SARNOFF Corporation
a subsidiary of SRI International

Empowering Solutions Through Innovation

H.264 Format/Bitrate/Quality Tradeoff Study

Dr. Arkady Kopansky, Dr. Michael Isnardi, Dr. Ali Chaudhry

Dr. Bing-Bing Chai, Pinghua Xiong

May 19, 2010

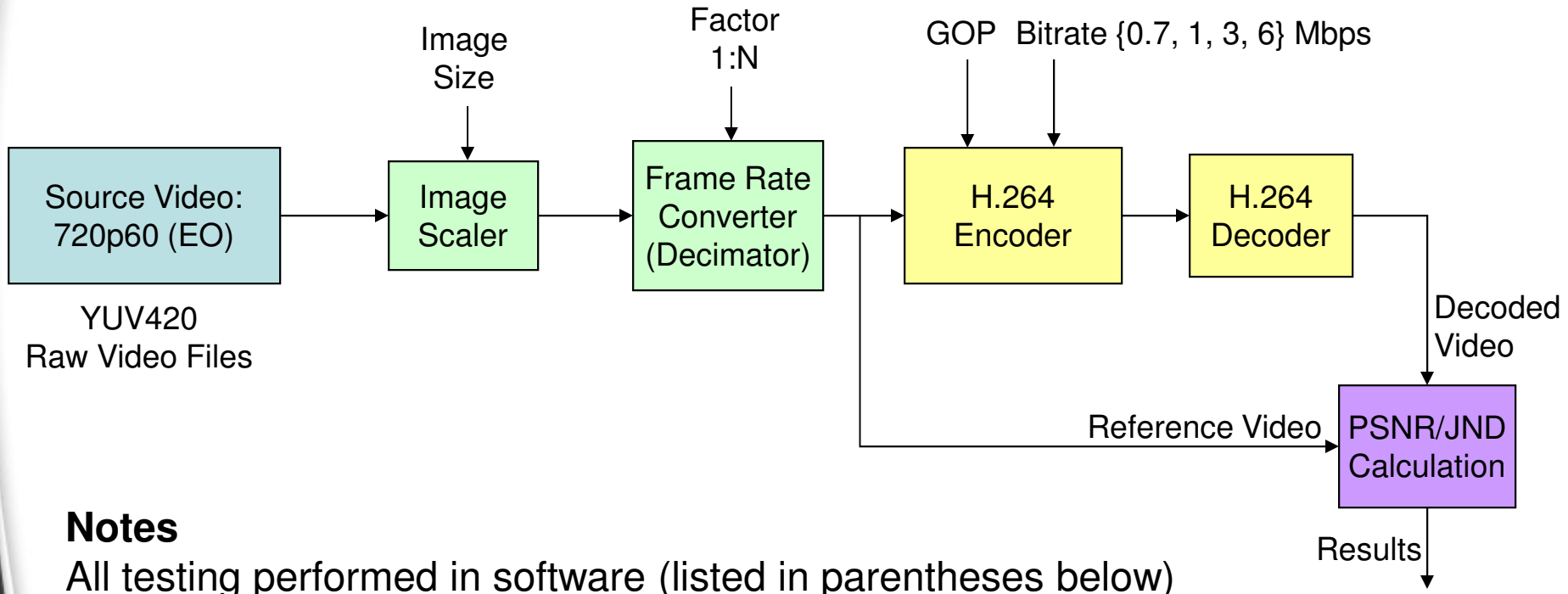
Outline

- Purpose of the study/Motivation
- Experimental Setup
- Metrics used
  - *PSNR*
  - *JND*
- Results
- Summary
- Conclusion

Purpose of the Study/Motivation

- Air-to-ground throughput can vary dramatically over short periods of time
  - *E.g., range from 6 Mbps to 0.7 Mbps*
- Maximizing transmitted video quality at the highest resolution and highest frame rate is desirable, whenever possible
- If available bandwidth does not allow transmission of full resolution video at native frame rate, multiple approaches can be employed to maximize transmission quality for the video at a given bitrate
  - *Lowering Frame Rate (maintains resolution at lower bitrates)*
  - *Lowering Frame Size (maintains frame rates at lower bitrates)*
  - *Lowering both Frame Size and Frame Rate (allows transmission through poor channels)*
- What is the best combination of frame size and frame rate at a given bitrate for maximum quality video transmission?
  - *Mission-dependent*
    - High Resolution needed for analysis
    - High Frame Rate needed for tracking

Block-diagram for Testing Format/Bitrate/Quality Tradeoffs



Notes

All testing performed in software (listed in parentheses below)

Image Scaler (ffmpeg): optionally downsamples input video

Frame Rate Converter (Sarnoff developed): optionally drops input frames

H.264 Encoder (MainConcept): GOP structure and Bitrate are coding parameters

H.264 Decoder (Tektronix MTS4EA): outputs raw decoded video

Sample Video Frame from the Aerial Video Clip Used for This Study

Video Source: SAIC

Encoder and Settings

- Encoder: MainConcept H.264 software encoder (version 2.01)
- Main Profile
- CABAC entropy coding
- I, P, stored and disposable B picture types
  - *All I pictures are IDR*
- Number of reference pictures: 4
- Nominal GOP length: 33 frames
  - *GOP length and structure is adaptive with scene change detection*
- Rate Control Mode: Variable bitrate, single pass

PSNR and Bits Per Pixel (bpp) Definitions

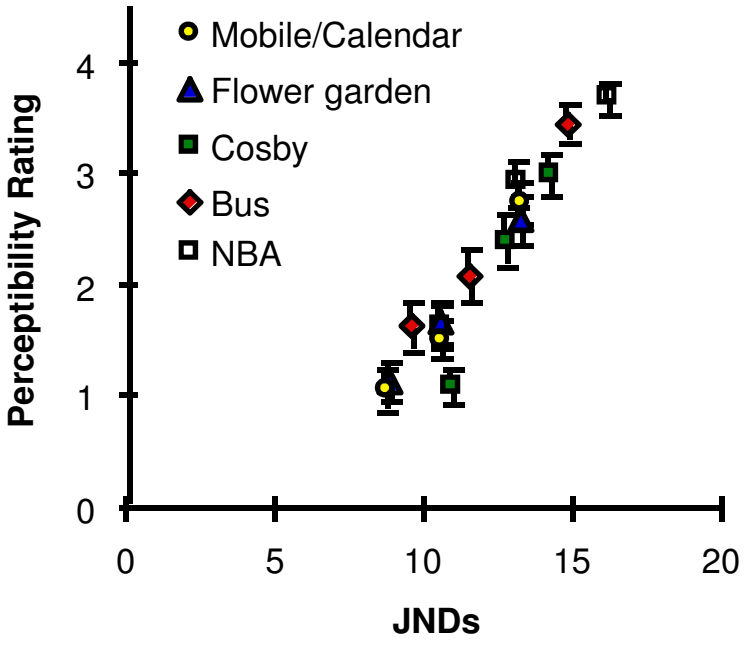
$$PSNRY[\text{dB}] = 10 \cdot \log_{10}\left(\frac{(2^b - 1)^2}{\frac{1}{W \cdot H}\sum_{h=1}^{H}\sum_{w=1}^{W}(I_1(w,h) - I_2(w,h))^2} \right),$$

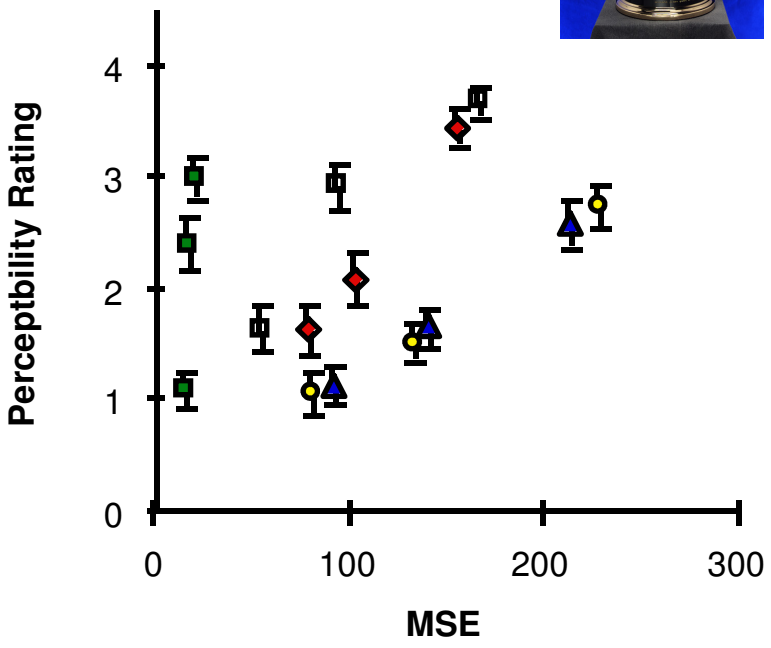
b denotes bits per pixel [bit depth; 8 bpp source data used for this study]
I_1 and I_2 are two images under consideration [one is a reference]
W and H are image width and height, respectively
w and h denote specific pixel locations

$$bpp = \frac{B}{W \cdot H \cdot S \cdot F},$$

B denotes bitrate in bits per second
W and H are image width and height, respectively
S is a scaling factor (1 for YUV400, 1.5 for YUV420, 2 for YUV422, 3 for YUV444)
F denotes frame rate

JND: A Proven Metric





- Sarnoff's Video Quality Measure – Just Noticeable Differences (JNDs) – correlates with subjective quality ratings
- Other measures, like PSNR and mean-squared error, do not
  - *PSNR does correlate for same sequence, same format*

Formats and Bitrates

Resolution	Frame Rate			
	60fps	30fps	15fps	10fps
1280x720	6Mbps 3Mbps 1Mbps	6Mbps 3Mbps 1Mbps	6Mbps 3Mbps 1Mbps 0.7Mbps	0.7Mbps
640x480	6Mbps 3Mbps 1Mbps	6Mbps 3Mbps 1Mbps	6Mbps 3Mbps 1Mbps	
320x240	6Mbps 3Mbps 1Mbps	6Mbps 3Mbps 1Mbps	6Mbps 3Mbps 1Mbps	

PSNR Results–Average Sequence PSNRY, dB

Format	Bitrate							
	6Mbps	bpp	3Mbps	bpp	1Mbps	bpp	0.7Mbps	bpp
1280x720@60fps	39.6	0.072	37.86	0.036	33.15	0.012		
1280x720@30fps	40.8	0.145	39.15	0.072	35.28	0.024		
1280x720@15fps	42.22	0.289	40.34	0.145	36.79	0.048	35.03	0.034
1280x720@10fps							35.46	0.051
640x480@60fps	42.41	0.217	40.11	0.109	36.02	0.036		
640x480@30fps	44.44	0.434	41.93	0.217	37.85	0.072		
640x480@15fps	47.08	0.868	43.89	0.434	39.52	0.145		
320x240@60fps	47.96	0.868	44.47	0.434	39.64	0.145		
320x240@30fps	52.33	1.736	47.35	0.868	41.86	0.289		
320x240@15fps	57.93	3.472	51.75	1.736	44.49	0.579		

Good Quality | Marginal Quality | Poor Quality | Below 0.04bpp

JND Results–Average Sequence JND

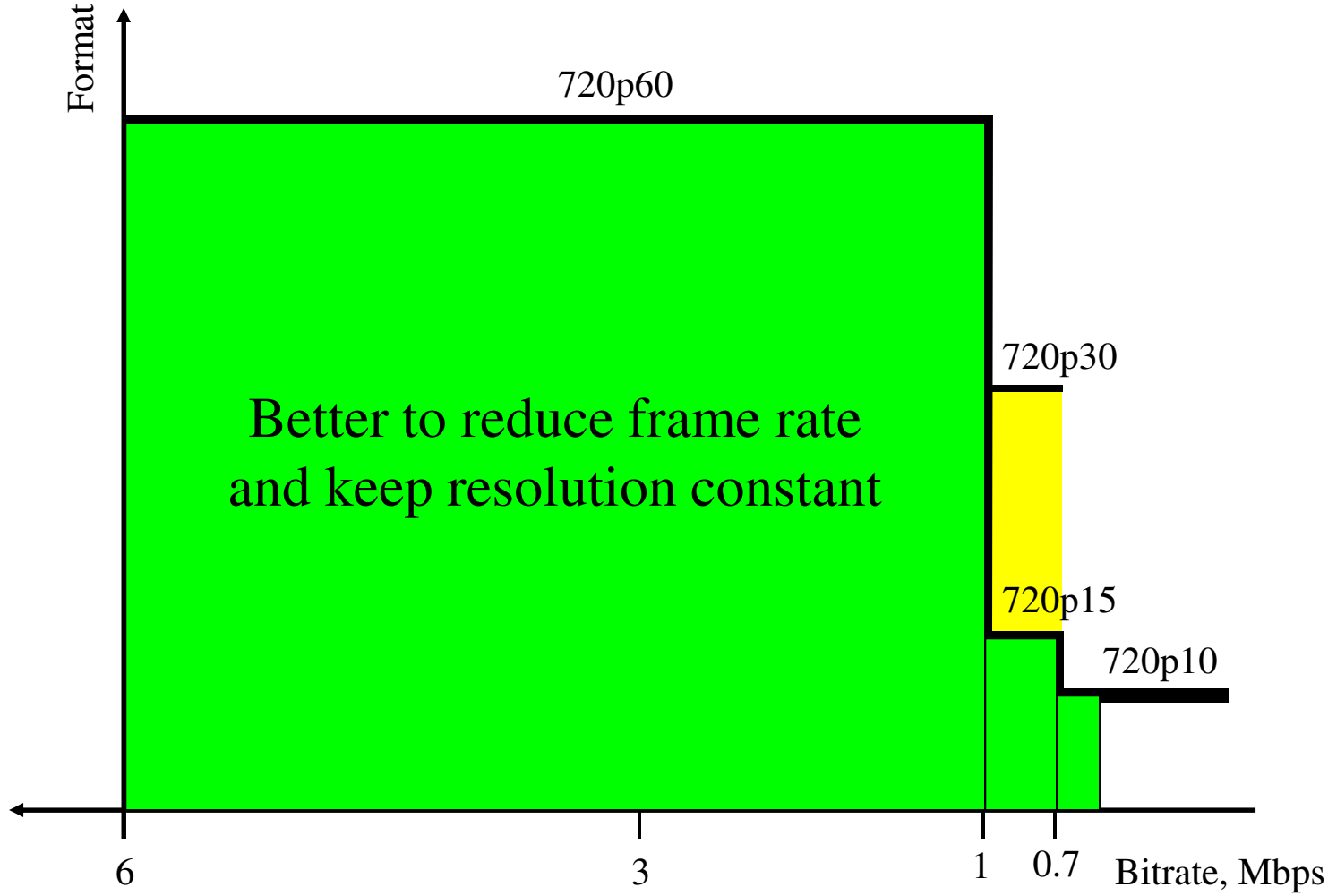
Format	Bitrate							
	6Mbps	bpp	3Mbps	bpp	1Mbps	bpp	0.7Mbps	bpp
1280x720@60fps	4.69	0.072	5.87	0.036	10.4	0.012		
1280x720@30fps	4.04	0.145	4.96	0.072	8.12	0.024		
1280x720@15fps	3.48	0.289	4.3	0.145	6.76	0.048	8.54	0.034
1280x720@10fps							7.75	0.051
640x480@60fps	3.09	0.217	3.94	0.109	6.48	0.036		
640x480@30fps	2.55	0.434	3.29	0.217	5.21	0.072		
640x480@15fps	1.98	0.868	2.7	0.434	4.32	0.145		
320x240@60fps	1.84	0.868	2.53	0.434	4.09	0.145		
320x240@30fps	1.21	1.736	1.94	0.868	3.25	0.289		
320x240@15fps	0.73	3.472	1.26	1.736	2.54	0.579		

Good Quality | Marginal Quality | Poor Quality | Below 0.04bpp

Demo

Resolution	Frame Rate			
	60fps	30fps	15fps	10fps
1280x720	6Mbps 3Mbps 1Mbps		1Mbps	0.7Mbps
640x480	1Mbps			
320x240				

Summary



Conclusions

- Aerial video can be transmitted with good quality at 3 Mbps maintaining 1280x720 resolution and 60fps frame rate
- At 1 Mbps, for good quality with 1280x720 resolution, frame rate has to be dropped to 15fps
  - *Quality is marginally acceptable at 30fps*
- At 0.7 Mbps, 1280x720 video can be transmitted with marginally acceptable quality at 10fps
- At 1 Mbps, to maintain 60fps frame rate at good quality, video has to be scaled down to 640x480
  - *Nearly visually lossless transmission with 640x480 resolution at 60 fps is achieved at 6 Mbps*
- H.264 allows reasonable aerial video quality at about 0.04 bpp
  - *For entertainment style video reasonable quality requires 0.1 bpp*

Thank you

Contact us for more information:

akopansky@sarnoff.com

misnardi@sarnoff.com

achaudhry@sarnoff.com

bchai@sarnoff.com

pxiong@sarnoff.com

SARNOFF Corporation

Empowering Solutions Through Innovation