دانشکده مهندسی برق و کامپیوتر

پایان نامه‌ی کارشناسی ارشد در رشته مهندسی برق-مخابرات (سیستم)

کدینگ توزیع شده ویدئو

به کوشش

وحید باستانی

استاد راهنما

دکتر مهران یزدی

بهمن ۱۳۹۰

بسم الله الرحمن الرحیم

به نام خدا

اظهارنامه

اینجانب وحید باستانی (880652) دانشجوی رشته‌ی برق-مخابرات گرایش سیستم دانشکده‌ی برق و کامپیوتر اظهار می‌کنم که این پایان‌نامه حاصل پژوهش خودم بوده و در جاهایی که از منابع دیگران استفاده کرده‌ام، نشانی دقیق و مشخصات کامل آن را نوشته‌ام. همچنین اظهار می‌کنم که پایان‌نامه و موضوع پایان‌نامه‌ام تکراری نیست و تعهد می‌نمایم که بدون مجوز دانشگاه دستاوردهای آن را منتشر ننموده و یا در اختیار غیر قرار ندهم. کلیه‌ی حقوق این اثر مطابق با آیین‌نامه‌ی مالکیت فکری و معنوی متعلق به دانشگاه شیراز است.

نام و نام خانوادگی: وحید باستانی
تاریخ و امضاء: ۱۳۹۰/۱۱/۱۱

به نام خدا

کدینگ توزیع شده ویدئو

به کوشش

وحید باستانی

پایان‌نامه

ارائه شده به تحصیلات تکمیلی دانشگاه شیراز به عنوان بخشی از فعالیت‌های تحصیلی لازم برای اخذ درجه کارشناسی ارشد

در رشته‌ی

مهندسی برق – مخابرات سیستم

دانشگاه شیراز

شیراز

جمهوری اسلامی ایران

ارزیابی شده توسط کمیته‌ی پایان نامه با درجه : **عالی**

دکتر مهران یزدی، دانشیار بخش مهندسی مخابرات و الکترونیک (رئیس کمیته) .....

دکتر عزیرالله جمشیدی، استادیار بخش مهندسی مخابرات و الکترونیک .....

دکتر علیرضا ذوالقدر اصلی، دانشیار بخش مهندسی مخابرات و الکترونیک .....

بهمن ماه ۱۳۹۰

سپاسگزاری

اکنون که این پایان نامه به پایان رسیده است بر خود لازم می‌دانم که از استاد ارجمندم آقای دکتر مهران یزدی به خاطر زحمات بی‌بدیل و راهنمایی‌های گران‌قدرشان که همیشه پرفایده و راه‌گشا بوده است تشکر و قدردانی به عمل آورم.
از اساتید مشاور آقایان دکتر عزیزالله جمشیدی و دکتر علیرضا ذوالقدر اصلی نیز به خاطر نظرات ارزنده شان سپاسگزارم.

چکیده

کدینگ توزیع شده تصاویر ویدئویی

به کوشش

وحید باستانی

محدودیت‌های موجود در کاربردهای جدید ویدئو دیجیتال از قبیل کمبود منبع انرژی و محدودیت در منابع پردازشگر و حافظه، استفاده از الگوریتم‌های معمول کدینگ ویدئو را با مشکل مواجه می‌کند. کدینگ توزیع شده ویدئو به عنوان یک جایگزین برای کدینگ‌های پیش‌بینی کننده معمول بسیار مورد توجه قرار گرفته است. انعطاف پذیریِ این روش کدینگ جدید، به کار بردن آن را در محیط‌های دارای محدودیت سخت افزاری میسر ساخته است. انکدر این روش‌ها به دلیل عدم نیاز به پیش‌بینی فریم می‌تواند بسیار ساده باشد. تکنیک‌های کدینگ توزیع شده ویدئو فعلی از روش‌های کدینگ کانال و یک کانال فیدبک بین انکدر و دیکدر برای کنترل نرخ بیت بهره می‌برند. مکانیزم کانال فیدبک، پیاده‌سازی عملی این کدک‌ها را در کاربردهای بلادرنگ مشکل می‌کند.

در این پایان‌نامه تلاش شده یک روش کدینگ توزیع شده جدید ارائه گردد که نیاز به کانال فیدبک نداشته باشد. در این روش از یک تکنیک کوانتیزاسیون دینامیک ضرایب تبدیل برای استخراج افزونگی‌های زمانی بین فریم‌های ویدئو بهره گرفته شده است. در دیکدر، یک روش تخمین و جبرانسازی حرکت وفقی با اندازه بلوک متغیر برای تولید اطلاعات جانبی استفاده شده است. این کار، پیش‌بینی کیفیت تصویر خروجی را در انکدر میسر کرده است و توانسته است نیاز به کانال فیدبک را از بین ببرد.

نتیاج بدست آمده نشان داده است که این روش می‌تواند علاوه بر حذف محدودیت‌های روش‌های معمول کدینگ توزیع شده بازدهی آن را نیز افزایش دهد. این کار با هزینه‌ی اندک افزایشی در پیچیدگی محاسبات انکدر نسبت به روش‌های معمول کدینگ توزیع شده صورت پذیرفته است.

فهرست مطالب

عنوان **صفحه**

فصل ۱. مقدمه ۲

فصل ۲. پیشینه تحقیق ۷

۲. ۱. تصاویر ویدئویی دیجیتال ۷

۲. ۱. ۱. تلویزیون آنالوگ ۷

۲. ۱. ۲. درهمآمیزی ۹

۲. ۱. ۳. تفکیک پذیری ۱۰

۲. ۱. ۴. فرمت رنگ ۱۳

۲. ۱. ۵. عمق بیت ۱۴

۲. ۲. اصول و مفاهیم فشرده سازی ویدئو ۱۶

۲. ۲. ۱. افزونگی در سیگنال ویدئو ۱۶

۲. ۲. ۲. استخراج افزونگی مکانی ۱۸

۲. ۲. ۳. استخراج افزونگی آماری ۲۱

۲. ۲. ۴. استخراج افزونگی زمانی ۲۳

۲. ۲. ۵. پیشبینی دو جهتی ۲۶

۲. ۲. ۶. کنترل کوانتیزاسیون ۲۷

۲. ۲. ۷. انتخاب حالت ۲۷

۲. ۲. ۸. کنترل نرخ ۲۸

۲. ۲. ۹. بلوک دیاگرام یک کدک ویدئو ۲۸

۲. ۳. تئوری کدینگ توزیع شده ۳۰

۲. ۳. ۱. انکدینگ مستقل، دیکدینگ مستقل ۳۱

۲. ۳. ۲. انکدینگ توأم، دیکدینگ توأم ۳۲

۲. ۳. ۳. انکدینگ مستقل، دیکدینگ توأم ۳۳

۲. ۳. ۴. انکدینگ همراه با اطلاعات جانبی در دیکدر ۳۴

عنوان **صفحه**

۲. ۴. روشهای کدینگ توزیع شده ویدئو ۳۵
۲. ۴. ۱. معماریهای اولیه کدینگ توزیع شده ویدئو ۳۷
۲. ۴. ۲. پیشرفتهای بعدی کدینگ توزیع شده ۴۳
۲. ۴. ۳. کدک وینر-زیو DISCOVER ۴۶
۲. ۵. کدینگ توزیع شده بدون کانال فیدبک ۵۵
۲. ۵. ۱. روش اول ۵۵
۲. ۵. ۲. روش دوم ۵۶

فصل ۳. روش پیشنهادی ۶۰
۳. ۱. معرفی ۶۰
۳. ۲. تبدیل DCT صحیح ۶۲
۳. ۳. کوانتیزاسیون ۶۴
۳. ۴. کدینگ آنتروپی ۶۵
۳. ۵. کنترل نرخ و بازسازی فریم توزیع شده ۶۷
۳. ۶. تولید SI و پیشبینی کیفیت آن ۷۱

فصل ۴. ارزیابی ۷۸
۴. ۱. دیتابیس و شرایط آزمایش ۷۹
۴. ۲. نتایج نرخ و اعوجاج ۸۱
۴. ۳. بررسی پیچیدگی انکدر ۸۶

فصل ۵. نتیجه گیری ۹۰

منابع ۹۳

فهرست جدول ها

عنوان **صفحه**

جدول ۲-۱: تفکیکپذیری VGA ۱۲
جدول ۲-۲: ابعاد تصویر در HDTV ۱۳
جدول ۲-۳: مثال ران-لول کدینگ ۲۳
جدول ۳-۱: پله و پارامترهای کوانتیزاسیون ۶۴
جدول ۳-۲: رابطه پارامترهای کوانتیزاسیون و مقدار *LevelScale* ۶۵
جدول ۴-۱: ساختار GOP در تستها ۸۰
جدول ۴-۲: روش کوانتیزاسیون ۸۱

فهرست شکل ها

عنوان **صفحه**

شکل ۲-۱: (a) الگوی جاروب یک فریم با ۱۱ خط. (b) شکل موجهای جاروب مربوطه ۱۰
شکل ۲-۲: فرمتهای مختلف تفکیکپذیری در سیستم PAL ۱۲
شکل۲-۳: (a) فرمت 4:2:2 (b) فرمت 4:1:1 (c) فرمت 4:2:0 ..... ۱۵
شکل۲-۴: تبدیل و عکس تبدیل DCT ۱۹
شکل ۲-۵: مثال تبدیل DCT و کوانتیزاسیون ۲۱
شکل ۲-۶ : نحوه جاروب ضرایب DCT برای کد کردن ۲۲
شکل ۲-۷ : نحوه جبران سازی حرکت برای یک ماکروبلوک ۲۵
شکل ۲-۸ : پیشبینی تصاویر B ۲۶
شکل ۲-۹ : ترتیب در تصاویر B ۲۷
شکل ۲-۱۰ : بلوک دیاگرام یک انکدر ویدئو نوعی ۲۹
شکل ۲-۱۱ : بلوک دیاگرام یک دکدر ویدئو نوعی ۳۰
شکل ۲-۱۲ : چهار حالت انکدینگ و دیکدینگ دو منبع ازاطلاعات وابسته ۳۱
شکل ۲-۱۳ : ناحیه نرخ بیت مجاز در حالت انکدینگ و دیکدینگ مستقل ۳۲
شکل ۲-۱۴ : ناحیه نرخ بیت مجاز در حالت انکدینگ و دیکدینگ توأم ۳۳
شکل ۲-۱۵ : ناحیه نرخ بیت مجاز در حالت انکدینگ مستقل و دیکدینگ توأم ۳۴
شکل ۲-۱۶ : معماری کدک استنفورد [12] ۳۸
شکل ۲-۱۷ : معماری کدک برکلی [15] ۴۱
شکل ۲-۱۸ : معماری کدک DISCOVER ۴۷
شکل ۲-۱۹ : تولید اطلاعات جانبی در DISCOVER ۵۲
شکل ۲-۲۰ : انکدر روش [28] ۵۵
شکل ۲-۲۱ :کنترل نرخ در انکدر روش [29] ۵۸
شکل ۳-۱ : بلوک دیاگرام کدک پیشنهادی ۶۱
شکل ۳-۲ : فرآیند کدینگ CABAC ۶۷
شکل ۳-۳ : فلوچارت تولید SI به روش VBSMC ۷۳
شکل ۳-۴ : نمودار پراکنده MSE ورودی و MSE خروجی بدست آمده توسط MEC با اندازه بلوکهای 4×4، 8×8 و 16×16 ..... ۷۵

عنوان **صفحه**

شکل ۳-۵ : نمودار پراکنده MSE ورودی و MSE خروجی بدست آمده توسط روش VBSMC ارائه شده ۷۶
Soccer..... ۸۰
شکل ۴-۱ : یک فریم از دنبالههای تست..... ۸۰
شکل ۴-۲ : نتایج نرخ – اعوجاج دنباله Foreman ۸۳
شکل ۴-۳ : نتایج نرخ – اعوجاج دنباله Hall-monitor..... ۸۴
شکل ۴-۴ : نتایج نرخ – اعوجاج دنباله Coastguard ۸۵
شکل ۴-۵ : نتایج نرخ – اعوجاج دنباله Soccer ۸۶
شکل ۴-۶ : کل زمان انکدینگ دنباله Foreman ۸۷
شکل ۴-۷ : کل زمان انکدینگ دنباله Hall-monitor ۸۸
شکل ۴-۸ : کل زمان انکدینگ دنباله Coastguard..... ۸۸

فهرست نشانه‌های اختصاری

BCH	Bose Chaudhuri Hocquenghem (Channel Coding)
CIF	Common Intermediate Format
CRC	Cyclic Redundancy Code
DCT	Discrete Cosine Transform
DISCOVER	Distributed Coding for Video Services
DF	Distributed Frame
DVC	Distributed Video Coding
DSC	Distributed Source Coding
EOB	End Of Block
FMCI	Fast Motion Compensated Interpolation
FPS	Frame Per Second
GB	Giga Byte
GOP	Group Of Picture
HDTV	High Definition TV
IDCT	Inverse Discrete Cosine Transform
IEC	International Electrotechnical Commission
ISO	International Standard Organization
KF	Key Frame
LDPC	Low Density Parity Code
LSB	Least Significant Bit
MB	Mega Byte
MEC	Motion Estimation/Compensation
MMSE	Minimum Mean Squared Error
MPEG	Motion Picture Experts Group
MSB	Most Significant Bit
MSE	Mean Squared Error
NTSC	National Television System Committee
PAL	Phase Alternating Line
PSNR	Peak Signal to Noise Ratio
QP	Quantization Parameter
RGB	Read Green Blue
SECAM	Séquentiel Couleur Avec Mémoire
SI	Side Information
SPA	Sum Product Algorithm
VGA	Video Graphic Array

VBSMC	Variable Block Size Motion Compensation
WZ	Wyner Ziv
WZF	Wyner Ziv Frame
WZVC	Wyner Ziv Video Coding

فصل اول

مقدمه

کدینگ‌های ویدئو سنتیِ مبتنی بر مدل هایبریدِ تبدیل-پیش‌بینی در استخراج افزونگی‌های موجود در دنباله تصاویر ویدئویی بسیار موثر هستند. انکدر این کدک‌ها از الگوریتم‌های پیچیده تخمین و جبرانسازی حرکت برای ناهمبسته کردن فریم‌های ویدئو و کاهش تعداد بیت مورد نیاز برای نمایش آن‌ها بهره می‌برند. این معماری کدک به نحو موثری در استاندارد H.264/AVC پیاده‌سازی شده است. مهمترین محدودیت در استفاده از این معماری کدک سنتی پیچیدگی بسیار زیاد انکدر آن به دلیل انجام فرآیند پیش‌بینی است. این کدک‌ها برای استفاده در کاربردهای یک-به-چند طراحی شده‌اند، یعنی کاربردهایی که یک ماشین پرقدرت و سریع برای یک بار انکدینگ ویدئو به کار گرفته می‌شود و تعداد زیادی وسیله ارزان و ضعیف، به دفعاتِ زیاد برای دیکدینگ آن استفاده می‌شود. مثالی از این کاربردها انتشار تلویزیونی و ذخیره سازی ویدئو است. اما این مدل با کاربردهای جدید ویدئو در مخابرات موبایل یا شبکه دوربین‌های نظارتی همخوانی ندارد. در این کاربردها کارایی ماشین انکدر کمتر و یا یکسان با ماشین دیکدر است. در مخابرات ویدوئی موبایل تعداد و قابلیت‌های انکدرها و دیکدرها تقریبا برابر است. بنابراین برای داشتن بهترین بازدهی باید پیچیدگی فرآیند انکدینگ و دیکدینگ نیز همسان باشد. در شبکه دوربین‌های نظارتی تعداد زیادی دوربین برای دریافت و انکدینگ ویدئو بکار می‌روند، در حالی که تنها چند سرور جهت دیکدینگ آن‌ها استفاده می‌شوند. در این حالت باید انکدر تا اندازه ممکن ساده باشد و پیچیدگی سیستم فشرده‌سازی ویدئو به دیکدر انتقال یابد.

بر اساس دو قضیه اسلیپین-ولف و وینر-زیو در تئوری اطلاعات که به سال‌های دهه ۷۰ برمی‌گردند، ثابت شد که دوگان مدل کدینگ پیش‌بینی کننده منابع نیز قابل دست یافتن است. قضیه اسلیپین-ولف درباره کدینگ توزیع شده بدون اتلاف دو منبع همبسته است. این

قضیه اثبات می‌کند که نرخ بیت مینیمم برای انکدینگ مستقل دو منبع وابسته آماری و دیکدینگ توأم آن‌ها برابر با آنتروپی توأم آن دو منبع است. قضیه وینر-زیو یک حد پایین برای کدینگ با اتلاف منبع X که از لحاظ آماری به منبع دیگر Y وابسته است را بدست می‌آورد. به منبع Y اطلاعات جانبی می‌گویند. Y تنها در دیکدر موجود است و در انکدر دسترسی به آن موجود نیست. این قضیه اثبات می‌کند که اگر منابع گوسی و بدون حافظه باشند و معیار اعوجاج MSE باشد، باند پایین نرخ بیت مورد نیاز برای انکدینگ X با یک اعوجاج مشخص، در حالتی که Y در هر دو سمتِ انکدر و دیکدر موجود باشد، یا تنها در دیکدر موجود باشد، تفاوتی نمی‌کند.

با توجه به نیاز موجود برای یک انکدر ساده ویدئو و بر اساس قضایای اسلیپین-ولف و وینر-زیو, در سال‌های اخیر کدینگ توزیع شده ویدئو یا کدینگ وینر-زیو ویدئو به وجود آمده است. در کدینگ توزیع شده ویدئو فریم‌های دنباله ویدئو به صورت منابع وابسته اطلاعات در قضایای اسلیپین-ولف و وینر-زیو فرض می‌شوند. این فریم‌ها به طور مستقل انکد شده و به طور توأم دیکد می‌شوند. بنابراین دیگر نیاز به اجرای فرآیند پیچیده تخمین و جبرانسازی حرکت در انکدر نیست و این کار در دیکدر برای تولید اطلاعات جانبی انجام می‌شود. بنابراین می‌توان یک انکدر ویدئو به سادگی انکدرهای ویدئو غیر پیش‌بینی کننده داشت.

اولین تلاش‌ها برای ارائه یک کدک ویدئو توزیع شده در حدود سال ۲۰۰۲ با معرفی دو معماری کدک استنفورد و برکلی به نتیجه رسیده است. در معماری استنفورد از کدهای کانال و کنترل نرخ در دیکدر استفاده شده و کدک برکلی از روش رده‌بندی بلوک‌های تصویر و کدینگ سندرم بهره می‌برد. مدل استنفورد بیشتر مورد توجه قرار گرفته است. اولین نسخه آن برای کار در دامنه پیکسل ارائه شده بود و بعد از آن بهبودهای زیادی روی آن انجام شده است. کدینگ در دامنه تبدیل، استفاده از کدهای کانال جدید، روش‌های پیشرفته تولید اطلاعات جانبی و مدل‌سازی بهتر نویز همبستگی از جمله بهبودهایی است که روی کدک اولیه استنفورد انجام شده است. یکی از پیشرفته ترین کدک‌های توزیع شده ویدئو مبتنی بر معماری استنفورد کدک DISCOVER است که در آن از بسیاری از بهبودهای اشاره شده استفاده شده است.

در روش‌های کدینگ توزیع شده ویدئو دنباله فریم‌های ویدئو به دو گروه فریم‌های کلیدی و فریم‌های وینر-زیو تقسیم می‌شوند. فریم‌های کلیدی با یکی از روش‌های کدینگ داخلی

(Intra-coding) فشرده شده و به دیکدر ارسال می‌شود. فریم‌های وینر-زیو بعد از تبدیل، کوانتیزاسیون و باینری شدن به وسیله یک انکدر کانال کد می‌شوند. بیت‌های سیستماتیک کد کانال حذف می‌شود و بیت‌های پریتی آن در بافر انکدر ذخیره می‌شوند تا برحسب نیاز به دیکدر ارسال شوند. در دیکدر اطلاعات جانبی که تخمینی از فریم وینر-زیو بر اساس فرم‌های کلیدی است تولید می‌شود. خطای بین اطلاعات جانبی با فریم وینر-زیو می‌بایست به وسیله بیت‌های پریتی که از انکدر ارسال می‌شود تصحیح گردد. کدهای کانال استفاده شده در کدینگ توزیع شده معمولا کدهای با نرخ تطبیقیِ Turbo یا LDPC هستند. تصحیح خطا در دیکدر با کمترین نرخ بیت پریتی شروع می‌شود. در صورتی که نتوان خطاها را به طور کامل تصحیح کرد، دیکدر از طریق کانال فیدبک تعداد بیشتری بیت پریتی را از انکدر درخواست می‌کند. دیکدر این کار را تا زمانی که دیکدینگ با موفقیت تمام شود انجام می‌دهد. در نهایت فریم وینر-زیو بازسازی شده و دنباله ویدئو در دیکدر تولید می‌شود.

بزرگترین محدودیت در کدک‌های توزیع شده ویدئو استفاده از کانال فیدبک است. اگرچه کانال فیدبک فرآیند کنترل نرخ را بسیار ساده می‌کند اما استفاده از آن در کاربردهای بلادرنگ ممکن نیست. چراکه در این حالت عملکرد انکدر و دیکدر مستقل نبوده و بازدهی انکدر می-تواند با به وجود آمدن مشکلات در دیکدر و کانال تحت تاثیر قرار گیرد. در کاربردهای بلادرنگ انکدر نمی‌تواند صبر کند تا کار دیکدر تمام شده و درخواست داده بیشتری کند. اگر عملکرد دیکدر کند باشد می‌تواند باعث پرشدن و سرریز شدن بافر انکدر شود. در این حالت ممکن است اطلاعات مهمی از دست برود. ذخیره سازی مستقیم داده‌های کد شده نیز با وجود کانال فیدبک ممکن نیست و باید واسطی وجود داشته باشد که کد توزیع شده را دیکد کرده و به فرمت قابل ذخیره سازی تبدیل کند.

در این پایان نامه یک روش جدید برای کدینگ توزیع شده ویدئو پیشنهاد می‌شود که از کدینگ کانال استفاده نمی‌کند و کانال فیدبک نیز در آن حذف شده است. با توجه به اینکه قضایای اسلیپین-ولف و وینر-زیو لزومی به استفاده از کدینگ کانال برای پیاده سازی کدینگ توزیع شده ندارند، در این پایان‌نامه روش دیگری مورد بررسی قرار گرفته که از خواص آماری سیگنال ویدئو بهره می‌برد. یک روش کوانتیزاسیون دینامیک در انکدر این کدک پیشنهاد شده تا بتوان بازدهی آن را را به بازدهی کدک‌های پیش‌بینی کننده نزدیک کرد. پله‌های

کوانتیزاسیون به طور وفقی با توجه به بازدهی نرخ-اعوجاج مورد نیاز و تخمین کیفیت اطلاعات جانبی تولید شده در دیکدر انتخاب می‌شوند. برای ایجاد قابلیت پیش‌بینی کیفیت اطلاعات جانبی در انکدر یک روش تطبیقی تخمین و جبرانسازی حرکت با اندازه بلوک متغیر پیشنهاد شده است. این روش به صورت دینامیک اندازه بلوک‌های تخمین و جبرانسازی حرکت را طوری تغییر می‌دهد تا بتواند به مقدار کیفیت پیش‌بینی شده توسط انکدر نزدیک شود.

فصل دوم