دانشکده فنی و مهندسی

گروه برق

گرایش:الکترونیک

پروژه کارشناسی

پیاده سازی پخش فایل های MP3 توسط میکروکنترلر AVR

استاد راهنما:جناب آقای مهندس احمدی

سحر علیمحمدی

تابستان ۱۳۸۷

فهرست

مقدمه ۱

فصل ۱ : شروع کار با MP3 ..... ۲-۷

۱- مقدمه ۲-۳

۲- معرفی لایه ها ۳-۴

۳-فشرده سازی بدون تلف یا تلفاتی ۴-۷

۱ – ۳ : پهنای باند بحرانی گوش ۵

۲ – ۳ : فرآیند هم پوشانی ۶-۷

فصل ۲ : هدر MP3 ..... ۸-۱۶

۱- مقدمه ۹

۲- CRC ۹

۳- اطلاعات حاشیه ای ۹

۴- داده اصلی ۱۰

۵- داده فرعی ۱۰-۱۱

۶- جدول هدر ۱۱-۱۵

۷- CBR یا VBR ۱۵

۸- Tag ۱۶

فصل ۳ : پروسهٔ رمزگذاری ۱۷-۲۲

۱- الگوریتم فشرده سازی هافمن..... ۲۱-۱۷
۲- رمز گذاری..... ۲۲-۲۱
فصل ٤ :پروسة رمزگشايي ۳۹- ۲۲
۱- مقدمه ۲۳
۲- هافمن ۲۵-۲۴
۳- اطلاعات حاشیه های..... ۳۲-۲۶
۴- دوباره تدریجی کردن ۳۴-۳۳
۵ – دوباره مرتب کردن..... ۳۵
۶ -Antialias ۳۶-۳۵
۷ -IMDCT..... ۳۷-۳۶
۸ -FilterBank..... ۳۸-۳۷
فصل ٥ ،پيلة سازئ..... ۵۰-۳۹
كليد واژة ۵۳-۵۱
پيوست الف
پيوست ب
منابع

مقدمه :

در ماه ژانویه سال ۱۹۸۹ (MPEG[1]) به وجود آمد که این کمیته ISO / IECJTCI / SC 29[2] W697 / در استاندار سازی الگوریتم فشرده سازی برای تصویر های محرک ، ویدئو شروع به کار کرد.حاصل کار این گروه MPEG – 1 در نوامبر سال ۱۹۹۲ بود که با هم دستی دانشگاه Erlangeno مورد هم کاری واقع شده و نتیجه کار آن ها MP3 بود . استاندارد لایه ۱ و ۲ و ۳ را توصیف می کند . که در نسبت فشرده سازی تفاوت دارند و پیچیدگی هم همراه با فشرده سازی افزوده می شود . لایه ۳ به عنوان MP3 شناخته شده که با خاطر کیفیت بالا و بیت ریت پایین معروف است.فشرده سازی به روش های مختلفی انجام می شود که به عنوان مثال فشرده سازی بدون تلف[3] همانندzip فقط۱۵٪ اندازه فایل را فشرده میکند ولی با استفاده از تکنیکی به نام(ادراک از طریق حواس ۵ گانه[4]) می توان ۹۰٪ فایل اصلی را حذف کرد بدون اینکه در شنیدن تفاوتی ایجاد شود .که در فایل های MP3از این نوع فشرده سازی استفاده می شود[14]

[1] Moving Picture Expert Group فرمتي استاندارد براي فشرده سازي صوت و تصوير
[2] اختصار عبارت internatonal standard organization (موسسه جهاني استاندارد)
[3] lossless
[4] perceptual .

شرح کار MP3 :

۱-۱مقدمه :

فایل های MP3 یکی از عجیب ترین پدیده هایی هستند که صنعت موسیقی تا به حال شاهد آن بوده است . فرمت MP3 یک سیستم فشرده سازی موسیقی می باشد که بدون صدمه زدن به کیفیت پخش آهنگ از تعداد بایت های آن می کاهد . هدف فرمت MP3 آن است که بدون تأثیر گذاشتن به کیفیت آهنگ آن را فشرده سازد .به کمک فرمت MP3 یک آهنگ 32مگا بایت به ۳ مگابایت فشرده شده و حجم آن کاهش می یابد .سوالی که این جا مطرح است این است که آیا عملیات فشرده سازی بدون آسیب رساندن به کیفیت آن ممکن است ؟

برای طراحی الگوریتم MP3 از خصوصیات گوش انسان الهام گرفته شده است این خصوصیات به ترتیب زیر هستند :

۱) صدا هایی در طبیعت وجود دارند که گوش انسان قادر به شنیدن آن ها نمی باشند .

۲) صدا هایی در طبیعت وجود دارد که گوش انسان بهتر از موجودات دیگر قادر به شنیدن آن ها است .

۳) اگر ۲ صدا به طور هم زمان در حال پخش باشند در این صورت گوش انسان تنها صدایی را می‌شنود که بلند تر از دیگری باشد .

به کمک اصول فوق بدون این که صدمه ای به کیفیت آهنگ وارد شود قسمت های خاصی از یک آهنگ را می توان حذف کرد . عملیات فشرده سازی با استفاده از تکنیک های شناخته شده ای انجام می گیرد . وقتی عملیات ساخت فایل MP3 به پایان می رسد در اصطلاح می گویند آهنگ به دست آمده(نزدیک به کیفیت CD)[5] می باشد . یعنی نسخه ی MP3 از یک آهنگ ، به علت حذف قسمت هایی از آن کاملا شبیه CD اوریجینال نمی باشند ولی خیلی نزدیک به آن است .

در رمزگذاری صوت[6] از فرمت WAV استفاده میشود. WAV غیر فشرده است، و به دسته های فراوانی تقسیم می شوند که یکی از آن ها PCM [7] است که برای رمزگذاری MP3 قابل قبول است. در رمزگذاری صوت در هر ثانیه ۴۴۱۰۰ عدد نمونه برداری می شود که هر یک از این نمونه ها فضایی معادل ۲ بایت یا ۱۶ بیت را اشغال می کنند . در سیستم های استریو نمونه های مربوط به باند

---

[5] Near CD Quality

[6] Audio Coding

[7] . Pulse Code Modulation یک استاندارد براي ذخيره و انتقال صوت غير فشرده ديجيتال كه سيگنال ورودي را بر حسب تعداد نمونه در واحد زمان نمايش مي دهد .

ها (چپ و راست) از هم جدا می شوند به این ترتیب یک فایل موسیقی می تواند تعداد انبوهی بیت در یک ثانیه ذخیره کنند .بنابراین در یک سیستم استریو داریم:

کیلو بایت در ثانیه 176 = 100 44 * 2 * 16 Bit = 1411200 Bits/s / 8 / 1000 = کانال

حال می خواهیم در باره ی آمار بالا بیشتر صحبت کنیم :

۱/۴ میلیون بیت در هر ثانیه معادل ۱۷۶ کیلو بایت در ثانیه است . بنابراین اگر یک موسیقی به طور متوسط ۳ دقیقه باشد ، در نتیجه این مقدار از موسیقی به طور متوسط فضایی معادل ۳۲ مگابایت اشغال می کند این مقدار بر یک آهنگ فضای زیادی است مخصوصا اگر بخواهیم با یک مودم K56 یک آهنگ را دانلود کنیم که برای این کار نزدیک ۲ ساعت زمان نیاز است .[13]

۲ - انواع لایه ها :

۱- MPEG : این فرمت در سال ۱۹۹۲ برای سیگنال های مونوو استریو عرضه شده و نرخ های نمونه برداری ورودی ۳۲ ، ۴۴/۱ یا ۴۸ KHz را قبول کرده و نرخ بیت خروجی ۳۲ تا ۳۸۴ $Kbit/S$ می باشد . و دارای ۳ لایه مستقل از هم می باشد که به تدریج پیچیده تر شده و بازده آن ها افزایش می یابد . لایه I (حالت اول) ساده ترین ساز و کار را دارد .

سیگنال ورودی به بسته های ثابت ۳۸۴ نمونه ای (معادل ۸ میلی ثانیه در فرکانس ۴۸ KHz) تقسیم شده و توسط فیلتر ۳۲ باندی تفکیک می شود . پهنای باند این فیلتر ها ثابت بوده (به دلیل سهولت) و بنابراین در فرکانسهای بم بیش از حد پهن ودر فرکانس های زیر بیش از حد لزوم باریک اندواحد تحلیل گر و مدل کننده از فوریه FFT[8] با دقت ۵۱۲ نقطه استفاده می کند . که دقت متوسطی به شمار می رود . در هر باند ۱۲ نمونه (۳۸۴ / ۳۲) وجود دارد که همه با هم پردازش می‌شوند و توسط مقیاس کننده ۶ بیتی کوانتیزه می گردند .

لایه II (حالت دوم) کمی پیچیده تر است . برش زمانی به ۱۱۵۲ نمونه معادل ۲۴ میلی ثانیه در فرکانس ۴۸ KHz و دقت مبدل FFT به ۱۰۲۴ نقطه افزایش می یابد ولی فیلتر ۳۲ باندی تغیر نمی کند . واحد تخصیص دهنده بیت نیز بهبود یافته و عمل کوانتیزه را برای فرکانس های بم و متوسط و زیر به تفکیک انجام می دهند .

لایه III که به MP3 معروف است رفتار باز هم پیچیده تری دارد . برش زمانی برای سیگنال های عادی ۱۱۵۲ نمونه است ولی برای سیگنال گذرا و ضربه ای به ۳۸۴ نمونه کاهش می یابد . واحد تخصیص دهنده بیت بسیار دقیق تری عمل می کند .

[8] Fast Fourier Transform

۲- MPEG : حالت BC این فرمت در سال ۱۹۹۴ عرضه شد و صدای سراند را نیز به سیگنال مونو و استریو هم چنین نرخ های نمونه برداری ۱۶ / ۲۲۰۵ / ۲۴ کیلو هرتز را به استاندارد MPEG 4 اضافه می کند ولی ساختار آن بسیار مشابه MPEG – 1 است . حالت Acc بسیار پیشرفته بوده که مبنای MPEG 4 نیز قرار گرفت . فیلتر ها پیچیده تر شده و برش زمانی بین ۲۵۶ و ۴۸ و ۲ نمونه متغیر است واحد تحلیل گر و مدل کننده اثرات هم پوشانی نیز کاملا بیهوده یافته و از یک الگوریتم پیش بینی کننده استفاده می نماید . در عمل این فرمت صدایی با کیفیت بهتر از فرمت های دیگر حاصل می کند نرخ داده های خروجی حداکثر برابر با ۲۸۸ کیلو بایت در ثانیه برای هر کانال است .

۳- AC : فرمتی که در سال ۱۹۹۲ توسط شرکت دالبی معرفی شد و اولین فرمت فشرده سازی است که امکان استفاده از صدای سرانه را فراهم ساخت . برش زمانی در حالت عادی برابر ۵۱۲ نمونه است . ورودی به طور منظم برای بررسی سیگنال ضربه ای کنترل می شود و در صورت وقوع آن ، برش زمانی به نصف کاهش می یابد .

واحد تخصیص بیت از روش ممیز شناور و حذف موارد متوالی یکسان برای صرفه جویی در حجم داده ها کمک می گیرد و حداکثر $Kbit/S$ ۶۴۰ است .

ATRAC [9]: فرمتی که توسط شرکت سونی در سیستم MD استفاده شده است در این روش سیگنال ورودی به ۲ نیمه فرکانس تقسیم شده و نیمه بالایی خود به ۲ نیمه تقسیم می شود . محدوده فرکانس های 0 – 5 / 5 KHz به ۲۰ باند و ۲ محدوده فرکانس های میانی و زیر هر یک به ۱۶ باند تقسیم می شود (مجموعا ۵۲ باند) برش زمانی نیز بر حسب ضربه ای بودن سیگنال بین ۵۴ / ۱ تا ۶ / ۱۱ میلی ثانیه متغیر است . سپس واحد تخصیص دهنده بیت عمل فشرده سازی داده ها را انجام می دهد به طوری که نرخ خروجی برابر با ۲۹۲ کیلو بیت بر ثانیه معادل تقریبا یک پنجم نرخ داده در دیسک فشرده می رسد.نسخه های جدید شامل (mdlp) نیز هستند.که با فشرده سازی بیشتر و البته افت کیفیت زمان ضبط و پخش را به ۲ تا ۴ برابر افزایش می دهد .

۳-۱ فشرده سازی بی و تلفاتی:

روش فشرده سازی به ۲ روش اصلی : بی تلف و تلفاتی [10] تقسیم می شود . در روش بی تلف پیام پس از رمزگشایی در گیرنده به طور کامل و بیت به بیت باز یافت می شود .ولی نسبت فشرده سازی آن کم و در حدود ۱ به ۲ است . (حجم داده ها نصف می شود .) در مقابل در روش تلفاتی نسبت فشرده سازی می تواند به یک به ده یا حتی بیشتر نیز برسد، که به بهای حذف مقداری از داده ها منجر، و در نتیجه افت کیفیت به دست می آید . این افت کیفیت در فشرده سازی با نسبت کم ،

[9] . یک فایل صوتی با اتلاف اطلاعات که توسط شرکت Sony ایجاد شده است .
[10] lossy

ناچیز بوده وقابل اغماض است، ولی در فشرده سازی با نسبت زیاد در مواردی مانند ارتباط تلفنی و اینترنتی که کیفیت صوتی کم تری مورد نیاز است ، مقدار قابل ملاحظه ای می باشد .

در روش تلفاتی برای حصول نتیجه ی بهتر از تأثیرات آکوستیک _ روانی استفاده می شود . بدین معنی که داده هایی حذف می شوند که در حالت عادی کمتر شنیده می شوند . و تأثیر خفیف تری در کیفیت صدای مورد نظر دارند . در حقیقت نوعی رفتار پوششی [11] در شنوایی انسان وجود دارد که باعث می شود تا برخی اصوات غالب ، برخی دیگر از اصوات کم اهمیت تر را پوشانیده و مانع شنیده شدن آن ها گردند . بنابراین حذف دقیق و پیش بینی شده اصوات اضافه توسط سیستم فشرده ساز، سبب کاهش حجم داده ها شده و در عین حال افت کیفیت ناشی از این حذف به حداقل می رسد . مهم ترین تأثیرات آکوستیک_ روانی که در فشرده سازی به کار می رود عبارتند از :

۱ - ۳ - پهنای باند بحرانی گوش :

شنوایی انسان دارای رفتاری است که به آن پهنای باند بحرانی [12] می گویند. در داخل حلزونی گوش غشایی به نام Basilar وجود دارد مانند یک تحلیل گر طیفی عمل می کند و فرکانس های مختلف در صدا را تفکیک می کند . فرکانس های مختلف باعث ارتعاش نقاط متفاوت این غشا می شوند، به این معنی که فرکانس های بسیار بم یک انتها و فرکانس های بسیار زیر در انتهای مقابل و فرکانس های دیگر نیز به تناسب در طول غشا درک می شوند . حال اگر فرکانس های خاصی با دامنه ی زیاد سبب ارتعاش ناحیه متناظر خود در غشا شوند ، و سیگنال دیگری با فرکانس نزدیک ولی دامنه ی کم تر در همان زمان وارد گوش شود نمی تواند آن ناحیه غشا را که هم اکنون در حال ارتعاش است به خوبی تحریک کند . بنابراین توسط سیگنال قبلی پوشانیده شده و شنیده نمی شود مگر آن که فرکانس این سیگنال ضعیف به قدر کافی (یعنی به اندازه پهنای باند بحرانی در آن فرکانس مرکزی) از فرکانس قوی تر فاصله بگیرد . برای مثال پهنای باند بحرانی گوش در فرکانس مرکزی ۱۰۰۰ هرتز در حدود ۱۶۰ Hz می باشد بنابراین در حضور فرکانس ۱۰۰۰ هرتز قوی ، فرکانس ۱۰۵۰ هرتز با دامنه ضعیف شنیده نخواهد شد . زیرا در محدوده پهنای باند بحرانی گوش در فرکانس مرکزی ۱۰۰۰ هرتز قرار دارد . در حالی که فرکانس ۱۲۰۰ هرتز با دامنه ضعیف تشخیص داده می شود چون در خارج از محدوده پهنای باند بحرانی گوش در فرکانس مرکزی ۱۰۰۰ قرار دارد .

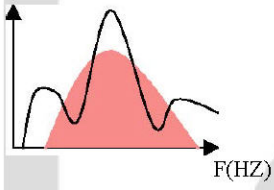
---

[11] Masking

[12] Critical – Band .

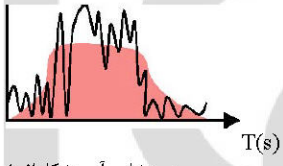
۲ - ۳ - ۱ فرآیند همپوشاندگی :

یک سیگنال صوتی از فرکانس هایی با دامنه متفاوت تشکیل شده است اگر فرکانس غالب وجود داشته باشد ، این فرکانس حضور فرکانس های دیگر، در مجاورت را حذف خواهد کرد ، در حقیقت به این معنی است که صدایی به خاطر حضور صدایی دیگر غیر قابل شنیدن باشد. همپوشانی فرکانسی زمانی اتفاق می افتد که یک صدای سنگین یک صدای سبک را که از لحاظ فرکانس به هم نزدیک اند را پوشش می دهد. آستانه ای[13] که نشان می دهد که فرکانس قابل حذف کردن است یا نه از روی قوی ترین صدا در سیگنال صوتی محاسبه می شود .در شکل ۱-۱شمایی از فرآیند همپوشانی دیده می شود.



شکل 1-1: همپوشانی فرکانسی

محدودیت دیگر این است که اگر یک فرآیندآنی[14] در سیگنال صوتی ایجاد شود، گوش انسان برای مدت کوتاهی به فرکانس کم تر از فرآیند آنی کم تر حساس خواهد بود، در این حالت فرآیندی اتفاق می افتد که به همپوشانی آنی[15] معروف است .شکل ۲-۱این فرآیند را نشان میدهد.



همپوشانی آنی :شکل۲-۱

همپوشاندگی فرکانسی :

این اثر در واقع ملاحظاتی درباره پهنای باند بحرانی است که توسط منحنی های هم پوشانی تن یا نویز بیان می شود .

---

[13] Masking Threshold
[14] Transient
[15] Temporal Masking

هم پوشانی زمانی :

اثر هم پوشانی به صورت زمانی هم عمل می کند بدین شکل که هر گاه ۲ سیگنال تقریبا هم زمان (ولی نه کاملا هم زمان) اتفاق بیافتند سیگنال ضعیف تر(مستقل از رفتار فرکانسی)تحت پوشش سیگنال قوی قرار گرفته و شنیده نمیشود.اگر سیگنال قوی قبل از سیگنال ضعیف اتفاق افتاده و پوشیده شود آن را پیش – پوشش [۱۶] و اگر بعد از سیگنال قوی اتفاق بیافتد و پوشیده شود آن را پس– پوشش [۱۷] می نامند . اثر هم پوشانی در حالت پس - پوشش شدید و موثر تر از حالت پیش - پوشش است .[1]

[16] Pre _ Masking
[17] Post – Masking

دانشجویان محترم:

جهت دسترسی به متن کامل پایان نامه‌ها به کتابخانه دانشکده مهندسی و یا آزمایشگاه پروژه گروه برق مراجعه فرمایید.

مدابع :

[1]Praveen Sripada and Blekinge Tekniska Hogskola ,"MP3 decoder in theory and practice" March 20006

[2] *ISO/IEC 11172-3:1993 Information technology – Coding of moving pictures and associated audio for digital storage media at up to about 1,5 Mbit/s*, 1993.

[3] S. Gadd and T. Lenart, "A hardware accelerated mp3 decoder with bluetooth streaming capabilities," Master's thesis, Lund Institute of Technology, November, 2001.

[4] B. G. Lee, "A new algoritm to compute the discrete cosine transform," IEEE transactions on acoustics, speech and signal processing, vol ASSP-32, No 6, December 1984.

[5] V. Britanak and K. R. Rao, "A new fast algorithm for the unified forward and inverse mdct/mdst computaion," Signal Processing, 2002.

[6] K. Brandenburg and H. Popp, "An Introduction to MPEG Layer-3", Fraunhofer Institute, EBU Technical Review, June 2000.

[7] Specification of the Bluetooth System, Version 1.1, Specification Volume 1 Core, February 2001.

[8] Specification of the Bluetooth System Version 1.1, Specification Volume 2 Profiles, February 2001.

[9] Rolf Johannesson and K. Sh Zigangirov ,"Fundamentals of Convolutional Coding",

[10] M. Ruckert, *Understanding MP3*, Vieweg, 2005, ISBN 3-528-05905-2.
[11] Raissi R., *"Theory behind MP3"*, 5th October 2005, <http://www.mp3-tech.org>.
[12] Mathew M., Bhat V.,Thomas S.M., Yim C., *"Modified MP3 Encoder using complex modified discrete cosine transform"* , ICME 2003.
[13] Fältmann I, Hast M, Lundgren A, Malki S, Montnemery E, Rångevall A, Sandvall J,Stamenkovic M., *"A Hardware implementation of an MP3 decoder"* , Digital IC project, LTH,Sweden , May 2003.
[14] Geoff Nicholson, *"MP3 Explained: A Beginners guide"* 18th January
[15] *"MP3 converter"*, 15th December 2005, <http://www.mp3-converter.com>.
[16] <http://www.wikipedia.com>.
[17] <http://www.myplace.nu/mp3/download>.
[18] http://www.mp3-tech.org/content/?Overview%20of%20the%20MPEG%20committee
[19] http://www.mp3-tech.org/tech.html
[20] http://www.mp3-tech.org/content/?Mp3%20Limitations
[21] http://www.mp3-tech.org/aac.html
[22] http://www.mp3-tech.org/programmer/docs/report_trimedia.pdf