Deep Learning for Image Super-resolution: A Survey

I: Dataset for Super-resolution

STT	Dataset	Amount	Avg.Resolution	Format	Category Keywords
1	BSDS300	300	(435, 367)	JPG	animal, building, food, landscape, people, plant, etc.
2	BSDS500	500	(432, 370)	JPG	animal, building, food, landscape, people, plant, etc.
3	DIV2K	1000	(1972, 1437)	PNG	environment, flora, fauna, handmade object, people, scenery, etc.
4	General-100	100	(435, 381)	BMP	animal, daily necessity, food, people, plant, texture, etc.
5	Manga109	109	(826, 1169)	PNG	manga volume
6	OutdoorScene	10624	(553, 440)	PNG	animal, building, grass, mountain, plant, sky, water.
7	PIRM	200	(617, 482)	PNG	environments, flora, natural scenery, objects, people, etc.
8	Set5	5	(313, 336)	PNG	baby, bird, butterfly, head, woman
9	Set14	14	(492, 446)	PNG	humans, animals, insects, flowers, vegetables, comic, slides, etc.
10	T91	91	(264, 204)	PNG	car, flower, fruit, human face, etc.
11	Urban100	100	(984, 797)	PNG	architecture, city, structure, urban, etc.

II: Image Quality Assessment

1. Peak Signal-to-Noise Ratio (PSNR) (**code**)

   - PSNR is defined via the maximum pixel value (denoted as L) and the mean squared error (MSE) between images.

$$\text{PSNR} = 10 \cdot \log_{10}(\frac{L^2}{\frac{1}{N} \sum_{i=1}^{N}(I(i) - \hat{I}(i))^2}),$$

2. Structural Similarity(**code**)

   - The structural similarity index (SSIM) is proposed for measuring the structural similarity between images, based on independent comparisons in terms of luminance, contrast, and

structures.

- For an image I with N pixels, the luminance µI and contrast σI are estimated as the mean and standard deviation of the image intensity, respectively.

$$\mu_I = \frac{1}{N} \sum_{i=1}^{N} I(i) \quad \sigma_I = \left(\frac{1}{N-1} \sum_{i=1}^{N} (I(i) - \mu_I)^2 \right)^{\frac{1}{2}}$$

- where I(i) represents the intensity of the i-th pixel of image I. And the comparisons on luminance and contrast, denoted as Cl(I, ˆI) and Cc(I, ˆI) respectively, are given by:

$$C_l(I, \hat{I}) = \frac{2\mu_I \mu_{\hat{I}} + C_1}{\mu_I^2 + \mu_{\hat{I}}^2 + C_1},$$

$$C_c(I, \hat{I}) = \frac{2\sigma_I \sigma_{\hat{I}} + C_2}{\sigma_I^2 + \sigma_{\hat{I}}^2 + C_2},$$

  - where C1 = (k1L)^2, C2 = (k2L)^2 are constants for avoiding instability, k1 << 1 and k2 << 1

- Structure comparison function Cs(I, ˆI) is defined as:

$$\sigma_{I\hat{I}} = \frac{1}{N-1} \sum_{i=1}^{N} (I(i) - \mu_I)(\hat{I}(i) - \mu_{\hat{I}}),$$

$$C_s(I, \hat{I}) = \frac{\sigma_{I\hat{I}} + C_3}{\sigma_I \sigma_{\hat{I}} + C_3},$$

  - where σI,Iˆ is the covariance between I and ˆI, and C3 is a constant for stability

- Finally, the SSIM is given by:

$$\text{SSIM}(I, \hat{I}) = [C_l(I, \hat{I})]^\alpha [C_c(I, \hat{I})]^\beta [C_s(I, \hat{I})]^\gamma,$$

  - where α, β, γ are control parameters for adjusting the relative importance.

3. The multi-scale structural similarity (MS-SSIM)(**code**)

- At the j-th scale, the contrast comparison (2) and the structure comparison (3) are calculated and denoted as cj (x, y) and sj (x, y), respectively. The luminance comparison (1) is computed only at Scale M and is denoted as lM(x, y).

$$\text{SSIM}(\mathbf{x}, \mathbf{y}) = [l_M(\mathbf{x}, \mathbf{y})]^{\alpha_M} \cdot \prod_{j=1}^{M} [c_j(\mathbf{x}, \mathbf{y})]^{\beta_j} [s_j(\mathbf{x}, \mathbf{y})]^{\gamma_j}$$

4. Mean Opinion Score (MOS)(**code**)

- Mean opinion score (MOS) testing is a commonly used subjective IQA method, where human raters are asked to assign perceptual quality scores to tested images. Typically, the scores are

from 1 (bad) to 5 (good). And the final MOS is calculated as the arithmetic mean over all ratings.

5. Mean Squared Error (MSE)(**code**)

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2.$$

   ○

6. Root Mean Sqaured Error (RMSE) (**code**)

$$\text{RMSD}(\hat{\theta}) = \sqrt{\text{MSE}(\hat{\theta})} = \sqrt{\text{E}((\hat{\theta} - \theta)^2)}.$$

   ○

7. Natural Image Quality Evaluator (NIQE)

8. The Feature Similarity (FSIM)

9. Information Fidelity Criterion (IFC)

10. Noise Quality Measure (NQM)

11. Universal Quality Image Index (UQI)(**code**)

   ○ Let x = {xi | i = 1,2, ... N} and y = {yi | i = 1,2, .. N} be the original and the test image signals, respectively. The proposed quality index is defined as.

$$Q = \frac{4\,\sigma_{xy}\,\bar{x}\,\bar{y}}{(\sigma_x^2 + \sigma_y^2)\left[(\bar{x})^2 + (\bar{y})^2\right]} \ ,$$

where

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i\,, \qquad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} y_i\,,$$

$$\sigma_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2\,, \quad \sigma_y^2 = \frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2\,,$$

$$\sigma_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})\,.$$

-
  - The dynamic range of Q is [-1, 1]. The best value 1 is achieved if and only if yi = xi.

12. Relative Average Spectral Error (RASE)(**code**)

- In order to estimate the global spectral quality of the merged images, we have used the following parameters.The relative average spectral error (RASE) index expressed as a percentage. It characterizes the average performance of the method in the spectral bands considered

$$RASE = \frac{100}{M}\sqrt{\frac{1}{N}\sum_{i=1}^{N} RMSE(B_i)^2}$$

-
  - Where M is the mean radiance of the N spectrl bands (Bi) of the original MS image, and the RMSE is the root mean square error computed in following the expression:

$$RMSE^2(B_i) = bias^2(B_i) + SDD^2(B_i).$$

-
  - Standard deviation of the difference image (SDD), relative to the mean of the original image, expressed as percentage. The lower the value of this parameter, the better the spectral quality of the merged image.

- Difference between the means of the original and the merged images (bias), in radiance as well as its value relative to the mean of the original image(RM) The smaller these differences are, the better the spectral quality of the merged image. They should be as close to 0 as possible

13. Erreur Relative Globale Adimensionnelle de Synthèse (ERGAS)**(code)**

   - The erreur relative globale adimensionnelle de synthèse (ERGAS) index (or relative global dimensional synthesis error) in the fusion

$$\text{ERGAS} = 100\frac{h}{l}\sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\frac{\text{RMSE}^2(B_i)}{M_i^2}\right)}$$

   -
   - Where h is the resolution of the high spatial resolution image and l the resolution of the low spatial resolution image and Mi the mean radiance of each spectral band involved in the fusion.

14. Spatial Correlation Coefficient (SCC)**(code)**

15. Block Sensitive - Peak Signal-to-Noise Ratio (PSNR-B)**(code)**

   - Propose a new block-sensitive image quality metric which we term peak signal-to-noise ratio including blocking effects (PSNR-B).As the quantization step size increases, blocking artifacts generally become more conspicuous. Blocking artifacts are gray level discontinuities at block boundaries, which are ordinarily oriented horizontally and vertically. They arise from poor representation of the block luminance levels near the block boundaries. Block-base 8x8 blocks.
   - We, therefore, consider blocking artifacts that occur along the horizontal and vertical orientations. Let Nh and Nv be the horizontal and vertical dimensions of the Nh x Nv image I.

$$\eta = \begin{cases} \frac{\log_2 B}{\log_2(\min(N_H, N_V))}, & \text{if } D_B(\mathbf{y}) > D_B^C(\mathbf{y}) \\ 0, & \text{otherwise} \end{cases}$$

   -
   - Let H be the set of horizontal neighboring pixel pairs in I. Let Hb ⊂ H be the set of horizontal neighboring pixel pairs that lie across a block boundary. Let Hcb be the set of horizontal neighboring pixel pairs, not lying across a block boundary, i.e., Hcb = H - Hb.Similarly, let V be the set of vertical neighboring pixel pairs, and Vb be the set of vertical neighboring pixel pairs lying across block boundaries. Let Vcb be the set of vertical neighboring pixel pairs not lying across block boundaries, i.e., . Vcb = V - Vb.

y_1	y_2	y_3	y_4	y_5	y_6	y_7	y_8
y_9	y_{10}	y_{11}	y_{12}	y_{13}	y_{14}	y_{15}	y_{16}
y_{17}	y_{18}	y_{19}	y_{20}	y_{21}	y_{22}	y_{23}	y_{24}
y_{25}	y_{26}	y_{27}	y_{28}	y_{29}	y_{30}	y_{31}	y_{32}
y_{33}	y_{34}	y_{35}	y_{36}	y_{37}	y_{38}	y_{39}	y_{40}
y_{41}	y_{42}	y_{43}	y_{44}	y_{45}	y_{46}	y_{47}	y_{48}
y_{49}	y_{50}	y_{51}	y_{52}	y_{53}	y_{54}	y_{55}	y_{56}
y_{57}	y_{58}	y_{59}	y_{60}	y_{61}	y_{62}	y_{63}	y_{64}

Fig. 2. Example for illustration of pixel blocks.

-
- If B is the block size, then

$$N_{H_B} = N_V \left(\frac{N_H}{B}\right) - 1$$
$$N_{H_B^C} = N_V(N_H - 1) - N_{H_B}$$
$$N_{V_B} = N_H \left(\frac{N_V}{B}\right) - 1$$
$$N_{V_B^C} = N_H(N_V - 1) - N_{V_B}.$$

-
- The sets of pixel pairs in this example are.

$$\mathcal{H}_B = \{(y_4, y_5), (y_{12}, y_{13}), \ldots, (y_{60}, y_{61})\}$$
$$\mathcal{H}_B^C = \{(y_1, y_2), (y_2, y_3), (y_3, y_4), (y_5, y_6)$$
$$\ldots, (y_{63}, y_{64})\}$$
$$\mathcal{V}_B = \{(y_{25}, y_{33}), (y_{26}, y_{34}), \ldots, (y_{32}, y_{40})\}$$
$$\mathcal{V}_B^C = \{(y_1, y_9), (y_9, y_{17}), (y_{17}, y_{25}), (y_{33}, y_{41})$$
$$\ldots, (y_{56}, y_{64})\}.$$

-
- Then define the mean boundary pixel squared difference (Db) and the mean nonboundary pixel squared difference (Dcb) for image y to be

$$D_B(\mathbf{y}) = \frac{\sum\limits_{(y_i, y_j) \in \mathcal{H}_B} (y_i - y_j)^2 + \sum\limits_{(y_i, y_j) \in \mathcal{V}_B} (y_i - y_j)^2}{N_{H_B} + N_{V_B}}$$

$$D_B^C(\mathbf{y}) = \frac{\sum\limits_{(y_i, y_j) \in \mathcal{H}_B^C} (y_i - y_j)^2 + \sum\limits_{(y_i, y_j) \in \mathcal{V}_B^C} (y_i - y_j)^2}{N_{H_B^C} + N_{V_B^C}}.$$

-
- Generally, as the quantization step size increases, Db will increase relative to Dcb, and blocking artifacts will become more visible. Also define the blocking effect factor

$$\mathrm{BEF}(\mathbf{y}) = \eta \cdot \left[D_B(\mathbf{y}) - D_B^C(\mathbf{y}) \right]$$

-

$$\eta = \begin{cases} \frac{\log_2 B}{\log_2(\min(N_H, N_V))}, & \text{if } D_B(\mathbf{y}) > D_B^C(\mathbf{y}) \\ 0, & \text{otherwise} \end{cases}$$

- Where
  emphasizes the BEF as a function of block size. The assumption here is that the visibility of blocking effects increases with block size.
- Let Dbk, Dcbk, BEFk and ηk modify for block size Bk. Then

$$\mathrm{BEF}_k(\mathbf{y}) = \eta_k \cdot \left[D_{B_k}(\mathbf{y}) - D_{B_k}^C(\mathbf{y}) \right]$$

-
- The BEF over all block sizes is defined as

$$\mathrm{BEF}_{\mathrm{Tot}}(\mathbf{y}) = \sum_{k=1}^{K} \mathrm{BEF}_k(\mathbf{y}).$$

-
- The mean-squared error including blocking effects (MSE-B) for reference image x and test image y is then defined as the sum of the MSE(x, y) and BEFtot. The MSE term measures the distortion between the reference image x and the test image y, while the BEF term specifically measures the amount of blocking artifacts just using the test image y.
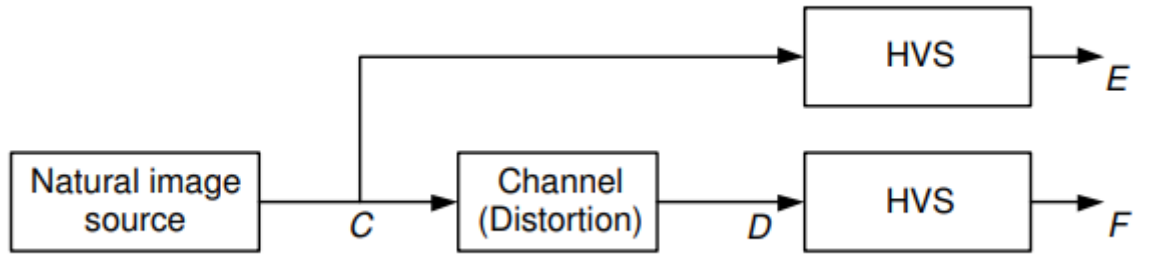
$$\mathrm{MSE}\text{-}B(\mathbf{x}, \mathbf{y}) = \mathrm{MSE}(\mathbf{x}, \mathbf{y}) + \mathrm{BEF}_{\mathrm{Tot}}(\mathbf{y}).$$

-
- Finally, we propose the PSNR-B as

$$\mathrm{PSNR}\text{-}B(\mathbf{x}, \mathbf{y}) = 10 \log_{10} \frac{255^2}{\mathrm{MSE}\text{-}B(\mathbf{x}, \mathbf{y})}.$$

-
- The associated quality index PSNR-B is obtained from the MSE-B by a logarithmic function, as is the PSNR from the MSE. The PSNR-B is attractive since it is specific for assessing image quality, specifically the severity of blocking artifacts.

16. [Visual Information Fidelity (VIF)](**code**)

- ○
- ○ The visual information fidelity (VIF) measure that we propose in this paper is derived from a quantification of two mutual information quantities: the mutual information between the input and the output of the HVS channel when no distortion channel is present (we call this the reference image information) and the mutual information between the input of the distortion channel and the output of the HVS channel for the test image. We discuss the components of the proposed method in this section.
- ○ A Gaussian Scale Mixtures (GSM) is a random field (RF) that can be expressed as a product of two independent RFs. That is, a GSM $\mathcal{C} = \{\vec{C}_i : i \in \mathrm{I}\}$, where I denotes the set of spatial indices for the RF, can be expressed as : $\mathcal{C} = \mathcal{S} \cdot \mathcal{U} = \{S_i \cdot \vec{U}_i : i \in \mathrm{I}\}$ where S is an RF of positive scalars and U is a Gaussian vector RF with mean zero and covariance Cu. Vector C and U are M dimensional vectors, and we assume that for the RF U, vector U is independent the steerable pyramid of an image as a GSM RF. We partition the subband coefficients into non-overlapping blocks of M coefficients each, and model block i as the vector Ci.
- ○ The purpose of a distortion model i to describe how the statistics of an image are disturbed by a generic distortion operator. The distortion model that we have chosen provides important functionality while being mathematically tractable and computationally simple. It is a signal attenuation and additive noise model in the wavelet domain:
- ○ $$\mathcal{D} = \mathcal{G}\mathcal{C} + \mathcal{V} = \{g_i \vec{C}_i + \vec{V}_i : i \in \mathrm{I}\}$$ where C denotes the RF from a subband in the reference signal, D denotes the RF from the corresponding subband from the test (distorted) signal, G is a deterministic scalar gain field, and V is a stationary additive zero-mean Gaussian noise RF with variance $C_V = \sigma_v^2 \mathbf{I}$. The RF v is white, and is indeoendent of S and U. We constrain the field G to be slowly-varying.
- ○ The Human Visual System (HVS) model that we use is also described in the wavelet domain. Since HVS models are the dual of natural scene statistics (NSS) models many aspects of the HVS are already modeled in the NSS description, such as a scale-spaceorientation channel decomposition, response exponent, and masking effect modeling.
- ○ The purpose of HVS model in the information fidelity setup is to quantify the uncertainty that the HVS adds to the signal that flows through it. We call this lumped HVS distortion visual noise, and model it as a stationary, zero mean, additive white Gaussian noise model in the wavelet domain. Thus, we model the HVS noise in the wavelet domain as stationary RFs img where vector Ni and vecotr N'i are zero-mean uncorrelated multivariate Gaussian with the same dimensionality as vector Ci:

$$\mathcal{E} = \mathcal{C} + \mathcal{N} \quad \text{(reference image)}$$

$$\mathcal{F} = \mathcal{D} + \mathcal{N}' \quad \text{(test image)}$$

- where E and F denote the visual signal at the output of the HVS model from the reference and the test images in one subband respectively, from which the brain extracts cognitive information. The RFs N and N' are assumed to be independent of U,S and V. We model the covariance of N and N' as:

- img where σ_n^2 is an HVS model parameter (variance of the visual noise)

- With the souce, distortion and HVS models as decribed above, the visual information idelity criterion that we propose can be derived. Let

$$\vec{C}^N = (\vec{C}_1, \vec{C}_2, \ldots, \vec{C}_N) \text{ denote } N \text{ elements from } \mathcal{C}. \text{ Let } S^N, \vec{D}^N, \vec{E}^N \text{ and } \vec{F}^N$$

be conrrespondingly defined.

- The mutual information $I(\vec{C}^N; \vec{E}^N)$ quantifies the amount of information that can be extracted from the output of the HVS by the brain when the test image is being viewed. the natural scene model to a specific reference image by treating

$$I(\vec{C}^N; \vec{D}^N | S^N = s^N) \text{ instead of } I(\vec{C}^N; \vec{D}^N),$$ where s^N denotes a realization of

S^N for particular reference image. The realization s N could be thought of as 'model parameters' for the associated reference image. The conditioning on S is intuitively in line with divisive normalization models for the visual neurons [1], and lends the VIF to analytical tractability as well.

- $$I(\vec{C}^N; \vec{E}^N | \vec{S}^N = s^N) \text{ as } I(\vec{C}^N; \vec{E}^N | s^N)$$. With the stated assumptions on C and the distortion model, get:

$$
\begin{aligned}
I(\vec{C}^N; \vec{E}^N | s^N) &= \sum_{j=1}^{N} \sum_{i=1}^{N} I(\vec{C}_i; \vec{E}_j | \vec{C}^{i-1}, \vec{E}^{j-1}, s^N) \\
&= \sum_{i=1}^{N} I(\vec{C}_i; \vec{E}_i | s_i) \\
&= \sum_{i=1}^{N} (h(\vec{C}_i + \vec{N}_i | s_i) - h(\vec{N}_i | s_i)) \\
&= \frac{1}{2} \sum_{i=1}^{N} \log_2 \left(\frac{|s_i^2 \mathbf{C}_U + \sigma_n^2 \mathbf{I}|}{|\sigma_n^2 \mathbf{I}|} \right)
\end{aligned}
$$

- Similarly we can show that for the test image

$$I(\overrightarrow{C}^N; \overrightarrow{F}^N | s^N)$$

$$= \sum_{i=1}^{N} (h(g_i \overrightarrow{C}_i + \overrightarrow{V}_i + \overrightarrow{N}_i | s_i) - h(\overrightarrow{V}_i + \overrightarrow{N}_i | s_i))$$

$$= \frac{1}{2} \sum_{i=1}^{N} \log_2 \left(\frac{|g_i^2 s_i^2 \mathbf{C}_U + (\sigma_v^2 + \sigma_n^2)\mathbf{I}|}{|(\sigma_v^2 + \sigma_n^2)\mathbf{I}|} \right)$$

-
- Since Cu is symmetric, it can be factored as Cu = QAQ^T, where Q is an orthonormal matrix, and A is a diagonal matrix of eigenvalues λk. One can use this matrix factorization to show:

$$I(\overrightarrow{C}^N; \overrightarrow{E}^N | s^N) = \frac{1}{2} \sum_{i=1}^{N} \sum_{k=1}^{M} \log_2 \left(1 + \frac{s_i^2 \lambda_k}{\sigma_n^2} \right)$$

$$I(\overrightarrow{C}^N; \overrightarrow{F}^N | s^N) = \frac{1}{2} \sum_{i=1}^{N} \sum_{k=1}^{M} \log_2 \left(1 + \frac{g_i^2 s_i^2 \lambda_k}{\sigma_v^2 + \sigma_n^2} \right)$$

-
- $I(\overrightarrow{C}^N; \overrightarrow{E}^N | s^N)$ and $I(\overrightarrow{C}^N; \overrightarrow{F}^N | s^N)$ represent the information that could ideally be extracted by the brain from a particular subband in the reference and the test images respectively. We call $I(\overrightarrow{C}^N; \overrightarrow{E}^N | s^N)$ the reference image information.
- We discovered that a simple ratio of the two information measures relates very well with visual quality. It is easy to motivate this choice of relationship between image information and visual quality. When a human observer sees a distorted image, he has an idea of the amount of information that he expects to receive in the image (modeled through the known S field), and it is natural to expect the proportion of the expected information actually received from the distorted image to relate well with visual quality. Also we have only dealt with one subband so far. One could easily incorporate multiple subbands by assuming that each subband is completely independent of others in terms of the RFs as well as the distortion model parameters. Thus, the VIF that we propose in this paper is given by:

$$\text{VIF} = \frac{\sum_{j \in \text{subbands}} I(\overrightarrow{C}^{N,j}; \overrightarrow{F}^{N,j} | s^{N,j})}{\sum_{j \in \text{subbands}} I(\overrightarrow{C}^{N,j}; \overrightarrow{E}^{N,j} | s^{N,j})}$$

- where we sum over the subbands of interest, and vector Cnj represent N elements of the RF Cj that describes the coefficients from subband j, and so on.

17. Spectral Distortion Index (D_lambda)(**code**)

- Spectral Distortion Index can be derived from the difference of inter-band QI values calculated from the fused MS bands, indicated as $\{\hat{G}_l\}_{l=1}^{L}$, and from the low-resolution MS bands, $\{\tilde{G}_l\}_{l=1}^{L}$. The terms $Q(\hat{G}_l, \hat{G}_r)$ and $Q(\tilde{G}_l, \tilde{G}_r)$ can be grouped into two LxL matrices. The two matrices are symmetrical and the values on the main diagonal are all equal to one.
- A spectral distortion index referred to as Dλ, is calculated as:

$$D_\lambda \triangleq \sqrt[p]{\frac{1}{L(L-1)} \sum_{l=1}^{L} \sum_{\substack{r=1 \\ r \neq 1}}^{L} \left| Q(\hat{G}_l, \hat{G}_r) - Q(\tilde{G}_l, \tilde{G}_r) \right|^p}$$

- with p being a
  positive integer exponent chosen to emphasize large spectral differences: for p = 1, all
  differences are equally weighted; as p increases, large components are given more relevance.
  The index (Equation 5) is proportional to the p-norm of the difference matrix, being equal to 0,
  if and only if the two matrices are identical. If negative values of
  $Q(\hat{G}_l, \hat{G}_r)$ and $Q(\tilde{G}_l, \tilde{G}_r)$, originated by anti-correlated bands, are clipped below zero;
  then, Equation 5 is always lower than or equal to one.

18. Spatial Distortion Index (D_S)(**code**)

  - Spatial Distortion Index is calculated as:

$$D_s \triangleq \sqrt[q]{\frac{1}{L} \sum_{l=1}^{L} \left| Q(\hat{G}_l, P) - Q(\tilde{G}_l, \tilde{P}) \right|^q}$$

  - in which P is the PAN image, and P~ a
    spatially degraded version of the PAN image obtained by filtering with a lowpass filter having
    normalized frequency cutoff at the resolution ratio between MS and PAN, followed by
    decimation. Analogously, Ds is proportional to the q-norm of the difference vector, where q
    may be chosen so as to emphasize higher difference values. The index Ds attains its minimum
    (equal to zero), when the two vectors are identical. Analogously to Equation 5, also Equation 6
    is upper bounded by one if clipping below zero of Q values is enabled.

19. Quality with No Reference (QNR)(**code**)

  - Quality with No Reference, which is the product of the one's complements of the spatial and
    spectral distortion indices, each raised to a real-valued exponent that attributes the relevance
    of spectral and spatial distortions to the overall quality.The two exponents jointly determine
    the non-linearity of response in the interval [0,1], same as a gamma correction, to achieve a
    better discrimination of the fusion results compared:

$$\text{QNR} \triangleq (1 - D_\lambda)^\alpha \cdot (1 - D_s)^\beta,$$

  - Thus, the highest value of QNR is one and is obtained when the spectral and spatial distortions
    are both zero. The main advantage of the proposed index is that, in spite of the lack of a
    reference data set, the global quality of a fusion product can be assessed at the full scale of
    PAN.

20. Spectral Angle Mapper (SAM)(**code**)

  - The result of the comparison is reported as the angular difference (radian) between the two
    spectra according to the equation:

$$\cos^{-1}\left(\frac{\vec{t} \cdot \vec{r}}{\|\vec{t}\| \cdot \|\vec{r}\|}\right)$$

which can also be written as,

$$\cos^{-1}\left(\frac{\displaystyle\sum_{i=1}^{nb} t_i \, r_i}{\sqrt{\displaystyle\sum_{i=1}^{nb} t_i^2}\sqrt{\displaystyle\sum_{i=1}^{nb} r_i^2}}\right)$$

- ○
- ○ Here nb is the numbers of bands in the image. Each pair of spectra is treated as a vecto in nb-space, allowing the similarity of the spectra to be determined without regard to their relative brightness values (displayed in brighter values) corresponding to a better match between reference and test spectra.