بسم الله الرحمن الرحيم

٣٩٧١٤

دانشگاه تهران
دانشکدهٔ فنی
گروه مهندسی برق وکامپیوتر

۱۳۸۱ /۲/ ۲۰
016956

عنوان:

پردازش در حوزه فشرده سیگنال ویدیوئی MPEG

نگارش :
مهدی ملک آرائی

استاد راهنما:
دکتر امید فاطمی

پایان‌نامه جهت دریافت درجهٔ کارشناسی ارشد در
رشتهٔ مهندسی برق
گرایش مخابرات
اسفند ۱۳۸۰

۳۹۷ ۱۴

دانشگاه تهران

دانشکدهٔفنی

گروه مهندسی برق وکامپیوتر

عنوان:

پردازش در حوزه فشرده سیگنال ویدیوئی MPEG

نگارش :

مهدی ملک آرائی

از این پایان نامه در اسفندماه ۱۳۸۰ در مقابل هیئت داوران دفاع به عمل آمده و مورد تصویب قرار گرفت.

سرپرست تحصیلات تکمیلی دانشکدهٔ فنی:	دکتر محمد علی بنی‌هاشمی
مدیر گروه آموزشی:	دکتر محمود کمره‌ای
سرپرست تحصیلات تکمیلی گروه:	دکتر جواد فیض
استاد راهنما:	دکتر امید فاطمی
عضو هیئت داوران:	دکتر فرخ آرزم
عضو هیئت داوران:	دکتر ناصریزدانی
عضو هیئت داوران:	دکتر محمدرضا پاکروان

چکیده:

در این پایان‌نامه به مبحث پردازش در حوزه فشرده می‌پردازیم و با استفاده از آن الگوریتمی برای وارد کردن لوگو در حوزه فشرده سیگنال ویدیوئی MPEG پیشنهاد می‌کنیم. پردازش روی تصاویر ساکن و متحرک را می‌توان پس از دیکد کردن آنها ودر حوزه مکان انجام داد. ولی انجام مستقیم پردازش در حوزه فشرده، با حذف مراحل دیکدکردن و کدکردن مجدد باعث افزایش سرعت و کاهش پیچیدگی می‌شود. آشنایی با پردازش در حوزه فشرده مستلزم آشنایی با نحوه فشرده سازی تصاویر ساکن و متحرک است. لذا در فصل‌های دوم و سوم به معرفی استانداردهای JPEG و MPEG پرداخته‌ایم. به مبحث پردازش در حوزه فشرده و معادل‌سازی عملیات مختلف حوزه مکان در حوزه فشرده در فصل چهارم اشاره شده است. در این فصل معادل عملیات مختلف حوزه مکان مانند جمع پیکسلی، ضرب پیکسلی، ضرب ماتریسی، U-Flip و جز آن در حوزه فشرده بررسی شده است. در فصل پنجم الگوریتمی پیشنهاد می‌شودکه لوگو را پس از اعمال پردازش‌های ابتدائی، در حوزه فشرده وارد سیگنال ویدیوئی MPEG می‌کند. اشاره خواهد شد که وارد کردن لوگو در فریم‌های I با عملیات کانولوشن در حوزه فشرده و وارد کردن لوگو در فریم‌های B و P مستلزم عملیات جبران‌سازی حرکتی است. به برتری انجام عملیات در حوزه فشرده ضمن مقایسه پیچیدگی عملیات حوزه‌های مکان و فشرده، اشاره شده است.

Abbreviations

اختصارات

BMA.....Block Matching Algorithm

DCT.....Discrete Cosine Transform

DFT.....Discrete Fourier Transform

DST.....Discrete Sine Transform

EOB.....End of Block

FDCT.....Fast Discrete Cosine Transform

GOP.....Group of Pictures

IBA.....Inner Block Algebra

IBT.....Inner Block Transformation

IDCT.....Inverse Discrete Cosine Transform

JPEG.....Joint Programming Expert Group

MAE.....Mean Absolute Error

MMEA.....Modified Motion Estimation Algorithm

MPEG.....Motion Picture Expert Group

MSE.....Mean Square Error

OSA.....Orthogonal Search Algorithm

SNR.....Signal to Noise Ratio

TSS.....Three Step Search

TDL.....Two-dimensional Logarithmic Method

مطلب صفحه

فصل اول: مقدمه

۱-۱- ورود به بحث ۱
۱-۲- ساختار پایان نامه ۲
۱-۲-۱- فصل اول،مقدمه ۲
۱-۲-۲- فصل دوم، آشنایی با JPEG؛ کدبندی تصاویر ساکن ۲
۱-۲-۳- فصل سوم، آشنایی با MPEG؛ کدبندی تصاویر متحرک ۲
۱-۲-۴- فصل چهارم، پردازش در حوزه فشرده ۲
۱-۲-۵- فصل پنجم، وارد کردن لوگو در حوزه فشرده سیگنال ویدیوئی M.PEG ۳
۱-۳- فرمت های RGB و YCrCb در نمایش تصویر ۳
۱-۴- معرفی مدهای Progreesive و Interlced ۵
۱-۵- کدبندی با تلف و بدون تلف ۶
۱-۶- انواع افزونگی در پردازش تصویر ۸
۱-۶-۱- افزونگی مکانی ۹
۱-۶-۲- افزونگی زمانی ۹
۱-۷- نتیجه گیری ۹

فصل دوم: آشنایی با JPEG؛ کدبندی تصاویر ساکن

۲-۱- مقدمه ۱۰
۲-۲- استاندارد JPEG در حالت Baseline Seqential ۱۰
۲-۳- مطالعه بیش‌تر ۱۷
۲-۴- نتیجه گیری ۱۷

فصل سوم: آشنایی با MPEG؛ کدبندی تصاویر متحرک

۳-۱- مقدمه ۱۸
۳-۲- انواع فریم‌ها در MPEG ۱۸
۳-۳- ساختار ورودی کدکننده ۱۹
۳-۳-۱- تعریف GOP ۱۹
۳-۳-۲- تعریف Slice ، ماکروبلاک و بلاک ۲۱

۳-۴- کدبندی تصاویر در MPEG ۲۲
۳-۵- تخمین حرکتی ۲۲
۳-۵-۱- الگوریتم بلاک متناظر (BMA) ۲۳
۳-۵-۲- جبران‌سازی حرکتی با دقت نصف پیکسل ۲۶
۳-۵-۳- جبران‌سازی حرکتی دو جهتی ۲۷
۳-۶- نحوه کدبندی فریم‌های مختلف در MPEG ۲۸
۳-۷- دیکد کننده MPEG ۲۸
۳-۸- مطالعه بیش‌تر ۲۹
۳-۹- نتیجه گیری ۲۹

فصل چهارم: پردازش در حوزه فشرده

۴-۱- مقدمه ۳۰
۴-۲- تبدیل گسسته کسینوسی و فشرده‌سازی تصویر ۳۱
۴-۲-۱- اهمیت تبدیل در فشرده سازی تصویر ۳۱
۴-۲-۲- تبدیل گسسته فوریه ۳۱
۴-۲-۳- تبدیل گسسته کسینوسی ۳۳
۴-۲-۳-۱- گذر از حالت یک بعدی به دو بعدی در تبدیل گسسته کسینوسی ۳۵
۴-۲-۴- مختصری درباره تبدیل گسسته سینوسی ۳۶
۴-۳- معادل‌سازی عملیات حوزه مکان در حوزه فشرده ۳۷
۴-۳-۱- عملیات IBA ۳۹
۴-۳-۱-۱- ضرب در اسکالر ۳۹
۴-۳-۱-۲- جمع با اسکالر ۴۰
۴-۳-۱-۳- جمع پیکسلی ۴۱
۴-۳-۱-۴- ضرب ماتریسی ۴۱
۴-۳-۱-۵- ضرب پیکسلی ۴۲
۴-۳-۲- عملیات IBT ۴۲
۴-۳-۲-۱- U-Flip ۴۴
۴-۳-۲-۲- V-Flip ۴۶
۴-۳-۲-۳- D-Flip ۴۷
۴-۳-۲-۴- OD-Flip ۴۹
۴-۳-۳- محل انجام عملیات در حوزه فشرده ۴۹
۴-۳-۴- اشاره به چند کاربرد الگوریتم‌های حوزه فشرده ۵۰
۴-۳-۴-۱- استخراج زیر بلاک ۵۰

۴-۳-۴-۲- تغییر تدریجی یا Dissolve ۵۱
۴-۳-۴-۳- Alpha Blending ۵۲
۴-۴- نتیجه‌گیری ۵۴

فصل پنجم، وارد کردن لوگو در حوزه فشرده سیگنال ویدیوئی MPEG

۵-۱- مقدمه ۵۵
۵-۲- مراحل مختلف وارد کردن لوگودر حوزه فشرده ۵۶
۵-۲-۱- پردازش اولیه روی لوگو ۵۷
۵-۲-۲- وارد کردن لوگو در فریم I ۵۹
۵-۲-۳- وارد کردن لوگو در فریم‌های P و B ۵۹
۵-۳- مقایسه پیچیدگی محاسبات در حوزه فشرده و حوزه مکان ۶۴
۵-۴- نتیجه‌گیری و پیشنهادات ۷۰

منابع و مراجع

ضمیمه الف. بخشی از برنامه کامپیوتری مربط به وارد کردن لوگو

ضمیمه ب. قسمتی از جدول کدهای هاف‌من در Run Lengh Coding

ضمیمه ج. برنامه کامپیوتری محاسبه تبدیل گسسته کسینوسی سریع

ضمیمه د. برخی مراجع مربوط به تبدیل گسسته کسینوسی و روش‌های محاسبه سریع آن

اشکال و جداول **صفحه**

فصل اول: مقدمه

شکل ۱-۱ : فشرده‌سازی با استفاده از فرمت YCrCb ۶
شکل ۲-۱ : فرمت‌های ۴:۲:۰ و ۴:۴:۴ ..... ۷

فصل دوم: آشنایی با JPEG؛ کدبندی تصاویر ساکن

شکل ۱-۲ : ماتریس کوانتیزه کننده لومینانس ۱۱
شکل ۲-۲ : ماتریس کوانتیزه کننده کرومینانس ۱۱
شکل ۳-۲ : تغییر فرکانسی مؤلفه‌ها در ماتریس DCT ۱۴
شکل ۴-۲ : تکنیک‌های اسکن کردن زیگزاگی ۱۷
شکل ۵-۲ : نمونه ماتریس تبدیل یافته و کوانتیزه شده ۱۵
شکل ۶-۲ : مراحل فشرده‌سازی در استاندارد JPEG ۱۶
شکل ۷-۲ : شمای کلی یک کدکنده JPEG ۱۷

فصل سوم: آشنایی با MPEG؛ کدبندی تصاویر متحرک

شکل ۱-۳ : مرتب سازی فریم‌های ورودی در MPEG ۱۹
شکل ۲-۳ : ساختار GOP در MPEG ۲۰
شکل ۳-۳ : ارتباط میانSlice ، ماکروبلاک و بلاک ۲۱
شکل ۴-۳ : ساختار کدکننده MPEG ۲۲
شکل ۵-۳ : میزان جابه‌جائی در دو فریم متوالی ۲۵
شکل ۶-۳ : تصویرمطلوب و خطای باقیمانده ناشی از جبران‌سازی حرکتی ۲۵
شکل ۷-۳ : یافتن بلاک متناظر ۲۶
شکل ۸-۳ : جبران‌سازی حرکتی با دقت نصف پیکسل ۲۷
شکل ۹-۳ : جبران‌سازی حرکتی دو جهتی ۲۷
شکل ۱۰-۳ : ساختار دیکد کننده MPEG ۲۸

فصل چهارم: پردازش در حوزه فشرده

شکل ۱-۴ : مراحل مختلف فشرده‌سازی و بازیابی تصویر ۳۸
جدول ۲-۴ : خلاصه عملیات IBA ۴۴
شکل ۳-۴ : خلاصه عملیات U-Flip ۴۵
شکل ۴-۴ : خلاصه عملیات V-Flip ۴۷
شکل ۵-۴ : خلاصه عملیات D-Flip ۴۸
شکل ۶-۴ : خلاصه عملیاتOD-Flip ۴۹
جدول ۷-۴ : خلاصه عملیات IBT ۴۹
شکل ۸-۴ : استخراج زیر بلاک با ضرب ماتریس‌های پیشین و پسین ۵۱

فصل پنجم، وارد کردن لوگو در حوزه فشرده سیگنال ویدیوئی MPEG

شکل ۱-۵ : نمونه یک لوگو ۵۸
شکل ۲-۵ : تخمین حرکتی دو جهتی با بلاک‌های بعدی و قبلی ۶۰
شکل ۳-۵ : تخمین حرکتی دو جهتی با بلاک‌های بعدی و قبلی ۶۲
شکل ۴-۵ : استخراج زیر بلاک ۶۲
جدول ۵-۵ : میزان صفرها در فریم‌های مختلف سیگنال MPEG ۶۵
جدول ۶-۵ : مقادیر وزنی $W_{m,n,u}$ ۶۸
جدول ۷-۵ : مقایسه میزان محاسبات حوزه مکان و فشرده ۶۹

فصل اول

مقدمه

۱-۱- ورود به بحث

دهۀ ۹۰ میلادی مصادف بود با تدوین استانداردهای MPEG-1 و MPEG-2. استانداردهای MPEG-1 و MPEG-2 امروزه تقریباً در تمام عرصه‌هائی که به نوعی احتیاج به ذخیره، فشرده‌سازی، پردازش و پخش سیگنال‌هائی ویدیوئی وجود دارد به کار می‌روند. در هر دو استاندارد با روشهای مختلفی سعی شده است تا سیگنال ویدیوئی خام دریافتی تا حد ممکن فشرده شود و در عین حال دارای کیفیت مطلوبی باشد.

در بسیاری موارد علاقه‌مندیم تا روی سیگنال ویدیوئی MPEG پردازش انجام دهیم. به عنوان یک مثال ساده فرض می‌کنیم، برنامه‌های یک شبکۀ تلویزیونی به صورت MPEG ذخیره شده است و سپس برای ارسال از طریق ماهواره به مرکزی فرستاده می‌شود. این مرکز در نظر دارد قبل از ارسال برنامه از طریق ماهواره، لوگوی[1] تبلیغاتی خود را در گوشه‌ای از برنامه وارد کند.

یک روش برای وارد کردن لوگو این است که سیگنال MPEG دریافتی را دیکد کنیم، تغییرات مورد نظر را (که همان وارد کردن لوگو است) در حوزۀ مکان اعمال نمائیم و سیگنال تغییر یافته را مجدداً کد کنیم. این روش مستلزم عملیات مفصل دیکد کردن و کد کردن مجدد است، اگر بتوانیم به طریقی تغییرات مورد نظر را روی سیگنال در حالتی که فشرده است اعمال کنیم، احتیاجی به عملیات مفصل دیکد کردن و کد کردن مجدد نخواهد بود. در واقع مسأله اصلی که در پایان‌نامه به آن می‌پردازیم "پردازش روی سیگنال‌های ویدیوئی MPEG در حوزۀ فشرده[2] است" و وارد کردن لوگو[3] به عنوان یک کاربرد خاص و مهم از پردازش در حوزه فشرده بررسی خواهد شد.

آشنائی با مبحث پردازش در حوزه فشرده مستلزم آشنائی با نحوۀ فشرده سازی تصاویر ساکن و متحرک است. لذا ساختار زیر برای پایان‌نامه در نظر گرفته شده است.

1- Logo
2- Compressed Domain
3- Logo Insersion

۱-۲- ساختار پایان‌نامه

۱-۲-۱- فصل اول مقدمه

در مقدمه پس از معرفی اجمالی مسأله پردازش در حوزهٔ فشرده به توضیح نکاتی پرداخته می‌شود که در تصویر و مطالب مرتبط به آن حائز اهمیت است و در سرتاسر پایان‌نامه مورد ارجاع قرار می‌گیرد. این موارد عبارتند از : فرمت‌های RGB و YCrCb در نمایش تصویر،معرفی مدهای Progressive و interlaced، کدبندی با تلف[1] و بدون تلف[2] و انواع افزونگی[3] در پردازش تصویر.

۱-۲-۲- فصل دوم، آشنائی باJPEG، کدبندی تصاویر ساکن

در این فصل به طور مختصر به نحوه کدبندی JPEG اشاره می‌کنیم، تبدیل گسسته کسینوسی، کوانتزیشن و استفاده از کدبندی منبع مورد بررسی قرار می‌گیرند. آشنائی با استاندارد JPEG خود مقدمه‌ای است برای آشنائی با نحوهٔ کدبندی در استاندارد MPEG. در بسیاری از بخش‌های این فصل از[۱] استفاده شده است.

۱-۲-۳- فصل سوم، آشنائی با MPEG، کدبندی تصاویر متحرک

در فصل سوم به معرفی استاندارد MPEG می‌پردازیم. انواع فریم‌ها در MPEG و نحوه کدبندی آنها، جبران‌سازی و تخمین حرکتی و دیکد کنندهٔ MPEG از مواردی است که در این فصل به آنها پرداخته می‌شود. از [۱] در بسیاری از بخش‌های این فصل استفاده شده است.

۱-۲-۴- فصل چهارم، پردازش در حوزهٔ فشرده

در فصل چهارم ابتدا به مزایا و خصوصیات پردازش در حوزه فشرده اشاره خواهد شد، اهمیت تبدیل‌ها و معرفی تبدیل گسسته کسینوسی از دیگر بخش‌های این فصل است. در انتهای فصل به

1- Lossy
2- Lossless
3- Redundancy

معادل سازی عملیات مختلف حوزهٔ مکان مانند جمع و ضرب پیکسلی، ضرب ماتریسی و مانند آن در حوزهٔ فشرده اشاره خواهد شد.

۱-۲-۵- فصل پنجم، وارد کردن لوگو در حوزه فشرده سیگنال ویدیوئی MPEG

در این فصل به یک کاربرد مهم از مبحث پردازش در حوزهٔ فشرده اشاره می‌کنیم و الگوریتمی را برای وارد کردن لوگو در این حوزه پیشنهاد می‌کنیم. در الگوریتم پیشنهادی، وارد کردن لوگو در فریم‌های I ، P و B مورد بحث قرار می‌گیرد و به برنامه‌های کامپیوتری در هر قسمت اشاره می‌شود. در پایان نیز به نتیجه‌گیری و نکاتی دربارهٔ ادامه کار در این زمینه پرداخته شده است.

در قسمت‌های ۱-۳ تا ۱-۶ به توضیح مسائلی کلی خواهیم پرداخت که در سراسر پایان‌نامه به آنها اشاره خواهد شد.

۱-۳- فرمت‌های RGB و YCrCb در نمایش تصویر

آنچه نور را از دیگر امواج الکترومغناطیسی متمایز می‌کند حساسیت چشم انسان به آن است. چشم انسان به طول موج‌های ۳۵۰nm تا ۷۵۰nm حساس است. برداشت و ادراک سیستم بینائی انسان از یک تصویر را می‌توان با سه پارامتر : روشنائی[1]، رنگ[2] و کروما[3] بیان کرد.

درحوزه دید انسان میزان روشنائی هر تصویر به λ بستگی دارد. اگر فرض کنیم $C(\lambda)$ نشان دهندهٔ تغییرات دامنهٔ موج برحسب طول موج است در آن صورت در یک نور تک رنگ[4] $c(\lambda)$ همه جا به جز در $\lambda=\lambda_r$ صفر است. حال در نظر بگیریم که از کسی می‌خواهیم روشنائی $C(\lambda_r)$ را با $C'(\lambda_t)$ مقایسه کند که λ_t طول موج آزمون است. فرض کنیم با تغییر $C'(\lambda_t)$ نهایتاً روشنائی دو نور از

1- Brightness
2- Hue
3- Chroma
4- Monochromatic Light

نظر بیننده یکسان شود، در این صورت $\frac{C(\lambda_r)}{C'(\lambda_t)}$ لومینانس[1] نسبی نامیده می‌شود. اگر لومنیانس نسبی را که تابعی از طول موج است با $V(\lambda)$ نمایش دهیم، می‌توان مقدار لومنیانس را با رابطه (۱-۱) حساب کرد که در آن k عددی ثابت است [۲، ص‌ص ۴۱۷-۴۱۳] و [۳].

$$\ell = k\int_{\lambda=0}^{\infty} c(\lambda)v(\lambda)d\lambda \qquad (۱-۱)$$

رنگ تصویر هم می‌تواند به $c(\lambda)$ مرتبط شود. برای مثال اگر $c(\lambda)$ در گستره بینائی انسان ثابت باشد نور سفید یا بی‌رنگ به نظر می‌رسد. در یک نور تک رنگ، رنگ به λ_i بستگی دارد. هر چه عرض باندی که نور در آن قرار دارد کمتر شود نور شفاف‌تر به نظر می‌رسد و هر چه پهن‌تر شود، کدرتر و مات‌تر به نظر می‌رسد. پارامترهای شفافیت و ماتی در واقع همان پارامتر کرومای تصویر است[۴].

رنگ‌های اصلی را قرمز (R)، سبز (G) و آبی (B) می‌دانیم که با ترکیب آنها می‌توان رنگ‌های مختلف را ساخت. مقادیر R، G و B را می‌توان به مجموعه دیگری از پارامترهای معرف تصویر تبدیل کرد و از این تبدیل در فشرده‌سازی تصویر استفاده کرد. برای مثال سیستم بینائی انسان به تغییرات مکانی در روشنائی نسبت به تغییرات مکانی در رنگ، حساس‌تر است. نمونه‌ای از این پارامترها YUV است که در سیستم PAL استفاده می‌شود و در آن Y بیانگر لومنیانس و U و V بیانگر دو مؤلفه رنگ تصویر است. محاسبه پارامترهای YUV از روی RGB با روابط (۲-۱) انجام می‌شود.

$$\begin{cases} Y = 0.299R + 0.587G + 0.114B \\ U = -0.147R - 0.289G + 0.436B \\ V = 0.615R - 0.515G - 0.100B \end{cases} \qquad (۱-۲)$$

1- Luminance