

FEATURE EXTRACTION: AN APPLICATION TO OBJECT AND IMAGE RECOGNITION

Niveditta Thakur¹, Sunidhi Raju², Aayush Gupta³

Abstract-Today technology is all about Artificial Intelligence. AI is a field that has brought the world up to a next level in the arena of innovation and automation. Feature extraction is a basic step in image processing and many other AI techniques. Without features, any image or subject is exactly same as a blank sheet of paper. Even a slightest part taken from an image can help extracting important information about it by just observing and extracting the features. No matter it is data mining, clustering, market analysis or any other field, without the knowledge of features none of them can work accurately. Feature extraction is that important technique which helps to detect every basic detail. In the presented paper, the main focus is on various feature extraction techniques and how those can be beneficial to identify different subjects in an image with clustered scene.

Keywords- Feature Extraction Technique (FET), Histogram of Oriented Gradients (HOG), Local Binary Pattern (LBP), Scale Invariant Feature Technique (SIFT), Speeded Up Robust Feature (SURF).

1. INTRODUCTION

The work of discovering [1] resemblance between two different images of the same arena or object is part of many computer vision functions and operations. Identifying human faces or objects in images is a daring task owing to their irregular look and the large sum of postures that they can embrace. Today technology has arrived which can clearly identify objects and different features at their own level and skill like a human could do. The similar thing was impossible to imagine a few decades ago but is possible now. One of the vital tricks behind this miracle is FET which is today used in every different field.

The general problem that arises is that, which FET must be used for accurate information extraction. But explanation to same relies on the quality and quantity of information to be extracted. Also, it counts upon the number of calculations a single technique includes and the quality of result it gives. Though not any technique is superior to the other as each technique has its own merit and demerit. Major applications [2] for feature extraction are ANN (Artificial Neural Networks), image analysis, data mining, expert systems, etc. It is really mesmerizing how feature extraction works in all these fields and has so many stupendous applications. How does it work, is what this paper aims to explain.

Feature extraction [3] is a technique where various features are extracted such as facial expressions in faces, edges, colors, textures, gradient changes, etc. The existing methods for feature extraction range from low-level edge detection to high-level shape extraction. Each and every technique has its own unique ability to extract a feature. In this paper, all the techniques mentioned are tested in MATLAB R2015a using TID 2008 image database for test images.

This paper is systemized in the manner as Section I presents an Introduction to the FET, Section II presents a brief Literature Survey of various FET, Section III presents Work Implementation of this paper, Section IV presents Conclusion and Future Scope and finally References section.

2. LITERATURE SURVEY

In the presented paper, various FETs are studied and tested. Features are an important part of every system and can help in extracting important knowledge about the system.

2.1 Feature Extraction Techniques(FET) Classification

FET are classified [4] in many forms where some of them are given as follow:

2.1.1 Low-Level FET

Low level techniques are the [5] basic FET where we don't have to extract the exact or all data about spatial relation or shapes. For example, thresholding is low-level FET which is performed with point operation. These all low-level [6] FET can further be used in high level FET.

Some of the low-level FET are as follow:

¹ Faculty University Institute of Information Technology, Himachal Pradesh University, Shimla-171005, India.

² Student University Institute of Information Technology, Himachal Pradesh University, Shimla-171005, India.

³ Student University Institute of Information Technology, Himachal Pradesh University, Shimla-171005, India.

- First Order Edge Detection: These are those edge detection techniques which observe and detect only signals which change. Here the 1st order differentiation operation result is observed. Some of the 1st order edge detection operators are Canny, Sobel, Robert Cross, Prewitt, etc.
- Second Order Edge Detection: These are those edge detection techniques which are applied to observe zero-crossing where the change in signal is not constant as it is in 1st order edge detection operator's peak values. Some of the 2nd order edge operators are Laplacian, Laplacian of Gaussian, Mar-Hildreth, etc.
- Phase Congruency: It detects features in such a method that broad range of features is detected and also it is less affected by the noise presence in the subject or change in contrast or different illuminations.
- Localized Feature Extraction: This technique basically detects features which can't be detected by other detectors properly. Corners and junctions are two of those features. An example is: Curvature Detection (Corner detection).
- Image Motion Detection: Other than static or still features there are others too which are not still. This technique detects motion. Here more than one image or subject are observed and differentiated for the outcome to be the exact motion or change.

2.1.2 FET by Shape Matching

These are high-level FET [7] where shapes are extracted. These generally consist of techniques where position, size, and orientation are observed for shape matching. These techniques are quite helpful in extracting basic or simple shapes when complex or complicated images are observed. Some Shape matching FET are given as below:

- Template Matching: As the name suggests itself, templates are used for shape extraction. A template or a sub-image which consists the exact shape we need is compared with the test image.
- Hough Transform: This technique helps in extracting and detecting basic shapes in an image where they can be circle, square, line or any other shape.
- Generalized Hough Transform (GHT): There do exist some shapes or patterns which are not basic or simple. They can be complex which have unknown orientation, size or position. These are called arbitrary shapes and this technique is used to find such arbitrary shapes. It is similar to template matching but a full mapping is done in GHT.

2.1.3 Flexible Shape Extraction

No template matching or [7] Hough transform can be applied to some shapes as they have an undefined architecture which can change or are flexible in nature. These techniques are used to extract such flexible shapes by evolving the target in an image or by adapting the results of other techniques.

Following are three such techniques discussed:

- Deformable Templates: In this technique, the similar method of template matching is used but for deformable shapes.
- Active Contour (snakes): This technique targets to enclose a feature which is needed to be searched. It is identical to trap an object inside a balloon and then slowly discharging the air from the balloon after which we are left with the exact object. As we stop the shrinking of the balloon at the exact shape, similarly technique is stopped at a point where the exact shape is found.
- Shape Skeletonizing: Not only by size, position or orientation but shapes can also be defined by their axis or we can say their skeleton. A skeleton (similar to center axis) which is at equidistant from the whole shape boundary is used for shape extraction by distance transform technique.

2.2 Existing Feature Extraction Methods

2.2.1 SIFT

Scale-invariant feature transform (SIFT) [8] is a formula in computer vision to discover and describe native features in pictures. Its applications embrace visual perception, robotic mapping, and navigation, image stitching, 3D modeling, gesture recognition, video trailing, individual identification of life and match moving.

Main steps in [9] SIFT formula are:

- i. Scale-space extrema detection
- ii. Keypoint localization
 - Interpolation of close information for correct position
 - Deleting low-contrast keypoints
 - Erasing edge responses
- iii. Orientation assignment
- iv. Keypoint descriptor

2.2.2 SURF

In computer vision, [9] speeded up robust features (SURF) is a proprietary native feature detector and descriptor. It may be used for tasks like visual perception, image registration, classification or 3D reconstruction. It's partially galvanized by the scale-invariant feature transform (SIFT) descriptor. To discover interest points, SURF uses number approximation. Its feature

descriptor is predicated on the total of the Haar circular wave response round the mark of interest. SURF descriptors are mainly used to find and identify objects, individuals or faces, to reconstruct 3D scenes, to trace objects and to extract points of interest. Main steps in SURF formula for feature matching are Detection, Descriptor, and Matching.

2.2.3 BRISK

The **Binary Robust Invariant scalable Keypoints** [10] pattern is used to get pairwise brightness comparison results that are assembled into the binary BRISK descriptor. With a robust target potency of computation, BRISK conjointly exploits the speed savings offered within the SSE (Streaming SIMD Extensions) instruction set widely supported on today's architectures.

2.2.4 FREAK

The Fast retina Keypoint [11] is a deluge of binary strings computed by fast examination pairs of image intensities over a retinal sampling pattern. Apparently, choosing pairs to cut back the spatial property of the descriptor yields a highly-structured pattern that mimics the saccadic search of the human eyes.

2.2.5 LBP

Local binary patterns (LBP) [12] is a style of visually designed descriptor applied for classification in computer vision. LBP is also used in texture classification. It's been perceived that once LBP is combined with descriptor of HOG, it enhances the detection work somewhat on some datasets.

The LBP [13] feature vector, in its easiest type, is made within the following manner:

- i. Bisect the analysed window into cells (e.g. 16x16 pixels for every cell).
- ii. For every component in a cell, correlate the pixel to every of its eight neighbors (on its left-top, left-middle, left-bottom, right-top, etc.). Follow the pixels in a circle, i.e. dextral or counter-clockwise.
- iii. iii Wherever the middle pixel's worth is bigger than the neighbour's worth, write "0". Otherwise, write "1".
- iv. The reason the bar-graph, over the cell, of the frequency of every "number" occurring (i.e., every combination of those pixels is smaller and that are larger than the center).
- v. Optionally normalize the bar-graph.
- vi. Concatenate bar-graphs of every cell.

2.2.6 HOG

The histogram of oriented gradients [14] is a descriptor for features that are employed mainly for detection of objects. In this technique, gradient orientation in native parts of a picture is counted. This technique is alike to that of descriptors of SIFT, edge aligned histograms, and form contexts. However, it only differs in calculation of an impenetrable grid of uniformly spaced cells. Also, it adopts overlapping native contrast standardization for enhanced accuracy. During this technique, the image is split into little-connected regions known as cells, and for the pixels inside every cell, a bar chart or histogram of gradient angles is compiled. The descriptor is that which concatenates those histograms.

2.2.7 PHOG

In Pyramid of histogram of Orientation Gradient [15] native shape is entrapped or enclosed by the distribution over edge orientations inside an arena, and spatial layout by covering the image into regions at multiple resolutions. The descriptor includes a bar chart of orientation gradients over every image sub-region at every resolution level. The space between 2 PHOG image descriptors then displays the degree to an extent that the pictures has identical shapes and complements in their spatial layout.

3. WORK IMPLEMENTATION AND DISCUSSION

There is a huge range of features in an image and for every unique feature, there is a separate feature detection method. Recognizing a small subject in a huge scene can be done using such techniques. For example, in an astronomical branch, it can help identifying clusters, stars, galaxies, etc. in a whole image. In this paper, a similar approach is made and implemented on detecting small subjects in a whole clustered image.

In this paper, SURF FET is used. SURF is a [10] technique in which there are three processes. First, the interest points are searched with the help of detector which is repeated over in the image for similar points in case of different viewing situations. Here the reliability of the detector lies in its repeatability. In the next process, the nearby points of interest points are observed by the descriptor. Descriptor has to maintain its robustness towards the distortions and noise and at the same time sensitivity towards distinction. The last process is composed of matching of the interest points in different images or subjects.

Steps of Algorithm are given as below:

- Step 1. Input the target image.
- Step 2. Input the test image.
- Step 3. Use SURF FET for detecting strongest points in images.
 - a) 100 strongest points for the target image.
 - b) 300 strongest points for the test image.

- Step 4. Compare both images and calculate putatively matching points.
- Step 5. Calculate perfectly matching point.
- Step 6. Mark the region of perfectly matching points with a polygon.

3.1 Object Recognition

Here two images are considered where one is the Test image 1 and other is the Target image 1. The test image 1 is of a scene of flowers in front of a window where the target image 1 (flower) is to be searched using SURF FET.

Step 1: Input target image 1(Flower).

Step 2: Input test image 1(scene with cluttered objects).



Fig. 1 a) Target image 1 b) Test image 1

Step 3: Using SURF FET detect 100 strongest points in target image 1.

Step 4: Using SURF FET detect 300 strongest points in test image 1.

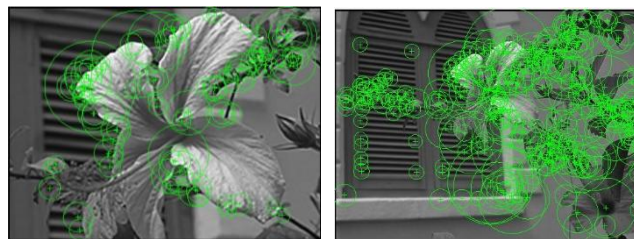


Fig. 2 a) 100 strongest points in Target image 1 b) 300 strongest points in Test image 1

Step 4. Compare both images and calculate putatively matching points.

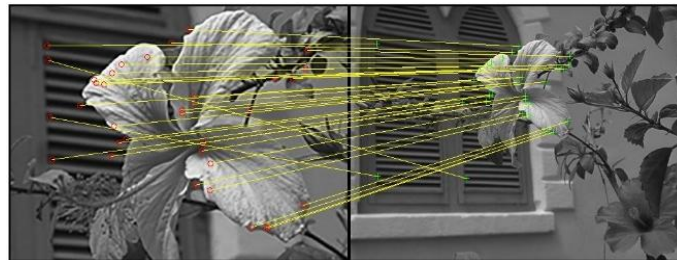


Fig. 3 Putatively Matching points in both images.

Step 5. Calculate perfectly matching point.

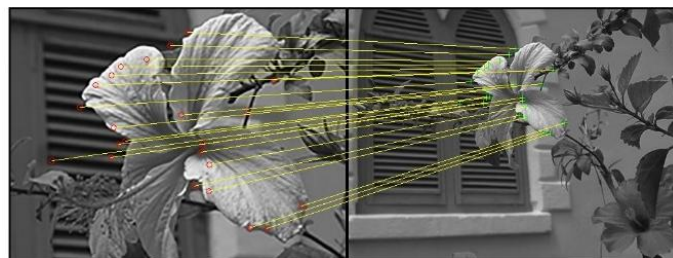


Fig. 4 Exactly Matching Points in both images.

Step 6. Mark the region of perfectly matching points with a polygon.



Fig. 5 Detected target in Test image 1.

3.2. Image Recognition

Again two images are taken for recognition, one is the Test image 2 and other one is the Target image 2. The test image 2 is of a scene where faces of many people are together and the target image 2 is an image of a single person which is to be searched using SURF FET.

Step 1: Input target image 2(Face of one person).

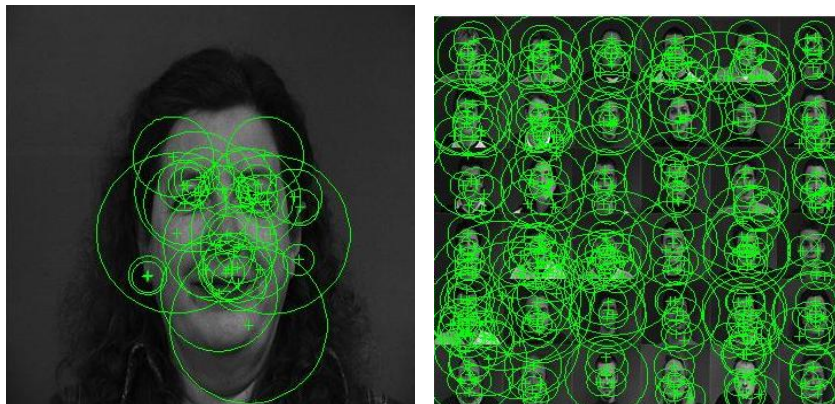
Step 2: Input test image 2(scene with many people).



(a) (b)
Fig. 6 a) Target image 2 b) Test image 2

Step 3: Using SURF FET detect 100 strongest points in target image 2.

Step 4: Using SURF FET detect 300 strongest points in test image 2.



(a) (b)
Fig. 7 a) 100 strongest points in Target image 2 b) 300 strongest points in Test image 2

Step 4. Compare both images and calculate putatively matching points.



Fig. 8 Putatively Matching points in both images.

Step 5. Calculate perfectly matching point.



Fig. 9 Exactly Matching Points in both images.

Step 6. Mark the region of perfectly matching points with a polygon.



Fig. 10 Detected target in Test image 2.

4. CONCLUSION AND FUTURE SCOPE

The overall intention of this paper is to know the fact that features are basic and vital part of every system and extraction of features is a very beneficial and important process. In this paper, various FETs are studied and tested. Each feature has its unique extraction method and all of them performed well in their job. No matter it is a low-level FET like thresholding, phase congruency or high-level like shape matching, every technique is designed and developed in such a way that it can obtain perfect features and shapes. Also, the SURF FET is successfully implemented in recognizing an object in a cluttered scene or a face of the person in a crowded scene.

Today such FETS exist which are not versatile. In image processing world, every different subject has different information conveying features. Extracting such features need different techniques or methods. Also, the results are somewhat near to accuracy. These issues could be worked upon and an effort could be made to design better and versatile FETs.

5. REFERENCES

- [1] M. Agrawal, K. Konolige, and M. Blas. "Censure: Center surround extremas for realtime feature detection and matching," 10th European Conference on Computer Vision, pp. 102–115, 2008.
- [2] M. Sahu, A. Saxena and M. Manoria, "Application of Feature Extraction Technique: A Review," International Journal of Computer Science and Information Technologies, Vol. 6 (3), pp. 3014-3016, 2015.
- [3] Isabelle Guyon and Andr e Elisseeff, "An Introduction to Feature Extraction," Journal of Machine Learning Research, pp. 18-23, 2003.
- [4] Seyyid Ahmed Medjahed , "A Comparative Study of Feature Extraction Methods in Images Classification," International Journal of Image, Graphics and Signal Processing, Vol. 3, D.O.I: 10.515, pp. 16-23, 2015
- [5] Xiang Sean Zhou, Ira Cohen, Qi Tian and Thomas S. Huang, "Feature Extraction and Selection for Image Retrieval," in ACM Multimedia 2000, pp. 1-7, 2000
- [6] Dilipsinh Bheda, Mahasweta Joshi, Vikram Agrawal, "A Study on Features Extraction Techniques for Image Mosaicing," International Journal of Innovative Research in Computer and Communication Engineering, Vol. 2, Issue 3, ISSN: 2320-9801, pp. 3432-3437, 2014.
- [7] Mark S. Nixon, Alberto S. Aguado, Feature Extraction and Image Processing on Computer Vision, 3rd ed., Academic Press, pp. 13-40, 2008.
- [8] Liang-Chi Chiu, Tian-Sheuan Chang, Jiun-Yen Chen, and Nelson Yen-Chung Chang, "Fast SIFT Design for Real-Time Visual Feature Extraction," IEEE Transaction on Image Processing, Vol. 22, Issue 8, pp. 3158-3167, 2013.
- [9] David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, Vol. 60, Issue 2, pp. 91 – 110, 2004.
- [10] Herbert Bay, Tinne Tuytelaars, and Luc Van Gool , "SURF: Speeded Up Robust Features," European Conference on Computer Vision, pp. 20-46, 2006.
- [11] Stefan Leutenegger, Margarita Chli and Roland Y. Siegwart, "BRISK: Binary Robust Invariant Scalable Keypoints," 11 Proceedings of the 2011 International Conference on Computer Vision, pp. 2548-2555, 2011.
- [12] Alexandre Alahi, Raphael Ortiz and Pierre Vandergheynst, "FREAK: Fast Retina Keypoints," IEEE Conference on Computer Vision and Pattern Recognition, ISSN: 1063-6919, pp. 1-8, 2012.
- [13] Md. Abdur Rahim, Md. Najmul Hossain, Tanzillah Wahid & Md. Shaful Azam, "Face Recognition using Local Binary Patterns (LBP)," Global Journal of Computer Science and Technology Graphics & Vision, Volume 13 Issue 4, ISSN: 0975-4172, pp. 1-7, 2013.
- [14] Timo Ojala, Matti Pietik inen and Topi M aenp a, "Multiresolution Gray Scale and Rotation Invariant Texture Classification with Local Binary Patterns," IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 24, Issue 7, pp. 971-987, 2002.
- [15] Takumi Kobayashi, "BoF meets HOG: Feature Extraction based on Histograms of Oriented Gradients for Image Classification," IEEE Conference on Computer Vision and Pattern Recognition, ISSN: 1063-6919, pp. 747-754, 2013.
- [16] A. Bosch, A. Zisserman, X. Munoz, "Representing shape with a spatial pyramid kernel", Proceedings of the International Conference on Image and Video Retrieval, pp. 401-408, 2007.