Self-similarity of images in the wavelet domain in terms of ℓ^2 and Structural Similarity (SSIM)

Devin Glew and Edward R. Vrscay

Department of Applied Mathematics

University of Waterloo, Waterloo, Ontario, Canada N2L 3G1

ervrscay@uwaterloo.ca

"Waterloo Fractal Coding and Analysis Group:"

http://links.uwaterloo.ca

Research supported by Natural Sciences and Engineering Research Council of Canada (NSERC)

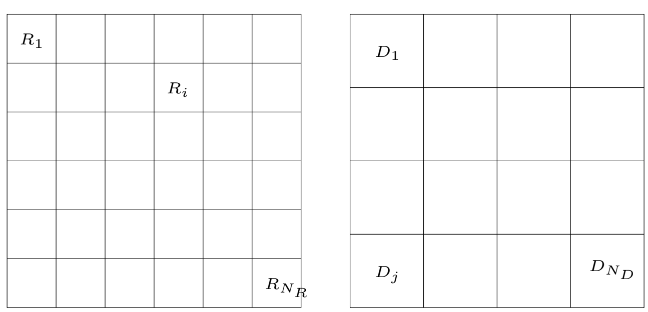
Introduction

Images generally possess a considerable degree of affine self-similarity, that is, their pixel subblocks are well approximated by a number of other subblocks with the help of affine greyscale transformations.

This accounts for the effectiveness of nonlocal image processing schemes such as **nonlocal means denoising** [3] and **fractal image coding** [5].

Model of affine image self-similarity in pixel domain [1]

1. Let $\mathcal{R}$ be a set of nonoverlapping $n \times n$-pixel **range** subblocks R_i, with $\cup_i R_i = X$, support of image.
2. Let $\mathcal{D}$ denote a set of $m \times m$-pixel **domain** subblocks D_i. where $m \geq n$ and $\cup_i D_i = X$.
3. Let $w_{ij} : D_j \to R_i$ denote affine geometric transformation (along with decimation, if necessary).



How well are subimages $u(R_i)$ approximated by subimages $u(D_j)$?

$$u(R_i) \approx \phi_i u(D_j) = \phi_i u(w_{ij}^{-1}(R_i)),$$

where $\phi_i : \mathbf{R} \to \mathbf{R}$ is a **greyscale transformation**.

We consider **affine greyscale transformations**

$$\phi(t) = \alpha t + \beta$$

and examine the distribution of L^2 (RMS) approximation errors Δ_{ij},

$$\Delta_{ij} = \inf_{\alpha,\beta \in \Pi} \| \, u(R_i) - \alpha u(w_{ij}^{-1}(R_i)) - \beta \, \|_2 \quad . \tag{1}$$

$\Pi \subset \mathbf{R}^2$: "feasible parameter space"

Four particular cases of self-similarity:

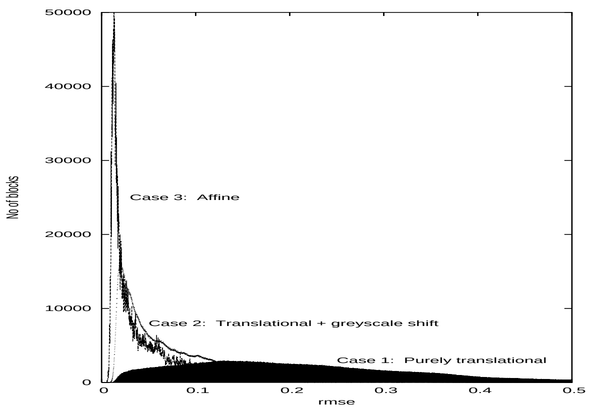
1. **No greyscale transformation:** $u(R_i) \approx u(D_j)$.
2. **Greyscale shift:** $u(R_i) \approx u(D_j) + [\bar{u}(R_i) - \bar{u}(D_j)]$.
3. **Affine, same scale:** $u(R_i) \approx \alpha_i u(D_j) + \beta_i$.
4. **Affine, cross-scale:** $u(R_i) \approx \alpha_i u(w_{ij}^{-1}(R_i)) + \beta_i$.

We expect that

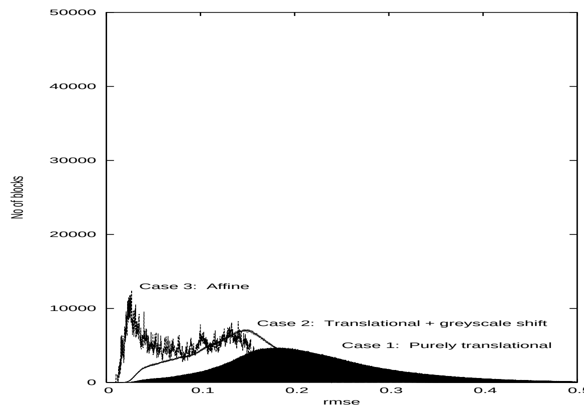
$$\Delta_{ij}^{(Case\ 1)} \geq \Delta_{ij}^{(Case\ 2)} \geq \Delta_{ij}^{(Case\ 3)}$$

since more parameters are involved as we move from Case 1 to Case 3.

Δ_{ij} error distributions for ***Lena*** and ***Mandrill***



(a) Lena



(b) Mandrill

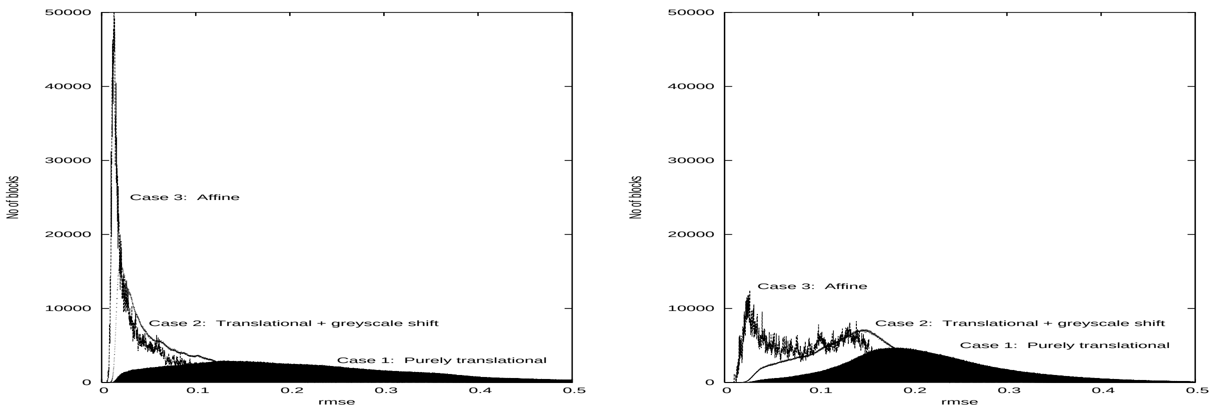
Same-scale (Cases 1,2 and 3) RMS self-similarity error distributions for normalized *Lena* and *Mandrill* images. 8×8-pixel blocks $R_i = D_i$ were used. Case 1 distributions are shaded.

The Δ-error distributions for *Lena* image are more concentrated toward zero error than for *Mandrill* image. From these statistics we may say that the *Lena* image is more self-similar than *Mandrill* image.

The Δ-error distributions for many natural images demonstrate the same characteristics as above. Some will have greater peaking near zero such as the *Lena* image, others will have more diffuse distributions such as the *Mandrill* image.

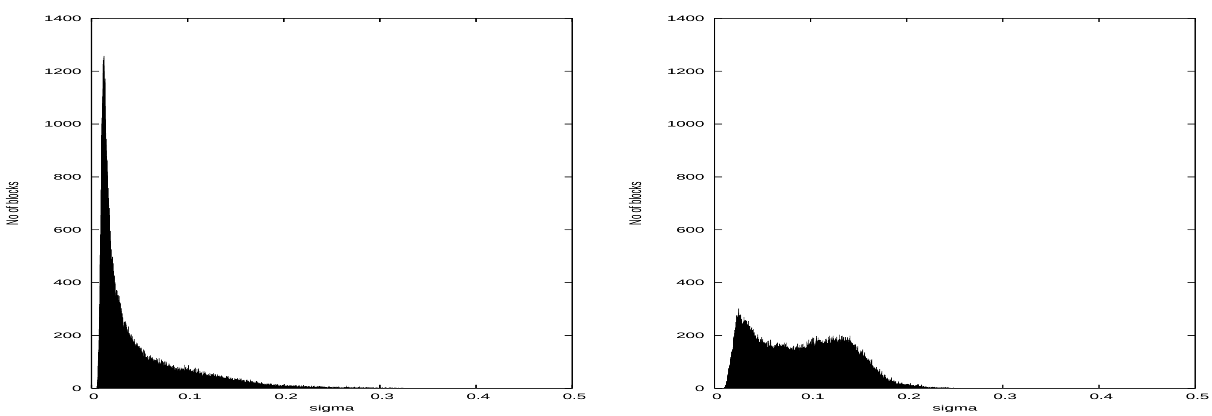
The means and variances of the Δ-error distributions may be used to assign relative degrees of self similarities of images, as done in [1].

The Δ_{ij} error distributions for images, e.g.,



(a) Lena (b) Mandrill

are similar to their $\sigma(u(R_i))$ distributions:



(a) Lena (b) Mandrill

Distributions of $\sigma(u(R_i))$ of 8×8-pixel blocks over the interval $[0, 0.5]$.

$\sigma(R_i)$ is the RMS error of approximation

$$u(R_i) \approx \bar{u}(R_i) \qquad \text{(best } L^2 \text{ fit with a constant).}$$

This corresponds to "clamping" the greyscale parameter, $\alpha = 0$, and optimizing only over β.

Since distributions of α parameters peak at zero, Case 3 distributions are slight perturbations of $\sigma(u(R_i))$ distributions toward zero error.

“Self-similarity” vs. “Approximability”?

The L^2 (RMSE) error in the constant approximation

$$u(R_i) \approx \bar{u}(R_i)$$

is the standard deviation $\sigma(R_i)$. As such, **flatter blocks**, i.e., $\sigma(R_i)$ near zero are **more approximable**.

Images with a higher proportion of flatter blocks will exhibit a greater degree of Δ-error peaking near zero.

In a sense, this may be viewed as **artificial self-similarity**, a consequence of using L^2 distance.

Question: Is there any way to correct for this **unfair advantage** possessed by flat blocks?

Answer: Yes, by examining the **structural similarity** (SSIM) [7, 8] between blocks instead. This was done in [2].

Briefly, the relevant "error" associated with the SSIM between two blocks $\mathbf{x}$ and $\mathbf{y}$ is

$$\sqrt{1 - S(\mathbf{x}, \mathbf{y})} = C \frac{\|\mathbf{x} - \mathbf{y}\|_2}{\sigma_{\mathbf{x}}^2 + \sigma_{\mathbf{y}}^2 + \epsilon_2}$$

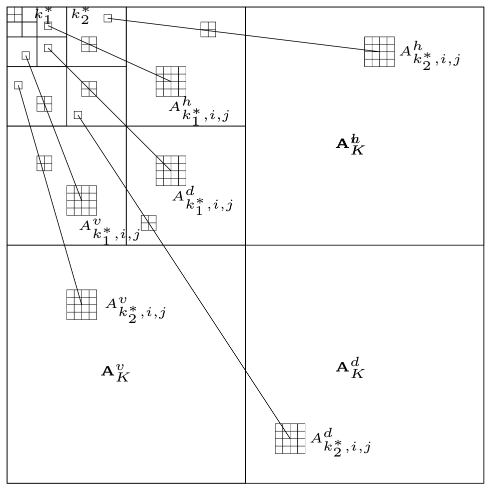
This is an **inverse variance-weighted** L^2 distance between $\mathbf{x} = u(R_i)$ and its approximation $\mathbf{y}$.

The low L^2 approximation errors of flatter blocks are increased by division by small denominators. In this picture, the "renormalized" SSIM-based errors in approximating range blocks R_i by domain blocks D_j are as follows:

The positions of the peaks have been pushed away from zero error.

Image self-similarity in the wavelet domain

Consider standard tensor-product (real) wavelet basis expansions of images. Wavelet coefficients are arranged in following standard matrix form [4],



Let A^λ_{kij}, $\lambda \in \{h, v, d\}$, denote quadtree rooted at coefficient a^λ_{kij}.

Self-similarity in wavelet domain: How well are "range" quadtrees approximated by "domain" quadtrees under affine transformations?

$$A^\lambda_{kij} \approx \alpha A^{\lambda'}_{k',i',j'}, \quad 0 \le k' \le k. \tag{2}$$

Noew that constant coefficient β is omitted in order to preserve ℓ^2 summability of (theoretically) infinite wavelet quadtree.

Three cases are considered:

- **Case 1, No transformation**, i.e., $\alpha = 1$: $A^{\lambda}_{kij} \approx A^{\lambda'}_{k',i',j'}$.
- **Case 3, Affine, same scale**, $k = k'$: $A^{\lambda}_{kij} \approx \alpha A^{\lambda'}_{k,i',j'}$.
- **Case 4, Affine, cross scale**, $k' < k$: $A^{\lambda}_{kij} \approx \alpha A^{\lambda'}_{k,i',j'}$.

Each of these cases may be viewed as a wavelet analogue of the corresponding pixel-based case. Case 4 is the basis of **fractal-wavelet transforms** [6].

Solution for optimal, non-clamped, α-values for approximation in Eq. (2) is simple. Rewriting as

$$\mathbf{a} = \alpha \mathbf{c},$$

optimal α for best ℓ^2 (least-squares) approximation is

$$\alpha = \frac{\langle \mathbf{a}, \mathbf{c} \rangle}{\langle \mathbf{c}, \mathbf{c} \rangle}.$$

Associated squared ℓ^2 approximation error,

$$\Delta^2 = \|\mathbf{a} - \alpha \mathbf{c}\|_2^2 = \langle \mathbf{a}, \mathbf{a} \rangle - \frac{\langle \mathbf{a}, \mathbf{c} \rangle^2}{\langle \mathbf{c}, \mathbf{c} \rangle} = \|\mathbf{a}\|_2^2 \sin^2 \theta,$$

where θ is angle between $\mathbf{a}$ and $\mathbf{c}$.

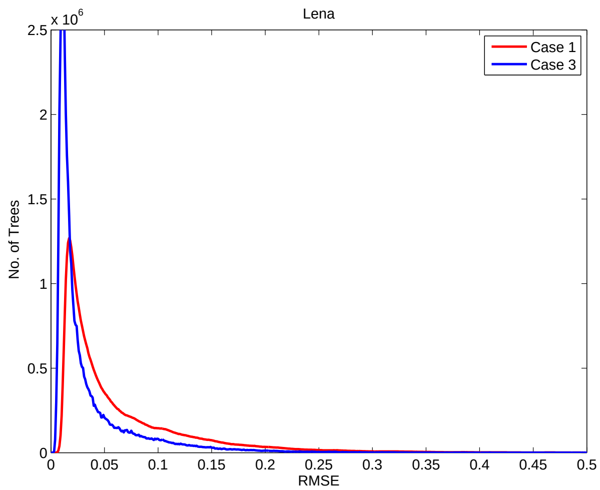
Consequently, the approximation error is bounded as follows,

$$\Delta^{\lambda,\lambda'}_{kij,k'i'j'} = \|A^{\lambda}_{kij} - \alpha A^{\lambda'}_{k'i'j'}\}\|_2 \leq \|A^{\lambda}_{kij}\|_2.$$

Consequence: Wavelet coefficient quadtrees with lower energy are more approximable.

Compare

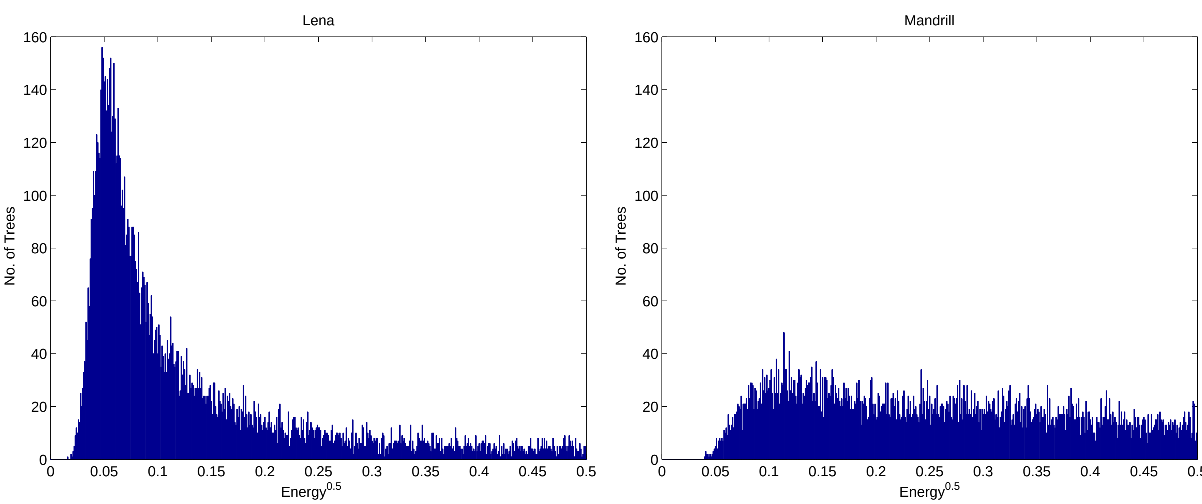
Wavelet Δ-error distributions for *Lena* and *Mandrill*



Case 1 and 3 wavelet-based Δ-error distributions for *Lena* and *Mandrill*.

with

Distributions of norms of wavelet quadtrees



Histogram distributions of norms $\|A^\lambda_{kij}\|_2$ of wavelet quadtrees, $k = 6$, in *Lena* and *Mandrill* images.

SSIM-based investigation of image self-similarity in the wavelet domain

Perhaps the "unfair advantage" of quadtrees with low energy can be removed, as was done in the pixel domain.

It remains to construct a wavelet-based version of the SSIM index. Details are presented in the published paper of the ICIAR 2012 Proceedings. Here we simply state the final results.

Given two wavelet coefficient quadtrees which are represented by the N-vectors $\mathbf{a}$ and $\mathbf{c}$, the SSIM index between them is given by

$$S_W(\mathbf{a},\mathbf{c}) = \frac{2\langle \mathbf{a},\mathbf{c}\rangle + C_2}{\|\mathbf{a}\|_2^2 + \|\mathbf{c}\|_2^2 + C_2},$$

where $C_2 = (M' - 1)\epsilon_2$.

We now seek to find the optimal SSIM-based approximation $\mathbf{a} = \alpha\mathbf{c}$, i.e., maximize

$$S_W(\mathbf{a},\alpha\mathbf{c}).$$

For $C_2 = 0$, the optimal coefficient is

$$\alpha_{SSIM} = \text{sgn}(\langle \mathbf{a},\mathbf{c}\rangle)\frac{\langle \mathbf{a},\mathbf{a}\rangle}{\langle \mathbf{c},\mathbf{c}\rangle}.$$

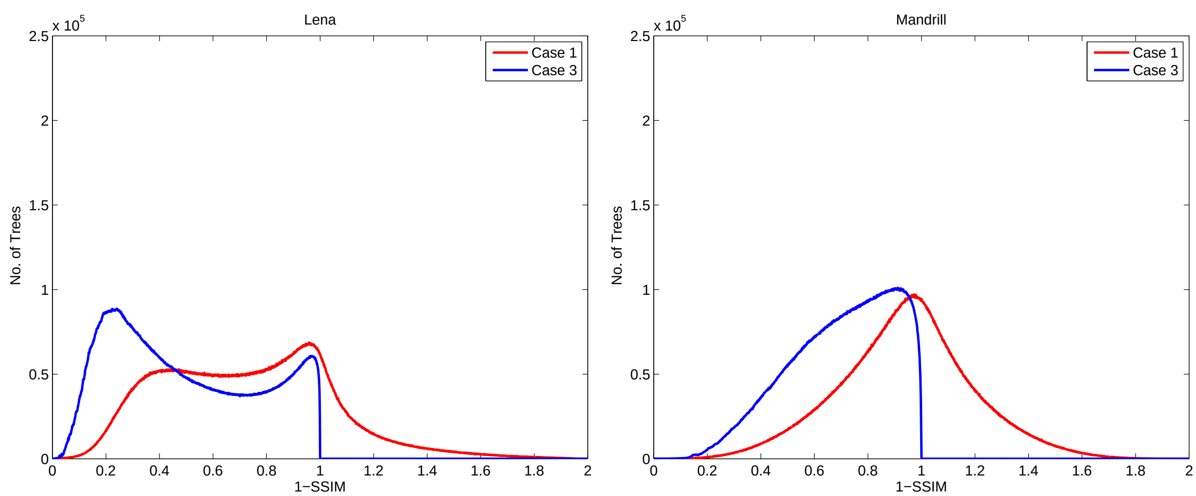
It is more instructive to consider the following quantities,

$$T_W(A^\lambda_{kij}, \alpha_{SSIM} A^\lambda_{k'i'j'}) = 1 - S_W(A^\lambda_{kij}, \alpha_{SSIM} A^\lambda_{k'i'j'}).$$

Since $T(\mathbf{x}, \mathbf{y}) = 0$ implies $\mathbf{x} = \mathbf{y}$, it represents a kind of SSIM-based approximation error. Its range is $[0, 2]$.

$$T_W(\mathbf{a}, \mathbf{c}) = \frac{\|\mathbf{a} - \mathbf{c}\|_2^2}{\|\mathbf{a}\|_2^2 + \|\mathbf{c}\|_2^2 + C_2}.$$

This shows that $T_W(\mathbf{a}, \mathbf{c})$ is an inverse energy-weighted squared ℓ^2 distance. The "unfair advantage" of quadtrees of low energy is lessened in the same way as in the pixel domain: their ℓ^2 distances are increased by division by small denominators.



Case 1 (red) and Case 3 (blue) distributions of the wavelet-based SSIM error $T_W = 1 - S_W$ *Lena* and *Mandrill* images.

References

[1] S.K. Alexander, E.R. Vrscay and S. Tsurumi, A simple model for the affine self-similarity of images, *Image Analysis and Recognition, ICIAR 2008*, Lecture Notes in Computer Science, **5112**, pp. 192-203. Springer-Verlag, Berlin Heidelberg (2008).

[2] D. Brunet, E.R. Vrscay and Z. Wang, Structural similarity-based affine approximation and self-similarity of images re-examined, in *Image Analysis and Recognition, ICIAR 2011*, Lecture Notes in Computer Science **6754**, pp. 264-275. Springer-Verlag, Berlin Heidelberg (2011).

[3] A. Buades, B. Coll and J.M. Morel, A review of image denoising algorithms, with a new one, Multiscale Modelling and Simulation, **4**, 490-530 (2005). An updated and expanded version of this paper appears in SIAM Review **52**, 113-147 (2010).

[4] I. Daubechies, *Ten Lectures on Wavelets*, SIAM Press, Philadelphia (1992).

[5] N. Lu, *Fractal Imaging*, Academic Press, New York (1997).

[6] E.R. Vrscay, A generalized class of fractal-wavelet transforms for image representation and compression, Can. J. Elec. Comp. Eng. **23**, 69-83 (1998).

[7] Z. Wang, A.C. Bovik, H.R. Sheikh and E.P. Simoncelli, Image quality assessment: From error visibility to structural similarity, IEEE Trans. Image Proc. **13**, 600-612 (2004).

[8] Z. Wang and A.C. Bovik, Mean squared error: Love it or leave it? A new look at signal fidelity measures, IEEE Sig. Proc. Mag., **26**, 98-117 (2009).