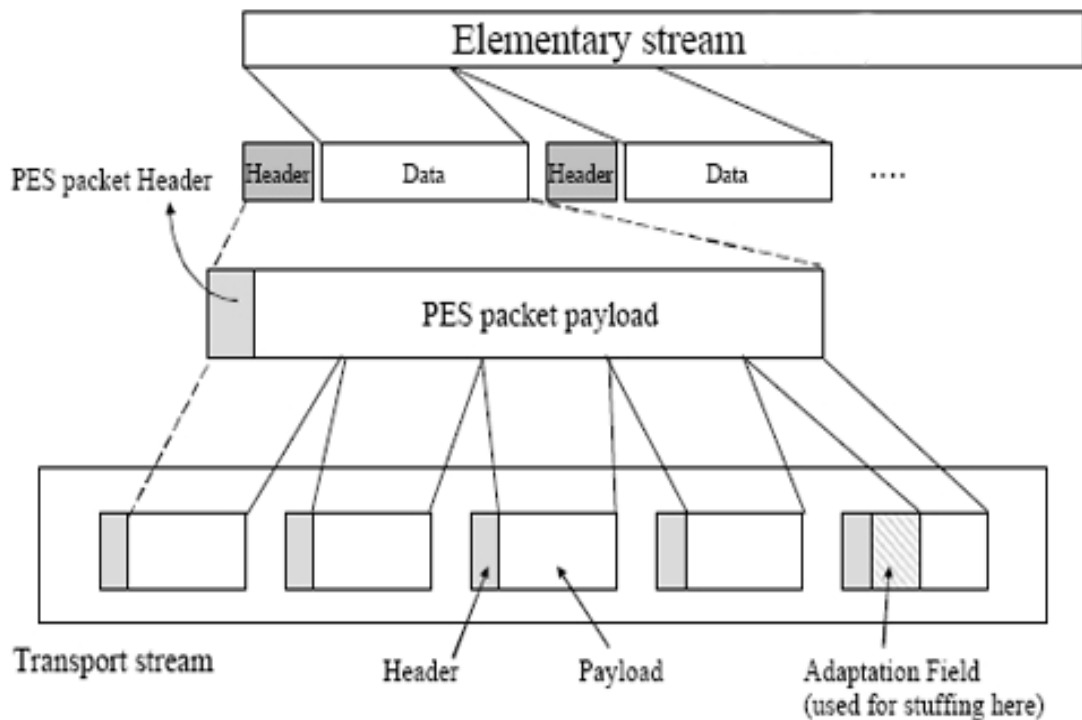
DTS : وظیفه آن اعلام زمان درست برای فرآیند کدگذاری اطلاعات است .

این نشانه ها برای هم زمانی صدا ، تصویر و سایر داده ها در گیرنده به کار برده می شود و معمولاً در فاصله های زمانی کمتر از 0.7 ثانیه برای گیرنده ارسال می شوند .

پس از این بخش ها اطلاعات اصلی و مفید ارسال می شوند که می توانند طول معادل 64 کیلوبایت را داشته باشند .

تعریف بسته های (TS) (Transport Stream)

بسته های TS حاصل ترکیب چند PES یا چند جریان برنامه (PS) (که به آن اشاره خواهیم کرد) می باشند. بسته هایی که در جریان انتقال (TS) قرار دارند، بسته های 188 بایتی می باشند که Header نیز شامل این 188 بایت می باشد. نحوه ایجاد TS از بسته های اولیه PES به صورت زیر است:



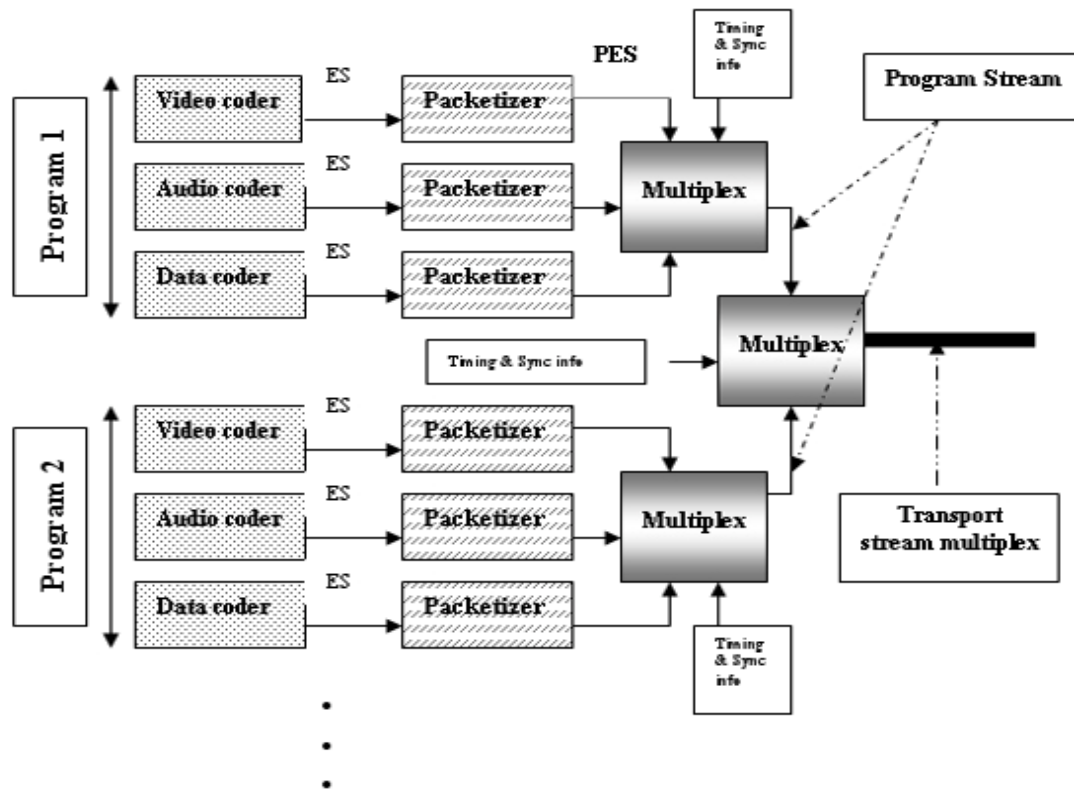
شکل ۲۷-۲ (نحوه ایجاد TS از بسته های اولیه PES)

محدودیت هایی که در شکل دادن و ساختن بسته های TS داریم به این صورت است:

- اولین بایت بسته های PES باید اولین بایت بسته های TS باشد.
- هر TS فقط شامل اطلاعاتی از یک PES می باشد.

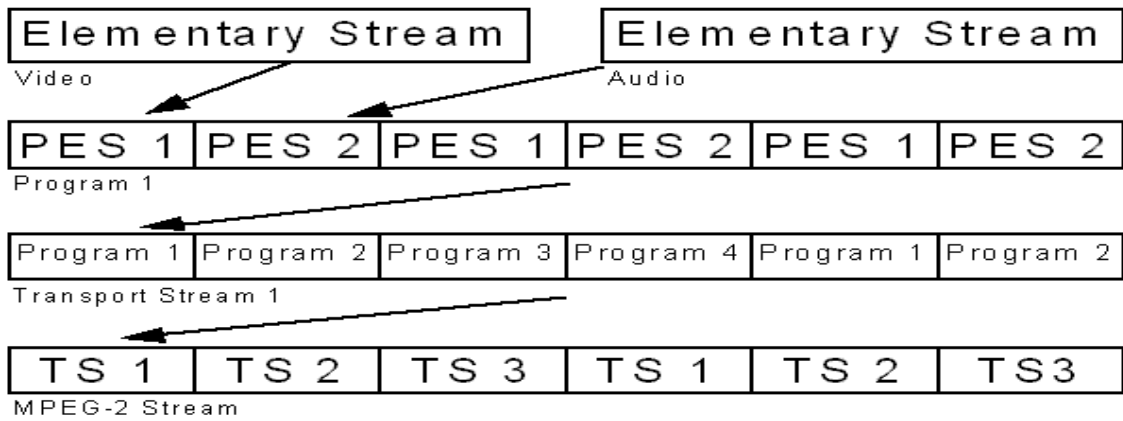
طرح دومی که برای فرآیند مالتی پلکس کردن استفاده می شود وجود PS (Program Stream) یا جریان داده های برنامه می باشد . هر PS از ترکیب چند PES که از منابع صوت و تصویر و دیتا ، که دارای مبنای زمانی یکسان می باشند به وجود می آید و دارای طول های متغیر هستند .

(مانند شکل)



شکل ۲۸-۲ (طرح دوم برای مالتی پلکس کردن استفاده از PS)

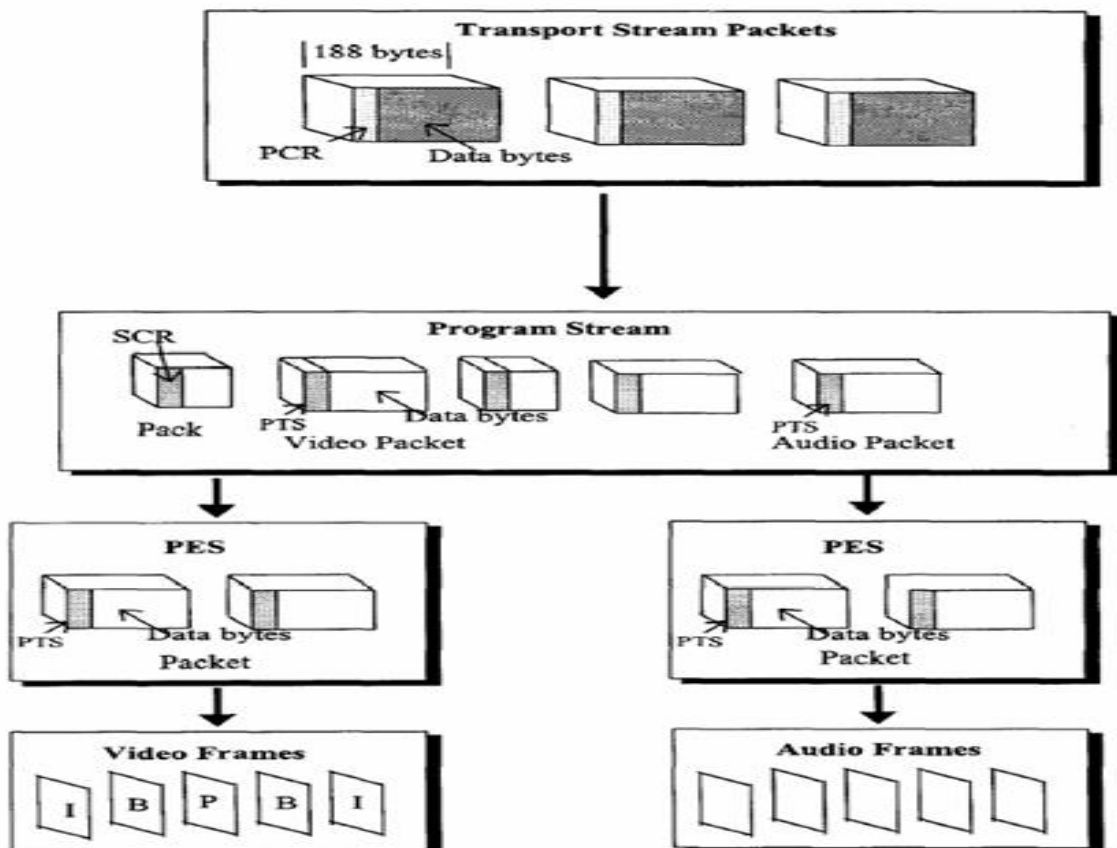
که سلسله مراتب آن به این صورت است که جریان های مقدماتی حاصل از یک منبع کد شده ES بسته بندی می شوند و PES را تشکیل می دهند ، سپس چند PES با مبنای زمانی یکسان یک PS را تشکیل می دهند . سپس PS های مختلف (گاهی حتی تا 20 برنامه) در یک مالتی پلکس کننده نهایی با هم در حوزه زمان تلفیق و به صورت مناسب لایه یکدیگر چیده می شوند و جریانی از بسته های ثابت درست می کنند که همان جریان انتقال یا TS نهایی MPEG-2 است .



شکل ۲-۲۸-۲ (Elementary stream فیلم و صدا)

نرخ ارسالی برای هر برنامه به میزان تقریبی 2 Mbit/s تا 6 می باشد که برای ویدئو ، صدا یا داده های دیگر ، بسته به نوع و محتوای برنامه می تواند ثابت یا متغیر باشد که البته معمولاً ثابت باقی می ماند ؛ که جریان داده ها (TS) حداکثر به نرخ داده ای 40 Mbit/s می رسد .

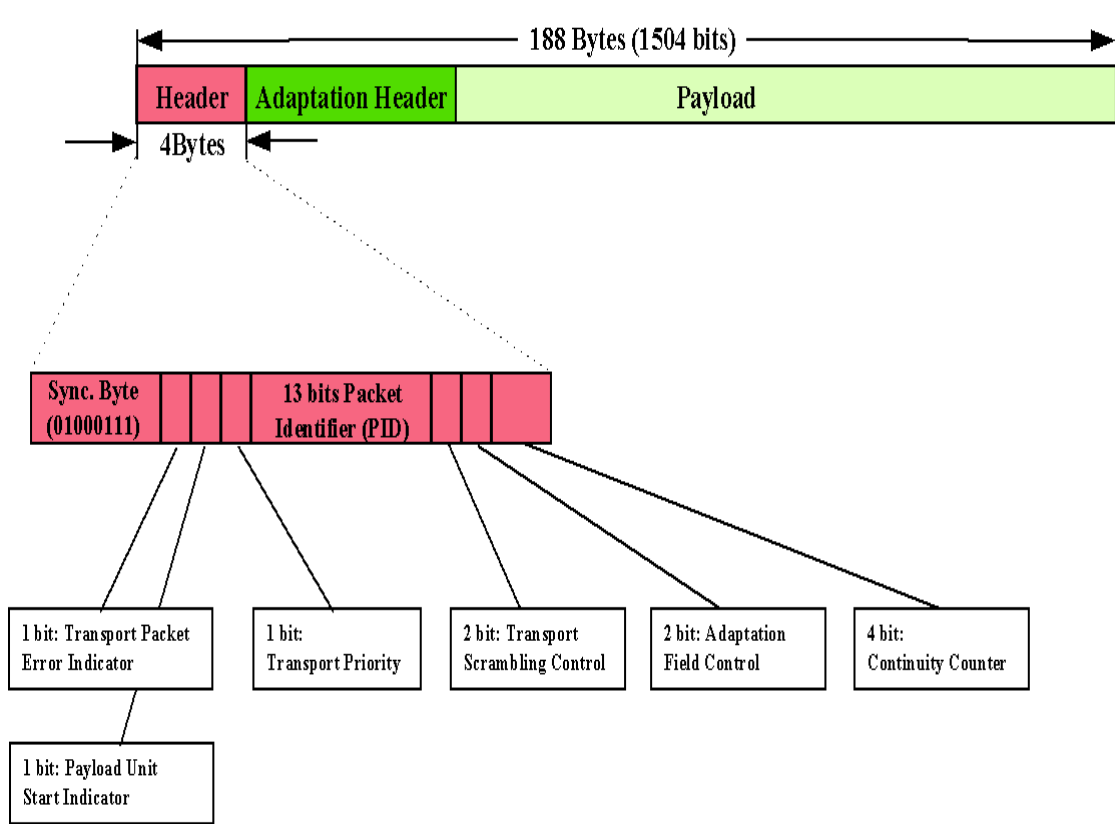
که در گیرنده عمل عکس آن به صورت زیر است :



شکل ۲-۲۹ (جریان انتقال پاکت های تصویر و صوت)

ساختار و اجزاء بسته های جریان داده های انتقالی (TS)

جریان های TS شامل بسته هایی از داده ها با طول ثابت به اندازه 188 بایت است که 4 بایت به Header تعلق دارد و 184 بایت بقیه متعلق به اطلاعات اصلی (Payload) می باشد ، که می تواند صدا ، تصویر یا داده را شامل شود . (مطابق شکل)



شکل ۳۰-۲ (ساختار و اجزا بسته های جریان داده های انتقالی)

- اولین بایت Header ، بایت سنکرون کننده می باشد . که دارای مقداری ثابت است و برابر است با (47 h) یا (01000111) باینری ؛ سنکرون سازی هر بسته از داده های TS توسط این بایت صورت می پذیرد ، که در دکدر برای همزمان سازی استفاده می شود

Melec.ir

- بخش مهم دیگر در Header یک بخش 13 بیتی تشخیص یا شناسایی محتوا و نوع داده هاست که PID(Packet Identification) نامیده می شود ، این بخش در حین مالتی پلکس کردن داده ها ساخته می شود . این بخش به جریان مقدماتی داده ها (ES) بستگی پیدا می کند و بر اساس آن تعریف می شود . PID در واقع معرف محتوای موجود در بخش اصلی اطلاعات است ؛ که عملکردی نظیر نشانی پستی را دارد . چون PID در مکان ثابتی است ، پس از همزمان سازی در دکدر ، جداسازی و استخراج بسته های گوناگون به سادگی صورت می پذیرد .
- بیتی که بلافاصله بعد از بایت سنکرون قرار می گیرد به عنوان بیت "شاخص خطای انتقال " Transport Error Indicator شناخته می شود . چنانچه در فرآیند انتقال احتمال بروز خطا به حدی افزایش یابد که امکان تصحیح خطا در گیرنده موجود نباشد ، با این بیت که در پایان مسیر ارسال می شود و در گیرنده فعال می شود و بسته های TS به عنوان حامل خطا مشخص و نشانه دار می شوند .
- بخش بعدی ، بیت "شاخص آغاز بخش اطلاعات اصلی " Payload Unit start Indicator است . این بیت ورود یک بخش جدید اطلاعات را اعلام می کند و این بسته جدید TS ، خود می تواند شامل بسته داده های اولیه TS به همراه Header بسته باشد و یا شروع یک جدول PID به عنوان اولین بایت بخش اطلاعات اصلی و سپس سایر داده هاست .
- بخش بعدی مربوط به بیت اولویت ارسال (Transport Priority) است . و بیانگر اولویت و اهمیت بیشتر برای یک بسته TS نسبت به سایر بسته ها با PID مشابه است .

- Transport Scrambling Control: بخش کنترل جریان در هم ریخته یک بخش دو

بیتی می باشد. این دو بیت نشان می دهند که آیا بخش اصلی اطلاعات از قبل در هم ریخته

شده اند یا نه! اگر هر دو بیت صفر باشند یعنی در هم ریخته نشده اند. اما اگر هر کدام از

بیت ها غیر صفر باشد یعنی داده ها به صورت در هم ریخته ارسال شده اند. در این

صورت یک جدول CAT(Conditional Access Table) یا دسترسی مشروط برای

بازگرداندن داده ها به وضعیت اول (غیر در هم ریخته) مورد نیاز است.

- Adaptation Field Control بیت های کنترل ناحیه تطبیق ، این دو بیت نشان می

دهند که آیا بخش Header همان 4 بایت است یا به صورت بخش تطبیق Adaptation

Field گسترش یافته است. اگر هر دو بیت صفر باشد یعنی چنین ناحیه ای موجود نیست ،

اما اگر این ناحیه موجود باشد طول header بیشتر و در عوض طول بخش اصلی کمتر

می شود که در کل طول TS همان 188 بایت می ماند .

- Continuity Counter یا شمارنده پیوستگی . آشکار سازی خطا در دریافت بسته در

دکدر توسط این بخش صورت می گیرد . یک بخش 4 بیتی برای هر کدام از بسته های TS

با PID مشابه دارای شمارنده هستند . در سمت فرستنده در جریان بسته های TS ، بسته

هایی که PID یکسان دارند از صفر تا 15 شمرده شده و در پایان دوباره از صفر شروع

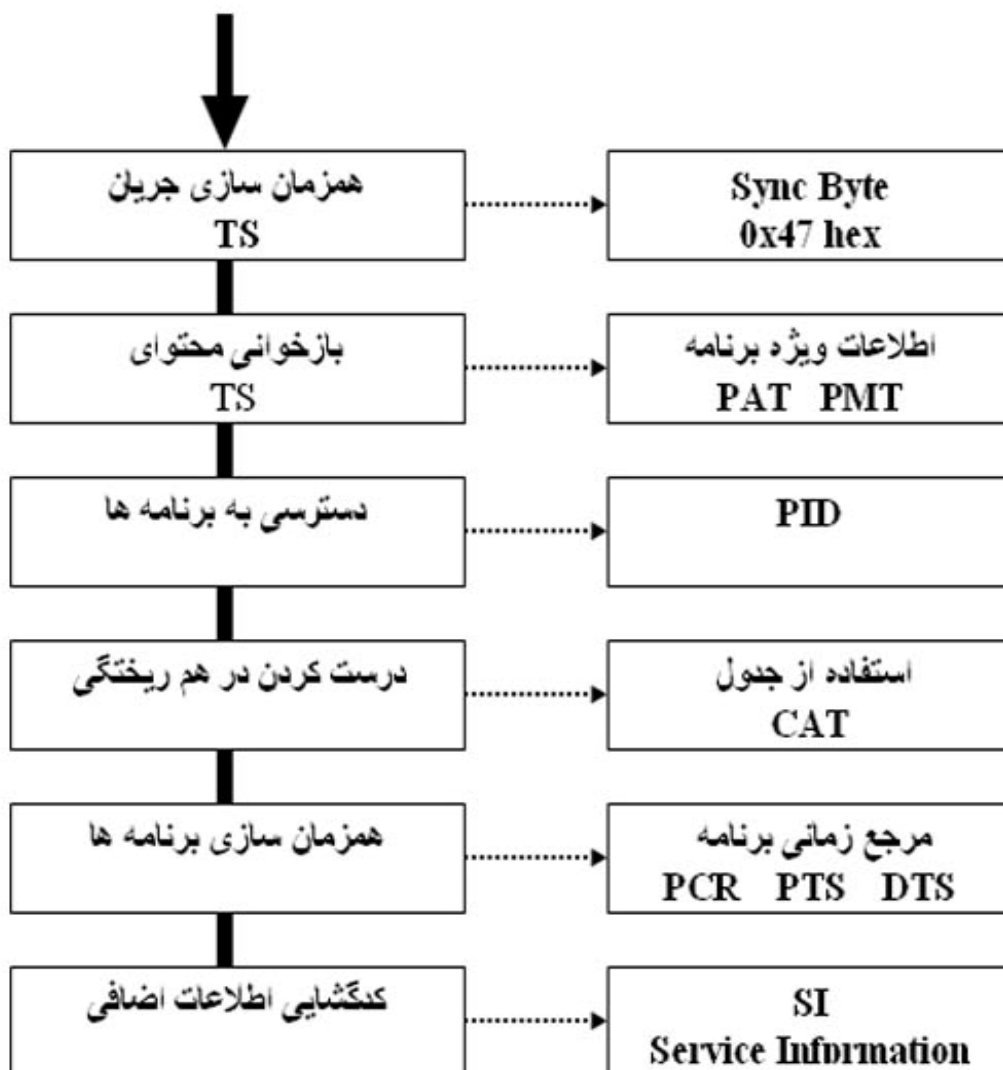
می شود . اگر در سمت گیرنده دریافت بسته ها با عدم پیوستگی همراه باشد ، نظم اعداد به

هم خورده و این به معنای بروز خطا بوده و نشان می دهد بسته داده ها در جریان انتقال گم

شده است .

اطلاعات ضروری برای گیرنده

اطلاعاتی هستند که گیرنده یا کدگشای MPEG به آن ها نیاز دارد تا بتواند اطلاعات یک برنامه مشخص را از مجموعه بزرگ جریان داده های دریافتی TS اخذ و از بقیه اطلاعات جداسازی کند که این اطلاعات به صورت زیر هستند :



شکل ۳۱-۲ (جریان داده های دریافتی TS)

2-3 سنکرون (Synchronization)

همزمان سازی جریان TS :

در مرحله اول ورود جریان های TS در گیرنده ، کدگشا باید عملیات پردازش خود را با جریان TS ورودی و داده های بسته بندی شده همگام و هماهنگ کند . بنابراین با شروع دریافت TS کدگشا بایت همزمانی را در بین رشته بیت دریافتی جستجو می کند . دکدر می داند که محتوای بایت همزمانی مقدار ثابت 47h و باید 188 بایت قبل و بعد از آن را نیز جستجو کند . اگر این بایت پیدا شد و 188 بایت بعدی نیز همین بود به عنوان بایت همزمانی شناخته می شود . در غیر این صورت به عنوان بایت عادی می باشد . در استاندارد MPEG همزمانی پس از 5 بسته 188 تایی TS صورت می گیرد یعنی 5 بایت 47h با فواصل 188 بایت از یکدیگر باید تشخیص داده شود تا همزمانی صورت گیرد و قطع آن در صورت فقدان 3 بسته TS در دکدر می باشد .

بازخوانی محتوای TS :

چون برنامه های فرستاده شده در جریان انتقال TS ساختار متغیری دارند، باید یک سری اطلاعات ویژه برای اینکه وضعیت ساختاری برنامه ها برای کدگشا معلوم شود همراه TS فرستاده شوند ؛ این اطلاعات به صورت جدول هایی بوده که همان قالب بسته های TS را دارند ، و این فهرست ها در بخش اصلی اطلاعات (Payload) بسته ها جای دارند . به این فهرست ها PSI (Program Specific Information) گویند .

اولین جدول PSI جدول PAT (Program Associate Table) یا جدول وابسته به برنامه می باشد . این جدول معمولاً در جریان TS تکرار می شود و معرف فهرست و تعداد برنامه های موجود در TS است . دوره تناوب تکرار این جدول بین 25ms تا 500ms می باشد .

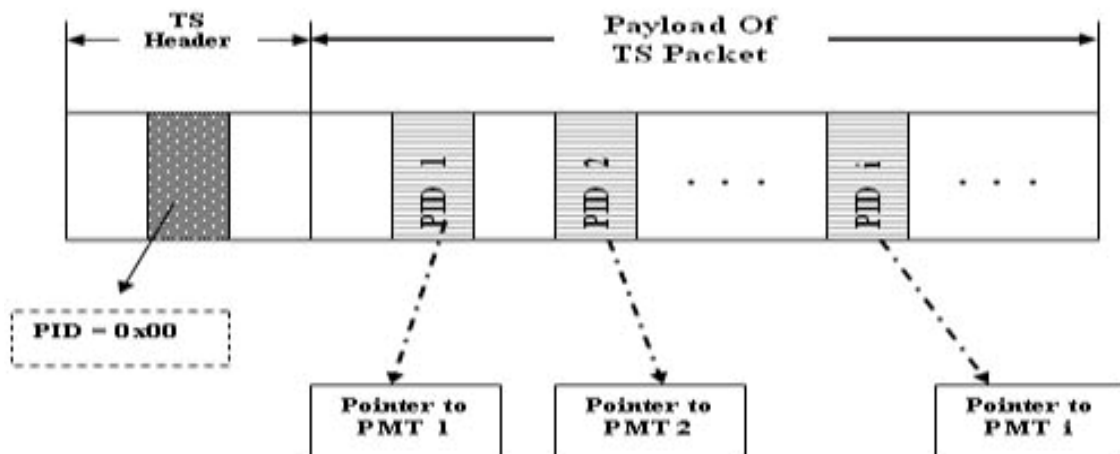
PID بسته های TS که جدول PAT را در خود دارند صفر می باشد و به همین دلیل به سادگی

شناسایی می شوند .

در بخش اطلاعات اصلی این بسته ها PID های ویژه ای ارسال می شوند ، که این PID ها نقش

نشانگر برنامه را ایفا می کنند (مانند شکل زیر) .

PAT (Program Associate Table)



برای هر برنامه یک PID داریم

شکل ۳۲-۲ (اجزا تشکیل دهنده PAT)

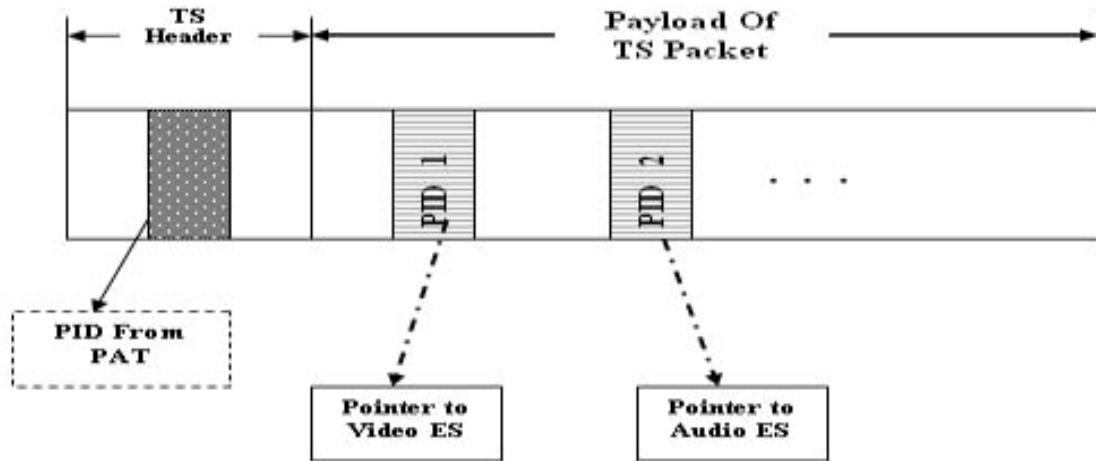
PMT (Program Map Table) یا جدول ترسیم برنامه ، این جدول ارائه کننده مشخصات

برنامه هاست و شامل نام هر برنامه ، وجود مجوز تکثیر ، ارجاع به بسته های ویژه ای که معرف

زمان برنامه هستند ، فهرستی از نام تولید کنندگان برنامه ها و فهرستی از اجزاء ساختاری برنامه

(صدا ، تصویر و داده های جانبی) می شود (مانند شکل) .

PMT (Program Map Table)



برای هر جریان مقدماتی ES یک PID داریم

شکل ۳۳-۲ (اجزا تشکیل دهنده PMT)

بنابراین اطلاعات دقیق درباره نوع بسته های TS در جدول PMT تعبیه شده ، که حداکثر زمان تکرار آن 500ms است ؛ باید این اطلاعات به طور متناوب تکرار شوند تا در صورتی که بیننده اقدام به تعویض کانال می کند ، محتوای صدا و تصویر برنامه به موقع فراهم شود .

دسترسی به برنامه ها :

پس از اینکه کاربر برنامه دلخواه خود را انتخاب کرد ، PID های وابسته به جریان مقدماتی داده ها که اکنون در TS قرار دارد و بر اساس اطلاعات موجود در PAT و PMT مشخص شده اند . به دو صورت قابل شناسایی هستند :

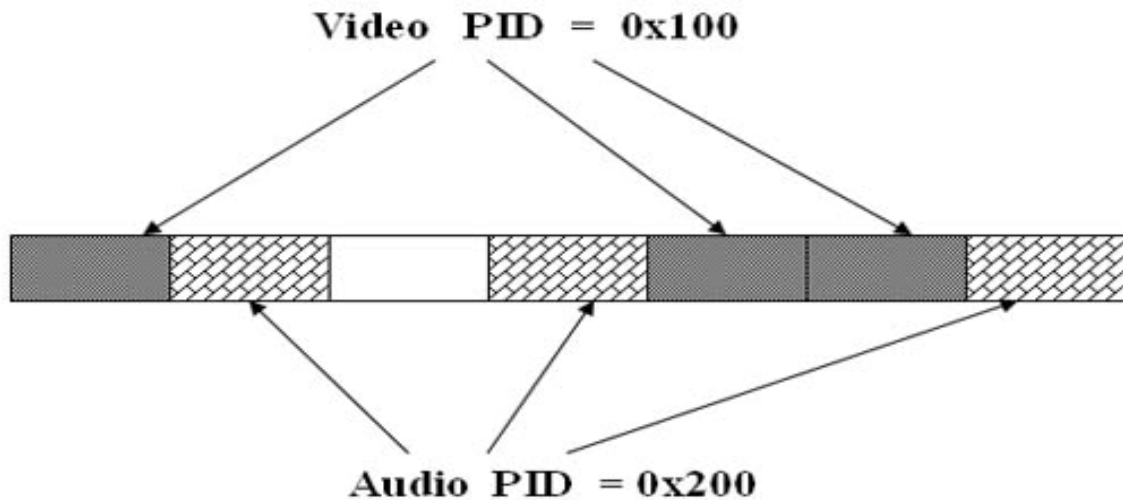
1- PID های مربوط به سیگنال تصویر

2- PID های سیگنال صدا

مثلاً اگر فرض کنیم PID مربوط به صدا 200hex و PID ویدئو 100hex باشد ، در فرآیند دی مالتی پلکس کردن، تمام بسته های TS که PID 100hex دارند به یکدیگر می

Melec.ir

پیوندند و PES ویدئو را بازسازی می کنند ، و وارد دکدر تصویر می شوند . همین عمل برای TS های با PID 200hex رخ می دهد که PES صدا را بازسازی می کنند (مانند شکل).



شکل ۳۴-۲ (بازسازی PES صدا و ویدئو)

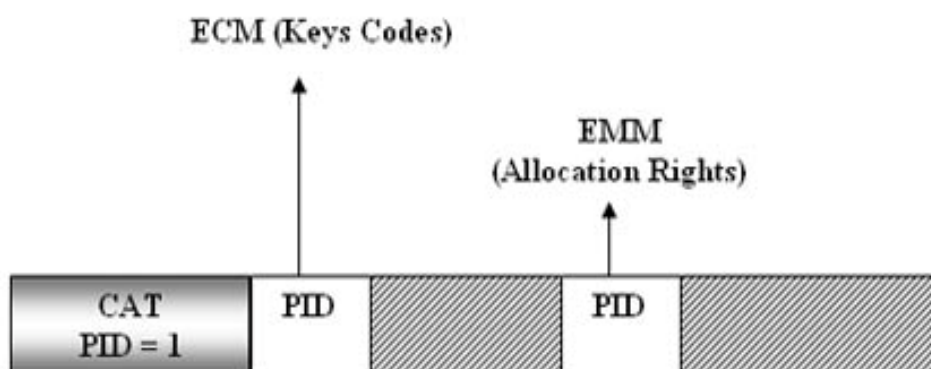
بنابراین اگر داده های جریان اولیه از ابتدا در حالت طبیعی بوده و درهم ریخته (Scrambled) نشده باشد ، مستقیماً و به سادگی استخراج می شوند .

دسترسی به برنامه های درهم ریخته :

به دلایل گوناگون گاهی تمام یا برخی از جریان های مقدماتی با روش های کدینگ ویژه محافظت شده و از دسترس عموم خارج می شوند . در هم ریختن اطلاعات به روش های متفاوتی انجام می شود که پس از اجرای آن گیرنده قادر به دریافت و آشکارسازی برنامه نیست ، مگر با وجود امکانات سخت افزاری و نرم افزاری لازم و داشتن مجوز و اعتبار قانونی برای این کار . این سخت افزار خاص نیازمند داده های لازم برای تأیید اعتبار و اجرای روش معکوس در هم ریختن به منظور آشکارسازی اطلاعات اولیه است . به همین دلیل در جریان TS یک جدول ویژه به نام

جدول دسترسی مشروط یا CAT(Conditional Access Table) برای گیرنده ارسال می شود که PID آن 1 می باشد . (PID = 1)

بسته CAT شامل PID هایی برای بسته هایی از TS هستند که در آن ها اطلاعات لازم برای معکوس کردن و در هم ریختن ارسال می شود . که این اطلاعات عبارتند از ECM (Entitlement Control Message) یا پیغام کنترل حق استفاده و دیگری (Entitlement Management Message) EMM یا پیغام مدیریت حق استفاده می باشد . داده های ECM برای ارسال کد های در هم ریخته و داده های EMM برای هدایت کاربر استفاده می شود . (مانند شکل)

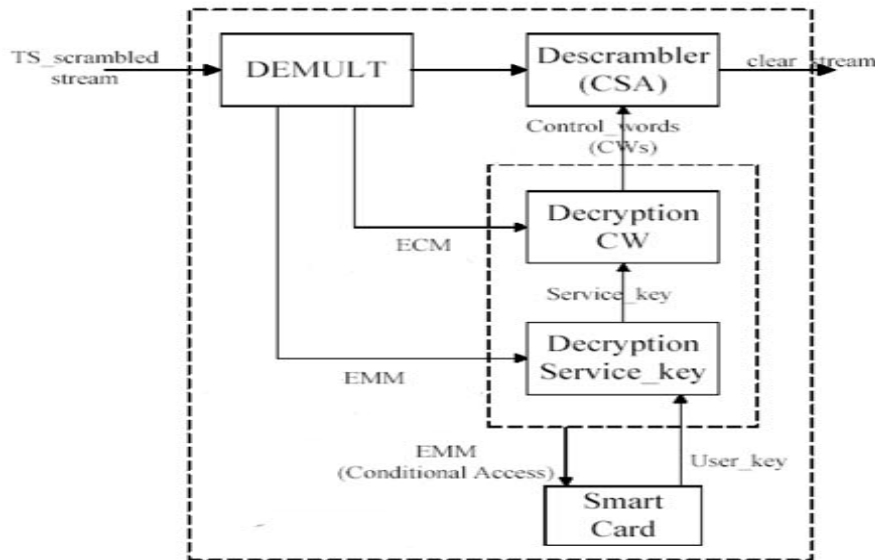


شکل ۳۵-۲ (اجزاء تشکیل دهنده پیغام مدیریت حق EMM)

نکته مهم اینکه فقط جریان اولیه داده ها می تواند در هم ریزد و این در هم ریختگی شامل بخش آغازین بسته ها یا مثلاً جدول ها نمی شود و در هم ریختگی بخش Header و ناحیه تطبیق مجاز نمی باشد .

خلاصه عمل عکس در هم ریختن يك صورت شکل می باشد ، که در آن ابتدا جریان در هم ریخته قبل از ورود به دکر MPEG وارد بخش Descrambler می شود . در اینجا با

اطلاعات EMM و ECM که از هم جدا شده اند ، قفل روی جریان TS برداشته می شود یا همان عمل عکس در هم ریختگی انجام می شود ؛ و خروجی آن جریان درست شده می باشد که وارد دکدر MPEG می شود .



شکل ۳۶-۲ (رمز گشایی قفل های TS)

همزمان سازی برنامه ها

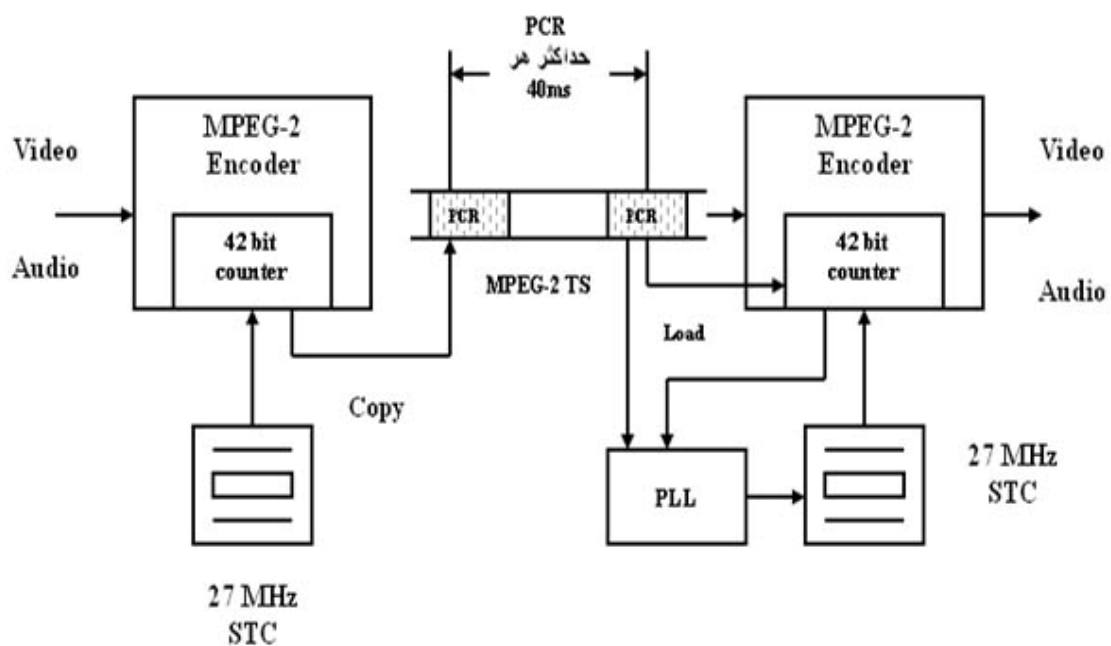
از آنجایی که در سیستم تلویزیون دیجیتال مقدار داده های تولید شده برای تصاویر کاملاً متغیر هستند، امکان استخراج اطلاعات زمانی از هنگام شروع تصاویر به طور مستقیم و پیوسته امکان پذیر نیست. با افزودن سیگنال های مرجع زمانی در مالتی پلکس کننده ، کاری می کنیم که ویدئو و صدای همراه آن در گیرنده همزمان و سنکرون نگه داشته شوند . به عبارت دیگر بسته های به خصوصی دارای مهر و نشان زمانی هستند و کدگشا موظف است ، تا این بسته ها را در زمان های مناسب با دقت زمانی کافی کد گشایی کند .

در قدم اول باید ساعت زمانی گیرنده و فرستنده همزمان و همگام گردند . در کد کننده

MPEG-2 معمولاً فرکانس مرجع برای تمامی پردازش ها 27Mhz است . که این نوسان ساز

27Mhz تأمین کننده "ساعت زمانی سیستم" یا STC(System Time Clock) که در واقع یک شمارنده 42 بیتی است که مرجع خود را از نوسان ساز 27Mhz می گیرد . در سمت گیرنده یک سیستم STC مشابه (شامل نوسان ساز 27Mhz و شمارنده) نیز وجود دارد .

کد کننده MPEG-2 اطلاعات زمانی را با نام "مرجع ساعت برنامه" یا PCR(Program Clock Reference) به سمت گیرنده ارسال می کند ، که این اطلاعات در واقع یک نسخه از آخرین وضعیت شمارنده در فرستنده هستند ؛ که در زمان های مناسب درون جریان داده ها تزریق می شوند . (مانند شکل)



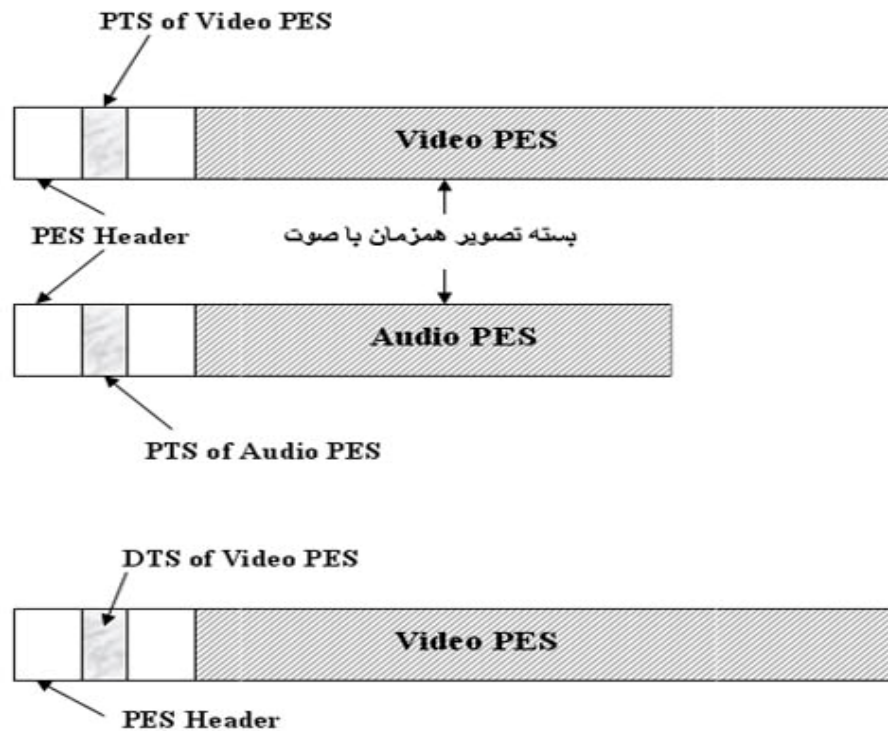
شکل ۳۷-۲ (مرجع زمان بندی ورود اطلاعات PCR(Program Clock Reference))

به این ترتیب جریان داده ها یک "ساعت درونی" خواهد داشت که تمامی پردازش های کدینگ به وسیله این مرجع انجام می شود که دگر MPEG-2 اطلاعات این ساعت را استخراج و با

سیستم زمانی خویش مقایسه می کند ، پس از این همزمان سازی مراحل کدگذاری صدا و تصویر قابل اجرا می باشد .

اگر PCR دریافتی با ساعت درونی دکدر همگام باشد آن گاه مرجع فرکانسی 27MHz درون گیرنده با فرستنده همگام و منطبق است . اما اگر انحراف و اختلافی وجود داشته باشد ، با یک سیستم قفل فاز نوسان ساز در گیرنده اصلاح می شود . در MPEG-2 حداکثر فاصله زمانی در ارسال مقادیر PCR 40ms است و حداکثر خطای جیتر (لرزش) قابل قبول 500ns است . در قدم بعدی برای همزمان نگه داشتن تصویر و صدای همراه آن ، اطلاعاتی را در گیرنده از سیستم STC اخذ می کنیم که این اطلاعات از ابتدا درون PES تزیق شده اند . در ابتدا 33 بیت از توالی بیت های با ارزش در STC جداگانه ذخیره می شود و در فواصل حداکثر 700ms در Header ، PES قرار می گیرند این اطلاعات همان طور که در بحث PES گفتیم "تشانگر های زمان نمایش " یا PTS می باشند ؛ که معرف زمان آغاز نمایش یک تصویر ویدئویی یا نواختن یک قطعه صوتی است . به این ترتیب با استفاده از PTS دکدر می داند در چه زمانی باید عمل ارائه و نمایش (زمان ورود) تصویر یا صدا را انجام دهد .

قبلاً درباره قاب های تصویری تولید شده در MPEG (گروه تصاویر یا GOP) صحبت کردیم که این قاب ها ، با قاب های تولید شده منبع تصویر متفاوت هستند . بر همین اساس برای دسترسی به این قاب ها ، به DTS یا "نشان های زمانی برای کد گذاری " نیاز داریم تا زمان درست کدگذاری را اعلام کنند . داده های DTS نیز درون بخش Header ، PES می باشند و 33 بیت نیز دارند . (مانند شکل)



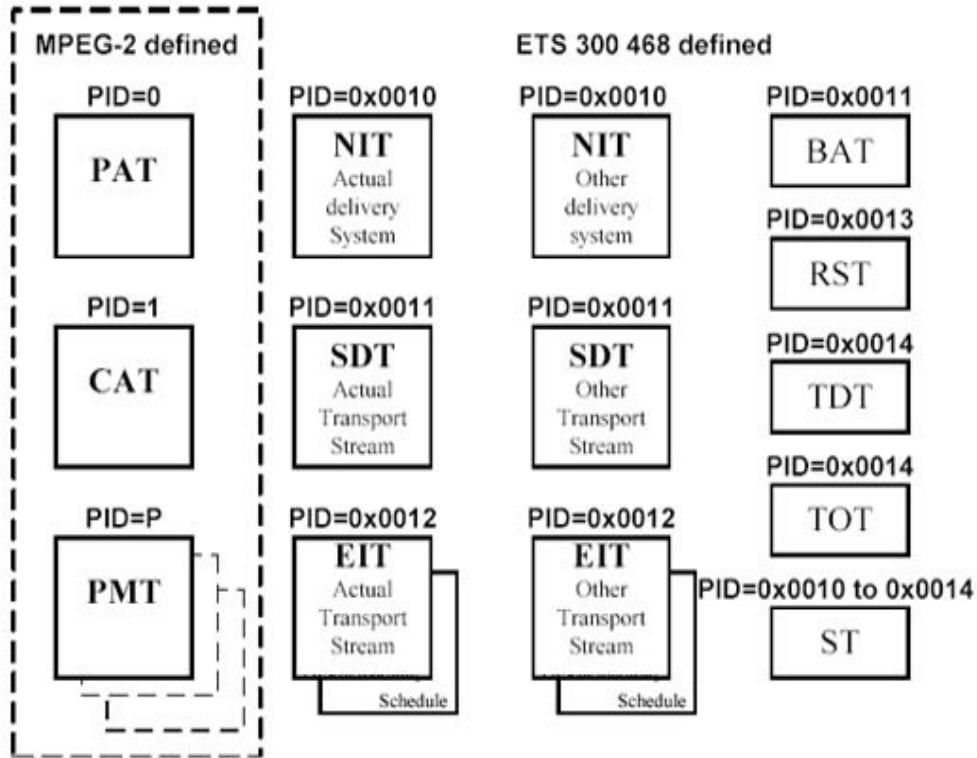
شکل ۳۸-۲ (بسته بندی همزمان تصویر و صوت)

کدگذاری اطلاعات اضافی

قبلاً جداولی تحت عنوان MPEG-2 PSI مطرح شد که شامل PAT ، PMT و CAT بود، که جداول اولیه بودند . هر دو سیستم پخش اروپایی DVB و آمریکایی ATSC ، اطلاعات ویژه اضافی را همراه اطلاعات اصلی صدا و تصویر هر برنامه ، برای گیرنده ارسال می کنند ، که نقش این اطلاعات ویژه ساده کردن عملکرد دستگاه های گیرنده و راحتی کاربر در استفاده از آن است . این اطلاعات به صورت جداول اضافی فرستاده می شوند . گروه DVB در استاندارد ETS 300 468 جدول هایی را معرفی کرده ، که اصطلاحاً به آنها اطلاعات خدماتی SI (Service Information) گفته می شود و به کاربر در انتخاب سرویس های موجود کمک می کند ؛ و همچنین شرایطی برای تطبیق گیرنده با داده های دریافتی فراهم می کند .

این جدول ها و PID های مخصوص آن ها به صورت زیر است :

Other tables in ETS 300 468



PAT Program Associate Table	MPEG-2
PMT Program Map Table	PSI
NIT Network Information Table	
SDT Service Descriptor Table	DVB SI
BAT Bouquet Associate Table	Service Information
EIT Event Information Table	
RST Running Status Table	

شکل ۳۹-۲ (جدول ها و PID های مخصوص (SI (Service Information))

Melec.ir

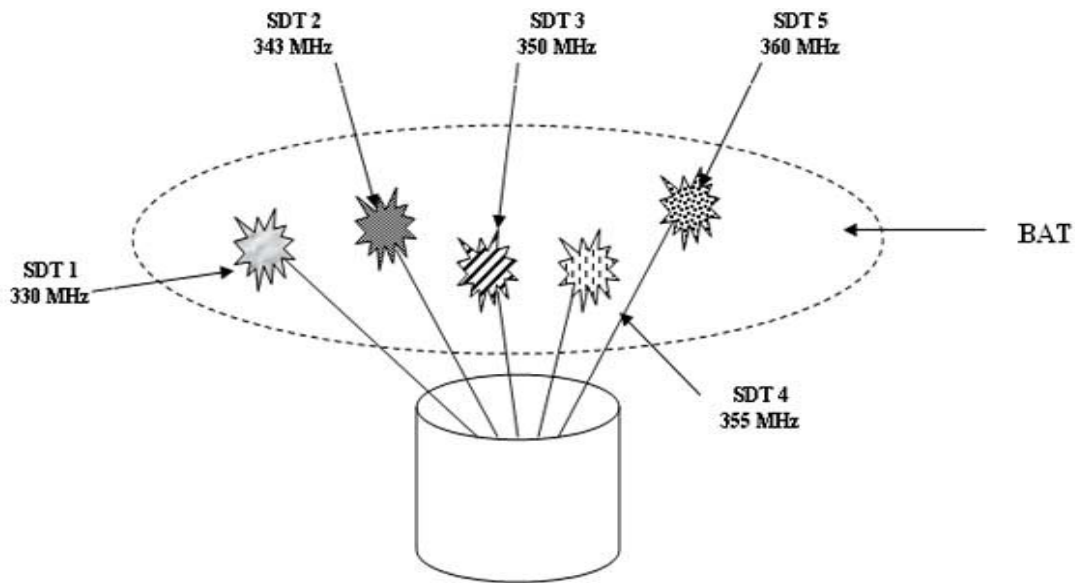
این جداول انواع متفاوتی دارند که البته همه آن ها جنبه ارسال اجباری ندارند ، برخی از آن ها

به شرح زیر هستند :

- NIT : جدول اطلاعات شبکه ، بیانگر پارامتر های فیزیکی کانال انتقال سیگنال DVB است .
این بخش شامل ، فرکانس سیگنال دریافتی و نوع سیستم ارسال (ماهواره ، کابل یا پخش زمینی) و همچنین برخی اطلاعات فنی و مشخصات ارسال است (نظیر چگونگی حفاظت در برابر خطا ، نوع مدولاسیون و ...) دستگاه های گیرنده قادر به ذخیره سازی این پارامتر های فیزیکی کانال هستند ؛ مهمترین نکته این که دستگاه های گیرنده در صورت منطبق نبودن اطلاعات NIT با پارامتر های واقعی ارسال ، رفتار ویژه ای را پیش می گیرند . مثلاً در صورت یکسان نبودن فرکانس دریافتی با فرکانس درون جدول NIT از بازسازی صدا و تصویر امتناع می کنند .

- SDT : جدول توصیف گر خدمات که شامل جزئیات توصیفی بیشتری از برنامه موجود در TS است ، نظیر نام برنامه مانند BBC و ... اطلاعات این جدول به موازات داده های PID وارد بخش PAT می شوند و اطلاعات متنی برای کاربر فراهم می کنند .

- BAT : جدولی اطلاعاتی نزدیک به SDT است و جدول مشتمل بر دسته ای از برنامه ها می باشد . که یک معرفی جامع از سرویس های موجود در یک گروه برنامه که در چند کانال مختلف (با فرکانس های مختلف) وجود دارند است . مانند مثال شکل زیر که BAT را به یک گلدان تشبیه کرده است .



شکل ۴۰-۲ (تشبیه کردن BAT به یک گلدان)

- EIT : جدول اطلاعات رویداد ها ، شامل جدول زمان بندی شروع و پایان رویداد ها با برنامه های پخش شده در طول روز یا هفته است . و تولید کنندگان برنامه آن را به عنوان " راهنمای الکترونیکی برنامه " EPG ارسال می کنند . متأسفانه تمامی گیرنده ها از این جدول پشتیبانی نمی کنند .
- RST : جدول وضعیت اجرا ؛ اگر یک دستگاه ضبط برنامه در موقعیت گیرنده داشته باشیم ، برای کنترل درست زمان شروع و پایان ضبط از این جدول استفاده می شود .
- TDT : جدول زمان و تاریخ ؛ که بر اساس آن تاریخ و ساعت جاری فرستاده می شود ، که جدول زمان مرجع گرینویچ GMT است .
- TOT : جدول تصحیح زمانی ؛ اگر اختلاف زمانی در نواحی مختلف جغرافیایی برای گیرنده مهم باشد ، آنگاه تصحیح زمانی با این جدول انجام می شود .

- ST : جدول انباشت ؛ با این جدول در نقاطی که TS مجدداً مالتی پلکسی می شود ، ممکن است لازم شود برخی از اطلاعات خاص از جریان دریافتی حذف شوند که با استفاده از این جدول ، این کار را انجام می دهند ، و روی اطلاعات مورد نظر دوباره نویسی می کنند .

MPEG-1

MPEG-1 اولین نسخه MPEG برای فشرده سازی صوت و تصویر می باشد و در سال 1993 به وجود آمد . بسیاری از اصول فشرده سازی تصویر MPEG-1 از روی فرمت JPEG گرفته شده است ، که در Video CD ها و CD های صوتی کاربرد دارد . با استفاده از این فرمت CD های معمولی 650 مگابایتی توانستند قابلیت ذخیره سازی ویدئو را داشته باشند . تصاویر MPEG-1 برای 25 Hz (25 فریم در ثانیه) برابر 352 x 288 و برای 30 Hz (30 فریم در ثانیه) اندازه ای برابر 352 x 240 پیکسل می باشد . MPEG-1 چون از نرخ بیت پایین پشتیبانی می کند ، کیفیت فیلمهای VCD در حد مطلوب نبوده و حتی به زحمت به فیلمهای VHS رقابت می کند . از طرف دیگر به دلیل آن که MPEG-1 فقط از یک سیگنال ویدئو پشتیبانی می کند ، این فرمت برای استفاده در ارسال فرکانس های ماهواره ای مناسب نیست ، چرا که در ارسال دیجیتالی امواج ماهواره ای هدف ارسال چندین برنامه ویدئویی گوناگون بر روی یک ترانسپاندر با یک فرکانس خاص می باشد . همگام با این فرمت ، فرمت H.261 توسط ITU ساخته شد که شباهت زیادی با MPEG-1 دارد .

MPEG-2

به دنبال پیشرفت H.261 و MPEG-1 ، نیاز به کدینگی قویتر بود تا کاربردهای وسیعی را پوشش دهد . به همین دلیل ITU-T و ISO/IEC با همکاری هم در سال 1995 توانستند MPEG-2/H.262 را که عموماً با نام MPEG-2 معروف است را ارائه دهند .

MPEG-1 از جریان بیت ثابتی برخوردار نبود ، ولی MPEG-2 چون در کاربرد های صدا و تصویر کارایی زیادی دارد ، باعث شده که MPEG-2 طول بیت انتقال ثابت و کوتاهی داشته باشد . MPEG-2 دو نوع جریان تعریف می کند ؛ PS جریان برنامه و TS جریان انتقال . که PS شبیه به جریان MPEG-1 می باشد (با طول متغیر و بلند) ولی با استفاده از توابع و امکانات پیشرفته تر ، به همین دلیل دگر MPEG-2 می تواند داده های MPEG-1 را نیز کدگشایی کند .

جریان برنامه (PS) MPEG-2 ویژگی هایی دارد که MPEG-1 آنها را پشتیبانی نمی کند مانند ، scrambling داده ها و یا بسته بندی کردن داده ها و ... بعلاوه دارای یک بخش اختیاری است که در آن اطلاعات مربوط به تست شبکه موجود ، اطلاعات تشخیص خطا و ... می باشد .

اما جریان انتقال (TS) MPEG-2 کاملاً متفاوت است ، چون طول ثابت و کوتاهی دارد (188 بایت) و در برابر نویز کانال احتمال خطا کمتری دارد .

(در بالا به طور مفصل در مورد مالتی پلکسی MPEG-2 بحث شده است .)

MPEG-2 از روی MPEG-1 با اضافه کردن تصاویر در هم بافته (interlaced) و همچنین گسترش دادن اندازه تصاویر و نرخ بیت آن ها ساخته شده است . در اصل MPEG-2 تعمیم یافته MPEG-1 می باشد .

MPEG-2 کاربرد های متفاوتی دارد ، به همین دلیل به پروفایل ها و سطوح مختلفی تقسیم می شود که پروفایل ها میزان پیچیدگی را بیان می کنند و LEVEL ها بیانگر اندازه تصاویر هستند ، همه Level ها همه پروفایل ها را پشتیبانی نمی کنند .

LEVELS		PROFILES				
	Simple main without B-frame 4:2:0	Main B-frames 4:2:0	4:2:2	SNR Scalability 4:2:0	Spatial Scalability 4:2:0	Professional
High 1920 x 1152	x	80 Mbit/s	x	X	x	100 Mbit/s
High-1440 1440x1152	x	60 Mbit/s	x	x	60 Mbit/s	80 Mbit/s
Main 720x576	15 Mbit/s	15 Mbit/s 90% users	50 Mbit/s	15 Mbit/s	x	20 Mbit/s
Low 352x288	x	4 Mbit/s	x	4 Mbit/s	x	x

جدول ۲-۲ (پروفایل ها وسطوح مختلف تقسیم MPEG-2)

پروفایل اصلی که 90% کاربران از آن استفاده می کنند در پخش تلویزیون های دیجیتال نیز

از آن استفاده می شود و با نام (main Profile and Main Level) MP@ML شناخته

می شود .

در پایان مراحل فشرده سازی تصویر را به وسیله نرم افزار MATLAB شبیه سازی می کنیم،

برنامه مربوطه به صورت زیر است :

```
>> o = imread('e:\img10.jpg');
w = size(o, 2);
samplesHalf = floor(w / 2);
samplesQuarter = floor(w / 4);
samplesEighth = floor(w / 8);
ci2 = [];
ci4 = [];
ci8 = [];
for k=1:3% all color layers: RGB
    for i=1:size(o, 1)% all rows
        rowDCT = dct(double(o(i,:,k)));
        ci2(i,:,k) = idct(rowDCT(1:samplesHalf), w);
        ci4(i,:,k) = idct(rowDCT(1:samplesQuarter), w);
        ci8(i,:,k) = idct(rowDCT(1:samplesEighth), w);
    end
end

h = size(o, 1);
samplesHalf = floor(h / 2);
samplesQuarter = floor(h / 4);
samplesEighth = floor(h / 8);
ci2f = [];
```

```

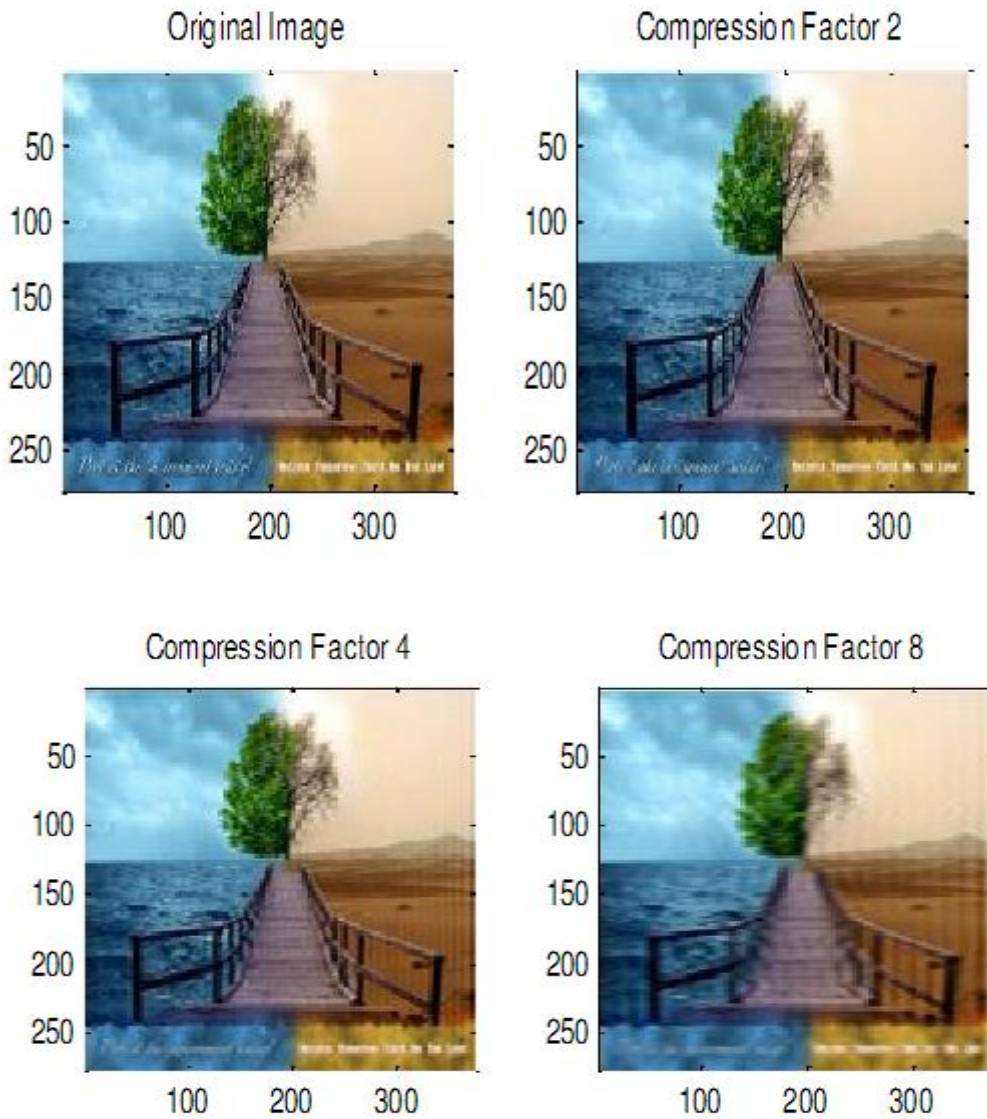
ci4f = [];
ci8f = [];
for k=1:3% all color layers: RGB
    for i=1:size(o, 2)% all columns
columnDCT2=dct(double(ci2(:,i,k)));
columnDCT4=dct(double(ci4(:,i,k)));
columnDCT8=dct(double(ci8(:,i,k)));
ci2f(:,i,k) = idct(columnDCT2(1:samplesHalf), h);
ci4f(:,i,k) = idct(columnDCT4(1:samplesQuarter), h);
ci8f(:,i,k) = idct(columnDCT8(1:samplesEighth), h);
    end
end
subplot(2,2,1), image(uint8(o)), title('Original Image');
subplot(2,2,2), image(uint8(ci2)), title('Compression Factor 2');
subplot(2,2,3), image(uint8(ci4)), title('Compression Factor 4');
subplot(2,2,4), image(uint8(ci8)), title('Compression Factor 8');
figure
subplot(2,2,1), image(uint8(o)), title('Original Image');
subplot(2,2,2), image(uint8(ci2f)), title('Compression Factor 2 * 2');
subplot(2,2,3), image(uint8(ci4f)), title('Compression Factor 4 * 4');
subplot(2,2,4), image(uint8(ci8f)), title('Compression Factor 8 * 8');

```

خروجی ها برای تصاویر با کیفیت های مختلف :

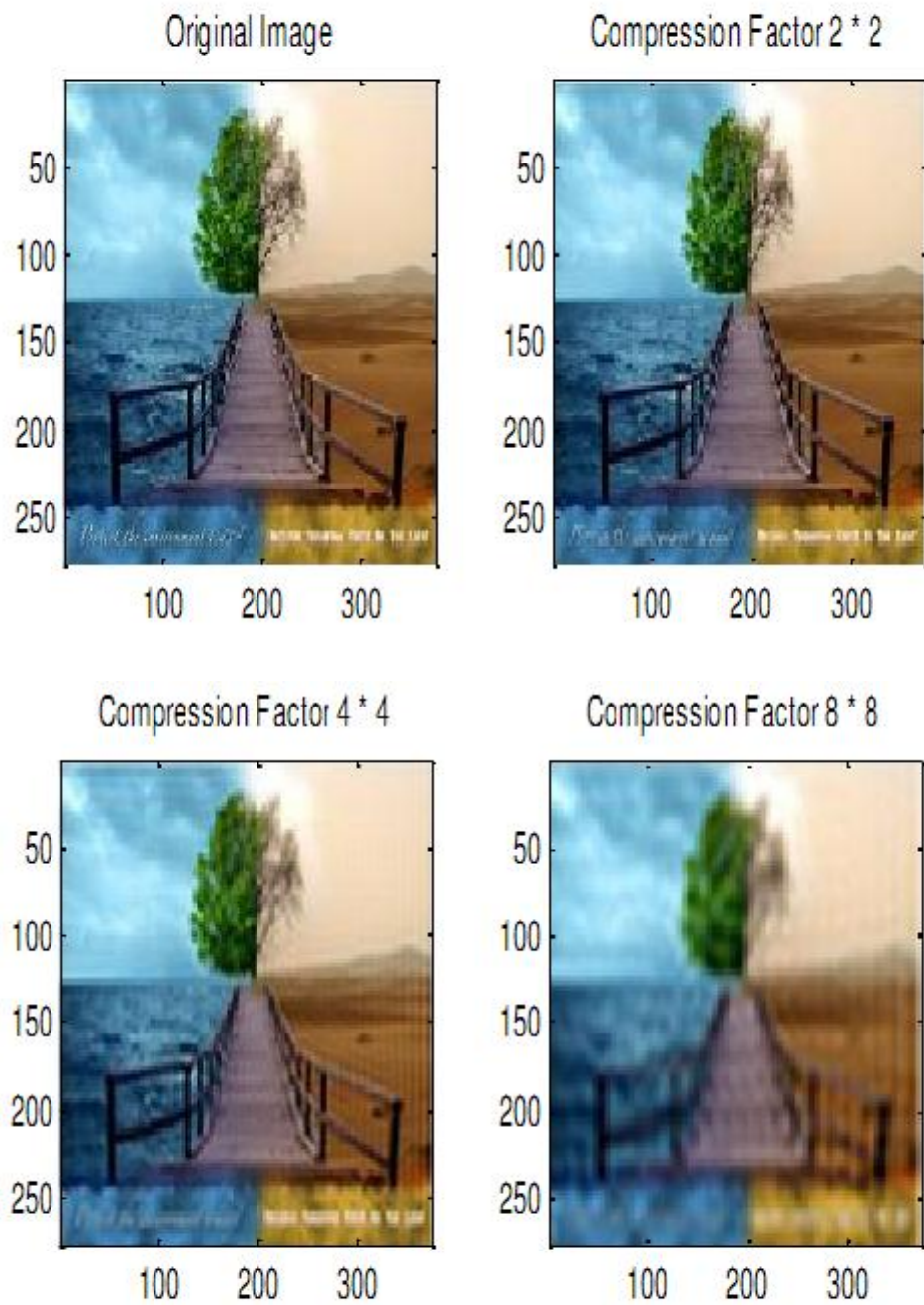
Image of size $375 \times 276 \times 3$ (Low Resolution Image)

a) DCT to rows only:



شکل ۳-۱ (Dct To Row Only تصویر اول)

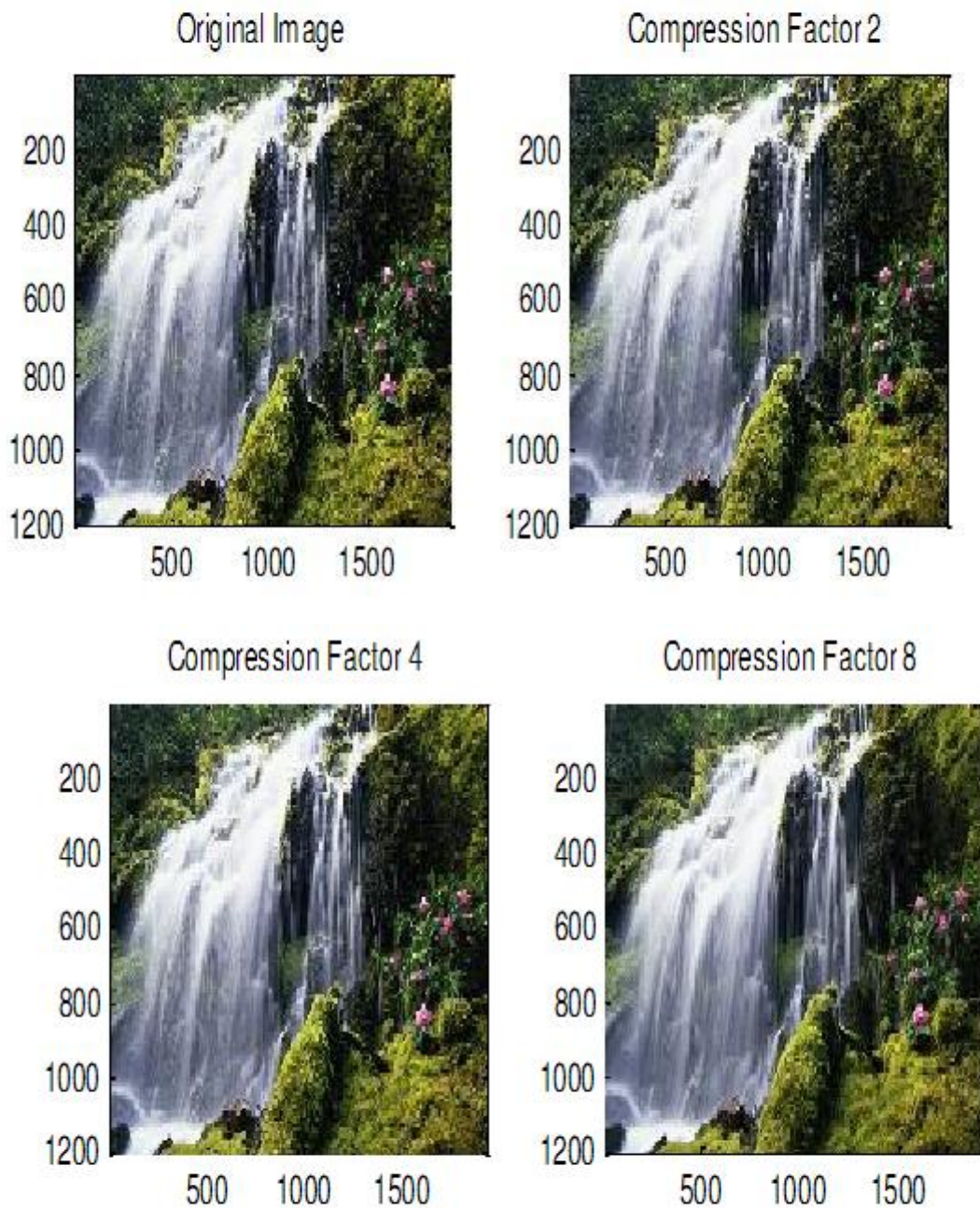
b) DCT to rows*column:



شکل ۲-۳ (Dct To Row*Column تصویر اول)

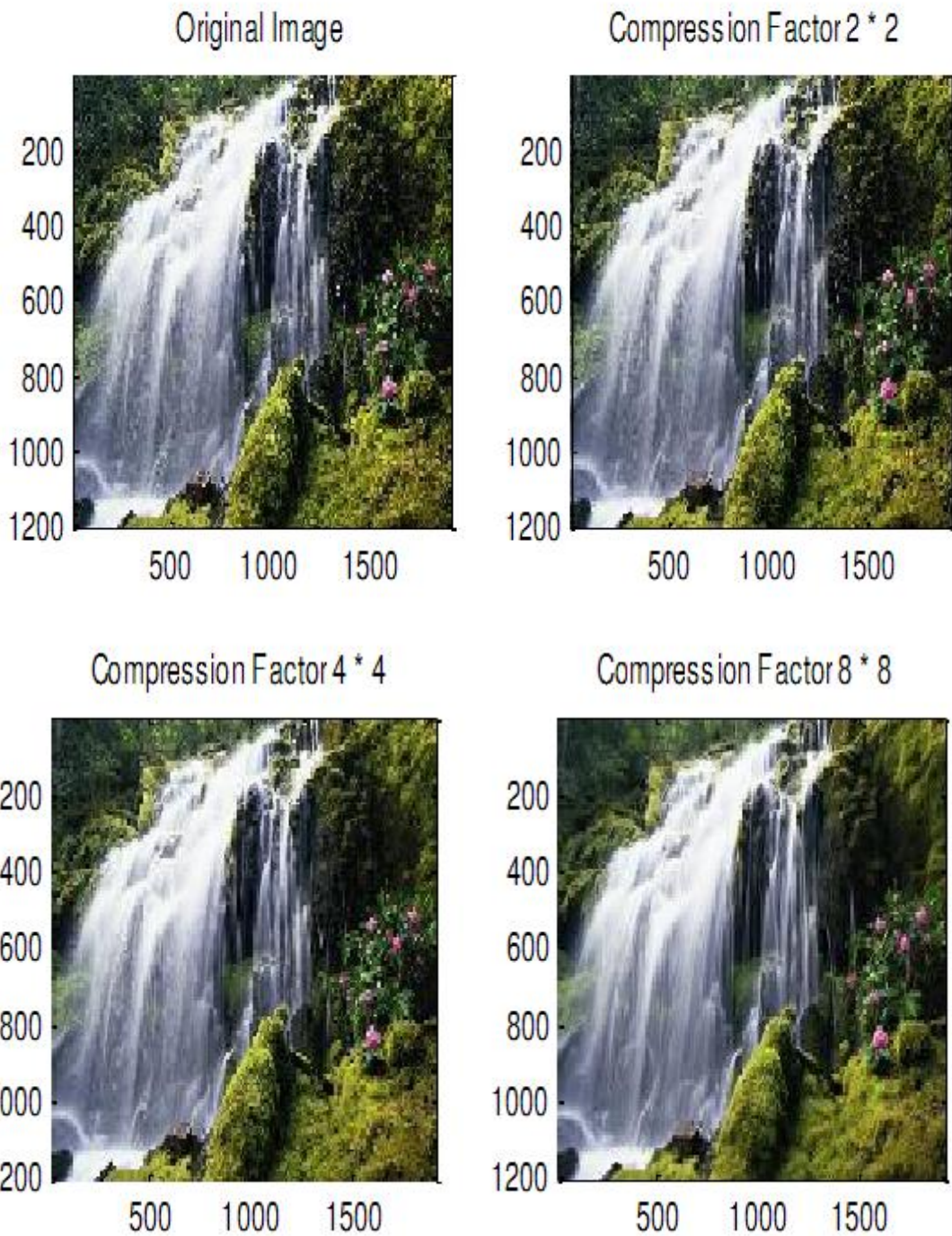
Image of size $1920 \times 1200 \times 3$ (High Resolution Image):

a) DCT to rows only:



شکل ۳-۳ (Dct To Row Only تصویر دوم)

b) DCT to rows*column:



شکل ۴-۳ (Dct To Row*Column تصویر دوم)

منابع :

- مجموعه مقالات جهاد خود کفایی
- کتاب پردازش تصویر گونزالس
- www.wikipedia.org
- Mathworks Matlab Central