Focus on Advances in Shape from Focus

Waqas Ahmad, Muhammad Usman
Department of Computing
Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology
Islamabad, Pakistan
http://www.szabist.edu.pk

Abstract: *Historically, the field of image processing grew from electrical engineering as an extension of the signal processing branch, whereas the discipline of computer science was largely responsible for developments in computer vision. An image processing operation typically define a new image in terms of existing image. On the other hand, computer vision deals with meaningful representation of an environment for decision making. Shape from Focus (SFF) is the combination of above mentioned (both) schemes for controlled environments such as automation and robot path planning in restricted environment. Different techniques have been used in SFF to retrieve spatial information from a sequence of images. Traditional SFF techniques are unable to perform satisfactorily on images which contain high contrast variation between different regions, shadows, defocus points, noise, and oriented edges. Mainly, SFF techniques utilize neighborhood information to deal with noisy gradient, and abrupt depth discontinuities. The neighborhood support and gradient detection complicate focus measure computation for oriented intensity variation. Various techniques have been proposed which range both in spatial and frequency domains to overcome inherent limitations and problems of depth estimation. In this paper, we present a brief overview of various methods along with critical evaluation, and recommendations for future work.*

Keywords: Shape from Focus, Robot Path, Noisy Gradient

Received: 28 July 2014, Revised 3 September 2014, Accepted 10 September 2014

1. Introduction

The field of image processing deals with enhancement and processing of multi-dimensional data. In other words, the images are to be examined and a acted upon by people and / or automated decision systems. Computer imaging can be separated into two primary categories, i.e., 1) computer vision 2) Image processing. These two categories are not totally separate and distinct but consist of some overlapping areas as shown in Figure 1.

$$1/f = 1/u + 1/v \qquad (1)$$

Where f is the focal length of the lens, u is the distance between lens and the image plane and v is the depth of the object. Shape From Focus can be implemented with any combination change in f, u, v. [1] Different approaches are used to measure the focus quality one of them is using window size to measure the focus quality of a pixel. Local window approach assumes that the surfaces of the object in the scene can be approximated by surface patches parallel to the image plane. This approach make the focus quality measurements practical, but this is not always valid because object surfaces can have

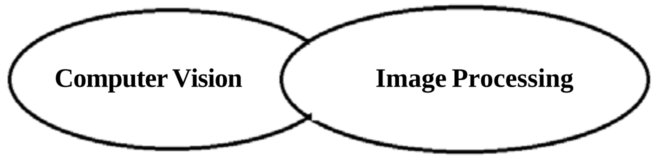

Figure 1. Relationship of Imaging and Vision Fields

complex structure including depth discontinuities. [2,3] Another major problem with traditional Shape From Focus is edge bleeding which is cause by sharp depth discontinuities. This problem cause focus measure ambiguities around the object position. Edge bleeding effects the image region it cannot be easily address by traditional Shape From Focus methods because it is visible in every image in image set. Errors due to edge bleeding can be reduced by increasing focus measure window as the window size support widest blur kernel expected in the setup. Means if the window size increase too much it can also increase the edge bleeding problem.

Depth map is a computation of distance between object point on focus plane and camera lens. Depth map estimation is a critical problem in computer vision with numerous application like in robot guidance, medical imaging, pose estimation, microscopic imaging, range segmentation, collision avoidance, 3D feature extraction and shape reconstruction. Sensors are used to compute the depth by sensing reflect light from the scene object are known as Time-of-flight (TOF) Sensor. Time of flight sensor are expensive and their range limitation are restricted their utilization to specific application. Noise is dependent on light reflected into sensor whereas light reflect from surrounding object complicates depth map computation. Object point appear sharp in image which are present on focus plane blur of image point are appear as they move away from the focus plane if we want a sharp scene with a great depth map than the point on focus plane must have the sharp appearance and rest of the scene is blurred and can be ignored during depth map estimation. Practically captured a sequence of images with different plane to acquire different point that are well focused and rest of the portion of the object is blur with low intensity variation among different image frame. Recently 3D summation in local window based focus measure has been proposed which is computationally expensive a new focus measure operator is proposed to search the frame number for the best focus object points. For proper orientation, steerable filter response with large amplitude to high frequency intensity variation in a image where other techniques are fall down therefore high amplitude is exploited to extract focused image points from different frames to construct a single well focused image. Such an image contain all point of interest on focus plane unlike

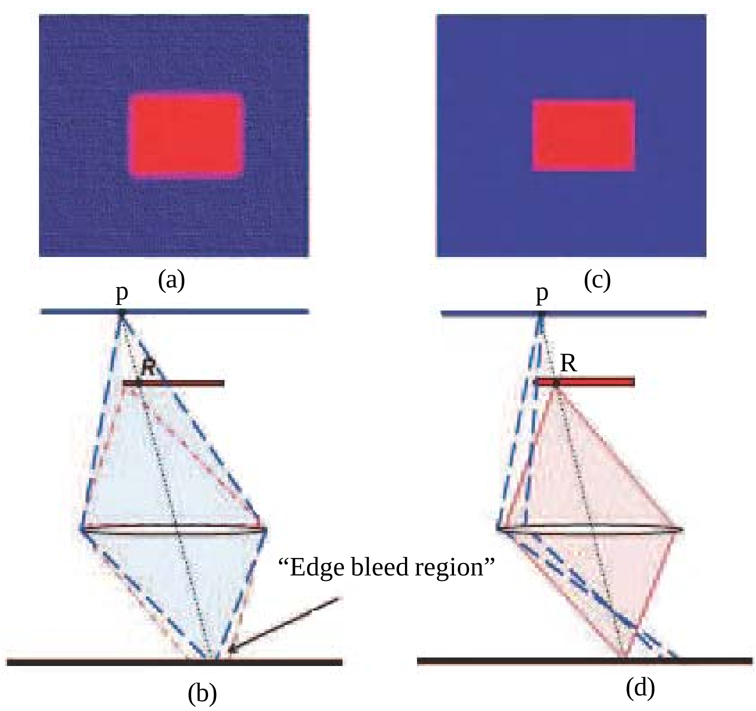

Figure 2. Effect of Window Size on SFF

conventional image scheme. Traditional shape from focus assume convex shaped object for accurate depth map estimation. More over as in shown in below fig the top portion of the object is well focused and the remaining portion is blurred due to its distance from the focus plane. Translation stage is move up and towards the CCD camera until translation stage touch the focus plane such translation is useful in capturing focused portion of an object. Translation of an object and number of images in a sequence depend upon available resources , required accuracy, object structure, illumination variation and reflection.

The objective of shape from focus is to reconstruct 3D shape and determine depth of every point of an object through the camera lens. Each construction yields a single matrix one for each of well-focused image and depth map similar to the size of input frame. For example if a pic consist 1024X786 pixels it mean it consist on 1024 numbers of rows and 786 numbers of columns of matrix and each cell in a matrix denote a pixel a cell of a matrix is combination of row and column. Shape From Focus techniques is further divided into two categories transform and non-transform based focus Transform based Shape from focus include application of domain, discrete cosine, wavelet and curvelet transform to extract focus point from different images to construct the composite image and estimate depth map of a scene. This technique is also useful for depth map estimation and 3D shape recovery in the presence of noise. Non transform based Shape From Focus is subdivided into three main categories 1) gradient 2) statistical 3) approximation based techniques. Subclass od non-transform based Shape From Focus that util..;izes statistical approach consist on focuss measure, mean focus measure, and entropy based fusion.

Moreover the depth can be recovered by binocular or trinocular stereo system, or by taking several images with a monocular camera whose extrinsic or intrinsic settings are changed for every frame. The local focused variation is then used as a depth cue. The algorithm utilize to measure the focus level for every image pixel is usually referred to as a focus measurement (FM) operator. Steerable filter also useful in other application of Shape From Focus like it improve the Shape From Focus using defocus cue, in 3D recovery and identification of 3D object position, real time using shadow to recover the depth discontinues, character recognition focused on off line Handwriting, etc. All these application based on the same principal of matrix and best focused images to obtain the depth map.

This paper consist of five section. Section 2-3 discuss the Literature Review, critical evaluation and theory of steerable filters respectively. Section 4 Proposed Method , followed by concluding remarks in section 5.

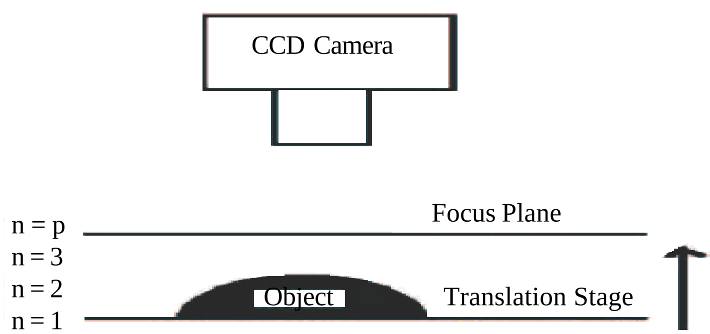

Figure 3. Traditional Image Acquisition Setup for SFF

2. Literature Review and Critical Evaluation

Tarkan Aydin [9] proposes a new focus measure operator for Shape From Focus to recover a dense depth map of a scene. The method can handle depth discontinuities effectively by using adaptively shaped and weighted support windows.This support window determined from the image characteristics of the all-focused image of the scene. Draw back of this approach is the large support window sizes also addresses the edge bleeding problem.

Rashid Minhas [6], [10] used steerable filter to overcome the window size problem. By using the steerable filter we can easily compute discontinuities effectively. For each pixel location 2D neighborhood sum on higher amplitude response, through convolution of oriented steerable filters with image frames, is exploited to locate focused image points for shape reconstruction and depth map computation. Better localization, directional specificity with high amplitude response only at focused points ensures reliable results with efficient processing.

Said Pertuz [8] states that Shape-from-focus (SFF) has widely been studied in computer vision as a passive depth recovery and 3D reconstruction method. One of the main stages in SFF is the computation of the focus level for every pixel of an image by means of a focus measure operator. In this work, a methodology to compare the performance of different focus measure operators for shape-from-focus is presented and applied. The selected operators have been chosen from an extensive review of the state-of-the-art. The performance of the different operators has been assessed through experiments carried out under different conditions, such as image noise level, contrast, saturation and window size. Such performance is discussed in terms of the working principles of the analyzed operators. Depth can be recovered by means of binocular (or trinocular) stereo systems, or by taking several images with a monocular camera whose extrinsic or intrinsic settings are changed for every frame. In the latter approach, shape-from-focus (SFF) has been proposed as a passive method for recovering 3D shapes.

K. S. Pradeep and A. N. Rajagopalan [7], [15] propose an improved SFF method that uses relative defocus blur derived from actual image data to arrive at the final estimates of the structure of the object. A space-variant image restoration scheme is also proposed to obtain a focused image of the 3-D object. The reconstructed 3-D structure as well as the quality of the restored image are superior for the proposed method in comparison to traditional SFF. The image sequence for improving the accuracy of the shape estimates. In the proposed method, for every point on the object, a pair of images is chosen from the sequence with respect to which its image is the sharpest. A spacevariant image restoration method is also proposed for recovering the focused image of the 3-D object.

Joanna Florczak [5] uses a method of 3D shape and depth estimation of an object from a sequence of pictures with changing focus settings. In this paper he propose a novel method of shape recovery, which was originally created for shape and position identification of glass pipette in medical hybrid robot. In the proposed algorithm, Sum of Modified Laplacian is used as a focus operator. Each step of the algorithm is tested in order to select operators with the best results. Reconstruction allows not only to determine the shape but also precisely define the position of the object.

Jean-Marc Hasenfratz [17] covers all current methods for real-time shadow rendering, without venturing into slower, high quality techniques based on ray casting or radiosity. Shadows are useful for a variety of reasons: first, they help understand relative object placement in a 3D scene by providing visual cues. Second, they dramatically improve image realism and allow the creation of complex lighting ambiances. Two main directions appear attractive to render high-quality soft shadows in real time: by programming graphics hardware, and by taking advantage simultaneously of both image-based and object-based techniques. Distributed rendering, using for instance PC clusters, is another promising avenue although little has been achieved so far. Interactive display speeds can be obtained today even on rather complex scenes. Continuing improvements of graphics technology in performance and programmability.

Seong-O Shim, Tae-Sun Choi [4] proposed algorithm tries to find the combination of pixel frames which produce s maxi mum focus measure computed over pixels lying on those frames. To reduce the high computational complexity , a local search method is proposed. After the estimate of the initial dept map solution of an object, the neighbor hood is defined, and an inter mediate image volume is generated from the neighbor hood. The proposed algorithm was compared with the previous SFF methods based on FIS(SFF-FIS) and dynamic programming (SFFDP). Moreover, the proposed algorithm was compared with the recent SFD algorithm based on diffusion (SFD-DFU). Experimental results show that the proposed technique achieves better quality of 3D reconstructed shape and lower computational time compared to SFF-FIS and SFF-DP. With respect to SFD-DFU, the proposed algorithm shows better performance in terms of accuracy with the disadvantage of requiring more image data.

J. Lorenzo et. al [2] work on Local Binary Patterns based focus measures a represented . Local Binary Patterns (LBP) have been introduced in Computer Vision tasks like texture classification or face recognition . In applications where recognition is based on LBP, a computational saving can be achieved with the use of LBP in the focus measures. The behavior of the proposed measures is studied to test if they fulfill the properties o f the focus measures and then a comparison with some well know focus measures is carried out in different scenarios.

Ng Kuang Chern,Nathaniel [16] states that Different autofocusing methods exist for many cameras today. While not ignoring commercially available methods requiring specialized hardware, this paper focuses mainly on pixel based autofocusing algorithms as applied to CCD camera systems. Different measures of image sharpness are compared. For each of these, different algorithms for searching the best lens setting are assessed in terms of performance as well as their applicability to various situations. In addition, several other factors potentially affecting camera focusing are also discussed. Based on the information obtained, this research attempts to formulate a robust autofocusing algorithm.

We present a critical evaluation of various SFF techniques in Fig. 4 with numerous important factors used in reliable computation of depth maps. It is clearly evident that steerable filters based technique [6] offers highest number of important factors in contrast with other techniques used in our analysis. In following section, we shall present steerable filters' technique for better understanding of this technique.

Author name	Efficient window size	Noise	Global optimal solution	Training data required	Directional selectivity	Frequency domain	Time domain
Tarkan Aydin [9]	✓			✓	✓		
Rashid minhas [6], [10]	✓	✓		✓	✓		✓
Said Pertuz[8]		✓		✓	✓		
K.S. Pradeep[7], [15]		✓		✓	✓	✓	
Joanna Florczak[5]		✓		✓	✓		
Jean –Marc Hasenfratz[17]				✓	✓		
Seong O Shim[4]				✓	✓		
J.Loenzo[2]		✓		✓	✓		
Ng Kuang Chern[16]				✓	✓		
Bilal[14]		✓	✓	✓	✓		

Figure 4. Critical Evaluation of Various Techniques on SFF

3. Detailed Review of Selected Method : Steerable Filters with Constant Time Computation[6]

Filter with orientation are used to examine filter response as a function of time and phase in many application. Oriented filters are used in numerous image processing and computer vision application such as image compression, segmentation, edge detection, texture and motion analysis and image enhancement. To find the response of a filter for different orientation is to analyze its output at various angular versions of the filter each successive filter is different from its preceding one with a small finite angle. Steerable filter refer to randomly oriented filter using a linear combination of the basis filters. Once the basis filter responses are known, the response of the filter to an arbitrary steered angle can be determined easily. Steerable filter offer significant advantages in image analysis over ad-hoc methods of combining randomly oriented filter at different orientations since its synthesis is analytic and exact. It is design as quadrature pair filters and are extremely useful in orientation analysis, adaptive filtering, enhancement of oriented structure, and contour detection. Quadrature pairs is used to mark lines and edges with a single response.

The cumulative row sum and integral image can be computed in one pass over the original image. Using the integral image, any rectangular sum can be computed with four array references. The sum of the pixel with in rectangular 4 can be computed using four array references. The value of the integral image at location *A* is the sum of the pixels in rectangular 1. The value at location *B* is the sum of pixel is the rectangular 1 and 2.similarly at location *C* we have sum of the sum of the pixels computed by and the location *D* it is . The sum of pixels within rectangle 4 can be computed as

$$sum = D + A - (B + C) \qquad (2)$$

3.1 Proposed Method for Constant Time Computation [6]

The focus measure computation is a two-step process 1) extraction of gradient information 2) embedding neighborhood information to minimize noise effects. The computational complexity in traditional SFF schemes is proportional to the area of neighborhood support for focus measure. In this paper, a new approach to search the best focused pixel values, also known as focus measure, in constant time is presented. Here two new focus measures based on steerable filters, FMS and 3D-FMS for 2D and 3D neighborhood support, respectively, which are proven to outperform traditional gradient detection schemes that fail to render robustness against noise, oriented intensity variations and low textured images. steerable filters are used for depth map estimation and reconstruction of a well-focused image from a sequence of images acquired at varying focus plane. Steerable filters have been successfully applied to different computer vision and image processing problems ranging from feature detection to texture analysis and image denoising. The underlying principle is to generate rotated versions of a filter from a linear combination of the basis filters; this sets some angular bandlimiting constraint on the class of admissible filters. Reliable extraction of focused image points is heavily dependent upon orthogonality constraint between orientation of a feature and gradient detector. The basis functions for steerable filters are directional derivatives which come in different size and orientations. The ability to interpolate such derivatives offers further freedom to attain translational and rotational invariant detection of image gradients, i.e., focus measure computation. Steerable filters have better orientation selectivity than the classical gradient or Hessian based detectors. For orientation analysis, steerable filters outperform other pyramid based schemes such as separable orthogonal wavelets, pyramid Laplacian, block discrete cosine transform (DCT) and Gabor (octaves). The steering theory presented in the previous section for continuous functions can also be extended to the discrete case through sampling since the order of spatial sampling and steering are interchangeable. The weighted sum of a set of spatially sampled basis functions is equivalent to the spatial sampling of the weighted sum of continuous basis functions [15]. The selected Gaussian function to steer due to its separable property, an important issue for efficient implementation, since a non-separable function f leads to higher computational load. Steerable filters, designed for improved control over phase and orientation, perform better for images with low texture where conventional SFF methods fail to accurately estimate depth map. The implementation of our proposed method, as shown in Figure 5, requires a sequence of images of equal size (i.e. Row Col) acquired at varying focus settings. To construct a focused image, all images belonging to the input sequence are assumed to be registered. SFF based techniques perform well for convex objects in a scene. Predefined sizes of basis filters along with a list of desired orientations ( ) is also input to our algorithm. Initially, Gaussian coefficients for two orientations (00 and 900 degrees) are computed and convolved with every image of an input sequence.After syntheses results in different gradient information obtained at varying orientations.

The steerable filters are synthesized at orientations 450 apart i.e. 00, 450, 900, and 1350. To determine gradient information of an image, our steerability spans over first half of the cartesian coordinates, whereas analogous response prevents the use of additional orientations from remaining quadrants. As per intuitive approach, a higher amplitude response is obtained for edges perpendicular to the orientation of a steerable filter. For focused image reconstruction and accurate depth map calculation, the amplitude responses for each pixel are MAX-pooled which leads to a matrix of size Row Col corresponding to the highest gradient information at one of the desired orientations.

Please refer to Figure 6 that represents the application of steerable filters using a sample frame from a simulated cone sequence at orientations 450 degrees apart. However, the number of desired orientations may vary depending upon available computing resources and structural variations of the probed object. The amplitude response greatly depends upon mutual interaction of oriented edges with the interpolated direction of a steerable filter. Ideally, the convolutions between an image and basis filters generate higher response to horizontal and vertical edges, whereas oriented gradient information (as shown amplitude response for orientation 450, 1350 in Figure 6) is observed using interpolation of earlier convolutions. Since convolution is a computationally intensive operation; the steerability strategy offers great computational savings by replacing convolution with an interpolation procedure for filters with directions other than the basis filters. It should be noted that number of amplitude responses are generated for each pixel of an input image and one that exhibits the highest value is retained for further processing. Such preservation of high amplitude responses is termed as MAX-pool operation. In a similar way, It compute gradient information for all input images.

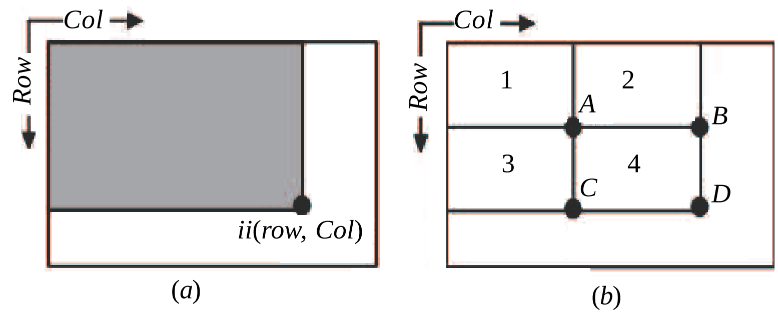

Figure 5. Integral Images for Constant Time Computation of Neighborhood Support [6]

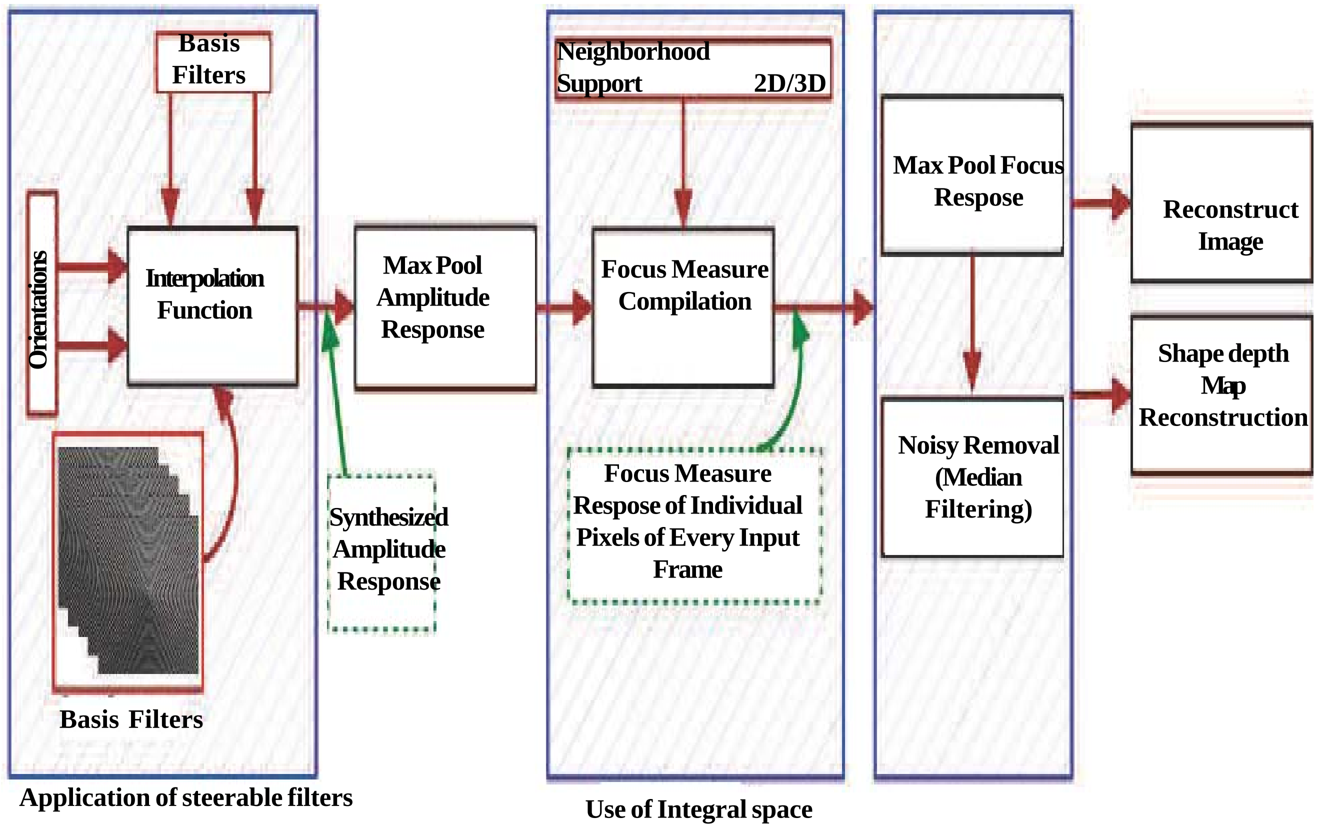

Figure 6: Algorithm Overview of Steerable Filter Based SFF Method [6],[10]

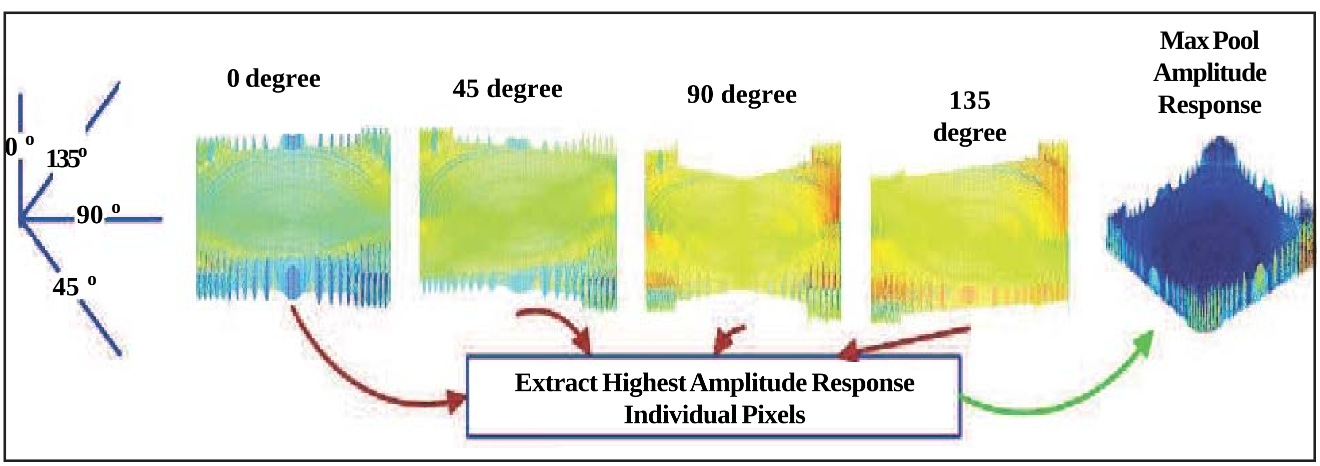

Figure 7. Direction Based Amplitude Response to Compute SFF [6],[10]

4. Future Work and Concluding Remarks

In this review analysis, we presented various techniques on SFF along with limitations and strengths of each technique. As per our analysis, bank of steerable filters offer superior performance and invariance to time for focus measure computation based on neighborhood support. Steerable filters applied at different orientation to the sequence of images with varying texture properties. In quantitative and qualitative analysis steerable filter reliably estimates depth maps and exhibit reliable and robust behavior in the presence of measurement noise, shadows, different illumination conditions.

For future work, it is still an open problem that how can we optimally select window size to include neighborhood information in depth map computation. Such optimal selection of window size can help us to avoid over-smoothing and/or noisy shape computation.

References

[1] Koerich, Alessandro L., Kalva, Pedro R. (2005). Unconstrained handwritten character recognition using metaclasses of characters. *In*: Image Processing, 2005. ICIP 2005. IEEE International Conference on. Vol. 2. IEEE.

[2] Lorenzo, J., et al. (2008). Exploring the use of local binary patterns as focus measure, *In:* Computational Intelligence for Modelling Control & Automation, 2008 International Conference on. IEEE.

[3]Arica, Nafiz, and Fatos T. Yarman-Vural. (2001). An overview of character recognition focused on off-line handwriting." Systems, *Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on* 31 (2) 216-233.

[4] Shim, Seong-O., Choi, Tae-Sun (2010). A novel iterative shape from focus algorithm based on combinatorial optimization, *Pattern Recognition* 43 (10) 3338-3347.

[5] Florczak, Joanna., Petko, Maciej. (2014). Usage of Shape From Focus Method For 3D Shape Recovery And Identification of 3D Object Position, *International Journal of Image Processing* (IJIP) 8 (3) 116.

[6] Minhas, Rashid., Abdul Adeel, Mohammed., Wu, QM Jonathan. (2012). An efficient algorithm for focus measure computation in constant time. *Circuits and Systems for Video Technology, IEEE Transactions on* 22 (1) 152-156.

[7] Pradeep, K. S., Rajagopalan, A. N.(2007). Improving shape from focus using defocus cue. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society* 16 (7) 1920-1925.

[8] Said, Pertuz., Puig, Domenec., Garcia, Miguel Angel. (2013). Analysis of focus measure operators for shape-from-focus. *Pattern Recognition* 46 (5) 1415-1432.

[9] Aydin, Tarkan., Akgul, Yusuf Sinan. (2008). A New Adaptive Focus Measure for Shape From Focus. BMVC.

[10] Minhas, Rashid, et al. (2009). 3D shape from focus and depth map computation using steerable filters. Image Analysis and Recognition. Springer Berlin Heidelberg, 573-583.

[11] Muhammad, Mannan Saeed, et al. (2009). Recovering 3D shape of weak textured surfaces. Computational Science and Its Applications, 2009. ICCSA'09. International Conference on. IEEE.

[12] Asif, Muhammad, Malik, Aamir Saeed., Choi, Tae-Sun. (2005). 3D shape recovery from image defocus using wavelet analysis, Image Processing, 2005. ICIP 2005. IEEE International Conference on. 1. IEEE.

[13] Malik, Aamir Saeed., Choi, Tae-Sun. (2007). Application of passive techniques for three dimensional cameras. *Consumer Electronics, IEEE Transactions on* 53 (2) 258-264.

[14] Ahmad, Muhammad Bilal., Choi, Tae Sun. (2007). Application of three dimensional shape from image focus in LCD/TFT displays manufacturing. *Consumer Electronics, IEEE Transactions on* 53 (1) (2007) 1-4.

[15] Pradeep, K. S., Rajagopalan, A. N. (2007). Improving shape from focus using defocus cue. *IEEE transactions on image processing: a publication of the IEEE Signal Processing Society* 16 (7) 1920-1925.

[16] Chern, Ng Kuang, N., Neow, Poo Aun., Ang, V. M. H. (2001). Practical issues in pixel-based autofocusing for machine vision. Robotics and Automation, 2001. Proceedings 2001 ICRA. IEEE International Conference on. 3. IEEE.

[17] Hasenfratz, JM., et al. (2003). A Survey of Realtime Soft Shadows Algorithms. *Computer Graphics Forum.* 22 (4). Blackwell Publishing, Inc.