



# **A Review:SIFT Hardware Implementation For Real Time Feature Extraction**

Shital S. Sakhare<sup>1</sup>, S. C. Wagaj, Professor<sup>2</sup>

PG Student, Dept. of E&TC, Rajarshi Shahu College of Engineering, Tathawade, Pune, Maharashtra, India <sup>1</sup>

Assistant Professor, Dept. of E&TC, Rajarshi Shahu College of Engineering, Tathawade, Pune, Maharashtra, India <sup>2</sup>

**ABSTRACT:** In pattern recognition and image processing, feature extraction is simple form of dimensionality reduction. The transformation of the input data's into a set of features is known as feature extraction. The large set of data's is to be analyzed and performed accurately from the features of the input data. By using SIFT (Scale Invariant feature transform) the hardware resources can be minimized and it could be performed as a process of parallel and the pipeline based VLSI architecture. We propose two parallel SIFT feature extraction algorithms using general multi-core processors, as well as some techniques to optimize the performance on multi-core. The proposed architecture led to a 6.7x faster speed on a dual-socket, quad-core system, which facilitated an average 45 frames/second for a VGA (640×480) video. Some implementations and accelerations of SIFT feature extraction on graphics processing units (GPUs) were introduced. With the parallelism and powerful computational ability, they achieved high processing speed.

**KEYWORDS:** FPGA (SPARTAN 3), Feature extraction, SIFT algorithm

## **I.INTRODUCTION**

In many computer visions application feature extraction is used to detect and describe local features in images. SIFT keypoints of objects are first extracted from a set of reference images and stored in a database. An object is recognized in a new image by individually comparing each feature from the new image to this database and finding candidate matching features based on Euclidean distance of their feature vectors. From the full set of matches, subsets of keypoints that agree on the object and its location, scale, and orientation in the new image are identified to filter out good matches. In SIFT image is transformed in to large collection of feature vectors, each of which is invariant to image translation, scaling, and rotation, partially invariant to illumination changes and robust to local geometric distortion. These features share similar properties with neurons in inferior temporal cortex that are used for object recognition in primate vision. Key locations are defined as maxima and minima of the result of difference of Gaussians function applied in scale space to a series of smoothed and resampled images. Low contrast candidate points and edge response points along an edge are discarded. Dominant orientations are assigned to localized keypoints. These steps ensure that the keypoints are more stable for matching and recognition. SIFT descriptors robust to local affine distortion are then obtained by considering pixels around a radius of the key location, blurring and resampling of local image orientation planes.

In existing there are many methods available which uses approximate nearest-neighbor lookup, a Hough transform for identifying clusters that agree on object pose, least-squares pose determination, and final verification. Other potential applications include view matching for 3D reconstruction, motion tracking and segmentation, robot localization, image panorama assembly, epipolar calibration, and any others that require identification of matching locations between images.

In this paper a new algorithm for descriptor generation is proposed with square sub-regions arranged in 16 directions to achieve rotation invariance, thus we can not only improve the parallelism of the algorithm, but also avoid floating calculation to save hardware resource consumption. With pipeline architecture to implement the descriptor generation module, the system achieves nearly 15 times higher processing speed than a recently developed solution.



# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

The rest of the paper is organized as follow; the section II describes literature survey in short. Section III illustrates the SIFT algorithm. Section IV describes the proposed hardware architecture of SIFT

## II.LITERATURE SURVEY

Many researchers have worked feature extraction techniques using SIFT such as at the first time Harris detector algorithm is evolved which is capable of immunity and invariance to intensity and rotation changes [1]. Then the SIFT algorithm proposed by Lowe [2]. Then Zhang introduced two parallel SIFT feature extraction algorithms using general multi-core processors, as well as some techniques to optimize the performance on multi-core. The proposed architecture led to a 6.7x faster speed on a dual-socket, quad-core system, which facilitated an average 45 frames/second for a VGA (640×480) video [3].

Kim et al introduced an object recognition processor, which was integrated with ten processing units for task-level parallelism; and it contained a single instruction multiple data (SIMD) instruction to exploit the data-level parallelism. However, the system can only detect the feature point without generating the descriptor, and the processing time for a QVGA (320×240) video was 42ms when the operation frequency was 200MHz, which was not sufficient for real-time application [4]. Large number of researchers have successfully applied FPGA to speed up the SIFT algorithm.

Bonato proposed a hardware architecture of SIFT which is able to detect features up to 30 frames per second for a QVGA video but to generate descriptor this system is get failed[5]. Yao proposed an architecture of optimized SIFT feature detection for an FPGA implementation of an image matcher. The feature detection module took 31 ms to process a typical VGA image [6]. E.S.Kim proposed a new hardware organization to implement SIFT with less memory and hardware cost but only 553 feature points/frame can be process for VGA (640 × 480) image at 30 frames/s [7]. Wang proposes a new FPGA based embedded system architecture for feature detection and matching but the robustness to rotation and scale change of the proposed method is weak [8]. Zhong presents a low-cost embedded system based on a new architecture that integrates FPGA and DSP [9]. Huang analyzed the time consumption of each part of SIFT by running the SIFT algorithm on a 2.1GHz Intel CPU and a soft-core of 100MHz 32-bit NIOS II CPU [10].

## III. SIFT ALGORITHM

SIFT algorithm is divided in to two parts, one for key point detection and other for descriptor generation. The steps for the algorithm are :-

1. Constructing DOG Pyramid which contains 2 octaves and 4 scales per octave. The DOG space is constructed by applying the subtraction of two nearby Gaussian scale images.
2. Key point detection in which feature points are chosen from the local maxima or minima in the DoG space. Each pixel in the middle scale of DoG space is compared with its 26 neighbor pixels, where there are 8 pixels in current scale image, and 9 neighbors in the scale above and below.
3. Calculation of Gradient Magnitude and Orientation The magnitude-orientation histogram is used to describe a feature point, which is computed from the gradient magnitude and orientation of the neighbor pixels around the candidate feature point.
4. Main Orientation of Feature Point where The orientation histogram has 36 bins, which cover the 360 degree range of orientations.
5. SIFT Descriptor Generation 16×16 window region around feature point is selected and rotated to the main orientation. Then, the direction of each pixel in this region is transferred to 8 bins, in the main orientation, and the gradient magnitude is weighted by Gaussian kernel with a scale equal to one half of the width of the descriptor window. After that, the 16×16 window is divided into sixteen 4×4 sub-regions. Finally, an orientation histogram of 8 bins is generated for each 4×4 sub-regions. Overall, the SIFT descriptor will be represented by a vector having  $16 \times 8 = 128$  elements.

## IV.SIFT SYSTEM

The SIFT algorithm mainly consists of two modules: the key point detection module and the descriptor generation module.



# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2015

In this paper, two symmetric RAMs are adopted with ping-pong operation to run the key point detection module and descriptor generation module in task level parallelism. The RAM block consists of an even RAM and an odd RAM to buffer locations of the key point, the gradient magnitudes and the directions. The key point detection module processes the image in data stream manner continuously and the results of even frame and odd frame are saved to the even RAM and odd RAM respectively. On the other hand, the descriptor generation module reads the result of the previous frame and generates a descriptor for each key point. With the parallel architecture, the processing speed of our system is determined by the longest time consumed by the two modules. In this paper, the RAM has an ability to save descriptors for up to 2900 key points, and the time consumed by descriptor module is less than the key point detection module. The procedures in key point detection module are processed with pipeline structure, and the descriptor generation module consists four blocks, running in pipeline-parallel structure.

## V. CONCLUSION

The table given below shows performance comparison of various recent works on FPGA such as Complexity, Timing parameter, frequency, image size .

TABLE 1: COMPARISON OF PARAMETERS OF VARIOUS TECHNIQUES

Parameter	Bonato	Yao	Huang	Proposed System
Size of image	320×240	320×240	320×240	320×240
Time , frequency of key point detection	33ms, 50MHz	31ms, 100MHz	3.4ms, 100MHz	6.55ms, 50MHz
Time , frequency of desriptor detection	11.7ms, 100MHz	Not Given	0.0331 ms, 100MHz	0.00223 ms, 100MHz
Platform for implementation	STRAT IX II	Virtex5	ASIC	SPATAN 3
Octave ,scale	3,6	2,4	3,6	2,4
Overall time	Not Given	Not Given	33ms 890keyp oints	6.55ms 2900keyp oints

A new algorithm for descriptor generation is proposed with square sub-regions arranged in 16 directions to achieve rotation invariance, thus we can not only improve the parallelism of the algorithm, but also avoid floating calculation to save hardware resource consumption. With pipeline architecture to implement the descriptor generation module, the system achieves nearly 15 times higher processing speed than a recently developed solution.

## REFERENCES

- [1] C.Harris and M.J.Stephens,"ACombined Coner and Edge Detector."in Proc. Of the Avley Vision Conf., Manchester, UK, 1988, pp. 147-152.
- [2] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Key points," Int'l Journal of Computer Vision, vol. 60, no. 2, pp. 91-110, Jan.2004
- [3] Q. Zhang and Y. R. Chen, "SIFT Implementation and Optimization For Multi-Core Systems," in Proc. of IEEE Int'l Symposium on Parallel and Distributed, pp. 1-8, 2008.
- [4] D. Kim, K. Kim, J. Y. Kim, S. Lee,S. J. Lee, and H.J. Yoo, "81.6 GOPS object recognition processor based on a memory-centric NoC," IEEE Trans. Very Large Scale Integr. (VLSI) Syst., vol. 17, pp. 370-383, Mar.2009
- [5] V. Bonato, "A parallel hardware architecture for scale and rotation invariant feature detection," IEEE Trans.Circuits Syst. Video Technol. vol. 18, pp. 1703-1712, Dec.2008.



ISSN (Print) : 2320 – 3765  
ISSN (Online): 2278 – 8875

# International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering

(An ISO 3297: 2007 Certified Organization)

**Vol. 4, Issue 4, April 2015**

- [6] L. F. Yao, "An Architecture of Optimised SIFT Feature Detection for an FPGA Implementation of an Image Matcher," in Proc. of Int'l Conf. on Field-Programmable Technology, pp. 30-37, 2009.
- [7] E. S. Kim, H. J. Lee, "A novel hardware design for SIFT generation with reduced memory requirement," *Journal of Semiconductor Technology And Science*, vol.13, no.2, Apr.2013.
- [8] J. Wang, Z. Sheng, L. Yan, Z. Cao, "An Embedded System-on-a-Chip Architecture for Real-time Visual Detection and Matching", *IEEE Trans.Circuits Syst. Video Technol.*, Accepted for publication, 2013.
- [9] S. Zhong, J. Wang, L. Yan, L. Kang, and Z. Cao, "A real-time embedded architecture for SIFT," *Journal of Systems Architecture*, vol. 59, no. 1, pp.16–29, Jan 2013.
- [10] F. Ch. Huang, "High-Performance SIFT Hardware Accelerator for Real-Time Image Feature Extraction," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 22, pp. 340-351, Mar.2012.