Multifocus Image Fusion Using Local Phase Coherence Measurement

Rania Hassen, Zhou Wang, and Magdy Salama

Department of Electrical and Computer Engineering, University of Waterloo,
200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1
`raniahassen@ieee.org, zhouwang@ieee.org, M.Salama@ece.uwaterloo.ca`

Abstract. Image fusion is the task of enhancing the perception of a scene by combining information captured by different imaging sensors. A critical issue in the design of image fusion algorithms is to define activity measures that can evaluate and compare the local information content of multiple images. In doing so, existing methods share a common assumption that high local energy or contrast is a direct indication for local sharpness. In practice, this assumption may not hold, especially when the images are captured using different instrument modalities. Here we propose a complex wavelet transform domain local phase coherence measure to assess local sharpness. A novel image fusion method is then proposed to achieve both maximal contrast and maximal sharpness simultaneously at each spatial location. The proposed method is computationally efficient and robust to noise, which is demonstrated using both synthetic and real images.

Key words: image fusion, local phase coherence, local energy, complex wavelet transform

1   Introduction

Image fusion is the process of combining two or more images to form one image. The main goal is to extract all the perceptually important features from all input images and integrate them to form a fused image in such a way that the new fused image is more informative and is more suitable for human visual perception or computer processing. The fusion of two or more images are often required for images captured using different instrument modalities or different camera settings of the same scene, e.g., at different focus levels. Image fusion has been used as an effective tool for many important applications, which include medical imaging, microscopic imaging, remote sensing, computer vision, and robotics.

The imaging properties of an optical system depend on the acquisition parameters such as focal length or focus and also depend on the distances to the objects imaged. Due to the limited depth-of-focus of optical lenses, it is often not possible to acquire an image that contains all relevant objects in-focus. Therefore, a multifocus fusion process is desirable to create a single image where all objects are in-focus.

Image fusion algorithms may be roughly classified as appearance-based or feature-based, depending on the representation format at which image information is processed. Feature-based algorithms require a feature extraction stage (which is done, for example, by image segmentation) followed by a combination of the feature descriptors. These algorithms are usually less sensitive to noise, but the fusion results highly depend on the accuracy of the feature extraction step [1].

In recent years, multiresolution analysis has become a widely adopted method for transform-domain fusion techniques. The basic idea is to perform multiresolution decomposition on each input image, integrate the decompositions to form a composite representation based on certain fusion rules, and then reconstruct the fused image by performing an inverse multiresolution transform. Various transformation approaches have been used in the literature, which include the Laplacian pyramid [4], the gradient pyramid [5], the ratio-of-low-pass pyramid [6], the discrete wavelet transform (DWT) [7] and the dual-tree complex wavelet transform (DT-CWT) [8, 2, 3]. There are also a variety of fusion rules proposed in the literature, where the most typical ones are:
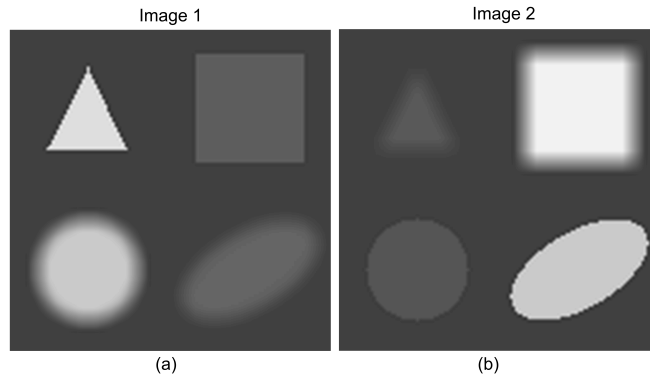
Fig. 1. Synthetic images with different sharpness and contrast in corresponding objects.

- *Choose Maximum (CM) Scheme*: This scheme works under the assumption that at each image location, only one of the source images provides the most useful information. The fused coefficient is obtained by selecting the coefficient with maximal magnitude of the corresponding coefficients in all images.
- *Weighted Average (WA) Scheme* [5]: This scheme uses a normalized correlation between the subband coefficients of the two images over a small local region. The fused coefficient is calculated as a weighted average of the corresponding coefficients.

− *Window-based Verification (WBV) Scheme* [7]: This scheme creates a binary decision map to choose between each pair of coefficients and applies consistency verification using a majority filter.

Since the goal of image fusion is to create a new image that is most informative at each spatial location, it is critical to provide a local measure of information content. In the case that optical out-of-focus is the major source of quality degradations, it is natural to assume that the image region that has higher local energy/contrast and higher sharpness is more active and thus more informative. In existing fusion algorithms, a common implicit assumption is that finding the high energy/contrast regions is equated with finding the high sharpness regions. While this might be true in many application environments, it may not always hold, especially when the images are captured using different instrument modalities. A synthetic example is given in Fig. 1, which demonstrates that local energy/contrast and local sharpness measurements can be two independent events. For example, the top-right square in the first image has lower contrast than that in the second image, but apparently it has higher sharpness. This motivates us to develop a novel image fusion algorithm that uses two different activity measures to assess local energy/contrast and local sharpness, respectively. We can then create a new fused image that is maximal in both local contrast and sharpness at each spatial location. At the core of our approach is a novel complex-wavelet domain image sharpness measure based on the local phase coherence theory [9], which shows that phases of complex wavelet coefficients exhibit a consistent relationship between coefficients at adjacent scales in the *vicinity* of sharp image features, such as edges and lines. This is one step beyond the phase congruency measure [10], which only evaluates cross-scale phase alignment at the *exact* locations of image features.

2   Local Phase Coherence in Complex Wavelet Transform Domain

In order to define the local phase coherence measurement, we will first introduce the concept in 1D which can be generalized to higher dimensions. Given a signal $f(x)$ localized near the position x_0, where $f(x) = f_0(x - x_0)$, the wavelet transform can be written as:

$$F(s,p) = \int_{-\infty}^{\infty} f(x) w_{s,p}^*(x) dx = \left[f(x) * \frac{1}{\sqrt{s}} g\left(\frac{x}{s}\right) e^{j\omega_c x/s} \right]_{x=p}, \qquad (1)$$

where $s \in R^+$ is the scale factor, $p \in R$ is the translation factor and the family of wavelets are derived from the mother wavelet $w_{s,p}(x)$ which is given by:

$$w_{s,p}(x) = \frac{1}{\sqrt{s}} w\left(\frac{x-p}{s}\right) = \frac{1}{\sqrt{s}} g\left(\frac{x-p}{s}\right) e^{j\omega_c (x-p)/s} . \qquad (2)$$

Using the convolution theorem, and the shifting and scaling properties of Fourier transform, Eq. (1) can be written as:

$$
\begin{aligned}
F(s,p) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\sqrt{s}\,G(s\,\omega - \omega_c)e^{j\omega p}d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{\infty} F_0(\omega)\sqrt{s}\,G(s\,\omega - \omega_c)e^{j\omega(p-x_0)}d\omega \\
&= \frac{1}{2\pi\sqrt{s}} \int_{-\infty}^{\infty} F_0\left(\frac{\omega}{s}\right) G(\omega - \omega_c)\,e^{j\omega(p-x_0)/s}d\omega\,,
\end{aligned} \tag{3}
$$

where $F(\omega)$, $F_0(\omega)$ and $G(\omega)$ are the Fourier transforms of $f(x)$, $f_0(x)$ and $g(x)$, respectively. From Eq. (3), we can see that the phase of $F(s,p)$ depends on the nature of $F_0(\omega)$. If $F_0(\omega)$ is scale invariant, meaning that $F_0(\omega/s) = K(s)F_0(\omega)$, where $K(s)$ is a real function of only s, but independent of ω, then from Eq. (3) we obtain:

$$
\begin{aligned}
F(s,p) &= \frac{K(s)}{2\pi\sqrt{s}} \int_{-\infty}^{\infty} F_0(\omega)G(\omega - \omega_c)e^{j\omega(p-x_0)/s}d\omega \\
&= \frac{K(s)}{\sqrt{s}}F(1, x_0 + \frac{p - x_0}{s})\,.
\end{aligned} \tag{4}
$$

Because $K(s)$ and s are real values, the phase of Eq. (4) can be expressed as:

$$
\Phi(F(s,p)) = \Phi(F(1, x_0 + \frac{p - x_0}{s}))\,. \tag{5}
$$

Eq. (5) indicates that there is a strong phase coherence relationship across scale and space, where equal phase contours in the (s,p) plane form straight lines that converge exactly at the location of the feature x_0, as illustrated in Fig. 2(a). These straight lines are defined by $x_0 + \frac{p-x_0}{s} = A$, where A is a constant. More generally, the phase at any given scale may be computed from the phase at any other scale by simply rescaling the position axis.

The above results can be extended for two-dimensional signals or images [9], where the phases of complex wavelet coefficients exhibit a consistent relationship between coefficients at adjacent scales in the vicinity of sharp image features, such as edges and lines. Therefore the fine-scale coefficients can be well predicted from their coarser-scale coefficients, provided that the local phase satisfies the phase coherence relationship defined in Eq.(5). In the case that the positions of the neighboring complex wavelet coefficients are aligned as in Fig. 2(b), the phase prediction expression from coarser scale coefficients $\{a, b_{11}, b_{12}, b_{21}, b_{22}\}$ to the finest scale coefficients $\{c_{ij}\}$ is given by:

$$
\hat{\Phi}(\{c_{ij}\}) = \Phi\left((a^*)^2 \cdot \begin{bmatrix} b_{11}^3 & b_{11}^2 b_{12} & b_{11}b_{12}^2 & b_{12}^3 \\ b_{11}^2 b_{21} & b_{11}^2 b_{22} & b_{11}b_{12}b_{22} & b_{12}^2 b_{22} \\ b_{11}b_{21}^2 & b_{11}b_{21}b_{22} & b_{11}b_{22}^2 & b_{12}b_{22}^2 \\ b_{21}^3 & b_{21}^2 b_{22} & b_{21}b_{22}^2 & b_{22}^3 \end{bmatrix}\right)\,. \tag{6}
$$
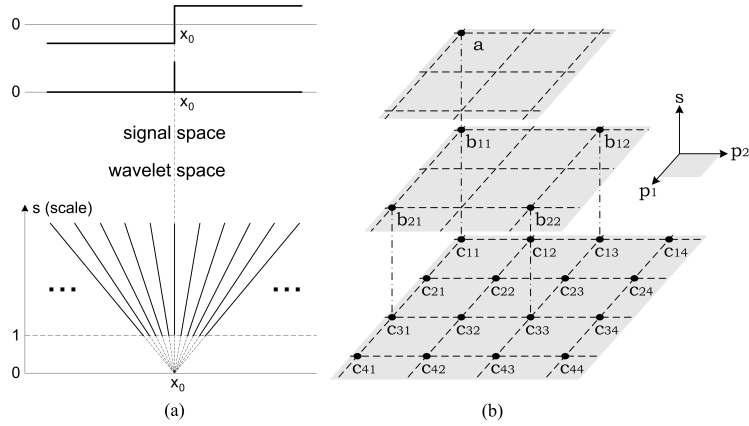
Fig. 2. (a) Local phase coherence of localized sharp features at x_0; (b) 2D sampling grid of wavelet coefficients.

3   Proposed Image Fusion Method

In the proposed fusion scheme, each input image is first decomposed into multi-orientation 3-scale sub-bands using the complex version of the steerable pyramid [11, 12]. The fusion process is carried out in the complex wavelet transform domain and is composed of three modules: *activity measure*, *decision rule*, and *combination rule*. Finally, an inverse transform is applied to the fused sub-band coefficients to create a fused image. In this section, we will focus on the three modules of the fusion process:
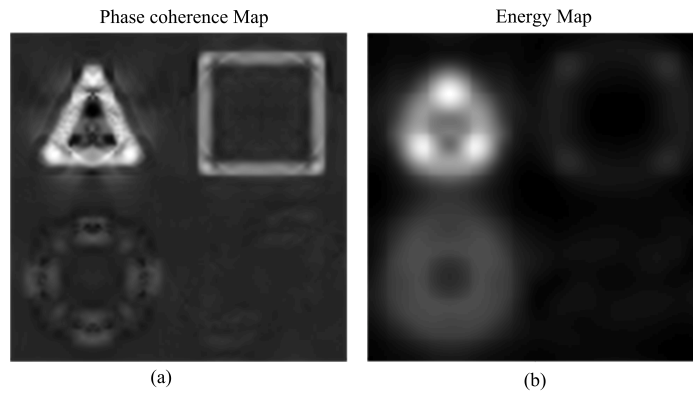
Fig. 3. (a) Local phase coherence map and (b) Local energy map of Fig. 1(a).

Activity measure: The purpose here is to provide an indication of informative content at each spatial location. Different from existing approaches, we compute two distinctive activity measures, one for local sharpness and the other for local energy/contrast. From the last section, we observe that local phase coherence is strong at sharp image features. This motivates us to use it as a measure of sharpness. In particular, we define:

$$P_i = \frac{\sum_o |c_i| \cos\left(\Phi\left(\{c_i\}\right) - \widehat{\Phi}\left(\{c_i\}\right)\right)}{\sum_o |c_i| + K} \,, \tag{7}$$

where K is a positive constant to avoid instability at small energy regions, and its value depends on the image dynamic range and the number of orientations used; $\Phi\left(\{c_i\}\right)$ is the phase of each coefficient in the finest subband; and $\widehat{\Phi}\left(\{c_i\}\right)$ is the predicted phase using Eq. (6). This measure achieves the maximal value when the phase prediction (and thus local phase coherence) is perfect. This is expected to occur in the vicinity of sharp image features. The measure is weighted by the magnitudes of the coefficients over orientations, so that the orientations that contains more energy are given higher weights. It is worth mentioning here that the proposed sharpness measure in Eq. (7) is invariant to image deformation such as translation, rotation and scaling, and is also robust to additive white noise. This is demonstrated in Fig. 4, which shows that regardless of image deformation or degradation, the proposed sharpness measure is able to detect sharp objects and responses weakly to blurred ones. This can be of great benefit if images are not previously aligned. The second activity measure is used to detect high contrast image structure. We use a local energy/contrast measure defined as

$$E_i = W * \sum_o |c_i| \,, \tag{8}$$

where W is a smoothing filter, which is convolved with the sum of the magnitudes of wavelet coefficients (at each spatial location but over all orientations) to provide a smooth local energy map. Fig. 3 shows an example of local phase coherence map and local energy map of Fig. 1(a). It is clear that they emphasize on different types of activities in the image.

Decision and Combination rules: This module is where the actual combination of multi-resolution coefficients is performed. The key idea of our approach is to maintain the phases of the coefficients with maximal local sharpness while boost their magnitudes to achieve the maximal local energy. By doing so, sharp and high contrast features from both images are combined. The decision and combination rules are given by

$$\alpha = arg \max_i P_i \,, \beta = arg \max_i E_i \,, c_{fused} = \frac{E_\beta}{E_\alpha} \cdot c_\alpha \,, \tag{9}$$

where P_i and E_i are the local phase coherence map and the local energy map of the i-th image, respectively. Although we are currently working on fusing two images, the above rule can also be directly applied for fusing three or more images.
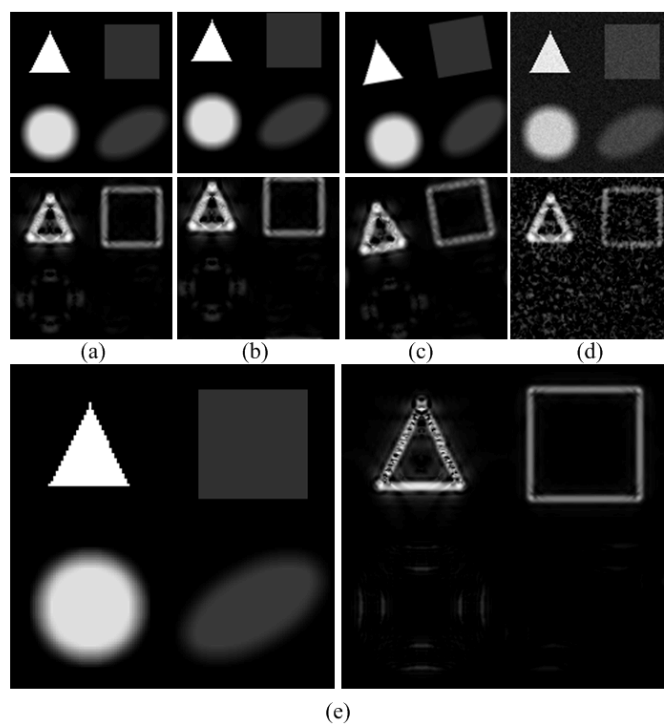
Fig. 4. Test images and their phase coherence maps computed for different image deformations and noise degradation. (a) Original image; (b) translated by (-5,10) pixels; (c) rotated by 10 degree in clockwise direction; (d) additive white noise with $\sigma = 5$; (e) scaled up to twice the size of the original image.
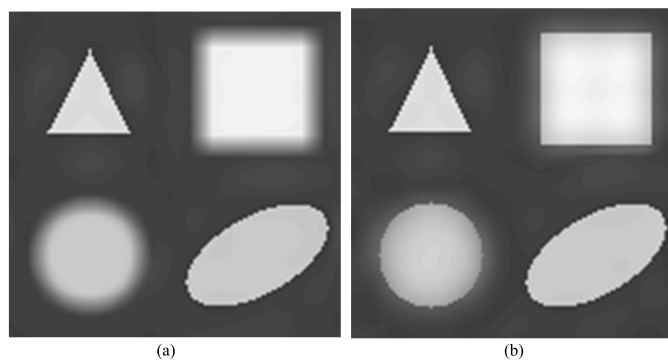
Fig. 5. Fusion results for the images in Fig. 1 using (a) local energy activity measure only and (b) the proposed fusion method.
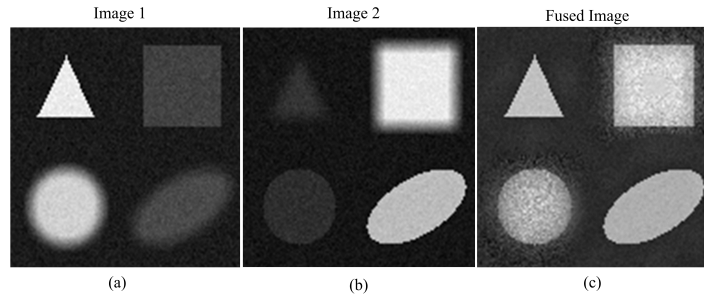
Fig. 6. (a),(b): Synthetic noisy images; (c): Fused image using the proposed method.

4   Experimental Results

In this section, the proposed fusion algorithm is demonstrated with four experiments. In the first experiment, we show the necessity of using separate activity measures for local contrast and local sharpness. Two synthetic test images are shown in Fig. 1, where the same four objects appear in both images at the same spatial locations. The only difference is on their contrast and sharpness. In particular, compared with those in Image 2, the top-left, top-right, bottom-left and bottom-right objects in Image 1 have higher-contrast/higher-sharpness, lower-contrast/higher-sharpness, higher-contrast/lower-sharpness, and lower-contrast/lower-sharpness, respectively. We apply the proposed fusion algorithm to these images and compare it with the result obtained by using the local energy activity measure (as in Eq. (8)) only. The fused images are shown in Fig. 5. It appears that using only the local energy measure, the higher sharpness of the top-right object in Image 1 and bottom-left object in Image 2 cannot be incorporated. By contrast, all useful information has been appropriately fused by the proposed method.

The purpose of the second experiment is to test the robustness of the proposed method in the presence of noise. The same images used in the first experiment are employed, but with added independent white Gaussian noise. The results are shown in Fig. 6. It can be observed that the performance of the proposed algorithm does not change, demonstrating the robustness of the local phase coherence and local energy measures as well as the fusion algorithm.

In the third experiment, we test the proposed fusion algorithm using microscopy images acquired by light microscope with varying focus settings. The result is shown in Fig. 7, where the microscopy images were obtained from the Vision Research lab web site at University of California, Santa Barbra [13]. It can be seen that high contrast and sharp features from both images are appropriately merged in the fused image.

In the last experiment, we test the proposed method using natural images. Fig. 8 demonstrates the results. In Fig. 8(a), two images, each with one object in-focus and another out-of-focus, are fused to create a new image with both objects

Image 1   Image 2   Fused image

(a)   (b)   (c)
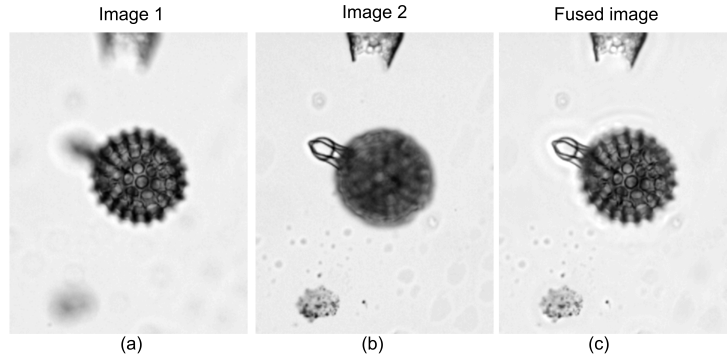
Fig. 7. (a),(b): Two images of radioleria acquired by optical microscope; (c): Fused Image using the proposed algorithm.

in-focus. In Fig. 8(b), by applying the proposed approach to a low contrast/sharp image and a high contrast/blurred image, we obtain a fused image with both high contrast and high sharpness.

5   Conclusions

In this paper, we propose a multifocus image fusion algorithm based on local phase coherence. The major contributions of our work include a novel activity measure of local phase coherence (Eq. (7)) and a new fusion algorithm (Eq. (9)) that combines maximal local energy and maximal local sharpness in the complex wavelet transform domain. Experimental results demonstrate the importance of using separate measures for local energy and local sharpness. The proposed method is robust to noise and computationally efficient, making it suitable for real-world applications.

Acknowledgement

References

1. Dubuisson, M.,Jain , A.K.: Contour extraction of moving objects in complex outdoor scenes. International Journal of Computer Vision, 14, pp. 83-105, (1995).
2. Nikolov, S.G., Hill, P., Bull, D.R., Canagarajah, C.N.: Wavelets for image fusion. In Wavelets in Signal and Image Analysis, A. Petrosian, F. Meyer (Eds.), Kluwer Academic Publishers, The Netherlands, pp. 213-244, (2001).
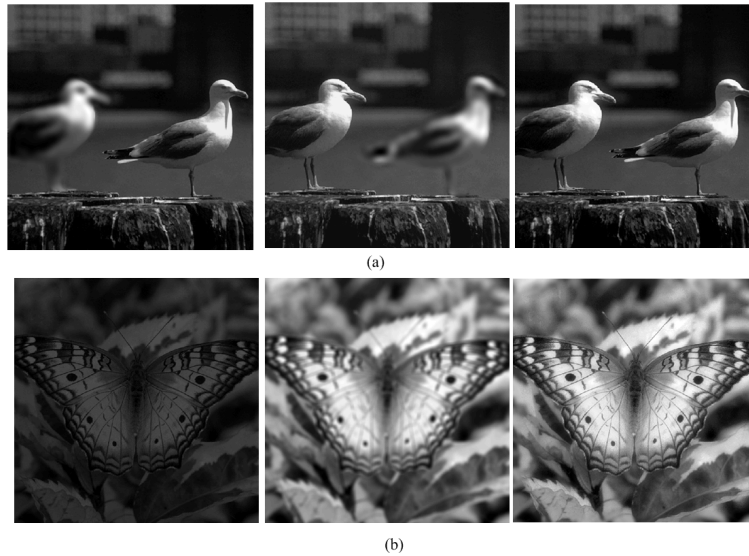
Fig. 8. Proposed fusion algorithm applied to natural images. Left and middle: input image pairs; right: fused images.

3. Lewis, J.J., OCallaghan, R.J., Nikolov, S.G. , Bull, D.R., Canagarajah, C.N.: Region-based image fusion using complex wavelets. In Proceedings of the 7th Inter. Conf. Information Fusion, Stockholm, Sweden, pp. 555–562, (2004).
4. Burt, P.J.: The pyramid as structure for efficient computation. In Multiresolution Image Processing and Analysis, A. Rosenfeld Ed., pp.6–35, Springer-Verlag (1984).
5. Burt, P.J., Lolczynski, R.J.:Enhanced image capture through fusion. in Proceedings of the Fourth Inter. Conf. Computer Vision, Berlin, Germany, pp. 173–182, (1993).
6. Toet, A., van Ruyven, L.J., Valeton, J.M.: Merging thermal and visual images by a contrast pyramid. Opt. Eng. 28(7), pp. 789–792, (1989).
7. Li, H., Manjunath, B.S., Mitra, S.K.: Multisensor image fusion using the wavelet transform. Graphical Models Image Processing, 57(3), pp. 235–245, (1995).
8. Kingsbury N.: Image processing with complex wavelets. In: B. Silverman, J. Vassilicos (Eds.), Wavelets: The Key to Intermittent Information, Oxford University Press, pp. 165-185, (1999).
9. Wang, Z., Simoncelli, E.P.: Local phase coherence and the perception of blur. In Adv. Neural Information Processing Systems NIPS03, MIT Press, pp. 786–792, (2004).
10. Kovesi, P. : Phase congruency: A low-level image invariant. Psych. Research, 64, pp. 136–148, (2000).
11. Simoncelli, E.P., Freeman W.T., Adelson, E.H., Heeger, D.J.: Shiftable Multi-scale Transforms. IEEE Trans Information Theory, 38 (2), pp. 587–607, March (1992).
12. Portilla, J., Simoncelli, E.P.: A Parametric Texture Model based on Joint Statistics of Complex Wavelet Coefficients. Int'l J Computer Vision, 40, pp. 49–71, (2000).
13. Vision Research Lab, University of California at Santa Barbara. http://vision.ece.ucsb.edu/