SURF: Speeded Up Robust Features.

1. Introduction

*Feature extraction* is an essential step in the process of classifying data. Its aim is to reduce the dimension of the data to be classified into a smaller dimension but whilst still retaining the important information to perform the task effectively. Object Recognition, Pattern Recognition and other classification algorithms rely heavily on feature extraction to obtain a feature set [1]. Before a feature set can be obtained, the data has to go through a pre-processing phase which involves presenting the data to be more appropriate for processing. Binarization, thresholding, resizing, etc. are some of the steps taken. After that has been done a feature extraction algorithm can now be applied to automatically determine which features of the data can be used. The result of this analysis is a meaning set of features that best represent the data to be process. This report will discuss the feature extraction process in images. There are several very well-known feature extraction methods for images, but for this report, I am going to discuss about **SURF (Speeded Up Robust Features)**. This method is partly based off **SIFT (Scale Invariant Feature Transform)**.

SURF is called a feature detection or description algorithm that automatically examines images and extract the most relevant information. This information can be later used for pattern or objection recognition. SURF algorithm can be divided into three steps; detection, description and matching. SURF utilises is based on the same principles and steps of SIFT [2].

1. Detection:

Automatically identify interesting features, *interest points* this must be done robustly. The same feature should always be detected regardless of viewpoint. In order to detect feature points or interest points, SIFT uses cascading filtering approach, where the Difference of Gaussians (DoG) is calculated on a progressively down-scaled image. The detector is based on the Hessian matrix, but uses a very basic approximation, just as DoG it relies on integral images to reduce the computation time. SURF uses a hessian matrix because of good computation time and accuracy. Given a point in an image *M*, the hessian matrix so that at scale is defined as follows:

(1)

Where is the convolution of the Gaussian second order derivative with the Image *M* in the point and the same applies for the other remaining points. In general, the technique hessian-based blob detector to find interest points. The heart of the SURF detection is non-maximal-suppression of the determinants of the hessian matrices [3]. The convolutions are very costly to calculate and it is approximated and speeded-up with the use of integral images and approximated kernels. To detect the interest points at each level of scale of the DoG pyramid every pixel is compared to its neighbours: if is a local extrema (local minimum or maximum) it is selected as a candidate key point each *candidate key point* is compared to the 9 neighbours in the scale above and below Only pixels that are local extrema in adjacent levels are promoted as key points. An Integral is an image where each point stores the sum of all pixels in a rectangular area between origo and .

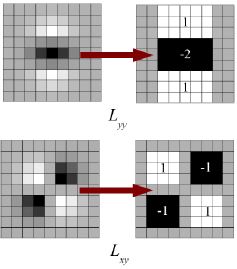
Using integrated images allows for calculating a response, in a rectangular shape with a random size. This can be shown with the image below:



Figure

Figure

As shown above, this algorithm detects features by using rectangular boxes. In that figure, the grey looking area can be represented with 0 and the white area is positive and black is negative. This makes it possible to find the convolution [4].

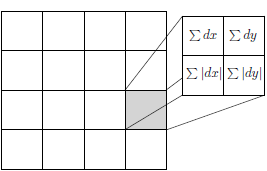


Figure

1. Description:

Each interest point should have a unique description that does not depend on the features scale and rotation. SURF uses wavelet responses in horizontal and vertical direction for a neighbourhood of size of 6s. Adequate gaussian weights are also applied to it. Then they are plotted in a space as given in below image. The dominant orientation is estimated by calculating the sum of all responses within a sliding orientation window of angle 60 degrees.

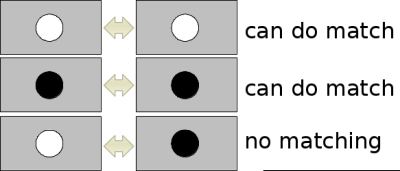
For feature description, SURF uses Wavelet responses in horizontal and vertical direction (again, use of integral images makes things easier). A neighbourhood of size 20sX20s is taken around the key point where s is the size. It is divided into 4x4 sub regions. For each sub region, horizontal and vertical wavelet responses are taken and a vector is formed like this, . This when represented as a vector gives SURF feature descriptor with total 64 dimensions. Lower the dimension, higher the speed of computation and matching, but provide better distinctiveness of features.



Figure

1. Matching:

Given and input image, determine which objects it contains, and possibly a transformation of the object, based on predetermined interest points. In matching, the algorithm, will described features with other features to return the most suitable ones. A common approach to matching is implementing a brute-force matching algorithm which will go through all the described features in one set and try to match them in the other set based on distance. If the distance is shorter, that feature is returned.



Figure

1. References:
   1. Herbert Bay, Andreas Ess, Tinne Tuytelaars, Luc Van Gool, "SURF: Speeded Up Robust Features", Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346--359, 2008
   2. Khan, Nabeel & McCane, Brendan & Wyvill, Geoff. (2011). SIFT and SURF Performance Evaluation against Various Image Deformations on Benchmark Dataset. 2011 International Conference on Digital Image Computing: Techniques and Applications. 501-506. 10.1109/DICTA.2011.90.
   3. Jacob Toft Pedersen SURF: Feature detection & description, 2010. http://cs.au.dk/~jtp/SURF/report.pdf
   4. <https://docs.opencv.org/3.0beta/doc/py_tutorials/py_feature2d/py_surf_intro/py_surf_intro.html>