

PHASE SINGULARITIES FOR IMAGE REPRESENTATION AND MATCHING

Yu Qiao¹, Wei Wang², Nobuaki Minematsu¹, Jianzhuang Liu³, and Xiaou Tang³

¹University of Tokyo, ²Heriot-Watt University, ³Chinese University of Hong Kong
{qiao, mine}@gavo.t.u-tokyo.ac.jp, w.wang@hw.ac.uk, jzliu@ie.cuhk.edu.hk, xitang@microsoft.com

ABSTRACT

Phase features are widely used in image processing and representation due to their stability to deformation and noise [1, 2]. However, phase singularities, where the signals vanish, are generally regarded as harmful and unreliable facts [3]. In this paper, on the contrary, we will show that phase singularities calculated by Laguerre-Gauss filter contain important information of input image and can provide a reliable representation for image matching. We show that the positions of phase singularities are invariant to translation and rotation. Usually, it is possible to recover the input image up to a constant scaling only from the positions of phase singularities. We study phase singularities in scale space, which allows us to determine the “intrinsic scales” of key phase singularities. We introduce three physical measures of the local structures of phase singularities and combine these measures with SIFT descriptor [4] for image matching. We execute experiments on benchmark database [5] to examine the proposed methods. The results indicate that the proposed method can achieve comparable performance with certain well-known methods [4, 5].

Index Terms— Phase Singularity, Image Representation, Scale Space, Image Matching

1. INTRODUCTION

Image representation is a fundamental problem in image processing and computer vision. A good representation should be stable to noise and deformations, while provide rich and distinctive information for image analysis and recognition. The local phase information, estimated by using complex filters, proves to be a robust representation with respect to noise additions and transformations [1, 3]. Phase can smooth shading and brightness variance due to its invariant nature to amplitude [2]. Moreover, phases include rich information of a signal. It was shown by [6] that one can recover the original image to a fairly large extent by using the phase information only (the magnitudes are set as a unity); on the other hand, if the magnitudes are retained and phases are set as zero, the recovered image is completely indiscernible. The phase features have found successes in many image applications, for example, disparity estimation of stereo [1], object matching [2].

Phase singularities (PSs) refer to the zeros in a complex signal. In most previous phase-related researches, phase singularities are generally regarded as awkward and unreliable events [7]. Usual treatments fail at these positions as phases change abruptly and cannot be estimated. In [7, 3], phase singularities are detected and discarded. Different from those works, in this paper, we take phase singularities estimated by using Laguerre-Gauss filter as important tools for image representation and matching. We show that phase singularities are invariant to shift and rotations. Usually phase singularities contain complete information for reconstructing the original

image up to a constant scaling. We study the core structure around phase singularity and introduce three physical measures for them. By using multi-scale space analysis, we generalize PS points into phase singularity (PS) curves. This allows us to select key PS points along PS curves and to determine their characteristic scales. We combine the physical measures with the famous Scalable Invariant Feature Transform (SIFT) descriptors [4] for image matching tasks. The experimental results on image matching task show that our approach can achieve comparable performance with some well known methods [4, 5].

We note that phase singularities have been a subject under intensive studies in the field of physics and optics in the last decades. Phase singularities proved to be an effective tool for describing and analyzing various physical phenomena, such as rainbow, tide, wave, and the optics of crystal [8, 9]. More recently, Wang et. al used phase singularities to measure the displacement of speckle patterns, for which was referred as the Optical Vortex Metrology [10].

2. PHASE SINGULARITY

In this section, we will give a brief introduction on phase and phase singularity (PS), and discuss the core structure of PS.

Mathematically, phase is defined as the argument of a complex signal. Phase singularity refers to a zero point ($0+i0$) in the complex signal. For 2D image $I(x, y)$, we can obtain its complex representation $\hat{I}(x, y)$ by convolving it with a 2D complex filter $T(x, y)$.

$$\begin{aligned}\hat{I}(x, y) &= I(x, y) * T(x, y) \\ &= I(x, y) * \text{Re}\{T(x, y)\} + iI(x, y) * \text{Im}\{T(x, y)\},\end{aligned}\quad (1)$$

where ‘*’ represents convolution. Let $\hat{I}(x, y) = \rho(x, y) \exp(i\theta(x, y))$ represent the polar form, where $\rho(x, y)$ is amplitude and $\theta(x, y)$ is phase.

In this paper, we use the Laguerre-Gauss (LG) filter for T ,

$$LG(x, y, \sigma) = -\frac{x + yi}{\sigma^4} \exp\left(-\frac{(x^2 + y^2)}{\sigma^2}\right),\quad (2)$$

where σ is a parameter of scale. This filter had been used in wang et al.’s work on Optical Vortex Metrology [10]. Although this paper focuses on LG-filter, many results might be generalized to other complex filters, such as, steerable filter, Gabor filter. The discussions on these are omitted in this paper for limited space.

The real and imaginary parts of LG-filter correspond to the partial derivatives of a Gaussian function $G(x, y) = 1/(2\sigma^2) \exp\{-(x^2 + y^2)/(\sigma^2)\}$, which have been widely used for edge detection and scale space analysis [11]. Let $E(x, y) = I(x, y) * G(x, y)$ denote the Gaussian smoothing image. We have

$$E_x(x, y) = I(x, y) * G_x(x, y),\quad (3)$$

$$E_y(x, y) = I(x, y) * G_y(x, y),\quad (4)$$

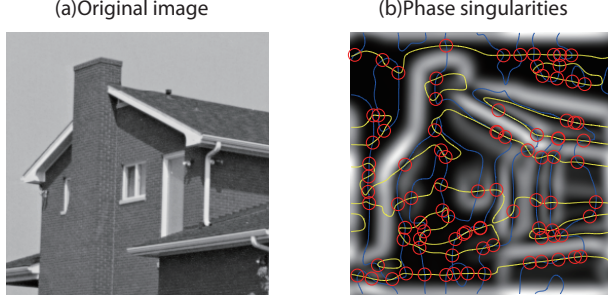


Fig. 1. Example of phase singularities. Red circles denote PS points. Yellow curves and blue curves represent zero-crossing curves of imaginary and real parts of $\hat{I}(x, y)$ respectively.

where subscript x or y denote partial derivatives. Thus,

$$\hat{I}(x, y) = E_x(x, y) + iE_y(x, y). \quad (5)$$

The formulation of Eq. 5 indicates an efficient calculation that the high order derivatives of $\hat{I}(x, y)$ can be obtained by convolving $I(x, y)$ with the high order derivatives of $G(x, y)$, for example,

$$\hat{I}_x(x, y) = I(x, y) * G_{xx}(x, y) + iI(x, y) * G_{xy}(x, y). \quad (6)$$

The Fourier transform of LG filter (Eq. 2) is represented by:

$$\mathcal{L}\mathcal{G}(u, v) = i\pi^2(u + vi) \exp(-\pi^2\sigma^2(u^2 + v^2)). \quad (7)$$

To reduce the computational cost of convolution in Eq. 1, one can calculate the Fourier transform of $I(x, y)$ at first and multiply it with $\mathcal{L}\mathcal{G}(u, v)$. Then the complex image $\hat{I}(x, y)$ can be obtained by applying inverse Fourier transform on the multiplication.

LG filter has several interesting characteristics. 1) LG filter is a band-pass filter. The pass band of the filter is determined by scale parameter σ . This allows us to control the density (number) of phase singularities in $\hat{I}(x, y)$ by changing scale σ . 2) LG filter is DC free, which makes it robust to homogeneous illumination variations.

Phase singularity points (for short, PS points or PSs) are zeros in $\hat{I}(x, y)$. Formally, we use $PS(I, \sigma) = \{(x, y) | \hat{I}(x, y) = 0\}$ to denote the set of PSs. An example of LG-transform and PSs is shown in Figure 1. It is noted that for PS calculation, $I(x, y)$ needs not have to be pixel density, and one can use features too.

It can be shown that the positions of phase singularities are invariant to translation and rotation of input image [12]. Although the invariance does not strictly hold for image zooming, this problem can be largely solved by using scale space analysis as we will show in Section 3.1. The positions of PS points contain nearly complete information of input image. Usually, it is possible to reconstruct the input image $I(x, y)$ up to a constant scaling only from the Positions of PS points. One can refer to [12] for a detailed discussion on the reconstruct method.

2.1. Local structure of phase singularity

We introduce three physical measures to characterize the local structure of a PS point. These measures have their roots in physics researches [8, 10] and are invariant to translation and rotation. They can provide preliminary description of PS points and can be used for rough PS matching. Phase singularities can be divided into two

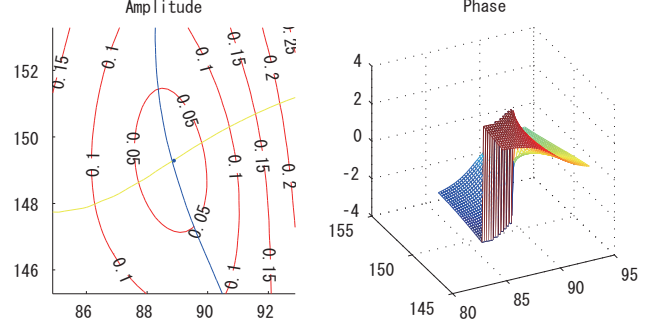


Fig. 2. Local structure of a PS point ((88.9, 149.3) in Fig.1).

classes: *extremes* and *saddles*, which correspond to the extremal and saddle points of $E(x, y)$, respectively.

1) Vorticity of PS p is defined as,

$$\Omega(p) = \nabla\{\text{Re}\{\hat{I}(x, y)\}\} \times \nabla\{\text{Im}\{\hat{I}(x, y)\}\} = \begin{bmatrix} E_{xx} & E_{xy} \\ E_{yx} & E_{yy} \end{bmatrix}. \quad (8)$$

The amplitudes of $\hat{I}(x, y)$ near a PS point exhibit a vortex structure Fig. 2. Ω measures the strength of the vortex [8, 10]. In fact, it equals to the Hessian matrix of Gaussian smoothing image $E(x, y)$.

2) The sign of PS p is defined as the sign of the determinant of vorticity matrix: $s(p) = \text{sgn}(E_{xx}E_{yy} - E_{xy}^2)$. Extreme points have positive signs, while saddle points are negative.

3) Directions of zero-crossing curves θ_1 and θ_2 at a PS point. For real zero-crossing line (yellow line of Fig. 2a)

$$\theta_1 = \arctan\left(\frac{dx}{dy}\right) = -\arctan\left(\frac{E_{xy}}{E_{xx}}\right), \quad (9)$$

and for imaginary zero-crossing curves (blue curve of Fig. 2a)

$$\theta_2 = \arctan\left(\frac{dx}{dy}\right) = -\arctan\left(\frac{E_{yy}}{E_{xy}}\right). \quad (10)$$

θ_1 and θ_2 can be used to evaluate the orientations of phase singularity. Note $\theta_1 - \theta_2$ is invariant to rotation.

3. IMAGE MATCHING USING PHASE SINGULARITIES

As an application of phase singularities, we will show how to use PS for image matching task in this section. Since PS points are stable to certain transformations, here the image matching problem is transformed to find the correspondences between the PS points of two images. Thus the first problem is to select key PS points for matching. We generalize PS points into PS curves in multi-scale space and select key points (scales) along PS-curves, which maximizes the normalized Laplacian function (Eq. 11). This selection also allows us to solve the zooming problem between matching images. For matching selected points, we need to calculate a local descriptor (features) for each point. A quick idea may be to use the three physical measures introduced in Section 2. However, these measures only provide rough descriptions and may lead wrong matchings when images contain complex textures. Therefore, these measures are only used for finding the candidate matching points. We use SIFT descriptor proposed by [4] as rich features of selected points. In the remainder of this section, we sketch our method at first, and then describe experimental results.

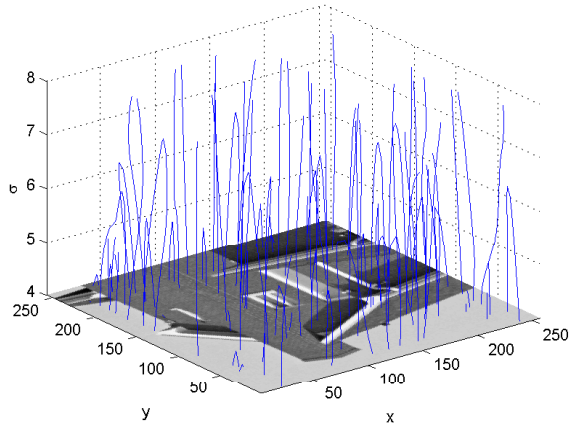


Fig. 3. Example of Phase Singularity curves in scale space.

3.1. Scale-space analysis and key PS points selection

The calculation of phase singularities depends on scale parameter σ of LG-transformation. Every different setting of the scale parameter leads to a different result: new points of phase singularities may appear, and existing ones may move or disappear. The pass-band of the LG filter increases as σ decreases. Thus, usually the larger σ is, the less number of PS points there are. There is a problem of how to select a good scale for the calculation of phase singularities. However, a single scale is seldom sufficient. In fact, the objects in a single image may come from different sizes (corresponding to different scales) and it is difficult to assert that one scale is better than another one. Therefore, other than using a single scale, we consider phase singularities in a multi-scale space coordinated by (x, y, σ) . In this way, PS points will connect into curves, named PS-curves. Some examples are shown in Fig.3.

It can be seen from Fig.3, that PS points of different scales along a PS curve are highly correlated (their x, y coordinates are near). So it is not necessary to use PS points in all the scales for image matching. We need to select interesting scales (points) along PS curves. The intrinsic problem here is to determine the *characteristic scale* of PS points, from which we can estimate the "physical length" of associated image pattern [11, 5] (i.e. in a image, a certain object takes how many pixels). The problem of automatic scale selection had been studied extensively by Lindeberg [11], where his basic idea is that local extremes of normalized derivatives likely correspond to interesting scales. Mikolajczyk [13] compared different normalized derivatives on image matching and showed that normalized Laplacian has the best performance. Following this result, we adopt the normalized Laplacian $NL(x, y, \sigma)$ for scale selection:

$$NL(x, y, \sigma) = \sigma^2 |E_{xx}(x, y, \sigma) + E_{yy}(x, y, \sigma)|, \quad (11)$$

where (x, y, σ) is a point in PS curves.

The main difference between our method and previous scale selection methods [13, 5, 4] is that our selection is constrained to PS curves. Unlike other selection methods, we do not need to compare the normalized Laplacian of every point with its neighbors in scale space for obtaining the extreme points. Moreover, the use of PS curves allows us to achieve sub-pixel precision.

3.2. Local features and image matching

The three physical measures introduced in Section 2.1 only provide rough description. For robust PS points matching, we need rich descriptors of local texture. In this work, we adopt the SIFT descriptor [4]. For each PS-point, we can crop a circled patch of image centered at this point, whose radius is determined by the associated scale. We calculate a SIFT descriptor for this image patch. Details are as follows. At first, the image patch is divided into 4×4 sub regions. Secondly, we calculate the image gradient magnitudes and orientations for every pixel in a sub region. Then, the magnitudes are weighted by a 2D Gaussian function centered at the PS point, and the gradient angles are quantized into eight orientations. Finally, we accumulate the weighted magnitudes for each quantized orientation to obtain a histogram representation in every sub region. The SIFT descriptor for the whole patch has a total dimensionality of $8 \times 4 \times 4 = 128$. SIFT descriptors are highly discriminant in the sense that it can be used to correctly identify two matching points among a large number of candidates [4]. They are robust to illumination changes, noise addition and occlusion. We calculate a SIFT descriptor for each PS in two matching images. Then for each PS, we can find its matching PS in another image by minimizing the difference of SIFT descriptor.

3.3. Experimental results

We evaluate the proposed matching methods on the benchmark database¹ constructed by [5]. The database contains 8 sets of structured and textured images. Each set of image includes a reference image and five compared images under different types of transformations, such as, viewpoint changes, scale changes, illumination changes, blur and JPEG compressions. The homographies (affine transformation matrixes) between the reference image and the compared images in a particular set are available, which allow us to examine the correctness of the matching points found. For each image, we calculate the key PS points, and determine their associate regions according to the characteristic scales.

Following the evaluation method of [5], we examine the repeatability of the key points under various transformations. That is we compare the the key points found in a reference image and a compared image, and count the number of correct correspondence of PS found in the two images. (Overlap error threshold is set as 40% [5].) The results are summarized in Table 1, which shows both the absolute correspondence number and the relative repeatability rate. Due to space limitation, we only list the comparison results with Lowe's key point detection method [4]². It is noted that both SIFT detector and our method are only scale invariant, not full affine invariant as those in [5]. To be fair for the two methods, we calculate PS points in the difference of Gaussian space (DoG) and use the same scale parameters and the refining methods [4]. For all the comparisons, the proposed method can find more correct correspondent points than Lowe's detector. This is an desirable fact which allows us to obtain more matching points (1.5-2 times). The average repeated rates of two methods are generally similar, but ours are sometimes slightly lower than Lowe's. This is partially because we use the key point refining method which has been developed for Lowe's detector. These results are also comparable to several well-known methods, which had been compared in [5]. We make use of the SIFT descriptors to match the key points detected. Some examples of image matching are shown in Fig 4. More results will be available in [12].

¹ Available at <http://www.robots.ox.ac.uk/~vgg/research/affine/index.html>

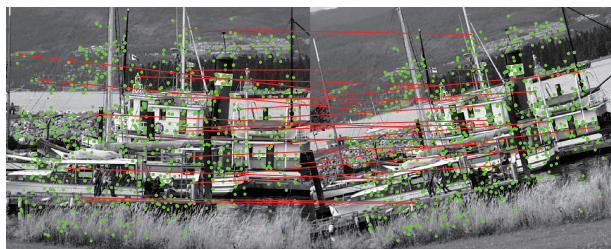
² We use the implementation at <http://vision.ucla.edu/~vedaldi/code/sift/sift.html>.

Table 1. Experimental results on repeatability. The first two and the second row show the types of variation and the names of data sets respectively. The third row indicates the level of changes (please refer to [5] for detailed meaning). Here we only show two levels for each set due to space limitation. CN represents for the correct Correspondence Number, and RR for Repeatability Rate

		Viewpoint				Zoom+Rotation				Blur				JPEG		Illumination	
data set		Graf		Wall		Boat		Bark		Bikes		Trees		UBC		Leuven	
level		20	30	20	40	1.1	1.4	1.2	1.8	3	4	2	3	80	90	3	4
CN	Ours	1100	711	2478	1587	1785	1361	1320	510	1706	1278	1629	1611	2141	1832	2141	1832
	Lowe	693	447	1429	930	940	755	802	331	1050	843	1046	1022	1173	994	1173	994
RR	Ours	74%	53%	70%	55%	66%	68%	70%	70%	60%	58%	55%	55%	75%	64%	73%	71%
	Lowe	76%	54%	72%	56%	68%	71%	69%	74%	62%	53%	57%	55%	75%	64%	73%	70%



(a) Viewpoint graf (672)



(b) Zoom+rotate boat (1023)



(c) Illumination leuven (841)

Fig. 4. Examples of image matching. The numbers in brackets denote the numbers for matching pairs found. Green points represents for all the matching points. For visibility, we only draw 20 matching pairs (red lines) for each example.

4. CONCLUSIONS

This paper shows that phase singularities (PSs) calculated by Laguerre-Gauss filter provide an efficient and effective representation for image matching. PSs are invariant to translation and rotation, and generally PSs contain complete information of an input image that it is possible to reconstruct the image up to a constant scaling only from the positions of PSs. Three physical measures are introduced to characterize the local structure of PS. We also study PSs in scale space and generalize PS points to PS curves. We show how to select key PS point with characteristic scales along PS curves by maximizing normalized Laplacian. As an application, we apply PSs on image

matching task. The experimental results on benchmark database [5] indicate that our method could achieve comparable results with several well-known methods [4, 5].

Finally, it should be noted that it is not our main objective in this paper to develop a high performance image matching method. Our main focus is to explore the usage of phase singularities for image representation. The results can be improved if using more precise methods to locate the extreme PS points and better refining methods, which will be our future work. We are also going to explore more theoretical properties of PS (including PS calculated by complex filters other than LG), and develop other possible applications of PS.

5. REFERENCES

- [1] D. J. Fleet and et. al, "Phase-based disparity measurement," *CVGIU*, vol. 53, no. 2, pp. 198–210, 1991.
- [2] G. Carneiro and A. D. Jepson, "Phase-based local features," in *ECCV*, 2002, pp. 282–296.
- [3] D.J. Fleet and A.D. Jepson, "Stability of phase information," *IEEE Trans. PAMI*, vol. 15, no. 12, pp. 1253–1268, 1993.
- [4] D.G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," *IJCV*, vol. 60, no. 2, pp. 91–110, 2004.
- [5] K. Mikolajczyk and et. al, "A Comparison of Affine Region Detectors," *IJCV*, vol. 65, no. 1, pp. 43–72, 2005.
- [6] AV Oppenheim and JS Lim, "The importance of phase in signals," *Proc. of the IEEE*, vol. 69, no. 5, pp. 529–541, 1981.
- [7] A.D. Jepson and D.J. Fleet, "Phase singularities in scale-space," *IVC*, vol. 9, no. 5, pp. 338–343, 1991.
- [8] M. Berry and M. Dennis, "Phase singularities in isotropic random waves," *Royal Society Proc.: Math., Phy. & Eng. Sci.*, vol. 456, no. 2001, pp. 2059–2079.
- [9] M. Berry, "Making waves in physics. Three wave singularities from the miraculous 1830s," *Nature*, vol. 403, no. 6765, pp. 21, 2000.
- [10] W. Wang and et. al, "Optical vortex metrology based on the core structures of PS in LG transform of a speckle pattern," *Optics Express*, vol. 14, no. 22, pp. 10195–10206, 2006.
- [11] T. Lindeberg, "Feature Detection with Automatic Scale Selection," *IJCV*, vol. 30, no. 2, pp. 79–116, 1998.
- [12] Y. Qiao, "A theory of phase singularity for image representation and its Applications," *Technical Report, The University of Tokyo*, 2008.
- [13] K. Mikolajczyk, *Detection of local features invariant to affine transformations*, Ph.D. thesis, 2002.