Full Reference Image Quality Assessment Based on Saliency Map Analysis

Journal of Image Science and Technology,

Vol. 54, No. 3, 2010

Yubing Tong and Hubert Konick, Fauzi A. Cheikh, Alain Tremeau

Presented by Bong Seok Choi

School of Electrical Engineering and Computer Science
Kyungpook National Univ.

Abstract

- Proposed method
  - Using saliency map to image quality assessment
    - Salient region
      - Accentuation between distorted and original image
    - Global quality measure of image
  - Using mixed saliency map model
    - Based on Itti`s model and face detection model
      - Composition of low-level feature
        - » Intensity, color, and orientation
      - Composition of high-level feature
        - » Using mixed model

Introduction

- Image quality assessment
  - Subjective image quality assessment model
    - Mean opinion score (MOS)
      - Numerical indication of perceived quality of image
    - Drawback of subjective image quality assessment
      - Requirement of large number of observer
      - Takes a lot of time

- Objective image quality assessment model
  - Objective image quality evaluation techniques
    - Mean squared error (MSE)
    - Peak signal-to-noise ratio (PSNR)
    - Universal Quality Index (UQI)
    - Structural Similarity Index (SSIM)
    - Noise quality measure (NQM)
    - Weight signal-to-noise ratio (WSNR)
    - Visual signal-to-noise ratio (VSNR)

- Objective image quality evaluation techniques for distortion image
  - Using PSNRHVS and PSNRHVSM
    - » Image divided fixed size blocks and same weights
    - » High performance on noise, noise2, safe, simple, and hard subsets of TID2008
    - » Low performance on Exotic and Exotic2 subset of TID2008
  - Drawback of PSNRHVS and PSNRHVSM
    - » Contradictory with Human visual system operate
      - » Dividing image into blocks of equal size
    - » Introducing discontinuity from original image

◆ Proposed method

- Using saliency information to mimic selectivity of HVS
  - Integrate into exiting objective image quality metrics
    - Contribution of salience regions of nonsalient regions
- Using image saliency map
  - Image reconstruction
    - Use to phase spectrum and Fourier transform

Analysis of previous work and primay conclusions

- Previous work
  - Using PSNR and MES for Assessment of distortion image quality
    - Dependent on absolute difference between original and distortion image

$$PSNR = \log_{10}\left(\frac{255^2}{MSE}\right) \tag{1}$$

$$MSE = \frac{1}{M \times N}\sum_{i=1}^{M}\sum_{j=1}^{N}\left[\delta(i,j)\right]^2 \tag{2}$$

$$\delta(i,j) = \left[a(i,j) - \hat{a}(i,j)\right] \tag{3}$$

where (i,j) is current pixel position, $a(i,j)$ and $\hat{a}(i,j)$ are original image and distortion image, M and N height and width of image

– Using PSNRHVS and PSNRHVSM

- Improving performance of PSNR and MSE
  - Considering visual perception
  - Defining PSNRHVS
    - » Dividing image into 8x8 pixel nonoverlapping block

$$\delta_{PSNRHVS}(i,j) = \delta(i,j) \cdot CSF_{cof}(i,j) \quad (4)$$

where (i,j) is calculated using DCT coefficients

  - Defining PSNRHVSM

$$\delta_{PSNRHVSM}(i,j) = \left[\delta(i,j) \cdot CM(i,j)\right] \cdot CSF_{cof}(i,j) \quad (5)$$

$$MSE_{PSNRHVS}(i,j,I,J) = \frac{1}{M \times N} \sum_{I=1}^{M/8} \sum_{J=1}^{N/8} \left\{ \sum_{i=1}^{8} \sum_{j=1}^{8} \left[\delta_{PSNR\_HVS}(i,j)\right]^2 \right\} \quad (6)$$

where (I,J) is position of 8x8 block in the imgae
(i,j) is position of pixel in 8x8 block

- ◆ Analysis of previous work
  - – Using PSNRHVS and PSNRHVSM
    - Processing by nonovelapping 8x8 blocks
      - – Equal to image quality metric at every 8x8 blocks
  - – Drawback of PSNRHVS and PSNRHVSM
    - Independent blocks with fixed size
      - – Producing blockiness result
      - – Drawback to sudden change region

- Comparison to different distortion number in image

PSNR	PSNRHVS	PSNRHVSM
46.3dB	33.74dB	36.3dB

Fig. 1. Reference Image I18.

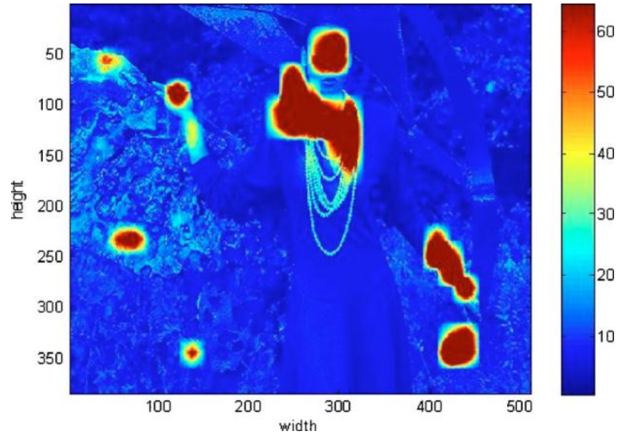


Fig. 2. Saliency map of I18 with face detection.

Fig. 3. I18 with noise in one salient region

PSNR	PSNRHVS	PSNRHVSM
41.6dB	32.4dB	35.8dB

Fig. 4. I18 with noise in nonsalient region

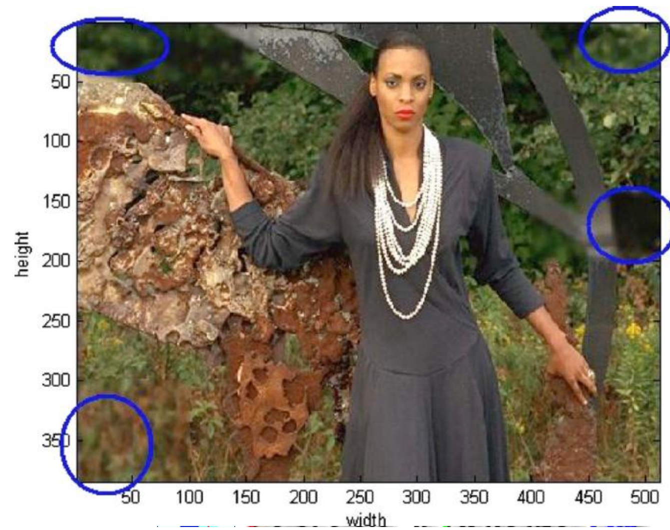


Fig. 5. I18 with distortion in four nonsalient region

- Equal quantity to different parts of image

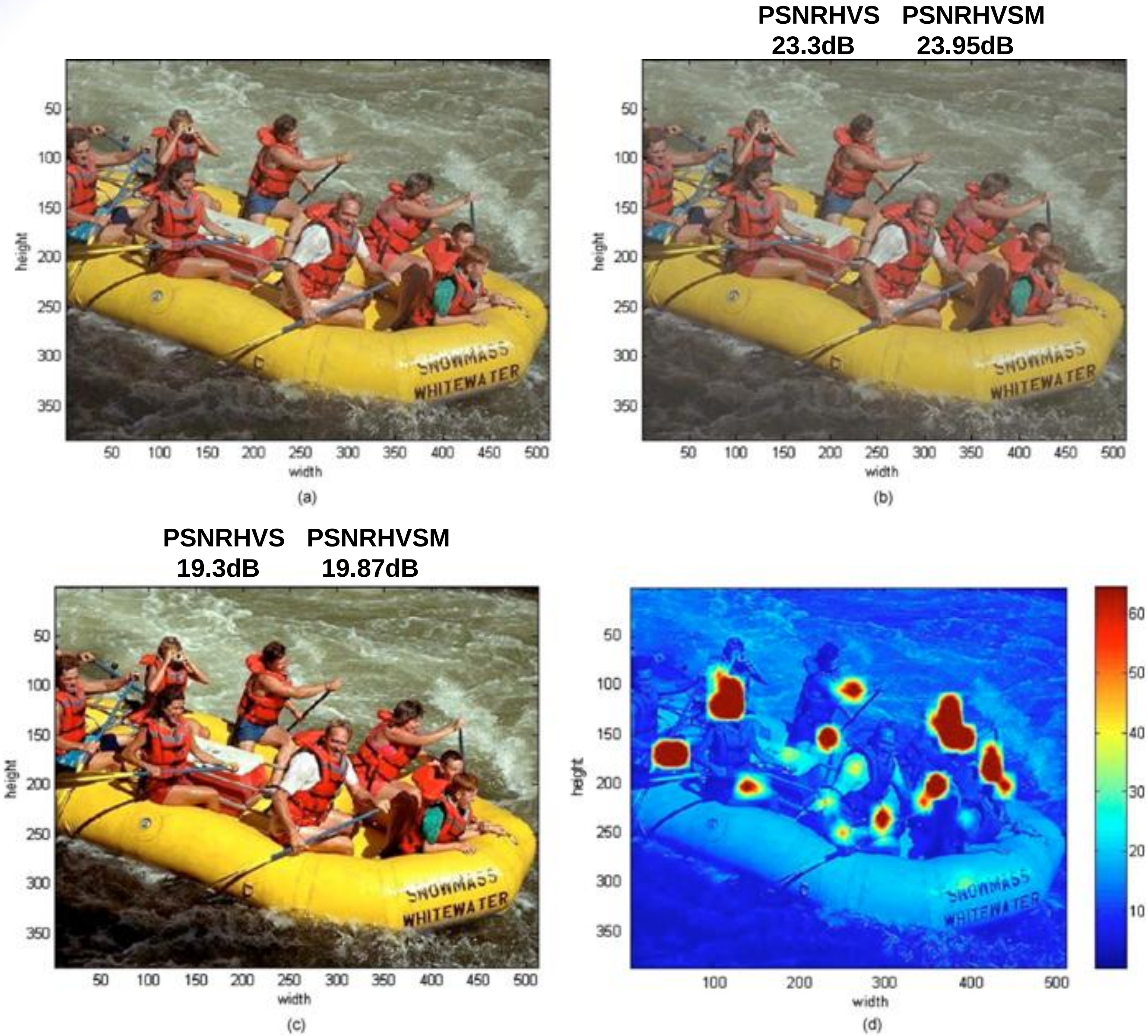


Fig. 6. I14 and corresponding image. (a) The reference I14. (b) The distorted image I14-17-2. (c) The distorted image I14-17-3. (d) The Saliency map of I14

Image quality assessment based on region sliency

- Calculating saliency map of image
  - Using Itti`s saliency map model
  - Using mixed saliency map model
    - Consist of Itti`s map model and Face detection
  - Process to calculate saliency map

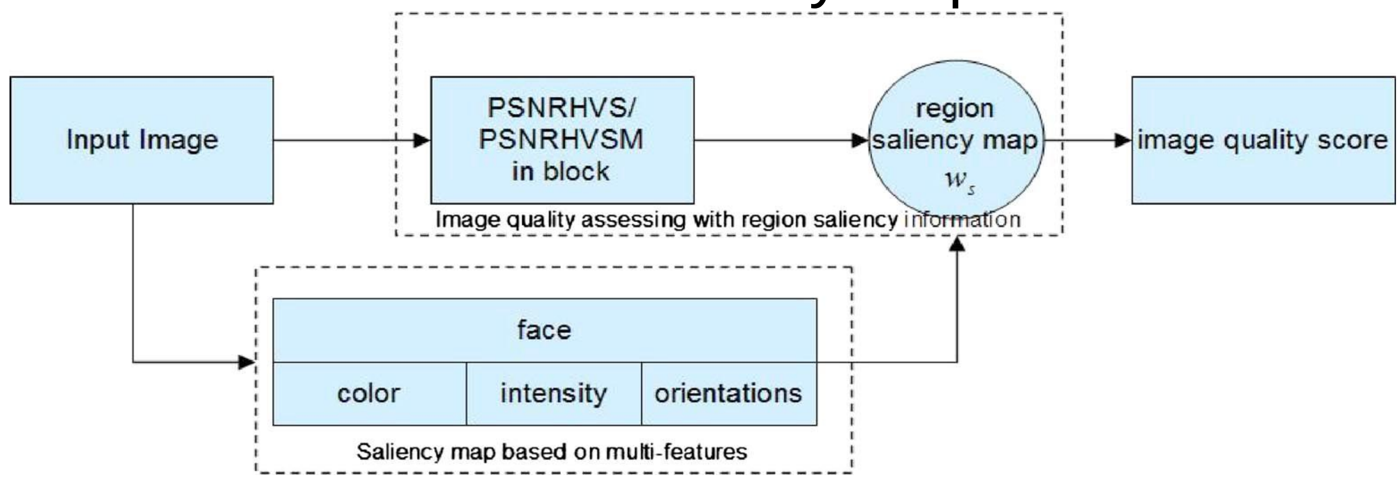


Fig. 7. Flowchart of the method based on region saliency used to assess the image quality

◆ Itti`s saliency map model

– Definition of bottom-up visual attention mechanism
  - Extraction saliency region from input image
– Procedure of saliency map modeling
  - Analyzing each feature by gaussian pyramid
    – One intensity, four orientation, two color opponencies
  - Normalization each features map
  - Combination of features map

$$\boldsymbol{S}_{Itti} = \frac{1}{3} \sum_{k \in i,c,o} C_k \qquad (7)$$

where C_i is intensity conspicuous map, C_c is color conspicous maps
C_o is orientation conspicuous map

– Application to image "I01" in TID2008

- Selectivity of HVS
  - Focus only on some part of image

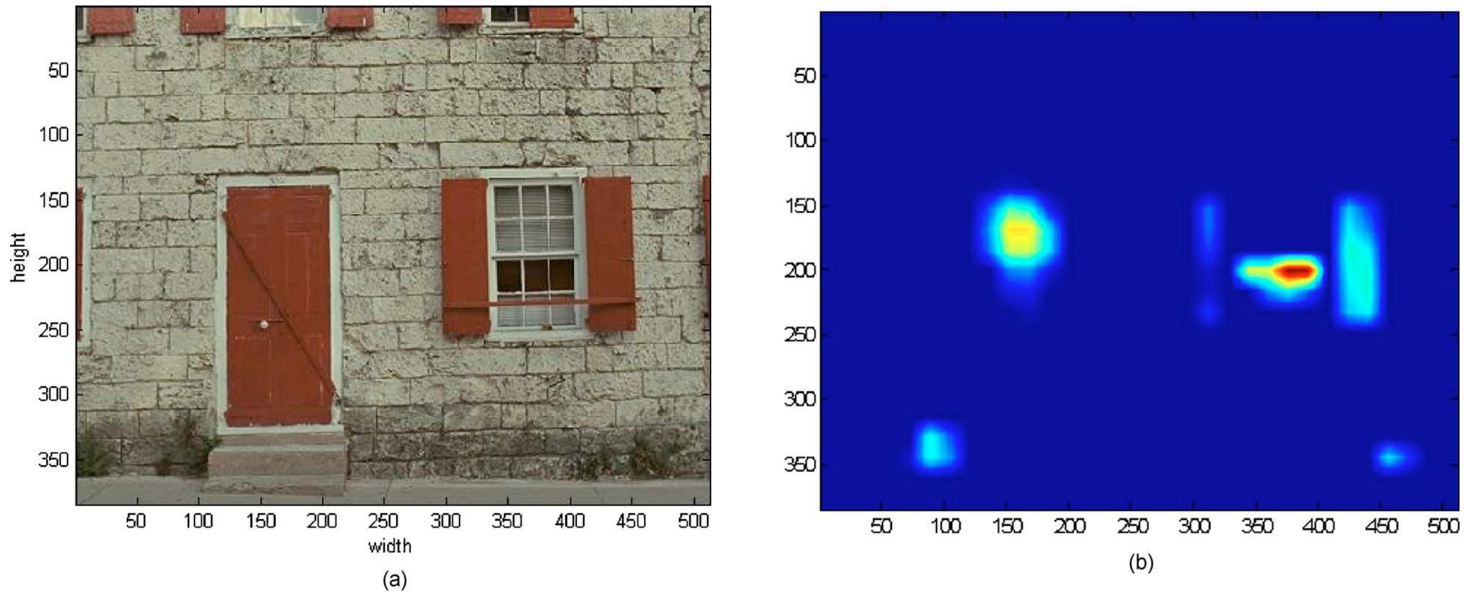


Fig. 8. Image I01 with its saliency map and corresponding surface plot. (a) Reference image I01. (b) saliency map of I01.

- Saliency map model based on face detection
  - Definition of saliency map in face detection
    - Use to face as high level feature
  - Detecting head and skin in image
    - Based on Gaussian model of skin hue distribution

– Procedure of saliency map modeling

- Definition of Hue response from pixel

$$h(r',g') = \exp\left\{-\frac{1}{2}\left[\frac{(r'-\mu_r)^2}{\sigma_r^2} + \frac{(g'-\mu_g)^2}{\sigma_g^2} + \frac{\rho(r'-\mu_r)(g'-\mu_g)}{\sigma_r\sigma_g}\right]\right\} \tag{8}$$

$$r' = \frac{r}{r+g+b} \quad and \quad g' = \frac{g}{r+g+b} \tag{9}$$

where μ_r, μ_g is average of skin hue distributions, σ_r^2 and σ_g^2 are the variances of r' and g' components, and ρ is correlation between components r' and g'

- Gaussian pyramid based on multiscale subsampling
- Calculating center-surround map from pyramid
- Normalization to saliency map

$$S_{face} = Norm(CS(GP(h(r',g')))) \tag{10}$$

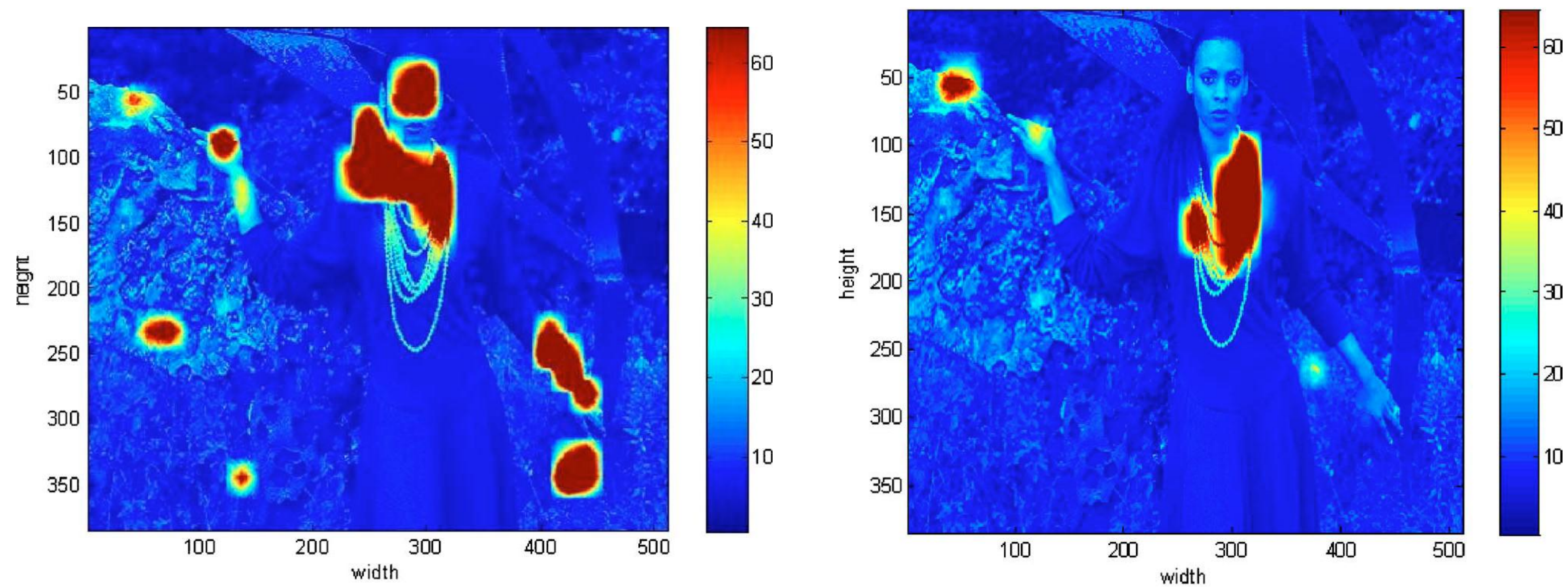
- Mixed model based on face detection
  - Combination of Itti`s model and gaussian face detection model

$$S_{MIX} = \alpha S_{Itti} + \ 1-\alpha \ \ S_{Face} \tag{11}$$

where α is constant, best result is $\alpha = 3/7$

– Comparison with Itti`s model and mixed model

- Human face in image
  - Reference image “I18” in In TID2008



(a) (b)

Fig. 9. Saliency maps for mixed model and Itti`s model on I18 reference image. (a) Saliency map from mixed model. (b) Saliency map from Itti`s model.

- nonhuman face in image
  - Reference image “I23” in In TID2008

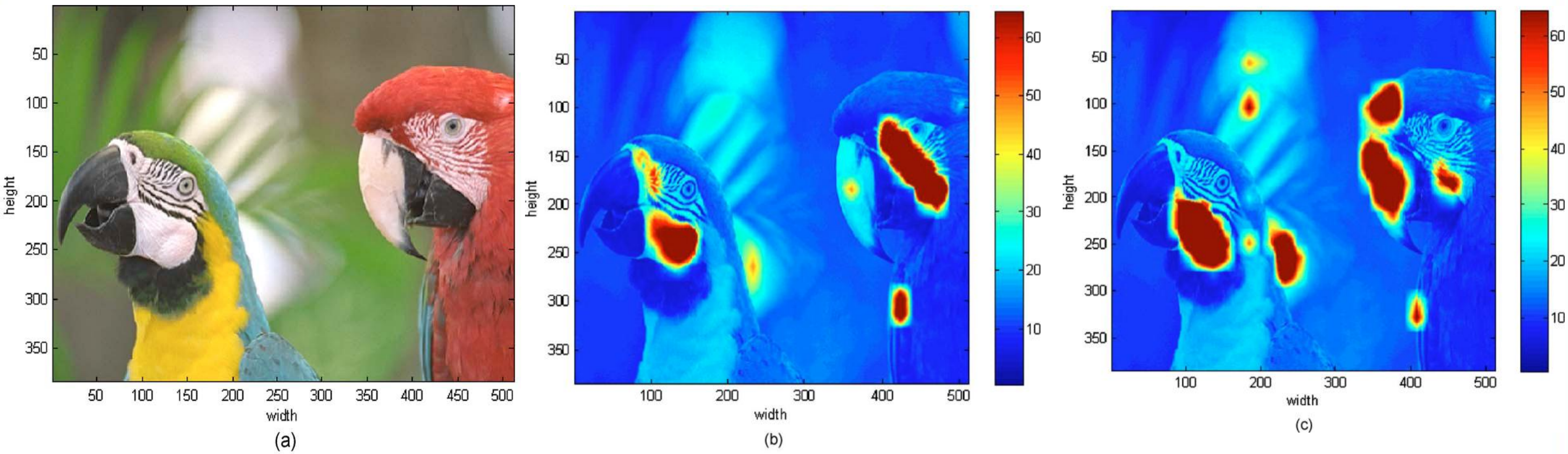


Fig. 10. Saliency map from mixed model and Itti`s model for I23 reference image. (a) I23 reference image (b) Saliency map from mixed model. (c) Saliency map from Itti`s model.

◆ Mixed model based on salient region

- Focus on saliency point
  - Defining saliency regions and neighboring filed associated given pixel
- Definition of saliency region
  - Computing binary mark metric

$$B_{i,j} = \begin{cases} 0 & S_{MAX}(i,j) < T_1 \\ 1 & otherwise \end{cases} \qquad (12)$$

where T_1 is experimental threshold that adaptive to average value of and $S_{MIX}(i,j)$ and $S_{MIX}(i,j)$ is saliency value computed from saliency map model considered and (i,j) is pixel position in image

- Computing block by block relative saliency degree

$$\phi = \begin{cases} false & if \sum_{i=1}^{8}\sum_{j=1}^{8} B_{block(I,J)}(i,j) < T_2 \\ true & otherwise \end{cases} \quad (13)$$

where T_2 is experimental threshold, and average of current block was used T_2; is pixel position in $Block(I,J)$

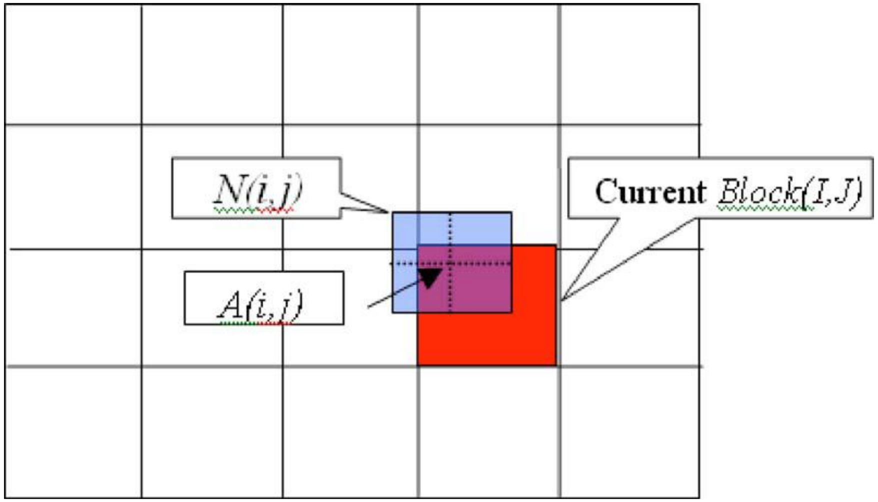


Fig. 11. Current block, current pixel, and its neighboring field.

– Relative saliency degree of current region and block

$$\rho_{Block}(I,J)=\frac{1}{\overline{\overline{S_{Grobal}}}}\left(\frac{1}{64}\sum_{i=1}^{8}\sum_{j=1}^{8}S_{MIX}(i,j)\right) \quad (14)$$

$$\rho_{region}(i,j)=\frac{\overline{\overline{S_{Local}}}}{S_{Global}} \quad (15)$$

with

$$\overline{S_{Local}}=\frac{1}{k\times k}\sum_{i=1}^{k}\sum_{j=1}^{k}S_{MIX}(i,j) \quad (16)$$

$$\overline{S_{Global}}=\frac{1}{M\times N}\sum_{i=1}^{M}\sum_{j=1}^{N}S_{MIX}(i,j) \quad (17)$$

– Relative saliency degree of current pixel and region

$$\rho_{pixel\_average}(i,j) = \max\left\{ \frac{S_{MIX}(i,j)}{\overline{S_{Local}}}, \frac{S_{MIX}(i,j)}{\overline{S_{Global}}} \right\} \tag{18}$$

$$\rho_{pixel\_\max}(i,j) = \frac{S_{MIX}(i,j)}{S_{Max\_Local}} \tag{19}$$

with

$$S_{Max\_Local} = \max\left\{ S_{MIX}(i,j) \middle| i \le k, j \le k \right\} \tag{20}$$

- Weighted saliency map

$$w_s(i,j) = \left\{ \max\left\{ \rho_{region}(i,j), \rho_{Block}(i,j) \right\} \middle| \rho_{region}(i,j) > T_3 \right\} \tag{21}$$

– Weighted saliency map of Fig. 9(a)

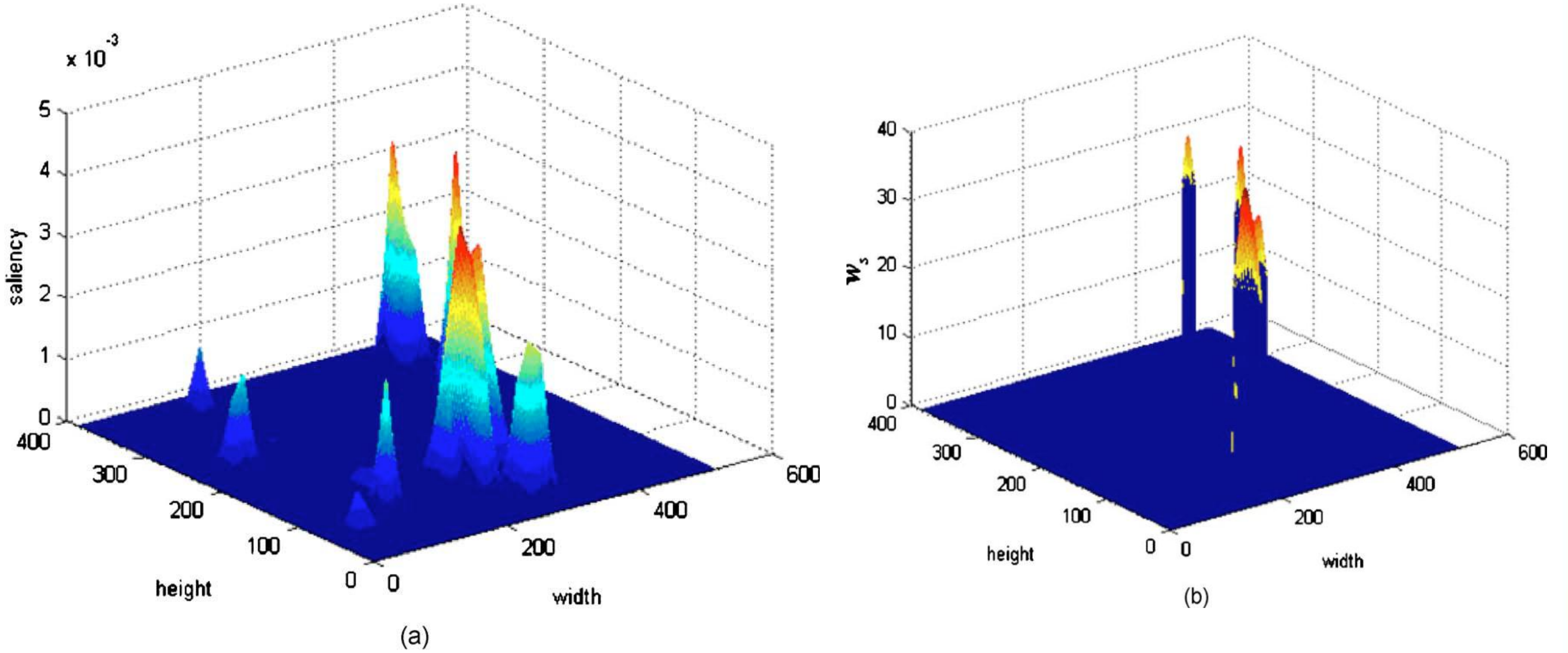


Fig. 12. Surface plot of saliency map and weighted saliency map w_s . (a) Surface plot of saliency map. (b) Surface plot of weighted saliency map w_s .

Image quality assessment weighted by saliency region

- Weight image different metrics
  - Application of weighted saliency map
    - Computing process

//for the pixels in a target block with 8×8

for $i=1:8$

  for $j=1:8$

    if ($\phi_{I,J}$ is false)

$$\delta_{PSNRHVS\_S}(i,j) = \delta(i,j) \cdot \left(\frac{\mathrm{CSF\ cof}(i,j)}{\mathrm{CSF\ cof}(i,j)+1} \right)$$

;

    Else

      if $[(\rho_{pixel\_max} > T_4)$ and $(\rho_{pixel\_average} > T_5)]$

$$\delta_{PSNRHVS\_S}(i,j) = \delta_{PSNRHVS}(i,j) \cdot w_s(i,j);$$

      Else

$$\delta_{PSNRHVS\_S}(i,j) = \delta_{PSNRHVS}(i,j);$$

      end

    end

  End

end

Experimental results and analysis

- Test image
  - Using TID2008 database
    - 1700 distorted image
      - 25 reference image, 17 type distortions, 4 level distortions
  - Using LIVE database
    - 779 distorted image
      - 5 type distortion image, 161 subjective experiments

Table 1. Distortion subsets in TID2008.

※ distortions that belong to a given subset are marked by +

No.	Distortion type	Noise	Noise2	Safe	Hard	Simple	Exotic	Exotic2	Full
1	Additive Gaussian noise	+	+	+	–	+	–	–	+
2	Different additive noise in color	–	+	–	–	–	–	–	+
3	Spatially correlated noise	+	+	+	+	–	–	–	+
4	Masked noise	–	+	–	+	–	–	–	+
5	High frequency noise	+	+	+	–	–	–	–	+
6	Impulse noise	+	+	+	–	–	–	–	+
7	Quantization noise	+	+	–	+	–	–	–	+
8	Gaussian blur	+	+	+	+	+	–	–	+
9	Image denoising	+	–	–	+	–	–	–	+
10	JPEG compression	–	–	+	–	+	–	–	+
11	JPEG2000 compression	–	–	+	–	+	–	–	+
12	JPEG transmission errors	–	–	–	+	–	–	+	+
13	JPEG2000 transmission errors	–	–	–	+	–	–	+	+
14	Non eccentricity pattern noise	–	–	–	+	–	+	+	+
15	Local blockwise distortions of different intensity	–	–	–	–	–	+	+	+
16	Mean shift (intensity shift)	–	–	–	–	–	+	+	+
17	Contrast change	–	–	–	–	–	+	+	+

Fig. 13. Examples of distortion in different subsets. (a) Original image. (b) Distortion 5: High frequency noise. (c) Distortion 8: Gaussian blur noise. (d) Distortion 12: jPEG transmission errors.

- Experimental results from TID2008
  - Comparison of image quality metrics
    - Using spearman correlation and kendall correlation coefficient

Table 2. Spearman correlation.

	PSNRHVS	PSNRHVS_S	Δ(%)	PSNRHVSM	PSNRHVSM_S	Δ(%)
Noise	0.917	0.914	−0.327	0.918	0.92	0.218
Noise2	0.933	0.863	−7.5	0.93	0.871	−6.344
Safe	0.932	0.92	−1.28	0.936	0.924	−1.282
Hard	0.791	0.814	2.908	0.783	0.816	4.215
Simple	0.939	0.933	−0.639	0.942	0.935	−0.743
Exotic	0.275	0.465	69.09	0.274	0.442	61.314
Exotic2	0.324	0.377	16.358	0.287	0.331	15.331
Full	0.594	0.622	4.71	0.559	0.595	6.44

Table 3. Kendall correlation.

	PSNRHVS	PSNRHVS_S	Δ(%)	PSNRHVSM	PSNRHVSM_S	Δ(%)
Noise	0.751	0.745	−0.799	0.752	0.752	0
Noise2	0.78	0.68	−12.82	0.771	0.689	−10.63
Safe	0.772	0.752	−2.59	0.778	0.757	−2.69
Hard	0.614	0.634	3.257	0.606	0.637	5.11
Simple	0.785	0.773	−1.52	0.789	0.777	−1.52
Exotic	0.195	0.313	60.51	0.194	0.294	51.55
Exotic2	0.238	0.254	6.72	0.21	0.22	4.76
Full	0.476	0.472	−0.8	0.449	0.455	1.34

Table 4. Spearman correlation comparison.

	WSNR	LINLAB	SNR	PSNR	PSNRHVSM	IFC	PSNRHVS	UQI	SSIM	PSNRHVS_S	PSNRHVSM_S
Noise	0.897	0.839	0.712	0.704	0.918	0.663	0.917	0.526	0.562	0.914	0.92
Noise2	0.908	0.853	0.687	0.612	0.93	0.743	0.933	0.599	0.637	0.863	0.871
Safe	0.921	0.859	0.699	0.689	0.936	0.775	0.932	0.638	0.632	0.92	0.924
Hard	0.776	0.761	0.646	0.697	0.783	0.736	0.791	0.759	0.812	0.814	0.816
Simple	0.931	0.877	0.794	0.799	0.942	0.817	0.939	0.784	0.769	0.933	0.935
Exotic	0.157	0.135	0.227	0.248	0.274	−0.269	0.275	0.292	0.385	0.465	0.442
Exotic2	0.059	0.033	0.29	0.308	0.287	0.276	0.324	0.546	0.594	0.377	0.331
Full	0.488	0.487	0.523	0.525	0.559	0.569	0.594	0.6	0.645	0.622	0.595

Table 5. Kendall correlation comparison.

	PSNR	SNR	LINLAB	WSNR	IFC	UQI	PSNRHVSM	SSIM	PSNRHVS	PSNRHVS_S	PSNRHVSM_S
Noise	0.501	0.512	0.652	0.714	0.477	0.363	0.752	0.388	0.751	0.745	0.752
Noise2	0.424	0.492	0.671	0.736	0.547	0.42	0.771	0.45	0.78	0.68	0.689
Safe	0.486	0.497	0.682	0.753	0.581	0.454	0.778	0.437	0.772	0.752	0.757
Hard	0.516	0.464	0.569	0.586	0.552	0.565	0.606	0.618	0.614	0.634	0.637
Simple	0.598	0.593	0.715	0.766	0.624	0.587	0.789	0.564	0.785	0.773	0.777
Exotic	0.178	0.154	0.084	0.107	−0.156	0.196	0.194	0.266	0.195	0.313	0.294
Exotic2	0.225	0.205	0.026	0.047	0.208	0.389	0.21	0.431	0.238	0.254	0.22
Full	0.369	0.374	0.381	0.393	0.426	0.435	0.449	0.468	0.476	0.472	0.455

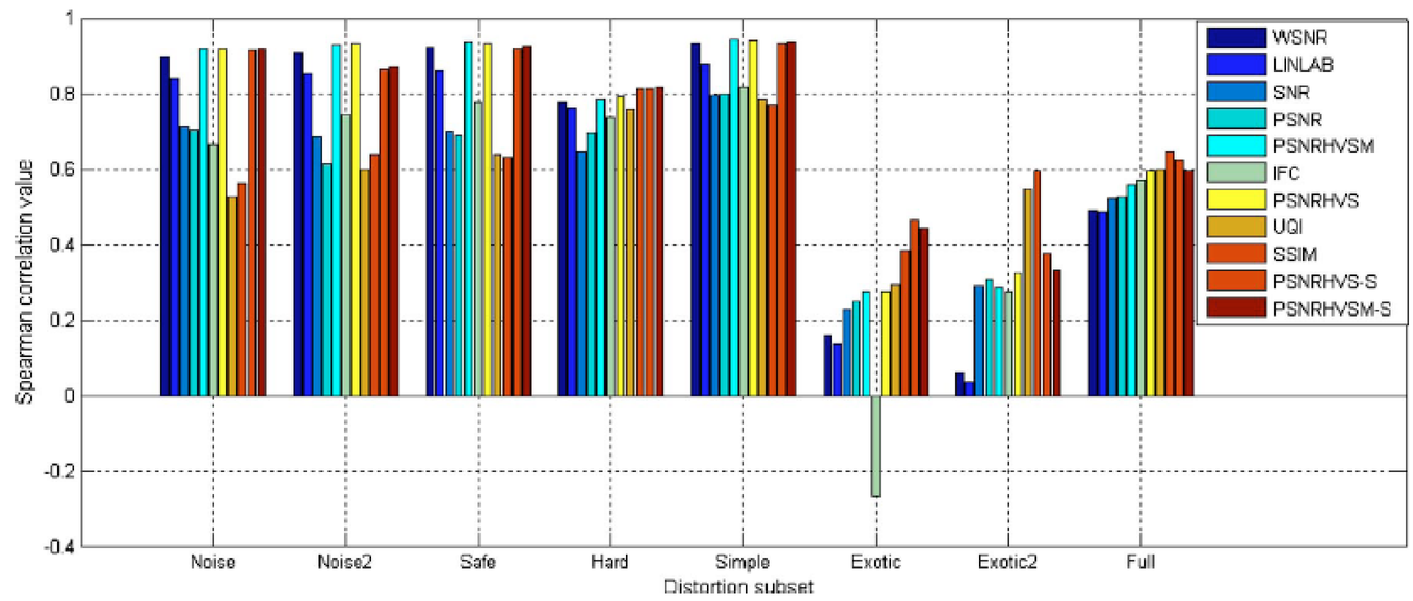


Fig 14. Spearman correlation Comparison.

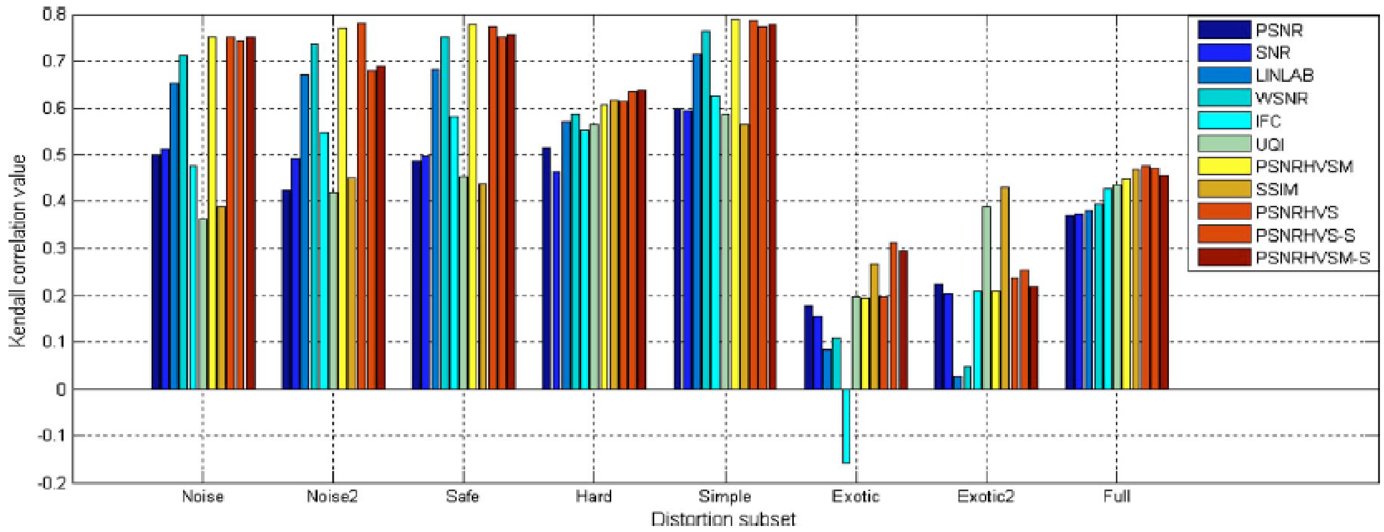


Fig 15. Kendall correlation comparison

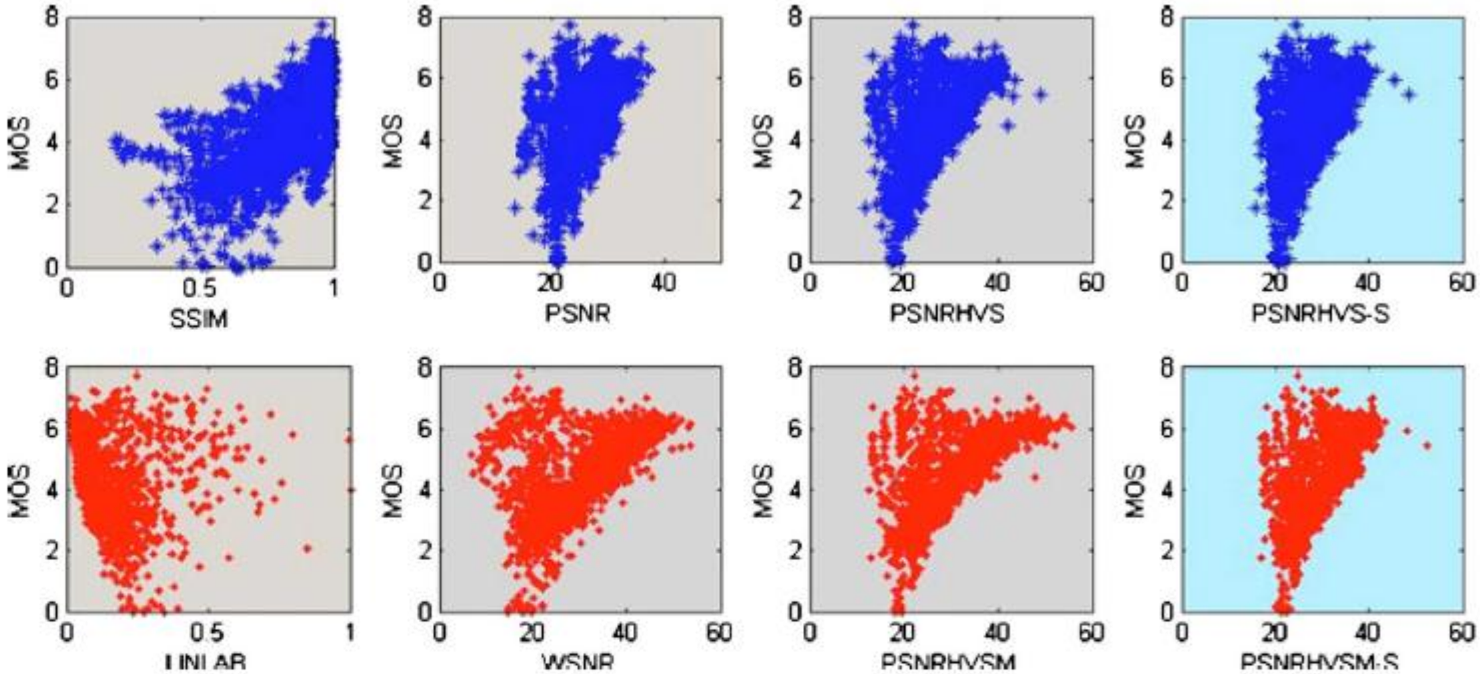


Fig 16. Scatter plots of the image quality assessment models, the plots with blue points are the results from the image quality assessment model based on weighted saliency map.

◆ Experimental results from LIVE database

Table 6. Spearman correlation and kendall correlation on LIVE database.

Correlation	SNR	PSNR	WSNR	UQI	IFC	SSIM	PSNRHVS_S	PSNRHVSM_S
Spearman	0.7811	0.8044	0.8479	0.802	0.8429	0.86	0.89	0.8963
Kendall	0.5922	0.6175	0.6883	0.6142	0.6677	0.7057	0.7179	0.7258

Conclusion

- Proposed method
  - Objective image quality assessment
    - Using weighted saliency map
      - Using mixed model
        - » Combination of Itti`s model and face detection model
      - Saliency region
        - » Including local contrast saliency and average
    - Enhancement of previous method
      - Contribution to perceived image quality

Table 7. Spearman correlation.

Distortion type	PSNRHVS_s			PSNRHVSM_s		
	$\rho_{region}(i,j)$	$\rho_{Block}(i,j)$	Max	$\rho_{region}(i,j)$	$\rho_{Block}(i,j)$	Max
Noise	0.913	0.913	0.914	0.920	0.920	0.92
Noise2	0.862	0.862	0.863	0.872	0.872	0.871
Safe	0.920	0.920	0.92	0.924	0.924	0.924
Hard	0.815	0.815	0.814	0.817	0.817	0.816
Simple	0.932	0.932	0.933	0.935	0.935	0.935
Exotic	0.463	0.463	0.465	0.440	0.440	0.442
Exotic2	0.377	0.377	0.377	0.331	0.331	0.331
Full	0.622	0.622	0.622	0.595	0.595	0.595

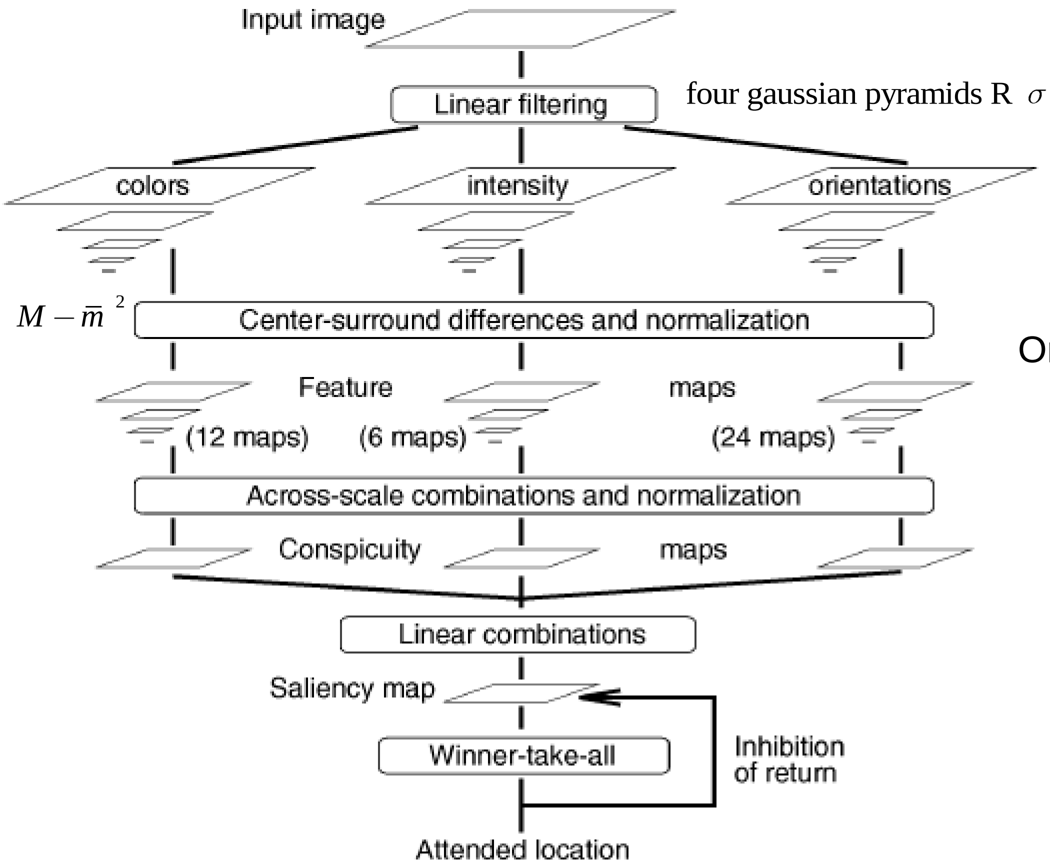
Table 8. Kendall correlation.

Distortion type	PSNRHVS_s			PSNRHVSM_s		
Distortion	$\rho_{region}(i,j)$	$\rho_{Block}(i,j)$	Max	$\rho_{region}(i,j)$	$\rho_{Block}(i,j)$	Max
Noise	0.743	0.743	0.745	0.752	0.752	0.752
Noise2	0.680	0.680	0.68	0.689	0.689	0.689
Safe	0.750	0.750	0.752	0.757	0.757	0.757
Hard	0.634	0.634	0.634	0.637	0.637	0.637
Simple	0.770	0.770	0.773	0.776	0.776	0.777
Exotic	0.313	0.313	0.313	0.293	0.293	0.294
Exotic2	0.255	0.255	0.254	0.220	0.220	0.22
Full	0.472	0.472	0.472	0.455	0.455	0.455

Table 9. PSNRHVS_S with different operator

	PSNRHVS_S				
	Spearman correlation			Kendall correlation	
Distortion type	$\rho_{Block}(i,j)$ without T_3	With T_3	Distortion type	$\rho_{Block}(i,j)$ Without T_3	With T_3
Noise	0.707	0.913	Noise	0.521	0.743
Noise2	0.657	0.862	Noise2	0.475	0.68
Safe	0.732	0.92	Safe	0.537	0.75
Hard	0.587	0.815	Hard	0.422	0.634
Simple	0.716	0.932	Simple	0.517	0.77
Exotic	0.228	0.463	Exotic	0.162	0.313
Exotic2	0.201	0.377	Exotic2	0.138	0.255
Full	0.446	0.622	Full	0.312	0.472

Itti`s saliency map



$(M-\bar{m})^2$

four gaussian pyramids R (σ), $G(\sigma)$, $B(\sigma)$, $Y(\sigma)$ $\quad \sigma \in [0..8]$

Intensity map $I(c,s) = |I(c) \ominus I(s)|$

Color map $\mathcal{RG}(c,s) = |(R(c) - G(c)) \ominus (G(s) - R(s))|$

$\mathcal{BY}(c,s) = |(B(c) - Y(c)) \ominus (Y(s) - B(s))|$

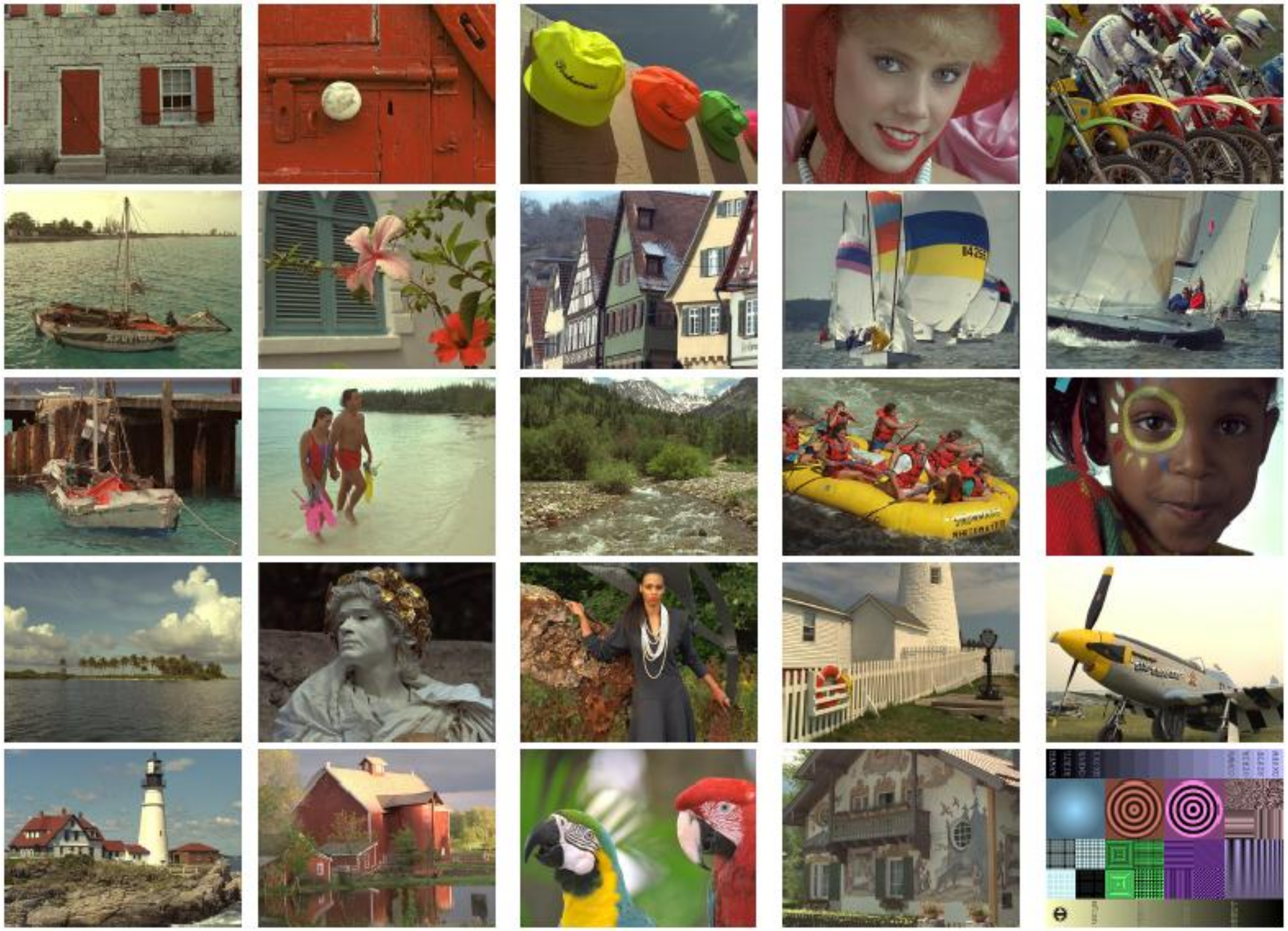
Orientation map $O(c,s,\theta) = |O(c,\theta) \ominus O(s,\theta)|$

$c \in \{2,3,4\}$

$s = c + \delta,\ \delta \in \{3,4\}$

General architecture of the Itti`s model

TID2008 database



Reference image

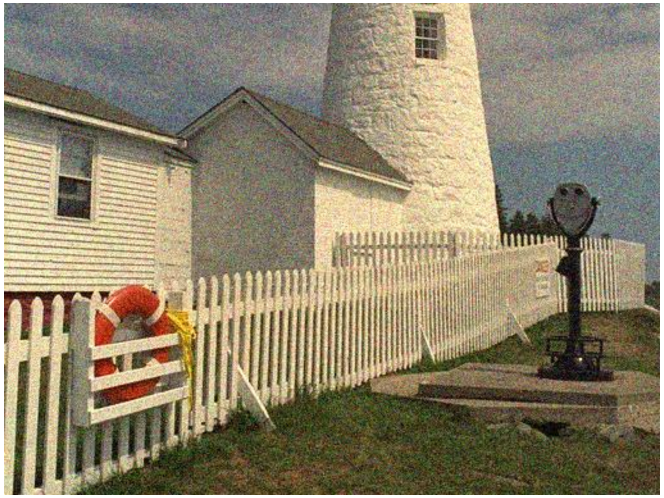
– Type of distortions image in TID2008 database

№	Type of distortion (four levels for each distortion)	Correspondence to practical situation	Accounted HVS peculiarities
1	Additive Gaussian noise	Image acquisition	Adaptivity, robustness
2	Additive noise in color components is more intensive than additive noise in the luminance component	Image acquisition	Color sensitivity
3	Spatially correlated noise	Digital photography	Spatial frequency sensitivity
4	Masked noise	Image compression, watermarking	Local contrast sensitivity
5	High frequency noise	Image compression, watermarking	Spatial frequency sensitivity
6	Impulse noise	Image acquisition	Robustness
7	Quantization noise	Image registration, gamma correction	Color, local contrast, spatial frequency
8	Gaussian blur	Image registration	Spatial frequency sensitivity
9	Image denoising	Image denoising	Spatial frequency, local contrast
10	JPEG compression	JPEG compression	Color, spatial frequency sensitivity
11	JPEG2000 compression	JPEG2000 compression	Spatial frequency sensitivity
12	JPEG transmission errors	Data transmission	Eccentricity
13	JPEG2000 transmission errors	Data transmission	Eccentricity
14	Non eccentricity pattern noise	Image compression, watermarking	Eccentricity
15	Local block-wise distortions of different intensity	Inpainting, image acquisition	Evenness of distortions
16	Mean shift (intensity shift)	Image acquisition	Light level sensitivity
17	Contrast change	Image acquisition, gamma correction	Light level, local contrast sensitivity

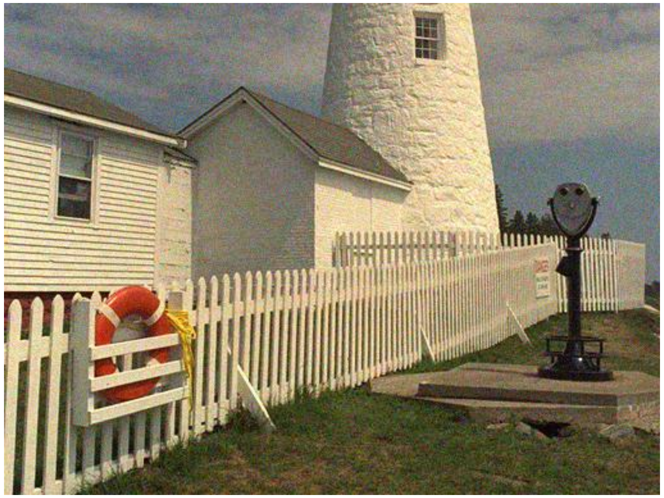
– Example of distortion type in TID2008 database

- 4 level distortion image for I19

Original image



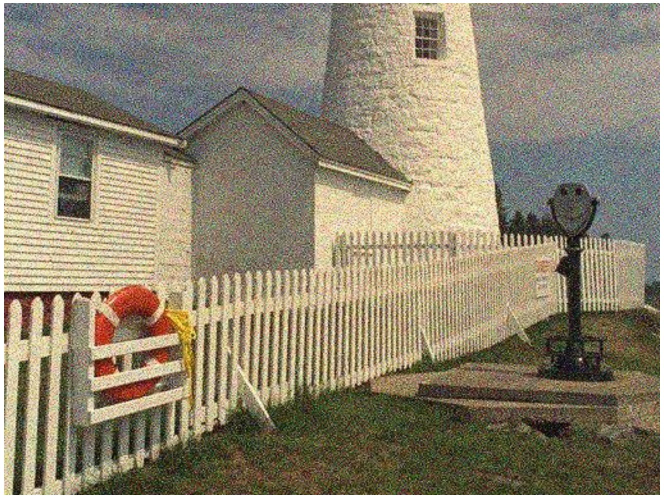
Additive Gaussian noise



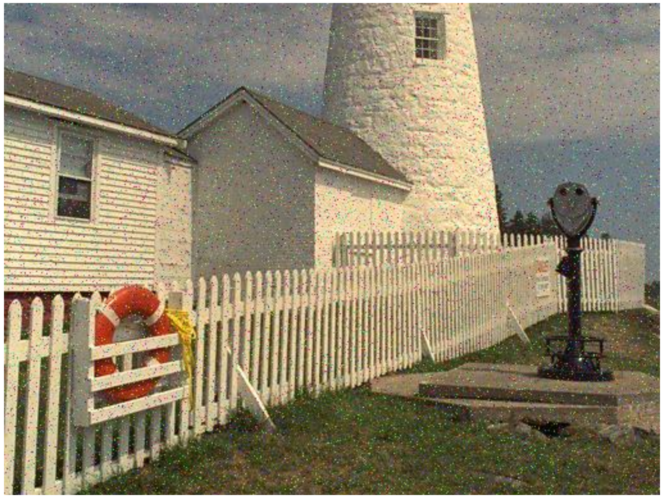
Different additive noise in color

Spatially correlated noise

Masked noise



High frequency noise



Impulse noise

Quantization noise

Gaussian blur

Image denoising

JPEG compression

JPEG2000 compression

JPEG transmission errors

JPEG2000 transmission errors

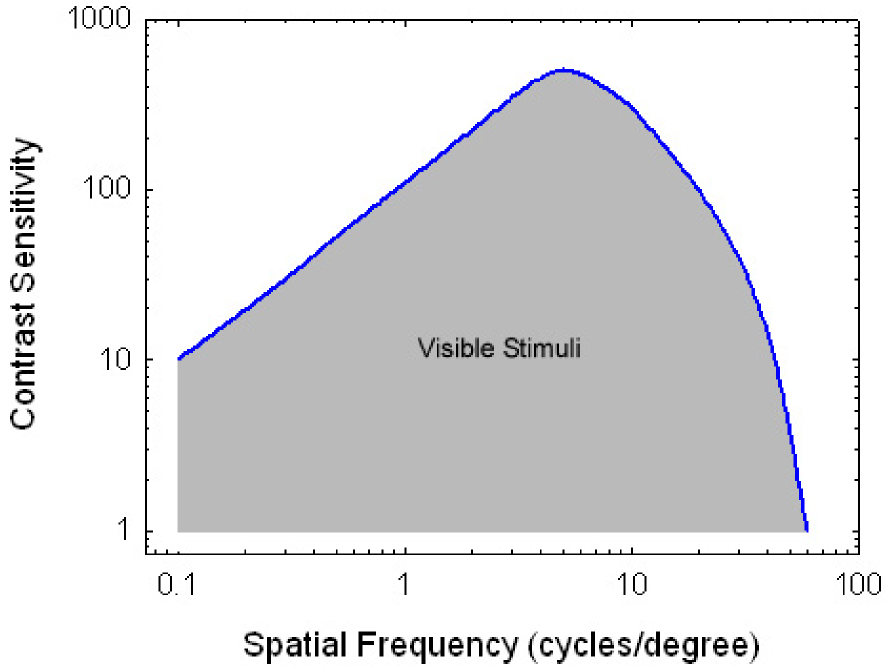
Non eccentricity pattern noise

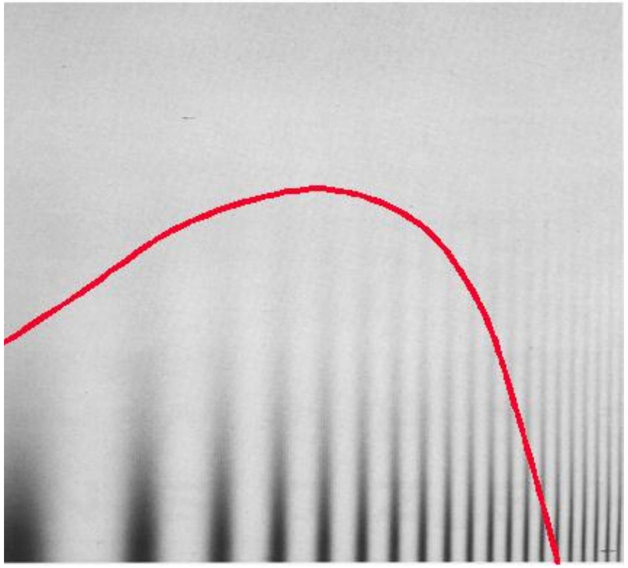
Local blockwise distortion of difference intensity

Mean shift (intensity shift)

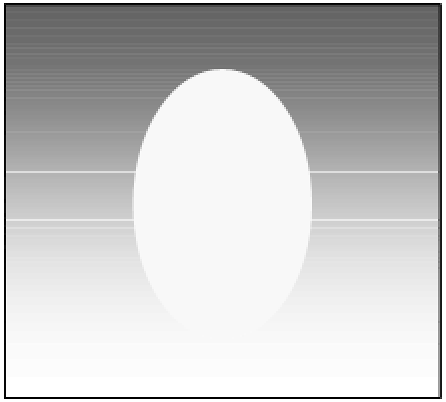
Contrast change

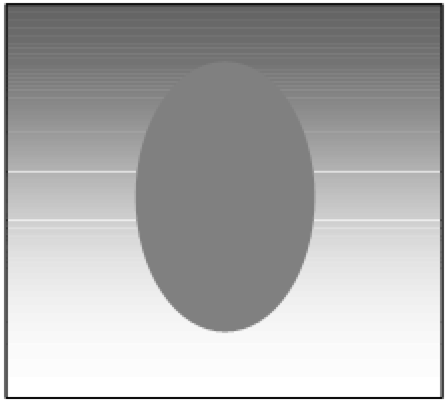
Contrast sensitivity function

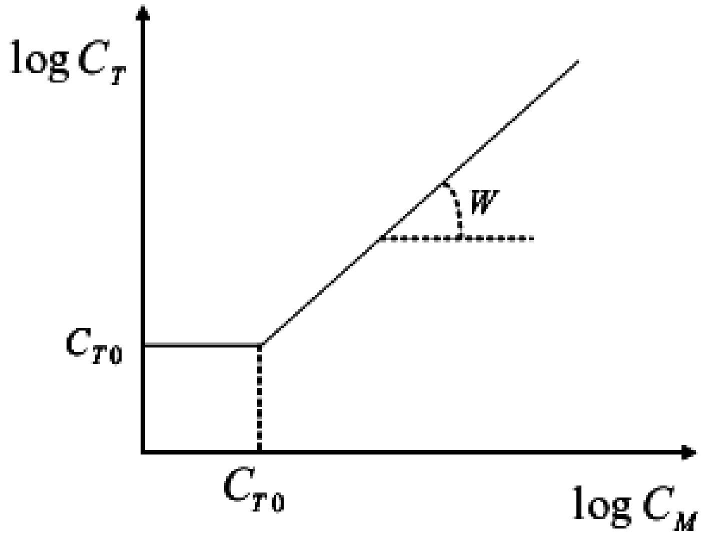




Contrast Masking







$$C_T = \begin{cases} C_{T0} & \text{if} \quad C_M < C_{T0} \\ C_{T0}\left(\dfrac{C_M}{C_{T0}}\right)^{\varepsilon} & \text{otherwise} \end{cases}$$