A SIFT-based Partial Processing Algorithm for Real Time Image Search

Neha Jayan, Computer Science

Mentor: Dr. John Korah

Kellogg Honors College Capstone Project

Abstract

Today's digital age can be defined by the massive amount of data and information readily accessible in various forms of media. While images have become a popular medium of information, current methods for relevant information search and retrieval face significant challenges. Due to rapidly growing data ingestion rates, current image retrieval methods have led to increased lag time as well as the demand for massive and expensive computational resources. Therefore, the need for an efficient process of searching data for immediate use has become critical. This research project emphasizes the partial processing paradigm for the design and implementation of algorithms for real time image search. Specifically, we will be focusing on content-based image retrieval methods, where the images are analyzed by their features such as color, texture, and shape. Scale Invariant Feature Transform (SIFT) [1], is one of the more robust feature detection algorithms that we have chosen to work with due to its ability to extract features amidst scaling, orientation, and illumination changes. We provide initial exploration of the aspects of SIFT based image search algorithms that can successfully leverage partial processing and can be successfully deployed on one or more parallel/distributed platforms.

Background: Image Retrieval

Image retrieval can be defined as a process of searching and retrieving images from a large database. The image retrieval research can be broadly classified into two subareas: Text Based Image Retrieval (TBIR) and Content Based Image Retrieval (CBIR). The earlier algorithms, known as text-based image retrieval (TBIR), utilizes keywords, captions, or descriptions associated with an image. CBIR methods try to retrieve images similar to a user-defined specification or pattern such as a sketch or query image. The objective is to extract visual content such as shape, color, and texture. One of the main advantages of CBIR is its possibility of being an automatic retrieval process, instead of the traditional keyword-based approach (TBIR), which usually requires time-consuming previous annotation of database images. While this could explain the transition from using more TBIR to CBIR, current image retrieval methods in general are not highly effective in light of rapidly growing databases and increased complexity of images. Current methods rely on pre-processing the data using a technique called indexing, where each document is processed entirely before similarities are calculated. Fully processing and searching through a growing database is the critical issue.

The Anytime Algorithm

In consideration of the problems dealing with current image retrieval methods, this prompts us to question if we can rather partially process data. From this, we may be able to achieve efficiency by reallocating computational resources to the relevant images according to the analysis of partial results. If we can partially process data, we may be able to prevent the wastage of computational resources on non-relevant images and speed up the process by identifying relevant images through partial similarity. To achieve this, we plan to consider algorithms with anytime property. Anytime algorithms have the four main properties [2]:

- Interruptibility: The algorithm can be stopped at any time and provide a result.
- Monotonicity: The quality of the result is a monotonically non-decreasing function of processing time and input quality.
- Measurability: The quality of an approximate result can be determined precisely. For example, when the quality reflects the distance between the approximate result and the correct result, it is measurable as long as the correct result can be determined.
- Preempt-ability: The algorithm can be suspended and resumed with minimal overhead.

As more execution time correlates with better solution quality, more processing will typically lead to better results. Considering that CBIR methods have the potential of becoming automatic retrieval processes, we have examined one of the prominent CBIR methods for our study.

Content-Based Image Retrieval

Content-based image retrieval (CBIR) can be defined as analyzing the visual features of an image rather than the keyword tags or descriptions associated with the image. CBIR methods can specifically be classified based on what they do. The kind of CBIR method we have selected for our study is key point extraction. This can be defined as the process of automatically computing a compact numerical representation of some qualitative attribute (feature) of digital images. These methods are feature-extraction based and involve a procedure of locating an area of interest and identifying key points that can be compared through images.

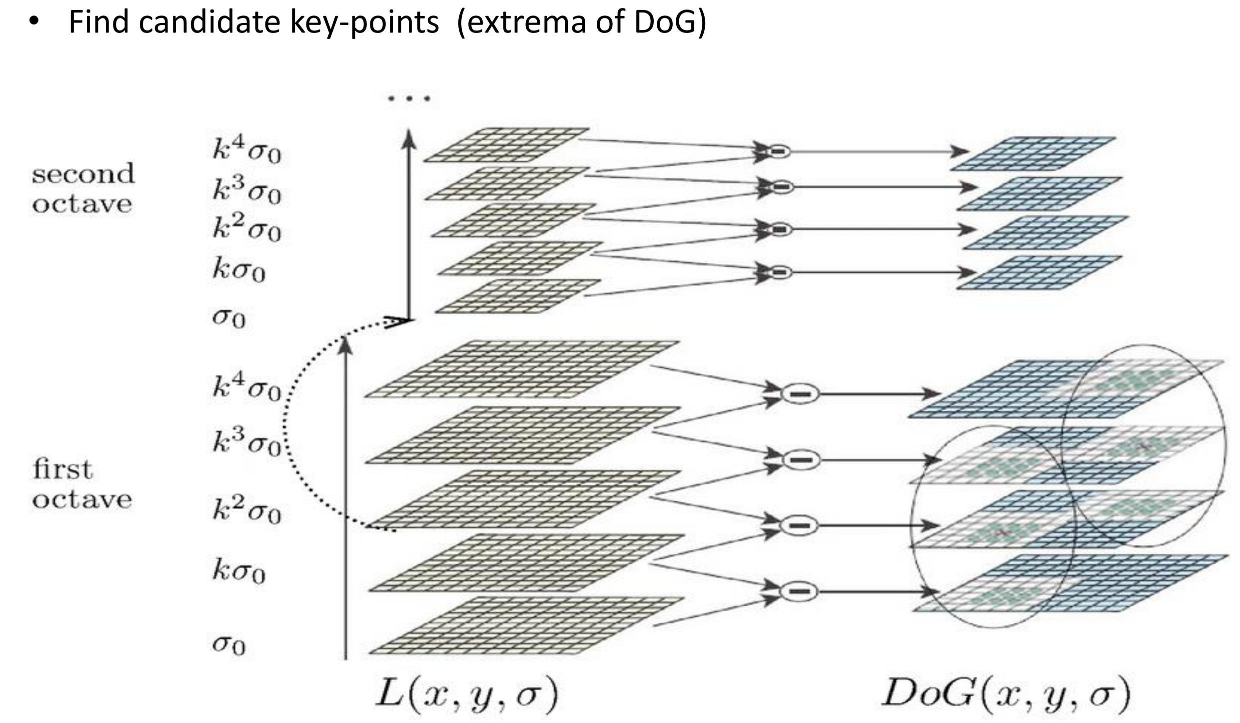
Therefore, an algorithm we plan to start working with is SIFT, or known as Scale Invariant Feature Transform [1]. This algorithm comes to our interest due to its ability to match features across different images that can be scaled, rotated, illuminated, or distorted. We will begin our investigation by exploring the steps of SIFT as well as potential ideas for formulating an anytime version of it.

SIFT

1) *Approximate Key-point Location:*

We want to find a potential location for finding features.
To do this:

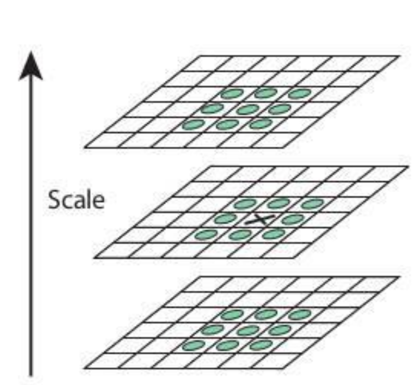
- Compute Gaussian scale-space
- Compute the Difference of Gaussians
- Find candidate key-points (extrema of DoG)



2) *Refining Key-point Location:*

We want to accurately locate the feature key points (by eliminating low contrast points).
To do this:

- We use a Taylor series equation of scale space to get a more accurate location of extrema.
- If the intensity at this extrema is less than a threshold value, it is rejected.

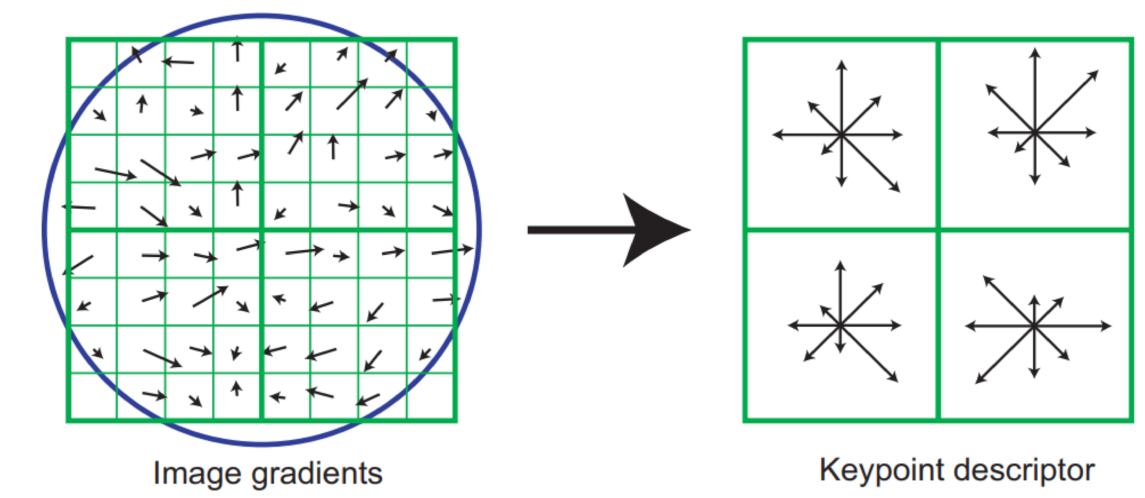


3) *Orientation Assignment:*

- An orientation is assigned to each key point to achieve invariance to image rotation.
- An orientation histogram will reveal the most critical key points

4) *Building Key-point descriptor:*

•Compute a descriptor for the local image region about each key point that is invariant to variations such as changes in viewpoint and illumination.



Our Objective and Proposal for Further Research

So far, we have introduced the background of Image Retrieval, the sub-areas within it, and SIFT. Challenges such as interpretation of partial information, computational complexity, and resource allocation will be considered when formulating and designing our algorithm. Therefore, we have begun to explore, design, and implement SIFT with anytime properties and study its performance advantages through experimental validation. To begin our implementation of SIFT, we chose to work with the Python programming language for simplicity as well as its available functions within the OpenCV libraries. OpenCV is a collection of programming functions aimed at real-time computer vision, and it includes functions that follow the steps of SIFT. As we formulate this algorithm, we have hypothesized that increasing the number of scales and blurs would generate better results for key-points. In the future, continued experimentation for implementing this as well as partially processing the data through parallelization would help us get closer to our goal of an anytime version of SIFT. But broadly speaking, our results are intended to reduce run time, computational resources, and hopefully be a step towards crafting a more efficient method for image retrieval.

References

[1] David G. Lowe. 2004. Distinctive Image Features from Scale-Invariant Keypoints. Int. J. Comput. Vision 60, 2 (November 2004), 91-110.

[2] Zilberstein, Shlomo. “Using Anytime Algorithms in Intelligent Systems.” AI Magazine 17 (1996): 73-83.

[3] Jayan, Neha. “A Sift-Based Partial Processing Algorithm for Real Time Image Search.” ScholarWorks, California State Polytechnic University, Pomona, 1 Jan. 1970, https://scholarworks.calstate.edu/concern/theses/1z40kw13r.