

1 State of the art

Since few years color image processing is a major problem, indeed lot of the colour texture discrimination has been explored in a marginal colour way. The problematic is that today we can do color image recognition on numerical images but we don't have good results on nature images.

The CLEF contest is an answer to that problematic, the aim of the challenge is to put in competition some university laboratories and company laboratories. In this contest each laboratories could compare its colour texture feature against all the other challengers.

In this way, we will make a little presentation of what is a Key-points and how we used them. In a first time with the classical descriptors, SIFT, SURF, and opponent SIFT. We choose to present the last one because this is the descriptors used by FINKI, which is the laboratory of the last year contest we choose as reference for their methodology and to compare our results in a first place. In a second time, we will use a new descriptors purposed by Noel Richard. This descriptors is the C_2O , which we going to describe in this part with his strengths and weakness.

1.1 Key-points

1.1.1 Classical

1.1.2 Dense Grid

The dense grid method is an easy way to extract Key-points. First we have to divide the image in k subimages, where k is the number of subimages choose by the user. After you just have to take as Key-points the corner of each subimages.

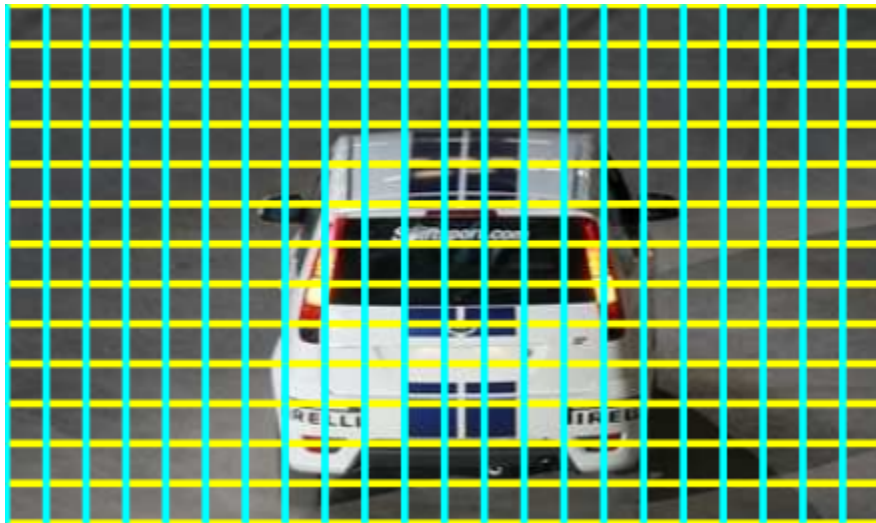


Figure 1: Dense grid

1.2 Descriptors

1.2.1 SIFT

Scale-invariant feature transform (or SIFT) is an algorithm in computer vision to detect and describe local features in images. The algorithm was published by David Lowe in 1999. Applications include object recognition, robotic mapping and navigation, image stitching, 3D modeling, gesture recognition, video tracking, individual identification of wildlife and match moving.

The algorithm is patented in the US; the owner is the University of British Columbia.

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. The determination of consistent clusters is performed rapidly by using an efficient hash table implementation of the generalized Hough transform. Each cluster of 3 or more features that agree on an object and its pose is then subject to further detailed model verification and subsequently outliers are discarded. Finally the probability that a particular set of features indicates the presence of an object is computed, given the accuracy of fit and number of probable false matches. Object matches that pass all these tests can be identified as correct with high confidence.

1.2.2 SURF

In computer vision, Speeded Up Robust Features (SURF) is a local feature detector that can be used for tasks such as object recognition or 3D reconstruction. It is partly inspired by the scale-invariant feature transform (SIFT) descriptor. The standard version of SURF is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT.

SURF uses an integer approximation of the determinant of Hessian blob detector, which can be computed with 3 integer operations using an integral image. For features, it uses the sum of the Haar wavelet response around the point of interest. These can also be computed with the aid of the integral image.

SURF descriptors can be used to locate and recognize objects, people or faces, make 3D scenes, track objects and extract points of interest.

SURF was first presented by Herbert Bay et al. at the 2006 European Conference on Computer Vision. An application of the algorithm is patented in the US.

1.2.3 SIFT vs SURF

1.2.4 Opponent SIFT

The opponent SIFT descriptors is an algorithm which use the same method than the classical SIFT descriptor. The only difference is that we calculate three descriptors for each Key-points. There are obtained from the color opponent channel, defined as

$$O_1 = \frac{R - G}{\sqrt{2}}, O_2 = \frac{R + G - 2B}{\sqrt{6}}, O_3 = \frac{R + G + B}{\sqrt{3}}. \quad (1)$$

The opponent SIFT describes the three opponent color spaces, we can see that the two first channels O_1 , O_2 contain some intensity information, but they are not invariant to changes of lights intensity. The last channel will contain the intensity information.

The strength of these method is that that use a color spaces, and we can see directly information of that with the algorithm of the SIFT descriptor. The weakness is that, in this approach we use the RGB space to make this computation.

1.2.5 C₂O