بسم الله الرحمن الرحیم

۱۳۰۷
دانشگاه صنعتی خواجه نصیرالدین طوسی
دانشکده مهندسی برق و کامپیوتر

پایان‌نامه دوره کارشناسی ارشد مهندسی برق–الکترونیک

تخمین حرکت
با استفاده از مدل انرژی و فیلتر کالمن

توسط:

امیر قهرمانی

استاد راهنما:

دکتر امیر موسوی‌نیا

زمستان ۱۳۹۲

تـأیـیـدیّـه هیـات داوران

(برای پایان نامه)

اعضای هیئت داوران، نسخه نهائی پایان نامه خانم / آقای:

را با عنوان:

از نظر فرم و محتوی بررسی نموده و پذیرش آن را برای تکمیل درجه کارشناسی/ کارشناسی ارشد تأیید می‌کند.

اعضای هیئت داوران	نام و نام خانوادگی	رتبه علمی	امضاء
۱ – استاد راهنما			
۲ – استاد مشاور			
۳ – استاد مشاور			
٤ – استاد ممتحن			
٥ – استاد ممتحن			
٦ – نماینده تحصیلات تکمیلی			

تقدیم به

پدر و مادر مهربانم که هرلحظه وجودم را از چشمه‌سار پرعشق چشمانشان سیراب می‌کنند

و

خانواده عزیزم که تمام تجربه‌های یکتا و زیبای زندگیم، مدیون حضور سبز آنهاست

تشکر و قدردانی

قبل از هرچیز، خداوند بزرگ را به خاطر لطفی که همواره شامل حال من نموده شاکرم. سپس، از استاد محترم راهنما، جناب آقای دکتر امیر موسوی‌نیا که وقت خود را به بحث و تبادل نظر در مورد پایان‌نامه بنده اختصاص داده‌اند، تشکر و قدردانی می‌کنم.

چکیده

هدف از انجام پایان‌نامه حاضر، اعمال مدل انرژی و فیلتر کالمن به بحث تخمین حرکت در تصاویر متحرک و ارائه نتایج حاصله از آن می‌باشد. فیلتر کالمن عبارتست از مجموعه روابط بازگشتی که پس از ارائه پیش‌بینی در مورد روندی خاص، با توجه به نتایج عملی به دست آمده از آن روند، اقدام به تصحیح پارامترهای پیش‌بینی خود می‌کند تا در مراحل بعد بتواند پیش‌بینی‌های مفیدتری ارائه نماید. این فیلتر از پارامترهای قابل اندازه‌گیری سیستم برای پیش بینی استفاده می‌نماید. در کاربردهایی مانند ردیابی، این فیلتر از پارامترهایی مانند سرعت و شتاب هدف بهره می‌گیرد. اعمال این روش در چهارچوب تخمین حرکت، نیازمند استفاده از شباهت‌های محلی بلوک‌های روش تطبیق بلوکی جهت ارائه پیش‌بینی برای بلوک‌های تحت بررسی می‌باشد. مدل انرژی که عملا بیانگر توصیفی از ویژگی‌های استاتیک و دینامیک تصاویر است، به همین منظور و برای بهبود دقت پیش‌بینی فیلتر کالمن به خدمت گرفته شده است. در این پایان‌نامه توصیف مذکور در قالب یک هیستوگرام گسسته از انرژی‌های جهت‌دار برای بلوک‌های روش الگوریتم تطبیق بلوکی در جهت‌های بالا، پایین، راست و چپ استخراج می‌گردد. در نهایت یک بین دیگر که نشان‌دهنده عدم وجود ساختار در تصویر می‌باشد به آن هیستوگرام اضافه می‌گردد. پس از استخراج مدل انرژی هر بلوک، شباهت‌های دینامیکی بلوک‌های همسایه با بلوک تحت بررسی استخراج و بردار حرکت پیش‌بین لازم برای فیلتر کالمن با توجه به آنها تولید می‌گردد. در نهایت با استفاده از درون‌یابی دوسویه، پیکسل‌های تعیین‌شده توسط بردار حرکت تصحیح‌شده به وسیله فیلتر کالمن، درون‌یابی شده و تصویر نهایی جبران‌سازی می‌گردد. نتایج حاصل از شبیه‌سازی‌های انجام شده توسط نرم‌افزار متلب نیز بیانگر بهبود کیفی نتایج حاصله می‌باشد. بر اساس این نتایج، روش ارائه شده در این پایان‌نامه نسبت به روش جستجوی سه مرحله‌ای، بهبودی ۰/۶۷ دسی‌بلی در PSNR ایجاد می‌نماید. این در حالی است که از نظر محاسباتی در مقابل آن دارای کاهش است. در مقایسه با روش جستجوی کامل، روش ارائه شده در این پایان‌نامه با انجام محاسباتی ۱۲ برابر کمتر، فقط به اندازه ۰/۴۹ دسی‌بل افت PSNR دارد. در مقایسه با سایر روش‌های تخمین حرکت پیش‌بین مانند فیلتر کالمن ارائه‌شده در مقالات نیز این روش بهبود در کیفیت را نتیجه می‌دهد. لازم به ذکر است الگوریتم پیشنهادی این پایان‌نامه در مواقعی که ویدیو دارای حرکت‌های چندگانه در نواحی کوچک می‌باشد، بهبود بسیاری در کیفیت تصویر بازسازی‌شده ارائه می‌کند.

کلید واژه: تخمین حرکت، فیلتر کالمن، مدل انرژی.

فـهـرست مـطـالـب

عنوان **صفحه**

فهرست جدول‌ها ..... د

فهرست شکل‌ها ..... ه

فصل ۱- مقدمه ۱

۱-۱-پیشگفتار ۱

۱-۲- نگاهی کلی به مبحث ۱

۱-۳- هدف از این پایان‌نامه ۴

۱-۴- فصل‌بندی پایان‌نامه ۴

۱-۵- جمع‌بندی ۶

فصل ۲- تخمین حرکت ۷

۲-۱- مقدمه ۷

۲-۲-الگوریتم‌های تطبیق بلوکی ۷

۲-۲-۱-نگاهی کلی بر مبحث ۷

۲-۲-۲-نحوه کار الگوریتم‌های تطبیق بلوکی ۸

۲-۲-۳- الگوریتم‌های تطبیق بلوکی سریع ۱۲

۲-۳- جمع‌بندی ۲۰

فصل ۳- فیلتر کالمن ۲۱

۳-۱- مقدمه ۲۱

۳-۲- مروری بر فیلتر کالمن ۲۲

۳-۳- تخمین حرکت در تصاویر متحرک با استفاده از فیلتر کالمن ۲۳

۳-۳-۱-نگاهی کلی بر مبحث ۲۳

۳-۳-۲-مدل‌های خود بازگشت برای بردار حرکت ۲۵

۳-۳-۳-فیلتر کالمن تطبیق‌پذیر ۳۷
۳-۴- جمع‌بندی ۴۱
فصل ۴- توصیف انرژی ۴۲
۴-۱- مقدمه ۴۲
۴-۲- مفاهیم اولیه ۴۳
۴-۲-۱- استخراج انرژی زمانی-مکانی ۴۳
۴-۲-۲- فیلترهای قابل پذیرش از لحاظ فیزیولوژیکی ۴۷
۴-۲-۳- تعارض حرکت ۵۰
۴-۳- نگاهی دقیق‌تر به نحوه استخراج انرژی زمانی-مکانی ۵۲
۴-۳-۱- چگونگی حذف فاز در خروجی فیلترهای انرژی ۵۵
۴-۳-۲- پیاده‌سازی فیلترها بصورت هدایت‌شونده و تفکیک‌پذیر ۵۵
۴-۴- توصیف انرژی ۵۷
۴-۴-۱-روش‌هایی برای مقایسه هیستوگرام‌های گسسته ۵۸
۴-۵- نمونه‌هایی از کاربردهای انرژی ۵۹
۴-۶- جمع‌بندی ۶۵
فصل ۵- الگوریتم طراحی‌شده ۶۶
۵-۱- مقدمه ۶۶
۵-۲- سیستم تخمین حرکت سریع طراحی‌شده ۶۷
۵-۲-۱-چالش‌های موجود در راه افزایش دقت پیش‌بینی ۶۷
۵-۲-۲-الگوریتم طراحی‌شده ۶۹
۵-۲-۳-تکنیکی برای کاهش حجم محاسباتی الگوریتم طراحی‌شده ۷۳
۵-۳- جمع‌بندی ۷۶
فصل ۶- نتایج شبیه‌سازی ۷۸
۶-۱- مقدمه ۷۸

۶-۲- کیفیت تصاویر جبران‌سازی‌شده ۷۸
۶-۳- مقایسه هزینه‌های محاسباتی ۸۰
۶-۴- جمع‌بندی ۸۱
فصل ۷- جمع‌بندی و پیشنهادات ۸۲
۷-۱- جمع‌بندی ۸۲
۷-۲- پیشنهادات ۸۳
مراجع: ۸۴
واژه نامه فارسی به انگلیسی ۸۶
واژه نامه انگلیسی به فارسی ۸۹

فهرست جدول‌ها

عنوان **صفحه**

جدول ۶-۱: مقایسه PSNR الگوریتم ارائه شده با پنج الگوریتم جستجوی مذکور. ۷۹
جدول ۶-۲: هزینه‌های محاسباتی لازم برای شش الگوریتم در حال مقایسه مذکور. ۸۱

فهرست شکل‌ها

عنوان **صفحه**

شکل ۱-۱، یک کد کننده ویدیوی DCT ترکیبی نمونه. ۲
شکل ۲-۱: پنجره جستجو و بردار حرکت مطرح‌شده در الگوریتم تطبیق بلوکی. ۱۰
شکل ۲-۲: نحوه جستجوی الگوریتم TSS در قالب یک مثال. ۱۴
شکل ۲-۳: الگوریتم جستجوی سه مرحله‌ای جدید در قالب یک مثال. ۱۷
شکل ۲-۴: الگوریتم جستجوی چهار مرحله‌ای در قالب یک مثال. ۱۹
شکل۳-۱: ساختار کلی تخمین حرکت با استفاده از فیلتر کالمن. ۲۵
شکل۳-۲: مدل خود بازگشت یک بعدی. ۲۸
شکل ۳-۳: مدل خود بازگشت دو بعدی. ۲۸
شکل۳-۴: بلوک‌های مورد استفاده توسط مدل محلی. ۳۱
شکل ۳-۵: بلوک‌های مورد استفاده در مدل زمانی. ۳۴
شکل ۳-۶: محدوده مدل محلی سه‌بعدی. ۳۵
شکل ۴-۱: سلول پایه یک آشکارساز انرژی حرکت و نحوه ساخت یکی از توابع تبدیل آن. ۴۵
شکل ۴-۲: نمونه‌ای از فیلترهای مورد استفاده در شکل ۴-۱. ..... ۴۵
شکل ۴-۳: بیانی دیگر از استخراج انرژی. ۴۷
شکل ۴-۴: نحوه تولید توابع تبدیل لازم برای استخراج انرژی. ۴۸
شکل ۴-۵: طیف انرژی زمانی-مکانی یک واحد حرکت حساس به حرکت در جهت راست. ۴۹
شکل ۴-۶: حالت کلی برای طیف فیلترهای جهت‌دار زمانی-مکانی. ۵۰
شکل ۴-۷: مروری بر انواع واحدهای آشکارساز حرکت مذکور. ۵۱
شکل ۴-۸: نمایه‌های برخی از فیلترهای ذکر شده. ۵۳
شکل ۴-۹: فریم شماره ۲۹ از ویدیوی ترافیکی MERL. ۵۴
شکل ۴-۱۰: نحوه پیاده‌سازی فیلترهای انرژی با استفاده از فیلترهای پایه. ۵۷
شکل ۴-۱۱: نمونه‌ای از هیستوگرام‌های انرژی. ۵۸
شکل ۴-۱۲: نحوه استخراج فریم‌های انرژی مختلف به منظور استفاده در الگوریتم گروه‌بندی. ۶۰
شکل ۵-۱: مثالی از مدل خودبازگشت مکانی استفاده شده توسط فیلتر کالمن. ۶۹
شکل ۵-۲: مدل زمانی و مکانی استفاده شده در این پایان‌نامه. ۷۲
شکل ۵-۳: طرح کلی الگوریتم جستجوی مورد استفاده در این پایان‌نامه. ۷۳

شکل ۵-۴: درون‌یابی دو سویه. ۷۳
شکل ۵-۵، الگوریتم پیشنهادی پایان‌نامه. ۷۷
شکل ۶-۱: نمودار PSNR ویدیوی نمونه bicycle. ۷۹
شکل ۶-۲: مقایسه کیفیت روشهای مورد بحث در قالب یک مثال.. ..... ۸۰

فصل ۱– مقدمه

۱–۱–پیشگفتار

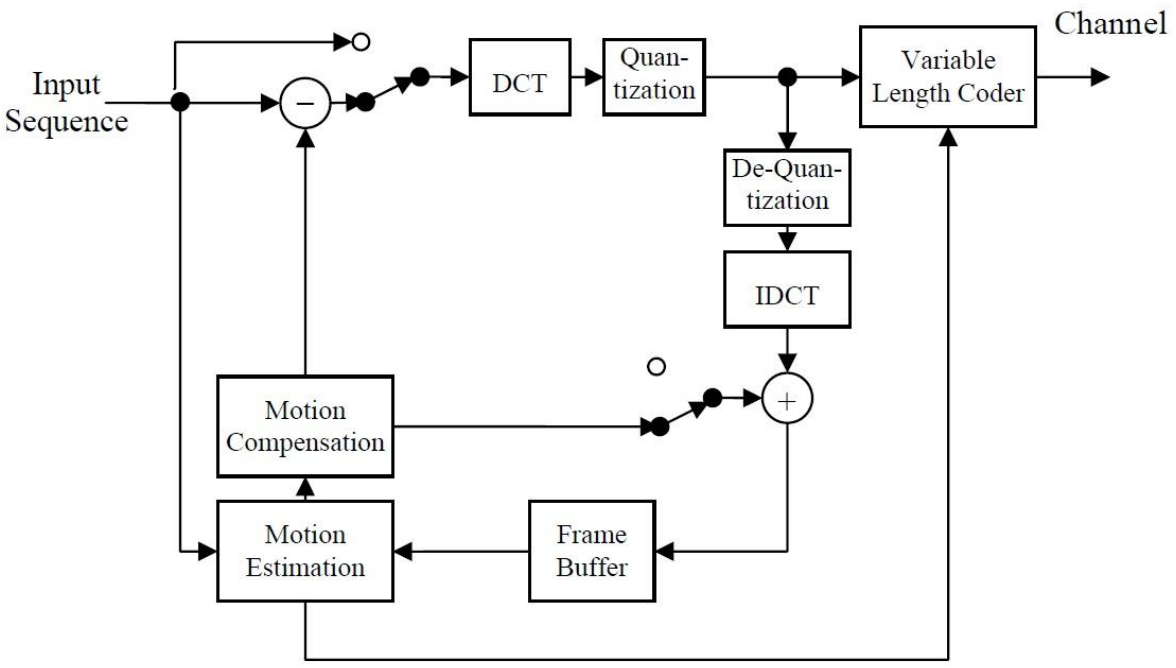
در دهه‌های اخیر، علم بینایی ماشین پیشرفت‌های بسیاری داشته است. این علم زمینه‌های گوناگونی را شامل می‌شود. می‌توان از تخمین حرکت در تصاویر متحرک که همواره به عنوان یکی از مهمترین زمینه‌های علم بینایی ماشین به شمار می‌رود، به عنوان مثالی از این زمینه‌ها یاد کرد. هدف اصلی روش-های تخمین حرکت، فشرده‌سازی ویدیوست. هرچه سیگنال‌های تصویری بیشتر فشرده‌سازی گردند، ذخیره‌سازی و ارسال آنها راحت‌تر صورت می‌گیرد. با توجه به محدودیت‌های سخت‌افزاری موجود در سیستم‌های تصویری، می‌توان به اهمیت ویژه مساله تخمین حرکت پی برد. روش‌های گوناگونی برای تخمین حرکت وجود دارد که در میان آنها، روش تطبیق بلوکی به عنوان ساده‌ترین و مفیدترین روش مورد توجه محققان قرار دارد.

۱–۲– نگاهی کلی به مبحث

اگر از صحنه‌ای با سرعت کافی و بصورت متوالی عکس بگیریم و این عکس‌ها را در بصورت پی‌درپی نشان دهیم، علی‌رغم گسسته بودن تصاویر گرفته شده، ذهن انسان آنها را بصورت پیوسته در نظر خواهد گرفت. اما در واقع تصاویر مذکور مستقل از یکدیگر بوده و دارای طبیعت گسسته‌ای هستند. ویدیوهای امروزی همه از این قاعده پیروی می‌کنند و در حقیقت متشکل از مجموعه‌ای از تصاویر متوالی می‌باشند. هر یک از این تصاویر، خود مجموعه‌ای از پیکسل‌ها می‌باشند که با نظم و ترتیب در کنار یکدیگر قرار گرفته‌اند. طبیعتا این ویدیوها که در کاربردهای گوناگونی مورد استفاده قرار می‌گیرند، باید ذخیره‌سازی و ارسال گردند. اما اگر حجم اطلاعات موجود در هر تصویر را در نظر بگیریم، به این نتیجه می‌رسیم که نیاز به امکانات سخت‌افزاری بسیار زیادی داریم. پس عملا ذخیره‌سازی و ارسال تمام تصاویر موجود در یک ویدیو که احتمالا متشکل از ده‌ها و حتی صدها هزار تصویر می‌باشد، امری ناممکن به نظر می‌آید.

با دقت در تصاویر ویدیویی، در می‌یابیم که در اکثر موارد اختلاف بسیار کمی بین دو تصویر (فریم) متوالی وجود دارد. پس می‌توان به منظور فشرده‌سازی و کاهش حجم ویدیوها در موقع ذخیره‌سازی و ارسال آنها، برخی از فریم‌ها را حذف و در هنگام استفاده از آنها، فریم‌های حذف شده را به کمک فریم-های موجود بازسازی نمود. این کار با استفاده از بردارهای حرکت انجام می‌پذیرد. بدین‌صورت که قبل از حذف فریم‌های انتخاب‌شده، ابتدا ارتباط آنها را با فریم‌های قبلی‌شان بدست آمده و سپس اطلاعاتی کمکی برای آنها استخراج می‌کنیم. این اطلاعات کمکی، همان بردارهای حرکت بوده و نشان می‌دهند که چطور می‌توان فریمی را از روی فریم قبلی‌اش تولید کرد. البته ممکن است بسته به روش مورد استفاده برای تخمین حرکت، نیاز به اطلاعات کمکی دیگری نیز باشد. این اطلاعات کمکی به همراه فریم‌های باقی‌مانده ذخیره‌سازی و ارسال می‌گردند. سپس در هنگام بازسازی تصاویر حذف شده، این اطلاعات مورد استفاده قرار خواهند گرفت.

ایده اصلی در فشرده‌سازی، حذف افزونگی‌های مکانی و زمانی موجود در توالی تصاویر می‌باشد. الگوریتم‌های فشرده‌سازی مورد استفاده در اغلب استانداردهای ویدیویی مانند MPEGx و H.26x، از تکنیک ترکیب روش‌های پیش‌بینی و کُدینگ تبدیلی استفاده می‌کنند. معمولا افزونگی‌های زمانی مذکور توسط سیستم تخمین‌گر حرکت پیش‌بین حذف می‌شوند و وظیفه حذف افزونگی‌های مکانی بر عهده یک سیستم کد کننده تبدیل کسینوسی گسسته (DCT) می‌باشد [۱]. شکل ۱-۱، یک کد کننده ویدیوی DCT ترکیبی را نشان می‌دهد.



شکل ۱-۱، یک کد کننده ویدیوی DCT ترکیبی نمونه [۱].

الگوریتم‌های تطبیق بلوکی که برای پیاده‌سازی ایده تخمین حرکت مطرح گردیده‌اند، همواره برای افزایش سرعت کارشان دستخوش تغییرات در نحوه عملکردشان بوده‌اند. این الگوریتم‌ها در ابتدا با استفاده از الگوهای ثابتی اقدام به جستجو می‌کردند. ابتدایی‌ترین آنها روش جستجوی کامل بود که با مشکل هزینه محاسباتی بسیار زیاد مواجه بود. برای رفع مشکل این الگوریتم، محققان تلاش بسیاری انجام دادند تا بتوانند با استفاده از الگوهای جستجوی سریع‌تری به جواب برسند. دسته الگوریتم‌های کاهنده نقاط جستجو مانند الگوریتم جستجوی سه مرحله‌ای، نتیجه این تلاش‌ها بود. البته چنین الگوریتم‌هایی همواره با مشکل افتادن در مینیمم‌های محلی درگیر بوده و هستند. در نتیجه کیفیت تصاویر بازسازی‌سده توسط آنها بعضا مناسب نیست.

دسته الگوریتم‌های پیش‌بین نمونه‌ای دیگر از الگوریتم‌هایی هستند که به منظور افزایش سرعت سیستم تخمین‌گر حرکت مطرح شده‌اند. این الگوریتم‌ها ابتدا سعی می‌کنند به نحوی بردارهای حرکت بلوک‌های همسایه‌ای که قبلا کُد شده‌اند را با یکدیگر ترکیب و به تخمین اولیه‌ای از بردار حرکت بلوک جاری دست یابند. دلیل آن است که بلوک‌های همسایه، چه از نظر مکانی و چه از نظر زمانی، اغلب دارای شباهت‌های حرکتی بسیاری هستند. در نتیجه می‌توانند حرکت‌های مشابه‌ای نیز داشته باشند. می‌توان بردارهای حرکت آنها را به طرق مختلفی با یکدیگر ترکیب کرد. مثلا می‌توان متوسط آنها را به عنوان بردار حرکت بلوک جاری در نظر گرفت. البته چنین الگوریتم‌هایی نیاز دارند تا در اطراف نقطه پیش‌بینی شده خود نقاطی را برای یافتن بهترین انتخاب جستجو نمایند. اما این نقاط جستجو می‌توانند به مراتب کمتر از نقاطی باشند که الگوریتم‌های دیگر جستجو می‌کنند.

الگوریتم تخمین حرکت توسط فیلتر کالمن نیز الگوریتمی است که از پیش‌بینی اولیه برای بردار حرکت بهره می‌گیرد. این فیلتر ابتدا با استفاده از بردارهای حرکت بلوک‌های همسایه بردار حرکت پیش-بین را استخراج و پس از اندازه‌گیری بردار حرکت، اقدام به تصحیح آن می‌کند. بردار حرکت تصحیح شده دقتی در حد زیر پیکسل دارد. پس می‌تواند منجر به بهبود عملکرد سیستم تخمین‌گر حرکت گردد. البته با توجه به محاسباتی که برای انجام عملیاتش نیاز دارد، مقداری نیز در هزینه محاسباتی الگوریتم افزایش ایجاد خواهد کرد.

سیستم‌های تخمین‌گر حرکتی که پیش از این مطرح گردیده‌اند، همواره از الگوی ثابتی در پیش‌بینی خود استفاده کرده‌اند. بدین معنی که بردارهای حرکت بلوک‌های همسایه خود را با ضرایب ثابتی با یکدیگر ترکیب کرده‌اند. مثل اگر بلوک بالایی بلوک nاُم به اندازه ۰/۲۵ در تعیین بردار حرکتش تاثیر گذار بوده، بلوک n+۱اُم نیز به همین میزان از بلوک بالایی خود بهره گرفته است. اما این در حالیست که

ممکن است بلوک‌های همسایه از بلوکی به بلوک دیگر دارای خصایص حرکتی متفاوتی باشند. مخصوصا اگر صحنه دارای حرکت‌های چندگانه نزدیک به هم باشد، قطعا بلوک‌ها نسبت‌های حرکتی متغیری خواهند داشت. به منظور رفع این مشکل، پایان‌نامه حاضر اقدام به استفاده از توصیف‌های حرکتی بلوک‌ها برای انتخاب بلوک‌های همسایه شبیه‌تر برای پیش‌بینی می‌نماید. توصیف‌های حرکتی مذکور، همان مدل‌های انرژی بلوک‌ها خواهند بود. این مدل‌ها، بیانی از دینامیک هر بلوک ارائه خواهند کرد. نتایج شبیه‌سازی نیز نشان از بهبود عملکرد سیستم تخمین‌گر حرکت با استفاده از تکنیک مطرح شده در این پایان‌نامه دارند.

۱–۳– هدف از این پایان‌نامه

پایان‌نامه حاضر به منظور ارائه روشی برای انجام بهتر الگوریتم تطبیق بلوکی، تعریف و انجام گرفته است. همانطوری که از نام آن نیز پیداست، این پایان‌نامه دارای سه موضوع کلی است که عبارتند از تخمین حرکت در تصاویر متحرک، فیلتر کالمن و مدل انرژی.

فیلتر کالمن، خود روشی برای اجرای ایده‌ی تخمین حرکت می‌باشد. این فیلتر با استفاده از شباهت-های زمانی و مکانی موجود در بین نواحی تصویر، ابتدا با ارائه یک پیش‌بینی اولیه، سعی در سرعت و دقت بخشیدن به الگوریتم تطبیق بلوکی داشته و سپس با استفاده از روابط بازگشتی، نتیجه به دست آمده از پیش‌بینی‌اش را تصحیح می‌کند. مدل انرژی، در حقیقت توصیفی دینامیکی از تصویر ارائه می‌کند. این توصیف با استفاده از فیلترهایی که به حرکت در جهت‌های گوناگون حساس می‌باشند، استخراج می-گردد.

در این پایان‌نامه از توصیف انرژی استخراج‌شده برای تصویر به منظور افزایش دقت بخش پیش‌بینی فیلتر کالمن استفاده می‌گردد. شباهت‌های دینامیکی موجود در بخش‌های مختلف صحنه، کمک شایانی به اجرای هر چه بهتر الگوریتم فیلتر کالمن می‌کنند. فیلتر کالمن با استفاده از این توصیف، دقت پیش-بینی خود را ارتقا داده و کیفیت تصویر بازسازی شده را بالا می‌برد. نتایج شبیه‌سازی‌شده نیز موید همین ادعا می‌باشند.

۱–۴– فصل‌بندی پایان‌نامه

فصل جاری به ارائه کلیاتی از آنچه در این پایان‌نامه مطرح خواهد شد، اکتفا می‌کند. در فصل دوم، تخمین حرکت و روش تطبیق بلوکی مورد بحث و گفتگو قرار خواهند گرفت. در این فصل، به ارائه کلیات تخمین حرکت و برخی روش‌های موجود در تطبیق بلوکی خواهیم پرداخت. سپس در فصل سوم فیلتر کالمن توضیح داده خواهد شد. انواع فیلترهای کالمن در این فصل به همراه روابطشان به تفصیل شرح

داده شده‌اند. فصل ۴، اختصاص به مدل انرژی و چگونگی استخراج خصایص آن دارد. در این فصل نگاهی دقیق به فیلترهای زمانی و مکانی مورد استفاده در استخراج انرژی خواهیم داشت. روش مطرح شده در این پایان‌نامه، در فصل ۵ توضیح داده شده است. نتایج شبیه‌سازی نیز در فصل پایانی پایان‌نامه ارائه خواهند شد.

۱–۵– جمع‌بندی

تخمین حرکت در تصاویر متحرک همواره از اهمیت بالایی در علم بینایی ماشین برخوردار بوده است. این روش که به منظور فشرده‌سازی ویدیوها مورد استفاده قرار می‌گیرد، تا کنون تلاش‌های علمی بسیاری را متوجه خود ساخته است. پایان‌نامه حاضر نیز با هدف ارتقا کیفیت روش‌های مورد استفاده برای تخمین حرکت تعریف و انجام گرفته است. در این پایان‌نامه از ترکیب فیلتر کالمن و توصیف انرژی استخراج شده برای فریم‌های ویدیو استفاده شده است.

فصل ۲– تخمین حرکت

۲–۱– مقدمه

با رشد روزافزون تکنولوژی‌های چندرسانه‌ای[1]، نیاز به ذخیره‌سازی فایل‌های صوتی و تصویری نیز رشدی قابل توجه داشته است. در نتیجه، تخمین حرکت[2] در تصاویر متحرک نیز که با هدف کاهش حجم موردنیاز برای ذخیره‌سازی و ارسال اطلاعات تصویری پا به میان گذارده است، همواره به عنوان یکی از زمینه‌های پرطرفدار در میان محققان مطرح بوده و تکنیک‌های متعددی برای بهبود پارامترهای تعیین کننده‌ی آن ارائه گردیده است. در این میان، الگوریتم‌های تطبیق بلوکی[3] به دلیل نتایج کیفی مناسبی که در عین انجام محاسبات کم ارائه می‌کنند، به عنوان مهم‌ترین روش پیاده‌سازی ایده‌ی تخمین حرکت به شمار می‌روند. هر یک از این الگوریتم‌ها برای خود مزایا و معایبی داشته و دائما از سوی محققان به منظور افزایش کیفیت کارشان مورد بحث و بررسی قرار دارند. برای مقایسه بهتر روش‌های ارائه شده توسط محققان، برخی پارامترهای تعیین‌کننده برای نشان دادن کیفیت ارائه‌شده، هزینه محاسباتی برای انجام الگوریتم و نرخ بیت[4] مورد نیاز برای ارسال داده‌های حاصل‌شده از الگوریتم استفاده می‌گردند. هر چه الگوریتمی با استفاده از محاسبات و نرخ بیت ارسالی کمتری بتواند کیفیت بالاتری را ارئه نماید، بیشتر مورد توجه قرار خواهد گرفت. در ادامه به بررسی روش‌های موجود در مبحث تطبیق بلوکی خواهیم پرداخت.

۲–۲–الگوریتم‌های تطبیق بلوکی

۲–۲–۱–نگاهی کلی بر مبحث

با توجه به وجود افزونگی‌های زمانی و مکانی بسیار در توالی تصاویر متحرک، الگوریتم‌های تخمین حرکت همواره در پی کاهش حجم مورد نیاز برای ذخیره‌سازی و ارسال داده‌های تصویری به کمک حذف چنین افزونگی‌هایی بوده‌اند. ایده‌ی کلی اینست که با توجه به حجم بالای مورد نیاز برای ذخیره‌سازی و ارسال تک تک فریم‌های یک ویدیو و با در نظر داشتن شباهت‌های زمانی و مکانی موجود در بین فریم‌ها،

[1] Multimedia
[2] Motion Estimation
[3] Block Matching Algorithm
[4] Bit rate

به جای ذخیره‌سازی و ارسال تمام فریم‌ها بصورت جداگانه، برخی از فریم‌ها را ارسال و برخی دیگر را از روی آنها تولید می‌کنند (این بازتولید را جبران‌سازی[5] می‌نامند). روش‌های متعددی برای انجام این مهم مورد استفاده قرار گرفته‌اند. برخی از این روش‌ها نیز در استانداردهای ویدیویی در حال استفاده می‌باشند. همانطور که ذکر شد، هدف نهایی ارائه روشی است که بتواند با انجام عملیات کمتر و حفظ کیفیت در حد قابل‌قبول، منجر به فشرده‌سازی بیشتری گردد. در این میان، یک دسته مورد توجه در میان روش‌های تخمین حرکت، الگوریتم‌های تطبیق بلوکی می‌باشند. این دسته از الگوریتم‌ها با توجه به عملکرد خوبی که در برقراری ارتباط مناسب بین پارامترهای تعیین‌کننده تخمین حرکت برقرار کرده‌اند، بیشتر مورد توجه قرار می‌گیرند.

۲-۲-۲-نحوه کار الگوریتم‌های تطبیق بلوکی

همانطور که ذکر شد، هدف حذف افزونگی‌های مکانی و زمانی به منظور کاهش حجم داده‌های تصویری است. روش در نظر گرفته شده نیز ارسال برخی فریم‌ها و بازسازی فریم‌های دیگر از روی آنها می‌باشد. در نتیجه مهم‌ترین بخش کار که باید طراحی گردد، بازسازی یک فریم از روی فریم قبلی‌اش است[۱][۲]. بدین منظور، الگوریتم‌های تطبیق بلوکی ابتدا کل فریم موردنظر را به بلوک‌های هم‌اندازه غیرهم‌پوشان مستطیلی‌شکل تقسیم‌بندی می‌کنند. سپس سعی می‌کنند هر یک از این بلوک‌ها را توسط بردارهایی موسوم به بردار حرکت[6] از روی فریم قبل بازسازی نمایند. بلوک‌ها می‌توانند سایزی مانند ۴×۴، ۸×۴، ۸×۸، ۱۶×۱۶ (و یا حتی ابعاد دیگر) داشته باشند[۳]. هر چند اغلب طول و عرض بلوک‌های فوق-الذکر مساوی در نظر گرفته می‌شود. البته در اکثر موارد (مانند این پایان‌نامه) ابعاد این بلوک‌ها بصورت پیش‌فرض برابر با ۱۶×۱۶ انتخاب می‌شود. از طرفی بصورتی تجربی و با توجه به نرخ فریم[7] در نظر گرفته شده برای ویدیو، معمولا بیشینه سرعت حرکت‌های موجود در تصاویر متوالی برابر با ۷ پیکسل/فریم در نظر گرفته می‌شود. منظور از نرخ فریم، سرعت عکس‌برداری از صحنه می‌باشد. مثلا اگر مخاطب ویدیو انسان باشد، سعی می‌شود این نرخ در حدی باشد که از نظر بیننده پیوسته به نظر برسد. با توجه به توانایی دید انسان، مقدار متداول این نرخ در اکثر ویدیوها ۲۵ فریم بر ثانیه می‌باشد.

پس از تقسیم‌بندی بلوک‌ها، مرحله بعدی محاسبه بردارهای حرکت گفته شده است. بدین منظور، بلوکی را در نظر گرفته و آن را بلوک جاری B_C می‌نامند. سپس در فریم قبل و حول مختصات بلوک جاری، پنجره‌ای بزرگتر در نظر می‌گیرند. نام این پنجره، پنجره جستجو[8] می‌باشد. در واقع به دنبال

[5] Compensation
[6] Motion vector
[7] Frame rate
[8] Search window