International Journal of Engineering Research and Development
***e**-ISSN: 2278-067X, **p**-ISSN: 2278-800X, www.ijerd.com*
Volume 2, Issue 12 (August 2012), PP. 07-12

Analysis and Classification of Feature Extraction Techniques: A Study

Shubha Bhat[1], Vindhya P Malagi[2], Ramesh Babu D R[3], Ramakrishna K A[4]
[1,2,3,4]Computer Science and Engineering Department, Dayananda Sagar College of Engineering, Bangalore, India

Abstract—A detailed study on feature extractors in spatial and transformed domain is carried out in this work. The survey in Spatial domain include most of the traditional detectors until recently the SIFT and its variants. In the transformed domain, the detectors developed using the Fourier transforms to wavelet transforms have been explored. The advantages and the limitations of each one of them is explained along with the results. Depending upon the application in hand together with time complexity and accuracy, an appropriate choice of the suitable detector has to be made.

Keywords—Localization; Keypoints; Scale Invariant Feature Transform; Fourier transform; Dual-Tree Complex Wavelet Transform

I. INTRODUCTION

Features are distinguishable properties or characteristics of an image. Distinct areas of interest such as an edge, corner or a contour can be considered as features in an image. There exist two groups of techniques in the literature for feature extraction. One is based on the localization of features in a sub image and the second global analysis of the entire image. The first approach derives the knowledge about the environment from geometric conditions for example obtained from the odometric measurements of the camera motion or information obtained from a pair of stereo cameras. In this case, the region of interest is a small patch or sub-image within the whole image. Whereas, the Global approach derives its information about the environment based on the information spread out over the entire image. The entire image as a whole is considered the region of interest.

Edge detection is an important task in feature extraction. It is a main tool in several applications for pattern recognition, image segmentation and scene analysis. An edge in an image is a contour across which the brightness of the image changes abruptly. In image processing, an edge is often interpreted as one class of singularities. In a function, singularities can be characterized easily as discontinuities where the gradient approaches infinity. However, image data is discrete, so edges in an image often are defined as the local maxima of the gradient. An edge detector is basically a high-pass filter that can be applied to extract the edge points in an image. This topic has attracted many researchers and many achievements have been made.

In addition to edges, the corners are also considered the best features that can be extracted from an image. Other than edges and corners, blobs are also the best candidates for extracting salient features in an image. Blobs are regions in the image that may contain objects of interest and are either brighter or darker than its surroundings. There are several techniques reported in the literature to detect blobs. Some of the approaches employed to detect blobs are Laplacian of Gaussian (LoG), Difference of Gaussian (DoG), Determinant of Hessian etc which are chosen aptly for the desired application.

In this study, the feature detection methods preferably the edges and the corners as point detectors in the spatial and transformed domains are explored from the vast literature.

The remainder of the paper is as follows: in Section II, a detailed survey in spatial domain is carried out. Section III explains the feature extraction techniques in transformed domain with results in the respective sections. Finally the study concludes with Section IV.

II. FEATURE DETECTION – SPATIAL DOMAIN

The following section explains in detail the traditional methods beginning with the work of canny and harris until recently the revolutionary work of Lowe [17] and its various variants in spatial domain.

A. Traditional feature extraction methods

The earliest work on feature extraction trace back to the year 1979 when Moravec first introduced the term 'interest points' [1]. Later many variations came into existence on the computation of interest points, followed with the pioneering work of Harris and Stephens [2]. The Harris-Laplace and Hessian-Laplace region detectors [3][4] are considered invariant to rotation and scale changes. Some moment-based region detectors [5][6] include Harris-Affine and Hessian-Affine region detectors [7][8]. Others include an edge-based region detector [9], an intensity- based region detector [10], an entropy-based region detector [11] and two independently developed level line-based region detectors called the MSER (Maximally Stable Extremal Region) [12] and LLD (Level Line Descriptor) [13] [14] [15]. These are designed to be invariant to affine transformations. These two methods stem from the Monasse image registration method [16] that uses well contrasted extremal regions to register images. It is reported that MSER is the most efficient one and has better performance than other affine invariant detectors [12]. However, as pointed out in [16], no known detector is actually fully affine invariant. All of them start with initial feature scales and locations selected in a non-affine invariant manner. The difficulty comes from the scale change from an image to another. This change of scale is actually an under-sampling, which means that the images differ by a blur.

It is found that the traditional methods in the spatial domain like canny, sobel, prewitt, roberts etc are simplistic and straight-forward in extracting and matching the image features. However it is observed that either too much of irrelevant information is provided making it slower or some of the useful information (prominent features) is lost. Further study in the spatial domain moves on to the famous state-of-the-art technique called the SIFT (scale-invariant feature transform) developed by David Lowe in 1999.

B. SIFT

Lowe in [17], has addressed the problem of affine invariance for feature extraction and proposed the so called scale-invariant feature transform (SIFT) descriptor, that is invariant to image translations and rotations, to scale changes (blur), and robust to illumination changes. It is also robust to orientation changes of the viewpoint up to 60 degrees. This approach has been named the Scale Invariant Feature Transform (SIFT), as it transforms image data into scale-invariant coordinates relative to local features. This methodology can perform the above mentioned steps either in spatial or in frequency domain. The study in the frequency domain is explained later in section III. The Fig. 1 below shows the features detected by traditional detectors starting from canny to SIFT in the spatial domain for an aerial image.

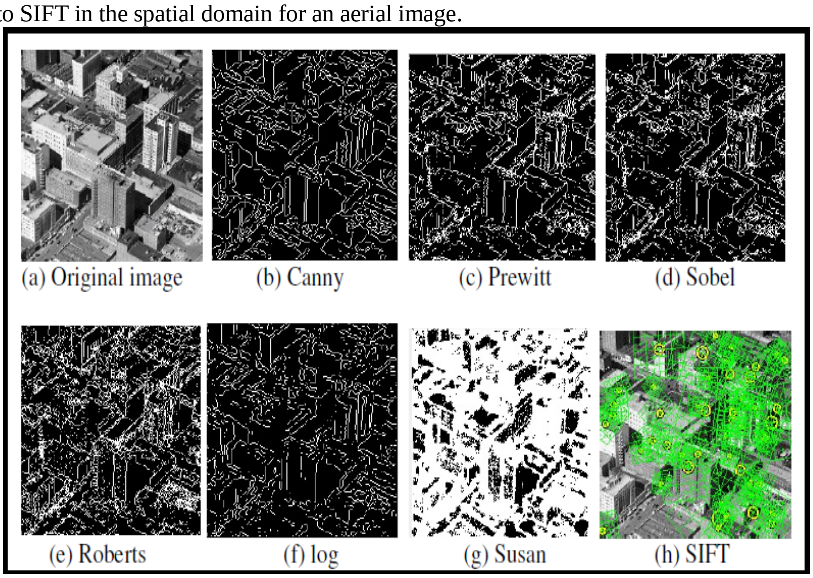

Fig. 1. Features detected using various traditional spatial domain detectors on an aerial image.

The edges and feature points extracted from various detectors shown in the Fig. 1 above reveal canny detector (Fig.1b) still remains the best edge detector. However, it is also seen that SIFT (Fig. 1h) detects local features of interest rather than extracting the entire continuum making it computationally efficient. Based on the scale space theory [18], the SIFT procedure simulates all Gaussian blurs and normalizes local patches around scale covariant image key points that are Laplacian extreme.

A number of SIFT variants and extensions including PCA-SIFT [19] and gradient location-orientation histogram (GLOH) [20] claim to have better robustness and distinctiveness with scaled-down complexity and have been improved with every version with respect to accuracy or time complexity [21] [22]. Several variants of SIFT are explained further.

*1)**PCA-SIFT***

Principal Component Analysis-SIFT [19]: This is an alternate representation for local image descriptors for the SIFT algorithm. Compared to the standard representation, PCA-SIFT is both more distinctive and more compact leading to significant improvements in matching accuracy and speed for both controlled and real-world conditions. Although PCA is ill suited for representing the general class of image patches, it is very well-suited for capturing the variation in the gradient image of a keypoint that has been localized in scale, space and orientation. The work in [19] is extended to the color images. Further exploration in the same is carried out by the authors of PCA-SIFT to other keypoint algorithms.

*2)**ASIFT***

The method proposed, affine-SIFT (ASIFT) [24], simulates all image views obtainable by varying the two camera axis orientation parameters, namely, the latitude and the longitude angles, left over by the SIFT method. Then it covers the other four parameters by using the SIFT method itself.

*3)**A^2SIFT***

With Lowe's implementation as the basis, Auto-Adaptive SIFT [25] improves the performance further. The technique allows extraction of homologous points not only in high geometric distortions but also over bad textured images, where the traditional implementation generally fails. A^2 SIFT can be effectively used in aerial photogrammetric applications.

*4)**SURF***

The Speeded Up Robust Features [23] developed by Bay et al, is a faster implementation compared to the other variants. It is also scale and rotation invariant interest point descriptor and detector. The important speed gain is due to the integration of images, which drastically reduce the number of operations for small box convolutions, independent of the chosen scale. Even without any dedicated optimizations, real time computation has been achieved without any loss in performance.

III. FEATURE DETECTION – TRANSFORMED DOMAIN

Further in the transformed domain, study of feature detection techniques using Fourier transforms and other transforms is carried out.

A. Fourier transforms

In the Fourier domain, the high frequency content are the edges and other significant features in the image. Normally a high pass filter is employed to extract out high frequency content of the signal. The filter allows the high frequency content to pass through while throwing out the low frequency content (less prominent features) of the image. Fourier transforms and its variants, have great ability to capture the frequency content of the image and convert it back to spatial domain using inverse transform very efficiently without losing any information. However, the whole image (global) is spread over the entire frequency axis limiting it from the localization of the image features both in space and frequency simultaneously.

The main drawback of Fourier analysis is that the function is defined from $-\infty$ to ∞ . The effects of each frequency are analyzed as if they were spread over the entire signal. In general, this is not the case. Usually an image is continuously varying in frequency (grey scale information in spatial domain). Fourier analysis done on the image tells us which frequencies exist, but not where they are.

B. Short term Fourier transforms

However, the short time Fourier transform (STFT) is slightly better over Fourier transforms. They often are used when the frequencies of the signal vary greatly with time using different windows but of fixed size. When larger windows are used, lower frequencies can be detected, but their position in time is less certain. With a smaller window, the position can be determined with greater accuracy, but lower frequencies will not be detected. This is the main disadvantage of STFT.

Here, the Wavelets solve this problem. Once applied to a function $f(t)$, it provides a set of functions $W_s f(t)$. Each function describes the strength of a wavelet scaled by factor s at time t. The wavelet extends for only a short period, so its effects are limited to the area immediately surrounding t. The wavelet transform will give information about the strengths of the frequencies of a signal at time t.

Significant contributions were done in the frequency domain [26] by Peter Kovesi, who proposed the concept of Phase Congruency to determine the features in the image. This technique is invariant to illumination and contrast. The image features such as step edges, lines, and Mach bands all give rise to points where the Fourier components of the image are maximally in phase. The use of phase congruency for marking features has found significant advantages over gradient-based methods.

Further, Luca Lucchese in [27] proposed an algorithm in frequency domain that efficiently determines the affine transformations so as to model the relations between pairs of images. This paper presents a new frequency domain technique for estimating affine transformations. It consists of two main steps, one the affine matrix is first estimated and second after compensating for the contribution of the affine matrix, the translation vector is then recovered by means of standard phase correlation. Experimental evidence of the effectiveness of this technique has also been reported and discussed.

C. Discrete Wavelet transforms

In the wavelet domain, Literature shows some interesting work for feature extraction and matching. In [28], the authors have proposed a method to detect edges of the given image using 2-D wavelet transform. This method uses the discrete wavelet transform (DWT) to decompose the image into sub-images, details and an approximation.

Further variants of wavelets and related families show significant advances in area of Feature Detection. In [29], the authors explore the directional extension of multidimensional wavelet transforms, called "contourlets", to perform pattern recognition. The general concept of a directional extension vs. a regular multidimensional wavelet transform is discussed along with the reasoning behind the directional extension. Then, a comparison is done using sample images between the contourlet transform and other edge detection methods for feature detection.

The authors in this paper [30] propose a new technique wherein the feature's (edge points) response is maximum in its neighborhood. The directions of the edges are also estimated from the edge outputs using a line-fitting model. The orientation (rotation angle) of the edges is estimated using angle histograms. The matching of the images is done based on this rotation angle. The authors have proven that translational and rotational changes do not cause much impact, whereas, scaling effect is tolerated up to 10%, beyond which the algorithm restricts itself. The authors claim that the algorithm is faster and more reliable than the conventional methods.

The initial study was conducted to explore all the variants of Wavelets available in the literature. The traditional Discrete Wavelet transform (DWT) was found well suited for image compression and denoising kind of applications due to its multi-scale and multi-resolution characteristics. However, it was seen that it suffered from poor directionality, shift sensitivity and lack of phase information. The other variants of wavelets such as the Wedgelets, Curvelets, Contourlets etc, in spite of their multiscaling features, lacked one or the other characteristics mentioned above. Moreover, what is reported so far in the literature is the usage of these techniques in the context of image compression, texture synthesis etc.

D. Complex Wavelet transforms

The above mentioned limitations of discrete wavelets were overcome by another variant of wavelets called the Dual-tree Complex Wavelet Transforms. The use of complex wavelets in image processing was originally set up in 1995 by J.M. Lina and Gagnon L in the framework of the Daubechies orthogonal filters banks. It was then generalized in 1997 by Prof. Nick Kingsbury of Cambridge University.

The fundamental paper [31] from Prof. Nick Kingsbury and his team on Keypoint detection using dual-tree complex wavelets is been a ground breaking. The paper shows that DTCWT is a well-suited basis for this problem, as it is directionally selective, smoothly shift invariant, optimally decimated at coarse scales and invertible (no loss of information). The authors claim that their scheme is fast because of the decimated nature of the DTCWT and yet provides accurate and robust keypoint localization, together with the use of the "accumulated energy map". Furthermore results show better robustness against rotation compared to the SIFT detector. Hence the choice of DTCWT would be a best option over any of its contemporary methods.

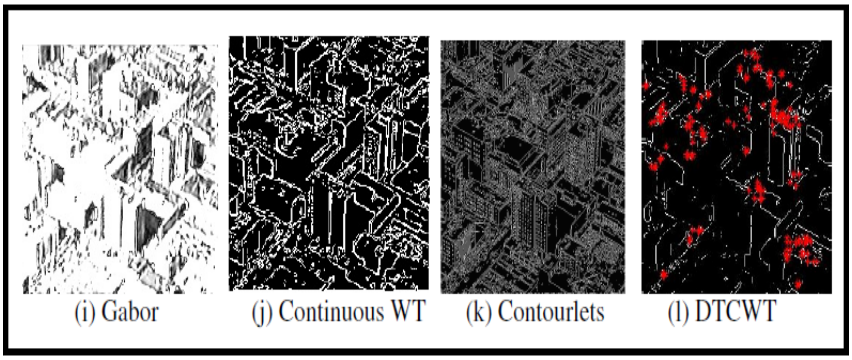

Fig. 2. Features detected using various transformed domain detectors on an aerial image.

Fig. 2 above shows features detected in the transformed domain. The DTCWT features detected here are predominant local interest points wherein their orientation is also taken of. Again, the number of feature keypoints detected are sufficient enough for performing correspondence and registration of images.

The comparison of the different feature extraction techniques explained in the entire paper are summarized in the Table 1 shown below. Here, the column heads T, R, S, L, TC and RE represents respectively the Translation invariance, Rotational invariance, Scale invariance, Localization, time complexity and Reliability as evaluation measures. Thus, one has to suitably choose the best detector as per the requirements of the application in hand.

TABLE 1 : EVALUATION OF EXISTING FEATURE EXTRACTION TECHNIQUES IN SPATIAL AND TRANSFORMED DOMAIN

Legend: ++ Very Good, + Good, - fair, - - poor

Technique	T	R	S	L	TC	RE
Spatial domain techniques (eg, SIFT)	++	+	++	++	-	-
Frequency Domain techniques (eg, Fourier)	+	+	+	- -	+	- -
Wavelet techniques (eg, Discrete Wavelet Transform)	-	-	++	++	-	-
Complex Wavelet techniques (eg, Dual-tree complex Wavelet Transform)	++	++	++	++	-	+

(Comparison based on our survey and experiments)

IV. OBSERVATIONS AND CONCLUSION

The study explores the different spatial and transformed domain approaches to feature extraction. In the literature, it is claimed that SIFT is one of the most robust technique used to detect and match features between images. It is invariant to image translations and rotations, to scale changes, robust to illumination changes and also robust to a certain extent to orientation changes of the viewpoint. Although it a robust method, most of the tasks are computationally intensive and cumbersome. Whereas in this study, the focus is on transformed domain techniques in order to speed up certain functionalities compared to spatial approaches. Further in the study, other frequency and wavelet domain methods to detect features are explored.

The Wavelet approach alternatively leverages the strengths of both spatial and frequency domain processing. The image information can be viewed and processed in both space as well as frequency domains simultaneously. It has an added advantage of representing the image in multiple scales and multiresolutions. Thus, working in the Wavelet domain becomes interesting to explore both spatio-frequency characteristics of the image.

The Complex Wavelets Transforms (CWT) use complex-valued filtering (analytic filter) that decomposes the real/complex signals into real and imaginary parts in transform domain. The real and imaginary coefficients are used to compute amplitude and phase information respectively in addition to the above mentioned characteristics of shift, rotation and scale invariance, just the type of information needed to accurately describe the energy localization of oscillating functions (wavelet basis).

Thus the choice of appropriate detector always should consider the application in hand. Further, appropriate tradeoff between computational time and accuracy in detecting the interest features should also be considered.

ACKNOWLEDGMENT

This work has been funded by ER&IPR Grand-in-Aid Research Project ERIP/ER/0904468/M/01/1181 of DRDO. The authors would like to thank Dr S Sankaran, Director ER&IPR HQ for all his encouragement and support. The authors would specially like to thank Shri P.S Krishanan, Director, Shri V.S Chandrashekar, Group Director and Ms Neeta Trivedi, Head, AIEL of ADE, DRDO for their constant help and support. The authors are thankful to Dr. Krishnan, professor at CSE, DSCE, Bangalore for his technical help and support.

REFERENCES

[1]. Moravec, “Visual mapping by a robot rover”. In International Joint Conference on Artificial Intellingence, pages 598–600, 1979.

[2]. C. Harris and M. Stephens, “A combined corner and edge detector,” in Alvey Vision Conference, pp. 147–151, 1988.

[3]. K. Mikolajczyk and C. Schmid. “Indexing based on scale invariant interest points”, Proc. ICCV, 1:525–531, 2001.

[4]. K. Mikolajczyk and C. Schmid, “Scale and Affine Invariant Interest Point Detectors”, International Journal of Computer Vision, 60(1):63–86, 2004.

[5]. T. Lindeberg and J. Garding, “Shape-adapted smoothing in estimation of 3-d depth cues from affine distortions of local 2-d brightness structure”, Proc. ECCV, pages 389–400, 1994.

[6]. A. Baumberg, “Reliable feature matching across widely separated views” Proc. IEEE CVPR, 1:774–781, 2000.

[7]. K. Mikolajczyk and C. Schmid, “An affine invariant interest point detector” Proc. ECCV, 1:128–142, 2002.

[8]. T. Tuytelaars and L. Van Gool. “Matching Widely Separated Views Based on Affine Invariant Regions”, International Journal of Computer Vision, 59(1):61–85, 2004.

[9]. J. Cooper, S. Venkatesh, and L. Kitchen, “The dissimilarity corner detectors,”International Conference on Advanced Robotics, pp. 1377–1382, 1991.

[10]. J. Cooper, S. Venkatesh, and L. Kitchen, “Early jump-out corner detectors,”IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 15,no. 8, pp. 823–828, 1993.

[11]. J. Matas, O. Chum, M. Urban, and T. Pajdla, “Robust wide-baseline stereo from maximally stable extremal regions”, Image and Vision Computing, 22(10):761–767, 2004.

[12]. P. Mus´e, F. Sur, F. Cao, and Y. Gousseau, “Unsupervised thresholds for shape matching. Image Processing, 2003” Proceedings. 2003 International Conference, 2, 2003.

[13]. P. Mus´e, F. Sur, F. Cao, Y. Gousseau, and J.M. Morel, “An A Contrario Decision Method for Shape Element Recognition”, International Journal of Computer Vision, 69(3):295–315, 2006.

[14]. F. Cao, J.-L. Lisani, J.-M. Morel, Mus´e P., and F. Sur, “A Theory of Shape Identification”, Number Vol. 1948 in Lecture Notes in Mathematics, Springer Verlag, 2008.

[15]. P. Monasse. Contrast invariant image registration. Proc. of the International Conf. on Acoustics, Speech and Signal Processing, Phoenix, Arizona, 6:3221–3224, 1999.

[16]. K. Mikolajczyk, T. Tuytelaars, C. Schmid, A. Zisserman, J. Matas, F. Schaffalitzky, T. Kadir, and L.V. Gool, A Comparison of Affine Region Detectors. International Journal of Computer Vision, 65(1):43–72, 2005.

[17]. D.G Lowe, “Distinctive image features from scale-invariant key points”, International Journal of Computer Vision, 60(2):91–110, 2004.

[18]. T. Lindeberg, “Scale-space theory: a basic tool for analyzing structures at different scales”, Journal of Applied Statistics, 21(1):225–270, 1994.

[19]. Y. Ke and R. Sukthankar, “PCA-SIFT: A more distinctive representation for local image descriptors”, Proc. CVPR, 2:506–513, 2004.

[20]. K. Mikolajczyk and C. Schmid, “A Performance Evaluation of Local Descriptors”, IEEE Trans. PAMI, pages 1615–1630, 2005

[21]. J.J. Foo and R. Sinha, “Pruning SIFT for scalable near-duplicate image matching”, Proceedings of the eighteenth conference on Australasian database-Volume 63, pages 63–71, 2007.

[22]. H. Lejsek, F.H. ´Asmundsson, B.T. J´onsson, and L. Amsaleg, “Scalability of local image descriptors: a comparative study”, Proceedings of the 14th annual ACM international conference on Multimedia, pages 589–598, 2006.

[23]. Herbert Bay, Tinne Tuytelaars, and Luc Van Gool, “SURF: Speeded Up Robust Features”, *Computer Vision ECCV*, Zurich, 2006, *Computer Vision and Image Understanding*, Vol 110, Issue 3, pages 346-359, June 2008.

[24]. Jean-Michel Morel and Goushen Yu, “ASIFT: A New Framework for Fully Affine Invariant Image Comparison” , Society for Industrial and Applied Mathematics Vol. 2, No. 2, pp. 438–469, 2009.

[25]. Andrea Lingua, David Marenchino, Francesco Nex, "Performance Analysis of the SIFT operator for Automatic Feature Extraction and Matching in Photogrammetric Applications", Article appeared in Sensors, 9(5), 3745-3766, doi: 10.3390/s90503745, ISSN 1424-8220, 2009.

[26]. Peter Kovesi, "Image Features from Phase Congruency", Department of Computer Science, the University of Western Australia, Nedlands, W.A. 6907, 1999.

[27]. Luca Lucchese, "Estimating Affine Transformations In The Frequency Domain", Department of Electrical and Computer Engineering, University of California at Santa Barbara, 2001.

[28]. Evelyn Brannock and Michael Weeks, "A Synopsis of RecentWork in Edge Detection using the DWT" *Proceedings of the IEEE SoutheastCon 2008*, Huntsville, Alabama, April 3-6, 2008, pages 515-520.

[29]. Wei-shi Tsai, "Contourlet Transforms for Feature Detection", 2008.

[30]. Jun–Wei Hsieh, Hong–Yuan Mark Liao, Kuo–Chin Fan‡, Ming–Tat Ko, and Yi–Ping Hung,: "Image Registration Using A New Edge–Based Approach", National Central University, Chung–Li, Taiwan, 1997.

[31]. Nick Kingsbury, "Multiscale Keypoint Detection Using the Dual-Tree Complex Wavelet Transform", *ICIP 2006: 1625-1628, 2006.*