Vincent Ng
vng1@jhu.edu
CIS II: 600.446 SPRING 2011
March 10 2011
Project: Retina Registration Project

Paper Critical Review

Paper: "SURF: Speeded Up Robust Features"

Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008

Authors: Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool,

Summary of motivation and background

Computer vision has been a novel field by using machines to extract information from images to do a specific task. One area is the case of image registration and tracking. Methods varying from estimating image movement to matching feature points have been used.

This paper presents “a novel scale- and rotation-invariant detector and descriptor” that claims to perform as well as prior work while maintaining a quicker run-time. One focus of the retina registration project is to provide real-time tracking of features. The speedup of SURF over its pseudo-predecessor SIFT(Scale-Invariant Feature Transform) along with its performance level for image deformations make it a valid focus for the project.

Summary of article

The article begins with a description of previous work and algorithms along with their flaws. The authors comment on the early-stages of this field, using the Harris corner-detector which is not scale-invariant. They proceed on highlighting the pros and cons of using Gaussians, Hessian and Laplacian for the feature detection portion before settling on using a Hessian-based approach, with approximation as small component in trying to speed up computation. For feature description, the authors reused ideas from SIFT, saving information of smaller-scale features nearby the feature point.

Vincent Ng
vng1@jhu.edu

For feature detection, three main mathematical tools are used. The target result is to get the determinant of the Hessian matrix, and accepting points where this determinant has a high value. Integral images are used to reduce computation time of a region from its four corners, box filters are used to approximate the determinant of the Hessian matrix, with some explanation of key constants in such calculations. The role of scale-invariant detection is achieved due to the usage of box filters. The trace of the Hessian is noted as an interesting point: features with the different sign of the trace have different contrasts and should be discarded straightaway, aiding in the speedup of matching.

For feature description, the generation of an orientation is done via summing the Haar wavelet responses and picking one with the longest summation vector. The description of the area around the feature point(separated into grids) is done with great detail and experimentation of grid size and wavelet features size. The authors note that SIFT uses the histogram of gradients, which is more prone to noise distortion than their spatial information approach.

The rest of the paper describes experiments in the SURF algorithm in terms of speed and accuracy when compared to varying parameters within SURF, and previous algorithms like SIFT. In the end, the authors concludes that the amount of speedup is tremendous with little data or accuracy loss, and is a viable option for a live video-streaming project.

Critique

The authors come up with a good design and idea for feature detection and description. This method is not entirely radical: it is based on observation and in-depth analysis of the flaws from previous algorithms, and implications proposed from such algorithms. The goals of having a repeatable feature detection with speedup are achieved in the end. The authors justify their choices, either from mathematical reasoning, or from testing and tweaking of the algorithm. They even propose a U-SURF algorithm that does away with the rotation-invariant component for a faster computation time, useful for conditions where rotation is not present. The second half of the paper deals with testing and showing how different SURF and SIFT are (despite the similarities in implementation) in terms of speed and accuracy, along with implications in object recognition and camera calibration.

There are two areas where the authors could improve on. They keep commenting on how good the speed of the algorithm is, but they could spend more time focusing on the performance comparisons, though their initial presentation of SURF(in a previous paper) had some information about the performance comparisons. Another way would be the testing procedure: the sample inputs of 48 images(blur, viewpoint changes, zoom/rotation, light, JPEG compression) is not a large subset of ways real-life images could differ. Different distortions might affect SURF differently than SIFT due to the approximation methods used.

Relevance

The paper highlights the speedup component of the algorithm. From other testing, it might be possible for a live implementation to be constructed (40ms for matching of 200 features). The presence of scale invariant is of tremendous help in this project, as the microscope can change magnification depending on the surgeon's need. Also, the amount of customization available (Hessian threshold for feature detection and description size) allows for more fine-tuning in the face of distortion.

Implications

Vincent Ng
vng1@jhu.edu
CIS II: 600.446 SPRING 2011
March 10 2011
Project: Retina Registration Project

There can be a lot of future work done from the results of this paper. In the second half, the paper only does a brief overview of how the algorithm performs in object recognition and camera calibration. More testing can be done in different areas of computer vision, along with more results of how well the SURF algorithm performs in different environments (lighting, distortion, exposure). Other work has been done on parallelizing this algorithm for GPU usage. One small portion of the algorithm that might be improved would be the use of color. Currently, the inputs are converted into grayscale images before running the algorithm, but color can provide more information in feature matching. Also, the authors only use a closest-neighbor matching algorithm: perhaps this algorithm might be better harmonized with a different matching algorithm like normalized cross correlation.