



دانشگاه صنعتی شریف

دانشکده مهندسی برق

پایان نامه دکتری

گرایش الکترونیک

عنوان

کدینگ چند توصیفه ویدئو مبتنی بر ترکیب لایه های پایه و بهبود کدینگ

مقیاس پذیر و بهینه سازی سازگار با کانال

نگارش

محمد کاظمی ورنامخواستی

اساتید راهنما

آقای دکتر خسرو حاج صادقی

آقای دکتر شروین شیرمحمدی

بهمن ۱۳۹۰

به نام خدا

دانشگاه صنعتی شریف

دانشکده مهندسی برق

عنوان: کدینگ چند توصیفه ویدئو مبتنی بر ترکیب لایه‌های پایه و بهبود کدینگ مقیاس پذیر و بهینه‌سازی سازگار با کانال.

نگارش: محمد کاظمی و رنامخواستی

کمیته ممتحنین:

..... امضاء: استاد راهنما:
..... امضاء: استاد راهنما:
..... امضاء: ممتحن داخلی:
..... امضاء: ممتحن داخلی:
..... امضاء: ممتحن داخلی:
..... امضاء: استاد مدعو:
..... امضاء: استاد مدعو:

تقدیر و تشکر

نگارنده این رساله بر خود واجب می‌داند از افراد زیر که بدون کمکهای بی دریغ آنها قادر به اتمام دوره دکتری نبودم، کمال تشکر و قدردانی را بنمایم.

اساتید محترم راهنمای بنده، آقایان دکتر خسرو حاج صادقی و دکتر شروین شیرمحمدی که با راهنماییها، حمایتها و تشویقهای خودشان همواره موجب دلگرمی و اطمینان خاطر اینجانب بودند.

هیأت محترم داوران، آقای دکتر فرخ مروستی، آقای دکتر محمد فنبری، آقای دکتر عمادالدین فاطمی زاده، آقای دکتر محمود رضا هاشمی و آقای دکتر مسعود بابایی زاده که با پیشنهادهای ارزشمندشان موجب بهبود غنای علمی این رساله گشتند.

دوستانم در دانشکده مهندسی برق و همچنین دوستان همخوابگاهی که بطور مستقیم و غیر مستقیم مرا در انجام این رساله یاری رساندند.

خانواده گرامیم که با محبتهای بی پایانشان، همیشه و همه جا پشتوانه و موجب دلگرمی من بوده‌اند. این رساله را به آنها تقدیم می‌کنم شاید که بخش ناچیزی از زحماتشان را جبران کرده باشم.

چکیده

کدینگ چند توصیفه ویدئو¹(MDC)، یک روش برای انتقال ویدئو در محیطهای نویزی است. در MDC، چند توصیف از یک ویدئو تشکیل می‌شود که اولاً توصیفها جداگانه قابل دیکود شدن هستند و ثانیاً توصیفها مکمل یکدیگر هستند به گونه‌ایکه با افزایش تعداد توصیفها کیفیت ویدئوی دریافتی بالاتر خواهد بود. به کیفیت منتج از تمام توصیفها، کیفیت مرکزی و در غیر این صورت به آن کیفیت جانبی گفته می‌شود. در حالت کلی، بین کیفیت جانبی و کیفیت مرکزی بده بستان وجود دارد و MDC می‌تواند این بده بستان را بهتر کنترل کند، مناسب‌تر است.

در این رساله، یک روش جدید MDC معرفی شده است که در مقایسه با روشهای مرسوم، مصالحه بهتری بین کیفیت جانبی و کیفیت مرکزی برقرار می‌کند. به عبارت دیگر، با کیفیت مرکزی یکسان، کیفیت جانبی بالاتر و یا بطور معادل، با کیفیت جانبی یکسان، کیفیت مرکزی بالاتری فراهم می‌کند. این روش مبتنی بر ترکیب لایه‌های پایه و بهبود در کدینگ² CGS یا مقیاس‌پذیر SNR کار می‌کند و به این خاطر آنرا MLMDC³ نام نهاده‌ایم. در دیکودر مرکزی، لایه‌ها از یکدیگر جدا شده و مانند CGS دیکد می‌شوند و به این ترتیب کیفیت مرکزی برابر کیفیت دیکدینگ دو-لایه‌ای خواهد بود. در دیکودر جانبی، یک توصیف در دسترس نیست و لایه‌ها قابل تفکیک نیستند. به این منظور، از تئوری تخمین استفاده شده و لایه‌ها تا حد امکان جدا می‌شوند. MLMDC دو توصیفه و چهار توصیفه در JM16.0، نرم افزار مرجع انکودر H.264/AVC، پیاده سازی شده است. نتایج عملی نشان می‌دهد که برای ویدئوهایی که محتوای پایینی (بافت و حرکت) ندارند، MLMDC در حالیکه کیفیت مرکزی در حدود سایر روشها

¹ Multiple Description Coding

² Coarse-Grain Scalable

³ Mixed Layer MDC

دارد، کیفیت جانبی به مراتب بهتری فراهم می‌کند. همچنین برای انتقال ویدئو روی کانالهای نویزی، با در نظر گرفتن ریزش پاکت، MLMDC بویژه برای کدینگ چهار توصیفه، کیفیت متوسط بالاتری ارائه می‌دهد.

در شرایطی که کانال بسیار نویزی است، به کیفیت جانبی بالا بیشتر نیاز است، حال آنکه برای کانالهای کمتر نویزی، اغلب اوقات همه توصیفهها در دسترس هستند و کیفیت مرکزی بالاتر مهم‌تر است. لذا برای داشتن بهترین کیفیت در شرایط مختلف کانال، لازم است که بهینه‌سازی صورت گیرد. به این منظور ابتدا مدلی برای پیدا کردن اعوجاج انتها-تا-انتهای^۱ ارائه شده است. برای پیدا کردن اعوجاج انتها-تا-انتهای، علاوه بر اعوجاج کوانتیزیشن، اعوجاج انتقال هم بایستی محاسبه شود که بیشتر ناشی از انتشار خطا^۲ است. انتشار خطا یا دررفت^۳ که از ناهمخوانی مرجع پیش‌بینی در انکودر و در دیکودر نتیجه می‌شود، بصورت تحلیلی بدست آورده شده است. مدل استخراج شده مدلی کلی است و برای سایر MDC های حوزه فرکانس قابل استفاده است. با آزمایشهای عملی، دقت مدل سنجیده شده و سپس برای طراحی بهینه MLMDC بکار گرفته شده است. نتایج عملی، عملکرد تابع بهینه‌ساز را تایید می‌کند و همچنین کیفیت بالاتر MLMDC نسبت به MDC های مرسوم، هنگامیکه همه بهینه طراحی شده‌اند، را نشان می‌دهد.

واژه‌های کلیدی: کدینگ چند توصیفه، ارسال ویدئو، ریزش بسته‌های داده ویدئو، انتشار خطا، اعوجاج انتها-تا-

انتهای، بهینه‌سازی مطابق با کانال، H.264/AVC

¹ End-to-End Distortion

² Error Propagation

³ Drift

فهرست مطالب

۱	۱. مقدمه
۱	۱.۱. مقدمه
۲	۲.۱. کدینگ چند توصیفه
۲	۱.۲.۱. کاربرد MDC
۵	۲.۲.۱. مفهوم پایه ای MDC
		3.2.1. دریافت ۷
۸	۳.۱. معرفی روش پیشنهادی و بهینه سازی آن
۸	۱.۳.۱. پیش زمینه: کدینگ مقیاس پذیر
۹	۲.۳.۱. روش پیشنهادی
۱۰	4.1. مرور مفاهیم مورد استفاده از استاندارد H.264/AVC
۱۱	۱.۴.۱. تقسیم بندی فریم به ماکرو بلاک
۱۱	۲.۴.۱. تخمین/جبران حرکت یا پیش بینی بین فریمی
۱۳	۳.۴.۱. پیش بینی داخل فریمی
۱۴	۴.۴.۱. کدینگ تبدیل کردن
۱۵	۵.۴.۱. کوانتیزیشن
۱۶	۶.۴.۱. آنتروپی کدینگ
۱۷	۵.۱. ساختار رساله
۱۸	2. مروری بر روشهای MDC
۱۹	۱.۲. دسته بندی طرحهای مختلف MDC
۱۹	1.1.2. حوزه مکانی
۲۴	2.1.2. MDC حوزه زمان
۲۷	۳.۱.۲. حوزه فرکانس
۳۶	4.1.2. MDC حوزه فشرده شده
۳۷	5.1.2. MDC حوزه هایبیرید
۳۸	6.1.2. روشهایی که مبتنی بر تقسیم بندی نیستند

۳۹	دسته بندی روشهای MDC بر اساس کارایی آنها	2.2.
۳۹	قابلیت تنظیم افزونگی	1.2.2.
۳۹	پیچیدگی محاسباتی	2.2.2.
۴۰	امکان افزایش تعداد توصیفها	3.2.2.
۴۳	MDC کدام حوزه؟	۳.۲
۴۵	۳. روش MLMDC	
۴۵	مقدمه	۱.۳
۴۶	معرفی روش پیشنهادی	2.3.
	انکودر ۴۶	۱.۲.۳
۴۹	دیکودر مرکزی	۲.۲.۳
۴۹	دیکودر جانبی	۳.۲.۳
۵۰	اعوجاج و نرخ در MLMDC	3.3.
۵۰	پیش زمینه	۱.۳.۳
۵۲	اعوجاج جانبی	۲.۳.۳
۵۹	اعوجاج مرکزی	۳.۳.۳
۶۰	افزونگی نرخ ناشی از ضریب بهبود	4.3.3.
۶۳	مقایسه عملکرد MLMDC	۴.۳
۶۹	نتایج عملی	۵.۳
۸۱	۴. بهینه سازی MLMDC سازگار با کانال	
۸۱	مقدمه	۱.۴
۸۳	کارهای مرتبط	2.4.
۸۵	مدل پیشنهادی برای اعوجاج انتها-تا-انتها	3.4.
۹۰	پیاده سازی مدل و تشکیل تابع هدف	4.4.
۹۳	تابع ناهمخوانی برای MDC های مورد مطالعه	1.4.4.
۹۷	نتایج شبیه سازی	2.4.4.
۱۰۰	نتایج عملی	5.4.
۱۰۰	اعوجاج انتها-تا-انتها	1.5.4.
۱۰۷	نتایج بهینه سازی	2.5.4.
۱۱۲	۵. جمع بندی و کارهای آینده	

۱.۵. نوآوریها..... ۱۱۲

۱.۱.۵. در فصل دوم..... ۱۱۲

۲.۱.۵. در فصل سوم..... ۱۱۲

۳.۱.۵. در فصل چهارم..... ۱۱۳

۲.۵. کارهای آینده..... ۱۱۴

۱.۲.۵. پیشنهاد اول..... ۱۱۴

۲.۲.۵. پیشنهاد دوم..... ۱۱۴

۳.۲.۵. پیشنهاد سوم..... ۱۱۵

۴.۲.۵. پیشنهاد چهارم..... ۱۱۵

۵.۲.۵. پیشنهاد پنجم: تخمین حرکت چند توصیفه..... ۱۱۵

۶. ضمیمه..... ۱۱۸

۱.۶. ضمیمه فصل سوم..... ۱۱۸

۲.۶. ضمیمه فصل چهارم..... ۱۱۹

۷. مراجع..... ۱۲۱

فهرست جداول

فصل اول:

جدول ۱-۲ دسته بندی و مقایسه MDC ها از نظر حوزه و کارایی ۴۱

فصل چهارم:

جدول ۱-۴ خروجی بهینه ساز برای GOP دوم از سکانس Mobile ۹۸

فهرست اشکال

فصل اول:

شکل ۱-۱ انتقال ویدئو در محیطهای نویزی با استفاده از شیوه MDC ۴

شکل ۱-۲ بلوک دیاگرام MDC ۶

شکل ۱-۳ نحوه انتشار خطا از فریم چهارم به فریمهای بعدی ویدئو [15] ۸

شکل ۱-۴ انکودر مقیاس پذیر SNR یا CGS ۹

شکل ۱-۵ نحوه ترکیب لایه ها در MLMDC ۱۰

شکل ۱-۶ بلوک دیاگرام H.264/AVC [16] ۱۱

فصل دوم:

شکل ۱-۲ MDC با شیوه زیرنمونه برداری چند فازه spatial ۲۰

شکل ۲-۲ اضافه کردن صفر قبل از زیرنمونه برداری، چپ: یک بعدی و راست: دو بعدی ۲۰

شکل ۲-۳ تقریب زدن توصیفهای از دست رفته و قتیکه (a) فقط یک توصیف رسیده باشد و (b) توصیفهای اول و دوم رسیده باشند [21].

..... ۲۱

- شکل ۲-۴ پیش‌بینی یک زیرتصویر از زیرتصویر دیگر ۲۲
- شکل ۲-۵ MDC با استفاده از زیرنمونه برداری چند فازه غیرخطی (a) تصویر اصلی (b) بعد از تبدیل غیرخطی (c) وقتی هر چهار توصیف رسیده باشند و (d) فقط یک توصیف رسیده باشد [26] ۲۳
- شکل ۲-۶ MDC چند نرخه حوزه spatial [27] ۲۳
- شکل ۲-۷ بلوک دیاگرام پایه‌ای MDC حوزه زمان ۲۴
- شکل ۲-۸ بلوک دیاگرام MDC حوزه زمان با شیوه نرخ فریم تطبیق پذیر [36] ۲۷
- شکل ۲-۹ ساده‌ترین MDSQ، دو خط بالایی: دیکودر جانبی و خط پایینی: دیکودر مرکزی ۲۸
- شکل ۲-۱۰ یک نمونه جدول تخصیص اندیس در MDSQ ۲۹
- شکل ۲-۱۱ جدول تخصیص اندیس در MDSQ اصلاح شده ۳۰
- شکل ۲-۱۲ چپ: بخش بندی فریم به بلوکهای هم پوشان و راست: تخصیص بلوکها به توصیفها [44] ۳۱
- شکل ۲-۱۳ بلوک دیاگرام MDTC [51] ۳۳
- شکل ۲-۱۴ MDC چند نرخه حوزه فرکانس [55] (a) انکودر و (b) دیکودر ۳۴
- شکل ۲-۱۵ در توصیف اول، MBهای گروه A در نرخ بالا و MBهای گروه B در نرخ پایین کد می شوند، و برعکس در توصیف دوم [56] ۳۵
- شکل ۲-۱۶ MDC به کمک اسلایسهای افزونگی [58] ۳۵
- شکل ۲-۱۷ MDC-FEC با تعداد مساوی توصیفها و لایه‌های SVC ۳۶
- شکل ۲-۱۸ MDC حوزه هایبرید: spatial و فرکانس [65] ۳۷
- فصل سوم:**

- شکل ۳-۱ بلوک دیاگرام روش پیشنهادی (الف) انکودر (ب) دیکودر مرکزی و (ج) دیکودر جانبی ۴۷
- شکل ۳-۲ توزیع سیگنالهای پایه و بهبود، $Q = 16$ ، $\lambda = 1/10$ و $c=2$ (الف) BC (ب) EC و Case I و $xb = 1$ (ج) EC و Case II ۵۳
- شکل ۳-۳ منحنی تخمین برای Case I و (الف) $c = 1.5$ (ب) $c = 3.5$ ۵۵
- شکل ۳-۴ منحنی تخمین برای Case II و (الف) $c = 1.5$ (ب) $c = 3.5$ ۵۶
- شکل ۳-۵ اعوجاج تخمین برای Case I و Case II، $Q=16$ و (الف) $\lambda = 0.05$ و (ب) $\lambda = 0.2$ ۵۷

شکل ۳-۶ اعوجاج جانبی برای Case I و Case II (الف) $\lambda = 1/20$ و $Q = 16$ (ب) $\lambda = 1/5$ و $Q = 16$ (ج) $\lambda = 1/20$ و $Q = 32$ و $\lambda = 1/5$ (د) $Q = 32$ و $\lambda = 1/5$ ۵۸

شکل ۳-۷. کیفیت مرکزی بر حسب c برای (الف) $Q = 16$ و (ب) $Q = 32$ ۶۰

شکل ۳-۸ ERn بر حسب c برای چند مقدار λQ ۶۲

شکل ۳-۹ $R2Layers - R1Layer$ بر حسب c برای چند مقدار λQ ۶۲

شکل ۳-۱۰ منحنی‌های RD برای Splitting ، Duplication ، MDTC و MLMDC برای چند مقدار $(\lambda 1, \lambda 2)$ چپ: rate/side-distortion و وسط: rate/central-distortion و راست: redundancy-rate-distortion و $Q0 = 16$ ۶۸

شکل ۳-۱۱ MLMDC: مقادیر cb بر حسب λQ ۶۹

شکل ۳-۱۲ منحنی‌های RD برای الگوریتم‌های Dup ، Dup4Sp ، Dup4Sp ، QMDC و MLMDC و سکانس Akiyo (الف) دیکودر جانبی و کدینگ دو توصیفه و (ب) دیکودر مرکزی و کدینگ دو توصیفه (ج) دیکودر جانبی و کدینگ چهار توصیفه و (د) دیکودر مرکزی و کدینگ چهار توصیفه ۷۲

شکل ۳-۱۳ منحنی‌های RD برای الگوریتم‌های Dup ، Dup4Sp ، Dup4Sp ، QMDC و MLMDC و سکانس Foreman (الف) دیکودر جانبی و کدینگ دو توصیفه و (ب) دیکودر مرکزی و کدینگ دو توصیفه (ج) دیکودر جانبی و کدینگ چهار توصیفه و (د) دیکودر مرکزی و کدینگ چهار توصیفه ۷۳

شکل ۳-۱۴ منحنی‌های RD برای الگوریتم‌های Dup ، Dup4Sp ، Dup4Sp ، QMDC و MLMDC و سکانس Mobile (الف) دیکودر جانبی و کدینگ دو توصیفه و (ب) دیکودر مرکزی و کدینگ دو توصیفه (ج) دیکودر جانبی و کدینگ چهار توصیفه و (د) دیکودر مرکزی و کدینگ چهار توصیفه ۷۳

شکل ۳-۱۵ منحنی نرخ-اعوجاج برای ویدئوی Akiyo (الف) کدینگ دو توصیفه و $PLR=5\%$ (ب) کدینگ دو توصیفه و $PLR=20\%$ (ج) کدینگ چهار توصیفه و $PLR=5\%$ (الف) کدینگ چهار توصیفه و $PLR=20\%$ ۷۶

شکل ۳-۱۶ منحنی نرخ-اعوجاج برای ویدئوی Foreman (الف) کدینگ دو توصیفه و $PLR=5\%$ (ب) کدینگ دو توصیفه و $PLR=20\%$ (ج) کدینگ چهار توصیفه و $PLR=5\%$ (الف) کدینگ چهار توصیفه و $PLR=20\%$ ۷۷

شکل ۳-۱۷ منحنی نرخ-اعوجاج برای ویدئوی Mobile (الف) کدینگ دو توصیفه و $PLR=5\%$ (ب) کدینگ دو توصیفه و $PLR=20\%$ (ج) کدینگ چهار توصیفه و $PLR=5\%$ (الف) کدینگ چهار توصیفه و $PLR=20\%$ ۷۸

شکل ۳-۱۸ تاثیر تخمین زدن برای (الف) $c = 1.3$ و $QP = 20$ و (ب) $c = 3.2$ و $QP = 32$ ۸۰

فصل چهارم:

شکل ۴-۱ منحنی نرخ-اعوجاج در DupSp برای مقادیر بهینه و مقادیر نوعی پارامتر MDC (الف) $P = 0.05$ ، (ب) $P = 0.10$ و (ج) $P = 0.20$

۹۹

شکل ۴-۲ منحنی نرخ-اعوجاج در QMDC برای مقادیر بهینه و مقادیر نوعی پارامتر MDC (الف) $P = 0.05$ ، (ب) $P = 0.10$ و (ج) $P = 0.20$

۹۹

شکل ۴-۳ منحنی نرخ-اعوجاج در MLMDC برای مقادیر بهینه و مقادیر نوعی پارامتر MDC (الف) $P = 0.05$ ، (ب) $P = 0.10$ و (ج) $P =$

۰.۲۰ ۹۹

شکل ۴-۴ سکانس Mobile و $QP0 = 20$ و $PLR = 0.05$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۲ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۵ سکانس Mobile و $QP0 = 20$ و $PLR = 0.20$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۲ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۶ سکانس Mobile و $QP0 = 28$ و $PLR = 0.05$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۳ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۷ سکانس Mobile و $QP0 = 28$ و $PLR = 0.20$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۴ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۸ سکانس Foreman و $QP0 = 20$ و $PLR = 0.05$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۴ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۹ سکانس Foreman و $QP0 = 20$ و $PLR = 0.20$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۵ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۱۰ سکانس Foreman و $QP0 = 28$ و $PLR = 0.05$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۶ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۱۱ سکانس Foreman و $QP0 = 28$ و $PLR = 0.20$ (الف) $\beta = 0$ و $P = \{P1\}$ ، (ب) $\beta = 0$ و $P = \{P2\}$ ، (ج) $\beta = 0.1$ و

۱۰۶ $P = \{P2\}$ و $\beta = 0.1$ (د) $P = \{P1\}$

شکل ۴-۱۲ سکانس Mobile و $\beta = 0$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و

۱۰۹ QMDC و MLMDC مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$

شکل ۴-۱۳ سکانس Mobile و $\beta = 0.1$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و

۱۰۹ QMDC و MLMDC مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$

شکل ۴-۱۴ سکانس Foreman و $\beta = 0$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و

QMDC و MLMDC مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$ ۱۱۰

شکل ۴-۱۵ سکانس Foreman و $\beta = 0.1$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و

QMDC و MLMDC مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$ ۱۱۰

۱. مقدمه

۱.۱. مقدمه

کدینگ چند توصیفه (MDC^۱) ویدئو به منظور انتقال ویدئو روی کانالهای نویزی استفاده می‌شود. سیگنالهای چندرسانه‌ای و بویژه ویدئو به این دلیل که افزونگی دارند اغلب بصورت فشرده شده مورد استفاده قرار می‌گیرند، چه برای ذخیره‌سازی و نگهداری در حافظه‌ها و چه به هنگام ارسال. اما ارسال ویدئوی فشرده‌شده در شبکه نیاز به تمهیداتی دارد، چونکه با فشرده‌سازی حساسیت ویدئو نسبت به خطا بیشتر می‌شود. علت این است که در سیگنال فشرده‌شده هر بیت نماینده چندین بیت در داده نافرده و یا بازنده است و لذا چنانچه یک بیت از داده فشرده شده دچار خطا شود، آن چندین بیت را در داده بازنده نخواهیم داشت. علاوه بر این به دلیل شیوه‌های استفاده شده در فشرده‌سازی، نداشتن یک بیت دیکد کردن بیت‌های دیگر که سالم رسیده اند را هم تحت تاثیر قرار می‌دهد. لذا داده فشرده شده نسبت به خطا حساسیت بالایی دارد. می‌توان گفت که بین نرخ فشرده‌سازی و حساسیت نسبت به خطا، بده-بستان وجود دارد؛ به عبارت دیگر باید بین نرخ بیت و خطاپذیری^۲ داده فشرده شده مصالحه کرد. این مهم، ارسال ویدئو را در محیط‌های نویزی، برای داشتن بهترین کیفیت در گیرنده، به یک مسئله برای مهندسان و پژوهشگران این حوزه تبدیل

^۱ Multiple Description Coding

^۲ Error Resiliency

کرده است. لذا روشها و تمهیدات مختلفی، به منظور مقابله با از دست دادن داده‌های فشرده شده به هنگام انتقال آنها، اندیشیده شده است که یکی از آنها MDC می‌باشد [1] و [2].

در ادامه با مرور مختصری بر روشهایی که برای ارسال ویدئو روی کانالهای نویزی استفاده می‌شوند، MDC را معرفی کرده و آنرا نسبت به سایر روشها ارزیابی می‌کنیم. مفاهیم پایه‌ای و دریافت¹ در MDC در این بخش توضیح داده شده، و در بخش سوم مرور مختصری بر روش پیشنهادی خود خواهیم داشت. در بخش چهارم، مروری مختصر بر نحوه فشرده‌سازی در استاندارد H.264/AVC خواهیم داشت. در این بخش فقط آن مفاهیمی از فشرده ساز H.264/AVC که در این رساله با آنها سر و کار خواهیم داشت، توضیح داده می‌شود، چونکه این استاندارد بسیار مفصل است و ارائه همه جزئیات آن ضرورت ندارد. در زیر بخش پایانی هم ساختار رساله گفته خواهد شد.

۲.۱. کدینگ چند توصیفه

۱.۲.۱. کاربرد MDC

ارسال ویدئو در محیطهای نویزی مسئله‌ای است که در دو دهه اخیر مورد توجه پژوهشگران قرار گرفته است. به این منظور، سه رویکرد وجود دارد: استفاده از پروتکل‌های با قابلیت ارسال مجدد (ARQ²)، استفاده از کدهای کدینگ کانال (FEC³) و یا استفاده از شیوه‌های کدینگ خطا پذیر (ERC⁴). دو روش اول یعنی ARQ و FEC در لایه کانال انجام می‌گیرند ولی ERC در لایه منبع صورت گیرد.

¹ Drift

² Automatic Repeat Request

³ Forward Error Correction

⁴ Error Resilient Coding

ARQ، همانطور که از نامش می‌توان حدس زد بر اساس ارسال مجدد داده‌هایی که سالم دریافت نشده‌اند کار می‌کند که مشخصاً برای شبکه‌هایی که قابلیت ارسال پس‌خورد به فرستنده را دارا نباشند قابل استفاده نخواهد بود. همچنین برای کاربردهای بی‌درنگ^۱ که تاخیر مسئله مهمی است و یا مواردی مانند ارسال از یک فرستنده به چندین گیرنده (multicast و broadcast) راه حل مناسبی نمی‌باشد. در روش FEC، اگر که خطای رخ داده بیشتر از آنی باشد که هنگام طراحی کدینگ در نظر گرفته شده است عملاً FEC استفاده شده فایده‌ای نخواهد داشت و داده‌ای بازیافت نمی‌شود. لذا در این حالت مجبوریم طراحی را بر اساس بیشترین خطای ممکن انجام دهیم که این به نوبه خود باعث کاهش کارایی این روش در محیط‌هایی می‌شود که میزان خطا کم و زیاد می‌گردد. ناگفته نماند که تغییرات درصد خطا در کانال مسئله‌ای معمول در شبکه‌ها است.

و اما از روشهای ERC می‌توان به Intra-refreshment، FMO^۲ که در استاندارد ویدئو H.264/AVC گنجانده شده‌اند اشاره کرد که اگر چه موثر هستند ولی برای کانالهای با خطاهای زیاد (مثلاً بیش از 5%) معمولاً کافی نیستند. در مواردی که شبکه به اصطلاح نویزی‌تر است یک روش کارآمد کدینگ چند توصیفه یا MDC است. در MDC چند استریم جداگانه از منبع ساخته می‌شود که این استریمها اولاً هر کدام جداگانه قابل دیکد کردن هستند و ثانیاً مکمل یکدیگر می‌باشند؛ به این معنی که دو استریم دریافتی کیفیت بالاتری را ارائه می‌دهند.

یک روش دیگر برای کدینگ چند استریمه، کدینگ مقیاس‌پذیر (SVC^۳) و یا کدینگ لایه‌ای^۴ است که منبع و یا ویدئو به یک لایه پایه و یک یا چند لایه بهبود کد می‌شود [3]^۵. با افزایش تعداد لایه‌های دریافتی، کیفیت

^۱ Real time

^۲ Flexible Macroblock Ordering

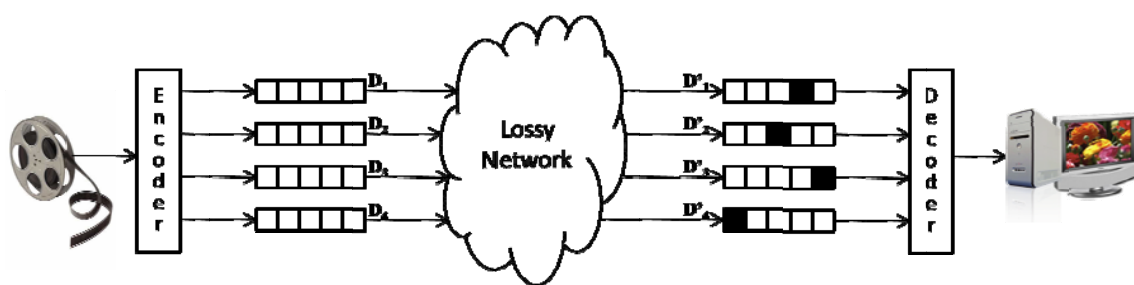
^۳ Scalable Video Coding

^۴ Layered Coding

^۵ مخترع این روش دانشمند ایرانی، پروفیسور محمد قنبری هستند.

هم افزایش می‌یابد ولی لایه‌ها به ترتیب روی یکدیگر قرار می‌گیرند. مثلاً لایه بهبود دوم در صورتی قابل استفاده است که لایه‌های قبلی یعنی لایه پایه و لایه بهبود اول دریافت شده باشند. به دلیل همین محدودیت، SVC نسبت به MDC در نرخ خطاهای بالا کارایی کمتری دارد همچنانکه در مقاله‌های [8]-[11] نشان داده شده است.

MDC معمولاً با ارسال چند مسیره (MPT¹) همراه است که در این صورت هر توصیف جداگانه روی یک مسیر فرستاده می‌شود. اگر یک بسته داده در یک توصیف دچار خطا شود، بسته معادل دیگری در یک توصیف دیگر وجود دارد که می‌تواند بکار آید و ویدئو قابل بازسازی است. این مسئله برای حالتیکه چهار استریم MDC داشته باشیم در شکل ۱-۱ نشان داده شده است. با استفاده از MDC، فقط در حالتی قادر به بازسازی ویدئو نیستیم که تمام مسیرها همزمان دچار خطا شوند و هیچکدام از توصیفها به گیرنده نرسد که احتمال رخداد آن بسیار کم است و لذا MDC همراه با MPT روشی کارآمد برای ارسال ویدئو در محیطهای نویزی محسوب می‌شود.



شکل ۱-۱ انتقال ویدئو در محیطهای نویزی با استفاده از شیوه MDC

علاوه بر این، با در دست داشتن چندین توصیف از ویدئو می‌توان چندین پهنای باند را نیز پوشش داد. با افزایش/کاهش تعداد استریمهای ارسالی، پهنای باند مصرف شده افزایش/کاهش می‌یابد که به این ترتیب

¹ Multi-Path Transmission

امکان استفاده از پهنای باند در اختیار به فرستنده داده می‌شود. از این ویژگی در شبکه‌های ناهمگن^۱ استفاده می‌شود. این ویژگی به همراه خط‌پذیری و همچنین اینکه توصیفها جداگانه قابل دیکد کردن هستند، MDC را به یک گزینه عملی انتقال ویدئو در شبکه‌های P2P نیز تبدیل کرده است [4]؛ مانند Hotstreaming [5]، Splitstreaming [6] و CoopNet [7].

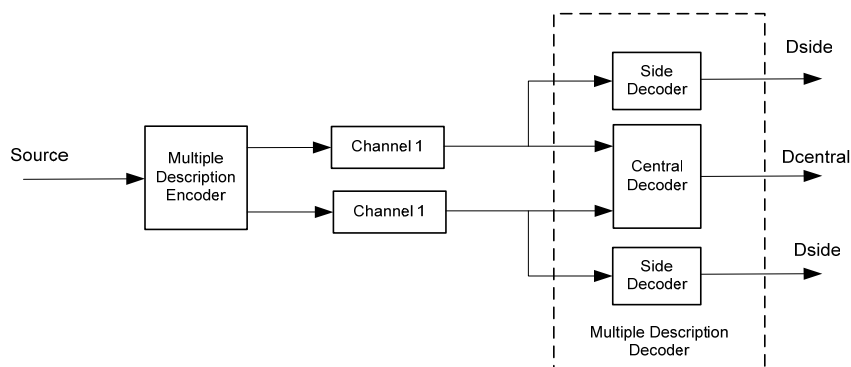
۲.۲.۱. مفهوم پایه ای MDC

در دهه ۱۹۷۰، MDC برای نخستین بار جهت مخابره تلفنی روی یک شبکه سوئیچ مداری^۲ مورد استفاده قرار گرفت. قبل از آن، برای افزایش قابلیت اطمینان سیستم، یک خط ارتباطی بصورت آماده بکار ولی خارج از مدار تعبیه شده بود که این خط ارتباطی در مواقعی که خط اصلی دچار اشکال می‌شد وارد مدار و مورد استفاده قرار می‌گرفت. ایده MDC همین‌جا بود که از هر دو خط همزمان استفاده شود؛ به گونه‌ای داده‌ها بین دو خط تقسیم شوند که اگر هر دو خط کار کردند کیفیت بالاتری داشته باشیم و در صورتیکه یک خط دچار اشکال شد، داده دریافتی از خط سالم بطور مستقل قابل استفاده باشد. مسئله طراحی بهینه این نحوه کدینگ در 1979 IEEE Information Theory Workshop توسط Gersho, Wistenhausen, Wolf, Wyner مطرح شد. اگر توصیف اول با نرخ بیت R_1 و اعوجاج D_1 و توصیف دوم با نرخ بیت R_2 و اعوجاج D_2 باشد، اعوجاج حاصل از دو توصیف D_0 است؛ مسئله، پیدا کردن نقاط بهینه $\{R_1, R_2, D_1, D_2, D_0\}$ بود که در مقاله‌های [12]-[14] به آن پرداخته شد. به این ترتیب بود که MDC مورد توجه پژوهشگران و همچنین مهندسان قرار گرفت.

¹ Heterogeneous Networks

² circuit-switched

شکل ۱-۲ بلوک دیاگرام پایه‌ای MDC را نشان می‌دهد. در این شکل حالت دو-توصیفه نشان داده شده ولی در حالت کلی تعداد بیشتر توصیفها هم امکان‌پذیر است. در دیکودر، اگر همه توصیفها رسیده باشند توسط "دیکودر مرکزی"^۱ دیکود می‌شوند و اگر نه توسط "دیکودر جانبی"^۲ این کار صورت می‌گیرد. در حالت اول به کیفیت (اعوجاج) حاصل کیفیت (اعوجاج) مرکزی و در حالت دوم به آن کیفیت (اعوجاج) جانبی گفته می‌شود. اگر همه توصیفها دچار خطا شده باشند تصویر با شیوه‌های اختفای خطا^۳ بازیافت می‌شود. اختفای خطا یعنی اینکه بجای داده از بین رفته داده مشابهی جایگزین شود به گونه‌ایکه با چشم کمتر قابل تشخیص باشد. به عبارت دیگر، قسمتهایی از تصویر که داده‌های مربوط به آنها دریافت نشده است به نحوی وصله خورده می‌شوند. البته، در دیکودر جانبی هم می‌توان از اختفای خطا استفاده کرد و اثر توصیفهای نرسیده را کاهش داد.



شکل ۱-۲ بلوک دیاگرام MDC

ویژگی خطا پذیری MDC در قبال افزونگی نرخ بیت بدست می‌آید. برای اینکه هر توصیف جداگانه قابل کد کردن باشد نیاز است که اطلاعات پایه‌ای را داشته باشد. این اطلاعات پایه‌ای در توصیفها مشترک هستند و در دیکودر مرکزی فقط یک نسخه از آنها استفاده می‌شود و بقیه افزونگی هستند. علاوه بر این، برای اینکه

^۱ Central decoder

^۲ Side decoder

^۳ concealment

بتوان توصیفهای خراب شده را به کمک توصیفهای سالم تا حدی بازیافت نمود مقداری از اطلاعات هر توصیف در توصیفهای دیگر گنجانده می‌شود؛ این اطلاعات فقط در دیکودر جانبی مورد استفاده قرار می‌گیرد و در دیکودر مرکزی اضافه است. افزونگی اگر چه کیفیت جانبی را بهبود می‌دهد، ولی در یک نرخ مشخص باعث کاهش کیفیت مرکزی می‌شود. در واقع، بین کیفیت جانبی و کیفیت مرکزی، به واسطه افزونگی، بده بستان وجود دارد. در کانالهای با نویز بالا چونکه احتمال دیکودینگ جانبی بالاست، به کیفیت جانبی بالاتر و در نتیجه افزونگی بیشتری نیاز است. ولی در شرایطی که نرخ خطا کم است و احتمال دیکودینگ جانبی پایین است این افزونگی نباید زیاد باشد.

۳.۲.۱. دریافت

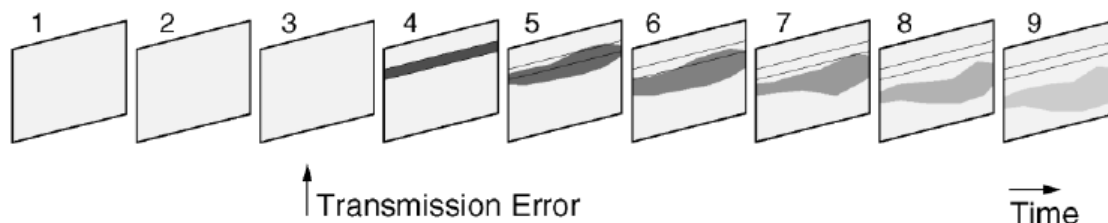
در ارسال ویدئو روی کانالهای نویزی با پدیده دیگری به نام دریافت^۱ و یا انتشار خطا^۲ مواجه هستیم که از ناهمخوانی مرجع پیش‌بینی در انکودر و دیکودر ناشی می‌شود. استانداردهای فشرده‌سازی ویدئو از شیوه پیش‌بینی استفاده می‌کنند که بازدهی فشرده‌سازی بالایی دارد اما باعث وابستگی بین فریمها می‌شود. وقتی که فریم مرجع بطور صحیح بازسازی نمی‌شود، فریم رجوع کننده هم بطور صحیح بازسازی نمی‌شود، این فریم جدید به نوبه خود مرجع فریمهای بعدی قرار خواهد گرفت و به این ترتیب انتشار خطا رخ می‌دهد، این روند که در شکل ۱-۳ نشان داده شده است تا انتهای GOP^۳ ادامه دارد. این پدیده مقوله‌ای مهم در ارسال ویدئو در شرایط نویزی، هم در SDC^۴ و هم در MDC است.

^۱ Drift

^۲ Error Propagation

^۳ Group Of Pictures

^۴ Single Description Coding



شکل ۱-۳ نحوه انتشار خطا از فریم چهارم به فریمهای بعدی ویدئو [15]

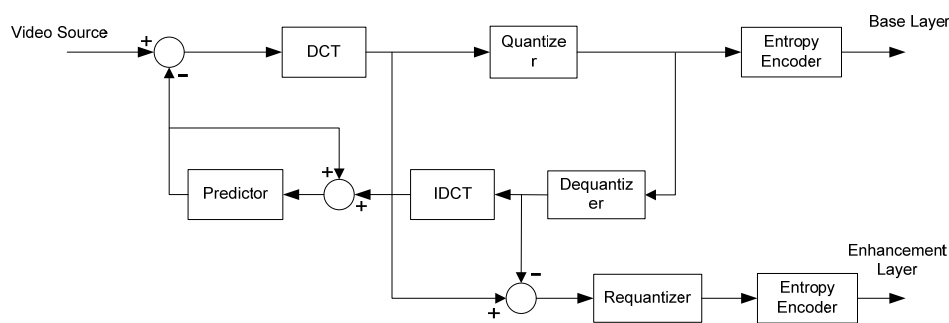
۳.۱ معرفی روش پیشنهادی و بهینه‌سازی آن

۱.۳.۱ پیش زمینه: کدینگ مقیاس پذیر^۱

در کدینگ مقیاس پذیر، همانطور که قبلاً هم توضیح داده شد ویدئو را لایه‌ای کد می‌کنند، یک لایه پایه و یک یا چند لایه بهبود. لایه‌ها به ترتیب روی یکدیگر قرار می‌گیرند و با افزایش تعداد لایه‌های دیکد شده، کیفیت حاصل هم افزایش می‌یابد. لایه پایه می‌تواند ویدئو با تصاویر کوچکتر باشد که با اضافه شدن لایه بهبود، اندازه تصاویر بزرگتر می‌گردد. به این نوع مقیاس پذیری، "مقیاس پذیری spatial" می‌گویند. در "مقیاس پذیری زمانی"، لایه پایه نرخ فریم پایینی دارد که با دیکد و اضافه شدن لایه بهبود، نرخ فریم زیاد می‌شود. و اما در مقیاس پذیری نوع سوم، لایه پایه لایه‌ای کیفیت پایین است که با اضافه شدن لایه بهبود، کیفیت تصویر بهتر و یا نویز کوانتیزیشن کمتری به ما می‌دهد. به این نوع مقیاس پذیری، مقیاس پذیری SNR و یا CGS^۲ گفته می‌شود. شکل ۱-۴ نحوه تشکیل لایه‌ها را در CGS نشان می‌دهد.

¹ Scalable coding

² Coarse-Grain Scalability



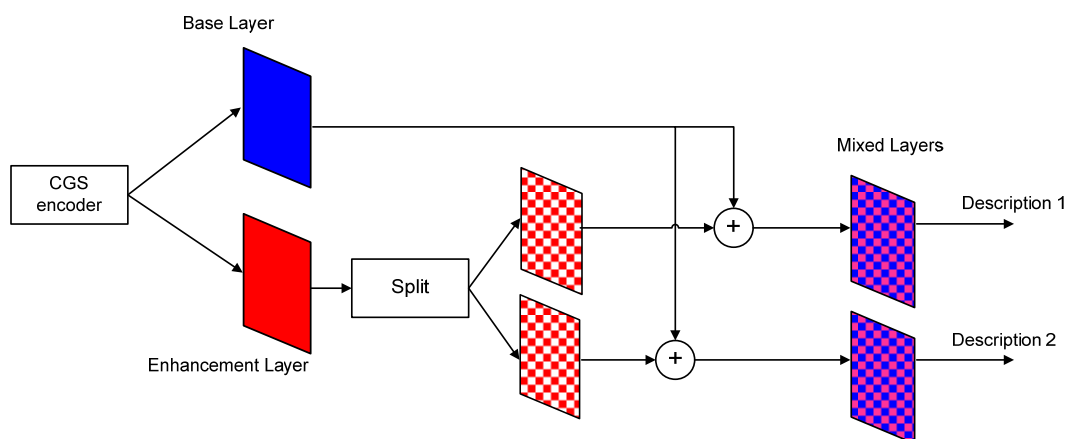
شکل ۱-۴ انکودر مقیاس پذیر SNR یا CGS

۲.۳.۱. روش پیشنهادی

روش MDC پیشنهاد شده بر مبنای ترکیب لایه‌های پایه و بهبود در CGS است و به همین خاطر آنرا MLMDC نامیده‌ایم. همانطور که شکل ۱-۵ هم نشان می‌دهد در MLMDC لایه بهبود به دو لایه تقسیم شده و هر قسمت با لایه پایه ترکیب شده و به عنوان یک توصیف ارسال می‌گردد. در دیکودر مرکزی، ابتدا لایه‌ها از یکدیگر جدا شده، هر کدام با Q مخصوص به خود دیکواتتایز شده و سپس روی همدیگر دیکد می‌شوند، درست مانند دیکودینگ دو لایه ای در CGS که طبیعتاً کیفیت بالاتری نسبت به یک لایه خواهد داشت. و اما در دیکودر جانبی، یکی از توصیفها در اختیار نیست و لذا نمی‌توان آنها را مانند دیکودر مرکزی جدا کرد، به این دلیل لایه پایه با استفاده از "تخمین" از لایه ترکیبی جدا می‌شود. با این کار گرچه خطای تخمین هم خواهیم داشت ولی همانطور که بعداً خواهیم دید MLMDC نسبت به روشهای مرسوم، مصالحه بهتری بین کیفیت مرکزی و کیفیت جانبی برقرار می‌کند که در نتیجه کیفیت متوسط بالاتری هنگام ارسال ویدئو در محیط‌های نویزی فراهم می‌شود.

همچنین MLMDC بر اساس مدل اعوجاج انتها-تا-انتها و با توجه به نرخ خطای کانال بهینه شده است. ابتدا مدلی برای پیدا کردن اعوجاج انتها-تا-انتها برای MDC های ارائه شده که نه تنها برای MLMDC بلکه برای

سایر MDC های حوزه فرکانس^۱ قابل کاربرد است. MDC های بهینه شده، بصورت قابل توجهی، عملکرد بهتری نسبت به غیر بهینه خود دارند. همچنین MLMDC بهینه از بهینه سایر MDC ها هم کیفیت بالاتری ارائه می دهد.

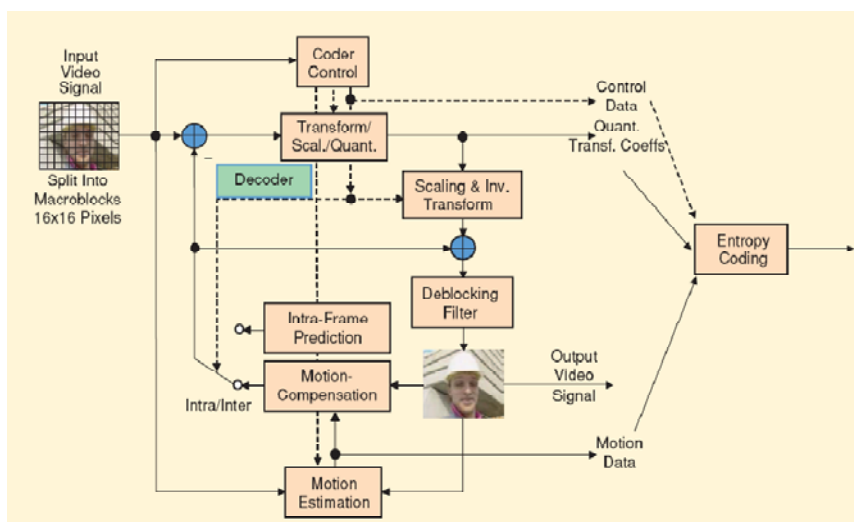


شکل ۱-۵ نحوه ترکیب لایه ها در MLMDC

۴.۱. مرور مفاهیم مورد استفاده از استاندارد H.264/AVC

روش پیشنهادی در استاندارد H.264/AVC پیاده سازی شده است به این خاطر که H.264/AVC جدیدترین کدک معرفی شده بود و گر نه کدکهای دیگر هم قابل استفاده است. در این زیربخش، مفاهیم اصلی فشرده سازی H.264/AVC که مورد رجوع در این رساله هستند را مرور کرده، و همچنین خواننده را با مفاهیم و اصطلاحات استفاده شده در رساله آشنا می کنیم. در شکل ۱-۶ بلوک دیاگرام انکودر H.264/AVC نشان داده شده است که اجزاء آن توضیح داده می شود.

^۱ در مورد حوزه های MDC در فصل بعد توضیح داده خواهد شد.



شکل ۱-۶ بلوک دیاگرام H.264/AVC [16]

۱.۴.۱. تقسیم بندی فریم به ماکروبلوک^۱

انکودرهای استاندارد ویدئو همه اصطلاحاً بلوکی عمل می‌کنند، یعنی یک بلوک از فریم را کد کرده سپس سراغ بلوک بعدی می‌روند. این واحد را یک ماکروبلوک (MB) گفته و اندازه آن 16x16 پیکسل است.

۲.۴.۱. تخمین/جبران حرکت^۲ یا پیش‌بینی بین فریمی

فریمهای ویدئویی تشابه زیادی با یکدیگر دارند، لذا کد کردن تفاضلی آنها از همان ابتدا مد نظر بود. روشن است که کد کردن اختلاف دو فریم متشابه تعداد بیت بسیار کمتری نیاز دارد تا اینکه خود فریم کد شود. تشابه دو فریم بستگی زیادی به میزان حرکت اشیاء در تصویر و یا حرکت دوربین دارد که غالباً هم در تصاویر وجود دارد. به این منظور با شیوهایی به نام جبران حرکت تلاش می‌شود که اختلاف را کمتر کنند. به عبارت دیگر، در جبران حرکت، فریم جاری (فریمی که قرار است کد شود) به جای اینکه مستقیماً از فریم قبلی کم شود، از جبران حرکت شده فریم قبلی کم می‌شود. از این پس به این فریم قبلی، فریم مرجع گفته

¹ Macroblock

² Motion Estimation/Compensation

می‌شود، چونکه فریم جاری را در قیاس با آن کد می‌کنیم. همچنین به سیگنال اختلاف بین بلوک جاری و بلوک مرجع، سیگنال باقیمانده^۱ می‌گوییم. به این قسمت از کدینگ، پیش‌بینی بین فریمی (چونکه یک فریم از یک فریم دیگر پیش‌بینی می‌شود) و یا اینتر کدینگ^۲ هم گفته می‌شود. به علت تشابه زیاد بین فریمها بویژه پس از جبران حرکت، این قسمت سهم زیادی در فشرده‌سازی دارد.

پیدا کردن نحوه حرکت قسمت‌های مختلف، بوسیله تخمین حرکت صورت می‌گیرد. تخمین و جبران حرکت هم بلوکی انجام می‌گیرد. اندازه بلوکها در استاندارد H.264/AVC نسبت به سایر کدکها متنوع تر است، 4x4، 4x8، ...، 16x16. برای هر بلوک با اندازه مشخص، با تخمین حرکت، یک بلوک از فریم مرجع که بیشترین تشابه را داشته باشد، به نام بلوک مرجع، پیدا شده و سپس اختلاف آنها کد و به گیرنده ارسال می‌شود. تفاوت مختصات بلوک جاری و بلوک مرجع تحت عنوان بردار حرکت^۳ به گیرنده ارسال می‌شود. در گیرنده، با در اختیار داشتن بردار حرکت، بلوک مرجع بازخوانی و سیگنال اختلاف با آن جمع شده که حاصل، بلوک جاری بازسازی شده خواهد بود. همانطور که شکل ۱-۶ نشان می‌دهد این عمل دقیقاً در انکودر هم انجام می‌شود. برای اینکه فریم مرجع در انکودر و دیکودر یکسان باشد، در انکودر یک دیکودر هم قرار داده شده که وظیفه آن ساختن فریم مرجع مورد استفاده در تخمین/جبران حرکت می‌باشد. به این ترتیب فریم مرجع در انکودر همانی است که دیکودر آنرا بازسازی خواهد کرد. ناگفته نماند فریم مرجع در استاندارد H.264/AVC می‌تواند تا چندین فریم قبل، چندین فریم بعد، و یا ترکیبی از یک فریم بعد و فریم قبل باشد ولی در بیشتر موارد فریم همسایه (قبل یا بعد) بیشترین سهم را دارد.

¹ Residual signal

² Inter coding

³ Motion Vector

۳.۴.۱. پیش‌بینی داخل فریمی^۱

قبل از آنکه پیش‌بینی داخل فریمی توضیح داده شود، اینترا کدینگ^۲ را در مقابل اینترا کدینگ معرفی می‌کنیم. همانطور که گفته شد در اینترا کدینگ بلوکهای فریم جاری از بلوکهای فریم مرجع، پیش‌بینی شده و تفاضل آنها کد می‌شود. اما در مواقعی که هیچکدام از بلوکهای مرجع به اندازه کافی با بلوک جاری تشابه نداشته باشند اینترا کدینگ کارآیی چندانی ندارد؛ لذا مناسب‌تر است که در این مواقع خود بلوک را، و نه تفاضلی، کد کنیم، که به این عمل اینترا کدینگ گفته می‌شود. اینترا کدینگ جدا از این که برخی مواقع از اینترا کدینگ بازدهی فشرده‌سازی بالاتری دارد، به دلیل اینکه مستقل از فریمهای دیگر کد می‌شود و کیفیت بازسازی آن به کیفیت بازسازی فریمهای مرجع وابسته نیست، باعث توقف انتشار خطا شده که در محیطهای نویزی مورد استفاده قرار می‌گیرد. این مطلب در آینده بیشتر توضیح داده خواهد شد. در کاربرد اینترا کدینگ، فریم اینترا را هم می‌توان نام برد که همه بلوکهای آن اینترا کد شده‌اند. فریمهای اینترا، هم از انتشار بیشتر خطا جلوگیری می‌کنند و هم امکان اصطلاحاً دسترسی تصادفی به فریمها را فراهم می‌کنند. در کدینگ فقط تفاضلی، برای دسترسی مثلاً به فریم ۱۱۰۰م بایستی ۹۹ فرم قبلی دیکد شوند چون که همه تفاضلی کد شده‌اند؛ با قرار دادن فریمهای اینترا با دوره مشخص، این مشکل نیز حل می‌شود. به فریمهای حد فاصل دو فریم اینترا یک GOP^۳ گفته می‌شود.

و اما در پیش‌بینی داخل فریمی، که فقط در استاندارد H.264/AVC وجود دارد، یک بلوک از یک سری پیکسلهای همان فریم کد می‌شود، یعنی تفاضل بلوک جاری از مثلاً یک ردیف از پیکسلهای بالایی کد و به گیرنده ارسال می‌شود. این کار باعث افزایش بازدهی فشرده‌سازی نسبت به اینترا کدینگ می‌شود. اما از طرف

¹ Intra frame prediction

² Intra coding

³ Group Of Pictures

دیگر استقلال بلوکها را که پیش از این بحث شد کمتر می‌کند و لذا خطا پذیری ویدئوی کد شده نیز کاهش می‌یابد.

۴.۴.۱. کدینگ تبدیل کردن^۱

بلوک تفاضل، چه در پیش‌بینی بین فریمی و چه در پیش‌بینی داخل فریمی، و خود بلوک در ایترا کدینگ، از یک مبدلی به نام تبدیل DCT^2 عبور می‌کنند. کار این تبدیل، ناهمبسته کردن عناصر بلوک است. مقدار پیکسلها در یک بلوک و همچنین مقدار تفاضل آنها در بلوک تفاضلی از یکدیگر مستقل نبوده و دارای اطلاعات مشترک و یا افزونگی هستند. از طرف دیگر چونکه این مقادیر بصورت تک تک و یا اسکالر کد و فرستاده می‌شوند، این افزونگی باعث کاهش فشردگی می‌گردد. لذا این تبدیل در واقع افزونگی مکانی^۳ را کم می‌کند، در مقایسه با اینتر کدینگ که افزونگی زمانی (بین فریمها) را کاهش می‌داد. نشان داده شده است در کوانتیزیشن اسکالر یک بلوک، اگر تبدیل KLT^4 را اعمال کنیم به کمترین اعوجاج می‌رسیم [17]. اما تبدیل KLT وابسته به داده ورودی است که این امر استفاده از آن را در استانداردها با مشکل مواجه می‌کند، به این دلیل از DCT^5 استفاده می‌شود. به خروجی تبدیل DCT ، بلوک DCT و به عناصر بلوک DCT ، ضرایب DCT گفته می‌شود.

تبدیل DCT یک ماتریس متعامد یکه است که درایه‌های آن توابع مثلثاتی و لذا اعداد گنگ هستند. در دیکودر و همچنین در انکودر، به تبدیل عکس DCT نیاز داریم که درایه‌های آن هم طبیعتاً اعدادی گنگ هستند. بنابراین، در کامپیوترهای دیجیتال که دقت اعداد محدود است نمی‌توان تبدیل DCT و عکس آنرا

¹ Transform Coding

² Discrete Cosine Transform

³ Spatial

⁴ Karhunen-Loeve Transform

⁵ Discrete Cosine Transform

کاملاً دقیق محقق نمود؛ نتیجه آن می‌شود که بلوک ورودی DCT و بلوک خروجی عکس DCT با یکدیگر منطبق نیستند. برای رفع این مشکل در H.264/AVC، تبدیل DCT به تبدیل IDCT (Integer DCT) تقریب زده می‌شود که عناصر آن با آسانی قابل پیاده‌سازی هستند.

۵.۴.۱. کوانتیزیشن^۱

آنچه تا بحال انجام شده است، کاهش افزونگی ویدئو بوده است، یعنی داده‌هایی که حاوی اطلاعات نبوده‌اند تا حد امکان حذف شده است. به عبارت دیگر، مراحلی که تا کنون انجام شده است فشردگی را بدون افت^۲ و یا بدون اعوجاج محقق می‌کنند. اما در کوانتیزیشن، با کوانتیزه کردن ضرایب DCT، که درست مشابه با مبدل‌های A/D عمل می‌کند، با حذف یک سری اطلاعات، اعوجاجی وارد می‌شود که در دیکودر قابل برگشت نیست. در عوض این کار، آنتروپی سیگنال کاهش یافته و سیگنال را می‌توان با تعداد بیت کمتری ارسال کرد. به عبارت دیگر، با کوانتایز کردن ضرایب، تشابه آنها بیشتر می‌شود، با افزایش تشابه و یا کاهش بی‌نظمی، طبق تعریف، آنتروپی کاهش می‌یابد و کاهش آنتروپی به معنای کاهش حداقل تعداد بیت برای ارسال سیگنال است.

کوانتیزیشن در H.264/AVC به فرم زیر انجام می‌گیرد:

$$\hat{x} = \text{sign}(x) \left\lfloor \frac{|x|}{Q} + f \right\rfloor \quad (1-1)$$

که در آن Q به معنای گرد کردن به سمت منفی بی‌نهایت است، f آفست گام کوانتیزیشن است. f آفست گردسازی است که بین صفر و نیم تغییر می‌کند که در مورد آن بعداً بیشتر توضیح داده خواهد شد. معادله فوق نشان می‌دهد که مقدار مطلق ضریب کوانتایز شده و بعد علامت آن وارد می‌شود. این فرم از کوانتیزیشن

^۱ Quantization

^۲ Loss less

با توجه به توزیع داده‌ها و با بهینه‌سازی بدست آمده است. عکس کوانتیزیشن و یا دیکوانتیزیشن هم به صورت زیر است:

$$\tilde{x} = Q \cdot \hat{x} \quad (2-1)$$

در H.264/AVC، برای ارسال مقدار Q از پارامتر دیگری به نام QP استفاده می‌شود که رابطه آن با Q تقریباً^۱ به فرم زیر است:

$$Q = 2^{(QP-4)/6} \quad (3-1)$$

چنانچه دیده می‌شود با استفاده از QP می‌توان Q را در نرخ بیت‌های بالا با فواصل کمتر (حتی کمتر از یک) و در نرخ‌های پایین با فواصل بیشتر تغییر داد.

ناگفته نماند که در H.264/AVC، DCT و کوانتیزیشن، برای صرفه جویی در عملیات با هم ترکیب شده‌اند، به همین خاطر شکل ۱-۶ آنها در یک بلوک نشان داده است.

۶.۴.۱. آنتروپی کدینگ^۲

در قسمت آخر، اطلاعات آماده شده از قسمت‌های قبل شامل داده‌های کنترلی (مثلاً QP)، ضرایب DCT، بردارهای حرکت و ... را کد می‌کنیم. نام این قسمت آنتروپی انکودر است چون تلاش می‌کند حداقل تعداد بیت، در حد نرخ آنتروپی که قبلاً بحث شد، را به اطلاعات ورودی تخصیص دهد. در H.264/AVC برای آنتروپی انکودر دو گزینه وجود دارد: CAVLC^۳ و CABAC^۴ اولی بر اساس کدهای هافمن و دومی بر

^۱البته در H.264/AVC از یک نگاشت یک به یک بین Q و QP استفاده می‌شود که مقداری با تابع فوق تفاوت دارد.

^۲ Entropy Coding

^۳ Context based Adaptive Variable Length Coding

^۴ Context based Adaptive Arithmetic Coding

اساس کدهای حسابی کار می‌کند. هر دو بسیار کارآمد هستند اما CABAC در ازای حجم محاسبات بالاتر نرخ بیت پایبندتری را فراهم می‌کند.

۵.۱. ساختار رساله

در فصل دوم مروری بر روشهای MDC موجود ارائه شده و در جدولی کارآیی آنها با یکدیگر مقایسه شده است. در فصل سوم MLMDC با جزئیات معرفی شده و عملکرد آن در مقایسه با MDC های مرسوم مورد مطالعه قرار گرفته است. در فصل چهارم، MLMDC برای داشتن بهترین کیفیت در دیکودر با توجه به شرایط کانال و محتوای ویدئو، بهینه سازی شده است. در این فصل هم MLMDC بهینه با MDC های دیگر که آنها هم بهینه شده‌اند مقایسه شده است. در فصل پنجم، جمع بندی رساله و پیشنهادهایی برای ادامه این کار را ارائه می‌دهیم.

۲. مروری بر روشهای

MDC

در این بخش، طرحهای MDC بر پایه حوزه‌ای که توصیفها تولید و از هم جدا می‌شوند گروه‌بندی می‌گردند. جزئیات این گروه‌ها تشریح می‌شود و روشهای خاصی که برای بهبود عملکرد MDC ارائه شده‌اند نیز آورده می‌شوند. تمرکز روی منبع ویدئو است ولی روشهایی که برای منبع تصویر ارائه شده و قابل اعمال روی ویدئو هستند نیز مرور شده است. ناگفته نماند که در مورد MDC دو مقاله مروری قبلاً به چاپ رسیده است، مقاله Goyal [1] و مقاله Wang و همکاران [2]. مقاله [1] بیشتر کاربرد MDC را در انتقال داده‌های چندرسانه‌ای در محیطهای نوپزی بررسی می‌کند و مقاله [2] هم که ویژه ویدئو نوشته شده است بیشتر حول مسئله دریافت بحث کرده است. متفاوت با و به‌روزتر از آنها، جمع آوری و دسته بندی روشهای MDC و در نهایت مقایسه آنها در این بخش ارائه شده است.

۱.۲. دسته بندی طرحهای مختلف MDC

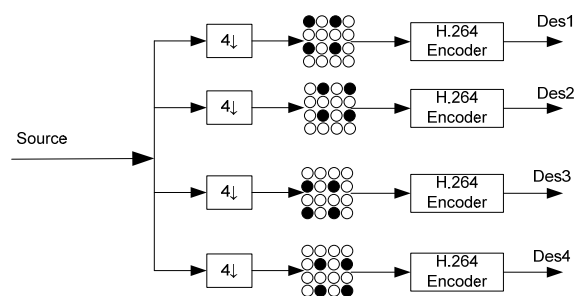
حوزه‌های گروه‌بندی MDC شامل spatial، زمانی، فرکانس، فشرده شده هستند که عمل MDC به ترتیب روی پیکسلها، فریمها، ضرایب یا بلوک DCT، و روی استریم نهایی فشرده شده صورت می‌گیرد. روشهایی هم هستند که عمل MDC در دو حوزه از حوزه‌های گفته شده انجام می‌گیرد. و همچنین روشهایی برای تولید MDC وجود دارند که بر خلاف روشهای فوق، بر پایه تقسیم داده‌ها نیستند (unpartitioning methods) که در انتها مورد بحث قرار می‌گیرند.

۱.۱.۲. حوزه مکانی

در این گروه از روشها، MDC در حوزه پیکسل انجام می‌گیرد. ساده‌ترین راه به این منظور تقسیم کردن تصویر/فریم به تکه‌های کوچکتر با شیوه زیرنمونه برداری چند فازه spatial^۱ (PSS) می‌باشد. هر تصویر کوچکتر را که زیرتصویر^۲ نامیده می‌شود مستقلاً انکود می‌کنیم، همان طور که در شکل ۲-۱ نشان داده شده است. در طرف گیرنده اگر که همه توصیفها دریافت شده باشند، تصویرهای کوچکتر با هم ادغام شده و تصویر اصلی بازسازی می‌شود. اگر یک توصیف نرسیده باشد با درون‌یابی از تصاویر دریافت شده بازسازی می‌گردد. اما در این روش به این صورت، امکانی برای افزایش افزونگی و در نتیجه کیفیت جانبی وجود ندارد. در ادامه راههایی که برای بهبود کیفیت جانبی ارائه شده است مرور می‌شوند. البته شایان ذکر است که زیرنمونه برداری چه در حوزه پیکسل و چه در حوزه زمان باعث کاهش همبستگی بین داده‌ها می‌شود که این خود فشرده شونده‌گی آنها را کم می‌کند. لذا نسبت به SDC افزونگی وجود دارد ولی به کیفیت جانبی کمکی نمی‌کند.

^۱ Polyphase spatial subsampling

^۲ Subimage

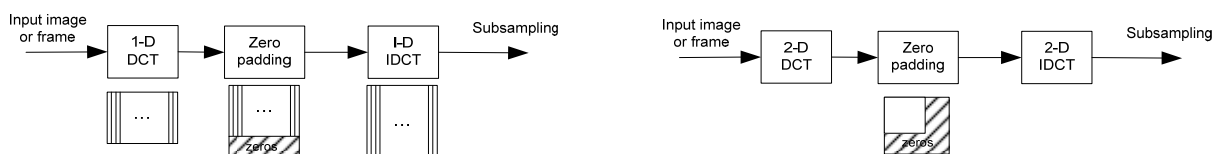


شکل ۲-۱ MDC با شیوه زیرنمونه برداری چند فازه spatial

zero padding

ابتدا تصویر به حوزه DCT و یا فرکانس برده می‌شود، سپس همانطور که در شکل ۲-۲ نشان داده شده است به کناره‌های ماتریس داده، صفر اضافه می‌شود. سپس از تصویر جدید عکس DCT گرفته می‌شود. حال به کمک PSS، از تصویر بزرگتر تصاویر کوچکتر ساخته می‌شود. همانطور که می‌دانیم اضافه کردن صفر در حوزه فرکانس به معنی نمونه‌برداری افزایشی در حوزه پیکسل است. ولی انجام این کار در حوزه DCT امکان تنظیم دقیق افزونگی را به ما می‌دهد که در حوزه پیکسل امکان پذیر نمی‌باشد. این ایده توسط Shirani و همکاران، ابتدا برای تصویر [18] و سپس برای ویدئو [19] ارائه شد.

همانطور که در شکل ۲-۲ نشان داده شده است صفرها را می‌توان یک بعدی و یا دوبعدی به تصویر DCT اضافه کرد. در [20] نشان داده شده است که اگر این کار، یک بعدی و در جهتی باشد که تصویر کمتر قابل تخمین است بازدهی بالاتری دارد. هر دو حالت یک‌بعدی و دوبعدی در این شکل نشان داده شده است.



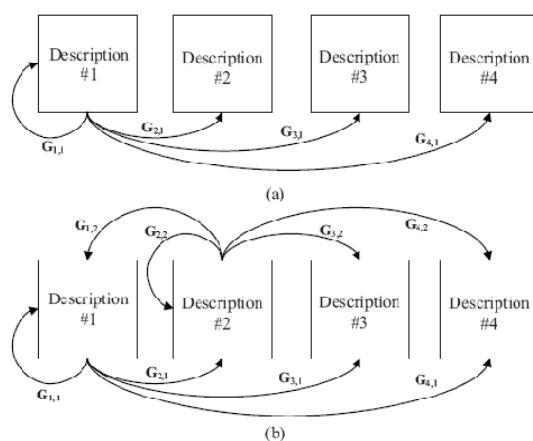
شکل ۲-۲ اضافه کردن صفر قبل از زیرنمونه برداری، چپ: یک بعدی و راست: دو بعدی.

کپی کردن داده‌های کمتر قابل بازیابی

همانطور که قبلاً گفته شد زیرتصاویری که دچار خطا شده‌اند از تصاویر دریافت شده، تخمین زده می‌شوند. در این روش خط‌هایی (افقی یا عمودی) از تصویر که تخمین آنها خطای زیادی دارد مستقیماً در همه توصیف‌ها گنجانده می‌شوند. روشن است که خطوط کپی شده باعث بهبود کیفیت جانبی شده گرچه موجب افزایش افزونگی هم می‌شوند. این روش هم در [20] پیشنهاد شده است.

فیلترینگ در انکودر یا دیکودر

بازیافت زیرتصویرها را می‌توان با فیلترهایی مناسبتر از درونیابی ساده انجام داد. همانطور که در شکل ۲-۳ نشان داده شده است ضرایب فیلتر مناسب برای بازیافت هر زیرتصویر از زیرتصویرهای دیگر، در انکودر بدست می‌آیند و به همراه توصیف ارسال می‌شوند. به این ترتیب می‌توان کیفیت جانبی را بسته به درجه فیلتر و حجم محاسبات قابل قبول، تا حد خوبی بهبود داد. شبیه‌سازیهای انجام داده شده در مرجع [21] عملکرد بهتر این روش را نسبت به درونیابی دوتایی خطی، درونیابی دوتایی مربعی و zero-padding (بجز در چند مورد) نشان می‌دهد. این روش را می‌توان با zero-padding ادغام کرد و نتایج بهتری گرفت [22].

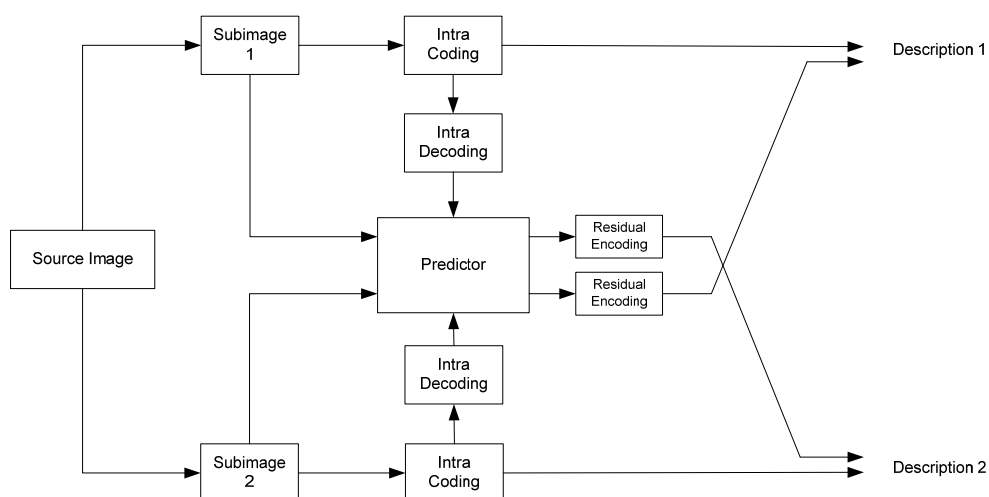


شکل ۲-۳ تقریباً زدن توصیف‌های از دست رفته و فیکه (a) فقط یک توصیف رسیده باشد و (b) توصیف‌های اول و دوم رسیده باشند

[21].

ارسال سیگنال باقیمانده از توصیف (های) دیگر

در این روش، زیرتصویر توصیف ۲ از زیرتصویر توصیف ۱ پیش‌بینی می‌شود و سیگنال اختلاف بعد از کوانتیزیشن همراه با توصیف ۱ ارسال می‌گردد؛ همانطور که در شکل ۲-۴ نشان داده شده است. در دیکودر جانبی، یک زیرتصویر داریم و سیگنال اختلاف آن با سایر زیرتصویرها، به این ترتیب با دریافت یک زیرتصویر، سایرین را می‌توان با دقت خوبی باز یافت کرد. میزان این دقت بسته به گام کوانتیزیشن است که برای کد کردن سیگنال اختلاف مورد استفاده قرار گرفته است که در واقع ابزار کنترل افزونگی هم بشمار می‌رود. این روش برای تصاویرهای JPEG و فریمهای اینترا پیشنهاد شده است [24] [23].



شکل ۲-۴ پیش‌بینی یک زیرتصویر از زیرتصویر دیگر

زیر نمونه برداری چندفازه غیرخطی

در روشهای مبتنی بر PSS خطی، تصویر را در راستای عمودی و افقی با نرخ یکسان نمونه برداری می‌کنند. اما با روشهای غیر خطی می‌توان بعضی از نقاط تصویر، مانند ROI^۱ را با نرخ بیشتر نمونه برداری کرد. به این ترتیب در توصیفها تعداد داده بیشتری از ROI قرار دارد که باعث بهبود کیفیت جانبی از نقطه نظر بیننده

^۱ Region Of Interest

می‌شود [25]. بطور معادل می‌توان ابتدا تصویر را تحت یک تبدیل غیر خطی قرار داد و سپس PSS خطی به

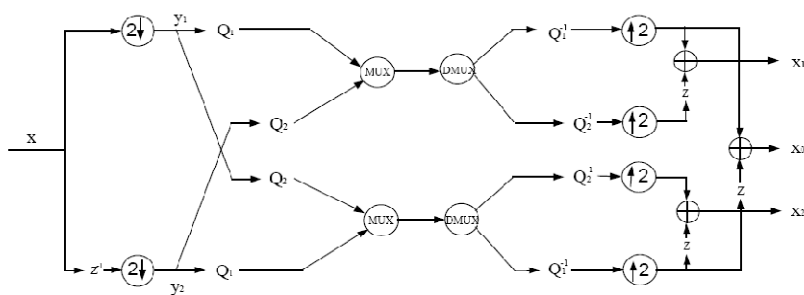
آن اعمال کرد. این کار در شکل ۲-۵ نشان داده شده است [26].



شکل ۲-۵ MDC با استفاده از زیرنمونه برداری چند فازه غیرخطی (a) تصویر اصلی (b) بعد از تبدیل غیرخطی (c) وقتی هر چهار توصیف رسیده باشند و (d) فقط یک توصیف رسیده باشد [26].

کدینگ چند نرخه

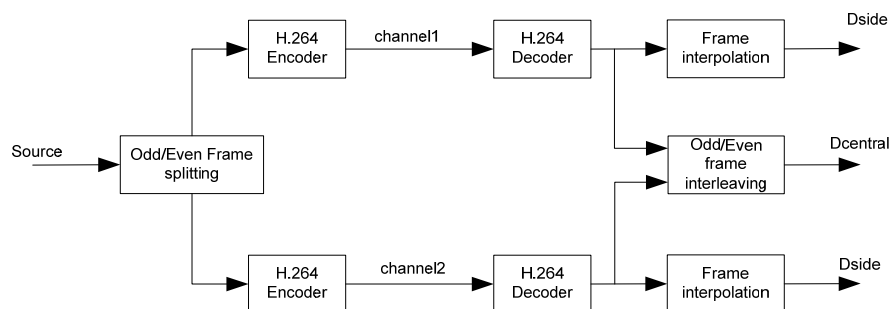
در این روش یک کپی از زیرتصویر دوم ولی با نرخ بیت کمتر تولید می‌شود و در توصیف اول قرار می‌گیرد. به همین ترتیب نرخ پایین زیرتصویر اول در توصیف دوم گذاشته می‌شود [27]. همانطور که در شکل ۲-۶ نشان داده می‌شود زیرتصویر دوم با Q_2 کوانتایز می‌شود و با زیرتصویر اصلی که با Q_1 کوانتایز شده، مالتی پلکس می‌شود و به عنوان توصیف اول فرستاده می‌شود. میزان افزونگی توسط Q_2 کنترل می‌گردد.



شکل ۲-۶ MDC چند نرخه حوزه spatial [27]

۲.۱.۲. MDC حوزه زمان

در MDC های حوزه زمان، توصیفها با پردازش یا عملیاتی که روی فریمها انجام می شود بدست می آیند. یک روش ساده که در شکل ۲-۷ نشان داده شده است این است که فریمها بین دو توصیف تقسیم شوند؛ فریمهای زوج در توصیف اول و فریمهای فرد در توصیف دوم [28]. در دیکودر جانبی، فریمهای نرسیده توسط شیوه های اختفای خطا بازیافت می شوند. کارآیی این روش بستگی به کیفیت فریم های بازیابی شده در دیکودر جانبی دارد که این خود وابسته به همبستگی بین فریمها دارد. به عبارت دیگر، اگر حرکت بین فریمها به اندازه کافی کم و یا قابل پیش بینی نباشد فریم بازیابی شده چندان مشابه فریم اصلی نیست و لذا اعوجاج جانبی خوبی نداریم. راه حلهایی که به این منظور ارائه شده اند را در ادامه مورد بررسی قرار می دهیم. همانطور که خواهید دید راه حلها، بجز اولی، منجر به اضافه کردن فریمهایی در توصیفهای تشکیل شده به فرم بالا می شوند؛ این فریمها ممکن است دقیقاً کپی فریم آن یک توصیف باشد، یا کیفیت پایینتر آن باشد و یا فریمهای مجازی میانی باشند.



شکل ۲-۷ بلوک دیاگرام پایه ای MDC حوزه زمان

کپی کردن اطلاعات حرکت در هر دو توصیف

بردارهای حرکت فریمهای توصیف اول (فریمهای زوج) در توصیف دوم قرار می‌گیرد و بالعکس [29]. به این ترتیب با نرسیدن توصیف اول و در نتیجه فریم زوج، از اطلاعات حرکت آن که در فریم فرد متناظر گنجانده شده، استفاده شده و با کیفیت بهتری بازیافت می‌شود.

در [30]، هر فریم از توصیف ۱، از دو فریم مجاور خود در توصیف ۲ تخمین حرکت می‌شود و بردارهای حرکت در فریمهای مرجع (توصیف ۲) قرار می‌گیرد. به این ترتیب اگر توصیف ۱ در دسترس نباشد، با بردارهای حرکت و فریمهای مرجع که در توصیف ۲ هستند، فریم با دقت خوبی بازیافت می‌شود. تفاوت آن با روش قبلی اینست که مرجع حرکت که فریمهای توصیف سالم هستند در اختیار است و لذا اختفای خطا بهتر صورت می‌گیرد. این روش با روشهای زیر مقایسه شده است:

”روش “average“: فریم نرسیده با متوسط گیری از دو فریم مجاور ساخته می‌شود.

روش “inplaceMC“: که در آن بردارهای حرکت بین دو فریم مجاور در توصیف فرد در $1/2$ ضرب شده و به عنوان بردار حرکت فریم زوج، که به گیرنده نرسیده است، مورد استفاده قرار می‌گیرد.

روش “MCinterp“: که در آن اطلاعات حرکت فریم از بین رفته از اطلاعات فریمهای مجاور به کمک الگوریتمی مبتنی بر هبستگی فاز بدست می‌آید [31].

نتایج شبیه سازی نشان می‌دهد که روش پیشنهاد شده در این مقاله ([30])، از همه بهتر است و بعد از آن به ترتیب “MCinterp“، “inplaceMC“ و “average“ قرار دارند.

کپی کردن داده‌های کمتر قابل بازیابی

همانطور که از نام این روش پیداست، فریمهایی که بازیابی آنها مشکل است (با خطای زیادی همراه است) در هر دو توصیف کاملاً یکسان کپی می‌شوند. این روش در [32] پیشنهاد گردیده است.

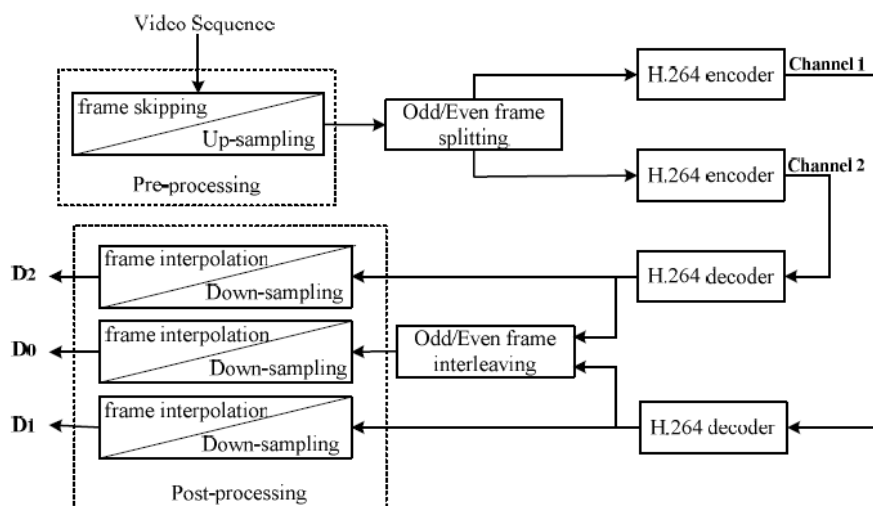
MDC چند نرخه

همانطور که قبلاً گفته شد در MDC چند نرخه، افزونگی از کد کردن مجدد داده‌های اصلی ولی با نرخ بیت کمتر حاصل می‌شود. در حوزه فریم، این داده‌های نرخ پایین می‌توانند یک GOP باشد [33] و یا یک فریم [34]. در [33]، یک GOP چندین بار ولی در نرخهای متفاوت کد می‌شود و هر GOP در یک توصیف قرار می‌گیرد. در دیکودر، توصیفها و یا GOPهای با نرخ پایتتر دور انداخته می‌شوند. سوئیچینگ بین توصیفها در انتهای GOP صورت می‌گیرد و لذا بسته به اینکه کدام توصیفها به دیکودر برسند کیفیت برای یک GOP در یک حد خواهد بود. در [34]، فریمهای توصیف اول با نرخ پایتتری مجدداً کد شده و در توصیف دوم قرار می‌گیرند. به همین ترتیب توصیف دوم شامل نرخ پایین فریمهای توصیف اول است و برعکس. افزونگی با گام کوانتیزیشن استفاده شده برای فریمهای نرخ پایین کنترل می‌شود.

کم و اضافه کردن تعداد فریمها

در روش پیشنهاد شده در [35]، قبل از اینکه فریمها را تقسیم کنند با اضافه کردن فریم در بازه‌هایی که حرکت سریعتر است حرکت بین آنها کندتر می‌شود. لذا در دیکودر جانبی، بهتر می‌توان فریمهای نرسیده را بازیافت کرد چونکه اختلاف ما بین فریمها در این حالت کم می‌شود. البته روشن است که برای حفظ حرکت واقعی ویدئو، فریمهای اضافی نمایش داده نمی‌شوند و فقط برای اختفا کردن فریمهای نرسیده مورد استفاده قرار می‌گیرند. همین ایده ولی با قابلیت کم کردن فریمها، علاوه بر قابلیت اضافه کردن، در [36] ارائه شده است. در این کار، در نواحی که سرعت خیلی کم است فریمها زیر نمونه برداری می‌شوند و در عوض در

نواحی که سرعت زیاد است فریمهای میانی اضافه می‌شوند. ساختار انکودر این روش در شکل ۲-۸ نشان داده شده است.



شکل ۲-۸ بلوک دیاگرام MDC حوزه زمان با شیوه نرخ فریم تطبیق پذیر [36]

۳.۱.۲. حوزه فرکانس

در MDC های حوزه فرکانس که حوزه DCT یا حوزه تبدیل هم گفته می‌شود، چندین روش وجود دارند که شامل کوانتایزر اسکالر چند توصیفه ($MDSQ^1$)، تقسیم کردن ضرایب^۲، $MDTC^3$ ، MDC چند نرخه می‌باشد. در ادامه جزئیات هر کدام توضیح داده می‌شود.

MDSQ

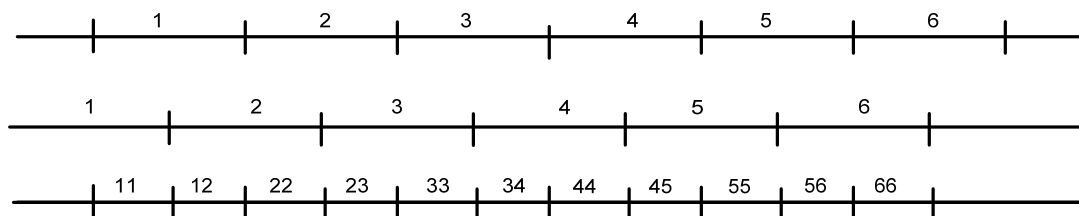
مفهوم MDSQ این است که در هر توصیف کوانتیزیشن را به گونه‌ای انجام دهیم که وقتی با یکدیگر در دیکودر مرکزی ادغام می‌شوند منجر به خطای کوانتیزیشن کوچکتری بشود. ساده ترین راه این است که

¹ Multiple Description Scalar Quantizer

² Coefficients partitioning

³ Multiple Description Transform Coding

نواحی کوانتیزیشن را به اندازه نصف بازه‌های کوانتیزیشن جابجا کنیم، همانطور که در شکل ۲-۹ نشان داده شده است.



شکل ۲-۹ ساده ترین MDSQ، دو خط بالایی: دیکودر جانبی و خط پایینی: دیکودر مرکزی

کوانتیزیشن را می‌توان با جابجایی نه لزوماً 1/2 انجام داد. در H.264/AVC گزینه ای به نام adaptive quantization offset وجود دارد که می‌تواند به این منظور استفاده شود. این ایده در [37] پیشنهاد شده است که در هر توصیف جابجایی‌های متفاوتی (Δ_{e1} و Δ_{e2}) استفاده می‌شود:

$$Z_{1,i,j} = \left\lfloor \frac{|W_{i,j}| + \beta + \Delta_{e1}}{\Delta} \right\rfloor \cdot \text{sign}(W_{i,j})$$

$$Z_{2,i,j} = \left\lfloor \frac{|W_{i,j}| + \beta + \Delta_{e2}}{\Delta} \right\rfloor \cdot \text{sign}(W_{i,j})$$

در دیکودر مرکزی، دو عدد بازسازی شده متوسط‌گیری می‌شوند، می‌توان نشان داد که در این صورت خطای کوانتیزیشن کمتر است.

میزان Δ_e ها می‌تواند ثابت باشد و یا با شرایط ورودی و کانال تغییر کند. شبیه سازی انجام شده نشان می‌دهد که در این حالت نسبت به حالت ثابت به کیفیت بالاتری می‌رسیم. با افزایش جابجایی که حد مشخصی دارد میزان افزونگی کاهش می‌یابد؛ و لذا در این روش میزان افزونگی را نمی‌توان از یک حدی کمتر کرد. به همین دلیل این روش برای کاربردهای با نرخ خطای پایین مناسب نیست.

در یک روش پیچیده تر MDSQ، تعداد سطوح در کوانتیزیشن مرکزی لزوماً دو برابر تعداد سطوح در کوانتیزیشن جانبی نیست. شکل ۲-۱۰ مثالی نشان می‌دهد که ۲۱ سطح کوانتیزیشن مرکزی به ۸ اندیس در هر توصیف نگاشته می‌شود. هنگامیکه هر دو توصیف دریافت شوند، هر دو اندیس در اختیار است و لذا طبق جدول، سطح کوانتیزیشن بطور یکتا مشخص می‌شود. در صورتیکه یکی از توصیفها نرسد دیکودر جانبی مجبور است یکی از سطوح داده شده در سطر یا ستون متناظر را انتخاب کند. اولین و مشهورترین مقاله در این زمینه، مرجع [38] است که جدول تخصیص اندیس را بطور بهینه برای منابع با توزیع یکنواخت بدست آورده است.

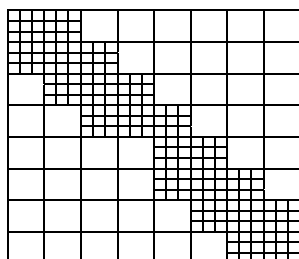
		Description 2 indices								
		1	2	3	4	5	6	7	8	
Description 1 indices	1	1	2							
	2	3	4	6						
	3		5	7	8					
	4			9	10	11				
	5					12	14			
	6					13	15	16		
	7						17	18	20	
	8							19	21	

شکل ۲-۱۰ یک نمونه جدول تخصیص اندیس در MDSQ

در MDSQ اصلاح شده^۱ [39]، دو مرحله کوانتیزیشن صورت می‌گیرد. در مرحله دوم، سطح کوانتیزیشن مرحله قبل دوباره کوانتایز می‌شود و اندیسهای مرحله دوم بطور یک در میان در توصیفها قرار می‌گیرند. در دیکودر جانبی، اندیسهای مرحله دوم کامل نیستند و دور ریخته می‌شوند ولی در دیکودر مرکزی مورد استفاده قرار گرفته و لذا کیفیت بالاتری خواهیم داشت. این کار در کدینگ مبتنی بر پیش‌بینی هم انجام شده است [40].

¹ Modified MDSQ

MDSQ در مقاله‌های تئوری اطلاعات بسیار مورد بحث قرار گرفته است اما بیشتر آنها مبتنی بر کوانتیزیشن اسکالر نیست و لذا قابل پیاده‌سازی در کدک‌های استاندارد ویدئو و تصویر نیستند. به همین دلیل، بیشتر به آن پرداخته نمی‌شود.



شکل ۲-۱۱ جدول تخصیص اندیس در MDSQ اصلاح شده

تقسیم کردن ضرایب

در این گروه، ضرایب DCT بین توصیفها تقسیم می‌شوند. فقط تبدیل DCT را در نظر می‌گیریم چونکه همه استانداردهای ویدئو بر مبنای DCT هستند. برای تقسیم بندی آنها دو رویکرد وجود دارد:

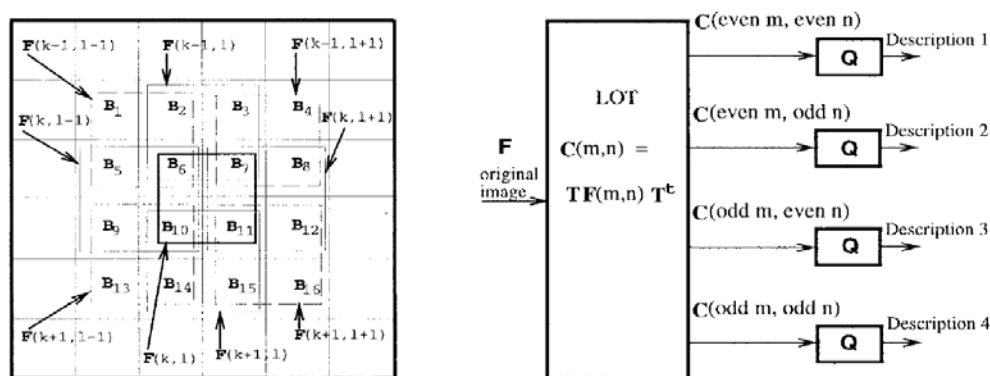
در رویکرد اول، ضرایبی که از یک حد آستانه بزرگتر هستند در همه توصیفها تکرار می‌شوند و بقیه بین آنها تقسیم می‌گردند [41]. جای خالی ضرایب کوچک در توصیفها صفر گذاشته می‌شود، مقدار متوسط ضرایب DCT در ویدئو. ارسال صفر به جای مقدار واقعی ضرایب سبب کاهش نرخ می‌شود چون آنتروپی را کم می‌کند. در عوض اعوجاج در دیکودر جانبی افزایش می‌یابد. هر چقدر که آستانه تقسیم کردن ضرایب کوچکتر باشد، اعوجاج جانبی کمتر و در عوض افزونگی بیشتر خواهد بود و برعکس. در مقاله [42]، مقدار بهینه آستانه تصمیم‌گیری به منظور متعادل کردن توصیفها از لحاظ نرخ بیت و اعوجاج بدست آمده است.

در رویکرد دوم، ضرایب ابتدایی در هر دو توصیف کپی می‌شوند ولی سایر ضرایب در یک توصیف کپی و در توصیف دیگر صفر می‌شوند [43]. فایده اصلی این رویکرد نسبت به رویکرد اول صرفه جویی بیشتر در

تعداد بیت است گرچه توصیفها نامتقارن هستند. تعداد بیت کمتر از نحوه کد کردن ضرایب که بر اساس run-of-zeros است ناشی می شود.

جداسازی ضرایب را می توان بلوکی هم انجام داد. اما نداشتن یک بلوک در دیکودر جانبی و بازیابی کردن آن از بلوکهای مجاور و یا بلوکهای فریم قبلی کیفیت جانبی خوبی به ما نمی دهد. یک راه حل تبدیل هم پوشان است. در مقاله [44]، به منظور افزایش همبستگی بین بلوکها از تبدیل LOT استفاده شده است که در آن بلوکها می توانند به عنوان مثال 50% (شکل ۲-۱۲) هم پوشانی داشته باشند. اطلاعات مشترک برای بازیافت بلوکهای از دست رفته مورد بهره برداری قرار می گیرد. اگر تبدیل استفاده شده DCT باشد به آن DCT-LOT گفته می شود که در واقع تبدیل بهینه از جهت فشرده سازی است. ولی تبدیل را هم می توان به منظور کیفیت بازیابی هم بهینه کرد به گونه ای که مقدار بهینه افزونگی، در توصیفها قرار داده شود.

این ایده در [45] هم مورد استفاده قرار گرفته است ولی با این تفاوت که سیگنال اختلاف بلوکهای توصیف ۲ (۱) نسبت به توصیف ۱ (۲) در توصیف ۱ (۲) قرار داده می شود که کیفیت جانبی را در ازای افزونگی بیشتر، بهتر می کند.



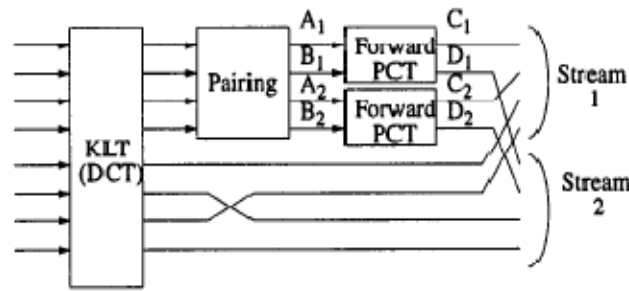
شکل ۲-۱۲ چپ: بخش بندی فریم به بلوکهای هم پوشان و راست: تخصیص بلوکها به توصیفها [44].

MDTC

ضرایب DCT یک بلوک را به این علت که ناهمبسته هستند نمی‌توان از یکدیگر تخمین زد. بجای آن باید از ضرایب بلوکهای مجاور استفاده کرد که اولاً چقدر تخمین خوبی است بستگی به میزان یکنواختی تصویر دارد و ثانیاً در منابع ویدئو که DCT روی تصویر اختلاف اعمال می‌شود و همبستگی بین بلوکها خیلی کم است این کار بازدهی چندانی ندارد. راه حل آن استفاده از یک تبدیل همبسته کننده است که در MDCT مورد استفاده قرار می‌گیرد.

در MDCT، بعد از تبدیل DCT یک تبدیل دیگر به نام PCT¹ اعمال می‌شود؛ این تبدیل همانطور که از نام آن پیداست یک جفت ضرایب ناهمبسته را به یک جفت ضرایب همبسته تبدیل می‌کند. حال این ضرایب همبسته بین توصیفها تقسیم می‌شوند؛ در صورتی که یک توصیف نرسید، ضرایب آن به کمک همبستگی که توسط PCT ایجاد شده است از ضرایب توصیف دیگر تخمین زده می‌شود. به این ترتیب کیفیت جانبی به مراتب بهتری نسبت به زمانی که خود ضرایب DCT را مستقیماً MDC کنیم بدست می‌آید. PCT از طرف دیگر، در خلاف جهت DCT که برای ناهمبسته کردن استفاده می‌شود، عمل می‌کند و باعث کاهش بازدهی فشرده‌سازی می‌شود، این افزایش نرخ همان افزونگی این روش است. به همین دلیل همانطور که شکل ۲-۱۳ نشان می‌دهد، PCT روی تمام ضرایب اعمال نمی‌شود. در واقع چه ضرایبی با هم ترکیب شوند هم مهم است؛ به عنوان مثال PCT روی ضرایبی که توزیع آماری یکسان دارند فقط افزونگی را بالا می‌برد و تاثیری در دقت تخمین ندارد.

¹ Pair-wise Correlating Transform



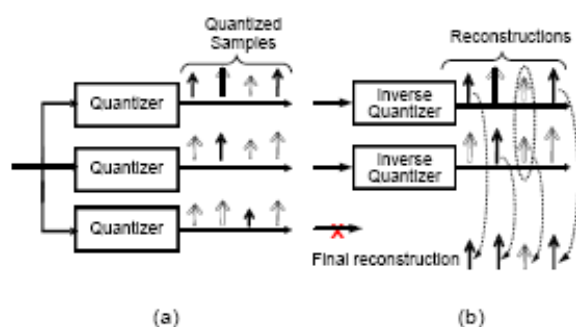
شکل ۲-۱۳ بلوک دیاگرام MDTC [51]

ایده اولیه این روش توسط Reibman و همکاران در [46] ارائه شد و همزمان و با کمک Goyal *et al.* کامل تر و جامع تر مورد مطالعه قرار گرفت. تحلیل نرخ-اعوجاج و بررسی تبدیلیهای اورتوگونال و نااورتوگونال در مقاله [47] بحث شده است، الگوریتم برای تعداد توصیفهای بیش از دو تا و توصیفهای نامتعادل و منابع گوسین در [48] تعمیم داده شده و در [49] روی منبع تصویر اعمال شده است. در [50]، با بهینه‌سازی افزونگی-نرخ-اعوجاج، بهترین تبدیل محاسبه شده است. تحلیل ژرف و مبسوط روش در [51] به چاپ رسیده است که نشان داده MDTC در افزونگیهای زیاد کارایی خود را از دست می‌دهد. این در واقع به این خاطر است که بهر حال از یک ضریب نمی‌توان دو ضریب اولیه را بدون خطا بدست آورد و تا حدی همبستگی ایجاد شده کمک می‌کند و همیشه مقداری خطای تخمین وجود دارد. این مسئله در [52] تا حد زیادی حل شده است به این گونه که آن قسمتی که بطور آماری تخمین آن مشکل است را در داخل توصیف قرار می‌دهند و به این ترتیب با افزایش افزونگی، خطای تخمین باز هم کمتر می‌شود. در مقاله [53] روی منبع ویدئو هم اعمال شد. مطالعه جامعی از کارهای Goyal در مقاله [54] به چاپ رسیده است.

MDC چند نرخه

کدینگ چند نرخه در MDC حوزه تبدیل را می‌توان هم روی ضرایب انجام داد و هم روی بلوکها. در مورد اول، ضرایب DCT را با QPهای متفاوت کوانتایز می‌کنند و نرخ بالاها و نرخ پایین‌ها را بین دو یا چهار توصیف تقسیم می‌کنند. به عبارت دیگر، همه ضرایب در همه توصیفها هستند ولی در هر توصیف نصف

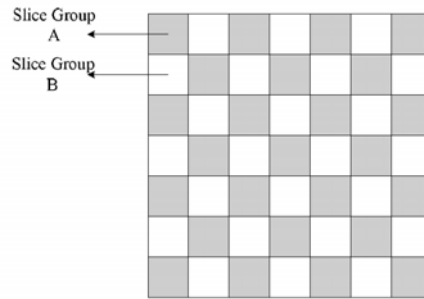
(یک چهارم) ضرایب در کدینگ دو (چهار) توصیفه با QP کوچک کوانتایز می‌شوند و بقیه با QP بالا کد می‌شوند. یک نمونه مقاله [55] است که در شکل ۲-۱۴ نشان داده شده است. در این کار، ضرایبی که قرار است با QP بالا کد شوند را با MDSQ کوانتایز می‌کنند. به این ترتیب از افزودن منتهج از وجود نرخ پایین‌ها در همه توصیفه‌ها، نه تنها در دیکودر جانبی، بلکه در دیکودر مرکزی هم استفاده می‌شود.



شکل ۲-۱۴ MDC چند نرخه حوزه فرکانس [55] (a) انکودر و (b) دیکودر

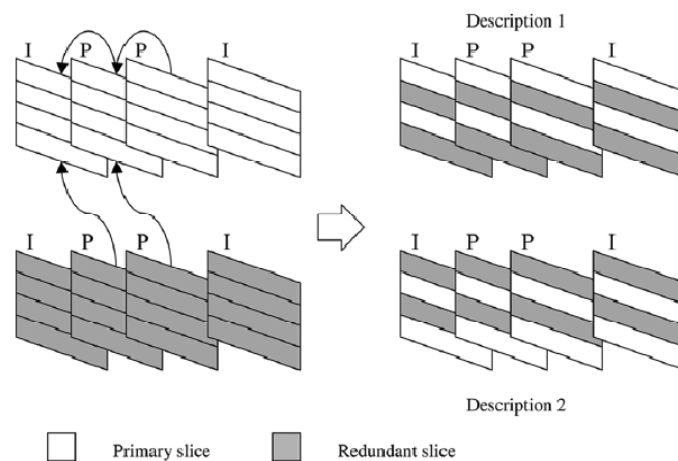
در MDC چند نرخه هنگامیکه بطور بلوکی اعمال می‌شود، در حالت کدینگ دو توصیفه، بلوکها یک در میان در توصیفه‌ها، نرخ بالا و نرخ پایین کد می‌شوند. به عبارت دیگر همه بلوکها در همه توصیفه‌ها هستند ولی یک بلوک فقط در یک توصیفه نرخ بالا کد می‌شود و در سایر توصیفه‌ها نرخ پایین است. این ایده در مقاله [56] ارائه شد که در آن یک فریم به دو SG^1 ، SGA و SGB تقسیم شده است. SGA در توصیف اول با QP پایین کد می‌شود و SGB با QP بالا، در توصیف دوم عکس این حالت است. به علاوه در این کار بردارهای حرکت بلوکهای نرخ پایین از بردارهای حرکت بلوکهای نرخ بالا بدست می‌آیند (بردار حرکتی برای آنها ارسال نمی‌شود) که باعث صرفه جویی در تعداد بیت می‌شود. همین کار با کمی اصلاحات در مقاله [57] هم انجام شده است.

¹ Slice-Group



شکل ۲-۱۵ در توصیف اول، MBهای گروه A در نرخ بالا و MBهای گروه B در نرخ پایین کد می شوند، و برعکس در توصیف دوم [56].

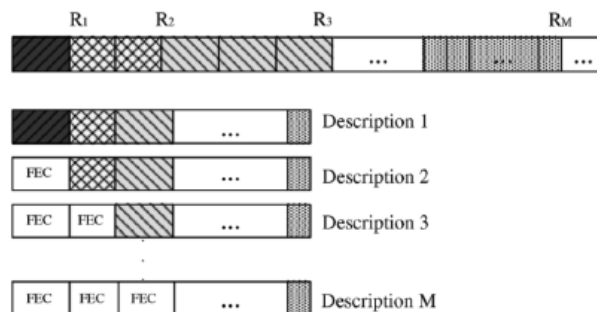
بلوکهای نرخ پایین را می توان به عنوان اسلایس افزونه یا مزاد هم کد کرد [58]. اسلایس مزاد یکی از گزینه های H.264/AVC است که به هنگام از دست رفتن اسلایسهای اصلی جایگزین آنها می شود. به ازای هر اسلایس اصلی، استاندارد اجازه استفاده از اسلایس افزونه را می دهد که همان اطلاعات اسلایس اصلی را دارد ولی با نرخ پایبتری کد شده است. اسلایسهای افزونه را می توان به همین منظور در MDC هم مورد استفاده قرار داد. کار ارائه شده در [58]، QP اعمال شده به اسلایسهای افزونه را با در نظر گرفتن وضعیت کانال و دریافت، بطور بهینه بدست آورده است. همچنین با در نظر گرفتن نقش هر اسلایس در انتشار خطا، میزان افزونگی را می توان باز هم بهینه تر کرد، این کار در [59] ارائه شده است.



شکل ۲-۱۶ MDC به کمک اسلایسهای افزونگی [58]

۴.۱.۲. MDC حوزه فشرده شده

همانطور که از نام آن می‌توان فهمید، این MDC به داده‌های فشرده شده، یعنی بعد از عمل فشرده‌سازی اعمال می‌شود. ایده اصلی این است که یک لایه را به K بخش تقسیم می‌کنیم و سپس $N-K$ کد FEC به آن اضافه می‌کنیم. حال این N بخش را با تقسیم بندی، به چندین توصیف MDC تبدیل می‌کنیم. می‌دانیم که با هر مجموعه K تایی از این N تا، می‌توان K داده اصلی را بازسازی کرد. این ایده از [60] الهام گرفته شده است که بعد از آن به منظور MDC در [61] و [62] بکار گرفته شد. چنانچه شکل ۲-۱۷ نشان می‌دهد، ویدئو یا تصویر ابتدا بصورت مقیاس‌پذیر^۱ کد می‌شود، سپس لایه l ام به l بخش تقسیم می‌شود و برای هر لایه کدهای $RS(N, l)$ تولید می‌شود (N تعداد توصیفها است). از هر لایه یک بخش (داده اصلی یا کد FEC) برداشته، کنار هم قرار داده و در یک استریم به عنوان یک توصیف ارسال می‌شود. کار ارائه شده در [61] هم چنین کاری را برای بایتها، و نه لایه‌های استریم مقیاس‌پذیر، انجام می‌دهد.



شکل ۲-۱۷ MDC-FEC با تعداد مساوی توصیفها و لایه‌های SVC

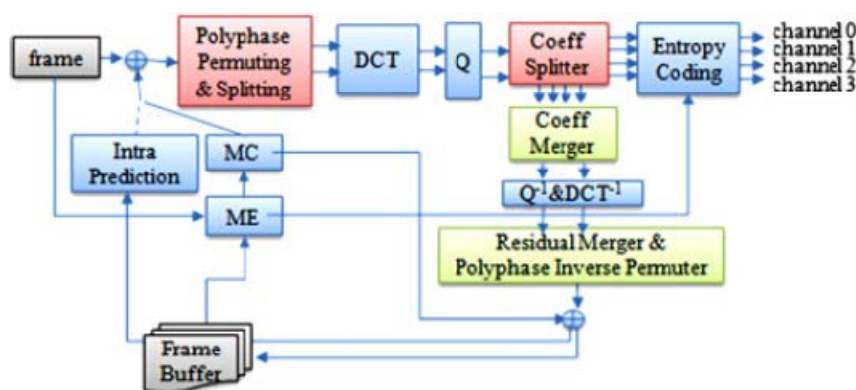
در حالت کلی، تعداد بخش‌بندیها در هر لایه مصالحه بین افزونگی و خطاپذیری را تعیین می‌کند. هر قدر تعداد بخش‌ها کمتر باشد، مقدار افزونگی بیشتری تخصیص داده می‌شود؛ با این ایده می‌توان MDC را برای کارایی بهینه در شرایط مختلف کانال تنظیم کرد [63]. این الگوریتم در واقع transcoding است، یعنی استریم

¹ Scalable

کد شده را به منظور انطباق بیشتر با شرایط، به گونه دیگر کد می‌کنند، در حالیکه داده خام اولیه در اختیار نیست. تعداد بخش‌ها در هر لایه با تابع بهینه‌سازی تعیین می‌شود.

۵.۱.۲. MDC حوزه هایبرید

در MDC حوزه هایبرید، MDC در دو حوزه از مواردی که اشاره شد انجام می‌شود. این روش بخصوص زمانی کارا است که تعداد بالای توصیفها را می‌خواهیم داشته باشیم که کار در یک حوزه تنها و با بعضی روشها مناسب نیست. به عنوان مثال، تولید چهار توصیف و استفاده از MDC های حوزه زمان روش خوبی نیست، چون دیکودر جانبی در صورتیکه یک توصیف در دسترس باشد باید سه توصیف را تخمین بزند، یعنی از یک فریم باید سه فریم دیگر ساخته شود. همانند [64] میتوان حوزه زمان و spacial را همزمان اعمال کرد و یا همانند [65] و [66] ترکیبی از spacial و فرکانس (اولی تقسیم ضرایب و دومی MDSQ) را داشت؛ که در شکل ۲-۱۸ نشان داده شده است. فایده اصلی کار کردن در دو حوزه، امکان بیشتر برای حفظ همبستگی بین توصیفها است، که به اختفای خطا و در نتیجه کیفیت جانبی بیشتر کمک می‌کند.



شکل ۲-۱۸ MDC حوزه هایبرید: spatial و فرکانس [65]

۶.۱.۲. روشهایی که مبتنی بر تقسیم بندی نیستند

در این گروه، دو توصیف به گونه‌ای تشکیل می‌شوند که خطای بازسازی آنها، ناهمبسته باشد، در این حالت دو توصیف را می‌توان طوری با هم ادغام کرد که خطای بازسازی نهایی در دیکودر مرکزی کمتر شود. هر چقدر خطاها، ناهمبسته‌تر باشند، مقدار اصلی پیکسلها را با دقت بیشتری می‌توان بدست آورد. در این روش، مقدار افزونگی توصیفها بالاست و آن را نمی‌توان به اندازه دلخواه کوچک کرد و لذا برای شرایطی که شبکه کمتر نویزی است مناسب نمی‌باشد.

توصیفهای این روش را می‌توان با بکار بردن پارامترهای انکدینگ متفاوت مانند مرجع پیش‌بینی در تخمین حرکت، جهت تخمین حرکت و یا QPهای مختلف، تولید کرد که در [67] انجام شده است. و یا در [68] که فریمهای هر دو توصیف با QPها و مراجع پیش‌بینی متفاوت کد می‌شوند. فریم بازسازی شده در دیکودر مرکزی متوسط وزن‌دار فریمهای هر توصیف خواهد بود؛ وزن‌ها با استفاده از تئوری تخمین، در انکودر تولید و همراه توصیفها ارسال می‌شوند. در این مقاله، برای کاهش بیت مصروف برای ارسال وزنها، آنها را کوانتایز و بعد ارسال می‌کنند.

به منظور مستقل کردن خطای بازسازی در توصیفها، یک روش دیگر تغییر دادن مرز بلوکها نسبت به هم در توصیفها است. یعنی اینکه، بلوک‌بندی در توصیفها با یکدیگر متفاوت است. تخمین و جبران حرکت در انکودرها بلوک به بلوک صورت می‌گیرد و تغییر مرزها به معنی تغییر در تصویر اختلاف است و لذا خطای کوانتیزیشن هم یکسان نخواهد بود که همین نایکسانی در دیکودر مرکزی مورد بهره برداری قرار می‌گیرد. این ایده در [69] ارائه شده است. با استفاده از تبدیل‌های متفاوت هم می‌توان توصیفها را درست کرد؛ مثلاً در [70]، یک توصیف با DCT تولید می‌شود و در توصیف دیگر هیچ تبدیلی استفاده نمی‌شود (یا تبدیل همانی استفاده می‌شود). گرچه به نظر می‌رسد یک جفت تبدیل با کارایی بهتر هم بتوان یافت.

۲.۲. دسته بندی روشهای MDC بر اساس کارایی آنها

در این بخش، سه عامل که در انتخاب یک MDC مناسب برای کاربرد خاص باید مورد ملاحظه قرار گیرد را مطالعه می‌کنیم. این عوامل، قابلیت تنظیم افزونگی، پیچیدگی محاسباتی و امکان افزایش تعداد توصیفها هستند. بر پایه این عوامل، عملکرد MDC های معرفی شده در جدول ۲-۱ مقایسه می‌شود.

۱.۲.۲. قابلیت تنظیم افزونگی

همانطور که قبلاً هم گفته شد، در MDC بعضی اطلاعات مشترک، علاوه بر اطلاعات سرآمد، باید در همه توصیفها وجود داشته باشند. این اطلاعات برای اینکه هر توصیف جداگانه قابل دیکود کردن باشد نیاز است. در کانالهای کم نویز، بیشتر اوقات همه توصیفها سالم هستند و کمتر پیش می‌آید که بخواهیم یک توصیف را دیکود کنیم، بنابراین نیازی به افزونگی بالا نیست. از سوی دیگر، در شبکه با نرخ خطای بالا احتمال استفاده از دیکودر جانبی بیشتر است و باید که کیفیت جانبی خوبی داشته باشیم که این خود به معنی داشتن افزونگی بالا است. بنابراین، کنترل مقدار افزونگی بطور پیوسته و در رنج وسیع یک ویژگی مهم است که برای داشتن کیفیت بهینه در شرایط مختلف کانال سودمند است. در جدول ۲-۱، سه سطح برای مقایسه MDC ها تعیین شده است، "پایین"، "متوسط" و "بالا" که قابلیت آنها را در تنظیم افزونگی نشان می‌دهد.

در کاربردها یا شبکه‌هایی که وضعیت کانال در نوسان است، استفاده از MDC های با قابلیت افزونگی بالا توصیه می‌شود و گرنه سایر MDC ها را می‌توان بکار برد.

۲.۲.۲. پیچیدگی محاسباتی

پیچیدگی محاسباتی هم موضوع قابل بحثی است که بویژه در کاربردهای بی درنگ اهمیت پیدا می‌کند. اگرچه توان پردازشی پردازنده‌ها روز به روز افزایش می‌یابد، نیاز به پردازشهای بیشتر هم فزونی می‌یابد. به عنوان مثال، استاندارد H.264/AVC حدوداً ده برابر حجم محاسبات بیشتری نسبت به استانداردهای قبل از

خود دارد [71]، و همچنین رزولوشن زمانی و مکانی ویدئوهای امروزی نسبت به قبل تقریباً دو برابر گشته است. از سوی دیگر، با رشد روز افزون ادوات باطری خور، هنوز هم الگوریتمهای با حجم محاسبات پایین مد نظر می‌باشند. در جدول ۱-۲، MDC هایی که پیچیدگی محاسباتی آنها به اندازه SDC است با سطح "پایین" مشخص شده‌اند، آنهایی که حجم محاسبات با تعداد توصیفها افزایش می‌یابد با سطح "بالا" و سایرین با "متوسط" نشان داده شده‌اند.

برای کاربردهایی که نیاز دارند بی درنگ کد کنند و یا با منابع HD و ارسال بیدرنگ آنها، مانند¹ HDVC سروکار دارند، MDC های با سطح "پایین" پیشنهاد می‌شوند.

۳.۲.۲. امکان افزایش تعداد توصیفها

امکان افزایش تعداد توصیفها در یک روش MDC هم قابلیت ویژه‌ای است. گرچه کدینگ دو-توصیفه برای غلبه بر کانالهای نویزی در بیشتر موارد مشکل‌گشا است و کانالها غالباً آنقدر نویزی نیستند که توصیفهای بیشتر از دو توجیه شود، اما دلایل دیگری برای تعداد توصیفهای بیشتر وجود دارد. برای نمونه، در ارسال ویدئوی P2P، به منظور فراهم کردن نرخها و رزولوشنهای متعدد برای تغذیه کردن گیرنده‌های مختلف، ممکن است هشت و یا شانزده توصیف نیاز باشد. به علاوه، MDC گاهی اوقات با انتقال چند مسیر (MPT) ترکیب می‌شود که هر توصیف روی یک مسیر جداگانه ارسال می‌شود. برخی از خواص MPT مانند پهنای باند مجموع، احتمال قطعی و تغییرات در تاخیر با تعداد مسیرها و بنابراین تعداد توصیفها بهبود می‌یابد [72]. در جدول، "بالا" نشان دهنده این است که تولید تعداد توصیفهای بیشتر از چهار تا برای روش MDC آسان است و بر عکس برای آنهایی که ساختن چهار توصیف آسان نیست و یا توصیه نمی‌شود با سطح "پایین" نشان داده شده‌اند؛ سایرین که بین این دو هستند با سطح "متوسط" بیان شده‌اند.

¹ High Definition Video Conferencing

در کاربردهایی که با نرخ خطاهای بالا مواجه هستند، و یا P2P و MPT، چنانچه در بالا توضیح داده شد، MDC هایی که براحتی قادر به تولید تعداد توصیفهای بالا هستند مناسبتر هستند و گزینه سایر MDC ها را می توان بکار برد.

جدول ۲-۱ دسته بندی و مقایسه MDC ها از نظر حوزه و کارایی

حوزه MDC	مراجع	قابلیت تنظیم افزونگی	پیچیدگی محاسباتی	امکان افزایش تعداد توصیفها	خلاصه روش
حوزه پردازش	[18] [19] [20]	بالا	متوسط	متوسط	zero-padding در حوزه فرکانس معادل upsampling در حوزه پیکسل است که به این ترتیب می توان افزونگی اضافه کرد.
	[20]	بالا	بالا	پایین	کپی کردن داده هایی که کمتر قابل پیش بینی هستند به دیکودر جانبی برای داشتن کیفیت بهتر کمک می کند.
	[21]	پایین	بالا	متوسط	با داشتن این فیلترها که در انکودر طراحی می شوند، می توان توصیفها را از یکدیگر با دقت بیشتری بدست آورد.
	[23] [24]	بالا	متوسط	پایین	توصیف دوم (اول) از توصیف اول (دوم) پیش بینی شده و سیگنال باقیمانده در توصیف اول (دوم) قرار می گیرد. کیفیت جانبی بهتر می شود.
	[25] [26]	متوسط	بالا	پایین	افزونگی بطور هدفمند به نقاط خاصی از تصویر اضافه می شود.
	[27]	بالا	متوسط	متوسط	نرخ پایین توصیف اول (دوم) در توصیف دوم (اول) قرار می گیرد.
	[29] [30]	متوسط	پایین	پایین	با داشتن اطلاعات حرکت، اختفای خطا خیلی بهتر صورت می گیرد و لذا کیفیت جانبی بهتری خواهیم داشت.
حوزه زمان MDC	[32]	متوسط	بالا	پایین	فریمهایی که براحتی قابل بازیابی نیستند در هر دو
					کپی کردن داده های کمتر قابل بازیابی

توصیف کپی می‌شوند.						
نرخ پایین فریمهای توصیف اول (دوم) در توصیف دوم (اول) قرار می‌گیرد.	متوسط	متوسط	بالا	[33] [34]	MDC چند نرخه	
به این ترتیب می‌توان حرکت بین فریمها را هموار کرد که به تخمین فریمهای دچار خطا شده کمک می‌کند.	پایین	بالا	متوسط	[35] [36]	کم و اضافه کردن تعداد فریمها	
خروجی کوانتایزر اسکالر به گونه ای است که در با ترکیب آنها در دیکودر مرکزی نویز کوانتیزیشن کمتر می‌شود.	پایین	پایین	بالا	[37] [38] [39] [40]	MDSQ	حوزه فرکانس
ضرایب مهم در هر دو توصیف تکرار می‌شوند ولی سایرین بین آنها تقسیم می‌شوند.	متوسط	پایین	بالا	[41] [42] [43] [44][45]	تقسیم کردن ضرایب	
ضرایب DCT قبل از اینکه بین توصیفها تقسیم شوند، توسط یک تبدیل دیگر همبسته می‌شوند. در دیکودر جانبی، بخاطر این همبستگی ایجاد شده، ضرایب نرسیده را می‌توان از ضرایب در دسترس تخمین زد.	متوسط	بالا	متوسط (بجز [52] که بالا است)	[46] [47] [48] [49] [50] [51] [52] [53] [54]	MDTC	
مشابه با MDC با تقسیم کردن ضرایب است ولی نرخ پایین ضرایب توصیف دوم (اول) در توصیف اول (دوم) قرار می‌گیرد.	متوسط	متوسط	بالا	[55] [56] [57] [58] [59]	MDC چند نرخه	
ویدئو بصورت لایه ای کد می‌شود، به لایه ها بر اساس اهمیت آنها کد FEC اضافه می‌شود، سپس توصیفها با ترکیبی از همه لایه ها تشکیل می‌شوند.	بالا	بالا	متوسط	[61] [62] [63]	MDC-FEC	
MDC دو بار اعمال می‌شود، مثلاً ابتدا در حوزه Spatial توصیفها را تولید می‌کنیم، سپس هر توصیف دوباره MDC زمانی می‌شود.	بالا	پایین	پایین	[64]	زمانی + spatial	حوزه هایبرد
	بالا	پایین	متوسط	[65] [66]	Spatial + فرکانس	

توصیفها به گونه ای تولید می‌شوند که خطای فشرده سازی در آنها مستقل باشد. به این ترتیب با در دست داشتن دو استریم با نوبزهای مختلف، به کمک تخمین می‌توان تا حدودی نوبز را کمتر کرد.	پایین	بالا	پایین	[67] [68] [69] [70]	تولید دو استریم با خطای فشرده سازی حتی المقدور مستقل	روشهایی که مبتنی بر تقسیم بندی نیستند
--	-------	------	-------	------------------------	--	---------------------------------------

۳.۲. MDC کدام حوزه؟

در بخش پیشین، روشهای MDC موجود که روی منابع ویدئو و در استانداردهای فشرده سازی قابل پیاده سازی بودند، مرور شدند. در جدول ۱-۲ کارایی آنها با معیارهایی چون تطبیق پذیری افزونگی، پیچیدگی و قابلیت افزایش تعداد توصیفها مقایسه شد. اما در مورد اینکه برای یک کاربرد خاص، MDC کدام حوزه مناسب تر است بحثی نشد، چونکه این بحث بیشتر کیفی و شهودی است و مشکل بتوان با اندازه‌گیری و بدون در نظر گرفتن ویژگیهای ویدئو و کاربرد مد نظر به قاعده خاصی در این مورد رسید. برای مثال اگر بخواهیم بدانیم MDC حوزه مکانی برای یک کاربرد خاص مفید است یا نه، با این مسائل مواجه هستیم: آیا داشتن تصویر کوچکتر از تصویر اصلی، اساساً در گیرنده یا گیرنده‌ها مهم است یا خیر؟ مثلاً در حالتیکه بیشتر گیرنده‌ها صفحه نمایش کوچکی دارند ارسال تصاویر با اندازه بزرگ کمتر ضرورت دارد. یا اینکه کاربرد به گونه ای است گیرنده‌ها تصویر را به اندازه واقعی می‌خواهند، مثلاً اگر تصویر اصلی کوچک باشد، با روش MDC مکانی تصویر باز هم کوچکتر می‌شود که شاید قابل قبول نباشد. آیا محتوای تصویر به گونه‌ای است که داشتن یک زیرتصویر از آن کاملاً به چشم آید؟ پهنای باند در اختیار چقدر است؟ آیا کد کردن در این نرخ، امکان داشتن بافت تصویر با کیفیت خوب را به ما می‌دهد یا خیر؟ در مورد MDC زمانی و یا سایرین هم به همین ترتیب می‌توان بحث کرد که آیا مناسب کاربرد مد نظر هستند یا خیر.

همانطور که دیده شد مباحث کاملاً کیفی است. گرچه در مقاله‌هایی مانند [74] و [73] نوع MDC را با توجه به شرایط کانال بطور بهینه پیدا می‌کنند ولی معمولاً اینطور است که از قبل نوع MDC انتخاب می‌شود. چنین

پرسشی در انتخاب نوع SVC هم مطرح است؛ در آنجا هم با همین مباحث، SVC مناسب که می‌تواند مکانی، زمانی و یا SNR باشد انتخاب می‌گردد که لایه های بهبود منتج به افزایش کیفیت (رزولوشن) به ترتیب مکانی، زمانی و ¹PSNR می‌شوند. بنابراین زیاد اتفاق می‌افتد که به دلایلی نیاز به حفظ رزولوشن مکانی و زمانی ویدئو در دیکودرهای جانبی است و MDC های حوزه فرکانس ترجیح داده شوند.

MDC ارائه شده در این رساله، یک MDC حوزه فرکانس است که بر اساس ترکیب لایه های پایه و بهبود در CGS کار می‌کند. خواهیم دید که روش ارائه شده نسبت به MDC های مرسوم، مصالحه بهتری بین کیفیت مرکزی و کیفیت جانبی فراهم می‌کند و در نتیجه، کیفیت متوسط بالاتری به هنگام ارسال ویدئو در محیطهای نویزی، در طرف گیرنده خواهیم داشت.

¹ Peak Signal to Noise Ratio

۳. روش MLMDC

۱.۳. مقدمه

در فصل پیش شیوه‌های MDC با جزئیات مرور شدند. همانطور که گفته شد ویژگی مشترک MDC ها این است که بین کیفیت مرکزی و کیفیت جانبی بده بستان وجود دارد؛ یعنی اینکه، در نرخ بیت ثابت، با افزایش کیفیت جانبی، کیفیت مرکزی کاهش می‌یابد و برعکس. به عنوان مثال اگر به کیفیتی جانبی برابر SDC نیاز باشد مجبوریم که همه‌ی ضرایب DCT را (در روش‌های حوزه DCT) در هر دو استریم تکرار کنیم ولی در حالتی که هر دو استریم در دسترس هستند یکی از آنها به کلی زیادی است و کاربردی ندارد لذا کیفیت مرکزی هم در حدود کیفیت جانبی می‌شود. در این فصل روشی معرفی می‌شود (MLMDC) که کیفیت جانبی در حدود SDC دارد اما کیفیت مرکزی آن تا حد زیادی بهتر است. به عبارت دیگر، مصالحه بهتری بین کیفیت مرکزی و کیفیت جانبی برقرار می‌کند. خواهیم دید که این امر مستقیماً منجر به کیفیت متوسط بهتر ویدئو به هنگام ارسال چند مسیره و در محیط‌های نویزی می‌شود.

در ادامه جزئیات MLMDC تشریح شده، رفتار نرخ-اعوجاج آن بررسی شده و در انتها عملکرد آن مورد مقایسه قرار گرفته است.

۲.۳. معرفی روش پیشنهادی

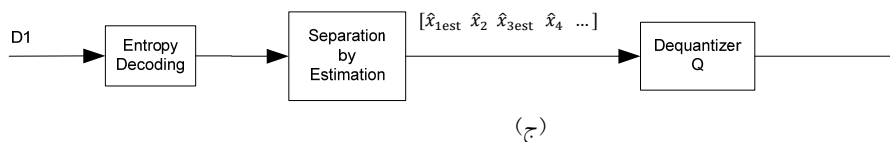
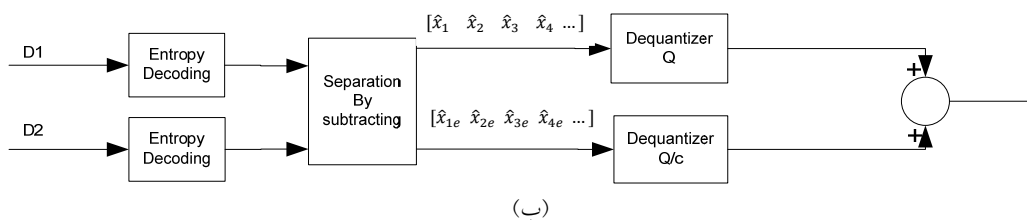
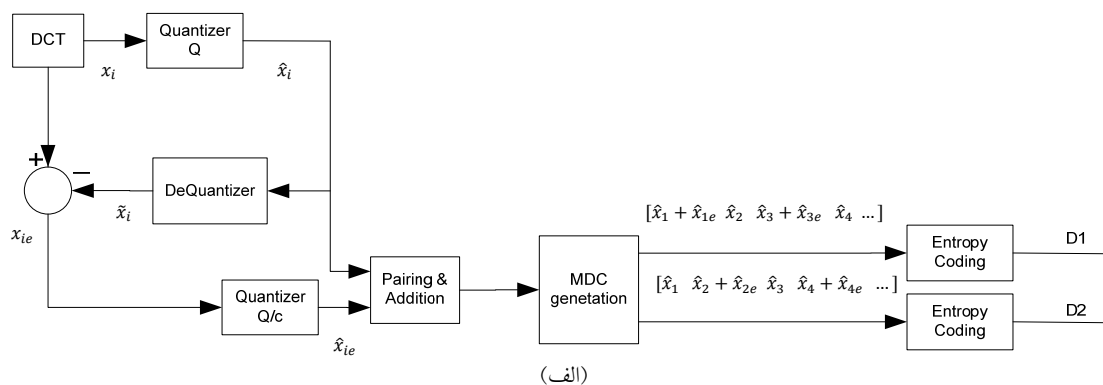
۱.۲.۳. انکودر

MLMDC از روش MDTC [رجوع شود به بخش ۳.۱.۲] گرفته شده با این تفاوت که به جای ترکیب ضرایب مختلف DCT، هر ضریب پایه (BC^1) با ضریب بهبود (EC^2) ترکیب شده است. همانطور که در شکل ۳-۱ (الف)، بلوک دیاگرام انکودر MLMDC، نشان داده شده است، ضریب BC از کوانتیزیشن مستقیم ضرایب DCT ساخته می‌شود و از کوانتیزه کردن اختلاف بین مقدار اولیه DCT و مقدار بازسازی شده آن، ضرایب EC بدست می‌آید. ضرایب ترکیبی (CC^3) از مجموع ضرایب BC و EC حاصل شده و به طور یک در میان در توصیف‌ها قرار می‌گیرند؛ به عبارت دیگر اگر ضریب اول DCT در توصیف اول، ترکیبی شود در توصیف دوم نیازی به این کار نیست و خود ضریب پایه ارسال می‌شود. به این ترتیب تعداد ضرایبی که ترکیبی هستند و همچنین افزونگی بیت ناشی از ترکیب شدن بین دو توصیف متعادل می‌شود. این امر در شکل ۳-۱ نشان داده شده است. در حالت کلی کافی است که هر CC فقط در یک توصیف قرار گیرد که در این صورت هزینه‌های ناشی از ترکیب آن ضریب DCT فقط در یک توصیف وجود خواهد داشت. به عنوان مثال، در فشرده‌سازی ۴-توصیفه، هر توصیف شامل یک‌چهارم کل ضرایب ترکیبی خواهد بود؛ به همین دلیل کارایی روش MLMDC با افزایش تعداد توصیف‌ها بهتر می‌شود.

¹ Base Coefficient

² Enhancement Coefficient

³ Combined Coefficients



شکل ۳-۱ بلوک دیاگرام روش پیشنهادی (الف) انکودر (ب) دیکودر مرکزی و (ج) دیکودر جانبی

در شکل ۳-۱، ضرایب BC و EC متعلق به یک ضریب DCT هستند ولی می‌تواند چنین نباشد، در این مورد بعداً بیشتر بحث خواهد شد.

فرض کنید x_i یک ضریب DCT باشد، که بعد از کوانتیزیشن با Q ، \hat{x}_i را به ما می‌دهد، ضریب بازسازی شده \tilde{x}_i است. با کم کردن \hat{x}_i از x_i نویز کوانتیزیشن بدست می‌آید که برای کوانتیزه کردن مجدد آن نیاز به Q کوچکتری داریم. به این منظور، کوانتیزیشن دوم با Q/c صورت می‌گیرد، c پارامتر MLMDC است. هر اندازه c بزرگتر باشد، نویز کوانتیزیشن بهبود کوچکتر می‌شود و در نتیجه کیفیت مرکزی بالاتری خواهیم داشت و از طرف دیگر خطای تخمین و افزونگی بیت هم بالاتر است. چگونگی این تغییرات با استخراج مدل ریاضی و شبیه سازی رفتار آن، در ادامه نشان داده خواهد شد.

فرض کنید که یک سیگنال از دو ضریب تشکیل شده باشد، $\{x_1, x_2\}$ که باید با MDC کد شود. ضرایب $\{x_1, x_2\}$ با Q کوانتیزه می‌شود و ضرایب $\{\hat{x}_1, \hat{x}_2\}$ بدست می‌آید. خطای کوانتیزیشن منتج $\{x_{1e}, x_{2e}\}$ دوباره با Q/c کوانتیزه شده و $\{\hat{x}_{1e}, \hat{x}_{2e}\}$ بدست می‌آیند، با جمع شدن ضرایب پایه و ضرایب بهبود، ضرایب ترکیبی حاصل می‌شوند. با کوانتیزیشن استفاده شده در H.264/AVC که در معادله (۱-۱) داده شده، نحوه تشکیل ضرایب BC و EC و CC در معادله زیر آورده شده است:

$$\begin{aligned} \text{BC: } \hat{x}_i &= \text{sign}(x_i) \left[\frac{|x_i|}{Q} + f \right], \quad \tilde{x}_i = Q \hat{x}_i \\ x_{ie} &= x_i - \tilde{x}_i \\ \text{EC: } \hat{x}_{ie} &= \text{sign}(x_{ie}) \left[\frac{|x_{ie}|}{Q/c} + f \right] \\ \text{CC: } \hat{z}_{ij} &= \hat{x}_i + \hat{x}_{je} \\ & i, j = 1, 2 \end{aligned} \quad (1-3)$$

دو راه ممکن برای ترکیب کردن وجود دارد: یکی اینکه ضرایب ترکیب شونده متعلق به یک ضریب DCT باشند و دیگری اینکه به دو ضریب مختلف. معادلات (۳-۲) و (۳-۳) این دو حالت را که به ترتیب CaseI و CaseII نامیده شده‌اند، نشان می‌دهد.

$$\text{CaseI: } \begin{cases} D_1 : \{(\hat{x}_1 + \hat{x}_{1e}), (\hat{x}_2)\} \\ D_2 : \{(\hat{x}_1), (\hat{x}_2 + \hat{x}_{2e})\} \end{cases} \quad (2-3)$$

$$\text{CaseII: } \begin{cases} D_1 : \{(\hat{x}_1 + \hat{x}_{2e}), (\hat{x}_2)\} \\ D_2 : \{(\hat{x}_1), (\hat{x}_2 + \hat{x}_{1e})\} \end{cases} \quad (3-3)$$

به این ترتیب، توصیفهای D_1 و D_2 پس از آنروپی انکدینگ به گیرنده ارسال می‌شوند.

۲.۲.۳. دیکودر مرکزی

در دیکودر مرکزی، هر دو توصیف در دسترس هستند و به راحتی می‌توان EC را از BC جدا نمود و بازسازی دو لایه‌ای (پایه و بهبود) انجام داد. معادله (۳-۴) این روند را نشان می‌دهد.

$$\begin{aligned}\hat{x}_{je} &= \hat{z}_{ij} - \hat{x}_i \\ x_{cen_j} &= Q\hat{x}_j + Q\hat{x}_{je}/c \\ D_{cen_j} &= E\left[(x_j - x_{cen_j})^2\right]\end{aligned}\quad (۳-۴)$$

که در آن x_{cen} و D_{cen} ، ضرایب بازسازی شده و اعوجاج حاصل در دیکودر مرکزی هستند. توجه شود که از نظر دیکودر مرکزی، هر دو Case I و Case II مشابه هستند.

۳.۲.۳. دیکودر جانبی

در دیکودر جانبی می‌بایست ضریب پایه از ضریب ترکیبی تخمین زده شود، مثلاً در Case I، اگر فقط D_1 سالم دریافت شده باشد، \hat{x}_1 در اختیار نیست و از $\hat{x}_1 + \hat{x}_{1e}$ تخمین زده می‌شود. برای بحث راحت‌تر، مستقل از اینکه کدام حالت (Case I یا Case II) استفاده شود، \hat{z} را به صورت زیر تعریف می‌کنیم.

$$\hat{z} = \hat{x}_b + \hat{x}_e \quad (۳-۵)$$

که در آن \hat{x}_b و \hat{x}_e به ترتیب ضرایب پایه و ضرایب بهبود هستند. از تئوری تخمین، می‌دانیم که بهترین تخمین \hat{x}_b از \hat{z} از رابطه زیر بدست می‌آید [75]:

$$\hat{x}_{est} = E\{\hat{x}_b|\hat{z}\} \quad (۳-۶)$$

که اگر توابع توزیع \hat{x}_b و \hat{x}_e را داشته باشیم، خواهیم داشت:

$$\hat{x}_{est} = \sum_{\hat{x}_b} \hat{x}_b P_{\hat{x}_b|\hat{z}}(\hat{x}_b|\hat{z}) \quad (۳-۷)$$

$$P_{\hat{x}_b|\hat{z}}(\hat{x}_b|\hat{z}) = \frac{P_{\hat{x}_b,\hat{z}}(\hat{x}_b,\hat{z})}{P_{\hat{z}}(\hat{z})} = \frac{P_{\hat{x}_b}(\hat{x}_b)P_{\hat{z}|\hat{x}_b}(\hat{z}|\hat{x}_b)}{P_{\hat{z}}(\hat{z})} \quad (۳-۸)$$

$$P_{\hat{z}|\hat{x}_b}(\hat{z}|\hat{x}_b) = P_{\hat{x}_e|\hat{x}_b}(\hat{z} - \hat{x}_b) \quad (9-3)$$

$$\begin{aligned} p_{\hat{z}}(\hat{z}) &= \sum_{\hat{x}_b} P_{\hat{x}_b, \hat{z}}(\hat{x}_b, \hat{z}) = \sum_{\hat{x}_b} P_{\hat{x}_b}(\hat{x}_b) P_{\hat{z}|\hat{x}_b}(\hat{z}|\hat{x}_b) \\ &= \sum_{\hat{x}_b} P_{\hat{x}_b}(\hat{x}_b) P_{\hat{x}_e|\hat{x}_b}(\hat{z} - \hat{x}_b) \end{aligned} \quad (10-3)$$

و در نتیجه

$$\hat{x}_{est} = \frac{\sum_{\hat{x}_b} \hat{x}_b P_{\hat{x}_b}(\hat{x}_b) P_{\hat{x}_e|\hat{x}_b}(\hat{z} - \hat{x}_b)}{\sum_{\hat{x}_b} P_{\hat{x}_b}(\hat{x}_b) P_{\hat{x}_e|\hat{x}_b}(\hat{z} - \hat{x}_b)} \quad (11-3)$$

مشاهده می‌شود که فرمول نهایی تابعی است از توزیع‌های جداگانه \hat{x}_e و \hat{x}_b و همچنین توزیع مشترک¹ آنها. از طرف دیگر \hat{x}_b ، \hat{x}_e در Case I وابسته و در Case II مستقلند، یعنی اینکه توابع توزیع پیوندی آنها در Case I و Case II متفاوت است؛ به عبارت دیگر به لحاظ دیکودر جانبی، Case I و Case II مشابه نیستند.

۳.۳. اعوجاج و نرخ در MLMDC

۱.۳.۳. پیش زمینه

MLMDC یک روش حوزه DCT است، یعنی اینکه پردازش‌ها در حوزه DCT صورت می‌گیرد. بنابراین مرور مختصری بر مفاهیم مربوط به ضرایب DCT و کوانتیزیشن آنها که مرتبط با MLMDC است ارائه می‌شود.

سه مدل برای توزیع ضرایب DCT ارائه شده است: گوسین [76]، لاپلاسین [77] و کوشی [78] که به همین ترتیب دقیقتر هم می‌شوند. اما مدل لاپلاسین، متداول‌تر است و از طرف دیگر فرم نمایی آن برای کارهای تحلیلی مناسب‌تر است. بنابراین، ما هم از این مدل استفاده می‌کنیم.

¹ Joint distribution

$$f_b(x_b) = \frac{\lambda}{2} e^{-\lambda|x_b|} \quad (12-3)$$

$$\sigma^2 = \frac{2}{\lambda^2}$$

که در آن λ و σ به ترتیب پارامتر لاپلاس و واریانس سیگنال x_b هستند. هر مولفه فرکانسی بلوک DCT، یک توزیع لاپلاسی با میانگین صفر و λ خاص خود دارد. مولفه‌های فرکانس پایین، واریانس بیشتری دارند و بنابراین λ کوچکتر، این خاصیت فشرده‌کنندگی تبدیل DCT است که بیشترین انرژی سیگنال را در مولفه‌های فرکانس پایین جمع می‌کند. بنابراین مولفه‌های فرکانس پایین نقش بسزایی هم در تعداد بیت و هم در اعوجاج دارند.

ضرایب DCT با قاعده گفته شده در فرمول (1-1) کوانتیزه می‌شوند. Sullivan نشان داد که برای یک منبع لاپلاسی، این نحوه کوانتیزیشن تقریباً بهینه است [80] [79]. در کوانتیزیشن نرمال¹، f برابر 1/6 برای بلوکهای ایتر و 1/3 برای بلوکهای ایتر قرار داده می‌شود. در حالت تطبیقی²، f به گونه‌ای تنظیم می‌شود که سطح‌های کوانتیزیشن غیرصفر در مرکز ثقل نواحی کوانتیزیشن قرار گیرند، با این روش تا حدود 1dB کیفیت بالاتری خواهیم داشت [81]. برای منابع لاپلاسی، این f برای تمام نواحی کوانتیزیشن یکسان است و کوچکتر از 1/2 اما با λ و Q تغییر می‌کند.

$$Q\hat{x}_b = E[x|\hat{x}_b=N] \quad (13-3)$$

$$N \neq 0$$

ناگفته نماند که برای کوانتیزاسیون به فرم (1-1) مینیمم اعوجاج در $f = 1/2$ بدست می‌آید ولی 1/2 از نظر نرخ بیت بهینه نیست؛ با $f < 1/2$ ناحیه کوانتیزاسیون حول صفر وسیعتر شده و در نرخ صرفه‌جویی بیشتری می‌شود به گونه‌ای که منحنی نرخ-اعوجاج نسبت به $f = 1/2$ بهتر خواهد بود.

¹ Normal Quantization

² Adaptive Quantization

اعوجاج کوانتیزیشن برای یک منبع لاپلاسین قبلاً در [82] بدست آمده است:

$$D_q = \frac{\lambda Q e^{f\lambda Q} (2 + \lambda Q - 2f\lambda Q) + 2 - 2e^{\lambda Q}}{\lambda^2 (1 - e^{\lambda Q})} \quad (14-3)$$

می توان ثابت کرد که برای λQ های به اندازه کافی کوچک، D_q به D_{qu} میل می کند.

$$D_{qu} = \left(f^2 - f + \frac{1}{3}\right) Q^2 \quad (15-3)$$

D_{qu} نویز کوانتیزیشن برای یک منبع با توزیع یکنواخت است.

۲.۳.۳. اعوجاج جانبی

در دیکودر جانبی، فقط یک توصیف در اختیار است و به همین علت قادر به جداسازی لایه پایه از بهبود نیستیم. همانطور که گفته شد از تخمین استفاده می شود که تخمین هم خود اعوجاج ایجاد می کند. بنابراین در دیکودر جانبی علاوه بر اعوجاج کوانتیزیشن، اعوجاج تخمین هم داریم. اعوجاج ناشی از فقط تخمین از رابطه زیر بدست می آید:

$$D_{est} = E[(\tilde{x} - Q\hat{x}_{est})^2] \quad (16-3)$$

اما برای یافتن اعوجاج جانبی، مرجع، مقدار اولیه سیگنال است و اعوجاج هم شامل کوانتیزیشن می شود و هم شامل تخمین:

$$D_{side} = E[(x - Q\hat{x}_{est})^2] \quad (17-3)$$

در قسمت های بعد، با پیدا کردن تخمین بهینه برای هر دو Case I و Case II عملکرد آنها به لحاظ اعوجاج تخمین و اعوجاج جانبی مطالعه می شود. به این منظور نیاز است که ابتدا توابع توزیع که در رابطه (۳-۱۱) استفاده شده است بدست آید.

بعد از اینکه کوانتیزیشن روی x_b اعمال شد، تابع توزیع ضرایب پایه به این صورت بدست می آید:

$$p_b(0) = p_b(\hat{x}_b = 0) = 2 \int_0^{(1-f)Q} f_b(x_b) dx_b = 1 - e^{-(1-f)\lambda Q} = P_{b0}$$

$$p_b(n) = p_b(\hat{x}_b = n) = \int_{(n-f)Q}^{(n+1-f)Q} f_b(x_b) dx_b = \frac{(1 - e^{-\lambda Q})e^{\lambda Q f}}{2} e^{-\lambda Q |n|} \quad (18-3)$$

$$= P_{bN} e^{-\lambda Q |n|}$$

در Case I توزیع x_e با تقسیم کردن توزیع x_b به احتمال وقوع \hat{x}_b متناظر (طبق تعریف) بدست می آید:

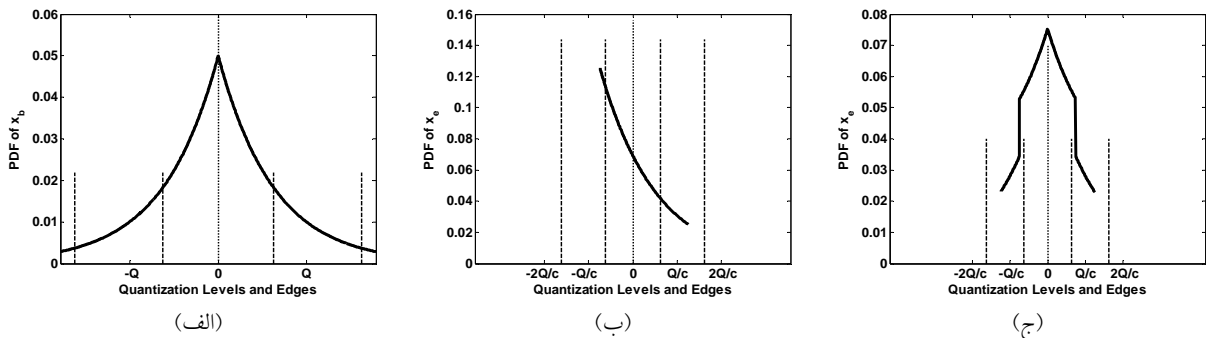
$$f_{x_e|\hat{x}_b}^I(x_e|\hat{x}_b) = \begin{cases} \text{if } \hat{x}_b > 0 & \begin{cases} (\lambda e^{-\lambda x_e})/P_{bN} & -fQ < x_e < (1-f)Q \\ 0 & \text{otherwise} \end{cases} \\ \text{if } \hat{x}_b < 0 & \begin{cases} (\lambda e^{\lambda x_e})/P_{bN} & -(1-f)Q < x_e < fQ \\ 0 & \text{otherwise} \end{cases} \\ \text{if } \hat{x}_b = 0 & \begin{cases} (\lambda e^{-\lambda |x_e|})/P_{b0} & -(1-f)Q < x_e < (1-f)Q \\ 0 & \text{otherwise} \end{cases} \end{cases} \quad (19-3)$$

که P_{b0} و P_{bN} در رابطه (۱۸-۳) تعریف شده اند. دقت شود که توزیع x_e با علامت \hat{x}_b معین می شود (و نه مقدار آن).

و اما در Case II، ضرایب پایه و بهبود از یکدیگر مستقلند. تابع توزیع f^{II} در ضمیمه بدست آمده و به این صورت می باشد.

$$f_{x_e|\hat{x}_b}^{II}(x_e|\hat{x}_b) = f_{x_e}^{II}(x_e) = \begin{cases} \frac{\lambda}{2} e^{-\lambda |x_e|} + \frac{\lambda}{e^{\lambda Q} - 1} & |x_e| \leq fQ \\ \frac{\lambda}{2} \frac{e^{\lambda Q}}{e^{\lambda Q} - 1} e^{-\lambda |x_e|} & fQ < |x_e| \leq (1-f)Q \end{cases} \quad (20-3)$$

تابع توزیع متغیرهای x_b و x_e و نواحی کوانتیزیشن در شکل ۳-۲ نشان داده شده است.



شکل ۳-۲ توزیع سیگنالهای پایه و بهبود، $\lambda=1/10$ ، $c=2$ (الف) BC (ب) EC و $\hat{x}_b = 1$ (ج) EC و Case II

تعداد سطح‌های ضریب بهبود بعد از کوانتیزیشن بستگی مستقیم به c (پارامتر MLMDC) دارد. هر چقدر c بزرگتر باشد طبیعتاً تعداد سطح‌ها هم بیشتر می‌شود. برای مثال در شکل ۳-۲، سه حالت ممکن $(-1,1,0)$ وجود دارد که تابع احتمال آنها به صورت زیر بدست می‌آید:

$$p_{e|\hat{x}_b}(1) = p_{e|\hat{x}_b}(\hat{x}_e = 1) = \int_{(1-f)Q/c}^{(1-f)Q} f_{x_e|\hat{x}_b}(x_e|\hat{x}_b) dx_e$$

$$p_{e|\hat{x}_b}(-1) = p_{e|\hat{x}_b}(\hat{x}_e = -1) = \int_A^{-(1-f)Q/c} f_{x_e|\hat{x}_b}(x_e|\hat{x}_b) dx_e$$

(۲۱-۳)

where A is $-fQ$ for *Case I* and $-(1-f)Q$ for *Case II*

$$p_{e|\hat{x}_b}(2) = p_{e|\hat{x}_b}(\hat{x}_e = 2) = 0, \quad p_{e|\hat{x}_b}(-2) = p_{e|\hat{x}_b}(\hat{x}_e = -2) = 0$$

$$p_{e|\hat{x}_b}(0) = p_{e|\hat{x}_b}(\hat{x}_e = 0) = 1 - (p_{e|\hat{x}_b}(1) + p_{e|\hat{x}_b}(-1) + p_{e|\hat{x}_b}(2) + p_{e|\hat{x}_b}(-2))$$

که در انتگرالها به جای f ، f^I و f^{II} به ترتیب برای *Case I* و *Case II* قرار می‌گیرد. به راحتی می‌توان ثابت کرد که برای *Case I*:

$$\forall m, n \in \mathbb{Z} \geq 0$$

$$p_{e|m}(1) = p_{e|n}(1) = p_{e|-m}(-1) = p_{e|-n}(-1) = P_{e1p}$$

$$p_{e|m}(2) = p_{e|n}(2) = p_{e|-m}(-2) = p_{e|-n}(-2) = P_{e2p}$$

$$p_{e|m}(-1) = p_{e|n}(-1) = p_{e|-m}(1) = p_{e|-n}(1) = P_{e1n}$$

$$p_{e|m}(-2) = p_{e|n}(-2) = p_{e|-m}(2) = p_{e|-n}(2) = P_{e2n}$$

$$p_{e|m}(0) = p_{e|n}(0) = 1 - (P_{e1p} + P_{e2p} + P_{e1n} + P_{e2n}) = P_{e0}$$

(۲۲-۳)

برای *Case II*، روابط (۲۲-۳) درست هستند و علاوه بر آن

$$P_{e1p} = P_{e1n}$$

$$P_{e2p} = P_{e2n}$$
(۲۳-۳)

با جایگذاری (۳-۱۸) و (۳-۲۱) در (۳-۱۱) خواهیم داشت:

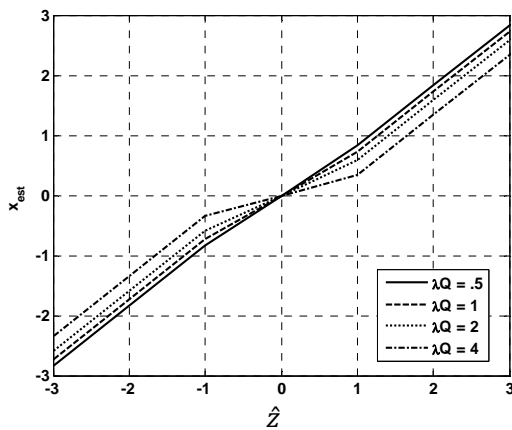
$$\hat{x}_{est} = \begin{cases} \hat{z} - \text{sign}(\hat{z})Z_0 & |\hat{z}| \geq 2 \\ N_0\hat{z} & |\hat{z}| < 2 \end{cases}$$
(۲۴-۳)

که در آن:

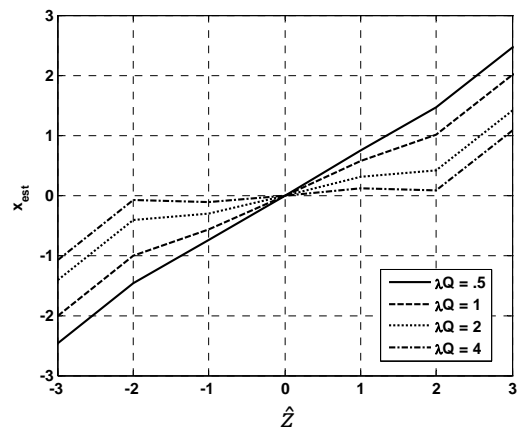
$$Z_0 = \frac{2P_{e2p}e^{2\lambda Q} + P_{e1p}e^{\lambda Q} - P_{e1n}e^{-\lambda Q} - 2P_{e2n}e^{-2\lambda Q}}{P_{e2p}e^{2\lambda Q} + P_{e1p}e^{\lambda Q} + P_{e0} + P_{e1n}e^{-\lambda Q} + P_{e2n}e^{-2\lambda Q}}$$

$$N_0 = \frac{3P_{e2n}e^{-2\lambda Q} + P_{e0} + 2P_{e1n}e^{-\lambda Q} - P_{e2n}}{P_{e2n}e^{-2\lambda Q} + P_{e0} + P_{e1n}e^{-\lambda Q} + P_{e2n} + P_{b0}P_{e1p}e^{\lambda Q}/P_{bN}}$$
(۲۵-۳)

تابع توزیع ضرایب پایه و بهبود توابعی λQ هستند و به همین علت Z_0 و N_0 توابعی از λQ (علاوه بر c) هستند و لذا تابع تخمین هم تابعی از λQ و c خواهد بود. در شکل ۳-۳ و شکل ۳-۴ این منحنی‌ها نشان داده شده‌اند. Case I و Case II منحنی‌های تقریباً مشابهی دارند؛ علت آنست که ترم‌های مربوط به BC که در هر دو مشترک هستند، غالب هستند.

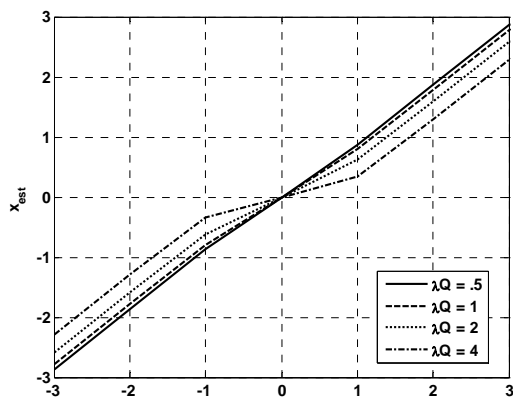


(الف)

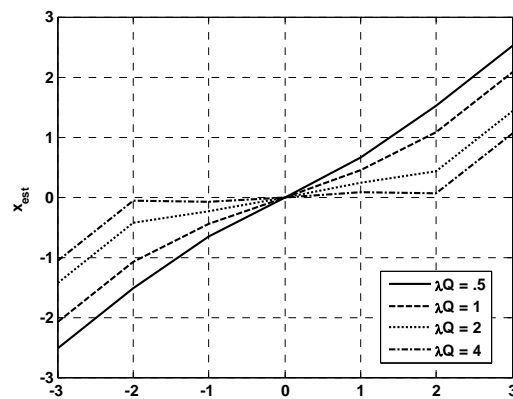


(ب)

شکل ۳-۳ منحنی تخمین برای Case I و (الف) $c = 1.5$ (ب) $c = 3.5$



(الف)



(ب)

شکل ۳-۴ منحنی تخمین برای CaseII و (الف) $c = 1.5$ (ب) $c = 3.5$

نکته دیگر اینکه برای مقادیر کوچک c ، توابع تخمین تقریباً شبیه تابع همانی $y = x$ است یعنی اینکه عملاً تخمین کاری انجام نمی‌دهد و همان ضریب ترکیبی (Z) را به عنوان \hat{x}_{est} می‌دهد. با مشاهده رابطه (۳-۲۴) مشخص می‌شود که برای Z_0 های کوچک این اتفاق می‌افتد. با روش‌های عددی، بدست می‌آید که وقتی c از حد زیر کوچکتر باشد، Z_0 هم از 0.2 کوچکتر است و با تقریب خوبی می‌توان تخمین را حذف کرد.

$$c \leq \left(\frac{6.425}{\lambda Q - 0.4274} \right)^{0.1131} \quad (۳-۲۶)$$

در شکل ۳-۵ و شکل ۳-۶، به ترتیب اعوجاج تخمین و اعوجاج جانبی نشان داده شده‌اند. CaseI و CaseII از نظر اعوجاج تخمین تقریباً با یکدیگر برابرند ولی CaseI اعوجاج جانبی کمتری را ارائه می‌دهد. این اختلاف در اعوجاج جانبی همانطور که گفته شد از همبستگی بین ضرایب ناشی می‌شود. فرض کنید منحنی تخمین بصورت تابع همانی در نظر گرفته شود، آنگاه داریم:

$$\begin{aligned} D_{est} &= E[(\tilde{x}_b - Q\hat{x}_{est})^2] = E[(\tilde{x}_b - Q\hat{z})^2] = E\left[(\tilde{x}_b - Q(\hat{x}_b + \hat{x}_e))^2\right] \\ &= Q^2 E[\hat{x}_e^2] \end{aligned} \quad (۳-۲۷)$$

جمله $E[\hat{x}_e^2]$ برای هر دوی CaseI و CaseII مشترک است و لذا اعوجاج تخمین یکسانی خواهیم داشت. و

اما برای

$$\begin{aligned} D_{side} &= E[(x_b - Q\hat{x}_{est})^2] = E[(x_b - Q\hat{z})^2] = E[(x_b - Q(\hat{x}_b + \hat{x}_e))^2] \\ &= E[x_b^2] + Q^2E[\hat{x}_b^2] + Q^2E[\hat{x}_e^2] - 2QE[x_b\hat{x}_b] + 2Q^2E[\hat{x}_b\hat{x}_e] \\ &\quad - 2QE[x_b\hat{x}_e] = A_c + A_d \end{aligned} \quad (28-3)$$

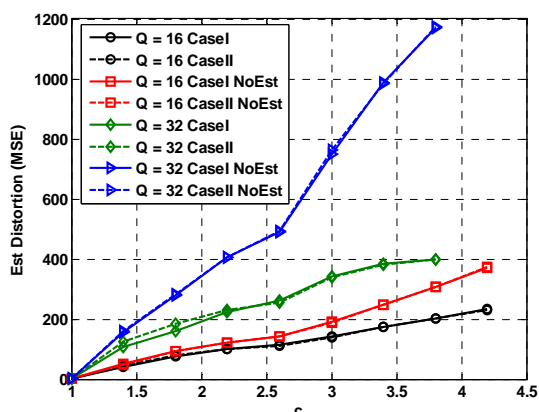
چهار جمله اول (A_c) برای هر دو (CaseI , CaseII) مشترک هستند و دو جمله آخر (A_d) در CaseII صفر است (چون ضرایب مستقلند) ولی در CaseI چنین نیست. برای بدست آوردن A_d در CaseI ، x_b را با بازسازی شده‌ی آن توسط دیکودر مرکزی تقریب می‌زنیم؛ در این صورت داریم:

$$E[x_b\hat{x}_e] \approx E[(Q\hat{x}_b + Q\hat{x}_e/c)\hat{x}_e] = QE[\hat{x}_b\hat{x}_e] + QE[\hat{x}_e^2]/c \quad (29-3)$$

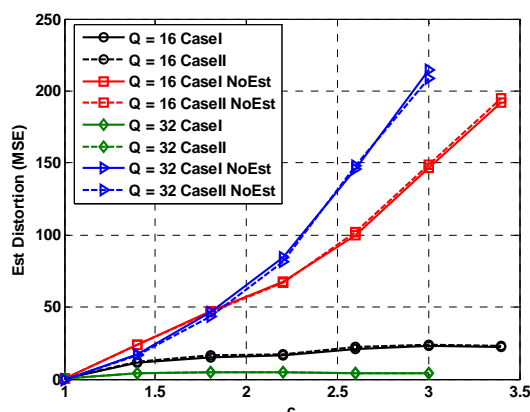
و لذا

$$A_d \approx 2Q^2E[\hat{x}_b\hat{x}_e] - 2Q(QE[\hat{x}_b\hat{x}_e] + QE[\hat{x}_e^2]/c) = -2Q^2E[\hat{x}_e^2]/c \quad (30-3)$$

در حالی که در A_d در CaseII صفر است مقدار آن در CaseI منفی است که همین، علت کمتر بودن اعوجاج جانبی است.

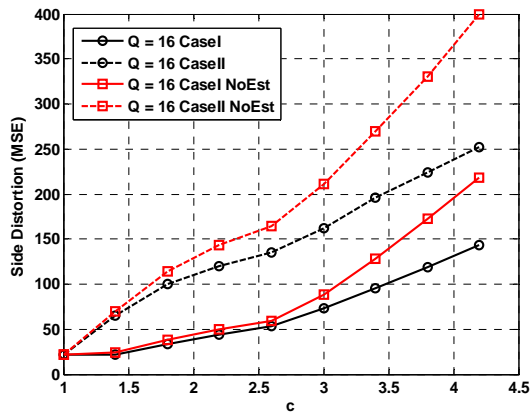


(الف)

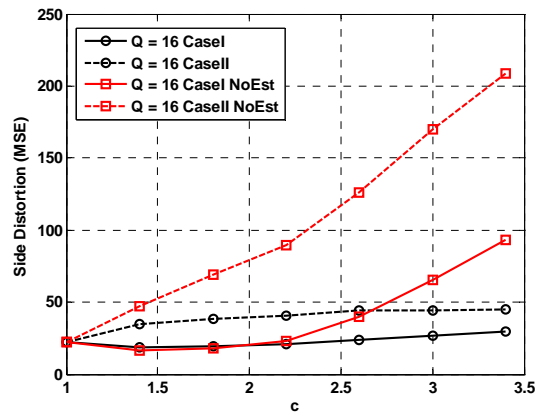


(ب)

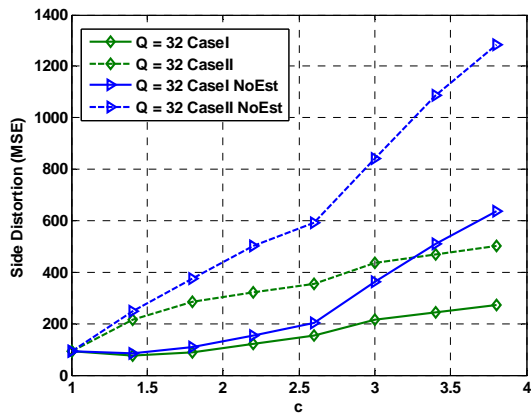
شکل ۳-۵ اعوجاج تخمین برای CaseI و CaseII ، Q=16 و (الف) $\lambda = 0.05$ و (ب) $\lambda = 0.2$



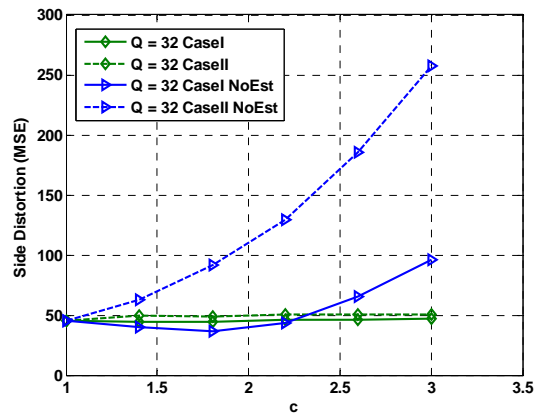
(الف)



(ب)



(ج)



(د)

شکل ۳-۶ اعوجاج جانبی برای CaseI و CaseII (الف) $\lambda = 1/20$ و $Q = 16$ (ب) $\lambda = 1/5$ و $Q = 16$ (ج) $Q = 16$ و $\lambda = 1/20$ و $Q = 32$ (د) $Q = 32$ و $\lambda = 1/5$

در این اشکال حالتی هم که تخمین استفاده نشود آورده شده است.

همانطور که شکل ۳-۵ نشان می‌دهد، اعوجاج تخمین برای $\lambda = 0.2$ هنگامی که $Q = 32$ است کمتر از حالتی است که $Q = 16$ است. در واقع برای هر λ ، یک مقدار مشخصی Q وجود دارد که بعد از آن خطای تخمین کاهش می‌شود؛ می‌توان نشان داد که ضرایب DCT که در محدوده $[-\frac{(1-f)Q}{c}, \frac{(1-f)Q}{c}]$ قرار گرفته‌اند بدون خطا تخمین زده می‌شوند و لذا افزایش Q کاهش اعوجاج تخمین را به دنبال خواهد داشت.

نکته جالب دیگر اینکه وقتی $\lambda = 0.2$ است، برای محدوده‌های پایین c ، اعوجاج جانبی کاهش می‌یابد؛ به عبارت دیگر بر خلاف آنچه انتظار داشتیم ترکیب کردن اعوجاج اضافی منفی ایجاد کرده است! همانطور که در شکل ۳-۲ نشان داده شد، به طور متوسط BC و EC علامت یکسانی دارند یعنی اینکه مقدار متوسط CC نسبت به BC از مبدأ دورتر است. این مانند وقتی است که نواحی کوانتیزیشن را به سمت مبدأ نزدیک‌تر کنیم و این همان افزایش f به سمت $1/2$ است. گفته شد که برای $f = 1/2$ نسبت به وقتی که f از رابطه (۳-۳) بدست آید اعوجاج کمتری خواهیم داشت. بنابراین اثر ترکیب مانند اثر افزایش f باعث کاهش اعوجاج می‌شود، البته این عمل برای c های کوچک و λ های بزرگ رخ می‌دهد. برای λ های کوچک f بهینه با $1/2$ چندان تفاوتی ندارد.

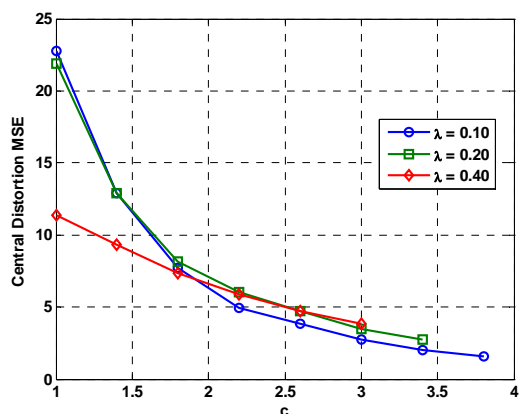
۳.۳.۳. اعوجاج مرکزی

اگر هر دو توصیف در دسترس باشند، برای همه ضرایب DCT، هم BC موجود است و هم CC که نتیجتاً از تفاضل آنها EC بدست می‌آید. اعوجاج مرکزی برابر است با نویز کوانتیزیشن EC وقتی که $Q = Q/c$ ؛ این نویز برای مقادیر صحیح c برابر با همان مقداری است اگر BC را با $Q = Q/c$ و همان f کوانتیزه می‌کردیم. یعنی اینکه کیفیت مرکزی برای مقادیر صحیح c معادل است با نویز کوانتیزیشن x_b وقتی $Q = Q/c$ ، یعنی:

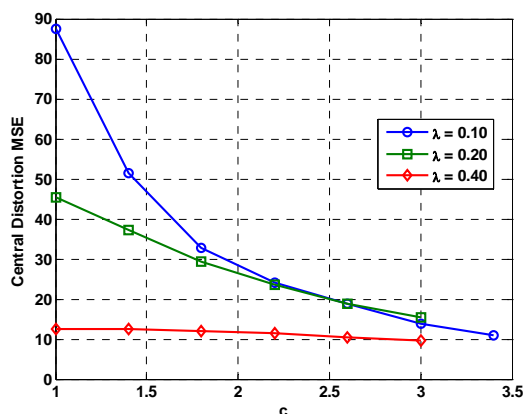
$$D_{cen} = D_q|_{Q=\frac{Q}{c}} \quad c \in \mathbb{Z} \quad (3-31)$$

برای مقادیر غیر صحیح c ، روش دقیق همان محاسبه انتگرال اعوجاج کوانتیزیشن برای سیگنال بهبود است؛ تابع توزیع در حالت CaseII (داده شده در (۳-۲۰)) برای این کار مناسب‌تر است (روشن است که CaseI و CaseII اعوجاج مرکزی یکسانی می‌دهند). اما شبیه سازیهای انجام شده نشان می‌دهد که با f داده شده در (۳-۱۳) و همچنین برای سیگنالهای با $\lambda Q > 1.4$ ، رابطه (۳-۳۱) اعوجاج مرکزی را با خطای کمتر از

10% می‌دهد. با این وجود، اعوجاج مرکزی از رابطه دقیق محاسبه شده ولی رابطه آن به علت طولانی بودن آورده نشده است. اعوجاج مرکزی بر حسب c برای چند مقدار λ و Q در شکل ۳-۷ نشان داده شده است.



(الف)



(ب)

شکل ۳-۷. کیفیت مرکزی بر حسب c برای (الف) $Q = 16$ و (ب) $Q = 32$

برای $Q = 16$ و $\lambda = 0.1, 0.2$ ، از رابطه D_q (۳-۱۵) بدست می‌آید و لذا مستقل از λ است. می‌توان دید که بهبود در اعوجاج مرکزی با افزایش λ یا Q ، کاهش می‌یابد. علت این است که در این صورت ضرایب عمدتاً صفر یا نزدیک به صفر هستند و لذا بازسازی دقیق‌تر آنها بهره‌ای در اعوجاج نخواهد داشت.

۴.۳.۳. افزونگی نرخ ناشی از ضریب بهبود

همانطور که گفته شد در روش MLMDC علاوه بر خطای تخمین، افزونگی بیت نیز به دلیل ترکیب شدن خواهیم داشت. این افزونگی بیت در Case II به دلیل بیشتر بودن واریانس سیگنال ترکیب (\hat{z}) روشن است اما در Case I، همبستگی میان ضرایب هم در واریانس موثر است. اما همانگونه که در ادامه نشان داده خواهد شد، این همبستگی چندان تأثیری در نرخ بیت ندارد.

ترکیب کردن EC و BC منجر به افزایش تعداد بیت لازم برای ارسال ضریب ترکیبی می‌شود. این تعداد بیت اضافی را نرخ بهبود نامیده‌ایم که با فرض اینکه تعداد بیت با آنتروپی تخمین زده شود، از رابطه‌ی زیر بدست می‌آید.

$$ER = \sum_{\hat{z}} P_{\hat{z}}(\hat{z}) \log_2\left(\frac{1}{P_{\hat{z}}(\hat{z})}\right) - \sum_{\hat{x}_b} P_{\hat{x}_b}(\hat{x}_b) \log_2\left(\frac{1}{P_{\hat{x}_b}(\hat{x}_b)}\right) \quad (3-32)$$

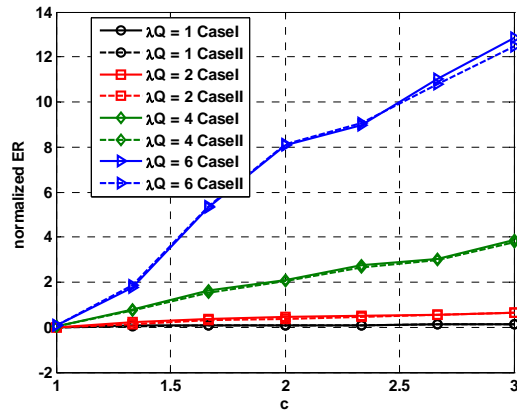
نرخ آنتروپی، گرچه ایده‌آل است ولی از آنجا که انکودرهای مورد استفاده در H.264/AVC، CABAC¹ و CAVLC²، تقریباً ایده‌آل هستند، تخمین قابل قبولی می‌دهد. برای مطالعه تأثیر ترکیب روی تعداد بیت، مقدار مطلق ER مناسب نیست؛ به همین منظور ER_n ، مقدار نرمال شده، به این صورت تعریف می‌شود.

$$ER_n = \frac{ER}{BR} \quad (3-33)$$

که در آن BR تعداد بیت لازم برای ارسال ضریب پایه است. ER_n تابعی از λQ است؛ شکل ۳-۸ تغییرات آن را نسبت به c برای چند مقدار λQ نشان می‌دهد. نکته اول اینکه هر دو Case I و Case II، منحنی تقریباً مشابهی دارند و دوم اینکه ضرایبی که λQ کوچکتری دارند، ER_n کوچکتری هم دارند. برای λQ های بزرگ، BR کاهش و ER افزایش می‌یابد که هر دو عامل افزایش ER_n هستند.

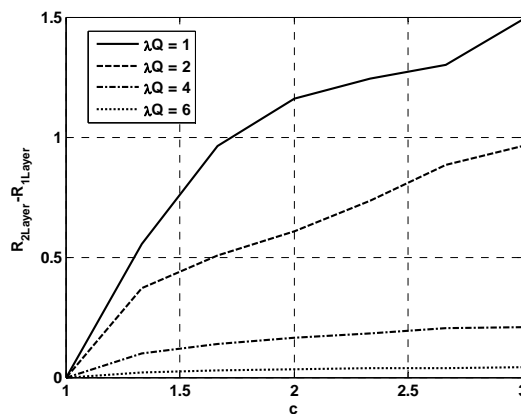
¹ Context based Adaptive Binary Arithmetic Coding

² Context based Adaptive Variable Length Coding



شکل ۸-۳ ER_n بر حسب c برای چند مقدار λQ

دیده شد که ترکیب باعث افزایش نرخ بیت می شود اما نرخ بیت لازم برای ارسال ضرایب ترکیبی از اینکه ضرایب پایه و بهبود را جداگانه در دو لایه ارسال کنیم کمتر است. در شبیه سازی شکل ۳-۹ این مطلب نشان داده شده است؛ در این شکل محور عمودی تفاوت تعداد بیت لازم برای کد کردن جداگانه ضرایب پایه و بهبود نسبت به کد کردن ضرایب ترکیبی را نشان می دهد. مشاهده می شود که این کمیت بزرگتر از صفر بوده و همچنین با افزایش λQ کوچکتر می شود. البته ناگفته نماند که ارسال جداگانه ضرایب علاوه بر اینکه بیت بیشتری نیاز دارد، اضافه نرخ ناشی از اطلاعات سرآمد هم دارد.



شکل ۹-۳ $R_{2Layers} - R_{1Layer}$ بر حسب c برای چند مقدار λQ

۴.۳. مقایسه عملکرد MLMDC

تاکنون عملکرد MLMDC را در اعوجاج و نرخ بیت بررسی کردیم؛ مشاهده شد که کیفیت مرکزی در ازای کاهش کیفیت جانبی و افزایش نرخ بهبود، زیاد می‌شود. در واقع این ویژگی مشترک الگوریتم‌های MDC است که باید بین کیفیت مرکزی، کیفیت جانبی و نرخ مصالحه برقرار کرد.

برای بررسی عملکرد MLMDC، منحنی نرخ-اعوجاج-افزونگی (RRD^1) آن را مورد مطالعه قرار می‌دهیم. RRD در [83] معرفی شد و طبق تعریف برابر است با کیفیت جانبی بر حسب نرخ افزونگی در یک کیفیت مرکزی ثابت. نرخ افزونگی در اینجا به معنی تعداد بیت‌های اضافی است نسبت به SDC، هنگامیکه کیفیت مرکزی برابر کیفیت SDC باشند. در یک کیفیت مرکزی ثابت، الگوریتمی که منحنی اعوجاج جانبی - نرخ بهتری داشته باشد مناسب‌تر است.

سه الگوریتم دیگر MDC: MDTC و Splitting (SP) و Duplication (Dup) نیز مورد بررسی قرار گرفته و منحنی‌های آنها با MLMDC مقایسه می‌شود. در SP، ضرایب بین توصیف‌ها تقسیم می‌شود و جای خالی هر ضریب در هر توصیف چنانچه معمول است صفر قرار داده می‌شود. از نظر کیفیت مرکزی، SP، بهترین عملکرد را دارد چون حداقل افزونگی را دارد. اما کیفیت جانبی آن، از همه بدتر است. در Dup همه‌ی ضرایب در هر دو توصیف کپی می‌شوند؛ روشن است که Dup کیفیت جانبی خوبی دارد ولی چونکه در دیکودر مرکزی یکی از توصیف‌ها کاملاً بدون استفاده می‌ماند کیفیت مرکزی مناسبی ندارد. SP و Dup الگوریتم‌های حدی MDC هستند. در MDTC که در بخش ۳.۱.۲ توضیح داده شد به جای اینکه دو ضریب را بین دو توصیف تقسیم کنیم، آنها را دو جور مختلف ترکیب کرده و ترکیب‌ها را بین دو توصیف قرار می‌دهیم. به این ترتیب در دیکودر مرکزی با بکار بردن عکس تابع ترکیب، ضرایب قابل بازیابی بوده و

¹ Redundancy Rate Distortion

کیفیت مرکزی در حد SP گر چه با نرخ بیت بیشتر حاصل می‌شود. ویژگی این روش اینست که در دیکودر جانبی با استفاده از ضریب ترکیب موجود می‌توان هر دو ضریب را تخمین زد نه این که مانند SP فقط یک ضریب را در اختیار داشته باشیم. با طراحی مناسب MDTC کیفیت جانبی مطلوب‌تری نسبت به SP بدست می‌آوریم.

حال دو ضریب x_1 و x_2 را در نظر بگیرید که با هر یک از روش‌های فوق به دو توصیف کد می‌شوند. SP و Dup که سراسر بوده و به راحتی اعمال می‌شوند. در MLMDC، توصیف‌ها از روابط (۳-۲) و (۳-۳) با $c = 2$ تولید می‌شوند. در طراحی MDTC، توزیع سیگنال‌ها گوسین در نظر گرفته شده ولی در اینجا لاپلاسین استفاده می‌شود؛ کوانتیزیشن به فرم رابطه (۱-۱) و با $f = 1/2$ اعمال می‌شود، ماتریس ترکیب کننده‌ی بهینه نامتعامل^۱ و روابط تخمین هم بصورت داده شده در زیر مورد استفاده قرار می‌گیرد [51].

$$\begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \end{bmatrix} = T \begin{bmatrix} \hat{x}_{1b} \\ \hat{x}_{2b} \end{bmatrix}$$

$$T = \begin{bmatrix} \sqrt{\frac{\cot(\theta)}{2}} & \sqrt{\frac{\tan(\theta)}{2}} \\ \sqrt{\frac{\tan(\theta)}{2}} & -\sqrt{\frac{\cot(\theta)}{2}} \end{bmatrix} \quad (3-34)$$

$$\text{if } \hat{z}_1 \text{ is available} \Rightarrow \begin{cases} \hat{x}_{1est} = \beta_1 \hat{z}_1 & \beta_1 = \frac{E[\hat{x}_{1b}\hat{z}_1]}{E[\hat{z}_1^2]} \\ \hat{x}_{2est} = \beta_2 \hat{z}_1 & \beta_2 = \frac{E[\hat{x}_{2b}\hat{z}_1]}{E[\hat{z}_1^2]} \end{cases}$$

¹ Nonorthogonal

برای منحنی‌های MDTC، $\theta = \pi/10$ قرار داده شده است، مقداری که افزونگی بالایی می‌دهد. منحنی‌های RD (Rate Distortion) و همچنین RRD، برای حالتی که λ ها برابر باشند و نباشند در شکل ۳-۱۰ رسم شده است.

همانطور که قبلاً هم اشاره شد، شکل‌ها نشان می‌دهند که اگر چه منحنی نرخ - اعوجاج مرکزی Case I و Case II مشابه هستند (منطبق بر یکدیگر) ولی اعوجاج جانبی Case I بسیار بهتر است و همچنین منحنی کیفیت جانبی آن به حالت ایده‌آل Dup (در بیشتر موارد) نزدیک‌تر از سایرین است حال آنکه اعوجاج مرکزی آن به طور قابل ملاحظه‌ای بهتر از Dup است. بهترین اعوجاج مرکزی متعلق به SP است و بعد از آن MDTC قرار گرفته است.

در [51] افزونگی به منظور کاهش اعوجاج جانبی مصرف می‌شود ولی در MLMDC به منظور کاهش اعوجاج مرکزی؛ اگر چه معادل هستند ولی برای رسم RRD، بایستی که آنها را در یک کیفیت مرکزی ثابت نگه داریم. برای نمونه در $Q = Q_0$ ، MDTC اعوجاجی برابر با D_0 می‌دهد ولی در MLMDC، اعوجاج $D_1 < D_0$ است به همین منظور $Q = Q_1 > Q_0$ را چنان پیدا می‌کنیم که اعوجاج مرکزی MLMDC برابر D_0 شود؛ به عبارت دیگر MDTC در Q_0 و MLMDC در Q_1 کار می‌کنند تا اعوجاج مرکزی یکسانی داشته باشند. SP و Dup پارامتری که با آن بتوان افزونگی را کنترل کرد ندارند و منحنی RRD آنها یک نقطه است.

در MLMDC، افزونگی‌های پایین منطبق است با مقادیر بزرگ c ، چونکه وقتی c بزرگ می‌شود Q_1 هم بزرگتر می‌شود که این خود باعث کاهش نرخ بیت می‌شود. بنابراین انتهای سمت راست منحنی‌های RRD در MLMDC مربوط است به مقادیر کوچک c . این قسمت به نقطه RRD الگوریتم Dup که معادل است با MLMDC وقتی $c = 1$ ، هم نزدیکتر است. مشاهده می‌شود اگر چه افزونگی MLMDC بیشتر است اما RRD

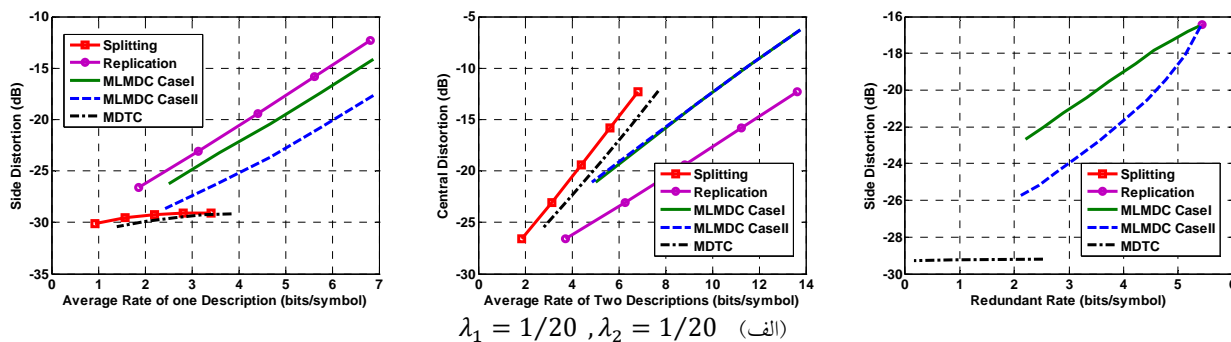
آن از سایرین بهتر است به ویژه برای λ های کوچک. از این شبیه‌سازی‌ها، نکات و تحلیل‌های زیر استخراج می‌شود.

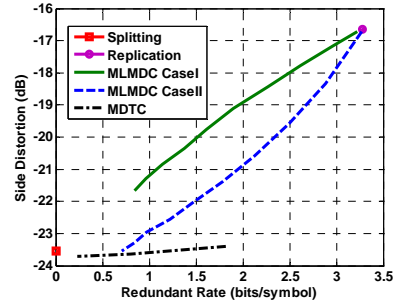
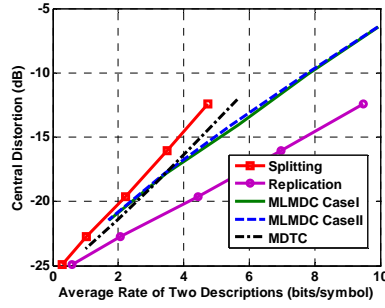
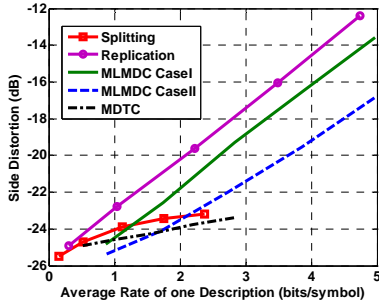
الف) عملکرد Case I از Case II بهتر است.

ب) MDTC برای منابع با λ های متفاوت پیشنهاد شده است؛ سیگنال با λ کوچک با سیگنال با λ بزرگ ترکیب می‌شود و عملاً وقتی λ ها یکسان باشند MDTC بی‌فایده است، چنانچه در شکل هم دیده می‌شود. اما در تصاویر/ویدئوهای طبیعی اینکه بعضی از ضرایب DCT توزیع حدوداً یکسان داشته باشند معمول است. MLMDC محدودیت گفته شده را ندارد و از MDTC بهتر عمل می‌کند به جز وقتی که λ ها خیلی بزرگ باشند.

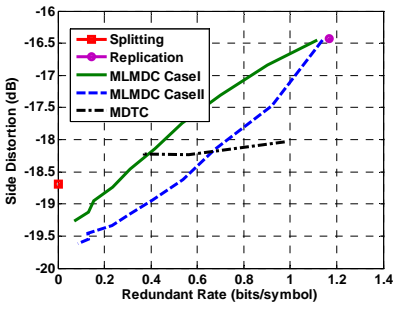
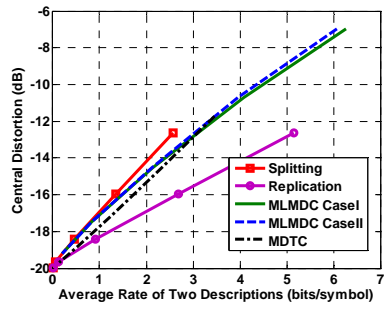
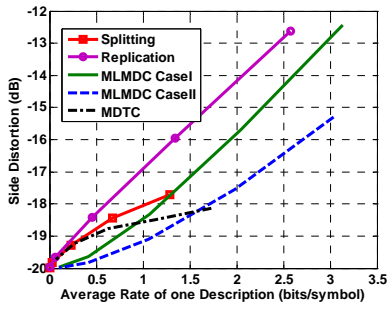
ج) کارایی MLMDC برای سیگنال‌های با λ بزرگ در یک Q ثابت، کاهش پیدا می‌کند. اول به خاطر اینکه ER_n زیادتر است و دوم به خاطر اینکه کیفیت مرکزی آنها با سیگنال بهبود چندان تغییر نمی‌کند. وضعیت بهتر می‌شود اگر که در Q پایین‌تر کدینگ را انجام دهیم.

د) شبیه‌سازی‌ها برای کدینگ دو توصیفه بود، در کدینگ مثلاً چهار توصیفه افزونگی و اعوجاج تخمین در هر توصیف کمتر شده و لذا کارایی MLMDC بهتر هم می‌شود.

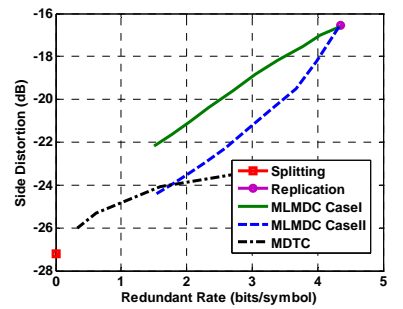
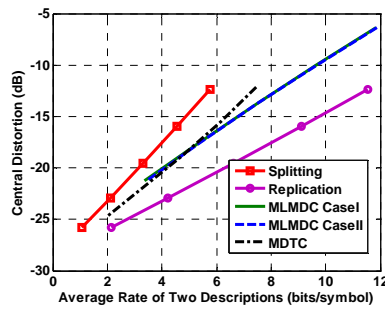
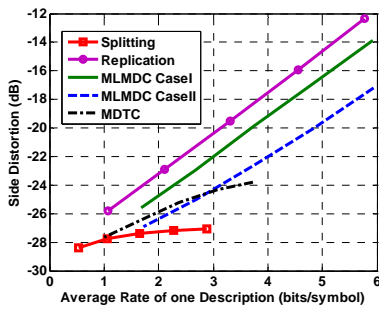




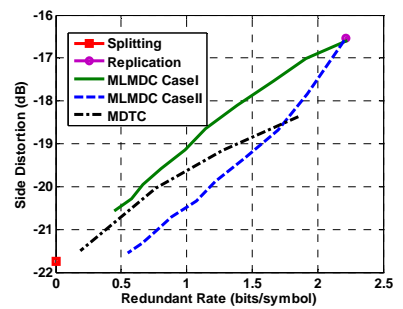
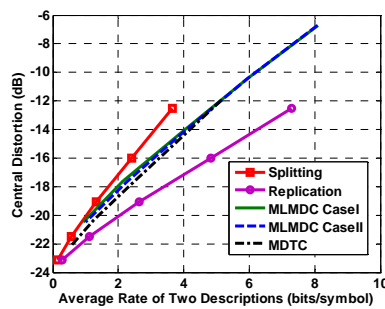
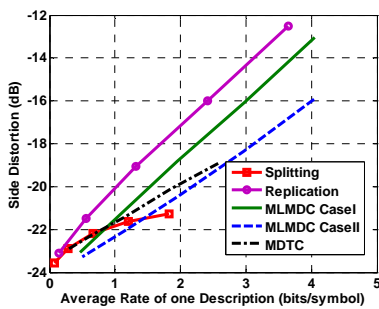
$$\lambda_1 = 1/10, \lambda_2 = 1/10 \quad (\text{ب})$$



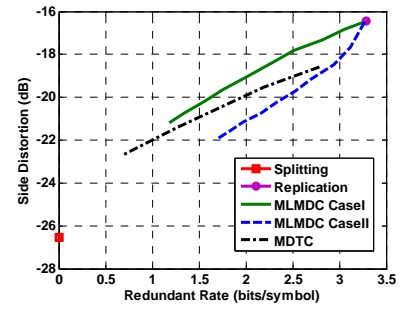
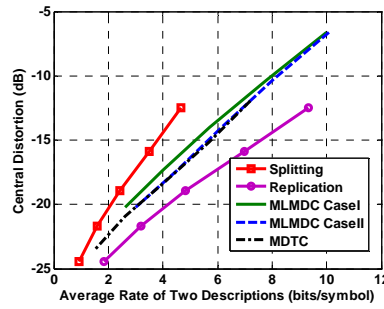
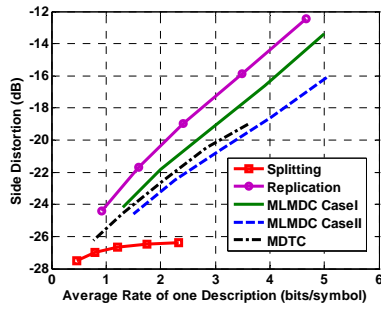
$$\lambda_1 = 1/5, \lambda_2 = 1/5 \quad (\text{ج})$$



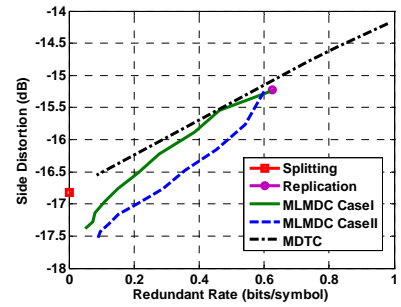
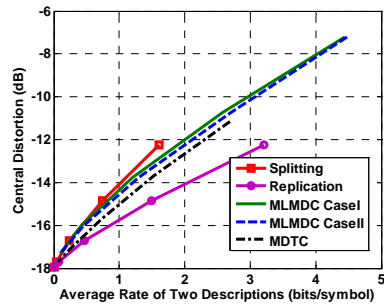
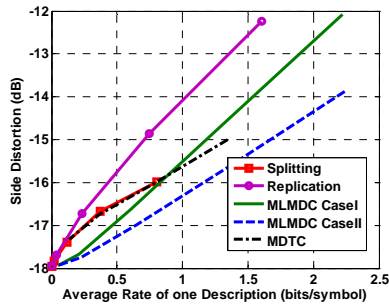
$$\lambda_1 = 1/20, \lambda_2 = 1/10 \quad (\text{د})$$



$$\lambda_1 = 1/10, \lambda_2 = 1/5 \quad (\text{ه})$$



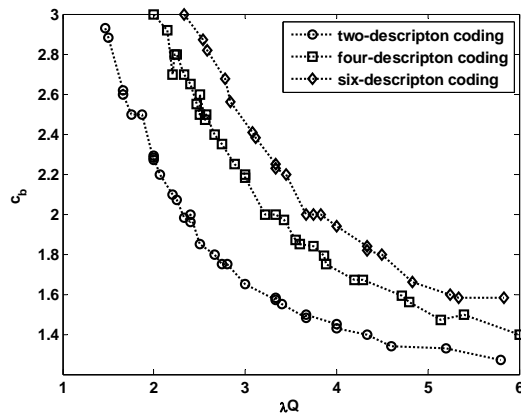
$$\lambda_1 = 1/20, \lambda_2 = 1/5 \quad (و)$$



$$\lambda_1 = 1/5, \lambda_2 = 1/2.5 \quad (ز)$$

شکل ۳-۱۰ منحنی‌های RD برای Splitting، Duplication، MDTC و MLMDC برای چند مقدار (λ_1, λ_2) چپ: rate/side-distortion و راست: rate/central-distortion و $Q_0 = 16$ redundancy-rate-distortion

می‌توان مشاهده کرد در افزونگی‌های پایین یا مقادیر بالای c ، منحنی RRD الگوریتم MLMDC به منحنی‌های سایرین نزدیک می‌شود. به عنوان نمونه در شکل ۳-۱۰ (ج) و (ز) عملکرد آن حتی از SP هم بدتر می‌شود؛ یعنی اینکه یک مقدار خاص c ، c_b وجود دارد که بعد از آن عملکرد MLMDC از SP بدتر می‌شود. این حد با استفاده از شبیه‌سازی متعدد بدست آمده، در شکل ۳-۱۱ نشان داده شده و از رابطه زیر به طور تقریبی بدست می‌آید.



شکل ۳-۱۱: مقادیر c_b برحسب λQ MLMDC

$$c_b \cong \frac{6.9 \left(1 + \frac{1}{3} \ln \left(\frac{n}{2} \right) \right)}{\lambda Q + 0.9} \quad (3-35)$$

که در آن n تعداد توصیف‌ها است. در واقع این رابطه نشان می‌دهد که چگونه تعداد توصیف‌ها و مقدار λQ در تعیین مقدار ماکزیمم پارامتر MLMDC نقش دارند.

۵.۳. نتایج عملی

MLMDC در نرم‌افزار مرجع H264/AVC ، JM16.0 ، پیاده‌سازی شده است؛ اگر چه هر کدک دیگری که بر پایه DCT باشد هم قابل استفاده خواهد بود.

CaseII، به خاطر اینکه عملکرد مناسبی نداشت از این بخش حذف می‌شود. ویدئوهای Akiyo ، Foreman و Mobile به عنوان نماینده سیگنال‌های با محتوای پایین، متوسط و بالا برای تست عملکرد MLMDC انتخاب شدند. فریم‌ها QCIF هستند و همه ۳۰۰ فریم کد شده‌اند. فقط 4x4DCT استفاده شده، برای فریم‌های اینترا، فقط حالت 4x4 مجاز شده است. GOP=10 و CABAC هم به عنوان انکودر آنتروپی فعال شده است. نرخ فریم ۳۰ فریم/ثانیه ، محدوده بردار حرکت ۳۲ پیکسل تنظیم و تخمین حرکت زیر یک پیکسل هم غیرفعال شده است. برای اسلایسهای نوع P ، مدهای 8x8 ، 8x16 ، 16x8 ، 16x16 فعال شده‌اند.

اسلایس نوع B هم نداریم. همه الگوریتم‌ها فقط روی مولفه‌های luma پیاده‌سازی شده و تعداد بیت‌های luma هم به عنوان نرخ بیت محاسبه شده است.

توزیع ضریب‌های DCT، فریم به فریم محاسبه می‌شود؛ در هر فریم QCIF، 1484 نمونه از هر فرکانس DCT وجود دارد که برای محاسبه توزیع ضرایب کافی است. با داشتن توزیع ضرایب، متغیرهای تابع تخمین برای هر فریم و در هر مقدار پارامتر c بدست می‌آید.

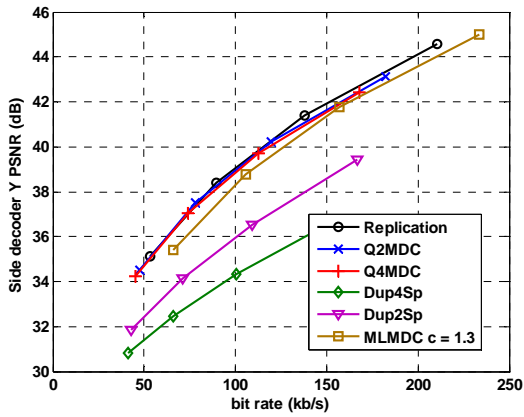
MLMDC با دو روش MDC دیگر حوزه فرکانس مقایسه شده است که در بخش قبل معرفی شدند. یکی MDC ارائه شده در [41] است که آنرا DupSp نامیده‌ایم و بر اساس تکرار کردن ضرایب بزرگ DCT در هر دو توصیف و تقسیم ضرایب کوچک بین آنها کار می‌کند [رجوع شود به بخش ۳.۱.۲]. روش‌هایی هم برای بهینه کردن اینکه کدام ضریب کپی شود و کدام تقسیم شود وجود دارد [41] و [42]، ولی به منظور وارد نشدن به مسائل بهینه‌سازی، DupSp را با دو سطح 2 و 4 پیاده‌سازی می‌کنیم که در آنها افزون بر ضریب DC (فرکانس صفر) ضرایب بزرگتر از به ترتیب 2 و 4 در هر دو توصیف تکرار می‌شوند و بقیه بین توصیف‌ها تقسیم می‌گردند. آنها را به ترتیب با Dup2Sp و Dup4Sp در شبیه‌سازیها نشان می‌دهیم. MDC دیگری که شبیه‌سازی شده، روش ارائه شده در [58] است که در همان بخش توضیح داده شده است. در این روش که QMDC نامیده شده است، هر توصیف ترکیبی از اسلایسهای اصلی و اسلایسهای افزونگی است؛ اسلایسهای افزونگی در نرخ بیت پایینتر کد می‌شوند. به عبارت دیگر QP در اسلایسهای افزونگی بزرگتر از مقدار آن در اسلایسهای اصلی است که البته اینکه چقدر بزرگتر باشد را می‌توان با توجه به مشخصات ویدئو و کانال بهینه کرد. در اینجا، به جای حل مسئله بهینه‌سازی، دو حالت Q2MDC و Q4MDC را در نظر گرفته‌ایم. در Q2MDC و کدینگ دو توصیفه، QP اسلایسهای افزونگی 2 واحد بیشتر از QP اسلایسهای اصلی است. در کدینگ چهارتوصیفه، به ازای هر اسلایس اصلی، سه اسلایس افزونگی داریم

که در Q2MDC با QP های 2 و 4 و 6 واحد بیشتر از QP اصلی کد می‌شوند. Q4MDC ، مشابه Q2MDC است اما واحد افزایش QP بجای 2، 4 است.

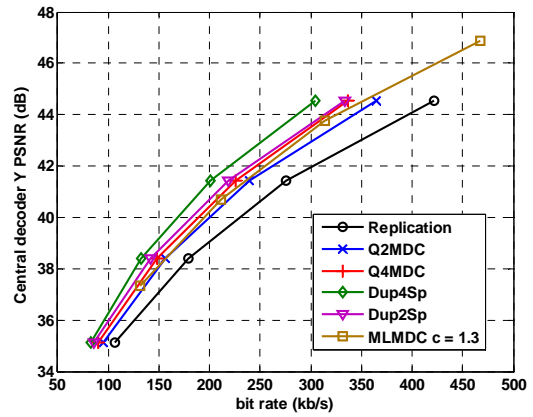
MDC های گفته شده و MLMDC و همچنین Replication که در آن استریم SDC روی هر دو کانال تکرار می‌شود، شبیه سازی شده و به ازای $QP = \{20, 24, 28, 32\}$ ، نرخ و اعوجاج حاصل اندازه گیری می‌شوند.

در شکل ۳-۱۲ تا شکل ۳-۱۴، برای سکانس‌های Akiyo و Foreman و Mobile، منحنی‌های نرخ - کیفیت جانبی و نرخ- کیفیت مرکزی برای هنگامیکه یک توصیف سالم می‌رسد و هنگامیکه تمام توصیفها در دسترس هستند رسم شده است. در واقع در این شبیه‌سازیها، کارکرد تک-توصیفه MLMDC را مورد سنجش قرار می‌دهیم. یکی از مسائل مطرح در MDC، دریفت^۱ است که به علت ناهمخوانی تصویر مرجع در انکودر و دیکودر رخ می‌دهد. به این منظور در انکودر، برای هر توصیف، حلقه پیش‌بینی جداگانه در نظر گرفته شده است؛ به این ترتیب تا زمانیکه سوئیچینگ بین توصیفها رخ ندهد، در دیکودر جانبی دریفت نخواهیم داشت چونکه مرجع پیش‌بینی استفاده شده در دیکودر جانبی قابل بازسازی است. نتایج نشان می‌دهد که MLMDC، به ویژه برای سکانس Mobile و Foreman که ضرایب DCT آن λ کوچکی دارند، کیفیت جانبی به مراتب بهتری از دیگران ارائه می‌دهد حال آنکه کیفیت مرکزی آن هم به نسبت خوب است. همانطور که قبلاً هم گفته شد با تعداد توصیف‌های بیشتر، عملکرد MLMDC بهتر هم می‌شود.

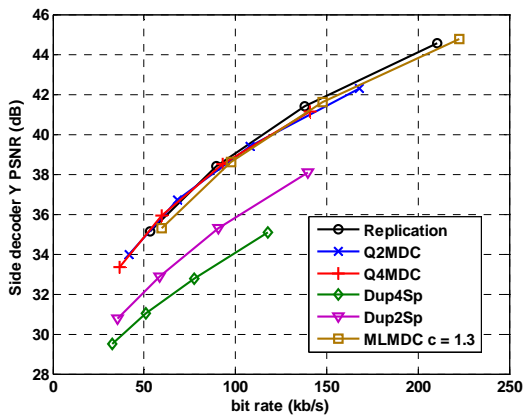
¹ Drift



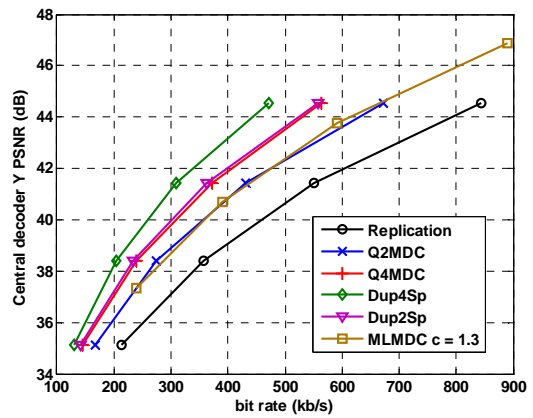
(الف)



(ب)

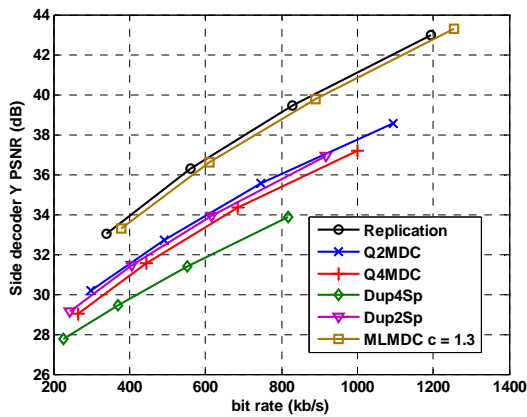


(ج)

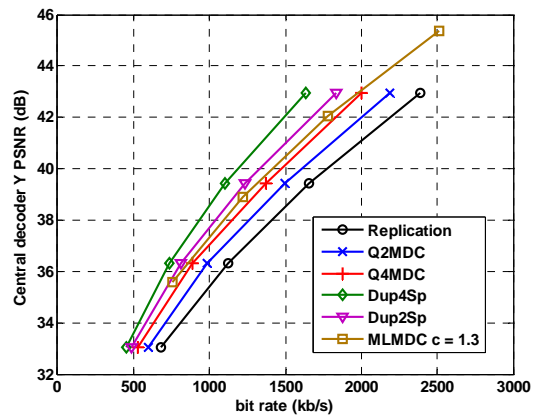


(د)

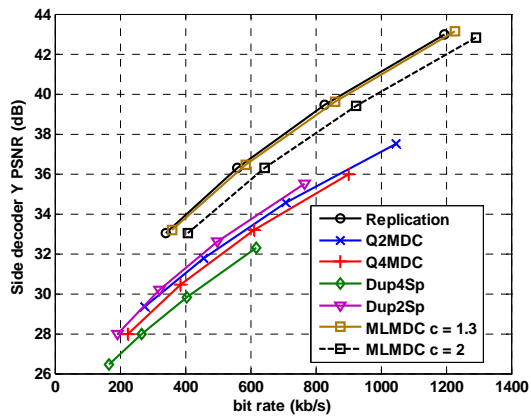
شکل ۳-۱۲ منحنی‌های RD برای الگوریتم‌های Dup، Dup4Sp، Dup4Sp، QMDC و MLMDC و سکانس Akiyo (الف) دیکودر جانبی و کدینگ دو توصیفه و (ب) دیکودر مرکزی و کدینگ دو توصیفه (ج) دیکودر جانبی و کدینگ چهار توصیفه و (د) دیکودر مرکزی و کدینگ چهار توصیفه



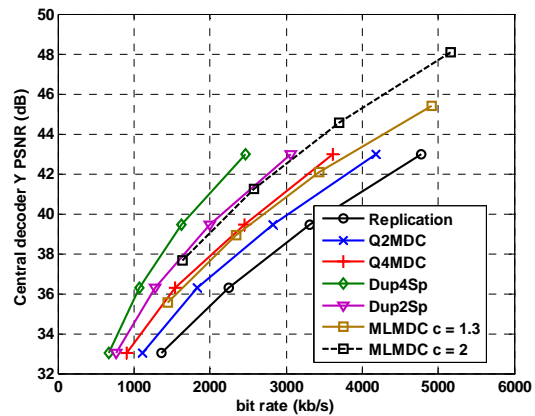
(الف)



(ب)

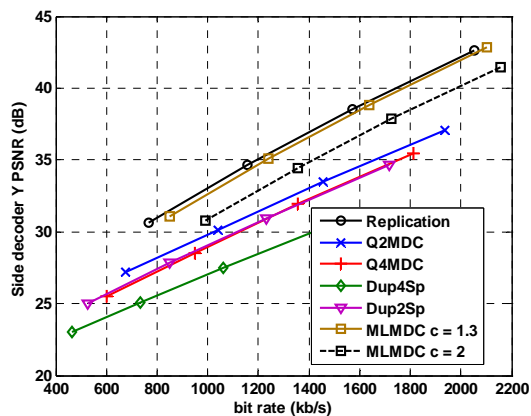


(ج)

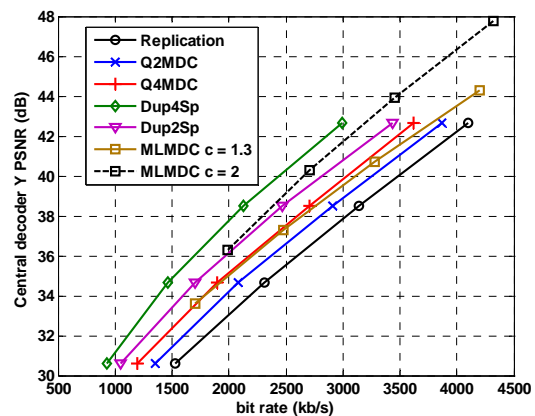


(د)

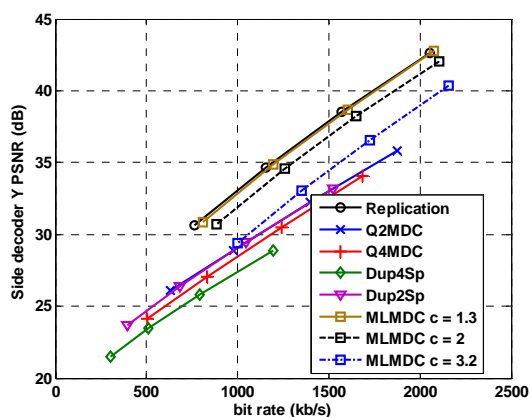
شکل ۳-۱۳ منحنی‌های RD برای الگوریتم‌های Dup، Dup4Sp، QMDC و MLMDC و سکانس Foreman (الف) دیکودر جانبی و کدینگ دو توصیفه و (ب) دیکودر مرکزی و کدینگ دو توصیفه و (ج) دیکودر جانبی و کدینگ چهار توصیفه و (د) دیکودر مرکزی و کدینگ چهار توصیفه



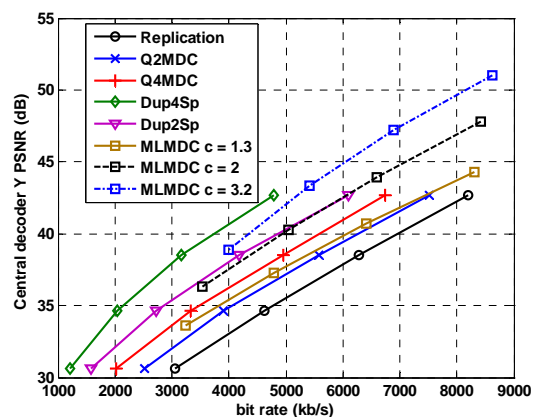
(الف)



(ب)



(ج)



(د)

شکل ۳-۱۴ منحنی‌های RD برای الگوریتم‌های Dup، Dup4Sp، QMDC و MLMDC و سکانس Mobile (الف) دیکودر جانبی و کدینگ دو توصیفه و (ب) دیکودر مرکزی و کدینگ دو توصیفه و (ج) دیکودر جانبی و کدینگ چهار توصیفه و (د) دیکودر مرکزی و کدینگ چهار توصیفه

از منحنی‌های فوق نکات زیر استنتاج می‌شود اگرچه قبلاً توضیح داده و پیش‌بینی شده بودند:

الف) منحنی کیفیت مرکزی MLMDC در حدود و یا در برخی موارد پایین‌تر از سایرین است، اما کیفیت جانبی آن به مراتب بهتر و نزدیک Replication (که کیفیت مرکزی خوبی نمی‌دهد) است، البته بجز برای سکانس Akiyo.

ب) برای سکانس Akiyo، منحنی نرخ - کیفیت جانبی نسبت به سایرین بدتر است؛ به این خاطر است که ضرایب DCT، واریانس کم و در نتیجه ER_n زیادتری دارند.

ج) منحنی‌ها نشان می‌دهد که برای $c = 1.3$ ، کیفیت جانبی از Replication بهتر است؛ این همان اثری که گفته شد مانند افزایش f است و در c های پایین رخ می‌دهد.

د) برای دیکودر مرکزی، منحنی‌های MLMDC در نرخ‌های پایین کمی افتاده‌اند. در نرخ‌های پایین، در یک λ و c ثابت، همانطور که قبلاً بحث شده بود کارایی MLMDC با افزایش Q کم می‌شود.

نتایج فوق عملکرد MLMDC را در حالتیکه فقط یک توصیف در دسترس باشد نشان داد. سناریوی دیگری که بررسی می‌شود، پاکت‌های ویدئو هنگام انتقال روی کانال بطور رندوم دچار خطا و یا گم شده، و به هر ترتیب در گیرنده امکان دیکد شدن ندارند. در واقع عملکرد MDC ها را در شبکه‌ی با وجود خطای ارسال بررسی می‌کنیم.

بسته به وضعیت کانال، پاکت‌ها ممکن است دچار خطا شوند و یا سالم برسند. هر پاکت برای اینکه بطور مستقل قابل دیکد شدن باشد بایستی شامل تعداد صحیحی اسلایس باشد و همچنین از بزرگترین واحد

ارسال (MTU¹) کوچکتر باشد تا به هنگام ارسال اصطلاحاً چند تیکه² نشود. حجم گسترده‌ای از end-pointها اینترنت³ است و اندازه MTU در اینترنت 1500 بایت می‌باشد. لذا برای بیشینه کردن حجم ارسالی داده‌های اصلی نسبت به داده‌های سرآمد، اندازه payload را 1460 بایت در نظر می‌گیریم که 40 بایت برای اطلاعات سرآمد RTP/UDP/IP رزرو شده است. در هر اسلایس، MBهای فریم به ترتیب raster قرار داده می‌شوند، که همراه با اطلاعات سرآمد MBها و اسلایس، 1460*8 بیت می‌شوند. در مواردی که تمام یک فریم نمی‌تواند payload پاکت را پر کند، اسلایسها به گونه‌ای خاتمه داده می‌شوند که تعداد صحیحی اسلایس در یک پاکت قرار داده شود، اگرچه اندازه یکی از آنها هم کمتر از 1460 بایت شود.

شایان ذکر است که در این حالت، در Q2MDC و Q4MDC، اسلایس بندی مانند قبل نیست که MBهای نرخ بالا در اسلایس اصلی باشند و MBهای نرخ پایین در اسلایس افزونگی. هر اسلایس تشکیل شده از تعداد مساوی MBهای نرخ پایین و نرخ بالا که تعداد MBها هم با قاعده‌ای که پیش از این برای اندازه اسلایسها گفته شد تعیین می‌شود.

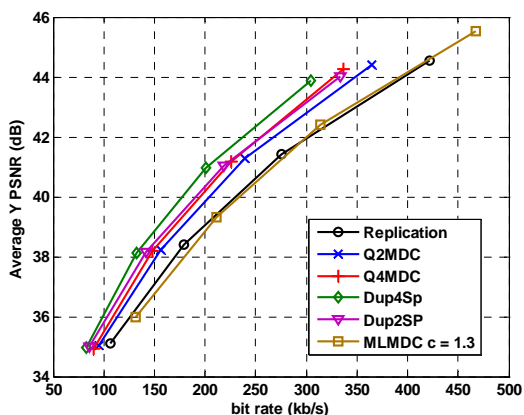
از نظر حلقه پیش‌بینی، MDCها در کلاس B (تعریف شده در [2]) پیاده‌سازی شده‌اند، یعنی مرجع پیش‌بینی همانی است که در SDC استفاده می‌شود. بنابراین با نرسیدن توصیفها، نمی‌توان مرجع را بدون خطا بازسازی کرد و لذا ناهمخوانی مرجع و در نتیجه دریافت خواهیم داشت. در MLMDC، هر ضریب ترکیبی در یک توصیف قرار داده می‌شود و ضریب متناظر آن در سایر توصیفها ضریب پایه است. بنابراین در کدینگ چهارتوصیفه با داشتن حداقل دو توصیف، می‌توان تصویر مرجع را بدون خطا بازسازی کرد.

¹ Maximum Transmission Unit

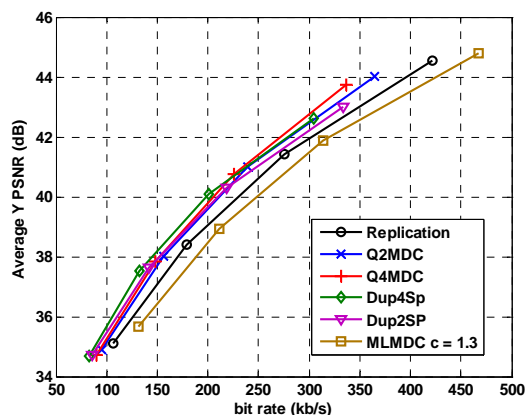
² Fragmentation

³ Ethernet

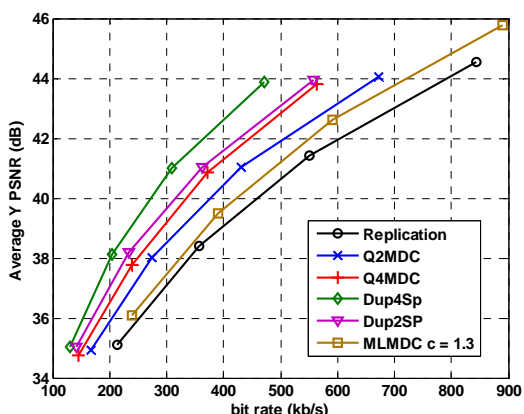
در آزمایشهای زیر، در PLR داده شده، برای هر کانال 50 الگوی خطای رندوم با توزیع برنولی تولید شده و اعمال می‌شوند. کانالها مستقل فرض می‌شوند. در شکل ۳-۱۵ تا شکل ۳-۱۷، نتایج برای ویدئوهای به ترتیب Akiyo، Foreman و Mobile آورده شده است.



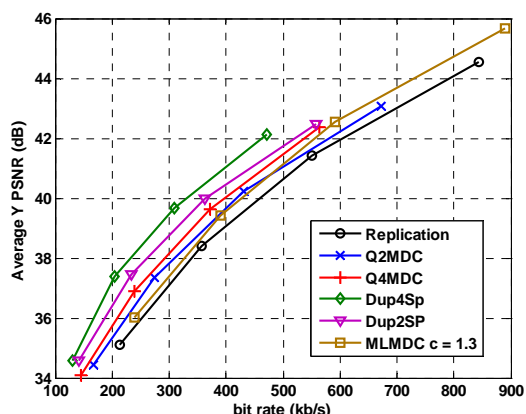
(الف)



(ب)

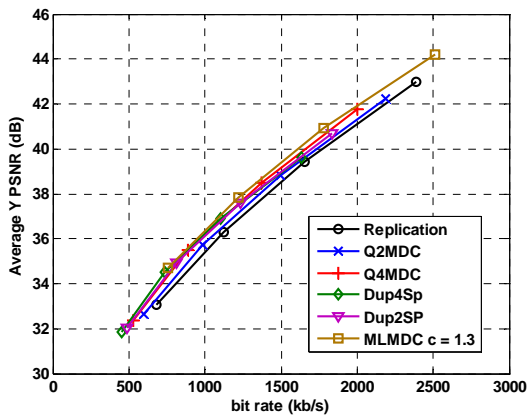


(ج)

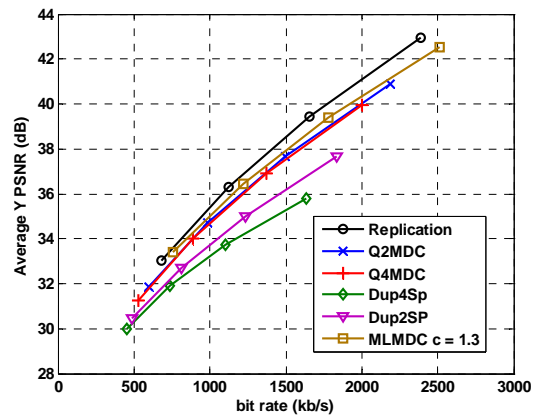


(د)

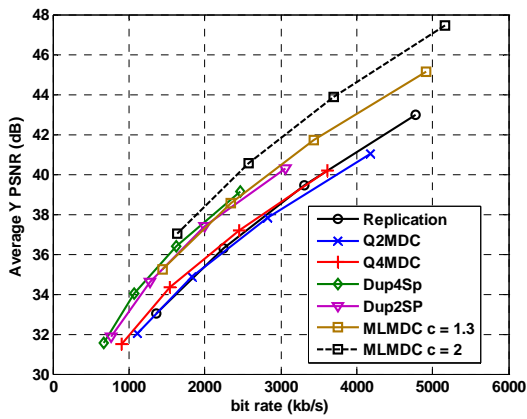
شکل ۳-۱۵ منحنی نرخ-اعوجاج برای ویدئوی Akiyo (الف) کدینگ دو توصیفه و $PLR=5\%$ (ب) کدینگ دو توصیفه و $PLR=20\%$ (ج) کدینگ چهار توصیفه و $PLR=20\%$ (د) کدینگ چهار توصیفه و $PLR=5\%$ (الف) کدینگ چهار توصیفه و $PLR=20\%$



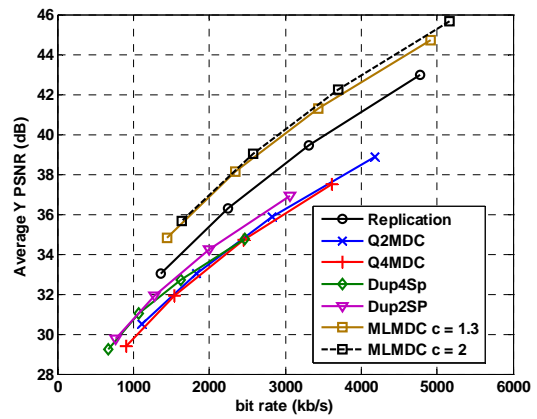
(الف)



(ب)

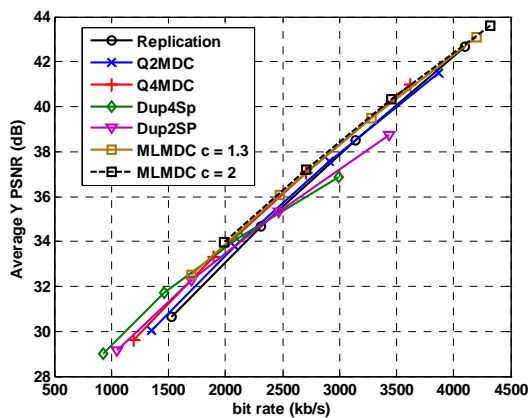


(ج)

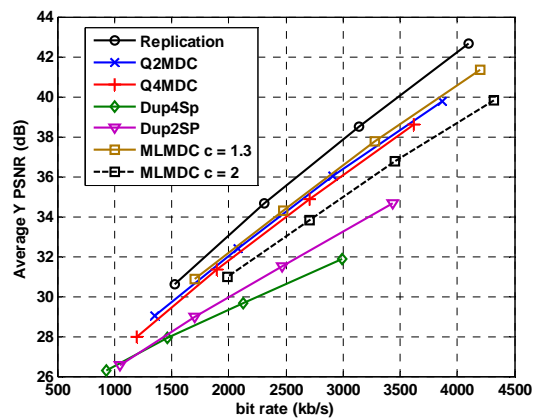


(د)

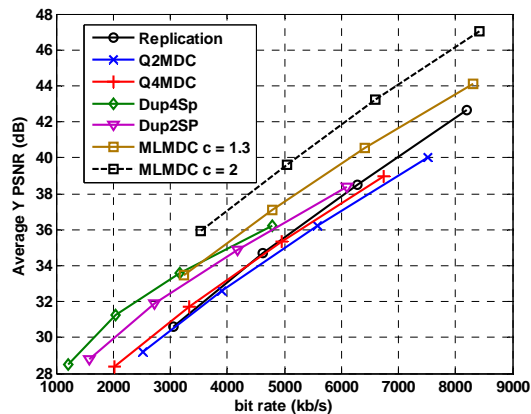
شکل ۳-۱۶ منحنی نرخ-اعوجاج برای ویدئوی Foreman (الف) کدینگ دو توصیفه و $PLR=5\%$ و (ب) کدینگ دو توصیفه و $PLR=20\%$ (ج) کدینگ چهار توصیفه و $PLR=5\%$ و (د) کدینگ چهار توصیفه و $PLR=20\%$



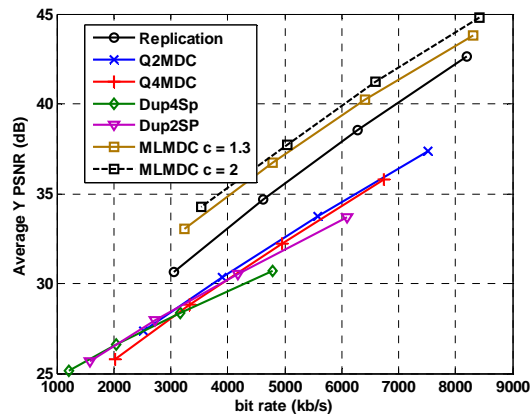
(الف)



(ب)



(ج)



(د)

شکل ۳-۱۷ منحنی نرخ-عوجاج برای ویدئوی Mobile (الف) کدینگ دو توصیفه و $PLR=5\%$ (ب) کدینگ دو توصیفه و $PLR=20\%$ (ج) کدینگ چهار توصیفه و $PLR=5\%$ (الف) کدینگ چهار توصیفه و $PLR=20\%$

از شبیه سازیهای فوق نتایج زیر استنتاج می شود:

۱- برای ویدئوهای Mobile و Foreman، MLMDC بهترین است بجز برای $PLR=20\%$ و کدینگ دو

توصیفه که MLMDC بعد از Replication قرار گرفته است. توجه شود که Replication وقتی نرخ خطا پایین باشد عملکرد خوبی ندارد.

۲- اگر یک MB در دیکودر جانبی، دیکد شود، برای تمام MBهای فریم بعدی که از این MB پیش بینی شده اند هم نمی توان کیفیت برابر کیفیت مرکزی داشت، چونکه بلوک مرجع سالم نبوده است و در واقع انتشار خطا رخ می دهد. این مسئله در نرخ خطاهای بالا و کدینگ دو توصیفه تعیین کننده شده و به همین دلیل Replication که عاری از انتشار خطا است در نرخ خطای بالا بهترین کیفیت را دارد.

۳- MLMDC برای ویدئوهای با محتوای پایین مانند Akiyo عملکرد خوبی ندارد، که البته علت آن قبلاً بحث شده بود.

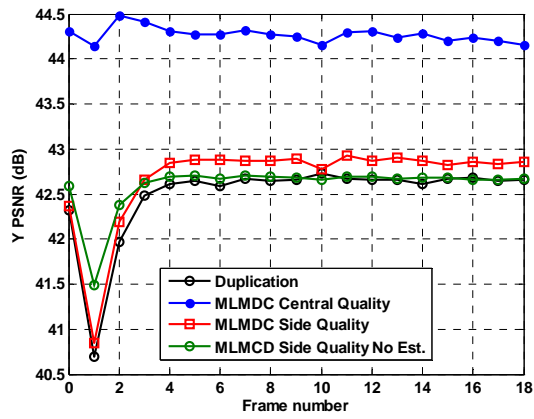
۴- مقایسه کدینگ دو توصیفه و چهار توصیفه نشان می دهد که از نظر عملکرد نرخ-عوجاج، در هر دو مقدار PLR ، کدینگ دو توصیفه بهتر عمل می کند. در نرخ خطای به اندازه کافی بالا، روشن است که کارایی چهار توصیفه بهتر از دو توصیفه خواهد شد اما کانالها معمولاً آنقدر نویزی نیستند. ولی دلایل دیگری همچنان

برای استفاده از چهار توصیف و حتی بیشتر وجود دارد. برای مثال، برای غلبه بر ناهمگنی^۱ پهنای باند، MDC با تعداد توصیفهای بیشتر، مقیاس‌پذیری بیشتری هم فراهم می‌کند. به علاوه، MDC معمولاً با MPT^۲ همراه است که هر توصیف روی یک مسیر جدا ارسال می‌شود. بعضی از ویژگیهای MPT مانند پهنای باند مجموع، کاهش احتمال قطع ارتباط، کاهش تغییرات تاخیر، با تعداد مسیرها و در نتیجه تعداد توصیفها بهبود می‌یابد [72]. با تعداد بیشتر توصیفها، پهنای باند مورد نیاز هر کانال هم کاهش می‌یابد. برای نمونه، برای داشتن کیفیت 40dB هنگامی که PLR برابر 20% است، در MLMDC و $c = 2$ ، سکانس Mobile، در کدینگ دو توصیفه حداقل پهنای باند کانال 2150 Kb/s است (شکل ۳-۱۷ ب)) ولی در کدینگ چهار توصیفه این مقدار به عدد 1500 Kb/s کاهش می‌یابد (شکل ۳-۱۷ د)).

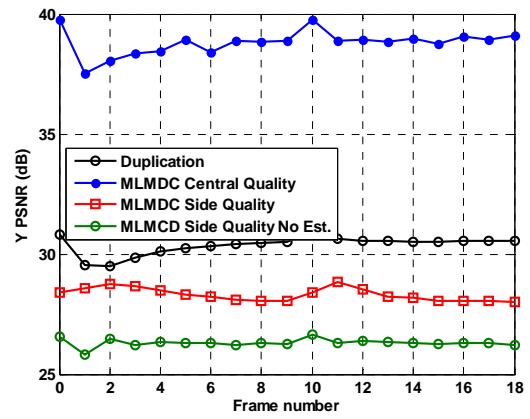
قبلاً نشان دادیم که برای مقادیر کوچک c ، به گونه‌ای که (۳-۲۶) ارضا شود، تابع تخمین تقریباً به صورت تابع همانی است و این یعنی که تخمین نیاز نیست. منحنی شکل ۳-۱۸ این ادعا را تأیید می‌کند، آنها نمونه‌هایی هستند که در آنها c بزرگتر و کوچکتر از حد داده شده در (۳-۲۶) است و تأثیر تابع تخمین را در PSNR نشان می‌دهد. این منحنی‌ها، نمایش کیفیت فریم به فریم شبیه‌سازی شکل ۳-۱۴ هستند که حالت بدون تخمین (No Estimation) هم اضافه شده است. همچنین یک کاهش شدید در PSNR فریم اول مشاهده می‌شود که به خاطر این است که f (rounding offset) هنوز به مقدار بهینه خود همگرا نشده است. در $QP = 32$ ، تأثیر این f نابهینه کمتر است.

¹ Heterogeneity

² Multi-Path Transmission



(الف)



(ب)

شکل ۳-۱۸ تاثیر تخمین زدن برای (الف) $c = 1.3$ و $QP = 20$ و (ب) $c = 3.2$ و $QP = 32$

لذا برای مقادیر کوچک c نسبت به $1/\lambda Q$ ، تخمین کاری انجام نمی‌دهد؛ یعنی اینکه یک دیکودر استاندارد بدون هیچ گونه اصلاحی قابل استفاده خواهد بود. اگر چه برای دیکودر مرکزی، الگوریتم ادغام کردن توصیف‌ها باید اضافه شود.

۴. بهینه‌سازی MLMDC

سازگار با کانال

۱.۴. مقدمه

همانطور که قبلاً گفته شد، در شرایطی که خطای کانال بالا است نیاز است که اعوجاج جانبی خوبی داشته باشیم چون احتمال از دست دادن توصیفها زیاد است. بر عکس، هنگامیکه نویز کانال کم است غالباً همه توصیفها در دسترس هستند و لذا بهتر است که کیفیت مرکزی بالایی داشته باشیم. بنابراین، برای داشتن کیفیت بهینه در شرایط مختلف، بایستی منطبق با وضعیت کانال MDC را طراحی کنیم. به عبارت دیگر، بایستی نحوه مصالحه بین کیفیت جانبی و کیفیت مرکزی با توجه به شرایط کانال بهینه شود. ابزار کنترلی آن پارامتر MDC است که توصیفها را از یکدیگر تفکیک می‌کند و مصالحه بین کیفیت جانبی و کیفیت مرکزی را تعیین می‌کند. بنابراین به منظور بهینه کردن کیفیت، باید ابتدا ارتباط بین کیفیت ویدئو و پارامترهای MDC و نرخ خطای کانال را پیدا کرد.

اعوجاج در دیکودر از دو چیز ناشی می‌شود: اعوجاج کوانتیزیشن و اعوجاج کانال. اعوجاج کوانتیزیشن از کوانتایز کردن ضرایب DCT حاصل می‌شود. اعوجاج کانال از دچار خطا شدن داده حین ارسال ناشی می‌شود، این اعوجاج در واقع اختلاف بین داده ارسال شده و آنی است که در دیکودر بازیافت می‌شود.

بنابراین، در هنگام انتقال ویدئو روی کانال، هر دو اعوجاج را باید در نظر بگیریم، به این منظور، اعوجاج انتها-تا-انتها^۱ را تعریف می‌کنند که در بر گیرنده اعوجاج کوانتیزیشن و اعوجاج کانال می‌باشد. از ملزومات اعوجاج انتها-تا-انتها، پیدا کردن یا فرموله کردن دریافت و یا انتشار خطا می‌باشد. دریف یا انتشار خطا نتیجه اعوجاج کانال است که باعث اختلاف یا ناهمخوانی^۲ مرجع پیش‌بینی در انکودر و دیکودر می‌شود. استانداردهای فشرده‌سازی ویدئو از کدینگ مبتنی بر پیش‌بینی استفاده می‌کنند، چون کارایی بسیاری در کاهش نرخ بیت دارد؛ ولی این کار از طرف دیگر باعث ایجاد وابستگی کیفیت یک فریم به کیفیت فریمی که مرجع پیش‌بینی است، می‌گردد. وقتی فریم جاری کاملاً صحیح دریافت می‌شود ولی مرجع آن بطور بی‌خطا بازسازی نشده باشد، فریم جاری را نمی‌توان بدون خطا و آنطور که در انکودر بوده است، بازسازی کرد. این فریم جدید که خود خطا دارد، مرجع جبران‌سازی حرکت فریمهای بعدی قرار می‌گیرد که باعث کاهش کیفیت آنها می‌شود. به همین ترتیب خطا منتشر می‌شود تا اینکه در انتهای GOP، فریم اینترا برسد. به این علت، شماره فریم در GOP که در آن خطا رخ داده است مهم است. به عنوان مثال، از دست دادن داده‌های فریم اول خیلی بیشتر از فریم آخر باعث کاهش کیفیت می‌گردد.

انتشار خطا ویژه MDC نیست و در SDC هم وجود دارد. در SDC، داده‌های نرسیده با شیوه‌های اختفای خطا جایگزین می‌شوند و ناهمخوانی از اختلاف بین داده اصلی و داده بازیافت شده ناشی می‌شود. در MDC هم وقتی حداقل یک توصیف در دسترس نیست، انکودر قادر به بازسازی دقیق داده‌ها نیست و لذا ناهمخوانی رخ می‌دهد. در این بخش فقط MDC دو-توصیفه را در نظر می‌گیریم و چون به ندرت اتفاق می‌افتد که هر دو توصیف دچار خطا شوند، ناهمخوانی بیشتر بین دیکودر جانبی و دیکودر مرکزی وجود دارد.

¹ End-to-End Distortion

² Mismatch

در این بخش بر پایه ناهمخوانی بین خروجی دیکودر مرکزی و دیکودر جانبی، خطایی که از فریمهای قبلی انتشار یافته است محاسبه و در اعوجاج فریم جاری لحاظ می‌شود. بیشتر کارهایی که MDC ها را بهینه کرده‌اند انتشار خطا را در نظر نگرفته‌اند، که مسئله مهمی در ارسال ویدئو روی کانالهای نویزی است. در این بخش، با در نظر گرفتن انتشار خطا و با دانستن نرخ خطای کانال، مقدار بهینه پارامتر MDC بدست می‌آید. ابتدا، مدلی برای پیدا کردن اعوجاج انتها-تا-انتها ارائه و سپس دقت این مدل در مقایسه با نتایج عملی سنجیده می‌شود. پس از تایید دقت مدل، MDC های مورد بررسی بهینه شده، نتایج بهینه با نتایج غیر بهینه مقایسه می‌گردند، همچنین کارایی MLMDC نسبت به دیگر MDC ها مشاهده می‌شود.

۲.۴. کارهای مرتبط

تعداد کمی از مقالات هستند که بهینه‌سازی MDC را با در نظر گرفتن انتشار خطا انجام داده‌اند. در [85] و [86]، اعوجاج انتها-تا-انتها محاسبه شده است اما برای تعیین مقدار کلی افزونگی مورد استفاده قرار گرفته است. این مقدار بطور مساوی بین فریمها تقسیم شده است که لزوماً بهینه نیست. اما در [87]، تعیین فریم به فریم افزونگی مطرح شده است، این کار بر اساس الگوریتم مشهوری به نام ROPE¹ صورت می‌گیرد [88]. در ROPE، با یک رفتار بازگشتی، ممان اول و دوم مقدار پیکسلهای بازسازی شده در دیکودر بدست می‌آید و اعوجاج بطور آماری برای هر پیکسل محاسبه می‌شود. با این رویکرد، در هر حالت اینترا یا اینتر، اعوجاجها با هم مقایسه و حالت بهتر انتخاب می‌شود. برای بهینه‌سازی پارامتر MDC، باید که برای هر مقدار آن، اعوجاج انتها-تا-انتها را حساب و سپس آن مقداری که کمینه اعوجاج را می‌دهد، بهینه است. روشن است که این روش از نظر حجم محاسبات مورد نیاز، روش جالبی نیست، بویژه اینکه خود ROPE در حالت پایه (انتخاب حالت اینترا-اینتر) پیچیدگی زیادی دارد به گونه‌ایکه یکی از معایب آن محسوب می‌شود. تاثیر

¹ Recursive Optimal Per-pixel Estimate

انتشار خطا در طراحی MDC در دو مقاله [84] و [58] هم مورد مطالعه قرار گرفته است. در [84]، بر اساس سهم هر ماکرو بلاک در جبران حرکت، اهمیت آن در انتشار خطا سنجیده شده و پس از آن، مقدار افزودنی تعیین می‌شود. اما، این کار مستلزم این است که برای هر بلوک، تحلیل حرکت در طی GOP انجام شود که طبیعتاً مشکل محاسبات زیاد وجود خواهد داشت. علاوه بر این، ارسال پارامتر MDC برای هر بلاک موجب افزایش نرخ بیت می‌شود. برای تطبیق افزودنی در این مقاله، QP هر بلوک تنظیم می‌شود که در MDC مورد بحث در آن مقاله این امر براحتی امکان پذیر است. اما برای سایر روشهای MDC، که پارامتر MDC آنها QP نیست، نمی‌توان این کار را به آسانی انجام داد، یعنی افزودنی را بلوک به بلوک تغییر داد. در مقاله [58]، افزودنی در هر فریم تعیین می‌شود اما بهینه‌سازی انجام شده برای این نوع خاص MDC است. به علاوه، از مدل ارائه شده در [89] برای وارد کردن انتشار خطا استفاده شده که یک مدل تجربی است و برای نرخ خطاهای بالا خوب جواب نمی‌دهد [90]. رویکردهای بازگشتی ارائه شده در [91] و [92] که انتشار خطا را در SDC مدل کرده‌اند دقیق‌تر و ملموس‌تر از مدل [89] هستند و به همین خاطر، ما هم از مدل بازگشتی استفاده می‌کنیم. بر اساس ناهمخوانی که در تصویر دیکود شده در دیکودرهای جانبی و دیکودر مرکزی اتفاق می‌افتد، و با در نظر گرفتن اثر اینترا کدینگ، انتشار خطا فرموله شده و در نهایت اعوجاج نهایی و یا همان انتها-تا-انتها در دیکودر محاسبه می‌شود. مدل ارائه شده، کلی است و نه فقط برای MLMDC، بلکه برای سایر MDC های حوزه DCT قابل استفاده است، چنانچه استفاده هم شده است. پس از آن، تابع هدف برای بهینه‌سازی اعوجاج انتها-تا-انتها با قید نرخ بیت تشکیل شده و MDC ها بطور بهینه طراحی می‌شوند. در دو بخش بعدی، نحوه فرموله کردن و بهینه‌سازی مسئله ارائه می‌شود.

۳.۴. مدل پیشنهادی برای اعوجاج انتها-تا-انتها

بر پایه ناهمخوانی بین بلوکهای دیکودر شده توسط دیکودر مرکزی و دیکودر جانبی، خطای منتشر شده از فریمهای قبلی محاسبه شده و با متوسط‌گیری روی اعوجاج بلوکها، اعوجاج هر فریم در GOP بدست می‌آید.

برای بلوک m ام از فریم n ، $B^n(m)$ در صورتیکه در حالت اینتر کد شده باشد، داریم:

$$B^n(m) = \hat{B}^{n-1}(k_m) + X^n(m) \quad (۱-۴)$$

که در آن $X^n(m)$ بلوک باقیمانده مربوطه است، $\hat{B}^{n-1}(k_m)$ بلوک مرجع پیش‌بینی در فریم $n-1$ است که توسط بردار حرکت به آن اشاره می‌شود. فریم $n-1$ در بافر انکودر موجود است. بلوک باقیمانده بعد از کد شدن، دوباره دیکد شده، بازسازی می‌شود و در بافر انکودر به منظور تخمین حرکت بلوکهای بعدی ذخیره می‌گردد:

$$\hat{B}^n(m) = \hat{B}^{n-1}(k_m) + \hat{X}_0^n(m) \quad (۲-۴)$$

$\hat{X}_0^n(m)$ در (۲-۴)، با اندیس 0 به این معنی است که در ساختن مرجع پیش‌بینی از هر دو توصیف استفاده شده است. اما به دلیل وجود خطای ارسال، چه در فریم جاری و چه در فریمهای قبلی، بلوک باقیمانده و بلوک مرجع می‌توانند در انکودر و دیکودر یکسان نباشند. بلوک بازسازی شده در دیکودر به این صورت است:

$$\tilde{B}^n(m) = \tilde{B}^{n-1}(k_m) + \tilde{X}^n(m). \quad (۳-۴)$$

که $\tilde{B}^{n-1}(k_m)$ بلوک بازسازی و ذخیره شده در بافر انکودر است؛ بسته به توصیف دریافتی، $\tilde{X}^n(m)$ می‌تواند به شکلهای زیر باشد:

$$\tilde{X}^n(m) = \begin{cases} \hat{X}_0^n(m) & \text{central decoder} \\ \hat{X}_1^n(m) & \text{side decoder 1} \\ \hat{X}_2^n(m) & \text{side decoder 2} \end{cases} \quad (4-4)$$

البته یک حالت دیگر هم امکان پذیر است؛ آن وقتی است که هر دو توصیف نرسیده باشند و بلوک بایستی با شیوه‌های اختفای خطا بازیافت شود. اما بجز در موارد خطای بالا، احتمال اینکه هر دو توصیف یک اسلایس، همزمان دچار خطا شوند بسیار ناچیز است و لذا در تحلیل ما این حالت چشم‌پوشی شده است.

بنابراین، خطای موجود در یک بلوک ایترا از دو چیز نشأت می‌گیرد: یکی از دست رفتن اطلاعات باقیمانده و دیگری خطای بلوکهای قبلی. اما برای بلوکهای ایترا، پیش‌بینی بین فریمی نداریم و لذا در سمت انکودر داریم که:

$$\hat{B}^n(m) = \hat{X}_0^n(m) \quad (5-4)$$

و در سمت دیکودر

$$\tilde{B}^n(m) = \tilde{X}^n(m) \quad (6-4)$$

که $\tilde{X}^n(m)$ همان است که در (4-4) تعریف شد. توجه شود که پیش‌بینی ایترا که در استاندارد H.264/AVC وجود دارد، را در نظر نمی‌گیریم. در حالت کلی پیش‌بینی ایترا بازدهی فشرده‌سازی بهتری نسبت به ایترا کدینگ بدون پیش‌بینی دارد ولی در عوض خطاپذیری کمتری دارد. اما از آنجا که در اسلایسهای نوع P¹ تعداد کمی بلوکها هستند که ایترا کد می‌شوند، لذا داشتن و یا نداشتن پیش‌بینی ایترا تاثیر چندانی در فشرده‌سازی ندارد (مگر در فریمهای ایترا). لذا برای وارد نشدن به مسائل پیش‌بینی ایترا در اعوجاج انتها-تا-انتها، کلاً بلوکهای ایترا را بدون پیش‌بینی کد می‌کنیم.

¹ در اسلایسهای نوع P، بلوکها می‌توانند ایترا یا ایترا کد شوند ولی حداقل یک بلوک ایترا است.

حال به منظور اندازه‌گیری خطای انتها-تا-انتها، $Y^n(m)$ را اختلاف بین بلوک با مقدار اصلی پیکسلها و بلوک با مقدار بازسازی شده پیکسلها در دیکودر، تعریف می‌کنیم:

$$\begin{aligned} Y^n(m) &= B^n(m) - \tilde{B}^n(m) = B^n(m) - \hat{B}^n(m) + \hat{B}^n(m) - \tilde{B}^n(m) \\ &= Y_Q^n(m) + Y_c^n(m), \end{aligned} \quad (۷-۴)$$

که در آن

$$Y_Q^n(m) = B^n(m) - \hat{B}^n(m) \quad (۸-۴)$$

معرف خطای کوانتیزیشن و

$$Y_c^n(m) = \hat{B}^n(m) - \tilde{B}^n(m) \quad (۹-۴)$$

معرف خطای کانال است. در مقاله [91] نشان داده شده است که اعوجاج کانال و اعوجاج کوانتیزیشن از یکدیگر مستقل هستند، یعنی اینکه می‌توان آنها را جداگانه حساب کرد و سپس برای پیدا کردن اعوجاج کل یا اعوجاج انتها-تا-انتها آنها را با هم جمع کرد. جایگزینی $\hat{B}^n(m)$ و $\tilde{B}^n(m)$ در معادله (۹-۴) از روابط (۲-۴) و (۳-۴) برای بلوکهای اینتر می‌دهد:

$$\begin{aligned} Y_c^n(m) &= \hat{B}^{n-1}(k_m) + \hat{X}_0^n(m) - \tilde{B}^{n-1}(k_m) - \tilde{X}^n(m) \\ &= \hat{B}^{n-1}(k_m) - \tilde{B}^{n-1}(k_m) + \hat{X}_0^n(m) - \tilde{X}^n(m) \\ &= Y_c^{n-1}(k_m) + \delta_X^n(m), \end{aligned} \quad (۱۰-۴)$$

و از روابط (۴-۵) و (۴-۶) برای بلوکهای اینترا می‌دهد:

$$Y_c^n(m) = \hat{X}_0^n(m) - \tilde{X}^n(m) = \delta_X^n(m), \quad (۱۱-۴)$$

که در آن، $\delta_X^n(m)$ ناهمخوانی بین بلوک دیکد شده در انکودر و دیکودر است؛ یعنی

$$\delta_X^n(m) = \hat{X}_0^n(m) - \bar{X}^n(m) = \begin{cases} 0 & \text{central decoder} \\ \delta_{01}^n(m) = \hat{X}_0^n(m) - \hat{X}_1^n(m) & \text{side decoder 1} \\ \delta_{02}^n(m) = \hat{X}_0^n(m) - \hat{X}_2^n(m) & \text{side decoder 2} \end{cases} \quad (12-4)$$

اگر β میزان نسبی تعداد بلوکهای اینترا در فریم باشد، اعوجاج کانال برای فریم n م به شکل زیر است:

$$\begin{aligned} D_c^n &= \frac{1}{M} \sum_{m=1}^M E \left[(Y_c^n(m))^2 \right] \\ &= \frac{1}{M} \sum_{m=1}^{(1-\beta)M} E \left[(Y_c^{n-1}(k_m) + \delta_{X_{inter}}^n(m))^2 \right] \\ &\quad + \frac{1}{M} \sum_{m=(1-\beta)M+1}^M E \left[(\delta_{X_{intra}}^n(m))^2 \right] \\ &= \frac{1}{M} \sum_{m=1}^{(1-\beta)M} E \left[(Y_c^{n-1}(k_m))^2 \right] \\ &\quad + \frac{1}{M} \left(\sum_{m=1}^{(1-\beta)M} E \left[(\delta_{X_{inter}}^n(m))^2 \right] + \sum_{m=(1-\beta)M+1}^M E \left[(\delta_{X_{intra}}^n(m))^2 \right] \right) \end{aligned} \quad (13-4)$$

که در آن M تعداد کل بلوکها در یک فریم است، و سیگنال ناهمخوانی (تعریف شده در (4-12)) برای بلوکهای اینترا و اینترا جداگانه و به ترتیب با $\delta_{X_{inter}}^n(m)$ و $\delta_{X_{intra}}^n(m)$ نشان داده شده است. همبستگی بین سیگنال ناهمخوانی بلوکهای اینترا، $\delta_{X_{inter}}^n(m)$ ، و اختلاف انتها-تا-انتها، $Y_c^{n-1}(k_m)$ ، صفر فرض شده است؛ گرچه معمولاً اینطور هم است و این فرض، فرضی منطقی است. هنگامیکه m بلوکهای فریم n را جاروب می‌کند، k_m بلوکهای فریم $n-1$ را جاروب می‌کند. با فرض اینکه این $(1-\beta)M$ بلوک، که احتمالاً هم پوشانی هم دارند، نماینده تمام بلوکها باشند می‌توانیم بگوییم که:

$$\frac{1}{M} \sum_{m=1}^{(1-\beta)M} E \left[(Y_c^{n-1}(k_m))^2 \right] = \frac{(1-\beta)}{M} \sum_{k=1}^M E \left[(Y_c^{n-1}(k))^2 \right] = (1-\beta)D_c^{n-1} \quad (14-4)$$

با جایگزینی (4-14) در (4-13) خواهیم داشت:

$$D_c^n = (1 - \beta)D_c^{n-1} + \Delta_X^n \quad (15-4)$$

که در آن

$$\Delta_X^n = \frac{1}{M} \left(\sum_{m=1}^{(1-\beta)M} E \left[\left(\delta_{X_{inter}}^n(m) \right)^2 \right] + \sum_{m=(1-\beta)M+1}^M E \left[\left(\delta_{X_{intra}}^n(m) \right)^2 \right] \right) \quad (16-4)$$

اعوجاج ناهمخوانی مرتبط با فریم n ام است. حال اگر فریم اول GOP که I-frame است سالم به دیکودر برسد، یعنی اینکه هر دو توصیف برای تمام اسلایسها رسیده باشد، داریم:

$$D_c^0 = 0 \quad (17-4)$$

با بکار بردن معادله بازگشتی (4-15) اعوجاج کانال از رابطه زیر بدست می‌آید:

$$D_c^n = \sum_{i=0}^n (1 - \beta)^{n-i} \Delta_X^i \quad (18-4)$$

از معادله (4-18) نتیجه می‌شود که به منظور پیدا کردن اعوجاج کانال در هر فریم، بایستی اعوجاج ناهمخوانی (معادله (4-16)) را برای فریم جاری و همچنین تمام فریمهای قبل که در جبران حرکت استفاده شده‌اند را بدست آورد. اعوجاج ناهمخوانی یک فریم، از متوسط‌گیری روی اعوجاج ناهمخوانی در تمام بلوکهای (چه اینترا و چه اینتر) آن فریم بدست می‌آید، همانطور که در (4-16) آورده شده است. از معادله (4-12)، اعوجاج ناهمخوانی مربوط به بلوک $X^i(m)$ برابر است با

$$E \left[\left(\delta_X^i(m) \right)^2 \right] = (1 - P)^2 0 + P(1 - P) E \left[\left(\delta_{01}^i(m) \right)^2 \right] + P(1 - P) E \left[\left(\delta_{02}^i(m) \right)^2 \right] \quad (19-4)$$

که در آن $\delta_{01}^i(m)$ و $\delta_{02}^i(m)$ در معادله (4-12) تعریف شده‌اند، و $(1 - P)^2$ و $P(1 - P)$ به ترتیب احتمال دیکدینگ مرکزی و دیکدینگ جانبی هستند. همانطور که قبلاً هم اشاره شد، حالتی که هر دو توصیف در دسترس نباشند را در نظر نمی‌گیریم. اگر که توصیفها را مقارن فرض کنیم، جمله‌های مربوط به دیکودرهای جانبی با هم برابر هستند و بنابراین رابطه (4-19) به فرم زیر تبدیل می‌شود:

$$E \left[\left(\delta_X^i(m) \right)^2 \right] = 2P(1-P)E \left[\left(\delta_{01}^i(m) \right)^2 \right]. \quad (۲۰-۴)$$

به این ترتیب با داشتن $E \left[\left(\delta_{01}^i(m) \right)^2 \right]$ می‌توان اعوجاج انتها-تا-انتها را بدست آورد. در بخش بعدی چگونگی اعمال این رویکرد نشان داده می‌شود.

۴.۴. پیاده‌سازی مدل و تشکیل تابع هدف

در این بخش، نشان می‌دهیم که چگونه مدل کلی ارائه شده در بخش فوق در MDC ها عملی شده و برای بهینه‌سازی مورد استفاده قرار گیرد. به این منظور، MLMDC و دو MDC دیگر را در مورد مطالعه قرار می‌دهیم. MDC ها، همه حوزه DCT هستند، به این معنا که عملیات MDC روی ضرایب DCT انجام می‌گیرد. در این حالت بهتر است که اعوجاج‌ها هم در حوزه DCT نوشته شود. تبدیل DCT یک تبدیل متعامد یکه^۱ است و لذا (۲۰-۴) را می‌توان بدون هیچ تغییری در حوزه DCT فرض کرد، اگر که \bar{X}_1 و \bar{X}_0 در معادله (۴-۱۲) خروجی‌های دیکودر مرکزی و جانبی اما در حوزه DCT باشند. همانطور که قبلاً هم بحث شده بود، ضرایب DCT را می‌توان با توزیع لاپلاسین مدل کرد. این خاصیت به هنگام محاسبه اعوجاج فریم سودمند است؛ بجای متوسط گرفتن از اعوجاج بلوکها، اعوجاج را برای یک بلوک شامل این ضرایب با توزیع آماری بدست آمده روی کل فریم، محاسبه می‌شود. بویژه، اگر بخواهیم اعوجاج ناهمخوانی داده شده در (۴-۱۶) را برای فریم m بدست آوریم، ضرایب بلوکهای اینتر را با ضرایب بلوکهای اینترا، با هم مشاهده کرده و یک توزیع برای آنها بدست می‌آوریم. حال اعوجاج را برای بلوکی از ضرایب با توزیع آماری حاصل بدست می‌آوریم. در بعضی بلوکهای اینتر به دلیل وجود عدم جبران حرکت خوب و همچنین در بلوکهای اینترا، ضرایب DCT توزیع پهن تری نسبت به ضرایب اینتر دارند، لذا در نظر گرفتن یک توزیع

¹ Orthonormal

خاص برای تمام بلوکهای فریم کار دقیقی نیست. اما شبیه‌سازیهای انجام شده نشان دادند که خطای این تقریب در کاربرد ما به اندازه کافی کوچک است. بنابراین (۴-۱۶) را می‌توان به صورت زیر نوشت:

$$\Delta_X^i \approx E \left[\left(\delta_{X[inter\ intra]}^i \right)^2 \right] \quad (۲۱-۴)$$

که در آن اندیس $[inter\ intra]$ یعنی اینکه ضرایب اینتر و اینترا با هم در نظر گرفته شده‌اند. برای سادگی، از این پس این اندیس نشان داده نمی‌شود و مانند قبل بصورت δ_X^i نوشته می‌شود. با دانستن Δ_X^i می‌توان از رابطه (۴-۱۸) اعوجاج کانال را بدست آورد.

اعوجاج کوانتیزیشن، جمله بعدی اعوجاج انتها-تا-انتها، را می‌توان دوباره در حوزه DCT بصورت زیر بدست آورد:

$$D_Q^n = \frac{1}{M} \sum_{m=1}^M E \left[\left(Y_Q^n(m) \right)^2 \right] \approx E \left[\left(\delta_{Q[inter\ intra]}^n \right)^2 \right] \quad (۲۲-۴)$$

که در آن $\delta_{Q[inter\ intra]}^n$ همان Y_Q^n تعریف شده در (۴-۸) است، و ضرایب اینترا و اینتر را با هم در نظر گرفته ایم که از پس با δ_Q^n نشان داده می‌شود. اکنون، اعوجاج کل بصورت زیر بدست می‌آید:

$$D_t^n = D_c^n + D_Q^n = \sum_{i=0}^n \left((1 - \beta)^{n-i} \Delta_X^i \right) + E \left[\left(\delta_Q^n \right)^2 \right] \quad (۲۳-۴)$$

معادله (۴-۲۳) مدلی کاملاً تحلیلی است که با داشتن تابع ناهمخوانی و همچنین نرخ خطای کانال، قادر به پیش‌بینی اعوجاج یک فریم در GOP است.

به منظور یافتن مقدار بهینه پارامتر MDC، می‌توان تابع هدفی به فرم زیر تعریف کرد:

$$\min_{\{P\}} \{D_{GOP}\} = \min_{\{P\}} \left\{ \sum_{n=0}^{N-1} D_t^n \right\} \quad (۲۴-۴)$$

$$s. t. \sum_{n=0}^{N-1} R_1^n + R_2^n \leq R_t$$

که در آن \mathbf{P} بردار پارامترهای MDC است، D_{GOP} اعوجاج کل روی یک GOP است، N اندازه GOP است، R_1^n و R_2^n تعداد بیت‌های فریم n ام به ترتیب در توصیف اول و دوم هستند، R_t حد بیشینه تعداد بیت‌های کل GOP مورد بهینه‌سازی است. با جایگزینی از (۴-۲۳) داریم:

$$\min_{\{\mathbf{P}\}} \{D_{GOP}\} = \min_{\{\mathbf{P}\}} \left\{ \sum_{n=0}^{N-1} \left(\gamma_n \Delta_X^i + E \left[(\delta_Q^n)^2 \right] \right) \right\}, \quad \gamma_n = \sum_{j=0}^{N-1-n} (1 - \beta)^j \quad (۴-۲۵)$$

$$s. t. \sum_{n=0}^{N-1} R_1^n + R_2^n \leq R_t$$

مسئله (۴-۲۵) یک مسئله بهینه‌سازی با قيود است که می‌توان آنرا با روش لاگرانژ حل کرد.

دیدیم که اعوجاج ناشی از کانال در طول GOP انتشار پیدا می‌کند، اما شدت آن با اینترا کدینگ (β) غیر صفر) کاهش پیدا می‌کند که این کاهش با ضریب γ در معادله (۴-۲۵) نمود پیدا کرده است. در عمل، عامل دیگری هم باعث کاهش اثر مخرب انتشار خطا می‌شود و آن فیلتر کردن پیکسلها است. این فیلترینگ که نتیجه جبران حرکت 1/2 پیکسل و یا 1/4 پیکسل و deblocking filtering است مانند یک متوسط‌گیر عمل کرده و نویز منتشر شده از فریم قبلی را کاهش می‌دهد. اما، مدل کردن تحلیلی این اثر بسیار مشکل است و لذا از روشهای تجربی، مانند آنچه در [89] و [92] پیشنهاد شده است، استفاده می‌کنند. پارامترهایی که مدل را کالیبره می‌کنند، برای هر ویدئوی خاص با روشهای منحنی-نشانند^۱ بدست می‌آید. در این کار، از آنجا که به دنبال یک مدل کاملاً تحلیلی که مناسب بهینه‌سازی باشد بودیم، با برداشتن عملگرهای زیر-پیکسلی اثرات فیلترینگ را هم در خروجی نمی‌بینیم.

^۱ Curve fitting

۱.۴.۴. تابع ناهمخوانی برای MDC های مورد مطالعه

دیدیم که برای یافتن اعوجاج انتها-تا-انتها، نیاز است که تابع ناهمخوانی را برای MDC ها بیابیم. در این قسمت، این کار را برای MLMDC و دو MDC دیگر به نام DupSp [41] و QMDC [58] بدست می آوریم. DupSp و QMDC از MLMDC سراسر تر هستند، به همین خاطر آنها را زودتر مورد بررسی قرار می دهیم.

DupSp: DupSp ضرایب AC که از یک سطح معینی بزرگتر هستند را در هر دو توصیف تکرار می شوند و سایر ضرایب بین دو توصیف تقسیم می گردند. ضریب DC در هر دو توصیف تکرار می شود. برای این نوع MDC، پارامتر MDC سطح جداسازی ضرایب تکرار شونده و ضرایب تقسیم شونده است. بنابراین

$$\mathbf{P} = \mathbf{t} = \{t^{(0)}, t^{(1)}, \dots, t^{(N-1)}\} \quad (۲۶-۴)$$

که در آن، $t^{(n)}$ سطح تصمیم گیری برای فریم n ام است. مقدار این سطح، تفاوت بین سیگنال جانبی و سیگنال مرکزی را تعیین می کند. برای مقادیر کوچک $t^{(n)}$ ، توصیفها اطلاعات مشترک بیشتری دارند و لذا ناهمخوانی آنها کمتر است، در حالیکه با بزرگتر شدن $t^{(n)}$ تفاوت دو توصیف بیشتر شده و سیگنال ناهمخوانی هم بیشتر می شود.

اگر x_Q ، مقدار x_k (ضریب k ام بلوک DCT) بعد از کوانتیزیشن باشد، ناهمخوانی در $t^{(i)}$ داده شده بصورت زیر است:

$$\delta_{01k}^D = \begin{cases} 0 & |x_Q| > t^{(i)} \\ x_Q & |x_Q| \leq t^{(i)} \end{cases} \quad (۲۷-۴)$$

بالانویس D نشان می دهد که این تابع برای DupSp در نظر گرفته شده است. اگر λ پارامتر لاپلاس سیگنال x باشد، توان سیگنال ناهمخوانی بصورت زیر بدست می آید:

$$E \left[(\delta_{01k}^D)^2 \right] = \sum_{T=-t^{(i)}}^{t^{(i)}} A e^{-\lambda_k Q |T|} (T-0)^2 = A \sum_{T=0}^{t^{(i)}} T^2 e^{-\lambda_k Q T} \quad (28-4)$$

که در آن

$$A = \frac{(1 - e^{-\lambda_k Q})}{2} e^{\lambda_k Q f}, \quad (29-4)$$

برای یک بلوک 4x4، ۱۶ ضریب با فرکانسهای مختلف وجود دارد و اعوجاج ناهمخوانی از مجموع آنها (بجز فرکانس DC، $k \neq 0$) بدست می آید:

$$\Delta_X^i = E \left[(\delta_X^i)^2 \right] = 2P(1-P) \left(\frac{1}{2} \right) \left(\frac{1}{16} \right) \sum_{k=1}^{15} E \left[(\delta_{01k}^D)^2 \right] \quad (30-4)$$

ضریب (1/2) در رابطه فوق از اینجا ناشی می شود که ضرایب کوچک تقسیم (و نه کلاً صفر) می شوند، یعنی یک ضریب در یک توصیف صفر است ولی در توصیف دیگر مقدار اصلی خود را دارد.

QMDC: در این MDC، برای هر بلوک، دو نسخه وجود دارد: یک نسخه اصطلاحاً نرم^۱ کوانتایز می شود و نسخه دیگر سخت^۲. به عبارت دیگر، در یک نسخه گام کوانتیزیشن کوچکتر از دیگری است. نسخه ای که با QP کوچک کد می شود داده اصلی است و آن یک، داده افزونگی می باشد. این دو نسخه از تمام بلوکها بطور یک در میان در دو توصیف قرار داده می شوند. در حالتیکه هر دو توصیف دریافت شده باشند، نسخه ای که با QP بالاتری کد شده بود، دور ریخته می شود. QP مورد استفاده برای داده های افزونگی (QP_r) به اندازه q_r از QP استفاده شده برای داده های اصلی (QP_p) بزرگتر است؛ q_r به عنوان پارامتر این MDC تعریف می شود:

¹ fine

² coarse

$$\mathbf{P} = \mathbf{q}_r = \{q_r^{(0)}, q_r^{(1)}, \dots, q_r^{(N-1)}\} \quad q_r = QP_r - QP_p \quad (31-4)$$

بنابراین، سیگنال ناهمخوانی مربوط به فرکانس k ام از بلوک DCT به این صورت است:

$$\delta_{01k}^Q = Q_p x_{Q_p} - Q_r x_{Q_r} \quad (32-4)$$

که در آن بالانویس Q بخاطر QMDC اضافه شده است، $Q_p x_{Q_p}$ و $Q_r x_{Q_r}$ به ترتیب داده‌های بازسازی شده اصلی و افزونگی هستند. Q_p و Q_r هم گام‌های کوانتیزیشن به ازای مقادیر QP_p و QP_r می‌باشند [رجوع شود به معادله (1-3)]. در نتیجه

$$\Delta_X^i = E[(\delta_X^i)^2] = 2P(1-P) \left(\frac{1}{2}\right) \left(\frac{1}{16}\right) \sum_{k=0}^{15} E[(\delta_{01k}^Q)^2] \quad (33-4)$$

دوباره ضریب (1/2) بخاطر تقسیم کردن داده‌های نرخ-پایین در توصیفها است. به عبارت دیگر، ناهمخوانی هنگامی ایجاد می‌شود که بلوک‌های نرخ-بالا دچار خطا شوند که این شامل نیمی از خطاها می‌شود.

MLMDC: همانطور که قبلاً بحث شده بود، پارامتر MDC در MLMDC، c است:

$$\mathbf{P} = \mathbf{c} = \{c^{(0)}, c^{(1)}, \dots, c^{(N-1)}\} \quad (34-4)$$

در MLMDC دو نوع ناهمخوانی وجود دارد: ناهمخوانی بین سیگنال تخمین زده شده و سیگنال دیکد شده از دو لایه (پایه و بهبود)، δ_{2L-est} ، و دوم ناهمخوانی بین سیگنال تخمین زده شده و سیگنال دیکد شده از یک لایه (پایه)، δ_{1L-est} . شایان ذکر است که برای دیکدینگ بدون خطای یک لایه‌ای، برای جدا کردن لایه پایه از لایه بهبود، وجود هر دو توصیف لازم است، ولی فقط لایه اول استفاده می‌شود. از آنجا که در MLMDC، تصاویر مرجع فقط از لایه پایه ساخته می‌شوند، δ_{1L-est} معرف ناهمخوانی در مرجع پیش‌بینی و عامل انتشار خطا است. در حالیکه δ_{2L-est} به هنگام محاسبه اعوجاج فریم جاری، خود را نشان می‌دهد. با وارد کردن این دو نوع ناهمخوانی، معادله (4-18) بصورت زیر در می‌آید:

$$D_c^n = \sum_{i=0}^{n-1} (1-\beta)^{n-i} \Delta_{X_{1L-est}}^i + \Delta_{X_{2L-est}}^n \quad (35-4)$$

که در آن، اعوجاج فریم m ($\Delta_{X_{2L-est}}^n$)، که ناشی از خطای داده‌های فریم جاری است و نه خطای منتشر شده از فریمهای قبلی، جدا شده است. در نتیجه، معادله (4-23) را برای MLMDC به شکل زیر می‌توان نوشت:

$$\begin{aligned} D_t^n &= \sum_{i=0}^{n-1} ((1-\beta)^{n-i} \Delta_{X_{1L-est}}^i) + \Delta_{X_{2L-est}}^n + D_Q^n \\ &= \sum_{i=0}^{n-1} \left((1-\beta)^{n-i} E \left[(\delta_{X_{1L-est}}^i)^2 \right] \right) + E \left[(\delta_{X_{2L-est}}^n)^2 \right] + E \left[(\delta_Q^n)^2 \right] \end{aligned} \quad (36-4)$$

دو جمله آخر رابطه (4-36) را می‌توان به شکل زیر نوشت:

$$\begin{aligned} E \left[(\delta_{X_{2L-est}}^n)^2 \right] + E \left[(\delta_Q^n)^2 \right] &= 2P(1-P)E \left[(\delta_{01}^i)^2 \right] + E \left[(\delta_Q^n)^2 \right] \\ &= 2P(1-P)E \left[(\hat{X}_0^n - \hat{X}_1^n)^2 \right] + E \left[(X^n - \hat{X}_0^n)^2 \right] \end{aligned} \quad (37-4)$$

به علت استقلال بین اعوجاج کانال و اعوجاج کوانتیزیشن، رابطه (4-37) را می‌توان به فرم زیر نوشت:

$$\begin{aligned} &= 2P(1-P)E \left[(X^n - \hat{X}_1^n)^2 \right] + (1-2P(1-P))E \left[(X^n - \hat{X}_0^n)^2 \right] \\ &= 2P(1-P)D_{side}^n + (1+2P^2-2P)D_{cen}^n \end{aligned} \quad (38-4)$$

که در آن D_{cen}^n و D_{side}^n اعوجاج‌های جانبی و مرکزی هستند که در روابط (3-17) و (3-4) تعریف شده‌اند. با استفاده از (4-38)، معادله اعوجاج فریم داده شده در (4-36) به این صورت نوشته می‌شود:

$$D_t^n = \sum_{i=0}^{n-1} \left((1-\beta)^{n-i} E \left[(\Delta_{X_{1L-est}}^i)^2 \right] \right) + 2P(1-P)D_{side}^n + (1+2P^2-2P)D_{cen}^n \quad (39-4)$$

مشاهده می‌شود که در انتها فقط δ_{1L-est} ظاهر شد. δ_{1L-est} را میتوان بصورت زیر بدست آورد:

$$\delta_{01k}^M = Qx_Q - Qx_{est} \quad (40-4)$$

که x_{est} در رابطه (۳-۲۴) داده شده است. بنابراین

$$\Delta_X^i = E \left[(\delta_{X_{1L-est}}^i)^2 \right] = 2P(1-P) \left(\frac{1}{2} \right) \left(\frac{1}{16} \right) \sum_{k=0}^{15} E \left[(\delta_{01k}^M)^2 \right]. \quad (۴-۴۱)$$

به این ترتیب روابط (۴-۳۰)، (۴-۳۳) و (۴-۴۱) اعوجاج ناهمخوانی را به ترتیب در DupSp، QMDC و MLMDC بیان می‌کنند. این روابط در بخش بعدی مورد استفاده قرار گرفته، MDC ها بهینه‌سازی شده و نتایج پارامترهای بهینه با نتایج پارامترهای نوعی (غیر بهینه) مقایسه می‌شوند.

۲.۴.۴. نتایج شبیه سازی

ده فریم دوم ویدئوی Mobile با $GOP = 10$ به عنوان داده تست مورد استفاده قرار گرفته و تابع هدف داده شده در رابطه (۴-۲۴) برای هر سه MDC حل شده است. برای $QP_0 = \{20, 24, 28, 32\}$ ، قید تعداد بیت کل در (۴-۲۴) را برابر با $R_t = 2R_{SDC}(QP_0)$ قرار می‌دهیم که $R_{SDC}(QP_0)$ تعداد بیت لازم برای کدینگ تک توصیفه است که در آن QP برابر QP_0 قرار داده شده است. به عبارت دیگر، حد تعداد بیت را آنقدر زیاد می‌گیریم که بتوان پارامترهای MDC را وقتی $QP = QP_0$ است در مقدار کمینه هم قرار داد، یعنی 0 و 1 برای به ترتیب DupSp، QMDC و MLMDC. برای مقادیر بالاتر پارامترهای MDC، افزونگی در MLMDC (DupSp و QMDC) افزایش (کاهش) پیدا می‌کند و بایستی QP را بزرگتر (کوچکتر) از QP_0 انتخاب کنیم. پارامترها می‌توانند در بازه [0 3] در DupSp، [0 8] در QMDC و [1 2] در MLMDC انتخاب شوند. تابع هدف با روش لاگرانژ حل شده و مقادیر بهینه پارامترها برای هر سه MDC در جدول ۴-۱ آورده شده است.

جدول ۴-۱ خروجی بهینه ساز برای GOP دوم از سکانس Mobile

		QP	MDC parameter vector (P)	QP	MDC parameter vector (P)	QP	MDC parameter vector (P)	QP	MDC parameter vector (P)
	DupSp	20	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]	23	[0, 0, 0, 0, 0, 0, 0, 1, 2, 2]	27	[0, 0, 0, 0, 0, 0, 1, 1, 3, 3]	31	[0, 0, 0, 0, 1, 1, 1, 1, 2, 2]
	QMDC	19	[0, 0, 0, 0, 0, 0, 8, 8, 8]	22	[0, 0, 0, 0, 0, 6, 6, 7, 8]	26	[0, 0, 0, 0, 0, 5, 6, 7, 8, 8]	31	[0, 0, 0, 0, 0, 0, 8, 8, 8, 8]
	MLMDC	21	[1, 1, 1, 1, 1, 1, 1.6, 1.8, 2, 2]	25	[1, 1, 1.1, 1.2, 1.3, 1.6, 1.8, 2, 2, 2]	29	[1, 1, 1, 1.5, 1.6, 1.8, 1.9, 2, 2, 2]	34	[1, 1.7, 1.7, 1.8, 1.8, 1.8, 1.9, 1.9, 1.9, 1.9]
	DupSp	20	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]	24	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]	28	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]	32	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]
	QMDC	20	[0, 0, 0, 0, 0, 0, 0, 0, 0, 1]	24	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0]	28	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0]	32	[0, 0, 0, 0, 0, 0, 0, 0, 0, 0]
	MLMDC	21	[1, 1, 1, 1, 1, 1, 1, 1, 1.1, 1.7]	25	[1, 1, 1, 1, 1, 1, 1, 1, 1.4, 1.8]	29	[1, 1, 1, 1, 1, 1, 1, 1, 1.4, 1.7, 1.9]	33	[1, 1, 1, 1, 1, 1, 1, 1, 1.7, 1.9, 2]

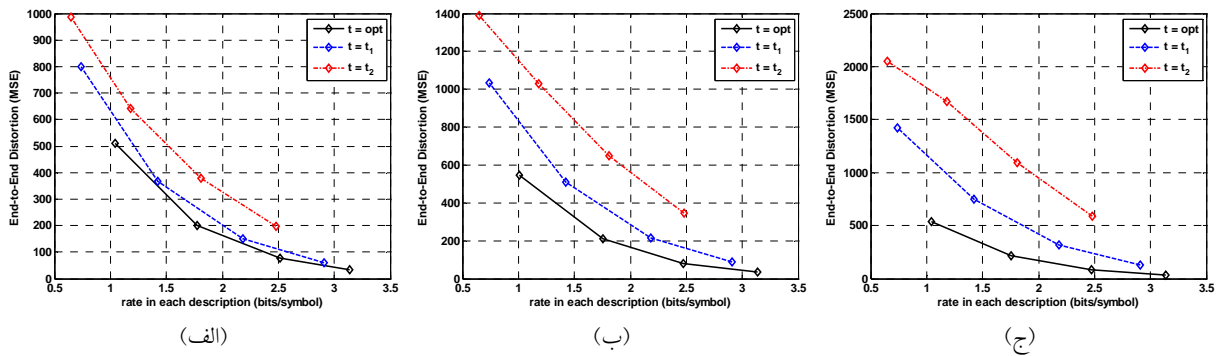
می توان مشاهده کرد که، در حالت کلی، برای فریمهای اول GOP، پارامترهای MDC کم هستند و با حرکت به سمت انتهای GOP افزایش می یابند. در $P = 0.20$ ، مقادیر آنها بیشتر همان مقدار کمینه هستند؛ یعنی اینکه در $P = 0.20$ خروجی بهینه ساز طوری است که انگار یک استریم تک توصیفه را در هر دو کانال کپی کرده ایم بویژه برای DupSp و QMDC.

در شکل ۴-۱، شکل ۴-۲ و شکل ۴-۳ منحنی RD برای هر سه MDC نشان داده شده است. β برابر صفر است و سه مقدار برای PLR در نظر گرفته شده است. تعداد بیت با استفاده از تابع آنتروپی محاسبه شده است. همچنین نتایج حاصله از اعمال مقادیر نوعی پارامترها هم در شکلها نشان داده شده اند. مقادیر نوعی به فرم زیر انتخاب شده اند:

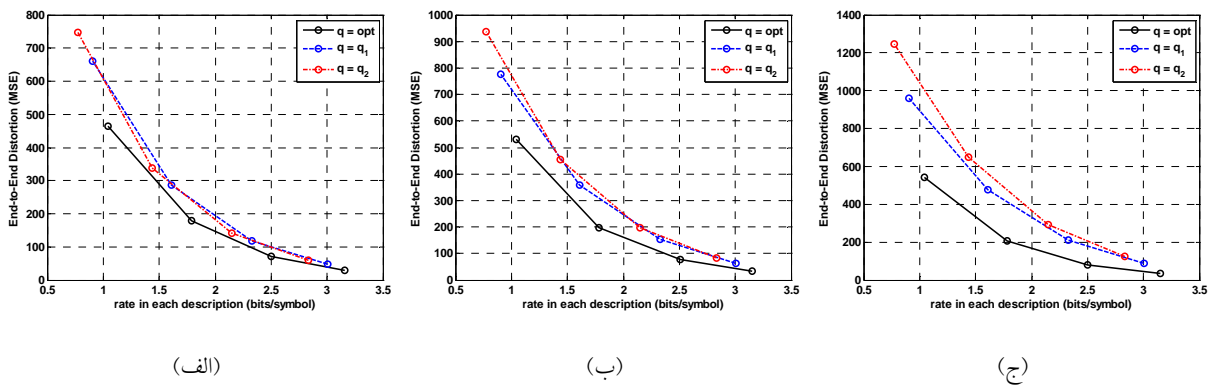
$$P_{DupSP_1} = t_1 = \{1, 1, \dots, 1\} \quad , \quad P_{DupSP_2} = t_2 = \{3, 3, \dots, 3\}$$

$$P_{QMDC_1} = q_{r_1} = \{2, 2, \dots, 2\} \quad , \quad P_{QMDC_2} = q_{r_2} = \{4, 4, \dots, 4\}$$

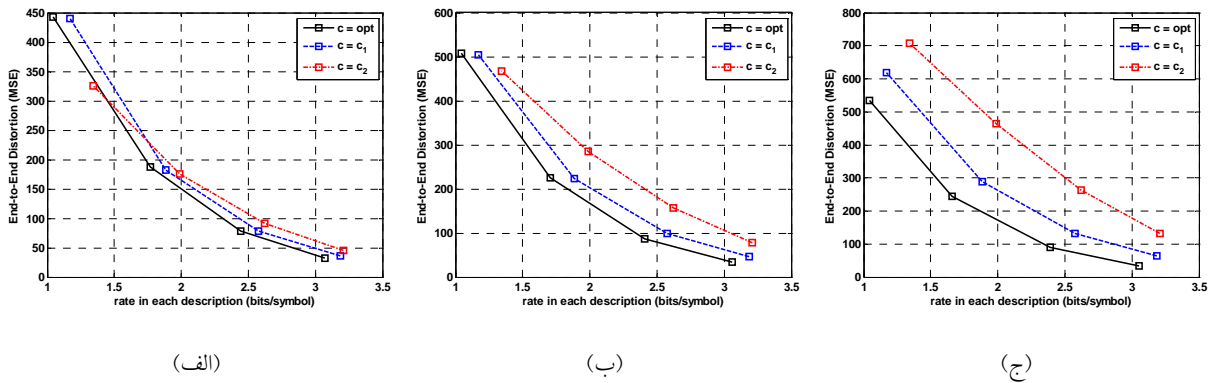
$$P_{MLMDC_1} = c_1 = \{1.3, 1.3, \dots, 1.3\} \quad , \quad P_{MLMDC_2} = c_2 = \{2, 2, \dots, 2\}$$



شکل ۴-۱ منحنی نرخ-اعوجاج در DupSp برای مقادیر بهینه و مقادیر نوعی پارامتر MDC (الف) $P = 0.05$ ، (ب) $P = 0.10$ و (ج) $P = 0.20$



شکل ۴-۲ منحنی نرخ-اعوجاج در DupSp برای مقادیر بهینه و مقادیر نوعی پارامتر QMDC (الف) $P = 0.05$ ، (ب) $P = 0.10$ و (ج) $P = 0.20$



شکل ۴-۳ منحنی نرخ-اعوجاج در DupSp برای مقادیر بهینه و مقادیر نوعی پارامتر MLMDC (الف) $P = 0.05$ ، (ب) $P = 0.10$ و (ج) $P = 0.20$

نتایج نشان می‌دهند که بجز در موارد با نرخ خطای پایین ($P=0.05$)، تاثیر بهینه‌ساز قابل توجه است.

۵.۴. نتایج عملی

به منظور نشان دادن عملکرد مدل ارائه شده برای اعوجاج انتها-تا-انتها و همچنین عملکرد تابع بهینه‌ساز، MLMDC و دو MDC دیگر در JM16.0، نرم افزار مرجع H.264/AVC، پیاده‌سازی شده‌اند. تنظیمات، مشابه آنچه در بخش ۵.۳ برای شبیه‌سازی عملکرد MLMDC در سناریوی ریزش پاکت انجام شد، است ولی ۱۰۰ فریم را شبیه‌سازی کرده‌ایم. در این بخش، دقت مدل اعوجاج انتها-تا-انتهای ارائه شده و سپس خروجی MDC ها بعد از بهینه‌سازی، اندازه‌گیری می‌شوند.

۱.۵.۴. اعوجاج انتها-تا-انتها

در این زیربخش، ویدئوهای مرجع توسط MDC ها کد می‌شوند و روی کانالهای نویزی ارسال می‌شوند، سپس اعوجاج هر فریم اندازه‌گیری و با مقدار پیش‌بینی شده توسط مدل مقایسه می‌شود. پارامترهای MDC بصورت زیر مقداردهی شده است:

$$\{P_1\} = \{t_1, q_{r_1}, c_1\} \text{ and } \{P_2\} = \{t_2, q_{r_2}, c_2\}$$

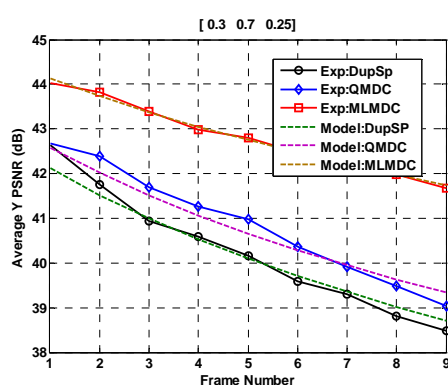
که t و c همان هستند که در زیربخش ۲.۴.۴ تعریف شدند اما به منظور پوشاندن رنج متغیرها که در بهینه‌سازی بکار می‌رود (برای QMDC این رنج بازه [0 8] است)، q را بصورت زیر قرار می‌دهیم:

$$P_{QMDC_1} = q_{r_1} = \{4, 4, \dots, 4\} \text{ , } P_{QMDC_2} = q_{r_2} = \{8, 8, \dots, 8\}$$

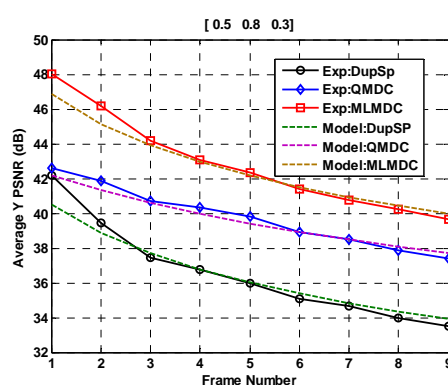
تمام فریمهای GOP مقدار پارامتر MDC یکسانی دارند. در شکل ۴-۴ تا شکل ۴-۱۱ منحنی کیفیت در فریمهای GOP دوم ویدئوهای Mobile و Foreman برای $QP = 20$ و $QP = 28$ ، $PLR = 0.05$ و $PLR = 0.20$ ، $\beta = 0$ و $\beta = 0.1$ ، مقدار اندازه‌گیری شده و آنچه توسط مدل پیش‌بینی شده است، آورده شده است. دقت شود که محور عمودی منحنی‌ها با واحد PSNR بیان شده است، اما برای محاسبه

کیفیت متوسط، متوسط گیری ریاضی از PSNR صحیح نیست. به این علت، مقادیر PSNR ابتدا به MSE^1 تبدیل شده، متوسط گیری انجام می گیرد و سپس دوباره به PSNR برگردانده می شوند. شبیه سازی های انجام شده نشان می دهد که مقدار PSNR پیش بینی شده توسط مدل و مقدار اندازه گیری شده اندکی جابجایی^۲ دارد، منحنی ها با افزودن این جابجایی تصحیح کننده رسم شده است. بردار جابجایی $S_0 = [S_{0DupSp} S_{0QMDC} S_{0MLMDC}]$ بالای شکلها داده شده است.

در تحلیل ارائه شده، تمام مباحث و محاسبات بر پایه لاپلاسین بودن توزیع ضرایب DCT بود که در هر فریم محاسبه می شد. اما همانطور که قبلاً هم گفته شد در ویدئوهای طبیعی یک توزیع مشخص نمی تواند توزیع ضرایب را در تمام بلوکهای یک فریم مدل کند، به عبارت دیگر ضرایب DCT در برخی مختصات تصویر، توزیع بازتر و در برخی دیگر توزیع بسته تر دارند و لذا فرض یک مدل توزیع در هر فریم، ایجاد خطا می کند. این ناشی از همان تقریبی است که در معادله های (۴-۲۱) و (۴-۲۲) اعمال شد. هر MDC، تابع ناهمخوانی ویژه خود را دارد و لذا حساسیت MDC ها نسبت به این تقریب یکسان نمی باشد؛ در نتیجه، جابجایی یکسانی هم ندارند.



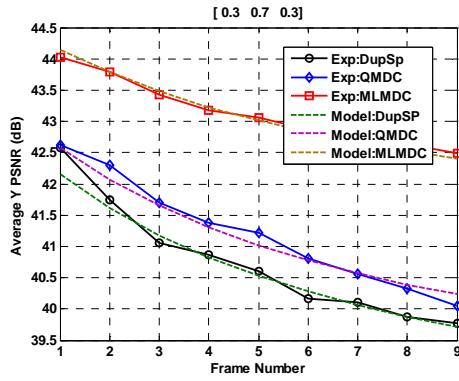
(الف)



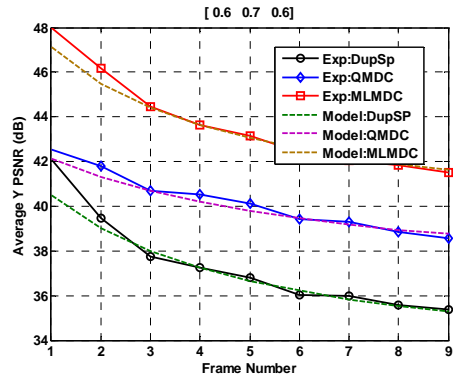
(ب)

¹ Mean Squared Error

² Shift

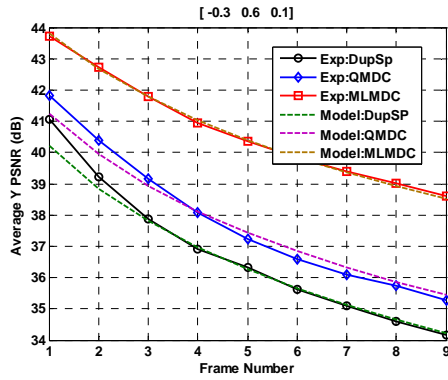


(ج)

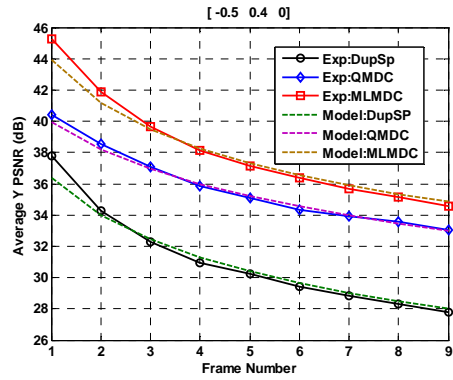


(د)

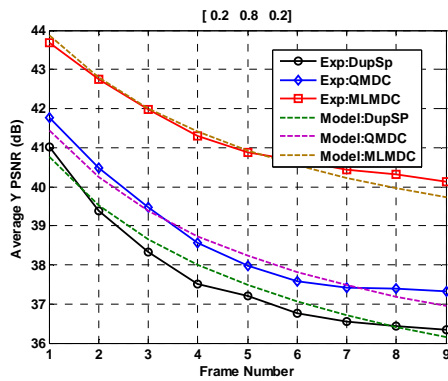
شکل ۴-۴ سکانس Mobile و $QP_0 = 20$ و $PLR = 0.05$ (الف) $\beta = 0$ و $P = \{P_1\}$ (ب) $\beta = 0$ و $P = \{P_2\}$ (ج) $\beta = 0.1$ و $P = \{P_1\}$ (د) $\beta = 0.1$ و $P = \{P_2\}$.



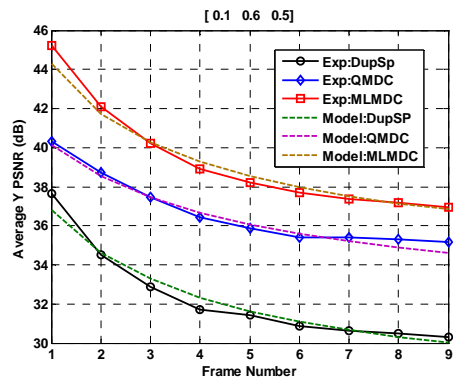
(الف)



(ب)

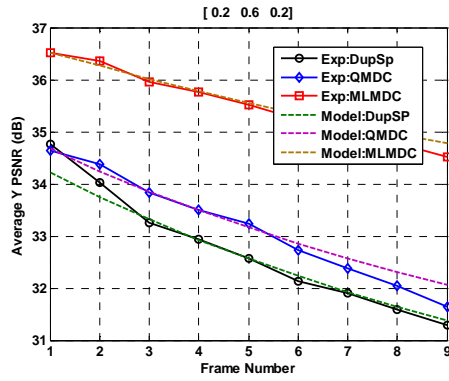


(ج)

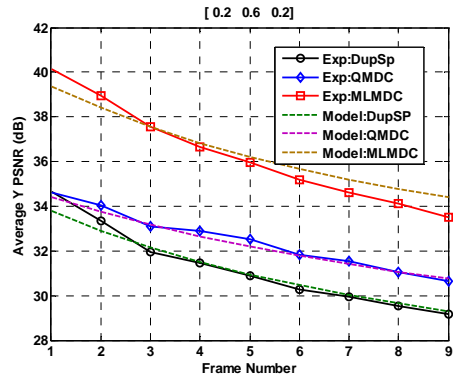


(د)

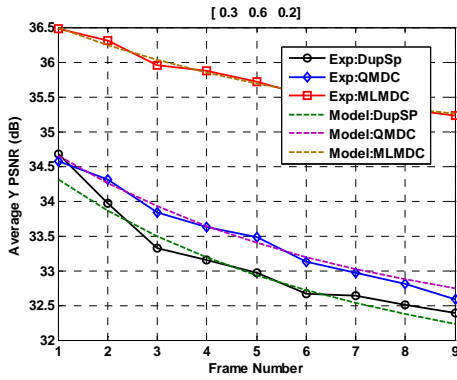
شکل ۴-۵ سکانس Mobile و $QP_0 = 20$ و $PLR = 0.20$ (الف) $\beta = 0$ و $P = \{P_1\}$ (ب) $\beta = 0$ و $P = \{P_2\}$ (ج) $\beta = 0.1$ و $P = \{P_1\}$ (د) $\beta = 0.1$ و $P = \{P_2\}$.



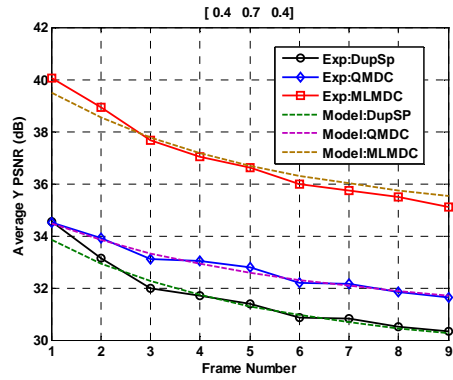
(الف)



(ب)

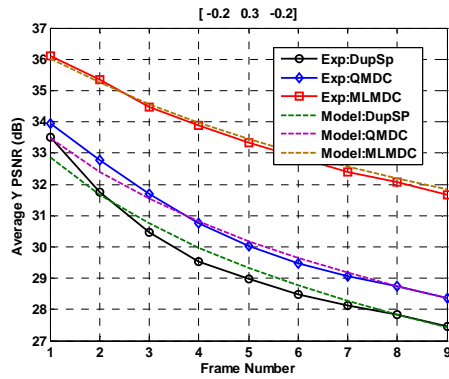


(ج)

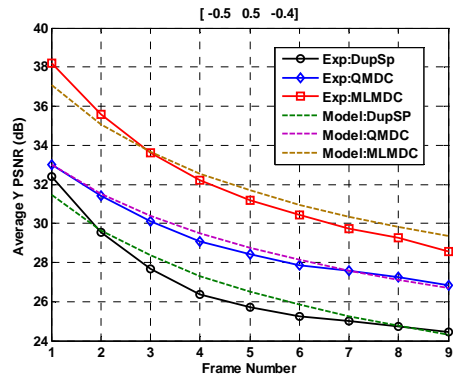


(د)

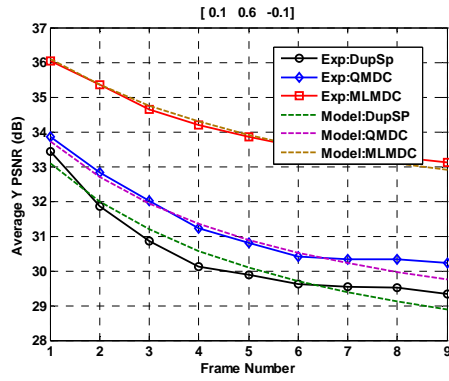
شکل ۴-۶ سکانس Mobile و $QP_0 = 28$ و $PLR = 0.05$ (الف) $\beta = 0$ و $P = \{P_1\}$ (ب) $\beta = 0$ و $P = \{P_2\}$ (ج) $\beta = 0.1$ و $P = \{P_1\}$ (د) $\beta = 0.1$ و $P = \{P_2\}$



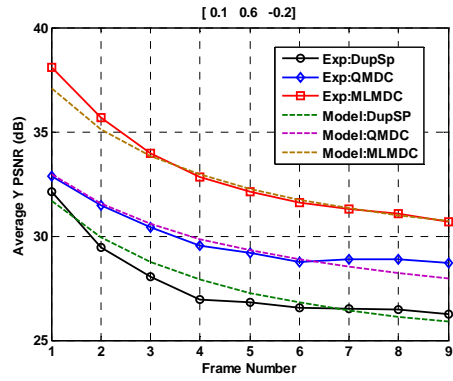
(الف)



(ب)

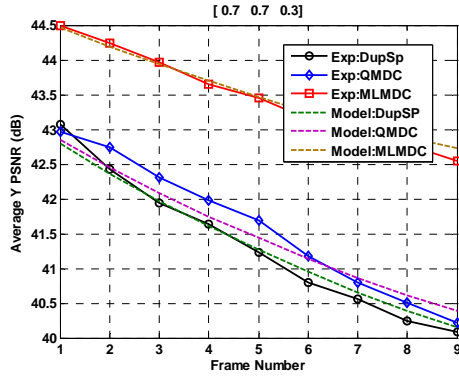


(ج)

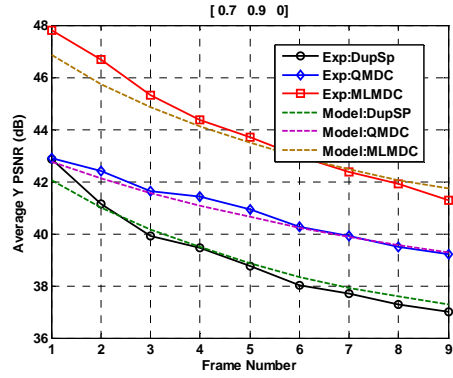


(د)

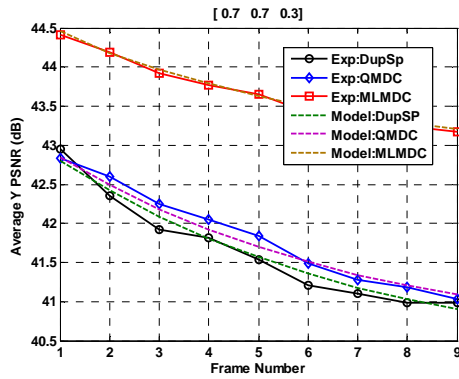
شكل ٤-٧ سكانس Mobile و $QP_0 = 28$ و $PLR = 0.20$ و (الف) $\beta = 0$ و $P = \{P_1\}$ و (ب) $\beta = 0$ و $P = \{P_2\}$ و (ج) $\beta = 0.1$ و $P = \{P_1\}$ و (د) $\beta = 0.1$ و $P = \{P_2\}$.



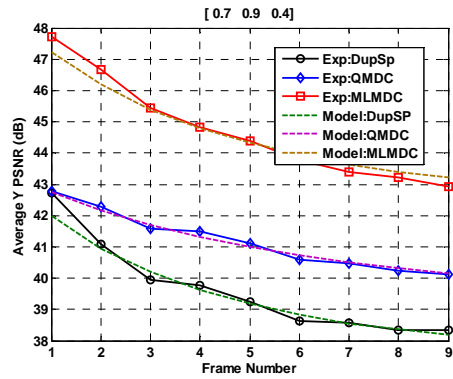
(الف)



(ب)

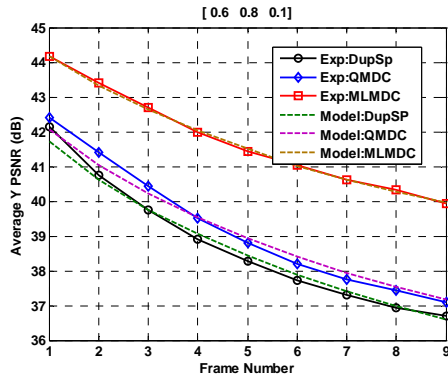


(ج)

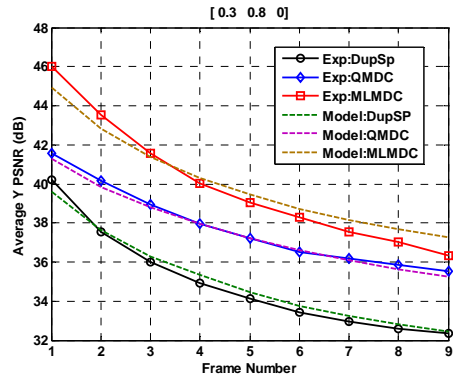


(د)

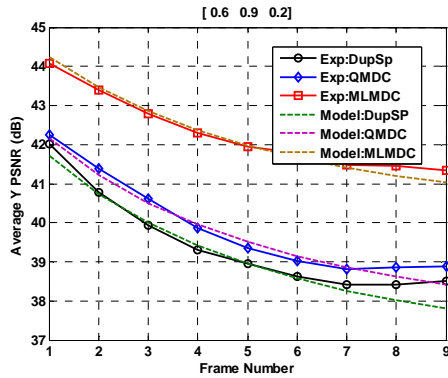
شكل ٤-٨ سكانس Foreman و $QP_0 = 20$ و $PLR = 0.05$ و (الف) $\beta = 0$ و $P = \{P_1\}$ و (ب) $\beta = 0$ و $P = \{P_2\}$ و (ج) $\beta = 0.1$ و $P = \{P_1\}$ و (د) $\beta = 0.1$ و $P = \{P_2\}$.



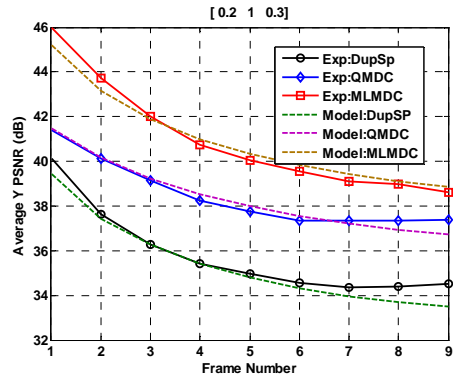
(الف)



(ب)

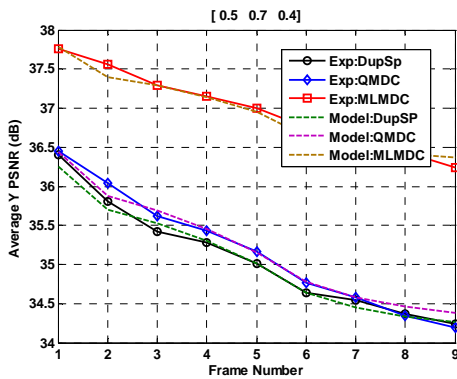


(ج)

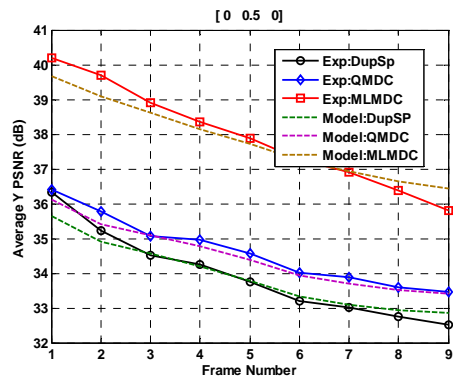


(د)

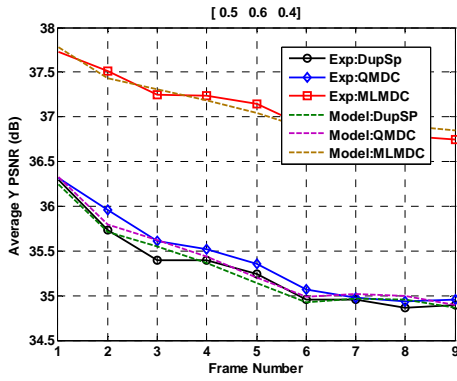
شكل ٤-٩ سكانس Foreman و $QP_0 = 20$ و $PLR = 0.20$ (الف) $\beta = 0$ و $P = \{P_1\}$ (ب) $\beta = 0$ و $P = \{P_2\}$ (ج) $\beta = 0.1$ و $P = \{P_1\}$ (د) $\beta = 0.1$ و $P = \{P_2\}$.



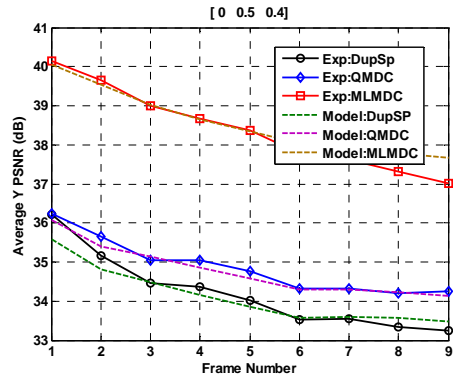
(الف)



(ب)

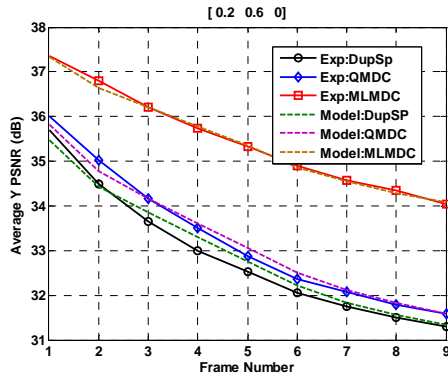


(ج)

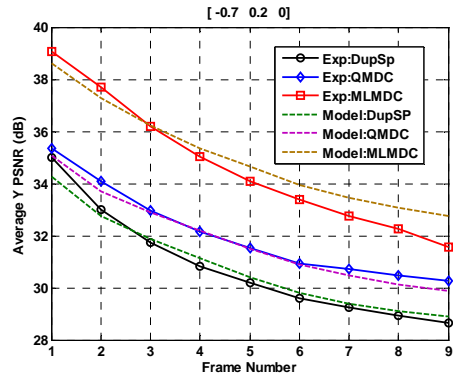


(د)

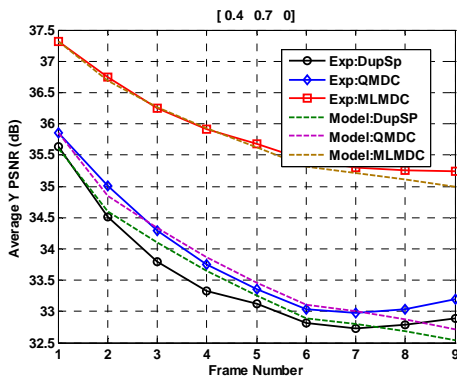
شکل ۴-۱۰ Foreman و $QP_0 = 28$ و $PLR = 0.05$ (الف) $\beta = 0$ و $\mathbf{P} = \{\mathbf{P}_1\}$ (ب) $\beta = 0$ و $\mathbf{P} = \{\mathbf{P}_2\}$ (ج) $\beta = 0.1$ و $\mathbf{P} = \{\mathbf{P}_1\}$ (د) $\beta = 0.1$ و $\mathbf{P} = \{\mathbf{P}_2\}$.



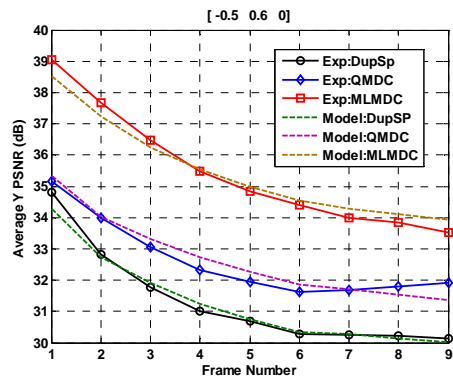
(الف)



(ب)



(ج)



(د)

شکل ۴-۱۱ Foreman و $QP_0 = 28$ و $PLR = 0.20$ (الف) $\beta = 0$ و $\mathbf{P} = \{\mathbf{P}_1\}$ (ب) $\beta = 0$ و $\mathbf{P} = \{\mathbf{P}_2\}$ (ج) $\beta = 0.1$ و $\mathbf{P} = \{\mathbf{P}_1\}$ (د) $\beta = 0.1$ و $\mathbf{P} = \{\mathbf{P}_2\}$.

می توان مشاهده کرد که فقط با یک تنظیم جابجایی، مدل تحلیلی ارائه شده با دقت خوبی اعوجاج واقعی را پیش بینی می کند. تاثیر β در کاهش انتشار خطا و در نتیجه کاهش افت کیفیت کاملاً روشن است، که البته مدل هم بخوبی این را نشان می دهد. در زیربخش بعدی، MDC ها را بر پایه این مدل بهینه کرده و با هم مقایسه می کنیم.

۲.۵.۴. نتایج بهینه سازی

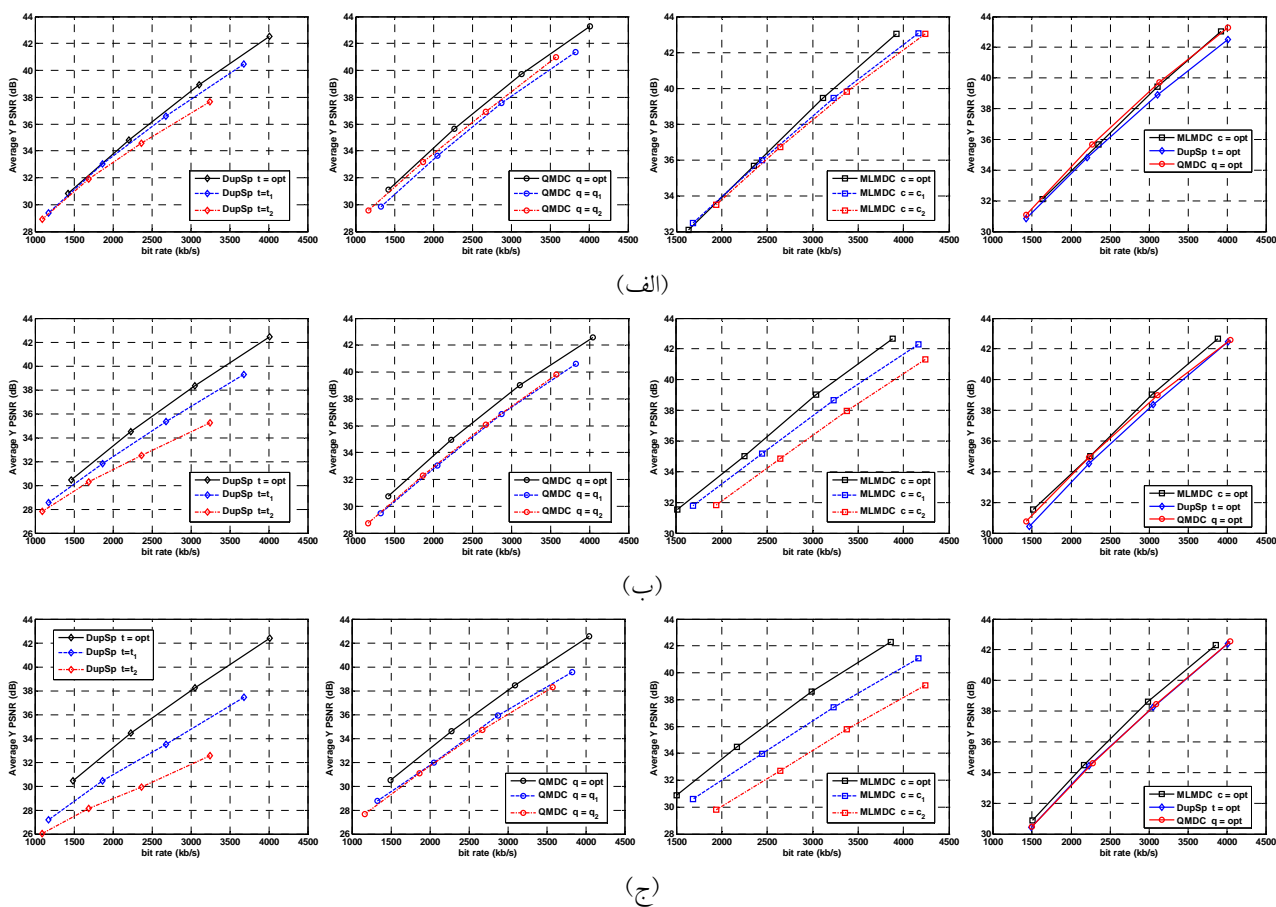
MLMDC و دو MDC دیگر بر پایه مدل اعوجاج داده شده در (۴-۲۴) بهینه می شوند. مقدار بهینه پارامترهای MDC به ازای هر فریم، در PLR و β و با توجه به خواص آماری فریمها (توزیع ضرایب DCT)، محاسبه می شود. مقادیر پارامترهای MDC در بازه [0 3] برای DupSp، [0 8] برای QMDC، [1 2] برای MLMDC تغییر می کنند.

همانطور که در نتایج زیربخش قبل دیده شد، مدل تحلیلی ارائه شده نیاز به تنظیم جابجایی داشت. به عبارت دیگر، مدل شیب را خوب پیش بینی می کرد ولی مقداری جابجایی نیاز داشت. اما توجه شود که تفاوت مقدار این جابجایی برای در یک β ، PLR و QP معین، در محدوده کوچکی تغییر می کند. متوسط قدر مطلق (MAD^1) مقدار این تفاوت برای DupSP، QMDC و MLMDC به ترتیب برابر 0.29 dB، 0.13 dB، 0.14 dB است. از طرف دیگر این میزان اختلاف در جابجایی، ناشی از تغییر متغیر بهینه سازی (پارامتر MDC) از ابتدای غیر صفر بازه تا انتهای آن است؛ برای پیدا کردن اکسترممها چنین رنج تغییری نخواهیم داشت و اگر هم داشته باشیم اعمال جابجایی تصحیح کننده تغییرات زیادی در نقاط بهینه ایجاد نمی کند. این مطلب در ضمیمه این فصل، برای نتایج شکل ۴-۴ (الف) و (ب) که اختلاف جابجایی در DupSp به 0.8 dB می رسد

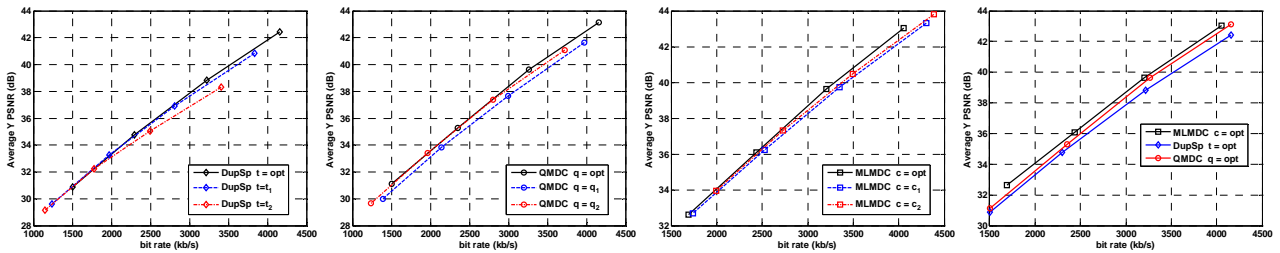
¹ Mean of Absolute Difference

بحث شده است. لذا می توان گفت که S_0 روی دامنه تغییرات متغیرهای بهینه سازی تقریباً ثابت است و تاثیری در بهینه سازی ندارد و نیازی به در نظر گرفتن آنها نیست.

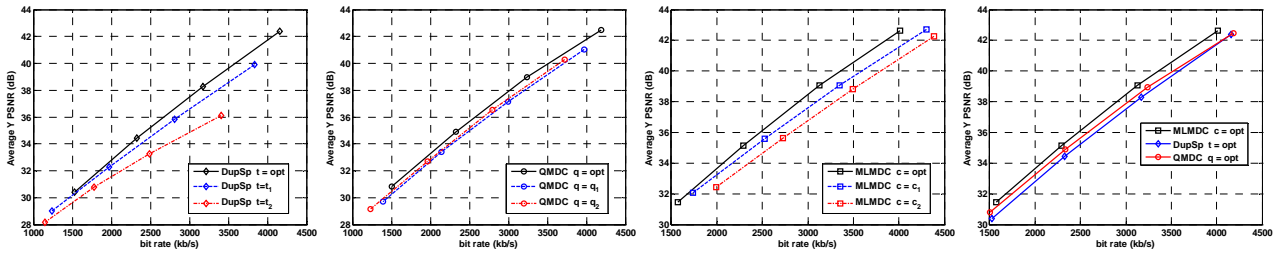
مشابه با تنظیمات بخش ۲.۴.۴، برای $QP_0 = \{20, 24, 28, 32\}$ ، تعداد کل بیت مجاز حساب می شود، برای هر GOP از ویدئو، مقدار بهینه QP و پارامترهای MDC مانند جدول ۴-۱ به ازای هر فریم بدست می آید و در JM قرار داده می شود. فرض می شود هر دو توصیف برای تمام پакتهای اولین فریم هر GOP که I-frame است کامل و بی عیب به گیرنده می رسند. همچنین حالت غیر بهینه هم که پارامترها ثابت هستند و مقادیر ذکر شده در زیربخش ۲.۴.۴ را دارند هم آزمایش شده است. در شکل ۴-۱۲ تا شکل ۴-۱۵ نتایج برای سکانس های Mobile و Foreman و دو مقدار $\beta = 0$ و $\beta = 0.1$ نشان داده شده است.



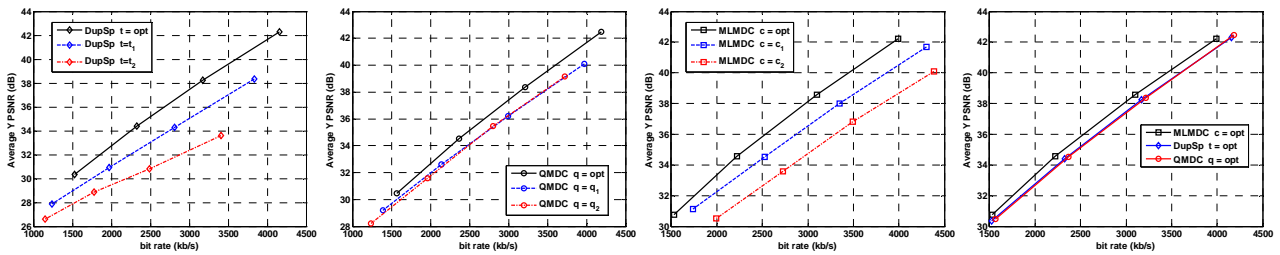
شکل ۴-۱۲ سکانس Mobile و $\beta = 0$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و QMDC و MLMDC مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$



(الف)

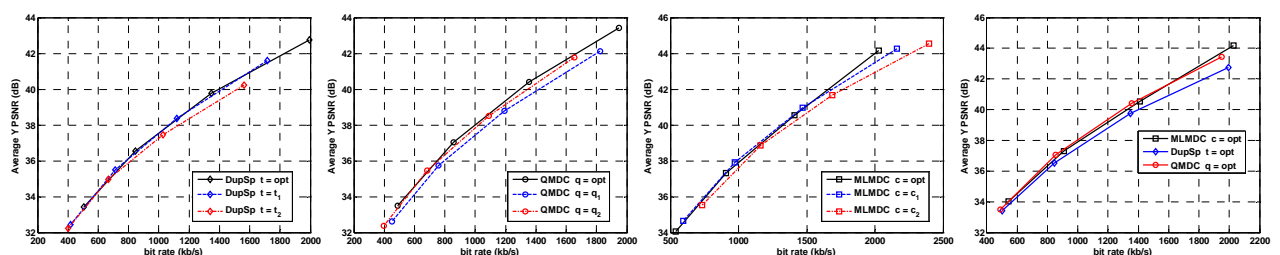


(ب)

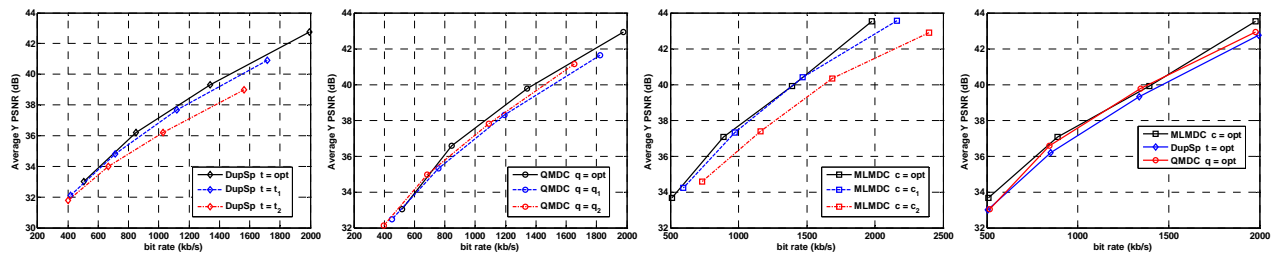


(ج)

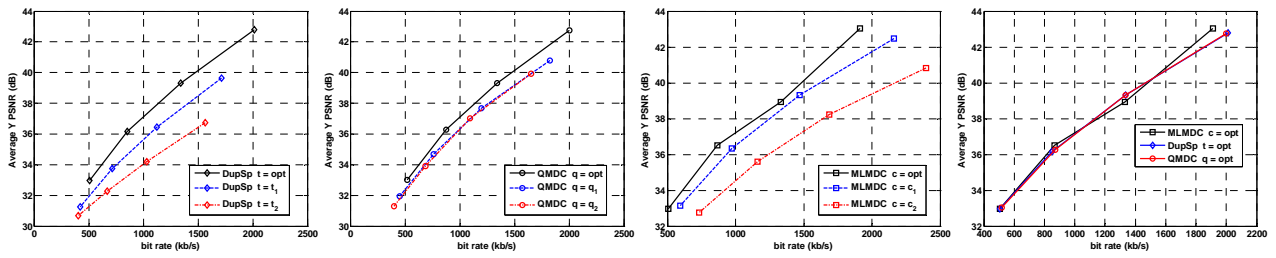
شکل ۴-۱۳ سکانس Mobile و $\beta = 0.1$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و QMDC و MLMDC مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$



(الف)

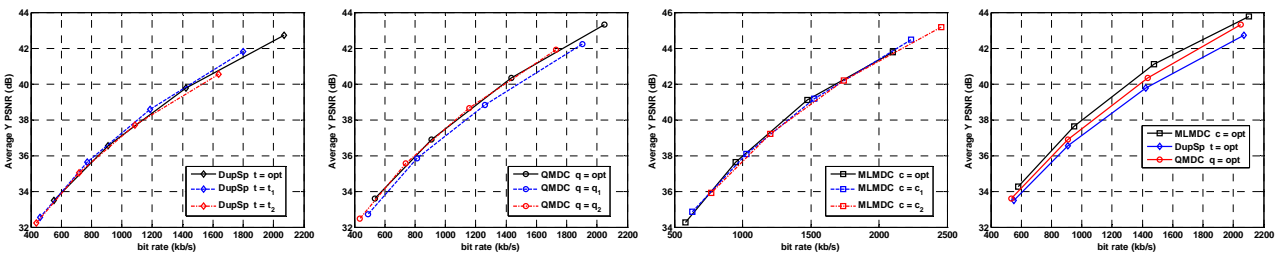


(ب)

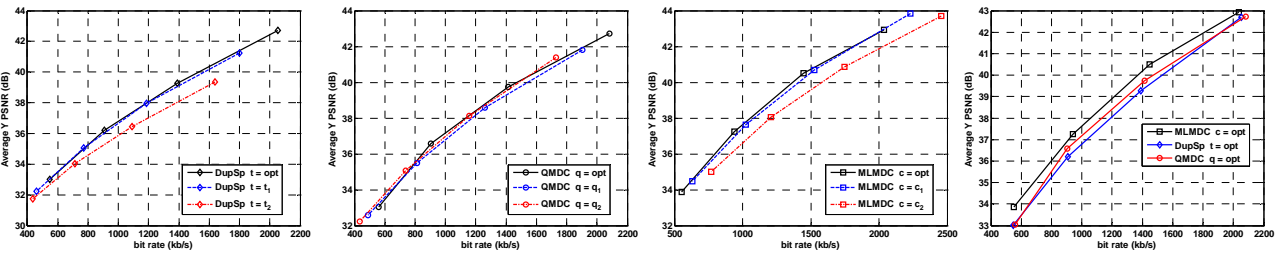


(ج)

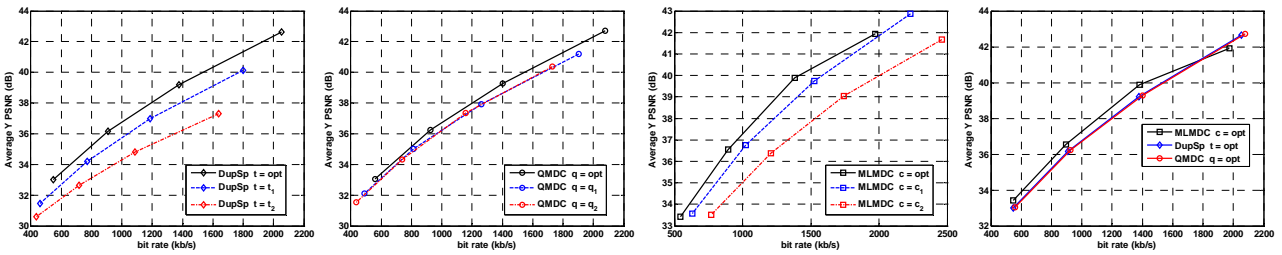
شکل ۴-۱۴ سکانس Foreman و $\beta = 0$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و QMDC و مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$ و MLMDC



(الف)



(ب)



(ج)

شکل ۴-۱۵ سکانس Foreman و $\beta = 0.1$ ، منحنی نرخ-اعوجاج برای مقادیر بهینه و غیر بهینه پارامترهای MDC در سه روش DupSp و QMDC و مقایسه نتایج (الف) $P = 0.05$ ، (ب) $P = 0.10$ ، (ج) $P = 0.20$ و MLMDC

شکل‌های بالا نشان می‌دهد که PSNR در حالتی که مقادیر بهینه پارامترها استفاده شود بالاتر است. همچنین عملکرد MLMDC در بیشتر اوقات از سایرین بهتر است. می‌توان مشاهده کرد که پارامترهای بهینه تاثیر β را در کیفیت نهایی حاصله از همه MDCها کاهش می‌دهد. قبلاً دیده بودیم که β تاثیر بسزایی در کاهش شیب افت کیفیت دارد ولی در حالت بهینه‌سازی و همچنین لحاظ کردن تعداد بیت مازاد برای ایترا کد کردن، باعث کم شدن اثر β در منحنی‌های RD می‌شود.

۵. جمع بندی و کارهای

آینده

در این فصل نوآوریهای رساله و سپس پیشنهادهایی برای ادامه این کار تحقیقاتی ارائه شده است.

۱.۵. نوآوریها

۱.۱.۵. در فصل دوم

در این بخش، بر پایه حوزه ای که عمل MDC انجام می شود، طرحهای MDC گروه بندی شده است. در هر گروه، روشهایی که برای بهبود کیفیت ارائه شده بوده معرفی و توضیح داده شده است. در انتها MDC ها را با توجه به کارایی آنها در مواردی چون تطبیق پذیری افزونگی، پیچیدگی و قابلیت افزایش تعداد توصیفها، با یکدیگر مقایسه کرده ایم که در جدول ۲-۱ آورده شده است. مقاله [96] از این فصل استخراج شده است.

۲.۱.۵. در فصل سوم

یک روش جدید MDC ارائه شد. در این روش هر توصیفی شامل ضرایب پایه و ضرایب ترکیب است، ضرایب ترکیب از جمع کردن ضرایب پایه و بهبود متناظر تشکیل می شوند. در دیکودر مرکزی، ضرایب ترکیب و بهبود از یکدیگر جدا شده، پس از dequantize کردن، مانند CGS دیکد می شوند. در دیکودر جانبی، ضرایب پایه از ضرایب ترکیب تخمین زده می شوند. دو حالت در نظر گرفتیم، ضرایب پایه و بهبود، از یک ضریب DCT باشند (Case I) و یا اینکه از دو ضریب مختلف (Case II). دیده شد که هر دو حالت

اعوجاج تخمین و افزونگی حدوداً یکسانی داشتند ولی Case I بخاطر همبستگی بین ضرایب ترکیب شونده، اعوجاج جانبی بهتری داشت. MLMDC برای ضرایب با واریانس بیشتر عملکرد بهتری داشت، که در مقایسه با سایر MDC های آزمایش شده، هم کیفیت جانبی و هم کیفیت مرکزی خوبی داشت. به عبارت دیگر برای ویدئوهایی که محتوای پایینی (بافت و حرکت) ندارند MLMDC مصالحه بهتری بین کیفیتهای جانبی و مرکزی فراهم می‌کرد؛ در نتیجه این مصالحه بهتر، کیفیت متوسط بالاتری به هنگام ارسال ویدئو در محیطهای نویزی فراهم می‌شود. عملکرد MLMDC در کدینگ چهار توصیفه، چشم‌گیرتر است. مقاله‌های [94] و [95] از این فصل استخراج شده است.

۳.۱.۵. در فصل چهارم

در این فصل MLMDC با توجه به میزان خطای کانال بهینه شده است. به این منظور ابتدا مدلی برای پیدا کردن اعوجاج انتها-تا-انتها ارائه شده است که نه تنها برای MLMDC، بلکه برای سایر MDC های حوزه فرکانس قابل استفاده است. مدل ارائه شده روی سه MDC (MLMDC، DupSp و QMDC) پیاده شده و نتایج مدل و اندازه‌گیری بیشتر مواقع بخوبی منطبق هستند، اگرچه یک کالیبراسیون جابجایی نیاز دارد. اما چنانچه بحث شد، این جابجایی تأثیری در نتایج بهینه سازی ندارد و لذا مدل را می‌توان بدون کالیبراسیون برای طراحی بهینه MDC ها بکار برد. نتایج، تأثیر تابع بهینه ساز را نشان می‌دهد. همچنین MLMDC نسبت به سایر MDC های تست شده، هنگامیکه همه بهینه شده‌اند، کیفیت متوسط بالاتری ارائه می‌دهد. مقاله [97] از این فصل استخراج شده است.

۲.۵. کارهای آینده

۱.۲.۵. پیشنهاد اول

معیاری که ما برای مقایسه MDC ها بکار بردیم و در مقالات دیگر هم بکار برده می‌شود، شبیه سازی یک انتقال ویدئو از کانالهای دارای ریزش پاکت و سپس اندازه‌گیری کیفیت ویدئو در دیکودر است. در این سناریو، برای اینکه خروجی مستقل از الگوی ریزش پاکت و ویدئوی خاص باشد، نیاز است که تعداد متنابهی شبیه سازی با الگوهای Loss مختلف و همچنین تعداد قابل قبولی فریم از ویدئوهای استاندارد انجام شود. به عبارت دیگر، معیاری وجود ندارد و فقط باید با نتایج عملی آنها را مقایسه کرد. لذا جای خالی چنین معیاری برای مقایسه MDC های ویدئو احساس می‌شود. ناگفته نماند، چنانچه در بخش ۴.۳ هم گفته شد، MDC های تصویر را می‌توان با معیار افزونگی-نرخ-عوجاج بخوبی مورد سنجش قرار داد؛ ولی در MDC های ویدئو به علت وجود انتشار خطا، این معیار قابل استفاده نیست و یا حداقل پاسخ صحیح به ما نمی‌دهد. بنابراین پروژه پیشنهادی یافتن معیاری برای مقایسه MDC های ویدئو است که انتشار خطا را هم در نظر گرفته است.

۲.۲.۵. پیشنهاد دوم

در این رساله، QP را چند مقدار خاص قرار داده و سپس کیفیت و نرخ بیت را اندازه گرفتیم؛ یعنی اینکه کنترل نرخ نداشتیم. در ادامه رساله می‌توان کنترل نرخ را هم اضافه کرد. در این رساله، نرخ بیت را با تابع آنروپی حساب کردیم که برای این کار (کنترل نرخ) به دلیل حجم محاسبات بالا و همچنین تفاوتی که با انکودرهای CAVLC و CABAC دارد، توصیه نمی‌شود. بنابراین ابتدا باید برای محاسبه افزایش نرخ بیت ناشی از ترکیب، مدلی پیدا شود. مدلهای متعددی در مقالات وجود دارند که نرخ بیت را با تابعی از QP و یا

اعوجاج بیان می‌کنند، ولی برای مدل کردن تعداد بیت لایه های ترکیبی مناسب نیستند و لذا به مدل جدیدی نیاز است. با بدست آوردن این مدل، کنترل نرخ را هم می‌توان در ادامه انجام داد.

۳.۲.۵. پیشنهاد سوم

شبه سازیه‌ها و همچنین مدل ارائه شده برای اعوجاج انتها-تا-انتها، همه بر اساس مدل برنولی کانال بود. اما در عمل کانالها دارای خطای burst هستند. گرچه با درهم‌تاباندن می‌توان اثر خطای burst را کاهش داد و همچنین در MDC بخاطر MPT این مساله کم رنگ‌تر است، ولی وارد کردن اثر خطای burst در فرمولهای بخش چهارم، مدل ارائه شده را جامع‌تر می‌نماید که در ادامه این رساله پیشنهاد می‌شود.

۴.۲.۵. پیشنهاد چهارم

همانطور که دیدیم به علت نایکنواخت بودن توزیع ضرایب در بلوکهای مختلف یک فریم، مدل ارائه شده برای اعوجاج انتها-تا-انتها مقداری خطا داشت. با دسته‌بندی بلوکها به گونه‌ایکه هر دسته λ مخصوص به خود داشته باشد، می‌توان این مشکل را حل کرد. البته این دسته‌بندی، پیچیدگی انکودر و اضافه اطلاعات سرآمد به دنبال خواهد داشت؛ چونکه، برای هر دسته بلوک، پارامتر MDC بایستی جداگانه تعیین، اعمال و ارسال شود.

۵.۲.۵. پیشنهاد پنجم: تخمین حرکت چند توصیفه

در تخمین حرکت یک توصیفه، با رویکرد بهینه سازی نرخ-اعوجاج، تخمین حرکت با کمینه سازی تابع J تعریف شده بصورت زیر، انجام می‌شود:

$$J = D + \lambda_{SDC} R_{mv} \quad (۱-۵)$$

که در آن R_{mv} تعداد بیت‌های مورد نیاز برای ارسال بردار حرکت است، و D اغلب بصورت مجموع مقدار مطلق اختلاف (SAD^1) بین بلوک جاری و بلوک مرجع تعریف می‌شود، یعنی

$$D = SAD[M_{curr} - M_{ref}] \quad (2-5)$$

پژوهش پیشنهادی تصحیح این رویکرد برای انکودرهای چند توصیفه است. مثلاً می‌توان رابطه (5-1) را به فرم زیر نوشت:

$$J = D_{MDC} + \lambda_{MDC} 2R_{mv} \quad (3-5)$$

بردارهای حرکت در هر دو توصیف تکرار می‌شوند و لذا تعداد بیت آنها دو برابر حالت SDC است. D_{MDC} را هم می‌توان بطور آماری بر اساس احتمال دیکودینگ مرکزی و دیکودینگ جانبی، در یک حالت ساده، به فرم زیر نوشت:

$$D_{MDC} = (1 - P)^2 SAD_{cen} + 2P(1 - P)SAD_{side} \quad (4-5)$$

که در آن SAD_{side} با توجه به الگوریتم MDC استفاده شده بدست می‌آید. به این ترتیب، افزونگی ناشی از تکرار بردار حرکت و همچنین اثر دیکودر جانبی در اعوجاج را هم در تخمین حرکت وارد می‌کنیم. براحتی می‌توان حدس زد، بلوکهای انتخاب شده به عنوان بلوک مرجع توسط رابطه (5-3) در چه مواقعی با بلوکهای انتخاب شده با (5-1) تفاوت دارند. به علاوه، λ_{MDC} هم ضرورتاً با λ_{SDC} یکسان نیست و البته می‌توان آن را هم با روندی مشابه با روند پیدا کردن λ_{SDC} بدست آورد. معیار بهتر آنست که اثر انتشار خطا در اعوجاج را هم در نظر گرفته و سپس تابع هدف را برای تعیین بهینه‌ترین بردار حرکت تشکیل دهیم. به

¹ Sum of Absolute Difference

این ترتیب تخمین حرکت با هدف داشتن بیشترین کیفیت متوسط در دیکودر، انجام می‌شود؛ اگرچه روشن است که با این کار حجم محاسبات بیشتری نیز نیاز است.

۶. ضمائ

۱.۶. ضمیمه فصل سوم

برای یک منبع لاپلاسیون و کوانتیزاسیون رابطه (۱-۱)، برای هر سطح کوانتیزیشن N توزیع نوین کوانتیزیشن

بصورت زیر است:

$$N > 0 \Rightarrow -fQ \leq x_e < (1-f)Q \quad \& \quad f_N(x_e) = \frac{\lambda}{2} e^{-\lambda Q N} e^{-\lambda x_e}$$

$$N < 0 \Rightarrow -(1-f)Q < x_e \leq fQ \quad \& \quad f_N(x_e) = \frac{\lambda}{2} e^{\lambda Q N} e^{\lambda x_e} \quad (۱-۶)$$

$$N = 0 \Rightarrow -(1-f)Q < x_e < (1-f)Q \quad \& \quad f_0(x_e) = \frac{\lambda}{2} e^{-\lambda |x_e|}$$

حال برای بدست آوردن $f^{II}(x_e)$ ، توابع خطا در تمام نواحی کوانتیزاسیون با یکدیگر جمع می‌شوند

$$fQ \leq x_e < (1-f)Q \Rightarrow$$

$$f^{II}(x_e) = \sum_{N=0}^{\infty} f_N(x_e) = \frac{\lambda}{2} e^{-\lambda x_e} \sum_{N=1}^{\infty} e^{-\lambda Q N} + \frac{\lambda}{2} e^{-\lambda x_e} = \frac{\lambda}{2} e^{-\lambda x_e} \left(\frac{e^{\lambda Q}}{e^{\lambda Q} - 1} \right) \quad (۲-۶)$$

$$0 < x_e < fQ \Rightarrow$$

$$\begin{aligned} f^{II}(x_e) &= \sum_{N=-\infty}^{\infty} f_N(x_e) = \frac{\lambda}{2} e^{-\lambda x_e} \sum_{N=1}^{\infty} e^{-\lambda Q N} + \frac{\lambda}{2} e^{-\lambda x_e} \\ &+ \frac{\lambda}{2} e^{\lambda x_e} \sum_{N=-\infty}^{-1} e^{\lambda Q N} = \frac{\lambda}{2} e^{-\lambda x_e} \left(\frac{e^{\lambda Q}}{e^{\lambda Q} - 1} \right) + \frac{\lambda}{2} e^{\lambda x_e} \left(\frac{1}{e^{\lambda Q} - 1} \right) \quad (۳-۶) \\ &= \frac{\lambda}{2} e^{-\lambda x_e} + \lambda \cosh(\lambda x_e) \left(\frac{1}{e^{\lambda Q} - 1} \right) \cong \frac{\lambda}{2} e^{-\lambda x_e} + \frac{\lambda}{e^{\lambda Q} - 1} \end{aligned}$$

$f^{II}(x_e)$ حول صفر متقارن است و بنابراین داریم

$$f_{x_e}^{II}(x_e) = \begin{cases} \frac{\lambda}{2} e^{-\lambda|x_e|} + \frac{\lambda}{e^{\lambda Q} - 1} & |x_e| \leq fQ \\ \frac{\lambda}{2} \frac{e^{\lambda Q}}{e^{\lambda Q} - 1} e^{-\lambda|x_e|} & fQ < |x_e| \leq (1-f)Q \end{cases} \quad (۴-۶)$$

۲.۶. ضمیمه فصل چهارم

در این ضمیمه نشان می‌دهیم که برای شکل ۴-۱ (الف) و (ب) و DupSp، گرچه اختلاف جابجایی آنها بیشترین مقدار، 0.8dB است اما اعوجاج ناهمخوانی نسبی فریمها را چندان تغییر نمی‌دهد و بنابراین جابجایی تصحیح کننده تاثیری در مقدار پارامترهای بهینه ندارد. جابجایی در PSNR معادل ضرب در MSE است، بنابراین رابطه اعوجاج پس از تنظیم جابجایی به صورت زیر در می‌آید:

$$D_t^n = \alpha_{t_0} \left(\sum_{i=0}^n ((1-\beta)^{n-i} \Delta_X^i) + E[(\delta_Q^n)^2] \right) \quad (۵-۶)$$

که در آن α_{t_0} ضریب تصحیح مرتبط با $t = t_0$ (پارامتر MDC) است. مقدار جابجایی وقتی $t_0 = 3$ است در این شکل به 0.8 dB بیشتر از هنگامیکه $t_0 = 1$ است، می‌رسد. براحتی می‌توان نشان داد که به ازای 0.8dB تفاوت در جابجایی، داریم:

$$\frac{\alpha_3}{\alpha_1} = 10^{-0.08} \quad (۶-۶)$$

برای فریم آخر GOP که اعوجاج عمدتاً ناشی از ناهمخوانی است، می‌توان از اعوجاج کوانتیزیشن در مقابل آن صرف‌نظر نمود. از سوی دیگر اختلاف بین کیفیت فریمها وقتی $t_0 = 3$ و $t_0 = 1$ است به 4.5 dB می‌رسد. یعنی اینکه:

$$\frac{\alpha_1(\sum_{i=0}^n((1-\beta)^{n-i}\Delta_X^i|_{t=1}))}{\alpha_3(\sum_{i=0}^n((1-\beta)^{n-i}\Delta_X^i|_{t=3}))} = 10^{-0.45} = 0.3548 \quad (7-6)$$

چونکه، حرکت و بافت فریمها در ویدئو خیلی سریع عوض نمی‌شوند، می‌توان گفت که توزیع ضرایب

DCT هم در فریمهای متوالی چندان متفاوت نیست و با این تقریب داریم:

$$\Delta_X^1|_{t=1} \approx \Delta_X^2|_{t=1} \approx \dots \approx \Delta_X^{N-1}|_{t=1} = \Delta_X|_{t=1}$$

$$\Delta_X^1|_{t=3} \approx \Delta_X^2|_{t=3} \approx \dots \approx \Delta_X^{N-1}|_{t=3} = \Delta_X|_{t=3}$$

بنابراین، از روابط (6-6) و (7-6) داریم:

$$\frac{\Delta_X|_{t=1}}{\Delta_X|_{t=3}} = 10^{-0.53} = 0.2951 \quad (8-6)$$

این رابطه به این معناست که در حالت تصحیح شده نسبت اعوجاجها از 0.3 به 0.35 تغییر می‌کند، یعنی کمی نزدیکتر به 1. اما این نسبت هنوز به اندازه کافی کوچک است و همانطور که شبیه سازیهای ما نشان داد، مقدار بهینه در هر دو حالت تصحیح شده و تصحیح نشده یکسان است.

٧. مراجع

- [1] V. K. Goyal, "Multiple description coding: Compression meets the network," *IEEE Signal Processing Mag.*, vol. 18, pp. 74–93, Sep. 2001.
- [2] Y. Wang, A. Reibman, and S. Lin, "Multiple description coding for video delivery," *Proc. IEEE*, vol. 93, pp. 57–70, 2005.
- [3] M. Ghanbari, "Two-layer coding of video signals for VBR networks," *IEEE J. Select. Areas Commun.*, vol. 7, pp. 801–806, June 1989.
- [4] V. N. Padmanabhan , H. Wang and P. Chou "Resilient peer-to-peer streaming", *Proc. IEEE ICNP 2003*.
- [5] J.-C. Wu , K.-J. Peng , M.-T. Lu , C.-K. Lin , Y.-H. Cheng , P. Huang , J. Yao and H. H. Chen "HotStreaming: Enabling scalable and quality IPTV services", *Proc. IPTV Workshop Conjunction 15th Int. World Wide Web Conf.*, pp. 2006.
- [6] M. Castro, P. Druschel, A.-M. Kermarrec, A. Nandi, A. Rowstron and A. Singh, "SplitStream: High-bandwidth multicast in cooperative environments", in *Proc. ACM SOSP'03*, New York, USA, October 2003.
- [7] V. N. Padmanabhan, H. J. Wang, P. A. Chou, and K. Sripanidkulchai, "Distributing streaming media content using cooperative networking", In *ACM/IEEE NOSSDAV*, pages 177–186, Miami, FL, USA, May 12-14 2002.
- [8] C. Lee, J. Kim, Y. Altunbasak, and R. M. Mersereau, "Layered coded vs. multiple description coded video over error-prone networks," *Signal Process.: Image Commun.*, May 2003

- [9] J. Chakareski, S. Han, and B. Girod, "Layered Coding vs. multiple descriptions for video streaming over multiple paths," *Multimedia Systems*, Springer, January 2005.
- [10] A. Reibman, Y. Wang, X. Qiu, Z. Jiang, and K. Chawla, "Transmission of multiple description and layered Video over an (EGPRS) wireless network", *IEEE Inter. Conf. Image Processing*, September 2000.
- [11] R. Singh, A. Ortega, L. Perret, and W. Jiang "Comparison of multiple description coding and layered coding based on network simulations," in *Proc. SPIE Conf. Visual Comm. Image Processing*, 2000.
- [12] J. Wolf, A. Wyner, and J. Ziv, "Source coding for multiple descriptions," *Bell Syst. Tech. J.*, vol. 59, no. 8, pp. 1417–1426, Oct. 1980.
- [13] H. S. Witsenhausen and A. D. Wyner, "Source coding for multiple descriptions II: A binary source," *Bell Labs., Tech. Rep. TM-80-1217*, Dec. 1980.
- [14] L. Ozarow, "On a source coding problem with two channels and three receivers," *Bell Syst. Tech. J.*, vol. 59, pp. 1909– 1921, Dec. 1980.
- [15] B. Girod, and N. Färber, "Feedback-based error control for mobile video transmission," *Proceedings of IEEE*, vol. 97, no. 10, pp. 1707-1723, Oct. 1999.
- [16] T. Wiegand, G. J. Sullivan, G. Bjøntegaard, and A. Luthra, "Overview of the H.264/AVC video coding standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 13, no. 7, pp. 560–576, Jul. 2003.
- [17] V. Goyal, "Theoretical foundations of transform coding," *IEEE Signal Process. Mag.*, vol. 18, no. 5, pp. 9–21, Sep. 2001.
- [18] S. Shirani, M. Gallant, and F. Kossentini, "Multiple description image coding using pre- and post-processing," in *Proc. Int. TCC*, Las Vegas, NV, Apr. 2001, pp. 35–39.
- [19] M. Gallant, S. Shirani, F. Kossentini, "Standard-compliant multiple description video coding," *IEEE International Conference on Image Processing*, Thessaloniki, Greece, October 2001.
- [20] T. Tillo, G. Olmo, "Data-dependent pre- and postprocessing multiple description of coding images", *IEEE Trans. on Image Proc.*, vol. 16, no. 5, pp. 1269-1280, May 2007

- [21] Y. Yapıcı, B. Demir, S. Ertürk, and O. Urhan, "Down-sampling based multiple description image coding using optimal filtering," *Journal of Electronic Imaging*, vol. 17, no. 3, Article No: 033018, Jul-Sep. 2008
- [22] C. Ates, Y. Urgan, B. Demir, O. Urhan, and S. Erturk, "Polyphase downsampling based multiple description image coding using optimal filtering with flexible redundancy insertion", *In International Conference on Signals and Electronic Systems (ICSSES)*, pages 193–196, Sept. 2008.
- [23] J. Wang, J. Liang, "H.264 intra frame coding and JPEG 2000-based predictive multiple description image coding," *Communications, Computers and Signal Processing, 2007. PacRim 2007. IEEE Pacific Rim Conference on*, Aug. 2007, pp 569-572.
- [24] Z. Wei, K. K. Ma, and C. Cai, "Prediction-compensated polyphase multiple description image coding with adaptive redundancy control," *IEEE Trans. Circuits Syst. Video Technol.*, accepted to published, DOI: 10.1109/TCSVT.2011.2168131.
- [25] M. Fumagalli, R. Lancini, and A. Stanzione, "Video transmission over IP by using polyphase downsampling multiple description coding." *In Proc. IEEE Int. Conf. on Multimedia and Expo*, August 2001. pp. 1095- 1098.
- [26] S. Shirani, "Content-based multiple description image coding," *IEEE Trans. Multimedia*, vol. 8, no. 2, pp. 411–419, Apr. 2006.
- [27] W. Jiang and A. Ortega, "Multiple description coding via polyphase transform and selective quantization," *in Proc. SPIE Conf. Visual Commun. And Image Processing*, San Jose, CA, 1999, vol. 3653, pp. 998-1008
- [28] J. G. Apostolopoulos, "Error-resilient video compression through the use of multiple states," *in Proc. IEEE Int. Conf. Image Processing*, vol. 3, 2000, pp. 352–355.
- [29] Refat Kibria, Joohee Kim, "H.264/AVC-based multiple description coding for wireless video transmission", *International Conference on Communications*, July 2008
- [30] J.G. Apostolopoulos, "Reliable video communication over lossy packet networks using multiple state encoding and path diversity," *Visual Communications and Image Processing (VCIP)*, January 2001.
- [31] G. A. Thomas, "Television motion measurement for DATV and other applications," *BBC Research Report BBC RD 1987/11*.

- [32] T. Tillo and G. Olmo, "A low complexity pre-post processing multiple description coding for video streaming," in *Proc. IEEE ICTTA*, Damascus, Syria, Apr. 2004, pp. 519–520
- [33] T. Tillo, E. Baccaglini, and G. Olmo, "Multiple descriptions based on multirate coding for JPEG 2000 and H.264/AVC", *IEEE TRANS. Image Processing*, vol. 19, pp. 1756–1767, July 2010.
- [34] I. Radulovic, P. Frossard, Y. K. Wang, M. M. Hannuksela, and A. Hallapuro, "Multiple description video coding with H.264/AVC redundant pictures," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 10, no. 1, pp. 144–148, 2010.
- [35] H. Bai, Y. Zhao, and C. Zhu, "Multiple description video coding using adaptive temporal sub-sampling," in *Proceedings of the IEEE International Conference on Multimedia and Expo. (ICME '07)*, pp. 1331–1334, Beijing China, July 2007.
- [36] M. Zhang, W. Liu, R. Wang, H. Bai, "A Novel Multiple description video coding Algorithm," *International Conference on Computational Intelligence and Security*, 2008.
- [37] V. Parameswaran, A. Kannur, B. Li, "Adapting quantization offset in multiple description coding for error resilient video transmission" *Visual Communication and Image Representation*, pp. 491-503, July 2009.
- [38] V. A. Vaishampayan, "Design of multiple description scalar quantizers", *IEEE Trans. Inform. theory*, vol. 39, pp. 821-834, May 1993.
- [39] C. Tian and S.S. Hemami, "A special class of multiple description scalar quantizers," in *Proceedings of IEEE Information Theory Workshop*, Oct. 2004, pp. 135–140
- [40] U. Samarawickrama, J. Liang, "A two-Stage algorithm for multiple description predictive coding", *ccece2008*
- [41] A. Reibman *et al.*, "Multiple description video using rate-distortion splitting," in *Proc. IEEE Int.Conf. Image Processing*, vol. 1, 2001, pp. 978–981
- [42] K. R. Matty and L. P. Kondi, "Balanced multiple description video coding using optimal partitioning of the DCT coefficients" *IEEE TRANS. Circuits Syst. Video Technol*, vol. 15, pp. 928-934, JULY 2005.

- [43] D. Comas, R. Singh, A. Ortega, and F. Marques, "Unbalanced multiple description video coding with rate-distortion optimization," *EURASIP J. Appl. Signal Process.*, vol. 2003, no. 1, pp. 81–90, Jan. 2003.
- [44] Doo-Man Chung and Yao Wang, "Multiple description image coding using signal decomposition and reconstruction based on lapped orthogonal transforms" *IEEE Trans. Circuits Syst. Video Technol.*, vol. 9, pp. 895-908, Sep. 1999.
- [45] G. Sun, U. Samarawickrama, J. Liang, C. Tian, C. Tu, and T. D. Tran, "Multiple description coding with prediction compensation," *IEEE Trans. Image Processing*, Vol. 18, No. 5, May 2009
- [46] Y. Wang, M. T. Orchard, and A. R. Reibman, "Multiple description image coding for noisy channels by pairing transform coefficients," in *Proc. 1997 IEEE 1st Workshop Multimedia Signal Processing*, Princeton, NJ, June 1997, pp. 419–424.
- [47] M. T. Orchard, Y. Wang, V. Vaishampayan, and A. R. Reibman, "Redundancy rate-distortion analysis of multiple description coding using pairwise correlating transforms", *Proc. IEEE Int. Conf. Image Process (ICIP97)*, vol. I, pp. 608 - 611, 1997.
- [48] V.K. Goyal and Jelena Kovacevic;, "Optimal multiple description transform coding of Gaussian vectors", *Proc. IEEE Data Compression Conf.*, pp. 388 - 397, 1998.
- [49] V.K. Goyal, Jelena Kovacevic, R. Arean, and M. Vetterli, "Multiple description transform coding of images", *Proc. IEEE Int. Conf. Image Processing*, vol. 1, pp. 674 - 678, 1998.
- [50] Y. Wang, M. Orchard, and A. Reibman, "Optimal pairwise correlating transforms for multiple description coding", *Proc. IEEE Int. Conf. Image Processing (ICIP'98)*, pp. 1998.
- [51] Y. Wang, et al., "Multiple description coding using pairwise correlating transforms", *IEEE Trans. Image Processing*, vol. 10, pp. 351 - 366, 2001.
- [52] Y. Wang , A. R. Reibman , M. T. Orchard and H. Jafarkhani "An improvement to multiple description transform coding", *IEEE Trans. Signal Process.*, vol. 50, pp. 2843-2854, 2002.

- [53] A. Reibman, H. Jafarkhani, Y. Wang, M. Orchard, and R. Puri, "Multiple-description video coding using motion-compensated temporal prediction", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, pp. 193 - 204, 2002.
- [54] Vivek K Goyal, and Jelena Kovacevic, "Generalized multiple description coding with correlating transforms," *IEEE Trans. Inform. Theory*, vol. 47, pp. 2199-2224, SEPTEMBER 2001
- [55] U. Samarawickrama, J. Liang, and C. Tian, "M-channel multiple description coding with two-rate coding and staggered quantization", *IEEE Trans. Circuits Syst. Video Technol.*, VOL. 20, pp. 933-944, JULY 2010
- [56] D. Wang, N. Canagarajah and D. Bull "Slice group based multiple description video coding using motion vector estimation", *Proc. IEEE ICIP*, pp. 3237 2004.
- [57] C.C. Su, J. Yao and H. Chen, "H.264/AVC based multiple description coding scheme", *Proceedings of International Conference on Image Processing*, pp. 265-268, September 2007.
- [58] Tammam Tillo, Marco Grangetto, and Gabriella Olmo "Redundant slice optimal allocation for H.264 multiple description coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 18, pp. 59-70, JANUARY 2008
- [59] L. Peraldo, E. Baccaglini, E. Magli, G. Olmo, R. Ansari, Y. Yao, "Slice-level rate-distortion optimized multiple description coding for H.264/AVC", *IEEE International Conference on Acoustics Speech and Signal Processing*, 2010
- [60] Andres Albanese, Johannes Blomer, Jeff Edmonds, Michael Luby, and Madhu Sudan, "Priority encoding transmission", *IEEE Trans. Inform Theory*, vol. 42, pp. 1737-1744, NOVEMBER 1996
- [61] A. E. Mohr, E. A. Riskin, and R. E. Ladner, "Unequal loss protection: Graceful degradation of image quality over packet erasure channels through forward error correction", *IEEE J. Select. Areas Commun.*, vol. 18, pp. 819 - 828, 2000.
- [62] R. Puri and K. Ramchandran, "Multiple description source coding through forward error correction codes," in *33rd Asilomar Conf. Signals, Systems and Computers*, CA, Oct. 1999

- [63] A. El. Essaili, S. Khan, W. Kellerer and E. Steinbach, "Multiple description video transcoding", *IEEE International Conference on Image Processing*, 2007.
- [64] M. T. Lu, J. C. Wu, K. J. Peng, P. Huang, J. J. Yao, and H. H. Chen, "Design and evaluation of a P2P IPTV system for heterogeneous networks," *IEEE Trans. Multimedia*, vol. 9, no. 8, pp. 1568–1579, Dec. 2007.
- [65] C. W. Hsiao and W. J. Tsai, "Hybrid multiple description coding based on H.264", *IEEE Trans. Circuits Syst. Video Technol*, vol. 20, pp. 76-87, January 2010.
- [66] Z. Xu, Z. Lin and A. Makur, "Multiple description image coding with hybrid redundancy" *IEEE Asia Pacific Conference on Circuits and Systems*, 2006
- [67] X. Wen, O. C. Au, J. Xu, Z. Liang, Y. Yang and W. Tang, "A novel multiple description video coding based on H.264/AVC video coding standard", *ISCAS 2009*.
- [68] C. Zhu and M. Liu, "Multiple description video coding based on hierarchical B pictures," *IEEE Transactions on Circuits Systems Video Technology*, vol. 19, no. 4, pp. 511–521, April 2009.
- [69] A.B. Zhao, W.S. Wang, H.J. Cui and K. Tang, "Efficient multiple description scalable video coding scheme based on weighted signal combinations," *Tsinghua Science and Technology*, vol. 12, no. 1, pp. 86-90, February 2007.
- [70] S. Tian and P. K. Rajan, "Multiple description coding using transforms and data fusion," in *Proc. Int. Conf. Information Technology: Coding and Computing*, 2005, vol. 1, pp. 85–90.
- [71] J. Ostermann, J. Bormans, P. List, D. Marpe, M. Narroschke, F. Pereira, T. Stockhammer, and T. Wedi, "Video coding with H.264/AVC; tools, performance, and complexity," *IEEE Circuits Syst. Mag.*, vol. 4, no. 1, pp. 7–28, 1Q, 2004.
- [72] J. G. Apostolopoulos and M. D. Trott "Path diversity for enhanced media streaming", *IEEE Commun. Mag.*, vol. 42, pp. 80-87, 2004.
- [73] O. A. Lotfallah, and S. Panchanathan, "Adaptive multiple description coding for internet video," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol.5, pp. 732-735.

- [74] Brian A. Heng, John G. Apostolopoulos, and Jae S. Lim, "End-to-end rate-distortion optimized MD mode selection for multiple description video coding," *EURASIP Journal on Applied Signal Processing*, 2006.
- [75] A. Papoulis, "Probability, Random Variables, and Stochastic Processes," New York: *McGraw-Hill*, 1984.
- [76] G. Yovanof and S. Liu, "Statistical analysis of the DCT coefficients and their quantization error," in *Proc. 30th Asilomar Conf. Signals, Syst. Comput.*, 1996, vol. 1, pp. 601–605.
- [77] E. Lam and J. Goodman, "Modeling DCT coefficients for fast video encoding," *IEEE Trans. Circuits Syst. Video Technol.* vol. 9, no.4, pp. 608–616, June. 1999.
- [78] Y. Altunbasak and N. Kamaci, "An analysis of the DCT coefficient distribution with the H.264 video coder," in *Proc. IEEE Int. Conf. coust., Speech, Signal Process. (ICASSP)*, 2004, vol. 3, pp. III-177–80.
- [79] G. J. Sullivan, "Efficient Scalar Quantization of Exponential and Laplacian Random Variables", *IEEE Transactions on Information Theory*, vol. 42, no. 5, pp. 1365-1374, Sept. 1996.
- [80] G. J. Sullivan and S. Sun, "On dead-zone plus uniform threshold scalar quantization," in *Proc. SPIE Visual Commun. Image Process.*, Beijing, China, Jul. 2005, pp. 1041–1052.
- [81] G. Sullivan, "Adaptive quantization encoding technique using an equal expected-value rule," *presented at the 14th Meeting of Joint Video Team, JVT-N011*, Hong Kong, China, Jan. 2005.
- [82] X. Li, N. Oertel, A. Hutter, A. Kaup, "Laplace distribution based lagrangian rate distortion optimization for hybrid video coding," *IEEE Trans. Circuits Syst. Video Technol.* Vol. 19, No. 2, pp. 193–205, February 2009.
- [83] M. Orchard, Y. Wang, V. Vaishampayan, and A. Reibman, "Redundancy rate distortion analysis of multiple description image coding using pairwise correlating transforms," in *Proc. Int. Conf. Image Process.*, Santa Barbara, CA, Oct. 1997.

- [84] C. Lin, T. Tillo, Y. Zhao, and B. Jeon, "Multiple description coding for H.264/AVC with redundancy allocation at macro block level," *IEEE Trans. Circuits Syst. Video Technol.*, VOL. 21, pp. 589-600, MAY 2011.
- [85] J. Kim, and J.C. Hong, "Channel-Adaptive Multiple Description Coding for Wireless Video Streaming", *In Proceedings of ICCCN 2007, Honolulu, Hawaii USA*, Aug. 13 - 16, 2007, pp. 474-478.
- [86] N. Kamnoonwatana, D. Agrafiotis and N. Canagarajah, "Flexible Adaptive Multiple Description Coding for Video Transmission," *IEEE Trans. Circuits Syst. Video Technol.*, Accepted to be published.
- [87] A. Reibman, "Optimizing multiple description video coders in a packet loss environment," *presented at the Packet Video Conf. 2002, Pittsburgh, PA*.
- [88] R. Zhang, S. L. Regunathan, and K. Rose, "Video coding with optimal inter/intra-mode switching for packet loss resilience," *IEEE J. Sel. Areas Commun.*, vol. 18, no. 6, pp. 966–976, Jun. 2000.
- [89] K. Stuhlmuller, N. Farber, M. Link, and B. Girod, "Analysis of video transmission over lossy channels," *IEEE J. Select. Areas Commun.*, vol. 18, pp. 1012–1032, June 2000.
- [90] J. Chakareski, J. Apostolopoulos, W.-T. Tan, S. Wee, and B. Girod, "Distortion chains for predicting the video distortion for general packet loss patterns," in *Proc. ICASSP*, vol. 5. Montreal, QC, Canada, May 2004, pp. 1001–1004.
- [91] Z. He, J. Cai, and C. W. Chen, "Joint source channel rate-distortion analysis for adaptive mode selection and rate control in wireless video coding," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 12, Special Issue on Wireless Video, no. 6, pp. 511–523, Jun. 2002.
- [92] Y. Wang, Z. Wu, and J. M. Boyce, "Modeling of transmission-loss-induced distortion in decoded video," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 16, no. 6, pp. 716–732, Jun. 2006.
- [93] H. Schwarz, D. Marpe, and T. Wiegand, "Overview of the scalable video coding extension of the H.264/AVC standard," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 17, no. 9, pp. 1103–1120, Sep. 2007.

- [94] M. Kazemi, K. Sadeghi, and S. Shirmohammadi, "A high video quality Multiple Description Coding scheme for lossy channels," *IEEE International Conference on Multimedia and Expo (ICME), 2011*
- [95] M. Kazemi, K. Sadeghi, and S. Shirmohammadi, "A Mixed Layer Multiple Description Video Coding Scheme," *IEEE Transactions on Circuits and Systems for Video Technology*, Accepted to be published, DOI: 10.1109/TCSVT.2011.2159431, 2011.
- [96] M. Kazemi, K. H. Sadeghi, S. Shirmohammadi, "A Review of Multiple Description Coding techniques for Video Transmission over Best-Effort Networks," submitted to *IEEE Trans. Circuits Syst. Video Technol.*
- [97] M. Kazemi, K. H. Sadeghi, S. Shirmohammadi, "End-to-End Distortion Optimization in DCT-Based Multiple Description Video Coding," submitted to *IEEE Trans. Multimedia*.

Abstract

Multiple distortion coding is a promising solution for video transmission over lossy channels. In MDC multiple descriptions of a source are generated which are dependently decodable and mutually refinable. When all descriptions are available, the corresponding quality is called central quality; otherwise it is called side quality. Generally, there exist a trade-off between side and central quality in all MDC schemes. The MDC methods which provide better central-side quality trade-off are of more interest to the designers.

In this thesis a new MDC scheme is introduced which has better trade-off between side and central quality. In other words, for the same central quality, it provides higher side quality; or equivalently for the same side quality, it has higher central quality. This method is based on the mixing of the base and enhancement layers of Coarse-Grain Scalable (CGS) coding and hence is called Mixed Layer MDC (MLMDC). At the central decoder the layers are separated and we have two-layer quality, such as in CGS decoder. At the side decoder, some descriptions are not available and hence we cannot separate the layers, directly. We propose to use estimation for this purpose. MLMDC for two-description coding and four-description coding is implemented in JM16.0, H.264/AVC reference software. The experimental results show that for videos which have high enough content (texture and motion activity), MLMDC in comparison to the conventional MDCs, provides higher side quality while they have the same central quality. In addition, when video transmission over channels with packet loss, MLMDC provides higher average video quality, in particular for four-description coding.

In error prone environments, we need to have higher side quality while in less noisy conditions the higher central quality is more important. Therefore, in order to have the best quality in the different channel conditions, an optimization is needed. For this purpose, a new model for end-to-end distortion is introduced which takes into account both quantization and transmission distortions for predicting the quality at the receiver side. The transmission distortion is the result of error propagation which in turn originated from the mismatch between side and central decoder outputs. The derived model is applicable for all DCT-domain MDCs. With experimental results, the model is verified first and then used for MDCs optimization. The results show the performance of the optimizer and also MLMDC higher video quality compared to that of conventional MDCs when they are designed optimally.

Index Terms- Multiple Description Coding, Video transmission, Packet Loss, Error propagation, End-to-end distortion, Channel adaptive optimization, H.264/AVC.



Sharif University of Technology

Electrical Engineering Department

Ph.D Thesis

**Multiple description video coding based on base and
enhancement layers of SVC and channel adaptive
optimization**

By:

Mohammad Kazemi Varnamkhasti

Supervisors:

Dr. Khosrow Haj Sadghi

Dr. Shervin Shirmohammadi

February 2012