Université du Québec
École de technologie supérieure
Département de génie logiciel et des TI

A NOVEL APPROACH FOR COMPUTING AND POOLING STRUCTURAL SIMILARITY INDEX IN THE DISCRETE WAVELET DOMAIN

Authors:

Soroosh Rezazadeh , Stéphane Coulombe

ÉTS le génie pour l'industrie

2009 IEEE International Conference on Image Processing

AGENDA

1. **Introduction to Structural SIMilarity**
   - Structural SIMilarity (SSIM) Index
   - Multi-scale & Multi-level SSIM
2. **The Proposed Method**
   - Description of the Proposed Framework
   - Computational Complexity of the Algorithm
3. **Experimental Results**
4. **Conclusion**

Université du Québec
École de technologie supérieure
Département de génie logiciel et des TI

1. Introduction to Structural SIMilarity

ÉTS
le génie pour l'industrie

Full-Reference Image Quality Metrics

- **PSNR:**

$$\text{PSNR} = 10 log_{10} \frac{L^2}{\frac{1}{\text{N}} \sum_{m,n} \left(X(m,n) - Y(m,n) \right)^2} = 10 log_{10} \frac{L^2}{\text{MSE}}$$

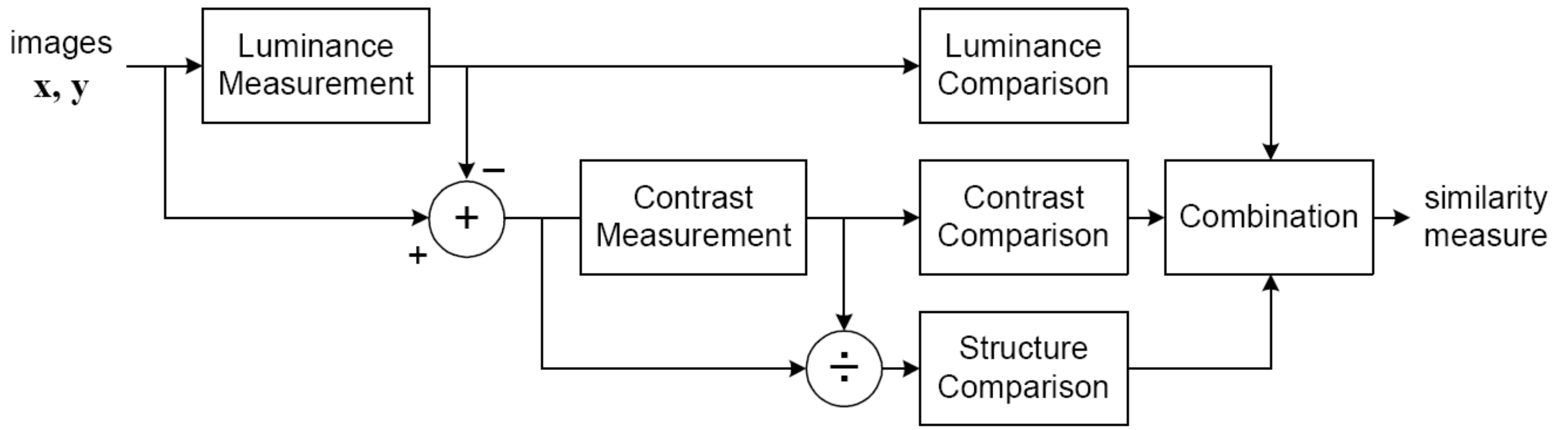
  - PSNR (MSE) is simple and has a clear physical meaning
  - PSNR does not accurately reflect the perceived image/video quality
  - A large gain in PSNR may result in a small improvement in visual quality

- **Structural SIMilarity (SSIM) Index [wang2004*]**

  - HVS is highly adapted to extract structural information from visual scenes
  - Structural similarity/distortion provides a good approximation of the perceptual quality
  - SSIM is mimicking the hypothesized functionality of the overall HVS

* Authors: Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli

Structural SIMilatity (SSIM) Index



$$\mathrm{SSIM}(\mathbf{x},\mathbf{y}) = l(\mathbf{x},\mathbf{y})\cdot c(\mathbf{x},\mathbf{y})\cdot s(\mathbf{x},\mathbf{y}) = \left(\frac{2\mu_x\mu_y + C_1}{\mu_x^2+\mu_y^2+C_1}\right)\cdot\left(\frac{2\sigma_x\sigma_y + C_2}{\sigma_x^2+\sigma_y^2+C_2}\right)\cdot\left(\frac{\sigma_{xy} + C_3}{\sigma_x\sigma_y+C_3}\right)$$

- SSIM properties:
- Symmetry SSIM(**x**,**y**) = SSIM(**y**,**x**)
- Boundedness: SSIM(**x**,**y**) ≤ 1
- Relatively accurate prediction: better than PSNR
- Good trade-off between accuracy and complexity

Multi-scale & Multi-level SSIM

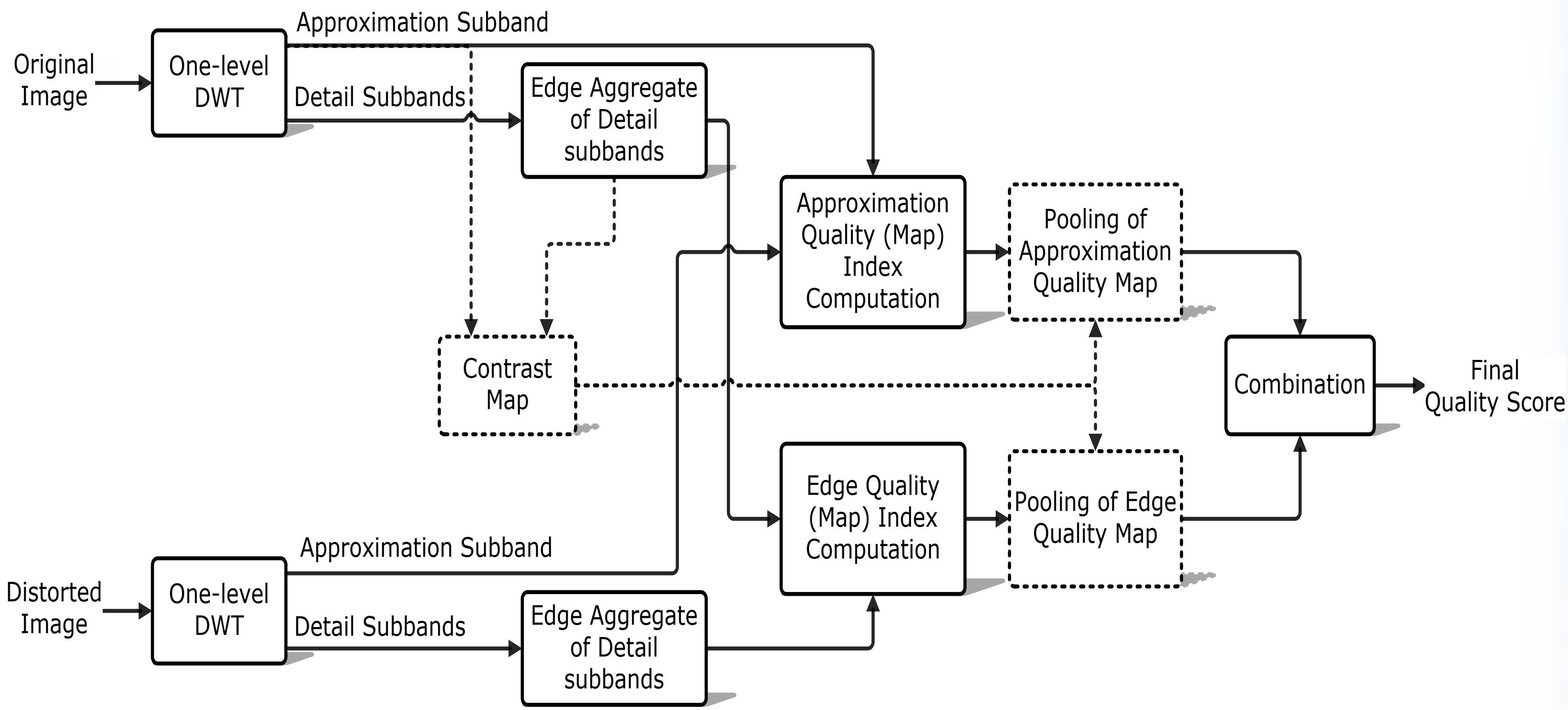
- **Multi-scale SSIM:**
  - incorporating image details at different resolutions by successive low-pass filtering
- **Multi-level SSIM**
  - Five-level decomposition using the Daubechies 9/7 wavelet
  - weighted mean of all SSIMs between corresponding subbands
- **Disadvantages:**
  - Finding the sensitivity of the HVS to different scales or subbands requires many experiments
    - effective combining of several subbands scores is also difficult
  - Multi-level decomposition of images makes the size of the approximation subband very small
    - no longer effective extraction of image statistics
  - Mean of the SSIM map is used in previous methods
    - various image areas have different impacts on the HVS

Université du Québec
École de technologie supérieure
Département de génie logiciel et des TI

2. The Proposed Method

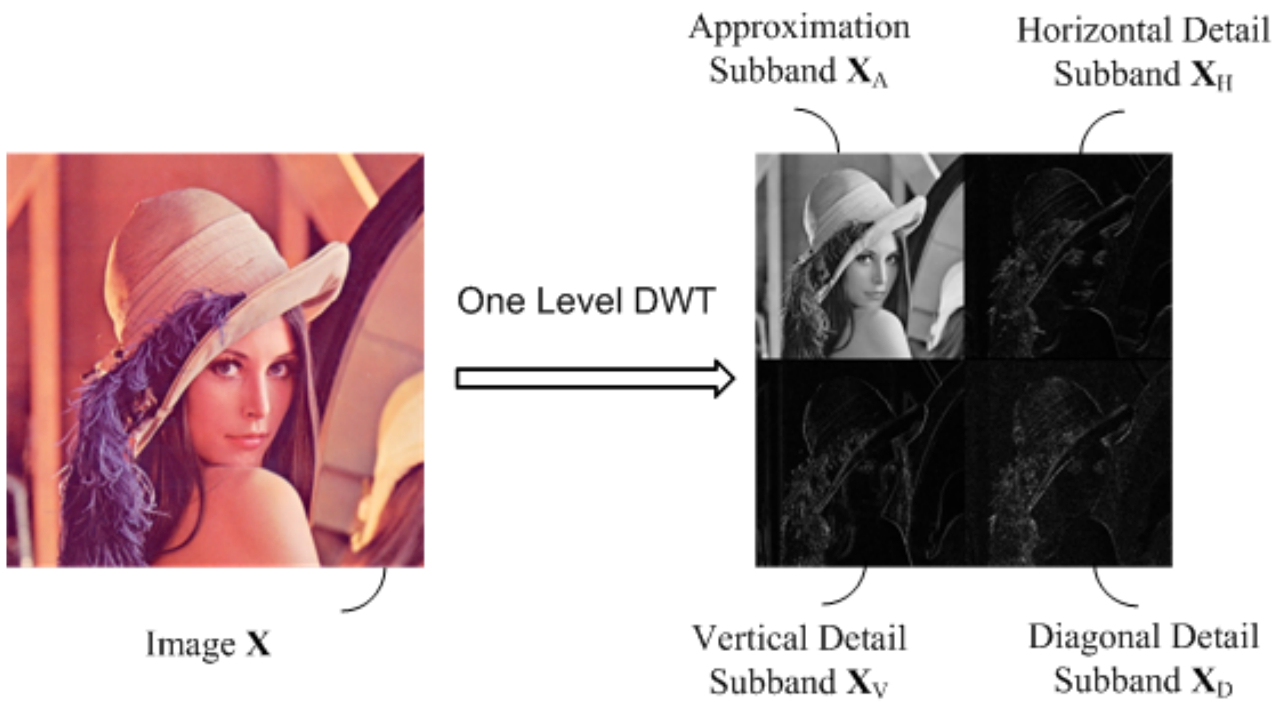
ÉTS
le génie pour l'industrie

The Proposed Framework



The Proposed Quality Metric

- With proposed framework, we can compute wavelet structural similarity
  - Excellent trade-off between complexity and accuracy
  - Computation performed in the discrete wavelet domain
    - Haar basis is used because of simplicity and symmetry



Edge Maps Computation and Comparison

- **Edge-map functions:**
  - mean square of detail subbands

$$\mathbf{X}_{\mathrm{E}}(m,n) = \frac{\mathbf{X}_{\mathrm{H}}^{2}(m,n) + \mathbf{X}_{\mathrm{V}}^{2}(m,n) + \mathbf{X}_{\mathrm{D}}^{2}(m,n)}{3}$$

$$\mathbf{Y}_{\mathrm{E}}(m,n) = \frac{\mathbf{Y}_{\mathrm{H}}^{2}(m,n) + \mathbf{Y}_{\mathrm{V}}^{2}(m,n) + \mathbf{Y}_{\mathrm{D}}^{2}(m,n)}{3}$$

  - only form image-edge structures and contain no luminance information

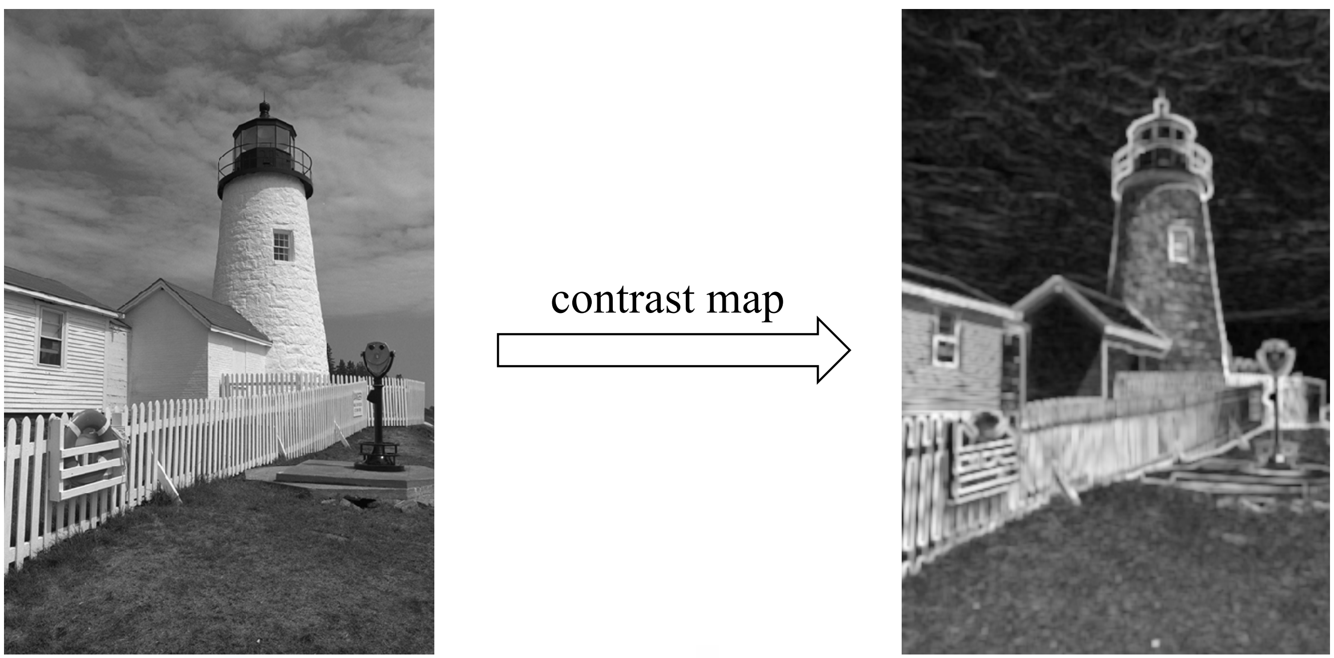
$$\mathrm{SSIM}_{\mathrm{E}}(\mathbf{x}_{\mathrm{E}}, \mathbf{y}_{\mathrm{E}}) = \frac{2\sigma_{x_{\mathrm{E}}, y_{\mathrm{E}}} + c}{\sigma_{x_{\mathrm{E}}}^{2} + \sigma_{y_{\mathrm{E}}}^{2} + c}$$

  - forming contrast map function for pooling

$$Contrast(\mathbf{x}_{\mathrm{E}}, \mathbf{x}_{A}) = (\mu_{x_{\mathrm{E}}} \sigma_{x_{A}}^{2})^{0.1}$$

The Contrast Map

- A novel contrast map for discrete wavelet domain pooling
  - Pixels near edges are given more importance
  - High energy regions are likely to contain more information
  - Brighter sample values in the contrast map indicate more important image structures



Final Quality Score Computation

- **Weighted pooling of approximation and edge maps**
  - for approximation map

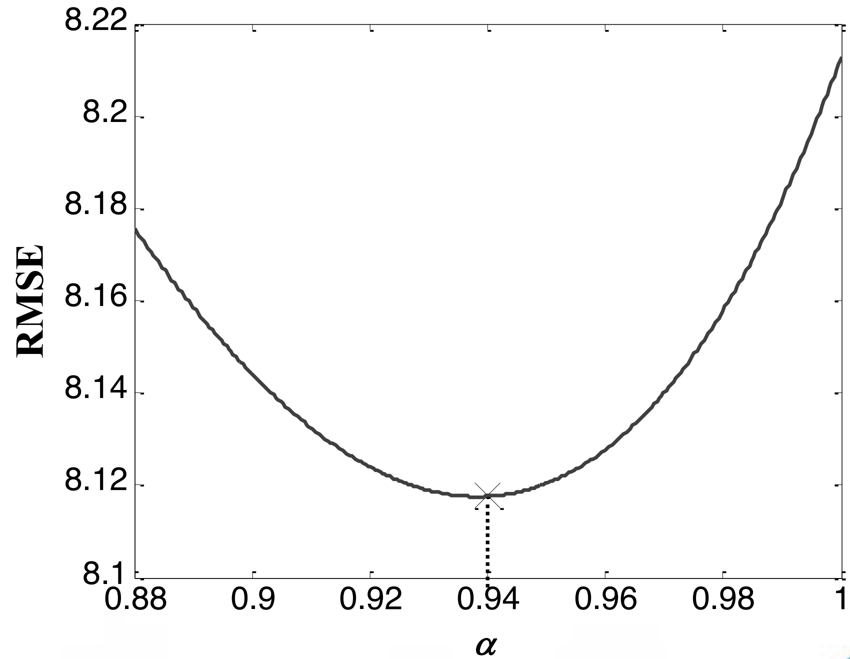
$$S_A = \frac{\sum_{j=1}^{M} Contrast(\mathbf{x}_{E,j}, \mathbf{x}_{A,j}) \cdot \mathrm{SSIM}_A(\mathbf{x}_{A,j}, \mathbf{y}_{A,j})}{\sum_{j=1}^{M} Contrast(\mathbf{x}_{E,j}, \mathbf{x}_{A,j})}$$

  - Pooling of edge map is similar to approximation map

- **Final quality score**

$$WSSI(\mathbf{X}, \mathbf{Y}) = \alpha S_A + (1-\alpha) S_E$$

$$0 < \alpha \leq 1$$



Computational Complexity of the Algorithm

- The algorithm provides excellent trade-off between accuracy and complexity
  - resolution of the approximation subband and edge map is a quarter of that of the original image
  - accurate local statistics can be extracted with a small sliding window size of 4×4.
  - Discrete wavelet transform has fast implementation
    - a simple wavelet (Haar wavelet) is used to reduce complexity
  - local statistics calculated for approximation and edge maps are used to form the contrast map.
    - no new parameters are needed for obtaining contrast map
  - luminance comparison part in the approximation SSIM map can be ignored (just 0.03% reduction in accuracy)
- The running time for calculating the WSSI is, on average, about 0.65 of running time for spatial domain SSIM

Université du Québec
École de technologie supérieure
Département de génie logiciel et des TI

3. Experimental Results

ÉTS
le génie
pour l'industrie





Simulation Results

- Performance evaluation carried out on LIVE Image Quality Assessment Database Release 2
  - 779 distorted images and 5 types of distortions
    - JPEG, JPEG2000, white noise, Gaussian blurring, fast fading channel model
- Three performance measures adopted
  - Correlation coefficients (CC) for prediction accuracy
  - Root mean square error (RMSE) for prediction consistency
  - Spearman rank order correlation coefficient (ROCC) for prediction monotonicity
    - Nonlinear regression between DMOS and output values of models before calculating performance measures
- Comparison with other metrics :
  - PSNR, spatial mean SSIM, DWT-SSIM, and visual information fidelity (VIF)

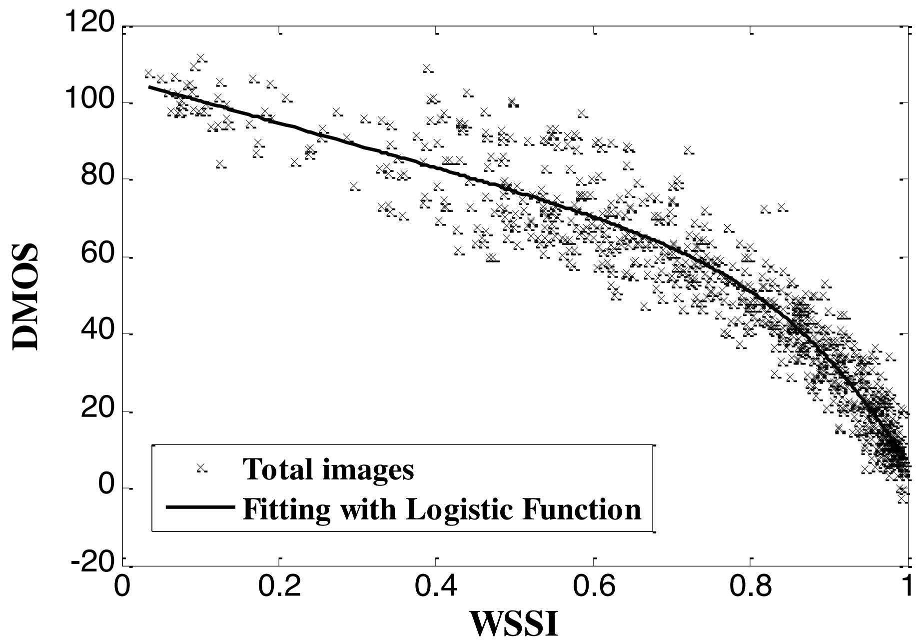
Simulation Results

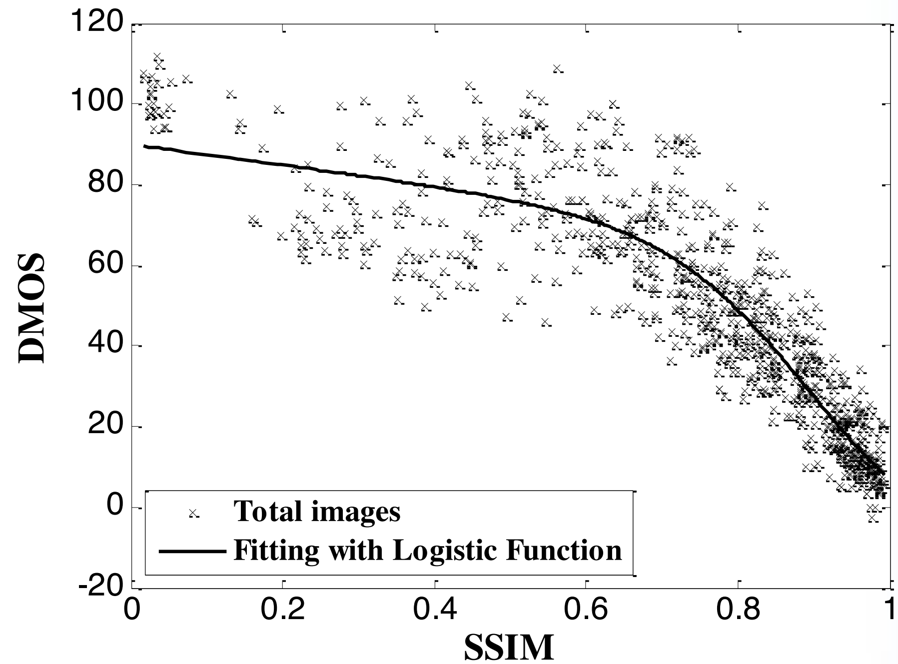
Model	CC	RMSE	ROCC
PSNR	0.8701	13.4685	0.8756
Mean SSIM (spatial)	0.9041	11.6736	0.9104
DWT-SSIM	0.9346	9.7201	0.9346
VIF	0.9593	7.7122	0.9635
Mean SSIM$_A$	0.9412	9.2270	0.9441
WSSI	**0.9548**	**8.1176**	**0.9586**

- WSSI uses: Haar wavelet, 4×4 Gaussian sliding window, and $\alpha = 0.94$
- SSIM$_A$ performs well as a separate quality metric
- ❑ The first-level approximation subband contains most of the image information
- Performance of WSSI is close to VIF but with much less complexity

Simulation Results: Scatter Plots

- Scatter plots of DMOS versus model prediction for all 779 distorted images
  - WSSI is more consistent with the subjective scores than the spatial mean SSIM



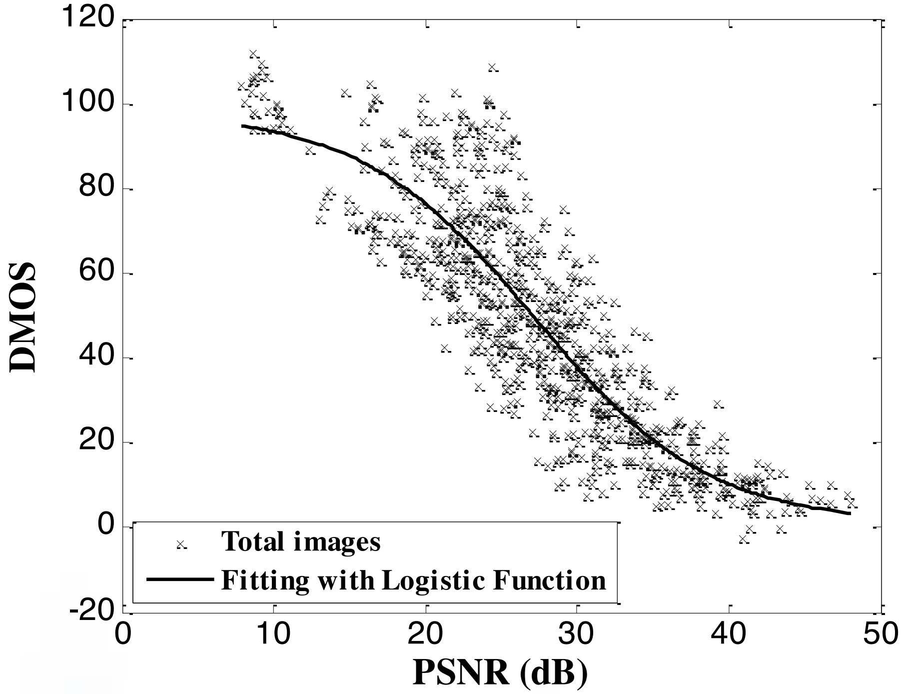


Simulation Results

➢ Wavelet basis has very little effect on performance of the algorithm:

- Contrast map is image-adaptive and changes with the wavelet
- For Daubechies 9/7 wavelet & previous parameters (α and window size) :

CC= 0.9489 , RMSE=8.6232 , ROCC=0.9529



Université du Québec
École de technologie supérieure
Département de génie logiciel et des TI

4. Conclusion

ÉTS
le génie pour l'industrie

Conclusion

- A DWT-based framework for computing image quality
  - excellent trade-off between accuracy and complexity

- The performance of wavelet structural similarity index (WSSI) is close to VIF index
  - The complexity of WSSI is less than spatial-domain SSIM index
  - The complexity of WSSI is much less than VIF index

- A novel contrast map for pooling of distortion map in the discrete wavelet domain

- First-level approximation subband has a significant role in improving quality assessment
  - Haar wavelet gives slightly better results compared to other wavelet bases

THANKS FOR YOUR ATTENTION