https://doi.org/10.2352/ISSN.2470-1173.2018.13.IPAS-218

Sharpening Image Details Using Local Phase Congruency Analysis

Andrey Shcherbinin, Konstantin Kolchin, Ivan Glazistov and Mikhail Rychagov, Samsung R&D Institute Russia, Moscow, Russia; Seung-Hoon Han, Samsung Research, Seoul, Korea

Abstract
We suggest a method for sharpening an image or video stream without using convolution, as in unsharp masking, or deconvolution, as in constrained least-squares filtering. Instead, our technique is based on a local analysis of phase congruency and hence focuses on perceptually important details. The image is partitioned into overlapping tiles, and is processed tile by tile. We perform a Fourier transform for each of the tiles, and define congruency for each of the components in such a way that it is large when the component's neighbours are aligned with it, and small otherwise. We then amplify weak components with high phase congruency and reduce strong components with low phase congruency. Following this method, we avoid strengthening the Fourier components corresponding to sharp edges, while amplifying those details that underwent a slight or moderate defocus blur. The tiles are then seamlessly stitched. As a result, the image sharpness is improved wherever perceptually important details are present.

Introduction
Sometimes the region of interest in a high-resolution image does not fit the field of depth when a scene has been shot, due to insufficient focus depth. In addition, defocus in high-resolution images is easily observed when displayed on large, high-resolution screens. At the same time, there is no necessity to remove defocus if the background (e.g. a wall or mountains) has been intentionally placed out of focus by the user.

The goal of our paper is to present a novel method for improving the sharpness of images and video that suffer from defocus blur.

Previous Work
The two most popular approaches to sharpening images are deconvolution [17] and unsharp masking [18]. Deconvolution was introduced to image processing by Enders Robinson in 1950 [19], and has been further developed by many researchers since then [17]. Unsharp masking was invented by German photographers in the 1930s [18]; it was successfully transferred to the digital realm in the 1990s and is widely used in graphical editing software today.

In the 1980s and 1990s, Morrone, Owens and other researchers suggested the local energy model of feature perception [8-16]. This postulates that features are perceived at points in an image where the Fourier components are maximally in phase. Peter Kovesi further developed this model, using the phase 'congruency term', and applied it to detect corners and edges. Our contribution is that we employ phase congruency to enhance perceptually important image details.

Phase congruency characterises coherence among image components in the frequency domain. It has been used for image symmetry, corner, edge and feature detection [1-4], defining sharpness metrics [4-6] and the detection of image splicing [7].

To our knowledge, the approach described below represents the first attempt to use phase congruency for image sharpness improvement and blur reduction.

Method Description
While phase congruency is usually used for edge and corner detection, our technique employs it for sharpness improvement and the removal of defocus blur. Our method consists of the following steps:

1. The RGB colour of each pixel of a video frame is converted to the YCbCr colour space (no conversion is needed in the case of a video stream).
2. We take only the luminance value, *Y*, at each pixel; the resulting Y image is broken up into overlapping tiles.
3. The Fourier transform is applied to the luminance of each tile after windowing using a Gaussian function.
4. The Fourier spectrum of the tile is rebalanced based using phase congruency analysis.
5. The tiles are merged after the inverse Fourier transform; again, Gaussian windowing is used to ensure seamless stitching of the tiles.
6. The new *Y* value is combined with the initial *Cb, Cr* into *Y'CbCr*, and *Y'CbCr* is converted to RGB (no conversion is needed in the case of a video stream).

The method workflow is illustrated in Figures 1 and 2.

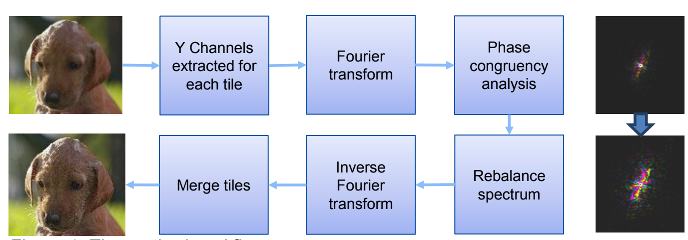


Figure 1. The method workflow

The *Y* component of the video signal is transferred to the chip, where our algorithm is implemented using a form of *System on Chip* (SoC) solution. It is processed there and then admixed with the original *CbCr* components.

The resulting *Y'CbCr* signal is displayed on a LED panel.

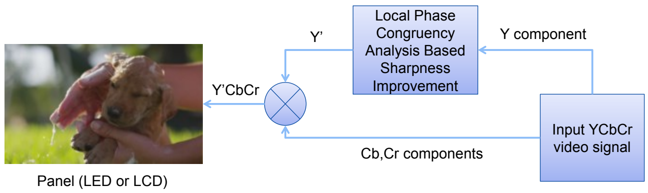


Figure 2. Colour processing.

Suppose we have a set of complex numbers $Y = \{y_i, i = 1..n\}$. We define *phase congruency* with a phase ω and a non-zero constant δ as

$$C_Y(\omega, \delta) = \sum_i \frac{|y_i|}{|y_i| + \delta} \cos\left(\arg(y_i) - \omega\right)$$

We process an image or video frame by applying the discrete Fourier transform that places the zero-frequency component at the spectrum center. For each frequency (l, k), we suggest that the following spectrum area be analyzed (see Figure 3):

$$D(l,k) = \{(i,j): |\varphi_{i,j} - \varphi_{l,k}| \le \Delta_\varphi \wedge r_{i,j} \in [r_{l,k} - \Delta_r, r_{l,k}]\}$$

Here, $(r_{i,j}, \varphi_{i,j})$ are the spherical coordinates of the point (i,j), and Δ_φ, Δ_r are the parameters of the algorithm. Such a choice of the areas to be analyzed is inspired by the following observation: if the image has a pronounced edge at the center, the spectrum acquires high-congruency regions of shapes similar to the one defined above. The areas $D(l,k)$ determine a manifold of the patterns like those shown in Figure 5.

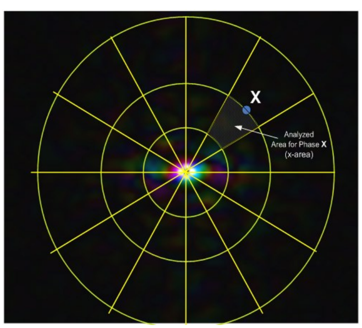
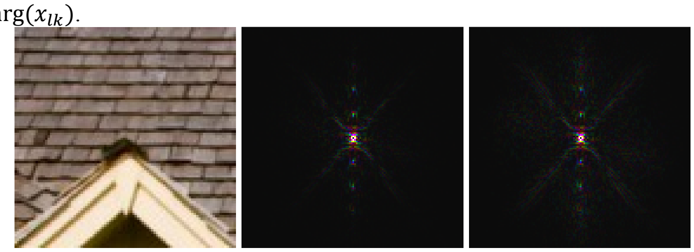


Figure 3. Elements of the phase maps chosen for the averaging of complex magnitudes

According to the local energy model [10, 11], perceptually important features produce the Fourier components that are maximally in phase [8-16]. To sharpen perceptually important image details, we first do the Fourier transform of a Gaussian-windowed tile. We then consider the Fourier components x_{lk} whose frequencies (l,k) have moderate values – that is, they are not too big and not too small. Too small components are of no interest to us because they are responsible for slowly varying image constituents rather than details. Big frequencies are excluded in order to ignore sharp edges and most of noise.

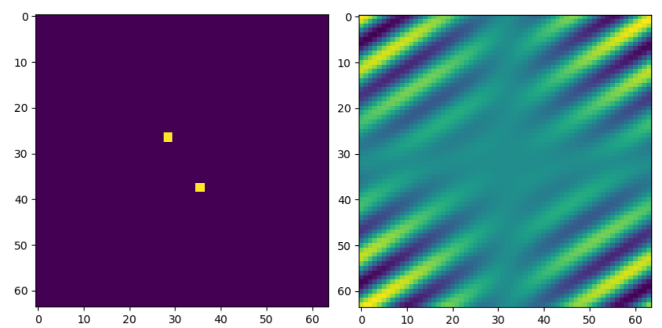
After we selected the Fourier components of interest, we compute phase congruency for them as $C_{D(l,k)}(arg(x_{lk}), \delta)$. Then we amplify weak components with high phase congruency and reduce strong components with low phase congruency. Notice that the parameter δ helps us segregate unreliable spectrum points (l,k). Namely, frequencies where the magnitude is noticeably smaller than δ almost do not contribute to the phase congruency.

We experimented with clusterization of points in $D(l,k)$ based on their phase congruency. We found that it is difficult to define robust criteria for such clusterization within $D(l,k)$. However, we implemented a variant of the algorithm, in which the computation of phase congruency includes only points where the phase $\arg(y_{ij})$ differs by not more than $\pm\pi/2$ from the phase of the point in question $\arg(x_{lk})$.



a. Source Image *b. Initial spectrum* *c. Enhanced spectrum*

Figure 4. Demonstration of typical congruency patterns and spectrum enhancement with our approach. On spectrum images, the color hue encodes the phase, the brightness encodes the magnitude.



a. Spectral area *b. Image pattern corresponding to spectrum area (a.) with congruence*

Figure 5. Demonstration of image pattern (b.) corresponding to spectrum area with phase congruence.

Results

To numerically evaluate our method, we generated in-focus and defocused images from light fields provided by Stanford University (http://lightfield.stanford.edu/lfs.html), using the Java applet available there. Tables 1 and 2 below show how PSNR and MSSIM change after the application of our algorithm. The computation time for images with a resolution of 1536x1280 was less than or approximately equal to 1 s on a PC with Intel Core i7-3770.

Table 1. Test results, PSNR measurements

PSNR, dB			
Test name	Defocused vs. ground truth	Sharpened vs. ground truth	Delta
Treasure Chest	22.283	22.403	0.121
Eucalyptus Flowers	31.256	31.899	0.643
Lego	24.181	24.44	0.26
Average PSNR	25.907	26.248	0.341

Table 2. Test results, MSSIM measurements.

MSSIM			
Test name	Defocused vs. ground truth	Refocused vs. ground truth	Delta
Treasure Chest	0.606	0.66	0.054
Eucalyptus Flowers	0.862	0.889	0.027
Lego	0.716	0.748	0.032
Average MSSIM	0.728	0.766	0.037

Figure 6. An example of the results of sharpness improvement obtained with our method: the upper image is the original frame from a video (courtesy of 4KSamples), and the lower image is the frame obtained by processing using our method

Figures 7 and 8 show a comparison of our method with certain deconvolution techniques [17, 19] and unsharp masking [18]. As can be seen from the above images, our method is the least prone to artefacts compared to other methods, and noticeably improves sharpness (notice the fur on the puppy's nose). The processing time is about 30 s for one UHD frame on a PC with an Intel Core i7-3770, 3.40GHz.

Figure 8. Comparison of our method with deconvolution methods. The upper left image is the original one; the upper right is the result of CLS filtering [17]; the lower left image is the result of the truncated CLS [20]; and the lower right is our result

Conclusion

We present a new method for improving the sharpness of high-resolution images and videos.

The differentiation points of our approach are as follows: (1) it uses image processing based on a model of human visual perception in the frequency domain, employing local phase congruency analysis; (2) it offers uniform multi-scale feature enhancement; (3) it has the capability of highly selective amplification of subtle, faint features, textures and details; (4) it has an artefact-free sharpening/clarifying effect; and (5) it has no over-contrast distortion.

To enhance detail in very dark and very bright areas, our company offers a technology called *Samsung Micro Dimming*. The proposed method improves sharpness in areas that lack sharpness due to defocus. The use of the suggested sharpening technique based on local phase congruency to complement *Micro Dimming* is worthy of interest. Other possible applications include synthesising high-quality still images from video sequences or shot bursts, content enhancement of UHD videos, and semi-automatic adaptive video/image editing.

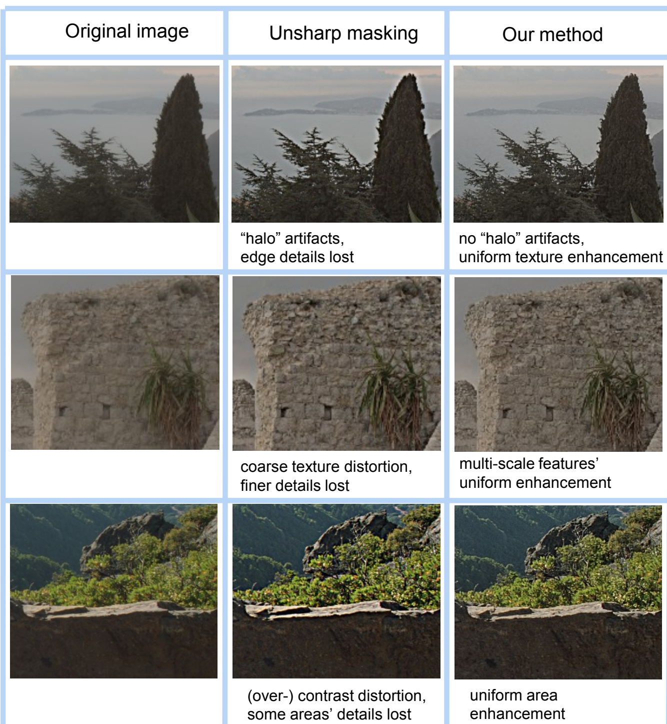


Figure 7. Comparison of our method with unsharp masking

References

[1] P. Kovesi, "Invariant Measures of Image Features from Phase Information," Ph.D. Dissertation, University of Western Australia, 1996.

[2] P. Kovesi, "Phase congruency: A low-level image invariant," *Psychol. Res.*, vol. 64, no. 2, pp. 136–148, Dec. 2000.

[3] P. Kovesi, "Phase Congruency Detects Corners and Edges," *Australian Pattern Recognition Society Conference: Digital Image Computing: Techniques and Applications*, *DICTA 2003*, Sydney. pp 309–318, December 2003.

[4] Z. Wang and E. Simoncelli, "Local phase coherence and the perception of blur," *Advances in Neural Information Processing Systems*, vol. 16, Cambridge, MA, USA: MIT Press, pp. 1–8, May 2004.

[5] R. Hassen, Z. Wang, and M. Salama, "No-reference image sharpness assessment based on local phase coherence measurement," *Proc. IEEE Int. Conf. Acoust. Speech Signal Process,* pp. 2434–2437, Mar. 2010.

[6] R. Hassen, Z. Wang, and M. Salama, "Image sharpness assessment based on local phase coherence," *IEEE Trans. Image Process*, vol. 22, no. 7, pp. 2798–2810, 2013.

[7] W. Chen, Y. Q. Shi, and W. Su, "Image splicing detection using 2-D phase congruency and statistical moments of characteristic function," *Proceedings of SPIE - The International Society for Optical Engineering*, vol. 6505, p. 26, February 2007.

[8] M. C. Morrone and D. C. Burr, "Feature detection in human vision: A phase-dependent energy model," *Proc. R. Soc. Lond. B*, vol. 235, pp. 221–245, 1988.

[9] M. C. Morrone, A. Navangione, and D. Burr, "An adaptive approach to scale selection for line and edge detection," *Pattern Recognition Letters*, vol. 16, pp. 667–677, 1995.

[10] M. C. Morrone and R. A. Owens, "Feature detection from local energy," *Pattern Recognition Letters*, vol. 6, pp. 303–313, 1987.

[11] M. C. Morrone, J. R. Ross, D. C. Burr, and R. A. Owens, "Mach bands are phase dependent," *Nature*, vol. 324, no. 6094, pp. 250–253, November 1986.

[12] R. A. Owens, "Feature-free images," *Pattern Recognition Letters*, vol. 15, pp. 35–44, 1994.

[13] R. A. Owens, S. Venkatesh, and J. Ross, "Edge detection is a projection," *Pattern Recognition Letters*, vol. 9, pp. 223–244, 1989.

[14] B. Robbins and R. Owens, "2-D feature detection via local energy," *Image and Vision Computing*, vol. 15, no. 5, pp. 353–368, May 1997.

[15] S. Venkatesh and R. Owens, "On the classification of image features," *Pattern Recognition Letters*, vol. 11, pp. 339–349, 1990.

[16] S. Venkatesh and R. A. Owens, "An energy feature detection scheme," In *International Conference on Image Processing*, pp. 553–557, Singapore, 1989.

[17] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*, 3rd Edition, Pearson, 2008.

[18] https://en.wikipedia.org/wiki/Unsharp_masking

[19] https://en.wikipedia.org/wiki/Deconvolution#Seismology

[20] S. Kim, E. Lee, M. Hayes, and J. Paik, "Multifocusing and depth estimation using color shift model-based computational camera," *IEEE Trans. Image Process*, vol. 21, pp. 4152–4166, 2012.

Author Biography

Andrey Shcherbinin received his MS (1997) in Applied Mathematics from the Moscow Institute of Steel and Alloys (MISiS). He then worked for Orgtekhdiagnostics Ltd. (1998-2004), Vildis Ltd.(2004-2011) and 1C Ltd. (2011-2013). Since 2013, he has been working at Samsung R&D Institute Russia, Moscow. His research interests include machine learning, pattern recognition, computer vision, video enhancement and content restoration/enrichment.

Michael N. Rychagov received his MSc in Acoustical Imaging and PhD from the Moscow State University (MSU) in 1986 and 1989, respectively. In 2000 he received a Dr.Sc. degree (Habilitation) from the same university. In 1991, he became involved in teaching and research at the Moscow Institute of Electronic Technology (MIET) as an associate professor in the Department of Theoretical and Experimental Physics (1998), a professor in the Department of Biomedical Systems (2008), and a professor in the Department of Informatics and SW for Computer Systems (2014). In 2004, he joined the Samsung R&D Institute in Moscow, Russia (SRR). Since then, Dr. Rychagov has been working on imaging algorithms for printing, scanning and copying, TV and display technologies, multimedia and tomographic areas. Currently, he is the Director of the Advanced Mobile Solution Division at SRR. His research interests are image and video signal processing, biomedical visualisation, engineering applications of machine learning and artificial intelligence.

Konstantin Kolchin received his MSc (1987) and PhD degree (1995) in physics from the Moscow Institute of Physics and Technology. From 1996 to

2003 he worked as a senior researcher at the Scientific Computational Centre within the Keldysh Institute of Applied Mathematics and did consignment research and development for Integra, Inc., Japan. From 2003 to 2013 he worked for Digital Media Professionals, Inc., NVIDIA, Align Technology, Inc., and TELECOM SudParis. In 2014, he joined the Samsung R&D Institute in Moscow, Russia (SRR). Konstantin Kolchin's research interests include computer vision, image enhancement, view synthesis, and physically based rendering.

Ivan V. Glazistov received his MS degree in applied mathematics and computer science from Moscow State University (MSU), Russia in 2011. Since 2013 Mr. I. Glazistov has joined Image Processing Group, Samsung RnD Institute Russia where he is engaged in image and video processing projects.

Seung-Hoon Han received his MS and PhD degrees in Electronics from Korea's Advanced Institute of Science and Technology in 1996 and 2004, respectively. He has been working for Samsung Electronics in Korea since 2004. His major research interests are image quality processing and computer vision. Recently he has also researched sensor fusion algorithms leveraging artificial intelligence for smart machines such as autonomous cars and robots.