Perceptual Quality Assessment and Processing for Visual Signals

Student: **Lin Ma**

Supervisor: **King Ngi Ngan**

Date: December 5, 2012

Outline

- Introduction & motivation
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

Introduction & motivation

- Perceptual quality assessment
  - Visual signal acquisition, compression, processing, transmission, communication, and so on.
- Two Approaches
  - Subjective evaluation
    - Most reliable
    - Time-consuming, expensive, viewer-dependent, and so on.
  - Objective evaluation
    - Full-reference (FR): watermarking, compression, and so on.
    - Reduced-reference (RR): quality monitoring, error concealment, and so on.
    - No-reference (NR): image/video denoising, enhancement, and so on.
    - New emerging forms of visual signals: retargeting image/video

Introduction & motivation

- Subjective quality database
  - Image:
    - LIVE, A57, IRCCy/IVC, MICT, TID2008, and so on.
  - Video:
    - LIVE, IVP, and so on.

	Image				Video
Databases	LIVE	IRCCyN/IVC	MICT	A57	LIVE
NOR	29	10	14	3	10
NDT	5	4	2	6	4
NOD	779	185	168	54	150
RES	768×512 or 512×768	512×512	512×512	512×512	768×432
SSF	DMOS	MOS	MOS	DMOS	DMOS
RNG	(0-100)	(1-5)	(1-5)	(0-1)	(0-100)

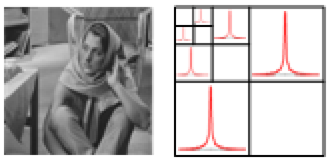
VQEG. (2000) Final Report From the Video Quality Experts Group on the Validation of Objective Models of Video Quality Assessment. [online] Available: http://www.vqeg.org.
H. R. Sheikh, M. F. Sabir, and A. C. Bovik, "A Statistical Evaluation of Recent Full Reference Image Quality Assessment Algorithms", IEEE Trans. Image Process., Vol. 15, pp. 3441-3452, 2006.

Outline

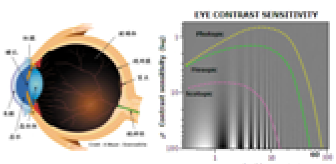
- Introduction & motivation
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

Overview of my research work

•DCT coefficient distribution modeling
•Video temporal information modeling



Statistical modeling



HVS property modeling

•HVS orientation sensitivity
•HVS saliency model
•HVS JND model
•HVS texture masking model

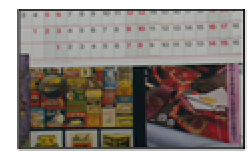
Perceptual quality assessment

- Full reference quality metric
- Reduced reference quality metric
- Retargeting image quality metric

Full reference quality metric:
- Image coding
- Video coding
- Consistent visual quality control

Reduced reference quality metric:
- Image restoration
- Image/video quality monitoring
- Video error resilience and concealment

Retargeting image quality metric:
- Image retargeting
- Video retargeting



Better quality of experiences (QoE) for users

Outline

- Introduction
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

P1-1. Full-reference image quality assessment

- HVS property modeling
  - Orientation sensitivity modeling
    - Horizontal Effect
      - For natural images:
        - » Oblique content is perceived to be the best;
        - » Horizontal content is the worst
      - For naturalistic broad-band stimuli:
        - » Oblique stimuli are perceived to be the best;
  - Saliency property

E. A. Essock et al., “Oblique Stimuli are seen best (not worst!) in naturalistic broad-band stimuli: a horizontal effect”, Vision Research, vol. 43, no. 12, pp. 1329-1335, Jun. 2003

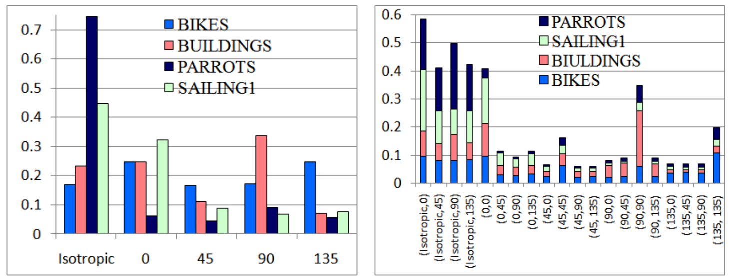
B. C. Hansen and E. A. Essock, “Influence of scale and orientation on the visual perception of natural scenes”, Vis. Cogn., vol. 12, no. 6, pp.1199-1234, Jun. 2003.

B. C. Hansen et al., “A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes”, J. Vis. Vol. 4, no. 12, pp. 1044-1060, 2004.

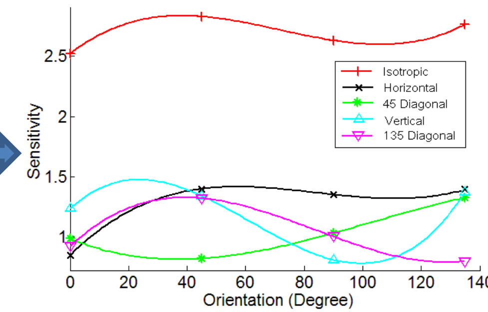
P1-1. Full-reference image quality assessment

- Visual horizontal effect modeling

Training Images



Statistical information

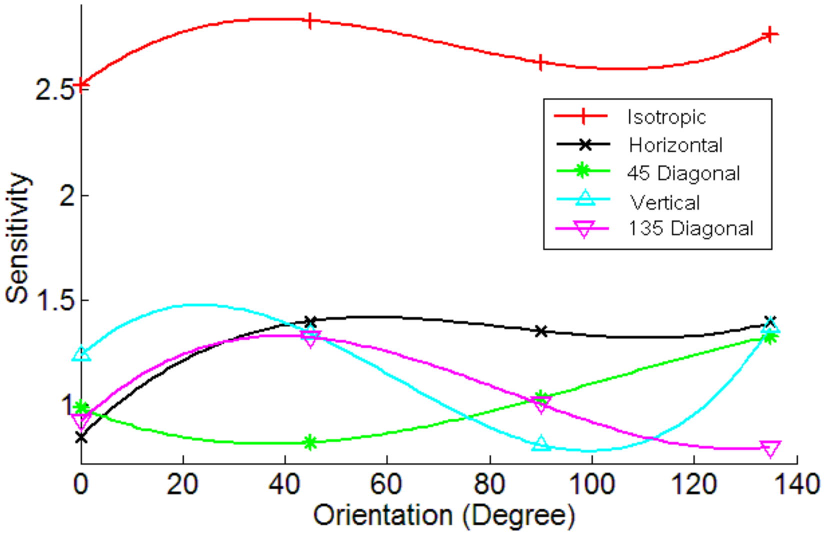


HVS HE Sensitivity Function

  - Content (reference image) and stimuli (distortion)
    - Orientation and energy
    - Gabor filtering
      - Maximum response determines the orientation and energy
      - Isotropic for some local smooth regions.

P1-1. Full-reference image quality assessment

•SSIM calculate structure distortion map;
•The HE sensitivity value is slightly modified by referring to the three aspects
•HE refined SSIM for IQA on the training database;
  ▪Performs better: tuning by following the same direction;
  ▪Otherwise: tuning by following the opposite direction;
•After several iterations, optimized HE sensitivity values are obtained.



$$S_{HE} = \varphi(\theta_I, \theta_S) = a_{\theta_I}\theta_S^3 + b_{\theta_I}\theta_S^2 + c_{\theta_I}\theta_S + d_{\theta_I}$$

E. A. Essock et al., "Oblique Stimuli are seen best (not worst!) in naturalistic broad-band stimuli: a horizontal effect", Vision Research, vol. 43, no. 12, pp. 1329-1335, Jun. 2003
B. C. Hansen and E. A. Essock, "Influence of scale and orientation on the visual perception of natural scenes", Vis. Cogn., vol. 12, no. 6, pp.1199-1234, Jun. 2003.
B. C. Hansen et al., "A horizontal bias in human visual processing of orientation and its correspondence to the structural components of natural scenes", J. Vis. Vol. 4, no. 12, pp. 1044-1060, 2004.

P1-1. Full-reference image quality assessment

- HVS saliency
  - HVS processes local regions with different visual acuities.
  - Spectral residual model is employed to detect the saliency:
    - Spectral residual R(f):

$$R(f) = L(f) - L_a(f)$$

    - Spectral residual contains important information related the HVS perception:

$$Sa_M = |\zeta^{-1}(\exp(R(f) + jP(f)))|^2$$

X. Hou and L. Zhnag, “Saliency Detection: A Spectral Residual Approach”, Proc. CVPR, 2007.

P1-1. Full-reference image quality assessment

- SSIM is performed to obtain the structural distortion map *SD*
- The HE refined structural distortion map *SMr*:

$$SM_r(i,j) = SD(i,j)\cdot\alpha_{SE}(i,j)/S_{HE}(\theta_I(i,j),\theta_S(i,j))$$

  - Horizontal effect modeling
  - Stimulus energy adaptation factor α_{SE}
- Saliency pooling

$$Index = \sum Sa_M \cdot SM_r \Big/ \sum Sa_M$$

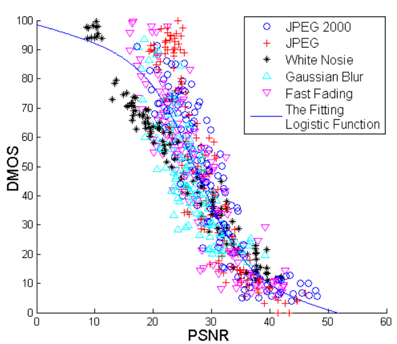
P1-1. Full-reference image quality assessment

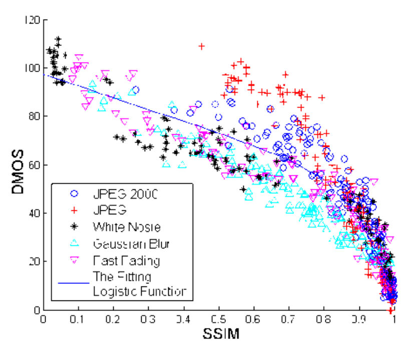
- Proposed method compares with
  - PSNR, S-SSIM, LDW-SSIM, ICW-SSIM, SMW-SSIM, VIF.
- Based on LIVE and A57 image database
- Standard evaluation procedure
  - Larger LCC and SROCC, better performance
  - Smaller RMSE and MAE, better performance

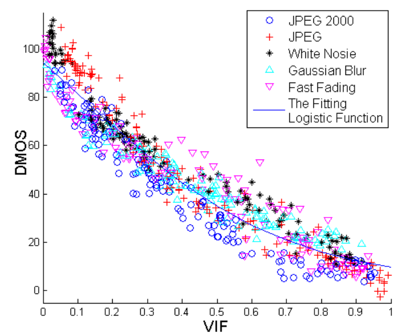
P1-1. Full-reference image quality assessment

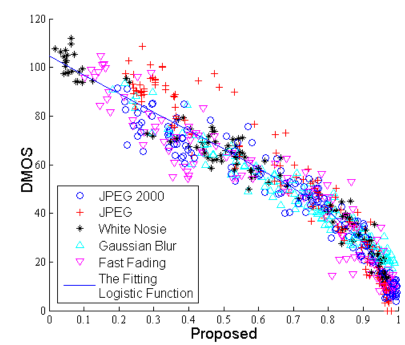
- Experimental results

		PSNR	S-SSIM	LDW-SSIM	ICW-SSIM	SMW-SSIM	VIF	Proposed
LIVE Database	CC	0.891	0.914	0.915	0.936	0.947	0.961	0.966
	SROCC	0.897	0.922	0.919	0.942	0.953	0.966	0.971
	RMSE	12.425	11.060	11.051	9.641	8.769	7.523	7.057
	MAE	9.765	8.727	8.655	7.478	6.921	6.043	5.374
A57 Database	CC	0.644	0.415	0.545	0.518	0.607	0.614	0.848
	SROCC	0.570	0.407	0.495	0.455	0.557	0.622	0.857
	RMSE	0.192	0.224	0.206	0.210	0.195	0.194	0.130
	MAE	0.151	0.186	0.160	0.166	0.145	0.141	0.102







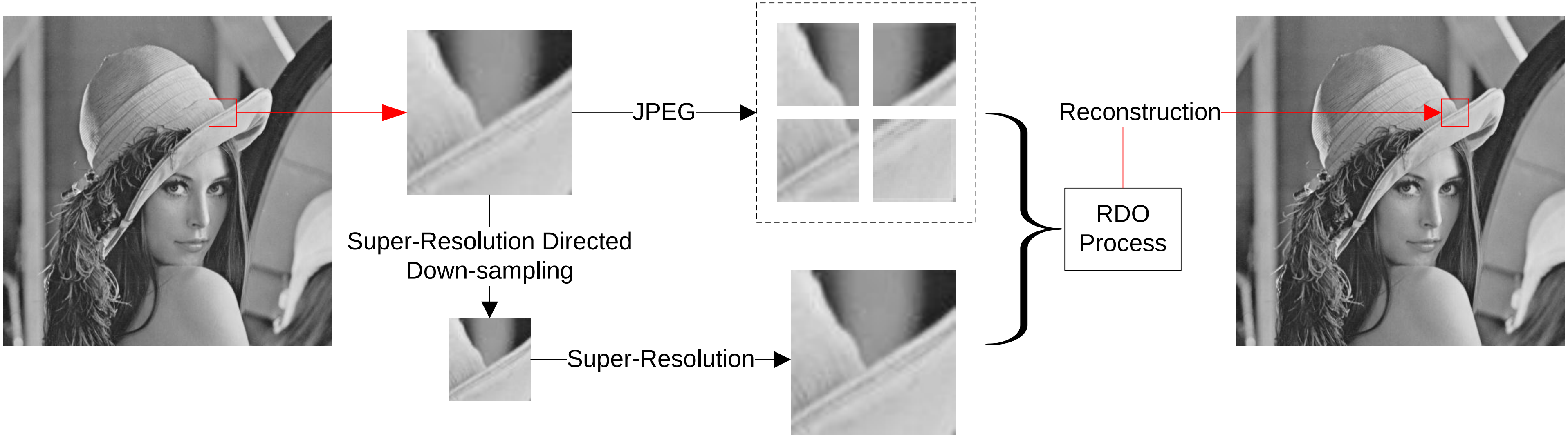


P1-2. Perceptual image compression by SRDDS

- Image compression via adaptive block-based super-resolution directed down-sampling (SRDDS).
- Encoder side
  - RDO determines the encode process at original or down-sampled resolution
    - MSE and SSIM depicts the perceptual distortion
  - Down-sampling directed by super-resolution method
    - Minimize the reconstruction error.
- Decoder side
    - Super-resolution in DCT domain together with inverse DCT.
    - Reduce complexity.

Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity", *IEEE Trans. Image Process.*, vol. 13, no.4, pp.600-612, 2004.

P1-2. Perceptual image compression by SRDDS



- For each 16×16 macroblock
  - Traditional JPEG (four 8×8 blocks)
  - Super-resolution directed down-sampling (SRDDS)
    - Down-sample marcoblock to 8×8 block
    - QP value is set as half of the JPEG mode

P1-2. Perceptual image compression by SRDDS

- Super-resolution in DCT domain:

$N \times N$ DCT coefficients $Coef_{N\times N}$ (N is equal to 8)

$Coef_{N\times N}$ extended to $Coef_{2N\times 2N}$ by

$$Coef_{2N\times 2N} = \begin{bmatrix} Coef_{N\times N} & 0_{N\times N} \\ 0_{N\times N} & 0_{N\times N} \end{bmatrix}$$

Inverse DCT is applied to $Coef_{2N\times 2N}$

$$P_{2N\times 2N} = D^T_{2N\times 2N} \times (Coef_{2N\times 2N}) \times D_{2N\times 2N}$$

$D_{2N\times 2N}$ is the DCT kernel for $2N$ samples

$$\begin{aligned} P_{2N\times 2N} &= D^T_{2N\times 2N} \times \begin{bmatrix} Coef_{N\times N} & 0_{N\times N} \\ 0_{N\times N} & 0_{N\times N} \end{bmatrix} \times D_{2N\times 2N} \\ &= D^T_{2N\times 2N} \times \begin{bmatrix} D_{N\times N} \times b_{N\times N} \times D^T_{N\times N} & 0_{N\times N} \\ 0_{N\times N} & 0_{N\times N} \end{bmatrix} \times D_{2N\times 2N} \\ &= (D^T_{N\times N} \times D_{N\times 2N})^T \times b_{N\times N} \times (D^T_{N\times N} \times D_{N\times 2N}) \\ &= V_{2N\times N} \times b_{N\times N} \times H_{N\times 2N} \end{aligned}$$

I. Shin, and H. W. Park, "Adaptive up-sampling method using DCT for spatial scalability of scalable video coding", *IEEE Trans. Circuits Syst. Video Technol.*, vol. 19, no. 2, pp. 206-214. Feb. 2009.

P1-2. Perceptual image compression by SRDDS

- The optimal low-resolution block in the encoder side is obtained by minimizing the reconstruction error:

$$\hat{b}_{N\times N} = \arg\min_{b}\{\|P_{2N\times 2N} - S(b_{N\times N})\|_2^2\}$$

$$\hat{b} = \arg\min_{b}\{\|P - VbH\|_F\}$$

- Kronecker product is employed to solve the problem:

$$\hat{b}_v = \arg\min_{b_v}\{\|(V \otimes H^T)b_v - P_v\|_2^2\}$$

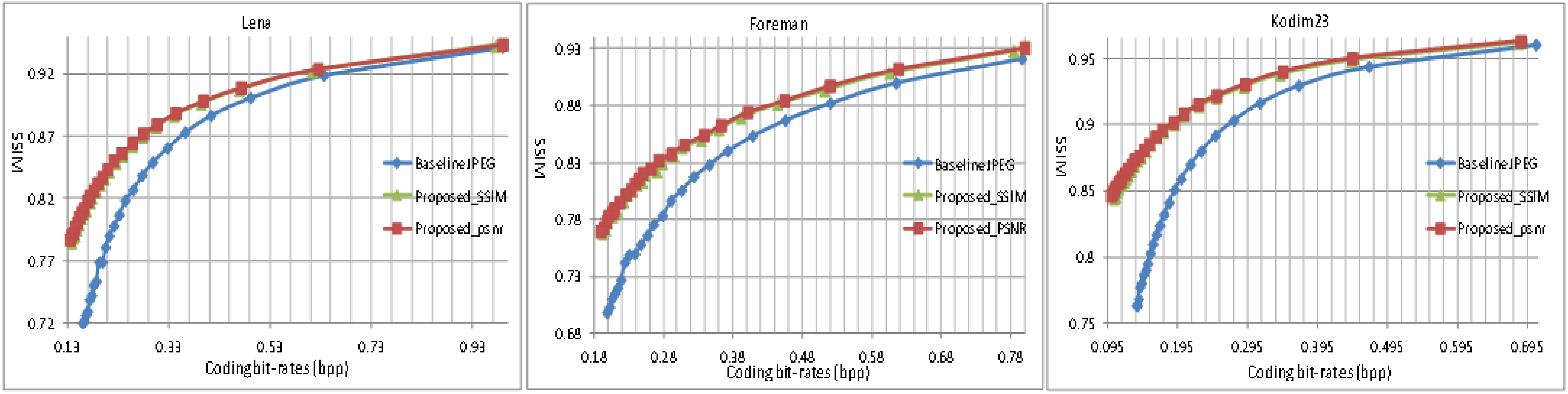
- The optimized low-resolution block:

$$\hat{b}_v = (M^T M)^{-1} M^T P_v$$

where $M = (V \otimes H^T)$

P1-2. Perceptual image compression by SRDDS

- 1-bit flag for each macroblock indicates the mode used:
  - Traditional JPEG for each 8×8 block
  - SRDDS mode for the low-resolution block



P1-2. Perceptual image compression by SRDDS

- Experimental results

bpp=0.1315, SSIM=0.75

bpp=0.2285, SSIM=0.6564

bpp=0.2122, SSIM=0.79

bpp=0.1308, SSIM=0.87

bpp=0.2264, SSIM=0.74

bpp=0.2115, SSIM=0.84

P1. Full-reference image quality assessment

- Contribution:
  - HVS orientation sensitivity modeling
    - Visual HE modeling
    - Correlate better with HVS perception
  - Perceptual image compression by SRDDS
    - Super-resolution directed down-sampling (SRDDS)
    - Low-complexity
    - Better visual quality of coded image

Outline

- Introduction & motivation
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

P2. Full-reference video quality assessment

- Just noticeable difference (JND)
  - Smallest detectable difference between a starting and secondary level of a particular sensory stimulus in psychophysics.
  - JND for video consists of spatial and temporal components

$$T_{JND} = T_{JND_s} \times T_{JND_T}$$

- Spatial JND
  - Basic JND value generated from HVS contrast sensitivity function (CSF)
  - Luminance adaptation
  - Contrast masking

$$T_{JND_s} = T_{basic} \cdot \alpha_{lum} \cdot \alpha_{cm}$$

- Temporal CSF is related to spatial frequency and the velocity of object motion on the retina plane.

P2. Full-reference video quality assessment

- Adaptive block-size transform (ABT) is proposed for video coding
  - Energy compaction
    - Larger block transform → exploit the correlation
    - Smaller block transform → adapt to local structural changes
  - Subjective benefits
    - Texture preservation of HD movie
    - Keep film details and grain noise

- Psychophysical experiment is carried out to obtain two different JNDs based on 8×8 and 16×16 DCT.

Y. Huh et al. "Variable Block Size Coding of Images with Hybrid Quantization", IEEE CSVT, Vol. 6, pp. 679-685, 1996.
C. Zhang et al. "The Technique of Prescaled Integer Transform: Concept, Design and Applications", IEEE CSVT, Vol. 18, pp. 84-97, 2008

P2-1. Full-reference video quality assessment

- Proposed adaptive block-size transform (ABT) JND
  - Determine which JND model is employed
    - Spatial frame content

$$SCS = \sum_{i=1}^{4} (Categ_{16} = \quad = Categ_8^i)$$

    - Temporal motion information

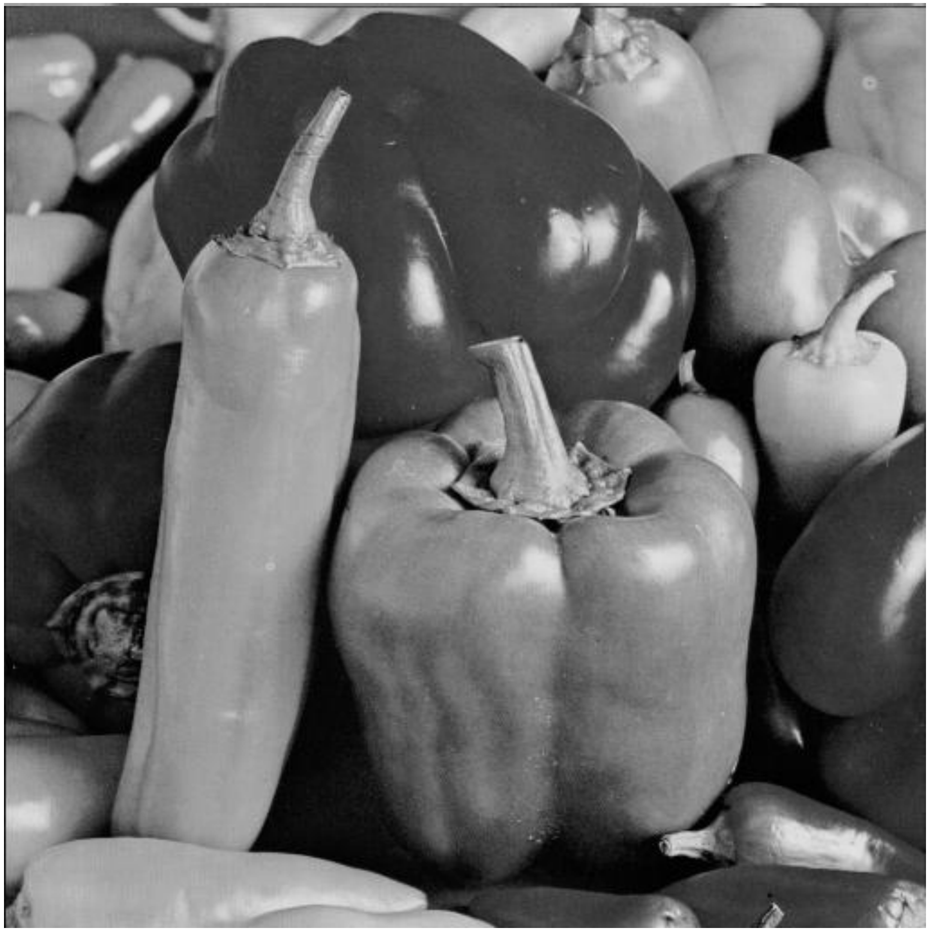
$$MCS = \sum_{i=1}^{4} ||Mv_8^i - Mv_{16}||_2^2 / 4$$

- JND can avoid generating values larger than the actual HVS threshold.

$$\hat{M}_{typ} = M_{typ} + r \cdot T_{JND\_typ}$$

P2-1. Full-reference video quality assessment

- HVS tolerance property by ABT-JND



Original image

Noise-inserted (psnr = 29.66dB)

P2-2. Full-reference video quality assessment

- Perceptual quality metric based on ABT-JND
  - HVS responses of the distortions are modeled by ABT-JND

$$Diff_{typ} = \begin{cases} 0, & if \left| I_{typ} - I_{typ}^{D} \right| \le T_{typ} \\ \left| I_{typ} - I_{typ}^{D} \right| - T_{typ}, & otherwise \end{cases}$$

$$P_{dist} = \tau_{typ} \cdot \frac{Diff_{typ}}{T_{typ}}$$

  - Frame-level $F_Q(k) = \log_{10}(mean(P_{dist}^2(k)))$
  - Sequence-level $V_Q = mean(F_Q(k))$

P2-2. Full-reference video quality assessment

- Perceptual quality metric performance

Database		PSNR	SSIM	VSNR	VIF	Proposed
LIVE image	CC	0.8716	0.904	0.637	0.956	0.933
	SROCC	0.8765	0.910	0.648	0.958	0.934
	RMSE	13.392	11.68	21.13	7.99	9.881
IRCCyN/IVC	CC	0.704	0.776	0.800	0.903	0.913
	SROCC	0.679	0.778	0.798	0.896	0.909
	RMSE	0.866	0.769	0.731	0.524	0.498
A57	CC	0.644	0.415	0.942	0.618	0.913
	SROCC	0.570	0.407	0.936	0.622	0.901
	RMSE	0.192	0.224	0.083	0.193	0.101

Database		PSNR	SSIM	VIF	VQM	MOVIE	Proposed
Live video	CC	0.5398	0.4999	0.5735	0.7160	0.8116	0.780
	SROCC	0.5234	0.5247	0.5564	0.7029	0.7890	0.761
	RMSE	9.241	9.507	8.992	7.664	–	6.935

P2-3. Full-reference video quality assessment

- Perceptual quality metric based on ABT-JND
  - Simple formulation
  - Perform in DCT domain
  - Correlate better with HVS perception
- Application on perceptual video coding:

$$\mathrm{Re}_{typ} = DCT_{typ}\{I - I_{pre}\}$$

$$\mathrm{Re}'_{typ} = \begin{cases} 0, & if\ \mathrm{Re}_{typ} \le JND_{typ} \\ sign(\mathrm{Re}_{typ}) \cdot (|\mathrm{Re}_{typ}| - JND_{typ}), & otherwise \end{cases}$$

$$P_{dist} = \tau_{typ} \cdot \frac{\mathrm{Re}'_{typ}}{T_{typ}}$$

P_{dist} will be employed during the RDO process.

P2-3. Full-reference video quality assessment

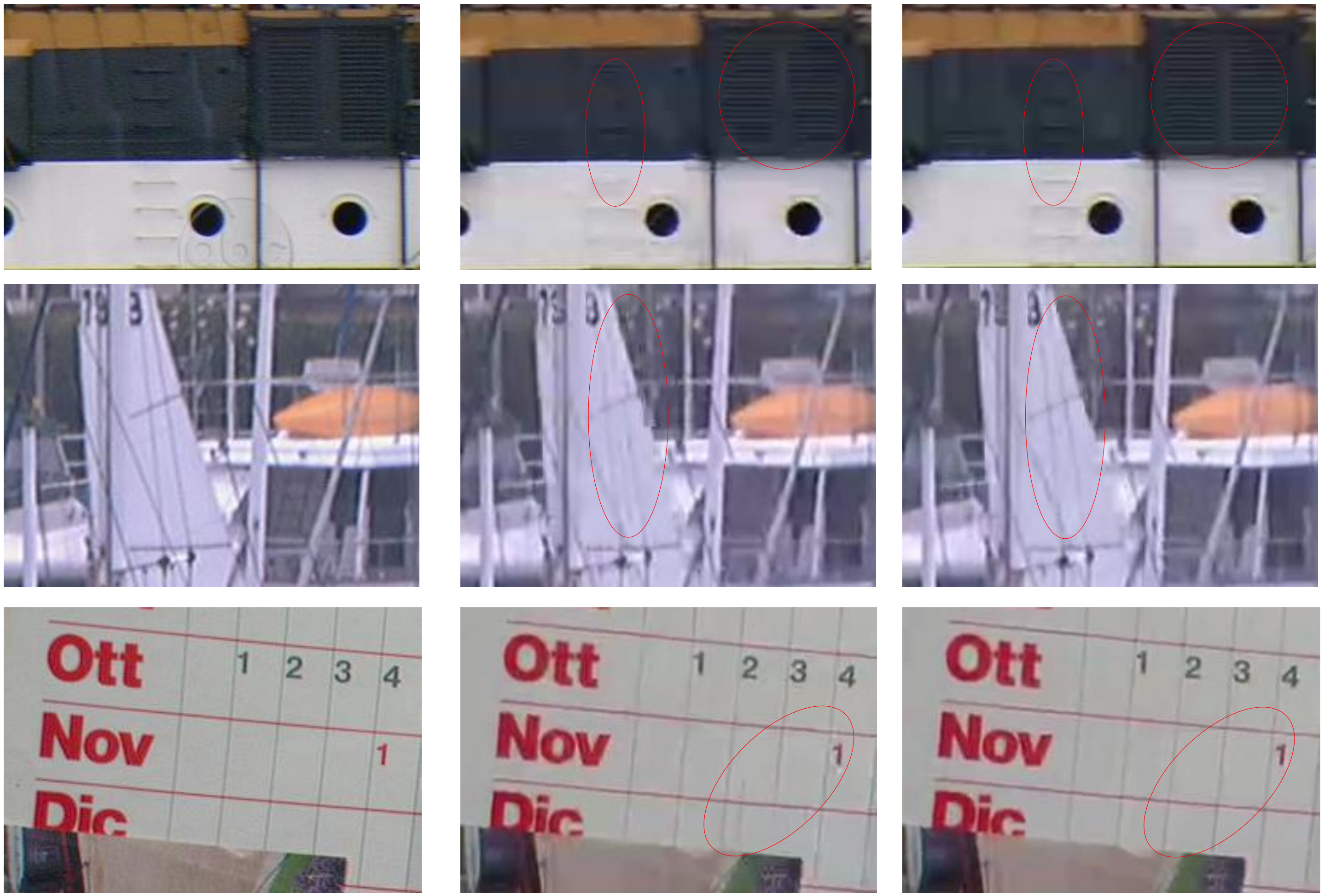
- Perceptual video coding experimental results:

Platform	JM 11 (H.264)
Sequence structure	IBBPBBP
Intra period	10 frames
Transform Size	8×8, and 16×16
Entropy coding	CABAC
Deblocking filter	On
R-D optimization	On
Rate control	Off
Reference frame	2
Search Range	± 32
Frame rate	30 frames/s
Total frame number	199

Video	ABT Codec			The proposed codec			ΔPSNR	ΔV_Q
	Bit-rates (Kbit/s)	PSNR (dB)	V_Q	Bit-rates (Kbit/s)	PSNR (dB)	V_Q		
Crew	807.79	36.68	2.88	806.28	36.42	2.76	−0.26	−0.12
Harbor	1068.34	30.05	13.32	1056.37	29.83	13.22	−0.22	−0.10
Sailormen	572.40	30.92	10.09	576.51	30.86	9.78	−0.06	−0.31
Spincalendar	683.91	31.23	8.40	688.70	31.05	8.23	−0.18	−0.17

J. Dong, K. N. Ngan, C. K. Fong, and W. K. Cham, "2D order-16 integer transforms for HD video coding," TCSVT, Vol.19, No.10, pp.1463-1474, Oct. 2009.

P2-3. Full-reference video quality assessment



Visual quality comparison of regions of the reconstructed frames generated by different video codec. Left: original frame; middle: reconstructed frame from ABT codec; Right: reconstructed frame for the proposed ABT-based JND codec.

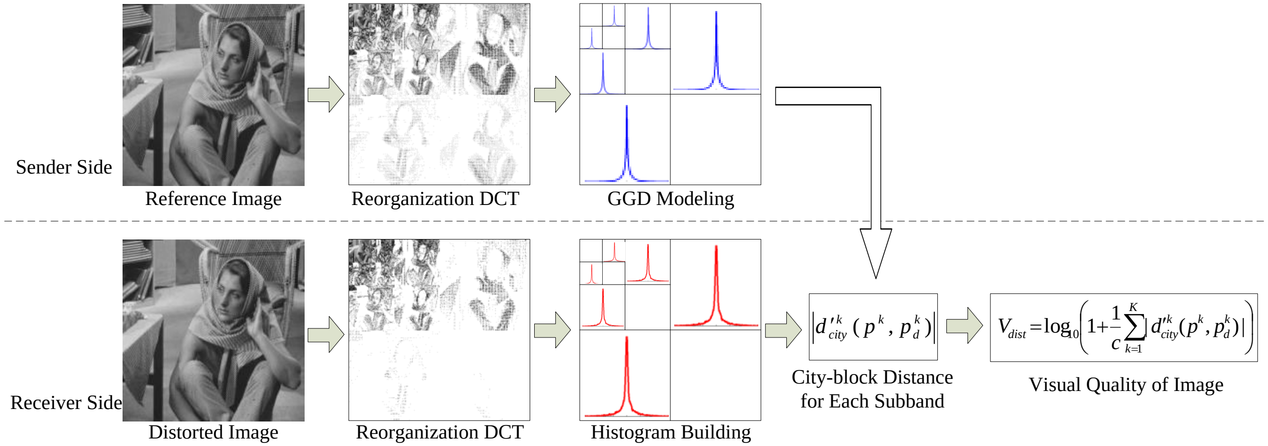
P2. Full-reference video quality assessment

- Contribution
  - ABT based JND model
    - Psychophysical experiment to extend JND from 8×8 to 16×16
    - Spatial and temporal selection strategy
    - Accurately modeling HVS error tolerance ability
  - Quality metric development
    - Simple formulation
    - Correlate better with HVS perception
  - Perceptual video coding
    - Perform in DCT domain
    - Better perceptual quality of coded video sequences

Outline

- Introduction & motivation
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

P3-1. Reduced-reference image quality assessment



- Sender side: **feature extraction and representation**
  - DCT coefficient reorganization
  - GGD modeling the DCT coefficient distribution
  - RR parameters represented in 162 bits.
- Receiver side: **quality analysis**
  - Histogram building for each reorganized DCT subband.
  - City block distance (CBD) measuring the difference between original and distorted image.

P3-1. Reduced-reference image quality assessment

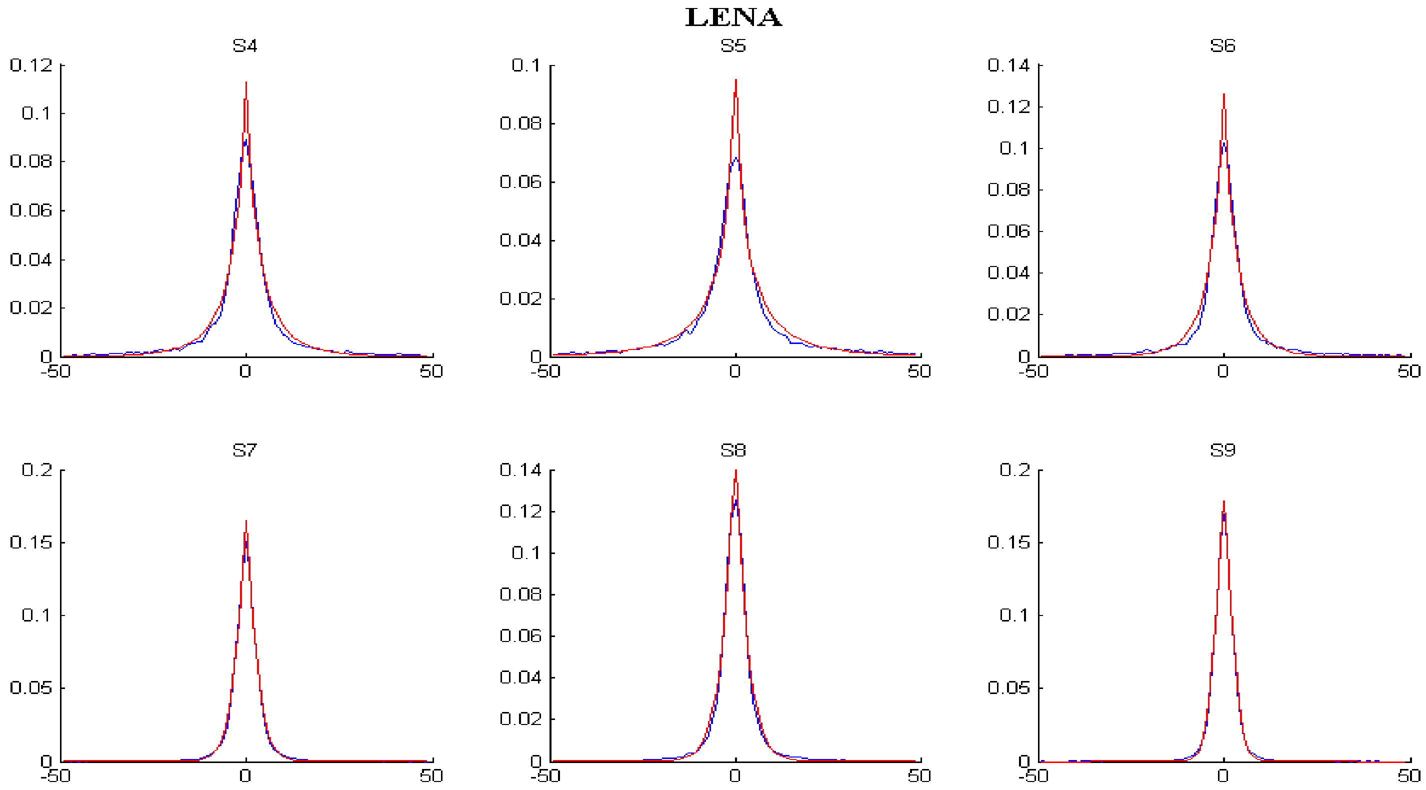
- DCT reorganization

(a) (b) (c) (d)

(a) one 8×8 DCT block with ten subband decomposition; (b): the reorganized DCT image representation taken as a three-level coefficient tree; (c): 8×8 DCT representation of Lena image; (d): RDCT representation of Lena image. (For better visualization, the DC components are rescaled to integers between 0 and 255, while the AC coefficients are obtained by 255-(5×|AC|).)

P3-1. Reduced-reference image quality assessment

- GGD modeling
  - Generalized Gaussian density (GGD) models the DCT coefficient distributions:



P3-1. Reduced-reference image quality assessment

- City-block distance:
  - In sender side, city-block distance (CBD) depicts the GGD modeling error:

$$d_{city}(p_{\alpha,\beta}, p) = \sum_{i=1}^{L} | p(i) - p_{\alpha,\beta}(i) |$$

  - In receiver side, city-block distance is obtained:

$$d_{city}(p, p_d) = \sum_{i=1}^{L} | p(i) - p_d(i) |$$

  - Distance is approximated:

$$d'_{city}(p, p_d) = d_{city}(p_{\alpha,\beta}, p_d) - d_{city}(p_{\alpha,\beta}, p)$$

P3-1. Reduced-reference image quality assessment

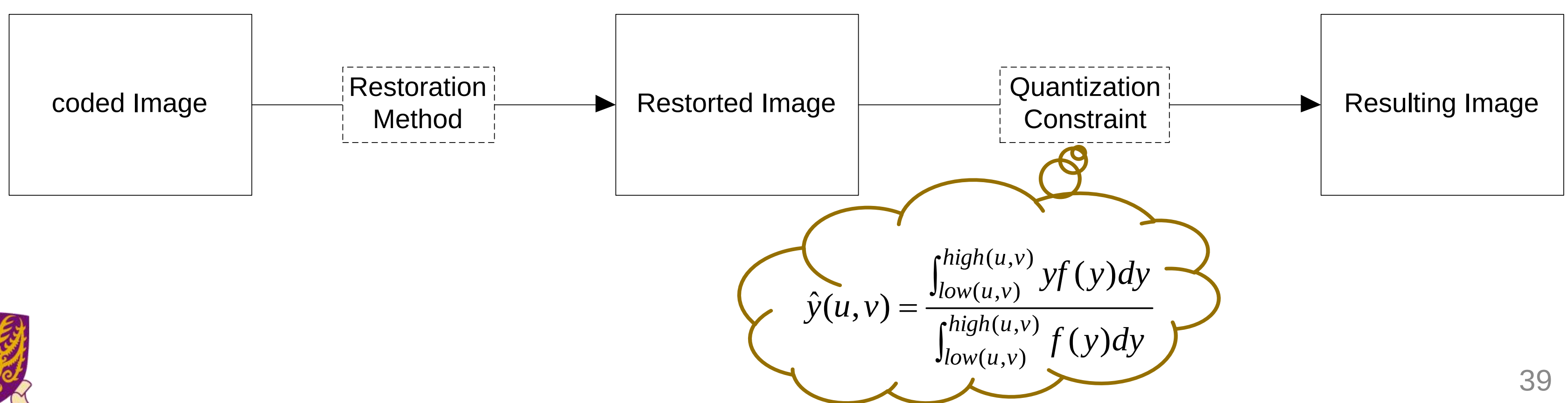
- Experimental results
  - Standard evaluation procedure
  - Over public image databases: LIVE
  - Better than FR PSNR and RR WNISM
  - Only 162 bits utilized for encoding RR features and representing the original image

	PSNR	SSIM	WNISM	The proposed RRIQA (with all 6 subbands)
CC	0.871	0.904	0.738	0.883
SROCC	0.876	0.910	0.779	0.879
RMSE	13.40	11.67	18.43	12.84

Z. Wang, and E. P. Simoncelli, "Reduced-Reference Image Quality Assessment Using A Wavelet-Domain Natural Image Statistic Model", HVEI, 2005.
Z. Wang, G. Wu, H. R. Sheihk, E. P. Simoncelli, E. Yang, and A. C. Bovik, "Quality Aware Images", TIP, 2006.

P3-2. Reduced-reference image quality assessment

- Application on image restoration of JPEG coded image
  - DCT coefficient lies within a reasonable range
    - Quantization constraint set (QCS).
  - GGD simulating the DCT coefficient distributions
  - Centroid of the QCS should be set to the DCT coefficient lying out of the reasonable range



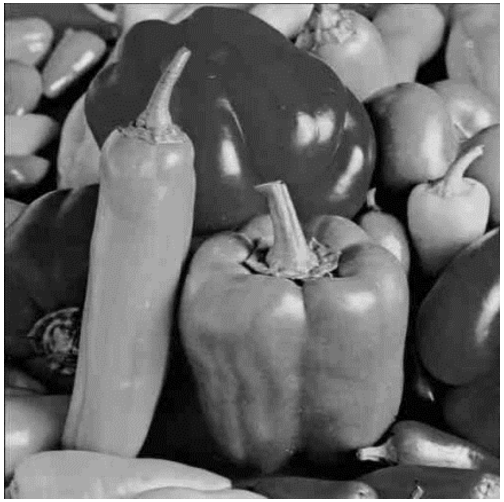
$$\hat{y}(u,v) = \frac{\int_{low(u,v)}^{high(u,v)} y f(y) dy}{\int_{low(u,v)}^{high(u,v)} f(y) dy}$$

P3-2. Reduced-reference image quality assessment

psnr=31.9469dB; SSIM=0.8528

psnr=32.1827dB; SSIM=0.8611

psnr=32.8117dB; SSIM=0.8746



psnr=32.4754dB; SSIM=0.8416

psnr=32.5863dB; SSIM=0.8510

psnr=33.0912dB; SSIM=0.8578

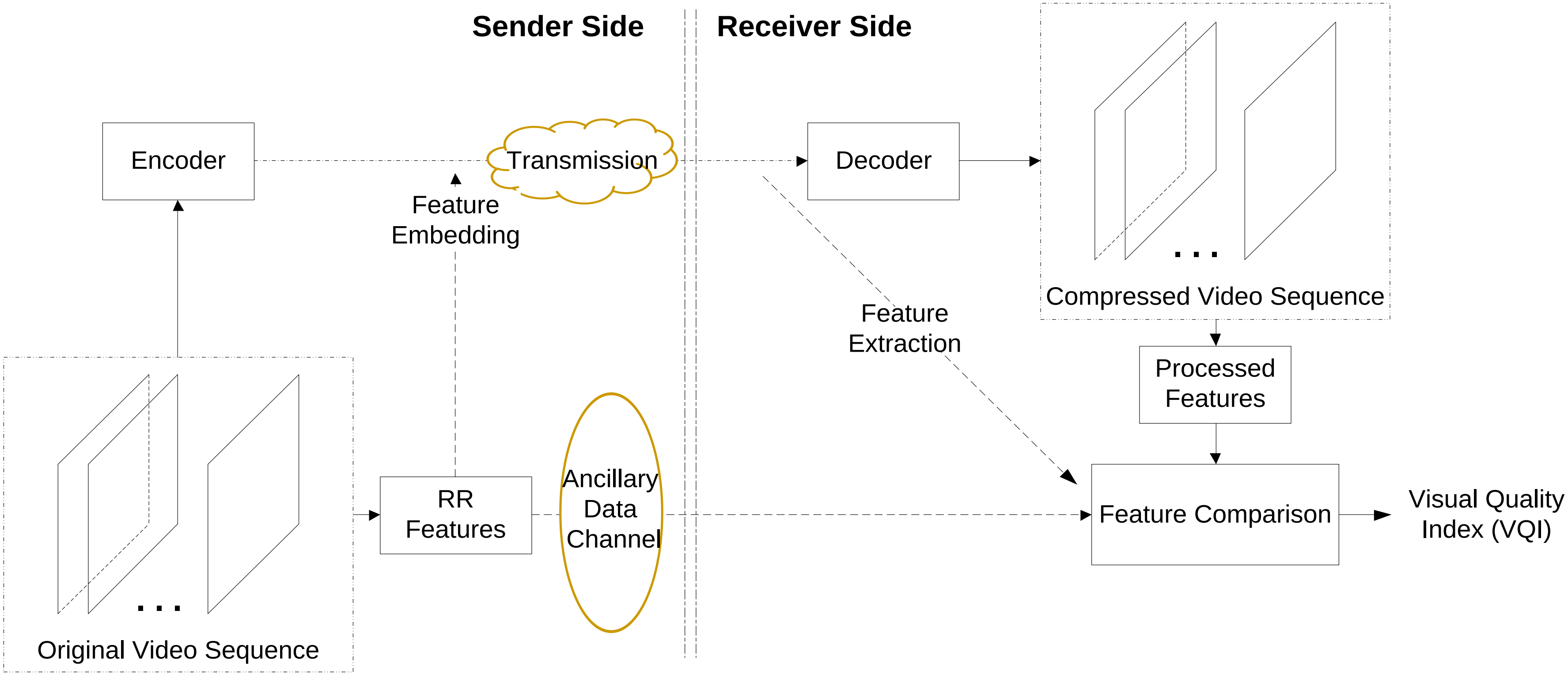
P3. Reduced-reference image quality assessment

- Contribution
  - DCT reorganization
  - GGD modeling the DCT coefficient distribution
  - CBD distance for quality analysis
  - Correlate better with HVS perception
  - Potential application on image restoration

Outline

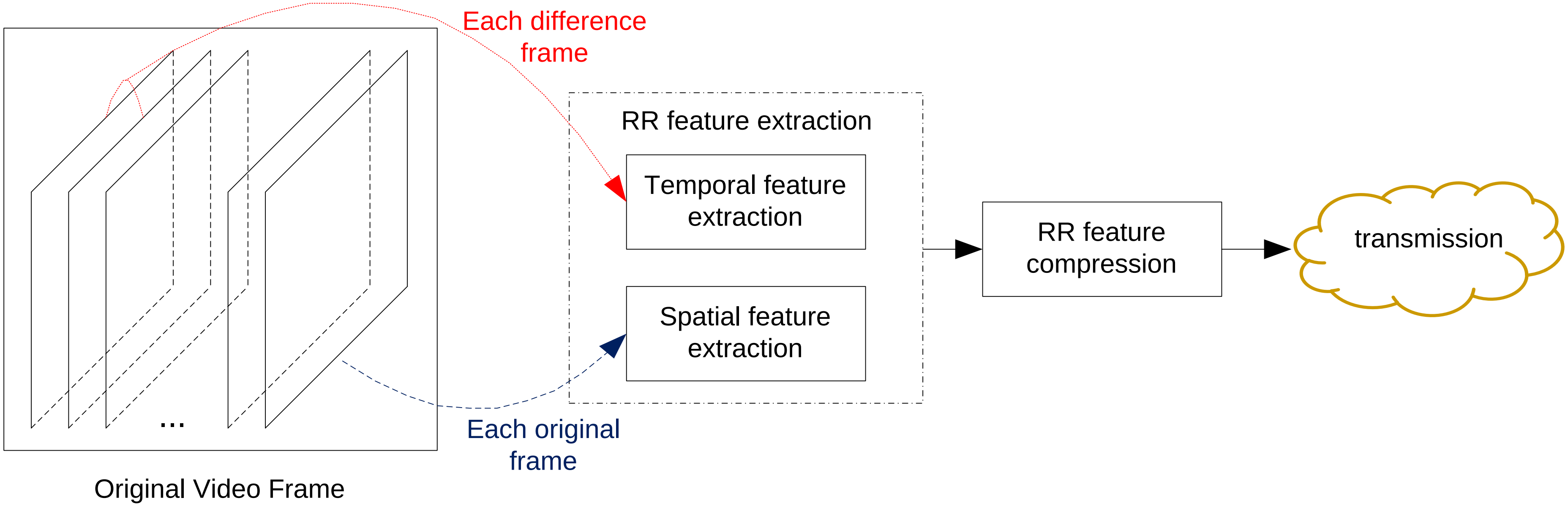
- Introduction & motivation
- Overview of my Ph.D research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

P4. Reduced-reference video quality assessment



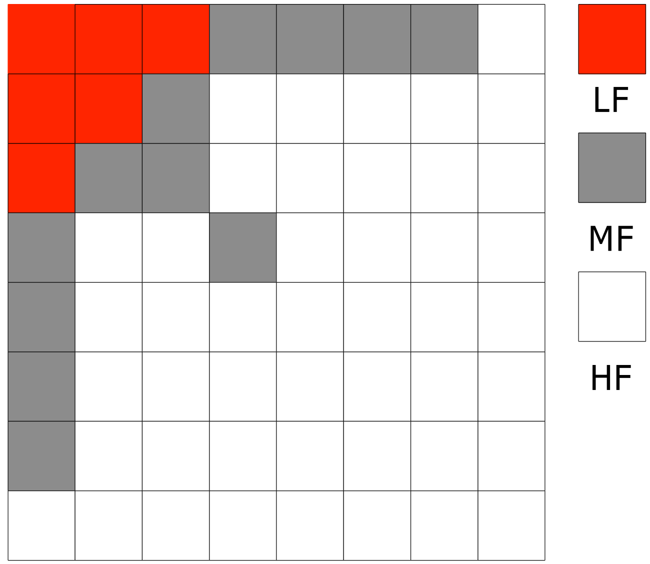
- Sender side:
  - Feature extraction and representation
- Receiver side:
  - Quality analysis

P4. Reduced-reference video quality assessment



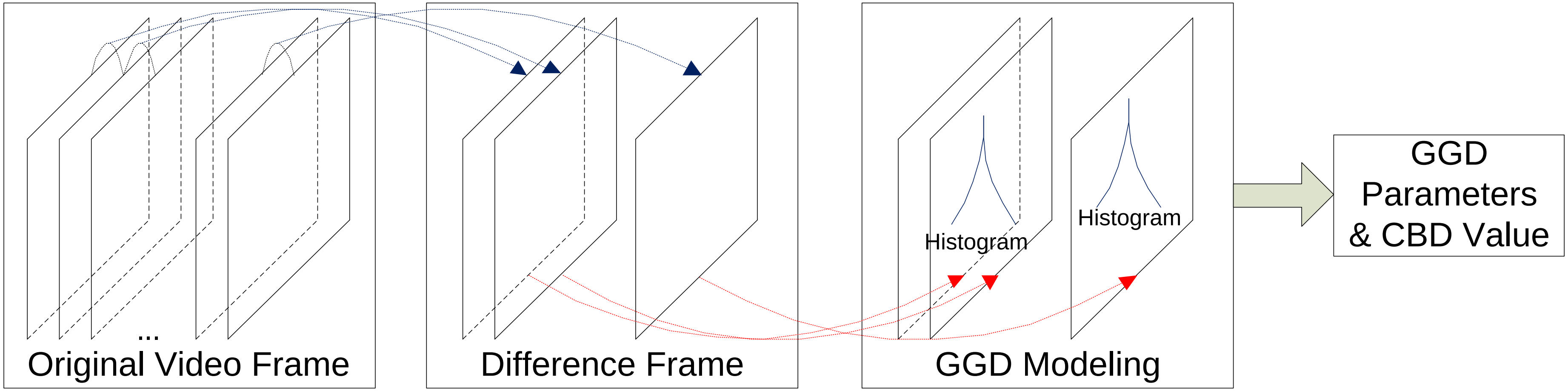
- Sender side: RR feature extraction and representation
  - Spatial feature extraction
  - Temporal feature extraction
  - RR feature representation (8 bits for spatial feature and 27 bits for temporal features), which results in 0.875 kbp/s for 25 fps video.

P4. Reduced-reference video quality assessment



- Spatial feature
  - Energy variation descriptor: $$EVD = \frac{(M+H)}{L}$$
    - Capturing energy variation caused by compression
    - Simulating **HVS masking** property
      - Higher EVD $\longrightarrow$ more texture information $\longrightarrow$ higher masking property

P4. Reduced-reference video quality assessment



- Temporal feature
  - GGD parameters

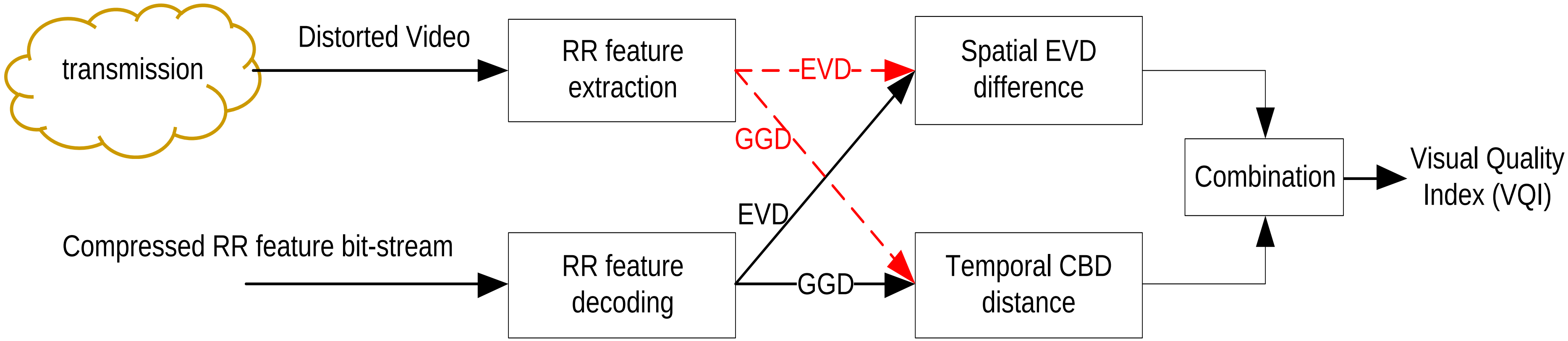
$$p_{\alpha,\beta}(x) = \frac{\beta}{2\alpha\Gamma\left(\frac{1}{\beta}\right)} exp\left\{-\left(\frac{|x|}{\alpha}\right)^{\beta}\right\}$$

$$\Gamma(x) = \int_0^{\infty} t^{x-1} e^{-t} dt$$

  - CBD value

$$d_{CBD}\left(p, p_{\alpha,\beta}\right) = \sum_{i=1}^{h_L} \left|p(i) - p_{\alpha,\beta}(i)\right|$$

P4. Reduced-reference video quality assessment



- Visual quality analysis in receiver side
  - Spatial EVD difference

$$EL_V = \frac{EL}{EVD_{ori}} = \frac{\left|EVD_{ori} - EVD_{pro}\right|}{EVD_{ori}}$$

  - Temporal CBD distance

$$d_{CBD}(p, p_d) \triangleq \left|d_{CBD}\left(p_{\alpha,\beta}, p_d\right) - d_{CBD}\left(p, p_{\alpha,\beta}\right)\right|$$

  - Quality combination

$$Q_s = EL_V \times log_{10}\left(1 + \frac{d_{CBD}(p, p_d)}{c}\right)$$

P4. Reduced-reference video quality assessment

- Experimental results
  - Over public video databases: LIVE
  - Better than FR PSNR, SSIM, VIF, and RR J.246, RR-LHS, and VQM.

	LCC	SROCC	RMSE	Reference type	Data rate (25fps)
PSNR	0.4488	0.4157	9.188	FR	-
SSIM	0.5946	0.5969	8.267	FR	-
MSSIM	0.6671	0.6944	7.717	FR	-
VSNR	0.3097	0.3041	9.777	FR	-
VIF	0.6447	0.6350	7.860	FR	-
J.246	0.5036	0.4460	8.883	RR	10 kbps
RR-LHS	0.4557	0.4082	9.152	RR	64 kbps
VQM	0.7003	0.6790	7.340	RR	150 kbps
Proposed metric	0.7567	0.7486	6.722	RR	0.875 kbps

ITU-T Recommendation J.246, “Perceptual visual quality measurement techniques for multimedia services over digital cable television networks in the presence of a reduced bandwidth reference”, Aug. 2008.

P. Gunawan, and M. Ghanbari, “Reduced-reference video quality assessment using discriminative local harmonic strength with motion consideration”, TCSVT, vol. 18, no. 1, pp. 71-83, Jan. 2010.

M. H. Pinson, and S. Wolf, “A new standardized method for objectively measuring video quality”, IEEE Trans. Broadcasting, vol. 50, no. 3, pp. 312-322, Sept. 2004.

P4. Reduced-reference video quality assessment

- Contribution
  - EVD depicts spatial energy variation
  - EVD simulates HVS texture masking property
  - GGD depicts temporal statistical property
  - Correlate better with HVS perception
  - Efficient RR feature representation (0.875kbp/s)
  - Low complexity

Outline

- Introduction & motivation
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

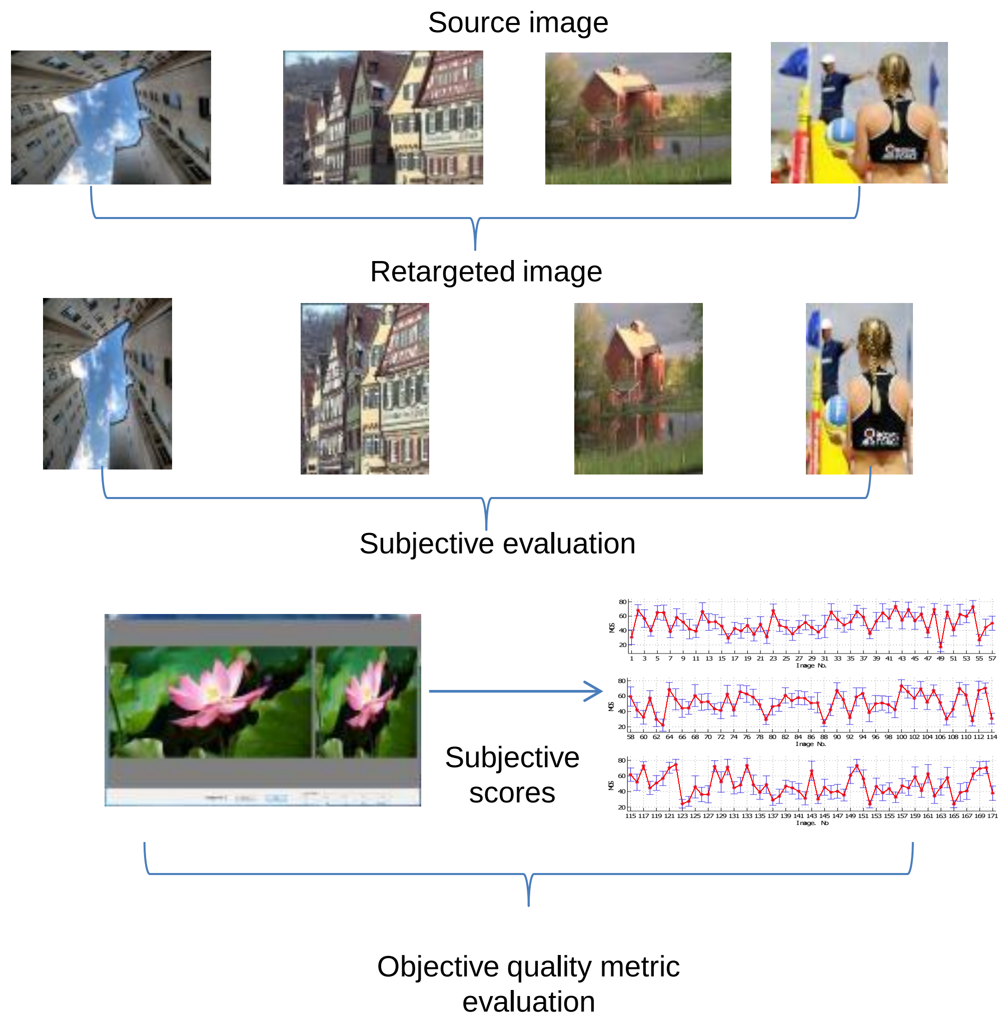
P5. Retargeting image quality assessment

- Content-aware image resizing – image retargeting
  - Reduction/Expansion of image
  - Better quality of experience (QoE)

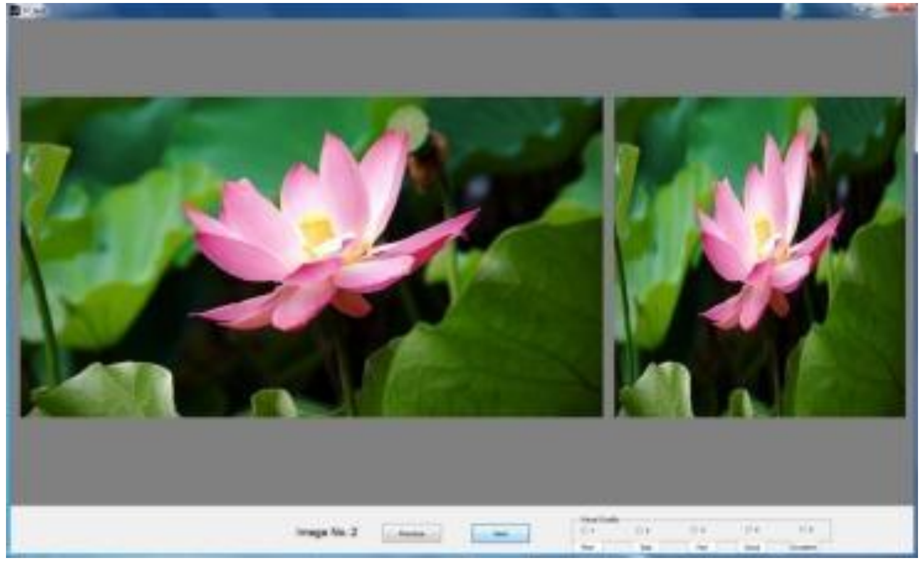
ent

P5. Retargeting image quality assessment

- Retargeting image quality assessment
  - Retargeted image
  - Subjective evaluation
  - Subjective score processing
  - Objective quality metric evaluation



P5. Retargeting image quality assessment



- Subjective testing
  - 57 source images $\xrightarrow{\text{10 retargeted methods}}$ 171 retargeted images
  - [Subjective testing]()
    - 30 subjects involved in session 1
    - 34 subjects involved in session 2

P5. Retargeting image quality assessment

- Subjective score processing
  - Check subjective agreement

$$NCC = \frac{a^t \cdot b}{\|a\|\|b\|} \quad EUD = \frac{\|a-b\|_2^2}{k}$$

  - Reject unreliable subjects

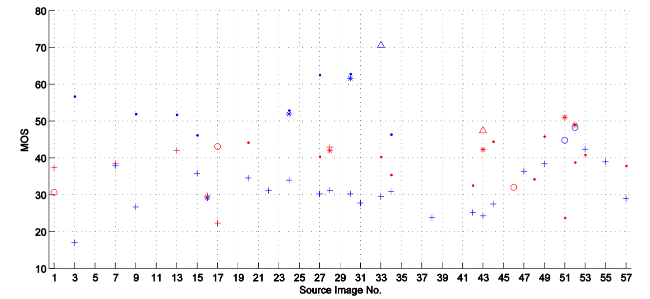
For each subject i, find the P_{ik} and Q_{ik}

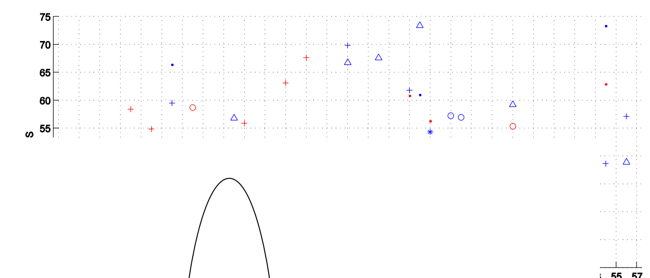
if $2 \leq \beta_j \leq 4$ **(normally distributed)**
    if $S_{ijk} \geq u_{ik} + 2\sigma_{ik}$, **then** $P_{ik} = P_{ik} + 1$;
    if $S_{ijk} \leq u_{ik} - 2\sigma_{ik}$, **then** $Q_{ik} = Q_{ik} + 1$;
else
    if $S_{ijk} \geq u_{ik} + \sqrt{20}\sigma_{ik}$, **then** $P_{ik} = P_{ik} + 1$;
    if $S_{ijk} \leq u_{ik} - \sqrt{20}\sigma_{ik}$, **then** $Q_{ik} = Q_{ik} + 1$;
end

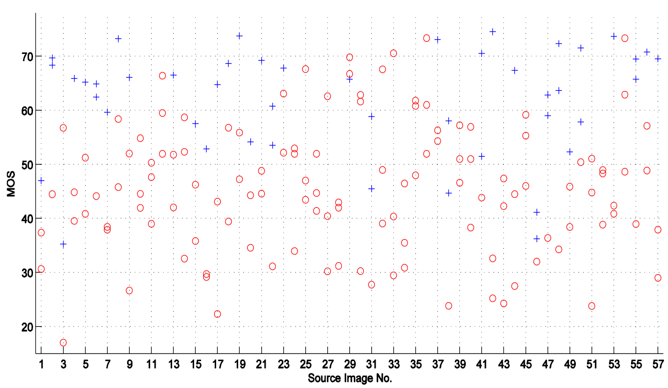
if $\frac{P_{ik}+Q_{ik}}{N_{ik}} > 0.05$ **and** $\left\|\frac{P_{ik}-Q_{ik}}{P_{ik}+Q_{ik}}\right\| < 0.3$, **then REJECT the subject** i.

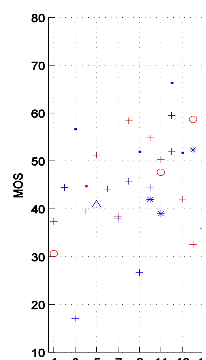
P5. Retargeting image quality assessment

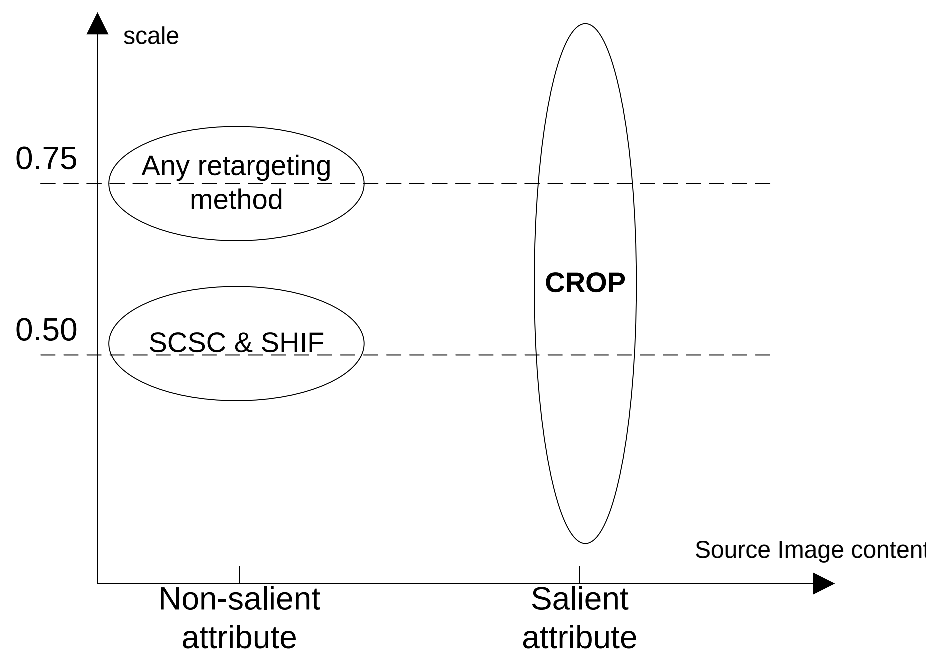
- Subjective score analysis
  - Retargeting scale
  - Retargeting method
  - Source image conten
    - Salient attributes(edg
    - Non-salient attribute











P5. Retargeting image quality assessment

- Objective quality metric evaluation
  - Earth mover's distance (EMD)
  - Bidirectional similarity (BDS)
  - Edge histogram (EH)
  - SIFT-flow
- Evaluation on our image retargeting database

	EH	EMD	BSD	SIFT-flow	Fusion (EH, EMD and SIFT-flow)	Fusion (EH, EMD, BSD, SIFT-flow)
LCC	0.3422	0.2760	0.2896	0.3141	0.4361	0.5217
SROCC	0.3288	0.2904	0.2887	0.2899	0.4203	0.4514
RMSE	12.686	12.977	12.922	12.817	12.149	11.484
OR	0.2047	0.1696	0.2164	0.1462	0.1462	0.1287

P5. Retargeting image quality assessment

- Possible candidate factors for retargeting image quality assessment
  - Shape distortion description
  - Content information loss evaluation
  - Fusion of shape distortion description and content information loss evolution
  - Source image quality and retargeting scale
  - Image content
  - HVS saliency

P5. Retargeting image quality assessment

- Contribution
  - Subjective testing for retargeting images
  - Retargeting image database
  - Subjective analysis of the database
  - Objective quality evaluation
  - Possible candidate factors

Outline

- Introduction & motivation
- Overview of my research work
  - Full-reference image quality assessment
  - Full-reference video quality assessment
  - Reduced-reference image quality assessment
  - Reduced-reference video quality assessment
  - Retargeting image quality assessment
- Conclusion

Conclusion

- Signal statistical and HVS properties are researched for quality evaluation of visual signals
- Quality metric design for visual signals
  - Full-reference image quality metric
  - Full-reference video quality metric
  - Reduced-reference image quality metric
  - Reduced-reference video quality metric
  - Subjective testing and objective evaluation of retargeted images
- Better quality of experience (QoE) can be provided for users.

Publications-1

- Journal
  - **Lin Ma**, Weisi Lin, Chenwei Deng, and King N. Ngan, "Image Retargeting Quality Assessment: A Study of Subjective Scores and Objective Metrics", *IEEE Journal of Selected Topics in Signal Processing, Special Issue on New Subjective and Objective Methodologies for Audio and Visual Signal Processing*, vol. 6, no. 6, pp. 626-639, Oct. 2012.
  - **Lin Ma**, Songnan Li, and King N. Ngan, "Reduced-Reference Image Quality Assessment in Reorganized DCT Domain", *Signal Processing: Image Communication, Special Issue on Biologically Inspired Approaches for Visual Information Processing and Analysis*. [Accepted]
  - **Lin Ma**, Songnan Li, and King N. Ngan, "Reduced-Reference Video Quality Assessment of Compressed Video Sequences", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 10, pp. 1441-1456, Oct. 2012.
  - **Lin Ma**, Songnan Li, Fan Zhang, and King N. Ngan, "Reduced-Reference Image Quality Assessment Using Reorganized DCT-Based Image Representation", *IEEE Transaction on Multimedia*, vol. 13, no. 4, pp. 824-829, Aug. 2011.
  - **Lin Ma**, King N. Ngan, Fan Zhang, and Songnan Li, "Adaptive Block-Size Transform Based Just-Noticeable Difference Model for Images/Videos", *Signal Processing: Image Communication*, vol. 26, no. 3, pp. 162-174, Mar. 2011.
  - **Lin Ma**, Songnan Li, and King N. Ngan, "Visual Horizontal Effect for Image Quality Assessment", *IEEE Signal Processing Letters*, vol. 17, no. 7, pp. 627-630, Jul. 2010.
  - **Lin Ma**, Debin Zhao, and Wen Gao, "Learning-based Image Restoration for Compressed Images", *Signal Processing: Image Communication*, vol. 27, no. 1, pp. 54-65, Jan. 2012.
  - Yaqing Niu, Matthew Kyan, **Lin Ma**, Azeddine Beghdadi, and Sridhar Krishnan, "Visual Saliency's Modulatory Effect on Just Noticeable Distortion Profile and Its Application in Image Watermarking", *Signal Processing: Image Communication*, Special Issue on Biologically Inspired Approaches for Visual Information Processing and Analysis. [Accepted]
  - Songnan Li, **Lin Ma**, and King N. Ngan, "Full-reference Video Quality Assessment by Decoupling Detail Losses and Additive Impairments", *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 22, no. 7, pp. 1100-1112, Jul. 2012.
  - Songnan Li, Fan Zhang, **Lin Ma**, and King N. Ngan, "Image Quality Assessment by Separately Evaluating Detail Losses and Additive Impairments", *IEEE Transaction on Multimedia*. vol. 13, no. 5, pp. 935-949, Oct. 2011.
  - Fan Zhang, **Lin Ma**, Songnan Li, and King N. Ngan, "Practical Image Quality Metric Applied to Image Coding", *IEEE Transaction on Multimedia*, vol. 13, no. 4, pp. 615-624, Aug. 2011.

Publications-2

- Conference
  - **Lin Ma**, Weisi Lin, Chenwei Deng, and King N. Ngan, "Study of Subjective and Objective Quality Assessment of Retargeted Images", *International Symposium on Circuits and Systems (ISCAS 2012)*, Seoul, Korea, May 20-23, 2012.
  - **Lin Ma**, Songnan Li, and King N. Ngan, "Reduced-Reference Image Quality Assessment via Intra- and Inter-Subband Statistical Characteristics in Reorganized DCT Domain", *Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2011)*, Xi'an, China, Oct. 18-21, 2011.
  - **Lin Ma**, Songnan Li, and King N. Ngan, "Motion Trajectory Based Visual Saliency for Video Quality Assessment", *International Conference on Image Processing (ICIP 2011)*, Brussels, Belgium, Sep. 11-14, 2011.
  - **Lin Ma**, Songnan Li, and King N. Ngan, "Perceptual Image Compression via Adaptive Block-Based Super-Resolution Directed Down-Sampling", *International Symposium on Circuits and Systems (ISCAS 2011)*, Rio de Janeiro, Brazil, May 15-18, 2011.
  - **Lin Ma**, Fan Zhang, Songnan Li, and King N. Ngan, "Video Quality Assessment Based on Adaptive Block-Size Transform Just-Noticeable Difference Model", *International Conference on Image Processing (ICIP 2010)*, Hong Kong, China, Sep. 26-29, 2010.
  - **Lin Ma**, and King N. Ngan, "Adaptive Block-Size Transform Based Just-Noticeable Difference Profile for Videos", *International Symposium on Circuits and Systems (ISCAS 2010)*, Paris, France, May 30 – Jun. 2, 2010.
  - **Lin Ma**, and King N. Ngan, "Adaptive Block-Size Transform Based Just-Noticeable Difference Profile for Images", *Pacific-Rim Conference on Multimedia (PCM 2009)*, Bangkok, Thailand, Dec. 15-18, 2009.
  - **Lin Ma**, Feng Wu, Debin Zhao, Wen Gao, and Siwei Ma, "Learning-based Image Restoration for Compressed Image through Neighboring Embedding", (Best Paper Award) *Pacific-Rim Conference on Multimedia (PCM 2008)*, Tainan, Taiwan, Dec. 9-13, 2008.
  - **Lin Ma**, Yonghua Zhang, Yan Lu, Feng Wu, and Debin Zhao, "Three-tiered Network Model for Image Hallucination", *International Conference of Image Processing (ICIP 2008)*, San Diego, California, USA, Oct. 12-15, 2008.
  - Long Xu, King N. Ngan, Songnan Li, and **Lin Ma**, "Video Quality Metric for Consistent Visual Quality Control in Video Coding", *Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC 2012)*, Hollywood, California, USA, Dec. 3-6, 2012.
  - Songnan Li, **Lin Ma**, and King N. Ngan, "Video Quality Assessment by Decoupling Additive Impairments and Detail losses", *International Workshop on Quality of Multimedia Experience(QoMEX 2011)*, Mechelen, Belgium, Sep. 7-9, 2011.
  - Yaqing Niu, Matthew Kyan, **Lin Ma**, Azeddine Beghdadi, and Sridhar Krishnan, "A Visual Saliency Modulated Just Noticeable Distortion Profile for Image Watermarking", *European Signal Processing Conference (EUSIPCO 2011)*, Barcelona, Spain, Aug. 29 - Sep. 2, 2011.
  - Songnan Li, **Lin Ma**, Fan Zhang, and King N. Ngan, "Temporal Inconsistency Measure for Video Quality Assessment", *Picture Coding Symposium (PCS 2010)*, Nagoya, Japan, Dec. 7-10, 2010.
  - Fan Zhang, Songnan Li, **Lin Ma**, and King N. Ngan, "Limitation and Challenges of Image Quality Measurement", *Visual Communications and Image Processing (VCIP 2010)*, Huang Shan, Anhui, China, Jul. 11-14, 2010. (Invited Paper)
  - Shaohui Liu, **Lin Ma**, Hongxun Yao, and Debin Zhao, "Universal Steganalysis Based on Statistical Models Using Reorganization of Block-based DCT Coefficients", *International Conference on Information Assurance and Security (IAS 2009)*, Xi'an, China, Aug. 18-20, 2009.

Thank You!
Q&A