

High Quality Digital Video Processing: Technology and Methods

**IEEE Computer Society
Invited Presentation**

Dr. Jorge E. Caviedes
Principal Engineer
Digital Home Group
Intel Corporation

LEGAL INFORMATION

THIS MATERIAL AND THE INFORMATION PRESENTED ARE PROVIDED "AS IS". NO LICENSE, EXPRESS OR IMPLIED, BY ESTOPPEL OR OTHERWISE, TO ANY INTELLECTUAL PROPERTY RIGHTS IS GRANTED BY THIS MATERIAL. INTEL ASSUMES NO LIABILITY WHATSOEVER AND INTEL DISCLAIMS ANY EXPRESS OR IMPLIED WARRANTY, RELATING TO THE MATERIAL INCLUDING LIABILITY OR WARRANTIES RELATING TO FITNESS FOR A PARTICULAR PURPOSE, MERCHANTABILITY, OR INFRINGEMENT OF ANY PATENT, COPYRIGHT OR OTHER INTELLECTUAL PROPERTY RIGHT.

Support for some formats may require the customer to obtain license(s) from one or more third parties that may hold intellectual property rights applicable to the media format, decoding, encoding, transcoding, and/or digital rights management capabilities.

All products, dates and specifications are based on current expectations and subject to change without notice.

Intel, the Intel logo and Intel Atom are trademarks of Intel Corporation in the U.S. and other countries.

*Other names and brands may be claimed as the property of others.

Copyright © 2009 Intel Corporation.

Part I: Video Processing

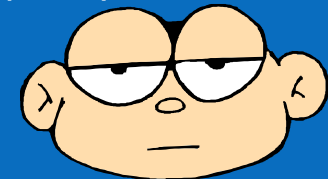
Summary

- Problem statement
- Video processing categories
- Corrective processing
- Conditioning
- Enhancement

Problem Statement: Interactive, Connected Multimedia Entertainment



Video processing is responsible for delivering content from any source to any display device with best visual quality



Satellite



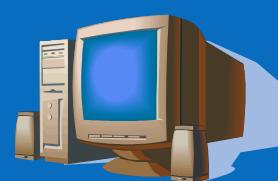
Cable



IPTV



User content

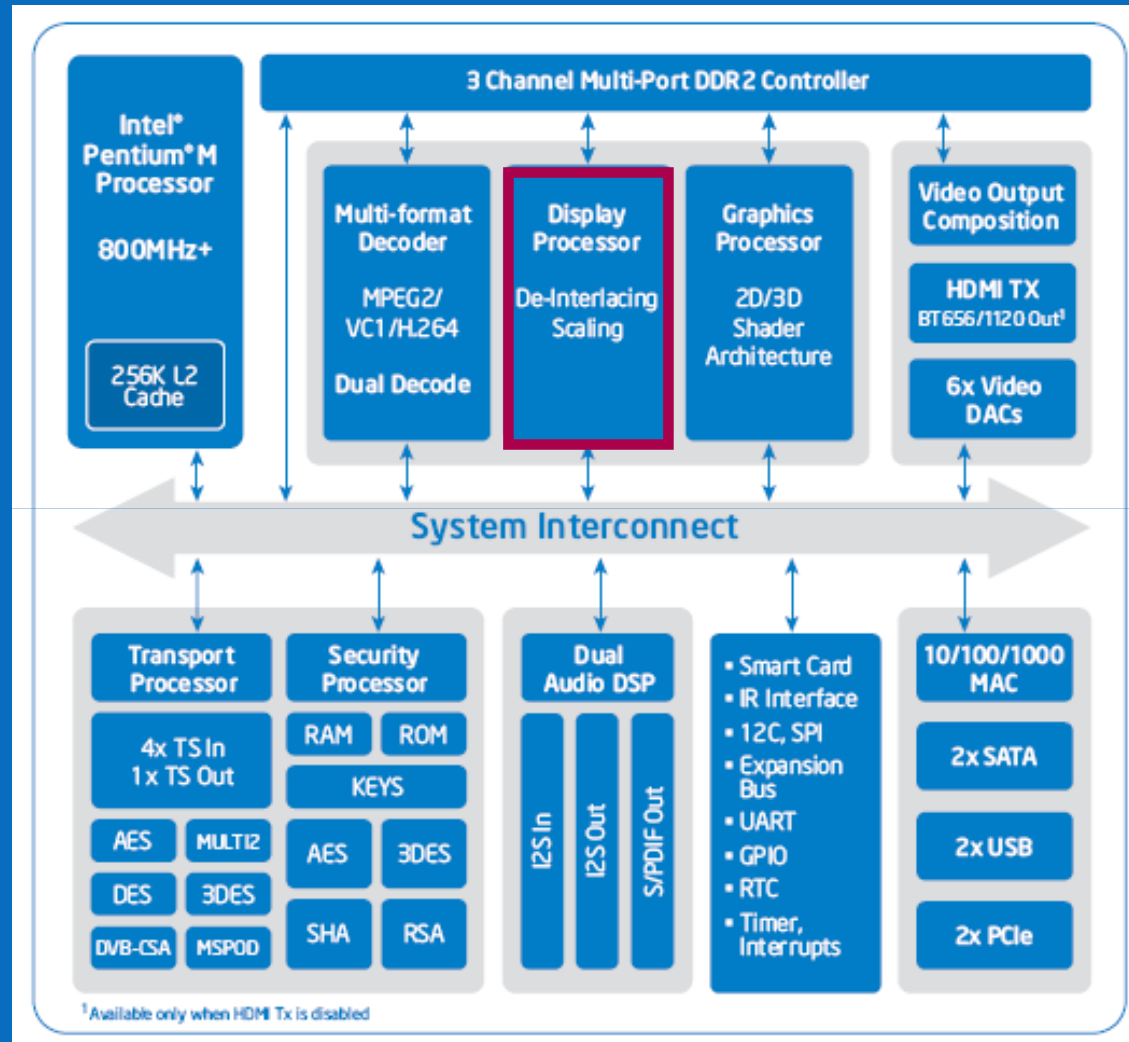


Intel® Media Processor CE3100

<http://download.intel.com/design/celect/downloads/ce3100-product-brief.pdf>

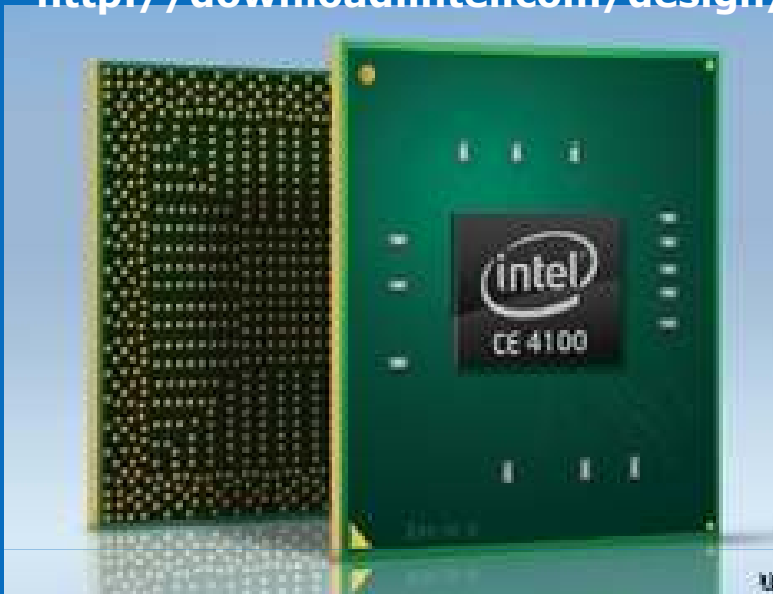


90nm SoC



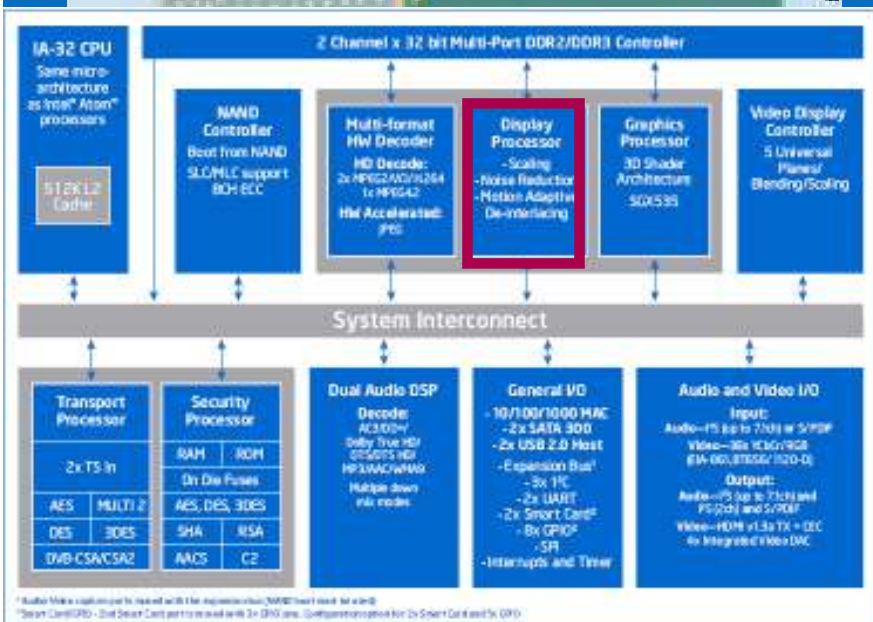
Intel® Atom™ Processor CE4100

<http://download.intel.com/design/celect/prodbrf/322572.pdf>



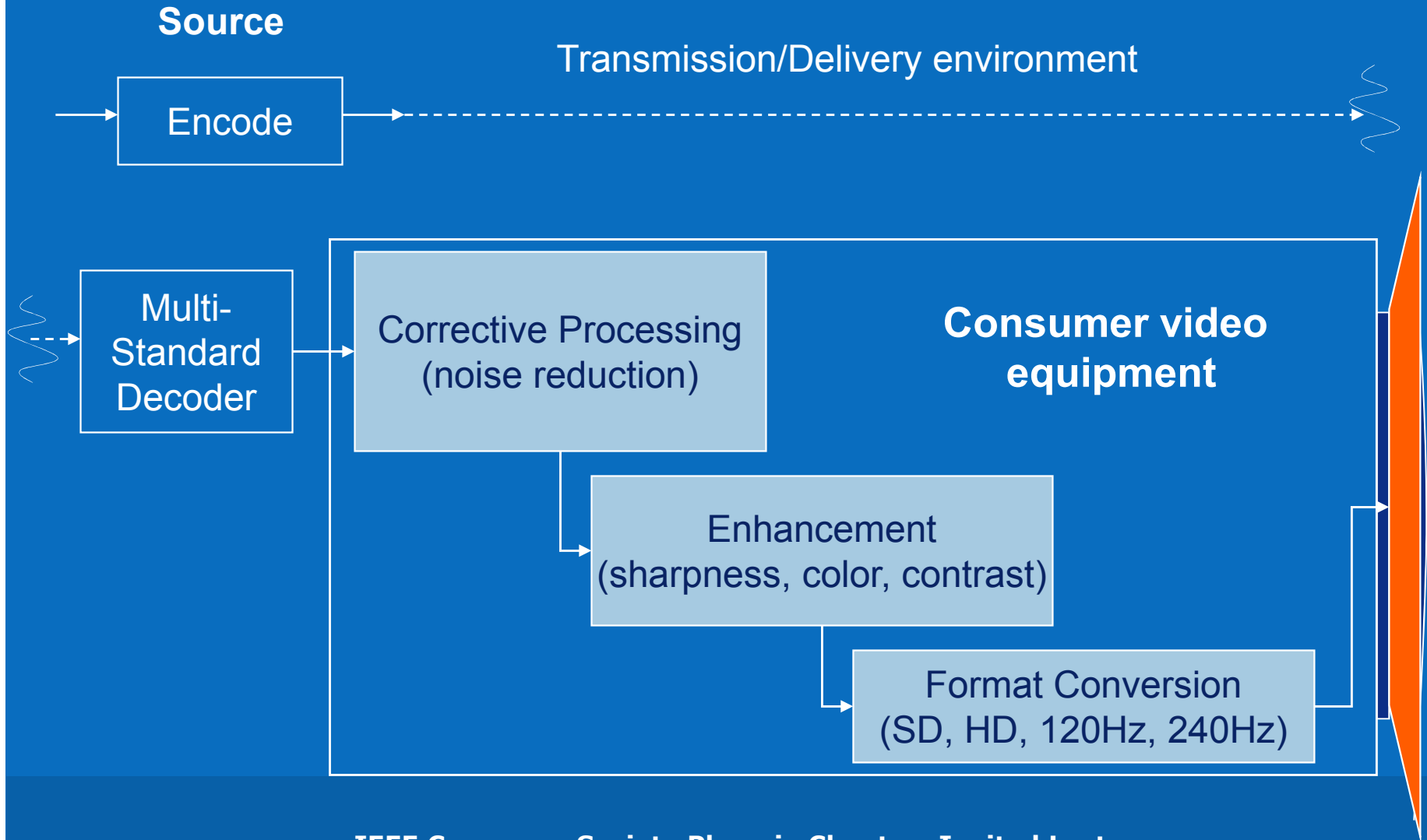
- Intel's latest CE4100 media processor is the latest generation, 45nm SoC with integrated processor and graphics controller
- Built on the low power Atom processor core, making it the ideal "brain" for set top boxes including cable boxes and Blu-ray players
- Capable of running at clock speeds up to 1.2GHz while featuring FSB speeds of 200MHz to 400MHz while supporting playback of 2 simultaneous 1080p video streams
- Supports H.264 video playback, 3D graphics and streaming media in Flash 10 format
- It does all that while consuming a mere 7 to 9 watts.

Support for some formats may require the customer to obtain license(s) from one or more third parties that may hold intellectual property rights applicable to the media format, decoding, encoding, transcoding, and/or digital rights management capabilities.



Phoenix Chapter- Invited Lecture

Video Processing Objectives



Correcting Analog and Digital Artifacts

MPEG Post-Processing (MPP)

- Deblocking
- Deringing
- Mosquito noise reduction

Analog Noise Reduction

- Gaussian noise reduction

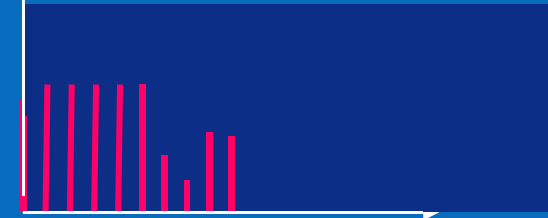
MPP: Frequency (DCT) Quantization



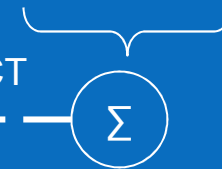
Block DCT



Frequency spectrum



Inverse DCT



MPEG Post-Processing is aimed at reducing the artifacts caused by DCT quantization, e.g. blocking, ringing, mosquito noise

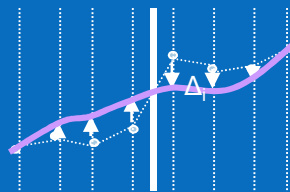
MPP: De-blocking



Original

$$p_i' = p_i + \Delta_i$$

$$q_i' = q_i + \Delta_i$$



Deblocked: pixels are aligned to reduce artifact

Deblocking Details

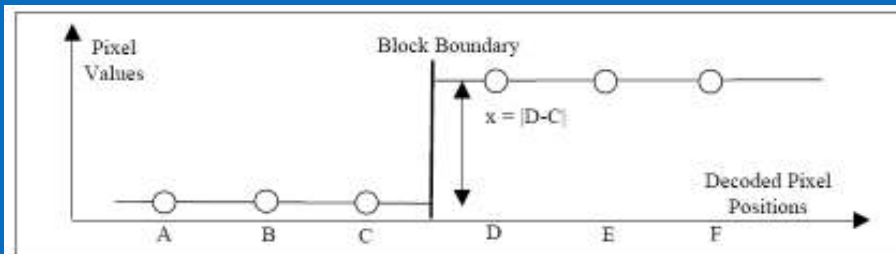


Figure 2 : One dimensional view of the block boundary after decoding

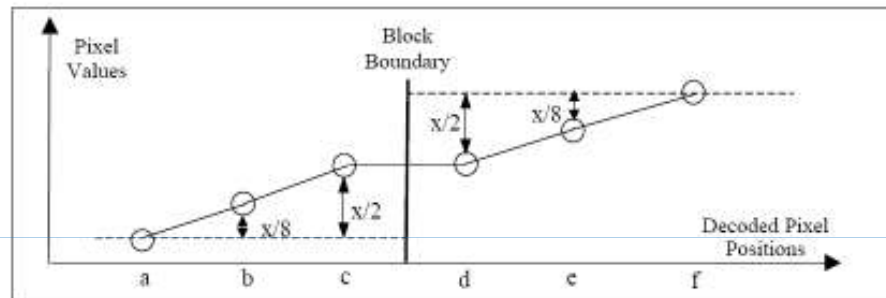


Figure 3 : One dimensional view of the block boundary after deblocking (Algorithm 1)

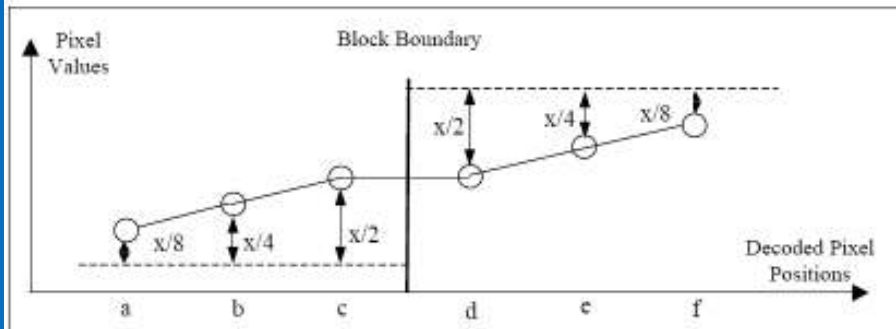


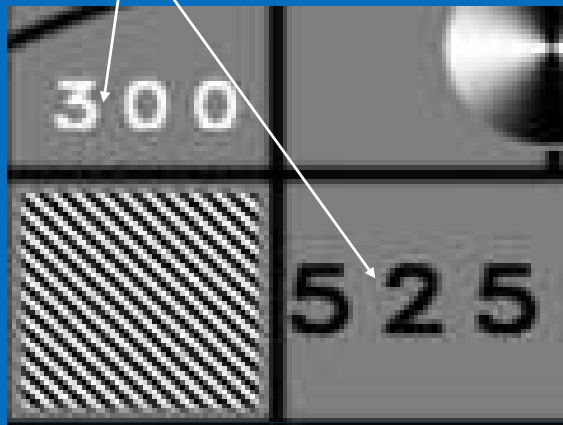
Figure 4 : One dimensional view of the block boundary after deblocking (strong filtering)

Filter strength depends on local content

[Ramkishor Korada and Pravin Karandikar, "Simple and Efficient Deblocking Algorithms for Low Bit-Rate Video Coding," IEEE International Symposium on Consumer Electronics, Hongkong, China, December 2000.]

MPP: De-ringing, Mosquito Noise Reduction Using non-linear, Adaptive Edge-preserving Filter

Ringing



Spatial filtering:
Apply adaptive
3x3 central
average filter



Spatio-Temporal
filtering:
Apply adaptive
3x3x2 S-T
median filter to
smooth regions



Mosquito noise

Gaussian (Analog) Noise Reduction:

Smooth region

σ^2 : noise variance

Lowpass Filter

Frequency Response of Box Filter

Frequency Response of Gaussian Filter

Method:

- (i) estimate noise variance in smooth regions and
- (ii) apply a filter matched to the noise characteristics without blurring the image.

Most practical solutions incorporate non-linear methods like central averaging and outlier exclusion.

Enhancing the Picture

Sharpness

- Peaking
- LTI/CTI (Luminance/Chrominance Transient Improvement)
- [*resolution and contrast affect sharpness perception*]

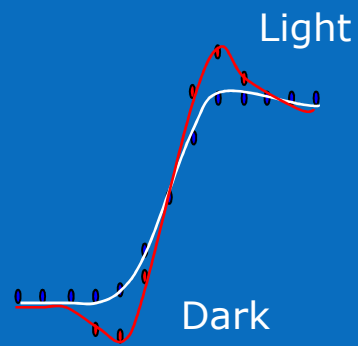
Color

- Skin, greens. blues

Contrast

- Adaptive contrast enhancement

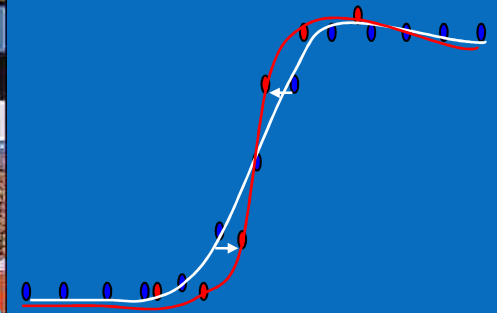
Sharpness Enhancement (1)



Peaking: add overshoot and undershoot to edge transitions



After enhancement



LTI/CTI: Make lightness/color transition steeper

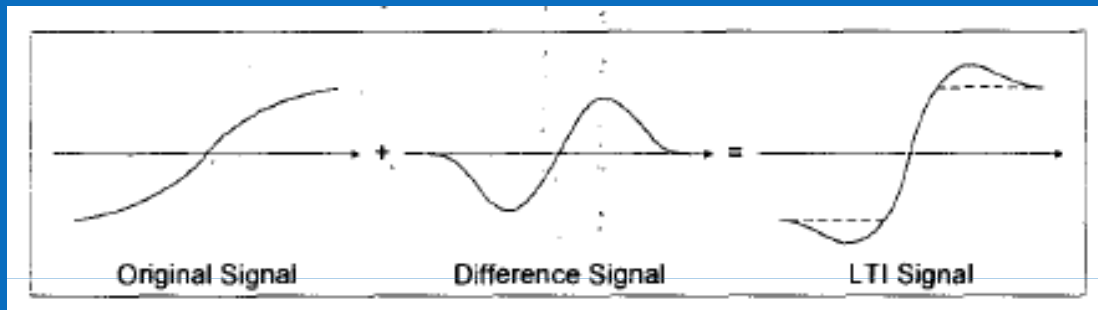
Edge transitions
Before ———
After ———

Sharpness Enhancement (2)

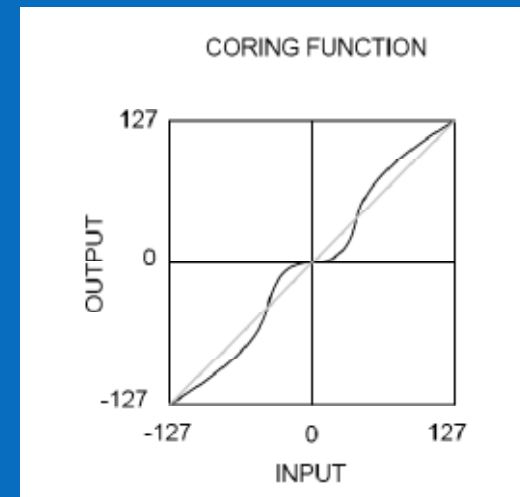
Before $\xrightarrow{\text{LTI/CTI}}$ After



LTI Details



Simple LTI approach



Coring used for noise reduction

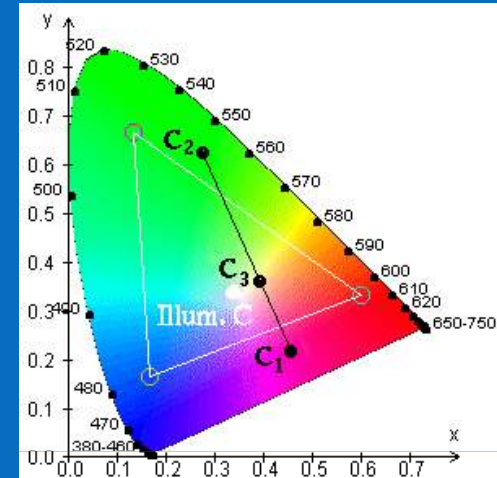
Color Spaces



RGB additive
color space



CMYK
subtractive color
space



Chromaticity Diagram

Emissive imaging systems (e.g. CRT) use additive colors, while absorptive systems use subtractive colors. The primary colors used by a system define a polygon or gamut in the chromaticity diagram.

Most color systems are defined by 3 primaries.

Color Space Conversion Example

ITU.BT-601 Y'CbCr International Standard Conversion:

$$Y' = 0.299 * R' + 0.587 * G' + 0.114 * B'$$

$$Cb = -0.169 * R' - 0.331 * G' + 0.500 * B'$$

$$Cr = 0.500 * R' - 0.419 * G' - 0.081 * B'$$

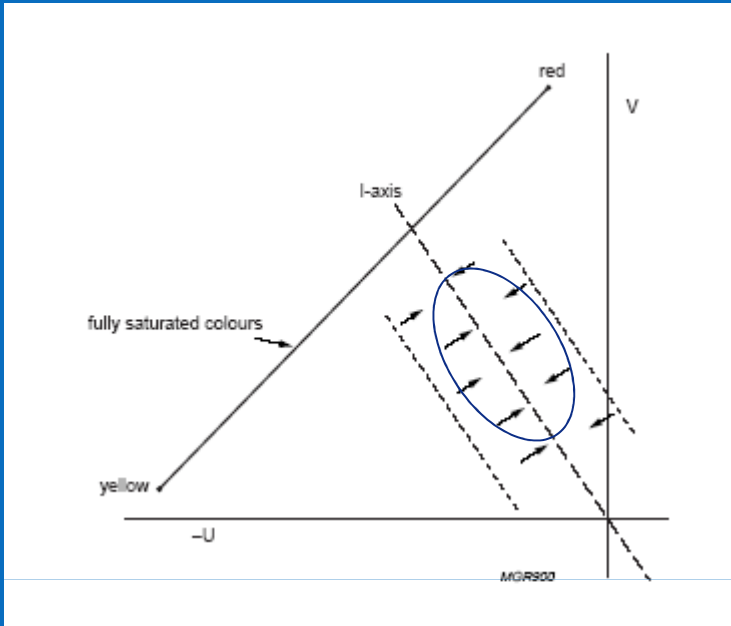
$$R' = Y' + 0.000 * U' + 1.403 * V'$$

$$G' = Y' - 0.344 * U' - 0.714 * V'$$

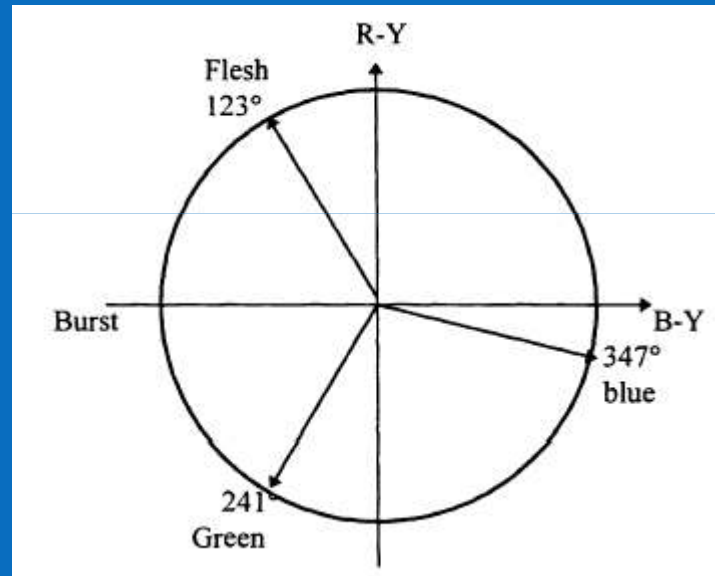
$$B' = Y' + 1.773 * U' + 0.000 * V'$$

RGB are CRT colors, YCrCb are DTV standard colors
(gamma correction indicated as primed values)

Color Enhancement (I)



In the U-V space, skin tones lay on the 123° line. Correction consists of bringing tones in the skin region closer to the flesh angle (123°)



Blue and green are also detected by the angle, and enhanced by increasing the amount of color



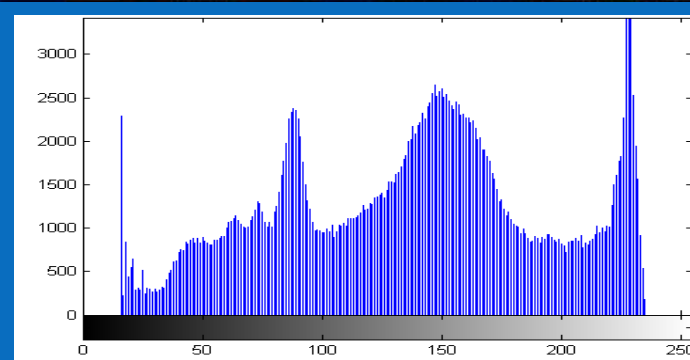
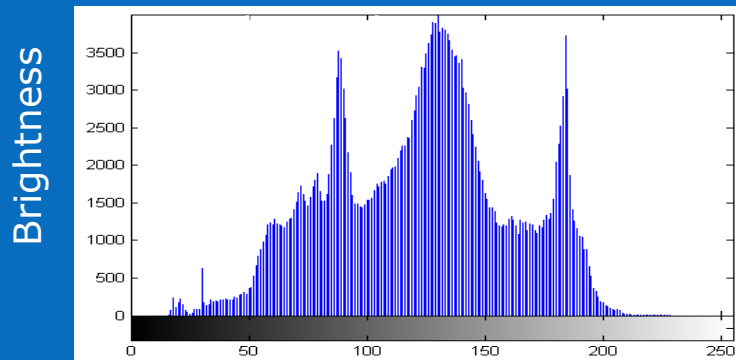
Color Enhancement (II)



Skin tone detected → Enhanced

Green color detected → Enhanced

Contrast Enhancement: Notice the change in luminance histograms



Combined Sharpness, Color and Contrast Enhancement



IEEE Consumer Society Phoenix Chapter- Invited Lecture

Format Conversion

Scaling (size and aspect ratio)

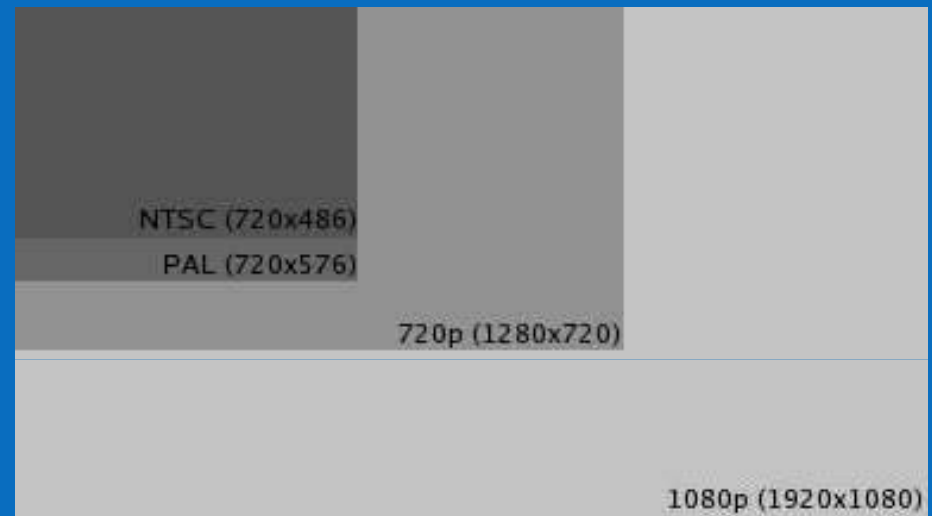
- SD to SD (different size)
- **SD to HD**
- HD to HD (different size)

Interlaced/Progressive

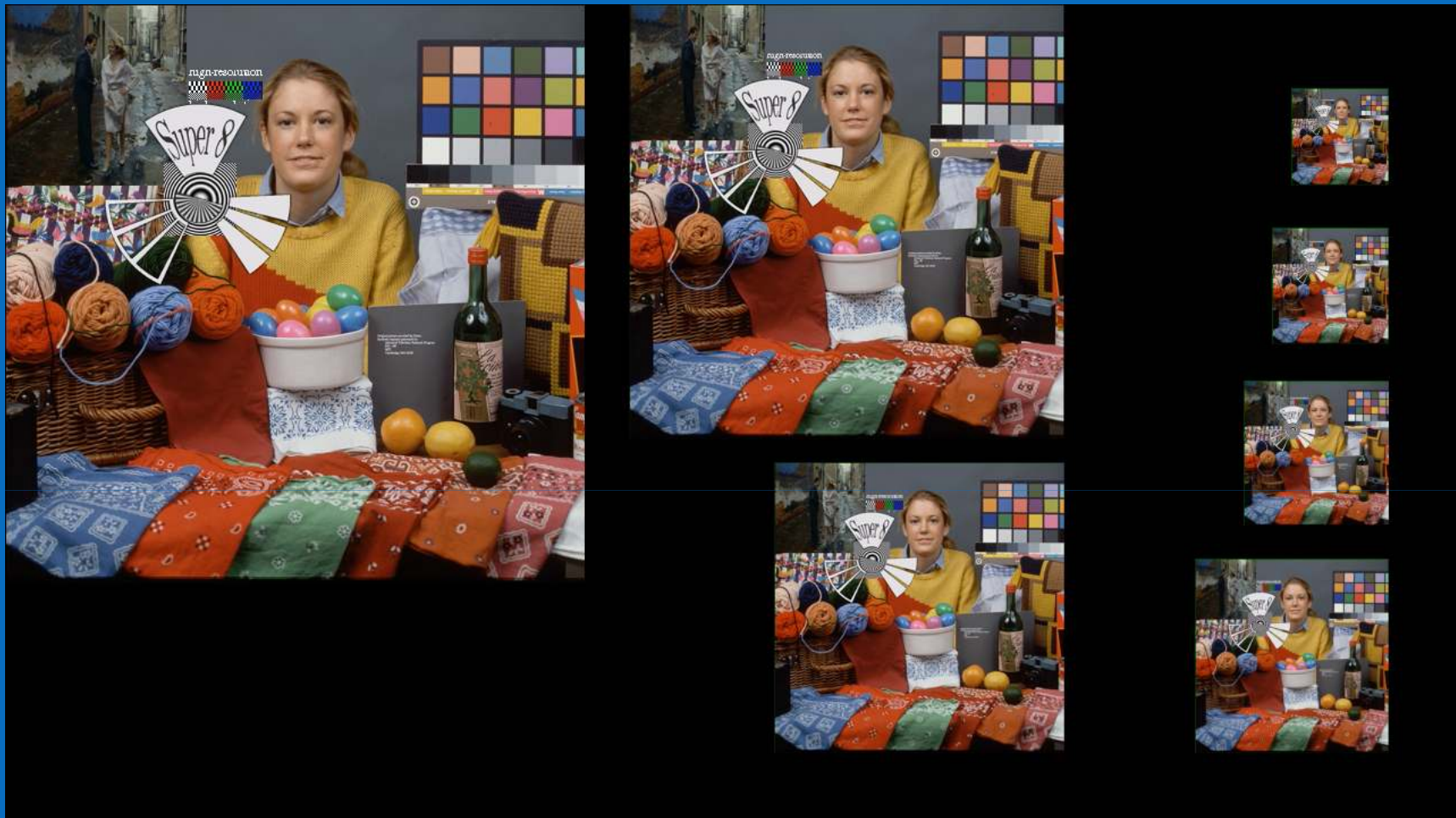
- Deinterlacing

Frame Rate

- Frame Rate Conversion



Down-Scaling for PIP



High quality down-scaling is important for PIP on large screens. Simple pixel dropping can be used but artifacts such as moiré may appear. Always need to lowpass in order to meet Nyquist sampling rate.

Up-scaling for Format Conversion SD->HD

4:3



Linear scaling



16:9



Non-linear
(anamorphic)
scaling



Scaling Methods

- Nearest neighbor interpolation
- Bilinear interpolation
- Bicubic interpolation

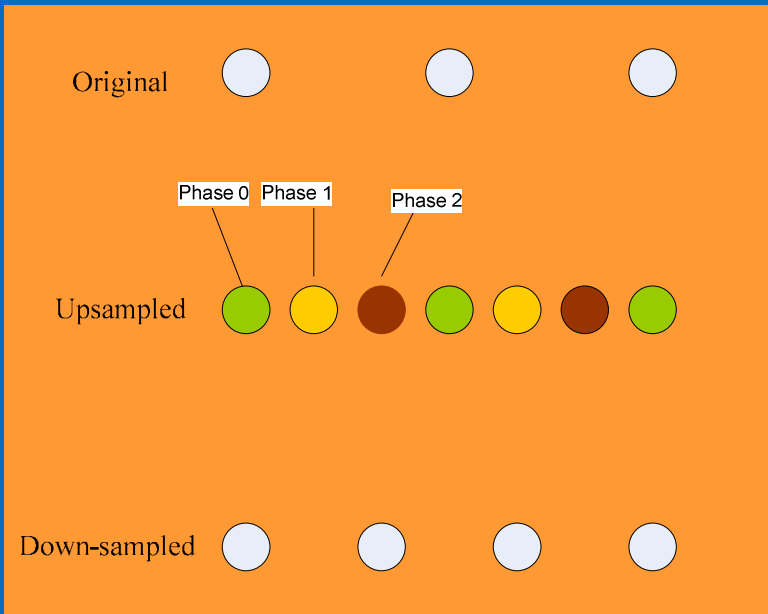
Commonly used implementation strategy:

- Polyphase interpolation

Advanced techniques:

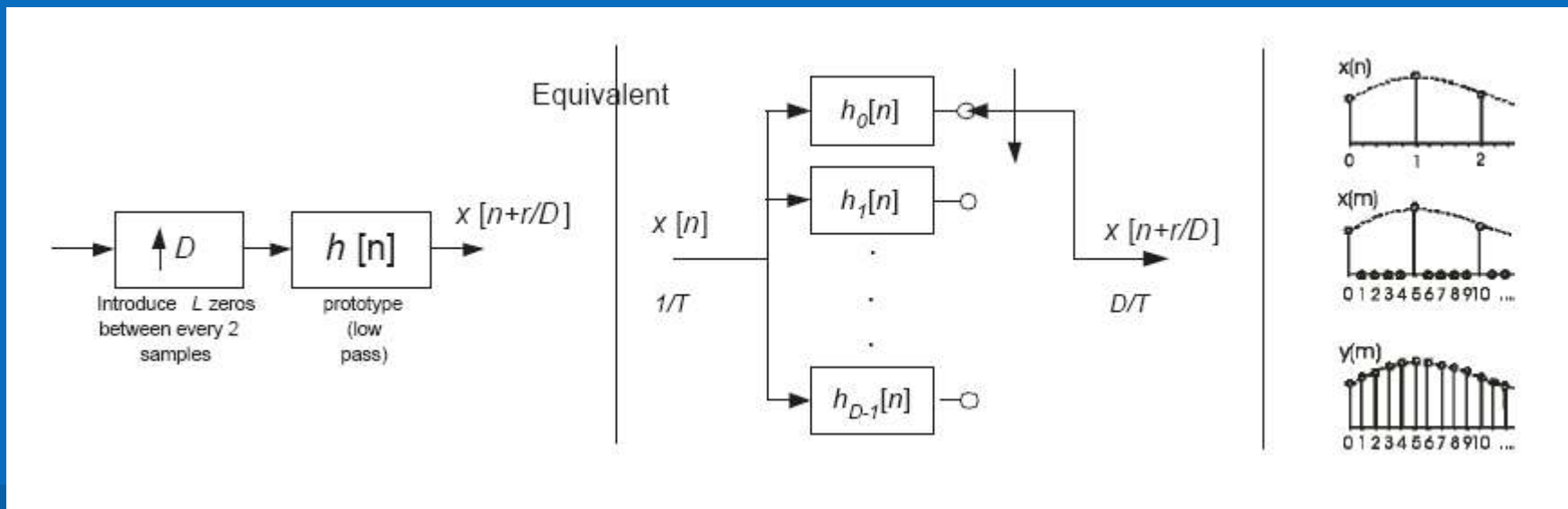
- Content-adaptive, non-linear scaling (e.g. EDI)
- Statistical up-conversion
- Super-resolution (up-scaling with resolution enhancement)

Scaling Methods



- Up-sampling (zero padding)
- Low-pass filtering
- Downsampling

Polyphase filter interpolation:



Extreme Scaling: Super Resolution

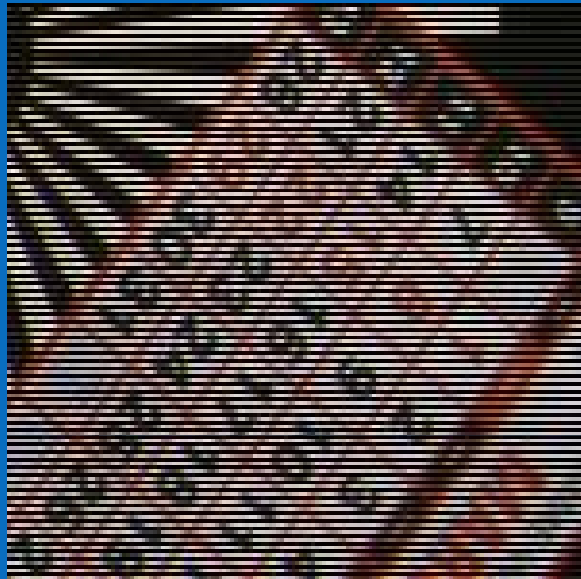


Low resolution (often sub-SD) images are not suitable for large LCD display.



Super-resolution computes pixels on a higher density sampling grid while reducing noise using multiple low-resolution frames as input.

Deinterlacing: from fields to frames



Odd field

Interlaced fields displayed one at a time

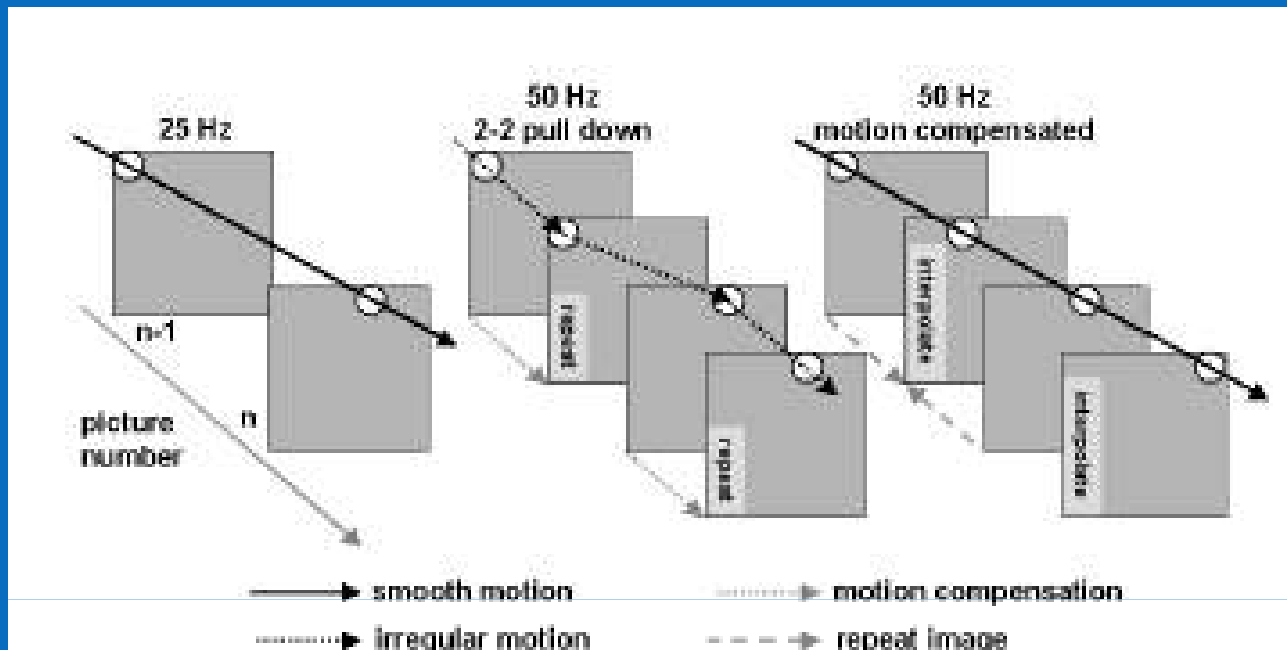


Frames can be made from merged (weaved) fields, but motion artifacts are visible



Advanced motion-adaptive and motion compensated deinterlacing are the most effective

Frame Rate Conversion (FRC)



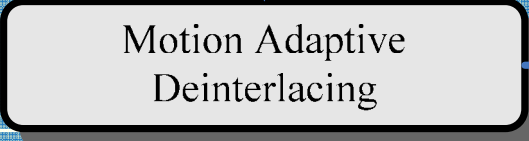
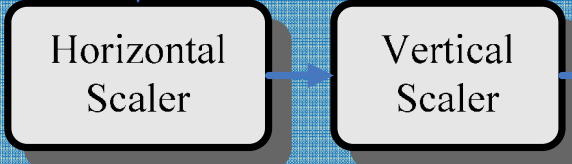



To display movie content shot at 24/25 Fps on a 50Hz (or 60Hz) TV it is necessary to do 2:2 (3:2) pull down.

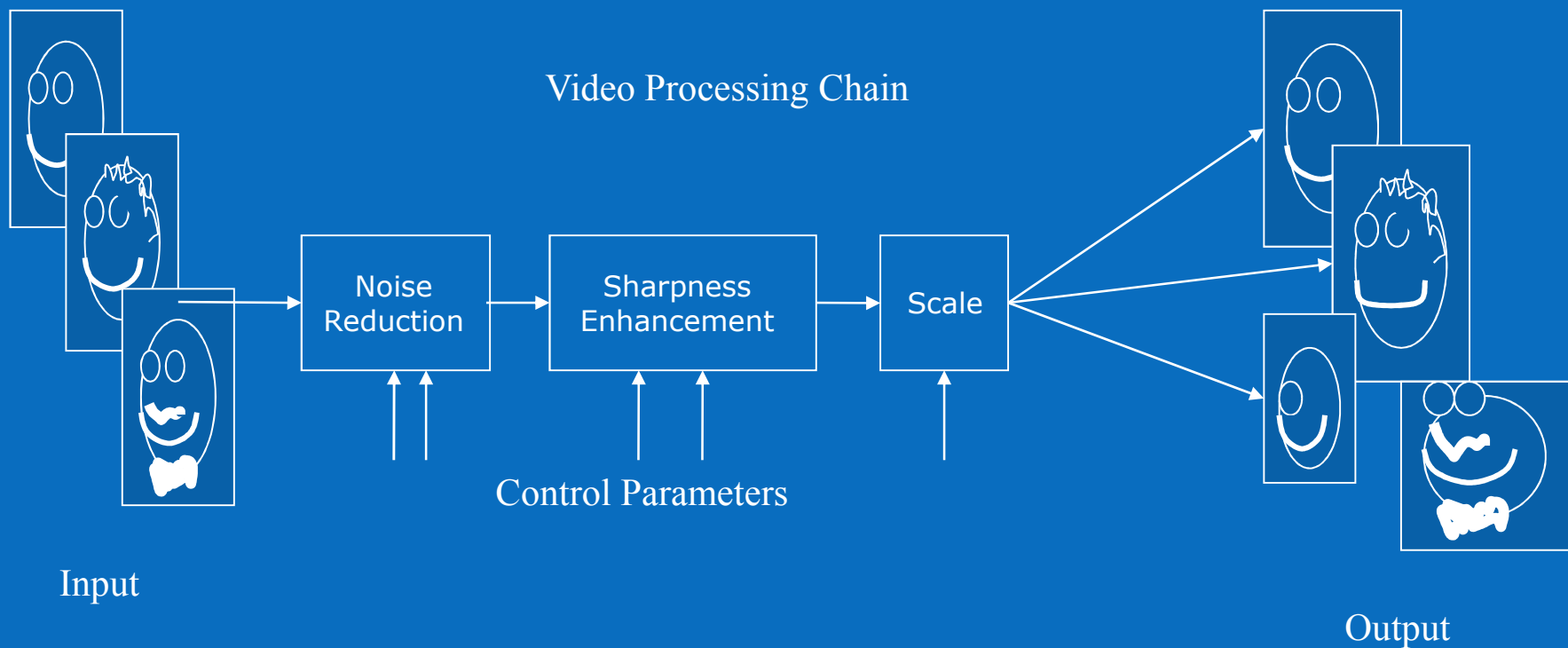
Possible methods are:

- *Picture repetition* -- gives irregular motion.
- *Motion adaptive temporal blend* -- intermediate, effective
- *Motion compensated FRC* -- gives smooth motion.

Video Processing Summary: State of the art and trends

Pipeline Modules	Description
 <pre> graph LR A[Deblocking] --> B[Dering] B --> C[Mosquito Noise Reduction] </pre>	<p>Compression artifacts tackled usually with multiple filters.</p> <p>Long term: generic noise filter.</p>
 <pre> graph LR A[Gaussian Noise Reduction] --> B[Sharpness Enhancement] </pre>	<p>Remove analog noise and improve sharpness.</p> <p>Long term: Joint sharpness-noise filter.</p>
 <pre> graph TD A[Motion Adaptive Deinterlacing] </pre>	<p>Highly competitive Motion Adaptive, and Motion Compensated solutions.</p> <p>Long term: superresolution or equivalent.</p>
 <pre> graph LR A[Horizontal Scaler] --> B[Vertical Scaler] </pre>	<p>High quality polyphase scalers. Advanced chroma upscaling (done separately) also possible.</p> <p>Long term: superresolution or equivalent.</p>
 <pre> graph LR A[Adaptive Contrast Enhancement] --> B[Color Enhancement] B --> C[Frame rate conversion] </pre>	<p>Skining, green, blue enhancement, ACE.</p> <p>Long term: total color management in CE video pipe.</p>

Tuning up the Video Chain



Present approach: trial and error
Long term: automated, driven by quality metrics

Part II: Visual Quality Optimization

Summary

- Quality scales
- Subjective and objective evaluation
- Video processing algorithm sequencing
- Algorithm interaction
- Perceptual interaction
- Picture quality optimization expertise and strategy
- Automation of picture quality optimization
- Quality evaluation cycle
- Research topics

Quality Scale: Perceived Difference and Preference

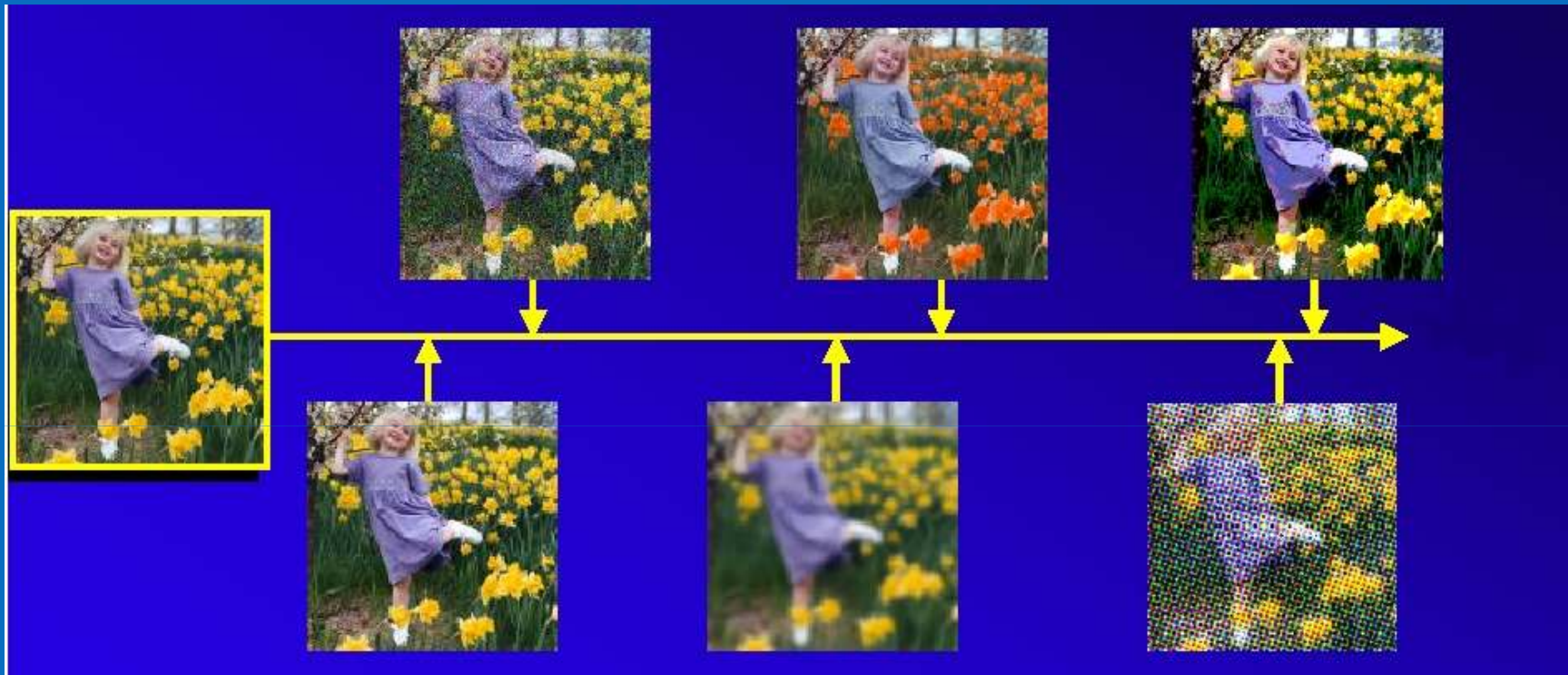
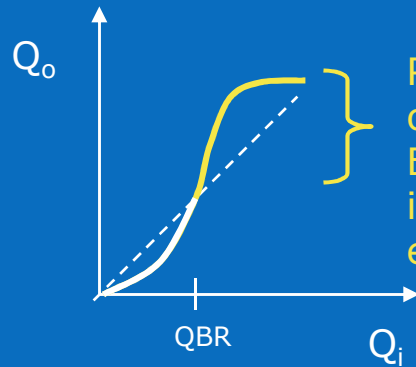
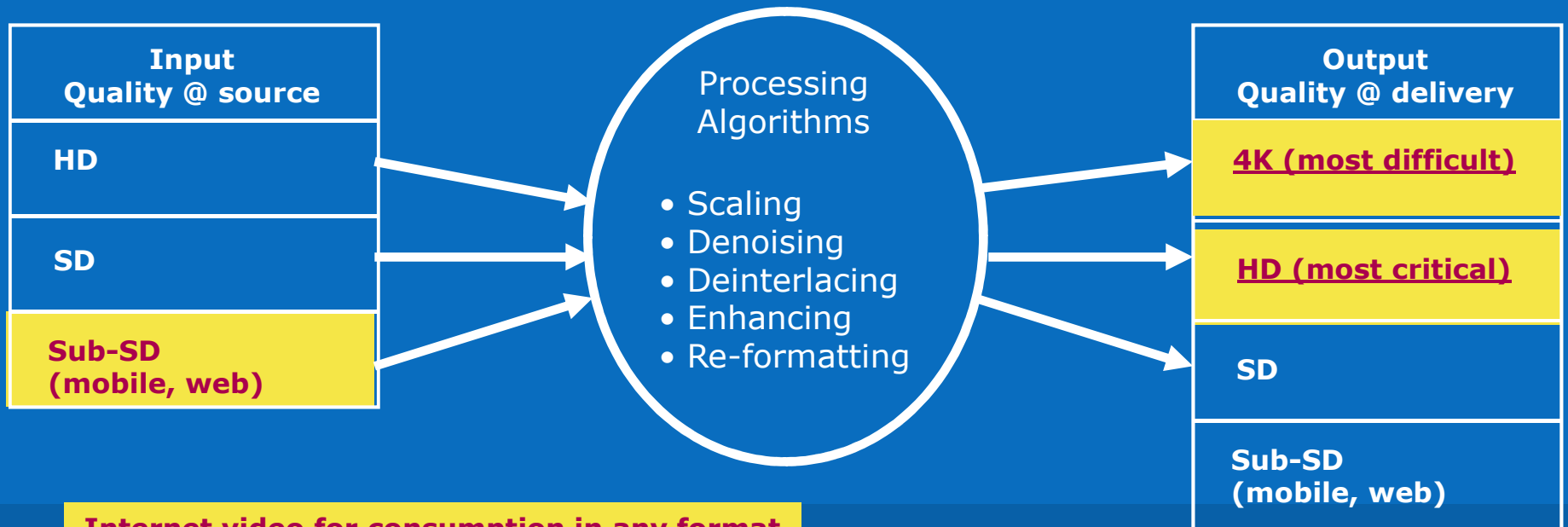


Image is modified as a result of processing or transmission prior to delivery to end user

Visual Quality and Processing Requirements

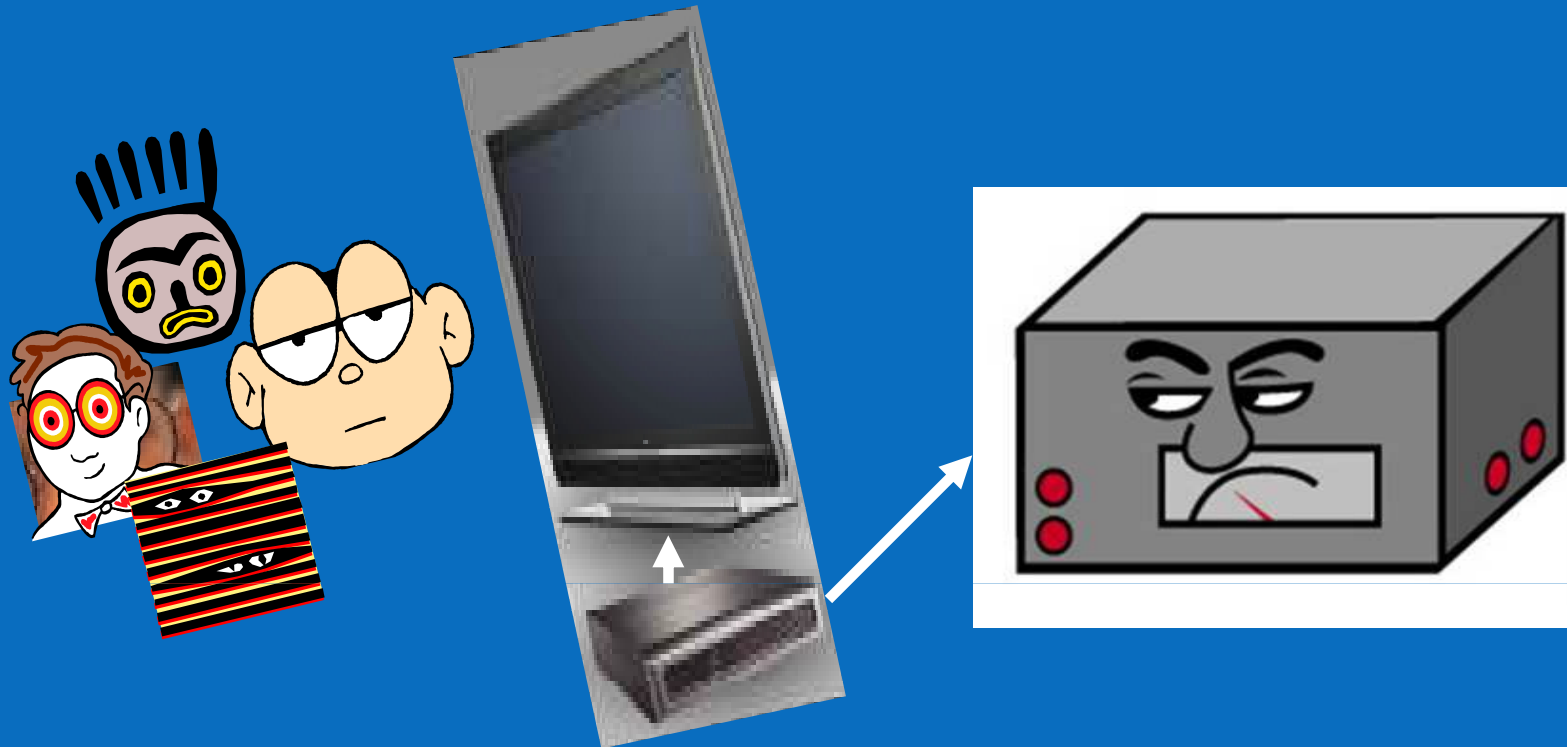


Processing from any format to any format creates opportunities for visual quality enhancement. Expected quality goes up with size and resolution, i.e. the larger the delivery format the higher the expected quality regardless of input.



Internet video for consumption in any format

Subjective and Objective Quality Assessment



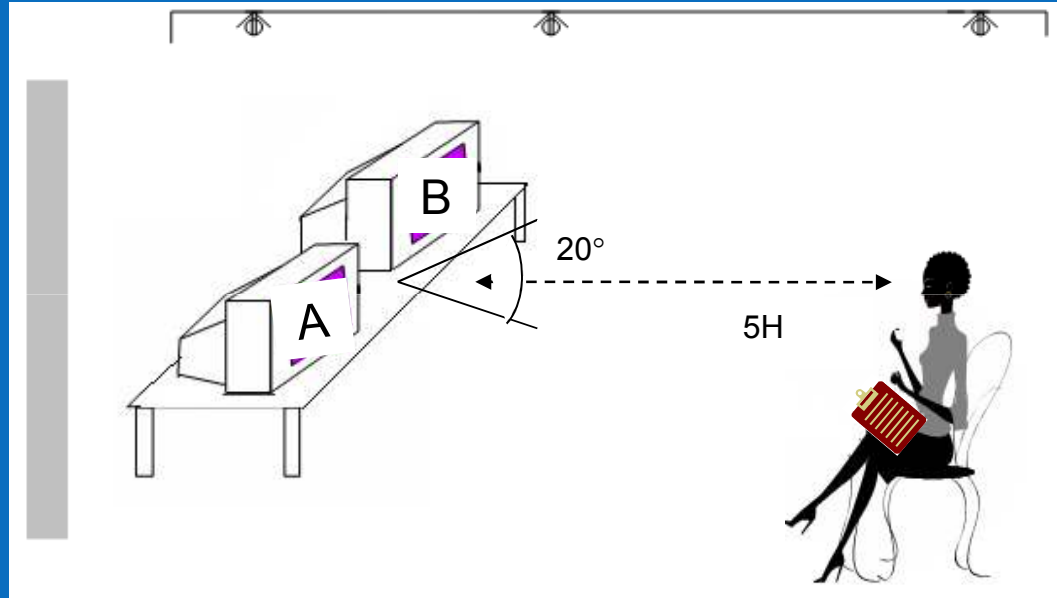
Subjective Assessments

- Statistical
- Uses Human Subjects
- Costly and time Consuming
- Most Reliable
- Expert & non-expert tests

No-Ref Objective Metrics

- Measure & Analyze Signal
- Analog & Digital Artifacts
- Image features, HVS modeling
- Fast and Cost Efficient
- Assume reference display

Subjective Quality Testing (double stimulus, comparative evaluation)



Standardized testing under controlled conditions for reliable, repeatable results.
Another modality is single-stimulus, continuous quality evaluation

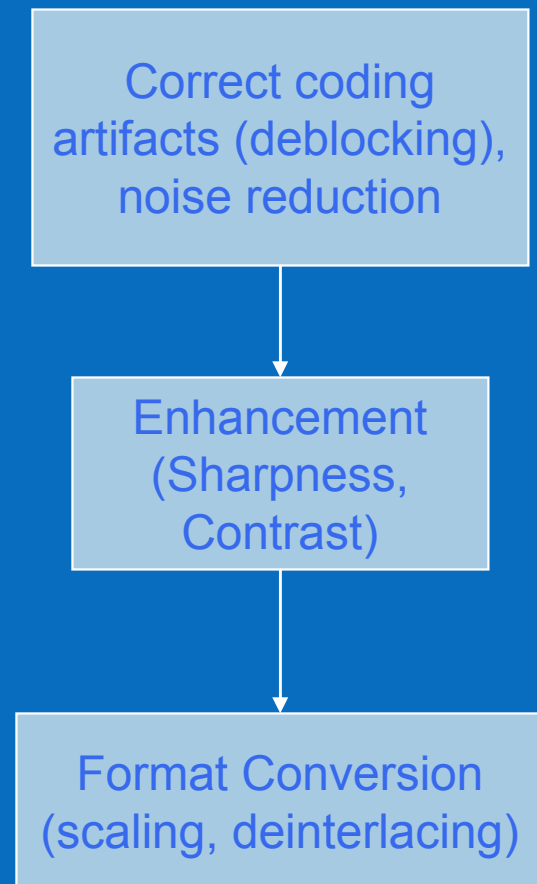
Optimizing Quality in the Design: Sequencing Principles

Correction precedes other processing types

Enhancement should follow, but is constrained by content losses in corrective processing

Format conversion is placed towards the end but, depending on the method, it may introduce blur (e.g. upscaling)

Post-formatting enhancement or correction may be necessary



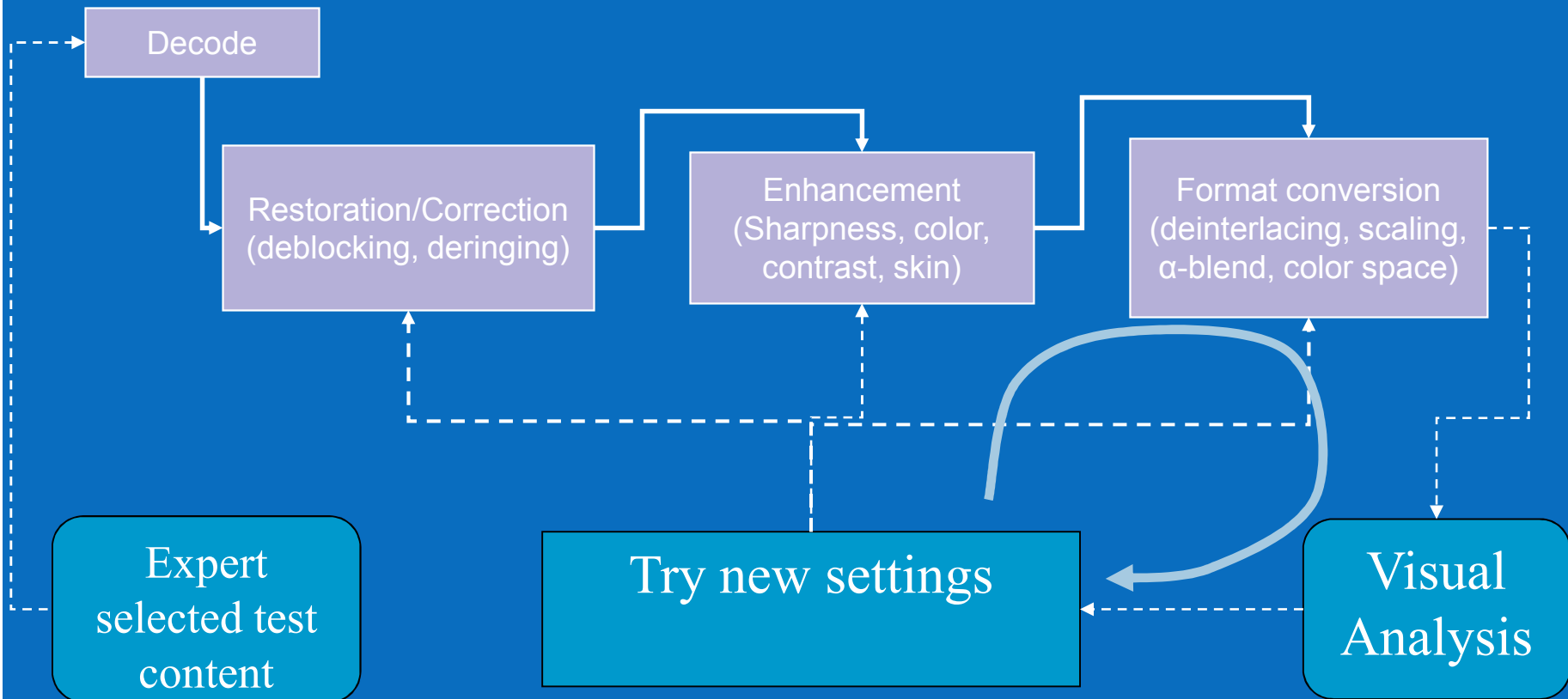
Algorithm interaction

- Sharpness enhancement and noise reduction
- Scaling and sharpness

Perceptual interaction

- Masking: sharpness and motion
- Masking: blockiness and ringing
- Facilitation: Sharpness, noise, contrast
- Mixed interaction: color, contrast, lightness

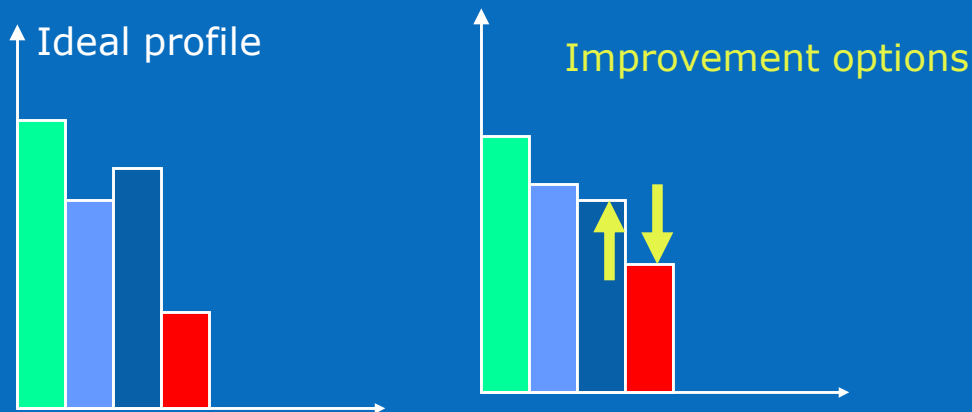
Picture quality Optimization: Expert Visual Analysis to Minimize Undesirable Features and Maximize Desirable Features



Trial and error is applied to tune-up for each block and for the entire video pipe

Intelligent Control: Content-Based Strategy

Natural scene case



- Sharpness
- Contrast
- Resolution
- Artifacts

A set of generic profiles (created by experts) would allow identification of enhancement potential and options

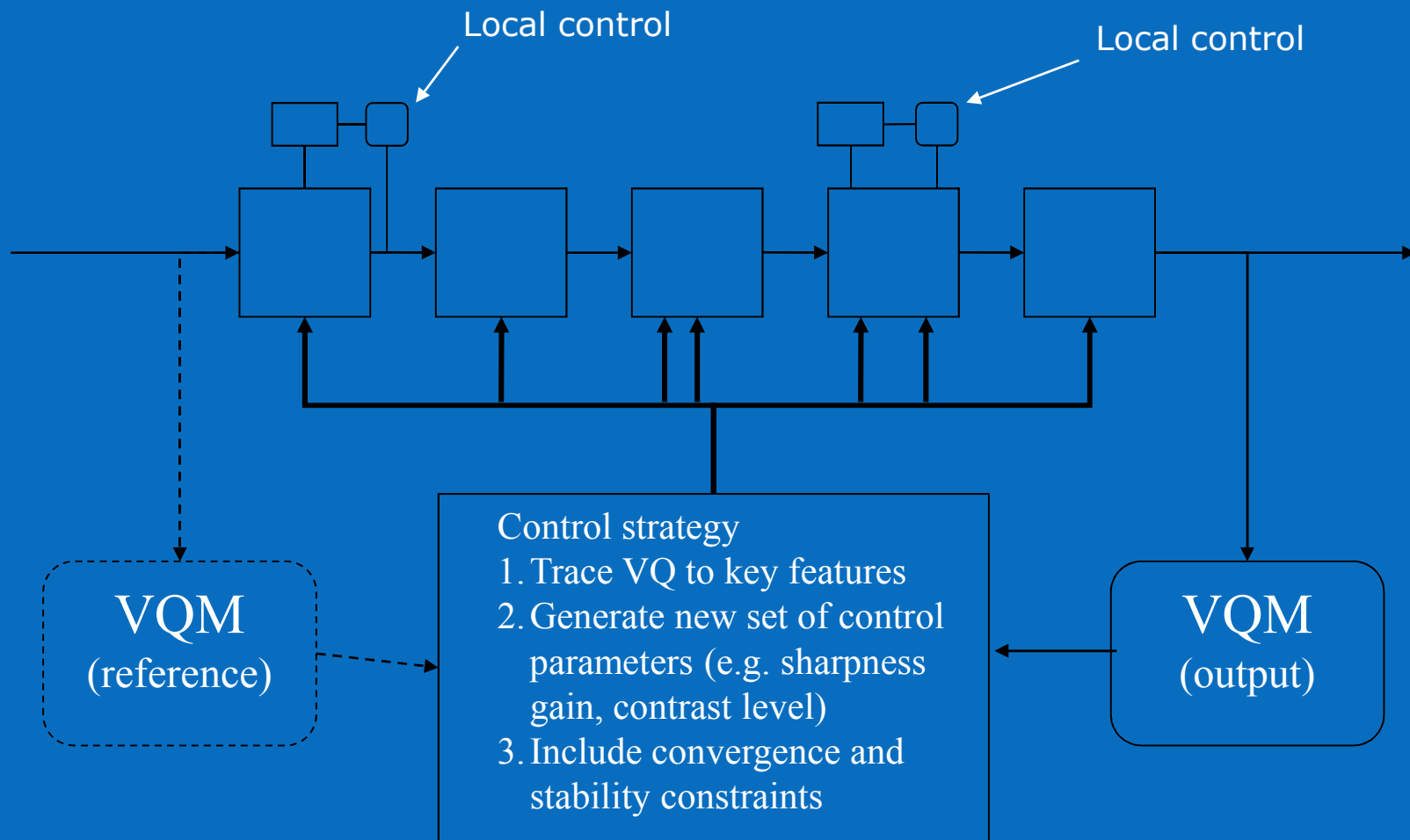
No-Reference Quality Metrics Design



Feature extraction computes the key inputs to the metric calculation. The NR metric is a perceptually calibrated computation. Features include:

- Edge pixels, gradients
- Contrast
- Artifacts (spatial, temporal)

Automated Picture quality Optimization



VQM: visual quality metric

Research Issues

Resolution-scalable NR metrics, to deal with all i/o formats

Perceptual calibration within and across metrics

Overall quality model (single vs. multiple dimensions)

Model scalability (local vs. global quality)

Modeling perceptual interactions (masking, facilitation)

Sensitivity analysis (per feature, display type)

Color, temporal quality metrics

Control system types suitable for real-time video processing

Questions?

Contact: jorge.e.caviedes@intel.com

IEEE Consumer Society Phoenix Chapter- Invited Lecture