Translation Insensitive Image Similarity in Complex Wavelet Domain

Zhou Wang & *Eero P. Simoncelli*
Laboratory for Computational Vision, New York University
zhouwang@ieee.org, eero.simoncelli@nyu.edu

Image Similarity Measures

Goal: evaluate the similarity between two images (or image patches) $\mathbf{x}$ and $\mathbf{y}$

Measure 1: mean squared error (MSE) ⟶ $E(\mathbf{x},\mathbf{y}) = \frac{1}{M}\sum_{i=1}^{M}(x_i - y_i)^2$

Measure 2: structural similarity index (SSIM)
[Wang, Bovik, Sheikh, Simoncelli, *IEEE Trans. Image Proc.*, 13(4), 2004]

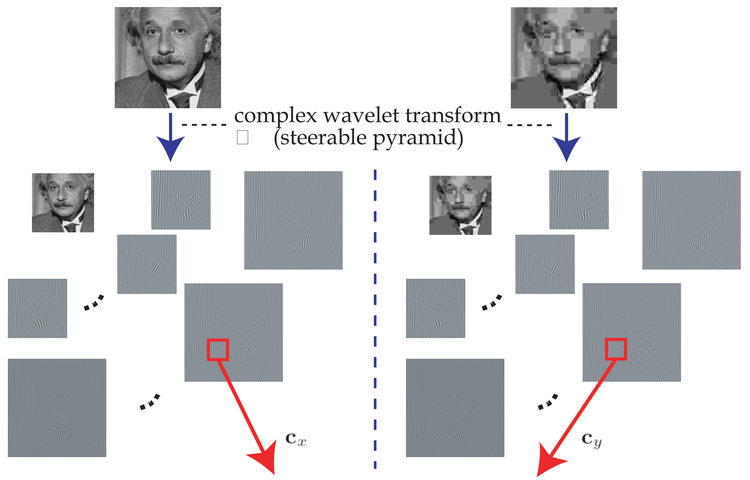
$$S(\mathbf{x},\mathbf{y}) = \frac{(2\,\mu_x\,\mu_y + C_1)\,(2\,\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)\,(\sigma_x^2 + \sigma_y^2 + C_2)}$$

C_1, C_2 : small positive constants $\quad \mu_x = \frac{1}{M}\sum_{i=1}^{M} x_i$

$\sigma_x^2 = \frac{1}{M}\sum_{i=1}^{M}(x_i - \mu_x)^2 \quad \sigma_{xy} = \frac{1}{M}\sum_{i=1}^{M}(x_i - \mu_x)(y_i - \mu_y)$

Drawback: sensitive to spatial translation, scaling and rotation

Complex Wavelet Structural Similarity (CW-SSIM)



$$\tilde{S}(\mathbf{c}_x,\mathbf{c}_y) = \frac{2\,|\sum_{i=1}^{N} c_{x,i}\,c_{y,i}^*| + K}{\sum_{i=1}^{N}|c_{x,i}|^2 + \sum_{i=1}^{N}|c_{y,i}|^2 + K}$$

$$= \frac{2\sum_{i=1}^{N}|c_{x,i}||c_{y,i}| + K}{\sum_{i=1}^{N}|c_{x,i}|^2 + \sum_{i=1}^{N}|c_{y,i}|^2 + K} \cdot \frac{2\,|\sum_{i=1}^{N} c_{x,i}\,c_{y,i}^*| + K}{2\sum_{i=1}^{N}|c_{x,i}\,c_{y,i}^*| + K}$$

magnitude comparison phase comparison

Intuition

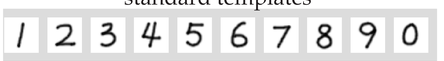
1. Image structural info. is contained in relative phase patterns of wavelet coefs.
2. Constant phase shift of all coefs. does not change image structure

Properties

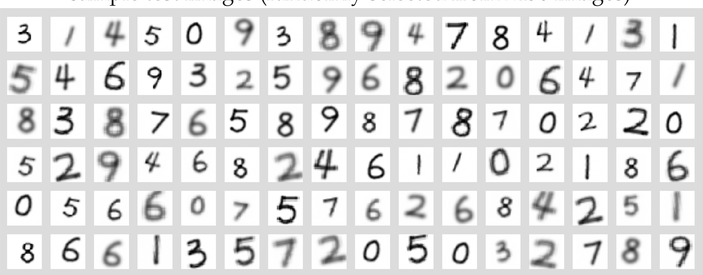
1. Insensitive to luminance and contrast changes
2. Insensitive to spatial translation, scaling and rotation
   (sclaing and rotation can be locally approximated by translation)
3. Sensitive to structural distortion

Image Matching without Precise Registration

standard templates



sample test images (randomly selected from 2430 images)



correct recognition rate (%)

digit	1	2	3	4	5	6	7	8	9	0	all
MSE	84.0	65.4	49.4	63.8	47.7	56.4	68.3	49.8	59.3	51.4	59.6
SSIM	76.1	45.3	47.7	41.6	18.5	42.0	60.9	39.1	51.4	46.5	46.9
CW-SSIM	100	98.4	97.1	100	96.3	97.9	94.2	99.6	100	93.0	97.7

Further Discussions

Computational complexity

1. Efficient: no estimation of registration parameters, no warping, interpolation ...

Connections to computational models of biological vision

1. Involvement of visual channels in pattern recognition [Solomon & Peli '94]
2. Representaion of phase info. in visual cortex [Pollen & Ronner '81]
3. Complex-valued product in visual cortex [Ohzawa *et al.* '90]
4. Sum-of-square computation by complex cells [Adelson & Bergen '85]
5. Divisive normalization by sensory neurons [Schwartz & Simoncelli '01]

Limitations

1. Does not provide correspondance information
2. Works only for small translation, scaling and rotation