H.264/AVC

Cem Yusuf Aydoğdu
Mehmet Taner Ünal

What will we cover?

- **Introduction**

  Definitions, background and history, improvements, methods, comparisons

- **System**

  Terminology, structure for the encoder and decoder of H.264, profiles

- **Experimental Results**

  Testing, PSNR, SSIM

Introduction

Background

- A video compression standard H.264, in the other words a codec, was cooperatively advanced by the International Telecommunications Union and International Organization for Standardization/International Electrotechnical Commission Moving Picture Experts Group.

- It was known as "H.264" by ITU and "MPEG-4 Part 10, Advanced Video Coding (AVC)" by ISO/IEC. So they both are basically the same thing.

What is x264?

- x264 is a free software library and application for encoding video streams into the H.264/MPEG-4 AVC compression format, and is released under the terms of the GNU GPL.
- It achieves dramatic performance, encoding 4 or more 1080p streams in realtime on a single consumer-level computer
- It has the most advanced psychovisual optimizations

Some Softwares using x264

- Avidemux
- ELDER
- Ffdshow
- ffmpeg
- GordianKnot
- Handbrake
- LiVES
- MeGUI
- Mencoder
- Bencos (RealAnime)
- StaxRip
- VLC Media Player

x264 forms the core of many web video services, such as Youtube, Facebook, Vimeo, and Hulu. It is widely used by television broadcasters and ISPs.

Main Objectives of H.264/AVC Standardisation

- to increase compression performance

- to boost provision of a network-friendly video representation addressing conversational (video telephony) and non-conversational (storage, streaming, or broadcasting) applications

History of Video Coding Standards till H.264

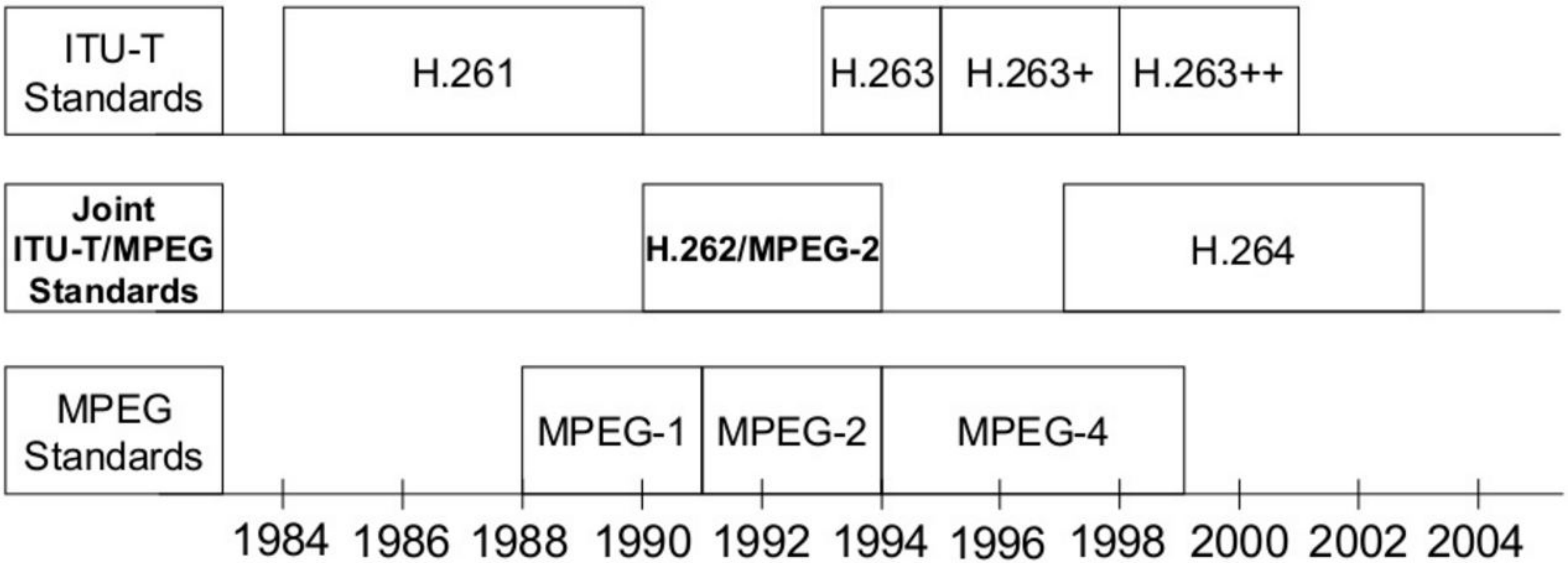


Figure: Evolution of the video coding standards

(Reference: http://www.slideshare.net/sajan45/h264-video-standard)

Improvements and Methods

- H.264 can convey MPEG-4 quality with a frame size up to four times greater.
- It provides MPEG-2 quality at a reduced data rate.

To achieve better compression and perceptual quality these are used:

- Motion estimation
- Intra estimation
- Transformation of motion estimation and intra estimation into the frequency domain
- Reduction of compression artifacts
- Entropy coding

Some more information

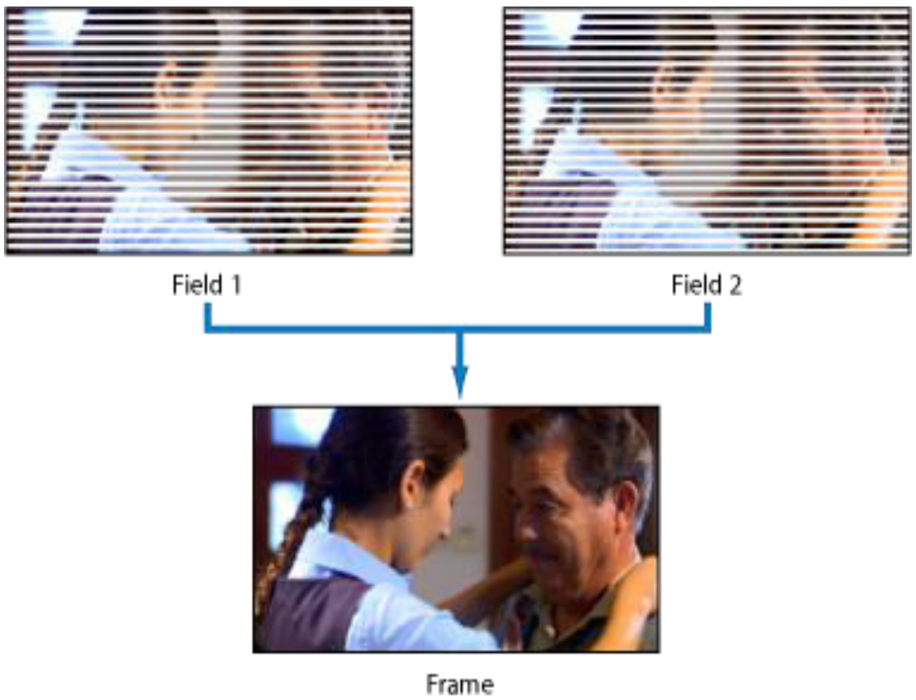
- To obtain very exact portrayal of the displacements of moving areas, H.264 grants quarter-pixel precision for motion compensation.
- For video network delivery and for delivery of high definition video, H.264 is more charming than MPEG-4

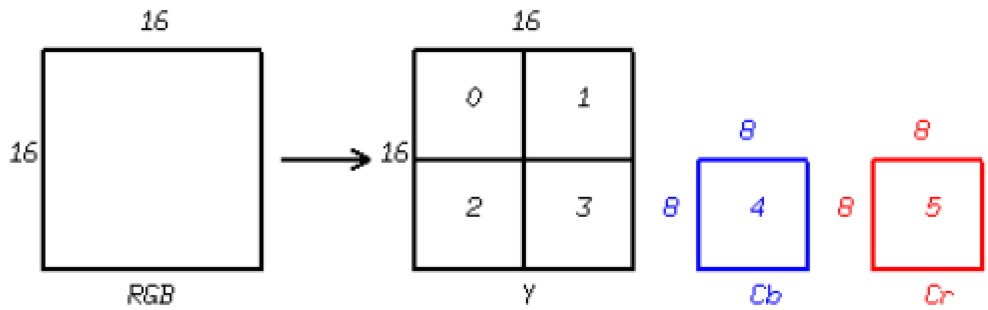
H.264 is promising to be the next standard for format convergence in the digital video industry disregarding of the platform and it is backed by big internet players like Google/YouTube, Adobe, and Apple iTunes

System

Terminology

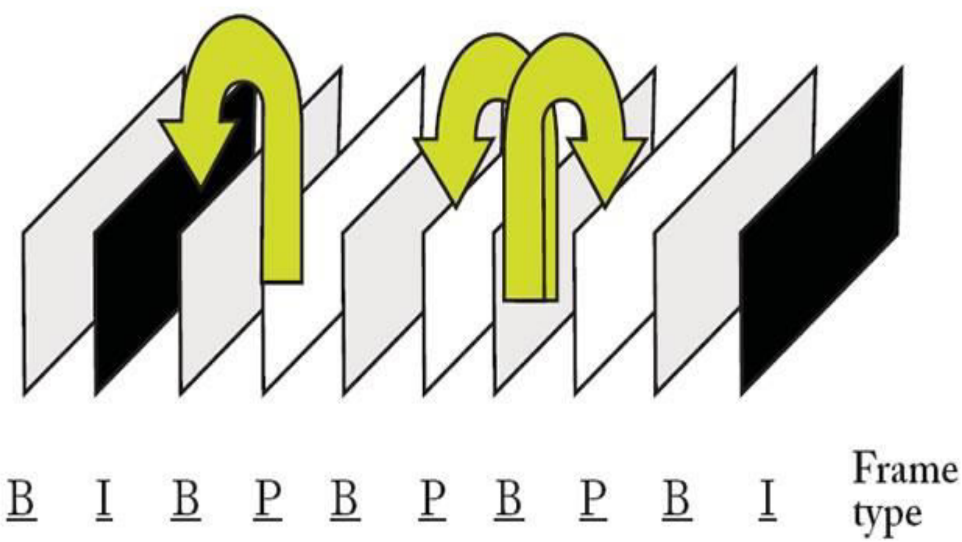
- Field: Partial frame
- Frame consists combined fields
- Macroblock: Fixed size region





Terminology

- Slice: Set of macroblocks



Different frame types:

I : Intra

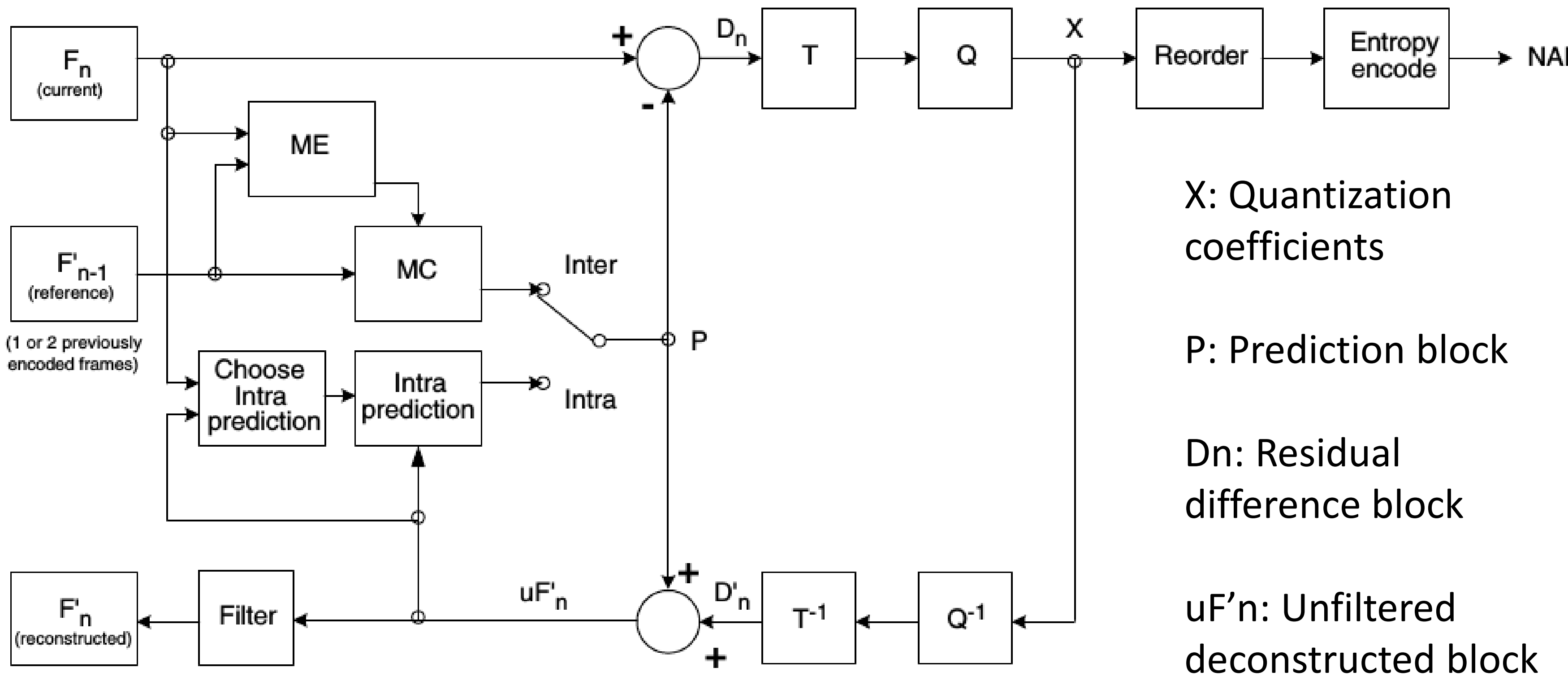
P : Predictive

B : Bi-predictive

SI : Switching I

SP : Switching P

Block Diagram of Encoder



X: Quantization coefficients

P: Prediction block

Dn: Residual difference block

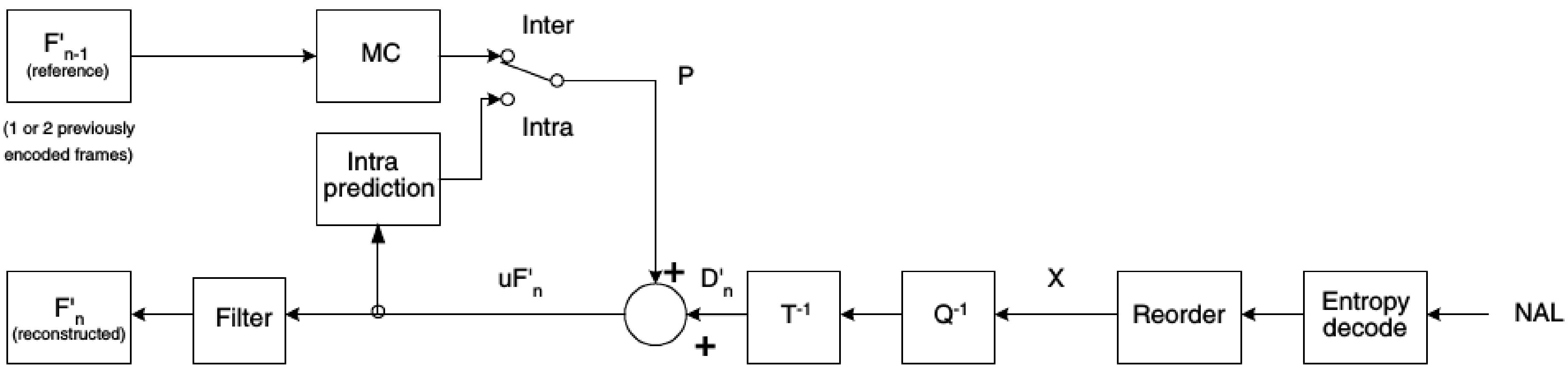
uF'n: Unfiltered deconstructed block

ME: Motion Estimation
MC: Motion Compensation

T: Transform
Q: Quantization

NAL: Network Abstraction Layer

Block Diagram of Decoder



MC: Motion Compensation

D'n: Residual difference block

uF'n: Unfiltered deconstructed block

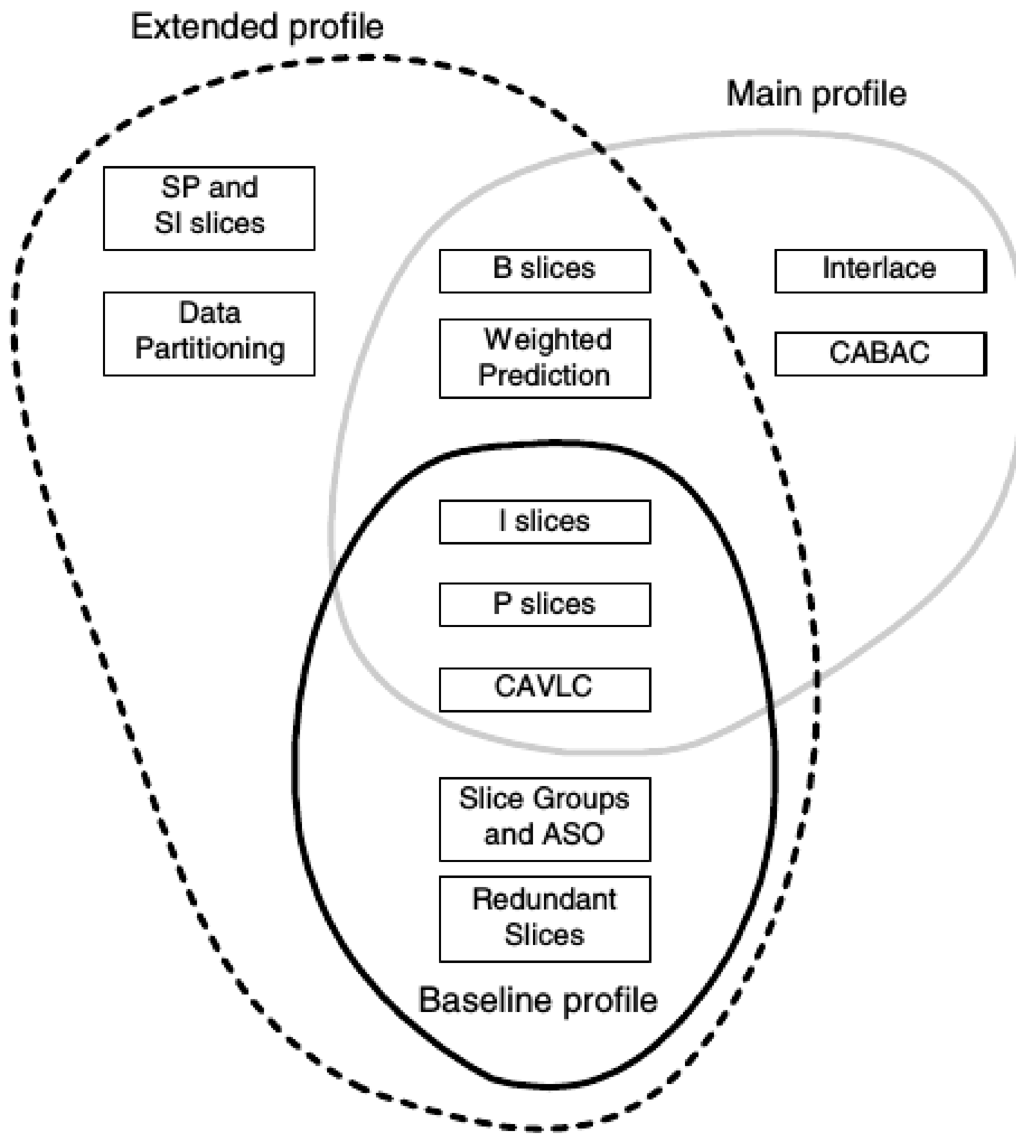
P: Prediction block

Profiles

1. Baseline Profile
   - Basic Functionalities

1. Main Profile
   - Storing, broadcasting

2. Extended Profile
   - Streaming



Profiles

4. High Profiles
   - For better quality

Coding Tools	High	High 10	High 4:2:2	High 4:4:4
Main Profile Tools	X	X	X	X
4:2:0 Chroma Format	X	X	X	X
8 Bit Sample Bit Depth	X	X	X	X
8x8 vs. 4x4 Transform Adaptivity	X	X	X	X
Quantization Scaling Matrices	X	X	X	X
Separate Cb and Cr QP control	X	X	X	X
Monochrome video format	X	X	X	X
9 and 10 Bit Sample Bit Depth		X	X	X
4:2:2 Chroma Format			X	X
11 and 12 Bit Sample Bit Depth				X
4:4:4 Chroma Format				X
Residual Color Transform				X
Predictive Lossless Coding				X

Experimental Results

Testing

Video Name	Resolution	Frame Per Second	Chroma Format
coastguard	352 x 288	29.97	4 : 2 : 0
container	352 x 288	29.97	4 : 2 : 0
news	352 x 288	29.97	4 : 2 : 0
riverbed	1920 x 1080	25	4 : 2 : 0

Metrics

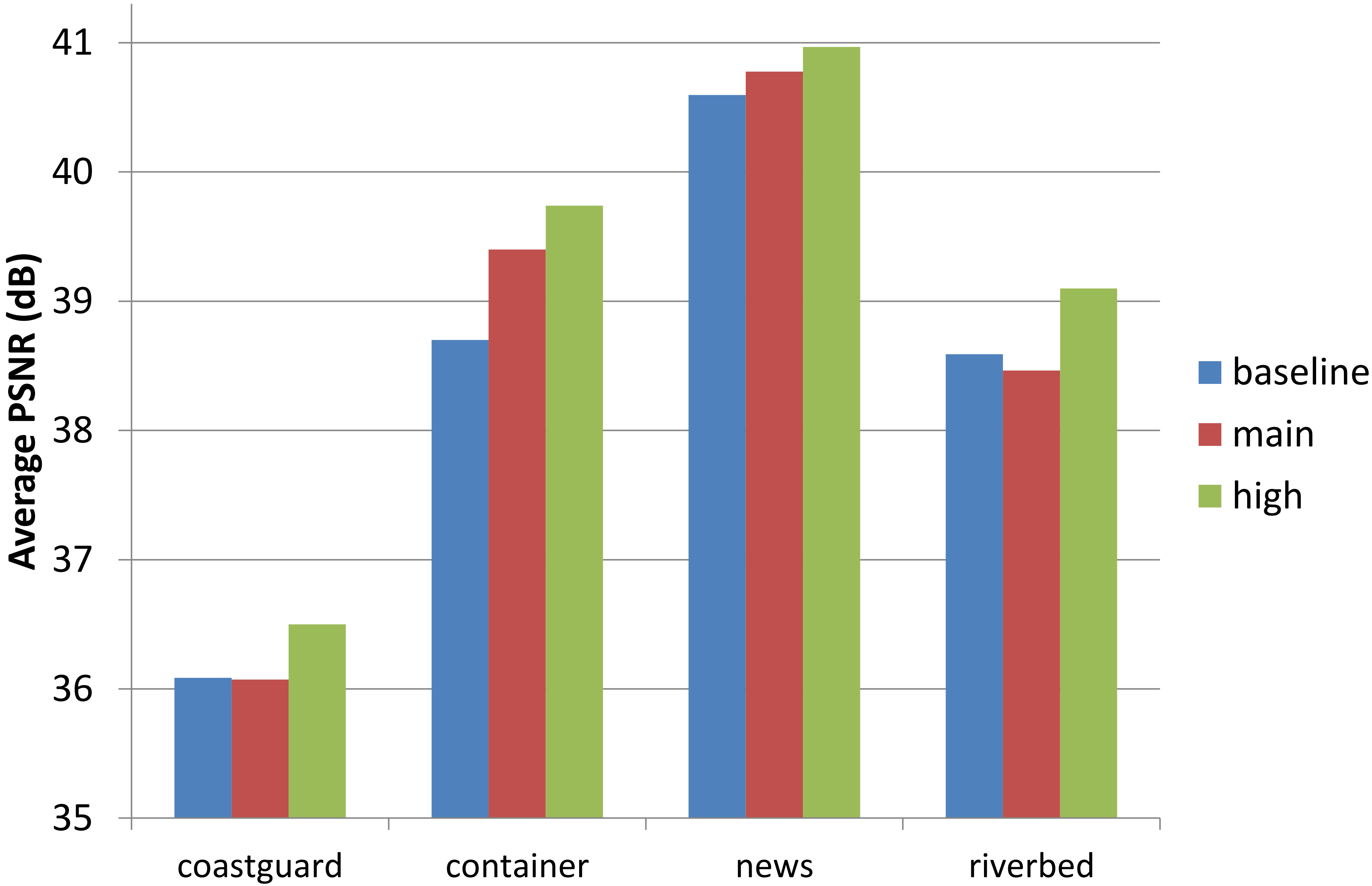
1. Peak Signal To Noise Ratio (PSNR):

$$PSNR = 10\ \log_{10} \frac{MAX^2}{MSE}$$

MAX is average of the maximum values in each image component

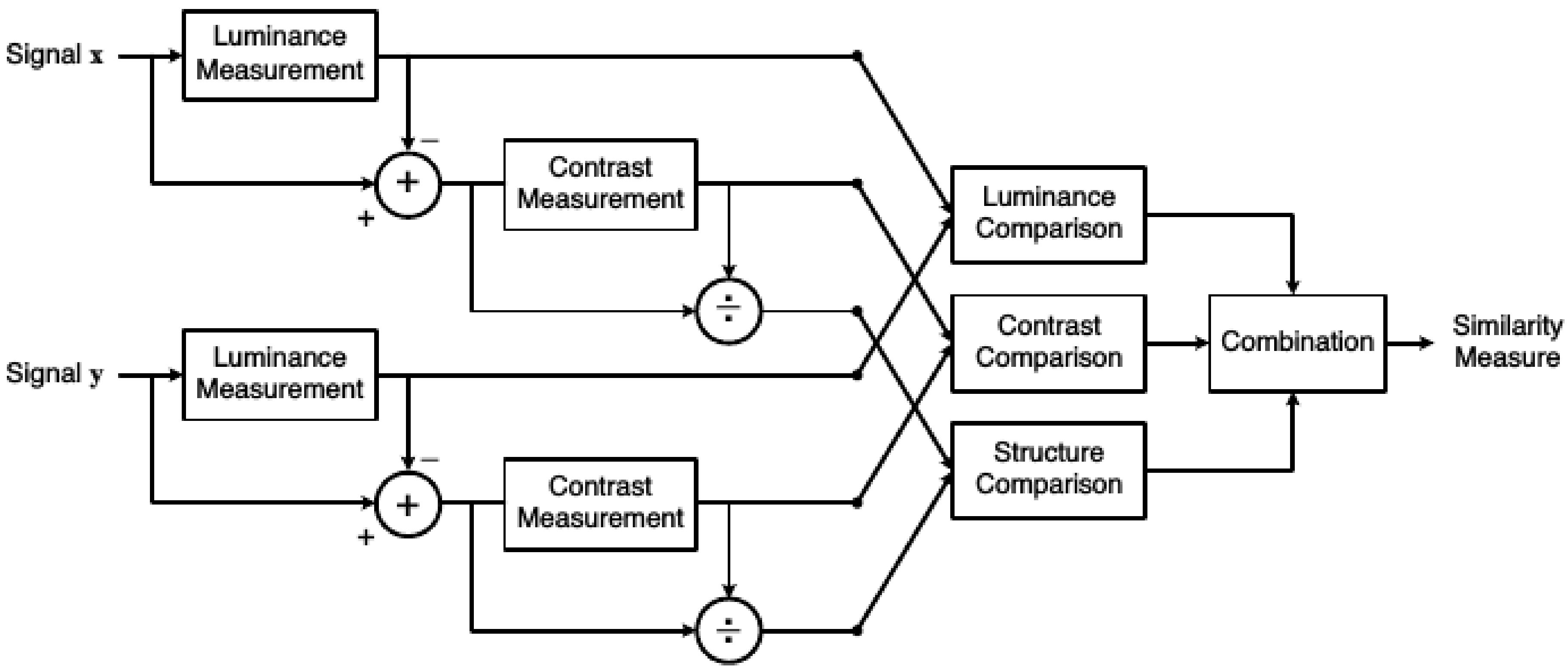
MSE is mean square error

Comparing Average PSNR Values



Metrics

2. Structural Similarity Metric (SSIM):



Metrics

$$SSIM(x, y) = [l(x, y)]^{\alpha} \cdot [c(x, y)]^{\beta} \cdot [s(x, y)]^{\gamma}$$

where

$$l(x, y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1},$$

$$c(x, y) = \frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2 + \sigma_y^2 + C_2},$$

$$s(x, y) = \frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y + C_3}$$

μ values denote local means

σ values denote standard deviations

Comparing SSIM Values