Analysis of Interest Point Distribution in SURF Octaves

Shoaib Ehsan, Nadia Kanwal, Erkan Bostanci, Adrian F. Clark and Klaus D. McDonald-Maier
School of Computer Science & Electronic Engineering
University of Essex
Colchester, United Kingdom
{sehsan, nkanwa, gebost, alien, kdm} @ essex.ac.uk

***Abstract*—Speeded-Up Robust Features (SURF) is a state-of-the-art, scale- and rotation-invariant feature extraction technique with the potential for real-time execution. Although SURF has been extensively employed for multi-scale computer vision applications since its inception, there are still some areas of this computationally complex algorithm that have not been fully explored and require detailed analysis to enable algorithm-level optimization of SURF for real-time execution. In particular, the distribution of interest points in SURF octaves is a topic that requires thorough investigation. Contrary to the present perception, this paper demonstrates that there is a possibility of higher octaves being more significant than the lower octaves in terms of detected interest points for real-life images. The paper also shows that variation of blob response threshold has a significant effect on interest point distribution. The results presented highlight the need of developing a systematic approach to SURF octave selection.**

Keywords- Image analysis, feature extraction, SURF, interest point distribution.

I. INTRODUCTION

Speeded-Up Robust Features (SURF) is a state-of-the-art computer vision technique which is focused on fast detection, description and matching of scale- and rotation-invariant image features [1]. Although incapable of achieving real-time performance with software-only implementations on modern desktop computers due to its high computational complexity, SURF is still attractive in terms of execution speed, and with comparable results, when contrasted with other contemporary algorithms for feature extraction, such as the Scale Invariant Feature Transform (SIFT) and Harris-Laplace feature detector [1, 2]. This speed advantage has been the real factor behind its popularity and has led to exciting SURF-based vision applications like an interactive museum guide, retina mosaicing and mobile augmented reality on a handheld platform [3, 4, 5].

It is interesting that, despite this popularity, there are still some areas of this computationally complex algorithm that have not been fully explored and require a more detailed analysis and understanding. Since algorithm analysis is the first step towards algorithm optimization, investigation of these unexplored areas is vital for algorithm-level optimization of SURF, in an effort to reduce its computational complexity and improve its performance both in terms of execution speed and accuracy. As the range of embedded vision applications is becoming broader and boarder, algorithm-level optimization is gaining more significance for a computationally complex algorithm like SURF to handle critical issue of power consumption in battery powered systems. SIFT, the main competitor of SURF, has undergone extensive algorithm-level optimizations to enhance its performance, and this has led to a number of variants, including PCA-SIFT, GLOH and RIFT [6, 7, 8]. In fact, SURF itself can be considered a descendant of SIFT. Thus, critical analysis of every stage of SURF is required to identify areas that can be optimized further, not only to improve its performance in terms of execution speed and accuracy but also to make it more suitable for low-power, embedded vision applications.

This paper is a step forward in this direction. It attempts to carry out an in-depth study of interest point distribution in SURF octaves – a topic that needs thorough investigation due to its significance for finding the optimal number of octaves required for any particular SURF-based vision application. The intention here is firstly to examine the general trend of interest point distribution in SURF octaves across a wide range of images which includes standard image data sets and images captured from real-life applications. The contribution of each octave is analyzed to identify octaves that are essential to the output of the algorithm in terms of detected interest points. The second part of this paper explores the relation between blob response threshold and interest point distribution as it is vital for determining the consistency of the distribution. To our knowledge, this is the first attempt to study the effect of variation of blob response threshold on interest point distribution.

The remainder of this paper is structured as follows. Section II provides a review of SURF and analyzes the general trend of the interest point distribution in SURF octaves. The effect of variation of blob response threshold on interest point distribution is investigated in Section III. Finally, conclusions are presented in Section IV.

II. ANALYSIS OF INTEREST POINT DISTRIBUTION

This section provides a brief overview of the SURF algorithm and then analyzes the general trend of interest point distribution in SURF octaves. The key stages of SURF are shown in Fig.1. Since SURF aims to extract scale-invariant image features, it pursues this objective by calculation of blob response maps at different scales to

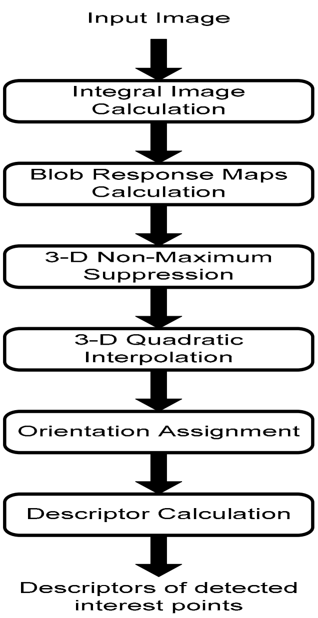

Fig. 1 The key stages of SURF

implement an image pyramid for scale-space analysis. The scale space is divided into number of octaves, formed by grouping blob response maps for adjacent scales. 3-D non-maximum suppression is then carried out to determine local maxima. In order to achieve sub-pixel, sub-scale accuracy, 3-D quadratic interpolation is done to provide interest points. A blob response threshold is normally applied to select high-contrast interest points. The descriptors for the detected interest points are calculated after orientation assignment to achieve rotation invariance.

SURF calculates a series of blob response maps at different scales by convolving the same input image with a filter of increasing size, as opposed to sub-sampling the input image. The creation of scale space starts by applying a 9 x 9 filter and then the size of the filter is increased with every increment in scale. Specifically, 9 x 9, 15 x 15, 21 x 21 and 27 x 27 filters are used to calculate blob response maps for the first octave. However, the filter size can only be increased as long as it is smaller than the input image size. Since each octave consists of a fixed number of scales, the total number of SURF octaves that may be processed is limited by the size of the input image. For example, a maximum of 6 octaves can be processed for an input image of resolution 640 x 480 pixels if the number of scales in each octave is 4.

According to [1], as we move from lower to higher octaves, the number of detected interest points per octave decays quickly. It is important to note here that this conclusion is based upon experiments performed on standard data sets [9] only, which are more suitable for performance evaluation of SURF under viewpoint, illumination, rotation and scale changes. A detailed analysis has not been carried out and there are some important questions which are unanswered:

a) Is this behavior of octaves consistent for all types of images?
b) Is a lower octave always more dominant in terms of detected interest points than higher octaves? More specifically, does octave 1 always detect more interest points than octaves 2, 3 and 4?
c) Assuming that the contribution of higher octaves becomes more and more negligible with increasing octaves, is it always justified to reject the higher octaves in favor of lower octaves when ever improvements in the execution speed is required?
d) Is there any effect of blob response threshold on interest point distribution?

These questions are significant for finding the optimal number of octaves for any SURF-based vision application. Therefore, a thorough investigation is required to determine the general trend of interest point distribution across a wide range of images.

This paper seeks to answer the above questions. For this particular analysis, more than 20,000 real-life images of different resolutions were tested using a MATLAB implementation of SURF to gain insight into the general trend of interest point distribution. The test images included images from standard data sets [9], Google image database [10] and images captured specifically for this work; some of them are shown in Fig. 2. The number of octaves and the number of scales per octave were set to four for this analysis where octave 1 was the lowest octave and octave 4 was the highest octave.

Fig. 2 Sample test images. Top left is image 1 which is followed by images 2, 3, 4 and 5 to the right. Bottom left is image 6 which is followed by images 7, 8 , 9 and 10 to the right.

TABLE I INTEREST POINT DISTRIBUTION WITHOUT APPLYING BLOB RESPONSE THRESHOLD

Image	Resolution	Octave 1	Octave 2	Octave 3	Octave 4
1.	640 x 480	3400	2326	991	224
2.	800 x 640	6172	3188	1124	322
3.	1000 x 700	7774	3604	1388	394
4.	1280 x 960	14854	7278	2696	785
5.	1280 x 960	15944	7609	3136	1040
6.	640 x 480	4061	2645	1300	432
7.	640 x 480	3082	2101	779	184
8.	640 x 480	2225	1867	646	84
9.	640 x 480	4107	2598	1021	229
10.	640 x 480	3851	2693	1008	234

TABLE II INTEREST POINT DISTRIBUTION WITH BLOB RESPONSE THRESHOLD OF 50,000

Image	Resolution	Octave 1	Octave 2	Octave 3	Octave 4
1.	640 x 480	217	126	90	49
2.	800 x 640	590	517	238	76
3.	1000 x 700	43	75	47	22
4.	1280 x 960	29	40	69	61
5.	1280 x 960	49	57	61	45
6.	640 x 480	29	19	22	10
7.	640 x 480	55	192	128	84
8.	640 x 480	0	23	15	18
9.	640 x 480	22	25	72	22
10.	640 x 480	9	17	20	19

The results for the test images shown in Fig. 2 are listed in Tables I and II. Table I provides information about interest point distribution without a blob response threshold applied, whereas Table II presents results with a blob response threshold of 50,000. A general trend of interest point distribution can easily be identified from Table I for the case when no blob response threshold is applied. The lowest octave, octave 1, appears to be the most dominant in terms of detected interest points, followed by octave 2; whereas octave 4 detects the least number of interest points.

The situation, however, does not remain the same when a threshold of 50,000 is applied to reject low-contrast interest points. In Table II, the interest point distribution for images 1 and 2 follows the same pattern as in the case of no threshold. However, for image 3 it can be seen that octaves 2 and 3 detect more interest points than octave 1. The distribution is more interesting for images 4, 8, 9 and 10, as the higher octaves (3 and 4) dominate the lower octaves 1 and 2. For image 4, more than 65% of the interest points are detected in octaves 3 and 4. Hence, rejecting octaves 3 and 4 in favor of octaves 1 and 2 for image 4 is certainly not justified. This particular case indicates that selecting the lower octaves only for enhancing execution speed is not always a good option and highlights the need for a more systematic procedure for selecting SURF octaves. Fig. 3 shows the interest point distribution for images 2, 4, 7 and 9. From these results, it can be concluded that there is a strong possibility of higher octaves dominating the lower octaves in terms of detected interest points when blob response threshold is applied to reject the low contrast interest points.

III. EFFECT OF BLOB RESPONSE THRESHOLD

This section investigates the relationship between blob response threshold and interest point distribution. From the results presented in Section II, it is evident that, when blob response threshold is applied, any octave can dominate in terms of detected interest points. However, it is important to probe whether interest point distribution is affected by any variation in blob response threshold or not.

For this particular analysis, the blob response threshold was varied from 0 to 60,000 with a step size of 5,000 and the resulting interest point distribution examined for every threshold value. For the purpose of discussion, the results for images 1 and 9 are utilized as they represent two dominant trends with respect to blob response threshold variation. Fig. 4 shows the break-up of interest point distribution for image 1 at different blob response thresholds. It can be observed that more than 70% of the interest points are detected in octaves 1 and 2 at every threshold value. The contribution of octaves 3 and 4 increases with increasing threshold values but the interest point distribution is largely consistent across threshold values.

Fig. 5 shows the break-up of the interest point distribution for image 9 at different blob response thresholds. This case is interesting: with a threshold value of 5,000, 85.42% of all detected interest points are in octaves 1 and 2. However as the threshold value increases, the contribution of octaves 3 and 4 becomes more and more significant and less than 40% of the interest points are detected in octaves 1 and 2 at threshold values beyond 30,000. This is a sharp contrast to the trend observed in the case of image 1. It can be concluded that for any image, lower octaves always detect

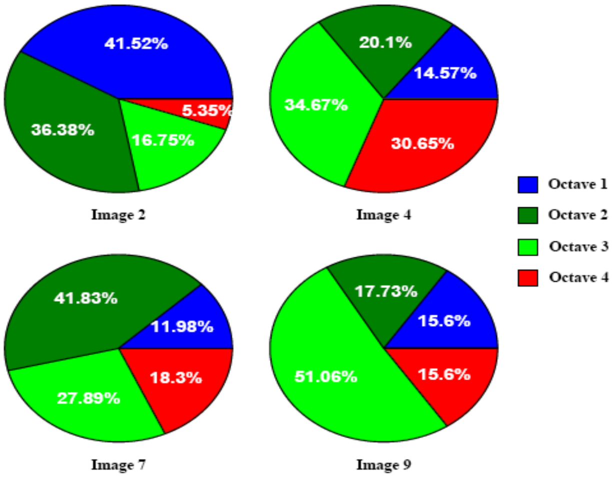

Fig. 3 Interest Point distribution with threshold of 50,000 for images 2, 4, 7 and 9

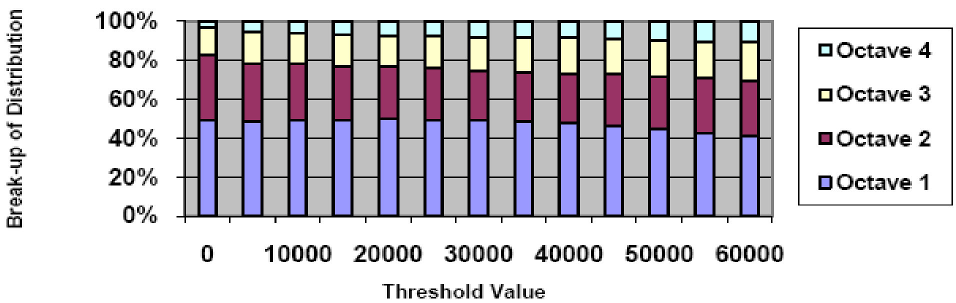

Fig. 4 Effect of variation of blob response threshold on interest point distribution for image 1

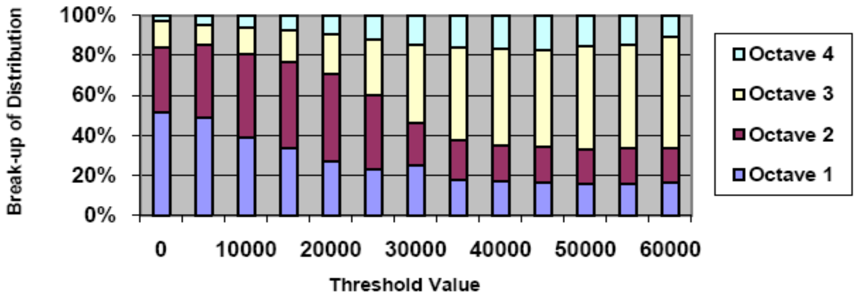

Fig. 5 Effect of variation of blob response threshold on interest point distribution for image 9

more interest points than higher octaves at lower thresholds and may continue their dominance at high threshold values too. However, there is also a possibility of higher octaves becoming more critical than lower octaves at high threshold values.

IV. Conclusions

This paper has examined the general trend of interest point distribution in SURF octaves by analyzing a wide range of real-life images. It has demonstrated that a lower octave always dominates a higher octave in terms of detected interest points if no blob response threshold is applied. However, the pattern is not the same when a threshold is applied for the rejection of low contrast interest points as there is a strong possibility of higher octaves dominating lower ones, contrary to current perception. The effect of variation of blob response threshold on interest point distribution has also been investigated by this paper, and it has been found that lower octaves always dominate higher ones at lower thresholds, a dominance that may continue at high thresholds. However, there is a strong possibility of the converse happening at high thresholds. The results presented highlight the need for a more systematic approach to SURF octave selection in order to improve performance, both in terms of accuracy and execution speed.

REFERENCES

[1] Bay, H., Tuytelaars, T. and Gool, Luc V., "Speeded Up Robust Features (SURF)," in Computer Vision and Image Understanding, Vol. 110, No. 3, pp. 346-359, June 2008.

[2] Tuytelaars, T. and Mikolajczyk, K., "Local Invariant Feature Detectors: A Survey," in Foundations and Trends in Computer Graphics and Vision, Vol. 3, No. 3, pp. 177-280, 2008.

[3] Bay, H., Fasel, B. and Gool, Luc V., "Interactive Museum Guide: Fast and Robust Recognition of Museum Objects," in Proceedings of the First International Workshop on Mobile Vision, May 2006.

[4] Cattin, P. C., Bay, H., Gool, Luc V. and Szekely, G., "Retina Mosaicing using Local Features," in MICCAI, No. 2, pp. 185-192, October 2006.

[5] Lee, S. E., Zhang, Y., Fang, Z., Srinivasan, S., Iyer, R. and Newell, D., "Accelerating Mobile Augmented Reality on a Handheld Platform," in Proceedings of the 27th IEEE International Conference on Computer Design,October 2009.

[6] Ke, Y. and Sukthankar, R., "PCA-SIFT: A More Distinctive Representation for Local Image Descriptors," in Proceedings of CVPR, 2004.

[7] Mikolajczyk, K., and Schmid, C., "A performance evaluation of local descriptors", IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 27, No. 10, pp. 1615--1630, 2005.

[8] Lazebnik, S., Schmid, C., and Ponce, J., "Semi-Local Affine Parts for Object Recognition," in Proceedings of the British Machine Vision Conference, 2004.

[9] http://www.robots.ox.ac.uk/~vgg/research/affine/,accessed September 22, 2010.

[10] http://images.google.co.uk/imghp?hl=en&tab=wi,accessed September 30, 2010.